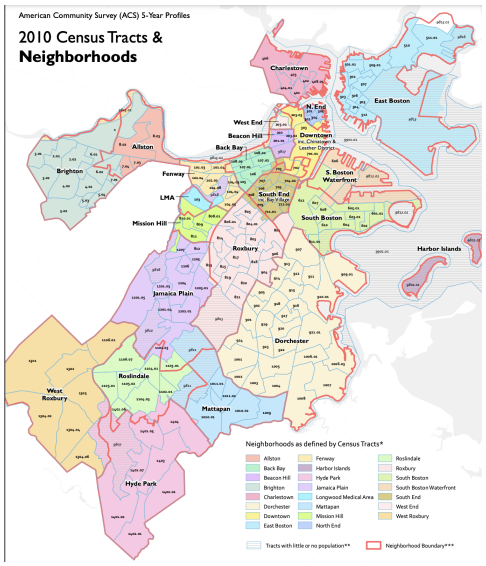
Project Deliverable 1

Livable Streets Bus Equity

Team member: Xiang Li, Xinyu Zhang, Bohan Wang, Yu Liang

1. Collect and pre-process a preliminary batch of data

Pre-Process(Xiang Li):
We find the corresponding data by using the US Census data. As the requirement from our client, we will focus on these three areas: **Roxbury, Dorchester, and Mattapan**. We follow the graph to access these three places by their census id.



The data we used include:
Question 1:
B08105A | MEANS OF TRANSPORTATION TO WORK (WHITE ALONE)
B08105B | MEANS OF TRANSPORTATION TO WORK (BLACK OR AFRICAN AMERICAN ALONE)

Question 2:
B08134: Means of Transportation to Work by Travel Time to Work

Question 3:
B08105IMEANS OF TRANSPORTATION TO WORK (HISPANIC OR LATINO) (data from 2011, 2016, 2021)

Question 4:
B08134: Means of Transportation to Work by Travel Time to Work (2016-2021)
B08105A: MEANS OF TRANSPORTATION TO WORK(WHITE ALONE) (2019-2021)

B08105B | MEANS OF TRANSPORTATION TO WORK (BLACK OR AFRICAN AMERICAN ALONE)
(2019-2021)

2. Perform a preliminary analysis of the data

Question 1:

Pre-Covid: 2010 (2010-2015); 2015 (2015-2020)

Covid Period: 2020 (2020-2024)

After-Covid: 2021 (2021-2025)

The data I use for these years are the population of riders on the bus for black and white riders of Roxbury, Mattapan and Dorchester , which includes the data for that year and the forecast data for the next five years.

Question 2:
After preprocessing the data, we have all means and bus data include: time["Less 10 mins","10-14 mins","14-19 mins","20-24 mins","25-29 mins","30-34 mins", "35-44 mins","45-59 mins","60 and more"] of Roxbury, Mattapan, Dorchester in 2021,2016,2011

```
df
```

Label (Grouping)	Roxbury_2021	Roxbury_2016	Roxbury_2011	Mattapan_2021	Mattapan_2016	Mattapan_2011	Dorchester_2021	Dorchester_2016	Dorchester_2011
Total:	11474	10640	8719	6925	6778	6129	27753	27321	23383
Less than 10 minutes	799	1088	816	217	105	240	1463	1197	1320
10 to 14 minutes	820	1096	718	278	379	271	2018	2068	2059
15 to 19 minutes	1321	1388	963	564	569	530	2734	2876	2265
20 to 24 minutes	1117	1375	1461	805	805	669	3466	3554	3050
...	...	...	...	...	...	...	...	...	...
25 to 29 minutes	51	0	51	0	45	0	46	31	23
30 to 34 minutes	43	30	28	11	15	0	66	63	164
35 to 44 minutes	17	8	21	29	0	38	6	10	15
45 to 59 minutes	15	0	4	64	7	8	158	38	0
60 or more minutes	132	34	8	39	0	0	229	111	56

120 rows × 9 columns

Question 3:
I use the data from 2011 to 2021 in every five years. The five years data means what will the feature data will look like based on the current information. After selecting the data from these three areas, we only need the data with the label “Public Transportation” and “Estimate”. There may be some potential information to show how the Covid-19 can affect the public transportation usage of Latinx communities in these three areas.

