BATTLE OF THE NEIGHBORHOODS

Applied Data Science Capstone

Background

Malaysia is a multi-ethnic and multi-cultural country, part of South East Asia with the largest population consists of Malay, Chinese and Indian. Chinese population is the second largest ethnic in Malaysia, as such there is high opportunity to invest in the Chinese restaurant business.

Problem Statement

The question here is how entrepreneurs can find the best locations (Kuala Lumpur or Penang) in Malaysia to open a Chinese restaurant. I will take into consideration two main factors for the location selection:

1) Chinese population by states
2) Total restaurant in the state

Data

1) The Foursquare location data will be used to explore the restaurant in Kuala Lumpur and Penang, in order to get the total restaurant in both states.
2) Use Geocoder API to extract longitude, latitude coordinates.
3) I will scrap the Wikipedia to derive the Malaysia demographic data.

Data source:

Wikipedia Malaysia postal code

https://ms.wikipedia.org/wiki/Senarai_poskod_di_Malaysia

Foursquare categories of Restaurant
https://developer.foursquare.com/docs/resources/categories

Malaysia demographic

Source: National Census 2010, Department of Statistics Malaysia

Target Audience

Entrepreneurs who are planning to open a new restaurant in Malaysia's two main states Kuala Lumpur and Penang.
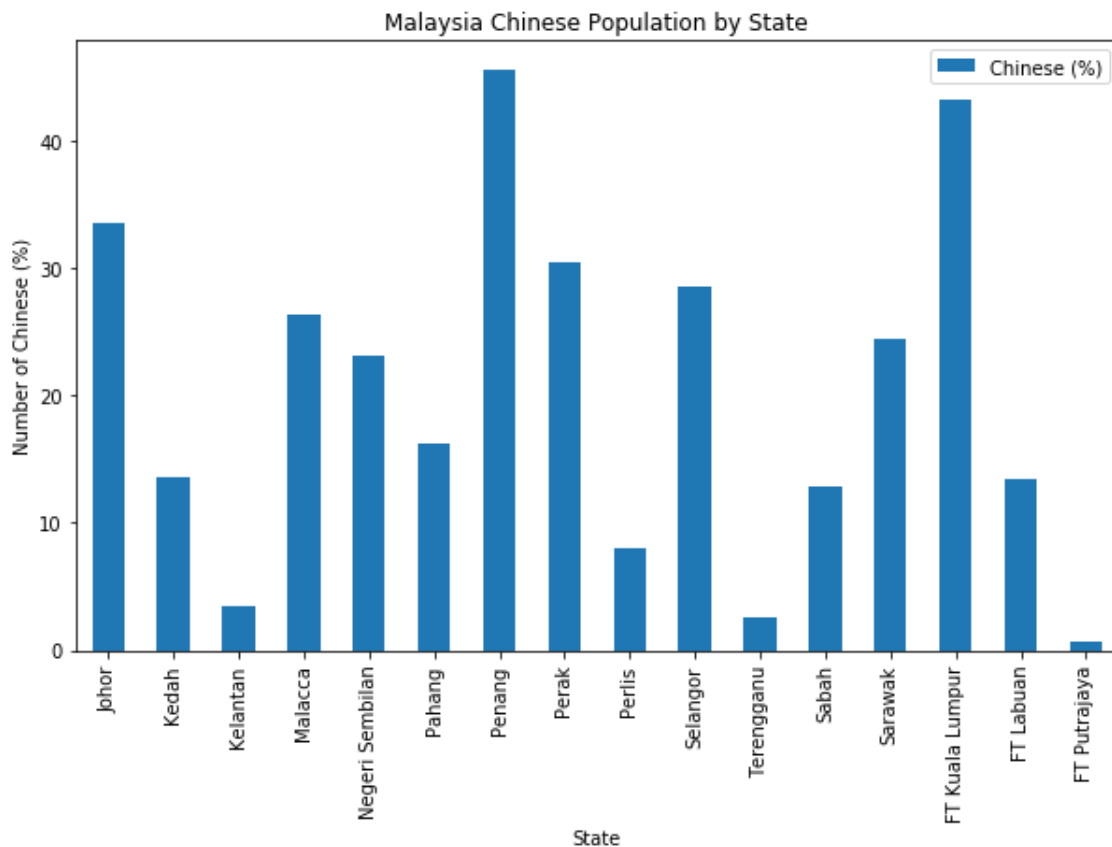
Methodology

