

Crime Analysis: San Francisco

STAT 5702

Report submitted by:

Vibhuti Mahajan (vm2486), Sheallika Singh (ss5136) , Tejas Dharamsi (td2520)

INTRODUCTION

Crime analysis is a law enforcement function, that involves systematic analysis for identifying and analyzing patterns and trends in crime and disorder. This information on patterns can help law enforcement agencies deploy resources, in a more effective manner, and assist detectives in identifying and apprehending suspects. In the recent years, the easy availability of open source police data, has added a dimension of public interest where crime rate inversely contributes to the living standards in an area. Such analysis also aids people in making informed decisions while they move to new places to live, by presenting a broader picture of civil disorder and police proactivity in that area.

Although the main aim of crime analysis is the formulation of crime solving strategies and general administrative policies, but we limit our research to exploratory data analysis and visualisation. We present our analysis and findings on San Francisco Police Department's Crime Report Data available on SF OpenData Website. We have exported the data starting from year 2003 to 2016 as a spreadsheet and have done entire analysis in R notebooks.

- Data Link: <https://data.sfgov.org/Public-Safety/Police-Department-Incidents/tmnf-yvry>
- Keywords: *patterns, trends, exploratory data analysis and visualisation, factors*

TEAM

Team Member	Contribution
Tejas Dharamsi	Spatial and time analysis of theft and larceny type of crimes (1.nb.html)
Vibhuti Mahajan	Spatial and time analysis of various drug related crimes reported (2.nb.html)
Sheallika Singh	Spatial and time analysis of vehicle theft (3.nb.html)

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

DATA QUALITY

The data we are working on is San Francisco Police Department crime reports data. Each recording in the data has following attributes:

IncidentNum: incident id

Category: Crime category

Descript: short description of the crime reported

Day of week, Date, Time, location of the crime place (in terms of longitude, latitude and address accurate upto the street level)

PdDistrict: police department district where the crime was reported

Missing data Analysis (Plot can be viewed in 3.nb.html):

The following plot shows that the data has one missing value for column PdDistrict, which records police department district where the crime was reported. The histogram to the left shows the proportion of the missing values.

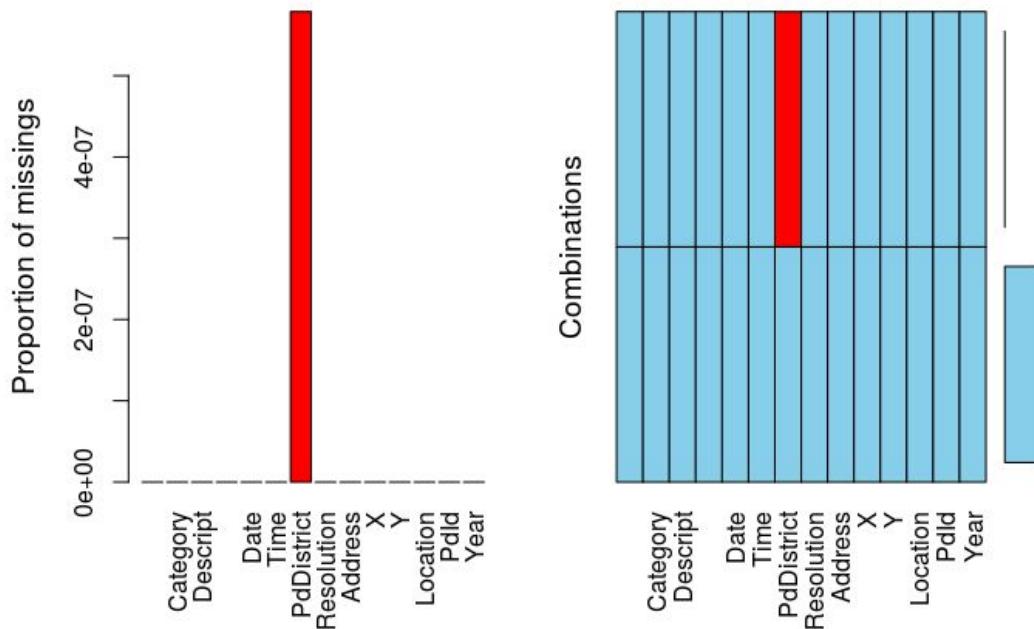


Fig 3.1.

The following table has the details for the proportions of missing values:

(Plot can be viewed in 3.nb.html)

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Fig 3.2

The data is updated to the current date that is why for year 2017 we have observations only till April. So we had to exclude the data for year 2017 as it would not give the correct interpretations And trends. There can be seen a dip in the crime reports for year 2017 but that is because the data is not for the whole year.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other browser.

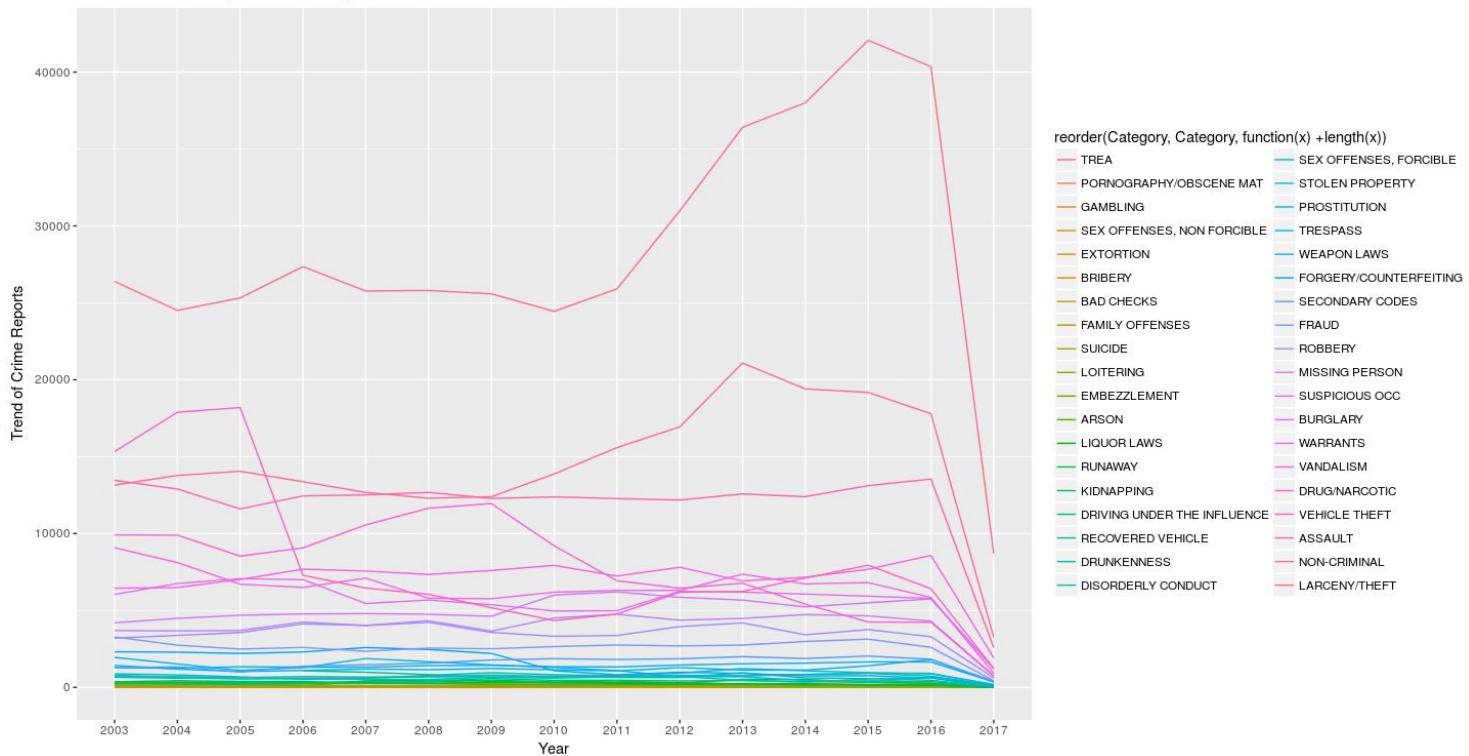
Link to Github Repository: <https://github.com/Dinaramsitas/Spring-EDAV2017/>

Link to notebook 1.nb.html: <https://dharamsiteas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html: <https://dharamsiteas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook Z.nb.html : <https://dharamsiteas.github.io/Spring-EDA-V2017/Z.nb.html>

Count of Crime reports over the years



(Plot can be viewed in 3.nb.html) Fig 3.3

Since the data is police records for the first information reports of the crime. The data does not have the information how the crime was resolved. Most of the reports have "None" as the resolution type. Therefore, it is difficult to tell apart that whether the report was false or considered into other crime category than that of original.

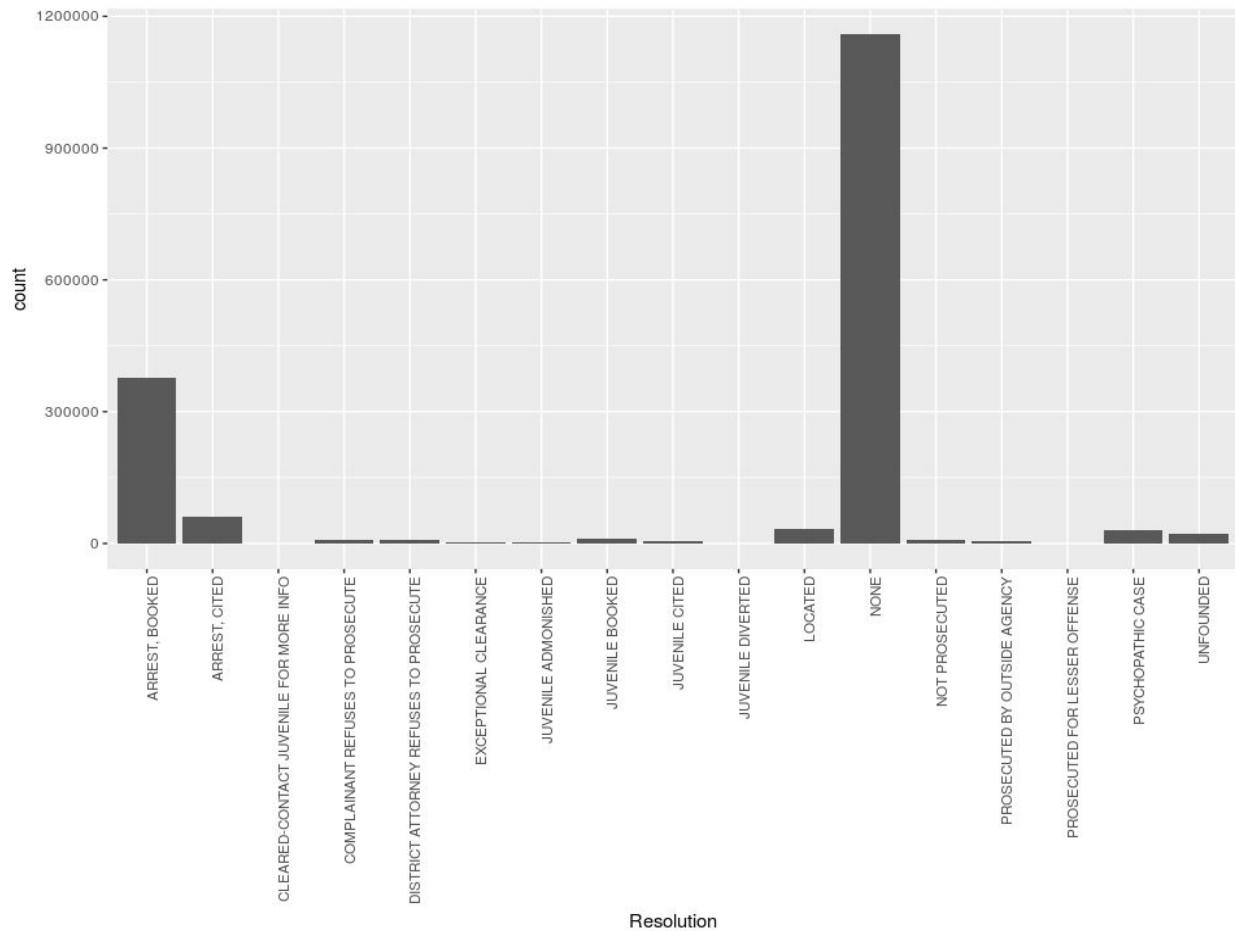
NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>



(Plot can be viewed in 3.nb.html) Fig 3.4

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

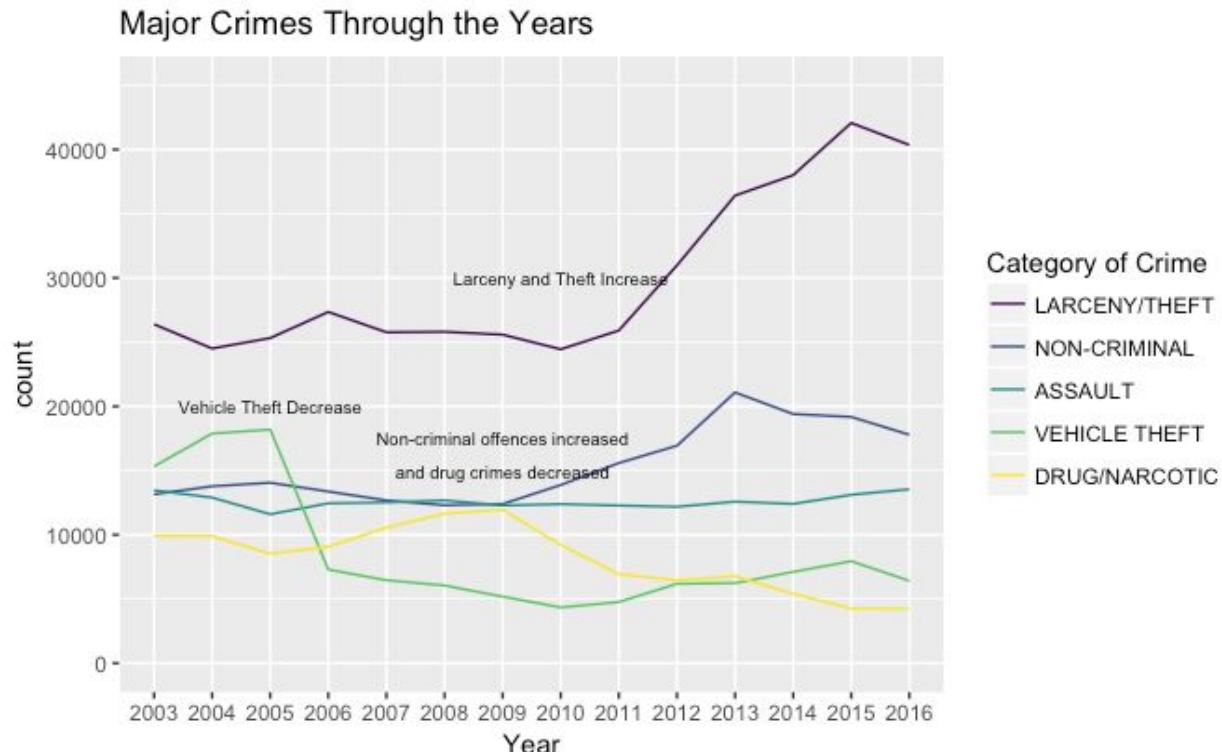
Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

EXECUTIVE SUMMARY

The nature of our data gives us a highly pliable precinct to perform time series analysis and heatmap analysis.

1. Time Analysis:



(Plot can be viewed in 2.nb.html) Fig 2.1

- This plot shows the trend in different categories of crime over the years. Here, we are showing only the major 5 contributors in the total count of crime for easy visualisation. The major 5 crime categories are: LARCENY/ THEFT, NON-CRIMINAL, ASSAULT, VEHICLE THEFT and DRUG/ NARCOTIC.
 - Larceny and Theft are the most common crimes reported. Such reports have increased by ~70% from 2010 to 2015.
 - Non-criminal offences also rose in 2009. The number of assaults reported have remained more or less similar over the years.
 - Vehicle theft and crimes related to drug/narcotics have shown more interesting trends where the reported number of felonies committed have decreased. Vehicle theft reduced in 2005 and drug crimes report reduced in 2009.
-

2. Heat Map Analysis:

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

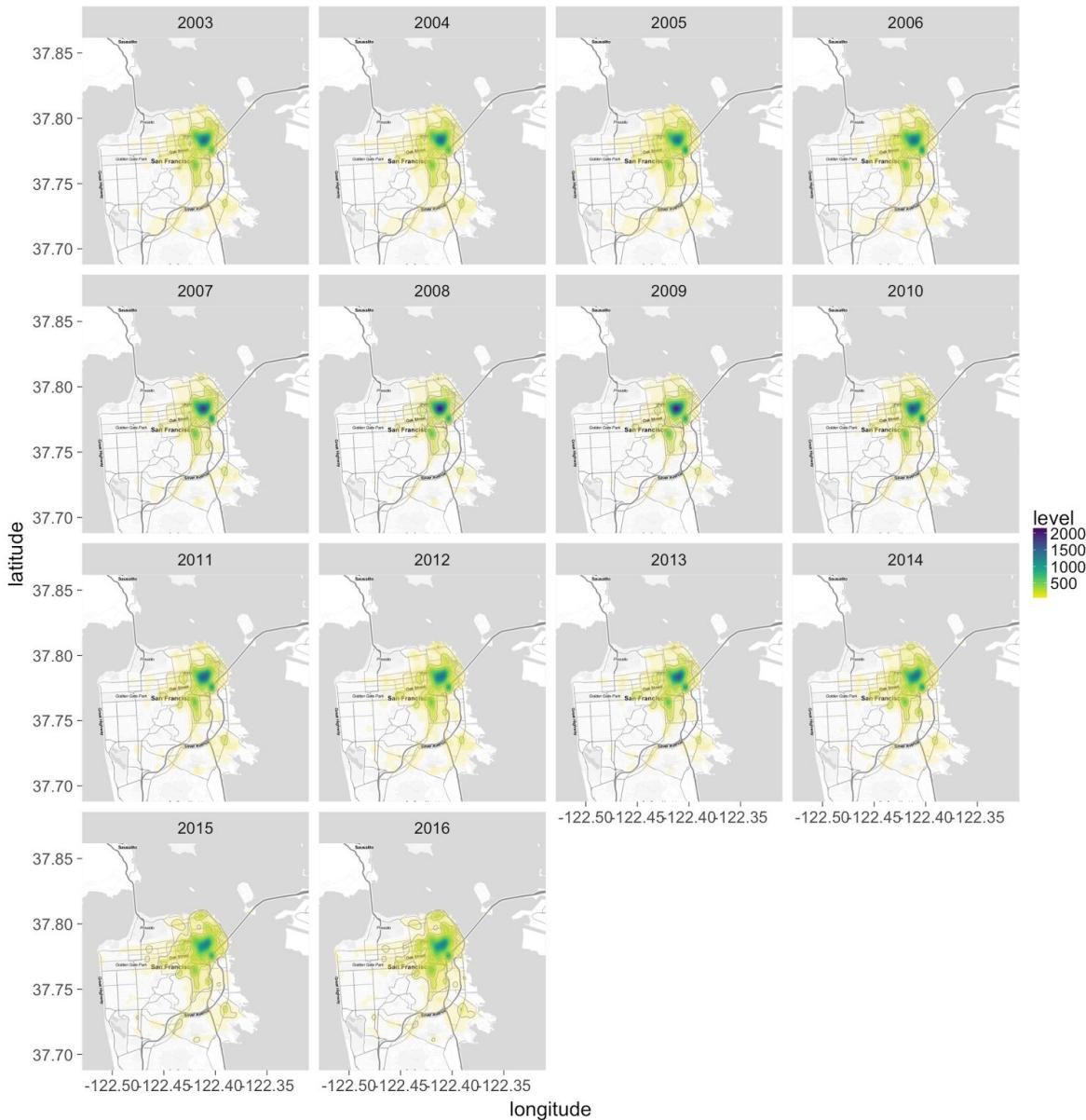
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Crime Report Density : Are Southern and Mission Districts Unsafe?



