Capstone Project
The Battle of Neighborhoods

Author: WX Lai
Nov 2020

1. Introduction

This capstone project is designed for me (the author) personally.

I am relocating to Shanghai, China in February 2021 to pursue my career goal. As I am not sure exactly which part of Shanghai I will be staying, I thought it is a good idea to take this capstone project as an opportunity to explore Shanghai in advance.

In this project, I will explore Shanghai (area where I will be settled down) following the criteria below:

    a. Find an area near to metro station

    b. Find an area near to hospital (health is very important during this pandemic!)

From the above criteria, I should be able to narrow down the area I'd like to stay.

Then I will explore the nearby area to look for:

    c. Chinese restaurant for food

2. Data

Raw data of Metro station and its respective location will be scraped from web. Next, I will impose the location to map using Folium map.

Similarly, raw data of hospital location will be scraped from web. The same procedure will be repeated, whereby I will impose hospital location to the SAME map that has the metro location.

As such, I will be able to look for a location that is close to a metro station AND a hospital. Once the location is decided, Foursquare API function is employed to explore Chinese restaurant around the location. Foursquare API will also be used to check for customer rating. From customer rating, Chinese restaurants are ranked and sorted based on Foursquare API customer rating.

3. Methodology & Results

As mentioned previously, this capstone project served as a practice project to explore Shanghai in advance. I acknowledged that the purpose of exploring Shanghai can be done through other website or apps. However, as this the final stage of the course, I believe it is still a good idea to put my idea into practice.

Step 1 – Scrap from Wikipedia for metro station data

Source: https://en.wikipedia.org/wiki/List_of_Shanghai_Metro_stations

Raw data is available at Wikipedia. The information is scrapped, processed and cleaned. Then, the data is saved as a csv file for further use.

Service Routes[7]	Station Name[8]		Opened	Location	Platform Level	Type of Platform[9]	Transfers[10]
	English	Chinese (S)					
• •	Xinzhuang	莘庄	28 December 1996[11]	Minhang	At-grade	Side platform	5 Jinshan ⊙ Xinzhuang
• •	Waihuanlu	外环路			At-grade	Side platform	
• •	Lianhua Road	莲花路			At-grade	Side platform	
• •	Jinjiang Park	锦江乐园	28 May 1993 (former) [12][13] 12 December 1996 (current)[11]		At-grade	Side platform	
• •	Shanghai South Railway Station	上海南站	28 May 1993 (former)[12][13] 30 October 2004 (current)[11]		Underground	Island platform	3 Jinshan [a] ⊙ Shanghainan[b]

Figure 1: List of Shanghai metro station in Wikipedia

Out[168]:

	Metro Station	Station Name	Line	Latitude	Longitude
0	七宝	Qibao	Line 9	31.155365	121.348824
1	三林	Sanlin	Line 11	31.143311	121.512324
2	三林东	Sanlin East	Line 11	31.146525	121.523234
3	三门路	Sanmen Road	Line 10	31.313091	121.507995
4	上南路	Shangnan Road	Line 6	31.149112	121.506413

