International Journal of Advances in Data and Information Systems
Vol. 4, No. 1, April 2023, pp. 51~61
ISSN: 2721-3056, DOI: 10.25008/ijadis.v4i1.1285

Prediction of Planning Value School Shopping Income Budget with Multiple Linear Regression

Cahyani Hana Bestari[1], Faisal Fajri Rahani[2]
[1,2]Department of Informatics, Universitas Ahmad Dahlan, Indonesia

Article Info

Article history:

Received Mar 16, 2023
Revised Apr 18, 2023
Accepted Apr 29, 2023

Keywords:

RAPBS
Prediction
Data Mining
Multiple Linear Regression

ABSTRACT

The School Expenditure Budget Plan or RAPBS is the pillar of school management for allocating the revenue budget and use of school funds to meet all school needs for one year. However, there are problems that occur in the management of the RAPBS, namely the difficulty of grouping the RAPBS data annually, making it difficult to predict the budget for the coming year. This research was conducted to study and implement the Multiple Linear Regression algorithm in predicting the value of data on income and expenditure budget plans which are a reference in planning future budgets. To support predictions of planned school budgets and income, BUMS data, Aid data, School Program Cost data, Original School Revenue data, Other Sources data, and Total Budget data are used. The prediction system method used consists of the planning stage, the analysis stage, the modeling stage, interface design, and implementation using the PHP and MySQL programming languages for database management and system testing and analysis. The results of testing the data analysis using the multiple linear regression method with SPSS software have a 100% result according to the manual calculations performed.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Faisal Fajri Rahani,
Department of Informatics,
Universitas Ahmad Dahlan, Indonesia,
Email: faisal.fajri@tif.uad.ac.id

1. INTRODUCTION

Education cannot be carried out alone without funding, and financing is needed to optimize the use of all aspects and resources in the teaching and learning process to achieve educational goals. Proper management of educational finances is the key to supporting the success of the educational goals themselves. There are three main problems in managing education finances, including financing regarding where to get funding sources, budgeting how to channel education funds, and accountability in the way the budget is used and accountability obtained. Permendikbud Number 06 of 2021 discusses "Technical Instructions for Managing Regular School Operational Assistance Funds" [1]. The School Revenue Budget Plan or RAPBS is the pillar of school management for allocating the revenue budget and using school funds to meet all school needs for one year.

The RAPBS includes the budgeting of funds for teaching and learning activities, school development, school renovation, and school facilities needed. Based on these needs, the RAPBS needs to be prepared in such a way that the budget allocation can optimally meet school needs. The function of the RAPBS is to plan the receipt and disbursement of school/madrasah funds so that they can be properly controlled. The contents of the RAPBS itself include the revenue budget and expenditure budget [2]. This budget data becomes the development of the school. If the RAPBS data has increased, it will make the school have good development. But on the contrary, if the RAPBS data on expenditure data is greater than the income data, it can be ascertained that the development of the school is not good. Therefore, to find out the development of a good or bad school, a budget prediction is needed which will be the initial picture for the upcoming RAPBS [3].

There are several researches on predicting income and expenditure budget plans, such as predicting the next year's APBD data using the K-Means Algorithm method. The results of the analysis can predict the following year's data. Researched that for APBD the K-Means clustering method produces an analysis that can produce a good classification but by using 3 centroids [4] [5].

From these several studies it can be concluded that to predict the RAPBS data, clustering must be used first in order to facilitate the management of RAPBS data [6]. Therefore, to strengthen the prediction of future budget values, use the forecasting method to predict future RAPBS data based on time series [7] [8]. The forecasting method has its own advantages in forecasting the future, namely taking into account the latest developments in the environment and internal information, and consistently and objectively considering the amount of information or data at once [9]. Where there are advantages, there are also disadvantages, namely biased forecasts and questionable accuracy, and quantitative data cannot be obtained only as well as the data entered [10]. This forecasting method is the right choice in this study because it predicts the value of future income and expenditure budget plans whose predictions are in the form of the value of the budget [11]. To support the Multiple Linear Regression Algorithm in predicting the budget [12]. This multiple linear regression method has the advantage of being a fairly simple method, but still producing maximum results and being able to identify how strong the influence is exerted by the independent variable on the dependent variable and being able to predict data in the future [13] [14]. On the other hand, there is a weakness of this method, namely the prediction results which are estimated values, so that the possibility does not match the actual data [15]. The multiple linear regression method is used here to support the estimated analysis of the predicted budget value under study [16] [17].

Therefore, in this study the authors will try to apply a multiple linear regression algorithm to predict the value of future budget plans. The results of the analysis can predict the next year's data. To predict the planned value of the income and expenditure budget in the future, the prediction is in the form of a budget value. To support the Multiple Linear Regression Algorithm in predicting the budget.

