PM566 Final Project: Analysis of burglaries and victims in LA

YINA LIU

Introduction

According to Crosstown, burglaries in Los Angeles have remained stable since the Los Angeles Police Department (LAPD) began diclosing its crime data in 2010, with a high of 17,465 incidents that year and a low of 15,125 incidents in 2014. Although LAPD have tried their best to safeguard the lives and the property of the citizens these years, burglary is still one of the most frequent crimes in Los Angeles city. In this project, my objective is to analyze Los Angeles' burglary-related data to explore the pattern of burglaries, which can not only enhance people's awareness of self-protection but also can help reduce such crimes.

The primary interest of this project are:
- To find out if there is an association between burglary and victims' gender, age or race.
- To find out when and where burglaries happened the highest and the loweset in LA in 2020.
- To look at the arrest rates of crime in the Los Angeles area.

Methods

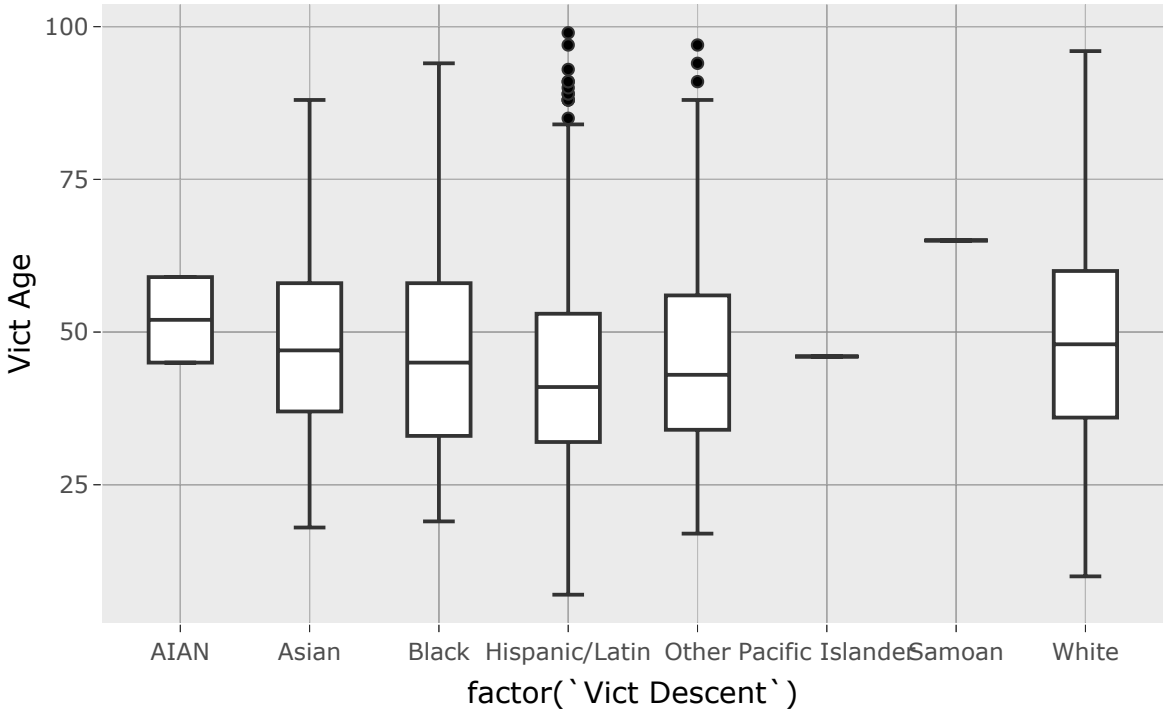
The original dataset is obtained online from 'Los Angeles Open Data' (https://data.lacity.org/A-Safe-City/Crime-Data-from-2020-to-Present/2nrs-mtv8 (https://data.lacity.org/A-Safe-City/Crime-Data-from-2020-to-Present/2nrs-mtv8)). Besides, I also introduced arrest data in the City of Los Angeles dating to 2020 as an assisted dataset (https://data.lacity.org/A-Safe-City/Arrest-Data-from-2020-to-Present/amvf-fr72 (https://data.lacity.org/A-Safe-City/Arrest-Data-from-2020-to-Present/amvf-fr72)). Data about crimes, victims and arrests of burglary from January 1st, 2020 to October 5th, 2020 were collected from these two datasets.

