

A nighttime photograph of a city skyline, likely Taipei, Taiwan. The image shows numerous illuminated skyscrapers and buildings against a dark blue sky with scattered clouds. The lights from the buildings create a warm, golden glow. The image is split into two main sections: the left side shows a closer view of the city lights, and the right side shows a wider view of the skyline.

IBM Applied Data Science Capstone 2021

**The battle of Neighbourhoods,
Taipei, Taiwan**

MARCH 4

Authored by: Catherine

Introduction

OPPORTUNITY IN TAIWAN

Despite the global economic impact of the Covid-19 Pandemic, Taiwan (Republic of China) has maintained robust economic growth. According to the Chung-Hua Institution for Economic Research (CIER), Taiwan's gross domestic product (GDP) has grown 2.38 percent in 2020 and will continue to grow by 3.73 percent in 2021.¹ Taiwan's economic freedom score is 77.1, making its economy the 11th freest in the 2020 Index of Economic Freedom jointly issued by the Heritage Foundation and the Wall Street Journal². These aspects provide positive economic and market growth in 2021.

Taiwan has a geographical area of 36,193 square kilometers, with a population of about 23,476,640. It leads to an average population density of 649 people per square kilometer. Making it the 17th most densely populated country in the world.³ Its capital city, Taipei, and New Taipei City are the two the most densely populated areas in the territory. In Taipei, 10,000 people live in every square meter. With such high density, and a culture of overworking, eating delicious food becomes the most preferred leisure activity of Taiwanese and resulting in the blossom of the restaurant.

This report will provide a glance for stakeholders who wish to open a new venue in northern Taiwan, where the most business opportunity gathered.

“Certain places seem to exist mainly because someone has written about them.” – William Faulkner

Problem and Discussion

When it comes to open a new venue, different factors need to take into consideration. However, economic activities hugely rely on the population, human activity. Does the area have enough labor and demand for business? What about the economic activities in that area, the diversification? Diversity is the basis of economic stability and growth. A diverse economic environment will be more ideal for a restaurant business. This report attempts to provide some semblance to these questions based on the analytics of the data collected from public platforms. The questions have been narrowed to the follows:

1. Which districts have a greater population density?
2. Which districts have a better economic diversity?
3. Which district has the most potential to open a venue?
4. What business can be a good idea to open.

¹ <https://focustaiwan.tw/business/202012180014>

² <https://www.heritage.org/index/country/taiwan#:~:text=Taiwan's%20economic%20freedom%20score%20is,monetary%20freedom%20and%20trade%20freedom.>

³ https://en.wikipedia.org/wiki/Demographics_of_Taiwan

DATA

DATA COLLECTION

To understand better the business activity in Greater Taipei Area, we need to collect not only basic demography data, but also geographic data:

Taipei and New Taipei geographic coordinate data can be found on Taiwan opening government data platform: <https://data.gov.tw/en>

New Taipei City Municipal administration data:

https://en.wikipedia.org/wiki/New_Taipei_City

Taipei City Municipal administration data: <https://en.wikipedia.org/wiki/Taipei>

Taipei and New Taipei Geojson data:

<https://github.com/littlebtc/geodata4appy/tree/master/simplified>

Exploratory data for venues in Taipei and new Taipei City: This part of data will be acquired by using the Foursquare API.

DATA PROCESSING

In this report, the `pandas.read_html()` function has been used as a scraper to obtain the municipal administration data of Taipei and New Taipei City from Wikipedia. After reading the files into their Dataframes and cleaning, the population density has been calculated. After the calculation, the two tables (Taipei and New Taipei) have been combined by using `append()` function. The geographic coordinate data then eventually joined with the Dataframe using the `join()` function. Result in the following dataframe:

	District	Postal Code	Longitude	Latitude	Population	Area_sq.km	Density_sq.km
0	Zhongzheng	100	121.519884	25.032405	162549	7.6071	21368.0
1	Datong	103	121.513042	25.063424	131029	5.6815	23062.0
2	Zhongshan	104	121.538160	25.069699	231286	13.6821	16904.0
3	Songshan	105	121.557588	25.059991	209689	9.2878	22577.0
4	Daan	106	121.543445	25.026770	312909	11.3614	27541.0
5	Wanhua	108	121.497986	25.028590	194314	8.8522	21951.0
6	Xinyi	110	121.571670	25.030621	229139	11.2077	20445.0
7	Shilin	111	121.550847	25.125467	290682	62.3682	4661.0
8	Beitou	112	121.517799	25.148068	257922	56.8216	4539.0
9	Neihu	114	121.592383	25.083706	287726	31.5787	9111.0
10	Nangang	115	121.609757	25.036009	122296	21.8424	5599.0
11	Wenshan	116	121.573608	24.988579	275433	31.5090	8741.0
12	Wanli	207	121.643931	25.175725	22634	63.3766	357.0
13	Jinshan	208	121.605264	25.217146	22273	49.2132	453.0
14	Banqiao	220	121.457968	25.011865	554008	23.1373	23944.0
15	Xizhi	221	121.654699	25.073313	196150	71.2354	2754.0
16	Shenkeng	222	121.620062	24.997675	23614	20.5787	1147.0
17	Shiding	223	121.647228	24.947141	7857	144.3498	54.0
18	Ruifang	224	121.823202	25.098129	40922	70.7336	579.0
19	Pingxi	226	121.757882	25.026071	4872	71.3382	68.0
20	Shuangxi	227	121.832982	24.996984	9233	146.2484	63.0
21	Gongliao	228	121.918247	25.024856	12858	99.9734	129.0
22	Xindian	231	121.531656	24.930390	300283	120.2255	2498.0
23	Pinglin	232	121.724223	24.910971	6503	170.8350	38.0
24	Wulai	233	121.541481	24.788243	6182	321.1306	19.0

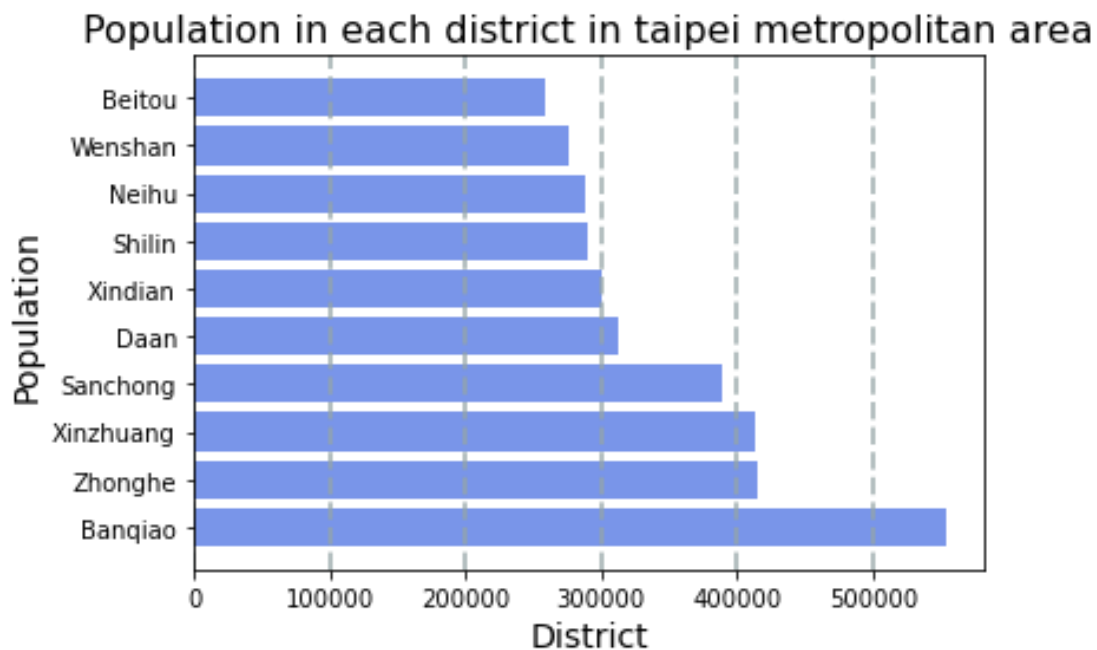
The Taipei and New Taipei City geojson data then been imported and merged for the visualization.

