ENHANCED KNOWLEDGE IN SCIENCES AND TECHNOLOGY
e-ISSN: 2773-6385

EKST

Vol. 5 No. 2 (2025) 424-432
https://publisher.uthm.edu.my/periodicals/index.php/ekst

Analysis of Fuzzy Linear Regression and Multiple Linear Regression on Household Income and Expenditure in Malaysia

Nurathirah Rodzi[1], Norziha Che Him[1], Mohd Saifullah Rusiman[1]*

[1] *Department of Mathematics and Statistics, Faculty of Applied Sciences and Technology, UTHM Kampus Cawangan Pagoh, Hub Pendidikan Tinggi Pagoh KM1, Jalan Panchor, 84600 Pagoh, Muar, Johor, MALAYSIA.*

*Corresponding Author: saifulah@uthm.edu.my
DOI: https://doi.org/10.30880/ekst.2025.05.02.050

Article Info

Received: 30 December 2024
Accepted: 18 January 2025
Available online: 19 December 2025

Keywords

Household Income, Expenditure, Multiple Linear Regression, Fuzzy Linear Regression, Economic Modelling, Model Selection

Abstract

This study examines the relationship between household income and expenditure in Malaysia using Multiple Linear Regression (MLR) and Fuzzy Linear Regression (FLR) models. Household financial data from the 2023 Household Income and Expenditure Survey (HIES) was analysed to assess model performance using the Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC). The results indicate a strong positive correlation (0.9117) between household income and expenditure, supporting the hypothesis that increased income leads to higher expenditure. Among the models tested, FLR using Tanaka's method demonstrated the best predictive performance, with the lowest AIC (2688.11) and BIC (2697.34) values, outperforming both MLR and FLR (Ni) in handling data uncertainty. In contrast, MLR exhibited significantly higher AIC (120064042) and BIC (120064048.2) values, suggesting a poorer fit for uncertain economic data. These findings highlight the advantages of FLR in modelling complex economic relationships and suggest its potential applications in economic forecasting. Future research should focus on refining FLR techniques further to enhance their accuracy and applicability in economic analysis.

1. Introduction

In this current economic situation, expenditure and household income are the key indicators of consumer behaviour, economic health, and social well-being [1]. To more knowledge, expenditure means financial transactions that individuals, households, and governments undertake to acquire goods and services. It represents the transactions of money and fulfils various needs or desires. This expenditure covers many expenses, from luxury products and entertainment needs items like food, housing, and healthcare [2]. Household income on the other hand, represents the entire amount of money the household members earn over a certain period, usually a year. This household income includes various revenue streams, including employment income, wages, salaries, and bonuses. Analysing the expenditure and household income provides a deep understanding of many social and economic aspects.

A study by [3], says that household income is a valuable economic factor that can develop Malaysia's economy. Hence, the study stated that the household income level is associated with the daily expenditure. The more household income there is, the more expenditure is spent. According to [4], household income and expenditure patterns change in line with the changing times. This is because changing times have become one of the big factors that can indirectly influence household expenditure, as the cost of goods and other services is due to inflation in the economy.

Up to now, Regression modelling becomes one of the popular statistical techniques that can be used to analyse the association between the variables. This method includes several techniques for analysing the data [5]. Multiple linear regression is a statistical technique that can be employed in the model's reliability and accuracy [6]. Broadly, the primary reasons for using the regression model are predicting and interpreting. Predicting is estimating the data value at any time while interpreting is crucial for understanding the relationship between variables. Multiple linear regression finds extensive application in economics, engineering, social sciences, and more [7]. Fuzzy linear regression (FLR) is a technique for the relationship between inputs and outputs in a fuzzy environment. It has been noted that FLR is an extension of classical fuzzy logic that can establish the link between variables in a fuzzy environment [8].

For model selection, selecting an appropriate model is essential for accurate data representation and prediction. Two prevalent criteria guiding the selection are the Akaike Information (AIC) and Bayesian Information Criterion (BIC). AIC estimates the relative quality of statistical models by balancing model fit and complexity [9]. While BIC incorporates sample size into its penalty for model complexity [10]

This study aims to identify the relationship between household income and expenditure in Malaysia using Pearson Correlation, to analyse household income using MLR and FLR by both Tanaka and Ni models and to evaluate the predictive accuracy of these models in understanding expenditure dynamics while addressing the limitations of data uncertainty and precision.

