Jurnal Galaksi (Global Knowledge, Artificial Inteligent and Information System) E-ISSN: 3048-2399
Vol 3, No 2, August 2026 , pp. 49-58
https://doi.org/10.70103/galaksi.v3i2.140

Retail Sales Forecasting Using ARIMA Based on Point-of-Sale Transaction Data

I Gede Sudiantara[1*], I Putu Noven Hartawan[2], I Kayan Herdiana[3], I Made Dwi Putra Asana[4]

[1*,2,3,4] Institut Bisnis dan Teknologi Indonesia, Denpasar, Bali, Indonesia
[1*]sudiantara@instiki.ac.id, [2]novenhartawan@instiki.ac.id, [3]herdikayan@instiki.ac.id, [4]dwiputraasana@instiki.ac.id

ARTICLE TEXT

Article history:
Received 28 June 2026
Revised 15 July 2026
Accepted 15 August 2026
Available Online 31 August 2026

Keywords:
ARIMA;
Sales Forecasting;
Point-of-sale;
Time Series;
Retail;

ABSTRACT

Retail businesses generate large volumes of Point-of-Sale (POS) transaction data that can be utilized for sales forecasting and inventory planning. However, transaction-level data need to be transformed into an appropriate time-series structure before forecasting can be performed. This study aims to apply the Autoregressive Integrated Moving Average (ARIMA) method to historical POS transaction data for retail sales forecasting. The study used 4,443,905 transaction records from a retail store covering the period from 2018 to 2021. The research process adopted the Cross-Industry Standard Process for Data Mining (CRISP-DM), with the analysis focusing on data preparation, ARIMA modeling, and model evaluation. Sales data were aggregated into product-level time series, and the ARIMA model was identified using stationarity analysis, Autocorrelation Function (ACF), and Partial Autocorrelation Function (PACF). The resulting ARIMA(1,1,1) model was applied to five selected products represented by M01–M05. Model evaluation used the Ljung–Box test and Root Mean Square Error (RMSE), with forecasting conducted for January and February 2022. The results show that forecasting errors varied across products, indicating differences in forecasting performance according to their historical sales patterns. The study demonstrates the applicability of ARIMA for product-level retail sales forecasting based on historical POS transaction data.

1. Introduction

The retail industry generates a large amount of transaction data through Point-of-Sale (POS) systems. These data contain historical information about product sales that can be utilized to support business analysis and operational decision-making. One of the important applications of historical sales data is sales forecasting, which can provide estimates of future product demand and support inventory planning (Fildes et al., 2022). In retail businesses, effective inventory management is important because the availability of products is directly related to business operations and customer service (Vaz et al., 2024).

Retail stores commonly face fluctuations in customer demand. Changes in demand can create an imbalance between available inventory and actual customer needs. When demand increases unexpectedly, insufficient inventory may result in stockouts and missed sales opportunities. Conversely, excessive inventory during periods of lower demand can increase storage costs and reduce the efficiency of working capital. Therefore, the use of historical sales data to estimate future demand can help retail businesses plan inventory and procurement more systematically.

Time-series forecasting is one approach that can be used to analyze historical sales patterns. In retail environments, forecasting is commonly required at different levels of aggregation, including product-level forecasting for operational decisions (Fildes et al., 2022). The Autoregressive Integrated Moving Average (ARIMA) method is a statistical approach that has been widely applied to time-series forecasting, including sales and demand forecasting. ARIMA can be applied by identifying the characteristics of historical observations, performing differencing when required to obtain stationarity, and determining appropriate model parameters (Kontopoulou et al., 2023). Previous studies have reported the application of ARIMA for forecasting product sales, demand, and other time-series data, indicating its applicability to short-term forecasting problems (Paç & Yakut, 2024; Almaraj & Al Ghamdi, 2024).

Several previous studies have applied ARIMA to sales forecasting in different business contexts. For example, applied ARIMA and SARIMA to forecast nutrition product sales and identified ARIMA as the better-performing model in their study based on the lowest Mean Square Error (MSE) (Hermawan & Suseno, 2022). Safitri (2021) applied ARIMA to forecast dodol salak sales and evaluated the resulting model using RMSE, MAE, and MAPE (Safitri, 2021). Farosanti et al. (2022) also applied ARIMA to forecast the sales of medical and laboratory equipment (Farosanti et al., 2022). These studies indicate that ARIMA can be used for forecasting problems with different sales characteristics.

