PRESENTATION SLIDES AND SPEAKER NOTES

- The following pages are not part of our report but are the slides that we used for our in class presentation. We are attaching them here as there is no separate slide submission. The text directly below each slide is the speaker notes/reminders for the slide.
- Note that not every slide has speaker notes and that these speaker notes were kept brief and are partially incomplete as they were only meant to prompt our memories as we presented and not serve as a complete script.

Speaker Notes:

- Here's a brief overview of our presentation.

01 INTRODUCTION

Introduction

- In a recent SF Chronicle poll *45% of participants* reported that they had an item stolen from them in the last 5 years. Specifically, theft from cars is of increasingly concern to SF residents.
- In this report, we ask the question: Do the trends in larceny from vehicles from 2018 to 2022 shift over the course the COVID-19 pandemic?
- We will look into:
  - **Weekly and monthly trends in larceny from vehicles.** In doing this, we can report which days of the week and which months or days of the month that this crime is more likely to occur.
  - **Specific police districts in the city** to see if there are any differences in the crime rates between districts that could be linked to the differences in the socioeconomic status
  - **The resolution rates of larceny from vehicles** will be explored from each district to see if more favoritism is given to wealthier, less diverse neighborhoods
- The results from this report can be used for better crime prevention strategies, better allotment of resources for different districts, and to perhaps shine a light on the lack of attention that some districts may receive from the police force.

Speaker Notes:

- For some more context, in a recent poll by the SF chronicle, 45% of participants reported that they had an item stolen from them in the last 5 years. Specifically, theft from cars is of increasing concern to SF residents.
  - Residents have taken creative and desperate measures to prevent this type of theft from happening to them, and understanding trends in this type of crime would be key in helping SF residents feel safe, and would give them insight into what steps they can take to protect themselves.
- In this report we ask the question: Do the trends in larceny from vehicles from 2018 to 2022 shift over the course of the COVID-19 pandemic?
- We will look into:
  - Weekly and monthly trends
  - Specific police districts in the city
  - The resolution rates of larceny from vehicles
- Results can be used for:
  - Crime prevention strategies
  - Allotment of resources
  - Shine light on possible lack of attention to specific districts

DATA STRUCTURE & CLEANING

02

Speaker Notes:

- The dataset we chose to use for our analysis comes from the DataSF initiative.
- The dataset is titled Police Department Incident Reports: 2018 to Present.
- Each row of the dataset represents an incident record that is filed by a police officer or by members of the public.
- The original full data set has more than 650,00 rows and 34 columns.

Cleaning and Sanity Checking

Index(['Incident Datetime', 'Incident Date', 'Incident Time', 'Incident Year',
       'Incident Day of Week', 'Report Datetime', 'Row ID', 'Incident ID',
       'Incident Number', 'CAD Number', 'Report Type Code',
       'Report Type Description', 'Filed Online', 'Incident Code',
       'Incident Category', 'Incident Subcategory', 'Incident Description',
       'Resolution', 'Intersection', 'CNN', 'Police District',
       'Analysis Neighborhood', 'Supervisor District', 'Latitude', 'Longitude',
       'Point', 'Neighborhoods', 'ESNCAG - Boundary File',
       'Central Market/Tenderloin Boundary Polygon - Updated',
       'Civic Center Harm Reduction Project Boundary',
       'HSOC Zones as of 2018-06-05', 'Invest In Neighborhoods (IIN) Areas',
       'Current Supervisor Districts', 'Current Police Districts'],
      dtype='object')

Incident Datetime	object
Incident Date	datetime64[ns]
Incident Time	object
Incident Year	int64
Incident Day of Week	object
Report Datetime	object
Row ID	int64
Incident ID	int64
Incident Number	int64
Report Type Code	object
Report Type Description	object
Incident Code	int64
Incident Category	object
Incident Subcategory	object
Incident Description	object
Resolution	object
Police District	object
Incident Month	int64
dtype: object

Speaker Notes:

