



# THE BATTLE OF THE NEIGHBORHOODS



# Contents

|   |   |
|---|---|
| 1. Introduction & Business Problem .....  | 2 |
| Locations for New Car Dealer with Showroom in New York.....   | 2 |
| 2. Criteria.....  | 2 |
| Data .....  | 2 |
| 3.Methodology.....  | 3 |
| Business Understanding .....  | 3 |
| Analytic Approach .....   | 3 |
| Exploratory Data Analysis.....  | 3 |
| 4.RESULTS.....  | 4 |
| Neighborhood K-Means clustering based on Frequency of each venue category of venue<br>category..... | 5 |
| 5.DISCUSSION.....   | 6 |
| 6.CONCLUSION:.....  | 6 |

# 1. Introduction & Business Problem

## Locations for New Car Dealer with Showroom in New York

The City of New York, is the multicultural, heavily populated and is the financial capital of USA. It is a great city for business opportunities and business friendly environment. The city stands out for its major centers of banking, finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, among others.

All opportunities and diversity also mean a market is highly competitive with higher costs of business, higher land value and higher salaries. Thus, any new business venture or expansion needs to be analyzed carefully. The insights derived from analysis will give good understanding of the business environment which help in strategically targeting the market and the best location for a new Showroom for a multi-brand of green car. The offer will be more than just a Toyota Prius hybrid. the show room will offer an entries stock of Mercedes-Benz, BMW, Porsche, and other high-end marques include hybrids, diesels, and unexpectedly high gas mileage.

## 2. Criteria

Qualitative data collected from previous studies, suggests that the best locations to open New Car Dealer selling green cars may not only be where other dealers are located. This data strongly suggests that the best places are in fact areas that are near Vegans Restaurant, Cafés and Wine Bars. Social, healthy people with highly concern about efficiency and environment frequent this place often, so opening new showroom near them is a smart decision.

The analysis and recommendations for new showroom locations will focus on general neighborhood with these establishments, not on specific store addresses. Narrowing down the best district options derived from analysis allows for either further research to be conducted, advising agents of the chosen district, or on the ground searching for specific sites by the company's personnel.

## Data

We will be using the below datasets for analyzing New York city.

**Data 1:** Neighborhood has a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the the latitude and logitude coordinates of each neighborhood.

This dataset exists for free on the web. Link to the dataset is: [https://geo.nyu.edu/catalog/nyu\\_2451\\_34572](https://geo.nyu.edu/catalog/nyu_2451_34572)

**Data 2:** New York city geographical coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each neighborhood. We will use the Foursquare API to explore neighborhoods in New York City and restaurant, cafe and wine bars locations inside each neighborhood to select the one to propose to the company.

### 3. Methodology

#### Business Understanding

The goal is provide recommendations for new showroom locations will focus on general neighborhood with these establishments, not on specific store addresses. Narrowing down the best district options derived from analysis allows for either further research to be conducted, advising agents of the chosen district, or on the ground searching for specific sites by the company's personnel.

#### Analytic Approach

New York city neighborhood has a total of 5 boroughs and 306 neighborhoods. In this project first part is clustering of Manhattan. This is done because of the following Exploratory data analysis.

