Finding: Larceny-thefts from vehicle are the most common crime incidents in the streets of Southern district near to downtown of the San Francisco

Introduction

This project analyzes the San Francisco Police Department crime data between June 1st 2014 and August 31st 2014. It visualizes the trends of larceny-thefts in the city across time and location. The primary purpose is to understand, the pattern of larceny-theft, when, where they occur most and to determine the most common type of larceny-theft incidents across all districts and the district where it occur most and where.

Methodology

Python Pandas, and Bokeh libraries were used for data analysis and visualization. There were 28993 entries in the original dataset. Duplicates that shared the same incident numbers were removed, leaving 23008 entries in the dataset.

Definition of Larceny:

According to FBI's Uniform Crime Reporting (UCR) Program (UCR Offense Definitions, 2017), larceny-theft is defined as *"the unlawful taking, carrying, leading, or riding away of property from the possession or constructive possession of another"*. Few example are such as, thefts of bicycles, motor vehicle parts and accessories, shoplifting, pocket-picking etc.

Exploratory Analysis of Crime Incidents

Table 1, a bar chart showing crime distribution by crime category suggests that during summer of 2014, Larceny-theft was the most common crime with a total of 9191 reported incidents i.e. 40% of total crime incidents. The next highest crime categories, excluding "Other offenses" (2415 incidents) and "Non-criminal" (2700 incidents), were "Assault" (2162 incidents) and "Vehicle Thefts" (1812 incidents)

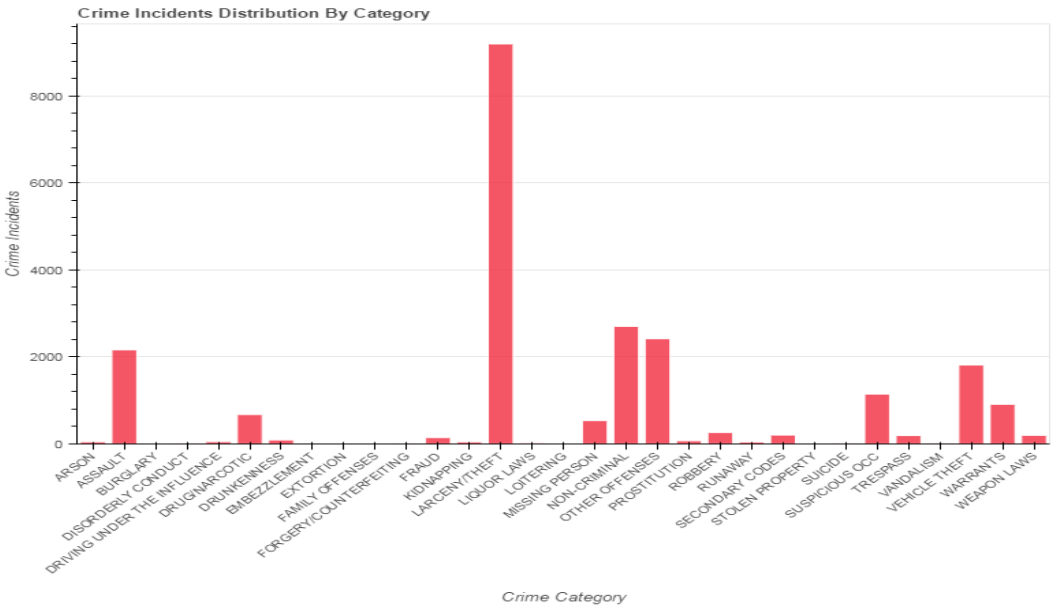


Table 1 Crime Distribution by Crime Category

What days, time and area are particularly unsafe?

To analyze the above question, heat map (Table 2) has been plotted to distribute crime incidents across police district and weekdays. Furthermore, bar chart (Table 3) and time series (Table 4) has been plotted to analyze the relationship between time period in a day and crime rate.

