

**FINAL PROJECT METODE PERAMALAN
OPTIMALISASI MODEL ARIMA DALAM MEMPREDIKSI
FLUKTUASI HARGA SAHAM PT BANK MANDIRI (PERSERO) TBK.**



UNIVERSITAS INDONESIA

**MATA KULIAH : Metode Peramalan
DOSEN : Dra. Ida Fithriani, M.Si.**

**Disusun Oleh:
KELOMPOK 5**

Amira Shohifa	(2206829130)
Safira Ramadhani	(2206026656)
Gian Sinar Katulistiwa	(2206031675)
Renata Shaula Alfino Ritonga	(2206815812)

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI STATISTIKA
DEPOK
JUNI 2024**

DAFTAR ISI

DAFTAR ISI.....	1
BAB I PENDAHULUAN.....	2
1.1 Latar Belakang.....	2
1.2 Rumusan Masalah.....	3
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
BAB II METODE PENELITIAN.....	5
2.1 Deskripsi Data.....	5
2.2 Metode Analisis Data.....	5
a) Identifikasi Model.....	5
b) Spesifikasi Model.....	6
c) Diagnostik Model.....	6
d) Forecasting.....	7
2.3 Flowchart.....	8
BAB III ANALISIS DATA.....	9
3.1 Pembersihan Data.....	9
3.2 Identifikasi Model.....	11
3.3 Spesifikasi Model.....	14
3.4 Diagnostik Model.....	16
3.5 Forecasting.....	19
BAB IV PENUTUP.....	23
4.1 Kesimpulan.....	23
LAMPIRAN.....	25

BAB I PENDAHULUAN

1.1 Latar Belakang

Saham adalah instrumen keuangan yang mewakili kepemilikan atau bagian kepemilikan dalam sebuah perusahaan. Ketika seseorang membeli saham perusahaan, mereka secara efektif menjadi pemilik sebagian dari perusahaan tersebut. Harga saham bersifat fluktuatif sehingga memberi peluang kepada investor atau trader untuk mendapatkan keuntungan ataupun kerugian dari selisih harga beli dan harga jualnya (capital loss atau capital gain). Oleh karena itu, diperlukan keterampilan dari investor atau trader untuk dapat memperhitungkan kapan dan di harga berapa harus membeli maupun menjual sahamnya untuk mendapatkan keuntungan maksimal dan menghindari kerugian.

Bank Mandiri (BMRI) adalah salah satu bank terbesar di Indonesia dan memainkan peran penting dalam sektor perbankan nasional. Sebagai perusahaan publik, saham Bank Mandiri diperdagangkan di Bursa Efek Indonesia (BEI) dan menjadi objek investasi bagi para investor yang tertarik dengan industri perbankan. Para investor atau trader dapat melakukan analisis harga saham historis Bank Mandiri yang dapat ditemukan melalui sumber-sumber finansial dan pasar saham yang terpercaya, seperti laporan keuangan, berita industri, dan platform perdagangan saham.

Untuk menganalisis harga saham Bank Mandiri di masa depan, laporan peramalan ini menitikberatkan pada metode untuk memprediksi harga saham dan tren pasar di masa depan dengan mempelajari grafik dan data historis harga saham di masa lalu. Dalam laporan peramalan ini, akan dibahas pemodelan forecasting harga saham Bank Mandiri dengan metode Time Series Forecasting, yaitu metode ARIMA (p,d,q). Data historis harga saham diambil secara langsung dari situs penyedia data dengan menggunakan metode Web Scraping.

Pasar saham adalah tempat terjadinya perdagangan saham atau efek sekuritas perusahaan berlangsung. Pasar saham merupakan lingkungan yang kompleks dan dinamis di mana harga saham terus berubah seiring waktu. Menganalisis tren saham menjadi krusial bagi para pelaku pasar karena tren dapat mengungkapkan informasi berharga mengenai kekuatan pasar,

sentimen investor, dan kinerja perusahaan. Salah satu metode yang dapat digunakan adalah time series analysis.

Time series analysis adalah metode statistik yang digunakan untuk mengidentifikasi pola dan tren dalam data seiring waktu. Metode time series analysis melibatkan penggunaan data historis harga saham untuk mengidentifikasi pola dan tren. Dengan memanfaatkan teknik dan alat statistik yang tepat, peneliti dapat mengidentifikasi tren jangka pendek, menengah, dan panjang dalam harga saham. Analisis ini dapat dilakukan dengan memperhatikan metode-metode seperti moving averages, exponential smoothing, ARIMA (Autoregressive Integrated Moving Average), dan model-model lanjutan seperti GARCH (Generalized Autoregressive Conditional Heteroskedasticity). Penelitian ini akan memberikan pemahaman yang lebih baik tentang perilaku harga saham dan dapat digunakan sebagai dasar untuk membuat keputusan investasi yang lebih efektif.

Dengan demikian, penelitian ini memiliki potensi untuk memberikan wawasan yang berharga bagi para pelaku pasar, termasuk investor, analis, dan manajer portofolio, dalam pengambilan keputusan investasi.

1.2 Rumusan Masalah

Dalam pasar saham yang fluktuatif, harga saham Bank Mandiri (BMRI) mengalami perubahan yang dinamis dan dipengaruhi oleh berbagai faktor ekonomi dan non-ekonomi. Untuk membantu investor membuat keputusan yang tepat, diperlukan metode yang dapat memprediksi pergerakan harga saham secara akurat. Oleh karena itu, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana mengidentifikasi model ARIMA yang tepat untuk memprediksi harga saham Bank Mandiri?
2. Sejauh mana model ARIMA dapat digunakan untuk memprediksi harga saham Bank Mandiri dengan akurasi yang tinggi?
3. Faktor-faktor apa saja yang mempengaruhi keakuratan prediksi harga saham menggunakan model ARIMA?

1.3 Batasan Masalah

Penelitian ini memiliki beberapa batasan yang perlu diperhatikan, yaitu:

1. Data yang digunakan dalam penelitian ini hanya mencakup harga penutupan saham harian Bank Mandiri (BMRI).
2. Periode data historis yang digunakan terbatas pada data yang tersedia dari sumber data finansial yang dapat diakses.
3. Analisis hanya dilakukan menggunakan metode ARIMA, tanpa mempertimbangkan model-model time series lainnya seperti SARIMA, GARCH, atau machine learning.
4. Faktor eksternal seperti kondisi ekonomi makro, berita pasar, dan kejadian-kejadian khusus tidak dimasukkan ke dalam model prediksi.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengidentifikasi model ARIMA yang paling sesuai untuk memprediksi harga saham Bank Mandiri.
2. Mengukur keakuratan model ARIMA dalam memprediksi harga saham Bank Mandiri.
3. Memberikan rekomendasi bagi investor tentang penggunaan model ARIMA dalam pengambilan keputusan investasi saham Bank Mandiri.

