



IBM Data Science
IBM



PROJECT: Find location to open an Italian restaurant in Ho Chi Minh, Viet Nam.

Create by: Hoang Tien Thuy

09/2019

1

INTRODUCTION

2

DATA AND CLEANING

3

METHODOLOGY

4

ANALYSIS

5

RESULTS&DISCUSSION

6

CONCLUSION

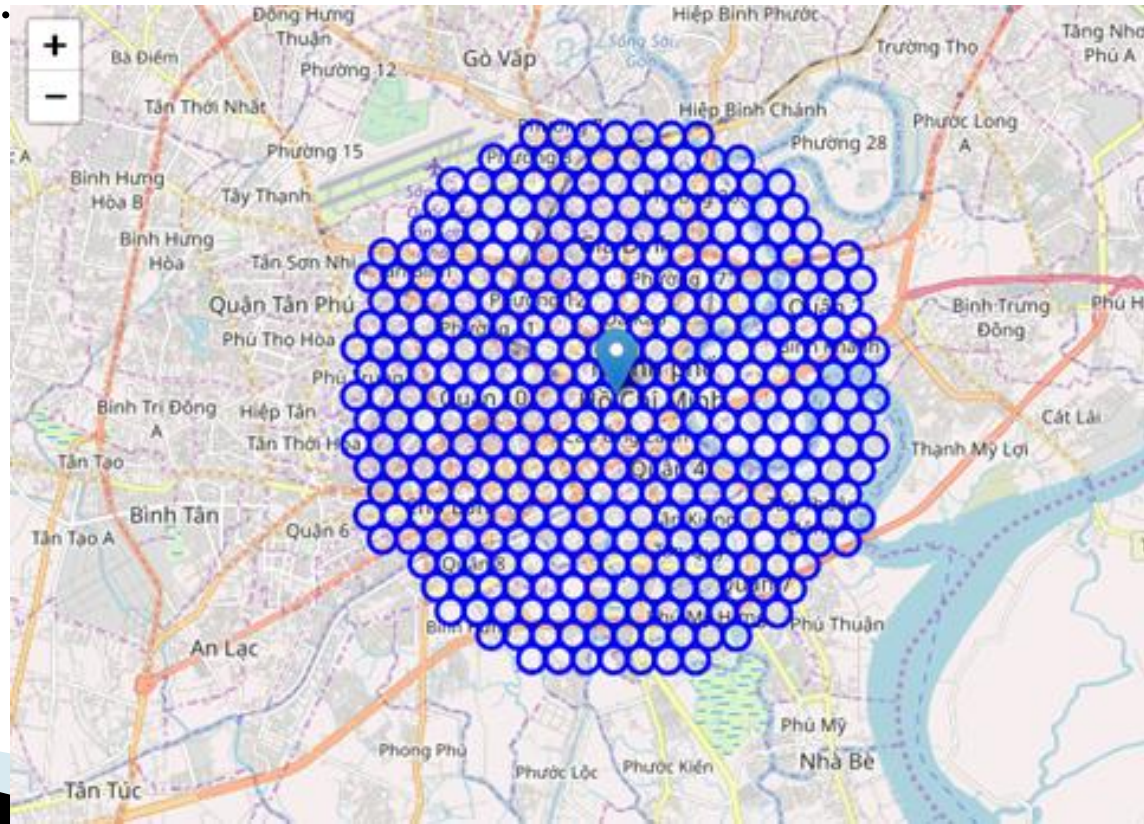
- We will build **Italian restaurants** in Viet Nam . we will try to find an optimal location for a **Italian restaurant in Ho Chi Minh city, Viet Nam**. Based on definition of our problem, factors that will influence our decision are:
 - ✓ Number of existing restaurants in the neighborhood (any type of restaurant)
 - ✓ Number of and distance to Italian restaurants in the neighborhood, if any
 - ✓ Distance of neighborhood from city center



- Following data sources will be needed to extract/generate the required information:
- ✓ Approximate addresses of centers of those areas will be obtained using library `geopy.geocoders`
- ✓ Number of restaurants and their type and location in every neighborhood will be obtained using **Foursquare API**
- ✓ Coordinate of Ho Chi Minh center will be obtained using library `geopy.geocoders` of well known as Bến thành Market.



- Create a grid of cells covering our area of interest which is approx. 12x12 kilometers centered around Ho Chi Minh city center. Zone (VietNam)= 48. Generated 364 candidate neighborhood centers.



