



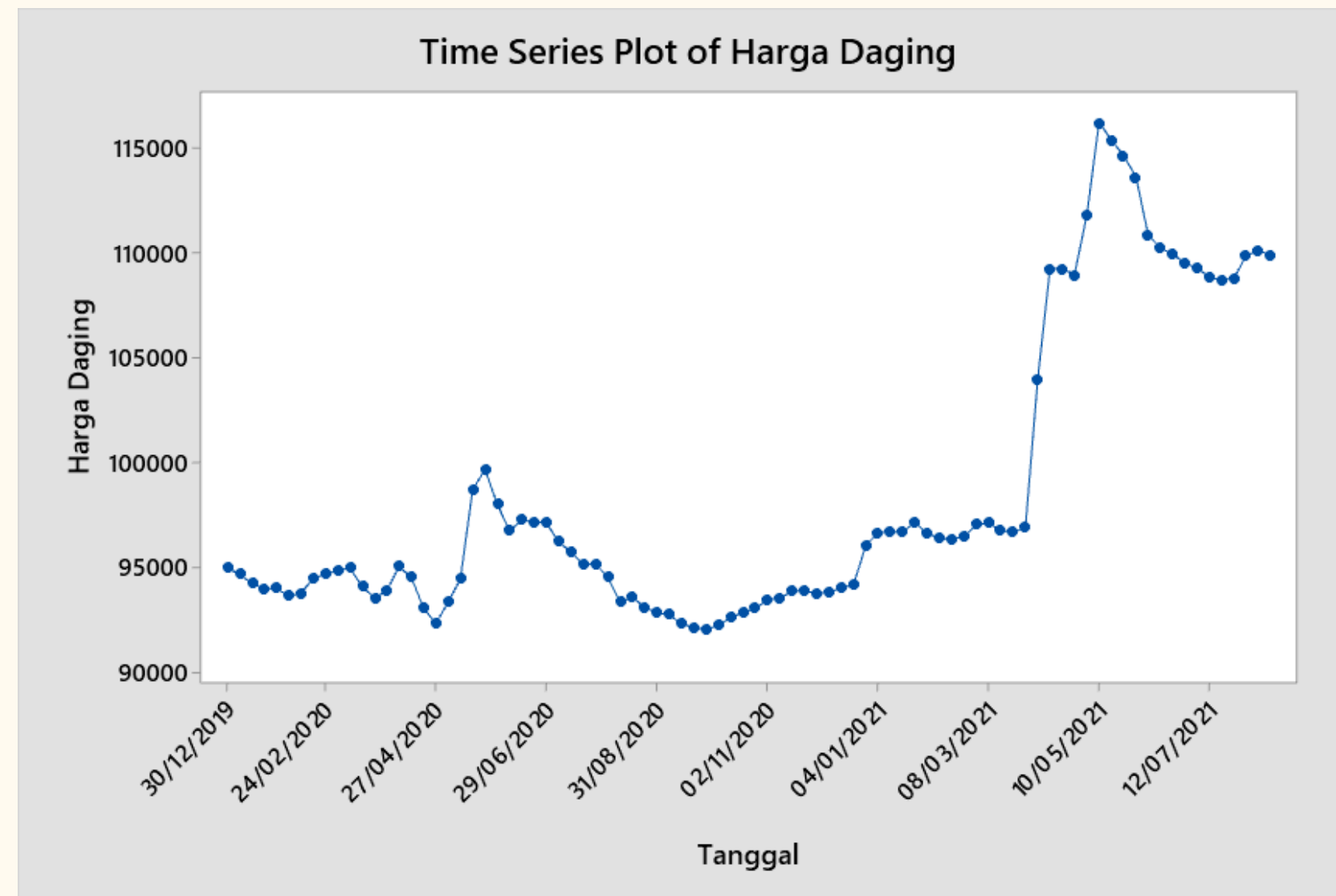
Analisis Data Statistik

# Analisis Time Series Harga Daging Jakarta

Latifatul Khumairoh (5003211032)

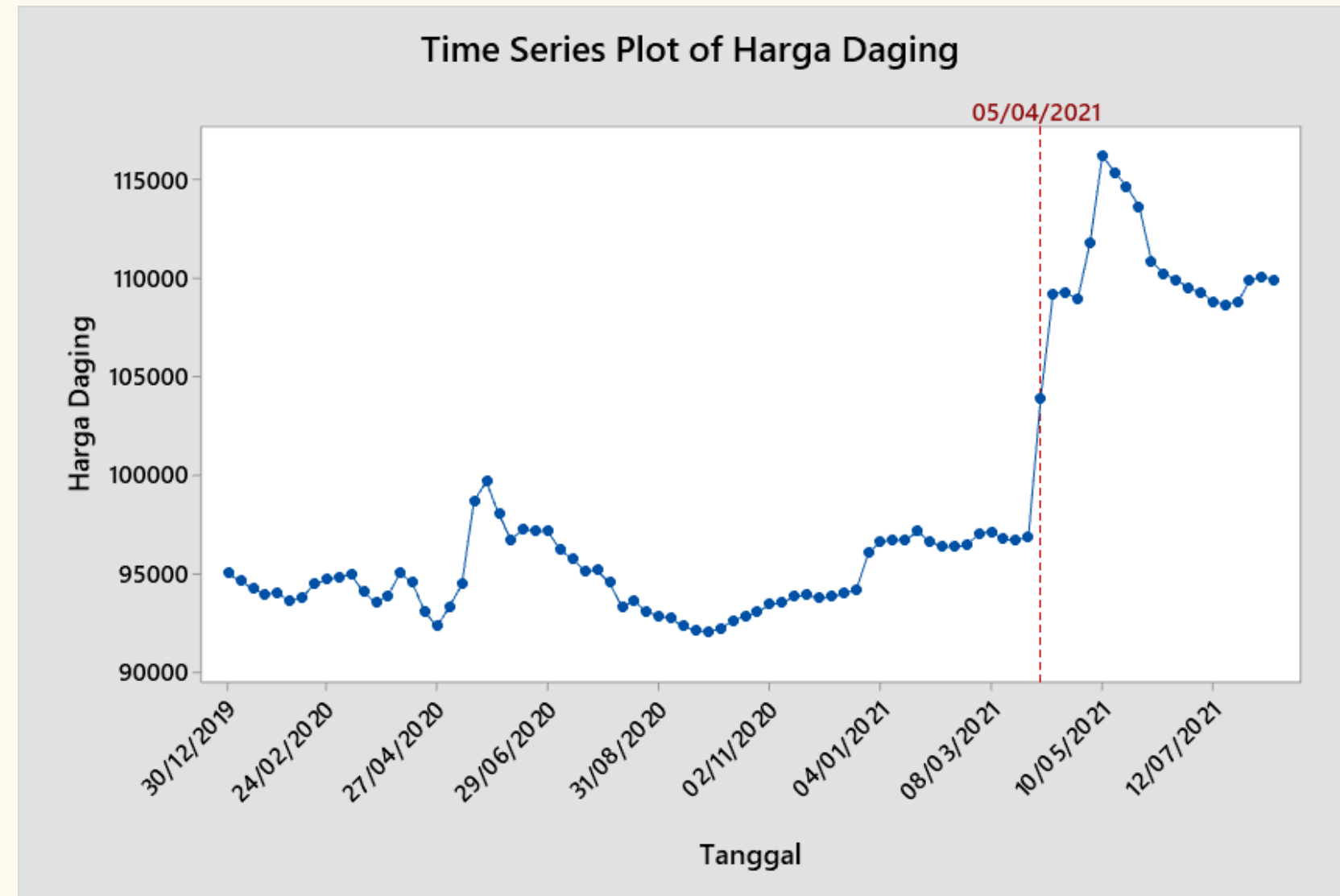


# Plot Time Series



Grafik ini menunjukkan pergerakan harga daging di Jakarta dari akhir Desember 2019 hingga pertengahan Agustus 2021. Pada sumbu Y ditampilkan harga daging dalam rupiah, sementara sumbu X menunjukkan tanggal observasi. Terlihat adanya lonjakan harga yang cukup drastis sejak **awal April 2021**, dengan peningkatan harga yang jauh lebih tinggi dibandingkan periode sebelumnya.

# Intervensi



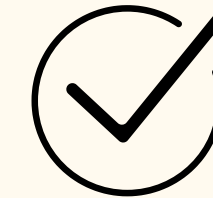
Kenaikan harga daging pada 5 April 2021 hingga Mei 2021 diduga disebabkan oleh meningkatnya permintaan sementara menjelang Ramadan dan Idulfitri, serta terus naiknya harga daging sapi dan sapi bakalan dari Australia. Dalam analisis time series, pola kenaikan harga ini dapat dikategorikan sebagai **Step Intervention**, di mana harga naik signifikan dan bertahan pada level yang lebih tinggi untuk beberapa waktu, mencerminkan perubahan permanen atau semi-permanen dalam level harga akibat faktor musiman dan pasar.

# Pembagian Data



## Data In Sample

Data sebelum terjadi intervensi,  
yaitu dari 30 Desember 2019  
hingga 29 Maret 2021,  
mencakup sebanyak 66 data.



