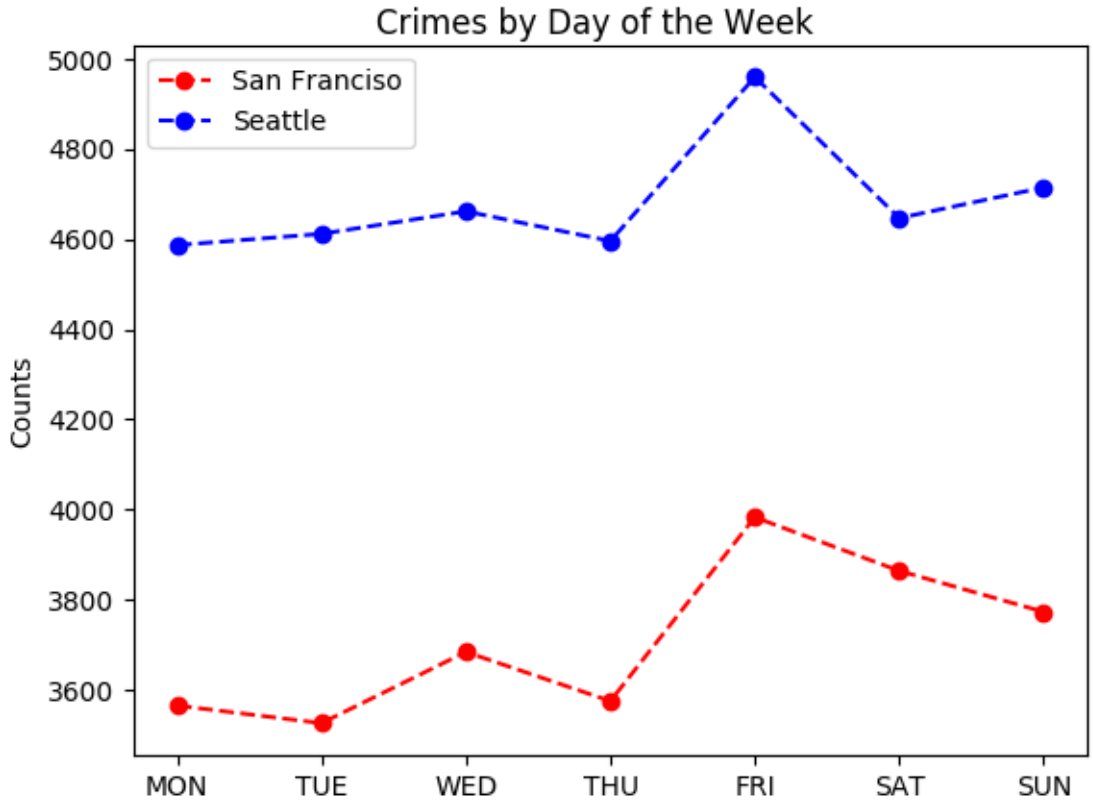
In this short study we analyse crimes reported during summer 2014 in Seattle and San Francisco. Our objective is to provide a clear overview of what is going on in both cities, unavailing the most significant patterns in data without providing too many details at the same time.

Crimes are recorded a bit differently in the two datasets in CSV format, but both of them contain location, time and a certain categorization (theft, assault, etc). The Seattle data has the specificity of having a start/end time and date for each crime as well as a reporting date. This is significant for instance when reporting a fraud that happened on a long period. On the other hand, the San Francisco dataset reports how crimes were resolved (e.g. person arrested, not prosecuted or unfounded). One needs to notice that the San Francisco data also contain non criminal events (e.g. reported missing person) that has to be removed for crime analysis.

