FINDING SUITABLE AND BEST LOCATION FOR ESTABLISHING EMERGENCY ROOMS IN TORONTO

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.

Target Audience

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.

Ideology

First, we should ensure that the location we are choosing must be considerably away from any other emergency rooms that are already available.

Secondly, we should ensure that the location is highly dense with collisions.This can be done by using any of the clustering model.

Data acquisition and data cleaning

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.

Data Source

The collisions 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 hospitals location 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.

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.

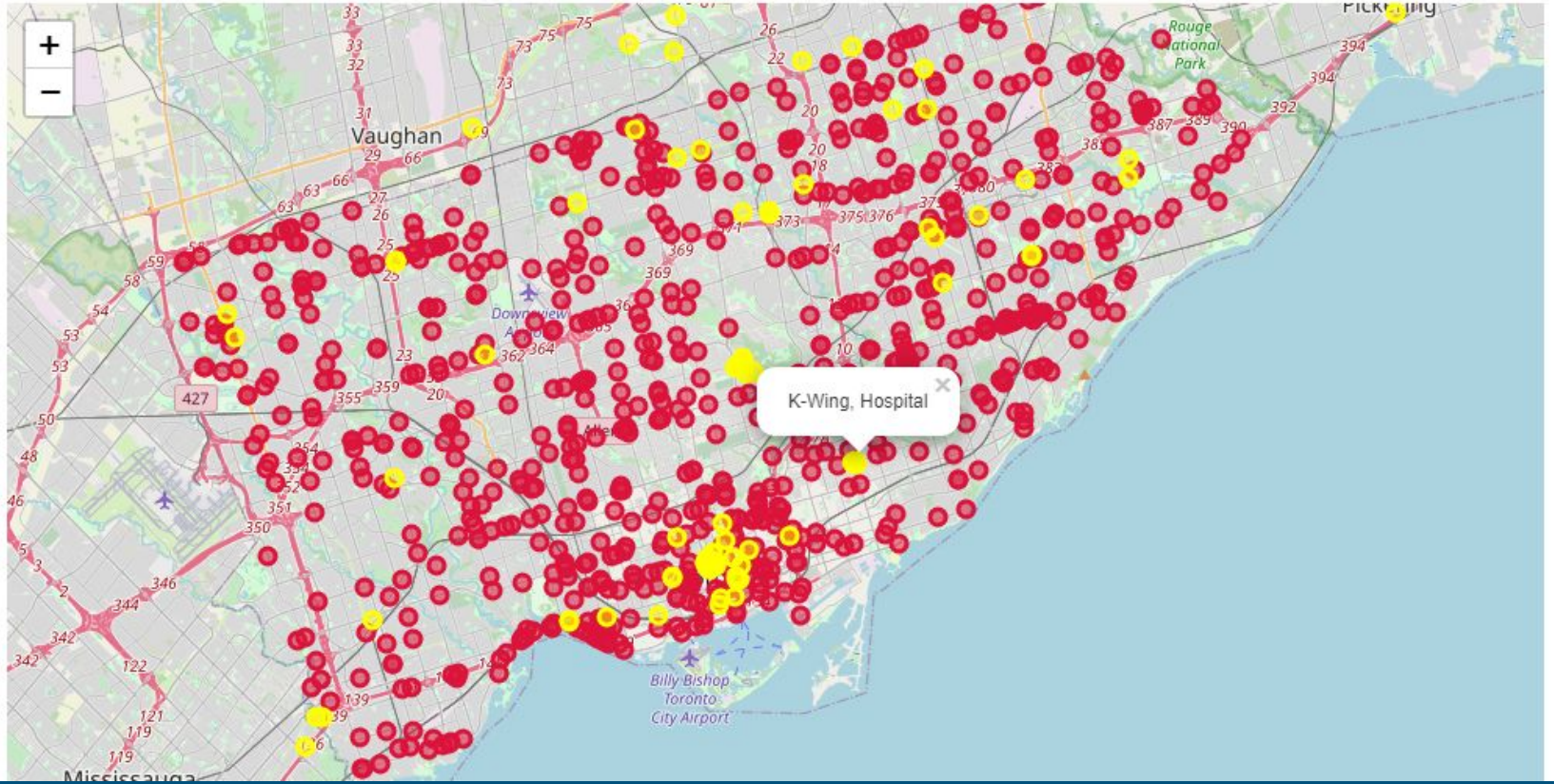
In the hospital dataset only the latitude and longitude of the hospital, venue name, and type of the venue(medical centre, emergency room, etc), venue id, are considered.

