In [ ]:

Capstone Project - The Battle of the Neighborhoods

Toronto Canada

Applied Data Science Capstone by IBM/Coursera ---- Valeriu Badea

Table of contents

--Text

In [ ]:

Introduction: Business Problem

Some investors from my native country, Romania asked me to analyze the investment possibilities in Toronto, Canada. The project can be also used by Toronto visitors, people who want to relocate in Toronto and investors interested in another kind of business Anyway their first preference is a traditional Romanian restaurant. The business requirements are large and the most important information come from financial data of the similar businesses around the chosen location. How many clients, how much they spend, what is the preferred menu are questions which will be analyzed in the followings studies. This initial project is going to respond to general problems like:

1. Position (how many similar businesses are around, landscape, parking, local traffic, etc.
    - Proximity to other businesses. Neighboring businesses may influence your store's sales, and their presence can work for you or against you.
    - Customer parking facilities. The site should provide convenient, adequate parking as well as easy access for customers.
    - Accessibility. Consider how easy it will be for customers to reach into your business. If you are mostly considering pedestrian traffic, consider whether or not nearby businesses will generate foot traffic for you.
    - The rent-paying capacity. If you have done a sales-and-profit projection for your first year of operation, you will know approximately how much revenue you can expect to generate, and you can use that information to decide how much rent you can afford to pay.
    - Traffic density. you have to check for food traffic, and to approximate sales potential of each pedestrian passing a given location. Two factors are very important: total pedestrian traffic during business hours and the percentage of it that is likely to patronize your food service business.
    - History of the site. Take out the recent history of each site into consideration before you make a final selection. Who were the previous tenants, and why are they no longer there?
    - Anticipated sales volume. Can the location contribute to your sales volume? Consider how easy it will be for customers to get into your business.
2. Demographic (nationality, income, age, etc.). Age is very important if you consider a restaurant business.
    - Generation Y (born between 1965 and 1980) is the most ethnically diverse generation yet and is more than three times the size of generation X. They are a prime target for a food-service business. About 25 percent of their restaurant visits are to burger franchises, follow by pizza restaurants at 12 percent.
    - Baby boomers. (born between 1946 and 1964) They represent a group of affluent professionals who can afford to visit upscale restaurants and spend money freely. During the 1980s, they were the main customer group for upscale, trendy restaurants. Today, those on the leading edge of the boomer generation are becoming grandparents, making them best customers for restaurants that offer a family-friendly atmosphere and those that provide an upscale, formal dining experience.
    - Seniors age 65 and older. Generally, on fixed incomes and may not often be able to afford upscale restaurants. They like to visit family-style restaurants that offer good service and reasonable prices. "Younger" seniors are likely to be more active and have more disposable income than "older" seniors, whose health may be declining. Seniors typically appreciate restaurants that offer early-bird specials and senior menus with lower prices and smaller portions, since their appetites are less hearty than those of younger people.

Go on top

In [ ]:

Data

- After understanding the business I started an extensive search process for the most suitable data.
- The data have been clear out from nonimportant information. Useful tables have been copied and meaningful information has been stored in pandas dataframes which could be visualized and analyzed
- The main idea is to use foursquare to find out the venues around some centroids,and after that to analyze the best opportunity to place a business,in our case a restaurant.
- In order to do this a grid of equidistance centroids should be created to cover the whole city and its borrows.
- In Toronto case this isn't necessary because only a part of Postal offices cover the whole city.
- A subroutine in my code shows that the distances between any different postal offices locations is less than 700 m, so the whole city can be covered by just a group of postal office location using a search radius of 500 m.
- I am happy about this discovery because I can use some work already done. Thank' s
- Information from 'Demographics_of_Toronto_neighbourhoods' help me to create comparations between Toronto neighborhoods pertinent to establishment of a new business

I found the following useful data:

- List of postal codes of Toronto Canada from Wikipedia 'http://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'
- IBM file 'http://cocl.us/Geospatial_data'
- Foursquare API
- Latitude/Longitude Distance Calculator 'https://www.nhc.noaa.gov/gccalc.shtml'
- Demographics of Toronto Neighbourhoods from Wikipedia 'https://en.wikipedia.org/wiki/Demographics_of_Toronto_neighbourhoods'

