Capstone Project - The Battle of the Neighborhoods - by Alexandre Maioli

Applied Data Science Capstone by IBM/Coursera

1- Introduction: Business Problem

In our final project "the Battle of the Neighborhoods", I will perform an optimal location analysis for a japanese restaurant in the city of Philadelphia.

Although the current COVID-19 situation is a negative impact to the majority of restaurants' owners accross USA (https://www.inquirer.com/food/philadelphia-restaurant-closings-coronavirus-farmicia-mad-river-vitarellis-20200518.html), food delivery has never been in a high trend, especially with the expansion of home officing. Therefore, a group of investors would like to take this opportunity to open a japanese venue focused on a fast pace delivery system.
As Philadelphia is well Known for its diverse and excellent food scene, I will detect locations that are not already crowded with restaurants nor with specifically japanese restaurants. I will impose two conditions for its optimal placement: no japanese restaurants within 1 km nor more than other 2 restaurants in the same 250m radius. We will try to adjust the location as close as possible to the city, assuming these previous two conditions are met.

I will be using similar adjusted thought and analysis as the example provided by the Coursera Platform (ref.https://cocl.us/coursera_capstone_notebook).

2- Data

Based on the above business problem, the follwoing conditions will guide my data analysis:

- number of existing restaurants in the neighborhood (any type of restaurant) - no more than 2 within 250 meter radius.
- number of and distance to Japanese restaurants in the vicinity - no other restaurant within 1 km.
- distance of neighborhood from city center
  We decided to use regularly spaced grid of locations, centered around city center, to define our neighborhoods.

Following data sources will be extracted/generate the required information:

- centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using **geopy.geocoders**
- number of restaurants and their type and location in Philadelphia will be obtained using **Foursquare API**
- coordinate of Philadelphia Center will be obtained using **geopy.geocoders**. The center is known to be its City Hall area.

Neighborhood Candidates and Restaurant Locations in Philadelphia Map

Based on the Philadelphia distinct geography, a rectangular grid of circular neighborhoods of 250m radius will cover the area of interest (4x10 killometers rectangle centered around City Hall). Please see as below:
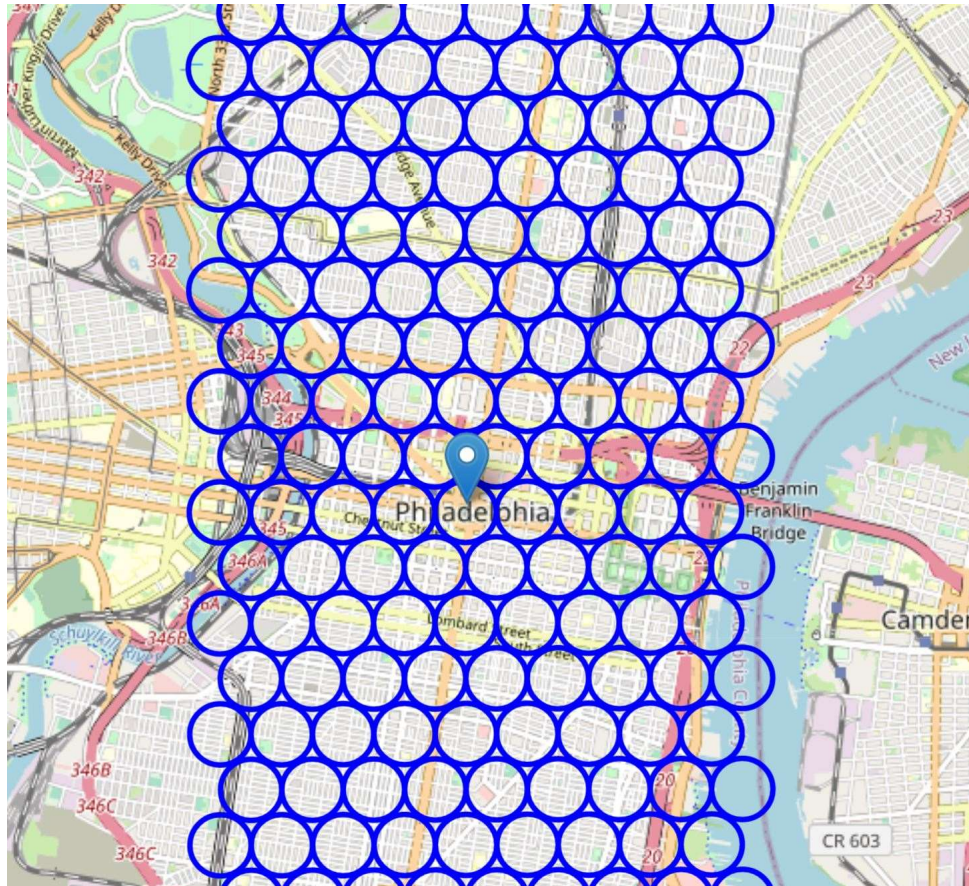
Using the folium package, 180 circular neighborhoods addresses were created. With the location data of each of the interesting neighborhoods in hands, we made use of the FOURSQUARE to start taking data to analyze my imposed conditions. Therefore, we may obtain all the restaurants and japanese ones within the circular neighborhoods.

