

<https://doi.org/10.33003/jaat.2026.1202.06>

USE OF REGRESSION IN SOCIAL SCIENCE RESEARCH: BEYOND LINEAR MODELS

¹Oyebamiji, B. A., ¹Oyebamiji, T. T., ¹Sodiya, C. I., ²Sanusi, R. A., ¹Abudusalami-Saghir, P. B., ¹Okunade, J. A., ¹Babalola, T. A., ³Akintobi, E. T.

¹Department of Agricultural Extension and Rural Development, Federal University of Agriculture, Abeokuta, Ogun State.

²Department of Agricultural Economics and Farm Management Federal University of Agriculture, Abeokuta, Ogun State

³Department of Animal Genetic Resources (AnGR), National Biotechnology Development Agency (NABDA)

Correspondence author’s e-mail: oyebamijiba@funaab.edu.ng, +234-806-124-2039

ABSTRACT

Research aims to systematically acquire knowledge through data collection, analysis, and interpretation to answer questions and enhance understanding. Therefore, this study seeks to provide insight into the use of regression in social science research, extending beyond linear models. Social research in agricultural extension aims to understand the dynamics, behaviours, and felt needs of rural communities to improve resource distribution and innovation, ultimately increasing agricultural productivity and informing policy decisions. A robust methodology was employed to establish a clear research framework, sampling methods, and data analysis techniques, ensuring the reliability and validity of the results. Social science research outcomes established relationships between dependent and independent variables, focusing on estimating or predicting outcomes as well as managing the effects of independent variables. Regression analysis is useful and powerful for interpreting, validating, reporting, and explaining covariates. This study highlights the uniqueness of regression compared to other test tools, and the study outlines the basic criteria for selecting appropriate regression models. In conclusion, it underscores the assumptions behind both linear and nonlinear regression and emphasizes the advantages of regression analysis over other testing tools in social science research. Additionally, the study provides insight into various types of nonlinear regression, including their strengths and limitations. Therefore the study recommends that researchers should strive to thoroughly understand test tools, seek further understanding, and employ both linear and nonlinear regression techniques for accurate data interpretation.

Keywords: Data, Regression, Research, Social research, Tool

INTRODUCTION

Research is an intensive and purposeful search for knowledge and understanding of social and physical phenomena. This systematic process involves collecting, analyzing, and interpreting data to answer unresolved questions, advance knowledge, promote progress, and foster effective relationships with the environment in order to solve scientific problems and achieve goals (Reed *et al.*, 2021). Research is a systematic, methodical process aimed at discovering, interpreting, and revising facts, theories, and applications in various fields to improve knowledge and understanding of the particular subject matter (Ghanad, 2023). Social research is more essential than other types of research because it focuses on the scientific foundation for understanding human behavior, providing knowledge for social progress, policy relevance, and sustainable development, ensuring scientific advancements meet societal needs (Confraria *et al.*, 2024).

Social research is a scientific endeavour involving logical and systematic techniques to discover new facts, verify and test old ones, analyze their sequences, and develop new scientific tools for reliable and valid human behavior studies (Dalong, 2022). Social research in agricultural extension aims to understand farming community dynamics, behaviours, and needs to improve innovation dissemination, education, and resource provision for agricultural productivity and sustainability. Figure 1 revealed the chains in research process and it shows the role of each components in the study. It involves formulating research questions, proving answers, developing hypotheses, conducting literature surveys, conducting methodological processes, analyzing data, presenting results, interpreting data, summarizing, concluding, and providing recommendations.. Furthermore, Creswell and Creswell (2018) emphasizes that objectives, a systematic approach, replicable, and critical analysis are the major characteristics of research.

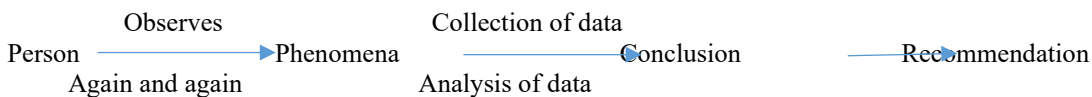


Figure 1: Conceptual Illustration of Social Research

Source: Zaid and Tsageem, 2022

Figure 1 illustrate the order in social research, regression is a fundamental technique or tool to analyze the data collected to promote decision making or policy. The techniques are ANOVA, correlation/unit weighting, classification methods, Structural Equation Modeling (SEM), and modern machine-learning approaches. Regression is a comprehensive framework for hypothesis testing, relationship quantification, and outcome prediction, applicable to continuous, binary, and count data. Regression demonstrate superior performance in prediction scenarios, especially when adjusted R^2 values are low, indicating greater robustness and interpretability. However, it has been observed that researchers misuse regression. Analysts often misunderstand the purpose of regression, whether descriptive, predictive, or causal. Researchers incorrectly interpret coefficients as causal effects. They advise selecting modeling approaches based on the substantive question, prioritizing this over data structure and assumptions.