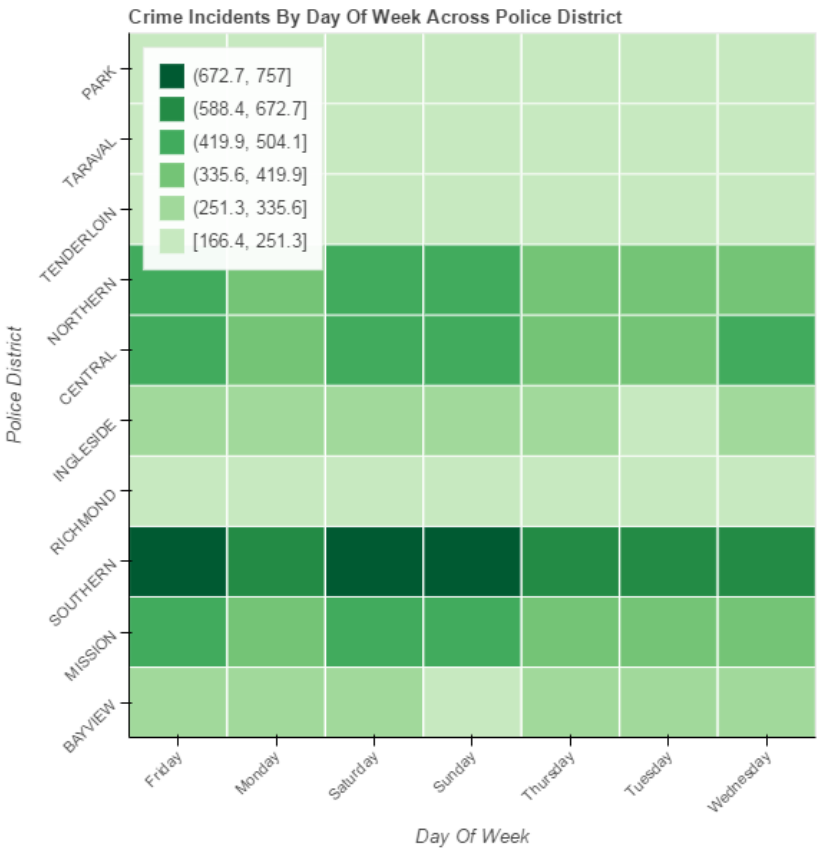


Table 2 Heat map of crime incidents in Police district across weekday

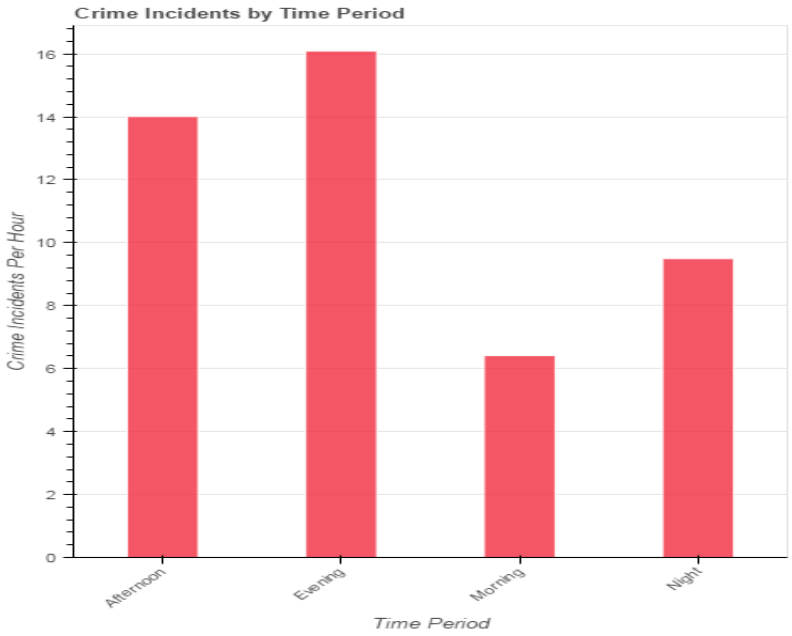


Table 3 Crime rate distribution across Time period

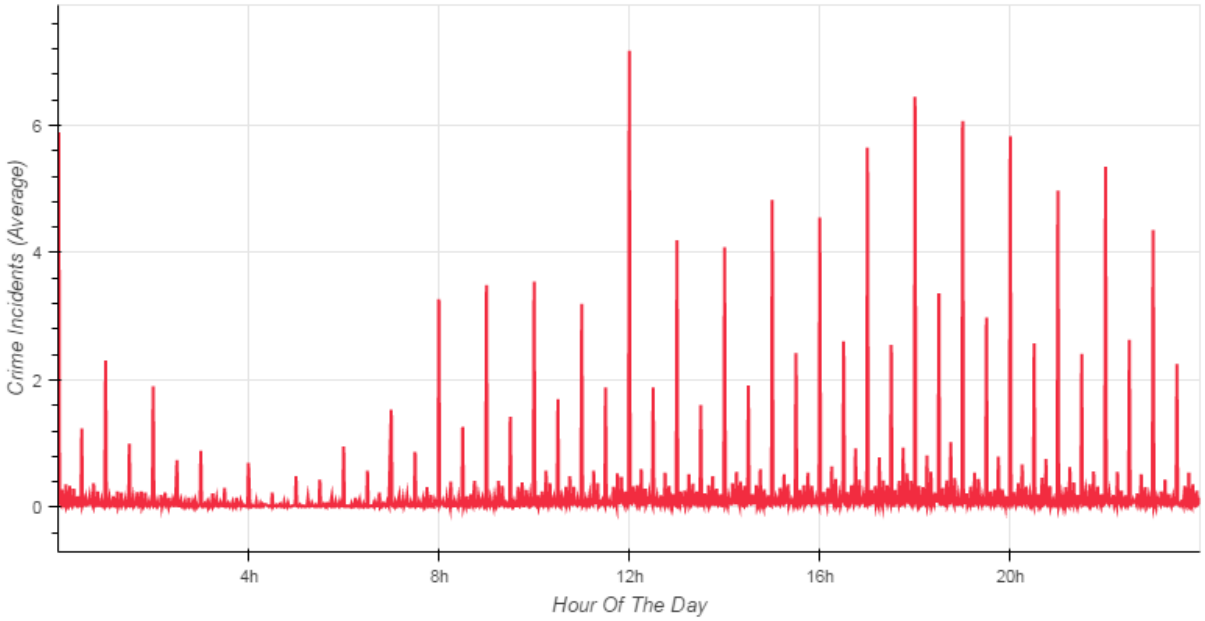


Table 4 Crime incidents across 24 hours

It seems crimes are fairly distributed around the week. However, there was a noticeable increase in the frequency of crimes during Weekend. On any day, most of the crime incidents happened during afternoon and evening. Night time seems also unsafe, while morning time seems pretty quiet. Time series of crime incidents across hour of the day also reasserts that crime incidents spike during afternoon and goes quite during early morning.

Data suggests weekends of summer 2014 were unsafe across all police districts. However Southern district was the crime magnet, where crime rate was quite high. Surprisingly, Tenderloin district, which is bordering Southern district seems to be quite safer. However since Tenderloin district is quite small, crime rate per square mile would have provided better insight, a candidate for future project. Interestingly, Bayview is the only district where Sunday is quite safe in comparison to other days.

How crimes are distributed over the police district?

While trying to answer the above question, low crime rate categories has been filtered out of dataset to get a better insight. Heat map (Table 5) of crime's category across police district suggests that most of the crime incidents were fairly distributed across district, however, Larceny-theft was quite significant in Southern District. Furthermore, Assault and Non-Criminal crimes also happens noticeably higher in Southern District.

