Popular Venues in San Francisco

Mandy Wong

1 INTRODUCTION

San Francisco is one of the popular cities all over the United States and the 4th largest city in California. According to the world population overview [1], San Francisco is a dense city with current population of 896,047 in 2020. Since San Francisco is a diverse city, there are variety of different race of people living in this prosperous city. It would be a great investment if there is a good business plan.

But, what shop or store should the businessman open in this area? This project plans to provide the rough idea or research for people who would like to do business in San Francisco. This project could also help the businessman or company to understand which type of business is too less or too much in certain zip code area, so that the better business planning could be developed.

2 DATA

All zip code in San Francisco are found from sanfranguide [2] website. In the zip code data, there are only one neighborhood corresponding to each zip code, it may not be enough for analysis. So, the data from healthsf [3] website was also used to diversify the neighborhood of each zip code. All the San Francisco zip code is also used to find the latitude and longitude by integrated the data from Opendatasoft [4] website.

The data described above will be used to find the top 10 most common venues in each neighborhood by using Foursquare API. Since there are different area (e.g. residential, commercial, tourist attraction) in San Francisco, the most popular business type in each area is different either. The data will be clustering into 5 clusters. These 5 clusters will be used to analyze the most popular type of business in certain area. This could help the investor or businessman to realize current popular venues around their intend-to-invest neighborhoods of San Francisco.

3 METHODOLOGY

3.1 Data preprocessing

San Francisco zip code data are originally from sanfranguide [2] website. The data from healthsf [3] website is used as supplemental neighborhood to each zip code. It makes the data more diverse for the information of each neighborhood. The zip code data are saved in JSON format and used to create the data frame by using python pandas library in jupyter notebook (Fig. 1). After the zip code dataframe is created, the data is integrated with the US zip code Latitude and Longitude from Opendatasoft [4] website (Fig. 2).

Neighborhood	zipcode	
Hayes Valley, Tenderloin, North of Market	94102	0
South of Market	94103	1
Financial District	94104	2
Embarcadero South	94105	3
Portrero Hill	94107	4
Chinatown	94108	5

Fig.1. SF zip code data frame

	zipcode	Neighborhood	Latitude	Longitude
0	94102	Hayes Valley, Tenderloin, North of Market	37.779329	-122.41915
1	94103	South of Market	37.772329	-122.41087
2	94104	Financial District	37.791728	-122.40190
3	94105	Embarcadero South	37.789228	-122.39570
4	94107	Portrero Hill	37.766529	-122.39577
5	94108	Chinatown	37.792678	-122.40793

Fig.2. SF zipcode data with lat. and long.

3.2 Find Top 10 most common venues and Clustering

Before finding out the top 10 most common venues in each neighborhood of San Francisco, exploring the nearby venues in each neighborhood is needed. All the nearby venues in each neighborhood are obtained by using Foursquare API. For each neighborhood, there is 100 nearby venues within radius of 500 meters. The next step is to analyze each neighborhood with 251 unique venues categories. This create the picture of the frequent venues in each neighborhood. Then, the above dataset is used to find the top 10 most common venues (Fig. 3, Fig. 4).

-	Bayview		Chinatown	
	venue	freq	venue	freq
0	Southern / Soul Food Restaurant	0.15	0 Hotel	0.10
1	Mexican Restaurant	0.12	1 Coffee Shop	0.06
2	Light Rail Station	0.08	2 Bakery	0.05
3	Bakery	0.08	3 Men's Store	0.04
4	Theater	0.04	4 Bubble Tea Shop	0.04
5	BBQ Joint	0.04	5 Chinese Restaurant	0.04
6	Bus Station	0.04	6 Tea Room	0.04
7	Garden		7 Italian Restaurant	0.04
8	Chinese Restaurant	0.04	8 Sushi Restaurant	0.02
9	Gym	0.04	9 Szechuan Restaurant	0.02

Fig.3 Top 10 common venues in Bayview

Fig.4. Top 10 common venues in Chinatown

The data of 10 most common veunes is also saved into a pandas dataframe (Fig. 5)

