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).**

	Assignment requirements	**Score (0, 1, 2)**	**Did code work/ what problems arose**	**Explanation of score and suggestions/feedback**
	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	N/A		N/A

Part 2	2.2: At the end of Part 2, include a summary of what you improved from the previous submission.	N/A		N/A
	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	1	There was some problem with the code in the beginning. The instructions I at first thought was to save the data as 'jupyter' but now reading back it is to save the data in a folder called 'jupyter' where the notebook and data will be located. When saving the dataset with the default name dataset also caused a problem, because the code is reading a csv that is named dataset.txt.csv. The rest of the code worked with no problem after adjusting the dataset name/code.	A score of 1 was given because the instructions provided a link to the dataset from the paper and instructed to download the file and slight typo from "placed in the a folder which I called jupyter". I mistook the directions as naming the dataset juypter, but it was the folder to save the dataset and notebook, but there was no instructions on what to name the dataset or save it as the default when downloaded. I would also include to name the dataset either the default "dataset.csv" or match your code "dataset.txt.csv". Was there a reason to name the csv with the txt in the title?

3.1 Load and clean data	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	Dataset was not named correctly from previous instructions, so the code to read in the files did not work until file was named correctly. The "head" of the data frame worked and showed the first 5 rows/columns.	A score of 2 was given because from the previous instructions the dataset was not named correctly therefore the code did not work to read in the relevant data files into data frames. One suggestion is to indicate what the dataset should be saved as in the previous step, but the code is correct and explained I would not change anything there.
	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.	0	No code for cleaning needed	A score of 0 was given because there is no explanation on how the data frame was checked and suitable for further analysis because no cleaning steps were needed, the columns of the dataset was separated by semicolons when the dataset was read in from the last step. My suggestion is to include the cleaning/checking of the dataset, it can be briefly explained saying this dataset did not need any cleaning.

	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	Code worked and the output was the shape of the dataframe, 862 rows by 18 columns.	A score of 2 was given because the code worked and the shape was explained in regards to the observations and features, observations being the collared flycatchers noticing mock competitors, the rows and each column corresponds to a feature. I do not have any suggestions.
	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?	1		A score of 1 was given because the feature was described and predicted that the majority of observations would be for Nestling. There was no explaination saying this feature would not give a describe output for the average across the observations. I would explain that this feature would not have an average given by the describe function.
3.2: Describe data numerically	**Feature 1**: Run the "describe" function and compare the results to what you predicted.	2	Code worked and the output described the feature breeding stage from the dataset.	A score of 2 was given because the code worked and the describe function gave an output for the breeding stage feature. They also compared their prediction with the output from the describe function and explained that their predictions were correct. I would not change anything from this code and description.

	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?	1		A score of 1 was given because the contact feature prediction for describe function was given, but it looks like they forgot to finish their prediction and only indicated which feature would have the higher frequency. One suggestion would be to finish the prediction for describing the contact feature, the mean, etc.
	Feature 2: Run the "describe" function and compare the results to what you predicted.	1	Code worked and the output described the feature contact? from the dataset.	A score of 1 was given because the code worked and the results of the describe function were described but there was no clear comparison of the results that were predicted. I think because the prediction was incomplete the comparison of the predicted and describe output was not clear and complete.
	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		A score of 2 was given because the description of the type of visualization was complete and appropriate for frequency providing insight into the different breeding stages. I do not have any suggestion for the description for this visualization.

Visualization 1: Provide code and explain your code to make the visualization.	2	Code worked and the visualization output was a barplot for the feature breeding stages.	A score of 2 was given because the code worked and the output was a visualization of a barplot for the breeding stages feature compared against the 'contact?' feature. I would not change any of the code or explanation of the code provided to create the barplot.
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	1		A score of 1 was given because the visualization was interpreted and they explained the majority of the observations in this plot was for the incubation breeding stage instead of nestling. One suggestion is to explain that when compared against the other feature contact the majority of observations was no longer in the nestlings stage of breeding which was seen in the describe function output from 3.2.

3.3 Visualizations.	**Visualization 1**: Describe how your visualization relates to one of the hypotheses or figures from the paper.	2		A score of 2 was given because they explained how the visualization, the barplot for the breeding stages related to the hypothesis given by the authors for displaying different levels of aggression during the breeding stages. I would not change the description given for the visualization relating to the hypothesis.
	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		A score of 2 was given because they explained what type of visualization they would make for the 'contact?' feature and that it would be appropriate to see the frequency of the feature. I would explain a little more on how the frequency seen in the histogram would provide more insight into the data.

Visualization 2: Provide code and explain your code to make the visualization.	1	Code worked and the output was the histogram based off of the feature 'contact?'	A score of 1 was given because the code worked and the output was a visualization for a histogram showing the contact feature. I would not change any of the code, but I would provide the explanation for the code you used to create the histogram.
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		A score of 2 was given because the visualization was interpreted and they explained that the majority of the observations in this plot were seen in no contact. One suggestion is to also explain that when compared against the describe function the majority of observations was seen with no contact and that only 9.6% had contact observed.
Visualization 2: Describe how your visualization relates to one of the hypotheses or figures from the paper.	1		A score of 1 was given because they described how the visualization, a histogram for the contact feature goes along with the paper. I would explain how it goes along with the paper in terms of their hypothesis or relating it to another figure/table found in the paper that uses similar data.