Finding best place for farmers market

Arpit Vasava

29/04/2019

1. Introduction

1.1 Background

New York, Which is considered as one of the most popular state in USA. New York is also very famous among tourist with 251 million visitors in year 2014.Visitors spent $117.5 billion in the state for food and other activities. Aside from tourist, New York is also a hub of giant offices. With this type of popularity it is hard to find food and place to sell. i will use data science tactics to find better and profitable place to sell farming vegies. We will go through each step of this project and address them separately. For this week I just describe the initial data preparation and future steps to start the battle of neighbourhoods in New York.

1.2 Problem

As discussed earlier with lots of population and skyscrapers it is hard to find places for selling. Indeed, population and other factors affects the selling market places.

2. Data acquisition and cleaning

2.1 Data sources

'https://query.data.world/s/c2nm5gs6r2exf2qt2hpneqxax7oxhe'. Is the link I will use data from. This data set consist of 11 columns which are FacilityName , Service Category , Service Type , Address , Address 2 , Borough , ZipCode , Latitude , Longitude , AdditionalInfo , StartDate. These columns consist of all information about selling place, facility information, etc.

2.2 Data cleaning and feature selection

Data downloaded or scraped from multiple sources were combined into one table. There were a lot of missing values from earlier seasons, because of lack of record

keeping. I decided to only use data which are available to me. There are lots of problem with data set one of which is unnecessary columns.

Second, multiple entries existed for multiple columns for example service category. This cause their seasonal data to represent multiple samples with incomplete data. I wrote script to unique data and discarded partial rows.

Another problem is null value. I add Nan string in place of null values.

The new, final and clean dataframe is as follows.

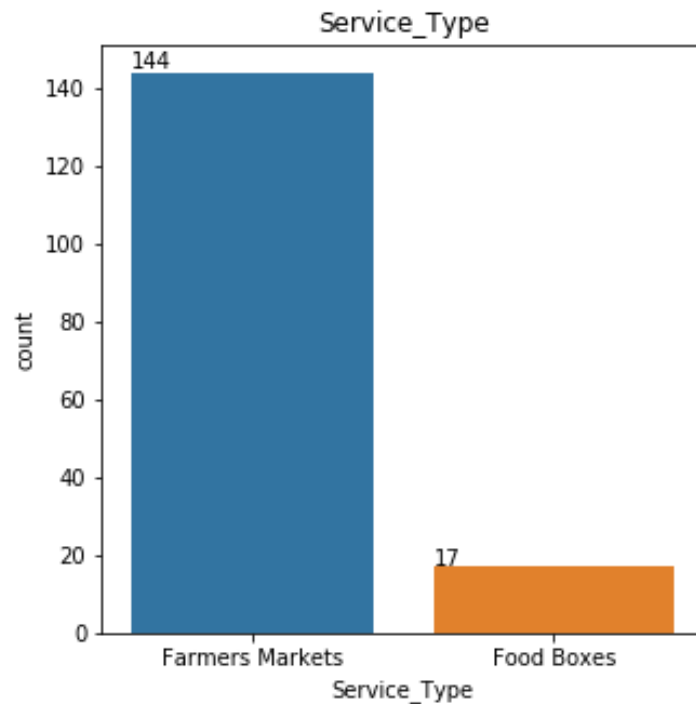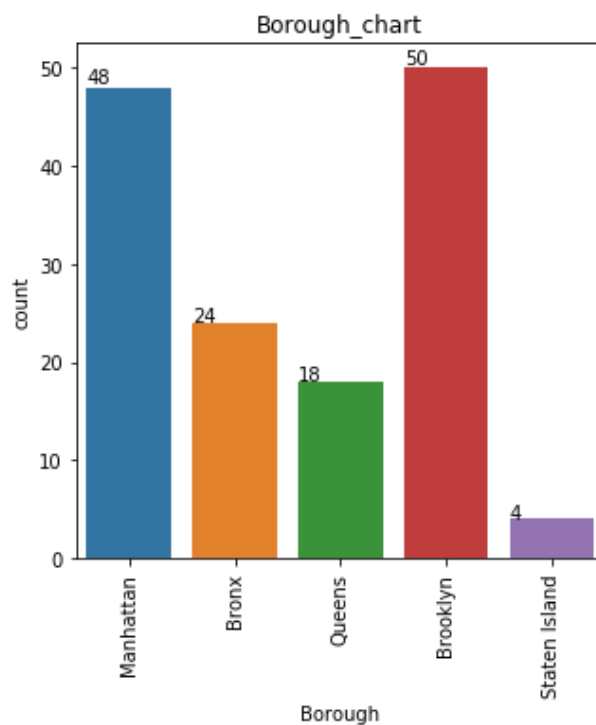|	FacilityName	Service Category	Service Type	Address	Address 2	Borough	ZipCode	Latitude	Longitude	AdditionalInfo	StartDate
0	Inwood Park Greenmarket	Farmers Markets and Food Boxes	Farmers Markets	Isham St bet Seaman & Cooper	NaN	Manhattan	10034	40.869009	-73.920320	Open year-round	NaN
1	82nd Street Greenmarket	Farmers Markets and Food Boxes	Farmers Markets	82nd St bet 1st & York Aves	NaN	Manhattan	10028	40.773448	-73.948954	Open year-round	NaN
2	1 Centre Street	Farmers Markets and Food Boxes	Food Boxes	1 Centre Street	South Building, 9th Floor	Manhattan	11101	40.713028	-74.003753	Open year-round	NaN

