PSET 8 Python Stock Portfolio Analysis TTest Assignment

To calculate the simple rate of return for a stock, you can use this formula:

rate of return = (closing price of today – closing price of yesterday) / closing price of yesterday

- today means the newest closing price
- yesterday means the previous closing price

For example, if IBM stock is $150 today, and yesterday's closing price was $145; then we can calculate the rate of return using this formula:

- Rate of return = (150 – 145) / 145 = 0.034 = 3%.

To calculate the average rate of return for stock data collected daily from March to June, we can first calculate the daily rate of return from March to June, then calculate the average rate of return by averaging the daily rate of return during this period.

If we have both IBM and HP stocks in our portfolio and we bought 1000 shares of IBM at $145 and 500 shares of HP at $45 yesterday, then we say that HP carries a weight of 33% and IBM carries a weight of 67% in our portfolio.

To calculate the rate of return for a portfolio, we multiply the average rate of return by the weight of each stock in the portfolio and then add them up. For example, Let's say that the average rate of return of IBM is 0.06 and the average rate of return for HP is 0.1, then we will use the following equation to calculate the portfolio's rate of return:

- Portfolio rate of return = (0.06 * 0.67) + (0.1 * 0.33)

For this assignment, we provide you with a CSV file (all_stocks.csv) containing the daily closing price for IBM (ibm), Walmart (wmt), Microsoft (msft), and Amazon (amzn) from 1/1/2019 to 4/18/2022. The stock ticker is the name of each column in the CSV file.

Here are the research questions you want to answer:

1) Are there any differences in the rate of return of IBM (ibm) vs. Walmart (wmt) based on the daily rate of return of these two stocks from 1/1/2019 to 4/18/2022?
    a. You need to store either "YES" or "NO" in the variable named iw_diff
    b. You need to store the p_value from your t-test analysis in the variable named iw_pvalue
2) Assume IBM is 40 percent, and Walmart is 60 % of the portfolio. What is the average rate of return for the portfolio containing IBM and Walmart from 1/1/2019 to 4/18/2022?
    a. You need to store the average rate of return in the variable named iwp_rr
3) Assume Microsoft is 40 percent, and Amazon is 60 % of the portfolio. What is the average rate of return for the portfolio containing Microsoft (msft) and Amazon (amzn) from 1/1/2019 to 4/18/2022?
    a. You need to store the average rate of return in the variable named map_rr
4) Which portfolio has a higher rate of return?
    a. Assume IBM and Walmart are Portfolio 1, and Microsoft and Amazon are Portfolio 2. You need to store the integer number 1 or 2 (whichever has the higher rate of return) in the variable named best_portfolio

5) Did Walmart (wmt) stock perform better pre-Covid (1/1/2019 to 3/14/2020) than during the pandemic (3/15/2020 to 5/25/2021) based on the daily rate of return of the Walmart stock from 1/1/2019 to 4/18/2022?
    a. You need to store either "YES" or "NO" in the variable named better_w_covid
    b. You need to store the p_value from your t-test analysis in the variable named better_w_pvalue

You will answer these questions by analyzing the data from the CSV file using Python techniques covered in class. This assignment aims to practice logical thinking, program design, pandas, NumPy array, and T-Test.

Hints and Tips:

1) Read CSV using pandas into the data frame
2) Create a 2-dimension array to store the daily rate of return of IBM and Walmart for the t-test analysis and calculate each stock's average rate of return. The same for Microsoft and Amazon.
3) The first daily return should be zero. You will change it to zero in your program since there is no "yesterday" for the first row of data in the CSV file. For example, there is no data for 12/31/2018 for you to calculate the daily rate of return when today is 1/1/2019. Therefore, you should store a zero value for the first daily rate of return. You also need to change the daily rate of return of 3/15/2020 to zero when you are comparing the pre-Covid and during-Covid Walmart stock performance since 3/15/2020 is considered the first day of the during-Covid period (you will only change it to zero for this pre-Covid vs during-Covid comparison scenario).