



METODE
PERAMALAN

ANALISIS RUNTUN WAKTU PADA PENUMPANG PESAWAT DI CHANGI AIRPORT SELAMA TAHUN 2009 - 2019

PROJECT UAS METODE PERAMALAN

KIRONO DWI SAPUTRO
LUTHFI ATHALLAH H. W.
ZAVIER RAIYAN DANA

2106656365
2206826980
2206829875

CONTENT

01

LATAR BELAKANG

02

LANDASAN TEORI

03

PRE-PROCESSING

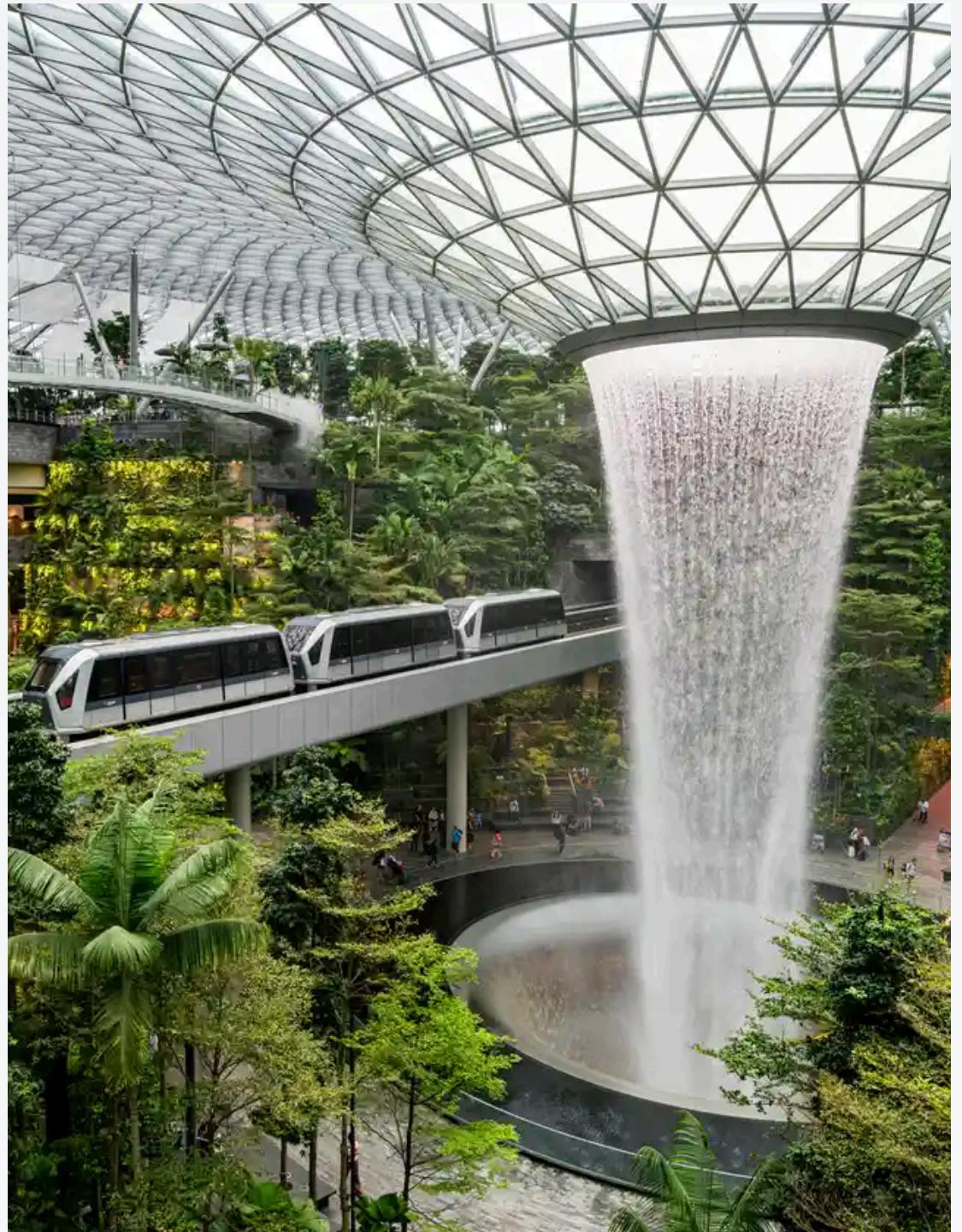
04

HASIL DAN PEMBAHASAN

05

KESIMPULAN

LATAR BELAKANG



Changi Airport di Singapura, sebagai salah satu hub transportasi udara tersibuk di dunia, tidak hanya menghubungkan Asia dengan berbagai belahan dunia tetapi juga berfungsi sebagai pusat transit utama global. Dari tahun 2009 hingga 2019, jumlah penumpang yang melewati bandara ini terus meningkat, mencerminkan peningkatan mobilitas global dan peran strategis Singapura dalam jaringan penerbangan internasional.

Meskipun menghadapi fluktuasi yang dipengaruhi oleh faktor musiman, ekonomi global, dan dinamika industri penerbangan, analisis data time series memainkan peran penting dalam memahami pola historis dan meramalkan tren masa depan jumlah penumpang. Dengan pendekatan pemodelan yang tepat, informasi ini memberikan wawasan berharga untuk pengambilan keputusan strategis dan perencanaan operasional bandara.

RUMUSAN MASALAH



1. Bagaimana pola jumlah penumpang pesawat bulanan di Bandara Changi dari tahun 2009 hingga 2019?
2. Apakah terdapat komponen musiman yang signifikan dalam data jumlah penumpang pesawat bulanan di Bandara Changi?
3. Model SARIMA apa yang paling cocok untuk memodelkan dan memprediksi jumlah penumpang pesawat bulanan di Bandara Changi?
4. Bagaimana akurasi model SARIMA yang telah dipilih dalam memprediksi jumlah penumpang pesawat pada periode tertentu di Bandara Changi?

TUJUAN PENELITIAN



1. Mengidentifikasi dan menganalisis pola jumlah penumpang pesawat bulanan di Bandara Changi dari tahun 2009 hingga 2019.
2. Menemukan komponen musiman yang signifikan dalam data jumlah penumpang pesawat bulanan di Bandara Changi.
3. Memilih dan mengevaluasi model SARIMA yang paling cocok untuk memodelkan data jumlah penumpang pesawat bulanan di Bandara Changi.
4. Mengevaluasi akurasi model SARIMA yang telah dipilih dalam memprediksi jumlah penumpang pesawat pada periode tertentu dan memberikan rekomendasi berdasarkan hasil analisis tersebut.

LANDASAN TEORI

TIME SERIES MODEL



Model Arima

$$\Delta^d Y_t = c + \phi_1 \Delta^d Y_{t-1} + \cdots + \phi_p \Delta^d Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

Model SARIMA

$$\Phi_P(B^s) \Delta_s^D \phi_p(B) \Delta^d Y_t = \theta_q(B) \Theta_Q(B^s) \epsilon_t$$

UJI ADF

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \delta_1 \Delta Y_{t-1} + \cdots + \delta_k \Delta Y_{t-k} + \epsilon_t$$

PRE-PROCESSING

untuk membersihkan, merapihkan, dan mempersiapkan data agar siap untuk proses analisis atau pemodelan lebih lanjut.

