



CITY DEVELOPMENT PLAN

COURSERA CAPSTONE PROJECT

IBM DATA SCIENTIST PROFESSIONAL

Authored by: Reza Sadeghi Jafari

PROBLEM

How we can optimize the distribution of facilities on city development plan

Most cities have a committee for planning the city development plan, they need to be ensured all the essential needs of the city's residents are well available in all parts of the city



INTERESTS:

- In this project, the city development team who play the role, as a customer, has requested that the necessary facilities be explored as follows throughout the city of **Toronto**, and that different parts of the city be segmented accordingly.

1- Medical centers

2- Police stations

3- Fire stations

DATA ACQUISITION AND CLEANING:

Needed data

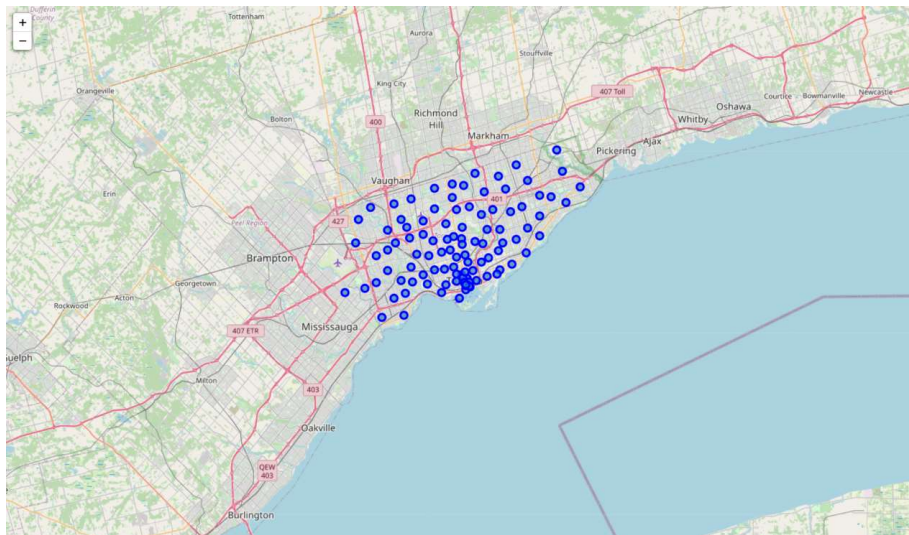
- 1- Urban areas
- 2- Facilities available in each area
- 3- the population on each area

Data Cleaning:

The data of Urban areas which provided on Wikipedia page had some invalid values that should be removed, We removed “Not assigned” borough from or data , then we need to group by and remove duplicate records.

DATA EXPLORATORY AND ANALYSIS:

The geographical information of urban areas will be added to base data, you can see the neighbors in the following map:



SOLUTION:

By clustering the data, we can group similar neighbors, I divide it to 5 groups of cluster which represent 0, 0-25%, 25%-50%, 50%-75%, 75%-100% based on describing of dataframe,

```
result_grouped.describe()
```

	Medical_Venue	Security_Venue
count	96.000000	96.000000
mean	0.000373	0.000713
std	0.000606	0.001284
min	0.000025	0.000025
25%	0.000098	0.000141
50%	0.000197	0.000312
75%	0.000361	0.000802
max	0.004489	0.009975

SOLUTION:

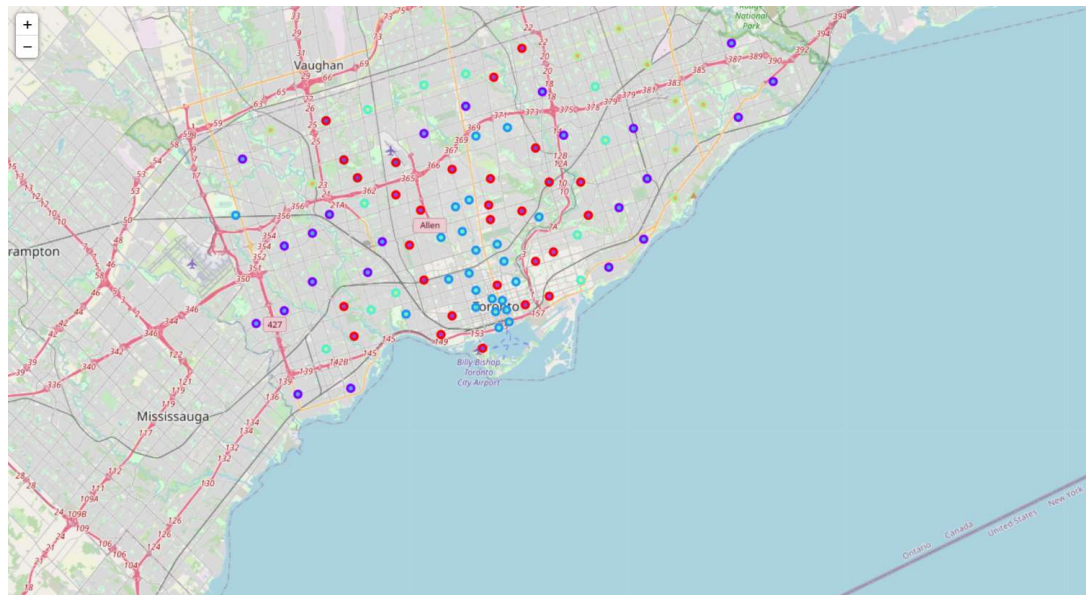
each neighbor will belong to a group number in each feature,

	Neighbourhood	Medical_Venue	Security_Venue	med_level	sec_level
0	Parkwoods	0.000087	0.000202	1	2

We use k-means model to cluster this 2 venues category based on med_level and sec_level features, to find similar neighbors and identify them. This help find the neighbors which are in same level of access to facilities.

SOLUTION:

Now the team can examine the distribution of facilities in each part of the city and prioritize the facilities that need further development based on each of the needs.





THANK YOU