(Plot can be viewed in 2.nb.html) Fig 2.2

- This graph shows the density of crime instances reported as they spread over San Francisco and faceted over the years.
- Most of the the reports (and simultaneously crimes) are mostly from 4 districts: Southern, Northern, Central and Mission.
- This is due to the fact that these areas are densely populated.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

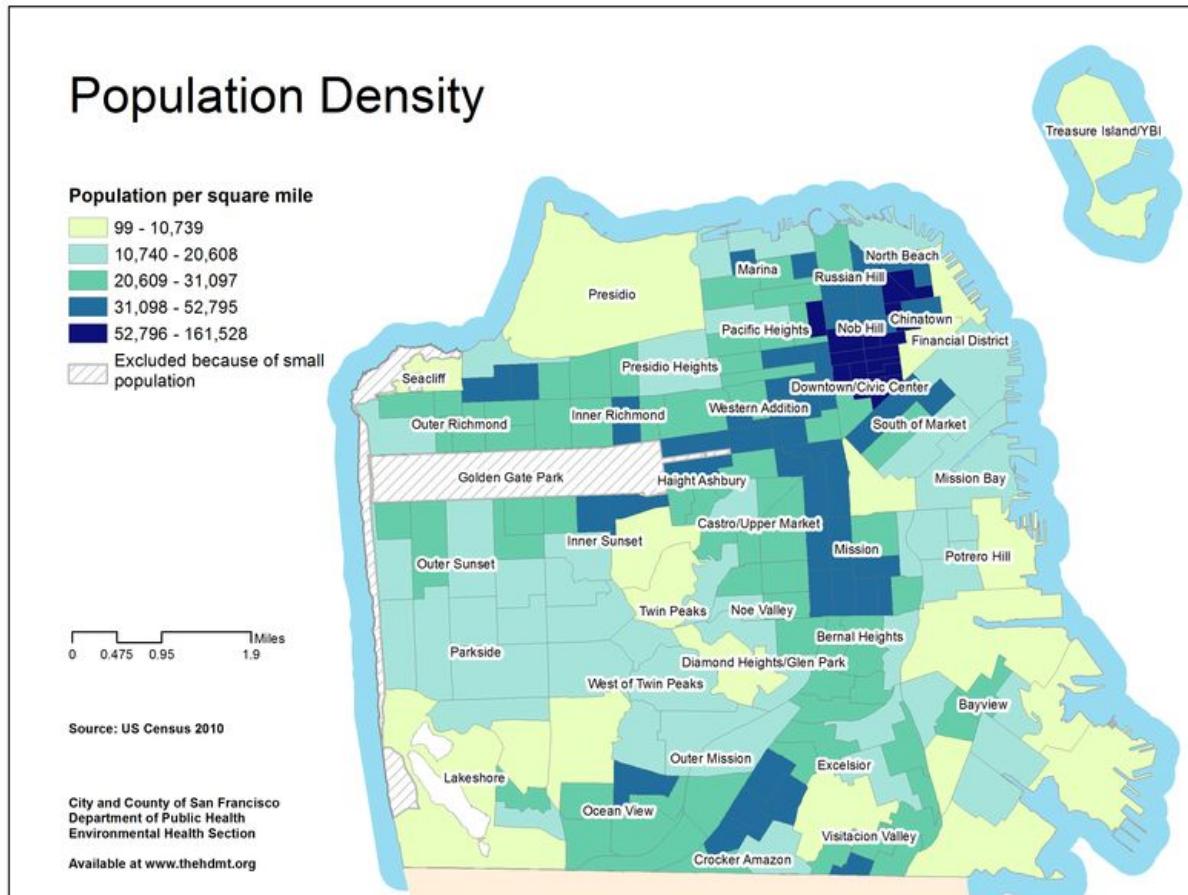
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Reference (link : <http://www.sfindicatorproject.org/indicators/view/183>):



3. Lesser Crimes on Christmas

- It has been observed from the bar plot below (faceted over month) that there is a dip in crime reportings on christmas eve.

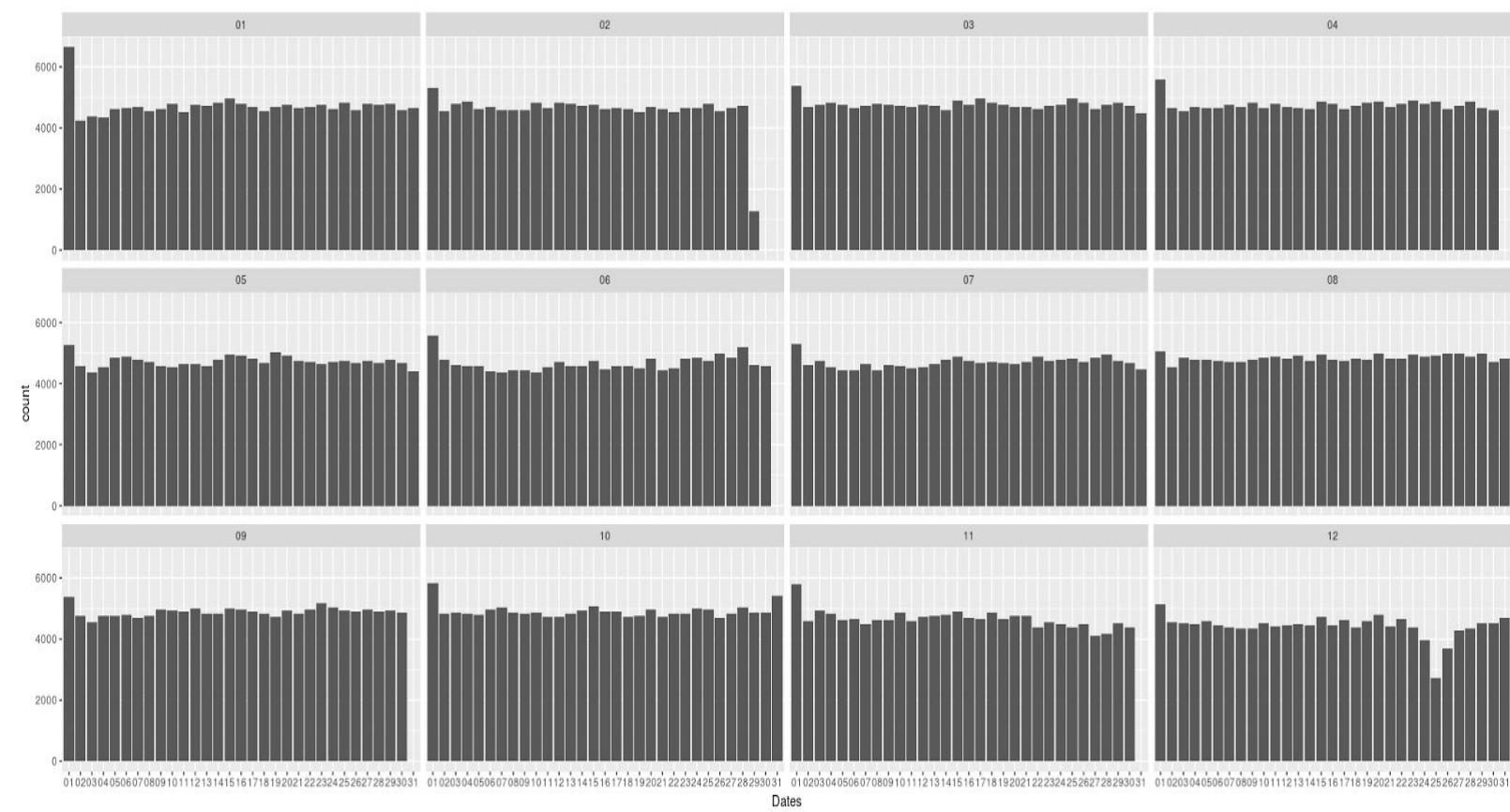
NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>



(Plot can be viewed in 3.nb.html) Fig 3.5

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

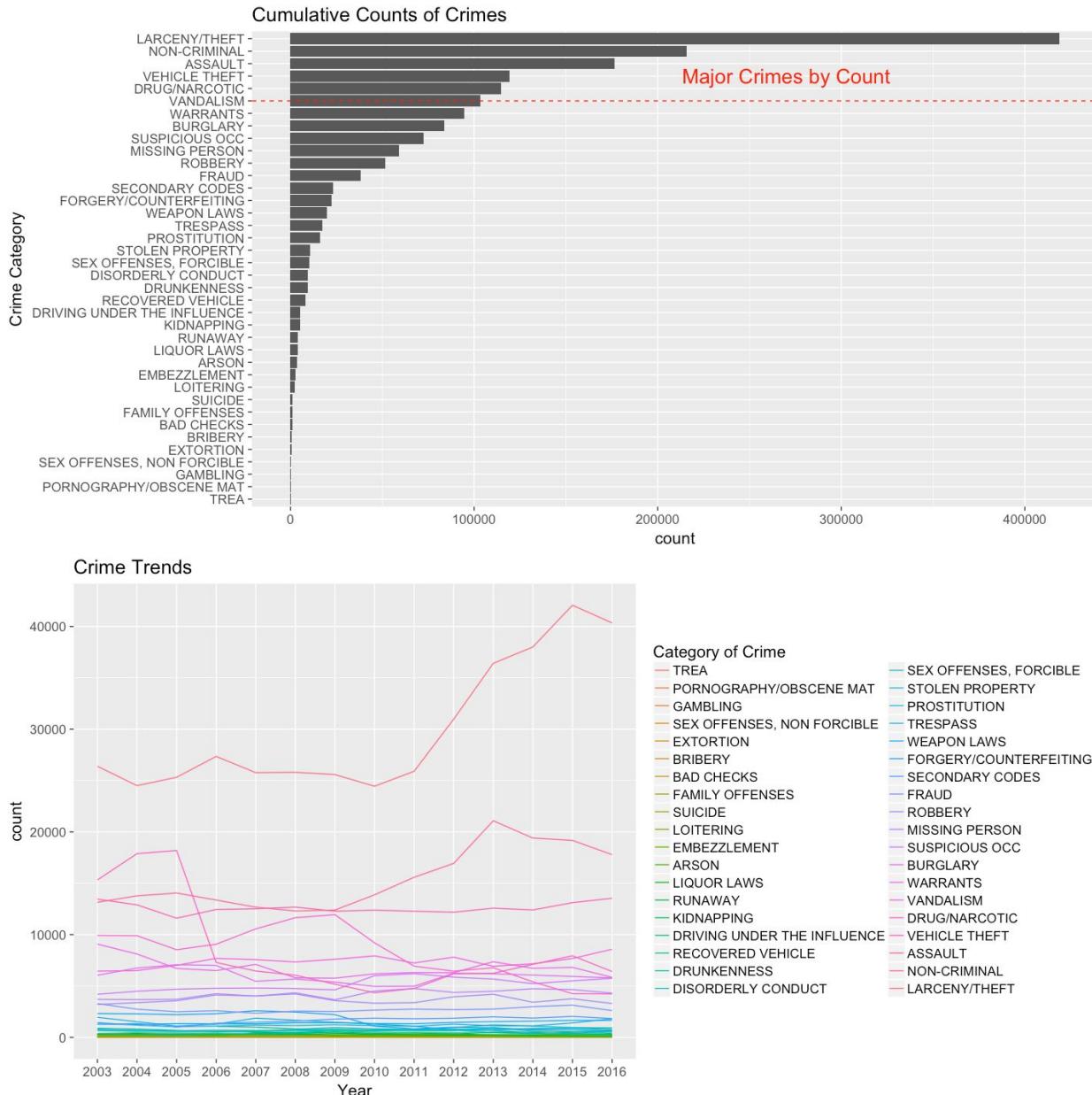
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

DETAILED ANALYSIS



(Plot can be viewed in 2.nb.html) Fig 2.3

- The bar plot above shows crime reports for different categories of crime. It can be observed that LARCENY/ THEFT, NON-CRIMINAL, ASSAULT, VEHICLE THEFT and DRUG/ NARCOTIC are the top 5 crimes reported.
- From the time series plot of the crime reports for different categories it can be seen that out of the top 5 crimes there is a peculiar trend from LARCENY/ THEFT, NON-CRIMINAL , VEHICLE THEFT and DRUG/ NARCOTIC.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

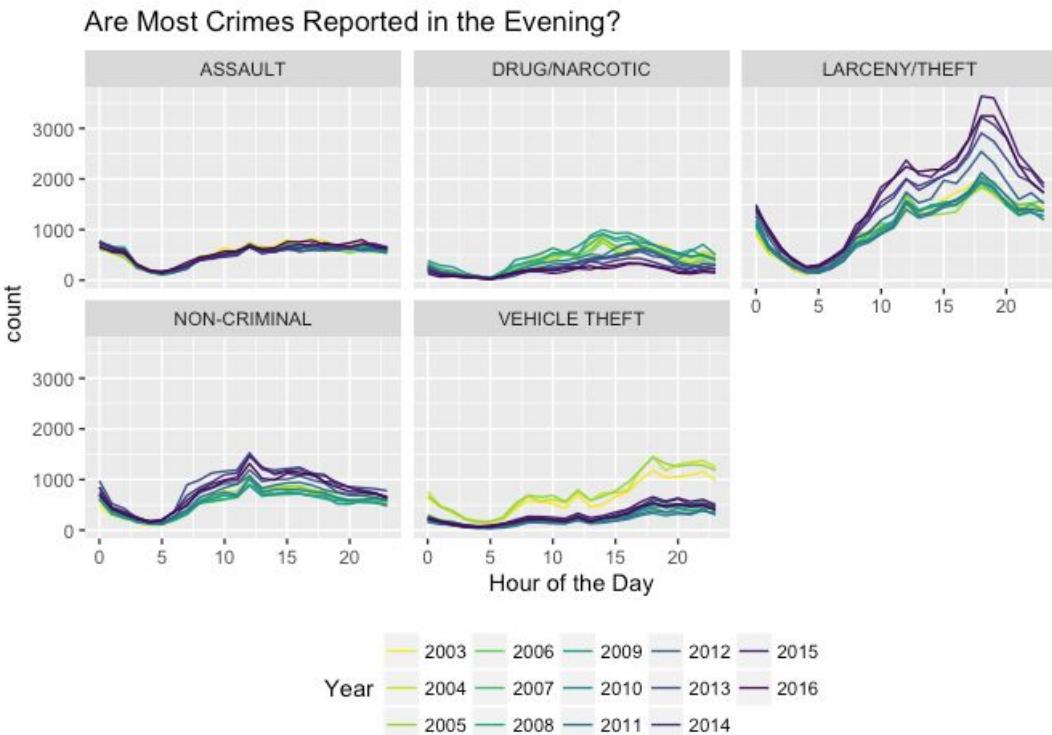
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

- LARCENY/ THEFT: there is an increase in the crime reports after 2011
 - NON-CRIMINAL: There is an increase in the crime reports after 2009
 - DRUG / NARCOTIC: Decrease in crime reports after 2009
 - VEHICLE THEFT: Drop in crime reports from 2006
- For other categories of crime no peculiar trend has been observed. We have taken Larceny/Theft, Vehicle Theft and Drug/ Narcotic for detailed analysis to explore the reason behind the trends.
-



(Plot can be viewed in 2.nb.html) Fig 2.4

- Here we present the daily trends of crime reports colour coded by year.
- Most of the reports for crimes involving a suspect like theft , assault, larceny and vehicle theft are reported in the evening (around 7-8 pm). This might imply that most of such crimes are committed during late afternoon or early evening hours when people usually commute from their workplaces.
- 5 in the morning sees the least number of reports.
- Non-criminal and drug related crimes have similar kind of reporting throughout the day.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

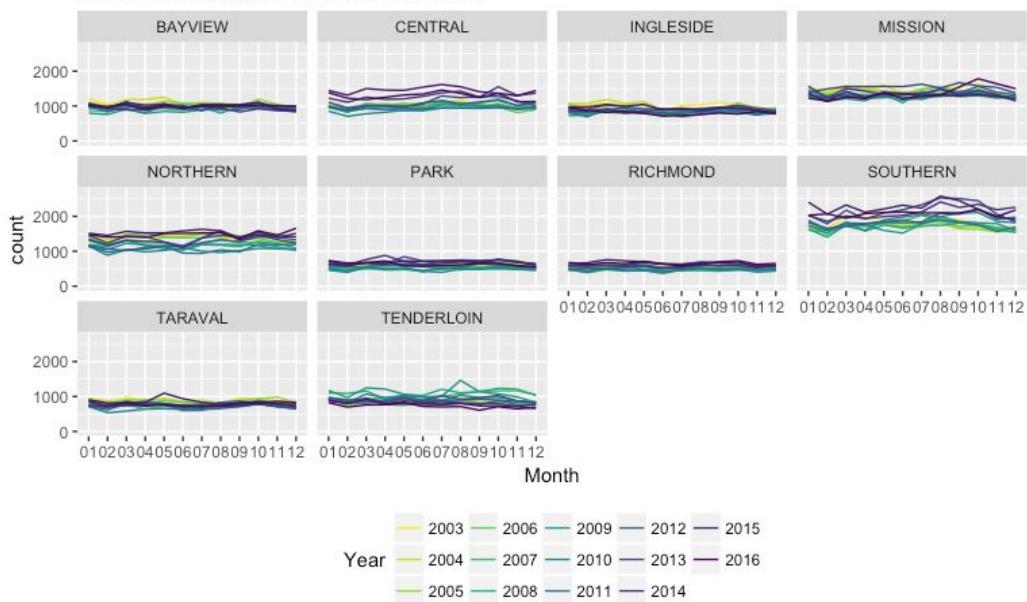
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

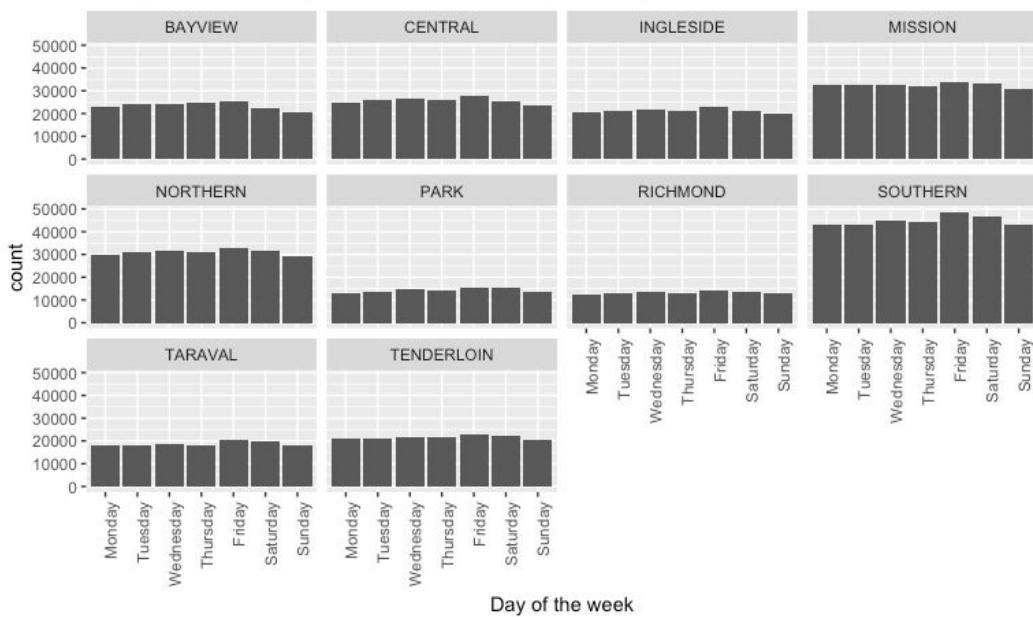
Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Crime Trends over Months are Similar!



