Readme:

One of the most exciting areas in all of data science right now is wearable computing - see for example this article . Companies like Fitbit, Nike, and Jawbone Up are racing to develop the most advanced algorithms to attract new users. The data linked to from the course website represent data collected from the accelerometers from the Samsung Galaxy S smartphone.

Steps to work on this course project:

1.Download and unzip the file (https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip) under current working directly. You will obtain data file “UCI HAR Dataset”.

2. Create run\_analysis.R under folder “UCI HAR Dataset”. Set “./UCI HAR Dataset” as working directly by setwd().

This run\_analysis.R script perform the following work about this data file:

a.Reformat the test and train data set separately, using measurement variables as column names. Two separate columns subject (30 volunteers) and activity (6 activities: WALKING, WALKING\_UPSTAIRS, WALKING\_DOWNSTAIRS, SITTING, STANDING, LAYING) are added to each data set.

b. Merge two data sets (train and test) into complete data set (completedata)

c. Subset completedata from step2 with column names containing activity, subject, mean(), and std(). The columns names with “meanFreq()” is not considered as mean value calculation since it is weighted average of the frequency components to obtain a mean frequency.

d. Replace column “ activity” with meaning content (6 activities: WALKING, WALKING\_UPSTAIRS, WALKING\_DOWNSTAIRS, SITTING, STANDING, LAYING). It creates a data file called completedata\_sub.

e. Use reshape2 library to melt and dcast the completedata\_sub file using “subject” and “activity” as ID variable. Calculate mean for the measurement variables.

3. Run source(run\_analysis.R), it will generate “tidy.data.txt” under folder UCI HAR Dataset.