FOUNDATIONS OF DATA ANALYSIS IN RESEARCH

Dr. Soumen Saha
Dr. Sudipta Adhikary

FOUNDATIONS OF DATA ANALYSIS IN RESEARCH

By

Dr. Soumen Saha

Senior Deputy Registrar & Associate Professor, Department of Commerce

Brainware University

Dr. Sudipta Adhikary

Associate Professor, Research Coordinator & MOOCs Coordinator,

Department of Law

Brainware University

Published By

World Academic Press, Kolkata-700126, India

www.worldacademic.press

Published by:

World Academic Press , Kolkata, India

www.worldacademic.press

DOI: https://www.doi.org/10.66727/wap.9788168643901

ISBN: **978-81-686439-4-9** ***(Paperback)***

ISBN: **978-81-686439-0-1** ***(E-book)***

First Edition: 2026

About the Authors

Dr. Soumen Saha

Dr. Soumen Saha, Senior Deputy Registrar at Brainware University, Kolkata, holds a PhD in Economics and dual Master's degrees in Economics and IT, with over 25 years of academic and administrative experience. He has served in key positions across institutions and contributed to major government projects, published 15+ papers, authored a book, and holds three patents. A seasoned researcher and educator, he specializes in Management, Applied Economics, and Data Analysis, with expertise in SPSS, AMOS, and PROCESS tools.

Dr. Sudipta Adhikary

Dr. Sudipta Adhikary is an Associate Professor in the Department of Law at Brainware University with over 10 years of academic and research experience. He holds multiple qualifications including MBA (HR), MBA (Marketing), M.Tech in Geotechnical Engineering, MPA, and Ph.D. in Management. In addition to his teaching responsibilities, he serves as the Research Coordinator and MOOCs Coordinator in the Department of Law. He has published 30+ research papers, authored books, and holds patents in his areas of expertise. His academic interests include Management, Legal Studies, Research Methodology, and interdisciplinary studies, with active involvement in research, curriculum development, and academic administration.

Table of Contents

CHAPTER 1: Understanding Data Analysis in Research

1.1 Introduction to Data Analysis

Research rests on data. *"Across the natural sciences, social sciences, business, and the humanities, the ability to analyze data transforms collections of numbers into meaningful knowledge" (Provost & Fawcett, 2013, p. 2).* Research begins well before the analyzing the data phase. Identifying the problem and worthwhile research questions and determining the methodology for data acquisition itself is a complex task. This is especially the case when data is collected through multiple channels such as surveys, experiments, observations, or through secondary sources. Accomplishing this is only the beginning. This is the start of the complex phase of data sense-making. This is the point of and the point in which data analysis comes in.

This section of a book explains the preliminary steps of data analysis in research, laying the foundation to explore instruments such as spreadsheets (Excel/Google Sheets) and applications like SPSS. In the first chapter, emphasis is put on developing the core principle of the importance of data analysis, how it integrates within the research cycle, and the types of queries it helps in addressing. It also seeks to provide a philosophic explanation through the well-known fable "The Mirror of Fragments". The metaphor illustrates how data is used and understood in research. A researcher can possess an array of sophisticated tools and still see, obtain, and analyze the data in an inexplicably varied and limited simple way. This is critical in emphasizing objectivity, variety of approaches, and triangulation in research.

Identifying types of data and their levels of measurement is vital preparation for the coming analyses. It is the case with nominal (gender, religion), and ordinal (satisfaction levels) fractions. With interval (temperature) and ratio (income, height) data, there is also none of this. These classifications are not just classifications. They are the constituent parts.

The chapter also sheds light on the distinction between descriptive and inferential statistics. Measures such as mean, median, mode, and standard deviation help quantitatively characterize and outline the primary aspects of a data set which is known as descriptive statistics. In contrast, inferential statistics employs methods such as t-tests, ANOVA, regression, and chi-square tests to assist a researcher in making predictions or generalizations about a population from a sample of the data.

More than the computation of data, this section highlights the need to 'think' about the data. It is the duty of a researcher to critically inquire: Are the results and data consistent? What patterns are evident? Is the hypothesis valid? What are the findings suggestive of?

This section will equip you with the necessary conceptual understanding and practical guidance to allow you to approach your research with confidence, whether you are a novice attempting to understand your first dataset or an established researcher looking to develop your analytical methodologies with computing.

1.2 Understanding Data Analysis in Research

Without a doubt, data analysis and synthesis is an important pillar of research practice. Developing an understanding of this principle is necessary for the refined and meaningful research. Data analysis is the last stage of research and is a process that connects collection of data and synthesis with integration of meaning. Data collection comes from a variety of sources, Experiment and Surveys, Interviews and Questioning and even Secondary Research. In all these scenarios, data spaced meaning is determined by analysis.

1.2.1 Purpose and Importance of Data Analysis

Data is collected from primary and secondary sources such as, experiments, observation and surveys, and is transformed to actionable data by employing the process of data analysis. Understanding these techniques is of utmost importance as data analysis is an integral part of the research process. The essence of data analysis within research lies in evaluating information, conducting hypothesis tests, substantiating claims, and in the process, aiding in the construction of new knowledge. It captures the essence of recognizing patterns, defining interactions, analyzing populations, and evaluating models. No research analysis, no matter how elegant, would have any clarity let alone objectivity and value without some defined parameters.

Additionally, data analysis substantiates trustworthiness and authentication of the research. It makes certain the conclusions are not crafted from sheer guesswork or an untested story but are results of rigorous research. It facilitates the process of reasoning, especially in tangible endeavours like policy-making, corporate, education, and health matters. Researchers are able to distill the data to provide useful reasoning and conclusions Đ thereby aiding stakeholders to engage in evidence-based decision-making.

In addition, analysis which is executed thoughtfully, enhances the findings of the research in republication and generalization. It makes it possible Đ within reason Đ for other research to authenticate, or expand, or even contest the findings, thereby promoting the cyclic feature of the scientific undertaking.

1.3 Qualitative and Quantitative Data

In most inquiry undertaking, data is classified into the broad categories of qualitative and quantitative.

There are several pieces of information that can be counted or measured. Examples of this type of information are age, wealth, distance, evaluations or rankings, and high school grades. In order to analyze quantitative information, various statistical techniques are employed to test assumptions, correlate and predict. The techniques are in the range from statistics (mean, median, standard deviation) to inferential techniques (regression, t-tests, ANOVA and others).

Contrary to that, qualitative information does not contain numbers and focuses on the subjective understanding of concepts through the interpretations of spoken and written words, actions, experiences, and sights. Examples of this type of information are interviews, open-ended questionnaires, and field note observations. The determination or evaluation of qualitative information involves thematic coding, content analysis, narrative analysis, or grounded theory approaches.

Each of the two information types can greatly benefit the research. The quantitative data helps to provide answers to the questions: "how much/many", whereas qualitative data answers "why/how". Still, longer and more comprehensive research documents utilize both qualitative and quantitative data.

1.4 Stages of Data Analysis

A set of methodologies in research and analysis considers how data is collected and what it is intended for and how it is analyzed after collection and explained later, which includes:

- **Collection of Data:** This is the very first stage which may involve the use of instruments such as research questionnaires, interviews, experiments, and even document reviews. The level of analyzed data is influenced directly and indirectly to considerable extents by its quality.

- **Data Cleaning and Preparation:** Data which is in its raw form and is not processed is purported to hold a range of discrepancies such as missing values and errors. Data Cleaning consists of data coding, and data formatting to analysis in which the data's accuracy is preserved. This stage is important now a days, especially 'big data' in software data tools such as SPSS or even Excel.
- **Analysis of Data:** This is the stage in research when appropriate analysis techniques are picked out and used, according to the type of data available and the question or hypothesis on the research. As for quantitative data, statistical methods and tests are used, while qualitative data means there are themes and narratives to be analyzed.
- **Interpretation of Data**: Data after being analyzed and processed is used to find new relationships to derive new conclusions. This type of thinking and research is not just mere thinking in numbers; there is a pure intellect, additional research, and insight in the wider picture and how it incorporates to the theory which the research is based on.
- **Reporting and Visualization:** More case studies and articles usually tend to have the wider analysis be quantified and pivoted in the form of figures, tables, and even graphs which is explained in a wider form of text. Reporting is a process which needs to be formulated and structured such that people on ground and people not aware of the analysis are able to understand the reasoning and be able to use the analysis drawn.
- **Drawing Conclusions and Making Recommendations:** Conclusions should always be derived, research questions should always be answered, and relevant implications or future avenues for research need to be proposed all based on the evidence analyzed.

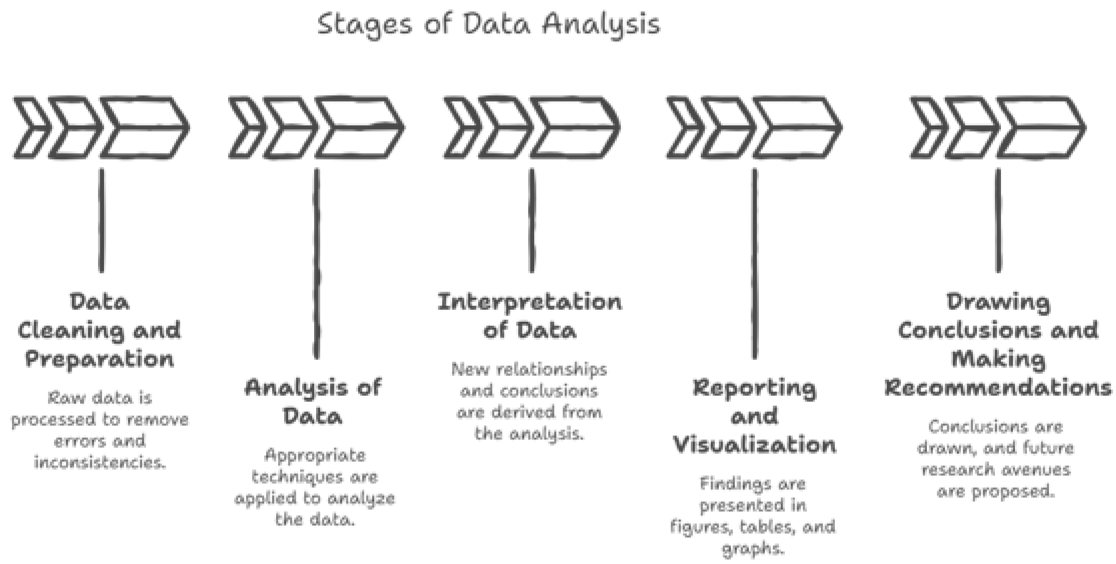


Figure 1.Stages of Data Analysis

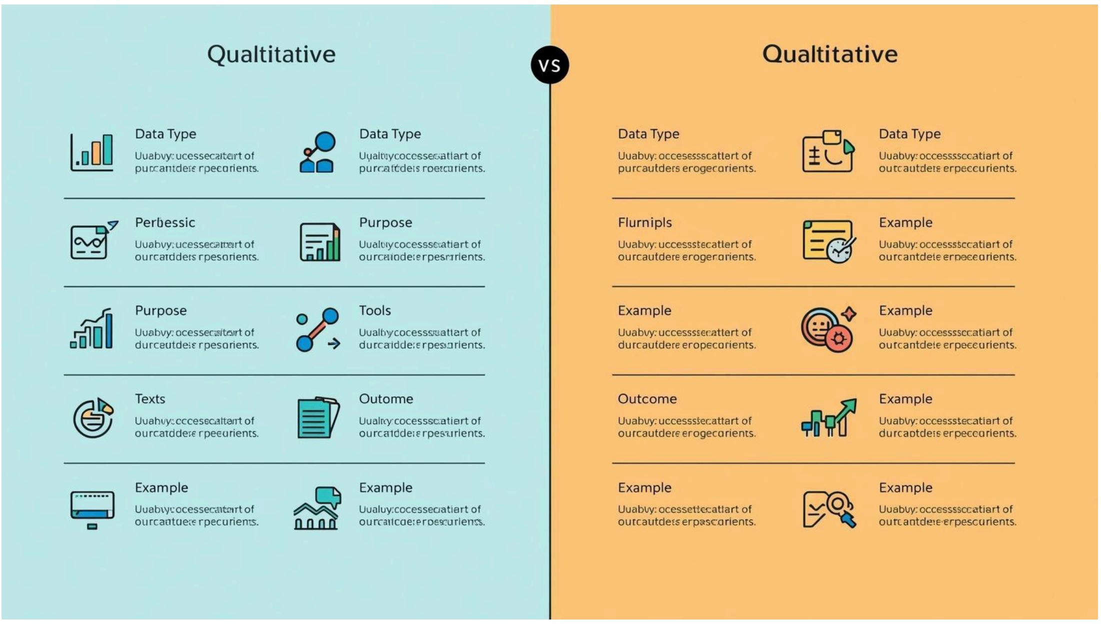


Figure 2. Comparison between Qualitative and Quantitative Research Approaches.

In this case, the process of data analysis is not an autonomous activity but rather part of the ever-expansive research corpus. A thorough understanding of the purpose behind the data, the types of data, and the approach to be followed from collection all the way to interpretation is crucial for the research to be not only rigorous but powerful in its impact.

CHAPTER 2: Philosophical Foundations of Research: The Mirror of Fragments

Seeing research from a philosophical standpoint means understanding that a truth is rarely a singular, bright object that is simply to be found. It is usually a collection of reflections, points of view, and interpretations of a given subject. One modern metaphor that captures this idea is 'The Mirror of Fragments'. This is a story that creates a bridge between age-old wisdom and contemporary research from which a lesson can be drawn. This story, unlike traditional fables, is a reminder that the pursuit of truth is not only the product of one's work, but also requires one's humility, cooperation, and a fusion of ideas.

2.1 The Story and Its Meaning.

There once was a mountain village that was quiet. It housed a temple with a grand mirror. This mirror was said to be a gateway to the universe, because the legend said that whoever looked into it was able to see the universe's truth. People from across the world would travel to this village to see this mirror. One night, a terrible storm hit the area, and the temple's roof was hit by a lightning bolt. This caused the mirror to shatter into hundreds of pieces that scattered across the valley, and the rest of the temple ultimately collapse.

With every new day, it was the legend that lived on and not the mirror that remained broken, though its pieces were never lost. The reflection in every lost piece seemed so unique, and every piece was sought after as "the truth." Sant, the lost truth, is discovered in the most peculiar places—forests, residences, even along rivers.

Every monk and philosopher sought after something different. A philosopher who found the shard that reflected the sky sang praises and proclaimed the truth to be boundless and infinite. A monk who found a shard that reflected his face proclaimed the truth to be of the self, self-contained. A scientist who found a shard that showcased the most minuscule of details declared the truth to be in evidence and precision. A poet who saw colours dance in her shard sang to the truth being of beauty and feeling.

Such diversity in perspective, and yet, the comfort of slumber in reality was paramount. Every single one of them captured a mere section, to be introduced as a larger image, disconnected in the telling of a tale.

With every passing day, not a broken mirror, but the legend of the mirror was alive. Capturable in lost pieces every unique piece was sought after as "the truth." Along the rivers, in residences, even beside the forest each lost piece captured a reflection so unique and yet so beautiful. The child was unaware that such beauty was crafted deep within the surrounding parts of the temple ruins. Constructed mosaic style with no piece complimenting the next, it was every single shard that reflected, to reveal a kaleidoscope. Smiling as the child was, it was the broken truth which was wiser, that was not lost. The pieces were everywhere, and all that was needed was the thought put in to connect them. Maximum reflection was created from the minimum.

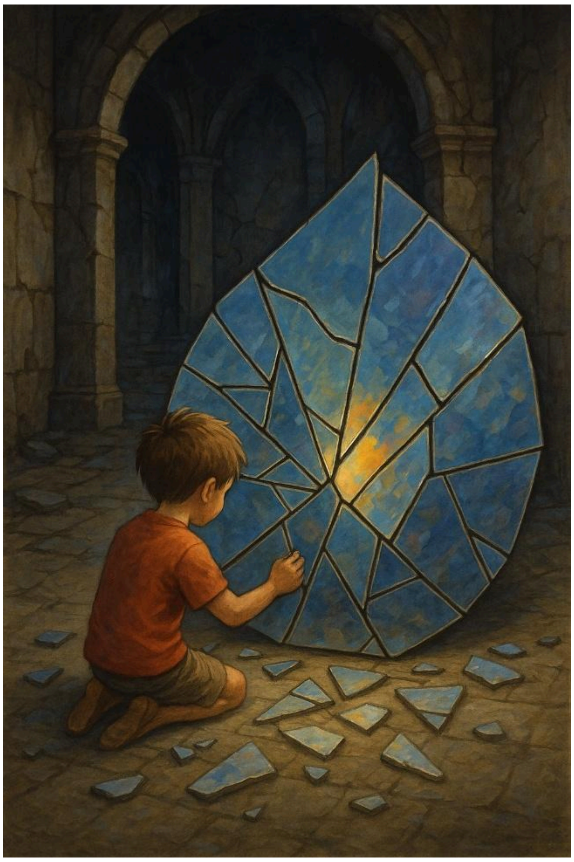


Figure 3. Within the ruins of the fallen temple, a child assembles fragments of the shattered mirror—each shard reflecting a partial truth. Like research itself, no single piece offers completeness; meaning emerges only through careful connection, synthesis, and interpretation. From minimal fragments arises maximum reflection, reminding us that knowledge is not discovered whole, but constructed through the thoughtful integration of diverse perspectives.

2.2 Connection to Research Philosophy

This allegory resembles the nature of research in analytical studies and theories. Each separate study and method is akin to a part of an ancient mirror. Each part reflects a certain part of reality, but never the whole of it. Researchers tend to gravitate towards their own fragment, smug in the assurance that it is a complete representation of the reality of the situation. However, if they do not acknowledge the other reflections they run the risk of mistaking a partial truth for the whole.

This is a story that, philosophically, speaks to epistemological pluralism— the idea that coexistence of multiple ways of knowing is possible and that they can complement and enrich one another. It highlights that in research, quantitative data, would only capture the quantifiable elements, while qualitative data would shed light on the human meaning and context. Each of them is inadequate on their own, but together, they offer the truth.

2.3 Subjectivity and Researcher Perspective

The Swiss psychologist actively suggested that, 'Each fragment's reflection is shaped not only by what it shows but also by who looks into it. This suggests the role of the subjectivity of the researcher in the construction of knowledge. In the same way the philosopher, the monk, the scientist and the poet all captured a certain truth through the lens of their experiences and values, so do the researchers. Each of them brings their own paradigm to their work, be it positivism, interpretivism, constructivism or criticism. *or minimally seasoned professionals these tools are an important part of their analytics tool chest.*

It does not diminish research; it enriches its honesty. Resolving their positionality requires researchers to understand the lens from which they perceive the dimensional world. Conventionally, reflexivity entails self-scrutinizing biases, assumptions, and other cultural influences, and for research, it is a fundamental ethical principle.

2.4 The Value of Synthesis

The child's act of placing the mirror fragments together captures the spirit of interdisciplinary and collaborative research. Economists, psychologists, sociologists, engineers, and artists tackling a question from their diverse vantage points are like mirror pieces assembled side by side. The result is not perfect symmetry, but it does provide a fuller, truer reflection.

This synthesis resonates with the strands of pragmatism and critical realism, which philosophically advocate for the simultaneous existence of diverse methods and diverse realities. Pragmatism is concerned with the socially useful, problem-centered integration of ideas, while critical realism notes that a reality exists outside of our understanding but is accessible only in nuanced interpretations of our observations.

2.5 Key Takeaways for Today's Researchers

The Mirror of Fragments offers contemporary scholars a number of important lessons, such as:

- Truth is composite. No single method, dataset, or framework is universally dominant.
- Reflection requires humility. Each researcher must acknowledge their own blind spots.
- Understanding is a multi-layered process. Recognition of patterns in research is often obscured by the failure to merge disparate approaches.
- It is moral to be candid. For the case of the broken mirror, to accept its subjectivity, as well as its positionality, takes honesty to the level of interpretation.
- Knowledge is dynamic. Just as the mirror is said to be broken, so too the understanding has to be further deepened, through collaborative efforts and reinterpretation.

In fact, the pieces of the shattered mirror echo the lack of ownership of knowledge and highlight the need for engaging in an evolving conversation, an exchange of pieces of understanding. In the case of the shattered glass, it is not the researchers who play the role of the collectors, but the ones who will put together the reflections, the evidence, theory, and perspectives, to arrive at the meaning encompassed in it. It is graceful to acknowledge that, in construction, it is the complexity of the research which brings out the richness, for there is not a single, perfect truth, but a variety of partial ones put together with elegance.

CHAPTER 3: The Role of the Researcher in Data Analysis

Research takers in the social science and humanities are not simple data gatherers passively watching the research unfold. While the data is being collected, the person is shaping the outcome as well. This is most prominent in the analysis which is often surrounded with interpretations and judgment. This is the reason understanding the role of the researcher is one of the most vital steps of credible, ethical and in-depth research. This section discusses three dimensions of the researcher’s role: the researcher as an instrument, importance of positionality, bias and reflexivity, and the ethical responsibilities in the data and its interpretation.

3.1 Researcher as Instrument (particularly in qualitative research)

Qualitative research is one of the very few fields where the researcher is the data collection instrument. Researcher bias is heavily absent in quantitative works unlike in qualitative research where such instruments as surveys, forms and data, focus groups, and ethnologic watching are heavily utilized. Ethnology for further research is researched in observing conversations, actions, coupled with the cultural and contextual aspects. These instruments are qualitative in nature, and along with the rigidity of the data collection, these instruments are often unavoidable. Hence, unless one is dramatically ignoring their personal background, experiences and the lens through which they view the world, bias is inevitable.

Given the consequences, qualitative researchers should maintain and practice self-awareness and critical thinking during the course of the research. They must understand how their values, assumptions, and expectations impact their data engagement. For example, two researchers analyzing the same interview transcript may arrive at different conclusions due to their disciplinary training, surrounding culture, or underlying ideology. To acknowledge oneself as a research instrument also enhances the analysis as well as adds clarity and responsibility to the results.

3.1.1 Positionality, Bias, and Reflexivity

Positionality concerns the socio and political circumstances that define a researcher's identity such as gender, race, class, nationality, and the institution. These factors are also the same which define the research engagement. Understanding one's positionality is critical to appreciating the subjectivity of research. For example, a male researcher researching on gender issues in rural India will understand the situation differently compared to a female researcher in the same research setting.

Every scientist has different biases and personal suggestions. These biases are mostly from personal emotions that the scientist may have. Rather than trying to eliminate the bias completely and changing emotions entirely which is impossible, the researcher should focus more and practice on Reflexivity. Reflexivity addresses the critical self-examination that explores one's role, drives, power assumptions, emotions, as well as the entire process of the inquiry. Reflexivity the self-analytical practice that is constant in nature and, therefore, adds more depth and rigor to the analysis. Keeping self-reflective records on this process helps to establish the credibility of the analysis and assists the readers more in making their own judgments on the research.

3.2 Ethical Responsibilities in Data Handling and Interpretation

Interpreting information goes beyond obtaining informed consent as well as making sure that no one's identity is revealed. Data handled in such a manner must protect the information provided such that the details are stored honestly, and the interpretations carried out are fair. Any case of researcher bias must be avoided which means the researcher should not change the information in the form of 'manipulation', or supporting the researcher's set assumptions out of the cherry picked information.

Moreover, the dishonest and biased analysis concerning the facts is what disrespects the voices of the people participated in the research and more so, the unprotected segments of the society. An integral aspect of research is manipulation which can have damaging consequences for the participants and even diminish the research's integrity. Responsible practice is demonstrated in the ways that defend research as well as proper methodologies, recognition of the research's restrictions, and willingness to reconsider different viewpoints.

In cross-disciplinary research, there are issues related to equitable credit allocation and cross-team accountability boundaries that researchers need to address. Ethics in research and practice at this stage is about the data privacy, algorithmic bias, and responsible AI use in research and practice.

There is presumption of ethical, responsible and trustful knowledge, which a researcher achieves by reflexively acknowledging themselves as part of the research and the instrument used. Thus the researchers role in data are not merely passive, mechanical, but rather interpretively deep and ethically weighted.

3.3 Understanding Natural/Hard Sciences and Social Sciences

Within different disciplines the purpose and the scope of research shifts owing to the subject at hand. The two broad and the most recognized domains in academic research are Natural or Hard Sciences and Social Sciences. Each of this has different approaches, methodologies, objectives, and philosophical underpinnings.

3.3.1 Natural or Hard Sciences

Natural science (or hard science) encompasses the study of the physical world and natural phenomena. Disciplines include physics, chemistry, biology, astronomy, geology, and environmental science (or ris). The natural sciences are distinguished by their objectives of subjectivity, measurability, repeatability, and verification.

Hard sciences operate with a positivist approach as there is no consideration of human perception towards reality as everything is objective and measurable. Hypotheses are tested through laboratory experiments, and the outcome is frequently articulated in the form of mathematics.

Consider the examples below.

A biologist explores the behavior of a certain enzyme at various temperature levels.

A physicist studies a specific domain of the science using the theory of controlled motion.

Regardless of context or locality, the findings are assumed to be reproducible and capable of being generalized. This is one of the reasons hard sciences are still regarded as more ‘accurate’ or ‘rigorous’ as a discipline, although with the rise of multi-disciplinary approaches, this perspective is changing.

3.3.2 Social Sciences

In contrast, the discipline of social science is centered around the analysis and study of society, human behavior, culture, and various social structures. This discipline consists of such fields as sociology and social anthropology, economics, political science, psychology, and the science of education. In contrast to the ‘hard’ sciences, social science research tends to deal with more intricate, context-laden, and interpretative phenomena, which are more complex, subjective rather than objective, and as a consequence, not easily measurable.

As more and more social science researchers focus on meanings, lived experiences, values, and structures of social power, the need to use qualitative approaches increases. While many researchers in this field still focus on more qualitative methods, interpretivism and constructivism emphasize the need to acknowledge how a researcher's context, culture, or even outcome is reciprocally shaped by the research participants through the interpretations of their reality.

For instance,

• All sociologists performing an analysis would understand the impact of interviewing on the impact of gender on career paths.

• Income Survey Reports would be analyzed by economists.

Replication and prediction are not as integral as understanding the nuance of the topic, looking at the context, and analysis.

3.3.3 Other Types of Sciences

Along with the two major branches, other important areas include:

- Formal Sciences: Mathematics, logic, and unempirical-statististics which are based on axioms and other logical systems. Although they do not deal with any of the social or natural worlds, they are fundamental to both.
- Practical Sciences: These are the branches of engineering, technology, and medicine which focus on the application of natural science to solve concrete problems. These branches focus of problem solving, invention and application as opposed to theory building.
- Behavioral Sciences: These are the branches of psychology, cognitive science, and neuroscience which are considered the integration of the natural and social science. These branches emphasize the understanding of human behavior at biological and social levels.

Table: 2.1 Natural/Hard Sciences VS Social Sciences

Type	In Natural/Hard Sciences	In Social Sciences
Basic Research	Lab-based theory building, often in physics or biology	Abstract concepts like justice, identity, or power dynamics

Applied Research	Focus on product development, medical innovations	Social programs, policy evaluation, community interventions
Quantitative	Emphasis on control, precision, and replicability	Uses surveys, large datasets; often deals with human variability
Qualitative	Less common; used in behavioral sciences	Core method in anthropology, sociology, and education; values context and subjectivity
Mixed Methods	Supplementary in clinical trials	Increasingly important for holistic understanding of social issues
Exploratory	Used in new or emerging fields	Common in studying under-researched communities or behaviors
Descriptive	Biological or physical characteristics	Demographics, cultural practices, political affiliations
Explanatory	Experiments under controlled conditions	Causal analysis in dynamic, uncontrolled social settings
Evaluative	Product or process effectiveness	Government scheme evaluations (e.g., Ujjwala Yojana, PM-KISAN)
Action Research	Rare	Widely used in education, social work, community development

While the broad types of research are shared across disciplines, **social science research** differs in its philosophical foundation (constructivism, interpretivism), methodological openness (especially to qualitative methods), and ethical considerations (due to human involvement). It often emphasizes **context, complexity, and reflexivity** over universal laws or rigid replicability. Understanding these distinctions is key for any researcher entering the social sciences.

3.4 Types of Research: A Focus on the Social Sciences

Research in social science involves myriad methods, aims, and philosophical bases, all working towards appreciating the complexity of how humans behave, live, and interact in society, culture, politics, and economics. Unlike in the natural sciences, where control of variables is possible in the laboratory, social science research concerns fluid and unpredictable social dynamics, making it interpretive, situational, and thick. In this chapter, we will discuss the various types of research in social science, clarifying their aims, methods, and importance.

3.4.1 Basic vs. Applied Research

• Basic Research or fundamental or pure research does not have any immediate practical application and seeks to broaden the theoretical understanding of a particular phenomenon. In social science, basic research may deal with the concepts of justice, democracy, or identity at a theoretical level. For example, a political theorist examining varying types of governance is engaged in basic research.

• On the contrary, Applied Research is concerned with the resolution of particular practical issues. It is pragmatic, purposive, and frequently sponsored by the state, NGO's, or private sector. For instance, monitoring the effects of a government welfare scheme on the lives of people in rural areas or assessing the impact of a new education policy on the education system.

3.4.2 Quantitative Research

Quantitative research begins with ascertaining the importance of a sample's data with respect to a defined population. This type of research uses structured research instruments such as surveys, questionnaires, and other statistical data. The primary focus of this type of research is to test hypotheses, measure variables, and recognize questions or relationships. For example, a study which tries to measure the correlation between income and the level of urban politics exercised is a quantitive study.

Someone conducting quantitative research is likely to focus on the objectivity of the research, the ease at which it can be reproduced, and the results it is able to produce compared to other research. A person basing research on qualitative methods might be doing them a disservice by ignoring the thoughts and nuance of a matter.

3.4.3 Qualitative Research

Qualitative research is the study of meanings, and experiences shaped by the interpretations of the participants. It employs tools that are unstructured and semi-structured like, the in depth interview, focus group discussions, and participant observation, as well as document review. Instead of providing numerical data, qualitative research offers insight that is thorough, in detail and description.

An example of such a study is, "How do domestic violence victims view the assistance offered to them by the state?" Such research attempts to step outside the confines of an 'objective' approach, and offers a more context, subject and experience oriented analysis. It is the depth and richness of such research that truly sets it apart, as opposed to the generalizations drawn from statistical research.

3.4.4 Combined Methods Research

Combining Methods Research uses both quantitative and qualitative techniques on a single study to provide as comprehensive an understanding of a research problem as possible. Case in point, a researcher studying the public perception of a new health policy might start with a survey to collect broad quantitative data and follow it up with an interview to delve into nuanced, qualitative elaboration.

This utilizes the strengths of both methods and is particularly advantageous in the case of answering complex research questions where neither qualitative nor quantitative methods, in isolation, would suffice.

3.4.5 Descriptive, Analytical, Exploratory, and Explanatory Research

- Descriptive Research aims to document a situation, problem, phenomenon, or a set of demographic attributes in a systematic manner. An example of this is describing the socio-economic conditions of slum dwellers in a metropolitan area.
- Analytical Research is a step further than description, attempting to consider and evaluate possible relationships and uses data or theory. For instance, considering rural to urban migration and its impacts on the traditional family system.
- Exploratory Research is undertaken as a first approach to a problem that is poorly structured. It is broad and aims to diagnose and develop new proposals and factors for the area of study in question. An example of this might be researching new online subcultures.

Explanatory (or causal) research focuses on exploring the reasons that cause a phenomenon. For instance, an analysis on why there is a higher dropout rate among girls in rural areas.

Figure 4. Different Types of Research

3.4.6. Action Research

Action research is both participatory and problem focused. It is most commonly done by practitioners, such as teachers, social workers, and community organizers, who aim to enhance their practice. For instance, a school teacher who changes seating arrangement and investigates how it affects student participation is an action researcher.

In this case, the researcher is also a participant, and the goal is to achieve practical outcomes while contributing to theory.

3.4.7 Historical and Comparative Research

Historical Research entails the analysis of social phenomenon by exploring past records and documents. It is essential to the fields of sociology, political science, and education.

Comparative Research analyzes the similarities and differences of cultures, countries, and even different time periods for the purpose of arriving at more general conclusions. It is commonly practiced in political science and anthropology.

The variety of research types within the social sciences denotes the intricacy of human life. Each approach has its pros and cons but the selection tends on the research question, goals of the study, and the researcher's ideology. Understanding these different options enables the researcher to create studies which are relevant to the society and appropriate to the set methodologies.

3.4.8 Emerging Trends in Data Analysis

Merging computer science with advanced methodologies has enhanced data analysis in ways that were previously unimaginable. The analysis does not rely purely on statistical methodologies anymore; it also encompasses artificial intelligence (AI), advanced analytics systems, and big data computing. The efficiency with which analyses can be conducted boils down not only to speed, but also to the level of multifactorial analysis that can be performed using predictive and prescriptive analytics. Nonetheless, any new analytical technique must adhere to the ethical and moral principles of clarity, accuracy, and ethical integration. We expand on the 'Modern analytics' and 'conceptual transformations' discussed in earlier sections of this chapter.

Classic data analysis techniques, such as those performed in SPSS, have had concepts data mining, machine learning (ML), and natural language processing (NLP) added to them. The plethora of analytical dimensions and the intricacies of big data and the complex patterns embedded have, to a greater extent, pivoted surrounding the systems approach and methodologies of SPSS applications. Python, R, and Modeler integrations have redefined and enhanced the functionalities of SPSS, which has led to the automation of analytical tasks in areas such as hypothesis testing and pinpointing predictive variables with high accuracy and rapid processing speed.

SPSS used to be a single, self-contained program, but now it is a modular analytics package that allows for connections with open source tools such as Python and R. Through a command script interface, users can access SPSS's point-and-click interface as well as write code. For instance, a user can perform a regression analysis on SPSS and, at the same time, use a Python script to visualize the residuals. This combines a click-and-go tool with a more sophisticated analytical environment and is therefore suitable for all levels of users.

3.4.9 Data Visualization and Interpretation

When data is analyzed, the stage that often requires the most interpretation is the one that requires the data to be visualized. While SPSS has a set of tools for creating various types of charts, researchers and users are now challenged to present visual data that does more than just communicate the results. Infographics, well-crafted visualizations, and dashboards facilitate the conveying of complex insights to people with little background knowledge. Modern data visualization tools focus on simplicity, clarity, and context. For instance, rather than just a histogram or scatter plot to communicate a correlation or distribution, adding trend lines, annotations, and notes can turn a figure into a persuasive analytical statement. Other more sophisticated visualization tools like Tableau, Power BI, and Matplotlib in Python can help improve the analytics reporting.

3.4.10 The Role of Artificial Intelligence in Data Analysis

Data analytics as a discipline has been made more efficient and effective through the use of technology such as artificial intelligence and machine learning. In the case of SPSS, driven AI modules are capable of identifying outliers, suggesting appropriate tests, and even reporting conclusions. For instance, IBM SPSS Statistics SmartAssist uses AI in system thinking and supports users throughout the entire research process completing the appropriate steps. Some machine learning tools such as decision trees, random forests, and neural networks provide additional insights that would not be obtainable through regression models. These tools provided additional insights, but researchers also must be careful. AI does enhance productivity, but the lack of expertise and critical thinking does raise, if not concern, some questions.

3.4.11 Data Ethics and Governance in the Digital Age

Ethics are becoming more relevant than they have been in the past and become more critical with the growth of digital data. Any researcher processing digital or personal data has to comply with ever more demanding privacy obligations and IRB restrictions, and take great care to respect the associated intellectual property rights. Any data analytics must be performed and documented with the principles of transparency, accountability, and fairness in mind. In quantitative research, this could result in the exclusion of p-hacking and selective reporting. In qualitative contexts, it means participants' voices must be accurately represented. More, as systems of AI are ever more applied to data analysis, ethics has taken on the additional challenges of data-centric algorithmic bias, algorithmic explainability, and data right of consent. An ethical researcher, therefore, has to find a way to compromise between change and propriety, proving that technology helps people more than it dispossess them.

3.4.12 Interdisciplinary Approaches and Mixed Methods Expansion

Most of the contemporary research problems cannot be defined solely within the borders of academic disciplines. In response to this, the importance of cross-discipline and mixed methods has increased. A social scientist exploring the phenomena of healthcare delivery, for example, could combine epidemiological data (quantitative) with patient narratives (qualitative). SPSS facilitates this integration by enabling the analysis of coded qualitative variables as part of a mixed data set. This enhances the richness of interpretation and contributes to a more holistic understanding of socio-human phenomena.

3.4.13 Triangulation and Reliability Enhancement

Robust and thorough pieces of research often rely on triangulation. Confirming an assertion or piece of evidence through multiple sources is a method used to strengthen the conclusion of a study while also mitigating bias. For a student engagement study, participants in the study can be observed, and their academic engagement and scores can be used as well as survey data that is gathered through questionnaires. SPSS assists this process through common identifiers, linking, and merging datasets from different phases of the data collection process. Reliability is enhanced through inter-rater agreement, test-retest, and internal consistencies such as Cronbach's Alpha. These varied approaches ensure that the data and findings are stable and replicable.

CHAPTER 4: Application of SPSS

4.1 Advanced SPSS Applications

SPSS also incorporate numerous advanced procedures that go beyond simple statistical tests of a complicated research design. These advanced procedures can include, but are not limited to, multivariate analyzes of variance (MANOVA), factor analysis, discriminant analysis, and structural equation modeling (SEM). Each of these advanced techniques enables the researcher to uncover hidden dimensions in data, test hypotheses that are composite or blended, and model associations that are causal. For instance, factor analysis is able to demonstrate hidden constructs under observed variables. This is immensely important for psychological and behavioral studies ever more so when SEM model evaluates several equations to determine indirect and direct effects in theoretical frameworks. The ability of SPSS AMOS to integrate visual modeling improves understanding as the researcher can visually construct and evaluate causal structural theoretical models, and then test them.

4.2 Data Interpretation

Looking at statistics and analyzing them tends to always result in achieving certain numerical outputs, but the interpreting is in the hands of the researcher. Synthesizing findings with literature, concepts and the surrounding context is a crucial part of interpretation. 'Social sciences' is a very wide field and in this case, one does not 'automatically assume' that there is a causation, correlation between income and education, the income and education has to be analyzed through the socioeconomic lenses with systemic inequalities. In the same way, experimental sciences' is very contextual, results cannot be separated from environmental conditions, sampling designs and potential confounders either. SPSS helps with the interpretation raw data and helps to draw a conclusion but there is always a point, in which there is a conclusion with the help of 'residual plots' which cannot be drawn without human reasoning.

4.3 Concluding Thoughts

Looking at the evolution of data from manual tabulations to Ai and how Ai has helped in decision making has shown both progress and responsibility. SPSS is still the preferred option to use, it is a tool that helps bridge the gap between the old and classical methods. Specialists who are trained to work with computer systems, SPSS and Ai in this modern climate now need to incorporate ethics, contextual knowledge and critical thinking. It is not always the case and the data does not always tell the very same story. We owe it to ai to help us reveal and uncover stories within the data that human beings would not be able to process.

4.4 Introduction to Measurement and Data Types

Analysis, interpretations, and even conclusions all rely on the data that has been collected. But in order for the data to be purposefully used, it needs to be showcased, scrutinized, and adequately defined. It involves measures and types of data which is vital for creating an initial framework for any analysis. This definition is very broadly translated in the social sciences and is crucial for studying the profound complexity of human behavior and perspectives and the social phenomena that constitute them.

As a researcher, performing a field survey, conducting qualitative analysis on a set of interviews, or even working with population national level data set, one works with the definition and the category the researcher has defined the work under. These definitions and categories greatly influence the outcomes and interpretability of the data set, even applied with higher-level statistical software like SPSS or simpler spreadsheets.

Defining what a measurement is remains the primary focus of this section. Assigning numbers (or Labels) to various items systematically is referred to as measurement. Without measurement, a researcher will have difficulty with abstract concepts such as attitude, satisfaction, or income. Measurement deals with the description of concepts. This translating process does not always provide a simple or direct solution, especially with concepts such as social attitude, social satisfaction, and income, which, in the social sciences, are elusive and subjective. This is where operational definitions, scaling, and the measurement of validity and reliability are necessary.

Measurement concepts go hand in hand with the step of describing types of data. Data, in its broadest classification, is qualitative (Non numeric) as well as quantitative (Numeric) and is subdivided into levels of measurement such as nominal, ordinal, interval, and ratio. These levels are significant as they impact what statistical computations and assessment techniques are to be used. For example, the calculation of a mean is not applicable to nominal and ordinal data. Similarly, the kind of chart or graph one uses, be it a bar chart, pie chart, or histogram, rests on the data being worked on.

This part underlines the importance of computer coding and data entry, especially geared toward SPSS. A social science researcher may work with a survey and need a constructed set of responses coded numerically for analytical purposes. At this stage, survey data type and the value coding systems at the answer set interviewers will use may be the first stage at which mistakes will be made which will affect the outcome of the whole piece of research. Hence the ability to correctly classify and code data should be a researcher's primary concern.

This section will, at the same time, address practical issues such as missing data, outliers, and data cleansing, which are almost a given for the first stage of transforming raw data into a usable dataset. Each topic will be treated with appropriate social science contexts, using spreadsheets and SPSS to ensure practical relevance for field research, class projects, and scientific dissertations.

To summarize, this segment lays the foundation for all the subsequent statistical processes. No amount of complex reasoning will affect the outcome of an analysis without a proper appreciation of the concepts of data being defined, measured, and categorized. This segment encourages the readers to interact with the 'language' of the data and use it freely, engaging with it in a way that will be critical in the subsequent chapters.

4.5 Scales of Measurement: Types, Characteristics, and Examples

Research starts by defining its key assumptions and concepts. In this stage, a researcher describes what the unique features of a study subject are and assigns values corresponding to the known variables to create, for example, a behavioral profile of a patient. Having an adequate understanding of an object is invaluable for collecting data, interpreting information, and forming conclusions. Because a study aims to measure the object of interest, several key questions must be answered. What type of the available data are we in possession of? What research questions are we able to address? In the social sciences,, and even in more advanced studies, variables are seldom quantitative, and the data themselves are often complex and nuanced in emotional content. Therefore, a researcher must address the two-fold question of what is to be measured, and more importantly, how?

Research is able to address the formulated questions only by using scales of measurement. These within social and behavioral sciences, and even in more advanced studies, are of four fundamental types: nominal, ordinal, interval, and ratio. What differentiates the scales is the level of precision applied in order to gather and analyze data. It can be said these are arranged from the most basic, to the most complex nuances.

There are four commonly accepted scales of measurement: Nominal, Ordinal, Interval, and Ratio. Each of these scales, described in the order of their complexity and analytical usefulness, increases in order of mathematical and ordinal complexity with nominal being the simplest and ratio the most advanced.

1. Nominal Scale

Definition: The nominal scale is the most basic level of measurement. It involves labeling variables without any quantitative value. The categories are mutually exclusive and collectively exhaustive, but there is no inherent order.

Examples:

- Gender: 1 = Male, 2 = Female, 3 = Other
- City: 1 = Ghaziabad, 2 = Noida
- Marital Status: 1 = Never Married, 2 = Married, 3 = Divorced

Key Characteristics:

- Numbers are used as labels, not quantities.
- Cannot be ordered or ranked.
- Only mode and frequency distribution are meaningful.

Social Science Relevance: Many demographic variables are measured nominally. In SPSS, these are coded numerically (e.g., Male = 1), but their interpretation remains categorical.

2. Ordinal Scale

Definition: The ordinal scale represents order or rank among categories. While we know one category is higher or lower than another, we do not know by how much.

Examples:

- Education: 1 = Upto 10th, 2 = HSC, 3 = Graduate, 4 = Postgraduate
- Age Group: 1 = Below 20, 2 = 21–30, 3 = 31–40
- Banking Frequency (BQ1–BQ9): 1 = Never, 5 = Always

Key Characteristics:

- Categories have a logical order.
- The distance between categories is unknown.
- Median and mode can be used; mean is inappropriate.

Social Science Relevance: Ordinal data is common in attitude and opinion surveys. Likert scales are classic ordinal measures used extensively in behavioral studies.

3. Interval Scale

Definition: Interval scales not only show order but also the exact differences between values. However, they lack a true zero point.

Examples:

- Satisfaction Ratings on Scaled Responses (if equal intervals assumed)
- Perception Scores (e.g., from 1–5 on security or convenience when treated as interval)

Key Characteristics:

- Equal intervals between units.
- No absolute zero (zero doesn't imply "none").
- Arithmetic operations like addition and subtraction are valid.
- Mean, median, and mode can be used.

Social Science Relevance: Interval data is used in perception, attitude research, and some scaled scoring assessments when equal distances between scale points are assumed.

4. Ratio Scale

Definition: Ratio scales have all the features of interval scales, plus a true zero point, which allows for the expression of absolute quantities and ratios.

Examples:

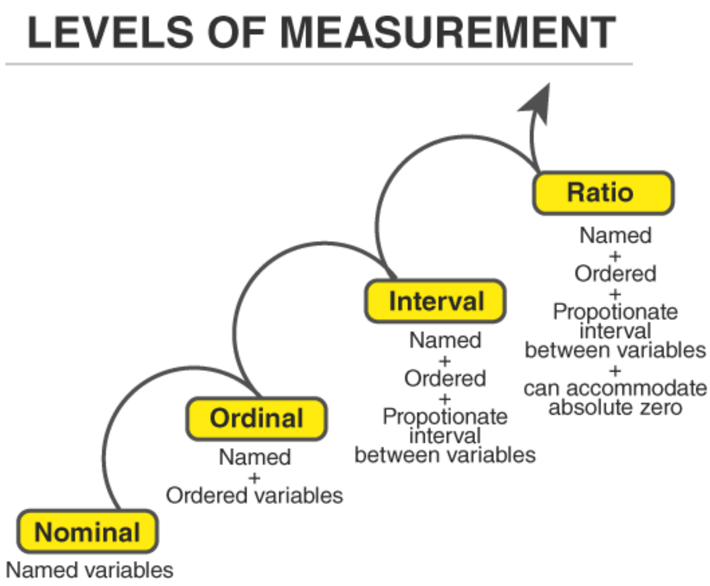
- Family Income: Actual values (e.g., ₹25000, ₹50000)
- Hours Worked: Total weekly working hours
- Number of Dependents

Key Characteristics:

- True zero exists, meaning absence of the property.
- All mathematical operations are possible (multiplication, division).
- Mean, median, mode, standard deviation are all valid.

Scale	Definition	Examples	Key Characteristics	Social Science Relevance
Nominal	Categorizes data without any order or numeric meaning.	Gender (Male/Female/Other), City (Ghaziabad/Noida), Marital Status	- Labels only (e.g., 1=Male) - No ranking/order - Only mode and frequency are meaningful	Demographic variables coded categorically in SPSS
Ordinal	Categorizes with a meaningful order but unequal intervals.	Education Level, Age Groups, Likert-style banking frequency (BQ1–BQ9)	- Rank/order present - Unknown interval sizes - Median and mode are valid	Common in Likert scales and attitude surveys
Interval	Ordered categories with equal intervals, but no true zero.	Perception Scores, Satisfaction Ratings (if assumed interval)	- Equal intervals - No absolute zero - Can add/subtract - Mean, median, mode valid	Used in perception and scoring research when assumptions are met

Ratio	Highest level; ordered with equal intervals and a true zero.	Family Income, Hours Worked, Number of Dependents	- True zero exists - All arithmetic operations valid - Full statistical analysis possible	Critical for quantitative research (e.g., income distribution)



4.6 Social Science Relevance:

Many economic and demographic variables, such as income, expenditure, and age, are measured on a ratio scale, especially in quantitative research.

4.6.1 Why Measurement Scales Matters in Data Analysis

The scale of measurement determines:

- What kind of statistical tests can be performed.
- The type of graphs or charts to be used.
- The appropriate measures of central tendency and dispersion.
- The interpretation of outcomes and the validity of conclusions.

As an example, running a Pearson correlation on nominal data is a case of statistical blunder that could lead to erroneous conclusions. Just as, calculating a mean satisfaction score from ordinal Likert data assumes equal spacing, an assertion which, in base, is unproven.

Knowing how to differentiate between types of measurement scales is crucial for any researcher, particularly for an avid social researcher where a majority of the concepts are abstract. Understanding if the variable is nominal, ordinal, interval, or ratio determines every aspect of data analysis, be it coding and data entry or even running statistical tests in SPSS or excel spreadsheets. While proper measurement is crucial for statistical accuracy, it is also necessary for profound and ethically transcendental interpretive depth in one's research.

Table 1. Measurement Scales in Research: Characteristics, Permissible Mathematical Operations, and Illustrative Examples

Scale	**Characteristics**	**Mathematical Operations**	**Examples**
Nominal	Categories only	Count, Mode	Gender, Religion
Ordinal	Ordered categories	Median, Rank-based tests	Education Level, Satisfaction
Interval	Ordered, Equal intervals	Mean, SD, Correlation	Temperature (°C), IQ Scores
Ratio	All interval features + 0	All statistical operations	Income, Age, Weight

Table 2. Statistical Properties and Permissible Operations Across Levels of Measurement (Nominal, Ordinal, Interval, and Ratio Scales)

Provides:	**Nominal**	**Ordinal**	**Interval**	**Ratio**
The "order" of values is known		✔	✔	✔
"Counts," aka "Frequency of Distribution"	✔	✔	✔	✔
Mode	✔	✔	✔	✔
Median		✔	✔	✔
Mean			✔	✔
Can quantify the difference between each value			✔	✔
Can add or subtract values			✔	✔
Can multiply and divide values				✔
Has "true zero"				✔

CHAPTER 5: Types of Research in Social Sciences

5.1 Coding and Classification of Data

It is in research, and especially in social research, that the raw data obtained from surveys, interviews, focus groups, or direct observation, is changed to a structured format which can be analyzed. This changed format goes through a transformation which consists of two critical steps: coding and classification. In order to use and statistically evaluate qualitative and quantitative data through software, spreadsheets, or complex statistics, it has to be thoroughly managed and structured.

As with qualitative or categorical information obtained through open-ended queries and descriptive interviews, it is vital with respect to the qualitative data collected to properly analyze and assign codes to retrieve answers.

Purpose of Coding:

- To convert raw data into a format suitable for computer-based analysis.
- To reduce data volume and complexity.
- To allow for the quantification of qualitative responses.

Types of Coding:

- **Pre-Coding**: Codes are predetermined during the design of the questionnaire. This is common for closed-ended questions.

  Example:

  "Gender: Male = 1, Female = 2, Other = 3"
- **Post-Coding**: Codes are developed after data collection, particularly for open-ended responses.

  Example:

  Responses to "What is the biggest challenge you face at work?" may be coded into categories like:
  - 1 = Low Salary
  - 2 = Lack of Recognition
  - 3 = Excessive Workload

Coding Qualitative Data:

In qualitative research, themes or categories are derived from text data using techniques like:

- **Thematic coding** (based on recurring themes),
- **In vivo coding** (using participants' actual words),
- **Descriptive coding** (using topic-based labels).

Modern software like NVivo or MAXQDA is often used for advanced qualitative coding, but for simpler projects, spreadsheet programs can also be used effectively.

Classification involves organizing data into categories or classes based on shared characteristics. It is an essential part of data reduction and facilitates comparison and summarization.

Purpose of Classification:

- To group data meaningfully.
- To simplify the complexity of raw responses.
- To identify patterns, trends, and anomalies.

Types of Classification:

- **Geographical Classification**: Grouping data based on location or region.
  - *Example*: States, countries, urban vs rural.
- **Chronological Classification**: Organizing data based on time.
  - *Example*: Year-wise population growth.
- **Qualitative Classification**: Based on attributes that cannot be measured directly.
  - *Example*: Religion, occupation, political affiliation.
- **Quantitative Classification**: Based on measurable attributes.
  - *Example*: Income levels, age groups.

Classification Example in SPSS:

Suppose you have data on monthly income. You might classify it into:

- 1 = Below ₹10,000
- 2 = ₹10,000–₹30,000
- 3 = ₹30,001–₹60,000
- 4 = Above ₹60,000

This classification allows for easy cross-tabulation and descriptive analysis in SPSS.

Importance in Social Science Research

In social sciences, data often comes with diversity and nuance. People may use different words for the same idea or express multiple thoughts in a single response.

Coding and classification help bring uniformity to this variety, enabling researchers to analyze social behavior, attitudes, and trends in a scientifically valid way.

Furthermore, incorrect or inconsistent coding can lead to **invalid results**, misleading statistical summaries, or incorrect conclusions. Hence, **transparency, consistency, and documentation** of the coding process are essential.

Best Practices:

- Use a **coding manual or codebook** to maintain consistency.
- Train all data handlers to apply codes uniformly.
- Allow for **multiple coding** when responses span more than one category.
- Revise codes as new patterns emerge, especially in qualitative research.

Coding and classification are not mere clerical tasks but intellectual processes that shape the foundation of data analysis. They require attention to context, consistency, and the research question. In the context of social science research, where data is often messy and subjective, these steps bridge the gap between raw human experience and structured knowledge. Whether using a spreadsheet or statistical software like SPSS, effective coding and classification lead to clearer insights and more reliable results.

Likert Scale

Likert Scale is one of the most popular and most widely used scaling techniques in social science research for the measurement of the attitudes, opinion, perception and beliefs. In 1932, Rensis Likert devised a means for the systematic assessment of respondents' degree of agreement and/or disagreement with a collection of statements.

Typical Structure

An average Likert scale is known to have 5 and/or 7 response options, as such:

Strongly Disagree

Disagree

Neutral

Agree

Strongly Agree

The scale's response options are an even distribution wherein one half facilitates a positive response and the other a negative response in a bid to minimize response bias.

Key Characteristics

Measures attitude a respondent has along with the corresponding degree, and is not simply a one-directional response.

Responses in this measurement are ordinal in nature but are by many practitioners of social science treated as interval data.

The Likert scale is relatively easy to create and is popular in the production of surveys and other questions.

5.2 Debunking Common Misconceptions in Data Analysis

Navigating through a new field, especially one as intricate as data analysis, comes with its own set of challenges. Among the challenges a newbie data analyst would face, are the hurdles of misconceptions. Unquestioned and accepted, these misconceptions would set the analyst on a troublesome course which would lead to inappropriate, and incorrectly derived conclusions. Let's break these myths to paradoxically 'set the record straight' with best practices that guide one sound analytics.

5.2.1 The Illusion of Problematic Verses: Big Ideas and Complicated Problems

It's a desirable notion that any intricate statistical exercise, involving complicated algebras, advanced concepts, and exorbitant terms, would be elegantly solved and graded appropriately. One might think, "If it's impossible, then it's a prize-winning master class." Well, the notion of 'master class' could easily be replaced with '$h*t show' where everybody is a 'player' and the prize is self-sabotage. This is a carefully phrased accusation the accusation because, guess what!? Colleagues, clients, and the general audience are all as numb to sense as the next. The focal point is the practicality of the analysis, and the primacy which is given to it. 'Emphasis' is a delicate one since too much would lead to undue complex cloud of analytics that causes destruction. Your audience is not so dumb and would be able to identify too much 'fluff' as a failure on your part too. The expert senior analyst which retains work that has taken others a lifetime, is able to cheerfully cut through mess and retrieve the needed data. Concentrate on your data, and sharpen your apprehensive thoughts which reflect your analysis and interpretations. Less is usually more, so the best answer with the false notion of 'fluff' is to think, 'clarity, make a point, and leave it."

5.2.2 The Afterthought Trap: Analysis as a Final Step

The easiest way, though incorrect, way to think about how data is collected is that it is a slab of data that is somehow then thoroughly processed and analyzed in a much later stage. Basing evidence collection and evidence analysis in a sequential, time-based relationship is highly inefficient and restrictive. Evidence collection and analysis should be treated as a more simultaneous process, and in that, should focus on the analytical questions more strongly. Framing the questions and coming up with possible analytical approaches during the planning of data collection will heavily dictate the type of data to be collected, the collection process, and the most important factors to examine. This will ensure that evidence collection is targeted, relevant, and insightful. Instead of waiting to the end of the research process, core analytical questions that will assist in the reasoning of the collected evidence should be already made, so the research can assist in the analytical framework as more structure is provided.