PRE-PROCESSING

data yang digunakan adalah [Air Passengers @ SGP Changi Airport for past 10yrs \(kaggle.com\)](#). dengan ukuran data nya adalah terdapat 2 variabel/kolom (Month dan value) dimana dengan 130 obesrvasi/baris

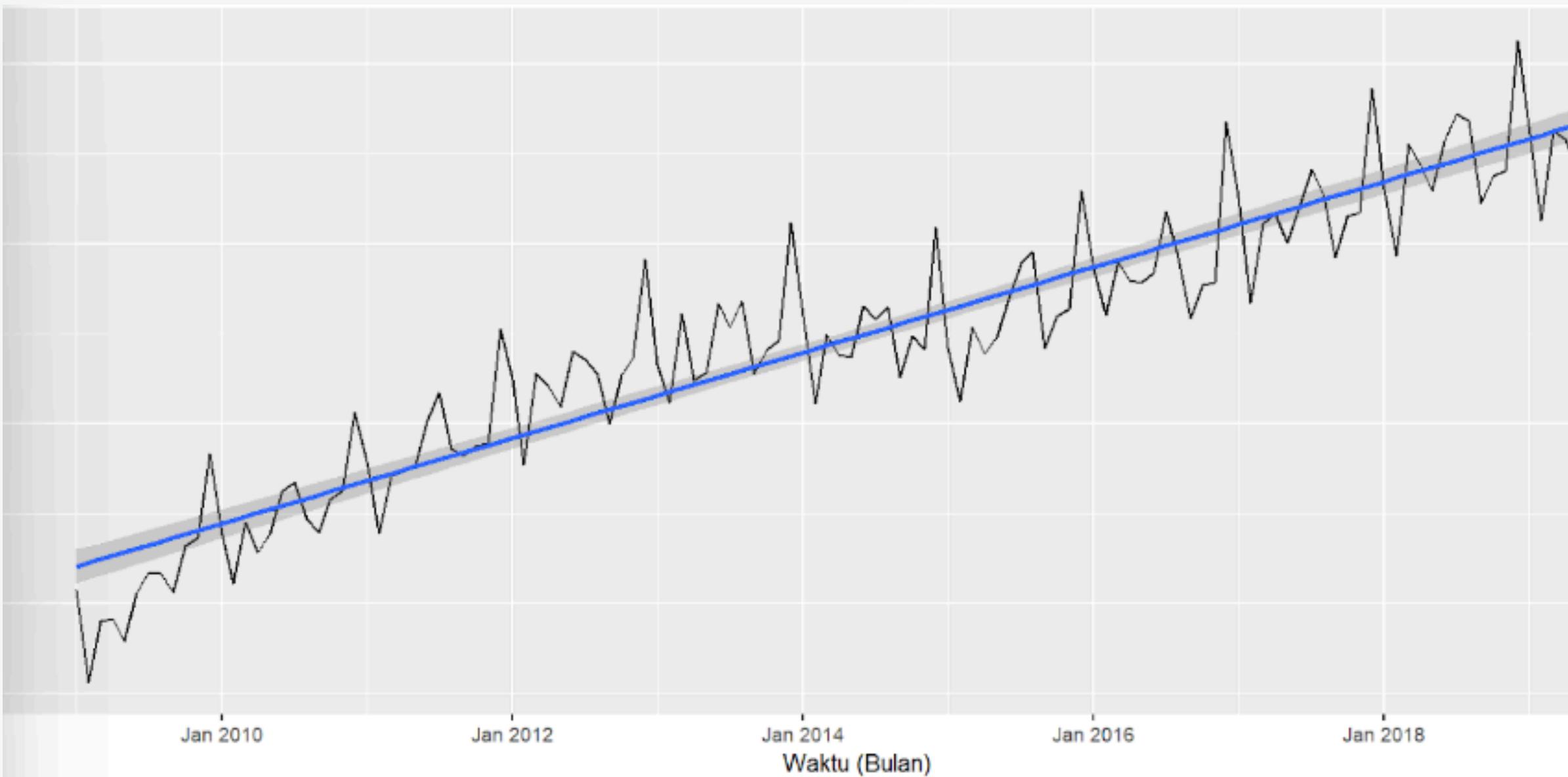
Konversi Format Kolom

```
# Konversi kolom 'month' menjadi objek 'yearmon'  
data <- data %>%  
  mutate(month = as.yearmon(month, "%Y-%m"))  
data
```

Missing value

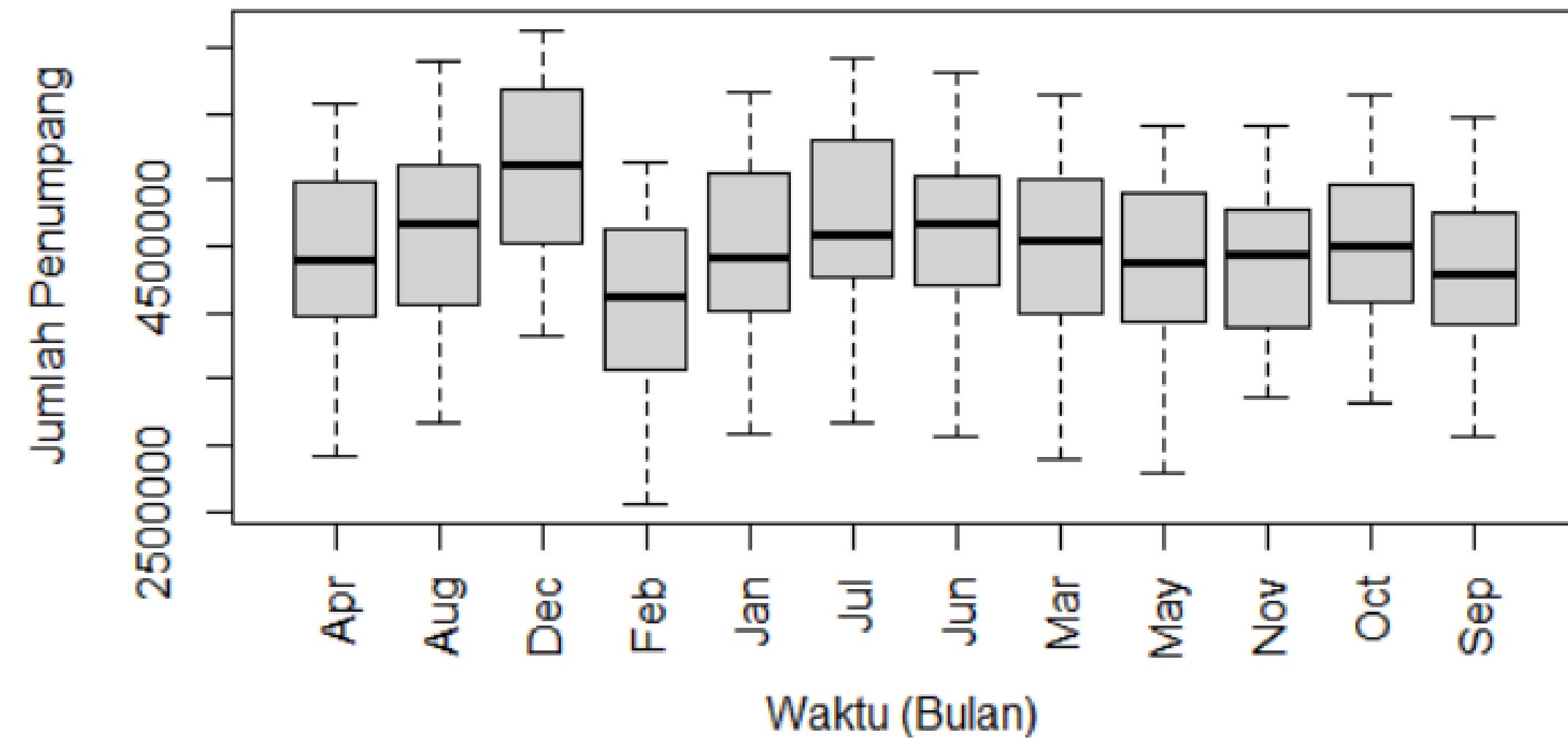
```
> sapply(df, function(x) sum(is.na(x)))  
month passenger  
      0          0
```

