Introduction and Business Problem

Introduction
The city of Hoboken, NJ is relatively small at ~1 square mile but it is packed with restaurants, night life and amazing people. For people that are new to Hoboken, despite its small geographic size, it can be daunting to figure out what restaurants are worth going to and where they are. For people that used to live in Hoboken or are visiting Hoboken, how do you know what the best places are to get something to eat?

Business Problem
For this project, I am going to put on my entrepreneur hat and create a simple guide on where to eat based on Foursquare likes, restaurant category and geographic location data for restaurants in Hoboken. 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 project, I will be utilizing the Foursquare API to pull the following location data on restaurants in Hoboken, NJ:
- Venue Name
- Venue ID
- Venue Location
- Venue Category
- Count of Likes

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

Methodology
The thought process behind this is that likes are a proxy for quality. The more likes there are, the better the restaurant is. This might be incorrect but API call issues (how many I can use for free) holds me back from getting price / rating data. 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 Mexican food or now what type of food might be best to eat in Hoboken if you are new to the area.

I will take the gathered data (see above in Data Acquisition Approach and Data Required sections) 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 Italian food or other

	name	id	categories	lat	lng	total likes	total likes_cat	categories_new	label
42	Hoboken Burrito	49ee57f6f964a5204f681fe3	Mexican Restaurant	40.741822	-74.031881	13	poor	mex southam food	0
43	Townhouse No 620	4bddbf6be75c0f47f171c503	Boutique	40.744128	-74.029070	4	poor	other	0
44	Jefferson's Coffee	5a888bc2c47cf91545daec08	Coffee Shop	40.743368	-74.029135	11	poor	other	0
46	Makai Poke Co	58c470fd37da1d593431c33a	Poke Place	40.742852	-74.029213	23	poor	euro asia indian food	0
47	Chicken Factory	4c41e1aee26920a1981e5fe7	Korean Restaurant	40.742975	-74.029083	16	poor	euro asia indian food	0
49	Midtown Philly Steaks	51070e83e4b0d92d935e04a3	Sandwich Place	40.742983	-74.029211	23	poor	other	0
51	Mr Wrap's	4ad89c0bf964a520d31221e3	American Restaurant	40.746308	-74.030003	21	poor	american food	0
64	Flatbread Grill	55d52901498ea18f871d5f9e	Restaurant	40.742767	-74.029281	10	poor	american food	0
66	Pita Pit	56f136ad498ecc5661aa49ce	Sandwich Place	40.745729	-74.028582	11	poor	other	0
70	Sauced	5ac2cec0419a9e0f679671f3	Sandwich Place	40.739274	-74.030205	9	poor	other	0
75	Napoli's Pizzeria	59fd168ec21cb1401894e47e	Pizza Place	40.739404	-74.034909	13	poor	italian food	0
78	Hidden Grounds Coffee	5b5b1e0bd4cc9800394e3256	Coffee Shop	40.745117	-74.030493	10	poor	other	0
80	Bangkok City	4acce0aaf964a520d8c920e3	Thai Restaurant	40.741110	-74.029734	20	poor	euro asia indian food	0
82	Frankie & Ava's Italian Eatery	567b01d8498e9310602c8d10	Italian Restaurant	40.739127	-74.030450	24	poor	italian food	0
83	Mario's Classic Pizza Café	4b44d83ef964a520f4fd25e3	Pizza Place	40.746364	-74.030152	20	poor	italian food	0
93	H&S Giovanni's	4af5d847f964a520aefd21e3	Pizza Place	40.743508	-74.029017	18	poor	italian food	0

Cluster 2
- Characteristics
  - o Below average quality food
  - o Mostly Europe / Asia inspired food

	name	id	categories	lat	lng	total likes	total likes_cat	categories_new	label
8	Kung Fu Tea	57168865498e9517f09fa03d	Bubble Tea Shop	40.743375	-74.029450	28	below avg	other	1
10	Karma Kafe	582dfc9565be5809f6a964ed	Indian Restaurant	40.742373	-74.029376	35	below avg	euro asia indian food	1
18	Cafe Michelina	4a7b5b6bf964a520c8ea1fe3	Italian Restaurant	40.742278	-74.030218	35	below avg	italian food	1
33	Moran's Pub	4ad12c5ef964a5203ddd20e3	Bar	40.742553	-74.031192	42	below avg	bars	1
35	Tutta Pesca	57f83f7acd10164c2ec1956f	Seafood Restaurant	40.740163	-74.031284	28	below avg	italian food	1
36	Ultramarinos	4c03f2fe39d476b0f5c530a7	Latin American Restaurant	40.740620	-74.033490	32	below avg	mex southam food	1
50	Mamoun's Falafel	5a6b6047f427de038c51031c	Falafel Restaurant	40.740090	-74.030374	31	below avg	euro asia indian food	1
58	Maroon	4a9703def964a520fb2720e3	Tea Room	40.745307	-74.032359	30	below avg	other	1
65	Dunkin' Donuts	4b191a58f964a520ffd723e3	Donut Shop	40.744813	-74.028812	39	below avg	other	1
74	Sri Thai	49e27ed2f964a5201c621fe3	Thai Restaurant	40.740103	-74.030955	44	below avg	euro asia indian food	1
79	Louise & Jerry's	4283ee00f964a52091221fe3	Dive Bar	40.740880	-74.029480	45	below avg	bars	1
81	Trattoria Saporito	4a9c695cf964a520153720e3	Italian Restaurant	40.740996	-74.029827	31	below avg	italian food	1
89	QDOBA Mexican Eats	49eabe84f964a52090661fe3	Mexican Restaurant	40.741413	-74.029926	31	below avg	mex southam food	1
97	Farside	49e8d1a7f964a52097651fe3	Bar	40.743198	-74.029145	29	below avg	bars	1
99	Pho Nomenon	529fc55011d2e9632e382ae3	Vietnamese Restaurant	40.742733	-74.029298	29	below avg	euro asia indian food	1