(Plot can be viewed in 2.nb.html) Fig 2.5

Fridays and Saturdays : Most Crime Susceptible Days



(Plot can be viewed in 2.nb.html) Fig 2.6

- The first plot shows how the crime reports differ in different districts over the months where different colours correspond to different years (darker being the latest). The reports over the months in different districts have almost the same pattern throughout the districts.
- The second plot shows how crime reports differ over various day of the week, faceted over districts. The trend is same over the days but reports on Fridays and Saturdays are somewhat more than the other days.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

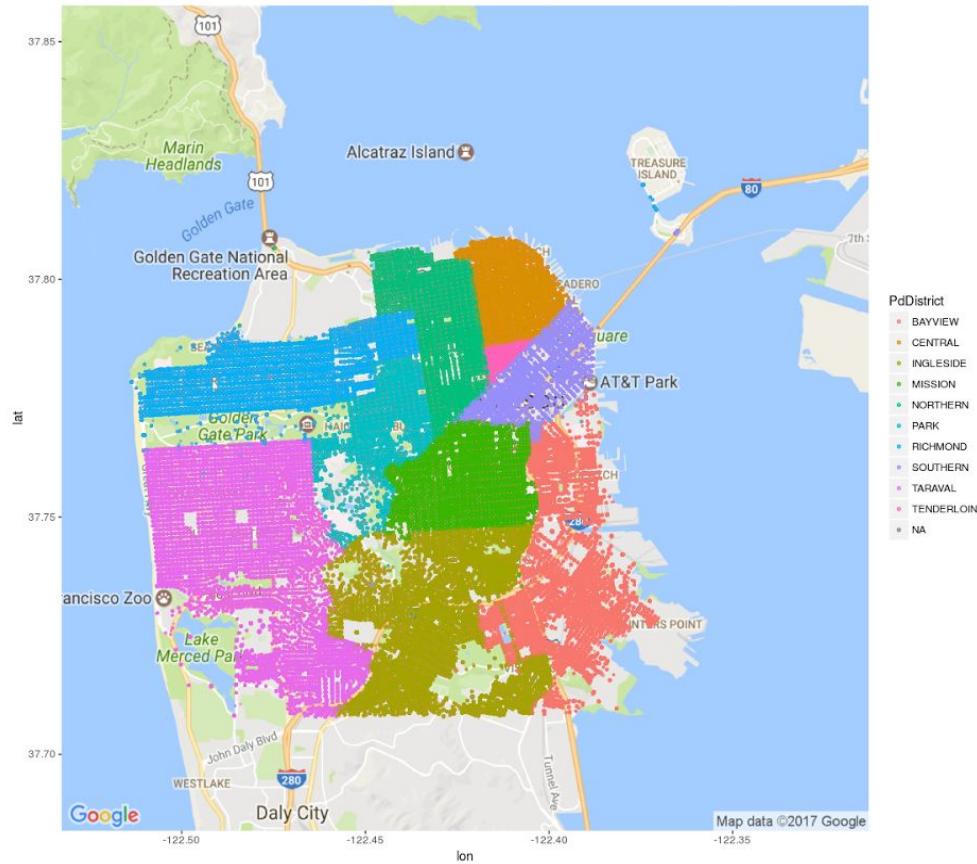
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Police Department Districts



(Plot can be viewed in 3.nb.html) Fig 3.6

Note : From here onwards on each section has plot in the respective nb.html file mentioned in the top of the section in the same order as here.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

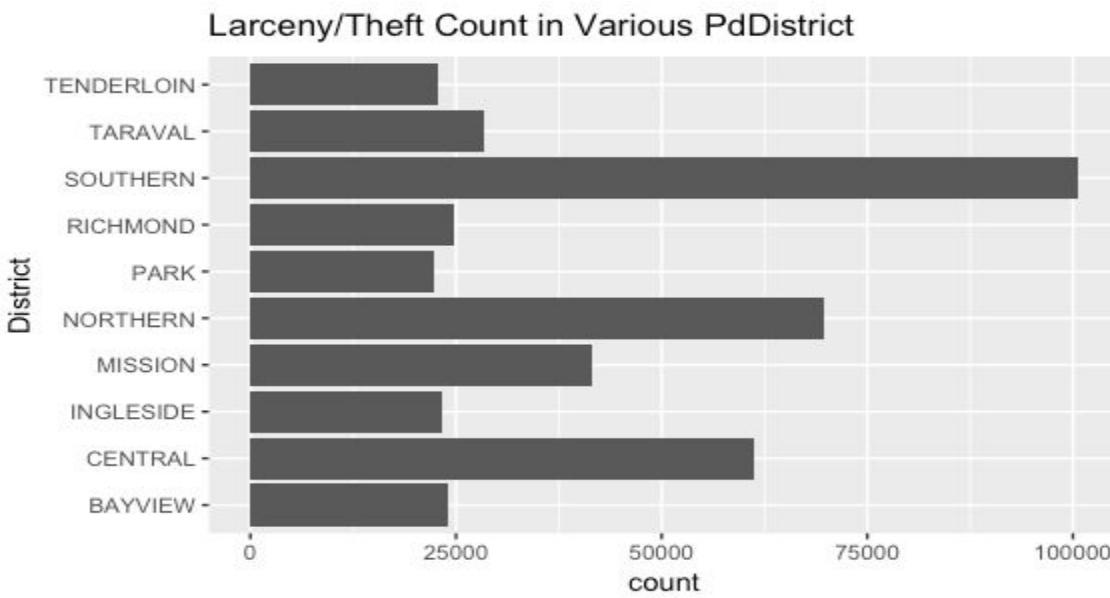
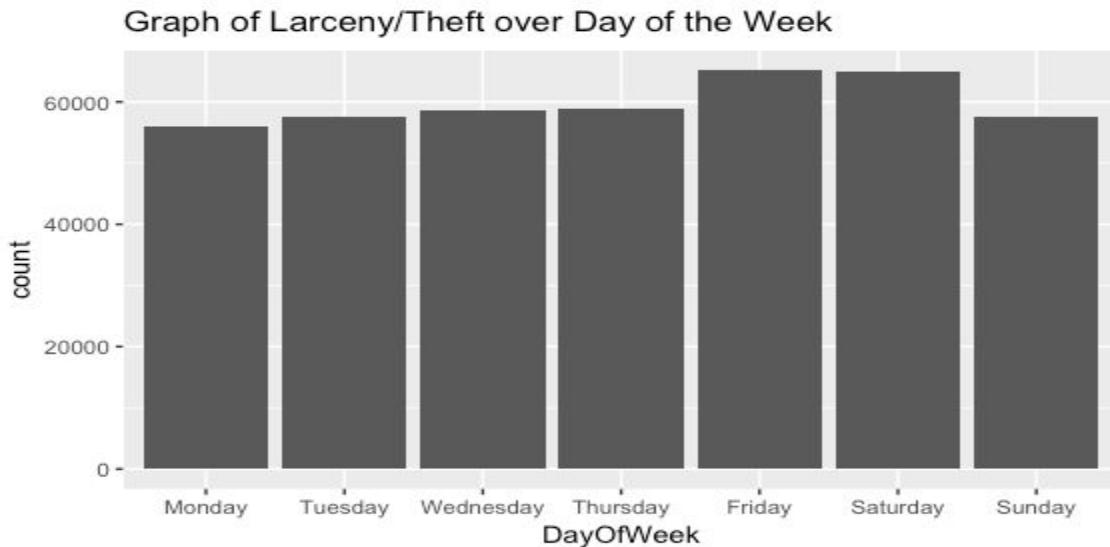
Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Section 1. Larceny/Theft Analysis (Plots in this section can be viewed in 1.nb.html)

As seen in the graphs above “Larceny/Theft” is the most reported crime category.

i) Trend of Larceny/Theft over Day of the Week and Police Department District.



- The above graph show that Larceny/Theft usually occur maximum on Friday and Saturday.
- Also, Southern District is the most affected region of all.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

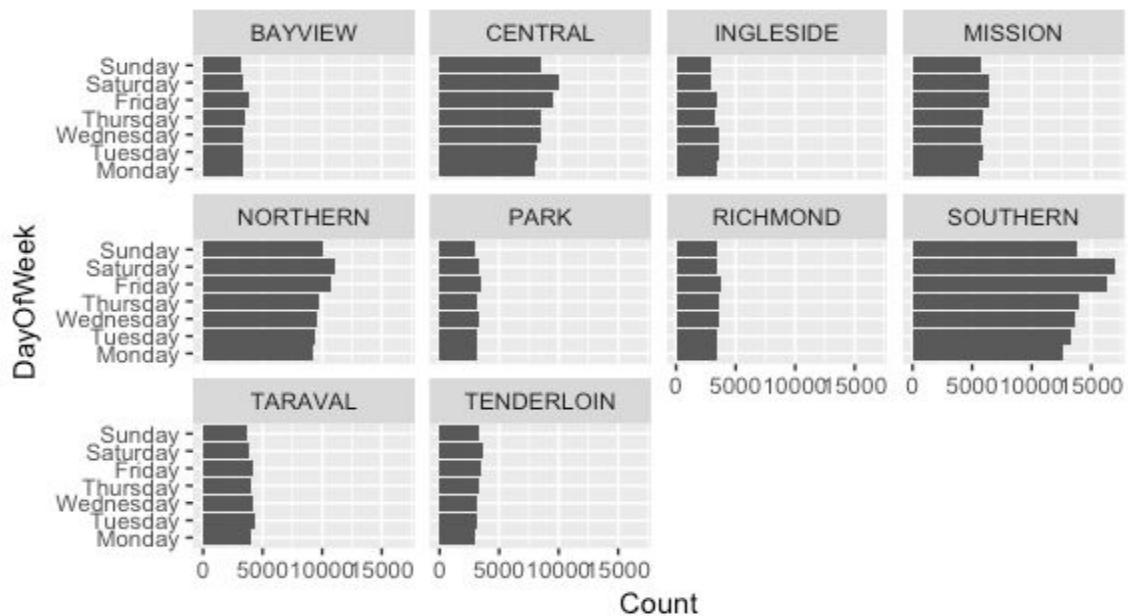
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

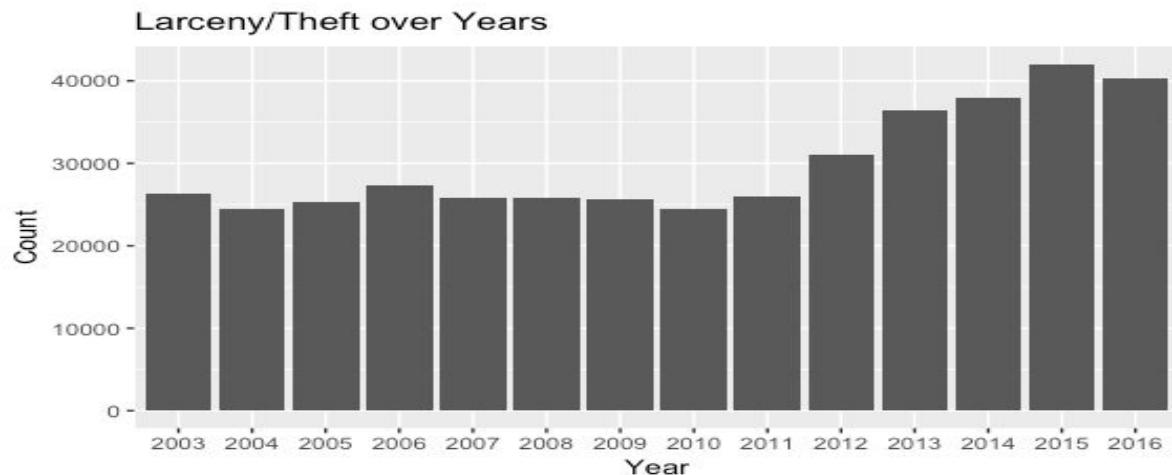
Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Larceny/Theft on Day of the Week faceted over PdDistrict



It can be observed that the trend of Larceny/Theft reported significantly high on Friday and Saturday is not prevalent in all the district but only in some like Southern, Northern, Central. But since, crime reported in this region on Friday and Saturday is so high that it skews the overall count on Friday and Saturday.

ii) Trend of Larceny/Theft over Years and District.



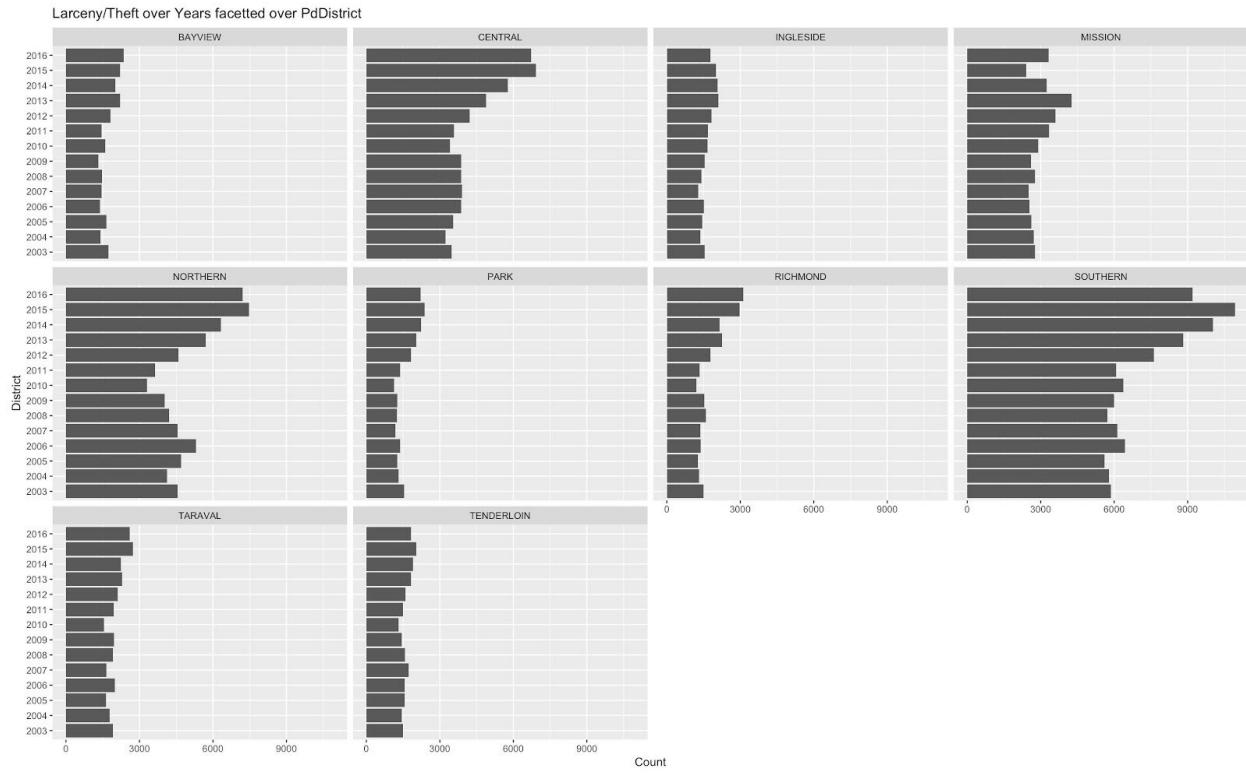
NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>



Except for Richmond And Bayview, it has been observed that there has been a dip in Larceny/Theft in 2016 unlike the increase over 2011-2015. Southern, Central and Northern seemed to be affected the most in the recent times.

iii) Time Series Analysis of Larceny/Theft over recent years.

It can be observed from the graphs above that in 2013, 2014 and 2016 more crime is observed in the first half of the year as compared to second half of the year.