Figure 2: CSV file is imported into a data frame

Step 2 – Impose metro station to map

Nominatim is used to locate latitude and longitude of metro station. Folium map library then impose each location to a map.
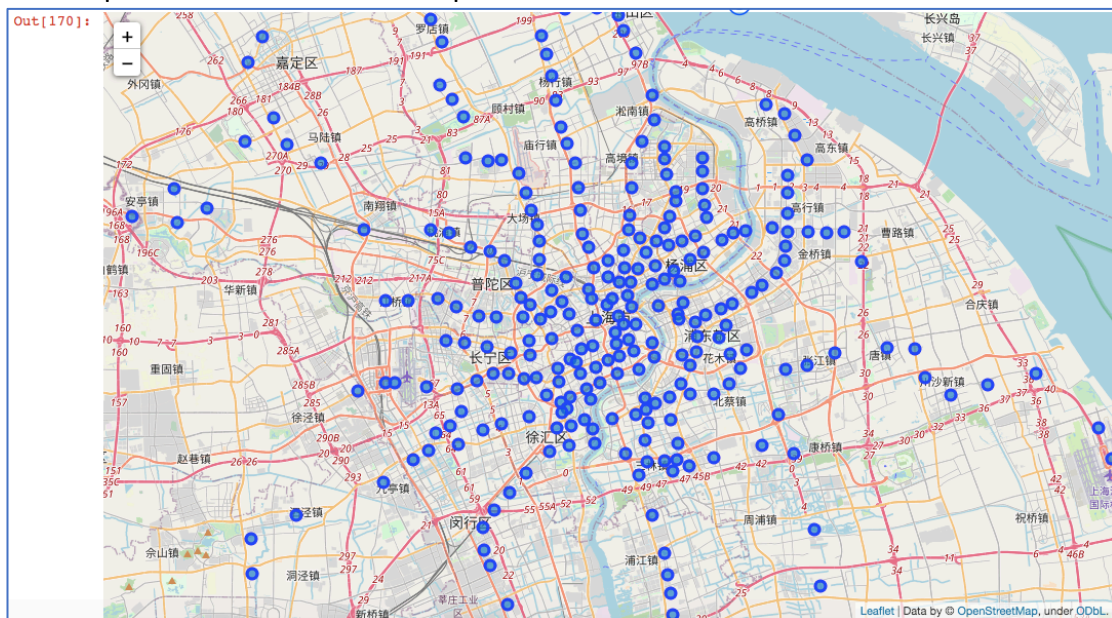

Figure 3: Blue dots are metro station in Shanghai

Step 3 – Raw data of hospital in Shanghai

Repeat Step 1 to obtain raw data of hospital location in Shanghai. Data is pre-processed, cleaned and save as csv file. CSV file is then imported to notebook as data frame.

```
In [171]: body = client_3695b1bee981          6.get_object(Bucket='courseraweek9capstoneproject-donotdelete-pr-gclzbwdl
          nvgigp',Key='SH_Hospital.csv')['Body']
          # add missing __iter__ method, so pandas accepts body as file-like object
          if not hasattr(body, "__iter__"): body.__iter__ = types.MethodType( __iter__, body )

          df_data_2 = pd.read_csv(body)
          df_data_2.head()
```

Out[171]:

	Hospital	Latitude	Longitude
0	Zhongshan Hospital	31.196552	121.456797
1	Huashan Hospital	31.212233	121.433824
2	Huadong Hospital	31.219098	121.440832
3	Children's Hospital of Fudan University	31.198649	121.454444
4	Red House Hospital (Fudan OBGYN Hospital)	31.259186	121.535251

Figure 4: Data frame of hospital in Shanghai

Step 4 – Location of hospitals on map

Nominatim is used to locate latitude and longitude of metro station. Folium map library then impose each location to a map.

Figure 5: Red dots are some of the hospitals in Shanghai
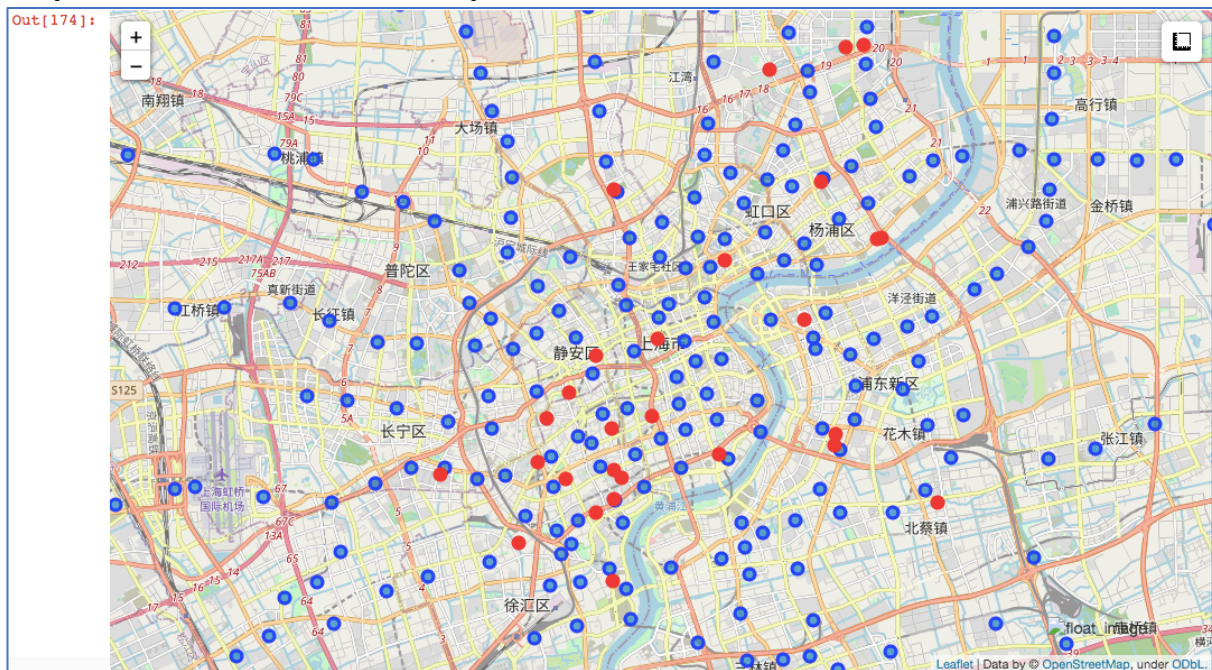
Step 5 – Combine into one map

