Computational Problem Solving I
CPET-121
Coding Challenge 2A : Class Statistics

Problem Overview:

Your high school physics teacher has asked you to write a program to perform a statistical analysis on the final averages for her three AP Physics classes. Students' final course averages for each class are stored in three data files, *class_1.dat*, *class_2.dat* and *class_3.dat*. She would like the ability to perform the statistical analysis on each class separately and to aggregate the data into one set and perform the analysis on the combined classes.

For the statistical analysis, she would like to know (i) **Mean**, (ii) **Median**, and (iii) **Standard Deviation** for the final averages. The definition of each of these terms follows:

Mean, is simply the average of the numbers. Written mathematically:

$$Mean = \frac{\sum Grade_i}{N}$$

Median is the value separating the higher half from the lower half of the data set. Simply, it's the middle number in a data set. To find the median:

- Arrange (sort) the numbers in ascending order.
- If there is an odd number in your data set, the median is the number that is in the middle of the set.
  - For example, if the data set is: { 78.2, 83.9, 85.2, 87.8, 93.2 }. The middle number, thus the median, is 85.2.
- If you have an even number in your data set, the median is the average of the two middle numbers.
  - For example, if the data set is: { 78.2, 83.9, 85.2, 87.8, 93.2, 97.3 }. The two middle numbers are 85.2 & 87.8, thus the median, is the average of these numbers, or 86.5.

Standard Deviation is a measure of the amount of variation in a set of values. Written mathematically:

$$StdDev = \sqrt{\frac{\sum((Grade_i - Mean)^2)}{N}}$$

Note, the formula for Standard Deviation requires that the Mean be calculated first.

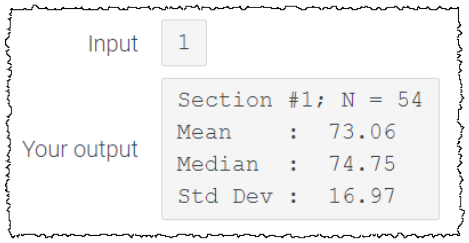
Please see https://www.mathsisfun.com/data/standard-deviation-formulas.html for more information on calculating Standard Deviation.

Code Design Specifications:

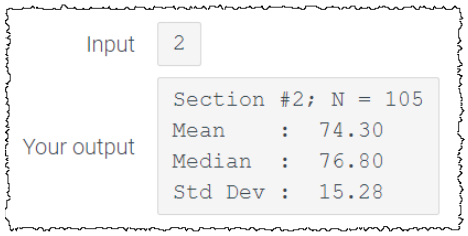
Design, code and test a procedural C++ program that performs the statistical analysis on the requested data.

- The program has one character input, 1, 2, 3, or A.
  - If a (1) is entered, the analysis is performed on the data file "*class_1.dat*"
  - If a (2) is entered, the analysis is performed on the data file "*class_2.dat*"
  - If a (3) is entered, the analysis is performed on the data file "*class_3.dat*"
  - If an (A) is entered, the analysis is performed on the aggregate of the data in files "*class_1.dat*", "*class_2.dat*", and "*class_3.dat*"
  - If any other value is entered, the program should display an error message and terminate.

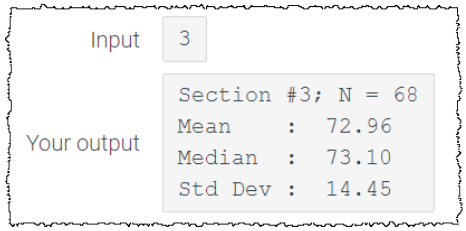
- Below are the expected outputs for the four possible inputs and one illegal input.
  - Example for (1) - data file "*class_1.dat*"



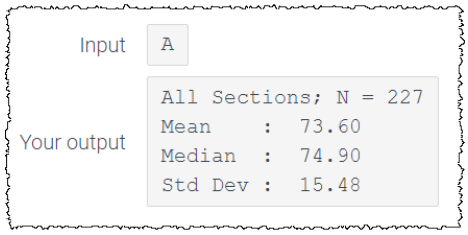
  - Example for (2) - data file "*class_2.dat*"



  - Example for (3) - data file "*class_3.dat*"



  - Example for (A) - aggregate of all three data files



- Example for (4) - an illegal input

- Note, each of the three files are of varying length. Your program must work with data files of any length, not just the examples shown.
- Note, all floating point outputs should be displayed with two-decimal points of accuracy.
- **Your program must use at least three user defined functions, one for each of the three statistical analysis.**
- If you prefer to do your code development outside the zyBook environment, the three data files are available for download.

Grading:

- Your grade for this Coding Challenge will be based on the complete and accurate implementation of the design specifications (80%) and adherence to proper coding style and commenting guidelines (20%).
- Any code that is found to be a fraudulent representation of your work, will receive a grade of zero.
- Any code that attempts to simply "match" the zyBook test-benches, will receive a grade of zero.
- Late assignments will be penalized 10% per day they are late. NO assignments will be accepted after August 6, 2021 @ 11:59 pm.