



## Linear Regression Analysis in Predicting the Amount of Stock of HP Sparepart Goods in GMT

Gilang Aryudha<sup>1\*</sup>, Wilda Rina Hasibuan<sup>2</sup>

<sup>1, 2</sup>*Program Studi Sistem Nformasi, Fakultas Ilmu Computer Dan Teknologi Informasi,  
Universitas Muhammadiyah Sumatera Utara  
Jl. Kapt. Mukhtar Basri No. 3 Medan Sumatera Utara 20238 Indonesia  
[aryudhagilang1@gmail.com](mailto:aryudhagilang1@gmail.com)*

---

### Abstract

The rapid advancement of the digital era has made smartphones an essential part of daily life, making the availability of high-quality spare parts crucial for their seamless operation. GMT, a store specializing in smartphone spare parts, faces challenges in predicting fluctuating consumer demand, often leading to either stock shortages or excesses. To address this issue, this research develops a stock prediction system based on linear regression, which analyzes sales data to accurately forecast stock needs. The implementation of this method has resulted in improved accuracy in stock management, enabling GMT to optimize inventory, minimize potential losses, and enhance both customer satisfaction and operational efficiency.

**Keywords:** *Linear Regression, Stock Prediction, Spare Parts Availability, Inventory Management, Operational Efficiency*

---

### 1. Introduction

In the digital era, smartphones are central to daily activities, making the availability and quality of spare parts crucial. Disruptions to spare parts, whether due to defects or user error, can impact smartphone functionality. Spare part distributors need to meet consumer demand quickly and accurately. GMT, as a smartphone spare parts store, faces challenges in managing fluctuating inventory. Unavailable stock can lead to lost customers, while excess stock can be financially detrimental, reduce storage space, and increase the risk of product damage.

To address these issues, GMT needs an efficient method for predicting stock levels, one of which is by developing an integrated Information System-based application. Information Systems combine system and information elements to produce data that is useful for users [1]. An Information System is designed to manage and process data and provide the information needed to support organizational activities and decision-making [2].

In developing information system-based applications, selecting the right algorithm is crucial for effectively processing, analyzing, and utilizing data. For GMT, an algorithm that can accurately predict spare part demand is required. One suitable algorithm is Linear Regression, which identifies the relationship between independent and dependent variables [3] and models the relationship between one or more predictors (independent variables) and response variables (dependent variables) by assuming a linear relationship between these variables [4]. Therefore, this method can be used to predict spare parts inventory more effectively.

Previous research has shown that linear regression is effective for forecasting sales and inventory. This method allows for the identification of patterns in sales data and the creation of more accurate predictions regarding stock needs. By using linear regression, GMT can optimize stock management, reduce the risk of overstocking, and enhance customer satisfaction through better product availability. The application of linear regression also improves the accuracy of demand predictions and aids in more efficient stock planning.

### 2. Method

Quantitative research utilizes statistical analysis with numerical data to address research problems. This method involves representing data numerically and follows a structured pattern. One approach is the case study, which involves an in-depth examination of a specific setting, subject, or event. The following is a framework of the stages to be undertaken in this research, as illustrated in Figure 1 below.

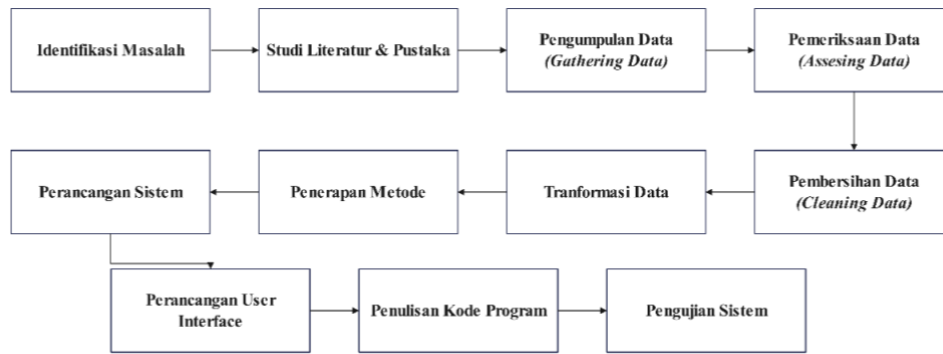


Fig. 1: Stages of System Completion

The following are the stages in completing the system:

1. At this analysis stage, data collection was carried out using three methods, namely observation, interviews, historical sales data and literature study.
2. At the design stage, the author carries out the system design process using UML (Unified Modeling Language) modeling, which includes use case diagrams, activity diagrams, sequence diagrams and class diagrams. The database design was created using the ERD (Entity Relationship Diagram) model.
3. At the coding stage, the database is implemented using MySQL. Next, to implement the system, the PHP (Hypertext Preprocessor) programming language is used in the CodeIgniter framework and uses the Linear Regression method. The linear regression method is a method for determining the relationship between independent variables and dependent variables by fitting a straight line to the data [5].

$$Y = a + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

Where,

- Y = Dependent Variable
- a = Intercept/Constant Value
- b = Regression Coefficient
- x = Independent Variable

The following formula is used to determine the value of a (constant):

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

Where,

- $\sum y$  = Sum of all Y values
- $\sum x^2$  = Sum of squares of all X values
- $\sum x$  = Sum of all X values
- $\sum xy$  = Number of products between all X and Y values
- n = Number of data

The following is the formula used to determine the b value (regression coefficient):

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

Where,

- b = Koefisien regresi
- n = Number of data
- $\sum xy$  = Number of multiplications between all X and Y values
- $\sum x$  = sum of all X values
- $\sum y$  = Sum of all Y values
- $\sum x^2$  = Sum of squares of all X values

The following error calculation formula is used to evaluate Linear Regression for MAD error calculations:

$$MAD = \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{n} \right|$$

Where,

- $y_i$  = nilai data aktual
- $\hat{y}_i$  = nilai data prediksi
- n = jumlah data

The following error calculation formula is used to evaluate Linear Regression for MSE error calculations:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Where,

- $y_i$  = actual data value
- $\hat{y}_i$  = predicted data value
- $n$  = number of data

The following error calculation formula is used to evaluate Linear Regression for MAPE error calculations:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$$

Where,

- $y_i$  = actual data value
- $\hat{y}_i$  = predicted data value
- $n$  = number of data

4. The system testing stage uses Black Box Testing and User Acceptance Testing (UAT) to ensure functionality and user satisfaction.

### 3. Result and Discussion

#### 3.1. Implementation

##### a. Regresi Linear

Linear regression calculations were carried out using weekly sales data of HP spare parts from GMT for November-December 2023, with a focus on full set LCDs, can be seen in table 1.

**Table 1:** Vivo Fullset LCD Sales Table

Sunday	Price	Sold	Initial stock	Promotion	Final Stock	Purchase
1	180000	18	150	3	65	0
2	180000	24	65	1	37	0
3	180441	34	100	4	60	0
4	182973	37	60	2	15	0
5	180000	39	100	2	62	0
6	182424	33	62	3	28	0
7	184474	38	100	4	50	0
8	183256	43	50	1	6	0
9	185000	25	170	1	145	0

The next step is to carry out multiple linear regression calculations based on the variables that have been explained. The following helper table will make the calculation process easier, can be seen in table 2.

**Table 2:** 3 Variable Helper Table Format

Y	X1	X2	X3	X1.Y	X2.Y	X3.Y	X1.X2	X1.X3	X2.X3	X1 <sup>2</sup>	X2 <sup>2</sup>	X3 <sup>2</sup>
---	----	----	----	------	------	------	-------	-------	-------	-----------------	-----------------	-----------------

The following are the results of multiple linear regression calculations with 3 variables for Vivo Fullset LCD spare parts, can be seen in table 3.