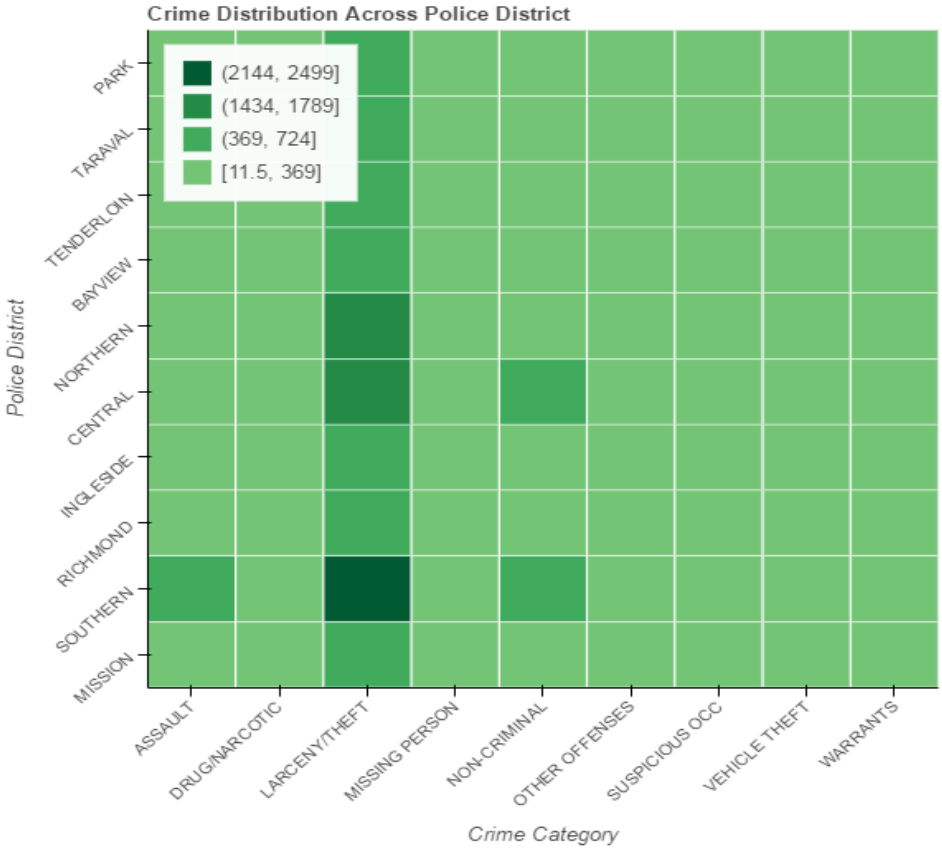


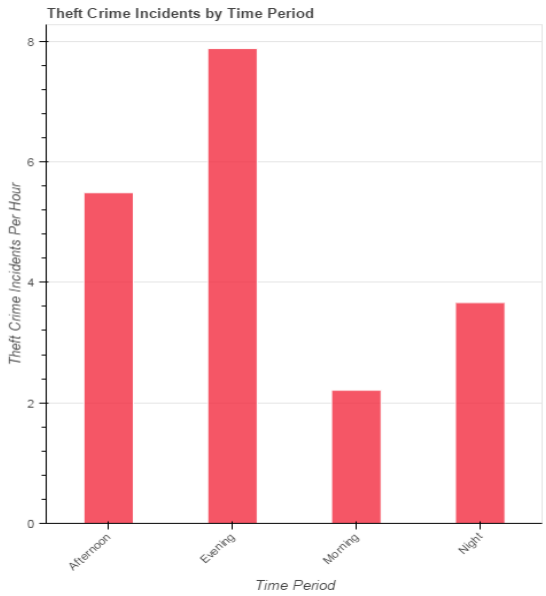
Table 5 Heat map of crime incidents for police district across crime categories

When Larceny-theft occur most? Is it different from the overall crime incidents trend?

To answer these questions, Bar chart (Table 6) has been plotted to analyze the relationship between time period in a day and larceny-theft crime rate

As Larceny-theft was the major part of all crimes incidents, it would be obvious that it will follow the similar pattern of overall crime incidents. During any day, similar to overall crimes, larceny-thefts happened mostly in evening, afternoon also seen an average of 5.5 thefts per hour, however, it was 70% of thefts that happened in the evening. Where as the overall crime incidents rate were nearly similar in afternoon and evening (ratio is 89%). Incidents of larceny-theft failry occurs daily, however weekends have seen a spike in these thefts.

Furthremore timeseries(Table 7) of overall crime and larceny-thefts over the summer, reasserts that larceny-theft is the primary crime, which structure the pattern of overall crime incidents, however there are periods where larceny-thefts dip and overall crime spikes and vice-versa.



