International Journal of innovation in Educational Management (IJIEM); Vol. 2, No.1; September, 2018.
Refereed Journal
www.ijiem.com.ng
ISSN (Print: 2598-6413, Online: 2598-6489)

Comparative Study of Manual and Computerized Software Techniques of Data Management and Analysis in Educational Research.

Owan, Valentine Joseph & Bassey, Bassey A.

Abstract
The study was conducted to compare manual and computerized software techniques of data management and analysis in educational research. Specifically, the study investigated whether there was a significant difference in the results of Pearson correlation, independent t-test and ANOVA obtained from using manual and computerized software technique of data analyses. Three null hypotheses were formulated accordingly to guide the study. The study adopted a quasi-experimental research design where several data were generated by the researchers and analyzed using manual and computerized software techniques. The data were generated to suit the required data of each statistical method of analysis. CASIO fx-991ES PLUS NATURAL DISPLAY scientific calculator and statistical tables were used for manual analysis; while data analysis tool pack of Microsoft Excel version 2013 were used for computerized software analysis. The results of the analysis revealed that both manual and computerized software techniques yielded the same results for Pearson correlation, independent t-test and ANOVA. It was concluded that though both manual and computerized techniques are reliable and dependable, computerized technique is faster and efficient in managing and analyzing data than manual technique. It was recommended, among other things, that any of the techniques should be used without fear when computing Pearson, independent t-test and one-way ANOVA as it is the same results that will be gotten.

Keywords: Manual; Computerized software; Data Management; Data Analysis; Educational Research.

Citation: Owan, V.J. & Bassey, B. A. (2018). Comparative Study of Manual and Computerized Software Techniques of Data Management and Analysis in Educational Research. *International Journal of Innovation in Educational Management (IJIEM), 2*(1), 35 - 46.

Introduction
The goal of every research is to solve a unique problem within the environment or to contribute to already existing knowledge by filling gaps that may exist in the literature. The means through which research reaches its end is through proper decisions that are made through inferences and deductions. Inferences are best made if there are empirical evidence to justify such decisions. In educational research, inferences are usually made through quantitative or qualitative data analysis or both (in mixed methods researches). Quantitative inferences are made through hypotheses testing with the help of inferential statistical techniques such as correlation for explaining relationships; t-test for finding differences that may exist between two dependent or independent samples; Chi Square which tests for independence; regression, analysis of variance and so on. Qualitative analysis on the other hand, is the use of words and sentences to explain observed phenomena on the basis of data generated from field experiments, or research instrument. The research design adopted for any study may not be as important as the methods in which data obtained will be managed, organized and analyzed.

According to Patel (2016), the role of data is central to any piece of research or research project. The research data is meticulously collected, analyzed, optimized, organized and made usable to carry out research studies. Research cannot happen and would not be complete in any area of study without authentic and objective data. Therefore, data sharing is increasingly encouraged and occasionally mandated by publishers and research institutions. Several journals now require the research data sets to be submitted along with the articles. For example, the Public Library of Science (PLoS), which publishes open access journals in the field of biological sciences, has a stringent data access policy which all prospective authors need to comply with.

Educational research data management (ERDM) refers to the process of collecting, processing, filing/storing, and retrieving and sharing educational research data, including all relevant research files and documents when needed. It is a term that describes the organization, storage, preservation, and sharing of data collected and used in a research project. It also involves decisions about how data will be preserved and shared after the project is completed.

However, managing research data is not without problems. Some of the challenges in research data management (RDM) that need to be addressed include: copyright, data licensing, erroneous interpretation of data, data security, data privacy, mindset, method of data storage/retrieval, etc. To address these problems, Patel (2016), suggested a framework which include: developing institutional policies for data-sharing; changing the mind-set of researchers; data collection from researchers rather than from publishers; defining clearly, every research project including the terms and conditions related to the ownership of the data; data repositories should have a mechanism to cross refer the data with the methodology used to collect or generate the data; data classification should be properly done; data anonymization; data description and identification; data organization; and developing interoperability framework for the research data.

