

The factors to get your car lost in Toronto and Vancouver

XiaoSheng (Mouder) Chen
20191125

1. Introduction

Thieves are every where. They may steal your wallet, your phone or your bag. No matter what kind of belonging get stolen, it will cause you trouble, big or small. Think about what makes you most trouble if you lose it. A key? Not really, you should always have a backup key. Money? No, it is unlikely that you will bring all of your money with you and get stolen. Your ID card? Yes, it could make you trouble, but you can still report loss and get a new card back. The answer could be your vehicle. If your vehicle gets lost, it could make you a lot of trouble even though the police can find your car back eventually or insurance company can compensate your lost.

1.1 Background

Why vehicles are stolen:

- **To sell the vehicle or parts** - Often thieves will take your car out of the Province or Country. Identification numbers may be altered or removed, ownership certificates forged, parts sold or the vehicle repainted.
- **For transportation** - Thieves will use your vehicle to get where they want to go and later abandon the vehicle. They may just drive it around to gain 'status' with their friends.
- **To commit other crimes** - Vehicles are stolen and used in other crimes such as robberies, break and enters, drug dealing or drive-by shootings.
- **For vehicle cloning** - Vehicles are stolen, given a fraudulent identification number and then sold to unsuspecting persons who are defrauded of their money. Fraudulent vehicles will be located and seized by the Police.

We can see there is a dark industry behind auto theft. As a normal citizen, what we can do is to stay alert in order to minimize the likelihood of losing your vehicle.

1.2 Problem

How can we avoid auto theft? If we are visiting a place that we know there could be higher chance for this kind of crime. We can be more careful or add extra protection when parking our car there. The problem is how we know if that area is safe or not? This assignment is to tell you what is the top ten venues to get your vehicle lost.

1.3 Target audience

The target audience is almost for everyone: e.g. citizen, traveller, owner of the venue, policemen or even government.

2. Data acquisition and cleaning

2.1 data source

Two big cities(Vancouver and Toronto) in Canada are chosen for this study. Fortunately, the auto theft data can be downloaded from the Vancouver police department website and Toronto Police service respectively. Our analyze will be base on the data from 2017-2018.

Vancouver police department : <https://vancouver.ca/police/>

Toronto police service : <http://data.torontopolice.on.ca/pages/open-data>

For the venues information, they can be retrieved from vendor source: Foursquare.

As a bonus section, this assignment will also try to consider if weather condition will impact the chance to lose your car.

2.2 data cleaning

Regarding the auto theft data of Toronto, it is in CSV(Comma Separated Value) and including the latitude and longitude for each incidents. As the data set contains more data than needed, extra records need to be filtered out and kept those data from year 2017-2018

Attributes

[Chart](#) • [Map Visualization](#)

Division Text	event_unique_id Text	Hood_ID Number	Index Number	Lat Number	Long Number	MCI Text	Neighbourhood Text
ObjectId Unique ID	occurrencedate Date or Time	occurredday Number	occurreddayofweek Text	occurreddayofyear Number			
occurrencehour Number	occurrencemonth Text	occurrenceyear Number	offence Text	premisetype Text	reporteddate Date or Time		
reportedday Number	reporteddayofweek Text	reporteddayofyear Number	reportedhour Number	reportedmonth Text	reportedyear Number		
ucr_code Number	ucr_ext Number	SHOW FEWER Attributes					

For the auto theft data of Vancouver, it is also in csv format. However, it only provides the X , Y coordinate values which are projected in UTM Zone 10. That requires a conversion from UTM format to latitude and longitude format.

- **X:** Coordinate values are projected in UTM Zone 10. All data must be considered offset and users should not interpret any locations as related to a specific person or specific property.
- **Y:** Coordinate values are projected in UTM Zone 10. All data must be considered offset and users should not interpret any locations as related to a specific person or specific property.

For the venue information, it can be retrieved from Foursquare. However, there is 5000 requests limit per hour. So it need to split into 3 batches at least in order to extract all of the venue data.

