Sigma Internship Coding Challenge

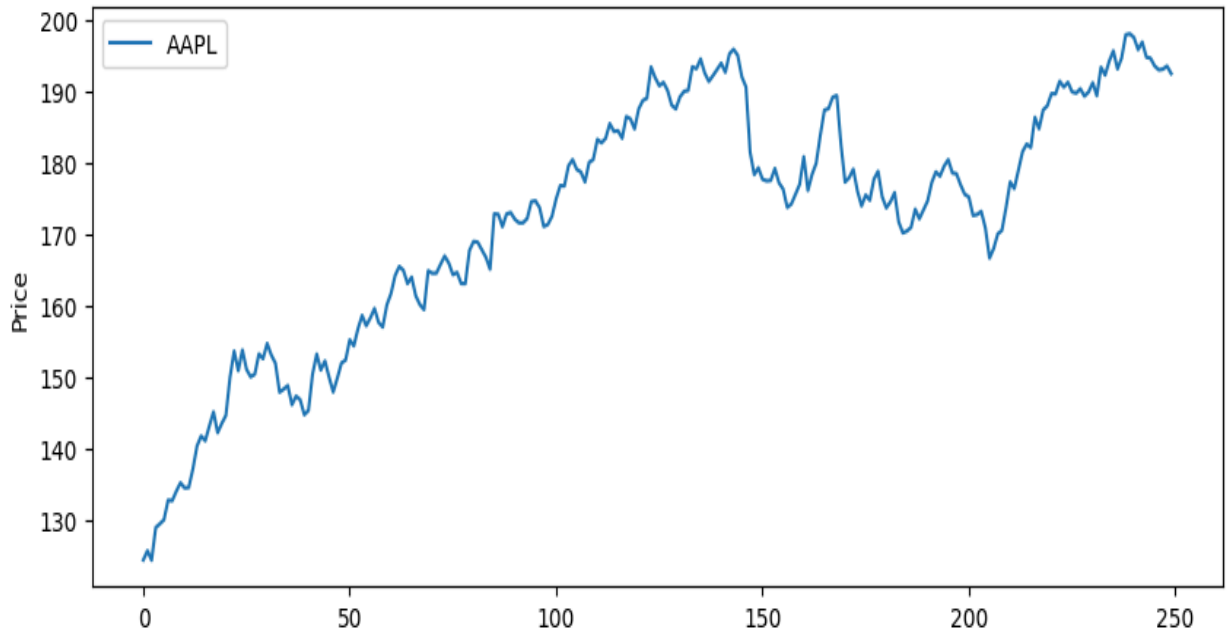
-Kaif Shaheem J

Objective is to build a simple model to make decisions on certain days using the below pre-specified logic and publish the output.

