EXPLORING HO CHI MINH CITY

1. Introduction

Ho Chi Minh City (HCMC) is the biggest city in Vietnam. It locates in the South and has a population of 9m people, contributing from 25% to 30% of the nation's GDP in 2019.

The city has 19 urban districts and 5 rural districts, which have an area of 494 sq km and 1,600 sq km. Annually, this place attracts 40m tourists, both local and international tourists, thanks to its beautiful historic buildings, nice coffee shops and restaurants.

This report will give tourists as well as local people a view on what kinds of venues are common in which districts so that they can have some information on where to go and what to do when visiting HCMC.
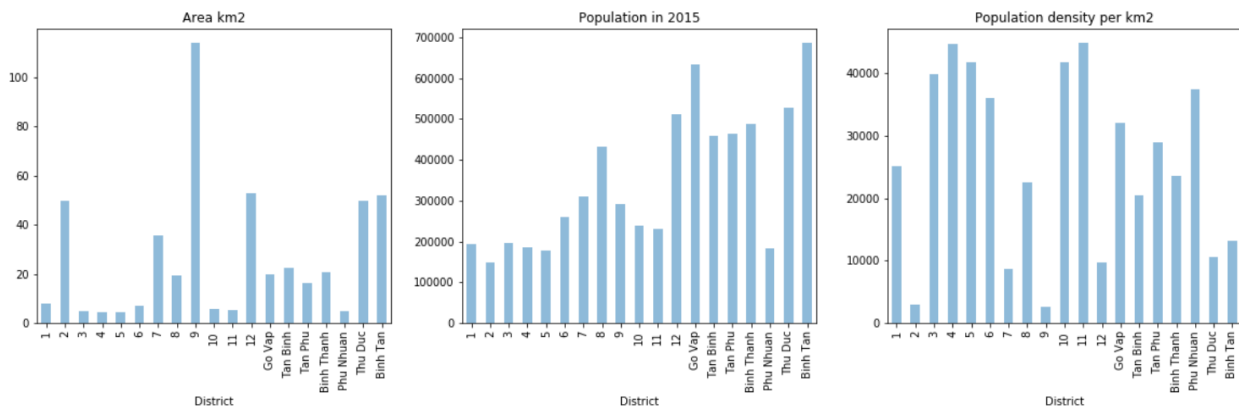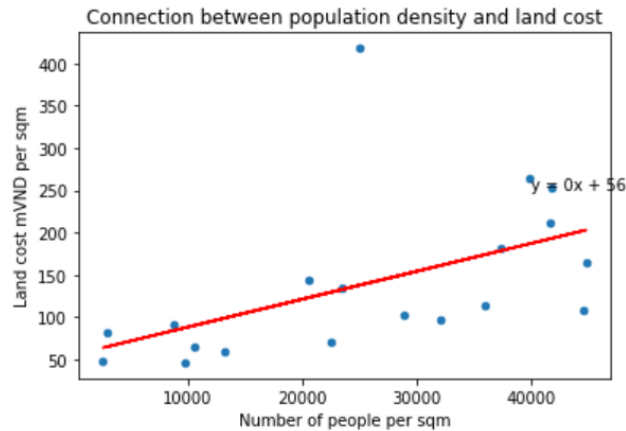
2. Data sources

The data comes from 3 sources:

- https://en.wikipedia.org/wiki/Ho_Chi_Minh_City for administration units
- https://mogi.vn/gia-nha-dat for land costs
- https://nominatim.openstreetmap.org/ for latitude and longitude
- Foursquare API for information about venues