Go on top

In [ ]:

In [13]:

```
# Use folium to generate a map of Toronto, Canada with postal codes included

map_Toronto = folium.Map(location =[latitude,longitude], zoom_start =10)

for lat, lng, postal_code  in zip( df1['Latitude'], df1['Longitude'], df1['Postal Code']):
    label = '{}'.format(postal_code)
    label = folium.Popup(label, parse_html= True)
    folium.CircleMarker([lat, lng], radius = 5, popup = label, color = 'blue', fill= True,
fill_color= '#3186cc', fill_opacity= 0.7, parse_htlm= False).add_to(map_Toronto)
map_Toronto
```

Out[13]:

The position of the selected postal codes on the Toronto map Go on top

In [ ]:

Methodology

- Determining the location of any business is very important especially in the case of a restaurant.
- The venues around can strongly influence positively or negatively the turnover of any business
- The price of buying or renting the location also depends on the position of the restaurant in different areas of the city.
- In my case, when you want to create a restaurant with a national profile, you need to pay a careful attention to the demographic analysis of the city.
- To accomplish this I will use foursquare. This software scans the area around a point on a certain radius.
- It retrieves a json file with information regarding to the venues located in that area.
- My job would be to form a network of equidistant nodes that would cover the whole city
- Fortunately, this is not necessary in my case. The file "List_of_postal_codes_of_Canada:_M" which I used in a previous project can be used as a grid.
- Some lines of code proves that the distance between any 2 locations is less than 1 km.

- Another problem could be that there is a possibility that a venue is framed at 2 nodes, but the the duplicates based on latitude, latitude and category are eliminated in the process of creating the final file(dataframe)
- A map of Toronto with the exploration points on it have been designed using Folium software.
- The collected data is processed for use of the k-mean cluster algorithm.
- I chose this k-mean algorithm to segment and group different areas of Toronto.
- The main idea is to see how and why this k-mean clustering algorithm divides Toronto from the venues business point of view.
- The analysis of these segmentation criteria would have helped me draw conclusions where it would be better to recommend the location of the restaurant.

Go on top

In [ ]:

In [67]:

```
# create map
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)

# set color scheme for the clusters
x = np.arange(kclusters)
print('x=', x)
ys = [i + x + (i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(toronto_merged['Latitude'], toronto_merged['Longitude'],
toronto_merged['Postal code'], toronto_merged['Cluster Labels']):
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
#        print('cluster= ', cluster),
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

```
x= [0 1 2 3 4]
```

Out[67]:

- Cluster no 2 has only 3 items and name it sport. It contains a lot of baseball fields, yoga places, etc.

In [73]:

```
print('DataFrame dimension: ', cluster1.shape)
cluster1
```

```
DataFrame dimension:  (85, 12)
```

Out[73]:

	Postal code	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	9 C
9	M1B	0	Fast Food Restaurant	Print Shop	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	D
18	M1C	0	Construction & Landscaping	History Museum	Bar	Yoga Studio	Eastern European Restaurant	Distribution Center	Dog Run	Doner Restaurant	Do
27	M1E	0	Rental Car Location	Breakfast Spot	Medical Center	Intersection	Electronics Store	Mexican Restaurant	Bank	Yoga Studio	Re
36	M1G	0	Coffee Shop	Korean Restaurant	Convenience Store	Yoga Studio	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	D
45	M1H	0	Hakka Restaurant	Athletics & Sports	Fried Chicken Joint	Bank	Bakery	Gas Station	Thai Restaurant	Caribbean Restaurant	Re
54	M1J	0	Playground	Yoga Studio	Drugstore	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Do
63	M1K	0	Discount Store	Department Store	Hobby Shop	Coffee Shop	Train Station	Drugstore	Diner	Distribution Center	
72	M1L	0	Bus Line	Bakery	Soccer Field	Ice Cream Shop	Bus Station	Metro Station	Intersection	Park	Re
81	M1M	0	American Restaurant	Motel	Eastern European Restaurant	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	D
90	M1N	0	College Stadium	Skating Rink	General Entertainment	Café	Doner Restaurant	Diner	Discount Store	Distribution Center	
99	M1P	0	Indian Restaurant	Brewery	Chinese Restaurant	Thrift / Vintage Store	Vietnamese Restaurant	Pet Store	Electronics Store	Eastern European Restaurant	D
108	M1R	0	Accessories Store	Auto Garage	Breakfast Spot	Shopping Mall	Sandwich Place	Middle Eastern Restaurant	Bakery	Smoke Shop	
117	M1S	0	Clothing Store	Lounge	Skating Rink	Breakfast Spot	Latin American Restaurant	Eastern European Restaurant	Distribution Center	Dog Run	Re
126	M1T	0	Pharmacy	Pizza Place	Bank	Fried Chicken Joint	Italian Restaurant	Thai Restaurant	Chinese Restaurant	Gas Station	Inte
144	M1W	0	Fast Food Restaurant	Chinese Restaurant	Pharmacy	Bank	Indian Restaurant	Electronics Store	Cosmetics Shop	Pizza Place	Coff
46	M2H	0	Mediterranean Restaurant	Athletics & Sports	Golf Course	Pool	Dog Run	Yoga Studio	Donut Shop	Diner	
55	M2J	0	Clothing Store	Coffee Shop	Fast Food Restaurant	Restaurant	Convenience Store	Shoe Store	Bakery	Food Court	Te
64	M2K	0	Café	Bank	Chinese Restaurant	Japanese Restaurant	Yoga Studio	Distribution Center	Dog Run	Doner Restaurant	Do
91	M2N	0	Ramen Restaurant	Coffee Shop	Restaurant	Café	Pizza Place	Sandwich Place	Juice Bar	Japanese Restaurant	F Re
109	M2R	0	Pizza Place	Home Service	Coffee Shop	Butcher	Bank	Pharmacy	Golf Course	Department Store	E Re
11	M3B	0	Caribbean	Gym / Fitness	Café	Athletics &	Japanese	Baseball	Donut	Distribution	

	Postal code	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	9... C...
			Restaurant	Center	Cafe	Sports	Restaurant	Field	Shop	Center	
20	M3C	0	Coffee Shop	Restaurant	Asian Restaurant	Gym	Beer Store	Clothing Store	Chinese Restaurant	Sporting Goods Shop	Sup
47	M3H	0	Coffee Shop	Bank	Fried Chicken Joint	Bridal Shop	Sandwich Place	Diner	Restaurant	Middle Eastern Restaurant	Sup
56	M3J	0	Miscellaneous Shop	Massage Studio	Caribbean Restaurant	Coffee Shop	Bar	Yoga Studio	Distribution Center	Dog Run	Re
74	M3L	0	Park	Grocery Store	Bank	Shopping Mall	Hotel	Yoga Studio	Donut Shop	Discount Store	Dis
83	M3M	0	Food Truck	Business Service	Baseball Field	Drugstore	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Yoc
