# LOCATING A NEW RESTAURANT

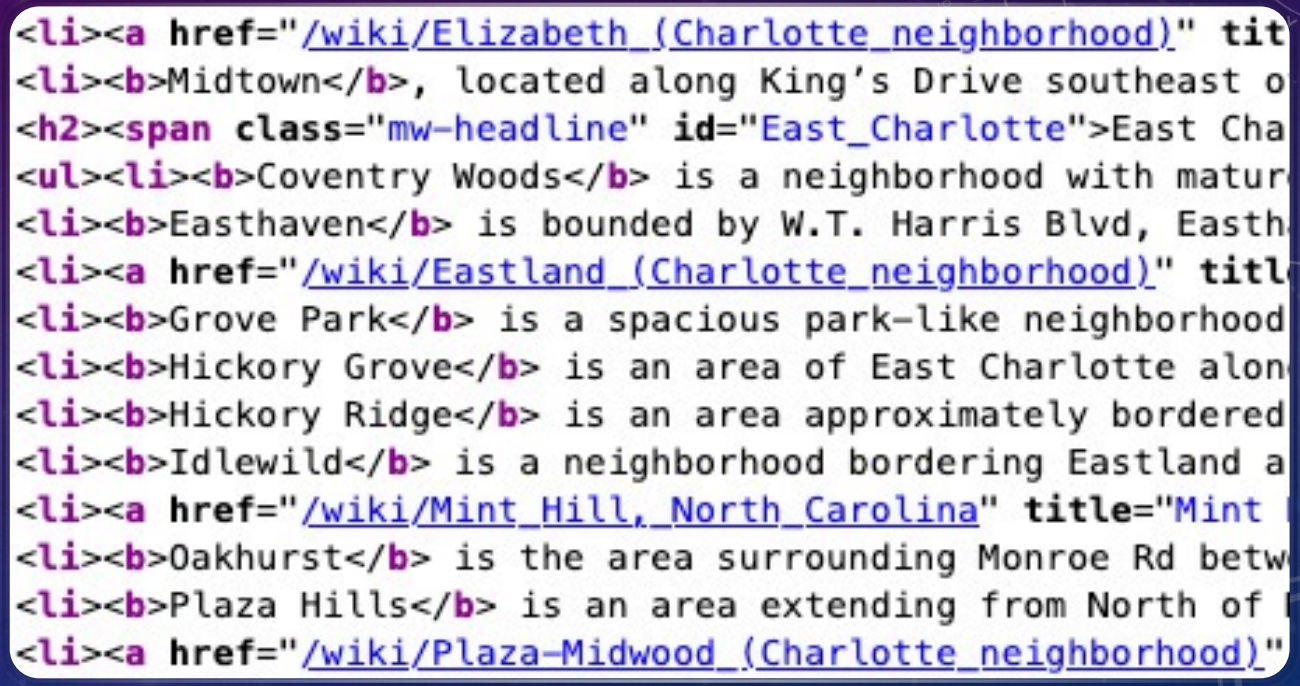
IN CHARLOTTE, NORTH CAROLINA

# THE PROJECT

- Customer is looking to invest in the restaurant business in Charlotte, North Carolina
- Belief that there is rising demand following the relaxing of quarantine following the pandemic
- Look to compare Charlotte neighborhoods that all belong to the same affluent cluster
- Establish what cuisines are most popular and discover what may be under-represented in the neighborhoods

# DATA

- Wikipedia provides a listing of Charlotte neighborhoods at: https://en.wikipedia.org/wiki/List_of_Charlotte_neighborhoods
- Duplicates, mis-tagged and invalid data are removed/edited