However, forecasting sales in a retail environment requires not only the selection of an appropriate forecasting model but also the transformation of transaction-level data into a suitable time-series representation. In addition, the forecasting results need to be presented in a form that can be utilized by retail businesses for inventory planning. The research underlying this article therefore focuses on applying ARIMA to historical POS transaction data from a retail store and evaluating its forecasting performance at the product level. The research adopts the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework, covering business understanding, data understanding, data preparation, modeling, evaluation, and deployment (Schröer et al., 2021).

Based on this background, this study aims to apply the ARIMA method to historical POS transaction data for retail sales forecasting. The study focuses on identifying an appropriate ARIMA model according to the characteristics of the sales time series and evaluating its forecasting performance using Root Mean Square Error (RMSE). The forecasting process is applied to five selected products represented by M01–M05.

2. Literatur Review

Sales forecasting using time-series methods has been widely applied in different business and industrial contexts. Previous studies have shown that historical sales and demand data can be transformed into forecasting information to support inventory management, production planning, and business decision-making (Aminullah et al., 2024; Kurlillah et al., 2024; Hasanudin & Mujiyono, 2026). Among the statistical forecasting methods, ARIMA has been applied to various types of time-series data with different characteristics (Kontopoulou et al., 2023).

Hermawan and Suseno (2022) compared ARIMA and Seasonal Autoregressive Integrated Moving Average (SARIMA) for forecasting nutrition product sales at PT Sapto Bumi Hidroponik. Their study identified ARIMA (2,3,0) as the selected model based on the lowest Mean Square Error (MSE). The study indicates that ARIMA can be applied to sales data when a strong seasonal pattern is not dominant (Hermawan & Suseno, 2022).

Safitri (2021) applied ARIMA to forecast dodol salak sales using historical sales data from 2016 to 2021. The study obtained ARIMA (1,0,0) as the selected model and evaluated the forecasting results using RMSE, MAE, and MAPE. The results demonstrated the application of ARIMA as a forecasting approach for supporting production and inventory planning (Safitri, 2021).

Farosanti et al. (2022) applied ARIMA to forecast medical and laboratory equipment sales at PT Tristania Global Indonesia. Their research utilized historical transaction data that had previously functioned mainly as business records and transformed the data into information for sales forecasting. The resulting ARIMA model produced forecasting errors considered sufficiently low for supporting decision-making (Farosanti et al., 2022).

Fitri et al. (2021) applied ARIMA to forecast rice prices in traditional markets in Indonesia. The study identified ARIMA (1,0,1) as the selected model for forecasting rice price changes. Although the research object was different from retail sales, the study demonstrates that ARIMA can be applied to various types of time-series data (Fitri et al., 2021).

Novyta and Alhazami (2022) investigated Nata de Coco demand forecasting in the context of supply chain management using ARIMA. Their study obtained ARIMA (0,2,1) based on the lowest MSE and used the model to forecast product demand. The study shows the relevance of time-series forecasting for supporting inventory and supply chain management (Novyta & Alhazami, 2022).

In addition to ARIMA-based studies, Winurputra and Ratnawati (2025) developed a sales forecasting system using Extreme Gradient Boosting (XGBoost) and the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework. Their study demonstrates the use of CRISP-DM as a systematic framework for processing sales data from business understanding through model development and implementation (Winurputra & Ratnawati, 2025). Although a different forecasting method was used, the study provides a methodological reference for applying CRISP-DM in sales forecasting system development.

Based on these previous studies, ARIMA has been applied to sales, demand, price, and other time-series forecasting problems. Retail forecasting research also emphasizes the importance of forecasting at the product and store levels to support operational and planning decisions (Fildes et al., 2022). The selected ARIMA parameters vary according to the characteristics of each dataset, indicating that model identification needs to consider the historical pattern of the data. Previous studies also demonstrate the use of forecasting results for inventory, production, supply chain, and business decision-making (Almaraj & Al Ghamdi, 2024; Paç & Yakut, 2024; Rachma P. et al., 2023). In contrast, this study applies ARIMA to a large-scale POS transaction dataset from a retail store and examines its forecasting performance for selected products. The study combines transaction-level data processing, time-series forecasting, residual diagnostics, and RMSE-based evaluation within a retail context.

