Codebook

* Source data:

Human Activity Recognition Using Smartphones Data Set

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

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>

* R script: run\_analysis.R

In the beginning the data are downloaded and unzipped and the necessary libraries for the analysis are loaded.

0. Initialize

0.1 Download data set

0.2 Unzip data set

0.3 Load libraries

The, the 5 steps as described in the course project are followed :

1. Merge training and test sets to create one data set
   1. Read files
      1. Read training set data, labels, subjects
      2. Read test set data, labels, subjects
      3. Read features data
      4. Read activity labels
   2. Make tidy data frames
      1. Make tidy activity labels: column names
      2. Make tidy train\_set: column names and binding with train\_labels and subject\_train
      3. Make tidy test\_set: column names and binding with test\_labels and subject\_test
   3. Merge tidy training and test sets
2. Extract only the measurements on the mean and standard deviation for each measurement
   1. Identify columns with mean and standard deviation
   2. Subset these measurements from tidy\_data\_set
   3. Add activity\_id and subject\_id
3. Use descriptive activity names to name the activities in the data set
4. Appropriately label the data set with descriptive variable names
5. From the data set in step 4, create a second, independent tidy data set
   1. with the average of each variable for each activity and each subject
   2. Make data set
   3. Write data set in txt format

* Variables :

train\_set : Training set (data from 70% of the volunteers)

train\_labels : Training labels

subject\_train : Each row identifies the subject who performed the activity for each window sample. Its range is from 1 to 30.

test\_set : Test set (data from 30% of the volunteers)

test\_labels : Test labels

subject\_test : Each row identifies the subject who performed the activity for each window sample. Its range is from 1 to 30.

features : List of all features

activity\_labels : Links the class labels with their activity name

(features are normalized and bounded within [-1,1])