Methodology

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".

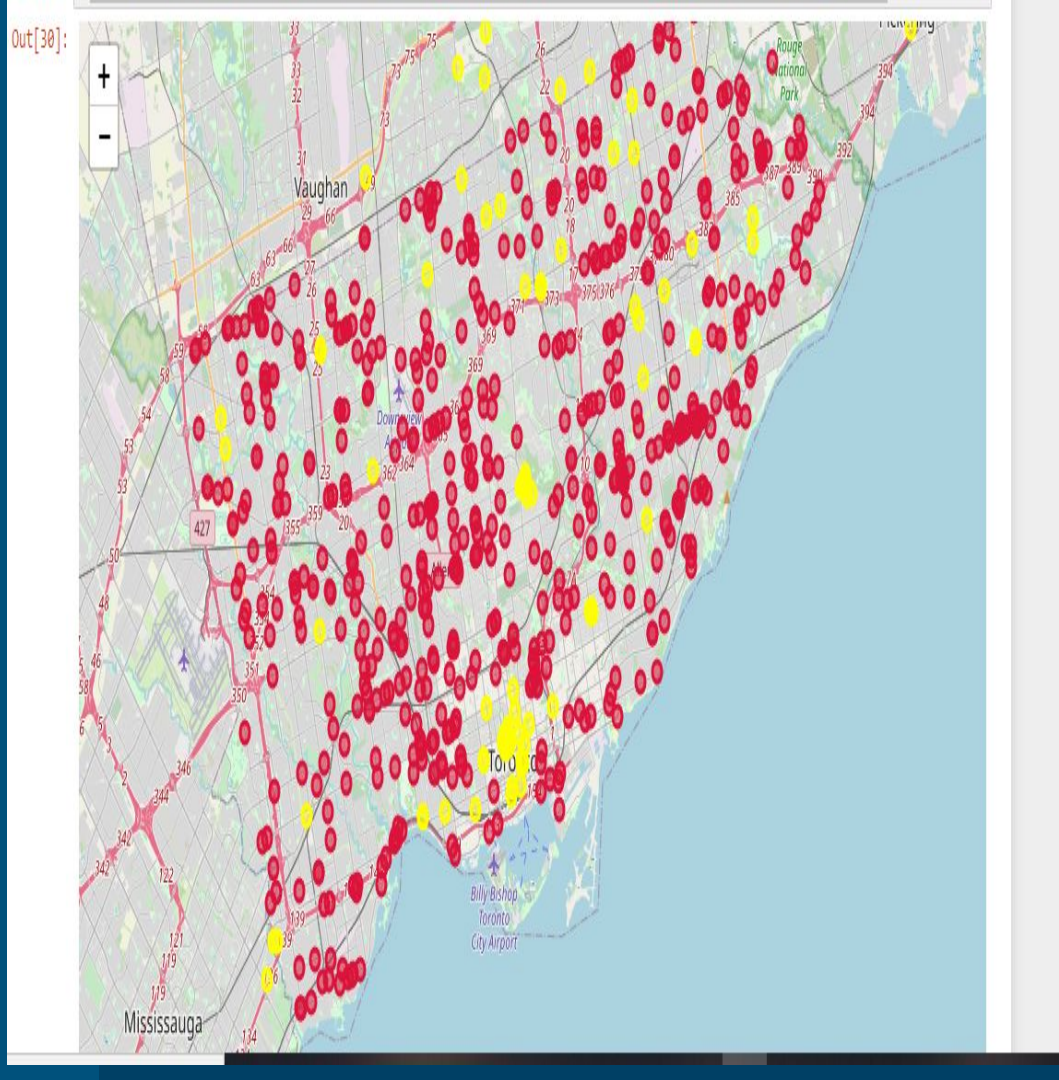
Out[24]:



Mapping hospitals locations along with collisions location in Toronto map

Removing the accident location records nearer to medical centres

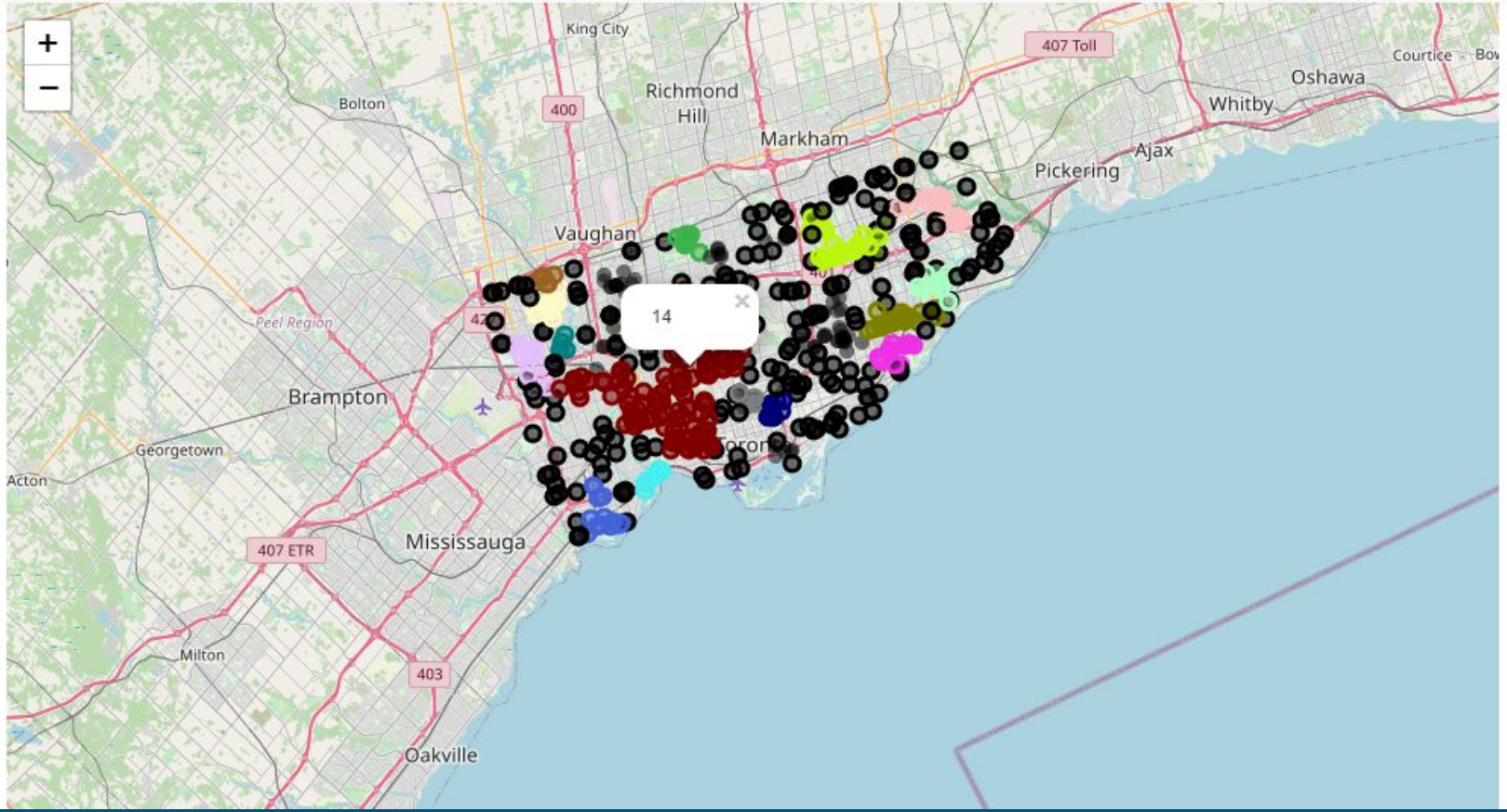
The accidents near the hospitals at a distance of 1km are removed from the dataset as they can reach the hospitals in time. The distance of the locations is calculated using the haversine formula.



