ANALYSIS OF LINEAR REGRESSION AND TREND MOMENT METHODS IN PREDICTING SALES USING MAPE

Adam Suhaidi Batubara*[1], Haida Dafitri[2], Ilham Faisal[3]
[1]Informatics Engineering Study Program
[2.3] Universitas Harapan Medan Indonesia
*adamsuhaidi10@gmail.com

ABSTRACK-Sales transaction data stored in the database stores a large number of transaction records, causing the amount of data to continue to increase every day. To explore sales transaction data, data mining techniques are used. One of the goals of data mining is prediction. Prediction is basically an assumption or estimate about the occurrence of an event or event in the future. Through prediction, it is expected to minimize the influence of uncertainty from the future, so that getting results that have the least prediction error is the goal of prediction. This shows that prediction is a very important tool in planning effectively and efficiently. The discussion method used to predict sales is the time series method by using a comparison of two types of prediction methods, namely the Linear Regression method and the Trend Moment method. The use of these two methods will be a better basis for making decisions to determine which method is suitable for predicting future sales. The result of a prediction cannot always be verified in absolute 100%. Therefore, the parameters used to determine the better method are based on the smallest error accuracy rate calculated using MAPE. Based on the results of the comparative prediction analysis of the Linear Regression method and the Trend Moment method, the recommended prediction result is to use the Trend Momnet method because the resulting MAPE error value is smaller, namely 0.439845%. Meanwhile, the MAPE error value with the Linear Regression method is 1.511509%.

Keywords: Prediction Accuracy, Sale, Linear Regression, Trend Moment, MAPE

1. PRELIMINARY

Sales transaction data stored in the database stores a large number of transaction records, causing the amount of data to continue to increase every day. To explore the sales transaction data used data mining techniques. Data mining uses data analysis to find patterns and relationships in the data, so that so much data can be used to predict sales in the future. The discussion method used to predict sales is the time series method by using a comparison of two types of prediction methods, namely the Linear Regression method and the Trend Moment method.

The use of these two methods will be a better basis for making decisions to find out which method is suitable for predicting future sales. The results of a prediction can not always be ascertained the truth in a matter of 100% absolute. Therefore, the parameter used to determine the better method is based on the smallest error accuracy rate calculated using MAPE.

Related research on sales predictions is used as supporting data in this study. Several studies related to using the Linear Regression method have been carried out, such as that conducted by[1], this research produces an application that is able to display prediction results based on the input trend data[2], the results obtained in this study are that the use of the Linear Regression method can be considered because the number of errors obtained in the prediction results of new student admissions for the next 1 year is not too large, namely 0.21%. Another research conducted by[3], the application of

the Linear Regression method can assist business owners in predicting the amount of raw material for the production of tofu that must be produced to meet consumer demand in order to avoid excess or shortage of stock.

Many studies related to using the Trend Moment method have also been carried out, such as that done by[4], prediction using the Trend Moment method gives an error value of 0.47% while the prediction of the Trend Moment method with the influence of the season index produces an error value of 0.45%. Thus, the influence of the season index value can reduce the error value in the prediction by 0.02%.[5]This research produces a system using the Trend Moment method that can produce fertilizer sales predictions with a success rate of above 75%. Another research conducted by[6], the error rate of the application system from the results of this study using the comparison value of predicted results with real data results has an accuracy rate of 98.25%.

Based on the background of the research problem that has been described, this research was conducted to compare the Linear Regression method with the Trend Moment method in determining the most appropriate method used to predict sales. The parameter used to determine the better method is based on the smallest error accuracy rate calculated using MAPE. From this description, the authors are interested in conducting research with the title "Analysis of Linear Regression and Trend Moment Methods in Predicting Sales Using MAPE".

2. RESEARCH CONTENT

The discussion method used to predict sales is the time series method by using a comparison of two types of prediction methods, namely the Linear Regression method and the Trend Moment method. The use of these two methods will be a better basis for making decisions to find out which method is suitable for predicting future sales. The results of a prediction can not always be ascertained the truth in a matter of 100% absolute. Therefore, the parameter used to determine the better method is based on the smallest error accuracy rate calculated using MAPE.

