INTRODUCTION

This in the introductory section of the final assignment of IBM Applied Data Science Capstone course on Coursera. Here, the students are asked to be as creative as they want to implement Foursquare API and clustering algorithm to explore or compare neighbourhoods or cities of their choice. Based on these instructions, I have decided to compare the neighbourhoods of New York City, particularly the borough of Manhattan in order to find out which neighbourhood is the best to open a restaurant.

Business Problem

Opening a restaurant requires a lot of factors to be considered. The weather of the location, crime data, cost of infrastructure, tastes of the people, and competition, just to name a few. In addition to this, the different values and data for different neighbourhoods makes the whole act of choosing between neighbourhoods to start a business, quite daunting and hectic.

However, this project aims to be of use to you through an in-depth analysis and comparison of the neighbourhoods in Manhattan. While this project's main attention is on restaurants, the analysis can also be used for other purposes. (Such as the ones pointed out below)

Manhattan has been chosen amongst the other boroughs of NYC due to its high population density and commercial lifestyle. It is my intention to delve deep into the restaurant statistics of the neighbourhoods of Manhattan and by the end of this report, recommend a few neighbourhoods that would be best to open a restaurant in with sufficient explanation.

Who can benefit from this project?

- Those who are planning to migrate to Manhattan

- A New Yorker trying to change neighbourhoods to find a more happening area

- Those who would like to expand or open their businesses in Manhattan, NYC

- Individuals who already own a restaurant or are looking to open a new one in Manhattan.

DATA ACQUISITION

This project aims to compare the neighbourhoods of Manhattan based on the following data:
- Crime rate data of Manhattan

- Venue data of Manhattan

- Rent data of the selected neighbourhoods of Manhattan

- Weather data of New York City (NYC)

Weather data

The reason for analysing weather data is because of its important effect on human behaviour. For example, on a rainy day, one might crave a hot beverage. On a warm sunny day, one might prefer desert. In addition to all this, having sound knowledge of the weather data and patterns can also help restaurant owners as they can make additions which will make their menus more appetizing!

It is imperative that the owners have sound knowledge of past annual weather data. Doing so will help them make a rough estimate as to what kind of weather they can expect per month.

In order to carry out this analysis of the weather data, the annual weather data from 2010-2019 for NYC was collected from the website [Current results].

	High_F	Low_F	Month	Days over 70F	Days over 90F	Days_below_32F	Days_below_20F
0	39	27	January	0	0	20	8
1	42	29	February	0	0	18	4
2	50	35	March	1	0	10	1
3	63	46	April	6	0	1	0
4	73	56	May	16	1	0	0

Table 1: Weather Data

High_F: The average of the highest temperatures in degree Fahrenheit per month from 2010-2019

Low_F: The average of the lowest temperatures in degree Fahrenheit per month from 2010-2019

Days over 70F and 90F: The total number of hot days in NYC per month. Here days are considered to be hot when the temperatures are above 70-degree Fahrenheit and 90-degree Fahrenheit

Days below 32F and 20F: The total number of cold days in NYC per month. Here days are considered to be cold when the temperatures are below 32-degree Fahrenheit and 20-degree Fahrenheit

Location and Venue data

Data acquisition for Manhattan's neighbourhoods was done through the NYCloc csv file, which contains all the data on the neighbourhoods of NYC. This file was obtained from the week 3 lab in the IBM 'Applied Data Science Capstone project' course. This csv file was further segmented into 5 different dataframes, with one dataframe per each borough. The Manhattan borough is of significant importance to us as all the analysis has been carried out on it.

The Manhattan dataframe was then further polished and all the venues such as restaurants, parks, pizza places, theatres, and malls were added to it with the help of the **Foursquare API.**

	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

Table 2: Manhattan Venue data

As you can see, the above table contains the name of the neighbourhood along with its latitude and longitude values, as well as all the neighbourhood's venues with their names, geographical coordinates and their categories.

Rent data and Crime data

Rent data and Crime data were utilized to further carry out the analysis midway through the project. Based on certain discoveries and results (which you will learn about later in this report), a select few neighbourhoods from Manhattan were selected. Upon their selection, their commercial rent data was collected from Property Shark.

For the rent data, the main focus was on the highest and lowest rent prices per square foot per year (USD) for each neighbourhood.

	Neighborhood	Low_Rent_sqft	High_Rent_sqft
0	Financial District	40	50
1	Tudor City+surroundings	68	78
2	Soho	75	90
3	Flatiron	60	77
4	Washinton Heights	35	55
5	Upper East Side	65	75
6	Sutton Place	70	80
7	Tribeca	50	60

Table 3: Rent data

Here the **Low_Rent_sqft** and **High_Rent_sqft** columns represent the lowest and highest rent prices in USD per square foot per year for each neighbourhood respectively.

Crime data for the neighbourhoods mentioned in Table 3 was also collected in order to determine the safety of the neighbourhoods. In order to do this, the data on the total seven major felony offenses was collected for each precinct for the years from 2014 to 2019.

	Neighbourhoods	CRIME	2014	2015	2016	2017	2018	2019
0	Financial District, Tribeca, Soho	TOTAL SEVEN MAJOR FELONY OFFENSES	1208	1446	1395	1337	1356	1386
1	Flatiron	TOTAL SEVEN MAJOR FELONY OFFENSES	2061	2056	1879	1694	1954	2014
2	Tudor City, Sutton Place	TOTAL SEVEN MAJOR FELONY OFFENSES	909	986	1000	1023	1021	1116
3	Upper East Side	TOTAL SEVEN MAJOR FELONY OFFENSES	1913	1904	1959	1972	2063	2273
4	Washington Heights	TOTAL SEVEN MAJOR FELONY OFFENSES	1151	1315	1323	1382	1377	1278

Table 4: Crime data

In the above table, certain neighbourhoods are grouped together. This is because these neighbourhoods belong to the same precinct. It is important to note that the crime data collected has been done so precinct-wise. Hence, the numbers above may be overestimated as a precinct contains multiple neighbourhoods. However, knowing the total number of crimes that have been committed in a neighbourhood's precinct helps in gauging the safety of the neighbourhood locality, which is why this data has been considered.

