Journal of Data Science, Technology, and Artificial Intelligences
ISSN XXXX-XXXX
Vol. 1, No. 1, July, 2024, pp. 13-17
Journal homepage: https://journal.itisd.org/index.php/ditech

Application of Linear Regression Method in Predicting Veil Sales (Case Study: Fauzan Kerudung Shop)

Fitri Amalini[1*], Ade May Luky Harefa[2]

[1] Department of Informatics Engineering, STMIK Pelita Nusantara, Medan 20153, Indonesia
[2] Department of Informatics Engineering, STMIK Pelita Nusantara, Medan 20153, Indonesia

Corresponding Author Email: fitriaamalini0@gmail.com

https://doi.org/xx.xxxxx/ditech.xxxxxx

Received : 20 March 2024
Revised : 22 July 2024
Accepted : 23 July 2024
Availabe online : 31 July 2024

Keywords:
application, sales, prediction, linear regression, veil

ABSTRACT

The sales prediction system can support business transactions, especially those engaged in the trading sector, as an operational process that forecasts products to be sold in the future. One of the stores involved in the trading business is Toko Kerudung Fauzan. This kerchief store relies solely on estimates to determine the quantity of goods to be purchased from suppliers. This reliance on estimates has led to difficulties for the store owner in predicting future sales without performing calculations to maintain inventory levels in the store. Therefore, there is a significant need for a prediction system to determine kerchief sales in the future. The method employed for sales prediction is the Linear Regression Method. The number of data samples used consists of sales data collected from January to December 2022. Consequently, it can be stated that the decision to predict sales in 2023 shows a decline.

1. INTRODUCTION

In the current era of globalization, the rapid development of information technology has penetrated into people's lives in various fields [1]. The development of information technology can not only be utilized in the fields of education and health, but also in various fields, including the business sector that can support and facilitate human work [2]. Today, many people have used information technology in their daily activities. One way to utilize information technology in business is to use a sales prediction system [3].

One method that can be applied in predicting future sales is the Linear Regression Method. The Linear Regression Method is a statistical tool to determine the effect of one variable on another [4]. Linear regression is one of the methods used to make forecasts or predictions in the future about quality and quantity characteristics [5].

Toko Kerudung Fauzan is one of the shops engaged in trading business. This shop only sells finished goods and does not produce goods. Purchases of goods are made once a month to suppliers. This Kerudung shop only uses estimates in determining the number of items to be purchased from suppliers. The shop owner must predict in advance how much to buy without doing the calculations in advance to fulfill the inventory of goods. Sales of goods that experience this increase and decrease make the Veil Shop unable to determine the number of sales in the future and plan in advance, to purchase veils so that the inventory of goods is fulfilled.

The purpose of this research is to build a system to predict sales of hoods in 2023 with web-based. Based on the explanation above, the research title is "Application of Linear Regression Method in Predicting Veil Sales (Case Study: Fauzan Veil Shop)".

2. LITERATURE REVIEW

Several previous studies have employed the same method as the one used by the author. Based on calculations using the linear regression method, an accuracy of 98.505% was obtained. This implies that the method can be implemented in predicting the stock of tablet-type medicine (ibuprofen) at Klinik Teluk [6]. The subsequent study that successfully utilized the linear regression method is as follows: The attributes used for prediction are the opening, change, highest, and lowest values of the exchange rate of the Indonesian Rupiah against the US Dollar. The data used in this research is sourced from investing.com. From the test results, the linear regression method achieved an accuracy of 95% with a threshold value of 30 rupiah. Additionally, the root mean squared error obtained was 14.951 [7]. The subsequent study that successfully utilized the linear regression method is as follows: The Application of Multiple Linear Regression Method for Population Estimation Gunung Malela District. This Multiple Linear Regression method aims to make the best predictions. The research data used is the population in the Gunung Malela sub-district in 2016-2020. Based on the research that has been done using the MultipleLinear Regression method, the results of the population growth are 40078 residents [8]. As the final part of this study, a conclusion can be made, namely: The results of the prediction of the number of new student admissions from the Management study program with the results of the test with an

error rate of 3.444% or an accuracy level of 96.556% [9]. The results of the study resulted in a linear regression formula y = 0.24743301702419157 X1+ 0.257319654680142 X2 + 0.2662633268986068 X3 + 0.2332290731664389 X4 + 9.697801443559612e-05 X5 - 1.4522339706890541e-06 X6 with a coefficient value of 0.997. Using Python with library support makes it easy to implement machine learning specifically for linear regression [10].

