IMPLEMENTATION OF MULTIPLE LINEAR REGRESSION METHODS AS PREDICTION OF VILLAGE SPENDING ON VILLAGE FINANCIAL MANAGEMENT SYSTEM

[a]Hanna Theresia Siregar, [b]Nisa Hanum Harani, [c]Cahyo Prianto

[a,b,c]Applied Bachelor Program of Informatics Engineering Politeknik Pos Indonesia
Jl. Sariasih No.54 – Bandung 40151, Indonesia
E-mail: hannatheresia6@gmail.com, nisahanum@poltekpos.ac.id, cahyoprianto@poltekpos.ac.id

Abstract

The realization of village welfare and improvement of village development can be started from the financial management aspects of the village. The village government has authority ranging from planning, implementation, reporting to accountability. There are two important variables as the financial aspects, there is village income, and village expenditure. The village budget process is a plan that will be compiled systematically. Planning has an association with predictions which is an indication of what is supposed to happen and predictions relating to what will happen. To provide a good village budget planning the village budget prediction feature is required. This prediction feature is done using data mining which is modeled i.e. multiple linear regression algorithm. The variable is selected using a purposive sampling technique and the sample count is 29 villages. Dependent variables are village Expenditure as Y, and independent variables i.e. village funds as X1 and village funding allocation as X2.   The best values as validation were gained in the 3rd fold with a correlation coefficient of 0.8907, Mean Absolute Error value of 87209395.37, the value of Root Mean Squared Error of 114867675.6, Roll Absolute Error (RAE) Percentage was 42 %, and Root Relative Squared Error was 44 %.

Key words: Multiple Linear Regression, Planning, Prediction, Village Expenditure.

.

INTRODUCTION

The village is the smallest unit in the government and has an important role in national development[1][2]. The village itself has the authority to manage its household, which means that the village has to regulate its village potential to improve the welfare and quality of life of the village community[3].

The realization of village welfare and quality improvement and village development can be started from the financial aspects of village management. In managing the finances of the village, the village government has authority ranging from planning, implementation, reporting to accountability. The good village financial management activities must be supported by competent human resources (HR) and have the quality and system and procedures of adequate village financial management[4]. In the financial aspects of the village, two things become the focus of the village, village income, and village expenditure. Village income is derived from various sources of village income. Village revenue is used to finance village spending, especially for village spending that is prioritized to improve the welfare and development of the village, where the village's shopping has been through deliberations or village agreements themselves[5].

In the Ministry of Interior regulation of the Republic of Indonesia No. 113 the year 2014 About Village Finance Management, the village revenue is all revenues in the form of money through a village account which is the right of the village in 1 year of unnecessary budget Re-payment by the village. Village income is derived from the village's original income consisting of the results of the business and the results of assets, self-participation, mutual assistance, and other indigenous village income. The village revenue also includes the allocation of State budget revenue and expenditure (APBN), part of local tax returns and district/city Levy, Village Fund Allocation, financial assistance of provincial and APBD districts, grants and donations, And others who are the legitimate income of the village. Meanwhile, village shopping is all form of village expenditure through village account which is the obligation of the village in the 1 year budget that will not be obtained by the village again [6].

Village spending can be the needs that will be fulfilled by the village itself, whether it is primary needs, basic service, environment, and community empowerment activities of the village. These needs have been through an agreement and aim to encourage the development of villages and the welfare of the village[7].

Every village in Indonesia has different income sources, it is adapted to the potential owned by the village itself. Village revenue and village spending are two interdependence, where village spending must adjust to village revenue. In drafting the government work plan, the village Government will verify and ratification of what kind of shopping will be included in the government work plan[8]. This budget planning has a close relationship to the predictions, planning a good program will affect the success of the village program. Planning plan is always related to the name of the prediction, predictions relate to what will happen, while planning is related to what should happen. The application of predictive methods in this budget planning will help the village government to develop good budget planning[9].

In this research, researchers conducted research using APBD data of villages in the city of Bandung obtained from the Bandung city of BPS. To gauge how the village's revenues affect village spending, researchers use statistical methods of multiple linear regression, which is one of the statistical methods used to test causal relationships as a result of one dependent variable and two or more independent variables. In general, the free factor is usually denoted by X and the fixed factor is denoted by the variable Y[10]. In this method, the fixed variable is village shopping, and the free variable is village funds and village Fund allocation both of these free variables are Income from the village. In village revenue sources, these two variables have a greater value among other sources. Note the table 1. Concerning table 1 the magnitude of the village funds, the allocation of village funds, and village spending showed that the relationship or influence of these three variables. If the value of village income and village funds allocation rises then village spending will also follow the value of the village income itself, this is the opposite as well if the village income decreases then the value of the village expenditure will also

decline[11].

Table 1. Village fund, village fund allocation, and village expenditure realization.

Name of Village	Village Funds	Village funds allocation	Village Expenditure
Cihideung	1,149,492,000	533,873,000	1,847,903,755
Cigugur Girang	1,138,071,000	523,246,000	1,799,470,606
Sariwangi	983,839,000	565,377,000	1,660,892,563
Cihanjuang	1,092,593,000	362,079,000	1,613,325,997
Sarijadi	1,005,661,000	400,112,000	1,786,901,294
Karyawangi	832,562,000	427,431,000	1,433,854,000
Ciwaruga	875,825,000	392,641,000	1,531,556,522