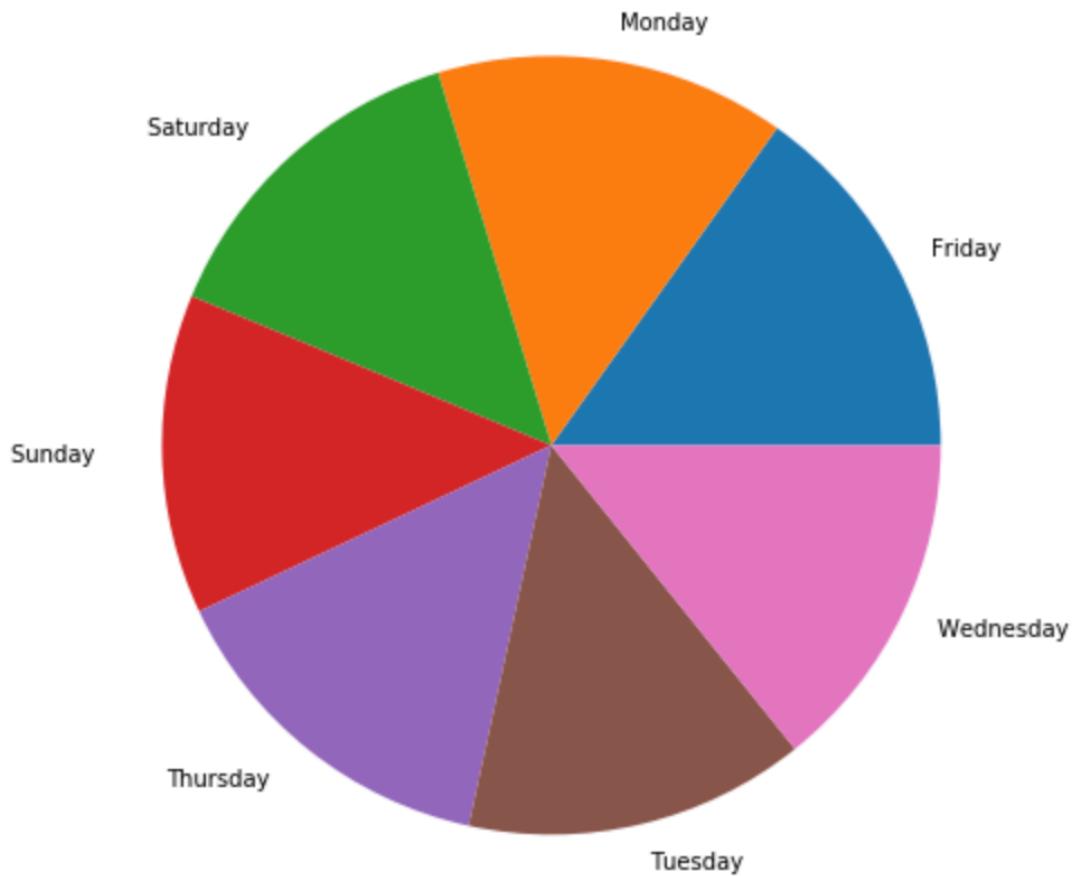
3. Explore the Data

The methodology using here is pretty straight forward. The target is to study exist auto theft data and venue information to see if there is any potential relationship.

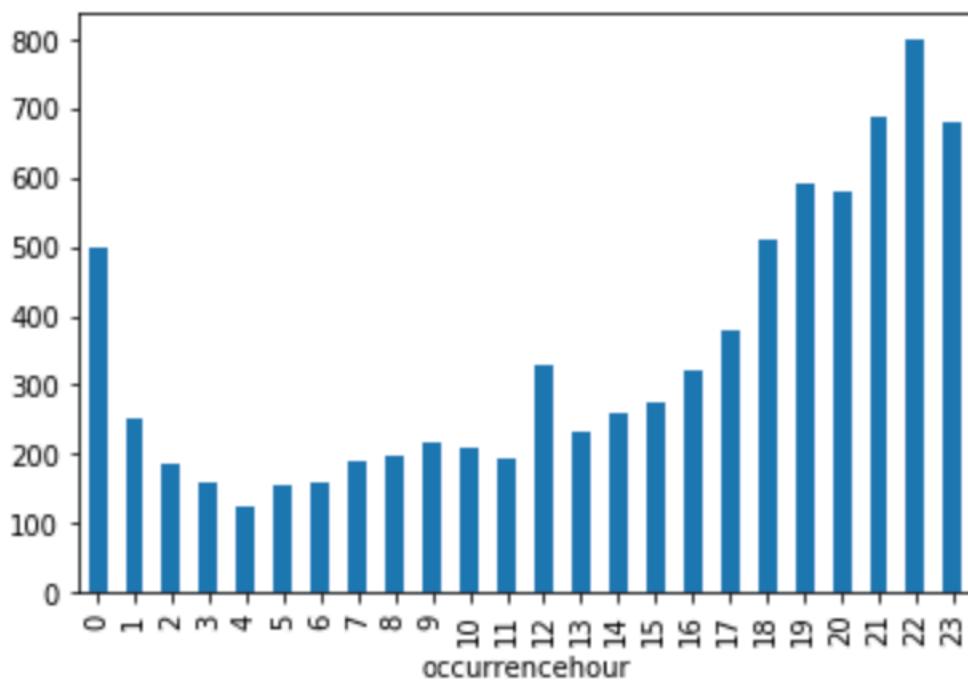
3.1 Auto theft data of Toronto

When looking into the auto theft data downloaded from Toronto police website, it contains a lot of information. e.g. date, hour, latitude and longitude, etc. The first thing we want to look at is to see if there will be more theft of vehicle incidents during weekend? By checking this, we is to compare the total number of each weekday. From