Question 4:

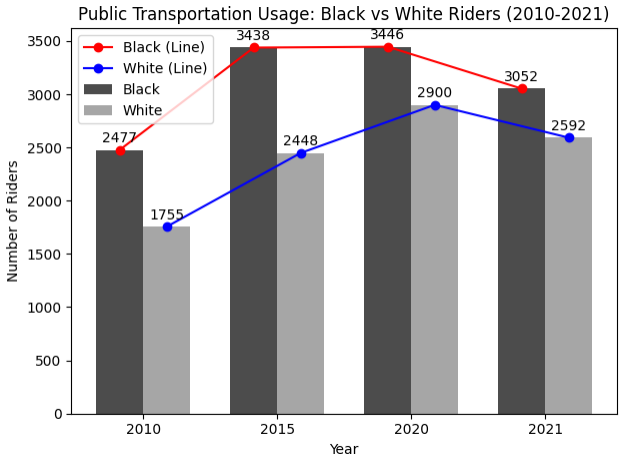
I mainly focus on how Covid-19 pandemic impacts the bus commute times. I'm looking for the change before and after covid, so I shortened the time span and used more datasets. Although this is a five-year forecast, it is released annually, and I think this can show a yearly change trend. I organized the data from 2016-2021 using the method mentioned in part 1, summed up the results, and obtained three neighborhoods.

3. Answer one key question

Question 1 (Yu Liang):

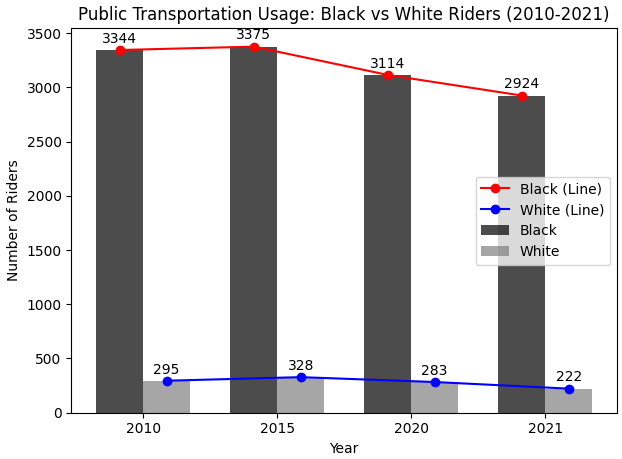
How has the 64-hour bus commute disparity between Black and White riders evolved over the past five years, considering both Pre-Covid and Covid-Affected periods, and are there discernible patterns or changes in the trend?

Roxbury:



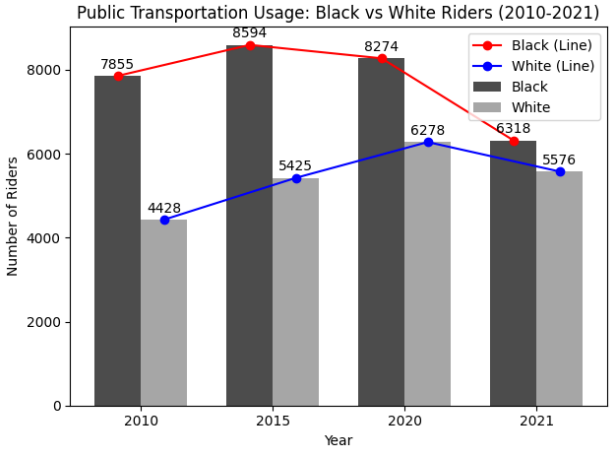
The number of black and white riders has been on the increasing trend from 2010 to 2020, but after the beginning of the Covid in 2020, the number of riders has a downward trend.