## Data Out Sample

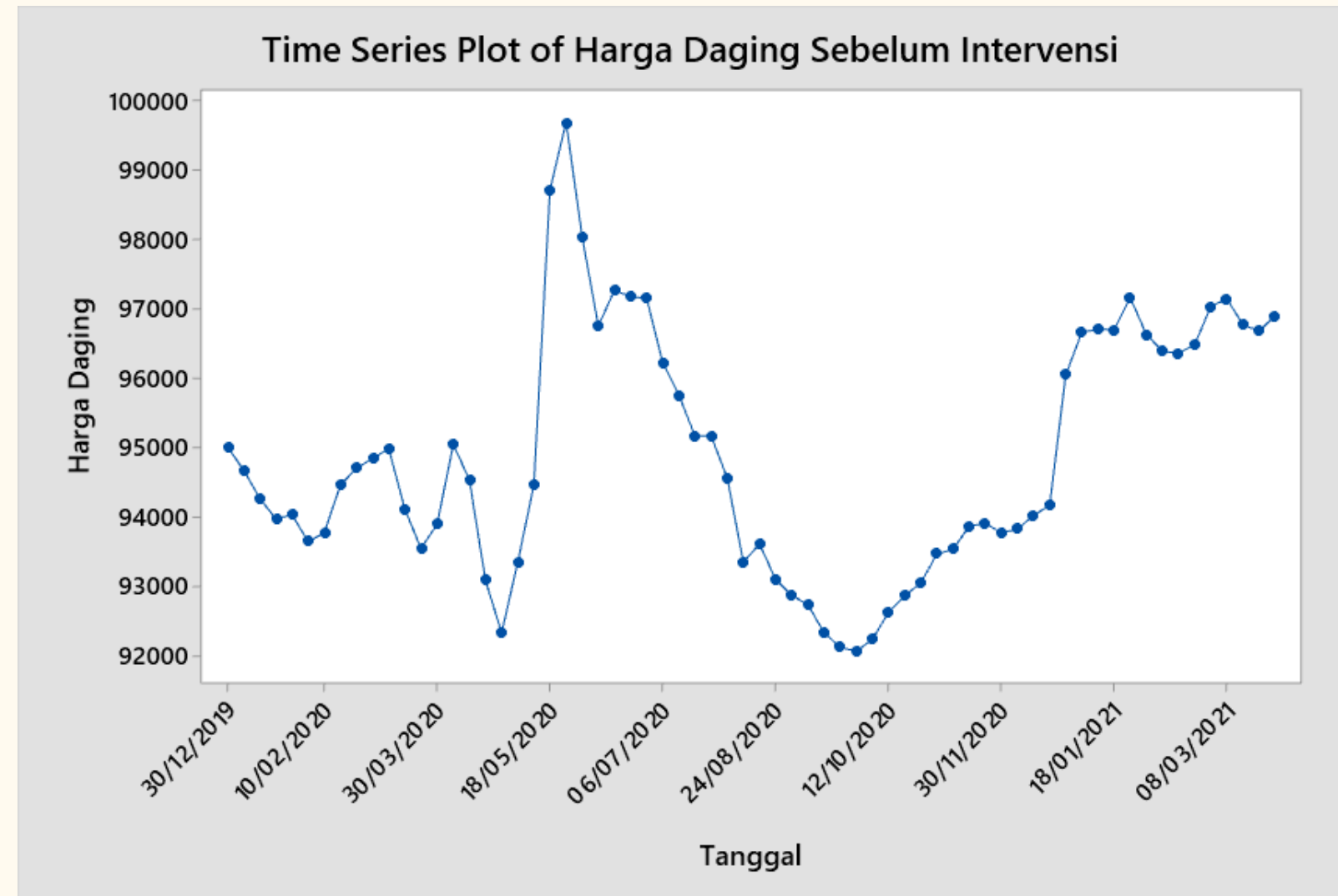
Data setelah terjadinya intervensi,  
yaitu dari tanggal 5 April 2021  
hingga 16 Agustus 2021,  
mencakup sebanyak 20 data.

Pembagian data ini memungkinkan analisis yang lebih baik terhadap efek intervensi,  
di mana data *in sample* digunakan untuk **model dasar** dan data *out sample* untuk  
**evaluasi model dan analisis perubahan yang terjadi setelah intervensi.**

# **Pemodelan ARIMA untuk Data Sebelum Intervensi**

**(In Sample)**

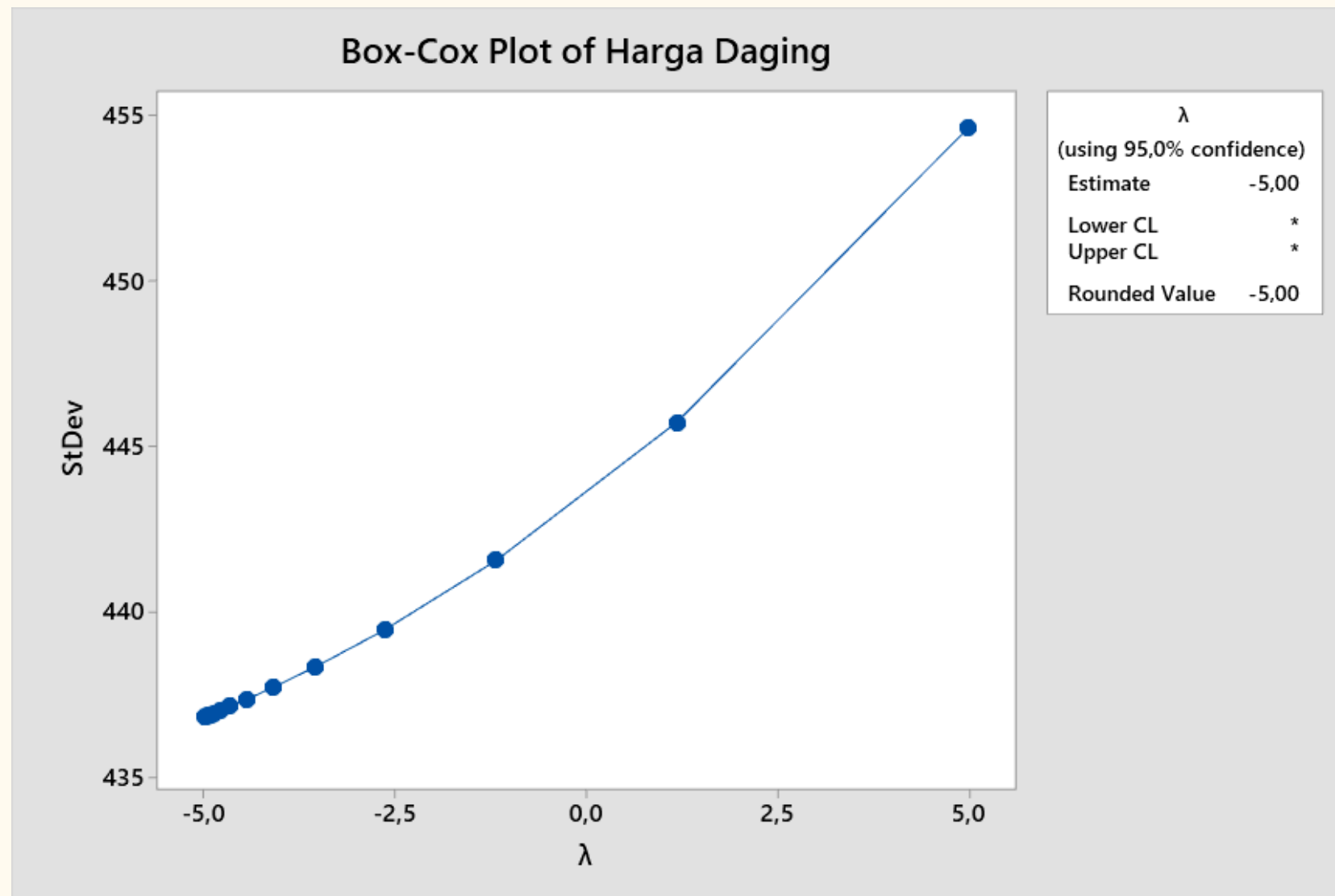
# Plot Time Series



Grafik menunjukkan pergerakan harga daging sebelum intervensi pada 5 April 2021, dengan lonjakan harga signifikan pada **Mei 2020** yang disebabkan oleh peningkatan permintaan menjelang Lebaran. Namun, lonjakan ini tidak dianggap sebagai intervensi karena **kenaikannya tidak terlalu tinggi dan bersifat sementara, dengan harga kembali turun setelahnya**. Jika dibandingkan dengan kondisi yang akan terjadi pada intervensi 5 April 2021, lonjakan Mei 2020 dianggap tidak cukup signifikan untuk mempengaruhi tren secara permanen.