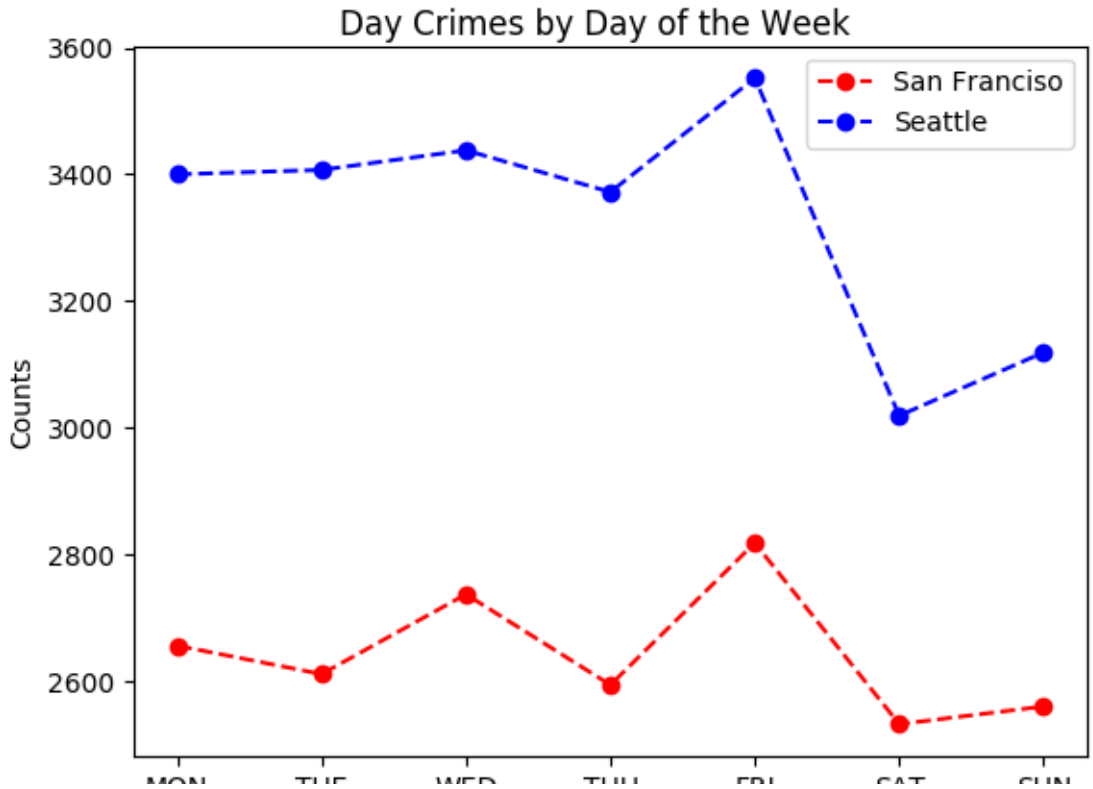
Overview

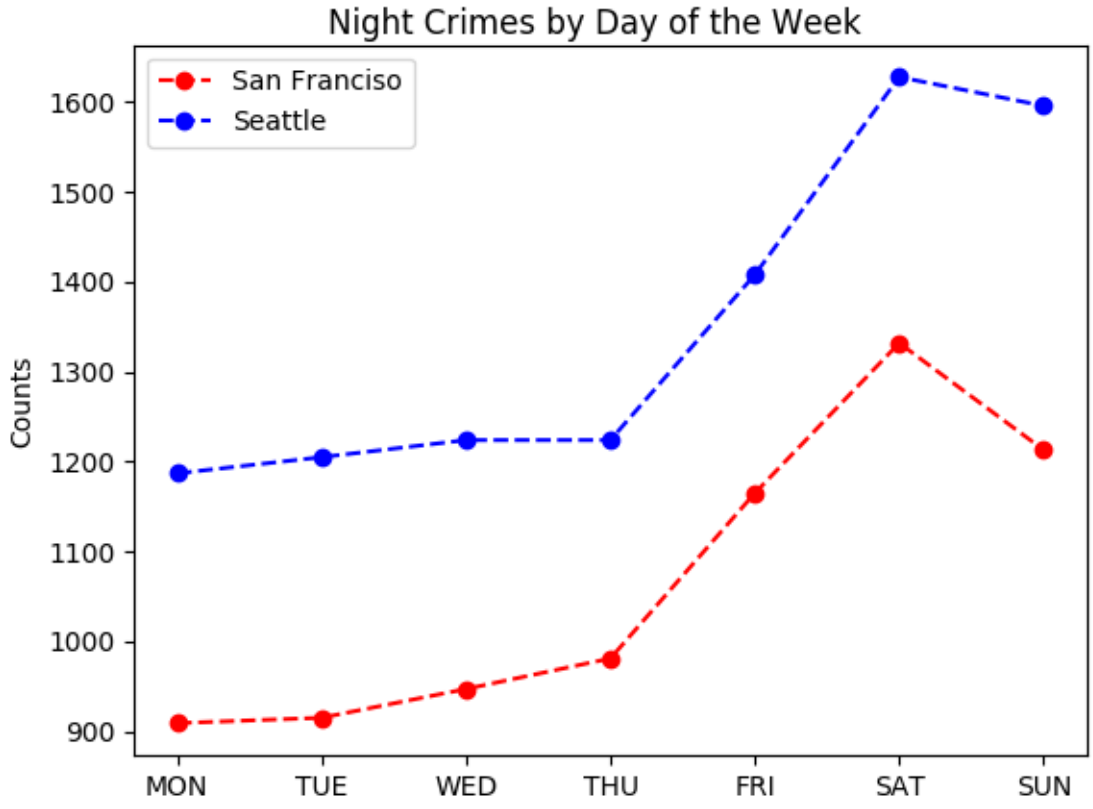
We start by a quick overview of crime in both cities. We first study how crimes are distributed on a typical week or a typical day. By plotting the total crime counts by day of the week for both cities, we find a singular common pattern: the beginning of the week is more quiet, then a pike of crime happens on Friday and it decays during the weekend. However one key difference between the two is the high level of crime tends to continue during the weekend in San Francisco while in Seattle the pike ends on Saturday.



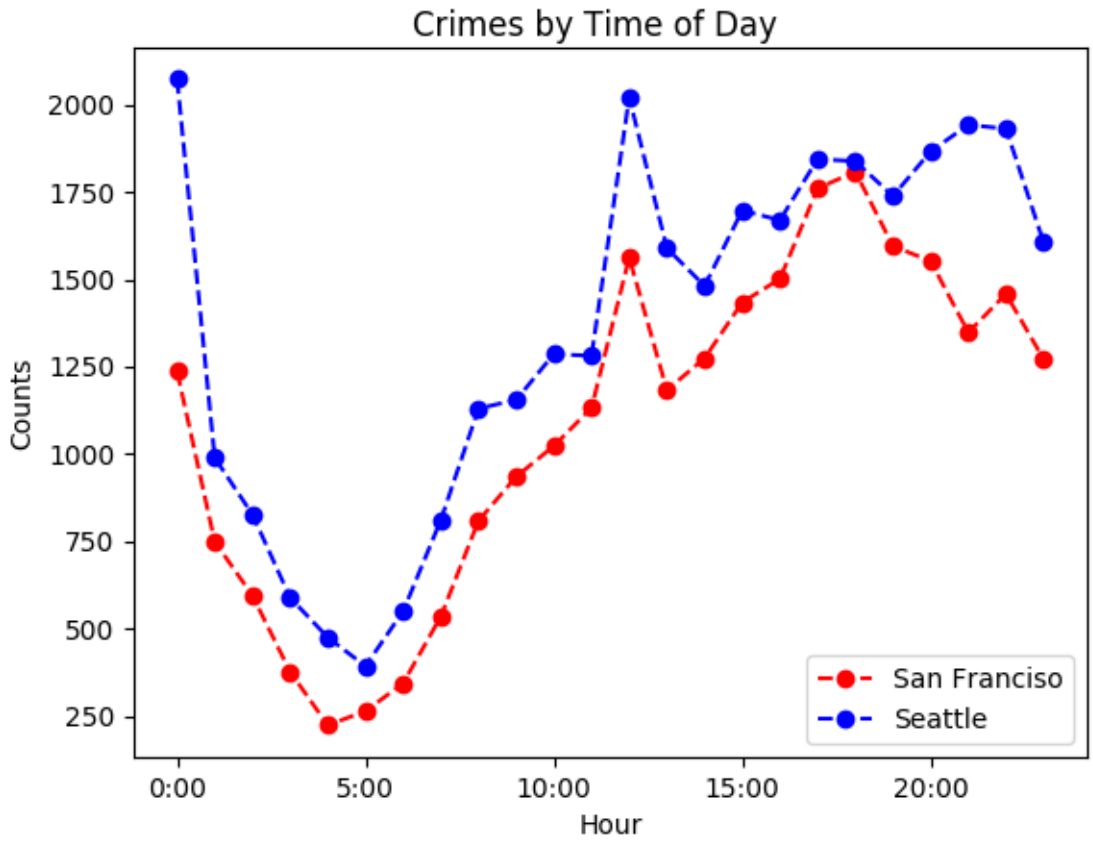
To get a clearer view, we separate crimes committed during the day and at night. To do that, we consider the definition of night as in Common Law: 30 minutes after sunset to 30 minutes before sunrise. External data containing sunrise/sunset times for various locations allows us to split day

and night crimes. If we plot the same curves as above for day and night crimes, and we find an even more accurate pattern: the spike of day crime happens on Friday, while night crime level is the highest during the weekend in both cities.





Another point of view consists in displaying crime counts as a function of time of the day. We discover another interesting pattern: Seattle and San Francisco crime counts as a function of time of the day are highly correlated. Their common trend is as follows: the number of crimes reported decays during the night then starts to increase in the early morning (around 5:00). It gradually increases and spikes at noon, then increase during the afternoon. On noticeable fact is Seattle crime has two spikes at noon and midnight which are equivalent, whereas in San Francisco the spikes occur at noon and around 16:00, the second being higher.



