

**PERBANDINGAN PERAMALAN SARIMA DAN *PROPHET* PADA  
JUMLAH PERJALANAN WISATAWAN NUSANTARA  
MENURUT PROVINSI TUJUAN (PERJALANAN)**

**(Skripsi)**

**Oleh**

**REIVALDO JUNIZAR PRAYITNO  
NPM 2217031079**



**JURUSAN MATEMATIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2026**

## **ABSTRACT**

### **COMPARISON OF SARIMA AND PROPHET FORECASTING ON THE NUMBER OF DOMESTIC TOURIST TRIPS BY DESTINATION PROVINCE**

**By**

**REIVALDO JUNIZAR PRAYITNO**

Time series analysis is a method that studies data based on chronological order for forecasting purposes, in which Seasonal Autoregressive Integrated Moving Average (SARIMA) and prophet represent two different approaches. SARIMA focuses on stochastic dependence structures and seasonality, while prophet emphasizes the separate decomposition of trend and seasonal components. This study aims to compare the performance of both methods to forecast the number of domestic tourist trips at the national level using monthly data from January 2019 to December 2025. Based on the analysis, the best model obtained is SARIMA(0,1,1)(2,1,1)<sub>12</sub>, which has fulfilled the residual diagnostic assumptions. The out-of-sample evaluation results show that SARIMA is more accurate with a Mean Absolute Percentage Error (MAPE) value of 6.71%, compared to prophet at 8.02%. Overall, the forecasting results project that the number of domestic tourist trips in 2026 will remain high, with peak travel occurring in April and December.

Key words : SARIMA, Prophet, Forecasting, Domestic Tourist Trips.

## ABSTRAK

### PERBANDINGAN PERAMALAN SARIMA DAN *PROPHET* PADA JUMLAH PERJALANAN WISATAWAN NUSANTARA MENURUT PROVINSI TUJUAN (PERJALANAN)

Oleh

REIVALDO JUNIZAR PRAYITNO

Analisis deret waktu merupakan metode yang mempelajari data berdasarkan urutan waktu untuk keperluan peramalan, di mana *Seasonal Autoregressive Integrated Moving Average* (SARIMA) dan *prophet* merepresentasikan dua pendekatan yang berbeda. SARIMA berfokus pada struktur ketergantungan stokastik dan musiman, sedangkan *prophet* menekankan dekomposisi komponen tren dan musiman secara terpisah. Penelitian ini bertujuan membandingkan kinerja kedua metode tersebut untuk meramalkan jumlah perjalanan wisatawan nusantara tingkat nasional menggunakan data bulanan periode Januari 2019 hingga Desember 2025. Berdasarkan hasil analisis, diperoleh model terbaik yaitu SARIMA(0,1,1)(2,1,1)<sub>12</sub> yang telah memenuhi asumsi diagnostik residual. Hasil evaluasi *out-of-sample* menunjukkan SARIMA lebih akurat dengan nilai *Mean Absolute Percentage Error* (MAPE) sebesar 6,71%, dibandingkan *prophet* yang sebesar 8,02%. Secara keseluruhan, hasil peramalan memproyeksikan jumlah perjalanan wisatawan nusantara pada tahun 2026 tetap tinggi, dengan puncak perjalanan terjadi pada bulan April dan Desember.

Kata kunci : SARIMA, *Prophet*, Peramalan, Perjalanan Wisatawan Nusantara.

**PERBANDINGAN PERAMALAN SARIMA DAN *PROPHET* PADA  
JUMLAH PERJALANAN WISATAWAN NUSANTARA  
MENURUT PROVINSI TUJUAN (PERJALANAN)**

Oleh

**REIVALDO JUNIZAR PRAYITNO**

Skripsi

**Sebagai Salah Satu Syarat Untuk Mencapai Gelar  
SARJANA MATEMATIKA**

Pada

**Jurusan Matematika  
Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2026**

Judul Skripsi

: **PERBANDINGAN PERAMALAN SARIMA  
DAN *PROPHET* PADA JUMLAH  
PERJALANAN WISATAWAN NUSANTARA  
MENURUT PROVINSI TUJUAN  
(PERJALANAN)**

Nama Mahasiswa

: **Reivaldo Junizar Prayitno**

Nomor Pokok Mahasiswa

: **2217031079**

Program Studi

: **Matematika**

Fakultas

: **Matematika dan Ilmu Pengetahuan Alam**



**Drs. Nusyirwan, M.Si.**

**NIP. 196610101992031028**

**Dina Eka Nurvazly, S.Pd., M.Si.**

**NIP. 199311062019032018**

**2. Wakil Dekan Bidang Akademik dan Kerjasama,  
FMIPA Universitas Lampung**

**Mulyono, S.Si., M.Si., Ph.D.**

**NIP. 197406112000031002**

**MENGESAHKAN**

I. **Tim Penguji**

**Ketua**

**: Drs Nusyirwan, M.Si.**

**Sekretaris**

**: Dina Eka Nurvazly, S.Pd., M.Si.**

**Penguji**

**Bukan Pembimbing**

**: Widiarti, S.Si., M.Si.**

**Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam**



**Dr. Eng. Heri Satria, S.Si., M.Si.**

**NIP.197110012005011002**

**Tanggal Lulus Ujian Skripsi : 18 Mei 2026**

## PERNYATAAN SKRIPSI MAHASISWA

Saya yang bertanda tangan di bawah ini :

Nama : Reivaldo Junizar Prayitno  
Nomor Pokok Mahasiswa : 2217031079  
Jurusan : Matematika  
Judul Skripsi : **Perbandingan Peramalan SARIMA dan  
*Prophet* Jumlah Perjalanan Wisatawan  
Nusantara Menurut Provinsi Tujuan  
(Perjalanan)**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri dan semua yang tertuang dalam skripsi ini telah mengikuti kaidah karya penulisan ilmiah Universitas Lampung.

Bandar Lampung, 18 Mei 2026

Penulis..



**Reivaldo Junizar Prayitno**

## RIWAYAT HIDUP

Penulis bernama lengkap Reivaldo Junizar Prayitno, lahir di Kota Jakarta Pusat pada tanggal 19 Juni 2004. Penulis merupakan anak kedua dari dua bersaudara yang lahir dari pasangan Bapak Benyamin Prayitno dan Ibu Sharlene Zein. Penulis memiliki seorang kakak perempuan bernama Evania Besariani Prayitno.

Penulis memulai pendidikan di TK Negeri Tegal Kota Jakarta Pusat pada tahun 2009-2011. Kemudian, penulis melanjutkan sekolah dasar di SD Negeri Menteng 02 Kota Jakarta Pusat pada tahun 2011-2016, sekolah menengah pertama di SMP Negeri 9 Jakarta pada tahun 2016-2019, dan sekolah menengah atas di SMA Negeri 99 Jakarta pada tahun 2019-2022. Pada tahun 2022 penulis diterima sebagai mahasiswa S1 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SBMPTN (Seleksi Bersama Masuk Perguruan Tinggi Negeri).

Selama menjadi mahasiswa, penulis telah melaksanakan kerja praktik (PKL) di BRI *Regional Office* Bandar Lampung pada bulan Desember 2024. Selain itu, penulis juga mengikuti program magang di Kantor Perwakilan Bank Indonesia Provinsi Lampung pada bulan Januari 2026. Sebagai bentuk pengabdian kepada masyarakat serta pelaksanaan Tri Dharma Perguruan Tinggi, penulis melaksanakan Kuliah Kerja Nyata (KKN) periode kedua di Kelurahan Beringin Raya, Kecamatan Kemiling, Kota Bandar Lampung pada bulan Juni hingga Agustus 2025.

## **KATA INSPIRASI**

*“Hard work is worthless for those that don’t believe in themselves“*

*(Naruto Uzumaki)*

*“If you’re not scared, you’re not taking a chance. And if you’re not taking a chance, then what the hell are you doing?“ (Ted Mosby)*

*“The important thing is not to stop questioning“ (Sheldon Cooper)*

## **PERSEMBAHAN**

*Dengan mengucap Bismillahirrahmanirrahim, ku persembahkan karya sederhana ini kepada :*

### ***Orang Tuaku Tersayang***

*Terima kasih Bunda tercinta atas segala kasih sayang, doa restu, pengorbanan, perjuangan dan segala hal yang tak dapat diucapkan.*

### ***Dosen Pembimbing dan Penguji***

*Terima kasih kepada dosen pembimbing dan penguji yang sudah memberikan bimbingan, motivasi, dan ilmu yang bermanfaat kepada penulis.*

***Almamater Tercinta  
Universitas Lampung***

## SANWACANA

Alhamdulillah, puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Perbandingan Peramalan SARIMA dan *Prophet* Jumlah Perjalanan Wisatawan Nusantara Menurut Provinsi Tujuan (Perjalanan)”. Shalawat serta salam semoga senantiasa tercurah kepada Nabi Muhammad SAW sebagai suri teladan bagi umat manusia.

Penyusunan skripsi ini banyak pihak yang telah membantu, untuk itu penulis mengucapkan terima kasih kepada :

1. Bapak Drs. Nusyirwan, M.Si. selaku pembimbing satu yang telah meluangkan waktunya untuk memberikan bimbingan, motivasi, dan saran selama proses penyelesaian skripsi.
2. Ibu Dina Eka Nurvazly, S.Pd., M.Si. selaku pembimbing dua yang telah memberikan bimbingan, motivasi, dan saran selama proses penyelesaian skripsi.
3. Ibu Widiarti, S.Si., M.Si. selaku pembahas yang telah memberikan saran, kritik, serta evaluasi yang membangun guna penyempurnaan skripsi.
4. Bapak Dr. Ahmad Faisol, S.Si., M.Sc. selaku dosen pembimbing akademik.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si. selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Bapak Dr. Eng Heri Satria, S.Si., M.Si. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Seluruh dosen, *staff*, dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

8. Bunda dan Kakak tercinta atas segala kasih sayang, doa, dan pengorbanan yang tidak pernah terputus, serta dukungan yang senantiasa mengiringi penulis.
9. Keluarga besar atas dukungan moral dan doa yang telah diberikan.
10. Titah Aura Maulia yang telah memberikan dukungan, semangat, serta menemani penulis dalam proses penyusunan skripsi.
11. Perkumpulan “Portugal Gank” atas kebersamaan, dukungan, dan semangat yang telah diberikan kepada penulis selama proses penyusunan skripsi.
12. Teman-teman yang tidak dapat disebutkan satu per satu, yang telah memberikan dukungan dan semangat kepada penulis dalam penyusunan skripsi.
13. Seluruh pihak yang telah memberikan bantuan dan dukungan kepada penulis dalam penyusunan skripsi, yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa dalam penulisan skripsi ini masih jauh dari kata sempurna dan terdapat banyak kekurangan, baik dalam segi penyajian maupun teknik penulisan. Oleh karena itu, penulis sangat mengharapkan saran, masukan, dan kritik yang membangun dari pembaca untuk dapat menyempurnakan laporan ini. Semoga skripsi ini dapat memberikan informasi yang bermanfaat.

Bandar Lampung, 18 Mei 2026

Penulis,

**Reivaldo Junizar Prayitno**

## DAFTAR ISI

Halaman

|  |           |
|--|-----------|
| <b>DAFTAR TABEL .....</b>  | <b>v</b>  |
| <b>DAFTAR GAMBAR.....</b>  | <b>vi</b> |
| <b>I. PENDAHULUAN .....</b>  | <b>1</b>  |
| 1.1 Latar Belakang dan Masalah .....   | 1         |
| 1.2 Tujuan Penelitian.....   | 4         |
| 1.3 Manfaat Penelitian.....  | 5         |
| <b>II. TINJAUAN PUSTAKA .....</b>  | <b>6</b>  |
| 2.1 Peramalan .....  | 6         |
| 2.2 Analisis Deret Waktu .....   | 6         |
| 2.3 Stasioneritas.....   | 8         |
| 2.4 <i>Autocorrelation Function</i> (ACF) dan <i>Partial Autocorrelation Function</i> (PACF) ..... | 10        |
| 2.5 Metode <i>Seasonal Autoregressive Integrated Moving Average</i> (SARIMA).....                  | 13        |
| 2.5.1 Model <i>Autoregressive</i> (AR) .....   | 14        |
| 2.5.2 Model <i>Moving Average</i> (MA).....  | 15        |
| 2.5.3 Model <i>Autoregressive Moving Average</i> (ARMA) .....                                      | 15        |
| 2.5.4 Model <i>Autoregressive Integrated Moving Average</i> (ARIMA) .....                          | 16        |
| 2.5.5 Model <i>Seasonal Autoregressive Integrated Moving Average</i> (SARIMA) .....                | 17        |
| 2.6 Metode <i>Prophet</i> .....  | 18        |
| 2.6.1 Komponen Tren.....   | 19        |
| 2.6.2 Komponen Musiman .....   | 20        |
| 2.6.3 Model <i>Prophet</i> .....   | 21        |
| 2.7 Kriteria Pemilihan Model .....   | 22        |
| 2.8 <i>Exact Maximum Likelihood</i> (EML).....   | 23        |
| 2.9 Pengujian Diagnostik Residual.....   | 25        |
| 2.9.1 Residual Bersifat Acak.....  | 26        |
| 2.9.2 Residual Bersifat Normal .....   | 27        |
| 2.10 <i>Maximum a Posteriori</i> (MAP).....  | 28        |

