Original Article *Journal of Multidisciplinary Research Advancements (JOMRA), Vol. 3, No. 2 (2025), 165 – 174*
e-ISSN 2990-7721 Available online at https://www.nepjol.info/index.php/jomra

Demystifying Statistical Tools and Treatment of Data in Research

Akaawase Alexander Mchi[1*], Moses Odigwe[1], Collins Ojoh[2]

[1]Department of Urban and Regional Planning, Southern Delta University, Ozoro, Nigeria
[2]Department of Urban and Regional Planning, Dennis Osadebay University, Asaba, Nigeria

**Corresponding author: mchia@dsust.edu.ng*

Abstract: This article focuses on demystifying statistical tools and data treatment in research. It began with a brief explanation of statistical concepts, data collection, and measurement scales, highlighting common errors in the use and selection of statistical tools. There are multiple statistical tools available for researchers. Some statistical tools are simple, whereas others are complex and highly purpose-specific. The article presents an overview of basic statistical tools necessary for research and explains when to use them to avoid manipulating analyses to achieve statistical significance. The article further identifies the conditions under which the choice of statistical tools for comparisons and relationships should be made if the results are to be meaningful.

Keywords: *Statistical tools, Treatment of data, Statistical assumptions, Parametric, Non-parametric tests*

Conflicts of interest: None
Supporting agencies: None

Received 25.03.2025 Revised 10.10.2025 Accepted 25.10.2025

Cite This Article: Mchi, A.A., Odigwe, M., & Ojoh, C. (2025). Demystifying Statistical Tools and Treatment of Data in Research. *Journal of Multidisciplinary Research Advancements*, 3(2), 165-174.

1. Introduction

Some novice researchers may have strong studies to present, but are afraid of the statistical methods involved, which hinders a realistic presentation of their results. The science of statistics concerns the collection, analysis, presentation, and interpretation of data (Ali and Bhaskar, 2016).

Similarly, researchers use data every day, in many forms, for planning, analysing, and interpreting information to inform fundamental scientific and technological decisions, with a view to advancing general knowledge. For this reason, statistics has been the foundation of research. It plays a vital role in several disciplines, including urban planning, engineering, environmental studies, applied science, natural science, social science, agriculture, medicine, and the humanities. It also serves as a significant link in developing logical and scientific thinking about the relationships of variables in research. The intelligent use of statistics provides a robust set of tools for gaining insight into the real world (Tomy et al., 2021).

Demystifying statistical tools in research has become paramount; they are the building blocks of high-quality research. This article aims to remove barriers to and fears about the use of statistical tools by researchers and to avoid the manipulation of analyses to achieve statistical significance. It is noteworthy that research grounded in statistical measures provides a strong foundation for sound judgment. Therefore, understanding the use and selection of relevant statistical tools in research cannot be overemphasised.

It implied that familiarity with statistical tools is essential to everyday operations. Everyone, irrespective of their status, needs a basic understanding of statistics to be informed, to understand issues, to make sound decisions based on available data, and to evaluate decisions that affect our lives in real-world situations.

On the global scale, the study of human-environment has assumed prominence due to the varied adverse effects of climate change and other vagaries we experience in our local environmental settings. The development of logical and scientific thinking to analyse statistical data for a fundamental decision requires a clear understanding of the research roadmap from its origin to its destination. This requires providing a clear picture, such as having good background knowledge of the study, clearly defined objectives, research questions, and the selection and application of relevant statistical tools/methods to achieve meaningful results. The application of statistical tools requires the statistical treatment of research data, beginning with collection, presentation, and analysis, and culminating in the interpretation of results in a logical sequence. Numerous statistical tools are commonly used for data analysis in research (Rajitha et al., 2020). The choice of statistical tools depends on the nature of the study, which could be parametric.

The most prominent statistical tools include charts and graphs, the mean, median, and mode, range, interquartile range, standard deviation, coefficient of variation, and dispersion. These tools can be presented or manipulated in many ways, ranging from manual calculations to the use of computer software packages. There are numerous software packages, such as the Statistical Package for the Social Sciences (SPSS), Microsoft Excel, Minitab, and the Statistical Analysis System (SAS), that are useful for interpreting results from large samples in research and for complex statistical data analysis.