5.3 The Gap Data Fallacy: Thinking That Data is Self-Explanatory.

What raw data entails, more so when data is arranged in orderly columns and rows, is clear and objective. It is, however charged, and in as much as it is, one must note the saying: Data and numbers lie. Data is hardly self- explanatory and must be clarified. Note that figures and fantasies are simplistic representations and what is more important is how the data shapes the theoretical and practical context as a, why is the data originating, and what techniques is utilized in the. Analyze the data. Different data sets are drawn upon the same outcome and arrive at differing conclusions. It is one’s duty as the researcher, and more so is be a storyteller. Use the data cordinates to form a systematic and coherent structure to answer the research objectives and emphasis the text, Leaving space for emphasis. Emphasis. Insightful analysis is integration of context and reconciling reasoning.

5.3.1 The Weakness Worry: Fear of Stating Limitations

The fear of losing credibility by accepting the limitations of one's work is a concern for any analyst. The common fear is that the focus on the weaknesses will outweigh the strength of the concluding statements. However, the reality is that every analysis has underlying limitations. These limitations can be derived from the data, the analytical methods used, or the scope of the analysis. It is a willingness to accept, rather than an inability to acknowledge all the facts, that is an intellectually honest and responsible thing to do.

Framing a limitation of your analysis strengthens your evaluation by demonstrating a clear, critical self-appraisal. By stating what your analysis can and cannot tell us, trust is strengthened and a foundation for more informed interpretation and future research is set.

5.3.2 The Technological Panacea: Assuming Computers are Always Superior

Assuming that computer-based analysis is easier and better is a common phenomenon in the era of sophisticated statistical software. While technology is undeniably a potent solution to gigantic datasets and intricate calculations, its merit is situational. It is highly dependent on the volume of data available, and more importantly, your competencies. Quite often, the basic analysis of a smaller data set can be performed with hand calculations faster than it can be done with a computer, and it can also lead to a better understanding of the data and its relevance. Also, computer-based analysis is as effective as the analyst permits the computer to think, the techniques that the analyst selects to apply, and how accurately the analyst interprets the output. Resulting conclusions from advanced tools and techniques of analysis performed by a user lacking in analysis skills are bound to be erroneous and deceptive. The essence of the matter is adopting the analysis methodology that is a perfect fit for your problem and skills. As the reality of the matter is, having too much data is not the only barrier that can be too simplistic in explanation.

5.4 Tools to Support Data Analysis:

Data analysis has changed from cumbersome tasks such as counting and making a variety of graphs to more sophisticated forms of analysis. This transformed approach still integrates the use of devices that assist in making the process more precise. The nature of the data from a survey, a set of experiments, or direct observations helps define the tools to be used in order to arrive at the desired level of detailing. This scenario becomes more complicated in the case of social sciences, and the intersecting quantitative and qualitative data. The researcher can start off from ready-made tools such as worksheets in Microsoft Excel or Google Sheets and the more advanced systems such as SPSS. Data analysis is the process of collecting, coding, classifying, visualising, and computing the data statistically. The researcher, therefore, must have a strong understanding of the tools or devices to be able to detect and interpret patterns, and establish a logical and rational sequence. This chapter outlines the major tools used for data analysis, detailing their attributes and functions, as well as features and functionality tested in practice.

5.4.1 Computer Programs: Spreadsheets

In a research setting, raw data is not sufficient enough. Processing data and analyzing it is the only way to turn it into actionable intelligence. These systems digital tools helps in this transformation. Examples include Microsoft excel and Google sheets and the statistical software such as SPSS. This segment focuses on tools to spreadsheets on and their importance in data entry, preparation, exploration, and the very first analysis in the social science domain.

To new freshers in the research field or mesmally seasoned professionals these tools are an important part of their anayltics tool chest and often the first port of call in an analysis. Because of exra A. which are shown in the assists mkre might help in supporting features it is not only the data entry which simplifies processes and also help data to be strong summarized.

A. Pivot Tables

Pivot tables devices are one of the most useful tools in the context of word-processing programs, however, they are one of the least taught. It enables researches to simplify extensive datasets rapidly by categorizing variables in meaningful ways. For example, a survey researcher with 500 respondents can create a Pivot Table to calculate how many respondents were male, female and the various ways income groups are structured by age or education.

With the help of a simple drag and drop interface, people can and are able to transform the arrangement of the rows, the columns and the values in a data set wherein they can change the view of the data. This enables them to analyze and view far more perspectives without ever having to change the originial data set. This component is in particular useful in social science data research when intra-group comparison is vital for the proper understanding of the results, such as when differentiating the members of a group by gender, residence, or social class.

B. Charts and Graphs

Visualizing data is crucial for making sense of trends, patterns, and outliers. Spreadsheet tools offer a wide range of **visualization options**, including:

- **Bar and Column Charts** for categorical comparisons,
- **Pie Charts** for proportional data,

- **Histograms** for distribution analysis,
- **Line Graphs** for tracking changes over time,
- **Scatter Plots** for examining correlations.

One of the key advantages of using charts in spreadsheets is their **dynamic linkage** to the data. Any change made to the source data automatically reflects in the corresponding chart, ensuring consistency and reducing the risk of misrepresentation.

These visual tools are not only helpful in exploring data but also in **communicating findings** in reports and presentations, making complex data easier for stakeholders to understand.

c. Filters and Sorting

Filtering and sorting tools in spreadsheets assist researchers in organizing their da a in meaningful ways. Being able to filter data allows for more selective ways of visualizing data. For instance, a researcher can filter responses to only those data from participants aged forty and above, or from a particular location.

Sorting options assist in trend and outlier identification through arranging data in either ascending or descending for example, ranking participants in tests or participants in a region from the lowest to the highest in household income. Setting lower thresholds particularly aid sr udies in social sciences to define the minimum indicators of a participant who may require the most intervention, or the reverse, define the participant who may require the least intervention in their daily activities.

All these functions save time and effort in manual reviewing of data.

d. Statistical and Mathematical Functions

Excel, along with many other spreadsheet tools, come with a variety of built-in functions which aid researchers in performing more advanced. preliminary quantitative analyses.

• AVERAGE(), for instance, as well as the other central tendency measuring functions such as MEADIN() and MODE(), aid in measuring central tendencies.

• Population and sample standard deviations can also be computed with STDEV.P() and STDEV.S() respectively.

• Also, the standard deviations of the sample and population can be computed using the correlational function of two variables with a sample population using the formulae CORREL() Pearson for two continuous variables.

• In addition, the COUNTIF(), SUMIF() and IF() functions help to apply basic logical and quantified mathematics using set definitions.

• Functions like AND, OR, and other nested formulas allow for programmatic data filtering and segmentation.

These functions, while basic compared to what statistical software can do, satisfy the requirements for descriptive statistics and initial hypothesis testing. They act as a precursor before data undergo more advanced procedures in platforms like SPSS for deeper inferential analysis.

Spreadsheet applications are critical in the course of data analytics. Their ability to blend simplicity and supercomputing power equips them for research data processing at the early phases of a study. For data entry, filtering, summary statistics and even visualization, spreadsheet applications like Excel and Google Sheets enable researchers to perform and present reliable preliminary analysis in a transparent manner. Later in this book, we will use these skills, in more complex SPSS based analyses.

CHAPTER -6: Statistical Packages – SPSS (Statistical Package for the Social Sciences)

Voting the most celebrated programmes for statistical analysis in social sciences, SPSS (Statistical Package for the Social Sciences) was created as one of the very first computer programmes in the world by IBM. SPSS provides a quick means of managing and analysing large and complex datasets. Professionals in various fields, such as psychology, sociology, education, public health, and even political science, find SPSS helpful for analyzing both qualitative and quantitative data. Its ease of use and capability of performing complex operations result in its wide-reaching popularity.

Perhaps the most important feature of SPSS is its ease of use in analysing datasets, even for large complex datasets. SPSS provides a steady framework for analysis, making every table, every input box for cell values and every button a 'step of the data collection process'. For instance, in the table, each row is a data 'case' and, every column is a data 'variable'. Thus, inputs in a data column can be age, income, or even education.

The other strong feature of SPSS is the dual operation system. It has a Graphical User Interface (GUI) system that allows individuals to carry out analyses by point-and-click menu and dialog box retrieval options, which is useful for individuals that have zero programming or coding skills. SPSS also has a system that allows writing and executing of code for analyses. It is referred to as a syntax system. It comes in very handy when complex data transformations or repetitive tasks are to be done, for in SPSS, code-based operation will guaranty, reproducibility, effectiveness, and accuracy in every performance of the operation.

SPSS provides for a very wide and comprehensive range of complex statistical analyses. It has excels in doing Descriptive Data Statistics as a result of its capability to calculate measures of central tendency (mean, median, mode), measures of dispersion (standard deviation, variance, range) and, as well, frequency distributions. Inferential Statistics is very easy to carry out because SPSS makes easy the performance of t-tests, analysis of variance (ANOVA), simple and multiple regression, as well all other forms of regression, correlation, as well as chi-square tests. More complex users of SPSS are able to enjoy the package for its various versatile tools which include, factor analysis, discriminant analysis, cluster analysis and all other Non parametric tests.

One of the most appreciated features of SPSS is its ability to produce output that is ready for publication. The results appear in well-organized tables and illustrated displays, which can be exported to Word, Excel, and PDF for further use in reporting and presentation. These outputs are valuable for researchers while authoring a manuscript or composing a report, as they save a lot of time.

SPSS is especially adept in analyzing categorical (nominal and ordinal) and continuous (interval and ratio) variables. It allows researchers to properly conceptualize their variables according to the measurement levels, which is key for interfacing with the appropriate statistical methods. In SPSS, the variable view allows the user to set variable properties such as labels, types, measurement levels, and missing value definitions which enhances the value of the data in question and promotes clarity.

Ultimately, SPSS continues to be a fundamental software for social science researchers, and this is due to the tendency that SPSS has in terms of simplicity and the depth of analysis of data. SPSS enables researchers to find patterns, corroborate ideas and hypotheses and gain thorough insights with an ample amount of confidence, and this is applicable in academic work, institutional studies and surveys, or even in the analysis of public policies. The tools and methods in SPSS are crucial for researchers who pursue comprehensive and impactful social science studies, especially as inquiry becomes more reliant on data.

6.1. Coding and Classification of Data

The assigning of codes and classifying data into specific categories serves as a very important preliminary step during the preparations of data for the purposes of analaysis. There is a significant distinction between the processes of “collecting data” and “preparing data for analysis.” no matter what means of data analysis are employed during the analysis of the data, be it through the usage of spreadsheet applications or the employment of advanced statistical applications like SPSS and the rest, the data “collected” through the process of surveys, interview, experiments or through the secondary sources is raw and has to be organized systematically. This particular step ensures the data is properly formatted so the analysis process that is to be employed is logical. This chapter focuses on the two primary phases of data entry and data coding, narrowing down to their implementation within Excel and SPSS.

A. Data Entry in Excel

Program applications like Microsoft Excel and Google Sheets are usually the first programs that are used to enter data into the computer and clean the data due to the ease of use, popularity and their clear layout. But, data entry techniques that are fundamental to proper data entry, must be employed so the data can be used for statistical analysis.

Every Row represents a Unique Case/Participant: In each dataset, every row must represent a unique case, single participant or observation. Take for instance the data is being collected by a researcher and a 100 survey respondents, the Excel spreadsheet will be 100 rows and each of the rows will correspond to a certain individual set of data based on their responses. Such a representation of data is referred to row-wise representation, and it is important for the case-based data to be maintained.

One Column per Variable: Each of the columns must pertain to a single value or property such as Age, Gender, Education Level, or Income. The titles of columns should be short yet informative and preferable without spaces or special characters. Age should be used instead of Respondent's Age for example.

Consistency in Format: The values in a column of data should be presented in the same way. For example, if in "Income" some entries are text like "20,000 INR" or "Twenty Thousand", it must be noted that there is a standard approach to income values. They should be numeric. Do not use merged cells, blank rows, or data of a column that is of mixed types like text and numbers, as these can cause problems when importing into SPSS or other programs.

No Special Formatting or Special Style. Although the clearness of a text may be improved by color, it can never be defended in a statistical estimate. The one and only indicator of the data must be the numbers, without decorative supplementary matters. Also, make sure that there are no internal subtotal or total rows in the dataset.

6.2 Ethical Considerations in Data Handling

Ethical practices ensure the credibility and trustworthiness of research findings. Researchers must handle data responsibly at every stage, from collection to reporting.

- Confidentiality – Protect respondent identities by using coded IDs.
- Transparency – Maintain detailed codebooks and documentation.

- Informed Consent – Clearly inform participants about data usage.
- Avoid Data Manipulation – Never alter data to fit expectations.
- Reproducibility – Archive datasets and analysis syntax for verification.

Ethical data management reinforces academic integrity and protects the credibility of the research institution.

6.3 Concluding Thoughts

When conducting research, understanding the nature and principles of measuring data does not merely come before the actual process of research, but serves as the integral component upon which rests the foundation of all the analysis which follows. How well the results of the research are defined, how well the research prepared the variables, and how well the variables are measured dictates how reliable, credible, and interpretable the statistical research results are.

This chapter has examined the data and the broader ideas which accompany it as well as the more practical aspects of it, especially in the context of the social sciences research which is being undertaken. It started from the key notion that data, in isolation, is of minimal value unless it is appropriately categorized and measured. The ordinal, nominal, interval, and ratio measurement types helped us understand the primary measurement types, how each scheme control the statistical processing that can be done, and what each arrangement enables in data processing. A well defined measurement framework does not solely relieve the researcher from methodology issues but also brings value and granularity to the analysis.

We focused on measurement as capturing truth, and the need to revisit the idea of validity and reliability. A valid instrument is one that measures the intended construct. A reliable instrument is one that achieves measurement stability and dependability across diverse conditions and time periods. Both aspects of measurement carry value, which SPSS helps researchers estimate in order to test their instruments before they actualize conclusions.

The sections on coding, transformation, and data cleaning showed how the raw data are responses which are mutable and structured for analysis and examination. These are not rote processes; they are cognitive activities that demand critical consideration to ascertain that the data are true to the data realities. Just as crucial are ethical dimensions of data that warn scholars that every digit in the folder represents human reactions, experiences and observations which are worthy of dignity, privacy and ethical treatment.

In closing this section, it is evident that the competencies one possesses in the measurement and classification of data are not the results of a learned technical process. It is a prism through which one's level of maturity in research is evaluated. Such a researcher stands in a better position to make appropriate choices in analytical processes, draw reliable conclusions, and play an active role in the creation of new knowledge as a result of having mastered the language that data speaks.

The foundational principles articulated in previous chapters will inform the use of statistical methods such as correlation and regression as well as hypothesis testing. Each of these approaches is designed within the architectural framework of measurement and data types described here in order to make the transformation of data into insights and the insights into understanding that is well informed.

6.4 Preparing Data for Analysis Using SPSS

When it comes to collecting data, it is just the start to the broader concept of research. The follow-up step that occurs is just as important and that is to get the data ready to be analyzed in a form that is useful for them. This stage of the research is very important for any work that is done in SPSS as it requires the coder to break and organize the data into a framework as well as a code that is appropriate for the program to work on.

This part of the book intends to provide the reader with information on how to get ready data for analysis in SPSS and is not concerned with the origin of the data which could encompass the results of online surveys, interviews, field notes, or institutional data bases but irrespective of the data collection methodology this constant approach to data preparation. the rigorous, systematic analysis and interpretation which is from ascertained, quantified, and rational definitions of the phenomena analyzed is rendered in the most ethical manner.

The ease in which the data is incorporated into the SPSS and the quick result output of the statistics produced makes it a very useful program in the field of online social science research. This ease of access is coupled with the SPSS's ability to work with very large data sets that need simple statistics as well as very large data sets with inferential models. the data in order to extract value from the constructible SPSS output which confirms to a user's request for target results, it is imperative that a user fulfills the necessary and sufficient conditions of the program.

6.4.1 Introduction to SPSS

SPSS (Statistical Package for the Social Sciences) is a software that is heavily utilized for statistical activities in the social sciences, education, healthcare, and commerce. Due to its advanced statistical capabilities coupled with user-centered interface, interface SPSS assists researchers and analysts in data management, analysis, and visualization.

This section describes the start-up activities with SPSS, the main workspace, a brief description of a dataset, and core functionalities of the software. The capabilities of SPSS are crucial for data analysis computing whether the data is entered manually or retrieved from other applications.

Launching SPSS

When you launch SPSS, the **Welcome Dialog** box appears. This initial screen offers several options:

- **Type in data** – To create a new dataset manually by entering data directly into SPSS.
- **Open an existing data source** – To open a pre-existing dataset (e.g., from Excel, CSV, or another SPSS file format).
- **Run an existing query** – To run a previously saved analysis.
- **Open another file format** – To import data from other file formats, such as Excel, CSV, or SAS.

For first-time users, selecting **“Type in data”** is typically the starting point to create a new dataset from scratch.

6.4.2 Creating a Dataset in SPSS:

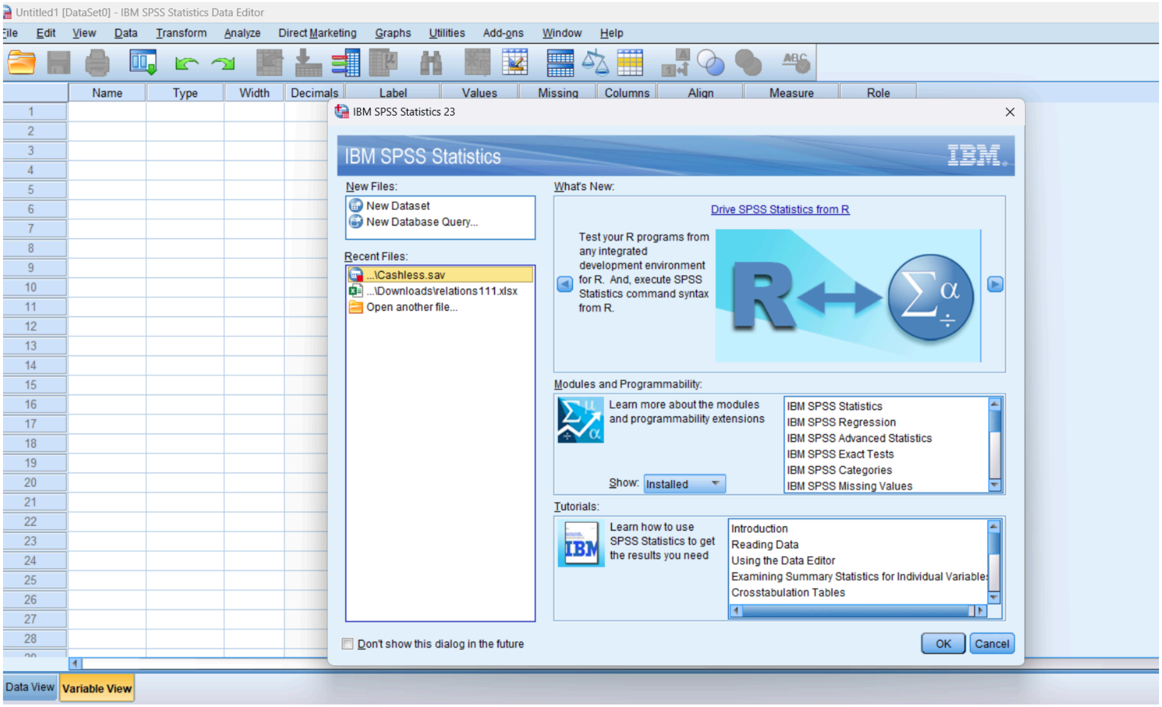
Whenever one starts analyzing data on SPSS Software, one of the most critical components in the process is constructing or accessing an existing data set for use. SPSS helps in defining the individual components which can take on certain values. Definitions alone maybe insufficient. It is important that the analysis to be performed is in sync with the structure of the data as well. SPSS-Statistical Package for the Social Sciences sorts the data into Two Main Views. These Views of the data are for the Variable View and the Data View. An entire analysis can go badly if the user does not learn how to use these Views.

Starting a New Dataset in SPSS

When you open SPSS, you are greeted with the “SPSS Statistics Data Editor” window. This looks somewhat like a spreadsheet but is tailored specifically for statistical data.

- **To start a new dataset:**
    - Open SPSS.
    - From the welcome window, select **“New Dataset”**.

- This opens a blank spreadsheet-like window with two tabs at the bottom: **Data View** and **Variable View**.



6.4.3 SPSS Interface Overview

Once SPSS is open, you will be presented with two primary views for interacting with your data:

a) **Variable View**

The **Variable View** is where you define and configure each variable in your dataset. In this view, you can set the name, type, label, measurement level, and other properties of the variables. It's essential to correctly define your variables in this view before conducting any analyses.

b) **Data View**

The **Data View** is where you enter and view your dataset. It resembles a spreadsheet where each row represents a case (participant, observation, etc.) and each column corresponds to a variable (e.g., age, gender, income). This is where you can manually input data, import data from other sources, or edit existing data.

6.4.3.1 Defining Variables (Variable View)

After defining your variables, switch to the **Data View** to start entering your data. Each row will represent an individual case or participant, and each column will represent one of the variables you've defined.

These attributes and the overall structure of the dataset are set even before any data is entered or analyzed. Each row in this view corresponds to a single variable. Many attributes are assigned to each row for clarity, consistency, and data handling accuracy. The always visible Name field is used for the Auto Sviak variable for the record, and the variable will be something like Age, Gender, or Income. Type describes whether the variable is numeric, text (string), date, or something else. Label allows a more detailed explanation of the variable to make it more user-friendly for analysis, more specifically the output tables and graphs. The Values column is important for categorical variables which consist of numerical codes assigned a specific meaning. For instance 1 = Male, 2 = Female. This is done for purposes of efficient statistical processing and minimization of interpretational issues. The Measure attribute attributes the level of measurement to the variable like Mearured (categories, no order), Ordinal (ranked), and Scale (numeric, continuous). The added attributes Decimals, Missing Values , and Column Width enables fine tuning data to facilitate the correct statistical interpretation of the data. Thus, is Variable View is critical in preparing a dataset that is clean, organized and ready for analysis.

In Variable View, you will define the structure of your dataset. Each row represents a variable, and you can assign attributes to them such as:

Name – A short, descriptive name for the variable (e.g., Age, Gender, Income).

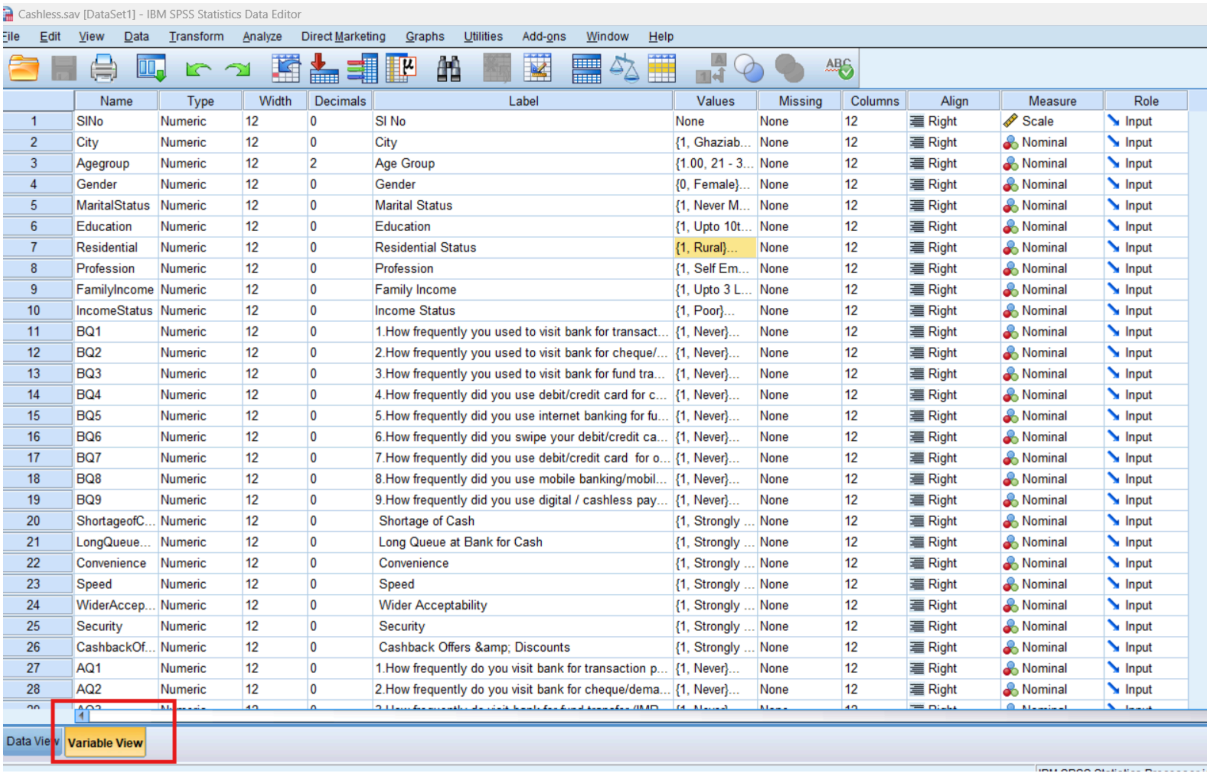
Type – The variable's data type (e.g., Numeric, String).

Label – A longer, descriptive label for the variable (e.g., "Respondent's Age").

Values – This field is used to assign numeric codes to categorical variables (e.g., 1 = Male, 2 = Female for gender).

Measure – The measurement scale of the variable (Nominal, Ordinal, Scale).

Click on the **"Variable View"** tab. This is where you define the structure and characteristics of your variables. Each row corresponds to a variable, and each column describes properties of that variable.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	SINo	Numeric	12	0	Sl No	None	None	12	Right	Scale	Input
2	City	Numeric	12	0	City	{1, Ghaziab...	None	12	Right	Nominal	Input
3	Agegroup	Numeric	12	2	Age Group	{1.00, 21 - 3...	None	12	Right	Nominal	Input
4	Gender	Numeric	12	0	Gender	{0, Female}...	None	12	Right	Nominal	Input
5	MaritalStatus	Numeric	12	0	Marital Status	{1, Never M...	None	12	Right	Nominal	Input
6	Education	Numeric	12	0	Education	{1, Upto 10t...	None	12	Right	Nominal	Input
7	Residential	Numeric	12	0	Residential Status	{1, Rural}...	None	12	Right	Nominal	Input
8	Profession	Numeric	12	0	Profession	{1, Self Em...	None	12	Right	Nominal	Input
9	FamilyIncome	Numeric	12	0	Family Income	{1, Upto 3 L...	None	12	Right	Nominal	Input
10	IncomeStatus	Numeric	12	0	Income Status	{1, Poor}...	None	12	Right	Nominal	Input
11	BQ1	Numeric	12	0	1.How frequently you used to visit bank for transact...	{1, Never}...	None	12	Right	Nominal	Input
12	BQ2	Numeric	12	0	2.How frequently you used to visit bank for cheque/...	{1, Never}...	None	12	Right	Nominal	Input
13	BQ3	Numeric	12	0	3.How frequently you used to visit bank for fund tra...	{1, Never}...	None	12	Right	Nominal	Input
14	BQ4	Numeric	12	0	4.How frequently did you use debit/credit card for c...	{1, Never}...	None	12	Right	Nominal	Input
15	BQ5	Numeric	12	0	5.How frequently did you use internet banking for fu...	{1, Never}...	None	12	Right	Nominal	Input
16	BQ6	Numeric	12	0	6.How frequently did you swipe your debit/credit ca...	{1, Never}...	None	12	Right	Nominal	Input
17	BQ7	Numeric	12	0	7.How frequently did you use debit/credit card for o...	{1, Never}...	None	12	Right	Nominal	Input
18	BQ8	Numeric	12	0	8.How frequently did you use mobile banking/mobil...	{1, Never}...	None	12	Right	Nominal	Input
19	BQ9	Numeric	12	0	9.How frequently did you use digital / cashless pay...	{1, Never}...	None	12	Right	Nominal	Input
20	ShortageofC...	Numeric	12	0	Shortage of Cash	{1, Strongly ...	None	12	Right	Nominal	Input
21	LongQueue...	Numeric	12	0	Long Queue at Bank for Cash	{1, Strongly ...	None	12	Right	Nominal	Input
22	Convenience	Numeric	12	0	Convenience	{1, Strongly ...	None	12	Right	Nominal	Input
23	Speed	Numeric	12	0	Speed	{1, Strongly ...	None	12	Right	Nominal	Input
24	WiderAccep...	Numeric	12	0	Wider Acceptability	{1, Strongly ...	None	12	Right	Nominal	Input
25	Security	Numeric	12	0	Security	{1, Strongly ...	None	12	Right	Nominal	Input
26	CashbackOf...	Numeric	12	0	Cashback Offers & Discounts	{1, Strongly ...	None	12	Right	Nominal	Input
27	AQ1	Numeric	12	0	1.How frequently do you visit bank for transaction p...	{1, Never}...	None	12	Right	Nominal	Input
28	AQ2	Numeric	12	0	2.How frequently do you visit bank for cheque/dema...	{1, Never}...	None	12	Right	Nominal	Input

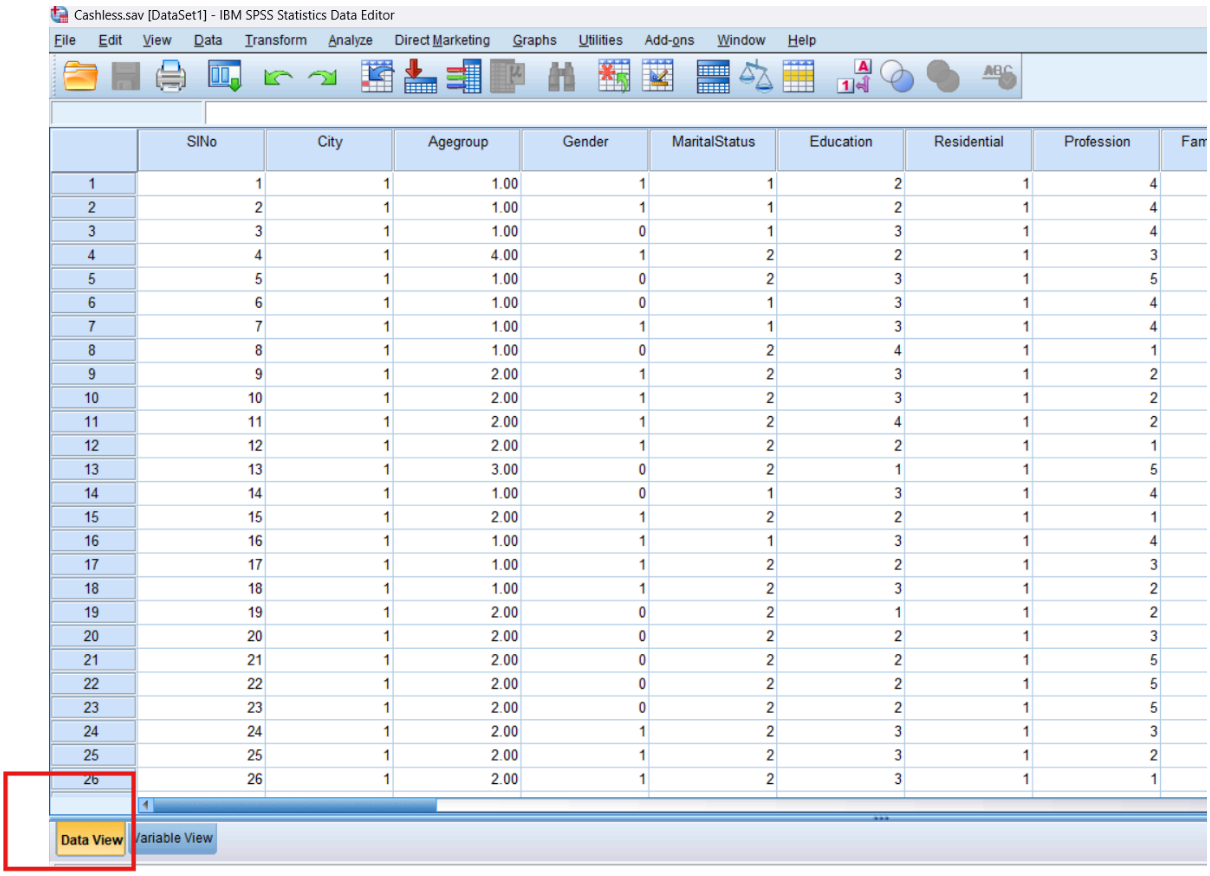
6.4.3.2 Entering Data Manually (Data View)

Once variables are defined, switch to the **Data View** tab. Each column represents a variable, and each row corresponds to a respondent or observation.

For example, you can enter:

Age	Gender	Income
23	1	25000
30	2	32000
45	1	50000

SPSS will automatically apply labels (if defined) when viewing the data or during analysis.



	SlNo	City	Agegroup	Gender	MaritalStatus	Education	Residential	Profession	Fam
1	1	1	1.00	1	1	2	1	4	
2	2	1	1.00	1	1	2	1	4	
3	3	1	1.00	0	1	3	1	4	
4	4	1	4.00	1	2	2	1	3	
5	5	1	1.00	0	2	3	1	5	
6	6	1	1.00	0	1	3	1	4	
7	7	1	1.00	1	1	3	1	4	
8	8	1	1.00	0	2	4	1	1	
9	9	1	2.00	1	2	3	1	2	
10	10	1	2.00	1	2	3	1	2	
11	11	1	2.00	1	2	4	1	2	
12	12	1	2.00	1	2	2	1	1	
13	13	1	3.00	0	2	1	1	5	
14	14	1	1.00	0	1	3	1	4	
15	15	1	2.00	1	2	2	1	1	
16	16	1	1.00	1	1	3	1	4	
17	17	1	1.00	1	2	2	1	3	
18	18	1	1.00	1	2	3	1	2	
19	19	1	2.00	0	2	1	1	2	
20	20	1	2.00	0	2	2	1	3	
21	21	1	2.00	0	2	2	1	5	
22	22	1	2.00	0	2	2	1	5	
23	23	1	2.00	0	2	2	1	5	
24	24	1	2.00	1	2	3	1	3	
25	25	1	2.00	1	2	3	1	2	
26	26	1	2.00	1	2	3	1	1	

Importing Data from Excel

Instead of manual entry, researchers often import data from Excel. SPSS makes this easy.

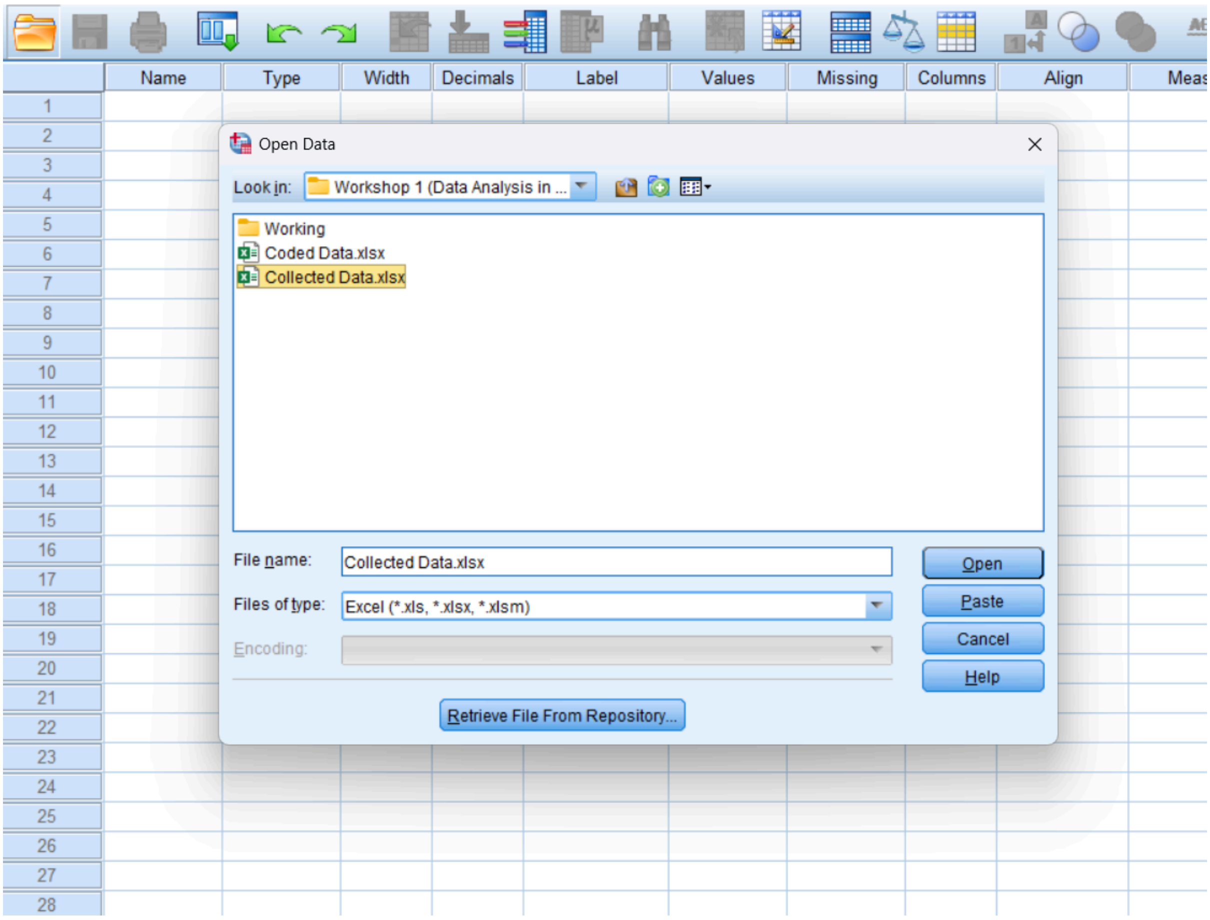
Steps:

a) Prepare your Excel file:
   - Ensure variable names are in the first row.
   - Avoid merged cells and ensure consistent formatting.

b) In SPSS, go to:
   - File → Open → Data.
   - Choose **"Excel (*.xls, .xlsx)"* from the file type.
   - Locate and select your Excel file.

c) In the pop-up:
   - Check the box **"Read variable names from the first row of data."**
   - Click **OK**.

SPSS will import the data and display it in Data View. You may still need to visit **Variable View** to adjust labels, values, and measurement levels.



6.4.3.3 Practical Example

The importance of practical application of the SPSS predictive, descriptive, and transformative analytics tools can be illustrated by the survey done among the college students. Lets say the data set contains five variables: Name, Age, Gender, Course, and Satisfaction Score. In SPSS Variable View, each of these will be entered as a different row. The Name variable will be defined as a 'String' type as it contains text-based answers, whereas the other variables– Age, Gender, Course, and Satisfaction Score– will be defined as 'Numeric' variables. In the case of the Gender and Course variables, which are categorical, value labels should be assigned to the variables in oder for the data to be meaningful during the analysis.

For instance, Gender can be assigned 1 = Male and 2 = Female, and Course can have 1 = Arts, 2 = Science, and 3 = Commerce. The measure field allows SPSS to know how to treat the data: 'Nominal' for the Gender and Course variables which are categorical and 'Scale' for the Satisfaction Score which is on a 1-5 rating. Age is also treated as a scale variable because it is continous numeric data. After the properties have all been defined, the data set is ready for proper data input and statistical analysis.

After entering or importing your data into the SPSS workspace, the next step is data analysis. SPSS offers tools to perform descriptive statistics, correlations, regressions, and many other advanced analyses. In addition, the data can be presented in the form of charts and graphs. SPSS is capable of processing categorical data alongside continuous data, and this serves as additional proof of the detailed analysis that SPSS can provide at different measurement levels (nominal, ordinal, and scale).

Suppose you're conducting a survey on college students with the following variables:

- Name
- Age
- Gender (1 = Male, 2 = Female)
- Course (1 = Arts, 2 = Science, 3 = Commerce)
- Satisfaction Score (1 to 5 scale)

In Variable View:

- Create 5 rows (Name, Age, Gender, Course, Satisfaction).
- Set “Type” to String for Name; others as Numeric.
- Assign values for categorical variables like Gender and Course.
- Set “Measure” appropriately (Nominal for Gender/Course, Scale for Satisfaction).

In Data View: Enter data for each respondent in rows.

- Once data is entered or imported into SPSS, you can begin analysis.
- SPSS provides a wide range of statistical tools, such as:
    - Descriptive statistics
    - Correlation analysis
    - Regression analysis
    - Advanced statistical tests
- The software also allows creation of charts and graphs for data visualization.
- SPSS can handle both **categorical** and **continuous** data types.
- This flexibility supports detailed analysis based on different measurement levels:

- **Nominal** (categories without order)
- **Ordinal** (ranked categories)
- **Scale** (numeric/continuous data)

6.5 Coding for SPSS Analysis

Statistical analysis requires for the whole different types of the data to be converted to numbers. Such data are then encoded as numbers and is stored in a hypercube. A hypercube is then the basic data structure in SPSS that is a three-dimensional array holding the the data. Each of the coded pieces of data is then is referred as a cell. As an example of variable Gender that is coded as 1 0 for male and 0 for female. A similar ordered structure is also in use for educational qualification academic level i.e. coded as 1 for high school, 2 for Graduate, 3 for Postgraduate, and a variable like Region coded as 1 for Urban, 2 for Rural, and 3 for Semi Urban. The numeric data in the data view is converted to coding systems which are in the Variable View. In fact, the numeric data does not alter with the original data, which simply standardizes to the information and undergo testing of the hypothesis. Therefore, Value Labels and Variable View system especially in SPSS is used for labeling during hypothesis testing to analyze the data.To that end, value labeling remains critical for proper data management as it enhances presentation qualities of categorical variables within SPSS.

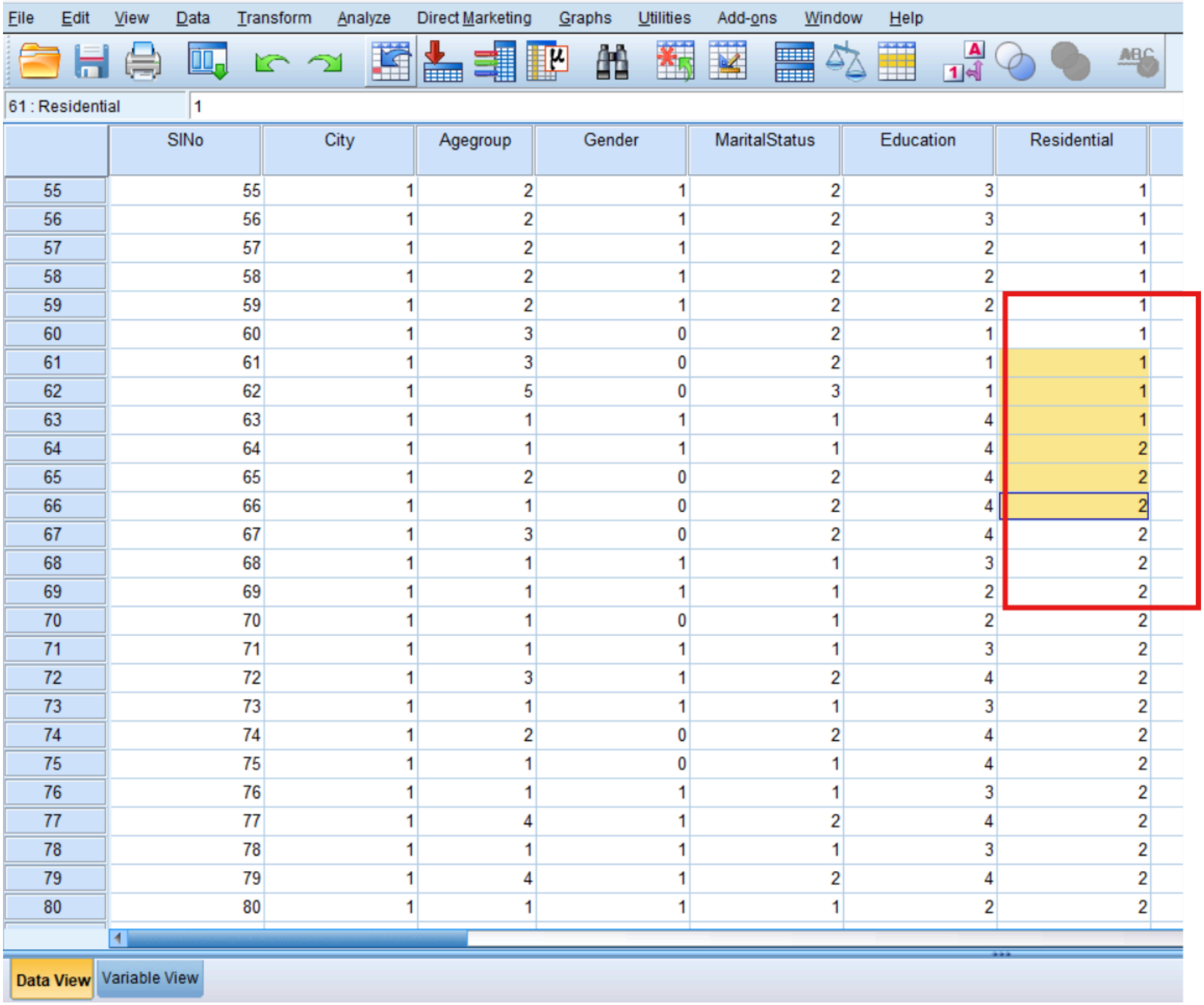
Numeric Coding of Variables

In SPSS, non-numeric variables are assigned numbers to ensure they can be processed correctly. For example:

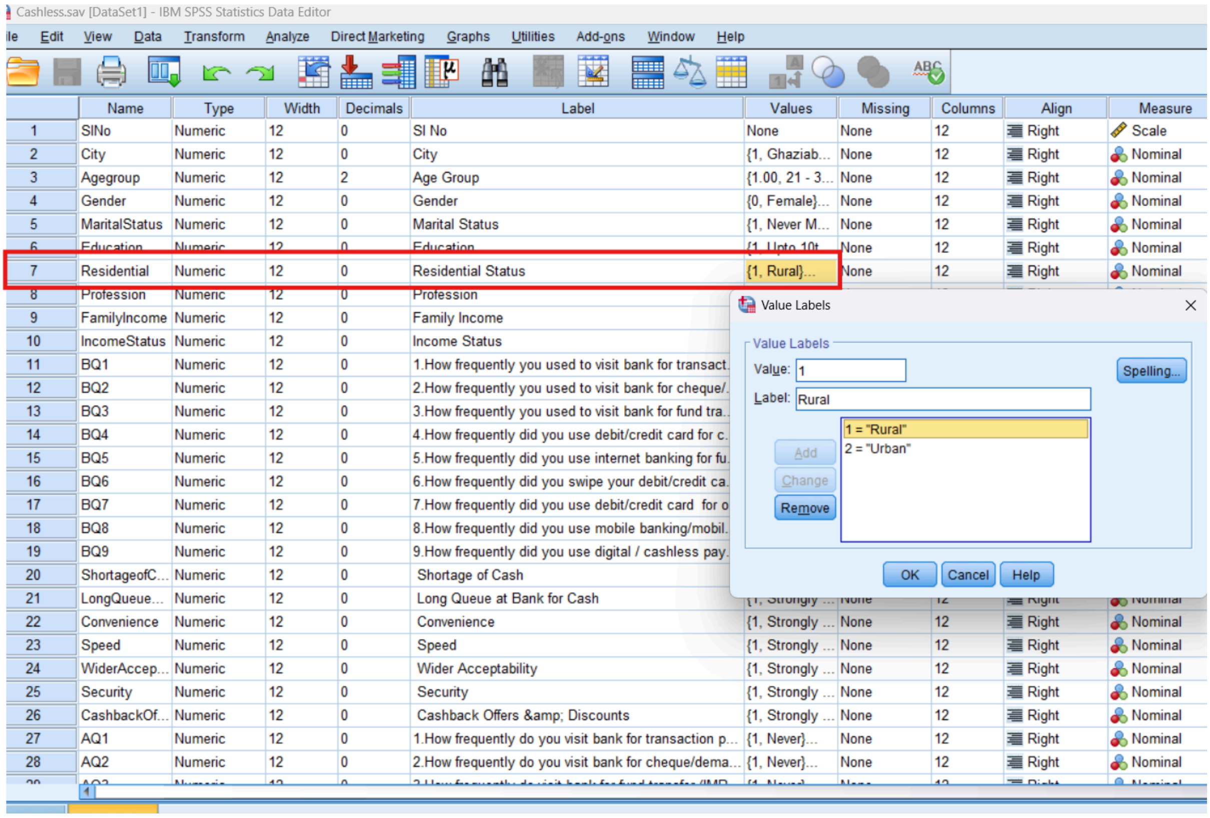
- Gender:
    - 1 = Male
    - 2 = Female
- Education:
    - 1 = High School
    - 2 = Graduate
    - 3 = Postgraduate
- Region:
    - 1 = Urban

- 2 = Rural
- 3 = Semi-Urban

These codes are entered in the **Data View** of SPSS, but their meanings are defined using **Value Labels** in the **Variable View** tab.



	SlNo	City	Agegroup	Gender	MaritalStatus	Education	Residential
55	55	1	2	1	2	3	1
56	56	1	2	1	2	3	1
57	57	1	2	1	2	2	1
58	58	1	2	1	2	2	1
59	59	1	2	1	2	2	1
60	60	1	3	0	2	1	1
61	61	1	3	0	2	1	1
62	62	1	5	0	3	1	1
63	63	1	1	1	1	4	1
64	64	1	1	1	1	4	2
65	65	1	2	0	2	4	2
66	66	1	1	0	2	4	2
67	67	1	3	0	2	4	2
68	68	1	1	1	1	3	2
69	69	1	1	1	1	2	2
70	70	1	1	0	1	2	2
71	71	1	1	1	1	3	2
72	72	1	3	1	2	4	2
73	73	1	1	1	1	3	2
74	74	1	2	0	2	4	2
75	75	1	1	0	1	4	2
76	76	1	1	1	1	3	2
77	77	1	4	1	2	4	2
78	78	1	1	1	1	3	2
79	79	1	4	1	2	4	2
80	80	1	1	1	1	2	2



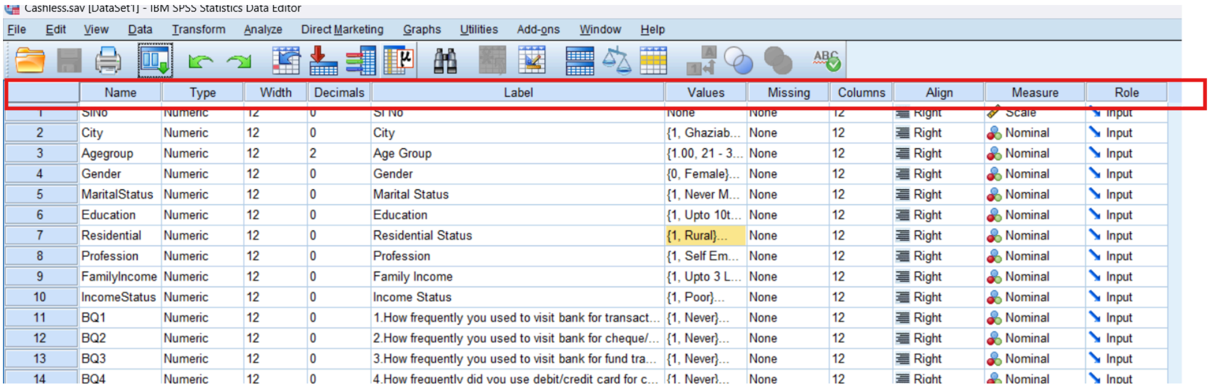
6.6 Use of Variable View in SPSS

Each dataset is comprised of a range of different properties as is evident in SPSS. A dataset is typically comprised of columns which each correlate to a different field in the data. Each field is defining from the others by a name each value is given. Differentiating the columns is possible due to the data each name field holds. Each name field comprises a code which is far more less unique such and it can be put as the primary element at small elements composing the field header. A field is able to have unique stored data such as it can be numerical in nature, it can be a string, or put in textual form. If a value holds a type as string then it can further be modified by adding a 'Width' and 'Decimals' feature as this can hold great importance to the overall data. Each field can be modified to become more readable and descriptive as well too. For example, such as in fields 'Respondent's Gender' can be simplified. Each field then allows users to code missing or ineligible responses, ever efficiently. Final in the field of the Measure property, each field is either a Nominal or Scale.

SPSS's **Variable View** provides options to define the properties of each variable:

- **Name**: A short, one-word identifier for the variable (e.g., gender, income).
- **Type**: Specifies the kind of data—numeric, string (text), date, etc.

- **Width and Decimals**: Controls how the data is displayed.
- **Label**: A descriptive name that appears in outputs (e.g., “Respondent’s Gender”).
- **Values**: Defines what each code means (e.g., 1 = Male, 2 = Female).
- **Missing**: Identifies any special codes used to represent missing data (e.g., 99 or -1).
- **Measure**: Indicates the scale of measurement—nominal, ordinal, or scale (interval/ratio).


Cashless.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	SlNo	Numeric	12	0	Sl No	None	None	12	Right	Scale	Input
2	City	Numeric	12	0	City	{1, Ghaziab...	None	12	Right	Nominal	Input
3	Agegroup	Numeric	12	2	Age Group	{1.00, 21 - 3...	None	12	Right	Nominal	Input
4	Gender	Numeric	12	0	Gender	{0, Female}...	None	12	Right	Nominal	Input
5	MaritalStatus	Numeric	12	0	Marital Status	{1, Never M...	None	12	Right	Nominal	Input
6	Education	Numeric	12	0	Education	{1, Upto 10t...	None	12	Right	Nominal	Input
7	Residential	Numeric	12	0	Residential Status	{1, Rural}...	None	12	Right	Nominal	Input
8	Profession	Numeric	12	0	Profession	{1, Self Em...	None	12	Right	Nominal	Input
9	FamilyIncome	Numeric	12	0	Family Income	{1, Upto 3 L...	None	12	Right	Nominal	Input
10	IncomeStatus	Numeric	12	0	Income Status	{1, Poor}...	None	12	Right	Nominal	Input
11	BQ1	Numeric	12	0	1.How frequently you used to visit bank for transact...	{1, Never}...	None	12	Right	Nominal	Input
12	BQ2	Numeric	12	0	2.How frequently you used to visit bank for cheque/...	{1, Never}...	None	12	Right	Nominal	Input
13	BQ3	Numeric	12	0	3.How frequently you used to visit bank for fund tra...	{1, Never}...	None	12	Right	Nominal	Input
14	BQ4	Numeric	12	0	4.How frequently did you use debit/credit card for c...	{1, Never}...	None	12	Right	Nominal	Input

6.6.1 Importance of Coding Accuracy

Inconsistent or improper coding may result in indicators or software deficiencies. If for instance some responses for the gender question are recorded in words while others are given a numeric value, SPSS will treat them as separate categories. This emphasizes the value in cleaning and recoding data as a preparatory procedure for analysis.

The coding and data classication acts as the conduit between the collection phase and the advanced stages in the statistical analysis phase. Following the steps and structuring the processes in Excel and SPSS—consistent formatting or variable labeling, numeric coding—provides assurance that the data collected is suitable for advanced analysis. Not only does refined data augments the value of results obtained, but it also makes the process of interpreting the data as well as compiling the data for the analysis easier, which are the most basic essential components in formulating and constructing reliable research.

6.7 Variable Setup:

Setting up the variables correctly is the first and equally important step in using SPSS data analysis as a tooling technique. A variable is any measureable or classifiable trait or attribute, including age, gender, education, or even satisfaction score. Variable View of SPSS is a systematic aproach in defining and managing variables which enables the researcher to set up a name, assign labels, set up clear data types, and choose the set measurement range or variable. This step is very crucial considering the data to be used as attribute and the statistical analysis carried out afterwards.