Mattapan:



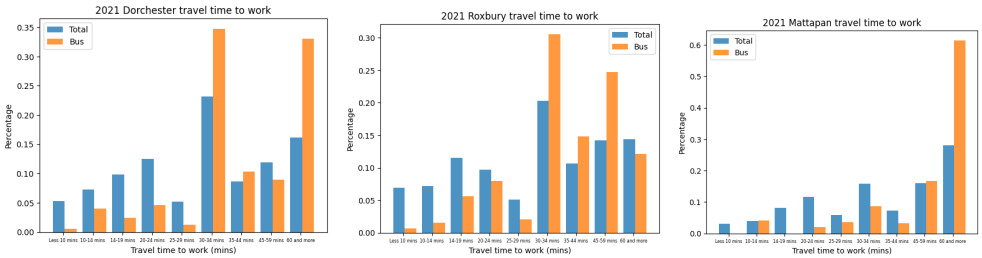
During 2010 and 2020, the number of white riders did not change significantly and was relatively flat. The number of black riders was relatively flat between 2010 and 2015, and showed a significant downward trend between 2020 and 2021 after the beginning of the Covid.

Dorchester:

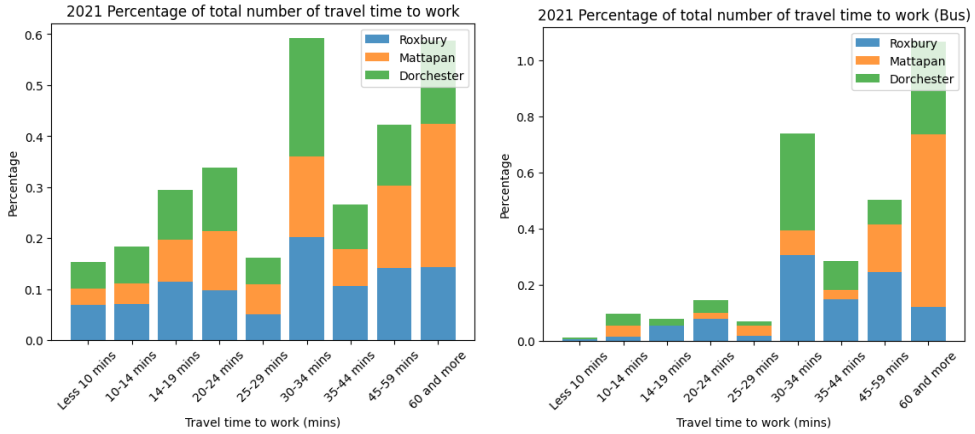


The number of black and white commuters increased from 2010 to 2020 but decreased significantly from 2020 to 2021 after the beginning of the Covid.

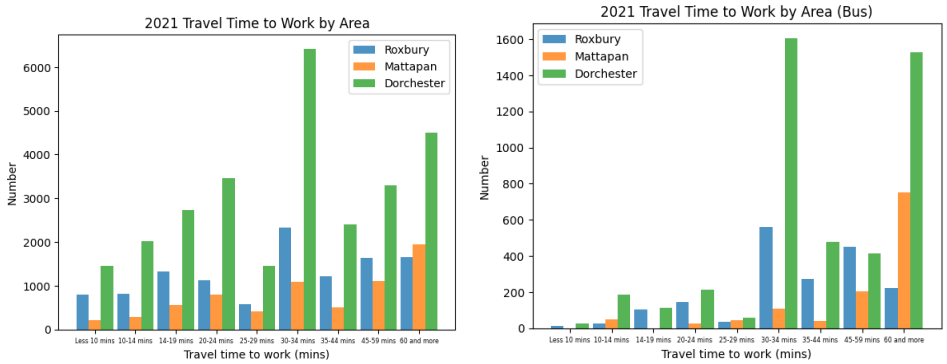
Question 2 (Bohan Wang):
Does the 64-hour gap in bus commute times hold consistent across different geographic areas or does it vary significantly, and are there specific regions where the gap is more pronounced?



