

Hong Kong Food Clusters analysis

Law Kwan Ping
May 2019


Analyzing Restaurant Clusters among Hong Kong

This project will analyze the geographic distribution of restaurants from a data science perspective and we will analyze the cuisine's clusters and the price levels around hong kong.

- What food choices do we have in the areas?
- What kinds of restaurants are more likely comes together?
- How the different restaurant categories form the districts?
- Hong kong restaurant divergence in cuisine styles and the price level



Data acquisition & adjustment

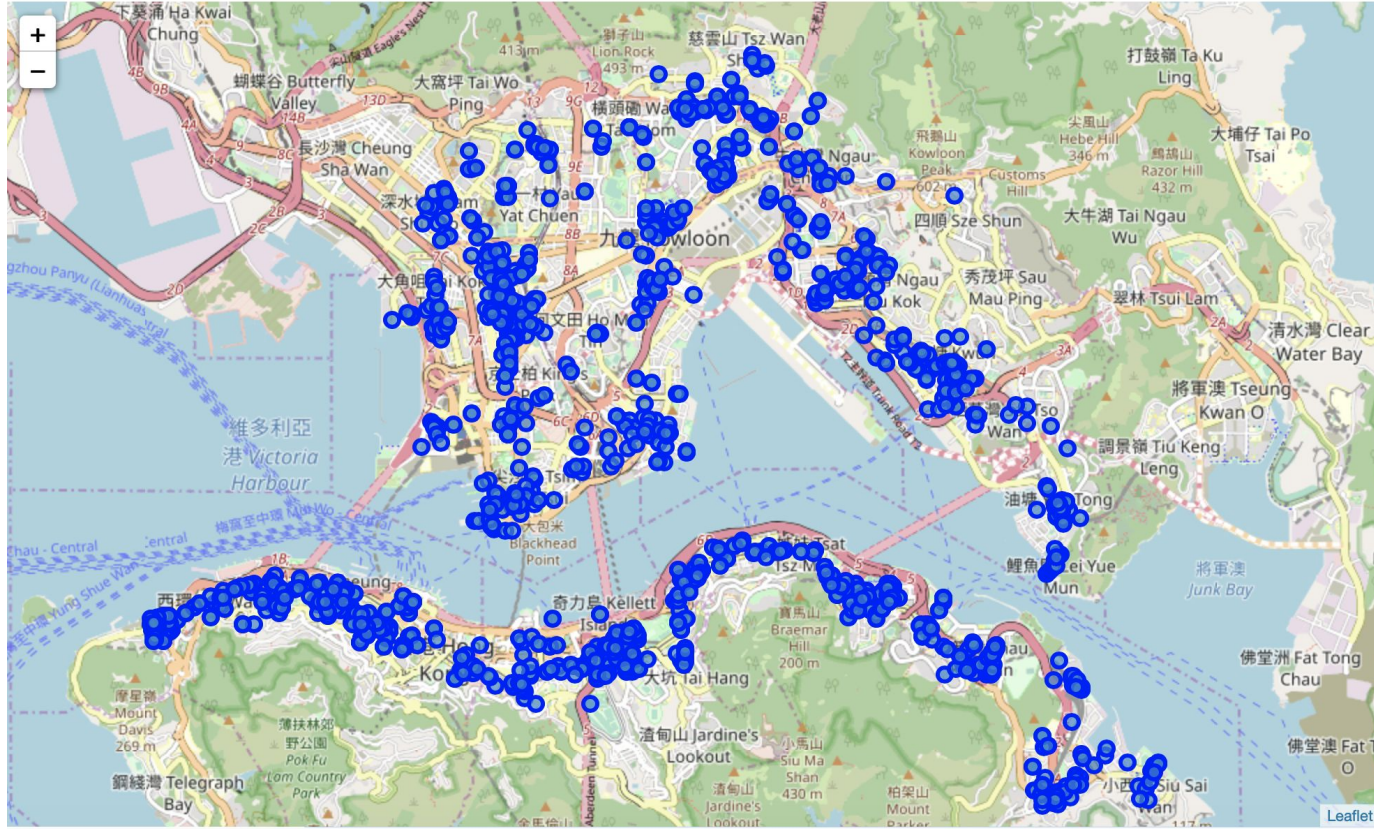
- Explore hong kong restaurant near MTR lines.
 - The city is well covered by the Kwun Tong Line, Island Line, part of Tsuen Wan Line and some popular stations.
 - Substitute those unrecognizable stations with the location name
 - - Whampoa -> HUNG HOM
 - - HKU -> SHEK TONG TSUI
 - - Wan Chai -> Wanchai
 - - Central -> CENTRAL,HK
 - - Wong Tai Sin -> CHUK YUEN
 - - Kowloon -> Austin
 - - Olympic-> Olympic,hk
- 

Understand the Data

1. We queried Foursquare site 35 times to get 1749 rows of data.
2. There are total 91 uniques Categories.
3. Lets see the data visualization.

