



Restaurant Analysis in Charlotte, NC, USA

By: Shyam Sridharan

Background

- Charlotte, North Carolina is a growing city, approximately 873,000+ residents
- Dramatic population increase
- Banking hub
- Surge of new Restaurant Business because of population boom and Food Delivery Services

Business Problem

- Given a city and its neighborhoods can we gather information on the restaurants in the area in order to be better informed on hotspots for multinational cuisine and locations in need of more variety using data driven analysis?

Target Audience

- Consumers looking to enjoy a variety of options for cuisines moving to Charlotte
- Entrepreneurs looking to open Restaurants in the city
- Drivers for Food Delivery apps trying to maximize amount of deliveries

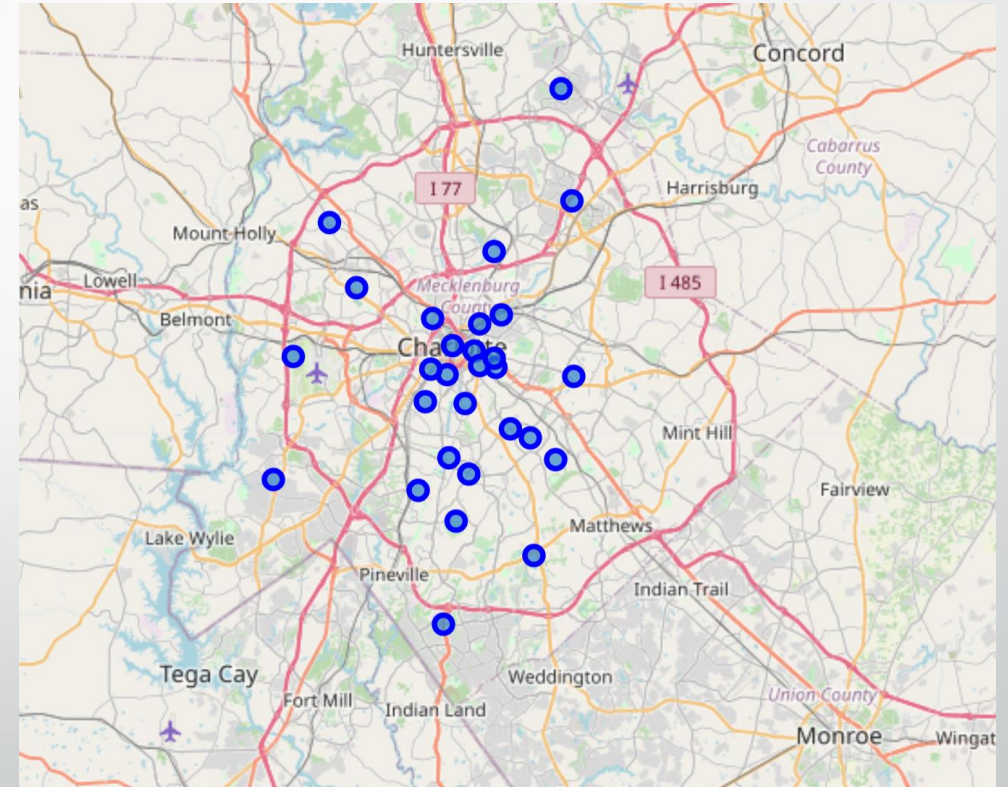
Data Preparation- Neighborhoods

- Initial Neighborhood data procured from Wikipedia page using BeautifulSoup some cleaning necessary

