



Where is the Best Coffee?

Business Problem

A business owner is looking to open up a coffee museum to pay tribute to coffee history. The owner would like to open the museum in the Mecca of coffee to ensure great coffee shops surround it. The owner has sought out multiple experts' opinions, but she has received numerous recommendations for the museum's local. The owner ultimately wants to open the museum where the best coffee is, and they would like to see how data can find this location. The only limitation the owner would like to limit is that the museum needs to be in the US, and they would like to open it in a major city.

Data Sources

To develop a dataset to answer this question, we will call the Foursquare API to collect data on coffee shops in the US's most populated cities. To achieve this, we will use a data set from <https://simplemaps.com/data/us-cities>, sorted by population, to call for coffee shops across the US, starting with the most populated cities. We will then analyze likes and rating data on these coffee shops to determine the city in which the best coffee shops are. Foursquare users contribute likes and ratings, so they should be a strong indicator of overall sentiment.

City Data

To start, We imported US city data from <https://simplemaps.com/data/us-cities> and sorted cities by population. The top 5 most populated cities in the US are New York, Los Angeles, Chicago, Miami, and Dallas.

	city	state_id	state_name	lat	lng	ll	population	density
0	New York	NY	New York	40.6943	-73.9249	40.6943, -73.9249	18713220.0	10715.0
1	Los Angeles	CA	California	34.1139	-118.4068	34.1139, -118.4068	12750807.0	3276.0
2	Chicago	IL	Illinois	41.8373	-87.6862	41.8373, -87.6862	8604203.0	4574.0
3	Miami	FL	Florida	25.7839	-80.2102	25.7839, -80.2102	6445545.0	5019.0
4	Dallas	TX	Texas	32.7936	-96.7662	32.7936, -96.7662	5743938.0	1526.0

Coffee Shop Data

Once we had city data, we took this geographic data on the most populated cities in the US and called the Foursquare API to find coffee shops within a 1700 meter proximity to the 20 most populated cities.

	id	coffee_shop	coffee_shop_address	cs_lat	cs_lng
0	53dc5f20498ec5526249f15f	Sunrise/Sunset	351 Evergreen Ave	40.693544	-73.922875
1	5818974a38fa5b066b4ee6ff	Hearts Coffee	1207 Broadway	40.692155	-73.926602
2	55e9fe71498e8edb034385a7	Dweebs	1434 Dekalb Ave	40.700507	-73.922795
3	594a997fd48ec17dba25d273	Knickerbocker Bagel	367 Knickerbocker Ave	40.700551	-73.921392
4	57b0cc38498e0f1b80d4e377	Cup of Brooklyn	1071B Gates Ave	40.689544	-73.922596

Data Points for Analysis

We then retrieved Price tier, rating, and like counts data on the delivered dataset, removing shops that did not have data. Descriptors for these datapoints are below as well as the resulting dataset.