Step 1 – Business Understanding

The goal is to create a profile of restaurants in both Kuala Lumpur and Penang. This profile will be the reference point for entrepreneurs who would like to open a restaurant business.
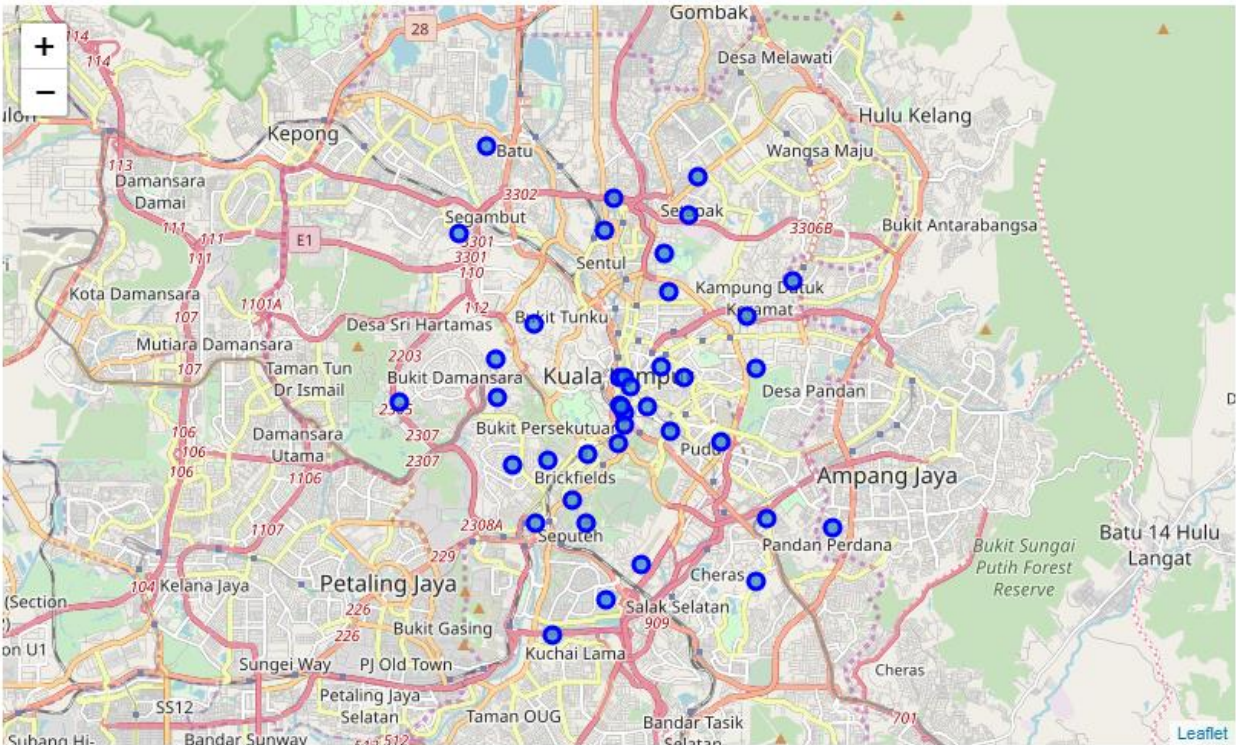
Step 2 – Data Requirement

As mentioned above section, data are collected from Wikipedia, Foursquare and Department of statistics Malaysia.