After importing, the data were carefully checked for variable names and types. Full descent names, provided in the data dictionary, were used instead of the victims' descent codes as to help illustration. Descendants who have experienced only a few burglaries were merged together. For example, 'Japanese' and 'Korean' were both classified as 'Asian'. A new variable was created called 'month_occurance', which refers to the month in which each burglary occurred. Finally, missing values and implausible values for key variables such as '0' in the age of the victims and '(0°,0°)' in the latitude and longitude were removed.

Following by data cleaning and wrangling, several graphs and plots were created for data visualization. The association between age, gender and descent of the victims were revealed through a boxplot and a barchart. I also generated a barchart to show the number of burglaries in different areas, along with a leaflet map showing the detailed location of each burglary. Furthurmore, two line graphs were used to illustrate the trend of number of burglaries by month. A scatter plot was also created to explore the arresting situations in different areas in LA. What's more, I generated three tables to display the complete information mentioned above.

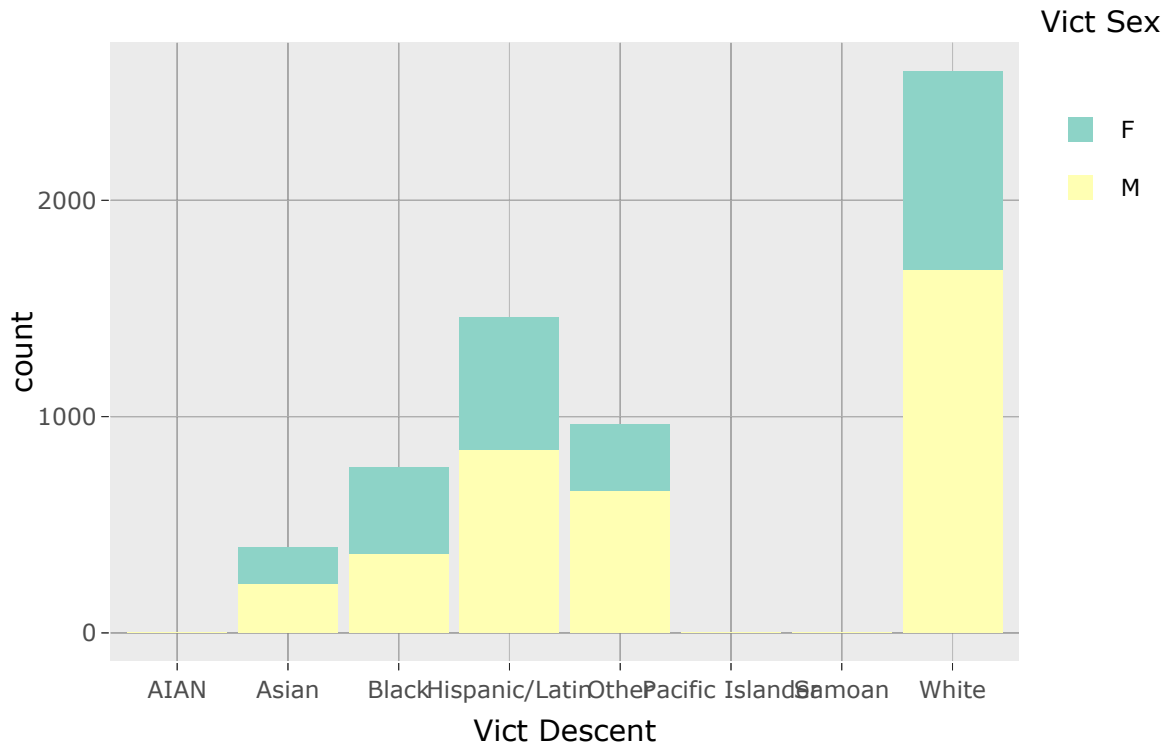
Results

Boxplot 1: The relationship between victims' age and descent



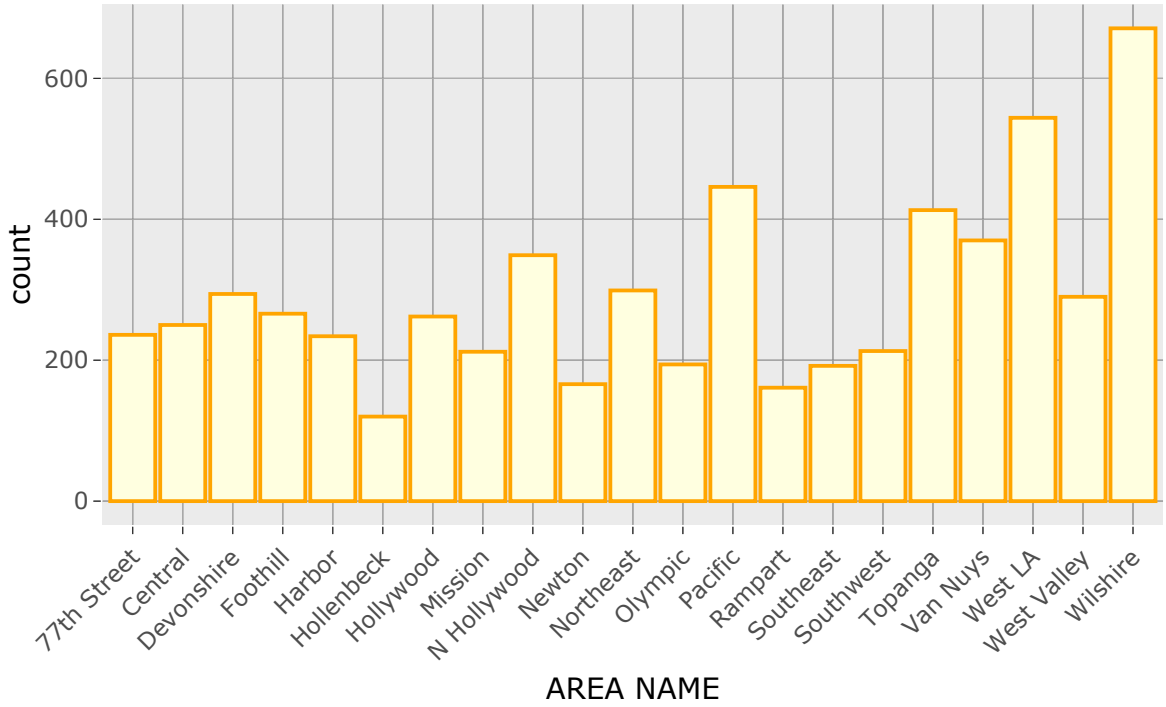
The Boxplot 1 shows that the age range of victims is 7 to 99 years old. The majority of the victims were between 30 to 60 years old, and both the youngest and the oldest victims were Hispanic/Latin. As there were very few AIAN, Pacific Islander and Somoan victims, their age patterns cannot be displayed properly. Overall, there is no significant difference in the median age of victims by different descents.