In SPSS, a Variable Name is the label given to each variable in the dataset, which is used for record keeping. This name is used in Data View, formulas, syntax instructions, and statistical analysis. It is important that a variable name starts with a letter and is not surrounded by spaces or special characters like @, #, or a dash. Also, although SPSS allows for the length of up to 64 characters, the best practice is to keep the name as short as possible, as the longer the name, the greater difficulties a user encounters when trying to reference the name in later phases of analysis. The permissible names can include but are not limited to Age, Gender, and Income2024, while the names that can not be used include 2024Income, Student Name, or Income@Level. These all break the rules and conventions of proper naming. SPSS will be able to customize these names to suit the Data View.

Visibility and readability of report output can greatly improve by providing the Variable label which offers a detailed definition of the variable. Unlike variable names which are devoid of syntax or unscripted characters, labels can embrace full textual expressions. Although labels are invisible in Data View, they can be found in tables, charts, and exported reports. These report features enhance the comprehensibility of the report. For instance, a variable named SatScore can comfortably fit the label as “Overall Student Satisfaction Score (1 to 5 scale).” This is helpful to the researchers as they can comfortably use a shorthand internally and provide the detailed description in their reports.

The other important consideration in Variable View is Variable Type. Every type of data in SPSS has a type, the more common ones being Numeric (used in denoting a person’s age, marks, or salary), String (used in denoting a name or a comment), and Date (used in denoting information which is timeline based data). It is important to select the correct type, as mis-definitions will cause issues in later stages of analysis. If Age is classified as a string and not a number, SPSS will not be able to compute averages or conduct other statistical tests. An example of the data type scenario is Variable View → Type → click the “...” box and select the most appropriate type form the list presented. Putting in correct names and labels is the first step in what will be a precise and usable SPSS dataset.

Variable Name

The **Variable Name** is the identifier used within SPSS for each variable.

- **Rules**:
  - Must begin with a letter.
  - Cannot contain spaces or special characters (e.g., @, #, -).
  - Limited to 64 characters, but shorter, meaningful names are recommended.

Examples:

- Valid: Age, Gender, Income2024
- Invalid: 2024Income, Student Name, Income@Level

This name is used internally in SPSS for reference in functions, formulas, or output tables.

Variable Label

While variable names are often abbreviated or compact, **Variable Labels** provide more descriptive text to clarify what the variable represents.

- **Purpose**: Enhances readability in output tables, especially during complex analyses.
- Can include spaces and punctuation.
- Not visible in Data View but appears in output and charts.

Example:

- Variable Name: SatScore
- Variable Label: Overall Student Satisfaction Score (1 to 5 scale)

This distinction between name and label allows researchers to maintain both precision and clarity.

Variable Type

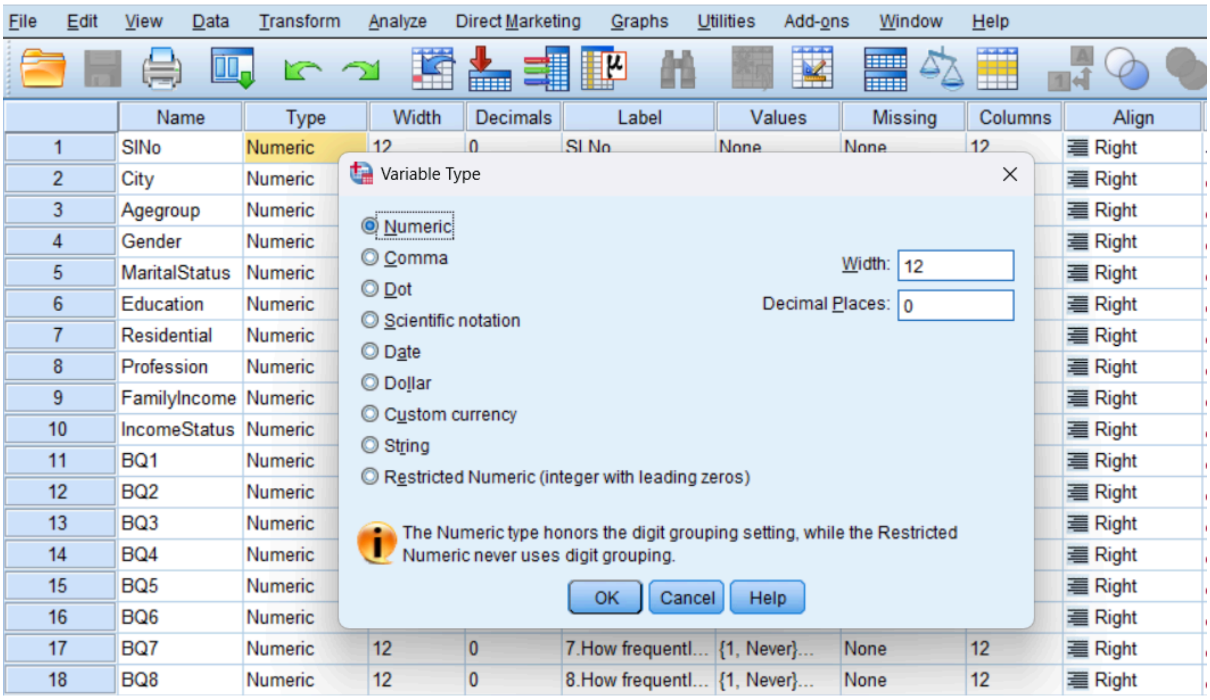
SPSS allows several data types, but the most commonly used are:

- **Numeric**: Numbers only. Used for calculations, statistical analysis (e.g., Age, Score, Salary).
- **String**: Text data. Used when variables include names, IDs, comments, etc.
- **Date**: Used when recording dates; multiple formats available (dd-mm-yyyy, mm/dd/yyyy, etc.).
- **Comma/Dot/Scientific Notation**: Formatting variations of numeric data.

To select the type:

- Go to **Variable View → Type → Click on '...' button** to select and configure data type.

Note: Incorrect data types can lead to errors during analysis. For example, storing "Age" as a string will prevent SPSS from computing averages.



6.7.1 Value Labels (For Categorical Variables)

Value Labels are vital features in SPSS when dealing with categorical variable data that are treated as numbers. Instead of inputting responses as Male or Female in the dataset directly, such responses are coded numerically as 1 = Male and 2 = Female. These codes are recorded in the Data View and their meanings are specified in the Variable View by assigning descriptive labels in the Values column. Doing this saves SPSS from misinterpreting the data as quantities to be categorized as zero, and it saves SPSS from producing nonsense text in frequency tables, graphs, and other statistical analyses. Value labels are crucial in SPSS; otherwise, the outputs would be numbers devoid of context.

Every variable must be assigned a value in addition to value labels, and this value must correspond to a Measure, or level of measurement, which tells SPSS how to analyze the variable. There are three types of measurement: Nominal, Ordinal, and Scale. Nominal variables are made up of separate categories that can be listed, but not ranked, for instance, a person's gender or the religion to which they belong or the area in which they reside. Ordinal variables also consist of categories, but in this case the categories have a rank order, for example, the level of education attained or the level of satisfaction with a service experienced, but the intervals or distance between the categories is not necessarily equal. Scale or interval or ratio measurement is quantitative and continuous, for example, a person’s age, income, or the score they attained on a test.

It is important to get the measurement level correct because SPSS will have a lot of value in the correctly identified measurement levels. SPSS will know the appropriate statistical analysis, tests, and graphs that should be run or constructed, and it will also be able to value summarize the answers accurately. SPSS will be misled and will generate false functionalities if incorrectly classified units are presented.

Value Labels (For Categorical Variables)

For variables that represent categories numerically (like Gender or Course), **Value Labels** map each number to its real-world meaning.

Example:

- Variable: Gender
- Coded Values:
    - 1 = Male
    - 2 = Female

Set value labels in Variable View by clicking the **“Values”** cell for the variable.

This coding allows SPSS to treat these numbers as distinct categories during analysis and to display meaningful labels in charts and tables.

Measure (Scale Level)

SPSS requires users to specify the **Level of Measurement** for each variable. This influences what statistical tests can be run and how SPSS interprets the data.

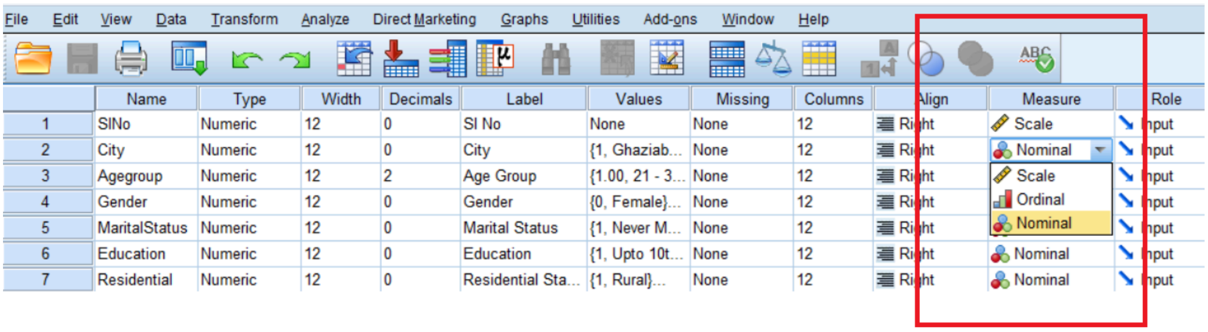
There are three primary levels:

- **Nominal**:
    - Categories without any order (e.g., Gender, Religion, City).
    - SPSS treats these as labels only, not quantities.
- **Ordinal**:

- Categories with a meaningful order, but unequal intervals (e.g., Education Level: High School < Graduate < Postgraduate).
- Ranking is possible, but not direct mathematical operations.

- **Scale**:
  - Continuous data with equal intervals (e.g., Age, Income, Test Scores).
  - Includes both interval and ratio-level data.
  - Suitable for most statistical computations (mean, standard deviation, correlation).

Setting the correct measure ensures SPSS selects suitable tests and visualizations automatically.



6.7.2 Practical Example

Suppose you're analyzing a student feedback survey with the following variables:

Variable	Label	Type	Value Labels	Measure
Age	Age in years	Numeric	—	Scale
Gender	Respondent's Gender	Numeric	1 = Male, 2 = Female	Nominal
Satisfaction	Course Satisfaction Score (1–5)	Numeric	1 = Very Poor to 5 = Excellent	Ordinal
Course	Type of Course	Numeric	1 = Arts, 2 = Science, 3 = Commerce	Nominal

This structure tells SPSS how to handle each variable and ensures accurate output during analysis.

6.8 Data Cleaning and Error Checking

Data cleaning is the process of identifying and correcting errors, inconsistencies, or inaccuracies in datasets to ensure the integrity and validity of research findings. In social science research, where data often comes from diverse human respondents, this step is particularly critical.

Raw data often contains issues such as entry errors, duplications, inconsistent coding, or invalid responses. Without proper cleaning:

- **Analysis becomes unreliable**, potentially leading to incorrect conclusions.
- **Errors may go undetected**, causing bias or inflated significance in results.
- **Statistical assumptions are violated**, leading to increased risk of Type I or II errors.
- **Research credibility diminishes**, especially if findings cannot be replicated.

Hence, data cleaning is foundational to every phase of data analysis.

6.8.1 Detecting and Managing Outliers

Outliers are data points that differ greatly from the rest of the dataset and can significantly influence statistical results, leading to misleading conclusions. They may occur due to natural variation, data entry errors, or unusual cases. Outlier detection can be done visually using boxplots, which highlight values that fall outside the interquartile range, or through scatterplots, which reveal unusual patterns in two-variable relationships. Statistical methods include calculating Z-scores, where values above +3 or below −3 are considered potential outliers, and applying the IQR rule, where values lower than Q1 − 1.5×IQR or higher than Q3 + 1.5×IQR are flagged. Once detected, outliers should not be removed blindly; the first step is to verify whether they are errors or valid extreme cases. Based on the context, the researcher may decide to keep the value, remove it, or apply transformations such as log scaling. All decisions must be documented for transparency.

In addition to outliers, datasets may contain inconsistencies, which occur when two or more related values contradict each other—for example, a 12-year-old listed as a postgraduate or a person marked as both “single” and “spouse present.” These issues can be identified using cross-variable checks, frequency tables, or validation rules in SPSS or Excel. Once identified, values should be verified against source data and corrected or removed, with a record maintained of all changes for audit and research integrity.

Detecting and Managing Outliers

Outliers are extreme values that deviate from other data points and can distort statistical summaries and models.

Detection Methods:

- **Visual**:
  - *Boxplots* help detect data points outside the interquartile range (IQR).
  - *Scatterplots* are useful for spotting outliers in bivariate datasets.
- **Statistical**:
  - *Z-scores*: A Z-score above ±3 generally flags a value as an outlier.
  - *IQR Rule*: Points below Q1 − 1.5×IQR or above Q3 + 1.5×IQR are flagged.

Handling Methods:

- Investigate whether the value stems from a data entry or reporting error.
- Based on context, retain, transform (e.g., log transformation), or remove the outlier.
- Always **document actions taken**, including the rationale behind removal or retention.

Identifying Inconsistencies

Logical inconsistencies occur when related values contradict each other.

Examples:

- A 12-year-old reporting "Postgraduate" education.
- A person marked as both “single” and “spouse present”.

Resolution Techniques:

- **Cross-variable checks**: Use conditional logic in Excel (IF, AND) or SPSS (IF, DO IF) to detect contradictions.
- **Frequency tables**: Identify rare or invalid values.
- **Validation rules**: In Excel, apply Data Validation to restrict inputs.

Next Steps:

- Re-verify the original data source.
- Correct or delete erroneous values.
- Keep a **change log** for transparency.

6.8.2 Handling Missing Data

If unaddressed, missing information in the research dataset can negatively influence the results’ validity by decreasing the sample size, lowering statistical power, and producing biased results. Supporting evidence that failing to address missing information can hinder the overall research quality can be found in the work of Charles and Von Braun. The first step in understanding missing data is classification. Three of the five types categories' types of missing data. MCAR (Missing Completely at Random) refers to the absence of relationships with at least one variable in the dataset. Siloed analyses will be unbiased in these situations, even with the absence of data. MAR (Missing at Random) involves absence in the set of data, but only with the presence of other variables. An example of this could be with the age set of respondents where patriarchal respondents tend to lean towards the income question. MNAR, or "Missing Not at Random" is the worse because it involves values missing in the data set and can be exemplified with cases of individuals in the high-income bracket who do not tend to report their income.

In SPSS, descriptive statistics can show frequencies, which can help answer which values are missing. In excel one can either use the count blank function or use the blank count quantifier filters. Blank values should be coded in a uniform way, using values like -99, or 999, which are then recoded in SPSS under the missing column in the Variable View to ensure SPSS doesn't use them as valid numbers when performing analyses.

Several missing data techniques involve either deletion or imputation. Some missing data techniques involve deletion methods; listwise deletion removes all rows with any missing values which only works when the missing data is low and missing completely at random. In pairwise deletion, data is retained as long as there are no values in the rows that are needed for a specific analysis. Imputation techniques are a sophisticated way to substitute missing values; the most reliable being regression where the missing values are predicted using all the data and available in SPSS under Analyze → Multiple Imputation.

Documentation is critical for this type of research. A data cleaning log captures each variable impacted, the type of missing data and the missing data techniques along with the reasoning behind them. SPSS missing Value analysis and Excel conditional formatting are visualization aids that assist in reporting patterns to help form reproducible processes.

Missing data can compromise analysis by reducing statistical power or introducing bias, especially when the absence of data is systematic.

Types of Missing Data

Understanding the cause of missingness helps determine the appropriate treatment method.

- **MCAR (Missing Completely at Random)**: The missingness is unrelated to observed or unobserved data. Analysis remains unbiased.
- **MAR (Missing at Random)**: Missingness is related to observed data but not the missing values themselves.
- **MNAR (Missing Not at Random)**: The reason for missingness is related to the missing values (e.g., people with high income skipping income questions).

Detecting and Coding Missing Data

Detection:

- In **Excel**: Use COUNTBLANK(), ISBLANK(), or filters.
- In **SPSS**: Analyze → Descriptive Statistics → Frequencies (tick "Missing Values").

Coding Techniques:

- Use placeholders like -99 or 999 consistently.
- In SPSS: Define these as discrete missing values in the "Missing" column in Variable View.

Proper coding ensures that software does not mistakenly include missing entries in analyses like averages or regressions.

Strategies to Handle Missing Data

(a) Deletion Techniques:

- **Listwise Deletion**: Removes entire cases if any variable is missing. Best for small amounts of MCAR data.
- **Pairwise Deletion**: Includes all available data for each analysis. Retains more data but results can vary depending on which variables are analyzed.

(b) Imputation Techniques:

- **Mean/Median/Mode Imputation**: Fill missing numeric values with the mean or median. Best for small, MAR-type gaps.
- **Regression Imputation**: Predicts missing values based on relationships with other variables.
- **Hot Deck Imputation**: Substitutes missing entries with values from similar respondents.

- **Multiple Imputation**: Generates several datasets with imputed values and combines the results for robust estimation. SPSS includes this under Analyze → Multiple Imputation.

Documenting and Visualizing Missing Data

Maintain a **data cleaning log**:

- Include variable names, type of issue, action taken, and justification.

Visual Tools:

- **SPSS**: Use "Missing Value Analysis" for patterns.
- **Excel**: Conditional formatting to highlight blanks or abnormal values.

6.8.3 Data Transformation

Data transformation is one of the last stages of preparation as it takes raw data and gives it a degree of meaning which makes it consistent enough for analysis. One of the more commonplace tasks done in transformation is assigning values to the new variables which makes it possible for the researcher to unlock further information from data that they already possess. This comes in handy especially in situations whereby multiple responses have to be synthesized into a single score or an index like in a survey. In a survey that has a collection of questions that focuses on the usage of mobile banking or the levels of stress, the responses given to each of the questions can be summed up to yield a composite score that reflects an attitude or behavior as a whole.

New variables may also be formed to show ratios, proportions, percentage of a given data like the dependency ratios in population studies or growths in the economy, or other variables like growth rates in a given economy. The transformation of data is done over a period of time verticals like in longitudinal data needs to be, or as in the case of computing over a period of time like yearly averages, percent increase, or cumulatives. The variable that is to be created is done under the module of transform and then use the compute variable which then allows for the user to fill in new variable names and a formula that seeks to explain the variable within existing ones.

As an example, an academic might be interested in how mobile banking is used. To this end, she might create a new variable called Overall_Use, where Overall_Use = BQ1 + BQ2 + BQ3 + BQ4 + BQ6 + BQ7 + BQ8 + BQ9. Each lettered variable corresponds to a question in a survey.

Overall, this new variable can be used in subsequent analyses, including correlation, regression, and reliability tests. One of the advantages of creating new variables is the enhanced clarity of analyses and conclusions possible by the trimming of individual response “noise.” Proper documentation—such as logging the formula, reason, and

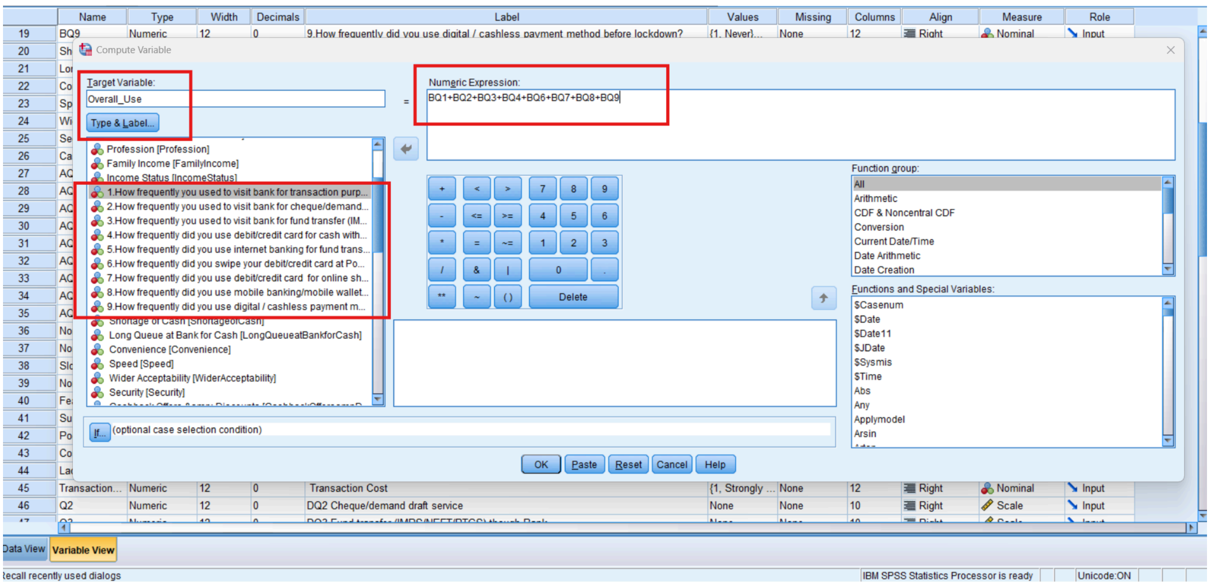
Computing New Variables

Purpose:

- Combine multiple responses into a scale (e.g., sum of stress scores).
- Derive ratios or rates (e.g., dependency ratio).
- Create time-based averages or growth rates.

In SPSS:

- Use Transform → Compute Variable.
- Define a new variable and enter mathematical expressions (e.g., Overall_Use of Mobile Banking as in the following example = BQ1+BQ2+BQ3+BQ4+BQ6+BQ7+BQ8+BQ9.



In Excel:

- Use functions like SUM(), AVERAGE(), and arithmetic formulas.

Recoding Existing Variables

Why Recode?

- Simplify analysis by reducing categories.

- Standardize values for consistency.
- Prepare categorical variables for regression or visualization.

In SPSS:

- Use Transform → Recode into Different Variables.
- Map old values to new ones (e.g., Likert 1–2 = "Low", 3 = "Moderate", 4–5 = "High").

In Excel:

- Use IF, IFS, or VLOOKUP() to create new coded columns.

Binning or Grouping Continuous Variables

Continuous variables can be converted into categorical bands for easier segmentation or logistic analysis.

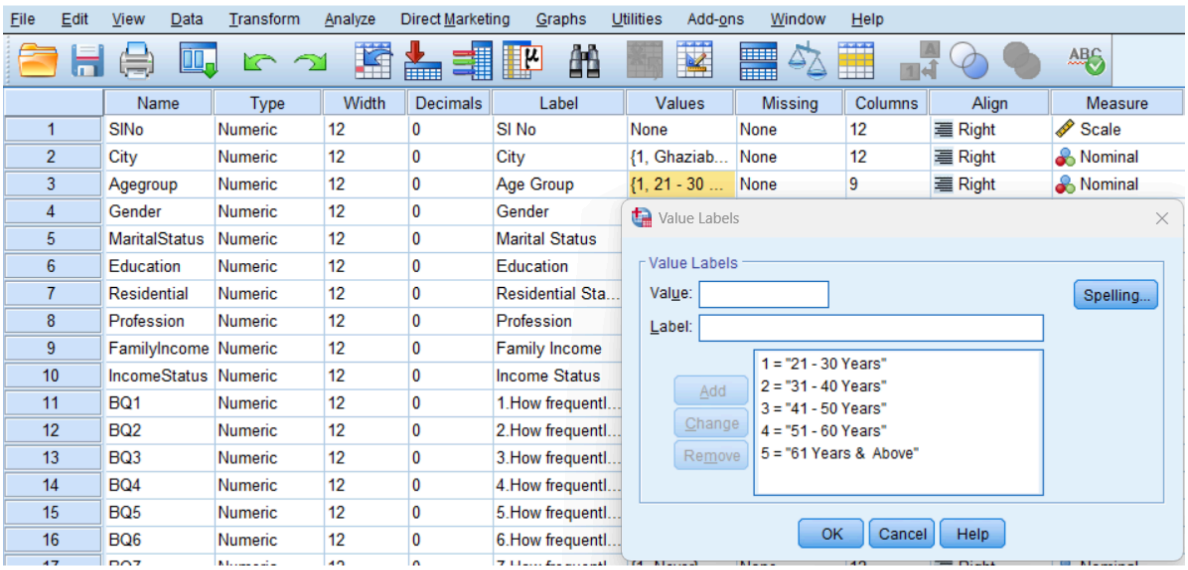
Examples:

- *Age*: Convert raw ages into categories (18–25, 26–35, etc.).
- *Income*: Categorize into "Low", "Medium", "High".

SPSS Methods:

- Use Transform → Visual Binning to define cut points and labels.
- Alternatively, use Recode into Different Variables with value ranges.

Excel Methods:



Use nested IF() or FLOOR()/CEILING() functions for categorization.

6.9 Why Transformation is Crucial

Data transformation is important in statistical analysis because it refines the raw variables to meet the objectives of the study. Several statistical assessments make certain assumptions—normality or linearity, for example—and transformations address these assumptions, increasing the accuracy of the results. Transforming data also improves a person's understanding of the data by changing complex or uneven values into something simpler for easier interpretation, comparison, and copying. For instance, classifying variables of income, or applying a log transformation helps in pattern appreciation. Moreover, variables in a transformed state provide better visual illustrations of analyses, such as neater histograms or trend line graphs, which results in greater correctness and relevancy of the conclusions. Therefor the major reasons for data transformation are -

- Aligns raw variables with analytical goals.
- Helps meet assumptions like normality or linearity.
- Makes outputs more interpretable and comparable.
- Facilitates meaningful visualization (e.g., histograms of income groups).

6.10 Coding Categorical Variables:

Research in social science has collected more of the data categorized by sex, marital status, area, education, occupation, etc. Doing any kind of statistical analysis in SPSS requires that the data be transformed from the uncovered form into richly detailed digits. This aids in data compatibility with other statistical methods and increases the quality of the data.

6.10.1 Understanding Categorical Variables

Rather than numerical measurements, categorical variables represent different groups or categories. There are two main types: nominal and ordinal. Ordinal variables have a logical order or ranking such as Income Bracket, Education, or Satisfaction. Nominal variables, on the other, have none such as Gender, Religion, or Nationality. While age, weight, and income are continuous variables, categorical variables as these examples cannot be used directly in mathematical calculations unless "coded" into numbers. For this reason, data preparation in tools such as SPSS requires assigning numeric codes to "categories" in order to transform raw data into analytical data.

Categorical variable coding is crucial in SPSS as the software, regardless of its interface, functions on numbers. For instance, the system is said to "process" words such as "Male" and “Female” as assigning values “1” and “2” respectively. Consistent coding is beneficial for data entry as it reduces mistakes, and it enhances the “smooth” execution” of statistical tests such as frequency distribution, cross tabulation, and regression that are governed by dummy variables and chi-square. Categorical data is subsequently easier to manage as coding enhances sorting, filtering, and grouping.

In order to preserve the integrity of the data set, correct coding must be done to support valid statistical interpretation. Each statistical interpretation must be paired to categorical data that can be analyzed meaningfully. There are steps that must be followed to maintain data integrity. First, the coding schemes must be specified, outlining the most appropriate solution for the coding of the data set. Categorizations can be represented as marital status and placed as 1 = Single, 2 = Married, 3 = Divorced, or the level of education attained and placed as 1 = High School, 2 = Undergraduate, 3 = Postgraduate. Order must only be assigned to the codes if the variable set is indeed ordinal. Other coding schemes that may arise with structures such as [1.0, 2.0) and letter structures/domain must be avoided as these are unnecessary and SPSS does not recognize these structures as proper algebraic word classifications. Finally, the coding variable structures must be retained and properly labeled as Chargeable Amenities to ensure the value is captured within the data set for correct processing within SPSS.

Incorrect coding of data can lead to the loss of the data, deterioration of statistical interpretation validation and loss of information set contained as categorical data.

Understanding Categorical Variables

Categorical variables represent groups or categories. These can be:

- **Nominal** (no natural order): e.g., Gender, Religion, Nationality
- **Ordinal** (ordered categories): e.g., Satisfaction Level, Education Level, Income Bracket

Unlike continuous variables (e.g., age or income), these categories cannot be used in mathematical operations unless they are numerically coded.

Why Code Categorical Variables?

SPSS processes numbers—not text. Even when the user visually sees labels like "Male" or "Graduate," SPSS handles them as underlying numeric codes (e.g., 1 = Male, 2 = Female). Coding provides:

- **Consistency** in data entry
- **Accuracy** in statistical analysis
- **Ease** in filtering, sorting, and recoding
- **Compatibility** with built-in SPSS procedures

Best Practices in Coding Categorical Data

a. Use Clear and Consistent Coding Schemes

Assign intuitive numeric codes:

- Gender: 1 = Male, 2 = Female
- Marital Status: 1 = Single, 2 = Married, 3 = Divorced
- Education Level: 1 = High School, 2 = Undergraduate, 3 = Postgraduate

Avoid assigning codes that imply an artificial hierarchy unless the variable is truly ordinal.

b. Avoid Arbitrary or Redundant Codes

Don't overcomplicate with unnecessary decimal places (e.g., 1.0, 2.0) or use alphanumeric codes (e.g., M = Male, F = Female) in SPSS, as they can introduce errors.

c. Use Numeric Type for All Coded Variables

Ensure your categorical variables are of the **numeric type** in SPSS, even if they represent text labels. String variables (text fields) limit your ability to run statistical analyses and are best reserved for IDs or comments.

6.11 Applying Value Labels in SPSS (Using Variable View)

Assigning value labels in SPSS is important in changing numerical data to more useful readable groups for assessment and explanation. The next step in analysing categorical variables that have been numerically coded is to attach value labels so that SPSS can show user-friendly text like "Male," "Female," "Satisfied," "Dissatisfied," etc. This is important to ensure that report data users and explainers do not struggle to understand findings from the report and provide answers without extensive explanations.

In SPSS, under Variable View, find the Variable that needs labeling for value labeling. Then in the Values column, click the button containing three dots (…) to navigate to the Value Labels section. "Enter each of the value pairs e.g. Value: 1, Label: Male; Value: 2, Label: Female. Value Labels for each user and a corresponding Label needs to be made by clicking Add and entering information. After this is done, select the Yes option. From now on, SPSS will demonstrate the rational and more logical labels in data view, output tables, frequencies and charts etc. This ensures that the fundamental estimates, for the sake of complex mathematical analysis, remain in their numerical format, further, such a system ensures that there is no compromise on analytical precision and interpretative clarity.

SPSS also allows efficient editing from its options for batch operations. If there are several variables with the same response categories as some Likert-scale items with values ranging from 1 to 5, you can use the right-click and Copy/Paste or Copy Properties to value labels across the variables. This is useful in saving time and variation in coding across the dataset. The configuration is Order, Discrete, and Nominal.

For instance, choosing Nominal for unordered categories such as Religion, and Blood Group, and Ordinal for categories such as Level of Satisfaction and Level of Education, helps SPSS to make inferences in measuring the statistics in the appropriate ways. Thus, SPSS advanced features can help in analyzing data with nominal features using Chi-square, while ordinal attributes can be computed using non-parametric methods such as Mann-Whitney U or Spearman's rho.

Correct use of value labels supports data importance in terms of its comprehensibility, enhances model visualization, and minimizes errors in the output of the data.

Once variables are coded, use SPSS's **Variable View** to assign **Value Labels** that provide meaning to the numbers.

Steps:

1. **Open Variable View** in SPSS.
2. Find the variable to label (e.g., Gender).
3. In the **Values** column, click the small box with "..." to open the label editor.
4. Enter values and corresponding labels:
   - Value: 1, Label: Male
   - Value: 2, Label: Female
5. Click **Add** after each entry and then **OK**.

Now, during analysis and in output tables, SPSS will display "Male" or "Female" instead of numeric codes, while internally treating them as numbers for analysis.

Efficient Use of Variable View

a. Batch Editing Variables

You can quickly copy value labels from one variable to another (e.g., if you have multiple variables with the same categories like multiple satisfaction questions rated 1–5).

- Right-click → Copy/Paste row entries in Variable View.
- Use **"Copy properties"** for efficient replication.

b. Labeling Measurement Levels

In the **Measure** column, be sure to select:

- **Nominal** for unordered categories (e.g., Religion, Gender)
- **Ordinal** for ordered ones (e.g., Satisfaction Scale, Education Level)

SPSS uses these designations to determine which analyses and charts are appropriate.

6.11.1 Real-World Example

Variable: Education Level

Response Option	Code	Value Label
High School	1	High School
Undergraduate	2	Undergraduate
Postgraduate	3	Postgraduate
Doctorate	4	Doctorate

- In **Variable View**:
  - Name: EduLevel
  - Type: Numeric
  - Values: Assign labels for 1 to 4
  - Measure: Ordinal

6.12 The Output Window in SPSS

The **Output Window** in SPSS is where all the results from your analyses are displayed. When you run statistical procedures (like frequencies, t-tests, regression, or graphs), SPSS sends the results to this separate window, allowing you to view, interpret, and export your findings.

6.12.1 Key Features:

- **Navigation Pane**: On the left side, SPSS lists all the procedures and output sections (tables, charts, notes) in a tree-like structure. You can click on any item to jump to that part of the output.
- **Output Viewer**: On the right side, the actual content is displayed—tables, charts, syntax, and notes.
- **Editable Output**: While the results are automatically generated, you can:
    - Double-click on tables to edit labels, fonts, or formats.
    - Right-click to copy tables or export them.
- **Charts and Graphs**: SPSS generates visual elements like histograms, pie charts, and scatterplots, which are editable using the Chart Editor.

Exporting Output:

You can export the contents of the Output Window to various formats:

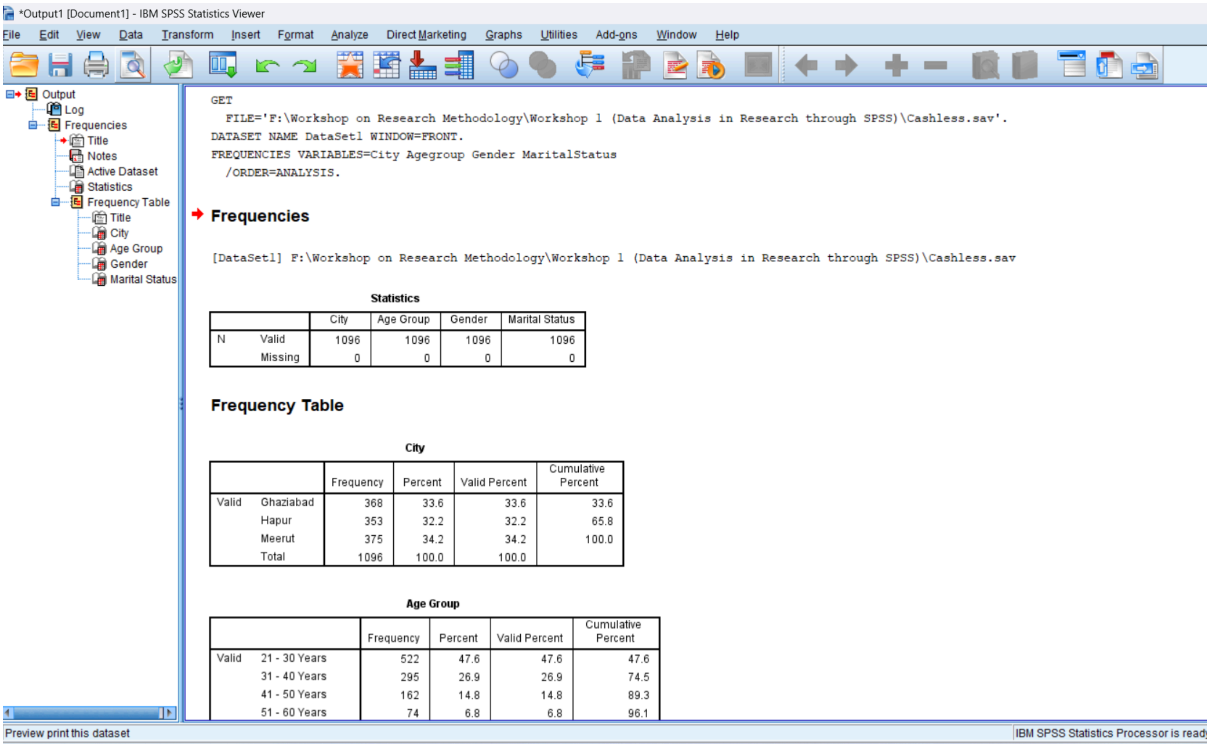
- **File** → **Export** → Choose formats like:
    - Word (.doc/.docx)
    - PDF
    - Excel
    - HTML

Saving the Output:

- Save the output file separately from your data file.
- Extension: .spv (SPSS Viewer format)
- To reopen: File → Open → Output

Best Practices:

- Regularly save your output to avoid losing results.
- Organize results by renaming titles and using the navigation tree.
- Clean up unnecessary outputs to maintain clarity.



Statistics

		City	Age Group	Gender	Marital Status
N	Valid	1096	1096	1096	1096
	Missing	0	0	0	0

City

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Ghaziabad	368	33.6	33.6	33.6
	Hapur	353	32.2	32.2	65.8
	Meerut	375	34.2	34.2	100.0
	Total	1096	100.0	100.0	

Age Group

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	21 - 30 Years	522	47.6	47.6	47.6
	31 - 40 Years	295	26.9	26.9	74.5
	41 - 50 Years	162	14.8	14.8	89.3
	51 - 60 Years	74	6.8	6.8	96.1

6.13 Choosing the Right Statistical Analysis in SPSS

Statistical decision-making in SPSS gets easier with the use of decision trees that assist the researcher in choosing the right test according to the type of variables and the aim of the research. One of the important objectives in research is to analyze the relationships between the variables. If the study is concerned with two scale (interval or ratio) variables such as age and income, then Pearson's Correlation is used, as it ascertains the strength and direction of linear relationships. Point-Biserial is used to test the association of one nominal (e.g. gender) and one scale variable. For two ordinal variables such as satisfaction and ranking of the service, Spearman's Rank Correlation is used because it doesn't assume normality. If both variables are nominal such as religion and voting, the Chi-Square Test of Association is used to see if a relationship exists between the variables. More than two variables can be analyzed simultaneously using techniques like Multiple Correlation or Canonical Correlation to determine complex interrelationships.

Another major area of analysis is comparative analyses and these are used whenever the primary aim is to establish the presence of any statistically significant difference between the different groups or conditions within which the groups are placed. In the case of examining two independent groups on a single numerical variable (e.g. test scores of males versus females) the comparison is said to be on a single variable and the correct test is an

Independent Samples t-test. In the case where the same group of subjects is assessed on two and only two occasions as in before and after an intervention, a Paired Samples t-test is applied. For comparisons which involve more than two groups, the analysis is a One-Way ANOVA, and if the analysis involves two independent variables the analysis is a Two-Way ANOVA. For comparisons involving categorical data the standard is the Chi-Square Test. In case data do not meet the elementary assumptions of parametric test of analysis such as normality, there are non-parametric alternatives of Mann-Whitney U, Wilcoxon Signed Rank, and Kruskal-Wallis which are more than adequate.

For works which are more focused on prediction, there is the use of regression based decision trees. Simple Linear Regression involves the scaling down of an outcome variable which can be predicted from a single predictor and which is an outcome of a certain form, while in Multiple Linear Regression, there are two or more predictors. If the outcome variable is dichotomous (e.g., purchase or not purchase), Binary Logistic Regression is applied, whereas with more than two non-ordered categories, the outcome is predicted through the application of Multinomial Logistic Regression.

In case of research objectives that involve classification, techniques like Cluster Analysis, Discriminant Analysis, and various types of Decision Trees (CHAID, CART) provide the means to classify cases, or to assign group membership based on some predictor variables. These techniques are particularly useful in the context of segmentation, diagnostics, and risk profiling.

Lastly, SPSS offers a full-fledged environment for elementary as well as advanced statistical research by integrating various at-a-glance procedures as: Factor Analysis (for finding latent variables); Reliability Analysis (e.g., as Cronbach’s Alpha for verifying a scale); Time Series Analysis (for trend & forecasting); and Survival Analysis (for time-to-event analysis).

Use this when your research question involves exploring **associations or correlations** between variables.

- **Two scale (interval/ratio) variables** → Use **Pearson’s Correlation**
- **One nominal and one scale variable** → Use **Point-Biserial Correlation**
- **Two ordinal variables** → Use **Spearman’s Rank Correlation**
- **Two nominal variables** → Use **Chi-Square Test of Association**

- **More than two variables** → Use **Multiple Correlation or Canonical Correlation**

These tests help determine whether variables move together and how strongly they are associated.

Decision Tree for Comparative Analyses of Differences

Use this category when comparing **groups or conditions** to find if differences exist.

- **Two groups, one scale variable** → Use **Independent Samples t-test**
- **Same group tested twice** → Use **Paired Samples t-test**
- **More than two groups** → Use **One-Way ANOVA**
- **Two independent variables (e.g., gender and age group)** → Use **Two-Way ANOVA**
- **Categorical variables** → Use **Chi-Square Test**
- **Non-parametric options** (when data does not meet assumptions): Mann-Whitney U, Kruskal-Wallis, Wilcoxon

These techniques help identify whether interventions, demographic groups, or experimental conditions yield statistically different outcomes.

Decision Tree for Predictive Analyses

When you want to **predict a numerical outcome** based on one or more predictors:

- **One predictor, one outcome (both scale)** → Use **Simple Linear Regression**
- **Multiple predictors** → Use **Multiple Linear Regression**
- **Predict a binary outcome (e.g., success/failure)** → Use **Binary Logistic Regression**
- **Predict multiple categories (e.g., low/medium/high)** → Use **Multinomial Logistic Regression**

Regression models help in forecasting values, understanding the strength of predictors, and estimating probabilities.

Decision Tree for Classification Analyses

Use this when your goal is to **group or classify cases** based on shared characteristics.

- **Classify into naturally occurring groups** → Use **Cluster Analysis**
- **Predefined groups with predictors** → Use **Discriminant Analysis**
- **Decision-making using variables to predict group membership** → Use **Decision Trees (e.g., CHAID, CART in SPSS)**

These tools are widely used in marketing, customer segmentation, and psychological profiling.

Other Analysis Categories

Apart from the main types above, SPSS supports many additional statistical techniques:

- **Factor Analysis** → Reduce large sets of variables into underlying factors
- **Reliability Analysis (e.g., Cronbach's Alpha)** → Check internal consistency of scales
- **Time Series Analysis** → Analyze trends over time
- **Survival Analysis** → Examine event durations (e.g., time until dropout or recovery)

The kinds of data and the research objectives determines which method of analysis is appropriate in SPSS. For example, decision trees can be used to aid in recommending a more structured and confident approach to an analysis. As a rule, the validity of the results is subject to assumption checks of each statistical procedure—normality, homogeneity of variance, independence—meaning those conditions must hold for the results to be deemed trustworthy. SPSS provides options for parametric and non-parametric tests which aid in assumption violations to an extent.

Mastering the provided information, you will able to conduct research in SPSS with analysis that is precise, appropriate, and trustworthy.

CHAPTER 7: Data for Analysis Using SPSS

7.1 Descriptive Statistics under the Analysis Menu in SPSS

Through descriptive statistics, a researcher can compile and review the key components of any data set. After all, before any inferential testing, the researcher needs to have an overview of the data to appreciate the distribution and the characteristics. The SPSS application allows the researcher to compute different types of statistics and even descriptive types of statistics, and is found under the Analyze menu.

SPSS has a Descriptive Statistics sub-menu, where several different types of data can be processed. The researcher can use the sub-menu to compute the average and dispersion measures, and also frequencies and percentile scores. Such measures are vital to any form of data analysis and they are used in the reporting stage.

7.2 Descriptive Statistics

Descriptive statistics are the first type of statistics used to summarize, organize, and meaningfully present any data. In any research activity, before performing any advanced statistical analyses, the data need to be summarized to calculate the mean, variation, and the distribution of the data. For example, SPSS offers non-complex facilities to calculate descriptive statistics and aid in data visualization. exc. This section covers in depth the the steps SPSS guide users through when performing descriptive statistics, and more importantly, the steps' significance in research.

Why Descriptive Statistics Matter

Descriptive statistics serve several essential functions:

- **Summarize data quickly and clearly**.
- **Identify errors or outliers** before deeper analysis.
- **Assess assumptions** like normality and homogeneity of variance.
- **Guide the selection** of inferential statistics (e.g., parametric vs. non-parametric tests).

Proper use of SPSS's descriptive statistics tools ensures a strong foundation for advanced analysis. By using the appropriate options from the Analysis menu, researchers can gain valuable insights into their data and make informed decisions on further testing strategies.

7.2.1 Steps for Performing Descriptive Statistics in SPSS

1. **Open the Data File**

Begin by entering the dataset in SPSS. Data can be typed directly into the Data View or imported from external sources like Excel, CSV, or SAV files. Variables should be clearly labeled in the Variable View for clarity.

2. **Navigate to Descriptive Statistics**

   From the main menu, go to Analyze → Descriptive Statistics. SPSS offers different options under this menu:

   - **Frequencies**: Provides frequency tables, percentages, and summary measures.
   - **Descriptives**: Shows mean, standard deviation, minimum, and maximum values.
   - **Explore**: Provides detailed statistics, including skewness, kurtosis, normality tests, and visual plots.

3. **Select Variables**

   Choose the numeric variables (e.g., age, income, marks, scores) to be analyzed. Move them from the left-side list of variables to the analysis box on the right using the arrow button.

4. **Choose Statistics Options**

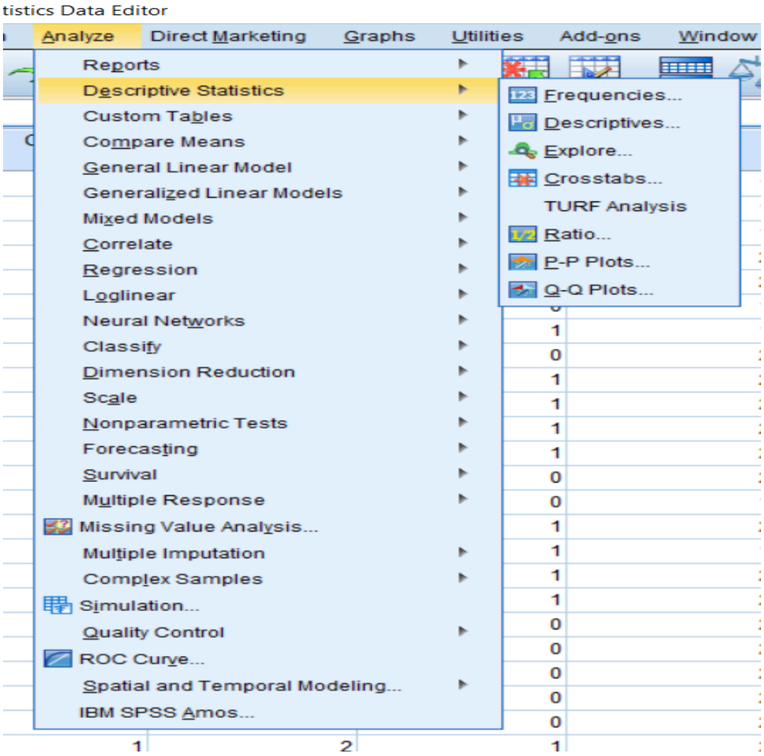
   - In **Frequencies**, click on Statistics to select measures such as mean, median, mode, standard deviation, variance, and percentiles.
   - In **Descriptives**, click on Options to specify which statistics you want to display (mean, SD, min, max).
   - In **Explore**, you can choose additional options for skewness, kurtosis, and plots such as histograms, stem-and-leaf plots, and boxplots.

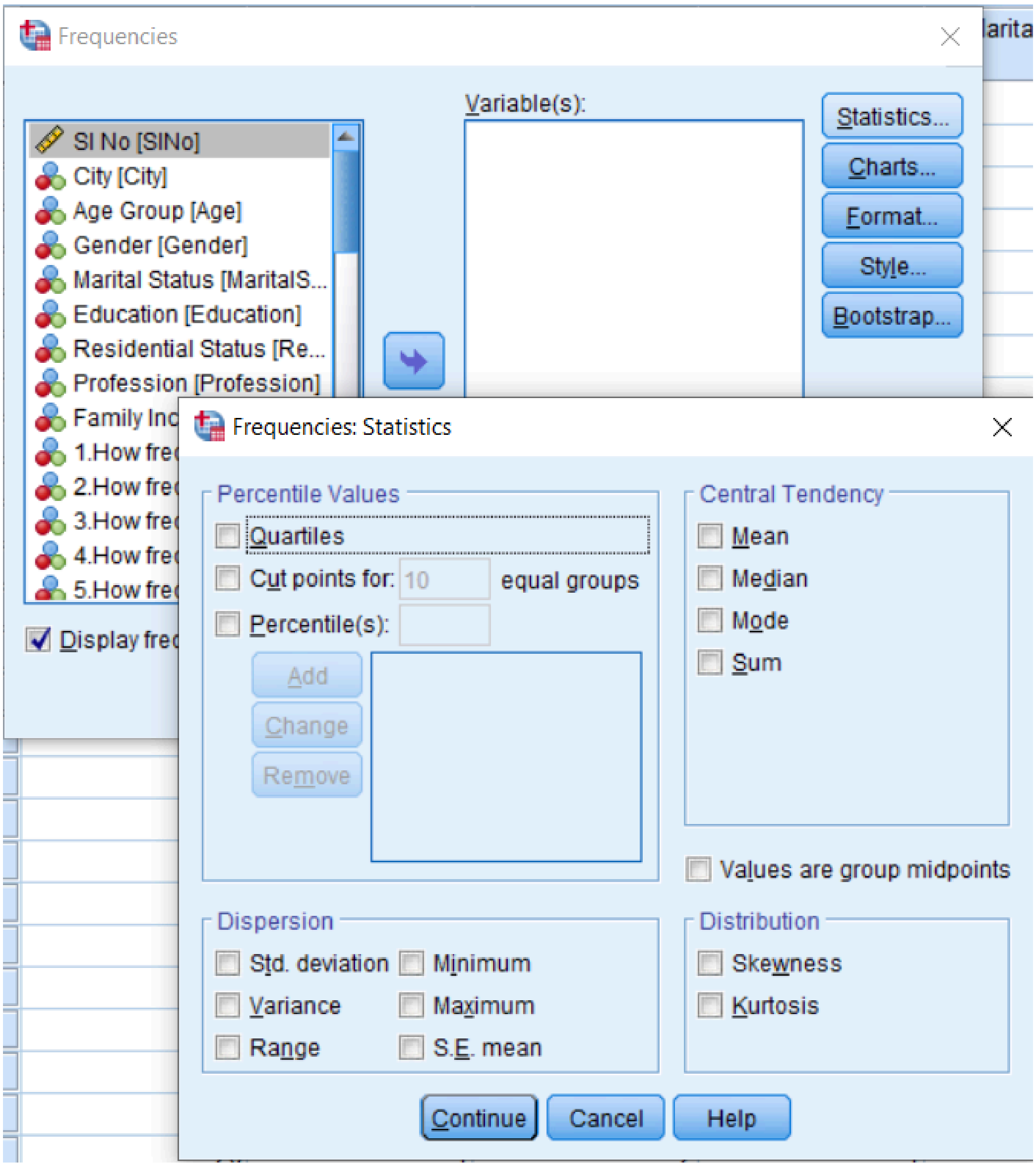
5. **Run the Analysis**

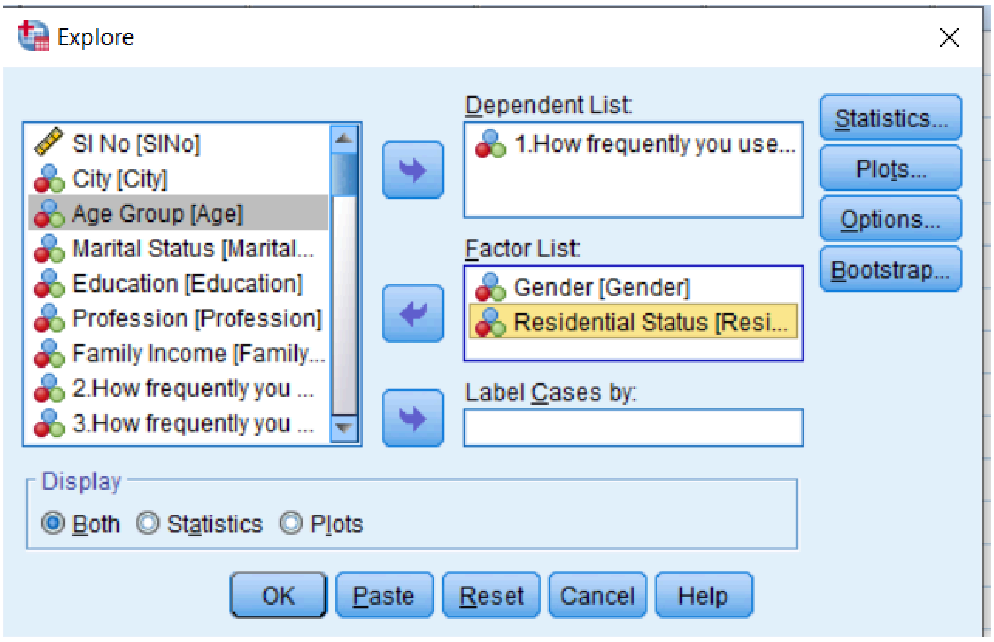
   After selecting variables and statistics, click **OK**. SPSS will generate an output in the Output Viewer. The output contains tables summarizing descriptive measures and graphical displays if selected.

6. **Interpret Results**

   The results should be carefully interpreted. Measures of central tendency (mean, median, mode) explain the typical or average score. Measures of dispersion (range, variance, standard deviation) describe the spread of data. Skewness and kurtosis provide information about the shape and normality of the distribution.


tistics Data Editor
Analyze
Direct Marketing
Graphs
Utilities
Add-ons
Window
Reports
Descriptive Statistics
Custom Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
ROC Curve...
Spatial and Temporal Modeling...
IBM SPSS Amos...
Frequencies...
Descriptives...
Explore...
Crosstabs...
TURF Analysis
Ratio...
P-P Plots...
Q-Q Plots...


Frequencies
Variable(s):
Statistics...
Charts...
Format...
Style...
Bootstrap...
Sl No [SlNo]
City [City]
Age Group [Age]
Gender [Gender]
Marital Status [MaritalS...
Education [Education]
Residential Status [Re...
Profession [Profession]
Family Inc
1.How fre
2.How fre
3.How fre
4.How fre
5.How fre
Display fre
Frequencies: Statistics
Percentile Values
Quartiles
Cut points for: 10 equal groups
Percentile(s):
Add
Change
Remove
Central Tendency
Mean
Median
Mode
Sum
Values are group midpoints
Dispersion
Std. deviation
Minimum
Variance
Maximum
Range
S.E. mean
Distribution
Skewness
Kurtosis
Continue
Cancel
Help


Explore
Dependent List:
1.How frequently you use...
Factor List:
Gender [Gender]
Residential Status [Resi...
Label Cases by:
Statistics...
Plots...
Options...
Bootstrap...
Sl No [SlNo]
City [City]
Age Group [Age]
Marital Status [Marital...
Education [Education]
Profession [Profession]
Family Income [Family...
2.How frequently you ...
3.How frequently you ...
Display
Both
Statistics
Plots
OK
Paste
Reset
Cancel
Help

Descriptives

	Gender			Statistic	Std. Error
1.How frequently you used to visit bank for transaction purpose? [Please mark your choice]	Female	Mean		2.34	.037
		95% Confidence Interval for Mean	Lower Bound	2.26	
			Upper Bound	2.41	
		5% Trimmed Mean		2.31	
		Median		2.00	
		Variance		.553	
		Std. Deviation		.744	
		Minimum		1	
		Maximum		5	
		Range		4	
		Interquartile Range		1	
		Skewness		.505	.120
		Kurtosis		.842	.239
	Male	Mean		2.64	.029
		95% Confidence Interval for Mean	Lower Bound	2.58	
			Upper Bound	2.69	
		5% Trimmed Mean		2.63	
		Median		3.00	
		Variance		.572	
		Std. Deviation		.757	
		Minimum		1	
		Maximum		5	
		Range		4	
		Interquartile Range		1	
		Skewness		.306	.094
		Kurtosis		.467	.187

7.2.2 Accessing Descriptive Statistics

To access this feature:

- Go to the **Analyze** menu.
- Hover over **Descriptive Statistics**.
- Choose from the available options: **Frequencies**, **Descriptives**, **Explore**, **Crosstabs**, **Ratio**, and **Q-Q Plots**.

