Coursera capstone project: Where to open a new chain restaurant

Business problem

As the owner of a famous chain restaurant in China, Mr. Lee is considering opening a brand-new oversea restaurant in Canada. He narrowed his choices of the locations into 3 major cities in Canada: Vancouver, Toronto and Montreal. All these 3 cities have a lot of Chinese immigrants and students, and show diversified development in almost all aspects, including food culture. They all contain a diverse array of cuisines which appeal to their diverse residents, tourists and business travelers. As a prospective restauranteur, Mr. Lee would be interested in knowing the optimal city and specific location for opening a new restaurant.

Due to the level of competition and density of the city, it would be worthwhile for restauranteurs to consider starting a restaurant where there are a high number of restaurants present from other cuisines, but few of the same cuisine to be in direct competition with.

In this project I want to analyze what kind of relevance does a location has for the restaurant's popularity and how can a person choose a perfect place for its new business. Location can be analyzed either by which neighborhood the venue belongs to or by what kind and how many other nearby businesses are. I will focus on the second part, which is quite important because by having more popular restaurants around, the competition increases, however your restaurant gets a bigger exposure and possibly more new customers.

Data

City geospatial data will be used to inform the neighborhoods to be analyzed. It contains a list of Neighborhoods with their Borough, Latitude and Longitude.

The FourSquare API will be used to provide details about surrounding venues in the neighborhoods obtained from the city geospatial data. It will provide a list of venues with their respective categories, latitudes and longitudes. They should include venues in restaurant categories which will be used as a data input for this investigation.

Required data:
- Types of venues
- Location of restaurants
- Names/ids of restaurant
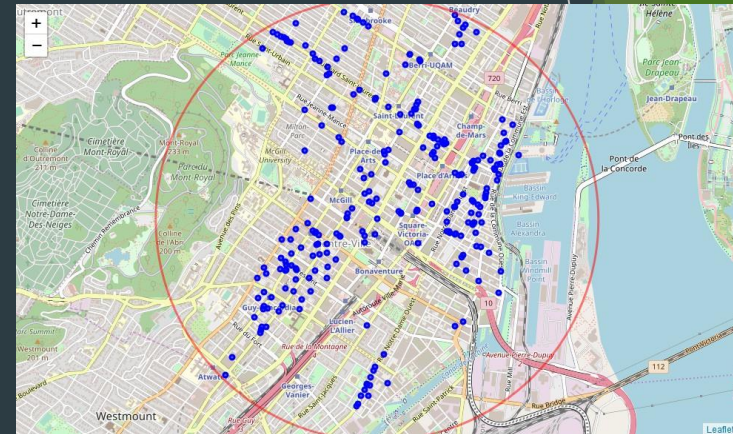- Number of likes