2. Method

2.1 Flowchart

The research framework of this study is shown in Fig. 1. The data were obtained from the Department of Statistics Malaysia (DOSM) from various sectors including household income, expenditure, Gini and poverty. This dataset analysed household-level income, expenditure, Gini and poverty.

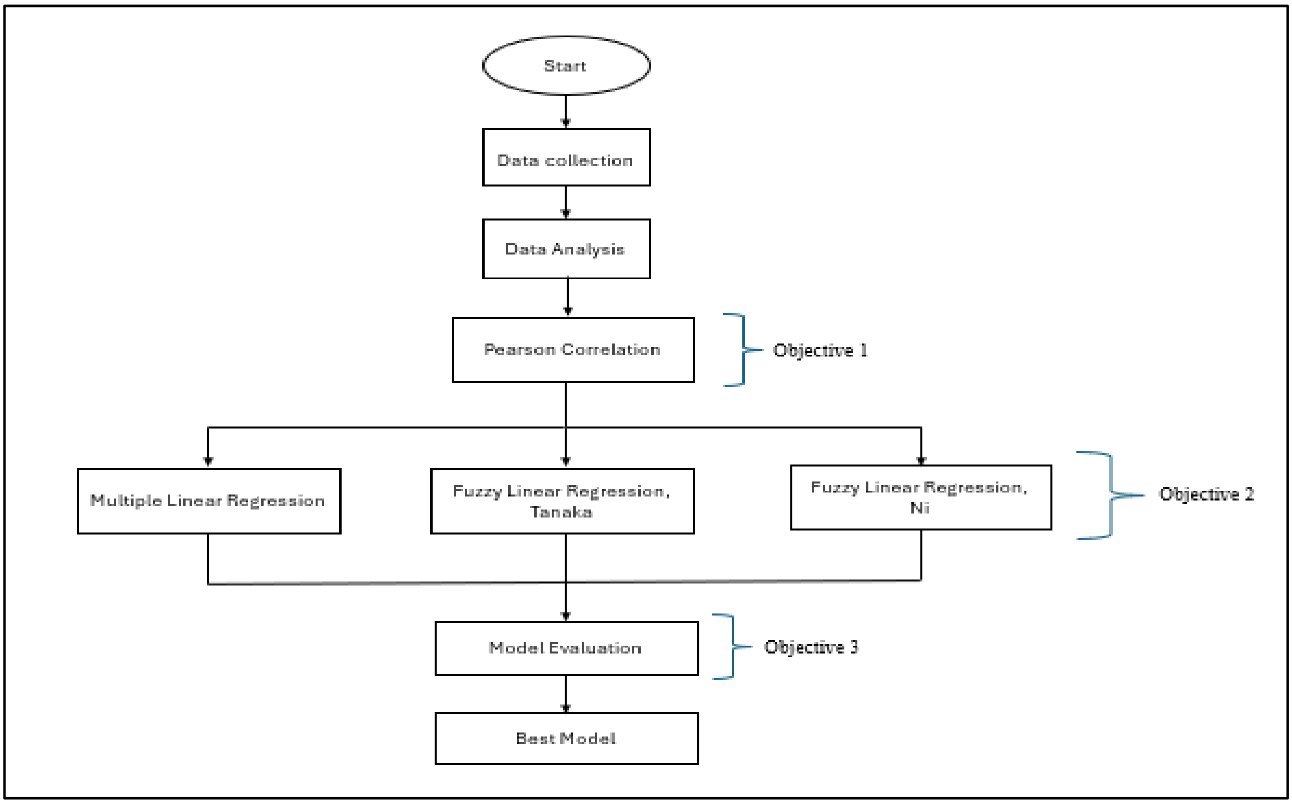

Fig. 1 Research Flowchart

Begin with the data selection, the data and plotted to see if there were any outliers before proceeding to another process of this research. The first method used is Pearson Correlation to evaluate the strength and direction of linear relationships between variables. The second is, Multiple Linear Regression (MLR). MLR helps to predict the outcomes based on the model. Next is, Fuzzy Linear Regression (FLR). FLR can help to represent the better relationship between the independent and dependent variables. It also provides more reliable and insightful results in complex real-world situations. Lastly is the model evaluation process. To ensure a comprehensive evaluation, AIC and BIC are used to find the lowest value between all the models indicated as the best between all the linear regression models.

