Capstone Project

THE BATTLE OF THE NEIGHBORHOODS

Table of contents

Introduction: Business Problem

Data

Methodology

Analysis

Results and Discussion

Conclusion

Introduction: Business Problem

The project will target stakeholders interested in opening an Italian restaurant in Ho Chi Minh City, Vietnam.

We will try to find optimal locations that are:
◦ Not already crowded with restaurants
◦ Without Italian restaurants in vicinity

Data science powers will be used to generate promising neighborhoods.

Data

Influential factors:

◦ Number of existing restaurants in the neighborhood (all type of restaurants)

◦ Number of and distance to Italian restaurants in the neighborhood (if any)

◦ Distance of neighborhood from city center

Data sources:

◦ Centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using GeoPy Nominatim API reverse geocoding.

◦ Number of restaurants, their type and location in every neighborhood will be obtained using Foursquare API.

◦ Coordinates of Ho Chi Minh City center will be obtained using GeoPy Nominatim API geocoding.

Methodology

First, collect location and type of every restaurant within 6km from city center. Italian restaurants will also be identified according to Foursquare categorization.

Second, calculate and explore 'restaurant density' across different areas of city center. Promising areas with low number of restaurants and no Italian restaurants in vicinity will be identified using heatmaps.
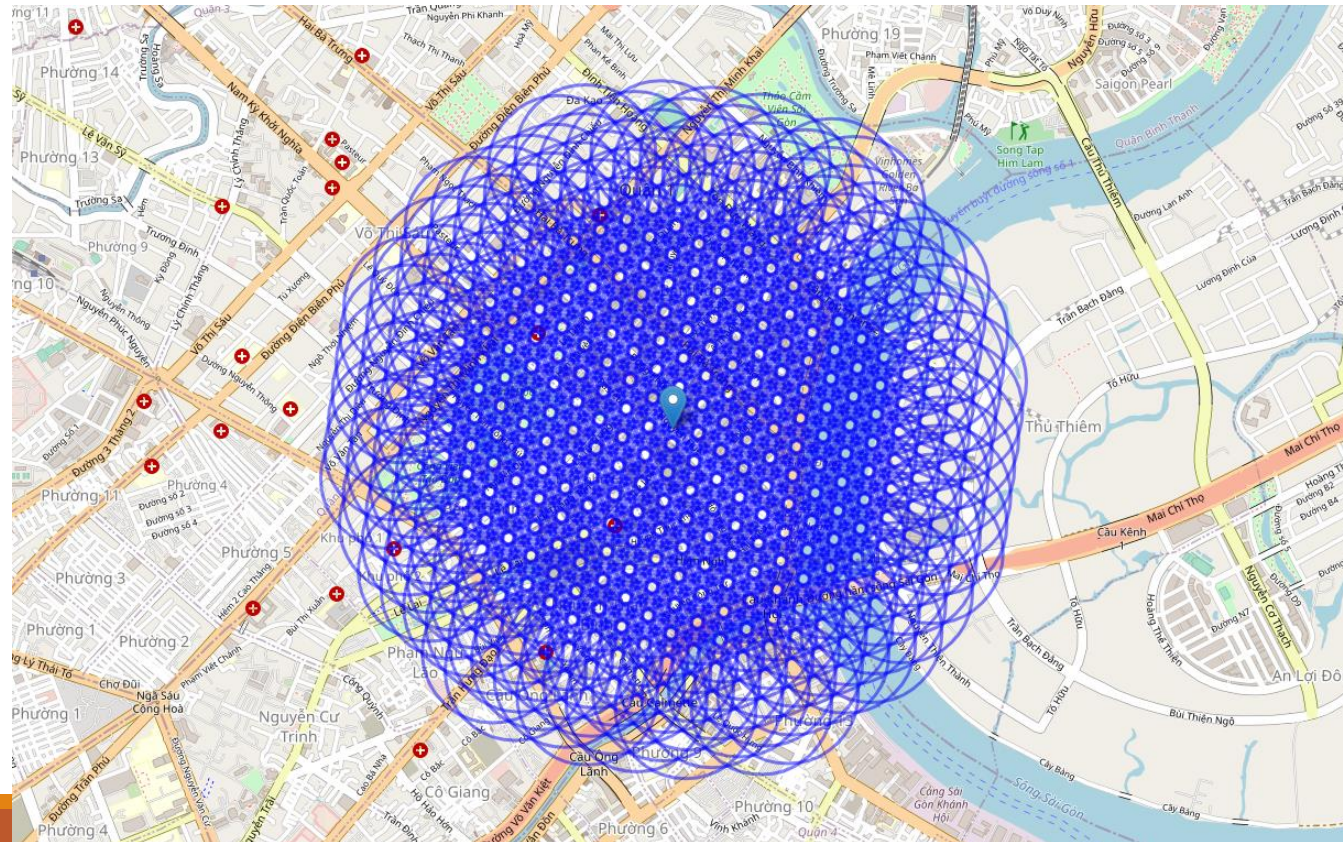
In final step, focus on most promising areas and within those create clusters (using k-means clustering) of locations that meet both following conditions:

◦ Locations with only one restaurant in radius of 250 meters.
◦ Locations without Italian restaurants in radius of 500 meters.

The map of all such locations should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

Analysis

A grid of equally spaced candidates within 6km around city center is created. Neighborhoods are defined as circular areas with a radius of 300 meters.
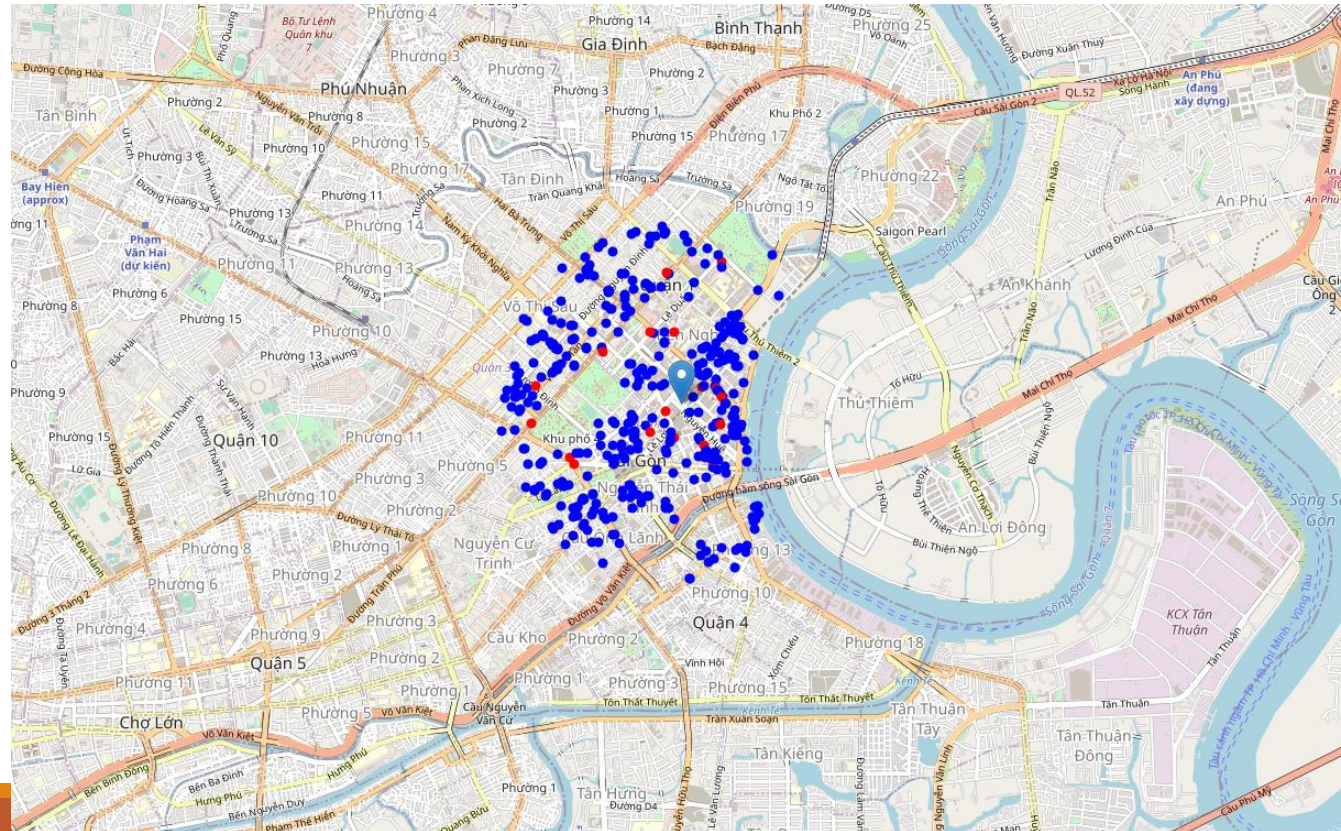
Analysis

