

Codebook for run_analysis.R

Created for Coursera Getting and Cleaning Data course project
created by Adrienne Lehnert: July 23, 2015

1. Load library needed to run script (dplyr)
2. Read and find the "mean" and "std" columns desired
 - a. read the "features.txt" file to get the names of variables and store in "features"
 - b. -read the names of activities from "activity_labels.txt" and store in "activity_labels"
 - c. -search for "mean" and "std" in labels and collect/sort column numbers to store in "cols"
3. Read training data, labels and subject information from "X_train.txt", "y_train.txt", and "subject_train" and store in "TrainData", "TrainLabels", and "TrainSubject"
4. Put training data together into single data frame "TrainData"
 - a. "features" are column names
 - b. add in column with activity labels ("TrainLabels")
 - c. add in column with subject information ("TrainSubject")
5. Repeat steps 3-4 with testing data to get the variable "TestData"
6. Merge "TrainData" and "TestData" into "Data"
7. Reassign the numbers in Data\$Activity to activity labels
8. Create tidy data set ("tidyData") that takes the average of every variable for every activity and subject
9. Write "tidyData" to text file "Tidy_Data.txt"