



NEW BUSINESS LOCATION ANALYSIS IN SAN FRANCISCO

Joye
December 25, 2019

Contents

 Introduction

 Data

 Methodology

 Results

 Discussion

 Conclusion

Background

Members of Generation Z have gradually become the main force of the consumption market. However, to win the Generation Z consumers is not an easy task. According to consumption characteristics, different companies should improve their products, services, marketing etc. While among all of those, location is a top priority, because a bad location will always discount consumption experiences and satisfaction. This is especially vital for a new opening business.

Audience

My target audience are those companies who want to win more market shares and have the plan to open several new businesses in cities. Since in this capstone project, we will choose San Francisco as our target city, so my audience are stakeholders who plan to open new businesses in San Francisco. To help them make a wise decision on location, I will take advantage of the Foursquare location data and provide valuable suggestions to them.

Data

To successfully finish this capstone project, we need the following kinds of data:

- The neighborhoods in San Francisco, and their latitude and longitude coordinates.
- The venue data of those neighborhoods.

Data source:

https://en.wikipedia.org/wiki/List_of_neighborhoods_in_San_Francisco

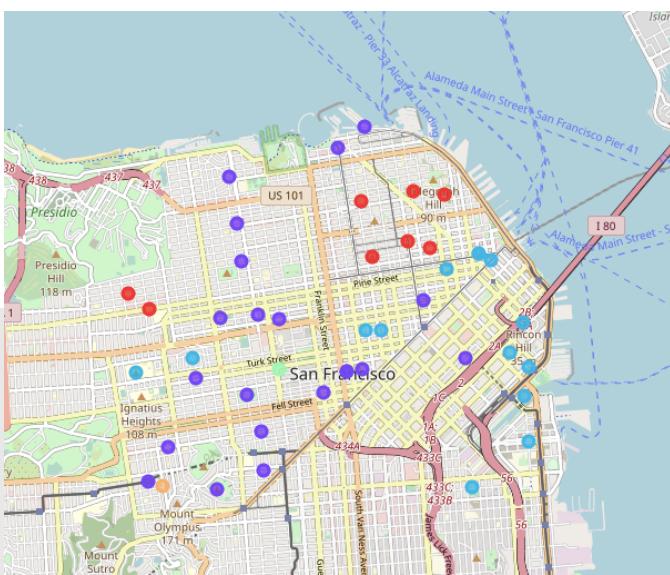
Methodology

- The first step is to check the size of resulting data. Next, we need to check the venue categories.
- Secondly, we convert the categorical data into numerical data, namely convert the venue categories into dummies. Then calculate the frequency of each category in all venues we got.
- Next, we use the unsupervised machine learning k-means clustering to cluster the neighborhoods.
- After clustering, we can visualize the neighborhoods on the map of San Francisco.
- In the final step, we will carefully review each cluster and its attributes, and use word cloud to visualize the most common business in clusters.

Results

----Alamo Square----

	venue	freq
0	Bar	0.05
1	Pizza Place	0.03
2	Sushi Restaurant	0.03
3	Hotel	0.03
4	Record Shop	0.03



Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0 Alamo Square	Bar	Hotel	Liquor Store	Seafood Restaurant	Café	Sushi Restaurant	Record Shop	Pizza Place	Ethiopian Restaurant	Indian Restaurant
1 Anza Vista	Café	Health & Beauty Service	Bus Stop	Bus Line	Coffee Shop	Tunnel	Grocery Store	Arts & Crafts Store	Big Box Store	Donut Shop
2 Belden Place	Coffee Shop	Gym	Cocktail Bar	Gym / Fitness Center	Café	Sushi Restaurant	Bubble Tea Shop	Food Truck	Men's Store	Boutique
3 Buena Vista	Park	Seafood Restaurant	Historic Site	Ice Cream Shop	Harbor / Marina	Chocolate Shop	Diner	Gift Shop	Clothing Store	Boat or Ferry
4 China Basin	Baseball Stadium	Coffee Shop	New American Restaurant	Gym / Fitness Center	Baseball Field	Bar	Outdoor Sculpture	Pier	Pizza Place	Athletics & Sports

Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0 Alamo Square	37.776360	-122.434689	1	Bar	Hotel	Liquor Store	Seafood Restaurant	Café	Sushi Restaurant	Record Shop	Pizza Place	Ethiopian Restaurant	Indian Restaurant
1 Anza Vista	37.780836	-122.443149	2	Café	Health & Beauty Service	Bus Stop	Bus Line	Coffee Shop	Tunnel	Grocery Store	Arts & Crafts Store	Big Box Store	Donut Shop
4 Belden Place	37.791744	-122.403886	2	Coffee Shop	Gym	Cocktail Bar	Gym / Fitness Center	Café	Sushi Restaurant	Bubble Tea Shop	Food Truck	Men's Store	Boutique
6 Buena Vista	37.806532	-122.420648	1	Park	Seafood Restaurant	Historic Site	Ice Cream Shop	Harbor / Marina	Chocolate Shop	Diner	Gift Shop	Clothing Store	Boat or Ferry
8 China Basin	37.776330	-122.391839	2	Baseball Stadium	Coffee Shop	New American Restaurant	Gym / Fitness Center	Baseball Field	Bar	Outdoor Sculpture	Pier	Pizza Place	Athletics & Sports

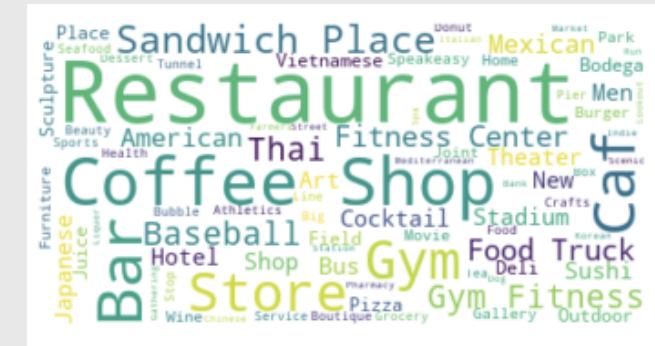
Wordclouds



Cluster 0



Cluster 1



Cluster 2

Discussion

The new business location analysis in this capstone is mainly from the view of venues, we assume that industry agglomeration plays an important role. While, in reality the situation is much more complex, and more factors can affect the operation of new business. In future, we should include more data, like demographic statistics and financial data.

Conclusion

Since our neighborhoods are near the downtown of San Francisco , it's not surprised that restaurants and coffee shops are the most common venue. Therefore, it's fine to choose cluster 0, 1 and 2 area to start catering business. From the analysis, we notice that cluster 0, 1 and 2 have their own characteristics. Clusters 0 is more to office zone, clusters 1 is more to shopping, entertainment and living, and cluster 2 is more to gym and fitness.

In sum, cluster 0 zone welcomes special restaurant, café and bar business. Cluster 1 zone welcomes catering business and living related shops. Cluster 2 welcomes catering services, healthy industry and sports shops.



Thank You

Merry Christmas