

BAB II

TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Dalam penelitian ini, terdapat beberapa studi yang telah dilakukan sebelumnya yang relevan dengan prediksi kebutuhan air menggunakan metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Penelitian-penelitian tersebut memberi wawasan tentang penerapan metode SARIMA pada berbagai bidang serta hasil akurasi model yang diperoleh. Berikut adalah beberapa studi yang relevan dengan topik ini.

Penelitian yang dilakukan oleh Agustin dkk., (2022) berjudul “Prediksi Kualitas Air Menggunakan Metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA)” penelitian ini berfokus pada prediksi parameter kualitas air seperti pH, suhu, dan kekeruhan. Hasil penelitian menunjukkan tingkat akurasi yang tinggi dengan nilai MSE sebesar 0.0 dan RMSE sebesar 0.05, yang menunjukkan bahwa model SARIMA mampu memberikan prediksi yang akurat.

Penelitian yang dilakukan oleh Suseno & Suryo Wibowo., (2023) berjudul “Penerapan Metode ARIMA dan SARIMA Pada Peramalan Penjualan Telur Ayam Pada PT Agromix Lestari Group” membandingkan akurasi metode ARIMA dan SARIMA dalam memprediksi penjualan telur ayam. Hasil penelitian menunjukkan bahwa metode SARIMA menghasilkan MSE sebesar 144.346, lebih rendah dibandingkan ARIMA, yang menunjukkan bahwa metode ini dianggap lebih akurat untuk data yang memiliki pola musiman.

Penelitian yang dilakukan oleh Tokan & Hermawan., (2023) berjudul “Implementasi Model SARIMA untuk Memprediksi Produksi Minyak Kelapa Sawit” penelitian ini memberikan hasil bahwa model SARIMA memiliki tingkat akurasi yang lebih baik dengan MAPE sebesar 5.54%, hasil ini menunjukkan bahwa SARIMA mampu menangani data musiman.

Penelitian yang dilakukan oleh Salsalina & Widyasari., (2023) berjudul “Prediksi Jumlah Pemakaian Air Bersih Menggunakan Metode *Hybrid Singular*

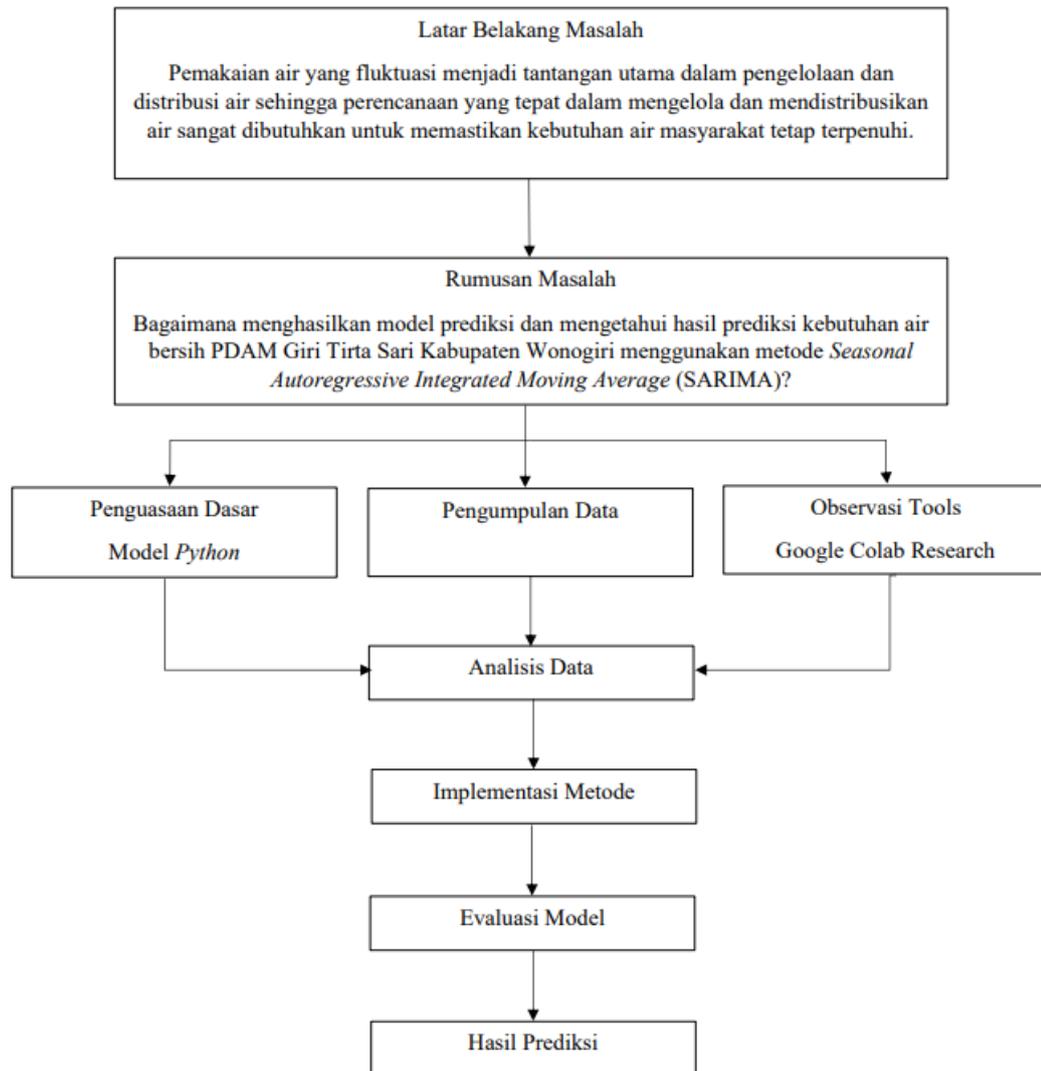
Spectrum Analysis (SSA) Dan SARIMA di PDAM Tirtanadi Sibolangit” penelitian ini menggunakan kombinasi metode SSA dan SARIMA dalam memprediksi kebutuhan air bersih. Hasil penelitian menunjukkan bahwa metode *hybrid* ini memberikan tingkat akurasi dengan MAPE 6.92%.