With the implementation of this multiple linear regression statistical method, researchers will measure the impact of village income and village funding allocation on village spending. In addition to determining the influence of variables, researchers will also implement a statistical method of multiple linear regression in the system, to measure the predicted number of village spending[12]. For now, the budget logging process in some villages still uses manual logging, noting what the budget is. In this system, researchers will design a system for the budget recording process and add predictive features that will help the village government to determine the proper use of the budget by predicting village spending as the basis for Specifies the government work plan. For the results of this study with validation accuracy of 94.47% precision and the regression model obtained is $Y = -217,067,261 + 1.3033363\ X_1 + 1.253114628\ X_2$

MATERIAL AND METHODS
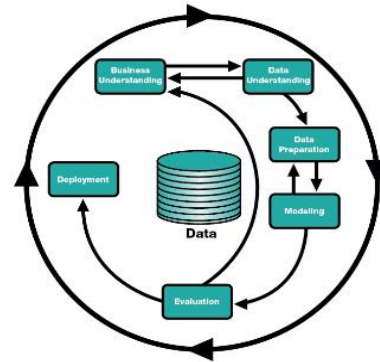
Research Methodology

Fig 1. Cross Industry Standard Process for Data Mining

The phase of the research methodology flowchart [13] contained in the Fig.1 will be described in general as follows:

1. Bussiness Understanding

Good and precise use of budgets will support Village development and the welfare of the village itself. Before using APBDes, the village will first prepare the (government work plan). These funding submissions must be completely accounted for and have valid documents. The relationship between government work plan and the village spending on this research is where the village's shopping comes from the expenses made by the village within 1 year. Planning alone relates to what is supposed to be done and has relevance to the predictions.

Predictions in the form of village spending can be used as a village guideline or reference in determining the use of budget anywhere, and the total prediction of village spending can also be used as an evaluation material to assist the Government in managing Budget. This will also help to avoid legal issues related to malpractice or improper use of village budgets. Imprecision of budget use goals can be a serious problem that the village has, the level of community confidence will decrease when there is an improper use of funds. Village budget preparation will adjust to each other's village spending plan by utilizing the amount of village income gained by the village.

2. Data Understanding

The data used is secondary data derived from the APBDes report in 2018. Data obtained from the website bps.go.id. and databandung.go.id. The data retrieval technique uses a purposive sampling technique, by taking certain criteria in the sampling. For the number of samples using as many as 29 samples of villages in West Java.

3. Data Preparation

Researchers will conduct testing using calculations manually using Microsoft Excel, Weka, and SPSS. SPSS is used for statistical analysis. The program has a fairly high analysis capability besides it also has a data management system on the graphical environment. The use of the program is also simple with a variety of descriptive menus and dialogue boxes.

4. Modelling

The researchers implemented the model that has been obtained by conducting various tests.

5. Evaluation

The results of the models already obtained will then be evaluated in the context of the research objectives set in the initial stage (business understanding). This will lead to the identification of other needs through the introduction of the acquired pattern. Acquisition of business understanding is a recurring procedure in data mining, which results from visualizations, statistic facts, and artificial intelligence methods towards new relationships. For self-evaluation, The research will be validated using the tools Weka, using these tools will find the validation result whether the model is good after performing the Test stage. In the tools, It will also measure the results of estimated performance from regression by measuring the values Correlation, MAE, RMSE, RAE, and RRSE.

6. Deployment

After going through all the research phases, researchers use the results of regression models obtained from the manual calculation process or through calculations using SPSS tools. Then all the design threads will be applied to the system, and the sample data is used on the system to determine the regression formula. The process of creating an APBDes report will begin from the village's income input, then to the shopping village, in this infiltration village spending will have a total village spending automatically/the maximum limit of village spending. At the

village spending, the related parties will adjust any budget or government work plan to be included. After that, the village financing will be inputted as the village income and village expenditure. The Output of the research itself is the APBDes report consisting of village revenue, village expenditure and village financing.
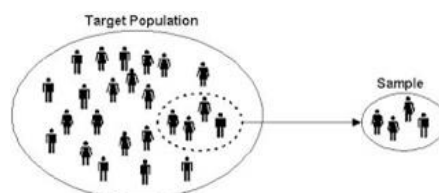
Data Retrieval Methods

Fig 2. Population and Sample

Purposive Sampling technique, the use of this data retrieval technique is used to determine non-random sampling retrievals, this data retrieval technique determines the special Citi characteristic by adjusting to the purpose of the research itself [14]. The data retrieval technique itself is part of the Non-random sampling technique by not providing the same opportunity in population members to be sampled on research. For the population in this study using village data in the province of West Java and for samples used as many as 29 samples of data.

Classic Assumption Test

In a double linear regression model, there is a classical assumption test to analyze whether the data used meets the criteria of BLUE (Best, Linear, Unbiased Estimator)[15]. Here are some stages of testing the data used on the double linear regression model:

1. Multicholinerity Classic Assumption Test

Fig 3. Multicholinerity Test

Based on the Fig.2 , following the rule at output SPSS, that if the value of tolerance in the output is obtained > 0.1 and the value of variance inflation factor (VIF) < 10 Then there is no symptom of multicollinearity[16]. The value of tolerance are both free variables in this study

of 0479 > 0.1 and the VIF value of 2,086 < 10 Then there are no symptoms of multicollinearity.