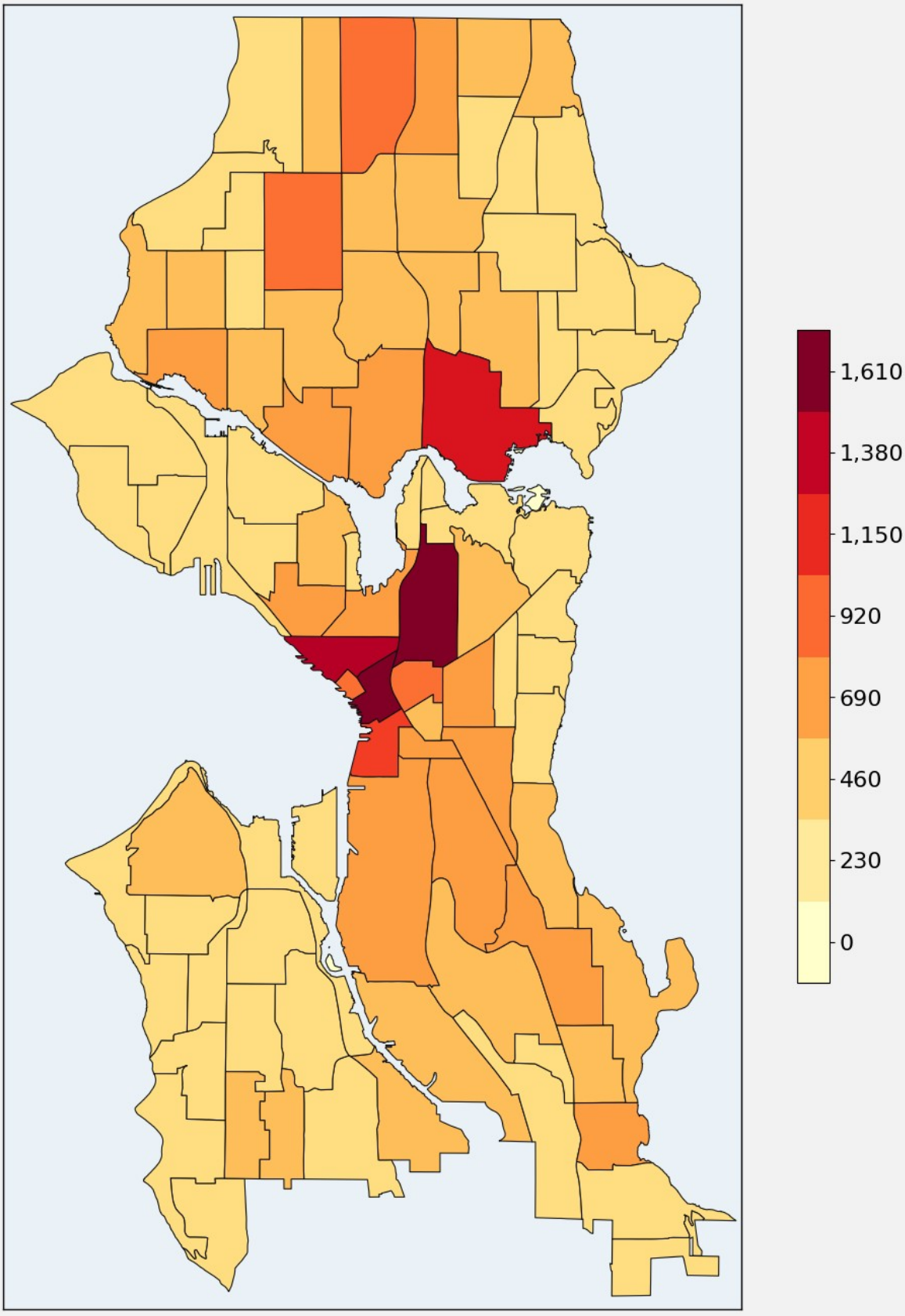
Now that let us dig into the details for each city separately.

Detailed analysis of crime in Seattle

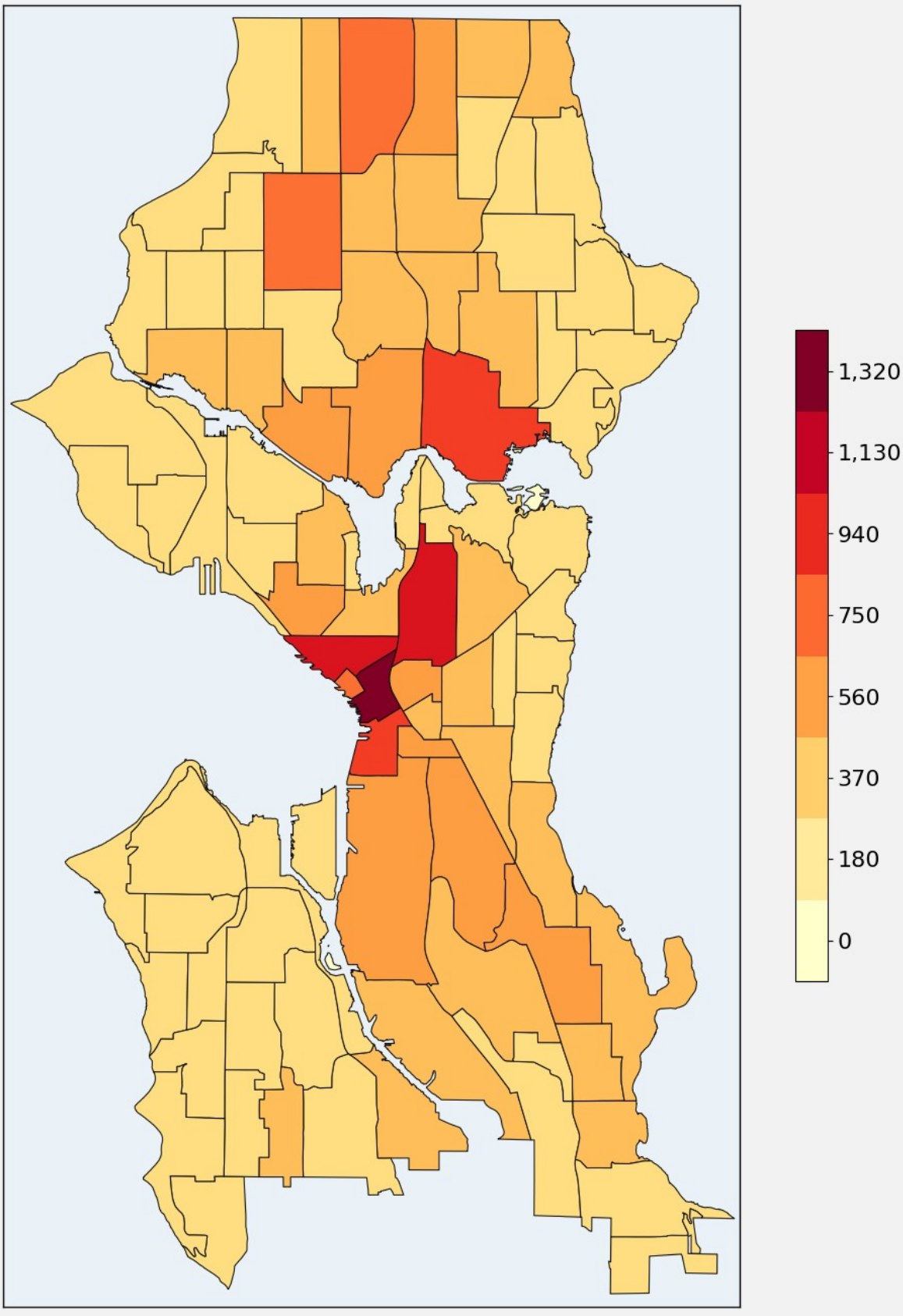
Thanks to latitude and longitude reported for each datum, we are able to split crimes into neighbourhoods. We display counts in a heatmap and observe that central neighbourhoods of Capitol Hill, Downtown and University District have the highest number of reported crimes.

If we split day and night crimes, we observe that the distribution is not too far, except that some neighbourhoods with high crime during the day tend to have less during the night. The most noticeable such area is Downtown (which makes sense since this is the business district of Seattle).

