

SAFETY VS. SERVICES

REPORTED CRIME INCIDENTS AND NEIGHBORHOOD VENUES IN SAN FRANCISCO



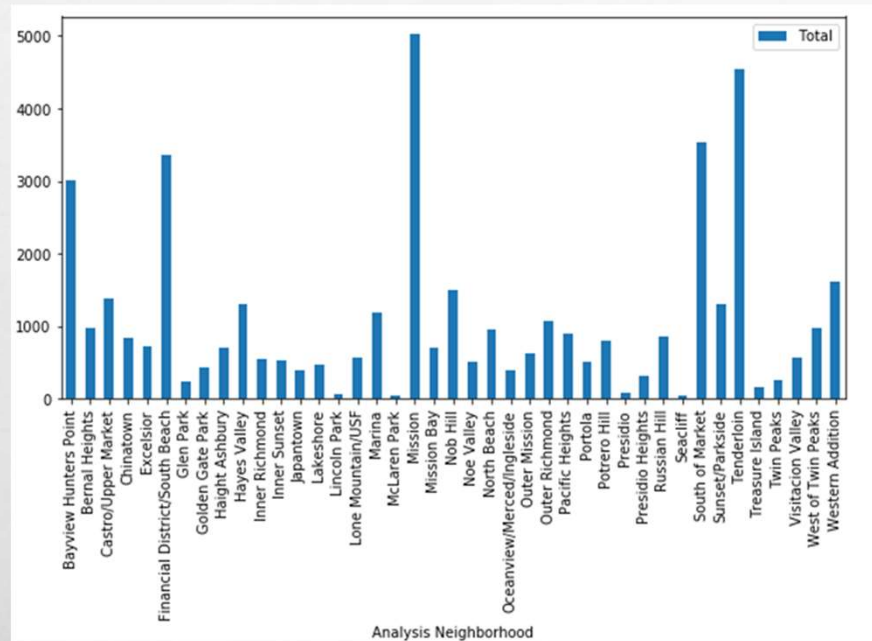
DO SOME VENUES BRING ALONG MORE CRIME?

- **BARS AND NIGHTCLUBS = CRIME. COFFESHOPS = GOOD PEOPLE? STEREOTYPES?**
- **REPUTATION OF A NEIGHBORHOOD CAN BE BASED ON PAST EVENTS, AND NOT REFLECT THE CURRENT REALITIES**
- **EXAMPLE: SOHO IN LONDON, CENTER OF NIGHTLIFE FROM 1960'S UNTIL 1990'S DUE TO ITS NIGHTCLUBS. CURRENTLY VERY GENTRIFIED, BUT CARRIES THE OLD SHADY REPUTATION**
- **DO CERTAIN VENUE TYPES REALLY BRING ALONG MORE CRIME TO THE NEIGHBORHOOD?**
- **NECESSARY INFORMATION FOR PEOPLE RELOCATING, REAL ESTATE INVESTORS, LANDLORDS, EVEN LOCAL POLITICIANS**

DATA ACQUISITION AND CLEANING

- **REPORTED INCIDENTS FROM SAN FRANCISCO POLICE DEPARTMENT FROM 1ST JANUARY 2020 UNTIL MAY 2020: [HTTPS://DATA.SFGOV.ORG/PUBLIC-SAFETY/POLICE-DEPARTMENT-INCIDENT-REPORTS-2018-TO-PRESENT/WG3W-H783](https://data.sfgov.org/public-safety/police-department-incident-reports-2018-to-present/wg3w-h783) LARGE DATABASE, CSV FORMAT. ABOUT 2000 ROWS DROPPED DUE TO MISSING VALUES, SEVERAL COLUMNS DROPPED (MAINLY POLICE DEPRATMENT'S OWN CASE CODES)**
- **44,044 CASES INCLUDED IN THE FINAL DATASET**
- **SAN FRANCISCO VENUE INFORMATION FROM LOADED FROM FOURSQUARE USING THEIR API. AMOUNT OF VENUES CAPPED BY FOURSQUARE AT 100 PER NEIGHBORHOOD**
- **41 SAN FRANCISCO NEIGHBORHOODS IN THE DATASETS**

