Data Capture/NLP Project on Rochester Restaurants using YELP.

Introduction

As an outsider to Rochester and an avid foodie, I wanted to identify the happening places in Rochester where I can take my friends and acquaintances to eat and relax. I have used Yelp.com to help me with this analysis. Yelp.com is a well-built website that can display information like names of the restaurants, the number of Yelp reviewers who have checked this place out, the number of ratings, cost per person, cuisines offered by the restaurant, where is it located, and the top review on this place. My goal is to understand local "Rochesterian" food culture and ways to delve into their views.
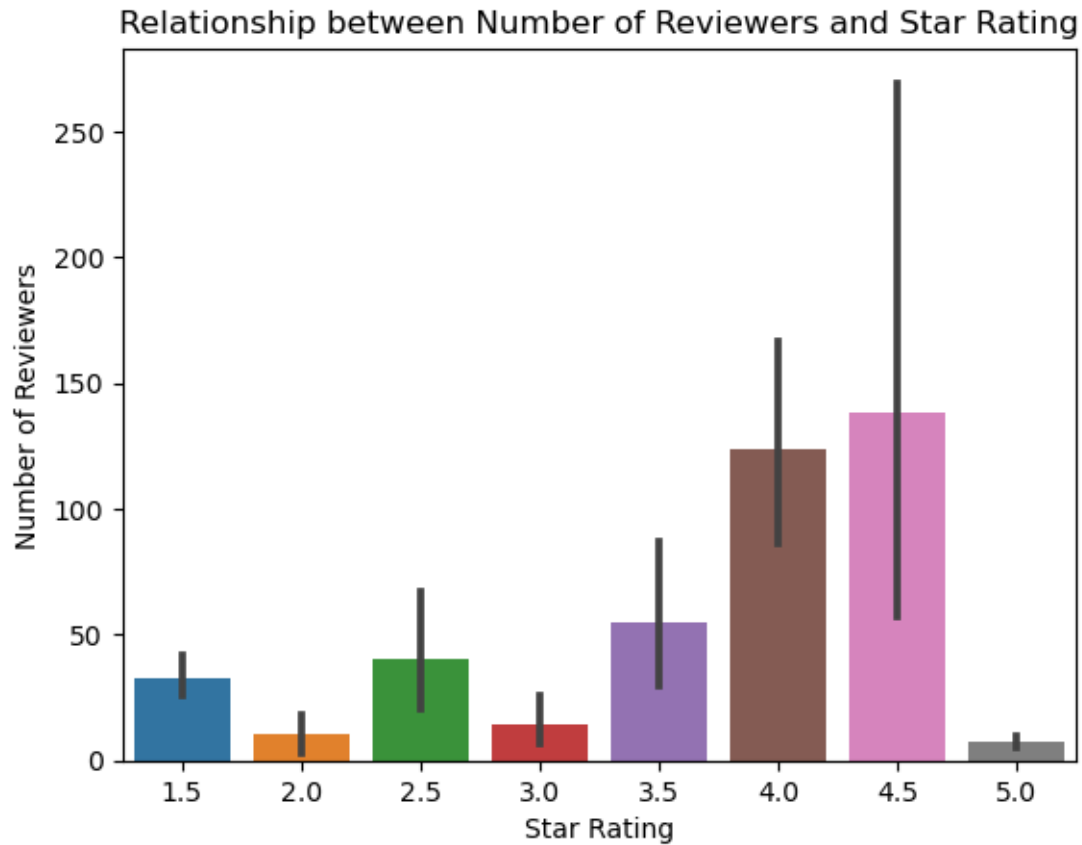
Initial Analysis

I have collected around 123 unique records for this assignment. I have used tools like Beautiful Soup and Selenium(for navigation only) to help me scrape data from their website. I made sure that the authentication was in place and not causing any harm to the website traffic.