3. RESEARCH METHOD

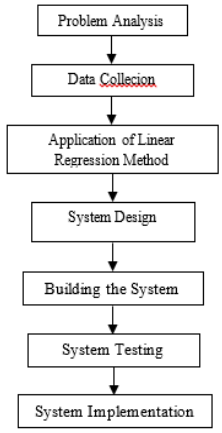


Figure 1. Research Framework

Based on the stages contained in the research framework in the figure above, the following is a description of the research framework:

3.1 Problem Analysis

Something that is done before starting aresearch. Analyzing the problem needs to be done in order to get a solution to the problem.

3.2 Data Collection

The process of collecting data in the form of information based on facts that actually occur by going directly to the Fauzan Veil Shop by conducting the observation and interview process.

3.3 Application of Linear Regression Method.

Researchers apply the Linear Regression method to solve research problems so that it is expected to provide the best solution and decision.

3.4 System Design

Drawing the framework of the system to be built. This is done to minimize errors in building the system.

3.5 Building the System

Researchers translate the program design that has been made and the system that has been designed will be implemented using the PHP programming language and MySQL as the database.

3.6 System Testing

The system that has been built will be tested to determine whether the system that has been built is running well as needed or not, so that it will produce a system that is suitable for use.

3.7 System Implementation

The final process of system design, where the system is ready for use.

.

4. RESULT AND DISCUSSION

4.1 Linear Regression Method Analysis

Linear regression is a statistical tool to determine the effect of one variable on another [11]. In linear regression analysis, the variables are divided into two, namely the independent variable and the dependent variable [12]. Independent variables or independent variables are variables that affect other variables, while dependent variables or dependent variables are variables that are affected. Simple linear regression is used to determine the effect of one independent variable on one dependent variable [13]. The general equation of simple linear regression is as follows [14]:

$Y = a + bX$.....1

Description:

Y: Non-independent variable (predicted value)

X: Independent variable

a: Constant

b: Regression coefficient (increase or decrease value)

The following formula is used to determine the values of a, and b, namely:

$a = \frac{\sum Yi - b \sum xi}{n}$.....2

$b = \frac{n \sum xy - \sum x . \sum y}{n \sum x^2 - (\sum x)^2}$.....3

According to [15] the steps for calculating linear regression are as follows:

- Identify the dependent (y) and independent (x) variables.
- Prepare the data that will be used as x and y variables.
- Calculate the value of and their respective totals.
- Calculate the values of a and b using formulas (2) and (3).
- Form a model for the linear regression equation using formula (1).

To solve the case in this study, the stages of completion are needed, including:

4.1.1 Determine the research data used

Table 1. Research Data

No.	Month and Year	Veil Sales (Strands)
1.	Januari 2022	190
2.	Februari 2022	189
3.	Maret 2022	200
4.	April 2022	50
5.	Mei 2022	45
6.	Juni 2022	78
7.	Juli 2022	10
8.	Agustus 2022	89
9.	September 2022	117
10.	Oktober 2022	178
11.	November 2022	49
12.	Desember 2022	120

4.1.2 Pre-Processing

After collecting data, the next stage is pre-processing the data that has been obtained. This stage includes the implementation of data filtering where the required variables are eliminated and grouping sales of hoods according to the order per month, starting from January 2022 to December 2022.

