Explore the neighborhood of Taipei - Taiwan

Data Science Capstone Final Project

Introduction - Overview

- Taipei is the capital of Taiwan ROC
- Area: 271.8 km2
- Population 2.674 million
- Main language: Mandarin
- Divided into 12 administrative districts

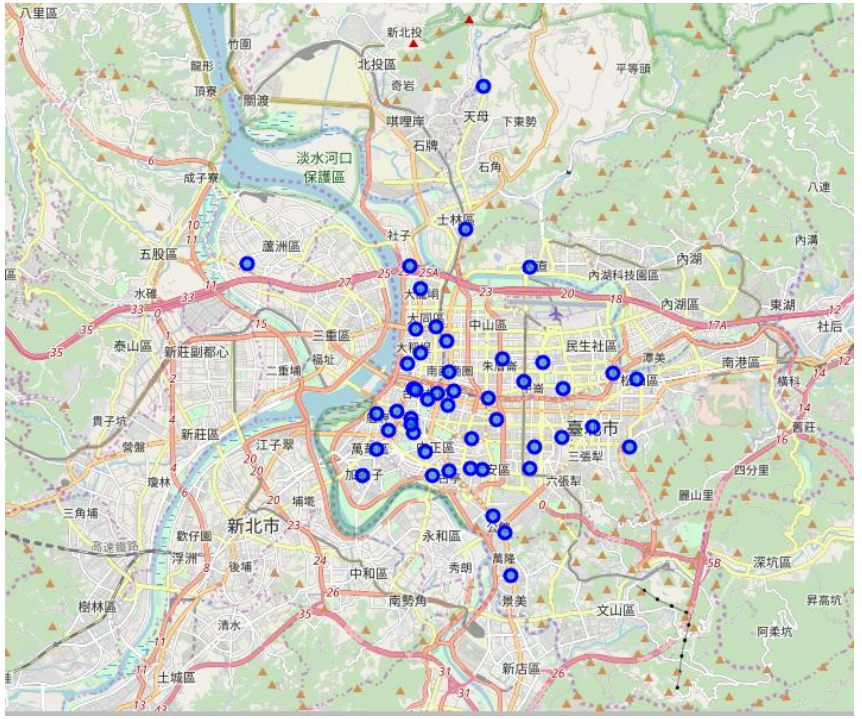
Introduction – Problem & Solutions

- An attractive but not-easy for non-speaking Mandarin people
- We will explore Taipei's inner cities using web scraping, Foursquare API and plotting
- Goal: explore distinguish characteristics of each areas
- Include two attempts:
  - For all features
  - Restaurant-only feature

Data

- 48 Taipei post office locations with latitude and longitudes that reprensents the neighborhood of different areas in the city.
- Foursquare explore method for each neighborhood

	Neighborhood	Latitude	Longitude
0	Taipei Beimen	25.04732	121.51179
1	Taipei Dongmen	25.03414	121.52856
2	Taipei Hanzhong Street	25.04133	121.50702
3	Taipei Xiyuan	25.04088	121.50121
4	Taipei Longshan	25.03658	121.50458



Process – Test with one sample neighborhood

```
dict_keys(['meta', 'response'])
61 venues were returned by Foursquare.
```

	name	categories	lat	lng
0	鄭記豬腳飯	Asian Restaurant	25.046989	121.511049
1	North Gate (台北府城北門)	Historic Site	25.047584	121.511179
2	Heritage Bakery & Cafe	Café	25.045171	121.511824
3	張家清真黃牛肉麵館 Chang's Halal beef Noodles	Noodle House	25.045718	121.510720
4	修圓素食	Vegetarian / Vegan Restaurant	25.046702	121.514228

- We test the Foursquare API with the first neighborhood in our Data Frame.
- The response seems very promising with 61 venues within a radius of 500 m from the location.

