```
!pip install yfinance statsmodels --quiet
```

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

import yfinance as yf

from statsmodels.tsa.stattools import adfuller
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from statsmodels.tsa.arima.model import ARIMA

import warnings
warnings.filterwarnings('ignore')
```

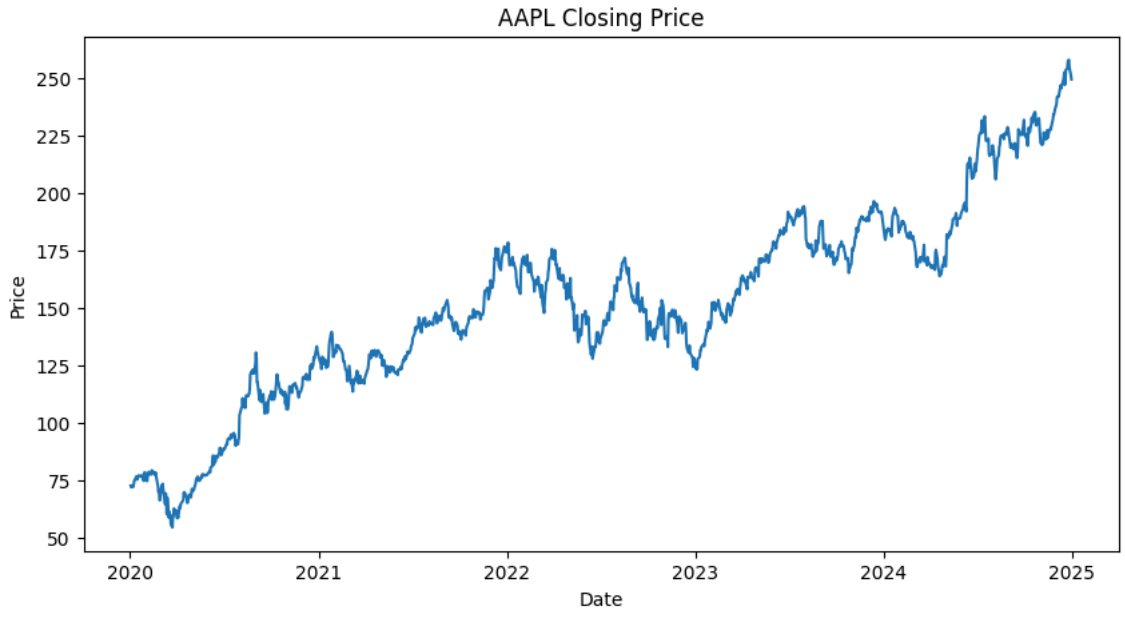
```
ticker = 'AAPL'  # Change to INFY.NS, RELIANCE.NS if needed
data = yf.download(ticker, start='2020-01-01', end='2025-01-01')

df = data[['Close']]
df.head()
```

```
[**********100%**********]  1 of 1 completed
```

Price	Close
Ticker	AAPL
Date	
2020-01-02	72.468254
2020-01-03	71.763725
2020-01-06	72.335564
2020-01-07	71.995354
2020-01-08	73.153488

Next steps: Generate code with df | New interactive sheet

```
plt.figure(figsize=(10,5))
plt.plot(df['Close'])
plt.title(f'{ticker} Closing Price')
plt.xlabel('Date')
plt.ylabel('Price')
plt.show()
```

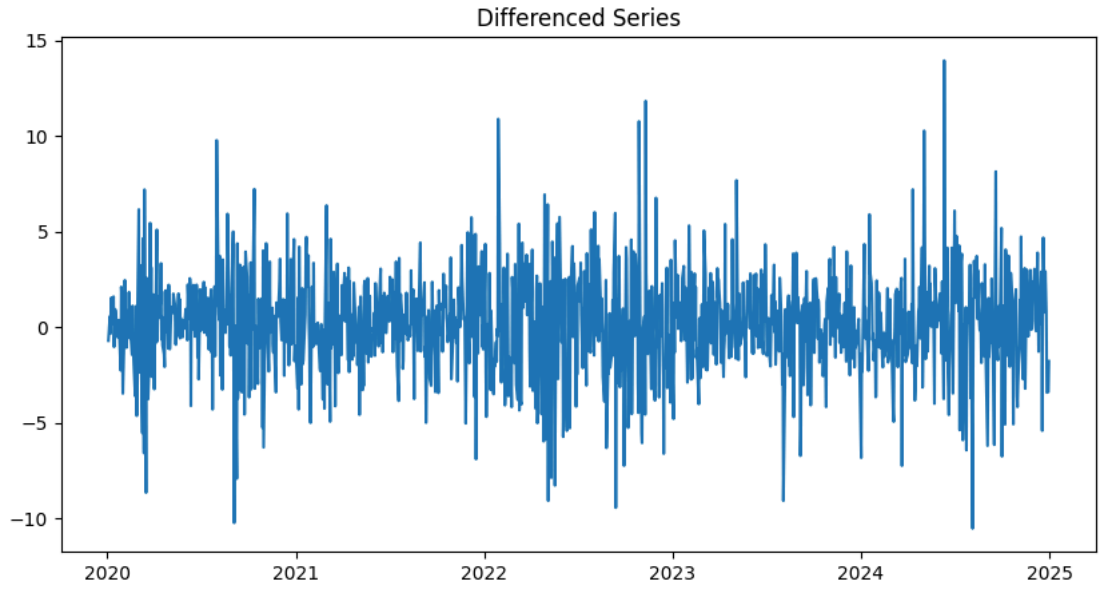



```
def adf_test(series):
    result = adfuller(series)
    print('ADF Statistic:', result[0])
    print('p-value:', result[1])
    if result[1] <= 0.05:
        print('Series is Stationary')
```

```
        print('Series is Stationary')
    else:
        print('Series is Non-Stationary')

adf_test(df['Close'])
```

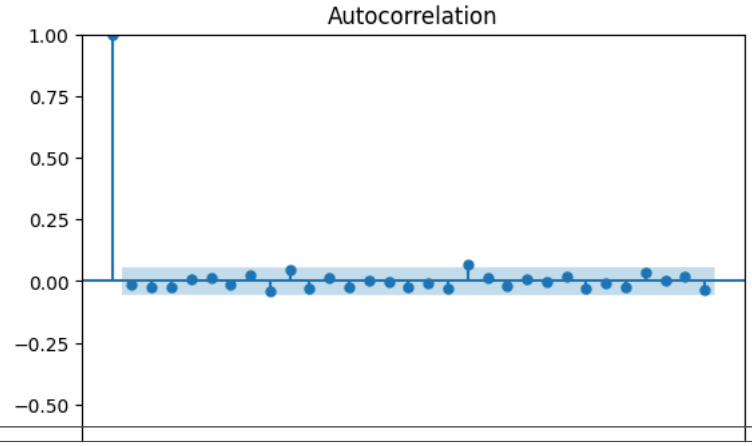
```
ADF Statistic: -0.7084056134720433
p-value: 0.8445602640910019
Series is Non-Stationary
```

```
df['Close_diff'] = df['Close'].diff()
df = df.dropna()

plt.figure(figsize=(10,5))
plt.plot(df['Close_diff'])
plt.title('Differenced Series')
plt.show()

adf_test(df['Close_diff'])
```

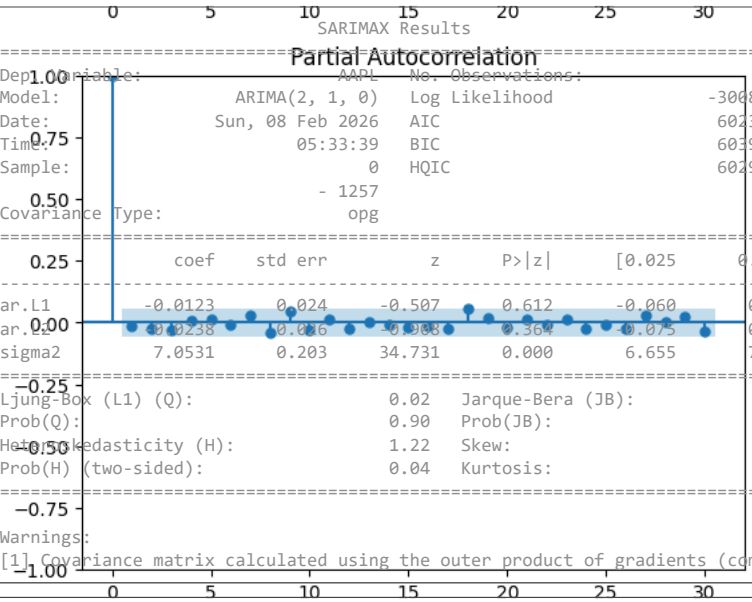



```
ADF Statistic: -35.93868949600309
p-value: 0.0
Series is Stationary
```

```
plot_acf(df['Close_diff'], lags=30)
plt.show()

plot_pacf(df['Close_diff'], lags=30)
plt.show()
```




```
ar_model = ARIMA(df['Close'], order=(2,1,0))
ar_result = ar_model.fit()
print(ar_result.summary())
```