4.1.3 Calculation of X, Y, XY, XX values

Table 2. Calculation of X, Y, XY, XX values

No.	Month and Year	Veil Sales (Pieces)	X(t)	Y(t)	XY	XX
1.	January 2022	190	1	190	190	1
2.	February 2022	189	2	189	378	4
3.	March 2022	200	3	200	600	9
4.	April 2022	50	4	50	200	16
5.	May 2022	45	5	45	225	25
6.	June 2022	78	6	78	468	36
7.	July 2022	10	7	10	70	49
8.	August 2022	89	8	89	712	64
9.	September 2022	117	9	117	1053	81
10.	October 2022	178	10	178	1780	100
11.	November 2022	49	11	49	539	121
12.	December 2022	120	12	120	1440	144
	Addition					

Description:

- Variable X (t) is the time period (month) used in the study. The number of time periods (months) is 12 using numbers 1 through 12.
- Variable Y is the number of veil sales.
- Variable XY is the multiplication between Variable X and Variable Y.
- Variable XX is the multiplication between Variable X and Variable X.
- Then the summation is done on each of the contents of the variable so that it is obtained, and.

4.1.4 Calculation of a and b

First, we calculate the value of b as follows:

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

With the formula:

$$b = \frac{12 \text{ x } 7655 - 78 \text{ x } 1315}{12 \text{ x } 650 - (78x78)}$$

$$b = \frac{91860 - 102570}{7800 - 6084}$$

$$b = \frac{-10710}{1716}$$

$$b = -6{,}24126$$

Then the value of b is obtained:

From the formula above, b = -6.24126 is obtained. Furthermore, after getting the value of b, we will look for the value of a with the formula below:

With the formula:

$$a = \frac{\sum Yt - b \sum Xt}{n}$$

Then the value of a is obtained:

$$a = \frac{1315 - \left((-6{,}24126) * 78\right)}{12}$$

$$a = \frac{1801{,}81828}{12}$$

$$a = 150{,}1515233$$

From the above calculation, the value of a = 150.1515233, and the value of b = -6.24126.

4.1.5 Obtain Linear Regression Equation

The coefficient a = 150.1515233, and the value of b = -6.24126 that has been obtained is used to obtain a linear regression equation $y = a + bx$. The coefficient obtained produces a simple linear regression equation, namely: Y = 150.1515233 + -6.24126 X, where the value of Variable X is the running time period from January to December 2022. The linear regression equation can be determined in the form of the table below:

Table 3. Regression Equation obtained from the coefficients a and b

No.	Month and Year	X (t)	Y = 150,1515233 + -6,24126 X
1.	January 2022	1	143,910263
2.	February 2022	2	137,669003
3.	March 2022	3	131,427743
4.	April 2022	4	125,186483
5.	May 2022	5	118,945223
6.	June 2022	6	112,703963
7.	July 2022	7	106,462703
8.	August 2022	8	100,221443
9.	September 2022	9	93,9801833
10.	October 2022	10	87,7389233
11.	November 2022	11	81,4976633
12.	December 2022	12	75,2564033

4.1.6 Making Predictions for 2023

To make predictions for the coming year, namely 2023, what will be used is the linear regression equation that has been obtained. This research will predict sales in 2023. The real data used is with the January X(t) variable labeled month 13 to month 24 (year 2023). As below:

Table 4. Results of Sales Prediction Decision for 2023 with Linear Regression

No.	Month and Year	X (t)	Y = 150,1515233 + -6,24126 X
1.	January 2023	13	69,0151433
2.	February 2023	14	62,7738833
3.	March 2023	15	56,5326233
4.	April 2023	16	50,2913633
5.	May 2023	17	44,0501033
6.	June 2023	18	37,8088433
7.	July 2023	19	31,5675833
8.	August 2023	20	25,3263233
9.	September 2023	21	19,0850633
10.	October 2023	22	12,8438033
11.	November 2023	23	6,6025433
12.	December 2023	24	0,3612833

After completing all calculations on the data sample, the results obtained based on tables 3 and 4 can be seen that the results of sales prediction decisions using the linear regression method show a decrease, so it can be stated that the results of sales prediction decisions in 2023 have decreased.

