Introduction and Business Problem

Introduction
New York City (NYC), also known as the City of New York or simply New York (NY), is the most populous city in the United States. With an estimated 2018 population of 8,398,748 distributed over a land area of about 302.6 square miles (784 km2), New York is also the most densely populated major city in the United States.

For people that are new to New York, it can be a problem to figure out what restaurants are worth going to and where they are. For people that used to live in New York or are visiting New York, how do you know what the best places are to get something to eat?

Business Problem
For this project, I am going to create a simple guide on where to eat based on Foursquare likes, restaurant category and geographic location data for restaurants in New York. I will then cluster these restaurants based on their similarities so that a user can easily determine what type of restaurants are best to eat at based on Foursquare user feedback.

Data Requirements and Methodology

Data Requirements
For this assignment, I will be utilizing the Foursquare API to pull the following location data on restaurants in New York:
- Venue Name
- Venue ID
- Venue Location
- Venue Category
- Count of Likes

To acquire the data mentioned above, I will need to do the following:
- Get geolocator lat and long coordinates for New York
- Use Foursquare API to get a list of all venues in New York
- Get venue name, venue ID, location, category, and likes

Methodology
The idea behind this is that likes are a proxy for quality. The more likes there are, the better the restaurant is. I will then bin this data into a quality categorical variables so we can cluster appropriately.

I am also going to create new categorical variables for the restaurants to better group them based on type of cuisine. This way you can look for good food or what type of food might be best to eat in New York if you are new to the area.

I will take the gathered data and will create a k-means clustering algorithm that groups restaurants into 4-5 clusters so that people looking to eat in Hoboken can easily see which restaurants are the best to eat at, what cuisine is available and where in Hoboken they can look to eat.

Results

Running my clustering algorithm, I was able to generate four clusters of restaurants. These are as follows:

Cluster 1

- Characteristics
  - o Poor quality food
  - o Mostly asia food or american food

	name	id	categories	Latitude	Longitude	total likes	total likes_cat	categories_new	label
6	Manhatta	5ada32224f0e584592628c1f	New American Restaurant	40.707654	-74.009138	178	poor	american food	0
10	Crown Shy	5c883f65f4b525002c0bf2ca	Restaurant	40.706187	-74.007490	106	poor	american food	0
25	Wayla	5cc4e9d0c876c8002c3010cb	Thai Restaurant	40.718291	-73.992584	91	poor	asia food	0
27	CAVA	59e0ee84f0ca95526b7fce9f	Mediterranean Restaurant	40.721928	-73.996512	105	poor	other	0
46	Raku	5aea422a033693002bf0c1cb	Udon Restaurant	40.727404	-74.002696	143	poor	asia food	0
47	Ramen Ishida	591b9e948496ca6e22cdd3f6	Asian Restaurant	40.719635	-73.988366	74	poor	asia food	0
56	Carmine Street Beers	5217e0dd11d2b33a2d55a11c	Beer Store	40.729945	-74.003728	91	poor	bars	0
80	Llama-San	5b95359481a0ea002cdce0ce	Peruvian Restaurant	40.732264	-74.000795	33	poor	american food	0
99	Lillo	5a0740a3e1f22816d11723d5	Italian Restaurant	40.690200	-73.996540	45	poor	euro food	0

Cluster 2

- Characteristics
  - o High quality food
  - o Mostly Europe food

	name	id	categories	Latitude	Longitude	total likes	total likes_cat	categories_new	label
21	The Dead Rabbit	50bcb93ae4b0efcb4c84d434	Cocktail Bar	40.703051	-74.010992	2016	great	bars	1
31	Rubirosa Ristorante	4cc6222106c25481d7a4a047	Italian Restaurant	40.722706	-73.995957	2466	great	euro food	1
38	Chobani	50002274e4b0deb8db4eea3b	Snack Place	40.725819	-74.000985	1027	great	other	1
39	Russ & Daughters Café	5244bd0e11d2d511de3e244e	Café	40.719515	-73.989724	1120	great	other	1
52	Cotenna	543075b3498e283e38ad9dec	Italian Restaurant	40.729014	-74.003539	870	great	euro food	1
82	Buvette	4d0bf7e3f29c236ac675bfe7	French Restaurant	40.732782	-74.004275	1831	great	euro food	1
87	L'Artusi	4a27289cf964a52062911fe3	Italian Restaurant	40.733888	-74.005114	1152	great	euro food	1
89	Joseph Leonard	4a78c865f964a52068e61fe3	New American Restaurant	40.733681	-74.001767	1142	great	american food	1
91	Cafe Mogador	41044980f964a520750b1fe3	Moroccan Restaurant	40.727277	-73.984505	1385	great	euro food	1

Cluster 3

- Characteristics
  - Below average quality food
  - Mostly other food

	name	id	categories	Latitude	Longitude	total likes	total likes_cat	categories_new	label
11	Black Fox Coffee Co.	574456e5498e39c73d234da1	Coffee Shop	40.706573	-74.008155	406	below avg	other	2
15	Le Coucou	57583641498e90001a32e13e	French Restaurant	40.719114	-74.000202	370	below avg	euro food	2
43	Champion Pizza - Ludlow	557a7031498e4837f11ba9e2	Pizza Place	40.719190	-73.988850	242	below avg	italian food	2
55	Court Street Grocers	5644c0cf498eda8d003cfa4b	Sandwich Place	40.728806	-73.998471	355	below avg	other	2
67	Faicco's Italian Specialties	4a74a36af964a520fede1fe3	Sandwich Place	40.731117	-74.003043	327	below avg	other	2
92	té (Té Company)	56292858498e5a271a288948	Tea Room	40.734537	-74.001828	187	below avg	other	2
93	Good Beer NYC	4cb34b64aef16dcb6287c754	Beer Store	40.727643	-73.983835	391	below avg	bars	2
94	4 Charles Prime Rib	585164b77220e62219c9aeb6	Steakhouse	40.735219	-74.000649	203	below avg	american food	2
95	Mah-Ze-Dahr Bakery	568c0ce238fafac5f5ffe631	Bakery	40.735117	-74.000052	234	below avg	other	2

Cluster 4

- Characteristics
  - Above average quality food

	name	id	categories	Latitude	Longitude	total likes	total likes_cat	categories_new	label
20	La Compagnie des Vins Surnaturels	53506fd9498ebca706a7c170	Wine Bar	40.720448	-73.997969	421	above avg	bars	3
22	Kiki's	5521c2ff498ebe2368634187	Greek Restaurant	40.714476	-73.992036	825	above avg	euro food	3
53	Fish Cheeks	57c169e3498e784947e307aa	Thai Restaurant	40.725861	-73.992533	420	above avg	asia food	3
58	Van Leeuwen Ice Cream	54e7da22498e3dd7390e3136	Ice Cream Shop	40.729840	-74.000433	435	above avg	other	3
75	Sushi Nakazawa	52155e9d11d2c369404df14e	Sushi Restaurant	40.731719	-74.004400	652	above avg	asia food	3
81	Upstate Craft Beer and Oyster Bar	4e41f0822271a90466a0e967	Seafood Restaurant	40.726331	-73.986453	627	above avg	american food	3
83	Via Carota	543c7c63498e0a3393da47a2	Italian Restaurant	40.733052	-74.003573	584	above avg	euro food	3
98	Alphabet City Beer Co.	4f5bcf37e4b09a78bfc29e19	Beer Bar	40.723753	-73.979043	693	above avg	bars	3

Map of Clusters for Users