Key concepts in social science research

Level of measurement: It is essential to understand the essence and importance of measuring variables, because it forms the basis for data analysis. Several literatures pointed to Steven's classification system of measurement, which consists of nominal, ordinal, interval, and ratio scales. Nominal scales are used to measure categorical data, while ordinal scales rank elements based on their grades or scores. Also, interval scales measure values equidistant from adjacent attributes, allowing comparison of attributes and central tendency measures, and ratio scales possess all qualities of nominal, ordinal, and interval scales but have a "true zero" point indicating non-availability of the underlying construct (Matthews, 2018; Thomas, 2018; Williams, 2021). Literatures affirmed that the degree of variable measurement plays a critical role in the development of hypotheses and suitable statistical methods and in improving the validity and reliability of research findings (Sürücü and Maslak, 2020). Hypotheses guide data analysis by providing testable predictions, selecting statistical methods, and interpreting results, aligning with research objectives for valid conclusions (Ayre and Scally, 2014).

Data analysis: This is a systematic method of using statistical and logical techniques to describe, illustrate, and evaluate data, aiming to uncover useful information, suggest conclusions, and aid decision-making (Calzon, 2023). The selection of statistical analysis in social sciences often relies on a classification of measures into continuous and discrete categories.

Continuous and categorical variables: Continuous variables are quantitative measures that can take infinite

values within a given range, are often measured rather than counted, and can assume any value along a continuum, while categorical variables are numerical values that can only take a finite number within a range and can be nominal, ordinal, or dichotomous (Thomas, 2022).

Homoscedasticity and Heteroscedasticity: McNeish and Silverman (2017) and Oscar *et al.* (2019) emphasizes the feature of a regression model where the variance of the residuals (errors) is constant across all levels of the independent variable, known as homoscedasticity, while heteroscedasticity is a statistical concept where the variance of residuals (errors) changes with the level of the independent variable (s), resulting in a systematic pattern or bias around the regression line.

Over-fitting: It occurs when a model is too complex, has inadequate regularization, or has irrelevant interaction and learns the noise in the training data rather than underlying patterns. It results in unreliable predictions, increases the risk of overconfidence, and makes it difficult to interpret model results (Aliferis and Simon, 2024).

In social science research, data analysis is an essential component, and regression modeling is among the data analysis methods that social scientists utilize the most frequently. Regression analysis is an important tool in social sciences research to analyze and express relationships between variables, assess association strength, and guide causality inference (Jilcha, 2020; Ali and Younas, 2021; Calzon, 2023). Regression analysis is very essential in social research, as it promotes understanding, quantifying, and predicting relationships between social variables. In addition, it offers more profound understanding of their influence on a dependent variable. When most scientists discuss regression, they typically think of the linear regression model. Field (2021) emphasized that independent variables, also known as predictors or explanatory variables, are manipulated to observe their effect on dependent variables, or regressed variables. Statistics is a vital tool for descriptive or inference purposes.

RESEARCH METHODOLOGY

There are different statistical tools used to analyze different hypotheses; in social science research and the major tools are Chi-square, Pearson Product Moment Correlation (PPMC), Spearman-rho, t-test, ANOVA, and regression (Poston *et al.* 2023). Literature for social scientists integrates several statistical methods such as regression, categorical predictors, polynomial and interaction terms, robust and symmetric regression, longitudinal and piecewise models, and dichotomous outcomes (Jocelyn, 2023). Recent research studies focus on regression implementation, providing applied workflows for estimation, diagnostics, and presentation that align with

contemporary social research standards. Regression is a fundamental tool in statistics and seeks to improve Ordinary Least Squares (OLS) technique by engaging different methods to fit social-science data, address assumption violations, and enhance prediction and inference. It systematically solves research problems by developing indices and appropriate techniques. It enhances reliability, validity, replication, and critical evaluation of findings, especially for researchers lacking a strong statistical background (Mohajan, 2020; O'Connor and Joffe, 2020). A distinct methodology in regression studies provide clear research framework, sampling method, instruments, and data analysis techniques, ensuring consistency and validity of findings. A good methodology summarizes data sources, collection techniques, and analysis methods, ensuring transparency and credibility, aligning research objectives with appropriate techniques, and justifying research designs, instruments, and procedures. In addition, methodology is the engine room of any social science research (Mohajan, 2020). It help design the research framework ad guides the choice of the most accurate and applicable statistical techniques, model specification and interpretation of result The level of measurement significantly influences research methodology, influencing statistical methods and data analysis approaches (Jilcha, 2020; Calzon, 2023).

Regression model

Regression analysis examines the relationship between dependent and independent variables, with an independent

variable being an input or factor influencing the dependent variable (Ramdurg, 2020; Ali and Younas, 2021). Regression analysis has four primary purposes: description, estimation, prediction, and control (Ramdurg, 2020; Younas, 2023). It explains the relationship between dependent and independent variables, estimates dependent variable values using observed values, predicts outcomes, and controls the effect of one or more independent variables while investigating the relationship between them. It captures complex relationships and indicates the direction of influence (Reza *et al.*, 2018; Algamal, 2020). It is an essential statistical method in social science research that reveals causal inference, predictive modeling, and influencing policy decisions, program evaluations, and theoretical understanding of social phenomena. Regression functions as an integrating language, connecting theory, various data types, complex structures, and causal inquiries within a unified framework, while requiring careful consideration of assumptions and interpretations.