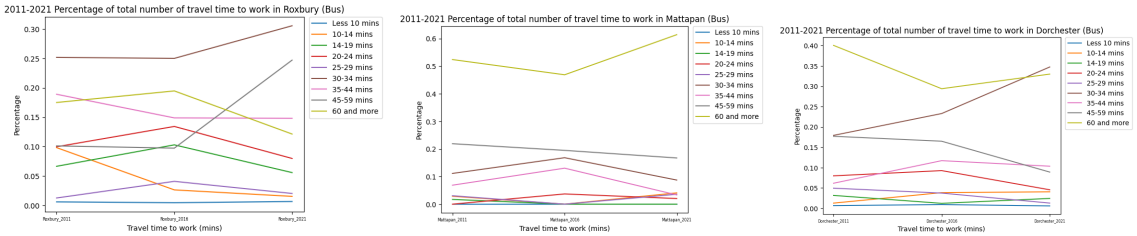
Bus cost more time compared to other means of transportation in traveling to work in different areas.



Mattapan cost most time about 60 mins and more. Dorchester also cost 30-34 mins and 60mins most.



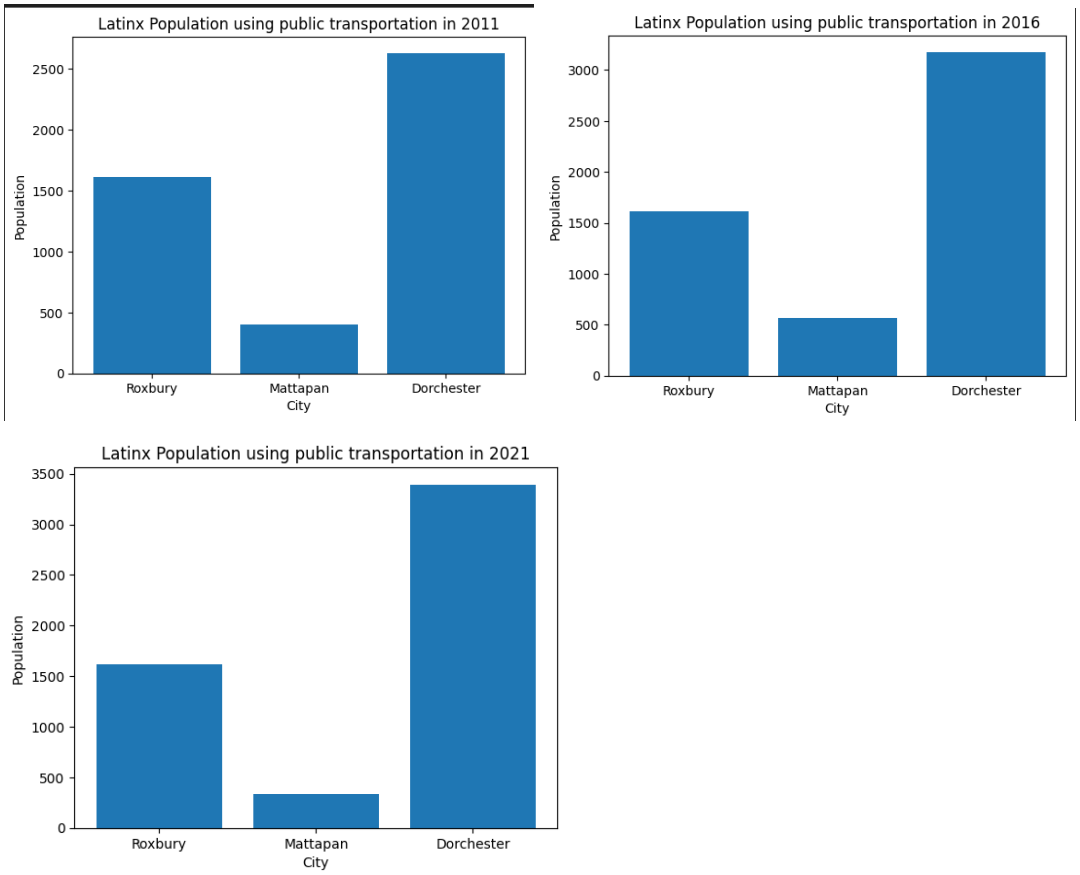
We can also see that 30-34 mins and 60 mins and more are two important time points of travel to work by bus.