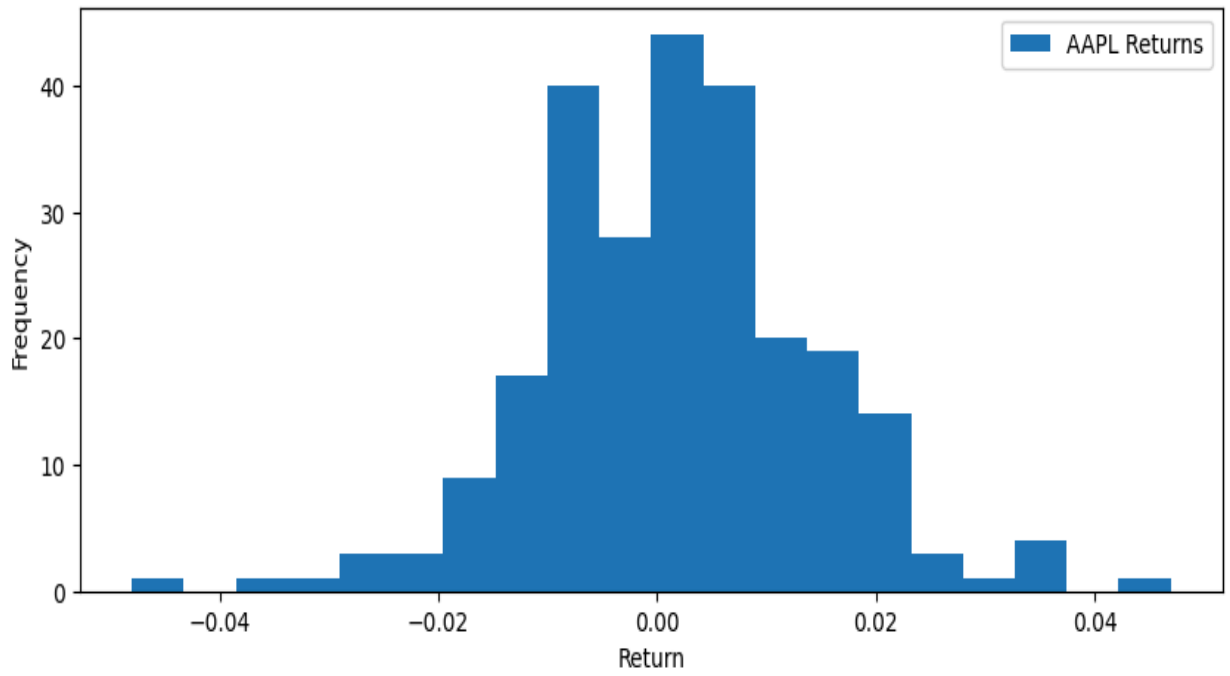
1) Data Collection Using Quantracket
2) Importing Libraries
3) Data Preprocessing
4) Data Manipulation
5) Indicators
6) Time Series Forecasting
7) Arima Model
8) Pre-Specified Logic
9) Validating using TimeGPT

Libraries Used:
⇒ Numpy
⇒ Pandas
⇒ Statsmodels
⇒ Scikit-Learn

AAPL Stock Plot Graph



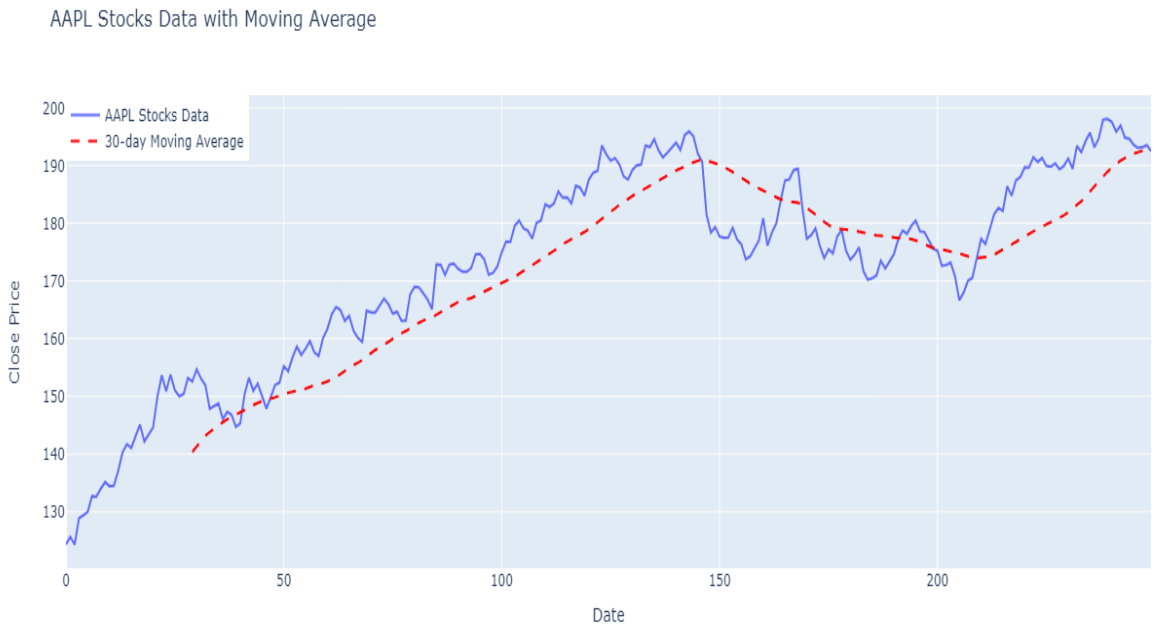
Histogram of Returns Values



Moving Average of the Stocks Data

The purpose of using a moving average, such as the 30-day moving average, is to filter out noise in the price data and provide a clearer picture of the underlying trend.

