EATERY DATA ANALYSIS OF FRDERICTON

Data Analysis Report

Abstract
This report includes the detail of the analysis for eatery category in a Canada city with data science methodologies

TARUN MENDIRATTA
Mendiratta.t13@gmail.com

Table of Contents

Introduction:

Background

Fredericton is the capital city of Canadian province of New Brunswick, with a population of around 60 thousand. One of main urban centers in New Brunswick, It is the third largest city in the province.

As a provincial capital, its economy is tied to the public sector; however, the city also contains a growing IT and commercial sector. The city has the highest percentage of residents with post-secondary education in the province and the highest per capita income of any city in New Brunswick [1].

Being a resident of this city, I decided to explore the venues of Fredericton. The city is fully diverse with lot of communities and it is becoming denser as it attracts more people around the globe. Indeed, the importance of food diversity cannot be denied here.

Problem

The aim of the project is to explore the neighborhoods of Fredericton. As Fredericton's neighborhoods are not dense and are less in numbers, the desire will be to find out the eatery category and their share in food diversity in Fredericton city.

Interest

This project will help to figure out the different cuisines available in the city. Any government official or food industry expert can use the outcome of this analysis to know the food diversity in the city. A vendor, who is planning to set up their restaurant or food outlet, can take the help of this report to make their decision.

Data:

To explore the city, different sources were used to gather the data. Our aim is to get the food venues of city neighborhood, to achieve that we need city data and their venue data. Following data sources were inclusive part of this analysis.

Wikipedia

List of Postal codes of Canada: E [2]

This page has list of postal codes of New Brunswick FSAs (Forward Sortation Area). The page data was uploaded and the table of postal code was scraped. This data consist postal code of all neighborhoods, I used this data to get the spatial coordinates of Fredericton neighborhoods.

Foursquare API

Foursquare is a technology company that built a massive dataset of location data. Currently its location data is most comprehensive and quite accurate. It feeds location data for many popular services like Apple Maps, Uber and Twitter.

Using the API [3], I got the data of venues of Fredericton neighborhoods. API provided the response in JSON and venues were retrieved along with their category and other details. I used this data to explore the food diversity in the city.

Methodology:

Scrap Wiki page and get Neighborhoods

In order to explore the city, first we need the city's data. The data was retrieved from the Wikipedia page; I scraped the page data using BeautifulSoup package of Python. Following library were imported to use it.

> *!pip install BeautifulSoup4*
> *from bs4 import BeautifulSoup*

Using *prettify* method, I checked the Wikipedia page content and table class in it. I got the postal code table scraped and read the cells into a data frame.

The table from wiki page was converted into below format in pandas data frame.

	PostalCode	Borough	Neighborhood
0	E1A	Moncton	East
1	E2A	Bathurst	Not Assigned
2	E3A	Fredericton	North
3	E4A	Chipman	Not Assigned
4	E5A	Moores Mills	Not Assigned

Figure 1 New Brunswick Postal Codes

Get Spatial Coordinates

To represent the neighborhoods on map and to get the venues from foursquare API, the spatial coordinates are required. To get the spatial coordinate of neighborhoods, I used the Geocode library of Python.

Using Geocode, I was able to get the longitude and latitude of all the neighborhoods that I scraped in last step. Geocode provide a method 'query_postal_code', which return the coordinates detail along with place name by just passing the postal code as parameter.

As a result, following data frame was created:

	postal_code	place_name	latitude	longitude
0	E1A	Dieppe Moncton East	46.0967	-64.7087
0	E2A	Bathurst	47.6181	-65.6511
0	E3A	Fredericton North	46.0401	-66.3862
0	E4A	Bathurst	46.1599	-65.8102
0	E5A	Moores Mills	45.3705	-67.1737

Figure 2 New Brunswick Coordinates

Now I got the two data frames; in next step, both data frame were merged and in result following data frame was created:

	PostalCode	Borough	Neighborhood	latitude	longitude
0	E1A	Moncton	Dieppe Moncton East	46.0967	-64.7087
1	E2A	Bathurst	Bathurst	47.6181	-65.6511
2	E3A	Fredericton	Fredericton North	46.0401	-66.3862
3	E4A	Chipman	Bathurst	46.1599	-65.8102
4	E5A	Moores Mills	Moores Mills	45.3705	-67.1737

Figure 3 New Brunswick Coordinates With Borough

The shape of the data frame is (111, 5), which means there are 111 neighborhood in New Brunswick. The postal codes, which are not assigned, have been cleaned out from the data frame.