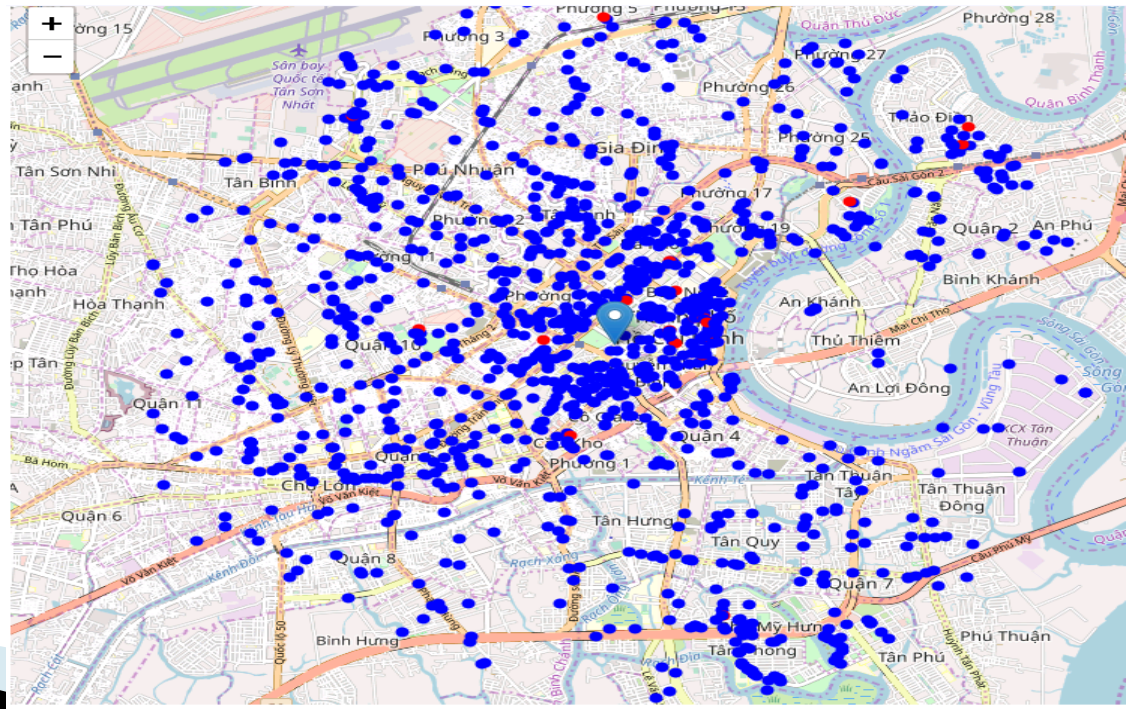
✓ **Use Foursquare for collect data and clean data:** get address about restaurants and Italian restaurants

Total number of restaurants: 1282 (Blue)

Total number of Italian restaurants: 20 (red)

Percentage of Italian restaurants: 1.56%

Average number of restaurants in neighborhood: 3.049



- ✓ **Step 1: Detecting areas of Ho Chi Minh that have low restaurant density, particularly those with low number of Italian restaurants**
- ✓ **Step 2 : Use heatmaps to identify a few promising areas close to center with low number of restaurants in general (and no Italian restaurants in vicinity) and focus our attention on those areas.**
- ✓ **Step 3: present map of all such locations but also create clusters (using k-means clustering) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.**

✓ **Detecting areas of Ho Chi Minh:** Average distance to closest Italian restaurant from each area center: 1759.86 m

```
In [95]: location_restaurants_count = [len(res) for res in location_restaurants]

df_locations['Restaurants in area'] = location_restaurants_count

print('Average number of restaurants in every area with radius=300m:', np.array(location_restaurants_count).mean())

df_locations.head(10)
```

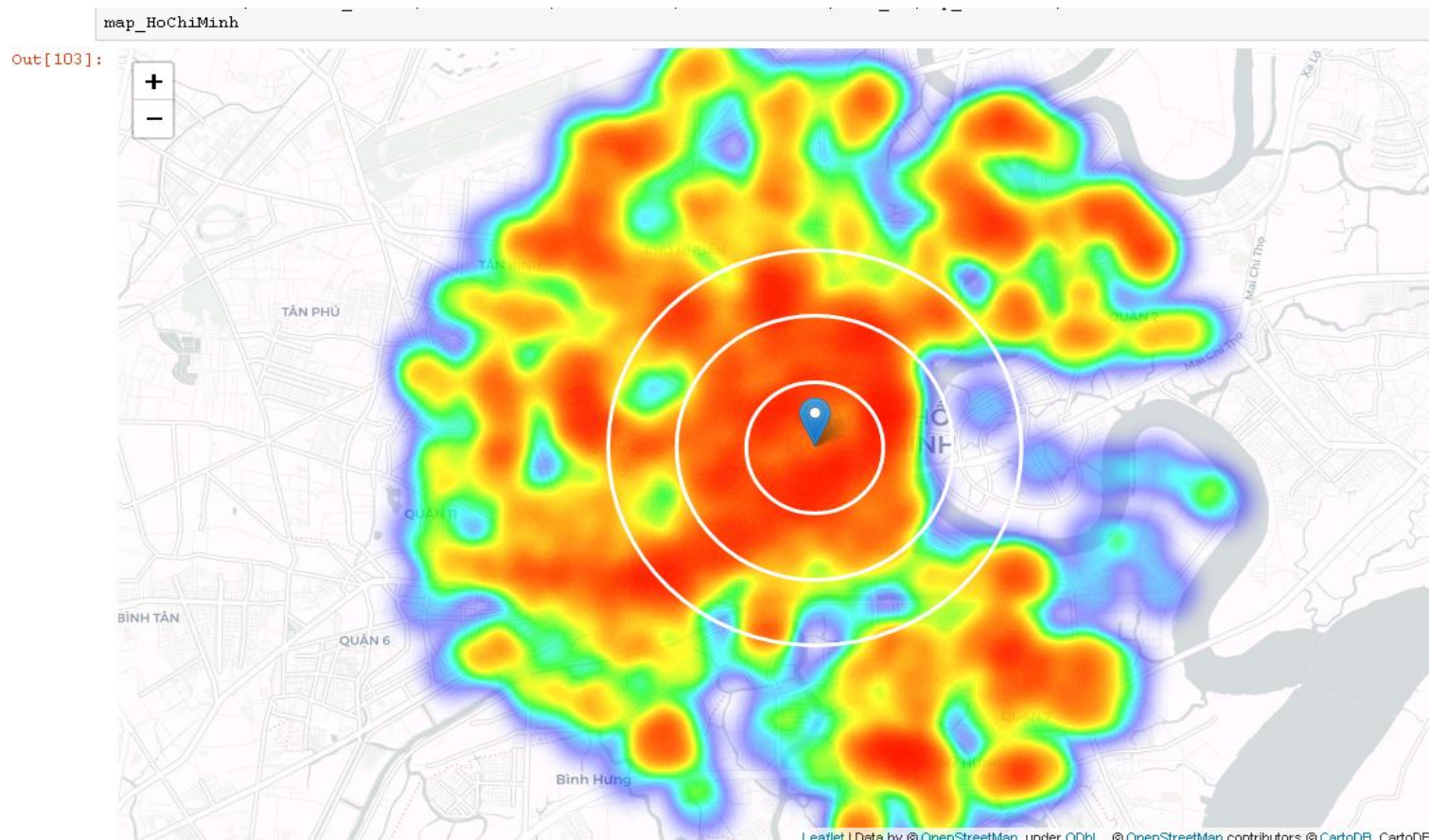
Average number of restaurants in every area with radius=300m: 3.0494505494505493