Ds *et al.* (2018) and Hama *et al.* (2020) affirmed that it is used for selecting, coding, interpreting, diagnosing, validating, revising, reporting, and explaining covariate differences. Shmueli (2010) categorized regression into linear and non-linear based on the nature of the variable, types of relationships, and number of regressed and repressor variables. Table 1 emphasized the uniqueness or advantages of regression over other statistical tools as shown below (Frost, 2020).

Table 1: Uniqueness of Regression over other test tools in research

Regression	Other research test tools
It can control for multiple variables and explore causal relationships.	Correlation analysis does not control for other variables or provide causal insights.
It provides estimates of the strength and direction of these relationships.	It does not provide the same level of detail about relationships between variables.
It can handle both continuous and categorical variables.	Chi-square is mainly used for categorical data to test associations between variables. It does not quantify the strength or nature of relationships.
It quantifies the impact of predictors.	
It directly examines and quantifies relationships between variables.	Factor Analysis: It is used for data reduction and identifying underlying structures, not for examining relationships between specific variables.

Linear Regression

Linear regression is a crucial statistical and machine learning tool for identifying relationships between variables and making informed predictions by fitting a linear equation to observed data and it is a powerful tool for understanding relationships and making predictions. It has various applications, including predictive modeling, trend analysis, forecasting, data visualization, and relationship

identification (Petropoulos *et al.*, 2022; Stevens, 2023). Linear regression can either be simple or multiple, depending on the number of independent variables. For example; a study testing the relationship between the cost of production of cassava and income. Maiwada *et al.* (2013); Maestrini *et al.* (2016) and Maxwell (2016) identified the five main assumptions underlying linear regression models that must be satisfied: there is a linear

relationship between the independent variables and the dependent variable, the value of the explanatory variables is decided by the researcher, explanatory variables should be measured without any experimental errors, the independent variables are normally distributed, the subpopulation of the dependent variable is homogeneous in nature, there is no multicollinearity, and the mean value of the subpopulation of 'y' is expected to lie on a straight line.

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_kX_k + \text{error}$$

Linear regression is frequently utilized by researchers due to its simplicity, interpretability and effectiveness. Despite potential violations of its assumptions, it cannot capture nonlinear relationships. It is sensitive to outliers, and

heteroscedasticity can cause biased standard errors. Correlations and high multicollinearity can compromise statistical tests and confidence intervals. Failing to capture these complexities can result in a poorly represented data structure. Regression coefficient is not suitable for complex models, categorical variables with multiple levels, time series data, temporal dependencies, or large data sets due to overfitting. Researchers overuse or abuse some linear regression tools like Chi-square and Spearman Rho due to their simplicity but do not consider the assumptions, deeper interpretation, and sensitivity of the data quality, whereas non-linear regression has the capacity to perform the operation.

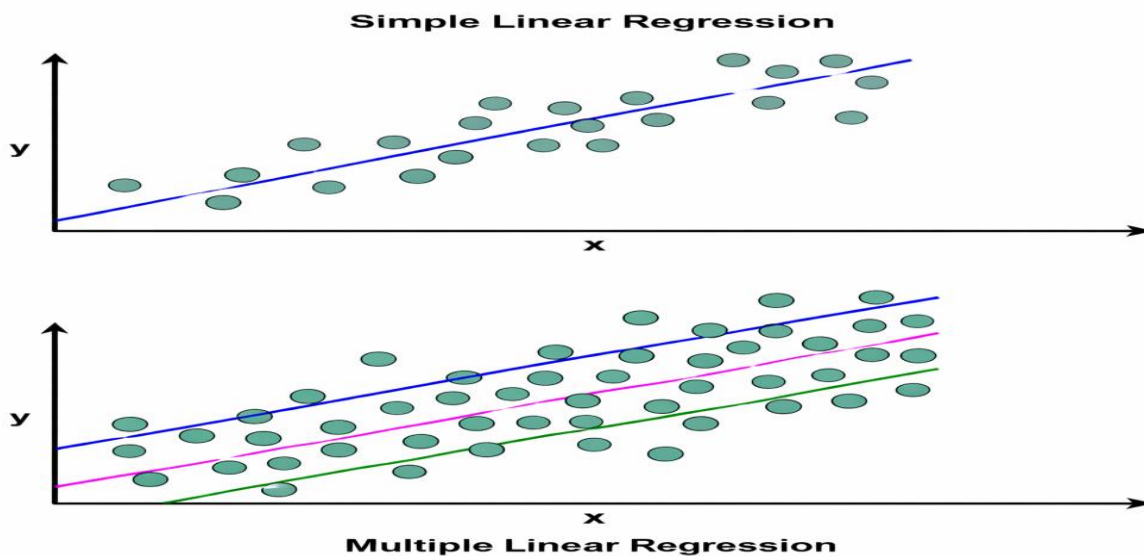


Figure 2:

Simple linear and Multiple regression

Source: <https://www.researchgate.net/figure/Simple-and-Multiple-Linear-Regression>

Non-linear regression used in social research

Non-linear regression models the relationship between dependent and independent variables, as it is a powerful tool for modeling complex variables, requiring understanding of model fitting and parameter estimation techniques to uncover meaningful insights (James *et al.*, 2021).

