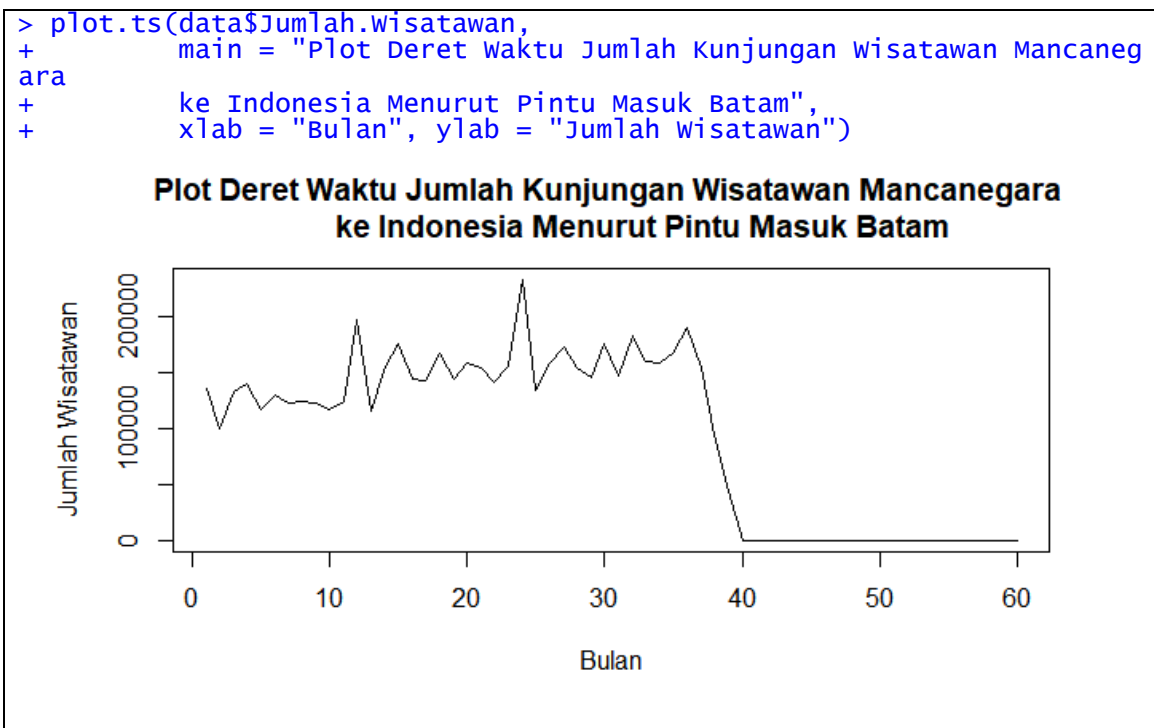


## Time Series Analysis

### 1. Import Data

```
> data <- read.csv("C:/Users/Indah/Documents/data jumlah wisatawan.csv",
+                 sep = ";", header = TRUE)
> head(data)
  No Jumlah.Wisatawan Tanggal
1  1      135046 01/01/2017
2  2      100329 01/02/2017
3  3      132551 01/03/2017
4  4      139361 01/04/2017
5  5      117372 01/05/2017
6  6      130541 01/06/2017
```

### 2. Standardisasi Data



Data memiliki pola tren linear negatif dengan ragam membesar atau tidak konstan dan data tidak stasioner.

### 3. Uji Stasioneritas

#### a. Uji Stasioneritas Ragam

```
> library(forecast)
> lambda <- BoxCox.lambda(data[-1])
> lambda
[1] 1
```

Lambda bernilai 1 sehingga data stasioner terhadap ragam, maka tidak perlu ditransformasi.

