

# Analisis Prediksi Pola Penjualan pada Kantin Sehat Universitas Majalengka menggunakan Data Mining Metode Time Series Analysis dan Algoritma SARIMA

Aa Herdi Prayoga<sup>1</sup>, Franciskus Antonius Alijoyo<sup>2</sup>

<sup>1,2</sup> School of business and Information Technology, STMIK LIKMI Bandung, 40132 - Indonesia  
e-mail: aaherdi832@gmail.com<sup>\*1</sup>, [Antonius.alijoyo@gmail.com](mailto:Antonius.alijoyo@gmail.com)<sup>\*2</sup>

## INFORMASI ARTIKEL

**Sejarah Artikel:**  
Diterima Redaksi : 16 Agustus 2024  
Revisi Akhir : 05 November 2024  
Diterbitkan Online : 30 November 2024

**Kata Kunci:**  
data mining, *time series analysis analysis*, SARIMA, pola penjualan.

**Korespondensi:**  
Telepon / Hp : +62 (0265) 272727  
E-mail : [email@afiliasi.ac.id](mailto:email@afiliasi.ac.id)

## A B S T R A K

Penerapan data mining memberikan dampak serta kontribusi besar dalam berbagai bidang dan sektor termasuk ritel dan layanan makanan. Data mining dalam konteks analisis telah menjadi salah satu kunci dalam memahami pola dan tren pada sebuah kumpulan data, pada kantin sehat Universitas Majalengka terdapat 406 hingga 686 transaksi perminggu nya, seringkali pada pengelolaan persediaan menjadi kendala mengakibatkan pembengkakan biaya operasional yang pada akhirnya mengurangi profitabilitas kantin. metode *time series analysis analysis* dan algoritma ARIMA dapat digunakan untuk memprediksi nilai bedasarkan data historis dengan tujuan untuk mengoptimalkan pengelolaan persediaan stok berbasis data. Data transaksi penjualan harian secara historis yang diambil dari satu tenant selama periode tertentu digunakan sebagai sampel analisis untuk memprediksi penjualan mingguan dimasa yang akan datang. Hasil prediksi pola penjualan diharapkan dapat dijadikan acuan pengambilan keputusan strategis pada manajemen operasional kantin memberikan rekomendasi pengadaan stok yang lebih efisien, mengurangi biaya operasional, dan meningkatkan profitabilitas kantin.

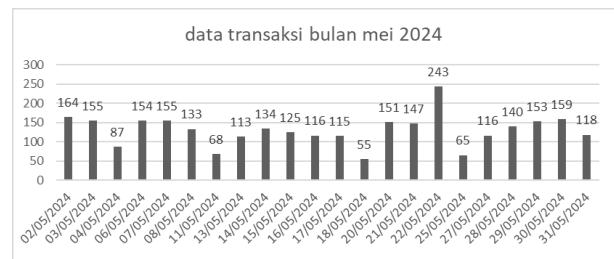
## 1. PENDAHULUAN

Dalam era digital, kemajuan teknologi semakin pesat, terutama dalam bidang data mining yang memberikan dampak besar di berbagai sektor. Secara umum, metode atau jenis data mining yang sering digunakan meliputi *clustering*, *classification*, *association rule mining*, dan *regression analysis*. *Clustering* digunakan untuk mengelompokkan data berdasarkan kemiripan tertentu, *classification* digunakan untuk mengkategorikan data ke dalam kelas-kelas yang telah ditentukan, *association rule mining* digunakan untuk menemukan hubungan antar item dalam suatu dataset, sedangkan *regression analysis* digunakan untuk memprediksi nilai numerik berdasarkan data historis [1].

Penerapan data mining dalam konteks analisis telah menjadi salah satu kunci dalam memahami pola dan tren pada sebuah kumpulan data. Data mining memungkinkan proses pengumpulan dan analisis data secara efektif, yang lebih mendalam mengidentifikasi pola, dan membuat prediksi yang lebih akurat, hasil dari analisis ini dapat digunakan untuk pengambilan keputusan yang lebih baik [2] yang pada akhirnya dapat mengoptimalkan dan mengefisiensikan operasional proses bisnis di berbagai bidang, termasuk sektor ritel dan layanan makanan. Salah satu penerapannya adalah dalam analisis penjualan, di mana data transaksi penjualan secara historis dapat dianalisis untuk memprediksi permintaan di masa depan. dikutip pada penelitian yang diterbitkan dari jurnal International Journal of Information Management Data Insights [3] penerapan data mining dapat membantu dalam manajemen persediaan yang lebih baik, meningkatkan

strategi penjualan, dan mengefisiensikan operasional proses bisnis.