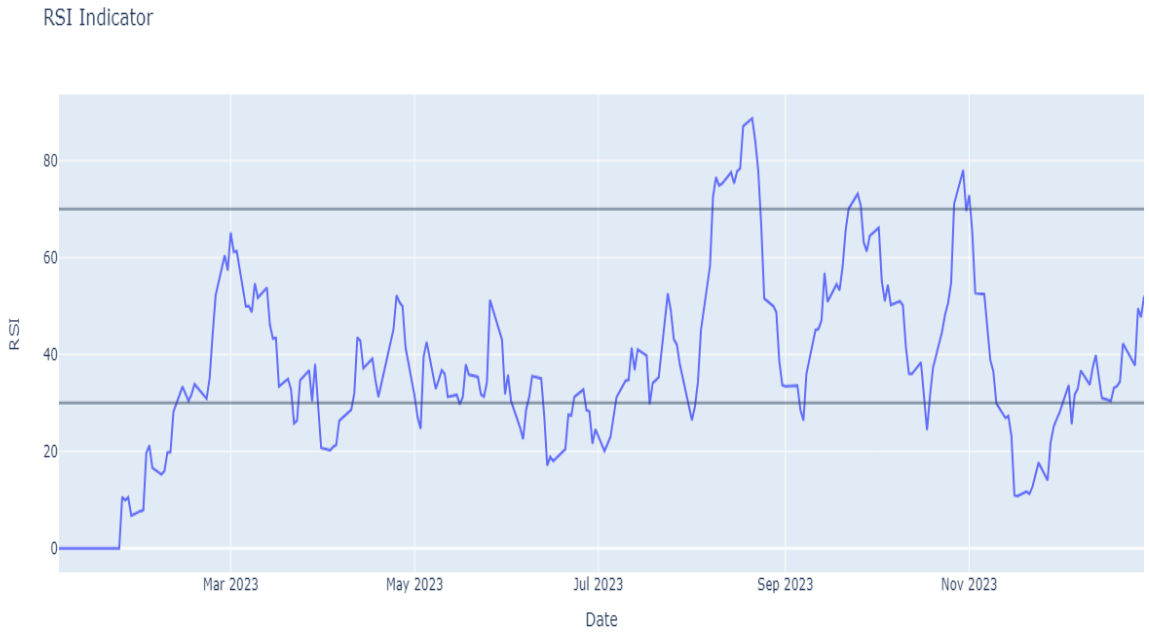
By taking the average of prices over a specific time period, it helps traders and analysts identify whether a security is in an uptrend, downtrend, or consolidating



