Comparing neighborhoods in New York and Toronto for a new Italian Restaurant

Saswat Sahoo

April 4, 2020

1. Introduction

1.1 Background

New York and Toronto are financial capitals of their respective countries, with a vibrant food scene, nightlife, numerous tourist hotspots and hosting world class universities with a sizable student population. They are perfect destinations for a restaurant chain to consider while opening a new chain.

1.2 Problem

Stakeholders wanted to figure out which localities and neighborhoods would suit better for opening a new Italian restaurant. With a wide variety and quantity of restaurants already populating the map of both the cities, it is imperative to use data analysis to create clusters of areas with low restaurant densities, particularly Italian, while considering a new location expanding. The stakeholders most importantly wanted to compare these chosen neighborhoods between New York and Toronto.

2. Data acquisition and cleaning

2.1 Data sources

Location coordinates were acquired using Google Maps API. Foursquare API was used to get nearby restaurants and other facilities. Toronto and New York restaurants around Times Square and Downtown Toronto were collected using the above.

2.2 Data cleaning

Data downloaded or scraped from multiple sources were combined into one dataframe. Since most data was accumulated from scratch from location APIs, there was no issue of missing data. New columns were created as our analysis progressed. Lastly, dataframe was reduced to contain only those locations with low number of restaurants, particularly Italian.

2.3 Feature selection

For New York we mined information of about 2565 restaurants in a radius of 1000m from city center. For Toronto we mined about 1244 restaurants. Italian restaurants constituted about 377 or around 14.7% of all restaurants in New York. In Toronto, they made up around 121 or 9.37%. Initially the dataframe had addresses, Latitude & Longitudes values and distance from city centers. In the next step number of restaurants in the area was added to the dataframe. Next distance from nearest Italian restaurant is calculate.

3. Exploratory Data Analysis

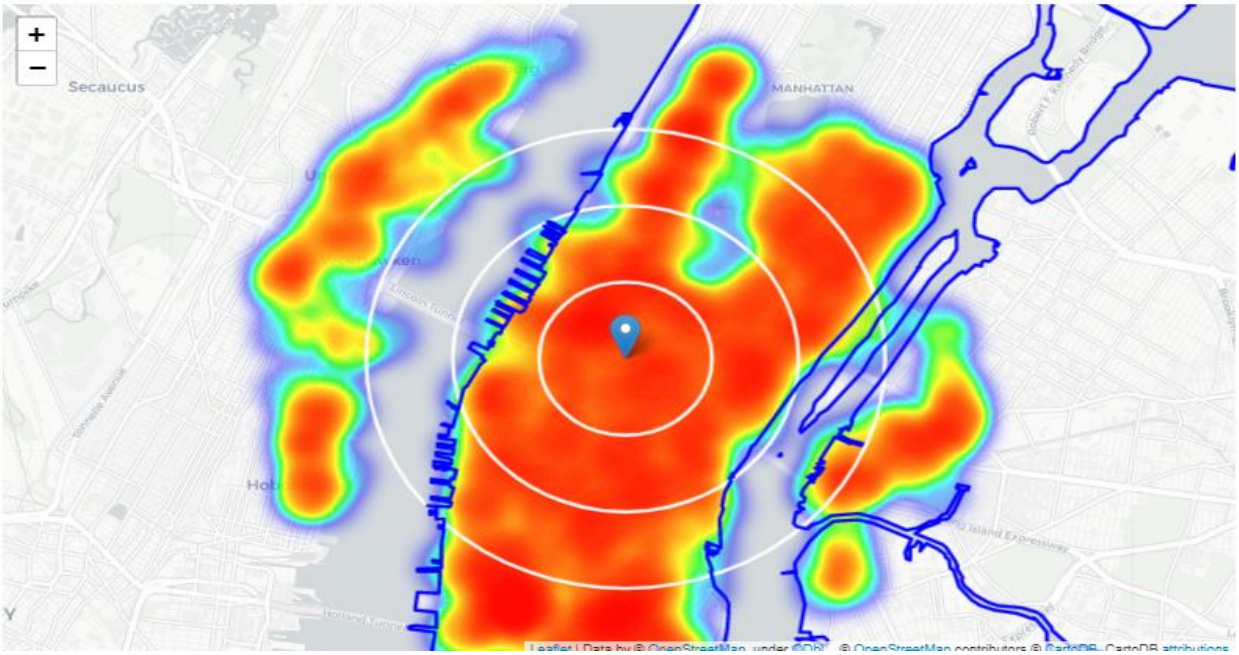
3.1 Calculation of target variable

Distance from city center was calculated from respective city centers. Further, number of restaurants in the neighborhood areas was calculated. Going forward distance from nearest Italian restaurant was also calculated. Finally, the dataframe was reduced to contain only those locations which did not have any restaurant in a radius of 250m and no Italian restaurant within 400m.