Universitas Majalengka merupakan perguruan tinggi swasta yang terletak di Kabupaten Majalengka, Provinsi Jawa Barat, terdiri dari 7 Fakultas (S1) dan 1 Program Pascasarjana (S2) dengan 23 Program Studi [4]. Sivitas akademika universitas ini terdiri dari mahasiswa dengan total 5282 serta tenaga pendidik dan kependidikan sekitar 345. Salah satu fasilitas di Universitas Majalengka adalah kantin sehat yang terdiri dari 12 tenant yang menyediakan berbagai jenis makanan dan minuman. Mekanisme transaksi di kantin ini setiap tenant beroperasi secara terpusat melalui sistem transaksi di kasir. Berdasarkan hasil pengumpulan data dari pihak manajemen pengelola kantin, terdapat 55-243 transaksi per harinya dan 406-686 transaksi per minggu, dengan total selama Mei 2024 sebanyak 2866 transaksi.



Gambar 1. Data transaksi bulan mei 2024

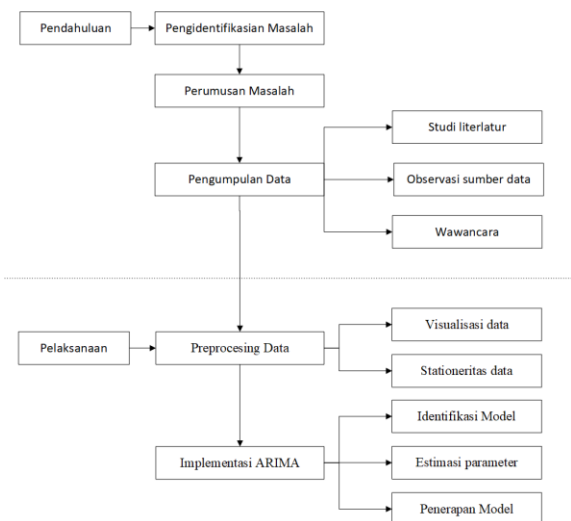
Namun, Kantin Sehat Universitas Majalengka menghadapi beberapa permasalahan terkait manajemen penjualan dan persediaan. Fluktuasi permintaan penjualan harian yang signifikan menyulitkan perencanaan stok persediaan, karena ketidakpastian ini membuat pengelola sulit memprediksi jumlah barang

yang harus disediakan setiap harinya. Pengelolaan persediaan yang kurang efisien juga menjadi tantangan besar karena sering kali mengakibatkan pembengkakan biaya operasional yang pada akhirnya mengurangi profitabilitas kantin. Selain itu, terdapat kebutuhan mendesak untuk pengambilan keputusan yang tepat dengan menggunakan analisis yang dapat memberikan rekomendasi pengadaan stok yang optimal berdasarkan prediksi penjualan.

Melalui penelitian ini, diharapkan informasi yang dihasilkan dapat diterapkan untuk pengambilan keputusan strategis dan manajemen operasional kantin. Penelitian ini mempunyai keterkaitan dengan ekonomi sirkular, dalam tujuan yang akan dicapai nanti yaitu pengelolaan sumber daya yang efisien dengan melakukan analisis prediksi pola penjualan melalui pendekatan data untuk selanjutnya menerapkan strategi keberlanjutan pengelolaan stok yang lebih efisien dan membantu perencanaan kebutuhan bahan. Hal tersebut dapat mengurangi pemborosan bahan makanan, memastikan bahan digunakan secara optimal, meminimalkan limbah yang dihasilkan, dan meningkatkan profitabilitas kantin dengan menerapkan metode *time series analysis* dengan algoritma SARIMA dalam konteks manajemen pengelolaan di sektor jasa makanan dan minuman. Dengan demikian, hasil dari penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam meningkatkan kinerja operasional [5] Kantin Sehat Universitas Majalengka.

**2. METODE PENELITIAN**

Penelitian ini menggunakan metode *time series analysis* dengan algoritma SARIMA untuk menganalisis dan memprediksi pola penjualan pada Kantin Sehat Universitas Majalengka. Langkah-langkah yang dilakukan dalam penelitian ini meliputi tahapan Pengumpulan data (sumber data, observasi lapangan, wawancara), Penerapan Algoritma SARIMA (visualisasi data, stationeritas data, identifikasi model, estimasi parameter, penerapan model) [6]



Gambar 2. Tahap Penelitian

**2.1. Time Series Analysis dan SARIMA**

Analisis *time series analysis* adalah metode yang digunakan untuk menganalisis sekumpulan data yang diambil dari waktu ke waktu untuk mengidentifikasi pola yang mungkin berguna dalam membuat prediksi. Dalam konteks ini, model SARIMA (Seasonal Autoregressive Integrated Moving Average) digunakan. SARIMA adalah model statistik yang memperhitungkan data yang tidak stasioner dengan pola musiman. persamaan model SARIMA dengan komponen musiman adalah:

$$\begin{aligned}
 y_t = & \Phi_1 y_{\{t-s\}} + \Phi_2 y_{\{t-2s\}} + \dots + \Phi_p y_{\{t-ps\}} \\
 & + \Theta_1 \epsilon_{\{t-s\}} + \Theta_2 \epsilon_{\{t-2s\}} \\
 & + \dots + \Theta_q \epsilon_{\{t-qs\}} + \epsilon_t
 \end{aligned}
 \tag{1}$$