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.

Out[34]:

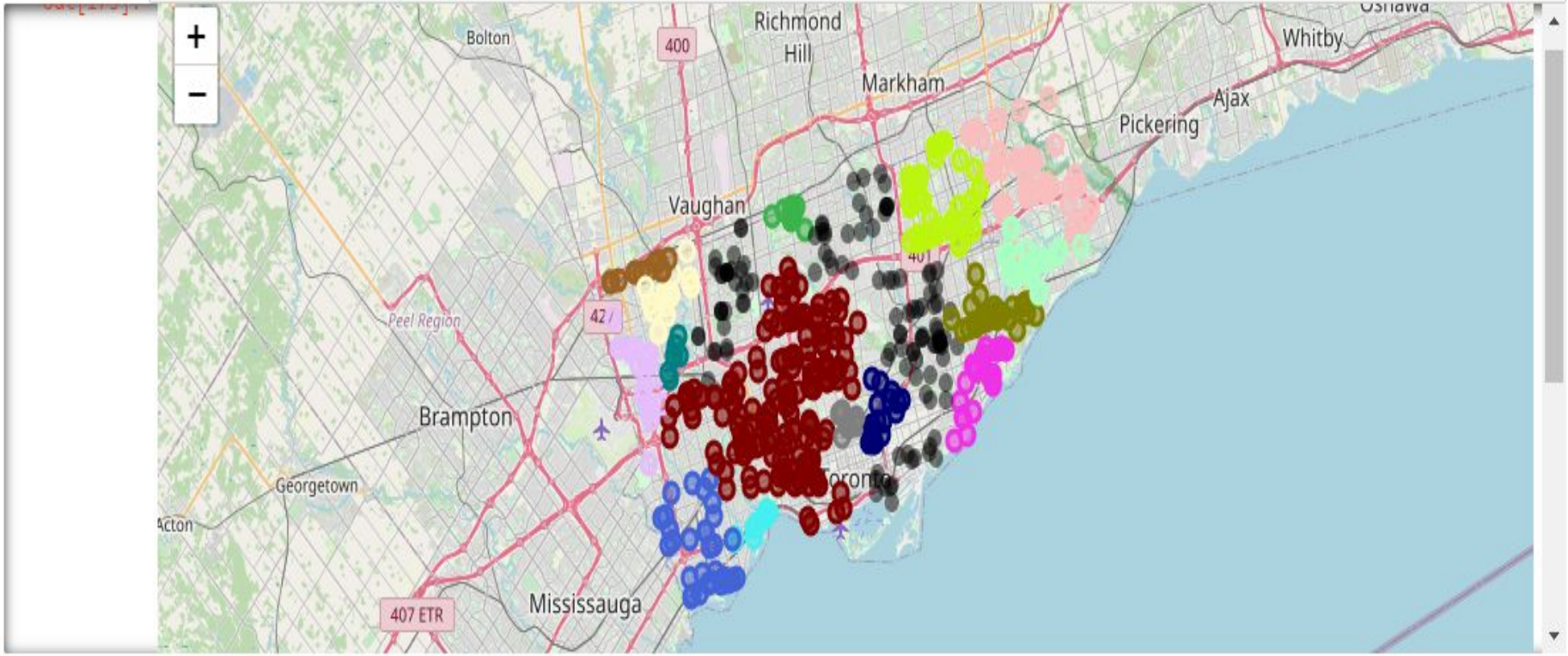


Map after clustering collision points based on density using hierarchical DBSCAN

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.

```
In [173]: create_map(new_acc,'Hybrid_label')
```

