CodeBook

Me

27/05/2020

#This is the code book for the project

##How to get to the tidydata.txt: 1. Download data from the link below and unzip it into working directory of R Studio. 2. Execute the R script.

About the source data

The source data are from the Human Activity Recognition Using Smartphones Data Set. A full description is available at the site where the data was obtained: http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones Here are the data for the project: https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip

About R script

File with R code "run_analysis.R" performs the 5 following steps (in accordance assigned task of course work):
1. Reading in the files and merging the training and the test sets to create one data set.
1.1 Reading files
1.1.1 Reading trainings tables
1.1.2 Reading testing tables
1.1.3 Reading feature vector
1.1.4 Reading activity labels
1.2 Assigning variable names
1.3 Merging all data in one set
2. Extracting only the measurements on the mean and standard deviation for each measurement
2.1 Reading variable names
2.2 Create vector for defining ID, mean and standard deviation
2.3 Making nessesary subset from merged data set
3. Using descriptive activity names to name the activities in the data set
4. Appropriately labeling the data set with descriptive variable names
5. Creating a second, independent tidy data set with the average of each variable for each activity and each subject
5.1 Making second tidy data set
5.2 Writing second tidy data set in txt file

The code assumes all the data is present in the same folder, un-compressed and without names altered.

About variables:

- `x_train`, `y_train`, `x_tst`, `y_tst`, `sub_train` and `sub_tst` contain the data from the downloaded files.

- `all_train`, `all_test` and `finaldata` merge the previous datasets to further analysis.
- `features` contains the correct names for the `finaldata` dataset, which are applied to the column names stored in.