3. Research Method

3.1. Research Framework

This study adopted the Cross-Industry Standard Process for Data Mining (CRISP-DM) as the research framework. CRISP-DM provides a systematic process for processing historical transaction data into information that can support business decision-making (Schröer et al., 2021; Martínez-Plumed et al., 2021). In this study, the framework was applied to organize the process of developing a retail sales forecasting model using the ARIMA method.

The research process consisted of six stages: business understanding, data understanding, data preparation, modeling, evaluation, and deployment (Schröer et al., 2021). The business understanding stage identified the need for sales forecasting to support inventory management. The data understanding stage examined the characteristics and structure of the historical transaction data. The data preparation stage focused on cleaning and transforming transaction records into time-series data. The modeling stage involved identifying and developing the ARIMA model. The evaluation stage assessed the model using residual diagnostics and forecasting error. The deployment stage was defined as the final stage of the CRISP-DM framework, although the analysis presented in this article focuses on the data preparation, modeling, and evaluation stages (Shimaoka et al., 2024).

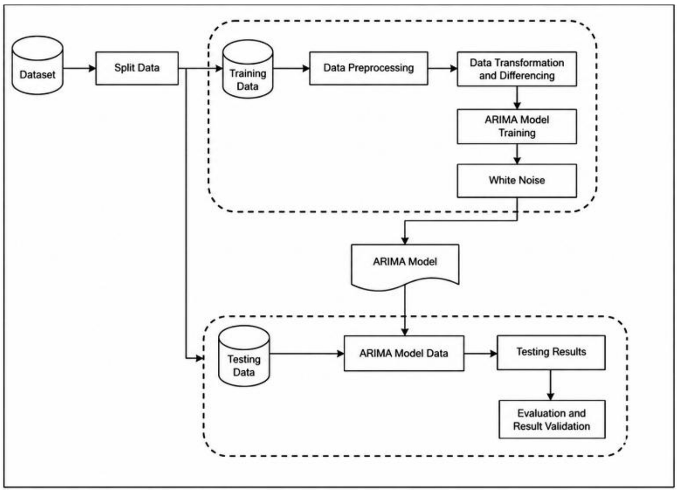

Fig. 1. Research Framework

3.2. Dataset

The dataset consisted of historical Point-of-Sale (POS) transaction records obtained from a retail store in Badung, Bali, Indonesia. The data covered the period from 2018 to 2021 and consisted of 4,443,905 transaction records. Each transaction record contained information including the transaction date, product code, product name, selling price, quantity sold, and transaction value. The quantity sold attribute was used as the main variable for the forecasting process because the objective of the study was to estimate future product demand. The original transaction-level data were subsequently transformed into time-series data by grouping sales based on product and time period. This transformation produced historical sales quantities that could be analyzed using the ARIMA method.

3.3. Data Preparation

Data preparation was conducted to transform the original transaction records into data suitable for time-series forecasting. The process consisted of data inspection, cleaning, outlier handling, aggregation, time-series transformation, and data splitting. First, the dataset was examined to identify missing values and understand the statistical characteristics of the transaction data. The distribution of sales quantities was also examined to identify observations that could affect the forecasting process.

Outlier analysis was performed using statistical analysis and boxplots. Observations identified during this process were handled during data cleaning. After the cleaning process, transaction records were aggregated based on product and month to obtain monthly sales quantities for each product. The resulting data were then transformed into a time-series structure. The time-series data were divided into training and validation datasets. The training dataset contained 80,854 observations, while the validation dataset contained 8,984 observations. The training data were used for model development, while the validation data were used to evaluate the forecasting results

3.4. ARIMA Modeling

The Autoregressive Integrated Moving Average (ARIMA) method was used to model the historical sales time series. Model identification was conducted by examining the stationarity of the time series and analyzing the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) (Kontopoulou et al., 2023). Differencing was performed when required to obtain a stationary time series. The ACF and PACF were subsequently examined to determine the appropriate autoregressive and moving-average orders.

