Food Destination- New York City or Houston, Texas

Matthew Kolakowski
July 7, 2020

1. Introduction

1.1 Background

Although the COVID-19 pandemic has halted non-essential travel now, planning a foodie vacation can still occur. Imagine traveling to New York City or Houston Texas and partaking in the diverse and nearly abundant food options. To narrow our food options and to prevent information overload, this project examined four types of food establishments:

1.   Pizza Places
2.   BBQ Joints
3.   Gastropub's
4.   Mexican Restaurants

Rather than solve a problem, it was the hope of this project to generate user-friendly lists and maps of each type of food establishment with Python (In IBM Watson Studio). You will find interesting maps and easily accessible information that not only is applicable to this final project, but if you enjoy food and are traveling to New York City or Houston Texas.

The diversity of offerings in both New York City and Houston Texas are impressive. While New York City can solidly lay claim to more pizza places and Gastropub's, Houston is a paradise for BBQ and Mexican cuisine enthusiasts.

1.2 Opportunity (Problem)

Rather than talk about a problem, this project discusses an opportunity to retrieve and display data that foodies (People who enjoy going out to eat) could utilize to determine the best locations for BBQ, Pizza, Gastropub's and Mexican Restaurants in New York City and Houston, Texas. This project aims to demonstrate that New York City is more favorable to Pizza and Gastropub enthusiasts and Houston, Texas is one of the top spots in the United States for BBQ and Mexican cuisine.

1.3 Interest

Food culture (Especially Post COVID-19) will again return to prominence. To demonstrate that Houston, Texas (While less accessible via public transportation compared to New York City) has robust BBQ and Mexican cuisine offerings. Additionally, New York City remains the home of finding that perfect slice of Pizza or interesting Gastropub.

2. Data Acquisition and Cleaning

2.1 Data sources

The source of data was Foursquare API. To pull the data, you simply would determine the location and then utilize the unique venue category: https://developer.foursquare.com/docs/build-with-foursquare/categories/ This page allows anyone to pull an unimaginable number of different venues that can be tailored to a diverse range of data questions and projects. An example of the Python code utilized to pull the data for Mexican Restaurants in New York City and Houston, Texas is as follows:

```python
# type your answer here
LIMIT = 500 # Maximum is 100
cities = ["New York, NY", 'Houston, TX']
results = {}
for city in cities:
    url = 'https://api.foursquare.com/v2/venues/explore?&client_id={} &client_secret={}&v={}&near={}&limit={}&categoryId={}'.format(
        CLIENT_ID,
        CLIENT_SECRET,
        VERSION,
        city,
        LIMIT,
        "4bf58dd8d48988d1c1941735") # MEXICAN RESTAURANT CATEGORY ID
    results[city] = requests.get(url). json()
```

2.2 Data cleaning
The fortunate portion of selecting Foursquare API data is that it comes across relatively clean and easy to examine in list form. Therefore, data cleaning techniques included organizing the display instead. A snippet of that code is below:

```python
df_venues= {}
for city in cities:
    venues = json_normalize(results[city]['response'] ['groups'][0]['items'])
    df_venues[city] = venues [['venue.name', 'venue. location. address', 'venue.location.lat', 'venue.location.lng']]
    df_venues[city]. columns = ['Name', 'Address', 'Lat', 'Lng']
```

2.3 Feature selection

To ensure consistency was achieve throughout the analysis, feature selection was standardized to location address, location latitude, location longitude, location city, and location name. This feature process was repeated for Pizza Places, BBQ Joints, Gastropub's, and Mexican Restaurants. To see the resulting maps that demonstrate these features, utilize this Python notebook viewer: Final Note: https://nbviewer.jupyter.org/

Link to GitHub Notebook:
https://github.com/MatthewK84/Coursera_Capstone/blob/master/Food%20Destination-%20New%20York%20City%20or%20Houston%2C%20Texas_.ipynb

3. Methodology

To find the appropriate venues, it was necessary to first find the Foursquare API venue category before proceeding to the analysis. The reason for matching the venue category to the desired variables: Pizza Place, BBQ Joint, Gastropub, and Mexican Restaurant was to present a uniform and easy to understand report. The idea behind this project was rather than demonstrating complexity through a myriad of code, was to utilize Python to show useful simplicity. Additionally, I found it interesting that Houston, Texas had robust dining opportunities spread across a vast area whereas New York City dining is densely packed into the five boroughs.

