



REGRESSION ANALYSIS AND RELEVANCE TO RESEARCH IN SOCIAL SCIENCES

¹Ajibade, Ayodele Temitope, ²Ademola, Ajayi and ¹Otitolaiye, Emmanuel Dare

¹Department of Accounting, School of Management Sciences

²Babcock University, Ilishan-Remo, Ogun State, Nigeria

Abstract: The study seeks to review regression analysis and its relevance to research in social sciences, the study relied on a review of various regression analyses being used in social sciences and the significance of regression analysis as a tool in the analysis of data sets. The study adopted a systematic exploratory research design, reviewing related articles, journals, and other prior studies in relation to regression analysis and its relevance in social sciences. After a careful systematic and contextual review, the study revealed that regression analysis is significant in providing a measure of coefficients of the determination which explains the effect of the independent variable (explanatory variable) on the explained variable otherwise known as regressed variables that give the idea of the prediction values of the regression analysis. Regression analysis provides a practical and strong tool for statistical analysis that can enhance investment decisions, business projections in manufacturing, production, stock price movement, sales, and revenue estimations, and generally in making future predictions. This review provides originality in a clear understanding of a comprehensive review of the relevance of regression analysis in social sciences, contributing to knowledge in this regard. The study recommends that researchers should adopt the required pragmatic and methodological steps when using regression analysis, unethical torturing of data should be avoided as this could lead to false results and wrong statistical predictions.

Keywords: Bivariate regression, Multiple regression, Investment projections, Predictive ability, Polynomial regression, Regression analysis.

1.1 Introduction

In the social sciences research, regression analysis is one of globally known methods being explored to analyze and establish the relationship between two or more variables and in expressing the variables in a mathematical equation that enhances a significant dimension required to carry out predictions and investment decisions (Algama, 2020). Some studies have shown that regression analysis plays an important role in social sciences, as in every research endeavors statistical regression is compulsory to enable researchers reach tested and estimated non-questionable

conclusions (Hama, Tanaka, Mochizuki, Tsuruoka & Kondoh, 2020; Das, Nair, Reddy & Venkatesh, 2018). Regression analysis according to Alsaawy, Alkhodre, Benaida and Khan (2020), aims to check whether there is a relationship between variables, identify the magnitude and extent of the impact of one or more variables on another (Aksan & Baki, 2017; Asseng, Zhu, Basso, Wilson, Cammarano, 2014). Studies have shown the complexities and rigorous process involved in research in social sciences, however, the complexities depend on the extent to which the research tends to go and the typical

Academic Journal of Accounting and Business Management

An official Publication of Center for International Research Development

Double Blind Peer and Editorial Review International Referred Journal; Globally index

Available www. <https://cirdjournal.com/index.php/ajabm> E-mail: journals@cird.online



research being undertaken in solving specifics and existing problems. Every research expedition is aimed to solve a problem, and when there is no specific problem or gap in any research, it becomes an exercise in futility (Constant & Roberts, 2017). The relevance of regression analysis to research in social sciences is essentially significant in providing procedures for regression analysis in a simpler manner, with each of the procedures having illustrations, instances, and models (Everingham, Sexton, Skocaj, Inman-Bamber, 2016). In consideration of the relevance of regression analysis to research in social science, it is pertinent to state that social science is an embodiment of a branch of science devoted to a methodological study of societies and the association among diverse individuals within the spheres of societies (Tuffour, 2017). Social science was revolved from the field of sociology, then the science of society in its originality (Reza, Na, Baek & Lee, 2018). Scholars have documented that in its spectrum, this consisted of a myriad of academic disciplines of anthropology, economics, archaeology, human geography, political science, linguistics, media studies, management science, psychology, history, and many more (Paul, 2017; Mohajan, 2017).