2. RESEARCH METHOD

Knowledge-based management planning is a data analysis process that is processed to become a support in all other subsystems acting as a main component in the system to be processed. The design of knowledge-based management is the calculation of multiple linear regression algorithms in predicting the future RAPBS. In calculating the multiple linear regression algorithm there are the following steps:

2.1. Determining Multiple Linear Regression Variables

This stage is the initial stage in the calculation of the multiple linear regression method to determine the dependent variable and independent variable [18]. The dependent variable, namely the dependent variable, is denoted by Y, while the independent variable is denoted by X [19]. At this stage, five independent variables are used and one dependent variable is used. In this study, the independent variables were grouped, namely BUMS, Aid, School Program Fees, School Original Income and Other Sources. As for the dependent variable, namely the total budget.

2.2. Forming a Multiple Linear Regression Equation Model

At this stage a new multiple linear regression equation model will be formed using the independent variables that have been determined in the previous stage [20] [21]. The form of the new multiple linear regression equation is as follows:

$$Y = a + b1X1 + b2X2 + b3X3 + b4X4 + b5X5 (1)$$

Information:

Y = Dependent Variable (predicted value)
a = Constant
b(1-5) = Regression coefficient
X(1-5) = Independent Variable

This study uses five independent variables that form a multiple linear regression equation formula above in formula (1). To form the formula, the following calculation steps are needed:

1. Finding Constants (a) and Coefficients (b)

In the multiple linear regression equation model above, the constants and regression coefficients are unknown. the constants in the multiple linear regression equation model are denoted by a, while the coefficients are denoted by b [22]. To find the constants and coefficients of the multiple linear regression equation model, you need the following matrix formula:

$$a = \begin{bmatrix} N & \sum X1 & \sum X2 & \sum X3 & \sum X4 & \sum X5 \\ \sum X1 & \sum(X1,X1) & \sum(X1,X2) & \sum(X1,X3) & \sum(X1,X4) & \sum(X1,X5) \\ \sum X2 & \sum(X2,X1) & \sum(X2,X2) & \sum(X2,X3) & \sum(X2,X4) & \sum(X2,X5) \\ \sum X3 & \sum(X3,X1) & \sum(X3,X2) & \sum(X3,X3) & \sum(X3,X4) & \sum(X3,X5) \\ \sum X4 & \sum(X4,X1) & \sum(X4,X2) & \sum(X4,X3) & \sum(X4,X4) & \sum(X4,X5) \\ \sum X5 & \sum(X5,X1) & \sum(X5,X2) & \sum(X5,X3) & \sum(X5,X4) & \sum(X5,X5) \end{bmatrix} \quad (2)\ [18]$$

$$b = \begin{bmatrix} a \\ b1 \\ b2 \\ b3 \\ b4 \\ b5 \end{bmatrix} \quad (3)$$

$$H = \begin{bmatrix} \sum(Y) \\ \sum(X1.Y) \\ \sum(X2.Y) \\ \sum(X3.Y) \\ \sum(X4.Y) \\ \sum(X5.Y) \end{bmatrix} \quad (4)$$

Information :
A = Matrix A
b = column vector (which is not yet known, namely the constants and coefficients)
H = column vector (which is already known)

From the basic matrix formula above, it will be combined into formula equation 5:

$$\begin{bmatrix} N & \sum X1 & \sum X2 & \sum X3 & \sum X4 & \sum X5 \\ \sum X1 & \sum(X1,X1) & \sum(X1,X2) & \sum(X1,X3) & \sum(X1,X4) & \sum(X1,X5) \\ \sum X2 & \sum(X2,X1) & \sum(X2,X2) & \sum(X2,X3) & \sum(X2,X4) & \sum(X2,X5) \\ \sum X3 & \sum(X3,X1) & \sum(X3,X2) & \sum(X3,X3) & \sum(X3,X4) & \sum(X3,X5) \\ \sum X4 & \sum(X4,X1) & \sum(X4,X2) & \sum(X4,X3) & \sum(X4,X4) & \sum(X4,X5) \\ \sum X5 & \sum(X5,X1) & \sum(X5,X2) & \sum(X5,X3) & \sum(X5,X4) & \sum(X5,X5) \end{bmatrix} \begin{bmatrix} a \\ b1 \\ b2 \\ b3 \\ b4 \\ b5 \end{bmatrix} = \begin{bmatrix} \sum(Y) \\ \sum(X1.Y) \\ \sum(X2.Y) \\ \sum(X3.Y) \\ \sum(X4.Y) \\ \sum(X5.Y) \end{bmatrix} \quad (5)$$

Information :

N = Total Regression Data
ΣX(1,2,3,4,5) = Total number of independent variables X(1,2,3,4,5) (adjust the name of the variable type)

