**Market Analysis - Opening a Restaurant in SF – Reno Vieira Neto**

# Introduction:

A friend is thinking about opening a restaurant in San Francisco, however, he's not sure about which Neighborhood might be the best place. Besides, he doesn't know what type of cuisine either, therefore, he'd like to know what type of restaurants are more frequent in different areas of the city.

# Data:

My idea is to use Foursquare's API to download different restaurants information about each Neighborhood, so I can easily see which restaurants are more popular in different areas of the city. Besides, I think it'd be great to know the population of each area, which is the 'potential market' for the restaurant and the race distribution, so we could maybe identify a good opportunity (a certain type of cuisine that is not being offered yet).

Note: The problem with the API is that it's limited to 100 venues per call, therefore, our analysis could be compromised by the fact that we don't have access to all the data.

# Methodology:

After extracting all the data, some cleaning was necessary, because most of the venues that I got from Foursquare were Coffee Shops, then I noticed that most of restaurants have the word 'Restaurant' as 'Category', therefore, I kept only records with the 'Restaurant' word.

Later on, I realized that one of the Neighborhoods was missing population information and this was due to the fact that the source that I've used to extract the population information was grouping different zip codes, so I went back and updated the script to make every zip code unique.

# Exploratory:

## Total number of unique restaurant types

- During the exploratory analysis I was curious about the **total number of unique restaurant types** that I had on my data frame, for my surprise, I had 61 different restaurants! Which was way more than I was expecting, probably I cannot even think about 10 different types ... But that's a consequence of San Francisco and its diversity.

## Most frequent type of cuisine that we have in the city

I was curious about the **most frequent type of cuisine that we have in the city** (picture on the right side) and I was surprised to see the total number of each category, San Francisco not only has a lot of different types, but also has a lot of options within those types.

| Venue Category | |
|---|---|
| Mexican | 20 |
| Sushi | 19 |
| Italian | 19 |
| Japanese | 19 |
| Vietnamese | 18 |
| American | 18 |
| Thai | 18 |
| Korean | 17 |
| Indian | 16 |
| Seafood | 16 |

## Neighborhoods with higher concentration of different cuisines

I thought that maybe it would be great to understand the market a bit more by understanding which **neighborhoods have higher concentration of different cuisines**. In order to do that, I grouped the data frame by Neighborhood and Cuisine, did a count of it and kept only records with 8+ count.

| | Neighborhood | Venue Category | Venue | Zip Code | Population | Total Venue | Pop/Venue |
|---|---|---|---|---|---|---|---|
| **5** | Bayview-Hunters Point | Chinese | 8 | 94124 | 35492 | 37 | 959.24 |
| **116** | Hayes Valley/Tenderloin/North of Market | Sushi | 8 | 94102 | 30140 | 62 | 486.13 |
| **133** | Ingelside-Excelsior/Crocker-Amazon | Mexican | 8 | 94112 | 85373 | 52 | 1641.79 |
| **153** | Inner Mission/Bernal Heights | Mexican | 11 | 94110 | 73737 | 59 | 1249.78 |
| **193** | Lake Merced | Chinese | 9 | 94132 | 31155 | 52 | 599.13 |
| **219** | Marina | Italian | 9 | 94123 | 25461 | 60 | 424.35 |
| **241** | North Beach/Chinatown | Italian | 9 | 94133 | 26942 | 60 | 449.03 |
| **257** | Parkside/Forest Hill | Chinese | 23 | 94116 | 47708 | 64 | 745.44 |
| **286** | Polk/Russian Hill (Nob Hill) | Italian | 8 | 94109 | 56587 | 65 | 870.57 |
| **299** | Polk/Russian Hill (Nob Hill) | Sushi | 8 | 94109 | 56587 | 65 | 870.57 |
| **347** | South of Market | Sushi | 8 | 94103 | 26990 | 64 | 421.72 |
| **377** | Sunset | Chinese | 9 | 94122 | 62516 | 63 | 992.32 |
| **410** | Twin Peaks-Glen Park | Italian | 8 | 94131 | 29056 | 60 | 484.27 |
| **441** | Visitacion Valley/Sunnydale | Mexican | 9 | 94134 | 43074 | 49 | 879.06 |
| **459** | Western Addition/Japantown | Italian | 9 | 94115 | 35751 | 60 | 595.85 |
| **470** | Western Addition/Japantown | Sushi | 8 | 94115 | 35751 | 60 | 595.85 |

