STAT 3250 **Assignment 8**

Directions: Please submit your Python code file, formatted as in previous assignments.

The grader should be able to run your code when it is placed in the same directory as the input data file. Be sure that your code loads any libraries you are using.

Background: The folder `Stocks.zip` contains nearly 300 files, each including daily data for a specific stock, with stock ticker symbol given in the file name. Each observation includes the following:

```
Date = date information recorded
Open = opening stock price
High = high stock price
Low = low stock price
Close = closing stock price
Volume = number of shares traded
Adj Close = closing price adjusted for stock splits (ignored for this assignment)
```

The time interval covered varies from stock to stock. There are some missing records, so the data is incomplete. Note that some dates are not present because the exchange is closed on weekends and holidays. Those are not missing records. Answer the questions below based on the data available in the files.

1. Find the mean for the Open, High, Low, and Close entries for all records for all stocks.
2. Find the top-5 and bottom-5 stocks in terms of their average Close price. Give tables showing the stock ticker symbol and the average Close price.
3. Find the top-5 and bottom-5 stocks in terms of the day-to-day volatility of the price, which we define to be the mean of the daily differences High - Low for each stock. Give tables for each, as in the previous problem.
4. Repeat the previous problem, this time using the relative volatility, which we define to be the mean of
$$\frac{\text{High} - \text{Low}}{0.5(\text{Open} + \text{Close})}$$
for each day. As above, provide tables.
5. For each day the market was open in October 2008, find the average daily Open, High, Low, Close, and Volume for all stocks.
6. For 2011, find the date with the maximum average relative volatility for all stocks and the date with the minimum average relative volatility for all stocks. (Consider only days when the market is open.)
7. For 2010-2012, for each day of the week, find the average relative volatility for all stocks. (Consider only days when the market is open.)
8. For each month of 2009, determine which stock had the maximum average relative volatility. Give a table with the month (number is fine), stock ticker symbol, and average relative volatility.
9. The "Python Index" is designed to capture the collective movement of all of our stocks. For each date, this is defined as the average price for all stocks for which we have data on that day, weighted by the volume of shares traded for each stock. That is, for stock values $S_1, S_2, \ldots$ with corresponding sales volumes $V_1, V_2, \ldots$, the average weighted by volume is
$$\frac{S_1V_1 + S_2V_2 + \cdots}{V_1 + V_2 + \cdots}$$
Find the Open, High, Low, and Close for the Python Index for each day the market was open in January 2013. Give a table the includes the Date, Open, High, Low, and Close, with one date per row.
10. For the years 2007-2012 determine the top-5 months and years in terms of average relative volatility of the Python Index. Give a table with the month, year, and average relative volatility.