## 3.1 Relationship between cuisine type and geographical location

To get an idea of the relationship between cuisine type and geographical location, I present maps of Pizza Places in New York City and Houston, Texas.

Figure 1. New York City Pizza Places
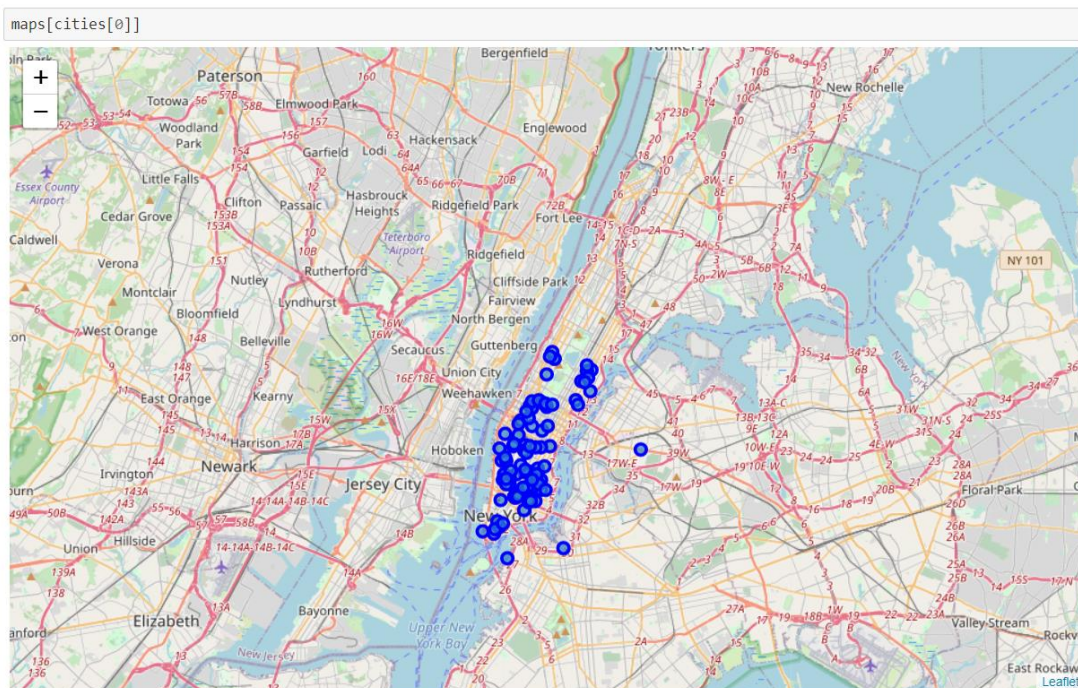
The New York City map has a total of 271 pizza places while Houston, Texas has a total of 191. The immediate difference from figure 1 versus figure 2 is the density. New York City's robust public transportation and tight locations would allow a pizza lover to have accessible options. Houston, Texas, on the other hand, is less accessible unless you had a rental car that afforded you increased mobility to travel to various parts of the city. While one example, it demonstrates a clear pattern throughout the analysis: New York City and its cuisine is more accessible and denser than Houston, Texas spread out culinary scene.