According to Reza, Na, Baek & Lee (2018) regression analysis in social sciences brings to fore the combination of methods of data regression analysis in a simpler manner, with one method having different illustrations and instances. The study further documented that a broad range of techniques are possible, and these include: Multiple clear-cut predictions, methods for curvilinear regressions, systematic regressions, linear regression, logistic regression, regression discontinuity, bivariate regression, simple regression, multiple regression. Onoyama, Ryu, Suguri and Lida (2018) stressed that various methods do exist in determining the average or best underlying association using regression analysis, this includes from the subjective of variables' assessment at the discretion of the researcher to more objective measurements that

determine the very unique link that minimizes variations. In addition, that when data is interval if the minimum variation is related to standard deviation or least square regression line of best fit, it becomes a regression analysis (Johnasen, Raharjo & McCabe, 2018).

The objective of this study was to review regression analysis and its relevance to research in social sciences. While previous studies have reported divergent opinions and mixed reactions that have followed the relevance of regression analysis, much has not been done in the Nigerian domain. This study in extending the frontiers and contributing to knowledge proposed to examine the relevance of regression analysis in social sciences. The objective of this study has been to examine regression analysis and its relevance to research in social sciences. In contributing to knowledge, the study was motivated to fill the gap created due to the dearth of studies that have specifically studied regression analysis and relevance to research in social sciences from the Nigerian literature domain. In addressing this gap in the literature, this study carried out a review and consequently examined the relevance of regression analysis to research in social sciences. The rest of the study was structured in this manner: Section 2 presented a literature review, in section 3, the methodology was considered, while the study concluded in section 4 with a conclusion and recommendations.

2. Methodology

Regression analysis and relevance to research in social sciences formed the basis of this study. A systematic exploratory research approach was adopted in carrying out the study, while a wide range of related journals, periodicals, and other material considered appropriate were considered in the review exercise.

3. Literature Review

Regression: The concept of regression is concerned with the ability to identify the landscape of the nexus between two or more variables of operational concepts in



literature. Miocic, Zekanovic-Korona and Hotti (2020) also posited that the issue of regression is basically to determine the statistically underlying association between two or more variables on the basis of observations. One of the major objectives of any regression analysis in the social sciences is to check if there be any relationship between variables, also to identify the strength of this association and then conclude a regression equation that is used to describe this association. More still, Khan (2014) noted that regression processes are expected to start by presupposing some form of underlying association, then try to identify exactly the association that subsists, and in the social sciences, these presumptions is regarded as a linear association (Kang, Nam, Kim, Lee, Seng, Jang & Ryu, 2021). Regressions can be simple, multiple, or multivariate regression analysis, polynomial regression, ridge regression, lasso regression, partial least square regression, regression discontinuity,

Bivariate Regression: Bivariate regression in some cases reflects a simple regression in which two variables kind of bivariate in relation to the association exactly specified as regression line for the reliant concepts otherwise the dependent variables usually denoted by a symbol (Y) on the explanatory variable also known to be (X), and this form the general form of a straight line of $Y = a + Bx$. According to the study of Lolli, Gasperini and Vannucci (2014), there are instances where bivariate becomes non-straight-line in nature, rather polynomial regression.

Multiple Regression: Unlike the simple regression or bivariate regression analysis, where Y is seen as a function of one independent variable, multiple regression, or multivariate regression is a the situation in which Y saw as a function of the number of independent variables or explanatory variables of $X = x_1, x_2, x_3, \dots, x_n$) and this requires a procedure explored in multivariate analysis. In econometrics form, the multivariate linear regression expression line takes the form of $Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots + b_nX_n$. Where n = independent variables and each has

coefficients (b_1, b_2, \dots), in other words, the multiple linear regression contains more than one regressor variables in a model form of $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_kX_k + \mu$. where Y is the dependent variable, β_0 = the constant, $\beta_1 - \beta_2$ = the coefficient of the explanatory variables, μ is the error term of any set of models (Maravelakis, 2019; Mohajan, 2017).

Basic assumptions of Linear Regression: Some studies (Maestrini, Luzzini, Shani & Canterion (2016); Maxwell (2016) documented that some of the fundamental assumptions for linear regression include: (i) the value of the explanatory variables 'x' is decided by the researcher (ii), the explanatory variables 'x' should be measured without any experimental errors (iii) In each case, the value of 'x' is the sub-population of the 'y' variables that are usually distributed up and down the Y-axis (iv) the operational concepts of the sub-population of 'y' are homogeneous in nature (v) and the mean value of the sub-population of 'y' is expected to lie on a straight line and this means that the assumptions do exist (Maiwada & Okey, 2015; Pierre, 2013).

