

1. The given file `stock_prices.csv` contains the daily closing price of a few stocks on the NYSE/NASDAQ. You are required to do the following analytics by PySpark:
  - a. Compute the average daily return (closing price – opening price) for every stock for every day. (10 marks)
  - b. Which stock had the greatest number of times of positive daily return? (10 marks)
  - c. Which stock was the most volatile as measured by annualized standard deviation of daily maximum price difference? (10 marks)
  - d. Compute the percentage of date while the daily return of AAPL is positive but the daily return of GOOG is negative. (10 marks)
  - e. List the date when each stock has the highest and lowest closing price thought the year. (10 marks)
  - f. If we measure the trading frequency as closing price \* volume on average, which stock was traded most frequently? (15 marks)
  - g. Which month had the highest trading frequency on average? (15 marks)
  - h. Which week had the highest trading frequency on average? (Hint: use [datetime.date.isocalendar\(\)](#) to get the week number) (20 marks)