Code Book

Felipe Haro

Sunday, January 11, 2015

This codebook describes the variables present in the tidy_data.txt file. It also describes the process the data took to transform from the raw data present in the samsung datasets to the tidy_data.txt file present in the respository.

Processing the data

The raw data can be found in the following link: https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip. That zip file contains training data and test data. To understand these datasets, there is a readme file if the raw data is downloaded. The readme file contains all the information that is not in this document. Both of those datasets were merged to create a dataset with all the subjects.
Among the 561 measurement present in the raw files, only the ones containing "mean()" or "std()" in the name were considered.
To create tidy_data.txt the data was grouped by activity and subject and the average value for all the measurements was calculated. In other words the average of all the measurements was calculated for each subject-activity combination.

Variables

subject
: Volunteer identifier
  : 1..30 .Unique identifier for the subject. Integer between 1 and 30

activity
: Activity identifier
  1. WALKING
  2. WALKING_UPSTAIRS
  3. WALKING_DOWNSTAIRS
  4. SITTING
  5. STANDING
  6. LAYING

The following variables are described in the features_info.txt file present in the raw files. In the data_tidy.txt file, the name of the variables were changed in the following way:
1. "-" was replaced by "_"
2. "()" was eliminated
The values for each of these variables corresponds to the average value of each combination subject-activity accross all the raw data.

tBodyAcc_mean_X

tBodyAcc_mean_Y

tBodyAcc_mean_Z

tGravityAcc_mean_X

tGravityAcc_mean_Y

tGravityAcc_mean_Z

tBodyAccJerk_mean_X

tBodyAccJerk_mean_Y

tBodyAccJerk_mean_Z

tBodyGyro_mean_X

tBodyGyro_mean_Y

tBodyGyro_mean_Z

tBodyGyroJerk_mean_X

tBodyGyroJerk_mean_Y

tBodyGyroJerk_mean_Z

tBodyAccMag_mean

tGravityAccMag_mean

tBodyAccJerkMag_mean

tBodyGyroMag_mean

tBodyGyroJerkMag_mean

fBodyAcc_mean_X

fBodyAcc_mean_Y

fBodyAcc_mean_Z

fBodyAccJerk_mean_X

fBodyAccJerk_mean_Y

fBodyAccJerk_mean_Z

fBodyGyro_mean_X

fBodyGyro_mean_Y

fBodyGyro_mean_Z

fBodyAccMag_mean

fBodyBodyAccJerkMag_mean

fBodyBodyGyroMag_mean

fBodyBodyGyroJerkMag_mean

tBodyAcc_std_X

tBodyAcc_std_Y

tBodyAcc_std_Z

tGravityAcc_std_X

tGravityAcc_std_Y

tGravityAcc_std_Z

tBodyAccJerk_std_X

tBodyAccJerk_std_Y

tBodyAccJerk_std_Z

tBodyGyro_std_X

tBodyGyro_std_Y

tBodyGyro_std_Z

tBodyGyroJerk_std_X

tBodyGyroJerk_std_Y

tBodyGyroJerk_std_Z

tBodyAccMag_std

tGravityAccMag_std

tBodyAccJerkMag_std

tBodyGyroMag_std

tBodyGyroJerkMag_std

fBodyAcc_std_X

fBodyAcc_std_Y

fBodyAcc_std_Z

fBodyAccJerk_std_X

fBodyAccJerk_std_Y

fBodyAccJerk_std_Z

fBodyGyro_std_X

fBodyGyro_std_Y

fBodyGyro_std_Z

fBodyAccMag_std

fBodyBodyAccJerkMag_std

fBodyBodyGyroMag_std

fBodyBodyGyroJerkMag_std