

# CodeBook

## Introduction.

The data is downloaded from the following url: <https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip>

## Steps.

- 1) The zip file is unzip
- 2) The data is taken from the folder “UCI HAR Dataset”. Inside that folder, we read the files: ‘activity\_labels’ and ‘features’. We also go deeper inside that folder, and read the data from: ‘X\_test’ and ‘y\_test’ files. These files are inside the test folder. We also go into the train folder, and read the data from: ‘X\_train’ and ‘y\_train’ files.
- 3) After reading the above mentioned files, we process to rename the columns of the X\_train file with the information from the features file.
- 4) We create new columns in x\_train with the information from the y\_train and subject\_train files.
  - The Same procedure is done with the information from the test file
- 5) We proceed to merge the training and the test sets to create one data set.
- 6) We only the measurements on the mean and standard deviation for each measurement.
- 7) We assign descriptive activity names to name the activities in the data set created.
- 8) We proceed to provide appropriately labels to the data set with descriptive variable names
- 9) We create a second, independent tidy data set with the average of each variable for each activity and each subject.