**JHU Data Science Specialization - Getting and Cleaning Data Project**

This repository contains my Getting and Cleaning Data project submission. Here you will find the script used to create the data set along with a file containing the associated data cleaning functions and a file detailing the entire process of creating this data set as well as information about the variables in the data set.

**Running the script**

Make sure that the following packages have been installed and loaded into the current R session:

* tidyverse
* magrittr
* zeallot

I have divided the script into two files:

* run\_analysis.R: this file is the script for performing the actual data wrangling.
* data\_wrangling\_functions.R: this file contains the functions called in the run\_analysis.R script.

The script uses data from the *Human Activity Recognition Using Smartphones Data Set* that can be found [here.](http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones) The data is stored in a zip file which, when unpacked, will contain a directory called *UCI HAR Dataset* along with some additional files.

This is important to note as the script assumes that this directory is a subdirectory of the current directory. In order to run the script copy the unpacked *UCI HAR Dataset* directory to the current directory. From here, make sure that the R files listed above are in the current working directory before setting this directory to be the current R working environment. You can then run the run\_analysis.R script by sourcing it and it will produce a tidy data frame (or more accurately, a tibble) called smartphone\_data. More information on the transformations applied to the data as well as on the variables contained in smartphone\_data can be found in the CodeBook.md file