LA Crime Analysis

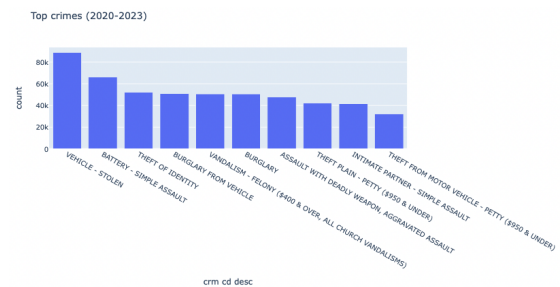Aim 1: Conduct EDA to identify trends
Aim 2: Investigate into crimes which require special attention from responding officers, i.e. high risk crimes (HRCs)
Aim 3: Identify actionable intelligence with regards to law enforcement and HRCs
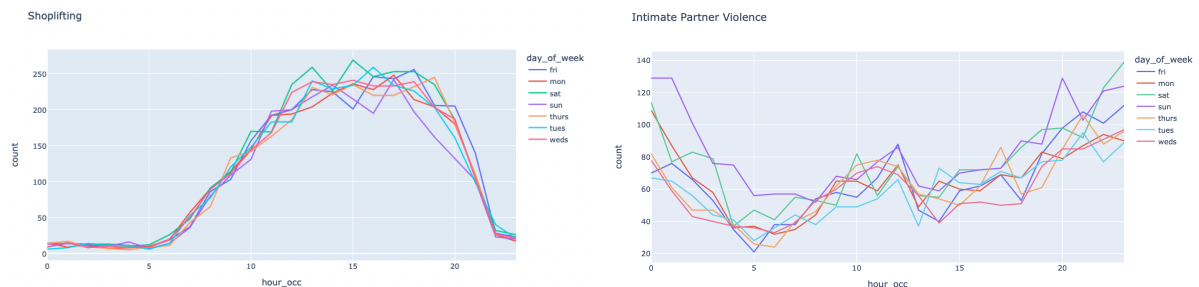
---

1.   EDA

1.1 Top crimes

By sorting the data by crime code, it was clear that vehicle thefts made up the highest number of crime occurrences for time period 2020-2023.
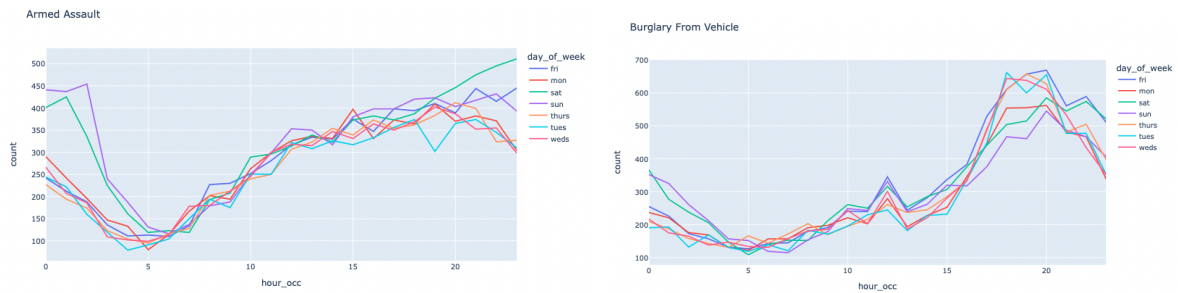

