Finding Out The Best Restaurants Around MRT(Mass Rapid Transit) In Taiwan

Wesley Wang

May 23, 2021

1. Introduction

1.1 Background

Taiwan is becoming the international country and visited by a lot of foreigners each year. That's why the need to find reliable and popular restaurants is getting important. And if we can also add other information, such as population, that will quickly become another visual tool for restaurant owners to make decisions about where they should run their business.

1.2 Problem

For those foreigners, the tools are familiar with could be Foursquares. If we can collect information from Foursquare, that would be more accessible for foreigners to use it. And the transportation in Taiwan is Mass Rapid Transit, if we can combine the two things together, it's will be convenient for users. For restaurant owners, where to run their business is always one of the most crucial decisions for them. If there is a visual tool for them to see population in each area, that can help them to make a better decision-making.

1.3 Interest

When everyone pays a visit to Taiwan first time, they can easily find out the best restaurants around MRT(Mass Rapid Transit). And restaurant owners can also find their crucial visual information about population.

2. Data acquisition and cleaning

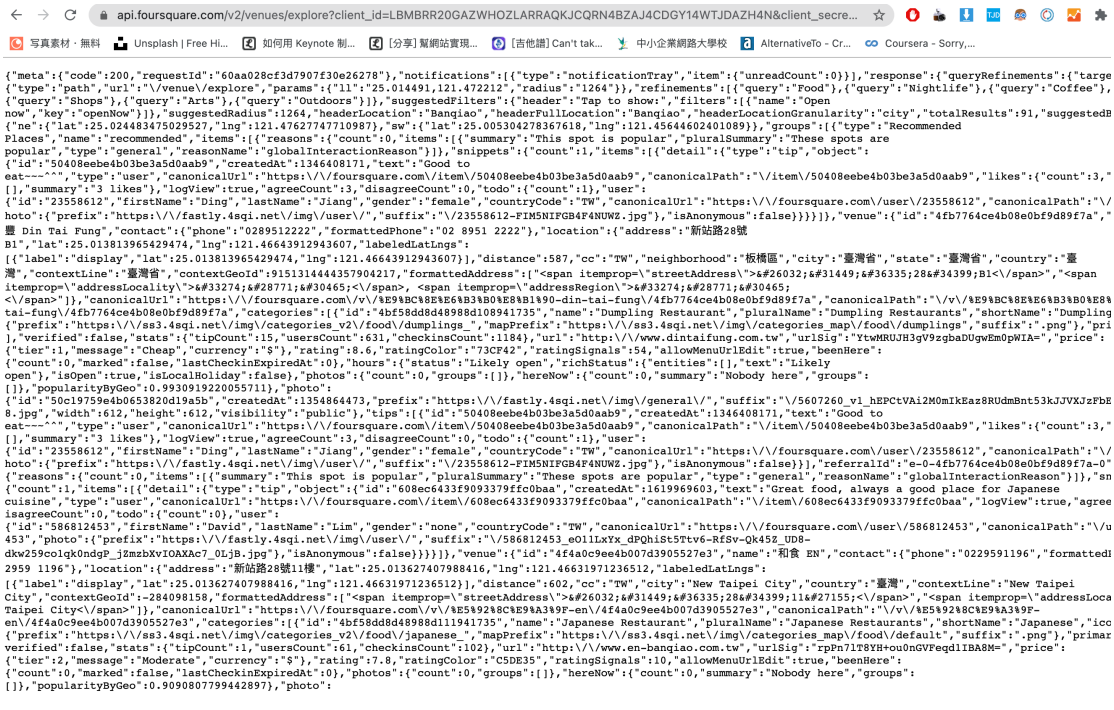
2.1 Data sources

The 1st step is to find out each locations/stops of MRT system. And those information is open for everyone, just need to download from it's official website (https://taipeicity.github.io/traffic_realtime/) as follows:

	Names	Latitudes	Longtitudes
0	新北產業園區站出口	25.061548	121.459926
1	幸福站出口2	25.049545	121.460593
2	幸福站出口1	25.050126	121.460146
3	新埔民生站出口	25.026125	121.466839
4	板橋站出口5	25.015502	121.464312
5	板橋站出口4	25.015105	121.464938
6	板新站出口	25.014491	121.472212
7	中原站出口	25.008410	121.484159
8	橋和站出口	25.004413	121.490461
9	中和站出口	25.001981	121.496052
10	景平站出口	24.991968	121.516553
11	秀朗橋站出口2	24.990549	121.525051
12	秀朗橋站出口1	24.990418	121.525740
13	十四張站出口	24.984467	121.527701
14	松山站出口5	25.050557	121.578130
15	松山站出口4	25.049799	121.577971
16	松山站出口3	25.049799	121.577464
17	松山站出口2	25.050171	121.576244
18	松山站出口1	25.050171	121.577022
19	南京三民站出口4	25.051588	121.564710

Luckily, we can find out it has latitudes and longitudes for each stops.