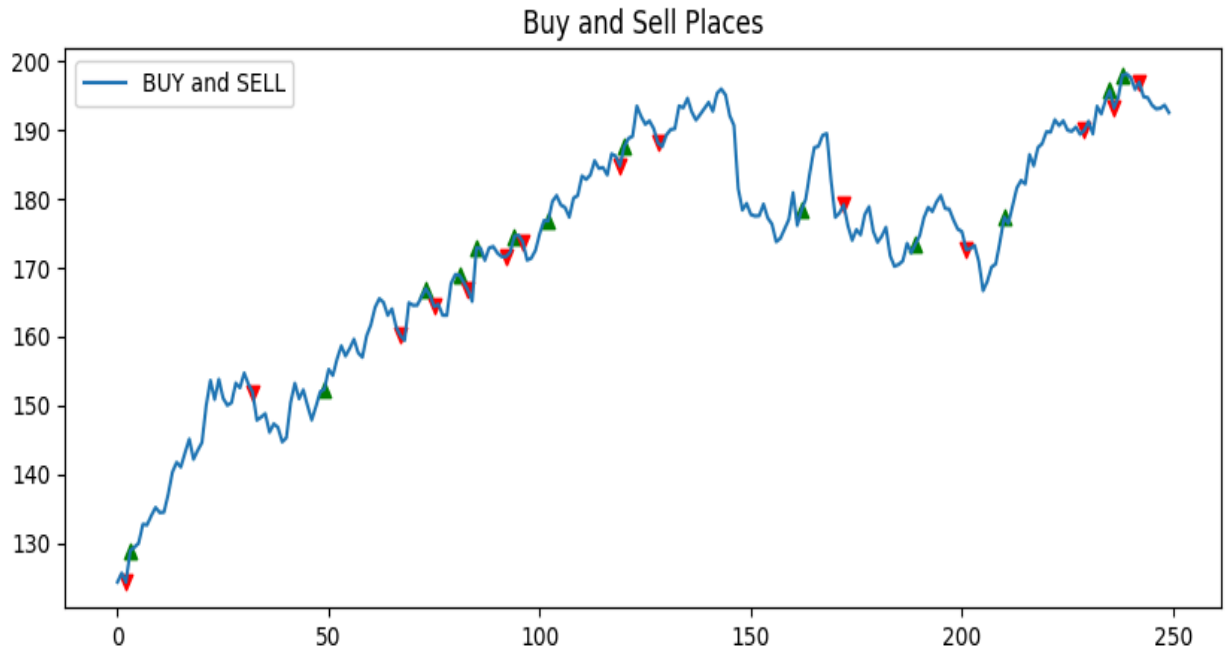
RSI indicator for the

RSI is a momentum oscillator used in technical analysis to measure the speed and change of price movements.

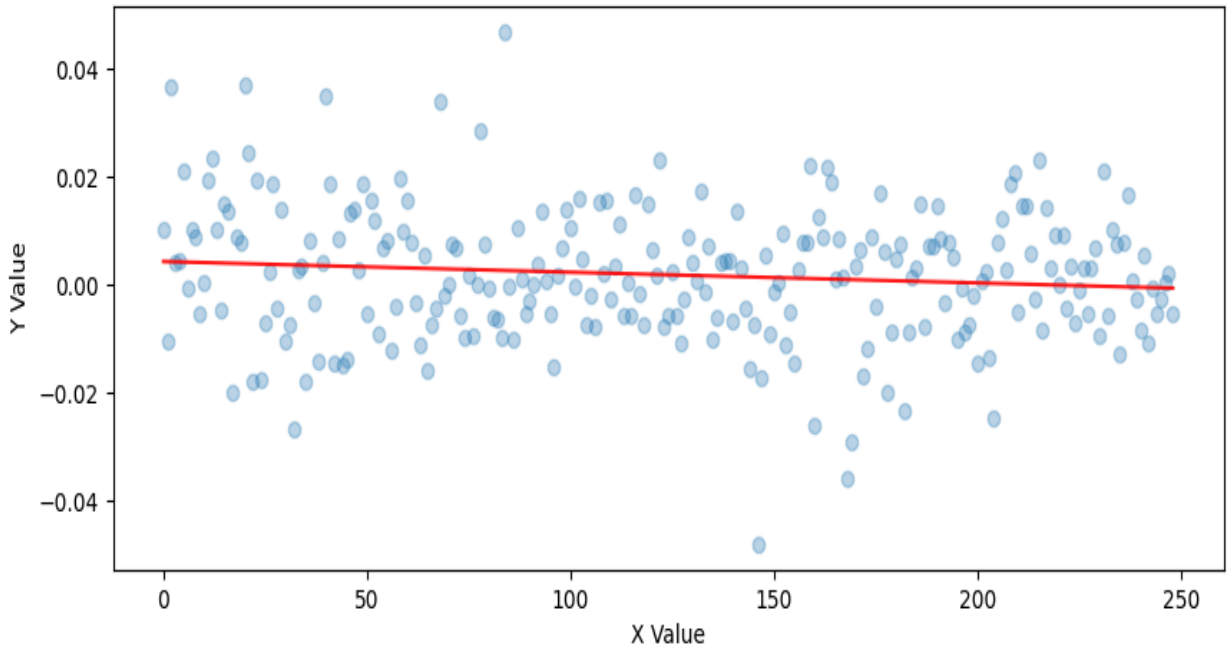
RSI compares the magnitude of recent gains and losses over a specified time period to determine whether a security is overbought or oversold.



