Finding the Perfect Location

November 19, 2019

Overview

For many budding restaurateurs, one of the most exciting and frightening moments is expanding to new locations. On the surface it might seem like a very natural and organic progression of replicating the initially successful recipe. However, as any established restaurateurs will tell you, there are many more factors. Of all these variables, location often surfaces to the top. People often say, it's all about location location location!
So how does one go about finding that perfect location that would replicate the success of the original restaurant?
Instead of basing on word of mouth, feels, emotion and other eyeballing approaches, the goal is to bring data science to a restaurateur who is looking to expand into other cities or neighborhoods. By using data science the proprietor will be presented with an analysis of the different neighborhoods in a defined area. The analysis will give a non emotional, objective picture of the area and pave the way toward choosing the right location. Furthermore, the analysis can validate or solidify a potential choice made from emotional or non scientific methodologies. Lastly, the analysis gives the restaurateur the ability to quantify why an "Unicorn" location is perfect.

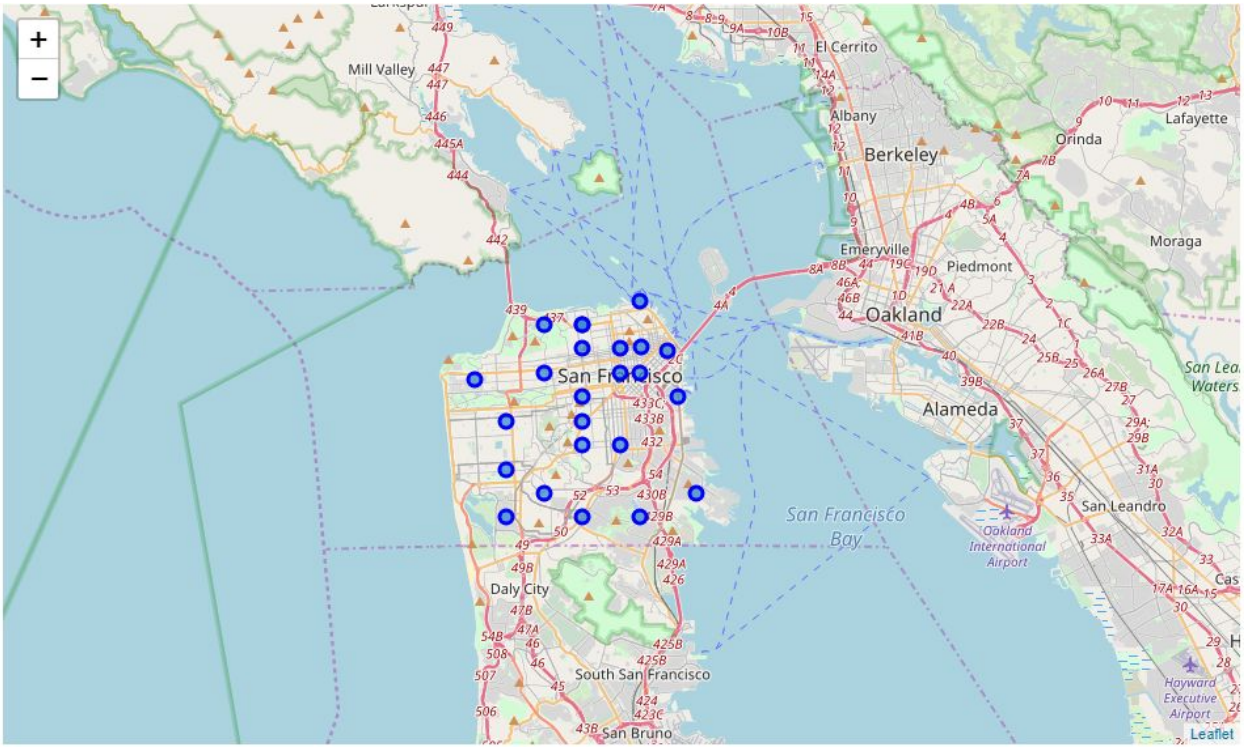
Data

The exercise was conducted on San Francisco with data from these sources.
http://www.healthysf.org/bdi/outcomes/zipmap.htm
http://zipatlas.com/us/ca/san-francisco/zip-code-comparison/population-density.htm
https://deansereni.com/newsletter/san-francisco-demographics-by-zip-code.