|   |           |
|---|-----------|
| 2.11 Kriteria Keباikan Model.....   | 31        |
| 2.11.1 <i>Mean Absolute Error</i> (MAE) .....                                     | 31        |
| 2.11.2 <i>Root Mean Square Error</i> (RMSE) .....                                 | 32        |
| 2.11.3 <i>Mean Absolute Percentage Error</i> (MAPE).....                          | 33        |
| 2.12 Perjalanan Wisatawan Nusantara Menurut Provinsi Tujuan<br>(Perjalanan) ..... | 34        |
| <b>III. METODOLOGI PENELITIAN.....</b>  | <b>35</b> |
| 3.1 Waktu dan Tempat Penelitian .....   | 35        |
| 3.2 Data Penelitian.....  | 35        |
| 3.3 Metode Penelitian.....  | 36        |
| <b>IV. HASIL DAN PEMBAHASAN.....</b>  | <b>42</b> |
| 4.1 Eksplorasi Data.....  | 42        |
| 4.2 Identifikasi Model SARIMA.....  | 45        |
| 4.2.1 Identifikasi Kestabilan Varians Data dan Stasioneritas Data.....            | 45        |
| 4.2.2 Transformasi dan <i>Differencing</i> Data .....                             | 48        |
| 4.2.3 Identifikasi Model SARIMA.....  | 52        |
| 4.2.4 Pendugaan Parameter dengan Metode EML dan Pemilihan Model<br>SARIMA .....   | 56        |
| 4.2.5 Uji Diagnostik Residual Model SARIMA .....                                  | 63        |
| 4.2.6 Peramalan <i>Out-of-Sample</i> Model SARIMA.....                            | 65        |
| 4.3 Pembentukan Model <i>Prophet</i> .....  | 67        |
| 4.3.1 Pembentukan Komponen Tren.....  | 67        |
| 4.3.2 Pembentukan Komponen Musiman .....  | 72        |
| 4.3.3 Model <i>Prophet</i> .....  | 77        |
| 4.3.4 Peramalan <i>Out-of-Sample</i> Model <i>Prophet</i> .....                   | 80        |
| 4.4 Perbandingan Akurasi Peramalan SARIMA dan <i>Prophet</i> .....                | 82        |
| 4.5 Peramalan 12 Bulan ke Depan (Januari–Desember 2026).....                      | 84        |
| <b>V. KESIMPULAN.....</b>   | <b>87</b> |
| <b>DAFTAR PUSTAKA.....</b>  | <b>88</b> |
| <b>LAMPIRAN.....</b>  | <b>92</b> |

## DAFTAR TABEL

| Tabel   | Halaman |
|---|---------|
| 1. Kriteria Interpretasi Nilai MAPE .....   | 33      |
| 2. Statistik Deskriptif Data .....  | 44      |
| 3. Hasil Uji ADF pada Data <i>in-sample</i> .....   | 48      |
| 4. Hasil Uji ADF pada Data <i>in-sample</i> setelah Transformasi dan<br><i>Differencing</i> Nonmusiman .....                              | 50      |
| 5. Hasil Uji ADF pada Data <i>in-sample</i> setelah Transformasi <i>Differencing</i><br>Nonmusiman, dan <i>Differencing</i> Musiman ..... | 52      |
| 6. Kandidat Model ARIMA .....   | 54      |
| 7. Estimasi Koefisien Kandidat Model SARIMA.....  | 57      |
| 8. Perbandingan Nilai <i>Log-Likelihood</i> dan AIC Kandidat Model SARIMA.....  | 57      |
| 9. Estimasi Koefisien Model SARIMA(0,1,1)(2,1,1) <sub>12</sub> .....  | 60      |
| 10. Hasil Uji <i>Ljung-Box</i> Residual Model SARIMA.....   | 63      |
| 11. Hasil Uji <i>Kolmogorov-Smirnov</i> Residual Model SARIMA.....  | 64      |
| 12. Hasil Peramalan <i>out-of-sample</i> Model SARIMA(0,1,1)(2,1,1) <sub>12</sub> .....   | 65      |
| 13. Estimasi Nilai Parameter Komponen Tren Model <i>Prophet</i> .....   | 69      |
| 14. Estimasi Komponen Musiman Tahunan Periodik ( $s = 12$ ) Model<br><i>Prophet</i> .....   | 75      |
| 15. Hasil Peramalan <i>out-of-sample</i> Model <i>Prophet</i> .....   | 80      |
| 16. Perbandingan Akurasi Peramalan Model SARIMA dan <i>Prophet</i> .....  | 83      |
| 17. Hasil Peramalan Jumlah Perjalanan Wisatawan Nusantara Tahun 2026.....   | 84      |

## DAFTAR GAMBAR

| Gambar  | Halaman |
|---|---------|
| 1. <i>Flowchart</i> Analisis Penelitian .....   | 38      |
| 2. <i>Flowchart</i> Analisis Penelitian (Cabang SARIMA) .....   | 39      |
| 3. <i>Flowchart</i> Analisis Penelitian (Cabang <i>Prophet</i> ) .....  | 40      |
| 4. <i>Flowchart</i> Analisis Penelitian (Lanjutan) .....  | 41      |
| 5. <i>Plot</i> Data .....   | 43      |
| 6. <i>Plot</i> Data <i>in-sample</i> .....  | 46      |
| 7. <i>Plot</i> ACF Data <i>in-sample</i> .....  | 47      |
| 8. <i>Plot</i> Data <i>in-sample</i> setelah Transformasi <i>Box-Cox</i> .....  | 49      |
| 9. <i>Plot</i> Data <i>in-sample</i> setelah Transformasi dan <i>Differencing</i> Nonmusiman... 50  |         |
| 10. <i>Plot</i> Data <i>in-sample</i> setelah Transformasi, <i>Differencing</i> Nonmusiman,<br>dan <i>Differencing</i> Musiman .....      | 51      |
| 11. <i>Plot</i> ACF Data <i>in-sample</i> setelah Transformasi dan <i>Differencing</i> .....  | 53      |
| 12. <i>Plot</i> PACF Data <i>in-sample</i> setelah Transformasi dan <i>Differencing</i><br>Nonmusiman .....                               | 53      |
| 13. <i>Plot</i> ACF Data <i>in-sample</i> setelah Transformasi, <i>Differencing</i><br>Nonmusiman, dan <i>Differencing</i> Musiman .....  | 55      |
| 14. <i>Plot</i> PACF Data <i>in-sample</i> setelah Transformasi, <i>Differencing</i><br>Nonmusiman, dan <i>Differencing</i> Musiman ..... | 55      |
| 15. <i>Plot</i> Data Aktual dan Peramalan <i>out-of-sample</i> Model SARIMA .....   | 66      |
| 16. Komponen Tren Model <i>Prophet</i> .....  | 71      |
| 17. Pola Komponen Musiman Tahunan Model <i>Prophet</i> .....  | 76      |
| 18. <i>Plot</i> Data Aktual dan Peramalan <i>out-of-sample</i> Model <i>Prophet</i> .....   | 81      |
| 19. Perbandingan Hasil Peramalan <i>out-of-sample</i> Model SARIMA<br>dan <i>Prophet</i> .....  | 82      |
| 20. <i>Plot</i> Data Aktual dan Peramalan Jumlah Perjalanan Wisatawan Nusantara<br>Tahun 2026 .....                                       | 85      |

# I. PENDAHULUAN

## 1.1 Latar Belakang dan Masalah

Deret waktu (*time series*) merupakan data yang dicatat secara berurutan berdasarkan waktu, sehingga nilai pada suatu periode memiliki keterkaitan dengan periode sebelumnya. Hubungan antarperiode ini memungkinkan pola perubahan dianalisis dan dimanfaatkan untuk keperluan peramalan (Shumway & Stoffer, 2017), oleh karena itu, analisis deret waktu banyak digunakan untuk memahami kecenderungan suatu fenomena dari waktu ke waktu serta memperkirakan nilai pada periode mendatang.

Tujuan utama analisis deret waktu adalah mengidentifikasi struktur data, khususnya komponen tren, musiman, dan variasi acak, sebagai dasar dalam membangun model peramalan (Chatfield, 2003). Dalam proses tersebut, pemeriksaan kestasioneran, autokorelasi, dan kestabilan ragam menjadi tahapan penting karena sangat memengaruhi ketepatan model dan akurasi hasil ramalan.

Seiring perkembangan metodologi deret waktu, berbagai pendekatan dikembangkan untuk menangkap ketergantungan data antarperiode. Model *autoregressive* (AR) dan *moving average* (MA), beserta kombinasinya dalam ARMA dan ARIMA, digunakan untuk menggambarkan hubungan dinamis antarwaktu, sedangkan *seasonal* ARIMA (SARIMA) dirancang untuk menangani pola musiman yang berulang. Di luar pendekatan *Box–Jenkins*, metode berbasis

komponen seperti *prophet* semakin banyak digunakan karena kemampuannya memodelkan tren dan musiman secara fleksibel (Box & Jenkins, 1970).

Dalam penelitian ini, metode SARIMA dan *prophet* digunakan dan dibandingkan karena keduanya merepresentasikan pendekatan yang berbeda dalam memodelkan deret waktu. SARIMA dipilih karena mampu menangkap pola musiman serta hubungan antarbulan melalui komponen *autoregressive*, *moving average*, dan *differencing* musiman, sehingga sesuai untuk data bulanan dengan pola berulang (Wei, 2006).

Sebagai pembanding, *prophet* menggunakan pendekatan berbasis dekomposisi komponen, yaitu memisahkan deret waktu ke dalam tren dan musiman. Tren dimodelkan sebagai garis linier bertahap yang dapat berubah kemiringannya pada titik tertentu (*change points*), sedangkan musiman direpresentasikan menggunakan deret *Fourier* (Taylor & Letham, 2018). Dalam penelitian ini, *prophet* diterapkan tanpa memasukkan efek hari besar agar evaluasi model sepenuhnya didasarkan pada pola yang terbentuk dari data historis.

Kinerja kedua metode kemudian dibandingkan melalui ukuran kesalahan peramalan, yaitu *mean absolute error* (MAE), *root mean square error* (RMSE), dan *mean absolute percentage error* (MAPE), serta melalui pemeriksaan residual untuk memastikan kelayakan model sebelum digunakan dalam peramalan (Hyndman & Athanasopoulos, 2021).

Penerapan analisis deret waktu dalam bidang pariwisata umumnya dilakukan pada data jumlah perjalanan wisatawan yang dicatat secara berkala. Data jumlah perjalanan wisatawan yang dipublikasikan oleh Badan Pusat Statistik disajikan dalam bentuk deret waktu bulanan dan menunjukkan pola musiman yang cukup jelas, seperti peningkatan perjalanan pada periode hari raya, libur sekolah, dan akhir tahun (Badan Pusat Statistik, 2025). Pola tersebut berpengaruh langsung terhadap

kebutuhan transportasi, akomodasi, serta layanan pendukung lainnya, sehingga peramalan menjadi penting sebagai dasar perencanaan.

Selain pola musiman tahunan, data jumlah perjalanan wisatawan nusantara periode 2019–2025 juga menunjukkan adanya perubahan struktur yang cukup signifikan, terutama pada masa pandemi COVID-19 dan fase pemulihan setelahnya. Periode tersebut ditandai dengan penurunan tajam yang kemudian diikuti peningkatan kembali secara bertahap. Kondisi ini menimbulkan tantangan dalam pemodelan karena model harus mampu menangkap pola musiman yang relatif stabil sekaligus menyesuaikan perubahan tren yang tidak berlangsung secara linier (Hyndman & Athanasopoulos, 2021).

Secara metodologis, SARIMA dan *prophet* merepresentasikan dua pendekatan yang berbeda dalam analisis deret waktu. SARIMA berfokus pada struktur ketergantungan stokastik melalui komponen *autoregressive* dan *moving average*, sedangkan *prophet* menekankan dekomposisi tren dan musiman dalam kerangka *Bayesian* dengan regularisasi parameter. Perbandingan kedua metode ini menjadi relevan untuk menilai apakah dinamika jumlah perjalanan wisatawan nusantara lebih dipengaruhi oleh ketergantungan antarperiode atau oleh perubahan tren dan pola musiman yang dapat dimodelkan secara terpisah (Taylor & Letham, 2018).

Beberapa penelitian sebelumnya telah membandingkan SARIMA dan *prophet* pada data pariwisata. Bouhaddour *et al.* (2023) menunjukkan bahwa kedua metode mampu menangkap pola musiman dengan baik pada data kunjungan wisatawan internasional di Singapura, meskipun perbedaan kinerja tidak terlalu signifikan pada data dengan musiman yang stabil. Pada tingkat regional di Indonesia, Rizal (2025) menemukan bahwa SARIMA dan *prophet* memberikan kinerja yang relatif sebanding dalam peramalan jumlah wisatawan di Provinsi Nusa Tenggara Barat, sementara itu, Nur Aziza *et al.* (2023) menunjukkan bahwa SARIMA memberikan akurasi yang lebih baik dibandingkan *prophet* pada data penumpang domestik yang dipengaruhi kebijakan mobilitas.

Penelitian di luar sektor pariwisata juga menunjukkan hasil yang beragam. Feng *et al.* (2022) menemukan bahwa meskipun metode *machine learning* tertentu memberikan akurasi lebih tinggi, SARIMA dan *prophet* tetap relevan sebagai model pembanding karena kemampuannya menangkap pola musiman dasar. Negre *et al.* (2024) menunjukkan bahwa pada data dengan perubahan tren bertahap, model berbasis komponen seperti *prophet* dapat memberikan performa yang lebih baik dibandingkan pendekatan *Box-Jenkins* klasik.

