BIOL 5860/6860 – Bioinformatics – GP Step 6   
Step 5 analysis evaluation   
Thank you for your thoughtful comments and suggestions   
Group evaluated: Salmonella Project

**Please provide a thorough review based on the group’s GitHub Repository from Step 5.**   
How useful is the main Readme.md file in aiding your understanding of the group’s analysis this   
semester?

> At first glance the main readme looks very well put together, but upon further investigation there are things that can be improved upon. For starters, the way that the introduction is set up with the description of the project is very hard to read. I would suggest editing the introduction with markdown to allow for an easier read for the viewer. As it is set up right now you must scroll back and forth to view the entire introduction. The steps of the pipeline are also not linked throughout so this would be also beneficial in helping me understand the analysis and the process that has been going on this semester. The Quality assessment section also seems very thrown together and the moment and could use some organization.

Based on the rubric for the final GitHub for the class, what elements need the most improvement for the final product?

> **The main readme**: My comments for the main readme can be found above my only comment here from the rubric would be to add some scientific references.

>**Graphical Quality**: The only graphs I could find were the ones on step one. I would suggest using R studio to produce histograms for percent mapped, percent duplicated, and overall coverage. The graphs on step one seemed to be very well produced. I would continue this effort to improve the quality of the other sections of the GitHub to allow for a more user-friendly experience when looking at the GitHub.

>**Separate Readme Pages**:

* Step 1: The GitHub page for step one is very barren and I had to dig the find the actual quality report/assessment. I would suggest possible writing up snippets under each section that is laid out and putting the information from the report onto the main readme for the page instead of having it in a link.
* Step 2: I like that the trimming parameters were included in this page. The rest of the page is still unfinished. I would give the same suggestions that I gave for step 1 when assembling this page on the GitHub.
* Step 3-5: these pages still need to be completed. I would just work towards completeness and getting the graphs done to allow the GitHub to have a more finished feel.

>**Data Analysis for Step 5: n/a**

**> Overall Organization:** I would make the different pages easier to find and link the actual individual steps to the main pages to make a clear path of what has been done over the course of this semester.

How detailed are the scripts provided? Do the comments help you to understand what they did?

>The scripts were hard to find when exploring the GitHub. I would also add comments throughout the scripts because sometimes there are just large chunks of script with nothing denoting what the script is trying to do. It would have been helpful especially on the sections where you had errors so that some reviewing might have been able to help.

How useful are the graphics that have been included? Are they effective?

> The graphs that were present for step 1 clarified the sequences and the need to trim. I would have liked to see more graphs throughout but I assume that the group is still working on perfecting the graphs before adding them into the GitHub.

Based on the Step 5 README page for this group, what are some analyses that the group did not do that you’d like to see added? Or perhaps some ways of visualizing their current results? Feel free to sketch them out below.

> There was nothing on the step 5 readme and it appears this group has not finished yet to be at a point of the step 5. I would love to see a map of where the different samples were found and how the sequences vary from place to place. I think it would be a cool conclusion to the project. I believe there was a diagram/map on the slide presented in class that could be cool to include within the GitHub results as a comparison from start to finish.

**On a scale of 1-10 with 5=average, 10=truly exceptional**

Rank the organization of the group’s GitHub repository: 6  
Rank the group’s overall analysis and results for Step 5: 1 (there was nothing really for me to look at so the 1 is there as a place holder because I am sure the final page will be beautiful when complete.)