Out[95]:

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area
0	Khu 6, Xã Bình Hưng, Huyện Bình Chánh, Thành p...	10.721284	106.677557	683457.385186	1.185662e+06	5992.495307	0
1	Phạm Hùng, Khu 6, Khu Dân cư T30, Xã Bình Hưng...	10.721254	106.683042	684057.385186	1.185662e+06	5840.376700	0
2	Đường số 1, Khu 6, Xã Bình Hưng, Huyện Bình Ch...	10.721225	106.688527	684657.385186	1.185662e+06	5747.173218	0
3	Đường A, Khu dân cư Phước Kiển, Khu Dân cư Ph...	10.721195	106.694012	685257.385186	1.185662e+06	5715.767665	0
4	Xã Phước Kiển, Huyện Nhà Bè, Thành phố Hồ Chí ...	10.721165	106.699497	685857.385186	1.185662e+06	5747.173218	4
5	Xã Phước Kiển, Huyện Nhà Bè, Thành phố Hồ Chí ...	10.721135	106.704981	686457.385186	1.185662e+06	5840.376700	0
6	Trường Đình Thiện Lý, 80, Nguyễn Đức Cảnh, Khu...	10.721105	106.710466	687057.385186	1.185662e+06	5992.495307	13
7	Cầu Xóm Củi, Đường Nguyễn Văn Linh, Xã Bình Hư...	10.726026	106.669356	682557.385186	1.186181e+06	5855.766389	0
8	Xã Bình Hưng, Huyện Bình Chánh, Thành phố Hồ C...	10.725996	106.674841	683157.385186	1.186181e+06	5604.462508	0
9	Khu 6, Xã Bình Hưng, Huyện Bình Chánh, Thành p...	10.725967	106.680326	683757.385186	1.186181e+06	5408.326913	2

ANALYSIS

✓ **Use heatmaps to identify a few promising areas close to center with low number of restaurants in general:** Looks like a few pockets of low restaurant density closest to city center can be found.