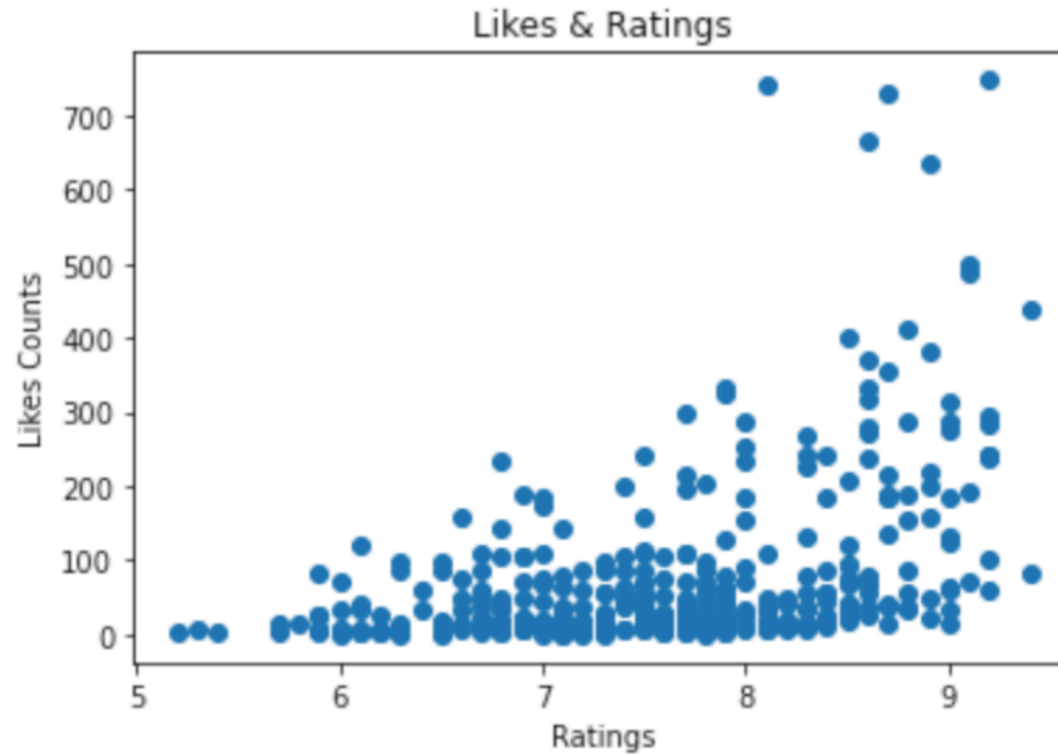
- **Price Tier:** descriptor of priciness of coffee shops on a 1 - 3 scale.
- **Rating:** overall rating of quality on a 1 – 10 scale.
- **Likes:** number of likes visitors have given the coffee shop in Foursquare's history.

Final Data Set

0	53dc5f20498ec5526249f15f	Sunrise/Sunset	351 Evergreen Ave	40.693544	-73.922875	3.0	8.5	119.0
1	5818974a38fa5b066b4ee6ff	Hearts Coffee	1207 Broadway	40.692155	-73.926602	1.0	8.7	39.0
2	55e9fe71498e8edb034385a7	Dweebs	1434 Dekalb Ave	40.700507	-73.922795	1.0	8.8	86.0
3	594a997fd48ec17dba25d273	Knickerbocker Bagel	367 Knickerbocker Ave	40.700551	-73.921392	1.0	9.0	63.0
4	57b0cc38498e0f1b80d4e377	Cup of Brooklyn	1071B Gates Ave	40.689544	-73.922596	1.0	8.2	42.0
...
379	561bbabd498e0094cb0e6a72	Misfit Coffee	2401 Lyndale Ave S	44.958956	-93.287652	1.0	7.7	5.0
380	5c8f9e3cf870fd002c853ce1	Caribou Coffee	50 S 10th St Ste 230	44.972964	-93.275810	1.0	7.2	0.0
381	4f901462e4b0d414c372ffe0	Starbucks	1000 Nicollet Mall	44.973337	-93.275716	1.0	7.2	5.0
382	5ac606f761e53b3252d39ed8	Starbucks	800 E 26th St	44.954472	-93.260524	1.0	7.1	0.0
383	438e8cedf964a520392b1fe3	Panera Bread	809 Nicollet Mall	44.975469	-93.273207	1.0	7.3	57.0