# Cek Stasioneritas

## Stasioner dalam Varians



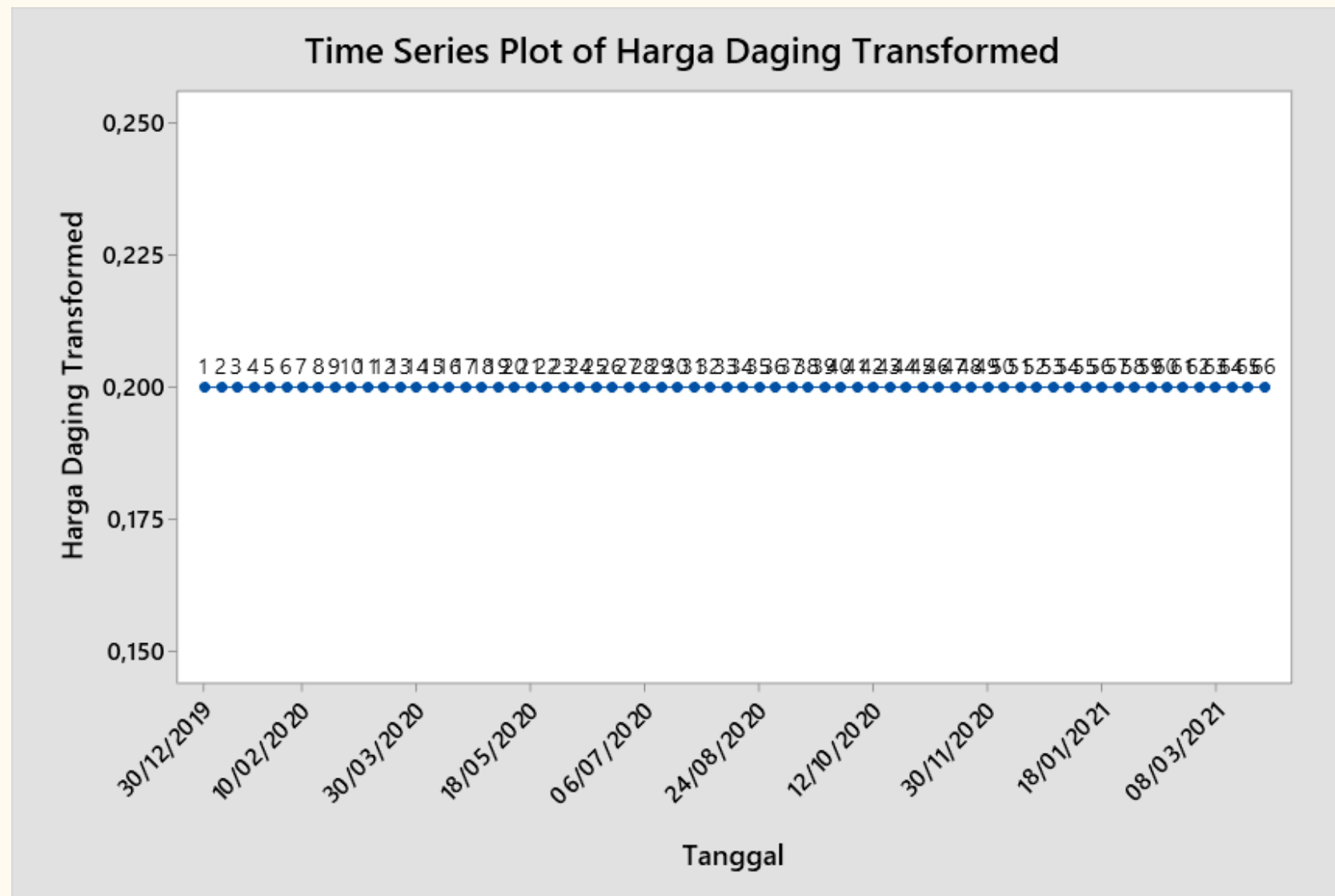
Pengujian stasioneritas terhadap varians dilakukan dengan grafik Box-Cox. Apabila hasil rounded value lebih dari sama dengan 1, maka data tersebut dapat dikatakan sudah stasioner. Namun, apabila syarat stasioner dalam varians belum terpenuhi, maka perlu dilakukan transformasi Box-Cox sesuai nilai lambda yang didapat menggunakan rumus:

$$w_t = \begin{cases} \log(y_t) & \text{if } \lambda = 0; \\ (y_t^\lambda - 1)/\lambda & \text{otherwise.} \end{cases}$$

Transformasi harus tetap memperhatikan pola data setelah diterapkan. Jika transformasi tidak mengubah pola data secara signifikan, maka data yang telah ditransformasi dapat digunakan untuk analisis selanjutnya. Namun, jika transformasi justru mengubah pola data, sebaiknya transformasi tidak perlu dilakukan.

# Cek Stasioneritas

## Stasioner dalam Varians



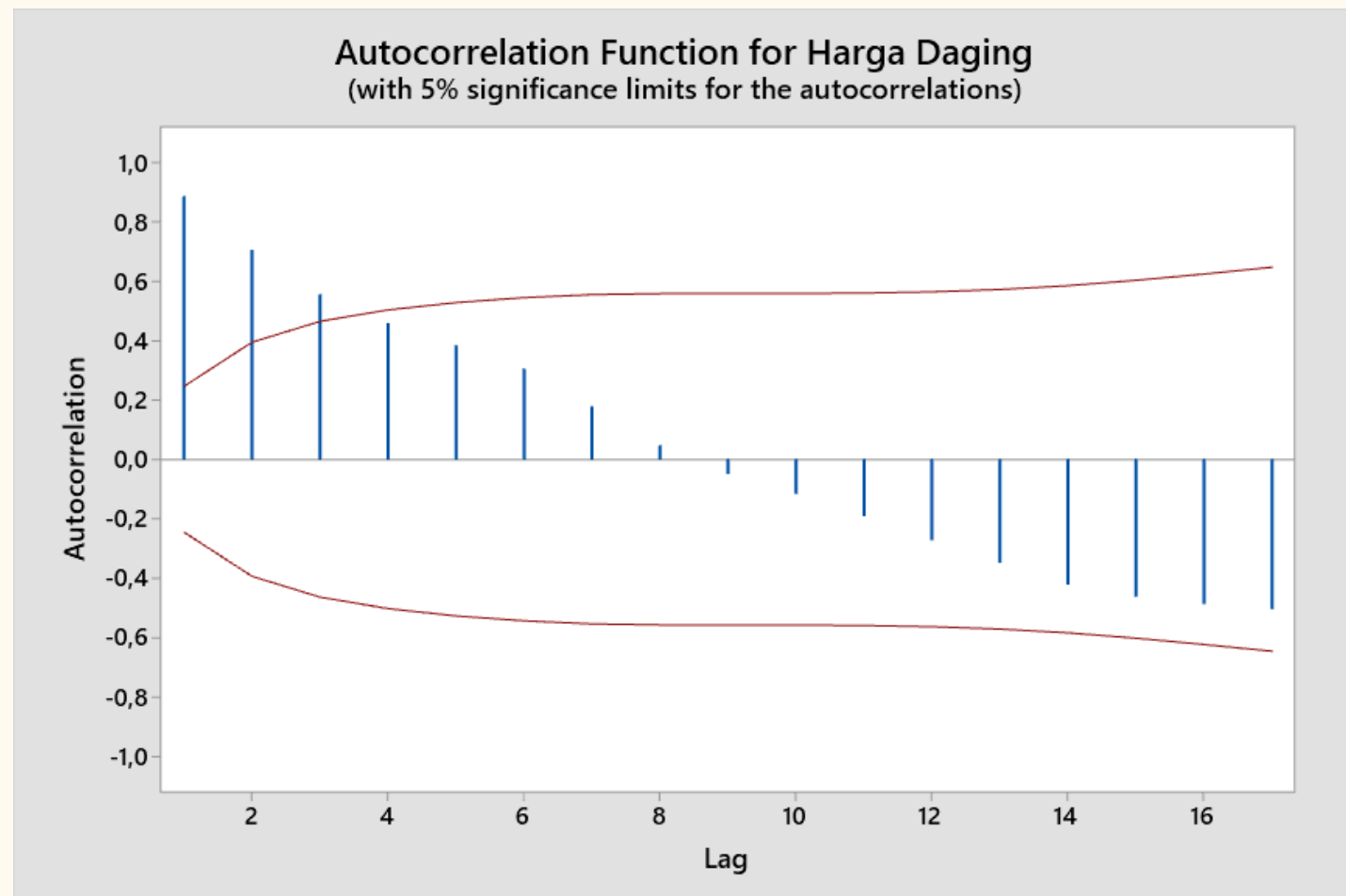
Setelah mensubtitusikan nilai lambda -5,00 ke dalam rumus transformasi Box-Cox, didapat data harga daging baru yang konstan bernilai 0,2 dengan plot time series seperti gambar di samping.

Karena transformasi ini secara signifikan mengubah pola data, peneliti memutuskan untuk **tidak melakukan transformasi** dan membiarkan **data tetap tidak stasioner dalam varians**.



# Cek Stasioneritas

## Stasioner dalam Mean



```
def cek_stationary(data):  
    result = adfuller(data)  
    print('p-value: {}'.format(result[1]))  
  
cek_stationary(train_data['Harga Daging'])
```