Cluster 3

- Characteristics
  - High quality food
  - Mostly Mexican and South American inspired food

	name	id	categories	lat	lng	total likes	total likes_cat	categories_new	label
1	Onieal's Restaurant & Bar	45e9482df964a52075431fe3	Pub	40.741608	-74.032304	151	great	bars	2
13	Court Street Bar & Restaurant	4a7eff1cf964a5206ff21fe3	American Restaurant	40.743322	-74.028615	98	great	american food	2
15	Empire Coffee & Tea	49f37b88f964a520a26a1fe3	Coffee Shop	40.741375	-74.030515	131	great	other	2
16	Mamoun's Falafel	4d9368407b5ea1437d14c8b8	Falafel Restaurant	40.742303	-74.029465	268	great	euro asia indian food	2
17	Zack's Oak Bar & Restaurant	49f26862f964a520296a1fe3	American Restaurant	40.740640	-74.033826	116	great	american food	2
19	Grimaldi's	4ca50f407334236a60ef1258	Pizza Place	40.741674	-74.029578	136	great	italian food	2
20	Robongi	4cdb36c1958f236a15a7ab03	Sushi Restaurant	40.742879	-74.029280	81	great	euro asia indian food	2
22	Bareburger	53ed3b37498e4151087521a9	Burger Joint	40.742694	-74.029070	81	great	american food	2
23	Cucharamama	49e2a407f964a52045621fe3	South American Restaurant	40.740807	-74.034582	101	great	mex southam food	2
24	Cork City Pub	4d4218cd607b6dcb31df08c6	Pub	40.740105	-74.030868	179	great	bars	2
27	The Cuban	4eb1b6859adfb95b77765bf9	Cuban Restaurant	40.741012	-74.029732	163	great	mex southam food	2
29	Benny Tudino's	4a9ac1b1f964a520813220e3	Pizza Place	40.744198	-74.028896	162	great	italian food	2
59	Chipotle Mexican Grill	5065c556e4b0a44a76b324ce	Mexican Restaurant	40.739599	-74.029981	103	great	mex southam food	2
60	East LA	49e37784f964a52083621fe3	Mexican Restaurant	40.742509	-74.029496	108	great	mex southam food	2
61	The Stewed Cow	5131669e582f0b9471213351	Bar	40.742461	-74.036385	119	great	bars	2

Cluster 4

- Characteristics
  - Above average quality food
  - Mostly Bars

	name	id	categories	lat	lng	total likes	total likes_cat	categories_new	label
4	Grand Vin	56d3b920498ec4e1c67c0907	Cocktail Bar	40.743209	-74.035099	62	avg avg	bars	3
6	O'Bagel	56daf06fcd107605ef3d86ea	Bagel Shop	40.743603	-74.029173	60	avg avg	other	3
7	Ayame Hibachi & Sushi	4dbc9859f7b1ab37dd636d12	Sushi Restaurant	40.743105	-74.029213	66	avg avg	euro asia indian food	3
25	Ben & Jerry's	49dfb562f964a52001611fe3	Ice Cream Shop	40.741430	-74.029484	62	avg avg	other	3
34	Zafra	4a453f9cf964a520f4a71fe3	Cuban Restaurant	40.740659	-74.033802	68	avg avg	mex southam food	3
37	Dozzino	4c60c4a1de6920a111ed9664	Pizza Place	40.744612	-74.035632	64	avg avg	italian food	3
38	Otto Strada	527f3d1711d2f7f001c656b2	Italian Restaurant	40.746604	-74.031161	50	avg avg	italian food	3
39	Illuzion	4a9578dff964a520562320e3	Japanese Restaurant	40.741229	-74.029582	76	avg avg	euro asia indian food	3
40	Bluestone Lane	58e7ed715f67173549fe6246	Coffee Shop	40.741601	-74.029410	63	avg avg	other	3
45	Vito's Italian Deli	4a8da189f964a520501020e3	Sandwich Place	40.746401	-74.028310	71	avg avg	other	3
53	Margherita's	4a775512f964a5202ee41fe3	Italian Restaurant	40.746034	-74.028412	71	avg avg	italian food	3
54	Mikie Squared Bar & Grill	4a3ad481f964a52057a01fe3	Sports Bar	40.744032	-74.028956	68	avg avg	bars	3
76	Hoboken Hot Bagels	4b76e28df964a52021672ee3	Bagel Shop	40.744448	-74.028822	48	avg avg	other	3
85	Finnegan's Pub	522cb84211d20c34ef82194f	Pub	40.746477	-74.032032	65	avg avg	bars	3
87	Willie McBride's	49cf90f2f964a520ae5a1fe3	Bar	40.745133	-74.034497	72	avg avg	bars	3

Map of Clusters for Users