2.2 Pearson Correlation

Pearson Correlation or Pearson's r measures the strength and direction of the linear relationship between two continuous variables [11]. The Pearson correlation is calculated using the formula in Eq. (1).

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (1)$$

Where:

x_i and y_i are the individual sample points,

$\bar{x}$ is the mean of the x values,

$\bar{y}$ is the mean of the y values,

n is the number of sample points

2.3 Multiple Linear Regression

Multiple linear regression extends beyond simple linear regression by including more than one variable [12]. The term 'linear' is retained because the response variable is associated with linear combination of the explanatory variables [13]. The multiple linear regression follows the same structure as simple linear regression but includes additional terms [14] as in Eq. (2).

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \epsilon \quad (2)$$

Where:

y_i is the dependent variable (the outcome or response variable)

β_0 is the intercept (the value of y when all independent variables are zero)

$\beta_1, \beta_2, \dots, \beta_n$ are the coefficients (slopes) for the independent variables $(x_1, x_2, \dots, x_n)$

$x_1, x_2, \dots, x_n$ are the independent variables (predictors and explanatory variables)

ϵ is the error term (residual), representing the differences between the observed and predicted values of y.

2.4 Fuzzy Linear Regression (Tanaka)

Tanaka first proposed FLR in 1982. According to [15], exploring possible models that fit observed data is the goal of FLR analysis. The fitting formula is typically the basis for the model difference. Eq. (3) display the FLR final model. As in Eq. (4), a linear programming problem is solved to estimate the fuzzy parameters, which yields the fuzzy model.

$$y_i \in [a + bx_i - h(1 + |x_i|), a + bx_i + h(1 + |x_i|)] \quad (3)$$

Where:

a: Intercept

b: Slope

h: Fuzziness (to be minimised)

x_i: Independent variable (input)

y_i: Dependent variable (output)

To estimates the fuzzy parameters $A_e^* = (\alpha_e, \varsigma_e)$ the following linear programming problem,

$$\min_{\alpha,\varsigma} \varsigma_1 + \cdots + \varsigma_g \tag{4}$$

Subject to $\varsigma \geq 0$ and

$$\boldsymbol{\alpha^T x_e} + (1 - \boldsymbol{H}) \sum_f \boldsymbol{\varsigma_f} |\boldsymbol{x_{ef}}| \geq \boldsymbol{y_e} + (1 - \boldsymbol{H})\boldsymbol{\varepsilon_e} \tag{5}$$

$$-\boldsymbol{\alpha^T x_e} + (1 - \boldsymbol{H}) \sum_f \boldsymbol{\varsigma_f} |\boldsymbol{x_{ef}}| \geq -\boldsymbol{y_e} + (1 - \boldsymbol{H})\, \boldsymbol{\varepsilon_e}$$

Where:

$\boldsymbol{A_\varepsilon^*}$= $(\boldsymbol{\alpha_\varepsilon}, \boldsymbol{\varsigma_\varepsilon})$: The estimated fuzzy parameters

$\boldsymbol{\alpha}$: Coefficient vector

$\boldsymbol{\varsigma}$: Slack variable (non-negative)

$\boldsymbol{g}$: Number of constraint or data points

$\boldsymbol{x_e}$: Input vector for a given data point $\boldsymbol{e}$

$\boldsymbol{y_e}$: Output or dependent variable for data point $\boldsymbol{e}$

$\boldsymbol{H}$: A given constant (possibly related to the confidence level in fuzzy regression)

$\boldsymbol{\varepsilon_e}$: Error term associated with data point $\boldsymbol{e}$

$\boldsymbol{c_f}$: Coefficient weight applied to the absolute values of $\boldsymbol{x_{ef}}$

$\boldsymbol{x_{ef}}$: Feature/component $\boldsymbol{f}$ of the input vector $\boldsymbol{x_e}$

$\sum_f$: Summation over all features $\boldsymbol{f}$

2.5 Fuzzy Linear Regression (Ni)

Fuzzy linear regression (FLR) by Ni differs from Tanaka's approach by minimising the total fuzziness while ensuring all data points satisfy the membership functions [16].

