ECO348-FAC413
SP 2

This project involves financial data and some elementary statistical analysis on it. It serves to refresh your memory of some basic statistical quantities you may already have learned earlier in other courses. You can use your favorite data analysis tool to do this SP – EXCEL; R; Python; STATA; SPSS; SAS; MATLAB; MATHEMATICA; or whatever (the course is, of course, **not** limited by any technology).

Write a short report on the following items:

Take the following four companies from the IT sector of the Indian financial market:

1. HCL
2. Infosys
3. TCS
4. Wipro

Collect their daily *closing prices* of common stock from 2023 (January 3) to 2024 (June 28). Make sure the data are perfectly *balanced*, meaning all four companies should have the data for the same days, and if a record for any day is missing, then you must delete the entry for that particular day for all four companies. Besides, you should ensure that none of the companies issued *stock split*[1] over the time window of this study. Do the following:

- For each company, compute the following quantities over the entire time frame:
  - Average daily price
  - SD
  - Median
  - Mode
  - Range
  - Quartiles
  - Interquartile range
  - Coefficient of variation
  - Scatterplot of daily price
  - Average slope of the daily price time series

[1] A *stock split* is a corporate decision to increase the # of outstanding shares by issuing further shares to the company's current shareholders. For example, a $3-1$ stock split issues three additional shares for each share owned by a shareholder.

- Divide the data into approximately quarters (a period of 4 months). Then, for each quarter and for each company, compute the following quantities:
    - Average daily price
    - Standard deviation
    - Median
    - Mode
    - Range
    - Quartiles
    - Interquartile range
    - Coefficient of variation
    - Scatterplot of daily price
    - Slope of the daily price time series

- Plot the average daily price and the SD of price across time measured in quarters. Describe the temporal behavior of these quantities.
- For each quarter between pairs of companies, compute the following quantities:
    - Pearson correlation coefficient (r)
- For the entire time window between pairs of companies (the pooled data set)
    - Pearson correlation coefficient (r)

Based on these different results, what can you conclude about each company and about similarities or dissimilarities between the companies over the given period? Use all your knowledge of the Indian market to enhance your report (you can Google a bit if you like but do not c/p results from some websites – the report must be your own intellectual work)

NOTE:

- In addition to the report, you must need to submit both your output. If you're using EXCEL, then submit the EXCEL file. Same for other software like R; Python; STATA; SPSS; SAS; MATLAB; MATHEMATICA.
- All submissions on BB
- Observe the due date and time on BB (**you already no late HWs are accepted**)
- If you need a little help with the data, contact the TA