

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
 - Poor quality food
 - Mostly Italian food or other

	name	id	categories	lat	lng	total likes	total likes_cat	categories_new	label
2	Grand Vin	56d3b920498ec4e1c67c0907	Cocktail Bar	40.743209	-74.035099	72	avg avg	bars	3
11	O'Bagel	56daf06fcd107605ef3d86ea	Bagel Shop	40.743603	-74.029173	69	avg avg	other	3
12	Ayame Hibachi & Sushi	4dbc9859f7b1ab37dd636d12	Sushi Restaurant	40.743105	-74.029213	71	avg avg	euro asia indian food	3
28	Bluestone Lane	58e7ed715f67173549fe6246	Coffee Shop	40.741601	-74.029410	71	avg avg	other	3
30	Moran's Pub	4ad12c5ef964a5203ddd20e3	Bar	40.742553	-74.031192	46	avg avg	bars	3
31	Ben & Jerry's	49dfb562f964a52001611fe3	Ice Cream Shop	40.741430	-74.029484	65	avg avg	other	3
33	Otto Strada	527f3d1711d2f7f001c656b2	Italian Restaurant	40.746604	-74.031161	57	avg avg	italian food	3
34	Illuzion	4a9578dff964a520562320e3	Japanese Restaurant	40.741229	-74.029582	76	avg avg	euro asia indian food	3
35	Dozzino	4c60c4a1de6920a111ed9664	Pizza Place	40.744612	-74.035632	66	avg avg	italian food	3
44	Mikie Squared Bar & Grill	4a3ad481f964a52057a01fe3	Sports Bar	40.744032	-74.028956	71	avg avg	bars	3
51	Margherita's	4a775512f964a5202ee41fe3	Italian Restaurant	40.746034	-74.028412	73	avg avg	italian food	3
58	Sri Thai	49e27ed2f964a5201c621fe3	Thai Restaurant	40.740103	-74.030955	46	avg avg	euro asia indian food	3
62	Hoboken Hot Bagels	4b76e28df964a52021672ee3	Bagel Shop	40.744448	-74.028822	50	avg avg	other	3
66	Louise & Jerry's	4283ee00f964a52091221fe3	Dive Bar	40.740880	-74.029480	47	avg avg	bars	3
70	Finnegan's Pub	522cb84211d20c34ef82194f	Pub	40.746477	-74.032032	69	avg avg	bars	3

Cluster 2

- Characteristics
 - Below average quality food
 - Mostly Europe / Asia inspired food