$$y_i = (a + bx_i) \pm h(|x_i| + c) \tag{6}$$

Where:

a: Intercept

b: Slope

h: Fuzziness parameter (to be minimised)

c: Constant to control fuzziness for small values of x_i

x_i: Independent variable (input)

y_i: Dependent variable (output)

The following linear programming

$$\min_{\alpha,\varsigma} \varsigma_1 + \cdots + \varsigma_g \tag{7}$$

Subject to $\boldsymbol{\varsigma} \geq 0$ and

$$\boldsymbol{\alpha}^{\boldsymbol{T}} \boldsymbol{x}_{\boldsymbol{e}} + (1 - \boldsymbol{H}) \sum_f \varsigma_f |\boldsymbol{x}_{\boldsymbol{ef}}| \geq \boldsymbol{y}_{\boldsymbol{e}} \tag{8}$$

$$-\boldsymbol{\alpha}^{\boldsymbol{T}} \boldsymbol{x}_{\boldsymbol{e}} + (1 - \boldsymbol{H}) \sum_f \boldsymbol{c}_f |\boldsymbol{x}_{\boldsymbol{ef}}| \geq -\boldsymbol{y}_{\boldsymbol{e}}$$

Where:

$\boldsymbol{\alpha}$: Coefficient vector

$\boldsymbol{\varsigma}$: Slack variable (non-negative)

$\boldsymbol{g}$: Number of constraint or data points

$\boldsymbol{x}_{\boldsymbol{e}}$: Input vector for a given data point $\boldsymbol{e}$

$\boldsymbol{y}_{\boldsymbol{e}}$: Output or dependent variable for data point $\boldsymbol{e}$

$\boldsymbol{H}$: A given constant (possibly related to the confidence level in fuzzy regression)

$\boldsymbol{\varepsilon}_{\boldsymbol{e}}$: Error term associated with data point $\boldsymbol{e}$

$\boldsymbol{c}_{\boldsymbol{f}}$: Coefficient weight applied to the absolute values of $\boldsymbol{x}_{\boldsymbol{ef}}$

$\boldsymbol{x}_{\boldsymbol{ef}}$: Feature/component $\boldsymbol{f}$ of the input vector $\boldsymbol{x}_{\boldsymbol{e}}$

$\sum_f$: Summation over all features $\boldsymbol{f}$

2.6 Model Selection

The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are statistical metrics used for model selection [17]. They assess a model's goodness of fit while penalising for model complexity to avoid overfitting [18]. The formula for AIC as in Eq. (9).

$$AIC = 2k - 2\ln(L) \tag{9}$$

Where:

k: The number of estimated parameters in the model (including the intercept)

L: The maximum likelihood of the model (the likelihood of the data given the model parameters)

A lower AIC value indicates a better fit relative to other models.

The formula for BIC as in Eq. (10). A lower BIC value indicates a better model fit relative to others [19]. BIC imposes a stricter penalty for complexity than AIC, especially as the number of observations (n) increases [20].

$$BIC = kln(n) - 2\ln(L) \tag{10}$$

Where:

k: The number of estimated parameters in the model.

n: The number of observations in the dataset.

L: The maximum likelihood of the model.

3. Results and Discussion

All the important variables have been included in modelling the suggested and other current approaches based on this research. The best model was also determined by the AIC and BIC which have the lowest value indicated as the best model.

3.1 Descriptive Statistics

Table 1 summarizes the descriptive statistics of key variables in the data. Shows the minimum, maximum, mean, and standard deviation for income, expenditure, Gini and poverty. The income means ranges from 3153 to 13673, with an average of 5973.34 and a standard deviation of 2115.77. Expenditure means between 2019 to 8897, with an average 3822.81 and a standard deviation of 1154.553. The Gini coefficient, a measure of income inequality, ranges from 0.22352 to 0.44869, with an average of 0.3437. Lastly, poverty levels range from 0 to 52.7, with an average of 12.11 and a standard deviation of 10.3.

