Introduction:

Nowadays, coffee plays an essentials role in most people's lives in a city like Toronto. Therefore, it will be competitive to build a coffee business, but because of the population, starting a coffee shop business will bring significant profit if it is in a good location. In this capstone project, I will leverage Foursquare APIs to fetch some data and use folium map to visualize the data analysis to provide an advise for a predictable location which may bring good profits to a new-starting coffee shop business in Toronto, Ontario, Canada.

Toronto is the capital city if the Canadian province of Ontario. With the most population in Canada, it brings many economic opportunities to whom wants to start their own small business. According to Hamilton Beach, specialty coffee consumption has increased by 20% per year. Therefore, a coffee shop business in Toronto could bring success to people who wants to start a small business.

We will use data science tools to fetch the raw data, visualize it then generate a suggested optima location to start a coffee business. In the meanwhile, we will also explain the advantage and traits for the stakeholders, so that stakeholders can make the final decision based on the analysis.

Data:

Based on the definition of our problem, factors that may impact our decision are:

- Demographic information, e.g. population, density, education, age, income
- the popular neighbor where locates most of the venue.
- Number of existing coffee shop in the neighborhood and nearby

We will implement Foursquare API to fetch the raw data and get the recommended data in Toronto city.

```
CLIENT_ID = 'your Foursquare ID' # your Foursquare ID
CLIENT_SECRET = 'your Foursquare Secret' # your Foursquare Secret
VERSION = '20180604'
LIMIT = 40
print('Your credentails:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET:' + CLIENT_SECRET)
```

And get requests near Toronto city.

```
request_parameters = {
    "client_id": CLIENT_ID,
    "client_secret": CLIENT_SECRET,
    "v": '20180605',
    "section": "coffee",
    "near": "Toronto",
    "radius": 1000,
    "limit": 50}

data = requests.get("https://api.foursquare.com/v2/venues/explore", params=request_parameters)
```

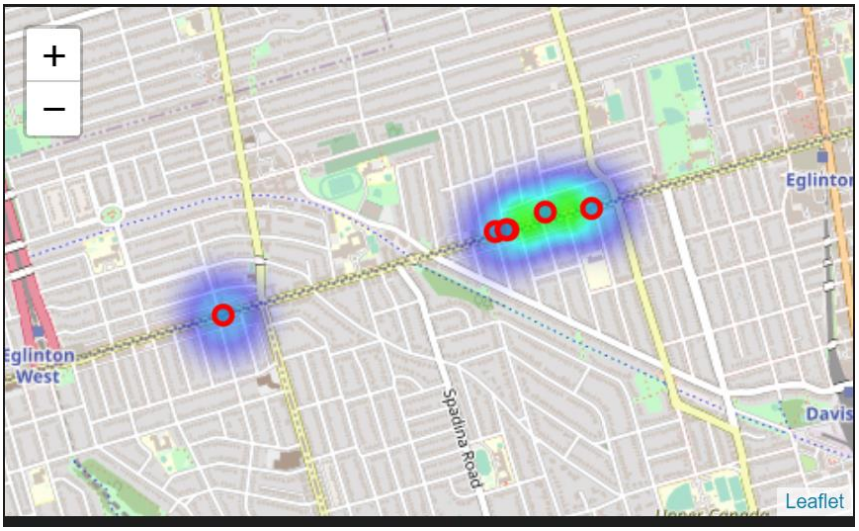
After the organizing and cleaning the data from requests, From the output we can identify necessary factors of what we will use later to consider the probability of launching our up-to-coming location.

```
df_raw = []
for item in items:
    venue = item["venue"]
    categories, uid, name, location = venue["categories"], venue["id"], venue["name"],
venue["location"]
    print(location)
    assert len(categories) == 1
    shortname = categories[0]["shortName"]
    address =  ''
    if hasattr(location, 'address'):
      address = location['address']
    if not "postalCode" in location:
        continue
    postalcode = location["postalCode"]
    lat = location["lat"]
    lng = location["lng"]
    datarow = (uid, name, shortname, address, postalcode, lat, lng)
    df_raw.append(datarow)
df = pd.DataFrame(df_raw, columns=["uid", "name", "shortname", "address", "postalcode",
 "lat", "lng"])
print("found %i cafes" % len(df))
df.head()
```

	uid	name	shortname	address	postalcode	lat	lng
0	4b896ec4f964a520bd3532e3	Hotel Gelato	Café		M5N 1B4	43.703478	-79.414311
1	4bdc79352a3a0f47a800b3b6	The Mad Bean Coffee House	Coffee Shop		M5N 1B1	43.703529	-79.413698
2	4b0feea1f964a5206a6623e3	Starbucks	Coffee Shop		M6C 2B6	43.700598	-79.427433
3	4ae99f04f964a52061b521e3	Starbucks	Coffee Shop		M5N 1A5	43.704171	-79.411887
4	4b17ed23f964a520c1c923e3	Second Cup	Coffee Shop		M5N 1A5	43.703583	-79.413824