#### b. Uji Stasioneritas Rata-rata

```
> deret <- ts(data$Jumlah.Wisatawan, start = 2017, frequency = 12)
> library(tseries)
```

```

'tseries' version: 0.10-53
'tseries' is a package for time series analysis and computation
al
finance.

See 'library(help="tseries")' for details.

Warning message:
package 'tseries' was built under R version 4.2.3
> adf.test(deret)

Augmented Dickey-Fuller Test

data: deret
Dickey-Fuller = -1.9397, Lag order = 3, p-value = 0.5992
alternative hypothesis: stationary

```

Karena  $p\text{-value}$  (0.599)  $>$   $\alpha$  (0.05), maka data tidak stasioner terhadap rata-rata pada taraf signifikan 5%. Oleh karena itu, perlu dilakukan *differencing* terhadap data.

#### 4. Penanganan Data Tidak Stasioner

##### a. Differencing 1

```

> differencing.1_deret <- diff(deret, differences = 1)
> differencing.1 <- ts(differencing.1_deret)
> plot.ts(differencing.1,
+         main = "Plot Deret Waktu Jumlah Kunjungan Wisatawan Manca
negara
+         ke Indonesia Menurut Pintu Masuk Batam",
+         xlab = "Bulan", ylab = "Jumlah Wisatawan")

```



Plot data *differencing* pertama menunjukkan bahwa data tidak stasioner.

##### Uji Augmented Dickey Fuller (ADF)

```

> adf.test(differencing.1)

Augmented Dickey-Fuller Test

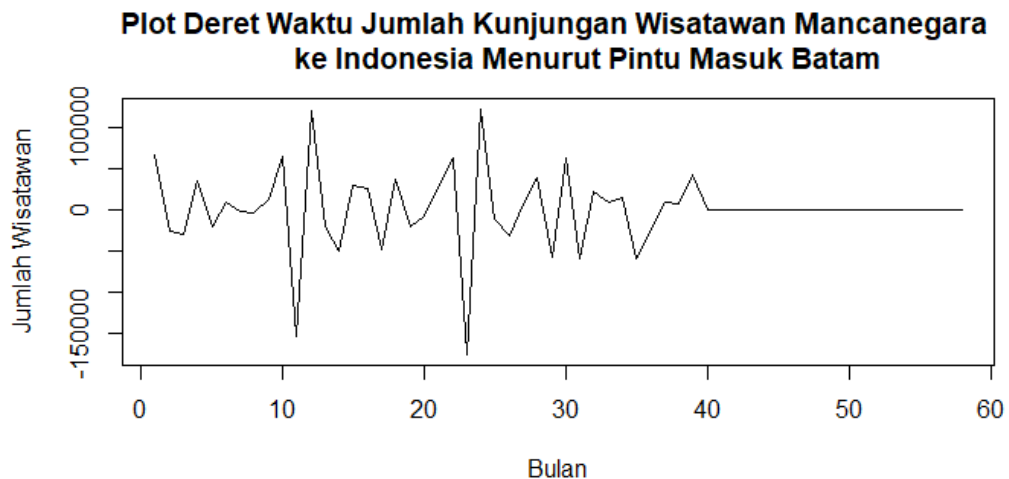
data: differencing.1
Dickey-Fuller = -3.1981, Lag order = 3, p-value = 0.09639
alternative hypothesis: stationary

```

Karena  $p\text{-value}$  (0.096)  $>$   $\alpha$  (0.05), maka data tidak stasioner terhadap rata-rata pada taraf signifikan 5%. Oleh karena itu, perlu dilakukan *differencing* kedua terhadap data.

## b. Differencing 2

```
> differencing.2_deret <- diff(deret, differences = 2)
> differencing.2 <- ts(differencing.2_deret)
> plot.ts(differencing.2,
+         main = "Plot Deret waktu Jumlah Kunjungan Wisatawan Manca
negara
+         ke Indonesia Menurut Pintu Masuk Batam",
+         xlab = "Bulan", ylab = "Jumlah Wisatawan")
```



Plot data *differencing* kedua menunjukkan bahwa data stasioner.