	name	id	categories	lat	lng	total likes	total likes_cat	categories_new	label
39	Townhouse No 620	4bddbf6be75c0f47f171c503	Boutique	40.744128	-74.029070	4	poor	other	1
45	Mr Wrap's	4ad89c0bf964a520d31221e3	American Restaurant	40.746308	-74.030003	21	poor	american food	1
46	Paper Source	5df2bca6775cb40008715de3	Paper / Office Supplies Store	40.742185	-74.029320	0	poor	None	1
47	Makai Poke Co	58c470fd37da1d593431c33a	Poke Place	40.742852	-74.029213	24	poor	euro asia indian food	1
48	Gong Cha	5c40edaaad910e0039b9fc01	Bubble Tea Shop	40.743034	-74.029034	7	poor	other	1
50	Chicken Factory	4c41e1aee26920a1981e5fe7	Korean Restaurant	40.742975	-74.029083	17	poor	euro asia indian food	1
55	Napoli's Pizzeria	59fd168ec21cb1401894e47e	Pizza Place	40.739404	-74.034909	16	poor	italian food	1
60	Pita Pit	56f136ad498ecc5661aa49ce	Sandwich Place	40.745729	-74.028582	11	poor	other	1
65	francesca's	56c5003ccd10e29c9c30a7f8	Boutique	40.740742	-74.030123	6	poor	other	1
73	Jefferson's Coffee	5a888bc2c47cf91545daec08	Coffee Shop	40.743368	-74.029135	16	poor	other	1
75	Flatbread Grill	55d52901498ea18f871d5f9e	Restaurant	40.742736	-74.028977	10	poor	american food	1
78	Hidden Grounds Coffee	5b5b1e0bd4cc9800394e3256	Coffee Shop	40.745082	-74.030464	17	poor	other	1
79	Bangkok City	4acce0aaf964a520d8c920e3	Thai Restaurant	40.741110	-74.029734	20	poor	euro asia indian food	1
80	Mario's Classic Pizza Café	4b44d83ef964a520f4fd25e3	Pizza Place	40.746364	-74.030152	23	poor	italian food	1
87	Midtown Philly Steaks	51070e83e4b0d92d935e04a3	Sandwich Place	40.742983	-74.029211	23	poor	other	1
90	H&S Giovanni's	4af5d847f964a520aefd21e3	Pizza Place	40.743508	-74.029017	18	poor	italian food	1
92	Soul Curry	5b857f112f97ec002c294c98	Indian Restaurant	40.744077	-74.028975	6	poor	euro asia indian food	1
98	Charrito's	587ab14c88641b157e49b764	Mexican Restaurant	40.742812	-74.029352	19	poor	mex southam 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
4	Karma Kafe	582dfc9565be5809f6a964ed	Indian Restaurant	40.742373	-74.029376	42	below avg	euro asia indian food	2
8	Kung Fu Tea	57168865498e9517f09fa03d	Bubble Tea Shop	40.743347	-74.029379	30	below avg	other	2
15	Cafe Michelina	4a7b5b6bf964a520c8ea1fe3	Italian Restaurant	40.742278	-74.030218	37	below avg	italian food	2
32	Tutta Pesca	57f83f7acd10164c2ec1956f	Seafood Restaurant	40.740163	-74.031284	33	below avg	italian food	2
42	Mamoun's Falafel	5a6b6047f427de038c51031c	Mediterranean Restaurant	40.740090	-74.030374	42	below avg	None	2
54	Dunkin'	4b191a58f964a520ffd723e3	Donut Shop	40.744912	-74.028920	38	below avg	other	2
72	Frankie & Ava's Italian Eatery	567b01d8498e9310602c8d10	Italian Restaurant	40.739127	-74.030450	27	below avg	italian food	2
83	Trattoria Saporito	4a9c695cf964a520153720e3	Italian Restaurant	40.740996	-74.029827	34	below avg	italian food	2
96	Farside	49e8d1a7f964a52097651fe3	Bar	40.743198	-74.029145	29	below avg	bars	2
97	Pho Nomenon	529fc55011d2e9632e382ae3	Vietnamese Restaurant	40.742733	-74.029298	31	below avg	euro asia indian food	2

Cluster 4

- Characteristics
 - Above average quality food
 - Mostly Bars

	name	id	categories	lat	lng	total likes	total likes_cat	categories_new	label
2	Grand Vin	56d3b920498ec4e1c67c0907	Cocktail Bar	40.743209	-74.035099	72	avg avg	bars	3
11	O'Bagel	56daf06fcd107605ef3d86ea	Bagel Shop	40.743603	-74.029173	69	avg avg	other	3
12	Ayame Hibachi & Sushi	4dbc9859f7b1ab37dd636d12	Sushi Restaurant	40.743105	-74.029213	71	avg avg	euro asia indian food	3
28	Bluestone Lane	58e7ed715f67173549fe6246	Coffee Shop	40.741601	-74.029410	71	avg avg	other	3
30	Moran's Pub	4ad12c5ef964a5203ddd20e3	Bar	40.742553	-74.031192	46	avg avg	bars	3
31	Ben & Jerry's	49dfb562f964a52001611fe3	Ice Cream Shop	40.741430	-74.029484	65	avg avg	other	3
33	Otto Strada	527f3d1711d27f001c656b2	Italian Restaurant	40.746604	-74.031161	57	avg avg	italian food	3
34	Illuzion	4a9578dff964a520562320e3	Japanese Restaurant	40.741229	-74.029582	76	avg avg	euro asia indian food	3
35	Dozzino	4c60c4a1de6920a111ed9664	Pizza Place	40.744612	-74.035632	66	avg avg	italian food	3
44	Mikie Squared Bar & Grill	4a3ad481f964a52057a01fe3	Sports Bar	40.744032	-74.028956	71	avg avg	bars	3
51	Margherita's	4a775512f964a5202ee41fe3	Italian Restaurant	40.746034	-74.028412	73	avg avg	italian food	3
58	Sri Thai	49e27ed2f964a5201c621fe3	Thai Restaurant	40.740103	-74.030955	46	avg avg	euro asia indian food	3
62	Hoboken Hot Bagels	4b76e28df964a52021672ee3	Bagel Shop	40.744448	-74.028822	50	avg avg	other	3
66	Louise & Jerry's	4283ee00f964a52091221fe3	Dive Bar	40.740880	-74.029480	47	avg avg	bars	3
70	Finnegan's Pub	522cb84211d20c34ef82194f	Pub	40.746477	-74.032032	69	avg avg	bars	3

Map of Clusters for Users

