Prediksi Tingkat Kesejahteraan Masyarakat Di Kelurahan North Wangurer Menggunakan Regresi Linear Berganda

Prediction of the Community Welfare in North Wangurer Village Using Multiple Linear Regression

Liza Wikarsa[1)], Steven Pandelaki[2)], Karen Gisella Sumajouw[3)]

[123] Universitas Katolik De La Salle Manado
[123] Kairagi I Kombos Manado, Belakang Wenang Permai II, Manado, 95000, Telp. 0431-813148, Fax. 0431-813160

lwikarsa@unikadelasalle.ac.id[1)], spandelaki@unikadelasalle.ac.id[2)], 18013101@unikadelasalle.ac.id[3)]

Diterima : 6 Juni 2023 || Revisi : 10 Agustus 2023 || Disetujui: 26 November 2023

Abstrak – Tingkat kesejahteraan masyarakat suatu negara dapat menentukan kualitas dan kondisi negara itu sendiri. Kelurahan Wangurer Utara berada di kecamatan Madidir Kota Bitung dengan jumlah penduduk sekitar 750 kepala keluarga. Tingkat kesejahteraan masyarakat di kelurahan ini hanya berdasarkan pendapatan bulanan yang diperoleh dari masing-masing masyarakat dianggap tidak efektif. Penelitian ini bertujuan untuk memprediksi tingkat kesejahteraan masyarakat di kecamatan tersebut dengan menggunakan metode regresi linier berganda. Penelitian ini diharapkan dapat memberikan wawasan bagi pihak kelurahan Wangurer Utara untuk membuat kebijakan/keputusan yang lebih efektif untuk meningkatkan tingkat kesejahteraan dengan harapan dapat mengentaskan kemiskinan dan pemerataan bantuan sosial kepada rumah tangga sasaran. Ada empat variabel independen yang digunakan dalam penelitian ini seperti pendapatan, pendidikan, pekerjaan, dan jumlah anggota keluarga. Sedangkan variabel terikatnya adalah tingkat kesejahteraan masyarakat yang terdiri dari Pra-KS, KS-I, KS-II, KS-III, dan KS-III Plus yang diakui oleh BKKBN. Hasil penelitian menunjukkan bahwa tingkat kesejahteraan masyarakat Kelurahan Wangurer Utara berada pada Keluarga Sejahtera III Plus (Level 5). Hampir semua keluarga (98,7%) di kelurahan ini dapat memenuhi semua kebutuhan dasar, psikologi sosial dan perkembangannya serta rasa tanggung jawab diri (*self-esteem*).

Kata Kunci: Kesejahteraan masyarakat, regresi linier berganda, keluarga sejahtera

Abstract *– The level of social welfare in a country can determine the quality and condition of the country itself. North Wangurer sub-district is in Madidir District, Bitung City, with a population of around 750 households. The level of community welfare in this sub-district is solely based on the monthly income obtained from each community, which is considered ineffective. Thus, this research aimed to predict the level of community welfare at this sub-district using the multiple linear regression method. It was hoped that this research could provide insights for the North Wangurer sub-district office to make more effective policies/decisions to increase the level of welfare to eradicate poverty and equitably distribute social assistance to the targetted households. Four independent variables were employed in this research: income, education, occupation, and the number of family members. Meanwhile, the dependent variable was the level of community welfare consisting of Pra-KS, KS-I, KS-II, KS-III, and KS-III Plus, as acknowledged by BKKBN. The results revealed that the North Wangurer sub-district community welfare level was in Prosperous Family III Plus (Level 5). Most families (98,7%) in this sub-district can meet all basic needs, social psychology and its development, and self-accountability (self-esteem).*

Keywords: *Community welfare, multiple linear regression, prosperous family*

INTRODUCTION

The level of social welfare in a country can determine the quality and condition of the country itself. Community welfare can be measured by the ability of each person to meet their needs. In government, people's welfare can be seen by the level of success in achieving goals in increasing the community's welfare. So far, the assessment of the

doi:

level of social welfare has always been associated with income or based on economic factors, which can be easily measured, observed, and compared (Kaluku & Pakaya, 2020).

North Wangurer sub-district is in Madidir District, Bitung City, with a population of around 750 households (Kepala Keluarga/KK) spread across 5 areas and has 21 neighborhood associations (Rukun Tetangga/RT). The level of community welfare in this sub-district is solely based on the monthly income obtained from each community, which is considered ineffective. In this regard, Iskandar et al. (2010) and Nofriani (2020) pointed out that there are at least 4 factors to determine the level of community welfare: income, education, employment, and the number of community members. Furthermore, the National Community Planning Population Agency (BKKBN) added that the level of community welfare can be grouped into 5, namely Pre-Prosperous Families (Pra-KS), Prosperous Families I (KS-I), Prosperous Families II (KS-II), Prosperous Families III (KS-III), Prosperous Community III Plus (KS-III Plus) (Masawah, 2018; Astuti, Adyatma, & Normelani, 2017; BKKBN, 2011; Cugung, 2020).

Cugung (2020), Masawah (2018), and BKKBN (2011) explained that there were several basic human needs to fulfill, such as food, clothing, housing, health, and education. They also further elaborated on the five levels of community welfare. Firstly, Pre-Prosperous Families do not meet one of the 6 indicators of Prosperous Family I (KS I) or the indicator of "basic needs of the family" (basic needs). Secondly, Prosperous Family I is for a family capable of fulfilling the 6 indicators of the KS I stage but not fulfilling one of the 8 indicators of a Prosperous Family II or the indicator of the "psychological needs" of the family. Thirdly, Prosperous Families II is for families that can fulfill 6 indicators of the KS I stage and 8 indicators of the KS II stages but do not fulfill one of the 5 indicators of Prosperous Family III (KS III), or the indicator "development needs" (developmental needs) of the family. Fourthly, Prosperous Family III is a family that can fulfill 6 indicators of the KS I stage, 8 indicators of KS II, and 5 indicators of KS III, but does not fulfill one of the 2 indicators of Prosperous Family III Plus (KS III Plus) or indicators of the family "self-actualization" (self-esteem). Lastly, Prosperous Families III Plus is for families that can fulfill all 6 indicators of the KS I stage, 8 of the KS II stage, 5 of the KS III stage, and 2 of the KS III Plus stage.