	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	Bayview	Southern / Soul Food Restaurant	Mexican Restaurant	Light Rail Station	Bakery	Theater	Park	Plaza	Dumpling Restaurant	Coffee Shop	Chinese Restaurant
1	Castro, Noe Valley	Gay Bar	Thai Restaurant	Coffee Shop	Indian Restaurant	Scenic Lookout	Yoga Studio	Pet Store	Deli / Bodega	Café	Convenience Store
2	Chinatown	Hotel	Coffee Shop	Bakery	Tea Room	Chinese Restaurant	Men's Store	Italian Restaurant	Bubble Tea Shop	Sushi Restaurant	Dim Sum Restaurant
3	Cow Hollow, Marina District	French Restaurant	Italian Restaurant	Wine Bar	Gym / Fitness Center	Sandwich Place	Sushi Restaurant	Deli / Bodega	Park	Cosmetics Shop	Coffee Shop
4	Embarcadero North	Food Truck	Scenic Lookout	Café	Coffee Shop	Men's Store	Italian Restaurant	Wine Bar	New American Restaurant	Park	Cosmetics Shop

Fig.5. Top 10 venues in each neighborhood

Lastly, all San Francisco neighborhoods are cluster into 5 clusters based on the data of top 10 common venues in each neighborhood and its location. The clustering results are visualized by using folium map. This clearly shows the similar neighborhoods so that it can help roughly dividing the whole area into different areas (eg. commercial, residential and tourist attract)

4 RESULTS

Based on the zip code and its neighborhoods data from sanfranguide [2] and healthsf [3] website, there are 25 neighborhoods can be concluded in the city of San Francisco. The following visualization showing is based on the zip code of San Francisco (Fig.6).



Fig.6. San Francisco neighborhood visualization

For each neighborhood of San Francisco, 100 nearby venues within radius of 500 meters are extracted through the Foursquare API. From counting the result of nearby venues, it is obvious that some neighborhoods don't have 100 common venues (Table 1). This may also help to divide neighborhoods into different areas, for examples, financial district, residential area and tourist attraction.

Neighborhood	Count	Neighborhood	Count
Bayview	26	Outer Richmond	50
Castro, Noe Valley	82	Outer Sunset, Parkside, Forest Hill	45
Chinatown	82	Pacific Heights, Western Addition, Japantown	98
Cow Hollow, Marina District	77	Polk Hill, Nob Hill, Russian Hill	81
Embarcadero North	100	Potrero Hill	31
Embarcadero South	82	Presidio	28
Financial District	100	Richmond District, Inner Richmond	60
Haight District, Cole Valley	39	South of Market	77
Hayes Valley, Tenderloin, North of Market	95	St Francis Wood, Miraloma, West Portal	2
Ingleside	34	Sunset District	7
Lake Merced	51	Twin Peaks, Glen Park	17
Mission District, Inner Mission, Bernal Heights	51	Visitacion Valley, Sunnydale	4
North beach, Fishermans Wharf, Chinatown	100		

Table. 1. Counts of exploring 100 nearby venues in each San Francisco neighborhood

After the 100 nearby common venues are found for each neighborhood in San Francisco, these data will be used as a reference to find the top 10 venues in each neighborhood. All the results for the top 10 most common venues in 25 neighborhoods of San Francisco is showing in Table 2:

Bayview	Castro, Noe Valley	Chinatown
venue freq 0 Southern / Soul Food Restaurant 0.15 1 Mexican Restaurant 0.12 2 Light Rail Station 0.08 3 Bakery 0.08 4 Theater 0.04 5 BBQ Joint 0.04 6 Bus Station 0.04 7 Garden 0.04 8 Chinese Restaurant 0.04 9 Gym 0.04	venue freq 0 Gay Bar 0.10 1 Thai Restaurant 0.05 2 Coffee Shop 0.05 3 Scenic Lookout 0.04 4 Indian Restaurant 0.04 5 Yoga Studio 0.02 6 Plaza 0.02 7 Pet Store 0.02 8 Deli / Bodega 0.02 9 Convenience Store 0.02	venue freq 0 Hotel 0.10 1 Coffee Shop 0.06 2 Bakery 0.05 3 Men's Store 0.04 4 Bubble Tea Shop 0.04 5 Chinese Restaurant 0.04 6 Tea Room 0.04 7 Italian Restaurant 0.04 8 Sushi Restaurant 0.02 9 Szechuan Restaurant 0.02
Haight District, Cole Valley venue freq Coffee Shop 0.10 Boutique 0.05 Park 0.05 Thrift / Vintage Store 0.05 Smoke Shop 0.03 Breakfast Spot 0.03 Supermarket 0.03 Historic Site 0.03 Grocery Store 0.03	Hayes Valley, Tenderloin, North o venue freq 0 Coffee Shop 0.05 1 Hotel 0.04 2 Wine Bar 0.03 3 Mexican Restaurant 0.03 4 Theater 0.03 5 Café 0.03 6 Boutique 0.02 7 Sushi Restaurant 0.02 8 Burger Joint 0.02 9 Concert Hall 0.02	Financial District venue freq 0 Coffee Shop 0.07 1 Food Truck 0.05 2 Men's Store 0.04 3 Hotel 0.03 4 Café 0.03 5 Italian Restaurant 0.03 6 Gym 0.03 7 Cocktail Bar 0.03 8 Japanese Restaurant 0.03 9 Sushi Restaurant 0.03