On the basis of formula (5) is used to form the formulation of matrices A0, A1, A2, A3, A4, A5. The matrices A0, A1, A2, A3, A4, A5 are the matrices used to determine the constants (a) and regression coefficients (b) contained in the equation of the multiple linear regression formula. Here's the matrix A0, A1, A2, A3, A4, A5:

$$A0 = \begin{bmatrix} \sum Y & \sum X1 & \sum X2 & \sum X3 & \sum X4 & \sum X5 \\ \sum(X1.Y) & \sum(X1,X1) & \sum(X1,X2) & \sum(X1,X3) & \sum(X1,X4) & \sum(X1,X5) \\ \sum(X2.Y) & \sum(X2,X1) & \sum(X2,X2) & \sum(X2,X3) & \sum(X2,X4) & \sum(X2,X5) \\ \sum(X3.Y) & \sum(X3,X1) & \sum(X3,X2) & \sum(X3,X3) & \sum(X3,X4) & \sum(X3,X5) \\ \sum(X4.Y) & \sum(X4,X1) & \sum(X4,X2) & \sum(X4,X3) & \sum(X4,X4) & \sum(X4,X5) \\ \sum(X5.Y) & \sum(X5,X1) & \sum(X5,X2) & \sum(X5,X3) & \sum(X5,X4) & \sum(X5,X5) \end{bmatrix} \quad (6)$$

$$A1 = \begin{bmatrix} N & \sum Y & \sum X2 & \sum X3 & \sum X4 & \sum X5 \\ \sum X1 & \sum(X1,Y) & \sum(X1,X2) & \sum(X1,X3) & \sum(X1,X4) & \sum(X1,X5) \\ \sum X2 & \sum(X2,Y) & \sum(X2,X2) & \sum(X2,X3) & \sum(X2,X4) & \sum(X2,X5) \\ \sum X3 & \sum(X3,Y) & \sum(X3,X2) & \sum(X3,X3) & \sum(X3,X4) & \sum(X3,X5) \\ \sum X4 & \sum(X4,Y) & \sum(X4,X2) & \sum(X4,X3) & \sum(X4,X4) & \sum(X4,X5) \\ \sum X5 & \sum(X5,Y) & \sum(X5,X2) & \sum(X5,X3) & \sum(X5,X4) & \sum(X5,X5) \end{bmatrix} \quad (7)$$

$$A2 = \begin{bmatrix} N & \sum X1 & \sum Y & \sum X3 & \sum X4 & \sum X5 \\ \sum X1 & \sum(X1,X1) & \sum(X1,Y) & \sum(X1,X3) & \sum(X1,X4) & \sum(X1,X5) \\ \sum X2 & \sum(X2,X1) & \sum(X2,Y) & \sum(X2,X3) & \sum(X2,X4) & \sum(X2,X5) \\ \sum X3 & \sum(X3,X1) & \sum(X3,Y) & \sum(X3,X3) & \sum(X3,X4) & \sum(X3,X5) \\ \sum X4 & \sum(X4,X1) & \sum(X4,Y) & \sum(X4,X3) & \sum(X4,X4) & \sum(X4,X5) \\ \sum X5 & \sum(X5,X1) & \sum(X5,Y) & \sum(X5,X3) & \sum(X5,X4) & \sum(X5,X5) \end{bmatrix} \quad (8)$$

$$A3 = \begin{bmatrix} N & \sum X1 & \sum X2 & \sum Y & \sum X4 & \sum X5 \\ \sum X1 & \sum(X1,X1) & \sum(X1,X2) & \sum(X1,Y) & \sum(X1,X4) & \sum(X1,X5) \\ \sum X2 & \sum(X2,X1) & \sum(X2,X2) & \sum(X2,Y) & \sum(X2,X4) & \sum(X2,X5) \\ \sum X3 & \sum(X3,X1) & \sum(X3,X2) & \sum(X3,Y) & \sum(X3,X4) & \sum(X3,X5) \\ \sum X4 & \sum(X4,X1) & \sum(X4,X2) & \sum(X4,Y) & \sum(X4,X4) & \sum(X4,X5) \\ \sum X5 & \sum(X5,X1) & \sum(X5,X2) & \sum(X5,Y) & \sum(X5,X4) & \sum(X5,X5) \end{bmatrix} \quad (9)$$

$$A4 = \begin{bmatrix} N & \sum X1 & \sum X2 & \sum X3 & \sum Y & \sum X5 \\ \sum X1 & \sum(X1,X1) & \sum(X1,X2) & \sum(X1,X3) & \sum(X1,Y) & \sum(X1,X5) \\ \sum X2 & \sum(X2,X1) & \sum(X2,X2) & \sum(X2,X3) & \sum(X2,Y) & \sum(X2,X5) \\ \sum X3 & \sum(X3,X1) & \sum(X3,X2) & \sum(X3,X3) & \sum(X3,Y) & \sum(X3,X5) \\ \sum X4 & \sum(X4,X1) & \sum(X4,X2) & \sum(X4,X3) & \sum(X4,Y) & \sum(X4,X5) \\ \sum X5 & \sum(X5,X1) & \sum(X5,X2) & \sum(X5,X3) & \sum(X5,Y) & \sum(X5,X5) \end{bmatrix} \quad (10)$$

$$A5 = \begin{bmatrix} N & \sum X1 & \sum X2 & \sum X3 & \sum X4 & \sum Y \\ \sum X1 & \sum(X1,X1) & \sum(X1,X2) & \sum(X1,X3) & \sum(X1,X4) & \sum(X1,Y) \\ \sum X2 & \sum(X2,X1) & \sum(X2,X2) & \sum(X2,X3) & \sum(X2,X4) & \sum(X2,Y) \\ \sum X3 & \sum(X3,X1) & \sum(X3,X2) & \sum(X3,X3) & \sum(X3,X4) & \sum(X3,Y) \\ \sum X4 & \sum(X4,X1) & \sum(X4,X2) & \sum(X4,X3) & \sum(X4,X4) & \sum(X4,Y) \\ \sum X5 & \sum(X5,X1) & \sum(X5,X2) & \sum(X5,X3) & \sum(X5,X4) & \sum(X5,Y) \end{bmatrix} \quad (11)$$