PRE-PROCESSING

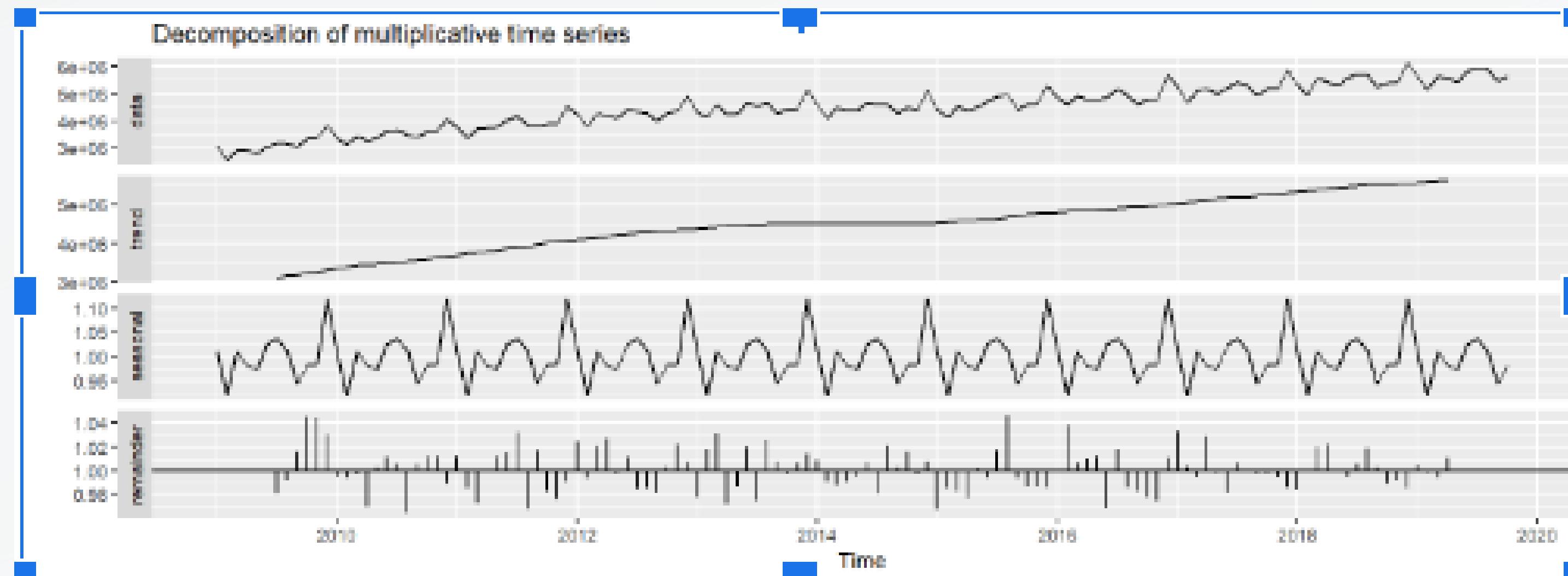


PRE-PROCESSING

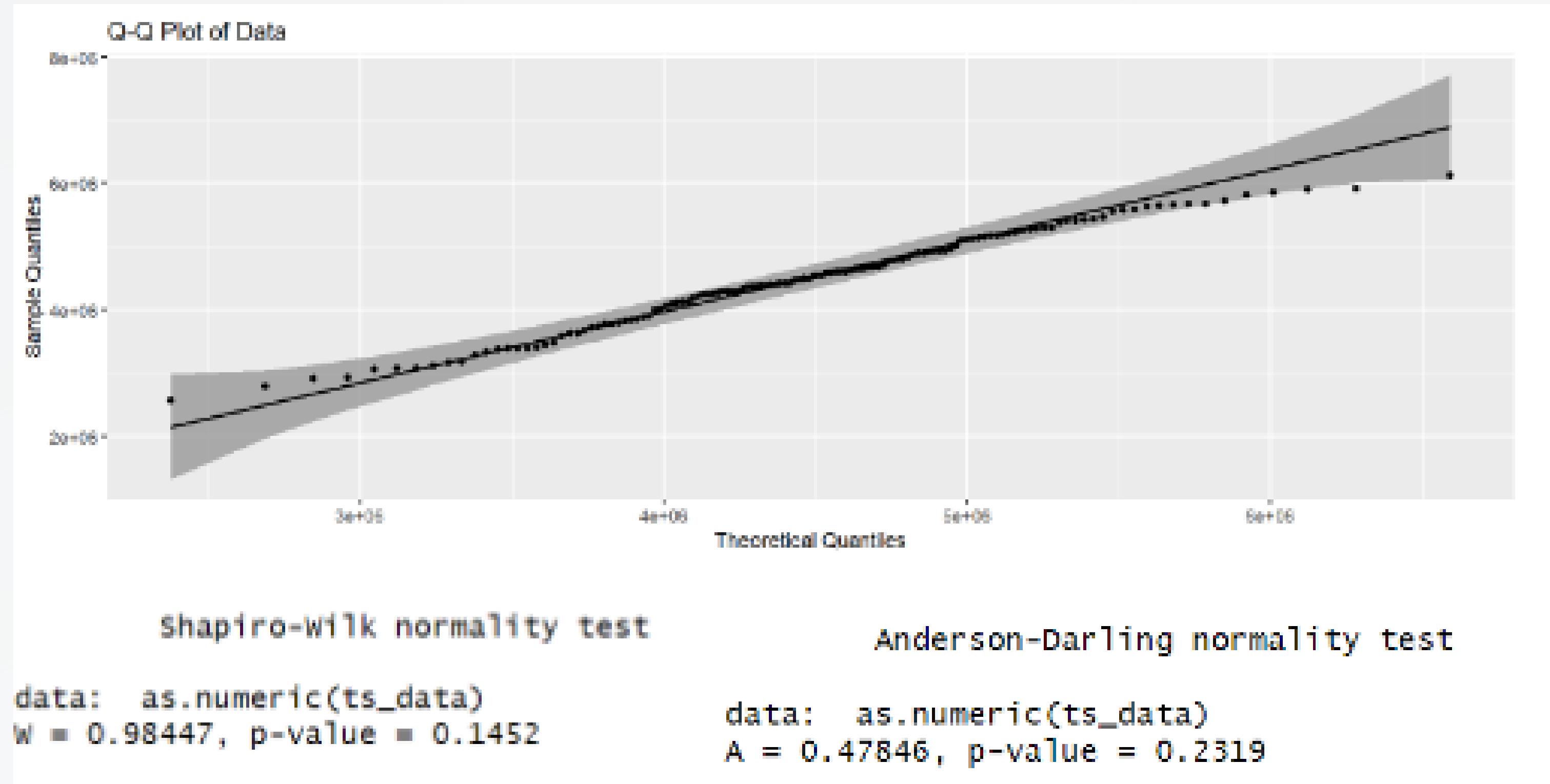
Jumlah Penumpang Pesawat Bulanan dari Tahun 2009-2019



PRE-PROCESSING



PRE-PROCESSING



PRE-PROCESSING

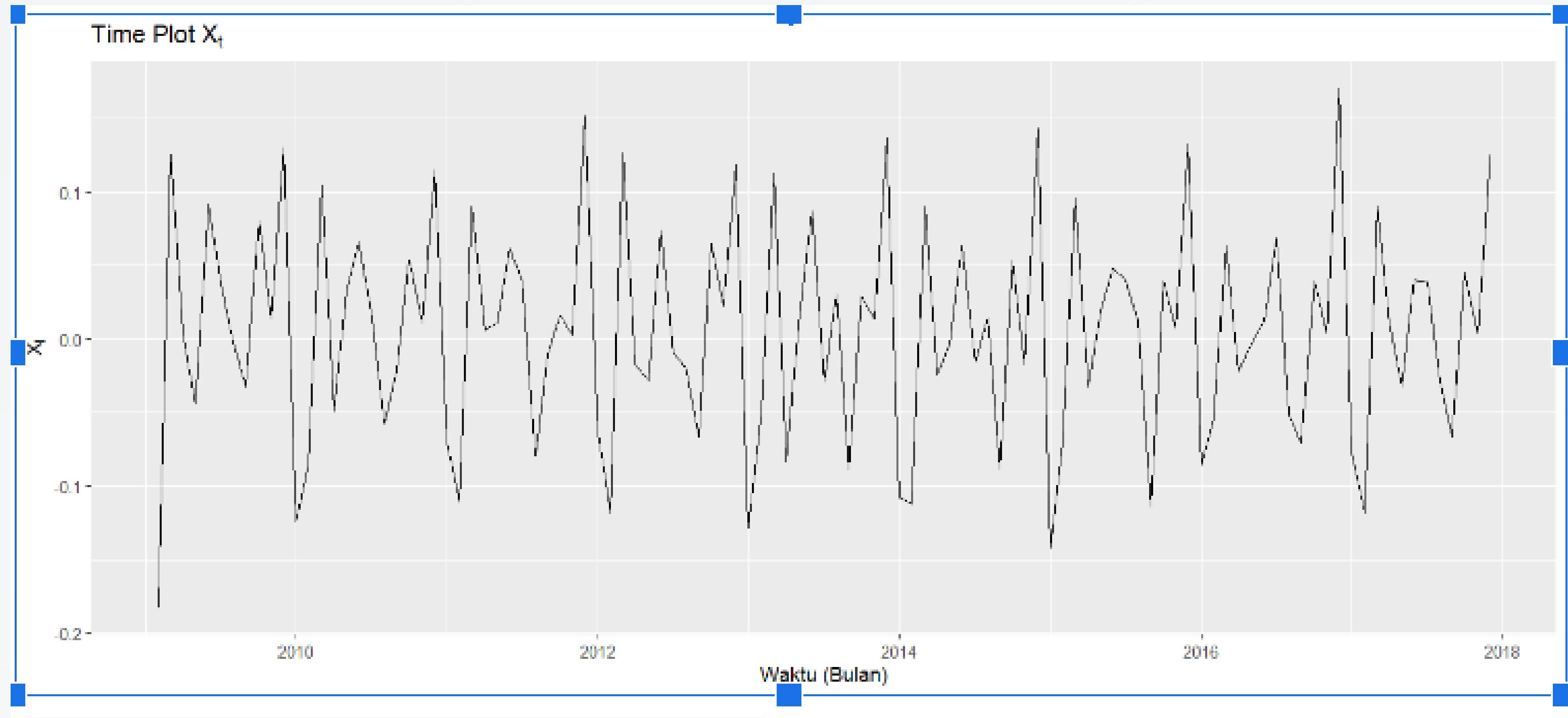
```
> print(adf_test)

Augmented Dickey-Fuller Test

data: ts_data
Dickey-Fuller = -3.4031, Lag order = 5, p-value = 0.05717
alternative hypothesis: stationary
```