Pages in category "Neighborhoods in Charlotte, North Carolina"		
The following 30 pages are in this category, out of 30 total. This list may not reflect recent changes (learn more).		
<ul style="list-style-type: none">• List of Charlotte neighborhoods	<ul style="list-style-type: none">• Derita (Charlotte neighborhood)• Dilworth (Charlotte neighborhood)	<ul style="list-style-type: none">• Parkdale (Charlotte neighborhood)• Paw Creek (Charlotte neighborhood)• Plaza-Midwood (Charlotte neighborhood)
A <ul style="list-style-type: none">• The Arboretum, Charlotte	E <ul style="list-style-type: none">• Eastland (Charlotte neighborhood)• Elizabeth (Charlotte neighborhood)	Q <ul style="list-style-type: none">• Quail Hollow (Charlotte neighborhood)
B <ul style="list-style-type: none">• Ballantyne (Charlotte neighborhood)• Biddleville (Charlotte neighborhood)• Brooklyn (Charlotte, North Carolina)	H <ul style="list-style-type: none">• Highland Creek (Charlotte neighborhood)	S <ul style="list-style-type: none">• Sedgefield (Charlotte neighborhood)• Sherwood Forest (Charlotte neighborhood)• South End (Charlotte neighborhood)• SouthPark (Charlotte neighborhood)• Starmount (Charlotte neighborhood)• Steele Creek (Charlotte neighborhood)• Stonehaven (Charlotte neighborhood)
C <ul style="list-style-type: none">• Chantilly (Charlotte neighborhood)• Uptown Charlotte• Cotswold (Charlotte neighborhood)• Coulwood (Charlotte neighborhood)	M <ul style="list-style-type: none">• Mountain Island, North Carolina• Myers Park (Charlotte)	
D	N <ul style="list-style-type: none">• NoDa (Charlotte neighborhood)• North Charlotte (Charlotte neighborhood)	U <ul style="list-style-type: none">• University City (Charlotte neighborhood)
	P	

Data Preparation- Coordinates

- Find Geographical coordinate data for each Neighborhood and attach to existing dataset of Neighborhoods

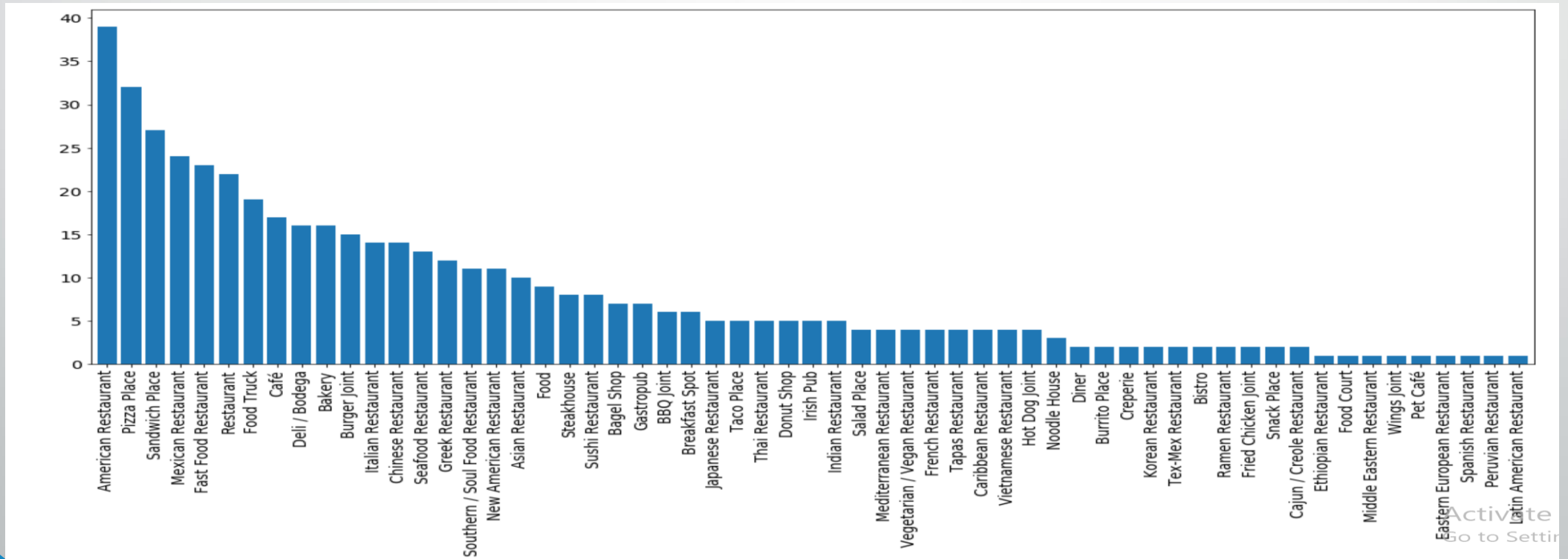


Data Preparation- FourSquare API

- FourSquare is a service with a large database consisting of venues of all type for countries across the world. What we must do is create a free, or paid if API is used extensively, developer account so we can connect to the API
- We use our developer credential to make a GET request for restaurant venues by specifying the categoryId for food: "4d4b7105d754a06374d81259" as well as setting a custom radius and limit of results
- Dataset is now a list of 481 venues

Exploratory Data Analysis

- View the value counts of distinct Venue Categories to see which Venues are most common



