The *run_analysis.R* script allows to analyze the content of two samples of measures from the Human Activity Recognition Smartphone Dataset. The two samples, test and train, include observations of different subjects included either in a test either in a train group.
Firstly the different columns are labeled according to the appropriate feature name. Then, the two datasets made available are merged in a unique dataset. Given that each row includes feature measures for a specific subject during a specific activity, an extra column including the combination "Subject_Activity" is created.
Then, only some columns of interest are isolated from the initial dataset, that are the ones including measures of the mean or the standard deviation. Using these columns a new dataset is obtained. In this the different observations (rows) are grouped according to different strata given by the combinations "Subject_Activity". Finally, using the *dplyr* package the mean for the values of each different feature is estimated within each different group. Those estimates are presented in the file final.txt

The *run_analysis.R* script can be broken down in **5 main subsections**:

A) The dataset set up

- This part includes all the steps necessary to load the data of interest from the different txt files

B) 1) the merge between the test and train dataset 4) the labeling of the different row and column variables

- In this section columns and rows are labeled taking from the text files including Subjects and Activities identifiers, that are *subject_test.txt* and *features.txt*
- Then the train and the test dataset after having been labeled are merged using the *rbind* command.

C) 3) The different activities are properly labeled within the dataset

D) 2) The columns including measures related to mean or standard deviation are identified within the dataset and a new dataset is created

E) 5) Finally in the new dataset observations are grouped according to the related "Subject-Activity" group and the mean within each of those is estimated. The resulting dataset is saved in *final.txt*.