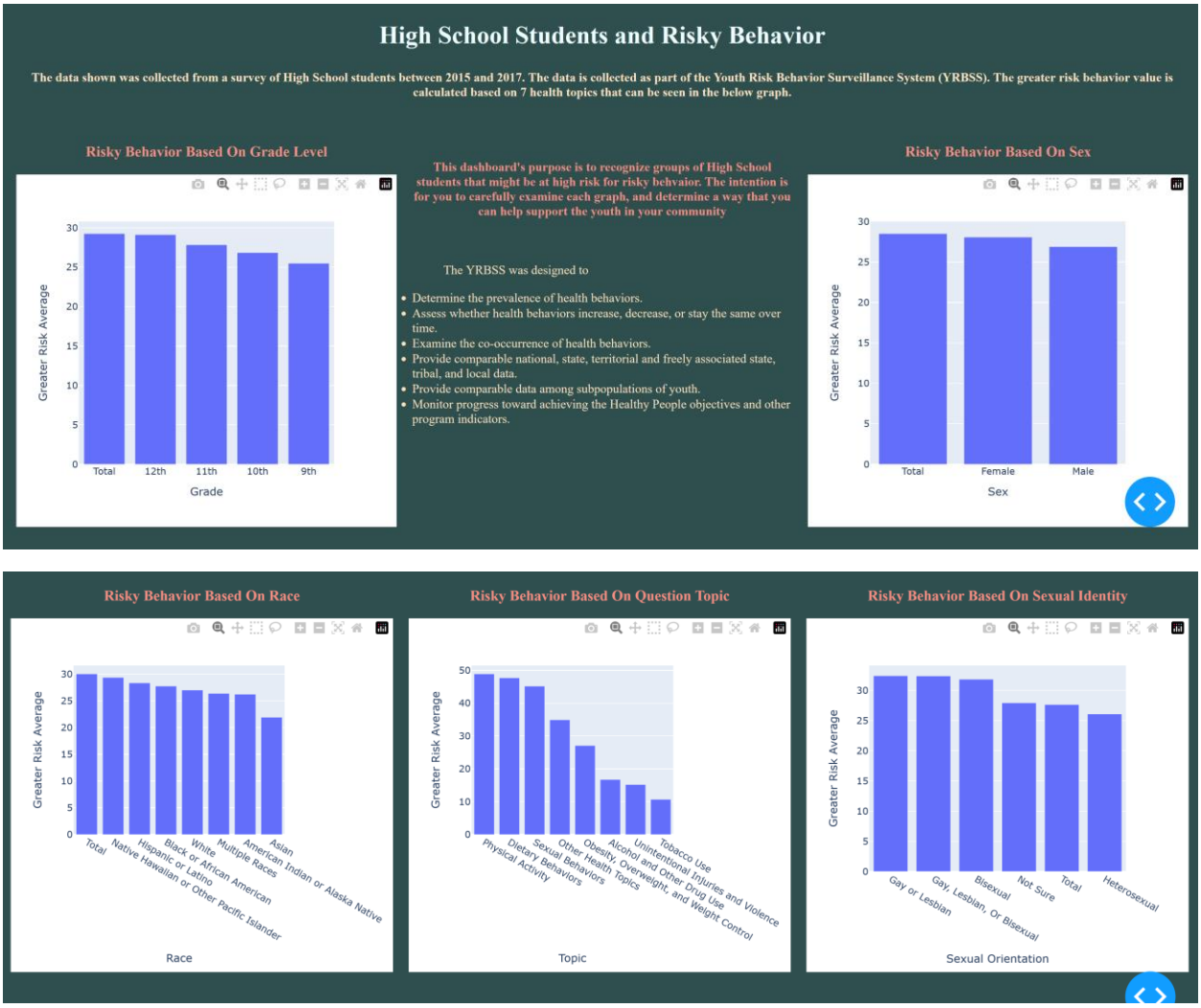
Emily West

Final Project Narrative

1.

My dashboard is focuses on the risk behavior averages for high school students in different locations, but mainly the United States. My dashboard only shows bar charts, the first 5 being non-interactive and show the greater risk averages for different demographics of high school students. The final graph is also a bar chart. This chart allows the user to select different locations from a dropdown, with a few being selected automatically. Then the user can select either greater risk average, lesser risk average, or both from a checklist. This checklist will filter the chart based on the value selected. If both are selected then it will be a grouped bar chart, if non are selected then it will be an empty figure. There is also a link at the very bottom of the dashboard that links to where the dataset is located online.



2.

The dataset for this dashboard can be found at https://chronicdata.cdc.gov/Youth-Risk-Behaviors/DASH-Youth-Risk-Behavior-Surveillance-System-YRBSS/q6p7-56au

3.

The infrastructure requirements are anaconda / python, dash, plotly express, and pandas. All of these frameworks are used in the project.

4.

A use case for my dashboard is that a user will be able to navigate to the dashboard, look at the beginning information, and then be able to filter and select different attributes for the final chart, where the chart will then update depending on that criteria. My hopes are that the user would be able to gain enough information to connect reasoning behind why different demographics might be more at risk or not.

5.

During my development process, I was frustrated that my computer was running so slowly, it would take around 5-10 minutes to run my project. Then I looked back at my dataset and realized that it was 16.5 million lines long. I knew it was a large dataset but I didn't realize how hard it would be on my computer. To counteract this, I took the average values from that dataset and put them into my own files. For ease of use I made a separate csv for each graph, so that I could easily organize the data how I wanted. This was by far the most frustrating part of the project, as I hand wrote each piece of data. I realize that I could automate this problem, however I thought that would most likely take me longer than doing it manually.

After I fixed that problem, it was mostly smooth sailing. Each of my first charts are simple once that data is prepared. I struggled a bit on formatting the page how I wanted it but

quickly began to understand how it worked. My next biggest challenge was creating the functionality for the interactive graph. I just took it one piece at a time, made sure the callback was working, then added a graph, then added a radio, then added the dropdown, then changed the radio to a checklist, then the grouped chart, and soon I was finished with that. I am glad that I did not try to do everything at once, as this would have lead to more errors and problems to fix.

6.

I learned that data visualization is much easier said than done. From reading the book, the choices made seem like easy decisions, until it is completely up to you on how to make a dashboard look or even the charts themselves. I was questioning some of my decisions and how to make the dashboard look the best. It is always easier to create something when all the tools are given to you, like the data, how it should look, what is included, etc. But when you have to design and implement an entire dashboard with data that you found, it becomes much harder. I really enjoyed this project however and I learned a lot about data visualization from this class and doing this project.