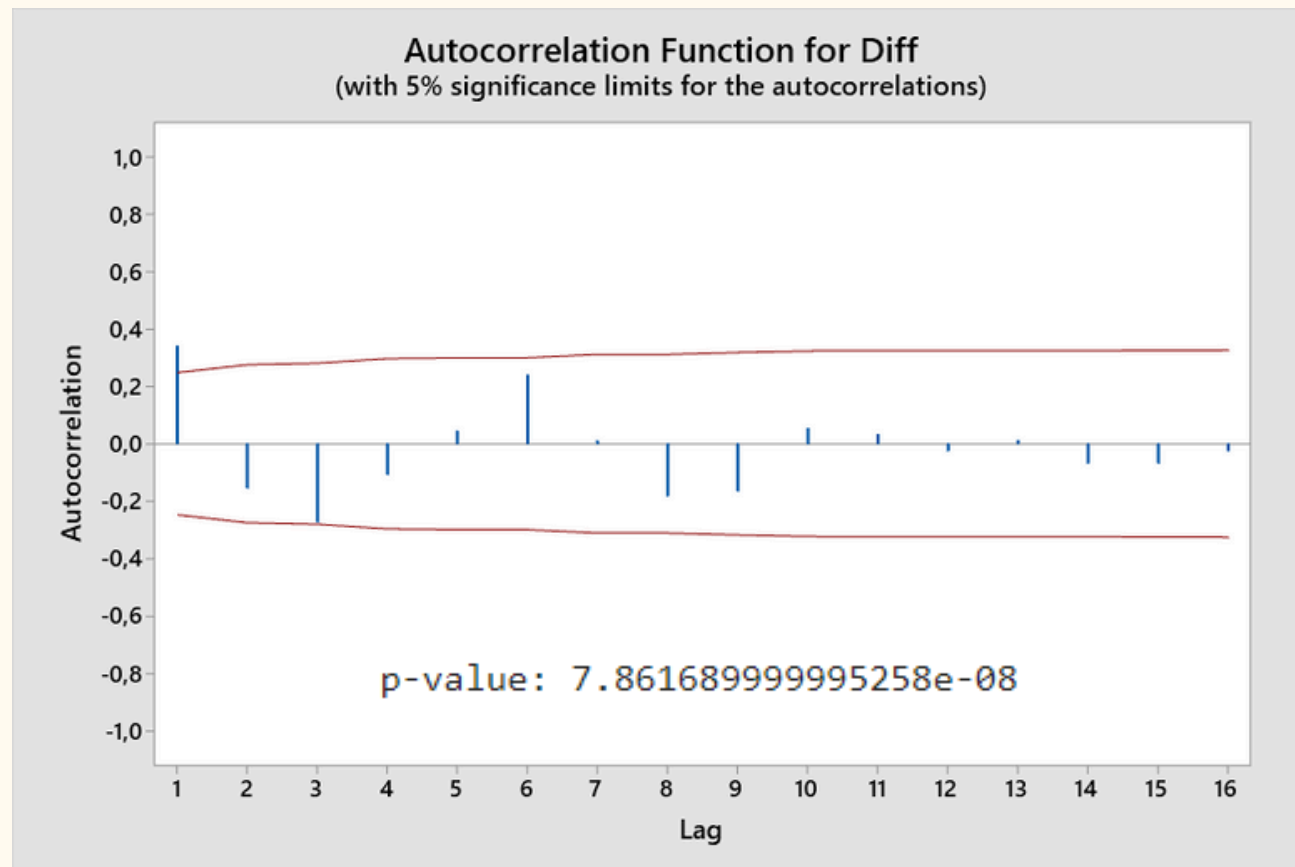
→ p-value: 0.34702035934679654

Pengujian stasioneritas terhadap mean dilakukan dengan melihat pola Autocorrelation Function (ACF) dari data. Data yang sudah stasioner dalam rata-rata ditunjukkan dengan ACF yang tidak dies down atau menurun secara lambat. Dapat juga diperiksa menggunakan Augmented Dickey-Fuller Test dengan kriteria stasioner dalam mean apabila p-value kurang dari 0.05.

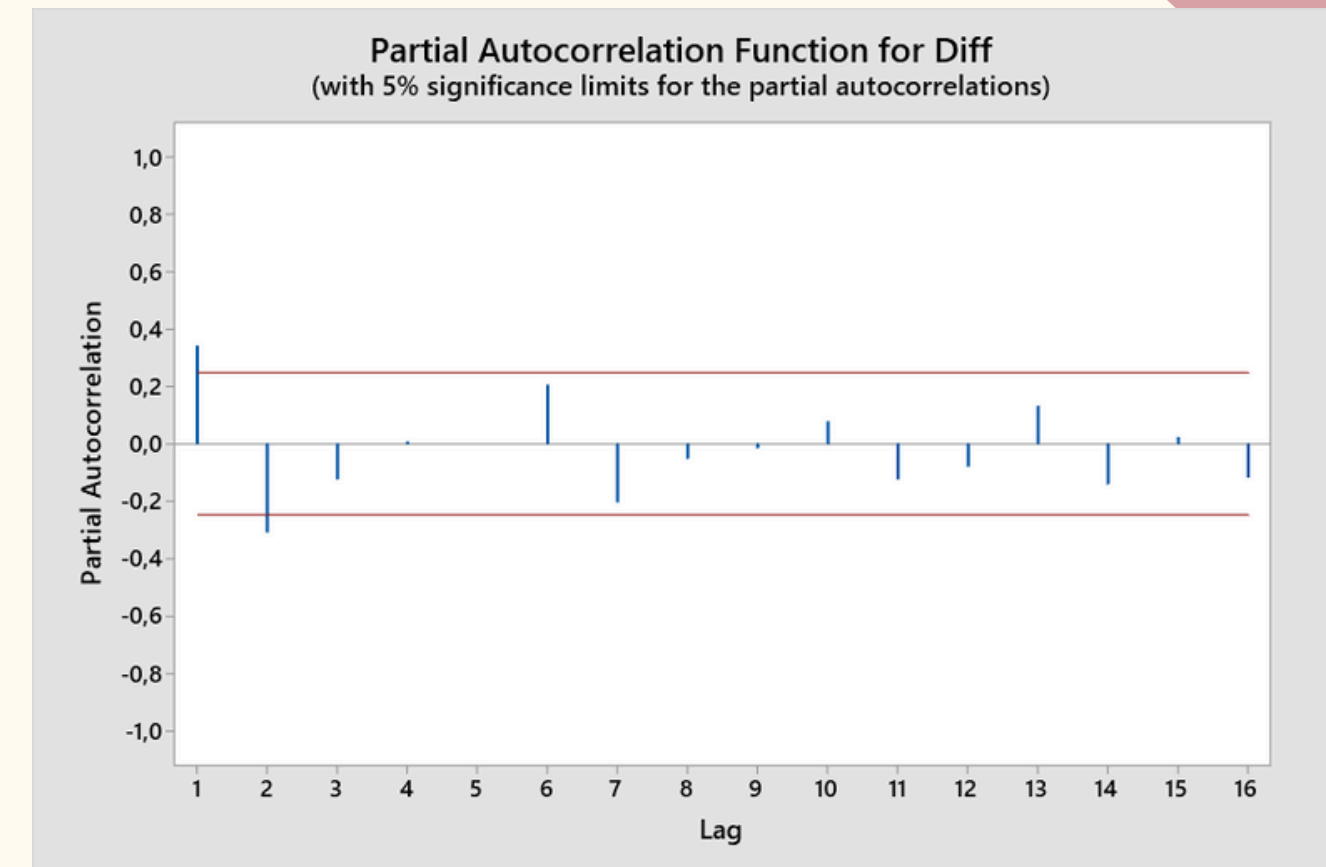
Dapat dilihat bahwa plot ACF menunjukkan pola yang dies down dan nilai p-value lebih dari 0,05 sehingga **data belum stasioner dalam mean**. Untuk itu perlu dilakukan differencing pada lag 1.

# Cek Stasioneritas

## Stasioner dalam Mean



Data sudah stasioner dalam mean karena pola ACF sudah tidak dies down dan p-value kurang dari 0,05. Grafik ACF ini digunakan untuk menentukan orde untuk moving average (MA). Terlihat bahwa lag yang signifikan adalah lag 1, sehingga MA dengan orde 1 dapat dijadikan acuan untuk model dugaan. Untuk mendapatkan model dugaan lainnya, perlu dilihat plot PACF dari data.



Grafik PACF di atas menunjukkan lag yang signifikan hingga lag 2. Grafik PACF digunakan sebagai acuan untuk penentuan orde autoregressive (AR), sehingga AR hingga orde 2 dapat dijadikan acuan untuk model dugaan. Dengan informasi-informasi yang ada, maka model-model dugaannya adalah **ARIMA(1,1,0)**, **ARIMA(2,1,0)**, **ARIMA(0,1,1)**, **ARIMA(1,1,1)**, dan **ARIMA(2,1,1)**. Dengan data tidak menunjukkan adanya pola musiman.

# Estimasi Parameter

## ARIMA(1,1,0)

### Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0,344	0,118	2,90	0,005
Constant	18,2	95,1	0,19	0,849

Differencing: 1 regular difference

## ARIMA(0,1,1)

### Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
MA 1	-0,418	0,114	-3,66	0,001
Constant	29	132	0,22	0,829

Differencing: 1 regular difference

## ARIMA(2,1,1)

### Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0,714	0,328	2,17	0,034
AR 2	-0,407	0,140	-2,91	0,005
MA 1	0,294	0,356	0,83	0,412
Constant	23,4	64,4	0,36	0,717

Differencing: 1 regular difference

## ARIMA(2,1,0)

### Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0,449	0,121	3,72	0,000
AR 2	-0,311	0,121	-2,57	0,013
Constant	27,3	91,1	0,30	0,765

Differencing: 1 regular difference

## ARIMA(1,1,1)

### Final Estimates of Parameters

Type	Coef	SE Coef	T-Value	P-Value
AR 1	0,054	0,296	0,18	0,855
MA 1	-0,382	0,274	-1,40	0,168
Constant	27	129	0,21	0,835

Differencing: 1 regular difference

Model ARIMA(1,1,1) dan ARIMA(2,1,1) tidak dapat digunakan untuk prediksi dikarenakan koefisiennya memiliki p-value lebih dari 0,05 sehingga koefisien tersebut tidak signifikan pada model.

