



Quantitative Data Analysis – SPSS a Crucial Tool

DR. SYED SHARIQUE ALI

Principal

K. M. ASGHAR HUSAIN COLLEGE OF EDUCATION, AKOLA

ABSTRACT: This paper aims at examining the utility of Statistical Package for social science research as an effective tool for quantitative data analysis which is widely known as SPSS. It elaborates the basic aspects of SPSS from its foundation as a tool for statistical analysis to its advancement as a priority in quantitative data analysis in various fields of research. For the fresh researchers it is must to understand the pros and cons of SPSS as to know to use it effectively to analyze the data. This study represents the details from personal experiences of using SPSS as well as referring to various reviews of earlier studies. Various works referring to SPSS have been discussed in the paper. The key features and advantages of using SPSS are presented in the paper so that the researchers can utilize the tool in analyzing their data to draw out the conclusions. Hence from the issued discussed in the paper we can reach up to the result that SPSS is one of the most important and crucial statistical tools for drawing the inferences from the quantitative data.

KEYWORDS: SPSS, Statistical tools, Quantitative Data Analysis, Data Analysis Tool, inferential analysis, Research Tool,

I. INTRODUCTION

Researches mean the collection of information from the field, through experiments or observations and provide a solution to different problems in a systematic manner. In this era of technology, computerization is playing an important role in social science research. It has now become an inseparable part in research. Social Science research also has been greatly influenced by the Technology. Earlier it was very difficult to analyse the quantitative data for which a researcher had to perform many analytical jobs. Taking this difficulty into consideration, Norman H. Nie, a Social Scientist, along with his two fellows, Dale H. Bent and C. Hadlai Hull in 1968 at the Stanford University developed Statistical Package for the Social Sciences widely known as SPSS. In the year 2009 IBM took over the software program and is now called IBM SPSS. The manual for SPSS written by the developers had become as one of the most influential books of the sociology of all time due to the ease of use features of statistical analysis for ordinary researchers. Data Management and data documentation are the key features of SPSS. The powerful and user friendly software package SPSS is beneficial for all sorts of statistical analysis of data. The researchers from the fields of sociology, psychology, economics, business studies, medicine, engineering, and other disciplines use SPSS for quantitative data analysis. Apart from that, SPSS is also used by various public, private, and non-government organizations also. Marketing and survey companies also use SPSS for analyzing consumer behavior and forecasting. This software is a full pack which helps the researchers to obtain statistics from simple descriptive numbers to the complex statistics. It also present the data in histograms, scatter plots, and other ways of presentation. The SPSS provides descriptive and inferential statistics on the data collected by the researcher from survey, experiments or through the observation.

II. METHODOLOGY

This paper is descriptive as well as an analytical. This paper is designed with a logical explanatory model. The paper is prepared on the basis of the experiment and on the secondary sources. The primary data source relies on the authors personal experiences of using SPSS for different social-science related research works done before. The secondary data sources like journal articles, seminar papers, and other related publications selected to advance this study are considered based on the properties such as authentic data source, accuracy and consistency of data, and author's background. The fair intention of this study is to draw direct explanations from the available data. An attempt is made to find out the pros and cons of the SPSS and its usability in social science research.

III. STATISTICS AND QUANTITATIVE ANALYSIS

According to Upton and Cook (2014), "statistics is the practice that collects, organize, and analyze numerical data and further interpret and present it to be applied in scientific, industrial, or social problem"

According to La Trobe, (2020) Statistics in another way is a mathematical procedure to interpret required information from gathered data, especially for using a group of samples from those in a representative population.

"Quantitative analysis is a statistical method of analyzing numerical data collected online or offline, polls, questionnaires, and surveys, or by considering pre-existing statistical data more mathematically with the help of computational techniques. The analyzed findings further generalized across different groups, or explain a particular phenomenon" (Muijs, 2010; Babbie, 2020).

“In social science research, within a given population, quantitative analysis is used to determine the relationship between an independent variable and a dependent variable. Values of the independent variable can be manipulated more precisely values can be controlled or changed to observe the effect on the dependent variable” (Niño-Zarazúa, 2012).

IV. SPSS: THE TOOL FOR STATISTICAL ANALYSIS

SPSS is a software which is used for the statistical analysis in social science research. Now a days quantitative data is predominantly used by most of the social science researchers for their studies. There are no. of software packages available to do statistical analysis such as R, PSPP, Stata, SAS, SPlus, JASP, and BMDP. Some of these are open source and the rest are under the commercial license. Excel, also allows quantitative data analysis through its add-on module widely used spreadsheet software by Microsoft also allows quantitative data analysis. Though these options are not so user-friendly, but SPSS being more user friendly as compared to its alternatives most educational and non-educational institutes use SPSS.

This software was acquired for the first time in 1968. From the version 10 of the software in 1983, data file was produced in such a way that it can analysis multiple records at a time. Now it can run in, MAC, Linux and Windows. It is developed by IBM and after 2015, all the versions named as ‘IBM SPSS Statistics (Bala. Jyoti, 2016).