Figure 2. Houston Texas Pizza Places

```
: maps[cities[1]]
```
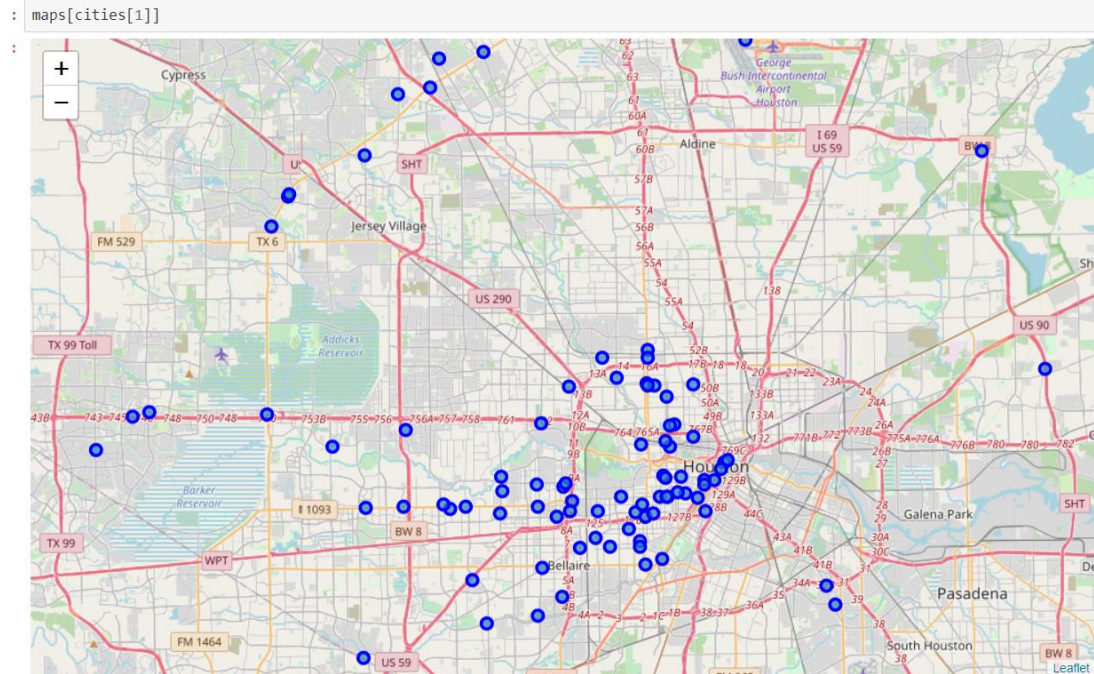

3.2 Exploratory Analysis

To examine nuance, in comparison to only one factor, I utilized four different types of cuisine: Pizza, BBQ, Gastropub, and Mexican. This allowed several visual displays and lists to be generated. If you chose to just read this report alone, you can find a great visual of this reports corresponding Python Notebook:
https://nbviewer.jupyter.org/github/MatthewK84/Coursera_Capstone/blob/master/Food%20Destination-%20New%20York%20City%20or%20Houston%2C%20Texas_.ipynb

One of the neat features on how this analysis explored the data was the generation of maps. The "Show Top 100" function is something that I utilized and wanted to share the Python code to:

```python
maps = {}
for city in cities:
    city_lat = np. mean([results[city]['response'] ['geocode'] ['geometry'
] ['bounds'] ['ne'] ['lat'],
                        results[city]['response'] ['geocode'] ['geometry']
 ['bounds'] ['sw'] ['lat']])
    city_lng = np. mean([results[city]['response'] ['geocode'] ['geometry'
] ['bounds'] ['ne'] ['lng'],
                        results[city]['response'] ['geocode'] ['geometry']
 ['bounds'] ['sw'] ['lng']])
    maps[city] = folium. Map (location= [city_lat, city_lng], zoom_start=1
1)

    # add markers to map
    for lat, lng, label in zip(df_venues[city]['Lat'], df_venues[city]['Ln
g'], df_venues[city]['Name']):
        label = folium. Popup (label, parse_html=True)
```

```
        folium. CircleMarker (
            [lat, lng],
            radius=5,
            popup=label,
            color='blue',
            fill=True,
            fill_color='#3186cc',
            fill_opacity=0.7,
            parse_html=False). add_to(maps[city])
    print (f"Total number of pizza places in {city} = ", results[city]['re
sponse'] ['totalResults'])

    print ("Showing Top 100")
    print("_____")
```

This function resulted in the easy to read maps in figures 1 and 2. Additionally, the code served as an exploratory function that built nuance and an ample level of detail into the analysis. The Foursquare API, if leveraged correctly, can be the building block to a robust yet user-friendly analysis of a city or cities features/venues.

4. Results

The results of this analysis are as follows:

1. Houston, Texas is a BBQ and Mexican cuisine destination. If lieu of the large population delta between New York City and Houston, Texas; each city's culinary scenes provide even the most discerning foodie robust dining options.

2. New York City has 271 pizza places while Houston, Texas has 191. It was surprising to find that even in Houston, you could eat Pizza twice a week every week for nearly 2 years and not dine at the exact same location twice. I found that to be interesting and something that a foodie would find insightful as well.