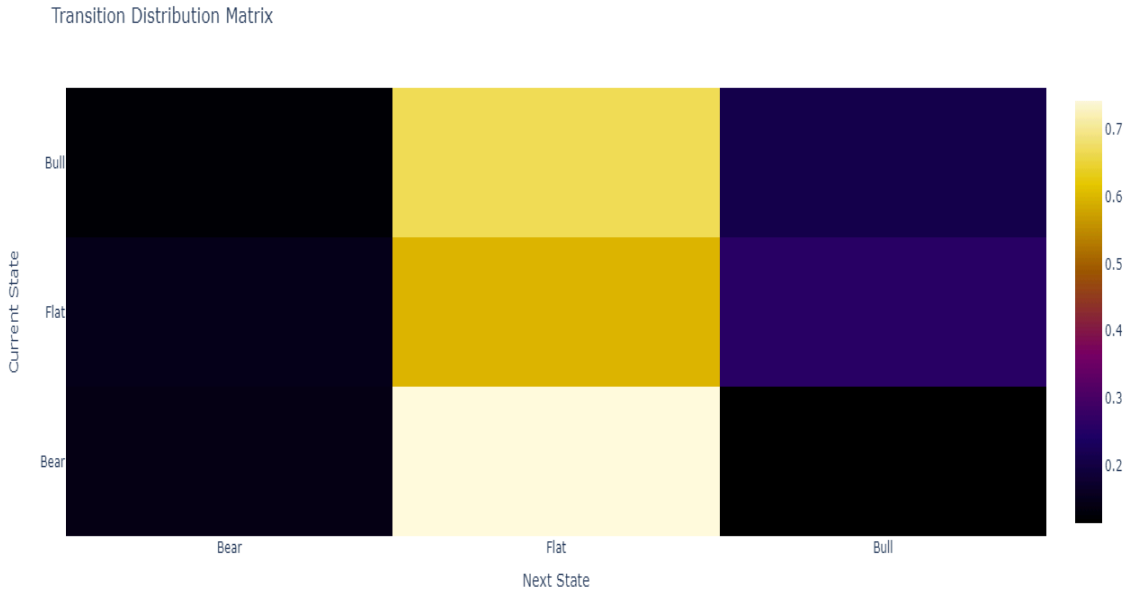
Buy and Sell Places



Linear Regression

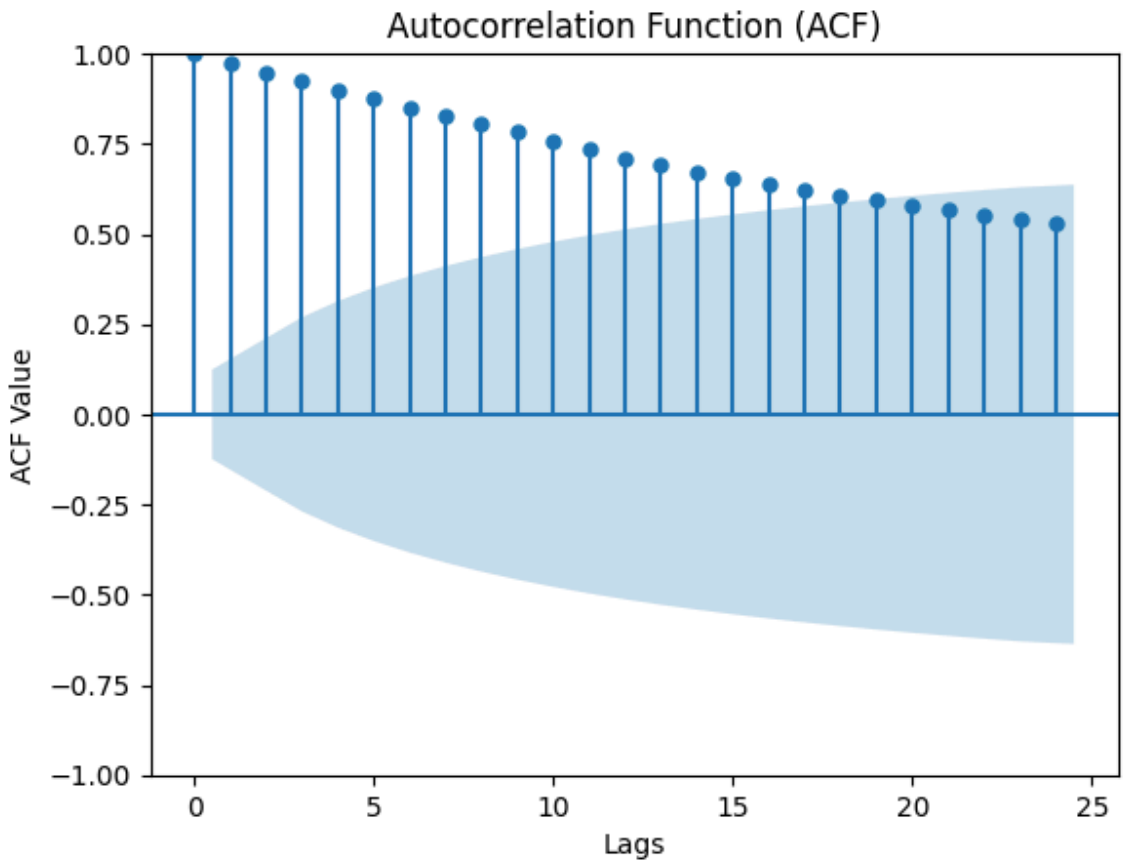


7) Transition Distribution



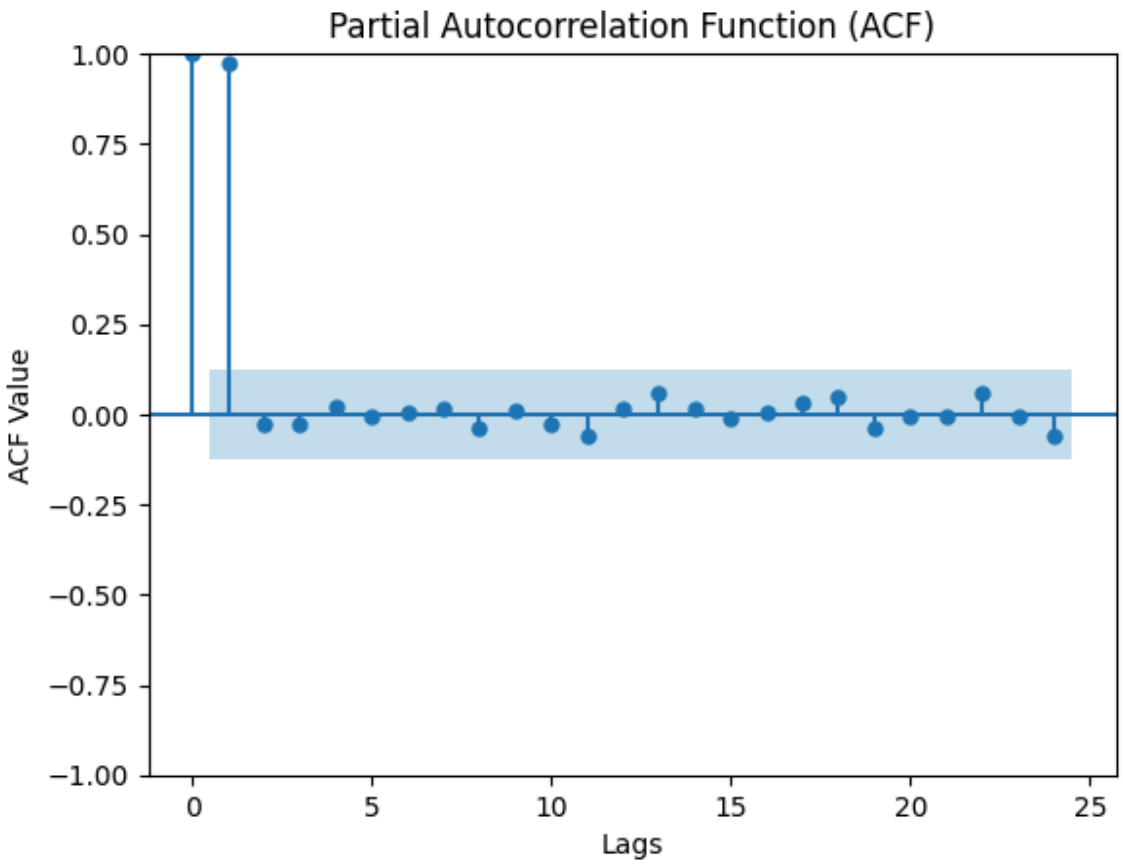
8) **Autocorrelation Function (ACF)**:

Measures the correlation between a series and its lagged values at different lags, helping identify the presence of serial correlation in the data.