4.2 Menu Display

The result of this research is the formation of a website application, namely the application of the linear regression method in predicting sales of hoods in 2023 which can be seen in the figure below:

4.2.1 Main Menu Display

Figure 2. Main Menu Display

4.2.2 Sales Period Data Form

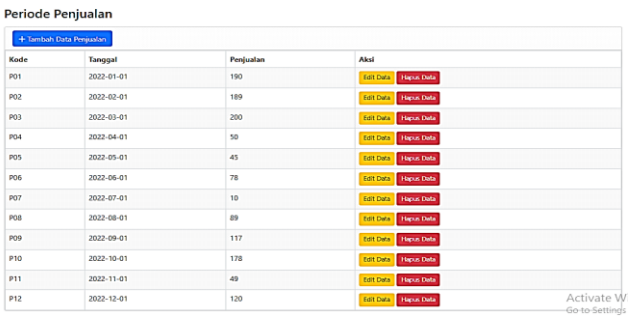
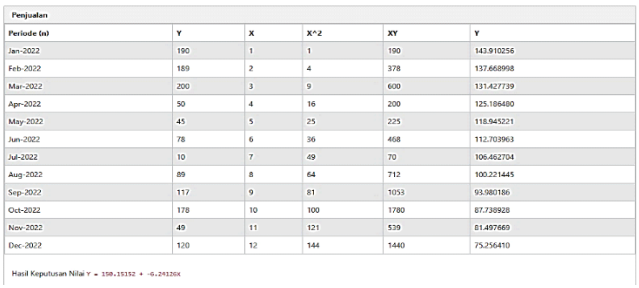


Figure 3. Sales Period Data Form

4.2.3 Linear Regression Calculation Data Form



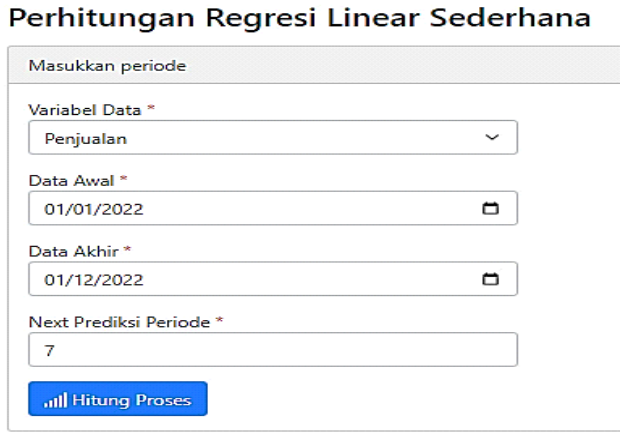


Figure 4. Linear Regression Calculation Data Form

4.2.4 Decision Result Form

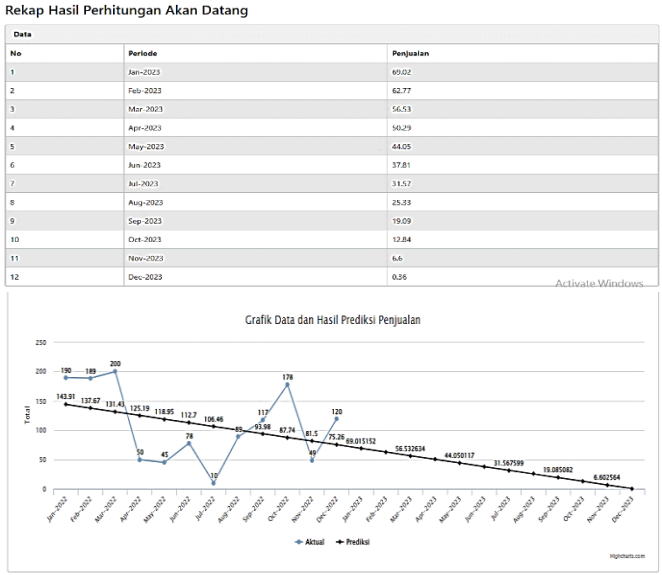


Figure 5. Decision Result Form

5. CONCLUSIONS

a. After completing the design and applying the linear regression method in predicting sales using sample sales data in 2022, the decision result is that the prediction of sales of hoods has decreased. This statement is obtained from the sales results in January 2022 with a value of 143.91 and January 2023 with a value of 69.01 which has been tested using a web-based application. So, it can be stated that the sales of hoods at Fauzan Shop in 2023 have decreased.

b. In this case, the application made is in accordance with the predetermined design. The system built can be used in helping to predict the sale of hoods at Fauzan Veil Shop.