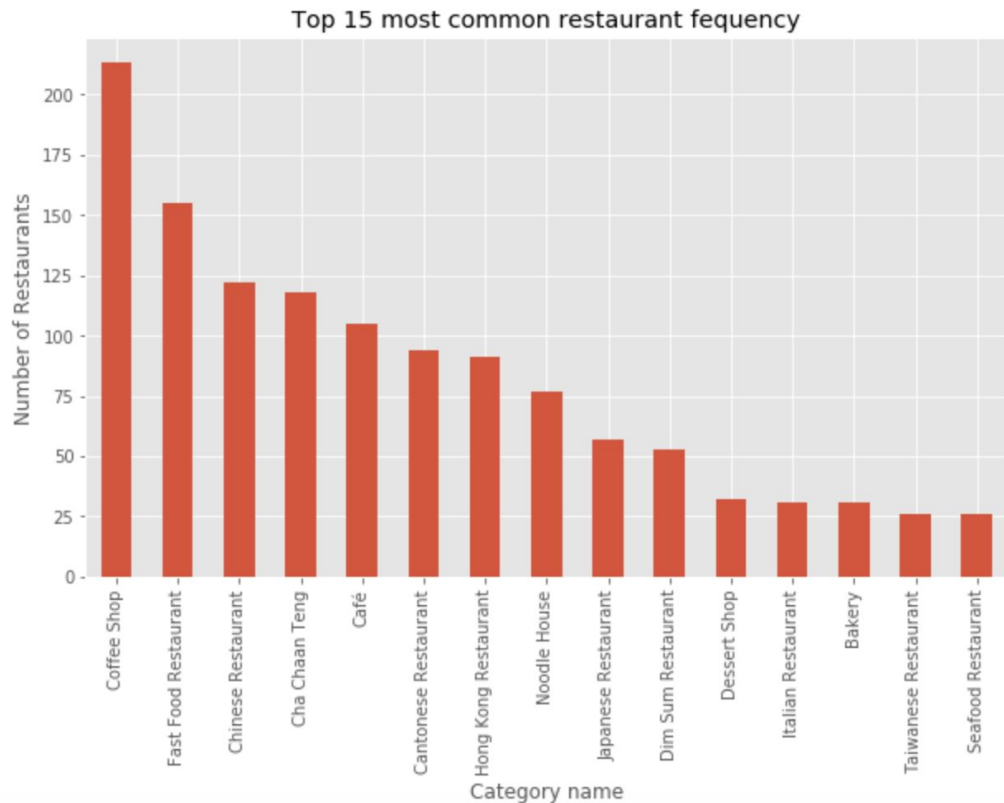


Map visualization



- Total 1749 data Points in the Map

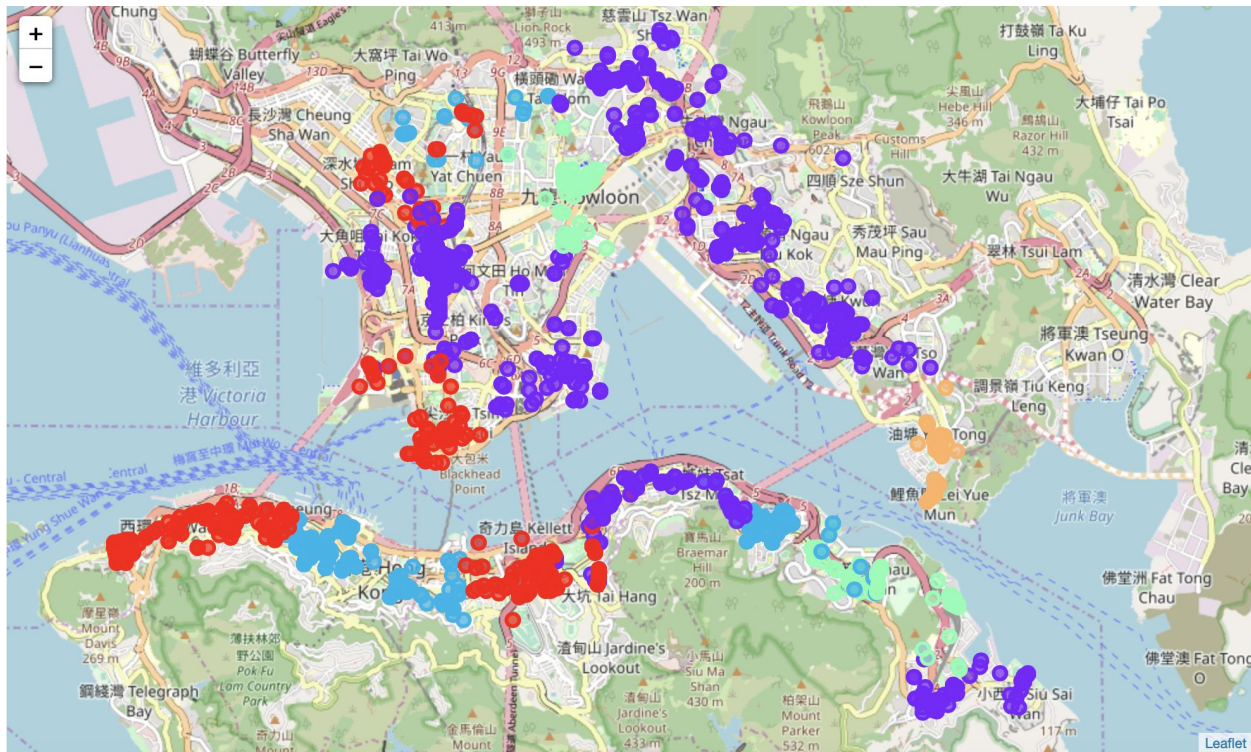
Distribution of the top 15 categories



- There are total 91 Categories
- This plot only listing 15 of them
- Coffee shop has highest frequency

Clustering

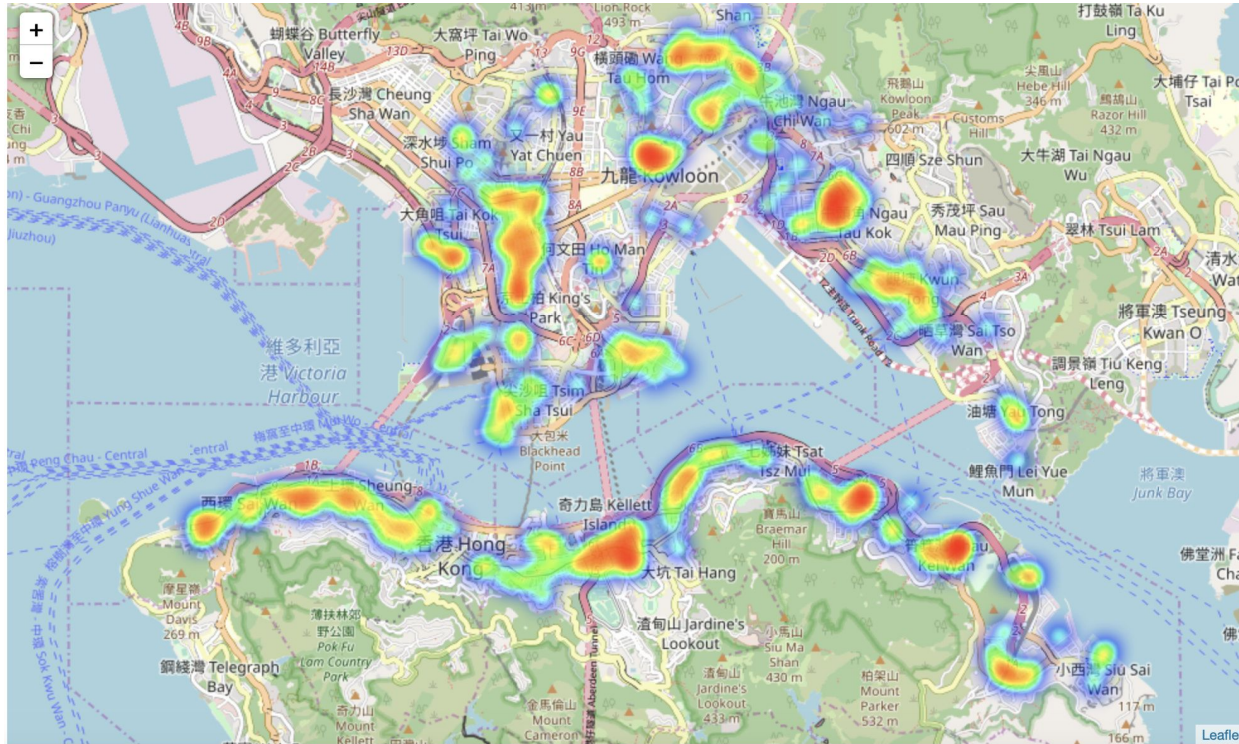
K-Means algorithm Clustering into 5 Clusters



- **Blue** Cluster 0 : Office Area with more western cuisine
- **Purple** Cluster 1: Fairly distributed Food Area
- **Red** Cluster 2 : Working or Shopping Area
- **Light Blue** Cluster 3: Asian Food Cluster
- **Yellow** Cluster 4 : Seafood intensive Area

Price levels analysis

Expensive restaurant
presenting in deep red color.



Insight : Red centers happens in the middle part of Mong Kok, Kowloon, Causeway Bay and the seaside area of hong kong Island.

Conclusion

- This report aiming to discover the restaurant cluster
- We can easily see the outcome of the clustering is very obvious and can be easily distinguished. It helps us find out those restaurant patterns.
- The price tier analysis helps us understand the price distribution of the restaurants.