Figure 6: Metro station + Hospital on one map

Step 6 – Choose a place that meet my requirement to stay

Based on the first two criteria that I set i.e. near to metro station and near to hospital, I am able to zoom the map the shortlist some possible candidate. After a thorough consideration, metro station "Jing An Temple" is chosen as the tentative place to stay when I relocate to Shanghai.
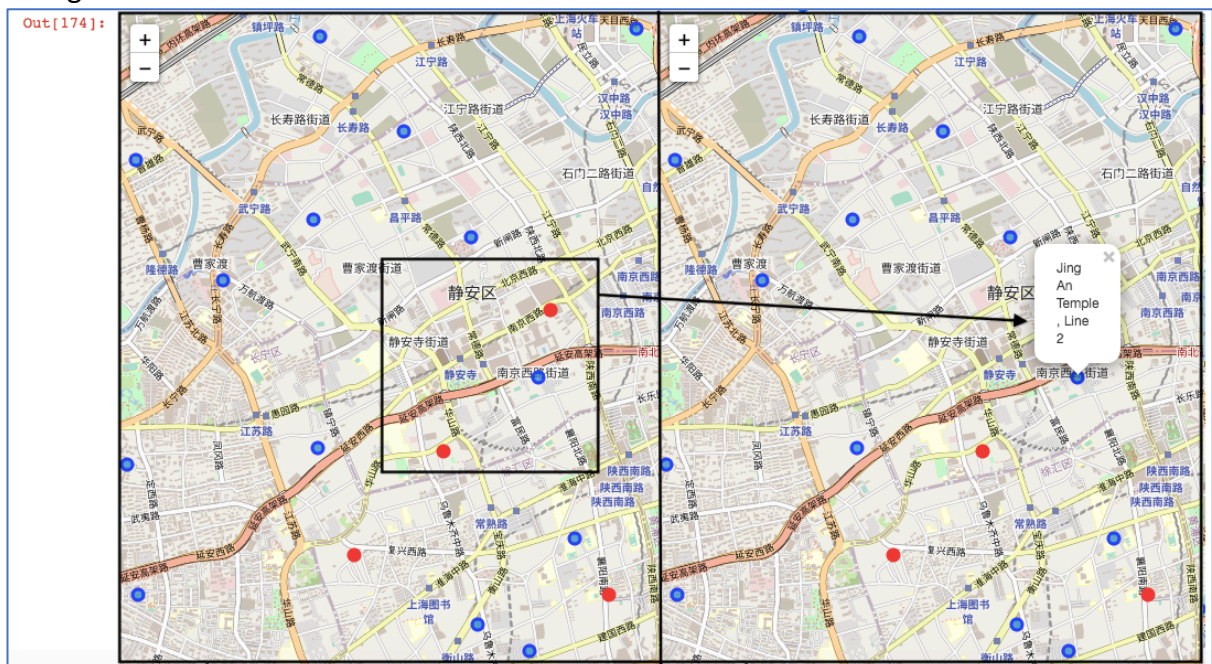