Regression Discontinuity: In statistics, some conditions could arise in regression analysis and effect of the intervention is noticed. In this scenario, regression discontinuity is applied to compute the impact of an intervention, to give an unbiased estimate. In regression discontinuity design, researchers use a rule to assign the intervention to a unit (Paul, 2017).

Polynomial Regression: According to Pistol and Bucea-Manea-Tonis (2017), polynomial regression is in the category of regression analysis in the social sciences that is employed in handling nonlinear analysis involving separable data. In this regression method, the best fit lines are not straight lines, instead, the line are curves that fit into the data known points and in the polynomial regression, the power of the explanatory variables are usually more than 1 (one).

Ridge Regression: Ridge regression serves as a unique regression that is used when there is high handling



collinearity among elements of variables, where standard linear or polynomial regression cannot be used. According to Tuffour (2017), ridge regression is used when data suffers from multicollinearity, which is the existence of near-linear association among the explanatory variables in a data set.

Lasso Regression: Sarkar, Ryu, Kang, Jun, Jan and Song (2018) submitted that lasso regression is a contraction and variable selection model for linear regression models. The aim of lasso regression is to gain the subset of predictors that would reduce predictive errors for quantitative response variables. In handling this task, lasso regression is employed by imposing a unique constraint on the model parameter that is capable to cause the regression coefficient for some of the elements in the model to be reduced to zero.

Partial Least Squared Regression: Tang (2019) noted that partial least squared regression is another type of multivariate linear regression method applicable where the least square model is applied to a mixture of both the dependent and independent variables to derive latent variables with high covariance. It is employed in statistics to establish optimal latent (dominate) squares that will reduce the explained variables in the dependent variable from explanatory variables and that could reduce the predictability residuals sum of squares and mean square errors (Wang, Cjang, Yang, Zhang & Li, 2018).

Multilevel Regression: Beyond the foregoing, multilevel regression modeling is one of the models used in analyzing data relating to clusters. Particularly, in social science research, there are often concerns with the association between two or more individuals or group that statistically related (Wei-Chih, Pao-Yuan, Chia-Sui, Jen-Chieh & Huang, 2020). This association actually leads to nested data within groups revealing that the outcome of one member could have an effect on the other members in the same group.

3.2 Statistical Tools and Softwares

In carrying out regression analyses, there various range of statistical tools and software at the disposal of the researcher. Apparently, since each offers slightly different services to the researcher, the type and what the researcher chooses depend on the range of factors, as well as the research questions, research expertise, and knowledge of statistics and capability of coding processes (Lin, Lei & Zhou, 2019). Basically, some notable statistical tools/and or software in use include the following:

- (i) **Statistical Package for the Social Sciences (SPSS).** SPSS sees the most widely statistical tool used by researchers in human behaviors research as it offers simple descriptive, parametric and non-parametric statistics coding, graphical and other advanced statistical (Neha, Viswanathan (2019).
- (ii) **R-Foundation (R):** R is a one of the free statistical tool that helps in learning curves application, which requires a certain degree of coding and plugins in the data process.
- (iii) **MatLAB:** Liu, Tan, Lim and Choi (2017) noted that MatLab is a powerful analytical platform and programing language tool used by researchers in projecting production data used by engineers. MatLab could be complex and difficult to use by non-experts but offers a wide range of flexibilities.
- (iv) **R-Studio:** R-Studio assists in building sustainable data and statistical computing in social sciences. It is among the brand of Rmarkdown, shiny software packages that help analysis make quick predictions in commercial product and project estimations.
- (v) **EViews:** This offers tremendous regression assistance to academic researchers, entities, government, agencies and students' access to statistical analysis, forecasting, and modeling using an object-oriented interface (Jagtiani & John, 2018).
- (vi) **Stata:** Stata is one of the contemporary statistical software that puts hundreds of statistical tools in



researchers' fingertips for data management, statistical analysis and publishing-quality graphics

(v) **JMP**: JMP comes to mind when there is a need to combine statistics with dynamic graphics in memory and on desktop. Uniquely, JMP's interactive and visual paradigm enable it to reveal insights that may seem impossible access to raw tables of number or static graphs.