(This data has been collected from the official New York Police Department website.

It is important to note that the total seven major felony offenses consist of **Grand Larceny, Burglary, Murder, Rape, Assault, Robbery, Grand Larceny of a vehicle.)**

METHODOLOGY

In this section the methods used to analyse the data will be discussed. Since multiple data points were used for analysis, there will be separate sub-sections for separate data points. While this section will be used to elaborate on the methods of analysis being used, the results will only be discussed in the next section.

Weather Data Analysis

The first analysis carried out in this project was the weather analysis. This analysis has been made under the assumption that Manhattan's weather will be more or less similar to the overall New York City weather.

Using the data show in **Table 1,** the temperature fluctuations throughout the year were visualized with the help of a line graph. This line graph was a plot showing the highest and lowest temperatures recorded per month. (See graph in results section)

Another important key point was to get the monthly comparison of the total number of hot and cold days. In order to visualize this, two bar graphs were plotted, one for the number of hot days and one for the number of cold days.

All of the above plotting was carried out through the Matplotlib package in Python.

Location Data Analysis

In order to understand how big a part of New York Manhattan is, two maps were plotted using the Folium package. The first map was a map of all the neighbourhoods of New York City and the second one was a map of all the neighbourhoods of Manhattan.

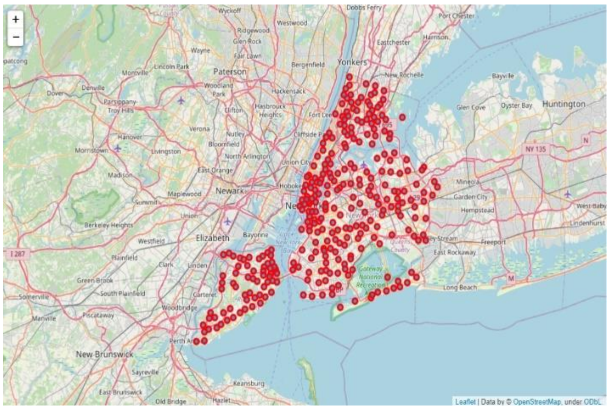


Fig 1: NYC Map

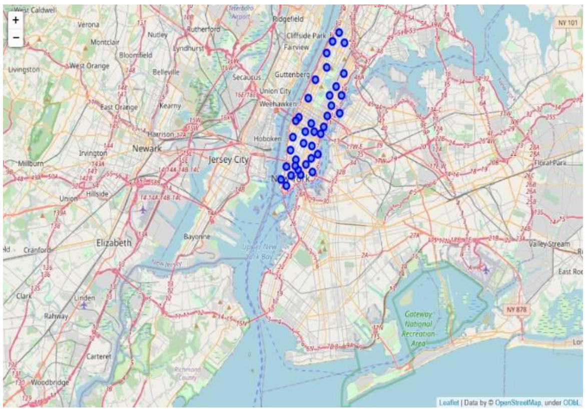


Fig 2: Manhattan Map

As evidenced by the above two maps, we can see how much of New York City comprises of Manhattan and its neighbourhoods.

Then the main dataframe containing all neighbourhood details was split into 5 different dataframes according to their boroughs. The Manhattan borough dataframe is the one that has been used in this project.

Using the Manhattan borough dataframe (Let's call it M1 in order to avoid unnecessary repetition) and my Foursquare credentials, a connection was made to the Foursquare database that allowed me to download all the venue details such as venue name, geographical coordinates, neighbourhood location, and venue category, for every single venue, and store it in a new dataframe (M2).

For easier analysis, all venues that weren't restaurants were dropped from the M2 dataframe. This resulted in a dataframe that had every single restaurant in Manhattan. Let this dataframe be named Final M.

In-depth Analysis on the Restaurants

From Final M, the most common restaurants were determined in Manhattan.

	Venue_Category	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude
164	Italian Restaurant	100	100	100	100	100	100
4	American Restaurant	69	69	69	69	69	69
194	Mexican Restaurant	54	54	54	54	54	54
59	Chinese Restaurant	45	45	45	45	45	45
166	Japanese Restaurant	43	43	43	43	43	43
290	Sushi Restaurant	42	42	42	42	42	42
121	French Restaurant	37	37	37	37	37	37
263	Seafood Restaurant	36	36	36	36	36	36
303	Thai Restaurant	31	31	31	31	31	31
191	Mediterranean Restaurant	30	30	30	30	30	30
158	Indian Restaurant	22	22	22	22	22	22

Table 5: Most common restaurants

The above table shows that there is a total of 100 Italian restaurants in the whole of Manhattan, followed by 69 American restaurants and so on. (Please ignore the latitude and longitude columns as their values are just reflecting the total number of restaurants as well).

This leads us to the inference that Italian restaurants are generally well received by the residents of Manhattan. (This inference is based on the assumption that the reason for the high number of certain restaurants is because the cuisines of these restaurants are well received by the residents of the neighbourhoods. In a hospitality business such as this, growth is only possible if the products being offered are well received and liked by the public.)