Cow Hollow, Marina District	Outter Sunset, Parkside, Forest H	Embarcadero North
venue freq	venue freq	venue freq
0 French Restaurant 0.05	0 Chinese Restaurant 0.11	0 Food Truck 0.05
1 Italian Restaurant 0.05	1 Dumpling Restaurant 0.07	1 Café 0.04
2 Sandwich Place 0.04	2 Light Rail Station 0.07	2 Scenic Lookout 0.04
Wine Bar 0.04	3 Korean Restaurant 0.04	3 Coffee Shop 0.04
4 Gym / Fitness Center 0.04	4 Liquor Store 0.04	4 Men's Store 0.04
5 Mexican Restaurant 0.03	5 Sandwich Place 0.04	5 New American Restaurant 0.03
6 Deli / Bodega 0.03	6 Café 0.04	6 Wine Bar 0.03
7 Cosmetics Shop 0.03	7 Hobby Shop 0.02	7 Italian Restaurant 0.03
8 Salad Place 0.03	8 Bubble Tea Shop 0.02	8 Park 0.03
9 Coffee Shop 0.03	9 Gym / Fitness Center 0.02	9 Cosmetics Shop 0.03
Mission District Inner Mission Bornel		Ingleside
Mission District, Inner Mission, Bernal venue freq	Lake Merced	venue freq
0 Mexican Restaurant 0.08	venue freq	0 Pizza Place 0.12
1 Pizza Place 0.06	0 Pizza Place 0.06	1 Mexican Restaurant 0.09
2 Grocery Store 0.06	Juice Bar 0.06	2 Vietnamese Restaurant 0.06
Coffee Shop 0.06	2 Café 0.06	3 Bus Station 0.06
4 Dive Bar 0.04	3 Sandwich Place 0.06	4 Sandwich Place 0.06
5 Gym / Fitness Center 0.04	4 Cosmetics Shop 0.06	5 Fried Chicken Joint 0.03
6 Deli / Bodega 0.04	5 Coffee Shop 0.04	6 Thai Restaurant 0.03
7 Park 0.04	6 Clothing Store 0.04	7 Hot Dog Joint 0.03
8 New American Restaurant 0.02	7 Mexican Restaurant 0.04	8 Gas Station 0.03
9 Hot Dog Joint 0.02	8 Snack Place 0.04 9 Candy Store 0.04	9 Chinese Restaurant 0.03
209 002 0102	9 Candy Store 0.04	
North beach, Fishermans Wharf, Chinatown-	Outter Richmond	Embarcadero South
venue freq	venue freq	venue freq
0 Pizza Place 0.07	0 Café 0.12	0 Coffee Shop 0.17
1 Coffee Shop 0.07	1 Pizza Place 0.08	1 Food Truck 0.06
2 Café 0.06	2 Convenience Store 0.06	2 Café 0.05
3 Italian Restaurant 0.05	3 Chinese Restaurant 0.06	3 Gym 0.04
4 Park 0.05	4 Bus Station 0.04	4 Art Gallery 0.04
5 Deli / Bodega 0.04	5 American Restaurant 0.04	5 Sandwich Place 0.04
6 Bakery 0.04	6 Japanese Restaurant 0.04	6 Street Food Gathering 0.02
7 Trail 0.03	7 Sandwich Place 0.04	7 New American Restaurant 0.02
8 Chinese Restaurant 0.03	8 Dessert Shop 0.04	8 Seafood Restaurant 0.02
9 Diner 0.02	9 Optical Shop 0.02	9 Italian Restaurant 0.02
Pacific Heights, Western Addition, Japan	South of Market	Portrero Hill
venue freq	venue freq	venue freq
0 Bakery 0.05	0 Nightclub 0.09	0 Wine Shop 0.06
1 Cosmetics Shop 0.04	1 Gay Bar 0.05	1 Café 0.06
2 Café 0.04 3 Spa 0.04	2 Cocktail Bar 0.05	2 Coffee Shop 0.06
Spa 0.04 4 Pizza Place 0.03	3 Thai Restaurant 0.04	3 Breakfast Spot 0.06