	AREA	PERIMETER	ROADLC_	ROADLC_ID	X	Y	ID	District	SECT_CODE	LIE_NAME	...	X_1	Y_1	CAT_ID	T_NAME	ALIAS	ADMIT_ID	ADMIV	ADMIV_ID	TM2X	TM
0	1.281570e+06	7845.48419	2528	0	305274.35890	2.774201e+06	1.0	Shimen	6300100	莊敬里	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	2.593486e+05	2158.12973	3369	0	305536.83356	2.772927e+06	2.0	Shimen	6300100	東榮里	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	1.345430e+05	1906.84462	3333	0	305835.64963	2.772981e+06	3.0	Shimen	6300100	三民里	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	1.151300e+05	1552.06662	3297	0	306387.36347	2.773104e+06	4.0	Shimen	6300100	新益里	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	9.981312e+04	1351.73546	3320	0	306152.31677	2.773055e+06	5.0	Shimen	6300100	富興里	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...
1027	7.845744e+07	46965.23981	1034	0	313433.34357	2.746730e+06	NaN	Wulai	NaN	NaN	...	313433.34357	2.746730e+06	A	李義里	李義里	29	李義里	005	313433.34357	2.746730e
1028	2.508367e+07	21792.93549	1035	0	302600.57815	2.746744e+06	NaN	Wulai	NaN	NaN	...	302600.57815	2.746744e+06	A	信賢里	信賢里	29	信賢里	004	302600.57815	2.746744e
1029	1.899392e+08	65153.74761	1036	0	301646.28110	2.738600e+06	NaN	Wulai	NaN	NaN	...	301646.28110	2.738600e+06	A	福山里	福山里	29	福山里	003	301646.28110	2.738600e
1030	1.910931e+05	2277.95429	241	0	313759.25905	2.774317e+06	NaN	Xizhi	NaN	NaN	...	313759.25905	2.774317e+06	A	湖蓮里	湖蓮里	11	湖蓮里	049	313684.97344	2.774301e
1031	2.599665e+05	3035.23962	80	0	295082.16516	2.784668e+06	NaN	Tamsui	NaN	NaN	...	295082.16516	2.784668e+06	A	幸福里	幸福里	10	幸福里	040	295070.51721	2.784387e

METHODOLOGY

Descriptive statistics analyse is used to answer the population density questions. The following chart illustrates the districts that have the great population in Greater Taipei Area. They are Banqiao, Zhonghe, Xinzhuang, Sanchong, Daan, Xindian, Shilin, Neihu, Wenshan, Beitou.

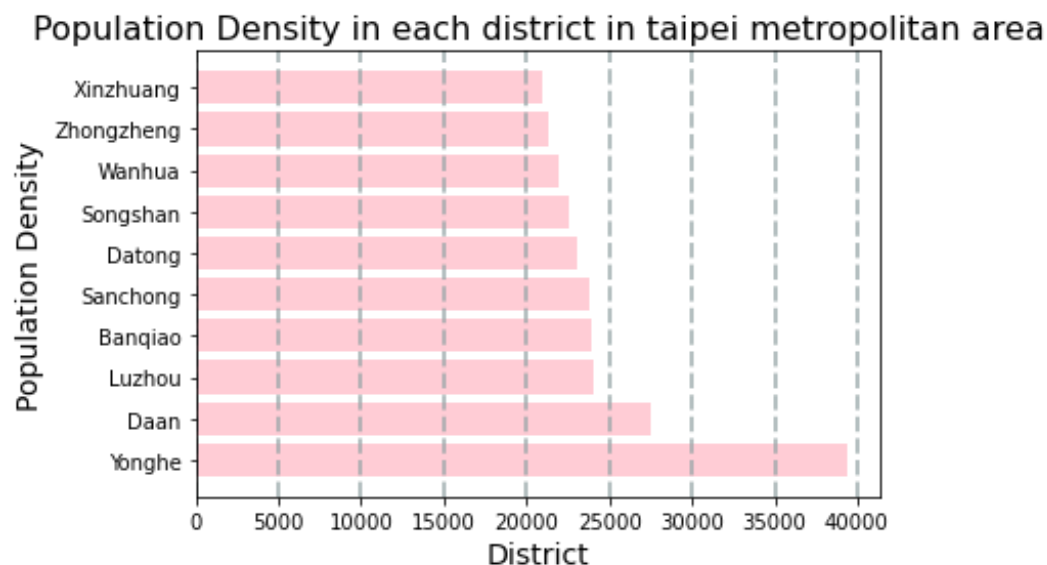
	District	Postal Code	Longitude	Latitude	Population	Area_sq.km	Density_sq.km
14	Banqiao	220	121.457968	25.011865	554008	23.1373	23944.0
26	Zhonghe	235	121.493674	24.990880	414356	20.1440	20570.0
32	Xinzhuang	242	121.436753	25.035830	413443	19.7383	20946.0
31	Sanchong	241	121.487098	25.062817	388386	16.3170	23803.0
4	Daan	106	121.543445	25.026770	312909	11.3614	27541.0
22	Xindian	231	121.531656	24.930390	300283	120.2255	2498.0
7	Shilin	111	121.550847	25.125467	290682	62.3682	4661.0
9	Neihu	114	121.592383	25.083706	287726	31.5787	9111.0
11	Wenshan	116	121.573608	24.988579	275433	31.5090	8741.0
8	Beitou	112	121.517799	25.148068	257922	56.8216	4539.0