Variable dalam persamaan (1) tersebut antara lain :  $y_t$  Nilai pengamatan dari *time series analysis* pada waktu,  $\phi_1, \phi_2 \dots \phi_p$  Parameter *autoregressive* (AR) yang menunjukkan hubungan antara nilai pengamatan saat ini dengan nilai pengamatan sebelumnya,  $y_{\{t-1\}}, y_{\{t-2\}} \dots y_{\{t-p\}}$  Nilai pengamatan pada waktu sebelumnya,  $\theta_1, \theta_2 \dots \theta_p$  Parameter *moving average* (MA) yang menunjukkan hubungan antara nilai pengamatan saat ini dengan kesalahan (error terms) dari nilai pengamatan sebelumnya,  $\epsilon_{\{t-1\}}, \epsilon_{\{t-2\}} \dots \epsilon_{\{t-p\}}$  Error terms pada waktu sebelumnya,  $\epsilon_t$  Error term pada waktu. Komponen dalam model SARIMA meliputi :  $p$  urutan dari komponen *autoregressive* (AR), menunjukkan jumlah lag nilai pengamatan sebelumnya yang dimasukkan dalam model,  $d$  derajat differencing yang diperlukan untuk membuat data stasioner, menunjukkan berapa kali data harus didifferensiasi,  $q$  urutan dari komponen *moving average* (MA) [7] menunjukkan jumlah lag dari error terms sebelumnya yang dimasukkan dalam model,  $P$  urutan dari komponen *autoregressive* musiman (seasonal AR) menunjukkan jumlah lag nilai pengamatan musiman sebelumnya yang dimasukkan dalam model,  $D$  derajat differencing musiman yang diperlukan untuk membuat data stasioner secara musiman,  $Q$  urutan dari komponen *moving average* musiman (seasonal MA) menunjukkan jumlah lag dari error terms musiman sebelumnya yang dimasukkan dalam model.  $s$  periode musiman, menunjukkan panjang siklus musiman (misalnya, 12 untuk data bulanan termasuk pola tahunan, 7 untuk data harian termasuk pola mingguan) [5].

Proses pemodelan dilakukan menggunakan pemrograman Python di Google Colab, yang menyediakan berbagai library seperti pandas, matplotlib, statsmodels, dan scikit-learn untuk memudahkan pengolahan data, visualisasi, pemodelan, dan evaluasi. Dengan pendekatan ini, setiap tahapan analisis mulai dari visualisasi data, uji stasioneritas, identifikasi model, hingga penerapan dan evaluasi model SARIMA dilakukan secara terstruktur dan efisien.

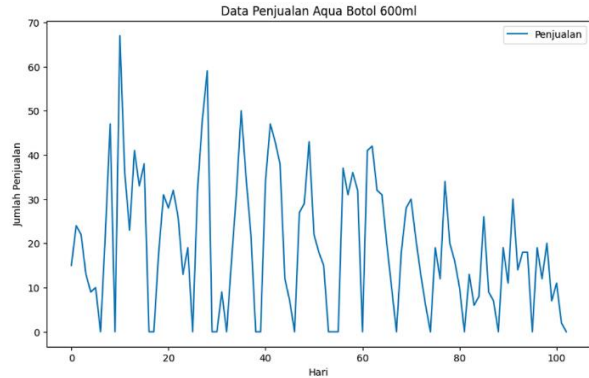
**2.2. Visualisasi Data**

Data yang digunakan dalam penelitian ini berasal dari penjualan harian salah satu menu di salah satu tenant di Kantin Sehat Universitas Majalengka yaitu penjualan aqua botol 600ml, dipilihnya data ini sebagai sampel karena merupakan data yang paling lengkap dibandingkan dengan menu lainnya. Setiap harinya terdapat transaksi untuk menu ini, menjadikannya ideal sebagai sampel dalam penerapan metode analisis *time series analysis*. Konsistensi dan kelengkapan data ini memungkinkan analisis yang lebih akurat dan andal dalam mengidentifikasi pola penjualan harian dan tren musiman selanjutnya dibentuk menjadi dataset. Dataset ini merupakan representatif dari pola penjualan yang terjadi pada menu tersebut selama periode waktu tertentu, secara garis besar diambil dari 4 bulan terakhir april, mei, juli, juni 2024 dari mulai tanggal 22 april 2024 sampai dengan tanggal 4 agustus 2024 dengan kuantitas data sebanyak 104. Dataset yang digunakan terdiri dari 3 kolom yaitu tanggal, nama menu, dan terjual (total penjualan dihari tersebut). Kolom terjual dipilih sebagai variable dependen yang menunjukkan jumlah penjualan setiap hari untuk digunakan sebagai acuan pola.