V. ROLE OF SPSS IN SOCIAL SCIENCE RESEARCH

In research Hypothesis testing is very important because conclusions are drawn on the basis of hypothesis testing. Hypothesis testing means examining the pre assumed statements by analyzing the information generated from the field. To do this task flawlessly the researcher has to collect a large data from the field and then draws a tangible and errorless conclusion and hence computer based statistical programs are used to lessen the calculation overload and technical error in analyzing with large quantum of data. In this milieu, SPSS is the most widely used software in social sciences. This software can be easily used with minimum technological knowledge. All the SPSS packages contain “Tutorial File” which is very useful for the beginners.

VI. FEATURES OF SPSS

People with non technical background can easily handle the software as knowledge of programming language is not required to operate it. Hence the software is end user friendly and is applicable to any kind of quantitative analysis. Minimum computer knowledge is required to operate the software but the tone and genuineness of the analysis depends on the profundity of knowledge of the researcher. The researcher should have a clear design about the research and should have the plan, what to do with the data for analysis. Data management ability makes it a first choice of researchers in social science. The data management system is for enormous information, it includes the data documentation including the file reshaping, creating derive data, case selection system etc. Report generation capability is outstanding. What make this program superior to its alternatives is the comprehensiveness and its flexibility. Social science researchers prefer to use SPSS because of its collaboration, authorization, and deployment of survey data for strategic data mining for statistical analysis. The researcher can effectively analyze the quantitative data by knowing basic concepts of operating SPSS. Identifying the set of variables is required and further creates cases by suitable input of data within these variables. Four types of variables are there in the datasheet of SPSS i.e. independent variables, dependent variables, intervening variables and moderator variables. An independent variable is a cause. Its value is independent of any other variables in a study whereas, a dependent variable is an effect, whose value depends on any changes in the independent variable. Intervening variable (sometimes called mediating variable), it refers to a hypothetical variable, that is usually used to explain causal links between other variables within research. Lastly, the moderating variable is that variable that can alter the association between independent and dependent variables.

There are various statistical tools available in the SPSS, which are very helpful for the analysis of different type of data. These are like:

Types of Statistical analysis	Tools
Descriptive	Descriptive, Explore, Cross tabulation, Descriptive Ratio Statistics
Identifying groups	Factor analysis, cluster analysis (two-step, K-means, hierarchical)
Numerical outcomes	Linear regression
Bivariate	Means, t-test, ANOVA, Correlation (bivariate, partial, distances), Nonparametric tests.

Table on the types and tools of statistical analysis

VII. SOME COMMON STATISTICAL ANALYSES BY SPSS:

Some of the common statistical analyses used in the social science research on the basis of the objectives of the research through SPSS are;

1. When the impact of the differences between the two groups of the same variable has to be tested, then ‘Independent t-test analysis’ or ‘Mann-Whitney test analysis’ through SPSS can be used.
2. When the differences among more than two groups of the same variable are to be found out, then ‘ANOVA analysis’ or ‘Wallis test analysis’ can be used.
3. When the importance of the differences between more than two different groups towards more than one variable has to be measured, then the tools of ‘MANOVA analysis’ can be used.
4. The importance of the relationship between two interested variables can be examined by ‘Correlation analysis’ and ‘Rank Correlation analysis’.
5. ‘Multiple Linear Regression’ tools can be used to assess the fundamental and effective connection between a set of independent variables paired with one dependent variable.

6. 'Logistic Regression analysis' or 'Multinomial Regression analysis' can be used to look at the fundamental and effective connection between a set of independent variables.

7. For modification or rebuild or to verify the variables' arrangement that reflect the same response the researcher can use 'Exploratory Factor Analysis' tools of SPSS.

VIII. TYPES OF SPSS

Three different types of licensing systems are there with the stable version of IBM SPSS 28. Subscription-based; it does not require an authorization code, Non-subscription based; It requires an authorization code and also comes with one-year technical support, and the last one is the academic license (IBM, n.d.). Some latest features of SPSS 28 like meta-analysis, power analysis, ratio statistics, relationship maps, statistics workbook, search, table side-pane editor, and high contrast support (IBM, 2021). SPSS 28 is the most stable SPSS version available on the market at present and has the most number of tests that can be performed.

