Determining a neighborhood to open a new Pizza place in San Francisco

Brett Tomlinson

Introduction

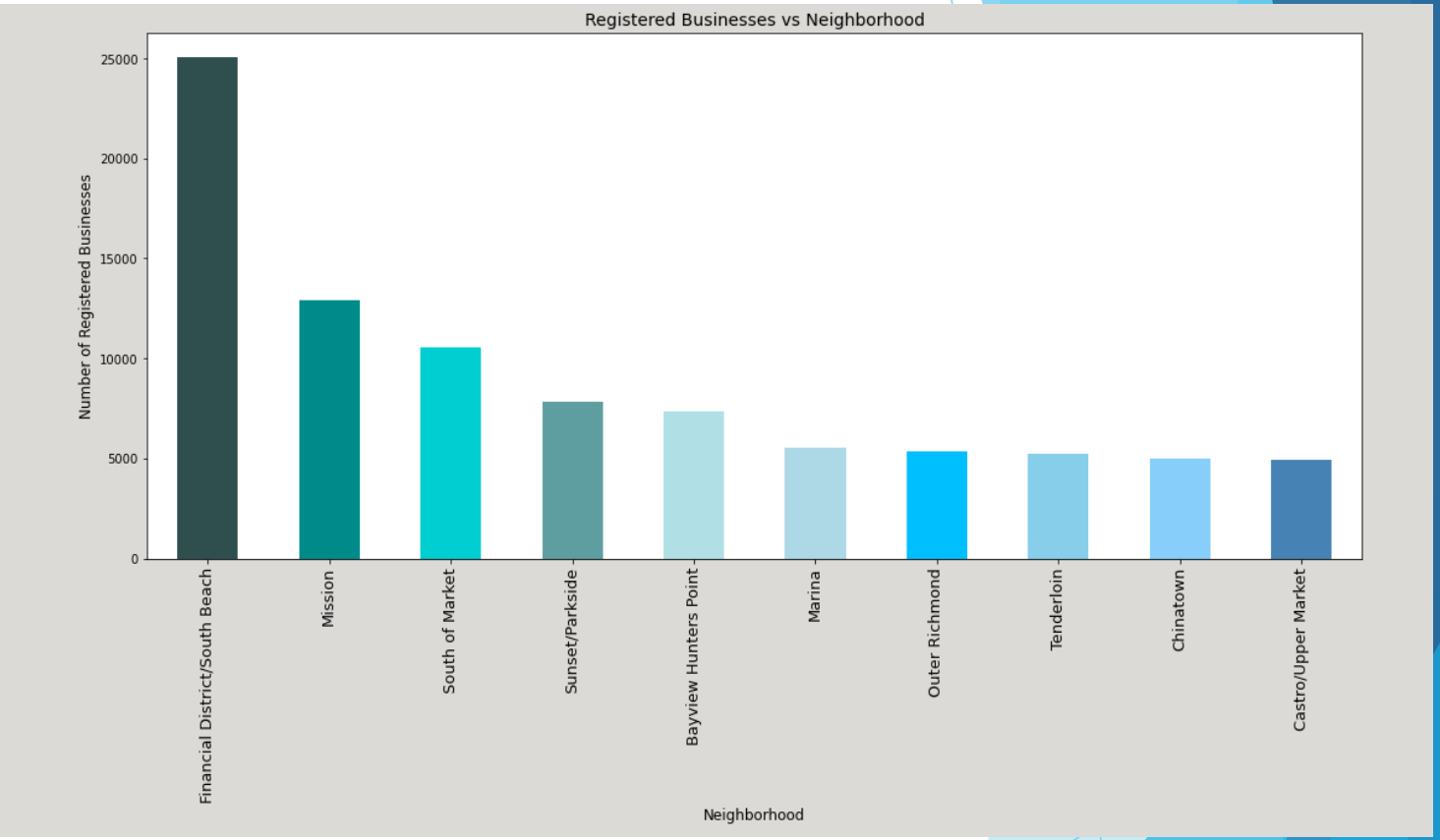
- Where to site a business can have a sizeable impact on the success or failure on the business after opening
- Odd Moe's Pizza has been regionally successful in Oregon and is looking to expand into California, more specifically San Francisco.
- To keep in line with their business strategy in Oregon, Odd Moe's will look for a neighborhood with a lower crime rate to open the new site.
- The site will need to be in a safe neighborhood, but also with a ready supply of patrons ready for a new pizza place.