Exploration of 3 city centers

- Primary exploration looks at city centers within 2 km radius
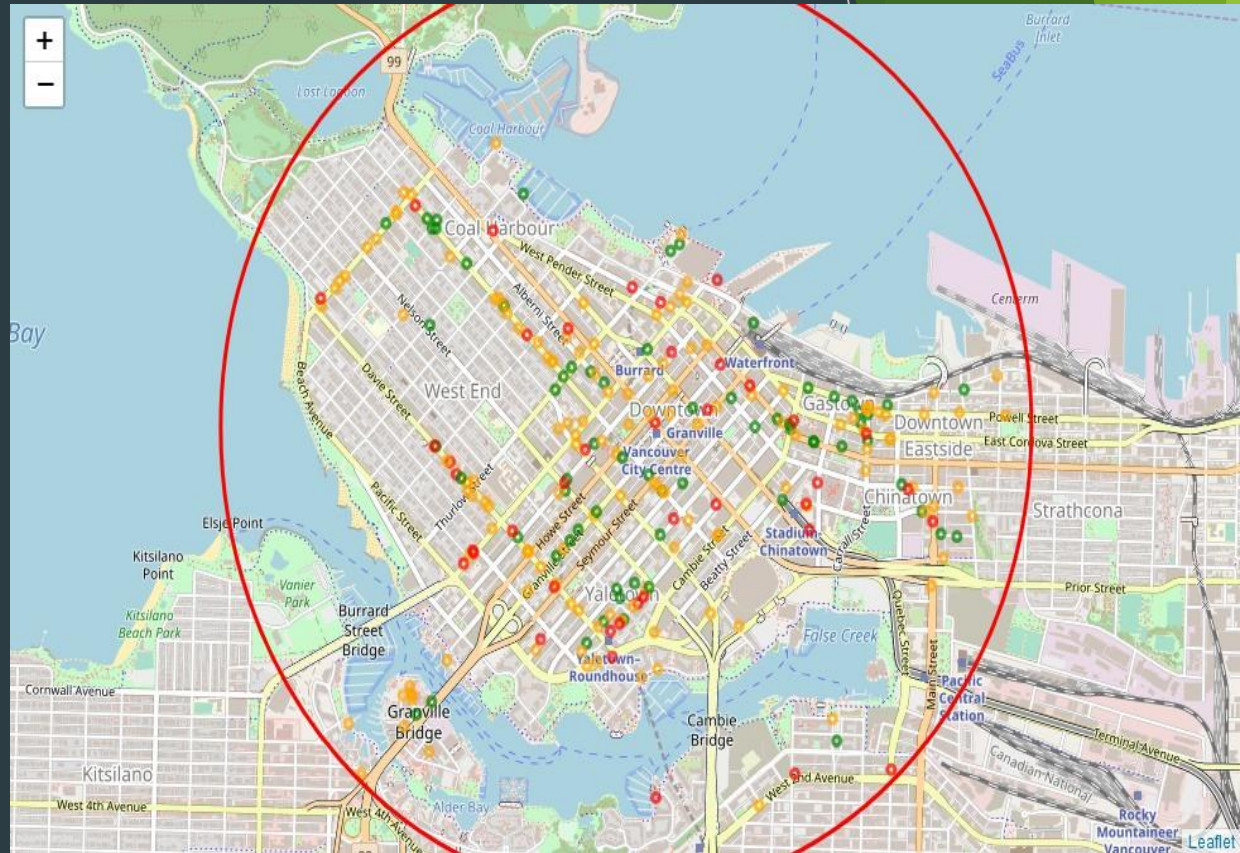- Restaurants are clustered into certain areas
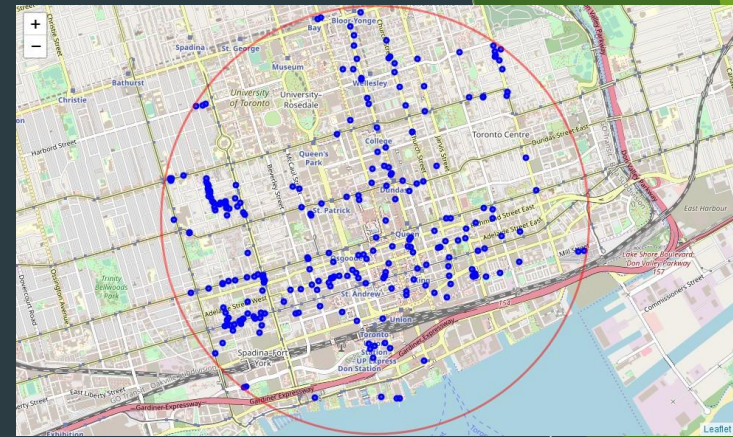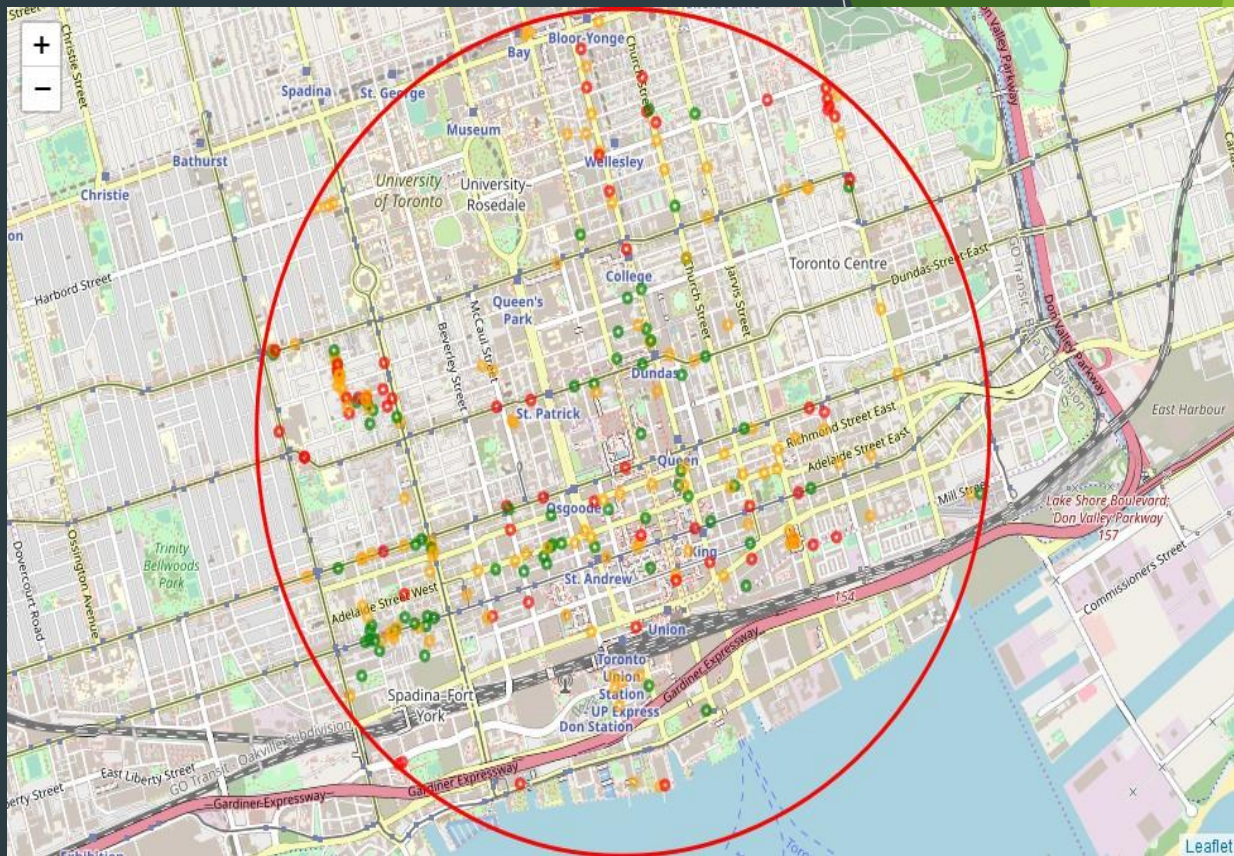