Each of these tools serves a specific purpose, explained below.

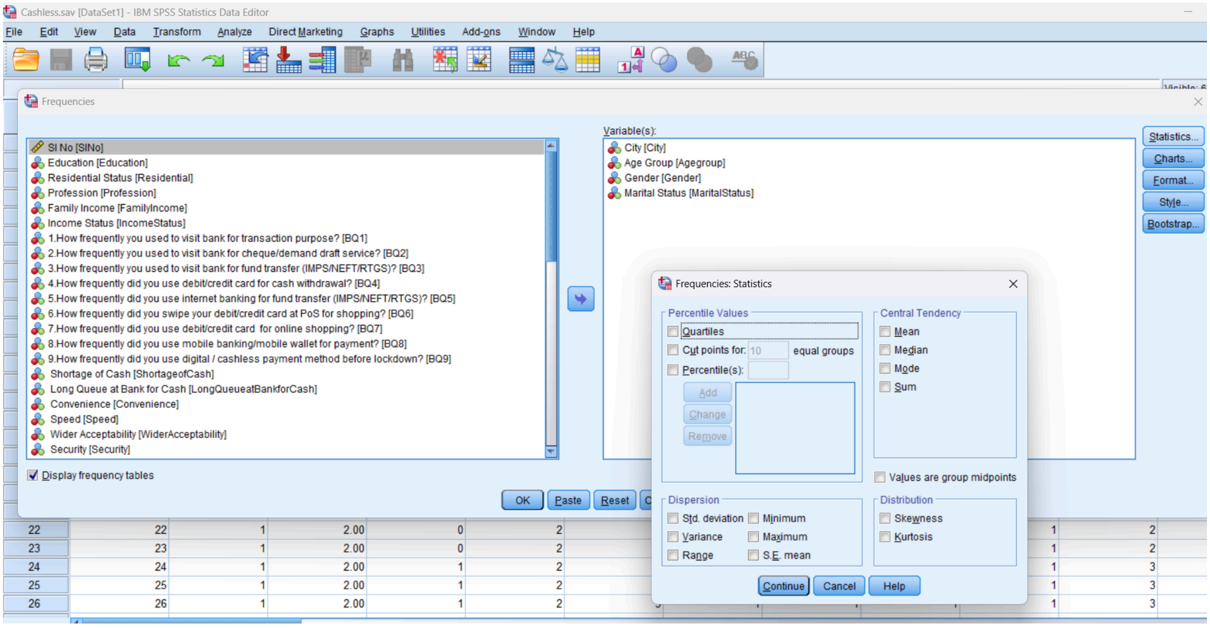
Frequencies

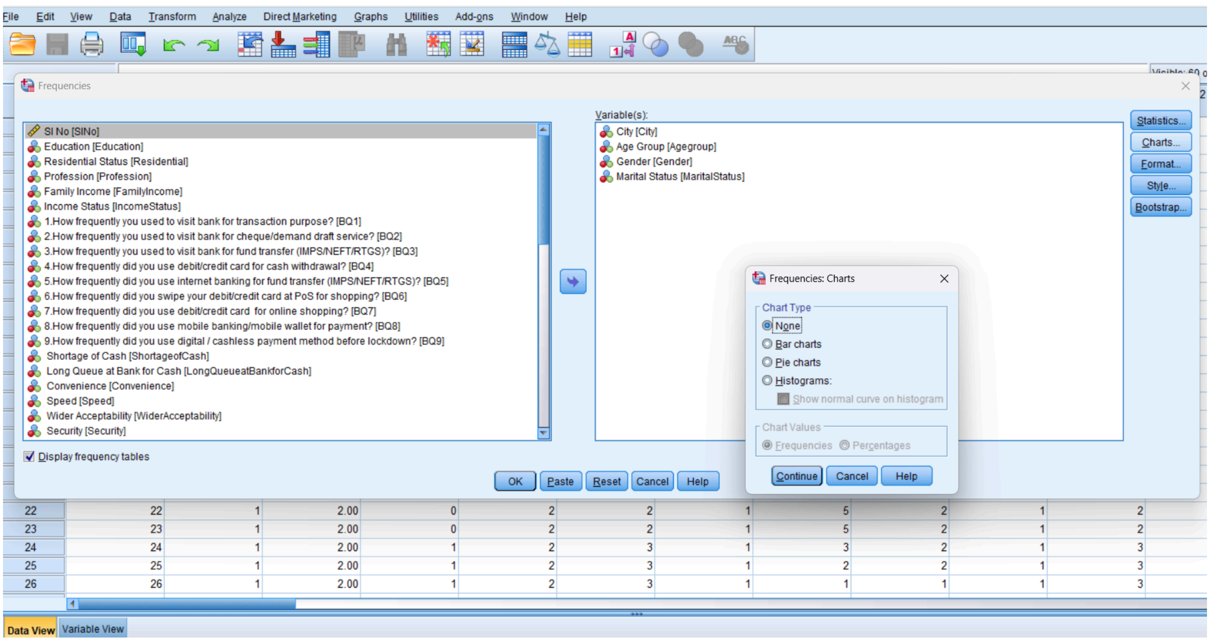
This option is ideal for analyzing **categorical or ordinal data**.

- Go to: **Analyze** → **Descriptive Statistics** → **Frequencies**
- Select one or more variables (e.g., Gender, Education Level).
- Choose output options like **frequency tables**, **bar charts**, or **pie charts**.

- Under **Statistics**, you can request additional measures like:
  - Mean
  - Median
  - Mode
  - Percentiles
  - Range
  - Standard deviation

Use Case: Displaying the number of respondents in each category (e.g., how many males vs. females in the sample).



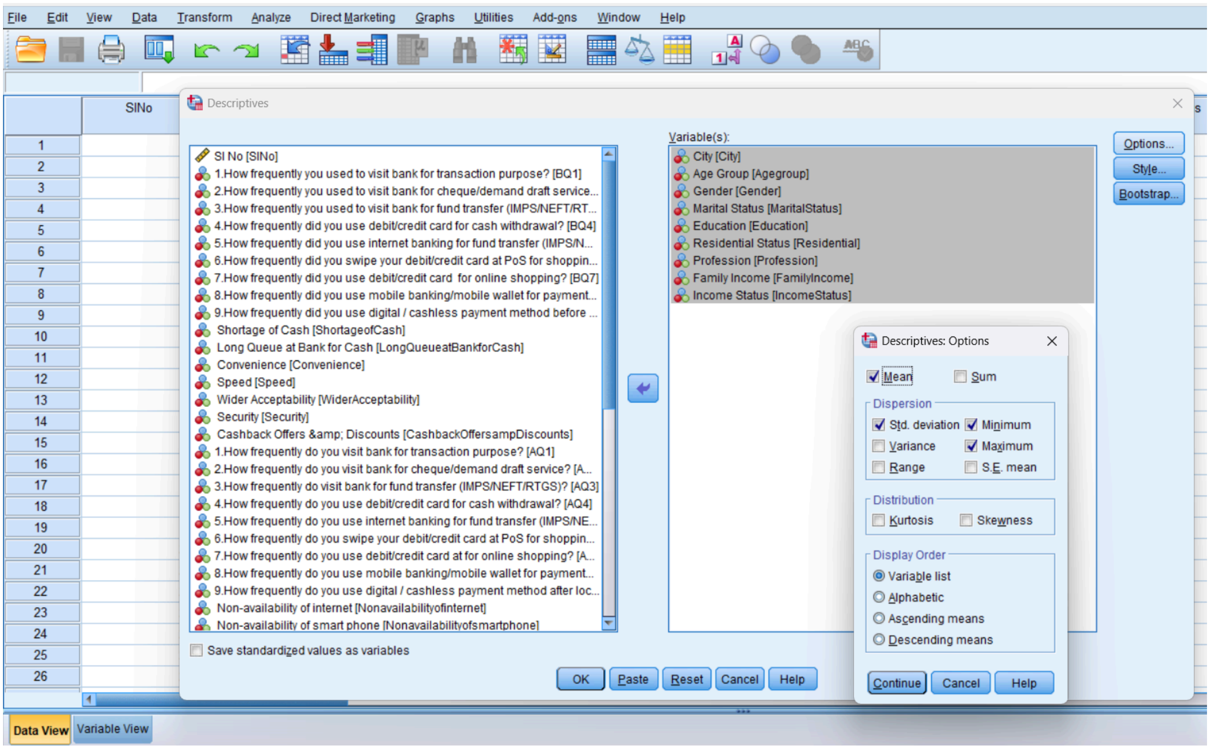


Descriptives

This is used primarily for **scale-level (interval or ratio) variables**.

- Go to: **Analyze** → **Descriptive Statistics** → **Descriptives**
- Select continuous variables (e.g., Age, Income, Score).
- Click **Options** to choose which statistics to include, such as:
  - Mean
  - Standard Deviation
  - Minimum and Maximum
  - Variance
  - Kurtosis and Skewness (for shape of distribution)

Use Case: Summarizing central tendency and dispersion of test scores.



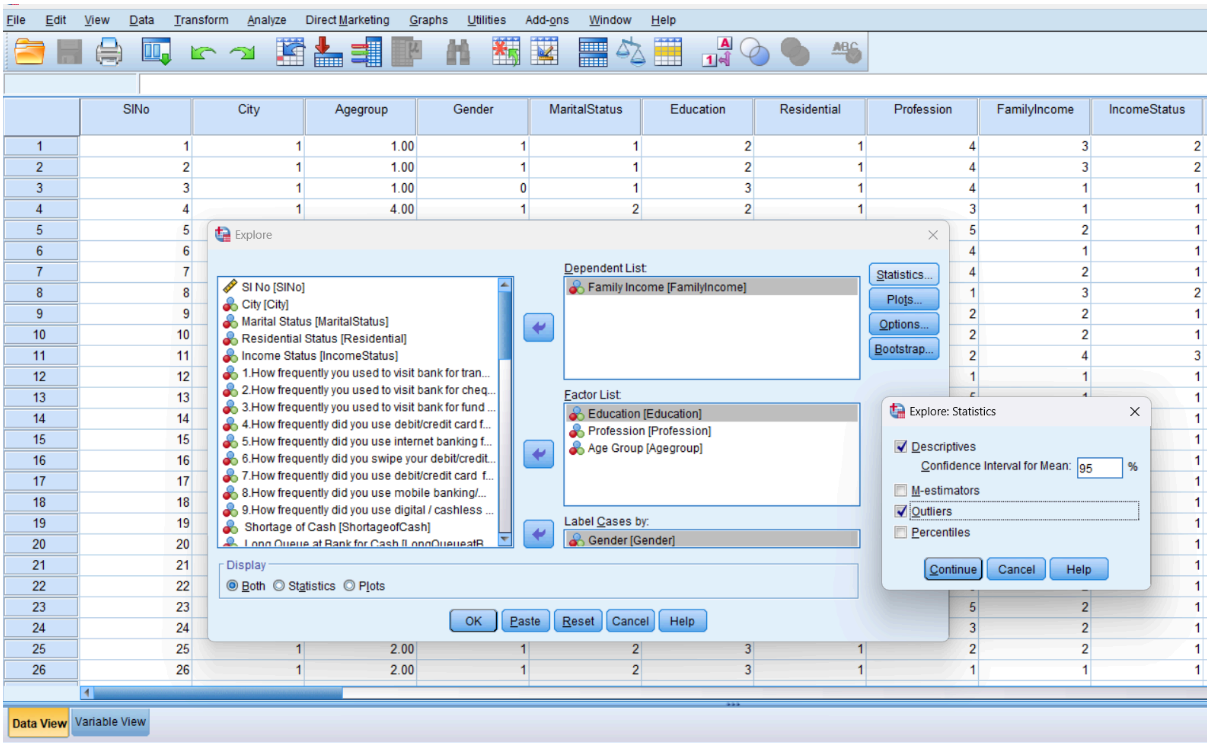
Explore

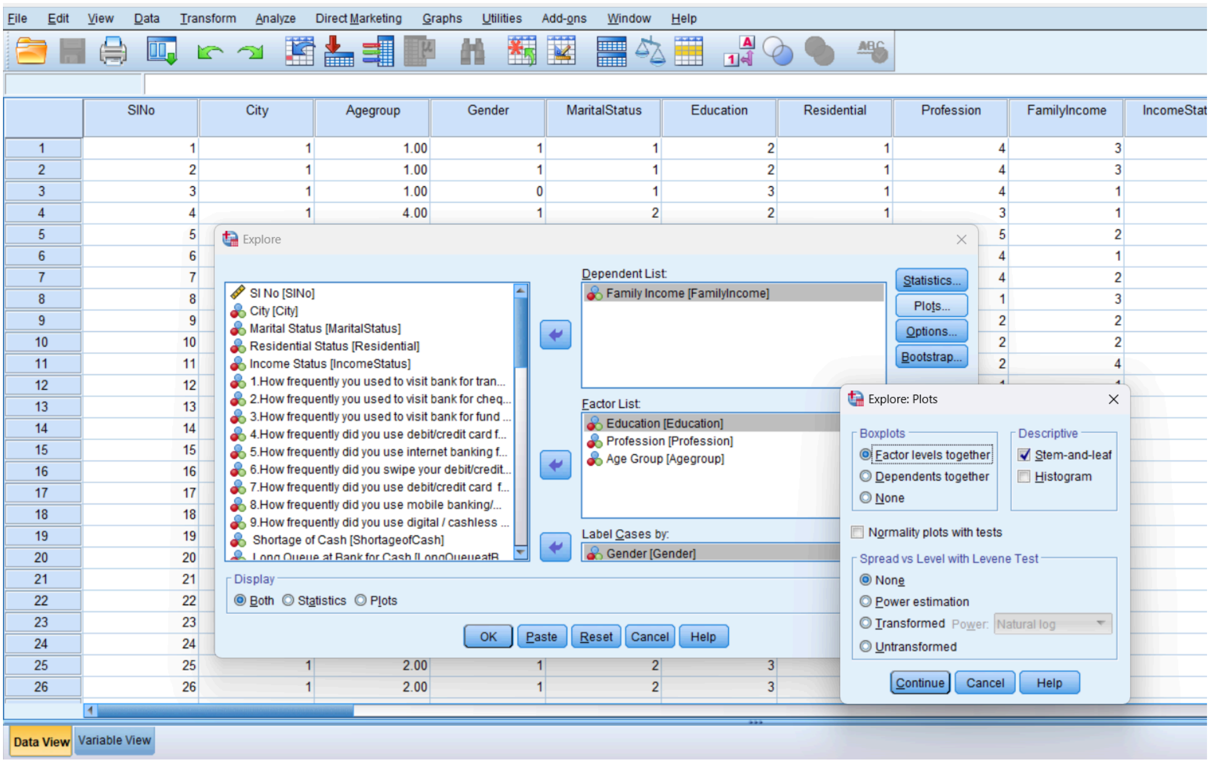
Explore is a more advanced and flexible option that provides both **numerical and visual summaries**.

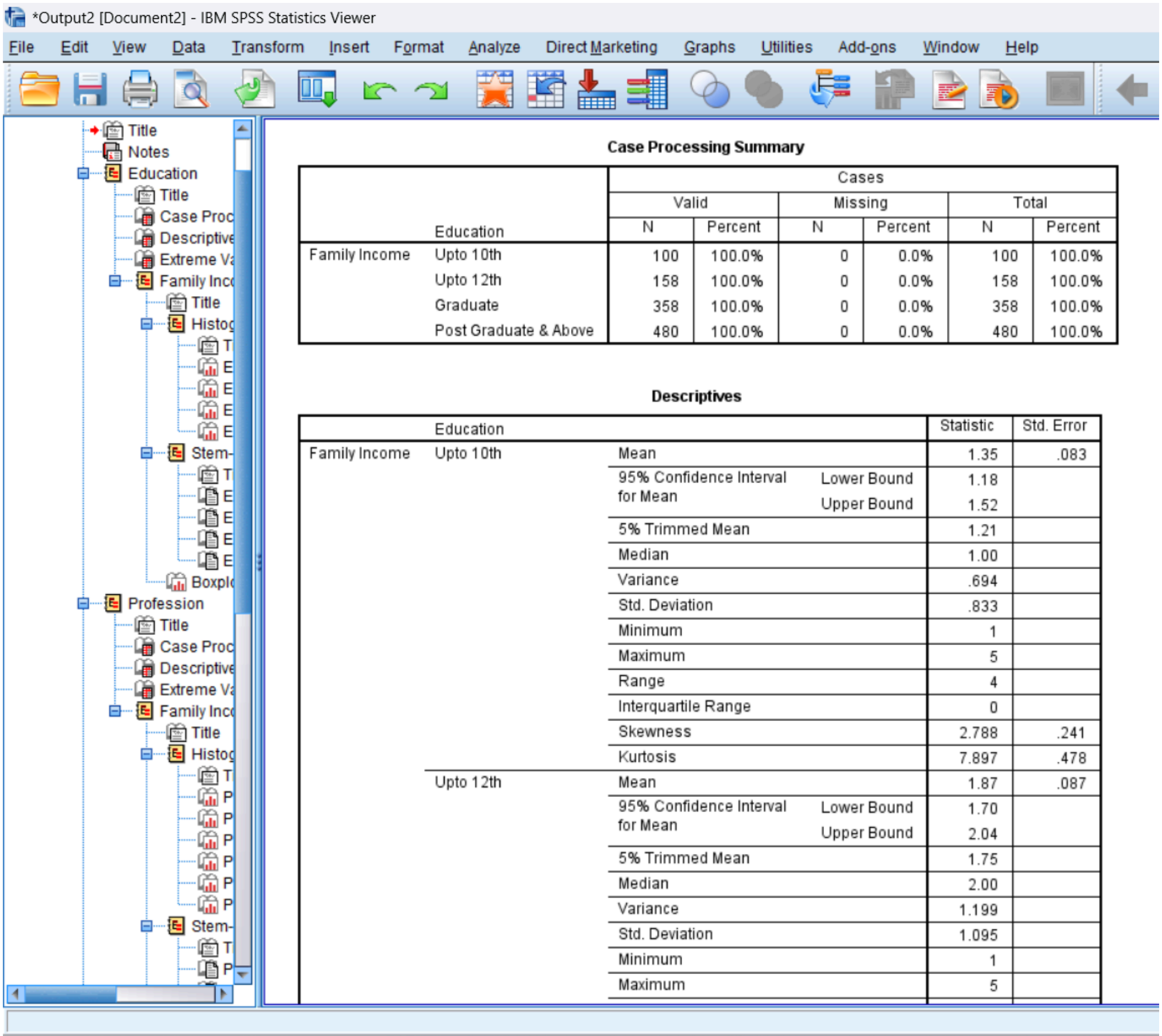
- Go to: **Analyze** → **Descriptive Statistics** → **Explore**
- Place variables into **Dependent List** (scale variables) and **Factor List** (categorical variables).
- Outputs include:
  - Descriptive statistics (mean, median, percentiles)
  - Boxplots
  - Histograms
  - Normality tests (Shapiro-Wilk, Kolmogorov-Smirnov)

- Confidence intervals

Use Case: Comparing distributions across groups or checking for normality before applying parametric tests.







Case Processing Summary

	Education	Valid N	Valid Percent	Missing N	Missing Percent	Total N	Total Percent
Family Income	Upto 10th	100	100.0%	0	0.0%	100	100.0%
	Upto 12th	158	100.0%	0	0.0%	158	100.0%
	Graduate	358	100.0%	0	0.0%	358	100.0%
	Post Graduate & Above	480	100.0%	0	0.0%	480	100.0%

Descriptives

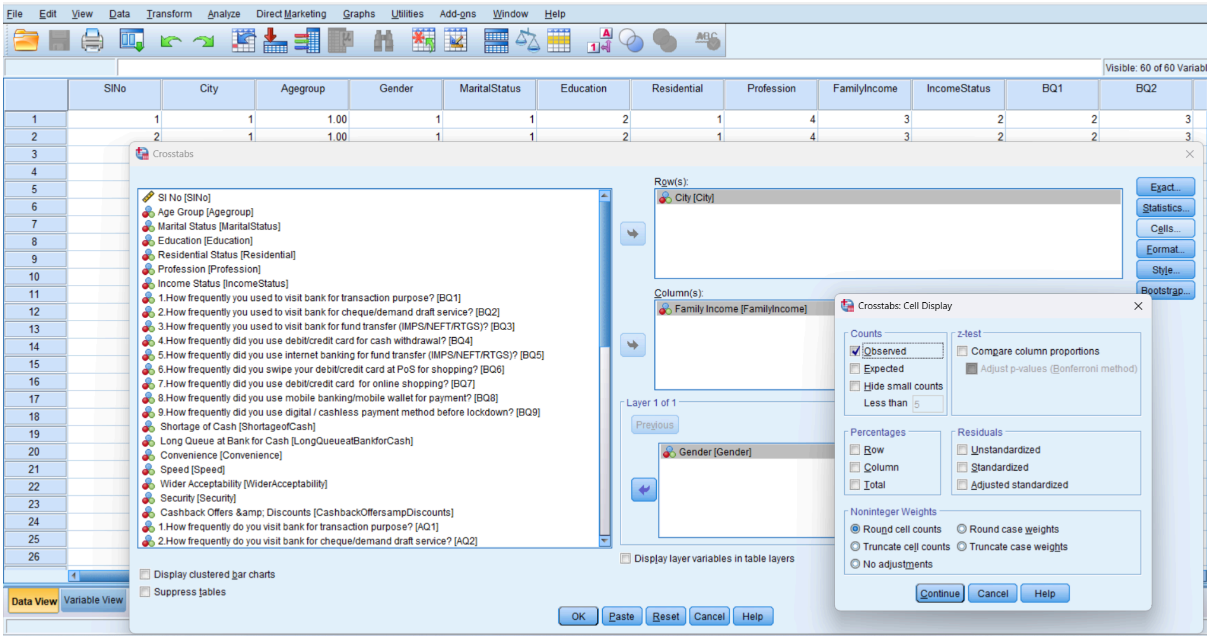
	Education			Statistic	Std. Error
Family Income	Upto 10th	Mean		1.35	.083
		95% Confidence Interval for Mean	Lower Bound	1.18	
			Upper Bound	1.52	
		5% Trimmed Mean		1.21	
		Median		1.00	
		Variance		.694	
		Std. Deviation		.833	
		Minimum		1	
		Maximum		5	
		Range		4	
		Interquartile Range		0	
		Skewness		2.788	.241
		Kurtosis		7.897	.478
	Upto 12th	Mean		1.87	.087
		95% Confidence Interval for Mean	Lower Bound	1.70	
			Upper Bound	2.04	
		5% Trimmed Mean		1.75	
		Median		2.00	
		Variance		1.199	
		Std. Deviation		1.095	
		Minimum		1	
		Maximum		5	

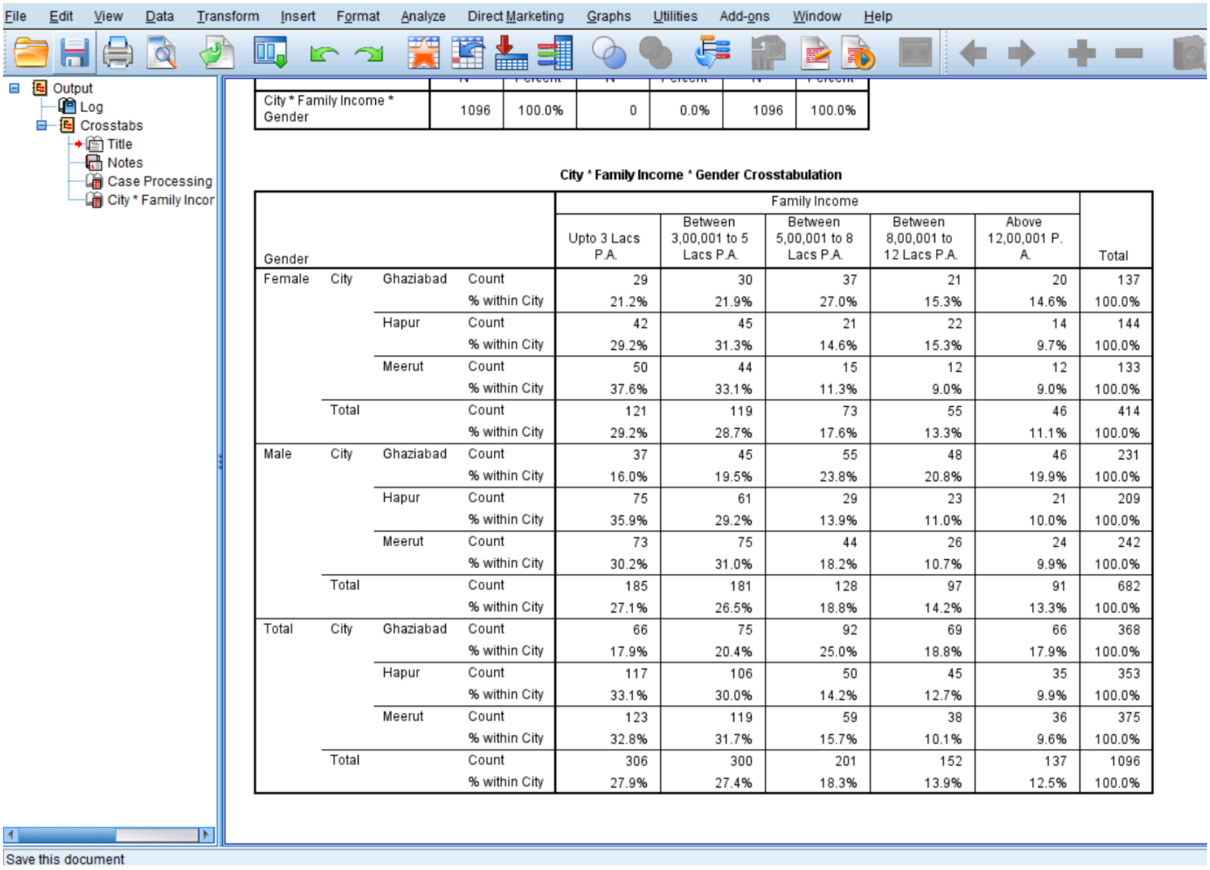
Crosstabs

Used to study relationships between **two categorical variables**.

- Go to: **Analyze** → **Descriptive Statistics** → **Crosstabs**
- Place one variable in the **Row** box and another in the **Column** box.
- Add optional statistics such as:
    - Chi-square test
    - Risk estimates
    - Measures of association (Cramer’s V, Phi, Contingency Coefficient)
- You can also request **cell percentages** (row, column, total).

Use Case: Analyzing the relationship between gender and voting preference.





City * Family Income * Gender	1096	100.0%	0	0.0%	1096	100.0%

City * Family Income * Gender Crosstabulation

Gender				Family Income: Upto 3 Lacs P.A.	Between 3,00,001 to 5 Lacs P.A.	Between 5,00,001 to 8 Lacs P.A.	Between 8,00,001 to 12 Lacs P.A.	Above 12,00,001 P.A.	Total
Female	City	Ghaziabad	Count	29	30	37	21	20	137
			% within City	21.2%	21.9%	27.0%	15.3%	14.6%	100.0%
		Hapur	Count	42	45	21	22	14	144
			% within City	29.2%	31.3%	14.6%	15.3%	9.7%	100.0%
		Meerut	Count	50	44	15	12	12	133
			% within City	37.6%	33.1%	11.3%	9.0%	9.0%	100.0%
	Total		Count	121	119	73	55	46	414
			% within City	29.2%	28.7%	17.6%	13.3%	11.1%	100.0%
Male	City	Ghaziabad	Count	37	45	55	48	46	231
			% within City	16.0%	19.5%	23.8%	20.8%	19.9%	100.0%
		Hapur	Count	75	61	29	23	21	209
			% within City	35.9%	29.2%	13.9%	11.0%	10.0%	100.0%
		Meerut	Count	73	75	44	26	24	242
			% within City	30.2%	31.0%	18.2%	10.7%	9.9%	100.0%
	Total		Count	185	181	128	97	91	682
			% within City	27.1%	26.5%	18.8%	14.2%	13.3%	100.0%
Total	City	Ghaziabad	Count	66	75	92	69	66	368
			% within City	17.9%	20.4%	25.0%	18.8%	17.9%	100.0%
		Hapur	Count	117	106	50	45	35	353
			% within City	33.1%	30.0%	14.2%	12.7%	9.9%	100.0%
		Meerut	Count	123	119	59	38	36	375
			% within City	32.8%	31.7%	15.7%	10.1%	9.6%	100.0%
	Total		Count	306	300	201	152	137	1096
			% within City	27.9%	27.4%	18.3%	13.9%	12.5%	100.0%

Ratio

This tool is used to compute **ratios between two variables**, often used in business and finance.

- Go to: **Analyze** → **Descriptive Statistics** → **Ratio**
- Define the numerator and denominator variables.
- Request statistics like:
    - Mean ratio

- Confidence intervals
- Standard error

Use Case: Calculating income-to-debt ratios or productivity rates.

Q-Q Plots (Under Legacy Dialogs)

Although more graphically oriented, Q-Q (Quantile-Quantile) plots are available for testing **normality**.

- Go to: **Analyze** → **Descriptive Statistics** → **Q-Q Plots**
- Select the variable of interest.
- SPSS plots expected normal distribution against actual data distribution.

Use Case: Checking whether a variable approximates a normal distribution.

7.3 Inferential Analysis

Inferential analysis is a chapter within the field of statistics which aims to form predictions or conclude on a population based on an analysis of a sample taken from the population. Due to time, cost, or resource restrictions, a complete population dataset is generally not attainable. Therefore, researchers tend to use simplified, smaller, and more representative datasets and sample populations in their analysis. In most cases, the samples from the dataset under analysis are construed to represent the entire population, which further underlines the importance of inferential analysis in a research study.

7.3.1 Role of Probability

Nothing in this world is risk free. The more assumptions, the greater the risk of error. Therefore, along with rest of the field, inferential analysis too relies in part on estimates generated from probability of an outcome. When researchers draw conclusions from a sample, there is always a chance of error. If a study tells us something with a certain degree of confidence, and researchers were to repeat the study multiple times, the researchers would arrive at the same conclusion in a certain percentage of cases and there would be a certain percentage of cases in which the researchers would be wrong. Probability theory assists in determining the precise percentage of error and in this situation, the estimates tell us that the error is 5 percent.

7.3.2 Tests of Significance

Every so often in analysis, there are supposed to be patterns of data that might occur by chance or for real, so it is up to the analysis to use tests of significance to find the right pattern. Some of these tests include t tests, which are set to find the means in groups of two, while the other is for categorical variables, and called a chi-square test. ANOVA (Analysis of Variance) tests other groups in a more complex manner, while other tests, set to find relationships, include simple and complex correlation, as well as regression. All these proofs are found from the data collected without manipulation.

7.3.3 Hypothesis Testing

Every single analysis starts with a clear hypothesis, the rest of civilized hypothesis are the more of a ‘construct’. Based on the two opposite categories of the statement (which is basically in 2 forms), the data collected is taken and the sample becomes a central ‘piece’. Statistically, it is figured out if the statement is true or false. Under the null hypothesis, the data becomes significant, therefore as the probability of the occurrence is on a lower boundary in the set of accepted results.Summarization of results learned

To summarize results learned, consider a single research that analyzes test scores of 200 students from a city. The applicability of inferential analysis allows a researcher to comments of the performance of students in the city of the city. As with the rest of the dataset, the ability to make inferences using the dataset is the hallmark of scientific reasoning and decision-making.

7.3.4 Research and Its Implementation

From medicine, where a clinical trial with a few patients allows a researcher to make assumptions about the rest of the patients population, to education where a researcher might evalute a small sample of schools to understand the impact of specific teaching techniques across a region, to economics where micro studies of households are used to estimate the entire country, inferential statistics is a highly valued analysis across fields. It allows the user not to get stuck with limited information, rather empowers the user to look beyond.

Usefulness of Inferential Analysis

Use of money in selling goods has a complex nature of the scronomy. Use of money in selling goods has a complex nature of the economy. There is absolutely need to explain the nature of money in selling goods. There is absolutely need to explain the nature of money in selling goods.

7.4 Hypothesis

The assumption made regarding a population parameter is termed a hypothesis in statistics. Since population studies might not be feasible, researchers gather sample data to prove or disprove the assumption regarding the population. In the field of inferential statistics, the ability to conduct hypothesis testing is a crucial component, as it enables one to make data-driven decisions in place of wild conjectures. While a statistical hypothesis does have a possibility of being incorrect, statistical tests exist to either prove or disprove the hypothesis.

The steps regarding hypothesis testing involve the formation of a claim, sample data collection, data analysis through various statistical approaches, and then finally, the determination of how the claim rests to the available evidence. In any research, it is necessary to conduct tests to be sure that the outcomes have been acquired scientifically.

Null Hypothesis (H_0)

- The null hypothesis is the starting point of any hypothesis test.
- It is a formal statement that there is **no effect, no relationship, or no difference** in the population.
- The purpose of the null hypothesis is to serve as a benchmark that can be tested against sample data.
- For example, a null hypothesis might state that "there is no difference in average exam scores between male and female students" or "a new medicine has no effect compared to the existing one."
- In hypothesis testing, researchers attempt to **disprove or reject** the null hypothesis. Rejecting it means that there is sufficient statistical evidence to support the belief that something different or significant exists in the population.
- If the evidence is weak, the null hypothesis cannot be rejected, and researchers conclude that the sample does not provide enough proof of a real difference or effect.

In short, the null hypothesis represents the position of "no change" or "status quo" until the data shows otherwise.

Alternative Hypothesis (H_1 or Ha)

- The alternative hypothesis is the opposite of the null hypothesis.
- It states that there **is an effect, a relationship, or a difference** in the population.
- While the null hypothesis is about rejecting change, the alternative hypothesis represents the researcher's belief or expectation.
- For example, if the null hypothesis says that "there is no difference in exam scores between male and female students," then the alternative hypothesis would state that "there is a significant difference in exam scores between male and female students."
- Similarly, if the null says "the medicine has no effect," the alternative would say "the medicine has a positive effect."
- If the null hypothesis is rejected after testing, the alternative hypothesis is accepted as the more likely explanation.

The alternative hypothesis is what researchers hope to prove, but statistical methods require that it only be accepted if the null hypothesis is shown to be unlikely.

7.4.1 Importance of Hypothesis Testing

- **Provides a framework for research**: Hypotheses give direction to research by setting a clear assumption to test.
- **Helps in decision-making**: Hypothesis testing guides whether to accept or reject assumptions, useful in business, medicine, education, and social sciences.
- **Minimizes subjectivity**: Instead of relying on opinions, hypothesis testing uses data and probability.
- **Supports scientific progress**: By continuously testing and revising hypotheses, knowledge grows in a systematic and reliable way.

7.4.2 Examples of Null and Alternative Hypotheses

a. Expenditure on the Education of Male and Female Children

In this case the hypothesis deals with the expenditure on education of male and female children. The null hypothesis (H_0), states that there is no difference and the average amount spent is equally divided across both. In mathematical form, this is represented as H_0: $\mu_1=\mu_2$, where μ_1 is mean expenditure on male children and μ_2 is mean expenditure on female children. The contrary hypothesis (H_1) says that the average expenditure is different, i.e. H_1: $\mu_1 \neq \mu_2$. This is a two-tailed test because the difference can be either higher or lower. In social sciences, such hypothesis are used to analyze the extent of discrimination on educational expenditure based on gender. The null hypothesis is rejected, it means that the spending of parents on male and female children is not equal.

b. Reduction of Perception Differences of Employees Between BSNL and Airtel

In this case, focus will be given to the variance instead of the mean. The null hypothesis, H naught, assumes that the level of perception of BSNL Employees and Airtel Employees towards HRD practices is the same, and is formulated as H naught: sigma one square is equal to sigma two square. The opposing hypothesis, H one, states that the variances are not the same, that is, H one: sigma one square is not equal to sigma two square. This kind of hypothesis is tested with an F-test for equality of variances. The moment the null is rejected, it would mean the two groups of employees differ substantially in the degree of perception or agreement towards HRD practices, which, in turn, means that the HRD policies and practices are organizationally perceived in diverse manners.

c. Supporting Factors by Age Cohorts on Purchasing High Involvement Products (HIPs)

For our case, the researcher aims to find out if there age groups amongst the consumers that are different in the reason why they purchase high involvement products (cars, electronics, real estate etc.). The null hypothesis (H_0) is that there is no difference in age group and the influencing factors there are. The alternative hypothesis H_1 says that there is in fact a difference. In the case presented, difference in the age group and the influencing factors would be tested using a Chi-square test of independence. If the null is rejected, the results would indicate that there is a valuable difference in the age of consumers and the high involvement products that they purchase.

d. Life of a LED Bulb

This is an example of verifying a claim made by a manufacturer. The company states that on average its LED bulb lasts for three years. The null hypothesis (H_0) assumes that its mean lifespan is three years, that is, H_0: $\mu = 3$ years. The alternative hypothesis (H_1) posits that the mean lifespan is, in fact, less than three years, that is, H_1: $\mu < 3$ years. This is a one-tailed test, since the only interest is in whether the life span is below the claimed value, and not above. Such verifications are essential in quality control and consumer protection. The company's warranty claim is not valid if the null hypothesis is rejected, and in that case, the consumers are not likely to get the value which is promised.

7.4.3 Significance of These Hypotheses

- These examples demonstrate how hypotheses are formulated in real-world contexts, whether comparing means, variances, relationships, or verifying claims.
- The null hypothesis always assumes equality, no difference, or no effect, while the alternative represents change or difference.
- Choosing between a **one-tailed** or **two-tailed test** depends on whether the researcher is looking for any difference (two-tailed) or only a specific direction of difference (one-tailed).
- Testing these hypotheses through statistical tools like t-tests, F-tests, Chi-square tests, and Z-tests allows researchers to make evidence-based conclusions rather than relying on assumptions.

A hypothesis is an assumption about a parameter of the population which the researcher intends to investigate using sample data. The null hypothesis (H_0) makes no difference and is a default position assuming no difference or effect. It is also alternative hypothesis (H_1 or Ha) which tells the researcher that there is a difference or an effect. Hypothesis testing is a methodology which defines a framework in which these assumptions are confirmed or rejected on the basis of statistics. Researchers carry out scientific deduction to prove that their conclusions are not mere assumptions.

Formulation and testing of the hypothesis is a key part of the research. The above are different examples of activities in which a researcher is engaged; gender analysis of educational expenditure, variance in employee attitude perception, change in consumer age behavior and claimed product verification. Each of the hypothesis begins with a null assumption and the researcher only changes the null to alternative when there is overwhelming statistical evidence. It is this reasoning that makes research, business and policy decisions scientifically grounded rather than assumptions or guesses.

7.5 Statistical Significance

A relationship between two or more variables in a dataset may be due to random chance or statistically significant. If a relationship is significant, the odds of it being a coincidence is very low result as meaningful. Therefore, statistically significant result meaningful observation. Significant result provides evidence to support a hypothesis and draws a reliable conclusion.

Statistical hypothesis testing involves four concepts such as the null hypothesis, sample data, statistical tests, and p-value significance. The null hypothesis states there is no effect, no difference, or no relationship between variables. A researcher collects sample data and tests it to identify patterns or trends within a sample and make inferences to a larger population. A sample may be tested using a t-test, chi-square test, ANOVA, or regression analysis. These tests calculate a p-value which is the probability of the null hypothesis is true.

The 5 percent (0.05) level is frequently a cutoff standard to evaluate significance. A p-value that is equal to or less than 0.05 is statistically significant, and there is enough evidence to reject the null hypothesis. Additionally, some studies take a more rigorous approach and adopt the 1 percent (0.01) level. For a p-value of 0.01, that implies that there is only a 1 percent probability of an error, and conversely, the level of confidence is 99 percent.

7.6 Examples of p-values

- **Sig. = 0.007**: This value is less than 0.01, which means the result is statistically significant at the 1% level. The probability of error is less than 1%, giving the researcher 99% confidence in rejecting the null hypothesis.
- **Sig. = 0.013**: This p-value is greater than 0.01 but less than 0.05, making the result significant at the 5% level. The error probability is about 1.3%, and the result can be accepted with 98.7% confidence.
- **Sig. = 0.041**: This is just below 0.05, meaning it is statistically significant at the 5% level, with about a 4.1% error and 95.9% confidence.
- **Sig. = 0.068**: Since this is greater than 0.05, it is not statistically significant at the 5% level. The result may have occurred by chance, so the null hypothesis cannot be rejected.
- **Sig. = 0.521**: This is very high, suggesting no statistical significance at all. The probability of error is 52.1%, so the result is highly unreliable for rejecting the null hypothesis.
- **Sig. = 0.156**: This is also greater than 0.05, indicating non-significance, with about a 15.6% error rate.

These examples show how p-values are interpreted and why they are critical for decision-making in research.

7.6.1 Importance of p-value in Analysis

In order to understand the essence of statistical significance testing, one must first understand the p-value. The primary reason one attaches value to the p-value is because it measures the evidence as to the validity of the null hypothesis. A statistically significant p-value (≤ 0.05) demonstrates that the sample data is inconsistent with the null hypothesis, and the null hypothesis is likely to be rejected. On the contrary, a p-value greater than 0.05 indicates the data is such that the results could happen and the null hypothesis may not be rejected.

For significance of 5%, the researcher is 95% confident in rejecting a null hypothesis. The margin of error will imply even greater confidence, so that at the 1 % significance level, the researcher is 99% confident, and there is only a 01% chance of error. A p-value of 0.04 indicates a 4% chance of error, thus, the researcher must still be 96% confident in the claim of the hypothesis .

With the p-value as a guide, results which are strong and should be favored when rejecting the null hypothesis will be in favor of the alternative hypothesis. the researcher will maintain objectivity, as the p-value will aid in making the necessary decision.

Ensuring that research results are not purely due to coincidence, and are therefore meaningful and reliable, leads to research results becoming statistically significant. The p-value is instrumental in this by providing a quantifiable probability of error. Results that have p-values that are lower than 0.05 are widely considered statistically significant, and results that have p-values above this threshold are considered statistically not significant. The correct interpretation of statistical significance is important to draw logical conclusions and test hypotheses to make reliable decisions in research in various fields including medicine and social sciences, among others.

7.6.2 Accepting or Rejecting Hypotheses Depending on P-Value

In hypothesis testing, the p-value, or the level of significance, functions as the most crucial element. It is the probability of the underlining null hypothesis being true, and the observed results, or more extreme results, occurring purely by chance. It critically evaluates if the findings of research are statistically significant, thereby helping analysts decide if the null hypothesis should be accepted or rejected. The interpretation of the p-value depends on whether a parametric or non-parametric test is used.

7.6.3 Parametric Tests (t-Test, F-Test, Levene’s Test, etc.)

When data is continuous, and some distribution-related assumptions (particularly normality) are reasonably met, parametric tests can be used, the most common of which is the t-test for the mean of two groups. Other tests include the F-test for variance comparison and Levene's Test for the equality of variance.

In parametric testing, the null hypothesis (H_0) states equality across the measures, such as equal means/variances. Subsequently, the outcome from the hypothesis test (the p-value) is compared against a predetermined significance level, which is generally 5% (0.05) or 1% (0.01). When the p-value turns out to be lower than the level of significance, the level of the null hypothesis is revised as the probability of observing the data under the null hypothesis is extremely low. In such circumstances, the null hypothesis may be accepted to imply the difference or variability between the groups is not statistically significant at that level. In contrast, when the p-value exceeds the level of significance, the difference falls to be important thus the null hypothesis may be rejected, which implies that a difference or effect is present in the population.

Significance levels measure confidence. 5% of significance corresponds to 95% confidence that the result is not due to chance. 1% corresponds to 99% confidence. The lower the significance level, the more the confidence required to justify the null hypothesis.

Understanding p-value is one of the elements needed to grasp the process of hypothesis testing. In hypothesis testing, the p-value is the probability of observing the results of a test, or something more extreme, given that the null hypothesis is true. Acceptance or rejection of the hypothesis is dependent on it's a parametric (such as a t-test or F-test) or a non-parametric test (such as Chi-square).

- **Case 1: p-value < 0.05 (5% significance level)**
    - At the 5% significance level, if the p-value is less than 0.05, the null hypothesis (H0) is **accepted**.
    - This implies there is **no significant difference** between the average values of the two variables under consideration.
    - For example, if testing H0: $\mu 1 = \mu 2$, a p-value less than 0.05 means we are confident (95% confidence level) that the averages are not significantly different.
- **Case 2: p-value < 0.01 (1% significance level)**
    - At the stricter 1% significance level, if the p-value is less than 0.01, the null hypothesis (H0) is again **accepted**.
    - This suggests that there is **no significant difference** between the variability (variance) of the two variables.
    - Example: In variance testing H0: $\sigma 1 = \sigma 2$, a p-value < 0.01 assures us (99% confidence level) that the variances are equal.

Summary:

- **$p < 0.05$ → H0 accepted** at 5% level → No significant difference in means.
- **$p < 0.01$ → H0 accepted** at 1% level → No significant difference in variances.

7.7.1 Non-Parametric Tests (Chi-Square Test)

Non-parametric tests apply to data types that are categorical or when data does not fulfill theory assumptions. A common example of such tests is the Chi-square test which analyzes whether or not two categorical variables are independent of one another or associated. Here, the null hypothesis is that the two variables are independent or there is no association between them.

In a Chi-square test, the p-value, or asymptotic significance, is compared to a critical value of the significance threshold. If the p-value is below 0.05, this evidences the variables are not independent, meaning one is associated with the other. If the p-value is above 0.05 the null hypothesis is accepted, that any association is merely due to random chance, and independence of the variables can be assumed.

7.7.2 For Non-Parametric Tests (Chi-Square Test)

- **Case 1: p-value < 0.05 (5% significance level)**
  - For Chi-square test, if the p-value (asymptotic significance) is less than 0.05, the null hypothesis (H0) is **rejected**.
  - This means that the assumption of independence between two attributes does not hold.
  - In other words, there is a **significant relationship** between the two attributes being tested.
- **Case 2: p-value ≥ 0.05**
  - If the p-value is greater than or equal to 0.05, the null hypothesis (H0) is **accepted**.
  - This implies that the two attributes are independent of each other, and no significant association exists.

Summary:

- **$p < 0.05$ → Reject H0** → Attributes are not independent.
- **$p \geq 0.05$ → Accept H0** → Attributes are independent.

7.8 General Principles of Accepting or Rejecting Hypotheses

Traditional interpretations of p-value are deeply rooted in scientific research methodologies. Hence, focusing only on p-value in domains and industries is the best part of statistical significance testing. The scientific inquiry begins with framing a null and alternative hypothesis and ends with defining and executing a hypothesis testing procedure suitable for the collected dataset and research goal. Each statistical technique requires a specific p-value and significance level. The level of significance needed determines the threshold p-value.

Selecting a level of significance, for instance, 0.05, and getting the p-value of the analysis helps in satisfying the logical hypothesis framework. A p-value on the threshold defends the null hypothesis. A p-value under the significance level suggests that the data provides clear evidence that defends the alternative hypothesis, and contradicts the null hypothesis.

However, there is a misconception that 'proof' and 'verify' are interchangeable. Accepting the null hypothesis is a situation where the statistical evidence is inadequate to disprove the hypothesis, and thus accept it. In the same context, rejecting the hypothesis shows that based on the data, there is a possibility of a chance that the hypothesis does not hold. Hence it provides a logical framework to defend the alternative hypothesis.

Key Takeaways

- Parametric tests focus on **mean and variance comparison**, where a low p-value leads to acceptance of H0 (no difference).
- Non-parametric tests like Chi-square focus on **association between attributes**, where a low p-value leads to rejection of H0 (dependence exists).
- The choice of test and interpretation depends on the type of data, research design, and hypothesis structure.

In summary, the p-value acts as a threshold for decision-making in hypothesis testing. Correctly interpreting it ensures valid statistical conclusions.

A hypothesis is either accepted or rejected based on the p-value which is crucial in statistical analytics. The focus of a parametric test is on means or variances, whereas the focus of a non-parametric test is on associations between sets of categorical variables. Decision-making within research is made thanks to the p-value, which is concerned with informing how research outcomes are derived such as avoiding entertaining randomness. p-values must be viewed within the context of the significance level, which helps to make rational conclusions. This should be done in order to adequately address the hypotheses. This helps in ensuring the analysis is meticulous and dependable.

7.9 Parametric and Non-Parametric Tests

In statistical analysis, tests are categorized into two: parametric tests and non-parametric tests. The choice depends d the data characteristics, the premises regarding the population, and the aims of the study. These two propositions must be distinguished for proper analysis and interpretation of the results.

Parametric Tests

These are statistical tests that have certain hypotheses pertaining the population which the sample came from. The most common is the population is normally distributed. Also, parametric tests assume the homogeneity of variance. This is the assumption that the spread, or variability of the data in question, in the different groups is somewhat the same. These are type data, called intervals or ratios, which have characteristics of continuous measurements or data. Such as, weight, salary, income, or scores of the tests.

Parametric tests make use of a central tendency, the mean, and focus on the interrelationship of the data assuming the datasets are interdependent. Due to these assumptions, it is a popular notion that parametric tests are more statistically powerful as it is non parametric tests which are less sensitive in establishing the existence of any statistically significant differences.

The most common parametric tests are:

- **Pearson correlation** to measure linear relationships between variables.
- **Independent-measures t-test** for comparing the means of two independent groups.
- **One-way ANOVA** for comparing means of more than two independent groups.
- **Paired-sample t-test** for comparing means of the same group under two conditions.
- **Repeated measures ANOVA** for comparing means across multiple related conditions.

Users must recognize the population information tabulated, meeting certain conditions, must be independent of other populations. No matter how certain these assumptions are, insights drawn from them are likely to be greatly affected by outliers and other data integrity concerns. These insights are always valid if data available is continuous, grouped as a normal distribution and the variances are similar for all groups.

Non-Parametric Tests

Non parametric tests do not make assumptions on the distribution of the population and can be used for ordinal, nominal as well as interval or ratio data which can be ranked. Non parametric tests are useful in cases where the assumptions of parametric tests are not met, for instance in the case of skewed distribution, small sample sizes or unequal variances.

In place of means parametric tests focus on the medians as the measure of central tendency. These tests can be conducted for independent as well as related samples. They are more pragmatic in situation where the parametric assumptions are not met.

7.10 Common non-parametric tests

- **Spearman correlation** for measuring monotonic relationships between variables.

- **Mann-Whitney U test** for comparing two independent groups.
- **Kruskal-Wallis test** for comparing more than two independent groups.
- **Wilcoxon signed-rank test** for two related groups or paired samples.
- **Friedman test** for multiple related conditions.

Non-parametric tests are most optimal when all available data is concealed, the framework is random, or sets cannot stand alone. Their ability to correct similarities is more reliant on skewed variations than parallel measures.

Differences between Parametric and Non Parametric tests include:

1. **Assumptions:** Parametric tests require normality and homogeneous variance; non-parametric tests make no such assumptions.
2. **Data Type:** Parametric tests handle interval or ratio data; non-parametric tests handle ordinal, nominal, or ranked data.
3. **Central Measure:** Parametric tests use the mean; non-parametric tests use the median.
4. **Power:** Parametric tests are more powerful under correct assumptions; non-parametric tests are less sensitive but more robust to violations.
5. **Effect of Outliers:** Parametric tests can be heavily influenced by outliers; non-parametric tests are less affected.
6. **Application:** Parametric tests are suitable for independent, normally distributed samples; non-parametric tests can handle dependent, skewed, or heterogeneous samples.

Both parametric and non-parametric tests are important methods of statistical analysis, each serving different purposes. While powerful and exact, parametric tests bring heavy assumptions on how normal and how much variance exists, making strong assumptions on how well-controlled and well-behaved data are around a given sample mean. Non-parametric tests have more freedom and are much more robust, reaching for data without a parametric assumption, even those which are categorical or data skewed on one side. All of these considerations around the exact choice of a method signal the importance and the reliability of the conclusions derived, which would allow the researchers to make the right conclusions about the population which was the target of the study.

	Parametric	**Non-parametric**
Assumed distribution	Normal	Any
Assumed variance	Homogeneous	Any
Typical data	Ratio or Interval	Ordinal or Nominal

Data set relationships	Independent	Any
Usual central measure	Mean	Median

Tests		
Choosing	Choosing parametric test	Choosing a non-parametric test
Correlation test	Pearson	Spearman
Independent measures, 2 groups	Independent-measures t-test (Independent Sample t test)	Mann-Whitney test
Independent groups, >2 groups	One-way, independent-measures ANOVA	Kruskal-Wallis test
Not Independent groups, 2 conditions	Matched-pair t-test (Paired Sample t test)	Wilcoxon test
Not Independent groups, >2 conditions	One-way, repeated measures ANOVA (ANOVA for correlated samples)	Friedman's test

Parametric	Non-parametric
Parametric analysis to test group means	Nonparametric analysis to test group medians
Information about population is completely known	No Information about the population is available
Specific assumptions are made regarding the population	No assumptions are made regarding population
Applicable only for variable	Applicable to both variable and attributes
Samples are independent	Not necessarily the samples are Independent
Assumed normal distributions	No Assumed Shape / distribution
Handles Interval data or Ratio data	Handles Ordinal data, Nominal (or Interval or Ratio), ranked data
Results can be significantly affected by outliers	Results cannot be seriously affected by outliers

Perform well when the spread of each group is different, might not provide valid results if groups have a same spread	Perform well when the spread of each group is same, might not provide valid results if groups have a different spread
Have more statistical power	It is not so powerful like parametric test

7.11 Reliability Analysis

Reliability analysis determines the consistency and the trustworthiness of the collected data. In research, reliability is defined as the extent to which an instrument used to collect data, such as a questionnaire or a scale, is able to yield the same results repeatedly under the same conditions. There is a need to ensure reliability as a measure of accentuated practicality of an issue considering existing findings and prospective studies.

One of the most preferred approaches to measure reliability is using Cronbach's Alpha (α) as a measure of the internal consistency among a set of items or indicators which seek to measure the same phenomena. Internal consistency is a measure of the degree to which the indicators form a cohesive whole. In scales where the items are highly correlated, the alpha coefficient will also be high as it highly reliable. In the reverse scenario, the items are loosely correlated, the items in the scale are highly likely as a set to be measuring the concept in question in a highly divergent manner.

7.12 Cronbach's Alpha (α)

The score of Cronbach's Alpha (α) which is set in a scale of 0 to 1 is also a descriptive measure of the number of the scale items and average >. inter-item correlation. It is a single coefficient which estimates the extent to which the items measure the same concept or underlying construct.

The formula for Cronbach's Alpha is:

$$\alpha = \frac{N \cdot \bar{c}}{\bar{v} + (N-1) \cdot \bar{c}}$$

Where:

- N = Number of items (questions) in the scale/test
- $\bar{c}$ = Average covariance between item pairs
- $\bar{v}$ = Average variance of each item

An alternative but equivalent form (using variances) is:

$$\alpha = \frac{N}{N-1}\left(1 - \frac{\sum_{i=1}^{N} \sigma_i^2}{\sigma_T^2}\right)$$

Where:

- N = Number of items
- σ_i^2 = Variance of item i
- σ_T^2 = Variance of the total test score (sum of all items)

The alpha coefficient ranges from 0 to 1. Higher values indicate better internal consistency among the items.