92	M3N	0	Gym / Fitness Center	Discount Store	Grocery Store	Liquor Store	Athletics & Sports	Yoga Studio	Drugstore	Distribution Center	
3	M4A	0	Pizza Place	Coffee Shop	French Restaurant	Hockey Arena	Portuguese Restaurant	Intersection	Eastern European Restaurant	Electronics Store	E Re
12	M4B	0	Pizza Place	Gastropub	Pharmacy	Gym / Fitness Center	Breakfast Spot	Fast Food Restaurant	Bank	Intersection	At
21	M4C	0	Park	Cosmetics Shop	Beer Store	Diner	Dance Studio	Athletics & Sports	Curling Ice	Skating Rink	Vid
30	M4E	0	Health Food Store	Neighborhood	Pub	Trail	Yoga Studio	Doner Restaurant	Diner	Discount Store	Dis
39	M4G	0	Coffee Shop	Sporting Goods Shop	Bank	Furniture / Home Store	Burger Joint	Fish & Chips Shop	Sushi Restaurant	Supermarket	Sp
48	M4H	0	Indian Restaurant	Yoga Studio	Park	Sandwich Place	Burger Joint	Restaurant	Coffee Shop	Pizza Place	P
66	M4K	0	Greek Restaurant	Coffee Shop	Italian Restaurant	Frozen Yogurt Shop	Ice Cream Shop	Bookstore	Furniture / Home Store	Fruit & Vegetable Store	Groce
75	M4L	0	Fast Food Restaurant	Sandwich Place	Burrito Place	Sushi Restaurant	Brewery	Restaurant	Pub	Italian Restaurant	Fish
84	M4M	0	Café	Coffee Shop	Bakery	American Restaurant	Gastropub	Brewery	Yoga Studio	Ice Cream Shop	Fisl
93	M4N	0	Park	Swim School	Construction & Landscaping	Bus Line	Yoga Studio	Donut Shop	Discount Store	Distribution Center	
102	M4P	0	Department Store	Hotel	Gym	Park	Breakfast Spot	Dog Run	Sandwich Place	Food & Drink Shop	
111	M4R	0	Clothing Store	Coffee Shop	Restaurant	Dessert Shop	Café	Diner	Mexican Restaurant	Furniture / Home Store	Re
120	M4S	0	Dessert Shop	Sandwich Place	Pizza Place	Sushi Restaurant	Gym	Italian Restaurant	Café	Coffee Shop	Danc
138	M4V	0	Pub	Coffee Shop	Bagel Shop	Pizza Place	Supermarket	Sushi Restaurant	Liquor Store	Fried Chicken Joint	Sp
156	M4X	0	Coffee Shop	Park	Pizza Place	Italian Restaurant	Bakery	Restaurant	Pub	Café	Pla
165	M4Y	0	Coffee Shop	Sushi Restaurant	Japanese Restaurant	Restaurant	Yoga Studio	Gay Bar	Grocery Store	Hotel	Medit Re
4	M5A	0	Coffee Shop	Bakery	Pub	Park	Breakfast Spot	Café	Theater	Yoga Studio	C
13	M5B	0	Clothing Store	Coffee Shop	Café	Japanese Restaurant	Restaurant	Middle Eastern Restaurant	Bubble Tea Shop	Cosmetics Shop	Re
22	M5C	0	Coffee Shop	Café	Gastropub	Cocktail Bar	American Restaurant	Seafood Restaurant	Theater	Creperie	De
31	M5E	0	Coffee Shop	Cocktail Bar	Café	Restaurant	Bakery	Beer Bar	Seafood Restaurant	Cheese Shop	
40	M5G	0	Coffee Shop	Café	Italian Restaurant	Sandwich Place	Ice Cream Shop	Middle Eastern Restaurant	Thai Restaurant	Bar	Bur
49	M5H	0	Coffee Shop	Café	Restaurant	Deli / Bodega	Gym	Hotel	Thai Restaurant	Clothing Store	Re
58	M5J	0	Coffee Shop	Aquarium	Hotel	Café	Fried Chicken Joint	Sporting Goods Shop	Scenic Lookout	Restaurant	Re

	Postal code	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	
67	M5K	0	Coffee Shop	Hotel	Café	Japanese Restaurant	Restaurant	Seafood Restaurant	Salad Place	Deli / Bodega	A Re
76	M5L	0	Coffee Shop	Café	Restaurant	Hotel	American Restaurant	Gym	Deli / Bodega	Italian Restaurant	J Re
85	M5M	0	Coffee Shop	Restaurant	Sandwich Place	Italian Restaurant	Sushi Restaurant	Greek Restaurant	Thai Restaurant	Cosmetics Shop	P
103	M5P	0	Jewelry Store	Park	Sushi Restaurant	Bus Line	Trail	Yoga Studio	Discount Store	Distribution Center	
112	M5R	0	Café	Sandwich Place	Coffee Shop	Cosmetics Shop	Indian Restaurant	Middle Eastern Restaurant	Pub	BBQ Joint	Bur
121	M5S	0	Café	Bookstore	Bar	Italian Restaurant	Japanese Restaurant	Restaurant	Bakery	Chinese Restaurant	Dess