Statistical data analysis is a process researchers use to convert raw data into an interpretable form to support better decision-making. Research data is any information obtained from the collection of facts or observations relating to the subject of study. It is the body of facts acquired from either measurements or observations, generated or created to validate original research findings. Datum is a singular of data that refers to an item of factual information derived from observation or measurement of a phenomenon. It is the information required to investigate a research problem after a proper study design (Mazhar et al., 2021).

2. Materials and methods

2.1. Data collection

The researcher must first decide on the type of data to collect, whether from existing data or a new data source. Even if a decision is made to use existing data, it is essential to understand how the data were collected and for what purpose, so that any resulting limitations are fully understood and judged acceptable (Gregory et al., (2005). If new data are to be collected, a careful plan must be developed, because the appropriate data analysis and the subsequent conclusions depend on how the data are collected.

Careless data collection methods will lead to errors. The error level in some data may be high, while in other cases it may be low. How good or bad a set of data depends on the following:

a) The accuracy of the observation;
b) The choice of sampling technique used;
c) How representative the samples collected are;
d) The level of error inherent in the data set.

However, six common identifiable errors can affect the quality of research. According to Rajitha et al. (2020), standard errors may arise before data collection or during the application of statistical tools. Such common errors include the following:

i) Collecting research data and thereafter trying to find out a statistical technique that can be fit in analysis.
ii) Selecting inappropriate statistical tools for the proposed analysis.
iii) Using only one statistical procedure, thereby neglecting several other methods that could be jointly applied to the available data. This leads to overlooking results that could have made a significant contribution to the study.
iv) Forcefully using statistical tools when data grossly fails to meet the assumptions upon which the tools are based.
v) Understating or overstating the importance of minor statistically significant differences.
vi) Applying the wrong significance statistical tables to interpret results.

This means that the decision on statistical tools for data analysis should be an integral part of the study from the planning stage, rather than after data collection. In other words, the type of research, the study's objectives, the research design and nature of the data, the study population, and the sample and sampling techniques are strong determinants of the statistical tools used for research data (Mweshi & Sakyi, 2020). Good initial planning in research anticipates and eliminates problems in data analysis. Good planning begins with an understanding of the scales or levels of measurement used in research. There are four basic types of measurement scales: nominal, ordinal, interval, and ratio. Each scale is better suited to certain types of variables because it enables or restricts the researcher to perform specific statistical operations or tests. According to Ofem et al. (2023), a variable is an unknown value, liable to or capable of change and marked by diversity or difference. A variable in research refers to a person, place, thing, situation, or phenomenon that the researcher is trying to measure. Variables provide focus for the study, serving as its basic framework. Therefore, faulty variable selection in research is bound to yield incorrect results or directly contribute to overall bias in the results.

The nominal scale is used to identify objects or to classify observations, using numbers as tags or labels. Data collected on a nominal scale can be presented in tabular, pictorial, or graphical form. In graphical form, a frequency distribution can be presented as a histogram, line graph, or polygon. The central tendency, a statistical measure, represents a single value of the entire dataset and is a useful way to visualise the classification. The chi-square test is used on categorical data to compare more than two categorical (nominal) variables to ascertain whether the data are significantly different from what you expect.

For example, if the researcher wants to classify a population into several categories with respect to two attributes, such as age and job performance, a chi-square test can be used to assess whether the two attributes are independent. The Cohen Kappa coefficient is a statistical tool used to measure inter

The ordinal scale addresses the ordering and ranking of data without specifying the degree of variation between them. Calculations such as frequency distributions, central tendency, and chi-square tests, which are typically performed on nominal scales, can also be applied to ordinal scales. Percentiles, quartiles, and median calculations are additional logical tests performed on an ordinal scale. Spearman's rank-order correlation is another helpful measure that assesses correlation between variables when the data are non-numeric and sufficient for ranking (Ofem et al., 2023).