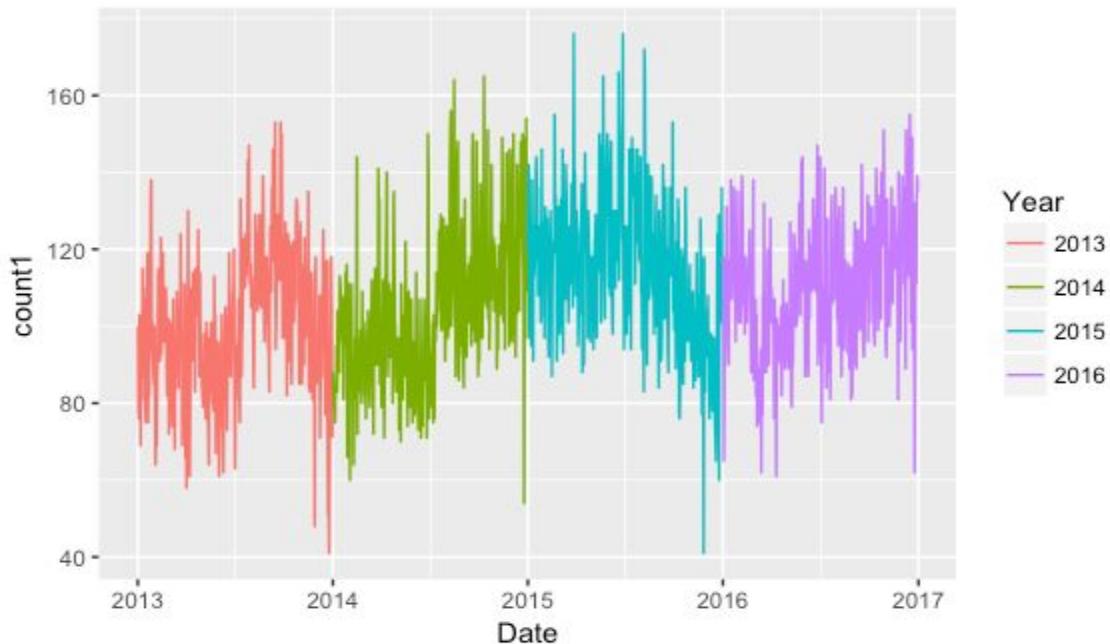
NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

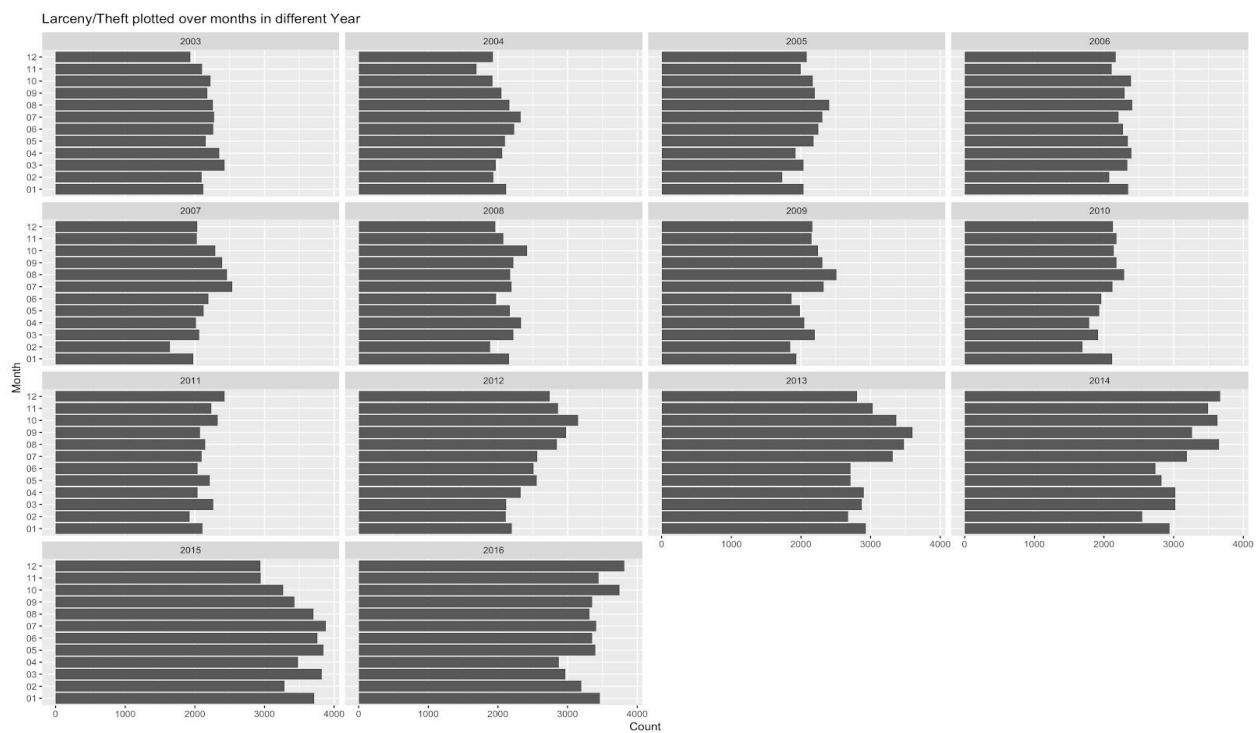
Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>



One thing to observe in all the 4 years is that there is a sudden decrease in crime count of Larceny/Theft on Christmas. We couldn't account for the sudden dip to zero in 2015.

iv) Trend of Larceny/Theft over various months in a given year.



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

It is observed that over the recent year except 2015, December has seen a rise in Larceny/Theft. February is usually low just because of lesser number of days. Other observations seemed a factor of randomization which we couldn't account for.

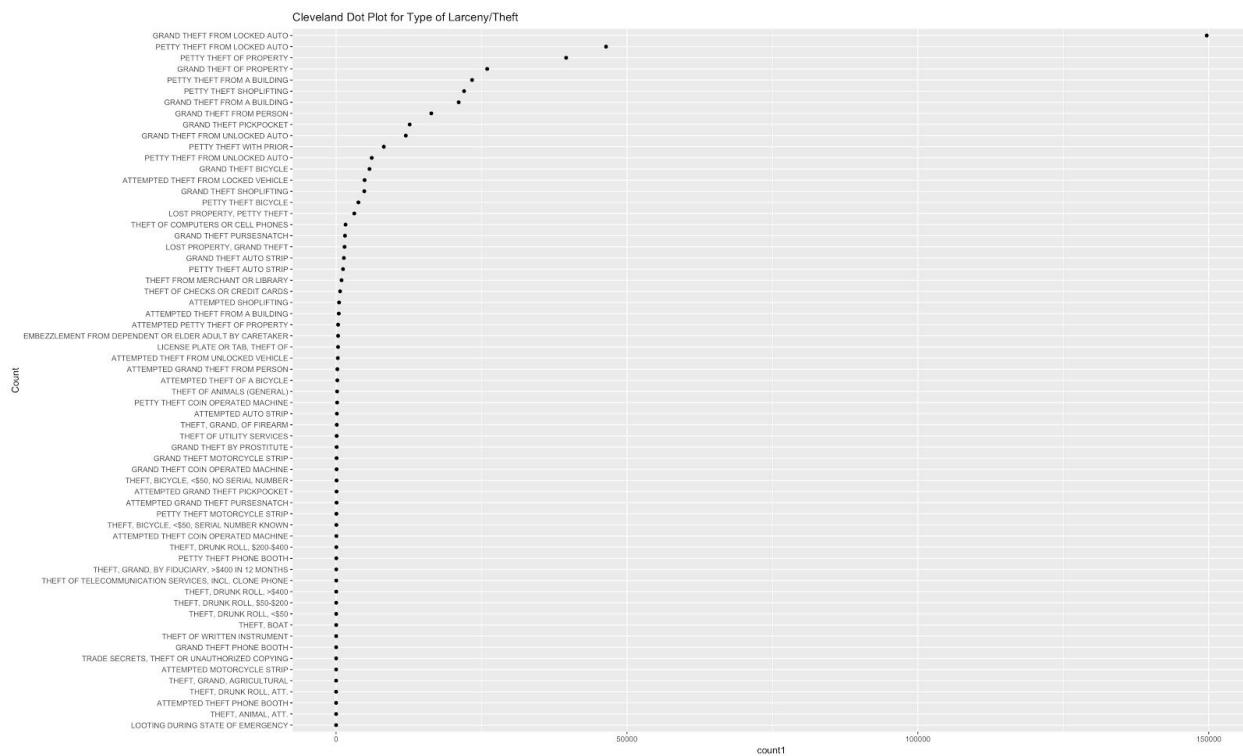
v) Count of type of Larceny/Theft and distribution of top types in various PdDistrict :

It is observed that thefts from Auto have been the major type of Larceny/Theft that have occurred. On further analysis of top two types resulted in bringing up something really interesting. There is a sudden dip in grand/petty thefts from auto in 2006 onwards. This dip can be attributed to the "special car-boosting task force" which was established in late 2007 to curb the menace of auto-mobile theft.

[\[http://www.sfexaminer.com/police-busting-san-franciscos-onslaught-of-car-burglaries/\]](http://www.sfexaminer.com/police-busting-san-franciscos-onslaught-of-car-burglaries/)

The plot below depicts the same. However we couldn't account for rise of these crime back in 2011.

Also, one can see in the third plot that Southern, Northern and Central account for major type of Larceny theft as indicated previously. However, Tenderloin shows that it has larceny reported for Pickpocket and theft from person making it unsafe for people to move around comparatively.



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

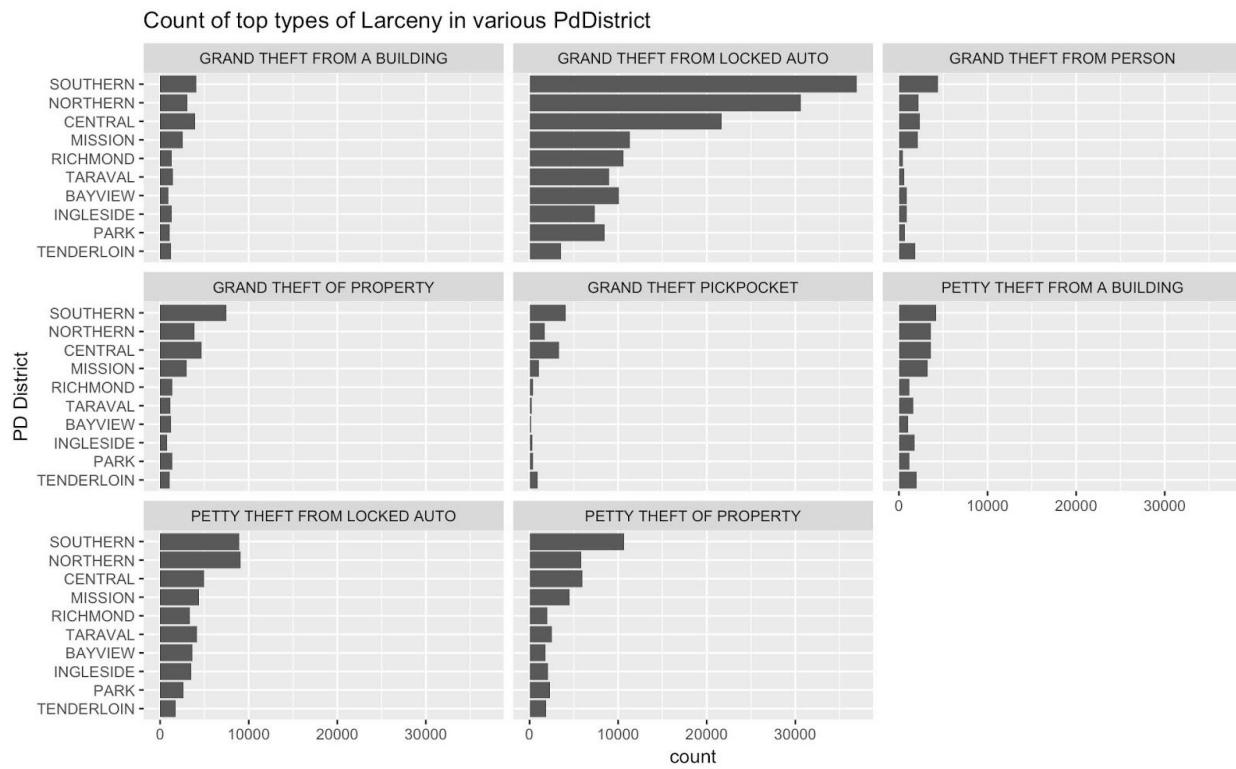
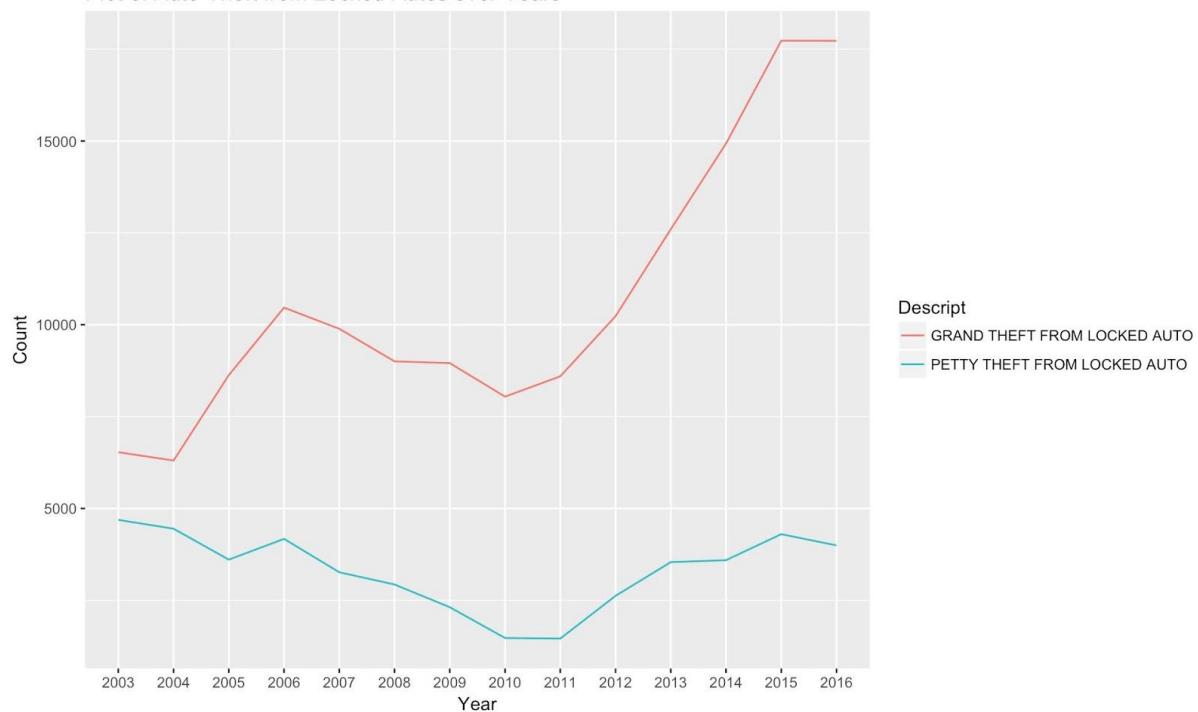
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Plot of Auto-Theft from Locked Autos over Years



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

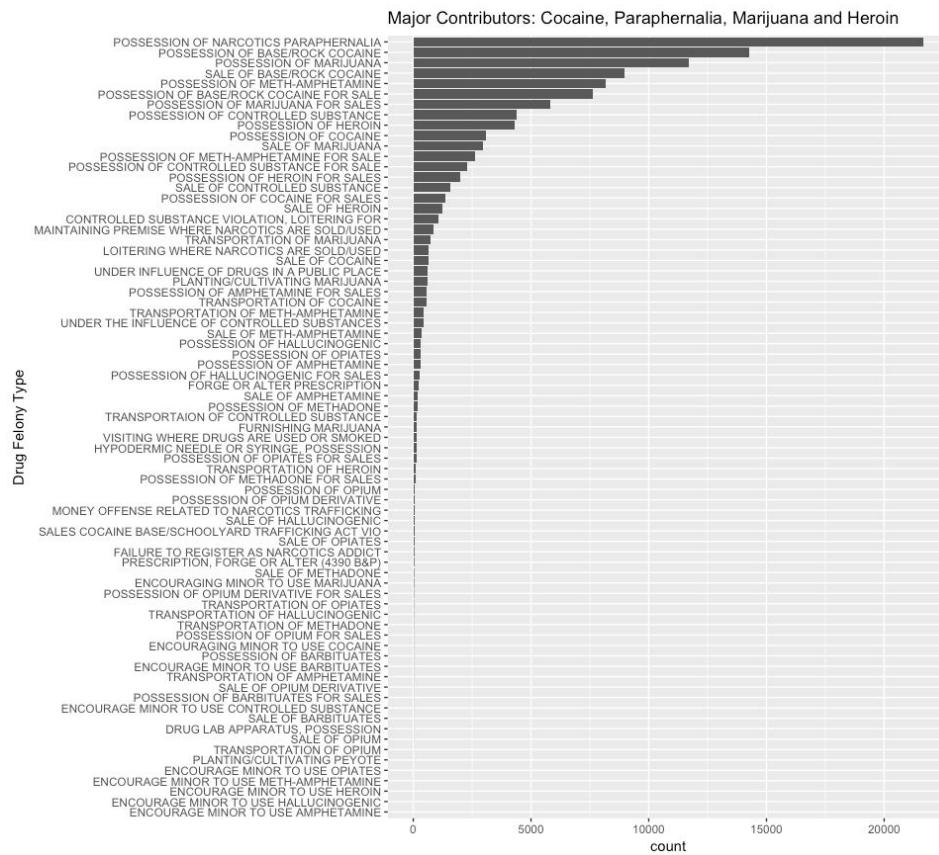
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Section 2. Drugs and Narcotics Analysis (Plots in this section can be viewed in 2.nb.html)



Major crime committed due to drugs and narcotics involve possession and sale of Cocaine, Paraphernalia, Marijuana and Heroin. Further analysis have been done on these major contributors of crime reportings.