2011 2016 before covid,2021 is covid period, in roxbury, 30-34 and 45-69 is increased after covid. In Mattapan, 60 and more is increasing. 30-34 is increased in Dorchester.

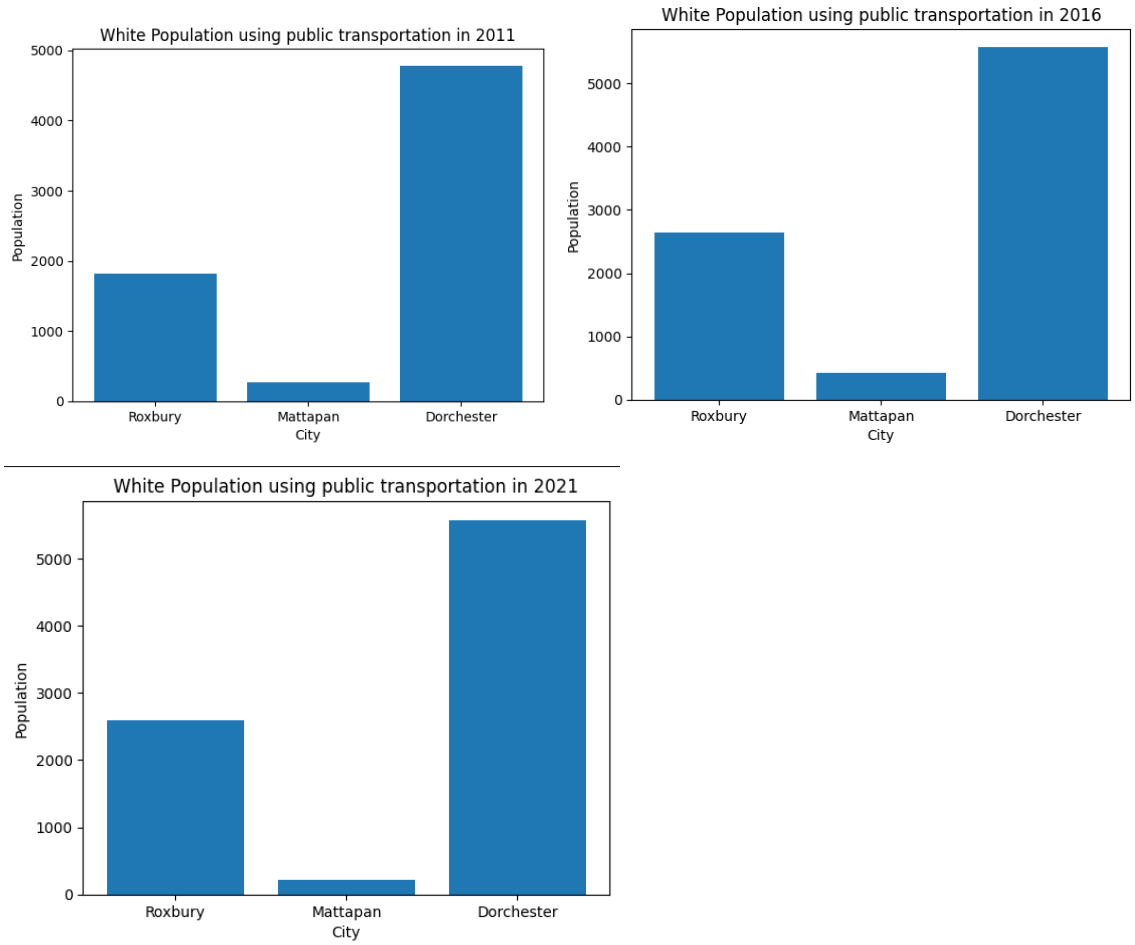
Question 3(Xiang Li):

How does the 10-hour bus commute disparity observed in LatinX communities compare with other racial groups, and are there common factors contributing to these discrepancies across different demographics?



