Project Number 3

Collaborators:

Diaz, Andreas Josef C.

Fernan, Simon Fredrick J.

Description: Project 3 is an update of Project 1, in which the major addition is the Line of Best Fit command.

Changes in Project 3:

- Added Line of Best Fit Command
- Added Correlation of Determination Command
- Added Switch Command
- Renovated the Read Command, it would now only read a text file with 2 columns separated by a tab
- Renovated the command system code, it is now using a system where it can accommodate beginners and experienced users.
    - Beginners can type the command and then they will be guided on what else to place.
    - Experienced users can also directly type the command with the required arguments for faster process.
- Renovated showall, head, and tail commands.
    - All of them would show x and y values side by side

- Removed the need for Argument 1
- Renovated look of HELP

List of Commands with syntax:

Command	Argument 1[1]	Argument 2	Description
read	(DataFileName.extension)[2]	N/A	reads the given file and places it in two vectors
Switch	N/A	N/A	switches x and y data
count	(x/y)	N/A	shows how many items are in the chosen data set
showall	N/A	N/A	shows all the data of the data frame[3]
head	(# of Items) [4]	N/A	shows the first few data items of the data frame
tail	(# of Items) [4]	N/A	shows the last few data items of the data frame
max	(x/y)	N/A	shows the maximum value in the chosen column/data set
min	(x/y)	N/A	shows the minimum value in the chosen column/data set

range	(x/y)	N/A	shows the range of the chosen column/data set
mean	(x/y)	N/A	shows the mean of the chosen column/data set
median	(x/y)	N/A	shows the median of the chosen column/data set
mode	(x/y)	N/A	shows only one mode of the chosen column/data set [5]
PVar	(x/y)	N/A	shows the population variance of the chosen column/data set
SVar	(x/y)	N/A	shows the sample variance of the chosen column/data set
PStdev	(x/y)	N/A	shows the population standard deviation of the chosen column/data set
SStdev	(x/y)	N/A	shows the sample standard deviation of the chosen column/data set
Hist	(x/y)	(Increment)[6]	shows the histogram of the chosen column/data set

cd	N/A	N/A	shows the correlation of determination of the data frame
lbf	N/A	N/A	shows the line of best fit and correlation coefficients of the data frame
HELP	N/A	N/A	shows all the possible commands in the program with syntax
END	N/A	N/A	Ends the program

Notes:

(1) Argument 1 is always about what data set to use (Whether x or y)

(2) Input only text files that are separated by a tab. The data in the files should only be on the first and second column.

(3) If data size is large, it is suggested not to use this command as it will take up most of the command prompt's space. The command is still possible to do on bigger data sizes.

(4) The number of items is the user-based

(5) The program would only output a single mode if there are multiple modes in the data set.

(6) The Increment is also user-based, so try to check the range first (max and min) before trying to choose a good increment number.

References

- Probability & Statistics for Engineers & Scientists; Ninth Edition; Walpole, Myers, Myers, and Ye
- How to read the lines with tab delimiters:
  - https://stackoverflow.com/questions/10617094/how-to-split-a-file-lines-with-space-and-tab-differentiation
- How to format tables:
  - https://stackoverflow.com/questions/14765155/how-can-i-easily-format-my-data-table-in-c
- Line of best fit:
  - https://www.varsitytutors.com/hotmath/hotmath_help/topics/line-of-best-fit#:~:text=A%20line%20of%20best%20fit%20can%20be%20roughly%20determined%20using,as%20many%20points%20as%20possible).
- User Input system:
  - https://stackoverflow.com/questions/9383324/nothing-at-input-cin-c
  - https://stackoverflow.com/questions/3946558/c-read-from-text-file-and-separate-into-variable