On the other hand, data analysis refers to a systematic process of breaking down research data into various units in a manner that each unit is understood independently. It also involves the process through which decisions can be made based on the results analyzed and presented and the hypothesis tested. Data analysis is very important in any educational research study because it is through it that results (which is the sole aim of researches) are gotten, and it also provides a basis for conclusions to be reached. According to Hellerstein (2008), data analysis is a process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, suggesting conclusions, and supporting decision-making. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, in different business, science, and social science domains. It is the process of systematically applying statistical and/or logical techniques to describe and illustrate, condense and recap, and evaluate data. According to Shamoo and Resnik (2003) various analytic procedures provide a way of drawing inductive inferences from data and distinguishing the signal (the phenomenon of interest) from the noise (statistical fluctuations) present in the data.

There are also a number of issues that researchers should be cognizant of with respect to educational research data analysis. These include: having the necessary skills to analyze; concurrently selecting data collection methods and appropriate analysis; drawing unbiased

inference; inappropriate subgroup analysis; following acceptable norms for disciplines; determining statistical significance; lack of clearly defined and objective outcome measurements**;** providing honest and accurate analysis; manner of presenting data; environmental/contextual issues; data recording method; partitioning 'text' when analyzing qualitative data; training of staff conducting analyses; reliability and validity and extent of analysis (Silver & Manson, 2003).

It must also be noted that with the global trends and technological evolutions, computerized software and manual methods of quantitative and qualitative data analyses are available to researchers for use. The choice of technique to use depends on the researcher, his knowledge, skills and mastery. Whether statistical or non-statistical methods of analyses are used, researchers should be aware of the potential for compromising data integrity. While statistical analysis is typically performed on quantitative data, there are numerous analytic procedures specifically designed for qualitative material including content, thematic, and ethnographic analysis. Regardless of whether one studies quantitative or qualitative phenomena, researchers use a variety of tools to analyze data in order to test hypotheses, discern patterns of behavior, and ultimately answer research questions. Failure to understand or acknowledge data management and analysis in educational research and it related issues presented above can compromise data integrity.

Statement of the problem

As earlier explained above, the importance of data management and analysis in educational research cannot be over-emphasized. It should be done with all amount of carefulness using the right tools to ensure that results gotten are not borne out of errors from the researcher or the instruments used in managing or analyzing such data. It is true that manual and computerized software systems of data management are in wide use by different researchers in different locations. Since research results are used to make valid decisions (usually about the population and its parameters), it is expected that one adopts a technique that will enable him not to make any mistake in the process of analysis.

However, it has been observed that many researchers adopt inappropriate statistical tools for analyzing data which often lead to a mismatch between the tool used and the result obtained. Also, most results obtained even where an appropriate statistical is used, contain errors. These errors most often, are traceable to the data analyst. This implies that many research data analysts make mistakes in their computations which can in part be attributed to their knowledge, in part to their methods, and in part to the techniques of data management and analysis. There has also been some debate in the literature with respect to the most reliable technique of data analysis. Some researchers and data analyst, favour the use of computers and its related analytical software for analysis, while others still believed that manual data management and analytical techniques are dependable and more reliable.

Given the importance of research results, and the role it plays in decision making, one will want to make use of an approach that will yield dependable results. Also, the arguments amongst researchers and scholars regarding the most reliable technique of data analysis, will be brought to fore if a study that will investigate and compare both methods of data management and analysis like this, is carried out.

Purpose of the study

The main purpose of this study was to compare between manual and computerized software data management and analysis techniques in educational research.

Specifically, this study sought to:

i. determine whether there is a difference in the results of Pearson correlation obtained from using manual or computerized software techniques of data management and analysis.
ii. determine whether there is a difference in the results of independent t-test obtained from using manual or computerized software techniques of data management and analysis.
iii. determine whether there is a difference in the results of One-Way ANOVA obtained from using manual or computerized software techniques of data management and analysis.

Statement of hypotheses

The following null hypotheses were formulated to guide this study

Ho_1: There is no significant difference in the results of Pearson correlation obtained from using manual or computerized software techniques of data management and analysis.

Ho_2: There is no significant difference in the results of independent t-test obtained from using manual or computerized software techniques of data management and analysis.

Ho_3: There is no significant difference in the results of One-Way ANOVA obtained from using manual or computerized software techniques of data management and analysis.