An interval scale provides more information than an ordinal scale because it indicates that differences between values are equal. This means that the interval scale is an extension of the ordinal scale, with equal intervals from low to high. However, the interval scale has a drawback: the absence of an actual zero point. Most statistical operations can be performed using interval scales. Mean, standard deviation, which is a measure of central tendency, and variability can be applied to this scale of measurement.

Ratio scales are the most informative scales because they provide rankings, assure equal differences between scale values, and have a valid zero point (Ofem et al., 2023). This implies that nominal, ordinal, and interval scales are embedded within ratio scales. Ratio scales are the most versatile scale, as all statistical tools can be applied to it. The most common of these tests are the parametric tests, regression, t-tests, Analysis of Variance (ANOVA), Analysis of Covariance (ANCOVA), Multiple Analysis of Variance (MANOVA), and Pearson correlation.

In general, each scale is suitable for specific types of variables that enable or limit us to perform specific statistical tests or operations. However, there are rules and guidelines for researchers to construct a composite scale by converting an ordinal scale to an interval scale. Researchers should select an appropriate statistical test based on the assumptions of the test they intend to use, the nature of the construct being measured, and the distribution of the data.

The summary of measurement scales is presented in Table 1.

Table 1: Summary of scales of measurement

Measurement scale	Statistical tool
Nominal	Frequency distribution Central tendency Chi-square test Kappa coefficient
Ordinal	All tests for nominal scale, Spearman's rank-order Correlation
Interval	All tests for ordinal scale, ANOVA, t-test
Ratio	All statistical tests can be performed

2.2 Choice of Appropriate Statistical Tool

Appropriate statistical tools can be selected to analyse the data for meaningful interpretation, depending on whether the researcher is interested in describing and summarising the information without inferring population parameters from the sample, or in generalising the findings to the population from which the sample was drawn (Dawit, 2020). If the intention is only to describe and summarise, then appropriate descriptive statistical tools should be selected for the analysis. Still, if the purpose is to draw inferences about the population, appropriate inferential statistics should be used.

Table 2 presents an overview of selected descriptive statistics and when to use them. These encompass the most frequently used descriptive tools in statistical analysis.

Table 2: Descriptive Statistical Tools and When to Use Them

Statistical Tools	When to use
Percentage (10%)	Used to determine the proportion in one-hundredths: A proportion stated in terms of one-hundredths that is calculated by multiplying a fraction by 100
Measures of Central Tendency: a. Mean	Arithmetical means are computed when: Greatest reliability is wanted. It usually varies less from sample to sample drawn from the same population. Other computations, as finding measures of variability are to follow.

	The distribution is symmetrical about the centre and especially when it is approximately normal. The researcher wishes to know the centre of gravity of a sample.
b. Median	Compute median when: i. There is no sufficient time to compute the mean. ii. Distributions are markedly skewed. These include the case in which one or more extreme measurements are at one side of the distribution. iii. The researcher is interested in whether cases fall within the upper or lower halves of the distribution and not particularly on how far from the central point. iv. An incomplete distribution is given.
c. Mode	Compute mode when: i. The quickest estimate of central value is wanted. ii. The researcher wishes to know the most typical case.
Measures of Dispersion a. Standard deviation and Variance	Use standard deviation when: i. Greatest dependability of the value is wanted. ii. Further computations that depend on it are likely to be needed. iii. Interpretation related to the normal distribution is desired.
b. Semi-Interquartile Range	Use semi-interquartile range when: i. The median is the only statistic of central value reported. ii. The distribution is truncated or incomplete at either end. iii. There are few very extreme scores or there is an extreme skewness. iv. The researcher wants the actual score limits of the middle 50 percent of the cases.
c. Range	Use range when: i. The quickest possible index of dispersion is wanted. ii. Information is wanted concerning extreme scores.

Table 3 summarises some inferential statistical tools and their appropriate applications, taking into account the essential assumptions of each test.