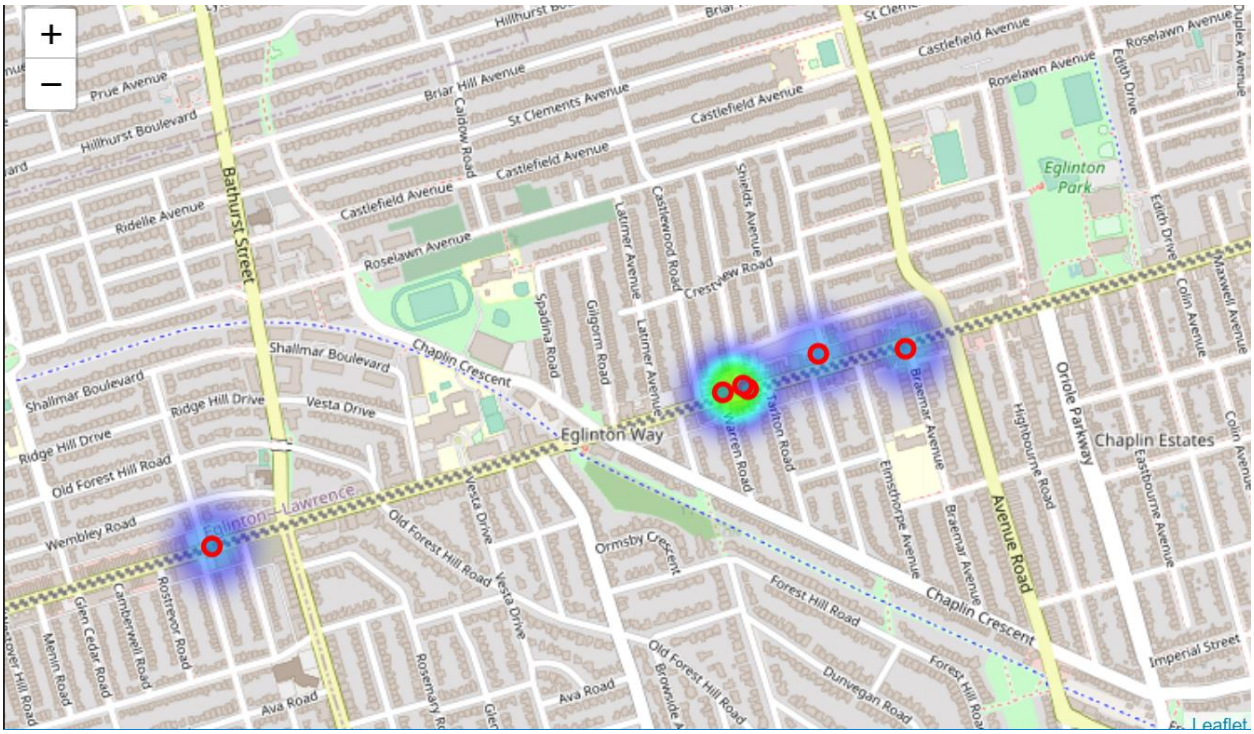
After we got the list of the “hot” coffee shop near center of Toronto. We apply FOLIUM to generate a heat map, which shows the most popular coffee shops in center of Toronto city.

```
from folium import plugins
map_Toronto = folium.Map(location=[43.70011, -79.4163]
, zoom_start=14)

def add_markers(df):
    for (j, row) in df.iterrows():
        label = folium.Popup(row["name"],
parse_html=True)
        folium.CircleMarker(
            [row["lat"], row["lng"]],
            radius=5,
            popup=label,
            color='red',
            fill=True,
            fill_color='#3186cc',
            fill_opacity=0.7,
            parse_html=False).add_to(map_Toronto)

add_markers(df)
hm_data = df[["lat", "lng"]].to_numpy().tolist()
map_Toronto.add_child(plugins.HeatMap(hm_data))

map_Toronto
```



From the heatmap, we can see the most popular coffee shops are mainly concentrated in downtown area. There are more office building and venues, which bring more customer to the coffee shops.

Next step we will get coordinates of Toronto and create folium map which will help visualize what we have got from data.



Methodology:

The purpose of this project is to find a satiable location to open a small coffee shop.

Now we retrieved the following data:

- The most popular coffee shops in Toronto
- The locations of the high-voted coffee shop in Toronto.