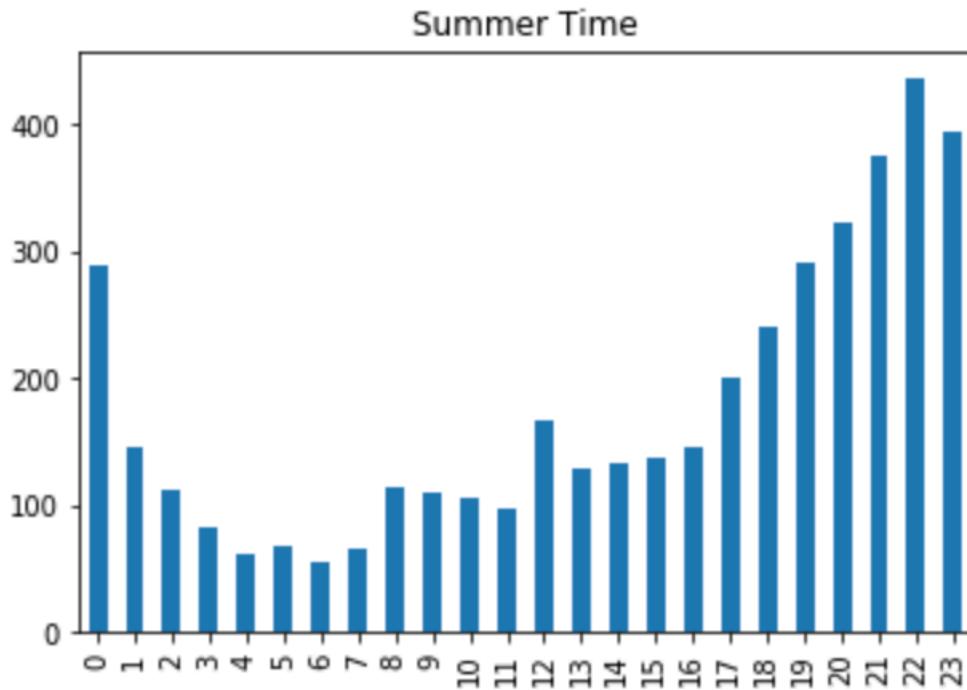
the chart below, we can see there is not much difference among weekday. So probably, the chance to lose your car is almost the same no matter it is weekday or a weekend.



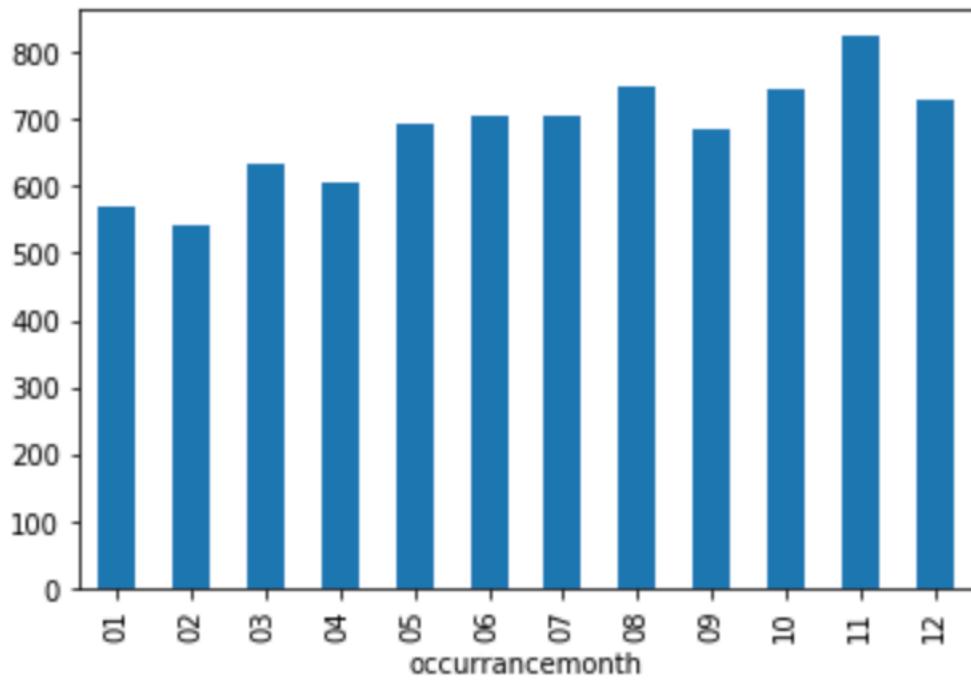
The next thing to check if the hour is a factor of losing your car. In order to do so, we will need to compare hour of the day and the total number of incidents each hour. Then following chart is generated. Now, we can see the number of incidents during midnight till afternoon is much lower than the period from evening till mid night.