Penelitian yang dilakukan oleh Heru Widiyanto dkk., (2023) berjudul “Implementasi *Time Series* Pada Data Penjualan Di Gaikindo Menggunakan Algoritma *Seasonal ARIMA*” penelitian ini bertujuan untuk menguji metode *time series* menggunakan algoritma SARIMA menunjukkan hasil tingkat akurasi model nilai MAPE sebesar 0,407% sehingga model ini memiliki kemampuan prediksi yang baik.

Berdasarkan dari penelitian terdahulu, metode SARIMA telah digunakan di berbagai bidang seperti kualitas air, prediksi penjualan, dan sektor industri. Namun belum banyak penelitian yang membahas prediksi kebutuhan air di PDAM menggunakan metode SARIMA. Selain itu, Sebagian besar penelitian terdahulu menggunakan metrik akurasi yang berbeda, seperti MSE, RMSE, dan MAPE, atau pendekatan *hybrid*. Oleh karena itu, penelitian ini bertujuan untuk mengisi celah tersebut dengan menerapkan metode SARIMA secara spesifik pada prediksi kebutuhan air di PDAM, serta membandingkan dengan metode ARIMA sebagai uji validasi untuk memastikan ketepatan model SARIMA dalam menangkap pola musiman yang ada pada data historis.

2.2 Kerangka Pemikiran

Kerangka pemikiran ini memberikan gambaran tahap-tahap dari penelitian. Alur penelitian terkait prediksi kebutuhan air bersih PDAM Giri Tirta Sari Kabupaten Wonogiri menggunakan metode *Seasonal Autoregressive Integrated Moving Average (SARIMA)* dapat dilihat pada Gambar 2.1.



Gambar 2.1 Kerangka Pemikiran

Penjelasan dari kerangka pemikiran di atas adalah:

1. Latar Belakang

Latar belakang memberikan penjelasan permasalahan yang menjadi fokus penelitian. Dalam hal ini kebutuhan air bersih yang fluktuasi menjadi tantangan utama dalam pengelolaan dan distribusi air sehingga perencanaan yang tepat dalam mengelola dan mendistribusikan air sangat dibutuhkan untuk memastikan kebutuhan air masyarakat tetap terpenuhi. Penelitian ini bertujuan untuk menghasilkan model prediksi dan mengetahui hasil prediksi kebutuhan air bersih PDAM Giri Tirta Sari Kabupaten Wonogiri menggunakan metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA).

2. Rumusan Masalah

Bagaimana menghasilkan model prediksi dan mengetahui hasil prediksi kebutuhan air bersih PDAM Giri Tirta Sari Kabupaten Wonogiri menggunakan metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA).

