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

[Author] : Yassine Sfaihi

Preparing the Data for Stock Market Analysis - Data Cleaning and Preprocessing

In this notebook, I performed data cleaning and preprocessing on a dataset of stock market data. The goal of this step was to prepare the data for further analysis and visualization. I started by importing the necessary libraries and loading the dataset into a pandas DataFrame.

I then performed various cleaning and preprocessing techniques on the data. This included handling outliers using z-score normalization, imputing missing values, and removing unnecessary columns. I also renamed columns to make the data more readable.

To handle outliers, I used z-score normalization. This technique calculates the z-score for each value in the dataset and checks if it is within a certain threshold. Values that fall outside of the threshold are considered outliers and are replaced with the median value. This helped to ensure that my data was more representative of the population.

I also handled missing values by imputing them with the mean value of the column. This helped to ensure that my data was complete and could be used for analysis.

After cleaning the data, I removed unnecessary columns that were not relevant to my analysis. I also renamed columns to make the data more readable and easier to work with.

By the end of this notebook, I had a clean and ready-to-use dataset for my next steps in analyzing the stock market. In the next notebook, I will be performing exploratory data analysis (EDA) on the cleaned dataset. This will include creating various plots such as line charts, histograms, and heatmaps, as well as calculating summary statistics. The EDA will give me a better understanding of the stock market and help inform my modeling decisions.