

README

Mark Edney

19/09/2020

R Code

The run_analysis code follows this procedure:

1. Load the dplyr packages
2. Load the different datafiles into the system as dataframes
3. Merge the train and test dataframes together
4. Load the featurelist and activity labels into the system as dataframes
5. Merge the activity list for the train and test data
6. Sort the merged data from for varialbes that only include std and mean
7. Change the variables of the merged data with the featurelist df
8. Amends the activity date to the merged data with mutate and changes the activity to the description with the activity label df
9. Amends the subject data to the merged df
10. Creates a second df grouped by activity and then by subject
11. Summarise the second df with the summarise_all by mean

Code book

train_x- data from the training group

train_y- activity data for the training group

test_x- data from the test group

test_y- activity data for the test group

subject_train- subject data for the training group

subject_test- subject data for the test group

datamerge- the mergeder data frame

featurelist- labels for the feature data

activitylabel- labels for the activity data

activity- merged activity data

sortfeatures- logical vector for sorting df by std and mean

datasecond- second dataframe grouped by activty and subject and summarised