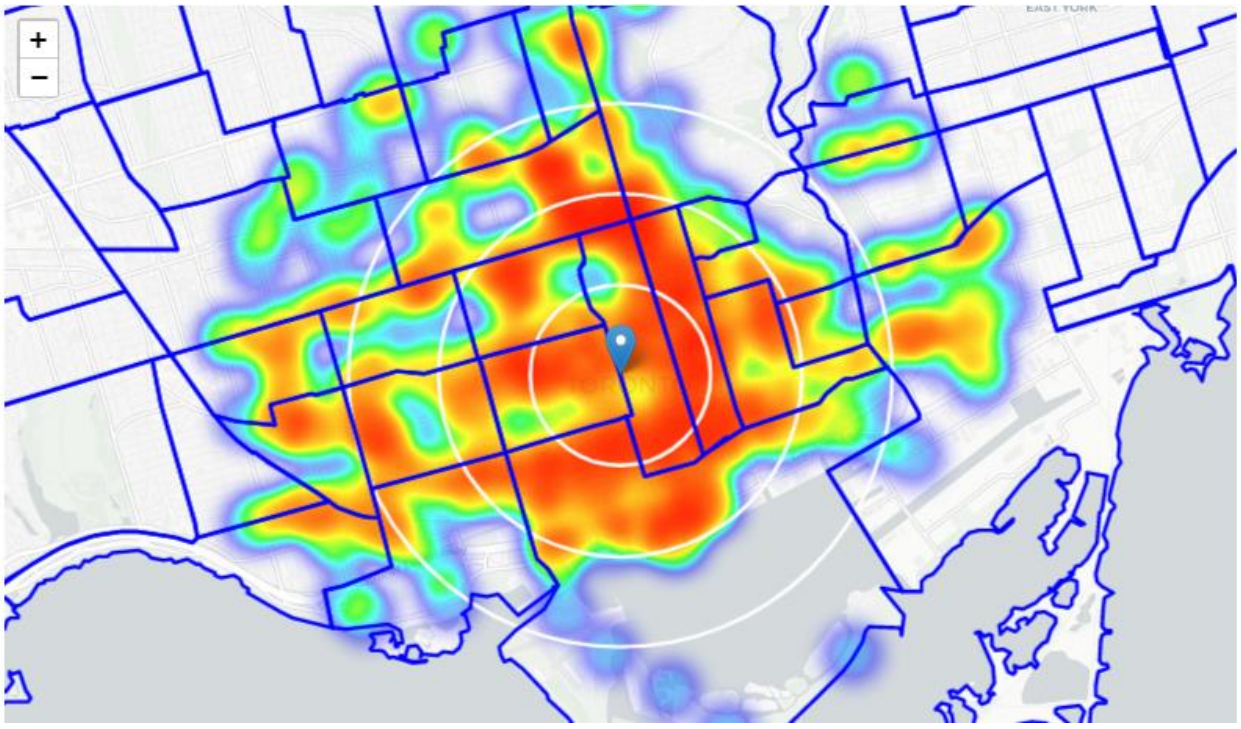
3.2 Difference between New York and Toronto by restaurants

New York is more densely populated and has higher per capita income than Toronto. Hence, intuitively it was expected and later found through analysis that there is higher restaurant density, as well as Italian, in New York than Toronto.

New York



Toronto



3.3 Difference between New York and Toronto by Italian restaurants

Heat map of Italian restaurants in New York and Toronto

New York

This map is not so 'hot' (Italian restaurants represent a subset of ~15% of all restaurants in New_York) but it also indicates higher density of existing Italian restaurants in all of Manhattan with closest pockets of **low Italian restaurant density positioned north-east and south-east from city center inside Manhattan and large patches of low density areas in North west and North East outside Manhattan**.

Based on this we will now focus our analysis on low density boroughs **Colombus Circle, Hell's Kitchen and Greenwich Village. **Colombus Circle** is a heavily trafficked busy intersection. **Greenwich Village** is a hub of popular cafes, bars and restaurants and **Hell's Kitchen** is condsidered home of actors and arts organizations where tourists from Times Square and workers from office high-rises jam the international restaurants, bars and pubs along 8th and 9th avenues.

Lenox Hill is not so popular among tourists or visited by locals. **Tudor City** is a residential area.**Koreatown** is heavily populated by Asians and Korean residents and restaurants. Hence they were not of any paticular interests to stakeholders

Colombus Circle, Hell's Kitchen and Greenwich Village

Analysis of popular travel guides and web sites often mention Colombus Circle, Hell's Kitchen and Greenwich Village as busy, heavily visited, tourist heavy New York neighborhoods popular with visitors and locals alike.