Dari hasil uji Dicky-Fuller di atas, p-value = 0.05717 > 0.05. Hal ini membuat kami akan menerima H₀, sehingga benar bahwa data belum stasioner

PRE-PROCESSING



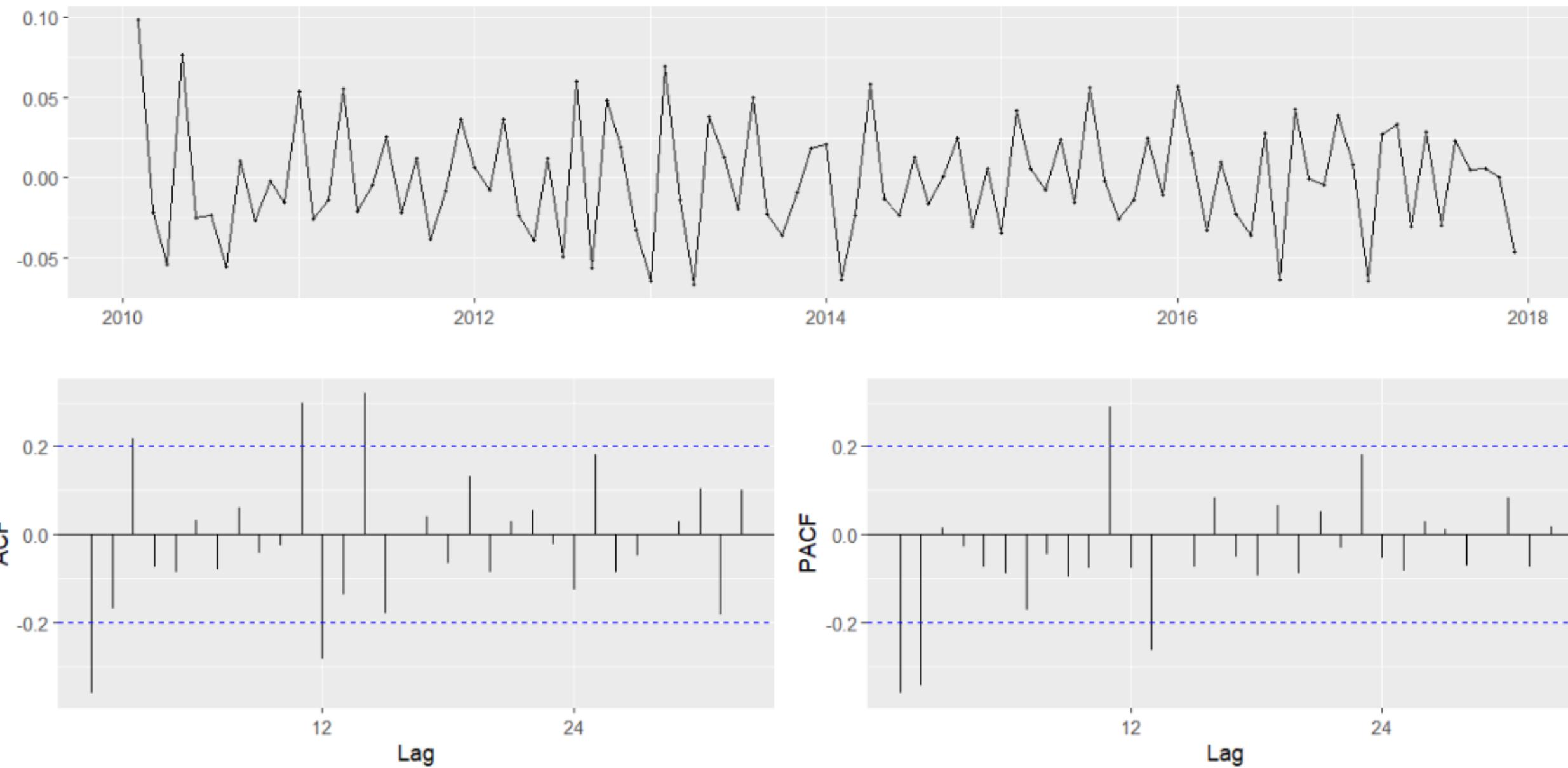


HASIL DAN PEMBAHASAN

untuk menganalisis, menginterpretasi, dan menyajikan temuan serta implikasi dari hasil penelitian secara komprehensif.

ACF DAN PACF

Dari plot sebelumnya kami dapat mendeteksi plot musiman. Oleh karena itu, akan digabungkan differencing regular (nonseasonal) dan differencing seasonal untuk spesifikasi model selanjutnya.



Akan dibuat plot
ACF dan PACF di
mana $d = 1$, $D = 1$,
dan $s = 12$.

EACF

```
> eacf(diff(diff(log(train_ts), lag = 12)))  
AR/MA  
 0 1 2 3 4 5 6 7 8 9 10 11 12 13  
0 x o x o o o o o o x x o x  
1 x x o o o o o o o o x o x  
2 o o x o o o o o o o o o o  
3 x o o o o o o o o o o o o  
4 x o o o o o o o o o o x o  
5 x x o o x o o o o o o o o  
6 x x x o o o o o o o o o o o  
7 o x x o o o o o o o o o x o
```

Berdasarkan plot EACF,
didapatkan pendugaan model
untuk uji coba yaitu:

- SARIMA(2, 1, 0)(0, 1, 1)12,
- SARIMA(0, 1, 0)(0, 1, 1)12,
- SARIMA(3, 1, 0)(0, 1, 1)12,
- SARIMA(4, 1, 0)(0, 1, 1)12,
- SARIMA(5, 1, 0)(0, 1, 1)12,