GeoPy Nominatim API reverse geocoding is used to obtain the approximate addresses of centers of neighborhoods.

	Address	Latitude	Longitude	X	Y	Distance from center
0	Lẩu Kichi Kichi, Lê Quý Đôn, Võ Thị Sáu, Phườn...	10.780574	106.691692	1.522905e+07	1.090728e+07	5992.495307
1	Võ Thị Sáu, Phường Võ Thị Sáu, Quận 3, 7000000	10.779478	106.691541	1.522965e+07	1.090728e+07	5840.376700
2	Phòng Khám Sản Nhi Diamond, 9, Trần Quốc Thảo,...	10.778382	106.691389	1.523025e+07	1.090728e+07	5747.173218
3	Maritime, Võ Văn Tần, Võ Thị Sáu, Phường Võ Th...	10.777287	106.691238	1.523085e+07	1.090728e+07	5715.767665
	...					

Analysis

Foursquare API is used to get information on restaurants in each neighborhood. Restaurants in our area of interest are in blue, Italian restaurants are in red:
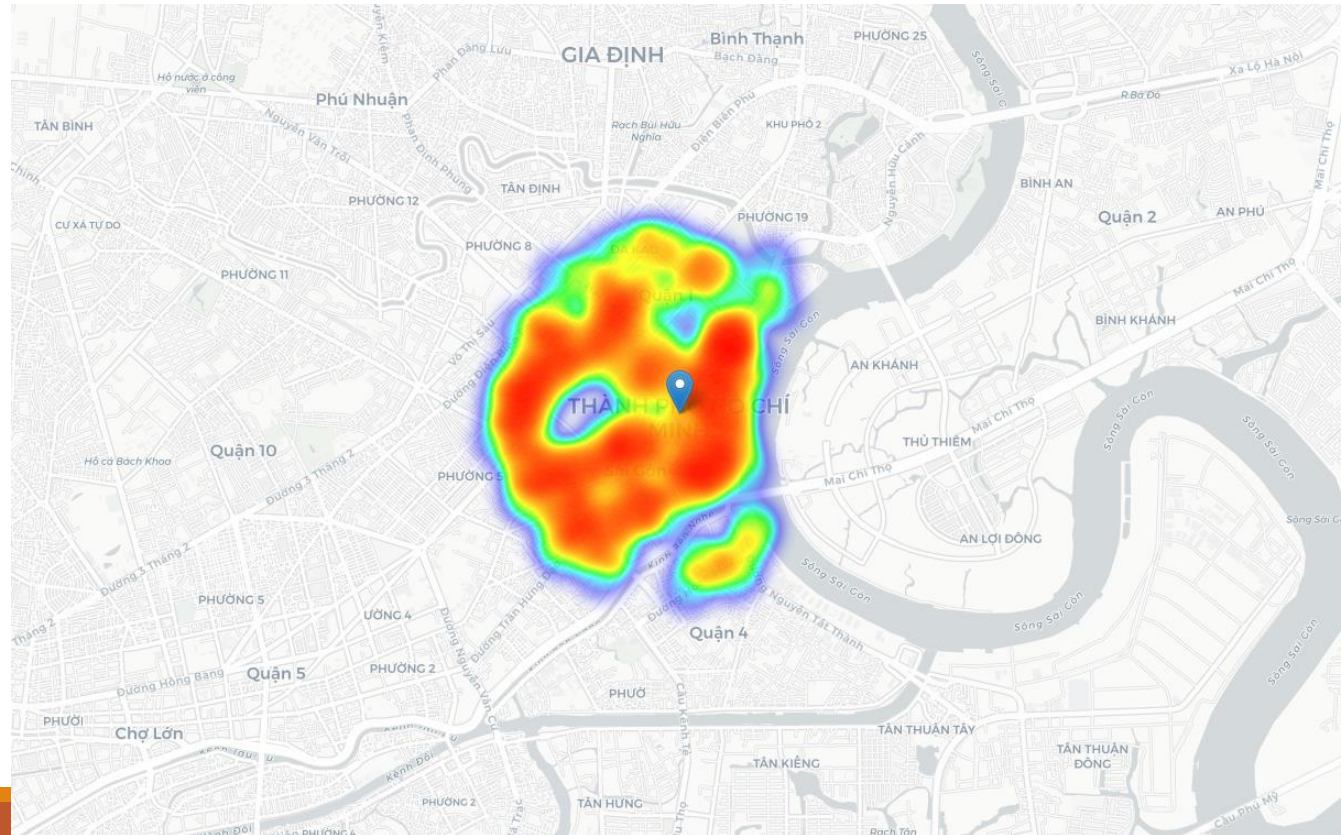
Analysis

Count the number of restaurants in every area candidate and calculate the distance to nearest Italian restaurant from every area candidate center.