3. Penguasaan Dasar

Tahap mempelajari dasar-dasar kode program *python* untuk meningkatkan pemahaman terkait implementasi model prediksi.

4. Pengumpulan Data

Sumber data yang digunakan dalam penelitian ini diambil langsung dari PDAM Giri Tirta Sari Kabupaten Wonogiri dengan menggunakan data pemakaian air bersih perbulan dalam satuan meter kubik (m^3) selama periode tahun 2020-2024.

5. Observasi Tools

Pengamatan terhadap fitur-fitur google colab serta mengkaji berbagai referensi dari jurnal, buku, dan karya ilmiah sebagai acuan dalam menerapkan model prediksi.

6. Analisis Data

Analisis data dilakukan untuk memastikan data siap untuk digunakan dalam pemodelan. Proses ini meliputi pengecekan kelengkapan data untuk

menemukan dan menangani data yang hilang atau tidak valid, normalisasi format data agar sesuai dengan kebutuhan prediksi, dan pembagian data menjadi data latih (*training set*) dan data uji (*test set*) untuk mengevaluasi hasil prediksi. Sebelum diterapkan dalam model prediksi, analisis pola data dilakukan dengan menggunakan uji stasioneritas *Augmented Dickey-Fuller* (ADF) dan menemukan pola musiman dan tren dari data historis.

7. Implementasi Metode

Implementasi metode meliputi beberapa langkah, pertama, parameter model diidentifikasi melalui analisis plot ACF dan PACF, kemudian, untuk memilih model terbaik, estimasi parameter didasarkan pada nilai *Akaike Information Criterion* (AIC), dan terakhir, uji kelayakan model untuk memastikan bahwa model yang digunakan sudah optimal. Setelah memperoleh model terbaik, peramalan kebutuhan air bersih dilakukan untuk Januari 2025 hingga Desember 2025 menggunakan data historis yang telah dianalisis sebelumnya.

8. Evaluasi Model

Evaluasi ini dilakukan dengan membandingkan model prediksi dengan data aktual menggunakan metrik *Mean Absolute Percentage Error* (MAPE), *Mean Absolute Error* (MAE), dan *Root Mean Square Error* (RMSE) sehingga dapat diketahui hasil akurasi model prediksi. Selain itu, evaluasi juga dilakukan dengan membandingkan performa model SARIMA dengan model ARIMA.

9. Hasil Prediksi

Model SARIMA yang telah dievaluasi digunakan untuk menghitung prediksi kebutuhan air bersih PDAM Giri Tirta Sari Kabupaten Wonogiri dari Januari 2025 hingga Desember 2025.

2.3 Landasan Teori

2.3.1 Prediksi

Prediksi adalah perkiraan permintaan untuk masa mendatang. Namun, prediksi hanyalah perkiraan, dengan metode tertentu prediksi dapat lebih akurat dari perkiraan sehingga dapat digunakan untuk membuat keputusan dalam perencanaan

(Andrian dkk. 2020). Prediksi dapat membantu menganalisis pola dari data sebelumnya dan meliputi peristiwa yang terjadi dalam rentang waktu tertentu seperti hari, minggu, bulan, atau tahun. Hasil prediksi tidak selalu memberikan jawaban yang tepat, tetapi dapat memberikan jawaban yang lebih dekat dengan apa yang akan terjadi di masa depan. Sehingga, prediksi dapat membantu dalam perencanaan dan pemecahan masalah secara sistematis serta memberi keyakinan yang kuat terhadap prediksi (Ainiyah & Bansori, 2021). Prediksi merupakan bagian penting pada pengambilan keputusan, sebagai dasar untuk perencanaan jangka pendek maupun jangka panjang. Berdasarkan hal tersebut tujuan prediksi adalah untuk mengoptimalkan pengambilan keputusan produksi.

Metode prediksi dapat diklasifikasikan berdasarkan rentang waktu yang digunakan dalam analisis yaitu:

1. Jangka pendek yaitu berfokus pada jangka waktu kurang dari satu tahun dan berguna untuk perencanaan keuangan sehari-hari, produksi, dan persediaan.
2. Jangka menengah yaitu digunakan untuk mengidentifikasi tren dan mengembangkan strategi bisnis untuk satu hingga tiga tahun ke depan.
3. Jangka panjang yaitu perencanaan masa depan yang dilakukan selama lebih dari tiga tahun untuk membantu dalam pengambilan keputusan jangka Panjang dan pembuatan strategi bisnis.

