FINDING SUITABLE AND BEST LOCATION FOR ESTABLISHING EMERGENCY ROOMS

November 12, 2020

1. Introduction

1.1. Business Problem

The route from accident zones to the medical or help centre might take long leading to more deaths due to shortage of time. This project aims to predict the highly dense accident zones in the state of **Toronto** and **establish** medical or **emergency rooms** in those areas, saving more lives and hopes.

1.2. Interest

Mainly the government of the state or country and the people who are prone to these kinds of accidents are of major concern here. The friends and families of those will be more happy and can lead a happy and joyful life.

2. Data acquisition and data cleaning

2.1. Data source

For solving this problem I need two different datas,

1. Locations of collisions or accidents that happened in previous years in Toronto.
2. Locations of Hospitals and Emergency rooms in Toronto.

The first data is found in the **toronto public service data portal**. This has column like Most accident stats, date and time, visibility, traffic control, neighborhood district, and locations(latitude, longitude) of the accidents, etc.

The second data needed for this project is the hospital or medical centre datas. These are retrieved using **Foursquare API**. The data is

requested using the URL for the API. It sends back a JSON file from which we make a dataframe.

2.2. Data Cleaning

In the accident dataset only the latitude and longitude, and neighborhood district of the accident is required as this project only considers the density of the accident.

In the hospital dataset only the latitude and longitude of the hospital, venue name, and type of the venue(medical centre, emergency room, etc), venue id, are considered.

3. Methodology

3.1. Mapping the two datasets

The view on the collisions of the accident's locations and the emergency room's locations are made. These are viewed in the map called "collisions_map".

3.2. Removing the accident location records nearer to medical centres

The accidents near the hospitals at a distance of 1km are removed from the dataset as they can reach the hospitals in time. The distance of the locations is calculated using the haversine formula.

$$d = 2r\sin^{-1}\left(\sqrt{\sin^2\left(\frac{\Phi_2-\Phi_1}{2}\right) + \cos(\Phi_1)\cos(\Phi_2)\sin^2\left(\frac{\lambda_2-\lambda_1}{2}\right)}\right)$$

3.3. Finding accident dense area using hierarchical dbscan

To find the density of the accident zones, the locations of the accidents are grouped together. Grouping is done using hierarchical density based spatial clustering of applications with noise clustering technique. This clustering is done with a minimal sample of 7. If we increase the sample size

the cluster becomes conservative so more outliers of cluster forms. The minimal number of clusters to be formed is given as 5. The HDBscan returns labels of each data which represents the cluster in which it is present. If the label is -1 it represents outliers.

3.4. Addressing outliers

Even here there will be outliers. Usage of k-nearest neighbors can help them group into the nearest cluster. The cluster without outliers formed using HDBscan are taken as a training set. Clusters with outliers formed using k-nearest neighbors are taken as a test set.