(vi). **OriginPro**: OriginPro is one of the good user-friendly and easy-to-learn software applications in regression analysis. It offers quality graphic abilities and extended analytical tools for peak fittings, surface fitting, signal process, and image handling flexibility in the hands of researchers (Hamire & Paula, 2019).

(vii). **Tlmi-Suite**: TlmiSuite consist of four basic tools of (a) Anatella, (b), Modeler, (c) StaDust (3D Segmentation) and (d) Kibella (BI Dashboarding solution). These are new innovation and most powerful analytics, predictive analytical, Artificial Intelligence (AI) and Big Data that offers a fast and friendly environment in financial accounting information computing.

(viii) **Scilab**: Scilab is one of the free open-access software for numerical computations and simulation of data in social science

(ix) **Minitab**: In regression analysis, Minitab software is one of the leading statistical software employed for quality improvement and in academics for statistical education and training globally.

(x) **GNU Octava**: GNU Octava is also one of the open access and open-source mathematical modeling and for data simulation software similar in nature to Matlab and Firemat software (Yao, Di, Zheng & Xu, 2018).

In addition, studies have shown that the emerging technologies, disruptive and non-disruptive technologies and digital competition are altering the landscape of financial reporting environment substantially and this change is accelerating, including artificial intelligence, robotics process automation, and bloc chain are all changing the way businesses are done and regression

analysis to research in social science is being equipped to face possibilities in meeting researchers and auditors in transforming their own processes to ensure quality financial reporting and other financial services delivery (Murthy, 2017; Yang, Stankevicius, Marozas, Deng, Liu, Lukosevicius, Dong, Xu & Min, 2018)

4.0 Theoretical Underpinning

Grounded theory: Ground theory was developed by Glaser and Strauss within the period 1965 to 1967 aimed at providing a scientific approach to the analysis of issues using qualitative models (Ellis, 1993; Fisher, 2009). The grounded theory suggested that after observing the way hospital staff attendee to dying patients, there was the reason to translate analyzed data into only organized, but many ideas. This detailed observation and analysis of data formed the principal line of the reasoning of grounded theory to organize many ideas from analysis of data (Straaauss & Corbin, 1990). According to Fisher (2009), the aim of grounded theory was to build a realistic theory that was capable to illuminate ideas under the area of study. The grounded theory suggested that the research model require a systematic set of procedures to improve and inductively drive grounded theory about a phenomenon (Strauss & Corbin, 1990). Procedurally, the grounded theory reflected the systematic analysis of sourced data, documents, and other raw data obtained from observations, interviews, or from secondary source with a continuous coding and comparison of data which in turn produces a strong theory.

Some of the proponents of ground theory suggested that ground theory stipulated that it pioneered a natural association between variables and established relationships between the data and the theory so that the end results will reveal processed and grounded ideas for predictions (Wilson & Hutchinson, 1991). The theory also assumes that problems to be discovered by the researcher should be resolved through a basic social-psychological process, that research is aimed to include discovering basic



social psychological problems through processes, phases and discovering properties using strategies (Pirolli & Card, 1995). The underlying significance of grounded theory to this study lies in the role and relevance of regression analysis to research in social sciences. In addition, grounded theory has been successfully used in many research areas, and its application in information management research in providing a rigorous insight into areas that are relatively unknown by the researchers within the field of social sciences and beyond.