#### Exploratory Data Analysis

##### Data 1- New York city Geographical Coordinates Data

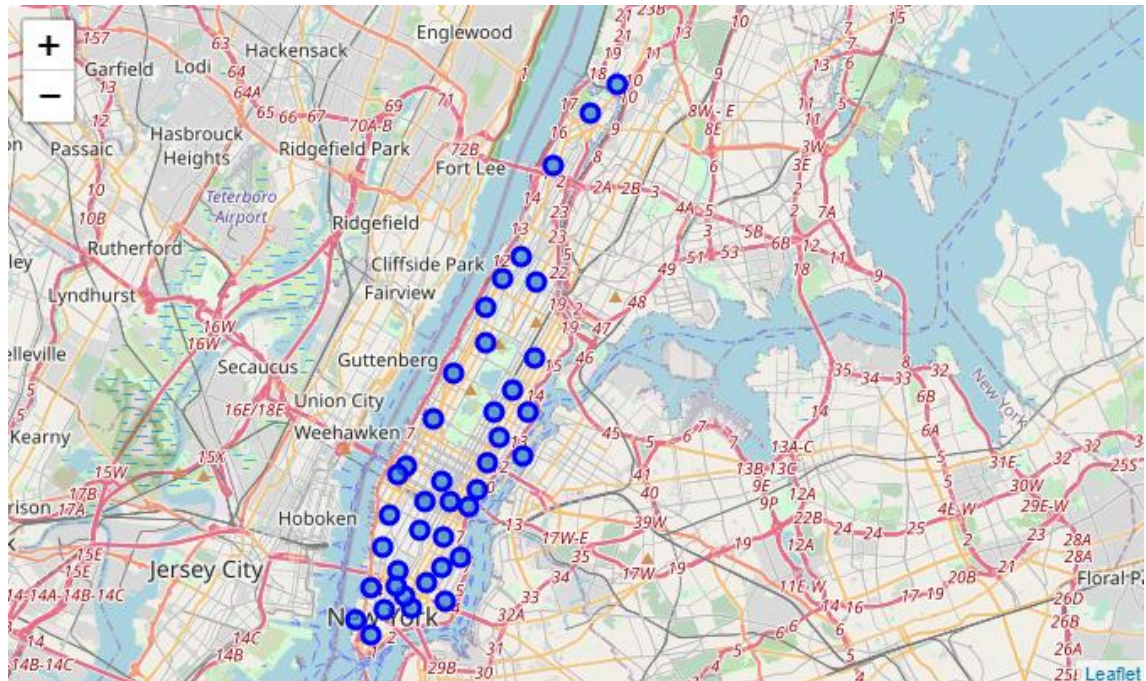
1. In this we load the data and explore data from newyork\_data.json file.
2. Transform the data of nested python dictionaries into a pandas data frame.
3. This data frame contains the geographical coordinates of New York city neighborhoods.
4. This data will used to get Venues data from Foursquare.
5. We used geopy and folium libraries to create a map of New York city with neighborhoods superimposed on top.

##### New York neighborhood visualization



##### Data 2- Focus on Manhattan

By filtering the data selecting Manhattan only we remain with a total of 39 neighborhoods in New York city.



**Data 3 – Venues in Manhattan**

#### Manhattan Venues:

|   | Neighborhood | Neighborhood Latitude | Neighborhood Longitude | Venue         | Venue Latitude | Venue Longitude | Venue Category |
|---|--------------|-----------------------|------------------------|---------------|----------------|-----------------|----------------|
| 0 | Marble Hill  | 40.876551             | -73.91066              | Arturo's      | 40.874412      | -73.910271      | Pizza Place    |
| 1 | Marble Hill  | 40.876551             | -73.91066              | Bikram Yoga   | 40.876844      | -73.906204      | Yoga Studio    |
| 2 | Marble Hill  | 40.876551             | -73.91066              | Tibbett Diner | 40.880404      | -73.908937      | Diner          |
| 3 | Marble Hill  | 40.876551             | -73.91066              | Starbucks     | 40.877531      | -73.905582      | Coffee Shop    |
| 4 | Marble Hill  | 40.876551             | -73.91066              | Dunkin'       | 40.877136      | -73.906666      | Donut Shop     |

## 4.RESULTS

From this venues data we filtered and used only the restaurant data for Manhattan clustering. As we focused only on restaurants with a specific cousin the data was filtered to analysis on venue category of Vegetarian / Vegan Restaurant, Wine Bar, Veterinarian, Whisky Bar.

