New York restaurant consumption analysis

Introduction

Research from Cornell University shows that 27% of new restaurants close in the first year, while nearly 60% close in three years. A key obstacle to success is yelp's lack of exposure. Good restaurants, customers will visit regularly, will have a higher number of scores and comments, which will improve the restaurant's position in the search ranking, and reserve buffer space for any bad luck. This project will discuss the status of New York restaurants.

Data

Our data is a collection of foursqure and yelp from restaurants in New York. The data includes the following items:

categories	coordinate s	display_ph one	distance	id	image_url	is_closed	location	name	phone	price	rating	review_co unt	transactio ns	url	cat
[{'alias': 'shanghain ese', 'title': 'Shanghai 0 n...	{'latitude': 40.714669 1442891, 'longitude' : -7...	(212) 233- 8888	1066.509	joes- shanghai- new-york- 2	https://s3- media3.fl.y elpcdn.co m/bphoto/ ekUo5e...	FALSE	{'address1' : '9 Pell St', 'address2': '', 'add...	Joe's Shanghai	1.21E+10	$$		4	4696	[]	https://ww w.yelp.co m/biz/joes -shanghai- new-yor...
[{'alias': 'chinese', 'title': 'Chinese'}, 1 {'a...	{'latitude': 40.7142, 'longitude' : - 73.99787}	(212) 791- 1817	1020.196	tasty- hand- pulled- noodles- new-york	https://s3- media1.fl.y elpcdn.co m/bphoto/ 3_3or7...	FALSE	{'address1' : '1 Doyers St', 'address2': '', 'a...	Tasty Hand- Pulled Noodles	1.21E+10	$		4	1280	['delivery', 'pickup']	https://ww w.yelp.co m/biz/tast y-hand- pulled- noo...
[{'alias': 'chinese', 'title': 'Chinese'}, 2 {'a...	{'latitude': 40.742227 5387194, 'longitude' : -7...	(212) 989- 6699	4183.124	buddakan- new-york	https://s3- media4.fl.y elpcdn.co m/bphoto/ Avw19V...	FALSE	{'address1' : '75 9th Ave', 'address2': None, '...	Buddakan	1.21E+10	$$$		4	3278	['restauran t_reservati on']	https://ww w.yelp.co m/biz/bud dakan- new- york?adj...
[{'alias': 'shanghain ese', 'title': 'Shanghai 3 n...	{'latitude': 40.759881 0063385, 'longitude' : -7...	(718) 321- 3838	14909.39	nan-xiang- xiao-long- bao- flushing	https://s3- media4.fl.y elpcdn.co m/bphoto/ ONP2Dh...	FALSE	{'address1' : '38-12 Prince St', 'address2': ''...	Nan Xiang Xiao Long Bao	1.72E+10	$		4	2664	[]	https://ww w.yelp.co m/biz/nan- xiang-xiao- long-b...

Because there are other special characters in the latitude and longitude information of the data, we need to clean them and readjust them:

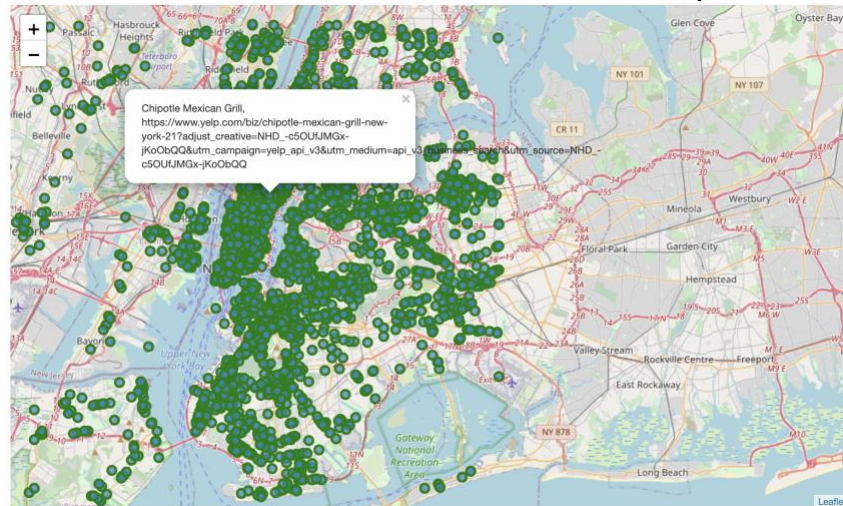|	latitude	longitude
0	40.7146691442891	-73.9977602206654
1	40.7142	-73.99787
2	40.7422275387194	-74.0047066478157
3	40.7598810063385	-73.8325853625656
4	40.7190148	-73.9968228
...	...	...
5496	40.8603060990572	-73.8908256590366
5497	40.7156225	-74.0073923
5498	40.6357082	-74.1349684
5499	40.81591	-73.9168399
5500	40.74225	-73.87199

