

***All user-defined functions could be at the TOP of your python file and each function should have a docstring.

Create a Python file called Lab4D_StdDev and copy and paste the dailyChanges list into your script file.

A. CREATE HELPER FUNCTIONS

1. You have made similar functions in the past, but create a `basicStats(...)` function that takes in a list of values called `data` and prints the min, max, mean, median and standard deviation of the data. Also, have your function return a tuple of these values after printing them.

Recall that you can use the statistics library's `statistics.stdev(...)` method to do this but you need to import statistics first. We will assume our data is sample data so please use `stdev(...)` instead of `pstdev(...)`.

2. Create a `printIntervals(...)` function that takes in a mean and a std deviation and an optional value of `isApproxNormal` that is by default set to `false`. This function should calculate the intervals corresponding to 1/2/3 SDs around the mean and print the ESTIMATED corresponding percentage of data that we can say are in these intervals per Chebyshev's (if `isApproxNormal` is `false`) or the Empirical Guidelines (if `isApproxNormal` is `true`). THEN your function should print the EXACT percentage of data within these intervals.

NOTE: Your print statement should say "Per Cheyshev's" or "Per Empirical Guidelines". Also, recall that Chebyshev's percentages tell you at least how much data is in these intervals whereas the Empirical Guidelines tell you about how much data is in these intervals. Your output should make this clear.

HINT: To analyze each data point for the exact percentages, use the most basic for loop:

for value in data ...

3. Create a `zscore` function that takes in a value, a mean, and a standard deviation and returns the zscore of the value.

B. ANALYZING WEATHER DATA

4. A few weeks ago, I heard my friend named Christina say, "I really just don't know how to dress these days. I'll wear a scarf and then be too hot, or I'll wear a tshirt and then freeze to death." So I wondered: do the daily temperatures really change that much? I went to weather.gov to get some temperature data to address this question.

The data in the `dailyChange` list (below) is based on temperature data from the month of October between the years 2015-2018.

[1, -4, 5, -8,.....]

The 1 indicates that for the first October day recorded in our data set, the max temperature was 1 degree higher than the previous day.

The -4 indicates that for the second October day in our data set, the max temperature was 4 degrees less than the day before.

And so on. So these values represent how many degrees the max temperature changes from the previous day.

5. Call your basicStats function on the dailyChanges list.
6. Call your printIntervals function on the dailyChanges list. You have not analyzed the shape of the data so you cannot assume the data is normal. Hence, you should not set isNormal to true.
7. What is the average temperature change for October? _____

Do the temperatures seem to change very much? How are you deciding?

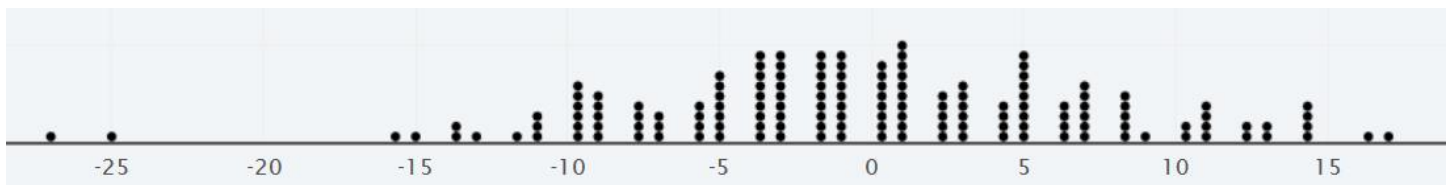
How much does it appear that the daily temperature typically changes from the previous day?

8. Below is a dot plot of the values in dailyChanges. Notice that there are two days where the max temperature is at least 25 degrees lower than that of the previous day. Use your zscore function to determine the zscores of these values.

-25 has a zscore of _____

-27 has a zscore of _____

Are these values typical or atypical. How are you deciding?



9. From the dot plot, we see that -25 and -27 are outliers in this data set. Create a second list called dailyChangesTrimmed, which is a copy of the first list but which no longer contains -25 and -27.
10. GUESS: Will the dailyChangesTrimmed's mean/standard deviation be larger/smaller/same than/as before?

The mean will be _____. The standard deviation will be _____.

Now rerun the basicStats function on dailyChangesTrimmed and confirm your guess.

11. Without the outliers, dailyChangesTrimmed is somewhat normal. Rerun printIntervals, now setting isApproxNormal to true. Record your output below.

12. Give an interval which describes the typical temperature change.

13. Below, write a paragraph (3+ sentences) to my friend Christina telling her if she can base her clothing choices for the next day based on the current day. Incorporate the mean/SD of the data and the typical interval from #12 into your discussion.

SUBMISSION INFO

TO GET CREDIT FOR THIS LAB, UPLOAD THESE 2 DOCUMENTS TO THE SUBMISSION AREA.

- Lab4D_StdDev.py (all functions defined up top; all conventions followed.)
- LabCh4D.pdf (with all blanks filled in)