Barchart 1: Number of victims' by descent and gender



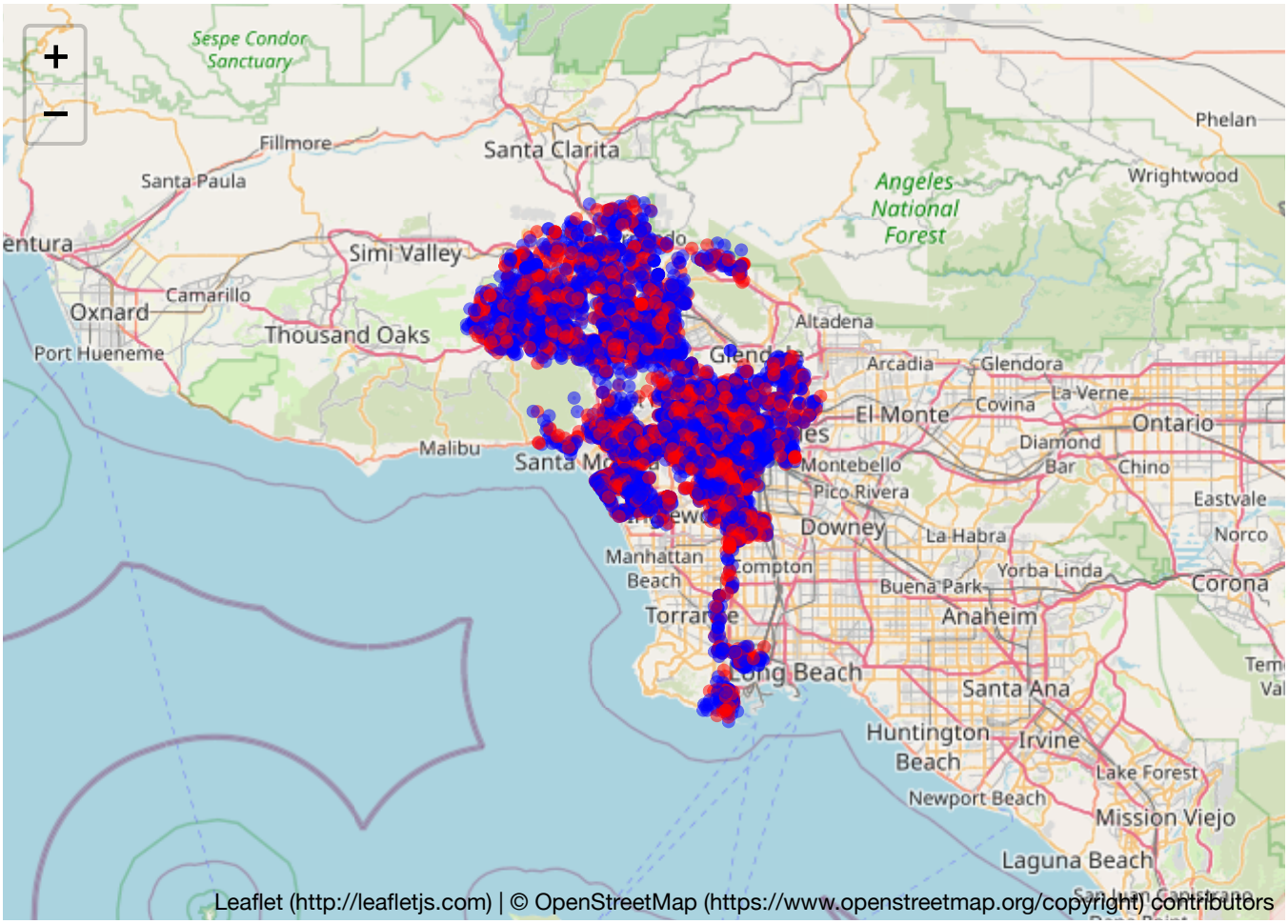
There were 6182 clearly recordered burglaries from January 1st, 2020 to October 5th, 2020 in LA city. It can be seen from the Barchart 1 that most victims were White (2598), followed by Hispanic/Latin (1457) and other (965). It also shows that men are more likely to experience burglary.

