"GETTING AND CLEANNING DATA" COURSE PROJECT

CODE BOOK

The Variables

Subject

Description: Unique identifier of each person who carried out the experiment.
Class: Factor
Levels: 30 (range between 1 and 30 identifying the person)

Activity

Description: Describes the type of action the subjects carried out wearing a smartphone on the waist.
Class: Factor
Levels: 6 (walking, walking upstairs, walking downstairs, sitting, standing and laying)

Measurement

Description: Name of the feature measured for the subject and activity. The features selected for this database come from the accelerometer and gyroscope 3-axial raw, and are related to the mean and/or the standard deviation of them.
Class: Factor
Levels: 79

Value

Description: Value of the feature measured for each subject carrying out a particular activity.
Class: Numeric
Values: Numbers range between -1 and 1 with 9 decimals

The Data

As a result of the experiment, the dataset show, group by activity and subject, the mean of all the measurements carried out where a measurement of mean or standard deviation was wrote off.

Transformations

From the raw data, it has been performed the following transformations to get the final dataset:

1. Group both training and test sets for the activities, subjects and values of the features in three complete sets (one for activities, another for subjects and the last for the values).
2. Merged the activity labels names file with the activities values file to link the id of each activity to its name. After this the id has been deleted.
3. Select just the features variables from the interest of the project (mean and std)
4. Bind the activities, subjects and features selected in the last step 3) to create a complete dataset.
5. Group the data by activity and subject and calculate the mean for each feature variable.
6. Melt the dataset result from the last step 5) to create a long style dataset.