Dugaan model yang digunakan untuk analisis selanjutnya adalah **model ARIMA(1,1,0), ARIMA(0,1,1), dan ARIMA(2,1,0).**

# Uji Asumsi Residual

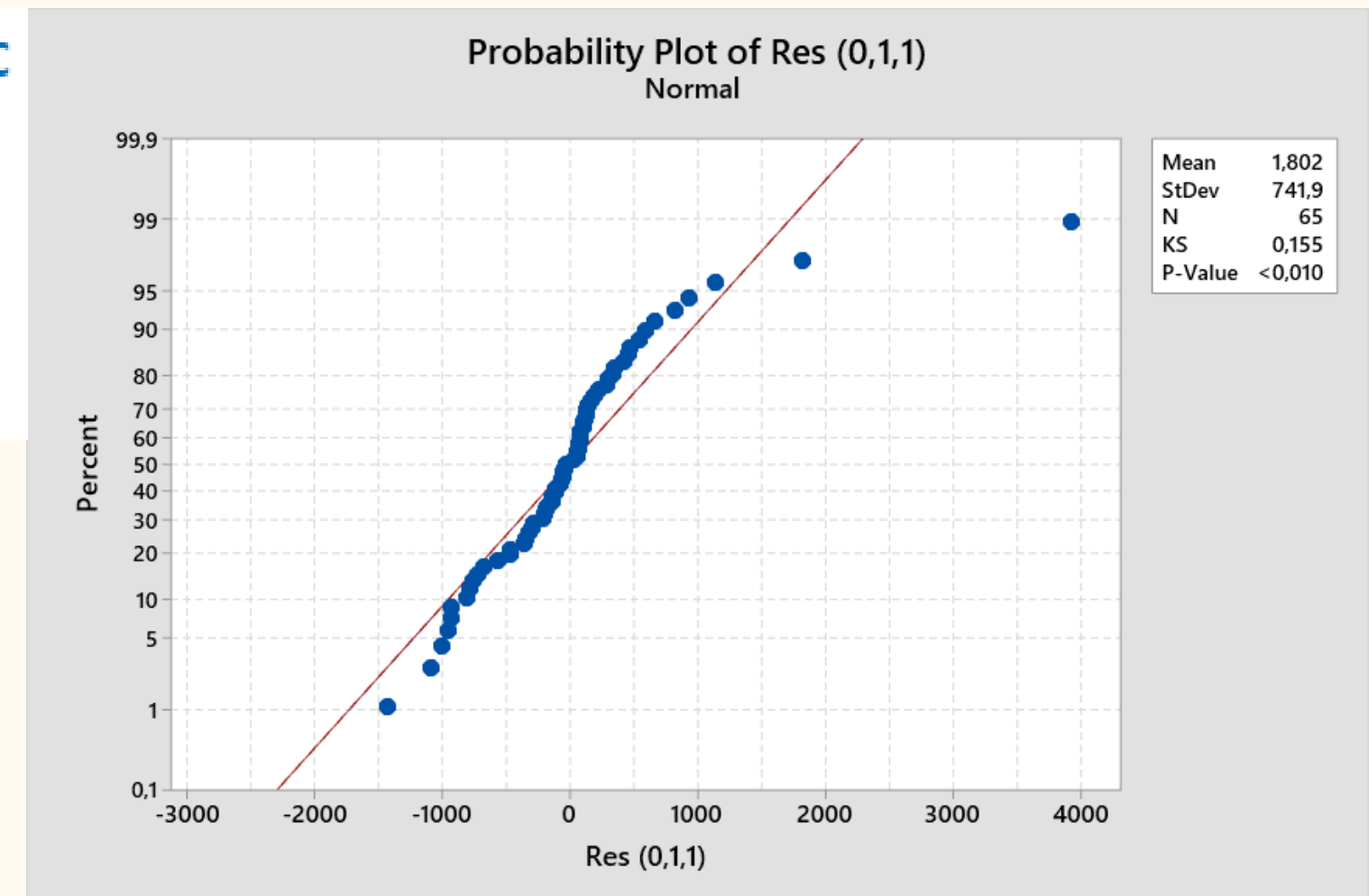
## White Noise dan Berdistribusi Normal

ARIMA(0,1,1)

### Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	15,46	18,21	30,29	33,87
DF	10	22	34	46
P-Value	0,116	0,693	0,650	0,907

- Ljung-Box Test menunjukkan bahwa bahwa p-value di semua lag sudah lebih besar dari 0,05, maka diputuskan gagal tolak  $H_0$  dan **residual memenuhi asumsi white noise**.
- Kolmogorov-Smirnov Test menunjukkan bahwa p-value bernilai kurang dari 0,05 sehingga tolak  $H_0$  dan **residual belum memenuhi asumsi distribusi normal**.



# Uji Asumsi Residual

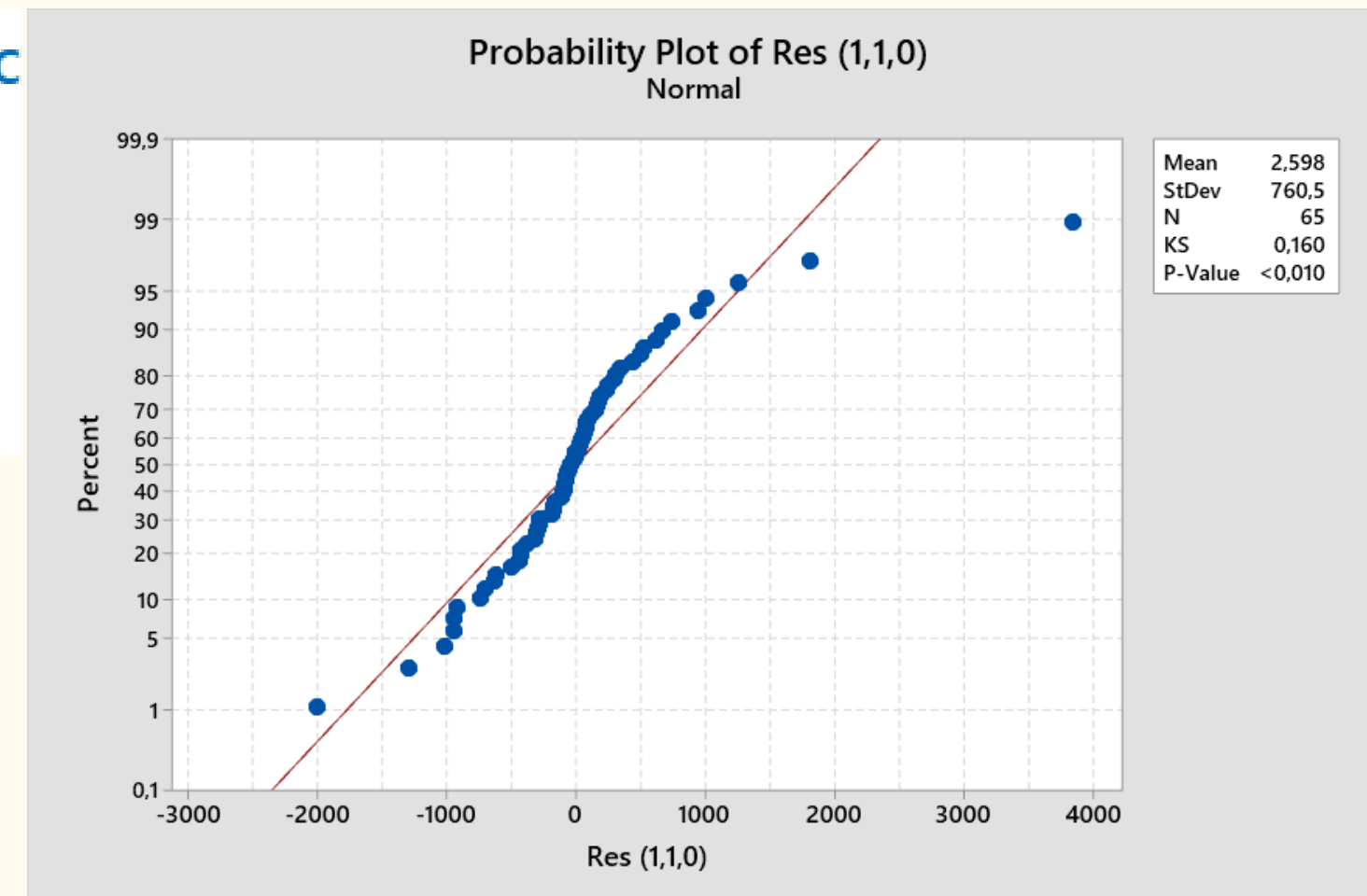
## White Noise dan Berdistribusi Normal

### ARIMA(1,1,0)

#### Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	20,39	23,26	33,75	37,04
DF	10	22	34	46
P-Value	0,026	0,387	0,480	0,824

- Ljung-Box Test menunjukkan bahwa p-value di lag 12 bernilai kurang dari 0,05, maka diputuskan tolak  $H_0$  dan **residual belum memenuhi asumsi white noise.**
- Kolmogorov-Smirnov Test menunjukkan bahwa p-value bernilai kurang dari 0,05 sehingga tolak  $H_0$  dan **residual belum memenuhi asumsi distribusi normal.**