To determine the level of community welfare in North Wangurer Village, this research will use the multiple linear regression method that involves the relationship of more than one independent variable. This method is widely used to identify the relationship intensity between the independent variables and the dependent variables using a straight line as well as to analyze the direction and influence of the independent variables on the dependent variable (Izzah & Widyastuti, 2017; Afkarina, Widodo, & Furqon, 2019). It is also suitable for use in this study because there is more than one independent variable. Additionally, this study will measure how much influence the independent variables have on the dependent variable.

The application of the multiple linear regression method to predict the dependent variable has been widely carried out in various studies (Hayuningtyas & Sari, 2022; Alfarizi & Andri, 2021; Herwanto, Widyaningtyas, & Indriana, 2019; (Sulardi, Hendro, & Umbara, 2017; (Kamal, Hendro, & Ilyas, 2017; (Marbun, Sihotang, & Nababan, 2018). The research conducted by Hayuningtyas and Sari discussed the prediction of diabetes using multiple linear regression with 8 independent variables, namely pregnancies, glucose, blood pressure, skin thickness, insulin, BMI, diabetes pedigree function, and age. Alfarizi and Andri also conducted another study on “predicting production of NPK Fertilizer Formula 15-15-15 PSO at PT. Pupuk Sriwidjaja” using the Multiple Linear Regression method (p.51). Furthermore, Herwanto, Widyaningtyas, and Indriana published their research on applying the multiple linear regression method to predict rice crop yields using 6 independent variables with 300 data.

This research differed from the previous studies as it built a web-based application that could predict the level of community welfare at the North Wangurer sub-district using the multiple linear regression method. Also, there were only four independent variables employed in this research such as income (x_1), education (x_2), occupation (x_3), and the number of family members (x_4). Meanwhile, the dependent variable was the level of community welfare (y) consisting of Pra-KS, KS-I, KS-II, KS-III, and KS-III Plus as acknowledged by BKKBN.

The hypotheses for this research were as follows:

H_0: The independent variables have no effect on the dependent variable.
H_1: There is an effect of income (x_1) on the level of welfare (y).
H_2: There is an effect of education (x_2) on the level of welfare (y).
H_3 : There is an influence of work (x_3) on the level of welfare (y).
H_4: There is an effect of the number of family members
(x_4) on the level of welfare (y).

The results of this research would provide insights for the North Wangurer sub-district office regarding the level of welfare in this sub-district. Hopefully, it can also help the sub-district office to make more effective policies/decisions to increase the level of welfare in the hope of eradicating poverty and equitably distributing social assistance to the targetted households.

This paper was organized into four parts: introduction, research methodology, results and discussion, and conclusions.

RESEARCH METHOD

Research Variables

Four independent variables were used in this research: income, education, occupation, and the number of family members. While the dependent variable is the level of community welfare that is differentiated into five levels with accord to BKKBN's prosperous family indicators, namely Pra-KS, KS-I, KS-II, KS-III, and KS-III Plus.

Research Sample

The population of the North Wangurer sub-district is 750 families. Based on Isaac and Michael's table with a 5% of significance level, the number of samples required for this research was at least 238 samples.

Data Collection

The data was gathered from interviews, questionnaires, and a study of documents The North Wangurer sub-district officers were interviewed to investigate the community welfare. Also, the questionnaires were distributed twice to the respondents using Google Forms. 30 respondents filled in the first questionnaires to find out the independent variables as well as how well the questionnaire was perceived. Meanwhile, the second questionnaire was distributed to 238 respondents to gather their responses regarding the community welfare in their district.

Statistical Testing of Research Instruments

The validity of the questionnaires was as follows:

Correlations

		Pendapatan	Pendidikan	Pekerjaan	Jumlah_anggota_keluarga
Pendapatan	Pearson Correlation	1	.757**	.817**	.803**
	Sig. (2-tailed)		.000	.000	.000
	N	30	30	30	30
Pendidikan	Pearson Correlation	.757**	1	.771**	.679**
	Sig. (2-tailed)	.000		.000	.000
	N	30	30	30	30
Pekerjaan	Pearson Correlation	.817**	.771**	1	.900**
	Sig. (2-tailed)	.000	.000		.000
	N	30	30	30	30
Jumlah_anggota_keluarga	Pearson Correlation	.803**	.679**	.900**	1
	Sig. (2-tailed)	.000	.000	.000	
	N	30	30	30	30

**. Correlation is significant at the 0.01 level (2-tailed).

Figure 1 The Validity of the Questionnaires

Figure 1 showed that the significance value (2-tailed) for each item in the questionnaires was less than 0,05 and the Pearson correlation was positive. The r-count for each item is also bigger than the r-table (0,3493) (Amanda, Yanuar, & Devianto, 2019). Therefore, the questionnaires were considered valid (Budiastuti & Bandur, 2018).

Figure 2 depicted the reliability of the questionnaire. The Cronbach's Alpha value for these questionnaires was 0.935 which is bigger than 0.6 and thus it was considered reliable and valid (Amanda, Yanuar, & Devianto, 2019; Budiastuti & Bandur, 2018).