## Uji Augmented Dickey Fuller (ADF)

```
> adf.test(differencing.2)

Augmented Dickey-Fuller Test

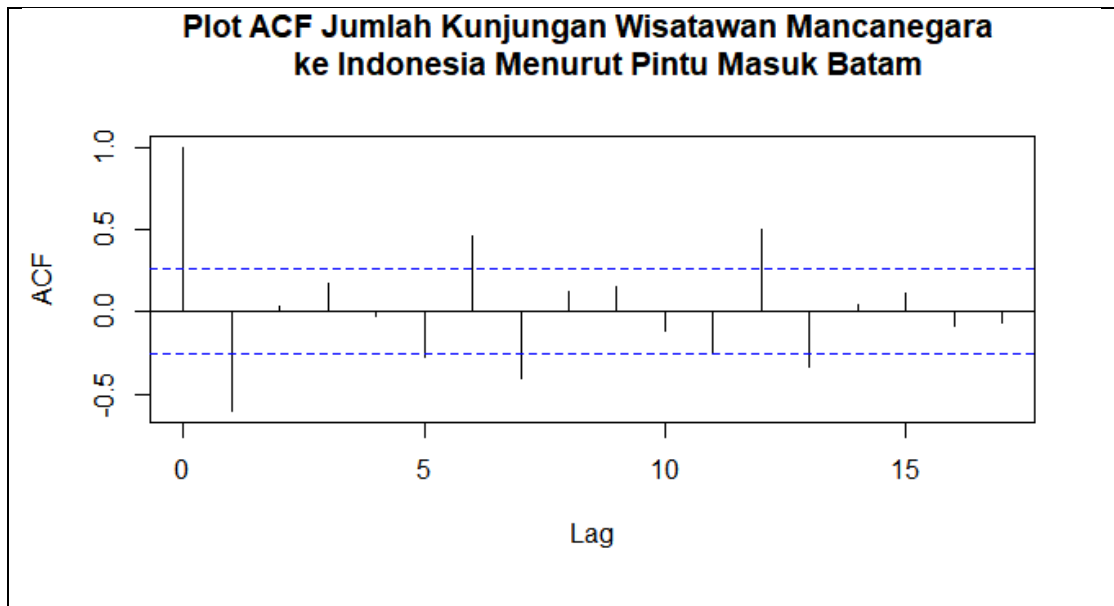
data: differencing.2
Dickey-Fuller = -5.1438, Lag order = 3, p-value = 0.01
alternative hypothesis: stationary
```

Karena  $p\text{-value}$  (0.01) <  $\alpha$  (0.05), maka data stasioner terhadap rata-rata pada taraf signifikan 5%. Oleh karena itu, model terbaik pada data dapat diidentifikasi.

## 5. Identifikasi Model

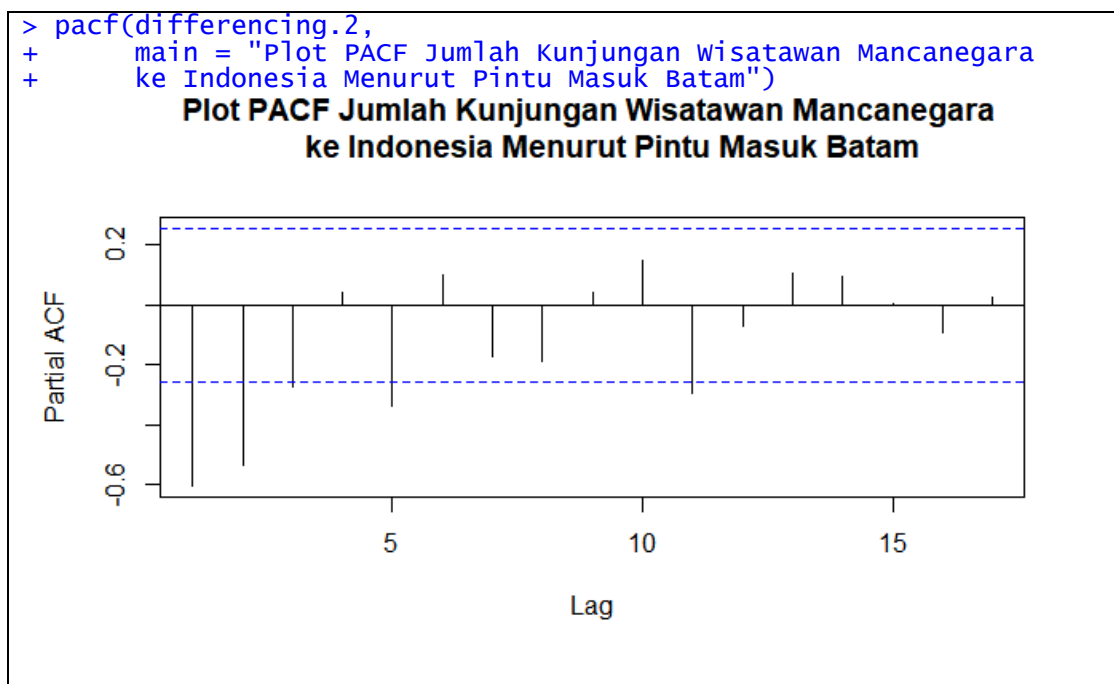
### a. Plot ACF

```
> acf(differencing.2,
+     main = "Plot ACF Jumlah Kunjungan Wisatawan Mancanegara
+     ke Indonesia Menurut Pintu Masuk Batam")
```