Information :

A0 = is an AO matrix to find constants (a)
A(1-5) = is the A matrix to find the coefficient (b)

After the matrix is determined, the next calculation is to determine the constants and regression coefficients, namely calculating the matrix determination A, A0, A1, A2, A3, A4 and A5. Calculation of the determinant of the matrix is used in the application of multiple linear regression because the resulting matrix has an order of 3x3 or more. The 3x3 order can be calculated using the cofactor method. In the case of the research data, the application of the matrix determinant formula using the cofactor method results in the calculation of Det(A), Det(A0), Det(A1), Det(A2), Det(A3), Det(A4) and Det(A5). The results of calculating the determinants of the matrix become the final result for determining the constants and coefficients of multiple linear regression. The formula for calculating constants (a) and coefficients (b) is as follows:

$$a = \frac{Det(A0)}{Det\ (A)} \quad (12)$$

$$b1 = \frac{Det(A1)}{Det\ (A)} \quad (13)$$

$$b2 = \frac{Det(A2)}{Det\ (A)} \quad (14)$$

$$b3 = \frac{Det(A3)}{Det\ (A)} \quad (15)$$

$$b4 = \frac{Det(A4)}{Det\ (A)} \quad (16)$$

$$b5 = \frac{Det(A5)}{Det\ (A)} \quad (17)$$

The formula above will produce a new multiple linear regression equation model which can be used as a reference in the research being conducted.

3. RESULTS AND DISCUSSION

The research data was collected using RAPBS sample data from SMP Muh 3 Yogyakarta with a period of 4 years back (2017-2021) obtained through an internship at PDM Yogyakarta city. The results of data collection are described in tables which are classified into several groups according to the type of data that you want to use for research, namely there are several attributes that are included, namely the date attribute, the BUMS attribute, the Assistance attribute, the School Program Cost attribute, the School Original Income attribute, the Other Sources attribute and attribute Total Budget. An example of RAPBS data for SMP Muh 3 Yogyakarta can be seen in Table 1.

Table 1 School RAPBS data

No	Date	BUMS	Donation	School Program Fees Sekolah	School Original Income	Other Source	Total Budget
1	Juli 2017	-	67.900.000	327.899.378	897.599.378	67.900.000	1.361.298.756
2	Agustus 2017	-	146.800.000	310.974.378	720.349.378	10.000.000	1.188.123.756
3	Sep-17	-	10.000.000	918.399.378	870.349.378	10.000.000	1.808.748.756
4	Oktober 2017	-	10.000.000	414.449.378	219.349.378	10.000.000	653.798.756
…	…	…	…	…	…	…	…
45	Maret 2021	-	263.140.000	713.772.000	156.000.000	850.000	1.133.762.000
46	Apr-21	-	22.900.000	418.282.000	156.000.000	850.000	598.032.000
47	Mei 2021	-	22.900.000	940.747.000	156.000.000	850.000	1.120.497.000
48	Juni 2021	-	828.270.000	653.259.500	156.000.000	850.000	1.638.379.500
	JUMLAH	28.700.000	5.538.996.508	24.200.333.036	18.455.442.536	1.391.717.408	49.615.189.488

Determining Multiple Linear Regression independent Variables

The first stage in the calculation of the multiple linear regression method is to determine the dependent variable and independent variable. The dependent variable, namely the dependent variable that binds the independent (independent) variable, is denoted by Y, while the independent variable is the independent variable, denoted by X. The variable data below is a grouping of data according to the variables, which can be seen in Table 2.

Table 2 Data Variables

Variabel	Data
X1	BUMS
X2	Donation
X3	School Program Fees
X4	Original School Income

X5	Other Source
Y	Total Budget Belanja

After grouping the variables used, the calculation process is carried out as follows by simplifying the complete data in table 1, simplified by dividing 10 million per available data. This simplification is carried out in order to simplify the process of calculating multiple linear regression.

Table 3 RAPBS Data Simplification

No	Date	BUMS (X1)	Donation (X2)	School Program Fees (X2)	Original School Income (X3)	Other Source (X5)	Total Budget (Y)
1	Juli 2017	-	6.79	32.79	89.76	6.79	136.13
2	Agustus 2017	-	14.68	31.09	72.03	1.00	118.81
3	Sep-17	-	1.00	91.84	87.03	1.00	180.87
4	Oktober 2017	-	1.00	41.44	21.93	1.00	65.38
…	…	…	…	…	…	…	…
45	Maret 2021	-	26.314	71.38	15.60	0.09	113.38
46	Apr-21	-	2.29	41.83	15.60	0.09	59.80
47	Mei 2021	-	2.29	94.07	15.60	0.09	112.05
48	Juni 2021	-	82.827	65.33	15.60	0.09	163.84
	JUMLAH	2.87	553.90	2.420.03	1.845.54	139.17	4.961.52

Forming a New Multiple Linear Regression Equation Formula Model

This stage will form a new multiple linear regression equation according to the equation formula (1). The following are the steps for applying the multiple linear regression calculation formula to form a new multiple linear regression equation which becomes a benchmark in the application of the system to be created.