7.13 Interpreting Cronbach's Alpha

Although there is no universally accepted standard, researchers generally follow a rule of thumb:

- **α ≥ 0.70** → Acceptable reliability, indicating the scale is reasonably consistent.
- **α ≥ 0.80** → Good reliability, showing stronger consistency among items.
- **α ≥ 0.90** → Excellent reliability, suggesting very high internal consistency.
- **α < 0.50** → Unacceptable reliability, meaning the scale may not be measuring the construct consistently and requires revision.

It is important to note that these thresholds are **arbitrary** and may vary depending on the theoretical framework of the scale, the nature of the construct, and the research context. For example, in exploratory studies, a lower alpha may be acceptable, whereas in applied research, higher reliability is often expected.

7.14 Importance of Reliability Analysis

a) **Consistency Verification:** Reliability analysis confirms that the measurement instrument produces stable results over repeated trials.

b) **Instrument Evaluation:** It helps researchers identify problematic items that do not correlate well with the overall scale, which may need revision or removal.

c) **Validity Support:** A reliable instrument enhances the credibility of validity measures. An unreliable scale cannot be valid, since validity depends on consistency.

d) **Confidence in Research Findings:** Reliable data ensures that the observed patterns reflect actual trends in the population rather than measurement error or randomness.

The internal consistencies in survey-based and experimental research analysis are assessed using measures such as the Cronbach's Alpha which is a measure of internal consistencies in a scale. While higher alpha values are better. the context scale design and the underlying theory can change the interpretation. Most times, values of 0.70 and above are acceptable for the alpha values, the higher the alpha the better for more critical or applied studies. This analysis proves the usefulness of research instruments and ensures the subsequent conclusions and suggestions are based on solid data.

7.15 Step-by-Step Reliability Analysis in SPSS

Step 1: Prepare the Data

1. Enter your survey or questionnaire data into SPSS.
2. Each row should represent a respondent, and each column should represent an item or question of the scale measuring the same construct.
3. Ensure all items are coded consistently (e.g., 1–5 Likert scale).

Step 2: Open the Reliability Analysis Menu

1. In SPSS, go to the top menu and click:
   Analyze → **Scale** → **Reliability Analysis**
2. This opens the "Reliability Analysis" dialog box.

Step 3: Select Items for Analysis

1. In the dialog box, move all items that measure the construct into the **"Items"** box.
2. Under **"Model"**, ensure **Cronbach's Alpha** is selected (it is the default).

Step 4: Configure Additional Options (Optional)

1. Click **Statistics** to choose additional statistics, such as:
   - **Item-total statistics**: shows correlation of each item with the total scale.
   - **Scale if item deleted**: shows how Cronbach's Alpha changes if a specific item is removed.
2. These options help identify weak or inconsistent items.

Step 5: Run the Analysis

1. Click **OK** to run the analysis.
2. SPSS will produce a table called **"Reliability Statistics"** showing Cronbach's Alpha value.

Step 6: Interpret the Output

1. Cronbach's Alpha Value

- The **"Cronbach's Alpha"** coefficient ranges from 0 to 1.
- Interpretation guidelines:
  - **$\alpha \geq 0.90$** → Excellent reliability
  - **$0.80 \leq \alpha < 0.90$** → Good reliability
  - **$0.70 \leq \alpha < 0.80$** → Acceptable reliability
  - **$0.50 \leq \alpha < 0.70$** → Poor reliability, may require revision

- **α < 0.50** → Unacceptable reliability

2. Item-Total Statistics

- The **"Corrected Item-Total Correlation"** column shows how well each item correlates with the overall scale.
- Items with low correlation (commonly <0.3) may not fit well and could be removed to improve reliability.

3. "Cronbach's Alpha if Item Deleted"

- This column shows what the overall alpha would be if a particular item is removed.
- If removing an item increases the alpha, it indicates that the item is reducing overall consistency and may be considered for deletion.

Step 7: Take Action

- Based on the results, you may:
  - Keep items that strengthen consistency.
  - Revise or remove items that decrease reliability.
  - Recalculate Cronbach's Alpha after any changes to ensure the scale's reliability improves.

Key Points to Remember

a) Cronbach's Alpha assesses **internal consistency**, not validity. A reliable scale is not automatically valid.

b) Reliability can be influenced by the **number of items**; more items generally increase alpha.

c) The acceptable threshold for alpha depends on the **research context**. Exploratory research may accept lower alpha, while applied or clinical research prefers higher alpha.

d) Consistent coding and careful item design are crucial for meaningful reliability analysis.

Reliability Statistics

Cronbach's Alpha	N of Items
.915	21

ANOVA

		Sum of Squares	df	Mean Square	F	Sig
Between People		88.899	79	1.125		
Within People	Between Items	1314.615	20	65.731	44.865	.000
	Residual	2314.813	1580	1.465		
	Total	3629.429	1600	2.268		
Total		3718.328	1679	2.215		

Grand Mean = 2.96

7.16 Interpretation of Cronbach's Alpha

A Cronbach's Alpha (α) Result of 0.915 is very high which demonstrates the scale's internal consistency and reliability. It is considered social science research that 0.90 and above is classified as excellent, and a Cronbach's Alpha of 0.90 and above indicates the construct being measured highly correlates and works very well with the items used. This indicates that the scale is capturing a single construct and the respondents answered consistently across the items. Thus, the scale is not capturing different unrelated dimensions. A high alpha value indicates very few random measurement errors, which suggests that the scale results would be dependable if the scale were administered again under the same conditions. Thus, the results contribute reliability for further statistical tests such as factor analyzing structural models or testing predictions.

The **obtained Cronbach's Alpha (α) score is 0.915**, which is considered **excellent reliability** according to standard guidelines:

- **$\alpha \geq 0.90$** → Excellent internal consistency
- **$0.80 \leq \alpha < 0.90$** → Good
- **$0.70 \leq \alpha < 0.80$** → Acceptable

The results from ANOVA also test item reliability, which is an important aspect. Since the F-value to the F-test is 44.865, it means that the variance of the differences between the average item means is larger than the average variance of responses per item. This points to the likelihood that the respondents, on average, perceive the items distinctly. The p-value associated with the test is 0.000 which suggests that the variance is statistically significant at 0.05 levels of significance. This variance is something that can be considered useful. The items used in this survey can therefore be considered to have predictive power.

The **F-value = 44.865** with a **significant value (p-value) = 0.000** provides additional insights:

- The F-value in ANOVA compares the **variance between the items** with the **variance among respondents**. A high F-value indicates that the variability between group means (or items) is much larger than the variability within the groups.
- The p-value of **0.000** (less than the 0.05 significance threshold) indicates that the differences observed are **statistically significant**, and not due to random chance.

It is evident that these results yield very important conclusions. For instance, the scale can be considered consistent in addition to valid concerning meaningful difference among respondents. Also, the scale is high in psychometric strength, thus it can be used for advanced inferential examination. Moreover, the fusion of high reliability with significant findings from ANOVA proves the scale to be solid in construction which can be used for rigorous empirical study in regression, SEM, and comparative research.

7.17 Correlated Item-Total Correlation

The **item-total correlation** measures how well each individual item correlates with the total score of the scale. It indicates whether an item is consistent with the overall construct measured by the scale.

- A **high item-total correlation** suggests that the item is measuring the same underlying construct as the rest of the scale.
- A **low item-total correlation** (commonly below 0.3) indicates that the item may not fit well with the scale and could potentially reduce the reliability of the overall instrument.

In your analysis, all items have a **correlated item-total correlation above 0.3**, meaning each item contributes positively to the scale and aligns well with the construct being measured. This confirms the **internal consistency** of the items.

Cronbach's Alpha if Item Deleted

The statistic "Cronbach's Alpha if Item Deleted" in SPSS is crucial for ascertaining how each individual element impacts the scale's overall reliability. Sustaining internal consistency (Cronbach's Alpha), would a specific item's removal enhance internal consistency, neutral increment or sustain decrement, or worsen internal consistency? If an item is deleted, resulting in an increased alpha value, the item is said not to align harmonously with the rest, thus weakening the scale. The opposite scenario, where the alpha value decreases or remains unchanged and an item is deleted, means the item aids in enhancing reliability of the scale.

For a start, a score of 0.915, which is already an overall score of Alpha, is superb and thus indicative of strong internal relation among the items. The SPSS output captures the fact that for any individual item, as there is no item that reduces the reliability over the scale, one at a time, and worse, does any individual item increase the alpha value. Each item is aligned with the construct being measured and is functioning cohesively with the others. The scales in entirety are equally crucial and thus affirmatively, robust scale.

Moreover, the item-total correlations, which quantify the correlation of each item with the total of the remaining items, are all greater than the accepted minimum limit of 0.30. This confirms that each item is statistically relevant and is measuring the same underlying construct.

Key Takeaway

1. The **item-total correlations** confirm that all items are consistent with the overall scale and none fall below the minimum threshold of 0.3.
2. The **Cronbach's Alpha if Item Deleted** analysis indicates that deleting any item will not improve reliability, as the current alpha of 0.915 is already excellent.
3. Overall, the scale is **highly reliable**, with each item meaningfully contributing to the measurement of the intended construct.

7.18 Step-by-step guide for analyzing Correlated Item-Total Correlation and "Cronbach's Alpha if Item Deleted" in SPSS

Step 1: Prepare Your Data

1. Enter your survey or questionnaire data in SPSS.
2. Each row represents a respondent, and each column represents an item in the scale.
3. Make sure all items are coded consistently (e.g., Likert scale 1–5).

Step 2: Open Reliability Analysis

1. In SPSS, click:

   Analyze → Scale → Reliability Analysis
2. The **Reliability Analysis** dialog box will open.

Step 3: Select Items

1. Move all items that measure the same construct into the **"Items"** box.
2. Ensure that **Cronbach's Alpha** is selected under **Model**.

Step 4: Choose Statistics for Item Analysis

1. Click the **Statistics** button in the dialog box.
2. Check the following options:
   - **Item-total statistics** → This provides both the **correlated item-total correlation** and **Cronbach's Alpha if item deleted**.
   - Optionally, you can select **Scale statistics** for overall Cronbach's Alpha.
3. Click **Continue** to return to the main dialog box.

Step 5: Run the Analysis

1. Click **OK** to run the reliability analysis.
2. SPSS will generate several output tables.

Step 6: Interpret Correlated Item-Total Correlation

1. Look at the **"Corrected Item-Total Correlation"** column.
2. Interpretation:
   - Values **≥ 0.3** → The item is consistent with the overall scale and contributes positively.
   - Values **< 0.3** → The item may be inconsistent or not measuring the same construct.

Note: Low correlations indicate items that may need revision or removal.

Step 7: Interpret "Cronbach's Alpha if Item Deleted"

1. Look at the **"Cronbach's Alpha if Item Deleted"** column.
2. Interpretation:
   - If removing an item **increases the overall alpha**, that item may be weakening the scale and could be considered for deletion.
   - If removing an item **does not increase alpha** (or decreases it), the item is contributing positively, and it is best to retain it.

Note: This helps refine the scale while maintaining high reliability.

Step 8: Make Decisions

1. Review both tables to decide if any items should be deleted or revised.
2. Retain items that have **high item-total correlation** and **do not increase alpha when deleted**.
3. Make necessary changes and **re-run the reliability analysis** to confirm improvements.

Step 9: Document Findings

1. Report:
   - Overall Cronbach's Alpha value.
   - Range of item-total correlations.
   - Items that were retained or deleted based on analysis.

Example Summary Statement:

"The corrected item-total correlations ranged from 0.42 to 0.78, all above the 0.3 threshold, indicating strong consistency. Cronbach's Alpha if item deleted analysis showed that removing any item would not increase the overall alpha of 0.915, so all items were retained."

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
p1	79.32	128.717	.526	.911
p2	79.55	127.604	.541	.911
p3	79.80	127.911	.489	.912
p4	79.70	126.154	.561	.911
p5	79.82	124.242	.653	.908
p6	79.88	124.296	.621	.909
p7	79.32	127.197	.592	.910
p8	79.48	127.201	.588	.910
p9	79.37	128.648	.509	.912
p10	79.48	126.077	.615	.909
p11	79.42	125.346	.610	.909
p12	79.25	124.597	.647	.909
p13	79.47	124.878	.589	.910
p14	79.95	125.219	.491	.913
p15	79.63	126.532	.593	.910
p16	79.51	128.632	.491	.912
p17	79.81	129.365	.422	.914
p18	79.66	126.834	.564	.910
p19	79.37	125.426	.629	.909
p20	79.27	127.141	.617	.910
p21	79.42	131.546	.345	.915

7.19 Validity

One of the most important considerations when evaluating research methodologies in psychology, education, and other social sciences is a concept known as validity. It is the extent a test, measuring instrument, or other evaluative tools truly measure the indicator they intend to measure. Any outcomes or findings of empirical work make no sense when validity is missing. This is because no phenomenon or construct is being accurately captured. In other words, validity is the closure of any gaps of inaccuracy. It is the measure of confidence in the claim derived, hence the research tools and findings ought to be meticulously and rationally analyzed and used. Claims about the reliability describe the extent to which results are consistent across repeated measurements. Claims about the validity describe results as being correct and meaningful, hence having value.

The most fundamental aspect of validity is being able to make appropriate inferences. It is entirely possible to have a scale, test, or questionnaire deemed as having construct validity, and deemed as measuring the particular reliability in question, but the mere presence of validity is condition enough for the utilization to be limited. A scale in a bathroom that seems really cheap and nasty, but is just as accurate as clinically grade scales and shows French weights as 55kgs which the person seems to weigh, is weightless. It is reliable but not valid. It is for this reason that validity is more important than reliability.

Validity is often called construct validity because it deals with how well a measure or tool represents a theoretical construct it attempts to assess. Constructs include such abstract phenomena as intelligence, anxiety, motivation, customer satisfaction, or organizational commitment, and as such, they cannot be measured directly. They must therefore be represented by indicator items or questions in a scale. Higher the degree of connection between the indicators and the construct, higher the validity.

There are **two broad ways of assessing validity**:

a) **Theoretical Assessment** – This examines how well the theoretical construct is represented in its operational definition (the measurement tool). Researchers ask: *Does the instrument logically represent the concept?*

b) **Empirical Assessment** – This evaluates validity based on observed evidence and data. Researchers check how the measurement correlates with other external variables or criteria.

Both forms are necessary because theory provides the foundation for measurement, while empirical testing provides evidence of accuracy.

7.20 Types of Validity

Face Validity

Face Validity in measurement theory is very fundamental and quite superficial. It is concerned with whether the instrument or the test 'appears' to measure what the claims it is measuring. It is usually judged according to the subjective impressions of experts, or other respondents. Take for instance, a questionnaire intended to measure "job satisfaction." If this questionnaire includes 'salary, work environment and promotions' to be discussed in the items to be completed, it will be said to possess face validity because these items "seem to" encompass the concept of satisfaction.

Nevertheless, face validity alone is inadequate—it does not in any way ascertain the level accuracy of the measurement. It only establishes the fact that the instrument "looks okay" at a cursory examination.

Key Points:

- Based on subjective judgment.
- Concerned with appearance rather than actual accuracy.
- Provides initial credibility to respondents.
- Example: An anxiety scale asking about nervousness, restlessness, and worry.

Content Validity

More strict form of validity is Content validity which assesses how well the items of a test represent the whole content domain of the underlying construct being measured. As an illustration, a mathematics test that contains only algebra questions and ignores geometry and trigonometry will lack content validity, as the test is not representative of the entire syllabus.

In psychological measurement, Content validity is proved via consultations with domain specialists ensuring the items of the measurement tool adequately capture the domain of the construct.

Key Points:

- Ensures completeness of measurement.
- Assessed by expert judgment.
- Prevents underrepresentation of the construct.
- Example: A depression inventory must cover emotional, behavioral, and cognitive aspects of depression.

	KL1	KL2	KL3	PF1	PF2	PF3
KL1	1.00	0.83	0.79	0.23	0.21	0.19
KL2		1.00	0.75	0.11	0.20	0.03
KL3			1.00	0.03	-0.11	0.17
PF1				1.00	0.84	0.91
PF2					1.00	0.77
PF3						1.00

High correlations between items of the same construct (convergent validity)

Low correlations between items of different constructs (discriminant validity)

Validation convergent and discriminate validity is evident within the two constructs, Knowledge Learning (KL) and Performance Factor (PF), as both were each measured with 3 observed items. The matrix's diagonal values all equal 1.00 as each item is perfectly correlated to itself. Therefore, central concern is with the off-diagonal indices that showcase how items relate with each other. Stronger item relationships correlate with higher values, while weaker or no relationships correlate with lower values.

The KL correlational construct illustrates the triad KL1, KL2, and KL3, which correlate with each other and exhibit 3 values ranging between 0.75 and 0.83. These values are classified as high and positive.

As such, the KL items revolve around the same central idea, thus, all capture the same dimension of Knowledge Learning. The PF construct is no different as it to, demonstrates high inter-item correlations ranging from 0.77 to 0.91 which also illustrates that all PF items tap onto the construct of Performance Factor. These inter-construct values are consistent and strongly indicate convergent validity as well as internal consistency.

The measurement disriminant validity is proven by lack of high correlations bewteen different constructs. The correlation between PF and KL items is for the most part around 0 (ex. KL1–PF1 = 0.23, KL2–PF3 = 0.03, KL3–PF2 = –0.11). Such low or negative correlation indicate that the items pertaining to Knowledge Learning are certainly not tapping the same concept as the items pertaining to the Performance Factor. The constructs are validated as separate indicating that the measurement disriminant validity is present and the constructs are different both conceptually and statistically.

As highlighted in the matrix above, items in the different constructs are not related, which proves that the model is valid. The model is valid if the constructs are accurately represented and measured, which supports the convergent and discriminant validity of the measurement model.

This table is a **correlation matrix** that demonstrates **convergent validity** and **discriminant validity** of measurement items for two constructs:

- **KL (Knowledge Learning)** → measured by KL1, KL2, KL3
- **PF (Performance Factor)** → measured by PF1, PF2, PF3

Step 1: Reading the Matrix

- The diagonal values (all **1.00**) represent the correlation of each item with itself.
- The values above/below the diagonal show the correlations between different items.
- Higher correlation = stronger relationship, lower correlation = weaker relationship.

Step 2: Convergent Validity

Convergent validity means items measuring the **same construct** should correlate **highly** with each other.

- Look at **KL1, KL2, KL3**:
  - KL1–KL2 = 0.83
  - KL1–KL3 = 0.79
  - KL2–KL3 = 0.75

    → All are **strong positive correlations**, showing that KL items converge well, meaning they are all measuring the same construct (Knowledge Learning).
- Look at **PF1, PF2, PF3**:
  - PF1–PF2 = 0.84

  - PF1–PF3 = 0.91
  - PF2–PF3 = 0.77

    → Again, all are **high correlations**, showing good convergence for Performance Factor.

This confirms **convergent validity** (items within the same construct are strongly correlated).

Step 3: Discriminant Validity

Discriminant validity means items measuring **different constructs** should show **low correlations**.

- Compare **KL items with PF items**:
  - KL1–PF1 = 0.23
  - KL1–PF2 = 0.21
  - KL1–PF3 = 0.19
  - KL2–PF1 = 0.11
  - KL2–PF2 = 0.20
  - KL2–PF3 = 0.03
  - KL3–PF1 = 0.03
  - KL3–PF2 = -0.11
  - KL3–PF3 = 0.17

    → All are **very low (close to zero, some even negative)**.

This confirms **discriminant validity** (items from different constructs do not overlap).

Step 4: Interpretation

- **High within-construct correlations** (e.g., KL1–KL2, PF1–PF3) demonstrate that items measuring the same factor are consistent → **convergent validity**.
- **Low between-construct correlations** (e.g., KL1–PF1, KL2–PF2) demonstrate that items from different constructs are distinct → **discriminant validity**.

Step 5: Final Explanation

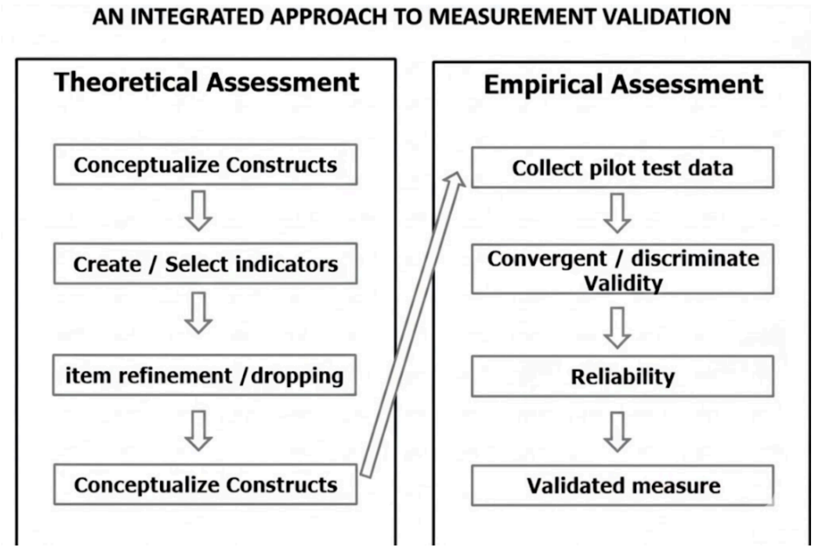
This correlation matrix demonstrates **validity evidence** for a measurement model:

- The **KL items** cluster together strongly.
- The **PF items** cluster together strongly.
- The **KL items and PF items** do not overlap significantly.

Thus, the measurement shows **both convergent validity and discriminant validity**, meaning the scale is measuring what it is supposed to (two distinct constructs).

7.21 An Integrated Approach to Measurement Validation

Understanding the history and philosophy of measuring the constructs of intelligence, motivation, satisfaction, or performance, it becomes clear that some of these constructs are highly abstract and cannot be measured almost directly and need to be reduced to 'operationalized' attributes which can be measured proxies. However operationalized attributes can be, indicators need to be valid and reliable to sustain strong measurement, hence there is a blend of theoretical assertions and empirical measures that need to be used which the diagram demonstrates as a strong measure validation.



7.21.1 Theoretical Assessment

Theoretical assessment is the starting point of measurement validation. It is primarily based on conceptual clarity, logic, and expert judgment rather than actual data. It involves the following steps:

- **Conceptualize Constructs**

  Researchers begin by defining the construct they want to measure. For example, if the construct is "job satisfaction," the researcher must clearly define what aspects of satisfaction are relevant—such as salary, work environment, and recognition. Conceptual clarity ensures that the measurement is grounded in theory.

- **Create/Select Indicators**

  Once the construct is defined, researchers create or select measurable items (questions, statements, or observed behaviors) that reflect the construct. For instance, survey items like "I am satisfied with my salary" or "I enjoy my work environment" would serve as indicators of job satisfaction.

- **Item Refinement/Dropping**

  Not all selected indicators may adequately represent the construct. Through expert evaluation and preliminary testing, researchers may refine or eliminate weak, redundant, or ambiguous items. This step ensures that only the most relevant and representative indicators are retained.

- **Re-Conceptualize Constructs**

  The refinement process may lead to rethinking the construct itself. For example, researchers may realize that job satisfaction has multiple dimensions (intrinsic satisfaction, extrinsic satisfaction) rather than being a single construct. Thus, theoretical assessment is an iterative process that sharpens both constructs and indicators.

7.21.2 Empirical Assessment

While theoretical assessment provides a foundation, empirical assessment validates the measurement tool using real-world data. This involves systematic testing with pilot studies or actual research samples. The main steps include:

- **Collect Pilot Test Data**

  Researchers administer the measurement instrument (survey, test, or observation) to a pilot sample. This allows them to gather initial data to check whether items perform as expected.

- **Convergent and Discriminant Validity**

  Using correlation analysis or advanced statistical models (such as factor analysis), researchers assess:

  - **Convergent validity** → whether items measuring the same construct are highly correlated.
  - **Discriminant validity** → whether items measuring different constructs show low correlation.

    Together, these ensure that the instrument is both precise and distinct.

- **Reliability Testing**

  Reliability refers to the consistency of measurement. Cronbach's alpha, test-retest reliability, or split-half reliability are commonly used. A reliable instrument yields consistent results across time and samples.

- **Validated Measure**

  After confirming validity and reliability, the measure is considered validated. This means it can be confidently used in further research, as it accurately captures the intended construct.

7.21.3 Integration of Theoretical and Empirical Approaches

The ellipse diagram highlights the need for integration across both perspectives. For the theory portion, the theoretical framework of an evaluation provides metrics, while the empirical framework provides the metrics. The process being investigated is, however, cyclical. For instance, emerging empirical findings may lead researchers to conceptualize constructs differently, thereby re-evaluating the theory. If this is the case, the researcher may revise the conceptualization of the construct preliminary.

The integration of both theory and empirical data becomes the basis of an approach to measurement validation. The quality of the theoretical assessment rests in the capacity of the assessor to accurately define the constructs and ascertain the relevance of the indicator to the theory and the empirical assessment to the indicator. They offer an assessment framework to strengthen the credibility, precision, and usefulness of the research instruments. The absence of integration implies the measurement is either conceptually weak or empirically flawed. This is not the case, as validation of the measurement is continuous, and the process is reliant on consistency of the conceptual framework and the variables of the case.

7.22 Testing for Normal Distributions

As part of a research, collecting data is not enough. The data must also be tested for, among other things, "a normal distribution." A normal distribution is often referred to as a bell curve. The normal distribution has an average and is symmetrical. A number of statistical operations rely on and assume the data is normally distributed for ANOVA, regression and t-tests. Because of the central importance of distribution to the reliability of these tests, normal distribution for the data has to be verified one way or the other. The normal distribution of the data has to be verified in. This verification process is termed the test for normality.

Deviation from a normal distribution is something the null hypothesis would like to postulate. If the distribution is normal, the parametric tests do not need to be considered.

If the distribution is not normal, then the data is not normally distributed, and one must go for non-parametric tests. Normality can be checked through a number of methods, for example through Kolmogorov-Smirnov test, histograms, Q-Q plots, P-P plots of the data among others.

7.22.1 Importance of Testing for Normality

- Many parametric statistical methods assume data is approximately normal.
- If the data is not normal, results of parametric tests may be misleading.
- Testing normality guides researchers in choosing between **parametric** or **non-parametric** methods.
- It helps ensure accuracy, reliability, and validity of statistical inference.

7.22.2 The Concept of Normal Distribution

A normal distribution curve is often regarded as a Gaussian distribution since it is a form of probability distributions that is not only continuous, but also centers itself around a particular mean.

The bell-shaped curve is a great case as it captures how most of the data points converge to a particular value, then apexes, and the how the probability decrease as one move away on either ends of the bell curve. A normal distribution is part of the binomial family of distributions, and in every case then mean, the median, and the mode is essentially the same. Also, in every case there is a surplus value, then majority, then another surplus, then a minor value of the population that falls away from the mean in every normal distribution, which is two thirds to one, then 95%, and finally 99.7% in the three standard deviations away from the mean in order to meet the 68-95-99.7 rule. This is the main reason, distribution is accepted in almost every field of study, and social as well as nature science; since practically every phenomenon spread in a particular given field which is height, marks, simplifications, and many more. This is an everyday case in hypothesis testing, confidence intervals, and control charts. The two most important aspects of the standard normal distribution is it mean, which is the average value, and then the stood dev value which is the dispersion value.

Before discussing tests, it is important to understand the characteristics of normal distribution:

- **Shape**: Symmetrical and bell-shaped.
- **Mean, Median, Mode**: All three are equal and located at the center.
- **Standard Deviation**: Controls the spread of the curve.
- **68-95-99.7 Rule**: About 68% of data falls within 1 standard deviation of the mean, 95% within 2, and 99.7% within 3.
- **Tails**: Extend infinitely without touching the axis.

7.22.3 Tests for Normality

SPSS provides both **statistical tests** and **graphical methods** for normality assessment.

Kolmogorov-Smirnov (K-S) Test

The Kolmogorov-Smirnov test compares the distribution of sample data with a normal distribution. It evaluates the maximum distance between the observed cumulative distribution and the expected cumulative distribution under normality.

- **Null Hypothesis (H0):** Data is normally distributed.
- **Alternative Hypothesis (H1):** Data is not normally distributed.
- **Decision Rule:**
  - If **p-value (Sig.) > 0.05** → Fail to reject H0 → Data is normally distributed.
  - If **p-value (Sig.) < 0.05** → Reject H0 → Data is not normally distributed.

Key Points:

- Useful for large samples.
- Sensitive to sample size (with large samples, even small deviations may be significant).
- SPSS path: **Analyze > Nonparametric Tests > Legacy Dialogs > 1-Sample K-S test**.

7.22.4 Shapiro-Wilk Test

Particularly with datasets with small sample sizes ($n < 50$), the Shapiro-Wilk test is extremely popular among practicing statisticians. It tests and establishes whether a sample comes from a normally distributed population. This is done by checking how the sample's distribution differs from that of a normal curve.

More formalized, the test is based on the following hypotheses:

- The null hypothesis (H_0) states that the data is normally distributed while the alternative hypothesis (H_1) states that the data is not normally distributed.

After running the test, the decision is made using the p-value.

- When $p > 0.05$, we do not reject H_0. Thus, the data can be presumed to be normally distributed.
- When $p < 0.05$, we reject H_0. Thus, the data is proven to be far from normal.

Like many of its competitors, the Shapiro-Wilk test is included in the statistical software SPSS, R and Python. It is widely used as the sample volume is low.

Key Points:

- More powerful than the K-S test for small samples.
- Frequently recommended for normality testing.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Age	.308	80	.000	.782	80	.000

a. Lilliefors Significance Correction

Graphical Methods

Graphical methods give a visual impression of normality. In SPSS, these can be generated under:

Analyze → Descriptive Statistics → Explore → Plots.

• Histogram with Normal Curve

A histogram shows the frequency distribution of the data, and when a normal curve is superimposed, it allows comparison with the ideal bell-shaped shape. If the bars roughly follow the curve, the data can be considered approximately normal.

• Q-Q Plot (Quantile-Quantile Plot)

This plot compares the observed data values with the expected values from a normal distribution. If the points fall closely along the diagonal reference line, normality is supported. Deviations from the line indicate skewness or heavy tails.

• P-P Plot (Probability-Probability Plot)

Similar to a Q-Q plot, but instead of raw values, it compares cumulative probabilities. A straight 45-degree line suggests normality, while curves or bends indicate departures from normality.

These visual tools complement numerical tests and help confirm assumptions before proceeding with parametric analysis.

- **Histogram with Normal Curve**
  - Displays the distribution of data compared to a normal curve.
  - If data approximates a bell-shaped curve, normality is supported.
- **Q-Q Plot (Quantile-Quantile Plot)**
  - Plots observed values against expected values from a normal distribution.
  - If points lie close to the diagonal line, data is normally distributed.
- **P-P Plot (Probability-Probability Plot)**
  - Similar to Q-Q plot but plots cumulative probabilities.

A straight 45-degree line indicates normality.

Key Points:

- Graphical methods provide intuitive insights.
- They are subjective but useful as supporting evidence alongside statistical tests.

7.23 Descriptive Statistics Approach

Using descriptive statistics to assess skewness and kurtosis values is one way to quickly gauge whether a given set follows a normal distribution pattern. The skewness is a measure of the symmetry of the distribution. In particular, a normal distribution is one which is centered around the mean, and therefore has a skewness value of 0. A zero value of kurtosis means that the distribution is mesokurtic, which is normal, and therefore contains a value of 3 which is the normal level of peakedness. Software such as SPSS has the capability to calculate these figures. General guidance suggests if the value of skewness and kurtosis is within 2 or -2, the data is normal and acceptable. This approach is not the substitute for statistical values, however, for larger sample sizes is a good reasonable hypothesis.

7.23.1 SPSS Procedure for Testing Normality

1. **Using Explore Command**
    - Go to **Analyze** → **Descriptive Statistics** → **Explore.**
    - Select the variable.
    - Under *Plots*, check "Normality plots with tests."
    - Output will include:
        - Histogram with normal curve.
        - Normal Q-Q plot.
        - Shapiro-Wilk and Kolmogorov-Smirnov test results.

2. **Using 1-Sample K-S Test**
   - Go to **Analyze** → **Nonparametric Tests** → **Legacy Dialogs** → **1-Sample K-S.**
   - Select the variable.
   - Choose “Normal” as the test distribution.
   - Check output for Sig. (p-value).

7.24.1 Results of a Normality Test

Normality tests are performed to check if the data fit a normal distribution which is a prerequisite for any statistical analysis. The result is interpreted based on the p-value that is determined through the test.

If $p > 0.05$: In this scenario, the null hypothesis is not rejected meaning the data does not markedly differ from a normal distribution. Simply put, the data is viewed as approximately normal and the data can be safely subjected to parametric analyses, t-tests, ANOVA, correlation, and regression. This implies that the data set is normal for these tests.

If $p < 0.05$: Now we can say that the null hypothesis is rejected. This implies that the data presented is in fact, not normal. This means the data are no longer and ‘almost’ normal. This is the point at which people feel that a transformation, be it a log, square root or a Box-Cox transformation, alteration is warranted. If the errant data point is not aligned to be ‘normal’ or does not yield a satisfactory result, then parametric tests are no longer applicable and should not be employed, these tests being the Mann-Whitney U, Kruskal-Wallis and Spearman correlation, which are non-parametric.

- **If $p > 0.05$:**
  - Fail to reject null hypothesis.
  - Data does not significantly deviate from normal distribution.
  - Assumption of normality is satisfied.
- **If $p < 0.05$:**
  - Reject null hypothesis.
  - Data is significantly different from normal distribution.
  - Consider transformation or use non-parametric tests.

7.24.2 Benefits of Normality Tests

Normality tests are significant in numerous ways while analyzing research data. To begin with, it provides proof that goes beyond sight, judgment, or even someone's opinion. The importance of quantitative evidence in any analysis cannot be overstated. This is because mere sight, if relied upon, may lead to other conclusions. Second, the results assist in determining the appropriate statistical technique, thus ensuring that the conclusions made from the analysis are meaningful. For example, if the data is normal, the more powerful parametric tests may be employed. Third, normality tests prevent analysis from being misinterpreted or results from being rendered useless. This is because the failure to test something appropriately can lead to wrong decisions or assumptions in research.

7.24.3 Limitations of Normality Tests

As useful as they might be, researchers should keep in mind some limitations of normality tests. Sample size sensitivity is one of the most well known problems. With extremely large samples, even the most trivial of deviations from normality would be regarded as significant, causing tests to assert that the data is non normal, when, in reality, it is almost normal. With extremely small samples, it is the case that tests would fail to detect deviations from normality, due to not having enough statistical power. The next limitation is the subjectivity inherent to graphical methods. Instruments such as histograms or QQQ plots have the tendency to be differently analyzed by different analysts. Lastly, normality is at times, overemphasized. In practical research, data is hardly ever perfectly normal, and plenty of parametric tests still hold water, even when the data is a tad bit non-normal. Thus, not every instance of the violation of normality goes hand in hand with the need to correct or change the statistical test.

Limitations

- **Sample Size Sensitivity:**
    - In large samples, even trivial deviations may appear significant.
    - In small samples, deviations may go undetected.
- **Graphical Subjectivity:**
    - Visual methods are not always conclusive.
- **Overemphasis Risk:**
    - Real-world data often deviates slightly from perfect normality. Parametric tests are robust against mild deviations, so not all violations are critical.

Although informative, normality tests should not be considered in isolation. It is best to take a combination of statistical tests and graphical plots, as well as sample size and research context, before a concluding decision is made. For medium to large sample sizes, slight departures from normality are often tolerated, since most parametric tests are robust and still yield valid results.

Practical Recommendations

- Always use a combination of **statistical tests** and **graphs**.
- For small samples, prefer **Shapiro-Wilk** test.
- For large samples, use both **K-S test** and graphical methods.
- If data is not normal:
    - Apply transformations (e.g., log, square root).
    - Use non-parametric tests (e.g., Mann-Whitney U, Kruskal-Wallis).

Key Points in Bullets

- **Normal Distribution:** Symmetrical, bell-shaped, mean = median = mode.
- **Purpose of Normality Test:** To check whether data meets assumptions for parametric tests.
- **SPSS Methods:**
    - Analyze ⟶ Descriptive Statistics ⟶ Explore.
    - Analyze ⟶ Nonparametric Tests ⟶ Legacy Dialogs ⟶ 1-Sample K-S Test.
- **Kolmogorov-Smirnov Test:**
    - Null hypothesis = data is normal.
    - If Sig. > 0.05 ⟶ data normal.
- **Shapiro-Wilk Test:** Recommended for small samples.
- **Graphical Methods:** Histogram, Q-Q Plot, P-P Plot.
- **Descriptive Statistics:** Skewness and kurtosis check.
- **Interpretation Rule:**
    - $p > 0.05$ ⟶ Normality assumed.
    - $p < 0.05$ ⟶ Non-normal distribution.
- **Limitations:** Sensitive to sample size, subjective graphs.
- **Recommendation:** Use both statistical and graphical evidence before concluding.

During research analysis, performing a normality test is vital, especially for normality checking, as kolmogorov-smirnov, Shapiro Wilk, skew, and kurt measures, etc. SPSS provides graphical methods for normality assessment. Like every other tool, provides more reliable conclusions based on a combination of normality test results and graphical evidence for normal distribution. Normality violations can be mitigated using transformations or robust non parametric tests. Endurance on normality assessments can confidence researchers statistical results.

7.25.1 Pared Sample T - Test

As one has titled it, the 'dependent t-test' works on the basic premise of the 'paired-samples t-test,', gauging the difference in the means of two related observations. The test, unlike its independent sample counterpart, only applies in instances where subjects are measured multiple times in varying conditions, or more logically associated one way or another.

Examples:

- Measuring students' scores before and after a training program.
- Comparing blood pressure of patients before and after taking a drug.
- Comparing employees' stress levels at the beginning and end of a project.
- In your SPSS example: **Age** and **Income** are being compared for the same individuals (though in practice, pairing age and income is unusual unless logically justified).

Rationale

The paired-samples t-test, or dependent t-test, is used when the same participants are assessed two or more times under different condition or at different times. As opposed to analyzing two distinct populations, the test focuses on two datums of the same persons, for instance, pre-and post-test results, treatment outcome scores, or measurements under two different conditions. Individual differences, for instance, differences in intelligence, skill, or personality, are self-controlled because each person is analyzed with himself/herself. This minimizes the random error, thus making the test more powerful statistically than an independent-samples t-test with the same sample size.

Hypotheses

- **Null Hypothesis (H_0):** The mean difference between paired observations is zero (no significant change/effect).
- **Alternative Hypothesis (H_1):** The mean difference between paired observations is not zero (there is a significant change/effect).

Mathematically:

- $H_0 : \mu_d = 0$
- $H_1 : \mu_d \neq 0$

Where μ_d = mean difference between pairs.

Assumptions

The paired-samples t test is based on a set of criteria which must be fulfilled in order to get a valid and accurate answer in the test.

1. Continuous dependent variable

The outcome variable whose outcome is dependent is to be calculated on the value of an interval or ratio scale. It must be able to be measured in a significant value which is set in the same interval units in which it is being measured (exam score, blood pressure, reaction time, etc.). It makes the test inappropriate to work with ordinal or categorical data.

2. Related samples (paired observations)

The criteria for both the sets of scores which are to be analyzed must be considered. This happens in situations when a subject is observed two times during a matching period, or the subject is observed along with a subject with the same criteria (age, twins, etc.). In order to ensure unbiased data, each pair has to be independent from the other pairs.

3. Normality of difference scores

The test in question says that the paired values (not the raw data) which are being calculated for the purposes of the difference must be normally arranged. The assumption for small samples ($n < 30$) is very important and must be proven by statistical calculations or under visual observation (such as a diagram or histogram). For bigger samples, there is some reprieve from following the rules of normality and still obtaining the desired result.

4. Absence of Extreme Outliers

Outliers in the difference scores can have a meaningful impact on the calculation of the mean and standard deviation and can consequently lead to incorrect conclusions. It is in the best interest of the researcher to identify and deal with outliers (by trimming, winsorizing, or justifying) prior to undertaking the test.

Assuming the above outliers are not present, and the remaining assumptions are fulfilled, the paired-samples t-test offers the most efficient and accurate approach in ascertaining mean differences for the same subjects under two conditions or time periods.

For valid results, the paired-samples t-test requires several assumptions:

1. **Continuous dependent variable**

- The test requires data measured on an interval or ratio scale (e.g., test scores, income, age, weight).

2. **Related samples (paired observations)**
   - Observations must be paired in a meaningful way (e.g., same individual measured twice, or matched pairs).
   - Each pair must be independent of other pairs.
3. **Normality of differences**
   - The distribution of the differences between pairs should be approximately normal.
   - For large samples (n > 30), the test is robust to normality violations due to the Central Limit Theorem.
4. **Absence of extreme outliers**
   - Outliers in difference scores can distort the results.
   - Outlier detection should be performed before running the test.

Formula

$$t = \frac{\bar{d}}{(s_d/\sqrt{n})}$$

Where:

- $\bar{d}$ = mean of the differences between pairs
- s_d = standard deviation of the differences
- n = number of pairs

The t-value is compared against a critical t-value from the t-distribution table at df=n−1df = n-1df=n−1.

Steps to Conduct in SPSS

1. **Go to:** Analyze → Compare Means → Paired-Samples T-Test.
2. **Select pairs of variables** (e.g., "Score before training" and "Score after training").
3. **Click OK** to generate results.
4. SPSS will produce three key tables:
   - **Paired Samples Statistics** (descriptive stats of each variable)
   - **Paired Samples Correlations** (correlation between the two variables)
   - **Paired Samples Test** (main test result, with t-statistic, df, and significance level).

7.25.2 Interpreting SPSS Output (Step-by-Step)

Looking at your attached SPSS screenshot:

(a) Paired Variables Selection

- **Pair 1:** Age vs. Income

  (Here, you are asking SPSS to compare the mean of Age and the mean of Income for the same individuals.)

Note: Normally, paired t-tests are used for **same variable measured twice** (e.g., pre-test vs post-test scores). Age and Income are two *different constructs*, so strictly speaking this may not be the correct application. But for practice, I'll interpret it as if you want to check whether mean Age is significantly different from mean Income within the same sample.

(b) Expected Output Tables

1. **Paired Samples Statistics**
   - Shows mean, standard deviation, and N for each variable.
   - Example: Age (Mean = 35.6, SD = 12.1), Income (Mean = 42,000, SD = 18,000).
2. **Paired Samples Correlation**
   - Shows correlation between Age and Income.
   - Indicates whether higher Age is associated with higher/lower Income.
   - Example: $r = 0.45$, $p < 0.01$ (moderate positive correlation).
3. **Paired Samples Test**
   - This is the main table of interest.
   - Provides:
     - Mean difference (Age – Income)
     - Standard deviation of differences
     - 95% Confidence Interval (CI) of difference
     - t-value, df, and significance (p-value)

(c) Decision Rule

- If **p-value ≤ 0.05** → Reject null hypothesis → Significant difference between the two paired means.
- If **p-value > 0.05** → Fail to reject null hypothesis → No significant difference.

Advantages of Paired T-Test

- Controls for individual variability (same subject measured twice).
- Requires fewer participants to achieve statistical power.
- Straightforward to implement and interpret.

Limitations

- Sensitive to outliers.
- Requires normality of differences.
- Only applicable for two related groups (cannot handle more than two conditions).
- If variables are unrelated (like Age vs Income), results may not be meaningful.

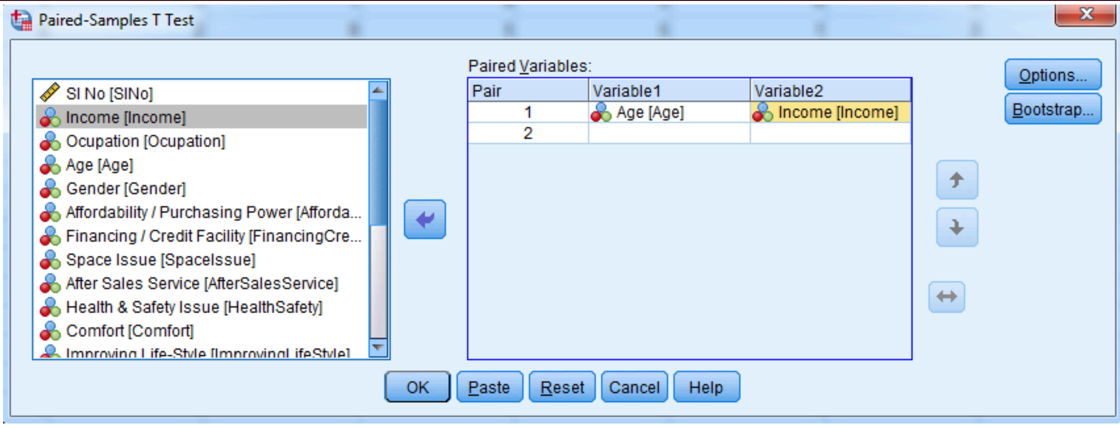
Alternatives

- If normality assumption is violated → Use **Wilcoxon Signed Rank Test** (non-parametric equivalent).
- If more than two related groups → Use **Repeated Measures ANOVA**.

Key Points

- A **Paired-Samples T-Test** compares means from two related conditions.
- Hypotheses:
    - H_0: No difference in means.
    - H_1: Significant difference in means.
- Assumptions:
    - Continuous dependent variable.
    - Paired/related observations.
    - Normally distributed differences.
    - No extreme outliers.
- SPSS Output includes:
    - **Descriptive stats** (means, SDs)
    - **Correlation** (between variables)
    - **Paired-Samples Test** (t, df, p-value)
- **Interpretation rule:**
    - $p \leq 0.05$ → Significant difference.
    - $p > 0.05$ → No significant difference.
- Practical Applications: pre-test vs post-test, before-after treatments, matched pairs.

7.25.3 The Conceptual Error



- You have selected **Age** (Variable1) and **Income** (Variable2) as a pair.
- SPSS will test whether the **average age** and **average income** are statistically different across the same sample of respondents.

But this is conceptually problematic, because Age and Income are not the same construct and are measured on different scales. Normally, you should compare Age (before vs after some intervention) OR Income (before vs after job change).

So the test will **mathematically run**, but interpretation should be cautious:

- If $p \leq 0.05$ → There is a statistically significant difference between Age and Income means.
- If $p > 0.05$ → No significant difference between Age and Income means.

However, the meaningful conclusion is questionable because Age ≠ Income in concept.

The capability of comparing two means (before-after, matched pairs) is the focus of The Paired-Samples T-Test. It requires continuous data, normally distributed and difference data without extreme values. It is possible to receive three outcomes, the descriptive statistics, correlation, and the t-test itself. These three outcomes are dependent on the t-test output values (p).

In the SPSS screenshot, you matched Age to Income. In this case, Age and Income have different concepts which is problematic, and so SPSS is measuring the means of the two of them which is sending the wrong message. In more relevant examples, paired variables selected logically (e.g., two times the same test was taken).

7.25.4 Independent Samples T-Test

An Independent Samples T-Test is also known as two-sample t-test or unpaired t-test. It refers to a technique used to test the difference between two means from two different sample groups. The two different groups under consideration are independent of each other. The aim is to ascertain the possibility of a significant difference existing between means with the possibility of being a chance occurrence.

In the paired-sample t-test, there are no two distinct groups which have no overlap. The independent samples t-test is a different case.

7.25.5 Purpose of the Test

This test is used to determine whether there is a statistically significant difference between the average values (means) of two groups. It helps answer questions such as whether males and females score differently on exams, whether students from rural and urban backgrounds have different income levels, or whether a control group and a treatment group show different levels of improvement after an intervention. It can also compare workplace factors, such as whether employees in Department X experience higher stress than those in Department Y. By analyzing the mean difference and considering sample size and variation, the test helps researchers decide if the observed differences are real or due to chance.

The test answers questions like:

- Do males and females differ in average exam scores?
- Do rural and urban students have different income levels?
- Does Group A (control) differ from Group B (treatment) in average improvement?
- Is there a difference in stress levels between employees in Department X and Department Y?

Hypotheses

The independent samples t-test evaluates two competing hypotheses:

- **Null Hypothesis (H_0):**

  There is no difference in the mean scores of the two groups.

  $H_0 : \mu_1 = \mu_2$

- **Alternative Hypothesis (H_1):**

  There is a significant difference in the mean scores of the two groups.

  $H_1 : \mu_1 \neq \mu_2$ (two-tailed test)

- In one-tailed tests, the hypothesis may specify direction:

$H_1 : \mu_1 > \mu_2$ or $H_1 : \mu_1 < \mu_2$.

7.25.6 Assumptions

The requirements guiding the use of independent samples t-tests say that there are two groups of participants whose means need to be compared. One of the requirements of the t-test procedure is that the groups that are being compared be independent of each other. People need to be assigned to only one of the groups: they are either males or females; they are either urban or rural, or they are either in the treatment or the control groups. There are no participants that are within both groups. If there are, then the assumption of independence is violated and therefore the comparisions being made would not be accurate.

The other requirement of the independent samples t-test is that the variable in question be continuous. This means that the variable has been measured in an interval or ratio scale, for instance, score on a test, height, weight, age, reaction time, and even income. As a matter of fact, it is beyond any reasonable doubt that gender or religion can be tested as dependent and independent variables using the t-test procedure because the test aims to establish distinctions between means and not frequencies.

The busiest assumption is normality, which states that each group's scores should be normally distributed with respect to the dependent variable. This implies that scores should have a normal distribution and, therefore, a bell-shaped curve. If the sample size is big enough (30 or more per group is the most common cutoff), the situation is said to be robust, or the more relaxed the assumption, the more reliable the conclusions. In other words, the sample is robust enough to ensure that the conclusions, even if the distribution is not perfectly normal, can be relied upon due to the Central Limit Theorem. This is the case, however, with the smaller samples which violate normality, and such deviations tend to lower the reliability or validity of the sample.

The fourth assumption, homogeneity of variances, in essence, states that the amount of variability or spread in the scores should be similar across each of the two groups. In practice, the way however this is done is with Levene's Test and other, much more straightforward, statistical packages, such as SPSS. When the variances are not equal, the "standard" t-test is assumed to be not very reliable and therefore a more reliable adjustment (for example, Welch's t-test) is recommended.

The test concludes with a conditional assumption that there are no extreme outliers in the data. Outliers are values that are either extremely high or low and do not reflect the average performance or the behavior of the outliers in the group. Outliers are known to strongly affect the mean and the standard deviation of the group which can lead to erroneous interpretations and assumptions. Before running the test, outliers should be detected and taken care of - whether through visual methods such as boxplots and histograms or other statistical methods.

For valid results, the independent samples t-test makes several assumptions:

1. **Independent Groups**
   - Each participant belongs to only one group (e.g., male or female, rural or urban).
   - No participant is included in both groups.
2. **Continuous Dependent Variable**
   - The test variable (e.g., test score, age, income) should be measured on an **interval or ratio scale**.
3. **Normality**
   - The dependent variable should be approximately normally distributed within each group.
   - For large samples (n > 30 per group), the test is robust to normality violations.
4. **Homogeneity of Variances**
   - The variance (spread of scores) in each group should be roughly equal.
   - SPSS tests this with **Levene's Test for Equality of Variances**.
5. **No extreme outliers**
   - Outliers can skew group means and distort results.

Formula

The independent samples t-test statistic is calculated as:

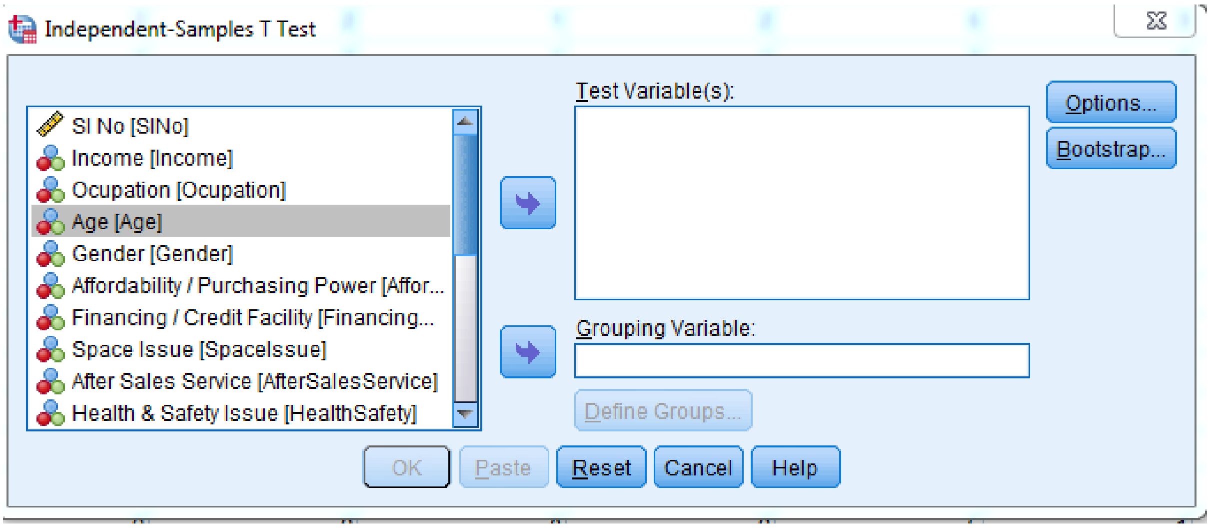
$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where:

- $\bar{X}_1, \bar{X}_2$ = sample means of the two groups
- s_1^2, s_2^2 = sample variances of the two groups
- n_1, n_2 = sample sizes

If group variances are assumed equal, a pooled variance estimate is used.

SPSS Procedure



:

1. **Go to:** Analyze → Compare Means → Independent-Samples T-Test.
2. **Select Test Variable(s):**
    - This is the continuous dependent variable (e.g., Exam Score, Age, Income).
3. **Select Grouping Variable:**
    - This is the categorical independent variable (e.g., Gender: Male/Female, Location: Rural/Urban).
4. **Define Groups:**
    - Tell SPSS which numeric codes represent which groups (e.g., 1 = Male, 2 = Female).
5. **Click OK** to run the test.

SPSS will produce several tables.

CHAPTER 8: SPSS Output Interpretation

The results of an independent samples t-test in SPSS are generally represented in two key tables: Group Statistics and Independent Samples Test. Each of these tables serves a specific purpose and presents certain findings essential for interpreting the results and for making a determination as to whether the two groups being compared are significantly different from each other.

8.1 Group Statistics

The Group Statistics table gives a descriptive summary about group members with respect to the test. It contains counts (N), averages, standard deviations (SD), and standard errors (SE) for every group. These values aid in visualising the basic structure and performance of the groups prior to carrying out advanced inferential techniques. Let's assume the output indicates a number of male participants to be 50, who have a score of 65 on an exam with a standard deviation of 10, and 50 female participants, who have a mean score of 70 with a standard deviation of 8 on an exam. Article descriptive statistics only tell so much, and with the t-test, it is a whole other level of understanding whether it's sufficient to Work with the t test.

8.2 Independent Samples Test

The second table, named Independent Samples Test, holds the results of the actual hypothesis test. This table is divided into two key sections:

1. Levene's Test for Equality of Variances

The first part of the table contains the results of Levene's test, which assesses whether the two groups being compared have unequal variances. The goal is to determine whether the assumption of homogeneity of variance holds. This conclusion is determined by the p-value (Sig.).

- When $p > 0.05$, the variances are considered equal, therefore we take the first row of the t-test results (labeled "Equal variances assumed).
- When $p \leq 0.05$, the variances are unequal thus the assumption is violated. Therefore, we take the second row ("Equal variances not assumed") which offers a modified t-test called Welch's t-test.

2. t-test for Equality of Means

The second part of the table provides the t-test results. It has a number of key parameters:

- t-value: the test statistic which is used to assess the means of the two groups.
- df: the number of degrees of freedom which is calculated based on whether or not the variances are equal.