PERBANDINGAN MODEL

```
> summary(model1)
Series: train_ts
ARIMA(2,1,0)(0,1,1)[12]
Box Cox transformation: lambda= 0

Coefficients:
      ar1     ar2     sma1
    -0.5530  -0.3128  -0.6093
  s.e.  0.1022   0.1020   0.1298

sigma^2 = 0.0007796: log likelihood = 205.63
AIC=-403.26  AICC=-402.81  BIC=-393.04
```

```
Training set error measures:
      ME     RMSE     MAE      MPE     MAPE     MASE     ACF1
Training set -15136.79 110367.7 84559.94 -0.3719493 1.971531 0.3126665 0.001655673
```

```
> summary(model3)
Series: train_ts
ARIMA(3,1,0)(0,1,1)[12]
Box Cox transformation: lambda= 0

Coefficients:
      ar1     ar2     ar3     sma1
    -0.5303  -0.2693  0.0812  -0.6221
  s.e.  0.1060   0.1168  0.1069   0.1285

sigma^2 = 0.000781: log likelihood = 205.91
AIC=-401.83  AICC=-401.15  BIC=-389.06
```

```
Training set error measures:
      ME     RMSE     MAE      MPE     MAPE     MASE     ACF1
Training set -14240.13 109980.3 84803.4 -0.3515243 1.972711 0.3135667 -0.01812974
```

```
> summary(model2)
Series: train_ts
ARIMA(0,1,1)(0,1,1)[12]
Box Cox transformation: lambda= 0

Coefficients:
      ma1     sma1
    -0.5251  -0.5960
  s.e.  0.0891   0.1332

sigma^2 = 0.0007994: log likelihood = 204.04
AIC=-402.08  AICC=-401.82  BIC=-394.42
```

```
Training set error measures:
      ME     RMSE     MAE      MPE     MAPE     MASE     ACF1
Training set -16616.71 112178.7 86190.22 -0.4074624 2.013174 0.3186946 -0.046785
```

```
> summary(model4)
Series: train_ts
ARIMA(4,1,0)(0,1,1)[12]
Box Cox transformation: lambda= 0

Coefficients:
      ar1     ar2     ar3     ar4     sma1
    -0.5317  -0.2638  0.0931  0.0231  -0.6230
  s.e.  0.1062   0.1193  0.1201  0.1056   0.1278
```

```
sigma^2 = 0.0007892: log likelihood = 205.94
AIC=-399.88  AICC=-398.92  BIC=-384.55
```

```
Training set error measures:
      ME     RMSE     MAE      MPE     MAPE     MASE     ACF1
Training set -13900.3 110002.8 84748.34 -0.3434422 1.970552 0.3133631 -0.01240721
```

PERBANDINGAN AIC DAN BIC

```
> AIC(model1, model2, model3, model4, model5)
      df      AIC
model1  4 -403.2550
model2  3 -402.0792
model3  5 -401.8280
model4  6 -399.8760
model5  7 -400.0167
> BIC(model1, model2, model3, model4, model5)
      df      BIC
model1  4 -393.0395
model2  3 -394.4175
model3  5 -389.0587
model4  6 -384.5527
model5  7 -382.1396
```

Dari perbandingan di atas, model terbaiknya adalah
model 1 SARIMA(2, 1, 0)(0, 1, 1)12

ESTIMASI PARAMETER TRANSFORMASI BOX-COX

Transformasi Box-Cox dengan $\lambda=0$ ini setara dengan mengambil logaritma natural dari data. Jadi, kami mulai dengan data asli (Y_t) dan menerapkan transformasi log:

$$Z_t = \ln(Y_t)$$

Differencing Pertama dan Differencing Musiman

$$W_t = Z_t - Z_{t-1} = \ln(Y_t) - \ln(Y_{t-1})$$

$$V_t = W_t - W_{t-12} = (\ln(Y_t) - \ln(Y_{t-1})) - (\ln(Y_{t-12}) - \ln(Y_{t-13}))$$

Dari model SARIMA(2, 1, 0)(0, 1, 1)12, maka kami peroleh:

$$V_t = \phi_1 V_{t-1} + \phi_2 V_{t-2} - \theta_{s1} \epsilon_{t-12} + \epsilon_t$$

Selanjutnya kami substitusi nilai V_t , V_{t-1} , dan V_{t-2} , sehingga model matematis yang kami dapatkan yaitu

$$\begin{aligned} (\ln(Y_t) - \ln(Y_{t-1})) - (\ln(Y_{t-12}) - \ln(Y_{t-13})) &= -0.5530((\ln(Y_{t-1}) - \ln(Y_{t-2})) - \\ &(\ln(Y_{t-11}) - \ln(Y_{t-12}))) + -0.3128((\ln(Y_{t-2}) - \ln(Y_{t-3})) - (\ln(Y_{t-10}) - \ln(Y_{t-11}))) \\ &- 0.6093\epsilon_{t-12} + \epsilon_t \end{aligned}$$