BAB II METODE PENELITIAN

2.1 Deskripsi Data

Data yang akan digunakan dalam penelitian ini adalah data historis bulanan harga saham PT Bank Mandiri (Persero) Tbk. (BMRI.JK) yang diperoleh dari situs <https://finance.yahoo.com/quote/BMRI.JK/history/?frequency=1mo&period1=1059696000&period2=1717200000>. Data diambil dari periode Agustus 2003 sampai Mei 2024. Data tersebut terdiri dari 7 kolom, yaitu: *date*, *open*, *high*, *low*, *close*, *adj close*, dan *volume*. Penjelasan mengenai kolom data tersebut adalah sebagai berikut:

- **Date:** Tanggal (Format Bulan / Hari / Tahun)
- **Open:** Harga Pembukaan
- **High:** Harga Tertinggi
- **Low:** Harga Terendah
- **Close:** Harga Penutupan
- **Volume:** Volume transaksi biasanya dalam jumlah lembar
- **Adjusted Close:** Harga penutupan yang disesuaikan dengan aksi korporasi seperti right issue, stock split atau stock reverse

Untuk melakukan peramalan harga saham, kolom data yang paling penting untuk analisis adalah *close* (harga penutupan). Maka hanya kolom data *close* yang akan digunakan untuk pengolahan data lebih lanjut.

2.2 Metode Analisis Data

a) Identifikasi Model

Identifikasi model adalah langkah awal dalam analisis time series untuk menentukan jenis model yang sesuai dengan data yang diamati. Proses identifikasi model melibatkan beberapa langkah, termasuk:

1. **Analisis Plot Data:** Langkah pertama adalah menganalisis plot data time series untuk memahami pola dan tren yang terlihat. Plot ini dapat berupa grafik garis, diagram titik, atau plot lainnya yang menggambarkan perilaku data dari waktu ke waktu.

2. **Analisis Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF):** Plot ACF dan PACF membantu mengungkapkan pola autokorelasi dan autokorelasi parsial yang dapat digunakan untuk memilih model yang sesuai.
3. **Identifikasi Pola Tren dan Musiman:** Pola tren dapat berupa peningkatan atau penurunan secara konsisten dari waktu ke waktu, sementara pola musiman adalah fluktuasi reguler dalam data yang terjadi dalam periode waktu tertentu.
4. **Uji Stasioneritas:** Stasioneritas adalah sifat data yang memiliki statistik yang tidak berubah seiring waktu, yaitu rata-rata dan varians konstan. Uji stasioneritas, seperti Augmented Dickey-Fuller (ADF) test, dapat membantu dalam menentukan apakah data time series stasioner atau tidak.
5. **Pemilihan Model:** Berdasarkan analisis sebelumnya, dapat dipilih model yang sesuai dengan data time series. Misalnya, jika data memiliki pola tren linier, model ARIMA dapat digunakan. Jika data memiliki pola musiman, model SARIMA atau model yang berbasis musiman dapat digunakan.

b) Spesifikasi Model

Setelah melakukan identifikasi model dan didapatkan model yang stasioner, perlu dilakukan spesifikasi model untuk menentukan nilai-nilai parameter yang spesifik untuk model yang telah diidentifikasi. Langkah ini dilakukan dengan memilih orde AR, orde MA, orde differencing, dan orde musiman yang tepat untuk memodelkan data. Spesifikasi model berfokus pada penentuan nilai-nilai yang sesuai untuk parameter-parameter model yang telah dipilih. Spesifikasi model dapat dilakukan dengan melakukan analisis ACF, PACF, ataupun EACF.

c) Diagnostik Model

Diagnostik model adalah proses pengujian goodness of fit dari model yang telah didapatkan. Apabila hasil pengujian model tersebut masih buruk, maka perlu diberikan modifikasi pada model. Proses diagnostik model terdiri dari dua pendekatan:

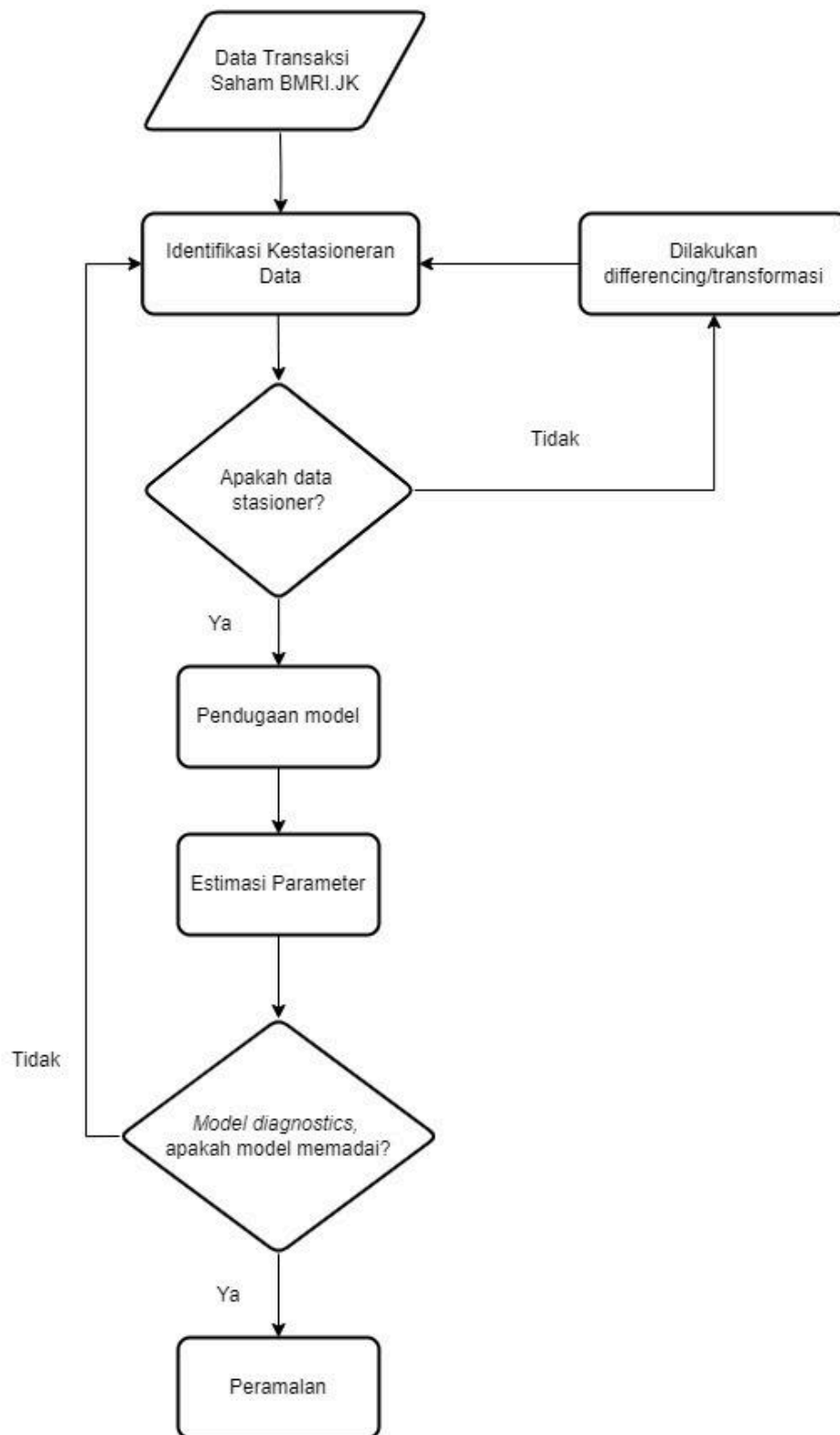
1. **Residual Analysis:** Bertujuan untuk menguji sifat independensi, normalitas, dan stasioneritas dari residual model melalui uji Ljung-Box, uji Jarque-Bera, uji Augmented Dickey-Fuller.

