Understanding COVID-19's Effect on Bicycle Route Usage in Seattle
(Fremont, Crown Hill, Belltown, Laurelhurst)

Authors: Annika Halvorson, Ani Ramadurai, Henry Ramstad
June 2023
CSE 163

Research Summary:

For the sake of our analysis, the pandemic began on January 1st, 2020, and ended on December 31st, 2021 (This is an oversimplified view of the situation which must be taken to define strict cut-offs between what is and isn't considered part of the pandemic period.)

1. **Among the four sensors, which one has the most traffic (regardless of the pandemic)?**
   We will analyze the bike route traffic of four different sensors in various neighborhoods in Seattle to determine if the location is a factor in route traffic and why some sensors might be more popular than others.
   Answer: The Fremont bridge sensor was by far the most popular pre-pandemic, during the pandemic, and post-pandemic. Regardless of pandemic status, the all-time high was in September 2019, when 142,414 cyclists passed the sensor going either direction on the bridge. The cycling culture of the area can explain a cultural impact on the exceptional results of the Fremont sensor. The Solstice parade, which runs during June, kicks off a season of activity in the area, increasing foot traffic and cyclist retention.

2. **What are the general trends of the data (e.g., trending up, certain months higher)?**
   It is important to get a general understanding of the overall trends of bike path usage and use this information in our analysis of bike trends pre-pandemic, during the pandemic, and post-pandemic.
   Answer: The general trend of the data followed two key patterns. First, there was an obvious dip during the pandemic, but the decline quickly corrected itself and did not disrupt the line trend increasing dramatically. The most likely cause of this result is the acceptance of cycling as an activity that did not increase the spread of COVID-19. The second trend noted among all four sensors was adherence to season patterns present during all three periods. The number of cyclists would peak in the late summer months and reach lows during the early year period. Poor weather conditions during Seattle winters disrupt all but the most committed cyclists and commuters. The summer and early fall months bring warmer temperatures, enabling school children, hobby cyclists, and competitive racers to be drawn onto the bike paths, setting off the sensors.

3. **How did the pandemic change the frequency of use overall?**
   This allows us to examine how Seattle bike path usage might have gained or lost popularity due to COVID-19.
   Answer: The pandemic significantly impacted the frequency of bike use overall. During the pandemic, there was a noticeable decrease in the frequency of bike use compared to the pre-pandemic period. This decline in bike activity can be attributed to various factors, such as lockdowns, social distancing measures, and people's concerns about safety and transmission of the virus. Many individuals were confined to their homes, offices shifted to remote work setups, and recreational activities were limited or restricted. These factors likely contributed to decreased overall bike usage during the pandemic.

4. **How did the pandemic change individual sensor data?**
   While the previous question examines how the pandemic might have impacted bike paths and allows us to see if certain regions' bike path usage was impacted more or less than others.
   Answer: The impact of the pandemic on individual sensor data varied depending on the specific location. However, in general, the pandemic resulted in a decrease in bike activity across all the sensors. During the pandemic, bike usage significantly dropped, especially during the initial stages of lockdowns and restrictions. As the pandemic progressed and restrictions were eased, bike activity gradually increased but remained lower than pre-pandemic levels. The specific patterns observed in the data, such as higher activity during the summer months and fluctuations based on weather conditions, still persisted during the pandemic, but the overall levels of bike usage were lower. It's important to note that individual sensors might have experienced slightly different impacts depending on local regulations, population density, and cycling infrastructure.

Motivation:

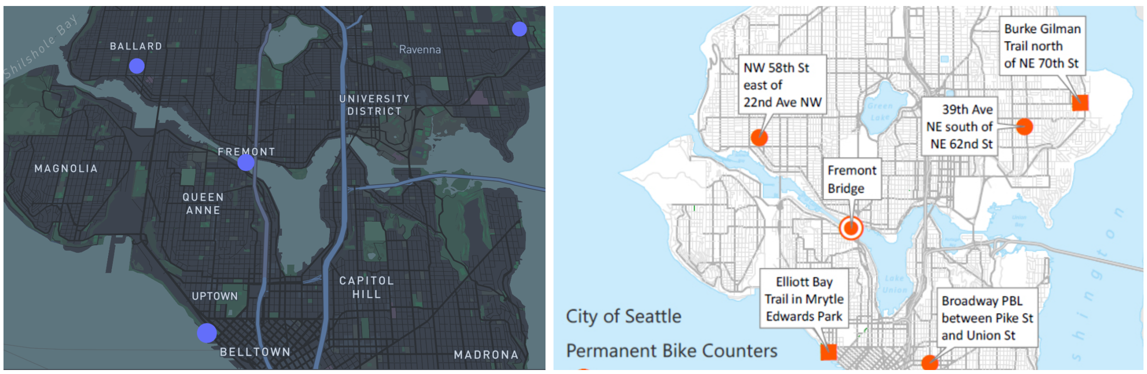
Our research questions were motivated by our shared interest in understanding biking culture within Seattle neighborhoods and how this culture was impacted or even shaped by the COVID-19 pandemic. We wanted to use the data from SDOTs bicycle counters that track ridership over 12 different bicycle counters that they use to assess ridership as a part of their bicycle master plan. This gives us hourly data for each sensor for many months and years that we can use to track the trends in ridership. We are using this information to compare it against the timing of the pandemic and looking at how the trends in ridership compared month to month over the three time periods that we determined, Pre-Pandemic (January 1st, 2018, to December 31st, 2019), Pandemic (January 1st, 2020 to December 31st, 2021), and Post-Pandemic (January 1st, 2022 to June 30th, 2022). Knowing this information, we can assess how the pandemic impacted health trends, as biking is a common transportation method and health method in Seattle. By focusing on Seattle and using it as an example of a major metropolitan area, we can generally predict how the pandemic impacted health trends nationwide.

Our research questions focus on looking at the biking trends in Seattle, regardless of the pandemic, and pandemic-specific information to compare which sensors might have been more popular at a given time. These questions allow us to perform an analysis as to what the impacts of the pandemic were and also how they impacted certain areas of Seattle more than others.

Data Setting:

https://www.seattle.gov/transportation/projects-and-programs/programs/bike-program/bike-counters

The datasets that we will be using come from the Bicycle Counter Project conducted by the Seattle Department of Transportation. In their project, there are 12 different counters, but we decided only to use the four that we found to be in good locations to give us data for all areas of Seattle while analyzing the least amount of datasets.



Bike Counter Image Source

These datasets give us hourly counts of bicyclists crossing the sensors in the Seattle neighborhoods of Fremont, Crown Hill, Belltown, and Laurelhurst. We decided not to use data from further south or east of the Elliot Bay Trail because those datasets did not include the data from the years we needed.

To analyze the data from each bike counter, we had to clean up the data and do some manipulations using Python to be able to join and graph the data. First, we had to reformat the date in the date column by using the built-in Pandas function, “Time Series,” to make the data easier to work with and use the dates as dates. We then used “resample” to group the data from the datasets into each month, as that is how we decided it would be best to analyze the data from the three different periods.

It should be noted that some of these datasets have missing data. The Laurelhurst dataset is missing data from June through September 2019, January through November 2021, and the Ballard dataset is missing data from February and August 2019. The Belltown dataset is missing data from August 2021. We determined that this missing data is likely caused by an error in the technology, such as a broken sensor.

Since we are referring to multiple datasets covering multiple neighborhoods, the bicycle usage patterns likely vary across each neighborhood due to the differences in infrastructure, demographics, and other factors. On top of that, factors including weather and the time of day may have different impacts on bicycle usage in these different neighborhoods. However, most importantly, the main factor that may complicate our analysis is the potential impact of COVID-19 on bicycle usage patterns.

To download the datasets, follow these steps:

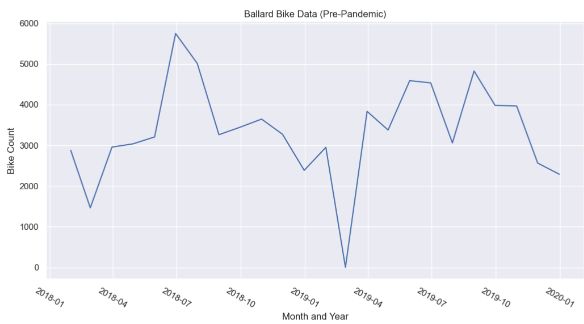
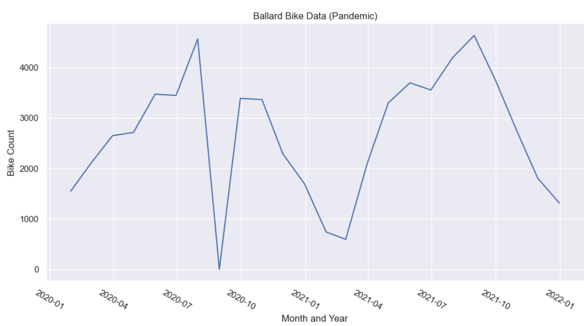
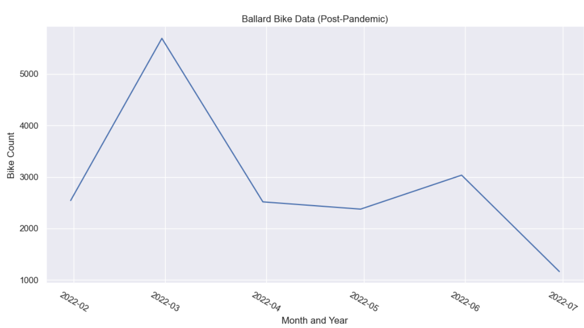
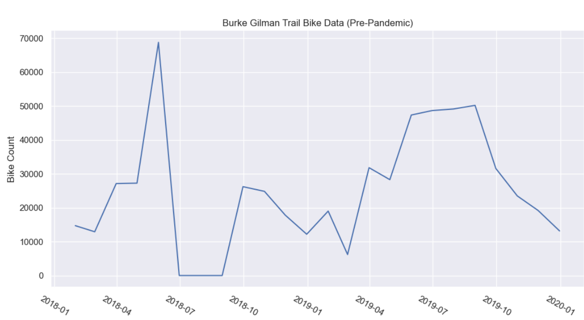
1. Go to the Seattle Government's Data Portal at: https://data.seattle.gov
2. In the search bar, search for "bicycle counter"
3. Select the four datasets titled "Fremont Bridge Bicycle Counter," "Burke Gilman Trail north of NE 70th St Bicycle and Pedestrian Counter", "Elliott Bay Trail in Myrtle Edwards Park Bicycle and Pedestrian Counter," and "NW 58th St Greenway at 22nd Ave NW Bicycle Counter"
4. Click the "Export" button and select the format you want to download the dataset in (CSV, JSON, or RDF)
5. The dataset will be downloaded to your computer.

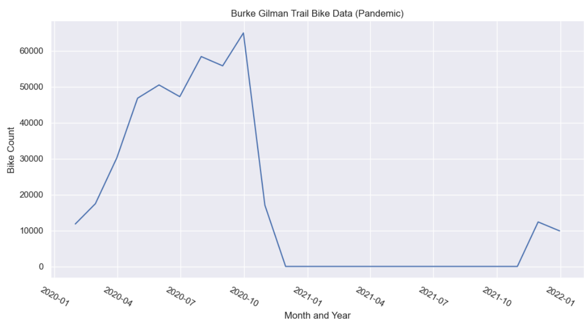
Method:

Change this much to make it more specific and detailed to what we did instead of just bullet points for each question.

1. Among the four sensors, which one has the most traffic? (regardless of the pandemic)
    a. Look at which of the four sensors had the most people using it during the five years.
2. What are the general trends of the data? (trending up, certain months higher)
    a. Look at a line of best fit for the data of the four combined sensors to get a sense of the general trend of bike path usage in Seattle, focusing on steepness and accounting for year-to-year changes.
    b. Look at the bike count by month and compare each, see if any have any specific results (higher or lower)
3. How did the pandemic change the frequency of use overall?
    a. Analyze how the frequency of the entire bike network changed from the pre-pandemic period to during and after. Highlight how certain parts of the data have large peaks or values and explain why they could exist.
4. How did the pandemic change individual sensor data?
    a. Calculate the monthly data usage trends at individual sensors and compare the individual analysis to each other. Look at how one sensor was used more, focusing on individual moments in a sensor's data and relating it contextually to pandemic situations (vaccine rollout, mask mandates, stimulus checks).