UJI SIGNIFIKANSI ESTIMASI PARAMETER

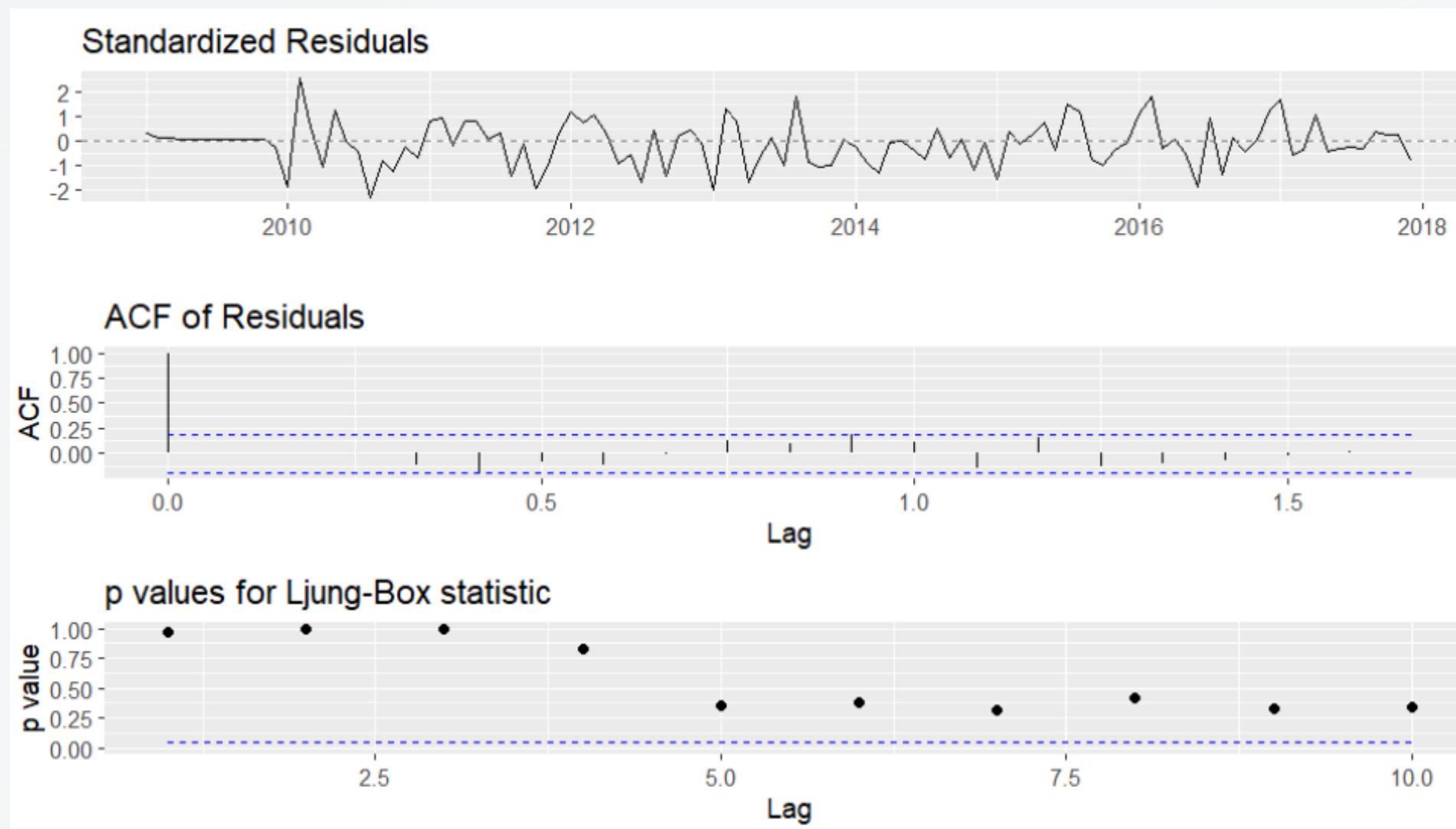
```
> coeftest(model1)

z test of coefficients:

    Estimate Std. Error z value Pr(>|z|)
ar1 -0.55301   0.10218 -5.4119 6.235e-08 ***
ar2 -0.31279   0.10201 -3.0664 0.002166 **
sma1 -0.60927   0.12976 -4.6953 2.662e-06 ***
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Semua koefisien (ar1, ar2, dan sma1) memiliki p-value yang kurang dari 0,05. yang berarti bahwa semua koefisien signifikan dan memberikan kontribusi yang signifikan terhadap model SARIMA.

MODEL DIAGNOSTIK



Plot Standarized Residual

Dari plot ini, residual tersebar secara acak dengan rata-rata mendekati nol, yang mengindikasikan model cukup baik dalam menangkap pola data.

ACF of Residuals

Sebagian besar nilai ACF residual berada dalam batas signifikan, yang menunjukkan tidak ada autokorelasi signifikan yang tersisa dalam residual. Namun, ada beberapa nilai ACF pada lag tertentu yang mendekati batas signifikan, tetapi mereka masih dalam batas wajar untuk model ARIMA.

PACF of Residuals

Sebagian besar nilai PACF berada dalam batas signifikan, mengindikasikan bahwa model telah menangkap sebagian besar pola dalam data dan residual tidak memiliki autokorelasi parsial yang signifikan.

UJI INDEPENDENSI (LJUNG-BOX TEST)

H₀: Residual independen

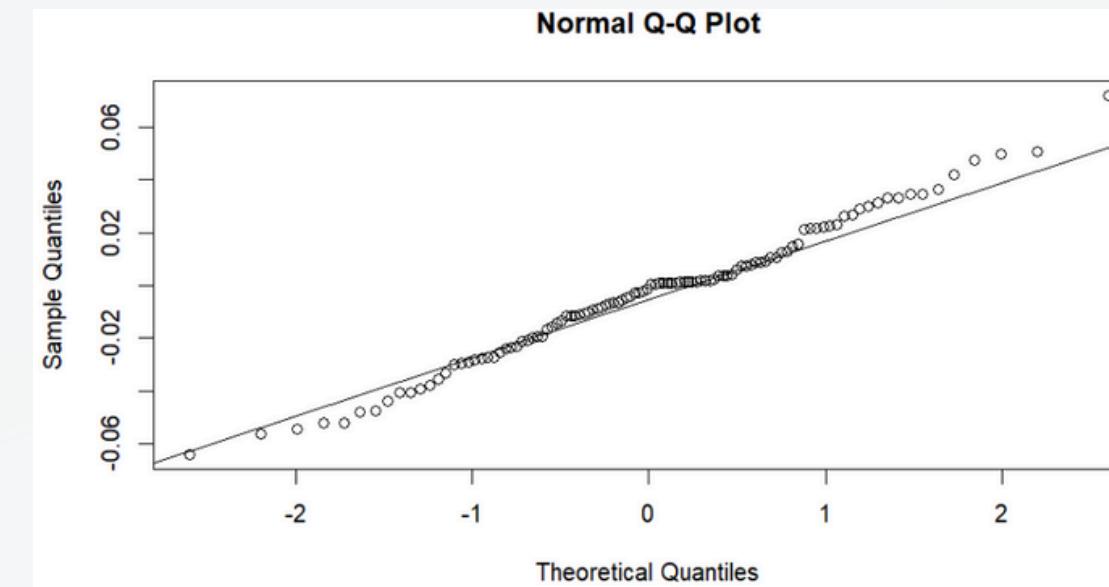
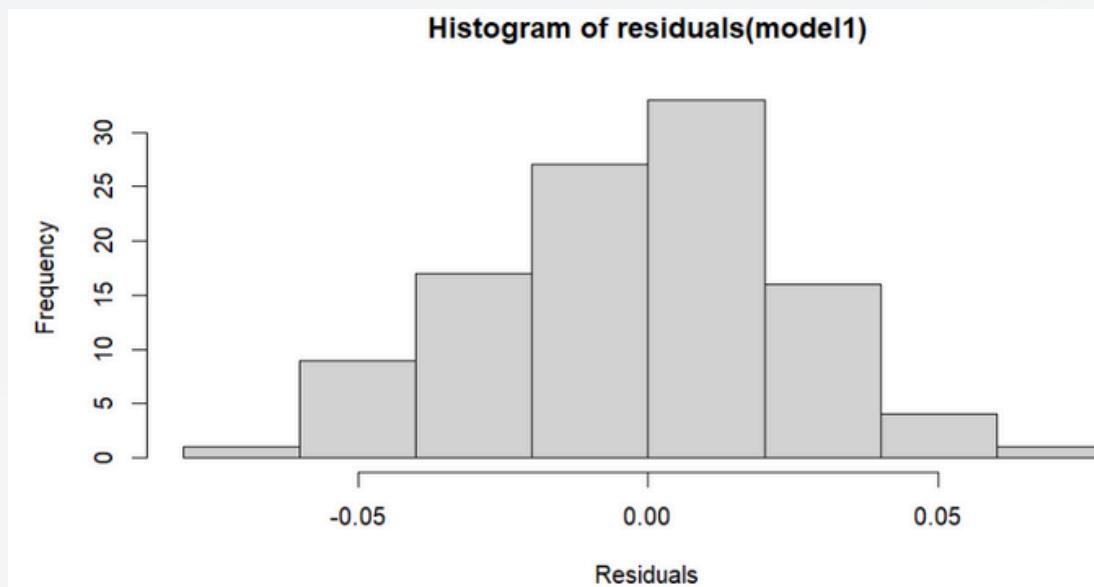
H₁: Residual tidak independen