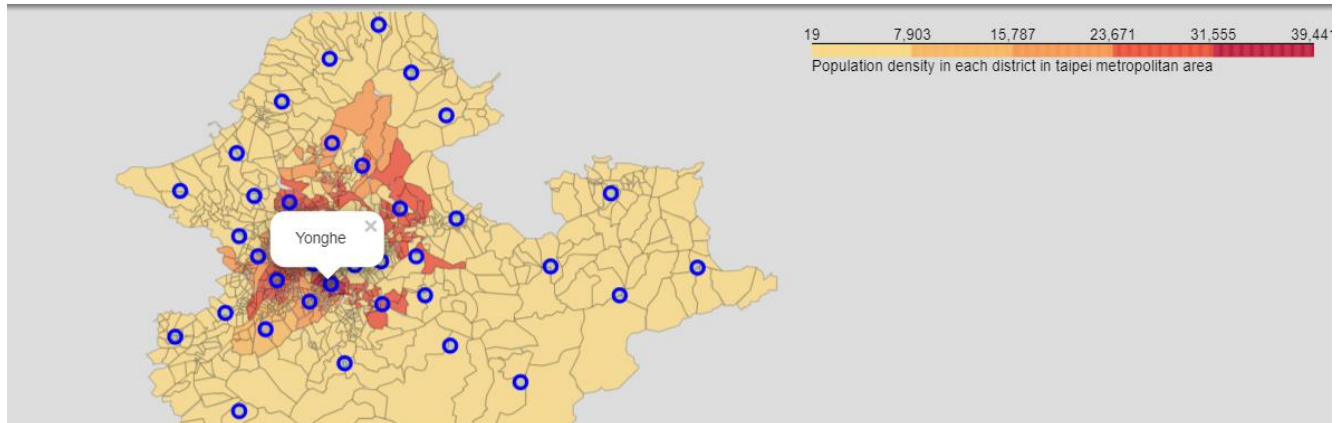
The population density

From the chart below we can see that the most densely populated district Yonghe has almost 40,000 people living every square kilometre, followed by Daan, Luzhou, Banqiao...etc.

	District	Postal Code	Longitude	Latitude	Population	Area_sq.km	Density_sq.km
25	Yonghe	234	121.516745	25.008102	225353	5.7138	39440.0
4	Daan	106	121.543445	25.026770	312909	11.3614	27541.0
35	Luzhou	247	121.471246	25.089272	200055	8.3210	24042.0
14	Banqiao	220	121.457968	25.011865	554008	23.1373	23944.0
31	Sanchong	241	121.487098	25.062817	388386	16.3170	23803.0
1	Datong	103	121.513042	25.063424	131029	5.6815	23062.0
3	Songshan	105	121.557588	25.059991	209689	9.2878	22577.0
5	Wanhua	108	121.497986	25.028590	194314	8.8522	21951.0
0	Zhongzheng	100	121.519884	25.032405	162549	7.6071	21368.0
32	Xinzhuang	242	121.436753	25.035830	413443	19.7383	20946.0



The statistics number is a good indication, however, we would like to see the population density in a geographical framework. Folium is an ideal tool for visualisation. The geopy library is used to get the latitude and longitude values of Taipei City. Then we define a user_agent, in our case, it's named tw_explorer. In the final step, a choropleth map is generated by using folium library and the geojson data.



From the thematic map above, it's easy to spot the population is gathered in the west part of northern Taiwan and it's highly centralised.

Foursquare API then be used to interact with the Foursquare platform where we can explore the districts by sending the GET request. Our starting point is Zhongzhen, within a radius of 500 meters, we've retrieved 32 venues in Zhongzhen area.

	name	categories	lat	lng
0	Kinfen Braised Pork Rice (金峰魯肉飯)	Taiwanese Restaurant	25.032194	121.518534
1	Chiang Kai-Shek Memorial Hall (中正紀念堂)	Monument / Landmark	25.034555	121.521835
2	樂田麵包屋 Gakuden Boulangerie	Bakery	25.032757	121.517534
3	National Theater (國家戲劇院)	Theater	25.035197	121.518188
4	臻味赤肉胡椒餅 烤地瓜	Bakery	25.033022	121.518246

Then we create a function to repeat the same process to all other districts in Greater Taipei area.