Reliability Statistics

Cronbach's Alpha	N of Items
.935	4

Figure 2 The Reliability of the Questionnaires

Multiple linear regression analysis is used to determine the linear relationship between the independent and dependent variables. F-test and T-test are needed to test whether there were the found model and estimated coefficients in the sample. The F-test is to test the overall model where the $R^2 = 0$, meaning that the independent variables do not influence the dependent variable (Zahra & Rina, 2018). This is to see how fit the model is. On the other hand, T-test would be performed to analyze the significance of

each coefficient and the intercept. Also, a t-test was used to test the null hypothesis whether the parameter's value is equal to zero.

F-test of the Research Instrument

ANOVA Table

			Sum of Squares	df	Mean Square	F	Sig.
Tingkat_kesejahteraan * Pendapatan	Between Groups	(Combined)	5.841	6	.973	13.495	.000
		Linearity	5.097	1	5.097	70.656	.000
		Deviation from Linearity	.744	5	.149	2.063	.107
	Within Groups		1.659	23	.072		
	Total		7.500	29			
			Sum of Squares	df	Mean Square	F	Sig.
Tingkat_kesejahteraan * Pendidikan	Between Groups	(Combined)	5.171	5	1.034	10.660	.000
		Linearity	4.410	1	4.410	45.454	.000
		Deviation from Linearity	.761	4	.190	1.962	.133
	Within Groups		2.329	24	.097		
	Total		7.500	29			
			Sum of Squares	df	Mean Square	F	Sig.
Tingkat_kesejahteraan * Pekerjaan	Between Groups	(Combined)	5.000	7	.714	6.286	.000
		Linearity	4.355	1	4.355	38.325	.000
		Deviation from Linearity	.645	6	.107	.946	.483
	Within Groups		2.500	22	.114		
	Total		7.500	29			
			Sum of Squares	df	Mean Square	F	Sig.
Tingkat_kesejahteraan * Jumlah_anggota_keluarga	Between Groups	(Combined)	5.321	6	.887	9.363	.000
		Linearity	4.784	1	4.784	50.502	.000
		Deviation from Linearity	.538	5	.108	1.136	.370
	Within Groups		2.179	23	.095		
	Total		7.500	29			

Figure 3 The F-test

Based on the significant value in Figure 3, the Deviation from the Linearity Sig value is greater than 0.05. Table 1 revealed that the values of the F-count were less than the values of the F-table. This means that the hypotheses of H_1, H_2, H_3, and H_4 were accepted as there was a significant linear relationship between the independent variables (x_1, x_2, x_3, and x_4) and the dependent variable (y) (Zahra & Rina, 2018).

Table 1 The Values of the F-test

Independent Variables	F-count	F-table
Pendapatan (x_1)	2.063	2,64
Pendidikan (x2)	1.962	2,78
Pekerjaan (x3)	0.946	2,55
Jumlah anggota keluarga (x4)	1.136	2,64

T-test of the Research Instrument

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pendapatan	4.1133	30	.46885	.08560
	Tingkat_kesejahteraan	4.5000	30	.50855	.09285
Pair 2	Pendidikan	4.3833	30	.39791	.07265
	Tingkat_kesejahteraan	4.5000	30	.50855	.09285
Pair 3	Pekerjaan	4.2833	30	.51165	.09341
	Tingkat_kesejahteraan	4.5000	30	.50855	.09285
Pair 4	Jumlah_anggota_keluarga	4.3667	30	.40329	.07363
	Tingkat_kesejahteraan	4.5000	30	.50855	.09285

Figure 4 The T-test For Paired Samples Statistics

The data of paired samples in Figure 4 suggested as follows:

a. In Pair 1, The mean of Pendapatan (4.1133) < the mean of Tingkat_Kesejahteraan (4.5000)
b. In Pair 2, The mean of Pendidikan (4.3833) < the mean of Tingkat_Kesejahteraan (4.5000)
c. In Pair 3, The mean of Pekerjaan (4.2833) < the mean of Tingkat_Kesejahteraan (4.5000)
d. In Pair 4, The mean of Jumlah_anggota_keluarga (4.3667) < the mean of Tingkat_Kesejahteraan (4.5000)

The results revealed a difference in the average results of the independent variables with the results of the dependent variable (Zahra & Rina, 2018). The results of the paired sample T-test shown in Figure 5 were provided to prove whether the difference is significant.

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Pendapatan & Tingkat_kesejahteraan	30	.824	.000
Pair 2	Pendidikan & Tingkat_kesejahteraan	30	.767	.000
Pair 3	Pekerjaan & Tingkat_kesejahteraan	30	.762	.000
Pair 4	Jumlah_anggota_keluarga & Tingkat_kesejahteraan	30	.799	.000

Figure 5 The T-test For Paired Samples Statistics

Figure 5 presented the correlation and significant values for Pair 1-4. Since the Sig value for each pair is less than 0.05, it can be concluded that H_0t was rejected while the hypotheses of H_1, H_2, H_3, and H_4 were accepted. This indicated that there was a relationship between the independent variables and the dependent variable.

	t	df	Sig. (2-tailed)
Pair 1 Pendapatan – Tingkat_Kesejahteraan	-7.250	29	0.000
Pair 2 Pendidikan – Tingkat_Kesejahteraan	-1.957	29	0.006
Pair 3 Pekerjaan – Tingkat_Kesejahteraan	-3.372	29	0.002
Pair 4 Jumlah_anggota_keluarga – Tingkat_kesejahteraan	-2.386	29	0.024

Figure 6 Paired Samples Test

In Figure 6, the Sig value is less than 0.05, this indicated that H_1, H_2, H_3, and H_4 were accepted. This means that the independent variables have influences on the dependent variable.

Steps in the Multiple Linear Regression

The following are the steps to implement the multiple linear regression method (Afkarina, Widodo, & Furqon, 2019):

1. Create a table calculation

This table must contain the independent and dependent variables.