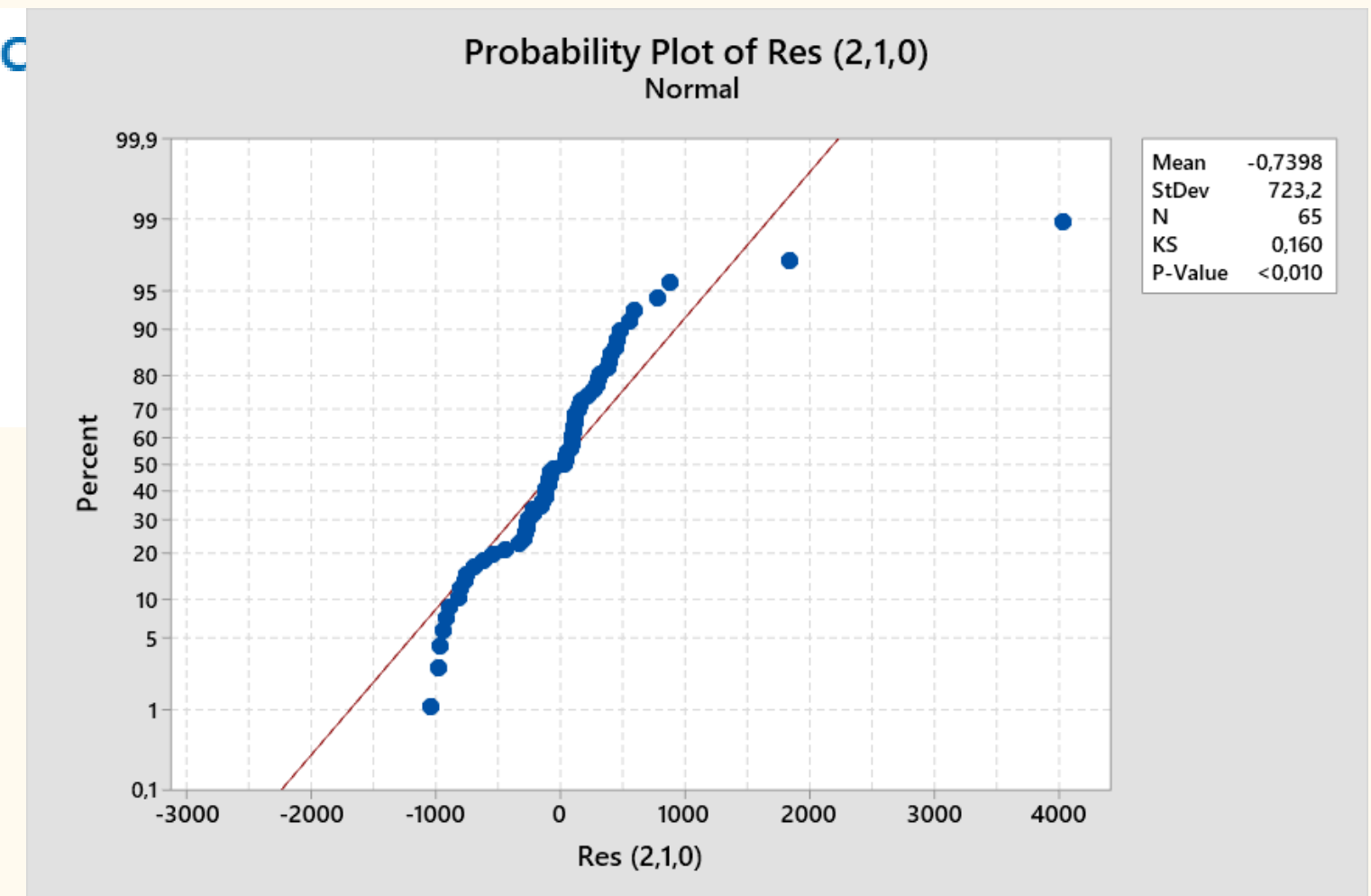
# Uji Asumsi Residual White Noise dan Berdistribusi Normal

## ARIMA(2,1,0)

### Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	10,50	14,55	28,33	32,27
DF	9	21	33	45
P-Value	0,311	0,845	0,699	0,923

- Ljung-Box Test menunjukkan bahwa p-value di semua lag sudah lebih besar dari 0,05, maka diputuskan gagal tolak  $H_0$  dan **residual memenuhi asumsi white noise**.
- Kolmogorov-Smirnov Test menunjukkan bahwa p-value bernilai kurang dari 0,05 sehingga tolak  $H_0$  dan **residual belum memenuhi asumsi distribusi normal**.



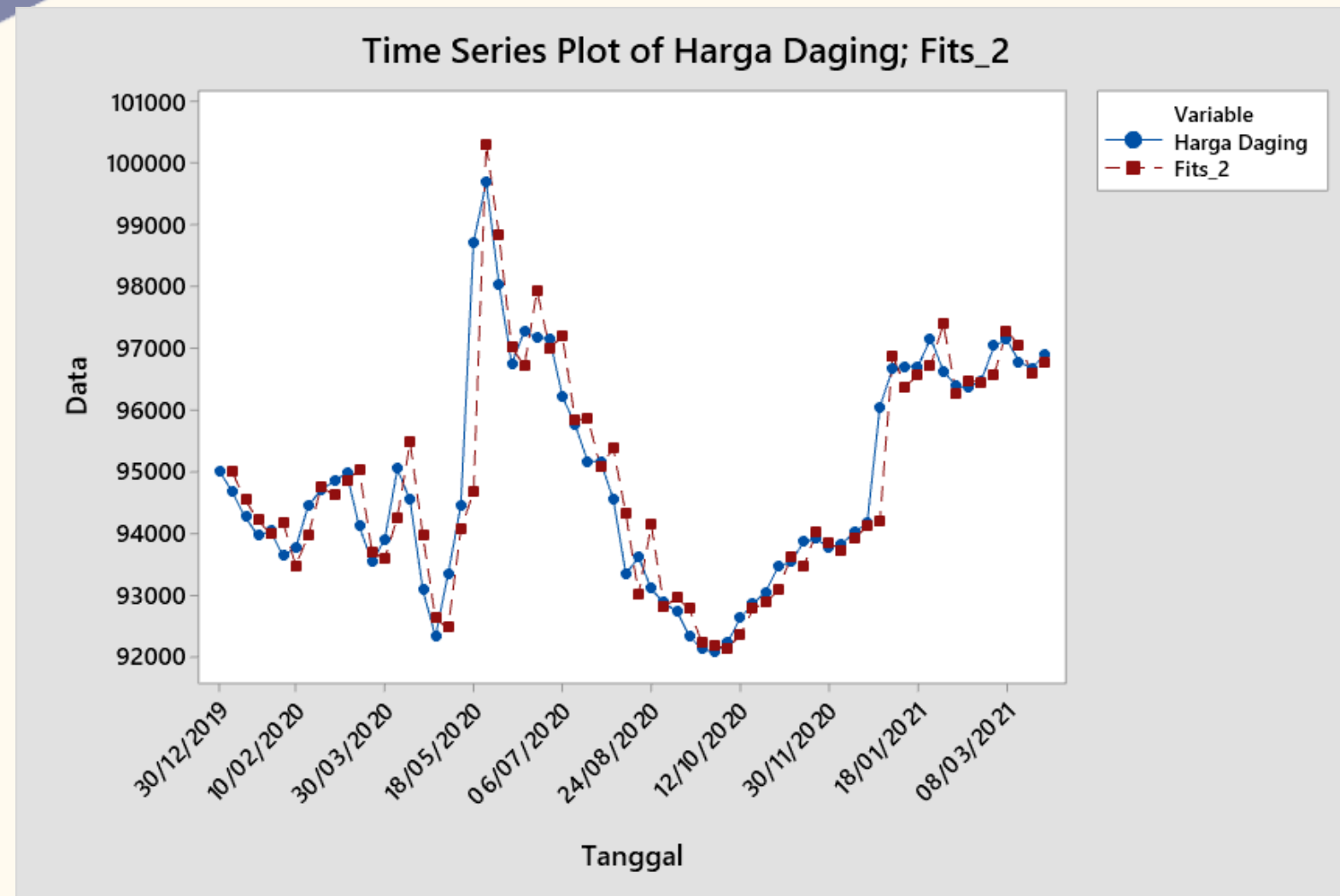
# Penentuan Model Terbaik

Berdasarkan uji asumsi pada slide sebelumnya, didapatkan hasil bahwa ketiga model tidak memenuhi asumsi residual berdistribusi normal, tetapi terdapat 2 model yang memenuhi asumsi residual white noise, yaitu model ARIMA(0,1,1) dan ARIMA(2,1,0). Selanjutnya dilakukan perhitungan MAPE (Mean Average Percentage Error) dengan Excel pada kedua model untuk menentukan model terbaik dengan melihat MAPE terkecil.