3.5. Finding core latitude and longitude and address of all clusters

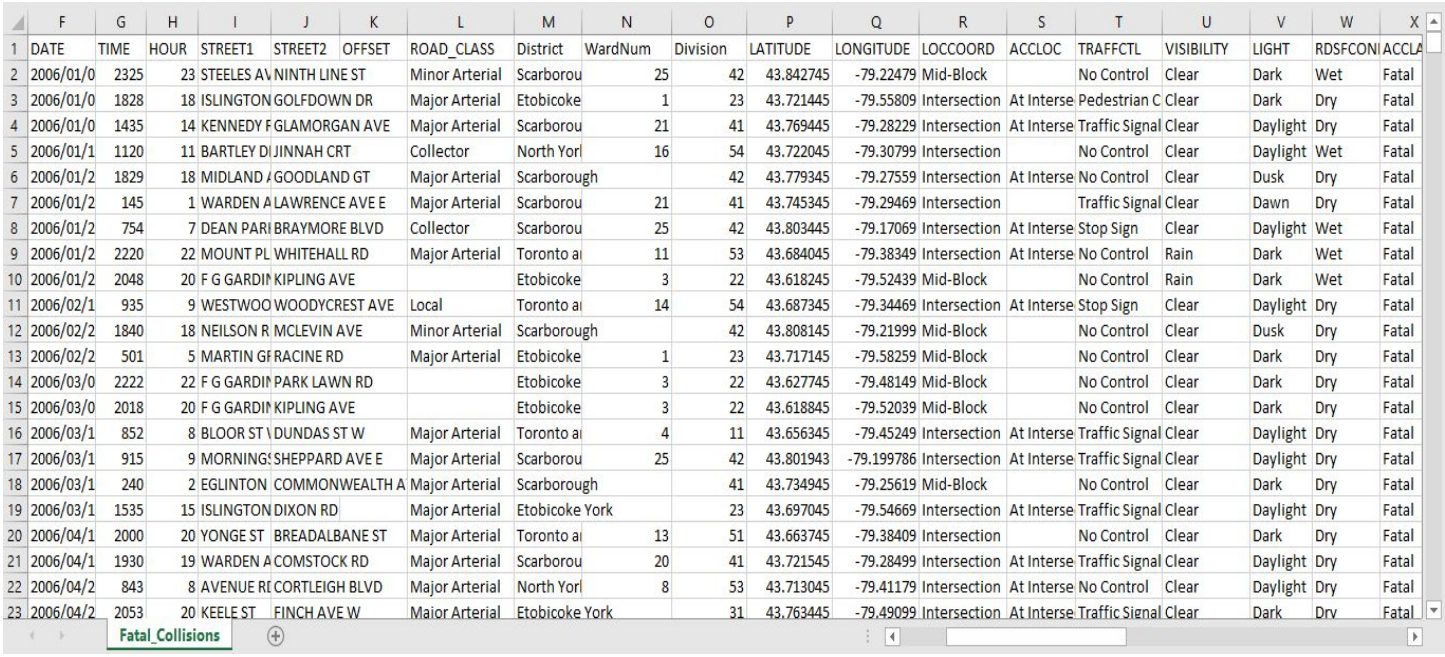
To find an appropriate location for the emergency room will be the mean of each cluster. This is considered as the core latitude and longitude of the cluster. Using the core latitude and longitude the address of that location is found. This address can be used as an ideal location for setting up an emergency room.

3.6. Calculating distance to the nearest medical centre

Now the distance of the core location and the nearest hospital is found. The longest distance of these is considered as the ideal one for establishing the hospital.

