Things To Be Considered by Millenials Dreamer Before Living Abroad in Canada

-Applied Data Science Project-

Introduction

- Milennials in my country are young,ambitious, & creative. They often have big dreams & hopes about what they want to achieve in their lives. Many of them have ever dreamed of living abroad in more developed countries.

- "Shoot the moon, at least it will drop among the stars" - is one of those inspirational quotes that they believe as one of their mottos.

- As energic & romantic as they can be, young people often forget to measure the distance between the earth, the moon, and the stars.

- Through data science, I believe everything needs to be measured,validized, clustered, analyzed, and planned, before it can be concluded (include those hopes & dreams) for a better result.

- In this project, I would like to show a brief about cities in Canada, one of those dreamy country for developing countries millenials.

- I hope young people would have a new perspective to measure their own capacities & select the best option to start their life, whether in one of their dream countries (Canada) or just being rational to stay at their home country & start to build their dreams from here.

Data

- **1. Postal Code of Canada**

- **2. Neighborhood Latitude & Longitude**

- **3. Venue Name, Latitude, Longitude, & Category**

- **4. Average House Pricing in Canada**

- **Postal Code, Latitude, Longitude are needed to plot the location of venue or public facilties. We can make cluster of venues based on their categories by their location to measure the venue/public facilities density on each area, How lively the areas will briefly defined by their facilities. Facilities would be correlated with cost of living. We can learn more about the cost of living starting from the house pricing. So, compare to our usual income or expense in our home country, have you prepared for those lives you dreamed about? Let us measure it together ^0^**

Methodology

- **Using K-Means Clustering Algorythm to explore and create grouping of those neighborhoods into clusters to define similar characteristics of the area**
- **Mapping nearby venues of the areas using Foursquare credential**
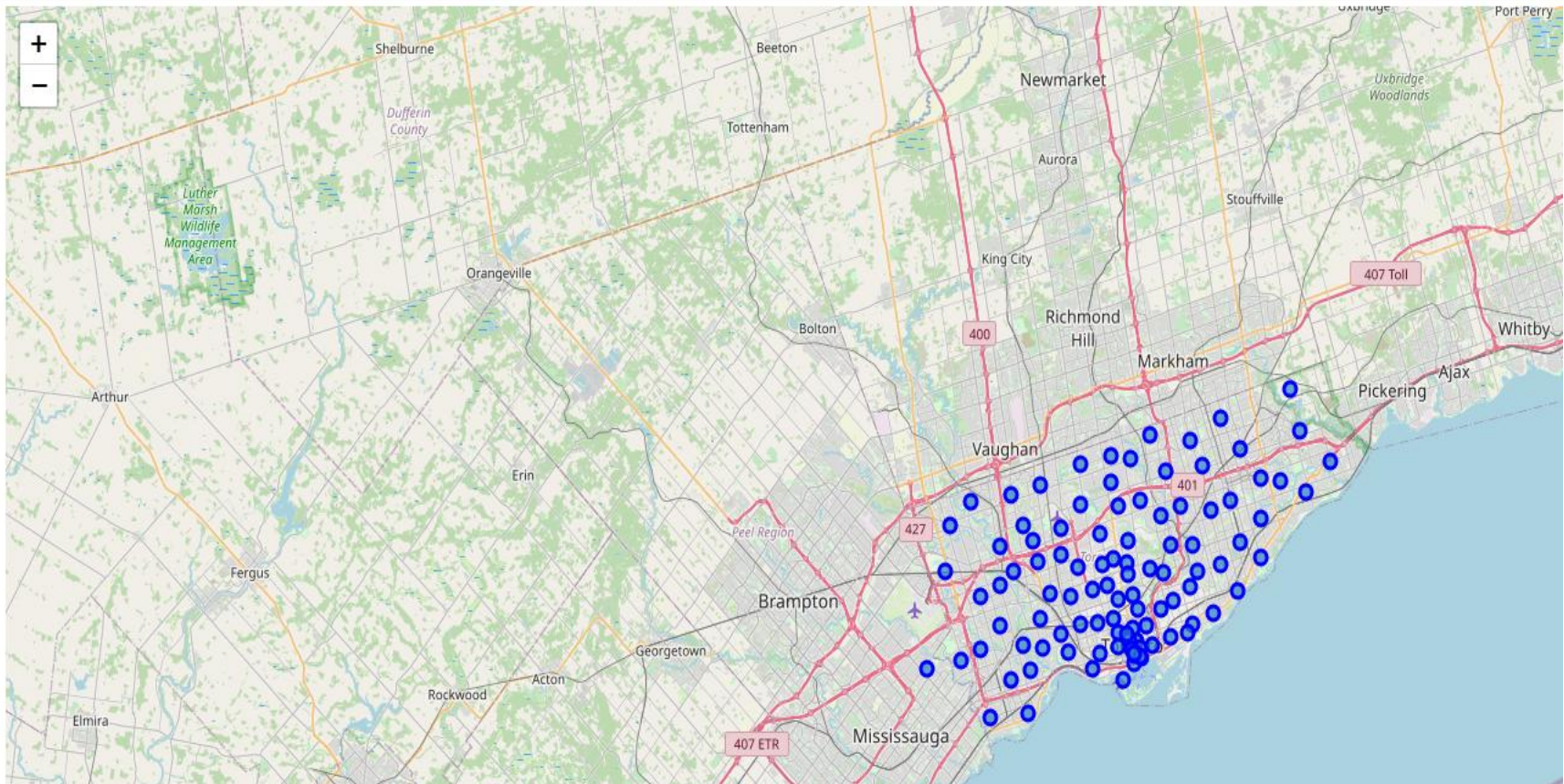
Result