Table 1. Dataset Awal

Tanggal	Nama Menu	Terjual
22/04/2024	Aqua Botol 600ml	15
23/04/2024	Aqua Botol 600ml	24
24/04/2024	Aqua Botol 600ml	22
25/04/2024	Aqua Botol 600ml	13
26/04/2024	Aqua Botol 600ml	9
27/04/2024	Aqua Botol 600ml	10
28/04/2024	Aqua Botol 600ml	0
.....	.....	.....

Grafik penjualan harian menunjukkan variasi yang signifikan dalam jumlah penjualan dari hari ke hari. Terdapat beberapa hari dengan puncak penjualan yang tinggi, diikuti oleh penurunan yang cukup drastis pada hari-hari tertentu, secara umum memang hari ke 1-6 hari senin-sabtu secara garis besar pasti ada nilai penjualan tetapi pada hari ke 7 yang merupakan hari libur yaitu minggu pasti nilai penjualannya 0, hal ini juga yang menjadi alasan digunakannya algoritma SARIMA karna sesuai dengan karakteristik dataset yang akan diolah. Grafik ini memberikan gambaran awal tentang volatilitas penjualan, yang penting untuk dipertimbangkan dalam model prediksi yang akan digunakan. Dari grafik yang dihasilkan, terlihat bahwa penjualan memiliki ketidakstabilan yang cukup tinggi, yang mungkin disebabkan oleh berbagai faktor. Ketidakstabilan ini menunjukkan pentingnya langkah-langkah selanjutnya dalam analisis, seperti uji stasioneritas, untuk memastikan bahwa model SARIMA yang akan digunakan dapat menangkap pola yang ada dengan akurat[8]



Gambar 3. Grafik Visualisasi Data Awal

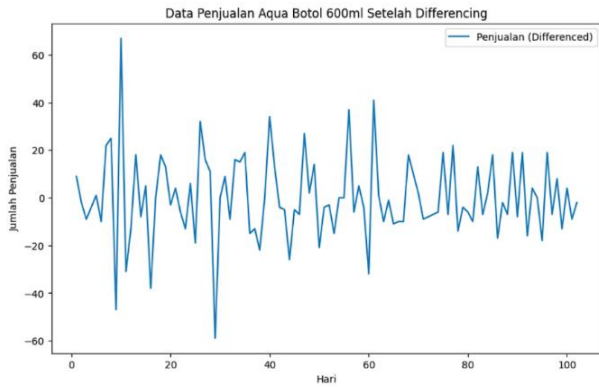
**2.3. Stationeritas Data**

Untuk menggunakan model SARIMA, data harus bersifat stasioner, karena pola musiman dan tren hanya dapat diidentifikasi secara akurat jika karakteristik statistik data, seperti mean dan varians, konstan sepanjang waktu. Uji stasioneritas dilakukan menggunakan Augmented Dickey-Fuller (ADF) test [9]. Persamaan ADF test adalah:

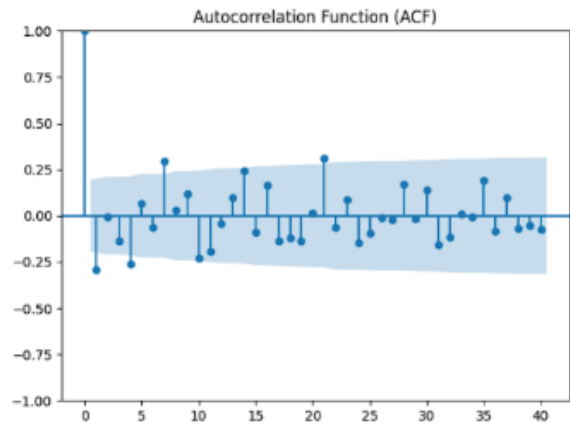
$$\Delta y_t = \alpha + \beta t + \gamma y_{\{t-1\}} + \delta_1 \Delta y_{\{t-1\}} + \delta_2 \Delta y_{\{t-2\}} + \dots + \delta_p \Delta y_{\{t-p\}} + \epsilon_t \tag{2}$$

Variable dalam persamaan (2) tersebut antara lain  $\Delta y_t$  nilai differenced dari  $y_t - y_{\{t-1\}}$ ,  $\alpha$  konstanta,  $\beta t$  tren linear,  $\gamma y_{\{t-1\}}$  koefisien yang diuji untuk melihat stasioneritas data,  $\delta_1, \delta_2 \dots \delta_p$  koefisien lag differenced,  $\epsilon_t$  error term pada waktu.