5 Chinese Restaurant 0.03	4 Motorcycle Shop 0.04 5 Mexican Restaurant 0.03	4 Park 0.06
6 Boutique 0.03		5 Deli / Bodega 0.06 6 Pool 0.03
7 Asian Restaurant 0.02		
8 Park 0.02	7 Bar 0.03 8 Cosmetics Shop 0.03	7 Peruvian Restaurant 0.03 8 Office 0.03
9 Gym / Fitness Center 0.02	9 Gym / Fitness Center 0.03	9 Gym 0.03
y cym / lichebb conseil vici	y dym / lichess center 0.03	9 Gym 0:03
Polk Hill, Nob Hill, Russian Hill	Presidio	Richmond District, Inner
venue freq	venue freq	venue fre
0 Grocery Store 0.05	0 Museum 0.14	0 Bakery 0.0
1 Gym / Fitness Center 0.05	1 Food Truck 0.11	1 Burmese Restaurant 0.0
2 Sushi Restaurant 0.04	2 Café 0.07	2 Chinese Restaurant 0.0
3 Deli / Bodega 0.04	3 General Entertainment 0.04	3 Japanese Restaurant 0.0
Wine Bar 0.04	4 Historic Site 0.04	4 Vietnamese Restaurant 0.0
5 Vietnamese Restaurant 0.04	5 Fried Chicken Joint 0.04	5 Massage Studio 0.0
6 Bar 0.02	6 Mexican Restaurant 0.04	6 Pizza Place 0.0
7 Bakery 0.02	7 Brewery 0.04	7 Café 0.0
8 Chinese Restaurant 0.02	8 Bed & Breakfast 0.04	8 Wine Shop 0.0
9 Mediterranean Restaurant 0.02	9 Street Food Gathering 0.04	9 Thai Restaurant 0.0
Visitacion Valley, Sunnydale	Sunset District	Twin Peaks, Glen Park
venue freq 0 Garden 0.25	venue freq	venue freq 0 Park 0.24
-	0 Chinese Restaurant 0.29 1 Electronics Store 0.14	1 Park 0.24 1 Trail 0.18
1 Baseball Field 0.25 2 Park 0.25	2 Café 0.14	2 Dim Sum Restaurant 0.06
Park U.75		3 Athletics & Sports 0.06
		4 Coffee Shop 0.06
3 Trail 0.25	A Tiguer Chara 0 14	COTTER SHOP U.UB
3 Trail 0.25 4 New American Restaurant 0.00	4 Liquor Store 0.14	
3 Trail 0.25 4 New American Restaurant 0.00 5 Noodle House 0.00	5 Pharmacy 0.14	5 Playground 0.06
3 Trail 0.25 4 New American Restaurant 0.00 5 Noodle House 0.00 6 Nightclub 0.00	5 Pharmacy 0.14 6 Opera House 0.00	5 Playground 0.06 6 Outdoors & Recreation 0.06
3 Trail 0.25 4 New American Restaurant 0.00 5 Noodle House 0.00 6 Nightclub 0.00 7 Newsstand 0.00	5 Pharmacy 0.14 6 Opera House 0.00 7 Office 0.00	5 Playground 0.06 6 Outdoors & Recreation 0.06 7 Korean Restaurant 0.06
3 Trail 0.25 4 New American Restaurant 0.00 5 Noodle House 0.00 6 Nightclub 0.00	5 Pharmacy 0.14 6 Opera House 0.00	5 Playground 0.06 6 Outdoors & Recreation 0.06

```
----St Francis Wood, Miraloma, West Portal----
          venue
                  freq
        Bus Line
         Garden
                   0.5
    Yoga Studio
                   0.0
   Optical Shop
                   0.0
  Moving Target
         Museum
   Music School
    Music Store
                   0.0
    Music Venue
                   0.0
     Nail Salon
```