Types of non-linear regression commonly use in social research

Binominal regression: It is known as binary logistic regression, is a statistical method use to predict the outcome

of a dichotomous dependent variable. Figure 3 illustrated the binary outcome variables, offers insights into predictor relationships handle continuous and categorical predictors, and calculates easy-to-interpret odds ratios. Example: testing a relationship between accessibility of extension services and utilization of extension services (utilized and not utilized). Adjei and Karim (2016) highlighted three (3) main assumptions of binominal aggression, as binary regression model assumes a dependent variable with two possible values, a large sample size, and no multicollinearity between explanatory variables.

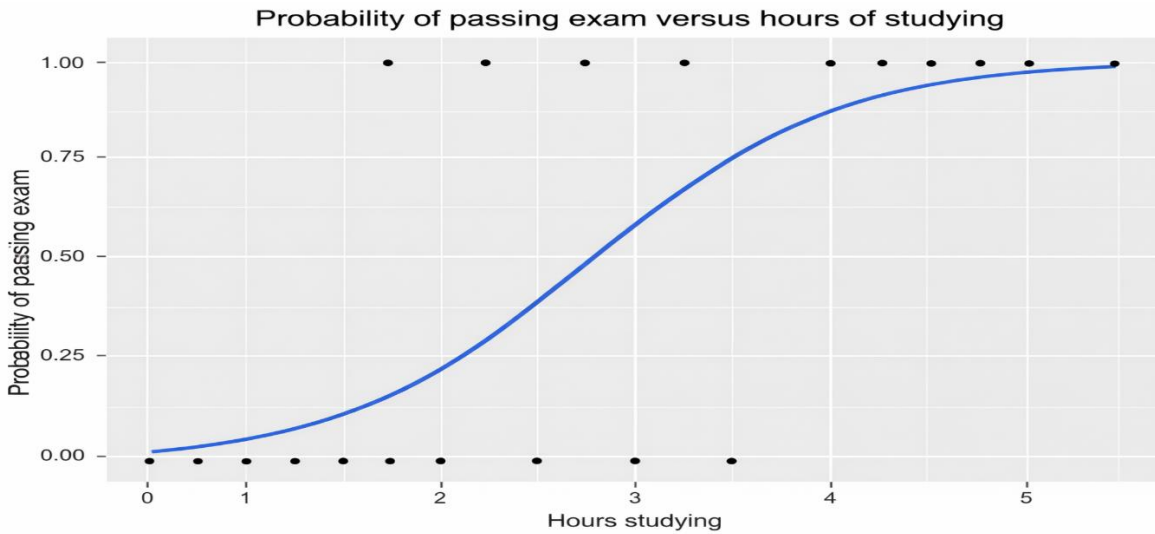


Figure 3: Binominal regression
 Source: https://en.m.wikipedia.org/wiki/File:Exam_pass_logistic_curve.svg

Polynomial regression: It is modified, even though it uses a linear model as an initial estimator. According to Pistol and Bucea-Manea-Tonis (2017), polynomial regression can be used in nonlinear analysis that involves separable data. It shows the connection between independent and dependent variables, but they are identified by the n th degree. The polynomial models can be used in those situations where the relationship between dependent and explanatory variables is curvilinear and consists of the power function features of the explanatory variables that are usually more than one variable. It is a crucial tool for modeling nonlinear data relationships, providing valuable insights and predictions in various fields by understanding its key concepts and analysis steps (James *et al.*, 2021). For example: testing a relationship between labour contribution and arable crop production. Helwig (2017) identified the assumption of polynomial regression that the dependent variable can be interpreted by a set of k independent variables ($x_i, i = 1$ to k), the relationship between the dependent variable and any independent variable is linear or curvilinear, independent variables are independent of each other, the errors are independent, normally distributed with a mean zero and a constant variance.

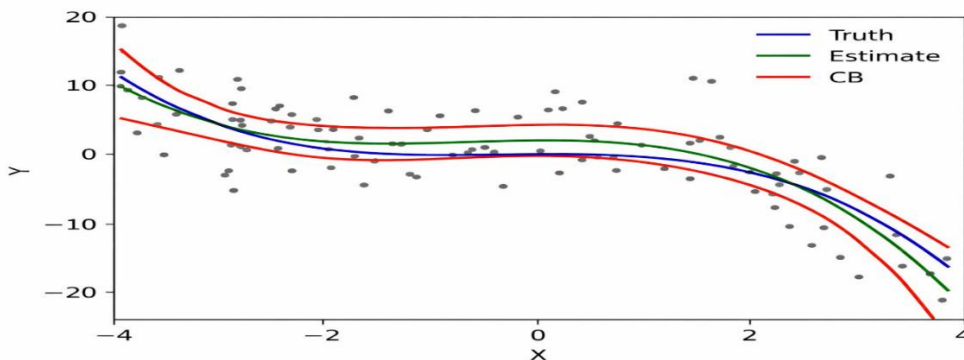


Figure 4 Polynomial Regression
 Source: https://en.wikipedia.org/wiki/Polynomial_regression#/media/File:Polyreg_scheffe.svg

Ridge regression: It is also known as Tikhonov regularization, or machine learning analysis. According to Hastie (2009), it is a method used to analyze multiple regression data that exhibits multicollinearity. This means that as one independent variable changes, others can change with it. Ridge regression effectively manages multicollinearity and prevents overfitting in regression models by adding a regularization term to the ordinary least squares objective for more reliable coefficient estimates. Example: relationship between multiple predictors (age, income, education level, hours worked) and a continuous outcome variable (average income or production). James *et al.* (2021) indicated the assumption of Ridge regression models that high-correlated

independent variables exist, over-fitting due to complexity, high-dimensional datasets, and regularization parameter (lambda-λ) for irregularities.

