## 

The stock market is a marketplace that allows for the seamless exchange of corporate stock purchases and sales. Every Stock Exchange has its own value for the Stock Index.

## 

Before working with non-stationary data, the Autoregressive Integrated Moving Average (ARIMA) Model converts it to stationary data. One of the most widely used models for predicting linear time series data is this one.

The ARIMA model has been widely utilized in banking and economics since it is recognized to be reliable, efficient, and capable of predicting short-term share market movements.

Now consider you have a certain value A that is influenced by another value B. Then you must determine the link between data points A and B in order to determine linear regression.

### **AR (Autoregressive)**

The AR algorithm determines the linear regression of (Present fitted values) vs. (Past fitted values).

### **MA (Moving Average)**

The linear regression of the (Present value of residuals) vs. MA is discovered by MA (Past value of residuals).

We employ a mixture of the two (ARIMA).

We would require libraries called pmdarima. This is used for .autoarima. The model that we use is a function called auto\_arima.

A time series is also regarded to include three systematic components: level, trend, and seasonality, as well as one non-systematic component termed noise.

### The following are the components’ definitions:

1. The average value in the series is called the level.
2. The increasing or falling value in the series is referred to as the trend.
3. Seasonality is the series’ recurring short-term cycle.
4. The random variance in the series is referred to as noise.

Because time series analysis only works with stationary data, we must first determine whether a series is stationary.

### ADF (Augmented Dickey-Fuller) Test

One of the most widely used statistical tests is the Dickey-Fuller test. It can be used to determine whether or not a series has a unit root, and thus whether or not the series is stationary. This test’s null and alternate hypotheses are:

Null Hypothesis: The series has a unit root (value of a =1)

Alternate Hypothesis: The series has no unit root.

If the null hypothesis is not rejected, the series is said to be non-stationary. The series can be linear or difference stationary as a result of this.

The series becomes stationary if both the mean and standard deviation are flat lines (constant mean and constant variance).

We can’t rule out the Null hypothesis because the p-value is bigger than 0.05.

Seasonality and trend may need to be separated from our series before we can undertake a time series analysis. This approach will cause the resulting series to become stagnant.

It’s time to choose the ARIMA model’s p,q, and d parameters. We chose the values of p,d, and q.p is the order of the AR term

q is the order of the MA term

d is the number of differencing required to make the time series stationary

Auto ARIMA: Automatically discover the optimal order for an ARIMA model.

The auto\_arima function returns a fitted ARIMA model after determining the most optimal parameters for an ARIMA model. This function is based on the forecast::auto. Arima R function, which is widely used.

The auro\_arima function works by performing differencing tests (e.g., Kwiatkowski–Phillips–Schmidt–Shin, Augmented Dickey-Fuller, or Phillips–Perron) to determine the order of differencing, d, and then fitting models within start p, max p, start q, max q ranges. After conducting the Canova-Hansen to determine the optimal order of seasonal differencing, D, auto\_arima also seeks to identify the optimal P and Q hyper-parameters if the seasonal option is enabled

**Graph interpretation 1.1**

Top left: The residual errors appear to have a uniform variance and fluctuate around a mean of zero.

Top Right: The density plot on the top right suggests a normal distribution with a mean of zero.

Bottom left: The red line should be perfectly aligned with all of the dots. Any significant deviations would indicate a skewed distribution.

Bottom Right: The residual errors are not autocorrelated, as shown by the Correlogram, also known as the ACF plot. Any autocorrelation would imply that the residual errors have a pattern that isn’t explained by the model. As a result, you’ll need to add more Xs (predictors) to the model.

As a result, the Auto ARIMA model assigned the values 0, 1, and 0 to, p, d, and q, respectively.

In conclusion, this model is great as an addition to our deciding factor but shouldn’t be our main factor because there are many other variables that affect stock prices. One data that is important to predict stock prices is the P/E ratio. This model is great because autoARIMA decides for us which p,d,q values are the best for our model and it uses historical data to give us insight on how the data would be in the future. This is only one of the factors that should be considered in stock analysis as there are many other factors besides historical data that would affect the analysis of stock’s market price and what is the best option to exercise in stocks.

Notes:

The website is outdated and there are changes to the newer version such as it’s not longer arima\_model library but arima.model. Disp parameter does not exist anymore.