4. Results

4.1 Data taken from Toronto public service data portal



	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	DATE	TIME	HOUR	STREET1	STREET2	OFFSET	ROAD_CLASS	District	WardNum	Division	LATITUDE	LONGITUDE	LOCCOORD	ACCLOC	TRAFFCTL	VISIBILITY	LIGHT	RDSFCONI	ACCLA
2	2006/01/0	2325	23	STEELES AV	NINTH LINE ST		Minor Arterial	Scarborou	25	42	43.842745	-79.22479	Mid-Block		No Control	Clear	Dark	Wet	Fatal
3	2006/01/0	1828	18	ISLINGTON	GOLFDOWN DR		Major Arterial	Etobicoke	1	23	43.721445	-79.55809	Intersection	At Interse	Pedestrian C	Clear	Dark	Dry	Fatal
4	2006/01/0	1435	14	KENNEDY F	GLAMORGAN AVE		Major Arterial	Scarborou	21	41	43.769445	-79.28229	Intersection	At Interse	Traffic Signal	Clear	Daylight	Dry	Fatal
5	2006/01/1	1120	11	BARTLEY DI	JINNAH CRT		Collector	North Yorl	16	54	43.722045	-79.30799	Intersection		No Control	Clear	Daylight	Wet	Fatal
6	2006/01/2	1829	18	MIDLAND A	GOODLAND GT		Major Arterial	Scarborough		42	43.779345	-79.27559	Intersection	At Interse	No Control	Clear	Dusk	Dry	Fatal
7	2006/01/2	145	1	WARDEN A	LAWRENCE AVE E		Major Arterial	Scarborou	21	41	43.745345	-79.29469	Intersection		Traffic Signal	Clear	Dawn	Dry	Fatal
8	2006/01/2	754	7	DEAN PARI	BRAYMORE BLVD		Collector	Scarborou	25	42	43.803445	-79.17069	Intersection	At Interse	Stop Sign	Clear	Daylight	Wet	Fatal
9	2006/01/2	2220	22	MOUNT PL	WHITEHALL RD		Major Arterial	Toronto ai	11	53	43.684045	-79.38349	Intersection	At Interse	No Control	Rain	Dark	Wet	Fatal
10	2006/01/2	2048	20	F G GARDI	KIPLING AVE			Etobicoke	3	22	43.618245	-79.52439	Mid-Block		No Control	Rain	Dark	Wet	Fatal
11	2006/02/1	935	9	WESTWOO	WOODYCREST AVE		Local	Toronto ai	14	54	43.687345	-79.34469	Intersection	At Interse	Stop Sign	Clear	Daylight	Dry	Fatal
12	2006/02/2	1840	18	NEILSON R	MCLEVIN AVE		Minor Arterial	Scarborough		42	43.808145	-79.21999	Mid-Block		No Control	Clear	Dusk	Dry	Fatal
13	2006/02/2	501	5	MARTIN GF	RACINE RD		Major Arterial	Etobicoke	1	23	43.717145	-79.58259	Mid-Block		No Control	Clear	Dark	Dry	Fatal
14	2006/03/0	2222	22	F G GARDI	PARK LAWN RD			Etobicoke	3	22	43.627745	-79.48149	Mid-Block		No Control	Clear	Dark	Dry	Fatal
15	2006/03/0	2018	20	F G GARDI	KIPLING AVE			Etobicoke	3	22	43.618845	-79.52039	Mid-Block		No Control	Clear	Dark	Dry	Fatal
16	2006/03/1	852	8	BLOOR ST \	DUNDAS ST W		Major Arterial	Toronto ai	4	11	43.656345	-79.45249	Intersection	At Interse	Traffic Signal	Clear	Daylight	Dry	Fatal
17	2006/03/1	915	9	MORNING	SHEPPARD AVE E		Major Arterial	Scarborou	25	42	43.801943	-79.199786	Intersection	At Interse	Traffic Signal	Clear	Daylight	Dry	Fatal
18	2006/03/1	240	2	EGLINTON	COMMONWEALTH A		Major Arterial	Scarborough		41	43.734945	-79.25619	Mid-Block		No Control	Clear	Dark	Dry	Fatal
19	2006/03/1	1535	15	ISLINGTON	DIXON RD		Major Arterial	Etobicoke York		23	43.697045	-79.54669	Intersection	At Interse	Traffic Signal	Clear	Daylight	Dry	Fatal
20	2006/04/1	2000	20	YONGE ST	BREADALBANE ST		Major Arterial	Toronto ai	13	51	43.663745	-79.38409	Intersection		No Control	Clear	Dark	Dry	Fatal
21	2006/04/1	1930	19	WARDEN A	COMSTOCK RD		Major Arterial	Scarborou	20	41	43.721545	-79.28499	Intersection	At Interse	Traffic Signal	Clear	Daylight	Dry	Fatal
22	2006/04/2	843	8	AVENUE RI	CORTLEIGH BLVD		Major Arterial	North Yorl	8	53	43.713045	-79.41179	Intersection	At Interse	No Control	Clear	Daylight	Dry	Fatal
23	2006/04/2	2053	20	KEELE ST	FINCH AVE W		Major Arterial	Etobicoke York		31	43.763445	-79.49099	Intersection	At Interse	Traffic Signal	Clear	Dark	Dry	Fatal

