**Code book:**

**Describes the variables, the data and steps performed to clean up the data**

**Variables and data**

Data used in this analysis were collected from Smartphone-Based Recognition of Human Activities and Postural Transitions Data Set. 30 volunteers that were randomly assigned in two groups (training group made up of 70% or 21 of the volunteers and 30% or 9 of the volunteers). The subjects wore Samsung Galaxy S II, equipped with an Accelerometer and a Gyroscope, on their waist-belt. The activities performed by the subjects were:

* WALKING
* WALKING\_UPSTAIRS
* WALKING\_DOWNSTAIRS
* SITTING
* STANDING
* LAYING

The raw signals captured by the accelerometer and gyroscope were the tAcc-XYZ and tGyro-XYZ respectively. The acceleration signal was then separated into body and gravity acceleration signals (tBodyAcc-XYZ and tGravityAcc-XYZ), while the Jerk signals (tBodyAccJerk-XYZ and tBodyGyroJerk-XYZ) were derived from the body linear acceleration and angular velocity. The magnitude of these three-dimensional signals (in the X, Y, Z directions) were calculated using the Euclidean norm (tBodyAccMag, tGravityAccMag, tBodyAccJerkMag, tBodyGyroMag, tBodyGyroJerkMag). Some of the data were also transformed using the Fast Fourier Transform (FFT) to derive fBodyAcc-XYZ, fBodyAccJerk-XYZ, fBodyGyro-XYZ, fBodyAccJerkMag, fBodyGyroMag, fBodyGyroJerkMag. The analyzed variables include:

tBodyAcc-XYZ

tGravityAcc-XYZ

tBodyAccJerk-XYZ

tBodyGyro-XYZ

tBodyGyroJerk-XYZ

tBodyAccMag

tGravityAccMag

tBodyAccJerkMag

tBodyGyroMag

tBodyGyroJerkMag

fBodyAcc-XYZ

fBodyAccJerk-XYZ

fBodyGyro-XYZ

fBodyAccMag

fBodyAccJerkMag

fBodyGyroMag

fBodyGyroJerkMag

The final tidy data was only the mean and standard deviation of all the above mentioned variables in each X, Y, Z directions as indicated.

**Steps taken to clean up data:**

 run\_analysis.R is an R- script that gets and cleans data collected from the accelerometers from the Samsung Galaxy S smartphonperforms in the following 5 steps described below.

* The first step uses the function rbind() to merge the X-train and X-test data into one; to merge the Y-train and Y-test data and well as merge the subject-train and subject-test.
* The second step extracts only the measurements on the mean and standard deviation for each measurement and names the extracted columns by their feature names from features.txt.
* Use descriptive activity names to name the activities in the data set.
* Appropriately label the data set with descriptive variable names. This is achieved by correctly naming columns with ambiguous names.
* The final step generates a tidy data of the average measures for each subject and activity type and outputs the result into averages\_data.txt.