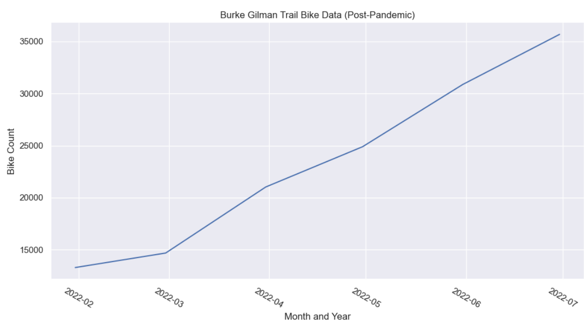
Results:

	Ballard Bike Data (Pre-Pandemic) Here, there was quite a bit of activity before the pandemic. Even though the sensor could not collect data in March of 2019, resulting in the sudden dip in the graph, there were as many as 5749 bikes ridden in this area, especially in the months with good weather, like July. As the seasons started to change, moderate activity remained, but slightly less than in the Summer months.
	Ballard Bike Data (Pandemic) Here, even though there was activity during the pandemic, there was not as much as before. The sensor was not able to collect data in September 2020. Still, like the pre-pandemic, there was more bike activity during the Summer than in other months. Still, due to possible precautions, people may have taken during the pandemic, there were only as many as 4637 bikes ridden here.
	Ballard Bike Data (Post-Pandemic) Here, activity has increased since the pandemic, but it is still slightly less than before the pandemic hit. Unlike the other two periods, where there was mostly activity during the Summer months, there seems to be a lot of activity during the Late Winter/Early Spring time, like in March, with as many as 5688 bikes ridden in the area just like the pre-pandemic.
	Burke Gilman Trail Bike Data (Pre-Pandemic) Here, there was constant activity before the pandemic, and this trail was quite popular among the rest of them. While there was no activity recorded from July 2018 to October 2018, in general, it is safe to say that most activities occurred in the mild and warm temperature months of the year. For instance, in June 2018, as many as 68,772 bikes were counted on this trail.



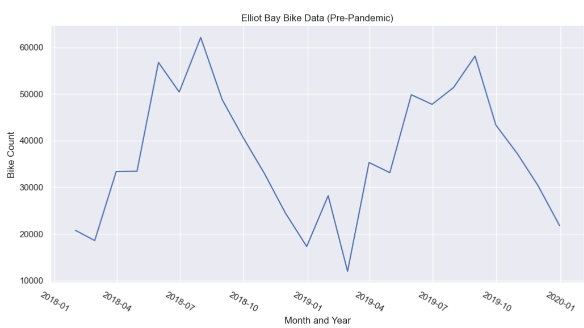
Burke Gilman Trail Bike Data (Pandemic)

Here, activity was mostly on this busy trail but dropped slowly during this time. It's also important to note that data was not collected for most of the year in 2021, except for December. From the data provided, there seems to be a steady increase as the months go by despite the pandemic and increased as protocols and rules became more clear and flexible, specifically in October 2020 with 64,970 bikes.



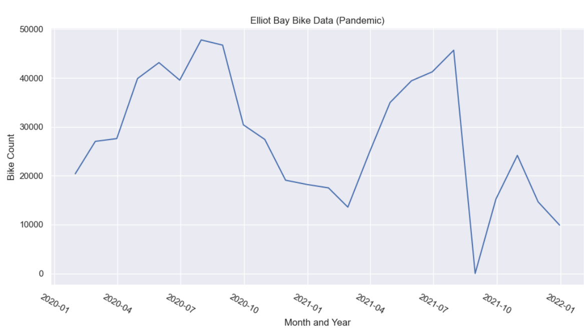
Burke Gilman Trail Bike Data (Post-Pandemic)

Here, activity was not as much as before and, surprisingly, during the pandemic. However, as the months passed, activity on the trail increased. From the beginning of 2022 until July 2022, the bike count increased from approximately less than 15,000 to 35,687, which is a little more than double the number of bikes counted at the start of the year.



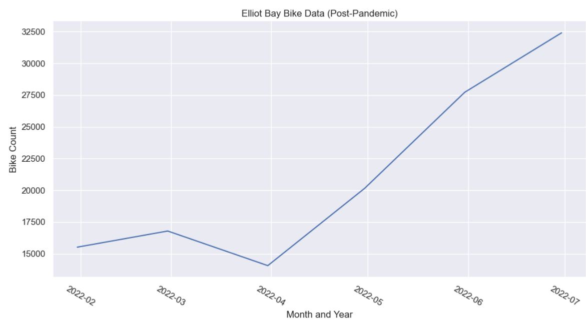
Elliot Bay Bike Data (Pre-Pandemic)

Here, there was always some kind of activity going on. Most of the activity seemed to occur between Spring and Fall, especially between June 2018 and September 2018, as well as in September 2019. Before the pandemic, there was a maximum of 62,098 bikes counted in September 2018, even though other months around this timeframe also had as many or fewer bikes passing by the trail.

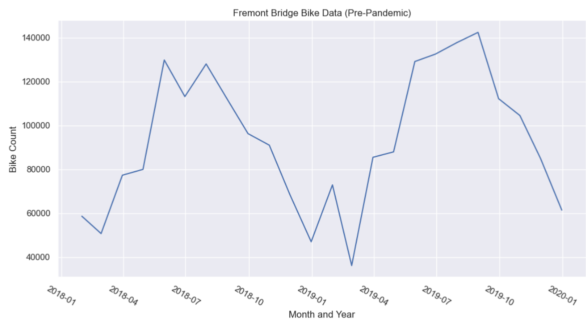


Elliot Bay Bike Data (Pandemic)

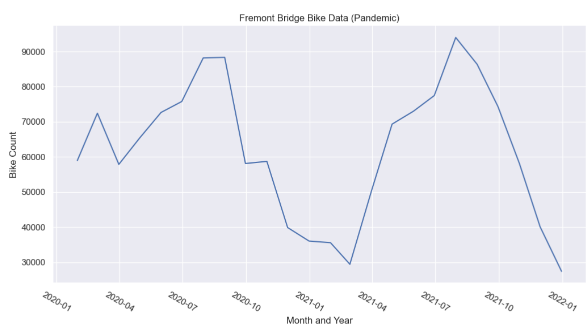
Here, activity decreased as expected, but not as much as seen in the data from other trails. It is important to note that there was not any data collected in August 2021, but other than that, most activities seemed to occur around July 2020 and July 2021 and other months around the same time, with a maximum of 47,780 bikes counted before there was a decrease when the colder months of the year hit.



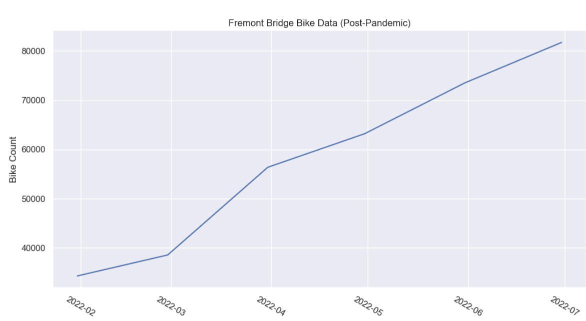
Elliot Bay Bike Data (Post-Pandemic)
Here, activity slowly started to pick up after the pandemic. The numbers are not as high as before and during the pandemic. As the weather started to become warmer in the Spring and Summer seasons, bike activity on the trail increased, going from just below approximately 15,000 to 32,402 bikes in July 2022, doubling the number of bikes in a span of just three months from April 2022 to July 2022.



Fremont Bridge Bike Data (Pre-Pandemic)
Here, the highest activity across all trails occurred in Fremont. Regardless of the month during this time, there were always at least approximately 40,000 bikes. The Spring and Summer seasons were when there was the most activity on the trail, especially in September 2019, with 142,414 bikes passing by, and they slowly started to drop when Fall and Winter came around.



Fremont Bridge Bike Data (Pandemic)
Here, there was still quite a bit of activity, and it only dropped when the weather became colder. Activity shot up significantly for most of the Summer and early Fall months, making two peaks, as seen in the graph. Also, the highest number of bikes passing by during the Pandemic was during this time, with 93,994 bikes riding through the trail in August 2021.



Fremont Bridge Bike Data (Post-Pandemic)
Here, activity grew steadily from the beginning of the year until July 2022, when 81,720 bikers were passing through the trail. In January and February, activity was not as high, possibly due to the cold weather and possibly unsuitable for many people to ride bikes. However, activity increased drastically from March onwards as the weather became warmer.

Among the slowly four sensors, which one has the most (regardless of the pandemic)?

The Fremont bridge sensor was popular pre-pandemic, during the pandemic, and post-pandemic. Regardless of pandemic status, the all-time high was in September 2019, when 142,414 cyclists passed the sensor going either direction on the bridge. The cycling culture of the area can explain a cultural impact on the exceptional results of the Fremont sensor. The Solstice parade, which runs during June, kicks off a season of activity in the area, increasing foot traffic and cyclist retention.

What are the general trends of the data (e.g., trending up, certain months higher)?

The general trend of the data followed two key patterns. First, there was an obvious dip during the pandemic, but the decline quickly corrected itself and did not disrupt the trend of the line increasing very dramatically. The most likely cause of this result is the acceptance of cycling as an activity that did not increase the spread of COVID-19. The second trend noted among all four sensors was adherence to season patterns present during all three periods. The number of cyclists would peak in the late summer months and reach lows during the early year period. Poor weather conditions during Seattle winters disrupt all but the most committed cyclists and commuters. The summer and early fall months bring warmer temperatures and enable school children, hobby cyclists, and competitive racers to be drawn onto the bike paths, setting off the sensors.

How did the pandemic change the frequency of use overall?

The pandemic significantly impacted the frequency of bike use overall. During the pandemic, there was a noticeable decrease in the frequency of bike use compared to the pre-pandemic period. This decline in bike activity can be attributed to various factors, such as lockdowns, social distancing measures, and people's concerns about safety and transmission of the virus. Many individuals were confined to their homes, offices shifted to remote work setups, and recreational activities were limited or restricted. These factors likely contributed to decreased overall bike usage during the pandemic.

How did the pandemic change individual sensor data?

The impact of the pandemic on individual sensor data varied depending on the specific location. However, in general, the pandemic resulted in a decrease in bike activity across all the sensors. During the pandemic, there was a significant drop in bike usage, especially during the initial stages of lockdowns and restrictions. As the pandemic progressed and restrictions were eased, bike activity gradually increased but remained lower than pre-pandemic levels. The specific patterns observed in the data, such as higher activity during the summer months and fluctuations based on weather conditions, still persisted during the pandemic, but the overall levels of bike usage were lower. It's important to note that individual sensors might have experienced slightly different impacts depending on factors such as local regulations, population density, and cycling infrastructure.

Impact and Limitations:

The data indicated that areas of Seattle have significantly higher numbers of cyclists. This data could impact how the city decides to allocate resources for new bike trails or how small businesses cater to the cycling community. There is an inherent bias in the cycling community of Seattle, specifically the northern portion of the city. This area is home to more affluent communities that may have more free time to commute to work by bike or ride for pleasure. An effect of the pandemic not mentioned in this analysis is the increase in bike cost due to supply chain issues. This increase in price and demand meant that individuals from lower socioeconomic backgrounds did not have access to cycling due to the price gauges. This situation meant that sensors outside of north Seattle saw a reduction in use while the northern sensors saw increases in some cases. The data limitations were based on the locations selected, removing analysis from some of these more southern sensors or sensors that fell outside the Seattle area.

Challenge Goals:

The Challenge goals completed by this group for this project involved two key elements. Firstly, we sought to visualize our data using an interactive map, allowing the user to modify and query the data at will and produce modified visuals that reflect the selection of the user. The second challenge goal was to write the project in a way that it could be deployed and maintained on a live server through GitHub pages. The Frameworks used to complete these goals were Dash, Plotly, PDFjs, Numpy, Pandas, Seaborn, and Matplotlib. The Plotly Mapbox function was used to generate a styled map in which our data was overlaid on top of using geospatial latitude and longitude coordinates. These goals stretched the group's understanding of the content of the course through our introduction to living service development and writing code using unknown packages.

Plan Evaluation:

We estimated around 30 hours total to complete the project but ran into issues that caused the development time to stretch past 60 hours. Some of the highlights which delayed development included issues with the underlying datasets, which required sorting and modification before the official wrangling began. Certain bugs plagued us, such as unorganized data entry, which crashed time series methods with little backtrace. The overall implementation of Dash also caused some delays due to a general lack of knowledge and trouble fitting the developed map element with the filtered data elements. When both sides of the project met in the middle, it was hard to bring cohesion due to some restrictions of the Dash library in relation to heightened interactivity. The group completed three four-hour sprints together alongside multiple remote meetings to complete the project and report.

Testing:

The testing for the project primarily examined how the filtering was done and ensured that what we planned to occur was coming out of the other side. We attempted to write tests for the website itself, but due to conflicts with interactivity, we decided to focus our effort on ensuring that our data was as tested as possible. A separate file was created using assert_equals from

CSE163_utils to look at certain elements of each dataset. Due to the nature of our filtering, all 4 datasets had the same number of columns and rows and were tested to ensure the new columns added for lat and lon had the correct values. Additional tests were all run to ensure the resampling was done correctly, and each method called in filter.py corresponded with a test for their effect.

Collaboration:

Besides the course staff and team members, nobody else was consulted during the project. However, websites like Stack Overflow and Code Exchange were occasionally referred to to double-check specific concepts and ideas implemented in our project, along with the package documentation for Dash, Plotly, Seaborn, Pandas, and Matploblib.