Fatal_Collisions

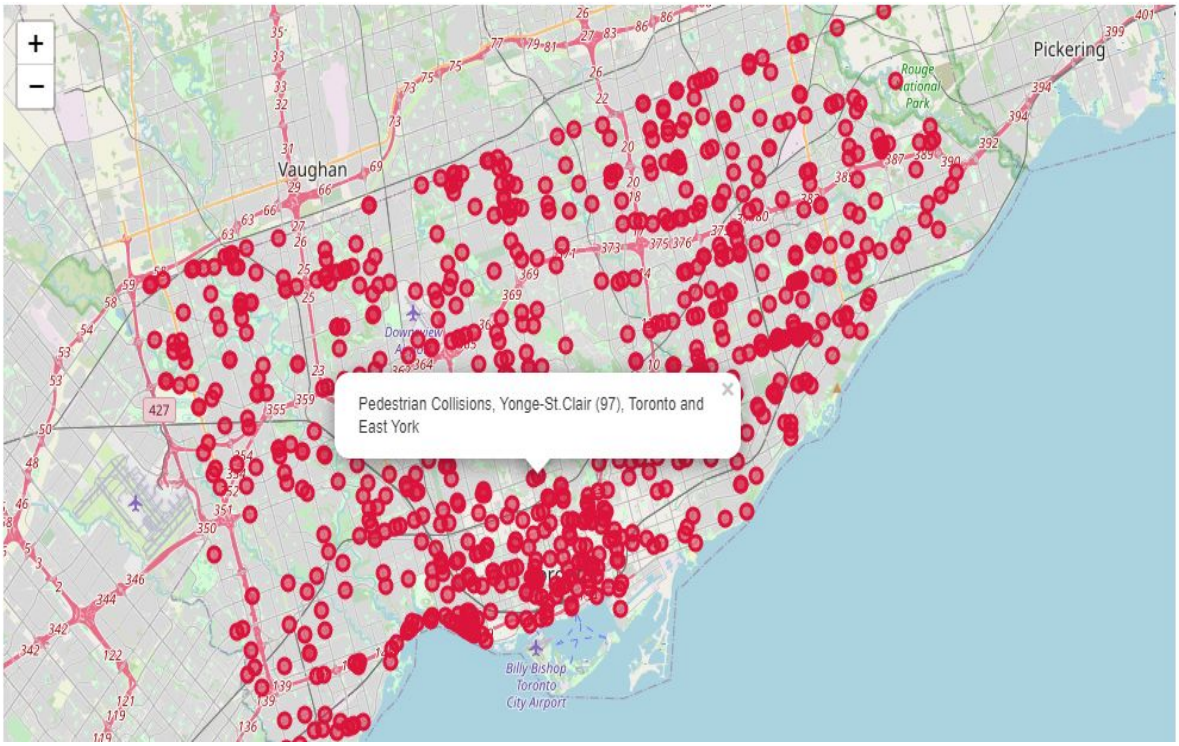
4.2 DataFrame after removing unwanted columns and null values:-

Out[5]:

	IMPACTYPE	Neighbourhood	District	LATITUDE	LONGITUDE
0	Approaching	Rouge (131)	Scarborough	43.842745	-79.22479
1	Pedestrian Collisions	Elms-Old Rexdale (5)	Etobicoke York	43.721445	-79.55809
2	Pedestrian Collisions	Dorset Park (126)	Scarborough	43.769445	-79.28229
3	SMV Other	Victoria Village (43)	North York	43.722045	-79.30799
4	Pedestrian Collisions	Agincourt South-Malvern West (128)	Scarborough	43.779345	-79.27559

4.3 Toronto Map in which collision locations are marked:-

Out[9]:



4.4 Hospitals location data when retrieved from Foursquare API:-

Out[170]:


```
[{'reasons': {'count': 0,
   'items': [{'summary': 'This spot is popular',
     'type': 'general',
     'reasonName': 'globalInteractionReason'}]},
  'venue': {'id': '4af2fb96f964a52086e921e3',
   'name': 'Toronto Western Hospital',
   'location': {'address': '399 Bathurst St.',
    'crossStreet': 'at Dundas St. West',
    'lat': 43.65343431584569,
    'lng': -79.40607359183444,
    'labeledLatLngs': [{'label': 'display',
      'lat': 43.65343431584569,
      'lng': -79.40607359183444}],
    'postalCode': 'M5T 2S7',
    'cc': 'CA',
    'city': 'Toronto',
    'state': 'ON',
    'country': 'Canada',
    'formattedAddress': ['399 Bathurst St. (at Dundas St. West)',
```

