Code book document

This code book gives the information about each variable used in the data set, information about the summary of each script choice and information about the experimental study design used.

**Study design:**

The experiments have been carried out with a group of 30 volunteers within an age bracket of 19-48 years. Each person performed six activities (WALKING, WALKING\_UPSTAIRS, WALKING\_DOWNSTAIRS, SITTING, STANDING, LAYING) wearing a smartphone (Samsung Galaxy S II) on the waist. Using its embedded accelerometer and gyroscope, we captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz. The experiments have been video-recorded to label the data manually. The obtained dataset has been randomly partitioned into two sets, where 70% of the volunteers was selected for generating the training data and 30% the test data.  
  
The sensor signals (accelerometer and gyroscope) were pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap (128 readings/window). The sensor acceleration signal, which has gravitational and body motion components, was separated using a Butterworth low-pass filter into body acceleration and gravity. The gravitational force is assumed to have only low frequency components, therefore a filter with 0.3 Hz cutoff frequency was used. From each window, a vector of features was obtained by calculating variables from the time and frequency domain.

Attribute Information:

For each record in the dataset it is provided:  
- Triaxial acceleration from the accelerometer (total acceleration) and the estimated body acceleration.  
- Triaxial Angular velocity from the gyroscope.  
- A 561-feature vector with time and frequency domain variables.  
- Its activity label.  
- An identifier of the subject who carried out the experiment.

**Source data:**

A full description is available at the site where the data was obtained:

<http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

Here is the data for the project:

<https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip>

**Purpose of R script**

Create one R script called run\_analysis.R that does the following.

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. 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

Code Book:

Variables

* x\_test, y\_test, subject\_test, x\_train, y\_train, subject\_train, data\_features, activity\_labels contains the data from the downloaded files.
* train\_combine, test\_combine are merged data sets
* data\_features contains the names of all features.

Functions

* Tidy data contains 180 observations of 480 variables of specifically mean and standard deviations.