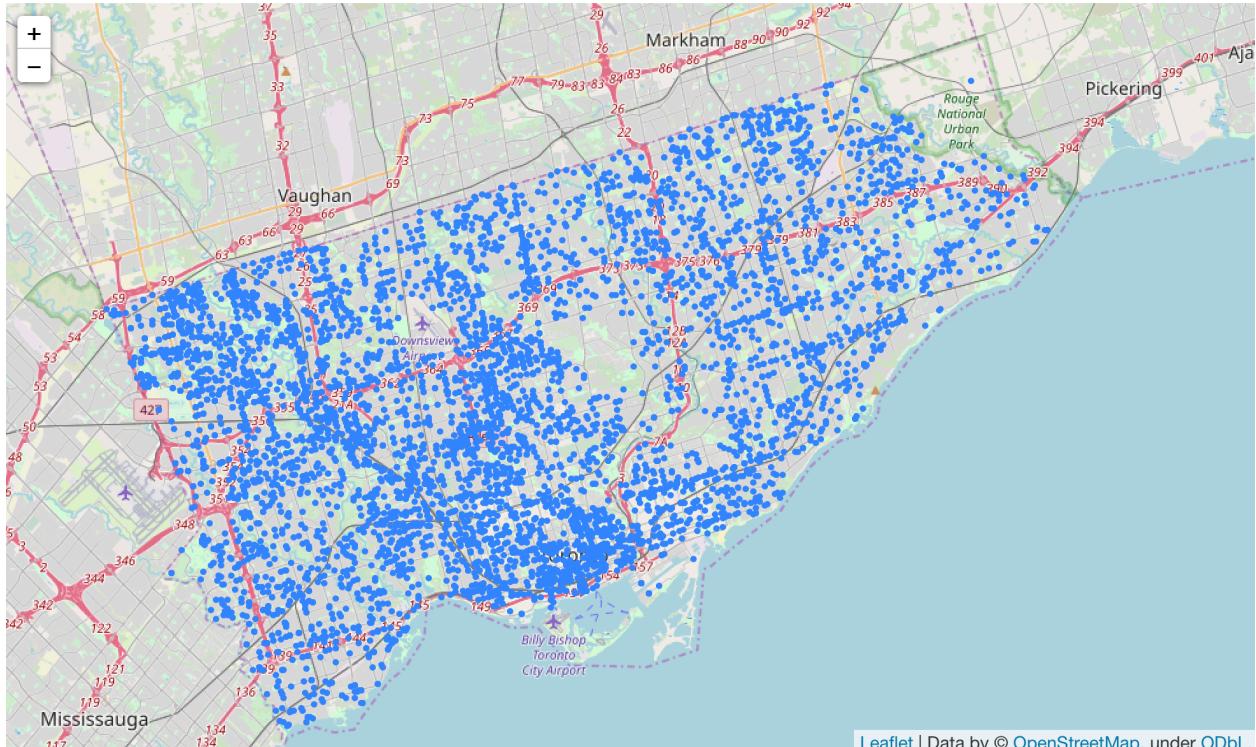
To go further, below chart is showing whether there is same case in summer given it is long sunshine hours. It is almost the same, so there is no direct relation between season and incident numbers.



When we are comparing with the month and incidents number, we do see there is slightly less incidents in first quarter than other quarters. But in terms of overall incident numbers between winter season and summer season, there is not much difference.

