README for run_analysis.R

JRB

README.md: this md file describes how to use the run_analysis.R script and its key output

---

- Pre-requisites: working internet connection and functional R install
- What the script does :
  1. Download, label and merge the train and test data sets
     - Create the data directory (if not already created) as a sub directory of the current working dir
     - Download the zip file from https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip using the curl method (works on the mac)
     - Unzip the downloaded file in the data subdirectory
     - Load the train and test data sets into R
     - Load the subject and activity id and attach them to the train and test datasets respectively
     - Merge the train and test datasets into a single R dataframe
     - Appropriately label the data sets column using the feature lists provided in the zip file
  2. Extract the values for mean and standard deviation and rename the variables to more explicit names (see codebook.md for a full description)
     - Extract the values for mean and standard dev using a standard expression
     - Make the variable explicit according to the following rules (yes names are longer but easy to type and understand)
       * Characters only (no symbols)
       * Lower case variable names (eliminate typing issue)
       * Spell out t abbreviation as time and f as frequency
       * std becomes stddev for standard deviation
       * acc becomes acceleration
       * gyro becomes gyroscope (to avoid any confusion with a greek sandwich)
       * mag becomes magnitude
     - Ta-da... here's our first tidy dataset called tidy1 and saved in the data sub directory as tidy1.txt
  3. Create the derived dataset
     1. Attempt to load dplyr, if not avaible install it form CRAN
     2. Transform the dataframe into table
     3. Use dplyr to group by the dataset by subject and activity and summarize the values using mean
     4. Export the derived dataset to the data subdirectory as summarybysubjectactivity.txt
- Outputs:
  - data/tidy1.txt: first tidy data dest
  - data/summarybysubjectactivity.txt: derived data set of summarized data