Time Analysis:

- Preliminary Analysis:
 - The trend for the days of the week for drug and narcotics crime reportings differ from the general trends as observed in other categories of crime. Most crime categories have maximum number of reportings during the end of the week. But most number of drug related crimes have been reported during middle of the week over the years (2008 and 2007 being different).
 - Before 2009, the number of reports for drug related crimes were more or less randomly distributed. But after 2009, there has been steady decline in the drug and narcotic based crimes.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

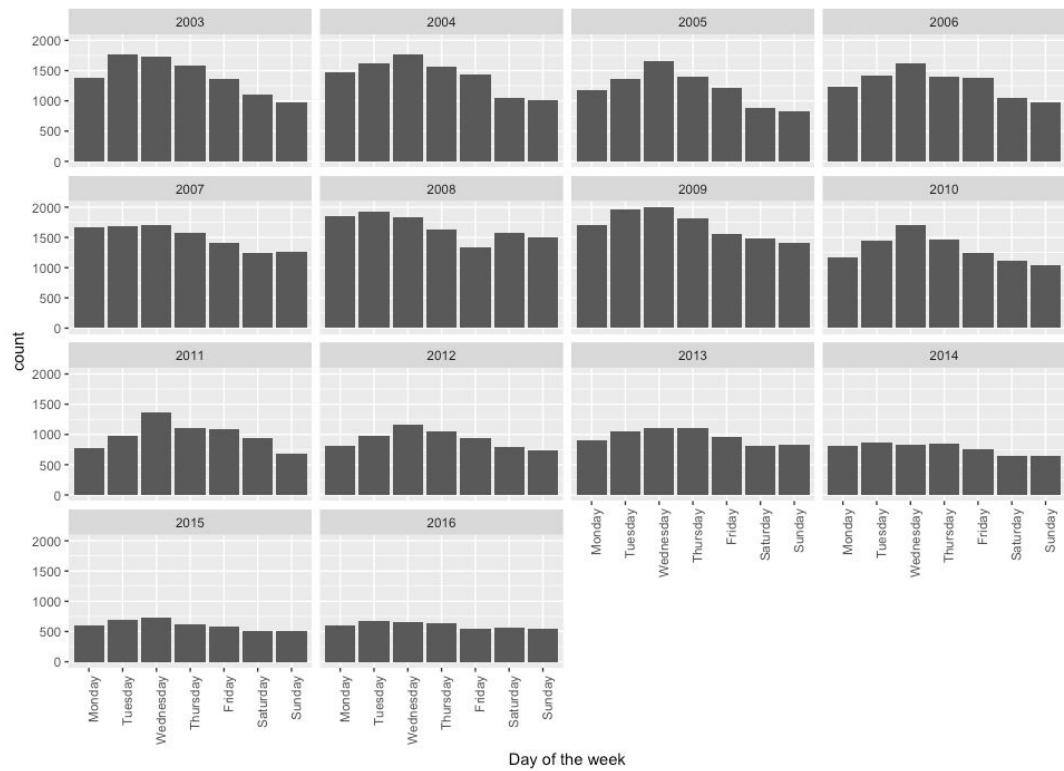
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

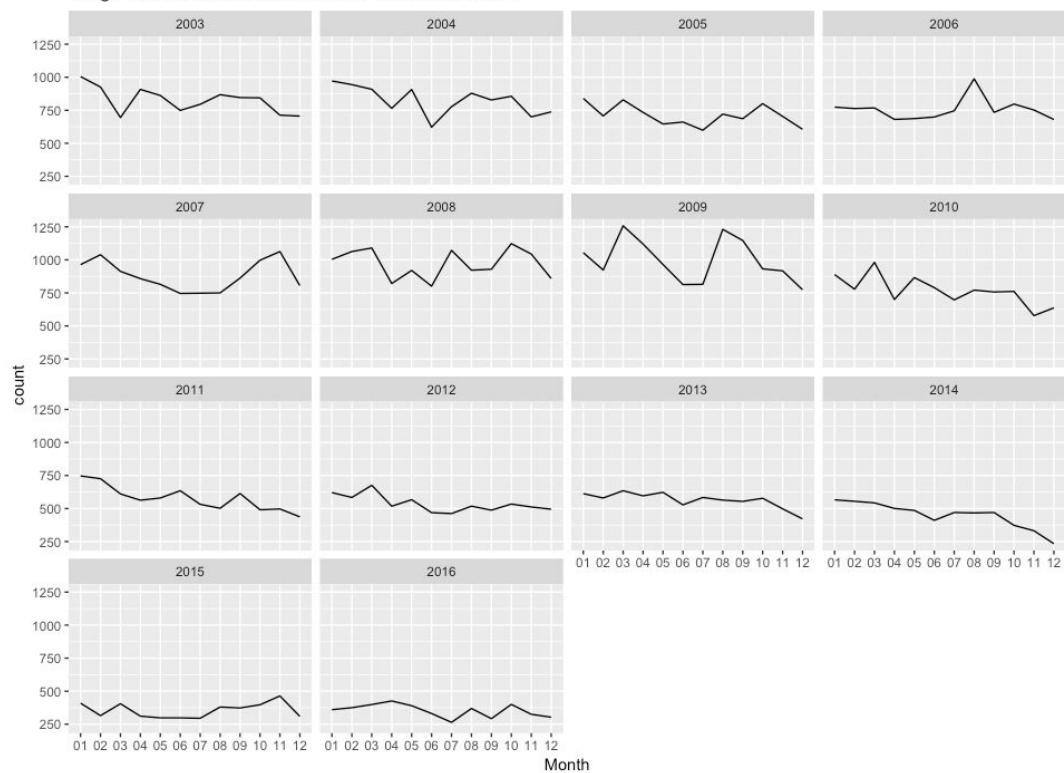
Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Drug/ Narcotic based Crimes Surge on Wednesdays!



Drug / Narcotics based crimes Decrease after 2009



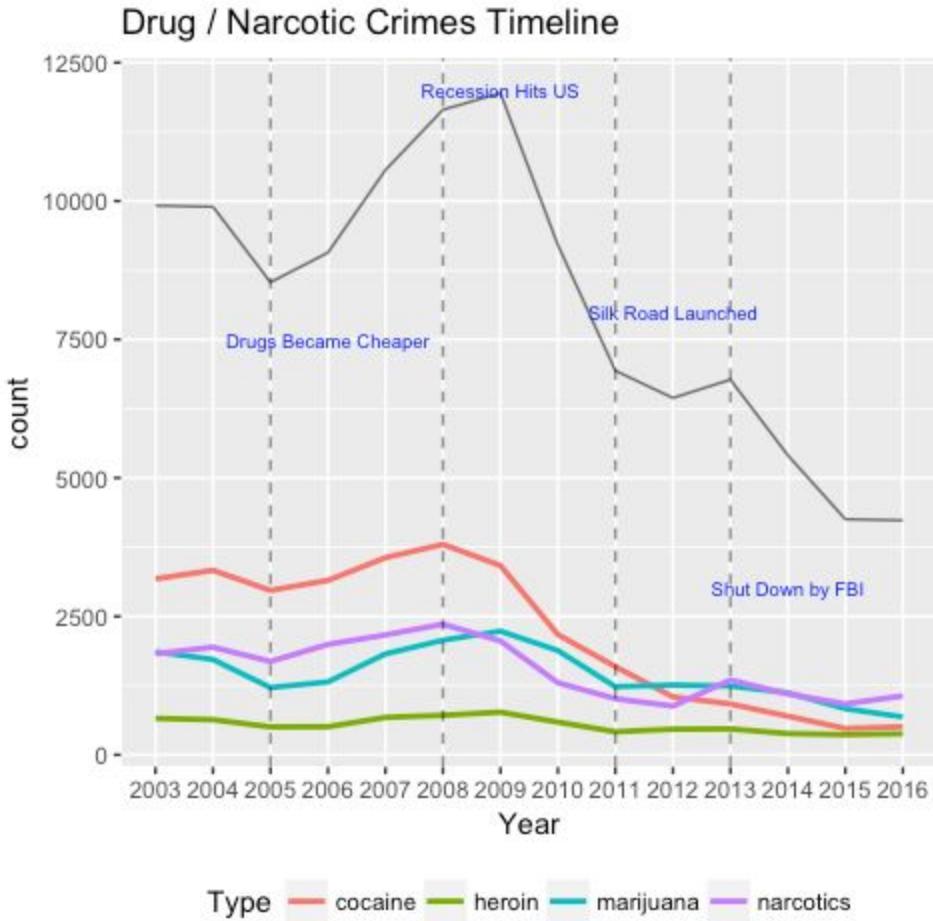
NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>



- Timeline Analysis:
 - As purity increased and drug prices decreased the number of crime reports also started increasing since 2005 . Reference Link : <https://www.drugabuse.gov/publications/drugfacts/nationwide-trends>
 - Many data analysts have reported how the usage of cocaine is positively correlated with the state of the economy. In the wake of 2007 recession, the cutbacks to the police force for drug enforcement and drug legalisation has seen a constant decline in the occurrences of drug related crime reporting especially coke. Since cocaine and base rock have the largest market, such a declining trend has been reflected in the total drug based crime reportings. Reference Link: <https://www.infowars.com/when-san-francisco-stopped-prosecuting-drug-users-violent-crime-went-down/>
 - Launch of Silk Road : the online marketplace for illicit drug users was launched in 2011 where Narcotics Paraphernalia, opiates and Marijuana had the dominant market share. Increased sales of these drugs resulted in the a slight variation in the general downward trend for all drugs combined. The site was shut down by FBI in 2013. (But other illegal deepweb sites still remain operational). Reference Link: [https://www.infowars.com/when-san-francisco-stopped-prosecuting-drug-users-violent-crime-went-down/](#)

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

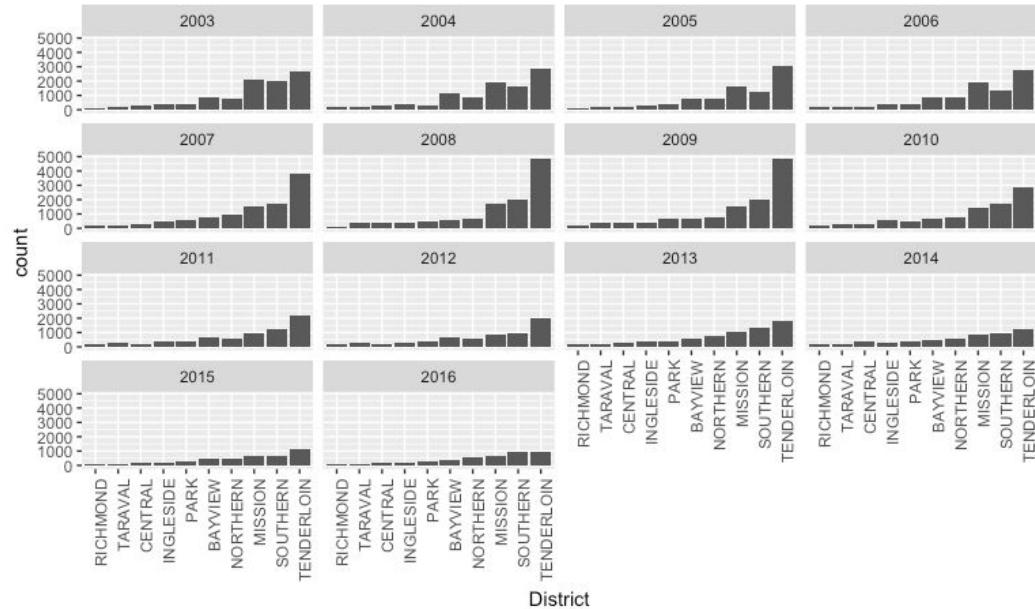
Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

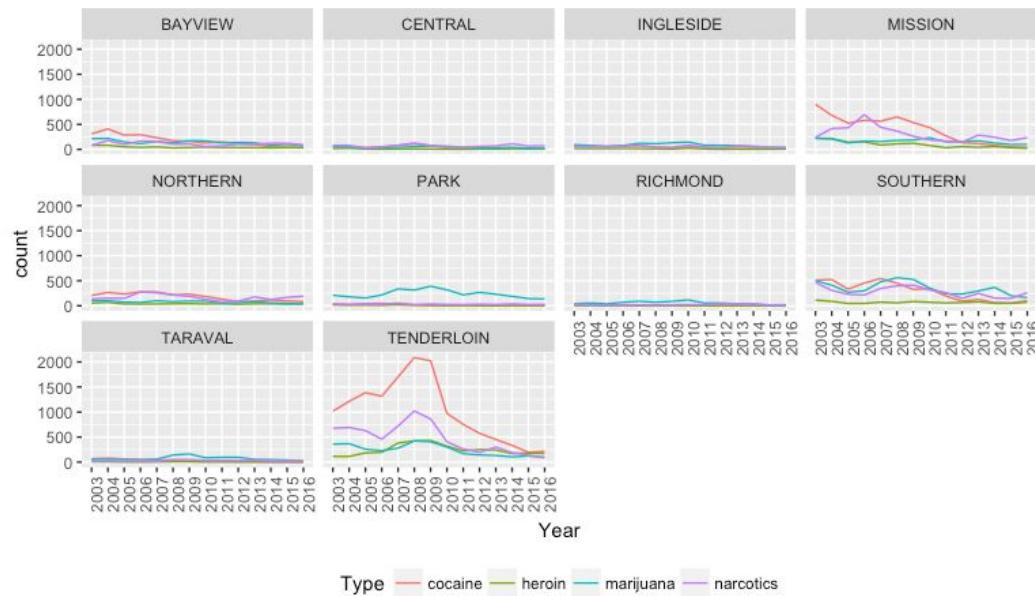
<http://www.nytimes.com/2013/10/03/nyregion/operator-of-online-market-for-illegal-drugs-is-charged-fbi-says.html>

Area Analysis:

Drug/ Narcotic based Crimes Distribution over Districts



Different drugs over Years and Districts



- Preliminary Analysis:

- The main districts that contribute to drug based crime reports are: TENDERLOIN, MISSION, SOUTHERN, NORTHERN and BAYVIEW. The other districts have a very small count of crime reports.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

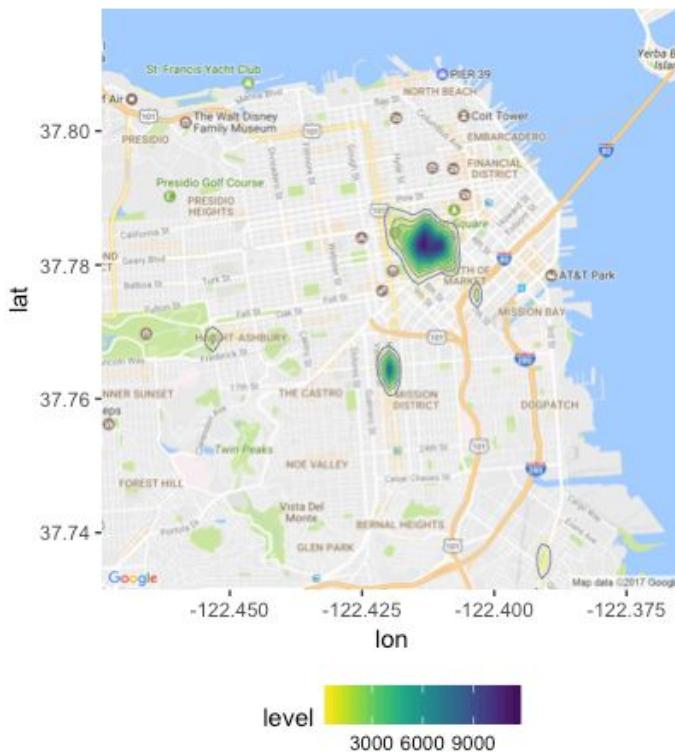
Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

- Tenderloin is the main district where the sale and purchase of cocaine takes place. Narcotics also have a huge market there.
- Density Analysis:
 - As is evident from the preliminary analysis, the major drug related crimes are centered along the small district of Tenderloin.
 - Second graph shows the closeup of the major areas of drug based crime reports. Market street is the main highway along which most of these reports originate. Considerable amount of reports originate from Mission Street as well.
 - The major contributors of the drug based crimes (cocaine, heroin, marijuana and narcotics) have similar density plots implying that no area has a specific preference when it comes to illicit drug usage .

Drug Related Crime Density



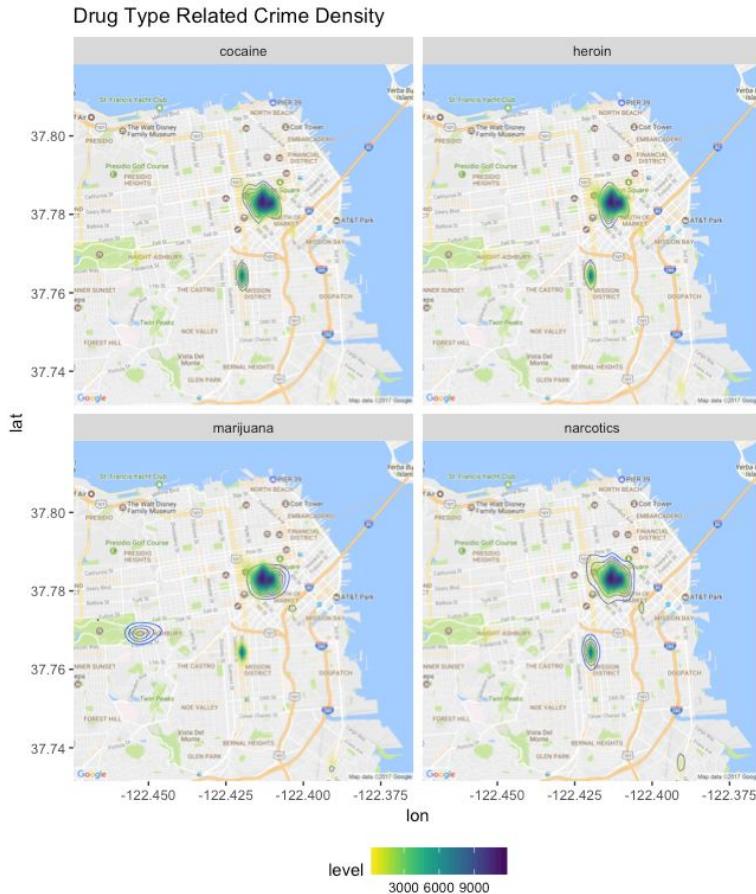
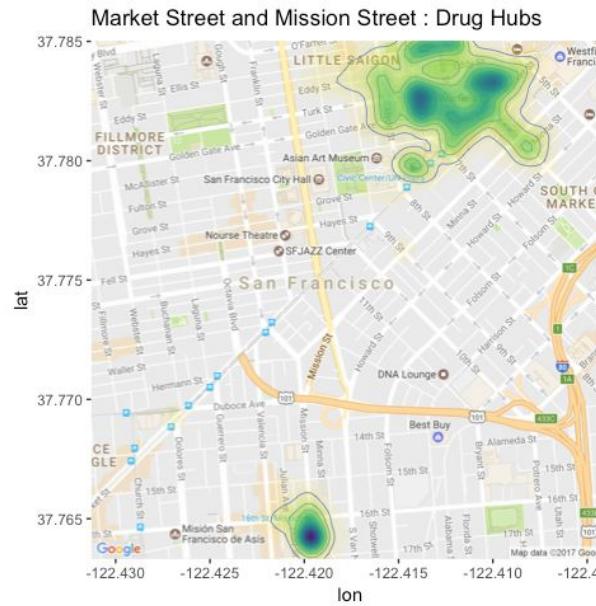
NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Section 3. Vehicle Theft Analysis (Plots in this section can be viewed in 3.nb.html)

Time Series Analysis

Yearly Trend

- It is observed from the crime reports for vehicle theft over years that there is a sudden decrease in the number of crime reports after year 2005.
- Vehicle theft decreases from year 2006 to year 2010 and then keeps in increasing (40% increase) upto year 2015.
- After 2015 there has been a decrease in the crime reports for vehicle theft.