The calculation:

$= x_1^2, x_2^2, x_3^2, x_4^2, x_1y, x_2y, x_3y, x_4y, x_1x_2, x_1x_3, x_1x_4, x_2x_3, x_2x_4, x_3x_4.$ (1)

2. After the calculation table for each variable was made, it is then arranged in the form of a matrix.

$$\begin{bmatrix} n & \sum x_1 & \sum x_2 & \sum x_3 & \sum x_4 \\ \sum x_1 & \sum x_1^2 & \sum x_2 x_1 & \sum x_3 x_1 & \sum x_4 x_1 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 & \sum x_3 x_2 & \sum x_4 x_2 \\ \sum x_3 & \sum x_1 x_3 & \sum x_2 x_3 & \sum x_3^2 & \sum x_4 x_3 \\ \sum x_4 & \sum x_1 x_4 & \sum x_2 x_4 & \sum x_3 x_4 & \sum x_4^2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} \sum y \\ \sum x_1 y \\ \sum x_2 y \\ \sum x_3 y \\ \sum x_4 y \end{bmatrix} \quad (2)$$

3. Find the determinant value of each matrix using the Sarrus method.

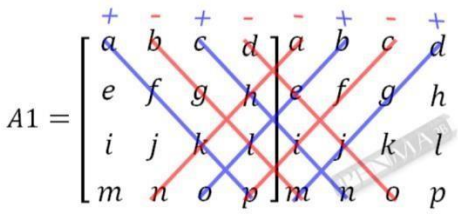

Figure 7 Calculation of Determinant Values

For example:

The value of A1 using the rules of Sarrus method.

= afkp – bglm + chin – dejo – ahkn + belo – cfip + dgjm.

4. Look for the values of b_1, b_2, b_3, b_4, and b_5 that will later be used for the values of the multiple linear regression equation.

$$b_n = \frac{\det An}{\det A} \quad (3)$$

$y = \alpha + b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \cdots b_n x_n$ (4)

Note:

y: dependent variable

x: independent variable

α: constant

b_0, b_1, b_2, b_3, b_n: regression coefficient value

5. Determine the range value of Y for the five levels of community welfare as shown in Table 2 below.

Table 2 The Value Range of Y

Level of the Community Welfare	Range Value
1	1 - 1,7
2	1,8 - 2,5
3	2,6 - 3,3
4	3,4 - 4,1
5	4,2 - 5

Analysis of the System

Identification of Problems

Data collection and determination of the level of welfare carried out by the Wangurer Utara sub-district office is still done manually. The level of welfare is only based on the monthly income each family earns. This method was ineffective and inaccurate to determine the level of community welfare in this district.

Analysis of Targeted Users

The targetted users for this application were the North Wangurer sub-district officers in hope of making more effective policies/decisions to increase the level of welfare and quality for those people who live in that sub-district.

Defined Requirement Specifications

There were several requirement specifications identified in this research that were as follows:

1. The users can upload a CSV file that contained the collected data from the questionnaires.
2. The application can display the data from the CSV file in the form of a table.
3. The application can show the results generated in each step of the multiple linear regression method.
4. The application can indicate the level of welfare for each respondent as well as calculate the total number of families belonging to each level of community welfare in the Wangurer sub-district.

Analysis of the Multiple Linear Regression Method

The following are the steps to implement the multiple linear regression method (Afkarina, Widodo, & Furqon, 2019):

1. Create a table calculation
   This table must contain several parameters, such as age, gender, income (x_1), education (x_2), occupation (x_3) number of family members (x_4), level of welfare (y), and value range for each respondent.
   a. The calculation would be:
      = x_1^2, x_2^2, x_3^2, x_4^2, x_1y, x_2y, x_3y, x_4y, x_1x_2, x_1x_3, x_1x_4, x_2x_3, x_2x_4, x_3x_4.
2. After the calculation table for each variable was made, it is then arranged in the form of a matrix.
   Using the formula in (2), the following matrix was generated.

$$A = \begin{bmatrix} n & \sum x_1 & \sum x_2 & \sum x_3 & \sum x_4 \\ \sum x_1 & \sum x1^2 & \sum x_2x_1 & \sum x_3x_1 & \sum x_4x_1 \\ \sum x_2 & \sum x_1x_2 & \sum x2^2 & \sum x_3x_2 & \sum x_4x_2 \\ \sum x_3 & \sum x_1x_3 & \sum x_2x_3 & \sum x3^2 & \sum x_4x_3 \\ \sum x_4 & \sum x_1x_4 & \sum x_2x_4 & \sum x_3x_4 & \sum x4^2 \end{bmatrix}$$

$$A1 = \begin{bmatrix} Y & \sum x_1 & \sum x_2 & \sum x_3 & \sum x_4 \\ \sum x_1y & \sum x1^2 & \sum x_2x_1 & \sum x_3x_1 & \sum x_4x_1 \\ \sum x_2y & \sum x_1x_2 & \sum x2^2 & \sum x_3x_2 & \sum x_4x_2 \\ \sum x_3y & \sum x_1x_3 & \sum x_2x_3 & \sum x3^2 & \sum x_4x_3 \\ \sum x_4y & \sum x_1x_4 & \sum x_2x_4 & \sum x_3x_4 & \sum x4^2 \end{bmatrix}$$

$$A2 = \begin{bmatrix} n & Y & \sum x_2 & \sum x_3 & \sum x_4 \\ \sum x_1 & \sum x_1y & \sum x_2x_1 & \sum x_3x_1 & \sum x_4x_1 \\ \sum x_2 & \sum x_2y & \sum x2^2 & \sum x_3x_2 & \sum x_4x_2 \\ \sum x_3 & \sum x_3y & \sum x_2x_3 & \sum x3^2 & \sum x_4x_3 \\ \sum x_4 & \sum x_4y & \sum x_2x_4 & \sum x_3x_4 & \sum x4^2 \end{bmatrix}$$