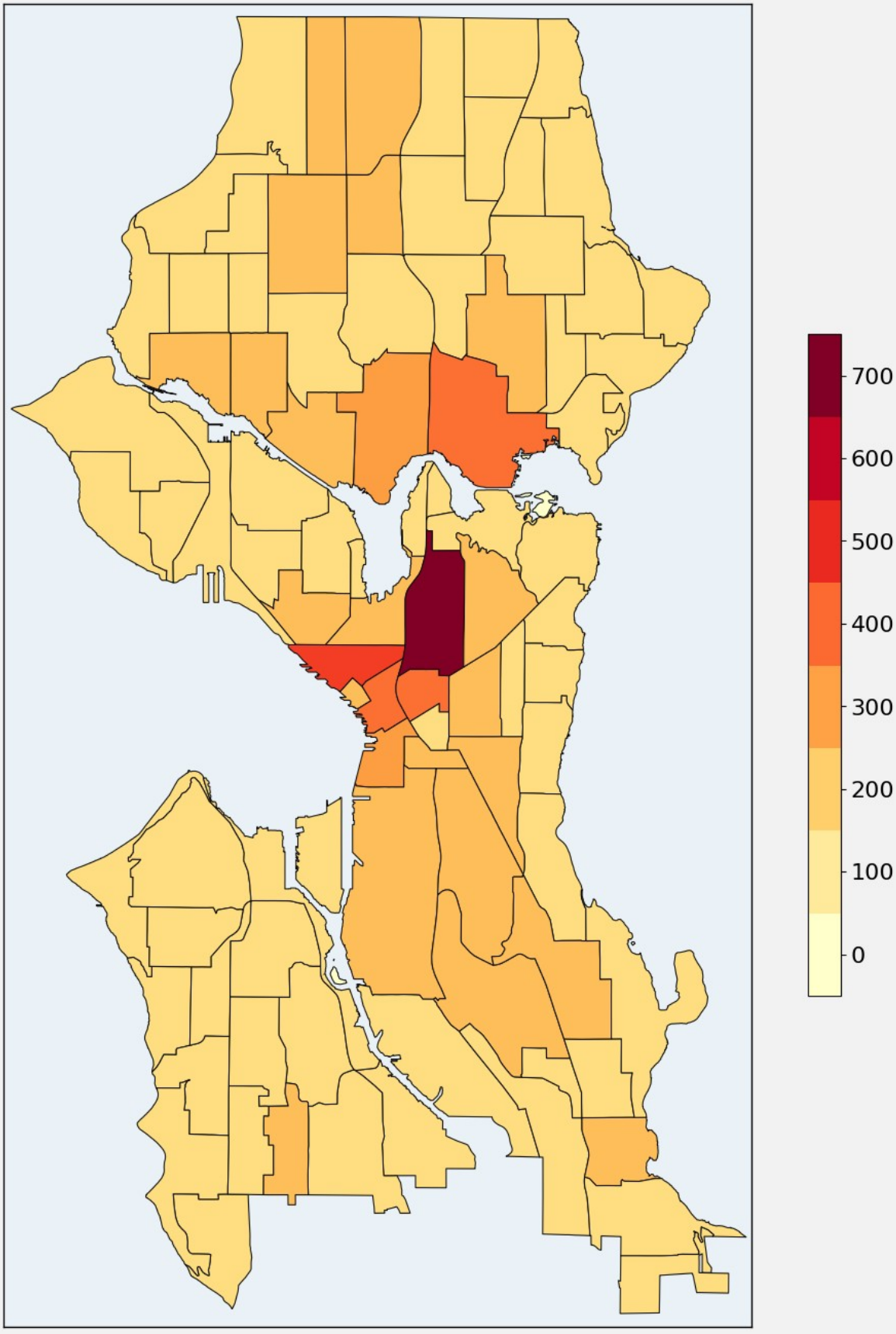
Crimes Reported in Seattle by Neighborhood



