**Peer-graded Assignment: Getting and Cleaning Data Course Project (Coursera)**

**Instructions for this assignment were:**

The purpose of this project is to demonstrate your ability to collect, work with, and clean a data set. Review criterialess

1. The submitted data set is tidy.
2. The Github repo contains the required scripts.
3. GitHub contains a code book that modifies and updates the available codebooks with the data to indicate all the variables and summaries calculated, along with units, and any other relevant information.
4. The README that explains the analysis files is clear and understandable.
5. The work submitted for this project is the work of the student who submitted it

=============================================================================================================

**The script required to generate the tidy data set is called run\_analysis.R.**

It has does several activities to complete the assignment.

1. It loads the libraries, gathers the datasets from the websites and unzips the files. All of these files are read into data.frames. This including pulling the variable names from the features.txt file and the activities names from activity\_labels.txt.
2. Next the script merges the training and the test sets to create one data set. This includes the X, y and subject files. X - contains the values for the various measurements y - contains the activity code (1-6) subject - contains the subject id
3. After that the script extracts only the measurements on the mean and standard deviation for each measurement. This was accomplished by using select to pull any column that had the word 'mean' or 'std' in the variable name.
4. A small function was created to mutate the activity code to the correct activity label.
5. Appropriately labels the data set with descriptive variable names. This task was accomplished by converting all labels to lower case and removing any non-alpha characters. The variable names were very descriptive in reference to the action that was measured. I chose not to decompose the X/Y/Z from the variables because these were individual measurements representing the 3-axial acceleration and velocity.
6. From the data set in step 4, creates a second, independent tidy data set with the average of each variable for each activity and each subject. This was accomplished by using group\_by and summerize\_all to provide the mean for each activity/subject group.

The tidy data set is available in this repository labelled tidy\_data.txt.

If you choose to download to view the data that can be accomplished by: data <- read.table(datafil, header = TRUE) # where datafil is the filename & location View(data)