Barchart 2: Number of burglaries in different areas



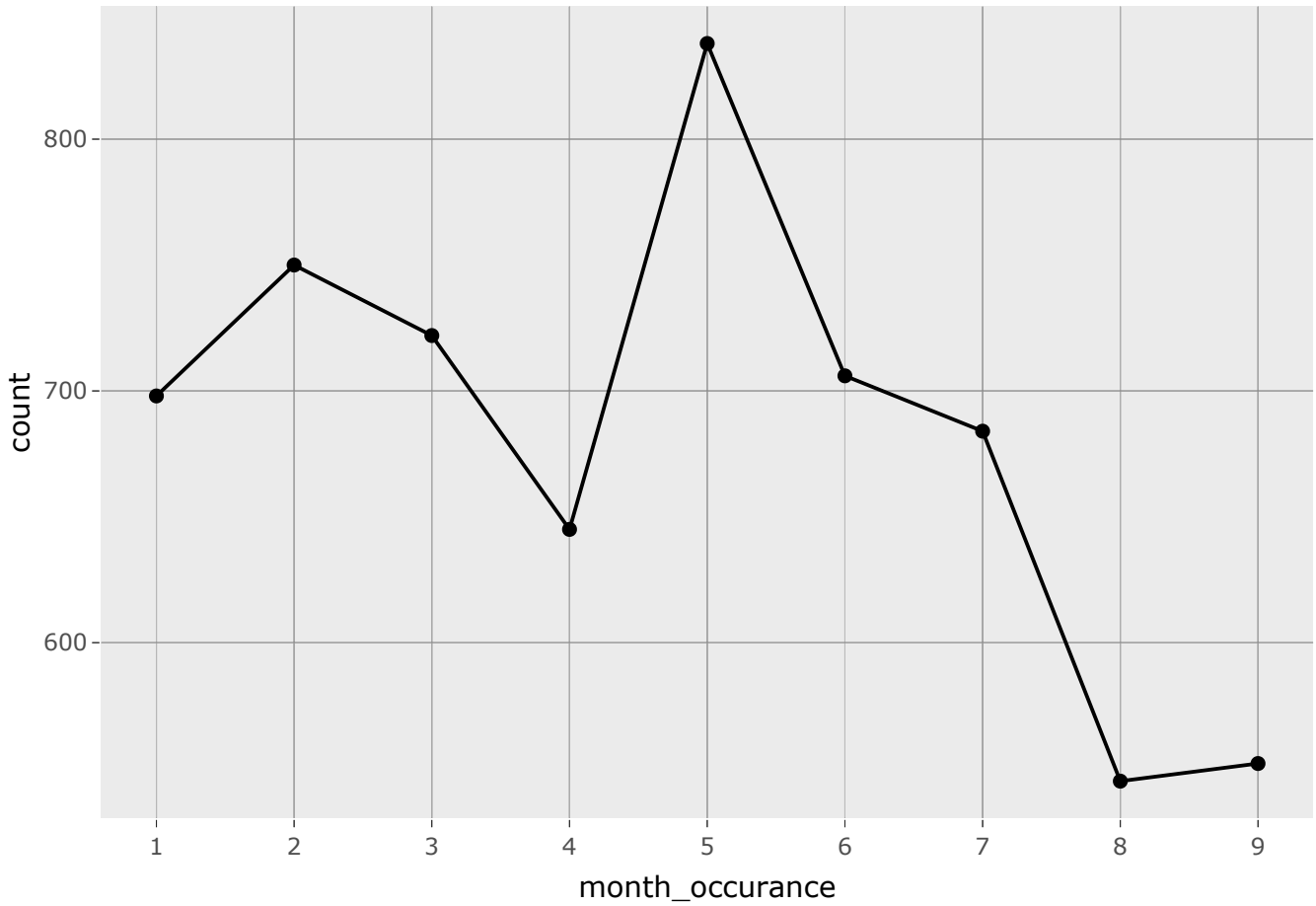
The barchart 2 reveals that Wilshire, West LA and Pacific are the top three places with the most burglaries (671, 544 and 446), which together accounted for over a quarter of the total number of burglaries in Los Angeles. Hollenbeck, Newton and Rampart are relatively safe compared with other areas, which only have burglaries around 150 respectively.