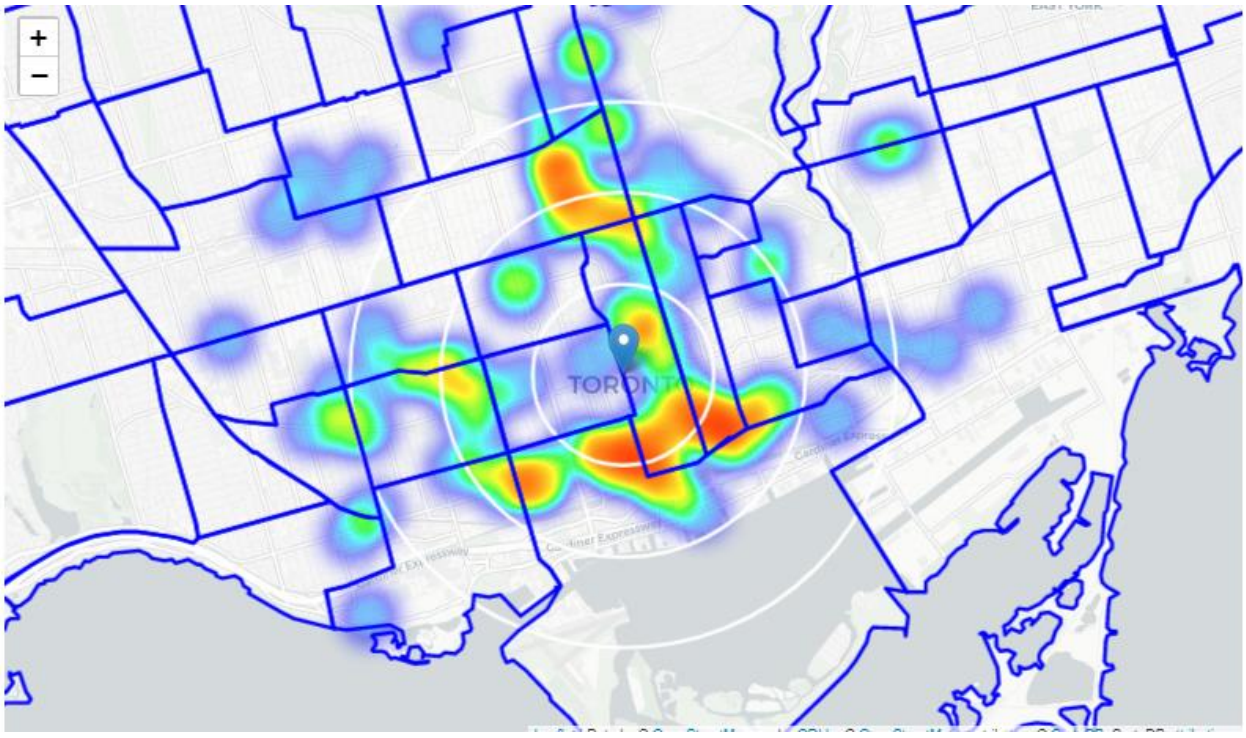
*"Like Times Square, Union Square and other bustling intersections found throughout New York City, Columbus Circle is full of zooming cars and quick-footed pedestrians. As one of the more

heavily trafficked areas in Manhattan, it's a noteworthy tourist attraction and busy commercial center, located at the southwest corner of Central Park. There are multiple places worth visiting nearby: the Time Warner Center, Jazz at Lincoln Center, the Shops at Columbus Circle, and the newly opened subterranean mall, TurnStyle. With such a wide array of retail shops and stunning landmarks, the traffic circle is much more than a channel to ferry cars and the occasional horse-pulled carriage through NYC's rush hour gridlock.(untappedcities.com)

"Hell's Kitchen is a place to do it up. There's cuisine from around the globe: Argentinean empanadas, Japanese ramen, Druze bourekas, Thai street food. There's all kinds of nightlife: jazz haunts, LGBTQ+ clubs, dive bars, rooftop lounges. Theater and dance thrive a stone's throw from Broadway. The High Line's northern entrance is at the rail yards on this neighborhood's southern tip. And by the glorious Hudson River Park, there's the Intrepid Sea, Air & Space Museum and the Circle Line Sightseeing Cruises departure point, so you can tour a ship or set sail on one."(nycgo.com)

*"Greenwich Village, NYC is home to a bevy of beloved institutions, dives and eateries. The spirit of rock & roll is alive at Cafe Wha?, where Jimi Hendrix and Bob Dylan used to jam. Laughter from comedy clubs can be heard on popular blocks like MacDougal Street. And folks with a case of the munchies can easily cure their sweet and salty cravings thanks to a bevy of excellent restaurants. Take a stroll down Washington Square Park—one of the best NYC parks—and see where the bohemian capital takes you. (timeout.com)

Toronto



This map is not so 'hot' (Italian restaurants represent a subset of ~10% of all restaurants in Toronto) but it also indicates higher density of existing Italian restaurants in South and North of Toronto with closest pockets of **low Italian restaurant density positioned north-west and east from city center**.

