**Getting-and-Cleaning-Data-Week-4-Assignment-Million Nzvuwu**

This repo was created to finish the assignment week 4 course-Getting and Cleaning data

1. Download and unzip the data file into your R working directory.

2. Download the R source code into your R working directory.

3. Finally, execute R source code to generate tidy data file.

**Data description**

The variables in the data X are sensor signals measured with waist-mounted smartphone from 30 subjects. The variable in the data Y indicates activity type the subjects performed during recording.

**Code explanation**

The code combined training dataset and test dataset, and extracted partial variables to create another dataset with the averages of each variable for each activity.

**New dataset**

The new generated dataset contained variables calculated based on the mean and standard deviation. Each row of the dataset is an average of each activity type for all subjects.

**The code was written based on the instruction of this assignment**

Read training and test dataset into R environment. Read variable names into R environment. Read subject index into R environment.

1. Merges the training and the test sets data to create one data set. Use command rbind to combine training and test set
2. Extracts only the measurements on the mean and standard deviation for each measurement. Use grep command to get column indexes for variable name contains "mean()" or "std()"
3. Uses descriptive activity names to name the activities in the data set Convert activity labels to characters and add a new column as factor
4. Appropriately labels the data set with descriptive variable names. Give the selected descriptive names to variable columns
5. 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. Use pipeline command to create a new tidy dataset with command group\_by and summarize\_each in dplyr package