I have identified key features that would help me in the initial analysis:-

1. Number of reviewers
2. Rating
3. Cost per person.
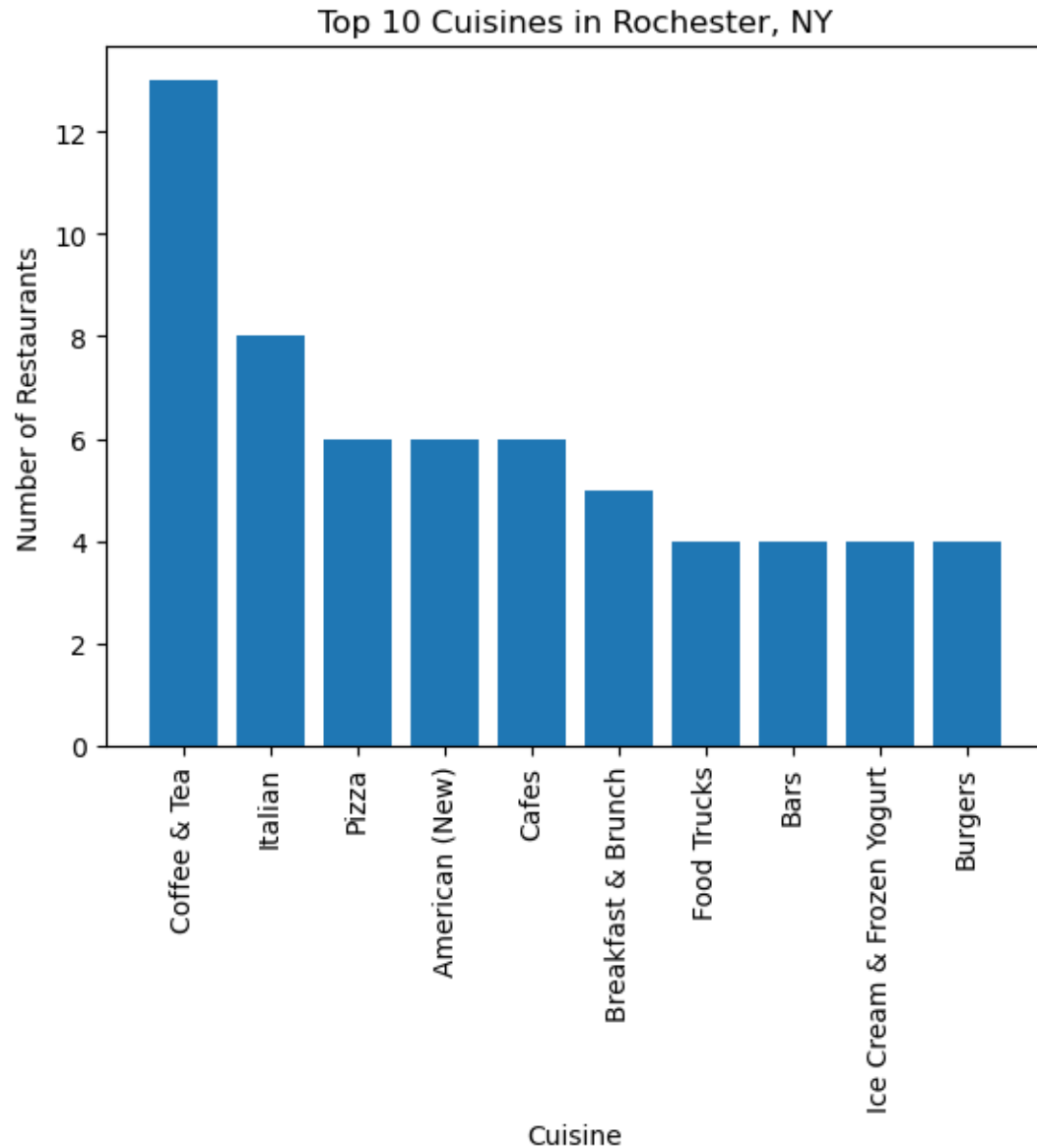4. Cuisines
5. Location

Number of Reviewers vs Rating

Rating impacts the goodwill and popularity of any business. The rating has become an important metric especially, in Gen-X, Gen-Y, and Gen-Z people. A good rating means more people would start visiting the place. More visitors would mean more ratings. If the experience was great, the ratings would end at the higher end of the spectrum. Here, we analyze how did rating impacted the number of reviewers and vice versa.