1.2 Crime trends by day and time

When plotting crime occurrences against day and time, some interesting patterns started to emerge. Some patterns were more explainable than others. For instance, shoplifting peaks once shops open, and drop once shops close. Intimate partner violence is generally lower during the day when people tend to be at work, and increases after office hours when people tend to be at home.
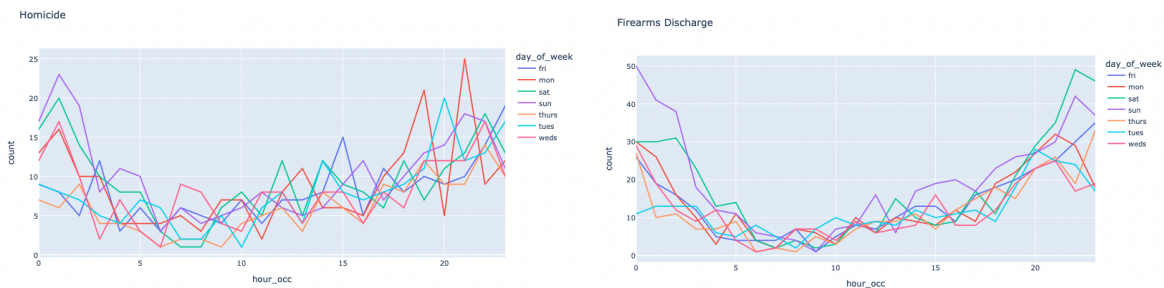

Crimes such as vehicle thefts, robberies,burglary from vehicles were also observed to follow a similar cosine curve pattern, with the lowest number of crime occurrences at about 5am and the peaks between about 6-10pm.

Other crimes such as Homicide and Firearms Discharge showed a different U shaped curve pattern, indicating that occurrences were lower during the day, and higher at night/early morning between 10pm to about 2am.
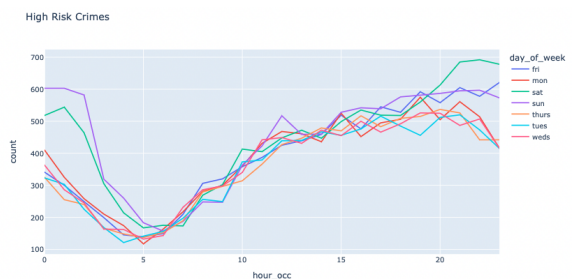

2.   high risk crimes

In order to achieve my second aim, I decided to zoom in on crime occurrences which may require special attention from responding officers. I created a dataframe out of crimes which involved the use of weapons and thus pose a higher risk to police, and named these as 'High Risk Crimes (HRCs). The specific crime codes, descriptions and count of HRCs can be found in the table on the right.

	crm cd	crm cd desc	count
0	230	ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT	47711
1	761	BRANDISH WEAPON	12934
2	753	DISCHARGE FIREARMS/SHOTS FIRED	2410
3	251	SHOTS FIRED AT INHABITED DWELLING	1506
4	110	CRIMINAL HOMICIDE	1415

2.1 Day and time trends for HRCs

When plotting by time, HRCs also exhibit a cosine-shaped curve, with higher occurrences between 8pm-2am, and especially on weekends.