$$A3 = \begin{bmatrix} n & \sum x_1 & Y & \sum x_3 & \sum x_4 \\ \sum x_1 & \sum x1^2 & \sum x_1y & \sum x_3x_1 & \sum x_4x_1 \\ \sum x_2 & \sum x_1x_2 & \sum x_2y & \sum x_3x_2 & \sum x_4x_2 \\ \sum x_3 & \sum x_1x_3 & \sum x_3y & \sum x3^2 & \sum x_4x_3 \\ \sum x_4 & \sum x_1x_4 & \sum x_4y & \sum x_3x_4 & \sum x4^2 \end{bmatrix}$$

$$A4 = \begin{bmatrix} n & \sum x_1 & \sum x_2 & Y & \sum x_4 \\ \sum x_1 & \sum x1^2 & \sum x_2x_1 & \sum x_1y & \sum x_4x_1 \\ \sum x_2 & \sum x_1x_2 & \sum x2^2 & \sum x_2y & \sum x_4x_2 \\ \sum x_3 & \sum x_1x_3 & \sum x_2x_3 & \sum x_3y & \sum x_4x_3 \\ \sum x_4 & \sum x_1x_4 & \sum x_2x_4 & \sum x_4y & \sum x4^2 \end{bmatrix}$$

$$A5 = \begin{bmatrix} n & \sum x_1 & \sum x_2 & \sum x_3 & Y \\ \sum x_1 & \sum x1^2 & \sum x_2x_1 & \sum x_3x_1 & \sum x_1y \\ \sum x_2 & \sum x_1x_2 & \sum x2^2 & \sum x_3x_2 & \sum x_2y \\ \sum x_3 & \sum x_1x_3 & \sum x_2x_3 & \sum x3^2 & \sum x_3y \\ \sum x_4 & \sum x_1x_4 & \sum x_2x_4 & \sum x_3x_4 & \sum x_4y \end{bmatrix}$$

The results could be seen as follows:

n = 30
y = 128,6
$\sum x_1$ = 123,4
$\sum x_2$ = 131,5
$\sum x_3$ = 128,5
$\sum x_4$ = 131

Table 3 Matrix A

n	**$\sum x_1$**	**$\sum x_2$**	**$\sum x_3$**	**$\sum x_4$**
30	123,4	131,5	128,5	131
123,4	513,96	545	534,25	543,25
131,5	545	581	567,81	577,38
128,5	534,25	567,81	558	566,5
131	543,25	577,38	566,5	576,75

Table 4 Matrix A_1

y	$\sum x_1$	$\sum x_2$	$\sum x_3$	$\sum x_4$
128,6	123,4	131,5	128,5	131
534,12	513,96	545	534,25	543,25
567,80	545	581	567,81	577,38
556,64	534,25	567,81	558	566,5
565,97	543,25	577,38	566,5	576,75

Table 5 Matrix A_2

n	y	$\sum x_2$	$\sum x_3$	$\sum x_4$
30	128,6	131,5	128,5	131
123,4	534,12	545	534,25	543,25
131,5	567,80	581	567,81	577,38
128,5	556,64	567,81	558	566,5
131	565,97	577,38	566,5	576,75

Table 6 Matrix A_3

n	$\sum x_1$	y	$\sum x_3$	$\sum x_4$
30	123,4	128,6	128,5	131
123,4	513,96	534,12	534,25	543,25
131,5	545	567,80	567,81	577,38
128,5	534,25	556,64	558	566,5
131	543,25	565,97	566,5	576,75

Table 7 Matrix A_4

n	$\sum x_1$	$\sum x_2$	y	$\sum x_4$
30	123,4	131,5	128,6	131
123,4	513,96	545	534,12	543,25
131,5	545	581	567,80	577,38
128,5	534,25	567,81	556,64	566,5
131	543,25	577,38	565,97	576,75

Table 8 Matrix A_5

n	$\sum x_1$	$\sum x_2$	$\sum x_3$	y
30	123,4	131,5	128,5	128,6
123,4	513,96	545	534,25	534,12
131,5	545	581	567,81	567,80
128,5	534,25	567,81	558	556,64
131	543,25	577,38	566,5	565,97

3. Find the determinant value of each matrix using the Sarrus method.

   The following is the result of each existing determinant value:

   Det(A) = 640,4597852

   $Det(A_1)$ = 4,05447E-10

   $Det(A_2)$ = 160,1149463

   $Det(A_3)$ = 160,1149463

   $Det(A_4)$ = 160,1149463

   $Det(A_5)$ = 160,1149463

4. Look for the values of b_1, b_2, b_3, b_4, b_5, and b_6 which will later be used for the values of the multiple linear regression equation. This could be done by using the formula in (3) and the following results were generated.

   b_1 =4,05447E-10/ 640,4597852 = 6,33056E-13

   b_2 = 160,1149463 / 640,4597852 = 0,25

   b_3 = 160,1149463 / 640,4597852 = 0,25

   b_4 = 160,1149463 / 640,4597852 = 0,25

   b_5 = 160,1149463 / 640,4597852 = 0,25

   Then, the multiple linear regression equation for the 1st data used the formula (4) was as follows:

   $Y = b_1 + b_2x_1 + b3x_2 + b4x_3 + b5x_4$

   Y = 6,33056E-13 + 1 + 1,125 + 1,0625 + 1,0625

   Y = 4,25

The level of community welfare was 5 based on the value range of Y that was shown in Table 1. This result implied that the first respondent belonged to the Prosperous Community III Plus (KS-III Plus). His family can fulfill all 6 indicators of the KS I stage, 8 indicators of the KS II stage, 5 indicators of the KS III stage, and 2 indicators of the KS III Plus stage.