- For our data cleaning and sanity checking process,
- The figure on the left shows the original columns in the dataset
  - We dropped 17 out of these 34 columns because many of these columns either had a large number of null values, or we didn't need them to answer our questions
- We used the "Incident Subcategory" variable to filter the data by "Larceny - From Vehicle" and the "Police District" variable to remove any districts that were not in San Francisco
- The table on the right shows the final columns that we kept for our analysis, as well as their data types
  - We made sure that there weren't any null values in these columns
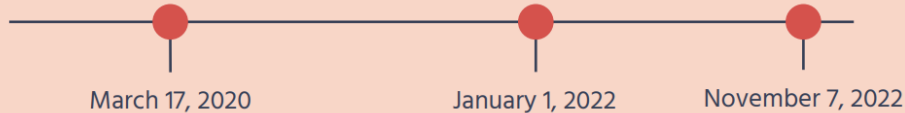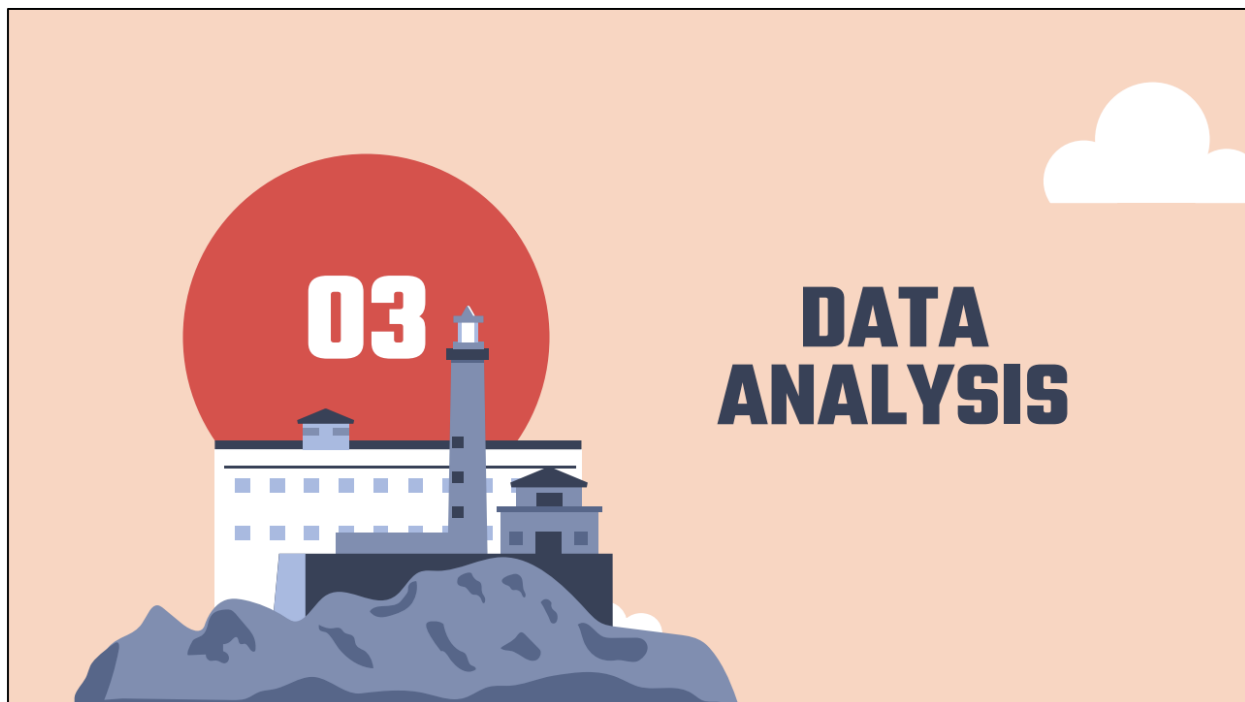
Speaker Notes:

- In order to understand the relationship between thefts from vehicles and COVID, we divided up our data into three time periods
- Pre-COVID was defined as the time period before March 17, 2020, which was the day San Francisco implemented the stay-at-home orders
- Peak-COVID was defined as the period between March 17, 2020 and December 31, 2021
- Post-COVID was the period beginning January 1, 2022 and ending on November 7, 2022, which was the day we downloaded the dataset
  - We define post-COVID as after most of the COVID restrictions were lifted, not the end of the pandemic

Questions of Interest

1. How did the COVID-19 pandemic affect the number of thefts from vehicles in San Francisco?
    a. Was there a difference in the monthly average number of thefts for our three time periods, pre-COVID, peak-COVID, and post-COVID?
2. How has the COVID-19 pandemic changed the trends in thefts from vehicles weekly and monthly?
    a. How do trends in thefts from vehicles change between COVID periods for each day of the week?
    b. How do trends in thefts from vehicles change between COVID periods for each day of the month?
3. Is there a socioeconomic trend in the districts that had higher or lower numbers of car thefts?
    a. Specifically, do districts with higher median incomes have high thefts than those with lower median incomes?
4. What are the trends of resolutions for theft from Vehicle in the different police district in San Francisco?
    a. Do districts with higher median incomes have higher resolutions rates than districts with lower median incomes?
    b. What is the overall percentage of reports resolved? Compared to all reports resolved?