- a. **Uji Ljung-Box (Uji Independensi):**
 - i. Hipotesis:
 1. H_0 = tidak terdapat korelasi residual antar lag
 2. H_1 = terdapat korelasi residual antar lag
 - ii. Daerah kritis: Tolak H_0 jika p-value $< \alpha$
 - b. **Uji Jarque-Bera (Uji Normalitas):**
 - i. Hipotesis:
 1. H_0 = residual berdistribusi normal
 2. H_1 = residual tidak berdistribusi normal
 - ii. Daerah kritis: Tolak H_0 jika p-value $< \alpha$
 - c. **Uji Augmented Dickey-Fuller (Uji Stasioneritas):**
 - i. Hipotesis:
 1. H_0 = data tidak bersifat stasioner
 2. H_1 = data bersifat stasioner
 - ii. Daerah kritis: Tolak H_0 jika p-value $< \alpha$
2. **Overfitting:** Setelah dihasilkan model yang memadai berdasarkan beberapa pengujian di atas, akan diciptakan beberapa kemungkinan calon model lain dengan menambah parameter atau dibawa ke model yang lebih umum. Kemungkinan calon model ini harus tetap mengandung model memadai yang telah didapatkan sebelumnya. Semua model tersebut pada akhirnya akan diuji satu-persatu untuk mendapatkan model terbaik.

d) Forecasting

Time Series Forecasting adalah proses meramalkan nilai-nilai masa depan dari suatu data time series berdasarkan pola dan perilaku historis data tersebut. Tujuan dari time series forecasting adalah untuk menghasilkan perkiraan yang akurat dan relevan mengenai nilai-nilai yang akan datang dalam deret waktu. Pemilihan metode forecasting yang tepat tergantung pada sifat data time series, termasuk pola, tren, musiman, dan faktor lain yang mempengaruhi perilaku data. Evaluasi dan validasi model forecasting juga penting untuk memastikan akurasi dan keandalan perkiraan.

2.3 Flowchart



BAB III ANALISIS DATA

3.1 Pembersihan Data

Melakukan import library yang diperlukan selama proses pengolahan data dan load dataset.

```
## Final Project Metode Peramalan
## Kelompok 5
# Import library yang diperlukan
library(forecast)
library(tseries)
library(MASS)
library(TSA)
library(lmtest)

# Load data
BMRI <- read.csv("C:/Users/LENOVO/Downloads/BMRI.csv")
View(BMRI)
```

Tampilan 15 data awal dari dataset BMRI.

	Date	Open	High	Low	Close	Adj.Close	Volume
1	8/1/2003	184.37	208.95	172.08	208.95	96.17	2826545386
2	9/1/2003	202.81	221.24	190.52	208.95	96.17	2786226200
3	10/1/2003	208.95	251.97	202.81	233.53	107.49	3791995007
4	11/1/2003	233.53	239.68	196.66	227.39	104.66	2612881960
5	12/1/2003	227.39	258.12	221.24	245.83	113.14	2683297653
6	1/1/2004	245.83	368.74	239.68	319.57	154.83	5136810717
7	2/1/2004	319.57	387.18	301.14	344.16	166.74	3327351088
8	3/1/2004	350.30	381.03	307.28	344.16	166.74	3566894626
9	4/1/2004	338.01	381.03	325.72	356.45	172.69	4337698343
10	5/1/2004	356.45	356.45	276.55	307.28	148.87	3736463835
11	6/1/2004	307.28	313.43	258.12	288.85	139.94	3426311377
12	7/1/2004	288.85	319.57	282.70	307.28	157.40	1899295565
13	8/1/2004	301.14	319.57	282.70	307.28	157.40	1897837215
14	9/1/2004	307.28	362.59	307.28	356.45	182.58	3723999728
15	10/1/2004	350.30	393.32	350.30	387.18	198.32	2887905892

Dataset BMRI terdiri data historis bulanan harga saham PT Bank Mandiri (Persero) Tbk. dari bulan Agustus 2003 sampai bulan Mei 2024 yang berkuantitas sebesar 250 data. Dataset akan dibersihkan dengan membuang kolom data lain dan hanya menggunakan kolom data close (harga penutupan) untuk pengolahan data selanjutnya.

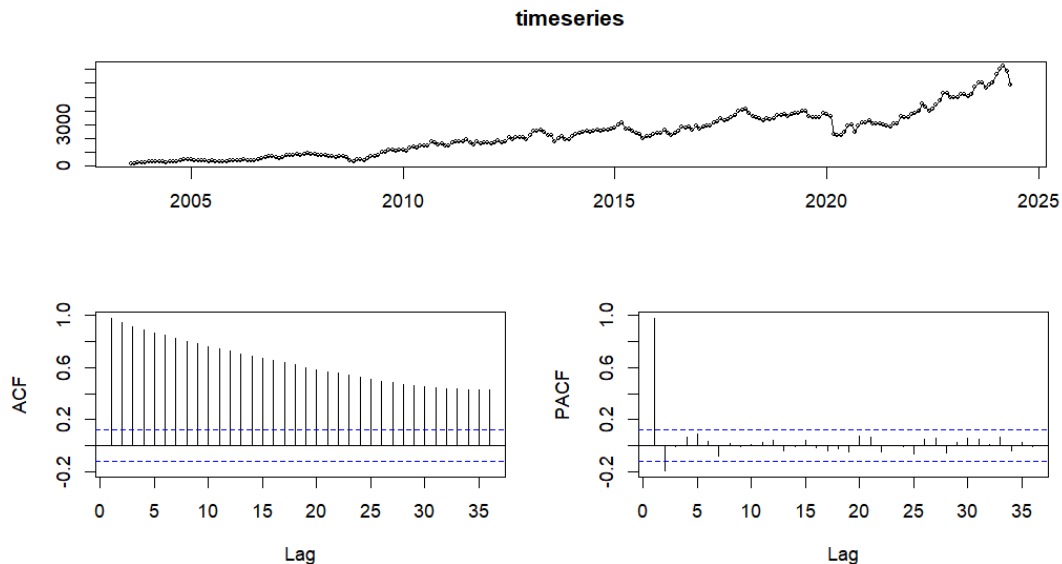
```
> # Menggunakan kolom close saja
> BMRI <- BMRI$close
> BMRI
```

