

## CodeBook for the final project in Getting and Cleaning Data

The `run_analysis.R` script is created according to the instructions in Getting and Cleaning Data Course Project. The project is composed by 5 steps.

### 1. Download step

- o Download and extract data in the “UCI HAR Dataset” folder

### 2. Assign data to variables step

- o `features <- features.txt` : 561 rows, 2 columns  
*3-axial raw signals tAcc-XYZ and tGyro-XYZ from the accelerometer and gyroscope.*
- o `activities <- activity_labels.txt` : 6 rows, 2 columns  
*Codes and descriptions of 6 typical activities*
- o `subject_test <- test/subject_test.txt` : 2947 rows, 1 column  
*test data of test subjects*
- o `x_test <- test/X_test.txt` : 2947 rows, 561 columns  
*recorded features test data*
- o `y_test <- test/y_test.txt` : 2947 rows, 1 columns  
*test data of activities*
- o `subject_train <- test/subject_train.txt` : 7352 rows, 1 column  
*train data of subjects under observation*
- o `x_train <- test/X_train.txt` : 7352 rows, 561 columns  
*train data features*
- o `y_train <- test/y_train.txt` : 7352 rows, 1 columns  
*activities' code train data*

### 3. Merging training and test sets to create a single data set

- o Merge `x_train` and `x_test` using `rbind()` function to obtain `x` (10299 rows, 561 columns)
- o Merge `y_train` and `y_test` using `rbind()` function to obtain `y` (10299 rows, 1 column)
- o Merge `subject_train` and `subject_test` using `rbind()` function to obtain `Subject` (10299 rows, 1 column)
- o `AllData` (10299 rows, 563 column) is created by merging `Subject`, `y` and `x` using `cbind()` function

### 4. Extracting only mean and standard deviation for each measurement

- o Select from `AllData`, only columns: `subject`, `code` and mean and *standard deviation* (`std`) to obtain `AllDataMeanStd` (10299 rows, 88 columns)

### 5. Using descriptive activity names to rename activities in the dataframe

- Replace codes in the data-frame `AllDataMeanStd`, column `code` with corresponding labels available in the second column of `activities` data-frame

## 6. Labelling the data set with descriptive variable names

- The name of the columns belonging to the data-frame `AllDataMeanStd` are renamed as follows

- `code` renamed as `Activities`
- All `Acc` renamed as `Accelerometer`
- All `Gyro` renamed as `Gyroscope`
- All `BodyBody` renamed as `Body`
- All `Mag` renamed as `Magnitude`
- All strings starting with character `f` renamed as `Frequency`
- All strings starting with character `t` renamed as `Time`

The final data-frame is then called `TidyDf1`

## 7. Creating a new data set from the previous one with the average of each variable for each activity and each subject

- Group in a temporary data-frame called `SumAllDataMeanStdTemp` the data by `Subject` and `Activity`
- Compute the means of each variable for each activity and each subject and store them in data-frame `SumAllDataMeanStd` (180 rows, 88 columns)

The final data-frame is then called `TidyDf2`

## 8. Exporting/writing data into a txt file

- Export `TidyDf2` into `FinalData.txt` file with the function `write.table`