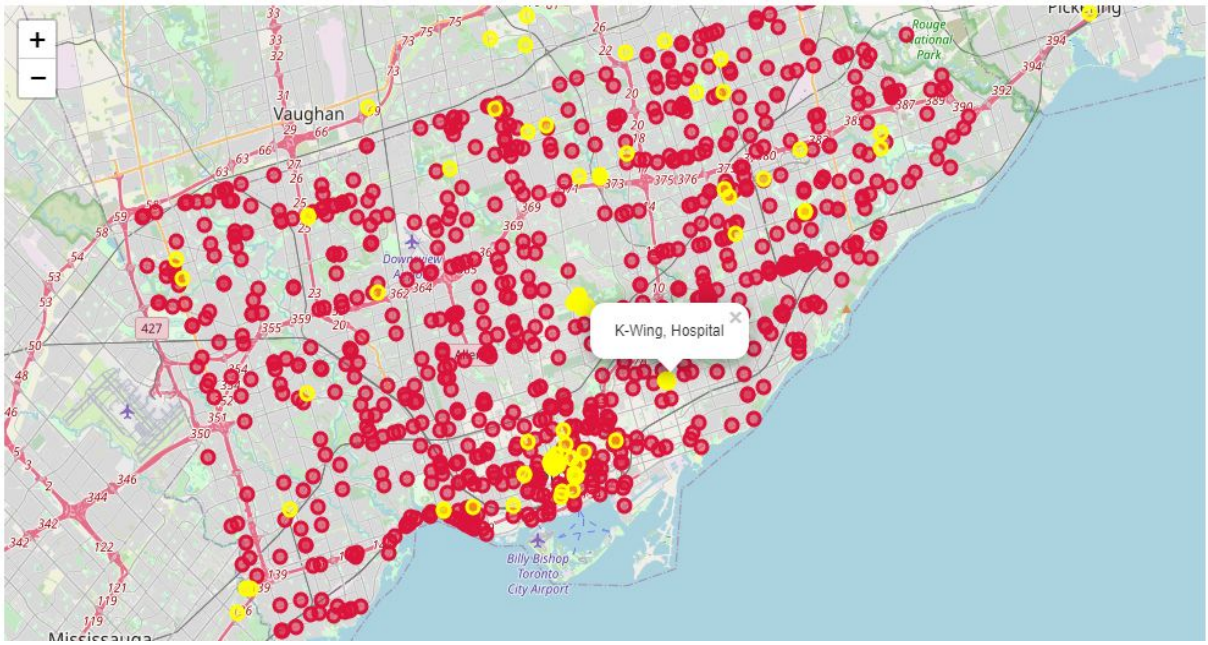
4.5 Hospitals data as data frame after performing Preprocessing:-

Out[15]:

	Venueid	Venue	Latitude	Longitude	Category
0	4af2fb96f964a52086e921e3	Toronto Western Hospital	43.653434	-79.406074	Medical Center
1	4b7d80d9f964a520d3c22fe3	Roswell Park	42.898945	-78.864902	Medical Center
2	4e023cdb6365ba98ee33dd32	Southlake Regional Health Centre	44.061136	-79.452311	Medical Center
3	4b66f884f964a520e6322be3	Buffalo Medical Group (295 Essjay)	42.974214	-78.736627	Medical Center
4	4bd057b177b29c7493848a82	Halton Family Health Centre	43.391872	-79.821928	Medical Center