[1]	208.95	208.95	233.53	227.39	245.83	319.57	344.16	344.16
[9]	356.45	307.28	288.85	307.28	307.28	356.45	387.18	448.63
[17]	473.21	476.90	444.94	420.36	393.32	408.07	368.74	395.78
[25]	336.78	356.45	324.49	317.12	403.15	437.57	398.24	427.74
[33]	471.99	415.45	422.82	432.65	516.23	571.55	669.88	682.17
[41]	712.90	633.00	571.55	614.56	755.91	798.93	768.21	866.54
[49]	798.93	866.54	927.99	872.68	860.39	817.37	805.08	774.35
[57]	706.75	712.90	639.15	731.33	694.46	651.44	383.49	366.28
[65]	497.80	447.40	427.74	534.67	682.17	731.33	780.50	1026.32
[73]	1007.89	1155.38	1149.24	1093.93	1155.38	1149.24	1100.07	1315.17
[81]	1425.79	1315.17	1474.96	1474.96	1450.37	1769.95	1720.78	1573.29
[89]	1597.87	1462.66	1450.00	1700.00	1787.50	1800.00	1800.00	1962.50
[97]	1712.50	1575.00	1787.50	1600.00	1687.50	1675.00	1612.50	1712.50
[105]	1850.00	1725.00	1800.00	2075.00	1950.00	2050.00	2062.50	2062.50
[113]	1950.00	2262.50	2512.50	2500.00	2625.00	2425.00	2250.00	2225.00
[121]	1775.00	1987.50	2150.00	1912.50	1962.50	2175.00	2275.00	2362.50
[129]	2456.25	2543.75	2431.25	2562.50	2593.75	2518.75	2587.50	2631.25
[137]	2693.75	2750.00	3000.00	3118.75	2687.50	2693.75	2512.50	2381.25
[145]	2275.00	1981.25	2175.00	2125.00	2312.50	2400.00	2387.50	2575.00
[153]	2412.50	2256.25	2381.25	2525.00	2806.25	2800.00	2868.75	2625.00
[161]	2893.75	2725.00	2825.00	2925.00	2925.00	3150.00	3187.50	3412.50
[169]	3275.00	3362.50	3525.00	3700.00	4000.00	4075.00	4150.00	3837.50
[177]	3562.50	3525.00	3425.00	3325.00	3450.00	3362.50	3425.00	3700.00
[185]	3687.50	3725.00	3562.50	3737.50	3862.50	3837.50	4012.50	3987.50
[193]	3625.00	3487.50	3512.50	3487.50	3837.50	3775.00	3637.50	2340.00
[201]	2230.00	2235.00	2475.00	2900.00	2975.00	2480.00	2887.50	3162.50
[209]	3162.50	3287.50	3075.00	3075.00	3087.50	3000.00	2950.00	2850.00
[217]	3050.00	3075.00	3587.50	3500.00	3512.50	3737.50	3850.00	3950.00
[225]	4475.00	4250.00	3962.50	4137.50	4425.00	4712.50	5275.00	5262.50
[233]	4962.50	4975.00	5000.00	5162.50	5175.00	5050.00	5200.00	5725.00
[241]	6025.00	6025.00	5675.00	5850.00	6050.00	6650.00	7000.00	7250.00
[249]	6900.00	5900.00						

3.2 Identifikasi Model

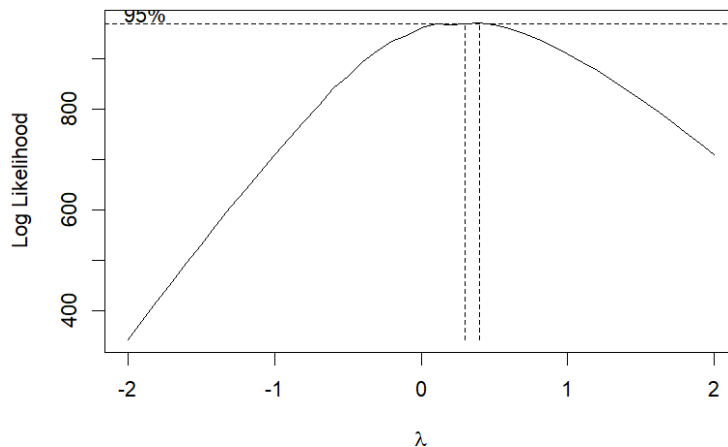
Melakukan konstruksi variabel time series.

```
# Konstruksi variabel time series  
timeseries <- ts(BMRI, start = c(2003, 8), frequency = 12)  
tsdisplay(timeseries)
```



Terlihat dari visualisasi grafik data variabel time series bahwa terdapat trend yang jelas atau terdapat kenaikan seiring waktu sehingga model time series di atas dianggap tidak stasioner. Grafik ACF menurun secara eksponensial dan pada grafik PACF terlihat tiang pancang pada lag ke-1. Karena model masing belum stasioner, maka diperlukan proses transformasi dan differencing.

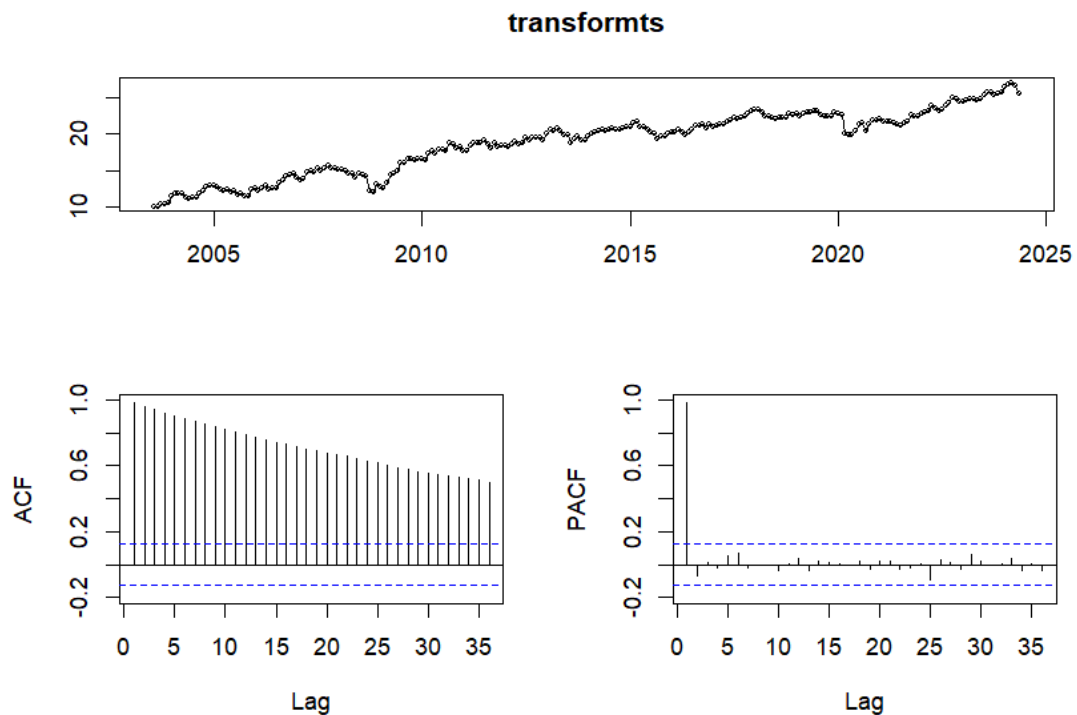
Model yang belum stasioner akan dimodifikasi dengan melakukan differencing dan transformasi Box-Cox agar model menjadi stasioner.