Hypothesis for decrease in vehicle theft after 2005:

- Many criminologists attribute the decline to technology especially cars. Smart keys and keyless ignitions did complicate the crime, and police gadgets such as automated license plate readers made it easier to find missing vehicles. All these technologies were newly introduced and stealing vehicles was not as easier as before. (<http://www.iii.org/issue-update/auto-theft>)
- Many times tampering with the car is reported in the vehicle theft. In 2006 to combat with the car burglaries/ smashing of car windows, mayor in that area assigned a special car-boosting task force. (<http://www.sfexaminer.com/police-busting-san-franciscos-onslaught-of-car-burglaries>)

Hypothesis for Increase in Vehicle thefts after 2010:

There are many arguments and facts that can support the fact that crime reports for vehicle theft have increased from 2011 to 2015.

- Due to continuous boom in technological sector and owing to the fact that bay area is the hub of technological advancement there grew the gap between the poor and the rich. Which then triggered the increase in crimes specially thefts (due to increased poverty). (<http://www.latimes.com/local/crime/la-me-aa2-snapshot-sf-crime-20141120-story.html>)

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

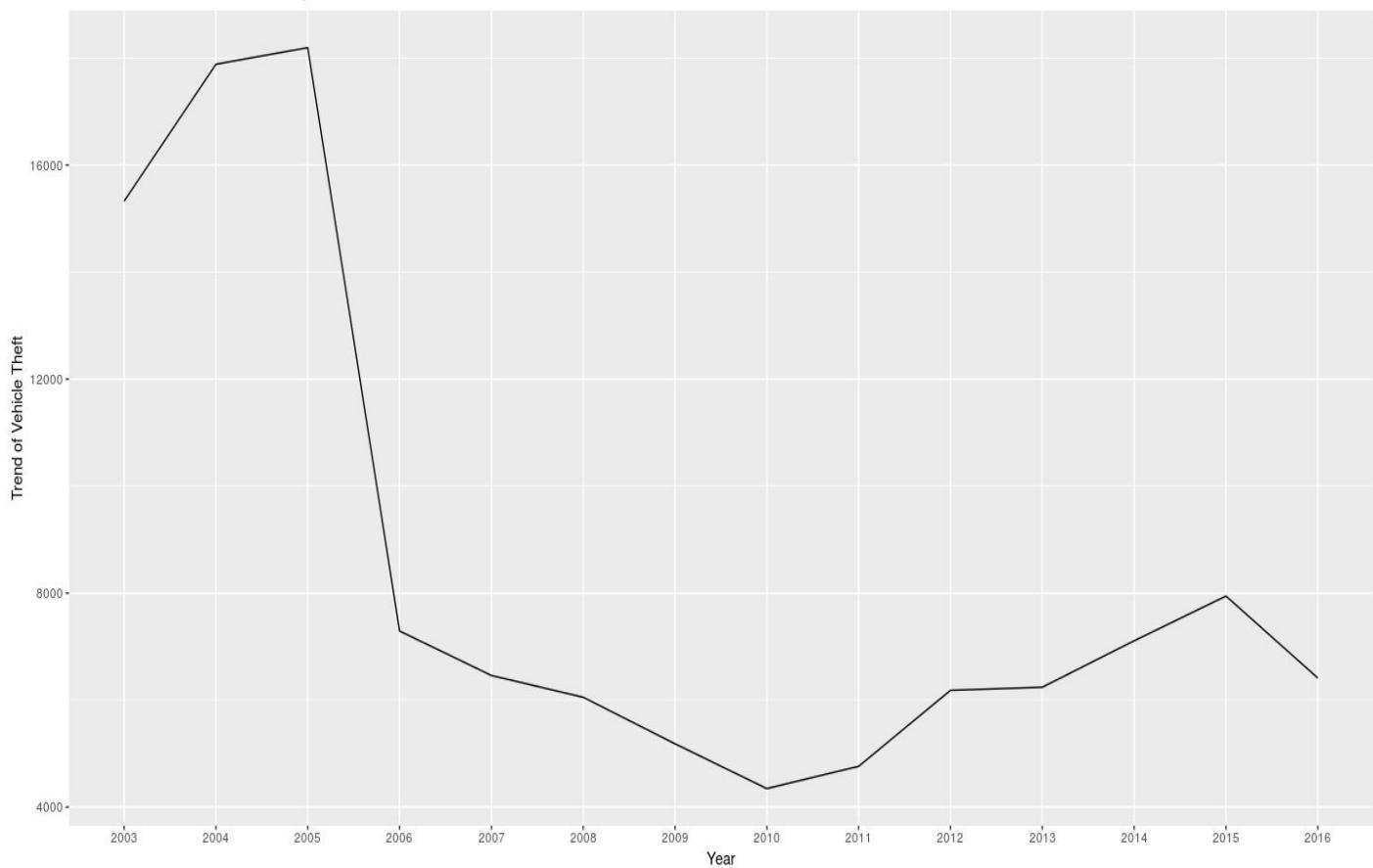
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Count of vehicle theft over the years



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

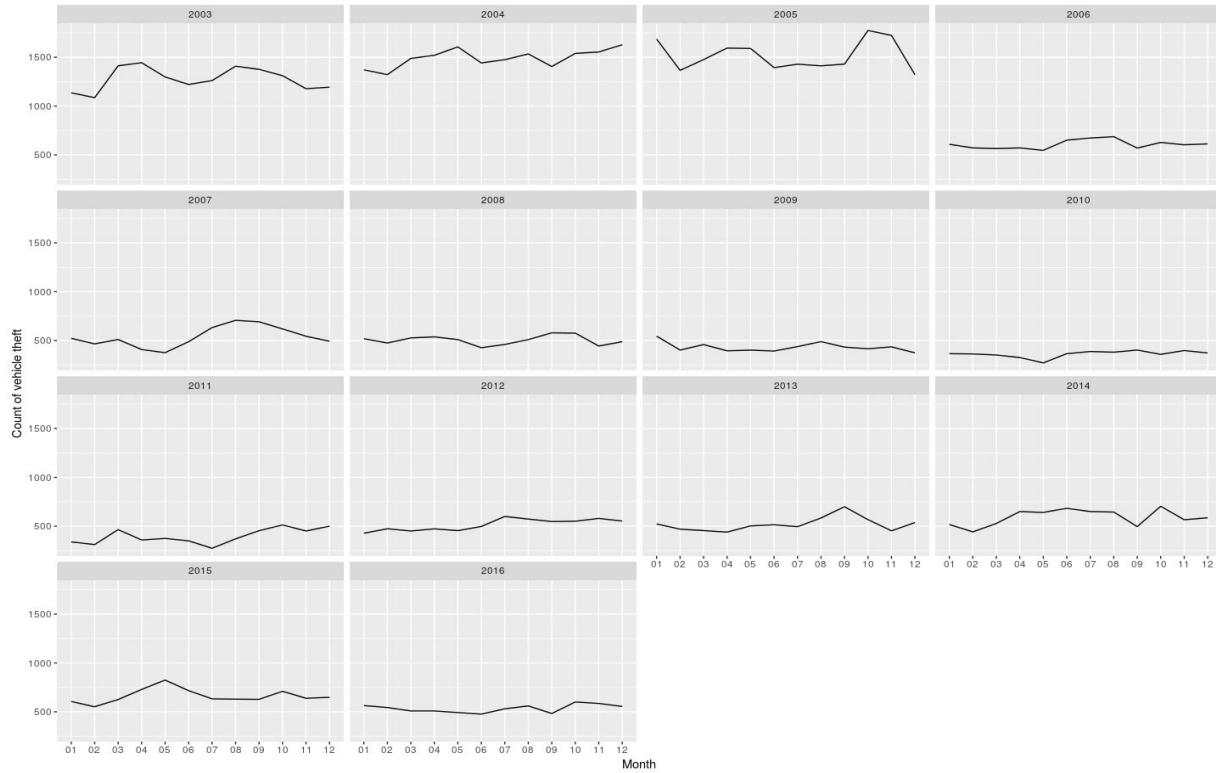
Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

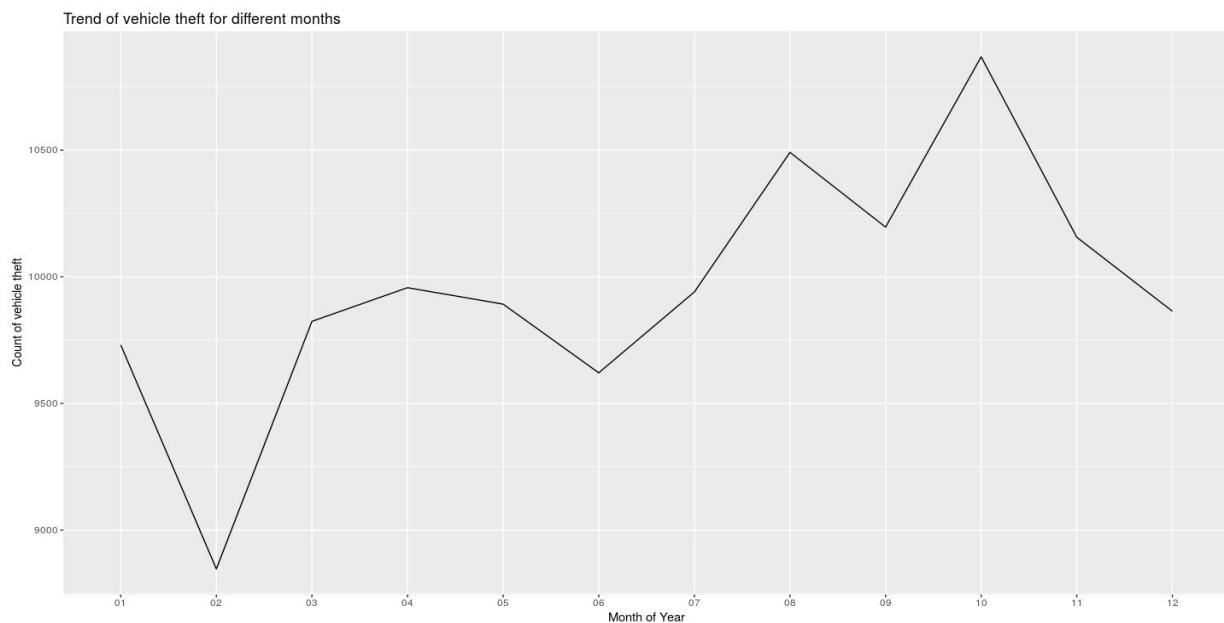
Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Monthly Trend

There is no particular trend being observed over the months for different years.



There is a dip in the crime reports for the vehicle theft for month of February and peak for the month of October. But since this is cumulative over all the years and also some months have lesser days than the others, no inference can be made based on this plot.



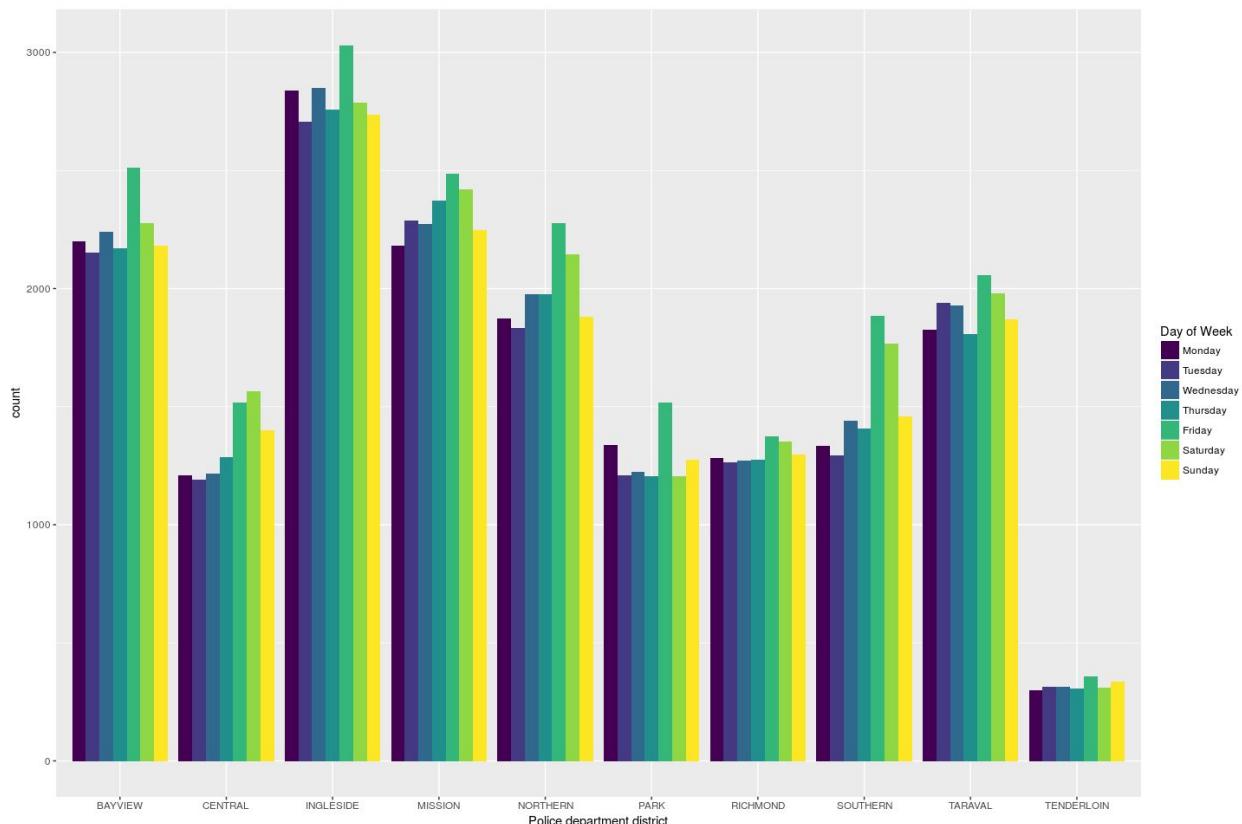
Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Weekly Trend

There is relatively less difference but vehicle theft (and also every other crime) is more on weekends in general. But when we look into police department district wise, some districts like Ingleside have higher vehicle thefts on Monday as compared to Saturday and Sunday. Though Friday is the day which observes highest vehicle theft for all districts.



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

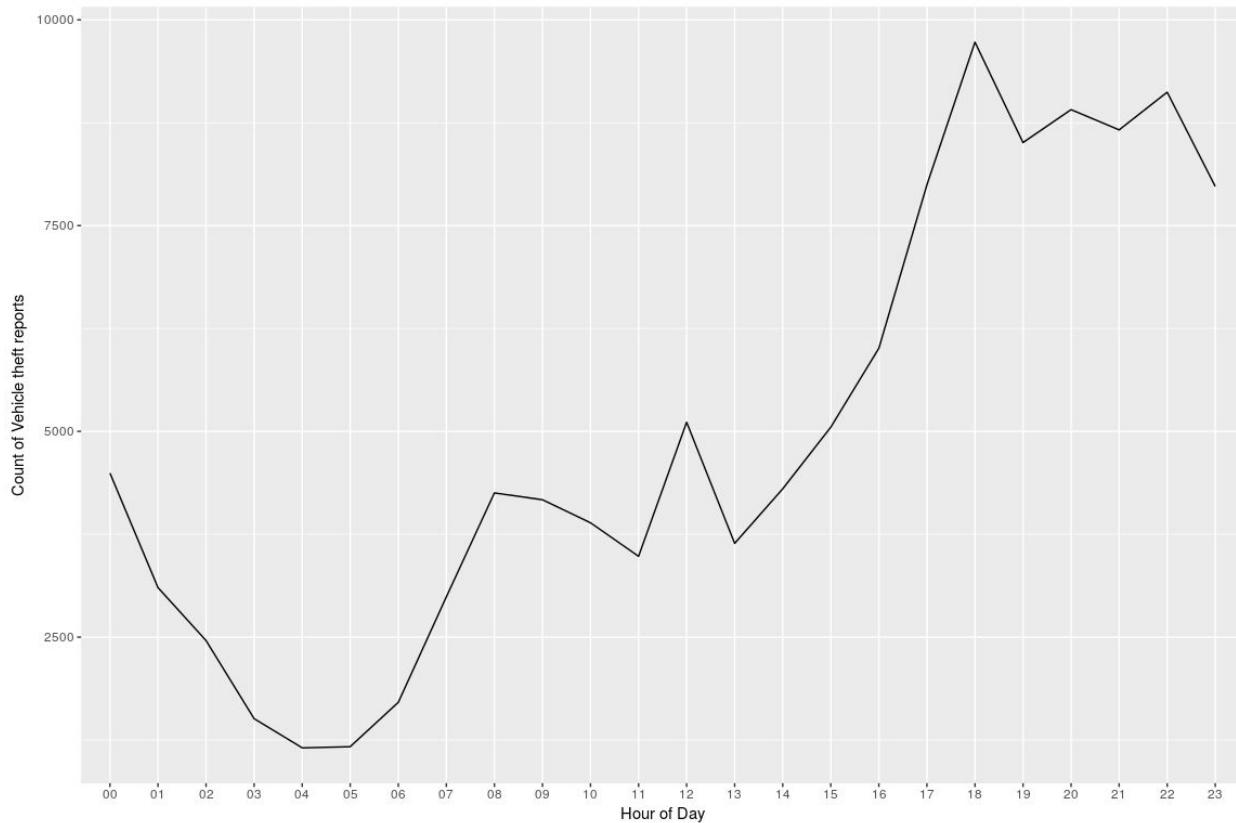
Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Hourly Trend

As observed for any other crime, vehicle theft reports increase after evening to around 3am in the morning and lower for morning hours (lowest being 4-5am)



Categorical Analysis

Under the vehicle theft there are different types of cases which are considered based on the type of vehicle stolen, tampering with vehicle, vehicles which are stolen but then recovered, vehicles stolen outside SF but recovered in SF.