Above data frame consist all the boroughs of New Brunswick province. It can be visualized on a map using folium library; following picture is the representation of New Brunswick with neighborhoods superimposed on the top, using *folium* library.
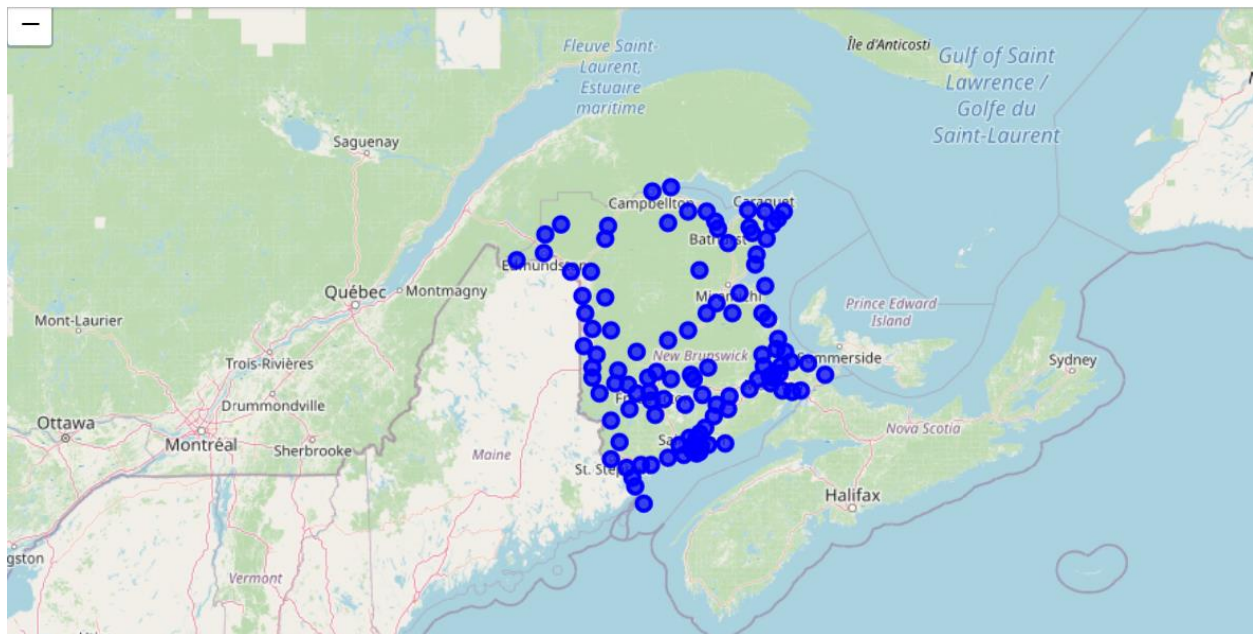
Figure 4 Neighborhood Superimposed on New Brunswick Map

As I am looking to explore only Fredericton neighborhoods, I separated out the data for borough 'Fredericton' in a new data frame.

	PostalCode	Borough	Neighborhood	latitude	longitude
0	E3A	Fredericton	Fredericton North	46.0401	-66.3862
1	E3B	Fredericton	Fredericton South New Brunswick Provincial Gov...	45.8119	-66.6763
2	E3C	Fredericton	Fredericton Southwest, New Maryland	45.9020	-66.7057
3	E3G	Fredericton	Fredericton	46.0546	-66.7344

Figure 5 Fredericton Neighborhoods

Get Venues from Foursquare:

Client ID and Secret are required to call the Foursquare API, which can be generated by signing-up on foursquare developer site. Foursquare provides API end points with different filter option. I first tried to get the venues for the entire neighborhood in Fredericton. The Fredericton neighborhoods are not dense, so I kept the radius as 10000 meter and limit of 100. After the data received from Foursquare, JSON was interpreted and transformed into following data frame:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Fredericton North	46.0401	-66.3862	Shell	46.022025	-66.190394	Gas Station
1	Fredericton North	46.0401	-66.3862	Garden's Gate Landscaping	46.144311	-66.555766	Construction & Landscaping
2	Fredericton North	46.0401	-66.3862	Casey's Diner	45.891275	-66.308779	Diner
3	Fredericton North	46.0401	-66.3862	McConnell's Welding Ltd	46.145906	-66.576328	Construction & Landscaping
4	Fredericton South New Brunswick Provincial Gov…	45.8119	-66.6763	Grant Harvey Centre	45.925002	-66.641004	Hockey Arena

Figure 6 Fredericton Neighborhood Venues

The shape of above data frame is (301, 7).To know the unique categories, I did group by on venue category and I found out there are 98 unique venues in the Fredericton neighborhood data frame.

Exploratory data Analysis:

The merged data frame has information for all the venues of neighborhoods. I did the one hot encoding of the venues; in this, I converted the venue categories into the columns of a new data frame and added neighborhood column into it. The resulting data frame was grouped based on neighborhoods.

The shape of the new data frame was (4, 53) that depicts there are 52 unique categories of venues. Using this data frame, I did further analysis on it, to find out three most common venues of Fredericton neighborhoods. Following are the results:

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Fredericton	Restaurant	Coffee Shop	Grocery Store
1	Fredericton North	Construction & Landscaping	Gas Station	Diner
2	Fredericton South New Brunswick Provincial Gov…	Coffee Shop	Restaurant	Grocery Store
3	Fredericton Southwest, New Maryland	Restaurant	Coffee Shop	Pub

Figure 7 Fredericton Most Common Places

The result tells us the three most common venues of Fredericton neighborhoods. As we see, the most common venues include restaurants for three neighborhoods. However, North Fredericton is the only neighborhood that does not have any restaurant in common venue. North Fredericton's population is very less as compare to the other parts of Fredericton. As a result, it has less number of outlets comparatively.

There are only four neighborhoods in Fredericton; hence clustering the neighborhoods is not a good option here. Thus, I did further analysis on the restaurant category to know the taste of Fredericton.

I used the Foursquare API again but this time category id was passed for food, to get venues only of food category. The JSON content was interpreted the same way as we did before and transferred into a data frame. There are 22 unique food categories that also contains general categories like Diner, Food and Food Court. I made a list of general categories and removed those venues from the data frame.

As Fredericton is a small city, instead of exploring the eatery category of neighborhoods of Fredericton, I decided to explore the eatery category in Fredericton city. Thus, I added one column in data frame for borough and inserted the value as Fredericton.

To make the data more structured for analysis; I made the venue categories as columns and performed a group-by on 'Borough' with mean values of categories multiply by 100. It gave me the percentage value of each category for the Fredericton city. The data frame with mean value of categories is depicted below:

	American Restaurant	Bakery	Café	Chinese Restaurant	Fast Food Restaurant	Gastropub	Greek Restaurant	Italian Restaurant	Japanese Restaurant	Korean Restaurant	Mexican Restaurant	Middle Eastern Restaurant
Borough												
Fredericton	3.08642	4.320988	6.790123	9.259259	24.074074	1.851852	1.851852	1.851852	3.08642	1.851852	1.851852	1.234568

Figure 8 Fredericton City Food Categories

To visualize the above data, I got the transposed data, rounded the percentage decimal value to two points and converted the data frame as following:

Borough	Count Percent
Middle Eastern Restaurant	1.23
Mexican Restaurant	1.85
Korean Restaurant	1.85
Italian Restaurant	1.85
Gastropub	1.85
Greek Restaurant	1.85
American Restaurant	3.09
Japanese Restaurant	3.09
Steakhouse	3.70
Bakery	4.32
Sushi Restaurant	4.94
Seafood Restaurant	5.56
Café	6.79
Chinese Restaurant	9.26

Figure 9 Transposed Data Frame

Above data frame is good to visualize. A horizontal bar chart would be a good choice to visualize it and show the percentage of eatery category in Fredericton.