Meskipun demikian, penelitian yang secara khusus membandingkan SARIMA dan *prophet* pada data bulanan jumlah perjalanan wisatawan nusantara tingkat nasional, terutama yang mencakup periode sebelum pandemi, masa pembatasan mobilitas, dan fase pemulihan secara simultan, masih terbatas. Padahal, karakteristik data yang mengalami perubahan tren signifikan sekaligus mempertahankan pola musiman tahunan yang kuat memerlukan evaluasi empiris untuk menentukan metode yang paling sesuai, oleh karena itu, penelitian ini dilakukan untuk membandingkan secara sistematis kinerja SARIMA dan *prophet* pada data nasional guna mengidentifikasi model yang paling akurat dan stabil dalam merepresentasikan dinamika perjalanan wisatawan nusantara.

## 1.2 Tujuan Penelitian

Adapun tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut:

1. Membandingkan model SARIMA dan *prophet* pada data deret waktu jumlah perjalanan wisatawan nusantara tingkat nasional.
2. Menilai akurasi peramalan kedua model menggunakan MAE, RMSE, dan MAPE, serta melakukan pengujian diagnostik residual pada model SARIMA untuk memastikan kelayakan model sebelum digunakan dalam peramalan.

### 1.3 Manfaat Penelitian

Adapun manfaat penelitian yang dilakukan adalah sebagai berikut:

1. Memberikan informasi hasil peramalan jumlah perjalanan wisatawan nusantara yang dapat digunakan sebagai bahan pertimbangan dalam pengambilan keputusan di sektor pariwisata.
2. Menyediakan dasar kuantitatif dalam perencanaan kapasitas transportasi, akomodasi, dan destinasi melalui model peramalan yang telah dievaluasi tingkat akurasinya.
3. Memberikan kontribusi metodologis melalui penerapan dan perbandingan model SARIMA dan *prophet* pada data pariwisata nasional, sehingga dapat menjadi referensi dalam analisis deret waktu serupa.

## **II. TINJAUAN PUSTAKA**

### **2.1 Peramalan**

Peramalan merupakan proses penaksiran nilai suatu variabel pada periode mendatang dengan memanfaatkan informasi yang terkandung dalam data historis melalui pendekatan statistik (Sudjana, 1986). Dalam konteks deret waktu, peramalan dilakukan dengan membangun model yang merepresentasikan struktur ketergantungan antarperiode sehingga pola perubahan yang telah terjadi dapat digunakan sebagai dasar untuk memperkirakan nilai pada periode berikutnya.

Tujuan utama peramalan adalah mengurangi ketidakpastian terhadap nilai di masa mendatang melalui pendekatan kuantitatif yang terukur. Pada data deret waktu, urutan waktu memungkinkan identifikasi komponen tren, pola musiman, serta variasi acak yang memengaruhi dinamika data. Setelah struktur tersebut dipahami dan dimodelkan, model digunakan untuk menghasilkan nilai ramalan pada periode di luar data pengamatan, yang selanjutnya dapat dievaluasi melalui ukuran kesalahan peramalan (Hyndman & Athanasopoulos, 2021).

### **2.2 Analisis Deret Waktu**

Analisis deret waktu mempelajari data yang dicatat secara berurutan menurut waktu, dengan nilai pada suatu periode umumnya dipengaruhi oleh nilai pada periode sebelumnya. Ketergantungan antarperiode ini menjadi ciri utama deret

waktu dan perlu diperhatikan karena menentukan struktur dinamika data serta menjadi dasar dalam penyusunan model peramalan (Shumway & Stoffer, 2017). Pada tahap awal analisis, pola dasar biasanya diamati melalui grafik dan ringkasan statistik untuk mengidentifikasi kecenderungan jangka panjang (*trend*), pola musiman yang berulang, variasi siklis, serta komponen acak yang tidak teratur.

Keempat komponen tersebut, yaitu tren, variasi musiman, variasi siklis, dan komponen acak, membantu menjelaskan bagaimana suatu deret berkembang dari waktu ke waktu. Tren menunjukkan arah umum pergerakan data dalam jangka panjang, sedangkan variasi musiman merepresentasikan pola berulang yang terikat pada periode kalender tertentu. Variasi siklis menggambarkan gelombang jangka panjang dengan durasi lebih dari satu tahun yang tidak selalu teratur, sementara komponen tidak teratur mencerminkan fluktuasi acak atau kejadian tak terduga. Dalam pemodelan, komponen acak inilah yang diharapkan tersisa sebagai residual setelah struktur utama data berhasil ditangkap oleh model (Chatfield, 2003).

Dalam keperluan pemodelan berbasis *Box-Jenkins*, deret waktu perlu memenuhi asumsi kestasioneran. Apabila rata-rata berubah dari waktu ke waktu, dilakukan *differencing* hingga deret mencapai kondisi stasioner dalam rata-rata, yang dapat diidentifikasi melalui pengamatan *plot* dan pengujian formal seperti uji *Augmented Dickey-Fuller*. Jika ragam meningkat seiring kenaikan nilai pengamatan, digunakan transformasi seperti logaritma atau *Box-Cox* untuk menstabilkan variansi. Setelah transformasi dan *differencing* dilakukan, pola ketergantungan jangka pendek maupun musiman diidentifikasi melalui *autocorrelation function* (ACF) dan *partial autocorrelation function* (PACF), yang memberikan indikasi mengenai keberadaan komponen *autoregressive* (AR), *moving average* (MA), serta komponen musiman (Box *et al.*, 2015).

Setelah model ditetapkan dan parameternya diestimasi, residual dianalisis untuk memastikan bahwa tidak terdapat pola sistematis yang tersisa. Pemeriksaan ini umumnya dilakukan melalui uji *Ljung-Box* untuk menilai keacakan residual serta

uji distribusi untuk melihat kesesuaiannya terhadap asumsi normalitas. Model dinilai layak digunakan untuk peramalan apabila residual bersifat *white noise* dan tidak menunjukkan struktur ketergantungan yang tersisa (Hyndman & Athanasopoulos, 2021).

### 2.3 Stasioneritas

Stasioneritas menggambarkan keadaan ketika sifat dasar suatu deret tidak berubah dari waktu ke waktu. Deret yang stasioner memiliki *mean* yang relatif stabil, ragam yang tidak meningkat, serta autokovarians yang hanya bergantung pada selang waktu (*lag*) dan bukan pada waktu pengamatan tertentu. Deret dengan karakteristik tersebut lebih mudah dianalisis karena hubungan antar nilai dari waktu ke waktu menjadi lebih teratur (Shumway & Stoffer, 2017).

Pemeriksaan awal biasanya dilakukan melalui *plot* deret untuk melihat adanya perubahan *mean*, peningkatan ragam, atau pola musiman yang kuat, selanjutnya, pengujian formal seperti *Augmented Dickey–Fuller* (ADF) digunakan untuk menilai keberadaan akar unit yang menunjukkan bahwa deret belum stasioner. Dengan mengetahui keberadaan akar unit, tindakan seperti transformasi atau *differencing* dapat diterapkan secara tepat agar asumsi yang diperlukan dalam pemodelan dapat terpenuhi (Hyndman & Athanasopoulos, 2021).

Dalam penelitian ini digunakan salah satu uji akar unit, yaitu uji *Augmented Dickey–Fuller* (ADF). Uji ADF dikembangkan untuk mengatasi autokorelasi pada galat dengan menambahkan sejumlah *lag* dari beda pertama. Bentuk umum uji ADF dituliskan pada persamaan berikut (Said & Dickey, 1984):

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \sum_{i=1}^L \alpha_i \Delta y_{t-i} + u_t \quad (2.1)$$

dengan:

|                              |  |
|------------------------------|--|
| $y_t$                        | = nilai deret waktu pada periode $t$ ,                   |
| $\Delta y_t = y_t - y_{t-1}$ | = beda pertama,  |
| $\beta_1$                    | = konstanta ( <i>drift</i> ),                            |
| $\beta_2 t$                  | = komponen tren deterministik,                           |
| $\delta$                     | = parameter utama yang diuji untuk mendeteksi akar unit, |
| $\alpha_i$                   | = koefisien <i>lag</i> dari beda pertama,                |
| $L$                          | = jumlah <i>lag</i> yang ditambahkan,                    |
| $u_t$                        | = galat acak regresi.                                    |

Parameter utama dalam pengujian ini adalah  $\delta$ . Jika  $\delta = 0$ , maka deret mengandung akar unit dan bersifat tidak stasioner, sebaliknya, jika  $\delta < 0$ , deret bersifat stasioner.

Pengambilan keputusan dilakukan melalui pengujian hipotesis:

$H_0: \delta = 0$  (deret memiliki akar unit, tidak stasioner)

$H_1: \delta < 0$  (deret stasioner)

Statistik uji yang digunakan adalah statistik  $\tau$ . Dalam konteks *Dickey-Fuller* dan ADF, statistik  $\tau$  merupakan *t-statistic* dari parameter  $\delta$ . Keputusan uji diperoleh dengan membandingkan nilai statistik  $\tau$  dengan nilai kritis yang dikembangkan oleh *MacKinnon* pada taraf nyata tertentu. Jika nilai statistik lebih kecil daripada nilai kritis, maka  $H_0$  ditolak dan deret dikatakan stasioner. Pemilihan jumlah *lag*  $L$  dilakukan secukupnya, misalnya menggunakan AIC, sehingga galat pada model ADF tidak lagi menunjukkan autokorelasi dan hasil uji menjadi lebih valid (*MacKinnon*, 1996).

Setelah dilakukan uji akar unit dan deret belum stasioner, beberapa langkah dapat diterapkan untuk membuat pola perubahannya lebih konsisten dari waktu ke waktu. Langkah yang dilakukan yaitu transformasi *Box-Cox* agar ragam stabil dan struktur data lebih mudah dianalisis. Bentuk umum transformasi tersebut diberikan pada persamaan berikut (*Box & Cox*, 1964):

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 0, \\ \log(y), & \lambda = 0 \end{cases} \quad (2.2)$$

dengan:

- $y$  = nilai asli (positif),
- $y^{(\lambda)}$  = hasil transformasi *Box-Cox*,
- $\lambda$  = parameter transformasi.

Selain transformasi, *differencing* dapat digunakan untuk mengatasi perubahan *mean* yang bergerak secara perlahan. *Differencing* nonmusiman dan *differencing* musiman diterapkan sesuai pola yang tampak pada data hingga deret mendekati stasioner. Setelah deret cukup stabil, proses identifikasi model dapat dilakukan melalui pola ACF dan PACF, kemudian dilanjutkan dengan tahap estimasi dan evaluasi model (Hyndman & Athanasopoulos, 2021).

#### 2.4 *Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF)*

Dalam analisis deret waktu, *autocorrelation function (ACF)* dan *partial autocorrelation function (PACF)* digunakan untuk memeriksa ketergantungan nilai deret pada berbagai *lag*. Keduanya merupakan alat utama dalam tahap identifikasi model *autoregressive (AR)* dan *moving average (MA)* sebagaimana dijelaskan dalam pendekatan *Box-Jenkins* (Box & Jenkins, 1970). Setelah deret mendekati kondisi stasioner, pola pada ACF dan PACF memberikan petunjuk awal mengenai struktur *lag* yang relevan.

Langkah pertama dalam memahami autokorelasi adalah mendefinisikan *autocovariance*. *Autocovariance* pada *lag k* menggambarkan kovarians antara nilai deret waktu pada periode  $t$  dan  $t + k$ , yang dinyatakan sebagai fungsi dari *lag k* dan tidak bergantung pada waktu  $t$  pada deret yang stasioner (Wei, 2006).

*Autocovariance* antara nilai deret waktu pada periode  $t$  dan  $t + k$  diberikan oleh persamaan berikut (Wei, 2006):

$$\gamma_k = E[(y_t - \mu)(y_{t+k} - \mu)] \quad (2.3)$$

dengan:

- $y_t$  = nilai deret waktu pada periode  $t$ ,
- $y_{t+k}$  = nilai deret waktu pada periode  $t + k$ ,
- $\mu$  = *mean* proses,
- $E[\cdot]$  = *expected value*,
- $\gamma_k$  = *autocovariance lag k*,
- $k$  = *lag*.

*Autocorrelation* diperoleh dengan menormalkan *autocovariance* terhadap variansi proses. *Autocorrelation* pada lag  $k$  didefinisikan sebagai berikut (Wei, 2006):

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad (2.4)$$

dengan:

- $\rho_k$  = *autocorrelation lag k*,
- $\gamma_0$  = variansi proses.

*Autocorrelation* sampel dihitung menggunakan persamaan berikut (Wei, 2006):

$$\hat{\rho}_k = \frac{\sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} \quad (2.5)$$

dengan:

- $\hat{\rho}_k$  = *sample autocorrelation lag k*,
- $y_t$  = observasi pada periode  $t$ ,
- $\bar{y}$  = *sample mean*,
- $k$  = *lag*,  $k = 1, 2, \dots, T - 1$
- $T$  = jumlah observasi pada deret waktu.

Plot  $\hat{\rho}_k$  digunakan untuk mengidentifikasi pola *cut-off* maupun peluruhan bertahap (*tailing off*) yang menjadi indikasi awal keberadaan komponen AR atau MA.