Prediction (prediction) is an attempt to guess or predict something that will happen in the future by utilizing various relevant information at previous times (history) through a scientific method.[7]. Through prediction, it is expected to minimize the influence of uncertainty from the future, so that getting a result that has the minimum prediction error is the goal of prediction. This shows that prediction is a very important tool in effective and efficient planning.

predictive usefulness is seen at the time of decision making. A good decision is a decision based on considerations that will occur when the decision is implemented[8]. Predictions can be classified based on the future time horizon they cover. In relation to the time horizon, predictions are divided into several categories, namely short-range forecasts, medium-range forecasts, and long-range forecasts.[9].

The Linear Regression method is based on the pattern of relevant data relationships in the past. In general, the predicted variable, such as inventory, is expressed as the variable sought. This variable is influenced by the magnitude of the independent variable. The relationship that occurs between the independent variable and the variable sought is a function [1]. Simple Linear Regression Notation which is a straight line pattern is expressed by using equation (1).

$$Y = a + bx \qquad (1)$$

Where:
Y : is the predicted variable (dependent)
x : is the independent variable
a and b : is a parameter or regression coefficient

To get value a and b, then equation (2) and equation (3) are used.

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \qquad (2)$$

$$a = \frac{\sum y - b(\sum x)}{n} \qquad (3)$$

Where is the value a is the slope, b is the intercept and n is the number of data used in the calculation[3].

Method $TrendMoment$ or often called Secular Trend is a Time-Series prediction method that adjusts the trend line on a set of past data and then is projected in a line to predict the future for short-term or long-term predictions. If the thing being studied shows symptoms of an increase, the trend that is owned shows an average increase, often called a positive trend, but if it shows symptoms that are decreasing, the trend that is owned shows an average decline or is also called a negative trend.[7]. The advantage of the Trend Moment method compared to other methods lies in the use of the X parameter that is used, so there is no difference whether the data used is an even or odd number of historical data, because the value in the X parameter always starts with a value of 0 as the first order.[10]. Trend Moment is calculated by equation (4).

$$Y = a + bX \qquad (4)$$

Where:
Y: value $trend$
a : constant number
b : $slope$ or skew coefficient
X: time index

To calculate the values of a and b, equation (5) and equation (6) are used.

$$b = \frac{\sum XY - (n * X_t * Y_t)}{\sum X^2 - (n * (X_t)^2)} \qquad (5)$$

$$a = \frac{\sum y - b \sum x}{n} \qquad (6)$$

Where:
Xt   : average demand per time period
Yt   : the sum of the time periods
n: number of data

After the predicted value that has been obtained from the prediction results with the method $Trending\ Moment$ will be corrected for seasonal effects using the seasonal index. Season index calculation[11] in equation (7).

$$Indeks\ Musim = \frac{Rata - rata\ penjualan\ bulan\ tertentu}{Rata - rata\ penjualan\ perbulan} \qquad (7)$$

To get the final prediction result after being influenced by the season index, equation (8).

$$Y^* = Indeks\ Musim * Y \qquad (8)$$

Where:
Y*:   prediction   results   using   the   method $Trending\ Moment$ influenced by the season index.
Y: prediction result using $Trending\ Moment$

In general, the error is calculated based on the difference between the actual (supposed) value and the value generated by the prediction method. Mean Absolute Percentge Error (MAPE) is the percentage calculated from the absolute value of the error in each period and divided by the actual amount

of data for the period then find the average error[11]. MAPE value can be calculated using equation (9).

$$MAPE = \frac{\sum \frac{|Y_1 - Y_t|}{Y_1}}{n} * 100\%$$

Where :
Y1   : data in period t
Yt   : forecast for period t
n: total number of periods
|| : absolute value

2.1 Results and Discussion

There is research on sales data, which is data that must be available to carry out the prediction process, therefore this prediction system will use actual sales data for the last 1 year starting from the period of January 2019 to the period of December 2019. The following is a representation of actual sales data drinks for a period of 1 year from January 2019 to December 2019 as shown in table 1.