Speaker Notes:

- Here are our specific questions of interest, but in the interest of time we'll go into them in more detail in our data analysis section.

03

DATA ANALYSIS
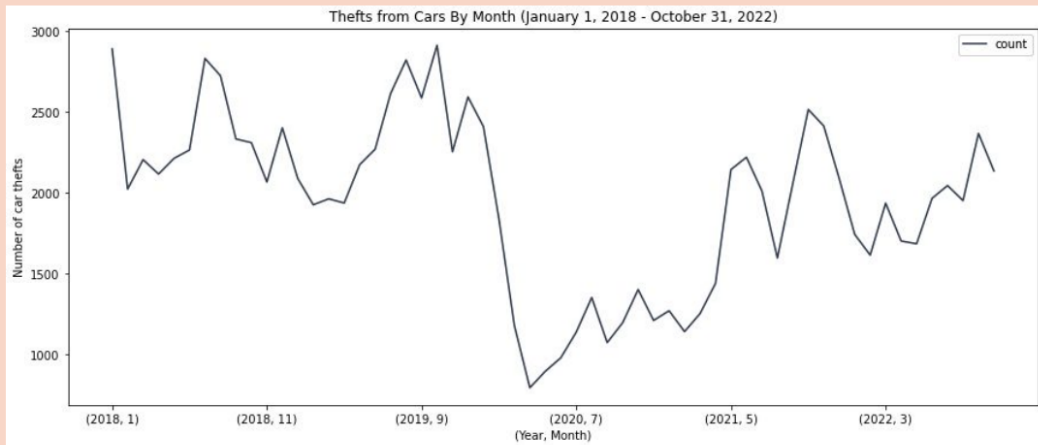
Speaker Notes:

- Our first question was whether or not there was a difference in the monthly average number of thefts for our three time periods
- Our hypothesis was that car thefts increased during the peak-COVID period and have decreased post-COVID, though not to pre-COVID levels
- We first created this graph to show the number of thefts from vehicles by month from January 1, 2018 to October 31, 2022
- November 2022 had the lowest number of thefts because we downloaded the original dataset early in the month, therefore we removed these incidents when creating this graph
- From this graph, we can see that the number of thefts from vehicles began to decline in February 2020 and reached an all-time low in April 2020
- We also see that thefts remained relatively low until around May 2021
- But most importantly, this figure clearly shows a big drop in the number of thefts from vehicles around spring and summer 2020, as well as the fact that thefts are on the rise once again
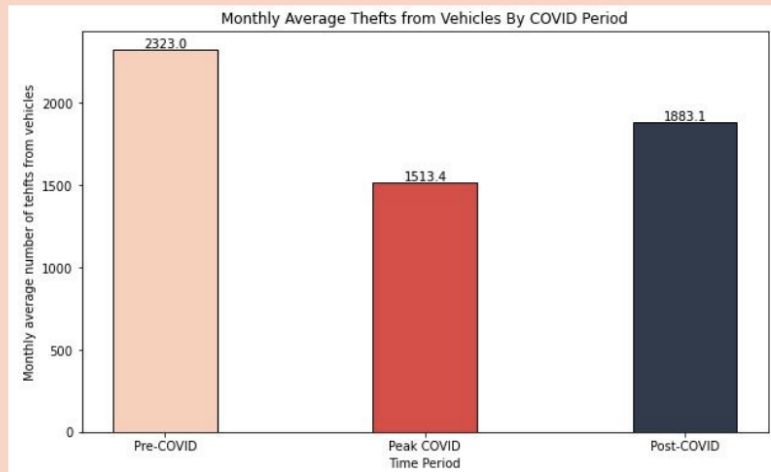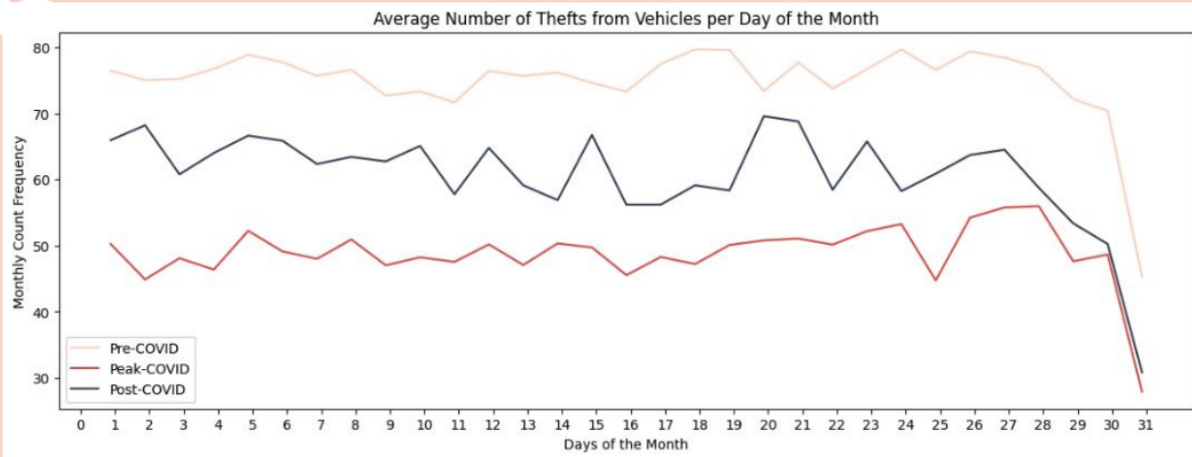
Speaker Notes:

- In order to compare the number of thefts from vehicles between the three time periods, we took the monthly average of each period
- This figure illustrates these monthly averages and from this, we can see that the number of thefts during peak-COVID was considerably lower than that of the pre-COVID period
- We also see that thefts are increasing again now that we are in the post-COVID period, but the average monthly number of thefts has not yet reached the pre-COVID average
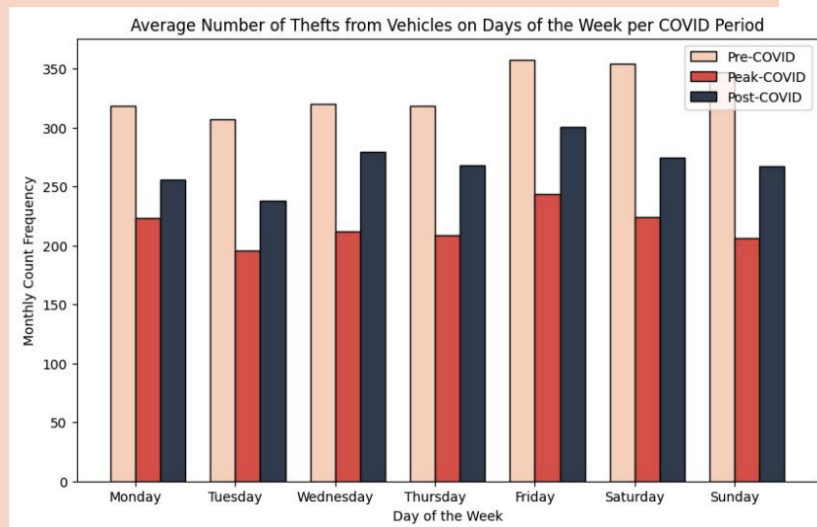
Speaker Notes:

- We wanted to further analyze the monthly and weekly trends in car thefts in San Francisco spanning the three COVID periods previously mentioned.
- First, we took a deeper look into the specific days of the month to observe which days of the month in each period that theft most commonly occurred.
  - We originally hypothesized that thefts from cars would occur more towards the end of the month due to upcoming bills and rent that need to be paid
- A graph of all the days of the month faceted by COVID period was created showing the average amount of thefts from vehicle of that day of the month
- As shown in the figure, the time period that had the highest upticks of crime towards the end of the month was during peak COVID as seen by the dark orange curve.
  - During this period, the highest rates of theft were on the 28th, 27th and 26th. This is most likely due to the reason previously stated in the hypothesis that the end of the month is when bills and rent are due and this most likely worsened due to the increased financial strain of the COVID-19 pandemic
  - Though pre-COVID did show a bit of a trend towards the end of the month, it was not as drastic as peak-COVID.
  - Post-COVID actually showed even less of a trend to the end of the month
- There is a big drop off in thefts on the 31st for all periods which is due to the fact that only 7 months out of the year have a 31st day.
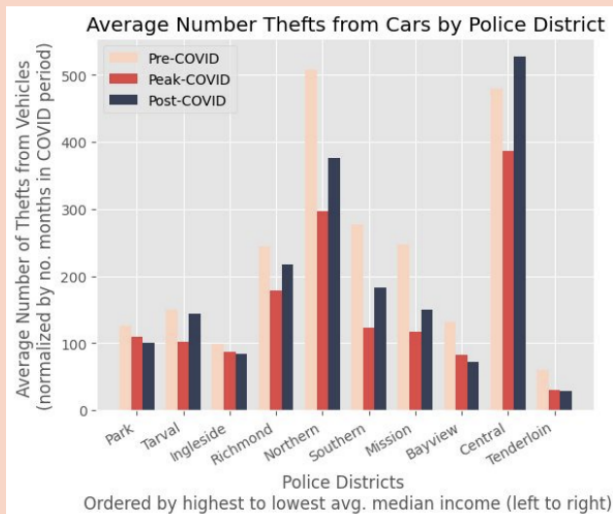
Speaker Notes:

- Next we delved into thefts on different days of the week.
    - We hypothesized that thefts in cars would increase on the weekends as more people and cars are in San Francisco leading to increased car thefts.
    - We also hypothesized that during peak COVID, the theft numbers may have decreased since people were not moving around, even on the weekends.
- This graph shows the average monthly thefts from vehicles over the course of the days of the week, faceted by COVID period
- For all three periods, the day with the highest average number of thefts from cars was Friday, which follows the hypothesis.
- For pre-COVID, all three days of the weekend had the top 3 highest average number of thefts, which was what we predicted. However, for peak-COVID and post-COVID we did not see this pattern, only Friday and Saturday were in the top 3.

Speaker Notes:

- To answer the questions of district trends in SF, we looked at the average no. of car thefts by Police district. Like with the other analysis we normalized the number of thefts by the number of months in each COVID period in order to be able to compare the averages.
- We initially hypothesized to see that districts that had a higher avg. median income would have a higher rate of thefts, but when we looked at the graph we found that there was no strong economic trend.
  - We can see this in the graph as the higher income districts to the left have a lower overall number of thefts, but also so does the tenderloin which is the lowest income district. It does seem like there's a weak trend where the districts with my middling incomes have a higher number of thefts but, this pattern is weak as the central district, one of the lower income districts, has the highest number of thefts overall.
- Across all districts, we see a pattern similar to what we found in the other analyses, where thefts decreased during the peak-COVID period, and then either decreased further post-COVID or increased again post-COVID but not to pre-COVID levels. There also did not appear to be much change in which districts were targeted more during the COVID pandemic.
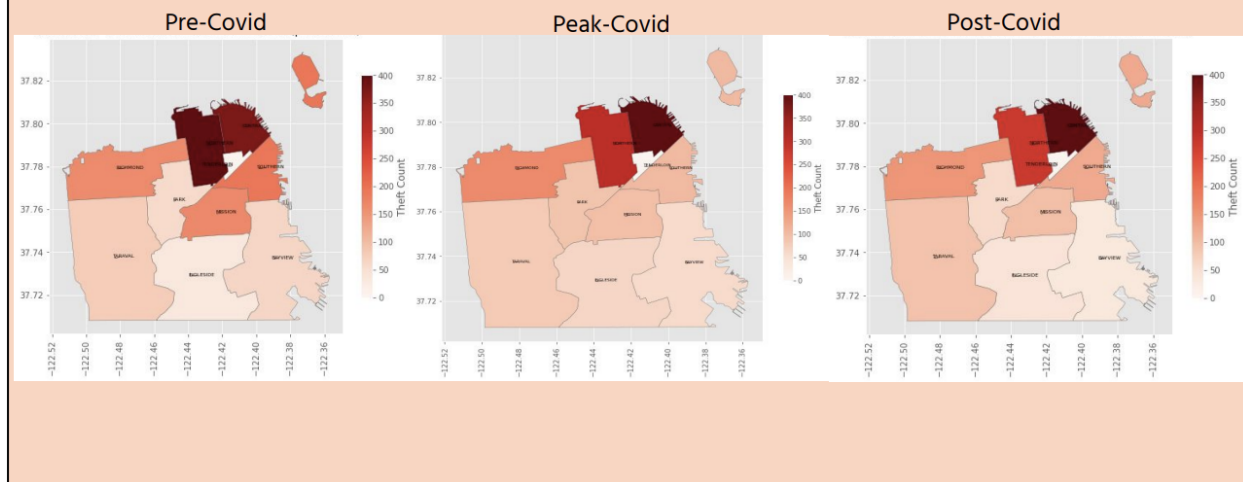
Speaker Notes:

- Here we again, through these heat maps, can see how there was little change across the COVID periods in which districts were targeted.
- Overall, we found there to be little to no socio-economic trend in which districts were targeted. And we have a few possible explanations for this. One that we consider most likely, is that this trend is being obscured by the fact that the districts are of different sizes.
    - Since we were unable to normalize the number of thefts by size of the districts, it stands to reason that the smaller districts, like the tenderloin, would have fewer thefts than the larger districts.
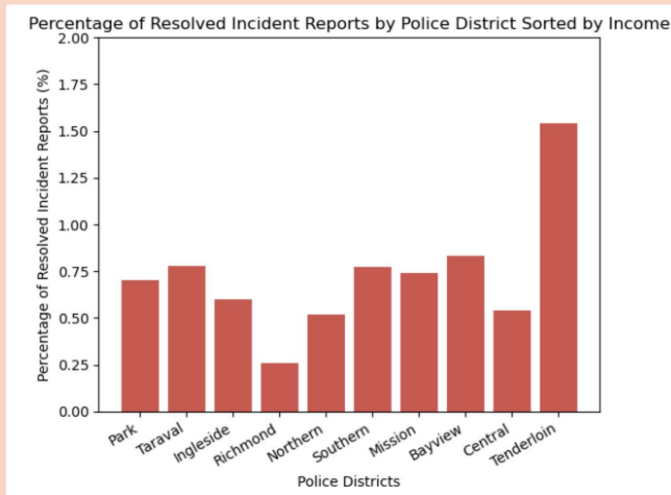
Speaker Notes:

- What are the trends of resolutions in theft from Vehicle by police district in San Francisco?
    - Do districts with higher median incomes have higher resolution rates than districts with lower median incomes?
    - What is the overall percentage of reports resolved? Compared to all reports resolved?
- Lastly, to address our final question, we looked at the percentage of "Resolved" incident reports by police district
- As we initially hypothesized that higher income police districts would have higher resolutions percentages, the bar chart, sorts police districts from highest to lowest median income, which does not reveal any clear relationship between median income and resolution rate for the districts
    - First off, the tenderloin with the highest percentage of resolved cases, around 1.5%, has the lowest median income.
    - Park, the district with the highest median income, at 90,000 dollars, has only .52 percent of resolved cases, which is lower than districts Bayview, Mission, Tenderloin, and Southerm that all have lower median incomes.
- Another important note to be made about the graph, is that you can see the range on the y axis goes from 0 to 2 %, this small range might not be meaningful to determine why one district has a higher percentage than the other. Therefore, it might be important to look further into trends of resolutions on a different scale other than income and theft from vehicles.

Speaker Notes:

- This takes me into my next point, that the percentage of overall arrests from theft from vehicles is only .58%, whereas the percentage of arrests for all crimes in SF is 19.32%.
- This could be due to the fact that larceny from vehicles is harder to resolve than most other crimes, as the discovery of the crime is delayed, which could explain why resolutions from district to district did not vary by a lot.

CONCLUSIONS & FUTURE EXPLORATIONS

04

Conclusion and Future Explorations

- Limitations
  - Timing of incidents is a rough estimate
  - Assumption: people notice and report car break-ins relatively quickly on average
- Local government can use studies like ours for crime prevention strategies, make plans for different types of surveillance, erect warning signs, etc.
- Future explorations
  - Types of cars being target and their shared features (makes/models, if the windows are tinted, if there were items in the backseat, etc.)
  - Location of crimes (residential versus commercial areas)
  - How much of an issue do SF residents perceive the crime to be?

Speaker Notes:

- And finally, to wrap up our presentation:
- The biggest limitation was that the
  - Timing of the incidents is a rough estimate
  - We are assuming that people notice and report car break-ins relatively quickly on average
- We believe that the local government can use studies like ours for crime prevention strategies, make plans for different types of surveillance, erect warning signs, etc.
- Some future areas of explorations include
  - Types of cars being targeted and their shared features (makes/models, if the windows are tinted, if there were items in the backseat, etc.)
  - We can also look into the location of these crimes (residential versus commercial areas) ○ Lastly, another area of interest would be how much of an issue SF residents perceive this crime to be