Autokorelasi mengalami *cut off* setelah lag 1 sehingga orde maksimum q adalah 1.

#### b. Plot PACF



Autokorelasi parsial mengalami *cut off* setelah lag 3 sehingga orde maksimum p adalah 3. Maka, model tentatif yang diperoleh adalah ARIMA (3,2,1).

#### 6. Pendugaan Parameter

```
> model <- arima(differencing.2, order = c(3,2,1), method = "ML")
> library(lmtest)
Loading required package: zoo
Attaching package: 'zoo'
The following objects are masked from 'package:base':
  as.Date, as.Date.numeric
```

```

Warning messages:
1: package 'lmtest' was built under R version 4.2.3
2: package 'zoo' was built under R version 4.2.3
> coeftest(model)

```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z )
ar1	-1.538786	0.107377	-14.3306	< 2.2e-16 ***
ar2	-1.322665	0.146232	-9.0449	< 2.2e-16 ***
ar3	-0.571481	0.104802	-5.4530	4.954e-08 ***
ma1	-0.999997	0.047204	-21.1847	< 2.2e-16 ***

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Model ARIMA (3,2,1):  $Z_t = -1.538Z_{t-1} - 1.322Z_{t-2} - 0.571Z_{t-3} + e_t - 0.999e_{t-1}$

