CAPSTONE PROJECT – THE BATTLE OF NEIGHBORHOODS

Introduction

- A chef wants to open his own restaurant in the city of San Francisco. Indian cuisine is his forte. He wants to find a neighborhood with restaurants around indicating there is customer volume. The chef also wants to make sure there are not many Indian restaurants in the area around. So that, there might be interest for a new cuisine.

Objective

- We will study and explore the neighborhoods of San Francisco using data from Foursquare. The aim of this project is to identify ideal locations to open a new Indian restaurant based on number of restaurants around, type of cuisines, number of Indian restaurants etc.

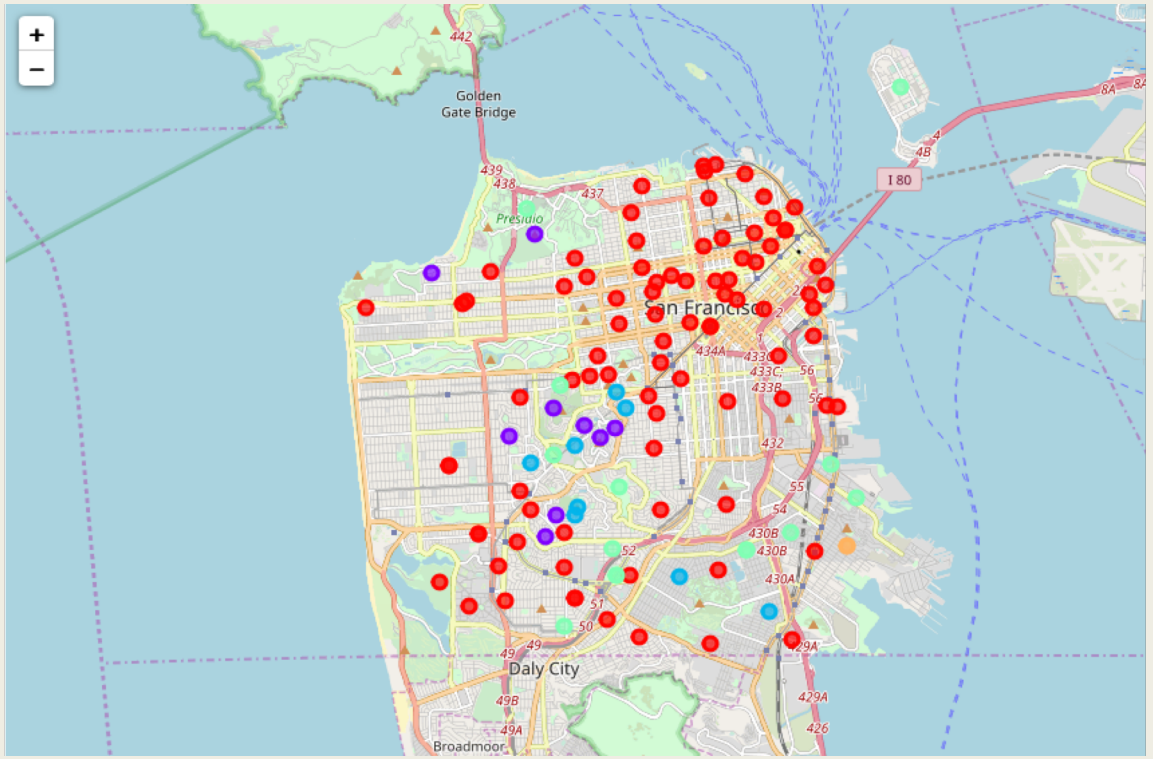
Data

- List of neighborhoods in San Francisco will be scraped from the following Wikipedia Page: https://en.wikipedia.org/wiki/List_of_neighborhoods_in_San_Francisco
- Google geocoder will be used to get the latitude and longitude for the neighborhoods.
- Foursquare API will be used to get venue data based on latitude and longitude.
- K-Means clustering will be used identify the ideal neighborhoods for a new Indian restaurant in San Francisco.

Methodology

- Used K-Mean clustering to segment and cluster the neighborhoods in San Francisco.
- Used the neighborhood characteristics to identify a suitable cluster.

Methodology



Results

- After examining each cluster, I found cluster 4 more suitable or ideal to open an Indian restaurant as it has good number of neighborhood options in it, good number of restaurants in its most common venues – indicating there is demand for restaurants in the area. But, also at the same time, there were not many Indian restaurants in the cluster, indicating there is scope to open a new one.

Results

```
sf_merged.loc[sf_merged['Cluster Labels'] == 3, sf_merged.columns[[1] + list(range(5, sf_merged.shape[1]))]].head(5)
```

	Latitude	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
3	37.724569	Park	BBQ Joint	Sandwich Place	Light Rail Station	Dessert Shop	Pool	Food & Drink Shop	Farm	Ethiopian Restaurant
22	37.742372	Video Store	Grocery Store	Salon / Barbershop	Athletics & Sports	Trail	Shopping Mall	Dim Sum Restaurant	Bus Line	Coffee Shop
40	37.740435	Food Truck	Thrift / Vintage Store	Brewery	Garden Center	Food	Non-Profit	Women's Store	Filipino Restaurant	Exhibit
45	37.746968	Light Rail Station	Moving Target	Bus Station	Theater	Wine Shop	Music Venue	Furniture / Home Store	Building	Pier
49	37.748976	Light Rail Station	Japanese Restaurant	Hotpot Restaurant	Bus Stop	Event Space	Park	Art Gallery	Lake	French Restaurant

Discussion

- Some of the challenging factors that could be considered along with these results are personal preferences of the chef as to the price width of the restaurant he is looking to open. For it, the economic status of the area would also have to be considered into the analysis which was missing here. Also, the area population would also play a part.

Conclusion

- So, after analyzing the data, I recommend neighborhoods in cluster 4 as suitable to open new Indian restaurant.
- Few neighborhoods in cluster 4 are: Balboa Park, Diamond Heights, India Basin, Islais Creek, Laguna Honda