- **Yassine Sfaihi**
- www.linkedin.com/in/yassinesfaihi
- www.github.com/yassinesfaihi
- www.kaggle.com/yassinesfaih

[Author] : Yassine Sfaihi

ARIMA Time Series Forecasting - The S&P 500 Stock Market

Time series forecasting is an important task in many fields, including finance, economics, and operations research. The **Autoregressive Integrated Moving Average (ARIMA)** model is a popular method for time series forecasting. In this notebook, we will walk through the entire process of using ARIMA for time series forecasting, including data visualization, stationarity tests, parameter tuning, model building, evaluation, and model selection.

The first step in the process is to import the necessary libraries and load the time series data. We will then perform a visual check of the time series to ensure it is stationary. Stationarity is an important assumption for ARIMA, as it ensures that the properties of the data do not change over time. We will use the Augmented Dickey-Fuller (ADF) test to formally test for stationarity. If the time series is not stationary, we will use the Differencing method to make it stationary.

Once we have a stationary time series, we can proceed to determining the values of 'p' and 'q' in the ARIMA model. The 'p' value is the number of autoregressive terms, and the 'q' value is the number of moving average terms. We will use the ACF and PACF plots to determine these values.

After determining the values of 'p' and 'q', we will use the grid search method to find the optimal values of 'p', 'd', and 'q'. The 'd' value is the number of differencing used to make the time series stationary. Once we have the optimal values, we will build the ARIMA model using these values.

We will then fit the model on the training data and use it to make predictions on the test data. We will plot the predictions against the actual values to evaluate the model's performance. We will also use the root mean squared error (RMSE) to evaluate the model's performance.

Finally, we will use the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) to compare different models and select the best one. These metrics allow us to compare the relative goodness of fit of different models, which is particularly useful when we have multiple models to choose from.