Results:

A horizontal bar chart was created on above data frame to visualize it. Matplot library was imported to use the visualization tools. The horizontal bar graph on this data can be seen here:
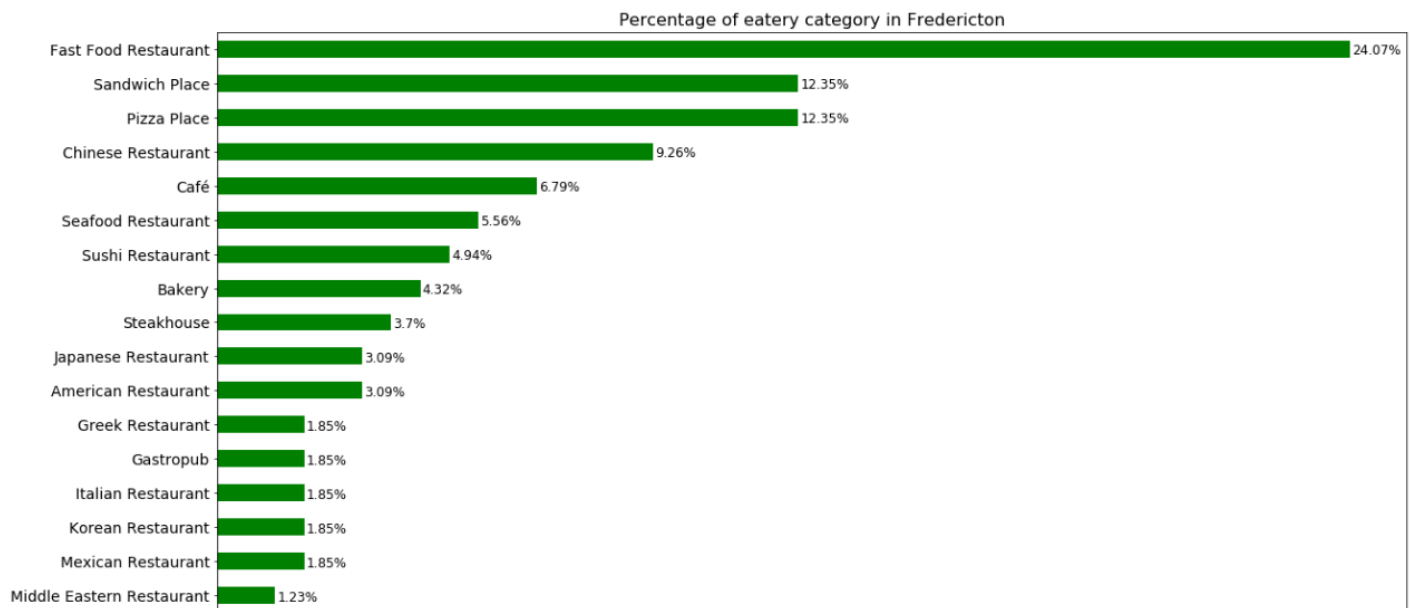

Figure 10 Bar Chart

The bar chart shows the percentage of eatery categories in Fredericton. The percentage of each bar is shown along with itself. The graph clearly shows that Fast Food Restaurants are leading here with around 25% share. Pizza and sandwich places are also running ahead from any kind of special cuisine in the city. Chinese Restaurants have their fair share in the food diversity of Fredericton. There are other cuisines as well in City that are enhancing the taste of the Fredericton.

Conclusion:

After considering the results, it can be said that fast food Restaurants are dominating the food market in Fredericton. The result shows the food diversity in the city and which is quite diverse by seeing the city's population. In addition, it cannot be denied that increasing population of the city will boost the demand of more cuisines in the city. Hence, it will bring more opportunities for the vendors who are planning to invest in food business.

Moreover, if we see the neighborhoods of Fredericton, the north Fredericton does not have the restaurant category in the top three most common places, yet construction is on top; it further grabs the attention of the stakeholders.

The scope of this analysis can be extended further by taking the data of community dominating locations, population figures and communities interests, to know the exact demand of the city; which can be a deciding factor for the vendors to make their decision.

References:

- **[1]** https://en.wikipedia.org/wiki/Fredericton
- **[2]** https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_E
- **[3]** https://developer.foursquare.com/