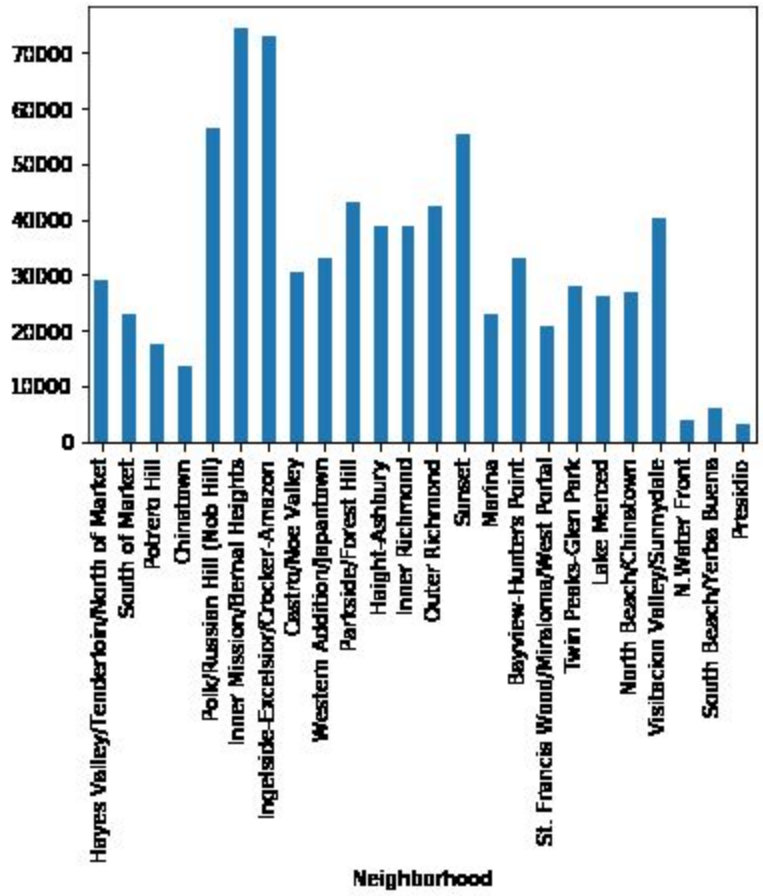
:

	Zip Code	Neighborhood	Population (Census 2000)
1	94102	Hayes Valley/Tenderloin/North of Market	28991
2	94103	South of Market	23016
3	94107	Potrero Hill	17368
4	94108	Chinatown	13716
5	94109	Polk/Russian Hill (Nob Hill)	56322
6	94110	Inner Mission/Bernal Heights	74633
7	94112	Ingelside-Excelsior/Crocker-Amazon	73104
8	94114	Castro/Noe Valley	30574
9	94115	Western Addition/Japantown	33115
10	94116	Parkside/Forest Hill	42958
11	94117	Haight-Ashbury	38738
12	94118	Inner Richmond	38939
13	94121	Outer Richmond	42473
14	94122	Sunset	55492
15	94123	Marina	22903
16	94124	Bayview-Hunters Point	33170
17	94127	St. Francis Wood/Miraloma/West Portal	20624
18	94131	Twin Peaks-Glen Park	27897
19	94132	Lake Merced	26291
20	94133	North Beach/Chinatown	26827
21	94134	Visitacion Valley/Sunnydale	40134
22	94111	N.Water Front	3700
23	94105	South Beach/Yerba Buena	6000
24	94129	Presidio	3200

To make sure the geographical coordinates were correct and the neighborhoods were reasonable map of San Francisco was created with the data points.