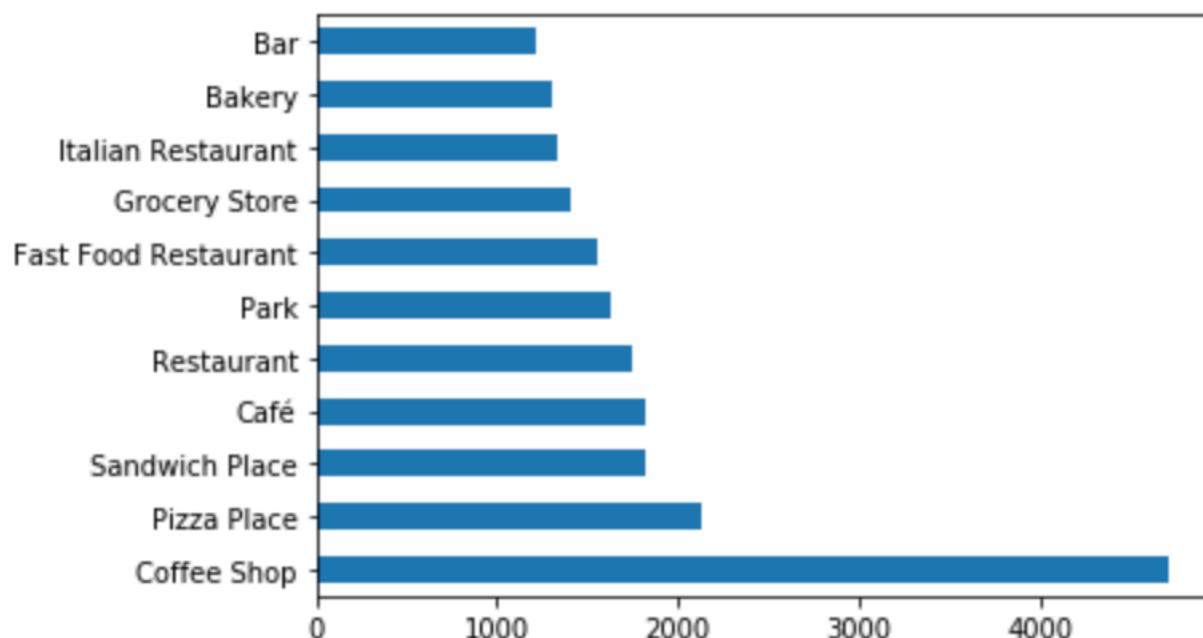


Now, it is time to see if the venues have anything to do with theft of vehicle. First, we can plot the geography chart for the place of each incident as below.



Then we can extract the top 11 occurrence venues which close to the incident

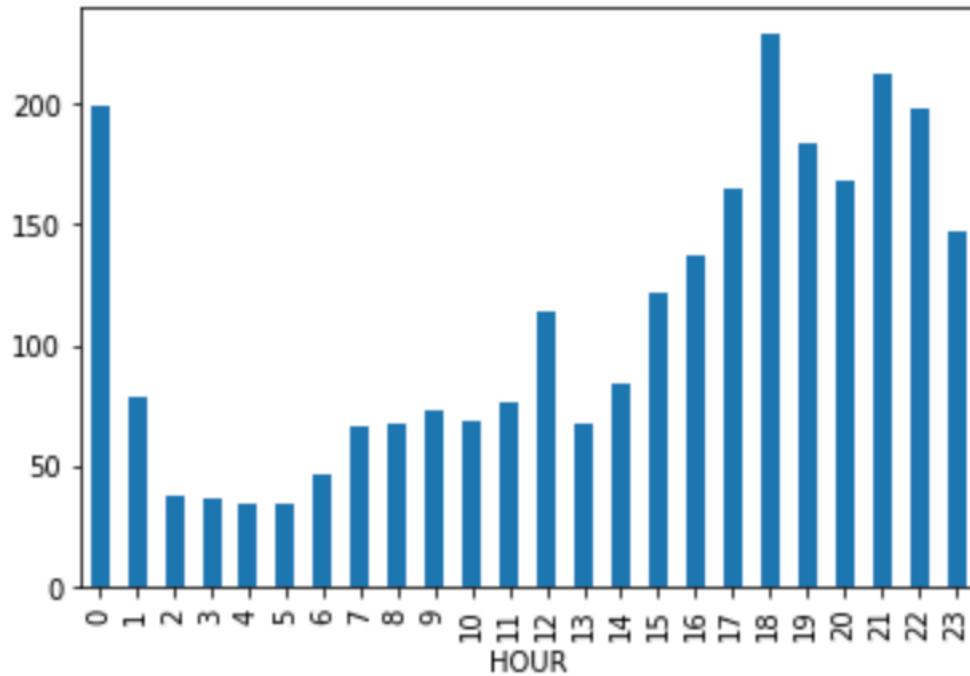
locations (with radius = 300metres). Obviously, coffee shop is so popular for auto theft. But it(coffee shop) maybe an outlier given there are so many coffee shops around Toronto. From the list, it is observed that restaurants are the highest chance avenue to get lose your car in Toronto.