	District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Zhongzheng	25.032405	121.519884	Kinfen Braised Pork Rice (金峰魯肉飯)	25.032194	121.518534	Taiwanese Restaurant
1	Zhongzheng	25.032405	121.519884	Chiang Kai-Shek Memorial Hall (中正紀念堂)	25.034555	121.521835	Monument / Landmark
2	Zhongzheng	25.032405	121.519884	樂田麵包屋 Gakuden Boulangerie	25.032757	121.517534	Bakery
3	Zhongzheng	25.032405	121.519884	National Theater (國家戲劇院)	25.035197	121.518188	Theater
4	Zhongzheng	25.032405	121.519884	臻味赤肉胡椒餅 烤地瓜	25.033022	121.518246	Bakery
...
336	Luzhou	25.089272	121.471246	三民路肯德基	25.087815	121.471632	Fast Food Restaurant
337	Luzhou	25.089272	121.471246	肯德基 KFC (蘆洲三民餐廳)	25.087043	121.471592	Fried Chicken Joint
338	Luzhou	25.089272	121.471246	FamilyMart_蘆洲民族店	25.091190	121.474040	Deli / Bodega
339	Luzhou	25.089272	121.471246	捷運三民高中站 MRT Sanmin Senior High School Station	25.085741	121.473080	Metro Station
340	Sanzhi	25.231599	121.515558	八達溪水車	25.233907	121.514404	Park

Use groupby and count ()function to know how many venues were returned for each district.

District	District Latitude	District Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Banqiao	31	31	31	31	31	31
Daan	38	38	38	38	38	38
Datong	25	25	25	25	25	25
Jinshan	2	2	2	2	2	2
Luzhou	4	4	4	4	4	4
Nangang	4	4	4	4	4	4
Neihu	25	25	25	25	25	25
Sanchong	6	6	6	6	6	6
Sanzhi	1	1	1	1	1	1
Shenkeng	2	2	2	2	2	2
Shilin	1	1	1	1	1	1
Shuangxi	1	1	1	1	1	1
Shulin	1	1	1	1	1	1
Songshan	53	53	53	53	53	53
Tucheng	1	1	1	1	1	1
Wanhua	7	7	7	7	7	7
Wanli	1	1	1	1	1	1
Wenshan	23	23	23	23	23	23
Xinyi	22	22	22	22	22	22
Xinzhuang	9	9	9	9	9	9
Xizhi	1	1	1	1	1	1
Yingge	2	2	2	2	2	2
Yonghe	17	17	17	17	17	17
Zhonghe	8	8	8	8	8	8
Zhongshan	10	10	10	10	10	10
Zhongzheng	32	32	32	32	32	32

In Greater Taipei Area, there are 96 unique venue categories. According to above, we know that Songshan, Dann, Zhongzeng, Banqiao, Datong and Neihu have more venues than other districts.

For a deeper exploration, we use the one hard encoding to make the representation of categorical data to be more expressive. Then calculate the mean of the frequency of occurrence of each category. So we can easily get the top 5 most frequent venues in all the districts.

----Songshan----

	venue	freq
0	Café	0.22
1	Convenience Store	0.10
2	Dumpling Restaurant	0.06
3	Taiwanese Restaurant	0.06
4	Bakery	0.06

----Daan----

	venue	freq
0	Café	0.13
1	Noodle House	0.11
2	Coffee Shop	0.11
3	Taiwanese Restaurant	0.08
4	Chinese Breakfast Place	0.05

----Zhongzheng----

	venue	freq
0	Café	0.12
1	Dumpling Restaurant	0.09
2	Breakfast Spot	0.09
3	Japanese Restaurant	0.06
4	Theater	0.06

----Banqiao----

	venue	freq
0	Hotel	0.10
1	Dessert Shop	0.07
2	Hotpot Restaurant	0.07
3	Café	0.07
4	Bookstore	0.07

----Datong----

	venue	freq
0	Taiwanese Restaurant	0.22
1	Convenience Store	0.13
2	Noodle House	0.09
3	Dessert Shop	0.09
4	Tea Room	0.09

----Neihu----

	venue	freq
0	Convenience Store	0.13
1	Italian Restaurant	0.09
2	Chinese Restaurant	0.09
3	Japanese Restaurant	0.09
4	Noodle House	0.09

Create a function to put the venues into a pandas dataframe. Following are the top 10 most common venues in Greater Taipei area.