However, we cannot conclude that Italian restaurants will find success in all the neighbourhoods of Manhattan equally. It is also possible that one or two neighbourhoods may have an overwhelmingly high number of Italian restaurants and that Italian food is only loved in these neighbourhoods, whereas the greater part of Manhattan may prefer American or Chinese food more.

In order to avoid such problems, the neighbourhoods were ranked according to the number of restaurants they had. By doing so, a clear picture was obtained on the spread of restaurants all over Manhattan. (Tables in next page)

Neighborhood	
Chinatown	39
East Village	36
Little Italy	33
Greenwich Village	32
Noho	32
Upper West Side	31
Turtle Bay	30
Yorkville	29
Midtown South	29
West Village	28
Civic Center	26
Lenox Hill	26
Murray Hill	26
Tudor City	25
Soho	24
Flatiron	23
Washington Heights	23
Upper East Side	23
Sutton Place	21
Financial District	20
Tribeca	20
Midtown	19
Carnegie Hill	19
Gramercy	18
Manhattanville	18
Inwood	18
Clinton	17
Hamilton Heights	17
Lincoln Square	17
East Harlem	15
Central Harlem	15
Chelsea	14
Lower East Side	13
Manhattan Valley	13
Morningside Heights	10
Hudson Yards	10
Battery Park City	3
Roosevelt Island	2
Marble Hill	2

Table 6: Neighbourhood Rankings

The above table shows that Chinatown has the greatest number of restaurants (39) and Marble Hill (2) has the least. This data is very important to this project as it is the driving force behind the selection of the neighbourhoods.

Going further, eight of the above neighbourhoods were selected in order to narrow down the best potential neighbourhoods.

The eight neighbourhoods selected were

Tudor City, Soho, Flatiron, Washington Heights, Upper East Side, Sutton Place, Financial District, Tribeca

These neighbourhoods were selected because they lie in the 20 to 25 restaurants per neighbourhood range. This range was selected based on the inference that these neighbourhoods have high potential for growth while having a decent restaurant ecosystem and network in place already.

The top neighbourhoods such as Chinatown and bottom neighbourhoods such as Marble Hill were avoided primarily due to the reasons mentioned below

- Suppose you wanted to open a restaurant in Chinatown. Opening a restaurant in such a saturated market would lead to heavy initial competition with the already other well-established restaurants. A saturated market also means that failure to meet the quality standards of the crowd would definitely result in heavy losses. While succeeding through such a move is definitely possible, it creates a very high stakes environment for the restaurant and also requires top class personnel and staff. Such personnel may be hard to acquire right at the start of a new business.

- On the other hand, opening a restaurant in any of the 8 selected neighbourhoods reduces the competition significantly. While the restaurant will experience a comparatively slower start by opting this method, there is a possibility that the restaurant may grow and the dynamic may shift towards these neighbourhoods in the near future. This decision will also give the restaurant ample breathing room and potential opportunities to expand its influence.

- A risky move would be opening a restaurant in neighbourhoods with just one or two restaurants. There may be a lot of reasons as to why these neighbourhoods have such low number of restaurants. It could be the population of these neighbourhoods, or the location itself. Until these reasons are clear, it's advisable to stay away from these neighbourhoods. There is definitely a possibility for success, as the room for growth is phenomenal, but this banks on the restaurant, its uniqueness, and the quality of food that it brings.

Rent and Crime data analysis

After the selection of the eight neighbourhoods, the rent and crime data for these neighbourhoods were collected. Then using the data shown in tables 3 and 4, bar graphs were plotted (see results section) that showed very clearly, the comparison of the prices as well as the crime rate in these neighbourhoods.

K Means Clustering

Two clusters were formed during this analysis. One cluster that clustered the above 8 neighbourhoods based on their most popular cuisines. The second cluster was a cluster of all the neighbourhoods in Manhattan based on the popularity of their cuisines. By doing so we were able to see how our eight neighbourhoods fared in comparison to the greater part of Manhattan. It gave us a clear visual of whether there were any predominant clusters that the 8 neighbourhoods were a part of.

What this means is that a neighbourhood with Italian restaurants as its most popular venue, and Chinese restaurants as its second most popular venue, will be in the same cluster as a neighbourhood with vice versa popular venues.

Regarding the procedure for carrying out K Means Clustering: All categorical variables were dropped from the Final M dataframe. This was done as K Means cannot be implemented with categorical variables in a dataframe. Then each restaurant was assigned a dummy numerical value based on its total popularity or common-ness per neighbourhood. This modified dataframe was further polished to produce a dataframe of all the neighbourhoods with their top ten most common restaurants. (see table next page)