```
                               SARIMAX Results
==============================================================================
Dep. Variable:                   AAPL   No. Observations:                 1257
Model:                 ARIMA(2, 1, 0)   Log Likelihood               -3008.967
Date:                Sun, 08 Feb 2026   AIC                           6023.935
Time:                        05:33:39   BIC                           6039.342
Sample:                             0   HQIC                          6029.725
                                - 1257
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1         -0.0123      0.024     -0.507      0.612      -0.060       0.035
ar.L2         -0.0238      0.026     -0.908      0.364      -0.075       0.028
sigma2         7.0531      0.203     34.731      0.000       6.655       7.451
===================================================================================
Ljung-Box (L1) (Q):                   0.02   Jarque-Bera (JB):               211.99
Prob(Q):                              0.90   Prob(JB):                         0.00
Heteroskedasticity (H):               1.22   Skew:                            -0.02
Prob(H) (two-sided):                  0.04   Kurtosis:                         5.01
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

```
ma_model = ARIMA(df['Close'], order=(0,1,2))
ma_result = ma_model.fit()
print(ma_result.summary())
```

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                   AAPL   No. Observations:                 1257
Model:                 ARIMA(0, 1, 2)   Log Likelihood               -3008.957
Date:                Sun, 08 Feb 2026   AIC                           6023.914
Time:                        05:33:40   BIC                           6039.321
Sample:                             0   HQIC                          6029.705
                                - 1257
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ma.L1         -0.0134      0.024     -0.552      0.581      -0.061       0.034
ma.L2         -0.0237      0.026     -0.904      0.366      -0.075       0.028
sigma2         7.0530      0.203     34.688      0.000       6.654       7.451
===================================================================================
Ljung-Box (L1) (Q):                   0.01   Jarque-Bera (JB):               211.55
Prob(Q):                              0.93   Prob(JB):                         0.00
Heteroskedasticity (H):               1.22   Skew:                            -0.02
Prob(H) (two-sided):                  0.04   Kurtosis:                         5.01
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

```
arma_model = ARIMA(df['Close'], order=(2,1,2))
arma_result = arma_model.fit()
print(arma_result.summary())
```