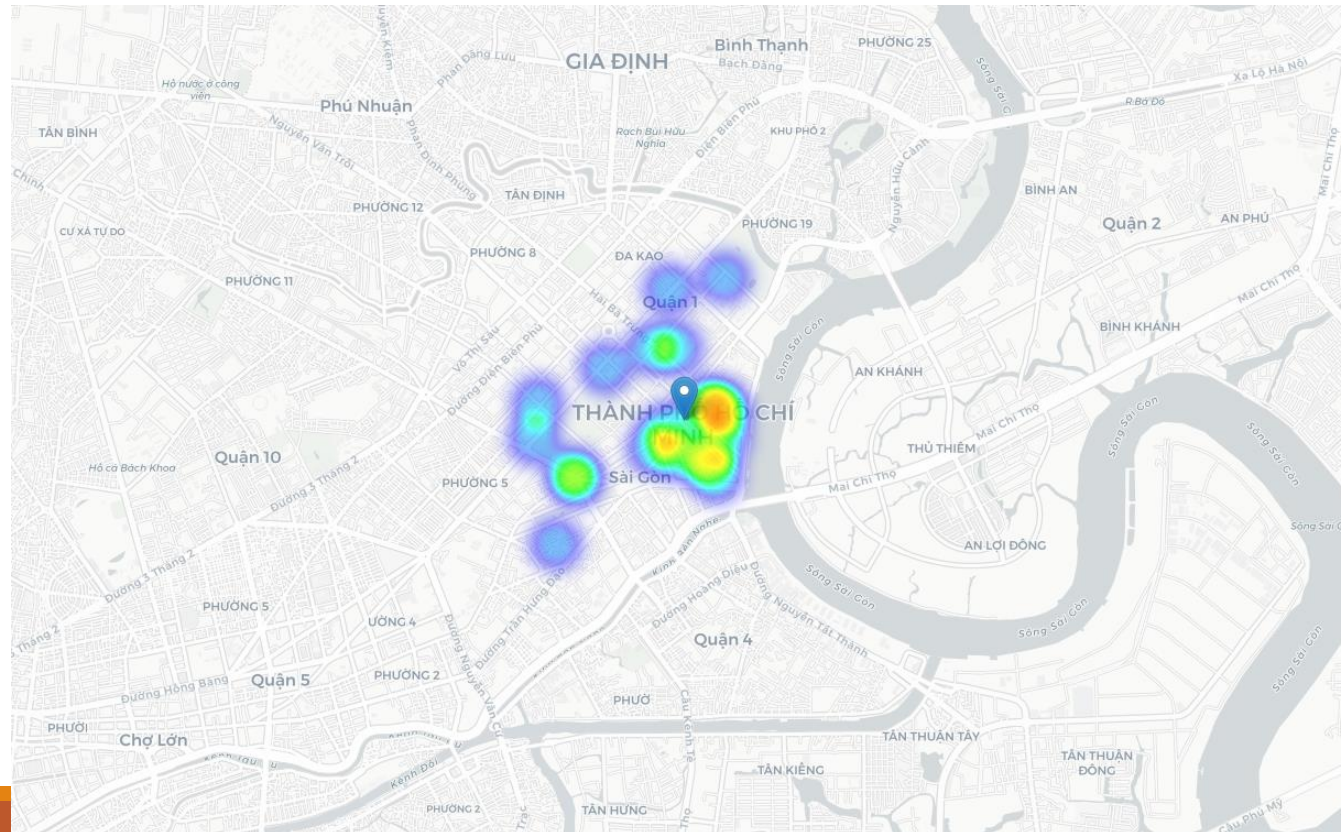|	Address	Latitude	Longitude	Distance from center	Restaurants in area	Distance to Italian restaurant
0	Lẩu Kichi Kichi, Lê Quý Đôn, Võ Thị Sáu, Phườn…	10.780574	106.691692	5992.495307	32	1921.587500
1	Võ Thị Sáu, Phường Võ Thị Sáu, Quận 3, 7000000	10.779478	106.691541	5840.376700	29	1339.665224
2	Phòng Khám Sản Nhi Diamond, 9, Trần Quốc Thảo,…	10.778382	106.691389	5747.173218	30	785.434468
3	Maritime, Võ Văn Tần, Võ Thị Sáu, Phường Võ Th…	10.777287	106.691238	5715.767665	23	398.888074
	…					