Design System

This section would demonstrate the database design and algorithm design that was required to build the prediction application.

Database Design

Figure 8 displayed the database design for the application to predict the level of community welfare in the North Wangurer sub-district. There were three

tables created, namely Responden, Tingkat_Kesejahteraan, and Perhitungan_MLR.

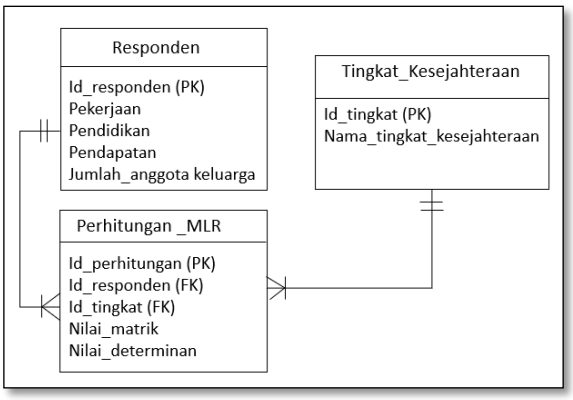

Figure 8 The Database Design

Algorithm Design

Figure 9 showed how the multiple linear regression would be applied in the prediction application. The detailed steps of the implementation of this method could be read in Steps in the Multiple Linear Regression section.

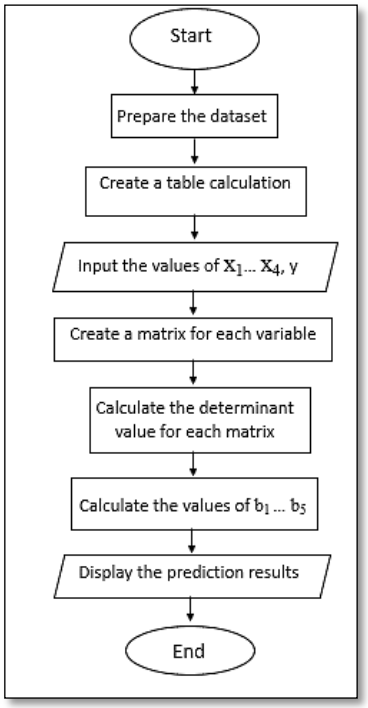

Figure 9 The Multiple Linear Regression Method

RESULTS AND DISCUSSION

Implementation

The implementation environment could be seen in Table 9.

Table 9 The List of Hardware and Software

Name	Description
a. Hardware	
1. *Processor*	Intel(R) Core(TM) i3-5005U CPU @ 2.00Hz 2.00GHz
2. *Memory* (RAM)	4.00 GB
3. *Mouse*	Standard
b. Software	
1. *Visual Studio Code*	A script editor to code
2. *Draw.io*	Drawing tool
3. Mockplus	Design user interface
4. *PostgreSQL*	*Database management system* (DBMS)
5. *Microsoft Word*	Text processing
6. SPSS 28	Processed statistical data
7. *Google Chrome* version 110.0.5481.77	Web browser
8. Windows 10	Operating system

User Interface

There were several user interfaces created for this application. Figure 10 demonstrated the front page to upload the CSV file that contained the dataset gathered from the questionnaires distributed to 238 respondents who live in the North Wangurer sub-district.

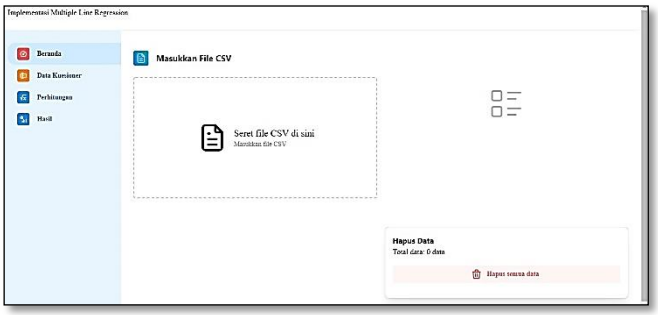

Figure 10 Upload CSV File

The columns from the table in Figure 11 consisted of income, education, occupation, and number of family members which were used as the independent variables while the y value was for the dependent variable.

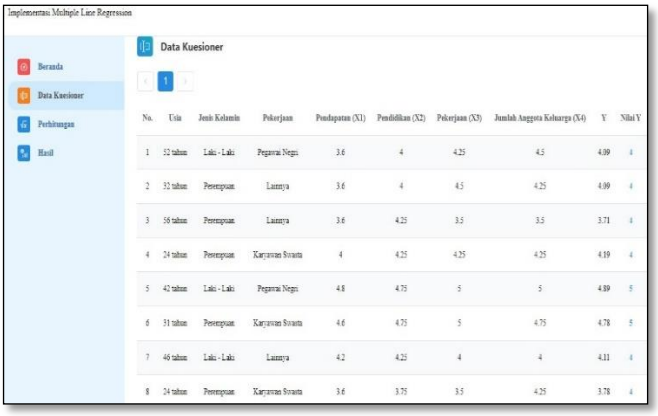

Figure 11 The Dataset from Questionnaires

Figures 12-15 depicted the steps of applying the multiple linear regression the application.