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                   AAPL   No. Observations:                 1257
Model:                 ARIMA(2, 1, 2)   Log Likelihood               -3008.566
Date:                Sun, 08 Feb 2026   AIC                           6027.131
```

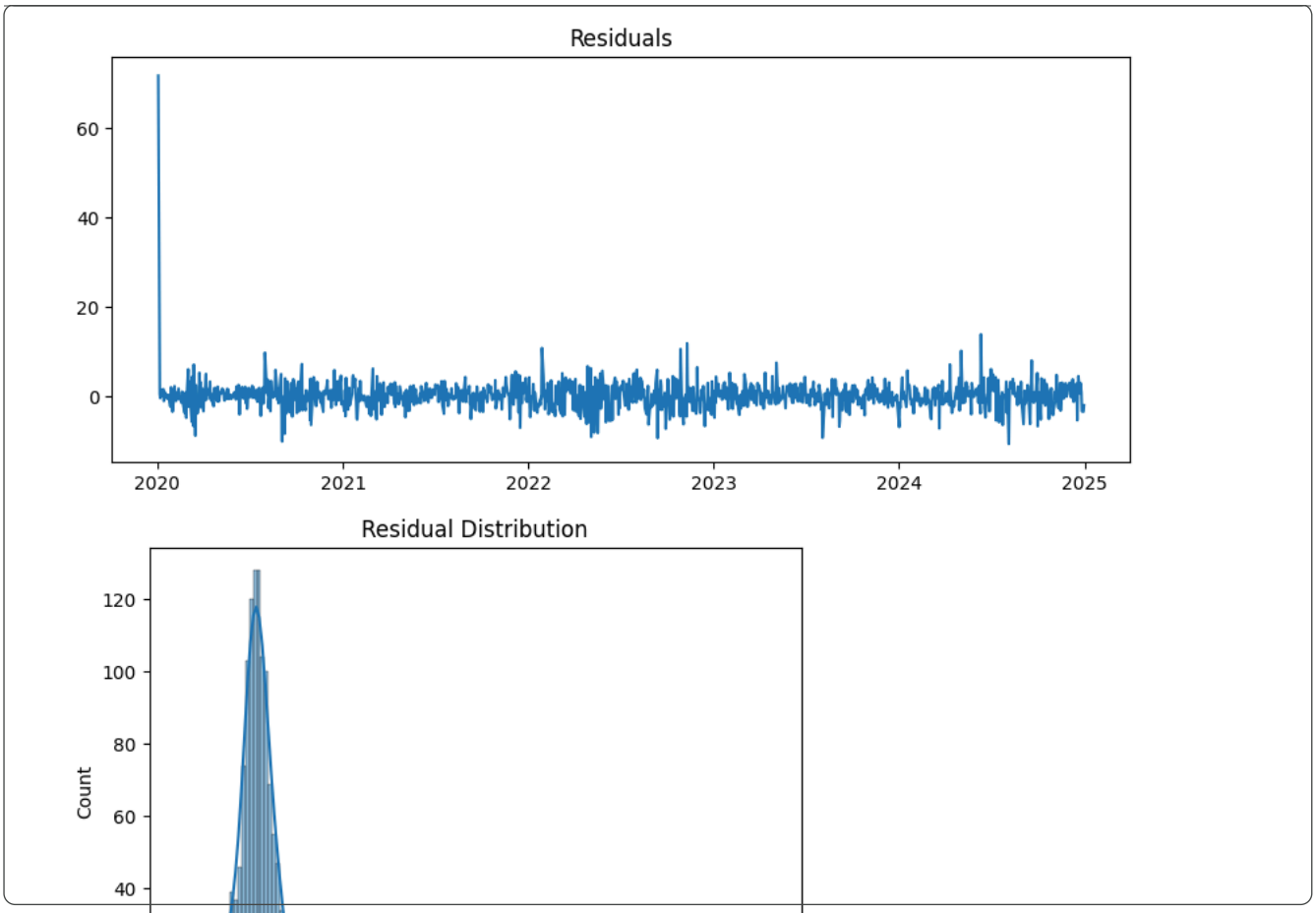
```
Time:                        05:33:41   BIC                          6052.810
Sample:                             0   HQIC                         6036.782
                                - 1257
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          0.7100      0.728      0.976      0.329      -0.716       2.136
ar.L2         -0.5001      0.584     -0.856      0.392      -1.645       0.645
ma.L1         -0.7218      0.736     -0.981      0.326      -2.164       0.720
ma.L2          0.4795      0.602      0.796      0.426      -0.701       1.660
sigma2         7.0486      0.202     34.835      0.000       6.652       7.445
===================================================================================
Ljung-Box (L1) (Q):                   0.02   Jarque-Bera (JB):               211.77
Prob(Q):                              0.88   Prob(JB):                         0.00
Heteroskedasticity (H):               1.22   Skew:                            -0.02
Prob(H) (two-sided):                  0.04   Kurtosis:                         5.01
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
```

```
comparison = pd.DataFrame({
    'Model': ['AR(2)', 'MA(2)', 'ARMA(2,2)'],
    'AIC': [ar_result.aic, ma_result.aic, arma_result.aic],
    'BIC': [ar_result.bic, ma_result.bic, arma_result.bic]
})

comparison
```

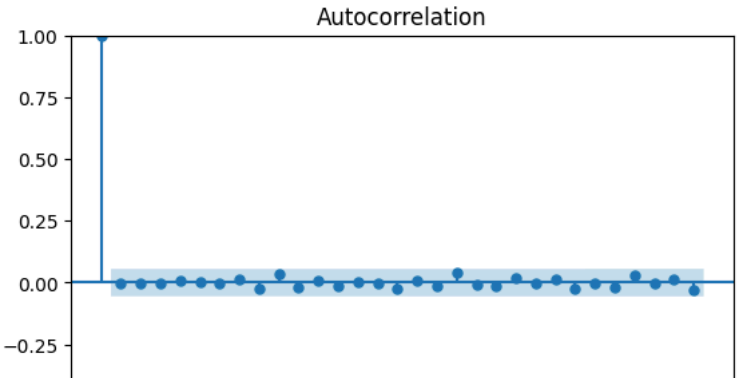
	Model	AIC	BIC
0	AR(2)	6023.934529	6039.341592
1	MA(2)	6023.914300	6039.321362
2	ARMA(2,2)	6027.131074	6052.809511

Next steps: Generate code with comparison | New interactive sheet

```
residuals = arma_result.resid

plt.figure(figsize=(10,4))
plt.plot(residuals)
plt.title('Residuals')
plt.show()

sns.histplot(residuals, kde=True)
plt.title('Residual Distribution')
plt.show()

plot_acf(residuals, lags=30)
plt.show()
```


Residuals
60
40
20
0
2020
2021
2022
2023
2024
2025

Residual Distribution
Count
120
100
80
60
40
20
0
0
20
40
60


Autocorrelation
1.00
0.75
0.50
0.25
0.00
−0.25