Leaflet 1: Distribtuion of burglaries in LA by gender



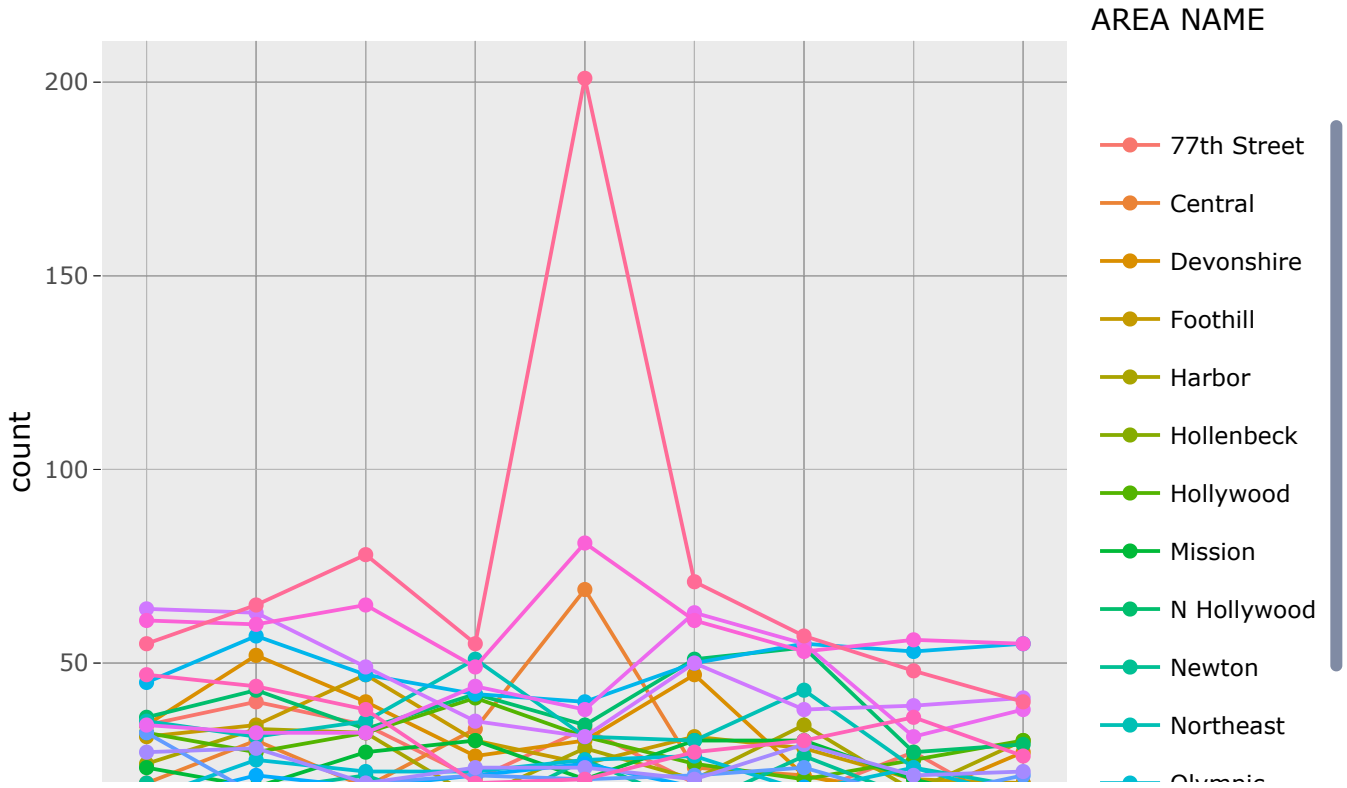
The leaflet map shows that male and female victims were evenly distributed in LA.

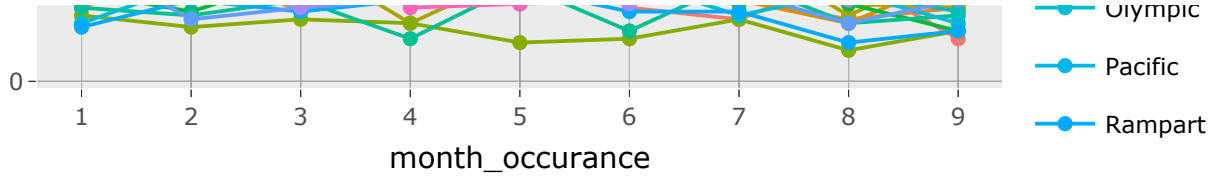
Line graph 1: Number of burglaries in LA by month



From Line graph 1, we found that in the first 4 months in 2020, the number of burglaries per month was about 700. This number reached a peak of 838 in May, then dropped sharply over the following months. In August and September, there were only 545 and 552 burglaries in Los Angeles.