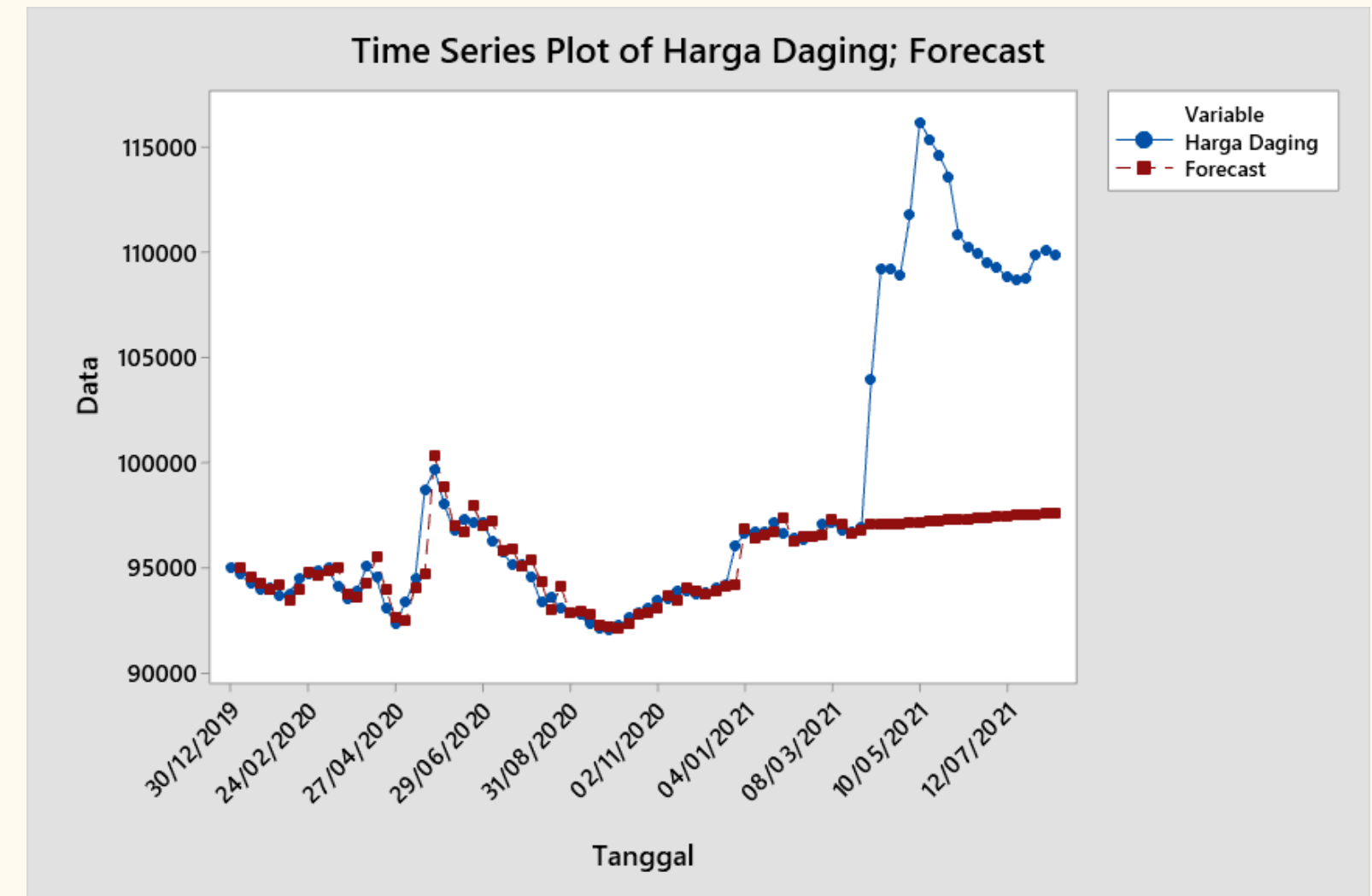
MODEL ARIMA	MAPE
ARIMA(0,1,1)	0,478%
ARIMA(2,1,0)	0,461%

Dipilih model ARIMA(2,1,0) karena memiliki MAPE terkecil. Residual yang tidak berdistribusi normal pada model ini dapat ditoleransi berdasarkan teorema limit pusat yaitu suatu sebaran dapat didekati dengan sebaran normal ketika ukuran contohnya besar, dalam kasus ini  $n = 66$ . Dengan begitu, model ARIMA(2,1,0) telah memenuhi asumsi residual white noise dan berdistribusi normal sehingga model ini dapat digunakan untuk memprediksi data setelah intervensi.

# Plot Actual vs Fits ARIMA (2,1,0)



Berikut adalah plot time series yang membandingkan harga daging sebenarnya dengan harga daging berdasarkan model ARIMA (2,1,0) sampai **sebelum intervensi**. Terlihat bahwa model ini mampu memberikan prediksi data yang dekat dengan data aktual, dibuktikan dengan nilai MAPE yang kecil yaitu 0,461%.



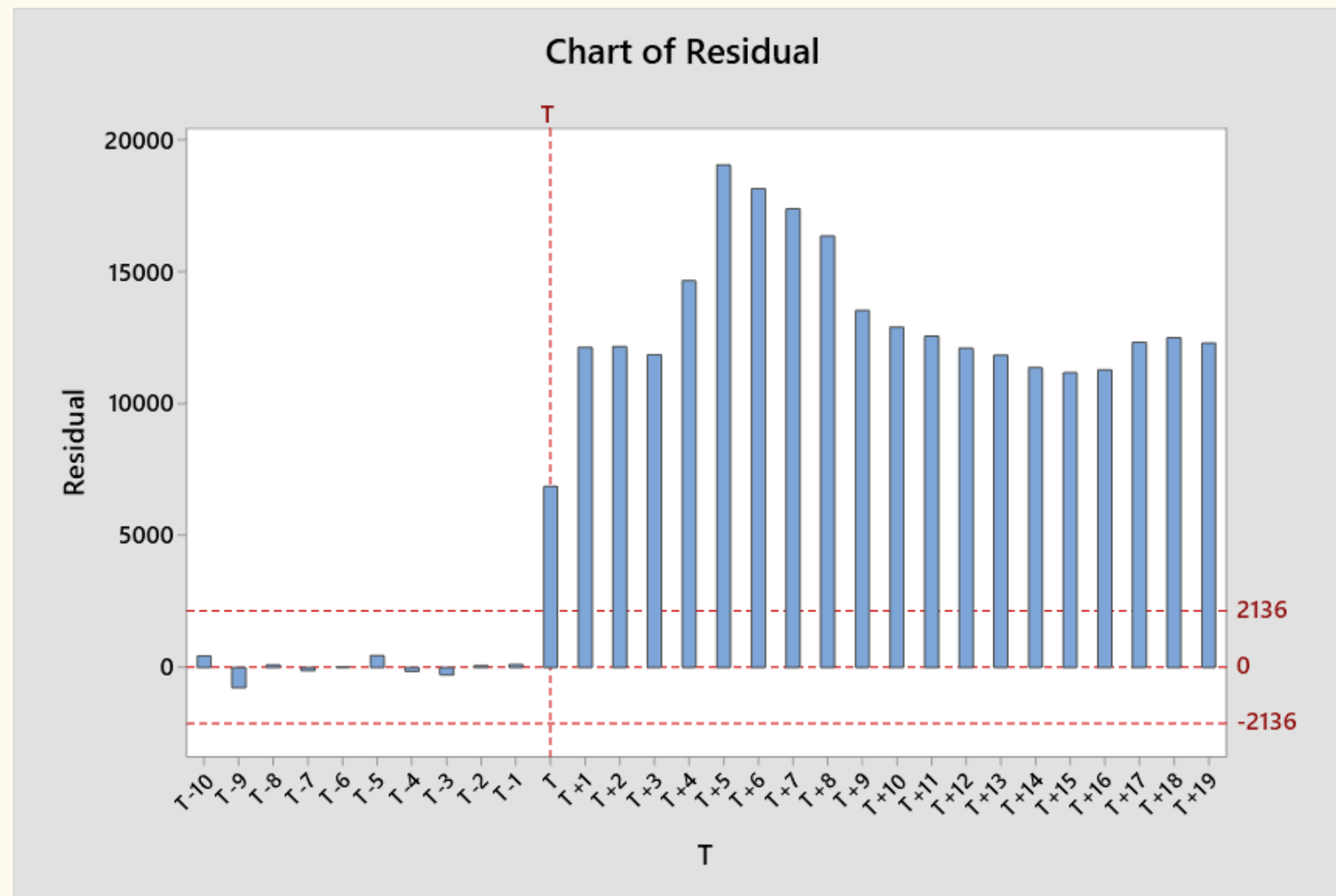
Tetapi model ini dirasa belum cukup sesuai untuk memprediksi data setelah intervensi dikarenakan rentang residual yang begitu jauh. Sehingga dengan begitu perlu dilakukan **analisis lebih lanjut untuk model ARIMA setelah intervensi**.



# **Pemodelan ARIMA untuk Data Setelah Intervensi**

**(Out Sample)**

# Plot Residual Respon Intervensi



Kemudian dibuat plot residual respon dari model ARIMA(2,1,0) dengan mengambil 10 data sebelum dan sesudah intervensi. Dengan menggunakan Excel, diperoleh nilai RMSE sebesar 712, sehingga batas  $\pm 3RMSE$  adalah  $\pm 2136$  yang menjadi batas bawah dan batas atas dari plot residual respon.

Plot residual respon memberikan informasi mengenai nilai  $b$ ,  $r$ ,  $s$  efek dari intervensi.

- Efek intervensi terjadi pada waktu ke- $T$ , atau pada tanggal 5 April 2021. Artinya, efek intervensi mulai terasa pada waktu tersebut, sehingga waktu tunda menjadi nol, yaitu  $b=0$
- Plot residual respon yang keluar dari garis signifikansi menunjukkan lamanya pengaruh dari intervensi yang terjadi hingga mulai normal kembali, sehingga diperoleh nilai  $s=5$ .
- Plot residual respon intervensi tidak membentuk pola tertentu, sehingga kemungkinan orde  $r=0$ .

Berdasarkan analisis tersebut, diperoleh model intervensi dengan dugaan  $b=0$ ,  $r=0$ , dan  $s=5$ . Selanjutnya, nilai orde  $b$ ,  $r$ , dan  $s$  dapat diperoleh melalui proses coba-coba untuk menghasilkan model terbaik.