Literature review

The importance of data management and analysis in educational research cannot be over-emphasized as it constitutes the basis for which decisions in educational researches are being made. Robson, (2002) noted that research data management (RDM) is gaining a lot of momentum in the present day and rightly so. Research data are the core of any research study. The findings and conclusions of a study are entirely dependent on the research data. Traditional publishing did not focus on the presentation of data, along with the publications such as research monographs and especially journal articles, probably because of the difficulties involved in managing the research data sets. The current day technology, however, has helped in making this task easier (Robson, 2002).

Studies have been conducted with various positions held regarding the best method of data management and analysis. In their study of the role of computer software in the analysis of qualitative data, Morrison and Moir (2008), in their paper illustrated the ways in which a package called NUD.IST facilitated analysis where grounded theory methods of data analysis were also extensively used. While highlighting the many benefits that ensued, their paper illustrated the limitations of such programs. The purpose of their paper was to encourage researchers contemplating the use of computer software to consider carefully the possible consequences of their decision and to be aware that the use of such programs can alter the

nature of the analytical process in unexpected and perhaps unwanted ways. They maintained that in the last 15 years there has been a proliferation of computer software packages designed to facilitate qualitative data analysis. The programs can be classified, according to function, into a number of broad categories such as: text retrieval; text base management; coding and retrieval; code-based theory building; and conceptual-network building. The programs vary enormously in the extent to which they can facilitate the diverse analytical processes involved. The decision to use computer software to aid analysis in a particular project may be influenced by a number of factors, such as the nature of the data and the researcher's preferred approach to data analysis which will have as its basis certain epistemological and ontological assumptions.

Denzin and Lincoln (2000), noted that it has sometimes been argued that computer programmes can distance the researcher from the data and thereby influence negatively, the analysis and interpretation of meaning. Notwithstanding these criticisms, he explained that several good computer programmes exist that can help the researcher to manage large data sets and the processes of filing, storage, coding, retrieval and presentation of data. These Computer Aided Qualitative Data Analysis Software (CAQDAS) programmes can replace or support some of the manual tasks that qualitative researchers need to undertake. Tesch (1993) suggests that these programmes can help with the different approaches to qualitative research (language orientated, descriptive or interpretative and theory building) by assisting with: Storing, annotating and retrieving texts; finding words, phrases and segments of data; naming or labeling data; sorting and organising the data into manageable sections; identifying sections (known as *units*) of data; preparing diagrams and maps on screen; and retrieving and extracting quotes.

However, a post by Cambridge University (2018), argued that it has sometimes happened that research students spend two or more years collecting large amounts of data before considering the problems of analysing such data and with little understanding of their statistical or computing needs. In a disturbing number of cases the proposed analysis has subsequently had to be severely restricted or even abandoned. The data may be too massive to handle; more often some essential measurements have been not made, survey questions are ambiguous or otherwise defective, or samples are of insufficient quality or size. Sometimes the standard statistical or numerical methods and the available computer programs are inadequate and the development of suitable techniques would of itself provide a research project with problems. Serious problems can also arise from the use of inappropriate computer software for data analysis. This can lead to results that are demonstrably incorrect or at least suspect, besides risking unfavourable comment from referees and reviewers. For example, statistical procedures provided by spreadsheet and database packages are best treated with caution. Though convenient for commercial use (for which they are chiefly intended), these might not meet the standards required for academic research. They might be based on poor numerical techniques leading to inaccurate statistics, their limitations might be poorly documented if at all, and essential supplementary tests might be omitted. Such packages can be very useful for data entry and management, but for statistical analysis it is always advisable (and often easier) to use software written for the purpose by a reputable specialist manufacturer. University staff, research workers, and supervisors of research students are urged to ensure that methodological and computing requirements for data analysis are thoroughly evaluated at an early stage in the

planning of any relevant project and certainly before any substantial resources are spent in collecting data. If necessary, advice should be sought. This is particularly important in the Arts and Social Sciences, where data are intrinsically very complex (University of Cambridge, 2018).