Toronto

Vancouver

Montreal

"Likes" map

- red - relatively low number of likes
- orange - medium number of likes
- green - relatively large number of likes

Vancouver

"Likes" map
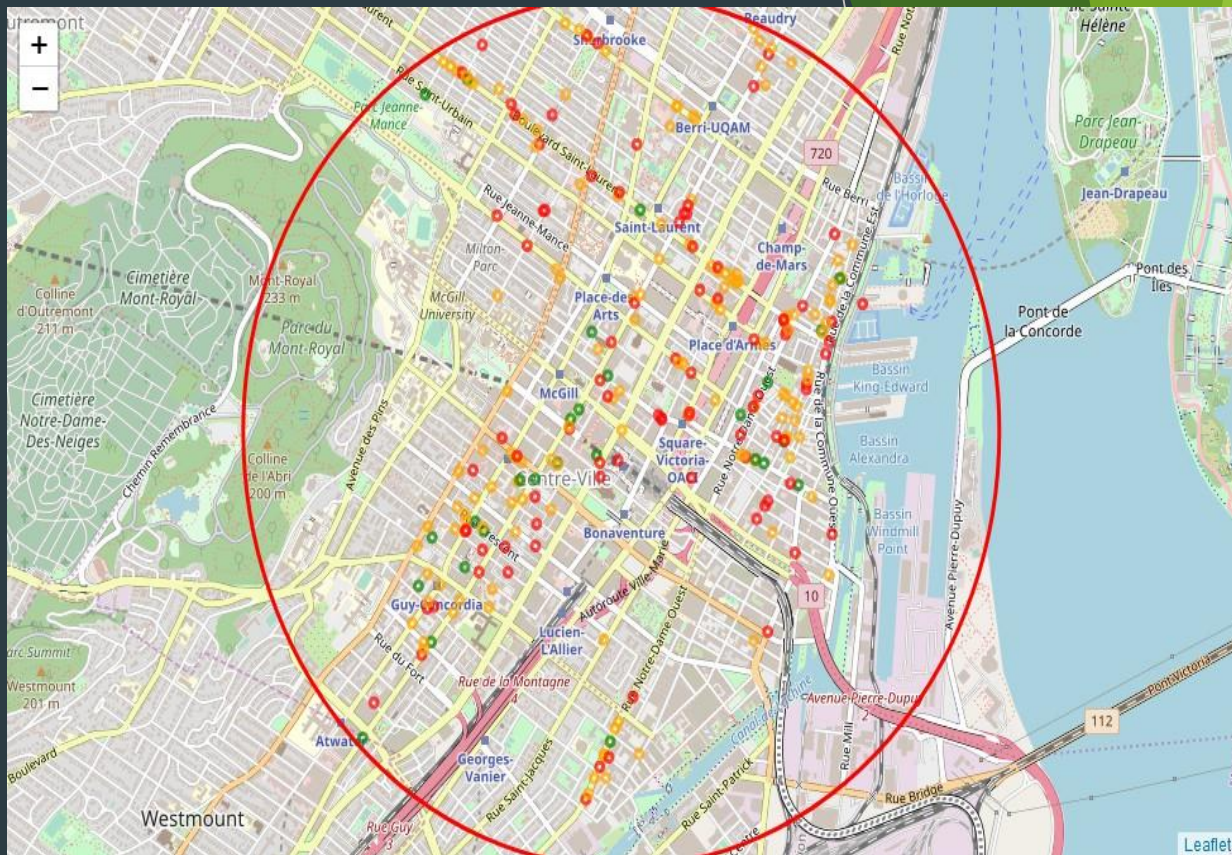
- red - relatively low number of likes
- orange - medium number of likes
- green - relatively large number of likes