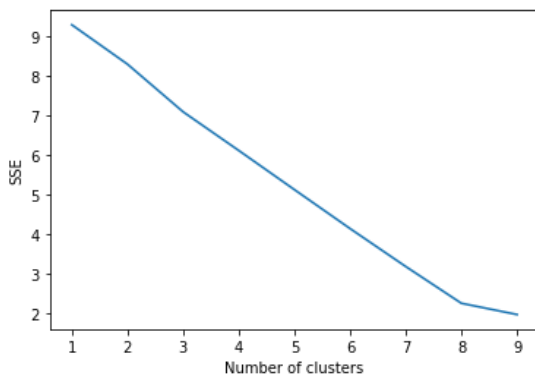
	District	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Banqiao	Japanese Restaurant	Bookstore	Taiwanese Restaurant	Hotel	Bus Stop	Café	Chinese Restaurant	Multiplex	Fast Food Restaurant	Convenience Store
1	Daan	Café	Coffee Shop	Noodle House	Taiwanese Restaurant	Chinese Breakfast Place	Farmers Market	Brewery	Ice Cream Shop	Dessert Shop	Cocktail Bar
2	Datong	Taiwanese Restaurant	Convenience Store	Food Truck	Asian Restaurant	Dessert Shop	Coffee Shop	Noodle House	Gym / Fitness Center	Bus Station	Flea Market
3	Jinshan	Farm	Hotel	Dumpling Restaurant	Concert Hall	Convenience Store	Cosmetics Shop	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant
4	Luzhou	Fried Chicken Joint	Deli / Bodega	Fast Food Restaurant	Metro Station	Women's Store	Discount Store	Concert Hall	Convenience Store	Cosmetics Shop	Department Store
5	Nangang	Convenience Store	Bus Station	Supermarket	Sake Bar	Breakfast Spot	Women's Store	Dive Bar	Cosmetics Shop	Deli / Bodega	Department Store
6	Neihu	Convenience Store	Coffee Shop	Chinese Restaurant	Japanese Restaurant	Italian Restaurant	Breakfast Spot	Bookstore	Noodle House	Metro Station	Market
7	Sanchong	Convenience Store	Chinese Restaurant	Restaurant	Breakfast Spot	Women's Store	Coffee Shop	Cosmetics Shop	Deli / Bodega	Department Store	Dessert Shop
8	Sanzhi	Park	Women's Store	Dive Bar	Concert Hall	Convenience Store	Cosmetics Shop	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant
9	Shenkeng	Taiwanese Restaurant	Farmers Market	Women's Store	Dive Bar	Concert Hall	Convenience Store	Cosmetics Shop	Deli / Bodega	Department Store	Dessert Shop
10	Shiding	Tea Room	Women's Store	Dumpling Restaurant	Concert Hall	Convenience Store	Cosmetics Shop	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant

Let's see what are the most popular and less popular business.

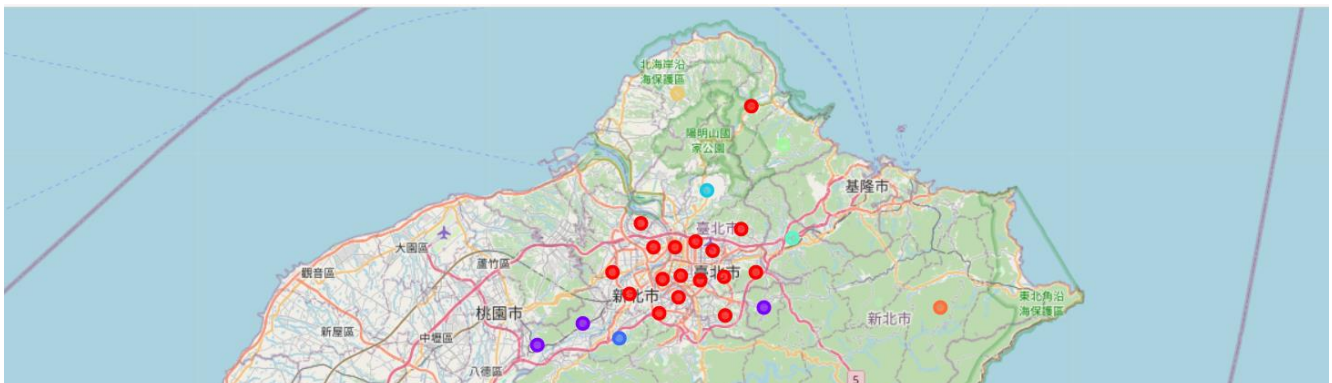
	Venue Category	Venue		Venue Category	Venue
28	Convenience Store	31			
20	Café	29	48	Indie Movie Theater	1
26	Coffee Shop	22	51	Jewelry Store	1
92	Taiwanese Restaurant	17	53	Karaoke Bar	1
69	Noodle House	13	55	Latin American Restaurant	1
14	Breakfast Spot	12	92	Women's Store	1
22	Chinese Restaurant	11			
36	Dumpling Restaurant	8			
7	Bakery	8			
57	Italian Restaurant	7			

From above, we can see that convenience stores are all the investors or business owners favorite, following by café and coffee shops. On the other hand, women's stores, Latin American restaurants, karaoke bars, Jewelry stores, and indie movie theaters are less favorable.