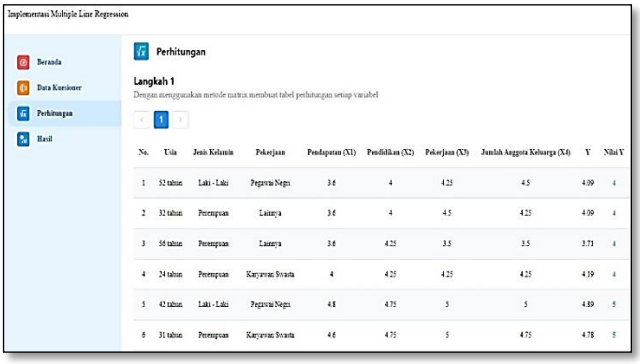

Figure 12 The 1st Step

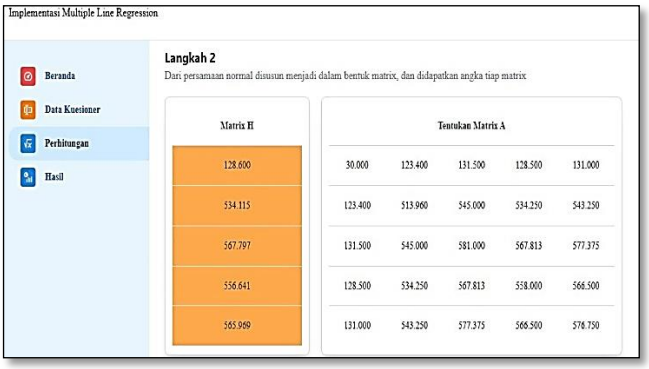

Figure 13 The 2nd Step

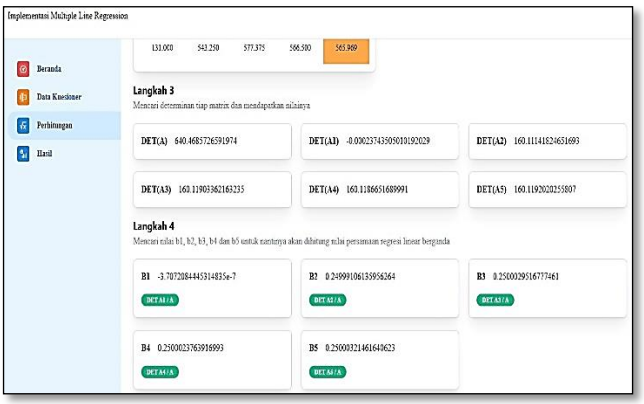

Figure 14 The 3rd and 4th Steps

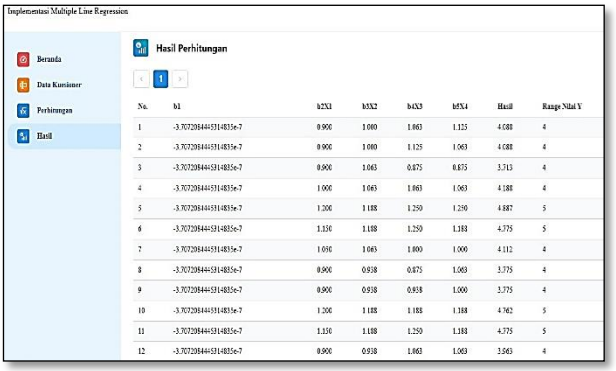

Figure 15 The Last Step

Testing

The test results revealed that the level of community welfare for the North Wangurer sub-district was in level five (5), KS III Plus, with a percentage of 98,7%. This meant that nearly all the families can fulfill all 6 indicators of the KS I stage, 8 indicators of KS II, 5 indicators of KS III, and 2 indicators of the KS III Plus stage.

Table 10 enlisted the number of respondents that belonged to the level of community welfare (y) using the multiple linear regression method. The results revealed that most respondents were in level 5, Prosperous Family KS III Plus (98,7%).

Table 10 Final Result

Value Range of Y (Level)	Number of Respondents
1	0
2	0
3	0
4	3
5	235

The families in this sub-district can meet all basic needs, social psychology and its development, and self-accountability (self-esteem). Two indicators of Prosperous Family III Plus (KS III Plus) or indicators of "self-actualization" (self-esteem) out of 21 family indicators, namely (BKKBN, 2011, section Indikator Tahapan Keluarga Sejahtera):

a. Families regularly voluntarily contribute materially to social activities.
   This means that a family has a great social sense by making material donations regularly (a certain time) and voluntarily, both in the form of money or goods, for the benefit of society (such as for orphans), houses of

worship, educational foundations, nursing homes, to finance activities at the RT/RW/Dusun, village level, and others) in this case, are not included as mandatory contributions.

b. There are family members who are active as administrators of social associations/foundations/communities. institutions.
It implied that a family has a great social sense by providing continuous energy, and mental and moral assistance for social interests by becoming administrators of various organizations/committees (such as administrators at foundations, traditional organizations, arts, sports, religion, youth, community institutions, RT/RW administrators, LKMD/LMD and others).

Due to the high level of community welfare achieved by the North Wangurer sub-district, the sub-district office must always encourage and support initiatives, movements, and participation of residents in their sub-district to develop their potential and assets for the common welfare. Furthermore, this office is urged to continuously seek community welfare development through various cross-sectoral, cross-district, and provincial cooperation. Equally important, this office must strive its best to sustain community welfare by conducting community development and empowerment fairly and equitably. Nevertheless, providing equitably distributing social assistance for households with a lower level of welfare is considered important and necessary.

These research findings can provide insights for the North Wangurer sub-district office to make more effective policies/decisions to sustain the current level of welfare as well as to promote public services that are professional, efficient and effective, open, and accountable for their residents to accelerate the realization of the community welfare,

CONCLUSIONS

Based on the F-test and T-test results, it was found that the independent variables used in this research, namely income, education, occupation, and the number of family members, did not only have influences on the dependent variable (level of community welfare) but also have a significant linear relationship between the independent variables and the dependent variable. Therefore, the hypotheses of H1, H2, H3, and H4 were accepted and H_0 was rejected.