```
> Box.test(residuals(model1), lag = 12, type = "Ljung-Box")
  Box-Ljung test

data: residuals(model1)
X-squared = 16.712, df = 12, p-value = 0.1608
```

Berdasarkan hasil Uji Ljung-Box, kami dapat bahwa p-value nya 0,1608 > 0,05. Hal ini mengindikasikan bahwa kami tidak menolak H₀ sehingga benar bahwa residual sudah independen.

UJI NORMALITAS (UJI SHAPIRO WILK)



Dari kedua plot ini, terlihat bahwa residual belum sepenuhnya memenuhi distribusi normal, terdapat sejumlah penyimpangan kecil dari kedua plot. Untuk memastikan apakah residual berdistribusi normal atau tidak akan dilakukan Uji Shapiro Wilk.

H₀: Residual berdistribusi normal

H₁: Residual tidak berdistribusi normal

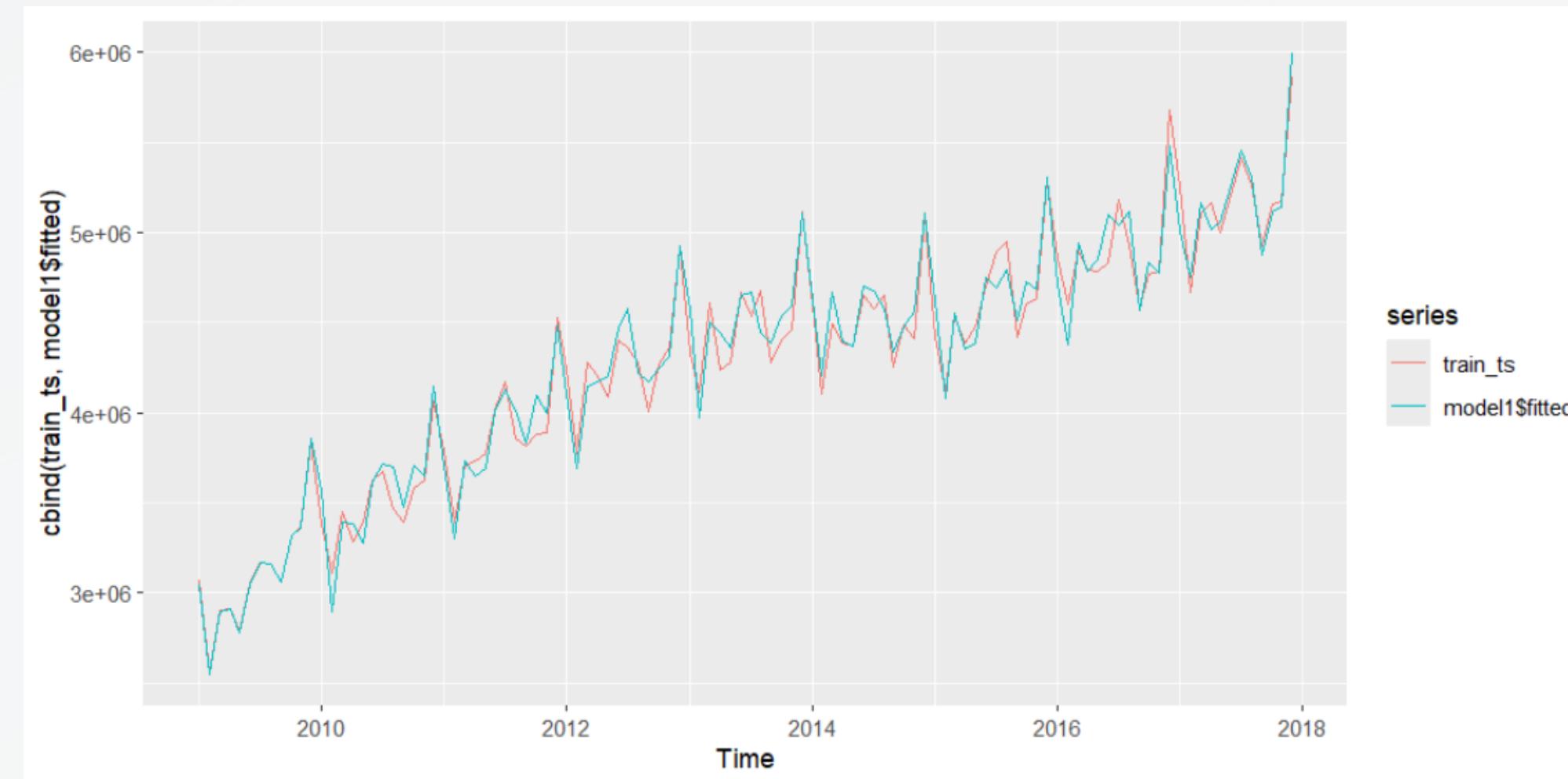
```
> shapiro.test(residuals(model1))  
Shapiro-Wilk normality test  
data: residuals(model1)  
W = 0.99, p-value = 0.6116
```

Dari hasil Uji Shapiro Wilk, kami dapat bahwa p-value nya $0,6116 > 0,05$. Hal ini mengindikasikan bahwa kami tidak menolak H₀ sehingga dapat kami simpulkan bahwa residual berdistribusi normal

Dari hasil-hasil uji model diagnostik di atas, didapat bahwa model SARIMA(2, 1, 0)(0, 1, 1)12 sudah cukup sesuai dan selanjutnya akan dilakukan forecasting menggunakan model SARIMA(2, 1, 0)(0, 1, 1)12.

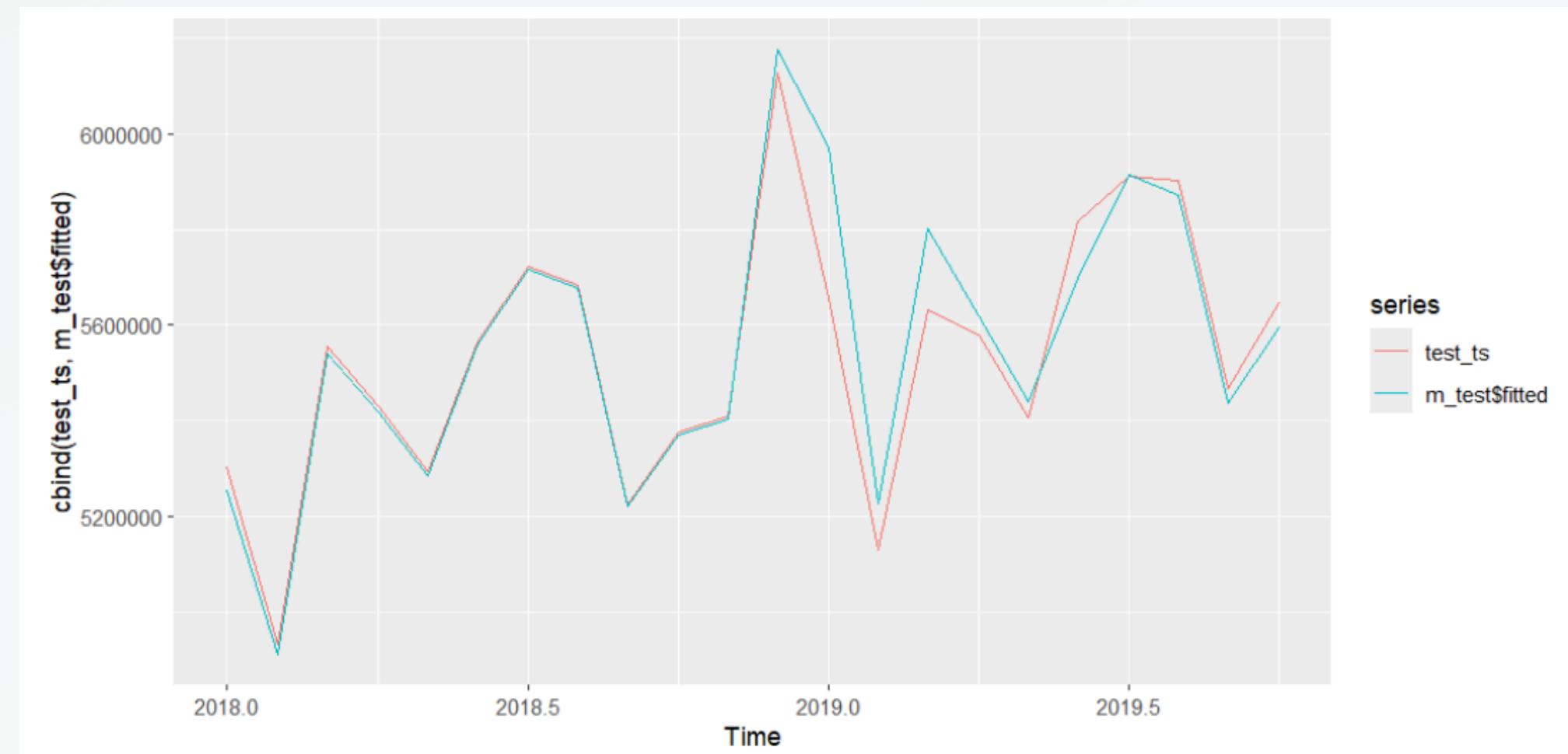
FORECASTING

FORECASTING



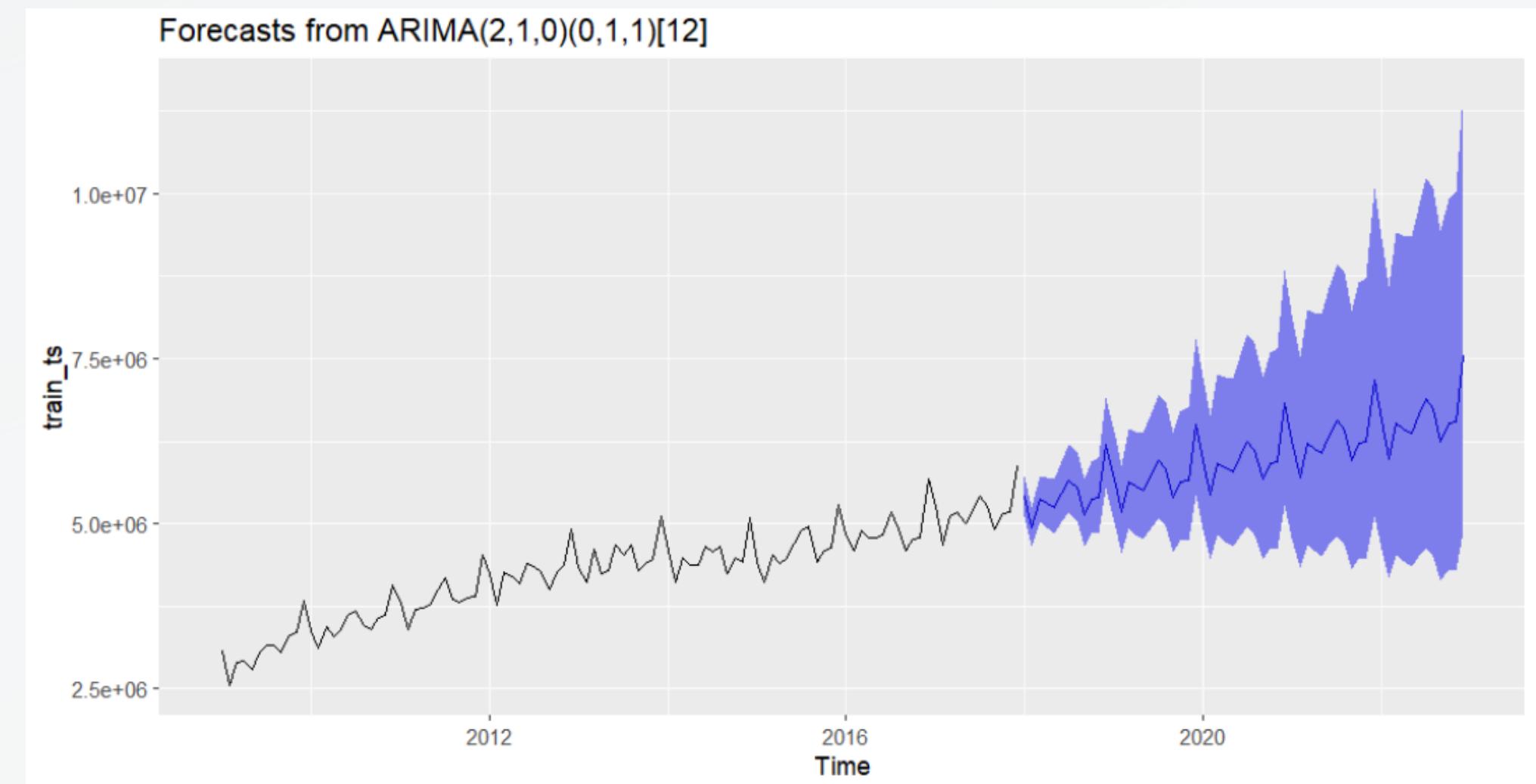
Hasil dari proses forecasting menggunakan model SARIMA(2, 1, 0)(0, 1, 1)[12] menunjukkan performa yang cukup baik. Berdasarkan metrik evaluasi, nilai Mean Absolute Percentage Error (MAPE) sebesar 8.68% menunjukkan bahwa model dapat melakukan prediksi jumlah penumpang dengan akurasi yang tinggi, yaitu rata-rata kesalahan prediksi kurang dari 10% dari nilai sebenarnya.

FORECASTING



Hasil dari plot ini menunjukkan bahwa data aktual (test set) dan nilai prediksi dari model SARIMA(2, 1, 0)(0, 1, 1)[12] saling berimpit secara baik. Hal ini mengindikasikan bahwa model yang telah dikembangkan mampu melakukan prediksi dengan baik terhadap data yang tidak digunakan dalam proses pelatihan (test set).

FORECASTING



Hasil dari plot forecast menunjukkan prediksi jumlah penumpang pesawat di Bandara Changi untuk periode yang akan datang (60 bulan ke depan) berdasarkan model SARIMA(2, 1, 0)(0, 1, 1)[12]. Garis biru menunjukkan nilai prediksi sedangkan area biru menunjukkan rentang prediksi dengan level kepercayaan 95%.

KESIMPULAN

KESIMPULAN

Analisis time series pada jumlah penumpang pesawat di Changi Airport dari tahun 2009 hingga 2019 telah memberikan sejumlah wawasan mengenai pola dan karakteristik data. Proses analisis melibatkan beberapa langkah, termasuk eksplorasi data, dekomposisi, pemodelan, dan evaluasi model. Berikut adalah kesimpulan dari analisis yang telah dilakukan:

1. Pola Musiman dan Tren: Hasil dari analisis menunjukkan adanya pola musiman yang kuat dalam data penumpang pesawat. Peningkatan jumlah penumpang terlihat pada bulan-bulan tertentu, Maret dan Desember, setiap tahun, yang konsisten dengan periode liburan dan musim liburan akhir tahun.
2. Transformasi dan Diferensiasi: Transformasi dan diferensiasi digunakan untuk menstabilkan varians dan menghilangkan tren. Hal ini bertujuan agar identifikasi komponen musiman dan tren lebih akurat.

KESIMPULAN

3. Pemodelan: Model SARIMA $(2,1,0)(0,1,1)[12]$ dipilih sebagai model terbaik. Model ini berhasil menangkap pola musiman dan tren dalam data dengan baik. Berikut taksiran untuk model terpilih