#### Number of venues by neighborhood

|                 | Vegetarian / Vegan Restaurant | Wine Bar | Veterinarian | Whisky Bar | Total_Venues |
|-----------------|-------------------------------|----------|--------------|------------|--------------|
| Neighborhood    |                               |          |              |            |              |
| West Village    | 0                             | 7        | 0            | 1          | 7            |
| Upper West Side | 2                             | 3        | 0            | 0          | 5            |
| East Village    | 2                             | 3        | 0            | 0          | 5            |
| Inwood          | 0                             | 2        | 1            | 0          | 3            |
| Civic Center    | 1                             | 2        | 0            | 1          | 3            |



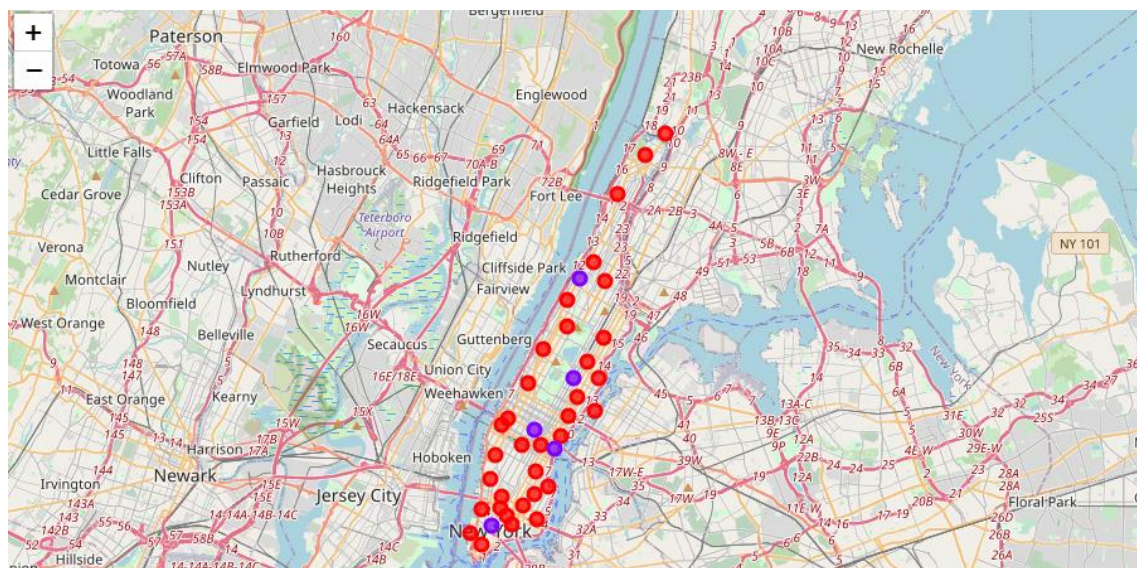
Neighborhood K-Means clustering based on Frequency of each venue category of venue category.

To cluster the neighborhoods into two clusters we used the K-Means clustering Algorithm. k-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. It uses iterative refinement approach.

**Cluster Labels represent the result of the K-means clustering**

|    | Cluster_Labels | Vegetarian / Vegan Restaurant | Wine Bar | Veterinarian | Whisky Bar | Total_Venues |
|----|----------------|-------------------------------|----------|--------------|------------|--------------|
| 36 | 1              | 0.024691                      | 0.037037 | 0.000000     | 0.000000   | 0.061728     |
| 15 | 1              | 0.000000                      | 0.035088 | 0.017544     | 0.000000   | 0.052632     |
| 8  | 1              | 0.020000                      | 0.030000 | 0.000000     | 0.000000   | 0.050000     |
| 32 | 1              | 0.000000                      | 0.040000 | 0.000000     | 0.013333   | 0.040000     |
| 5  | 1              | 0.010417                      | 0.020833 | 0.000000     | 0.010417   | 0.031250     |
| 31 | 0              | 0.020408                      | 0.010204 | 0.000000     | 0.000000   | 0.030612     |
| 29 | 0              | 0.012500                      | 0.012500 | 0.000000     | 0.000000   | 0.025000     |
| 26 | 0              | 0.012346                      | 0.012346 | 0.000000     | 0.000000   | 0.024691     |
| 10 | 0              | 0.020202                      | 0.000000 | 0.000000     | 0.000000   | 0.020202     |
| 3  | 0              | 0.010000                      | 0.010000 | 0.000000     | 0.000000   | 0.020000     |
| 6  | 0              | 0.000000                      | 0.020000 | 0.000000     | 0.000000   | 0.020000     |
| 18 | 0              | 0.010000                      | 0.010000 | 0.000000     | 0.000000   | 0.020000     |
| 34 | 0              | 0.000000                      | 0.020000 | 0.000000     | 0.000000   | 0.020000     |

