### Code Book

The run\_analysis.R script performs the data preparation and then followed by the 5 steps required as described in the course project’s definition.

1. **Download the dataset**
   * Dataset downloaded and extracted under the folder called UCI HAR Dataset
2. **Assign each data to variables**
   * features <- features.txt : 561 rows, 2 columns   
     *The features selected for this database come from the accelerometer and gyroscope 3-axial raw signals tAcc-XYZ and tGyro-XYZ.*
   * activities <- activity\_labels.txt : 6 rows, 2 columns   
     *List of activities performed when the corresponding measurements were taken and its codes (labels)*
   * subject\_test <- test/subject\_test.txt : 2947 rows, 1 column   
     *contains test data of 9/30 volunteer test subjects being observed*
   * x\_test <- test/X\_test.txt : 2947 rows, 561 columns   
     *contains recorded features test data*
   * y\_test <- test/y\_test.txt : 2947 rows, 1 columns   
     *contains test data of activities’code labels*
   * subject\_train <- test/subject\_train.txt : 7352 rows, 1 column   
     *contains train data of 21/30 volunteer subjects being observed*
   * x\_train <- test/X\_train.txt : 7352 rows, 561 columns   
     *contains recorded features train data*
   * y\_train <- test/y\_train.txt : 7352 rows, 1 columns   
     *contains train data of activities’code labels*
3. **Merges 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
   * Merged\_Data (10299 rows, 563 column) is created by merging Subject, Y and X using **cbind()** function
4. **Extracts only the measurements on the mean and standard deviation for each measurement**
   * TidyData (10299 rows, 88 columns) is created by subsetting Merged\_Data, selecting only columns: subject, codeand the measurements on the mean and *standard deviation* (std) for each measurement
5. **Uses descriptive activity names to name the activities in the data set**
   * Entire numbers in code column of the TidyData replaced with corresponding activity taken from second column of the activities variable
6. **Appropriately labels the data set with descriptive variable names**
   * code column in TidyData renamed into activities
   * All Acc in column’s name replaced by Accelerometer
   * All Gyro in column’s name replaced by Gyroscope
   * All BodyBody in column’s name replaced by Body
   * All Mag in column’s name replaced by Magnitude
   * All start with character f in column’s name replaced by Frequency
   * All start with character t in column’s name replaced by Time
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 sumarizing TidyData taking the means of each variable for each activity and each subject, after groupped by subject and activity.
   * Export FinalData into FinalData.txt file.

README

Peer-graded Assignment: Getting and Cleaning Data Course Project

This repository is a **Nunno Nugroho** submission for Getting and Cleaning Data course project. It has the instructions on how to run analysis on Human Activity recognition dataset.

Dataset

[Human Activity Recognition Using Smartphones](http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones)

Files

* CodeBook.md a code book that describes the variables, the data, and any transformations or work that I performed to clean up the data
* run\_analysis.R performs the data preparation and then followed by the 5 steps required as described in the course project’s definition:
  + Merges the training and the test sets to create one data set.
  + Extracts only the measurements on the mean and standard deviation for each measurement.
  + Uses descriptive activity names to name the activities in the data set
  + Appropriately labels the data set with descriptive variable names.
  + 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.txt is the exported final data after going through all the sequences described above.