Step 3 – Data Analysis

Above Bar graph shown that two states with highest percentage of Chinese population are Penang and Kuala Lumpur. Both are having more than 50% of Chinese.
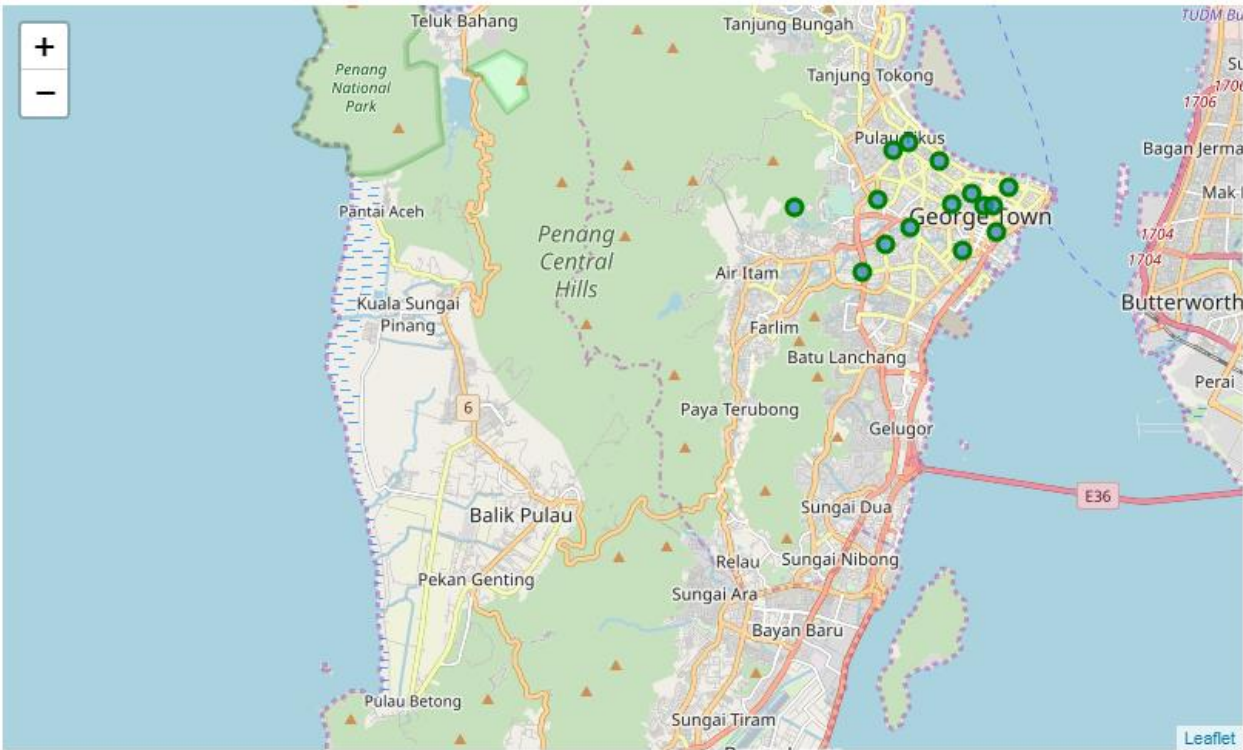
Neighborhoods of Kuala Lumpur by Post Code

Coordinates of Kuala Lumpur by Post Code

	PostalCode	State	State Abbreviation	Latitude	Longitude
0	50000	Kuala Lumpur	KUL	3.1433	101.6955
1	50050	Kuala Lumpur	KUL	3.1451	101.6945
2	50100	Kuala Lumpur	KUL	3.1513	101.6947
3	50150	Kuala Lumpur	KUL	3.1406	101.6955
4	50200	Kuala Lumpur	KUL	3.1447	101.7007
5	50250	Kuala Lumpur	KUL	3.1447	101.6951

Neighborhoods of Penang by Post Code

Coordinates of Penang by Post Code

	PostalCode	State	State Abbreviation	Latitude	Longitude
0	10000	Pulau Pinang	PNG	5.4166	100.3305
1	10050	Pulau Pinang	PNG	5.4195	100.3278
2	10100	Pulau Pinang	PNG	5.4166	100.3326
3	10150	Pulau Pinang	PNG	5.4067	100.3259
4	10200	Pulau Pinang	PNG	5.4208	100.3361
5	10250	Pulau Pinang	PNG	5.4265	100.3205