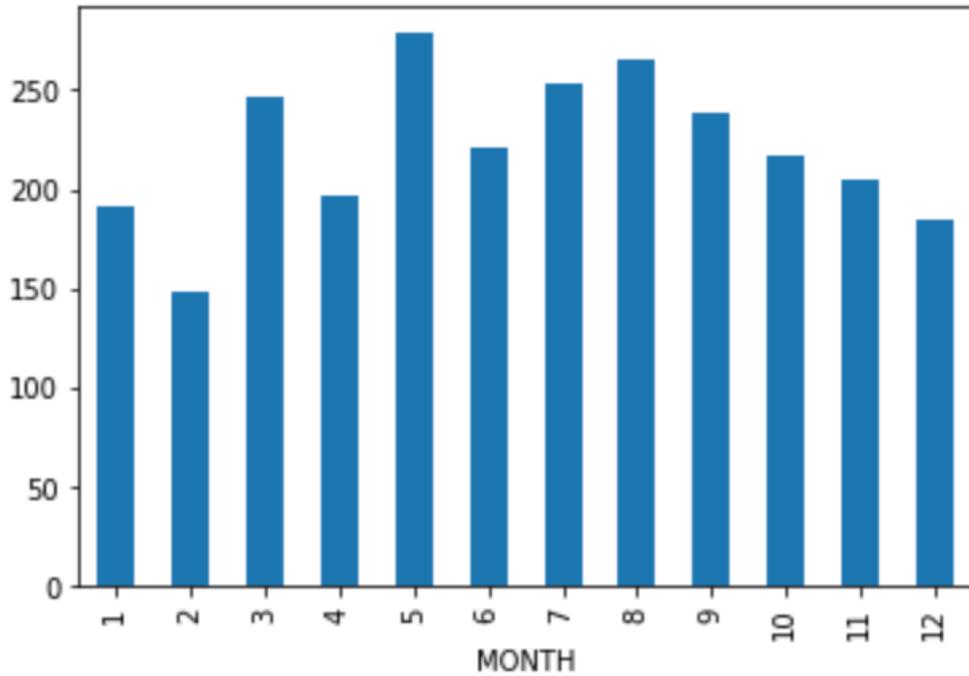
3.2 Auto theft data of Vancouver

The same analyze will be applied to Vancouver.

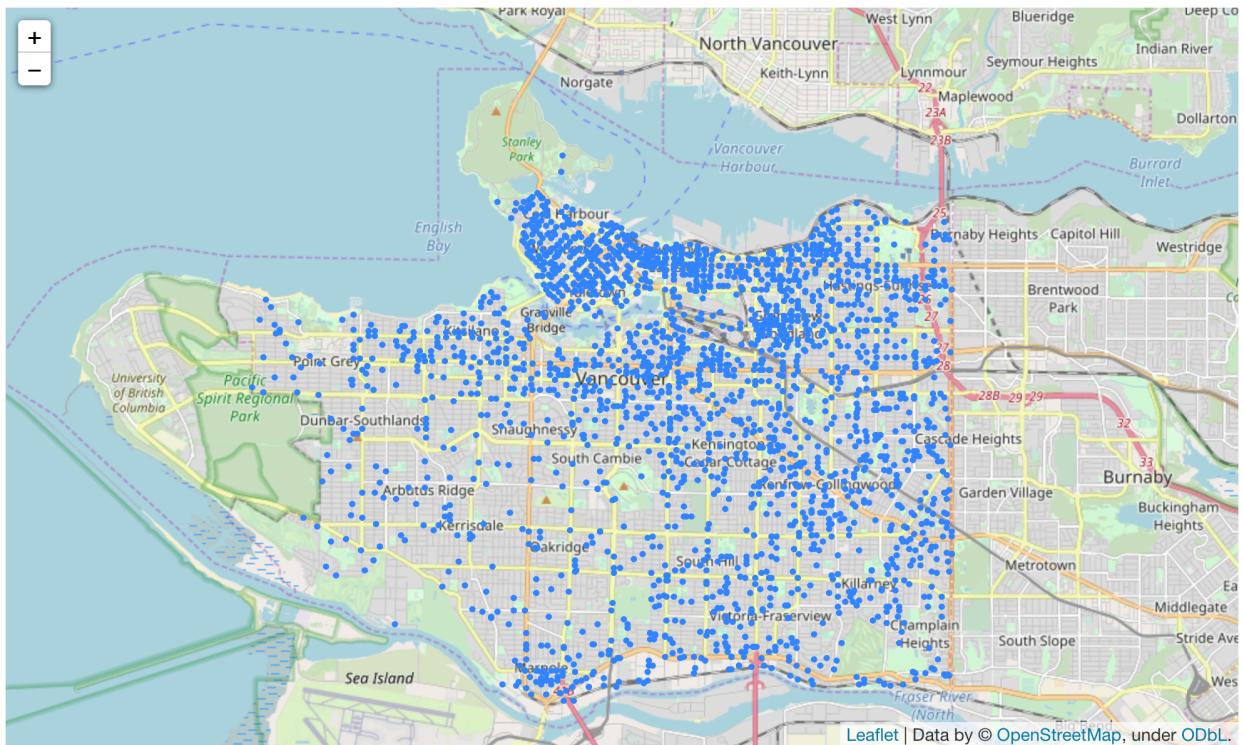
Hours vs incident numbers: The trend aligns with Toronto. There is higher chance to lose your car from evening till mid night than mid night till afternoon.