Table 2. Top 10 most common venues in 25 neighborhoods of San Francisco

The top 10 common venues in each neighborhoods are also saved in pandas dataframe(Fig.7).

	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
C	Bayview	Southern / Soul Food Restaurant	Mexican Restaurant	Light Rail Station	Bakery	Theater	Park	Plaza	Dumpling Restaurant	Coffee Shop	Chinese Restaurant
1	Castro, Noe Valley	Gay Bar	Thai Restaurant	Coffee Shop	Indian Restaurant	Scenic Lookout	Yoga Studio	Pet Store	Deli / Bodega	Café	Convenience Store
2	Chinatown	Hotel	Coffee Shop	Bakery	Tea Room	Chinese Restaurant	Men's Store	Italian Restaurant	Bubble Tea Shop	Sushi Restaurant	Dim Sum Restaurant
3	Cow Hollow, Marina District	French Restaurant	Italian Restaurant	Wine Bar	Gym / Fitness Center	Sandwich Place	Sushi Restaurant	Deli / Bodega	Park	Cosmetics Shop	Coffee Shop
4	Embarcadero North	Food Truck	Scenic Lookout	Café	Coffee Shop	Men's Store	Italian Restaurant	Wine Bar	New American Restaurant	Park	Cosmetics Shop

Fig.7. Top 10 venues in each neighborhood in pandas dataframe

Apart from analyzing the top 10 most common venues individually, it may also be beneficial to explore the popular venues in similar neighborhoods. According to the results obtained in Table 2, it can be clustered into 5 clusters (Fig.8) and its visualization is showing below (Fig.9). In Fig.8, the Cluster Labels columns indicates which neighborhood belongs to which clusters. In this case, the first 4 neighborhoods belong to the fifth cluster (label 4). And, Potrero Hill is belongs to the label 0 which is the first cluster.

	zipcode	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	
0	94102	Hayes Valley, Tenderloin, North of Market	37.779329	-122.41915	4	Coffee Shop	Hotel	Theater	
1	94103	South of Market	37.772329	-122.41087	4	Nightclub	Gay Bar	Cocktail Bar	Λ
2	94104	Financial District	37.791728	-122.40190	4	Coffee Shop	Food Truck	Men's Store	F
3	94105	Embarcadero South	37.789228	-122.39570	4	Coffee Shop	Food Truck	Café	
4	94107	Portrero Hill	37.766529	-122.39577	0	Breakfast Spot	Coffee Shop	Deli / Bodega	٧

Fig.8. example of the clustering results

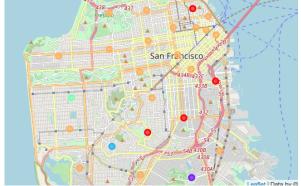


Fig.9. Clustering Visualization

Now, let's explore each cluster and analyze each cluster's top 10 most common venues. The following shows the information of 5 clusters (Table 3):

Cluster 1		Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Mos Common Venue	n Common	Common	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
	4	Portrero Hil	Breakfast Spot	Coffee Shop	Deli / Bodega	Wine Shop) Café	Park	French Restaurant	Gym / Fitness Center	Bookstore	Rock Club
	7	Mission District, Inner Mission, Berna Heights	Restaurant	Coffee Shop	Pizza Place	Grocery Store		Dive Bar	Deli / Bodega	Park	Dry Cleaner	Food & Drink Shop
	21			Trail	Grocery Store	Athletics & Sports		Korean Restaurant	Outdoors & Recreation	Playground	Shopping Mall	Dim Sum Restaurant
	23	North beach, Fishermans Wharf, Chinatowr	Coffee Shop	Pizza Place	Café	Park	Italian Restaurant		Deli / Bodega	Chinese Restaurant	Trail	Yoga Studio
Cluster 2					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
	24	Visitacion Valley, Sunnydale	Garden	Baseball Field	Trail	Park	Wings Joint	Fish Market	Fast Food Restaurant	Field	Filipino Restaurant	Flower Shop
Cluster 3		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
	18	St Francis Wood, Miraloma, West Portal	Garden	Bus Line	Wings Joint	Frozen Yogurt Shop	French Restaurant	Fountain	Food Truck	Food Stand	Food Court	Food & Drink Shop
Cluster 4					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
	19	Sunset District Re	Chinese Pestaurant	harmacy E	lectronics Store	Café	Liquor Store	Playground	Food Stand	Food Court	Food & Drink Shop	Eye Doctor
Cluster 5		Neighborhood	1st Mo Commo Venu	on Commo	on Commo	on Comm	on Commo	on Common	Common	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
	0	Hayes Valley, Tenderloin, North of Market	Coffee Sho	op Hot	tel Theat	er Wine E	Bar Ca	fé Mexican Restaurant		Cocktail Bar	Poke Place	Park
	1	South of Market	Nightch	ub Gay B	ar Cocktail B	ar Motorcy Sh	cle Th op Restaura			Restaurant	Coffee Shop	Food Truck
	2	Financial District	Coffee Sho	op Food True	ck Men's Sto	re Japane Restaur			Café	Hestaurant	Sandwich Place	Gym
	3	Embarcadero South	Coffee Sho	op Food True	ck Ca	fé Sandw Pla	ich Art Galle	ry Gym	Salad Place	New American Restaurant	Lounge	Street Food Gathering
	5	Chinatown	Hot	tel Coffee Sho		ry Tea Ro	om Chines Restaura		Italian Restaurant		Sushi Restaurant	Dim Sum Restaurant
	6	Polk Hill, Nob Hill, Russian Hil		Gym ore Fitne Cent	SS Pactaura		Bar Vietnames Restaura			Diner	Italian Restaurant	Coffee Shop
	8	Embarcadero North	Food Tru	ck Scen Looko		fé Coffee Sh	op Men's Sto	re Italian Restaurant		New American Restaurant	Park	Cosmetics Shop
	9	Ingleside	Pizza Pla	Hestaura	nt Restaura	nt Pla	ace Dus Static	nestaurani			Furniture / Home Store	Metro Station
	10	Castro, Noe Valley	Gay B	ar Restaura		p Restaur			Pet Store	Deli / Bodega	Café	Convenience Store
	11	Pacific Heights, Western Addition, Japantown	Bake	ery Ca	fé Cosmetic Sho		Spa Chines Restaura		Boutique	Yoga Studio	Arts & Crafts Store	Bubble Tea Shop
	12	Outter Sunset, Parkside, Forest Hill							Café	Pet Store	Bubble Tea Shop	Shoe Store

