Data analysis and visualization on SAS studio

Name of student

Name of professor

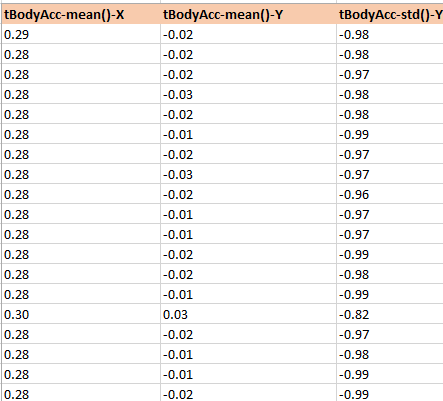
University

Course

Date

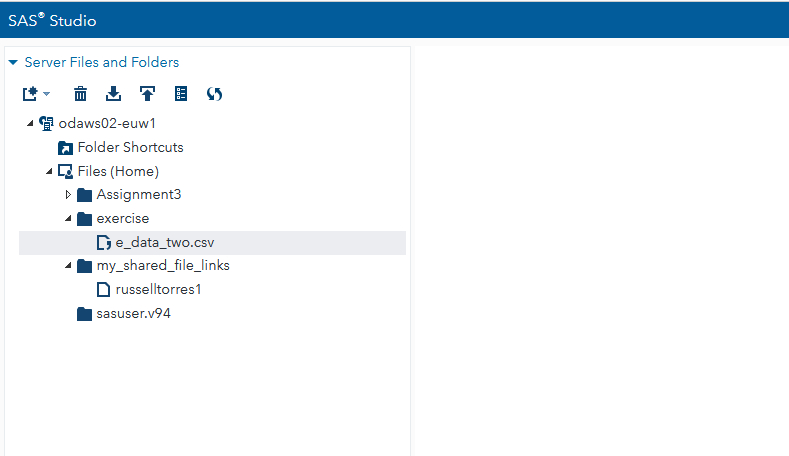
**Introduction**

The analysis and presentation of the data in this task relied on sample selections of accelerations and gyroscope data of 30 participants selected for this particular exercise. For the purpose of this analysis, a presentation of the accelerometer mean and Y deviations have been taken and presented for further analyses. A quick overview of the dataset in analysis looks like below. The last column represents the standard deviation of these data variables.

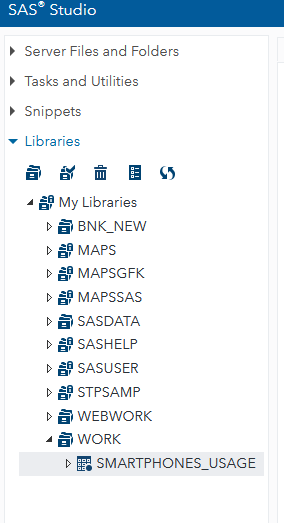


The analysis is done on SAS studio and the relevant summary statistics, descriptive statistics, graphical representations and time series forecasting of trend line data is equally established.

The first step is to load the CSV data file of the dataset into the SAS studio as shown below:



The next step is to convert the CSV data file into the SAS data type that SAS studio can recognize and use. This equally stores the data in the correct server file directory as below:

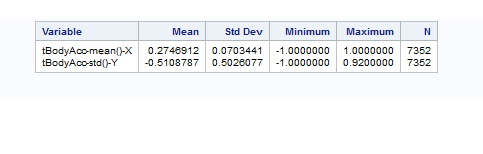


Once, this is achieved, data can be viewed as tabular format:



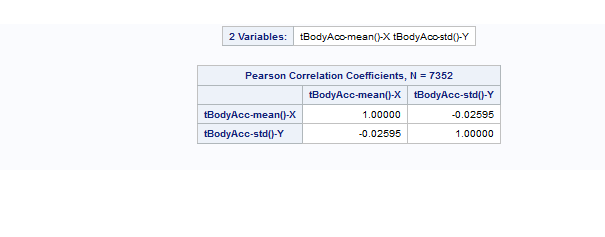
**Summary statistics**

Getting a summary overview of the datasets in this selection looks like below with means, standard deviations, minimum and maximum values well laid out for the 7352 instances of data that is to be worked on.



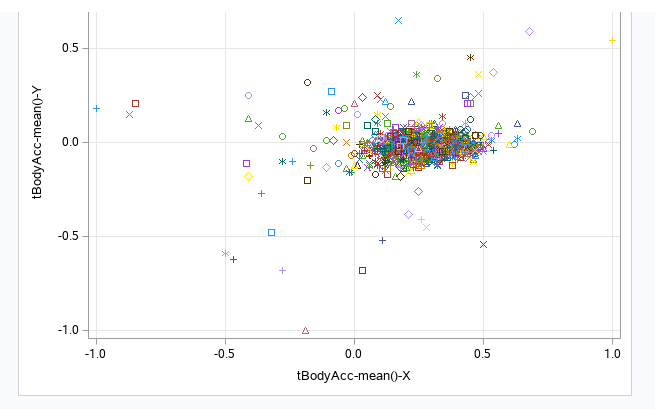
**Correlation analysis**

The findings from the accelerometer mean and its equivalent standard deviation. The relationship between the two is that of a negative one. Which means that there is no direct relationship between the accelerometer and its corresponding Y standard deviation. A change on one element does not affect the other.



**Graphical distribution**

The graphical representation shown below further proves the relationship between the factors in play. Instead of a more linear movement and distribution, the r distribution of this dataset is more of nuclear. Meaning that the change in one variable does not fact the other.



**Time series exploration and forecasting**

A time series analysis and predictive analytics supplied on this dataset reveals that the status of the X and Y input variables supplied are constant and almost uniform over time and do not seem change with change in sample population or size. This is a clear indication that the relationship between the two variables is negative and in no way related to each other. The graphical summary of these are shown below:

