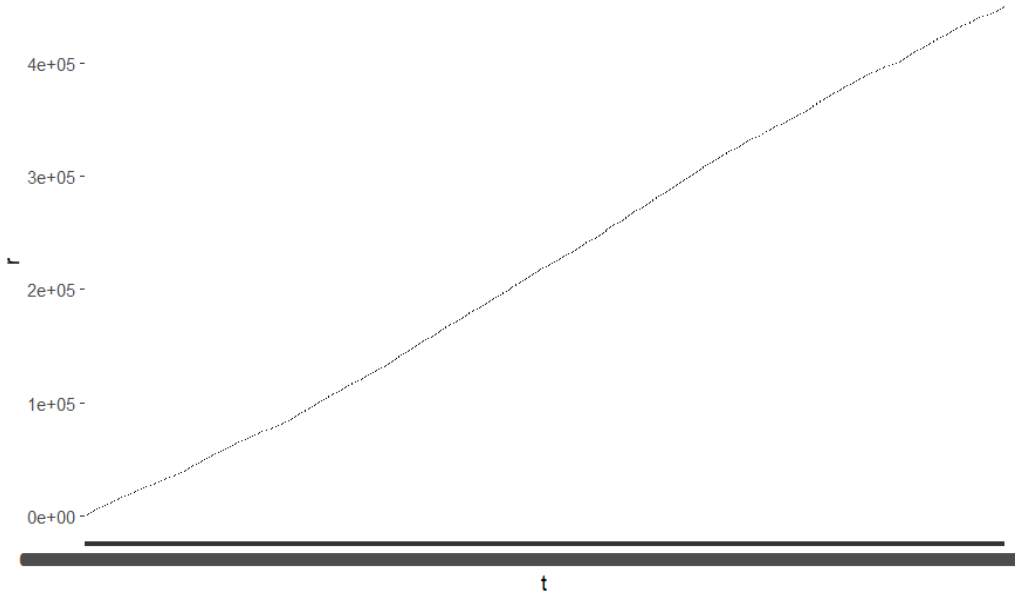
Our idea is to collect the bikeshare data and taxi rider data for the years 2020 and 2022. The 2020 will serve as our year where travel has been affected by COVID-19 and 2022 will be our year where the effects from COVID-19. We will view each month individually and specifically in the peak months of COVID-19 in 2020 to see how travel has changed while people are trying to avoid cramped public transportation. Based on these findings, we will conclude how COVID-19 affected “for hire” transportation in DC.

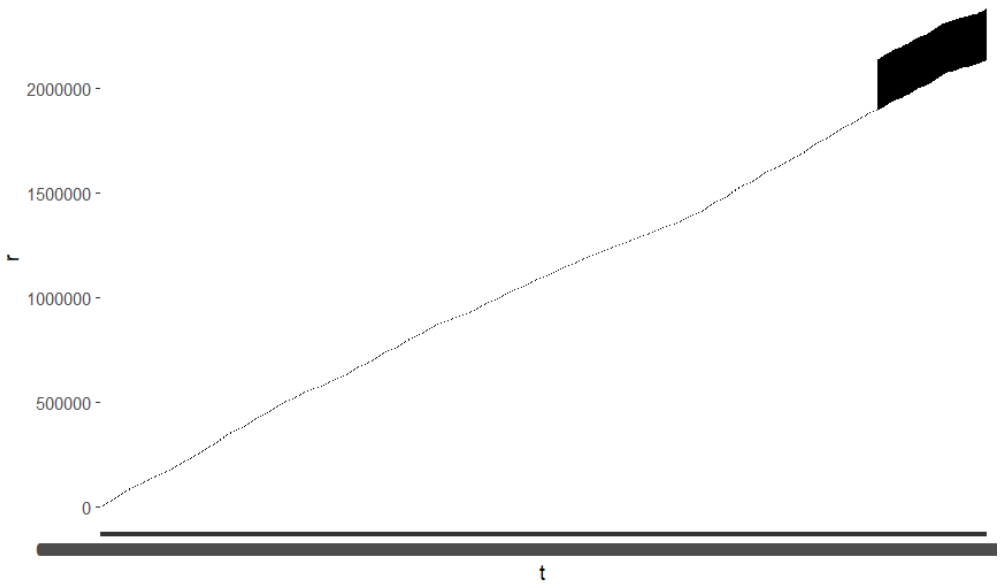
While working with our data, we realized that the months January-March 2020 were missing the start times and end times for the bikes. Due to this complication and no way to add these times, we excluded January-March 2020 and January-March 2022 data from our graphs for both the taxi and the bikeshare data. We deemed this appropriate because, around April 2020, COVID-19 was strongly prevalent, which should be reflected in the data.

For the analysis of the bikeshare data, we included the casual riders and members and plotted the cumulative amount of riders against the time. For the taxi data we were going to use the ObjectID, but we discovered that the variable objectid does not refer to the number of passengers while working with the taxi data. We decided to filter out the distance that the taxi traveled to make sure that the taxi trip counted and used that as our riders and we calculated the cumulative sum of the riders. We utilized the function mutate to construct a new column that calculates the number of taxi rides based on mileage. If the mileage is zero, no ride occurred. Below are the graphs for the 2020 and 2022 data files for the taxi data:

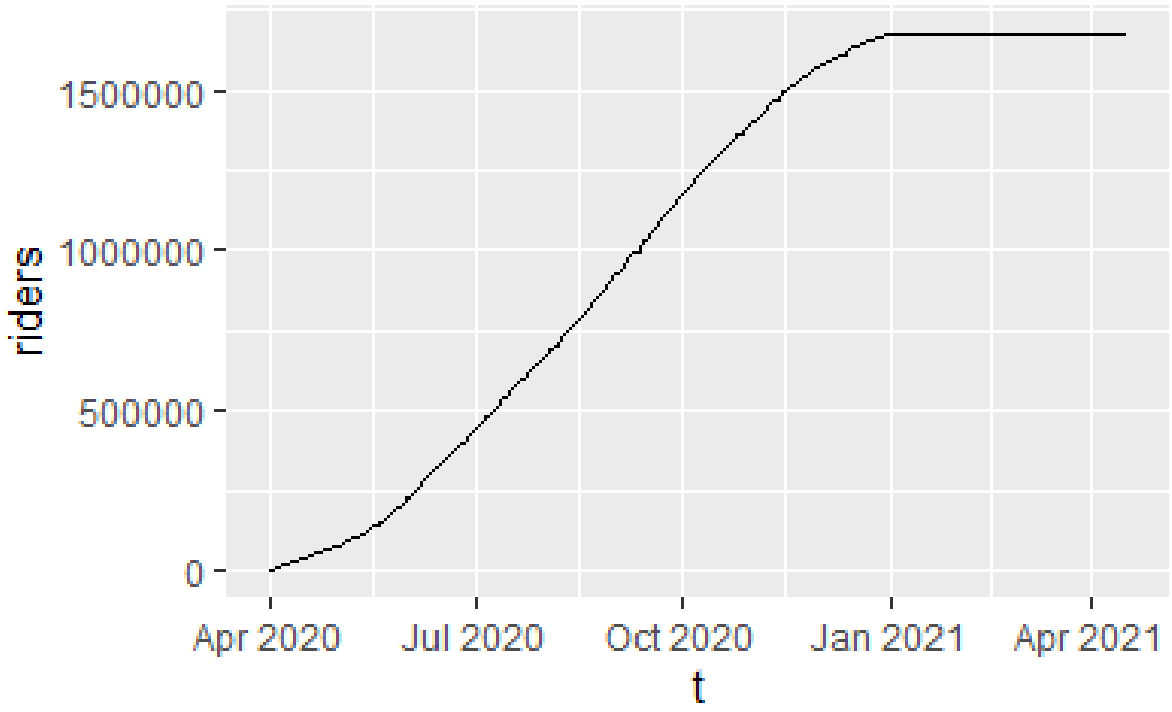
2020 (taxi):



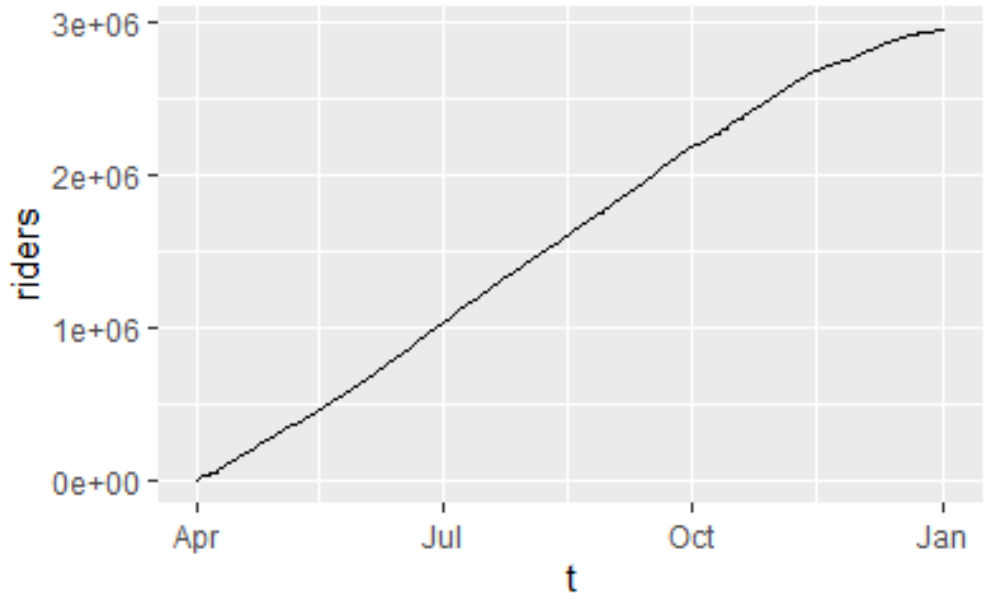
2022 (taxi):



2020 (bikes):



2022 (bikes):



The taxi data showed less trips and usage compared to the bikeshare program. However, the taxi data seemed to have a more linear pattern than the bikeshare data. Right after COVID-19 in April 2020, there was less usage of the bikeshare program compared to later in the year. We can observe the slope of the 2020 graph for the bikeshare program to find out if people often used the program or not. In April 2020 to June 2020 and December 2020, the slope of the graph shows that there was not a lot of increase in ridership. The middle of the year had a lot more increase in rides than the beginning and end of our analysis. On the other hand, 2022's bikeshare data has a more consistent increase throughout the year and a slower increase in ridership near the end of the year in around November. Additionally, from April 2020 to December 2020 had around 1,525,000 rides and April 2022 to December 2022 had around 3 million rides.

Taxi data seems to be consistent all throughout the year in 2020 and 2022. However, taxi usage seems to have increased by 5 times from 2020 to 2022. Bikeshare data seems to have only increased by 2 times, which can show that COVID-19 didn't affect the bikeshare program as much as the taxi program.

We can continue to do more research about why the bikeshare program wasn't as affected as other industries by looking into the types of people that use the taxi versus the types of people that use the bikeshare program. Furthermore, we can calculate the change in riders to see more accurate analysis about the bikeshare program to see if it has recovered and grown or not. Some problems we had were with the taxi data. It was messed up near the end of 2022 and that may have affected the analysis. Additionally, the taxi data didn't have a date time in the POSIX format so we couldn't make a very good looking graph.