130	M5T	0	Café	Coffee Shop	Vietnamese Restaurant	Mexican Restaurant	Bakery	Bar	Vegetarian / Vegan Restaurant	Dessert Shop	Gam
139	M5V	0	Airport Service	Airport Lounge	Airport Terminal	Harbor / Marina	Sculpture Garden	Rental Car Location	Plane	Boutique	Boat
148	M5W	0	Coffee Shop	Café	Japanese Restaurant	Restaurant	Seafood Restaurant	Beer Bar	Hotel	Cocktail Bar	Re
157	M5X	0	Coffee Shop	Café	Hotel	Japanese Restaurant	Restaurant	Gym	Seafood Restaurant	Salad Place	Ste
5	M6A	0	Clothing Store	Furniture / Home Store	Accessories Store	Coffee Shop	Event Space	Shoe Store	Sporting Goods Shop	Miscellaneous Shop	Arts
14	M6B	0	Metro Station	Pizza Place	Park	Japanese Restaurant	Pub	Doner Restaurant	Diner	Discount Store	Dis
23	M6C	0	Field	Dog Run	Hockey Arena	Trail	Donut Shop	Diner	Discount Store	Distribution Center	Re
41	M6G	0	Grocery Store	Café	Park	Restaurant	Candy Store	Baby Store	Italian Restaurant	Diner	Cof
50	M6H	0	Bakery	Pharmacy	Grocery Store	Gas Station	Café	Recording Studio	Middle Eastern Restaurant	Bar	Sup
59	M6J	0	Bar	Restaurant	Vietnamese Restaurant	Café	Vegetarian / Vegan Restaurant	Asian Restaurant	Men's Store	Yoga Studio	A Re
68	M6K	0	Café	Nightclub	Coffee Shop	Breakfast Spot	Grocery Store	Bakery	Performing Arts Venue	Pet Store	
77	M6L	0	Bakery	Park	Construction & Landscaping	Basketball Court	Yoga Studio	Eastern European Restaurant	Dog Run	Doner Restaurant	Do
86	M6M	0	Fast Food Restaurant	Sandwich Place	Coffee Shop	Drugstore	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Do
95	M6N	0	Caribbean Restaurant	Bus Line	Brewery	Yoga Studio	Drugstore	Distribution Center	Dog Run	Doner Restaurant	Do
104	M6P	0	Café	Thai Restaurant	Mexican Restaurant	Grocery Store	Flea Market	Diner	Bar	Fried Chicken Joint	
113	M6R	0	Gift Shop	Bookstore	Coffee Shop	Restaurant	Dessert Shop	Movie Theater	Dog Run	Eastern European Restaurant	Re
122	M6S	0	Café	Pizza Place	Coffee Shop	Sushi Restaurant	Pub	Italian Restaurant	Tea Room	Yoga Studio	Dess
6	M7A	0	Coffee Shop	Sushi Restaurant	Diner	Yoga Studio	Park	Bar	Beer Bar	Sandwich Place	Bur
114	M7R	0	Hotel	Coffee Shop	Gym	American Restaurant	Intersection	Fried Chicken Joint	Middle Eastern Restaurant	Sandwich Place	Burr
168	M7Y	0	Yoga Studio	Auto Workshop	Comic Shop	Park	Pizza Place	Recording Studio	Restaurant	Burrito Place	
142	M8V	0	Pharmacy	Fast Food Restaurant	Fried Chicken Joint	Café	Restaurant	Coffee Shop	American Restaurant	Seafood Restaurant	Liqu
151	M8W	0	Pizza Place	Pharmacy	Athletics & Sports	Dance Studio	Coffee Shop	Pub	Sandwich Place	Skating Rink	
160	M8X	0	Park	Smoke Shop	Pool	River	Yoga Studio	Dog Run	Dessert Shop	Dim Sum Restaurant	
			Hardware	Tanning		Fast Food	Discount	Convenience	Burrito		S

	Postal code	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	9 C
178	M8Z	0	Hardware Store	Tanning Salon	Grocery Store	Fast Food Restaurant	Discount Store	Convenience Store	Burrito Place	Burger Joint	S
26	M9C	0	Pizza Place	Beer Store	Convenience Store	Coffee Shop	Cosmetics Shop	Café	Shopping Plaza	Liquor Store	F
107	M9P	0	Pizza Place	Coffee Shop	Chinese Restaurant	Sandwich Place	Intersection	Electronics Store	Eastern European Restaurant	Ethiopian Restaurant	D
116	M9R	0	Pizza Place	Bus Line	Sandwich Place	Mobile Phone Shop	Doner Restaurant	Discount Store	Distribution Center	Dog Run	Do
143	M9V	0	Grocery Store	Pizza Place	Fried Chicken Joint	Coffee Shop	Sandwich Place	Fast Food Restaurant	Beer Store	Pharmacy	C
152	M9W	0	Drugstore	Rental Car Location	Bar	Donut Shop	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Yog

In [ ]:

In [74]:

```
print('DataFrame dimension: ', cluster2.shape)
cluster2
```

```
DataFrame dimension:  (1, 12)
```

Out[74]:

	Postal code	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
94	M5N	1	Garden	Yoga Studio	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Drugstore	Donut Shop

In [ ]:

In [75]:

```
print('DataFrame dimension: ', cluster3.shape)
cluster3
```

```
DataFrame dimension:  (1, 12)
```

Out[75]:

	Postal code	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
80	M9L	2	Pizza Place	Dessert Shop	Farmers Market	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Drugstore	Donut Shop

In [ ]:

In [76]:

```
print('DataFrame dimension: ', cluster4.shape)
cluster4
```

```
DataFrame dimension:  (9, 12)
```

Out[76]:

	Postal code	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 Cor V
135	M1V	3	Playground	Park	Coffee Shop	Yoga Studio	Donut Shop	Diner	Discount Store	Distribution Center	Dog Run	Resta
100	M2P	3	Park	Bank	Convenience Store	Bar	Yoga Studio	Eastern European Restaurant	Dog Run	Doner Restaurant	Donut Shop	Drug
2	M3A	3	Food & Drink Shop	Park	Eastern European Restaurant	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Drugstore	
65	M3K	3	Airport	Park	Yoga Studio	Eastern European Restaurant	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Drug
57	M4J	3	Park	Convenience Store	Coffee Shop	Yoga Studio	Drugstore	Discount Store	Distribution Center	Dog Run	Doner Restaurant	
129	M4T	3	Park	Yoga Studio	Drugstore	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	E Eur Resta
147	M4W	3	Park	Playground	Trail	Yoga Studio	Donut Shop	Diner	Discount Store	Distribution Center	Dog Run	Resta
32	M6E	3	Park	Women's Store	Spa	Yoga Studio	Drugstore	Discount Store	Distribution Center	Dog Run	Doner Restaurant	
98	M9N	3	Park	Yoga Studio	Drugstore	Diner	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	E Eur Resta

In [ ]:

```

```

In [77]:

```
print('DataFrame dimension: ', cluster5.shape)
cluster5
```

```
DataFrame dimension:  (2, 12)
```

Out[77]:

	Postal code	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
169	M8Y	4	Baseball Field	Yoga Studio	Drugstore	Discount Store	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Eastern European Restaurant	Filipino Restaurant
89	M9M	4	Paper / Office Supplies Store	Baseball Field	Yoga Studio	Eastern European Restaurant	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Electronics Store

Results

- The k-mean clustering model found only 5 clusters.
- Cluster No 1 has 85 items.I will call it general.
- Cluster no 2 has only 1 items and I name it M5N as the code of postal code
- Cluster no 3 has only 1 item and name it M5L
- Cluster No 4 has 9 items and with a lot of parks in it.I will call it recreational
- Cluster No 5 has 2 items. I will call it outdoor, for its 3 playgrounds, 3 dog run places and 1 park.
- This practically means (with over 85 % of items in one cluster) that no enough differences were found between the collection points
- Another main aspect of the analysis is the designation of spots where similar businesses are located.
- In my case it would have been to determine the number of Romanian restaurants located around the centroids

- In my case it would have been to determine the number of Romanian restaurants located around the centroids.
- In Toronto is only one traditional Romanian restaurant named 'Moldova' located in East York.
- This situation makes the determining of the number and location of Romanian restaurants around the city of Toronto inappropriate.
- Anyway, just with several lines of codes this improvement can be added to the project
- Now I got the feeling that the data and information I have at this time are insufficient and inconclusive.
- A deeper analysis supported by additional data must be tried.
- So as the theory says I have to go back from the evaluation to the modeling phase.

[Go on top]

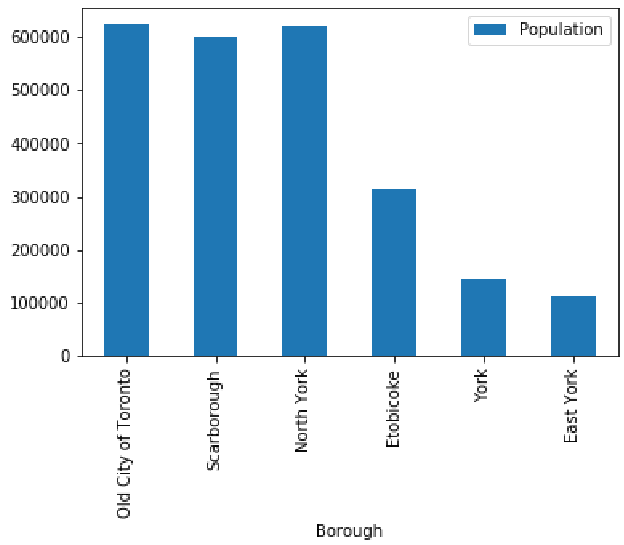
In [ ]:

In [78]:

```
dd_agg.plot.bar(x= "Borough", y= "Population")
```

Out[78]:

<matplotlib.axes._subplots.AxesSubplot at 0x7f74ab9f7748>



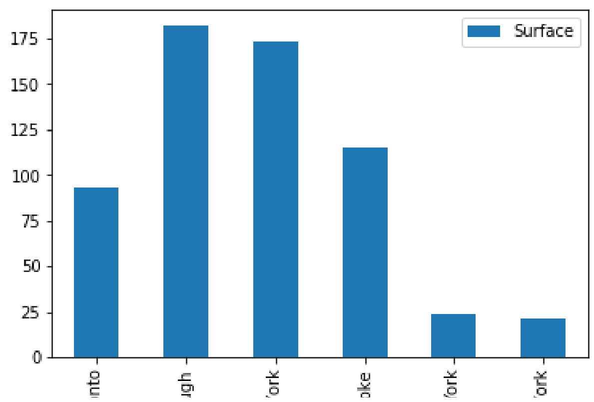
In [ ]:

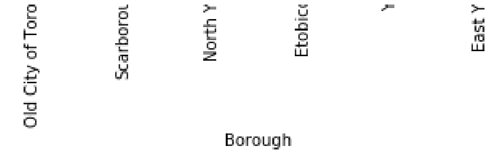
In [79]:

```
dd_agg.plot.bar(x= "Borough", y= "Surface")
```

Out[79]:

<matplotlib.axes._subplots.AxesSubplot at 0x7f74abbf07f0>





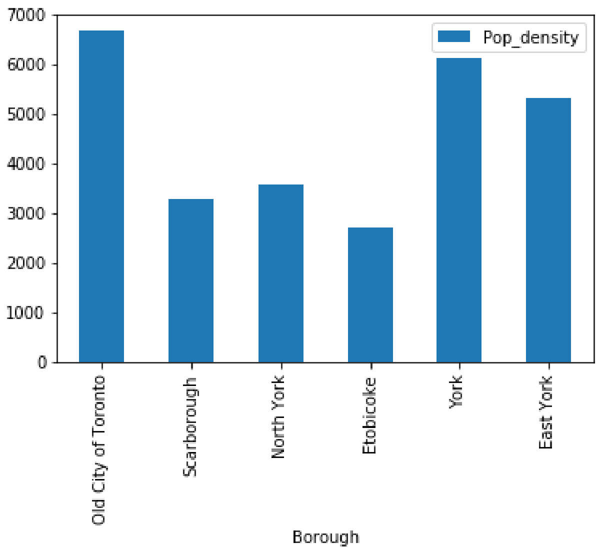
In [ ]:

In [80]:

```
dd_agg.plot.bar(x= "Borough", y= "Pop_density")
```

Out[80]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f74ab5f4c88>
```



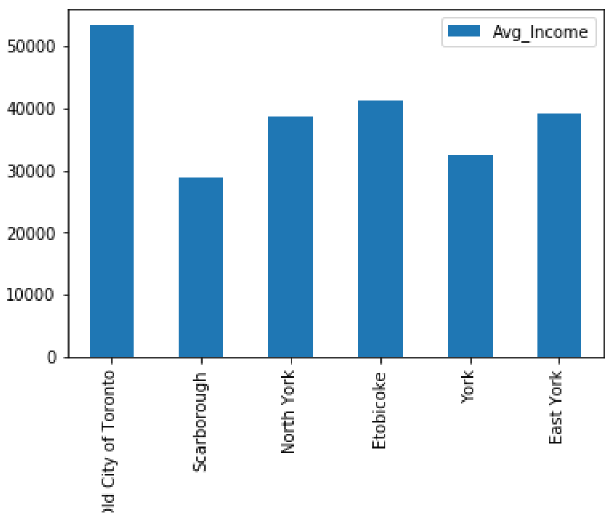
In [ ]:

In [81]:

```
dd_agg.plot.bar(x= "Borough", y= "Avg_Income")
```

Out[81]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f74ab5d42e8>
```



Borough

[Go on top]

Disscusion

- Given the fact that I am looking to optimize the location of a national restaurant, I will continue to focus more on the demographic aspect of the city of Toronto
- I am encouraged in this direction by the fact that the results from the k-mean clustering model are quite evasive and inconclusive.
- I found a file in Wikipedia that includes information about the population, average income, area, second language spoken after English, etc. The file gives all these info for the Toronto city and its borough and neighborhood
- The file is called "Demographics_of_Toronto_neighbourhoods"
- The file has been downloaded into a pandas dataframe, clean out, etc.
- It has been processed so that the meaningful data can be visualized and analyzed.

[Go on top]

In [ ]:

```

```

[Go on top]

Conclusion

- After examining the demographic data, it can be seen that "Old Toronto" is the city borough with the highest population but also, with the highest population density
- In terms of average earnings Old Toronto is again in first place with an average of 53407 dollars
- After than come Etobicoke with 41251 and East York with 39186 dollars.
- The last ones are North York with 38537, York with 32431 and at last Scarborough with 28806 dollars
- From the above mentioned file I didn't get any information regarding the Romanian community in Toronto.
- I searched the internet and I found out that in Toronto are only 30,000 Romanians.
- No any information on how they are spread in the city was found.
- Consequently, I looked for information on Eastern or Southeast European populations that may become future customers.
- I found out that in Old Toronto and East York there are 11 respective 21 neighborhoods where the population has a east or south east native European language.
- The only Romanian restaurant in Toronto, 'Moldova', is located in East York and is a 3.5 star restaurant under Russian management
- From their list of remarks I noticed that people with Romanian names leave an unsatisfactory message like 'I didn't find what I expected', the other people leave good or very good messages.
- I conclude from here that there is still a requirement of an authentic Romanian restaurant.

------ Finally, my recommendations are:

- to continue this study with a statistical survey among the Romanian population in Toronto
- to consider the establishment of a restaurant with a mixed profile (Romanian but also general European / North American)
- a careful examination of 'Old Toronto' and 'East York' boroughs which looks like the best candidates being on top of average income and population density and the most populated with east and south-east european people.

[Go on top]