|  |  |
| --- | --- |
| **Full name of student** | Muhammad Afzal Ismail |
| **Student ID number** | 170061238 |
| **Date of meeting (Time/Room)** | 02/04/19 12.00pm  G13 |
| **Date of previous meeting** | 19/03/19 |

**Meeting Agenda:**

|  |
| --- |
| a. Discuss progress of project: - Correlation coefficient  - Scatter graph  - Median for numerical data  b. Discuss use of external library for ANOVA |

**Action Item/s from the previous meeting**

|  |
| --- |
| 1. Implement a method to read one column of a file to get the number of records it contains. It this exceeds 500,000 ask the user if they want to continue with analysing this data set. 2. Research and work on correlation coefficient. Make this appear in a new window. 3. Research on graphs in javaFX, as well as other packages.   **For supervisor:**   1. Research if the mean of a data set should be modified if there are missing values. 2. Research on ANOVA. 3. Research on sorting data by groups of columns. |

**Matters Arising from the Minutes**

|  |
| --- |
| a. Discussed the addition of Pearson’s correlation coefficient in the data visualizer. The user simply has to choose 2 columns of data type integer or double. The results will appear in a separate window, along with a scatter graph to show the relationship of the data in respect to the two variables. The results are tested with the help of an online calculator, which can be found on this link: <https://www.socscistatistics.com/tests/pearson/default2.aspx>. (Appendix A - Figure 1).  b. Demonstrated the addition of the median value for numerical data. The program will sort the data in ascending order before finding the median for the data set.  c. Discussed with the supervisor the use of an external library to calculate ANOVA. The external library is the Maths common library by Apache Commons. The website and a guide can be found on this link:  <http://commons.apache.org/proper/commons-math/userguide/stat.html>  Using this library, the program simply needs to get an arraylist of double arrays, each array containing the data for a particular column and a OneWayAnova object needs to be created.  A test program using this library was demonstrated to the supervisor. The results were verified using an online calculator, which can be found on this link:  <https://www.socscistatistics.com/tests/anova/default2.aspx>  d. Demonstrated how graphical representation of the data would be displayed by the program. The user will have to selected columns that he/she wishes to display on graphs. This will cause another window to be displayed, giving the choice of the graphs the user can select. The buttons will only be functional depending on the type of data that have been selected. (Appendix A – Figure 2).  e. Demonstrated how, upon closing the program, the actions of the user are saved in a text file in a separate folder. The text files are named according to the date and time the program was closed. The folder containing the files will be created if it is not already present. (Appendix A – Figure 3).  f. Discussed the future enhancements that can be done to the project, after the submission deadline. These include:   * Two way anova. * Save a copy of statistical information in a pdf or word document. * Make the program available as a web application. |

**Action Item/s**

|  |
| --- |
| a. Integrate the ANOVA program code in the main program.  b. Work on displaying the data in graphs.  c. Work on the final report. |

|  |  |  |
| --- | --- | --- |
| **Signature of Student:** |  | **Date: 08/04/19** |
| **Signature of Supervisor:** |  | **Date:** |

**Appendix A**

Figure 1

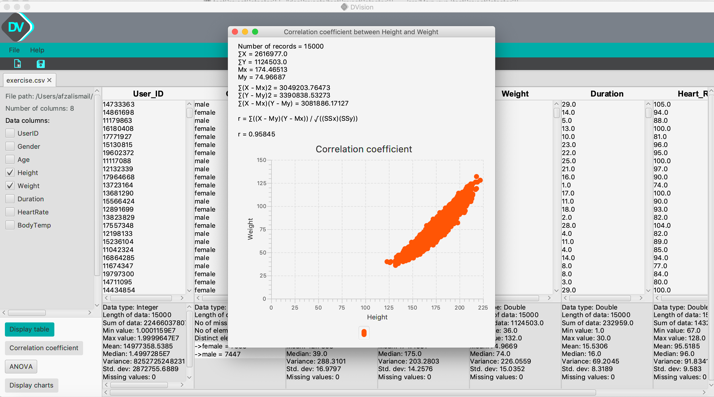


Figure 2

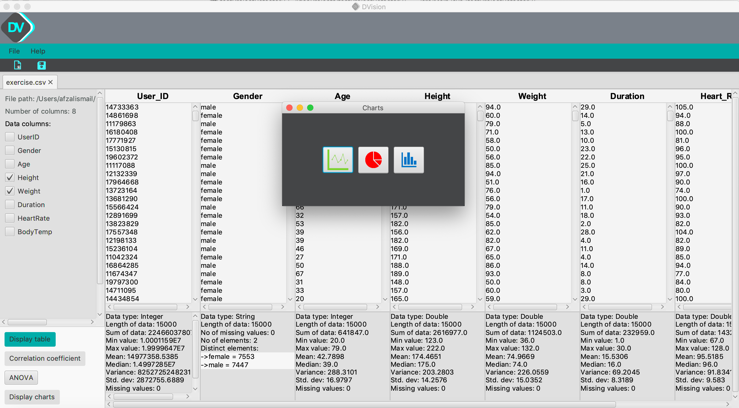




Figure 3