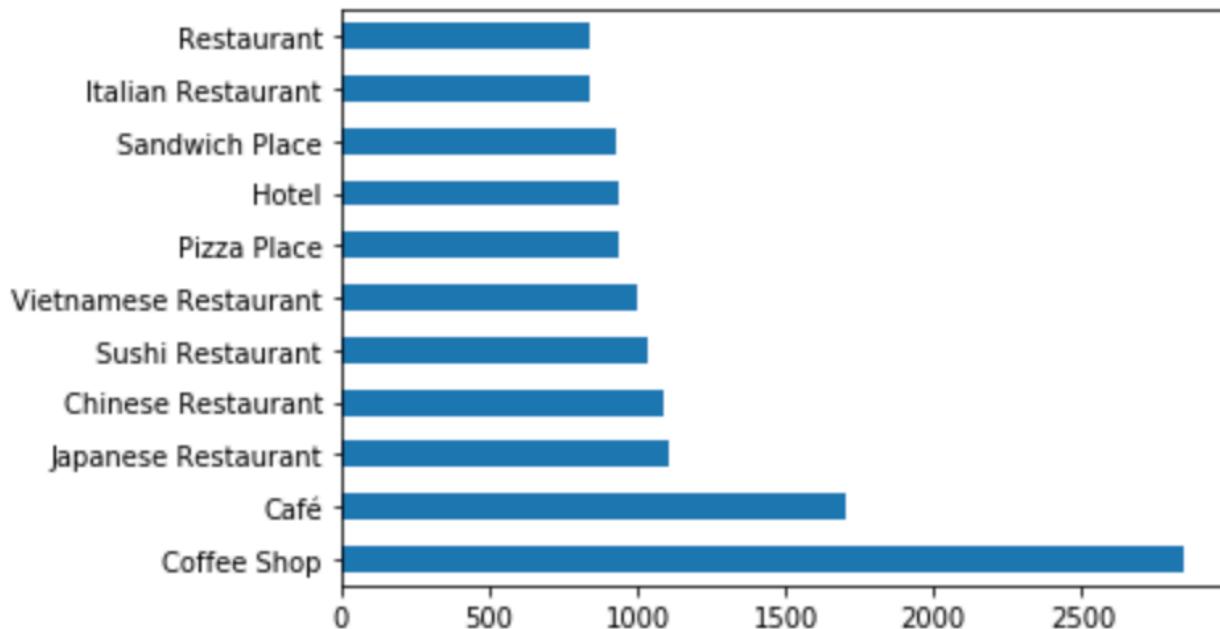
Months vs incident numbers: It is interesting to see that winter has slightly less chance to lose your cars.



Geography map for all the incidents happened in Vancouver from 2017-2018. It is more dangerous in North of Vancouver than south of Vancouver.



The chart below is showing top 11 venues which within in 300meters of each incidents. Same as Toronto, Coffee shop is again at the first place with much higher occurrences than other venues. Another similarity as Toronto is that restaurant is most dangerous places, but it is interesting that Asian restaurants are on the top of the list.



4. What this report found

Though there are only handful features are selected to analyze if they have any correlation with the chance of losing vehicle, we do have some interesting finding:

- The chance to lose your car is pretty much the same for all the day of the week, no matter it is weekday and weekend (Toronto)
- Evening till mid night has 2-3 times chance to lose vehicle than other hours during the day (Toronto & Vancouver)
- Places near restaurants are higher chance (Toronto & Vancouver)
- Asian restaurants are even higher chance than other restaurants in Vancouver
- Park is also a potential place on top 11 list (Toronto)
- For unknown reason, February is the lowest months for theft of vehicle. (Toronto & Vancouver)
- There are no significant discrepancy for the incidents happen at each month
- North of Vancouver has higher chance than south of Vancouver

For the traveler/resident, they will need to raise their alert level when visiting restaurants at night. Especially for the people who visiting north of Vancouver during that period.

For police, probably, they can increase patrol on those high risk locations in order to

mitigate this kind of criminal.

For the venue owner, they can think how to improve their security and that can be a selling point to attract customers.