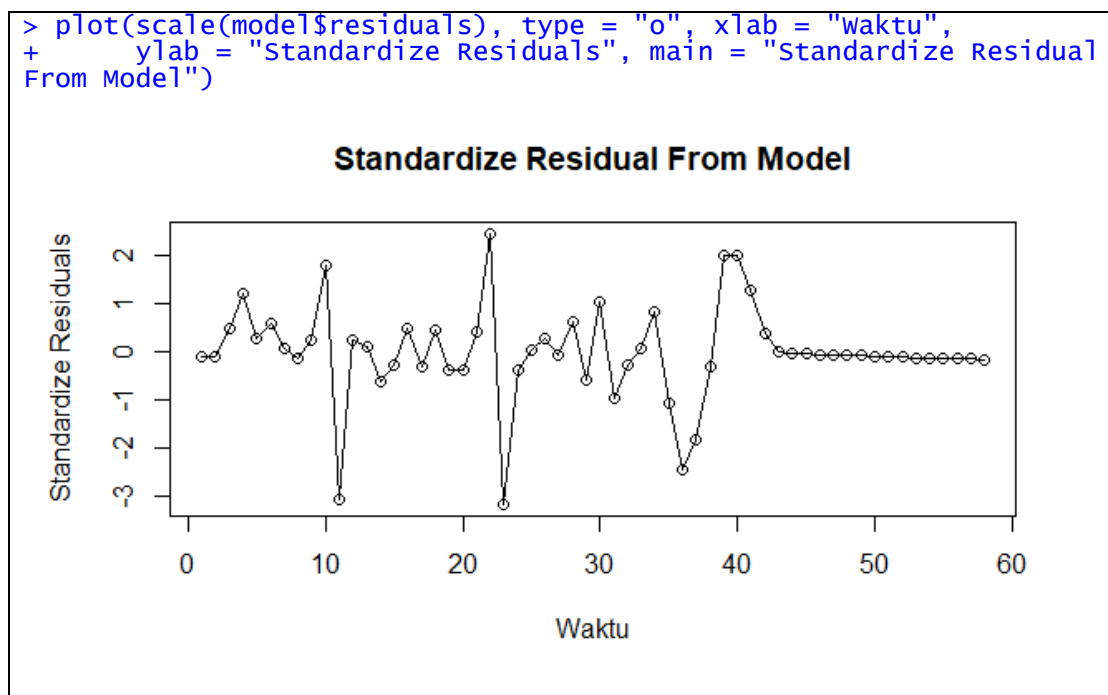
## 7. Uji Diagnostik

### a. Plot Residual

```

> plot(scale(model$residuals), type = "o", xlab = "waktu",
+       ylab = "Standardize Residuals", main = "Standardize Residual
+       From Model")

```



Terdapat variasi yang kecil pada pertengahan tahun 2017-2021 dan variasi konstan setelah bulan ke-44.

### b. Normalisasi Sisaan

```

> sisaan <- model$residuals
> shapiro.test(sisaan)

```

Shapiro-wilk normality test

```

data:  sisaan
W = 0.85769, p-value = 7.035e-06

```

Karena  $p\text{-value}$  (0.000)  $< \alpha$  (0.05), maka data tidak berdistribusi normal pada taraf signifikan 5%. Oleh karena itu, perlu dilakukan *overfitting*.

### c. Autokorelasi Residual

```

> Box.test(sisaan, type = "Ljung")

```

### Box-Ljung test

data: sisaan  
X-squared = 0.15985, df = 1, p-value = 0.6893

Karena  $p\text{-value}$  (0.902)  $> \alpha$  (0.05), maka tidak terdapat autokorelasi residual atau bersifat *white noise* pada taraf signifikan 5%.

## 8. Overfitting Model

### a. Model ARIMA (4,2,1)

```
> ARIMA421 <- arima(differencing.2, order = c(4,2,1), method = "ML")
> coeftest(ARIMA421)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
ar1	-1.601536	0.135013	-11.8620	< 2.2e-16	***
ar2	-1.462107	0.234463	-6.2360	4.489e-10	***
ar3	-0.729908	0.233594	-3.1247	0.00178	**
ar4	-0.099854	0.132201	-0.7553	0.45006	
ma1	-0.999997	0.047536	-21.0368	< 2.2e-16	***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Model ARIMA (3,2,1) masih lebih baik daripada model ARIMA (4,2,1) karena model tersebut memiliki parameter yang tidak signifikan.

### b. Model ARIMA (3,2,2)

```
> ARIMA322 <- arima(differencing.2, order = c(3,2,2), method = "ML")
> coeftest(ARIMA322)
```

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
ar1	-1.094128	0.130037	-8.4140	< 2.2e-16	***
ar2	-0.796467	0.164133	-4.8526	1.219e-06	***
ar3	-0.266060	0.128447	-2.0714	0.03833	*
ma1	-1.993678	0.079815	-24.9786	< 2.2e-16	***
ma2	0.999994	0.079689	12.5487	< 2.2e-16	***

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Seluruh parameter pada model ARIMA (3,2,2) signifikan. Oleh karena itu, dilakukan perbandingan pada nilai AIC dengan model ARIMA (3,2,1) untuk menentukan model ARIMA terbaik.

## 9. Model ARIMA Terbaik

