CS 121 – Lab #6 – Lists of numbers

Today we will experiment with lists, specifically lists of real numbers. To begin, create a new folder on your USB drive or account called lab06, and do all your work in there. Open Anaconda Spyder as usual. You will write one Python program today, whose purpose is to calculate the mean, standard deviation, and median of a list of numbers entered by the user. You may find it helpful to refer to the class handout you received on lists. If you forgot to bring it, you can find it online on the class Web site.

1. The first thing to do is input. Create a new program called `stat.py`. At the beginning of the program, create a new empty list. You may call the list `L` or some other suitable name.

   You need to prompt the user for input. The user will enter (real) numbers, one at a time, and one number per line. When the user is done entering input, the special sentinel value –1 is used to tell your program that there is no more input. This number –1 does not go into the list.

   You should write a loop similar to the one shown in the list handout in order to accomplish the input. After the user has finished entering the input, it would be a good idea for your program to print the list, to verify that it has been stored correctly. Be sure the list is correct before continuing.

2. Your program should calculate and print the *mean* of the list. We talked about how to do this in class. Print the mean to exactly 2 decimal places. You may need to refer to earlier programs to recall how to format output. Your program should print a line of output like this:
   `The mean is  5.00`

3. The next task is to compute the standard deviation of the list. The purpose of the standard deviation is to indicate how spread out the numbers are. For example, if all the numbers in the list are the same (e.g. 5,5,5,5,5,5,5), then the standard deviation would be zero. In general, the standard deviation is defined as follows.

$$standard\ deviation = \sqrt{\frac{\sum_{i=1}^{n}(x_i - mean)^2}{n}}$$

   The formula for the standard deviation is a little complicated. So, it's wise to break up the procedure into steps. So, here is a convenient way to determine the standard deviation in Python.

   a. Create an empty list called `dev`.

b. For each number in L, let's say it's called `num`. We need to calculate the *square* of the deviation. What is a deviation? That's simple: the deviation is `num - avg`. Put the squared deviation into the `dev` list.

c. After the loop, find the average of the `dev` list, and then take the square root of this average. Now you have the standard deviation. Print the standard deviation to 2 decimal places. You should have a line that looks something like this:

```
The standard deviation is 3.54
```

4. Our next task is to calculate the median number in the list. The median means the number "in the middle" of the list, after it has been sorted.

a. First, write a Python statement in your program that says to sort the list L. Then, print the list L. Which way is the list sorted, ascending or descending?

b. The location of the median depends on whether the list has an odd or even length. We need to handle each case separately.

c. Let's suppose the list has an odd length. Consider some examples. The list could have a length of 5, 7 or 9.

0	1	2	3	4

0	1	2	3	4	5	6

0	1	2	3	4	5	6	7	8

If the list has a length of 5, what is the index of the middle cell?

If the list has a length of 7, what is the index of the middle cell?

If the list has a length of 9, what is the index of the middle cell?

Using these answers, you should be able to write a formula, in terms of the length of the list, for the index of the middle cell of the list. Write the appropriate formula for the median in your program.

d. Let's suppose the list has an even length. Consider some examples. The list could have a length of 4, 6 or 8. In this case, calculating the median is a little more interesting. There technically is no "middle cell" because of the even length. In this case, we must find the two middle cells, and average their contents. For example, in the list 1, 2, 4, 8, the median is the average of 2 and 4, which is 3.

0	1	2	3

0	1	2	3	4	5

0	1	2	3	4	5	6	7

If the list has a length of 4, what is the index of the two middle cells?

If the list has a length of 6, what is the index of the two middle cells?

If the list has a length of 8, what is the index of the two middle cells?

Finally, write two formulas. One that gives the index of the "first" middle cell, and one that gives the index of the "second" middle cell of a list of even length. Your formulas should refer to the length of the list.

Use these formulas to calculate the average of the two numbers contained in those cells. That is the median.

e. Print the median to 2 decimal places. Your output should look like this.
```
The median is  5.00
```

5. Lastly, your program should print the 3 largest numbers in the list. Your output should look like this:
```
The 3 largest numbers are  9, 8 and 7.
```

6. Please have your work checked by the instructor or lab aide. We will ask you to run your program with various test input values. Here are some test cases to run.

Input	Mean	Standard Deviation	Median	Top 3
1, 2, 8, 9	5.00	3.54	5.00	9, 8 and 2
8, 1, 2, 3, 4				
8, 15, 1, 2, 3, 4				
4.1, 4.1, 0.1, 0.1				
2, 9, 4, 2, 0, 9, 7				