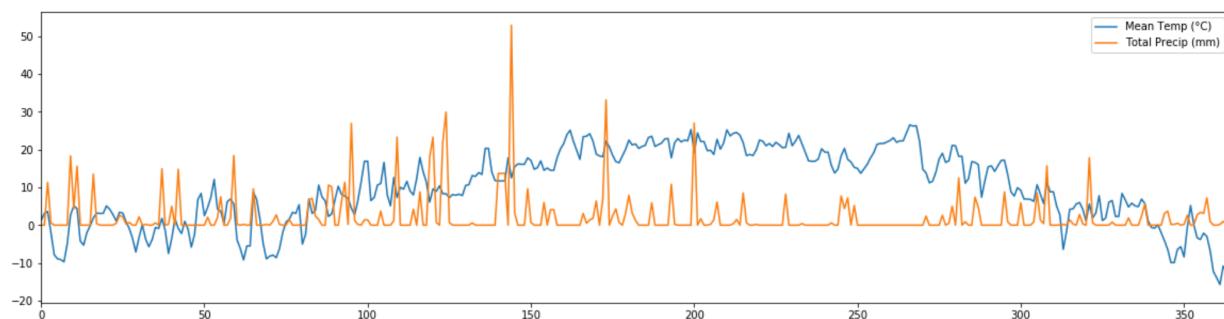
5. Conclusion

Even though there is no complicated method using for the analysis, we can still identify some of relationships between the features we selected and occurrence of the incidents. I believe that it could help for the individual to understand the situation and raise their alert and cooperation can seek improvement on how to provide safer environment for their business. Moreover, it could help the society to build a better place for residents who live in the cities. The same approach can be also applied to other cities.

6. Future direction

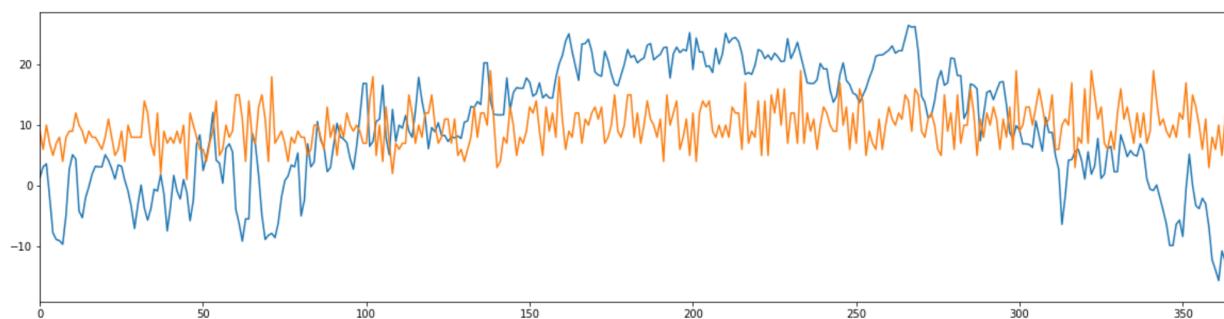
More data and features can be added for the study in order to find out if there are more factors that have correlation with the occurrence of the incident.

For example, if the temperature or has anything to do with the incidents. To validate that we can download the weather data from government website. Then we extract the temperature and precipitation in Toronto for 2017 as following:

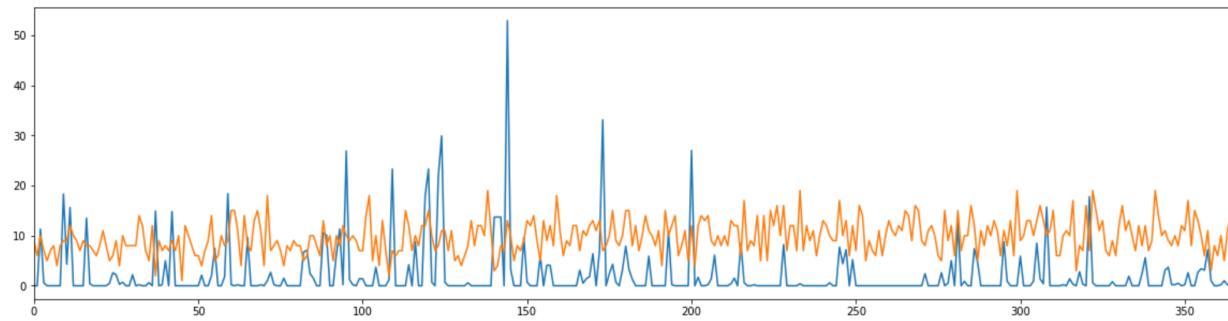


And then to compare with the daily number of incidents:

Incident vs temp.



incident vs precipitation



One interesting thing we found is that the incidents will rise for those days when there are rain fall. So traveller and resident should also aware of this.