Table 1. Beverage Sales Data for 2019

Month	Beverage Type				
	Dalgona Coffee	Spice Coffee	Coffee Milk Boba	Coffee Avocado	Cinnamon Mocha Coffee
January	304	129	234	108	132
February	315	120	256	113	131
March	394	132	246	117	135
April	408	141	234	119	135
May	420	142	267	123	129
June	417	132	268	120	126
July	410	156	291	117	131
August	356	169	267	119	137
September	378	145	278	173	137
October	375	148	283	199	128
November	410	182	287	189	133
December	421	180	281	190	138

Predictive analysis in this study uses a comparison of the Linear Regression method and the Trend Moment method. The prediction results will then calculate the error accuracy level using MAPE (Mean Absolute Percentage Error), this is done to determine the most appropriate method used to predict sales. The parameter used to determine the better method is based on the smallest error accuracy rate calculated using MAPE.

2.2 Prediction Analysis of Linear Regression Method

Completion method *Linear Regression* starting with inputting sales data and month period as input data to be processed. Next, the constant a and the regression coefficient b are calculated to get the regression equation. From the results of the regression equation is used to obtain the results of sales predictions in the next period. The prediction results will also calculate the error value using the MAPE method to determine the level of accuracy of the prediction results.

In the calculation of the results of this sales prediction, the results of manual calculations will be exemplified by taking one of the types of food and beverage data in table 2

Table 2. Beverage Sales Data Pieces for the Year 2019

Dalgona Coffee	304	315	394	408	420	417	410	356	378	375	410	421

From table 2 you can find the values for X², Y², XY and the total of each value in table 3.

Table 3. Calculation Results of X2, Y2 and XY Values of Dalgona Coffee

Month	Period (X)	Amount Sales (Y)	X2	Y2	XY
January	1	304	1	92416	304
February	2	315	4	99225	630
March	3	394	9	155236	1182
April	4	408	16	166464	1632
May	5	420	25	176400	2100
June	6	417	36	173889	2502
July	7	410	49	168100	2870
August	8	356	64	126736	2848
September	9	378	81	142884	3402
October	10	375	100	140625	3750
November	11	410	121	168100	4510
December	12	421	144	177241	5052

Total (Σ)	78	4608	650	17873 16	3078 2

After obtaining the total value of X2, Y2 and XY, then the next step is to calculate the value of the regression coefficient (b) using equation (2).

$$b = \frac{12(30782) - (78) * (4608)}{12(650) - (78)^2} = 5{,}804196$$

Then calculate the value of constant (a) using equation (3).

$$a = \frac{(4608) - 5{,}804196(78)}{12} = 346{,}2727$$

Create a linear regression equation model using equation (1), so that it becomes:

$$y = 346{,}2727 + 5{,}804196\,(x)$$

After the linear regression equation model is obtained, the next step is to make predictions. The following is the result of the prediction calculation using the Linear Regression method.

$$y = 346{,}2727 + 5{,}804196\,(13 = 421{,}7273)$$

Do the same thing until x=24

Table 4. Prediction Results of Dalgona Coffee Beverage Sales Using Linear Regression Method

No.	Moon Period (X)	Actual Data (At)	Prediction (ft)	Results Rounded
1.	January	304	421,7273	422
2.	February	315	427.5314	428
3.	March	394	433.3356	433
4.	April	408	439,1398	439
5.	May	420	444.944	445
6.	June	417	450.7482	451
7.	July	410	456.5524	457
8.	August	356	462.3566	462
9.	September	378	468,1608	468
10.	October	375	473,965	474
11.	November	410	479,7692	480
12.	December	421	485.5734	486
Total (Σ)		4608	5443,804	5445

The last step is to calculate the accuracy value *error* MAPE by using equation (9).

$$MAPE = \left(\frac{1}{12}\right) * \left|\frac{4608 - 5443{,}804}{4608}\right| = 1{,}511509\%$$

Based on the results of calculations using the Linear Regression method, the prediction error value based on MAPE is 1.511509%.