- **1. Map of Scarborough (result as belows)**
- **2. Cluster of Area (result as belows)**
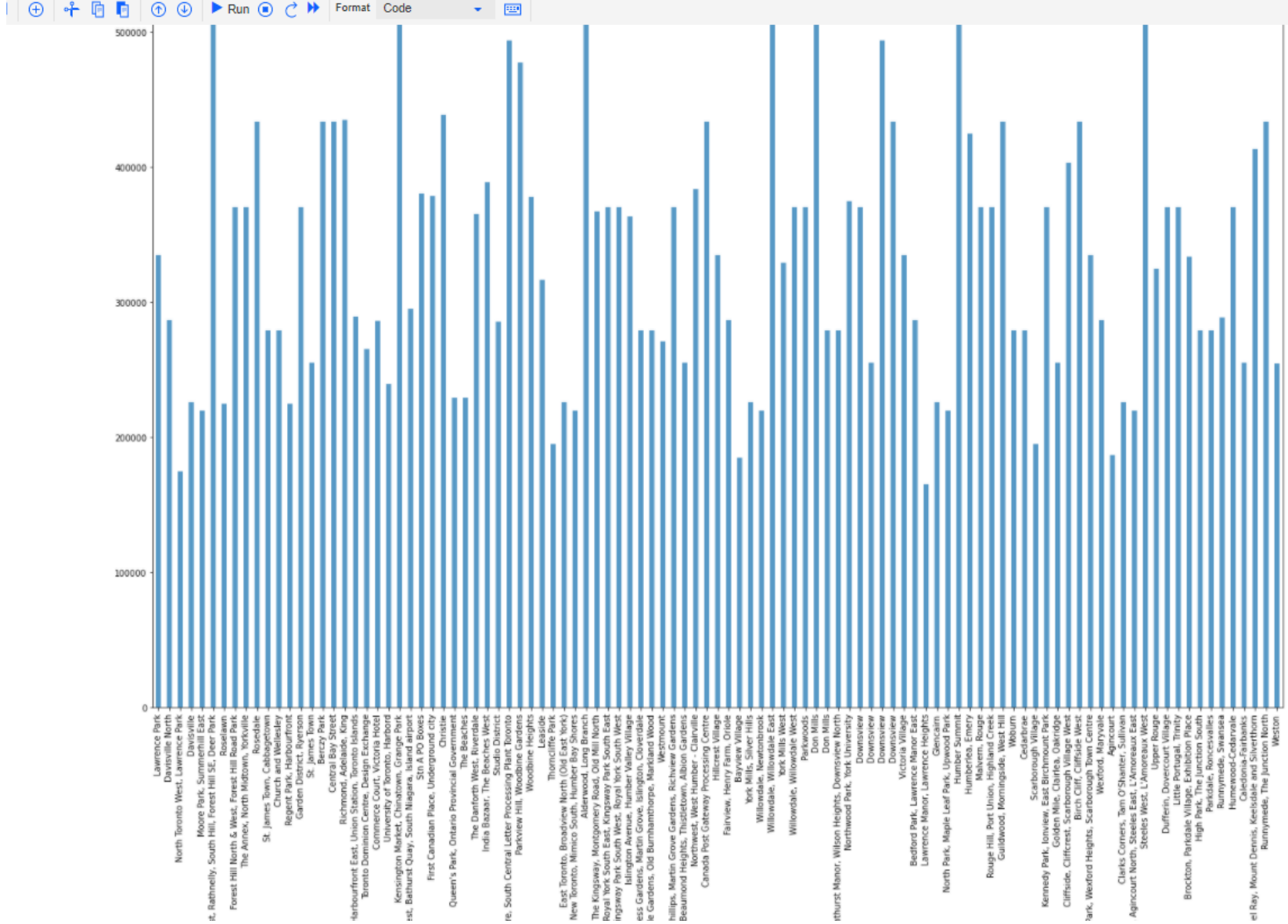- **3. Average Housing Price per Area (result as belows)**

```
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_Scarborough)

map_Scarborough
```

Out[115]:

```
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_venues(Scarborough_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head()
```

