

# The Battle of Neighbourhoods

Yiteng Zhang

August 2020

## 1 Introduction

The San Francisco Bay Area, popularly referred to as the Bay Area or simply the Bay, is a populous region surrounding the San Francisco, San Pablo, and Suisun Bay estuaries in Northern California.

The Bay Area has the second-most Fortune 500 companies in the United States, after the New York metropolitan area, and is known for its natural beauty, liberal politics, entrepreneurship, and diversity.

Among the Fortune 500 companies located in the region include technology companies Google, Facebook, Apple Inc., Hewlett Packard, Intel, Adobe Inc., Applied Materials, eBay, Cisco Systems, Symantec, Oracle, Netflix, Sony Interactive Entertainment, Electronic Arts, and Salesforce; energy companies Chevron and PG&E; financial service companies Charles Schwab Corporation, Visa Inc., and Wells Fargo; apparel retailers Gap Inc., Levi Strauss & Co., and Ross Stores; aerospace and defense contractor Lockheed Martin; local grocer Safeway; pharmaceutical company McKesson; and biotechnology companies Genentech and Gilead Sciences. The largest manufacturers include Tesla Inc., Lam Research, Bayer, Chevron, and Coca-Cola. Oakland is the site of the fifth-largest container shipping port in the United States and is also a major rail terminus. In research, NASA's Ames Research Center and the federal research facility Lawrence Livermore National Laboratory are based in Mountain View and Livermore respectively. In the North Bay, Napa and Sonoma counties are known for their famous wineries, including Fantesca Estate & Winery, Domaine Chandon California, and D'Agostini Winery.

### 1.1 Problem Statement

Nearly all those technology companies have headquarters located in the Bay Area. During the lunch time, most of the staffs will either grab a quick lunch in the nearby restaurants or have a lunch at company's canteen. It turns out there are already many restaurant chains in the valley and near the offices of almost every technology company. However, this still attracts some adventurers to start a new restaurant business in the Bay Area. This capstone tries to use data science technique to analyse the restaurants and in particular the cuisine

of the restaurant near the Bay Area to find out the best possible neighbourhood for opening a new business.

## 1.2 Objectives

The main objective is to analyse the restaurant chains in the Bay Area near the technology companies and to look for the areas where a successful restaurant business can be started.

The second objective is to pay special attention to the restaurant cuisine. It aims to analyse the restaurants and in particular the cuisine of the restaurant near the Bay Area and to find out the best possible neighbourhood for opening a new business.

## 1.3 Beneficiaries

This capstone is beneficial to anyone who is looking to start a new restaurant business in the Bay Area. The restaurants are in operations already can utilise the analysis to adjust their business prospects also.

# 2 Data Integration

The data requirements of this capstone include:

- List of all the technology companies;
- The coordinates of the offices;
- The list of all restaurants near the offices using Foursquare API.

The Bay Area companies list is an open source list of San Francisco Bay Area. According to the author, Connor Leech, the focus in creating this list is to make a resource for candidates to find tech jobs in the Bay Area. The .csv file has been uploaded to Github and will be used for the analysis.

data.head()							
	Company Name	Tags	Location	Investors	Description	Website	Foursquare
0	3scale	B2B Software and Services, Developer Tool	San Francisco	NaN	Unlock the power of your APIs (owned by RedHat)	<a href="https://www.3scale.net/">https://www.3scale.net/</a>	20
1	8tracks	Music, Consumer Goods and Services	San Francisco	NaN	8tracks is the best place to make and discover...	<a href="https://8tracks.com/">https://8tracks.com/</a>	20
2	10 by 10	B2B Software and Services, Talent Acquisition	San Francisco	Y Combinator	We help independent recruiters and agencies ea...	<a href="https://www.10by10.io/">https://www.10by10.io/</a>	20
3	15Five	Employee Engagement	San Francisco	NaN	15Five software elevates the performance and e...	<a href="https://15five.com">15five.com</a>	20
4	21Tech	Real Estate and Construction, Financial Technol...	East Bay	NaN	Enterprise Asset Management: 21Tech has broad ...	<a href="https://21tech.com">21tech.com</a>	19

Figure 1: A glimpse of the dataset.

## 2.1 Methodology

This subsection explains how the data will be used to solve the problem. Some highlights are given here:

- The data/list is extracted from Github.
- The data consists 750+ rows. The first 200 rows were extracted from the data source.
- Feature selection and cleansing were carried out.
- Map for all the offices in the Bay Area was plotted.
- Foursquare API was used to find the restaurants near the offices.
- The restaurants were classified using k-means algorithm.
- The plot of all the classified restaurants and the offices was made on the map of the Bay Area.