In order to quickly discover and segment the economic activities in Greater Taipei Area. K-means is used to group the districts based on their venue categories. The elbow method is used to determine the optimal K. We get a clear elbow, k=8, which we used to run K-means clustering.



Finally, we plotted the clusters on the map.



Following are the clusters:

Cluster 1

	District	Area_sq.km	Density_sq.km	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Zhongzheng	7.6071	21368.0	0.0	Café	Dumpling Restaurant	Breakfast Spot	Theater	Bakery	Chinese Breakfast Place	Concert Hall	Latin American Restaurant	Korean Restaurant	Noodle House
1	Dalong	5.6815	23062.0	0.0	Taiwanese Restaurant	Convenience Store	Dessert Shop	Gym / Fitness Center	Supermarket	Chinese Restaurant	Night Market	Coffee Shop	Snack Place	Noodle House
2	Zhongshan	13.6821	16904.0	0.0	Seafood Restaurant	Convenience Store	Fish Market	Scenic Lookout	Breakfast Spot	Market	Farmers Market	Sushi Restaurant	Hotpot Restaurant	BBQ Joint
3	Songshan	9.2878	22577.0	0.0	Café	Convenience Store	Taiwanese Restaurant	Dumpling Restaurant	Italian Restaurant	Bakery	Furniture / Home Store	Chinese Restaurant	Clothing Store	Pastry Shop
4	Daan	11.3614	27541.0	0.0	Café	Coffee Shop	Noodle House	Taiwanese Restaurant	Chinese Breakfast Place	Farmers Market	Breakfast Spot	Bookstore	Dessert Shop	Cocktail Bar
5	Wanhua	8.8522	21951.0	0.0	Convenience Store	Coffee Shop	Noodle House	Bakery	Supermarket	Fast Food Restaurant	Dive Bar	Concert Hall	Deli / Bodega	Department Store
6	Xinyi	11.2077	20445.0	0.0	Scenic Lookout	Park	Café	Convenience Store	Mountain	Chinese Restaurant	Southern / Soul Food Restaurant	Bakery	Coffee Shop	Trail
9	Neihu	31.5787	9111.0	0.0	Coffee Shop	Japanese Restaurant	Convenience Store	Italian Restaurant	Bookstore	Noodle House	Metro Station	Department Store	Pharmacy	Pizza Place
10	Nangang	21.8424	5599.0	0.0	Convenience Store	Bubble Tea Shop	Supermarket	Sauna Bar	Jewelry Store	Women's Store	Discount Store	Coffee Shop	Concert Hall	Deli / Bodega
11	Wenshan	31.5090	8741.0	0.0	Convenience Store	Coffee Shop	Chinese Restaurant	Asian Restaurant	Café	Japanese Restaurant	Italian Restaurant	Bus Stop	Playground	Breakfast Spot
13	Jinshan	49.2132	453.0	0.0	Snack Place	Hotel	Farm	Dumpling Restaurant	Coffee Shop	Concert Hall	Convenience Store	Deli / Bodega	Department Store	Dessert Shop
14	Banqiao	23.1373	23944.0	0.0	Bookstore	Café	Hotel	Taiwanese Restaurant	Hotpot Restaurant	Stadium	Bus Stop	Deli / Bodega	Movie Theater	Chinese Restaurant
25	Yonghe	5.7138	39440.0	0.0	Coffee Shop	Convenience Store	Noodle House	Breakfast Spot	Hotpot Restaurant	Massage Studio	Italian Restaurant	Paper / Office Supplies Store	Dumpling Restaurant	Ice Cream Shop
26	Zhonghe	20.1440	20570.0	0.0	Coffee Shop	Pet Store	Taiwanese Restaurant	Convenience Store	Bowling Alley	Park	Fish Market	Food Court	Cocktail Bar	Fried Chicken Joint
31	Sanzhong	16.3170	23803.0	0.0	Convenience Store	Chinese Restaurant	Café	Restaurant	Breakfast Spot	Women's Store	Discount Store	Concert Hall	Deli / Bodega	Department Store
32	Xinzhuan	19.7383	20946.0	0.0	Pet Store	Furniture / Home Store	Climbing Gym	Convenience Store	Food Court	Café	Bus Station	Metro Station	Diner	Concert Hall
35	Luzhou	8.3210	24042.0	0.0	Coffee Shop	Fried Chicken Joint	Noodle House	Metro Station	Dive Bar	Concert Hall	Convenience Store	Deli / Bodega	Department Store	Dessert Shop