2.3 Prediction Analysis of Trend Moment Method
In the analysis of the process, the use of the Trend Moment prediction method requires input of past sales data (historical data). After obtaining these data, predictions are then calculated based on the Trend Moment method. In the calculation of the results of this sales prediction, the results of manual calculations will be exemplified by taking one of the drink type data in table 5.

Table 5. Amount and Average Sales of Dalgona Coffee

Month	Year	Total Sales (Y)	Time Index (X)	XY	X2
January	2019	304	0	0	0
February	2019	315	1	315	1
March	2019	394	2	788	4
April	2019	408	3	1224	9
May	2019	420	4	1680	16
June	2019	417	5	2085	25
July	2019	410	6	2460	36
August	2019	356	7	2492	49
September	2019	378	8	3024	64
October	2019	375	9	3375	81
November	2019	410	10	4100	100
December	2019	421	11	4631	121
Quantity (Σ)		4608	66	26174	506
Average		384	5.5	2181.167	42.16667

Based on the data that has been obtained previously in table 5, then the sales prediction results will be calculated using the Trend Moment method with equation (2.4).

To find the values of a and b in the Trend Moment formula, equations (5) and .6) are used.

$$a = \frac{4608 - 5{,}804196(66)}{12} = 352{,}0769$$

$$b = \frac{12(26174) - (66)(4608)}{12(506) - (66)^2} = 5,804196$$

After the values of a and b are known, the next step is to enter the process of determining the value of Y or Trend with equation (4). It is known that the value of a = 352.0769 and b = 5.804196 and the value of x = 12 which is a time index calculated from January 2019 to December 2019.

$$Y = a + bX = 352,0769 + (5,804196 * 13 = 427,5315)$$

rounded to 428

Lakukan hal yang sama sampai x=24

After that the prediction results obtained from the trend values above will be calculated using the season index. Based on the season index formula in equation (7) the calculation results will be obtained:

$$\text{Indeks Musim} = \frac{304}{384} = 0,791667$$

To get the final prediction results after being influenced by the season index, equation (8) is used, the calculation results will be obtained as follows:

$$Y^* = 0,791667 * 427,5315 = 338,4264.$$

Lakukan hal yang sama sampai Desember 20

Table 6. Dalgona Coffee Beverage Sales Prediction Results Using the Trend Moment Method

No.	Moon Period (X)	Actual Data (At)	Prediction (ft)	Results Rounded
1.	January	304	338,4264	338
2.	February	315	343.0575	343
3.	March	394	347.6525	348
4.	April	408	352.2475	352
5.	May	420	356.8425	357
6.	June	417	361.4375	361
7.	July	410	366.0325	366
8.	August	356	370.6275	371
9.	September	378	375,2224	375
10.	October	375	379,8174	380
11.	November	410	384,4124	384
12.	December	421	389,0074	389
Total (Σ)		4608	4364.7835	4364

After the prediction results are obtained, the error accuracy level is calculated using MAPE using equation (9) so that the following results are obtained: The error value of the prediction accuracy with MAPE is:

$$MAPE = \left(\frac{1}{12}\right) * \left|\frac{4608 - 4364,7835}{4608}\right| = 0,439845\%$$

Based on the results of calculations using the Trend Moment method, the error value of the prediction results based on the MAPE method is 0.439845%.

After calculating the predictions using the Linear Regression method and the Trend Moment method and calculating the error value of the prediction results using MAPE, then a comparison of values is carried out to determine the recommended method in determining the sales prediction results for the coming period.