IX. ADVANTAGES OF SPSS:

It is Graphical User Interface (GUI) and thus easy to learn and use by the end user. The statistical capacity of SPSS is very high. It has the unique quality of creating variables from existing information. Most of the complex statistical tests are available built-in. One can open a variety of file formats in SPSS such as Excel, SAS, Stata, tab-delimited, and few other known formats. Platform compatibility for software is a big issue today and here SPSS has advantages over other statistical packages. SPSS runs on Windows, macOS, and LINUX platforms. SPSS is an excellent feature cluster analysis as well. SPSS deals with all facets of the analytical process from data preparation and management to data analysis and reporting. Forecasts and plans are also improved by imputing missing values with expected values with the use of regression and expectation-maximization. Automated methods to identify anomalies and statistical transformations to address outliers if any make the program distinguished. Much significant features of SPSS are tables and effective visualizations of results. It can also predict values of target variables based on values of predictor variables. Accurate modeling of linear and non-linear relationships enables another door for researchers. SPSS is very useful for the analysis of large data. Analysis of the huge data-based research like election results, census results etc can be easily done through this tool. This software utilizes the proprietary 4GL command syntax language (Bala. Jyoti, 2016). The benefit of the language is that it aids in simplifying and analyzing the complex information. Because of this language command many shortcuts like 'copy', 'paste' etc. button are there. SPSS provides the facilities to directly export the analysis (output) to a Word file, PDF, Excel sheet, etc. there is another significant feature is there, Researcher can convert their graphical analysis report into JPG, PNG, BMP etc. format.

X. LIMITATIONS OF SPSS

Price of the SPSS software is very high. Commercial license of the program is very costly. The students are not even getting the program free. Discounts with some cost attached to them are there. Editing the graphs is difficult. Not many options for graphs as well. Graphic quality could also be a major nay. The output is not easy to edit. The program can run slow depending on the machine it is installed. Certain add-on modules are not as easy as they should be. This software needs minimum basic knowledge on statistical or mathematic. Sometime it's become as a barrier for the researcher while using the SPSS. It is also found in many times difficult to start an analysis without the help of the expert on statistics.

XI. CONCLUSION

IBM SPSS is one of the most used statistical analysis software by social scientists all over the world. It has undergone several improvements over the years. It has been improved in accuracy and perfection in data presentation. It has been made feature rich and the user friendly as well. It has developed as a major statistical analysis tool for other genres of studies too. The dimensions of its use are so vast that people from various fields of studies like medical, engineering, business, education, the government is using SPSS. The scope of SPSS is huge. It has the quality of going deep into analyst's data and making it a much more efficient. SPSS has excellent advantage in drawing conclusions and predictions of user data. It is indeed one of the fastest in handling statistical procedures. It is not only confined to statistical analysis but data management and documentation is also its distinguishing feature. It Enables the researchers to achieve any complex analyses and further plotting the same to various charts and graphs available within the program. It has user-friendly interface. Availability of online tutorials and documentation from the parent company and online user communities has made it one of the most easy-to-learn statistical analysis software. In spite of its few disadvantages, its unique features make it distinct in comparison to other statistical tools for quantitative data analysis. SPSS has the capability conducting all major tests required for quantitative data analysis. In today's time SPSS is not only the choice but in some cases considered crucial for social researchers to use it as their quantitative data analysis and representation tool.

XII. RECOMMENDATIONS

All the educational and research institutions should make sure the accessibility of the SPSS software. Training programs in all the research institutions should be conducted on the SPSS concerning its usage in statistical analysis of data. As the cost of the software is high, hence the availability of this software should be ascertained by the institutions. There should be a free version for the registered researcher under the university and valued research institutions with a minimum limitation. Or the price of the SPSS software for the registered researcher should be less than the ordinary people or institution, etc. The government also should provide the SPSS software to the researcher at a subsidized price. SPSS should be compulsorily taught as a part of syllabus at PG level.

REFERENCES:

1. Aljandali, A. (2016). Quantitative analysis and IBM® SPSS®statistics: A guide for business and finance.
2. Arkkelin, D. (2014). Using SPSS to Understand Research and Data Analysis. Psychology Curricular Materials.
3. Arifa Rahman1, Md. Golam Muktadir2 (2021) International Journal of Research and Innovation in Social Science (IJRISS) |Volume V, Issue X, October 2021|ISSN 2454-6186
4. Babbie, E. R. (2020). The practice of social research. Cengage AU.
5. Bala, J. (2016). Contribution of SPSS in Social Sciences Research. International Journal of Advanced Research in Computer Science, 7(6), 250-254.

6. Devare. Suresh (2015); Important of SPSS for Social Sciences Research, accessed from <http://ssrn.com/abstract=2663283> on 21st September, 2019.
7. IBM. (2020). IBM SPSS Statistics: What's New
8. IBM. (2021). What's New in SPSS Statistics 28.
9. IBM. (n.d.). IBM SPSS Statistics: Features and modules
10. Joyjit Hazarika (2019); International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075 (Online), Volume-8 Issue-12, October 2019
11. Mehta, C. R., & Patel, N. R. (1989). IBM SPSS Exact Tests (2011 ed.). IBM Corp.
12. Puteh. Fadilah (2017), Quantitative Data Analysis: Choosing Between SPSS, PLS and AMOS in Social Science Research, International Interdisciplinary Journal of Scientific Research, Vol. 3 No. 1 July.
13. <https://surveysparrow.com/blog/what-is-spss/>

