CSCI-160 *Spring 2019*

Stock Portfolio Management

BACKGROUND

Publicly traded stocks are one of the many investment instruments available to people with surplus income. Although it can be very volatile in a short period, over periods of many years the stock market will generally outperform many other alternative investments, as measured by the percentage increase in value.

For this project you will read in a data file that contains a number of stock price quotations for a collection of companies, calculate a number of statistics for that data, and display the results in an attractive tabular form.

PROJECT KIT You will need to create your own project for this one. The project kit files (all of them) should be placed inside the package folder on the same level with the src folder. Your program must run correctly with each of the 3 input files (test0.dat, test1.dat, and test2.dat) the correct output is shown in three other text files (out0.txt, out1.txt, and out2.txt).

INPUT DATA FILE FORMAT You must structure your program to read the name of the input file from the String[] args using the Run Arguments from the Project Properties.

The input file format will be a sequence of paired lines, the first line holding the name of the company (which may have embedded blanks, as in "Apple Inc.") and the second line holding a sequence of positive real values representing the share price quotations(pq_n) over some number of periods. For example,

```
Apple Inc.
114.5 120.6 130.2 128.1 126.7 129.3
```

So the generic structure of the data for a single company is

```
<company name>
<pq0> <pq1> <pqn-1>
```

where there are n price quotations.

You should write your program so that if the file is empty, it fails gracefully, that is, without throwing an exception. Just have it report that there is no data in the file. (test with data0.dat)

The number of prices given can vary from file to file, but within a file it will be the same for each company and will be at least one price. All the prices will be positive.

You will need to design and implement a Stock class and a Portfolio class. The Stock class will model a single Stock and the Portfolio class will contain an ArrayList of Stocks. Some of the statistics below are done for a Stock and should be part of that class, the extra credit statistics are all done on the entire Portfolio, so it should be responsible for those

calculations. You should not calculate anything twice. If you need it in two places, it should be a property and should be stored.

STATISTICS TO CALCULATE For each company, you should calculate the following statistics.

- Low price: the minimum price given for the company's shares, 114.5 for our example above
- High price: the maximum price given for the company's shares, 130.2 in our example
- Net change over the whole sequence: the difference between the last price and the first price, 14.8 for our example
- Average price for the whole sequence: the sum of the prices divided by the number of quotations, 124.9 for our example above
- Standard deviation for prices: this is defined as the square root of the average of the squares of the differences between each price quotation and the average. For our example, this would be the square root of
  $\sqrt{((114.5 - 124.9)^2 + (120.6 - 124.9)^2 + \ldots + (129.3 - 124.9)^2)/6}$
  which turns out to be 5.59 when rounded to two decimal places.
- Longest upward trend: if a quoted price is not LESS than its predecessor, then we say the stock is on an upward trend for that period(we could define this more strictly, but we will do it this way). A maximal length upward trend would be a sequence of price quotes $q_i, q_{i+1}, \ldots, q_j$such that:

  either i is 0 so pq_i is the first quotation given, or i >0 and $pq_{i-1} > pq_i$

  either j is n - 1 so pq_j is the last quotation given, or j <n and $pq_{j+1} < pq_j$

  The length of such a trend is one less than the number of quotations in it, j - i. Our example has two upward trends of positive length, one from 114.5 to 130.2, of length 2, and one from 126.7 to 129.3, of length 1. The longest upward trend is the largest length of all the maximal length upward trends, which is 2 in our example.
  Note, under this definition, if the stock prices declined over the entire represented interval, each individual quotation would be a maximal length upward trend of length 0.
- Best growth rate of upward trend: with a maximal length upward trend defined as above, we define the growth rate of the upward trend as $(pq_j - pq_i)/(j - i)$, provided j >i. So for our first trend, the growth rate is (130.2 - 114.5)/2, or 7.85, and the growth rate of the second is (129.3 - 126.7)/1, or 2.6. The best growth rate of an upward trend is the highest value for all of the upward trends, 7.85 for our example. This value represents the best average per period gain for the stock during its upward trends. It is undefined for a stock that declines over the entire represented interval, since such a sequence will not have positive length upward trends.

OUTPUT DISPLAY After calculating the values for each listed company, you should display them in a table with labeled columns for the company name, and each of the seven numerical statistics described above. The values for each column should line up attractively. The company names should be left justified in their column, and the numerical values should be right justified in theirs. Double values should be rounded to the hundredths place.

You can use the String class's format method to construct appropriate strings. Company names will be no longer than 20 characters, all double values will be no longer than 8 characters, which includes the minus sign(when needed), the decimal point, and the two digits of precision after the decimal point. The integer value will need no more than 3 digits. Of course, there should be at least one space character between the value of a statistic and whatever data is on the same row to its left.

Column labels need not be as long as in our list above, but should be clear. If you use a long description string, such as "Best Upward Trend Rate", if the column itself is not long enough to accommodate all the characters of the label, you should break it up into more than one line, as in:

Best Upward
Trend Rate

The last statistic may be undefined for some companies. In such cases, just put "n/a" for the entry, right justified in the column. Refer to out2.txt for formatting and correct statistics.

GRADING CRITERIA

First Five Statistics correctly calculated	
Min	12
Max	12
Net Change	12
Average	12
Standard Deviation	12
Longest Upward Trend	10
Best Growth Rate	10
Formatting of Result output	20

PROGRAMMING STYLE Your programs will be evaluated for formatting, use of meaningful identifiers, and documentation. Although no points are allocated for that on this assignment, somewhere between twenty and thirty percent of your grade will derive from the programming style on this and all future assignments. Note, the NetBeans environment can take care of formatting for you and it can generate skeleton JavaDoc which you MUST complete. Up to 30 points can be deducted for failure to adhere to the coding conventions, failure to correctly format your code, and failure to complete the documentation.

WHAT TO HAND IN Compress the entire project folder with the data files in the correct place and upload it to Blackboard.

EXTRA CREDIT You may select either or both of the choices below.

(for 10 points) We can define the rate of increase(or decrease, if the stock declines in value) for a company from period i to period i + 1 as $pq_{i+1}/pq_i = r_i$. It represents the investment multiplier for that period. If a person invested X dollars in the company at period i, that money would purchase X/pq_i shares. At the next period, those shares would be worth $pq_{i+1} * X/pq_i$ dollars, which is $r_i * X$. Add two statistical columns to the output

Best rate: the largest value of r_i for that company, for i ranging over 0, 1, . . . , n-2. Note this statistic is undefined if n is 1, that is, when there is only one price quotation in the list.

Period of best rate: the smallest i for which r_i achieves the best rate value, that is the earliest period in the list in which the best rate value is achieved. Again, this is undefined if there is only one quotation in the list.

(for 10 points)Suppose we had X dollars to invest in the market, and we could change the stocks we held at every period. Calculate and print the multiplier of X that we would have at the end of represented interval if we could go back in time and invest the value of our holdings in any of the stocks at each period. At each period we would invest all the money we have in the stock that will grow the fastest in the ensuing period. Note, if the entire market is declining, the multiplier will be less than 1. Print your calculation for the multiplier below the table, and list companies that would be chosen at each of the periods for 0, 1, . . . , n - 2 to obtain that maximum growth.

For example, if the file had

```
A
10 20 15 20
B
5 5 10 15
C
20 15 10 20
```

The the ratio's for the first period for A, B, and C are 2, 1, and 0.75, respectively and A's is the best, so we would want to buy A. The second period ratios are 0.75, 2, and 0.67, and we would want to buy B. The third period ratios are 1.33, 1.5, and 2, and we would want to buy C. The multiplier for the whole interval would be $2^3 = 8$.