

Data Analysis Project

This assignment is designed to give you practice with program design, control and data structures, and files.

The Project

In this assignment, you will create an interactive program for data analysis. The program should ask the user whether the data will be entered directly or from a file. If entered directly the program will read one number per line until a blank line is entered. If from a file, the program will prompt the user for the name of a file that contains a set of numbers, one number per line. The program should open the file and read the data from the file into a list. Be sure to check to see if the file exists and respond accordingly. While the number of entries in the file is unspecified, you may assume the file is well formed.

Once you have read in the data, your program should sort it. Once sorted, your program should display summary statistics including the number of items in the file, the largest three items in the data, the smallest three items in the data, and the mean, the median, and the mode for the data. (If there are multiple modes, you need only report one, but it would be better to note that fact.)

Finally your program should give the user the option of writing the sorted data to a new file. Ask the user whether they want the data sorted ascending or descending and proceed accordingly. If the user wishes to save the data, the program should read in a new file name, test to see if the file exists, respond accordingly, and then write the data to the new file.

Finally, the program should give the user the option of processing more data.

Pair Programming

For this assignment, you will work in pairs. Please find someone you haven't worked with before. Your pair (team) will submit one program and you will each receive the same grade.

What to Hand In

1. Turn in *all* code files used for your project by attaching the files to an email. This should be done by the start of class on the project due date, one copy per team. Be sure to keep archival copies.
2. At the start of class on the due date, turn in one copy per pair of the printout of sample output.
3. Each of you should also turn in the answers to the following questions.

Questions

Answer the following questions and turn in your answers with your program.

1. How much time did you spend planning before you started coding?
2. How much time did you spend coding?
3. How much time did it take to correct the errors in your program?
4. What problems did you have getting your code to work?
5. How did pair-programming work for you? What worked well? What problems did you have?

Grading

The program will be graded on the following basis:

1. Error-free. The program should run without errors.
2. Documentation. The main program and every function should be properly documented. The comment for the main program should include your name, the date, and a brief description of what the program does. Each function should include a comment that includes a brief description of what the function does. Each function should also include a comment listing preconditions and post conditions.
3. Programming style. Did you use appropriate indentation and white space?
4. Completeness and correctness. Does your program fulfill all the requirements of the assignment correctly? Did you turn in everything including the questions?

Due: Start of class, Wednesday, December 5.

Sample Run:

Data Analysis Program

This program prints summary statistics for data entered directly or stored in a file.

Select input source (1 for file, 2 for keyboard): **1**

The file should contain one number per line.

Enter the name of the file: **mydata.txt**

File does not exist.

Enter the name of the file: **mydata.dat**

Data Summary:

There are 250 numbers in your data file.

The largest three are: 923, 921, 920

The smallest three are: 12, 15, 15

Mean: 512.2

Median: 504

Mode: 82 occurs 6 times

Would you like to save the sorted data? (y/n): **y**

Enter A for ascending or D for descending sort: **d**

Enter the new file name: **mysortdata.dat**

The file exists. Overwrite? (y/n): **n**

Enter the new file name: **mysortdata2.dat**

Data Saved.

Another file? (y/n): **n**

Done!