Toronto

"Likes" map

- red - relatively low number of likes
- orange - medium number of likes
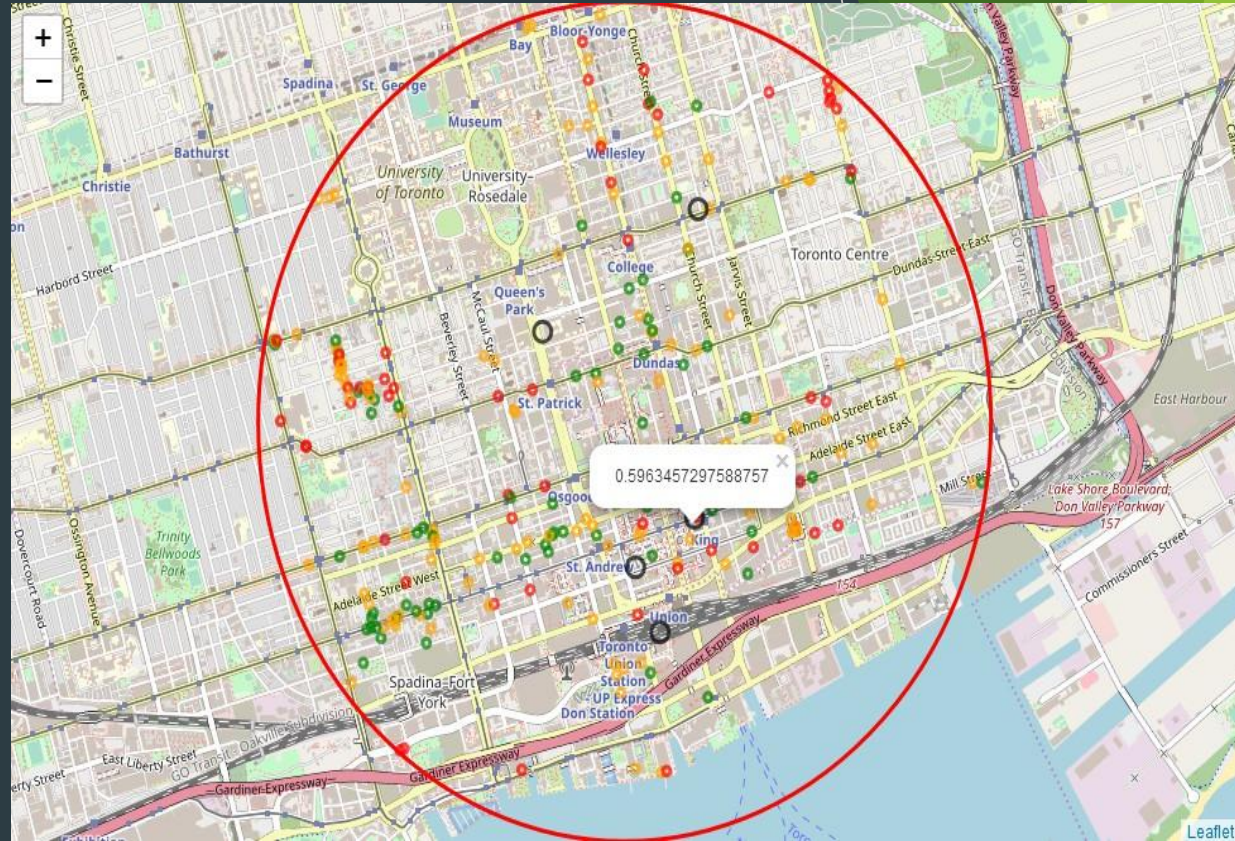- green - relatively large number of likes

Montreal

Vancouver map analysis

- Possible restaurant locations (black circle) and their "likes" score (from 0 to 1), with 1 being as popular as the most popular restaurant in the area.