Table 3: Inferential Statistical Tools and When to Use Them

Statistical Tools	Statistical Assumptions	When to use
Comparison and relationships between numbers: Correlation:		It is used when the degree of relationship between two variables is required.
i. Simple Correlation	A linear relationship exists between the variables. The variables are normally distributed (Normality). No significant outliers exist.	The correlation coefficient measures the degree of linear association between two variables.
ii. Partial Correlation	The dependent and independent variables, as well as the control variables, must be measured on an interval or ratio scale. There should be a linear relationship between all pairs of variables. The data should contain no significant outliers, as partial correlation is sensitive to outliers, which can distort results. The variance is consistent across levels of variables (Homoscedasticity).	Use partial correlation when the degree of relationship between two variables is of interest in a situation with three or more variables, holding one or more variables constant while allowing the others to vary.
iii Multiple Correlation	Independent variables should not be highly correlated with one another to avoid inflated errors and difficulties in interpreting results (Multicollinearity). The residuals should have constant variance across all levels of the independent variables (Homoscedasticity). There should be a linear relationship between dependent and independent variables.	Use multiple correlations when the degree of relationship between a dependent variable and an optimally weighted combination of two or more independent (predictor) variables are involved.

iv. Canonical Correlation. This is a multivariate technique use to investigate two different sets of data from a single population. The goal is to reveal combinations of the first dataset and the linear combinations of the second dataset.	There must be adequate sample size since small sample size can affect reliability. The data should be collected through random sampling method. The relations between variables in both sets should be linear. Variables in both sets should follow a multivariate normal distribution. The variance-covariance matrices of the variable sets should be relatively consistent to support reliability (Homogeneity).	Used to extract linear combinations for the criteria and linear combinations for the predictor in such a way that when these two sets of liner combinations are correlated, a maximum amount of correlation between these two datasets is obtained. It simultaneously correlates several independent variables and several dependent variables.
Regression:		It is used to model how one dependent variable varies with the values of a set of independent variables. The dependent variable must be interval or ratio, whereas the independent variables may be on any scale. Researchers should note that some categorical or ordinal independent variables may be recoded according to rules or guidelines specified in the analysis.
i. Simple/ Multiple Linear Regression	The dependent and independent variable(s) should be measured on an interval or ratio scale. The residuals should be approximately normally distributed. There should be no significant outliers.	It is used to investigate causal relationships between variables, to measure the effect of one variable on another, and to predict values of the dependent variable given values of the independent variable(s).
ii. Logistic Regression	Unlike linear regression, logistic regression does not assume linearity between predictors and the outcome variable, normality of the outcome variable or its residuals, or homoscedasticity of the errors. It requires the response variable to be binary for binary logistic regression or approximately categorical for multinomial logistic regression. There should be no repeated measurements or related cases within the datasets (Independence of observations).	It is a nonlinear regression model that predicts the presence or absence of an outcome based on predictor variables. Instead of predicting the value of the dependent variable from the independent variable, logistic regression estimates the probability of the dependent variable occurring as a function of the independent variable.
Factor Analysis	These include a large sample size, factorability of the correlation matrix (i.e., correlations of 0.3 or above), and linearity, as factor analysis is based on correlations.	Used to reduce the number of variables to a manageable level. It could be used for subsequent analysis by forming combinations of the original measured variables that account for as much of the original variance as possible, while also facilitating more straightforward naming or interpretation of the results. It is also used to analyse patterns of intercorrelation among many variables, to isolate the dimensions that account for these patterns, and, in a well-designed study, to draw inferences concerning the psychological nature of the constructs represented by these dimensions.
Chi-square	It requires a categorical variable that is nominal or ordinal, with mutually exclusive categories, such that each subject contributes to exactly one cell. Random sampling is preferred. Data	This is a measure of squared deviations between observed and theoretical frequencies in the categories or cells of a table, used to determine whether these deviations are attributable to sampling