Based on the model identification process, ARIMA(1,1,1) was selected for the forecasting process. The model was applied to five selected products represented by M01, M02, M03, M04, and M05. These codes represent the products being forecasted and do not represent different model configurations. The ARIMA model is represented by three parameters:

$$ARIMA(p, d, q) \tag{1}$$

where p represents the autoregressive order, d represents the differencing order, and q represents the moving-average order (Almaraj & Al Ghamdi, 2024). The model used in this study was:

$$ARIMA(1,1,1) \tag{2}$$

which consists of one autoregressive term, first-order differencing, and one moving-average term.

3.5. Model Evaluation

Model evaluation was performed through residual diagnostics and forecasting error measurement. The Ljung–Box test was used to examine whether the residuals produced by the ARIMA model behaved as white noise (Kurlillah et al., 2024). A significance level of 0.05 was used as the evaluation criterion. The forecasting accuracy was evaluated using Root Mean Square Error (RMSE). RMSE measures the difference between actual and predicted values, where a smaller error indicates a smaller difference between the forecast and actual observation (Aminullah et al., 2024; Mustapha & Sithole, 2025).

4. Result and Discussions

4.1. Data Preparation Results

The initial dataset consisted of 4,443,905 transaction records collected from the retail store. The dataset contained transaction information such as transaction date, product code, product name, selling price, quantity sold, and transaction value. The data preparation process began with an examination of the dataset structure and missing values. Statistical analysis was then performed to understand the distribution of the sales data. Boxplot analysis was used to identify observations considered outliers in the transaction data.

Fig. 2 shows the initial POS transaction dataset used in the study. The figure illustrates the transaction-level structure of the data, where each record represents a sales transaction containing product and sales-related information. This structure provides the basis for the subsequent data preparation and aggregation processes.

	tgl	kode	nama_barang	harga_jual	jumlah	total
0	1/1/2021	22598	Mentos Roll Limited Edition	26000	1	26000
1	1/1/2021	22581	Mentos Mint Limeted Edition	26000	1	26000
2	1/1/2021	811126	Chocopie Lotte	20000	2	40000
3	1/1/2021	154	Kacang Shanghai Boby 1000	16500	1	16500
4	1/1/2021	5250	Top Triple Choco 16g	19500	1	19500
5	1/1/2021	218	Goriorio 14g	9000	2	18000
6	1/1/2021	1523	Roti Jordan Sisir	8500	2	17000
7	1/1/2021	96004	Sampoerna Mild 16	22100	10	221000
8	1/1/2021	370199	Camel Yellow	23000	10	230000
9	1/1/2021	141417	KANJI 1KG PLASTIK	8000	1	8000

Fig. 2. Transaction Dataset

The completeness of the dataset was then examined by identifying missing values. Fig. 3(a) presents the missing-value analysis and shows the distribution of missing data across the dataset attributes. This analysis was used to determine the data fields that required attention during the preparation process. After examining data completeness, statistical analysis was performed to understand the characteristics of the sales transaction data. As shown in Fig. 3(b), the statistical summary provides information about the distribution and variation of the sales-related variables. This analysis was used as an initial reference for identifying observations that could affect the forecasting process.

```
tgl                0
kode               0
nama_barang    21604
harga_jual         0
jumlah             0
total              0
dtype: int64
```

	jumlah
count	1.292736e+06
mean	3.478009e+00
std	5.035571e+00
min	1.000000e+00
25%	1.000000e+00
50%	1.000000e+00
75%	4.000000e+00
max	4.000000e+02

Fig. 3. (a) Missing Value Analysis, (b) Statistical Analysis of Sales Transaction Data

The identified outliers were handled during the data cleaning process. Fig. 4(a) shows the distribution of the transaction data before the cleaning process and indicates the presence of observations outside the main data distribution. These observations were examined and handled during data cleaning. After the cleaning process, the resulting distribution is presented in Fig. 4(b). Compared with the previous distribution, the figure shows the condition of the sales data after the identified observations were handled. The cleaned data were then used in the subsequent transformation process.