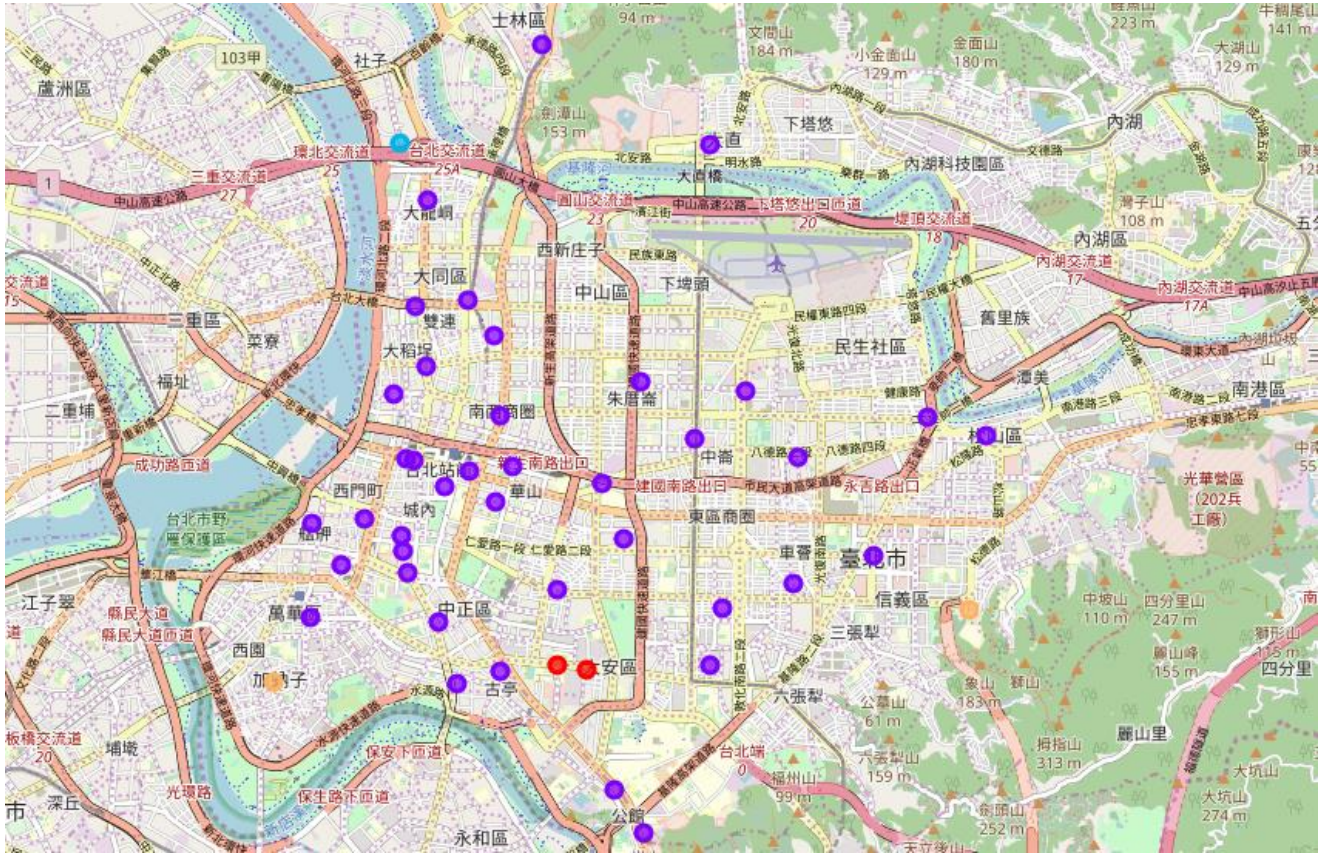
Solution – Explore all available neighborhoods

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Academia Historica	Hotel	Noodle House	Chinese Restaurant	Convenience Store	Café
1	Executive Yuan	Convenience Store	Café	Hotel	Japanese Restaurant	Seafood Restaurant
2	Legislative Yuan	Hotel	Noodle House	Chinese Restaurant	Art Gallery	Fast Food Restaurant
3	National Taiwan University	Other Nightlife	BBQ Joint	Bike Rental / Bike Share	Taiwanese Restaurant	Music Venue
4	Presidential Office Building	Noodle House	Garden	Convenience Store	Café	Szechuan Restaurant

- We apply the same method for all available neighborhoods and save it as a Data Frame consists of the most common venues

Solution – KNN clustering of the neighborhood

- We use KNN clustering to classify the neighborhood based on the common venues in each areas and make a report from the results.



