Final report – Capstone project

INTRODUCTION

The purpose of my project is to help people in exploring better facilities around their neighborhood. The idea is to select a great neighborhood out of numbers of other neighborhoods in Downtown Toronto. I want to help people decide which is the best place to live, considering factors such as nearby places to eat, schools, etc. It is ideal for people who want to feel safe before making the decision to move.

DATA SECTION

Data Link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

I will use Downtown Toronto dataset which we scrapped from wikipedia. Dataset consisting of latitude and longitude, zip codes.

I will also use Foursquare and the information i will obtain is: Neighborhood, Neighborhood Latitude, Neighborhood Longitude, Venue, Name of the venue, Venue Latitude, Venue Longitude and Venue Category.

METHODOLOGY SECTION

Using K-Means Clustering Approach

	Postalcode	Borough	Neighborhood	Latitude	Longitude	Cluster Label	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
0	M1B\n	Scarborough\n	Malvern, Rouge	43.81153	-79.19552	1	1	Zoo Exhibit	Fast Food Restaurant	Climbing Gym	Falafel Restaurant	Dumpling Restaurant	Eastern European Restaurant	Elec Stor
1	M1C\n	Scarborough\n	Rouge Hill, Port Union, Highland Creek	43.78564	-79.15871	1	1	Moving Target	Fish & Chips Shop	Bar	Yoga Studio	Eastern European Restaurant	Electronics Store	Elem Sch
2	M1E\n	Scarborough\n	Guildwood, Morningside, West Hill	43.76575	-79.17520	2	2	Park	Gym / Fitness Center	Athletics & Sports	Doner Restaurant	Dumpling Restaurant	Eastern European Restaurant	Elec Stor
3	M1G\n	Scarborough\n	Woburn	43.76820	-79.21761	2	2	Coffee Shop	Chinese Restaurant	Fast Food Restaurant	Park	Event Space	Dumpling Restaurant	East Eurc Rest
4	M1H\n	Scarborough\n	Cedarbrae	43.76969	-79.23944	1	1	Thai Restaurant	Indian Restaurant	Bakery	Flower Shop	Gas Station	Caribbean Restaurant	Athl Spo

METHODOLOGY SECTION

Most Common venues near

	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	Chinese Restaurant	Shopping Mall	Pizza Place	Pool Hall	Asian Restaurant	Restaurant	Bank	Bakery	Sandwich Place	Latin American Restaurant
1	Alderwood, Long Branch	Sandwich Place	Print Shop	Coffee Shop	Gym	Pub	Gas Station	Pizza Place	Convenience Store	Field	Electronics Store
2	Bathurst Manor, Wilson Heights, Downsview North	Park	Convenience Store	Other Great Outdoors	Event Space	Dumpling Restaurant	Eastern European Restaurant	Electronics Store	Elementary School	Ethiopian Restaurant	Falafel Restaurant
3	Bayview Village	Dog Run	Golf Driving Range	Asian Restaurant	Trail	Park	Doner Restaurant	Donut Shop	Dumpling Restaurant	Eastern European Restaurant	Electronics Store
4	Bedford Park, Lawrence Manor East	Italian Restaurant	Pizza Place	Coffee Shop	Restaurant	Sandwich Place	Breakfast Spot	Pet Store	Juice Bar	Sushi Restaurant	Sports Club