2. Autocorrelation Classic Assumption Test

Model Summary^b					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.915^a	.837	.825	107630251.5	1.665
a. Predictors: (Constant), ADD (X2), Dana (X1)					
b. Dependent Variable: Belanja (Y)					

Fig 4. Autocorrelation Test

Following the provisions of testing in SPSS, based on output obtained in Fig 3. using the value of Durbin Watson by 1,665, where the value of du at k = 2 and N = 29 are 1,563. Following the provisions that Du < Durbin Watson < 4-du or 1, which means 1,563 < 1,665 < (4-1.563 = 2,437). It can be concluded that there are no symptoms of autocorrelation [17].
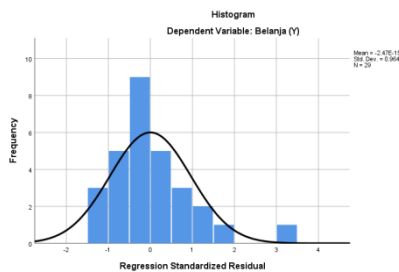
3. Normality Classic Assumption Test

Fig 5. Test the normality by Histogram

This image is the output of SPSS tools, the above diagram is a histogram to indicate normality. In multiple linear regression, the normality test will test the residual instead of the variable data. According to the output rule of the normal distributed data SPSS if the Histogram diagram forms a normal curve then the residual is expressed normal and the assumption of normality is fulfilled[18].
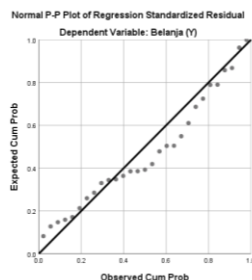

Fig 6. Test the normality by Normal P-Plot

In this SPSS output, it can be noted that the plots follow a straight line groove, thus it can be assumed that the normal distributed data.

One-Sample Kolmogorov-Smirnov Test			
			Unstandardized Residual
N			29
Normal Parameters^{a,b}	Mean		-.0000003
	Std. Deviation		103715105.5
Most Extreme Differences	Absolute		.151
	Positive		.151
	Negative		-.084
Test Statistic			.151
Asymp. Sig. (2-tailed)			.092^c
a. Test distribution is Normal.			
b. Calculated from data.			
c. Lilliefors Significance Correction.			

Fig 7. Test the normality by Kolmogrov Smirnov

In this SPSS output, it can be noted that the plot follows a straight line groove, so it can be assumed that the distributed data is normal.

Hypothesis Testing

1. Statistical Test F

ANOVA^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1.548E+18	2	7.739E+17	66.806	.000^b
	Residual	3.012E+17	26	1.158E+16		
	Total	1.849E+18	28			
a. Dependent Variable: Belanja (Y)						
b. Predictors: (Constant), ADD (X2), Dana (X1)						

Fig 8. Statistical Test F

Based on the output of multiple linear regression in Figure 6, the statistical test result of F with a significance value of 0.000 which means the sig value of < 0.05, i.e. 0,000 < 0.05, it can be deduced village funds (X1) and village funds allocation (X2) simultaneously Significant effect on village spending (Y1).

2. Statistical Test t

Coefficients^a										
		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	-217067261	156387772.4		-1.388	.177	-539526931	104392409.1		
	Dana (X1)	1.303	.240	.620	5.427	.000	.810	1.797	.479	2.086
	ADD (X2)	1.253	.398	.360	3.149	.004	.435	2.071	.479	2.086
a. Dependent Variable: Belanja (Y)										

Fig 9. Statistical Test t

Based on the results of the regression in Figure 7, the result of the test value is statistical significance t in column P > — t — = 0.00. Then it can be concluded that:
Variable significance value X1 (village Fund) P > — t — = 0.00 < = 0.05. It was concluded that

the X1 (village Fund) variable has a significant significance and positive impact on the Y variable (village expenditure). Variable significance value X2 (village Fund Allocation) P > — t — 0.04 < = 0.05. It can be concluded that the variable X2 (village fund allocation) is a significant point towards Y (village spending)..

Multiple linear regression

Multiple linear regression methods are one of the statistical methods used to test causal relationships from one variable dependent and two or more independent variables[19]. In general, the causal factor is usually denoted by X and the resulting factor is denoted by the variable Y. Multiple linear regression methods are also often used in production to perform production forecasting related to quality as well as Quantity[20]. In general, the formula of double linear regression is:

$$Y = a + b1X1 + b2X2 + \cdots + bnXn \qquad (1)$$

Where:
Y : Variabel dependent
a : Constant
b1 : The value of the regression coefficient or the size of the first variable value
b2 : The value of the regression coefficient or the size of the second variable value
X1 : First Variabel independent
X2 : Second Variabel independent

To obtain a constant value of a and a regression coefficient value of b1 and b2, can use the formula below:

$$b1 = \frac{(\sum X2^2)(\sum X1Y) - (\sum X1X2)(\sum X2Y)}{(\sum X1^2)(\sum X2^2) - (\sum X1X2)^2} \qquad (2)$$

$$b2 = \frac{(\sum X1^2)(\sum X2Y) - (\sum X1X2)(\sum X1Y)}{(\sum X1^2)(\sum X2^2) - (\sum X1X2)^2} \qquad (3)$$

To get a constant value, then use the following formula to make use of the values b1 and b2 are obtained first.

$$a = \bar{Y} - b1\overline{X1} - b2\overline{X2} \qquad (4)$$

To perform the analysis using multiple linear regression, for the stage used using a small quadratic method. The first stages are as follows:
1. Determine what are the objectives in conducting multiple linear regression analyses.
2. Identify the variables to be used, e.g. which x variables, and which y variables and collect the data associated with the variables
3. Calculate the $X1^2$, $X2^2$, and Y^2, $X1Y$, $X2Y$, $X1X2$, as well as the average value & the total of each of the sample data variables used.
4. Then calculate the values of constants A, b1 and b2.
5. Make a model or multiple linear regression equation.