	should be presented as frequency counts, not as percentages.	error or to interdependence or correlation among the frequencies. It involves comparing frequencies across one or more response groups and is particularly useful for tables of Yes-No frequencies.
Cluster Analysis	It assumes a representative sample, relevant and minimally correlated variables, and the presence of natural groupings that can be meaningfully separated into several clusters using statistical criteria.	Used to classify subjects with similar characteristics according to the values of variables measured. The purpose of classification is to find out whether variables can be formed into any natural system of groups.
Discriminant Analysis	It assumes multivariate normality, homogeneity of variance-covariance matrices across groups, independent of observations, no significant multicollinearity among predictors, and linear relationships between predictor variables within each group.	Used to identify those variables which best discriminate between known groups of subjects. The test results could improve the practical assignment of new subjects to known groups based on their values of the discriminating variables. Discriminant analysis is useful for classifying cases based on predictive functions; it assesses whether cases are correctly classified and examines how the derived groups differ. It helps us discard variables that are not significantly related to group discrimination and interpret the discriminant dimensions that distinguish groups based on their discriminant loadings.
Analysis of Variance (ANOVA)	Assumptions for both ANOVA and MANOVA include normality of the independent variables, independence of observations, and homogeneity of variances across groups.	This is used to test one or more (null) hypotheses that the means of all groups sampled come from a population with equal means and differ only because of sampling error. The F-test is a technique used in Analysis of Variance, which compares the Between-group variance to the Within-group variance.
Analysis of Covariance (ANCOVA)	It assumes that the dependent variable and covariate(s) are continuous, that the dependent variable has two or more categories, and that observations are independent, with no significant outliers.	Comparison of three or more groups' independent means for the significance of their differences on two or more dependent variables based on pre-test and post-test mean data.
z-ratio or t-ratio	Z-test assumes that the sample is random, representative and consists of independent, continuous data that are approximately normally distributed.	These are used to test the hypothesis that two samples come from populations with the same mean and differ only due to sampling error. The z-test applies to populations with or without equal variances, whereas the t-test assumes equal population variances.
Time Series Analysis	It assumes stationarity, in which statistical properties such as the mean, variance, and autocorrelation remain constant over time.	Used to investigate patterns and trends in a variable measured regularly over a period of time. It could also be used to identify and adjust for seasonal variations.

From Table 3, it is evident that the various statistical tools for Comparison and relationship analysis are based on basic assumptions that must be satisfied for the results to be meaningful (Dawit, 2020).

2.3 Parametric and Non-Parametric Tests

Parametric tests involve variables that have been assigned numerical values according to specific rules. The assigned numerical value describes the characteristics of the population of interest under certain assumptions and is meaningful only when those assumptions hold. Parametric assumptions include normality, the absence of measurement error, the scale of measurement, and independence across datasets. In the same way, a parametric hypothesis test assumes that observed data are distributed according to distributions of a well-known form as in normal and binomial, up to some unknown parameter(s) on which we want to make inferences like the mean, or the success probability (Claudia, 2019).

Nonparametric tests focus on the relative ranks of sample observations rather than their numerical values. Nonparametric tests are helpful when the dataset is not numeric but can be ranked in order of importance. Nahm (2016) recorded that nonparametric tests are the alternative methods available, because they do not require the normality assumption. Nonparametric tests are the statistical methods based on signs and ranks. To a reasonable extent, non-parametric tests are the reverse of their parametric counterparts.

Table 4: Non-Parametric Tools and Conditions for Using Them

Purpose of test	Type of Data	No. of Groups	No. of Dependent Variables	Statistical Tool
Differences between two frequency distributions	Nominal	≥2	≥1	chi-square
Two correlated means to indicate the direction of differences between pairs of scores.	Ordinal scale when samples are matched	2	1	Sign test
Finding the magnitude of the difference between two pairs.	Ordinal-scaled variables	2	1	Wilcoxon’s matched - pairs signed rank test
Showing the significance of the difference between pairs of scores	Ordinal rather than interval	2	1	Mann-Whitney U test

Table 4 presents an overview of nonparametric statistical tools for comparison, relationships, and conditions for their use.

