**The Code Book assignment 3 Coursera Tidy dataset**

**Data Source**

The data linked to the course website represent data collected from the accelerometers from the Samsung Galaxy S smartphone. 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](http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones" \t "_blank)

Here is the data for this project:

[https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip](https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip" \t "_blank)

**Purpose of this project:** the purpose of his project is to demonstrate my ability to collect, work with and clean a data set. The goal is to prepare a tidy data that can be used for later analysis.

You should 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.

**Read Data**

In the run analysis.R script, each step is programmed

This file 'run\_analysis.R' contains all the code to perform the analyses described in the 5 steps. They can be launched in RStudio by just importing the file.

First the files were downloaded and unzipped

* File download, verification of the download to working directory
* File unzip verification of unzip the downloaded file.

**Data Analysis**

1. Merge the training and the test sets to create one data set.

* Reading Subject test tables
* Reading subject train tables
* Reading values on variables in test
* Reading values on variables in train
* Reading activity ID in test tables
* Reading activity ID in train tables
* Assign activity labels
* Merge Data sets training and test tables

1. Extracts only the measurements on the mean and standard deviation for each measurement.

* Create a vector of only mean and standard deviation; use the vector to subset dataset.
* The vector will originate a data set with the mean values and standard deviation of each table that we read

1. Changing column names in the dataset for a more descriptive names

* Crate a vector that clean the data set taking off () and use the descriptive activity names to name the activities in the data set.
* Apply it to Merged data set

1. Appropriately labels the set with descriptive activity names.

* Use a vector to combine test data of subject and activity, then add the respective names
* Bind with Merged dataset to add 2 additional columns in the left side of the table, the tables are assigned to subject and activity names.

1. Create a new data set with the average of each variable of each activity and each subject

* Create a vector that melts the data into a table
* Find the average of each variable and each activity
* Print second tidy data set to the working directory.