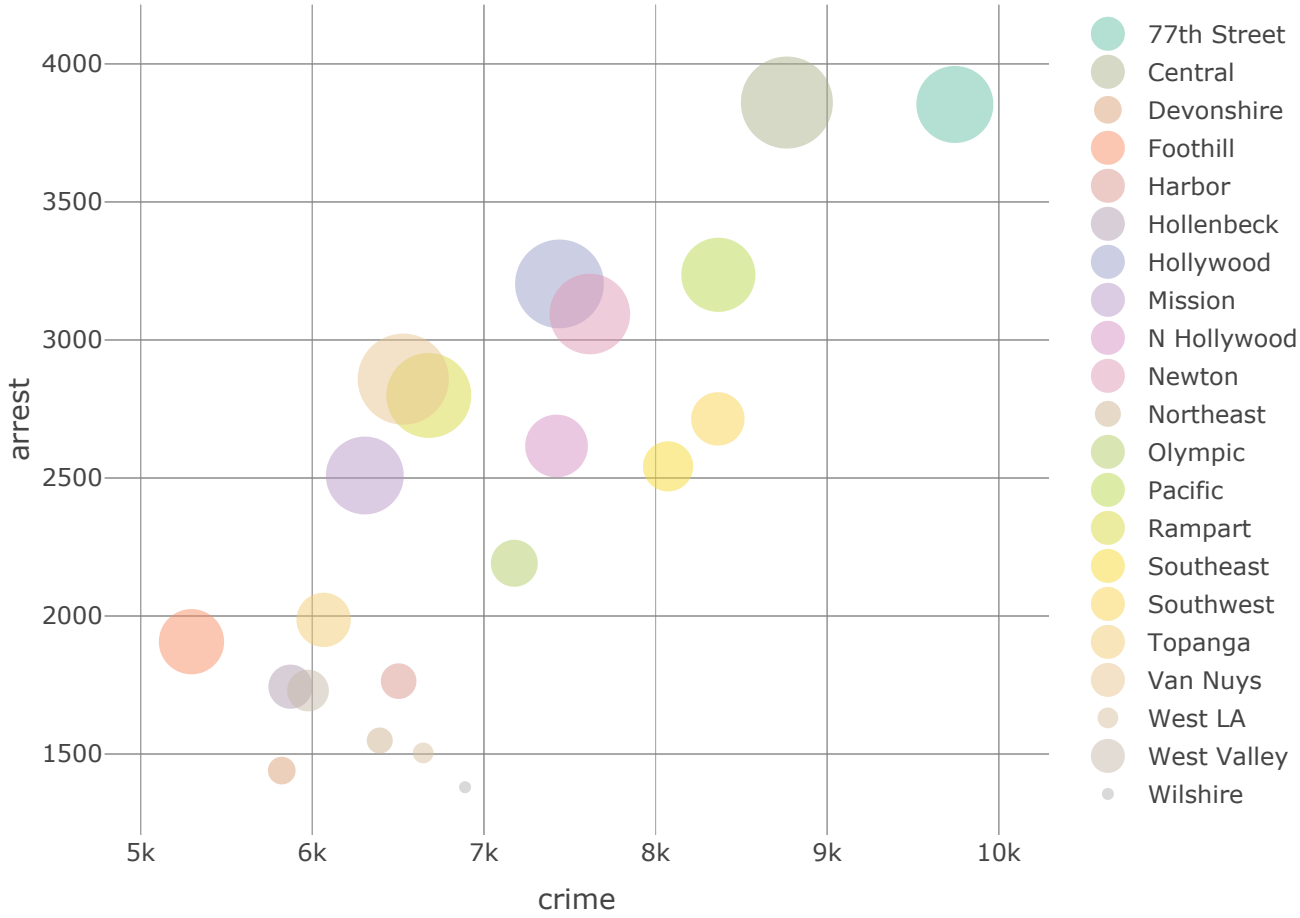
Line plot 2: Number of burglaries in different areas by month





According to Line graph 2, the amount of burglaries were generally constant in each area through these months. However, it's obvious that the number of burglaries has tripled in Wilshire in May.

