

Battle of Neighbourhoods

Chirag Patel March 2020

1 Introduction

1.1 Background

It's trend of globalization and technology. Data is generated at a tremendous rate. Data is a great tool which can be used to solve several problems. Data can be used to predict outcome of certain event which can help humanity to a greater extend. Here also data will be used to generate an investment plan which will help an organization economically.

1.2 Business Problem

An investment group is looking for investing in Two biggest cities of North America, New York and Toronto respectively. They need an investment plan where they can invest their money. They are not sure which business are flourishing in that cities. They need to which neighbourhood would be best for their investment in those cities and which business would be best to invest in.

Data

2.1 Data Sources

Data needed for such business problem can be found on the web. The data for Toronto neighbourhood is available on Wikipedia and can be gathered using data scrapping. Data for New York neighbourhood is available in JSON format and is needed to be converted into tabular. Foursquare API will also be used to get venues by particular neighbourhood.

2.2 Data Cleaning and Formatting

Data for Toronto neighbourhood can be gathered using Beautiful Soup from Wikipedia and needs to be cleaned for further use and data for New York Neighbourhood need to be converted from JSON format to tabular format

3 Exploratory Data Analysis

It can be seen by maps of both cities that there are more neighbourhood in New York compared to Toronto.

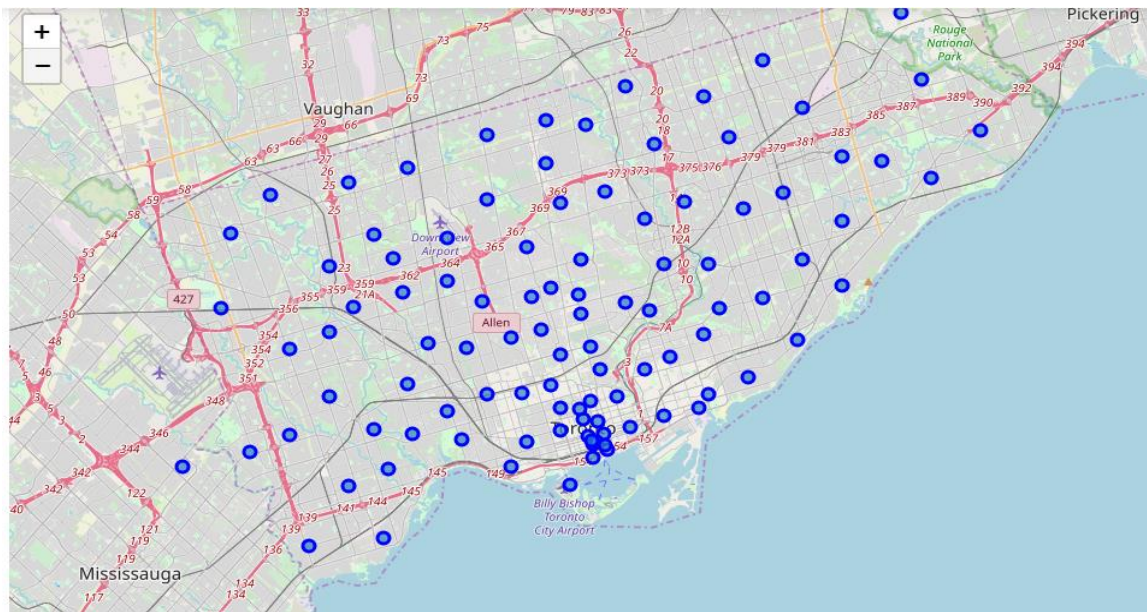


Fig 3.1.1 Map of New York by Neighbourhood

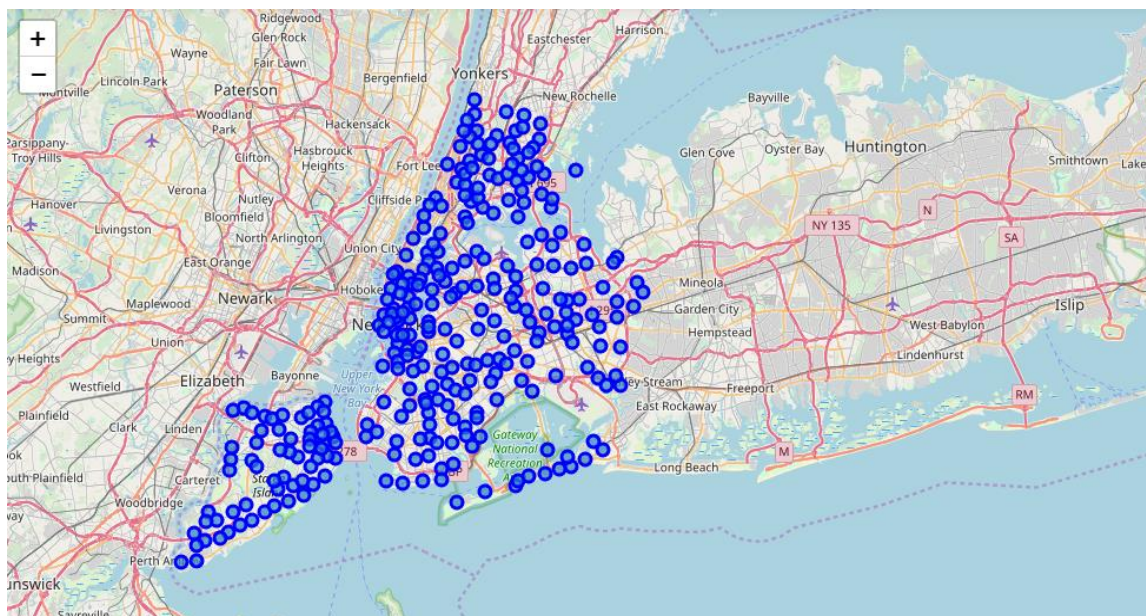
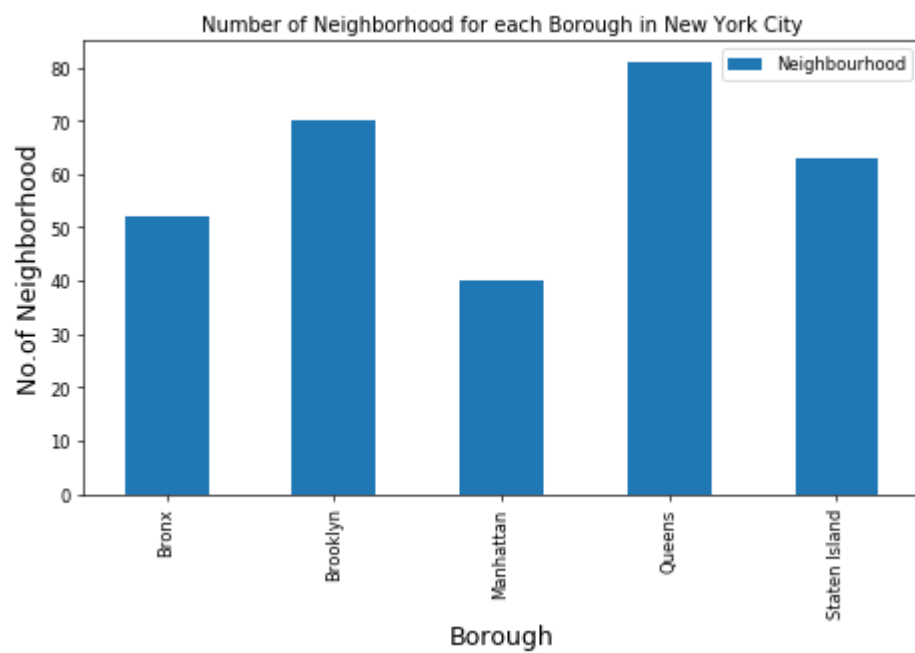
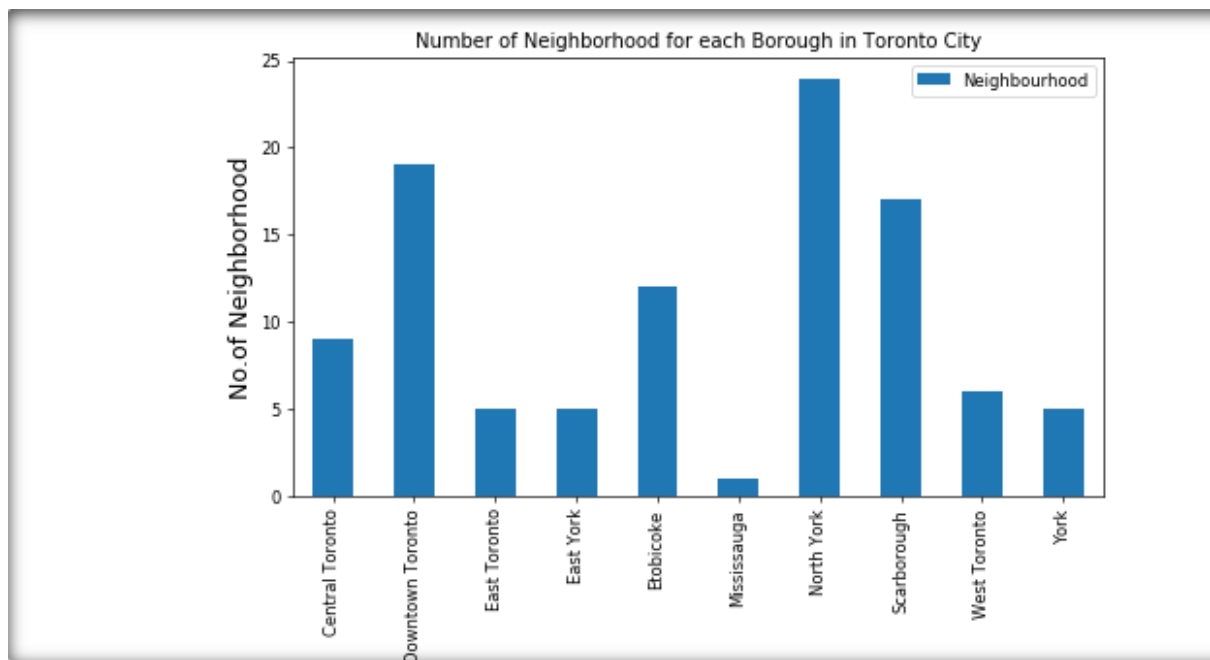