Table 1 Descriptive Analysis

	Minimum	Maximum	Mean	Std. Deviation
Income Mean	3153	13673	5973.34	2115.765
Expenditure Mean	2019	8897	3822.81	1154.553
Gini	0.22352	0.44869	0.3437172	0.03946197
Poverty	0.0	52.7	12.107	10.0338

Table 2 provides additional descriptive statistics for key variables. It shows the first (Q1), second (Q2), and third (Q3) quartiles, interquartile range (IQR), skewness, and kurtosis. For income, the quartiles range from 4633 to 6712, with an IQR of 2079. Income is positively skewed (1.8045) and has a kurtosis of 3.7410, indicating a long right tail and a peaked distribution. Expenditure has quartile ranging from 3042.75 to 4383.75, an IQR of 1341, positive skewness (1.4922), and a kurtosis of 3.0405. The Gini coefficient has an IQR of 0.052273, slight negative skewness (-0.334630, and a low kurtosis (0.2785). Poverty quartiles range from 0.975 to 16.525 with an IQR of 11.55, positive skewness (1.653), and kurtosis of 3.4551, reflecting some extreme values.

Table 2 Descriptive Statistics

	Quartile 1	Quartile 2	Quartile 3	IQR	Skewness	Kurtosis
Income Mean	4633	5324.5	6712	2079	1.8045	3.7410
Expenditure Mean	3042.75	3564.5	4383.75	1341	1.4922	3.0405
Gini	0.316945	0.346585	0.369218	0.052273	-0.33463	0.2785
Poverty	4.975	9.05	16.525	11.55	1.653	3.455129

3.2 Pearson Correlation

Table 3 displays the Pearson Correlation results between household income and expenditure. A strong positive linear relationship is evident, with a correlation coefficient of 0.9117. This mean that as household income increases, expenditure tends to increase proportionally, and vice versa. The diagonal values (1.000) represent the perfect correlation of each variable itself, confirming data consistency.

Table 3 Result of Pearson Correlation

Predictor	Income Mean	Expenditure Mean
Income Mean	1.0000	0.9117
Expenditure Mean	0.9117	1.0000

3.3 Multiple Linear Regression

The multiple linear regression analysis results insight into the relationship between the dependent and independent variables. The regression equation, $\hat{Y} = -412.35 + 1.67x_i$, indicates that for every unit increase in the expenditure mean (x_i), the predicted dependent variable increases by 1.67 units, while the intercept of -412.35 represents the predicted value when $\boldsymbol{x_i}$ is zero. The coefficient for expenditure mean is highly significant, with a very low p-value (1.86×10^{-62}), indicating a strong relationship. The intercept is not statistically significant, as its p-value is above 0.05.

$$\hat{Y} = -412.35 + 1.67x_1 \tag{11}$$

Table 4 shows the statistical details of the regression model.

Table 4 Result of Multiple Linear Regression Analysis

	Coefficient	Standard Error	t-Stat	p-value
Intercept	-4123494593	239.8923	-1.7889	0.087605
Expenditure Mean	1.670090538	0.060147	27.76668	1.86E-62

3.4 Fuzzy Linear Regression (Tanaka)

This equation and Table 5 are part of a fuzzy linear regression (FLR) model based on Tanaka's approach. The equation shows the predicted value $(\hat{Y})$ as a linear function of x_i, where the coefficients (A_0 *and* A_1) are represented by fuzzy numbers. Each fuzzy number has a center (a_q), which is the main value, and a width (c_q), representing uncertainty. In this case, A_0 has a center of 943.5007 with no uncertainty (width = 0), and A_1 has a center of 1.3954 and a width of 1.1661, indicating some fuzziness or variability in this parameter.

$$\hat{Y} = (943.5007{,}0) + (1.3954{,}0)x_i \tag{12}$$

Table 5 Detailed of the antecedent parameter of FLR (Tanaka)

Fuzzy Parameters	Fuzzy Center, a_q	Fuzzy Width, c_q
A_0	943.5007	0
A_1	1.3954	1.1661

3.5 3.4 Fuzzy Linear Regression (Ni)

The parameters of the model are detailed in the table, where A_0 is the fuzzy intercept with a center value of -943.50 and no uncertainty (fuzzy width = 0), and A_1 is the fuzzy slope with a center value of -1.41 and no uncertainty as well. This indicates that the model describes a precise linear relationship between x_i and $\hat{Y}$, as no fuzziness or variability is present in the parameters.