**Table 3:** Multiple Linear Regression Helper Table 3

Sunday	Final Stock (Y)	Initial Stock (X1)	Promotion (X2)	Sold (X3)	X1.Y (Initial Stock x Final Stock)	X2.Y (Promotion x Final Stock)	X3.Y (Sold x Final Stock)	X1.X2 (Initial Stock x Promotion)	X1.X3 (Initial Stock x Sold)	X2.X3 (Promotion x Sold)	X1 <sup>2</sup> (Initial Stock <sup>2</sup> )	X2 <sup>2</sup> (Promotion <sup>2</sup> )	X3 <sup>2</sup> (Sold <sup>2</sup> )
1	132	150	2	18	19800	264	2376	300	2700	36	22500	4	324
2	108	132	3	24	14256	324	2592	396	3168	72	17424	9	576
3	74	108	3	34	7992	222	2516	324	3672	102	11664	9	1156
4	37	74	5	37	2738	185	1369	370	2738	185	5476	25	1369
5	148	37	4	39	5476	592	5772	148	1443	156	1369	16	1521
6	115	148	4	33	17020	460	3795	592	4884	132	21904	16	1089
7	77	115	3	38	8855	231	2926	345	4370	114	13225	9	1444
8	170	77	4	43	13090	680	7310	308	3311	172	5929	16	1849
9	145	170	1	25	24650	145	3625	170	4250	25	28900	1	625
<b>TOTAL</b>	<b>1006</b>	<b>1011</b>	<b>29</b>	<b>291</b>	<b>113877</b>	<b>3103</b>	<b>32281</b>	<b>2953</b>	<b>30536</b>	<b>994</b>	<b>128391</b>	<b>105</b>	<b>9953</b>

The next step is to calculate the multiple linear regression process and determine the value of each coefficient needed, based on the results of the table above, the next step is to form a matrix with the format as below.

$$A = \begin{matrix} n & \sum X_1 & \sum X_2 & \sum X_3 \\ \sum X_1 & \sum X_1^2 & \sum X_1X_2 & \sum X_1X_3 \\ \sum X_2 & \sum X_1X_2 & \sum X_2^2 & \sum X_2X_3 \\ \sum X_3 & \sum X_1X_3 & \sum X_2X_3 & \sum X_3^2 \end{matrix} \dots\dots\dots(3.1)$$

$$H = \begin{matrix} \sum Y \\ \sum X_1 Y \\ \sum X_2 Y \\ \sum X_3 Y \end{matrix} \dots\dots\dots(3.2)$$

The matrix results can be seen in the format as below.

$$A1 = \begin{matrix} 2053 & 1989 & 26 & 948 \\ 462086 & 624405 & 6270 & 212459 \\ 5470 & 6270 & 88 & 2848 \\ 232013 & 212459 & 2848 & 109230 \\ 9 & 2053 & 26 & 948 \end{matrix}$$

$$A2 = \begin{matrix} 1989 & 462086 & 6270 & 212459 \\ 26 & 5470 & 88 & 2848 \\ 948 & 232013 & 2848 & 109230 \\ 9 & 1989 & 2053 & 948 \end{matrix}$$

$$A3 = \begin{matrix} 1989 & 624405 & 462086 & 212459 \\ 26 & 6270 & 5470 & 2848 \\ 948 & 212459 & 232013 & 109230 \\ 9 & 1989 & 26 & 2053 \end{matrix}$$

$$A4 = \begin{matrix} 1989 & 624405 & 6270 & 462086 \\ 26 & 6270 & 88 & 5470 \\ 948 & 212459 & 2848 & 232013 \end{matrix}$$

The next step is to calculate the determinant of each matrix, where the results of the determinant of each matrix above are as follows:

1. Det (A) = 161.930.025.100
2. Det (A1) = 19.648.813.023.200
3. Det (A2) = 29.972.263.500
4. Det (A3) = -10.253.241.172.600
5. Det (A4) = 382.460.403.700

The next step is to calculate each required coefficient, the calculation is as follows:

$$a = \frac{Det A(1)}{Det (A)} = \frac{19.648.813.023.200}{161.930.025.100} = 121.34$$

$$b1 = \frac{Det A(2)}{Det (A)} = \frac{29.972.263.500}{161.930.025.100} = 0.19$$

$$b2 = \frac{Det A(3)}{Det (A)} = \frac{-10.253.241.172.600}{161.930.025.100} = -63.32$$

$$b3 = \frac{Det A(4)}{Det (A)} = \frac{382.460.403.700}{161.930.025.100} = 2.36$$

The next step is to analyze the regression prediction results by entering calculations into the regression equation that has been obtained.