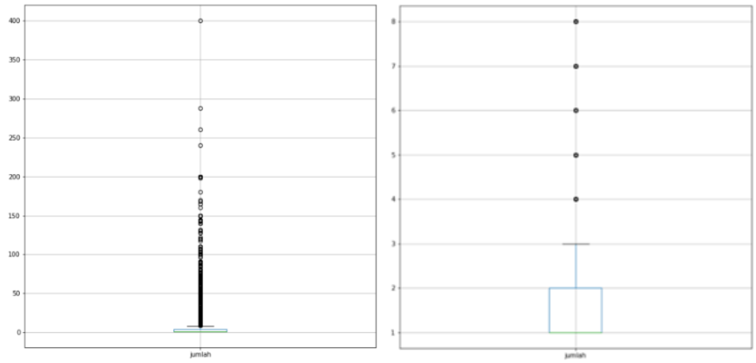

Fig. 4. Boxplot of Sales Transactions (a) Before Data Cleaning (b) After Data Cleaning

The cleaned transaction data were then aggregated based on product and month. Fig. 5 illustrates the resulting monthly sales aggregation. This transformation converted the original transaction-level records into a time-series structure that could be used for ARIMA modeling.

kode	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
LILIN PDK	2	0	3	4	5	6	7	9	0	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
BAA	1	2	3	4	6	0	8	0	9	10	11	12	13	15	0	24	0	0	0	0	0	0	0	0	25	26
BAAA	0	2	3	4	5	6	7	8	9	10	11	12	13	14	17	0	0	21	0	0	0	23	0	0	0	0
BERAS LOKAL	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
BIR	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

Fig. 5. Monthly Sales Aggregation

Finally, the resulting time-series data were divided into training and validation datasets. Fig. 6 presents the division of the data used in the forecasting process. The training dataset contained 80,854 observations, while the validation dataset contained 8,984 observations. The training data were used to develop the forecasting model, while the validation data were used to evaluate its forecasting performance.

```
Train set (80854, 12)
Validation set (8984, 12)
```

Fig. 6. Training and Validation Data Split

Overall, the data preparation process transformed the original large-scale POS transaction records into structured time-series data suitable for ARIMA modeling. The prepared data were subsequently used in the model development and forecasting stages.

4.2. ARIMA Model Development

After the data preparation process, the ARIMA model was developed using the resulting time-series data. Stationarity analysis, differencing, ACF, and PACF analysis were used to identify the appropriate model configuration. Based on the model identification process, ARIMA(1,1,1) was selected for the forecasting process. The model was applied to five selected products represented by M01, M02, M03, M04, and M05. The parameter estimation results for the five products are presented in Table 1.

Table 1. Parameter Estimation Results

Product Name	Product Code	ARIMA Model	Parameter	p-value
Merries Pants XL 1	M01	ARIMA(1,1,1)	AR(1)	0.009
			MA(1)	0.772
Super Zuper Renceng	M02	ARIMA(1,1,1)	AR(1)	0.727
			MA(1)	0.997
Bamboe Kare	M03	ARIMA(1,1,1)	AR(1)	0.861
			MA(1)	0.142
Fancy Telur	M04	ARIMA(1,1,1)	AR(1)	0.960
			MA(1)	0.848
ABC Kecap Manis 135 ml	M05	ARIMA(1,1,1)	AR(1)	0.759
			MA(1)	0.999

The parameter estimation results show that the parameter significance varied among the five products. The results were retained as reported in the original analysis and were not used to introduce a different model configuration. Residual diagnostics were subsequently performed using the Ljung–Box test. The results are presented in Table 2.

Table 2. Ljung–Box Test Results

Product Code	ARIMA Model	p-value	Residual Conditions
M01	ARIMA(1,1,1)	0.33	White Noise
M02	ARIMA(1,1,1)	0.05	Not White Noise
M03	ARIMA(1,1,1)	0.12	White Noise
M04	ARIMA(1,1,1)	0.18	White Noise
M05	ARIMA(1,1,1)	0.05	Not White Noise

The Ljung–Box results show differences in residual characteristics among the five products. M01, M03, and M04 satisfied the white-noise criterion based on the reported p-values, whereas M02 and M05 were reported as not satisfying the criterion. This result indicates that the historical sales behavior differs among products even though the same ARIMA(1,1,1) configuration was used. Fig. 7 illustrates the ARIMA model training process for the selected product data. The output shows the estimated model parameters and diagnostic information used in the model evaluation process.

