**Introduction**

The script run\_analysis.Rperforms the 5 steps described in the course project's definition.

* First, all the similar data is merged using the rbind() function. By similar, we address those files having the same number of columns and referring to the same entities.
* Then, only those columns with the mean and standard deviation measures are taken from the whole dataset. After extracting these columns, they are given the correct names, taken from features.txt.
* As activity data is addressed with values 1:6, we take the activity names and IDs from activity\_labels.txt and they are substituted in the dataset.
* On the whole dataset, those columns with vague column names are corrected.
* Finally, we generate a new dataset with all the average measures for each subject and activity type (30 subjects \* 6 activities = 180 rows). The output file is called averages\_data.txt, and uploaded to this repository.

**Variables**

* x\_train, y\_train, x\_test, y\_test, subject\_train and subject\_test contain the data from the downloaded files.
* x\_data, y\_data and subject\_data merge the previous datasets to further analysis.
* features contains the correct names for the x\_data dataset, which are applied to the column names stored inmean\_and\_std\_features, a numeric vector used to extract the desired data.
* A similar approach is taken with activity names through the activities variable.
* all\_data merges x\_data, y\_data and subject\_data in a big dataset.
* Finally, averages\_data contains the relevant averages which will be later stored in a .txt file. ddply() from the plyr package is used to apply colMeans() and ease the development.