$$\hat{Y} = (-943.50) - (1.41{,}0)x_i \tag{13}$$

Table 6 Detailed of the antecedent parameter of FLR (Ni)

Fuzzy Parameters	Fuzzy Center, a_q	Fuzzy Width, c_q
A_0	-943.50	0
A_1	-1.41	0

3.6 Summary of Results

Table 7 summarises the results for the MLR, FLR of Tanaka and the FLR of Ni models using AIC and BIC. Lower AIC and BIC values indicate a better model fit with fewer complexities. The MLR has extremely high AIC (12006404.2) and BIC (12006404.8) values, showing a poor fit compared to the FLR models. Tanaka's FLR has lower AIC (2688.11302) and BIC (2697.33854), suggesting it balances accuracy and simplicity well. Ni's FLR also performs better than MLR but has a slightly higher AIC (2692.77494) and a much higher BIC (3406.1999). Overall, Tanaka's FLR is the most efficient model in this comparison.

Table 7 Summary Result for MLR, FLR Tanaka and FLR Ni

Model of Linear Regression	AIC	BIC
MLR	120064042	120064048.2
FLR (Tanaka)	2688.11302	2697.33854
FLR (Ni)	2692.77494	3406.1999

4. Conclusion and Recommendations

This study examined the relationship between household income and expenditure in Malaysia using MLR and FLR models. The analysis revealed that FLR by Tanaka method performed better than both MLR and FLR (Ni) as it lowered values for AIC and BIC. In contrast, MLR exhibited significantly higher AIC and BIC values, indicating a poorer fit in accounting for uncertainty within the dataset. Although traditionally regression models such as MLR are widely used, FLR by Tanaka's approach proved more effective in handling data uncertainty and imprecise relationships. These findings suggest that FLR can be a valuable tool in modelling complex economic relationships, especially when precise data is not always available. For improvement, future studies should explore enhancements to FLR methodologies to improve accuracy, particularly in handling uncertainty in economic data.

Acknowledgement

The research was supported by Universiti Tun Hussein Onn Malaysia (UTHM) through Multidisciplinary Research Grant (MDR) Vot Q702 and GPPS Grant Vot Q668.

Conflict of Interest

The authors declare no conflict of interest in the paper's publication

Author Contribution

The authors confirm their contribution to the paper as follows: ***study conception and design:*** *Nurathirah Rodzi, Mohd Saifullah Rusiman;* ***data collection:*** *Nurathirah Rodzi;* ***analysis and interpretation of results:*** *Nurathirah Rodzi, Mohd Saifullah Rusiman, Norziha Che Him;* ***draft manuscript preparation:*** *Nurathirah Rodzi, Mohd Saifullah Rusiman. All authors reviewed the results and approved the final version of the manuscript*

References

[1] Yigzaw Tilahun, M., Seyoum Tsehay, A., & Kebede Teklie, D. (2021). Rural Household Poverty and its Determining Factors: A Poverty Analysis Using Alternative Measurement Approaches. *International Journal of Advanced Research*, *9*(07), 148–163. https://doi.org/10.21474/ijar01/13111

[2] Serván-Mori, E., Islam, M. D., Kaplan, W. A., Thrasher, R., & Wirtz, V. J. (2022). Out-of-pocket expenditure on medicines in Bangladesh: An analysis of the national household income and expenditure survey 2016-17. *PLoS ONE*, *17*(9 Septamber). https://doi.org/10.1371/journal.pone.0274671

[3] May, T. K., Saharan, S., & Rusiman, M. S. (2023). *A*nalysis of Income and Expenditure of Households in Peninsular *Malaysia* (pp. 187–200). https://doi.org/10.1007/978-981-99-2850-7_17

[4] Othman, Z. A., Bakar, A. A., Sani, N. S., & Sallim, J. (2020). Household overspending model amongst B40, M40 and T20 using classification algorithm. *International Journal of Advanced Computer Science and Applications*, *11*(7), 392–399. https://doi.org/10.14569/IJACSA.2020.0110751

[5] Etemadi, S., & Khashei, M. (2021). Etemadi multiple linear regression. *Measurement: Journal of the International Measurement Confederation*, *186*. https://doi.org/10.1016/j.measurement.2021.110080