5. Relevance of Regression Analysis in Social Sciences

In the social sciences, the regression analysis techniques are simplified and carried out step by step, assisting students and many other researchers analyze chosen data from social and behavioral sciences, making the best predictive information required by the existing and potential investors in making investment decisions and/or portfolio divestment options, the analysts in providing predictive information to its clients and the general public who desire critical and reliable information. From the statistical assumptions, regression analysis in statistics is one of the vital tool used in the social sciences by the social scientist in conducting research studies and making predictions (Maiwada & Okey, 2015). Other relevance includes the following:

Firstly, regression analysis provides assistance in establishing the functional relationship between two or more connected variables towards a prediction of unknown values of one or more variables from the known values of another variables. Secondly, regression analysis also provides the desired measure of errors of the estimate made from end to end of the regression lines. Thirdly, regression analysis in social sciences helps in measuring coefficients of correlation between two or more variables which can be calculated by taking a possible square root of the sum of the two regression coefficients. Fourthly, regression analysis helps in providing a measure of coefficients of the determination which explains the effect of the independent

variable (explanatory variable) on the explained variable is otherwise known as regressed variables that give the idea of the prediction values of the regression analysis. Fifthly, it provides a strong tool for statistical analysis for investment decisions, business analysis, and in manufacturing and production, stock price movement, sales and revenue estimations, and generally in making future predictions. Every successful investor strongly depend on the degree of accurate projection arising from regression analysis. Sixthly, it provides a veritable tool to measure and estimate the cause-and-effect association among useful economic and accounting variables that constitute the crux of economic principle and theory, such as estimation of demand and supply curves, production tasks, production, and overheads cost functions, consumption estimations, and functions, using regression analysis for the number of inputs and possible outputs (Reza, Na, Baek & Lee, 2018).

Lastly, regression analysis in social sciences is practically relevant in day-to-day life activities, giving possible outcome and variations in all human endeavors. Linear regression is one of the most used regression analyses in the social sciences (Tuffour, 2017). The regression in all purposes is significant in many ways of descriptive analysis, in adjusting for the impact of covariates or the confounders; it enhances predictions as a portent predictor, by helping in estimating the significant risk factors that impact the explanatory variables; it determined the extent of prediction, by assisting in analyzing the possible changes in the independent variable in units variations that would naturally impact the dependent variables and enhances prediction by qualifying the new cases (Paul, 2017). Besides, regression analysis is considered relevant, important as it provides the students in the social sciences with the basic requisite skills and knowledge they need to estimate, interpret, present, and in addition, publish basic regression models using international best practices in making predictions and decisions.



5.2 Significance of Regression Analysis in Business Operations

The following are some the relevance of regression analysis as reported in literature: In social sciences, regression analysis offers researchers a detailed insight that can be employed in future improvement of product and services. According to Witte and Didelez (2019), regression analysis is a powerful statistical tool that allows researchers ascertain the nexus between different variables, the predictors and explanatory variables in social science. It enhances the determination the influence of one phenomenon.

Predictive Analysis: Regression analysis with analytical abilities using statistical formulas and models in making risks, assessment and risks management as well as forecasting of future opportunities (Vatcheva, Lee, McCormick & Rahbar, 2016).

Operations Efficiency: According to Gregorich, Strohmaier, Dunkler and Heinze (2021), corporate organizations use application of regression analysis in optimizing business opportunities. For instance in factory oven temperature monitoring of loaves of bread baked in bread oven, call centers using regression analysis in creating optimal operation efficiency of time management.

Supporting Decisions: Globally, high technologies sensitive corporations uses regression analysis in form of data analytical and 'Big Data' to make informed business decisions and also enhance elimination of gut intuitions in scientific management strategies (Miocic, Zekanovic & Hotti, 2020).

Correcting Errors: Regression analysis is relevant in helping managers and other corporate organizations in recognizing, and correcting errors. Huebner, LeCessie, Schmidt and Vach (2018) posited that regression analysis assist managers in determining that increasing number of hours may not lead to increase in profits.

New Insights: Regression analysis is relevant in indicating a spike in sales during certain days of the week or drops in sales, this could lead to taking necessary steps and adjustments.

5.2 Possible Irrelevance of Regression Analysis in Social Sciences: While there have been vast advantages of regression analysis in social sciences, some studies have debated whether there are some possible disadvantages or irrelevances of regression analysis in social sciences. According to the study of Devi and Radhika (2019), there are no known disadvantage of regression analysis as a statistical tool, rather the reliability of the data being used and the skill of the researchers in carrying out regression analysis. Regression analysis in the hands of researchers will determine the results and the acquired skill of the researchers to carry out the appropriate regression analysts, interpretation ability, and application of the results in making relevant predictions (Ds, Nair, Reddy & Venkatesh, 2018). Apparently, faulty and unreliable data will ultimately reveal a faulty and incredible results and decisions made using faulty predictions and advice will not add any economic value rather zero value to such investment decisions. Everingham, Sexton, Skocaj and Inman-Bamber (2016) opined that using an incomplete or defective data could lead to false conclusions, defective findings and not necessarily the fault of regression analysis as a tool in research in social sciences.