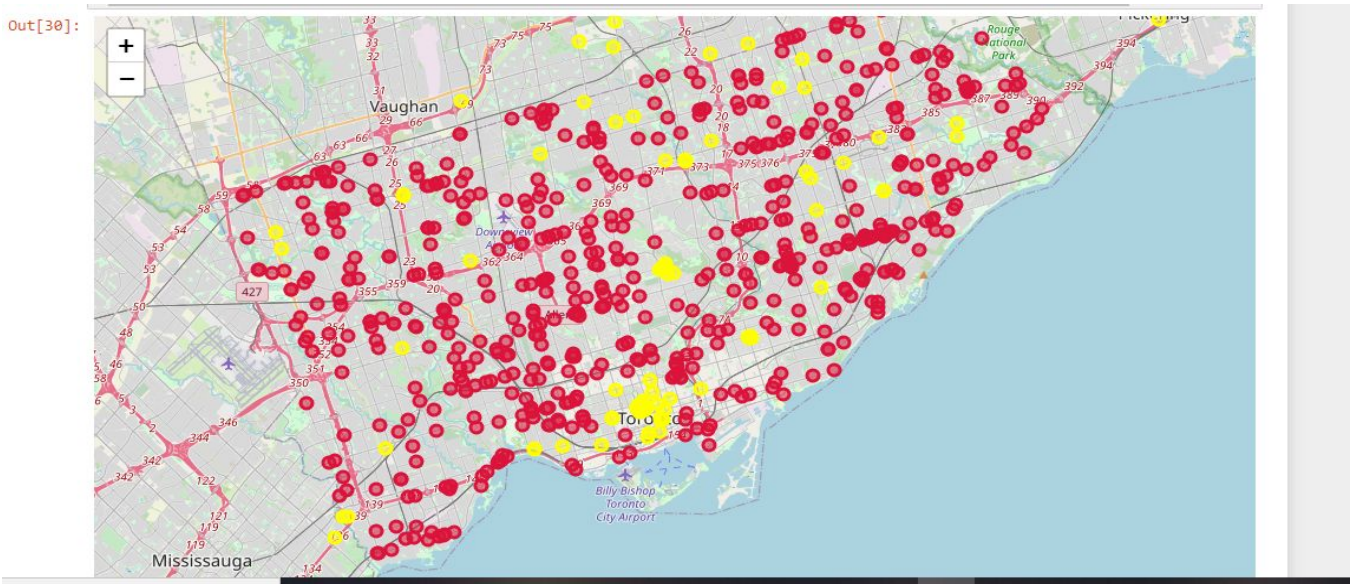
4.6 Mapping hospitals locations along with collisions location in Toronto map:-

Out[24]:



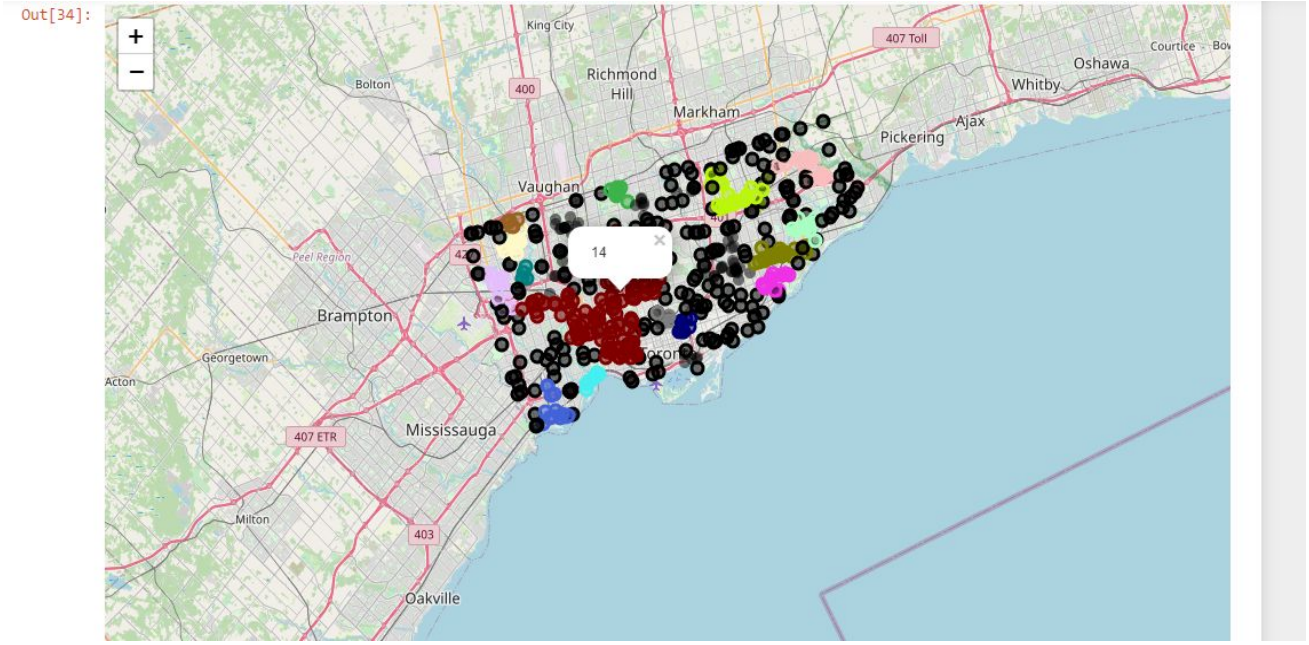
4.7 Mapping after removing all collisions which are in the radius of 1 KM from any hospital:-