Population for each Neighborhood



Exploration

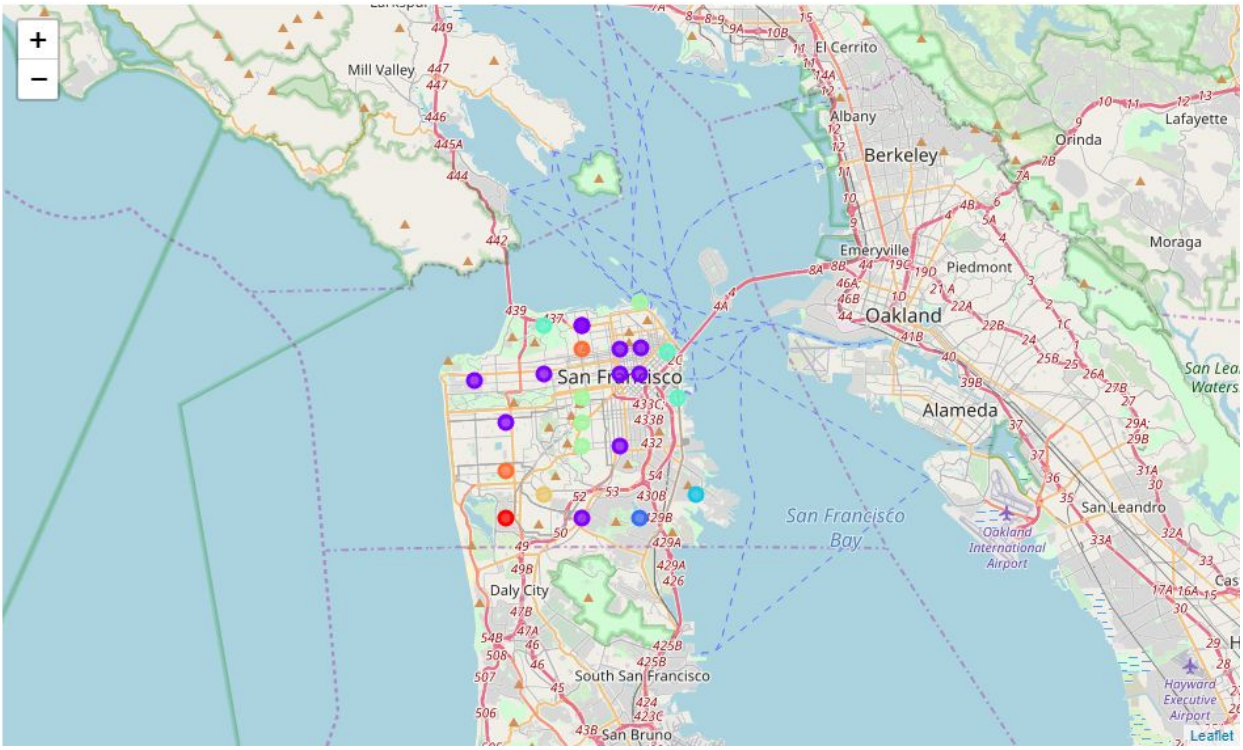
Foursquare Venue data

:

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Bayview-Hunters Point	4	4	4	4	4	4
Castro/Noe Valley	62	62	62	62	62	62
Chinatown	100	100	100	100	100	100
Haight-Ashbury	30	30	30	30	30	30
Hayes Valley/Tenderloin/North of Market	100	100	100	100	100	100
Ingelside-Excelsior/Crocker-Amazon	37	37	37	37	37	37
Inner Mission/Bernal Heights	64	64	64	64	64	64
Inner Richmond	65	65	65	65	65	65
Lake Merced	25	25	25	25	25	25
Marina	90	90	90	90	90	90
N.Water Front	52	52	52	52	52	52
North Beach/Chinatown	90	90	90	90	90	90
Outer Richmond	49	49	49	49	49	49
Parkside/Forest Hill	51	51	51	51	51	51
Polk/Russian Hill (Nob Hill)	100	100	100	100	100	100
Potrero Hill	75	75	75	75	75	75
Presidio	36	36	36	36	36	36
South Beach/Yerba Buena	100	100	100	100	100	100
South of Market	100	100	100	100	100	100
St. Francis Wood/Miraloma/West Portal	6	6	6	6	6	6
Sunset	60	60	60	60	60	60
Twin Peaks-Glen Park	16	16	16	16	16	16
Visitacion Valley/Sunnydale	5	5	5	5	5	5
Western Addition/Japantown	35	35	35	35	35	35