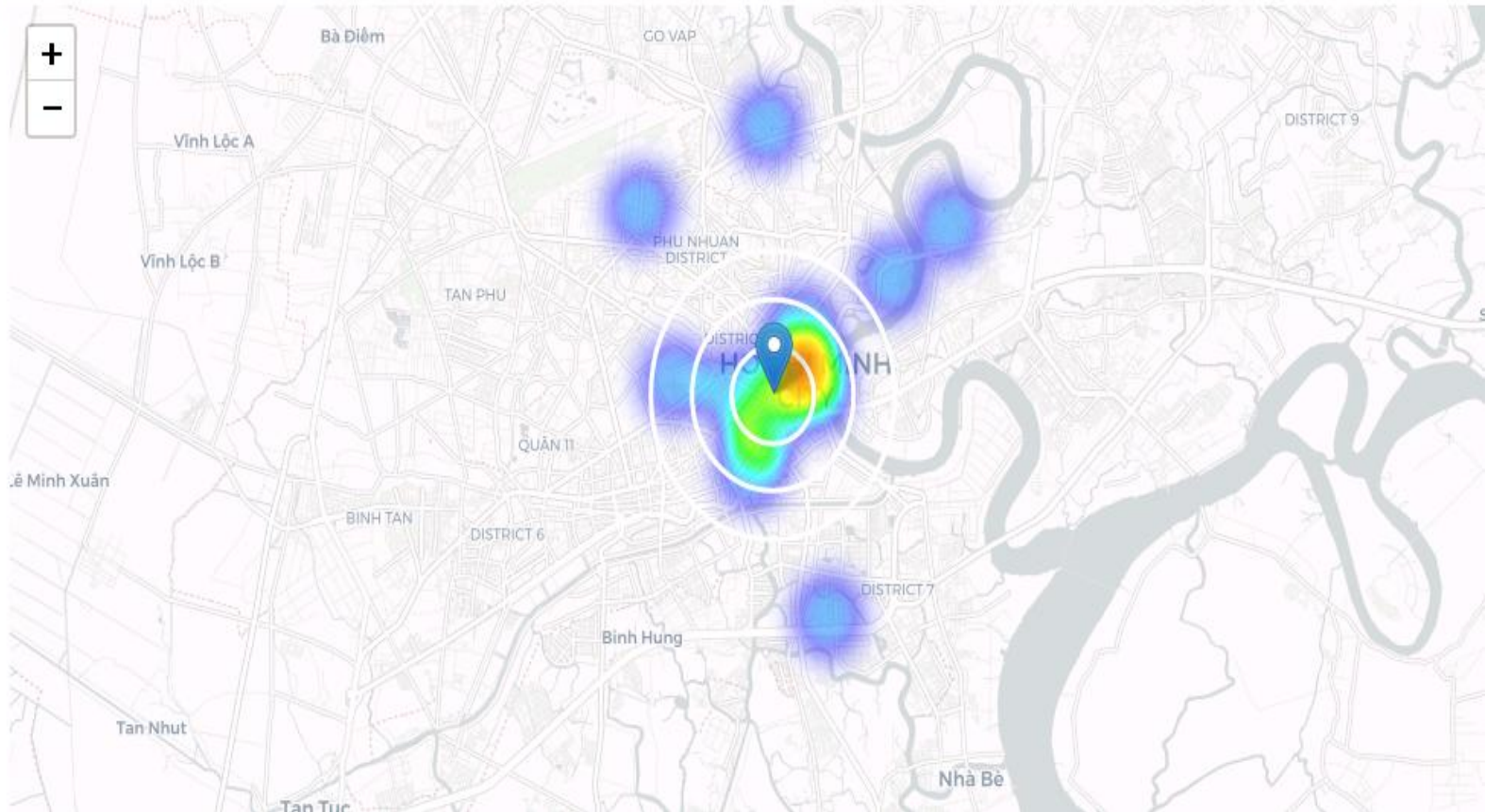


4

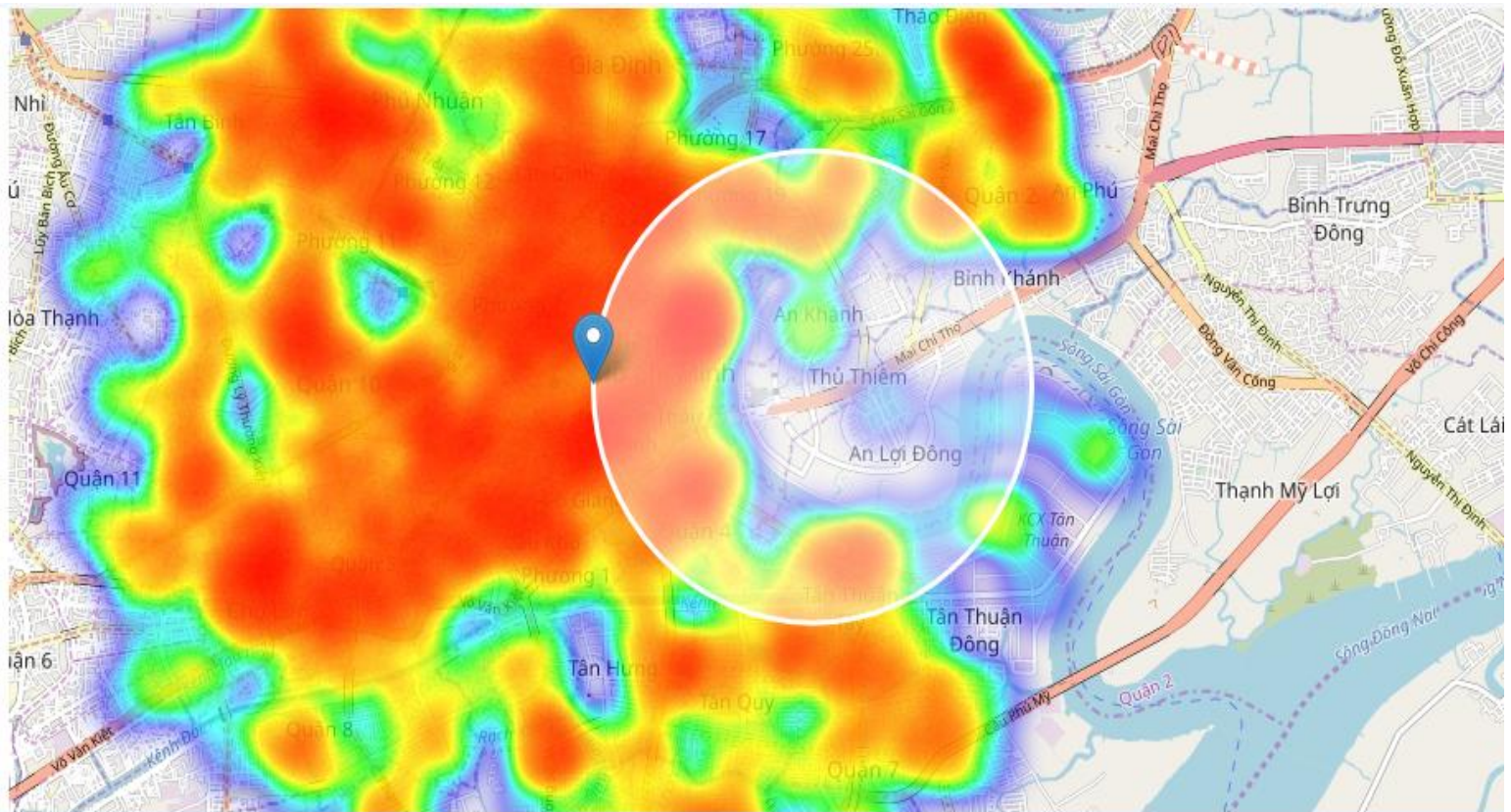
ANALYSIS

✓ Use heatmaps showing heatmap/density of Italian restaurants only.