Cluster 2

	District	Area_sq.km	Density_sq.km	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
16	Shenkeng	20.5787	1147.0	1.0	Café	Taiwanese Restaurant	Farmers Market	Women's Store	Dive Bar	Coffee Shop	Concert Hall	Convenience Store	Deli / Bodega	Department Store
29	Shulin	33.1288	5564.0	1.0	Food Truck	Women's Store	Dumpling Restaurant	Coffee Shop	Concert Hall	Convenience Store	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant
30	Yingge	21.1248	4162.0	1.0	Asian Restaurant	Food Truck	Women's Store	Dumpling Restaurant	Concert Hall	Convenience Store	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant

Cluster 3

	District	Area_sq.km	Density_sq.km	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
27	Tucheng	29.5578	8074.0	2.0	Garden	Women's Store	Dumpling Restaurant	Coffee Shop	Concert Hall	Convenience Store	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant

Cluster 4

	District	Area_sq.km	Density_sq.km	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
7	Shilin	62.3682	4661.0	3.0	Bus Stop	Women's Store	Dumpling Restaurant	Concert Hall	Convenience Store	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant	Diner

Cluster 5

	District	Area_sq.km	Density_sq.km	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
15	Xizhi	71.2354	2754.0	4.0	Breakfast Spot	Women's Store	Dumpling Restaurant	Concert Hall	Convenience Store	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant	Diner

Cluster 6

	District	Area_sq.km	Density_sq.km	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
12	Wanli	63.3766	357.0	5.0	Seafood Restaurant	Women's Store	Dive Bar	Coffee Shop	Concert Hall	Convenience Store	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant

Cluster 7

	District	Area_sq.km	Density_sq.km	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
39	Sanzhi	65.9909	355.0	6.0	Park	Women's Store	Dive Bar	Coffee Shop	Concert Hall	Convenience Store	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant

Cluster 8

	District	Area_sq.km	Density_sq.km	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
20	Shuangxi	146.2484	63.0	7.0	Women's Store	Dumpling Restaurant	Coffee Shop	Concert Hall	Convenience Store	Deli / Bodega	Department Store	Dessert Shop	Dim Sum Restaurant	Diner

DISCUSSION

This report aims to provide an insight for stakeholders who are interested in opening a new venue in Greater Taipei City. As we can see in the report, the most densely populated district is Yonghe, followed by Daan, Luzhou, Banqiao and Sanchong. Among them, Dann and Banqiao have a good diversity of economic activities. They are both included in cluster 1 and centralized in most populated area. That indicates that Greater Taipei Area is a highly competitive place to open a venue.

Two ideas can be discussed. First, we go with long tail. Offer an obscure product/service. Opening a karaoke bar or a Latin American restaurant. Or a business that has a mature market but with a stable demand like a breakfast place or a café. To be even more innovative, maybe a Latin American brunch place with a decent menu can attract some attention. Taiwanese people like dining out and discover new restaurants. Now, about the location, cluster 1 has the most stable market with enough population to develop business. Banqiao and Daan both have some Taiwanese, Chinese hotpot, fastfood restaurants , but Latin American isn't on the list. Therefore, we can have our first step result. Opening a Latin American brunch place in Banqiao or Daan.

LIMITATION

We had our first conclusion. However, there exist still a lot of factors that could affect the result. There are still some aspects need to be considered. The more detailed demographic data is defiantly needed, the income, education level, age range...etc. Also the rent, transportation facility, and the lifestyle of the population.

Not every business/venue has been listed on Foursquare. An official statistic data is recommended if available.

CONCLUSION

It's fun to use the tools that IBM offers on Coursera and it provides a gateway for a new career. In this report, we practiced the knowledge that we learned from Coursera's courses. We learned how to construct the methodology, how to process data, how to use machine learning models to categorize data or predicate results. In this report, the K-

means model has been used to cluster the districts. However, a different cluster method can be considered for further exploration.