13	Haight District, Cole Valley	Coffee Shop	Park	Pizza Place	Boutique	Thrift / Vintage Store	Playground	Shoe Store	Supermarket	Gastropub	Bookstore
14	Richmond District, Inner Richmond	Bakery	Japanese Restaurant	Burmese Restaurant	Chinese Restaurant	Pizza Place	Vietnamese Restaurant	Massage Studio	Thai Restaurant	Wine Shop	Café
15	Outter Richmond	Café	Pizza Place	Chinese Restaurant	Convenience Store	Japanese Restaurant	Bus Station	Sandwich Place	American Restaurant	Dessert Shop	Pharmacy
16	Cow Hollow, Marina District	French Restaurant	Italian Restaurant	Wine Bar	Gym / Fitness Center	Sandwich Place	Sushi Restaurant	Deli / Bodega	Park	Cosmetics Shop	Coffee Shop
17	Bayview	Southern / Soul Food Restaurant	Mexican Restaurant	Light Rail Station	Bakery	Theater	Park	Plaza	Dumpling Restaurant	Coffee Shop	Chinese Restaurant
20	Presidio	Museum	Food Truck	Café	Art Gallery	Bed & Breakfast	Gift Shop	Field	Movie Theater	Sporting Goods Shop	Bowling Alley
22	Lake Merced	Juice Bar	Pizza Place	Cosmetics Shop	Sandwich Place	Café	Candy Store	Bakery	Snack Place	Coffee Shop	Clothing Store

Table 3. Top 10 most common venues of each neighborhood in each cluster

5 DISCUSSION

San Francisco is one of the densely populated cities in the United States and the 4th largest city in California. There are total 251 unique venues categories from the result of pull request through Foursquare API. From Table 1 shown in result section, there are some neighborhoods that don't have 100 common venues, this indicates that the neighborhood is mainly for residential purpose.

Neighborhood	Count	Neighborhood	Count
North beach, Fishermans Wharf, Chinatown	100	Mission District, Inner Mission, Bernal Heights	51
Embarcadero North	100	Outer Richmond	50
Financial District	100	Outer Sunset, Parkside, Forest Hill	45
Pacific Heights, Western Addition, Japantown	98	Haight District, Cole Valley	39
Hayes Valley, Tenderloin, North of Market	95	Ingleside	34
Castro, Noe Valley	82	Potrero Hill	31
Chinatown	82	Presidio	28
Embarcadero South	82	Bayview	26
Polk Hill, Nob Hill, Russian Hill	81	Twin Peaks, Glen Park	17
Cow Hollow, Marina District	77	Sunset District	7
South of Market	77	Visitacion Valley, Sunnydale	4
Richmond District, Inner Richmond	60	St Francis Wood, Miraloma, West Portal	2
Lake Merced	51		

Table 4. sorted count of nearby venues in each SF neighborhood

From the table shown above (Table 4), it is clear that the neighborhood with less nearby venues is more likely to be concluded as residential area. And in fact, for examples, Sunset District and Bayview are also residential neighborhoods in San Francisco. On the other hand, the neighborhood with more nearby venues are mostly tourist attractions or financial district in San Francisco, for examples, Fishermans Wharf, Chinatown and Tenderloin.