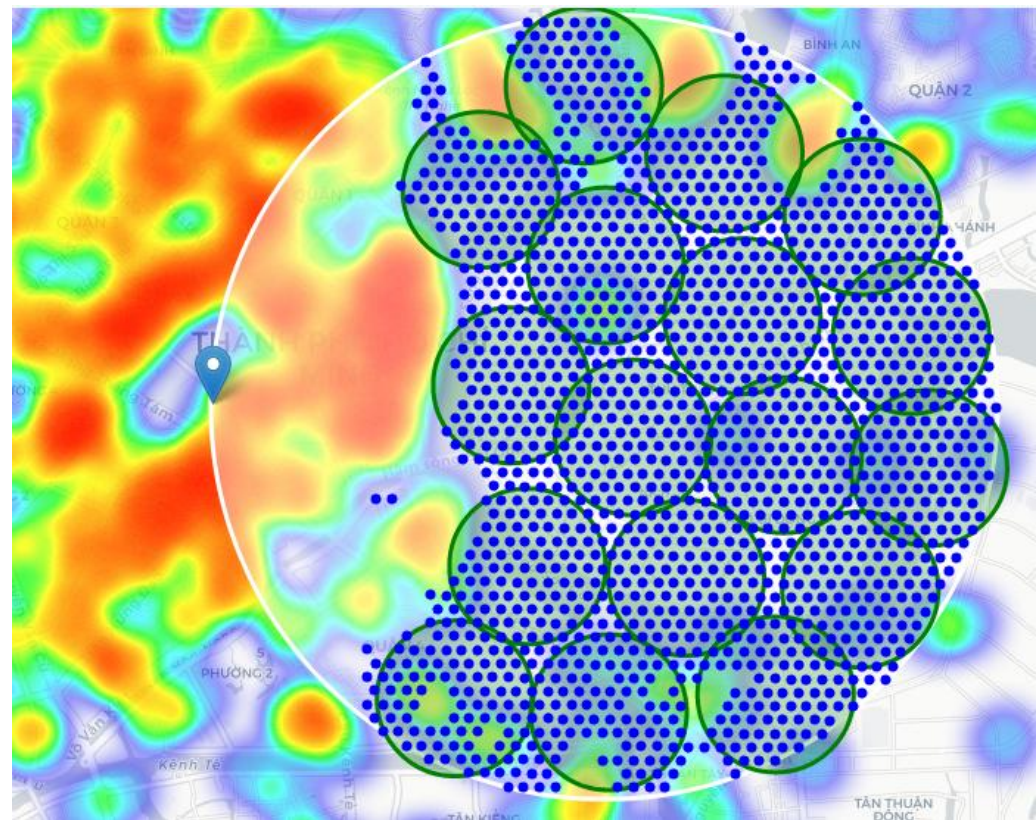
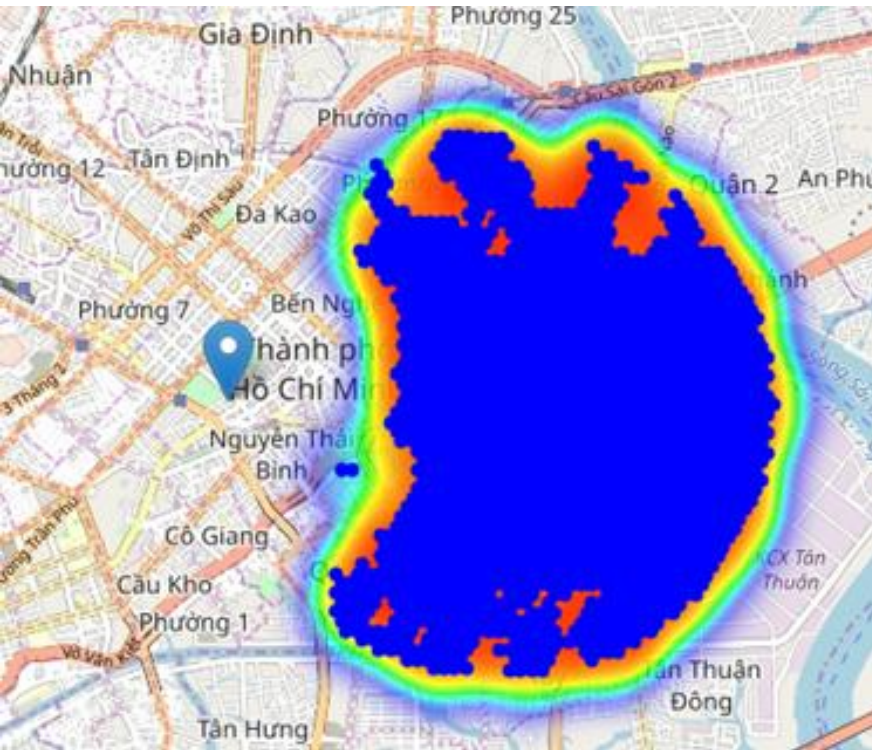
Out[104]:



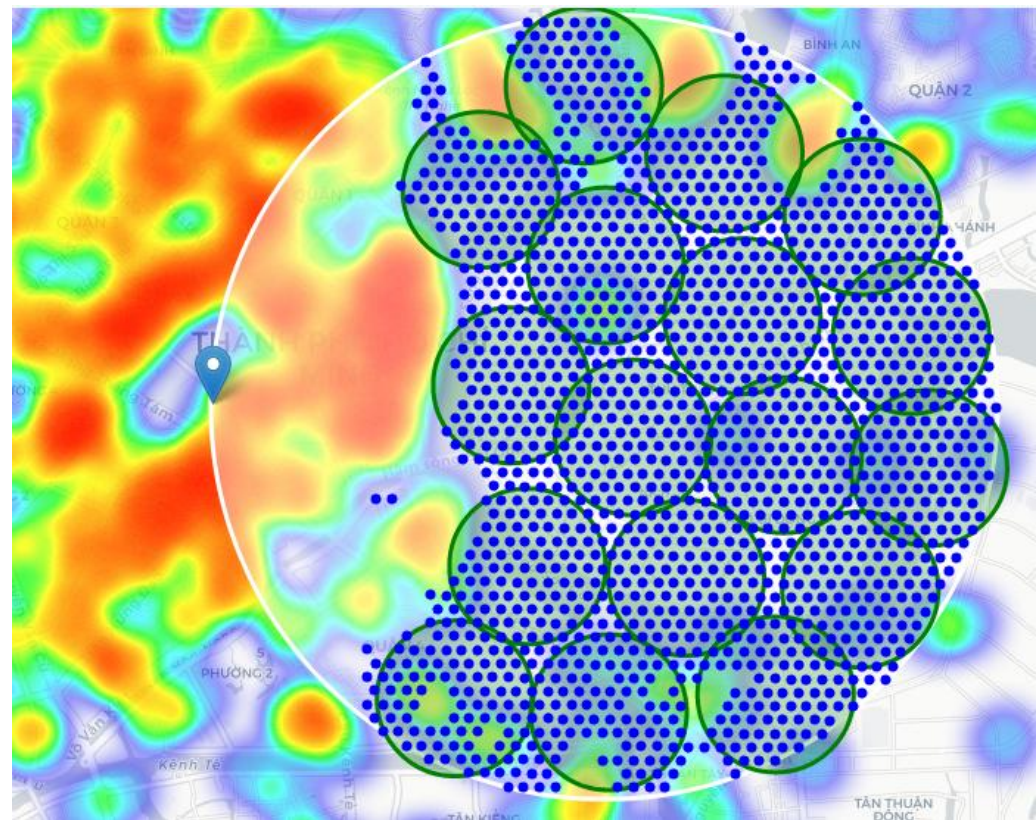
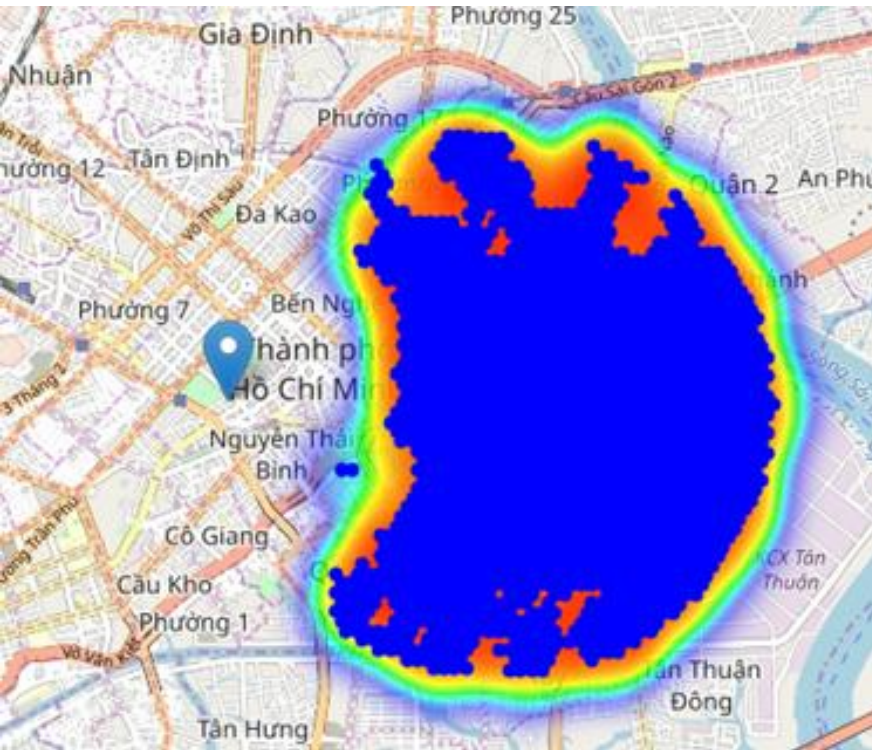
- ✓ Use **heatmaps** :pockets of low restaurant density in An Khanh and Thu Thiêm closest to Ben Thanh Market .



