Peer review assignment. In order to get credit you must fulfill the following components:
1. Submit your project to gradescope on time
2. Also post on github and email Professor Huang the link to the repository, also on time.
3. For each of your two peers make a copy of this spreadsheet and fill in using instructions below. You get credit if you follow all instructions. This includes providing **specific** feedback on **every category,** and checking that their **code runs.**
4. You will submit on Gradescope two PDFs one for each peer (see notes below about converting this to PDF)
5. *Also* upload your PDF to **this link** . Name your PDF like "row2_peer17.pdf" if you reviewed the submission on row2 and your peer number is 17.

Instructions: For each question give a score (column C) using the scale below, and for each score, provide feedback on what they did well or could improve (1 sentence per question is fine). Important: for each question in column E, if they got full credit, be **specific** what they did to get full credit. If they get less than full, again be **specific** what they need to change to get full credit.

This includes running the code for code-related questions. (the ones not grayed out). If the code did not work or if you had to change something, just say what problem you saw in column D.

When complete, download as PDF (File --> Download --> PDF). **Before you submit open your PDF to see if that all columns for a row are on the same page (if you drag the columns to be wider, they might not be on the same page, and I will have to ask you to resubmit).**

Scale for column C:
2 - Satisfactory: This is approximately equivalent to "B" work, where the thought process is logical and justified for most answers, even if not all answers are correct.
1 - Progressing: This work may not be correct and not all answers are logically justified, but the work shows thoughtful engagement with the assignment.
0 - Incomplete: This would describe missing or hastily performed work that shows little effort.

	Assignment requirements	Score (0, 1, 2)	Did code work/ what problems arose	Explanation of score and suggestions/feedback
Part 2	2.1: Include the contents of Part 2. You must address all comments where you lost points. You do not need to get a perfect score on Part 2 to get a Satisfactory grade on this project but you must work to address all comments.	2		Part 2 is included, no need for comments or addressing them for the review
	2.2: At the end of Part 2, include a summary of what you improved from the previous submission.	2		Same as above, not needed at this moment

3.1 Load and clean data	Describe steps someone can take to obtain the data files you are working with. For example, provide the link to a the paper's supplementary Excel file, and explain that you opened it in excel and then saved it as a CSV file called "frogs.csv" in the same directory as the notebook (or whatever, as long as what you instruct them to do matches what your code is so your code works). ONLY IN RARE CASES: Only if you check with Dr. Melamed first, and you have some data that requires special permission to access, or some other big hassle to access, you can do this some alternate way. Your data should be directly downloadable from the source and not your personal shared folder.	2	Works	There is a link to data that doesn't work right above the main link, which you might want to remove since it could potentially cause confusion.
	Provide code to read in all relevant data files into data frames. Explain your code and why you did it that way. Show the "head" (first few lines/rows/columns) of each data frame.	2	Works	This is good, correctly reads in.
	If any cleaning steps were needed at this point, explain these cleaning steps. Otherwise, explain how you checked that the data frames were suitable for the further analyses.	2	Works	Not sure if doing pd.DataFrame on the transposed lipids dataframe is necessary, since read_csv already reads in the data as a dataframe. Regardless, no issues.
3.2: Describe data numerically	Provide code to obtain the shape of the data files. Describe how this shape relates to the number of observations and the number of features. Be precise, such as "This data frame has 6000 rows which is the number 500 mice times the 2 treatments times the 6 time points per treatment".	2	Works	I'm pretty sure this is good, but since the observations names start with C for some of them and P for the others, are there groups the observations fall into?
	Feature 1: Explain what you expect the "describe" function would output, based on your understanding of that features. How many observations have a recorded value of that feature and what is the average across observations?	2		Yeah this works, count and averages have good reasoning for why they were guessed to be that
	Feature 1: Run the "describe" function and compare the results to what you predicted.	2	Works	Thorough explanation, good.

	Feature 2: Explain what you expect the "describe" function would output, based on your understanding of that features. How many observations have a recorded value of that feature and what is the average across observations?	2		Same as feature 1
	Feature 2: Run the "describe" function and compare the results to what you predicted.	2	Works	Same as feature 1
3.3 Visualizations.	**Visualization 1:** Describe what kind of visualization you want to make, why this is appropriate for this feature and data set, and how the visualization will provide insight into the data.	2		Good reasoning, you could also do a histogram and make a few bins, plotting the counts for each bin. The reason I say this is: making 73 bars might slightly lessen the readability of the plot.
	Visualization 1: Provide code and explain your code to make the visualization.	2	Works	This is good
	Visualization 1: Interpret the visualization: compare it to the "describe" function output from 3.2, and explain what insight into the data you can make with the visualization	2		This covers pretty much everything in the describe function
	Visualization 1: Describe how your visualization relates to one of the hypotheses or figures from the paper.	1		I cant tell which ones are the control and PTEN groups, perhaps because of the 73 bars. You could perhaps use a hue or an entirely separate barplot.
	Visualization 2: Describe what kind of visualization you want to make, why this is appropriate for this feature and data set, and how the visualization will provide insight into the data.	2		Same as above
	Visualization 2: Provide code and explain your code to make the visualization.	2	Works	^
	Visualization 2: Interpret the visualization: compare it to the "describe" function output from 3.2, and explain what insight into the data you can make with the visualization	2		^
	Visualization 2: Describe how your visualization relates to one of the hypotheses or figures from the paper.	1		^