International Journal Of Health, Engineering And Technology (IJHET) *E-ISSN 2829 - 8683*
Volume 5, Number 2, July 2026, *Page. 2596 - 2601*
Email : editorijhess@gmail.com

Web-Based Prediction Of Routine And Seasonal Office Supply Demand Using Linear Regression

Winky Eprilianto[1)], Iis Aisyah[2)]
[1,2)] Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Universitas Pamulang, Tangerang Selatan, Indonesia

*Corresponding Author
E-mail: winkyeprilianto@gmail.com

Abstract

Inventory management of office stationery (ATK) is crucial for maintaining operational continuity and sales performance. CV Sumber Nyata still relies on manual records and past experience to plan product demand, which often leads to overstock or stockouts. This study aims to develop a web-based prediction system for routine and seasonal ATK product needs that can assist the company in determining stock requirements more systematically based on historical sales data. The research employs a quantitative method with a predictive approach using the Linear Regression algorithm, utilizing historical sales data from March 2025 to March 2026. The system was built using PHP and MySQL, and prediction accuracy was evaluated using Mean Absolute Error (MAE) and Symmetric Mean Absolute Percentage Error (SMAPE). The results show that the system successfully generates product demand predictions based on available sales data and presents the information informatively through a web interface. Black Box and White Box testing confirm that all system functions operate as required. Furthermore, user response testing from 10 respondents yielded a satisfaction rate of 89%, categorized as very good. In conclusion, the developed system can assist companies in managing ATK inventory and support decision-making for future product demand planning.

Keywords: Product Demand Prediction, Office Stationery, Linear Regression, MAE, SMAPE

INTRODUCTION

The development of information technology has driven significant changes in various business sectors, including data management and decision-making processes in the trading sector. The utilization of information technology enables companies to process data more effectively, which can be used as a basis for business analysis and sales strategy planning. In the current digital era, sales data analysis has become one of the important approaches used by companies to improve operational efficiency and optimize inventory management (Ardhanur et al., 2025). Various studies indicate that the main problem in inventory and sales management lies in the inability of business actors to predict product needs accurately. Errors in determining stock quantities often result in excess inventory that ties up capital or stock shortages that lead to lost sales opportunities. Therefore, a prediction approach is needed that can utilize historical sales data as a basis for more objective and measurable decision-making (Fatman et al., 2023).

CV Sumber Nyata, as one of the office stationery (ATK) distributors, still plans its inventory based on manual records and past experience, without optimal utilization of historical sales data. This condition potentially leads to discrepancies between stock quantities and market demand, especially in dealing with the different characteristics of routine and seasonal products. Routine products have relatively stable demand patterns, while seasonal products experience increased demand during certain periods, such as brown envelopes and map folders which typically surge towards the end of the year. The inability to distinguish the characteristics of these two product types causes the company to struggle in determining the right stock quantity, thus risking either excess stock that ties up capital or stock shortages that reduce sales potential.

Linear regression has become one of the most widely used statistical methods in sales and inventory prediction research due to its ability to model the relationship between independent and dependent variables based on historical data. Research by (Handoko et al., 2025) showed that the linear regression algorithm can predict the number of tourism visitors with an R-Squared value of 0.9346, meaning the model can explain 93.46% of the data variation. Meanwhile, research by

(Hurifiani et al., 2024) applied linear regression to predict ATK sales at BUMDes and produced a model with a reasonably good RMSE value of 7802.802. Other research by (Ardhanur et al., 2025) also successfully implemented linear regression for a brownies sales prediction model with a Relative Error of 5.40%. However, these studies have not specifically addressed the differentiation of routine and seasonal product characteristics within a single web-based prediction system, particularly in the context of ATK distribution. Additionally, research by (Anggrawan et al., 2022) successfully applied simple linear regression to predict Unilever product sales with MAPE values ranging from 1% to 10%, indicating that this method is effective for data with relatively stable patterns. Nevertheless, its application is still limited to general consumer products and has not accommodated companies with seasonal product characteristics and high fluctuations.

Based on this gap, the novelty of this research lies in the development of a web-based prediction system that explicitly distinguishes the calculation and presentation of prediction results for routine and seasonal products, utilizing historical sales data at CV Sumber Nyata. This system not only generates prediction values using the Linear Regression method but also includes accuracy evaluation through Mean Absolute Error (MAE) and Symmetric Mean Absolute Percentage Error (SMAPE) for each product, allowing users to assess the reliability of the prediction results. Furthermore, the system is designed with an informative and user-friendly interface, equipped with sales and prediction report features that can be printed or exported, thus supporting more structured monitoring and business decision-making processes.