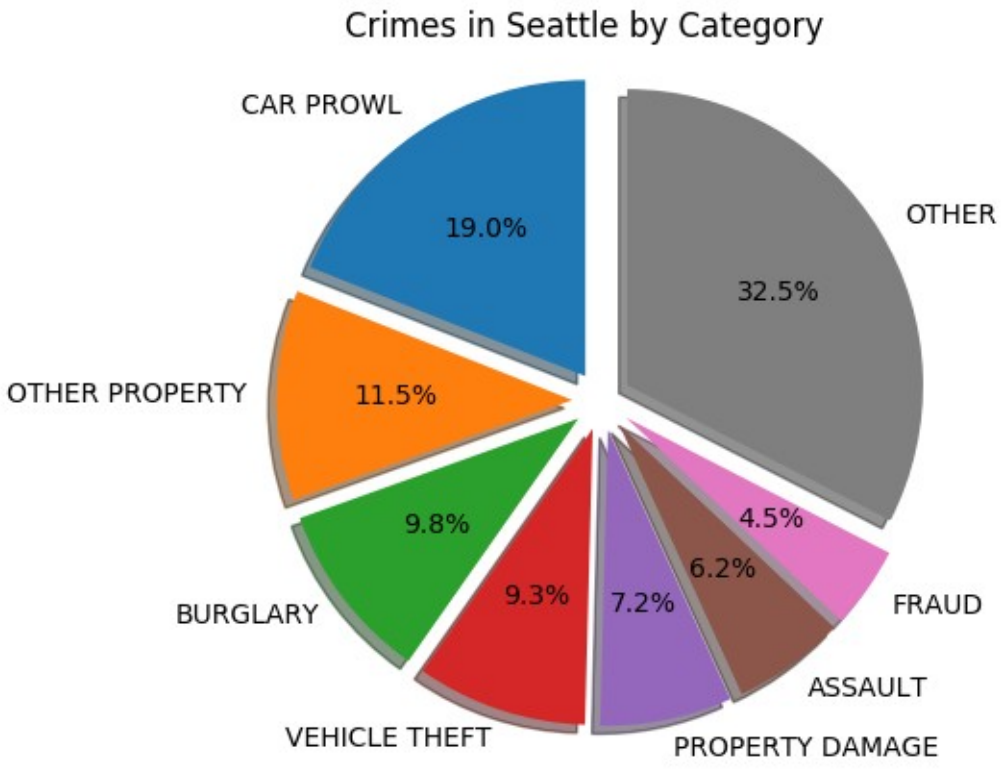
Day Crimes in Seattle by Neighborhood



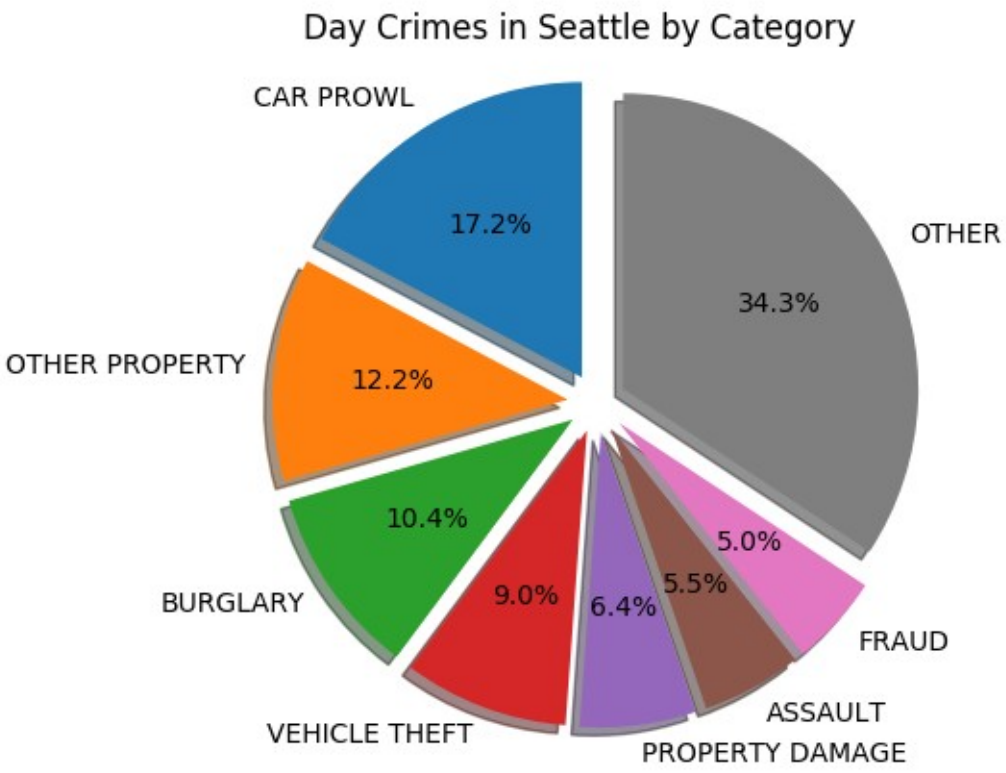
Night Crimes in Seattle by Neighborhood

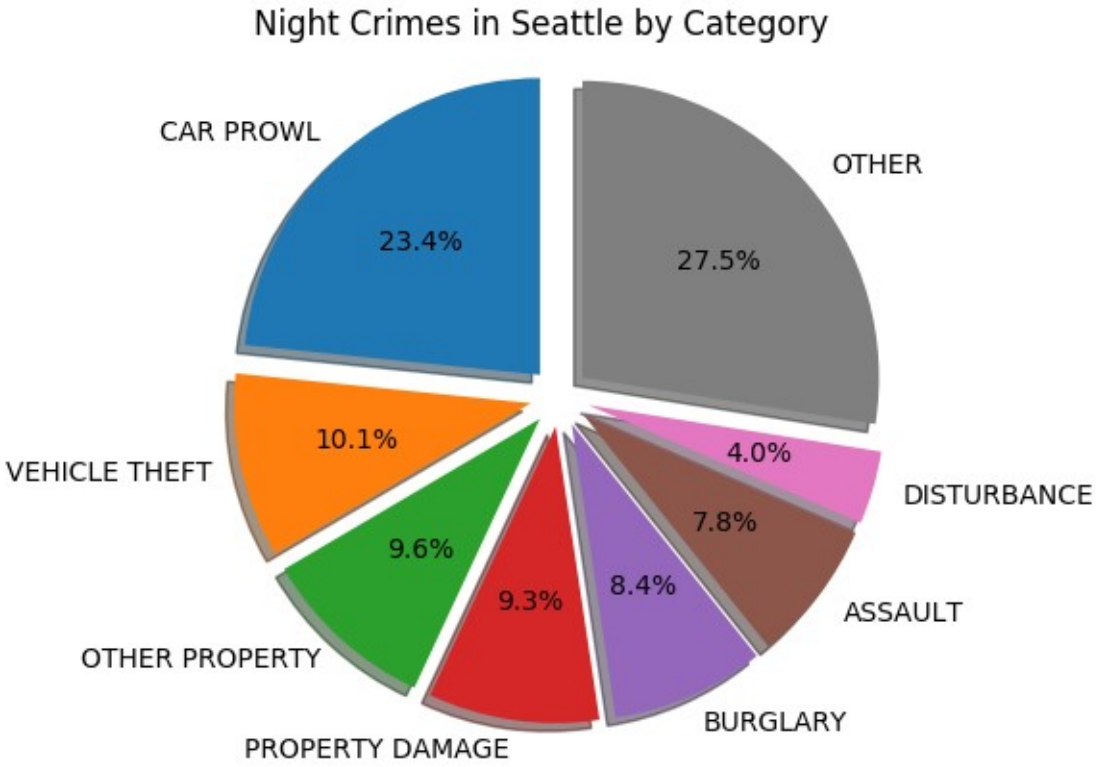


Now we take a look at the most common crimes in Seattle. Overall, the distribution of crimes are represented in the pie chart below.

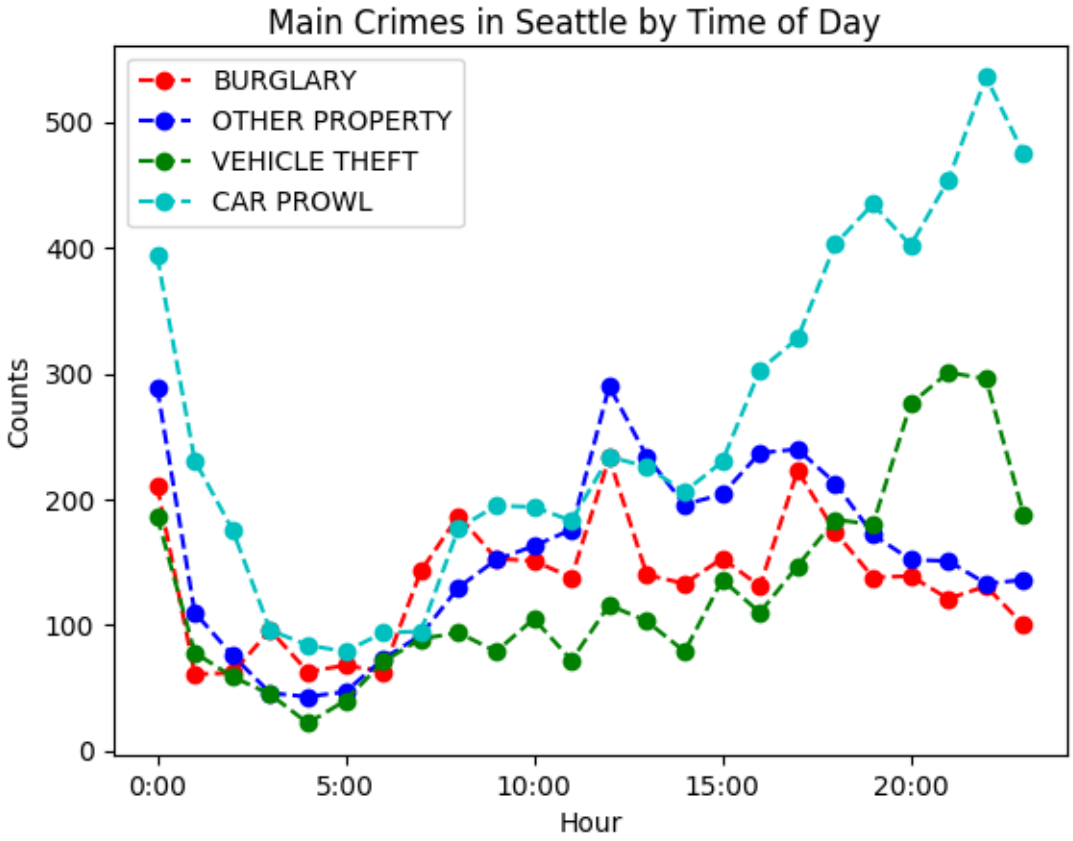


CAR PROWL is the largest issue in Seattle, followed by OTHER PROPERTY (which corresponds to various kinds of theft), BURGLARY and VEHICLE THEFT. If we separate by day and night, we can see the crime overview is mostly driven by day crime, which makes sense because if accounts for the largest part. However, VEHICLE THEFT becomes the second highest crime at night and BURGLARY tends to decrease at the same time.