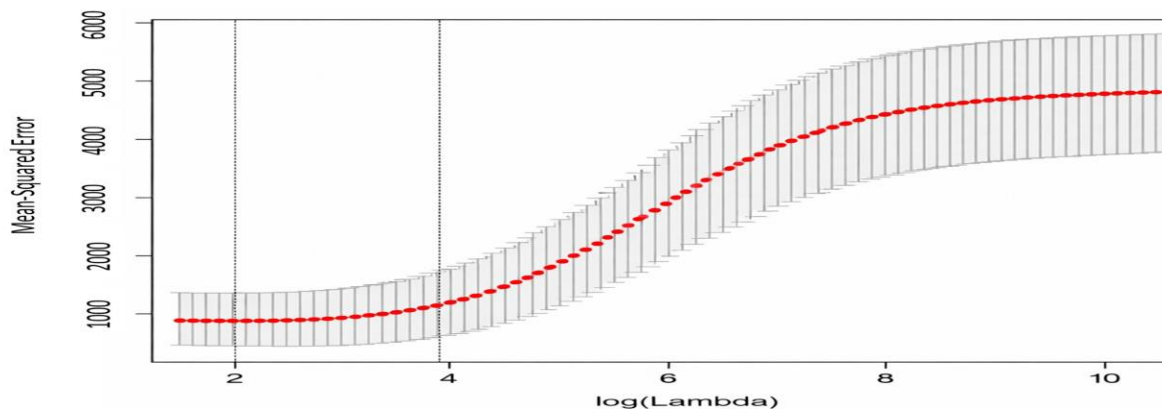


Figure 5: Ridge regression. Source: <https://www.statology.org/ridge-regression-in-r/>

Multinomial regression: This is known as multinomial logistic regression. It is a statistical method used to predict outcomes of categorical dependent variables, particularly in fields like social sciences, healthcare, and marketing. It is applicable when the dependent variable is categorical with more than two levels with continuous and categorical predictors (Hosmer *et al.*, 2013). For example: testing a relationship between socio-economic characteristics of artisanal fishers' and livelihood diversification (categorical). Agresti (2018) identified that the dependent variable is categorical with more than two categories, the independent variable can be continuous or categorical, and independent variables have the capacity to predict the probabilities of the different outcomes as the main assumption of multinomial regression.

$$\text{Log} \left(\frac{P(Y=K)}{P(Y=\text{reference})} \right) = \beta_{0k} + \beta_{1k}X_1 + \beta_{2k}X_2 + \dots + \beta_{pk}X_p$$

Where: YYY is the categorical dependent variable with k categories.

β_{0k} is the intercept for category k.

$\beta_{1k}, \beta_{2k}, \dots$, are the coefficients for the predictors for category k

$P(Y=k)$ is the probability of the outcome being in category k.

$P(Y=\text{reference})$ is the probability of the outcome being in the reference category.

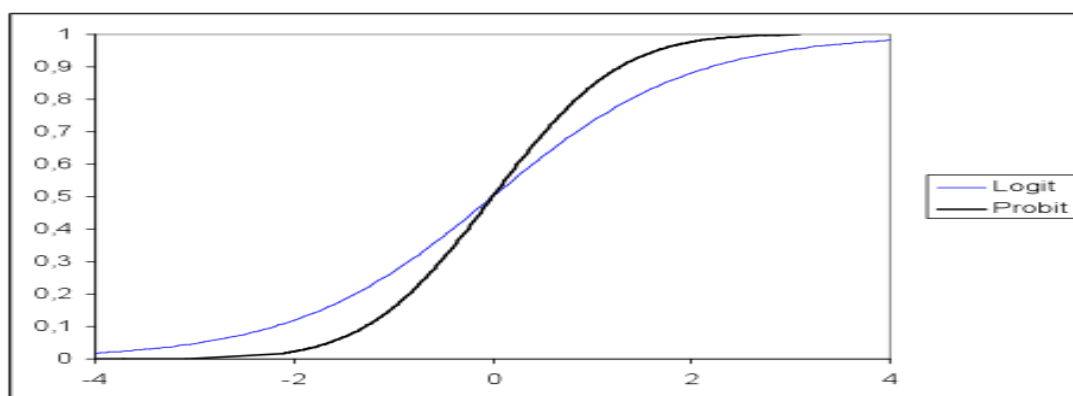


Figure 6: Multinomial regression.

Source: <https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/mlr/>

Decision tree regression:

It is a versatile method for handling diverse data types. Figure 2 affirms that decision tree regression is a non-parametric supervised learning method that partitions data into subsets based on input features, recursively modeling the relationship between input variables and outputs. Breiman *et al.* (1984) introduce decision trees for classification and regression tasks, discussing algorithmic details, splitting criteria, and applications in various fields. This method is particularly useful in showing non-linear relationships between variables or to capture interactions between predictors. Smith and Draper (2014) affirmed that

the assumptions of decision tree are that they are non-parametric models that do not assume data distribution, the existence of homogeneity independent variables, handling missing data values and complex data, having the ability to over-fit trains, and possessing less sensitivity to outliers than linear regression models.