Table 7. Prediction Result Comparison

No.	Moon Period	Actual Data	Prediction of Linear Regression Method	Trend Moment Method Prediction Results
1.	January	304	421,7273	338,4264
2.	February	315	427.5314	343.0575
3.	March	394	433.3356	347.6525
4.	April	408	439,1398	352.2475
5.	May	420	444.944	356.8425
6.	June	417	450.7482	361.4375
7.	July	410	456.5524	366.0325
8.	August	356	462.3566	370.6275
9.	September	378	468,1608	375,2224
10.	October	375	473,965	379,8174
11.	November	410	479,7692	384,4124
12.	December	421	485.5734	389,0074
MAPE Error Value			**1.511509 %**	**0.439845 %**

Based on table 7, it can be concluded that the best recommended method for predicting sales is the Trend Moment method based on the smallest MAPE error value.

System implementation is the execution stage of the system design that has been made into program code (source code) so that a sales prediction system application can be produced using a comparison of the Linear Regression method with the Trend Moment method which is ready to be used in accordance with the functions that have been set at the system analysis and design stage.
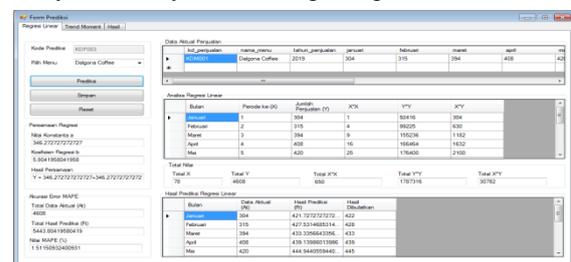
Image 1. Linear Regression Method Prediction Results Display

Figure 1 is a display of sales prediction results using the Linear Regression method. The results of sales predictions for the following year period are obtained with a MAPE error value of

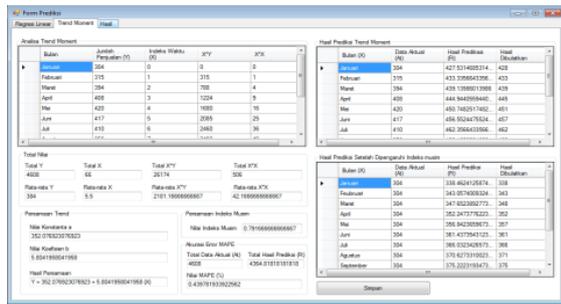1.511%. While the prediction results using the Trend Moment method can be presented in Figure 2.

Figure 2. Trend Moment Method Prediction Results Display

Figure 2 is a display of sales prediction results using the Trend Moment method. The results of sales predictions for the following year period are obtained with a MAPE error value of 0.439%.

After the prediction results have been obtained for each of the Linear Regression and Trend Moment methods, the next step is to compare the results of the two methods with the aim of finding out which prediction results are more effective to apply based on the smallest MAPE error value.
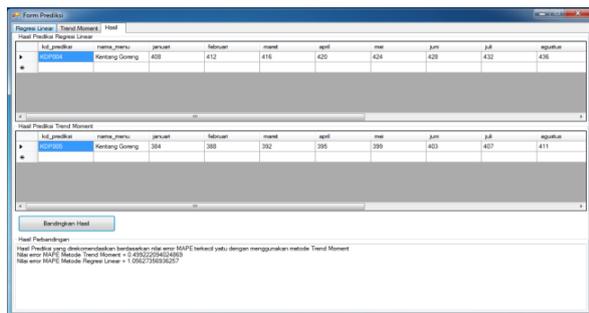

Figure 3. Comparison Result Analysis Display

Based on Figure 3, it can be explained the results of the comparison of the sales prediction method, where the results of the recommendation for the best method to make predictions use the Trend Moment method because the resulting MAPE error value is smaller than the Linear Regression method.

System testing in this study was conducted to determine the comparison of the results of sales predictions using the Linear Regression method and the Trend Moment method, so that it can be seen which prediction method will be used as a reference in making decisions to determine the results of sales predictions in the coming period.