Model Selection

- Project goal is to find distinction between Neighborhoods so the best approach is some way to group together similar Neighborhoods
- Perfect task for Clustering more specifically K-means Clustering

Data Preparation

- Convert dataset with 481 observations
- First use One-Hot Encoding and then use group by to find distribution of Venue Categories for each Neighborhood, then trim off the least common Venue Categories by only listing the top 10 most common venue categories for each Neighborhood

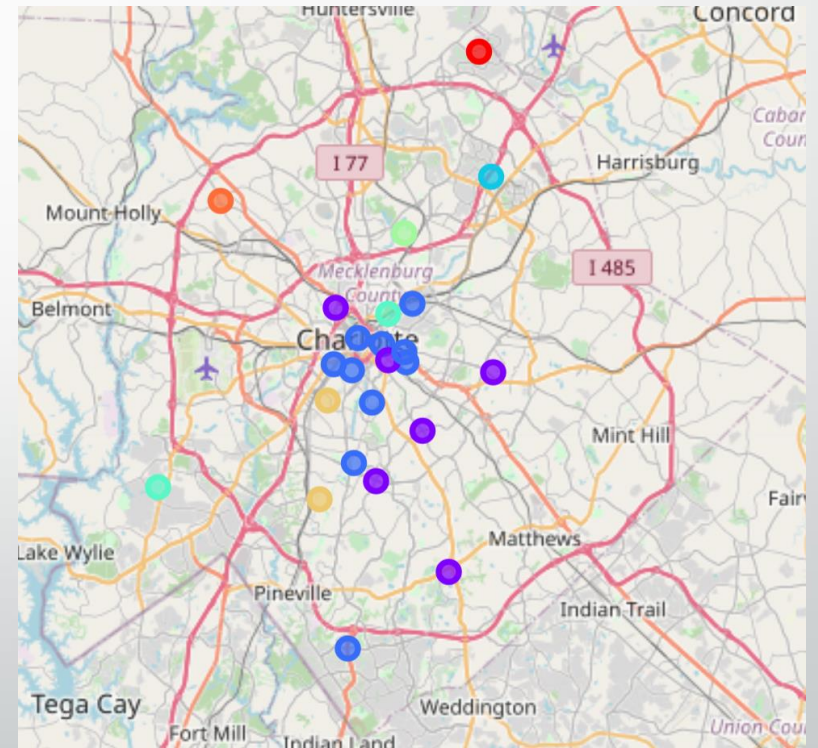
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Ballantyne	American Restaurant	Pizza Place	Asian Restaurant	Bakery	Burger Joint	Food Truck	Italian Restaurant	Mexican Restaurant	Indian Restaurant	Wings Joint
1	Biddleville	Café	Fried Chicken Joint	Mediterranean Restaurant	Pizza Place	Wings Joint	Diner	Gastropub	French Restaurant	Food Truck	Food Stand
2	Brooklyn	Pizza Place	Food	Deli / Bodega	Burger Joint	Food Truck	Food Stand	Taco Place	Restaurant	Southern / Soul Food Restaurant	Mexican Restaurant

Model Deployment

- Sklearn is a widely used package for model deployment and machine learning, K-means clustering is part of this package
- Set up your K-means Clustering, only argument is number of clusters which you can experiment with to find correct one

Results

- Of the 8 clusters; clusters 0, 3, 4, 5, 6, 7 were made up of 2 or less neighborhoods while clusters 1 and 2 were made up of many. Clusters 1 and 2 are purple and blue respectively



Conclusions

- Clusters 0, 3, 4, 5, 6, 7 were composed of neighborhoods with many of the common venue categories such as American restaurants, Fast Food, Mexican restaurants, Burger joints, Hot Dog Joints etc. Suggest that those wishing to stand out with multicultural food should open here
- Cluster 2 was more distinct than previous and neighborhoods were close together, recommended area for those wishing to make most money on as Deliverers through Delivery applications
- Cluster 1 was home to multiple multicultural culinary options such as French, Indian, and different European and Asian dining.
- This is the most popular destination for those looking for multicultural dining

Future Improvements

- Find other APIs to connect and find more data such as a ratings API
- Do more proximity analysis into neighborhoods and connect to Housing listings to in theory find houses near multiple venues
- Create a web application allowing for the functionality to be interactive
- Procure more venues to look at to get more specific results
- Looking into menus for a more in-depth look concerning dishes similar or dissimilar
- Scrape venues from social media