Analysis

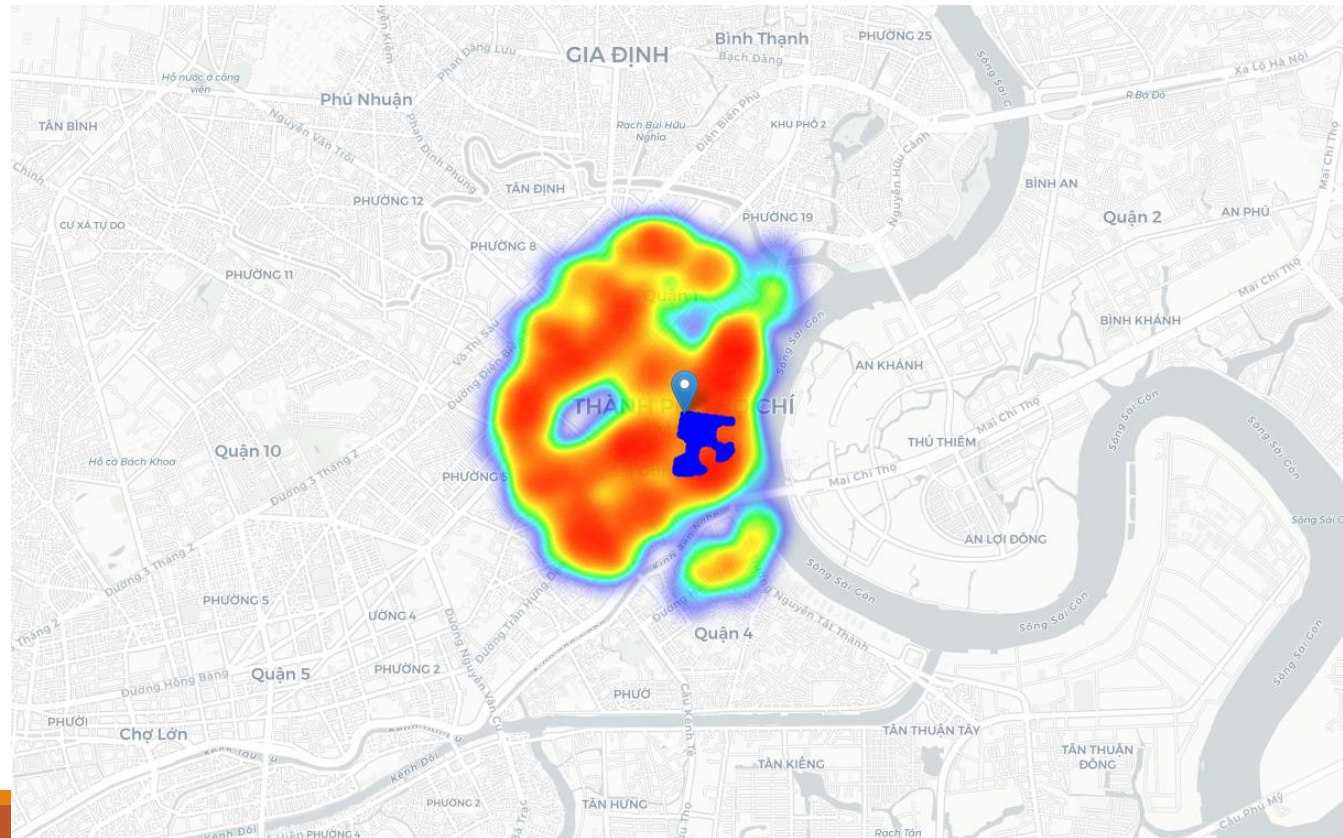Map showing density of all restaurants

Analysis

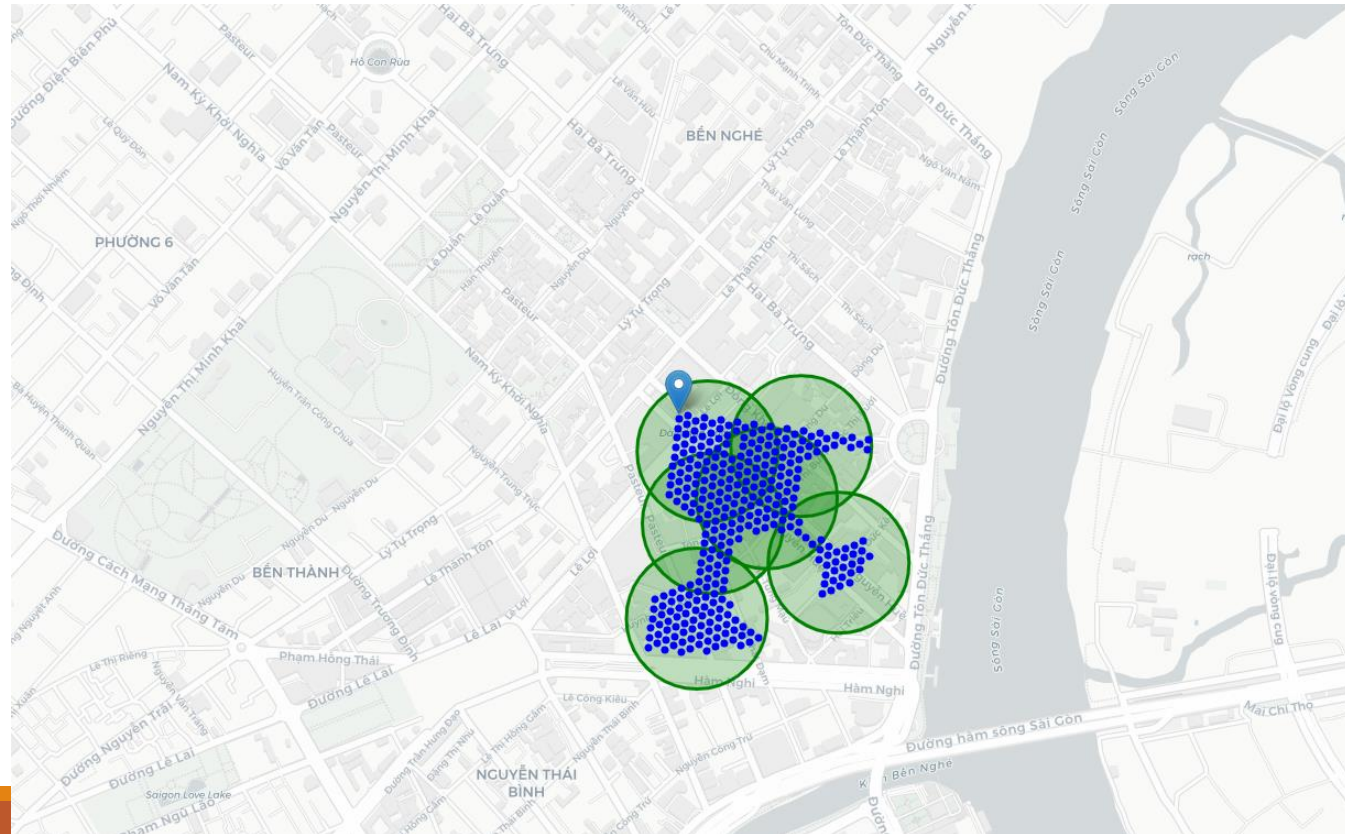Map showing density of Italian restaurants only

Analysis

Map showing locations in blue with one restaurant in radius of 250 meters, and no Italian restaurants in radius of 500 meters