An increasing number of researches are being conducted in education. It is therefore important that educational researchers become aware of the possibilities of using computer data analysis software packages. Analysing data is always time-consuming and any software package that can reduce the amount of time spent in this activity should be welcomed. However, the process of learning to make the best use of such packages is itself time-consuming. It is also best to learn about these packages before actually using one. Traditionally, computer software packages designed for the management and analysis of research data have been associated with quantitative methods. However, analytical packages are increasingly being designed for use in qualitative research studies. Researchers say that these packages are now being routinely used and are revolutionising the analysis of qualitative data (St John & Johnson, 2000; Fielding & Lee, 1998).

It would be useful if more people who are already using computer software analysis packages were to discuss and disseminate information on how effective the package actually was in relation to managing their data. This would enable novice researchers to benefit from the increased amount of information available to them, which in turn would allow them to make a more informed choice of software package to suit their research. Despite the difficulties associated with choosing the most appropriate software package, there are real benefits to their use, some of which are considered here. Jemmott (2002) claims that data preparation and management are much easier when using computer software packages rather than manual techniques. Software packages allow for effective and efficient coding of themes and categories, and for easy retrieval and movement of data between documents when compared with manual handling. This allows for the straightforward attachment of codes to segments of text that can even be colour coded for ease of identification and retrieval.

While software can be extremely helpful to researchers, it does have a number of disadvantages. Learning how to use a software package is a steep learning curve and can be time-consuming. Time must therefore be set aside in order to become familiar with the software (St John & Johnson, 2000; Woods & Roberts, 2000). The language and jargon used within these packages will become more familiar with time and use. Some software packages include an interactive tutorial that users are encouraged to complete. However, from a novice's perspective, this is not a straightforward process. There is an assumption that users are already familiar with other packages and this may not necessarily be the case.

From the above, it can be seen that there are arguments and differences in the positions held by different researchers with respect to using manual or computerized software techniques in data management and analysis in research generally. Also, all the works cited in this study are non-Nigerian studies which implies that little or no such study has been conducted in Nigeria. No foreign study has also been found which compare both techniques, meaning that this area has not really been explored by researchers resulting in a gap in the literature. An attempt to fill these gaps is what gave rise to this study.

Methods

This study adopted quasi-experimental design. Series of data that suit each statistical technique of analysis were generated by the researchers and both techniques (manual and computerized software) were used to analyze the data in order to determine the results obtained. The tools used for calculation and analysis include: the CASIO fx-991ES PLUS NATURAL DISPLAY scientific calculator and statistical tables for manual analysis; while data analysis tool pack of Microsoft Excel version 2013 were used for computerized software analysis. The results of Pearson correlation, independent t-test and one-way ANOVA that were obtained from both manual and computerized software analyses were compared in order to determine whether there is a difference in the results.

Presentation of results

Ho_1: There is no significant difference in the results of Pearson correlation obtained from using manual or computerized software techniques of data management and analysis.

Table 1: Scores of 17 students from the Departments of Educational Administration and Educational Foundations respectively, from a research methods examination results and analyzed using manual and computerized software techniques.

Edu Admin.	60	49	88	76	34	34	26	93	34	21	43	23	65	40	58	21	66
Env. Educ.	38	67	43	55	89	38	21	92	34	56	65	73	45	37	45	78	39

Table 2: Summary of Pearson correlation results obtained from analyzing the data in table 1 using manual and computerized software techniques of data management and analysis (N=17).

Manual	$\sum X_1$	$\sum Y_1$	$\sum X_1^2$	$\sum Y_2^2$	$\sum XY_1$	*Calc.* r_1
	831	915	49119	55947	44980	.0335
	831	915	49119	55947	44980	.0335
Computerized	$\sum X_2$	$\sum Y_2$	$\sum X_2^2$	$\sum Y_2^2$	$\sum XY_2$	Calc. r_2

The results in table 2 indicates that the r – value of .0335 gotten from the calculation of both manual and computerized analyses is the same. Therefore, the null hypothesis stating that there is no significant difference in the results of Pearson correlation obtained from using manual or computerized software techniques of data management and analysis is retained. This implies that there is no difference in the results of both manual and computerized software techniques.

Ho_2. There is no significant difference in the results of independent t-test obtained from using manual or computerized software techniques of data management and analysis.

Table 3: Scores of 25 male and 25 female students of Guidance and Counselling in a statistics test in University of Calabar.

