Business Problem

New York City is the largest city in the United States and one of the world's largest tourist destinations. There are 65 Million Tourists to New York City a year. The heart of the destination in New York City is centered around Times Square and Broadway. This area is known as the theater district. In 2019, 15 million attended a broadway show. There are numerous bars and restaurants in the theater district. I want to know what cuisine the restaurants are serving, the number of types of restaurants, and the price point.

Questions

1) How many restaurants are in the theater district?
2) Where are they located
3) What cuisine are they serving?
4) How expensive are the restaurants?

Data

Broadway and the theater district, there are many restaurants with diverse food and different price points. So there will be something for everyone.

But in order to find it successfully the project will need location data, Foursquare Data, and Geospatial Data.

New York City Borough

This data set contains the boundries of the boroughs, neighborhoods with their longitude, and latitude.

Data set- https://cocl.us/new_york_dataset

Description: This Data Set will allow me to expore the theater district.

Restaurants in Theater Square

Data Set: Foursquare API

The Data that will be used will be Foursquare for resturants and theaters, because it will populate the data and we can filter by resturant type and find how many resturants are in the neighborhood.

GeoSpace Data

This data source will be useful in determining where the boundaries of the theater district are and how we can determine where the restaurants are and how many

are in the neighborhood and help us visit a chlorophenol map of the theater district.

https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm

Methodology

The First step was to determine what the locations of Time Square. I decided to use the location of the Marriott Marque in New York City. I used this code to get the Latitude and Longitude using the Geolocator Library.

In [5]:
```
address='1535 Broadway, New York, NY 10036'
print('Done')
```

```
Done
```

In [6]:
```
address = '1535 Broadway, New York, NY'

geolocator = Nominatim(user_agent="foursquare_agent")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print(latitude, longitude)
```

```
40.7586132 -73.98620679523486
```

Then I used the Foursquare API to get the venues and I filtered the data.

[10]:
```
LIMIT = 900 # limit of number of venues returned by Foursquare API
radius = 10000 # define radius
# create URL
url = 'https://api.foursquare.com/v2/venues/explore?client_id={}&client_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.forr
    CLIENT_ID,
    CLIENT_SECRET,
    latitude,
    longitude,
    VERSION,
    search_query,
    radius,
    LIMIT)
url # display URL
```

t[10]:
```
'https://api.foursquare.com/v2/venues/explore?client_id=UGVUZN2V5GGFR2X520DBTUD0KLD3AJ5PUSQPDT5L3LFHALMT&client_secret=UUPEWJ4U
HDQFZENLEVDE0BNQPVVNKDPSZLJXSTP3Y4SYMWZK&ll=40.7586132,-73.98620679523486&v=20200711&query=resturants&radius=10000&limit=900'
```