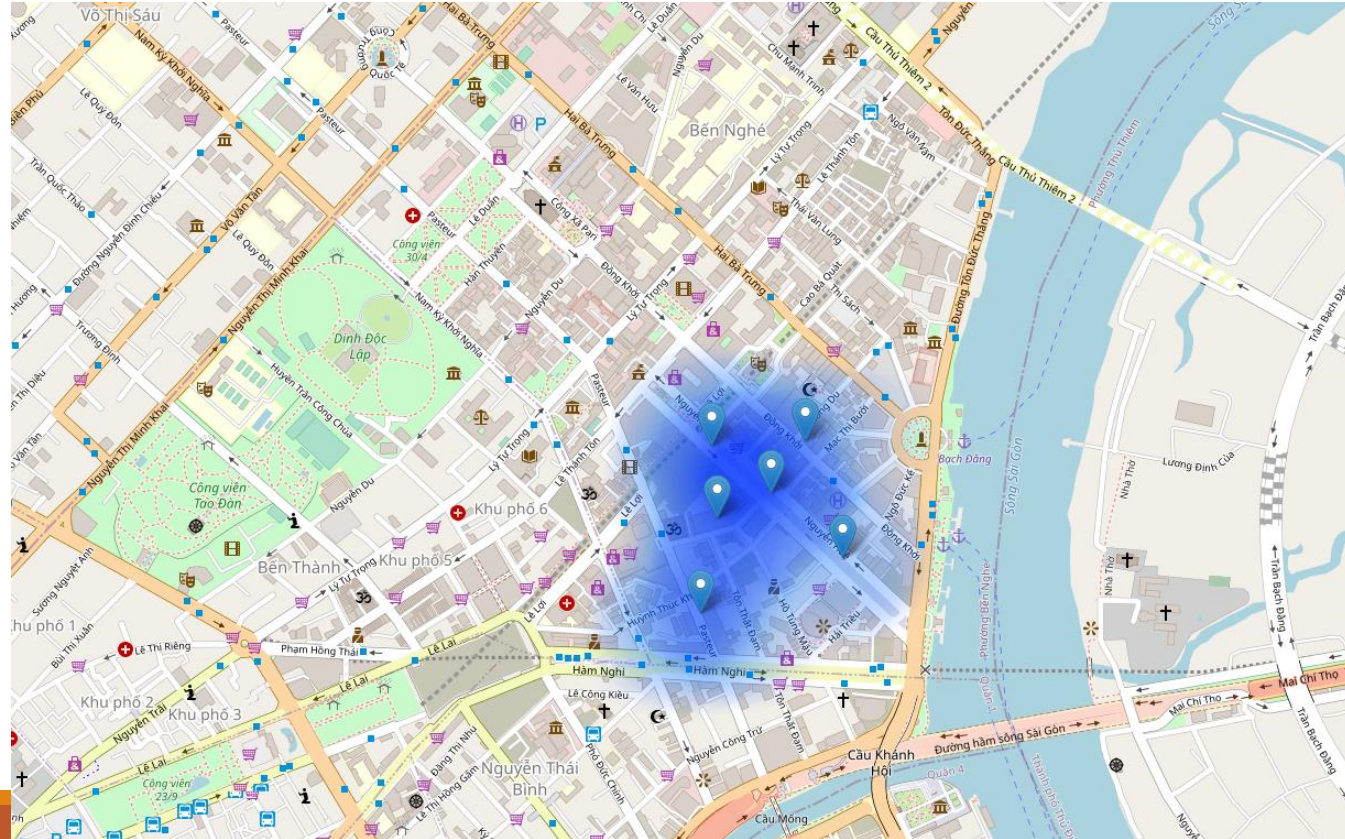# Analysis

Map showing cluster locations

Analysis

Addresses of centers of recommended areas for further exploration:

◦ Chung cư 42 Tôn Thất Thiệp, 42, Tôn Thất Thiệp, Phường Bến Nghé, Quận 1, 00084
    => 1.2km from Ho Chi Minh City center

◦ Phố đi bộ Nguyễn Huệ, Nguyễn Huệ, Phường Bến Nghé, Quận 1, 71006
    => 0.5km from Ho Chi Minh City center

◦ Phố đi bộ Nguyễn Huệ, Nguyễn Huệ, Phường Bến Nghé, Quận 1, 71006
    => 1.3km from Ho Chi Minh City center

◦ Phố đi bộ Nguyễn Huệ, Nguyễn Huệ, Phường Bến Nghé, Quận 1, 71006
    => 2.3km from Ho Chi Minh City center

◦ 86, Pasteur, Phường Bến Nghé, Quận 1, 00084
    => 2.1km from Ho Chi Minh City center

◦ Trung Nguyen Coffee, Đông Du, Phường Bến Nghé, Quận 1, 00084
    => 1.3km from Ho Chi Minh City center

Analysis

Map showing 6 addresses of centers of recommended areas

Results and Discussion

Although there are nearly 400 restaurants in Ho Chi Minh City center, according to our analysis there are pockets of low restaurant density.

We have identified 6 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues.

The recommended zones should be considered only as a starting point for more detailed analysis.

Conclusion

The purpose of this project is to identify areas close to Ho Chi Minh City center with low number of restaurants to aid stakeholders in the search for optimal location for a new Italian restaurant.

By calculating restaurant density distribution from Foursquare, we generated collection of locations which satisfy the requirements regarding existing nearby restaurants.

We clustered those locations to create major zones of interest. The addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal restaurant location should be made by stakeholders taking into consideration additional factors like attractiveness of each location, levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.