2.3.2 Kebutuhan Air Bersih

Kebutuhan air adalah jumlah air yang diperlukan untuk memenuhi kebutuhan pokok manusia. Kebutuhan air menentukan besaran sistem dan ditetapkan berdasarkan pemakaian air. Kebutuhan air dibagi menjadi kebutuhan domestik (proyeksi penduduk) dan kebutuhan non-domestik/fasilitas umum. Kebutuhan air domestik digunakan untuk memenuhi kebutuhan air rumah tangga yang ditentukan berdasarkan jumlah dan laju pertumbuhan penduduk. Sedangkan, kebutuhan non-domestik digunakan untuk fasilitas umum (Aminuddin dkk. 2023).

Kebutuhan air bagi manusia meningkat seiring dengan pertumbuhan penduduk. Jumlah ketersediaan air sangat penting diperlukan untuk memfasilitasi masyarakat, karena manusia menggantungkan hidupnya kepada air dalam

keberlangsungan hidup. Jumlah kebutuhan air berbeda-beda berdasarkan dari jumlah pemakaian dan jumlah penduduk setempat (Primandani dkk., 2022).

Oleh karena itu, jumlah kebutuhan air memerlukan prediksi untuk menghindari kekurangan atau kelebihan produksi, optimalisasi sumber daya, dan memastikan distribusi yang memenuhi terutama pada musim kemarau atau saat peningkatan jumlah pelanggan.

2.3.3 Data dan Analisis Deret Waktu (*Time Series*)

Analisis deret waktu adalah metode prediksi menggunakan analisis pola hubungan antara variabel yang akan diperkirakan dengan variabel waktu. Data *time series* adalah suatu data runtun waktu dalam periode tertentu, yang kemudian digunakan untuk membuat model dan selanjutnya model tersebut digunakan untuk memprediksi di periode selanjutnya.

Prediksi suatu data *time series* memperhatikan tipe atau pola data (Baidowi & Buniarto, 2020). Secara umum terdapat empat macam pola data *time series* yaitu:

1. Tren (*Trend*)

Pola data yang menunjukkan perubahan data cenderung naik atau turun. Pada masa puncak (*prosperity*) mengalami kenaikan sedangkan, pada masa *declined* pola data mengalami penurunan.

2. Musiman (*Seasonality*)

Pola data dengan permintaan yang dipengaruhi oleh perubahan musim, seperti cuaca dan liburan (misalnya kuartal tahun tertentu) kebutuhan air bersih menunjukkan pola ini. Biasanya terjadi pada rentang waktu tahunan, namun bulanan dan mingguan juga dapat terjadi pola permintaan musiman.

3. Siklis (*Cycle*)

Pola data yang terjadi pada rentang waktu yang lebih panjang, misal pola musiman dalam waktu beberapa tahun atau *decade*.

4. Horizontal

Pola ini terjadi perubahan pada nilai rata-rata yang tidak membentuk pola seperti pola musiman, tren, atau siklus.

Analisis deret waktu adalah metode kuantitatif yang khusus untuk data yang dikumpulkan pada interval tertentu. Metode ini mengidentifikasi tren yang terjadi,

perubahan musiman, dan perubahan siklus. Data pemakaian air dari bulan ke bulan dapat dianalisis dengan deret waktu, sehingga metode ini tepat untuk digunakan.

2.3.4 Stasioneritas Data

Stasioneritas data dapat diuji menggunakan *Augmented Dickey Test*. Pengujian ADF memiliki pertimbangan tentang auto-korelasi pada error jika data tidak stasioner. Data dinyatakan stasioner menggunakan ADF jika nilai $\alpha < 0.05$, jika data dengan hasil uji ADF menghasilkan nilai $\alpha > 0.05$ maka harus dilakukan *differencing* data (Sunariadi dkk., 2022).

2.3.5 Plot ACF dan PACF

Plot ACF dan PACF metode yang digunakan untuk menentukan parameter model dugaan sementara dengan melihat *lag* yang keluar dari batas garis pada plotnya. Plot ACF untuk membentuk model MA pada *lag* musiman (Q) dan non-musiman (q), sedangkan plot PACF untuk membentuk model AR *lag* musiman (P) dan non-musiman (p) (Muryanto, 2021).