Furthermore, by analyzing the most 10 common venues of each neighborhood (Fig. 10), the most popular venue in San Francisco is Coffee Shop, Café and Park. Because San Francisco is one of the popular tourist destinations, there are variety of cuisine. The popular restaurants in San Francisco are Chinese, Italian, Mexican, Japanese, Thai, French and Vietnamese. Other popular eat places in San Francisco are pizza place, sandwich place and bubble tea shop. Other than the place to eat, for business side, cosmetics shop, men store, and boutique are also a popular venue in San Francisco. Moreover, the most common venues for activity and recreation are wine bar, gym, fitness center, playground and trail.

From Table 4, it is noticeable that half of the 25 neighborhoods in San Francisco have more than 50 common nearby venues. The more the venues in one neighborhood, the more prosperous the neighborhood is. The tourist attraction and financial district should have more venues than the residential area. Also, because of different cuisine in San Francisco, it is obvious that San Francisco is a diverse city that variety race of people living in San Francisco.

San Francisco										
Venues	Count		Venues	Count						
CoffeeShop	16		ThaiRestaurant	3						
cafe	14		YogaStudio	3						
Park	10		"MensStore"	3						
ChineseRestaurant	8		BubbleTeaShop	3						
PizzaPlace	8		FrenchRestaurant	3						
SandwichPlace	7		Gym/FitnessCenter	3						
FoodTruck	7		Gym	3						
Bakery	6		CocktailBar	3						
ItalianRestaurant	6		Boutique	3						
Deli/Bodega	5		Playground	3						
SushiRestaurant	5		VietnameseRestaurant	3						
MexicanRestaurant	4		GroceryStore	3						
WineBar	4		Food&DrinkShop	3						
CosmeticsShop	4		Trail	3						
JapaneseRestaurant	4	·								

Fig. 10 popular venues in San Francisco

By finding out the most common venues in San Francisco, it gives the businessman or investor the basic idea of popular venues in this tourist destination. It could help them to plan ahead if they would like to operate their business in tourist attraction, financial district or residential area.

In addition, clustering 25 neighborhoods into 5 clusters could also help to understand the trend in San Francisco. As mentioned in results section, the clustering result is visualized in Fig. 9.

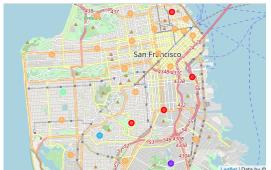


Fig.9. Clustering Visualization

In Fig. 9, it is obvious that the neighborhood in same cluster are mostly close to each other.

The red dots indicate the cluster 1; The purple dot indicates the cluster 2; The blue dot indicates the cluster 3; The turquoise dot indicates the cluster 4; and The orange dots indicate the cluster 5.

For cluster 2,3 and 4, there is only one neighborhood in each cluster. Cluster 2 is showing the Visitacion Valley neighborhood; Cluster 3 is showing the Saint Francis Wood neighborhood; and Cluster 4 is showing the Sunset District. These 3 neighborhoods are residential area.

Cluster2		Cluster 3			Cluster 4	
Venues	Count	Venues	Count		Venues	Count
Garden	1	Garden	1		Chinese Restaurant	1
Baseball Field	1	Bus Line	1		Pharmacy	1
Trail	1	Wings Joint	1		Electronics Store	1
Park	1	Frozen Yogurt Shop	1		cafe	1
Wings Joint	1	French Restaurant	1		Liquor Store	1
Fish Market	1	Fountain	1		Playground	1
Fast Food Restaurant	1	Food Truck	1		Food Stand	1
Field	1	Food Stand	1		Food Court	1
Filipino Restaurant	1	Food Court	1		Food & Drink Shop	1
Flower Shop	1	Food & Drink Shop	1		Eye Doctor	1

Fig. 11. Most 10 common venues in cluster 2, cluster 3 and cluster 4

According to the roadsnacks[5] website, Visitacion Valley is one of the worst neighborhoods in San Francisco. From the results of top 10 most common venues in cluster 2 (Fig. 11), it is obvious that common venues in this area are other than restaurants or food places. One of the most common venues in cluster 2 is fast food restaurant, this may indicate that the residents living in this neighborhood have the lower average income than others. Visitacion Valley indeed has low median income and high crime rate. Another interesting insight is the Filipino restaurant. If Filipino restaurant is the common venue in Visitacion Valley neighborhoods, it indicates that Filipino is the major residents in this area.