In this study using more than two independent variables, so the value of the constants and regression variables for each independent variable can be obtained using the determinant matrix. The following research data obtained has 5 equations with 5 independent variables whose values are unknown, therefore constants (a) and coefficients (b1, b2, b3, b4, and b5) will be processed to find out their values using these equations. The following is the result of calculating the application of formula (5) which has been processed in the matrix formulas (6), (7), (8), (9), (10), and (11) which will produce matrices A, A0, A1, A2, A3, A4, and A5 for calculating the determinant of the matrix. The steps for the calculation are as follows:

The initial step in the formula part (2) must determine the calculation of X12, X1(X2- X5), X22, X2(X3-X4). The following are the results and total formula calculations that were obtained from July 2017 to June 2021. Can be seen in Table 4.

Table 4 Matrix (2) formula calculation

Date	X1X1	X1X2	X1X3	X1X4	X1X5	X2X2	X2X3	X2X4
Juli 2017	-	-	-	-	-	46,10	222,64	609,47
Agustus 2017	-	-	-	-	-	215,50	456,51	1.057,47
Sep-17	-	-	-	-	-	1,00	91,84	87,03
Oktober 2017	-	-	-	-	-	1,00	41,44	21,93
…	…	…	…	…	…	…	…	
Maret 2021	-	-	-	-	-	692,43	1.878,22	410,50
Apr-21	-	-	-	-	-	5,24	95,79	35,72
Mei 2021	-	-	-	-	-	5,24	215,43	35,72
Juni 2021	-	-	-	-	-	6.860,31	5.410,75	1.292,10
Jumlah	4,38	2,45	141,69	82,38	0,67	21.909,88	33.314,01	32.129,76

The initial step in part 2 of the formula should determine the calculations X2X5), X32, X3(X4-X5), X42, X4X5, and X52. The following are the results and total formula calculations that were obtained from July 2017 to June 2021. Can be seen in Table 5.

Table 5 Matrix (3) formula calculation

Date	X2X5	X3X3	X3X4	X3X5	X4X4	X4X5	X5X5	X2X5
Juli 2017	46,10	1.075,18	2.943,22	222,64	8.056,85	609,47	46,10	46,10

Agustus 2017	14,68	967,05	2.240,10	31,10	5.189,03	72,03	1,00	14,68
Sep-17	1,00	8.434,57	7.993,28	91,84	7.575,08	87,03	1,00	1,00
Oktober 2017	1,00	1.717,68	909,09	41,44	481,14	21,93	1,00	1,00
…	…	…	…	…	…	…	…	…
Maret 2021	2,24	5.094,70	1.113,48	6,07	243,36	1,33	0,01	2,24
Apr-21	0,19	1.749,60	652,52	3,56	243,36	1,33	0,01	0,19
Mei 2021	0,19	8.850,05	1.467,57	8,00	243,36	1,33	0,01	0,19
Juni 2021	7,04	4.267,48	1.019,08	5,55	243,36	1,33	0,01	7,04
Jumlah	3.587,31	146.901,14	113.685,61	9.462,05	168.932,13	10.563,40	5.407,01	3.587,31

Next, the calculation of the H matrix contained in formula (4) is carried out, namely the calculation of $X1Y, X2Y, X3Y, X4Y, X5Y$. The following are the results and total formula calculations that were obtained from July 2017 to June 2021. Can be seen in Table 6.

Table 6 Matrix (4) formula calculation

No	Date	YX1	YX2	YX3	YX4	YX5
1	Juli 2017	-	924,32	4.463,69	12.219,01	924,32
2	Agustus 2017	-	1.744,17	3.694,76	8.558,64	118,81
3	Sep-17	-	180,87	16.611,54	15.742,43	180,87
4	Oktober 2017	-	65,38	2.709,66	1.434,10	65,38
…	…	…	…	…	…	…
45	Maret 2021	-	2.983,38	8.092,48	1.768,67	9,64
46	Apr-21	-	136,95	2.501,46	932,93	5,08
47	Mei 2021	-	256,59	10.541,04	1.747,98	9,52
48	Juni 2021	-	13.570,21	10.702,87	2.555,87	13,93
	JUMLAH	231,58	90.943,42	303.504,51	325.393,28	29.020,44

In Table 3 it will be applied to the calculation of the H value used in the calculation of the matrix which is a column vector. This H value formula will be processed into the matrix formula A0, A1, A2, A3, A4, A5. The H value table is the result of calculating formula (4) in Table 7.