Data analysis for week 1 of the regression prediction results for Vivo Fullset LCD spareparts is as follows.

Data Minggu 1  
 Data Stok Awal (X1) = 150  
 Promosi (X2) = 2  
 Terjual (X3) = 18  
 Stok Akhir (Y) = 132  
 $Y_{pred} = 203.81 - 0.33 X1 - 24.54 X2 + 0.76 X3$   
 $Y_{pred} = 203.81 - 0.33 x 150 - 24.54 x 2 + 0.76 x 18$   
 $Y_{pred} = 203.81 - 49,5 - 49,08 + 13,68$   
 $Y_{pred} = 118,22$

The next step is to calculate the Mean Absolute Deviation (MAD), Mean Squared Error (MSE), and Mean Absolute Percentage Error (MAPE) to measure the regression's ability to predict data accurately. The calculations are as follows.

1. MAD Calculation

$$MAD = \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{n} \right|$$

$$MAD = \frac{|(132 - 118,22) + (108 - 104,29) + (74 - 119,95) + (37 - 84,55) + (148 - 123,01) + (115 - 81,25) + (77 - 120,66) + (170 - 112,66) + (145 - 141,40)|}{9}$$

$$MAD = \frac{13,78 + 3,71 + 45,95 + 47,55 + 24,99 + 33,75 + 43,66 + 57,34 + 3,60}{9}$$

$$MAD = \frac{274,33}{9} = 30,48$$

2. Perhitungan MSE

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$MSE = \frac{(132 - 118,22)^2 + (108 - 104,29)^2 + (74 - 119,95)^2 + (37 - 84,55)^2 + (148 - 123,01)^2 + (115 - 81,25)^2 + (77 - 120,66)^2 + (170 - 112,66)^2 + (145 - 141,40)^2}{9}$$

$$MSE = \frac{189,88 + 13,76 + 2111,60 + 2260,80 + 624,01 + 1138,06 + 1905,63 + 3287,63 + 12,96}{9}$$

$$MSE = 1282,96$$

3. Perhitungan MAPE

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$$

$$MAPE = \frac{10,44\% + 3,44\% + 62,09\% + 128,51\% + 16,89\% + 29,35\% + 56,70\% + 33,73\% + 2,48\%}{9}$$

$$MAPE = 38,18\%$$

The summary results of the calculations above can be seen in table 6.

Table 4: LCD Spare Parts Prediction Accuracy Table

Sunday	Final Stock (Y)	Initial Stock (X1)	Promotion (X2)	Sold (X3)	Final Stock Prediction	Absolute Error (MAE)	Squared Error (MSE)	Absolute Percentage Error (MAPE)
1	132	150	2	18	118,22	13,78	189,89	10,44%
2	108	132	3	24	104,29	3,71	13,76	3,44%
3	74	108	3	34	119,95	45,95	2111,40	62,09%
4	37	74	5	37	84,55	47,55	2261,00	128,51%
5	148	37	4	39	123,01	24,99	624,50	16,89%
6	115	148	4	33	81,25	33,75	1139,06	29,35%
7	77	115	3	38	120,66	43,66	1906,20	56,70%
8	170	77	4	43	112,66	57,34	3287,88	33,73%
9	145	170	1	25	141,4	3,6	12,96	2,48%
						<b>30,48111</b>	<b>1282,961</b>	<b>38,18%</b>
						<b>MAD</b>	<b>MSE</b>	<b>MAPE</b>

b. System

The "Item Data" menu was implemented to facilitate management, including adding, editing and deleting item data efficiently, can be seen in Figure 2.

