The Battle of Neighborhoods

Introduction

A businessman wants to open a new restaurant, but he does not know where is the best place. He has information that New York City (EEUU) and Toronto (Canada) are good options to start, but he only has a budget to open only one.

This businessman wants to get better profits reducing cost and receiving many clients.

Problem

Use data to identify the best place where a new restaurant get better profits.

Interest

The businessman who wants to open a new restaurant.

Data

Data sources

Information of Toronto and New York is on Internet. In this case, I collected data from FourSquare, Wikipedia and Coursera.

- https://cocl.us/new_york_dataset
- https://geo.nyu.edu/catalog/nyu_2451_34572
- https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Data cleaning

Steps:

- Data was store in dataframes using pandas library.
- Missing values were dropped.
- To collect information about localization I used FourSquare API.
- Categorical features were transformed to numerical, and then were transformed to One-hot encoding.

Methodology

In this work, I decided to use unsupervised methodology like K-means clustering because I wanted to look for common behaviors in the data.

To work with K-means clustering I use categorical transformation, because it is easier to work with clustering.

New York data:

```
manhattan_merged.head() # check the last columns!
```

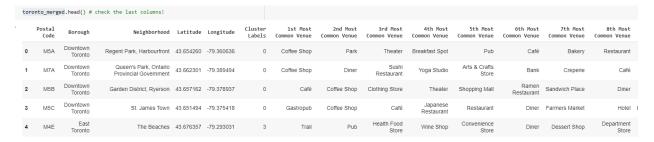|	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue
0	Manhattan	Marble Hill	40.876551	-73.910660	1	Sandwich Place	Coffee Shop	Gym	Yoga Studio	Tennis Stadium	Pizza Place	Deli / Bodega	Department Store
1	Manhattan	Chinatown	40.715618	-73.994279	1	Chinese Restaurant	Sandwich Place	Spa	Hotpot Restaurant	Bike Shop	Boutique	Bubble Tea Shop	English Restaurant
2	Manhattan	Washington Heights	40.851903	-73.936900	0	Café	Wine Shop	Park	Deli / Bodega	Tapas Restaurant	Indian Restaurant	Scenic Lookout	Frozen Yogurt Shop
3	Manhattan	Inwood	40.867684	-73.921210	0	Restaurant	Mexican Restaurant	Wine Bar	Deli / Bodega	Park	Frozen Yogurt Shop	Bakery	Café
4	Manhattan	Hamilton Heights	40.823604	-73.949688	0	Coffee Shop	Yoga Studio	Pizza Place	Mexican Restaurant	Caribbean Restaurant	Cocktail Bar	Italian Restaurant	Pub

Toronto data:

```
toronto_merged.head() # check the last columns!
```

	Postal Code	Borough	Neighborhood	Latitude	Longitude	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
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	0	Coffee Shop	Park	Theater	Breakfast Spot	Pub	Café	Bakery	Restaurant
1	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	0	Coffee Shop	Diner	Sushi Restaurant	Yoga Studio	Arts & Crafts Store	Bank	Creperie	Café
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	0	Café	Coffee Shop	Clothing Store	Theater	Shopping Mall	Ramen Restaurant	Sandwich Place	Diner
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	0	Gastropub	Coffee Shop	Café	Japanese Restaurant	Restaurant	Diner	Farmers Market	Hotel
4	M4E	East Toronto	The Beaches	43.676357	-79.293031	3	Trail	Pub	Health Food Store	Wine Shop	Convenience Store	Diner	Dessert Shop	Department Store

Results

The results show that most common venues in both cities are Coffee Shops. In case of restaurants, there are few of them, but they are more common in New York.

Toronto	New York
`[103] toronto_merged['1st Most Common Venue'].value_counts()`	`[118] manhattan_merged['1st Most Common Venue'].value_counts()`

Toronto:
```
Café                    11
Coffee Shop              7
Park                     4
Trail                    3
Restaurant               1
Airport Service          1
Gay Bar                  1
Pharmacy                 1
Mexican Restaurant       1
Dessert Shop             1
Garden Center            1
Garden                   1
Grocery Store            1
Gastropub                1
Clothing Store           1
Greek Restaurant         1
Bar                      1
Gift Shop                1
Name: 1st Most Common Venue, dtype: int64
```

New York:
```
Coffee Shop              5
Gym / Fitness Center     3
Sandwich Place           2
Chinese Restaurant       2
Italian Restaurant       2
Japanese Restaurant      2
Korean Restaurant        2
Theater                  2
Park                     2
French Restaurant        1
Café                     1
Ice Cream Shop           1
Cocktail Bar             1
Spa                      1
Hotel                    1
Greek Restaurant         1
Restaurant               1
Memorial Site            1
Bakery                   1
Karaoke Bar              1
Mexican Restaurant       1
American Restaurant      1
Gym                      1
Yoga Studio              1
Women's Store            1
Bagel Shop               1
Boat or Ferry            1
Name: 1st Most Common Venue, dtype: int64
```

Discussion

There is no any observation.

Conclusion

Showing the results, a better place to open a restaurant is in Toronto, because there less competitors.