# Estimasi Parameter ARIMA Setelah Intervensi

	Pr(> z )
ar1	1.612e-07
ar2	0.004261
intervensir0s1-MA0	0.036222

	Pr(> z )
ar1	1.079e-07
ar2	0.004423
intervensir0s2-MA0	0.957728

	Pr(> z )
ar1	1.078e-07
ar2	0.004385
intervensir0s3-MA0	0.866552

	Pr(> z )
ar1	1.074e-07
ar2	0.004399
intervensir0s4-MA0	0.935430

	Pr(> z )
ar1	1.083e-07
ar2	0.004409
intervensir0s5-MA0	0.998255

Didapatkan hasil bahwa model ARIMA intervensi yang seluruh parameternya signifikan ( $p\text{-value} > 0,05$ ) adalah model **ARIMA(2,1,0)** dengan ordo intervensi **b=0**, **r=0**, dan **s=1**

# Uji Asumsi Residual

## White Noise dan Berdistribusi Normal

### Box-Ljung test

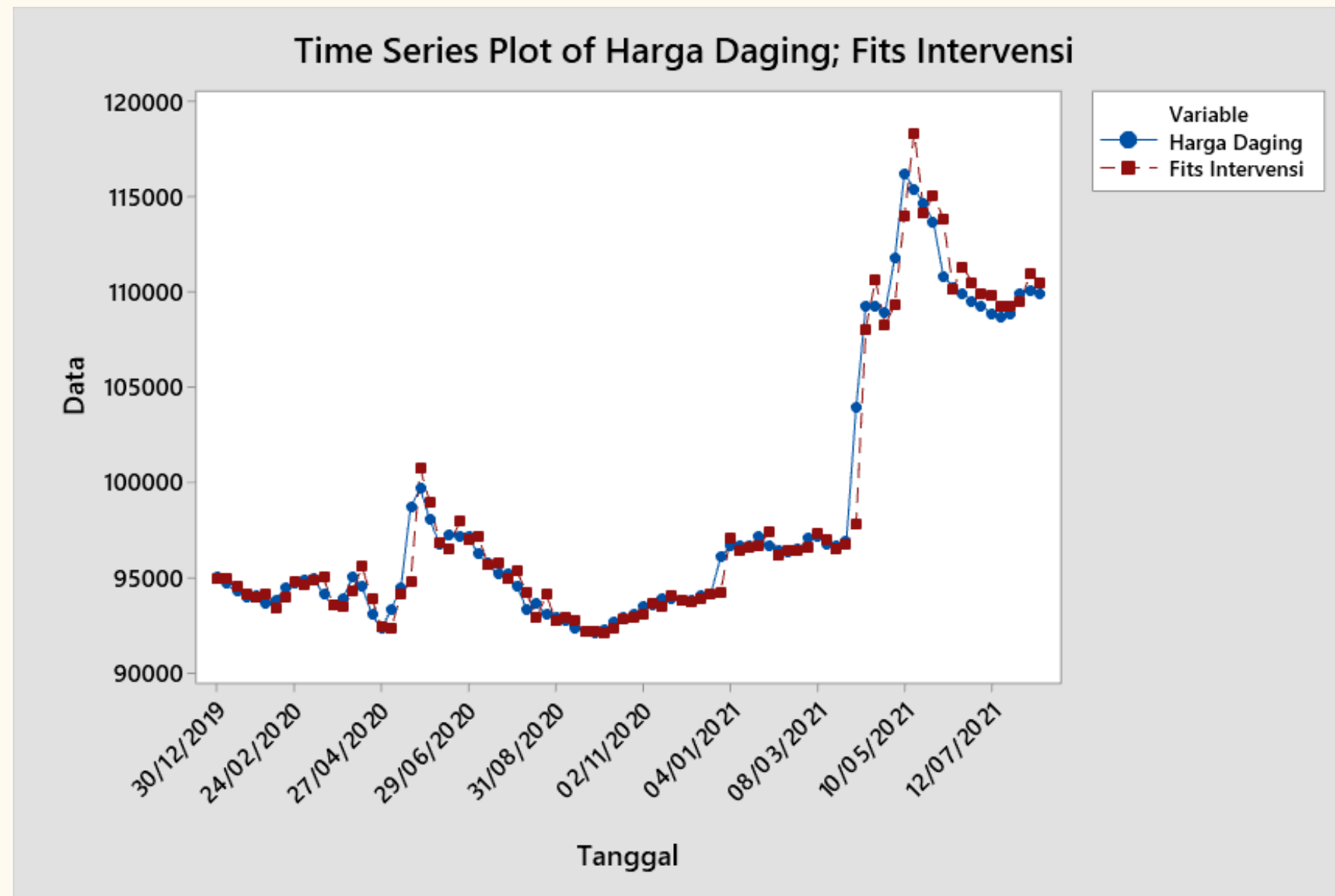
```
data: arimax_intervensi0s1$residuals  
X-squared = 13.943, df = 16, p-value = 0.603
```

### Exact one-sample Kolmogorov-Smirnov test

```
data: arimax_intervensi0s1$residuals  
D = 0.20327, p-value = 0.001362  
alternative hypothesis: two-sided
```

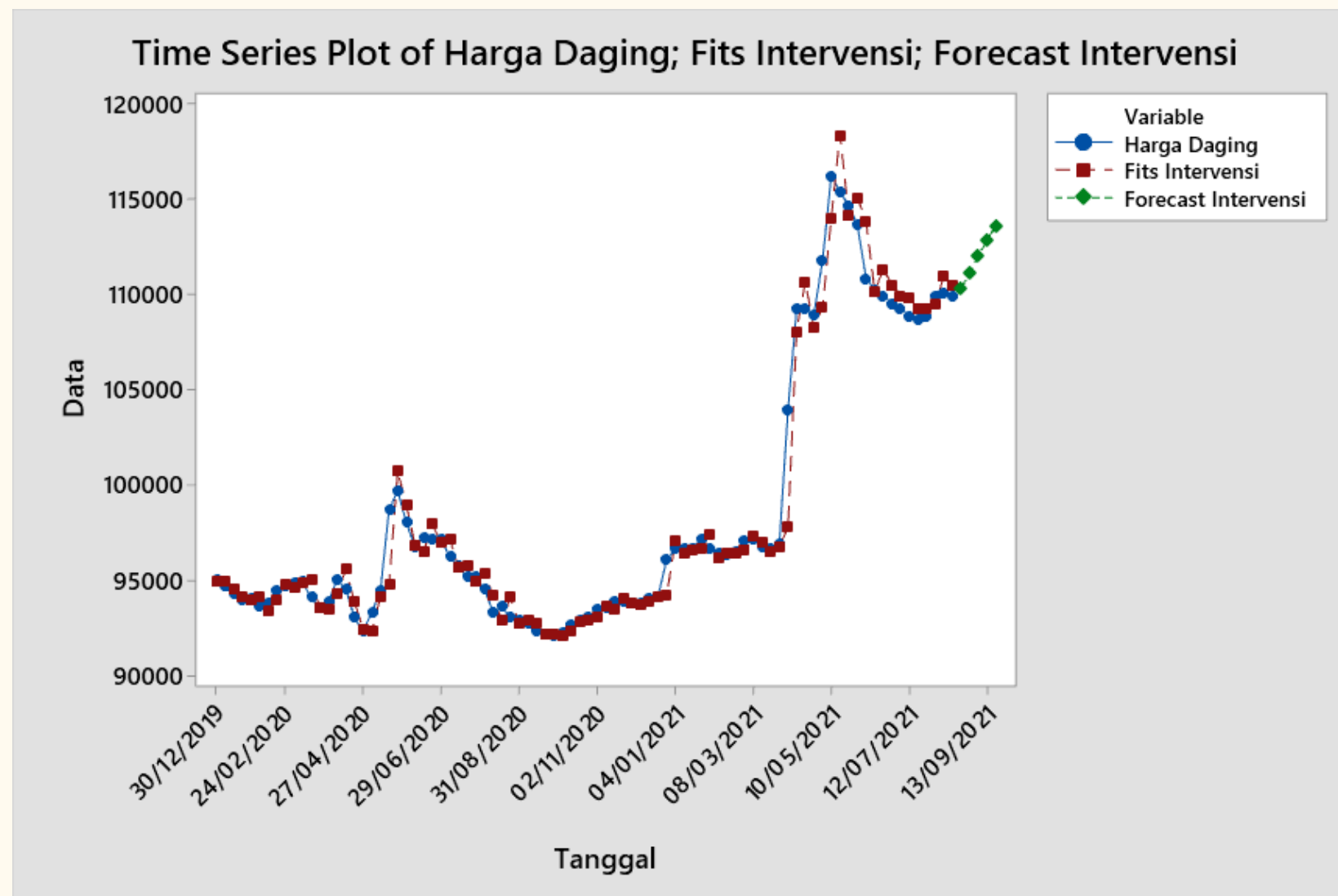
- Ljung-Box Test menunjukkan bahwa p-value sudah lebih besar dari 0,05, maka diputuskan gagal tolak  $H_0$  dan **residual memenuhi asumsi white noise**.
- Kolmogorov-Smirnov Test menunjukkan bahwa p-value bernilai kurang dari 0,05 sehingga tolak  $H_0$  dan residual belum memenuhi asumsi distribusi normal. Hal ini dapat ditoleransi berdasarkan teorema limit pusat yaitu suatu sebaran dapat didekati dengan sebaran normal ketika ukuran contohnya besar, dalam kasus ini  $n = 66$ . Sehingga dapat dikatakan bahwa **residual memenuhi asumsi berdistribusi normal**.

# Plot Actual vs Fits ARIMA Intervensi



Setelah melakukan analisis data menggunakan RStudio, fitted value dari model ARIMA intervensi kemudian dipindahkan ke Minitab untuk dibuat plot timeseries dan dibandingkan dengan data asli. Terlihat bahwa model mampu menangkap pola data dengan cukup baik, dibuktikan dengan nilai evaluasi residual **MAPE** sebesar **0,663%**

# Forecast Harga Daging



Tanggal	Forecast
23/08/2021	110327
30/08/2021	111159
06/09/2021	112055
13/09/2021	112859
20/09/2021	113595

- Model ARIMA intervensi mampu menangkap pola data dengan baik karena secara konsisten mengikuti data aktual sebelum dan sesudah intervensi. Model ini mampu menangkap lonjakan harga selama periode intervensi dan memberikan proyeksi yang masuk akal untuk periode berikutnya.
- Proyeksi untuk 5 langkah ke depan dipilih untuk memprediksi harga daging. Pendekatan ini penulis anggap cukup efektif untuk meminimalkan potensi penyimpangan dari pola sebenarnya.



**Terima kasih!**