Table 7 H Value

NO	H
ΣY	4.961,52
Σ(X1.Y)	231,58
Σ(X2.Y)	90.943,42
Σ(X3.Y)	303.504,51
Σ(X4.Y)	325.393,28
Σ(X5.Y)	29.020,44

Once the value of H is known, it can be entered into the calculation matrix A0, A1, A2, A3, A4. The following is the entire matrix obtained, namely there are matrices A, A0, A1, A2, A3, A4, A5 which are in Table 8 to Table 14.

Table 8 Matrix A

A	48	2,87	553,90	2.420,03	1.845,54	139,17
	2,87	4,38	2,45	141,69	82,38	0,67
	553,90	2,45	21.909,88	33.314,01	32.129,76	3.587,31
	2.420,03	141,69	33.314,01	146.901,14	113.685,61	9.462,05
	1.845,54	82,38	32.129,76	113.685,61	168.932,13	10.563,40
	139,17	0,67	3.587,31	9.462,05	10.563,40	5.407,01

Table 9 Matrix A0

A0	4.961,52	2,87	553,90	2.420,03	1.845,54	139,17
	231,58	4,38	2,45	141,69	82,38	0,67
	90.943,42	2,45	21.909,88	33.314,01	32.129,76	3.587,31
	303.504,51	141,69	33.314,01	146.901,14	113.685,61	9.462,05
	325.393,28	82,38	32.129,76	113.685,61	168.932,13	10.563,40
	29.020,44	0,67	3.587,31	9.462,05	10.563,40	5.407,01

Table 10 Matrix A1

A1	48	4.961,52	553,90	2.420,03	1.845,54	139,17
	2,87	231,58	2,45	141,69	82,38	0,67
	553,90	90.943,42	21.909,88	33.314,01	32.129,76	3.587,31
	2.420,03	303.504,51	33.314,01	146.901,14	113.685,61	9.462,05
	1.845,54	325.393,28	32.129,76	113.685,61	168.932,13	10.563,40
	139,17	29.020,44	3.587,31	9.462,05	10.563,40	5.407,01

Table 11 Matrix A2

A2	48	2,87	4.961,52	2.420,03	1.845,54	139,17
	2,87	4,38	231,58	141,69	82,38	0,67
	553,90	2,45	90.943,42	33.314,01	32.129,76	3.587,31
	2.420,03	141,69	303.504,51	146.901,14	113.685,61	9.462,05
	1.845,54	82,38	325.393,28	113.685,61	168.932,13	10.563,40
	139,17	0,67	29.020,44	9.462,05	10.563,40	5.407,01

Table 12 Matrix A3

A3	48	2,87	553,90	4.961,52	1.845,54	139,17
	2,87	4,38	2,45	231,58	82,38	0,67
	553,90	2,45	21.909,88	90.943,42	32.129,76	3.587,31
	2.420,03	141,69	33.314,01	303.504,51	113.685,61	9.462,05
	1.845,54	82,38	32.129,76	325.393,28	168.932,13	10.563,40
	139,17	0,67	3.587,31	29.020,44	10.563,40	5.407,01

Table 13 Matrix A4

A4	48	2,87	553,90	2.420,03	4.961,52	139,17
	2,87	4,38	2,45	141,69	231,58	0,67
	553,90	2,45	21.909,88	33.314,01	90.943,42	3.587,31
	2.420,03	141,69	33.314,01	146.901,14	303.504,51	9.462,05
	1.845,54	82,38	32.129,76	113.685,61	325.393,28	10.563,40
	139,17	0,67	3.587,31	9.462,05	29.020,44	5.407,01

Table 14 Matrix A5

A5	48	2,87	553,90	2.420,03	1.845,54	4.961,52
	2,87	4,38	2,45	141,69	82,38	231,58
	553,90	2,45	21.909,88	33.314,01	32.129,76	90.943,42
	2.420,03	141,69	33.314,01	146.901,14	113.685,61	303.504,51
	1.845,54	82,38	32.129,76	113.685,61	168.932,13	325.393,28
	139,17	0,67	3.587,31	9.462,05	10.563,40	29.020,44

The matrix tables above are the results of matrix calculations which produce six matrices, namely matrices A, A0, A1, A2, A3, A4, A5. From this matrix, it is possible to calculate the determinants of the matrix, namely Det(A), Det(A0), Det(A1), Det(A2), Det(A3), Det(A4) and Det(A5). The following are the results of the calculation of the matrix determinants obtained from calculations using the matrix determinant formula in excel which are listed in Table 15.

