Codebook of Getdata project

Americo

30 giugno 2015

Overview

The dataset contains the average (calculated grouping by the subject and type of activity) of the measurements of mean and standard deviation of the datasets available here (https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip)

Origin

For information about the original dataset please take a look at the **readme.txt** file and the **features_info.txt** file that you can download with the datasets.

Pre-processing

Our dataset has been created merging the train and the test dataset, keeping only variables that were measurements of mean (not meanFreq) and standard Deviation and calculating the mean of them grouping by subject and activity. It is in a so called *long format*.

Variables

Our dataset is composed by 180 rows and 68 columns. They contains:

- Column 1-2: the activity and the subject whose average measurement where computed;

- Column 3-35: the average of the measurements of mean;

- Column 36-68: the average of the measurements of standard deviation.

The naming convention of the variables is quite simple: I added the word AVERAGE to the original name of the variable. For specification of the meanining of original variables please take a look at the *features_info.txt* file.

Notes

For further information about the dataset were constructed and the variables selected and calculated please see the README.md uploaded here (https://github.com/AmeCostantini/getdata)