DADV Final Exam

Marks: 50 Time: 4 hours

All problems carry equal marks.

A typical link given below fetch the daily historical data for the stock Apple Inc (AAPL is ticker symbol) from March,01,2018 to December 13,2018.
https://finance.yahoo.com/quote/AAPL/history?period1=1519842600&period2=1544639400&interval=1d&filter=history&frequency=1d
By changing ticker symbols and start and end dates you get the historical data for other stocks. You can also change frequency to weekly monthly by wk and mo.
The following links give all S&P 100 and S&P 500 companies list.
https://en.wikipedia.org/wiki/S%26P_100
https://en.wikipedia.org/wiki/List_of_S%26P_500_companies

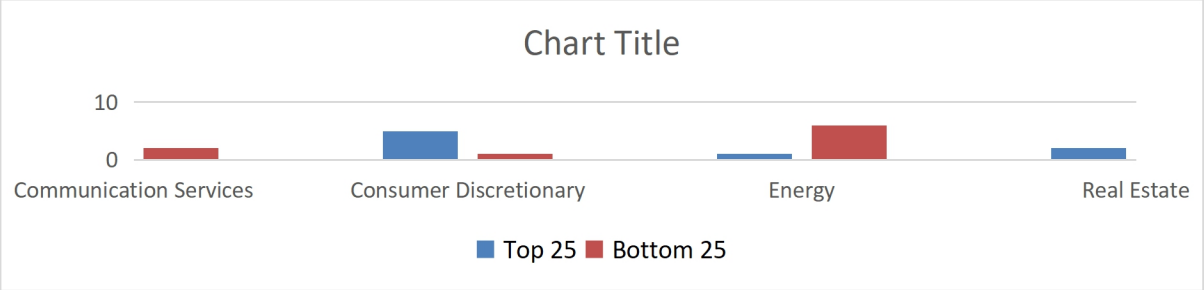
Problems:

1. By using scraping tools download the last 1000 trading days historical data (daily, weekly and monthly) for all S&P500 companies in to your system. Use parallelization to make download faster. (Note: Saturdays and Sundays and some festival days etc. are not trading days. The NYSE and NASDAQ average about 253 trading days a year)
2. If we define **daily gain or loss of a stock** is the % of change of the stock closing price to previous day closing price. Compute daily gain or loss for each S&P500 company. Do the same thing for weekly and monthly gains or losses.
3. Top 25 gainers of S&P 500 for 1000 trading days are the companies which have max % gain during those 1000 trading days. Bottom 25 gainers of S&P 500 for 1000 trading days are the companies which have least % gain (it can be negative gain too which is loss) during those 1000 trading days. Compute and draw the correlation matrix for the top 25 and bottom 25 companies for daily gains or losses. This is a matrix with both axes holding the stock names. The color of each cell is the

correlation between that pair of stocks: red for -1 and blue for 1. Do the same thing for weekly and monthly gains or losses.

4. For each company in top 25 and bottom 25, identify their GICS sectors from https://en.wikipedia.org/wiki/List_of_S%26P_500_companies and draw a bar graph depicting number of top 25 and bottom 25 belongs to that sector. Do the same thing for weekly and monthly data.

   Example:



5. For the top 2 companies, find the maximum negatively correlated companies in the remaining 48 companies from the correlation matrix. Using Tornado templates, draw sparkline plots of the 4 companies of their daily gains or losses in the same chart. Do the same thing for weekly and monthly gains or losses. Do the same thing for bottom 2 companies. Convert these files into PDF without using a browser.