Relationship between Number of Reviewers and Star Rating

This graph is evenly spread on all ratings from 1.5 to 5.0 We see a large concentration of 4.5 ratings. We see a pattern where the number of reviewers fluctuates to 3.5 and keeps increasing to 4.5 ratings.
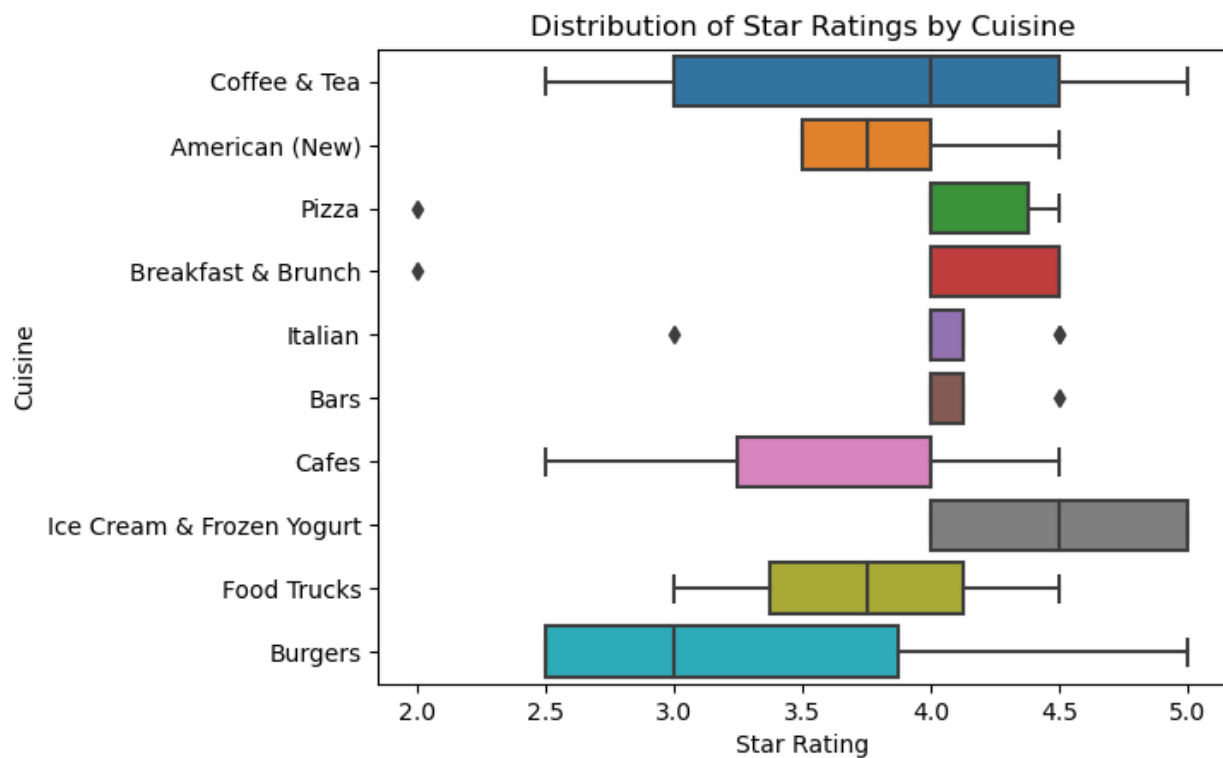
Identify which is the topmost popular cuisine here in ROC.