Hasil uji ADF pada data penjualan harian "Aqua Botol 600ml" menunjukkan nilai ADF Statistic sebesar -2.552976, dengan p-value sebesar 0.103110. Selain itu, critical values untuk tingkat signifikansi 1%, 5%, dan 10% masing-masing adalah -3.500378, -2.892152, dan -2.583099. Mengingat bahwa nilai ADF Statistic lebih besar dari critical value pada tingkat signifikansi 5% dan 1%, serta p-value yang lebih besar dari 0.05 menunjukkan bahwa tidak ada cukup bukti untuk menolak hipotesis nol bahwa data memiliki unit root, dapat disimpulkan bahwa data ini tidak stasioner pada tingkat kepercayaan yang umum digunakan. Karena data tidak stationer maka harus dilakukan differencing dengan mengurangi nilai data pada waktu  $t$  dengan nilai data pada waktu  $t - 1$  [10]. Berikut merupakan grafik hasil dari differencing telah dilakukan sekali untuk mengubah data menjadi lebih stasioner yang terlihat dari variabilitas data yang lebih konsisten dalam grafik hasil differencing dan memungkinkan model SARIMA untuk menangkap pola dalam data dengan lebih baik serta bisa menghasilkan prediksi yang lebih akurat



Gambar 4. Visualisasi Data Hasil Differencing



Gambar 5. Grafik ACF

### 2.4. Identifikasi Model

Untuk mengidentifikasi parameter terbaik ( $p, d, q$ ) dan parameter musiman ( $P, D, Q, s$ ), dilakukan plot Autocorrelation Function (ACF) membantu menentukan parameter  $q$  (*moving average*) dan Partial Autocorrelation Function (PACF) membantu menentukan parameter  $p$  (*autoregressive*) dalam model SARIMA [10].

ACF (autocorrelation function) digunakan untuk mengukur korelasi antara observasi dalam *time series analysis* yang dipisahkan oleh lags. Persamaan ACF adalah:

$$\rho_k = \frac{\sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (3)$$

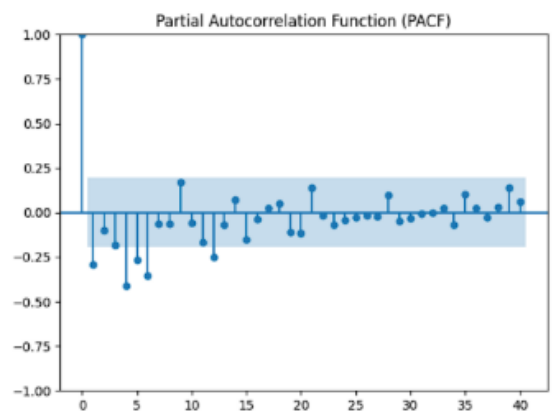
Variable dalam persamaan (3) tersebut antara lain :  $\rho_k$  Autocorrelation pada lag  $k$ ,  $y_t$  nilai pengamatan dari *time series analysis* pada waktu,  $\bar{y}$  rata-rata dari *time series analysis*,  $n$  jumlah total observasi.

PACF (*partial autocorrelation function*) mengukur korelasi antara observasi dalam *time series analysis* yang dipisahkan oleh lags, dengan memperhitungkan efek lags di antara mereka. Persamaan PACF adalah:

$$\phi_{kk} = \rho_k - \sum_{j=1}^{k-1} \phi_{kj} \phi_{k-j} \quad (4)$$

Variable dalam persamaan (4) tersebut antara lain :  $\phi_{kk}$  Partial autocorrelation pada lag  $k$ ,  $\rho_k$  Autocorrelation pada lag  $k$ ,  $\phi_{kj}$  Partial autocorrelation pada lag  $k$  dan  $j$

Hasil dari indentifikasi grafik ACF menunjukkan autocorrelation yang signifikan pada lag 1, sementara grafik PACF juga menunjukkan *partial autocorrelation* yang signifikan pada lag 1. Signifikansi pada lag 1 menunjukkan bahwa nilai penjualan pada hari sebelumnya memiliki pengaruh yang kuat terhadap nilai penjualan hari ini menunjukkan bahwa model SARIMA yang cocok kemungkinan memiliki komponen Moving Average ( $q$ ) dan Autoregressive ( $p$ ) dengan nilai lag 1 atau 2.



Gambar 6. Grafik PACF

Interpretasi awal dari ACF dan PACF memberikan panduan awal untuk memilih parameter, tetapi *grid search* akan digunakan untuk memastikan kombinasi parameter terbaik dengan mengevaluasi kombinasi parameter  $p, d, q, P, D, Q,$  dan  $s$  berdasarkan nilai *Akaike Information Criterion* (AIC) [11]. nilai AIC terendah dipilih karena menyeimbangkan antara kompleksitas model dan kualitas fit, sehingga model yang dipilih tidak hanya akurat tetapi juga efisien.

### 2.5. Estimasi Parameter

Pencarian parameter SARIMA terbaik dilakukan dengan menggunakan metode *grid search* berdasarkan nilai AIC (*Akaike Information Criterion*). Nilai AIC terendah dianggap sebagai model terbaik karena menyeimbangkan keakuratan dan kompleksitas model [12], persamaan AIC adalah:

$$AIC = 2k - 2\ln(L) \quad (5)$$

Variable dalam persamaan (5) tersebut antara lain :  $k$  jumlah parameter dalam model,  $L$  likelihood dari model.