Table 15 Determinan Matrix

Det A	25.143.095.708.547.500.000
Det A0	2.502.245
Det A1	25.143.095.708.547.500.000
Det A2	25.143.095.708.547.500.000
Det A3	25.143.095.708.547.500.000
Det A4	25.143.095.708.547.500.000
Det A5	25.143.095.708.547.500.000

From the calculation of the matrix determinants that have been produced above, it can be used to determine the values of a, b1, b2, b3, b4, b5. The calculation formula is formula (12) to (17). Following are the results of the calculations that have been carried out:

$$a = \frac{2.502.245}{25.143.095.708.547.500.000} = 0 \quad (18)$$

$$b1 = \frac{25.143.095.708.547.500.000}{25.143.095.708.547.500.000} = 1 \quad (19)$$

$$b2 = \frac{25.143.095.708.547.500.000}{25.143.095.708.547.500.000} = 1 \quad (20)$$

$$b3 = \frac{25.143.095.708.547.500.000}{25.143.095.708.547.500.000} = 1 \quad (21)$$

$$b4 = \frac{25.143.095.708.547.500.000}{25.143.095.708.547.500.000} = 1 \quad (22)$$

$$b5 = \frac{25.143.095.708.547.500.000}{25.143.095.708.547.500.000} = 1 \quad (23)$$

From this model it is known that the constant a is 0, and the coefficients b1 (1), b2 (1), b3 (1), b4 (1). So as to produce a new multiple linear regression equation model as follows:

$$Y = 0 + 1(X1) + 1(X2) + 1(X3) + 1(X4) + 1\,(x5)$$

This model can later be used for prediction calculations that will be carried out in the future with $X1, X2, X3, X4$ and $X5$ being independent variables. An example of its application can be done as follows: an example for the July 2021 Period

X1(BUMS) = 0
X2(Donation) = 22.900.000
X3(School Program Fees) = 414.400.000
X4(Original School Income) = 156.000.000
X5(Other sources) = 850.000
its application:
Y= 0 + 1(x1) + 1(X2) + 1(X3) + 1(X4) + 1(X5)
Y= 0 + 1(0)+ 1(22.900.000) + 1(414.400.000) + 1(156.000.000) + 1(850.000)
Y= 0 + 0 + 22.900.000 + 414.400.000 + 156.000.000 + 850.000
Y= 594.150.000

The future prediction results in budget management are approximately Rp. 594,150,000 for the July 2021 period. These results can later be used to measure how much of the budget can be used for school activities.

Data testing

The data testing phase is the testing phase carried out to test whether the data is in accordance with the final results of the prediction calculation model which is calculated manually. At this testing stage using SPSS testing. The following is a display of the results of testing prediction calculation data in Table 16.

Table 16 SPSS Testing

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Constant	-0.004	0.002		-1.525	0.135
BUMN (X1)	1.001	0.003	0.004	312.252	0.000
Donation (X2)	1.000	0.000	0.254	17710.969	0.000
School Program Fees (X3)	1.000	0.000	0.325	21446.552	0.000
Original School Income (X4)	1.000	0.000	0.644	42390.692	0.000
Other sources (X5)	1.000	0.000	0.146	10307.712	0.000

System Testing

Checking the results of manual analysis with those in the system for matching between the manual and the system is appropriate, namely the coefficients for data from July 2017 – June 2021, namely in the table 17.

Table 17 Testing system

KETERANGAN	EXCEL	SISTEM
A	0	0
B1	1	0,99 = 1
B2	1	1
B3	1	1
B4	1	1
B5	1	1,001= 1

4. CONCLUSION

Based on the research that has been done, the following conclusions are obtained:

1. From the results of analysis calculations with the Multiple Linear Regression Algorithm on the previous RAPBS data, a new multiple linear equation model can be produced, namely
$$\boldsymbol{Y = 0 + 1(X1) + 1(X2) + 1(X3) + 1(X4) + 1\,(x5).}$$
2. The results of the new multiple linear equations generated can be used as a reference for future budget prediction systems.
3. The Multiple Linear Regression Algorithm has a good correlation as evidenced from the data analysis test through the SPSS test, which is in accordance with the analysis calculations performed.

REFERENCES

[1] D. Pembinaan *et al.*, "PERATURAN MENTERI PENDIDIKAN DAN KEBUDAYAAN REPUBLIK INDONESIA NOMOR 1 TAHUN 2018".

[2] I. Chugunov and V. Makohon, "BUDGETARY PROJECTION IN THE SYSTEM OF FINANCIAL AND ECONOMIC REGULATION OF SOCIAL PROCESSES," *Baltic Journal of Economic Studies*, vol. 6, no. 1, Art. no. 1, Mar. 2020, doi: 10.30525/2256-0742/2020-6-1-130-135.

[3] X. Chen, W. Yu, and Y. Zhao, "Financial Risk Management and Control Based on Marine Economic Forecast," *Journal of Coastal Research*, vol. 112, no. SI, pp. 234–236, Nov. 2020, doi: 10.2112/JCR-SI112-065.1.

[4] A. Fahim, "K and starting means for k-means algorithm," *Journal of Computational Science*, vol. 55, p. 101445, Oct. 2021, doi: 10.1016/j.jocs.2021.101445.

[5] E. Zhang, H. Li, Y. Huang, S. Hong, L. Zhao, and C. Ji, "Practical multi-party private collaborative k-means clustering," *Neurocomputing*, vol. 467, pp. 256–265, Jan. 2022, doi: 10.1016/j.neucom.2021.09.050.