2nd step is to find out restaurant information from Foursquares. As long as we register, we can access necessary information from it's website as following json files :



3rd step is to make population information and blend into the above data. And this information can be easily accessible from government`s website as follows:

	site_id	people_total	area	population_density	Latitudes	Longtitudes
0	台北市松山區	204193.0	9.29	21985.0	25.0542	121.5639
1	台北市信義區	220021.0	11.21	19631.0	25.0409	121.5720
2	台北市大安區	307631.0	11.36	27077.0	25.0249	121.5434
3	台北市中山區	227387.0	13.68	16619.0	25.0792	121.5427
4	台北市中正區	158014.0	7.61	20772.0	25.0421	121.5199
5	台北市大同區	126043.0	5.68	22185.0	25.0627	121.5113
6	台北市萬華區	187076.0	8.85	21133.0	25.0263	121.4970
7	台北市文山區	271806.0	31.51	8626.0	24.9929	121.5713
8	台北市南港區	120297.0	21.84	5507.0	25.0312	121.6112
9	台北市內湖區	285795.0	31.58	9050.0	25.0689	121.5909
10	台北市士林區	283459.0	62.37	4545.0	25.0950	121.5246
11	台北市北投區	253319.0	56.82	4458.0	25.1152	121.5150

2.2 Data cleaning

This can be divided into 3 parts since our data is from 3 different sources.

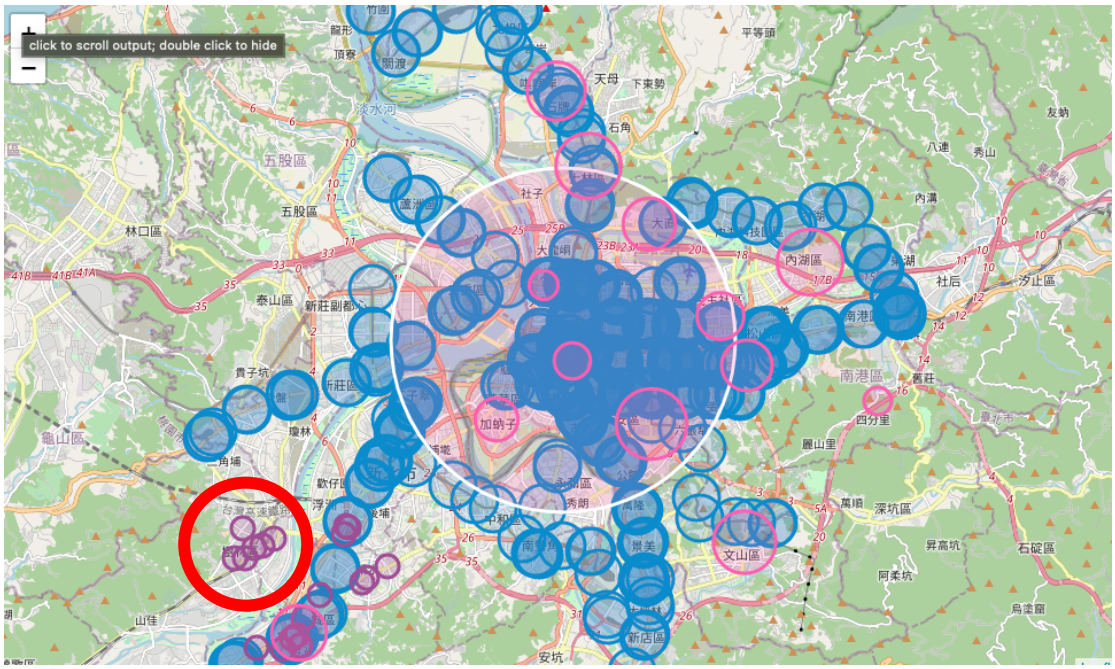
Part 1 – This could be easier because it's downloadable Json file. We just need to classify it and make it as a dataframe. Pandas can process it quickly.