Top 10 Cuisines in Rochester, NY

Coffee and tea seem to be among the top priorities here in ROC. Although Pizza belongs to the Italian cuisine(We have an American variant too), Yelp shows them they are popular individually. American(New) cuisine is the next most popular cuisine.
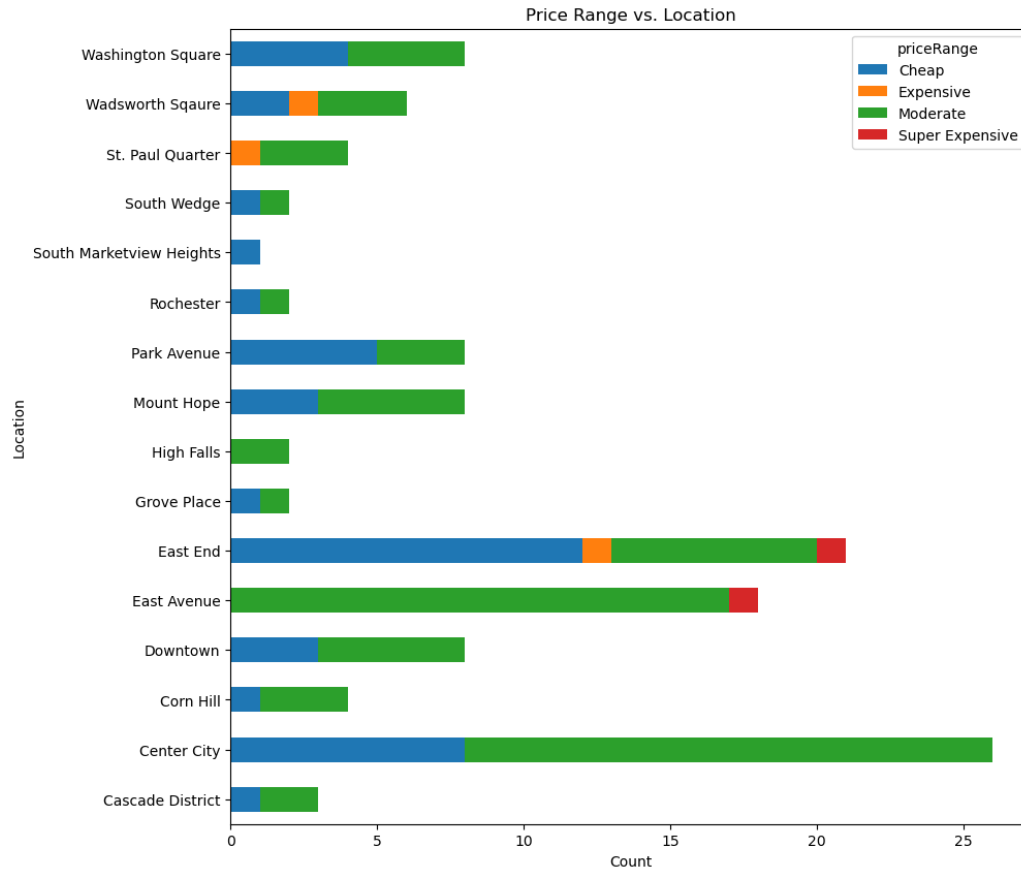
Distribution of star ratings by Cuisine

We can also identify the distribution of the star rating based on each cuisine. Below is the graph drawn to depict the same.

Distribution of Star Ratings by Cuisine

We can see there is a smaller number of observations for Italian restaurants to comment on. Coffee and tea are one such cuisine that is widely distributed.

Price range vs location

Price Range vs. Location

Center City is the most popular destination for eateries, and it is a budget-friendly area among all of them. East-End and East Avenue have the most expensive places of all the other places in Rochester. Wadsworth Square and Washington Square have a few expensive places as well. South MarketView Heights has the least places to eat.

Word cloud using Word2VEC.

We use a tool called Word2vec to create beautiful plots to identify interesting word patterns called Word Clouds. Word Clouds are a good way to display the most popular cuisines or to highlight the most popular sentiment people have.

Here is my output of word cloud I have:-

Ice cream seems to be a popular word. It also coincides with the most popular cuisines. People use the word "**great service**" to describe the place doing a good job in the serving section. **Amazing**, **well**, and **delicious** gets highlighted in this word cloud.

<u>Identify if there is any similarity between the top reviews!</u>

Although we don't see any highlighting patterns, we can say confidently that there are at least two reviews that are like each other based on the proximity of the data points.

Conclusion

➜ Coffee & Tea, Italian, and American(New) are the most common cuisines among the local Rochester crowd.

➜ There are decent places with 4.0 and 4.5 if you feel like exploring.

➜ We need to select burger places carefully as there are some not-good burger places around.

➜ Center City, East End, and East Avenue are happening areas in Rochester.

➜ You can go to Center City if you are on a budget.

➜ East End and East Avenue have some expensive places to explore.

➜ **Great Service** and **Amazing Place** are some of the common words people have used to describe these restaurants.

Closing thoughts

Some of the drawbacks of this project are that the number of restaurants on Yelp was quite limited. This analysis can be used as a food guide blog to help choose places for occasions or quiet evening dinners.