6.0 Summary, Conclusion and Recommendations

Summary: Regression analysis is phenomenal and has played significant roles in assisting researchers in social sciences. The review reveals that previous studies have made good efforts in studying the relevance of regression analysis to research in social science, however, in spite of the numerous merits and relevance benefits attributable to regression analysis in social science, those studies have equally shown inconsistent in reported results using the same range of data, yet arriving at different results (Kang, Jang, Park, Song, Ryu, Jun & Kim, 2020). The regression



analysis assumed that the effect and cause association between variables remain unchanged. Evidence has shown that assumptions may not all the time hold, many regression analyses seemed to be erroneously reported and in most cases misleading (Levitt, Motulsky, Wertz, Morrow & Ponterotto, 2017). Lin, Lei and Zhou (2019) opined that in spite of the usefulness of regression analysis in social sciences, some researchers have abused and tortured data, to arrive at predetermined results thereby reporting distorted and unreliable information. Besides, some results had been ambiguous, the process involved in regression analysis in some instances seemed lengthy, and that requires a complicated procedure of establishing the validity and normality of data before analysis

Conclusion: This study examined regression analysis and relevance to research in social sciences. The review exercise cut across the spectrum of regression analysis within the social sciences and relevance in a research exercise. The study revealed that regression analysis had played a significant role in social sciences. In every research exercise effort, statistical regression analysis is compulsory to enable the researcher to reach a non-questionable conclusion as an outcome analysis.

Recommendations: The study recommended that researchers should be patient to methodological sequence and process of data regression analysis procedures to harvest non-questionable and reliable results capable of enhancing investment decisions for existing and potential investors and production predictions for manufacturing companies and for the analysis for an accurate forecast abilities.

References

Akşan, E., & Baki, A. (2017). Content analysis of curriculum-related studies in turkey between 2000 and 2014. *Educational Sciences: Theory & Practice*, 17, 877–904

Alsaawy, Y., Alkhodre, A., Benaida, M., Khan, R. A. (2020). A comparative study of multiple regression analysis and back propagation neural network approaches for predicting financial strength of banks: an Indian perspective. *Wseas Transactions on Business and Economics*, 17(60), 627-637. DOI: 10.37394/23207.2020.17.60

Aguguom, T. A., Ajayi, A., & Otilaiye, E. D. (2020). Covid-19 and Accounting Education in Sub-Sahara Africa. *European Journal of Business, Economics and Accountancy*, 8(3), 1-11.

Algamal, Z. Y. (2020). Shrinkage parameter selection via modified cross-validation approach for ridge regression model. *Commun. Stat.-Simul. Comput*, 49(7), 1922–1930.

Asseng, S., Zhu, Y., Basso, B., Wilson, T., Cammarano, D. (2014). Simulation modeling: applications in cropping systems. *Encycl Agric. Food System*, 14(5), 102–112.

Che, Shulu, Ronald, C. S., Ravinesh, K., & Stauvermann, J. P. (2021). Taxation of land and economic growth. *Economie,s* 9(3), 42-61.

Choy, L. T. (2014). The strengths and weaknesses of research methodology: comparison and complimentary between qualitative and quantitative approaches, *IOSR Journal of Humanities and Social Science*, 19(4), 99–104.

Constant, N., & 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.