*Partial autocorrelation function* (PACF) mengukur hubungan antara  $y_t$  dan  $y_{t-k}$  setelah pengaruh semua *lag* di antaranya dihilangkan. Nilai PACF pada *lag*  $k$  didefinisikan sebagai berikut (Levinson, 1947):

$$\phi_{kk} = \text{Corr}(y_t, y_{t-k} \mid y_{t-1}, y_{t-2}, \dots, y_{t-k+1}) \quad (2.6)$$

dengan:

$$\phi_{kk} = \text{partial autocorrelation lag } k.$$

Nilai PACF pada data sampel, diperoleh dengan menggunakan rekurensi *Levinson–Durbin* sebagaimana diperkenalkan oleh Levinson (1947). Penduga PACF pada *lag*  $k$  diberikan oleh persamaan berikut:

$$\hat{\phi}_{kk} = \frac{\hat{\rho}_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} \hat{\rho}_j} \quad (2.7)$$

dengan:

$$\hat{\phi}_{kk} = \text{penduga partial autocorrelation lag } k,$$

$$\hat{\rho}_k = \text{sample autocorrelation lag } k,$$

$$\hat{\phi}_{k-1,j} = \text{koefisien rekurensi orde sebelumnya,}$$

$$j = \text{indeks penjumlahan}$$

Koefisien lainnya pada orde  $k$  diperbarui menggunakan persamaan berikut (Levinson, 1947):

$$\hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk} \hat{\phi}_{k-1,k-j}, \quad j = 1, \dots, k-1 \quad (2.8)$$

dengan:

$$\hat{\phi}_{k,j} = \text{koefisien AR pada orde } k,$$

$$\hat{\phi}_{k-1,j} = \text{koefisien orde sebelumnya,}$$

$\hat{\phi}_{k-1,k-j}$  = koefisien simetris orde sebelumnya.

ACF dan PACF kemudian dianalisis bersama untuk mengidentifikasi model. Pola *cut-off* pada PACF menunjukkan dugaan model AR, sedangkan *cut-off* pada ACF menunjukkan dugaan model MA. Pada deret musiman, puncak berulang di kelipatan periode seperti *lag* 12 sering menunjukkan keberadaan komponen musiman (Box *et al.*, 2015).

## 2.5 Metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA)

Metode *seasonal autoregressive integrated moving average* (SARIMA) merupakan pengembangan dari ARIMA yang digunakan untuk memodelkan deret waktu dengan pola musiman yang berulang. Pendekatan ini diperkenalkan dalam kerangka *Box-Jenkins* dengan menambahkan komponen musiman ke dalam struktur ARIMA sehingga pola yang muncul setiap  $s$  periode dapat dimasukkan ke dalam pemodelan (Box & Jenkins, 1970).

Model SARIMA menggabungkan komponen *autoregressive* (AR), *moving average* (MA), dan *differencing* nonmusiman dengan komponen musiman berupa *autoregressive* musiman dan *moving average* musiman. Komponen AR menggambarkan ketergantungan nilai saat ini terhadap beberapa nilai sebelumnya, sedangkan komponen MA merepresentasikan pengaruh galat dari periode terdahulu. Jika deret belum stasioner, dilakukan *differencing* nonmusiman; dan apabila terdapat pola berulang setiap  $s$  periode, diterapkan *differencing* musiman (Wei, 2006).

Pembentukan model SARIMA mengikuti prosedur *Box-Jenkins*, yaitu identifikasi, estimasi, dan pemeriksaan diagnostik. Pada tahap identifikasi, pola ACF dan PACF diamati untuk memperoleh dugaan awal struktur nonmusiman dan musiman. Tahap berikutnya adalah estimasi parameter, kemudian residual diperiksa untuk

memastikan tidak terdapat autokorelasi yang tersisa. Jika residual masih menunjukkan pola tertentu, struktur model perlu ditinjau kembali (Hyndman & Athanasopoulos, 2021).

Model SARIMA digunakan ketika data menunjukkan pola musiman yang relatif konsisten dari tahun ke tahun, seperti pada deret waktu bulanan yang memiliki puncak pada bulan tertentu. Dengan memasukkan *differencing* musiman serta komponen AR dan MA musiman, model ini mampu merepresentasikan ketergantungan jangka pendek sekaligus pola musiman secara simultan (Box *et al.*, 2015).

### 2.5.1 Model *Autoregressive* (AR)

Model *autoregressive* (AR) menyatakan bahwa nilai deret waktu pada periode  $t$  dipengaruhi oleh beberapa nilai sebelumnya. Model ini digunakan untuk menggambarkan ketergantungan jangka pendek antarperiode dalam suatu deret yang telah stasioner. Bentuk umum model AR orde  $p$ , atau  $AR(p)$ , diberikan oleh (Wei, 2006):

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t \quad (2.9)$$

dengan:

$$\begin{aligned} y_t &= \text{nilai deret pada waktu } t, \\ \phi_1, \dots, \phi_p &= \text{koefisien } \textit{autoregressive}, \\ a_t &= \text{galat acak (} \textit{innovation error} \text{) yang diasumsikan } \textit{white noise}. \end{aligned}$$

Dalam bentuk operator *backshift*  $B$ , dengan  $By_t = y_{t-1}$ , model  $AR(p)$  dapat dituliskan secara ringkas sebagai:

$$\phi(B)y_t = a_t \quad (2.10)$$

dengan  $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ .

### 2.5.2 Model *Moving Average* (MA)

Model *moving average* (MA) menyatakan bahwa nilai deret waktu pada periode  $t$  ditentukan oleh *innovation* saat ini dan beberapa *innovation* pada periode sebelumnya, di mana *innovation* merupakan galat acak (*white noise*) yang tidak dapat dijelaskan oleh nilai masa lalu. Model ini digunakan untuk merepresentasikan pengaruh kejadian acak jangka pendek terhadap nilai observasi. Bentuk umum model MA orde  $q$ , atau MA( $q$ ), diberikan oleh (Wei, 2006):

$$y_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (2.11)$$

dengan:

$$\begin{aligned} y_t &= \text{nilai deret waktu pada periode } t, \\ \theta_1, \dots, \theta_q &= \text{koefisien } \textit{moving average}, \\ a_t &= \text{galat acak (} \textit{innovation error} \text{) yang diasumsikan } \textit{white noise}. \end{aligned}$$

Dalam bentuk operator *backshift*  $B$ , model MA( $q$ ) dapat dituliskan sebagai:

$$y_t = \theta(B)a_t \quad (2.12)$$

dengan  $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$ .

### 2.5.3 Model *Autoregressive Moving Average* (ARMA)

Model *autoregressive moving average* (ARMA) menggabungkan komponen *autoregressive* (AR) dan *moving average* (MA) untuk merepresentasikan ketergantungan nilai deret waktu terhadap nilai sebelumnya maupun terhadap *innovation* pada beberapa periode terdahulu. Model ini digunakan untuk deret

waktu yang telah memenuhi asumsi stasioneritas dan tidak memerlukan proses *differencing* tambahan. Model ARMA orde  $(p, q)$ , atau ARMA( $p, q$ ), diberikan oleh (Wei, 2006):

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + a_t + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} \quad (2.13)$$

Dalam bentuk operator *backshift*, model tersebut dapat dituliskan secara ringkas sebagai (Box & Jenkins, 1970):

$$\phi(B)y_t = \theta(B)a_t \quad (2.14)$$

dengan:

- $y_t$  = nilai deret waktu pada periode  $t$ ,
- $a_t$  = galat acak (*innovation error*) yang diasumsikan *white noise*,
- $B$  = operator *backshift* (*lag operator*),
- $p, q$  = orde AR dan MA.

Polinomial  $\phi(B)$  memuat koefisien *autoregressive*  $\phi_1, \dots, \phi_p$ , sedangkan polinomial  $\theta(B)$  memuat koefisien *moving average*  $\theta_1, \dots, \theta_q$  (Box & Jenkins, 1970).

#### 2.5.4 Model *Autoregressive Integrated Moving Average* (ARIMA)

Model *autoregressive integrated moving average* (ARIMA) digunakan untuk menganalisis deret waktu yang belum stasioner, tetapi dapat dibuat stasioner melalui proses *differencing*, setelah *differencing* diterapkan, struktur ketergantungan terhadap nilai pada beberapa periode sebelumnya serta terhadap *innovation* pada beberapa *lag* dapat dimodelkan melalui kombinasi komponen *autoregressive* (AR) dan *moving average* (MA). Bentuk umum ARIMA orde  $(p, d, q)$  dinyatakan dengan operator *backshift* sebagai berikut (Box & Jenkins, 1970):

$$\phi(B)\nabla^d y_t = \theta(B)a_t \quad (2.15)$$

dengan:

- $y_t$  = nilai deret waktu pada periode  $t$ ,
- $a_t$  = galat acak (*innovation error*) yang diasumsikan *white noise*,
- $B$  = operator *backshift* (*lag operator*),
- $\nabla = 1 - B$  = operator *differencing* nonmusiman orde satu,
- $d$  = orde *differencing* nonmusiman.

Fungsi  $\phi(B)$  merupakan polinomial *autoregressive* yang dibentuk dari koefisien  $\phi_1, \dots, \phi_p$ , sedangkan  $\theta(B)$  merupakan polinomial *moving average* yang dibentuk dari koefisien  $\theta_1, \dots, \theta_q$  (Box & Jenkins, 1970).

Dengan demikian, ARIMA( $p, d, q$ ) dapat dipandang sebagai model ARMA( $p, q$ ) yang diterapkan pada deret hasil *differencing* sebanyak  $d$  kali. Pendekatan ini memungkinkan pola perubahan yang tidak stasioner pada deret awal dimodelkan secara lebih terstruktur setelah komponen tren atau *unit root* dihilangkan melalui *differencing* (Box & Jenkins, 1970).

### 2.5.5 Model *Seasonal Autoregressive Integrated Moving Average* (SARIMA)

Model *seasonal autoregressive integrated moving average* (SARIMA) merupakan pengembangan dari ARIMA yang menambahkan komponen musiman sehingga pola yang berulang setiap  $s$  periode dapat dimodelkan bersamaan dengan pola nonmusiman. Model ini digunakan ketika deret waktu menunjukkan pola musiman yang jelas, seperti pola tahunan pada data bulanan. Dengan menerapkan *differencing* nonmusiman dan *differencing* musiman, SARIMA mampu merepresentasikan ketergantungan jangka pendek sekaligus pola musiman yang muncul secara teratur.

Bentuk umum SARIMA  $(p, d, q)(P, D, Q)_s$  dinyatakan sebagai berikut (Box & Jenkins, 1970):

$$\phi(B) \Phi(B^s) \nabla^d \nabla_s^D y_t = \theta(B) \Theta(B^s) a_t \quad (2.16)$$

dengan:

- $y_t$  = nilai deret waktu pada periode  $t$ ,
- $a_t$  = galat acak (*innovation error*) yang diasumsikan *white noise*,
- $B$  = operator *backshift* (*lag operator*),
- $\nabla = 1 - B$  = operator *differencing* nonmusiman orde satu,
- $\nabla_s = 1 - B^s$  = operator *differencing* musiman orde satu,
- $d$  = orde *differencing* nonmusiman,
- $D$  = orde *differencing* musiman,
- $s$  = panjang periode musiman.

Fungsi  $\phi(B)$  dan  $\theta(B)$  masing-masing merupakan polinomial *autoregressive* dan *moving average* nonmusiman yang dibentuk dari koefisien  $\phi_1, \dots, \phi_p$  dan  $\theta_1, \dots, \theta_q$ . Fungsi  $\Phi(B^s)$  dan  $\Theta(B^s)$  merupakan polinomial *autoregressive* dan *moving average* musiman yang dibentuk dari koefisien  $\Phi_1, \dots, \Phi_P$  dan  $\Theta_1, \dots, \Theta_Q$  (Box & Jenkins, 1970).

## 2.6 Metode *Prophet*

*Prophet* merupakan metode peramalan deret waktu yang memodelkan data sebagai kombinasi beberapa komponen utama, yaitu tren jangka panjang, pola musiman, dan pengaruh tanggal khusus. Dengan pendekatan berbasis dekomposisi ini, perubahan dari waktu ke waktu direpresentasikan melalui pemisahan komponen yang menggambarkan arah tren dan pola berulang yang relatif stabil (Taylor & Letham, 2018).

Dalam penelitian ini, komponen tanggal khusus tidak disertakan karena daftar hari libur tidak dimasukkan sebagai variabel tambahan, oleh karena itu, model yang digunakan terdiri atas komponen tren berbentuk *piecewise linear* dengan jumlah kandidat *changepoints* yang ditentukan sebelum proses estimasi, serta komponen musiman tahunan yang dibangun menggunakan deret *Fourier* (Taylor & Letham, 2018).

Berbeda dengan pendekatan *Box-Jenkins* yang memerlukan identifikasi orde melalui analisis ACF dan PACF, *prophet* membentuk model secara langsung melalui penyesuaian komponen tren dan musiman terhadap data. Secara umum, model *prophet* dapat dinyatakan dalam bentuk aditif sebagai kombinasi komponen tren, musiman, dan galat acak. Pendekatan ini sesuai untuk deret waktu yang memiliki pola musiman relatif stabil namun memungkinkan perubahan arah tren pada waktu tertentu (Hyndman & Athanasopoulos, 2021).