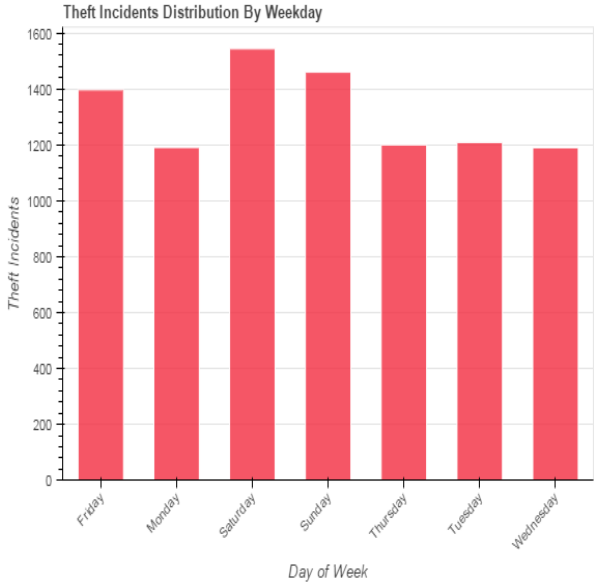


Table 6 Larceny-thefts crime distribution by time period and weekday

Table 7 Total Crimes vs Larceny-thefts Time series

What is the most common type of Larceny-theft and where it occur most

Larceny-theft incidents distribution by theft type (Table 8) suggests that most of the theft were 'GRAND THEFT FROM LOCKED AUTO', second most common type of theft was 'GRAND THEFT FROM UNLOCKED AUTO' were property related. Other auto thefts incident were also frequent and property related incidents were also noticeable.