Dilakukan evaluasi terhadap beberapa kombinasi parameter SARIMA yang meliputi paramter  $p, d, q$  untuk komponen non-musiman, serta  $P, D, Q, s$  untuk komponen musiman. Kombinasi parameter yang diuji melibatkan berbagai nilai yang dihasilkan dari interpretasi grafik ACF dan PACF sebelumnya. Tabel

berikut menunjukkan hasil *grid search*, dengan nilai AIC yang dihasilkan untuk setiap kombinasi parameter

Table 2. Hasil *Grid Search*

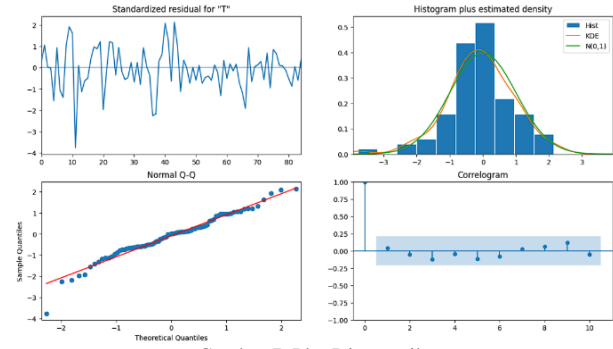
p,d,q	P,D,Q	AIC
0, 1, 0	1, 1, 1, 7	739.026
0, 1, 1	1, 1, 1, 7	699.787
0, 1, 2	1, 1, 1, 7	677.567
1, 1, 0	1, 1, 1, 7	718.773
1, 1, 1	1, 1, 1, 7	697.242
1, 1, 2	1, 1, 1, 7	679.559
2, 1, 0	1, 1, 1, 7	720.580
2, 1, 1	1, 1, 1, 7	698.010
2, 1, 2	1, 1, 1, 7	672.218

Berdasarkan hasil *grid search* yang didapatkan model SARIMA terbaik memiliki kombinasi parameter (2,1,2) untuk komponen non-musiman dan (1,1,1,7) untuk komponen musiman. Model ini memberikan nilai AIC terendah yaitu 672.218, yang menunjukkan bahwa model ini adalah yang paling optimal untuk memprediksi data penjualan harian dari dataset menu aqua botol 600ml karena memberikan keseimbangan terbaik antara keakuratan prediksi dan kompleksitas model[12]. Model SARIMA dengan parameter terbaik ini akan digunakan dalam tahap penerapan prediksi pada bagian Hasil dan Pembahasan. Dengan parameter yang telah diidentifikasi, model akan diterapkan untuk memprediksi penjualan harian selama 7 hari ke depan, dan hasilnya akan dievaluasi untuk mengukur keakuratan dan efektivitas model.

### 3. HASIL DAN PEMBAHASAN

Pemaparan hasil analisis menggunakan model SARIMA untuk memprediksi penjualan harian "Aqua Botol 600ml" di Kantin Sehat Universitas Majalengka. Sebelumnya, metode penelitian telah memastikan data memenuhi asumsi stasioneritas melalui differencing, diikuti oleh *grid search* untuk menemukan parameter SARIMA terbaik berdasarkan nilai AIC terendah. Hasil prediksi penjualan untuk 7 hari ke depan akan disajikan, disertai evaluasi keakuratan model menggunakan metrik seperti MAE, MSE, dan RMSE. Selain itu, diagnostik model akan dibahas untuk memastikan pemenuhan asumsi-asumsi penting [13]. Bagian ini juga akan mengeksplorasi implikasi hasil prediksi dalam pengelolaan stok, memberikan panduan praktis bagi penerapan hasil penelitian ini.

pada parameter (2, 1, 2)x(1, 1, 1, 7), dilakukan analisis diagnostik untuk memastikan bahwa model yang digunakan memenuhi asumsi-asumsi penting dari analisis *time series analysis*. Grafik plot diagnostik di bawah ini menunjukkan empat aspek kunci dari model residuals.



Gambar 7. Plot Diagnostik

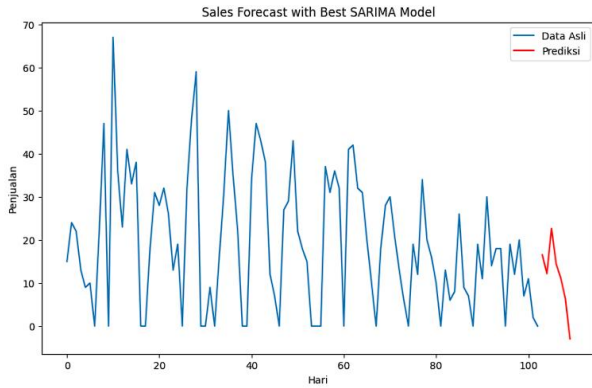
Residuals terdistribusi secara acak di sekitar nol, menunjukkan bahwa model berhasil menangkap pola utama tanpa meninggalkan tren atau musiman yang signifikan. pada histogram residuals mendekati distribusi normal, meskipun ada sedikit deviasi di ekor, mengindikasikan bahwa asumsi normalitas sebagian besar terpenuhi, begitu juga pada Normal Q-Q plot distribusinya normal dengan titik-titik yang sebagian besar berada di sepanjang garis referensi, serta pada correlogram tidak ada autocorrelation yang signifikan hal ini menunjukkan bahwa model telah efektif menghilangkan korelasi serial dalam data