### 2.6.1 Komponen Tren

Dalam *prophet*, perubahan jangka panjang pada deret waktu dimodelkan melalui komponen tren berbentuk *piecewise linear*. Tren direpresentasikan sebagai garis linier yang kemiringannya dapat berubah pada titik waktu tertentu yang disebut *changepoints*. Dengan pendekatan ini, model dapat menyesuaikan perubahan arah pertumbuhan tanpa menetapkan bentuk tren yang berbeda secara terpisah untuk setiap periode (Taylor & Letham, 2018).

Bentuk tren potongan linier dalam *prophet* dinyatakan sebagai:

$$g(t) = (b + \mathbf{z}(t)^\top \boldsymbol{\kappa}) t + (m + \mathbf{z}(t)^\top \boldsymbol{\eta}) \quad (2.17)$$

dengan:

- $g(t)$  = nilai komponen tren pada waktu  $t$ ,
- $b$  = laju perubahan awal (*slope* awal),
- $m$  = intersep awal,

- $\mathbf{z}(t)$  = vektor indikator titik perubahan,  
 $\boldsymbol{\kappa}$  = perubahan laju tren pada masing-masing titik perubahan,  
 $\boldsymbol{\eta}$  = penyesuaian intersep untuk menjaga kontinuitas tren.

Setiap elemen indikator didefinisikan sebagai:

$$z_j(t) = \begin{cases} 1, & t \geq s_j \\ 0, & t < s_j \end{cases} \quad (2.18)$$

dengan  $s_j$  menyatakan lokasi titik perubahan ke- $j$ .

Parameter  $b$  menggambarkan kecenderungan awal pertumbuhan atau penurunan deret, sedangkan  $\boldsymbol{\kappa}$  menangkap perubahan laju tren setelah melewati titik perubahan tertentu. Penyesuaian intersep melalui  $\boldsymbol{\eta}$  memastikan bahwa fungsi tren tetap kontinu meskipun terjadi perubahan kemiringan.

## 2.6.2 Komponen Musiman

*Prophet* memodelkan pola musiman menggunakan deret *Fourier*. Pendekatan ini menghasilkan bentuk gelombang yang fleksibel sehingga dapat mengikuti pola berulang dalam data tanpa menentukan bentuk musiman secara eksplisit. Representasi *Fourier* membantu menggambarkan puncak dan lembah musiman secara teratur. Pada deret waktu dengan pola yang relatif konsisten dari tahun ke tahun, bentuk komponen musiman dengan periode  $s$  dan orde *Fourier*  $N$  dinyatakan sebagai berikut (Taylor & Letham, 2018):

$$s(t) = \sum_{n=1}^N \left[ a_n \cos\left(\frac{2\pi n t}{s}\right) + b_n \sin\left(\frac{2\pi n t}{s}\right) \right] \quad (2.19)$$

dengan:

- $s(t)$  = nilai komponen musiman pada waktu  $t$ ,  
 $s$  = panjang periode musiman,  
 $N$  = orde *Fourier* yang menentukan tingkat kelenturan pola musiman,

- $n$  = indeks urutan komponen sinus dan kosinus yang menyusun pola musiman,
- $a_n, b_n$  = koefisien amplitudo komponen kosinus dan sinus pada harmonik ke- $n$  yang menentukan kontribusi masing-masing gelombang terhadap bentuk pola musiman.

Pemilihan  $N$  dilakukan agar komponen musiman mampu mengikuti pola berulang tanpa menyesuaikan variasi acak secara berlebihan. Orde *Fourier*  $N$  menentukan tingkat kelenturan pola musiman, yaitu kemampuan model dalam menangkap kompleksitas bentuk pola musiman melalui jumlah komponen sinus dan kosinus yang digunakan. *Prophet* menerapkan regularisasi pada koefisien musiman untuk mencegah model memetakan fluktuasi acak sebagai pola musiman (Taylor & Letham, 2018).

Pada penelitian ini, data bersifat bulanan sehingga periode musiman ditetapkan  $s = 12$ . Pasangan fungsi kosinus dan sinus hingga orde tertentu dibentuk, kemudian parameter-parameter tersebut diestimasi dari data. Dengan pendekatan ini, pola musiman seperti peningkatan pada masa Ramadan dan Idulfitri, libur sekolah, serta puncak akhir tahun dapat direpresentasikan tanpa perlu menentukan bentuk pola tersebut secara manual.

### 2.6.3 Model *Prophet*

Setelah komponen tren dan musiman dibentuk, *prophet* menyusun model dengan menjumlahkan kedua komponen tersebut bersama galat acaknya. Secara umum, model *prophet* dapat dinyatakan dalam bentuk aditif sebagai berikut:

$$y(t) = g(t) + s(t) + \varepsilon_t \quad (2.20)$$

dengan:

$$y(t) = \text{nilai deret pada waktu } t,$$

$$\begin{aligned}
 g(t) &= \text{komponen tren,} \\
 s(t) &= \text{komponen musiman,} \\
 \varepsilon_t &= \text{galat acak model.}
 \end{aligned}$$

Estimasi parameter dalam prophet dilakukan menggunakan pendekatan *Bayesian* melalui pemberian *prior* pada parameter tren dan musiman, sehingga proses penyesuaian model terhadap data dilakukan dengan regularisasi parameter. Dengan pendekatan ini, ketidakpastian komponen tren dan musiman dapat dihitung secara bersamaan dalam proses *fitting* (Taylor & Letham, 2018).

Setelah estimasi selesai, *prophet* menyediakan komponen tren dan musiman secara terpisah, sehingga arah pergerakan jangka panjang dan pola berulang dapat dianalisis secara eksplisit. Model ini kemudian digunakan untuk menghasilkan nilai ramalan pada periode di luar data pelatihan, dan akurasi hasil peramalan dievaluasi menggunakan ukuran kesalahan yang umum dalam analisis deret waktu.

## 2.7 Kriteria Pemilihan Model

Dalam penentuan model SARIMA, kriteria informasi digunakan untuk menilai keseimbangan antara kecocokan model terhadap data dan kompleksitas model yang diukur melalui jumlah parameter yang diestimasi. Salah satu kriteria yang umum digunakan adalah *Akaike Information Criterion* (AIC). Kriteria ini didasarkan pada nilai *log-likelihood* maksimum yang dikoreksi dengan penalti terhadap jumlah parameter, sehingga model yang dipilih tidak hanya memiliki kecocokan yang baik tetapi juga tidak terlalu kompleks (Akaike, 1974).

Nilai AIC pada model yang diestimasi dengan metode *maximum likelihood* dirumuskan sebagai:

$$AIC = -2\ell + 2r \quad (2.21)$$

dengan:

- $\ell$  = nilai *log-likelihood* maksimum model,
- $r$  = jumlah parameter yang diestimasi.

Model dengan nilai AIC yang lebih kecil lebih disukai dibandingkan model dengan nilai yang lebih besar. Namun, nilai AIC bersifat relatif dan hanya bermakna ketika digunakan untuk membandingkan beberapa model yang diestimasi pada data yang sama.

## 2.8 *Exact Maximum Likelihood (EML)*

Estimasi parameter pada model *seasonal autoregressive integrated moving average* (SARIMA) dalam penelitian ini dilakukan menggunakan metode *exact maximum likelihood* (EML). Pendekatan ini membentuk fungsi *likelihood* berdasarkan distribusi probabilitas penuh dari seluruh vektor pengamatan (*joint density*), yaitu distribusi gabungan dari semua observasi dalam deret waktu yang mempertimbangkan ketergantungan antarperiode, sehingga seluruh struktur kovarians proses dimanfaatkan dalam proses estimasi. Berbeda dengan pendekatan *conditional likelihood* yang mengondisikan pada sejumlah observasi awal, *exact maximum likelihood* tidak mengabaikan informasi awal proses, sehingga secara teoritis menghasilkan *estimator* yang lebih efisien (Hamilton, 1994). Dalam implementasinya, *likelihood* tersebut dapat dinyatakan dalam bentuk *innovation*, yaitu galat prediksi satu langkah ke depan yang merepresentasikan informasi baru pada setiap periode.

Misalkan deret waktu yang telah melalui transformasi dan *differencing* dinyatakan sebagai vektor pengamatan:

$$\mathbf{y} = (y_1, y_2, \dots, y_T)^\top \quad (2.22)$$

dengan  $T$  menyatakan jumlah observasi. Parameter model SARIMA dinyatakan sebagai vektor:

$$\boldsymbol{\psi} = (\psi_1, \psi_2, \dots, \psi_r)^\top \quad (2.23)$$

yang mencakup seluruh parameter *autoregressive* dan *moving average*, baik nonmusiman maupun musiman. Dengan asumsi *innovation*  $a_t \sim N(0, \sigma^2)$ , maka vektor pengamatan memenuhi distribusi normal multivariat  $\mathbf{y} \sim N(\mathbf{0}, \boldsymbol{\Omega}(\boldsymbol{\psi}))$  dengan matriks kovarians  $\boldsymbol{\Omega}(\boldsymbol{\psi})$  yang bergantung pada parameter model (Wei, 2006).

Fungsi *likelihood exact* dapat dituliskan sebagai:

$$L(\boldsymbol{\psi}, \sigma^2 | \mathbf{y}) = (2\pi)^{-\frac{T}{2}} |\boldsymbol{\Omega}(\boldsymbol{\psi})|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2} \mathbf{y}^\top \boldsymbol{\Omega}(\boldsymbol{\psi})^{-1} \mathbf{y}\right\} \quad (2.24)$$

sehingga fungsi *log-likelihood* diperoleh dalam bentuk:

$$\ell(\boldsymbol{\psi}, \sigma^2) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Omega}(\boldsymbol{\psi})| - \frac{1}{2} \mathbf{y}^\top \boldsymbol{\Omega}(\boldsymbol{\psi})^{-1} \mathbf{y} \quad (2.25)$$

representasi *likelihood* tersebut ekuivalen dengan bentuk *innovation* yang dihitung melalui algoritma *Kalman filter*, yang secara komputasi lebih efisien (Shumway & Stoffer, 2017). Misalkan *innovation* pada waktu ke- $t$  dinyatakan sebagai:

$$a_t(\boldsymbol{\psi}) \quad (2.26)$$

yang diperoleh sebagai selisih antara nilai observasi dan prediksi satu langkah ke depan berdasarkan informasi hingga waktu sebelumnya. Dengan asumsi  $a_t(\boldsymbol{\psi}) \sim N(0, \sigma^2)$ , fungsi *likelihood* dapat dinyatakan sebagai (Shumway & Stoffer, 2017):

$$L(\boldsymbol{\psi}, \sigma^2) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{a_t(\boldsymbol{\psi})^2}{2\sigma^2}\right) \quad (2.27)$$

sehingga *log-likelihood* menjadi:

$$\ell(\boldsymbol{\psi}, \sigma^2) = -\frac{T}{2} \log(2\pi) - \frac{T}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=1}^T a_t(\boldsymbol{\psi})^2 \quad (2.28)$$

turunan pertama terhadap  $\sigma^2$  menghasilkan:

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{t=1}^T a_t(\boldsymbol{\psi})^2 \quad (2.29)$$

dengan menyamakan turunan tersebut dengan nol diperoleh estimator ragam galat:

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^T a_t(\boldsymbol{\psi})^2 \quad (2.30)$$

substitusi kembali menunjukkan bahwa proses maksimum *likelihood* terhadap  $\boldsymbol{\psi}$  ekuivalen dengan meminimumkan jumlah kuadrat *innovation*:

$$S(\boldsymbol{\psi}) = \sum_{t=1}^T a_t(\boldsymbol{\psi})^2 \quad (2.31)$$

Karena *innovation* bergantung secara rekursif pada parameter melalui struktur *autoregressive* dan *moving average*, fungsi  $S(\boldsymbol{\psi})$  bersifat nonlinier terhadap parameter sehingga estimasi dilakukan melalui algoritma optimisasi numerik. Estimator yang diperoleh melalui *exact maximum likelihood* bersifat konsisten dan efisien secara asimtotik, serta memiliki distribusi mendekati normal pada ukuran sampel besar (Shumway & Stoffer, 2017). Nilai *log-likelihood* maksimum yang dihasilkan kemudian digunakan dalam pembentukan kriteria informasi seperti *Akaike Information Criterion* (Akaike, 1974), sehingga pemilihan model dilakukan berdasarkan pendekatan *likelihood* yang konsisten.

## 2.9 Pengujian Diagnostik Residual

Dalam pendekatan *Box-Jenkins*, model yang baik ditandai oleh residual yang bersifat *white noise*, yaitu memiliki nilai harapan nol, ragam konstan, serta tidak menunjukkan autokorelasi pada berbagai *lag*. Kondisi ini menunjukkan bahwa struktur ketergantungan dalam data telah dijelaskan oleh model, sehingga tidak terdapat pola sistematis yang tersisa pada residual (Box *et al.*, 2015).

Pada penelitian ini, pengujian diagnostik residual dilakukan terhadap model SARIMA sebagai bagian dari prosedur evaluasi model. Pemeriksaan difokuskan pada dua aspek utama, yaitu keacakan residual dan kesesuaiannya terhadap asumsi distribusi normal. Uji keacakan dilakukan untuk memastikan tidak adanya autokorelasi yang tersisa, sedangkan pengujian distribusi dilakukan untuk menilai apakah residual mendekati distribusi normal. Pemeriksaan ini diperlukan agar model yang dipilih tidak hanya memiliki kecocokan yang baik terhadap data pelatihan, tetapi juga memenuhi asumsi dasar dalam pemodelan deret waktu (Box *et al.*, 2015).