384 rows × 8 columns

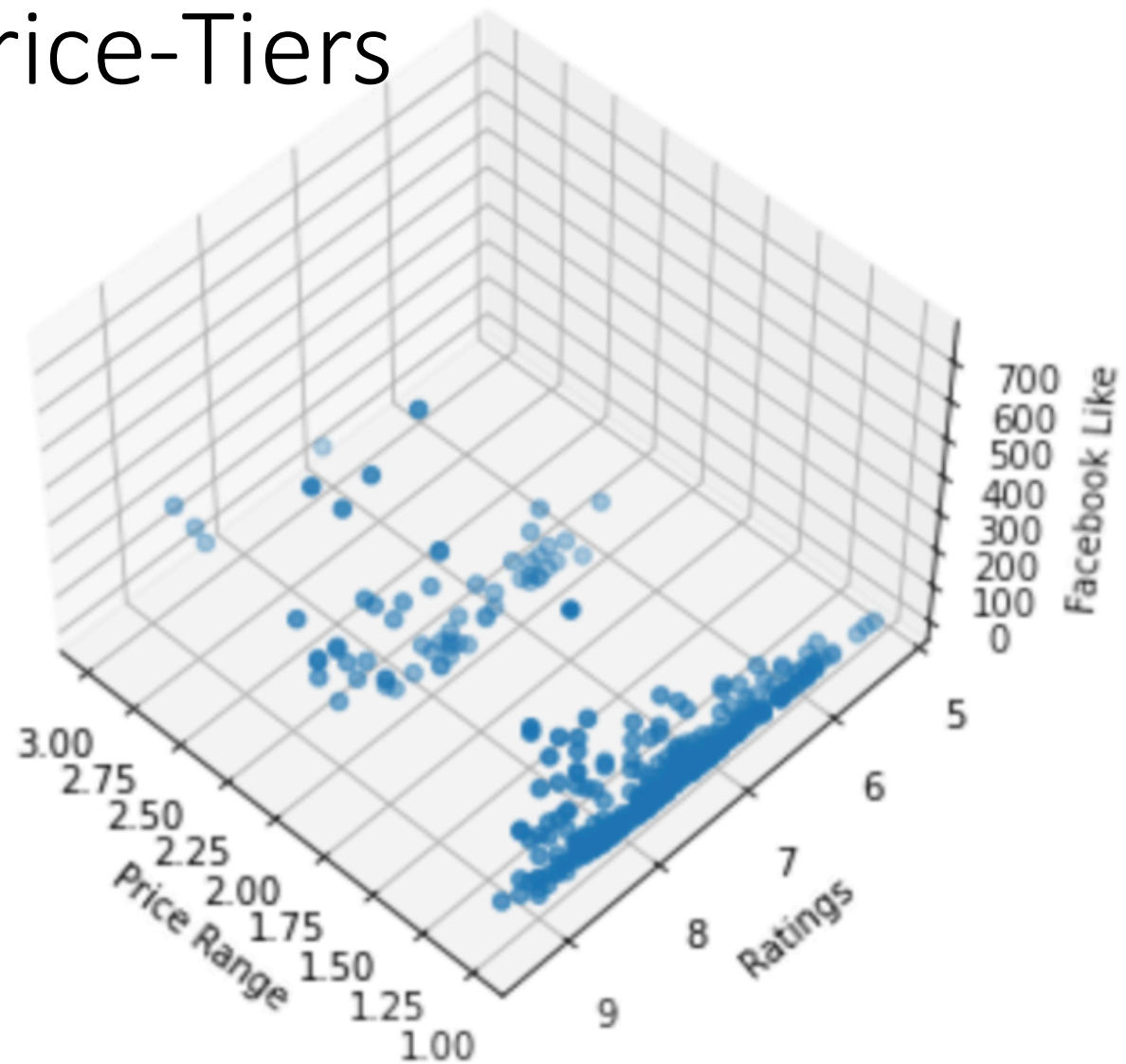
Ratings and Likes



From initial analysis, there appeared to be a loose linear correlation between likes and ratings. The highest rated coffee shops also had the most likes. Most coffee shops had less than 100 likes, with the highest rated ones have upwards of 700.