	Neighborhood	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	9th Most Common Venue	10th Most Common Venue
0	Financial District	American Restaurant	Falafel Restaurant	Italian Restaurant	Mexican Restaurant	Japanese Restaurant	Mediterranean Restaurant	Seafood Restaurant	Restaurant	French Restaurant	New American Restaurant
1	Flatiron	Italian Restaurant	Japanese Restaurant	Mediterranean Restaurant	American Restaurant	Vegetarian / Vegan Restaurant	New American Restaurant	Cuban Restaurant	Fast Food Restaurant	Indian Restaurant	Kebab Restaurant
2	Soho	Italian Restaurant	Mediterranean Restaurant	French Restaurant	Sushi Restaurant	Australian Restaurant	Vegetarian / Vegan Restaurant	Falafel Restaurant	Egyptian Restaurant	Mexican Restaurant	American Restaurant
3	Sutton Place	Italian Restaurant	Vegetarian / Vegan Restaurant	Mexican Restaurant	American Restaurant	Cambodian Restaurant	French Restaurant	Latin American Restaurant	Lebanese Restaurant	Greek Restaurant	Persian Restaurant
4	Tribeca	Italian Restaurant	American Restaurant	Greek Restaurant	French Restaurant	Modern European Restaurant	Sushi Restaurant	Indian Restaurant	Korean Restaurant	Seafood Restaurant	Chinese Restaurant
5	Tudor City	Mexican Restaurant	Sushi Restaurant	Vietnamese Restaurant	Thai Restaurant	Asian Restaurant	Greek Restaurant	Seafood Restaurant	French Restaurant	Hawaiian Restaurant	Japanese Restaurant
6	Upper East Side	Italian Restaurant	American Restaurant	French Restaurant	Sushi Restaurant	Mexican Restaurant	Chinese Restaurant	Vegetarian / Vegan Restaurant	Latin American Restaurant	Mediterranean Restaurant	North Indian Restaurant
7	Washington Heights	Chinese Restaurant	New American Restaurant	Tapas Restaurant	Spanish Restaurant	Latin American Restaurant	Mexican Restaurant	Ramen Restaurant	Indian Restaurant	Italian Restaurant	Caribbean Restaurant

Table 7: The selected 8

	Neighborhood	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	9th Most Common Venue	10th Most Common Venue
0	Battery Park City	Chinese Restaurant	Mediterranean Restaurant	Mexican Restaurant	Vietnamese Restaurant	German Restaurant	English Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	Filipino Restaurant
1	Carnegie Hill	Japanese Restaurant	Italian Restaurant	French Restaurant	Vietnamese Restaurant	Thai Restaurant	American Restaurant	Indian Restaurant	Kosher Restaurant	Mexican Restaurant	Fast Food Restaurant
2	Central Harlem	African Restaurant	Chinese Restaurant	American Restaurant	Seafood Restaurant	French Restaurant	Southern / Soul Food Restaurant	Ethiopian Restaurant	Caribbean Restaurant	Vietnamese Restaurant	Falafel Restaurant
3	Chelsea	American Restaurant	Italian Restaurant	Seafood Restaurant	Restaurant	Sushi Restaurant	Mediterranean Restaurant	Middle Eastern Restaurant	Ramen Restaurant	Chinese Restaurant	Japanese Restaurant
4	Chinatown	Chinese Restaurant	American Restaurant	Vietnamese Restaurant	Shanghai Restaurant	Vegetarian / Vegan Restaurant	Greek Restaurant	Malay Restaurant	Mexican Restaurant	Dumpling Restaurant	Dim Sum Restaurant
5	Civic Center	French Restaurant	American Restaurant	Sushi Restaurant	Vietnamese Restaurant	Falafel Restaurant	Italian Restaurant	Restaurant	Molecular Gastronomy Restaurant	Modern European Restaurant	Cuban Restaurant

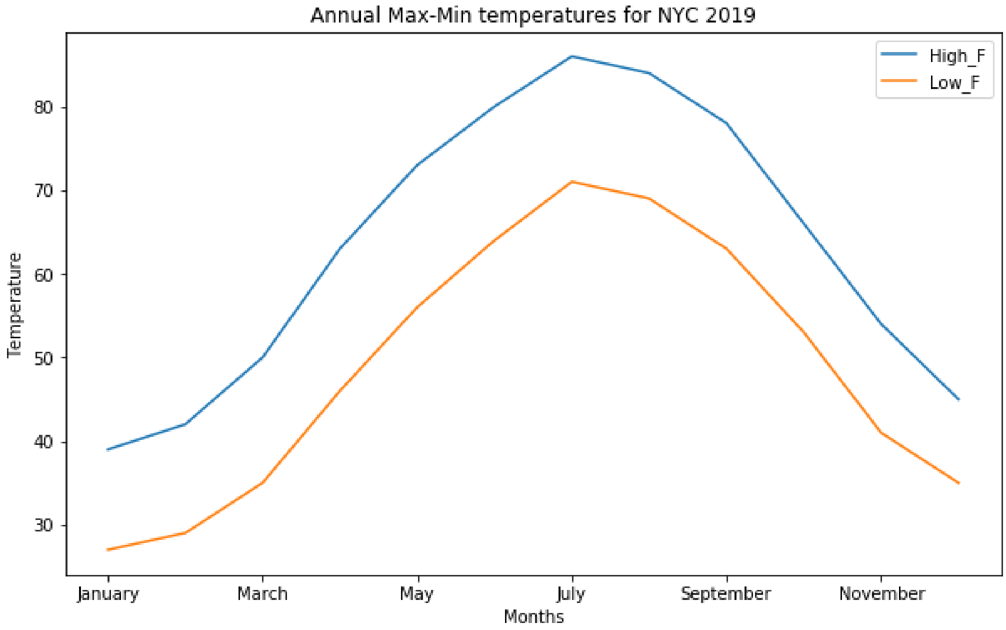
Table 8

Table 7 and 8 were then used for creating the two clusters. How the clusters look and the inferences that can be made from them will be discussed in the next section.