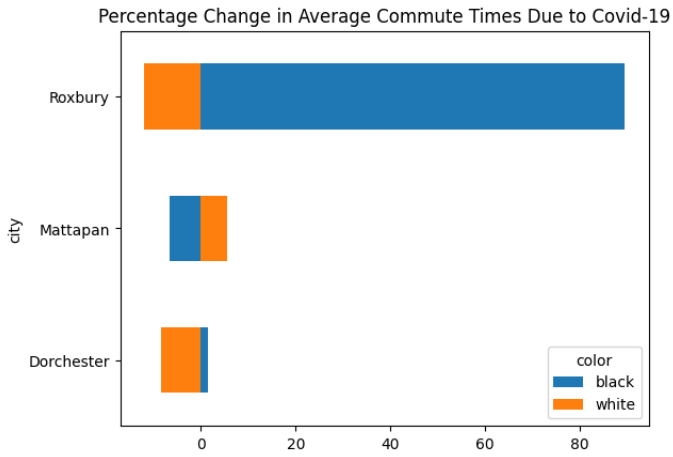
We do not have the data measured about the Latinx and other communities' commute time in different census track locations, so I am not able to answer this question at this moment. What I can do is find some visualizations about the demographics of different racial groups. We can see from the above plots that Dochester alway has more Latinx who use public transportation than other two areas and it hass been continue increase in recent years.

I compare the demographics data of Latinx group to white group and I will continue this step when I find more data to support this question. The porportation of each city has the same structure compare to the Latino population.



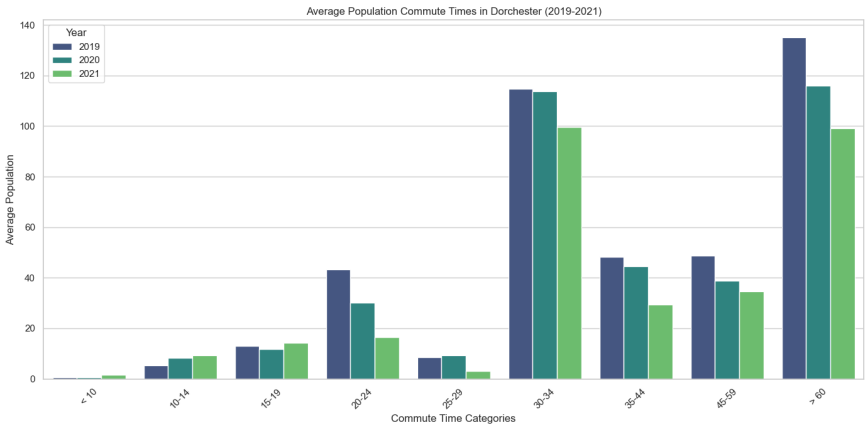
Question 4(Xinyu Zhang):
To what extent did the Covid-19 pandemic impact the bus commute times for both Black and white riders, and how has this impact contributed to the existing disparity in hours spent on buses?

The following graph is the percentage change in average commute times for Black and White populations due to COVID-19. The black population in Roxbury using the bus for commuting has increased dramatically. The white population has decreased a little. In Dorchester, the number of black people taking the bus to work has increased slightly, while the number of white people has decreased a lot. The trend in Mattapan is opposite to the other two cities, which may be due to a different population structure for those commuting by bus in Mattapan.