Day Crimes in Seattle by Category
CAR PROWL
17.2%
OTHER
34.3%
OTHER PROPERTY
12.2%
10.4%
BURGLARY
9.0%
6.4%
5.5%
5.0%
FRAUD
ASSAULT
VEHICLE THEFT
PROPERTY DAMAGE

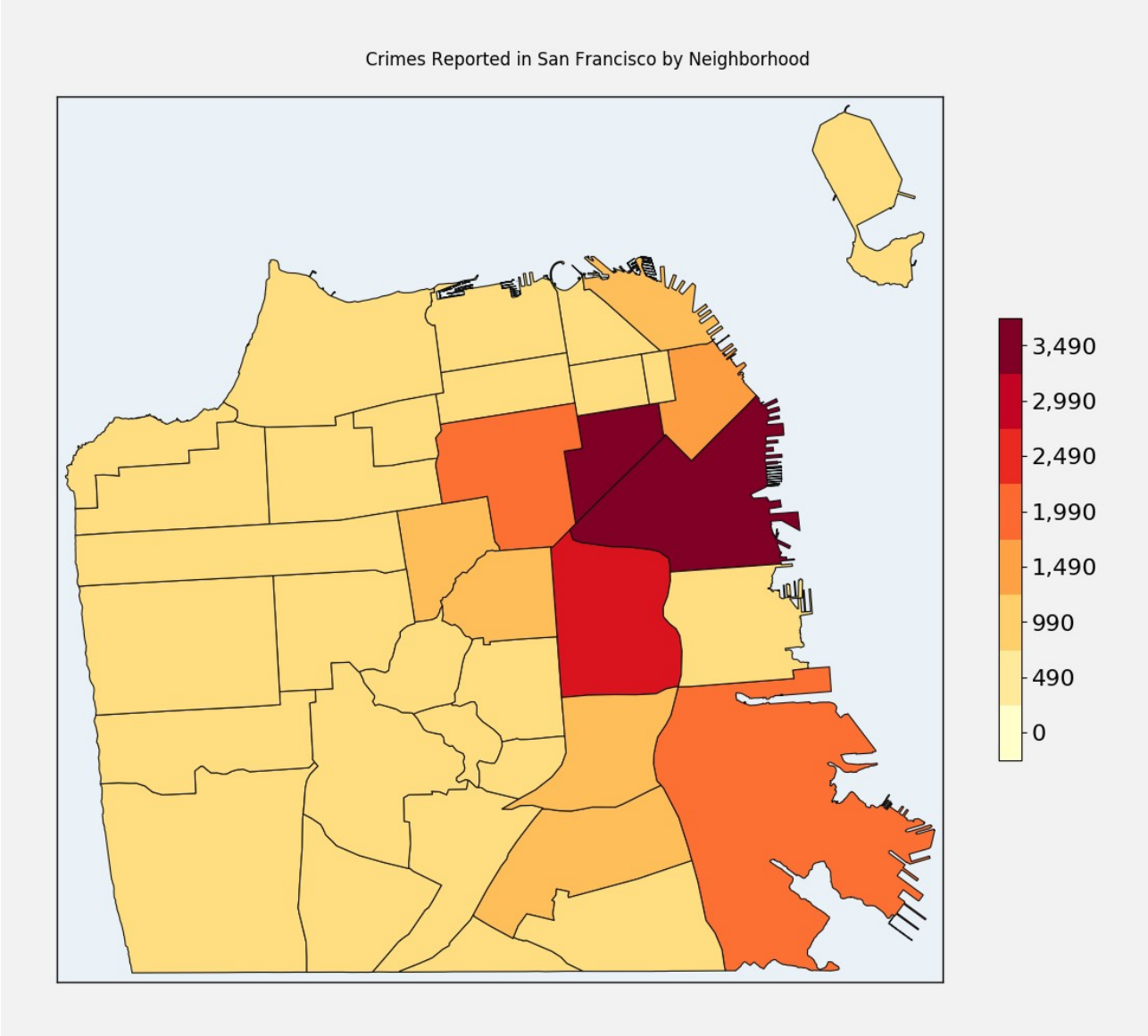

Night Crimes in Seattle by Category
CAR PROWL
23.4%
OTHER
27.5%
10.1%
VEHICLE THEFT
4.0%
DISTURBANCE
7.8%
9.6%
8.4%
9.3%
ASSAULT
OTHER PROPERTY
BURGLARY
PROPERTY DAMAGE

We can get a better visual impression of this difference between day and night crimes by plotting the crime counts for the main categories (see below). This clearly show that the VEHICULE THEFT category explodes around 20:00 while OTHER PROPERTY and BURGLARY decay below this category.



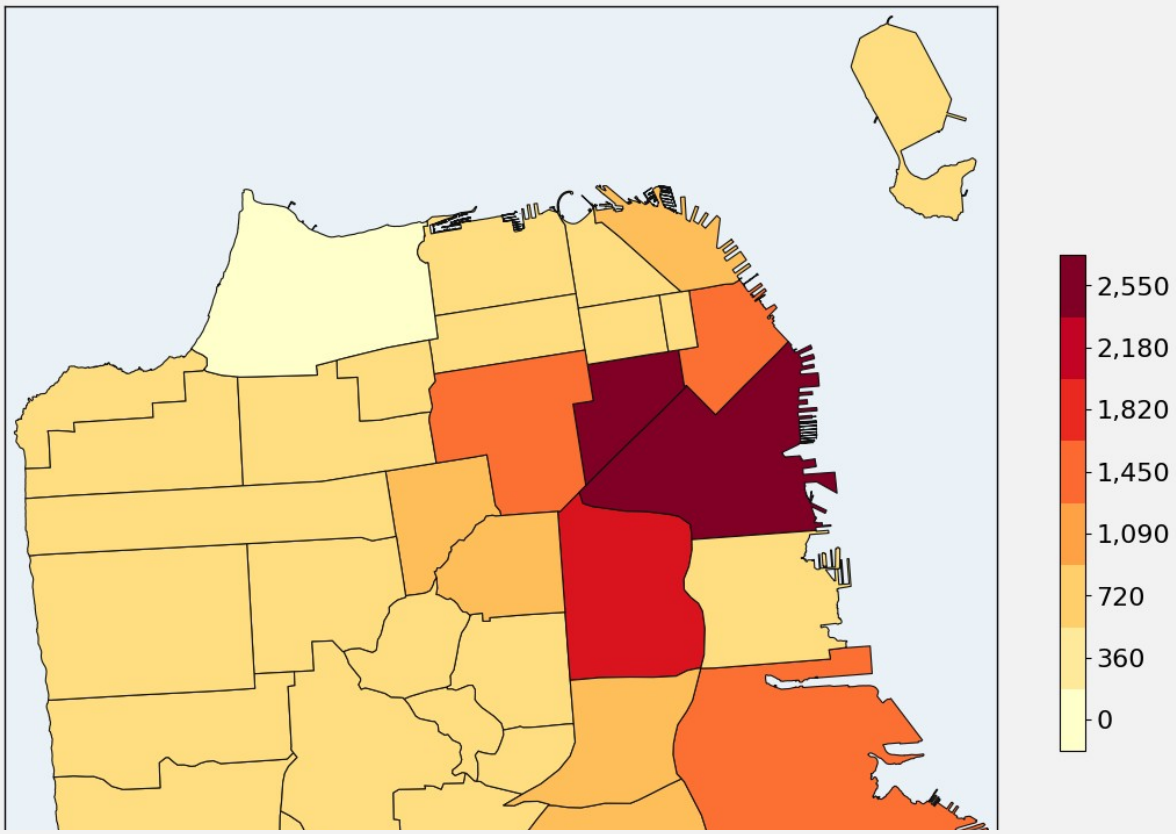
Detailed analysis of crime in San Francisco