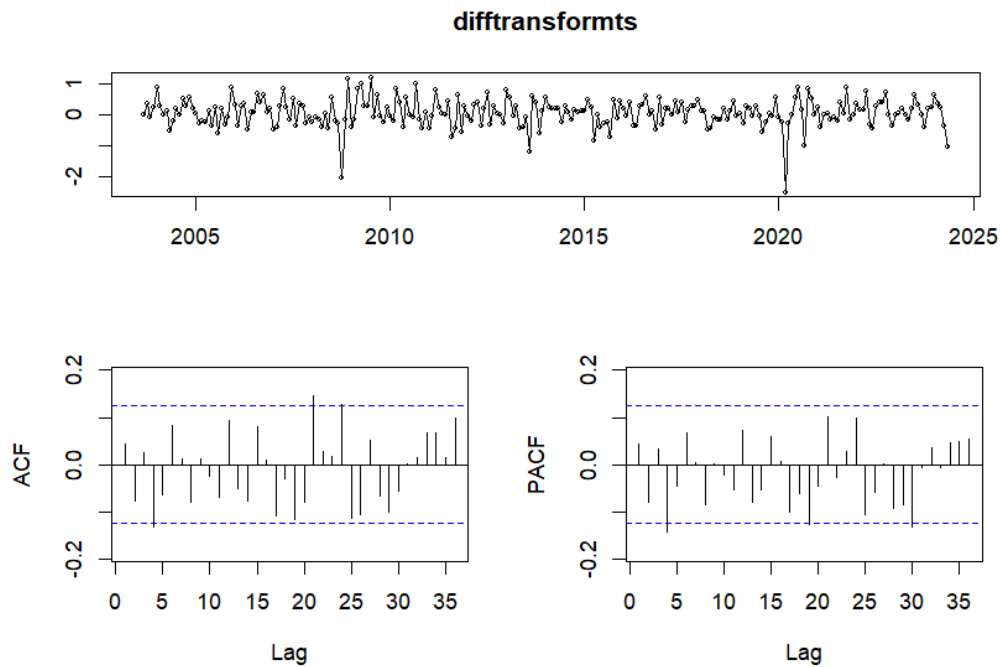


Box-Cox Transformation adalah sebuah teknik dalam analisis statistik yang digunakan untuk mentransformasi data yang tidak berdistribusi normal menjadi lebih mendekati distribusi normal. Lambda (λ) merupakan parameter transformasi yang akan digunakan untuk menentukan jenis transformasi. Interpretasi dari grafik Box-Cox di atas adalah bahwa transformasi logaritmik ($\lambda \approx 0$) adalah yang paling sesuai untuk meningkatkan distribusi data sehingga model yang digunakan dapat lebih baik dalam menggambarkan data tersebut.

```
> # Menentukan parameter Box-Cox (transformasi)
> BoxCox.lambda(timeseries)
[1] 0.216672
> transformts <- BoxCox(timeseries, BoxCox.lambda(timeseries))
```

Estimasi lambda yang didapatkan untuk transformasi Box-Cox adalah 0.216672. Sehingga akan dibuat model dengan transformasi Box-Cox dan dicek kembali kestasionerannya.





Dari visualisasi data setelah transformasi (transformts) masih terlihat tren yang jelas sehingga belum dapat dikatakan stasioner, namun apabila dilakukan differencing sebanyak satu kali setelah transformasi (difftransformts), terlihat bahwa deret waktu tampaknya berfluktuasi di sekitar nilai rata-rata yang konstan, tanpa tren yang jelas, maka dapat dikatakan bahwa model sudah bersifat stasioner.

```
> # Differencing setelah transformasi
> difftransformts <- diff(transformts, 1)
> tsdisplay(transformts)
> tsdisplay(difftransformts)
> # Melakukan uji
> adf.test(difftransformts)
```

Augmented Dickey-Fuller Test

```
data: difftransformts
Dickey-Fuller = -6.0884, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary
```

Pernyataan tersebut diperkuat dengan uji Augmented Dickey-Fuller dari model yang di-differencing satu kali setelah transformasi (difftransformts) dimana hasil uji menyatakan bahwa nilai $p\text{-value} = 0,01 < 0,05$ sehingga hipotesis nol ditolak yang artinya model difftransformts merupakan model yang bersifat stasioner.

3.3 Spesifikasi Model

Melakukan spesifikasi model, model yang akan dicari adalah model yang berbentuk ARIMA sehingga akan ditinjau melalui tabel EACF.

```
> # Spesifikasi model
> eacf(difftransformts)
AR/MA
  0  1  2  3  4  5  6  7  8  9 10 11 12 13
0 o o o x o o o o o o o o o o
1 x o o o o o o o o o o o o o
2 x x o o o o o o o o o o o o
3 x x o x o o o o o o o o o o
4 x x x x o o o o o o o o o o
5 x x x x o o o o o o o o o o
6 o x x x o x o o o o o o o o
7 o x x x x o o o o o o o o o
```

Berdasarkan hasil dari tabel EACF, akan dikonstruksikan calon model-model ARIMA. Dari tabel EACF diatas didapatkan beberapa calon model yang mungkin dapat digunakan, yaitu:

- Model 1 = ARIMA(0,1,3)
- Model 2 = ARIMA(2,1,0)
- Model 3 = ARIMA(3,1,1)
- Model 4 = ARIMA(3,1,2)
- Model 5 = ARIMA(4,1,2)
- Model 6 = ARIMA(5,1,2)

```
# Evaluasi model berdasarkan hasil EACF
# Model yang diajukan:
#ARIMA(0,1,3); ARIMA(2,1,0); ARIMA(3,1,1); ARIMA(3,1,2); ARIMA(4,1,2); ARIMA(5,1,2)
model_1 <- Arima(transformts, order = c(0, 1, 3), include.constant = TRUE)
model_2 <- Arima(transformts, order = c(2, 1, 0), include.constant = TRUE)
model_3 <- Arima(transformts, order = c(3, 1, 1), include.constant = TRUE)
model_4 <- Arima(transformts, order = c(3, 1, 2), include.constant = TRUE)
model_5 <- Arima(transformts, order = c(4, 1, 2), include.constant = TRUE)
model_6 <- Arima(transformts, order = c(5, 1, 2), include.constant = TRUE)
```

Masing-masing dari model tersebut kemudian diperiksa nilai AIC dan BIC nya dan akan dipilih model terbaik dengan nilai AIC dan BIC terendah.

```
# Menampilkan ringkasan model untuk membandingkan
summary(model_1)
summary(model_2)
summary(model_3)
summary(model_4)
summary(model_5)
summary(model_6)
cbind(model_1, model_2, model_3, model_4, model_5, model_6)

> cbind(model_1, model_2, model_3, model_4, model_5, model_6)
      model_1      model_2      model_3      model_4      model_5      model_6
coef      numeric,4      numeric,3      numeric,5      numeric,6      numeric,7      numeric,8
sigma2    0.1985655    0.198429    0.1989062    0.1939465    0.1955816    0.1962766
var.coef  numeric,16    numeric,9      numeric,25    numeric,36    numeric,49    numeric,64
mask      logical,4    logical,3    logical,5    logical,6    logical,7    logical,8
loglik    -150.0418    -150.4571    -149.7495    -147.1702    -146.6865    -146.6268
aic       310.0837     308.9143     311.499     308.3405     309.3729     311.2535
arma      integer,7    integer,7    integer,7    integer,7    integer,7    integer,7
residuals ts,250      ts,250      ts,250      ts,250      ts,250      ts,250
call      expression  expression  expression  expression  expression  expression
series    "transformts" "transformts" "transformts" "transformts" "transformts" "transformts"
code      0           0           0           0           0           0
n.cond    0           0           0           0           0           0
nobs      249        249        249        249        249        249
model     list,10     list,10     list,10     list,10     list,10     list,10
aicc      310.3306     309.0782     311.8462     308.8052     309.9729     312.0067
bic       327.671     322.9841     332.6038     332.9626     337.5125     342.9106
xreg      integer,250  integer,250  integer,250  integer,250  integer,250  integer,250
x         ts,250      ts,250      ts,250      ts,250      ts,250      ts,250
fitted    ts,250      ts,250      ts,250      ts,250      ts,250      ts,250
```

	model_1	model_2	model_3	model_4	model_5	model_6
loglik	-150.0418	-150.4571	-149.7495	-147.1702	-146.6865	-146.6268
aic	310.0837	308.9143	311.499	308.3405	309.3729	311.2535
bic	327.671	322.9841	332.6038	332.9626	337.5125	342.9106

Untuk memilih model terbaik, kita biasanya mempertimbangkan kriteria seperti AIC, BIC, dan log-likelihood. Nilai yang lebih rendah untuk AIC dan BIC menunjukkan model yang lebih baik. Log-likelihood yang lebih tinggi (lebih positif) juga menunjukkan model yang lebih baik.

Evaluasi Model Berdasarkan Kriteria

1. AIC

- Model 2 memiliki nilai AIC terendah: 308.9143
- Model 4 memiliki nilai AIC kedua terendah: 308.3405

2. BIC

- Model 2 memiliki nilai BIC terendah: 322.9841
- Model 1 memiliki nilai BIC kedua terendah: 327.671

3. Log-Likelihood

- Model 4 memiliki nilai log-likelihood tertinggi: -147.1702
- Model 5 memiliki nilai log-likelihood kedua tertinggi: -146.6865

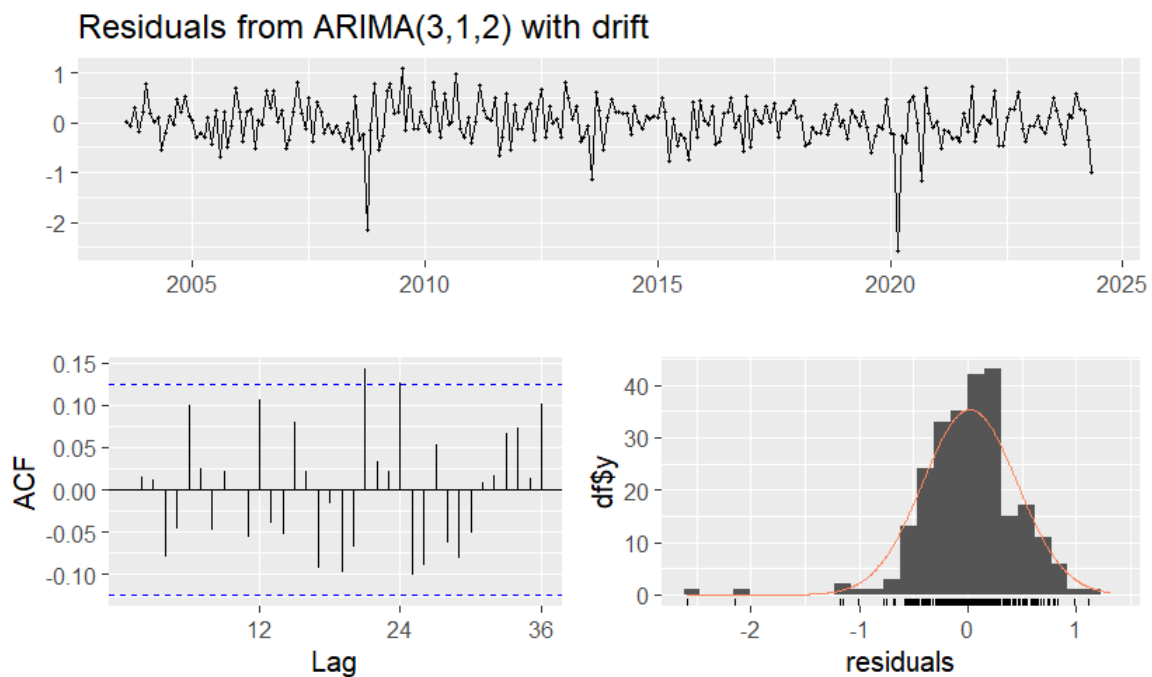
Dari evaluasi di atas, model yang tampaknya terbaik adalah **Model 4 (ARIMA(3,1,2))** karena memiliki nilai AIC kedua terendah dan nilai log-likelihood tertinggi. Model ini memberikan keseimbangan yang baik antara kecocokan model dan kompleksitasnya.

3.4 Diagnostik Model

1. Analisis Residual

Model terbaik dari tahap spesifikasi model akan diuji lebih lanjut dengan pengujian residual melalui uji Ljung-Box, uji Augmented Dickey-Fuller, dan uji Jarque-Bera

Plot residual analysis ARIMA(3,1,2).



```
> # Residual
> checkresiduals(model_4)
```

Ljung-Box test

```
data: Residuals from ARIMA(3,1,2) with drift
Q* = 29.071, df = 19, p-value = 0.06488
```

```
Model df: 5.    Total lags used: 24
```

Uji autokorelasi dilakukan untuk menguji apakah residual antar data memiliki korelasi atau tidak. Model yang baik memiliki nilai residual yang saling tidak berkorelasi satu dengan yang lainnya. Non-autokorelasi dapat diuji dengan Uji Ljung-Box. Dari output dapat dilihat bahwa $p\text{-value} = 0.06488 > 0.05$, artinya H_0 diterima, maka pada data tersebut residualnya independen.

```
> adf.test(model_4$residuals)
```

Augmented Dickey-Fuller Test

```
data: model_4$residuals
Dickey-Fuller = -5.6674, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary
```

Dilakukan pengecekan apakah residual stasioner atau tidak dengan menggunakan uji ADF (Augmented Dickey-Fuller). Dari output dapat dilihat bahwa $p\text{-value} = 0.01 < 0.05$, artinya H_0 ditolak, maka pada data residual tersebut stasioner.

```
> # Normality
> jarque.bera.test(model_4$residuals)
```

Jarque Bera Test

```
data: model_4$residuals
X-squared = 516.01, df = 2, p-value < 2.2e-16
```

Uji normalitas dilakukan untuk melihat kenormalan dari residual. Model baik jika residualnya mengikuti distribusi normal. Dari output dapat dilihat bahwa $p\text{-value} < 2.2e-16 < 0.05$, artinya H_0 ditolak, maka pada data tersebut residualnya tidak berdistribusi normal.

Dari pengujian di atas dapat disimpulkan bahwa residual dari model ARIMA (3,1,2) tidak mempunyai korelasi yang signifikan dan residual tidak berdistribusi normal sehingga residual mengikut proses white noise.

2. Overfitting

Selanjutnya akan dilakukan overfitting untuk melihat apakah model yang diajukan sudah model yang terbaik atau masih ada model lain yang lebih baik. Pada tahap ini akan dilakukan peningkatan orde pada AR maupun MA. Namun peningkatan orde tidak dilakukan bersamaan pada kedua model. Peningkatan orde dilakukan secara bertahap. Overfitting akan dilakukan dengan mengajukan model ARIMA (4,1,2) dan ARIMA (3,1,3). Model ARIMA (4,1,2) didefinisikan sebagai overfit_1. Sedangkan model ARIMA (3,1,3) didefinisikan sebagai overfit_2. Didapatkan output perbandingan ketiga model sebagai berikut.

```
> cbind(overfit_1, overfit_2, model_4)
```

	overfit_1	overfit_2	model_4
coef	numeric,7	numeric,7	numeric,6
sigma2	0.1955816	0.1944778	0.1939465
var.coef	numeric,49	numeric,49	numeric,36
mask	logical,7	logical,7	logical,6
loglik	-146.6865	-147.0015	-147.1702
aic	309.3729	310.003	308.3405
arma	integer,7	integer,7	integer,7
residuals	ts,250	ts,250	ts,250
call	expression	expression	expression
series	"transformts"	"transformts"	"transformts"
code	0	0	0
n.cond	0	0	0
nobs	249	249	249
model	list,10	list,10	list,10
aicc	309.9729	310.603	308.8052
bic	337.5125	338.1426	332.9626
xreg	integer,250	integer,250	integer,250
x	ts,250	ts,250	ts,250
fitted	ts,250	ts,250	ts,250

Nilai AIC dan BIC:

- Model ARIMA(3,1,2) (disebut model_4) memiliki AIC = 308.3405 dan BIC = 332.9626.
- Model ARIMA(4,1,2) (disebut overfit_1) memiliki AIC = 309.3729 dan BIC = 337.5125.
- Model ARIMA(3,1,3) (disebut overfit_2) memiliki AIC = 310.003 dan BIC = 338.1426.