Figure 7: "Jing An Temple" is chosen as a tentative location

Step 7 – Explore Jing An temple and look for any Chinese Restaurant

Out[178]:

	id	name	categories	referralId	hasPerk	location.address	location.crossStreet	location.l
0	4b0588c2f964a52085d822e3	Ai Mei Chinese Restaurant (艾美轩)	[{'id': '4bf58dd8d48988d1f5931735', 'name': 'D...	v-1605341131	False	南京东路789号8F | 789 Nanjing Dong Lu	上海皇家艾美酒店内	31.236212
1	5386909f498e49f2a511ee9b	Chinese Breakfast	[{'id': '4bf58dd8d48988d143941735', 'name': 'B...	v-1605341131	False	NaN	NaN	31.234238
2	4b0588c1f964a52020d822e3	ROOF 325 Rooftop Restaurant & Bar	[{'id': '52e81612bcbc57f1066b79f9', 'name': 'M...	v-1605341131	False	325 Nanjing Xi Lu, 5/F, Shanghai Art Museum | ...	Huangpi Bei Lu | 近黄陂北路	31.232848
3	4cde7f61c4f6a35d802bca6c	Charme Restaurant (港丽餐厅)	[{'id': '52af3a7c3cf9994f4e043bed', 'name': 'C...	v-1605341131	False	268 Xizang Rd | 西藏中路268号	6/F Raffles City | 来福士广场	31.234173

Figure 8: List of Chinese Restaurant is obtained through Foursquare API command

Step 8 – Rank and sort according to customer's rating

Out[181]:

	categories	ratings
0	Ai Mei Chinese Restaurant (艾美轩)	5.6
1	Chinese Breakfast	0.0
2	ROOF 325 Rooftop Restaurant & Bar	7.1
3	Charme Restaurant (港丽餐厅)	5.6
4	TAI HING Restaurant (太兴餐厅)	5.5

Figure 9: Rank according to customer rating

4. Discussion

I. Foursquare API command returns Chinese restaurant around Jing An Temple

Code from Coursera lab is taken as guide to generate a list of Chinese restaurant around 500m of Jing An Temple.

```
In [175]: #Explore neighbourhood by setting Jing An Temple as centre point (my first location to consider)
          address = 'Jing An Temple'
          geolocator = Nominatim(user_agent="SH_explorer")
          uot_location = geolocator.geocode(address)
          uot_latitude = location.latitude
          uot_longitude = location.longitude
          print('The geograpical coordinate of Jing An Temple are {}, {}.'.format(latitude, longitude))

          The geograpical coordinate of Jing An Temple are 31.2322758, 121.4692071.

In [176]: #let's get 50 venues that are in Jing An Temple within a radius of 500 meters.
          LIMIT = 50 # limit of number of venues returned by Foursquare API
          radius = 500 # define radius
          search_query = 'Chinese Restaurant'
          url = 'https://api.foursquare.com/v2/venues/search?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limi
          t={}'.format(
              CLIENT_ID,
              CLIENT_SECRET,
              latitude,
              longitude,
              VERSION,
              search_query,
              radius,
              LIMIT)

          url
```

Figure 10: Code to employ Foursquare location data

The results generated contains many unwanted and unnecessary data. Those data was filtered to show only the important message such as Name, Address, Longitude and Latitude.

Out[179]:

	name	categories	address	crossStreet	lat	lng
0	Ai Mei Chinese Restaurant (艾美轩)	Dim Sum Restaurant	南京东路789号8F | 789 Nanjing Dong Lu	上海皇家艾美酒店内	31.236212	121.470667
1	Chinese Breakfast	Breakfast Spot	NaN	NaN	31.234238	121.463569
2	ROOF 325 Rooftop Restaurant & Bar	Modern European Restaurant	325 Nanjing Xi Lu, 5/F, Shanghai Art Museum | ...	Huangpi Bei Lu | 近黄陂北路	31.232848	121.466538

Figure 11: List of Chinese restaurant (after data is filtered)