Figure *1*

3. Exploratory Data analysis

3.1 Relationship between different variables

For my project it is necessary to compare different variables from the data frame. Various comparison is given below:

1) Comparison between farmers market and food boxes. It it clear from the figure that Farmers market has way higher values than food boxes.
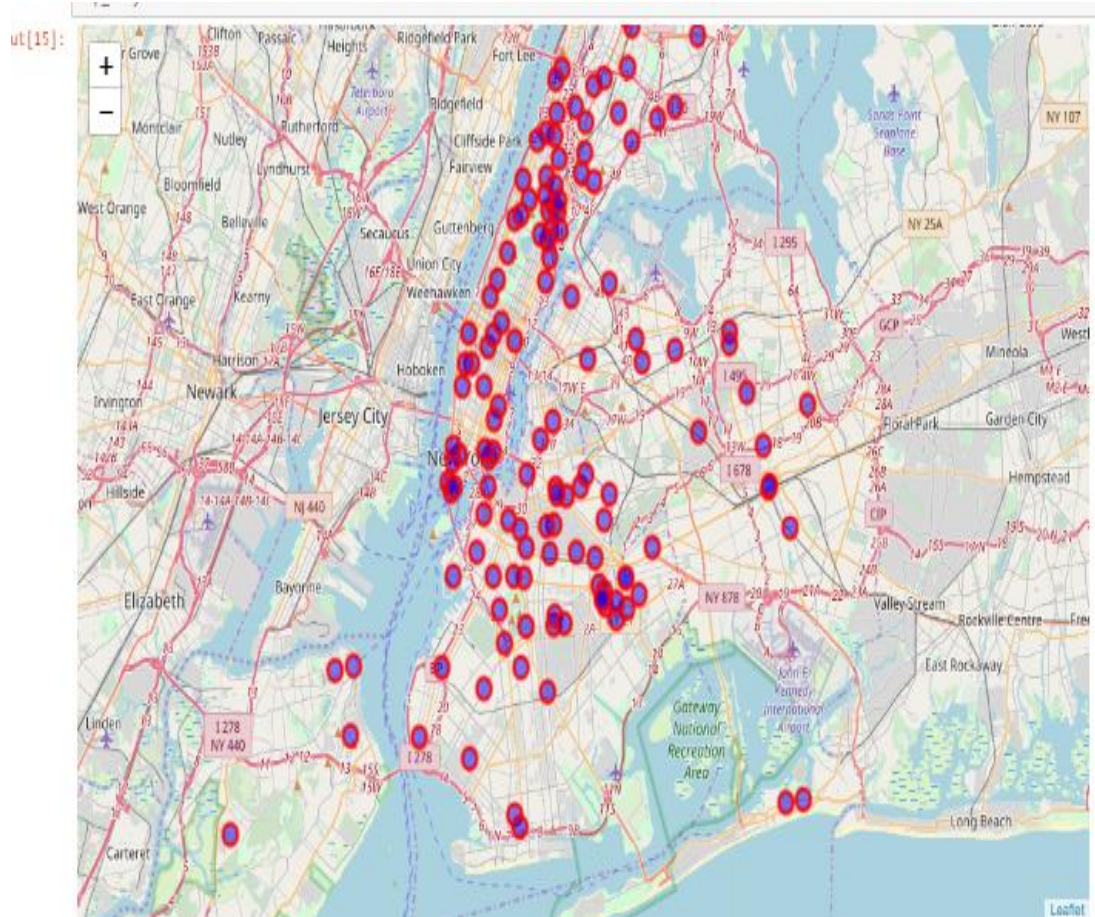
Service_Type

2) Second comparison I did between 4 city areas and their hit counts. From the given figure it is clear that Bronx, Queens has medium hit counts. Furthermore, Staten island has lowest hit score and is least preferable spot for farmers market. It is not very popular. In addition Manhattan and Brooklyn has highest hit rates and most preferable for farmers market.

Borough_chart

4. Conclusive Map

Finally I created a map for newyork city in manhatten area. This area of map shows the preferable places for farmers market.

This map is the final evaluation of my project.

5. Conclusion

New York is a big city for everything, it is very difficult to find places for your comfortness as individual and selling place as a wandor. And if you are a farmer you have to think about various factor like population, and buildings etc. From the above analysis it is clear that for farmers market Manhattan and Brooklyn are the most preferable places. If you want to be successful in farmers selling you have to place your stall in one of these cities.