Dari nilai AIC dan BIC, model ARIMA(3,1,2) memiliki nilai AIC dan BIC yang lebih rendah dibandingkan dengan dua model lainnya. Ini menunjukkan bahwa model ARIMA(3,1,2) lebih baik dalam hal keseimbangan antara kesesuaian dan kompleksitas model. Maka, model akhir yang akan digunakan untuk forecasting adalah Model 4, yaitu model ARIMA(3,1,2).

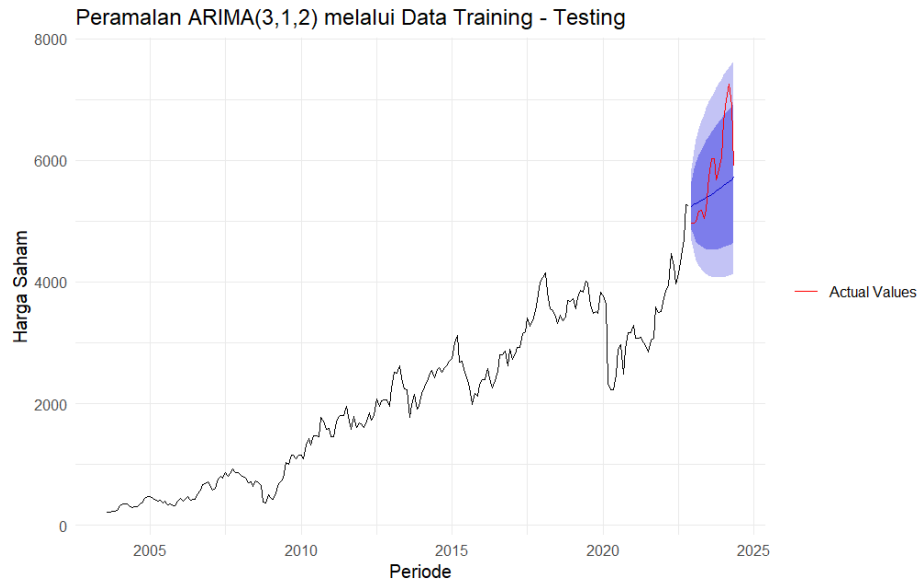
3.5 Forecasting

Langkah terakhir dalam analisis runtun waktu adalah peramalan atau prediksi untuk periode berikutnya. Akan dilakukan Cross Validation dengan membagi data menjadi 2, yaitu data testing dan data training. Pada dataset saham BMRI ini, data 12 bulan terakhir digunakan sebagai data testing.

```
# Cross Validation
# Membagi data menjadi training dan testing
train_end <- length(timeseries) - 12 # Misalnya 12 bulan terakhir sebagai data testing
train <- window(timeseries, end = c(2003 + (train_end %% 12), train_end %% 12 + 1))
test <- window(timeseries, start = c(2003 + (train_end %% 12), train_end %% 12 + 2))

# Melatih model ARIMA pada data training
model <- Arima(train, order = c(3, 1, 2), include.constant = TRUE, lambda = BoxCox.lambda(train))

# Menggunakan model untuk meramalkan data testing
forecast_result <- forecast(model, h = length(test))
```



Pada hasil plot di atas, garis hitam pada plot menunjukkan nilai aktual dari data saham BMRI, garis merah menunjukkan nilai aktual pada periode testing, dan area bayangan biru menunjukkan interval prediksi dengan tingkat kepercayaan 80% dan 95%.

Kinerja model ARIMA(3,1,2) jika dilihat secara visual, nilai peramalan (forecast) model ARIMA cenderung mengikuti pola dari data aktual, namun terdapat beberapa deviasi. Selain itu, hasil prediksi tampaknya mengikuti tren kenaikan harga saham, tetapi terkadang tidak menangkap fluktuasi yang lebih tajam pada data aktual.

```
> # Menampilkan hasil forecast dan data aktual
> print(cbind(Actual = test, Forecast = forecast_result$mean))
```

	Actual	Forecast
Dec 2022	4962.5	5245.787
Jan 2023	4975.0	5284.696
Feb 2023	5000.0	5292.542
Mar 2023	5162.5	5320.064
Apr 2023	5175.0	5338.600
May 2023	5050.0	5364.098
Jun 2023	5200.0	5387.508
Jul 2023	5725.0	5413.802
Aug 2023	6025.0	5440.098
Sep 2023	6025.0	5467.922
Oct 2023	5675.0	5496.337
Nov 2023	5850.0	5525.792
Dec 2023	6050.0	5555.964
Jan 2024	6650.0	5586.968
Feb 2024	7000.0	5618.678
Mar 2024	7250.0	5651.107
Apr 2024	6900.0	5684.194
May 2024	5900.0	5717.923

Dari kedua output di atas, dapat diambil beberapa kesimpulan dengan melihat hasil perbandingan nilai aktual dan prediksi dari model.

- Pada beberapa titik, seperti di bulan Desember 2022, Januari 2023, dan Februari 2023, peramalan cukup dekat dengan nilai aktual, tetapi pada beberapa periode seperti Juli 2023 dan Agustus 2023, ada deviasi yang cukup signifikan.
- Nilai prediksi untuk bulan Juli 2023 (5413.802) jauh lebih rendah dibandingkan dengan nilai aktual (5725.0). Hal ini menunjukkan bahwa model mungkin tidak cukup responsif terhadap perubahan tiba-tiba dalam data.
- Model ARIMA (3,1,2) cukup baik dalam menangkap tren umum dari data, namun kurang baik dalam menangkap variasi musiman atau perubahan mendadak.
- Sebagai contoh, pada bulan Januari 2024 dan Februari 2024, peramalan menunjukkan nilai yang lebih rendah dari nilai aktual yang menunjukkan kenaikan tajam.
- Interval prediksi memberikan gambaran tentang ketidakpastian dalam peramalan. Area biru yang lebih lebar menunjukkan ketidakpastian yang lebih besar.
- Nilai aktual sering berada di luar interval prediksi, menunjukkan bahwa model mungkin kurang akurat dalam meramalkan beberapa fluktuasi jangka pendek.

Secara keseluruhan, dapat dilihat bahwa semua nilai aktual (pada test) berada pada interval kepercayaan. Dengan demikian model yang digunakan (Model ARIMA (3,1,2)) untuk forecasting sudah baik karena semua nilai aktual sudah berada pada interval kepercayaan.