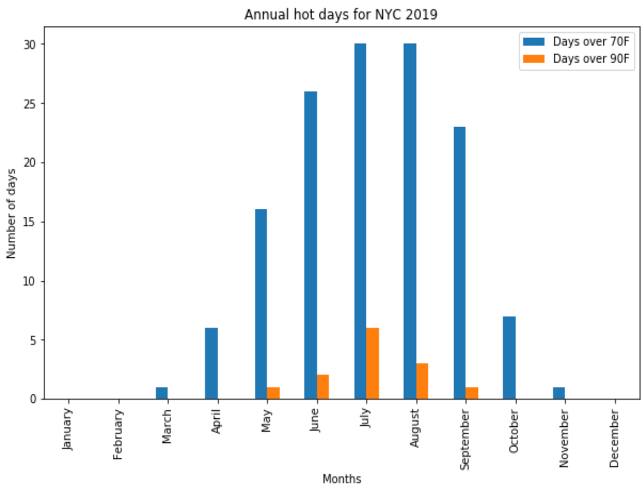
RESULTS AND CONCLUSION

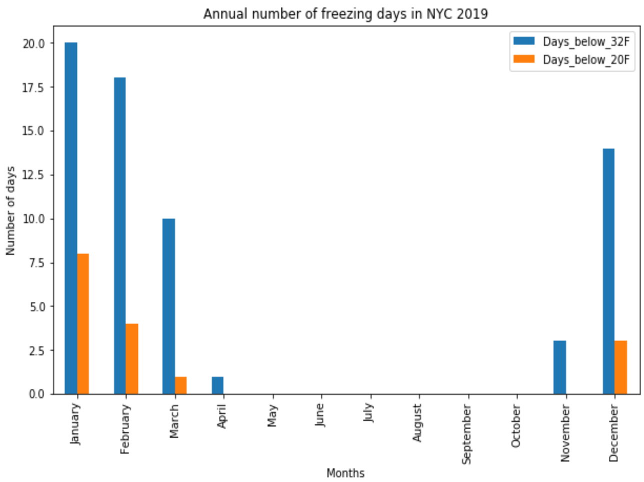
In this section, all the plots, clusters, and final tables will be displayed.

Weather Data Analysis Results

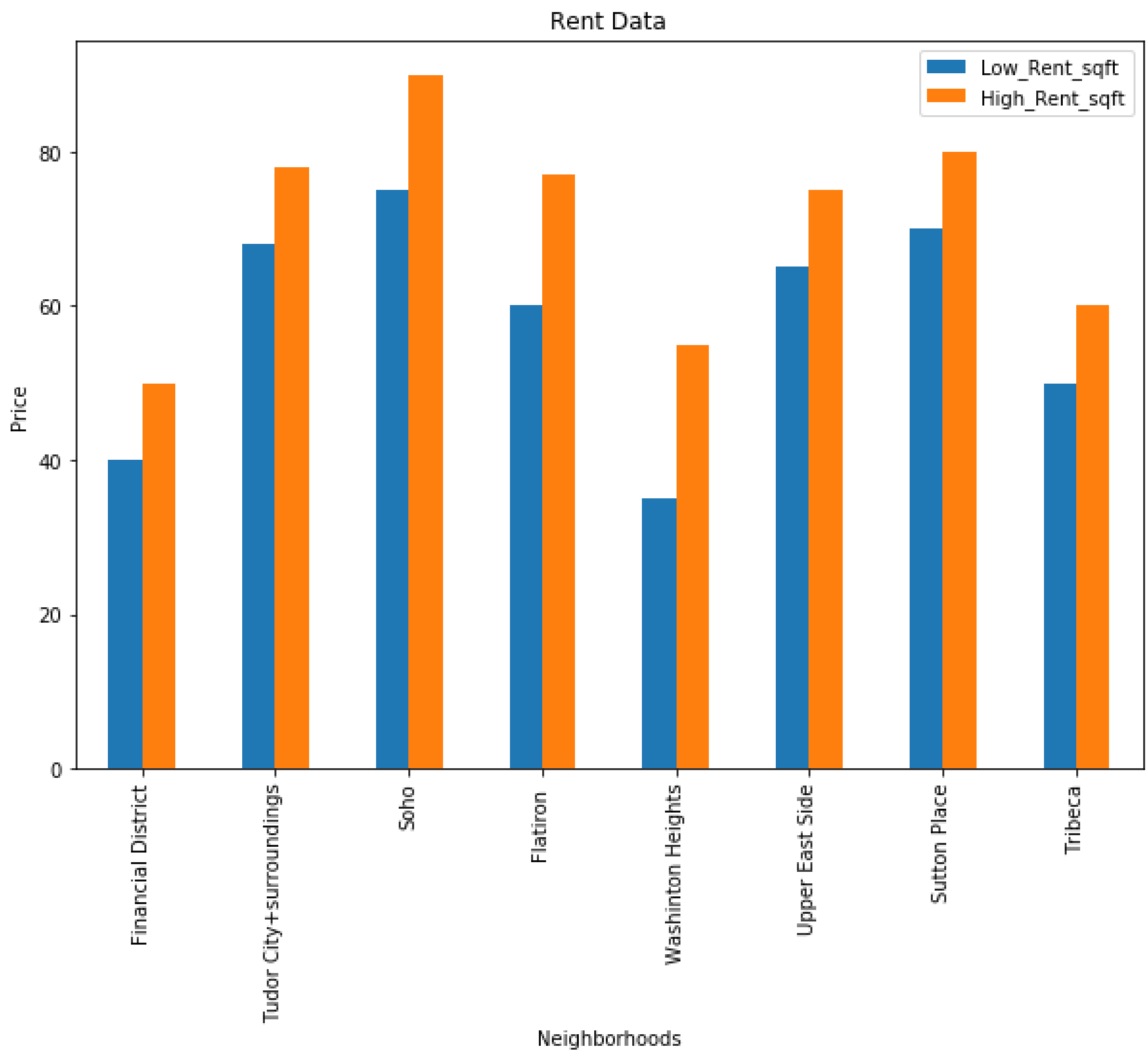


The graph displays the rise of both average Hot and Cold temperatures throughout the year. We can conclude that NYC experiences its hottest days between May to September, and its coldest days from November to March