Out[104]:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Agincourt	Latin American Restaurant	Badminton Court	Coffee Shop	Pool Hall	Breakfast Spot	Lounge	Motorcycle Shop	Skating Rink	Women's Store	Dog Run
1	Alderwood, Long Branch	Pizza Place	Convenience Store	Gym	Sandwich Place	Gas Station	Pub	Athletics & Sports	Pharmacy	Coffee Shop	Community Center
2	Bathurst Manor, Wilson Heights, Downsview North	Coffee Shop	Park	Bank	Ice Cream Shop	Middle Eastern Restaurant	Pharmacy	Bridal Shop	Shopping Mall	Supermarket	Sushi Restaurant
3	Bayview Village	Bank	Chinese Restaurant	Japanese Restaurant	Intersection	Grocery Store	Skating Rink	Café	Women's Store	Distribution Center	Dive Bar
4	Bedford Park, Lawrence Manor East	Italian Restaurant	Coffee Shop	Sandwich Place	Women's Store	Toy / Game Store	Sushi Restaurant	Bank	Bakery	Bagel Shop	Restaurant

In [105]:
```
# Using K-Means to cluster neighborhood into 3 clusters
Scarborough_grouped_clustering = Scarborough_grouped.drop('Neighborhood', 1)
kmeans = KMeans(n_clusters=3, random_state=0).fit(Scarborough_grouped_clustering)
kmeans.labels_
```

Out[105]: array([0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 0, 0, 2,
        0, 0, 2, 0, 2, 0, 0, 1, 0, 0, 2, 0, 0, 0, 2, 2, 2, 1, 0, 0, 0,
        0, 2, 0, 0, 0, 0, 0, 2, 0, 1, 0, 0, 0, 0, 0, 0, 2, 1, 0, 0, 0,
        1, 0, 0, 0, 0, 2, 2, 0, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2,
        2, 2, 0, 1, 0, 2, 1, 0, 1], dtype=int32)

In [106]:
```
neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

Scarborough_merged =df_2.iloc[:16,:]

# merge toronto_grouped with toronto_data to add latitude/longitude for each neighborhood
Scarborough_merged = Scarborough_merged.join(neighborhoods_venues_sorted.set_index('Neighborhood'), on='Neighborhood')

Scarborough_merged.head()
```

Out[106]:

	Postalcode	Borough	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	M1B\n	Scarborough\n	Malvern, Rouge	0	Fast Food Restaurant	Coffee Shop	Hobby Shop	Spa	Women's Store	Dumpling Restaurant	Distribution Center	Dive Bar	Dog Run	Doner Restaurant
1	M1C\n	Scarborough\n	Rouge Hill, Port Union, Highland Creek	0	Breakfast Spot	Bar	Burger Joint	Electronics Store	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Dumpling Restaurant	Eastern European Restaurant
2	M1E\n	Scarborough\n	Guildwood, Morningside, West Hill	0	Restaurant	Park	Fast Food Restaurant	Sports Bar	Rental Car Location	Electronics Store	Pharmacy	Breakfast Spot	Beer Store	Mexican Restaurant
3	M1G\n	Scarborough\n	Woburn	1	Park	Coffee Shop	Business Service	Eastern European Restaurant	Distribution Center	Dive Bar	Dog Run	Doner Restaurant	Donut Shop	Drugstore
4	M1H\n	Scarborough\n	Cedarbrae	0	Indian Restaurant	Coffee Shop	Gas Station	Fried Chicken Joint	Caribbean Restaurant	Lounge	Athletics & Sports	Flower Shop	Gym / Fitness Center	Chinese Restaurant

In [109]:
```
kclusters=10
# create map
map_clusters = folium.Map(location=[latitude_x, longitude_y], zoom_start=11)

# set color scheme for the clusters
x = np.arange(kclusters)
colors_array = cm.rainbow(np.linspace(0, 1, kclusters))
rainbow = [colors.rgb2hex(i) for i in colors_array]
print(rainbow)
# add markers to the map




map_clusters
```

Edit   View   Insert   Cell   Kernel   Help

⊕   ✂   ▤   ▥   ⊕ ⊗   ▶ Run   ■   ↻   ▶▶   Format   Code   ▾   ⌨

Discussion

- **1. Scarborough geographical location : Scarborough is contained within the borders of Victoria Park Avenue on the west, Steeles Avenue to the north, Rouge River and the city of Pickering to the east, and Lake Ontario to the south. It borders Old Toronto, East York and North York in the west and the city of Markham in the north.**

- **2. Food & Beverages Venue is the most common categories in Scarborough**

- **3. Average Housing value per area starting from 1650000-5790000**

Recommendation

- 1. Food & beverages business in Scarborough is promising, yet challenging based on the categories of venue nearby. If someone want to start a food & beverages business, it would be highly recommend to have a specialization and differentiation of the menu to compete with the Western Food Restaurant and The Coffee Shop.

- 2. Food & beverages needs in Scarborough can be fulfilled, as long as the new immigrant can adapt with the Western Food and accept the rarity of Asian or African food,

- 3. The housing cost in the area are for about 10-20 times more expensive than the average housing cost in our home country

- 4. It is highly recommended that you should have for at least 10 times higher income per month compare to our home country GDP to safely starting a new life in Canada