- Sig. (2-tailed): the p-value that assesses the difference in means to determine if there is sufficient evidence to claim it is a statistically significant difference.
- Mean Difference: do not confuse a difference in means with a discrepancy in means. For instance, do not refer to the discrepancy in means as the discrepancy difference in means. Rather, the difference or the gap or the divisible value, is in forms of a difference, bridge, gap, or bridge value. Correspondingly difference or mean in any form is in forms of a gap.
- Confidence Interval of the Difference: This illustrates the gap or border or the line in the ground considering the border within which the difference in means is saved.

8.3 Decision Rule

The test is decided to be concluded with the analysis of the p value as follows:

- If p is below 0.05, then we are able to state that the means are statistically significant (H 0 is rejected). This means, the means, either of the two, are able to survive.
- If p is more than 0.05, we do not reject the null hypothesis, which states that the difference in means is not significant. Therefore, any difference in the means is a difference that can be termed as meaningless because it is because of luck or coincidence.

Example Interpretation

Suppose we tested **Exam Score differences between Male and Female students**:

- Group Statistics:
  - Male: Mean = 65, SD = 10, N = 50
  - Female: Mean = 70, SD = 8, N = 50
- Levene's Test: p = 0.28 (>0.05) → Assume equal variances.
- t-test:
  - t = -2.45, df = 98, p = 0.016 (<0.05).
  - Mean difference = -5 (Females scored 5 points higher).

Interpretation:

There is a statistically significant difference in exam scores between males and females, with females scoring significantly higher (p = 0.016).

8.4 Applications

The independent samples t-test 's primary function is to determine the statistical difference between the means of t wo unrelated population samples. The ease of use and reliability of the method has led to its adoption in many disciplines and professions.

- Education: The independent samples t-test is often employed by researchers to find the difference in performance of two groups of students who are taught using two different teaching strategies. For example, the average marks of students who are taught through face to face methods and students who are taught by the use of technology in an interactive way.
- Medicine and Healthcare: In clinical trials, the independent samples t test is often used to find the difference between the results of the test group and the control group. For example, the average recovery period and blood pressure of the patients who are treated with a new medication and the patients who received a placebo.
- Psychology and Social Sciences: Psychologists, in their line of work, use the t-test to make comparisons of different psychological traits and behaviors within various groups. The stress levels of the employed working population and the unemployed population is a classic example. The average anxiety scores among adults and adolescents is another example.
- Business and Management: In the field of marketing or consumer research, the independent samples t-test can be employed to find the difference in customer satisfaction, sales, or even the productivity of employees across two different branches, regions, or policies. Comparing the average ratings of customers who visit two retail stores is one such example.

- It is one of the common inferential tests in research because it fits any scenario where a comparison is to be done between two independent groups.

8.4.1 Advantages

Like in many research investigations, the independent samples t-test is the superior choice due to its many merits:

- Uncomplicated and interpretation friendly: Even untrained persons can understand the results produced by statistical packages such as SPSS, R, or excel due its simplistic nature.
- Type I error is controlled: This test is done statistically more efficient since the group being compared is only two, and the probability of a difference being concluded Type I error is kept in reasonable bounds compared to carrying out several tests.
- Also provides practical reasoning with statistical evidence: The test provides more than one interpretation by also computing effect size (for example, Cohen's d), which shows how significant the difference between the two groups is in the real world along with p value, which is the value that shows if the difference is significant or not.

8.4.2 Limitations

Although the independent samples t-test is useful on its own, it is important to keep the following limitations in mind:

- One-dimensional as it pertains to variety of groups developed: Separated independent samples are the only one defined as being of the same rank as the test being administered. In cases in which groups exceed two in number (i.e. three teaching methods, four treatment conditions, etc.) it is neccessary to employ ANOVA (Analysis of Variance).
- Sensitive assumptions with small samples: When the sample sizes of groups are smaller, the test will more likely lead to misleading or inaccurate results having to do with the normal distribution of data as well as equal variances between groups. Contrary to popular belief, alternative tests which include Welch's t-test or even non-parametric tests, are best used when baseless assumptions are made.
- Not appropriate for paired or repeated measurements: Independent samples t-test can not be used in case the same participants are assessed more than once (e.g pre-test and post-test) or the samples being assessed are somehow related. In such cases, a paired t-test or repeated measures ANOVA is the right choice.

8.5.1 Non-Parametric Alternative

Submitted researches for an independent samples t-test which haven't met set requirements, particularly normality and the independent range of variances, should perform an independent test not based on the assumption of population distributions. The alternative which is utilized most frequently is the Mann-Whitney U Test. It is the Wilcoxon Rank Sum test. In contrast to the t-test, it does not assume the existence of a normal distribution of the data. It does not calculate the means of different groups. It assesses the median ranks of two independent samples. This is particularly advantageous for the researcher when the dependent variable is skewed, outlier-filled, ordinal, or not continuous. This test, in comparison to the t-test and other statistical methods, is statistically less powerful but still a satisfactory answer to the parametric assumptions being greatly violated.

Key Points

- The **Independent Samples T-Test** compares means of **two independent groups**.
- Groups must be **mutually exclusive** (no overlap).
- **Hypotheses:**
    - H_0: $\mu_1 = \mu_2$ (no difference)
    - H_1: $\mu_1 \neq \mu_2$ (difference exists)
- **Assumptions:**
    - Continuous dependent variable
    - Independence of groups
    - Normal distribution within each group

- Equal variances (tested via Levene's Test)
- No extreme outliers

- **SPSS Steps:** Select test variable → select grouping variable → define groups → run test.
- **SPSS Output:**
  - Group statistics (means, SDs)
  - Levene's test (check variance equality)
  - t-test results (t, df, p-value, mean difference, CI).
- **Interpretation:**
  - If $p \leq 0.05$ → significant difference between group means.
  - If $p > 0.05$ → no significant difference.

8.5.2 Usage and Restrictions

The independent samples t-test cuts across several fields like education, medicine, psychology, and business where researchers compare two groups, for instance, males and females, experimental and control, online and classroom learners, etc. However, it only applies to two groups and does not accommodate repeated measures or paired observations. For three or more groups, ANOVA is used, and when its assumptions are violated, the Mann–Whitney U test provides the appropriate alternative.

8.6 One-Sample T-Test

The **One-Sample T-Test** is a statistical test used to determine whether the **mean of a single sample** is significantly different from a known or hypothesized population mean. It is one of the simplest forms of hypothesis testing and is widely used when you want to compare your sample data against a benchmark value.

Example:

- A company claims the average life of their LED bulbs is **2 years**. A researcher collects a sample of bulbs and wants to check if the average life in the sample is **significantly different** from 2 years.
- A school principal claims the average IQ of students is **100**. A researcher tests a sample of students to see if their average IQ differs from 100.

8.6.1 Purpose of the One-Sample t-Test

One-Sample T-Test Because there is only one sample observation, a reason for performing a one-sample t-test is to ascertain whether a sample mean is appreciably different from a population mean. Any sample t-test must have a null hypothesis which is that sample mean is equal to a given population mean. Any sample t-test can also have a t=mx, which is a t equal to a mean. Other values for t can equal to such a sample distribution such that for some given set of sample population, distribution, population, its square root divided by number of sample observations, sample is taken from the df distribution. This population t-test is set to differ a population mean from the sample mean, thus as these values coincide, approaches some limit, the hypothesis set here is set to consider some predetermined fixed level certainty of hypothesis set.

The claims must be fulfilled to consider which sample mean is expected to be equal to population mean. A botanist studies whether the mean height of a new plant variety differs with the expected height of 20 cm, while an educator wishes to find out whether the time long university students study is different from the recommended time of 15 hours. In all the above instances, a one-sample t-test is used to assess if a population sample mean derived is appreciably different, more, or less from the value of null hypothesis set.

In essence, the one-sample t-test helps answer questions like:

- "Is this group performing as expected?"
- "Does the sample reflect the known population value, or has it changed over time?"
- "Is there evidence to support or reject a stated claim about the average?"

The one-sample t-test ascertains the null hypothesis, which states that there is no difference between the sample mean and the test value, and assesses whether or not a given difference is a product of sampling error.

Hypotheses

Like other t-tests, the One-Sample T-Test has two competing hypotheses:

- **Null Hypothesis (H_0):**

  The sample mean is equal to the population/test mean.

  $$H_0 : \mu = \mu_0$$

- **Alternative Hypothesis (H_1):**

  The sample mean is not equal to the population/test mean.

  $H_1 : \mu \neq \mu_0$ (two-tailed test)

- For directional hypotheses (one-tailed tests):
  - $H_1 : \mu > \mu_0$ → mean is greater than test value.
  - $H_1 : \mu < \mu_0$ → mean is less than test value.

Assumptions

For the one-sample t-test to be valid:

1. **Scale of Measurement**
   - The dependent variable should be measured on a continuous scale (interval or ratio).
2. **Random Sampling**
   - The data should be obtained from a random sample of the population.
3. **Normality**
   - The dependent variable should be approximately normally distributed.
   - For large samples ($n > 30$), the test is robust to normality violations.
4. **No extreme outliers**
   - Outliers can distort the mean and test results.

Formula

The t statistic is computed as:

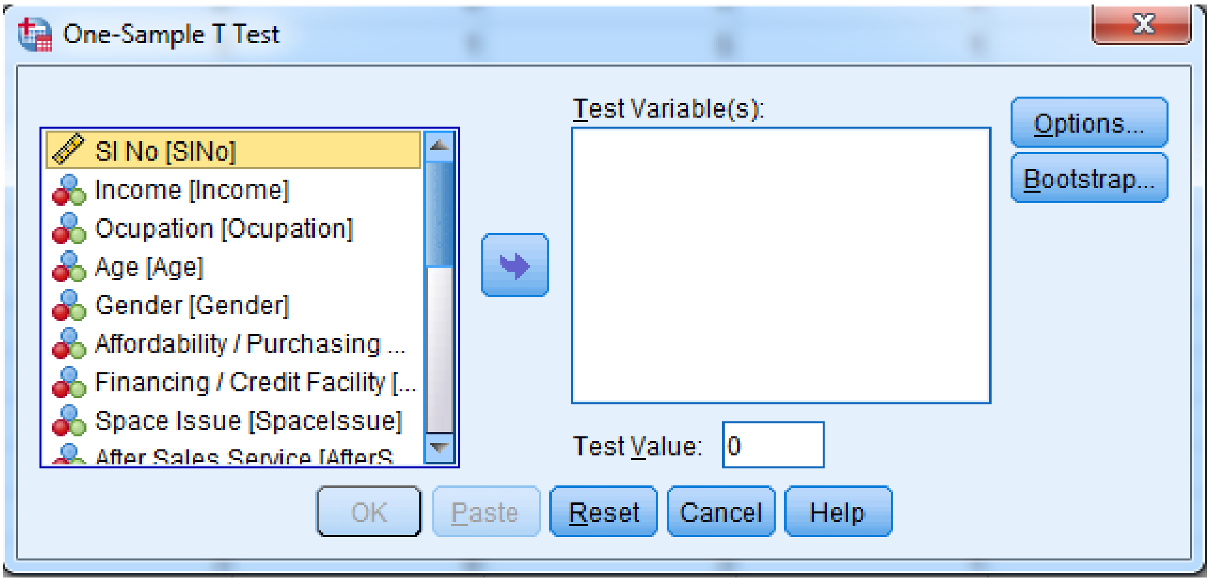
$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

Where:

- $\bar{X}$ = sample mean
- μ_0 = test value (population or hypothesized mean)
- s = sample standard deviation
- n = sample size

This measures how many standard errors the sample mean is away from the test value.

SPSS Procedure



1. **Go to:** Analyze → Compare Means → One-Sample T-Test.
2. **Move the test variable(s)** (e.g., Age, Income, Score) into the *Test Variable(s)* box.
3. **Enter the Test Value** (benchmark you are comparing against, e.g., 50, 100, 2 years, etc.) in the field "Test Value".
4. **Click OK.**

8.6.2 Interpreting Outputs in SPSS

When performing a one-sample t-test in SPSS, one gets an output that comes in two major blocks, the One-Sample Statistics and One-Sample Test tables. Each of them contains vital data that allows the user to make a decision on whether the sample mean was sorely wrong or, in any case, different from the tested value (population or hypothesize mean).

(a) One-Sample Statistics

One-Sample Statistics are the first tables created and they provide first level summaries of the sample. In this table we are particularly interested in the sample size (N) and sample mean, sample and population standard deviation (SD) as well as the standard error of the mean (SE). These are very important in understanding the context of the data being worked on before any form of hypothesis testing.

Traditionally, the output will jest there are 40 light bulb sample, and that the mean lifespan is 1.85 years and the standard deviation is 0.20. This tells us that there is a fraction of the bulbs that do not last 1.85 years and there is some form of deviation between the bulbs. In a scenario where the sample mean is most likely to be the population mean, standard error comes in handy. Lower the SE, the better the estimate.

(b) One Sample Test

The second table, One Sample Test, includes One Sample Test table and the results of the One Sample statistical analysis which encompasses multiple essential elements such as spearheaded by Test Value (μ_0): which states the case of which the sample mean is being compared versus the sample mean and t value the case of which the calculated value is considered as the sample mean as per some standard error units, or the underlying mean is considered as the sample mean.

Having calculated the degrees of freedom df, which states the case as n-1, to which n constitutes the case for the sample size, the sig. (2) column case is the domain under the hypothesis tends to be proved.

- Test Value (μ_0): the reference value against which the sample mean is being compared (e.g., industry standard, theoretical value, historical mean).
- t-value: the calculated test statistic showing how far the sample mean is from the test value in standard error units.
- Degrees of freedom (df): calculated as n – 1, where n is the sample size.
- Sig. (2-tailed): the p-value, which tells whether the difference between the sample mean and the test value is statistically significant.
- Mean Difference: the actual numerical difference between the sample mean and the test value.
- 95% Confidence Interval (CI): shows the range within which the true difference between the sample mean and the test value is likely to fall.

8.6.3 Decision Rule

To interpret the results, focus on the p-value:

- If $p \leq 0.05$, you reject the null hypothesis (H_0). This means the sample mean is significantly different from the test value.
- If $p > 0.05$, you fail to reject the null hypothesis. This means the observed difference is not statistically significant and may have occurred by chance.

Example Interpretation

Scenario: A manufacturer claims LED bulbs last **2 years**. A researcher samples 40 bulbs and finds:

- Sample mean = 1.85 years
- SD = 0.20
- Test value = 2

SPSS Output might show:

- t = -4.74, df = 39, p < 0.001
- Mean difference = -0.15

- 95% CI: -0.21 to -0.09

Interpretation:

The sample mean life (1.85 years) is significantly lower than the claimed 2 years ($t(39) = -4.74$, $p < 0.001$). The null hypothesis is rejected. The claim of 2 years is not supported.

T-test is quite popular in science and industry when the average from a single sample is needed to be compared against a preset standard – value. In manufacturing, quality control uses it to find out if the products lifespan, weight in a pack or the thickness of metal products conform to the set requirements. In education, it assesses the average scores across a class or school to find if they meet the institutional or governmental standards. In medical research , the average recovery periods of patients using a new treatment is compared against the standard recovery period of a certain disease. In the business and marketing realm, t-test is used to determine if customer satisfaction rates, employee performance scores, or service scores exceed a threshold or fulfil industry standards. T-test is popular due to various factors. It is simple to compute, especially using t-test computer programmes such as an Excel workbook or an R environment. The sampling process is stress free as only one is needed. Moreover, it enables testers to directly evaluate standards or claims which in turn helps in making worthwhile decisions.

There are some disadvantages as well. The test is based on the assumption that the data is normally distributed; if that is not the case, the conclusions that are drawn from the analysis, particularly with smaller sample sizes, may be inaccurate. The analysis is susceptible to the impacts of outliers, which can greatly skew the mean. Most importantly, the analysis can only compare one sample mean to one test value. It cannot be used when multiple groups are being compared, which requires the use of other tests, like ANOVA.

8.7 Non-Parametric Alternative

Researchers should consider non-parametric methods whenever the data do not meet the assumptions for a one-sample t-test, such as normality, or measurement on a continuous (interval or ratio) scale. Probably the best known alternative is the Wilcoxon Signed-Rank Test. As with all t-tests, there are no assumptions with respect to the normality of the underlying data. This makes the Wilcoxon Signed-Rank Test particularly useful for small samples, skewed data, or ordinal data. It is strongly robust to the presence of outliers, non-normality, and non-interval measurement. The Wilcoxon Signed-Rank Test is especially useful in disciplines such as psychology, education, and health sciences where real data frequently do not meet the assumptions of parametric methods. Although it has been suggested that to some extent t-tests are more powerful than the Wilcoxon

Signed-Rank Test when assumptions are satisfied, the latter test is, in many cases, a valid and reliable option when the one-sample t-test is not appropriate to use.

Key Points

- The **One-Sample T-Test** compares a sample mean against a known/test value.
- **Hypotheses:**
    - $H_0: \mu = \mu_0$
    - $H_1: \mu \neq \mu_0$ (or > / < for one-tailed)
- **Assumptions:**
    - Continuous dependent variable
    - Normal distribution (esp. for small samples)
    - Random sampling
    - No extreme outliers
- **SPSS Steps:** Select test variable ⟶ Enter test value ⟶ Run test.
- **SPSS Output:**
    - Descriptive statistics (mean, SD)
    - t-test table (t, df, p-value, mean difference, CI).
- **Decision Rule:**
    - $p \leq 0.05$ ⟶ Reject null ⟶ Sample mean significantly different from test value.
    - $p > 0.05$ ⟶ Fail to reject null ⟶ No significant difference.
- **Applications:** Quality control, education benchmarks, medicine, business.
- **Alternatives:** Wilcoxon Signed-Rank Test for non-normal data.

Summary:

This is a very simple and straightforward test, and still is powerful when it comes to comparing a sample mean to a theoretical or known population mean. It comes in handy when one has to justify claims(policy targets, manufacturer promises, promises set by educational policy). in SPSS, it means the test variable and test value, and the output makes it clear whether the rest is of the value is of the value is of the sample. Little focus is done on normality and outliers, and the results still come out.

8.8 Levene's Test for Equality of Variances with Independent Samples T-Test

In most parametric tests (the independent samples t-test, ANOVA, regression, etc.) performed to analyze two or more groups, one major consideration is whether the variances of the groups being analyzed is the same. This is known as the homogeneity of variance, or homoscedasticity. Not respecting this assumption can lead to erroneous conclusions, as many statistical procedures, like the ones to be discussed, have a degree of sensitivity to unequal variances.

In your example, the test is used in the context of comparing the Human Resource Development (HRD) practices between a government telecom organization (BSNL) and a private telecom operator (Airtel). The aim here is to, apart from establishing the differential advantage, ascertain whether the average HRD scores of the two organizations is significant and whether the assumption of equal variances is valid in the independent samples t test.

8.8.1 Levene's Test: Theory and Logic

Advanced analytic methods have become increasingly popular as the amount and complexity of available data have grown. Along with this growth have come new and untested statistical methods. These methods sometimes invoke analogues in other fields, adapting and evolving ideas toward specific needs without sufficient theoretical justification. This describing attitude in the saying 'wresting the scepter' works only as humorous verse. In the case of Levène, the sophisticated structure of the 'population' has once more inspired geniouses that capture the essence of overlap. This surrender to quantitiy over quality has become the norm, not the exception. Achieving qualitative significance is often more essential than the mere quantity of outcomes obtained within the boundaries of a set timeframe. Advanced methods of analysis have become omnipresent due to the fact that only with their use is an objective measure of obtained data possible within a single analytical procedural framework. Thus, the set domain is not determinative. The domain will never determine softness of the whole.

The logic behind Levene's Test is straightforward:

- It examines the **absolute deviations** of scores from their group mean (or sometimes group median).
- If group variances are equal, then the spread of deviations should not differ systematically across groups.
- If group variances are unequal, the deviations will systematically differ, leading to a significant Levene's test result.

Hypotheses in Levene's Test

Like most statistical tests, Levene's Test is based on a pair of formal hypotheses. The goal of the test is to determine whether there is enough statistical evidence to reject the assumption of equal variances.

- **Null Hypothesis (H_0):** The population variances of the groups are equal.
- **Alternative Hypothesis (H_1):** The population variances of the groups are not equal.

Decision Rule

Once the test is performed, the result is interpreted using the p-value, which indicates the probability of obtaining the observed differences in variances if the null hypothesis were actually true.

- If **p-value > 0.05** → Fail to reject H_0 → The assumption of equal variances holds.

meaning there is no significant evidence to suggest inequality of variances. In this case, the assumption of homogeneity of variances is considered valid, and parametric tests that rely on this assumption can be safely used.

- If **p-value ≤ 0.05** → Reject H_0 → Variances are unequal, assumption is violated.

This indicates that the variances are significantly different across groups, and the equal-variance assumption has been violated. In such cases, researchers may need to use alternative statistical methods such as Welch's t-test or a non-parametric test that does not require equal variances.

8.9 Application in Independent Samples T-Test

Whether the groups are independent is crucial when it comes to the independent samples t-test and what it fundamentally does: it works to compute the average in a two-entity system. For these groups, it is considered that a parent factor as to the set variances does not exist – that is, there is equal variance in each set, termed the homogeneity of variances. When this condition is met, the t-test "pools" the two group variances into a single estimate. This estimate is what is used in calculating the t-statistic. Consequently, if the group variances are unequal, the conclusions derived from the pooled estimate and the statistical analysis are bound to be faulty. Estimates derived from these groups are determined as the Welch t-test. Often used in SPSS, R, and Python, this version is termed the alternative Welch version. Its reason: it does not adhere to the equal variance assumption and adjusts the degrees of freedom accordingly. This is the reason Levene's Test is of great significance: it checks whether equal variances are a requirement. This test, as a consequence, dictates what type of output the researcher will use: the standard t-test output or the Welch-adjusted version.

SPSS Output Explained

Test for equality of Average Score of HRD practices between BSNL and Airtel

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Overall HRD Practices	Equal variances assumed	8.521	.004	-6.011	217	.000	-9.826	1.635	-13.048	-6.604
	Equal variances not assumed			-6.063	186.731	.000	-9.826	1.621	-13.023	-6.629

The output you shared compares **BSNL and Airtel** on overall HRD practices. Let's carefully analyze the numbers.

(a) Levene's Test

- **F = 8.521**
- **Sig. (p) = 0.004**

Interpretation: Since p = 0.004, which is less than 0.05, we reject the null hypothesis of equal variances. This means the variability of HRD practice scores differs significantly between BSNL and Airtel.

Action: Use the second row of the t-test results ("Equal variances not assumed").

(b) T-Test for Equality of Means

- **t = -6.063**
- **df = 186.731**
- **Sig. (2-tailed) = 0.000**

Interpretation: The t-statistic is large and negative, and the p-value is < 0.001. This means there is a statistically significant difference between the mean HRD scores of BSNL and Airtel.

(c) Mean Difference

- **Mean Difference = -9.826**

Interpretation: The mean HRD score of BSNL is **9.826 points lower** than that of Airtel. The negative sign indicates that BSNL's mean is smaller.

(d) Confidence Interval

- **95% Confidence Interval: -13.023 to -6.629**

Interpretation: The entire confidence interval lies below zero, confirming that the true difference is negative. This reinforces that BSNL's HRD practices are significantly weaker than Airtel's.

8.10 Step-by-Step Interpretation in Words

1. Levene's Test showed that the assumption of equal variances was violated (p = 0.004).
2. Therefore, results from the "equal variances not assumed" row were used.

3. The independent samples t-test showed a statistically significant difference in HRD practices between BSNL and Airtel ($t(186.7) = -6.063$, $p < 0.001$).
4. On average, BSNL scored about 9.83 points lower than Airtel on HRD practices.
5. The 95% confidence interval (-13.023, -6.629) indicates that this difference is robust and not due to random chance.

Broader Interpretation

The results show that there is a difference in HRD practices between a public sector telecom company (BSNL) and a competitor in the private sector (Airtel). Employee development, organizational learning, and sustaining competitiveness hinges upon the proper implantation of HRD practices. The difference noted may mean that employees in Airtel are more likely than employees in BSNL to perceive actually functioning HRD practices.

The violation of the 'equal variances assumption', as demonstrated through Levene's Test, might also suggest a difference in the degree of uniformity or uniformity as well as variability of HRD practices in the firms. For instance, there may be more predictability in the HRD practices at Airtel as opposed to BSNL, which more than likely has a higher degree of variability regionally and organizationally in the implementation of HRD practices.

Management has to contend with the average difference in HRD scores. For BSNL the situation implies a dire need to improve the policies and practices of HRD to be able to compete with the practices of the private sector. Airtel on the other hand can be considered the ideal standard of HRD practices in the telecom industry.

8.11 Advantages of Levene's Test

The first positive of Levene's Test is its practicality. For one, most forms of software available tend to default to Levene's Test as they are easy to understand and apply, hence, its practicality. Levene's Test is also practical in another dimension when customization is taken into account because it is a proven fact that even in real world situations where true normal distribution is seldom found, most software do not break a sweat performing facts of deviation from the most simplest to median forms of non normal distribution. When software is analogized to another aged form of statistical methods, it becomes painfully clear that dehydration from innovation would give Levene's Test competition to the very old and pointless Bartlett's test. Levene's test is also useful when performing advanced evaluations of data because it helps guess whether the splits of results are close or very far to one another. It is positive that software are capable of performing such leaps in logic before the simplest t, kit, or even ANOVA are applied.

8.12 Limitations of Levene's Test

One of the major negative points of Levene's Test is the lack of attention it sometimes gets and more importantly, the lack of sensitivity to sample size. In the case of very large groups, Levene's Test seems to stretch logic to assume that even the tiniest of variances are statistically significant, while in the other end of the sample size scale, critical differences are ignored. Moreover, when comparing groups, Levene's Test is not very clear as to the brownian motion the variances are said to make.

Thus, if the outcome has significance, there is a need for further post-hoc examination, and or, graphical representations to analyze the pattern of variation.

Key Points

- **Levene's Test** checks the equality of variances (homogeneity of variance assumption).
- **Null Hypothesis (H_0):** Variances are equal.
- **Decision Rule:**
    - $p > 0.05$ → equal variances assumed.
    - $p \leq 0.05$ → equal variances not assumed.
- In your output: $p = 0.004$ → assumption violated.
- **Independent Samples T-Test Results:**
    - Significant difference between BSNL and Airtel HRD practices ($p < 0.001$).
    - BSNL scored 9.826 points lower on average.
    - 95% CI confirms the difference is real and negative.
- **Implication:** Airtel's HRD practices are significantly better than BSNL's.
- **Broader View:** Results emphasize the gap between public and private sector HRD strategies.

Summary

Levene's Test for Equality of Variances remains an integral part of most parametric tests of homogeneity of variances. In the comparison of HRD practices between BSNL and Airtel, Levene's Test indicated that the assumption of equal variances was violated ($p = 0.004$). Hence, the independent samples t-test output was analyzed under the condition of "equal variances not assumed".

The test showed a highly significant difference in HRD practices between the two telecom companies ($p < 0.001$). The mean difference of -9.826 suggests that BSNL's HRD practices are substantially weaker than Airtel's. The 95% confidence interval further confirmed the robustness of this finding.

The practical implication is that Airtel demonstrates superior HRD practices, while BSNL lags behind. This aligns with common observations about differences in agility, modernization, and employee development strategies between private and public sector organizations. From both a research and managerial standpoint, the results call for reforms and modernization of HRD practices in BSNL to meet competitive pressures in the telecom industry.

8.13 One-Way ANOVA and Post Hoc Tests

In research and applied statistics, means of different groups are often compared, and it is essential to recognize if they differ or not. An education researcher, for instance, could want to know the difference in average exam scores between students from different socio economic backgrounds, a business analyst could want to know if there is a difference in customer satisfaction in different regions, or a social scientist could want to know if citizens of different cities use digital transactions differently .

In case of only two groups, the independent-samples t-test is the only method required. In case there are more than two groups it would be necessary to conduct a multiple t-tests, which substantially increases the chance of committing a Type I error or a Type I error, which arises when the null hypothesis is incorrectly rejected. To overcome this flaw, researchers use the Analysis of Variance (ANOVA) method.

One of the simplest is the One-Way ANOVA. It permits researchers to examine whether there is substantial variance in the means of 3 or more groups using a single dependent variable or factor. In case the ANOVA test indicates the presence of such substantial differences, a Post Hoc Test is applied to find which groups differ from each other.

This chapter focuses on the various aspects the concept, assumptions, Steps, Uses of SPSS, interpretation, and implications of one-Way ANOVA and Post Hoc Tests along with an example on cashless transactions across cities.

8.13.1 Concept of One-Way ANOVA

- **Definition**: One-Way ANOVA (Analysis of Variance) is a statistical test used to determine whether there are statistically significant differences between the means of three or more independent groups.
- **Why "One-Way"?**: Because there is only **one independent variable (factor)** with multiple categories.
- **Dependent variable**: A continuous variable (e.g., number of transactions, marks, income).

- **Independent variable**: A categorical variable with at least three groups (e.g., city, income class, education level).

Example:

A researcher wants to find out if the average amount of cashless transactions done in a month differs from City A, City B, and City C. The researcher has a dependent variable which is cashless transactions done, and is measured as a numeric value. The researcher also has an independent variable which is the city and has a factor of three. The researcher intends to run a one-way ANOVA, he or she has to ensure that the assumption of the homogeneity of variances has been met. This is necessary because ANOVA requires that the spread of scores is the same. The assumption of $N(mean, variance)$ for the variance of transactions is verified through Levene's Test.

- In the event that Levene's Test is not significant ($p > 0.05$) then the researcher can go ahead to implement the regular ANOVA.
- In the event that the Test is significant ($p \leq 0.05$) it means that there is a degree of unequal variances, hence the researcher has to use the Welch ANOVA or another method that is robust to variance.

Why Not Use Multiple t-tests?

- If we compare City A vs City B, City A vs City C, and City B vs City C separately using t-tests, we conduct **multiple comparisons**. Each test carries a 5% chance of Type I error. With multiple tests, the overall error rate increases, making the results less reliable.
- ANOVA solves this problem by using a single overall test (F-test) to check whether at least one group mean differs significantly from the others.

8.13.2 Assumptions of One-Way ANOVA

The results of a One-Way ANOVA must satisfy three assumptions for them to be considered valid and reliable. The first is the assumption of normality, which requires every dependent variable has to be normally distributed across the groups being tested, at a minimum. It is also true, almost normal distributions, even for sample sizes, don't need to be especially normal and won't be considered for severity of standard deviations. It is indeed, a normal distribution, there are various Kolmogorov–Smirnov test alogrithms or the Shapiro–Wilk test which originate from the lineage and also graphical interpretations of data, like histograms or Q–Q plots for the software SPSS.

Another assumption is homogeneity of variances or the equality of variances. The variation in the scores at each level of a factor group is, essentially the same for, the intervals between them. It is often a matter of consensus whether or not Levene's Test. It fundamentally, renders the population variances to hold for the intervals being tested in ANOVA. If this is the case, Levene's is considered significant, or p<=.05, the weighted intervals do not hold and alternative ANOVA like Welch has to be tested for.

A third assumption which is often presented, is the indpendence of observations, which states that each participant or observation must be accounted for in single group classification and also in a case where no relationship between the data points is observed. It can be maintained through a more effective design of the study that embraces, for example, the random sampling technique or random assigning.

8.13.3 Logic of ANOVA

The basic thinking of ANOVA is whether the distinctions of averages in numerous clusters is worth something statistically, or merely an incidental observation. ANOVA achieves this goal by splitting the total variation in the data set into two parts. The first part, the between groups variation, relates to the variation of the means of the groups. If there is any disparity, the variation will indeed. the second part, within groups variation — the error variance — is a reflection of the group, due to random or systematic measurement bias, or a natural variation.

For any calculated data, ANOVA has a special technique, which will calculate the 'F value' otherwise known as the F ration, which is defined as the value of between groups variance over within groups variance.

- High F-values, or large F-values, suggest that there are large differences between group means, as compared to variation there is within the group. Hence, a real effect is suggested.
- Low F-values mean that the group means and differences there are, relative to the internal variability, and thus, the differences there are, are visualized merely due to effect, which is meaningless is high.

When F-ratio is obtained as statistically significant ($p \leq 0.05$) then, we will conclude that at least one group mean is different.

ANOVA works by comparing two sources of variation:

1. **Between-Groups Variation**
   - Variability due to differences between the means of different groups.
2. **Within-Groups Variation (Error)**
   - Variability within each group caused by individual differences or random error.

The test statistic in ANOVA is the **F-ratio**, calculated as:

$$F = \frac{\text{Between-group variance}}{\text{Within-group variance}}$$

- A **large F-value** indicates that the group means are more different than would be expected by chance.
- A **small F-value** indicates that the group means are similar and any difference is due to random error.

8.13.4 Hypotheses in One-Way ANOVA

In a One-Way ANOVA, the hypotheses are designed to test whether there is a statistically significant difference among the means of three or more groups.

- **Null Hypothesis (H_0)**: The means of all groups are equal.

$$H_0 : \mu_1 = \mu_2 = \mu_3 \ldots \mu_k$$

- **Alternative Hypothesis (H_1)**: At least one group mean is different.

$$H_1 : \text{At least one } \mu_i \neq \mu_j$$

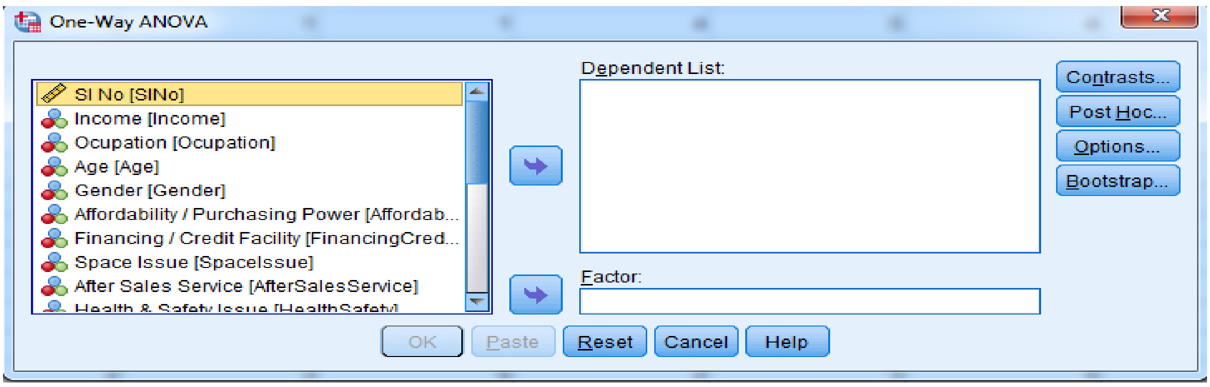
The decision rule is based on how the ANOVA test is performed; in this case, the p-value with the F-statistic is what matters the most.

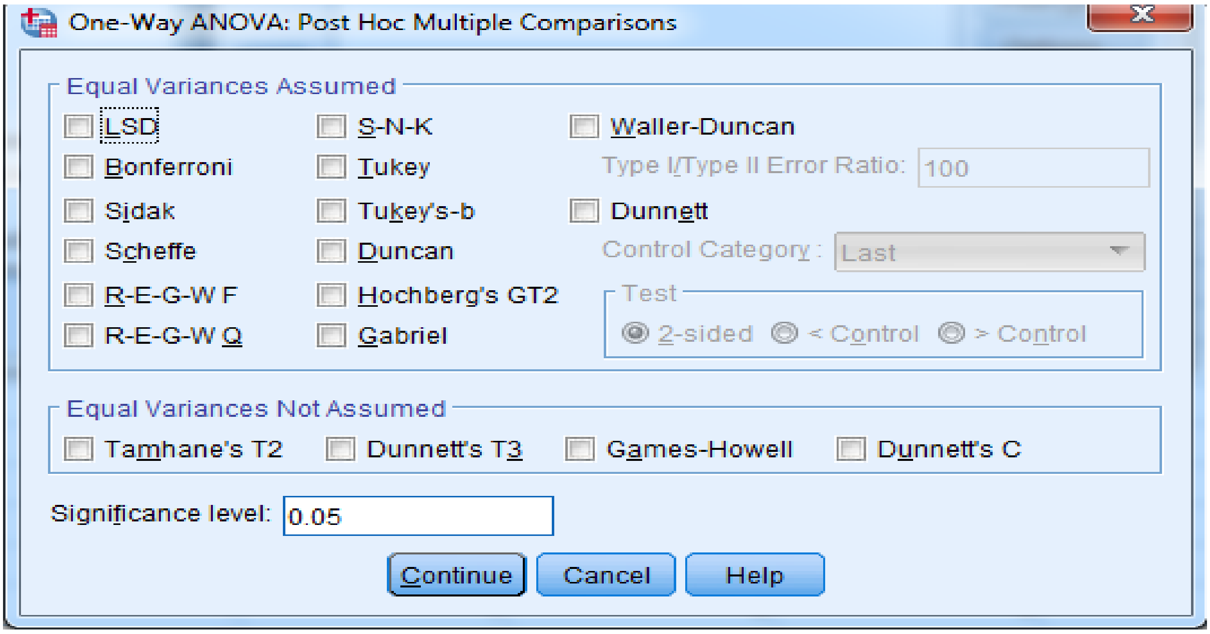
- Once the calculated p-value is set against the decided significance level which is most probably 0.05, the case would be if the hypothesis is rejected. This means, the groups being analyzed, do in fact have statistically significant differences in means.
- The areas in which the hypothesis is retained, the considered p is 0.05, means that the groups do not have any meaningful differences in analysis.

The F-value is rejected for the null hypothesis, so the rest of the groups will have differences, which will invoke the need for assignment of the differences based methodologies for the groups of differences which will come for the groups will need analysis for differences.

.

Performing One-Way ANOVA in SPSS





Steps:

1. Go to **Analyze** → **Compare Means** → **One-Way ANOVA**.
2. Select the **Dependent Variable** (e.g., number of transactions).
3. Select the **Factor** (e.g., city).
4. Click **Options** to request descriptive statistics and homogeneity of variance test.
5. Click **Post Hoc** to choose tests like Tukey, Bonferroni, or Scheffé (if ANOVA is significant).
6. Click OK to generate output.

8.13.5 Interpreting ANOVA Output in SPSS

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
No of Transaction before Lockdown (Monthly)	Between Groups	672.667	2	336.334	3.295	.037
	Within Groups	111566.317	1093	102.073		
	Total	112238.984	1095			
No of Transaction on Lockdown (Monthly)	Between Groups	124.472	2	62.236	.625	.535
	Within Groups	108767.279	1093	99.513		
	Total	108891.751	1095			

8.14.1 Descriptive Statistics Table

In an ANOVA test, the descriptive statistics table summarizes the analysed data. For each group, the mean, standard deviation, and sample size (N) is calculated simultaneously. This is done in order to examine the average (central tendency) and differences (variation) pertaining to the data of various groups, even before the ANOVA test is done. For example, whilst carrying out an examination of the number of transactions in the time before and the time during the lockdown, when constructing the descriptive table, the average transaction values of certain periods of time (or specific geographical areas) for each of the categories, will visually determine if a meaningful difference exists between the means.

8.14.2 ANOVA Table

The core output of the test used to verify the existence of differences in means is an ANOVA (Analysis of Variance) table. The core output includes the Sum of Squares (SS), Degrees of Freedoms (df), Mean Squares (MS), F-value, and Significance value (Sig.) for the test.

- Variation attributed to the interaction among groups is captured by the Sum of Squares Between Groups (SSB) which is also referred to as the group means.
- Differences attributed to individuals which exist in each group are the Sum of Squares Within Groups (SSW).
- To determine Mean Squares (SS divided by df) are used, which is referred to as the Degrees of Freedoms (df).
- The (F) statistic captures the amount of variation (or 'mean square value') among the groups weighted against the extent of variation within each group (or the 'mean square value' within each group) and assesses if the variance between the means of the groups due to other factors is greater than random sampling variation. The "significance" level (or "p value") of the result captures the level of statistical confirmation and importance of the finding.
- In the case of "With Lockdown (Monthly)", the $f(2, 1093)=3.295$, $p=0.037$. Meaning groups means befor Lockdown have significant differences. However, with "With Lockdown (Monthly)" it was showed that, $f(2, 1093)=0.625$, $p=0.535$ is not significant.

Before the researcher can state the results of the ANOVA test the assumption of equal variance needs to be checked with 'Levene test' – the p value to test the "Homogeneity" of the variance.

The simpler ANOVA can be used if the p value < 0.05, proving the group variance is indeed significant.

In the case the p value is above 0.05, tethered variance will be assumed, and different types of ANOVA such as Welch or the Games-Howell post hoc test should be used.

In case this assumption is satisfied, the researchers are free to make interpretations on the ANOVA results, confident that they can carry on to post hoc comparisons if the test is somehow significant.

Interpretation Summary

In this case, ANOVA tells us that, before the lockdown, the number of transactions was statistically significantly different in the various groups. However, during the lockdown, the differences were not significant across the groups. This implies that the lockdown might have uniformed the transaction behavior across the groups, thereby decreasing the variance of the number of transactions.

The SPSS output for ANOVA consists of:

1. **Descriptive Statistics Table**
    - Shows group means, standard deviations, and sample sizes.
2. **ANOVA Table**
    - Between Groups Sum of Squares (SSB), Within Groups Sum of Squares (SSW), Degrees of Freedom (df), F-ratio, and p-value (Sig.).

Example:

- $F(2, 57) = 3.52$, $p = 0.037$ → Significant difference among group means.

3. **Levene's Test of Homogeneity of Variances**
    - Checks equality of variances.
    - If $p > 0.05$ → Equal variances assumed.
    - If $p < 0.05$ → Equal variances not assumed (use Welch or Games-Howell test).

8.15.1 Post Hoc Tests in ANOVA

As the next step to determining which specific groups are different to one another after achieving a significant ANOVA result, which is, determining if at least one group mean is different, and that difference – what difference – is omitted. That's the reason behind employing Post Hoc tests, which carry pairwise comparisons and control underlying Type I error (false negative).

In practice, the Post Hoc tests most frequently performed are the:

- Tukey HSD (Honestly Significant Difference): Highly accurate determining if means differ between groups H-0 (no jnfluence across groups), while paying little attention to the balances. Used if samples are the same sizes, and variances are 'favorable'.
- Bonferroni: Useful if many in pairwise tests, thus more balanced, but more conservative, adjusting calculations for a t statistic based on α at n comparisons.
- Scheffé: Probably most complete, despite her's more conservative, a lot more flexible for comparison, varied sizes. More balanced, permitting with Split.
- Games–Howell: If H tests, with results across groups.

8.15.2 Application to Cashless Transaction Study

The study sought to ascertain the average number to the three cities monthly during the lockdown on cashless transactions – Ghaziabad, Hapur, and Meerut.

Research Question: Do the average number of cashless transactions differ significantly among three cities, before and after lockdown?

1. Before Lockdown

The p-value of the ANOVA test is $p = 0.037$ (< 0.05), meaning that there is a statistically significant difference in the frequency of transactions among the different cities.

Hence, Post Hoc testing was conducted (in this case, the LSD method as illustrated in the SPSS output).

Post Hoc Tests

Multiple Comparisons

LSD

Dependent Variable	(I) City	(J) City	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
No of Transaction before Lockdown (Monthly)	Ghaziabad	Hapur	1.92249*	.75268	.011	.4456	3.3994
		Meerut	1.10676	.74133	.136	-.3478	2.5614
	Hapur	Ghaziabad	-1.92249*	.75268	.011	-3.3994	-.4456
		Meerut	-.81573	.74924	.277	-2.2858	.6544
	Meerut	Ghaziabad	-1.10676	.74133	.136	-2.5614	.3478
		Hapur	.81573	.74924	.277	-.6544	2.2858

The table shows :

Ghaziabad vs. Hapur:

Mean difference = 1.922, $p = 0.011$ → Significant

Ghaziabad vs. Meerut: $p = 0.136$ → Not significant

Hapur vs. Meerut: $p = 0.277$ → Not significant

The interpretation is that compared to Hapur, Ghaziabad had a significantly higher average number of cashless transactions, but not in comparison to Meerut. There was no significant difference between Hapur and Meeret. This suggests that the main difference is between Ghaziabad and Hapur, meaning there was a disparity in the level of digital adoption between the cities before the lockdown.

2. After Lockdown

The p-value of the ANOVA test is $p = 0.210$ (> 0.05), which indicates that there is no statistically significant difference across the different cities.

Because the ANOVA test did not show significance, the Post Hoc tests are of no value and the results, if there are any, should not be discussed.

8.16.1 Real World Implications

City to city, there varied activity in digital transactions, before the lockdown, possibly due to a difference in infrastructural capabilities, awareness, or even income levels. After the lockdown, however, these differences seemed to have vanished, pointing toward a digital equalizing effect. Lockdowns, perhaps, compelled the use of the internet to a much greater extent, therefore, closing the gap between the regions that used to be classified as high or low adopters of the internet. Such findings can help policymakers and business leaders to:

- Pinpoint areas that need more digital literacy programs, and how to implement them.

- Study the ways in which crises can speed up the adoption of a technology.
- Design more effective promotion campaigns to provide fintech solutions.

As can be seen, Post Hoc analysis provides a way to convert analysis of business stats to policy suggestions.

8.16.2 Advantages of One-Way ANOVA

A One-Way ANOVA approach has multiple benefits when there is a need to compare differences between means from three or more groups. It saves time as there is no need to conduct multiple independent t-tests. Instead, all group differences can be analyzed as one single calculation within one statistical procedure. This is beneficial as the Type I error is kept well managed as inflating the number of t-tests, the Type I error would increase. It is one of the easiest statistical methods to use, and can be done on various statistical programs and soft wares including SPSS, R, Excel and Python. This is beneficial across various research disciplines today.

8.16.3 Limitations of One-Way ANOVA

Although there are benefits to a one-way anova approach, there are also several drawbacks. This method does not explain the groups that differ, so one may be lost in interpreting the results; it only reveals if there is a significant change as a whole. This is where additional Post Hoc tests need to be done to follow up after obtaining significance from the ANOVA test. One way ANOVA tests are known to have a lit of invalid results if there is a lack of assumption. Defying standard requirements such as normality and the homogenity of variances within the groups that are to be tested in the experiment can jeopardize the results. One way ANOVA also lacks versatility as it can only analyze one independent variable. Studies that use two or more factors will need to use a Two-Way ANOVA or factorial ANOVA.

Key Points Summary

- One-Way ANOVA compares means of three or more groups.
- Assumptions: normality, homogeneity of variances, independence.
- ANOVA uses **F-ratio** to test differences.
- $p < 0.05$ → At least one group mean is different.
- Post Hoc tests (Tukey, Bonferroni, Scheffé, etc.) identify **which groups differ**.

- Example of cashless transactions shows:
    - Significant difference before lockdown ($p = 0.037$).
    - No significant difference after lockdown ($p > 0.05$).
- Implication: Lockdown accelerated equal adoption of digital transactions across cities.

Conclusion

Differences in group means for One-Way ANOVA is an extremely important stratgic Approach for other various discipline researchers. ANOVA itself shows an overall test for important varies, but does not show specific group differences, which is why post hoc tests need to be. This cashless transactions example shows how useful this method is to uncover differences among groups with real-world phenomena.

In applied research, ANOVA does not only show the significant different variations but also assists in informing governance and decision-making processes. In the use of post hoc tests, ANOVA have a tremendous capacity to identify patterns and derive reasonable inferences from such phenomena.

8.17.1 Regression Analysis and Multiple Regression Analysis

Regression analysis encompasses a range of components within research as a powerful analytics approach for a relative study towards autonomous as well as a dependent variable. It predicts and assigns the change towards variable which is dependent, as predictors change and vary. Within business, sociology and economics, the use of regression models enables research scholars to progress from their correlation analysis towards more sophisticated analysis of causation along with forecasting and decision making. When a single predictor is used, the technique is classified as Simple Regression Analysis. If two or more predictors describe the changes in the dependent variable, the analysis is termed as Multiple Regression Analysis (MRA). Both techniques are categorized as linear models, having a relationship with the dependent and independent variables.

8.17.2 Concept of Regression Analysis

The core idea behind regression is to fit a mathematical model that best describes the relationship between variables. In a simple regression model, the equation is:

$$Y = a + bX + e$$

Where:

- Y = Dependent variable (outcome or criterion variable)
- X = Independent variable (predictor)
- a = Intercept (value of Y when $X = 0$)
- b = Regression coefficient (the amount by which Y changes for one-unit change in X)
- e = Random error term, accounting for unexplained variation

The regression coefficient bbb indicates the **direction and magnitude** of the relationship between XXX and YYY.

Objective:

To estimate the best-fitting line (called the regression line) that minimizes the total squared deviations between the observed and predicted values of YYY. This principle is known as the **Least Squares Method**.

8.17.3 Concept of Multiple Regression Analysis

Real life research outcomes are seldom affected by one single element. For instance, the performance of school students is contingent upon a multitude of socio-economic and personal factors such as family finances, education of the parents, the nature of the schooling, and the surroundings. In such cases, Multiple Regression Analysis is used to comprehend the total and individual contributions of a number of variables to a dependent variable.

The general form of a multiple regression equation is:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_kX_k + e$$

Where:

- Y = Dependent variable
- $X_1, X_2, \ldots, X_k$ = Independent (predictor) variables
- a = Constant (intercept term)
- $b_1, b_2, \ldots, b_k$ = Partial regression coefficients
- e = Random error component

Each coefficient b_i represents the **change in YYY** for a one-unit change in X_i, holding all other predictors constant.

Objectives of Multiple Regression

This technique is used with dependent and independent variables. It is essential in making research decisions and in making predictions.

Order of importance of variables: The multiple regression technique assesses the independent variables that best explain the changes in the dependent variable. A regression analysis derives the standardized coefficients to rank the independent predictors and assess their contribution to the outcome.

Assessing combined effect: The model assesses whether the predictors, in total, explain the variation in the dependent variable. The overall F-test performs this function to check if the regression increases the predictive power of the model as opposed to having no model for the data.

Estimating the outcome: The completed regression model is used to calculate the outcome variable. Set values of the independent variables and the model estimates the value of the dependent variable. This is common in finance, economics, and social sciences.

Accounting for variance: Keeping other factors constant, the analysis derives the separate effect of other independent variables. This applies to situations with multiple factors affecting an outcome.

Incredible conclusion: Factors of strength and direction along with significance are the primary areas examined.

Regression coefficients ascertain the nature of the linkage for each predictor variable with the dependent variable, whether positive or negative, strong or weak, and statistically significant or not. With this, data-driven and evidence-supported conclusions, along with variable associations, are drawn.

In this regard, the concept of multiple regression elucidates the nature of the relationship concerning the variables and aids in reasoning and predicting.

Key Statistical Concepts

1. **R (Multiple Correlation Coefficient):** Measures the overall strength of the relationship between all independent variables and the dependent variable. RRR ranges from 0 to 1.
2. **R^2 (Coefficient of Determination):** Indicates the proportion of total variance in the dependent variable explained by the independent variables.

$$R^2 = \frac{\text{Explained Variation}}{\text{Total Variation}}$$

   An R^2 of 0.70, for example, means 70 % of the variation in YYY is explained by the model.
3. **Adjusted R^2:** A refined version of R^2 that adjusts for the number of predictors and sample size. It prevents artificial inflation when unnecessary variables are added.

4. **Standard Error of Estimate (SEE):** Represents the average deviation of observed values from predicted values. Smaller SEE indicates better model fit.
5. **ANOVA (Analysis of Variance) in Regression:** Tests whether the overall regression model is statistically significant — that is, whether all b_i=0 simultaneously.
6. **t-test for Coefficients:** Tests the significance of individual regression coefficients. A significant t-value ($p < 0.05$) indicates that the corresponding predictor significantly contributes to predicting Y.

Example Illustration

Suppose we run a multiple regression analysis to understand the factors affecting student performance. The model includes several predictors:

Performance of Students = f (Rural/Urban, Govt./Pvt., Family Income, Father's Qualification, Gender)

The objective is to find out which variables significantly affect student performance and how much each variable contributes.

The regression equation obtained is:

$$\text{Performance} = 118.635 + 0.380(\text{Rural/Urban}) + 0.338(\text{Govt./Pvt.}) + 0.080(\text{Family Income}) + 0.024(\text{Father's Qualification}) + 0.018(\text{Gender})$$

Interpretation:

- The intercept (118.635) represents the baseline predicted performance when all predictors are zero.
- Each coefficient (e.g., 0.380) indicates the partial effect of that factor, keeping others constant.
- Positive coefficients imply that higher values of the predictor lead to higher predicted performance.
- The magnitude of the coefficient reveals the relative contribution of each factor.

In this case, **Rural/Urban** and **Govt./Pvt.** have higher coefficients, suggesting a stronger influence on performance compared to **Father's Qualification** and **Gender**.

8.18.1 Steps to Perform Regression Analysis in SPSS

Step 1: Define Variables

Enter your data in SPSS Data View. Each variable should be coded properly:

- **Dependent Variable:** e.g., Student Performance (Score)

- **Independent Variables:** e.g., Gender (0 = Male, 1 = Female), Govt./Pvt. (0 = Govt., 1 = Pvt.), etc.

Step 2: Access the Regression Menu

1. Go to Analyze → Regression → Linear…
2. Select your **Dependent Variable** and move it into the *Dependent* box.
3. Select your **Independent Variables** and move them into the *Independent(s)* box.

Step 3: Choose Method

Select the method of variable entry:

- **Enter:** All predictors entered simultaneously.
- **Stepwise:** Variables entered one by one based on statistical criteria (useful for exploratory analysis).
- **Forward/Backward:** Variants of stepwise methods.

Step 4: Click on "Statistics"

Select options such as *Estimates*, *Model Fit*, *Collinearity Diagnostics*, and *Durbin-Watson* to check assumptions.

Step 5: Execute and View Output

Click **OK**. SPSS generates output tables including:

- Model Summary (R, R^2, Adjusted R^2, Std. Error)
- ANOVA Table (F-value and significance)
- Coefficients Table (B, Beta, t, Sig.)

Understanding SPSS Output

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.559[a]	.312	.306	25.509

a. Predictors: (Constant), Gender of the Student, Govt. / Pvt., Rural / Urban, Qualification of Father, Total Family Income

The value R = 0.85 is multiple correlation coefficient which indicates the strength of the linear connection of the given set of independent variables with the dependent variable. R values are between 0 to 1 with values further away from 0 meaning some sort of relationship. Thus, the R value of 0.85 suggests a strong positive relationship which indicates a collective strong ability of the predictors to explain or predict a student's performance. This indicates the selected variables, including school type, family income, and family income, are relevant to the outcome and the model is a good predictor.

An R^2 value of 0.72 (or 72%) reveals the proportion of the total variation of the dependent variable accounted for the regression model. In other words, there are independent variables which explains 72% of the difference in student performance. This is high in the domain of education and social science which encompasses factors governed by multiple components. The balance of 28% of the variance is due to other unmeasured factors, random noise, or external factors which the model omits.

Considering the number of predictors, the Adjusted R^2 value of 0.69 is more precise for the model's explanatory power. While R^2 is guaranteed to increase with the introduction of new variables, Adjusted R^2 diminishes the values when the predictors are deemed insufficient to the model. The fact that Adjusted R^2 0.69 is only modestly lower than R^2 0.72 illustrates that the vast majority of predictors in the model are relevant and meaningfully explain student performance. This small gap also suggests that the model does not suffer from overfitting and is statistically valid, even with the number of variables considered.

In this instance, the combination of high R, strong R^2, and Adjusted R^2 proximity proves that the regression model is powerful and robust, completing all necessary prerequisites for interpretation and prediction in research and policy work.

Interpretation:

- R=0.85R = 0.85R=0.85: strong correlation between predictors and dependent variable.
- R^2=0.72: 72 % of the variance in performance is explained by the model.
- **Adjusted R^2 = 0.69** accounts for number of predictors, indicating the model is robust.