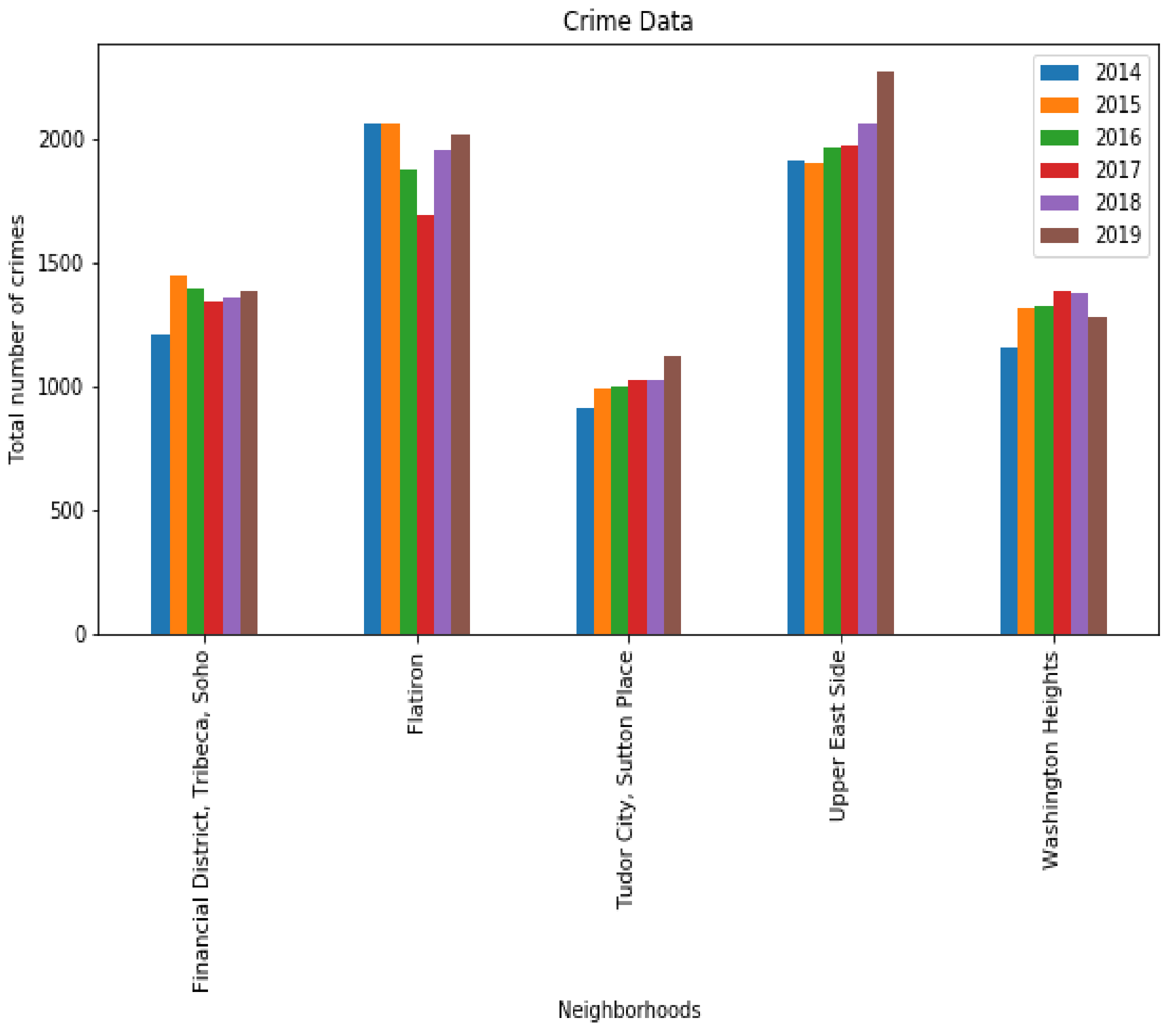


Rent Data and Crime Data Results



The above graph was plotted from the data given in Table 3.

From this graph we can infer that **Financial District, Washington Heights,** and **Tribeca** have some of the best and cheapest commercial rent rates.