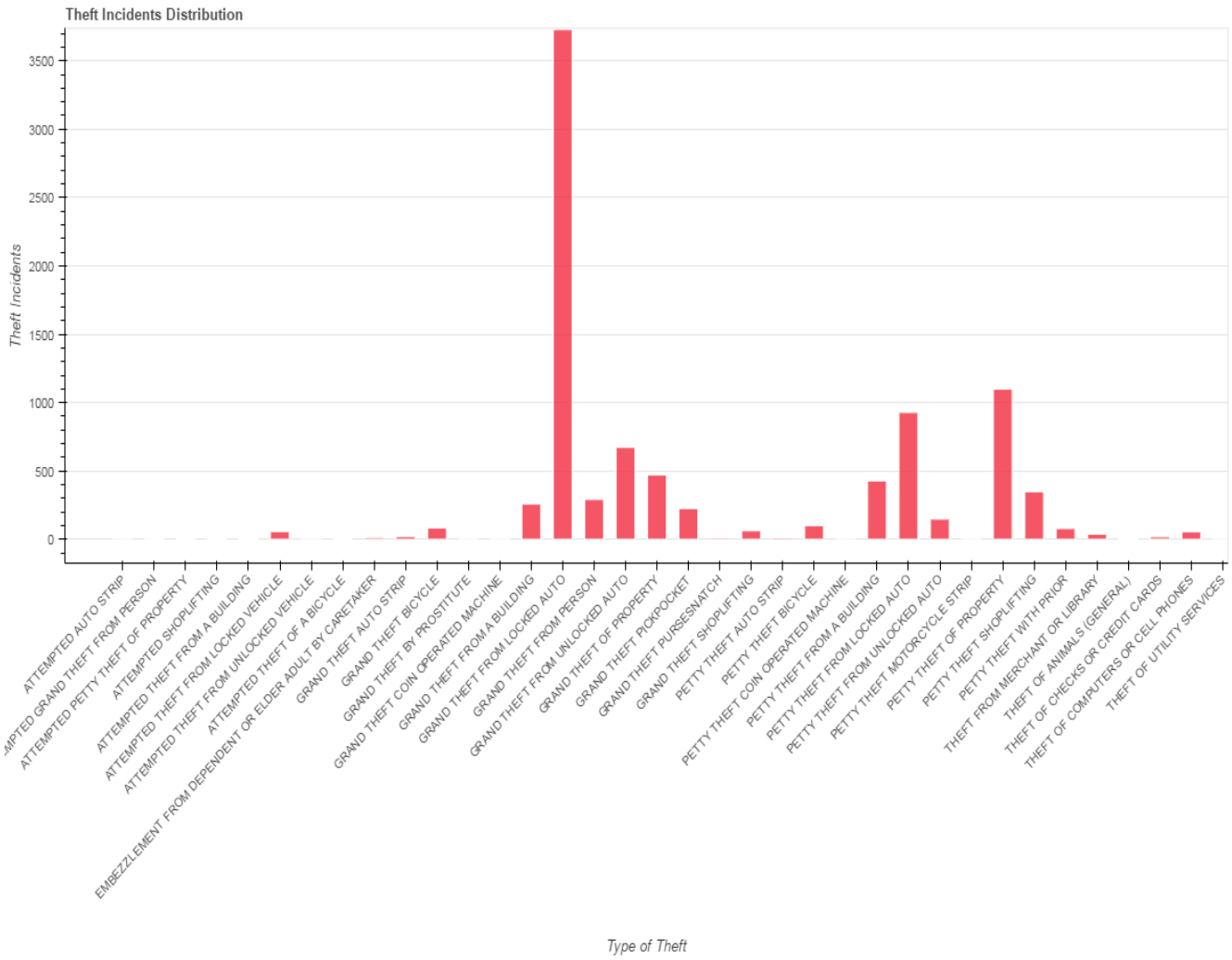


Table 8 Larceny-theft Distribution by theft type

In the effort to further analyze the distribution, data has been grouped so that related thefts are grouped under one description, such as ('GRAND THEFT FROM LOCKED AUTO','GRAND THEFT FROM UNLOCKED AUTO') etc. are grouped under Auto. Following categories are created:

a) Auto
b) Building
c) Shoplifting
d) Person
e) Pick Pocket
f) Others (all other larceny-thefts incident)

Pie chart (Table 9) has been plotted, which reasserted that auto related thefts dominated the Larceny-thefts and second most frequently occurring crimes were property related.

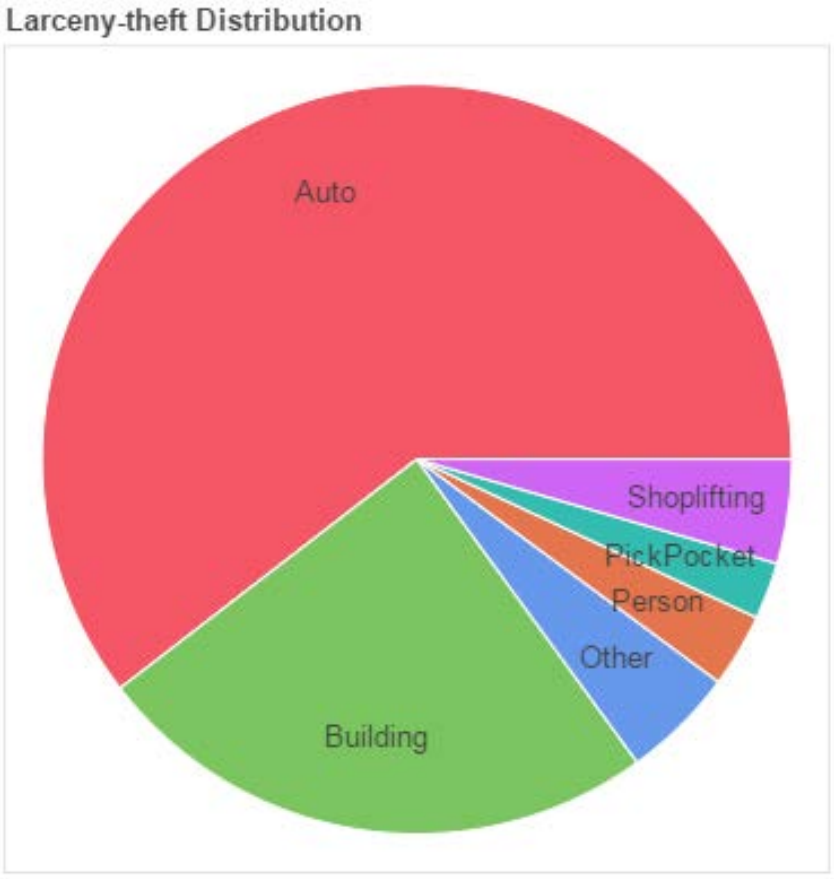


Table 9 Larceny-theft distribution by theft type

Where auto related larceny-thefts occurred most and what kind of auto theft dominates overall auto thefts?

To respond to the above query, heat map (Table 10) had been plotted for auto related larceny-theft across police district. Analysis of this heat map suggests that most of the auto related larceny thefts occur in Southern district and it seems to be the most unsafe place for vehicles, as interestingly, major portion of auto thefts incidents include locked vehicles, that means criminals broke either lock or windows to break in.