Table 8. System Test Results

No.	Menu Type	MAPE Linear Regression	MAPE Trending Moment	Best Method Results
1.	Dalgona Coffee	1.511 %	0.439%	Trending Moment
2.	Fried Enoki Mushroom	1.322%	0.266 %	Trending Moment
3.	French fries	1.056 %	0.499 %	Trending Moment
4.	Coffee Milk Boba	1.706%	0.623 %	Trending Moment

Based on the results of the system test in table 8 using the four types of menus tested, the conclusion is that the best method that will be used as a reference for predicting the level of sales in the coming period is by using the Trend Moment method. This is based on the MAPE error value of the Trend Moment method which is smaller than the MAPE error value of the Linear Regression method.

3. CONCLUSION

Based on the analysis of sales predictions with a comparison of the Linear Regression method and the Trend Moment method, the following conclusions can be drawn:

1. The accuracy of sales predictions can be measured by the standard error of prediction, from which the prediction method will be chosen which has a value of *error* the smallest between the Linear Regression method and the Trend Moment method used. The parameter used to determine the better method is based on the smallest error accuracy rate calculated using MAPE.

2. Based on the results of the prediction analysis method comparison *Linear Regression* and the Trend Moment method, the recommended prediction result is to use the Trend Momnet method because the resulting MAPE error value is smaller, namely 0.439845%. While the MAPE error value with the Linear Regression method is 1.511509%.

3. The study resulted in a prediction system application with a comparison of the Linear Regression method and the Trend Moment method that can be used to predict the number of sales for the next one year period, so that the prediction results can be used as recommendations for improvements made by management and as a means of determine future business strategy

4. CLOSING

Based on the research that has been done as a closing some suggestions for further application development include the following.

1. This application needs to be developed by adding other prediction methods so that it can get the best prediction results

2. The calculation of the level of error accuracy in the application only uses MAPE, so it is expected to use other methods such as MSE or MAD.

BIBLIOGRAPHY

[1] DA Trianggana, "Forecasting the Number of Students Through the Linear Regression Method Approach," *J. Media Infotama*, vol. 16, no. 2, 2020.

[2] TN Putri, "Forecasting University of Samudra's New Student Admission Using Simple Linear Regression Method. J-TIFA," *J. Teknol. information.*, vol. 2, no. 1, 2019.

[3] E. Kwok and W. Susanti, "Application of Linear Regression Method in the Application of Forecasting Amount of Raw Materials for Tofu Production,"*J. Mhs. app. Technol. computer. and Inf.*, vol. 1, no. 2, 2019.

[4] I. Yulian, "Application of Trend Moment Method in Forecasting CV Product Sales," *JURTEKSI (Journal of Technol. and Information Systems)*, vol. 6, no. 2, 2020.

[5] U. Ulfa, "Forecasting Fertilizer Sales Using the Trend Moment Method," *J. Inform. icon. Business*, vol. 1, no. 4, 2019.

[6] Ilyas, "Implementation of the Trend Moment Method (Forecasting) for New Students at Widyagama University, Malang," *J. Inf. Technol. Comput. science.*, vol. 3, no. 2, 2018.

[7] SIH and SAR, "Predicting the Amount of Chicken Egg Demand Using the Trend Moment Method," *information. Mulawarman J. Ilm. Computer Science.*, vol. 14, no. 2, 2019.

[8] M. Zunaidi, "Profit/Loss Forecasting Application To Increase Sales With The Single Moving Average (SMA) Method," *J. SAINTIKOM*, vol. 15, no. 3, 2016.

[9] MEDS, "Predicting the Rupiah Exchange Rate Against the US Dollar Using the Trend Moment Method," *J. Inovtek Polbeng - Inform Series.*, vol. 4, no. 2, 2019.

[10] AAF Dliya'ul Izz, "Trend Moment Method for predicting Multimedia Equipment RentalNeeds," *inf. J. Ilm. bids. Technol. inf. and Commune.*, vol. 5, no. 1, 2020.

[11] PI, "Implementation of Trend Moment Method in Sales Forecasting for Procurement of Drugs and Plant Seeds at Tani Makmur Agricultural Stores," *Simki-Techsain*, vol. 2, no. 1, 2018.

Contact person Author: Adam Suhaidi
Mobile phone: 089518726040