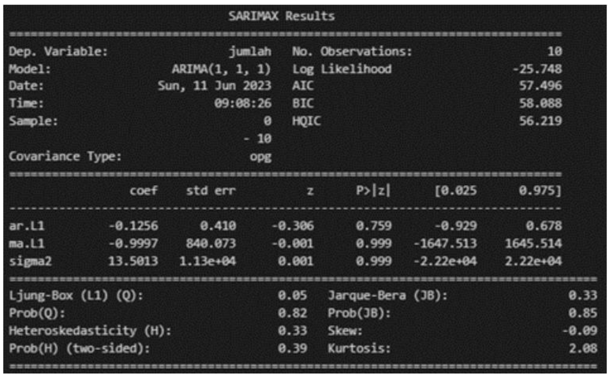
```
                               SARIMAX Results
==============================================================================
Dep. Variable:                 jumlah   No. Observations:                   10
Model:                 ARIMA(1, 1, 1)   Log Likelihood                 -25.748
Date:                Sun, 11 Jun 2023   AIC                             57.496
Time:                        09:08:26   BIC                             58.088
Sample:                             0   HQIC                            56.219
                                 - 10
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.1256      0.410     -0.306      0.759      -0.929       0.678
ma.L1         -0.9997    840.073     -0.001      0.999   -1647.513    1645.514
sigma2        13.5013   1.13e+04      0.001      0.999   -2.22e+04    2.22e+04
===================================================================================
Ljung-Box (L1) (Q):                   0.05   Jarque-Bera (JB):                 0.33
Prob(Q):                              0.82   Prob(JB):                         0.85
Heteroskedasticity (H):               0.33   Skew:                            -0.09
Prob(H) (two-sided):                  0.39   Kurtosis:                         2.08
===================================================================================
```

Fig. 7. ARIMA Model Training

4.3. Sales Forecasting Results

The ARIMA(1,1,1) model was used to generate sales forecasts for five selected products represented by M01, M02, M03, M04, and M05. The forecasting process was conducted for two periods, namely January 2022 and February 2022. For the January 2022 forecasting period, the model used the sales history of the previous 12 months as input. The resulting forecasting input data are presented in Fig. 8. The figure contains the historical monthly sales quantities for the five selected products, while the label column represents the target sales value for the forecasting period.

	Kode	Nama Barang	0	1	2	3	4	5	6	7	8	9	10	11	label
0	5425	Merries Pants X1 L	50	14	38	3	12	11	2	0	0	6	20	12	29
1	2002	Super Zuper Renceng	16	9	27	6	2	9	3	0	0	9	20	9	19
2	110029	Bamboo Kare	4	6	4	8	4	0	4	1	0	1	7	21	21
3	15874	Fancy Telur	17	18	11	20	0	0	18	21	21	12	30	11	13
4	110113	Abc Kecap Manis 135ml	10	8	0	0	1	13	13	9	10	6	11	1	6

Fig. 8. Forecasting Input Data for January 2022

Fig. 8 shows that each product has a different sales pattern during the observed period. For example, M01 has relatively higher sales quantities than several other products, while M05 has comparatively lower sales quantities. These differences provide different time-series patterns for each product in the forecasting process.

Table 3. Forecasting Results for January 2022

Product Code	Actual	Predicted	RMSE
M01	29	24.79910	17.64755
M02	19	13.36934	31.70424
M03	21	3.88237	293.01309
M04	13	29.99694	288.89621
M05	6	9.08042	9.48899

The results show that the prediction error varied among the five products. The predicted value for M01 was relatively close to its actual sales value, while larger differences were observed for M03 and M04. This variation indicates that the forecasting performance differed according to the historical sales pattern of each product.

The forecasting process was subsequently performed for February 2022. In this period, the previous 12 months of sales data were used as the forecasting input. Fig. 9 presents the corresponding forecasting input data, with the monthly columns representing the historical sales quantities and the label column representing the target sales value.