Data Calculation Process

With the use of 29 samples of village data in West Java, researchers will analyze the relationship or influence of village income to village expenditure. For variable X or variable due is village spending and variable Y or variable free is village fund and village fund allocation. In the calculation of data, researchers first calculate the total of variables X1, X2 and variable Y. Note on Fig. 8

N ke	Dana Desa(X1)	ADD(X2)	Belanja(Y)
1	1,149,492,000	533,873,000	1,847,903,755
2	1,138,071,000	523,246,000	1,799,470,606
3	983,839,000	565,377,000	1,660,892,563
4	1,092,593,000	362,079,000	1,613,325,997
5	1,005,661,000	400,112,000	1,786,901,294
6	832,562,000	427,431,000	1,433,854,000
7	875,825,000	392,641,000	1,531,556,522
8	918,327,000	333,968,000	1,356,131,147
9	841,171,000	396,428,000	1,354,011,540
10	807,027,000	347,856,000	1,283,926,796
11	761,443,000	300,340,000	1,238,437,240
12	787,930,000	334,291,000	1,222,862,095
13	763,118,000	289,737,000	1,141,614,542
14	1,167,914,000	528,452,000	2,306,565,844
15	1,060,114,000	460,962,000	1,828,673,765
16	1,127,498,000	467,993,000	1,753,151,737
17	973,339,000	389,430,000	1,659,584,039
18	1,058,727,000	405,190,000	1,562,994,760
19	957,470,000	484,092,000	1,567,962,000
20	949,955,000	401,049,000	1,462,439,437
21	909,738,000	448,263,000	1,594,475,472
22	878,761,000	425,948,000	1,463,221,461
23	899,416,000	351,086,000	1,351,900,395
24	812,180,000	390,996,000	1,293,917,750
25	921,316,977	360,741,000	1,286,057,977
26	852,847,000	336,749,000	1,285,168,947
27	867,484,000	320,860,000	1,284,178,164
28	810,758,000	367,515,000	1,352,374,471
29	848,628,000	321,102,000	1,262,018,832
Total	27,053,204,977	11,667,807,000	43,585,573,148
Rata2	932,869,137.14	402,338,172.41	1,502,950,798.21

Fig 10. Calculation Data

The value of each sample data exists in the X1, X2 and Y variables, and then in the search for the total squared value of each sample. Once found the quadratic value of each new sample will be equalized. After that, the variable

multiplication value is X1, X2 and the variable Y. Multiply the values of the X1, X2 and Y variables in each sample. Then in total, the value of all of these totals will be used to determine the values a, b1, and b2.

For value N in research is the amount of data sample used in research, for sample data as many as 29 villages. The following formulas are used to specify the values of the variables to be used in the formula to calculate the value of regression coefficient b1, b2 and a constant.

$$\sum Y^2 = \sum \bar{Y}^2 - N\bar{Y}^2 \tag{5}$$

$$\sum X1^2 = \sum X1^2 - N\overline{X1}^2 \tag{6}$$

$$\sum X2^2 = \sum X2^2 - N\overline{X2}^2 \tag{7}$$

$$\sum X1Y = \sum X1Y - N\overline{X1}\bar{Y} \tag{8}$$

$$\sum X2Y = \sum X2Y - N\overline{X2}\bar{Y} \tag{9}$$

$$\sum X1X2 = \sum X1X2 - N\overline{X1X2} \tag{10}$$

Once the searched total value is obtained then specify a constant value a and the regression value b1 and b2 by using the formula (2) and the formula (3). Then the values of b1 and b2 that have been obtained are used in the general formula (4) of multiple linear regression, after which the specified value.

X1Y	X2Y	X1X2
2,124,150,583,142,460,000	986,545,921,393,115,000	613,682,742,516,000,000
2,047,925,312,041,030,000	941,565,796,707,076,000	595,491,098,466,000,000
1,634,050,878,289,360,000	939,030,454,591,251,000	556,239,942,303,000,000
1,762,708,691,040,220,000	584,151,463,667,763,000	395,604,980,847,000,000
1,797,016,942,225,330,000	714,960,650,544,928,000	402,377,034,032,000,000
1,193,772,353,948,000,000	612,873,649,074,000,000	355,862,808,222,000,000
1,341,375,490,880,650,000	601,351,884,354,602,000	343,884,803,825,000,000
1,245,371,847,831,070,000	452,904,406,901,296,000	306,691,831,536,000,000
1,138,955,241,113,340,000	536,768,086,779,120,000	333,463,737,188,000,000
1,036,163,590,395,490,000	446,621,639,549,376,000	280,729,184,112,000,000
942,999,367,337,320,000	371,952,240,661,600,000	228,691,790,620,000,000
963,529,730,513,350,000	408,791,792,599,645,000	263,397,907,630,000,000
871,186,606,061,956,000	330,767,972,555,454,000	221,103,519,966,000,000
2,693,870,541,129,420,000	1,218,909,333,393,490,000	617,186,489,128,000,000
1,938,602,659,709,210,000	842,949,116,061,930,000	488,672,269,668,000,000
1,976,675,077,164,030,000	820,462,740,853,841,000	527,661,171,514,000,000
1,615,337,868,936,220,000	646,291,812,307,770,000	379,047,406,770,000,000
1,654,784,753,270,520,000	633,309,846,804,400,000	428,985,593,130,000,000
1,501,276,576,140,000,000	759,037,860,504,000,000	463,503,567,240,000,000
1,389,251,655,375,330,000	586,509,873,769,413,000	380,978,502,795,000,000
1,450,554,926,946,340,000	714,744,358,505,136,000	407,801,885,094,000,000
1,285,821,954,289,820,000	623,256,254,870,028,000	374,306,490,428,000,000
1,215,920,845,669,320,000	474,633,302,078,970,000	315,772,365,776,000,000
1,050,894,118,195,000,000	505,916,664,579,000,000	317,559,131,280,000,000
1,184,867,047,616,380,000	463,933,840,680,957,000	332,356,807,599,957,000
1,096,052,480,942,110,000	432,779,357,733,303,000	287,195,374,403,000,000
1,114,004,010,419,380,000	412,041,405,701,040,000	278,340,916,240,000,000
1,096,448,421,359,020,000	497,017,903,709,565,000	297,965,726,370,000,000
1,070,984,517,362,500,000	405,236,770,992,864,000	272,496,148,056,000,000
41,434,554,089,344,200,000	17,965,316,401,924,900,000	11,067,051,226,755,000,000