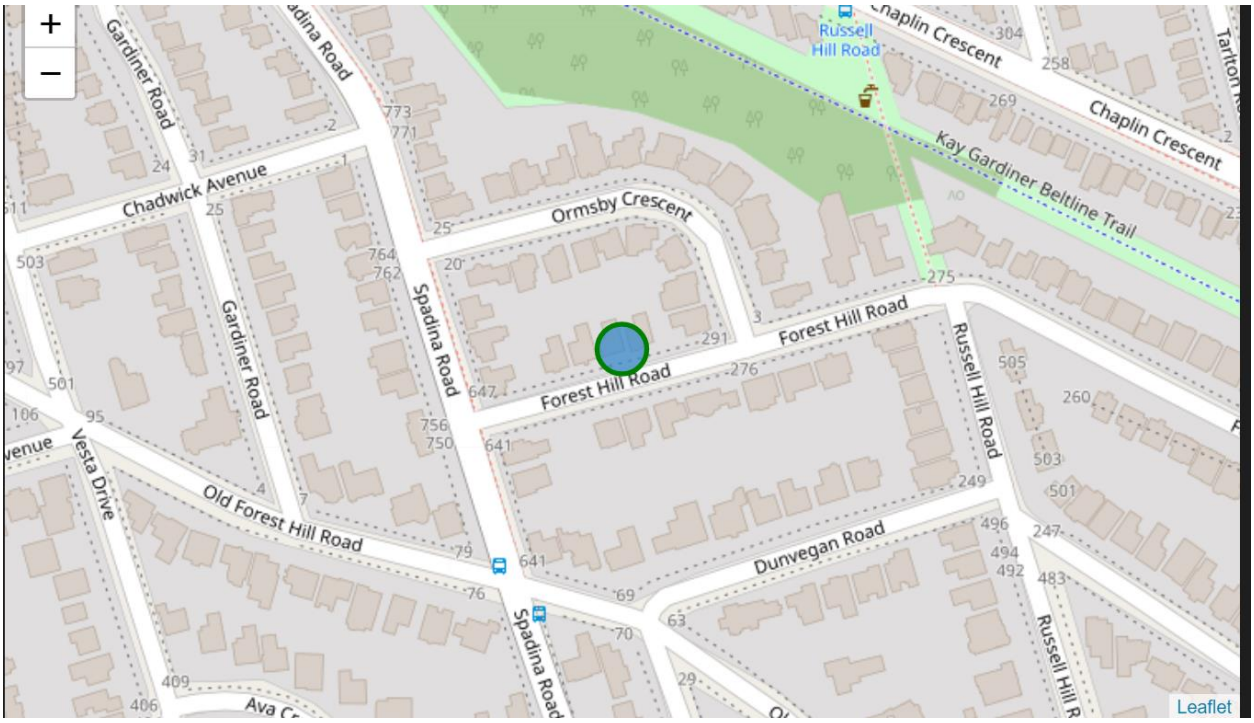
We generated the maps to show the “hot” area in the center of Toronto.

The final step will be considering different factors to provide a good advice of the location to open a coffee shop.

Analysis:

Based on the data we have above, the center of the Toronto downtown will bring more customers and opportunities to a small coffee shop. Because of the building density, a small coffee will be recommended in this area. The locations showed in following map will be the recommend location to open a small coffee shop, which is near to the competitive area, but not compete to them, and there are enough customers to bring the profits.

```
lat = 43.70011
lng = -79.4163
map_Toronto = folium.Map(location=[lat, lng], zoom_start=17)
add_markers(df)
folium.CircleMarker(
    [lat, lng],
    radius=15,
    color='green',
    fill=True,
    fill_color='#3186cc',
    fill_opacity=0.7,
    parse_html=False).add_to(map_Toronto)
map_Toronto
```



Result and Discussion:

From data analysis and visualization, we can see the most popular coffee shops are always located in the most population, such as near to the shopping mall or office building usually, which inspired us to find out the area with more population and heavy human traffic. Less competition will also be a important factor to be considered.

Based on the heatmap, we observe that the coffee shops which have more customer and sales are all near to the center of Toronto downtown. Therefore, if a coffee shop opens near to that area will very likely bring a good profit to the owner. A few blocks away from the most “hot area” will effectively reduce the completion rate, in the meantime, it will not greatly reduce the passenger flow.

Conclusion:

The purpose of this project is to find an area in Toronto to open a coffee shop.

After fetching data from Foursquare API and process them into a clean data frame, generating heatmaps with folium. We picked the cluster with more population flow and more popular coffee shops on average. We narrow down that the center of Toronto downtown would be a good place to open a small coffee shop.

The final decision on optimal location of the coffee shop will be made by stakeholders based on specific characteristics of neighborhoods and locations in every recommended zone, taking into consideration additional factors like the parking lot of each location, traffic of existing coffee shops in the cluster, and current revenue of them, etc.