3. Methodology

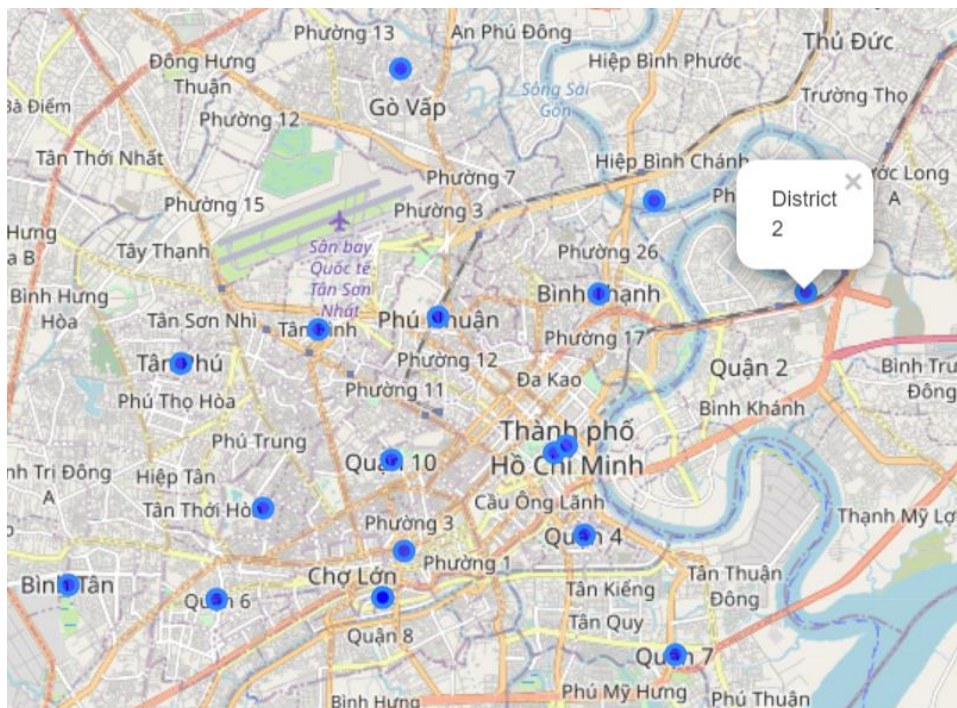To answer the first question, I use **Matplotlib**, particularly **bar charts**, to compare total population, area and density among urban districts. Instead of having three separate charts, I apply **subplots** so that the charts can be viewed side by side.

Next, a **scatter chart** illustrates the correlation between population density and land cost. Correlation coefficient is calculated by using **numpy.polyfit**.

Connection between population density and land cost

I manage to use **folium** and visualize urban districts in HCMC on a map. They are marked in blue circles with pop-up labels when being clicked.
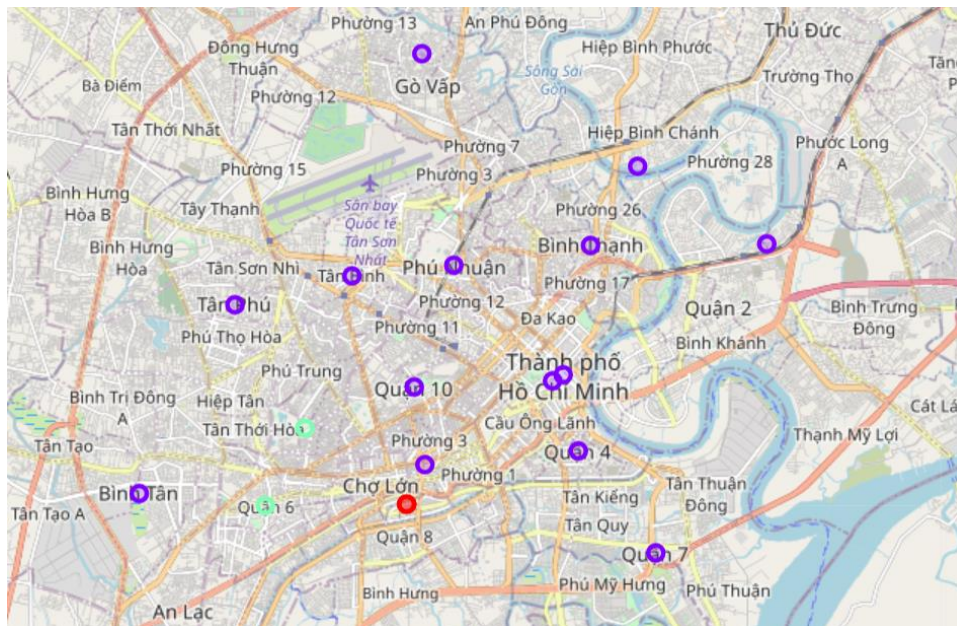