Given the relatively small city size, we found the dataset to be much smaller for San Francisco. Several neighborhoods had extremely few venues.

The Key was to identify similar neighborhoods and K means clustering with K clusters of 8 was conducted.



It was very apparent as to which cluster contained heavily commercialized neighborhoods Cluster 2 with top 10 most common venue.

	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
1	Hayes Valley/Tenderloin/North of Market	Coffee Shop	Cocktail Bar	Café	French Restaurant	Furniture / Home Store	Wine Bar	Vietnamese Restaurant	Beer Bar	Boutique	Thai Restaurant
2	South of Market	Coffee Shop	Music Venue	Theater	Sandwich Place	Café	Art Gallery	American Restaurant	Vietnamese Restaurant	Wine Bar	Italian Restaurant
4	Chinatown	Hotel	Coffee Shop	American Restaurant	Men's Store	Spa	Boutique	Cocktail Bar	Bar	Salon / Barbershop	Bubble Tea Shop
5	Polk/Russian Hill (Nob Hill)	Grocery Store	Sushi Restaurant	Bar	Bakery	Breakfast Spot	Vietnamese Restaurant	Massage Studio	Diner	Sandwich Place	Coffee Shop
6	Inner Mission/Bernal Heights	Mexican Restaurant	Dive Bar	Coffee Shop	Italian Restaurant	Grocery Store	New American Restaurant	Massage Studio	Cocktail Bar	Bakery	Gym / Fitness Center
7	Ingelside-Excelsior/Crocker-Amazon	Pizza Place	Mexican Restaurant	Pharmacy	Vietnamese Restaurant	Coffee Shop	Bar	Sandwich Place	Fried Chicken Joint	Pool Hall	Restaurant
12	Inner Richmond	Japanese Restaurant	Sushi Restaurant	Pizza Place	Bakery	Wine Shop	Burger Joint	Vietnamese Restaurant	Italian Restaurant	Korean Restaurant	Burmese Restaurant
13	Outer Richmond	Café	Chinese Restaurant	Bakery	Sporting Goods Shop	Japanese Restaurant	Vietnamese Restaurant	Dessert Shop	Pizza Place	Intersection	Seafood Restaurant
14	Sunset	Bubble Tea Shop	Bakery	Vietnamese Restaurant	Chinese Restaurant	Dim Sum Restaurant	Deli / Bodega	Cosmetics Shop	Japanese Restaurant	Dumpling Restaurant	Grocery Store
15	Marina	Italian Restaurant	French Restaurant	American Restaurant	Gym / Fitness Center	Thai Restaurant	Salad Place	Wine Bar	Mexican Restaurant	Sushi Restaurant	Taco Place
20	North Beach/Chinatown	Italian Restaurant	French Restaurant	American Restaurant	Gym / Fitness Center	Thai Restaurant	Salad Place	Wine Bar	Mexican Restaurant	Sushi Restaurant	Taco Place

Conclusion

The study allows a restauranteur to quantify the desirability of a location. Furthermore, by clustering, one can quickly find related neighborhoods that share similar characteristics as the original location. The study is not meant to be the end of the process but the beginning. It helps to chanel and narrow the scope for the proprietor in hope of going down the right direction. It allows the owner to establish the characteristics of the current location as the baseline and quickly find similar potential locations. Lastly, the study can be conducted in different cities and the data easily compared. The ability to describe what “perfect” location means will be an awesome tool for a successful restaurateur who’s about to build an empire