Table 5: Parametric Tools and Conditions for Using Them

Purpose of the test	Type of Data	No. of Groups	No. of Dependent Variables	Statistical Tool
Test of significance: Differences between means of two groups obtained from post-testing two randomly composed groups of subjects.	ratio data	2	1	t-test of independent means and critical ratio (Design: True experiment). Assumptions include a normal distribution of the data, independent or paired samples depending on the test type, and equal variances across groups for the independent-samples test.
Test of significance Differences between means of two groups obtained from post-testing two randomly composed groups of subjects, except that the two mean comparison is done with the gain (post-test/pre-test) scores.	ratio data	2	1	t-test for selected or non-independent means and critical ratio (Design: Quasi-experimental). The t-test compares the calculated value to a critical value from a t-distribution table based on the degrees of freedom to reject or accept the null hypothesis.
Comparison of two, three or more groups' independent means for the significance of the	The dependent variable must be continuous (ratio	≥2	1	ANOVA or MANOVA (Design: True experimental) ANOVA provides a method for testing

differences on the dependent variables	data). It means data should be measurable quantities like temperature, height, and distance.			hypothesis that the means of many sampled populations are equal.
Comparison of three or more groups' independent means for the significance of the differences on two or more dependent variables.	The independent variable must be categorical data. It defines levels or groups being compared, like experimental conditions or racial groups.	≥2	≥2	ANOVA or MANOVA (Design: True experimental). MANOVA is the extension of ANOVA that tests multiple continuous dependent variables at the same time against categorical independent variables. Assumptions include the use of large samples to address data anomalies, linearity, the absence of homoscedasticity, and the homogeneity of variances across groups.
Comparison of three or more groups' independent means for the significance of their differences on two or more dependent variables based on pre-test and post-test mean data.	ratio data	≥2	≥2	Scheffe test (Design: True or Quasi-experimental). The Scheffe test is an a posteriori comparison similar to performing a t-test on every pair of groups. It is used to determine which groups differ. It is suitable for many types of contracts but with less power, and can accommodate unequal sample sizes. The Scheffe test is based on the F-table.
Comparison of three or more groups' independent means for the significance of the differences on two or more dependent variables based on pre-test and post-test mean data, in which F-cal is significant.	ratio data	≥2	≥2	Turkey test (Design: True or Quasi-experimental). The Turkey test is more powerful for pairwise comparisons but requires equal sample sizes and is less flexible.
Relationships between variables	Interval or ratio data	≥2	≥2	Product-Moment Correlation (Design: Correlational)

Table 5 presents an overview of some parametric statistical tools for comparison, the relationships, and conditions for using them.

However, Shukla et al. (2024) confirmed that for virtually every parametric type of test, there is at least one equivalent non-parametric test and in general, parametric tests fall into three major categories:

i) Test of differences between two groups (Independent samples);
ii) Test of differences between dependent groups; and
iii) Test of relationships between variables.

Table 6: Test Categories in Basic Statistics and their Statistical Tools/Methods

Purpose of the test	**Nature of Data**	**Parametric Test**	**Non-Parametric Test**
i. Differences between two groups (Independent samples)	Differences between the means of two groups	t-test	Mann-Whitney U test,
	Multiple groups	ANOVA or MANOVA	Kruskal-Wallis analysis of ranks and the Median test
ii. Differences between dependent groups	Comparison of two variables measured in the same sample,	t-test of dependent samples	Sign test, Wilcoxon's test, chi-square if data is dichotomous

	e.g. pre-test and post-test scores		
	More than two variables that were measured in the same sample	ANOVA or MANOVA	Friedman's two-way Analysis of Variance; Cochran Q-test for proportions and frequencies
iii. Relationships between variables	Relationships between variables Two variables of interest are categorical in nature, like ‘passed’ or ‘failed’ by ‘male’ or ‘female’	Product-Moment Correlation	Spearman Rho, Kendall Tau, chi-square, Phi coefficient, Fisher's exact test

Table 6 provides clearer information on the test categories in basic statistics and their corresponding statistical tools/methods.