Foursquare API allows me to collect information about places in each district, including their names, categories and coordinates. Below is information about some venues in district 1:

	District	District_Latitude	District_Longitude	Venue	Venue_Latitude	Venue_Longitude	Venue_Category
0	1	10.77454	106.699184	Pizza 4P's	10.773301	106.697599	Pizza Place
1	1	10.77454	106.699184	Liberty Central Saigon Citypoint Hotel	10.774758	106.700795	Hotel
2	1	10.77454	106.699184	Pasteur Street Brewing Company	10.775220	106.700894	Brewery
3	1	10.77454	106.699184	CGV Cinemas Liberty CityPoint	10.774763	106.700766	Multiplex
4	1	10.77454	106.699184	Saxn'art Jazz Club	10.774068	106.700905	Jazz Club

Afterwards, **one hot encoding** will be run to return the most common venues in each district. A slice of the resulting table is below:

	District	1th Most common venue	2th Most common venue	3th 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	1	Vietnamese Restaurant	Hotel	Café	Coffee Shop	Asian Restaurant	Clothing Store	BBQ Joint	Multiplex	Burger Joint	Bar
1	10	Café	Vietnamese Restaurant	Coffee Shop	Bookstore	Multiplex	Arts & Crafts Store	Hotpot Restaurant	Sushi Restaurant	Korean Restaurant	Steakhouse
2	11	Café	Pizza Place	Eastern European Restaurant	Cocktail Bar	Coffee Shop	Convention Center	Cupcake Shop	Department Store	Dessert Shop	Dim Sum Restaurant
3	12	Vietnamese Restaurant	Food Truck	Dessert Shop	Chinese Restaurant	Whisky Bar	Eastern European Restaurant	Cocktail Bar	Coffee Shop	Convention Center	Cupcake Shop

Finally, **KMeans** is adopted to cluster and categorize the urban districts into 3 distinct groups based on their similarity of common places. The result of clustering is visualized by a folium map.

4. Results

After running the codes, there are some interesting findings:

- District 9 has the largest area while Binh Tan District has the biggest population. District 5 and 11 are the ones that have highest population density.
- There is a positive correlation between population density and the land cost. However, the correlation coefficient is not highly positive. With a formula of y = 0x + 56, it indicates a weakly positive connection.
- The most common venues in HCMC urban districts are Vietnamese restaurants and cafe/coffee shops.
- Among 3 clusters, 13 out of 19 urban districts share the same characteristics and are grouped in cluster no.1. District 8, 9 and 12 are in cluster no.0 due to the popularity of food trucks, desert shops and Chinese restaurants. Finally, district 6 and 11 are in cluster no.2.
  Cluster 0:

	District	Latitude	Longitude	Cluster Labels	1th Most common venue	2th Most common venue	3th 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	8	10.747343	106.666116	0	Vietnamese Restaurant	Food Truck	Dessert Shop	Chinese Restaurant	Whisky Bar	Eastern European Restaurant	Cocktail Bar	Coffee Shop	Convention Center	Cupcake Shop
8	9	10.747343	106.666116	0	Vietnamese Restaurant	Food Truck	Dessert Shop	Chinese Restaurant	Whisky Bar	Eastern European Restaurant	Cocktail Bar	Coffee Shop	Convention Center	Cupcake Shop
11	12	10.747343	106.666116	0	Vietnamese Restaurant	Food Truck	Dessert Shop	Chinese Restaurant	Whisky Bar	Eastern European Restaurant	Cocktail Bar	Coffee Shop	Convention Center	Cupcake Shop