9) **Partial Autocorrelation Function (PACF):**

Measures the correlation between a series and its lagged values, controlling for the influence of intervening lags, aiding in identifying the direct relationship between variables and determining the order of autoregressive terms in time series modeling.



10) The Dickey-Fuller test is a statistical test used to determine whether a unit root is present in a time series dataset. Its purpose is to assess the stationarity of a time series by testing the null hypothesis that the series possesses a unit root (i.e., it is non-stationary) against the alternative hypothesis that the series is stationary.

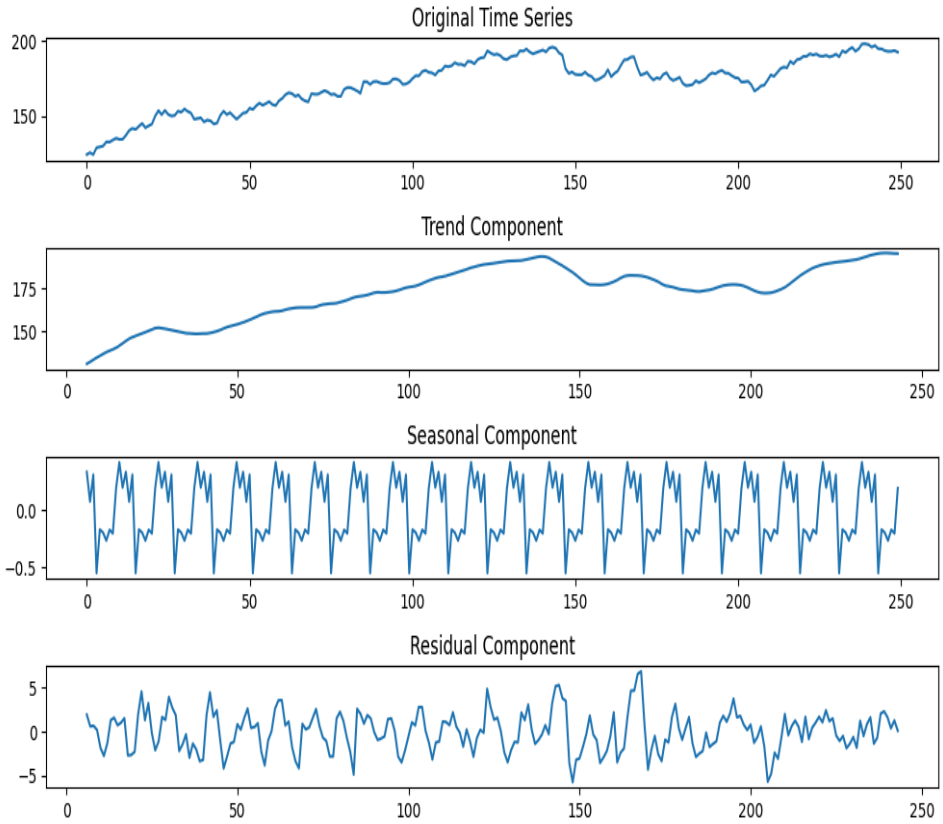
Dickey-Fuller Test Results:
Test Statistic: -2.5861110326138066
p-value: 0.09590194595133555
Critical Values:
1%: -3.4568881317725864
5%: -2.8732185133016057
10%: -2.5729936189738876

11) The **KPSS (Kwiatkowski-Phillips-Schmidt-Shin)** test is a statistical test used to assess the stationarity of a time series.