Das, B., Nair, B., Reddy, V. K., Venkatesh, P. (2018). Evaluation of multiple linear, neural network and penalised regression models for prediction of rice yield based on weather parameters for west coast



- of India. *International Journal of Biometeorol*, 62(4), 1809–1822
- 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.
- Devi, S. S., & Radhika, Y. (2019). An analytical approach for bankruptcy prediction using big data and machine learning technique. *Journal of Theoretical and Applied Information Technology*, 97(6), 1823-1832.
- Ellis, D. (1993). Modeling the information seeking patterns of academic researchers: A grounded theory approach. *The Library Quarterly*, 63(4), 469-486.
- Everingham, Y., Sexton, J., Skocaj, D., Inman-Bamber, G. (2016). Accurate prediction of sugarcane yield using a random forest algorithm. *Agro Sustainable Development* 36, 27-42.
- Gasperini, P., Lolli, B., & Castellaro, S. (2015). Comparative analysis of regression methods used for seismic magnitude conversions. *Bulletin of the Seismological Society of America*, 105(3), 1-5. doi: 10.1785/0120150018
- Gregorich, M., Strohmaier, S., Dunkler, D., & Heinze, G. (2021). Regression with highly correlated predictors: variable omission is not the solution. *International Journal Environmental Reserves Public Health*, 18(5), 1-12, 4259.
- Hama, A., Tanaka, K., Mochizuki, A., Tsuruoka, Y., Kondoh, A. (2020). Estimating the protein concentration in rice grain using uav imagery together with agroclimatic data. *Agronomy*, 10, 431.
- Huebner, M., Le Cessie, S., Schmidt, C. O., Vach, W. (2018). A contemporary conceptual framework for initial data analysis. *Observational Studies*, 4(3), 171–192.
- Itamir, C., & Paula, A. M. S. (2019). Academic production and technological emergence in finance Bibliometric study on FinTechs. *Innovation and Management Review- Published by Emerald Publishing Limited*, 1(2), 421-435.
- Jagtiani, J., & John, K. (2018). Fintech: The impact on consumers and regulatory responses. *Journal of Economics and Business*, 100, 1–6.
- Johansen, K., Raharjo, T., McCabe, M. F. (2018). Using multi-spectral uav imagery to extract tree crop structural properties and assess pruning effects. *Remote Sens.* 10, 854.
- Kang, Y., Nam, J., Kim, Y., Lee, S., Seong, D., Jang, S., Ryu, C. (2021). Assessment of regression models for predicting rice yield and protein content using unmanned aerial vehicle-based multispectral imagery. *Remote Sens.* 2021, 13, 1508.
- Kang, Y. S., Jang, S. H., Park, J. W., Song, H. Y., Ryu, C. S., Jun, S. R., & Kim, S. H. (2020). Yield prediction and validation of onion. Using key variables in narrowband hyperspectral imagery and effective accumulated temperature. *Computer Electron*, 178, 1-21.
- Khan, S. N. (2014). Qualitative research method: Grounded theory. *International Journal of Business and Management*, 9(11), 224–233.
- Levitt, H. M., Motulsky, S. L., Wertz, F. J., Morrow, S. L., & Ponterotto, J. G. (2017). Recommendations for designing and reviewing qualitative research in psychology: promoting methodological integrity. *Qualitative Psychology*, 4(1), 2–22.
- Lin, S. B., Lei, Y., & Zhou, D. X. (2019). Boosted Kernel Ridge Regression: Optimal Learning