2.3.6 *Seasonal Autoregressive Integrated Moving Average (SARIMA)*

Model *Seasonal-ARIMA (SARIMA)* merupakan model gabungan komponen ARIMA yaitu *Auto-regressive* dan *Moving Average* serta komponen musiman atau *seasonal*. Model ini sesuai untuk data yang memiliki data historis panjang dan lengkap, karena model ARIMA akan menghasilkan model yang fit jika memiliki data historis yang panjang. ARIMA adalah salah satu metode peramalan deret waktu yang menggabungkan tiga komponen yaitu:

1. *Autoregressive (AR)*: hubungan antara data saat ini dengan data masa lalu dalam satu deret waktu.
2. *Differencing (I – Integrated)*: proses transformasi data agar menjadi stasioner dengan mengurangi tren.
3. *Moving Average (MA)*: Model yang memperhitungkan kesalahan dari observasi sebelumnya dalam proses prediksi.

Model ARIMA dinotasikan sebagai ARIMA (p, d, q) dengan:

p : orde komponen AR

d : jumlah *differencing* yang dilakukan agar stasioner

q : orde komponen MA

Model ARIMA cocok digunakan untuk data deret waktu yang tidak memiliki pola musiman tetapi memiliki tren yang kuat (Khoiri, 2022).

Metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA) adalah pengembangan dari ARIMA yang menambahkan komponen musiman pada proses prediksi. Secara matematis membuat model AR(p), MA(q), ARMA(p,q), ARIMA(p,d,q), lalu SARIMA(p,d,q) (P,D,Q). Metode ini, p adalah nilai AR dari plot PACF data non musiman, d adalah nilai *differencing* dari data non musiman, q adalah nilai MA dari plot ACF data non musiman, dan P adalah nilai AR dari plot PACF data musiman, D adalah nilai *differencing* data musiman, Q adalah nilai MA dari plot ACF data musiman, dan s adalah jumlah musimannya (Sunariadi dkk., 2022).

Notasi umum untuk model SARIMA adalah:

$$\text{SARIMA}(p, d, q)(P, D, Q)^s \quad (2.1)$$

Dengan:

p = orde AR untuk data non-musiman

d = *differencing* untuk data non-musiman

q = orde MA untuk data non-musiman

P = orde AR untuk data musiman

D = *differencing* untuk data musiman

Q = orde MA untuk data musiman

S = periode musiman

Komponen utama model SARIMA yang mencakup musiman dan non-musiman terdiri dari:

1. AR (*Autoregressive*)

Model *autoregressive* adalah bagian dari metode ARIMA, yang memiliki orde p, sehingga dapat dituliskan dengan $(p, 0, 0)$. Persamaan matematis untuk model AR adalah:

$$\hat{z}_t = \emptyset_1 \hat{z}_{t-1} + \emptyset_2 \hat{z}_{t-2} + \dots + \emptyset_p \hat{z}_{t-p} + \alpha_t \quad (2.2)$$

Dengan:

\emptyset_p : parameter *autoregressive* ke-p

α_t : *white noise* nilai kesalahan pada saat t

\hat{z}_{t-p} : variabel bebas

Variabel bebas dalam persamaan ini sebagai model untuk prediksi, sedangkan antara \hat{z}_t dengan \hat{z}_{t-p} adalah variable yang sama tetapi beda deret waktunya.

Pada model AR ini penentuan orde berdasarkan *lag cut off* pada plot PACF.

Berdasarkan model umum AR (p) maka untuk AR musiman didefinisikan sebagai berikut:

$$\hat{z}_t = \emptyset_1 \hat{z}_{t-s} + \emptyset_2 \hat{z}_{t-2s} + \dots + \emptyset_p \hat{z}_{t-ps} + \alpha_t \quad (2.3)$$

Dengan:

\emptyset_i : koefisien AR musiman, $i=1,2,3,\dots,p$

P : orde AR musiman

S : jumlah periode permusim

2. MA (*Moving Average*)

Model *moving average* atau metode rata-rata bergerak adalah yang mencari rata-rata dari data aktual pada periode sebelumnya untuk memprediksi apa yang akan terjadi pada periode mendatang. Model ini dilambangkan dengan MA(q) dengan q adalah orde dari MA dituliskan dengan $(0,0,q)$. Pada model MA ini penentuan orde berdasarkan *lag cut off* pada plot ACF. Persamaan matematis untuk model MA adalah:

$$\hat{z}_t = \alpha_t - \emptyset_1 \alpha_{t-1} - \emptyset_2 \alpha_{t-2} - \dots - \emptyset_q \alpha_{t-q} \quad (2.4)$$