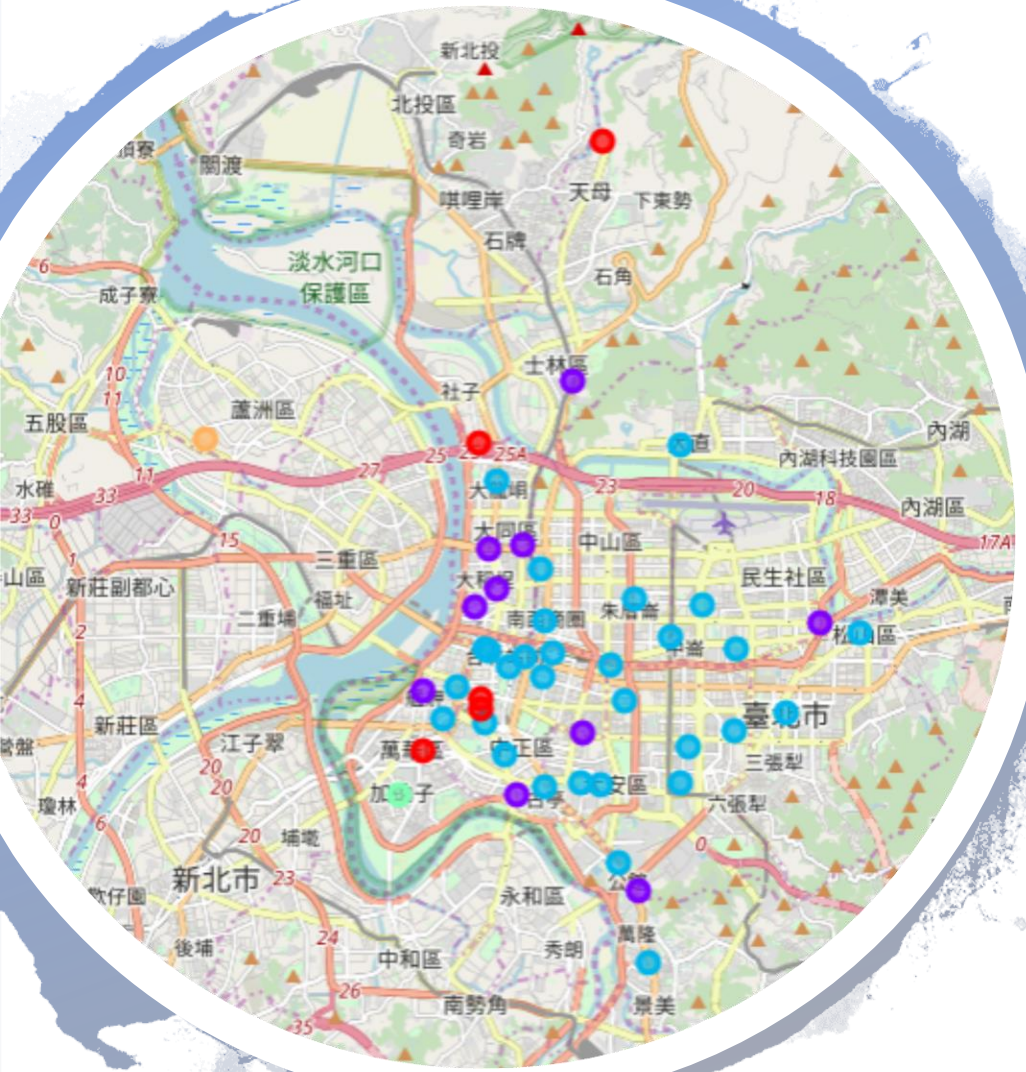
	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Group 1	Café	Noodle House	Sushi Restaurant	Chinese Restaurant	Tea Room
Group 2	Café	Convenience Store	Noodle House	Coffee Shop	Chinese Restaurant
Group 3	Intersection	Burger Joint	Bus Station	Chinese Restaurant	Yunnan Restaurant
Group 4	Fast Food Restaurant	Department Store	Burger Joint	Planetarium	Flea Market
Group 5	Convenience Store	Park	Coffee Shop	Bus Station	Fast Food Restaurant

Solution – KNN clustering of the neighborhood

- It seems like Taipei people love drinking and eating so much
- Group 1, 2 are inner cities filled with Cafe and restaurant
- Group 3,4,5 are outside areas with intersection, convenience store, parks and so on

It seems like the clustering is not very distinguish as Taipei is filled with Café and restaurant.

How about a clustering based only on Restaurant features?



Solution – KNN clustering of the neighborhood (Restaurant)

- I make a new Data Frame that only contains features including the word “Restaurant”

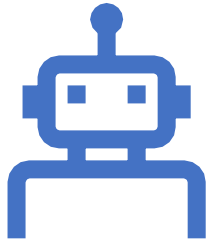
```
res = taipei_venues[taipei_venues["Venue Category"].str.contains("Restaurant")].reset_index(drop=True)
```

- The new clustering is more interesting!

Solution – KNN clustering of the neighborhood (Restaurant)

	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Group 1	Chinese Restaurant	Japanese Restaurant	Greek Restaurant	Japanese Restaurant	Japanese Curry Restaurant
Group 2	Taiwanese Restaurant	Japanese Restaurant	Hotpot Restaurant	Japanese Restaurant	Greek Restaurant
Group 3	Chinese Restaurant	Chinese Restaurant	Hotpot Restaurant	Asian Restaurant	Fast Food Restaurant
Group 4	Hotpot Restaurant	Yunnan Restaurant	Greek Restaurant	Japanese Restaurant	Japanese Curry Restaurant
Group 5	Fast Food Restaurant	Yunnan Restaurant	Greek Restaurant	Japanese Restaurant	Japanese Curry Restaurant

- Group 1, 2 are mostly inner cities areas filled with Chinese Restaurant or Taiwanese restaurant
- Group 3,4,5 contain more exotic restaurant like Hotpot, Yunnan or Greek Restaurant



Thank you and have fun with Data Science!