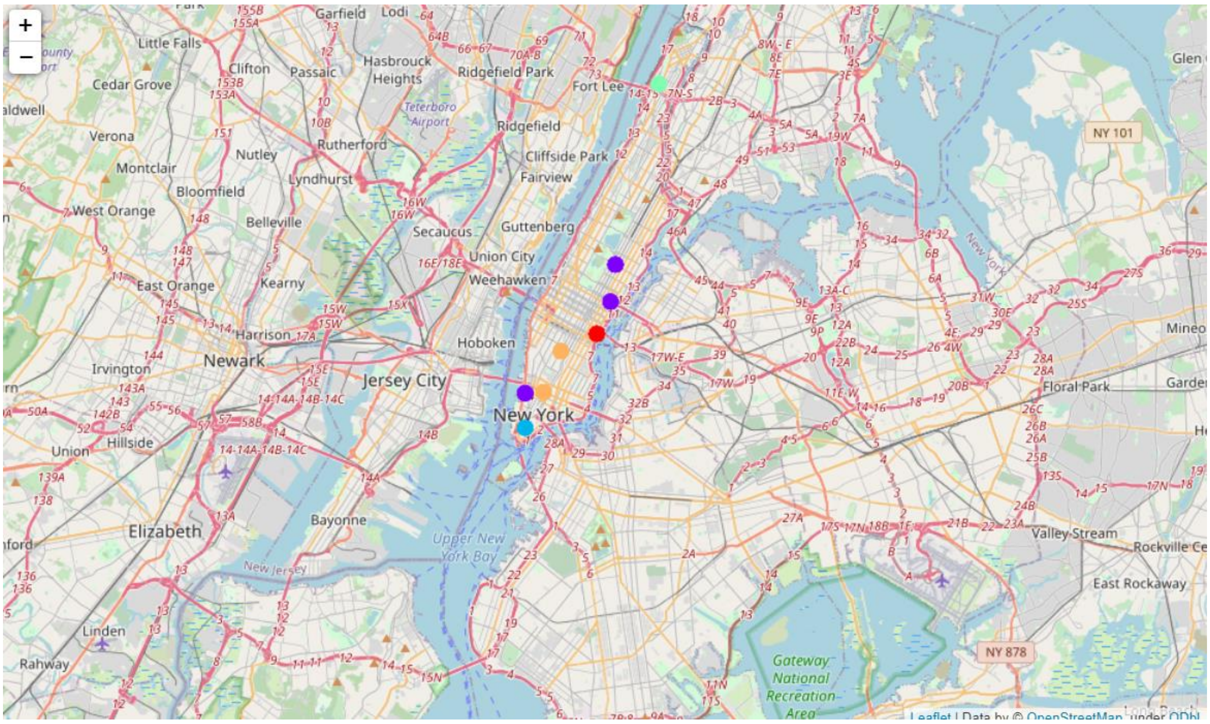
This graph was plotted from the data given in Table 4.

With the help of this graph we can see that **Tudor City, Sutton Place, Washington Heights**, and **Financial District** have low crime rates. It is also encouraging to see that Washington Heights has experienced a drop in the total crimes committed in 2019.

Upper East side and Flatiron have both experienced an increase in the number of crimes committed.

Considering only the above rent and crime data we can see that both Washington Heights and Financial District are good options to start a restaurant in. However, it is also important to see the clusters these districts fall into.

K MEANS CLUSTERING RESULTS



Clustering the 8 selected neighbourhoods.

In the above map the clusters are as follows

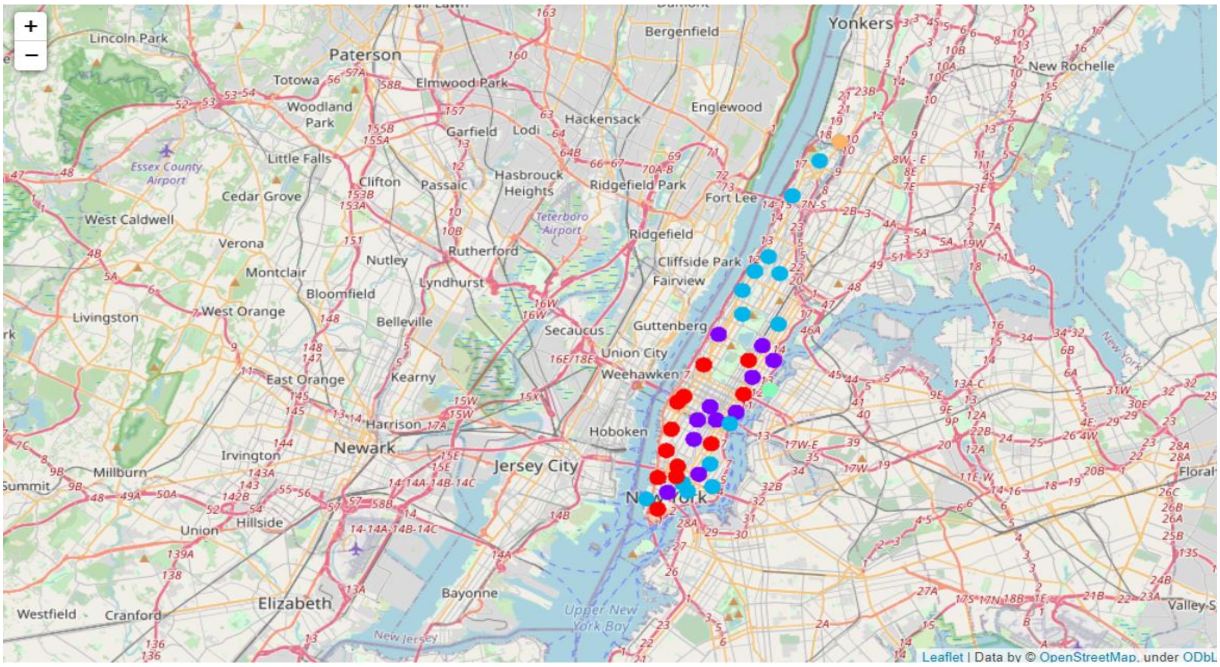
Cluster 1: Flatiron, Soho, Sutton Place, Upper East Side – Colour Purple

Cluster 2: Tudor City – Colour Red

Cluster 3: Financial District – Colour Blue

Cluster 4: Tribeca – Colour Orange

Cluster 0: Washington Heights – Colour Green



Clustering all the neighbourhoods of Manhattan

Cluster 1: Flatiron, Soho, Sutton Place, Upper East Side – Colour Purple

Cluster 2: Washington Heights, Tudor City – Colour Blue

Cluster 3: – Colour Green

Cluster 4: Tribeca – Colour Orange

Cluster 0: Financial District – Colour Red

From the above images we can see that the dominant clusters are Cluster 0 (Red colour), Cluster 1(Purple colour), Cluster 2 (Blue colour). It can be inferred that the residents staying in the neighbourhoods belonging to any of these three clusters share more or less similar restaurant tastes i.e. residents of neighbourhoods belonging to the same cluster may share similar preferences.