There were other tests performed by inserting the data collected from 238 respondents in the prediction application and later being processed using the multiple linear regression method. This method could calculate and generate the results that showed the level of welfare for each respondent as well as the level of community welfare altogether. In this regard, it seemed that the North Wangurer sub-district managed to reach the 5^{th} level of community welfare as families in this sub-district can cater to all the basic needs, social psychology and its development, and self-accountability (self-esteem). This means that the residents in the North Wangurer sub-district require less social assistance from the respective sub-district office. However, this office must maintain the level of community welfare by conducting community development and empowerment fairly and equitably. Also, this office must have the capabilities to solve its own social and economic problems by optimizing the use of existing facilities and infrastructure available and strengthening the sub-district community as the subject of development.

For future works, adding more independent variables from external factors, namely human factors, natural factors, and state economic factors, is highly recommended. From the human factors, it might include physical threats and violation of norms. At the same time, the natural factors are viral diseases, natural hazards, and riots. The state economic factors include income per resident and inflation.

ACKNOWLEDGMENT

We would like to convey our gratitude toward North Wangurer officers and residents who helped with this research. Also, we were grateful for the support given by Universitas Katolik De La Salle Manado.

REFERENCE

Afkarina, N. K., Widodo, A. W., & Furqon, M. T. (2019). Implementasi Regresi Linear Berganda Untuk Prediksi Jumlah Peminat Mata Kuliah . *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, 3(11)*, 10462-10467.

Alfarizi, A. D., & Andri. (2021). Pemanfaatan Data Mining Dalam Memprediksi Produksi Pada PT Pupuk

Sriwidjaja Palembang Menggunakan Metode Multiple Linear Regression. *JNIK, 2(1)*, 51-63.

Amanda, L., Yanuar, F., & Devianto, D. (2019). Uji Validitas Dan Realibitas Tingkat Partisipasi Politik Masyarakat Kota Padang. *Jurnal Matematika UNPAD, 8(1)*, 179 – 188.

Astuti, Adyatma, S., & Normelani, E. (2017). Pemetaan Tingkat Kesejahteraan Keluarga Di Kecamatan Banjarmasin Selatan. *Jurnal Pendidikan Geografi, 4(2)*, 20-34.

BKKBN. (2011). *Batasan dan Pengertian MDK*. Retrieved from BKKBN Pemutakhiran Data Keluarga: http://aplikasi.bkkbn.go.id/mdk/BatasanMDK.aspx

Budiastuti, D., & Bandur, A. (2018). *Validitas Dan Reliabilitas Penelitian*. Jakarta: Mitra Wacana Media.

Cugung, P. D. (2020, Mei). *Klasifikasi Keluarga Menurut 5 Tahapan Keluarga*. Retrieved from Sistem Informasi Desa Cugung: https://cugung-rajabasa.desa.id/artikel/2020/5/8/klasifikasi-keluarga-menurut-5-tahapan-keluarga

Hayuningtyas, R. Y., & Sari, R. (2022). Implementasi Data Mining Dengan Metode Multiple Linear Regression Untuk Memprediksi Penyakit Diabetes. *Jurnal Teknik Komputer AMIK BSI, 8(1)*, 40-44.

Herwanto, H. W., Widyaningtyas, T., & Indriana, P. (2019). Penerapan Algoritma Linear Regression untuk Prediksi Hasil Panen Tanaman Padi. *JNTETI, 8(4)*, 364-370.

Iskandar, Hartoyo, Sumarwan, U., & Khomsan, A. (2010). Faktor – Faktor Yang Mempengaruhi Kesejahteraan Keluarga. *Jurnal Departemen Gizi Masyarakat Fakultas, 5(3)*, 63-72.

Izzah, A., & Widyastuti, R. (2017). Prediksi Harga Saham Menggunakan Improved Multiple Linear Regression Untuk Pencegahan Data Outlier. *KINETIK, 2(3),*141-150.

Kaluku, M. R., & Pakaya, N. (2020). Penerapan Metode AHP-Topsis Untuk Mengukur Tingkat Kesejahteraan Masyarakat Pesisir. *ILKOM Jurnal Ilmiah, 12*(3), 191-199.

Kamal, I. L., Hendro, T., & Ilyas, R. (2017). Prediksi Penjualan Buku Menggunakan Data Mining. *Seminar Nasional Teknologi Informasi dan Multimedia 2017* (pp. 49-54). Yogyakarta: STMIK AMIKOM Yogyakarta.

Marbun, M., Sihotang, H. T., & Nababan, M. A. (2018). Perancangan Sistem Peramalan Jumlah Wisatawan Asing. *Manajemen dan Informatika Komputer, 2(1)*, 41-49.

Masawah, P. D. (2018, Mei 18). *Tingkat Kesejahteraan Keluarga dan Indikator Tahapan Keluarga Sejahtera*. Retrieved from Masawah Desa: https://masawah.desa.id/2018/05/30/tingkat-kesejahteraan-keluarga-dan-indikator-tahapan-keluarga-sejahtera/

Nofriani. (2020). Machine Learning Application for Classification Prediction of Household's Welfare Status. *Journal of Information Technology and Computer Engineering, 4(2)*, 72-82.

Sulardi, P., Hendro, T., & Umbara, F. R. (2017). Prediksi Kebutuhan Obat Menggunakan Regresi Linier. *Seminar Nasional Teknologi dan Informatika (SNATIF)* (pp. 57-62). Kudus, Jawa Tengah: Fakultas Teknik, Universitas Muria Kudus.

Zahra, R., & Rina, N. (2018). Pengaruh Celebrity Endorser Hamidah Rachmayanti Terhadap Keputusan Pembelian Produk Online Shop Mayoutfit Di Kota Bandung. *Lontar: Jurnal Ilmu Komunikasi, 6*(1), 43-57.

Halaman ini sengaja dikosongkan