Unlike the Dickey-Fuller test, which tests for the presence of a unit root (non-stationarity), the KPSS test checks for the presence of a trend in the data. Its purpose is to determine whether a time series is stationary around a deterministic trend, as opposed to exhibiting stochastic trends or random fluctuations

KPSS Test Results:
Test Statistic: 1.7159551343844486
p-value: 0.01
Critical Values:
10%: 0.347
5%: 0.463
2.5%: 0.574
1%: 0.739

12) Seasonal Decompose:



13) Differencing the data

Where the AAPL dataset has Trend, so for predicting the values we perform the Differencing.

Pre-Specified Logic:

RMSE: 2.161124019245345

Predicted Values:

Last 15 values are predicted values for the specific date.

[124.375, 125.658, 124.325, 128.9, 129.427, 130.003, 132.748, 132.668, 134.011, 135.184, 134.458, 134.518, 137.104, 140.326, 141.738, 141.071, 143.16, 145.119, 142.205, 143.488, 144.622, 149.982, 153.641, 150.887, 153.79, 151.076, 150.031, 150.4, 153.228, 152.581, 154.702, 153.089, 151.934, 147.88, 148.308, 148.796, 146.117, 147.322, 146.814, 144.723, 145.321, 150.42, 153.209, 150.988, 152.252, 149.982, 147.9, 149.862, 151.974,

152.372, 155.22, 154.374, 156.764, 158.636, 157.192, 158.288, 159.603, 157.641, 157.013, 160.12, 161.704, 164.234, 165.499, 164.961, 163.098, 163.995, 161.375, 160.15, 159.453, 164.891, 164.543, 164.562, 165.797, 166.953, 165.977, 164.353, 164.662, 163.108, 163.098, 167.73, 168.994, 168.905, 167.859, 166.773, 165.12, 172.869, 172.799, 171.076, 172.854, 173.048, 172.111, 171.612, 171.612, 172.23, 174.584, 174.694, 173.736, 171.103, 171.382, 172.529, 174.963, 176.828, 176.778, 179.611, 180.468, 179.102, 178.733, 177.347, 180.089, 180.478, 183.301, 182.822, 183.46, 185.515, 184.428, 184.517, 183.47, 186.502, 186.183, 184.777, 187.559, 188.746, 189.085, 193.454, 191.948, 190.821, 191.299, 190.172, 188.108, 187.579, 189.265, 190.033, 190.182, 193.474, 193.214, 194.581, 192.616, 191.429, 192.237, 193.104, 193.982, 192.706, 195.309, 195.927, 195.084, 192.067, 190.661, 181.505, 178.374, 179.321, 177.716, 177.496, 177.556, 179.224, 177.217, 176.338, 173.771, 174.26, 175.609, 176.997, 180.882, 176.148, 178.375, 179.953, 183.878, 187.403, 187.623, 189.211, 189.45, 182.669, 177.326, 177.946, 179.124, 176.068, 173.981, 175.509, 174.78, 177.736, 178.834, 175.259, 173.701, 174.56, 175.848, 171.734, 170.206, 170.465, 170.985, 173.521, 172.173, 173.432, 174.68, 177.256, 178.754, 178.155, 179.563, 180.472, 178.615, 178.485, 176.917, 175.609, 175.229, 172.653, 172.772, 173.212, 170.875, 166.67, 167.999, 170.066, 170.545, 173.741, 177.336, 176.418, 178.994, 181.581, 182.649, 182.17, 186.4, 184.8, 187.44, 188.01, 189.71, 189.69, 191.45, 190.64, 191.31, 189.97, 189.79, 190.4, 189.37, 189.95, 191.24, 189.43, 193.42, 192.32, 194.27, 195.71, 193.18, 194.71, 197.96, 198.11, 197.57, 195.89, 196.94, 194.83, 194.68, 193.6, 193.05, 193.15, 193.58, 192.53, **192.13831001443083, 191.46161817753003, 191.69733396232434, 192.43813614537612, 191.69903166528383, 190.98688840014083, 191.02468596395883]**

Maximized Portfolio Value : 17

Transition Distribution :
[[0.14285714 0.74285714 0.11428571]
[0.1402439 0.61585366 0.24390244]
[0.12280702 0.66666667 0.21052632]]

Validating Using TimeGPT (LLm)