## Peer Assessments (https://class.coursera.org/getdata-005/human\_grading/)

/ Getting and Cleaning Data Course Project

<u>Help (https://class.coursera.org/getdata-005/help/peergrading?url=https%3A%2F%2Fclass.coursera.org%2Fgetdata-005%2Fhuman\_grading%2Fview%2Fcourses%2F972582%2Fassessments%2F3%2Fsubmissions)</u>

due in 1wk 1d	
Submission Phase	
1. Do assignment ☐ (/getdata-005/human_grading/view/courses/972582/assessments/3/submissions)	
Evaluation Phase	
2. Evaluate peers	ts)
Results Phase	
3. See results	
In accordance with the Honor Code, I certify that my answers here are my own work, and that I have appropriately acknowledged all external sources (if any) that were used in this work.  Save draft  Submit for grading	

The purpose of this project is to demonstrate your ability to collect, work with, and clean a data set. The goal is to prepare tidy data that can be used for later analysis. You will be graded by your peers on a series of yes/no questions related to the project. You will be required to submit: 1) a tidy data set as described below, 2) a link to a Github repository with your script for performing the analysis, and 3) a code book that describes the variables, the data, and any transformations or work that you performed to clean up the data called CodeBook.md. You should also include a README.md in the repo with your scripts. This repo explains how all of the scripts work and how they are connected.

One of the most exciting areas in all of data science right now is wearable computing - see for example this article (http://www.insideactivitytracking.com/data-science-activity-tracking-and-the-battle-for-the-worlds-top-sports-brand/). Companies like Fitbit, Nike, and Jawbone Up are racing to develop the most advanced algorithms to attract new users. The data linked to from the course website represent data collected from the accelerometers from the Samsung Galaxy S smartphone. A full description is available at the site where the data was obtained:

http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones (http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones)

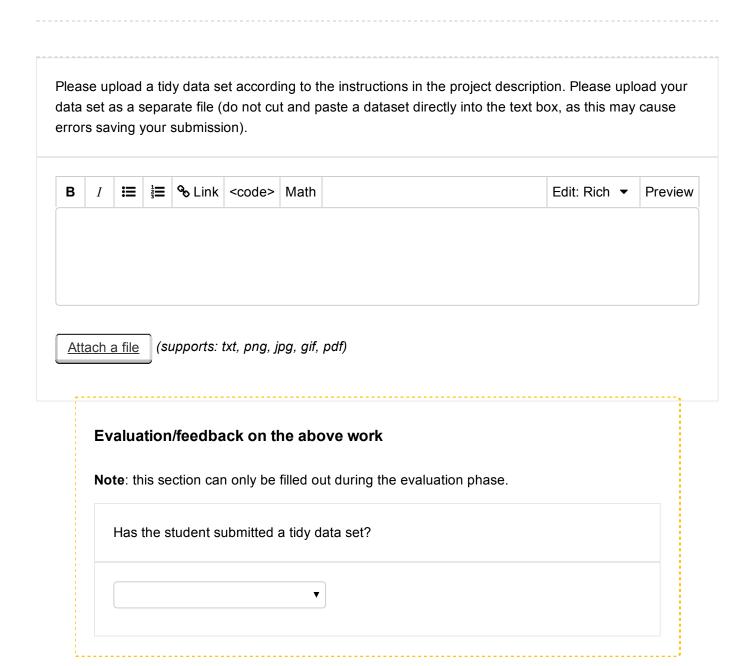
Here are the data for the project:

https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip (https://d396qusza40orc.cloudfront.net/getdata%2Fprojectfiles%2FUCI%20HAR%20Dataset.zip)

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.
- 2. Extracts only the measurements on the mean and standard deviation for each measurement.
- 3. Uses descriptive activity names to name the activities in the data set
- 4. Appropriately labels the data set with descriptive variable names.
- 5. Creates a second, independent tidy data set with the average of each variable for each activity and each subject.

Good luck!



Please submit a link to a Github repo with the code for performing your analysis. The code should have a

file run\_analysis.R in the main directory that can be run as long as the Samsung data is in your working directory. The output should be the tidy data set you submitted for part 1. You should include a README.md in the repo describing how the script works and the code book describing the variables. В Edit: Rich ▼ Preview Evaluation/feedback on the above work **Note**: this section can only be filled out during the evaluation phase. Did the student submit a Github repo with the required scripts? Was code book submitted to GitHub that modifies and updates the codebooks available to you with the data to indicate all the variables and summaries you calculated, along with units, and any other relevant information? I was able to follow the README in the directory that explained what the analysis files did.

## Overall evaluation/feedback

**Note**: this section can only be filled out during the evaluation phase.

	▼
the wor	use the space below to provide constructive feedback to the student who submitted rk. Point out the submission's strengths as well as areas in need of improvement. You so use this space to explain your grading decisions.
	You've written 0 words
	rdance with the Honor Code, I certify that my answers here are my own work, and tha