Fig 11. Calculation Data II

X1²	X2²	Y²
1,321,331,858,064,000,000	285,020,380,129,000,000	3,414,748,287,743,100,000
1,295,205,601,041,000,000	273,786,376,516,000,000	3,238,094,461,858,010,000
967,939,177,921,000,000	319,651,152,129,000,000	2,758,564,105,828,710,000
1,193,759,463,649,000,000	131,101,202,241,000,000	2,602,820,772,596,040,000
1,011,354,046,921,000,000	160,089,612,544,000,000	3,193,016,234,498,870,000
693,159,483,844,000,000	182,697,259,761,000,000	2,055,937,293,316,000,000
767,069,430,625,000,000	154,166,954,881,000,000	2,345,665,380,080,740,000
843,324,478,929,000,000	111,534,625,024,000,000	1,839,001,687,863,540,000
707,568,651,241,000,000	157,155,159,184,000,000	1,833,347,250,453,170,000
651,292,578,729,000,000	121,003,796,736,000,000	1,648,468,017,486,830,000
579,795,442,249,000,000	90,204,115,600,000,000	1,533,726,797,418,820,000
620,833,684,900,000,000	111,750,472,681,000,000	1,495,391,703,387,790,000
582,349,081,924,000,000	83,947,529,169,000,000	1,303,283,762,505,870,000
1,364,023,111,396,000,000	279,261,516,304,000,000	5,320,245,992,707,430,000
1,123,841,692,996,000,000	212,485,965,444,000,000	3,344,047,738,799,280,000
1,271,251,740,004,000,000	219,017,448,049,000,000	3,073,541,012,946,120,000
947,388,808,921,000,000	151,655,724,900,000,000	2,754,219,182,503,550,000
1,120,902,860,529,000,000	164,178,936,100,000,000	2,442,952,619,787,460,000
916,748,800,900,000,000	234,345,064,464,000,000	2,458,504,833,444,000,000
902,414,502,025,000,000	160,840,300,401,000,000	2,138,729,106,892,880,000
827,623,228,644,000,000	200,939,717,169,000,000	2,542,352,030,809,620,000
772,220,895,121,000,000	181,431,698,704,000,000	2,141,017,043,930,970,000
808,949,141,056,000,000	123,261,379,396,000,000	1,827,634,678,001,160,000
659,636,352,400,000,000	152,877,872,016,000,000	1,674,223,143,765,060,000
848,824,972,108,419,000	130,134,069,081,000,000	1,653,945,120,205,330,000
727,348,005,409,000,000	113,399,889,001,000,000	1,651,659,222,333,090,000
752,528,490,256,000,000	102,951,139,600,000,000	1,649,113,556,894,410,000
657,328,534,564,000,000	135,067,275,225,000,000	1,828,916,709,812,530,000
720,169,482,384,000,000	103,106,494,404,000,000	1,592,691,532,322,640,000
25,656,183,598,750,400,000	4,847,063,126,853,000,000	67,355,949,280,193,000,000

Figure 12. Calculation Data III

Table 2. Variable Value calculations

Symbol	Value
ΣY²	1,848,977,327,101,330,000
ΣX1²	419,083,615,040,844,000
ΣX2²	152,658,982,396,137,000
ΣYX1	774,918,075,107,222,000
ΣYX2	429,176,557,950,917,000
ΣX1X2	182,514,178,373,048,000

RESULT AND DISCUSSION

Determination of multiple linear regression models

Specify a value of b1 and b2 by using the value that is in table 2. following calculation results:

b1= (152,658,982,396,137,000)(774,918,075,107,222,000) - (182,514,178,373,048,000) (429,176,557,950,917,000) / (419,083,615,040,844,000)(152,658,982,396,137,000) - (182,514,178,373,048,000) = **1.3033363**

b2 = (419,083,615,040,844,000)(429,176,557,950,917,000) - (182,514,178,373,048,000)(774,918,075,107,222,000)/ (419,083,615,040,844,000)(152,658,982,396,137,000) - (182,514,178,373,048,000) = **1.253114628**

Once the values of b1 and b2 are obtained, next specify the constant values of the calculation

a = 1,502,950,798.21- (1.3033363*
932,869,137.14) - (1.253114628*
402,338,172.41)
a = **-217,067,261**

The value a = -217,067,261, the value b1 = 1.3033363, and the value b2 = 1.253114628, the regression then model is as follows:

$$Y = -217,067,261 + 1.3033363X1 + 1.253114628X2$$

Next, this stage will demonstrate the process of testing the data using the model already acquired. Using the X value will then be able to find what the predicted value is from the Y value.