- The possible restaurant with the highest "likes" score is located in downtown, surrounded by various of restaurant.

- When the restaurant is located in the area where there's only a few or no restaurant, the restaurant itself will not have a high "likes" score.
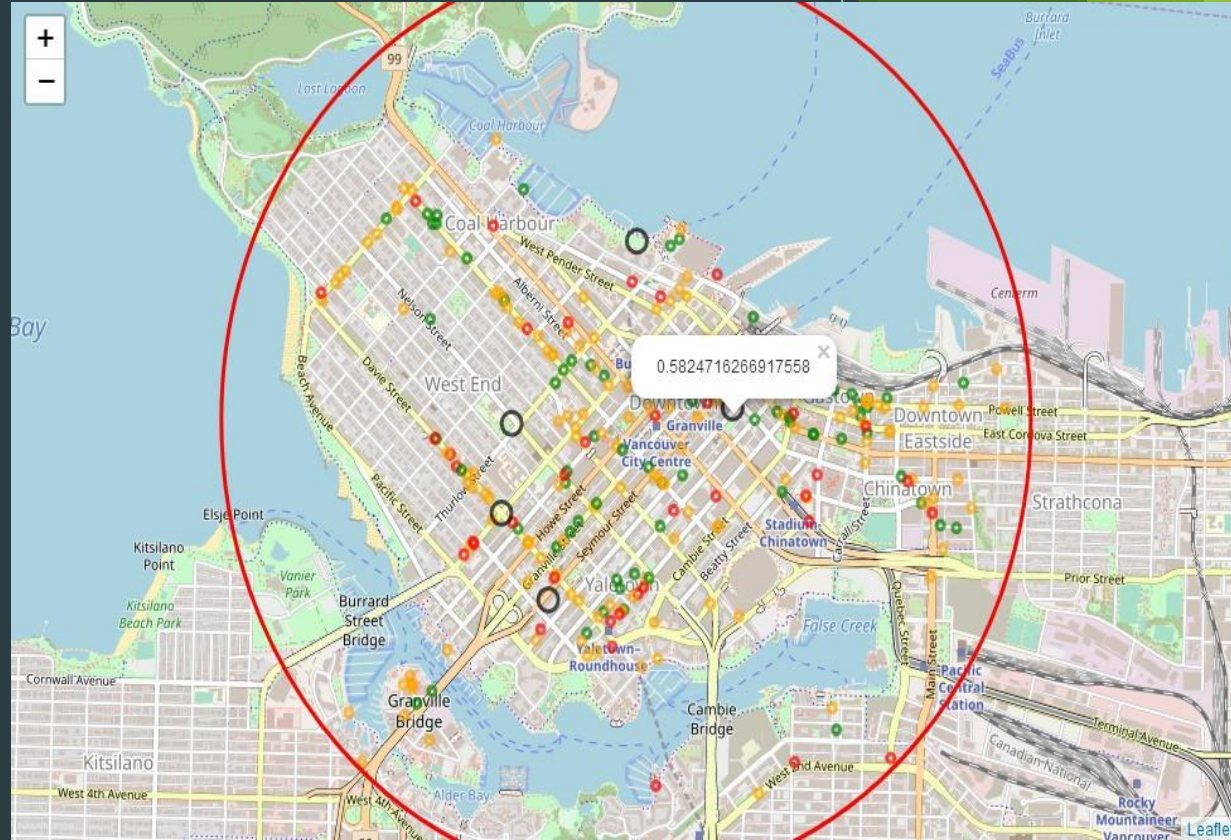
Toronto map analysis

- Possible restaurant locations (black circle) and their "likes" score (from 0 to 1), with 1 being as popular as the most popular restaurant in the area.

- The possible restaurant with the highest "likes" score is located in downtown, more specifically, in the financial district just next to a restaurant with low "likes" score.