	Kode	Nama Barang	0	1	2	3	4	5	6	7	8	9	10	11	label
0	5425	Merries Pants X1 L	14	38	3	12	11	2	0	0	6	20	12	29	14
1	2002	Super Zuper Renceng	9	27	6	2	9	3	0	0	9	20	9	19	23
2	110029	Bamboo Kare	6	4	8	4	0	4	1	0	1	7	21	21	10
3	15874	Fancy Telur	18	11	20	0	0	18	21	21	12	30	11	13	22
4	110113	Abc Kecap Manis 135ml	8	0	0	1	13	13	9	10	6	11	1	6	7

Fig. 9. Forecasting Input Data for February 2022

Fig. 9 shows that the historical sales patterns also vary among the five products. Some products show relatively stable sales quantities, while others exhibit greater variation across the observed months. These differences affect the resulting forecasting values for each product.

Table 4. Forecasting Results for February 2022

Product Code	Actual	Predicted	RMSE

M01	14	28.19726	201.56213
M02	23	12.41229	112.09965
M03	10	3.66389	40.14633
M04	22	29.43248	55.24174
M05	7	8.56791	2.45834

The February 2022 results also show different forecasting errors among the products. M05 produced a predicted value relatively close to its actual sales value, whereas M01 and M02 showed larger differences between predicted and actual values. These results further indicate that the forecasting performance of ARIMA(1,1,1) varies according to the sales characteristics of individual products.

5. Conclusion

This study applied the ARIMA method to historical Point-of-Sale (POS) transaction data for retail sales forecasting. The dataset consisted of a large number of transaction records that were processed and transformed into product-level time-series data before the forecasting process. Based on the model identification process using stationarity analysis, ACF, and PACF, the ARIMA(1,1,1) model was selected for forecasting the sales of five selected products represented by M01–M05. The model was applied to the historical sales patterns of each product and evaluated using residual diagnostics and RMSE. The forecasting results for January and February 2022 showed that the prediction performance varied among the five products. This indicates that differences in historical sales patterns can affect the forecasting accuracy of the ARIMA model. The forecasting results provide information on estimated future product sales that can be used as supporting information for retail inventory planning. Overall, the study demonstrates the application of ARIMA to POS transaction data for retail sales forecasting. The resulting forecasting process can transform historical transaction data into information that supports inventory-related decision-making

Acknowledgment

The authors thank the Institute of Business and Technology Indonesia, through the Directorate of Research, Community Service, and Innovation, for supporting the 2026 INSTIKI Research and Development Program.

References

Almaraj, I. I., & Al Ghamdi, M. S. (2024). Adopting new ARIMA double layering technique in make-to-stock production policy. Decision Science Letters, 13(2), 315–328. https://doi.org/10.5267/j.dsl.2024.2.005

Aminullah, A. A., et al. (2024). Forecasting sales using SARIMA models at the Sinar Pagi building materials store. JIKO (Jurnal Informatika dan Komputer), 7(2), 71–78. https://doi.org/10.33387/jiko.v7i2.8266

Farosanti, L. (2022). Analisa peramalan penjualan alat kesehatan dan laboratorium di PT. Tristania Global Indonesia menggunakan metode ARIMA. JIMP: Jurnal Informatika Merdeka Pasuruan, 7(1). https://ejurnal.unmerpas.ac.id/index.php/informatika/article/view/428

Fildes, R., Ma, S., & Kolassa, S. (2022). Retail forecasting: Research and practice. International Journal of Forecasting, 38(4), 1283–1318. https://doi.org/10.1016/j.ijforecast.2019.06.004

Fitri, A. E., Nabila, D., Melenia, D., Guntoro Aji, L., Salwa Azmah, N., & Fadilah, S. (2021). Peramalan harga beras pada pasar tradisional di Indonesia dengan menggunakan model ARIMA. DwijenAGRO, 11(1), 12–16. https://doi.org/10.46650/dwijenagro.11.1.1079.12-16