3. The density of New York City was also surprising. While the boroughs are spread out over a large land mass, the dining options are quite tightly confined as shown in Figure 3 below

Figure 3. New York Pizza Places Density Map

```
maps[cities[0]]
```
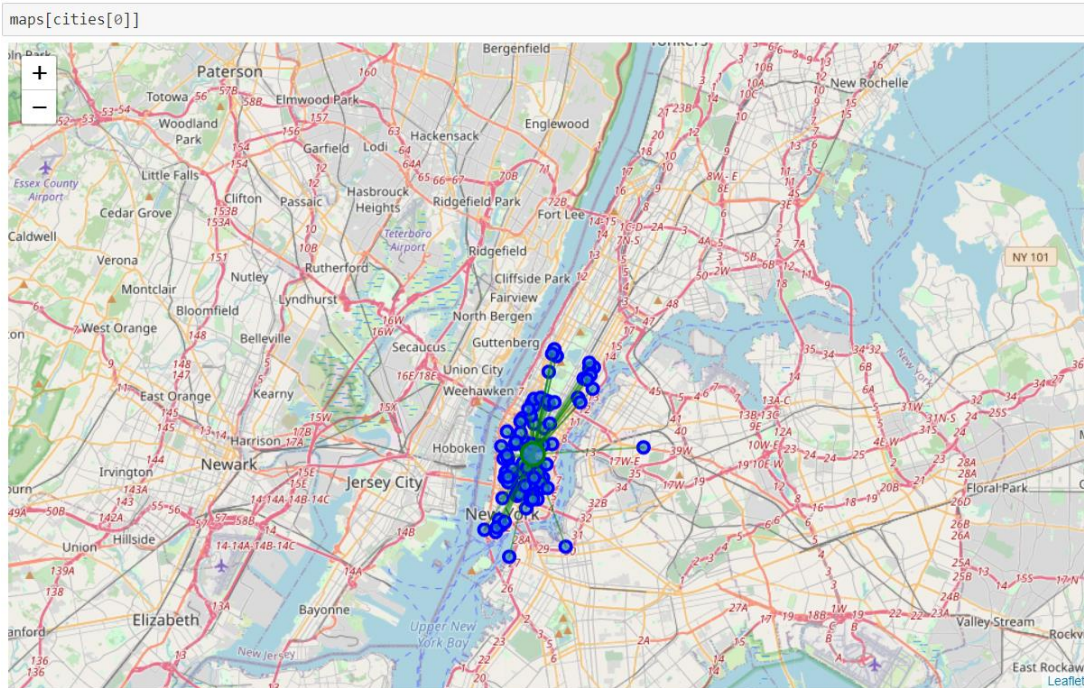

4. If I were going to take this information and create a "foodie guide" focused on Pizza Places, BBQ Joints, Gastropub's, and Mexican Restaurants, I would ensure to create alphabetical listings for increased readability as the number of venues per category is quite large. The alphabetized listing would also have the address and link to menu if possible, to drive virtual traffic to said guide.

5. Discussion

This analysis was particularly interesting as the number of dining destinations in Houston, Texas are larger in two categories (BBQ and Mexican Restaurants) than New York City. New York City, on the other hand, has a clear advantage for foodies that enjoy Pizza and Gastropub's. The power of the Foursquare API cannot be fully appreciated until completing an analysis like this. The ease of accessing the data, the accuracy, and clarity of displays makes it a go-to product if comparing and contrasting features/venues in major cities across the United States. It was the aim of this analysis to compare at least four types of cuisine, which I believe was quite successful after reviewing my Python code and resulting information.

6. Conclusion

After reading this report, I hope you have seen a reasonable demonstration of Python skills that resulted in results that are simplistic in design yet meaningful. For future data projects, more robust methodology will be undertaken as my confidence as a data scientist/analyst increases. To review the complete project, please go to my GitHub or just follow this link where I already put it into the viewer so you only had to click one thing:
https://nbviewer.jupyter.org/github/MatthewK84/Coursera_Capstone/blob/master/Food%20Destination-%20New%20York%20City%20or%20Houston%2C%20Texas_.ipynb

GitHub Project Link:

https://github.com/MatthewK84/Coursera_Capstone/blob/master/Food%20Destination-%20New%20York%20City%20or%20Houston%2C%20Texas_.ipynb