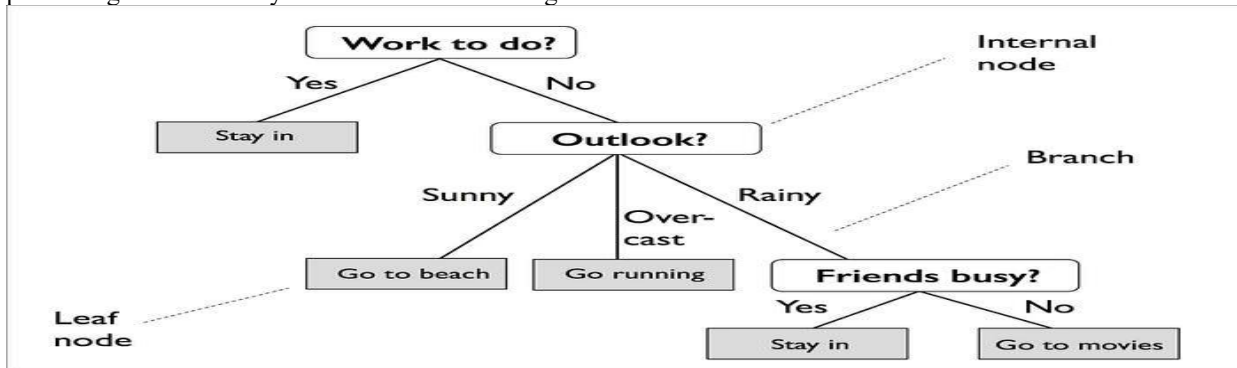


Figure 7: Decision tree regression

Source: <https://towardsdatascience.com/https-medium-com-lorli-classification-and-regression-analysis-with-decision-trees-c43cdbc58054>

Table 2: Strengths and limitation of different types of nonlinear regression in social science research

Non-Linear Regression	Strengths	Limitations
Binominal	It handles the non-linear relationship of the variables. It is easy to compute and implement the large datasets.	It is not suitable for multivariable issues compared to multinomial. It needs a substantial sample size to generate accurate estimates.
Polynomial	It can be used in various fields (education, health and finance, etc.). It is easy to implement and compute It makes the interpretation of the coefficient easy to understand.	It results to unreliable result due to high multicollinearity It is not suitable for high-dimensional data. It is not suitable for data with complex relationships.
Ridge	It can be used for non-linear relationships of any degree It reduces overfitting and difficult to implement on large data sets. It handles multicollinearity better than ordinary least squares (OLS)	It is expensive to compute a large data set. It may not capture complex interactions. It may not be suitable for non-nominal data.
Multinomial	It is flexible for both linear and nonlinear regression operations. It effectively predicts both continuous and categorical variables. The odd ratios and coefficients are easy to understand. The prediction outcome can be a decision-making process.	It is not prone to overfitting and expensive to compute. It requires more processing power for numerous variables The interpretation of model coefficients becomes more difficult due to large variables It may suffer overfitting if the categories are large.
Decision tree	It provides more flexibility in handling real-world data. It is easy to understand and simple to explain the predict variable. It handles both numerical, categorical data. and missing values It captures non-linear relationships between variables.	Not suitable for continuous outcomes It is not suitable for complex relationships. It is sensitive to hyper parameters.

Table 3: Criteria for choosing appropriate regression models

Types of regression	Dependent variable	Independent variable	References
Simple Linear regression	One variable (continuous)	One variable (continuous)	Petropoulos <i>et al.</i> , 2022; Stevens, 202
Multiple linear regression	One variable (continuous)	It involves more than one variable (continuous and categorical)	Maxwell, 2016
Polynomial regression	One variable (outcome)	Indefinite (many) variables and not constant	Ostertagová, 2012; Helwig, 2017
Ridge regression	Continuous outcome variable (production, income)	Continuous or categorical (multiple predictors that are highly correlated)	Hastie, 2009; James <i>et al.</i> , 2021
Multinomial regression	Categorical variable	Continuous and categorical (income, age, or sex) or both	Hosmer <i>et al.</i> , 2013; Agresti, 2018
Binomial regression	It is binary in nature (yes or no or high or low)	Continuous and categorical (income, age or sex) or both	Hosmer <i>et al.</i> , 2013; Adjei and Karim, 2016
Decision tree regression	Continuous variable (income, production)	Continuous (years of experience) or categorical (types of livelihood activities).	Smith and Draper, 2014

CONCLUSION AND RECOMMENDATIONS

The study emphasizes the assumptions of linear and non-linear regression and highlights the advantages of regression over other test tools in social science research. The study seeks more insight on the different types of non-linear regression analysis, its strengths and limitations in social sciences research, revealing their significant roles in achieving non-questionable conclusions and highlighting its importance in achieving research outcomes. Therefore, the study recommends that researchers should seek deeper knowledge and utilize the nonlinear regression techniques when applicable for better and deeper interpretation of data for decision makers and address the limitations of linear regression