- Below is the bar plot showing the count of different types of cases. It can be observed that most of the vehicle thefts reported are for automobile being stolen and then is truck being stolen. The trend in overall vehicle theft is due to The possible reason for this could be that cars/ automobiles are more in number as compared to any other vehicle.

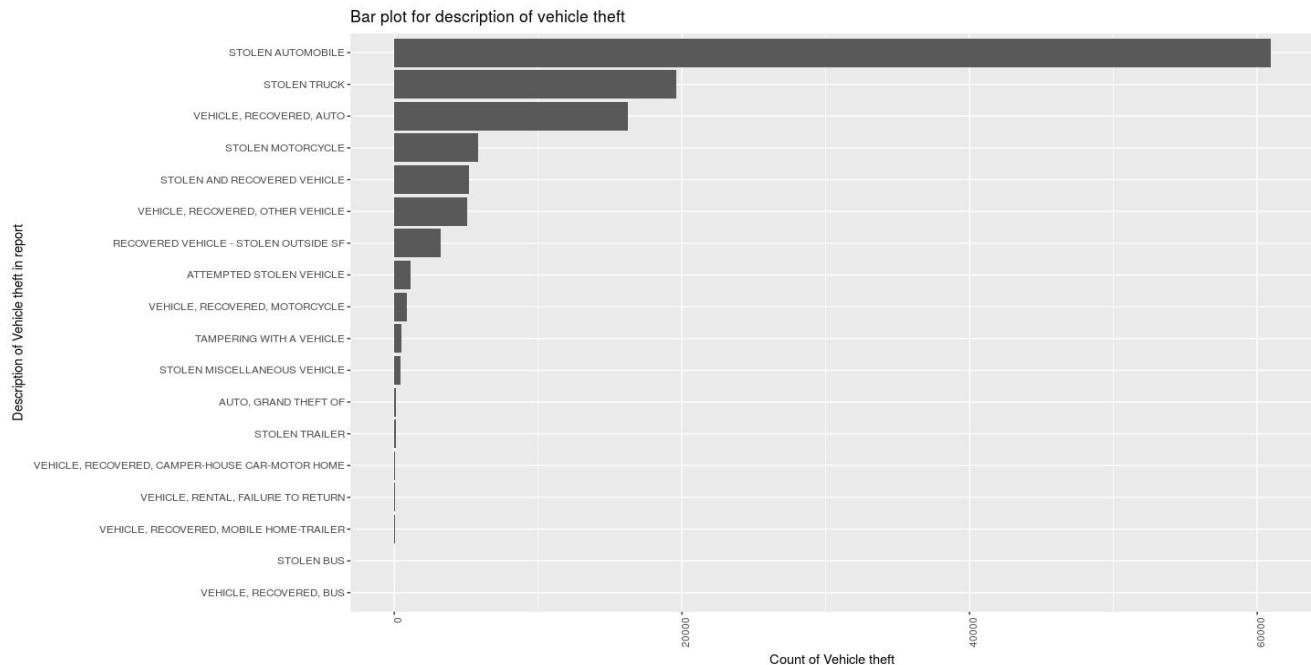
NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

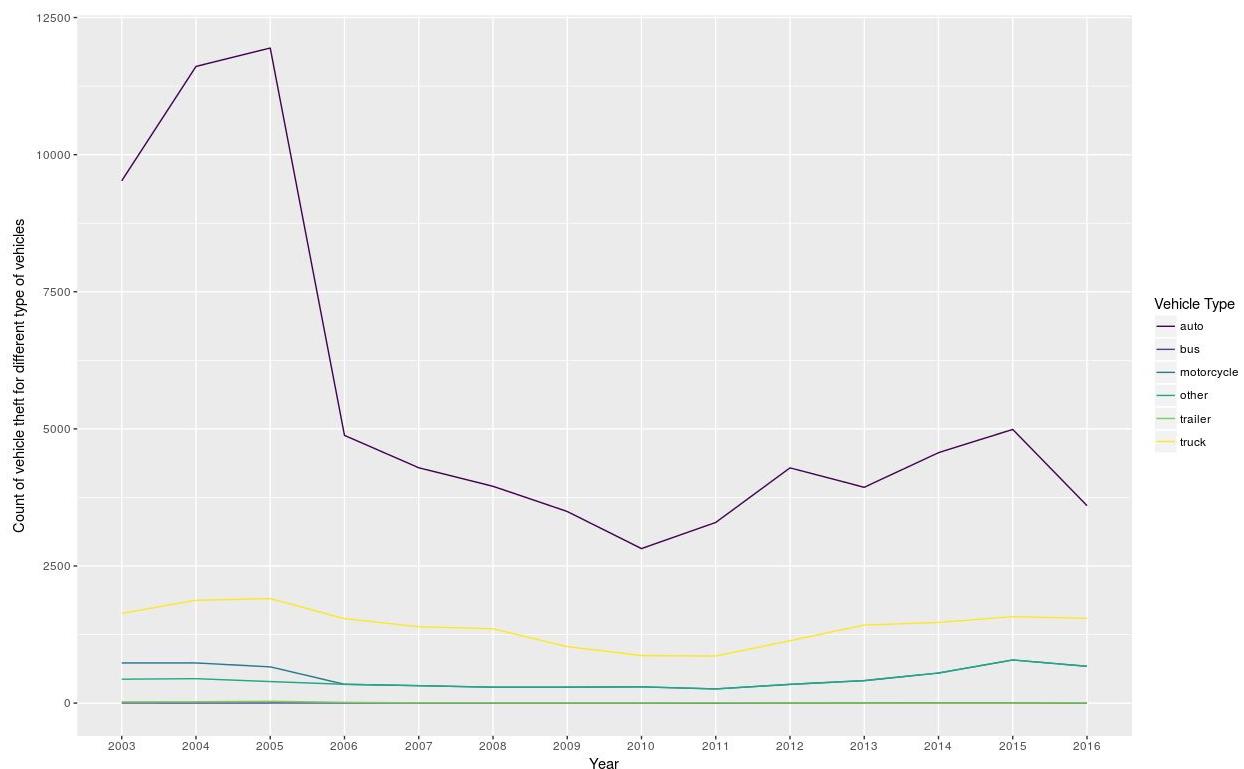
Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>



Yearly trend for vehicle theft for different types of vehicles over the years.



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

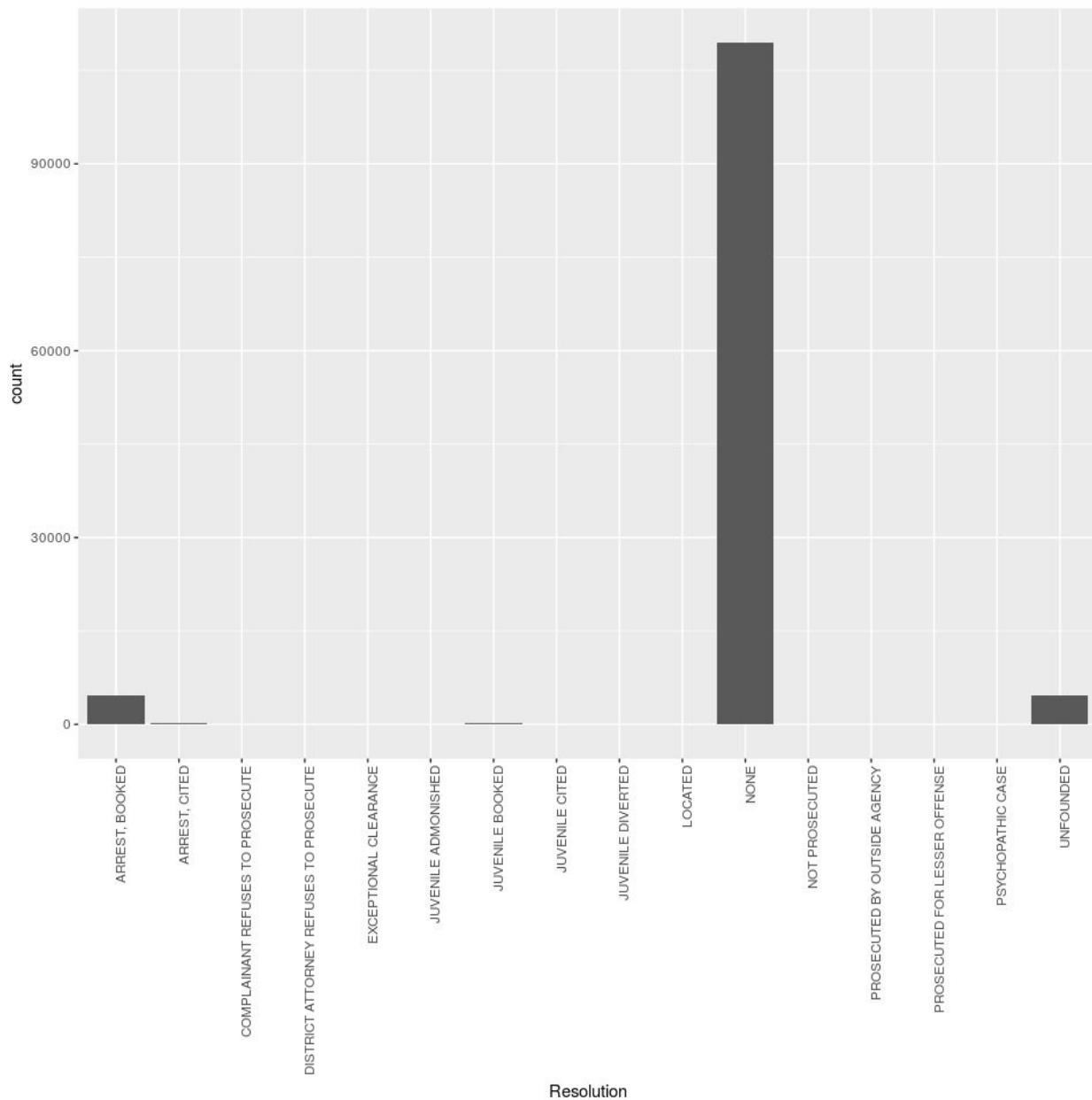
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

As mentioned in the beginning of the analysis most of the crimes resolution are not mentioned in this data, as this is the first information report. The plot below supports this argument



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

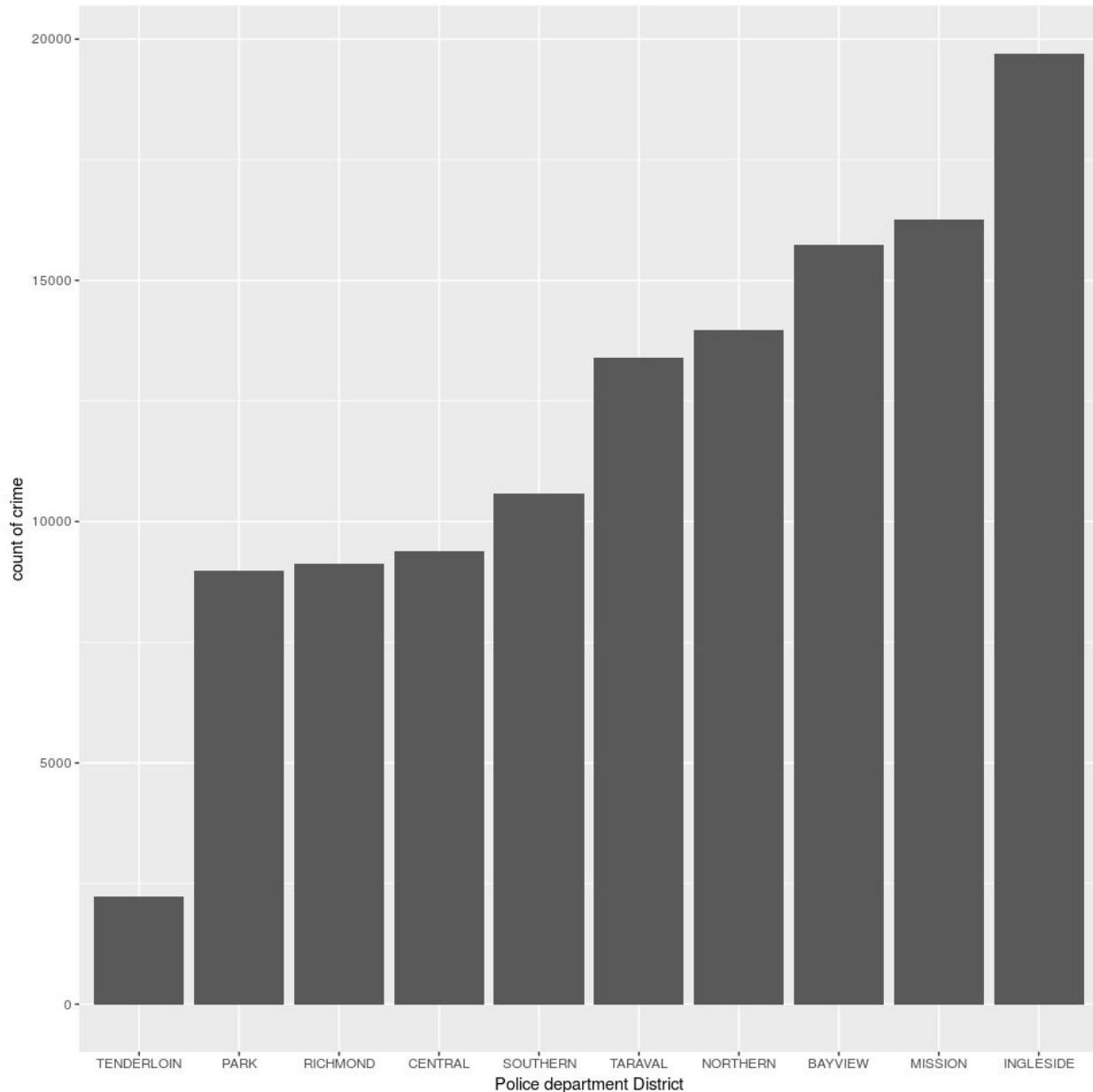
Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Geographical Analysis

The below bar plot shows that most vehicle thefts are reported in the Police department District of Ingleside.



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

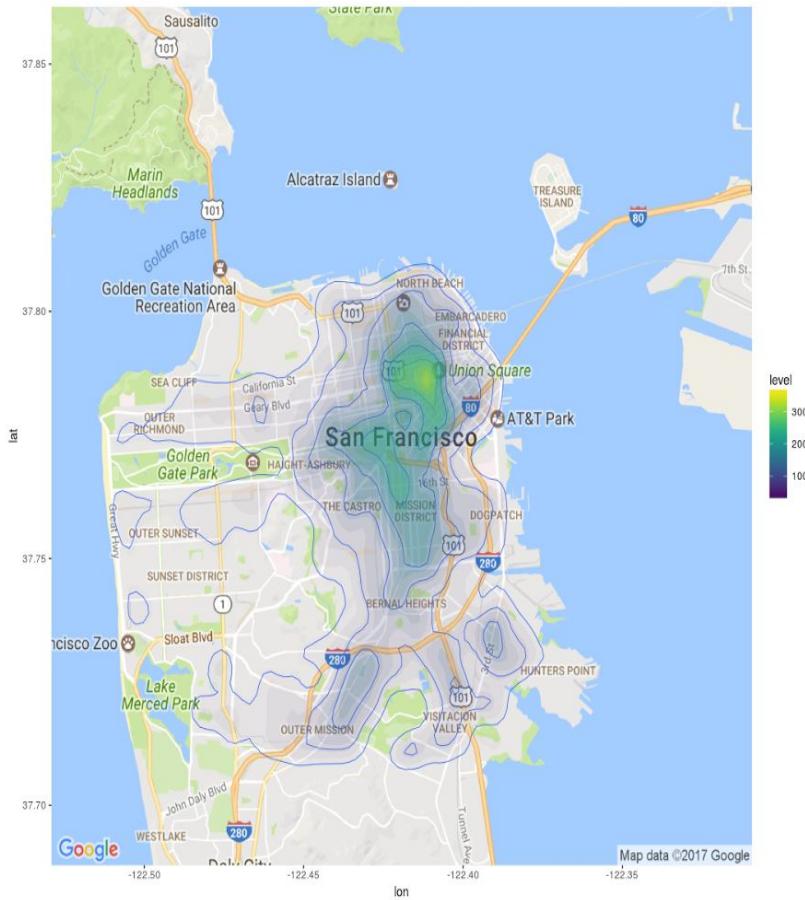
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

The count of reports is highest for Ingleside district but the density of vehicle theft reports is maximum in the districts around market street (Southern, Mission, Tenderloin)



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

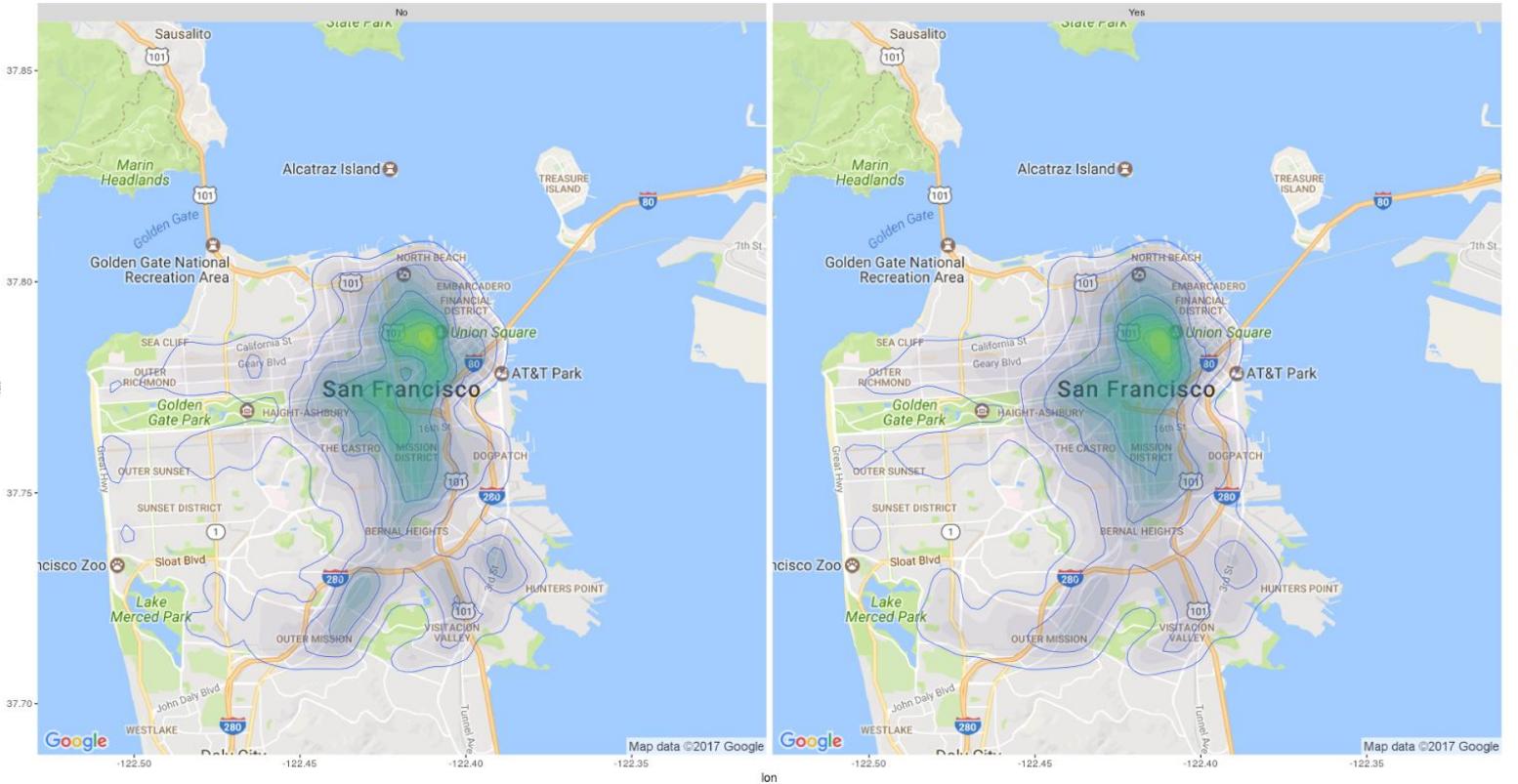
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

The map on the left is for the reports of stolen vehicle which were not recovered at the time of report and right is the map for vehicle stolen and recovered. The area where vehicles were stolen and where the stolen vehicles are recovered is about the same.