Hasanudin, M., & Mujiyono, S. (2026). Peramalan penjualan mitra konsinyasi menggunakan SARIMA berbasis grid search dan evaluasi akurasi. Jurnal Algoritma, 23(1). https://doi.org/10.33364/algoritma/v.23-1.3147

Hermawan, R., & Suseno. (2022). Analisis peramalan penjualan produk nutrisi dengan metode ARIMA dan SARIMA pada PT Sapto Bumi Hidroponik. JURITI Prima (Jurnal Ilmiah Teknik Industri Prima), 5(2), 17–25. http://jurnal.unprimdn.ac.id/index.php/juriti/article/view/2253

Kontopoulou, V. I., Panagopoulos, A. D., Kakkos, I., & Matsopoulos, G. K. (2023). A review of ARIMA vs. machine learning approaches for time series forecasting in data driven networks. Future Internet, 15(8), 255. https://doi.org/10.3390/fi15080255

Kurlillah, H. A., Anggita, A. T., Agustin, N., & Prahesti, D. (2024). Peramalan permintaan produk beras pandan wangi asli menggunakan metode ARIMA dan SARIMA. JUPRIT, 3(4). https://doi.org/10.55606/juprit.v3i4.4374

Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernández-Orallo, J., Kull, M., Lachiche, N., Ramírez-Quintana, M. J., & Flach, P. (2021). CRISP-DM twenty years later: From data mining processes to data science trajectories. IEEE Transactions on Knowledge and Data Engineering, 33(8), 3048–3061. https://doi.org/10.1109/TKDE.2019.2962680

Mustapha, O. O., & Sithole, T. (2025). Forecasting retail sales using machine learning models. American Journal of Statistics and Actuarial Sciences. https://doi.org/10.47672/ajsas.2679

Novyta, N., & Alhazami, L. (2022). Peramalan permintaan produk nata de coco dalam supply chain management dengan model ARIMA. Jurnal Teknologi dan Manajemen, 7(2). https://doi.org/10.36665/theorems.v7i2.655

Paç, A. B., & Yakut, B. (2024). Assessing the profit impact of ARIMA and neural network demand forecasts in retail inventory replenishment. International Journal of Computational and Experimental Science and Engineering, 10(4). https://doi.org/10.22399/ijcesen.439

Rachma P., S. M., Sholihah, S. A., Poncotoyo, W., Maulana, A., & Jonathan, J. (2023). Sosialisasi penerapan metode ARIMA untuk peramalan permintaan dan MRP untuk perencanaan bahan baku pengemasan minyak goreng. Jurnal Abdi Masyarakat Transportasi & Logistik, 3(1). https://doi.org/10.54324/j.atl.v3i1.1143

Safitri, E. (2021). Analisis forecasting penjualan dodol salak di UD. Salacca menggunakan metode ARIMA [Skripsi, UIN Syekh Ali Hasan Ahmad Addary Padangsidimpuan]. http://etd.uinsyahada.ac.id/7439/

Schröer, C., Kruse, F., & Marx Gómez, J. (2021). A systematic literature review on applying CRISP-DM process model. Procedia Computer Science, 181, 526–534. https://doi.org/10.1016/j.procs.2021.01.199

Shimaoka, A. M., Ferreira, R. C., & Goldman, A. (2024). The evolution of CRISP-DM for data science: Methods, processes and frameworks. SBC Computing Reviews, 4(1), 28–43. https://doi.org/10.5753/reviews.2024.3757

Vaz, C. B., Sena, I., Braga, A. C., Novais, P., & Pereira, A. I. (2024). Predicting retail store transaction patterns: A comparison of ARIMA and machine learning models. In Optimization, Learning Algorithms and Applications (OL2A 2024), CCIS vol. 2280, 268–283. https://doi.org/10.1007/978-3-031-77426-3_18

Winurputra, R., & Ratnawati, D. E. (2025). Peramalan Penjualan Produk Menggunakan Extreme Gradient Boosting (XGBOOST) Dan Kerangka Kerja Crisp-Dm untuk Pengoptimalan Manajemen Persediaan (Studi Kasus: Ub Mart). (2025). Jurnal Teknologi Informasi Dan Ilmu Komputer, 12(2), 417-428. https://doi.org/10.25126/jtiik.2025129451