[6] Semwal, J., Bahuguna, A., Uniyal, A., & Vyas, S. (2023). A study to analyse COVID-19 outbreak using multiple linear regression: a supervised machine learning approach. *National Journal of Community Medicine*, *14*(02), 82–89. https://doi.org/10.55489/njcm.140220232656

[7] Dhaval, B., & Deshpande, A. (2020). Short-term load forecasting with using multiple linear regression. *International Journal of Electrical and Computer Engineering*, *10*(4), 3911–3917. https://doi.org/10.11591/ijece.v10i4.pp3911-3917

[8] Gkountakou, F. I., Konguetsof, A., Souliotis, G., & Papadopoulos, B. K. (2023). A Novel Approach to Fuzzy Implication Through Fuzzy Linear Regression. *Mathematical Modelling of Engineering Problems*, *10*(6), 1927–1936. https://doi.org/10.18280/mmep.100602

[9] Zhang, J., Yang, Y., & Ding, J. (2023). Information criteria for model selection. *Wiley Interdisciplinary Reviews Computational Statistics*, *15*(5). https://doi.org/10.1002/wics.1607

[10] Sukumar, G. (2024, January 15). AIC, BIC, Penalties and the Science behind Model selection. *Medium*. https://medium.com/%40gopal.sukumar/aic-bic-penalties-and-the-science-behind-model-selection-bd0d4d06d335

[11] Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ane.0000000000002864

[12] Trunfio, T. A., Scala, A., Giglio, C., Rossi, G., Borrelli, A., Romano, M., & Improta, G. (2022). Multiple regression model to analyze the total LOS for patients undergoing laparoscopic appendectomy. *BMC Medical Informatics and Decision Making*, *22*(1). https://doi.org/10.1186/s12911-022-01884-9

[13] Yee, G. P., Rusiman, M. S., Ismail, S., Suparman, Hamzah, F. M., & Shafi, M. A. (2023). K-means clustering analysis and multiple linear regression model on household income in Malaysia. *IAES International Journal of Artificial Intelligence*, *12*(2), 731–738. https://doi.org/10.11591/ijai.v12.i2.pp731-738

[14] Hijji, Mohammad & Alam, Gulzar. (2021). A Multivocal Literature Review on Growing Social Engineering Based Cyber-Attacks/ Threats During the COVID-19 Pandemic: Challenges and Prospective Solutions. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.3048839.

[15] Ammar Shafi, M., Saifullah Rusiman, M., & Syuhada Abdullah, S. N. (2021). Application of fuzzy linear regression with symmetric parameter for predicting tumor size of colorectal cancer. *Mathematics and Statistics*, *9*(1), 36–40. https://doi.org/10.13189/ms.2021.090106

[16] Ramly, N., Rusiman, M. S., Him, N. C., Nor, M. E., Man, S., Basri, N. Z. A., & Mohamad, N. (2018). A new hybrid of Fuzzy C-Means Method and Fuzzy Linear Regression Model in Predicting Manufacturing Income. *International Journal of Engineering & Technology*, *7*(4.30), 473. https://doi.org/10.14419/ijet.v7i4.30.22371

[17] Brewer, M. J., Butler, A., & Cooksley, S. L. (2016). The relative performance of AIC, AICC and BIC in the presence of unobserved heterogeneity. *Methods in Ecology and Evolution*, *7*(6), 679–692. https://doi.org/10.1111/2041-210x.12541

[18] Chakrabarti, A., & Ghosh, J. (2011). *AIC, BIC and recent advances in model selection*. https://www.semanticscholar.org/paper/AIC%2C-BIC-and-Recent-Advances-in-Model-Selection-Chakrabarti-Ghosh/791c825fd4ee0cf790b794329204da6f73c6405b

[19] Banerjee, S. (2023, October 20). Model Magic with AIC & BIC: Navigating Fit and Elegance | Medium. *Medium*. https://shekhar-banerjee96.medium.com/model-magic-aic-bic-mdl-navigating-fit-and-elegance-726c784edf9b

[20] Kumar, A. (2023, November 10). *AIC & BIC for Selecting Regression models: Formula, examples*. Analytics Yogi. https://vitalflux.com/aic-vs-bic-for-regression-models-formula-examples/