Project Data

- The city of San Francisco has multiple datasets useful for this project which are available to the public
- Datasets for Registered Business and Crimes were used:
  - Businesses: https://data.sfgov.org/Economy-and-Community/Registered-Business-Locations-San-Francisco/g8m3-pdis/data?no_mobile=true
  - Crime: https://data.sfgov.org/Public-Safety/Police-Department-Incident-Reports-2018-to-Present/wg3w-h783
- Foursquare API will be used to get venue and location information

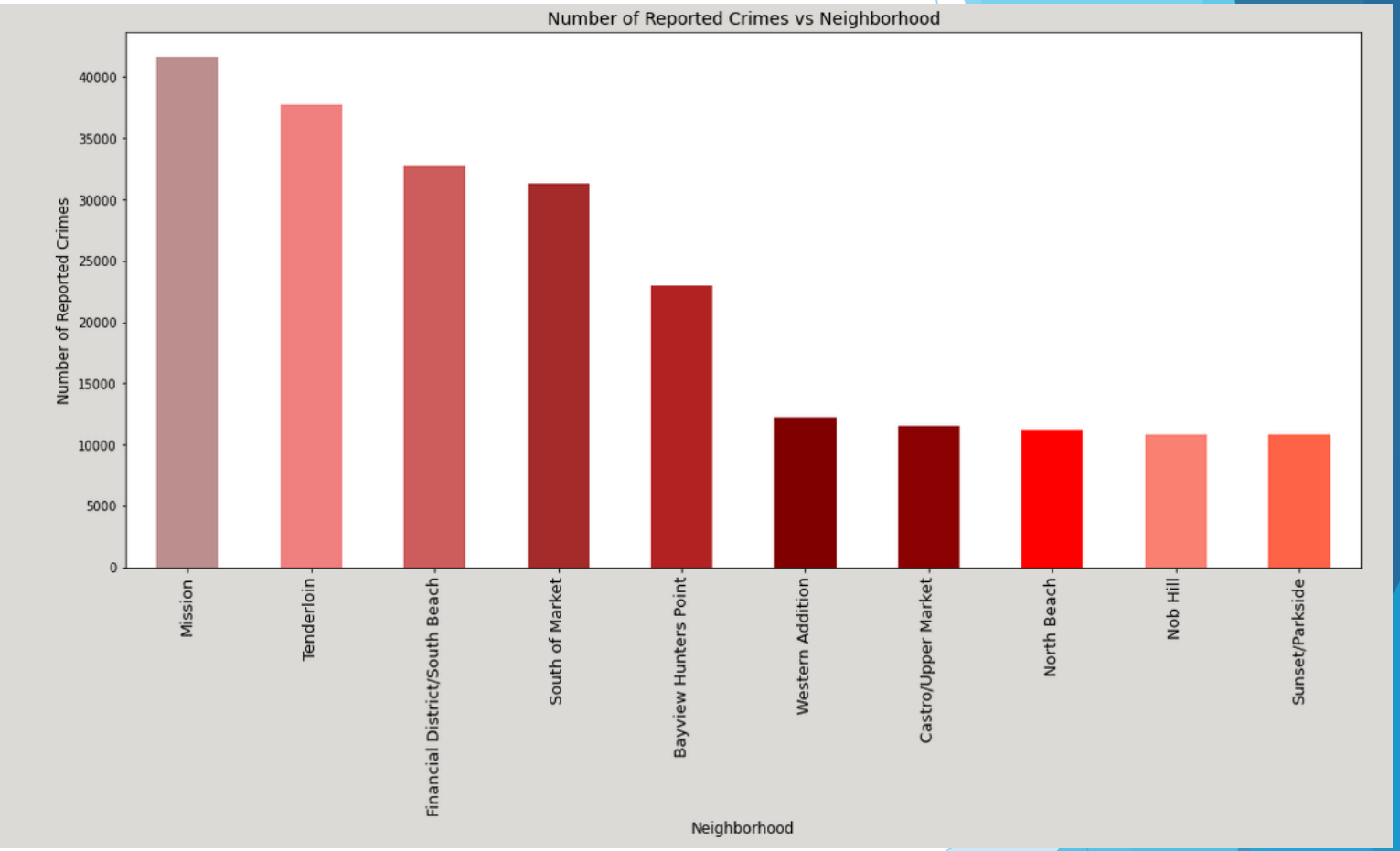
Registered Business