It could be deduced from Tables 4, 5 and 6 that the decision on an appropriate statistical tool or treatment of research data is dependent on several attributes, namely:

a) Purpose of the test that shows the differences between two groups (dependent or independent), or relationships between variables;
b) Nature/type of data which could be nominal, ordinal, interval or ratio;
c) Number of groups;
d) Number of dependent variables; and
e) The Research design may be truly experimental, quasi-experimental or non-experimental.

Researchers must, therefore, ensure that the aforementioned attributes are carefully considered before selecting a statistical tool for any given data analysis.

5. Conclusion

Selecting an appropriate statistical technique is essential in any research, as failing to do so can lead to erroneous inferences and conclusions. When deciding on the choice of statistical technique for the analysis of any dataset, the researcher must carefully consider the purpose of the study, the research design, the scale of measurement in which the dataset falls, and the number of groups involved (independent and dependent), the statistical assumptions and the final selection of relevant statistical tests. It is recommended that researchers select statistical tools that align with the assumptions underlying the measurements.

References

Ali, Z., & Bhaskar, S. B. (2016). Basic statistical tools in research and data analysis. Indian Journal of Anaesthics, 60(9): 662-669. DOI: 10.4103/0019-5049.190623. Erratum in: Indian Journal of Anaesthics. 60(10):790. PMID: 27729694; PMCID: PMC5037948.

Claudia, A. (2019). Hypothesis Testing. In *Encyclopedia of Bioinformatics and Computational Biology*, (Eds): Ranganatha, S., Gribskov, M., Nakai, K., & Schönbach, C., Academic Press, Pages 691-697 https://doi.org/10.1016/B978-0-12-809633-8.20356-7.

Dawit, D. A. (2020). An Overview of Data Analysis and Interpretations in Research. *International Journal of Academic Reseach in Education & Review*, 8(1): 1-27. http://www.academicresearchjournals.org/IJARER/Index.htm

Gregory, R., Fischhoff, B. & McDaniels, T. (2005). Acceptable Input: Using Decision Analysis to Guide Public Policy Deliberations. *Decision Analysis*, 2(1):4-16. https://doi.org/10.1287/deca.1050.0035

Mazhar, S. A., Anjum, R., Anwar, I. A. & Khan, A. A. (2021). Methods of Data Collection: A fundamental tool of Research. Journal of Integrated Community Health, 10(1): 6-10. DOI: https://doi.org/10.24321/2319.9113.202101

Ofem, B. I. & Mchi, A. A. (2023). Variable Conceptualisation and Measurement in Environmental Research, International Journal of Methodology, 2(1): 2-11. DOI : https://doi.org/10.21467/ijm.2.1.5991

Mweshi, G. K., & Sakyi, K. (2020). Application of Sampling Methods for the Research Design. *Archives of Business Research*, 8(11). 180- 193. DOI: 10.14738/abr.811.9042.

Nahm, F. S. (2016). Nonparametric statistical tests for the continuous data: the basic concept and the practical use. *Korean Journal of Anesthesiology*, 69(1):8-14. DOI: 10.4097/kjae.2016.69.1.8. Epub 2016 Jan 28. PMID: 26885295; PMCID: PMC4754273.

Rajitha, N. & Madhav, B. K. (2020). The Choice of an Appropriate Statistical Technique for Research. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3560867 or http://dx.doi.org/10.2139/ssrn.3560867

Shukla, R. M., Pandey, Y., Faisal, S., Ali, M. and Muzamil, M. (2024). Parametric and Non-Parametric Analysis**.** In *Advances in Agricultural Research Methodology (Volume 2, First Edition),* Eds.: Gopichand, S., Pavan, S., Amit, K. P., & Simeet, S. R**.** SP Publishing, India, Pages 260 – 272. https://www.researchgate.net/publication/395451689_Advances_in_Agricultural_Research_Methodology_Volume_2

Tomy, L., Chesneau, C. & Madhav, A. K. (2021). Statistical Techniques for Environmental Sciences: A Review Mathematical Computations and Applications, 26(4): 74; https://doi.org/10.3390/mca26040074.