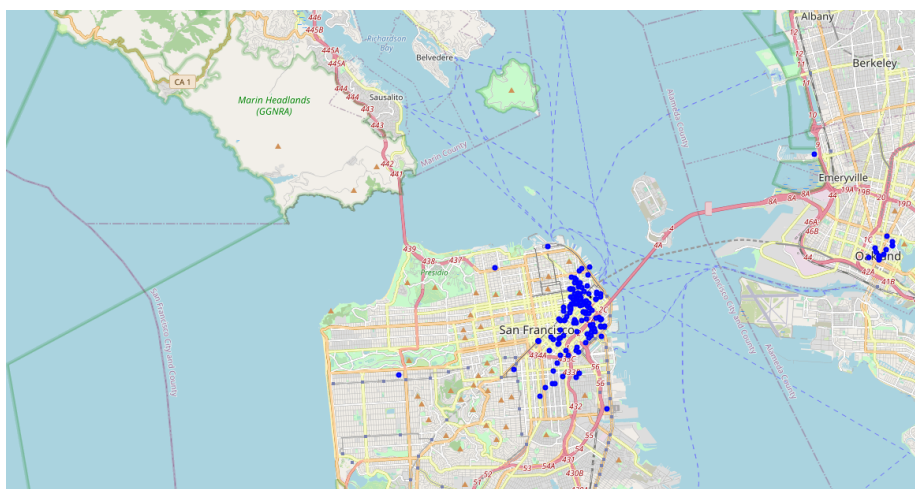


Figure 2: Map of offices in the Bay Area.

## 3 Results

The restaurants near the offices listed have been split into five clusters using k-means algorithms. Below is the map of all five clusters. The companies are plotted on the same map for evaluation purpose.

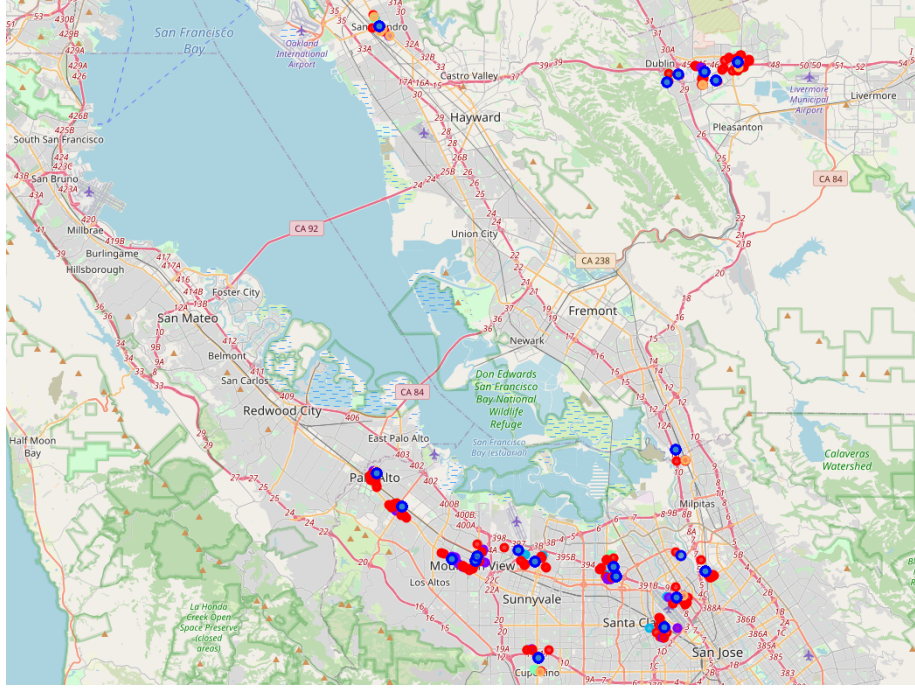


Figure 3: Map of offices along the clusters of restaurants.

The colour code of five clusters is described as follows

Clusters	Colour	Cuisine
0	Red	Japanese, Indian, Sushi
1	Purple	Mexican
2	Light Blue	American
3	Green	Chinese
4	Orange	Vietnamese

The cluster 0 contains the most number of restaurants. Cluster 1 contains 85 restaurants of Mexican cuisine. Cluster 2 has 58 American restaurants. Cluster 3 and Cluster 4 have 69 Chinese restaurants and 65 Vietnamese restaurants, respectively.

## 4 Discussion

Based on 3, there are numerous observations we can draw. Below lists a few selected observations:

- North Bay has very few restaurants. But interestingly these few restaurants cover various cuisines. Simply judging by the number of restaurants,

just opening one another restaurant regardless of cuisines might not be a bad idea.

- In East Bay area near Art.com, there is only one restaurant nearby. Several other restaurants in Cluster 0 are a bit far away. So opening a new restaurant as long as a Mexican restaurant shall be a good idea.
- In Chinatown near Okaland City Center, there are quite many Chinese restaurants and few Far East restaurants. This is a good at least. As other cuisines might not be successful in Chinatown area.
- Around Oakland City Center, there are no Mexican restaurants.
- In the San Lendro region, there is no American restaurant, so opening one there might lead to a good business.
- In San Ramon Marketplace Shopping Center, there is no Mexican restaurant in their food court.
- Many restaurants are located in South Bay area, thus a lot more observations can be drawn here.

## 5 Conclusion

The report documents how to use data technique to analyse the restaurants and in particular the cuisine of the restaurant near the Bay Area to find out the best possible place for opening a new business. The restaurants have been clustered into five groups and the cluster map can offer insights to advise where might be a good place to start a restaurant business in the Bay Area.