# DATA

GEOPY PROVIDES THE LATITUDE AND LONGITUDE DATA WE NEED FOR THE NEIGHBORHOODS

```
for item in neighborhoods:
    loc = item + ", Charlotte, NC"
    geolocator = Nominatim(user_agent="my_request")
    location = geolocator.geocode(loc)
    print(item, location.latitude, location.longitude)
    lat_list.append(location.latitude)
    long_list.append(location.longitude)
```

```
Derita 35.2833128 -80.8192863
Starmount 35.1382582 -80.8683834
Quail Hollow 35.1186739 -80.8399701
Steele Creek 35.14409795 -80.97896658564757
Paw Creek 35.263798 -80.916153
Eastland 35.2088208 -80.7513256
```

```
# Prepare inputs to retrieve top 100 venues from this first neighborhood from Foursquare
limit = 100
radius = 500
url = 'https://api.foursquare.com/v2/venues/explore?client_id={}&client_secret={}&v={}&l

# Request URL from Foursquare
results = requests.get(url).json()
results
```

```
'venue': {'id': '4b105607f964a520276e23e3',
 'name': 'Cafe Monte',
 'location': {'address': '6700 Fairview Rd',
  'crossStreet': 'Allen Tate Building at Phillips Place',
  'lat': 35.14715397463186,
  'lng': -80.82872611013036,
  'labeledLatLngs': [{'label': 'display',
    'lat': 35.14715397463186,
    'lng': -80.82872611013036}],
  'distance': 244,
  'postalCode': '28210',
  'cc': 'US',
  'neighborhood': 'SouthPark',
  'city': 'Charlotte',
  'state': 'NC',
```

# DATA

FOURSQUARE RETURNS THE LIST OF VENUES FOR THE NEIGHBORHOODS

# DATA

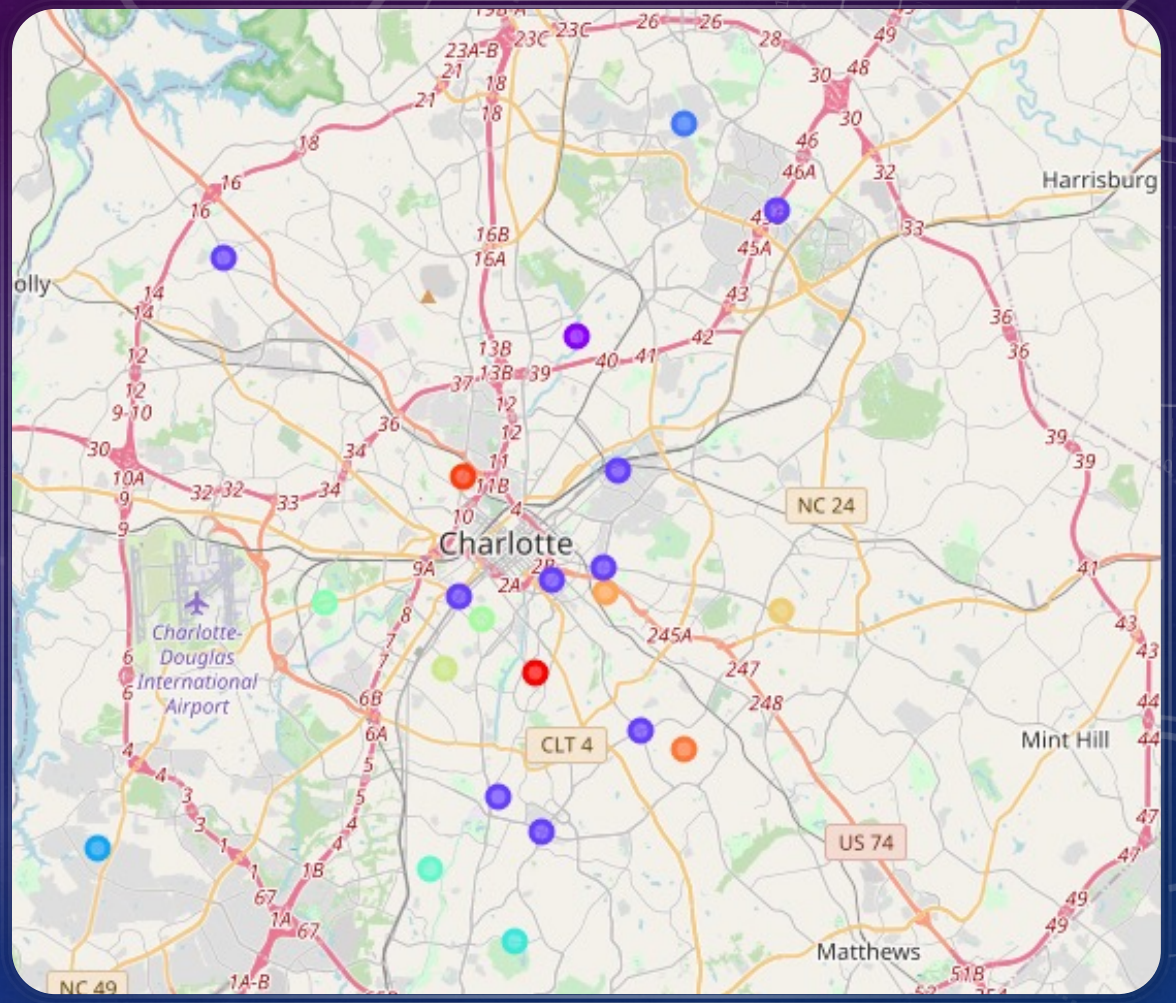
```
# Cluster Neighborhoods

# set number of clusters
# Having tried kclusters at 7 and 10 which had most of the neighborhoods in one category
# I found that 15 works well for this exercise, giving a large enough cluster to focus on
# for further processing
kclusters = 15

charlotte_grouped_clustering = charlotte_grouped.drop('Neighborhood', 1)

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(charlotte_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:15]
```