Male	6	5	7	5	8	4	8	7	3	7	6	0	2	1	9	6	8	3	4	5	8	6	9	3	9
Female	8	7	2	7	5	8	9	5	6	3	1	5	5	0	6	4	6	2	5	8	9	8	4	3	3

Table 4: Summary of independent t-test results obtained from analyzing the data in table 3 using manual and computerized software techniques of data management and analysis.

Manual	$Mean_1$	$Mean_2$	S_1^2	S_2^2	*d f*	*Calc. t*	*t – crit.*
	5.56	5.16	6.507	6.307	48	0.559	t < 2.014 > 2.009
	5.56	5.16	6.507	6.307	48	0.559	2.011
Computerized	$Mean_1$	$Mean_2$	S_1^2	S_2^2	d *f*	Calc. t	t-crit.

The results in table 4 indicates that both manual and computerized techniques of data analyses yielded the same results across board. Therefore, the null hypothesis that there is no significant difference in the results of independent t-test obtained from using manual or computerized software techniques of data management and analysis is retained. This implies that in independent t-test, there is no difference in the results of both manual and computerized software techniques.

Ho_3. There is no significant difference in the results of One-Way ANOVA obtained from using manual or computerized software techniques of data management and analysis.

Table 5: Scores of 60 students taught using three teaching methods – Lecture, Demonstration, and Discussion methods and their relative effectiveness on students' performance in Statistics.

Lecture	6	9	20	12	10	9	14	6	13	18	12	4	5	8	8	6	5	17	12	7
Demonstration	8	13	15	5	6	9	3	9	6	12	14	20	11	9	11	17	9	9	2	6
Discussion	9	19	12	14	15	6	9	9	20	10	3	2	9	1	20	18	8	8	7	9

Table 6: Manual technique ANOVA results of data in table 5

Source of Variation	*SS*	*d f*	*MS*	*F*	*F crit.*
Between groups	4.9	2	2.45	0.098	F < 3.18 > 3.15
Within groups	1419.95	57	24.911		
Total	1424.85	59			

Table 7: Computerized software ANOVA results of data in table 5

Source of Variation	*SS*	*d f*	*MS*	*F*	*p-value*	*F crit.*
Between Groups	4.9	2	2.45	0.098	0.906	3.159
Within Groups	1419.95	57	24.911			
Total	1424.85	59				

The results in table 6 and table 7 revealed that the results of ANOVA are the same for both manual and computerized techniques of data management and analysis. Therefore, the null hypothesis (H_o3) is retained which implies that; there is no difference in the results of One-Way ANOVA obtained from using manual or computerized software techniques of data management and analysis.

Discussion

The results obtained for Pearson product moment correlation was not different at all from both techniques of analysis. This means that both approaches will always yield the same results. Also, during the process of computation, more time was spent calculating manually, it took just some few minutes to enter the data into Microsoft excel, while it took just a few seconds for the software to analyze the data.

The findings of this study revealed that the results obtained from manual and computerized software independent t-test analysis were not different. However, computerized technique yielded exact value for 48 degrees of freedom, while in the statistical tables used in manual analysis, there was no exact values for 48 degrees of freedom; instead the values 2.014 and 2.009 for 45 and 50 degrees of freedom respectively were used to provide an estimate of the actual degrees of freedom. The Computerized technique also yielded the answer very much faster than manual, saving a lot of time and energy as opposed to the manual technique.

The findings of this study also revealed that the ANOVA results obtained from both manual and computerized techniques of data management and analyses were the same. Meaning that there was no difference in results obtained from both techniques. However, the difference was that the computerized technique of data management and analysis, provided a p – value as well as exact critical values 3.159 for 57 degrees of freedom; something the manual technique could not be used to do. The manual technique F-values $F < 3.18 > 3.15$ was an estimated F-value for 57 degrees of freedom. Actually, 3.18 is for 50 degrees of freedom, while 3.15 is for 60 degrees of freedom. Since 57 falls between 50 and 60, their degrees of freedom were used to provide an estimate. The computerized technique provided its ANOVA results in just seconds while it took over 50 minutes to compute the ANOVA, manually.