Out[30]:

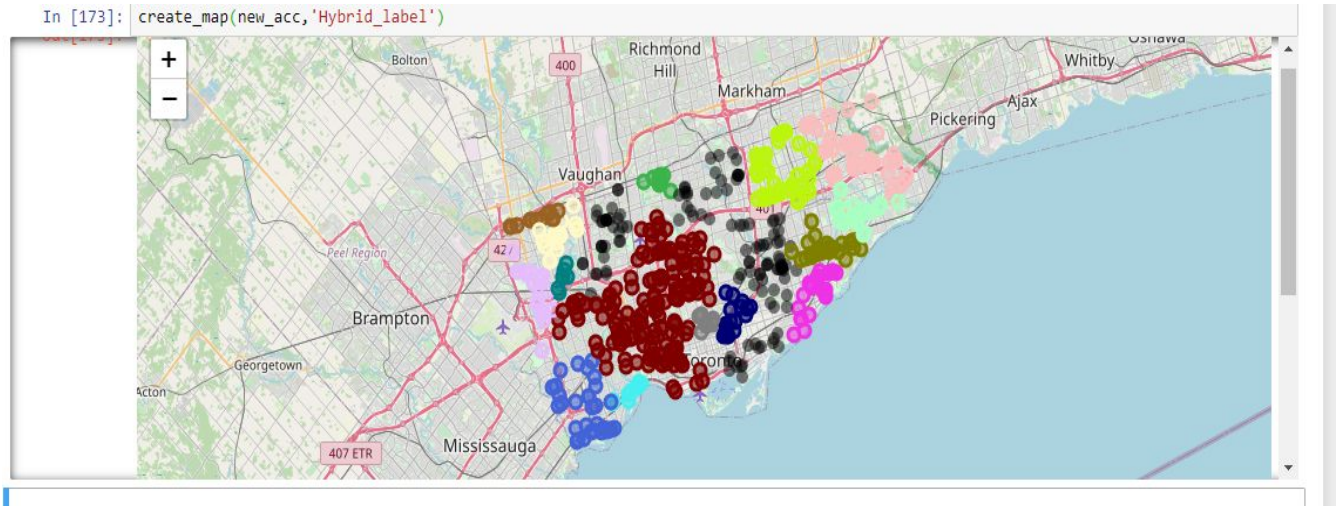


4.8 Map after clustering collision points based on density using hierarchical DBSCAN:-

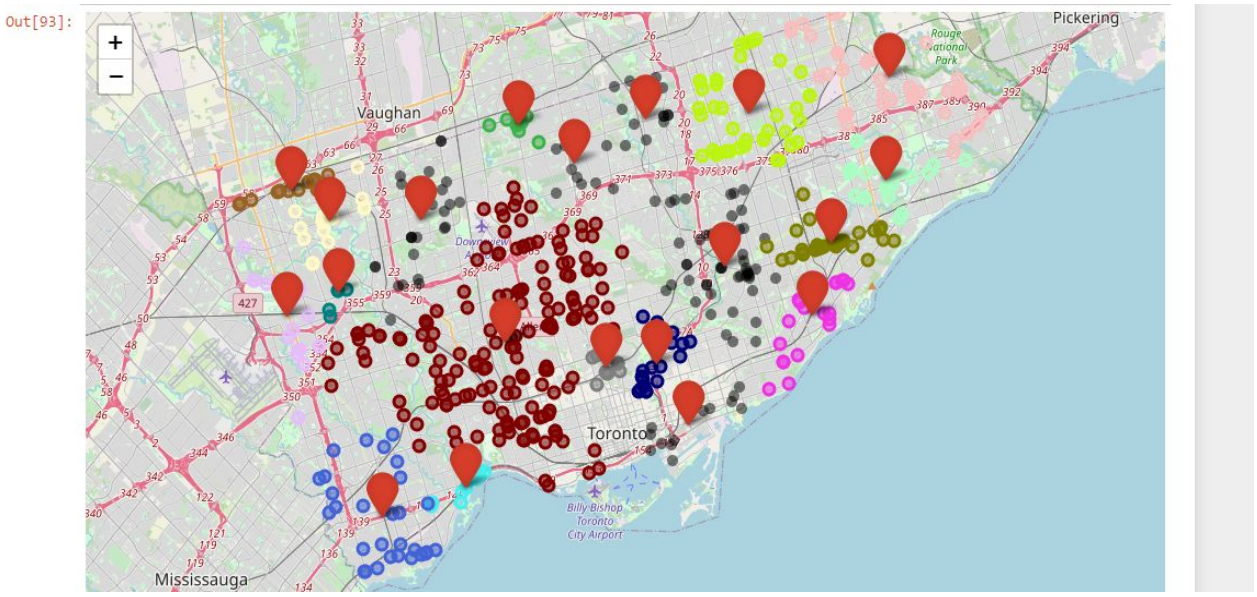
Out[34]:



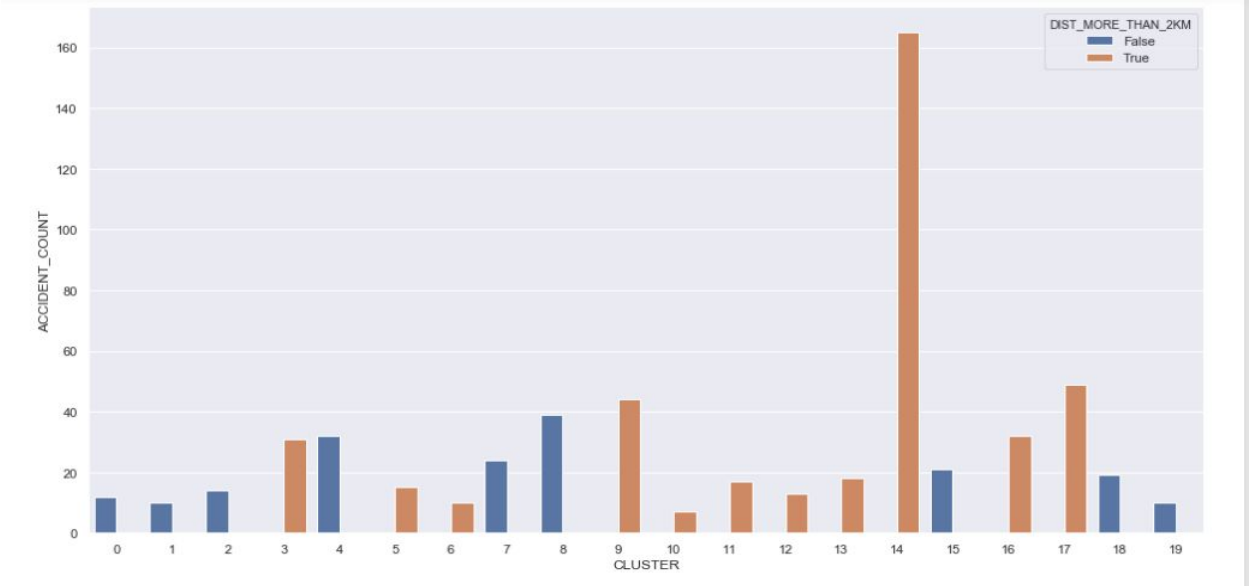
4.9 Map after clustering the outliers using KNearestNeighbors:-



4.10 Cores points along with their clusters members:-



4.11 Bar plot showing number of accidents in each clusters and distance from each cluster center to the nearest hospital:-



5. Discussions

By considering the number of accidents in each cluster and also the distance to the hospitals, it can be suggested that the cluster with the longest distance and largest number of accidents can be an ideal one for establishing the emergency room.

From the above Bar visualization if we have to select one location then we can choose cluster 14, as it has large number of collisions and also the distance to the nearest hospital from the cluster center is more than 2 Kilometers.

6. Conclusion

In this study, I analyzed the sufficiency and necessity of emergency rooms in accident zones. I built a clustering model to group the accidents into different clusters. Analyzing these clusters can be very useful in helping the person in interest to establish a suitable place for setting up an emergency room.