For example a village with X1 = 1,149,492,000 and X2 = 533,873,000, what is the value Y?
Y= -217,067,261 + (1.3033363 * 1,149,492,000)+(1.253114628*533,873,000) = **1,950,111,455**

Double correlation coefficient

The multiple correlation coefficient (R) is the relative size of the association between two variables to study the relationship between two quantitative variables based on the number. The strong link between the two variables expressed in a range or size is -1 ≤ R ≥ 1. The formula used is::

$$R = \sqrt{\frac{b1 \sum X1Y + b2 \sum X2Y}{\sum Y^2}} \qquad (11)$$

Then the value R = 0.914933867 which means **-1 ≤0.914933867 ≥1.**

Coefficient of determination

Coefficient of determinations are used to measure how far the ability of the model in describing the variation of dependent variables by a regressor or independent variable. The value of the efficiency determination has an interval between 0 and 1 (0 ≤ R^2 ≥1). The formulas used are:

$$R = \frac{b1 \sum X1Y + b2 \sum X2Y}{\sum Y^2} \qquad (12)$$

Then the value R = 0.83710398which means **-1 ≤ 0.83710398≥1.**

Validate using Cross Validation

Validation testing on regression models will be used to measure the regression model obtained by using 29 samples of data already obtained. Cross-Validation is how to find the best parameter of a model by testing the amount of error in the test data[21]. In the cross-validation, the data will be divided into k samples of the same size. In the cross-validation researchers must set the number of partitions or fold, using the validation test to use the K-fold cross-validation. For the standard use of fold is to use 5, 10, 15, 20 to 30 fold. Here's the result::

Fold	Correlation	MAE	RMSE	RAE	RRSE
2	0.8656	110614600.6	132146345	51%	49%
3	0.8907	87209395.37	114867676	42%	44%
4	0.8809	91022065.95	121189914	44%	47%
5	0.8791	91448635.39	122274111	44%	48%
6	0.8707	95561542.9	126627153	46%	50%
10	0.8778	90679226.12	122530016	44%	48%
15	0.8841	88185807.61	119036647	42%	46%
20	0.8833	88883803.52	119567684	42%	46%
30	0.8789	89,688,529	121404809	43%	47%

For validation values with the best value obtained in the 3rd fold with a correlation coefficient of 0.8907, and the Mean Absolute error rate of 87209395.37, and the Root Mean Squared error value was 114867675.6, the Absolute Error (RAE) Percentage of 42% and Root Relative Squared Error of 44%. Here's a graph of regression validation results.
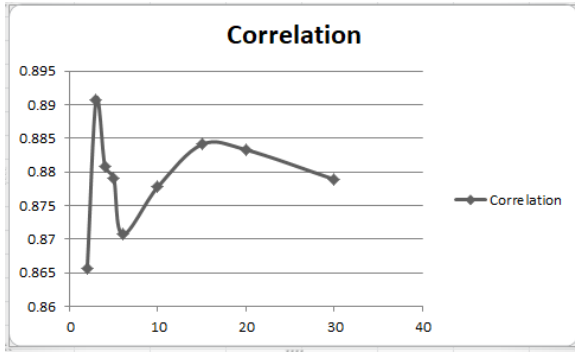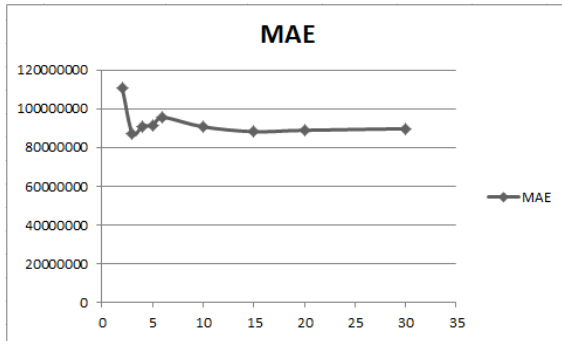
Fig 13. Graph of Correlation Result
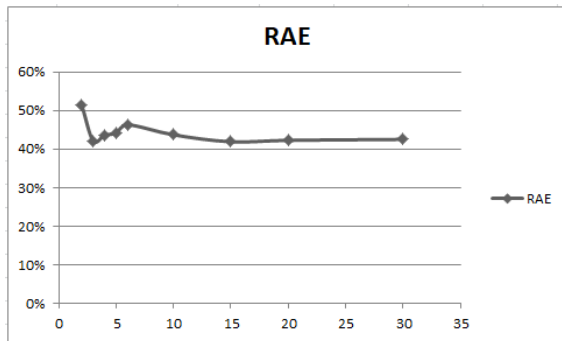

Fig 14. Graph of MAE Result
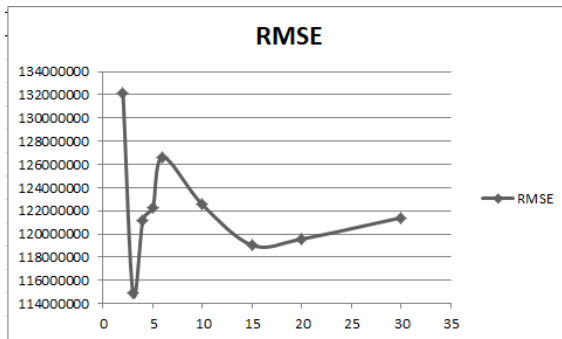

Figure 15. Graph of RAE Result
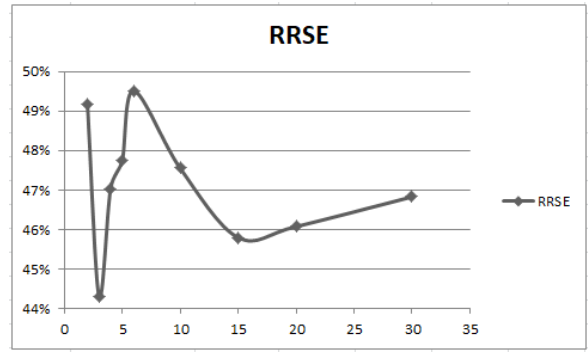

Fig 16. Graph of RMSE Result

Fig 17. Graph of RRSE Result

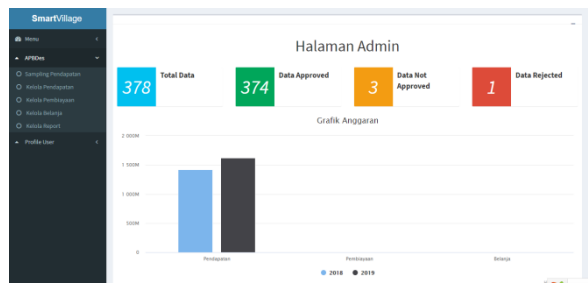## Implementation of methods on the system