```
> Model.ARIMA <- c("ARIMA (3,2,1)", "ARIMA (3,2,2)")
> Nilai.AIC <- c(model$aic, ARIMA322$aic)
> perbandingan <- data.frame(Model.ARIMA, Nilai.AIC)
> perbandingan
  Model.ARIMA Nilai.AIC
1 ARIMA (3,2,1) 1358.025
2 ARIMA (3,2,2) 1344.369
```

Model ARIMA (3,2,2) merupakan model terbaik karena memiliki nilai AIC yang lebih kecil daripada model ARIMA (3,2,1). Maka, model yang digunakan adalah ARIMA (3,2,2).

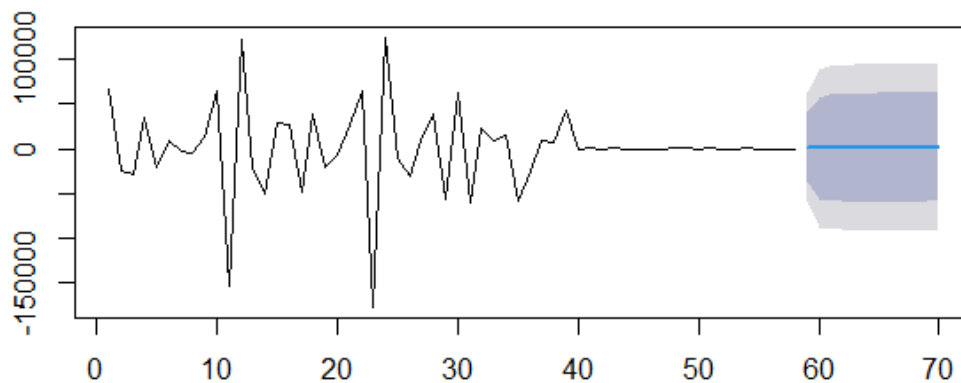
## 10. Peramalan

```
> model_forecast <- arima(differencing.2, order = c(3,2,2))
> ramalan_model <- forecast(model_forecast, h = 12)
> ramalan_model
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
59	2051.72672	-37492.64	41596.09	-58426.16	62529.61
60	76.18616	-57336.32	57488.69	-87728.67	87881.04
61	847.59681	-58738.16	60433.35	-90280.96	91976.15
62	1262.19127	-58776.06	61300.44	-90558.39	93082.77
63	949.85224	-59522.93	61422.63	-91535.29	93434.99
64	986.30115	-59488.46	61461.07	-91501.87	93474.47
65	1315.00698	-59477.22	62107.24	-91658.68	94288.70
66	1239.57154	-59586.42	62065.57	-91785.76	94264.90
67	1280.73565	-59548.16	62109.63	-91749.03	94310.50
68	1438.46223	-59465.03	62341.95	-91705.39	94582.31
69	1483.30659	-59423.74	62390.35	-91665.98	94632.59
70	1527.79960	-59384.73	62440.33	-91629.87	94685.47

```
> plot(ramalan_model)
```

Forecasts from ARIMA(3,2,2)



Hasil peramalan jumlah kunjungan wisatawan mancanegara ke Indonesia menurut jalur masuk Batam pada bulan Januari hingga Desember 2022 berada pada kisaran 2051 hingga 1527 kunjungan. Awal periode peramalan terjadi fluktuasi dan mengalami kenaikan yang konstan pada periode akhir peramalan.

## Kesimpulan

Kesimpulan yang dapat diperoleh dari hasil dan pembahasan pada analisis peramalan jumlah kunjungan wisatawan mancanegara ke Indonesia menurut jalur masuk Batam adalah sebagai berikut:

1. Model terbaik yang terbentuk adalah ARIMA (3,2,2) dengan bentuk persamaan sebagai berikut:

$$Z_t = -1.538Z_{t-1} - 1.322Z_{t-2} - 0.571Z_{t-3} + e_t - 0.999e_{t-1}$$

2. Hasil peramalan pada bulan Januari hingga Desember 2022 berada pada kisaran 2051 hingga 1527 kunjungan. Kenaikan jumlah kunjungan yang signifikan terjadi pada bulan Agustus hingga Desember 2022.

Saran yang dapat saya berikan apabila terjadi penurunan jumlah kunjungan wisatawan mancanegara yaitu pihak pengelola membuat strategi untuk meningkatkan kunjungan dengan memperhatikan daya dukung kawasan melalui penguatan promosi dan menambah fasilitas penunjang lainnya.