[6] R. A. Poldrack, G. Huckins, and G. Varoquaux, "Establishment of Best Practices for Evidence for Prediction: A Review," *JAMA Psychiatry*, vol. 77, no. 5, pp. 534–540, May 2020, doi: 10.1001/jamapsychiatry.2019.3671.

[7] W. Lu, J. Li, J. Wang, and L. Qin, "A CNN-BiLSTM-AM method for stock price prediction," *Neural Comput & Applic*, vol. 33, no. 10, pp. 4741–4753, May 2021, doi: 10.1007/s00521-020-05532-z.

[8] M. Rashid, B. S. Bari, Y. Yusup, M. A. Kamaruddin, and N. Khan, "A Comprehensive Review of Crop Yield Prediction Using Machine Learning Approaches With Special Emphasis on Palm Oil Yield Prediction," *IEEE Access*, vol. 9, pp. 63406–63439, 2021, doi: 10.1109/ACCESS.2021.3075159.

[9] N. Altay and A. Narayanan, "Forecasting in humanitarian operations: Literature review and research needs," *International Journal of Forecasting*, vol. 38, no. 3, pp. 1234–1244, Jul. 2022, doi: 10.1016/j.ijforecast.2020.08.001.

[10] L. Zhang *et al.*, "A review of machine learning in building load prediction," *Applied Energy*, vol. 285, p. 116452, Mar. 2021, doi: 10.1016/j.apenergy.2021.116452.

[11] C. B. Santiago, J. Y. Guo, and M. S. Sigman, "Predictive and mechanistic multivariate linear regression models for reaction development," *Chemical Science*, vol. 9, no. 9, pp. 2398–2412, Feb. 2018, doi: 10.1039/C7SC04679K.

[12] C. Chu *et al.*, "Modeling the impact of some independent parameters on the syngas characteristics during plasma gasification of municipal solid waste using artificial neural network and stepwise linear regression methods," *Renewable and Sustainable Energy Reviews*, vol. 157, p. 112052, Apr. 2022, doi: 10.1016/J.RSER.2021.112052.

[13] D. H. Maulud and A. M. Abdulazeez, "A Review on Linear Regression Comprehensive in Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 01, no. 04, pp. 140–147, 2020, doi: 10.38094/jastt1457.

[14] P. L. Bartlett, P. M. Long, G. Lugosi, and A. Tsigler, "Benign overfitting in linear regression," *Proceedings of the National Academy of Sciences*, vol. 117, no. 48, pp. 30063–30070, Dec. 2020, doi: 10.1073/pnas.1907378117.

[15] B. Khemet and R. Richman, “A univariate and multiple linear regression analysis on a national fan (de)Pressurization testing database to predict airtightness in houses,” *Building and Environment*, vol. 146, pp. 88–97, Dec. 2018, doi: 10.1016/J.BUILDENV.2018.09.030.
[16] X. Lin, G. Zhang, S. Wei, and Y. Yin, “Energy consumption estimation model for dual-motor electric vehicles based on multiple linear regression,” *https://doi.org/10.1080/15435075.2020.1763358*, vol. 17, no. 8, pp. 488–500, Jun. 2020, doi: 10.1080/15435075.2020.1763358.
[17] V. Thamizhmani, S. R. Kumar, V. Pranesh, S. Ravikumar, and S. Padakandla, “Multiple Linear Regression Theory Based Performance Optimization of Bakken And Eagle Ford Shale oil Reservoirs,” *Journal of Engineering Research and Application www.ijera.com*, vol. 8, pp. 66–88, 2018, doi: 10.9790/9622-0801026688.
[18] M. Ali, R. Prasad, Y. Xiang, and R. C. Deo, “Near real-time significant wave height forecasting with hybridized multiple linear regression algorithms,” *Renewable and Sustainable Energy Reviews*, vol. 132, p. 110003, Oct. 2020, doi: 10.1016/J.RSER.2020.110003.
[19] F. M. S. Al-Zwainy, F. M. Al-Zwainy, H. Ala, and H. F. Ibraheem, “Predicting productivity in construction industry utilizing multiple linear regression technique and artificial neural network technique: A review for research and applications,” *International Journal of Research in Advanced Engineering and Technology 7 International Journal of Research in Advanced Engineering and Technology*, 2020, [Online]. Available: https://www.researchgate.net/publication/359704984
[20] S. Liu, M. Lu, H. Li, and Y. Zuo, “Prediction of Gene Expression Patterns With Generalized Linear Regression Model,” *Frontiers in Genetics*, vol. 10, 2019, Accessed: Apr. 20, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fgene.2019.00120
[21] S. Rath, A. Tripathy, and A. R. Tripathy, “Prediction of new active cases of coronavirus disease (COVID-19) pandemic using multiple linear regression model,” *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, no. 5, pp. 1467–1474, Sep. 2020, doi: 10.1016/j.dsx.2020.07.045.
[22] A. A. Babatunde and S. Abbasoglu, “Predictive analysis of photovoltaic plants specific yield with the implementation of multiple linear regression tool,” *Environmental Progress & Sustainable Energy*, vol. 38, no. 4, p. 13098, Jul. 2019, doi: 10.1002/EP.13098.