[11]:
```
results = requests.get(url).json()
```

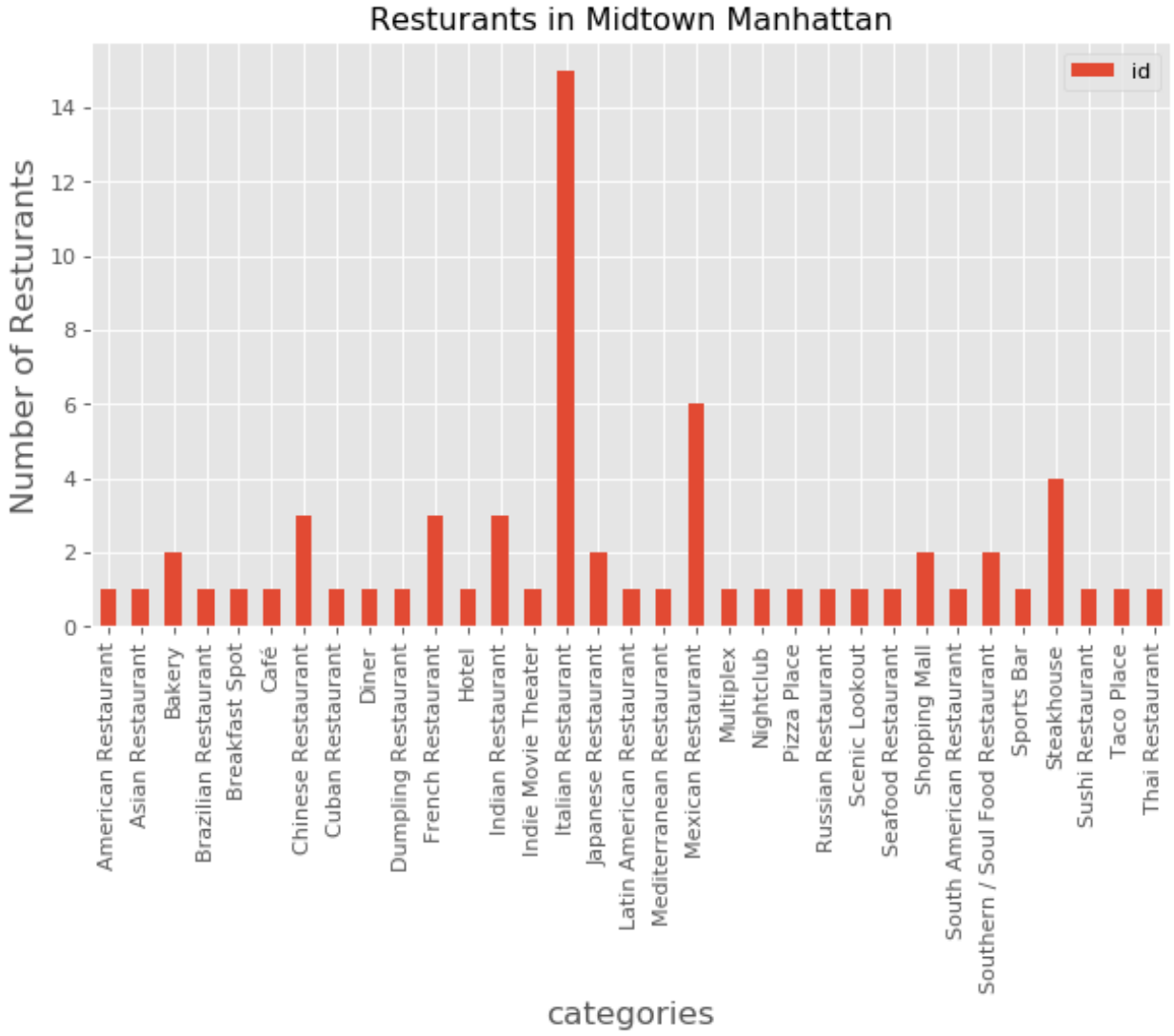
[12]:
```
items = results['response']['groups'][0]['items']
items[0]
```

t[12]: {'reasons': {'count': 0

Then filtering the data and seeing what the categories where during this meal.

13]:
```
dataframe = pd.json_normalize(items) # flatten JSON

# filter columns
filtered_columns = ['venue.name', 'venue.categories'] + [col for col in dataframe.columns if col.startswith('venue.location.')] +
dataframe_filtered = dataframe.loc[:, filtered_columns]

def get_category_type(row):
    try:
        categories_list = row['categories']
    except:
        categories_list = row['venue.categories']

    if len(categories_list) == 0:
        return None
    else:
        return categories_list[0]['name']

# filter the category for each row
dataframe_filtered['venue.categories'] = dataframe_filtered.apply(get_category_type, axis=1)

# clean columns
dataframe_filtered.columns = [col.split('.')[-1] for col in dataframe_filtered.columns]

dataframe_filtered.head(10)
```

I did some exploratory data analysis to determine what the most common restaurants and create a visualization to see the number of restaurants in Midtown Manhattan.



Finally, I used the Folium library to visual the restaurants locations to see where to open one up.

Results

The results were that there are mostly Italian restaurants in Manhattan with 14. There are more restaurants and most others would be effective.

Conclusions

The conclusions is that it wise to open up any other restaurant except Italian in New York City. Though Midtown Manhattan might be a little though to open up a restaurant in that neighborhood.