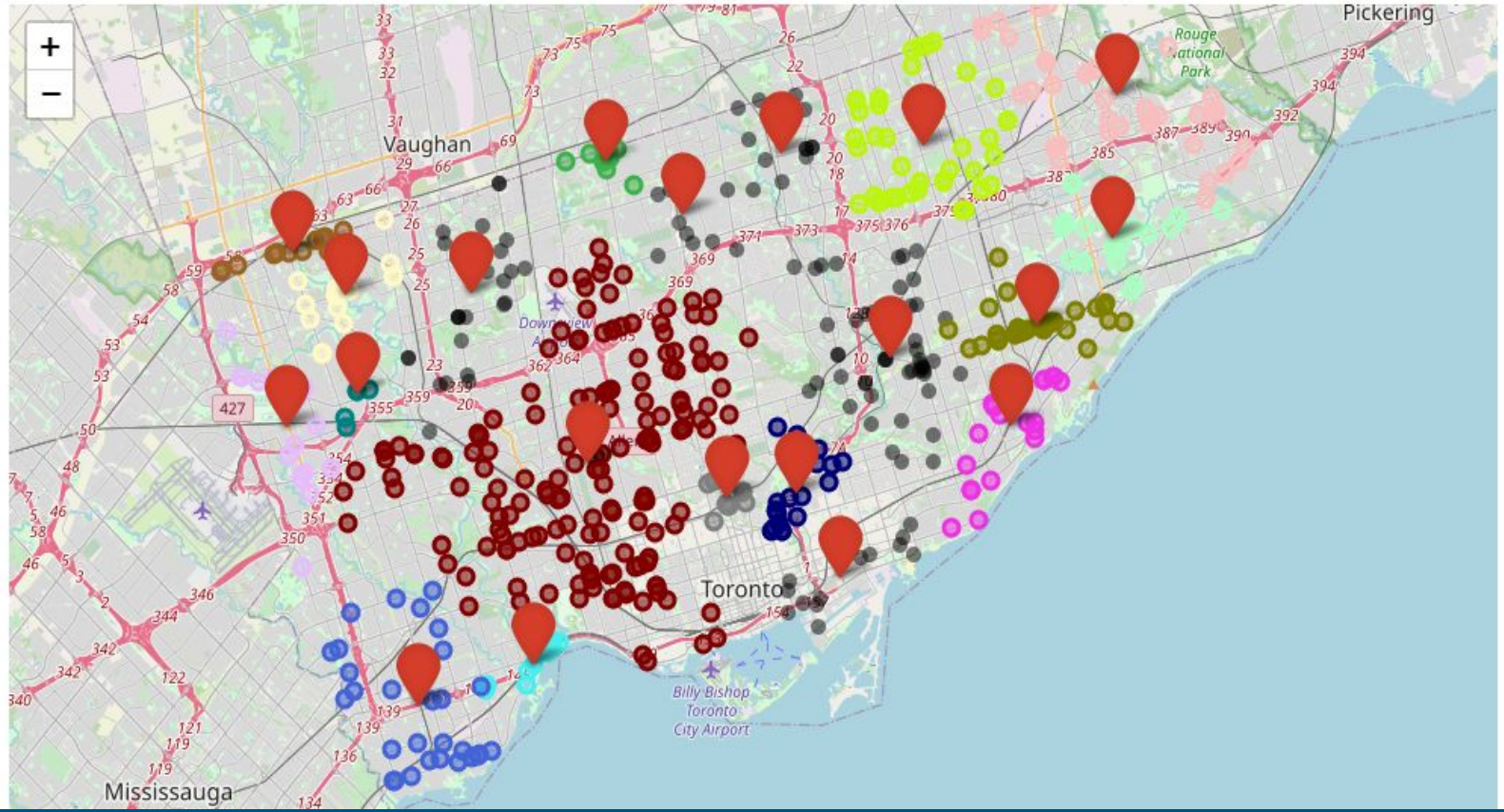


Map after clustering the outliers using KNearestNeighbors

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.