RESULTS SECTION

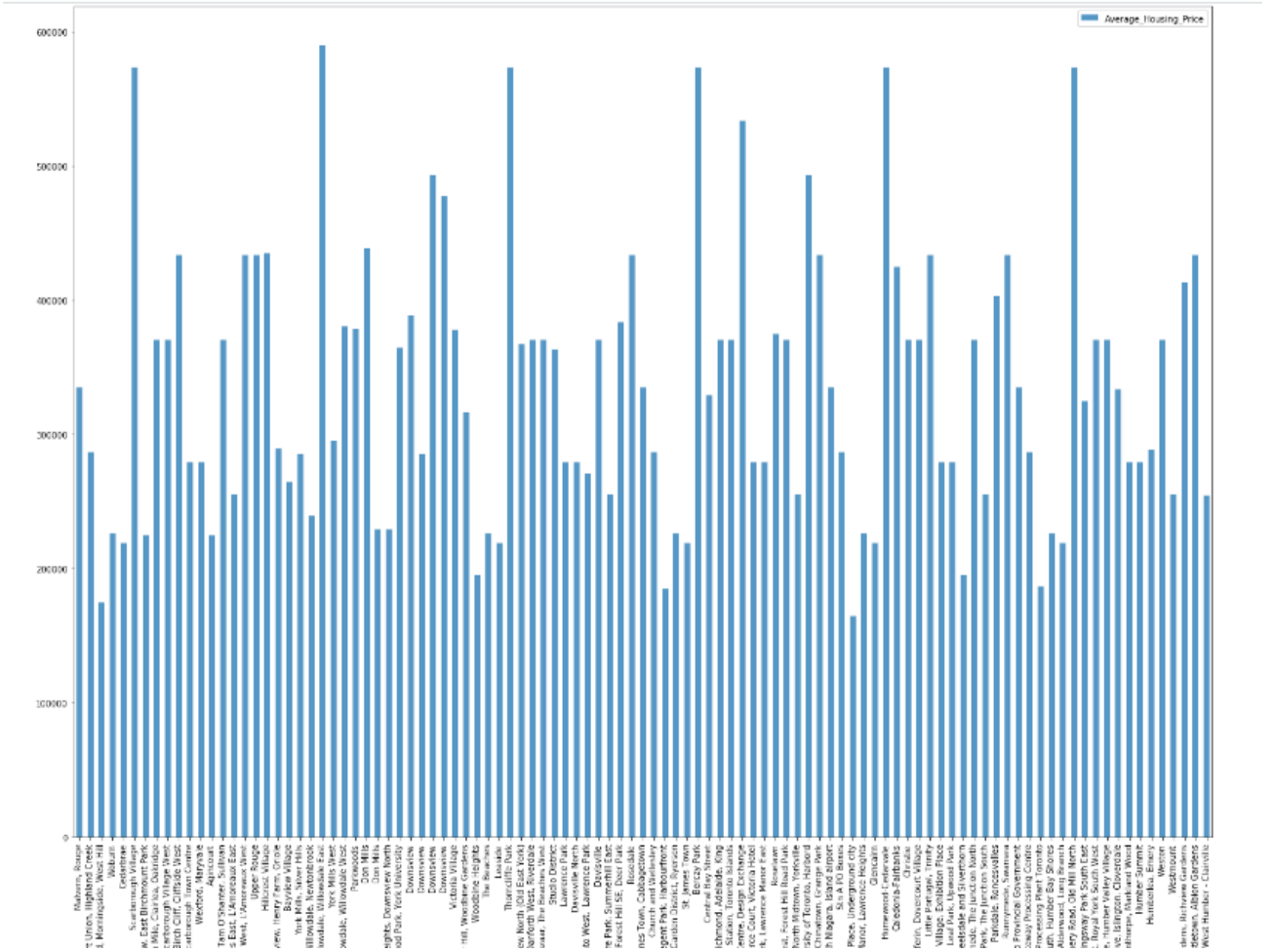
Map

```
# 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

markers_colors = []
for lat, lon, nei , cluster in zip(DowntownToronto_merged['Latitude'],
                                   DowntownToronto_merged['Longitude'],
                                   DowntownToronto_merged['Neighborhood'],
                                   DowntownToronto_merged['Cluster Label']):
    label = folium.Popup(str(nei) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

```
['#8000ff', '#4856fb', '#10a2f0', '#2adddd', '#62fbc4', '#9cfba4', '#d4dd80', '#ffa256', '#ff562c', '#ff0000']
```

RESULTS SECTION

Average housing price



DISCUSSION SECTION

Problem Which Tried to Solve:

The purpose of my project is to help people in exploring better facilities around their neighborhood. The idea is to select a great neighborhood out of numbers of other neighborhoods in Downtown Toronto.