Cluster 1:

	District	Latitude	Longitude	Cluster Labels	1th Most common venue	2th Most common venue	3th 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
0	1	10.774540	106.699184	1	Vietnamese Restaurant	Hotel	Café	Coffee Shop	Asian Restaurant	Clothing Store	BBQ Joint	Multiplex	Burger Joint
1	2	10.804963	106.747470	1	Café	Asian Restaurant	Restaurant	Shopping Mall	Buffet	Department Store	Movie Theater	Coffee Shop	Residential Building (Apartment / Condo)
2	3	10.775844	106.701756	1	Hotel	Coffee Shop	Café	Vietnamese Restaurant	Thai Restaurant	Bar	French Restaurant	Multiplex	Massage Studio
3	4	10.759243	106.704890	1	Seafood Restaurant	Vietnamese Restaurant	Snack Place	Furniture / Home Store	Japanese Restaurant	Whisky Bar	Dumpling Restaurant	Cocktail Bar	Coffee Shop
4	5	10.756129	106.670376	1	Noodle House	Chinese Restaurant	Coffee Shop	Dim Sum Restaurant	Vietnamese Restaurant	Convention Center	Asian Restaurant	Hotel	Dessert Shop
6	7	10.736573	106.722432	1	Vietnamese Restaurant	Café	Flea Market	Coffee Shop	Eastern European Restaurant	Cocktail Bar	Convention Center	Cupcake Shop	Department Store
9	10	10.773198	106.667833	1	Café	Vietnamese Restaurant	Coffee Shop	Bookstore	Multiplex	Arts & Crafts Store	Hotpot Restaurant	Sushi Restaurant	Korean Restaurant
12	GoVap	10.847072	106.669609	1	Grocery Store	Food Truck	Boutique	Warehouse Store	Furniture / Home Store	French Restaurant	Cocktail Bar	Coffee Shop	Convention Center
13	Tan Binh	10.797979	106.653805	1	Gym / Fitness Center	Sushi Restaurant	Fast Food Restaurant	Multiplex	Breakfast Spot	Stadium	Steakhouse	Supermarket	Bubble Tea Shop
14	Tan Phu	10.791640	106.627302	1	Café	Vietnamese Restaurant	Coffee Shop	BBQ Joint	Electronics Store	Cocktail Bar	Convention Center	Cupcake Shop	Department Store
15	Binh Thanh	10.804659	106.707848	1	Soup Place	Travel Agency	Whisky Bar	Eastern European Restaurant	Cocktail Bar	Coffee Shop	Convention Center	Cupcake Shop	Department Store
16	Phu Nhuan	10.800118	106.677042	1	Coffee Shop	Café	Asian Restaurant	Bike Rental / Bike Share	Hotel	Juice Bar	Electronics Store	Convention Center	Bookstore
17	Thu Duc	10.822023	106.718302	1	Café	Vietnamese Restaurant	Seafood Restaurant	Food Truck	Pool	Dim Sum Restaurant	Chinese Restaurant	Clothing Store	Cocktail Bar
18	Binh Tan	10.749809	106.605664	1	Video Store	Asian Restaurant	Brazilian Restaurant	Whisky Bar	Electronics Store	Coffee Shop	Convention Center	Cupcake Shop	Department Store

Cluster 2:

	District	Latitude	Longitude	Cluster Labels	1th Most common venue	2th Most common venue	3th 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
5	6	10.746928	106.634495	2	Café	Electronics Store	Cocktail Bar	Coffee Shop	Convention Center	Cupcake Shop	Department Store	Dessert Shop	Dim Sum Restaurant	Dumpling Restaurant
10	11	10.764208	106.643282	2	Café	Pizza Place	Eastern European Restaurant	Cocktail Bar	Coffee Shop	Convention Center	Cupcake Shop	Department Store	Dessert Shop	Dim Sum Restaurant

- To closely examine the codes, please click here .

--END--