Total number of restaurants: 740
Total number of Japanese restaurants: 51
Percentage of Japanese restaurants: 6.89%
Average number of restaurants in neighborhood: 3.6166666666666667

6.89% seems a low and satisfactory result so far. In addition, each neighborhood has only 3.6 restaurants as average within its 250m radius . In the following page, there is a good overview of the restaurant locations (blue dots for restaurants and red dots for Japanese ones).
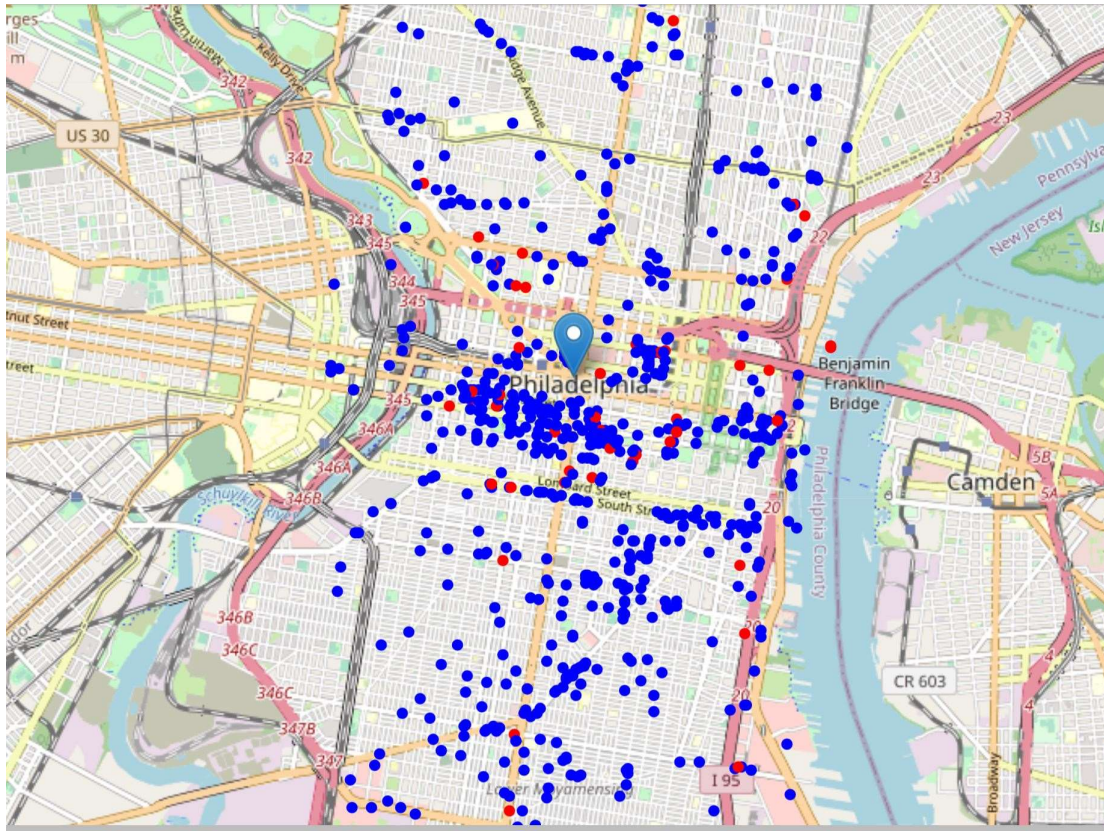
3- Methodology

As previously stated, we will focus on neighborhoods with low restaurant density, particularly those with few Japanese restaurants. I will follow the same good methodology as per Coursera example, but here limiting our analysis to area ~3km around city center.