KMEANS IS USED TO GROUP THE CHARLOTTE NEIGHBORHOODS INTO CLUSTERS

KLUSTERS OF 15 GIVE US A WORKABLE CLUSTER SIZE

# DATA

FOLIUM IS USED TO MAP THE LOCATION OF THE NEIGHBORHOODS AND SHOW THE CLUSTERS THEY BELONG TO



# METHODOLOGY

A UNIQUE LIST OF NEIGHBORHOODS FROM THE WIKIPEDIA PAGE IS GENERATED

```
# Remove any duplicates from the list
neighborhoods = list(set(neighborhoods))
neighborhoods
```

```
['Dilworth',
 'Cotswold',
 'Coulwood',
 'Myers Park',
 'Highland Creek',
 'Paw Creek',
 'Ballantyne',
 'Mallard',
 'Starmount',
 'Parkdale',
 'Sedgefield',
 'Steele Creek',
 'South End',
 'Sherwood Forest',
 'Quail Hollow',
 'Elizabeth',
 'Eastland',
 'Reid Park',
 'NoDa',
 'University City',
 'Biddleville',
 'Chantilly',
 'SouthPark',
 'Plaza-Midwood',
 'Derita']
```

# METHODOLOGY

A DATAFRAME IS GENERATED THAT TIES THE NEIGHBORHOOD TO ITS LATITUDE AND LONGITUDE

| | Neighborhood | Latitude | Longitude |
|---|---|---|---|
| 0 | Dilworth | 35.206612 | -80.850914 |
| 1 | Cotswold | 35.175924 | -80.798330 |
| 2 | Coulwood | 35.304209 | -80.936693 |
| 3 | Myers Park | 35.191735 | -80.833489 |
| 4 | Highland Creek | 35.386623 | -80.760504 |
| 5 | Paw Creek | 35.263798 | -80.916153 |
| 6 | Ballantyne | 35.054659 | -80.850246 |
| 7 | Mallard | 35.340920 | -80.783890 |
| 8 | Starmount | 35.138258 | -80.868383 |
| 9 | Parkdale | 35.158287 | -80.845661 |
| 10 | Sedgefield | 35.192921 | -80.863405 |

