## General Notes concerning using this routine and the requirements of this effort.

# This routine wants to see a data directory called /data in the working directory
# its probably best to set the working directory to a directory of your choice then place
# the jhu data set in a directory called /data
# assuming the data is in the correct place this routine will work and produce
# three files. A text file output, a csv file output and a csv of the output std and mean values

# this data did not require a great deal of cleaning. I used na.omit to remove any
# na's from the data. The std function ignores them and the mean can be made to ignore
# them but I wanted to make sure that the data was tidy so I used na.omit()

# the only variables that I changed were the column names. I used the gsub function
# to make the changes to something more meaningful

# you may or may not have to set the working directory, your choice as long as the data
# is located in a directory called /data located directory underneath the current directory

## Output file names are:

course3_wk4_project_DF_output.txt
course3_wk4_project_DF_output.csv
course3_wk4_project_stats_output.csv

## Learning Requirements are:

1. The submitted data set is tidy. – It was very clean to start with but I made sure that it was free of na's by using na.omit()

2. The Github repo contains the required scripts. – There is only one script

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. - The codebook will describe the changes that I made to the original data

4. The README that explains the analysis files is clear and understandable. – There will be a README.md and a README.docx file

5. The work submitted for this project is the work of the student who submitted it. – Of course!

You should create one R script called run_analysis.R that does the following.

1. Merges the training and the test sets to create one data set. – Done and verifiable via the code

2. Extracts only the measurements on the mean and standard deviation for each measurement. – Done and verifiable via the code

3. Uses descriptive activity names to name the activities in the data set – I changed the names of the columns to reflect that its people that we are measuring

4. Appropriately labels the data set with descriptive variable names. – Same as #3

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. – Done as an output csv file. The name of this file is course3_wk4_project_stats_output.csv