Berdasarkan *grid search* yang telah dilakukan, model SARIMA terbaik yang dipilih adalah SARIMA(2, 1, 2)x(1, 1, 1, 7), dilakukan prediksi selama 7 hari ke depan disajikan pada tabel 3. menunjukkan estimasi jumlah penjualan harian yang bervariasi antara 6 hingga 22 unit, yang mencerminkan pola musiman dan tren yang terdeteksi dalam data historis.

Table 3. Hasil Prediksi 7 Hari Kedepan

Tanggal	Nama Menu	Terjual
05/08/2024	Aqua Botol 600ml	16
06/08/2024	Aqua Botol 600ml	12
07/08/2024	Aqua Botol 600ml	22
08/08/2024	Aqua Botol 600ml	14
09/08/2024	Aqua Botol 600ml	11
10/08/2024	Aqua Botol 600ml	6
11/08/2024	Aqua Botol 600ml	0

berikut bentuk grafik hasil prediksi yang disandingkan dengan historis. Garis biru pada grafik menunjukkan data penjualan historis yang digunakan untuk melatih model, sementara garis merah menunjukkan hasil prediksi. Dari grafik gambar 8, terlihat bahwa model SARIMA dapat mengikuti pola umum dari data penjualan historis



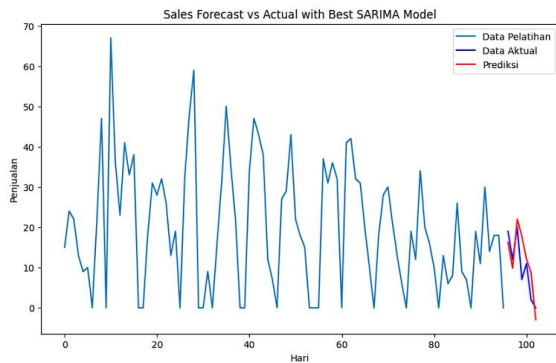
Gambar 8. Grafik Hasil Prediksi

Evaluasi akurasi prediksi dilakukan menggunakan tiga metrik utama: Mean Absolute Error (MAE), Mean Squared Error (MSE), dan Root Mean Squared Error (RMSE) [14]. Hasilnya adalah sebagai berikut:

Table 4. Hasil Evaluasi

Metrik	Nilai hasil
MAE	4.057564018100818
MSE	26.745518462089063
RMSE	5.171606951624327

Nilai evaluasi MAE 4.06 yang menunjukkan bahwa rata-rata kesalahan absolut dalam prediksi adalah sekitar 4 unit penjualan, MSE: 26.75 yang mengindikasikan bahwa ada beberapa perbedaan besar antara prediksi dan data aktual yang meningkatkan nilai kuadrat kesalahan. RMSE: 5.17 yang menunjukkan bahwa rata-rata kesalahan prediksi berada pada kisaran 5 unit penjualan.



Gambar 9. Grafik Data Awal, Aktual Dan Prediksi

Grafik yang dihasilkan menunjukkan perbandingan antara data pelatihan, data aktual, dan prediksi untuk 7 hari ke depan. Garis merah pada grafik mewakili prediksi yang dihasilkan oleh model SARIMA, sementara garis biru dan hitam mewakili data pelatihan dan data aktual, masing-masing. Dari grafik ini, terlihat bahwa model SARIMA dapat mengikuti pola umum penjualan, meskipun terdapat beberapa variasi antara prediksi dan data aktual, menunjukkan bahwa model SARIMA yang dipilih memiliki performa yang baik dalam memprediksi penjualan harian dengan tingkat kesalahan yang dapat diterima

**4. KESIMPULAN**

Penelitian ini telah berhasil mengaplikasikan model SARIMA (2, 1, 2)x(1, 1, 1, 7) untuk memprediksi penjualan harian "Aqua Botol 600ml" di Kantin Sehat Universitas Majalengka. Hasil prediksi menunjukkan tingkat akurasi yang memadai, dengan evaluasi metrik MAE, MSE, dan RMSE yang menunjukkan kesalahan prediksi berada dalam batas yang dapat diterima. Analisis diagnostik residuals mengkonfirmasi bahwa model ini memenuhi asumsi-asumsi penting seperti normalitas, tidak adanya autocorrelation, dan acaknya residuals, yang menegaskan validitas model dalam memprediksi pola penjualan.