	Post Code	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	10000	Chinese Restaurant	Vegetarian / Vegan Restaurant	Thai Restaurant	American Restaurant	Italian Restaurant
1	10050	Chinese Restaurant	Vegetarian / Vegan Restaurant	Asian Restaurant	Thai Restaurant	Mamak Restaurant
2	10100	Chinese Restaurant	Thai Restaurant	Malay Restaurant	Indian Restaurant	Japanese Restaurant
3	10150	Chinese Restaurant	Asian Restaurant	Mamak Restaurant	Vietnamese Restaurant	Japanese Curry Restaurant
4	10200	Chinese Restaurant	Restaurant	Indian Restaurant	Japanese Restaurant	Asian Restaurant

Top 5 Restaurants of Penang

	Post Code	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	50000	Chinese Restaurant	Asian Restaurant	Malay Restaurant	Restaurant	Thai Restaurant
1	50050	Asian Restaurant	Malay Restaurant	Chinese Restaurant	Restaurant	Thai Restaurant
2	50100	Indian Restaurant	Malay Restaurant	South Indian Restaurant	Restaurant	Vegetarian / Vegan Restaurant
3	50150	Chinese Restaurant	Asian Restaurant	Malay Restaurant	Restaurant	Halal Restaurant
4	50200	Restaurant	Asian Restaurant	Chinese Restaurant	Malay Restaurant	Halal Restaurant

Top 5 Restaurants of Kuala Lumpur

Based on above table, residents of Penang prefer Chinese restaurant as their top choice, whereas Kuala Lumpur residents has a more varieties of taste by different Post Code/Area.

```
#to get total venues of Kuala Lumpur
print('{} Kuala Lumpur restaurant were returned by Foursquare.'.format(KL_venues.shape[0]))

517 Kuala Lumpur restaurant were returned by Foursquare.

#check how many venues were returned for each Post Code, sort by descending order
KL_venues.groupby('Post Code').count().sort_values(by='Venue Category', ascending=False)
```

Post Code	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
50400	34	34	34	34	34	34
55100	34	34	34	34	34	34
50350	30	30	30	30	30	30
50100	28	28	28	28	28	28
50450	26	26	26	26	26	26

```
#to get total venues of Penang
print('{} Penang restaurant were returned by Foursquare.'.format(Penang_venues.sha

227 Penang restaurant were returned by Foursquare.

#check how many venues were returned for each Post Code, sort by descending order
Penang_venues.groupby('Post Code').count().sort_values(by='Venue Category', ascend
```

Post Code	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
10050	35	35	35	35	35	35
10350	35	35	35	35	35	35
10200	33	33	33	33	33	33
10100	26	26	26	26	26	26
10470	19	19	19	19	19	19

Total restaurants of Kuala Lumpur and Penang and total by Top 5 Post Code

Kuala Lumpur is having total of 517 restaurants compare to Penang only has 227 which is less than 50% of Kuala Lumpur.

State	Capital	Royal Capital	Population[1]
Penang	George Town	—	1,663,000
Kuala Lumpur	Kuala Lumpur	—	1,768,000

However, looking at the total population of Kuala Lumpur which is higher than Penang, there is still potential to setup a restaurant there.

Discussion

In this study, I analyzed the neighborhoods of Kuala Lumpur and Penang by Post Code. By using the Foursquare API, I derived the nearby restaurants of each Post Codes for further analysis. It is recommended to open Chinese restaurant in either Kuala Lumpur or Penang, looking at the high population of Chinese residents in both States.

Kuala Lumpur is having more restaurants compare to Penang. This is due to the bigger land area of Kuala Lumpur and there is still potential business to open a new restaurant in Kuala Lumpur.

Conclusion

Based on the analysis, it is recommended to open a new restaurant in Kuala Lumpur.