- Rates and Early Stopping. *Journal of Machine Learning Research*, 20(3), 1–36.
- Liu, F., Tan, C., Lim, E., Choi, B. (2017). Traversing knowledge networks: an algorithmic historiography of extant literature on the Internet of Things (IoT). *Journal of Management Analysis*, 1(1), 3–34.
- Lolli, B., Gasperini, P., & Vannucci, G. (2014). Empirical conversion between teleseismic magnitudes (mb and Ms) and moment magnitude (Mw) at the Global, Euro-Mediterranean and Italian scale, *Geophys. Journal of International*, 199, 805–828, doi: 10.1093/gji/ggu264.
- Maestrini, V., Luzzini, D., Shani, A. B., & Canterino, F. (2016). The action research cycle reloaded: conducting action research across buyer-supplier relationships. *Journal of Purchasing & Supply Management*, 22, 289–298.
- Maiwada, S., & 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
- Maxwell, J. A. (2016). Expanding the history and range of mixed methods research. *Journal of Mixed Methods Research*, 10(1), 12–27.
- Miocic, J., Zekanovic-Korona, L., & Hotti, L. (2020). Importance of regression analysis in sports information systems at evaluation of sports and sports associations. *Acta Kinesiologica* 14(1), 40-45
- 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.
- Murthy, J. J. (2017). Managing innovation and developing intellectual property strategies for firms. *Journal of Management*, 13(1), 82-88.
- Neha, C. R., & Viswanathan, T. (2019). Disruptive technologies and its effect on the workforce in banks: A framework of assessment for mitigation. *The Empirical Economics Letters*, 18(3), 267-281.
- O'Brien, R. M. (2017). Dropping highly collinear variables from a model: Why it typically is not a good idea. *Social Science Quarterly*, 98(5), 360–375.
- Onoyama, H., Ryu, C., Suguri, M., Iida, M. (2018). Estimation of rice protein content before harvest using ground-based hyperspectral imaging and region of interest analysis. *Precis. Agric.* 19, 721–734.
- Paul, K. B. (2017). Introducing interpretive approach of phenomenological research methodology in environmental philosophy: a mode of engaged philosophy in the anthropocene. *International Journal of Qualitative Methods*, 16, 1–10.
- Pierre, E. A. S. (2012). A brief and personal history of qualitative research. *Journal of Curriculum Theorizing*, 30(2), 1–19.
- Pistol, L., & 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.
- Pirolli, P., & Card, S. (1995). Information foraging in information access environment, *Insight*, 34(3), 1-12.



- Reza, N., Na, I. S., Baek, S. W., Lee, K. H. (2018). Rice yield estimation based on K-means clustering with graph-cut segmentation using low-altitude UAV images. *Biosyst. Eng.* 177, 109–121.
- Sarkar, T. K., Ryu, C. S., Kang, J. G., Kang, Y. S., Jun, S. R., Jang, S. H., & 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.
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thought Oaks, CA: Sage Publications
- Tang, Y. (2019). A differential evolution-oriented pruning neural network model for bankruptcy prediction. *Complexity*, 13(2), 1-21.
- Tuffour, I. (2017). A critical overview of interpretative phenomenological analysis: *A Contemporary Qualitative Research Approach*. 2(4), 1-12.
- Vatcheva, K. P., Lee, M., McCormick, J. B., Rahbar, M. H. (2016). Multicollinearity in regression analyses conducted in epidemiologic studies. *Epidemiol. Sunnyvale Calif.* 2016, 6, 227.
- Wang, L., Liu, D., Pu, H., Sun, D. W., Gao, W., Xiong, Z. (2018). Use of Hyperspectral Imaging to Discriminate the Variety and Quality of Rice. *Food Anal. Methods*, 8, 515–523.
- Wang, L., Chang, Q., Yang, J., Zhang, X., Li, F. (2018). Estimation of paddy rice leaf area index using machine learning methods based on hyperspectral data from multi-year experiments. *Publi-Medics*, 34(4), 1-16.
- Wei-Chih, H., Pao-Yuan, C., Chia-Sui, W., Jen-Chieh, H., & Huang, W. (2020). Application of regression analysis to achieve a smart monitoring system for aquaculture. *Information*, 11(2), 1-9.
- Wilson, H. S., & Hutchinson, S. A. (1991). Pearls, pith and provocation. Traingular of qualitative methods: Heideggerian hermeneutics and grounded theory: *Qualitative Health Research*, 1(2), 263-276.
- Witte, J., & Didelez, V. (2019). Covariate selection strategies for causal inference: Classification and comparison. *Biomedical Journals*, 61(5), 1270–1289.
- Yang, B., Stankevicius, D., Marozas, V., Deng, Z., Liu, E., Lukosevicius, A., Dong, F., Xu, L., & Min, G., 2018. Lifelogging data validation model for internet of things enabled healthcare system. *IEEE Trans. Syst. Man Cybern. Part Syst. Hum.* 48 (1), 50–64.
- Yao, M., Di, H., Zheng, X., & Xu, X. (2018). Impact of payment technology innovations on the traditional financial industry: A focus on China. *Technological Forecasting & Social Change* 13(5), 199–207.