**The Highlights from the analysis were the following:**

- Mexican Restaurants are mostly located in Inner Mission/Bernal Heights. Opening a new Mexican place would be really challenging over there.

- Parkside/Forest Hill look like a good place for Chinese food, however, not a good place to open a new one.

## Population/Venue Analysis

I had an idea of creating a ratio of Population/Venue, basically the plan was to identify populated neighborhoods with not lots of restaurants, which could represent potential places to open new restaurants.

| | Neighborhood | Pop/Venue | Total Venue | Population |
|---|---|---|---|---|
| **120** | Ingelside-Excelsior/Crocker-Amazon | 1641.79 | 52 | 85373 |
| **141** | Inner Mission/Bernal Heights | 1249.78 | 59 | 73737 |
| **375** | Sunset | 992.32 | 63 | 62516 |
| **0** | Bayview-Hunters Point | 959.24 | 37 | 35492 |
| **427** | Visitacion Valley/Sunnydale | 879.06 | 49 | 43074 |
| **277** | Polk/Russian Hill (Nob Hill) | 870.57 | 65 | 56587 |
| **254** | Parkside/Forest Hill | 745.44 | 64 | 47708 |
| **70** | Haight-Ashbury | 726.83 | 60 | 43610 |
| **166** | Inner Richmond | 668.02 | 62 | 41417 |
| **190** | Lake Merced | 599.13 | 52 | 31155 |
| **451** | Western Addition/Japantown | 595.85 | 60 | 35751 |
| **303** | Potrero Hill | 575.38 | 52 | 29920 |
| **13** | Castro/Noe Valley | 566.57 | 61 | 34561 |
| **96** | Hayes Valley/Tenderloin/North of Market | 486.13 | 62 | 30140 |
| **401** | Twin Peaks-Glen Park | 484.27 | 60 | 29056 |
| **232** | North Beach/Chinatown | 449.03 | 60 | 26942 |
| **351** | St. Francis Wood/Miraloma/West Portal | 448.79 | 47 | 21093 |
| **210** | Marina | 424.35 | 60 | 25461 |
| **325** | South of Market | 421.72 | 64 | 26990 |
| **42** | Chinatown | 232.12 | 64 | 14856 |

With that analysis I was able to conclude the following:

- Mexican Restaurants are mostly located in Inner Mission/Bernal Heights. Opening a new Mexican place would be really challenging over there.

- Parkside/Forest Hill look like a good place for Chinese food, however, not a good place to open a new one.

**And my recommendation was to choose neighborhoods with less restaurants to reduce the number of competitors: Bayview-Hunters Point and Visitacion Valley/Sunnydale.**

## Race Analysis

After merging the race information with both potential neighborhoods that I was able to identify:

- Restaurant type in Bayview-Hunters Point:

| | Zip Code | Race | Population |
|---|---|---|---|
| 0 | 94134 | Asian - Chinese | 41.122952 |
| 1 | 94134 | White | 16.918398 |
| 2 | 94134 | Some Other Race | 12.167446 |
| 3 | 94134 | Asian - Filipino | 10.386159 |
| 4 | 94134 | Black or African American | 9.974699 |

The recommended cuisine in this area would be African cuisine, given that Black or African American represents 35% of the area and I see only 4 Southern/Soul Food restaurants.

Filipino cuisine could be also a possibility, despite the Filipino population not being really big in the area, I don't see any Filipino restaurants in the area.