Dengan:

\emptyset_q : parameter *moving average*

α_t : *white noise* kesalahan pada saat t

Berdasarkan model umum MA (q) maka untuk MA musiman didefinisikan sebagai berikut:

$$\hat{z}_t = \alpha_t - \emptyset_1 \alpha_{t-s} - \emptyset_2 \alpha_{t-2s} - \dots - \emptyset_q \alpha_{t-qs} \quad (2.5)$$

Dengan:

\emptyset_i : koefisien MA musiman, $I = 1, 2, 3, \dots, Q$

q : orde MA musiman

s : jumlah periode permusim

3. ARMA (*Auto Regressive Moving Average*)

Model *auto regressive moving average* adalah gabungan dari model AR dan MA, dengan penggabungan dua model ini agar mendapat output yang lebih baik dan nilai error yang lebih kecil. Model ARMA dituliskan dengan $(p, 0, q)$. persamaan matematis untuk model ARMA adalah:

$$\hat{z}_t = \emptyset_1 \hat{z}_{t-1} + \dots + \emptyset_p \hat{z}_{t-p} + \alpha_t - \emptyset_1 \alpha_{t-1} - \dots - \emptyset_q \alpha_{t-q} \quad (2.6)$$

Pada model ini jenis data yang digunakan adalah data yang memiliki plot ACF dan PACF menurun secara eksponensial.

Berdasarkan model umum MA (q) maka untuk MA musiman didefinisikan sebagai berikut:

$$\hat{z}_t = \emptyset_1 \hat{z}_{t-s} + \dots + \emptyset_p \hat{z}_{t-ps} + \alpha_t - \emptyset_1 \alpha_{t-s} - \dots - \emptyset_q \alpha_{t-qs} \quad (2.7)$$

Dengan:

\emptyset_i : koefisien MA musiman, $I = 1, 2, 3, \dots, Q$

\emptyset_i : koefisien AR musiman, $I = 1, 2, 3, \dots, P$

q : orde MA musiman

p : orde AR musiman

s : jumlah periode permusim

4. ARIMA (*Auto Regressive Integrated Moving Average*)

Model ARIMA adalah model yang tidak memperhatikan perhitungan faktor dari variabel bebas untuk proses prediksi. Persamaan matematis untuk model ARIMA adalah:

$$\emptyset_p(B)(1 - B)^d \hat{z}_t = \theta_q(B) \alpha_t \quad (2.8)$$

Dengan:

$$\emptyset_p(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \dots - \emptyset_p B^p$$

$$\emptyset_q(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \dots - \emptyset_q B^q$$

$(1 - B)^d$: *differencing* orde non – musiman

$\phi_1, \phi_2, \dots, \phi_p$: koefisien orde p

$\theta_1, \theta_2, \dots, \theta_q$: koefisien orde q

α_t : nilai galat pada waktu t dengan asumsi *white noise*

Model musiman pada ARIMA dapat diidentifikasi melalui plot ACF pada data stasioner yang signifikan selain pada tiga lag awal, misalkan signifikan pada lag 12 maka terdapat indikasi adanya musiman dengan periode 12. Orde dalam model SARIMA terdiri dari orde untuk ARIMA dan orde untuk komponen musiman. Dalam menentukan orde yang tepat untuk model SARIMA, langkah-langkah yang dilakukan sama seperti penentuan orde ARIMA, perbedaannya adalah pada metode SARIMA dilakukan dua kali identifikasi orde yaitu untuk orde non musiman dan orde musiman.

2.3.7 Mean Absolute Percent Error (MAPE)

MAPE adalah akurasi persentase absolut antara nilai hasil prediksi dan nilai aktual. Hasil prediksi mengandung ketidakpastian, penting untuk mengetahui seberapa akurat prediksi tersebut MAPE adalah metrik untuk mengukur akurasi prediksi dengan menghitung rata-rata kesalahan dari nilai prediksi dan nilai aktual dalam nilai persentase (Kesuma, 2024).

Secara matematis dapat dituliskan sebagai berikut:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right| \times 100\% \quad (2.9)$$

Dengan:

n : jumlah data

A_i : data aktual

F_i : data hasil peramalan

Kategori akurasi MAPE ditunjukkan pada Tabel 2.1.