REFERENCES

- Adjei, I. A. and Karim, R. (2016). An application of bootstrapping in logistic regression model. *Open Access Library Journal*, 3(9): 24 -31.
- Agresti, A. (2018). An Introduction to Categorical Data Analysis, Chapter 7: Multi category Logit Models, 3rd Edition, Published by Wiley Series in Probability and Statistics. 193 P.
- Ali, P. and Younas, A. (2021). Understanding and interpreting regression analysis: Research made simple. *Evidence-Based Nursing*, 24: 116-118.
- Aliferis, C. and Simon, G. (2024). Over-fitting, Under-fitting and General Model Overconfidence and Under-Performance Pitfalls and Best Practices in Machine Learning and AI. In: Simon, G. J., Aliferis, C. (edition) Artificial Intelligence and Machine Learning in Health Care and Medical Sciences. Health Informatics. Springer, Cham. https://doi.org/10.1007/978-3-031-39355-6_10.
- Ayre, C. and Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 47(1): 79-86
- Byrne, B. M. and Cahyono, S. T. (2022). Structural Equation modelling with AMOS. Publisher by Taylor and Francis Group. SBN 978-0-8058-6372-7.
- Byrne, B. M. (2016). Structural Equation Modeling With AMOS: Basic Concepts, Applications, and Programming, Third Edition. Routledge. <https://doi.org/10.4324/9781315757421>.
- Calzon, B. (2023). Your modern business guide to data analysis methods and techniques. The datapine FUDMA Journal of Agriculture and Agricultural Technology, Volume 12 Number 2, June 2026, Pp 45-56
- Blog. Retrieved from <https://www.datapine.com/blog/data-analysis-methods-and-techniques>.
- Constant, N., and Roberts, E. (2017). Narratives as a mode of research evaluation in citizen science: understanding broader science communication impacts. *Journal of Science Communication*, 16(4), A03:1-18.
- Creswell, J. W., and Creswell, J. D. (2018). Research Design: *Qualitative, Quantitative, and Mixed Methods Approaches* (5th edition). SAGE Publications.
- Das, R., H. R. Wason, and M. L. Sharma (2014). Unbiased estimation of moment magnitude from body- and surface-wave magnitudes, *Bulletin Seismological Society of America*, 104 (4), 1802: doi: 10.1785/0120130324.
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics* (5th edition) SAGE Publications India. 9781526419514, 1526419513.
- Frost, J. (2020). *Regression Analysis: An Intuitive Guide for using and interpreting linear models*. Statistics Jim Publishing.
- Hama, A., Tanaka, K., Mochizuki, A., Tsuruoka, Y. and Kondoh, A. (2020). Estimating the protein concentration in rice grain using Uav imagery together with Agro-climatic data. *Agronomy*, 10, 431.
- Helwig, N. E. (2017). *Regression with Polynomials and Interactions*. Psychology and Statistics University of Minnesota (Twin Cities). Retrieved from <http://users.stat.umn.edu/~helwig/notes/polyint-Notes.pdf>.
- Hosmer Jr., D. W., Lemeshow, S. and Sturdivant, R. X. (2013). *Applied Logistic Regression*. 3rd Edition, John Wiley and Sons, Hoboken, NJ. <https://doi.org/10.1002/9781118548387>.
- James, G., Witten, D., Hastie, T. and Tibshirani, R. (2021). *Introduction of statistical learning: Application in R*. Published by Springer Texts in Statistics (STS), pp. 553-595.
- Jilcha S. K. (2020). *Research Design and Methodology*. Chapter Metrics Overview IntechOpen. doi: 10.5772/intechopen.85731.
- Jocelyn H.B. 2023. A Comprehensive View for the Social Sciences Get access Regression Analysis in R: A Comprehensive View for the Social

Sciences. V Kalyani V Kalyani School of Social Work, DMI St. Eugene University, Lukasa, Africa kalsmsw@gmail.com Retrieved from <https://doi.org/10.1093/jrsssa/qnad081>