In first step we have collected the required data: location and type (category) of every restaurant within 5km from Philadelphia center (City Hall). We have also identified japanese restaurants (according to Foursquare categorization).

Second step in my analysis will be the calculation and exploration of 'restaurant density' across different areas of Philadelphia - using heatmaps to identify a few promising areas close to center with low number of restaurants in general (and no Japanese restaurants in vicinity) and focus our attention on those areas.

In third and final step we will focus on most promising areas and within those create clusters of locations that meet BOTH of the following conditions:
1- Locations with no more than two restaurants in radius of 250 meters.
2- Locations without Japanese restaurants in radius of 1000 meters.

Finally, a map will be presented of all such cluster locations (using k-means clustering with k=5) 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.

4 – Analysis

Let's perform some basic explanatory data analysis and derive some additional info from our raw data. First let's count the number of restaurants in every area candidate and create a table (see below example for first 10 neighborhoods) with how many restaurants are located within each and its distance to nearest Japanese restaurant.

Average number of restaurants in every area with radius=250m: 3.6

Average distance to closest Japanese restaurant from each area center: 707m

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area	Distance to Japanese restaurant
0	3205, Napoli Way, Packer Park, South Philadelp...	39.912098	-75.187504	483974.176098	4.418018e+06	4970.160963	0	1565.174474
1	3110, South Uber Street, Packer Park, South Ph...	39.912108	-75.181654	484474.176098	4.418018e+06	4790.876746	3	1132.334128
2	1612, Packer Avenue, Packer Park, South Philad...	39.912117	-75.175804	484974.176098	4.418018e+06	4658.594209	1	783.957919
3	Schuylkill Expressway, Packer Park, South Phil...	39.912125	-75.169954	485474.176098	4.418018e+06	4577.390086	1	668.580183
4	Saint Maris Convent, South 10th Street, Whitma...	39.912134	-75.164104	485974.176098	4.418018e+06	4550.000000	1	882.841380
5	524, Bigler Street, Whitman, South Philadelphi...	39.912142	-75.158254	486474.176098	4.418018e+06	4577.390086	1	1269.574080
6	Oregon Market, South 3rd Street, Whitman, Sout...	39.912150	-75.152404	486974.176098	4.418018e+06	4658.594209	3	1166.180486
7	East Oregon Avenue, Whitman, South Philadelphi...	39.912157	-75.146554	487474.176098	4.418018e+06	4790.876746	1	1058.880413
8	East Oregon Avenue, Whitman, South Philadelphi...	39.912165	-75.140704	487974.176098	4.418018e+06	4970.160963	0	1175.788473
9	Brite Star Manufacturing Company, Oregon Avenu...	39.916157	-75.184590	484224.176098	4.418468e+06	4457.858230	0	1186.946589

Therefore, on average a Japanese restaurant can be found within ~700m from every area center candidate. That is close even though not as close as Berlin or New York. Nevertheless, we will later impose a condition for no Japanese restaurant within 1 km.

Let's create a map showing heatmap / density of restaurants and try to extract some meaningfull info from that. Also, I am circles indicating distance of 1km, 2km and 3km from City Hall. We can see that the north area does not have as many 'competitors' as the south zone.
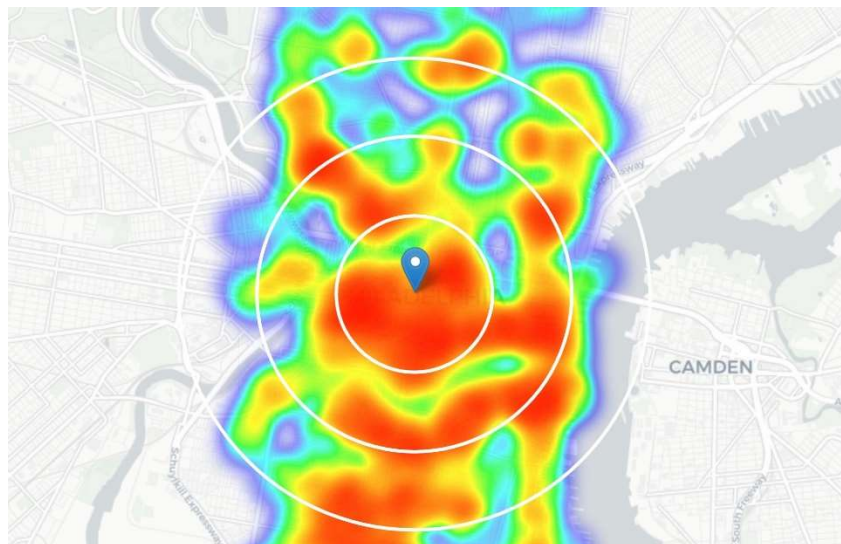
Few pockets of low restaurant density closest to city center can be found Northeast from City Hall (Philly City Center). Below, another heatmap map showing heatmap/density of japanese restaurants only. With that fact, the northeast is an interesting area to focus.
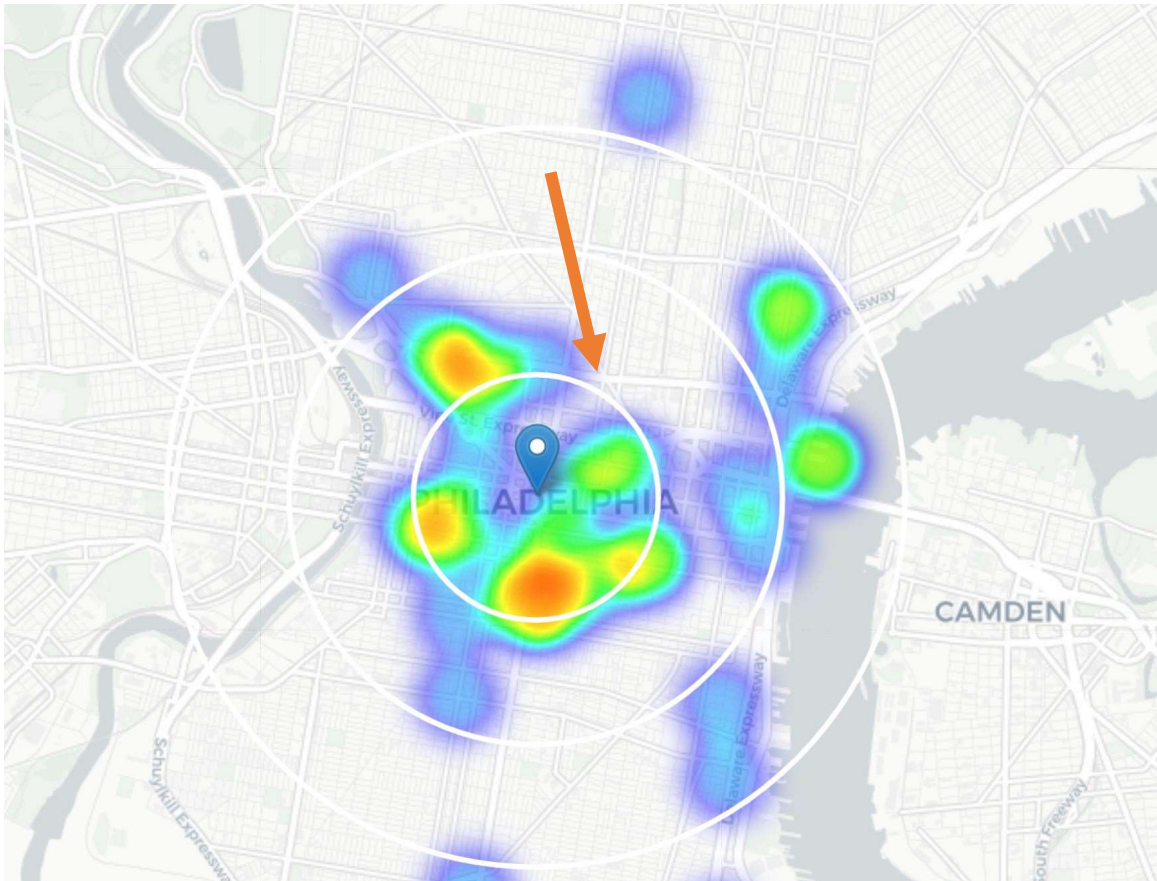

This map is definetly 'not hot' (Japanese restaurants represent a subset of ~7% of all restaurants in Philadelphia) but it also indicates higher density of existing Italian restaurants directly south and northwest from City Hall, with closest pockets of low Italian restaurant density positioned northeast from city center.

Based on this we will now focus our analysis on NORTHEAST from Philly center - we will move the center of our area of interest and reduce it's size to have a radius of 750m. This places our location candidates mostly in YORKTOWN/GIRARD/CAMBRIDGE PLAZA.

YORKTOWN/GIRARD/CAMBRIDGE PLAZA

Analysis of the sorroundings amplifies the good location of this areas to stakeholders as they are mostly residential and nearby big university campus as TEMPLE and DREXEL Univeristies. Ultimately, it is a big surplus as schools tend to go online for at least a year prompting the demand of food delivery. In addition, its relatively close to city center in walking distance. This borough appear to justify further analysis.

A new and narrower region of interest will be created, which will include low-restaurant-count parts of YORKTOWN/GIRARD/CAMBRIDGE PLAZA, closest to City Hall.
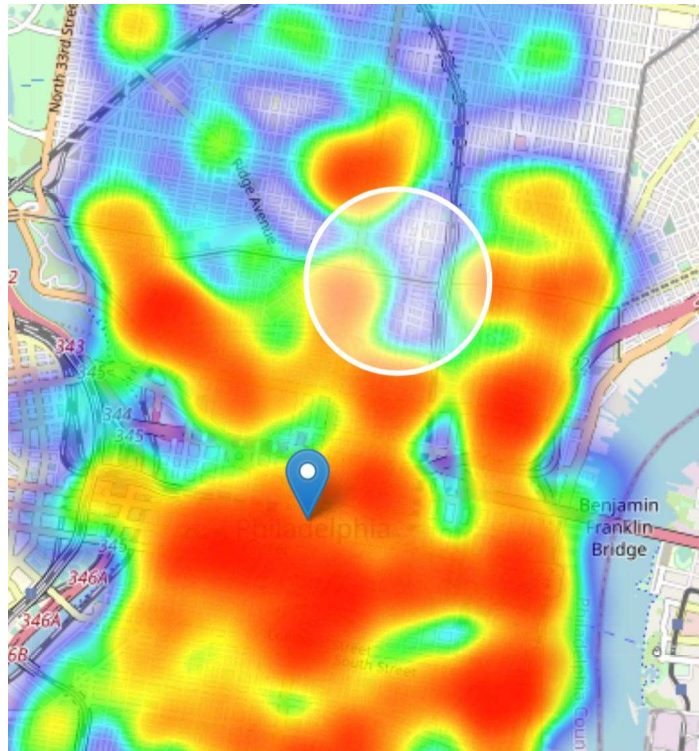
This above area nicely covers all the pockets of low restaurant density in an area of YORKTOWN/GIRARD/CAMBRIDGE PLAZA, closest to City Hall. A new, denser grid of location candidates restricted to our new region of interest (let's make our location candidates 100m appart).
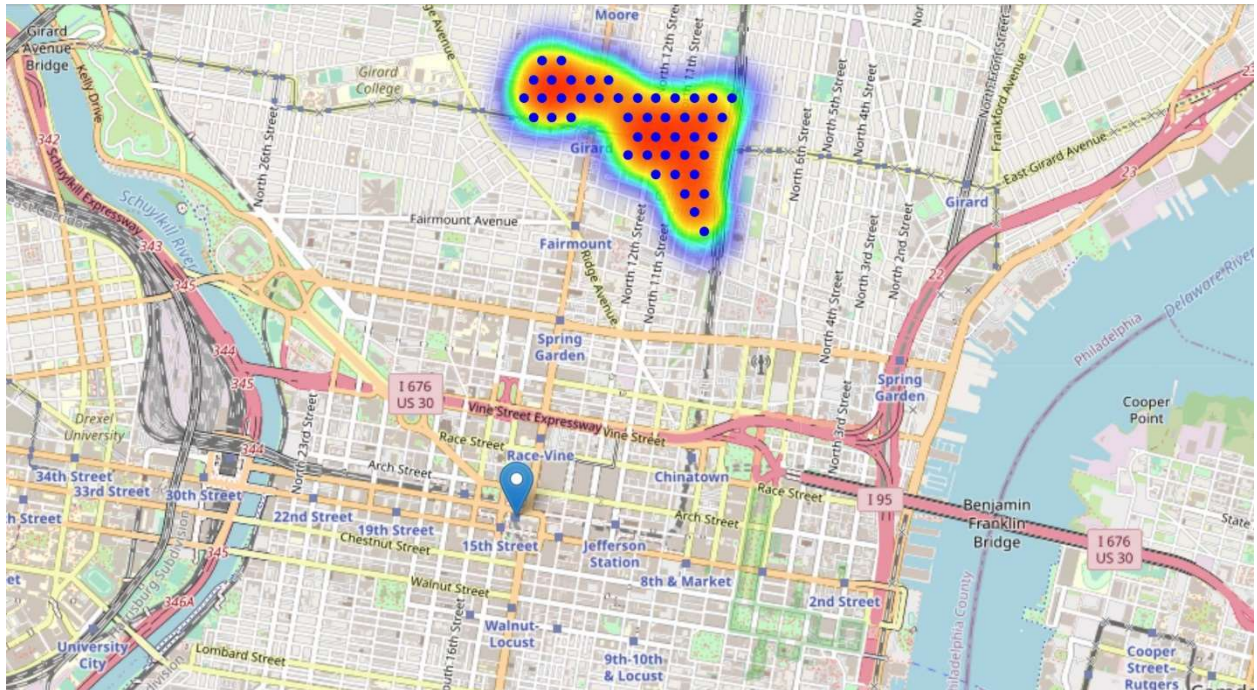
Locations with no more than two restaurants nearby (250m): 127
Locations with no Japanese restaurants within 1000m: 49
Locations with both conditions met: 45
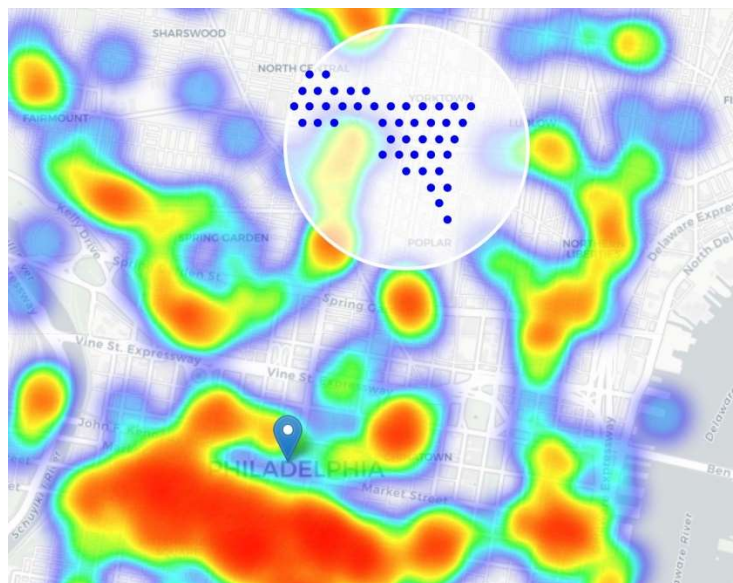
We now have locations fairly close to City Hall (near Girard, mostly south of Yorktown). Let us now show those good locations in a form of heatmap:


There is clear indication of zones with low number of restaurants in vicinity, and no Japanese restaurants at all nearby.

Let us now cluster those locations to create 5 centers of zones containing good locations. Those zones, their centers and addresses will be the final result of our analysis.
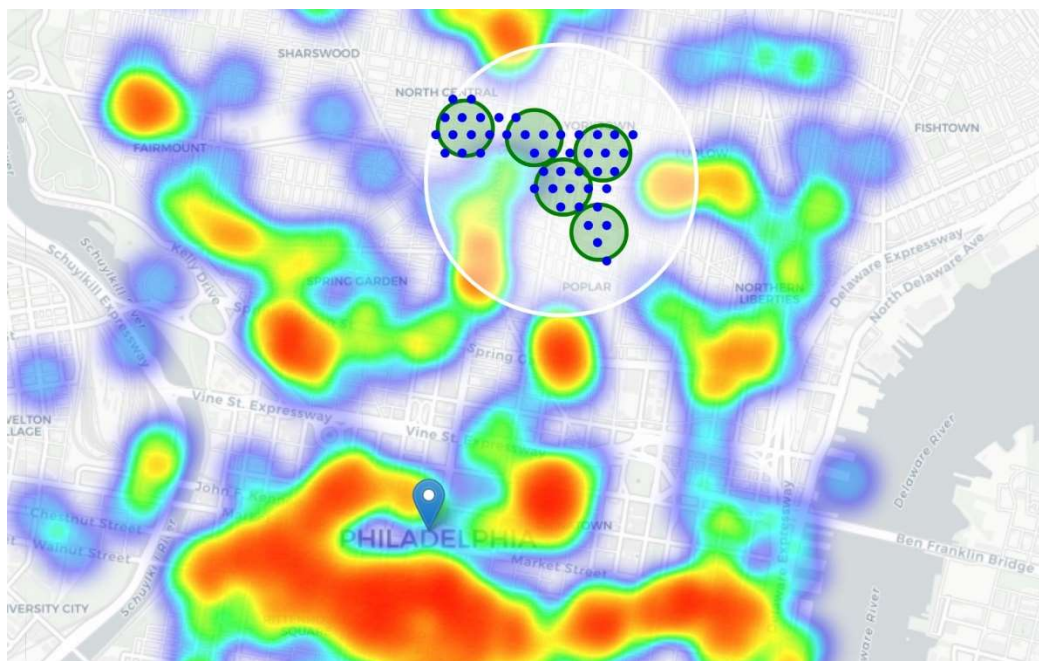
Addresses of those cluster centers are a good starting point for exploring the neighborhoods to find the best possible location based on neighborhood specifics.
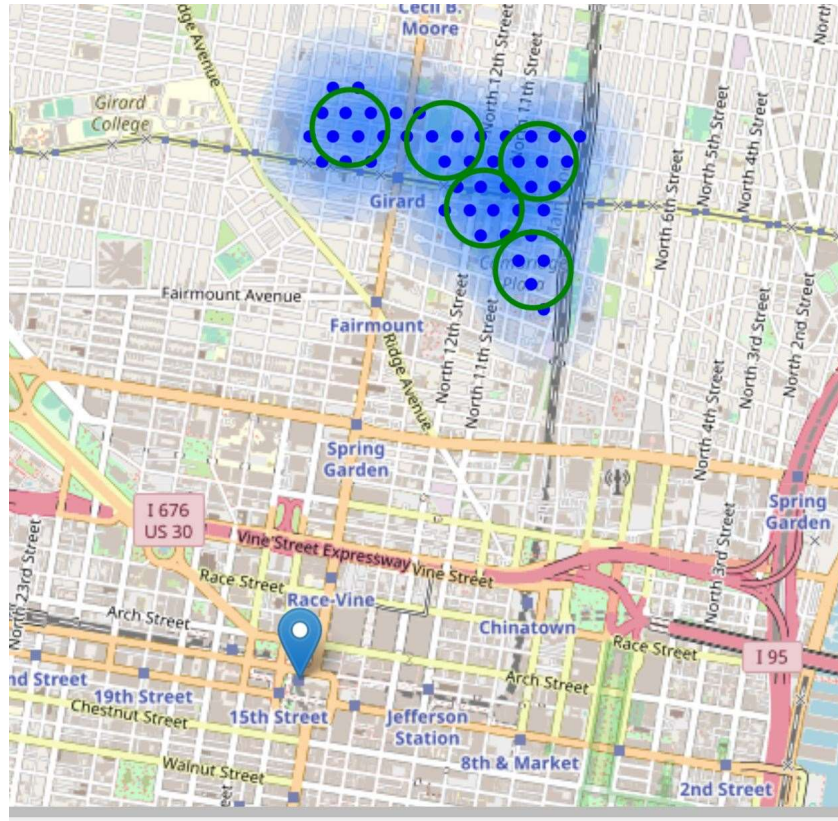

Our analysis is concluded with the below 5 addresses representing centers of zones containing locations with low number of restaurants and no japanese restaurants nearby, all zones being fairly close to city center (all less than 3km from City Hall):

```
================================================================
Addresses of centers of areas recommended for further analysis
================================================================
    ⇨ Temple University Sports Complex, North 13th Street, Cambridge Plaza,
      19133,
    ⇨ ST Malachy Catholic School, West Flora Street, Yorktown, 19122
    ⇨ 1559 West Cabot Street, North Central, 19121
    ⇨ 901 North 10th Street, Harrison Plaza, 19123
    ⇨ Girard Plaza, West Harper Street, Yorktown, 19123
```

These addresses should be considered only as a starting point for exploring area neighborhoods in search for potential restaurant locations. Most of the zones are located in Girardi/Yorktown/Poplar, which we have identified as interesting due to being popular with university students, home owners and to being fairly close to city center and well connected by public transport.
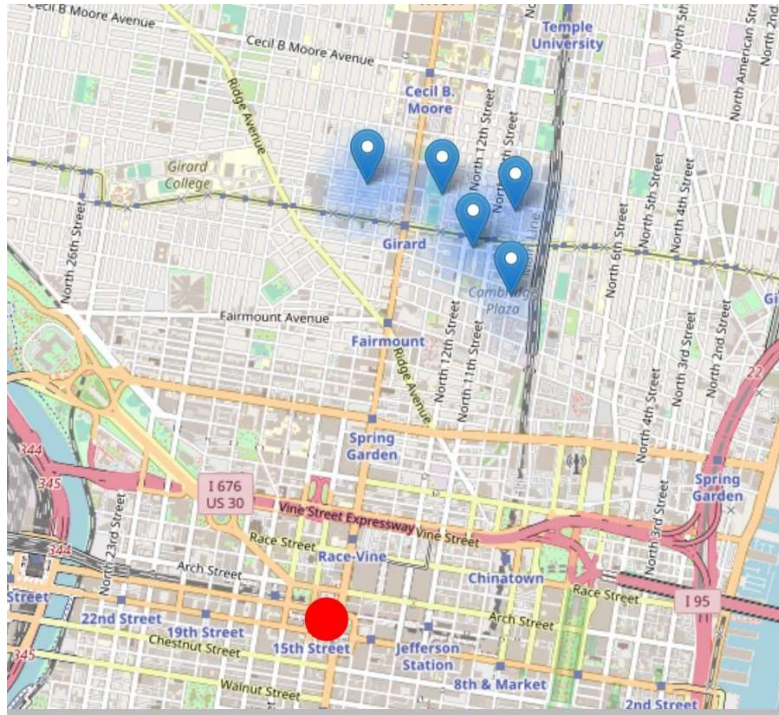
5- Results and Discussion

My analysis shows that there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected south from City Hall, so we focused our attention to areas northeast, corresponding to GIRARDI, YORKTOWN and CAMBRIDGE PLAZA. These addresses should be considered only as a starting point for exploring area neighborhoods in search for potential restaurant locations. Most of the zones are located in Girardi/Yorktown/Poplar, which we have identified as interesting due to being popular with university students, home owners and to being fairly close to city center and well connected by public transport.

Those location candidates were then clustered to create 5 zones of interest which contain greatest number of location candidates. The conditions initially stated in the problem were imposed: **locations with no more than two restaurants in radius of 250 meters**, and **no Japanese restaurants in radius of 1000 meters**. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Also good to point out that this was a simple analysis and should be considered a starting point for further development of a location search for a japanese restaurant.

6- Conclusion

This project objective is to identify Philadelphia areas close to center with low number of restaurants (particularly Japanese restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Japanese restaurant, taking as a guide and reference the Coursera example from Berlin.

Optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, 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.