Methodology

Yelp's Restaurant ranking is globally recognized and is one of the best examples of using public experience and comments. The answer we want to seek is: what are the key characteristics that

affect the success of restaurant management? We believe that from the official data related to scoring provided by yelp, we will be able to identify important features that can be identified.
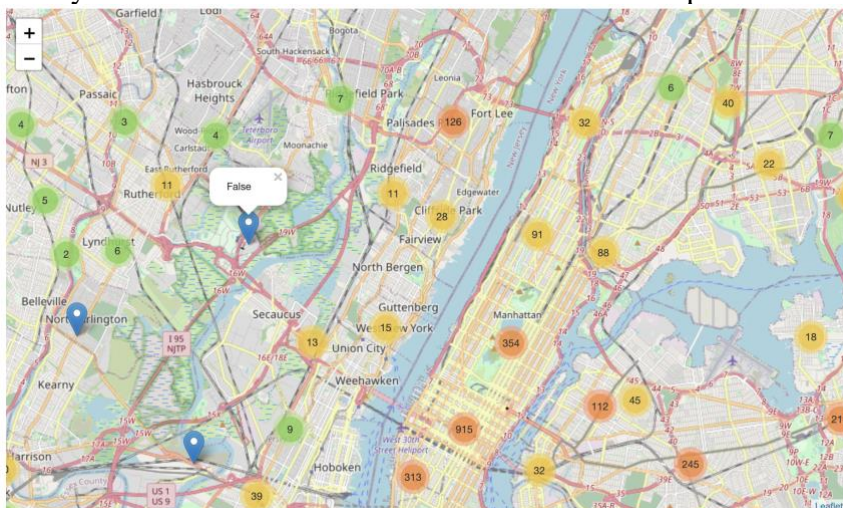
These key characteristics can be the inherent attributes of business operation, such as opening time, noisy environment, or subjective factors of customers. Through data visualization and prediction model, we explore the key features of yelp to achieve high scores.
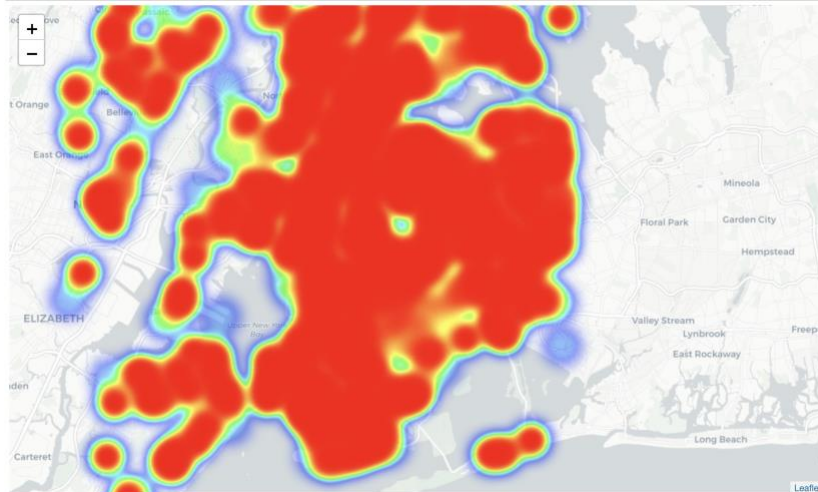
Results

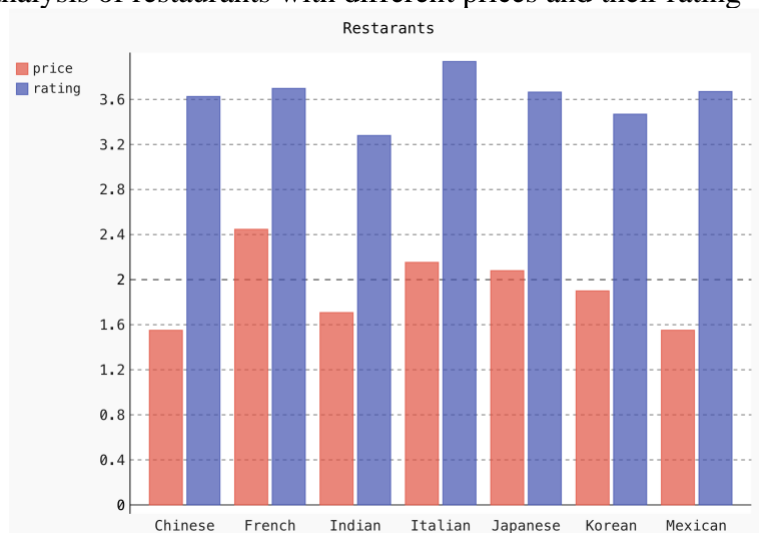Restaurant distribution: label the restaurant name and URL on the map