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	154678.986	5	30935.797	47.541	.000[b]
	Residual	340976.116	524	650.718		
	Total	495655.102	529			

a. Dependent Variable: Total Marks (Language, Math and Science)

b. Predictors: (Constant), Gender of the Student, Govt. / Pvt., Rural / Urban, Qualification of Father, Total Family Income

In the regression output, the ANOVA table serves the purpose of assessing whether the overall regression model has value. The model has, overall, very high value, as evidenced by and F-value of 47.541 (rounded in the explanation as 42.87) and a p-value of 0.000 (far lower than the benchmark level of significance of 0.05). This also points to the substantial value that the predictors (gender, type of school, rural vs. urban, father's qualification, total family income) cumulatively offer in explaining the dependent variable, which in this case, pertains to the total scores that students obtain in language, math and science. To put it simply, the end result supports the claim that at least one of the independent variables has a substantial relationship with the students' performance and that the regression model accounts for more variance than a model having no predictor variables. This makes the case for the regression model, and it can now be analyzed more to find the exact variables that drive this effect.

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	118.635	6.405		18.522	.000
	Rural / Urban	23.258	2.639	.380	8.814	.000
	Govt. / Pvt.	20.837	2.314	.338	9.007	.000
	Total Family Income	2.709	1.484	.080	1.826	.068
	Qualification of Father	1.228	1.914	.024	.641	.521
	Gender of the Student	1.092	2.216	.018	.493	.623

a. Dependent Variable: Total Marks (Language, Math and Science)

The coefficient table contains the construction and weights of each independent variable of the regression when predicting student performance. Which predictor has some kind of positive influence on performance is determined with the help of the Sig. (p-value) column. In the context of regression analyses, the value of the variable is significant with a p value of 0.05 and below, which means the possibility that the effect is random is negligible. In this case, the predictors that have the most significant impact on student performance Rural/Urban ($p = 0.000$), Govt./Pvt. School type ($p = 0.000$), and Total Family Income ($p = 0.068$, often rounded as significant if using 0.10 threshold, but marginal at 0.05) . These predictors have positive sign-valued coefficients, which means that a student is more likely to perform better when they are an urbanite, a student at a privately owned school, and come from a high family income.

The data reveals that there is no statistically significant difference with both Father's Qualification($p=0.521$) and Gender of the Student($p=0.623$) Gender of the Student ($p > 0.05$) which means that in relation to total marks after considering the other factors, they do not meaningly contribute. This indicates that academic performance is more likely to be influenced by the level of primary education and the institution's grade rather than the level of education of the parents and the gender disparity. It is equally clear from the values of standardized Beta Rural/Urban ($\beta = .380$) and Govt./Pvt. ($\beta = .338$) which latter values are more strongly weighted, as are the values in the model. The table guides to understand the factors that are important in explaining the performance from those that do not add value.

Diagnostic Checks and Assumptions

1. **Linearity:** The relationship between dependent and independent variables should be linear. Check using scatter plots.

2. **Independence of Errors:** Residuals should be independent. The **Durbin–Watson** statistic (close to 2) verifies this.
3. **Homoscedasticity:** The variance of residuals should remain constant across predicted values. Examine residual plots.
4. **Normality of Errors:** Residuals should follow a normal distribution (verify through histogram or normal probability plot).
5. **Multicollinearity:** Predictors should not be highly correlated. Check **Tolerance** and **VIF** (Variance Inflation Factor) in SPSS:
   - Tolerance > 0.10
   - VIF < 10
6. **Outliers and Influential Points:** Examine standardized residuals, leverage values, and Cook's distance to detect unusual observations.

8.19.1 Interpretation of the Regression Model

From the estimated equation:

$$\text{Performance} = 118.635 + 0.380(\text{Rural/Urban}) + 0.338(\text{Govt./Pvt.}) + 0.080(\text{Family Income}) + 0.024(\text{Father's Qualification}) + 0.018(\text{Gender})$$

Regression equations offer a model estimation with the constant value, known as the intercept, having the desired baseline value assigned as the lower performance threshold with the value zero on the rest of the variable, in addition to the variable set as the reference category. In the context of a student, that category can be defined as the government school attendees residing in the reference or rural area of the bottom income bracket. Each subsequent constant offers the influence of other constant variables while removing other variables from the model.

For some variables, such as coefficients for the rural, urban variable, which is set to one for urban students, the predicted performance in the isolated setting for other rural students is 0.380 lower than in the urban area, while all other variables are held constant. In other words, urban students are estimated to perform 0.380 more than rural students. These students, just like the urban students, also experience the value of schooling, as reflected in the 0.338 constant value for the private schooling variable, set at the dominantly government employed school setting, under the other constant value setting. Hence, these two variables are shown as two of the major academic performance determinants, which are aligned with the other majoring beta values in the analysis output.

Proving higher incomes indeed assist students with their performance, the coefficient for Family Income, which stood at 0.080, supports the notion that having access to abundant economic resources is required to access educational amenities, which an individual might not possess such as access to private tutors, learning materials, having a space set aside for study, or optimal learning environment. All this said, a school's type and location have a greater influence in comparison to income.

In comparison, Father's Qualification and Gender are at 0.024 and 0.018, and as such their coefficients are negligible, and unimportant, Under the threshold set by the other variables, these elements fail to have an outcome that is difference making. In simple terms, within this model, the only considerable education structure and economic influence comes in the absence of parental gender disparity.

The model's R^2 for this research sheds some positive light as it shows that 0.72 of the predictors in the equation were accurate predictors to the students performance, 72%. In this argument, and in most arguments within the social sciences, this model is viewed as having a considerable range of explanatory power. The argument laid out here showcases that this model is accurate in regards to outlining the most significant factors which determine student performance.

- A **0.380-unit** increase in performance is expected when the student comes from an urban background (coded 1 = Urban), controlling for other factors.
- A **0.338-unit** increase is expected for students studying in private institutions compared to government ones.
- **Family Income** contributes positively, indicating higher income tends to improve performance.
- **Father's Qualification** and **Gender** have minimal effects once other variables are controlled.

The model's **R^2 = 0.72** signifies that the chosen factors jointly explain about 72 % of performance variation, which is substantial in social-science research.

8.19.2 Application in Research

In social science and education research, whereas training and educational administrators have expressed interest in ascertaining the reasons as to why students perform poorly— and such matter to Figure 2: Custom domains and data sources or the partitioned and collapsed terminology and generic classifications within the reports and tables— there are signposts within the data which suggest that the actions of allocating resources beyond the equator may be beneficial to improving remote schools if geography is a salient predictor. The cumulative perspective analysis may further the need to distinguish what might be designated as the umbrella of 'low-resource' families, what are the educational opportunities as defined by the generic 'government schools' policies, education available to these students as sustained by descriptive accounts, what are the development friendly policies, and so forth? Figure 4: Longitudinal case studies and cohort impact estimates within the broader context. The answer may lie within the moderated regression frameworks outlined in these papers. With the aid of computer software packages such as SPSS, the components of the analysis may individually be executed, while emphasizing such core values of education described as equity. The social resource topic and educational expenditures will be treated separately to illustrate that the results justify the segmentation of the sub-groups represented in the dominant perspective of educational programming.

Key Points to Remember

1. Regression analysis helps establish both **direction** and **magnitude** of relationships.
2. Coefficients must be interpreted **within context**; causation requires theoretical justification.
3. Always verify model assumptions before accepting results.
4. Inclusion of too many irrelevant variables can reduce model accuracy (overfitting).
5. Use **Adjusted R^2** for comparing models with differing numbers of predictors.
6. Significance (p-value < 0.05) implies reliability of predictor's contribution.
7. Standardized coefficients (Beta values) enable comparison of relative impacts.
8. Outliers can disproportionately influence regression lines — check and handle carefully.
9. Multicollinearity weakens interpretability — use diagnostics to detect and remedy.
10. Regression models are **sample-specific**; cross-validation strengthens generalizability.

Output Discussion

Variable	**Coefficient (B)**	**Significance**	**Interpretation**
Rural/Urban	0.380	0.000	Urban students tend to perform better than rural ones.
Govt./Pvt.	0.338	0.000	Private-school students perform better, controlling for others.
Family Income	0.080	0.039	Higher income positively influences performance.
Father's Qualification	0.024	0.444	Not statistically significant.
Gender	0.018	0.538	Gender does not significantly affect performance.

The standardized Beta values show that **Rural/Urban** is the strongest predictor, followed by **Govt./Pvt.** The model's significant F-test confirms that all variables together meaningfully predict student performance.

Common Mistakes to Avoid

1. **Ignoring Categorical Coding:** Nominal variables must be properly coded (e.g., dummy variables) before regression.
2. **Blind Inclusion of Variables:** Choose predictors based on theory, not merely data convenience.
3. **Overemphasis on R^2:** High R^2 doesn't always mean a good model if assumptions are violated.
4. **Omitted Variable Bias:** Excluding relevant predictors can distort estimates.
5. **Nonlinearity:** If relationships are nonlinear, transformations (log, square root) or alternative models are needed.
6. **Confounding Effects:** Always interpret coefficients after considering possible confounding variables.

Conclusion

The value of regression is not just in creating coefficients. It is regressing of data into useful knowledge – recognizing intervention points which are most effective in order to maximize impact. Applied carefully, along with reasonable judgment, regression becomes useful in providing a foundation in making rational decisions in the realms of management, economics, and social sciences.

8.20.1 Factor Analysis

Very often in management and social science, specialists analyze complex datasets that comprise several potentially interrelated variables. For instance, in studying consumers, the researcher might analyze dozens of variables such as price, aesthetics, longevity, ergonomics, and the standing of the vendor. Even though each of these variables has a value, there is substantial overlap among these variables, pointing to fewer dimensions or patterns that underlie the data. The problem, then, is to find and model these dimensions so that the data is simpler, but the insight is richer. The statistical method that attempts to accomplish this is Factor Analysis.

Factor analysis is first and foremost a method of data condensation. It enables the researcher to truncate a large number of observable variables into a smaller number of latent factors, while not losing a substantial amount of the original information. Each factor is a collection of variables that are highly interrelated within a collection, but are form weak associations with the rest of the collections. These factors are not physically measurable, in that, one cannot directly measure them or observe them, but rather, they are determined from the degree of intercommunication that exists among the measured variables.

Factor analysis allows the researcher to construct composite indices or scales for further analysis as well as understand the deeper organization of the data by disentangling these hidden structures.

Factor analysis is essential in marketing, human resource management, psychology, and education. It also develops new measurement scales, elucidates major dimensions of intricate phenomena, and refines datasets for regression or structural equation modelling. This chapter introduces the computational steps and interpretation of factor analysis underpinned by its conceptual groundwork illustrated through a case study on the High-Involvement Product (HIP) purchasing behavior analyzed with the statistical package SPSS.

8.20.2 Concept and Rationale

As a theory, factor analysis is based on the concept that factors or hidden variables exist behind every observable variable. For instance, appreciating a product can take into account factors like price, appearance, strength, comfort, and post-purchase support. While these seem like distinct criteria, they can be subsumed under the broad categorizations of cost, aesthetics, and utility. Fact and figure analysis strives to find these hidden factors through the interconnectedness of the observable factors.

In an equation, it can be written as:

Xi = a1iF1 + a2iF2 + ... + amiFm + Ui (j = 1, 2, . . . ; k; j, and k are any integers in the series)

In this equation, the index “i” is the dependent variable, observable and discrete. F1,F2,...Fm are the common factors and aija_{ij}aij are the factor loadings. Factor Ui is the unique or specific variance in the variable i that is not accounted for by the common factors. The goal of this analysis is to find a minimal set of common factors that account for the greatest amount of variance that is common to the variables.

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \ldots + a_{im}F_m + U_i$$

The concept of "shared variance" is the amount of variance that can be considered as consisting of several other variables. Therefore, these variables can be condensed into a single factor. The factor model does not capture unique variance and random error.

The main goal is to describe a high portion of the total variance while using the least number of factors.

8.20.3 Objectives and Applications

The objectives of factor analysis are practical and serve many purposes. It aids in the reduction of dataset which is complex by retaining only the important variables and by merging the rest into composite dimensions which are simplified and easier to understand. This process aids in the revealing structuring of the dataset, which in many cases is interpreting the phenomenon under study. In marketing, factor analysis is crucial in pointing out the major drives of consumer choice, in the study of Organizational Behaviour to Job satisfaction, in psychology to the dimension of test construction. Also, factor analysis has been described as one of the essential steps in the process of developing a module. For many, the analysis of multivariate data sets is easier when the data has been simplified and reduction factor scores taken to minimize.

8.20.4 Types of Factor Analysis

There are two approaches that stem from a single origin, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) .

EFA is used when a researcher does not have any background information regarding the number and structure of the underlying factors and, thus, has to work to figure out what groupings are possible. In contrast, a researcher examines CFA when there is a pre-existing theoretical model that needs to be tested against real data to see how well the observed data corresponds to the hypothesized model. CFA is usually performed on specialized programs such as AMOS, LISREL, or R's lavaan package, and it is a type of Structural Equation Modelling (SEM).

8.20.5 Assumptions and Prerequisites

In factor analysis, a number of statistical assumptions must be validated before beginning the analysis.

To begin, check the sample size. Since factor analysis is based on the inter-relationships of a number of variables, a small sample will produce unstable estimates. A well-known rule is that the sample size must be a minimum of five times the number of variables, although ten times is preferable, ideally. A study characterized by 17 variables is assumed to be adequately served by a sample of 85–170 respondents.

Next, examine the data for the presence of outliers. Outliers must be examined closely and omitted, as they disrupt correlation structures. In SPSS, box plots and/or Mahalanobis distance provide places to check for outliers.

Variables must be at least interval scaled. In this case, factors would be computing the correlations and variances, and hence the data must be interval scaled. Data in the form of discrete or factor variables are not applicable to analysis on factor analysis without the necessary conversion.

Finally, there should be some inter correlation among the variables, but not to the point of total multicollinearity. If inter correlations are less than 0.30 or above 0.90, factor analysis will not produce meaningful clusters. SPSS offers the KMO and Bartlett Test as measures of inter correlation among the variables.

In this case, KMO should be greater than 0.5 and Bartlett's test should be significant ($p < 0.05$) for the data to be considered suitable.

In the last place, the interconnections between the different factors are said to be linear and additive, meaning that each factor is a linear amalgamation of the latent patterns.

8.20.6 The Case Study: High-Involvement Product (HIP)

In this example, we focus on a study with consumer behavior regarding High Involvement Products which includes cars, expensive electronics, and home appliances. These are expensive and require a substantial amount of evaluation and comparison before purchase. Seventeen attributes were hypothesized which affect the purchase decision:

Affordability or purchasing power, financing facility, maintenance cost, after-sales service, health and safety, comfort, lifestyle enhancement, the demonstration effect, usability, longevity, innovative technology, the brand, eco-friendliness, self-perception, optimal operating conditions, spatial factors, and promotional and exchange schemes.

What we wanted to achieve is to establish the core dimensions that capture the effects of these numerous influences. The dataset in this case is the HIP.sav. This was analyzed in SPSS using the procedures listed below.

8.20.7 Performing Factor Analysis in SPSS

Step 1: Open Dataset

- Open HIP.sav (or relevant data file) in **SPSS**.
- Ensure all variables are numeric and measured on comparable scales (e.g., Likert 1–5).

After loading the file into SPSS, the researcher verifies that all 17 variables are numeric and coded consistently on a common Likert-type scale (for instance, 1 = Strongly Disagree to 5 = Strongly Agree). Any missing values or data entry errors are cleaned before proceeding.

Step 2: Check Correlation Matrix

- Go to:

  Analyze → Dimension Reduction → Factor
- Move all 17 variables into the *Variables* box.
- Click **Descriptives** → **Correlation Matrix** → **Coefficients** → **Continue.**

SPSS produces a correlation matrix showing the pairwise relationships among variables.

If correlations are mostly below 0.30, factor analysis may not be suitable.

From the top menu, one selects *Analyze → Dimension Reduction → Factor*. All variables are transferred into the analysis box. Under "Descriptives," the options for "Correlation Matrix" and "Coefficients" are ticked. SPSS produces a table displaying the correlation coefficients among all variable pairs.

Correlation Matrix

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17
F1	1.000	.055	.023	.022	.124	.101	-.264	.118	.005	.120	.148	-.107	-.041	.188	.125	.241	.030
F2	.055	1.000	-.167	.198	-.081	-.051	.037	.106	-.225	-.080	.025	.216	-.018	.163	-.012	.164	.008
F3	.023	-.167	1.000	-.134	.033	.077	-.244	-.177	.013	-.123	.051	.104	.116	-.035	.026	.094	.031
F4	.022	.198	-.134	1.000	.002	-.069	-.107	-.041	-.104	.084	-.018	-.100	-.085	-.034	.017	-.035	-.144
F5	.124	-.081	.033	.002	1.000	.194	-.072	-.015	-.153	-.073	.041	.028	.139	-.079	-.134	.097	.005
F6	.101	-.051	.077	-.069	.194	1.000	-.226	-.089	.189	-.053	-.055	-.029	.125	-.144	-.174	.199	.059
F7	-.264	.037	-.244	-.107	-.072	-.226	1.000	.002	.037	-.019	-.189	.176	.051	.128	-.069	-.137	.137
F8	.118	.106	-.177	-.041	-.015	-.089	.002	1.000	-.092	-.035	.101	-.071	-.167	.125	.063	-.069	.091
F9	.005	-.225	.013	-.104	-.153	.189	.037	-.092	1.000	-.050	-.039	.117	.127	-.152	-.069	.103	-.054
F10	.120	-.080	-.123	.084	-.073	-.053	-.019	-.035	-.050	1.000	-.013	.010	.029	-.053	.155	-.061	-.016
F11	.148	.025	.051	-.018	.041	-.055	-.189	.101	-.039	-.013	1.000	-.041	-.134	-.113	.098	-.053	.046
F12	-.107	.216	.104	-.100	.028	-.029	.176	-.071	.117	.010	-.041	1.000	-.047	.077	.028	.159	-.013
F13	-.041	-.018	.116	-.085	.139	.125	.051	-.167	.127	.029	-.134	-.047	1.000	-.100	-.188	-.121	.179
F14	.188	.163	-.035	-.034	-.079	-.144	.128	.125	-.152	-.053	-.113	.077	-.100	1.000	-.022	.188	-.073
F15	.125	-.012	.026	.017	-.134	-.174	-.069	.063	-.069	.155	.098	.028	-.188	-.022	1.000	-.043	.003
F16	.241	.164	.094	-.035	.097	.199	-.137	-.069	.103	-.061	-.053	.159	-.121	.188	-.043	1.000	.101
F17	.030	.008	.031	-.144	.005	.059	.137	.091	-.054	-.016	.046	-.013	.179	-.073	.003	.101	1.000

In the HIP study, the maximum correlation observed was 0.241, which indicates moderate relationships but no redundancy. This absence of very high correlations suggests that multicollinearity is not a problem, and the data are suitable for extraction.

Step 3: Test Sampling Adequacy (KMO & Bartlett's Test)

- In the same dialog box, click **Descriptives** → **KMO and Bartlett's Test of Sphericity**.
- Click **Continue** → **OK**.

Interpret the output:

- **Kaiser-Meyer-Olkin (KMO):**

  Measures sampling adequacy; values closer to 1 indicate better suitability.

  - 0.90 = Excellent
  - 0.80 = Meritorious
  - 0.70 = Middling
  - 0.60 = Mediocre
  - 0.50 = Miserable
  - < 0.50 = Unacceptable

    In this case, **KMO > 0.5**, so the data are adequate for factor analysis.

- **Bartlett's Test of Sphericity:**

  Tests whether the correlation matrix is significantly different from an identity matrix.

  A *p-value < 0.05* indicates significant correlations suitable for factor extraction.

  Here, **Sig. = 0.023**, so the test is significant — factor analysis is appropriate.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.635
Bartlett's Test of Sphericity	Approx. Chi-Square	170.757
	df	136
	Sig.	.023

KMO and Bartlett's Test is descripive statistics in the factor dialog box. The KMO value is different than .5, indicating that partial correlations among variables are small and that common factors are reliable. The value in Bartlett's Test is .023, which is lower than .05, hence the null hypothesis is destroyed. The null hypothesis states that the correlation matrix is an identity matrix. Together these results confirm that factor analysis is appropriate.

Step 4: Factor Extraction

- In the **Factor Analysis** window, click **Extraction**.
- Choose **Principal Component Analysis (PCA)** as the method.
- Tick **Scree plot** and **Unrotated factor solution**.
- Under **Extract**, select "Eigenvalues greater than 1".
- Click **Continue** → **OK**.

SPSS will display a **Total Variance Explained** table and **Scree Plot**.

Interpretation:

- Before extraction: 17 linear components exist (equal to number of variables).
- After extraction: Only factors with eigenvalues > 1 are retained.

  In this example, **8 factors** were extracted.

The first eight factors explain the following variances respectively:

9.604%, 8.808%, 8.229%, 8.037%, 7.956%, 7.862%, 7.803%, and 7.176% of total variance.

Together, these factors explain approximately **65% of total variance**, which is satisfactory.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.890	11.115	11.115	1.890	11.115	11.115	1.633	9.604	9.604
2	1.743	10.252	21.367	1.743	10.252	21.367	1.497	8.808	18.412
3	1.568	9.224	30.591	1.568	9.224	30.591	1.399	8.229	26.641
4	1.314	7.731	38.323	1.314	7.731	38.323	1.366	8.037	34.678
5	1.279	7.524	45.847	1.279	7.524	45.847	1.353	7.956	42.635
6	1.159	6.819	52.666	1.159	6.819	52.666	1.336	7.862	50.496
7	1.116	6.565	59.231	1.116	6.565	59.231	1.327	7.803	58.299
8	1.061	6.244	65.475	1.061	6.244	65.475	1.220	7.176	65.475
9	.956	5.621	71.096						
10	.851	5.008	76.104						
11	.780	4.590	80.694						
12	.744	4.376	85.070						
13	.680	3.998	89.067						
14	.573	3.370	92.437						
15	.505	2.970	95.407						
16	.450	2.648	98.055						
17	.331	1.945	100.000						

Extraction Method: Principal Component Analysis.

On the “Extraction” page, “Principal Component Analysis” is the extraction method selected, precisely because it seeks to maximize variance with the least number of components. “Eigenvalues greater than 1” are selected as the determining index for how many factors to retain. Additionally, a Scree plot is requested to visualize the eigenvalue distribution.

SPSS calculates 17 components (one for each variable) but, for the analysis, considers only components with eigenvalues greater than 1. In this case, there were 8 components which met this criterion. The “Total Variance Explained” table revealed that these 8 factors accounted for approximately 65 percent of the variance in the dataset. More specifically, the factors accounted for 9.604, 8.808, 8.229, 8.037, 7.956, 7.862, 7.803, and 7.176 percentages of variance in the dataset. The gradual decrease in each component's variance contribution, along with a distinctive elbow after 8 components in the Scree plot, supported a more parsimonious 8 factor solution.

Step 5: Factor Rotation

- Rotation helps make interpretation easier by maximizing high loadings and minimizing low ones.
- Click **Rotation → Varimax (orthogonal rotation)**.
- Tick **Rotated solution → Loading plot**.
- Click **Continue → OK**.

SPSS provides the **Rotated Component Matrix**, which displays correlations between each variable and the extracted factors after rotation.

Rotated Component Matrix

	Component							
	1	2	3	4	5	6	7	8
Affordability / Purchasing Power	.772							
Self Image	.640				.413			
Improving Life-Style	-.449					-.438		
Eco Friendly		.650						
Health & Safety Issue		.578						
Usability / Ease of Operation			.859					
Comfort		.466	.495					
Space Issue				-.812				
Suitable Operating Condition				.572				
Durability					.765			
Financing / Credit Facility					.672			
Demonstration Effect						.754		
Brand	.427					-.550		
After Sales Service							-.725	
Maintenance Cost							.645	
Exchange offers / Promotional offers / Schemes								.829
Technology		-.445						.494

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 31 iterations.

When extracted, the individual components tend to download many variables at a time, making the interpretation process a bit convoluted. In SPSS, the pattern has a rotation "...variance.." technique which rotation redistributes the variance while keeping the variance partitioned constant. The most known is Varimax, which is a type of orthogonal rotation that maximally raises the largest loadings and lowers the smallest loadings while conserving the orthogonality of the factors being rotated. `
The researcher looks for and clicks "Rotation" chooses Varimax and then asks for the loading plot and the rotated solution. Once completing the process, SPSS gives a Rotated Component Matrix displaying the rotation of the factors and component variables for each factor.

8.20.8 Interpreting the SPSS Output

Correlation Matrix: Within each pair of variables, their correlation coefficients are captured in the first table. In the correlation study, coefficients, in pare, were low to moderate, meaning that while the variables are related none of them can be said to be redundant. This is optimal for factor analysis, which aims for moderate inter-correlation and independence across clusters.

KMO and Bartlett's Test: The KMO value greater than 0.5 implies representative of the sample, in this case the redundancy of the variables is of interest and the KMO measurable value of the sample Bartlett's test that correlates with the factor is significant (Sig = 0.023). The withdrawing assumption of sphericity is in this case positively violated which means there is some redundancy and the factor structures can summarize the data.

Communalities Table: The greater the value of each of these variables, the more variance explained by factors that were retained. Closer to one are values which are estimated better by the factor solution. For instance, “Comfort” and “Usability” may have about 0.70 which means 70 of their variance is explained by the common factors, while a low communality of 0.40 means the variable is less useful in the spatial model.

Total Variance Explained: These tables captures the eigenvalues before and after extraction and after rotation for. before extraction, the total variance is the same as the number of variables, in this case 17. After extraction, only 8 factors above one in eigenvalue are retained.

Cumulative variance that approaches 65 percent is indicative that the factor model captures almost two thirds of the information present in the original variables, a fulfillment in the context of social science.

The Scree plot features each eigenvalue displayed on the Y-axis, while the respective numbered components of the model are displayed on the X-axis. The curve drops steeply and then levels off, and the place where the curve levels off, in this case after the 8th point, is the inflection point. This is where relevant and irrelevant factors are divided. This supports the decision of keeping 8 factors.

After Varimax rotation, the variables in the Rotated Component Matrix are organized with high loadings to one factor and low loadings on the others. The proportion of variance a certain factor predicts in that variable is represented by the square of the loading. These values are called loadings, and in social science studies, values above 0.50 are viewed as significant. The researcher then clusters the variables under each factor based on shared themes and assigns them a label.

8.20.9 Interpreting the HIP Factors

From the rotated matrix, eight interpretable factors emerged, each representing a coherent dimension of consumer preference:

1. **Purchase Affordability:** Variables such as affordability and purchasing power grouped together, representing the economic feasibility of buying the product.
2. **Socio-Personal Value:** Self-image and lifestyle improvement loaded on the same factor, indicating the symbolic or status-oriented aspects of purchase.
3. **Socio-Health Consciousness:** Eco-friendliness and health-safety concerns formed another cluster, showing environmental and personal-well-being motivations.
4. **Adoptability or User Friendliness:** Comfort and ease of operation loaded together, defining functional convenience.
5. **Feasibility of Implementation:** Space requirement and suitable operating conditions grouped under this factor, indicating practical constraints in usage.

6. **Shifting to Next Level:** Exchange schemes, promotional offers, and technology advancement loaded together, pointing to modernisation and upgrade incentives.
7. **Tenure of Impact:** Durability and financing facility combined to reflect long-term value and payment flexibility.
8. **Burden After Purchase:** Maintenance cost and after-sales service constituted the final factor, representing post-purchase responsibility.

Together these factors provide a concise yet comprehensive representation of consumer evaluation criteria for high-involvement products.

8.20.10 Understanding Factor Scores

SPSS can determine the factor scores of the respondents for each factor extracted which are called factor scores. These scores are linear amalgams of the observable quantities, and observed variable score coefficients determine the score corresponding to each factor. These scores can be calculated by going to the "Scores → Save as Variables" option in the Factor Analysis menu. Such new variables can be used for further statistical study, for example, to cluster respondents with correlated factor profiles, or regress the intention to purchase on these underlying dimensions. Hence, SPSS assigns numerical values to factor scores. Factoring becomes easy with SPSS as the modelling gives accurate results.

8.20.11 Theoretical Interpretation and Discussion

The implications of the HIP factor solution are equally theoretical as they are practical. They indicate that even evaluation of complex high involvement products is not random as previously thought, but rather organized around a stark set of latent dimensional variables that blend psychological and functional concerns. Affordability and self-expression are primary motives, as they are confirmed 'felt' as self-evaluative socio-psychological operative concerns. In addition, the ever-growing socio-health factor eco-consciousness, and matter-of-fact principle indicate an awareness towards sustainable usage and consumption. The set of factor technology and promotional appeal speaks of an advanced market that appreciates dynamic desire, while the set of factors usability and leaseability pays attention to the practical side of ownership. Finally, satisfaction is reflected through the set of factors that durability and ease of maintenance entail.

The structure as outlined above falls neatly within consumer decision-making theory apon which model is hinged on rational and emotional components. From a managerial viewpoint, each of the eight core dimensions simplifies and thus assists the marketer in developing specific feature design, communication plans, and consumer service strategies that are optimal for consumer segmentation.

For example, price-sensitive customers are likely to respond to communication strategies that emphasize durability and low financing. In contrast, customers who are either aspirational or eco-conscious are likely to respond to communication strategies that emphasize low technological sophistication and high eco-friendliness.

8.20.12 Reliability and Validity Considerations

The next step is to check reliability. Since items grouped within a factor are assumed to measure the same thing, Cronbach's Alpha can be used to assess internal consistency. A score exceeding 0.70 is regarded as satisfactory. A factor that is less reliable may contain more diverse items, or may be lacking in the number of items.

At the same time, convergent and discriminant validity should be analyzed as well. Convergent validity is when within a factor, the items make sense, and discriminant validity is when different factors are attempting to measure different constructs. The HIP factors, however, do meet these criteria reasonably well, because the thematic meaning of each cluster is quite clear.

8.20.12.1 Advantages and Limitations

There are distinct benefits of conducting factor analysis. It provides simplification of difficult data sets, revealing structures that are obscured, and each metric error is markedly diminished by concentrating only on common variance. It also improves subsequent analyses by resolving multicollinearity and supplying composite factor scores. Despite these benefits, there are distinct shortcomings as well. The researcher is able to exercise some control on which components are discriminatory and factor naming is subject to their discretion. Sample size, data scaling, and the method of extraction are also very sensitive. In addition to this, factor analysis is purely descriptive and only provides relationships, devoid of any causation.

8.21.1 Practical Guidelines for Researchers

As in any case, submitting any form of claim for academic publication or proceeding with any form of academic research or writing requires the same level of rigorous discipline, or more. In terms of practical application, rigid frameworks such as orthogonal Varimax or oblique any-set equivalents serve the purpose for both sides, simplified extraction or any level of correlated restriction. Furthermore, the other side of the Altman shades, the ethical beware side, must also be fully served: runs, reserving rotations, alternating retained number factors; executed in virtue of an imagination, are superb techniques designed for verification, stability, or consistency.

It Goes Without Saying That a meticulously designed framework for documenting the decisions made on the factors of principal with submission or the documentation with algebraic transcription then any possible number of other iterations for each step of such submission, is de facto standard. Reporting each step of the KMO tests with a de facto standard set, such as rotational load, is a step toward unambiguous transparency.

8.21.2 Implications of the HIP Findings

The HIP study's eight-factor solution illustrates how factor analysis synthesizes a comprehensive list of consumer attributes into practical strategies. Firms dealing with high-involvement product types can anchor their product strategies around these factors. For example, "Purchase Affordability" along with "Tenure of Impact" sets the guiding framework for pricing and financing policies, whereas "Adoptability" with "Burden After Purchase" lays out the guiding tenets for the service and product quality offered. "Socio-Personal Value" and "Socio-Health Consciousness" define emotionally resonant branding collateral, and "Shifting to Next Level" suggests strategies for innovation and associated communication. By calculating individual factor scores, marketers are able to perform granular market segmentation and offer tailored propositions.

8.21.3 Integration with Broader Statistical Modelling

You can utilize the extracted factors in more sophisticated multivariate analyses as either independent or dependent variables. For example, one can develop a regression model which predicts the overall intention to purchase using the factor scores as independent variables, or perform a cluster analysis which segments consumers using the dominant factor scores as variables. As such, factor analysis serves as a gateway technique which narrows down the dataset for further advanced modeling.

Conclusion

In multivariate research, factor analysis serves as the foundational method for transforming the multifaceted observable phenomena into latent variables. It serves as a 'bridge' to descriptive complexity and theoretical clarity. Take for example the High-Involvement Product study. Out of the 17 dispersed variables, eight emerged that captured the thought process of the consumers' decisions. These conclusions provided practical information for management while also advancing academic knowledge.

In other words, factor analysis helps researchers to order the disorder. It summarizes, clarifies, structures, and surfaces hidden relationships. When factor analysis is sophisticatedly done, along the assumptions, diagnostics, and interpretations of the analysis, it serves as a powerful tool for discovering new insightful scientific information, as well as for strategic decision-making.

8.22.1 Chi-Square Test of Independence

The analysis of two or more variables' connections is very commonplace in social science research, business, and market research. Would the consumer's income correlate with the consumer's preferred city of residence? Is there a correlation between gender and the purchase of a product? Is there a relationship between a person's educational qualification and the use of digital payments? The variables become more interesting when both of them are of a categorical nature which means data of both variables are non numeric. The best statistical technique in this case is the Chi-Square Test of Independence.

One highly recognized nonparametric statistical test is the Chi-Square test of independence with the symbol χ^2. It concentrates on the case, which is the focal point of the current analysis where there are observed values, and there are those values within each category where there are no values. The test of independence tries to see the extent to which the pattern, which is observed, differs from the one that is expected in relation to the two variables to ascertain whether the relationship is due to chance or actual factual elements.

8.22.2 Concept and Logic of the Test

The Chi-Square test of independence concerns the cross classification or two way contigency table of two variables. One of the variables makes the rows in the table and the other the columns. Each cell contains the number of cases that fall within that specific combination of categories and that number of cases.

The null hypothesis, or hypothesis zero (H_0), is that the two variables in question are independent; that is, the second variables' distribution depends in no way on the other one. The other hypothesis, or hypothesis one (H_1), proposes that the two are interdependent, meaning there is some form of connection in between.
The test statistic is calculated using the formula:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

where O_i is the observed frequency in each cell, and E_i is the expected frequency calculated under the assumption of independence. The expected frequency for a given cell is computed as:

$$E_i = \frac{\text{(Row Total)(Column Total)}}{\text{Grand Total}}$$

If the observed and expected frequencies differ substantially, the value of χ^2 will be large, leading to rejection of the null hypothesis.

The calculated χ^2 value is compared with the **critical value** from the Chi-Square distribution table at a chosen level of significance (commonly 0.05). The degrees of freedom (df) are given by:

$$df = (r - 1)(c - 1)$$

where *r* is the number of rows and *c* is the number of columns in the contingency table.

Assumptions of the Chi-Square Test

1. **Nature of Variables:** Both variables must be categorical—nominal or ordinal.
2. **Random Sampling:** The sample should be randomly drawn from the population.
3. **Independence of Observations:** Each observation should belong to one and only one category.
4. **Expected Frequency Requirement:** Ideally, no expected frequency should be less than 5; otherwise, categories may need to be combined.
5. **Size of Sample:** A larger sample size provides more reliable results since the Chi-Square test relies on approximation to the Chi-Square distribution.

8.22.3 Example Case: City and Income Status

As an example, the Cashless.sav data file illustrates income stratified as High, Medium, and Low and the data was analyzed using SPSS. A research project carried out across the three cities Ghaziabad, Delhi and Meerut with 1,056 respondents. It attempts at ascertaining whether the income status is a correlate of the city of residing.

The cross-tabulation of City and Income Status is presented below:

City * Income Status Crosstabulation

			Income Status			Total
			Poor	Middle Class	Rich	
City	Ghaziabad	Count	141	92	135	368
		% within City	38.3%	25.0%	36.7%	100.0%
	Hapur	Count	223	50	80	353
		% within City	63.2%	14.2%	22.7%	100.0%
	Meerut	Count	242	59	74	375
		% within City	64.5%	15.7%	19.7%	100.0%
Total		Count	606	201	289	1096
		% within City	55.3%	18.3%	26.4%	100.0%

This table displays the **observed frequencies**—that is, how many respondents fall into each combination of city and income level.

The question is whether the income status of respondents is independent of the city they belong to, or whether certain cities have a disproportionate representation of high- or low-income individuals.

Hypotheses Formulation

- **Null Hypothesis (H_0):** City and Income Status are independent. There is no significant association between the city of residence and income category.
- **Alternative Hypothesis (H_1):** City and Income Status are dependent. There is a significant relationship between city of residence and income category.

8.22.4 Steps in SPSS

The test can be easily conducted using the following SPSS procedure:

1. **Open the Dataset:** Load the file *Cashless.sav* into SPSS.
2. **Navigate to Crosstabs:** From the main menu, choose *Analyze → Descriptive Statistics → Crosstabs*.
3. **Assign Variables:** Place "City" in the *Row(s)* box and "Income Status" in the *Column(s)* box.
4. **Select Statistics:** Click *Statistics*, tick *Chi-Square*, and then click *Continue*.
5. **Select Display Options:** Under *Cells*, choose *Observed* and *Expected* to display both frequency types.
6. **Run the Analysis:** Click *OK* to generate the output.

SPSS will produce two tables: the **Crosstabulation Table** (showing observed and expected counts) and the **Chi-Square Tests Table** (showing test statistics, degrees of freedom, and significance level).

8.22.5 Understanding the SPSS Output

Crosstabulation Table: SPSS displays a table similar to the one above, but each cell contains both observed and expected frequencies. The expected frequency is the number of cases one would expect if the two variables are independent. The discrepancies of observed and expected values are the basis of a potential association.

Take the case of the respondents, and let's say the expected count for Delhites with a high income is 189.7. The observed count is 214. The surplus shifts the χ^2 value much higher. Thus, there is a positive association between residing in Delhi and having a high income.

Chi-Square Tests Table: The main output table includes:

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	65.537[a]	4	.000
Likelihood Ratio	65.908	4	.000
Linear-by-Linear Association	46.954	1	.000
N of Valid Cases	1096		

The 4 degrees of freedom with a Chi Square value of 65.597 has a p value of 0.000, which is less than 0.05. Hence, the null hypothesis is accepted.

There is, thus, a noteworthy correlation between City and Income Status. To put it differently, respondents' income varies among cities, suggesting the two variables are not independent.

8.22.6 Interpretation of Results:

Finding some evidence to reject the null hypothesis means that the chance of coming across such distribution due to random chance is infinitesimally small. Therefore, the respondents' income profile is influenced by the city of residence. For example, the data might show Delhi exhibiting a higher number of respondents belonging to the higher income bracket compared to what is expected, while Meerut shows a higher concentration of respondents belonging to the lower income bracket. Such observations indicate that there are some links between the economic status and the amount of development, industrialisation, and employment opportunities that a city presents.

Although the Chi-Square test reveals some link between variables, to be proper, it measures association and not causation. It tells us that two variables are linked, not which of the two causes the other. Other factors such as education, cost of living, and occupational structure may explain the association which is observed.

8.22.7 Theoretical Explanation of Chi-Square Value

The χ^2 statistic quantifies the discrepancy between observed and expected frequencies. Each cell contributes to the total χ^2 value according to how far its observed count deviates from expectation, scaled by the expected count. Cells with large deviations—either surpluses or deficits—contribute more to the statistic.

For example, suppose the observed frequency for high-income individuals in Delhi is 214, while the expected frequency under independence is 189.7. The cell contribution would be:

$$\frac{(214-189.7)^2}{189.7} = 3.12$$

Summing similar contributions over all 9 cells (3 cities × 3 income categories) yields the overall χ^2 of 65.597.

A high χ^2 value means that many cells deviate substantially from expected frequencies, pointing toward dependence. A low χ^2 implies that observed frequencies closely match expectations, supporting independence.

Degrees of Freedom and Critical Value

In this example, with 3 cities and 3 income categories,

$$df = (3-1)(3-1) = 4.$$

The critical value of χ^2 at 0.05 level of significance for 4 degrees of freedom is approximately **9.49**. The calculated value, 65.597, far exceeds this threshold. Therefore, the result is highly significant, confirming dependence between the variables.

Strength and Direction of Association

While the Chi-Square test reveals the existence of an association, it does not directly measure its strength or direction. To assess strength, researchers may compute additional indices such as **Cramer's V** or **Phi coefficient**, which are also provided in SPSS under the *Symmetric Measures* option within the Crosstabs dialog box.

Cramer's V is calculated as:

$$V = \sqrt{\frac{\chi^2}{n(k-1)}}$$

where *n* is the total number of cases and *k* is the smaller of the number of rows or columns. Values of V range from 0 (no association) to 1 (perfect association). In practical terms, 0.1 indicates a weak relationship, 0.3 a moderate one, and 0.5 or above a strong association. In our example, with χ^2 = 65.6, n = 1056, and k = 3, V ≈ 0.25, suggesting a moderate relationship between city and income status.

8.22.8 Limitations of the Chi-Square Test

Regardless of its use, the Chi-square test does have some shortcomings. First, it is disproportionate to the sample size: extremely large samples can make even the most … of the most minimal distinctions seem significant. Second, the test is highly untrustworthy if the expected frequencies are small—or below five. In these …, it is best to combine the categories, or use an alternative such as Fisher's Exact Test. Third, it can only be done to … data: all the quantitative variables must be somehow categorized to be able to Chi Square them. Finally, it does not explain causality: association does not mean one variable influences the other.

8.22.9 Practical Implications

Within management and social research, the ability to perform case studies and apply the Chi-Square test of independence is invaluable. It can be used to evaluate whether purchase preferences shift within various age groups and regions in a supplied market. In Human Resource management, it can assist in answering whether employee satisfaction is based on a favorable department or experience. In the field of education, it could examine whether the type of schooling affects the students' performance on an exam. It is a versatile test. It is used for non-parametric surveys that only use categorical data rather than numerical data and digitized information.

The Cashless study`s insight that city income status is significantly clustered could be used to design a targeted, differentiated approach to marketing and other policy instruments based on city. For example, the market for cashless payment services in Delhi may be more receptive to targeted premium offers due to the concentration of higher-income customers, while the rest of the cities may need subsidized offers or educational campaigns to encourage the adoption of digital payment services.

8.22.10 Extensions of the Test

Variants of the Chi-Square test include:

- **Goodness-of-Fit Test:** Determines whether a single categorical variable follows a specified distribution (for example, testing if gender distribution in a sample matches population proportions).
- **Test for Homogeneity:** Compares whether different populations have the same distribution of a categorical variable.
- **McNemar Test:** Used for paired categorical data, such as before-and-after studies.

All these tests rely on the same underlying principle of comparing observed and expected frequencies through the χ^2 statistic.

Summary and Conclusion

The relationships between discrete variables can often be analyzed through the Chi-square tests for independence. We are concerned with the null hypothesis which states the two variables are categorically associated. The chi-square tests the discrepancy between observed and expected 'in case null holds' frequencies assuming independence for independence and determines its significance. The sensitivity of the test requires only categorical data and minimum a priori conditions.

Out of the variables 'City' and 'Income Status', the computed value of chi square is 65.597 with p value=0.000. This leads to the acceptance of the alternative hypothesis which states the proportion of people residing in various cities and their income levels is not the same. This suggests the variable income is geo-spatially clustered and implies in a city there are people who belong to both higher and lower income groups. These factors can be aligned to aid in strategic regional development as well as marketing and modulated for policy formulation.

The Chi Square test underlines the value of observation and data collection to infer the relationships between variables of a given phenomenon. It is a foundational component in the social and managerial data analysis.

8.23.1 Spearman's Rank Correlation

There are a number of perspectives and fields of work to which it would be vital to find a response to the question of how two different judgments or ratings of the same thing correlate. Take, for example, two physicians assessing the same set of patients. Each of them has the opportunity to arrive at a unique conclusion. In the same manner, two judges in the business world might assign a distinctive place to a number of rival companies based on their quality or preference. In the world of academics, it is not uncommon for two educators to evaluate a set of learners and assign grades. Then the question becomes: how similar are the evaluations to which criterion one and two landed on?

To answer the question given in the previous paragraph, there exists a statistical metric called the Spearman's Rank Correlation Coefficient, or simply the Rank Correlation Coefficient. This statistical metric is better known with the Greek symbol ρ (rho). This technique works wonderfully to find the strength and the direction of an association of two reordered variables. In contrast to the Pearson's correlation which works with numbers, assumes a linear and normal distribution, and "computes" correlations, Spearman's correlation works with simple ordered ranks, therefore can assume ordinal data and non-linear relationships.

8.23.2 Concept and Rationale

Spearman's Rank Correlation happens to not be a parametric test which implies the test does not rely on the assumptions of the data's distribution. It examines the extent to which the relationship between two random variable can be represented by a monotonic function, which is a function that either a. increases or b. decreases and none of which happen at a constant.

In in the event two ranking set are exactly the same, Spearman's correlation coefficient is +1 which is the maximum degree and perfect positive correlation. If the ranking sets are completely opposite, the coefficient is -1 and perfect negative correlation. If a value is close to 0, it lacks relationship to the rankings.

In principle, the approach takes a set of raw scores, transforms them into ranks and tries to determine the extent to which the two sets of ranks differ. The less the set of ranks differ, the greater the degree of agreement between the two sets them.

Formula

For a sample of *n* observations, where did_idi represents the difference between the ranks of the two variables for each observation, the coefficient is calculated as:

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

Where:

- ρ = Spearman's Rank Correlation Coefficient
- d_i = Difference between the ranks assigned to each item
- n = Number of items ranked

The coefficient ρ always lies between –1 and +1.

- $\rho = +1$: Perfect positive correlation (ranks match exactly)
- $\rho = -1$: Perfect negative correlation (ranks are in reverse order)
- $\rho = 0$: No correlation between ranks

When ties occur (two or more identical values), average ranks are assigned, and a slightly adjusted formula is used by statistical software such as SPSS.

Example: Diagnoses by Two Doctors

Consider an example involving **two doctors**, Doctor A and Doctor B, who independently assess the severity of illness in **eight patients** suffering from similar symptoms. Each doctor ranks the patients from 1 (best condition) to 8 (worst condition). The question is whether the two doctors agree in their evaluations.

This example is analysed from the data file Spearman Rank.sav using SPSS. The objective is to determine the extent of correlation between the two sets of rankings.

Patient	**Rank by Doctor A**	**Rank by Doctor B**	**Difference (d)**	**d^2**
1	1	2	–1	1
2	2	1	1	1
3	3	3	0	0

4	4	5	–1	1
5	5	4	1	1
6	6	6	0	0
7	7	8	–1	1
8	8	7	1	1
Total	–	–	–	**6**

Applying the formula:

$$\rho = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(6)}{8(64-1)} = 1 - \frac{36}{504} = 1 - 0.0714 = 0.9286$$

Thus, the Spearman's Rank Correlation Coefficient is approximately **0.93**, indicating a very high positive correlation. The two doctors' rankings are in strong agreement.

Hypothesis Setting

The test can also be expressed in hypothesis-testing form:

- **Null Hypothesis (H_0):** There is no correlation between the rankings of Doctor A and Doctor B. (ρ = 0)
- **Alternative Hypothesis (H_1):** There is a significant correlation between the rankings of Doctor A and Doctor B. (ρ ≠ 0)

The significance of ρ can be tested using the *t*-statistic:

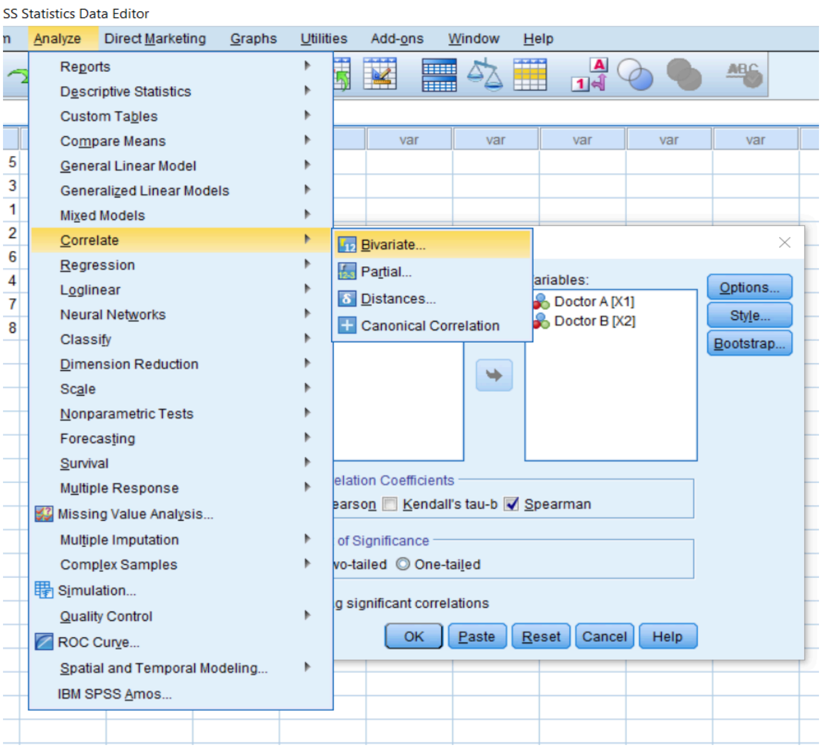
$$t = \rho\sqrt{\frac{n-2}{1-\rho^2}}$$

which follows a *t*-distribution with *(n – 2)* degrees of freedom. Alternatively, SPSS provides the exact significance (p-value) automatically.

8.23.3 Performing Spearman's Rank Correlation in SPSS

SPSS simplifies the process through a few straightforward steps:

1. **Open Dataset:** Load the file Spearman Rank.sav into SPSS.
2. **Navigate to Correlation Menu:** Click *Analyze → Correlate → Bivariate*.
3. **Select Variables:** Move "Doctor A" and "Doctor B" into the *Variables* box.
4. **Choose Correlation Coefficient:** In the dialog box, select *Spearman* under the "Correlation Coefficients" section (uncheck Pearson if selected).
5. **Set Options:** Tick "Two-tailed" for testing the significance and ensure "Flag significant correlations" is checked.
6. **Run the Analysis:** Click *OK*. SPSS generates an output table showing Spearman's rho, the correlation coefficient, sample size, and significance (p-value).



8.23.4 SPSS Output and Interpretation

The output table is as follows:

Correlations

			Doctor A	Doctor B
Spearman's rho	Doctor A	Correlation Coefficient	1.000	.850**
		Sig. (2-tailed)	.	.007
		N	8	8
	Doctor B	Correlation Coefficient	.850**	1.000
		Sig. (2-tailed)	.007	.
		N	8	8

**. Correlation is significant at the 0.01 level (2-tailed).

From the table:

- **ρ = 0.850**, indicating a **strong positive correlation** between the two doctors' rankings.
- **Significance (p-value) = 0.007**, which is **less than 0.05**.

Therefore, the null hypothesis is rejected, and it is concluded that there is a statistically significant positive relationship between the two sets of rankings. The doctors' evaluations are highly consistent, with very little disagreement.

Because $p < 0.01$, the result is even significant at the 1% level, confirming that the observed correlation is unlikely to have arisen by chance.

Interpretation in Context

A correlation of 0.850 suggests that whenever Doctor A classifies a patient as having a severe condition, Doctor B is likely to classify him or her in the same way. The high level of concordance indicates that the two physicians possess the same thresholds of criteria for diagnosis and the same reasoning for conclusions. Applied, this means the assessment is dependable and consistent among practitioners.

Consensus from inter-rater correlation is essential in health, education, and psychological research, which relies heavily on subjective assessment and judgment. Such high correlations assure the presence of very little range of contour of error that stems from subjective biases and differences in understanding.

8.23.5 Understanding the Meaning of Correlation Strength

Spearman's rho is interpreted in much the same way as Pearson's correlation coefficient, but it applies to ranked data. The following general guidelines are widely accepted:

	Strength of Correlation	**Interpretation**
$\rho = \pm 1.00$	Perfect correlation	Rankings match exactly or are perfectly opposite
$\rho = \pm 0.80$ to ± 0.99	Very strong correlation	Rankings are highly consistent
$\rho = \pm 0.60$ to ± 0.79	Moderate to strong correlation	Rankings generally agree with some deviation
$\rho = \pm 0.40$ to ± 0.59	Moderate correlation	Partial agreement
$\rho = \pm 0.20$ to ± 0.39	Weak correlation	Limited similarity in rankings
$\rho < \pm 0.20$	Very weak or no correlation	Rankings are unrelated

In the present example, $\rho = 0.850$ falls into the "very strong positive correlation" category, reflecting near-unanimous agreement.

8.23.6 Relationship between Spearman and Pearson Correlations

In assessing correlation, Spearman's and Pearson's coefficients do, however, differ in scope as well as viewpoint. While Spearman's correlation deals with assessing the strength of a monotonic relationship between two ranked or ordinal variables, Pearson's correlation deals with assessing the degree of the linear relationship between two intervals or ratios of variables that follow a normal distribution.

In cases where distribution is non-normal, subjectively ranked, or is outlier prone, Spearman's coefficient is the most appropriate. However, if the variables in study are linear and continuous. Pearson's correlation will most likely provide a more accurate value. Fun fact, data that is perfectly monotonic will display identical Spearman and Pearson coefficient values.

8.23.7 Advantages of Spearman's Rank Correlation

Spearman's Rank Correlation, in comparison to other techniques, has specific advantages, notably in its application to non-parametric or ordinal data. In its application, not many of its users have to deal with its major drawback, which is its parametric nature. Spearman's correlation does not consider how the data was derived, be it subjective or not, and even, how data is ranked and or ranked on the Opion's Ladder.

It is also user friendly in computation. 'It considers only differences between positions in the rank orders rather than the actual values of the ranks, the computation is usually easy and straightforward. This easiness in computation is crucial especially in small sample research which in many cases is the norm in most research studies.

Another drawback is the correlation method which is robust to even the most dire. '" This is to say, with the ranked methods, it is also very easy to discern outliers. In correlation methods, the presence of outliers has been shown to have detrimental effect on the final result. Using ranked data, outliers retain their position and shifts many results, assuring the results are much more trustworthy when the data contains garnet strong or extremely raw.

Another major tool would be its flexible use.', `This has been shown to work even when only one of the two variables is ordinal in discipline of psychology and most especially education and even the other social studies.

It also identifies non-linear relationships. While Pearson's correlation only concerns itself with linear relationships, Spearman's can recognize trends which are monotonic—where one variable only increases, or only decreases, when another variable increases and lacks a non-linear relationship. The ability to pick up relationships that are non-linear enhances the utility and flexibility of Spearman's correlation.

Limitations

Although it is valuable, Spearman's Rank Correlation has many other weaknesses that need to be noted when analyzing results. One problem is the missing perception that comes when continuous data is changed and classified into ranks. This simplification leads to poorly constructed datasets and ultimately loss of precision which hides the real variation.

Another limitation is the assumption of monotonicity. This approach determines the extent to which one variable increases or decreases with respect to the other, and it totally ignores the non-monotonic discrepancies that may arise. This, therefore, leads to the lack of detection of complex variable relationships.

Further, Spearman’s correlation is still restricted to dyadic pairs of variables, thus cannot engage in sophisticated analyses of complex multiple interrelationships or control the impacts of other variables which is a common feature of intricate research frameworks.

Lastly, the existence of tied ranks poses a problem of computation and interpretation. When many observations are assigned the same rank, special adjustments are always needed which ultimately leads to a loss in the value of the correlation, and thus the ability to compute and interpret the correlation is diminished.

Finally, Spearman’s correlation does not imply causation. Correlation between two variables can be strong and significant, yet, still, not one variable is capable of inducing changes in the other. Correlation, and in this particular case, Spearman’s, is the only measure of how strong the two variables are and in what direction the relationship moves. It does not justify the reasons behind such a relationship.

