The Code Book for John Hopkins Getting and Cleaning Data course project

Karol Dorniak

Sunday, January 25, 2015

### 1. Overview

The objective of the project is to create a tidy.csv data set based on the raw data collected from the accelerometers from the Samsung Galaxy S smartphone.

To fulfil the project's requirements, the R script has been developed: run\_analysis.R, which purpose was to collect, transform and tidy raw data. The final results were saved in the tidy.csv file.

**See Readme.md file for an explanation of the run\_analysis.R script.**

### 2. Study Design

A full description of the raw data (collected from the accelerometers from the Samsung Galaxy S smartphone) is available at the site where it was obtained: <http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>

The raw data for the project is available here: <https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip>

The raw data was downloaded and extracted to "UCI HAR Dataset" folder. The run\_analysis.R script and "UCI HAR Dataset" folder are located in the same folder (working directory).

The following source files from the "UCI HAR Dataset" folder were used by run\_analysis.R to produce tidy.csv:

* UCI HAR Dataset/activity\_labels.txt - an activity's user friendly desciption of a given activity id
* UCI HAR Dataset/features.txt - list of all features (variables in X\_train.txt, X\_test.txt). The label of the column Nth of the X\_train.txt/X\_test.txt can be found in the row N of features.txt.
* UCI HAR Dataset/train/X\_train.txt - training data set
* UCI HAR Dataset/train/subject\_train.txt - each row identifies the subject who performed the activity for each window sample. The data in the Nth row of training data set relates to a subject identified in the Nth row of subject\_train.txt
* UCI HAR Dataset/train/y\_train.txt - data in the Nth row of the training data set relates to an activity id identified in the Nth row of the y\_train.txt. An activity's user friendly desciption of a given activity id can be found in the activity\_labels.txt
* UCI HAR Dataset/test/X\_test.txt - test data set
* UCI HAR Dataset/test/subject\_test.txt - each row identifies the subject who performed the activity for each window sample. The data in the Nth row of test data set relates to a subject identified in the Nth row of subject\_test.txt
* UCI HAR Dataset/test/y\_test.txt - data in the Nth row of the training data set relates to an activity id identified in the Nth row of the y\_test.txt. An activity's user friendly desciption of a given activity id can be found in the activity\_labels.txt

### 4. Code Book

The tidy.csv (the final output file produced by run\_analysis.R) contains the following variables:

* Activity - activity measured
* SubjectID - identifier of a subject
* fBodyAcc.mean...X - Fast Fourier Transform of tBodyAcc.mean...X
* fBodyAcc.mean...Y - Fast Fourier Transform of tBodyAcc.mean...Y
* fBodyAcc.mean...Z - Fast Fourier Transform of tBodyAcc.mean...Z
* fBodyAcc.std...X - Fast Fourier Transform of tBodyAcc.std...X
* fBodyAcc.std...Y - Fast Fourier Transform of tBodyAcc.std...Y
* fBodyAcc.std...Z - Fast Fourier Transform of tBodyAcc.std...Z
* fBodyAccJerk.mean...X - Fast Fourier Transform of tBodyAccJerk.mean...X
* fBodyAccJerk.mean...Y - Fast Fourier Transform of tBodyAccJerk.mean...Y
* fBodyAccJerk.mean...Z - Fast Fourier Transform of tBodyAccJerk.mean...Z
* fBodyAccJerk.std...X - Fast Fourier Transform of tBodyAccJerk.std...X
* fBodyAccJerk.std...Y - Fast Fourier Transform of tBodyAccJerk.std...Y
* fBodyAccJerk.std...Z - Fast Fourier Transform of tBodyAccJerk.std...Z
* fBodyAccMag.mean.. - Fast Fourier Transform of tBodyAccMag.mean..
* fBodyAccMag.std.. - Fast Fourier Transform of tBodyAccMag.std..
* fBodyBodyAccJerkMag.mean.. - Fast Fourier Transform of tBodyBodyAccJerkMag.mean..
* fBodyBodyAccJerkMag.std.. - Fast Fourier Transform of tBodyBodyAccJerkMag.std..
* fBodyBodyGyroJerkMag.mean.. - Fast Fourier Transform of tBodyBodyGyroJerkMag.mean..
* fBodyBodyGyroJerkMag.std.. - Fast Fourier Transform of tBodyBodyGyroJerkMag.std..
* fBodyBodyGyroMag.mean.. - Fast Fourier Transform of tBodyBodyGyroMag.mean..
* fBodyBodyGyroMag.std.. - Fast Fourier Transform of tBodyBodyGyroMag.std..
* fBodyGyro.mean...X - Fast Fourier Transform of tBodyGyro.mean...X
* fBodyGyro.mean...Y - Fast Fourier Transform of tBodyGyro.mean...Y
* fBodyGyro.mean...Z - Fast Fourier Transform of tBodyGyro.mean...Z
* fBodyGyro.std...X - Fast Fourier Transform of tBodyGyro.std...X
* fBodyGyro.std...Y - Fast Fourier Transform of tBodyGyro.std...Y
* fBodyGyro.std...Z - Fast Fourier Transform of tBodyGyro.std...Z
* tBodyAcc.mean...X - mean of body acceleration signal - X axis
* tBodyAcc.mean...Y - mean of body acceleration signal - y axis
* tBodyAcc.mean...Z - mean of body acceleration signal - Z axis
* tBodyAcc.std...X- standard deviation of body acceleration signal - X axis
* tBodyAcc.std...Y- standard deviation of body acceleration signal - Y axis
* tBodyAcc.std...Z- standard deviation of body acceleration signal - Z axis
* tBodyAccJerk.mean...X - mean of body linear acceleration signal - X axis
* tBodyAccJerk.mean...Y - mean of body linear acceleration signal - Y axis
* tBodyAccJerk.mean...Z - mean of body linear acceleration signal - Z axis
* tBodyAccJerk.std...X - standard deviation of body linear acceleration signal - X axis
* tBodyAccJerk.std...Y - standard deviation of body linear acceleration signal - Y axis
* tBodyAccJerk.std...Z - standard deviation of body linear acceleration signal - Z axis
* tBodyAccJerkMag.mean.. - mean of the magnitude of the body linear acceleration signal
* tBodyAccJerkMag.std.. - standard deviation of the magitude of the body linear acceleration signal
* tBodyAccMag.mean.. - mean of the magnitude of body acceleration signal
* tBodyAccMag.std.. - standard deviation of the magnitude of body acceleration signal
* tBodyGyro.mean...X - mean of the body gyro signal - X Axis
* tBodyGyro.mean...Y - mean of the body gyro sugnal - Y Axis
* tBodyGyro.mean...Z - mean of the body gyro sugnal - Z Axis
* tBodyGyro.std...X - standard deviation of the body gyro sugnal - X Axis
* tBodyGyro.std...Y - standard deviation of the body gyro sugnal - Y Axis
* tBodyGyro.std...Z - standard deviation of the body gyro sugnal - Z Axis
* tBodyGyroJerk.mean...X - mean of the body angular velocity - X axis
* tBodyGyroJerk.mean...Y - mean of the body angular velocity - Y axis
* tBodyGyroJerk.mean...Z - mean of the body angular velocity - Z axis
* tBodyGyroJerk.std...X - standard deviation of the body angular velocity - X axis
* tBodyGyroJerk.std...Y - standard deviation of the body angular velocity - Y axis
* tBodyGyroJerk.std...Z - standard deviation of the body angular velocity - Z axis
* tBodyGyroJerkMag.mean.. - mean of the magnitude of the body angular velocity
* tBodyGyroJerkMag.std.. - standard deviation of the magnitude of the body angular velocity
* tBodyGyroMag.mean.. - magnitude of mean of the body gyro signal
* tBodyGyroMag.std.. - magnitude of standard deviation of the body gyro signal
* tGravityAcc.mean...X - mean of gravity acceleration - X axis
* tGravityAcc.mean...Y - mean of gravity acceleration - Y axis
* tGravityAcc.mean...Z - mean of gravity acceleration - Z axis
* tGravityAcc.std...X - standard deviation of gravity acceleration - X axis
* tGravityAcc.std...Y - standard deviation of gravity acceleration - Y axis
* tGravityAcc.std...Z - standard deviation of gravity acceleration - Z axis
* tGravityAccMag.mean.. - magnitude of mean of gravity acceleration
* tGravityAccMag.std.. - magnitude of standard deviation of gravity acceleration