Scatterplot 1: Pattern of crime and arrest by area



In Scatterplot 1, every bubble represents an area in Los Angeles while the size of the bubble represents the arrest rate of corresponding area. It can be seen that Central area and Van Nuys have the highest arrest rate in 2020, at about 44%. But Wilshire has the lowest arrest rate, at about 20%.

Table 1: Summary of victims' age, sex and races

	AIAN (N = 2)	Asian (N = 393)	Black (N = 765)	Hispanic/Latin (N = 1457)	Other (N = 965)	Pacific Islander (N = 1)	Samoan (N = 1)	White (N = 2598)
Victim Age								
min	45	18	19	7	17	46	65	10
max	59	88	94	99	97	46	65	96

	AIAN (N = 2)	Asian (N = 393)	Black (N = 765)	Hispanic/Latin (N = 1457)	Other (N = 965)	Pacific Islander (N = 1)	Samoan (N = 1)	White (N = 2598)
mean (sd)	52.00 ± 9.90	47.88 ± 15.08	46.31 ± 15.79	43.27 ± 14.78	45.19 ± 14.93	46.00 ± NA	65.00 ± NA	48.91 ± 16.18
Victim Gender								
Male	1 (50)	226 (58)	362 (47)	846 (58)	654 (68)	1 (100)	1 (100)	1,677 (65)
Female	1 (50)	167 (42)	403 (53)	611 (42)	311 (32)	0 (0)	0 (0)	921 (35)

Table 2: Number of burglaries by month and area

Show 10 entries Search:

	rn	1	2	3	4	5	6	7	8	9	10	total
1	1: Central	19	30	18	33	69	23	21	16	19	2	250
2	10: West Valley	47	44	38	19	20	27	30	36	26	3	290
3	11: Northeast	35	31	35	51	31	30	43	23	18	2	299
4	12: 77th Street	34	40	34	20	33	19	16	27	11	2	236
5	13: Newton	19	17	21	11	25	13	26	15	17	2	166
6	14: Pacific	45	57	47	42	40	50	55	53	55	2	446
7	15: N Hollywood	36	43	33	42	34	51	54	27	29	0	349
8	16: Foothill	31	34	47	30	24	31	28	20	19	2	266
9	17: Devonshire	34	52	40	26	30	47	21	15	27	2	294
10	18: Southeast	32	16	19	21	20	21	23	15	21	4	192

Showing 1 to 10 of 21 entries Previous 1 2 3 Next

Table 3: Arrest rate in each area

Show 10 entries Search:

	area	crime	arrest	arrest_rate
1	77th Street	9744	3853	0.4
2	Central	8765	3860	0.44
3	Devonshire	5822	1440	0.25
4	Foothill	5296	1907	0.36

	area	crime	arrest	arrest_rate
5	Harbor	6503	1764	0.27
6	Hollenbeck	5873	1744	0.3
7	Hollywood	7440	3203	0.43
8	Mission	6306	2509	0.4
9	N Hollywood	7423	2616	0.35
10	Newton	7617	3094	0.41

Showing 1 to 10 of 21 entries Previous 1 2 3 Next

Conclusion

We found that there were no significant association between burglary and victims' age. Generally, men are more likely to be burgled than women. White are also more likely to experience burglary compared with others. Most burglaries occurred in May during this year. Wilshire, West LA and the Pacific are the three most dangerous areas of burglary in the city of Los Angeles. In terms of arrest situation, Central area and Van Nuys have the highest arrest rate in 2020, at about 44%. But Wilshire has the lowest arrest rate, at about 20%.