### 2.9.1 Residual Bersifat Acak

Keacakan residual perlu diperiksa untuk memastikan bahwa model tidak meninggalkan pola ketergantungan antarwaktu. Apabila residual masih menunjukkan autokorelasi, maka struktur data belum sepenuhnya dijelaskan oleh model, yang bisa dinilai menggunakan uji *Ljung-Box*, yaitu uji yang memeriksa secara simultan apakah autokorelasi residual hingga sejumlah *lag* tertentu dapat dianggap sama dengan nol (Ljung & Box, 1978).

Statistik uji *Ljung-Box* dinyatakan sebagai:

$$Q = T(T + 2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{T - k} \quad (2.32)$$

dengan:

- $Q$  = statistik uji *Ljung-Box*,
- $T$  = jumlah observasi pada deret waktu,
- $h$  = banyaknya *lag* yang diperiksa,
- $k$  = *lag*,
- $\hat{\rho}_k$  = autokorelasi residual pada *lag* ke- $k$ .

Hipotesis yang diuji adalah:

$$H_0: \rho_1 = \rho_2 = \dots = \rho_h = 0$$

yang menyatakan bahwa tidak terdapat autokorelasi residual hingga *lag*  $h$ . Secara asimtotik, statistik  $Q$  mengikuti distribusi  $\chi_{h-r}^2$ , dengan  $r$  menyatakan jumlah parameter AR dan MA yang diestimasi dalam model, baik nonmusiman maupun musiman.

Keputusan pengujian dilakukan dengan membandingkan nilai  $Q$  dengan nilai kritis distribusi *chi-square* pada taraf nyata  $\alpha$ . Jika  $Q$  lebih besar dari nilai kritis (atau *p-value*  $< \alpha$ ), maka hipotesis nol ditolak dan residual dianggap masih mengandung autokorelasi, sebaliknya, jika tidak ditolak, residual dinilai cukup acak pada *lag* yang diuji.

### 2.9.2 Residual Bersifat Normal

Kenormalan residual diperiksa untuk menilai apakah sebarannya mendekati distribusi normal, karena asumsi ini sering digunakan dalam evaluasi model deret waktu. Salah satu metode yang dapat digunakan adalah uji *Kolmogorov–Smirnov* (KS). Uji ini membandingkan fungsi distribusi kumulatif empiris residual dengan fungsi distribusi kumulatif teoritis dari distribusi normal (Kolmogorov, 1933).

Statistik uji *Kolmogorov–Smirnov* didefinisikan sebagai

$$D_T = \sup_x | F_0(x) - S_T(x) | \quad (2.33)$$

dengan:

- $D_T$  = statistik uji *Kolmogorov–Smirnov*,
- $F_0(x)$  = fungsi distribusi kumulatif teoritis (normal),
- $S_T(x)$  = fungsi distribusi kumulatif empiris residual,
- $T$  = jumlah observasi pada deret waktu.

Hipotesis yang diuji adalah:

$$H_0: F(x) = F_0(x)$$

yang menyatakan bahwa residual berasal dari distribusi normal. Setelah taraf nyata  $\alpha$  ditentukan, nilai  $D_T$  dibandingkan dengan nilai kritis yang sesuai atau dinilai melalui *p-value*. Jika  $D_T$  melebihi nilai kritis (atau *p-value*  $< \alpha$ ), maka hipotesis nol ditolak dan residual dianggap tidak mengikuti distribusi normal, sebaliknya, jika tidak ditolak, residual dinilai masih konsisten dengan asumsi normalitas pada taraf nyata yang dipilih.

### 2.10 Maximum a Posteriori (MAP)

Pendekatan *Maximum a Posteriori* (MAP) merupakan metode estimasi dalam kerangka *Bayesian* yang menentukan parameter sebagai nilai yang memaksimalkan distribusi *posterior* (Bernardo & Smith, 1994). Dalam pendekatan Bayesian, parameter dipandang sebagai peubah acak yang memiliki distribusi *prior*, dan informasi tersebut diperbarui melalui data pengamatan menggunakan Teorema Bayes (Gelman *et al.*, 2014). Dalam proses estimasi, galat didefinisikan sebagai selisih antara nilai observasi dan hasil prediksi model yang mencerminkan ketidaksesuaian terhadap data. Sementara itu, penalti merupakan kontribusi dari distribusi *prior* yang berfungsi membatasi kompleksitas parameter agar tidak terlalu menyesuaikan data. Oleh karena itu, estimasi MAP dilakukan dengan mempertimbangkan keseimbangan antara galat yang berasal dari data dan penalti yang berasal dari *prior* melalui maksimisasi distribusi *posterior*.

Misalkan vektor pengamatan dinyatakan sebagai:

$$\mathbf{y} = (y_1, y_2, \dots, y_T)^\top \quad (2.34)$$

dan parameter model dinyatakan sebagai:

$$\boldsymbol{\vartheta} = (\vartheta_1, \vartheta_2, \dots, \vartheta_p)^\top \quad (2.35)$$

hubungan antara data dan parameter dapat dituliskan dalam bentuk regresi linier:

$$\mathbf{y} = \mathbf{Z}\boldsymbol{\vartheta} + \boldsymbol{\varepsilon} \quad (2.36)$$

dengan:

$$\begin{aligned} \mathbf{Z} &= \text{matriks desain berukuran } T \times p, \\ \boldsymbol{\varepsilon} &\sim N(\mathbf{0}, \sigma^2 \mathbf{I}). \end{aligned}$$

Dengan asumsi kenormalan galat, fungsi *likelihood* diberikan oleh (Gelman *et al.*, 2014):

$$p(\mathbf{y} | \boldsymbol{\vartheta}) = (2\pi\sigma^2)^{-\frac{T}{2}} \exp\left(-\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{Z}\boldsymbol{\vartheta})^\top (\mathbf{y} - \mathbf{Z}\boldsymbol{\vartheta})\right) \quad (2.37)$$

Misalkan distribusi *prior* terhadap parameter berbentuk normal dengan rata-rata nol dan ragam  $\tau^2$ ,  $\boldsymbol{\vartheta} \sim N(\mathbf{0}, \tau^2 \mathbf{I})$  maka fungsi *prior* dinyatakan sebagai (Bernardo & Smith, 1994):

$$p(\boldsymbol{\vartheta}) = (2\pi\tau^2)^{-\frac{p}{2}} \exp\left(-\frac{1}{2\tau^2} \boldsymbol{\vartheta}^\top \boldsymbol{\vartheta}\right) \quad (2.38)$$

Distribusi *posterior* diperoleh melalui Teorema *Bayes*:

$$p(\boldsymbol{\vartheta} | \mathbf{y}) = \frac{p(\mathbf{y} | \boldsymbol{\vartheta})p(\boldsymbol{\vartheta})}{p(\mathbf{y})} \quad (2.39)$$

dengan  $p(\mathbf{y})$  sebagai konstanta normalisasi (Gelman *et al.*, 2014).

Dengan mengabaikan konstanta yang tidak bergantung pada parameter, *log-posterior* dapat dituliskan sebagai:

$$\log p(\boldsymbol{\vartheta} | \mathbf{y}) = -\frac{1}{2\sigma^2} (\mathbf{y} - \mathbf{Z}\boldsymbol{\vartheta})^\top (\mathbf{y} - \mathbf{Z}\boldsymbol{\vartheta}) - \frac{1}{2\tau^2} \boldsymbol{\vartheta}^\top \boldsymbol{\vartheta} + C \quad (2.40)$$

Estimator MAP didefinisikan sebagai:

$$\hat{\boldsymbol{\vartheta}}_{MAP} = \arg \max_{\boldsymbol{\vartheta}} \log p(\boldsymbol{\vartheta} | \mathbf{y}) \quad (2.41)$$

Memaksimalkan *log-posterior* ekuivalen dengan meminimumkan fungsi:

$$Q(\boldsymbol{\vartheta}) = \frac{1}{\sigma^2} (\mathbf{y} - \mathbf{Z}\boldsymbol{\vartheta})^\top (\mathbf{y} - \mathbf{Z}\boldsymbol{\vartheta}) + \frac{1}{\tau^2} \boldsymbol{\vartheta}^\top \boldsymbol{\vartheta} \quad (2.42)$$

definisikan  $\lambda = \frac{\sigma^2}{\tau^2}$

maka fungsi yang diminimumkan menjadi:

$$Q(\boldsymbol{\vartheta}) = (\mathbf{y} - \mathbf{Z}\boldsymbol{\vartheta})^\top (\mathbf{y} - \mathbf{Z}\boldsymbol{\vartheta}) + \lambda \boldsymbol{\vartheta}^\top \boldsymbol{\vartheta} \quad (2.43)$$

turunan pertama terhadap  $\boldsymbol{\vartheta}$  adalah:

$$\frac{\partial Q}{\partial \boldsymbol{\vartheta}} = -2\mathbf{Z}^\top (\mathbf{y} - \mathbf{Z}\boldsymbol{\vartheta}) + 2\lambda \boldsymbol{\vartheta} \quad (2.44)$$

dengan menyamakan turunan tersebut dengan nol diperoleh:

$$\mathbf{Z}^\top \mathbf{Z}\boldsymbol{\vartheta} + \lambda \boldsymbol{\vartheta} = \mathbf{Z}^\top \mathbf{y} \quad (2.45)$$

atau

$$(\mathbf{Z}^\top \mathbf{Z} + \lambda \mathbf{I})\boldsymbol{\vartheta} = \mathbf{Z}^\top \mathbf{y} \quad (2.46)$$

Sehingga estimator MAP diberikan oleh:

$$\hat{\boldsymbol{\vartheta}}_{MAP} = (\mathbf{Z}^\top \mathbf{Z} + \lambda \mathbf{I})^{-1} \mathbf{Z}^\top \mathbf{y} \quad (2.47)$$

hasil tersebut menunjukkan bahwa estimasi MAP dengan *prior* normal ekuivalen dengan regresi linier terpenalti (*ridge regression*), dengan parameter  $\lambda$  mengontrol tingkat regularisasi (Bernardo & Smith, 1994). Parameter *prior* dengan ragam kecil menghasilkan penalti lebih besar terhadap besarnya parameter, sehingga kompleksitas model dapat dikendalikan.

Karena *likelihood* dan *prior* sama-sama berbentuk normal, distribusi *posterior* juga berdistribusi normal dengan parameter (Gelman *et al.*, 2014):

$$\boldsymbol{\vartheta} \mid \mathbf{y} \sim N((\mathbf{Z}^\top \mathbf{Z} + \lambda \mathbf{I})^{-1} \mathbf{Z}^\top \mathbf{y}, \quad \sigma^2 (\mathbf{Z}^\top \mathbf{Z} + \lambda \mathbf{I})^{-1}) \quad (2.48)$$

Dalam implementasi komputasi, estimasi parameter dilakukan dengan mengoptimalkan *log-posterior* secara numerik. Pendekatan ini menghasilkan estimasi deterministik yang menyeimbangkan kecocokan terhadap data dan pembatasan kompleksitas melalui penalti kuadrat pada parameter (Taylor & Letham, 2018).

## 2.11 Kriteria Kebaikan Model

Kinerja model peramalan dinilai dengan mengukur seberapa besar perbedaan antara nilai ramalan dan nilai aktual. Pengukuran ini bertujuan untuk menilai kemampuan model dalam merepresentasikan pola data dan menghasilkan ramalan yang akurat pada periode evaluasi. Dalam penelitian ini, digunakan tiga ukuran kesalahan yang umum dalam analisis deret waktu, yaitu *mean absolute error* (MAE), *root mean squared error* (RMSE), dan *mean absolute percentage error* (MAPE) (Hyndman & Athanasopoulos, 2021).

Ketiga ukuran tersebut menghitung selisih antara nilai aktual  $y_t$  dan nilai ramalan  $\hat{y}_t$  dengan pendekatan yang berbeda. MAE mengukur rata-rata besarnya kesalahan absolut, RMSE memberikan penalti lebih besar terhadap kesalahan yang berukuran besar melalui kuadrat, sedangkan MAPE menyajikan kesalahan dalam bentuk persentase relatif terhadap nilai aktual. Perbedaan karakteristik ini memberikan sudut pandang yang lebih lengkap dalam mengevaluasi dan membandingkan model peramalan (Hyndman & Koehler, 2006).

### 2.11.1 Mean Absolute Error (MAE)

*Mean absolute error* (MAE) merupakan ukuran yang digunakan untuk menghitung rata-rata besarnya selisih antara nilai aktual dan nilai ramalan. Karena kesalahan dinyatakan dalam nilai absolut, ukuran ini tidak membedakan apakah ramalan berada di atas atau di bawah nilai sebenarnya. MAE memiliki satuan yang sama dengan satuan data sehingga hasilnya mudah diinterpretasikan dalam konteks praktik. Secara matematis, MAE dinyatakan sebagai (Hyndman & Athanasopoulos, 2021):

$$e_t = y_t - \hat{y}_t, \quad MAE = \frac{1}{H} \sum_{t=1}^H |e_t| \quad (2.49)$$

dengan:

- $y_t$  = nilai aktual pada waktu ke- $t$ ,
- $\hat{y}_t$  = nilai ramalan pada waktu ke- $t$ ,
- $e_t$  = galat peramalan,
- $H$  = jumlah observasi pada periode evaluasi.