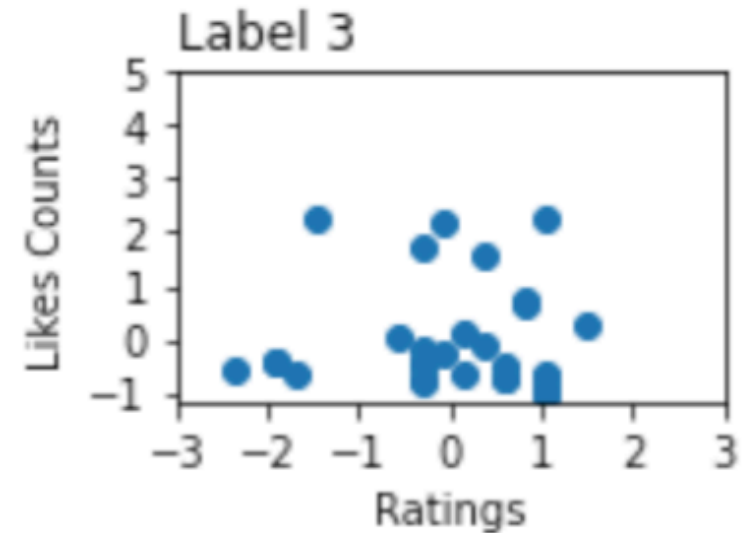
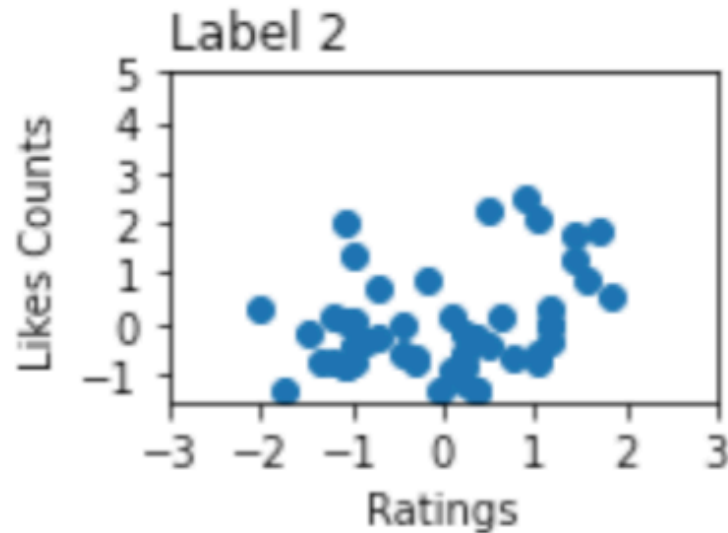
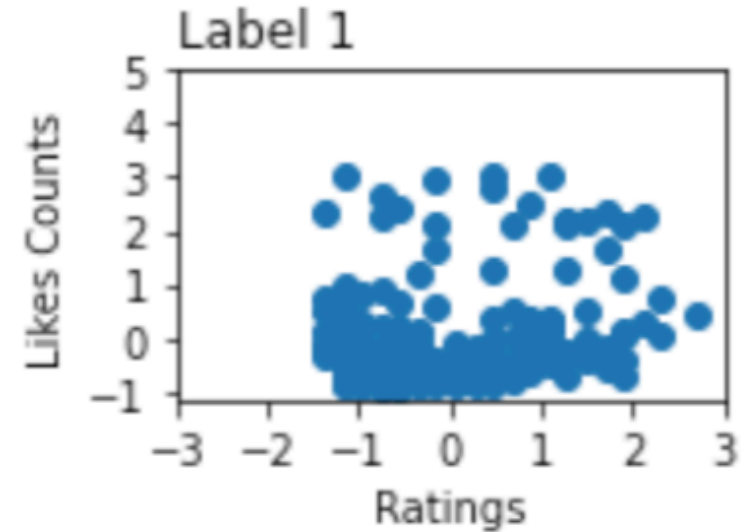
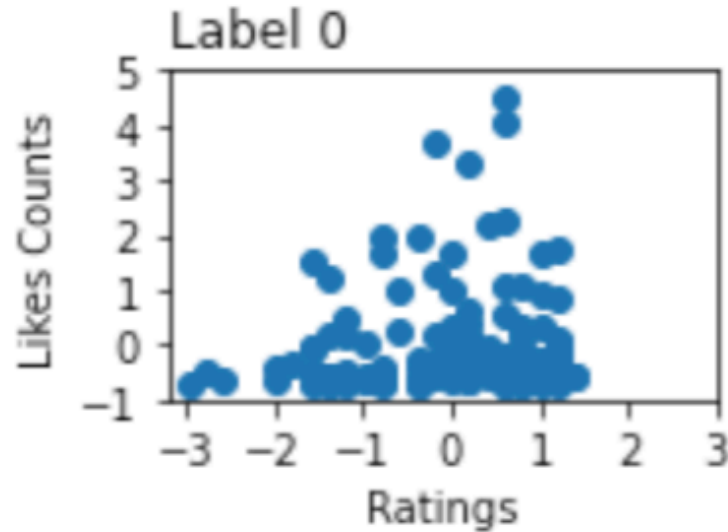
Ratings, Likes, and Price-Tiers

When bringing price range into consideration, We found that the majority of coffee shops were low price, but there was no significant correlation between price and ratings. A level 3 price tier was not predictor of rating or likes.