- Restaurant type in Visitacion Valley/Sunnydale:

| | Neighborhood | Venue Category | Venue | Zip Code | Population | Total Venue | Pop/Venue |
|---|---|---|---|---|---|---|---|
| **441** | Visitacion Valley/Sunnydale | Mexican | 9 | 94134 | 43074 | 49 | 879.06 |
| **450** | Visitacion Valley/Sunnydale | Vietnamese | 5 | 94134 | 43074 | 49 | 879.06 |
| **447** | Visitacion Valley/Sunnydale | Southern / Soul Food | 4 | 94134 | 43074 | 49 | 879.06 |
| **440** | Visitacion Valley/Sunnydale | Latin American | 4 | 94134 | 43074 | 49 | 879.06 |
| **429** | Visitacion Valley/Sunnydale | Asian | 3 | 94134 | 43074 | 49 | 879.06 |
| **442** | Visitacion Valley/Sunnydale | New American | 2 | 94134 | 43074 | 49 | 879.06 |
| **448** | Visitacion Valley/Sunnydale | Sushi | 2 | 94134 | 43074 | 49 | 879.06 |
| **435** | Visitacion Valley/Sunnydale | Hunan | 2 | 94134 | 43074 | 49 | 879.06 |
| **437** | Visitacion Valley/Sunnydale | Italian | 2 | 94134 | 43074 | 49 | 879.06 |
| **428** | Visitacion Valley/Sunnydale | American | 2 | 94134 | 43074 | 49 | 879.06 |
| **449** | Visitacion Valley/Sunnydale | Thai | 1 | 94134 | 43074 | 49 | 879.06 |
| **446** | Visitacion Valley/Sunnydale | Seafood | 1 | 94134 | 43074 | 49 | 879.06 |
| **445** | Visitacion Valley/Sunnydale | Salvadoran | 1 | 94134 | 43074 | 49 | 879.06 |
| **444** | Visitacion Valley/Sunnydale | Restaurant | 1 | 94134 | 43074 | 49 | 879.06 |
| **443** | Visitacion Valley/Sunnydale | Peruvian | 1 | 94134 | 43074 | 49 | 879.06 |
| **427** | Visitacion Valley/Sunnydale | African | 1 | 94134 | 43074 | 49 | 879.06 |
| **438** | Visitacion Valley/Sunnydale | Japanese | 1 | 94134 | 43074 | 49 | 879.06 |
| **436** | Visitacion Valley/Sunnydale | Indian | 1 | 94134 | 43074 | 49 | 879.06 |
| **434** | Visitacion Valley/Sunnydale | Filipino | 1 | 94134 | 43074 | 49 | 879.06 |
| **433** | Visitacion Valley/Sunnydale | Fast Food | 1 | 94134 | 43074 | 49 | 879.06 |
| **432** | Visitacion Valley/Sunnydale | Dim Sum | 1 | 94134 | 43074 | 49 | 879.06 |
| **431** | Visitacion Valley/Sunnydale | Chinese | 1 | 94134 | 43074 | 49 | 879.06 |
| **430** | Visitacion Valley/Sunnydale | Caribbean | 1 | 94134 | 43074 | 49 | 879.06 |
| **439** | Visitacion Valley/Sunnydale | Korean | 1 | 94134 | 43074 | 49 | 879.06 |

A Chinese restaurant could be a good idea in the area, if you look at the population data frame, Chinese represents 41% of the population while we have only one Chinese restaurant in the area.

Filipino could also be a possibility here, given the size of the population and the fact that there's only one Filipino restaurant in the area.

## Conclusion:

More data would be necessary to make a precise recommendation, it would be also interesting to see the rent price in each area, so we could estimate how expensive it would be to have a restaurant in each area. Another point is that we don't have all the data, the API limits to 100 venues per call, therefore, it's possible that the data is not complete.

Also, the population data is based off people who resides in the area, therefore, it's possible that during the day the distribution is different. Downtown areas will be busier during the week and it's important to know exactly what type of restaurant we are trying to create.

Regardless, I think that this is already good for my friend, he will be happy with my suggestions.