Tabel 2. 1 Kategori Akurasi MAPE

Range MAPE	Kategori MAPE
<10%	Akurasi prediksi sangat baik

10%-20%	Akurasi prediksi baik
20%-50%	Akurasi prediksi layak (cukup baik)
>50%	Akurasi prediksi buruk (tidak akurat)

Sumber: Nirmala dkk., (2023)

2.3.8 Mean Absolute Error (MAE)

MAE mengukur rata-rata selisih absolut antara nilai prediksi dengan nilai aktual. Dengan rumus:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2.10)$$

Dengan:

y_i : nilai aktual

\hat{y}_i : nilai prediksi

n: jumlah data

MAE memberi gambaran langsung rata-rata kesalahan dalam satuan yang sama dengan data asli, sehingga mudah dipahami dan digunakan dalam evaluasi model prediksi (Amansyah dkk., 2024)

2.3.9 Root Mean Square Error (RMSE)

RMSE mengukur akar kuadrat dari rata-rata kuadrat selisih antara nilai prediksi dan nilai aktual. Dengan rumus:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.11)$$

RMSE menghitung seberapa besar kesalahan kuadrat rata-rata model dengan penalty yang lebih besar terhadap kesalahan yang lebih besar (Made dkk., 2024)

2.3.10 Google Colabs

Google colab singkatan dari google collaboration yang merupakan platform berbasis cloud computing yang disediakan oleh Google. Platform ini memungkinkan pengguna mengeksekusi kode python tanpa perlu menginstal atau mengatur lingkungan local mereka sendiri. Peneliti, pengembang, dan ilmuwan

data sering menggunakan Google Colab untuk melakukan berbagai tugas, seperti pemrosesan data, analisis data, dan pelatihan model mesin.

Menurut (Yanuar, 2024), beberapa fitur utama dari Google Colab adalah:

1. *Python di Cloud*: dapat menulis dan mengesekusi kode *python* langsung di browser web tanpa perlu menginstal Pustaka di komputer.
2. Gratis: Google Colab adalah layanan gratis yang disediakan oleh Google.
3. GPU Gratis: Google Colab menyediakan akses ke GPU (*Graphics Processing Unit*) secara gratis. Ini sangat berguna untuk pelatihan model mesin yang memerlukan daya komputasi tinggi.
4. Akses penyimpanan Google Drive: dapat mengakses dan menyimpan file langsung di Google Drive, sehingga memudahkan dalam menyimpan dan berbagi file.
5. Notebook Interaktif: Google Colab menggunakan format “notebook” yang memungkinkan untuk menggabungkan kode, teks, gambar, dan hasil dalam satu file interaktif. Ini sangat memudahkan untuk dokumentasi dan melakukan berbagai analisis.
6. Pustaka Tersedia: Goggle Colab menyediakan banyak pustaka umum yang dapat diimpor seperti *numpy*, *pandas*, *tensorflow*, *pytorch*, dan sebagainya.
7. Kerjasama Tim: dapat berbagi notebook colab dengan anggota tim dan berkolaborasi.
8. Fleksibel dan Mudah Digunakan: google Colab adalah layanan yang ramah bagi pemula, memudahkan untuk memulai pemograman *python* dan eksplorasi data tanpa kerumitan konfigurasi lokal.

2.3.11 Python

Python adalah bahasa pemrograman dengan kode program yang mudah dipahami, *open-source*, dan dapat diinterpretasikan. *Python* sering digunakan dalam berbagai bidang analisis data, pembelajaran mesin, dan peramalan deret waktu. *Python* memiliki Pustaka yang luas, seperti *numpy*, *pandas*, *matplotlib*, dan *statsmodels*, yang mendukung analisis data dan prediksi (Alfarizi dkk., 2023).

2.3.12 Akaike Information Criterion (AIC)

AIC diperkenalkan pertama oleh Hirotugu Akaike pada tahun 1973 sebagai kriteria informasi yang digunakan untuk pemilihan model statistik terbaik dari sekumpulan dugaan model. AIC adalah kriteria untuk memilih model terbaik yang dapat dihitung dengan rumus:

$$AIC = \frac{\sum_{i=1}^n \theta_i^2}{n} + \frac{2k}{n} \quad (2.12)$$

AIC adalah teknik pemilihan model terbaik dari beberapa model yang telah ditentukan, dipilih berdasarkan kriteria yang digunakan. AIC adalah suatu kriteria yang menyeimbangkan *goodness of fit* model terbaik berdasarkan nilai likelihood dari dugaan parameter yang ada. Sehingga, model yang terbaik adalah model dengan nilai AIC terkecil (Utama & Hajarisman, 2021).