Out[93]:

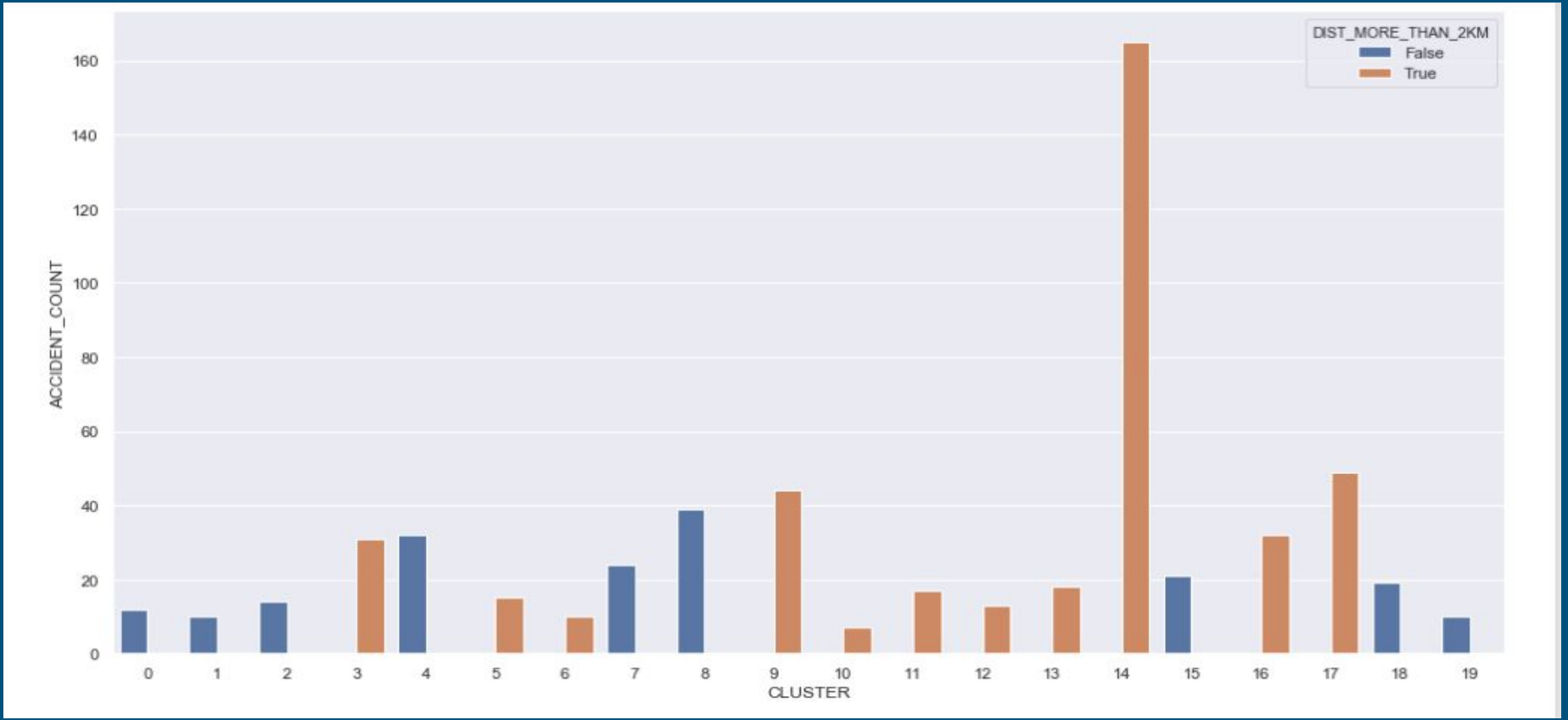


Cores points along with their clusters members

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.

Below we are plotting a bar against three attributes which are cluster, accident counts in each cluster and a boolean attribute which tells whether distance to the nearest hospital is more than 2 KM or not.



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

Result Discussion

From the above graph the suitable Cluster in which we can establish the emergency rooms are 14,17 and 9 since they have more accident counts compared to others and also they have their nearest hospital in the distance of more than 2 KM which will result in risk of people lives. thus these three can be chosen in prior to other 16 clusters.

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.