Fig 3.1.2 Map of New York by Neighbourhood



4 Results

It can be noted that Coffee shop is the most famous venue in North York followed by Clothing store. Pizza place and Deli are the most common venue in Queens Neighbourhood.

```
In [25]: north_york_venues['Venue Category'].value_counts()
```

```
Out[25]: ['Coffee Shop',  
          'Clothing Store',  
          'Restaurant',  
          'Park',  
          'Japanese Restaurant',  
          'Sandwich Place',  
          'Pizza Place',  
          'Grocery Store',  
          'Fast Food Restaurant',  
          'Sushi Restaurant']
```

```
In [53]: queens_venues['Venue Category'].value_counts()
```

```
Out[53]: ['Pizza Place',  
          'Deli / Bodega',  
          'Chinese Restaurant',  
          'Bakery',  
          'Donut Shop',  
          'Pharmacy',  
          'Bank',  
          'Bar',  
          'Grocery Store',  
          'Sandwich Place']
```

5 Recommendations

Top 5 venues varies in North York while all top 5 venues in Queens are related to food. So something involving food should be open in Queens while that's not case with the people of North York.

6 Conclusion

It can be concluded that most famous neighbourhood in Toronto is North York and Queens in New York. Hottest investment for Toronto is coffee shop while, for New York is Pizza Place. This comparison also shows that people in Toronto have more inclination towards coffee while, Population in New York prefer Pizzas. Thus investment should be made in coffee place in Toronto while in New York, Pizza place should be opened for maximum profit.