Prediksi yang dihasilkan dapat diimplementasikan secara langsung dalam pengelolaan stok di kantin, memungkinkan manajemen untuk lebih tepat dalam merencanakan pengadaan dan mengurangi risiko kelebihan atau kekurangan stok. Model ini menawarkan alat yang efektif untuk mendukung keputusan manajerial, terutama dalam menghadapi fluktuasi permintaan yang tidak menentu. Dengan menerapkan prediksi ini, Kantin Sehat Universitas Majalengka dapat meningkatkan efisiensi operasional, mengurangi biaya, dan memastikan tingkat kepuasan pelanggan yang lebih tinggi.

**DAFTAR PUSTAKA**

- [1] M. Mustika et al., DATA MINING DAN APLIKASINYA. Bandung: CV WIDINA MEDIA UTAMA, 2021. Accessed: Jun. 27, 2024. [Online]. Available: <https://repository.penerbitwidina.com/publication/s/351768/data-mining-dan-aplikasinya>
- [2] N. S. Represa, A. Fernández-Sarría, A. Porta, and J. Palomar-Vázquez, "Data Mining Paradigm in the Study of Air Quality," *Environmental Processes*, vol. 7, no. 1, pp. 1–21, Mar. 2020, doi: 10.1007/s40710-019-00407-5.
- [3] Y. Ensafi, S. H. Amin, G. Zhang, and B. Shah, "Time-series forecasting of seasonal items sales using machine learning – A comparative analysis," *International Journal of Information Management Data Insights*, vol. 2, no. 1, p. 100058, Apr. 2022, doi: 10.1016/j.ijime.2022.100058.
- [4] Aa Herdi Prayoga, "Implementasi Load Balancing Menggunakan Metode Per Connection Classifier (PCC) dengan Failover pada Server Jaringan Mikrotik (Studi Kasus Universitas Majalengka)," Universitas Majalengka, Majalengka, 2023.
- [5] A. D. Supriatna and F. A. Aljjoyo, "Increasing The Competitiveness Of Sukasari District UMKM Through Digital Technology-Based Cooperatives," *Jurnal Info Sains: Informatika dan Sains*, vol. 13, no. 03, pp. 754–759, 2023.
- [6] M. Heru Widiyanto, R. Mayasari, and G. Garno, "IMPLEMENTASI TIME SERIES PADA DATA PENJUALAN DI GAIKINDO MENGGUNAKAN ALGORITMA SEASONAL ARIMA," *JATI (Jurnal Mahasiswa Teknik*

- Informatika), vol. 7, no. 3, pp. 1501–1506, Oct. 2023, doi: 10.36040/jati.v7i3.6879.
- [7] D. Pratiwi, S. M. U. Agustini, W. Windasari, and E. N. Kencana, “Forecasting Farmer Exchange Rate in Bali Province Using Seasonal Autoregressive Integrated Moving Average (SARIMA) Method,” *J Phys Conf Ser*, vol. 1503, no. 1, p. 012002, Jul. 2020, doi: 10.1088/1742-6596/1503/1/012002.
- [8] R. Saputro and S. Novani, “Data Analytics for Decision-Making in Evaluating the Top-Performing Product and Developing Sales Forecasting Model in an Oil Service Company,” *International Journal of Current Science Research and Review*, vol. 07, no. 02, Feb. 2024, doi: 10.47191/ijcsrr/V7-i2-14.
- [9] V. P. Ariyanti and Tristyanti Yusnitasari, “Comparison of ARIMA and SARIMA for Forecasting Crude Oil Prices,” *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 2, pp. 405–413, Mar. 2023, doi: 10.29207/resti.v7i2.4895.
- [10] Hedi, A. Suryani, and A. Binarto, “Forecasting the Number of New Cases of COVID-19 in Indonesia Using the ARIMA and SARIMA Prediction Models,” 2021. doi: 10.2991/aer.k.211106.011.
- [11] A. Natayu, Q. J. H. Clarke, and M. Fikri, “Benchmark of Holt-Winters and SARIMA Methods in Predicting Jakarta Climate,” 2022.
- [12] S. K. Prilistya, A. E. Permanasari, and S. Fauziati, “The Effect of The COVID-19 Pandemic and Google Trends on the Forecasting of International Tourist Arrivals in Indonesia,” in *2021 IEEE Region 10 Symposium (TENSYPMP)*, IEEE, Aug. 2021, pp. 1–8. doi: 10.1109/TENSYPMP52854.2021.9550838.
- [13] D. K. Silalahi, “Forecasting of Poverty Data Using Seasonal ARIMA Modeling in West Java Province,” *JTAM | Jurnal Teori dan Aplikasi Matematika*, vol. 4, no. 1, p. 76, Apr. 2020, doi: 10.31764/jtam.v4i1.1888.
- [14] H. U. MY, S. L. Setyowati, K. A. Notodiputro, Y. Angraini, and L. N. A. Mualifah, “Comparison of Seasonal ARIMA and Support Vector Machine Forecasting Method for International Arrival in Lombok,” *Jambura Journal of Mathematics*, vol. 6, no. 2, pp. 212–219, Aug. 2024, doi: 10.37905/jjom.v6i2.26478.