We create the same geographical heatmap as above for San Francisco. We observe that central neighbourhoods of South of Market, Downtown, Mission District have the highest number of reported crimes. The scale is the number of crimes reported during summer 2014,

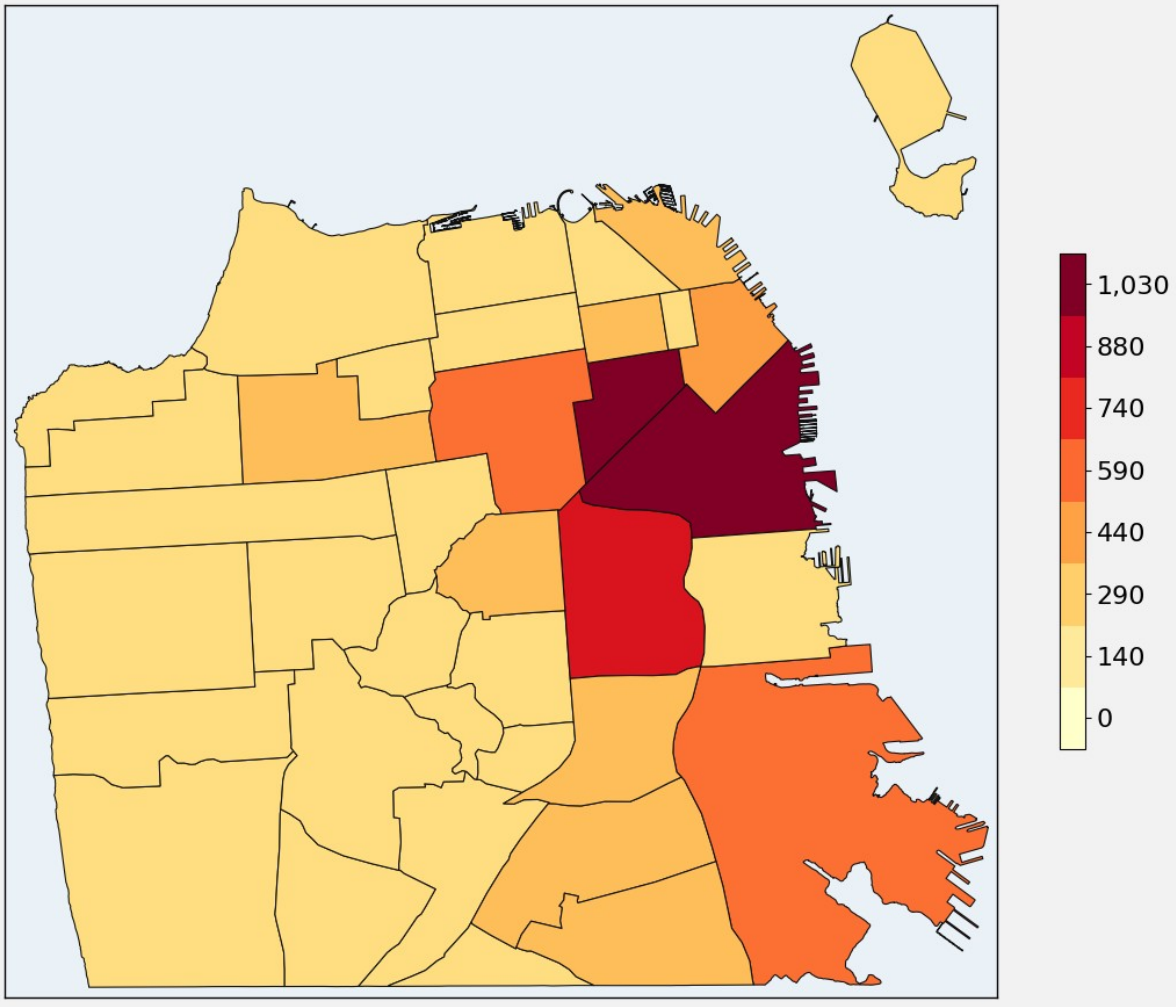


When spiting between day and night crimes, we observe that the same phenomenon as for Seattle: the distribution of crimes at night is not too far from during the day, except for some neighbourhoods (see below).

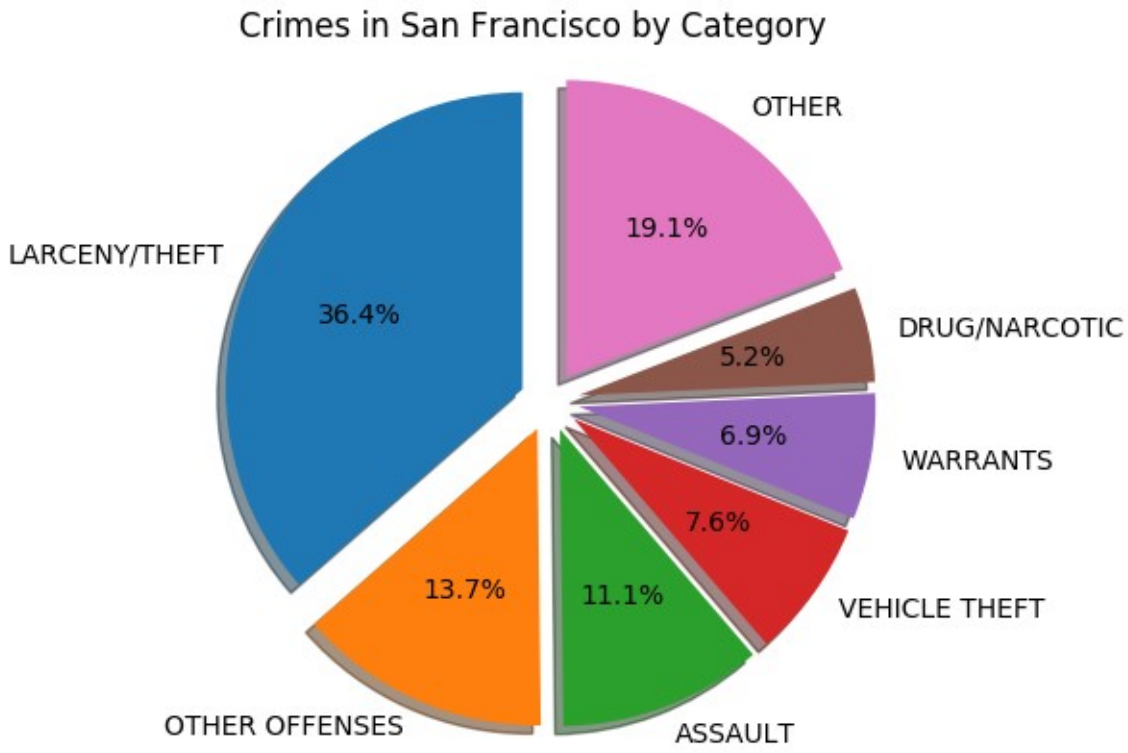
Day Crimes in San Francisco by Neighborhood