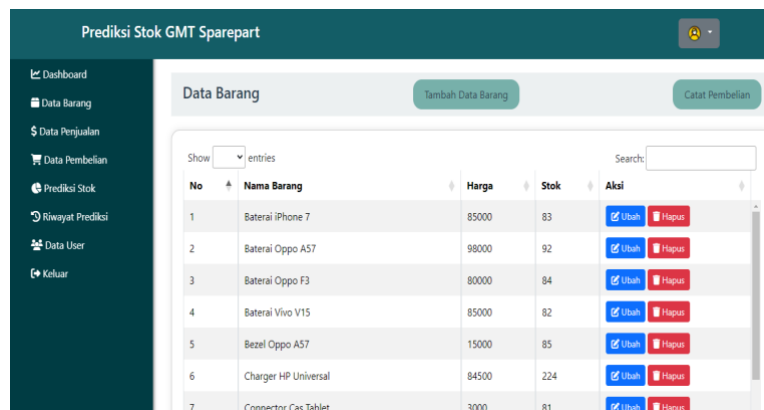


Fig. 2: Item Data Menu Display

The "Sales Data" page makes it easier to manage spare parts sales with add, edit, delete, as well as search and organize entry features. Implementation of this feature allows users to manage sales data more efficiently and accurately, can be seen in Figure 3.

No	Tanggal	Nama Barang	Jumlah Terjual	Harga	Diskon	Total	Aksi
1	2023-11-01	LCD Fullset Vivo	2	180000	0	360000	Ubah Hapus
2	2023-11-01	Charger HP Universal	16	85000	20000	1360000	Ubah Hapus
3	2023-11-02	LCD Fullset Vivo	2	180000	0	360000	Ubah Hapus
4	2023-11-02	Charger HP Universal	12	85000	0	1020000	Ubah Hapus
5	2023-11-03	LCD Fullset Vivo	4	180000	17000	720000	Ubah Hapus
6	2023-11-03	Charger HP Universal	15	85000	25000	1275000	Ubah Hapus
7	2023-11-04	LCD Fullset Vivo	10	180000	18000	1800000	Ubah Hapus

Fig. 3: Sales Data Menu Display

The stock prediction menu using multiple linear regression in the GMT system is designed to make it easier for users to calculate stock predictions based on sales data. Users select items from the available list, then define a date range by selecting start and end dates via the calendar icon or typing manually. For example, users can select the period November 1, 2023 to December 31, 2023. Can be seen in figure 4.

Minggu	Nama Barang	Harga	Terjual	Stok Awal	Promosi	Stok Akhir	Pembelian
1	LCD Fullset Vivo	180000	18	150	2	132	0
2	LCD Fullset Vivo	180000	24	132	3	108	0
3	LCD Fullset Vivo	180441	34	108	3	74	0
4	LCD Fullset Vivo	182973	37	74	5	37	0
5	LCD Fullset Vivo	180000	39	37	4	148	150
6	LCD Fullset Vivo	182424	33	148	4	115	0
7	LCD Fullset Vivo	184474	38	115	3	77	0
8	LCD Fullset Vivo	183256	43	77	4	170	136
9	LCD Fullset Vivo	185000	25	170	1	145	0

Fig. 4: Display of Sales Data After Processing

After the sales and purchase data has been processed, press the "Perform Regression Calculation" button to start the analysis. The system will calculate stock predictions and evaluate accuracy using MAD, MSE, and MAPE. Multiple linear regression results, including equations and coefficients, will be displayed to provide accurate stock predictions, which can be seen in Figure 5.

Minggu	Harga	Terjual (X3)	Stok Awal (X1)	Promosi (X2)	Stok Akhir (Y)	Pembelian	Prediksi Stok Akhir	Error	Absolute Error	Percentage Error (%)
Minggu 1	180000	18	150	2	132	0	118.22	13.78	13.78	10.44
Minggu 2	180000	24	132	3	108	0	104.29	3.71	3.71	3.44
Minggu 3	180441	34	108	3	74	0	119.95	-45.95	45.95	62.10
Minggu 4	182973	37	74	5	37	0	84.55	-47.55	47.55	128.52
Minggu 5	180000	39	37	4	148	150	123.01	24.99	24.99	16.89
Minggu 6	182424	33	148	4	115	0	81.25	33.75	33.75	29.35
Minggu 7	184474	38	115	3	77	0	120.66	-43.66	43.66	56.70
Minggu 8	183256	43	77	4	170	136	112.66	57.34	57.34	33.73
Minggu 9	185000	25	170	1	145	0	141.40	3.60	3.60	2.48