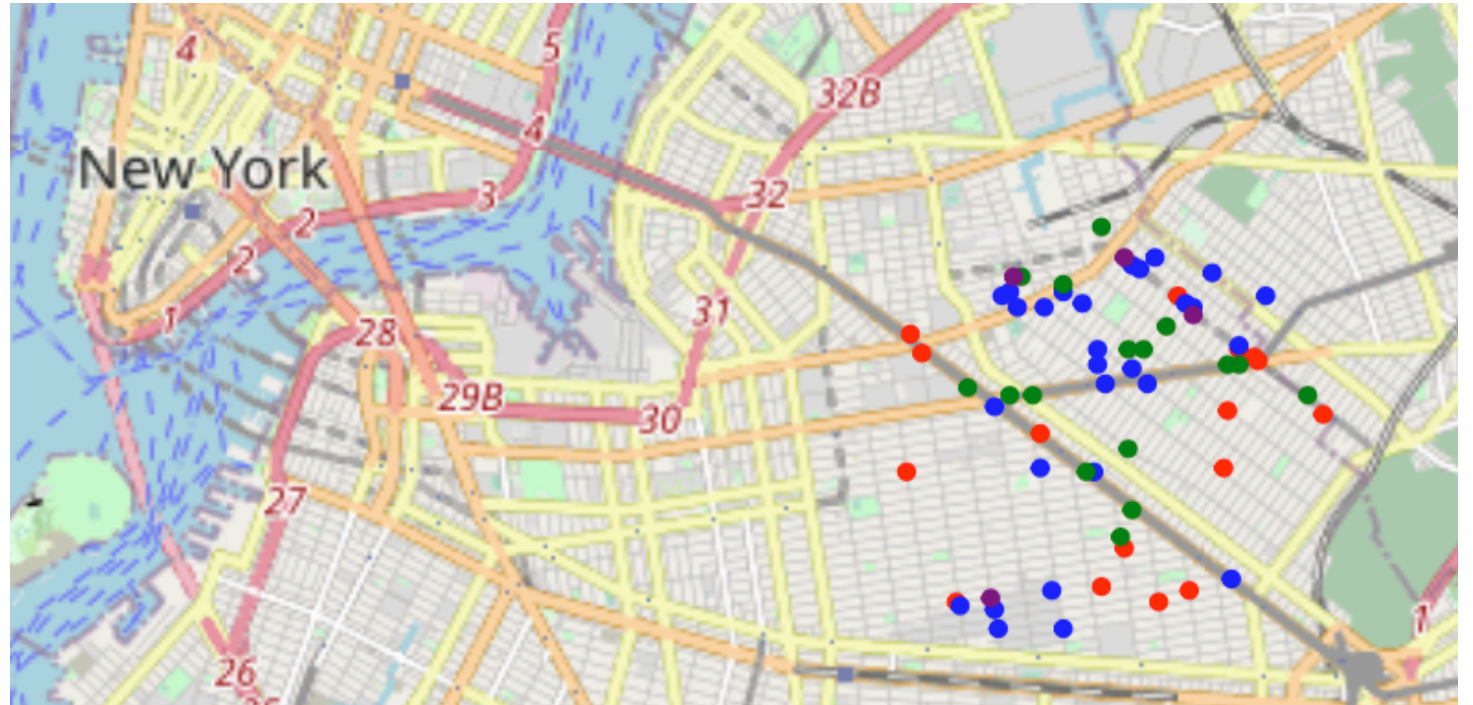


Clusters

Exploring the data using k-means, 4 clusters were identified across price tier, prating, and likes. Looking across clusters, the cluster labeled 1 showed the highest rated coffee shops at a medium price tier.



Results



The major city with the most coffee shops in cluster 1 was New York, followed by Washington and Chicago. This leads us to the recommendation that the best place to open the museum would be in New York on Myrtle Avenue.