- Data was pulled of registered business in San Francisco for the past 15 years. This was done to enough relevant data on Pizza Places
- Financial District neighborhood has the largest share of businesses in San Francisco by a wide margin



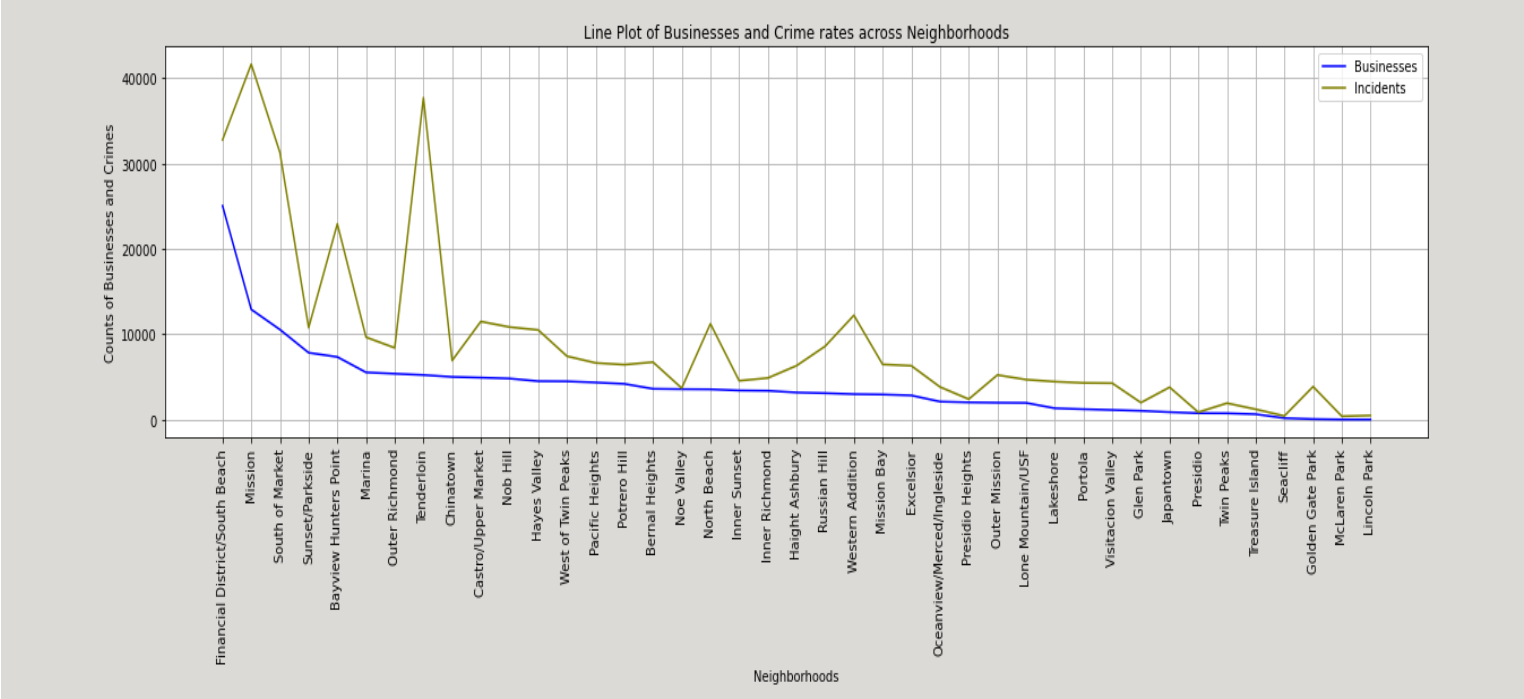
Crime

- Data was pulled for Crimes in the past 5 years to ensure that we have the most relevant data for the scenario
- Plotting number of crime incidents across neighborhoods shows that Mission neighborhood has the highest crime rates
- There is a clear divide between the top and bottom 5 neighborhoods with regard to crime rates