On the other hand, most of the common venues in cluster 3 and cluster 4 are food places. It is noticeable that most popular venues are frozen yogurt shop and French restaurant in cluster 3. This shows that the Saint Francis Wood neighborhood should be mainly developed by European. Similarly, the most common venue is Chinese restaurant in

cluster 4, this presents that the major residents in this area should be Chinese. Analyzing the top 10 most common venues individually could help the businessman or investor to discover the "trend" and popular thing within the neighborhood.

Besides, there are 4 neighborhoods in cluster 1, they are Potrero Hill, Mission District, Twin Peaks and North beach. This cluster is mixed with tourist attractions and residential areas. According to the top 10 most common venues showing in Fig. 12, coffee shop and park are also the common venues in these 4 neighborhoods.

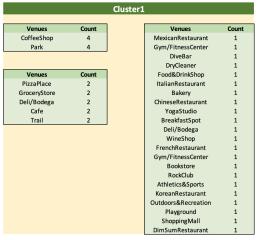




Fig. 12. Most 10 common venues in cluster 1

Fig. 13 Most 10 common venues in cluster 5

Lastly, for the cluster 5, there are 18 neighborhoods. Cluster 5 is mixed with the financial district, tourist attractions and residential areas. For those residential neighborhoods that is closed to the tourist attractions and financial district are benefited. Also, according from homesnacks[6] website, some of the neighborhoods in cluster 5 are one of the richest neighborhoods in San Francisco (eg. Pacific Heights, Noe Valley and Presidio Heights). Since this is the largest cluster out of 5 clusters, the most 10 common venues are more diverse. Coffee shop and café are the most common venues within these 18 neighborhoods. It is also obvious that there are various cuisines in cluster 5. Not to mention, the majority of common venues in cluster 5 are restaurant and eat places.

6 CONCLUSION

San Francisco is a densely populated city with diversity. As a prosperous city, there are many opportunities everywhere. The key of doing the business successfully is to develop a great business plan. It is advantageous for the businessman or the investor to find out the neighborhood that they are proposed to invest before they start planning. Researching for the best place to start their business is one of the strengths to develop the satisfactory business plan.

This project should help businessman or investor to have a brief idea of each neighborhood in San Francisco. All neighborhoods are found from the corresponding zip code in San Francisco. The 100 nearby venues for each neighborhood were found by using the Foursquare API. The results obtained from Foursquare API were used to found out the top 10 most common venues for each neighborhood.

In Discussion section, there are multiple analysis of the top 10 most common venues of each neighborhood either individually or in groups. For considering the city of San Francisco, the most popular venues are Coffee Shop, Café and Park. Also, other popular venues in San Francisco are variety of cuisine. Moreover, the most common venues also analyzed into groups. 25 neighborhoods are clustered in 5 groups. The insights for business planning could also be found in these 5 clusters. When the businessman or investor have a rough idea of which area, they would like to start their business, then they can take a look their intended neighborhoods deeply. The top 10 most common venues for each neighborhood could help them to have a better understanding to their intended neighborhoods.

To summaries, all of the analysis in the discussion section could help the businessman or the investor to understand and discover the "trend" and popular thing in each neighborhoods of San Francisco. Analyzing the neighborhoods of San Francisco in groups could also help them to get well understanding on how tourist attractions and financial district could give advantages to nearby neighborhoods.

Reference

- [1] "San Francisco, California Population 2020" worldpopulationreview.com. [Online]. Available: https://worldpopulationreview.com/us-cities/san-francisco-ca-population. [Accessed Jul.13, 2020].
- [2] "San Francisco Zip Codes San Francisco Guide" sanfranguide.com. [Online]. Available: http://sanfranguide.com/zip-codes. [Accessed Jul.13, 2020].
- [3] "San Francisco Neighborhoods as ZIP Codes" healthysf.org. [Online]. Available: http://www.healthysf.org/bdi/outcomes/zipmap.htm. [Accessed Jul.13, 2020].
- [4] "US Zip Code Latitude and Longitude" public.opendatasoft.com. [Online]. Available: https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/table/?refine.state=CA&q=san+francisco. [Accessed Jul.13, 2020].
- [5] "THE 10 WORST Neighborhoods In San Francisco For 2020" roadsnacks.net. [Online]. Available: https://www.roadsnacks.net/worst-neighborhoods-in-san-francisco-ca/. [Accessed Jul.21, 2020].
- [6] "Richest Neighborhoods In San Francisco, CA For 2020" homesnacks.net. [Online]. Available: https://www.homesnacks.net/richest-neighborhoods-in-san-francisco-129014/. [Accessed Jul.21, 2020]