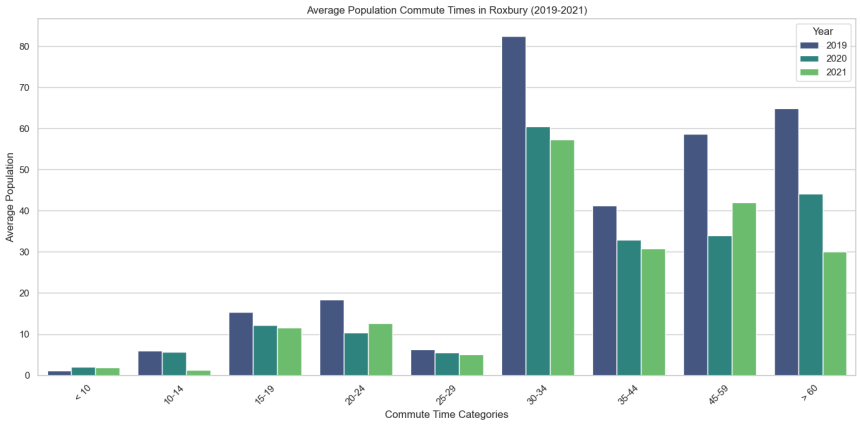


City	black	white
Roxbury	89.44%	-11.90%
Mattapan	-6.66%	5.56%
Dorchester	1.42%	-8.29%

The following graph is the average commute time in three neighborhoods.

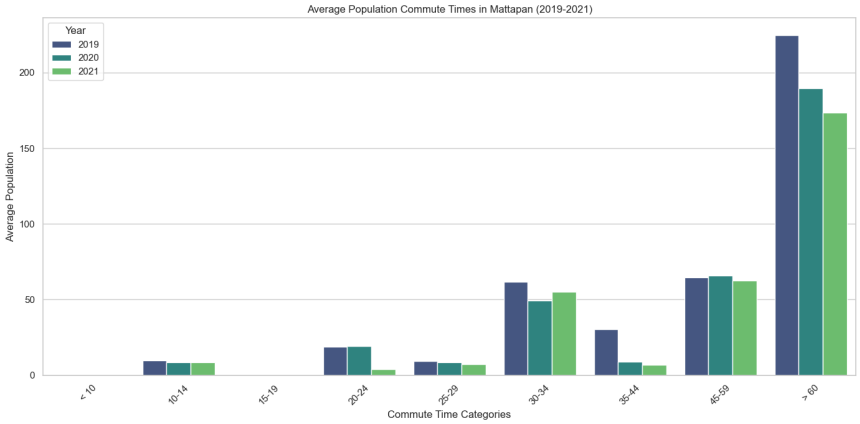


From the graph, we can see that the commute time decreased from 2019 to 2021 in 20 minutes or more category, and increased at <20 min category. There could have been a trend towards more remote working options, resulting in more people reporting shorter commute times and fewer long commute times.



There’s a noticeable decrease in the 35-44 minute category from 2019 to 2020, which then slightly increases in 2021 but does not return to 2019 levels.

Also a general decrease in the number of people commuting for longer times (45 minutes or more) from 2019 to 2020, with a slight increase or stabilization in 2021.



Have a different structure than the other two neighborhoods. Most people commute >60 min. Similar to Roxbury, there was a decrease in 2020, then stabilized at 2021.

Conclusion: By 2020, there are fewer people in shorter commutes and more people in longer commutes. This could mean that people weren't going to work as much because of lockdowns or working from home, and only essential workers were commuting, and they often have longer trips.

By 2021, the numbers look a bit more like 2019, but still with more people having longer commutes. Maybe more jobs went remote permanently, or people moved farther away from work because they didn't need to go in as often.

4. Refine project scope and list of limitations with data and potential risks of achieving project goal

- **Project scope**:
    1. Find the differences in commute times and choice ratios for using transit across regions, and across ethnic groups. And determine trends over time, focusing on pre- and post-covid changes. Then make policy recommendations.

- **Limitation**:
    1. 1. Data spans from 2010 to 2021 but lacks post-2021 data, which is crucial for analyzing trends in the post-COVID period.
    2. 2&3. Data includes estimated commute times for all racial groups in different areas for the next five years but lacks actual recorded data.
    3. 4. Having the population data for Black and White and LatinX riders but not their specific commute times.

- **Risk:**
    1. Without recent data, it's challenging to assess how the pandemic has affected the commute disparity. This gap might lead to outdated or incomplete conclusions.
    2. Relying on estimates can introduce uncertainty and potential inaccuracies in the analysis. Real-time or historical data is more reliable for understanding current trends.
    3. Without detailed travel time data, it's difficult to make direct comparisons and identify factors contributing to disparities.

5. Submit a PR with the above report and modifications to original proposal
   Shane to submit