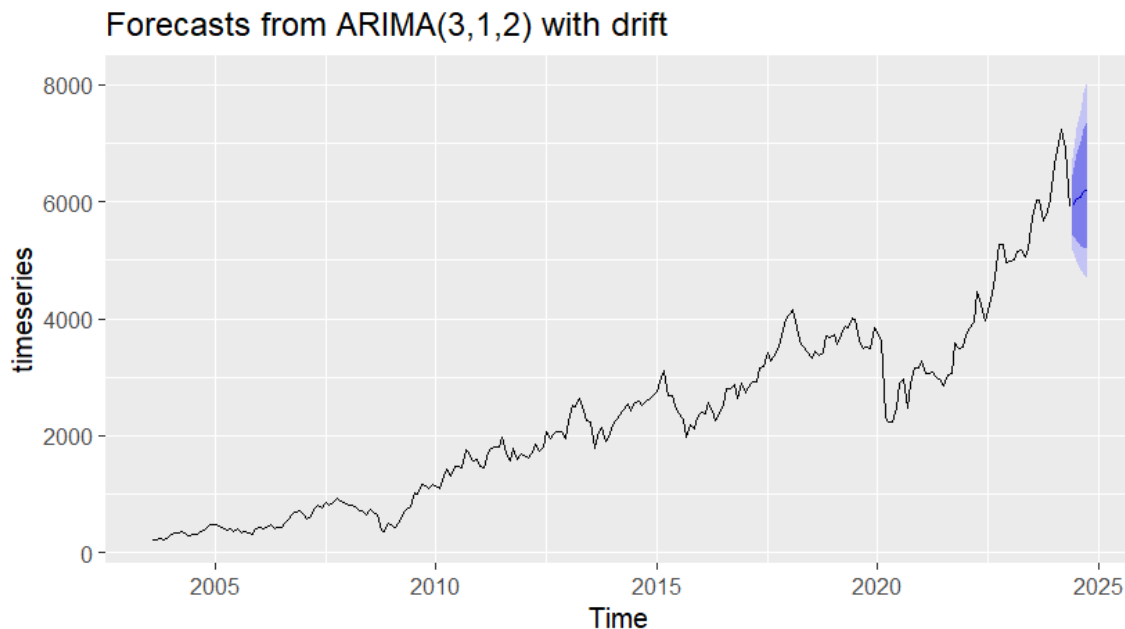
Kemudian, akan dilakukan proses forecasting 5 bulan ke depan dengan menggunakan model ARIMA(3,1,2)

```
> # Forecasting
> model_forecast <- Arima(timeseries, order = c(3,1,2), include.constant = TRUE,
+                          lambda = BoxCox.lambda(timeseries))
> model_forecast
Series: timeseries
ARIMA(3,1,2) with drift
Box Cox transformation: lambda= 0.216672

Coefficients:
      ar1      ar2      ar3      ma1      ma2      drift
    0.3847  0.4957  0.0251 -0.3588 -0.6412  0.0580
s.e.  0.2751  0.2975  0.0766  0.2687  0.2685  0.0056

sigma^2 = 0.1939:  log likelihood = -147.17
AIC=308.34  AICc=308.81  BIC=332.96
```

Model ARIMA(3,1,2) yang akan digunakan merupakan model random walk with drift dengan koefisien drift sebesar 0,0580. Model ini akan digunakan untuk melakukan forecasting harga penutupan dari saham PT Bank Mandiri (Persero) Tbk. untuk 5 bulan kedepan.



```
> # Peramalan 5 bulan kedepan
> forecast_result <- forecast(model_forecast, h = 5)
> autoplot(forecast_result)
> # Menampilkan hasil forecast
> print(forecast_result)
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jun 2024	5924.165	5431.107	6451.616	5183.454	6745.333
Jul 2024	6044.665	5336.115	6824.879	4988.364	7268.642
Aug 2024	6084.684	5251.839	7017.600	4848.718	7554.823
Sep 2024	6166.734	5223.448	7238.517	4772.155	7862.039
Oct 2024	6226.995	5198.800	7408.056	4711.280	8100.507

Dari output di atas dapat dilihat bahwa seluruh nilai peramalan berada pada interval kepercayaan 80% dan 95%, maka hasil forecasting pada output tersebut dapat dipercaya. Sehingga telah didapatkan hasil nilai forecasting harga penutupan dari saham PT Bank Mandiri (Persero) Tbk. selama lima bulan ke depan (Juni 2024 - Oktober 2024) dengan interval kepercayaan 80% dan 95%.

BAB IV PENUTUP

4.1 Kesimpulan

Berdasarkan hasil analisis runtun waktu pada data harga penutupan saham PT Bank Mandiri (Persero) Tbk., model ARIMA (3,1,2) telah diidentifikasi sebagai model yang paling sesuai untuk meramalkan harga saham BMRI berdasarkan data historis yang tersedia. Model ini menunjukkan performa yang baik dalam menangkap tren umum dari pergerakan harga saham, meskipun ada beberapa fluktuasi yang tidak terprediksi dengan akurat. Dari hasil pengujian menggunakan 12 bulan terakhir sebagai data testing, model ARIMA (3,1,2) mampu memberikan hasil peramalan yang cukup dekat dengan nilai aktual, terutama pada periode yang tidak mengalami perubahan tajam.

Proses analisis data runtun waktu menggunakan model ARIMA adalah:

1. Melakukan uji stasioneritas pada data dengan pembuatan plot, lalu melakukan differencing dan transformasi apabila data tersebut tidak stasioner.
2. Menentukan spesifikasi model yang tepat dengan menggunakan plot ACF, PACF, dan EACF.
3. Melakukan estimasi parameter dengan menggunakan metode momen, metode least square, dan metode maximum likelihood.
4. Melakukan diagnostik model, yaitu evaluasi kualitas model melalui analisis residual dan uji asumsi N.I.I.D untuk memastikan bahwa model tidak overfitting dan residualnya memenuhi asumsi distribusi normal.
5. Forecasting (peramalan atau prediksi untuk periode berikutnya).

Model runtun waktu terbaik untuk melakukan peramalan (forecasting) data harga penutupan saham PT Bank Mandiri (Persero) Tbk. adalah model ARIMA (3,1,2) dengan persamaan sebagai berikut:

$$\begin{aligned}
 W_t &= 0.0580 + 0.3847Y_{t-1} + 0.4957Y_{t-2} + 0.0251Y_{t-3} + e_t \\
 &\quad - 0.3588e_{t-1} - 0.6412e_{t-2} \\
 Y_t - Y_{t-1} &= 0.0580 + 0.3847Y_{t-1} + 0.4957Y_{t-2} + 0.0251Y_{t-3} + e_t \\
 &\quad - 0.3588e_{t-1} - 0.6412e_{t-2} \\
 Y_t &= 0.0580 + 1.3847Y_{t-1} + 0.4957Y_{t-2} + 0.0251Y_{t-3} + e_t \\
 &\quad - 0.3588e_{t-1} - 0.6412e_{t-2}
 \end{aligned}$$

Hasil penerapan forecasting untuk memprediksi harga penutupan saham PT Bank Mandiri (Persero) Tbk. untuk 5 bulan ke depan dengan interval kepercayaan (CI) 80% dan 95% adalah sebagai berikut.

Periode Ramalan	Hasil Peramalan	CI 80%	CI 95%
Juni 2024	5924.165	[5431.107, 6451.616]	[5183.454, 6745.333]
Juli 2024	6044.665	[5336.115, 6824.879]	[4988.364, 7268.642]
Agustus 2024	6084.684	[5251.839, 7017.600]	[4848.718, 7554.823]
September 2024	6166.734	[5223.448, 7238.517]	[4772.155, 7862.039]
Oktober 2024	6226.995	[5198.800, 7408.056]	[4711.280, 8100.507]

LAMPIRAN

Seluruh lampiran dari project akhir mata kuliah ini, terlampir pada tautan google drive berikut.

📁 Final Project Metode Peramalan Kelompok 5

<https://drive.google.com/drive/folders/1ugcQqyFBAX0k8GJOb-dTdQhq4m6aAXGI?usp=sharing>

[g](#)