# METHODOLOGY - CLUSTERING

KMEANS GIVES US A CLUSTER OF 10 NEIGHBORHOODS TO ANALYZE

| | Neighborhood | Cluster Labels | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue |
|---|---|---|---|---|---|---|
| 1 | Cotswold | 2 | Pizza Place | Cosmetics Shop | Furniture / Home Store | Fast Food Restaurant |
| 2 | Coulwood | 2 | Baseball Field | Gas Station | Gastropub | Pharmacy |
| 6 | Ballantyne | 2 | Pizza Place | Asian Restaurant | Coffee Shop | Pub |
| 9 | Parkdale | 2 | Spa | Grocery Store | Italian Restaurant | American Restaurant |
| 12 | South End | 2 | Coffee Shop | American Restaurant | Bakery | Thai Restaurant |
| 15 | Elizabeth | 2 | New American Restaurant | Park | Stadium | Bike Shop |
| 18 | NoDa | 2 | Bar | Brewery | Gastropub | Sports Bar |
| 19 | University City | 2 | Brewery | Warehouse Store | Furniture / Home Store | Gym / Fitness Center |
| 22 | SouthPark | 2 | Clothing Store | Women's Store | American Restaurant | Jewelry Store |
| 23 | Plaza-Midwood | 2 | Bar | Pizza Place | Gastropub | Brewery |

# METHODOLOGY

- A unique list of restaurant types is generated

```
# Get a unique list of Restaurant Types/Cuisines
rest_list = list(set(charlotte_cluster_restaurants["Venue Category"]))
rest_list
```

```
['Ramen Restaurant',
 'Greek Restaurant',
 'New American Restaurant',
 'Vegetarian / Vegan Restaurant',
 'Mexican Restaurant',
 'Japanese Restaurant',
 'Peruvian Restaurant',
 'Thai Restaurant',
 'Indian Restaurant',
 'Italian Restaurant',
 'Asian Restaurant',
 'Chinese Restaurant',
 'Seafood Restaurant',
 'Eastern European Restaurant',
 'Sushi Restaurant',
 'French Restaurant',
 'Southern / Soul Food Restaurant',
 'Tapas Restaurant',
 'Restaurant',
 'Caribbean Restaurant',
 'Tex-Mex Restaurant',
 'American Restaurant']
```

```
# What are the top 5 cuisines that we should consider
top5 = popular_restaurants.head(5).reset_index(drop=False)
top5
```

| | Cuisine | Neighborhood |
|---|---|---|
| 0 | American Restaurant | 12 |
| 1 | Chinese Restaurant | 7 |
| 2 | Asian Restaurant | 6 |
| 3 | Mexican Restaurant | 6 |
| 4 | Italian Restaurant | 4 |

# METHODOLOGY

THE TOP 5 MOST POPULAR RESTAURANT CUISINES ARE GENERATED

| Neighborhood<br>Cuisine | Ballantyne | Cotswold | Elizabeth | Parkdale | Plaza-Midwood | South End | SouthPark | All |
|---|---|---|---|---|---|---|---|---|
| American Restaurant | 2 | 0 | 0 | 1 | 1 | 3 | 5 | 12 |
| Asian Restaurant | 3 | 0 | 0 | 0 | 1 | 2 | 0 | 6 |
| Chinese Restaurant | 1 | 1 | 1 | 0 | 1 | 1 | 2 | 7 |
| Italian Restaurant | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 4 |
| Mexican Restaurant | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 6 |
| All | 8 | 2 | 2 | 3 | 3 | 8 | 9 | 35 |

# METHODOLOGY

THE TOP 5 CUISINES ARE MATCHED TO THE NEIGHBORHOODS IDENTIFIED BY THE CLUSTER

# CONCLUSION

| Neighborhood / Cuisine | Ballantyne | Cotswold | Elizabeth | Parkdale | Plaza-Midwood | South End | SouthPark | All |
|---|---|---|---|---|---|---|---|---|
| American Restaurant | 2 | 0 | 0 | 1 | 1 | 3 | 5 | 12 |
| Asian Restaurant | 3 | 0 | 0 | 0 | 1 | 2 | 0 | 6 |
| Chinese Restaurant | 1 | 1 | 1 | 0 | 1 | 1 | 2 | 7 |
| Italian Restaurant | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 4 |
| Mexican Restaurant | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 6 |
| All | 8 | 2 | 2 | 3 | 3 | 8 | 9 | 35 |

- Utilize the results to find popular cuisines that are underrepresented in neighborhoods with otherwise large number of restaurants