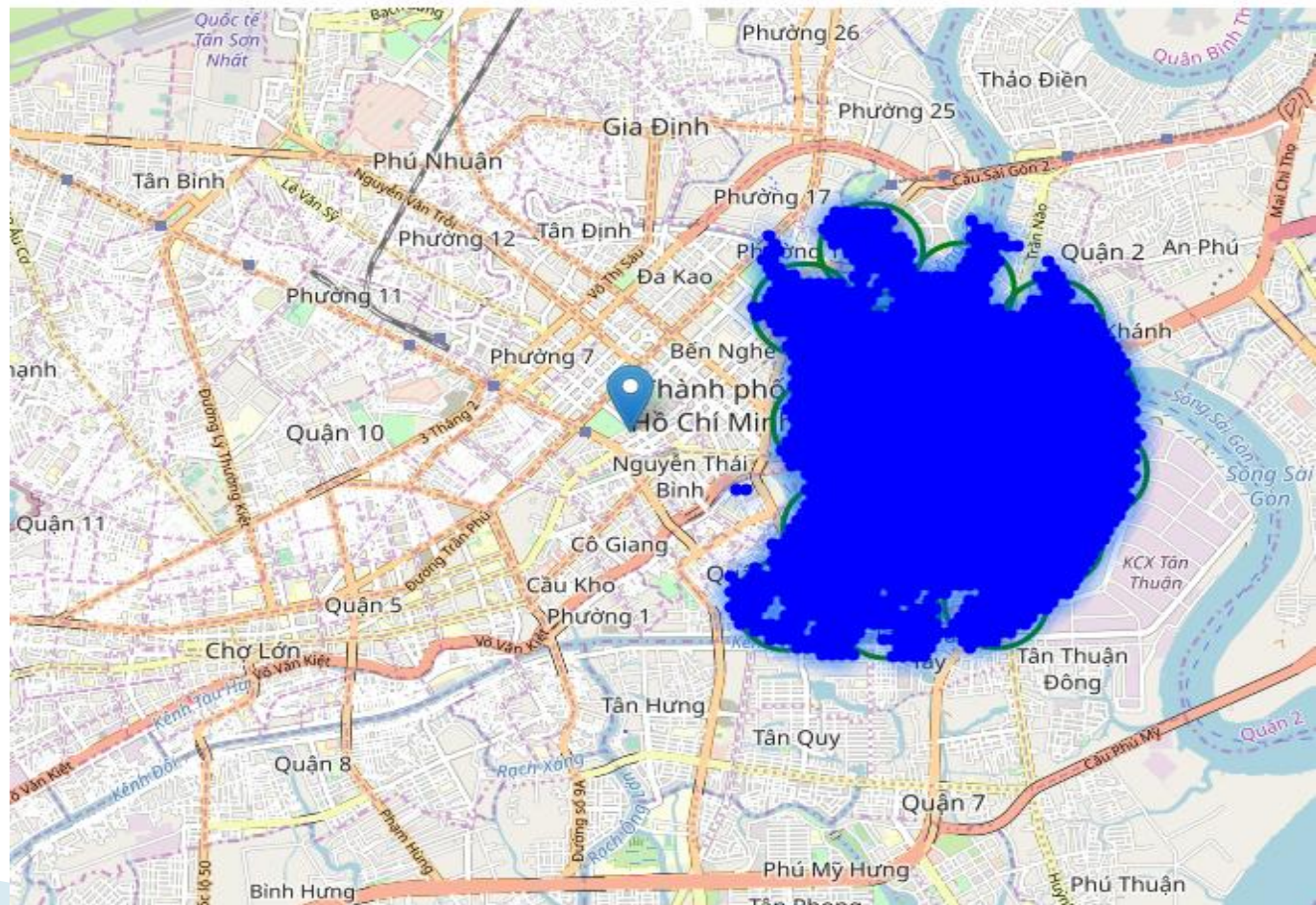
✓ Create clusters (using k-means clustering) : 17 Cluster



✓ Create clusters (using k-means clustering) : 17 Cluster



✓ Most of the zones are located in ****Phường An Lợi Đông, Quận 2, Thành phố Hồ Chí Minh, 77000 => 3.4km from Ben Thành Market****