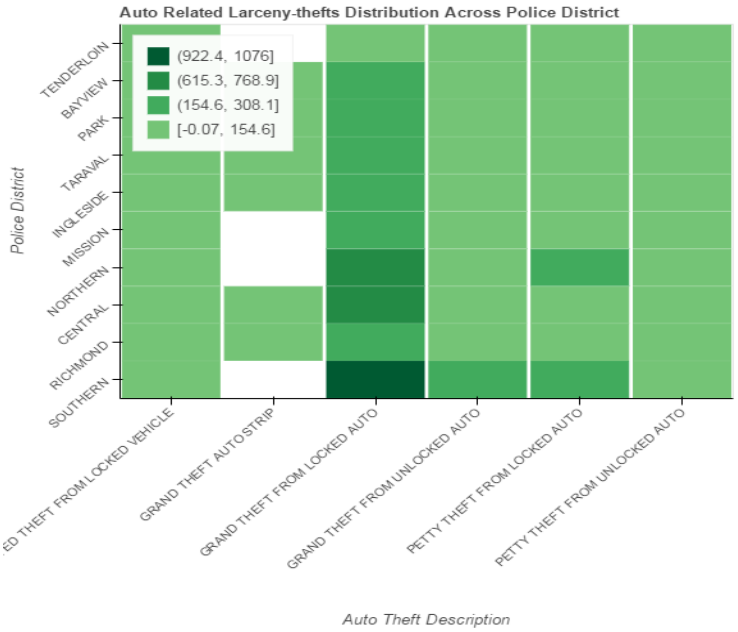


Table 10 Auto related larceny-thefts distribution across district

Now it is clear that auto related thefts happened frequently in Southern district, another question arise, whether there is any specific area in Southern district where this kind of thefts happened frequently. To understand the same, auto related larceny-thefts data has been plotted upon Southern district map (Table 11) using Google Map APIs, it is now clear that most of the auto thefts happened in the streets near the downtown area.

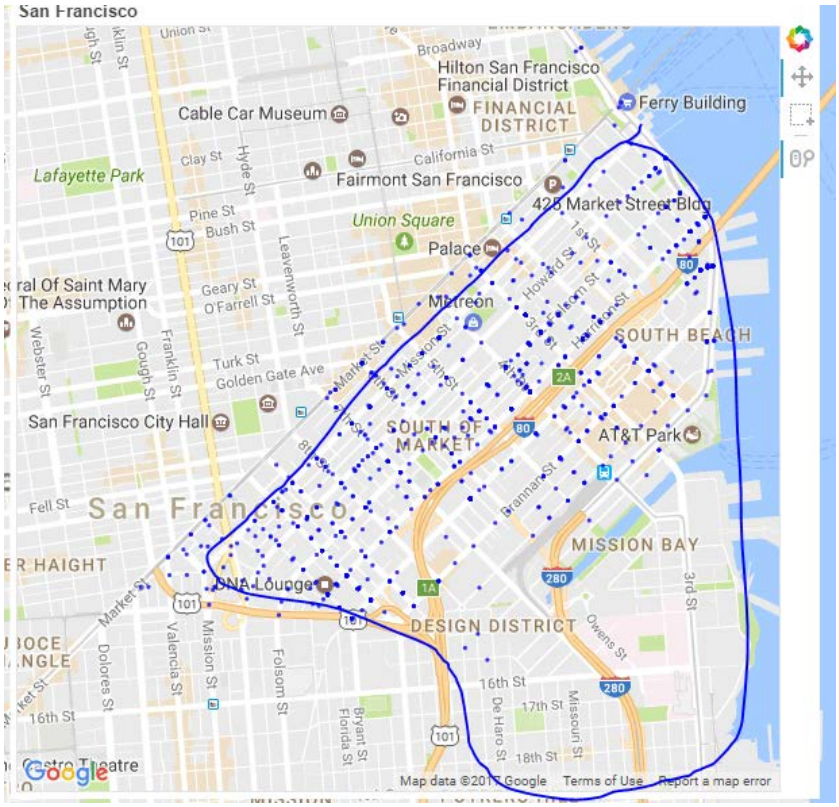


Table 11 Auto larceny thefts distribution in Southern District

In summary, analysis of San Francisco crime data (summer 2014) suggests that southern district is a dangerous area with a high crime rate, people with vehicle needs to be extra cautious as auto related thefts are the most common crime incidents in this area and they should be smart and vigilant while parking their vehicle.

References

UCR Offense Definitions (Jan 26, 2017). Retrieved Feb 5, 2017, from https://www.ucrdatatool.gov/offenses.cfm