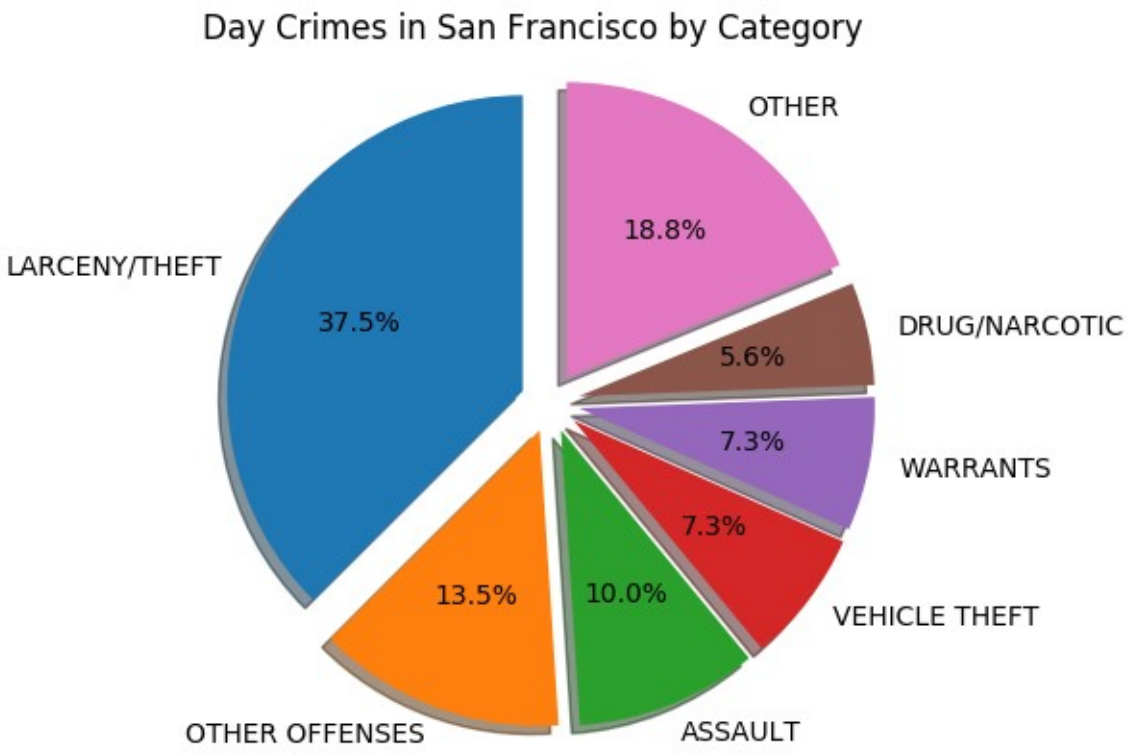
Night Crimes in San Francisco by Neighborhood

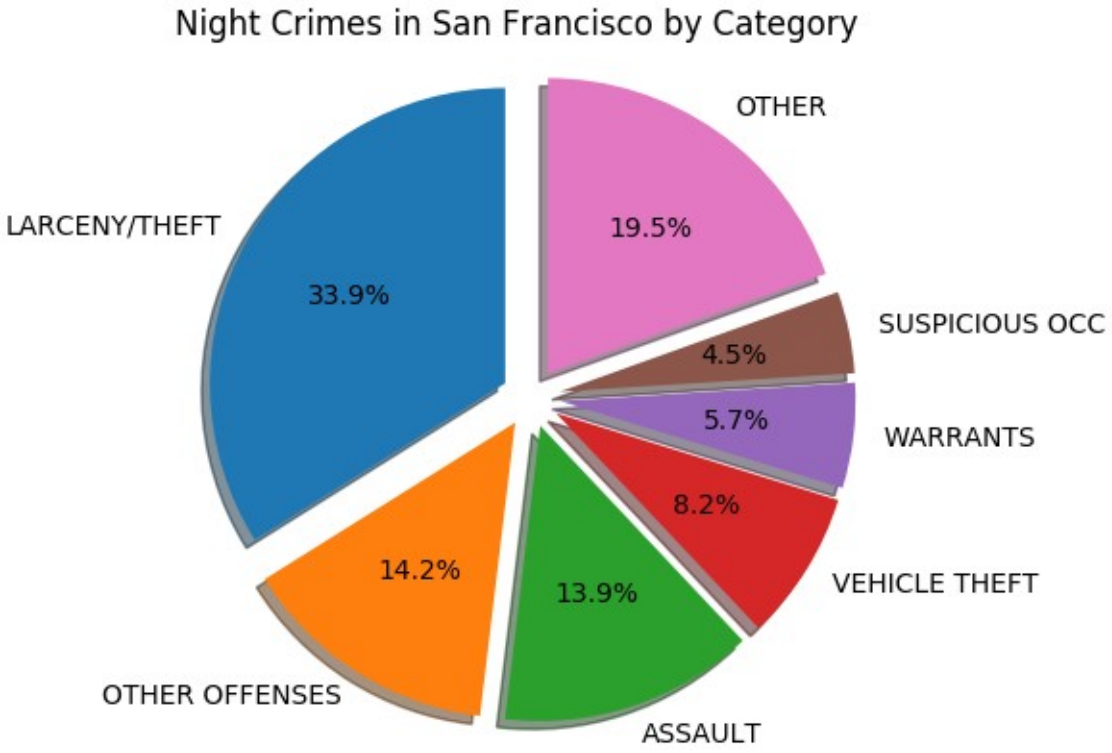


Now let us take a look at the most common crimes in San Francisco.



LARCENY/THEFT is the largest issue in San Francisco, followed by OTHER OFFENSES, ASSAULT, and VEHICLE THEFT. If we split by day and night, we can see very small variation in the distribution of crimes. We do not need to go further and investigate how main categories behave like we did for Seattle.


Day Crimes in San Francisco by Category
OTHER
18.8%
LARCENY/THEFT
37.5%
DRUG/NARCOTIC
5.6%
7.3%
WARRANTS
7.3%
VEHICLE THEFT
13.5%
10.0%
OTHER OFFENSES
ASSAULT


Night Crimes in San Francisco by Category
OTHER
19.5%
LARCENY/THEFT
33.9%
SUSPICIOUS OCC
4.5%
5.7%
WARRANTS
8.2%
VEHICLE THEFT
14.2%
13.9%
OTHER OFFENSES
ASSAULT

Takeaways

- We observe a typical pattern of the level of crime as a function of day of the week for both Seattle and San Franciso
- When analysing crime levels intra-day, we find that again Seattle and San Francisco follow a common pattern
- Plotting crime counts on a map reveals which neighbourhoods are the most dangerous.
- Whereas Seattle has two different crime regime during day and night, San Francisco is more stable

That study was the occasion to develop a set of tools we can reuse for more seeking other patterns, or for analysing longer periods.