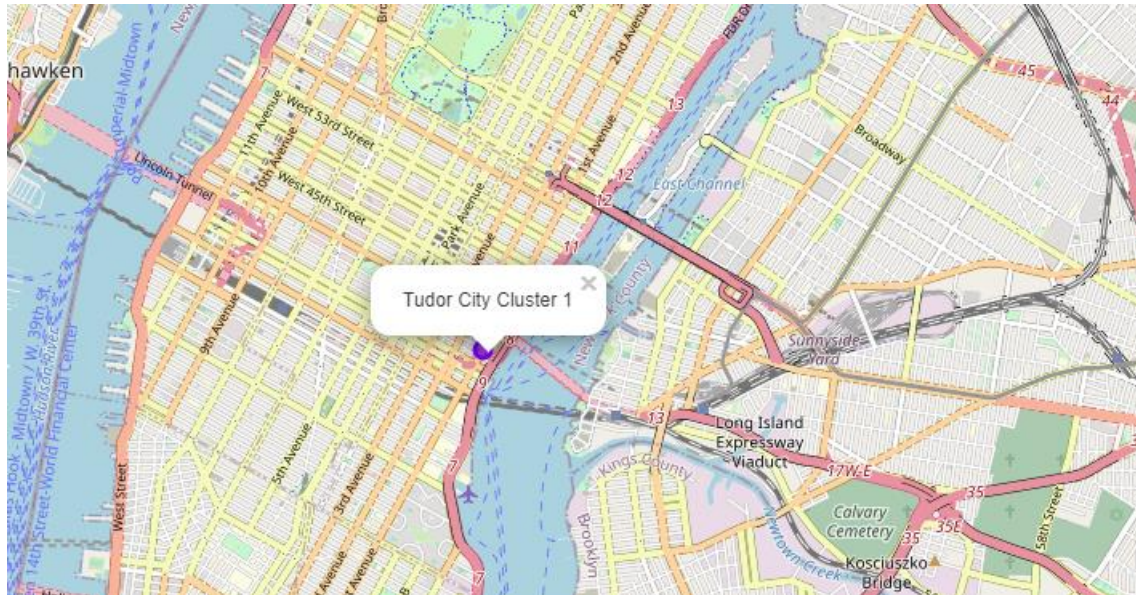
In the below Map Visualization, we can see the different types of clusters created by using K-Means Manhattan.



**Cluster 0:** The Total and Total Sum of cluster0 has smallest value. It shows that in this neighborhood the number of selected venues is lower.

**Cluster 1:** The Total and Total Sum of cluster1 has highest value. Number of restaurants are very high.

**The Neighborhood with the highest number of venues is Tudor City.**



## 5.DISCUSSION

1. The select restaurant in Manhattan represents a smaller number of the total venues making difficult the analysis.
2. There is scope to explore to open New Car Dealer selling green cars.

## 6.CONCLUSION:

This analysis is performed on limited data. Manhattan has high concentration of restaurant business. Very competitive market.

Using the success criteria that the best places to open New Car Dealer selling green cars are in fact areas that are near Vegans Restaurant, Cafés and Wine Bars. Social, healthy people with highly concern about efficiency and environment frequent this place often, so opening new showroom near them is a smart decision.

The recommendation is to select a location in **Tudor City**.