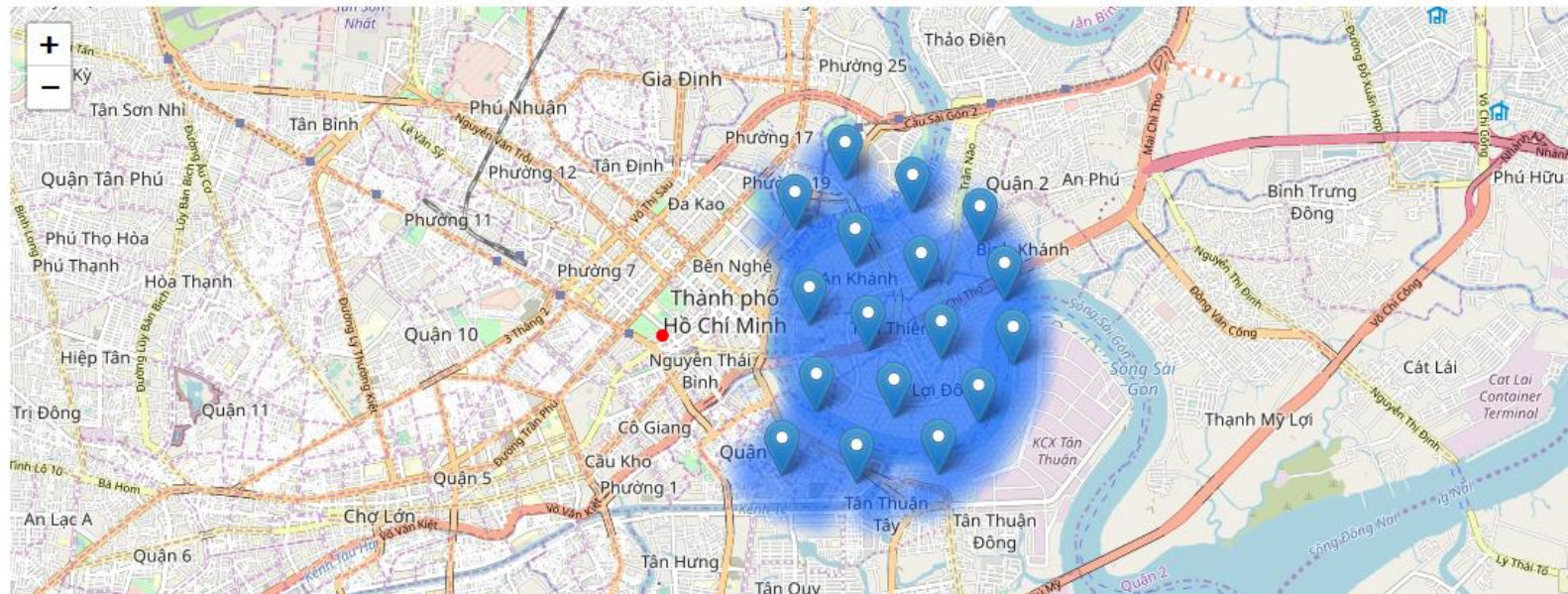


5 RESULTS&DISCUSSION

✓ Result of all this is 17 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Italian restaurants particularly.

```
folium.Circle(HoChiMinh_center, radius=50, color='red', fill=True, fill_color='red', fill_opacity=1).add_to(map_HoChiMinh)
for lonlat, addr in zip(cluster_centers, candidate_area_addresses):
    folium.Marker([lonlat[1], lonlat[0]], popup=addr).add_to(map_HoChiMinh)
for lat, lon in zip(good_latitudes, good_longitudes):
    folium.Circle([lat, lon], radius=250, color='#0000ff00', fill=True, fill_color='#0066ff', fill_opacity=0.05).add_to(map_HoChiMinh)
map_HoChiMinh
```

Out[66]:



✓ Purpose of this project was to identify Hồ Chí Minh areas close to center with low number of restaurants and I show 17 location

```
In [64]: candidate_area_addresses = []
print('=====')
print('Addresses of centers of areas recommended for further analysis')
print('=====\\n')
for lon, lat in cluster_centers:
    addr = get_address( lat, lon).replace(', Việt Nam', '')
    candidate_area_addresses.append(addr)
    x, y = lonlat_to_xy(lon, lat)
    d = calc_xy_distance(x, y, HoChiMinh_center_x, HoChiMinh_center_y)
    print('{}{} => {:.1f}km from Ben THành Market'.format(addr, ' '* (50-len(addr)), d/1000))
```

```
=====
Addresses of centers of areas recommended for further analysis
=====
```

```
Phường An Lợi Đông, Quận 2, Thành phố Hồ Chí Minh, 77000 => 3.4km from Ben THành Market
Phường An Lợi Đông, Quận 2, Thành phố Hồ Chí Minh, 8408 => 3.7km from Ben THành Market
Cây Bàng, Phường Thủ Thiêm, Quận 2, Thành phố Hồ Chí Minh, 8408 => 2.3km from Ben THành Market
Phường Bến Nghé, Quận 1, Thành phố Hồ Chí Minh, 77000 => 2.2km from Ben THành Market
Đường số 11, Phường An Lợi Đông, Quận 2, Thành phố Hồ Chí Minh, 71108 => 4.5km from Ben THành Market
Phường Bình An, Quận 2, Thành phố Hồ Chí Minh, NGŨ TẮT TỔ => 3.6km from Ben THành Market
Lương Định Của, Phường Thủ Thiêm, Quận 2, Thành phố Hồ Chí Minh, 8408 => 1.9km from Ben THành Market
Phường Tân Thuận Đông, Quận 7, Thành phố Hồ Chí Minh, NGŨ TẮT TỔ => 4.6km from Ben THành Market
Phường An Lợi Đông, Quận 2, Thành phố Hồ Chí Minh, 8408 => 3.2km from Ben THành Market
Hẻm 200C Xóm Chiếu, Phường 15, Quận 4, Thành phố Hồ Chí Minh, 710200 => 2.4km from Ben THành Market
Hẻm 108 Võ Duy Ninh, Phường 22, Quận Bình Thạnh, Thành phố Hồ Chí Minh, NGŨ TẮT TỔ => 3.1km from Ben THành Market
Trần Văn Khánh, Phường Tân Thuận Đông, Quận 7, Thành phố Hồ Chí Minh, 8408 => 4.0km from Ben THành Market
Phường Bình Khánh, Quận 2, Thành phố Hồ Chí Minh, NGŨ TẮT TỔ => 4.3km from Ben THành Market
Cầu Kênh 1, Mai Chí Thọ, Phường Thủ Thiêm, Quận 2, Thành phố Hồ Chí Minh, 8408 => 2.7km from Ben THành Market
Đại lộ vòng cung, Phường An Khánh, Quận 2, Thành phố Hồ Chí Minh, 77000 => 2.7km from Ben THành Market
```

✓ Link Data :

https://github.com/hoangtienthuy83/Coursera_Capstone/blob/master/restaurants_hochiminh.csv

✓ Link project Notebook :

Thank You for Reading

A blue ballpoint pen is shown in the process of writing the words "Thank you!" in a cursive script on a white, slightly textured surface. The pen is positioned at the end of the word "you!", with its tip touching the paper. The lighting is soft, creating a slight shadow of the pen on the surface.

Thank you!