1. Home Page

Fig 18. Home Page

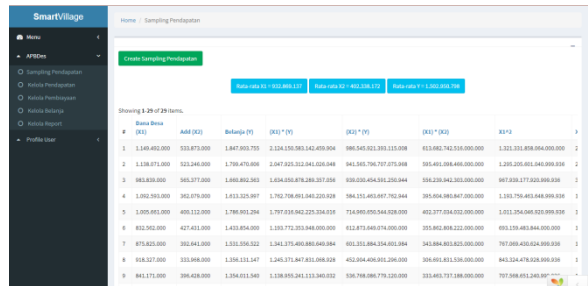In Figure. 18 is the starting page of the system.

2. Calculation Page

Fig 19. Calculation Page

In Fig. 19, this page serves to manage calculations using multiple linear regression methods.
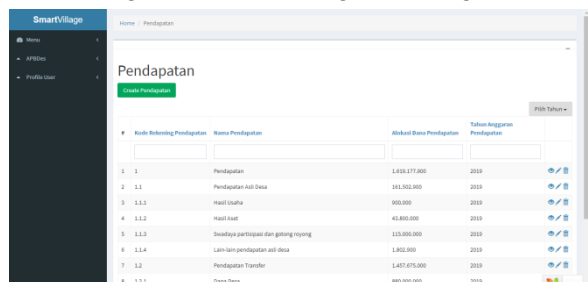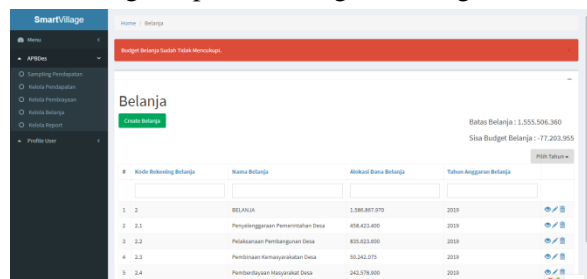
3. Village Revenue Management Page

Fig 20. Village Revenue Management Page

In Figure 20, this page serves to manage the revenue of the village, so that the value of the village funds variable and the allocation of village funds can be used on the calculation page to form the model.
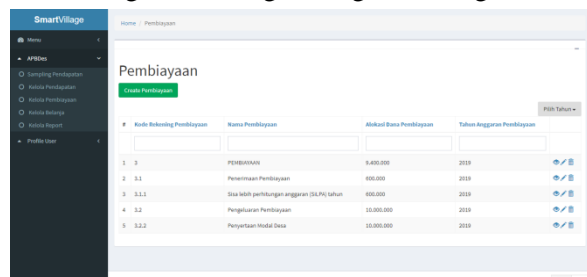
4. Village Expense Management Page

Fig 21. Village Expense Management Page

In Figure 21, this page serves to manage village spending, and there are a number of how many village spending limits. So it can be used as a reference for filling village expenditure.

5. Village Financing Management Page

Fig 22. Village Financing Management Page

In Figure 22, this page serves to manage village financing

CONCLUSION

Based on research conducted by researchers with the background of the problem regarding the management of village finances with the aim to analyze the linkages of the village fund and village funds allocation to the village expenditure, as well as manufacturing APBDes Management System and the addition of village spending prediction features to help use the government work plan budget more precisely on target. Researchers found that research could answer or provide solutions to existing problems. With the use of methods as a feature for village spending predictions, will help the village to have a picture of the village budget usage will be addressed anywhere. Total expenditure village which so predicted to be the limit of budget use and can also be used as a village driver to further improve the sources of income thus village spending will also increase and support the development of the village and progress.

A multiple linear regression method is a statistical analysis method that can be used as a predictive method and also as a method for measuring the linkage between the variables used in the study. The use of methods as a village expenditure prediction feature helps research in conducting village spending predictions using the village funds and village fund allocation variables, two variables that have gone through the Test stage on the SPSS also escaped the various Symptoms on testing or meeting the criteria of BLUE (Best, Linear, Unbiased Estimator). For validation values with the best value obtained in the 3rd fold with a correlation coefficient of 0.8907, and the Mean Absolute error rate of 87209395.37, and the Root Mean Squared error value was 114867675.6, the Absolute Error (RAE) Percentage of 42% and Root Relative Squared Error of 44%.

REFERENCES

[1] M. Anwar And B. Jatmiko, "Kontribusi Dan Peran Pengelolaan Keuangan Desa Untuk Mewujudkan Anggaran Pendapatan Dan Belanja Desa Yang Transparan Dan Akuntabel (Survey Pada Perangkat Desa Di Kecamatan Ngaglik, Sleman, Yogyakarta)," *J. Univ. Muhammadiyah Yogyakarta*, 2012.

[2] I. Ibrahim, S. Rejeki, And S. Sawaludin, "Persepsi Masyarakat Desa Bajur Kecamatan Labuapi Kabupaten Lombok Barat Terhadap Undang-Undang No 6 Tahun 2014 Tentang Desa," *Civ. Pendidikan-Penelitian-Pengabdian Pendidik. Pancasila Dan Kewarganegaraan*, Vol. 5, No. 2, Pp. 1–9, 2017.