From all the findings of this study, it can be said that the computerized software technique is a lot faster, and more efficient in data management and analysis than the manual technique. However, both manual and computerized software techniques were found to be reliable and dependable. These findings are supported by the position held by Jemmott (2002) who explained that data preparation and management are much easier when using computer software packages rather than manual techniques. Software packages allow for effective and efficient coding of themes and categories, and for easy retrieval and movement of data between documents when compared with manual handling. This allows for the straightforward attachment of codes to segments of text that can even be colour coded for ease of identification and retrieval.

Conclusion

Based on the findings of this study, it was concluded that there is no significant difference in results of Pearson product moment correlation, independent t-test and one-way analysis of variance statistics obtained from using manual technique or computerized software techniques of data management and analysis. Computerized software technique is a lot faster and efficient in data management and analysis than the manual technique. However, both manual and computerized software techniques were found to be reliable and dependable.

Recommendations

Based on the conclusion of this study, it was recommended that:

1. Both manual and computerized software techniques should be used in data management and analysis without fear when computing Pearson, independent t-test, and one-way analysis of variance, as it is the same results that will be gotten,
2. Data analyst should be careful when coding using both manual and computerized approaches to avoid inputting errors and make mistakes in the final output.
3. Computerized technique should be used when the data is very large, as this will save time and increase efficiency in the process of analysis.

Suggestions for further research

Other statistical methods such as Chi-Square, ANCOVA, Regression, dependent t-test, Spearman Rank correlation and so on, were not compared; meaning that conclusions reached in this study, perhaps, may have been different if all of these were included. This study did not explore other computerized software of data management and analysis such as SPSS, R, Q1 macros, Stata, Minitab, Scilab, SAS, JMP, and so on. Only Microsoft Excel 2013 was used. Who knows, maybe other software might yield different or the same results. Further researches need to be conducted using other packages and including other statistical methods.

REFERENCES

Cambridge University (2018). Analysis of data by computer. Retrieved on 16th April, 2018; 7:05pm WAT from https://goo.gl/ciHFLF

Denzin, N.K. & Lincoln, Y.S. (2000). *Handbook of qualitative research*. 2nd edition. Thousand Oaks, CA: Sage.

Fielding, N. & Lee, R. M. (1998). *Computers analysis and qualitatave research*. London: Sage.

Hellerstein, J. (2008). *Quantitative Data Cleaning for Large Databases.* EECS Computer Science Division: 3.

Jemmott, W. D. (2002). Data preparation and management using computer software packages and manual techniques. *Research Journal of Statistical Innovations; 5*(7): 22 – 45.

Morison, M. & Moir, J. (2008). The role of computer software in the analysis of qualitative data: efficient clerk, research assistant or Trojan horse? *Journal of Advanced Nursing; 28*(1):106 – 116.

Patel, D. (2016). Research data management: A conceptual framework. *Accounting Research Journal; 65*(4); 222 – 241.

Robson, C. (2002). *Real world research: A resource for social scientists and practitioner researches*, 2nd edition. Oxford: Blackwell

Shamoo, A. E & Resnik, B. R. (2003). *Responsible conduct of research.* London: Oxford University Press.

Silverman, S; & Manson, M. (2003). Research on teaching in physical education doctoral dissertations: a detailed investigation of focus, method, and analysis. *Journal of Teaching in Physical Education; 22*(3): 280 – 297.

St John, L. G. & Johnson, H. P. (2000). *Collecting and interpreting qualitative materials* (2nd ed.). Thousand Oaks, CA: Sage

Tesch R. (1993). Software for qualitative researchers: Analysis needs and program capabilities. In N.G. Fielding & Lee, R. M (Eds), *Using computers in qualitative research* (pp. 16 - 37). London: Sage.

Woods, F. D. & Roberts, C. J. 2000). Conducting and presenting social work research: some basic statistical considerations. *British Journal of Social Work; 33*(1): 567 – 573.

About The Authors

Owan, Valentine Joseph
Department of Educational Administration and Planning, University of Calabar, Calabar.
07062914623
owanvalentine@gmail.com

Bassey A. Bassey, PhD
Department of Educational Foundations, University of Calabar, Calabar.
08035407639
babssey67@gmail.com