Nilai MAE yang lebih kecil menunjukkan kinerja peramalan yang lebih baik.

### 2.11.2 Root Mean Square Error (RMSE)

RMSE digunakan untuk mengukur rata-rata besarnya galat dengan memberi bobot lebih pada galat besar karena dikuadratkan terlebih dahulu. Ukuran ini umum dipakai ketika kesalahan besar dianggap lebih penting dalam evaluasi model. RMSE dapat dinyatakan dengan persamaan sebagai berikut (Hyndman & Athanasopoulos, 2021):

$$e_t = y_t - \hat{y}_t, \quad RMSE = \sqrt{\frac{1}{H} \sum_{t=1}^H e_t^2} \quad (2.50)$$

dengan:

- $y_t$  = nilai aktual pada waktu ke- $t$ ,
- $\hat{y}_t$  = nilai ramalan pada waktu ke- $t$ ,
- $e_t$  = galat peramalan,
- $H$  = jumlah observasi pada periode evaluasi.

Nilai RMSE yang lebih kecil menunjukkan kinerja peramalan yang lebih baik. Karena melibatkan kuadrat galat, RMSE lebih sensitif terhadap kesalahan yang berukuran besar dibandingkan MAE.

### 2.11.3 Mean Absolute Percentage Error (MAPE)

MAPE menyatakan besar kesalahan dalam bentuk persentase sehingga hasilnya mudah dibaca dan dapat digunakan untuk membandingkan akurasi pada data dengan skala berbeda. Namun, ukuran ini tidak terdefinisi ketika  $y_t = 0$  dan dapat menghasilkan nilai sangat besar apabila  $y_t$  berada sangat dekat dengan nol. MAPE dapat dinyatakan dalam persamaan berikut (Hyndman & Koehler, 2006):

$$e_t = y_t - \hat{y}_t, \quad MAPE = \frac{100}{H} \sum_{t=1}^H \left| \frac{e_t}{y_t} \right| \quad (2.51)$$

dengan:

- $y_t$  = nilai aktual pada waktu ke- $t$ ,
- $\hat{y}_t$  = nilai ramalan pada waktu ke- $t$ ,
- $e_t$  = galat peramalan,
- $H$  = jumlah observasi pada periode evaluasi.

Interpretasi nilai *mean absolute percentage error* (MAPE) dapat digunakan sebagai pedoman untuk menilai tingkat akurasi hasil peramalan. Secara umum, semakin kecil nilai MAPE menunjukkan hasil ramalan yang semakin baik. Klasifikasi tingkat akurasi berdasarkan nilai MAPE dapat disajikan pada tabel 1 sebagai berikut (Lewis, 1982).

Tabel 1. Kriteria Interpretasi Nilai MAPE

| Nilai MAPE              | Interpretasi |
|-------------------------|--------------|
| $MAPE \leq 10\%$        | Sangat baik  |
| $10\% < MAPE \leq 20\%$ | Baik         |
| $20\% < MAPE \leq 50\%$ | Cukup        |
| $MAPE > 50\%$           | Kurang baik  |

## **2.12 Perjalanan Wisatawan Nusantara Menurut Provinsi Tujuan (Perjalanan)**

Menurut *International Recommendations for Tourism Statistics (IRTS 2008)*, perjalanan wisata adalah perpindahan seseorang keluar dari lingkungan kebiasaannya di dalam negeri dengan durasi kurang dari satu tahun, dan tujuan utamanya bukan untuk bekerja di tempat yang dikunjungi (United Nations, 2010). Dalam konteks ini, yang dihitung bukan jumlah orangnya, melainkan jumlah perjalanan yang dilakukan. Satu orang dapat melakukan lebih dari satu perjalanan dalam satu periode, sehingga angka yang tercatat merepresentasikan frekuensi perjalanan.

Badan Pusat Statistik (BPS) secara rutin menerbitkan data bulanan jumlah perjalanan wisatawan nusantara menurut provinsi tujuan. Data tersebut kemudian dapat dijumlahkan untuk memperoleh total perjalanan tingkat nasional. Karena setiap perjalanan berkaitan dengan kebutuhan transportasi, akomodasi, serta layanan destinasi lainnya, indikator jumlah perjalanan sering digunakan sebagai gambaran permintaan pariwisata pada skala bulanan (Badan Pusat Statistik, 2025). Penyajian dalam bentuk deret waktu bulanan memungkinkan pola tren dan musiman dianalisis lebih lanjut dalam penelitian ini.

### III. METODOLOGI PENELITIAN

#### 3.1 Waktu dan Tempat Penelitian

Penelitian ini dilaksanakan pada semester ganjil tahun ajaran 2025/2026 di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

#### 3.2 Data Penelitian

Data penelitian yang digunakan adalah deret waktu bulanan “Jumlah Perjalanan Wisatawan Nusantara Menurut Provinsi Tujuan (Perjalanan)” yang dipublikasikan oleh Badan Pusat Statistik (BPS), tertera pada *link* berikut:

<https://www.bps.go.id/id/statistics-table/2/MjIwMSMy/jumlah-perjalanan-wisatawan-nusantara-menurut-provinsi-tujuan--perjalanan-.html>.

Data diunduh dari situs resmi BPS, disatukan dalam *file* Excel (.xlsx), kemudian nilai perjalanan dari seluruh provinsi dijumlahkan untuk memperoleh total nasional per bulan. Deret ini menjadi dasar pemodelan SARIMA dan *prophet*.

Rentang data yang digunakan adalah Januari 2019 sampai Desember 2025. Periode ini dipilih karena seri bulanan dari BPS tersedia secara berurutan dan dapat digunakan untuk melihat perubahan dari waktu ke waktu (Badan Pusat Statistik, 2025). Rentang tersebut juga mencakup masa sebelum pandemi, masa pembatasan kegiatan, dan masa pemulihan, sehingga perubahan pola perjalanan selama periode

tersebut ikut terlihat (Kementerian Pariwisata dan Ekonomi Kreatif, 2021), selain itu, panjang deret telah mencakup beberapa tahun, sehingga pola musiman bulanan dalam data lebih mudah dikenali untuk pemilihan model (Hyndman & Athanasopoulos, 2021).

### 3.3 Metode Penelitian

Data penelitian dibagi menjadi dua bagian, yaitu *in-sample* untuk proses pembangunan model dan *out-of-sample* untuk mengevaluasi akurasi peramalan. Pembagian ini mengikuti praktik umum dalam analisis deret waktu agar kemampuan peramalan dapat diuji pada data yang tidak digunakan selama proses estimasi, sehingga hasilnya lebih objektif dan tidak terpengaruh *overfitting* (Hyndman & Athanasopoulos, 2021). Periode Januari 2019 sampai Desember 2024 digunakan sebagai *in-sample* karena panjang data mencukupi untuk mengenali pola *trend* dan musiman tahunan sesuai kebutuhan pendekatan *Box–Jenkins* (Box *et al.*, 2015). Periode Januari 2025 sampai Desember 2025 digunakan sebagai *out-of-sample* untuk menilai kinerja model pada konteks peramalan aktual (Chatfield, 2003). Seluruh pengolahan data dilakukan menggunakan *RStudio*.

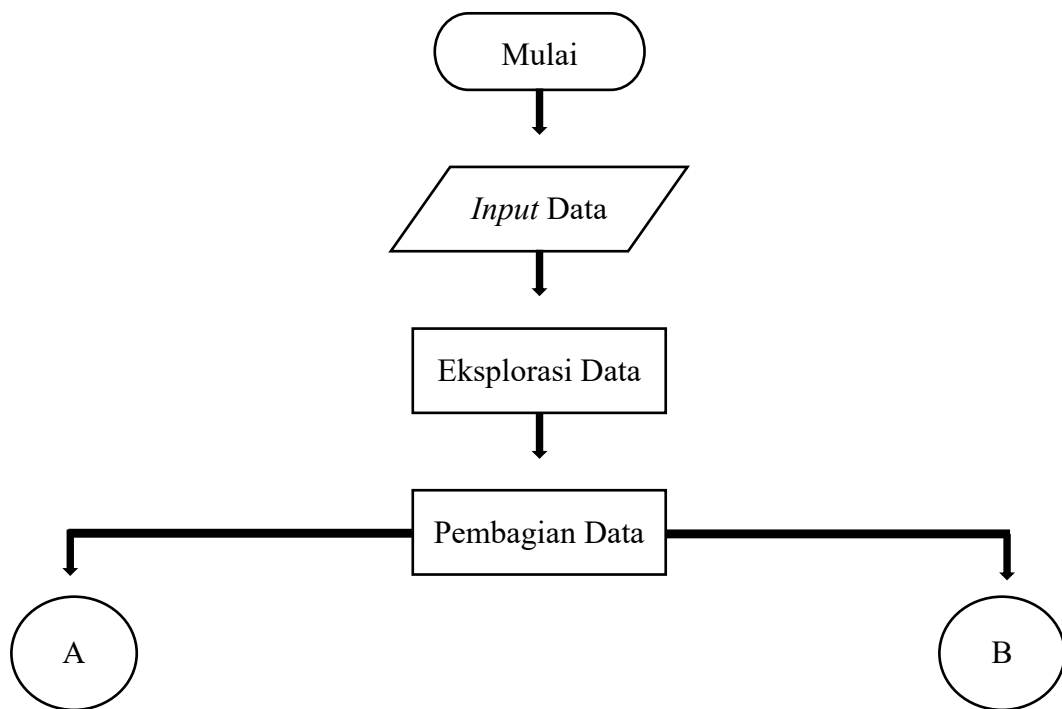
Tahapan analisis:

1. Eksplorasi data melalui *plot* dan statistik deskriptif untuk melihat pola tren, musiman, serta fluktuasi yang muncul pada data.
2. Pemeriksaan kestasioneran dilakukan melalui pengamatan visual (*plot*) serta pengujian *Augmented Dickey–Fuller* (ADF). Apabila deret menunjukkan ketidakstabilan dalam *mean* atau variansi, diterapkan transformasi *Box–Cox* dan *differencing* hingga deret memenuhi asumsi kestasioneran yang diperlukan dalam pemodelan SARIMA.
3. Identifikasi struktur deret waktu dilakukan dengan menganalisis pola *autocorrelation function* (ACF) dan *partial autocorrelation function* (PACF). Pola yang muncul digunakan untuk menentukan dugaan awal komponen

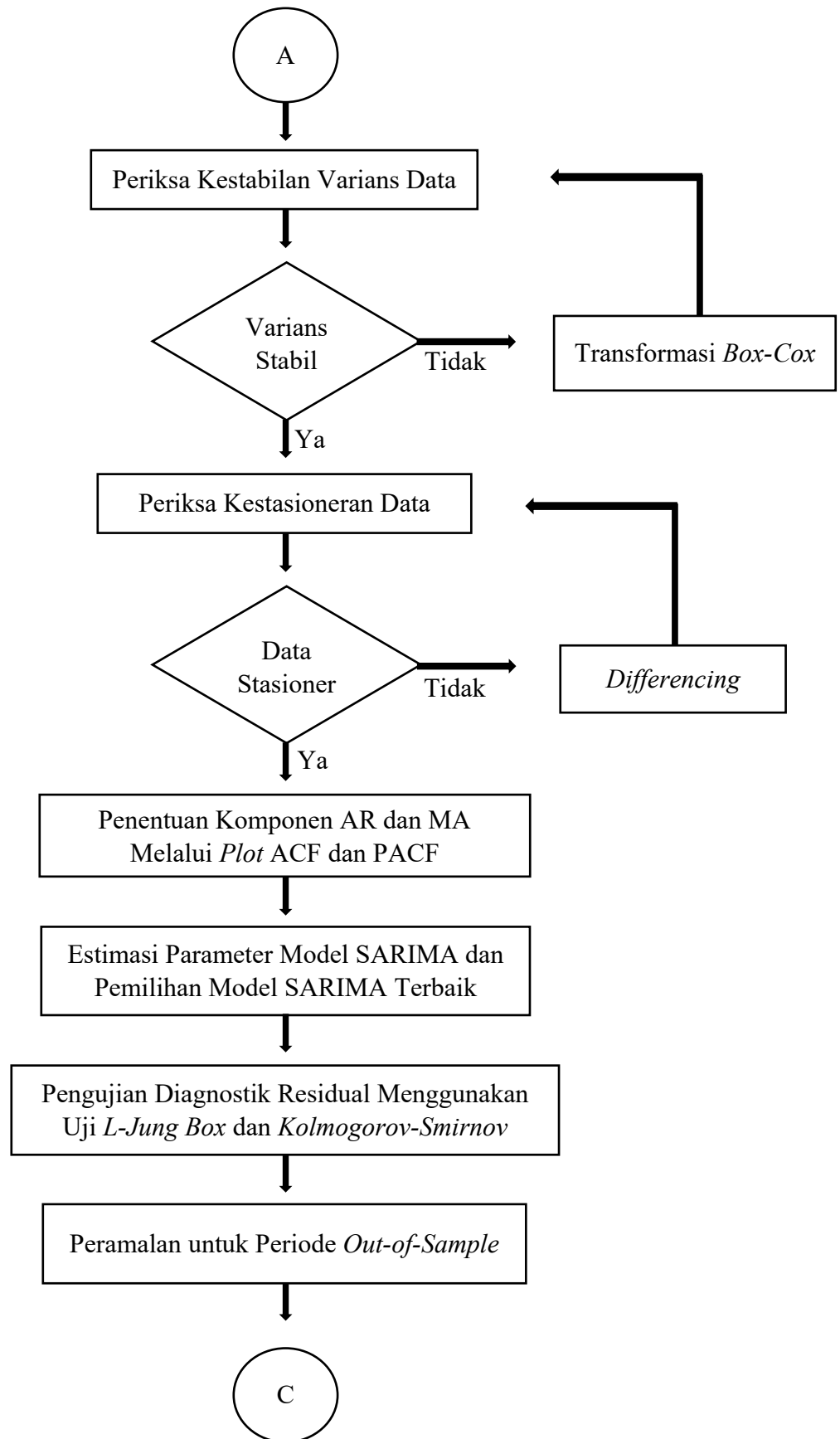
*autoregressive* (AR), *moving average* (MA), serta komponen musiman yang berpotensi membentuk struktur model SARIMA.