- Kiernan, D. (2014). Multiple linear regression: Natural Resources Biometric. MILNE library. Retrieved from <https://milnepublishing.geneseo.edu/natural-resources-biometrics/chapter/chapter-8-multiple-linear-regression/>.
- Ilgamal, Z. Y. (2020). Shrinkage parameter selection via modified cross-validation approach for ridge regression model. *Communication. Statistical-Simulation Computation*, 49(7), 1922–1930.
- Maestrini, V., Luzzini, D., Shani, A. B., and Canterino, F. (2016). The action research cycle reloaded: conducting action research across buyer-supplier relationships. *Journal of Purchasing and Supply Management*, 22, 289–298.
- Maiwada, S. and Okey, L. E. (2015). The relevance and significance of correlation in social science research. *International Journal of Sociology and Anthropology Research*, 1(3), 22-28.
- Maravelakis, P. (2019). The use of statistics in social sciences. *Journal of Humanities and Applied Social Sciences*, 1(2), 87-97.
- Matthews, N. L. (2018). Measurement, levels of variables. The Media School, Indiana University. Retrieved from file:///C:/Users/ba_oy/Downloads/MatthewsN.L.2017.LevelsofMeasurement.pdf on June 11, 2024.
- Maxwell, J. A. (2016). Expanding the history and range of mixed methods research. *Journal of Mixed Methods Research*, 10(1), 12–27.
- McNeish, D., Stapleton, L. M., and Silverman, R. D. (2017). On the unnecessary ubiquity of hierarchical linear modeling. *Psychological Methods*, 22: 114-140. Doi: 10.1037/met0000078
- Mohajan, H. K. (2017). Two criteria for good measurements in research: Validity and Reliability *Annals of Spiru Haret University Economic Series*, 17(3), 58–82.
- Mohajan, H. K. (2020). Quantitative Research: A Successful Investigation in Natural and Social Sciences. *Journal of Economic Development, Environment and People*, 9 (4). 52-79.
- Montgomery, D. C., Peck, E. A. and Vining, G. G. (2015). *Wiley Series in Probability and Statistics Introduction to Linear Regression Analysis*. (5th edition). Library of Congress Cataloging-in-Wiley Publication Data, pp. 1-872.
- O'Connor, C. and Joffe, H. (2020). Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. *International Journal of Qualitative Methods*, 19: 1-13. <https://doi.org/10.1177/1609406919899220>.
- Ostertagová, E. 2012. Modelling using polynomial regression. Sciverse Science direct, *Procedia Engineering*, published by Elsevier, pp. 1-8.s
- Petropoulos, F., Apiletti, D., Assimakopoulos, V., Zied Babai, M., Barrow, D. K. and Taieb, S. B. (2022). Forecasting: theory and practice, *International Journal of Forecasting*, 38(3): 705 - 871.
- Pistol, L., and Bucea-Manea-Tonis, R. (2017). Model of simulation for optimizing marketing mix through conjoint analysis case study: launching a product on a new market. *Economics World*, 5(4), 311 315.
- Poston, Jr, D. L., Conde, E. and Field, L. M. (2023). *Applied Regression Models in the Social Sciences*. Cambridge: Cambridge University Press. Pp.158.
- Ramdurg, A. I. (2020). Correlation (Unit 4): Business Statistics II. Business Communication 4th Semester, Rani Channamma University, Belagavi, 1-27.
- Reza, N., Na, I. S., Baek, S. W., and Lee, K. H. (2018). Rice yield estimation based on K-means clustering with graph-cut segmentation using low-altitude UAV images. *Biological system Eng*. 177, 109–121.
- Sauerbrei, W., Abrahamowicz, M., Altman, D.G., leCessie, S. and Carpenter J. (2014). Initiatives strengthening analytical thinking for observational studies: the STRATOS initiative. *Statistical Medicine*, 33(30):5413–5432. <https://doi.org/10.1002/sim.6265> PMID: 25074480.
- Smith, H and Draper, N. R. (2014). *General and introduction statistics: Applied regression analysis*. John Wiley and Sons publication, Hoboken, 326: 736P.
- Snijders, T. A. B., and Bosker, R. J. (2012). *Multilevel Analysis: An Introduction to Basic and Advanced*

- Multilevel Modeling* (2nd edition). Sage Publications.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines, *Journal of Business Research*, 104: 333-339.
- Song, H. Y. (2018). Artificial neural network-based model for predicting moisture content in rice using UAV remote sensing data. *Korean Journal of Remote Sens.* 34(6), 611–624.
- Stevens, M. (2023). The seven (7) most useful data analysis methods and techniques. Careers foundry Blog. Retrieved from <https://careerfoundry.com/en/blog/data-analytics/data-analysis-techniques>.
- Sürücü, L. and Maslakçı, A. (2020). Validity and reliability in quantitative research. *Business and Management Studies: An International Journal* 8(3): 2694-2726, doi: <http://dx.doi.org/10.15295/bmij.v8i3.1540>
- Tang, Y. (2019). A differential evolution-oriented pruning neural network model for bankruptcy prediction. *Complexity*, 13(2), 1-21.
- Thomas, M.A. (2019). Mathematization, Not Measurement: A Critique of Stevens' Scales of Measurement. *Journal of Methods and Measurement in the Social Sciences*, 10(2): 76-94.
- Thomas, S. (2022). Discrete vs. Continuous variables: Differences Explained. Outlier article Home Statistics. 2240 Kent Avenue; Brooklyn, New York, 11249, United States Oscar L. Olvera Astivia, O.L. O and Zumbo, B.D. 2019. Heteroskedasticity in Multiple Regression Analysis: What it is, How to detect it and How to solve it with Applications in R and SPSS. Practical Assessment, Research and Evaluations.
- Wei-Chih, H., Pao-Yuan, C., Chia-Sui, W., Jen-Chieh, H., and Huang, W. (2020). Application of regression analysis to achieve a smart monitoring system for aquaculture. *Information*, 11(2), 1-9.
- Williams, M. N. (2021). Levels of measurement and statistical analyses. *Meta-Psychology*, 5: 1-14. <https://doi.org/10.15626/MP.2019.1916>.
- Zaid, A.I. and Tsagem, S.Y. (2022). Summary, Conclusion and Recommendation. In Bagudo, A. A., Yusuf, M. A. and Gumbi, S. U. (Edition.), Report Writing for Educational Research: A Guide. Sokoto, Nigeria: *9ice Plus* Prints. ISBN: 978-978-59429-9-6.