Part 2 – This is much harder than the previous one because the information from the website is a little bit messy (you can see from the above picture). In order to access the important information, I need to find out and try it layer by layer. So try several layers, we could finally find out the information (restaurants' latitudes and longitudes).

Part 3 – Population information would be easier to clean. After remove the unnecessary words, such as N, comma and E, this part of work will be done, quickly.

2.3 Feature selection & choices of visual tools

In order to present the best visual population information for restaurant owners, we need to use K-means and find the best part on the map as follows:



Pink circles mean the population density and purple circles mean restaurant locations(only present an area to make it clear). And the blue circles mean the areas people can reach after getting off from MRT system. That restaurant owners can easy the know where is not suitable to run their business(Red circle).

3. Exploratory Data Analysis

3.1 Calculation and relationship between circles

To make a population information more visual, I need to calculate the population and density, and make it on the map. After several adjustments and trials, I think the best present is like the above picture as you can see. And I also need to describe what does that mean for each circle:

Pink circles: the population density. The bigger circle, the more population.

Purple circles: restaurant locations.

Blue circles: the areas people can reach after getting off from MRT system.

Red circle: That restaurant owners can easy the know where is not suitable to run their business, because its population and public transportation is not that convenient.

4. Predictive Modeling

4.1 K-means to find out the best and worse places for restaurants owners

With K-means classification, we can easily find out the while circle is the most populated-place and its public transportation is also the most convenient. That is why white circle is the suitable place for owners to run their restaurant. And the red circle as you can see on the map, its population and public transportation is not that

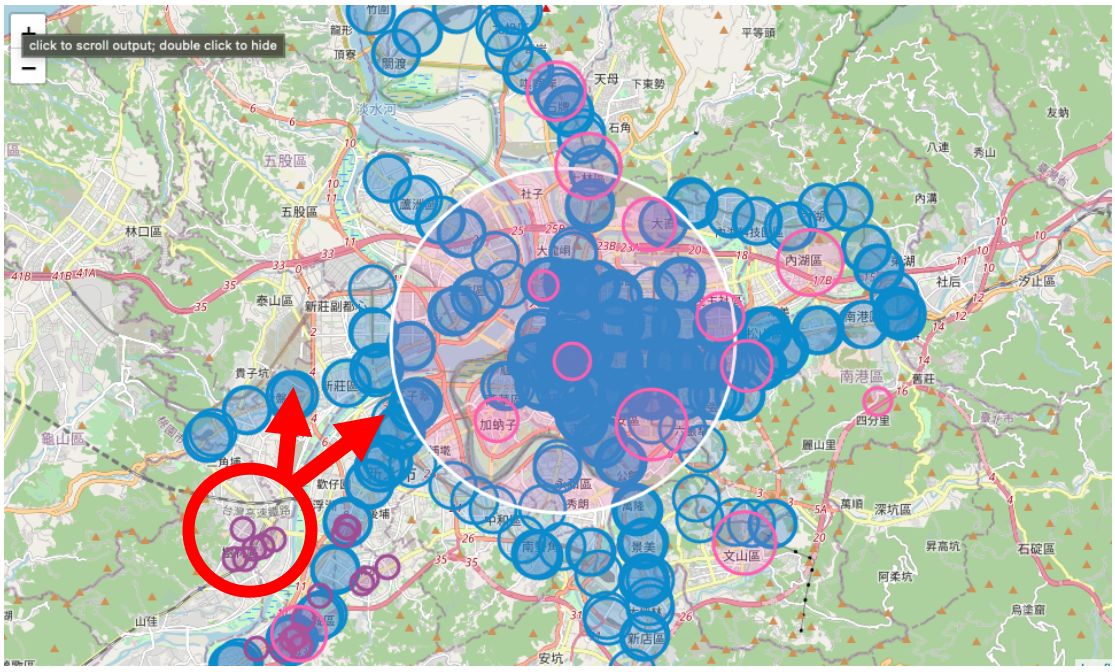
convenient. That is why restaurant owners may lose their money if they run their restaurant in this area.

5. Conclusions

From the picture below, I recommend 2 ways for owners to make the better decision-making.

1st way is to move their restaurants closer to MRT system (blue circles) because MRT system brings a lot of people and foreigners. This will generate business rewards accordingly.

2nd way is to move the restaurants to the downtown (white circle) because that is the center of the city and is also the center of the public transportation. That is why there are so many blue circles and pink circles around this area.



6. Future directions

In the future, this map can also add more information, such as hotel information, other business revenues, business types in order to see the whole picture about those areas. The more information, the better decision-making we can do.