2.2 Examining location hot spots for HRCs

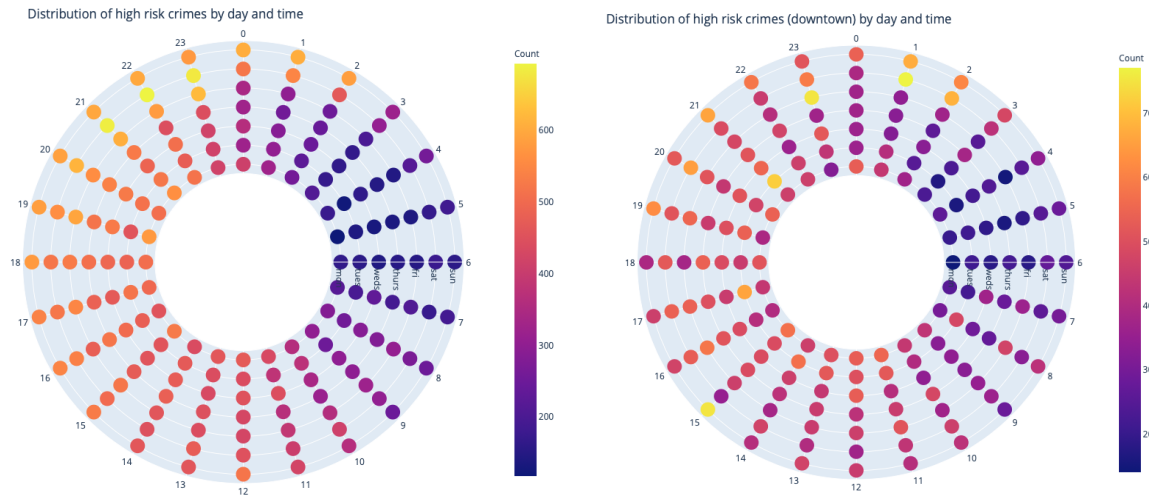|	location	lat	lon	count
0	600 S ALVARADO ST	34.0595	-118.2749	93
1	800 N ALAMEDA ST	34.0561	-118.2375	83
2	2200 W 6TH ST	34.0601	-118.2761	64
3	500 S SAN PEDRO ST	34.0437	-118.2455	58
4	600 S SAN PEDRO ST	34.0421	-118.2469	50
5	5TH	34.0442	-118.2439	43
6	400 E 7TH ST	34.0421	-118.2469	39
7	500 S LOS ANGELES ST	34.0460	-118.2493	39
8	600 S ALVARADO ST	34.0593	-118.2750	37
9	500 SAN JULIAN ST	34.0446	-118.2450	37
10	500 S SPRING ST	34.0467	-118.2520	36
11	5TH	34.0435	-118.2427	33
12	600 S FIGUEROA ST	34.0508	-118.2586	33
13	500 S SAN PEDRO ST	34.0442	-118.2439	32
14	600 S SPRING ST	34.0460	-118.2510	31
15	HOLLYWOOD	34.1016	-118.3387	31
16	3100 WILSHIRE BL	34.0652	-118.2878	31
17	6TH	34.0423	-118.2453	30
18	00 ROSE AV	33.9946	-118.4805	30
19	6600 TOPANGA CANYON BL	34.1904	-118.6059	30

Using a Python mapping package called Folium to plot coordinates on a heatmap, it became observable that there were 6 main hotspots for HRCs in LA(circled), with the highest number of occurrences being in Downtown LA(green circles). The locations of the nearest LAPD police stations to HRC hotspots are indicated by the blue markers, which would be where the first responders are likely to be based. These stations may thus require better equipped/experienced officers to handle a higher occurrence of HRCs.

Nearest police stations:
- LAPD Rampart Station
- LAPD Headquarters
- Central Community Police Station

2.3 Examining HRCs in Downtown LA

HRCs were plotted on a custom polar scatter template I created with Plotly below, which compared the distribution by day and time for the general population of HRCs, as well as for HRCs occurring in Downtown LA. When examining the general HRC population, the peak HRC occurrence can be observed on Saturdays between 9-11pm, with occurrences remaining especially high on weekends from about 4pm to about 2am. However, the distribution of HRCs in the downtown area can be observed to be less consistent. Occurrences are more sporadic, with peaks observed on Saturdays between 1-2am, Sundays between 3-4pm, and Fridays between 11pm-12mn. That Sundays between 3-4pm was a peak was surprising, as it broke away from the usual trend of high occurrences being late at night. This anomaly seems to be worth examining in more detail (although outside the scope of this project).

Distribution of high risk crimes by day and time

Distribution of high risk crimes (downtown) by day and time

3. Final conclusions

- In order to better protect responding officers, LAPD offices situated closest to HRC hotspots, namely LAPD Rampart Station, LAPD Headquarters, and Central Community Police Station, might be better off having officers who are better equipped/experienced to deal with HRCs
- With regards to HRCs occurring within Downtown LA, certain time anomalies were observed which broke away from usual trends involving HRCs in general, such as the high occurrence of HRCs on Sundays between 3-4pm, which may require higher police patrols/manpower on standby.