To open a Vietnamese restaurant in Osaka

CAPSTONE PROJECT – The Battle of Neighborhoods

APPLIED DATA SCIENCE

Introduction & Business Problem

- Osaka is also known as the “Nation’s Kitchen,” and is well known for delicious foods. You can also find a variety of cuisines in Osaka street, from Western to South East Asia.
- In recent years, the number of foreign people especially Vietnamese in Osaka keep raising every year.
- Besides, the majority of Japanese people nowadays, are searching for more deep flavor in South East Asian cuisine.
- The business question: *if someone is looking to open a new Vietnamese restaurant in Osaka area, where should they consider to open it?*

Data

We will need data from reliable sources for analysis. To understand our problem and quantify result we will use the following data.

- List of wards/boroughs in Osaka.

  Wikipedia page https://en.wikipedia.org/wiki/Osaka.

- Latitude and longitude coordinates of those wards.

  Using Python Geocoder package.

- Venue data, particularly data related to restaurant.

  Foursquare Developer Account at https://foursquare.com/

- Foreign population, particularly Vietnamese in Osaka.

  Osaka office website https://www.city.osaka.lg.jp

Methodology

- Collected the required data: location and type (category) of every restaurant within 2km from each Ward's center (24 wards).
- Identified Vietnamese restaurants (according to Foursquare categorization)
- Explored 'restaurant density' across different areas of Osaka
- Took into consideration locations with high Vietnamese density and locations without Vietnamese restaurants in radius of 200 meters.
- Created clusters (using k-means clustering) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

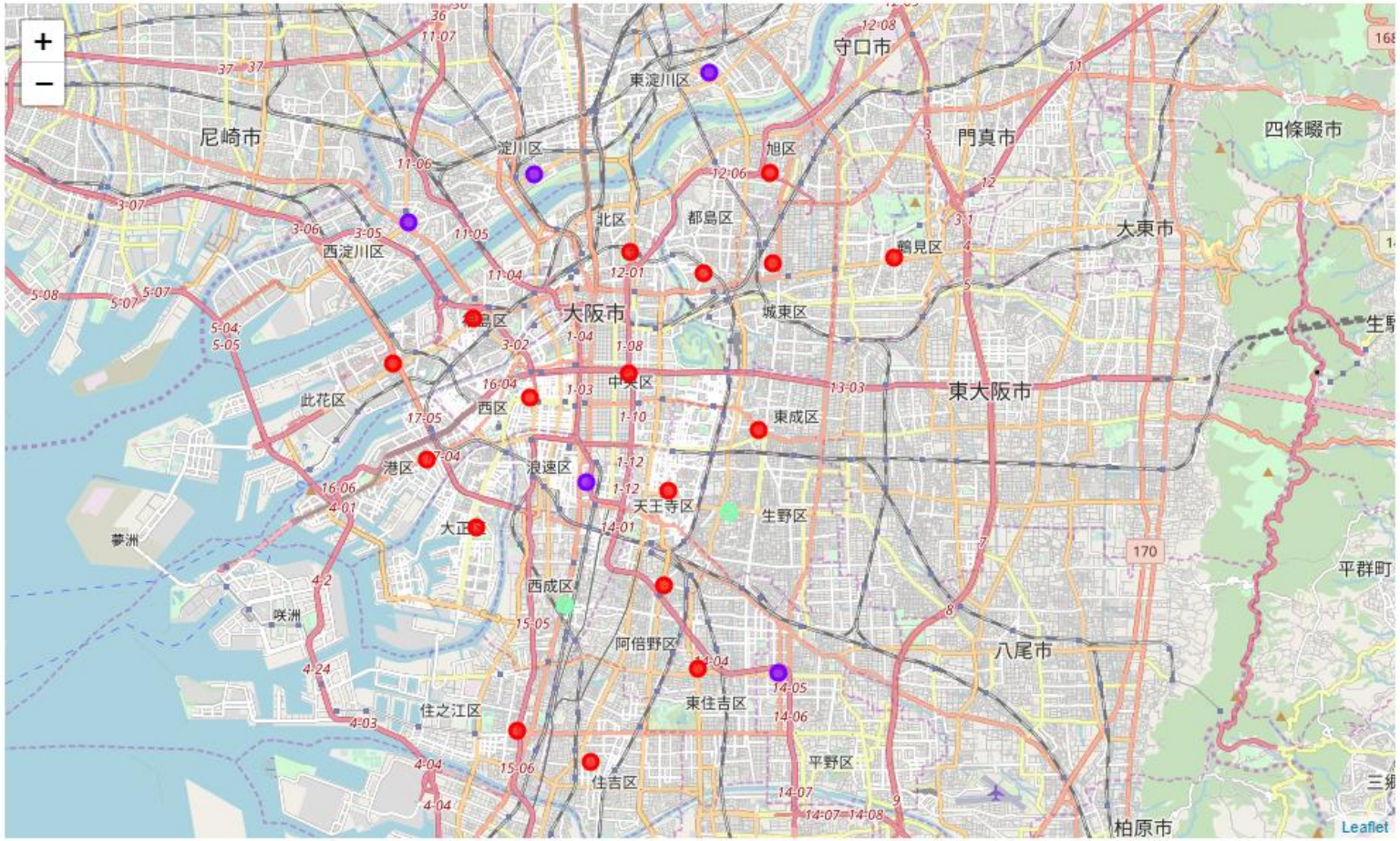
Result

The results from the k-means clustering show that we can categorize the wards into 3 clusters based on the frequency of occurrence for Vietnamese restaurant and number of Vietnamese:

- Cluster 0: wards with moderate number of Vietnamese restaurant and low density of Vietnamese
- Cluster 1: wards with low number of Vietnamese restaurant and moderate density of Vietnamese
- Cluster 2: wards with low number of Vietnamese restaurant and high density of Vietnamese

Result

The results of the clustering are visualized in the map below with cluster 0 in red color, cluster 1 in purple color, and cluster 2 in mint green color.



Conclusion and further discussion

- By calculating restaurant density distribution from Foursquare data we have first identified general wards that justify further analysis (生野区 and 西成区), and then generated extensive collection of locations which satisfy some basic requirements regarding density of Vietnamese.

- Taking into consideration additional factors
  - like attractiveness of each location
  - accessibly to train station
  - real estate availability, prices
  - social and economic dynamics of every ward