Based on this we will now focus our analysis on low density boroughs **The Annex and Liberty Village**.**The Annex** draws students to its casual eateries, indie bookshops, bars and cafes on Bloor Street West. **Liberty Village** is a hub of bars, restaurants, shops and services.

Cityplace, Rosedale and Ossington is not so popular among tourists or visited by locals as they are residential areas. **Koreatown** is heavily populated by Asians and Korean residents and restaurants. Hence, they were not of any paticular interests to stakeholders

The Annex and Liberty Village

Analysis of popular travel guides and web sites often mention The Annex and Liberty Village as beautifull, interesting, rich with culture, 'hip' and 'cool' Toronto neighborhoods popular with tourists, students and locals alike.

*"In addition to brunch, beer and patio options to choose from, Liberty Village is home to a wide variety of restaurants representing many cuisines and dietary preferences, depending on what you're in the mood for. (tripsavvy.com)

"The Annex is one of Toronto's oldest and most prominent neighborhoods; it is most recognized for its turn-of-the-century housing plus its vicinity to the University of Toronto. But outside of the residential area, on the main streets, is a wealth of independent restaurants and businesses that are rooted in a bohemian lifestyle, which the area maintains."(theculturetrip.com)

*"With its tree-lined streets, great indie stores and student energy, the Annex - bordered in the north by Dupont, in the south by Bloor and by Bathurst and Avenue Road on the west and east respectively - is a dynamic neighbourhood with a kinetic nightlife and vibrant street scene. (nowtoronto.com)

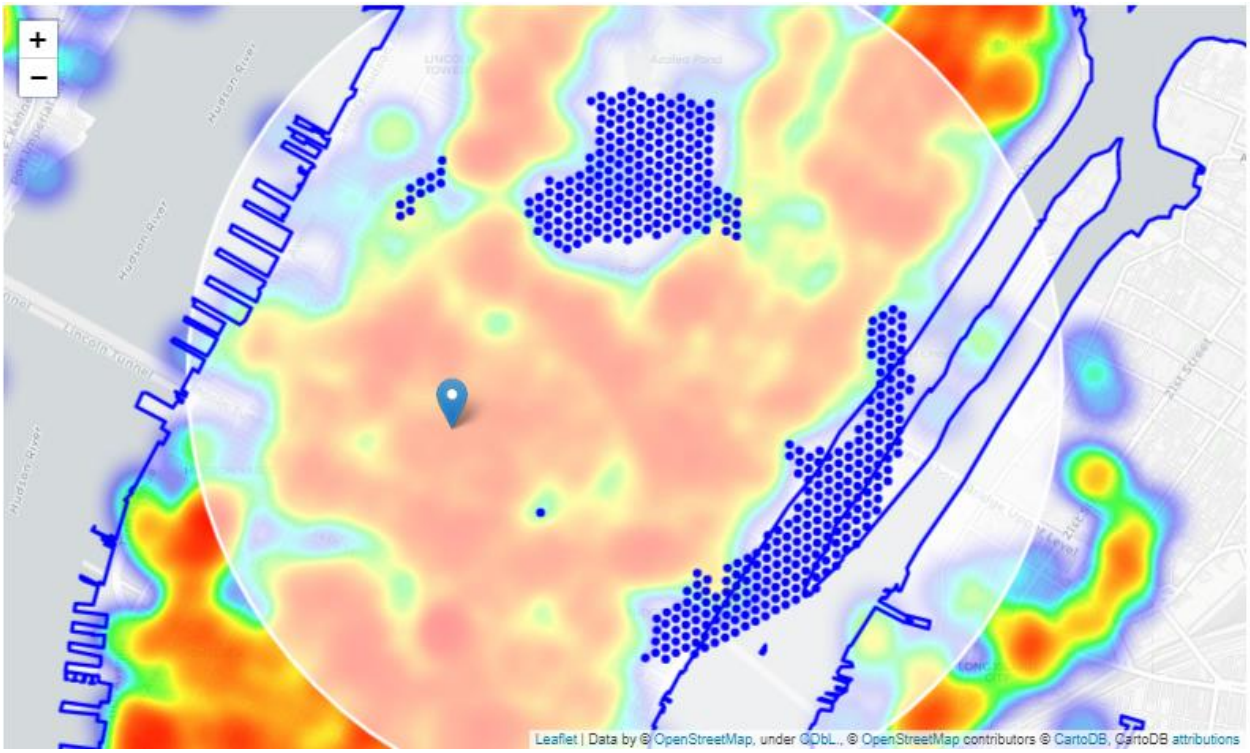
"Liberty Village is a 43-acre master-planned community,combining res- idential,commercial and retail uses.This neighbourhood is attracting young professionals in media, high-tech and design businesses who live and work in the urban core." (torontoneighbourhoods.net)

Popular with tourists, alternative and bohemian but booming and trendy, relatively close to city center and well connected, those boroughs appear to justify further analysis.

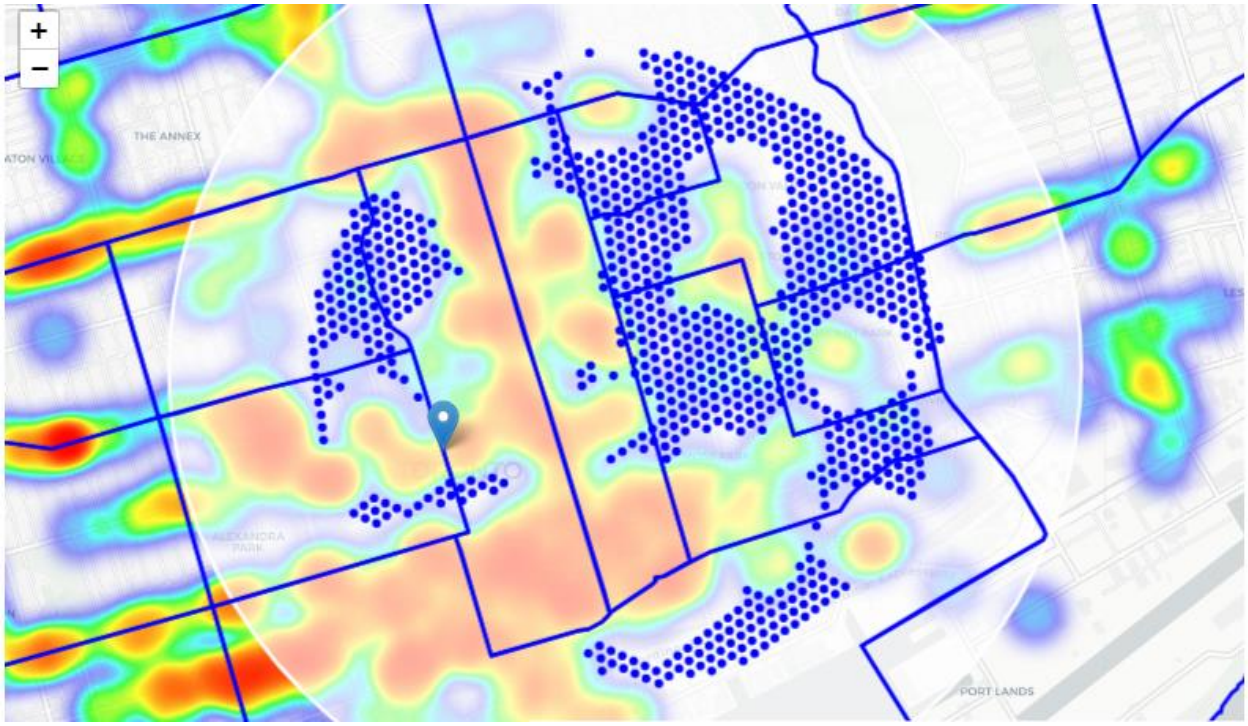
Let's define new, more narrow region of interest, which will include low-restaurant-count parts of The Annex and Liberty Village closest to Downtown Toronto.

3.4 Selected locations in low density spots plotted

New York

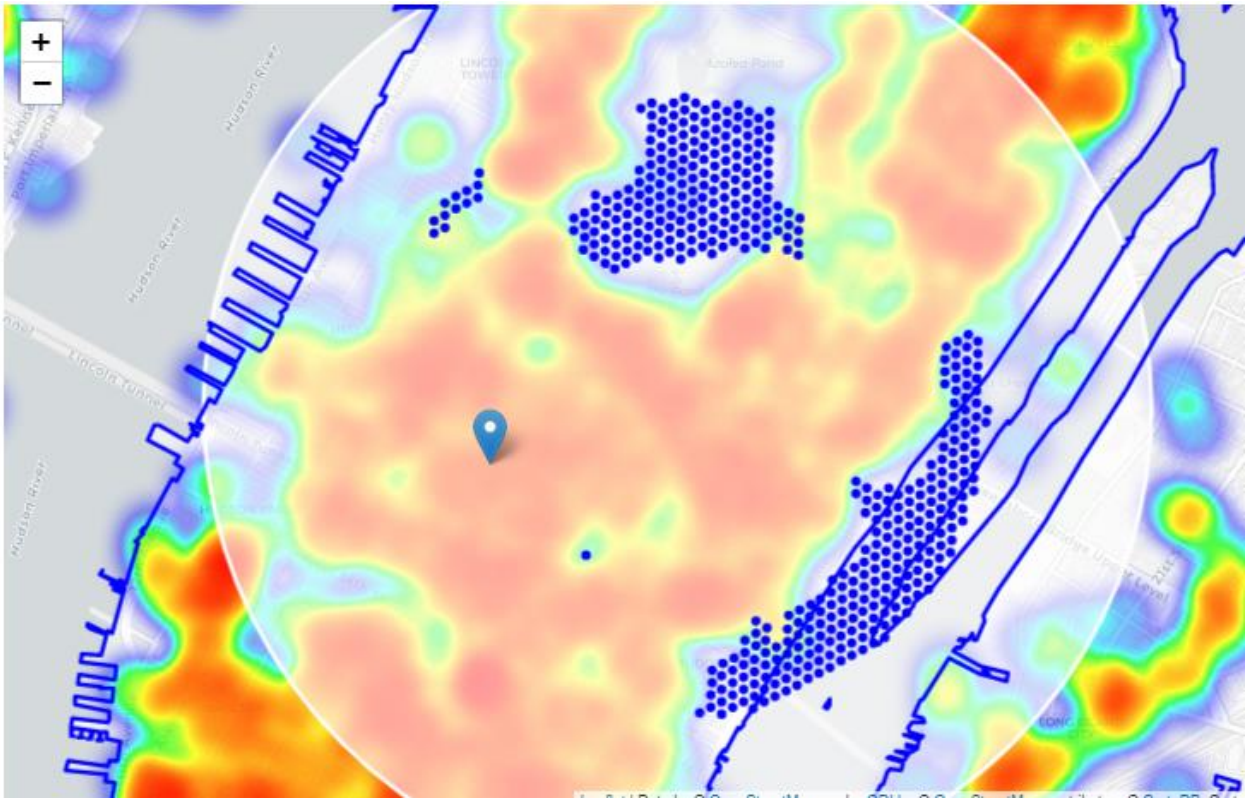


Toronto

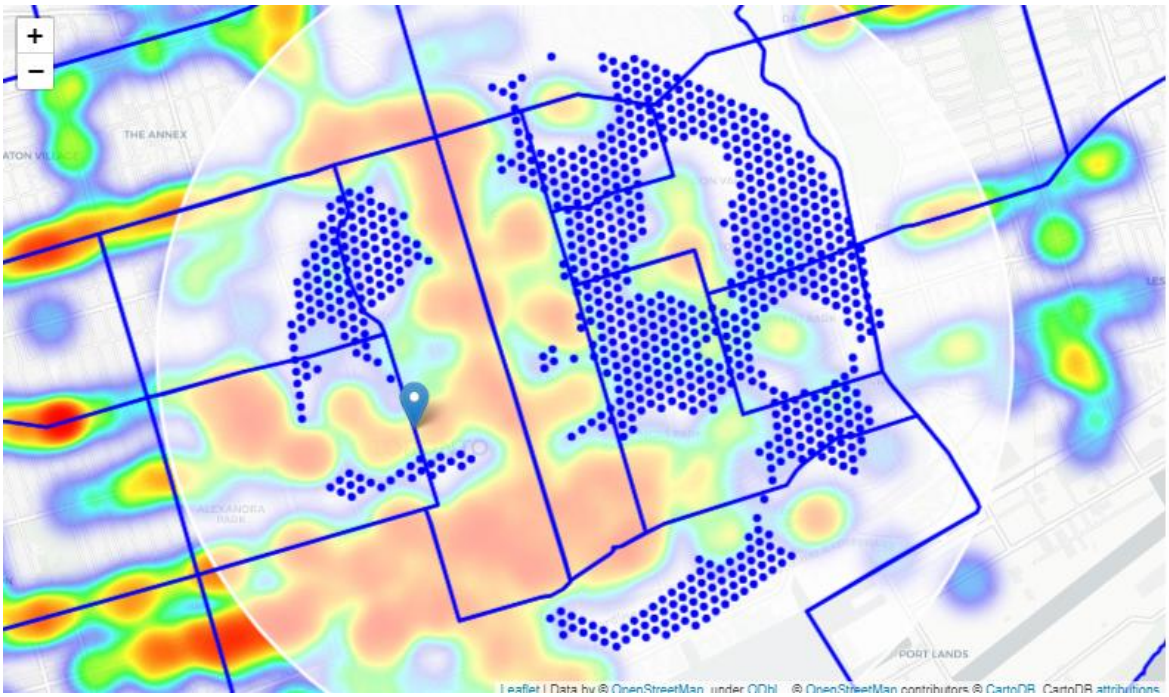


3.4 Locations with no more than two restaurants in radius of 250 meters, and no Italian restaurants in radius of 400 meters.

New York



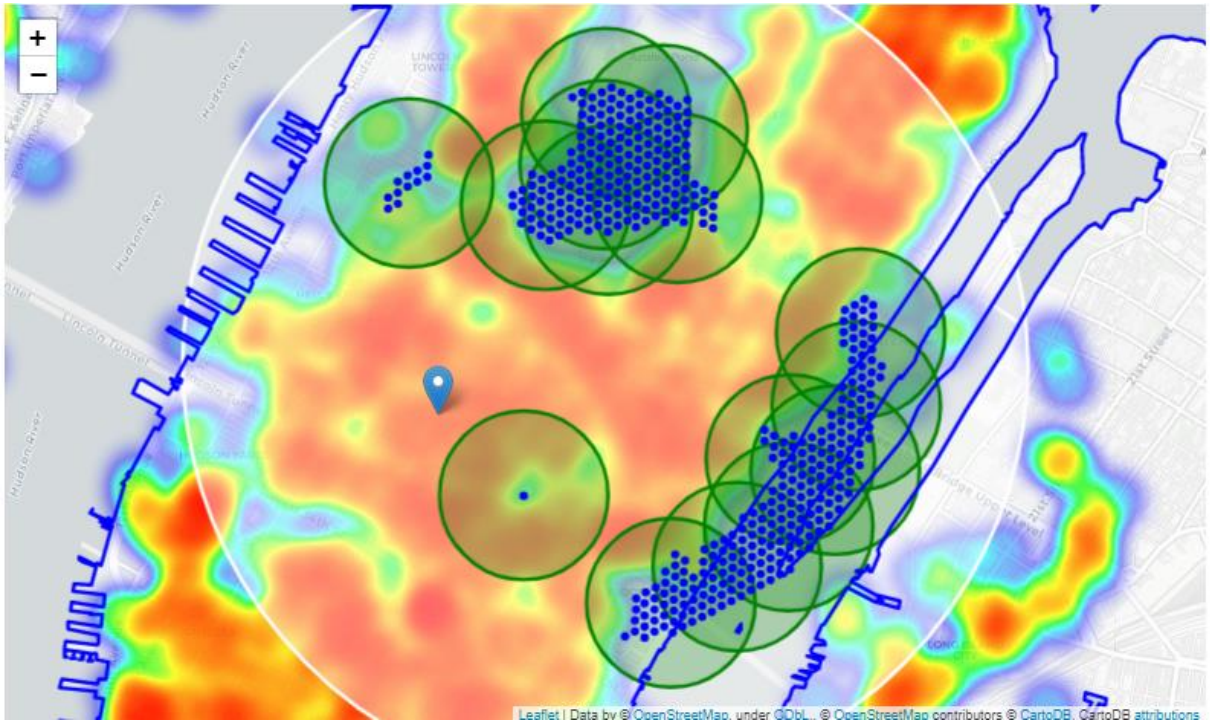
Toronto



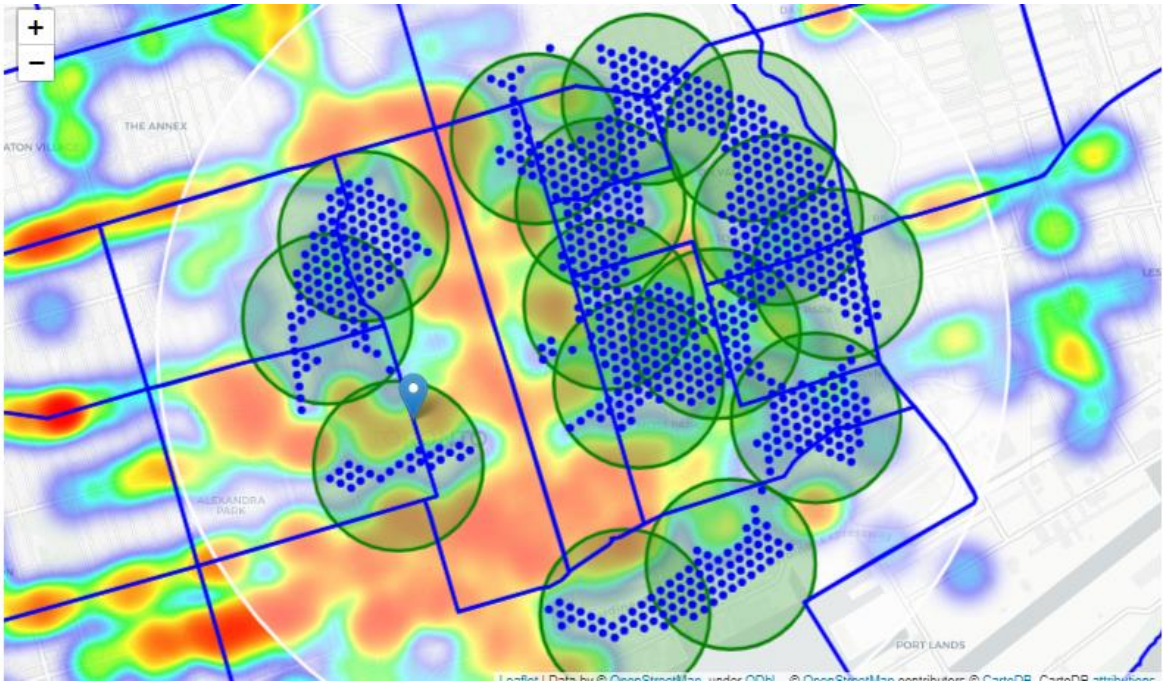
4. K-Means Cluster Modeling

The selected locations are **clustered** to create **centers of zones containing good locations**. Those zones, their centers and addresses will be the final result of our analysis.

New York



Toronto



Our clusters represent groupings of most of the candidate locations and cluster centers are placed nicely in the middle of the zones 'rich' with location candidates.

Addresses of those cluster centers will be a good starting point for exploring the neighborhoods to find the best possible location based on neighborhood specifics.

This concludes our analysis. We have created 15 addresses each from Toronto and New York representing centers of zones containing locations with low number of restaurants and no Italian restaurants nearby, all zones being close to city center. Although zones are shown on map with a radius of ~500 meters (green circles), their shape is very irregular, and their centers/addresses should be considered only as a starting point for exploring area neighborhoods in search for potential restaurant locations.

Results and Discussion

Our analysis shows that although there is a great number of restaurants in **Toronto**, there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected north, east and south from Downtown Toronto, so we focused our attention to areas north-west and south-west, corresponding to boroughs The Annex and Liberty Village. Other boroughs were also identified as potentially interesting, but our attention was focused on The Annex and Liberty Village which offer a combination of popularity among tourists, closeness to city center, strong socio-economic dynamics and a number of pockets of low restaurant density.

Our analysis shows that although there is a great number of restaurants in **New York**, there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected east and west from Times Square, so we focused our attention to areas south, south-east and north-west, corresponding to boroughs Colombus Circle, Hell's Kitchen and Greenwich Village.

Other boroughs were also identified as potentially interesting, but our attention was focused on Colombus Circle, Hell's Kitchen and Greenwich Village which offer a combination of tourist heavy traffic, closeness to city center, strong socio-economic dynamics and a number of pockets of low restaurant density.

New York has a higher a overall restaurant and Italian restaurant density that is higher than that of Toronto. But NYC also has a higher tourist footfall, higher per capita income and higher overall population and population density than Toronto. Therefore, the stakeholders should take these fatcors into consideration while making their decisions to set up new Italian restaurants.

"Expatistan, a resource for comparing the cost of living in different cities, shows that living in Toronto is 29% cheaper than in New York City, based on 10,000 prices entered by 2,392 people. In almost all categories, from food to entertainment, Toronto ranks as the more affordable place for expats and tourists alike" (narcity.com)

After directing our attention to this narrower area of interest we first created a dense grid of location candidates (spaced 100m apart); those locations were then filtered so that those with more than two restaurants in radius of 250m and those with an Italian restaurant closer than 400m were removed.

Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. 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.

Result of all this is 15 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. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to city centers but not crowded with existing restaurants (particularly Italian) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

5. Conclusions

Purpose of this project was to identify Toronto and New York areas close to center with low number of restaurants (particularly Italian restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Italian restaurant. By calculating restaurant density distribution from Foursquare data we have first identified general boroughs that justify further analysis, and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centers were created to be used as starting points for final exploration by stakeholders.

Final decision on 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 (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.