II.   Visualization of Chinese restaurant

Recall that my third requirement in this study is to find Chinese restaurant near Jing An Temple. Code from Coursera lab was taken as a guide to check customer ratings on Foursquare. For easier visualization of Chinese restaurant returned from Foursquare API were plotted.

```
In [183]:  import matplotlib as mpl
           import matplotlib.pyplot as plt

           df_data_3.plot(kind='bar', figsize=(8, 6))
           plt.title('Ratings of Chinese Restaurant near Jing An Temple') # add title to the plot

           plt.xticks(np.arange(1,24), df_data_3['categories'], rotation=90)

           plt.xlabel('Restaurant') # add to x-label to the plot
           plt.ylabel('Ratings') # add y-label to the plot


           plt.show()
```

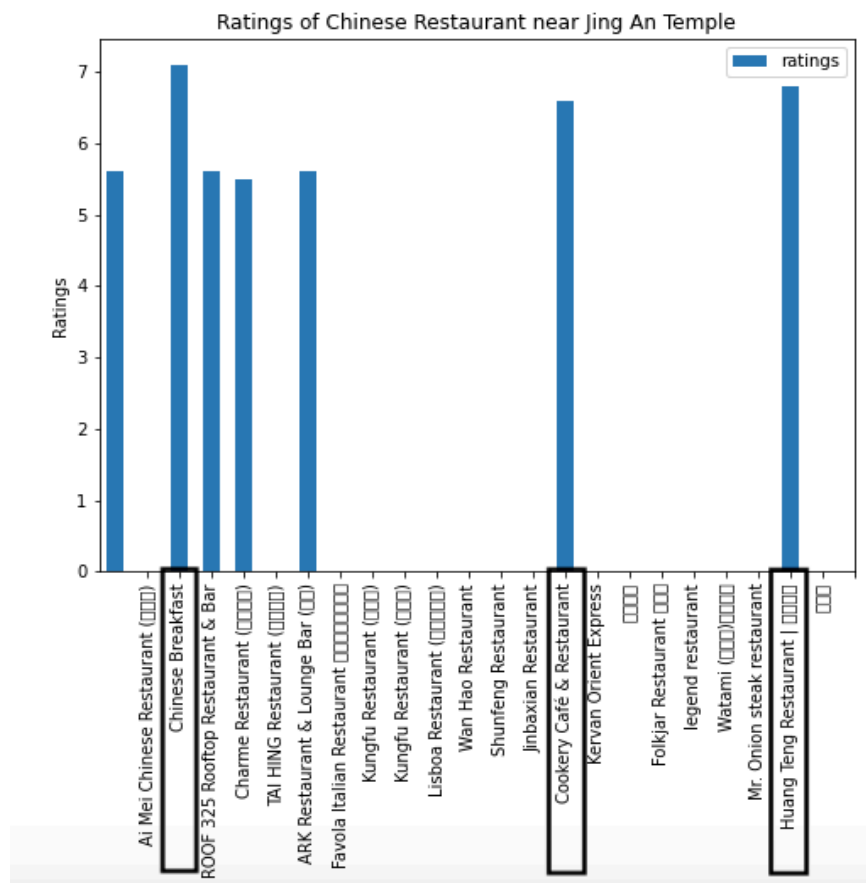*Figure 12: Find customer ratings from Foursquare API*

Figure 13: Customer ratings on Chinese restaurant

III. **Many restaurant has not yet been rated in Foursquare**

One observation from this project is that most of the Chinese restaurants return no ratings when Foursquare command is run and asked for ratings data. In csv file, restaurants with no ratings on Foursquare were given 0.0 score. In the bar chart prepared to visualise customer ratings, some restaurant has no bar. This does not indicate that the restaurant has zero ratings given by customer. Instead, no bar in the chart simply means that the restaurant have not been given any rating in Foursquare yet.

5. Conclusion

Criteria	Description	Status
Near to metro station	Location of metro station is imposed to a map	Fulfilled and completed. Metro Station: Jing An Temple
Near to hospital	Location of hospital is imposed to a map (individually and combined).	Fulfilled and completed. Hospital: Huadong Hospital and Huashan Hospital
Explore Chinese restaurant	Foursquare API location data is used to explore Chinese restaurant around	Listed and ranked. Restaurant with the highest ranking: 1st – Chinese Breakfast 2nd – Cookery Café & Restaurant 3rd – Huang Teng Resturant