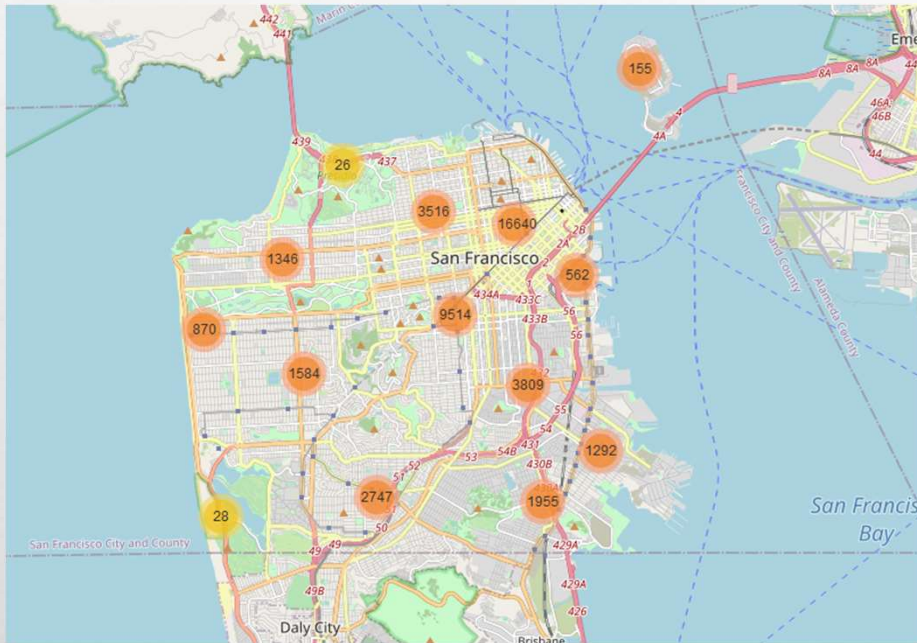
SORTING TO FIND INCIDENTS PER NEIGHBORHOOD



Neighborhood	Incidents
Mission	5018
Tenderloin	4547
South of Market	3539
Financial District/South Beach	3353
Bayview Hunters Point	3015
Western Addition	1606
Nob Hill	1492
Castro/Upper Market	1385
Hayes Valley	1310
Sunset/Parkside	1304

- The top 3 incident neighborhoods cover 29.75% of all reported incidents in San Francisco
- The top 3 incident neighborhoods neighbours with each other

CLUSTERING TO FIND MORE EVIDENCE



Neighborhood	Incidents
Mission	5018
Tenderloin	4547
South of Market	3539
Financial District/South Beach	3353
Bayview Hunters Point	3015
Western Addition	1606
Nob Hill	1492
Castro/Upper Market	1385
Hayes Valley	1310
Sunset/Parkside	1304

- **North Eastern San Francisco has more incidents than other parts of the city**
- **Mission, Tenderloin and South of Market located there**

FOURSQUARE VENUES IN HIGH-INCIDENT NEIGHBORHOODS

----Mission----		
	venue	freq
0	Mexican Restaurant	5.0
1	Café	5.0
2	Art Gallery	4.0
3	Music Venue	4.0
4	Italian Restaurant	2.0
5	Boxing Gym	2.0
6	Ice Cream Shop	2.0
7	Cocktail Bar	2.0
8	Theater	2.0
9	Bakery	2.0
10	Arts & Crafts Store	2.0
11	Bar	2.0
12	Yoga Studio	2.0
13	American Restaurant	2.0
14	Wine Bar	2.0
15	Electronics Store	1.0
16	Performing Arts Venue	1.0
17	Furniture / Home Store	1.0
18	Park	1.0
19	Cheese Shop	1.0

----South of Market----		
	venue	freq
0	Coffee Shop	5.0
1	Vietnamese Restaurant	3.0
2	Bar	3.0
3	Bakery	3.0
4	Pizza Place	2.0
5	Marijuana Dispensary	2.0
6	American Restaurant	2.0
7	Gym / Fitness Center	2.0
8	Dance Studio	2.0
9	Art Gallery	2.0
10	Sports Bar	1.0
11	Café	1.0
12	Accessories Store	1.0
13	Burger Joint	1.0
14	Brewery	1.0
15	Breakfast Spot	1.0
16	New American Restaurant	1.0
17	Hotel	1.0
18	Gym	1.0
19	Performing Arts Venue	1.0

----Tenderloin----		
	venue	freq
0	Coffee Shop	8.0
1	Vietnamese Restaurant	8.0
2	Sandwich Place	6.0
3	Thai Restaurant	5.0
4	Theater	5.0
5	Cocktail Bar	4.0
6	Burger Joint	3.0
7	Bar	3.0
8	Music Venue	3.0
9	Korean Restaurant	3.0
10	Hotel Bar	2.0
11	Speakeasy	2.0
12	Breakfast Spot	2.0
13	Vegetarian / Vegan Restaurant	2.0
14	Mexican Restaurant	2.0
15	Southern / Soul Food Restaurant	2.0
16	Wine Bar	2.0
17	American Restaurant	2.0
18	Spa	1.0
19	Food Truck	1.0

- **The top 3 incident neighborhood venues not out of the ordinary**
- **Restaurants and coffee shops dominate the results**
- **Venue amount in top 3 incident neighborhoods 11.8% of all San Francisco venues**

CONCLUSION

1. **With the information available, it was not possible to come to the conclusion that a certain venue type would be causing an increased amount of reported crime incidents**
2. **Incidents seem to cluster: crime is more regional than neighborhood-specific**

CAVEATS AND SUGGESTIONS

1. Foursquare data is more directed towards social media: the users reporting on Foursquare might not report all places they visit- the dataset might be biased towards more socially acceptable venues
2. Limit of 100 venues per neighborhood probably has an effect on the results, but of limited impact, as Mission is below 100 venues.
3. The venue dataset should be tried with the new dataset of registered business from San Francisco city: (<https://data.sfgov.org/Economy-and-Community/Registered-Business-Locations-San-Francisco/g8m3-pdis>). This is likely to be less biased than Foursquare
4. The analysis should include socioeconomic data as per neighborhood. A good starting point would be (https://default.sfplanning.org/publications_reports/SF_NGBD_SocioEconomic_Profiles/2010-2014_ACS_Profile_Neighborhoods_v3AH.pdf)

THANK YOU FOR YOUR ATTENTION!

Jani Peurakoski / Coursera Capstone Project