================================================ Coursera Course Project for Getting and Cleaning Data ================================================ ## A. Required tasks: 1. Merges the training and the test sets to create one data set. 2. Extracts only the measurements on the mean and standard deviation for each measurement. 3. Uses descriptive activity names to name the activities in the data set 4. Appropriately labels the data set with descriptive variable names. 5. Creates a second, independent tidy data set with the average of each variable for each activity and each subject

## B. Script task explanation:

* Read activity labels (./activity\_labels.txt) to create the activity data table consisting of activity Code (1-6) and equivalent six activity labels
* Read all 561 measurement features from (./features.txt)
* Record the index of measurement features which are only relating to the mean or standard deviation values.
* Use the index vector record to extract only 66 "mean/stdev"-type measurement features
* Modify the descriptions of column names of measurement features (e.g. tBodyAcc-mean()-X become Time.BodyAcc.Mean.X)
* Merge two data sets including x\_test (./test/x\_test.txt) of 2947 observations and x\_train (./train/X\_train.txt) of 7352 observations into a big data set with 10299 observations and 561 measurement features. The outcome of this process is called merged data.
* Using the record of "mean/stdev"-type measurement features to filter this merged data. The outcome of this process is called filtered merged data.
* Merge test subject data (./test/subject\_test.txt) and train subject data (./train/subject\_train.txt)
* Merge test activity data (./test/y\_test.txt) and activity data(/train/Y\_train.txt)
* Substitute the activity codes with the relevant descriptive activity labels
* Merge all three data sets: merged subject data, merged activity label data and filtered merged data. THe outcome of this process is a data frame with 10299 observations and 68 variables.
* Create tidy data set by data reshape, listing the average of 66 "mean/stdev"-type measurement data for each activity and subject. This tidy data have only 35 observations and 68 variables.
* Write down the tidy data to the hard disk