- Other possible restaurant locations with high "likes" score are also in the financial district or next to the Union train station

Montreal map analysis

- Possible restaurant locations (black circle) and their "likes" score (from 0 to 1), with 1 being as popular as the most popular restaurant in the area.

- The possible restaurant with the highest "likes" score is located in downtown commercial area and next to the McGill University campus, with several restaurants with low "likes" score around.

- The three restaurants next to the Montreal central station and Lucien station are rated relatively low, mainly because of such areas are neither commercial areas nor main tourist attraction areas.
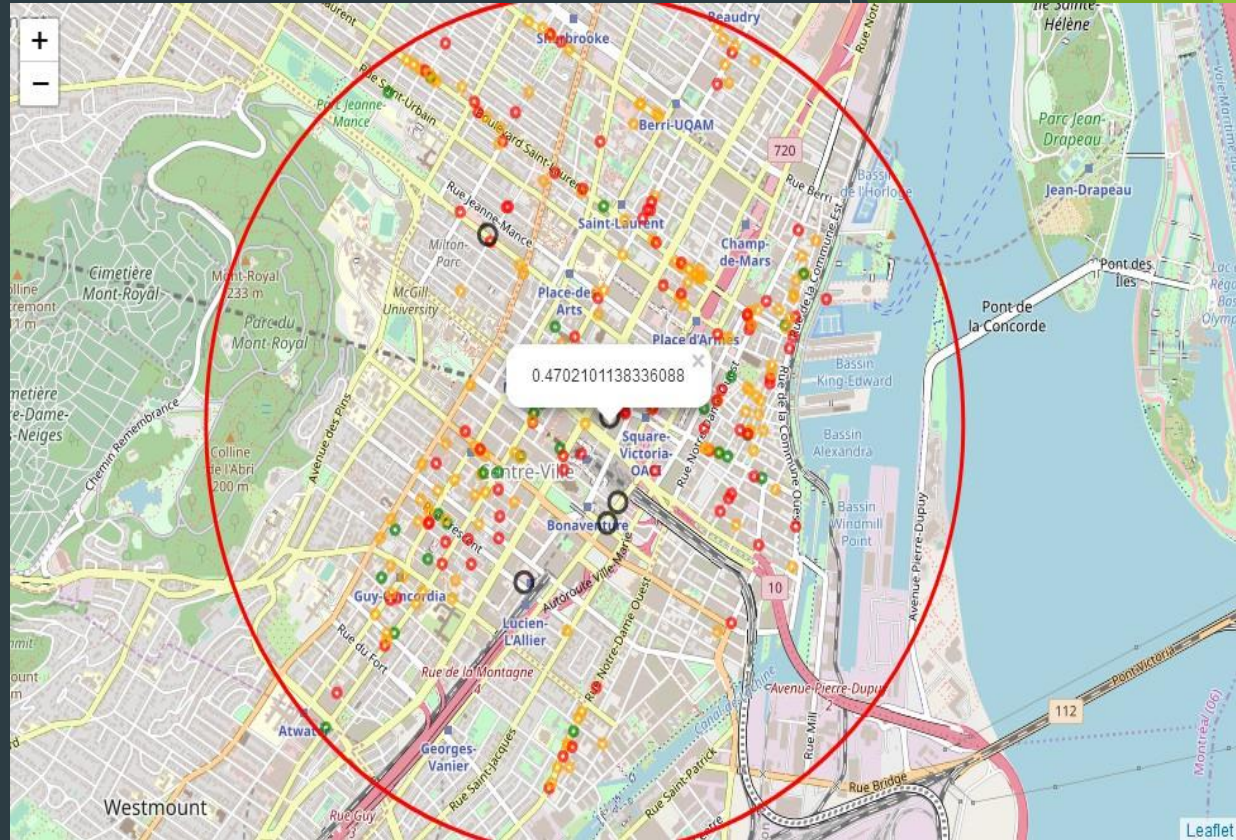
Conclusions and discussion

- Primary insights shows that restaurant has higher relative number of likes when the place is closer to other venues and gives lower score when it is a single restaurant in the area.

- Restaurants close to other low "likes" score restaurants may help, as we may see in the Toronto and Montreal example. Because people may more willing to try out the new restaurant instead of those already-known not-so-good restaurants, especially when the new restaurant is just next door.

- Toronto seems to be the city where Mr. Lee should set up his oversea restaurant, given by the highest "likes" score predicated.

- Approach can be further improved by adding more features and using more data

- This is also the outline of the tool, that can be used by businesses to choose a best location for their venues