Business status: analyze whether restaurants in different blocks are open
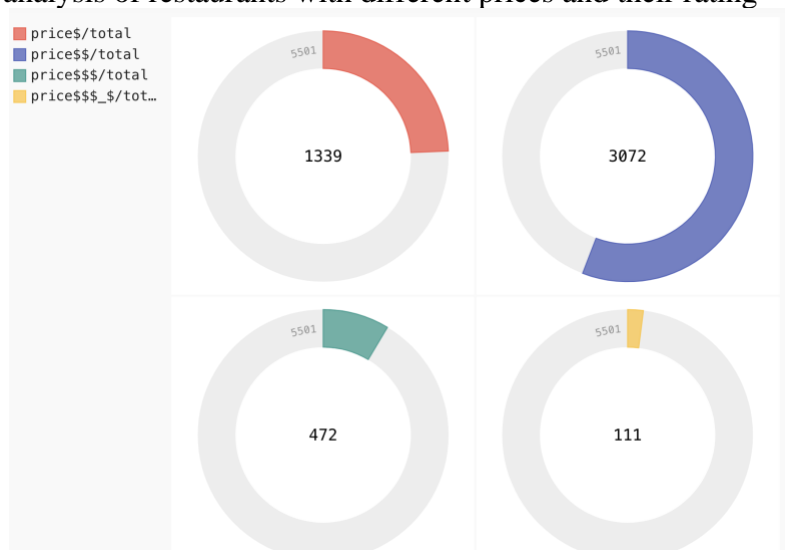

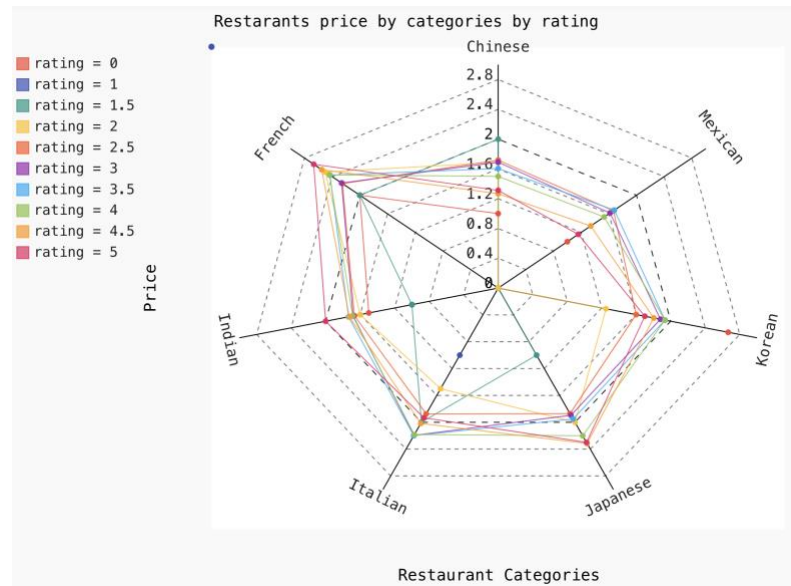Regional distribution of restaurant popularity: reviews counts heat map

Price and rating: analysis of restaurants with different prices and their rating

Price and rating: analysis of restaurants with different prices and their rating

Cuisine Classification Library: Analyze the types and ratings of different types of cuisines

Restarants price by categories by rating

Discussion

Cuisine Classification Library: Analyze the types and ratings of different types of cuisines - The general trend is that better rating leads to higher price. in Chinese and Mexican categories, high priced restaurants are poorly rated.

Conclusion

Conclusion for the price and cuisines: on average, Chinese restaurants have the lowest prices and French restaurants have the highest prices. The reviews of Indian restaurants are relatively poor, while the reviews of Italian restaurants are relatively high. In addition, there is no obvious relationship between price and rating. In addition, if people want reasonably priced and well-reviewed food, Chinese and Mexican food is recommended.