4. Pembentukan model SARIMA dilakukan dengan menentukan orde  $(p, d, q)(P, D, Q)_s$  berdasarkan hasil *differencing* dan pola ACF–PACF. Parameter model diestimasi menggunakan metode *exact maximum likelihood*. Pemilihan kandidat model dilakukan dengan nilai AIC terendah. Model diestimasi menggunakan data *in-sample*.
5. Pengujian diagnostik model dilakukan melalui analisis residual. Uji *Ljung–Box* digunakan untuk memastikan residual bersifat *white noise*, sedangkan uji *Kolmogorov–Smirnov* (KS) digunakan untuk menilai apakah residual mendekati distribusi normal. Pemeriksaan diagnostik residual difokuskan pada model SARIMA karena model ini mengikuti kerangka *Box–Jenkins*.
6. Pembentukan model *prophet* dilakukan menggunakan *piecewise linear trend* dengan titik perubahan (*change points*) yang ditentukan secara manual, serta komponen musiman tahunan berbasis *Fourier series*. Model diestimasi menggunakan data *in-sample*.
7. Peramalan dan evaluasi akurasi dilakukan dengan menggunakan kedua model untuk memprediksi periode *out-of-sample*, yaitu periode Januari 2025 sampai Desember 2025. Akurasi peramalan dibandingkan menggunakan ukuran kesalahan *mean absolute error* (MAE), *root mean square error* (RMSE), dan *mean absolute percentage error* (MAPE). Model dengan nilai kesalahan yang lebih kecil dinilai memiliki kinerja peramalan yang lebih baik.
8. Pemilihan model terbaik ditentukan berdasarkan kinerja akurasi pada data *out-of-sample*. Model terpilih selanjutnya digunakan untuk menghasilkan peramalan 12 bulan ke depan dari periode data terakhir.
9. Penarikan kesimpulan disusun berdasarkan perbandingan kinerja model SARIMA dan *prophet*, hasil evaluasi akurasi, serta hasil peramalan lanjutan. Temuan penelitian digunakan untuk memberikan gambaran mengenai kemampuan model dalam memprediksi jumlah perjalanan wisatawan nusantara sebagai dasar pendukung perencanaan dan pengambilan keputusan.

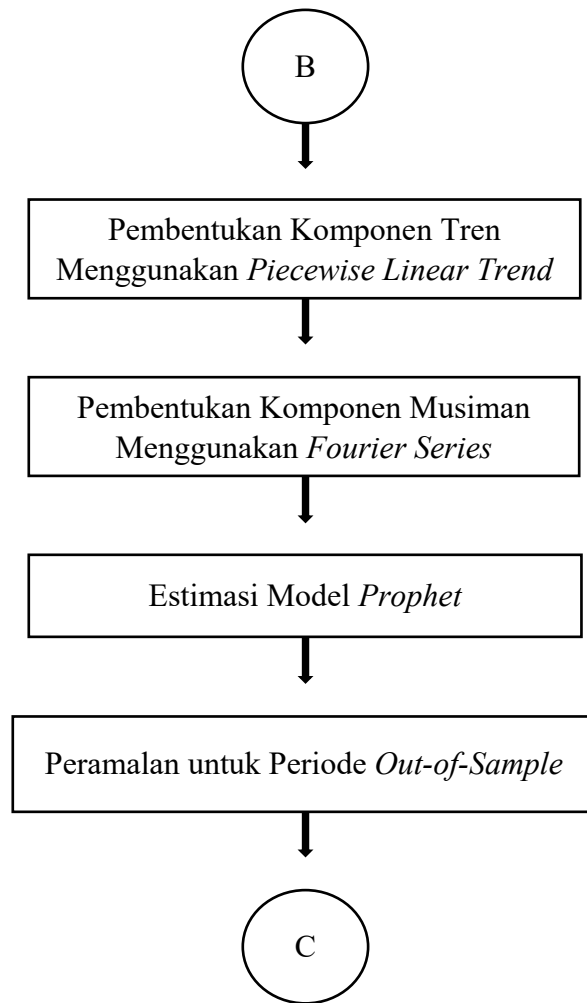
Berikut diberikan *flowchart* dari penelitian ini:



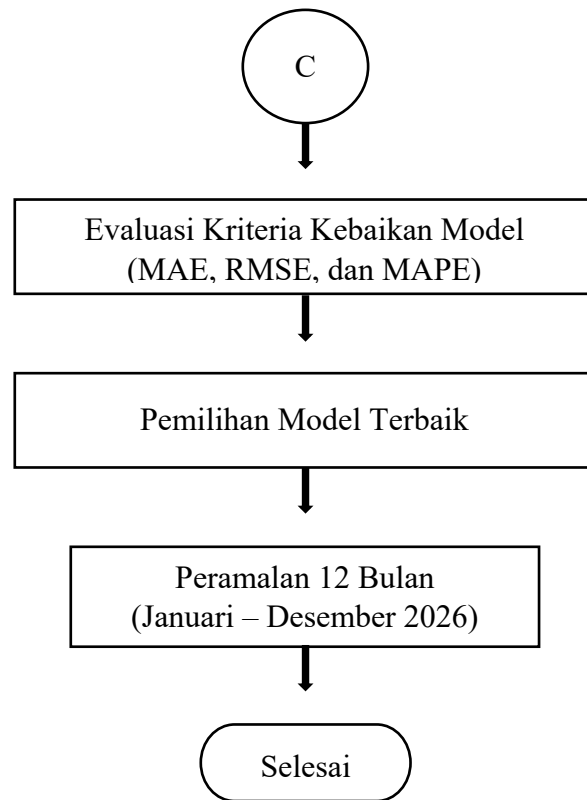
Gambar 1. *Flowchart* Analisis Penelitian



Gambar 2. Flowchart Analisis Penelitian (Cabang SARIMA)



Gambar 3. *Flowchart* Analisis Penelitian (Cabang *Prophet*)



Gambar 4. *Flowchart* Analisis Penelitian (Lanjutan)

## V. KESIMPULAN

Pemodelan deret waktu jumlah perjalanan wisatawan nusantara periode Januari 2019 hingga Desember 2025 menunjukkan bahwa baik SARIMA maupun *prophet* mampu merepresentasikan struktur tren dan musiman tahunan yang terdapat dalam data. Setelah dilakukan transformasi *Box-Cox* serta *differencing* nonmusiman satu kali dan *differencing* musiman satu kali dengan periode 12, model terbaik yang diperoleh berdasarkan kriteria AIC adalah model SARIMA(0,1,1)(2,1,1)<sub>12</sub>. Model tersebut menghasilkan residual yang memenuhi asumsi *white noise* dan asumsi residual berdistribusi normal, sehingga tidak terdapat autokorelasi yang tersisa dan distribusi residual tidak berbeda secara signifikan dari normal.

Pada periode evaluasi *out-of-sample* Januari 2025 hingga Desember 2025, model SARIMA menghasilkan nilai MAE sebesar 7.177.901, RMSE sebesar 10.112.054, dan MAPE sebesar 6,71%, sementara itu, model *prophet* menghasilkan MAE sebesar 8.905.072, RMSE sebesar 12.718.798, dan MAPE sebesar 8,02%. Kedua model memiliki tingkat kesalahan relatif di bawah 10 persen, yang menunjukkan tingkat kesalahan relatif yang rendah. Namun secara konsisten, seluruh ukuran kesalahan pada model SARIMA lebih kecil dibandingkan *prophet*.

Hasil tersebut menunjukkan bahwa pada data dengan pola musiman tahunan yang relatif stabil dan ketergantungan antarperiode yang kuat, model berbasis *Box-Jenkins* seperti SARIMA memberikan kinerja yang lebih optimal dibandingkan pendekatan berbasis komponen seperti *prophet*.

## DAFTAR PUSTAKA

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*. **19**(6): 716–723.
- Badan Pusat Statistik. 2025. *Jumlah Perjalanan Wisatawan Nusantara Menurut Provinsi Tujuan (Perjalanan)*. Jakarta: BPS.  
<https://www.bps.go.id/id/statistics-table/2/MjIwMSMy/jumlah-perjalanan-wisatawan-nusantara-menurut-provinsi-tujuan--perjalanan-.html>. Diakses 8 Februari 2026
- Bernardo, J. M., & Smith, A. F. M. 1994. *Bayesian Theory*. Chichester: John Wiley & Sons.
- Bouhaddour, S., Saadi, C., Bouabdallaoui, I., Guerouate, F., dan Sbihi, M. 2023. Tourism in Singapore, prediction model using SARIMA and PROPHET. *AIP Conference Proceedings*. **2508**(1): 020042.
- Box, G. E. P., & Cox, D. R. 1964. An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*. **26**(2): 211–252.
- Box, G. E. P., & Jenkins, G. M. 1970. *Time Series Analysis: Forecasting and Control* (Ed. ke-1). San Francisco: Holden-Day.

- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. 2015. *Time Series Analysis: Forecasting and Control* (Ed. ke-5). Hoboken: John Wiley & Sons.
- Chatfield, C. 2003. *The Analysis of Time Series: An Introduction* (Ed. ke-6). Boca Raton: Chapman & Hall/CRC.
- Feng, T., Zheng, Z., Xu, J., Liu, M., Li, M., Jia, H., dan Yu, X. 2022. The comparative analysis of SARIMA, Facebook Prophet, and LSTM for road traffic injury prediction in Northeast China. *Frontiers in Public Health*. **10**: 946563.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. 2014. *Bayesian Data Analysis* (Ed. ke-3). Boca Raton: CRC Press.
- Hamilton, J. D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.
- Hyndman, R. J., & Athanasopoulos, G. 2021. *Forecasting: Principles and Practice* (Ed. ke-3). Melbourne: OTexts.
- Hyndman, R. J., & Koehler, A. B. 2006. Another look at measures of forecast accuracy. *International Journal of Forecasting*. **22**(4): 679–688.
- Kementerian Pariwisata dan Ekonomi Kreatif. 2021. *Informasi kebijakan pembatasan kegiatan masyarakat pada sektor parekraf*. Jakarta: Kemenparekraf.  
<https://kemenparekraf.go.id/ragam-ekonomi-kreatif/Kebijakan-PPKM-Darurat-Sektor-Parekraf>. Diakses 7 November 2025.
- Kolmogorov, A. N. 1933. Sulla determinazione empirica di una legge di distribuzione. *Giornale dell'Istituto Italiano degli Attuari*. **4**: 83–91.

- Levinson, N. 1947. The Wiener RMS error criterion in filter design and prediction. *Journal of Mathematics and Physics*. **25**: 261–278.
- Lewis, C. D. 1982. *Industrial and Business Forecasting Methods*. London: Butterworths.
- Ljung, G. M., & Box, G. E. P. 1978. On a measure of lack of fit in time series models. *Biometrika*. **65**(2): 297–303.
- MacKinnon, J. G. 1996. Numerical distribution functions for unit root and cointegration tests. *Journal of Applied Econometrics*. **11**(6): 601–618.
- Negre, P., Alonso, R. S., Prieto, J., García, Ó., dan de-la-Fuente-Valentín, L. 2024. Prediction of footwear demand using Prophet and SARIMA. *Expert Systems with Applications*. **255**: 124512.
- Nur Aziza, V., Moh'd, F. H., Maghfiroh, F. A., Notodiputro, K. A., dan Angraini, Y. 2023. Performance comparison of SARIMA intervention and Prophet models for forecasting the number of airline passenger at Soekarno Hatta International Airport. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*. **17**(4): 2107–2120.
- Rizal, A. A. 2025. Comparative evaluation of SARIMA, Prophet, and XGBoost for tourist arrivals in NTB. *COGITO Smart Journal*. **11**(1): 152–166.
- Said, S. E., & Dickey, D. A. 1984. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*. **71**(3): 599–607.
- Sudjana. 1986. *Metoda Statistika*. Bandung: Tarsito.

Shumway, R. H., & Stoffer, D. S. 2017. *Time Series Analysis and Its Applications: With R Examples* (Ed. ke-4). Cham: Springer.

Taylor, S. J., & Letham, B. 2018. Forecasting at scale. *The American Statistician*. 72(1): 37–45.

United Nations. 2010. *International Recommendations for Tourism Statistics 2008 (IRTS 2008)*. New York: United Nations.

Wei, W. W. S. 2006. *Time Series Analysis: Univariate and Multivariate Methods* (Ed. ke-2). Boston: Pearson/Addison-Wesley.