2.3.13 Z-Score

Z-Score adalah perhitungan statistika yang menunjukkan seberapa jauh hubungan sebuah nilai dengan rata-rata dari kumpulan data dalam satuan standar deviasinya. *Z-Score* dapat mengidentifikasi outlier dengan nilai Z jika lebih besar dari +3 atau kurang dari -3 dianggap sebagai *outlier* (Henderi, 2021).

Z-Score dihitung dengan rumus:

$$Z = \frac{(x - \mu)}{\sigma} \quad (2.13)$$

Di mana:

X : titik data yang ingin di cari z-score-nya

μ : rata-rata dari data

σ : standar deviasi data

Analisis data menggunakan *Z-Score* sangat penting dalam proses pra-pemrosesan data untuk deteksi anomali yang mempengaruhi akurasi dalam pemodelan prediksi. *Z-Score* sangat penting untuk mengidentifikasi dan menangani *outlier* dengan memberi gambaran yang objektif tentang penyebaran data dalam proses analisis statistik.

2.3.14 Uji *White Noise*

Uji white noise pada analisis deret waktu digunakan untuk menguji data apakah residual dari model bersifat acak atau tidak berkorelasi antar waktu. Model prediksi dianggap baik apabila residualnya merupakan white noise, yaitu tidak memiliki pola tertentu dan tidak korelasi secara signifikan antar waktu. Data yang white noise adalah yang historis yang sudah sepenuhnya ditangkap oleh model, sehingga residual (tingkat kesalahan) hanya pada fluktuasi acak. Salah satu metode yang umum digunakan untuk uji white noise adalah Ljung-Box Test, yang menguji autokorelasi dari residual pada lag (Nugroho dkk., 2025)

Hipotesis dalam uji Ljung-Box:

- a. H_0 : tidak terdapat autokorelasi dalam residual (white noise).
- b. H_0 : terdapat autokorelasi dalam residual (bukan white noise).

Kriteria uji:

- a. Jika nilai p-value $> 0,05$, maka H_0 tidak ditolak, sehingga residual dianggap white noise.
- b. Jika nilai p-value $\leq 0,05$, maka H_0 ditolak, sehingga residual mengandung autokorelasi.

Uji white noise merupakan uji yang sangat penting dalam evaluasi model deret waktu, karena kemampuan model dalam menangkap pola data sangat dipengaruhi oleh keacakannya. Model yang residualnya bersifat white noise dianggap mampu merepresentasikan pola data dengan baik (Azmi & Syaifudin, 2020)

2.3.15 Uji Distribusi Normal

Uji distribusi normal merupakan salah satu tahap penting dalam analisis deret waktu, yang digunakan untuk mengevaluasi apakah data residual bersifat normal. Uji distribusi normal bertujuan untuk mengetahui apakah data yang diperoleh berasal dari populasi yang berdistribusi normal, meskipun dalam data deret waktu, asumsi normalitas tidak bersifat mutlak. Salah satu metode yang umum digunakan untuk uji normalitas adalah uji Jarque-Bera. Uji ini berdasarkan pada

nilai skewness (kemiringan) dan kurtosis (keruncingan) dari distribusi data (Erdin, 2020)

Hipotesis dalam uji Jarque-Bera:

- a. H_0 : data residual berdistribusi normal.
- b. H_1 : data residual tidak berdistribusi normal.

Kriteria uji:

- a. Jika nilai p-value $> 0,05$, maka H_0 tidak ditolak, sehingga residual dianggap berdistribusi normal.
- b. Jika nilai p-value $\leq 0,05$, maka H_0 ditolak, sehingga residual tidak berdistribusi normal.

Asumsi normalitas residual dalam penerapan deret waktu bukan syarat wajib untuk validitas model, apabila syarat keacakan (white noise) terpenuhi maka model dapat digunakan meskipun residual tidak terdistribusi normal. Uji distribusi normal dapat menjadi informasi tambahan dalam evaluasi model SARIMA karena kemampuannya dalam menangkap pola musiman dan tren (Judul dkk., 2025)