Business-Crime Combination

- Top crime and business data is combined to see how far apart crime and business counts are for each neighborhood
- Mission neighborhood has the worst combination of business and crime. The rate of businesses in Mission is behind Financial District, however the crime rate is much higher in Mission than Financial District



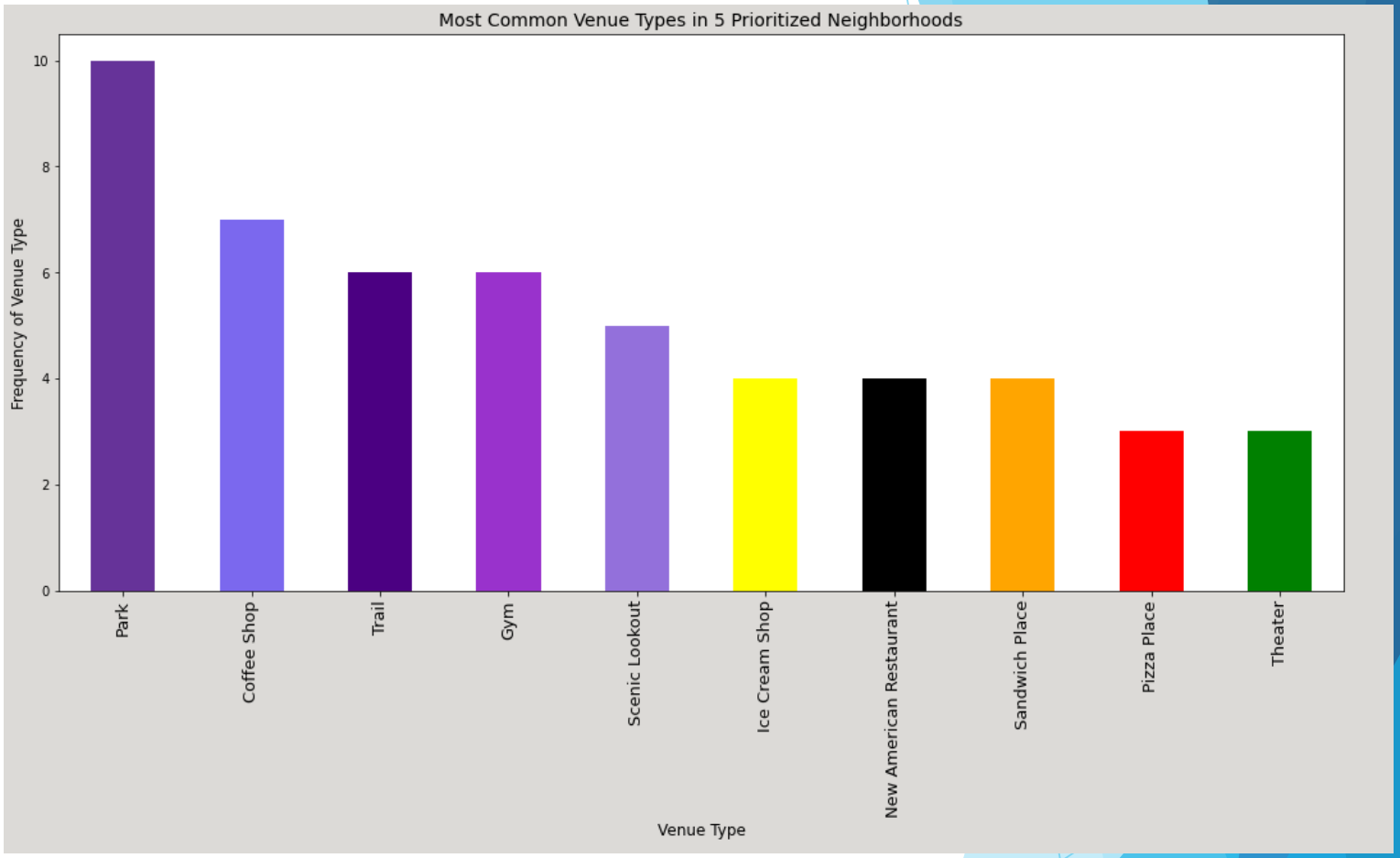
Geolocation Data

- Coordinates for each of the selected neighborhoods were gathered using https://latlong.net which provides the longitude and latitude.
- Businesses, Crime Rates, and coordinate data was then merged

	Neighborhood	Businesses	Crimes	Coordinates	Latitude	Longitude
0	Sunset/Parkside	7839	10767	(37.749690, -122.494880)	37.749690	-122.494880
1	Marina	5538	9657	(39.820580, -84.887010)	39.820580	-84.887010
2	Outer Richmond	5383	8410	(37.803850, -122.450350)	37.803850	-122.450350
3	Chinatown	5011	6925	(37.765730, -122.435470)	37.765730	-122.435470
4	Castro/Upper Market	4924	11495	(37.758430, -122.448440)	37.758430	-122.448440

Foursquare API

- Using a radius of 1 mile from the prioritized neighborhood, venues with the most frequency were found using Foursquare
- The most common venues are Parks, Coffee Shops, Trails, and Gyms



Common Venues

- The 2 neighborhoods that have Pizza Places in the top spots of common venues are Marina and Sunset/Parkside neighborhoods.
- Marina has Pizza Places among the top venue frequency while Sunset/Parkside has Pizza Places tied in the second spot
- All other prioritized neighborhoods do not have Pizza Places in the most frequent venues

----Marina----

	venue	freq
0	Sandwich Place	0.07
1	Theater	0.07
2	Bagel Shop	0.07
3	American Restaurant	0.07
4	Gym	0.07
5	Video Store	0.07
