**Hospitals in Our Neighbourhoods**

* **Business Problem/Problem Statement:**

The main focus of my problem statement will be based on health care, now we are living in an era where accidents, health emergencies, diseases happen almost everyday, now lets say you are new to a neighbourhood and you don’t know a lot about the area that you are living in, how would you know about the hospitals present in the neighbourhood if in case of an emergency, you have to go the browser and look up hospitals near you which would take time and yeah time is really valuable in these cases. A good solution would be leveraging the power of foursquare api and data analysis techniques to make an application where the user just has to open his app and the app will automatically show top 10 or top 20 hospitals in your area. That would be really time saving and also it would be really helpful in case of accidents, because accident can happen in areas which you are not even familiar with.

Now developers can use our script to develop web or mobile based application which would be the plan in action. I believe this idea would really be helpful in solving some issues of common people.

Later we can upgrade our services to include the ratings, comments of the hospitals and can also provide a section for customer feedback on the application.

* **Data**

For now we would use the readily available neighbourhood data of Toronto for this project but later we can modify it so we can provide api calls to developers so they would send a neighbourhood as a query and we would provide a list of hospitals as a response.

We would also use the foursquare api to find the hospitals in our neighbourhood.

The data is present in form of csv comma separated data which we would process using pandas data frame and manipulate it using a lot of techniques

* **Methodology**

What I did is I took the data that is readily available on Wikipedia, that is a neighbourhood data of Toronto, scraped that data using pandas library and then converted it into a dataframe.

Now there were some borough’s that were unassigned so we filtered them by using the below technique:

df\_canada = df\_canada[df\_canada['Borough']!='Not assigned']

The above lines of code will filter out any rows that are unassigned

Now after that we have to get the latitude and longitude of the postal codes using the geocoder package, we have to pass in the postal code name along with the area in the url and we would get the latitude and longitude using the below code:

import geocoder # import geocoder

lat\_lung = {'Latitude':[], 'Longitude':[]}

for postal\_code, borough in zip(df\_canada['Postal Code'], df\_canada['Borough']):

lat\_lng\_coords = None

# loop until you get the coordinates

while(lat\_lng\_coords is None):

g = geocoder.google('{}, Toronto, Ontario'.format(postal\_code))

lat\_lng\_coords = g.latlng

print(lat\_lng\_coords[0],lat\_lng\_coords[1])

lat\_lung['Latitude'].append(lat\_lng\_coords[0])

lat\_lung['Longitude'].append(lat\_lng\_coords[1])

Then I sorted the postal codes according to the latitudes and longitude and rearranged them.

Now the real objective behind our project starts now we have to find hospitals in our neighbourhood and show them to the people who are searching for and also we would also show the nearst five hospitals based on the distances from the neighbourhood, now this will be really good application if developers would use api calls with their latitudes and longitudes and will get the top nearest hospitals according to the distances.

The first step will be exploring the neighborhoods in Toronto:

We can do it using the below code, what we are actually is we are constructing a foursquare api containing the clientid, client secret, category id = **'4bf58dd8d48988d196941735'** which foursquare takes and gives the details about the hospitals in that area within a specific radius and a specific limit, here we have specified the radius as 500 meters and the limit is 5.

The below code is being used:

def getnearbyhospitals(names, latitudes, longitudes, radius=500):

hospital\_list=[]

for name, lat, lng in zip(names, latitudes, longitudes):

print(name)

# create the API request URL

url = 'https://api.foursquare.com/v2/venues/search?ll={},{}&categoryId={}&client\_id={}&client\_secret={}&limit={}&v={}'.format(

latitude, longitude, category\_id, CLIENT\_ID, CLIENT\_SECRET, LIMIT, VERSION)

# make the GET request

results = requests.get(url).json()["response"]['venues']

#extracing data that is required for a person to find the hospital like its latitude, longitude, its distance from the point and so on

hospital\_list.append([(

name,

lat,

lng,

v['name'],

v['location']['lat'],

v['location']['lng'],

v['location']['distance']) for v in results])

hospitals = pd.DataFrame([item for hospita\_list in hospital\_list for item in hospita\_list])

hospitals.columns = ['Neighborhood',

'Neighborhood Latitude',

'Neighborhood Longitude',

'Hospital',

'Hospital Latitude',

'Hospital Longitude',

'Hospital Distance']

return(hospitals)

Then we sort the data according to the distance of each neighbourhood to get the nearest 5 hospitals:

This can be done using the below code:

nearest\_hospitals = {'Neighborhood':[],'1stNearest':[],'2ndNearest':[],'3rdNearest':[],'4thNearest':[],'5thNearest':[]}

last = hospitals['Neighborhood'][0]

list\_hospital = {'Hospital':[],'distance':[]}

last = hospitals['Neighborhood'][0]

for i,name in enumerate(hospitals['Neighborhood']):

if name == last:

list\_hospital['Hospital'].append(hospitals.loc[i,'Hospital'])

list\_hospital['distance'].append(hospitals.loc[i,'Hospital Distance'])

else:

temp\_df = pd.DataFrame(list\_hospital)

temp\_df.sort\_values(by = 'distance', axis = 0, inplace = True)

temp\_df.reset\_index()

print(temp\_df)

nearest\_hospitals['Neighborhood'].append(name)

nearest\_hospitals['1stNearest'].append(temp\_df.iloc[0,0])

nearest\_hospitals['2ndNearest'].append(temp\_df.iloc[1,0])

nearest\_hospitals['3rdNearest'].append(temp\_df.iloc[2,0])

nearest\_hospitals['4thNearest'].append(temp\_df.iloc[3,0])

nearest\_hospitals['5thNearest'].append(temp\_df.iloc[4,0])

last = name

list\_hospital = {'Hospital':[],'distance':[]}

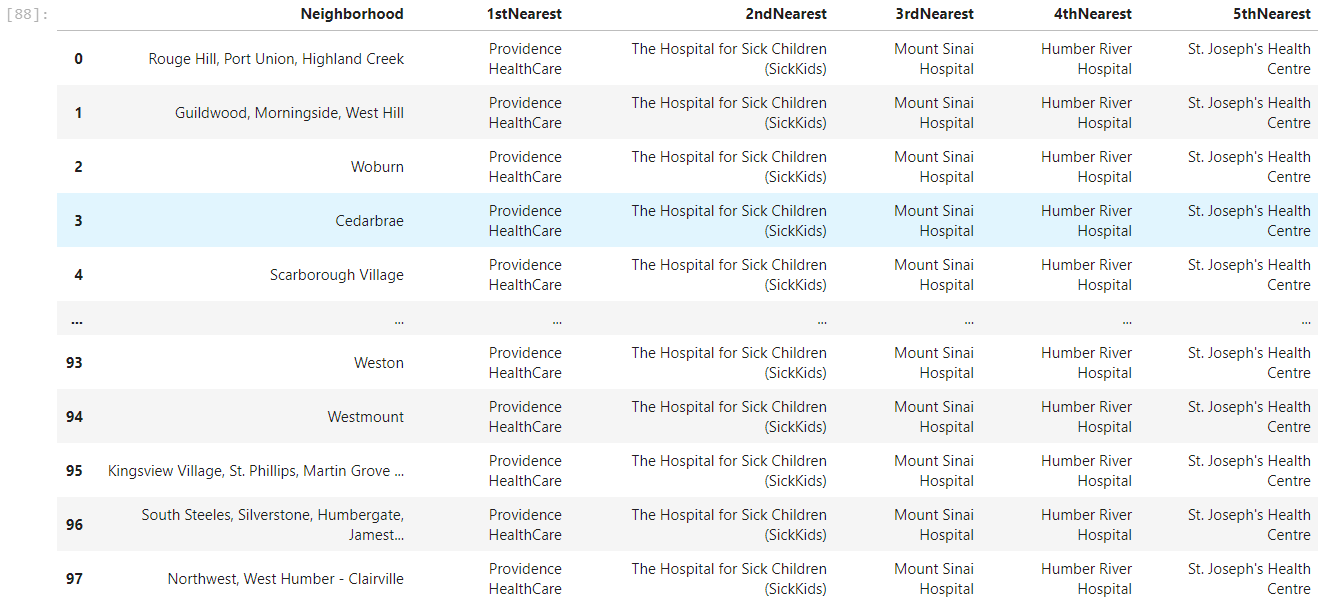
list\_hospital['Hospital'].append(hospitals.loc[i,'Hospital'])

list\_hospital['distance'].append(hospitals.loc[i,'Hospital Distance'])

Then we can display the data frame. Or if we modify our code we can send responses to api calls made by developers.

* **Results**

The results of the above analysis is displayed by the dataframe below:



As we can see in the above data frame the neighborhoods along with the 5 nearest hospitals are shown in the picture.

This results are based on the data analysis on the data that we have used and we can also modify it for specific calls to neighborhoods.

* **Discussion:**

This would be a very comprehensive project if done in a large scale proportion with a lot of developers, analysts and testing people. This project done by me just shows how we can help people by doing a little bit of data analysis and it can benefit a lot of people.

Now this project really focuses on the data provided on Toronto that was previously available by the course team. But this can be really useful if we can provide the hospital data on specific api calls like a person would just send the address of a location and we would fetch the latitude and longitude of the data then find out the hospital in the data and show it to the people. We can also show the data related to the hospitals but due to limitations on the premium api calls I was not able to experiment much.

* **Conclusion**

In the conclusion I want to say that I just shown an example on how we can leverage the technologies of foursquare and python data analysis techniques to benefit the people and we can achieve a lot much.