The run\_analysis.R script performs data cleaning and preparing the data set

1 - Download dataset - Download and extract the data set from the url under the folder called UCI HAR Dataset

2 - Assign each data set to variables

* features <- features.txt: 561 rows, 2 columns

The features data captures the accelerometer and gyroscope 3-axial raw signals tAcc-XYZ and tGyro-XYZ.

* activities <- activity\_labels.txt: 6 rows, 2 columns

The activities data lists all the activities performed when the corresponding measurements were taken and its codes (labels)

* subject\_test <- test/subject\_test.txt: 2947 rows, 1 column

The subject\_test data set contains test data 9/30 volunteer test subjects being observed

* x\_test <- test/X\_test.txt: 2947 rows, 561 columns

The X\_test data set contains recorded features test data

* y\_test <- test/y\_test.txt: 2947 rows, 1 columns

The Y\_test data set contains test data of activities ‘code labels

* subject\_train <- test/subject\_train.txt: 7352 rows, 1 column

The subject\_train data set contains train data of 21/30 volunteer subjects being observed

* x\_train <- test/X\_train.txt: 7352 rows, 561 columns

The X\_train data set contains recorded features from training data

* y\_train <- test/y\_train.txt: 7352 rows, 1 columns

The Y\_test data set contains training data of activities ‘code labels

3 - Merge the training and the test sets to create one data set

* X (10299 rows, 561 columns) is created by merging x\_train and x\_test using rbind() function
* Y (10299 rows, 1 column) is created by merging y\_train and y\_test using rbind() function
* Subject (10299 rows, 1 column) is created by merging subject\_train and subject\_test using rbind () function
* MergedData (10299 rows, 563 column) is created by merging Subject, Y and X using cbind() function

4 - Extract only the measurements on the mean and standard deviation for each measurement

TidyData (10299 rows, 88 columns) is created by subsetting MergedData to select columns **subject**, **code** and **the measurements** which containthe mean and standard deviation for each measurement

5 - Assign descriptive activity names to name the activities in the data set

Entire numbers in code column of the TidyData set is replaced with corresponding activity name taken from second column of the activities variable

6 - Appropriately label the data set with descriptive variable names

Code column in TidyData renamed into activities

All Acc in column name replaced by Accelerometer

All Gyro in column name replaced by Gyroscope

All BodyBody in column name replaced by Body

All Mag in column name replaced by Magnitude

All column name that start with character f is replaced by Frequency

All column name that start with character t is replaced by Time

-mean() is replaced with Mean

-std() is replaced with STD

-freq() is replaced with Frequency

7- From the data set in step 4, creates a second, independent tidy data set with the average of each variable for each activity and each subject

FinalData (180 rows, 88 columns) is created by summarizing TidyData taking the means of each variable for each activity and each subject, on the independent tidy data set grouped by subject and activity.

Export FinalData into FinalData.txt file.