The objectives of this research are :

1) to develop a prediction system for ATK product needs that can distinguish the characteristics of routine and seasonal products.
2) to analyze and implement ATK inventory planning that no longer relies on manual records and past experience by utilizing systematic data processing.
3) to improve ATK product inventory management to reduce the risk of overstock and stockouts, thereby helping to maintain sales stability and company cash flow.
4) to optimally utilize historical sales data as a basis for predicting ATK product needs.

to develop a web-based application that can present ATK product need predictions informatively and easily understood by users.

RESEARCH METHODS

This research employs a quantitative method with a predictive approach. The main objective is to analyze historical sales data of office stationery (ATK) products and generate demand predictions for future periods. Overall, the research process consists of four main stages: data collection, data preprocessing, model building using Linear Regression, and model evaluation using MAE and SMAPE. Subsequently, the system was implemented into a web-based application to facilitate data management and visualization of prediction results.

Research Design and Data Collection

This research was conducted at CV Sumber Nyata, an office stationery distributor company. The dataset used is historical sales data from March 2025 to March 2026, covering 13 monthly periods. The data includes product names, sales periods (months), sales quantities, and product categories (routine or seasonal). Routine products have relatively stable demand patterns, while seasonal products show significant fluctuations during certain periods, such as envelopes and maps which usually increase towards the end of the year. This classification is important as it allows the system to present predictions with appropriate contextual notes for users.

Pre-processing Data

Before modeling, the raw sales data underwent a preprocessing stage to ensure data quality and consistency. The preprocessing steps included:

(1) data cleaning to remove incomplete or duplicate records.

(2) format standardization to ensure uniform date formats.

(3) data aggregation by converting daily sales into monthly totals for each product,

(4) product classification by labeling each product as "Routine" or "Seasonal" based on historical fluctuation patterns (Daniswara & Nuryana, 2023) These steps were performed to improve data quality, ensuring that the subsequent modeling stage produces reliable and accurate predictions.

Linear Regression Model

The prediction model was built using Simple Linear Regression. This method was chosen due to its simplicity, interpretability, and effectiveness in modeling the linear relationship between time as the independent variable and sales quantity as the dependent variable (Duran et al., 2024).In this study, the independent variable (X) is the time period (month sequence), and the dependent variable (Y) is the sales quantity. The linear regression equation is defined as:

$$Y = a + bX$$

Where:

Y = predicted sales value

a = intercept (constant)

b = regression coefficient (slope)

X = time period

The values of **a** and **b** are calculated using the least squares method based on historical data, with the following formulas:

$$a = \frac{(\Sigma Y)(\Sigma X^2) - (\Sigma X)(\Sigma XY)}{n(\Sigma X^2) - (\Sigma X)^2}$$

$$b = \frac{n(\Sigma XY) - (\Sigma X)(\Sigma Y)}{n(\Sigma X^2) - (\Sigma X)^2}$$

Where **n** is the number of data points. This model was applied separately to each product, generating specific regression equations for each product. The resulting equation was then used to predict sales for the subsequent period (period 14).

Model Evaluation Metrics

To measure the accuracy of the prediction model, two evaluation metrics were used: Mean Absolute Error (MAE) and Symmetric Mean Absolute Percentage Error (SMAPE).

MAE measures the average magnitude of errors between actual and predicted values, calculated using the formula (Anggrawan et al., 2022):

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i$$

Where:

Y = the actual value

$\hat{Y}$ = the predicted value

n = the number of data points.

SMAPE is used to measure relative error in percentage form and is more stable when handling data containing zero or near-zero values. SMAPE is calculated using the formula (Ahaggach et al., 2024):

$$\text{SMAPE} = \frac{1}{n}\sum_{i=1}^{n} \frac{|y_i - \widehat{y_i}|}{(|y_i| + |\hat{y}_i)/2} X\ 100\%$$

Where :

Yi = actual value at data point i

$\hat{Y}$i = predicted value at data point i

n = number of data points used in testing

Both metrics were calculated for each product to assess model performance. The lower the MAE and SMAPE values, the higher the prediction accuracy.

System Implementation

The prediction model was then integrated into a web-based application to facilitate data input, processing, and visualization. The system was developed using the PHP programming language (version 8.0.30) with a MySQL database as the data storage, running on an Apache web server (XAMPP environment). The application architecture follows a standard client-server model, where users interact with the system through a web browser. The system has three user roles: Admin (full access), Employee (sales input access), and Owner (report viewing access). The user interface is designed to be informative and easy to understand, displaying prediction results in tables and graphs, complete with MAE and SMAPE values to help users assess the reliability of the predictions.

RESULTS AND DISCUSSION

The Linear Regression model was applied to 10 office stationery (ATK) products consisting of routine and seasonal products. Based on historical sales data from March 2025 to March 2026, the system generated regression equations for each product, which were then used to predict product demand for the subsequent period (period 14). The prediction results are presented in Table 1.

Table 1. Prediction Results and Evaluation of the Linear Regression Model on 10 ATK Products

No	Product Name	Category	Regression Equation	Prediction (Units)	MAE	SMAPE (%)
1	Brown Envelope Plain	Seasonal	Y = 7.38 + 12.70X	185	100.34	126.87
2	Small Notebook 38 Sheets	Routine	Y = -2.15 + 1.41X	18	8.09	102.59
3	Folio Lined Paper	Seasonal	Y = 4.08 + 1.53X	26	24.10	132.60
4	HVS A4 75 Gsm Paper	Routine	Y = 5.58 - 0.24X	2	4.38	79.53
5	HVS F4 75 Gsm Paper	Routine	Y = -3.77 + 1.70X	20	9.61	102.42
6	Map Business	Seasonal	Y = 2.46 + 0.77X	13	12.94	140.15
7	Stamp Duty 10k	Routine	Y = 15.73 + 1.13X	32	10.78	48.18
8	Joyko Retractable Pen	Routine	Y = 3.85 + 0.82X	15	4.55	50.32
9	Kenko Black Pen	Routine	Y = -1.23 + 2.01X	27	11.62	81.25

Source: CV Sumber Nyata Sales Data (2026), processed.

Based on Table 1, the prediction results show that seasonal products such as Brown Envelope Plain have a very high predicted value (185 units), driven by significant sales spikes in the final period (February 2026). In contrast, routine products such as Ring Kacu and Joyko Retractable Pen have more stable predictions that are closer to their historical sales averages.

Model evaluation was conducted by comparing predicted values against actual values using MAE and SMAPE. Products with stable sales patterns such as Joyko Retractable Pen, Ring Kacu, and HVS A4 75 Gsm Paper showed relatively low error rates (MAE < 7, SMAPE < 80%). For example, Joyko Retractable Pen has an MAE of 4.55 and a SMAPE of 50.32%. This indicates that the model can follow relatively stable data patterns quite well. Similarly, Ring Kacu has an MAE of 6.26 and a SMAPE of 42.51%, which is the lowest SMAPE among all products. This indicates that Linear Regression works optimally on data with less extreme fluctuations, which aligns with the findings of Anggrawan et al. (2022) who successfully applied Linear Regression with MAPE values between 1% and 10% on products with stable patterns.

Conversely, seasonal products such as Brown Envelope Plain, Folio Lined Paper, and Map Business showed very high error rates. Map Business has an MAE of 12.94 and a SMAPE of 140.15%, which is the highest error among all products. This high error is caused by highly fluctuating data patterns, with many periods having zero values and extreme spikes in December 2025 (65 units) and June 2025 (17 units). This condition causes the regression line to be unable to capture these spikes accurately. This finding is consistent with research by Tiurma and Samosir (2025) which showed that Linear Regression produces higher errors on products with unstable data fluctuations. This is a major

limitation of the simple Linear Regression method, namely its inability to capture non-linear data patterns with extreme spikes.

System testing was conducted through three approaches: Black Box Testing, White Box Testing, and User Response (questionnaire). Black Box Testing was performed on all system functions (login, dashboard, product data, sales, prediction, reports, user management, and logout). Based on the test results, all tested functions were declared successful and ran according to the specified scenarios. No functional failures were found in the system for admin, employee, or owner roles. White Box Testing used the basis path testing technique by calculating Cyclomatic Complexity (V(G)) on 20 main function modules. All modules had V(G) values between 1 and 3, indicating that the program logic complexity is still in the simple and maintainable category. All identified independent paths were successfully tested and produced outputs that matched the specifications.