Perhitungan Matematis Detail	
<b>Mean Squared Error (MSE):</b> Rata-rata dari squared error	
$MSE = (\sum(y - \hat{y})^2) / n$	
$MSE = 11,546.61 / 9 = 1,282.96$	
<b>Mean Absolute Deviation (MAD):</b> Rata-rata dari absolute error	
$MAD = (\sum y - \hat{y} ) / n$	
$MAD = 274.33 / 9 = 30.48$	
<b>Mean Absolute Percentage Error (MAPE):</b> Rata-rata dari absolute percentage error	
$MAPE = (\sum( y - \hat{y}  / y) * 100) / n$	
$MAPE = 343.63 / 9 = 38.18 \%$	

Fig. 5: System Calculation Results

### 3.2. Testing

Linear regression analysis in predicting HP spare parts stock at GMT uses several important parameters. Mean Absolute Deviation (MAD) produces a value of 33.84, which indicates a significant deviation between the predicted value and the actual value, indicating a large prediction error. Mean Squared Error (MSE) with a value of 1441.71 shows a large difference between the predicted value and the actual value, as well as some predictions that are very deviant. Mean Absolute Percentage Error (MAPE), which shows a result of 42.64%, indicates a fairly high level of prediction error, indicating that the model needs improvement to increase its accuracy.

In addition.

Black Box testing shows that the system functions according to specifications, with 100% functional results from 5 respondents. Although linear regression analysis provides some accurate predictions, test results show that there are prediction errors that need to be corrected to improve overall system efficiency.

### 4. Conclusion

Based on linear regression analysis in predicting the stock of HP spare parts in GMT, we can draw several important conclusions. Linear regression methods, which use variables such as starting stock, promotions, and quantity sold, have proven useful in predicting ending stock by relying on historical sales data. The analysis results show that although this model can provide insights for optimizing stock management, there are some significant prediction errors. The Mean Absolute Deviation (MAD) value is 33.84, Mean Squared Error (MSE) is 1441.71, and Mean Absolute Percentage Error (MAPE) is 42.64%, indicating that there are quite large deviations and prediction errors. Web-based application development with PHP and MySQL allows GMT to carry out stock predictions efficiently, with features such as item selection, date ranges, sales data processing, and linear regression calculations. While this model needs refinement, the application provides a useful tool to better manage stock and increase product availability as needed.

### Acknowledgement

Thank you to the colleagues at the Faculty of Computer Science and Information Technology, Universitas Muhammadiyah Sumatera Utara, Medan, for your support and assistance during this research. Special thanks for the valuable collaboration and suggestions. Your support has been crucial to the success of this research.

### References

- [1] Dewi Putri, R., & Andri. (2022). "Prediksi Pembelian Produk Elektronik Yang Terlaris Pada CV.Istana Komputer Palembang Menggunakan Algoritma Regresi Linear Sederhana". *Jurnal Mantik*, 6(2).
- [2] Davis, G. B., & Olsson, T. (2020). "Management Information Systems: Managing the Digital Firm." *Journal of Management Information Systems*, 37(2), 330-350.
- [3] Harsiti, Muttaqin, Z., & Srihartini, E. (2022). "Penerapan Metode Regresi Linier Sederhana Untuk Prediksi Persediaan Obat Jenis Tablet". *JSil*, 9(1), 12–16.
- [4] Friedman, J., Hastie, T., & Tibshirani, R. (2020). "The Elements of Statistical Learning: Data Mining, Inference, and Prediction". Springer.
- [5] Montgomery, D. C., Peck, E. A., & Vining, G. G. (2020). *Introduction to linear regression analysis*. Wiley.