$$\begin{aligned} (\ln(Y_t) - \ln(Y_{t-1})) - (\ln(Y_{t-12}) - \ln(Y_{t-13})) &= -0.5530((\ln(Y_{t-1}) - \ln(Y_{t-2})) - \\ (\ln(Y_{t-11}) - \ln(Y_{t-12}))) + -0.3128((\ln(Y_{t-2}) - \ln(Y_{t-3})) - (\ln(Y_{t-10}) - \ln(Y_{t-11}))) \\ &- 0.6093\epsilon_{t-12} + \epsilon_t \end{aligned}$$

KESIMPULAN

4. Evaluasi Residual: Semua hasil analisis residual telah memenuhi asumsi diagnostik model. Hal ini menunjukkan bahwa model telah mengakomodasi struktur data dengan baik.
5. Peramalan: Model terpilih memiliki keakuratan yang cukup baik dalam memprediksi. Untuk hasil peramalan 5 tahun ke depan, diperolah hasil bahwa tren data yang naik akan terus berlanjut. Selain itu batas prediksi juga terus melebar seiring berjalannya waktu yang menunjukkan bahwa prediksi semakin tidak pasti seiring berjalannya waktu. Namun secara keseluruhan, model memberikan wawasan yang berguna untuk perencanaan masa depan dengan mempertimbangkan tren dan pola musiman, meskipun dengan tingkat ketidakpastian yang meningkat seiring waktu

KESIMPULAN

6. Rekomendasi: Berdasarkan hasil analisis, model SARIMA $(2,1,0)(0,1,1)[12]$ dapat digunakan untuk membuat prediksi jangka pendek mengenai jumlah penumpang pesawat di Changi Airport. Namun, penting untuk melakukan evaluasi berkala dan menyesuaikan model dengan data terbaru untuk menjaga akurasi prediksi.

6.2 Saran dan Rekomendasi

Berdasarkan kesimpulan analisis time series kami pada jumlah penumpang pesawat di Changi Airport, berikut adalah rekomendasi dari kami:

1. Implementasi Model SARIMA: Disarankan untuk mengimplementasikan Model SARIMA $(2,1,0)(0,1,1)[12]$ sebagai alat prediksi untuk memperkirakan jumlah penumpang pesawat di masa mendatang. Model ini terbukti efektif dalam menangkap pola musiman dan tren yang ada dalam data historis.

KESIMPULAN

2. Pemantauan dan Evaluasi Berkala: Penting untuk melakukan pemantauan dan evaluasi berkala terhadap kinerja model. Hal ini akan membantu dalam mengidentifikasi perubahan pola atau tren baru yang mungkin terjadi, serta memastikan bahwa model tetap relevan dan akurat.
3. Penyesuaian Model dengan Data Terbaru: Model SARIMA perlu disesuaikan dengan data terbaru secara teratur untuk meningkatkan keakuratan prediksi. Penyesuaian ini mencakup pemilihan ulang parameter dan evaluasi ulang struktur model sesuai dengan dinamika baru dalam jumlah penumpang pesawat.

KESIMPULAN

4. Kolaborasi dengan Pihak Terkait: Kerja sama yang erat dengan pihak terkait seperti maskapai penerbangan, otoritas bandara, dan lembaga terkait lainnya penting untuk memperbaiki perencanaan kapasitas, manajemen antrian, dan strategi pelayanan pelanggan di Changi Airport.
5. Inovasi Teknologi dan Analisis Data: Investasi dalam teknologi canggih untuk analisis data, seperti kecerdasan buatan dan big data, dapat meningkatkan kemampuan prediksi dan pengambilan keputusan yang lebih baik di masa depan.

Dengan menerapkan rekomendasi ini, diharapkan Changi Airport dapat terus memperkuat posisinya sebagai salah satu hub penerbangan terkemuka di dunia dengan layanan yang handal dan responsif terhadap perubahan pasar dan kebutuhan penumpang.

THANK YOU

Lampiran Link Video:

[https://drive.google.com/drive/folders/
126XNzLKFhfeVFgEoZPt9L3LishyTB1Zq?
usp=drive_link](https://drive.google.com/drive/folders/126XNzLKFhfeVFgEoZPt9L3LishyTB1Zq?usp=drive_link)