REFERENCES

[1] I. Mialiawati, "The Effect of Globalization on Culture, Information Technology, and Education," in *PROCEEDING ICTESS (Internasional Conference on Technology, Education and Social Sciences)*, 2020.

[2] A. Sheikh *et al.*, "Health information technology and digital innovation for national learning health and care systems," *Lancet Digit. Heal.*, vol. 3, no. 6, pp. e383–e396, 2021.

[3] O. Wisesa, A. Adriansyah, and O. I. Khalaf, "Prediction analysis for business to business (B2B) sales of telecommunication services using machine learning techniques," *Majlesi J. Electr. Eng.*, vol. 14, no. 4, pp. 145–153, 2020.

[4] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to linear regression analysis*. John Wiley & Sons, 2021.

[5] W. N. Putri, M. H. P. Swari, and R. Mumpuni, "PENERAPAN METODE REGRESI LINEAR UNTUK PREDIKSI PENJUALAN SUKU CADANG," *J. Inform. Teknol. dan Sains*, vol. 5, no. 4, pp. 679–685, 2023.

[6] Harsiti, Z. Muttaqin, and E. Srihartini, "Penerapan Metode Regresi Linier Sederhana Untuk Prediksi Persediaan Obat Jenis Tablet," *JSiI (Jurnal Sist. Informasi)*, vol. 9, no. 1, pp. 12–16, 2022, doi: 10.30656/jsii.v9i1.4426.

[7] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono, and C. Lauw, "Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier," *J. Nas. Teknol. dan Sist. Inf.*, vol. 7, no. 1, pp. 8–17, 2021.

[8] W. A. L. Sinaga, S. Sumarno, and I. P. Sari, "The Application of Multiple Linear Regression Method for Population Estimation Gunung Malela District," *JOMLAI J. Mach. Learn. Artif. Intell.*, vol. 1, no. 1, pp. 55–64, 2022, doi: 10.55123/jomlai.v1i1.143.

[9] A. A. Muhartini, O. Sahroni, S. D. Rahmawati, T. Febrianti, and I. Mahuda, "Analisis Peramalan Jumlah Penerimaan Mahasiswa Baru Dengan Menggunakan Metode Regresi Linear Sederhana," *J. Bayesian J. Ilm. Stat. dan Ekon.*, vol. 1, no. 1, pp. 17–23, 2021.

[10] M. Sholeh, "Penerapan Regresi Linear Ganda Untuk Memprediksi Hasil Nilai Kuesioner Mahasiswa Dengan Menggunakan Python," *J. Din. Inform.*, vol. 11, no. 1, pp. 13–24, 2022.

[11] D. Alita, A. D. Putra, and D. Darwis, "Analysis of classic assumption test and multiple linear regression coefficient test for employee structural office recommendation," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 15, no. 3, pp. 295–306, 2021.

[12] A. Gelman, J. Hill, and A. Vehtari, *Regression and other stories*. Cambridge University Press, 2021.

[13] S. Adiguno, Y. Syahra, and M. Yetri, "Prediksi Peningkatan Omset Penjualan Menggunakan Metode Regresi Linier Berganda," *J. Sist. Inf. Triguna Dharma (JURSI TGD)*, vol. 1, no. 4, p. 275, 2022, doi: 10.53513/jursi.v1i4.5331.

[14] D. Maulud and A. M. Abdulazeez, "A review on linear regression comprehensive in machine learning," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 2, pp. 140–147, 2020.

[15] N. L. W. Arya Della, R. A. N. Diaz, and K. D. P. Novianti, "Penerapan Metode Regresi Linier untuk Memprediksi Permohonan ITAS," *J. Eksplora Inform.*, vol. 10, no. 2, pp. 92–100, 2021, doi: 10.30864/eksplora.v10i2.380.