The mapping above shows the density of vehicle thefts which were stolen and the ones which were stolen but recovered later.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

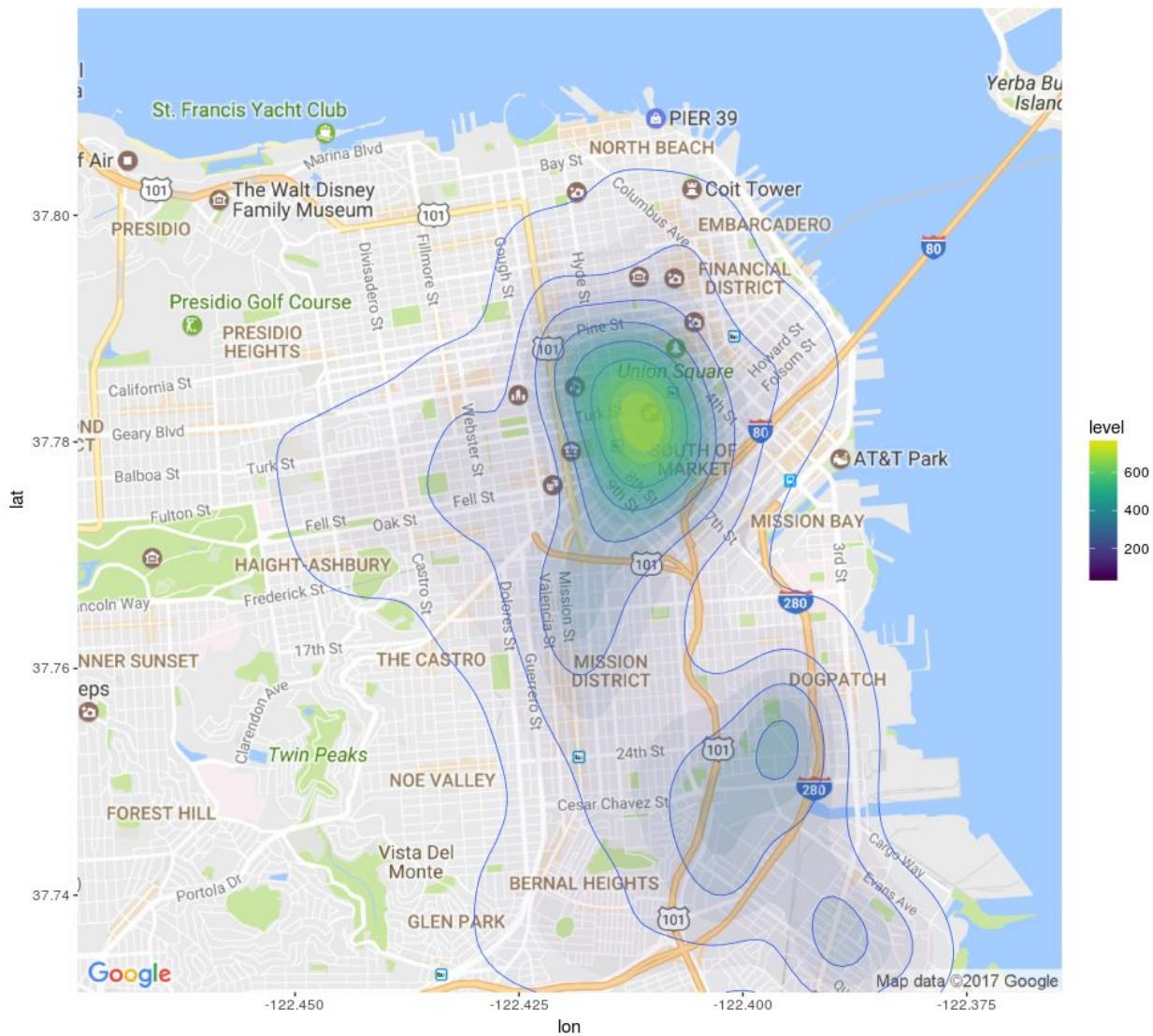
Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

The below map shows the area where the vehicles stolen outside SF are densely recovered.

Recovery of Vehicle stolen outside SF



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

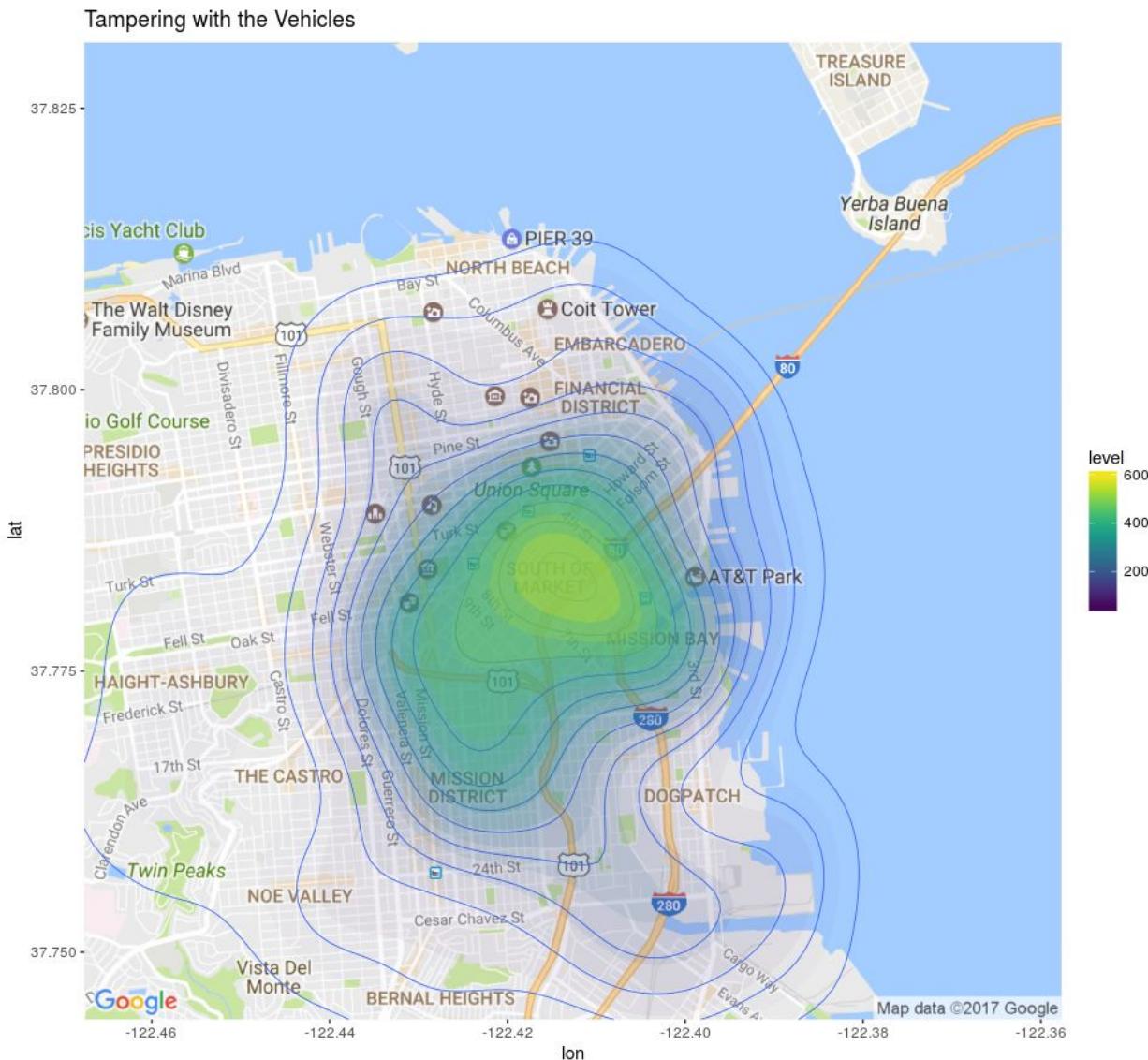
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

The map below shows density and the location where the reports of tampering with vehicles are recorded.



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

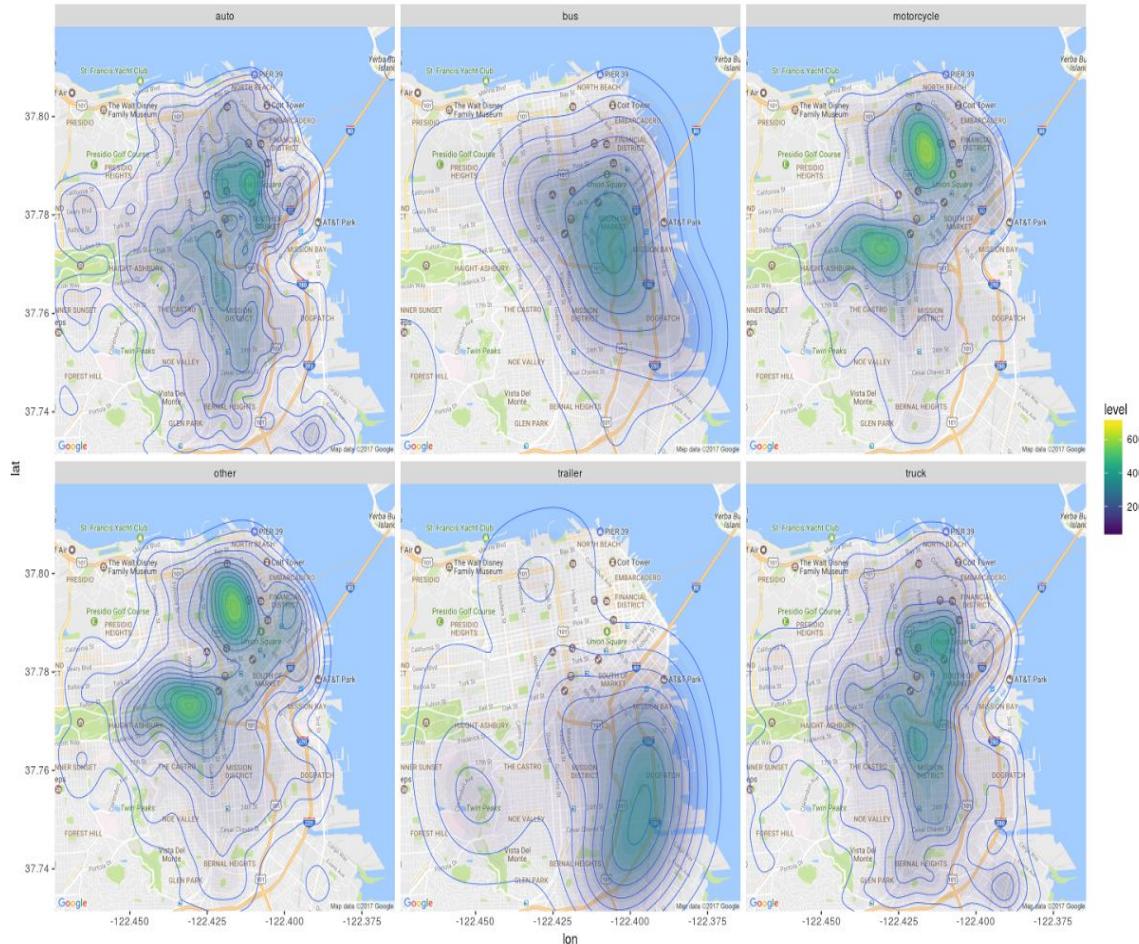
Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

Given below is the density plot of reports for vehicle thefts for different type of vehicles. Key observations are as below:

- Auto: The density of the reports for vehicle theft is most scattered. The maximum number of vehicle thefts are auto thefts and it is maximum for the Ingleside district but the density is high near Market street as well.
- Truck: This is the second highest reported type of vehicle theft and has the similar region as that of the auto theft.
- Trailer: The area of trailer theft is very different from the other vehicles.



Market street area:

- From all the above geographical analysis it is clear that the Market street area has a higher density for vehicle theft.
- One of the probable reason being that this area is most densely populated. So this area has a lot of garages which can attract a lot of vehicles thieves.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

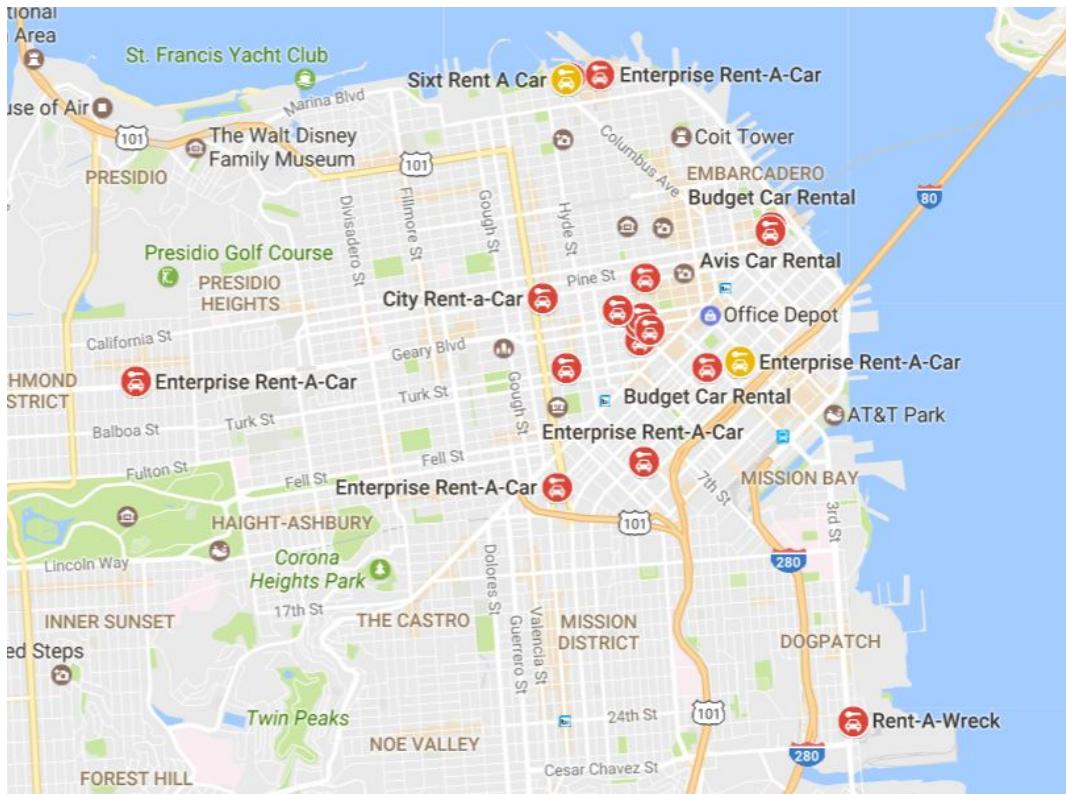
Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

- This area also has a larger number of car rentals and often even a slight delay in return of car rental is reported as a theft. This could also add to the more reports for vehicle theft (below is the image for google map result for car rentals in SF)



NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>

CONCLUSION

Limitations and Shortcomings:

We could have used external metadata like population data and economic data and data for the entire USA to hypothesise claims based on the graphical analysis. But the most recent census data for these variables is available for 2010. Using this 2010 data would have given incorrect inferences for the years after that. Hence we could not derive more interesting results in the context of a bigger picture and limited our analysis to the data available.

Future Directions:

As mentioned, the analysis could be easily extended to other dimensions by using economic and population data. We could also do a comparative analysis with the crime reports in the rest of the America, provided such kind of data is publically available.

We also wanted to make an interactive app like Trulia, but which gave more insightful findings rather than just aggregates and trends. But such a vision is more ambitious and more research time.

Lessons Learned:

- Line charts are very useful for time analysis.
- Heat Maps are very useful for geographical analysis
- Faceted bar charts give a good representation for aggregates like counts for categorical variables.

NOTE : Links may not work in Adobe Acrobat Reader. Please use any other pdf viewer or paste the link in browser.

Link to Github Repository: <https://github.com/Dharamsitejas/SpringEDAV2017>

Link to notebook 1.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/1.nb.html>

Link to notebook 2.nb.html : <https://dharamsitejas.github.io/SpringEDAV2017/2.nb.html>

Link to notebook 3.nb.html: <https://dharamsitejas.github.io/SpringEDAV2017/3.nb.html>