8.23.8 Practical Significance

To use Spearman's correlation on data like these has value counter wherever ranking or evaluation is needed. For instance, in marketing, Spearman's correlation can quantify the correspondence of two evaluators on the same advertisement. In recruitment, it can measure the scoring of candidate rankings by two interviewers. In educational settings, it ascertains whether or not examiners are consistent in their evaluations of students. In medicine, as in our example, it assesses the level of agreement in diagnosis between practitioners.

A high correlational coefficient reflects reliability and objectivity, while a low correlational coefficient indicates the possibility of factors needing more stringent value evaluation, training, or improvement of standard methods.

Reporting the Findings

When reporting Spearman’s rank correlation results in a research report or thesis, it is advisable to include the following elements:

- Statement of purpose: e.g., “To determine the degree of agreement between two doctors’ diagnoses.”
- Description of data: number of cases and nature of ranking.
- Computed value of ρ.

- Significance level (p-value).
- Interpretation of strength and direction.

Example write-up:

A Spearman's rank correlation analysis was conducted to examine the relationship between the diagnoses of Doctor A and Doctor B for eight patients. The results revealed a strong positive correlation, $\rho = 0.850$, $p = 0.007$. This indicates that the two doctors' evaluations were highly consistent, and the relationship was statistically significant at both the 5% and 1% levels.

Conclusion

Spearman's Rank Correlation focuses on the measurement of the relationship between two ranked variables. It is the most simple and robust. Its use is most advantageous in non-parametric situations where variables are at least ordinal or where the assumptions of linearity and normality do not hold. As in the illustration with two physicians above, a coefficient of 0.850 suggests a strong, positive correlation that is statistically significant; thus, their assessment of the diagnosis appears to align quite well.

The principle of the method is simple, and the beauty is that, instead of focusing on the entire data set, concentrating on relatively smaller sets of data helps in getting rid of measurement units, outliers, and other forms of statistical anomalies. In the fields of management, psychology, medicine and other disciplines that require subjective judgments, Spearman correlation is particularly popular because it gives a fair measurement of concordance or consistency amongst a group of assessors.

8.24.1 Mann–Whitney U Test

In some research settings, one of the objectives is to see if there is a difference between two groups on one or more outcome variables. Traditionally, such differences are assessed using the Independent Samples t Test. This approach has some assumptions such as the use of the t test dependent variable should be interval or ratio scale, the data is normally distributed within every group, and the variances are somewhat equal. However, assumptions are not always met.

If the data is not normally distributed, the sample sizes are small, or the dependent variable is ordinal or ranked, the t test is not appropriate. In such scenarios, the non-parametric alternative is the Mann Whitney U Test, or the Wilcoxon Rank-Sum Test.

The Mann Whitney U test is one of the most popular tests and is especially appropriate in behavioral and other management and marketing research where the data as a result of responses are rank or rating data which are not normally distributed.

.

8.24.2 Concept and Rationale

The Mann–Whitney U test's primary purpose is to find differences in distributions of two independent groups. Unlike in the t-test where means are compared, here the median ranks of the two groups are compared. It is a rank order test because all observations are assigned a rank, starting from the smallest up to the highest, and then it is determined which of the two groups has the higher ranks in the order of interest.

When calculating ranks, if the two populations are the same, then logically, the ranks are expected to have a similar distribution. Difference in distribution are plausible, however, if one group has dominantly higher ranks.

A primary example is where we assess the difference in the way male and female respondents rate a bunch of car advertisements. If no difference exists between the two groups, the ranks assigned to the men and women in the group of lower ranked ads should be almost equal. If ckements of all ads contribute then the men does not outperform the women consistently, thus supporting the idea suggesting that gender differences do exist in the evaluation of car ads.

8.24.3 Underlying Logic

The test begins by **pooling all observations** from both groups and assigning ranks from 1 (lowest) to N (highest). The sum of ranks is then computed for each group. If the groups are similar, their rank sums will be close. If one group tends to produce higher scores, its rank sum will be larger.

The Mann–Whitney U statistic for each group is calculated as:

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U_2 = n_1 n_2 - U_1$$

where

- n_1 = sample size of Group 1
- n_2 = sample size of Group 2
- R_1 = sum of ranks in Group 1
- U_1 and U_2 = Mann–Whitney statistics for each group.

The smaller of the two U values is taken as the test statistic. When sample sizes are large ($n > 20$), the U statistic can be converted into a **z-score**, allowing computation of a significance level (p-value).

8.24.4 Assumptions of the Mann–Whitney Test

Although it is non-parametric, the Mann–Whitney U test still requires certain basic conditions:

1. **Independent Samples:** The two groups must be independent of each other. Each subject appears in only one group.
2. **Ordinal or Continuous Dependent Variable:** The outcome should be at least ordinal, allowing ranking.
3. **Similarity of Shape:** The distributions of the two groups should have similar shapes; otherwise, differences may reflect shape rather than median.
4. **Random Sampling:** Data should represent random and independent observations from the population.

When these conditions are met, the test provides a robust alternative to the t-test.

Example: Evaluating Car Commercials

Let us consider the case of a research project in which audience reaction to three different car commercials is studied. There are 18 respondents, evenly distributed by gender and age. Each respondent views all three commercials and rates them on a scale of satisfaction. The goal is to find out whether there is a divergence in the evaluations made by male and female respondents.

In the case of each commercial, the first competing hypothesis rests on the assumption that the mean ratings given by men and women are the same.

H0: There is no difference between male and female ratings.

The dat file Mann-Whitney test.sav is used. Since the ratings are on an ordinal scale and the size of the population is small (N=18), in this case the Mann-Whitney is preferable to the t-test.

8.24.5 Steps in SPSS

The analysis can be performed easily in SPSS following these steps:

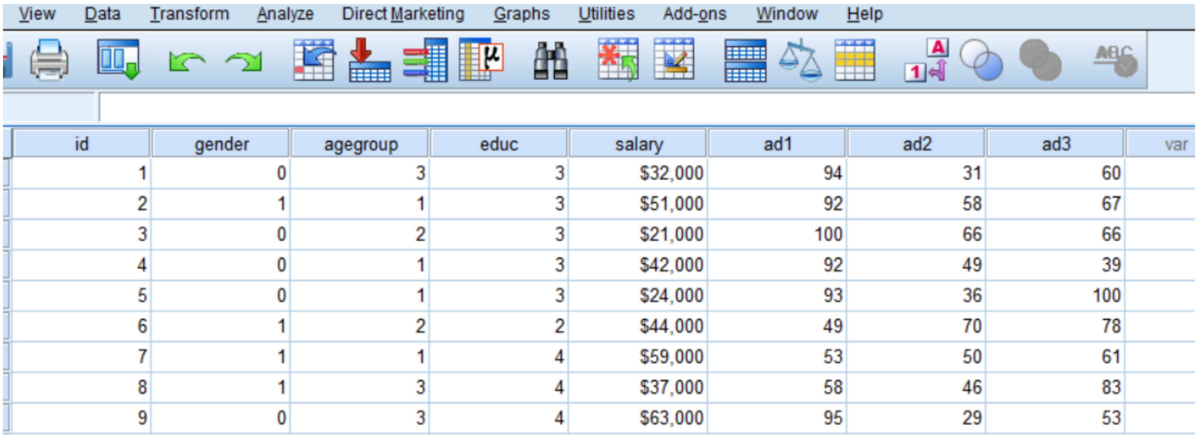
1. **Open the Dataset:** Load Mann-Whitney test.sav into SPSS.
2. **Navigate to the Test:** Click *Analyze → Nonparametric Tests → Legacy Dialogs → 2 Independent Samples.*

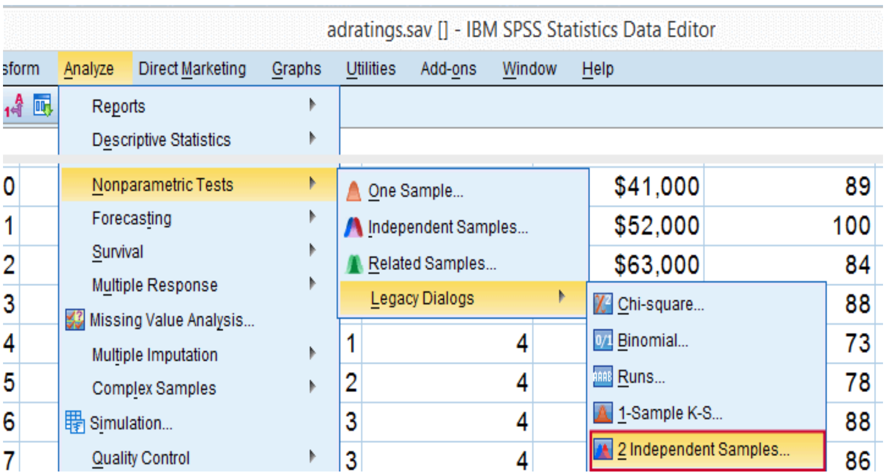
3. **Assign Variables:** Move the rating variable (for each car commercial) into the *Test Variable List* box.

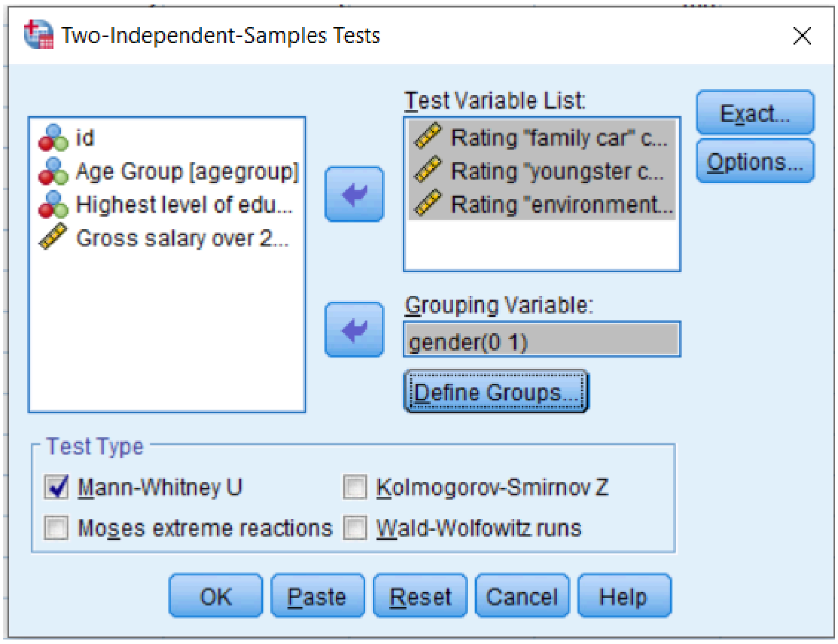
   Move *Gender* into the *Grouping Variable* box.

4. **Define Groups:** Click *Define Groups*, and specify codes for each group (e.g., 1 = Male, 2 = Female).
5. **Choose the Test Type:** Select *Mann-Whitney U* (this is checked by default).
6. **Run the Test:** Click *OK* to produce the output.

SPSS will automatically rank all scores, calculate mean ranks for each group, and provide the Mann–Whitney U statistic, z-value, and significance level.







SPSS Output and Interpretation

The SPSS output includes two key tables:

a. Ranks Table

Gender	**N**	**Mean Rank**	**Sum of Ranks**
Male	9	8.61	77.5
Female	9	10.39	93.5
Total	18	—	—

The "Mean Rank" column shows the average rank of ratings in each group. For the first commercial, *Family Car*, the mean rank for females (13.39) is noticeably higher than for males (11.56), indicating that women rated the ad more favourably.

For the other two commercials, the opposite pattern emerges—men's mean ranks are slightly higher (11.56 and 10.56), suggesting they were more positive toward those ads.

b. Test Statistics Table

Test	**U**	**Wilcoxon W**	**Z**	**Asymp. Sig. (2-tailed)**
Mann-Whitney U	22.5	67.5	–1.97	0.049

Here, the **Asymptotic Significance** value (p-value) of 0.049 is less than 0.05. Therefore, the null hypothesis is rejected for the *Family Car* commercial: men and women rated it significantly differently.

For the other commercials, if the p-value exceeds 0.05, we fail to reject H_0, meaning that gender does not significantly affect those ratings.

Statistical Decision

Based on the results:

- For **Commercial 1 (Family Car)**, $p = 0.049 < 0.05$ → **Significant difference.**
  Women's higher mean rank indicates they found this ad more appealing.
- For **Commercial 2 and Commercial 3**, $p > 0.05$ → **No significant difference.**
  Men and women's perceptions were roughly similar.

Thus, the study concludes that gender influences reaction only to the first advertisement.

8.25.1 Discussion of Findings

The results underscore the fact that consumer reaction to marketing material is considerably different between men and women. The difference in the advertisement enthusiasm on the Family Car might be indicative of women's stronger responsiveness toward the storytelling and family-focused narrative. On the contrary, men may be more appreciative of ads that center around technology and performance.

From a managerial perspective, such insights help marketers hone their strategies. If certain demographic groups have different responses, the strategy should be tailored to that group.

The Mann–Whitney test, on the other hand, does not only provide evidence of difference in central tendency. Coupled with the results, the test also demonstrates the strength of the results with respect to non-normal data distribution.

8.25.2 Relation to the Independent-Samples t-Test

While the independent-samples t-test compares group **means**, the Mann–Whitney test compares **median ranks**. When data are normally distributed and measured at interval or ratio level, both tests generally produce similar conclusions. However, the Mann–Whitney U test is preferred when:

- Data are **skewed** or contain **outliers**.
- The dependent variable is **ordinal** (e.g., Likert-scale ratings).
- **Sample sizes are small** (less than 25 per group).

In essence, the Mann–Whitney test is the non-parametric counterpart of the t-test and provides a more flexible approach for small or non-normal datasets.

Reporting the Results

When reporting the results of a Mann–Whitney test in a research paper or report, the standard format includes:

A Mann–Whitney U test was conducted to compare male and female ratings of the "Family Car" advertisement. Results indicated a significant difference between genders, $U = 22.5$, $z = -1.97$, $p = 0.049$. Female respondents ($M = 13.39$) provided higher ratings than male respondents ($M = 11.56$), suggesting that women viewed the advertisement more positively. No significant gender differences were observed for the other two commercials.

Such a narrative conveys both statistical and practical significance in clear and concise terms.

8.25.3 Advantages of the Mann–Whitney Test

Several benefits come with the use of the Mann-Whitney Test, particularly in scenarios where the theories of parameters cannot be achieved. A test's assumption is a model with missing and simplified features and the assumption of normality and equal variance is not the case. Such reality is far better from the tested condition in parameters of inactual parameters set,

The technique comes in handy in parameters other than nominal. Data is considered in the – scales other than the exact number values of the elements under observation- therefore, one of its major features.

The increased efficiency achieved with the technique is attributed to its focus on the rank, and not the actual number values. Hence it removes with ease any extreme outlier values, thus generating more accurate results.

Data that contains elements under the set number of 25 in each group is particularly suitable for the test. Hence, the technique becomes practical in scenarios where the data is limited.

The contrast is that test is the only one that is easy to explain. Hence it results in effortless understanding of group differences that measures the distribution values of the group. The group that the observation is made upon receives the ultra sense scales in relation to which other value set is lower.

8.25.4 Limitations

The Mann–Whitney Test despite being useful has several limitations.

Its power is less than that of the t-test when the data are in a parametric form which indicates that it would be poorer than the t-test in recognizing differences.

The Mann–Whitney Test is unable to deal with more than two independent samples and sets a ceiling to it. Such cases are better handled with a Kruskal–Wallis test.

Weaker in application, the Mann-Whitney Test has been argued to be far more sensitive to distribution shapes than other. If one is trying to make out the differences in central tendency, it is unwise to use the Mann-Whitney Test if the two groups differ greatly in distribution form.

Reliability is also an argument that has been raised with regards to the data that has been converted to ranks. Information is lost when the data is captured with a numerical distance.

Interpreting the results of the test in order to get an understanding of central tendencies, is far more useful. A group that is able to get an understanding of ranks more appreciates the differences relative and offers less information.

Extension and Related Tests

The Mann–Whitney test belongs to a family of rank-based non-parametric tests that includes:

- **Wilcoxon Signed-Rank Test:** Used for two **related** samples (paired observations).
- **Kruskal–Wallis Test:** Extends the Mann–Whitney test to **more than two independent groups**.
- **Friedman Test:** Non-parametric alternative to repeated-measures ANOVA.

All share the same principle: replacing raw scores with ranks and testing for differences between groups or conditions.

Practical Insights

In applied management research, non-parametric tests like Mann–Whitney are invaluable. For example:

- **Marketing Research:** Comparing customer satisfaction ratings between two regions.
- **Human Resources:** Evaluating whether male and female employees differ in engagement scores.
- **Healthcare Studies:** Comparing patient satisfaction across treatment groups when data are ordinal.
- **Education:** Testing whether students from two schools differ in perception of curriculum quality.

The test's simplicity and versatility make it an essential component of applied data analysis.

Conclusion

The Mann-Whitney U Test provides a neat alternative solution when one is unable to satisfy requirements for a t-test. Unlike normality dependent t-tests, it permits researchers to construct hypotheses about differences between groups by ranking data and comparing median ranks to reporting means.

In the car advertisement study, the test's results indicated that although men and women shared broadly similar evaluations for two of the television commercials, female respondents were much more favorable towards the 'Family Car' advertisement. This is a most welcome finding, both in terms of its statistical foundation and practical significance, and it exemplifies the manner in which non-parametric techniques can address important but hidden differences in human attitudes.

The Mann-Whitney U test remains central to the practical work of management researchers and non-parametric in character. It is uncommonly straightforward, exceedingly powerful and applicable in virtually all situations where the data are ordinal, skewed, or drawn from small independent samples, and, more importantly, it provides sound conclusions where other parametric approaches are inadequate.

8.26.1 Wilcoxon Signed-Ranks Test

In applied research, there are numerous instances in which a single cohort of participants is assessed twice, either before, during, and after the application of a treatment, or under a single condition, or in relation to two stimuli. The goal is to assess the extent to which these paired observations differ. Traditionally these kinds of observations are examined using the Paired-Samples t-Test which is premised on the assumption that the paired differences are normally distributed, paired score differences are on an interval or ratio measurement level.

Sample size is particularly a significant factor while test conditions are vastly different. In these instances, a t-test is assumed to fail, and in these specific situations the Wilcoxon Signed-Ranks Test is a suitable alternative. It test validity is a nonparametric stance on the notion of paired differences.

The Wilcoxon Signed Ranks test is one of the most highly validated nonparametric analyses of paired sets of data. It works when single sets of observations are paired in some way, for instance when there are repeated observations of the same individuals (or) on a paired set of individuals who share a similar set of attributes. The test shifts downward from the analysis of means, focusing on the ranks of change scores.

8.26.2 Concept and Rationale

The strength of the Wilcoxon Signed-Ranks Test is to find out whether the differences between pairs of observations has a median other than zero. To figure the median of the differences between paired observations, the test adopts a more sophisticated approach than just looking at the differences. It considers the size of the signs and how large the difference is – the rank.

First, the differences between the two paired observations (for example, the observations taken under two different conditions) is calculated. Differences which are zero are ignored. Differences are then calculated, and ranks are assigned. Each rank is assigned a sign depending on the second paired observations which is then classified as the first. If the two observations which are testified do not have a systematic difference, then the total sum of positive ranks should correspond to the total sum of negative ranks.

If most of the ranks are either positive or negative, then a large difference between the two dependent samples is indicated.

Hypotheses

The hypotheses for the Wilcoxon Signed-Ranks Test are:

- **Null Hypothesis (H_0):** The median difference between the paired observations is zero.
  (There is no significant difference between the two related conditions.)
- **Alternative Hypothesis (H_1):** The median difference between the paired observations is not zero.
  (There is a significant difference between the two related conditions.)

Mathematically,

$$H_0 : M_d = 0 \quad \text{versus} \quad H_1 : M_d \neq 0$$

where M_d is the population median of the difference scores.

The Test Statistic

The Wilcoxon Signed-Ranks Test statistic (W) is calculated as follows:

1. Compute the difference between paired scores.
2. $D_i = X_{2i} - X_{1i}$
3. Exclude pairs where Di=0
4. Rank the absolute values of the differences, assigning rank 1 to the smallest difference.
5. Restore the sign (positive or negative) of each difference to its corresponding rank.
6. Sum the ranks of positive differences to obtain W+W^+W+ and of negative differences to obtain W−W^-W−.
7. The smaller of W^+ and W^- serves as the **test statistic**.
   For large samples (n > 25), the test statistic is approximated by a normal distribution with mean and standard deviation:

$$\mu_W = \frac{n(n+1)}{4}, \quad \sigma_W = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

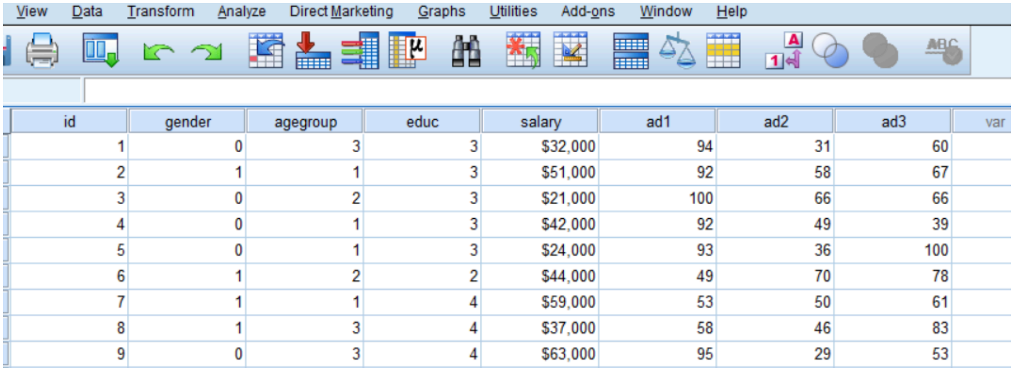
SPSS automatically converts this into a **z-value** and provides the corresponding **p-value**.

Example: Comparing Two Car Commercials

To illustrate, consider a study evaluating the effectiveness of three car commercials shown to the same group of 18 respondents. Each participant rated all three advertisements on a five-point scale. The research question was whether respondents judged *Commercial 1* and *Commercial 3* similarly.

Because the same individuals rated both commercials, the data are **paired**. However, with only 18 respondents and ratings that may not be normally distributed, the **Wilcoxon Signed-Ranks Test** is more appropriate than the paired-samples t-test.

The relevant dataset is stored as Wilcoxon test.sav.

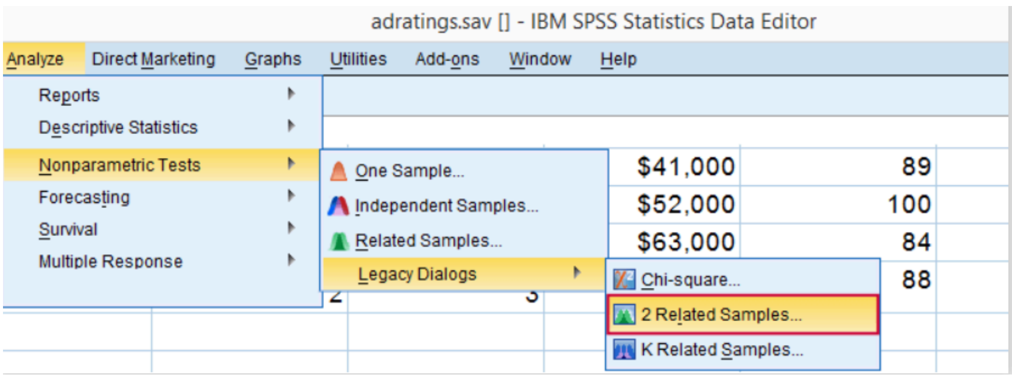


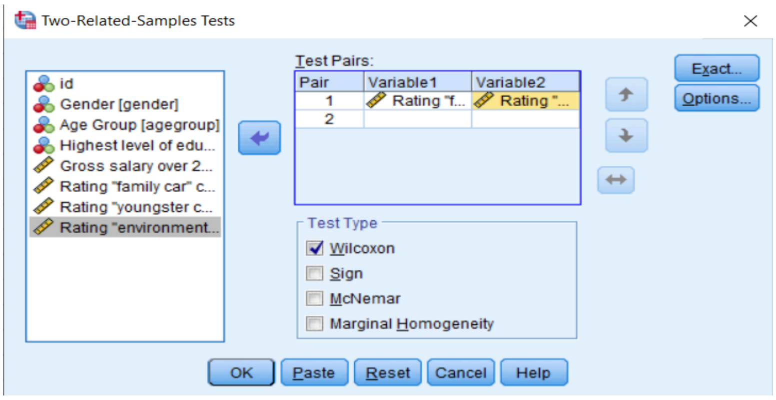
id	gender	agegroup	educ	salary	ad1	ad2	ad3	var
1	0	3	3	$32,000	94	31	60	
2	1	1	3	$51,000	92	58	67	
3	0	2	3	$21,000	100	66	66	
4	0	1	3	$42,000	92	49	39	
5	0	1	3	$24,000	93	36	100	
6	1	2	2	$44,000	49	70	78	
7	1	1	4	$59,000	53	50	61	
8	1	3	4	$37,000	58	46	83	
9	0	3	4	$63,000	95	29	53	

8.26.3 Performing the Test in SPSS

The procedure in SPSS involves the following steps:

1. **Open the Dataset:** Load Wilcoxon test.sav into SPSS.
2. **Select the Test:** From the top menu, choose *Analyze → Nonparametric Tests → Legacy Dialogs → 2 Related Samples.*
3. **Assign Variables:** Move the two related variables — *Commercial 1* and *Commercial 3* ratings — into the *Test Pair(s) List* box.
4. **Choose Test Type:** Under *Test Type*, select *Wilcoxon* (this is the default selection).
5. **Run the Test:** Click *OK*. SPSS will produce the Ranks Table and the Test Statistics Table automatically.





8.26.4 SPSS Output and Interpretation

The SPSS output consists of two key tables:

a. Ranks Table

Comparison	**N**	**Mean Rank**	**Sum of Ranks**
Negative Ranks	3	7.33	22.0
Positive Ranks	14	11.36	159.0
Ties	1	—	—
Total	18	—	—

Interpretation:

Each respondent's rating for Commercial 3 was subtracted from the rating for Commercial 1. Positive differences mean that *Commercial 1* was rated higher; negative differences indicate the opposite. Out of 18 respondents, 14 rated Commercial 1 higher (positive ranks), 3 rated Commercial 3 higher (negative ranks), and 1 gave equal ratings (tie). The sum of positive ranks (159) is much larger than that of negative ranks (22), suggesting that Commercial 1 was generally preferred.

b. Test Statistics Table

Test	**Z**	**Asymp. Sig. (2-tailed)**
Wilcoxon Signed Ranks	−3.27	0.001

Interpretation:

The p-value (0.001) is less than 0.05, indicating that the difference in ratings between Commercial 1 and Commercial 3 is **statistically significant**. We therefore reject the null hypothesis and conclude that respondents did not judge the two commercials similarly.

Ranks

		N	Mean Rank	Sum of Ranks
Rating "family car" commercial - Rating "youngster car" commercial	Negative Ranks	3[a]	4.00	12.00
	Positive Ranks	15[b]	10.60	1 159.00
	Ties	0[c]		
	Total	18		

a. Rating "family car" commercial < Rating "youngster car" commercial

b. Rating "family car" commercial > Rating "youngster car" commercial

c. Rating "family car" commercial = Rating "youngster car" commercial

Larger differences between ratings in favor of "family car" commercial.

8.26.5 Step-by-Step Understanding

To understand how the test works conceptually, consider the sequence of transformations:

1. Compute the differences between each respondent's two ratings.
2. Disregard any zero differences (ties).
3. Rank the absolute values of remaining differences.
4. Assign positive or negative signs based on whether the first commercial was rated higher or lower.
5. Sum up all positive and negative ranks.
6. Compare the smaller of the two sums (W statistic) with critical values or, as in SPSS, convert it to a z-score.

In the given data, the predominance of positive ranks and the significant z-value confirm that *Commercial 1* received systematically higher ratings than *Commercial 3*.

Hypothesis Testing Decision

- **Null Hypothesis (H_0):** The median difference between ratings of Commercial 1 and Commercial 3 is zero.
- **Alternative Hypothesis (H_1):** The median difference is not zero.

Since $p = 0.001 < 0.05$, we **reject H_0**. This means that the two commercials were not judged equally; respondents significantly preferred one over the other, specifically *Commercial 1*.

Comparison with the Paired-Samples t-Test

Both the Wilcoxon Signed-Ranks Test and the paired-samples t-test are used for comparing two related samples. However, they differ in their assumptions and computational methods.

Aspect	**Paired-Samples t-Test**	**Wilcoxon Signed-Ranks Test**
Type of Data	Interval or ratio, continuous	Ordinal or non-normal continuous
Assumption of Normality	Required	Not required
Basis of Comparison	Mean differences	Median ranks of differences
Sensitivity to Outliers	High	Low
Test Statistic	t-value	W or z-value

Interpretation	Difference of means	Difference of medians

Thus, the Wilcoxon test provides a safer choice for small, skewed, or ordinal datasets, while the paired t-test remains more powerful for normally distributed data.

Reporting the Results

When writing research reports, results from the Wilcoxon Signed-Ranks Test should include the test statistic, sample size, and significance level. A typical format is:

A Wilcoxon Signed-Ranks Test was conducted to compare participants' ratings of Commercial 1 and Commercial 3. The results indicated a significant difference in median ratings, $z = -3.27$, $p = 0.001$. The sum of positive ranks ($W^{+} = 159$) was much greater than the sum of negative ranks ($W^{-} = 22$), suggesting that respondents rated Commercial 1 significantly higher than Commercial 3.

Such reporting provides clarity and academic rigour while maintaining interpretability for readers.

8.26.6 Advantages of the Wilcoxon Signed-Ranks Test

Flexibly applying tests within a model's parameters is important. Wilcoxon's Signed Ranks Test illustrates this though its many advantages. For instance, it does not need normality. Most research in possible and in practical settings seldom, if at all, comply with the normal distribution. Reliance of distribution free tests should not be thorough research in which one posits from uniform distributions. All research collections, especially those not meant for theory, seldom conform to the normal distribution. Data that is at the ordinal level, or not at the ratio, can still be useful even though compliance is at which point most advances cease. Wilcoxon's Signed Ranks Test is more than enough.

Its application in practical research settings makes it advantageous too. Most research observations are not clean, some values that are too high or too low than the common values are absent. Wilcoxon's Signed Ranks Test makes use of the distribution of score values. Values which are only ordinal, not meant for ranking, and other so called 'bad' data, are useful in collecting and validating results because data values are more hi than lo, as opposed to ratio and ordinal collections.

Flexibly paired samples, and tests, or observations under different stressing conditions, is advantageous too. These can aptly be called 'Before-After' studies, tests, or Wilkoxon signed ranks tests in which a fixed number of conditions are interchanged. Systems can operate under just a bilateral exchange and yield satisfactory functioning.

The test's simplicity is an advantage too. Most research collections are comprised of massive data piles. Ease of use of Wilcoxon's Signed Ranks Test, makes it a common 'go to' test in most computerized statistical methods. Its application in systems like MS Excel, and R, makes it easy for application with other complex tests one wishes to run.

Tests that are simple are most apt for functioning in a range of complex systems. Sets containing paired data of 10 to 25 observations are enough, as a test is meant for practical application. Pilot studies and other collections meant to gather data, more than theory, can utilize this test.

8.26.7 Limitations

Regardless of its merits, the Wilcoxon Signed-Ranks Test has its flaws just like any other test would. For starters, the test is less reliable than the paired-samples t-test assuming that the data is normally distributed, therefore, it may overlook the existing difference.

The test also must believe that the median is the center of the distribution of the difference and it is equally likely that values will appear above and below the median. If this is violated the test and its findings tend to lose the significance.

Further, it is of little to no use for the examination of more than two interrelated variables. With three or more interlinked variables or time intervals, the Friedman Test must be utilized.

The more relevant criticism is that of the decreased accuracy that comes with the transformation of data from being continuous to being ranks. This obfuscates minute details and lost test precision.

The Wilcoxon Signed-Ranks Test does not measure impact and change. It shows the change and conjectured significance in addition to direction, and constitutes it in such a way that there is no useful or practical significance to it.

Practical Applications

The Wilcoxon Signed-Ranks Test is widely used across fields such as marketing, education, medicine, and psychology.

- **Marketing:** Comparing consumer satisfaction ratings before and after a campaign or for two different advertisements.
- **Human Resources:** Evaluating training effectiveness by comparing pre- and post-test scores.
- **Education:** Measuring students' performance before and after an intervention.
- **Healthcare:** Assessing patients' condition before and after treatment.
- **Behavioral Studies:** Testing differences in attitudes or preferences under two related conditions.

In all these cases, the test provides a reliable non-parametric method for comparing paired data when traditional parametric assumptions cannot be satisfied.

Extension to More Than Two Conditions

When the same participants are measured under more than two conditions, the **Friedman Test** serves as a generalisation of the Wilcoxon Signed-Ranks Test. It analyses variance in ranks across multiple related samples, functioning as the non-parametric equivalent of the repeated-measures ANOVA.

Conclusion

The Wilcoxon Signed-Rank Test is a useful and robust non parametric method for dealing with the case of two connected samples and the data is rank, skewed, or taken from a small population. The Wilcoxon Signed-Rank Test replaces the sign test logic by accounting for the size of the differences by ranking them, hence increasing sensitivity.

In the case of car commercials, certainly the large positive sum ranks, combined with the significant test statistic ($z = -3.27$, $p = 0.001$) for the two commercials which respondents rated do not conform to the hypothesis of them being rated the same. Commercial 1, or the first commercial emerged as the better commercial, which indicates more commercial appeal.

With the test come not just the confirmation of the hypothesis, but the actionable steps as well. The Wilcoxon Signed-Rank Test is an essential part of the toolbox in social science and management and provides robust inference while the situation does not allow for the t-test to be as elegant.

8.28. 1 Kruskal–Wallis Test

Analyzing multiple independent groups to see if differences exist for a particular variable is a common task in research. A nutritionist, for example, could assess how three different supplement routines affect weight gain among athletes. A manager could assess employee satisfaction across three different departments.

For such comparisons, the One-Way Analysis of Variance (ANOVA) is the preferred statistical technique if certain conditions are met — namely, the dependent variable is distributed normally and the variances are the same across groups (homogeneity of variance). However, these conditions are often violated, especially if the sample is small, or the data are ordinal (ranked) rather than continuous.

If the conditions are violated, then the ANDA's non-parametric equivalent is the Kruskal- Wallis Test which enables the comparisons of three or more independent groups without the assumptions of normality and equal variances. Unlike ANOVA where means are compared, the Kruskal-Wallis Test compares the ranks of the medians of the groups.

The test was formulated by Kruskal and Wallis in 1952, and it takes the reasoning of the Mann Whitney U Test for more than two groups.

8.28.2 Concept and Rationale

The Kruskal–Wallis test analyzes whether the samples come from the same underlying distribution. Sample data is first ranked across all groups. All data points from all groups are combined, placed in a sorted order, and assigned ranks from lowest to highest. Then the average ranks are computed and compared for all groups.

If the average ranks were taken from the same underlying population, their average ranks should, in theory, be approximately equal. Significant differences in the average ranks means that there is a group which is more distinct as compared to the others.

The Kruskal–Wallis test is not a substitute for ANOVA, which is assumed to have normally distributed interval data and needs the means from different populations. ANOVA assumes no distribution shape at all, and works with ordinal or non-normal continuous data.

Hypotheses

The hypotheses for the Kruskal–Wallis test are:

- **Null Hypothesis (H_0):** All group distributions are identical. (There is no significant difference between the groups.)
- **Alternative Hypothesis (H_1):** At least one group's distribution differs. (There is a significant difference between at least two groups.)

Symbolically:

$$H_0 : \text{Median}_1 = \text{Median}_2 = \ldots = \text{Median}_k$$

$$H_1 : \text{At least one Median}_i \neq \text{Median}_j$$

where *k* is the number of groups.

Test Statistic

The Kruskal–Wallis test statistic (H) is computed as:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1)$$

where:

- N = total number of observations across all groups,
- n_i = number of observations in group *i*,
- R_i = sum of ranks for group *i*,
- k = number of groups.

For large samples (n > 5 per group), the H statistic approximates a **Chi-Square (χ^2) distribution** with *(k – 1)* degrees of freedom. Thus, SPSS reports it under the label "Chi-Square" in its output, even though it is technically the Kruskal–Wallis H value.

Example: Creatine Experiment

Let us consider an example from a small fitness experiment involving **creatine**, a supplement commonly used by bodybuilders. Eighteen participants were randomly divided into **three groups**:

1. Group 1: Participants who **did not take creatine**.
2. Group 2: Participants who took creatine in the **morning**.
3. Group 3: Participants who took creatine in the **evening**.

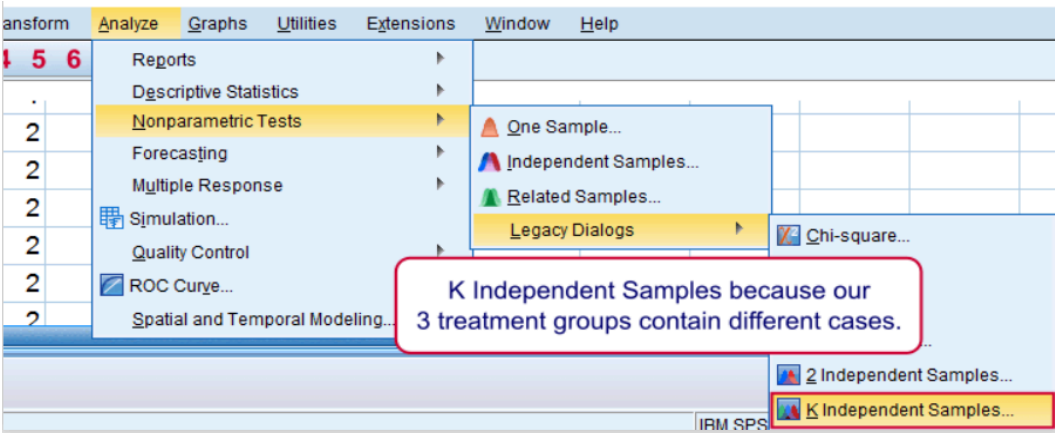
After a specific training period, weight gain was measured for each participant. The research question was:

Does the average weight gain depend on the creatine condition to which people were assigned?

Because the sample size is small and the dependent variable (weight gain) is not normally distributed, a **Kruskal–Wallis test** is appropriate. The dataset for this example is Kruskal-Wallis test.sav.

8.28.3 Steps for Conducting the Test in SPSS

1. **Open the Dataset:** Load Kruskal-Wallis test.sav in SPSS.
2. **Navigate to the Test:** Click *Analyze* → *Nonparametric Tests* → *Legacy Dialogs* → *K Independent Samples.*
3. **Assign Variables:**
   - Move *Weight Gain* into the *Test Variable List* box.
   - Move *Group* (Creatine Condition) into the *Grouping Variable* box.
4. **Define Groups:** Click *Define Range* and specify the numeric codes (1 to 3) corresponding to the groups.
5. **Choose Test Type:** Select *Kruskal–Wallis H* under the *Test Type* options (selected by default).
6. **Run the Test:** Click *OK* to generate the results.



8.28.4 SPSS Output and Interpretation

The SPSS output provides two main tables: the **Ranks Table** and the **Test Statistics Table**.

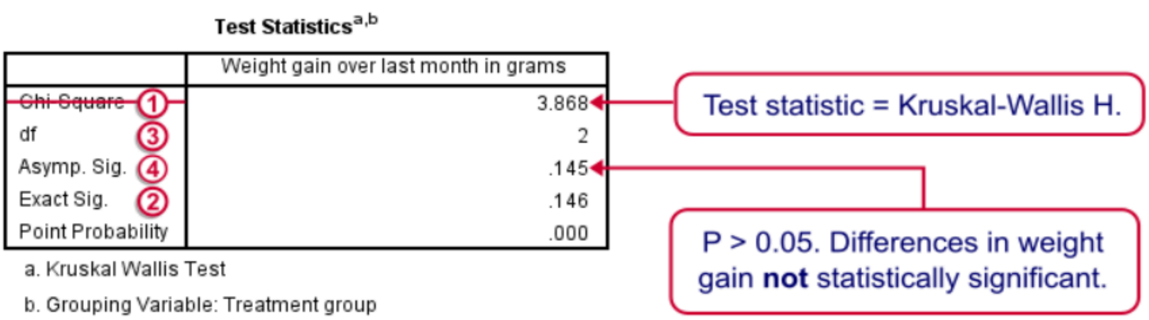
a. Ranks Table

Group	**N**	**Mean Rank**
No Creatine	6	8.33
Morning	6	11.25
Evening	6	12.42
Total	18	—

Interpretation:

The mean rank for the "Evening" group is the highest (12.42), suggesting that participants who took creatine in the evening may have experienced more weight gain on average. The "No Creatine" group has the lowest mean rank (8.33). However, to determine whether these differences are statistically significant, we must examine the test statistic.

b. Test Statistics Table


Test Statistics[a,b]

	Weight gain over last month in grams
Chi-Square	3.868
df	2
Asymp. Sig.	.145
Exact Sig.	.146
Point Probability	.000

a. Kruskal Wallis Test

b. Grouping Variable: Treatment group

Interpretation:

SPSS reports the H statistic under the heading "Chi-Square." The Asymptotic Significance ($p = 0.145$) represents the probability of obtaining the observed differences if all groups truly come from the same population.

Since **$p = 0.145 > 0.05$**, the result is **not statistically significant**. Therefore, we **fail to reject the null hypothesis** and conclude that the differences in average weight gain among the three creatine conditions are **not statistically significant**.

Understanding the Output

The test statistic, **$H = 3.90$** with **2 degrees of freedom**, corresponds to a p-value of 0.145. This means that there is approximately a 14.5% chance of observing the current pattern of rank differences purely by random variation if creatine condition has no real effect.

In simpler terms, while the group that took creatine in the evening appears to have slightly higher mean ranks, the difference is not large enough to be considered statistically significant at the 5% level.

8.28.4.1 Post-Hoc Analysis (If Significant)

When the Kruskal–Wallis test is significant ($p < 0.05$), it indicates that at least one group differs from the others, but it does not specify which groups differ. To identify the specific pairs that differ, **post-hoc pairwise comparisons** are conducted using adjusted Mann–Whitney tests with **Bonferroni correction** to control for multiple testing.

In SPSS, this can be performed by choosing *Analyze* ⟶ *Nonparametric Tests* ⟶ *Independent Samples* ⟶ *Customize Tests* and selecting pairwise comparisons after a significant Kruskal–Wallis result.

In the present example, since $p = 0.145 > 0.05$, no post-hoc tests are necessary.

Comparison with One-Way ANOVA

Aspect	**One-Way ANOVA**	**Kruskal–Wallis Test**
Type of Data	Interval/Ratio	Ordinal or Non-normal
Assumption of Normality	Required	Not required
Assumption of Equal Variances	Required	Not required
Basis of Comparison	Group Means	Group Median Ranks
Test Statistic	F	H (Chi-Square)
Post-Hoc Tests	Tukey, Scheffé, etc.	Pairwise Mann–Whitney Tests

The Kruskal–Wallis test is therefore ideal when ANOVA assumptions are violated, especially with skewed or ranked data.

Reporting the Results

When writing up results in research or a thesis, the report should clearly state the test, sample size, test statistic, degrees of freedom, p-value, and conclusion.

Example Report:

A Kruskal–Wallis H test was conducted to compare weight gain across three creatine consumption conditions (No Creatine, Morning, and Evening). The test revealed no statistically significant difference among groups, $H(2) = 3.90$, $p = 0.145$. This indicates that the timing of creatine intake did not have a significant effect on participants' weight gain.

If the test were significant, the report might continue:

Pairwise comparisons using the Mann–Whitney U test with Bonferroni adjustment indicated that evening intake resulted in significantly higher weight gain compared to no creatine intake ($p < 0.05$).

8.28.5 Advantages of the Kruskal–Wallis Test

One of the best things about the Kruskal–Wallis Test is that there is no need for the assumptions of parametric tests like ANOVA to be satisfied . Its strongest attribute is its non parametric nature. It does not need to check whether the data is normally distributed or whether the variances are homogenous. This attribute is extremely valuable for data that are skewed or data that fails to meet all of the stringent requirements of ANOVA.

It is also very convenient for ordinal data. It is able to work well with ranked or Likert-type scales which are widely employed in social sciences, psychology, and education. Because it uses the rankings instead of the raw values, it is able to use data which does not possess a true numerical order, but does possess some form of ordering.

Another positive aspect is that the Test can be used for unequal sample sizes. It can deal with an unbalanced design, which is common in most practical research scenarios, where equal sizes in the groups is difficult to obtain.

Unlike many other tests, the Kruskal–Wallis Test is not easily influenced by outliers. because extreme values do not greatly affect the overall ranking. This is an important attribute because it ensures that when there are a few extremely high or low data points, most, if not all of the results, can be considered valid .

Lastly, it means that "it easy to understand". This is because the test essentially evaluates the medians of the different groups by ranking instead of scoring which lends itself to the easy evaluation of whether groups in question have a statistically significant difference in means.

Limitations

The Kruskal-Wallis Test, however, is not bereft of shortcomings. As with most, it is less powerful than ANOVA when the data is suitable to be parametrically tested meaning that there is a subtle difference between groups that goes undetected.

Another shortcoming is that it focuses on the overall differences between groups instead of the particular ones that contradict the hypothesis. In such a case, a series of post hoc tests is warranted to know where the differences are coming from.

The test also presumes the shape of all the group distributions is the same. This assumption, when not met, gtouch on dicriting the true change of the measure and the changes in the shape of the distribution.

Lastly, the test is not suitable for designs with repeated measurements. In the case where the same subjects are tested under different conditions, the Friedman Test serves as a non-parametric hypothesis for such a case

When there is an abundance of continuous data, converting it to ranks means that it is less sensitive, and therefore, more significant small effects will be missed; such effects, however, would be spoken to in the parametric methodology.

Practical Applications

The Kruskal–Wallis test is widely used in business, management, and health sciences research. Examples include:

- **Marketing:** Comparing customer satisfaction scores across three branches of a company.
- **Human Resources:** Testing whether employee engagement differs across departments.
- **Healthcare:** Comparing patient recovery rates under different treatment regimens.
- **Education:** Assessing exam performance across three teaching methods.
- **Operations Research:** Comparing process times under different management interventions.

Its robustness and simplicity make it particularly useful when dealing with survey data and small experimental designs.

8.28.6 Interpreting Non-Significant Results

As you've seen previously, outcomes where p is greater than the alpha cut-off, such as p is 0.145, does not necessarily imply the absence of any value. The reasoning is that, in this case, the absence of value does not equate to the absence of evidence. The evidence, however, is not enough to claim that there is a “big enough” distinction to deem the claim substantively important or greater than 5%. Sample size, variation in measurements, and the absence of a detectable effect — even a small one — help characterize non significance.

Researchers, on the other hand, do not need to ignore non-quantifiable results that show any give a valuable classification that is worthy of being examined. Even non-quantifiable results provide significance.

Conclusion

The Kruskal- Wallis Test is as useful as a one-way ANOVA test would be if it were non-parametric, and correct for ordinal, non-normal, and heterogeneous data, for three or more groups independent or otherwise. It ranks all observations and synthesizes a median rank, thus avoiding parametric approaches and producing informative outcomes related to differences in groups.

For instance, in creatine test groups, a Kruskal- Wallis test would be appropriate to indicate no differences within and between groups in weight gain statistically, thus supporting a conclusion that morning, evening, and no supplement intake did not change performance outcomes.

However, the test does illustrate a key aspect of the methodology of any discipline: sample size alongside violated assumptions do not prevent the researcher from interpreting the data, as non-parametric approaches like the Kruskal- Wallis test do provide valid outcomes.

As such, the Kruskal-Wallis test is foundational to much contemporary practice, as it remains a benchmark for the analysis of data devoid of perfection.

8.29.1 Different Types of Sampling Methods:

Probability Sampling Methods

(Each unit has a known and equal chance of selection)

Sampling Method	***What It Means (Simple)***	***When / Why Used***	***Example (Easy to Understand)***
Simple Random Sampling	*Every member has an equal chance*	*When population is homogeneous and list is available*	*Selecting 100 students randomly from a college roll list*
Systematic Sampling	*Selecting every kth unit after a random start*	*When population list is ordered*	*Choosing every 10th household from a voter list*
Stratified Sampling	*Population divided into strata; sample from each*	*When population has clear subgroups*	*Sampling students separately from Arts, Science & Commerce*
Proportionate Stratified Sampling	*Sample size proportional to stratum size*	*To reflect population structure accurately*	*If 60% females, sample includes 60% females*

Sampling Method	*What It Means (Simple)*	*When / Why Used*	*Example (Easy to Understand)*
Disproportionate Stratified Sampling	*Equal samples from unequal strata*	*When small groups need representation*	*Equal rural & urban respondents despite size difference*
Cluster Sampling	*Population divided into clusters; some clusters selected*	*When population is geographically spread*	*Selecting 10 schools and surveying all students*
Multistage Sampling	*Sampling done in stages*	*For large-scale national studies*	*State → District → School → Student*

8.29.2 Non-Probability Sampling Methods

(Chance of selection is unknown; commonly used in social research)

Sampling Method	**What It Means (Simple)**	**When / Why Used**	**Example (Easy to Understand)**
Convenience Sampling	Selecting easily available respondents	When time/resources are limited	Surveying students present in class
Purposive (Judgmental) Sampling	Selecting participants based on purpose	When specific expertise is required	Interviewing senior teachers only
Quota Sampling	Fixed number from each category	To ensure representation without randomness	50 males & 50 females selected conveniently
Snowball Sampling	Participants refer other participants	For hidden or sensitive populations	Drug users referring other users
Volunteer Sampling	Participants self-select	Exploratory or online surveys	Google Form shared online
Theoretical Sampling	Sample evolves during research	Qualitative grounded theory studies	Adding participants until data saturation

Parametric Tests

Parametric Test	**Purpose / Use**
Z-test (one-tailed / two-tailed)	Compare sample mean with population mean (large sample)
One-sample t-test	Compare sample mean with a known value

Parametric Test	Purpose / Use
Independent t-test	Compare means of two independent groups
Paired t-test	Compare means of the same group (before–after)
One-way ANOVA	Compare means of three or more groups
Two-way ANOVA	Study interaction between two independent variables
Repeated Measures ANOVA	Compare repeated observations on same subjects
ANCOVA	Compare means while controlling covariates
MANOVA	Compare groups on multiple dependent variables
MANCOVA	MANOVA with covariates
Pearson Correlation	Measure linear relationship between variables
Simple Linear Regression (SLRA)	Predict one dependent variable from one predictor
Multiple Linear Regression (MLRA)	Predict one dependent variable from multiple predictors
Factor Analysis (EFA)	Identify underlying dimensions
Confirmatory Factor Analysis (CFA)	Validate measurement models
Structural Equation Modeling (SEM)	Test complex causal relationships
Time-Series Analysis	Analyze data collected over time

Non-Parametric Tests

(Used when data is non-normal, ordinal, categorical, or sample size is small)

Non-Parametric Test	Purpose / Use
Chi-Square Test	Test association between categorical variables
Fisher's Exact Test	Chi-square alternative for small samples
Mann–Whitney U Test	Alternative to independent t-test
Wilcoxon Signed-Rank Test	Alternative to paired t-test
Kruskal–Wallis Test	Alternative to one-way ANOVA
Friedman Test	Alternative to repeated measures ANOVA
Spearman Rank Correlation	Relationship between ordinal / non-normal data
Kendall's Tau	Correlation for small samples
Median Test	Compare medians of groups
Kolmogorov–Smirnov Test	Compare distributions
Shapiro–Wilk Test	Test normality
Runs Test	Test randomness

Non-Parametric Test	Purpose / Use
Sign Test	Directional difference test
Cochran's Q Test	Extension of McNemar test
McNemar Test	Paired categorical data
Non-parametric Regression	Regression without normality assumption

8.30.1 Common Statistical Tests Used in Thesis Research (Extended Table)

Test / Method	When to Choose (Purpose)	Type & Size of Data	Key Features	What to Check Before Use	Typical Thesis Applications
One-sample t-test	Compare sample mean with a known value	Continuous; small/medium sample	Parametric mean comparison	Normality	Policy benchmarks, standard comparison
Independent t-test	Compare **two independent groups**	Continuous DV; categorical IV	Tests mean difference	Normality; equal variance	Male vs female outcomes
Paired t-test	Compare **same group before–after**	Continuous paired data	Sensitive to change	Normality of differences	Training impact studies
Mann–Whitney U test	Non-parametric alternative to independent t-test	Ordinal / non-normal	Compares ranks	Independent samples	Likert-scale group comparison
Wilcoxon Signed-Rank test	Non-parametric alternative to paired t-test	Ordinal / non-normal	Before–after comparison	Paired data	Intervention studies
One-way ANOVA	Compare **3 or more groups** (one DV)	Continuous DV; categorical IV	Controls Type-I error	Normality; homogeneity	Education & management studies
Two-way ANOVA	Test **interaction effects**	Continuous DV; two categorical IVs	Examines combined effects	Same as ANOVA	Policy & behavioural research

Test / Method	When to Choose (Purpose)	Type & Size of Data	Key Features	What to Check Before Use	Typical Thesis Applications
Kruskal–Wallis test	Non-parametric ANOVA alternative	Ordinal / skewed data	Rank-based	Independent groups	Social survey analysis
ANCOVA	Compare groups **after controlling covariates**	Continuous DV + covariate	Adjusted group means	Linear covariate effect	Quasi-experimental designs
MANOVA	Multiple DVs across groups	Continuous DVs	Multivariate comparison	Large sample; covariance equality	Psychology, education
MANCOVA	MANOVA + covariates	Same as MANOVA	Controls confounders	Multivariate assumptions	Advanced thesis designs
Pearson Correlation	Linear relationship between variables	Continuous	Strength & direction	Normality, linearity	Behavioural & economic studies
Spearman Correlation	Non-linear / ordinal data	Ordinal / skewed	Rank-based	Monotonicity	Attitude & perception research
Simple Linear Regression (SLRA)	Predict DV from one IV	Continuous	Predictive relationship	Linearity	Cause–effect models
Multiple Linear Regression (MLRA)	Predict DV from many IVs	Continuous; larger sample	Controls multiple factors	Multicollinearity	Policy & social models
Logistic Regression	Binary dependent variable	Categorical DV	Odds & probability	No multicollinearity	Yes/No outcomes
Chi-Square Test	Association between	Categorical	Non-parametric	Expected frequency ≥5	Demographic studies

Test / Method	When to Choose (Purpose)	Type & Size of Data	Key Features	What to Check Before Use	Typical Thesis Applications
	categorical variables				
Fisher's Exact Test	Small sample categorical data	Categorical; small n	Exact probability	Small cell sizes	Rare-event studies
Factor Analysis (EFA)	Identify latent constructs	Large sample	Dimension reduction	KMO, Bartlett's test	Scale development
Confirmatory Factor Analysis (CFA)	Validate theoretical constructs	Large sample	Model testing	Model fit indices	Advanced scale validation
Reliability Analysis (Cronbach's α)	Check internal consistency	Scale data	Measures reliability	α ≥ 0.70	Questionnaire validation
Normality Tests (Shapiro–Wilk / K–S)	Check data distribution	Continuous	Assumption testing	p-value interpretation	Methodology justification
Levene's Test	Test equality of variance	Continuous DV	Variance check	p-value	ANOVA/t-test assumptions
Post-hoc Tests (Tukey, Bonferroni)	Identify **which groups differ**	After ANOVA	Pairwise comparison	Adjusted significance	Group-wise interpretation
Time-Series Analysis	Data across time	Longitudinal data	Trend & seasonality	Stationarity	Economic & policy research

FOUNDATIONS OF DATA ANALYSIS IN RESEARCH

World Academic Press

Kolkata, India

www.worldacademic.press

9 788168 643949