The Battle of Neighbourhoods

IBM Data science Professional Certificate Capstone Report

Which type of cuisine restaurant would be the best option to open in Bay Area and near which company?

Naman Manchanda – 28 July 2019

Introduction

Silicon Valley needs no introduction. It is considered as the cradle of innovation. Many tech companies have their offices and head quarters situated in the Bay Area. But what if someone wants to start a restaurant business in the Bay Area? There are already many restaurant chains in the valley and near the offices of almost every tech company. So how can someone address this problem?

BUSINESS PROBLEM

The main idea behind the project is to analyse the restaurant chains in the Bay Area near the tech companies and to look for the areas where a successful restaurant business can be started. The other objective of the project is to pay special attention to the restaurant cuisine. Some neighbourhoods might already have a large number of same cuisine restaurants. Starting a restaurant in such an area won't be much profitable. So to explain the project in an abstract, it aims at analysing 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 restaurant.

Target Audience

The project is beneficial to anyone who is looking to start a new restaurant business in the Bay Area. The audience is quite large and includes the big food supply brands like McDonalds and the Cheesecake Factory. The project can also benefit already running restaurants to shift to some other cuisine for better business prospects.

Data Description

The data requirements include:
➡ List of all the tech companies
➡ The coordinates of the offices
➡ The list of all the restaurants near the offices using Foursquare API

How the data will be used to solve the problem?

Firstly, the data will be cleaned to get the list of some of the companies of the area. After the pre-processing, the list of the offices will be used to explore the venues near the offices within the radius of 1000 using the Foursquare API. The venues will be filtered to obtain the list of the restaurants in the region. Then using KMeans clustering and data visualisation using Folium, a clear idea will be obtained on where to start the next restaurant business. The clusters' evaluation will describe the best possible regions to open which particular cuisine restaurants.

Data acquisition source

https://github.com/namanmanchanda09/tech-companies-bay-area/blob/master/Bay-Area-Companies-List.csv

The Bay Area companies list is an open source list created by Mr. Connor Leech (https://github.com/connor11528) and is available on Github. The repo has been forked from Mr. Connor for the purpose of analysis and visualisation. The .csv file has been uploaded on IBM cloud and will be used for the analysis.
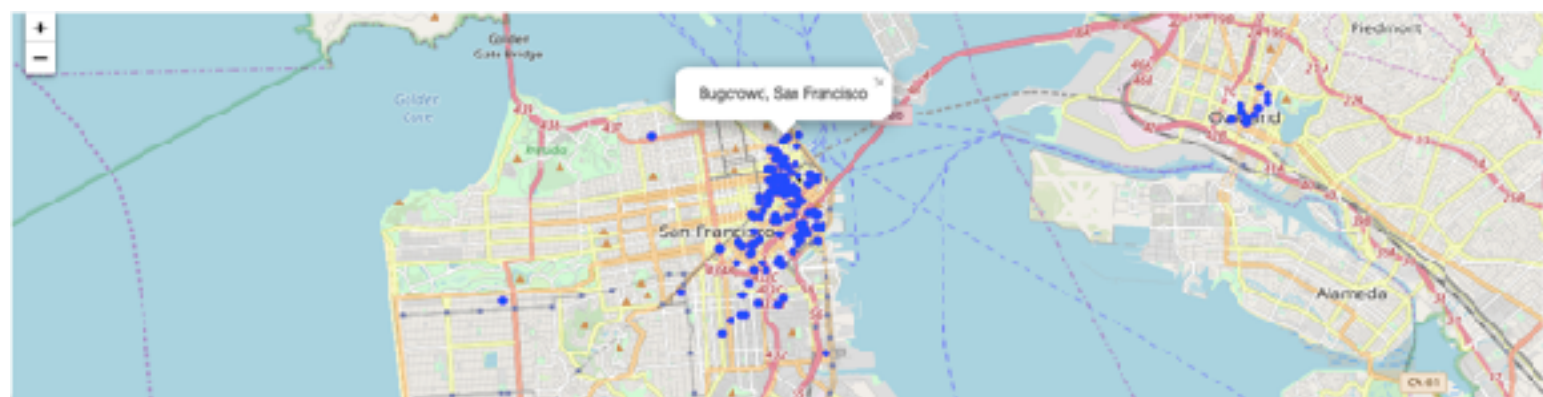
A glimpse of the dataset

Company Name	Location	Investors	Description	Address	Lat	Long
3scale	San Francisco	Nan	Unlock the power of APIs	450 Townsend St, San Francisco, Ca	37.33	-122.39
8tracks	San Francisco	Nan	8tracks is the best place to make and discover…	51 Sharon St, SF, CA	37.76	-122.42