User response testing was conducted on 10 respondents consisting of admins, employees, and owners. The questionnaire consisted of 10 questions using a Likert scale of 1–5. The calculation results showed a total score of 448 out of a maximum score of 500, resulting in a user satisfaction percentage of 89.6%. Based on the feasibility category, this figure falls into the Very Good / Very Satisfied category.

Although the Linear Regression method has limitations on fluctuating data, the developed system still provides significant added value for CV Sumber Nyata. With this system, the company no longer fully relies on experience and manual records, but can use historical data as a basis for more measurable decision-making. The evaluation features (MAE and SMAPE) included in the prediction report help users assess the reliability of the prediction results, so restocking decisions can be made with more mature consideration. The system testing results, which show that all functions run well and the very high user satisfaction rate, indicate that this system is feasible to be implemented and can support more efficient ATK product inventory management.

CONCLUSION

This research successfully developed a web-based prediction system for office stationery (ATK) product needs using the Linear Regression method, capable of distinguishing routine and seasonal products based on historical sales data at CV Sumber Nyata. This system enables more structured and data-driven inventory planning, no longer relying solely on subjective experience. The evaluation results show that the Linear Regression method works quite accurately on products with stable sales patterns, but is less optimal on fluctuating seasonal products. Nevertheless, the system includes MAE and SMAPE values as indicators of prediction reliability, enabling the company to minimize the risk of overstock or stockouts that impact cash flow. Based on Black Box and White Box testing, all system functions operate according to specifications, and the user satisfaction rate reached 89.6% in the very good category, confirming the system's feasibility to be implemented as a more efficient ATK inventory management tool. For future development, it is recommended to use alternative prediction methods such as Random Forest or Double Exponential Smoothing to improve accuracy on products with high fluctuations, as well as to add automatic stock notification features and integration with real-time cashier systems.

REFERENCES

Ahaggach, H., Abrouk, L., & Lebon, E. (2024). Systematic Mapping Study of Sales Forecasting: Methods, Trends, and Future Directions. *Forecasting*, *6*(3), 502–532. https://doi.org/10.3390/forecast6030028

Anggrawan, A., Hairani, H., & Azmi, N. (2022). Prediksi Penjualan Produk Unilever Menggunakan Metode Regresi Linear. *Jurnal Bumigora Information Technology (BITe)*, *4*(2), 123–132. https://doi.org/10.30812/bite.v4i2.2416

Ardhanur, I., Martanto, Dikananda, A. R., & Mulyawan. (2025). IMPLEMENTASI ALGORITMA REGRESI LINEAR UNTUK MODEL PREDIKSI PENJUALAN DI TOKO AMANDA BROWNIES. *Jurnal Informatika Dan Teknik Elektro Terapan*, *13*(2), 836–845. https://doi.org/10.23960/jitet.v13i2.6337

Daniswara, A. A. A., & Nuryana, K. D. (2023). Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru. *Journal of Informatics and Computer Science*, *05*, 97.

Duran, P. A., Vitianingsih, A. V., Riza, Moch. S., Maukar, A. L., & Wati, S. F. A. (2024). Data Mining Untuk Prediksi Penjualan Menggunakan Metode Simple Linear Regression. *Teknika*, *13*(1), 27–34. https://doi.org/10.34148/teknika.v13i1.712

Fatman, Y., Majapahit, S. A., & Ramadhan, M. (2023). A Linear Regression Model for Deploying a Cognitive Web for an Inventory Prediction System. *Journal of Intelligent Computing & Health Informatics*, *4*(1), 1–8. https://doi.org/10.26714/jichi.v4i1.11258

Handoko, Y., Satria, E., & Siti Nursifa, F. (2025). Prediksi Jumlah Pengunjung Pariwisata di Kabupaten Garut Menggunakan Algoritma Regresi Linear. *Jurnal Algoritma*, *22*(1), 569–581. https://doi.org/10.33364/algoritma/v.22-1.1807

Hurifiani, A., Purnamasari, A. I., & Ali, I. (2024). PENERAPAN ALGORITMA REGRESI LINEAR UNTUK PREDIKSI PENJUALAN ALAT TULIS KANTOR (ATK) DI BUMDES. In *Jurnal Mahasiswa Teknik Informatika* (Vol. 8, Number 1). https://www.kaggle.com/.