My Recommendations

Why is it important that our selected neighbourhood should share a similar taste or restaurant profile as the dominant clusters? It's because a good restaurant should have the power to not only wow the residents of its own neighbourhood, but it should also have the strength to attract customers from other neighbourhoods. Having a cuisine that falls into the category which is loved by the greater Manhattan will only help expedite the restaurant's attraction power. However, choosing a cuisine is not so straightforward. (More on this later)

Given that the neighbourhoods belonging to cluster 1 have either high rent rates or crime rates when compared to the other neighbourhoods of the 8. The choices come down to

Washington Heights (Cluster 2)

Tudor City (Cluster 2)

Financial District (Cluster 0)

If you look at the rent and crime data graphs in the Results section, you will see that these three neighbourhoods have very low crime rates, which is a good quality.

However, the rent in Tudor City is quite high when compared to the likes of Washington Heights and Financial Districts.

	Neighborhood	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	9th Most Common Venue	10th Most Common Venue
0	Financial District	American Restaurant	Falafel Restaurant	Italian Restaurant	Mexican Restaurant	Japanese Restaurant	Mediterranean Restaurant	Seafood Restaurant	Restaurant	French Restaurant	New American Restaurant
1	Flatiron	Italian Restaurant	Japanese Restaurant	Mediterranean Restaurant	American Restaurant	Vegetarian / Vegan Restaurant	New American Restaurant	Cuban Restaurant	Fast Food Restaurant	Indian Restaurant	Kebab Restaurant
2	Soho	Italian Restaurant	Mediterranean Restaurant	French Restaurant	Sushi Restaurant	Australian Restaurant	Vegetarian / Vegan Restaurant	Falafel Restaurant	Egyptian Restaurant	Mexican Restaurant	American Restaurant
3	Sutton Place	Italian Restaurant	Vegetarian / Vegan Restaurant	Mexican Restaurant	American Restaurant	Cambodian Restaurant	French Restaurant	Latin American Restaurant	Lebanese Restaurant	Greek Restaurant	Persian Restaurant
4	Tribeca	Italian Restaurant	American Restaurant	Greek Restaurant	French Restaurant	Modern European Restaurant	Sushi Restaurant	Indian Restaurant	Korean Restaurant	Seafood Restaurant	Chinese Restaurant
5	Tudor City	Mexican Restaurant	Sushi Restaurant	Vietnamese Restaurant	Thai Restaurant	Asian Restaurant	Greek Restaurant	Seafood Restaurant	French Restaurant	Hawaiian Restaurant	Japanese Restaurant
6	Upper East Side	Italian Restaurant	American Restaurant	French Restaurant	Sushi Restaurant	Mexican Restaurant	Chinese Restaurant	Vegetarian / Vegan Restaurant	Latin American Restaurant	Mediterranean Restaurant	North Indian Restaurant
7	Washington Heights	Chinese Restaurant	New American Restaurant	Tapas Restaurant	Spanish Restaurant	Latin American Restaurant	Mexican Restaurant	Ramen Restaurant	Indian Restaurant	Italian Restaurant	Caribbean Restaurant

The neighbourhoods that I recommend finally are Washington Heights and Financial District. They both have very good crime and rent rates. They each have 23 and 20 restaurants respectively. Another thing to note is that these two neighbourhoods lie in the higher population density ranges according to the data provided by www.heath.ny.gov.

Regarding the cuisine, if you notice the above table, you will see that the residents of both of these places have quite different taste profiles. So, it comes down to the owner's or business person's decision regarding which cuisine he prefers his restaurant to have.

A Chinese, Spanish, Italian, Mexican restaurant may find success in Washington Heights

A Japanese, Mediterranean, French, Seafood, Falafel restaurant may find success in Financial District.

From the table in the previous page that contains the most common venues per each neighbourhood, I believe that starting a restaurant with a cuisine that falls into the 4th or 5th most common type would be a good move.

In order to further explain my above point.

An example:

- Suppose you wanted to open a restaurant in the Financial District. Opening another American restaurant would lead to heavy initial competition with the already well-established American Restaurants. A saturated market means that failure to meet their quality standards would definitely result in heavy losses. While succeeding through such a move is definitely possible, it creates a very high stakes environment for the restaurant itself.

- On the other hand, opening an Italian restaurant or Japanese restaurant reduces the competition significantly. While the restaurant will experience a comparatively slower start by opting this method, there is a possibility that the restaurant may grow and the dynamic may shift to these types of cuisines in the near future. This decision will also give the restaurant ample breathing room and potential opportunities to expand its type of cuisine. The analysis conducted clearly shows that Italian restaurants are well liked by the borough of Manhattan.

- A risky move would be opening a Chinese restaurant. While Financial District residents clearly do not prefer Chinese, good quality food as well as the overall popularity of Chinese restaurants in other parts of Manhattan may lead to its success.

If you have noticed by now, yes, a similar logic was applied during selecting neighbourhoods as well!

CONCLUSION

A lot of work has gone into this analysis and I have tried to make sure that the final result is as accurate as possible. However, the recommendations provided by this report are not the only right answers. You may definitely find success if you decide to open a restaurant in a neighbourhood that has not been advised in this project as well.

The results and inferences were derived using the logic that I considered to be the best, but if there is one thing that we know about Data Science, it’s that there are always multiple right answers to a single question.

That being said, I do hope you enjoyed reading this report and you found the insights useful.

Thank you for reading this far.