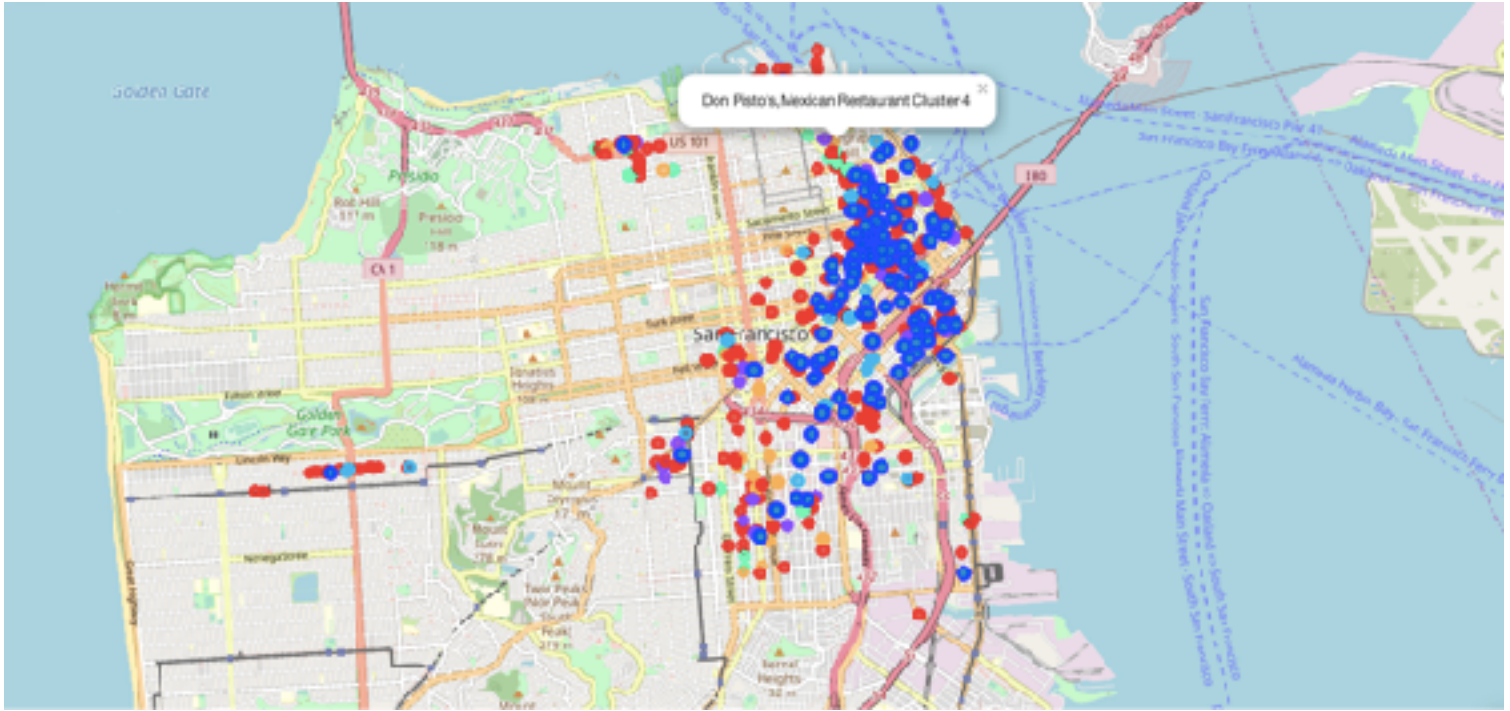| Company Name	Location	Investors	Description	Address	Lat	Long
10 by 10	San Francisco	Y Combinator	We help independent recruiters and agencies…	12 Gallagher Ln, SF, CA	37.77	-122.41

Data Science Methodology

- The data was extracted from the open source version control platform Github.
- The data consisted of 750+ rows. The first 200 rows were extracted from the datasource.
- Feature Selection and cleaning was carried out on the data.
- Map for all the offices in the Bay Area was plotted.

<u>Map of offices in the bay area</u>

- Foursquare API was used to find the restaurants near the offices.
- OneHotEncoding was used to create dummy variables for the VenueCategory as it is a categorical variable.
- The restaurants were classified using KMeans ML algorithm.
- The plot of all the classified Restaurants and the Offices was made on the map of Bay Area.

Map of offices along with the clusters of restaurants

Results

CLUSTERS		COLORS	RESTAURANT CUISINE
	0	Red	American, Japanese, Sushi
	1	Purple	New American
	2	Light Blue	Vietnamese
	3	Green	Italian
	4	Italian	Mexican

Observations and Discussions

- **1. The Mexican and New American cuisine restaurants are very less near the Cadre Talent office, San Francisco. So a Mexican or a New American restaurant can be started near the area.**

- 2. In the area near Canvas Medical, San Francisco, there is not even a single New American, Mexican or Italian Restaurant. Starting one of these restaurant in the area might be a good idea.

- 3. In the East Bay region near Art.com company, there is neither an Italian Restaurant nor a New American one.

- 4. In the San Leandro region near the Bay, there is no New American Restaurant, so opening one there might lead to really good business.

- 5. Many other observations can be made from the map.

CONCLUSION

The report has been concluded on the account that all restaurants in the Bay Area have been plotted on the maps along with the offices. The restaurants have been clustered and can help anyone looking forward to start a restaurant business in Bay Area.