6	Pizza Place	0.07
7	Cosmetics Shop	0.03
8	Restaurant	0.03
9	Discount Store	0.03
10	Pharmacy	0.03
11	ATM	0.03
12	Fast Food Restaurant	0.03
13	Rock Club	0.03
14	Gas Station	0.03
15	History Museum	0.03
16	Ice Cream Shop	0.03
17	Food	0.03
18	Pub	0.03
19	Coffee Shop	0.03

----Sunset/Parkside----

	venue	freq
0	Coffee Shop	0.10
1	Chinese Restaurant	0.07
2	Playground	0.07
3	Trail	0.07
4	Ice Cream Shop	0.07
5	Pizza Place	0.03
6	Grocery Store	0.03
7	Record Shop	0.03
8	Empanada Restaurant	0.03
9	Dumpling Restaurant	0.03
10	Donut Shop	0.03
11	Liquor Store	0.03
12	Dive Bar	0.03
13	Sandwich Place	0.03
14	Dim Sum Restaurant	0.03
15	Cocktail Bar	0.03
16	Korean Restaurant	0.03
17	Beer Bar	0.03
18	Vietnamese Restaurant	0.03
19	Bakery	0.03

K-means Clustering

- Marina and Sunset/Parkside neighborhoods have some commonality with Pizza Places being popular venues.
- Pizza Places are tied as the most frequent venue in the Marina neighborhood which suggest that there is a large market of pizza lovers in the neighborhood.
- Using K-means clustering we can see if the clustering algorithm agrees with the initial theory we gained from analysis.
- Cluster labels from K-means gets mapped to the neighborhoods and it was found that Marina and Sunset/Parkside neighborhoods were placed in the same cluster.

Conclusion

- In this project, common python libraries and Foursquare API were used to manipulate data and explore the neighborhoods of San Francisco.
- Using K-means clustering, we verified the initial theory that supported our observation.
- In the future we would attempt a time series analysis of businesses and crimes in neighborhoods so we can observe the trend and guard against basing our judgement on trends overall history which could have changed.