[3] S. Suharto, "Pemberdayaan Masyarakat Desa Dalam Implementasi Uu Desa (Analisis Implementasi Uu No. 6 Tahun

2014 Tentang Desa)," *Pros. Senaspolhi*, Vol. 1, No. 1, 2018.

[4] A. Riyanto, A. Suherman, And D. Prayudi, "Akuntansi Dalam Perspektif Pengelolaan Keuangan Desa," In *Seminar Nasional Ilmu Pengetahuan Dan Teknologi Komputer*, 2016, Pp. 71--Ekm.

[5] S. Hanifa And A. Wibowo, "Faktor Yang Memengaruhi Kemiskinan Di Provinsi Jawa Timur Tahun 2014," *J. Biometrika Dan Kependud.*, Vol. 6, No. 2, Pp. 136–143, 2018.

[6] M. I. Walukow, L. Kalangi, And S. Pinatik, "Analisis Perencanaan Pengelolaan Keuangan Desa Sesuai Dengan Peraturan Menteri Dalam Negeri Nomor 113 Tahun 2014 Di Desa Kauneran I Kecamatan Sonder Kabupaten Minahasa," *J. Ris. Akunt. Going Concern*, Vol. 12, No. 2, 2017.

[7] M. Tahir, A. Anwar, And S. Dinakhir, "Analisis Pengelolaan Keuangan Dana Desa Di Desa Bululoe Kecamatan Turatea K abupaten Jeneponto," Universitas Negeri Makassar, 2019.

[8] U. Nain, "Relasi Pemerintah Desa Dan Supradesa Dalam Perencanaan Dan Penganggaran Desa." Pustaka Pelajar, 2017.

[9] Y. Wu, M. Lan, S. Sun, Q. Zhang, And X. Huang, "A Learning Error Analysis For Structured Prediction With Approximate Inference," In *Advances In Neural Information Processing Systems*, 2017, Pp. 6129–6139.

[10] J. B. Darmayasa, T. Mulyana, And Others, "Ethnomathematics: The Use Of Multiple Linier Regression Y= B 1 X 1+ B 2 X 2+ E In Traditional House Construction Saka Roras In Songan Village," In *Journal Of Physics: Conference Series*, 2018, Vol. 948, No. 1, P. 12076.

[11] D. A. Rokhmawan And D. Wahyono, "Analisis Pengaruh Pendapatan Asli Desa (Padesa), Dana Desa (Dd), Alokasi Dana Desa (Add), Dan Bagi Hasil Pajak Dan Retribusi Terhadap Belanja Desa Bidang Pertanian Tahun 2017 (Studi Empiris Di Seluruh Desa Se-Kabupaten Sragen),"

Universitas Muhammadiyah Surakarta, 2018.

[12] M. Iqbal, "Pengaruh Kepemimpinan, Dokumen Perencanaan, Dam Kompetensi Sdm Terhadap Realisasi Anggran Pendapatan Dan Belanja Desa (Apb-Desa) Pada Pemerintah Desa Di Kabupaten Maros," In *Seminar Nasional Hasil Penelitian & Pengabdian Kepada Masyarakat (Snp2m)*, 2018.

[13] F. Schäfer, C. Zeiselmair, J. Becker, And H. Otten, "Synthesizing Crisp-Dm And Quality Management: A Data Mining Approach For Production Processes," In *2018 Ieee International Conference On Technology Management, Operations And Decisions (Ictmod)*, 2018, Pp. 190–195.

[14] I. Etikan, S. A. Musa, And R. S. Alkassim, "Comparison Of Convenience Sampling And Purposive Sampling," *Am. J. Theor. Appl. Stat.*, Vol. 5, No. 1, Pp. 1–4, 2016.

[15] M. Syafruddin, L. Hakim, And D. Despa, "Metode Regresi Linier Untuk Prediksi Kebutuhanenergi Listrik Jangka Panjang (Studi Kasus Provinsi Lampung)," No. 1.

[16] J. L. Garcia-Balboa, M. V Alba-Fernandez, F. J. Ariza-López, And J. Rodriguez-Avi, "Homogeneity Test For Confusion Matrices: A Method And An Example," In *Igarss 2018 - 2018 Ieee International Geoscience And Remote Sensing Symposium*, 2018, Pp. 1203–1205.

[17] A. Fadholi, "Persamaan Regresi Prediksi Curah Hujan Bulanan Menggunakan Data Suhu Dan Kelembapan Udara Di Ternate," *Stat. J. Theor. Stat. Its Appl.*, Vol. 13, No. 1, 2013.

[18] F. E. Harrell Jr, *Regression Modeling Strategies: With Applications To Linear Models, Logistic And Ordinal Regression, And Survival Analysis*. Springer, 2015.

[19] T. A. Munandar And S. Sumiati, "Implementasi Linier Regresi Untuk Prediksi Curah Hujan Bulanan," In *Prosiding Seminar Nasional Riset*

Terapan| Senasset, 2017, Pp. 63–66.

[20]   S. Y. Fraticasari, D. E. Ratnawati, And R. C. Wihandika, "Optimasi Pemodelan Regresi Linier Berganda Pada Prediksi Jumlah Kecelakaan Sepeda Motor Dengan Algoritme Genetika," *J. Pengemb. Teknol. Inf. Dan Ilmu Komput.*

E-Issn, Vol. 2548, P. 964x, 2018.

[21] S. Yadav, "Analysis Of K-Fold Cross-Validation Over Hold-Out Validation On Colossal Datasets For Quality Classification," No. Cv, 2016
.