**NSE IT Sector Analysis**

**Data Source:**

NSE IT CSV file– [Nifty Indices Dataset | Kaggle](https://www.kaggle.com/sudalairajkumar/nifty-indices-dataset?select=NIFTY+IT.csv)

This dataset contains data of Nation stock exchange information technology sector stocks details like date, open and close price, high and low price, close price, volume, turnover, P/E ratio, P/B ratio, and dividend yield.

NSE IT website - [NIFTY IT Stock Price, Nifty IT Market Indices, Nifty IT Price, Stock Performance & Comparison (moneycontrol.com)](https://www.moneycontrol.com/indian-indices/nifty-it-19.html)

This website contains data of Nation stock exchange information technology sector stocks details like date, open and close price, high and low price, close price, company names, sectors, CMP, Changes, Volume, technical rating.

NSE API Calls - [maanavshah/stock-market-india: API for Indian Stock Market's NSE and BSE. (github.com)](https://github.com/maanavshah/stock-market-india)

This API contains data of Nation stock exchange information technology sector stocks details like date, open and close price, high and low price, close price, company names, sectors, CMP, Changes, Volume, technical rating, turnover, P/E ratio, P/B ratio, and dividend yield.

**Relationship between data:**

The relationship between each dataset here is the date. We are going to prepare a dataset that helps in finding which sector performed well on which date.

**Summary:**

The next step in my project is to clean the flat file data by replacing headers, formatting data, finding duplicates, identifying the outliers and bad data, and managing the missing data. Then in the third milestone, we need to perform the second milestone steps for the website data. It's a kind of website scraping. Using python like beautiful soup, we can scrap the website and get all the data and we can perform all the operations to clean and format the data.

In the four milestones, we need to do API calls to get the data. Since there is no API found on the internet for the API calls on the stock market, I found this project on Github and I am planning to host this project somewhere in the cloud and use it for this project. Using the python library called “request” we can establish the API calls in our project. Then we need to do a data cleaning and formatting process for the pulled data.

Finally, for the last milestone, we need to load all three datasets in a database and convert them into a single dataset. So, finally, we will get a dataset containing details of the National stock exchange of the IT sector having details like date, open and close price, high and low price, close price, company names, sectors, CMP, Changes, Volume, technical rating, turnover, P/E ratio, P/B ratio, and dividend yield, etc. Then we need to create a visualization for the data which we have using python libraries like matplotlib, seaborn, etc. The visualization will be mostly line charts and box plots because these plots are very helpful mainly for analyzing the stock price and the market situation on any particular day.