Improve Restaurants Rating/Performance on Yelp Final Write-up

Team Members: Zhangcheng Guo, Jiazheng Li, Xuanqi Liang, Ying Zhang

Business Background

Yelp is a leading application in sharing restaurant information. However, there are still improvements Yelp can make to maximize its purpose. We are particularly interested in how restaurant owners on Yelp can utilize the app's features and achieve higher business values. Our goal is to provide valuable suggestions and improve their ratings by predicting restaurant stars rating and analyzing user reviews.

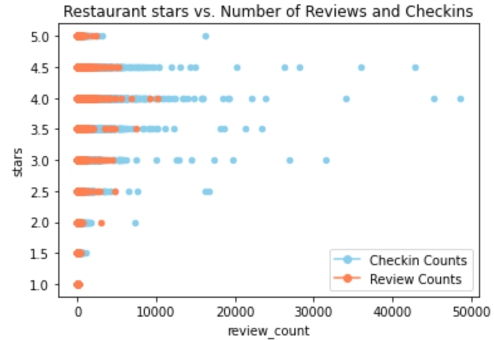
Dataset Cleaning

Source Link: https://www.kaggle.com/yelp-dataset/yelp-dataset

Our dataset is acquired from Kaggle, originated from Yelp Dataset Challenge, records information about businesses, user reviews, tips on Yelp. Datasets are original, therefore, messy. Due to equipment limitations, we have filtered only companies containing "restaurant" in category columns, attaining 49947 business information. The "Attribute" feature is extracted from JSON format and converted to columns for each feature, a total of 56 features. We have dropped them for machine learning analysis for those features containing more than 50% of missing values (mostly binary features describing business infrastructures). All missing values are assumed to be False for machine learning purposes for the rest of the binary variables recording information such as parking, breakfast, etc.

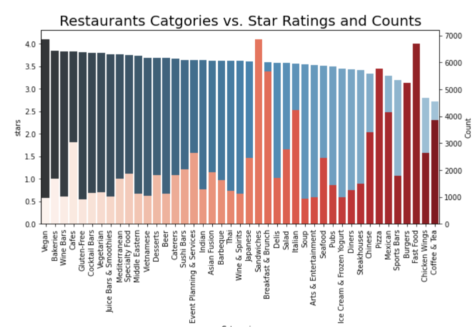
Exploratory Data Analysis

We compared restaurant star ratings with business attributes to observe what affects ratings.
Business attributes record and various information on infrastructure, restaurant attires, style, etc. Below are some highlighted findings.



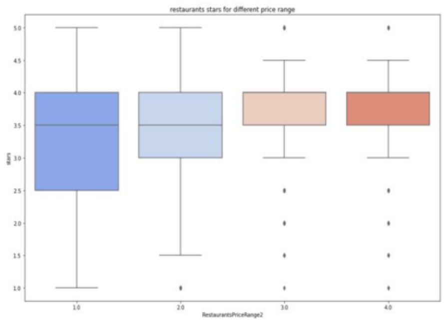
The scatter plot records review counts and check-in counts compared with stars for each restaurant. It is observed that restaurants around 3.5 - 4.5 stars record the highest review counts.

From the scatter plot of restaurant stars with average review stars and the trend line, we can notice the positive linear relationship between those two variables.



From the bar plot of the top 43 category count (red bars), We observed these restaurant stars vary from 3.0 - 4.0.

Moreover, Fast food and sandwiches have the top two count numbers here.



From the boxplots of restaurants with different price ranges (1.0 means the lowest and 4.0 means the highest), we noticed that restaurants with a higher range of prices tend to have higher scores on their stars.

Predicting Star Ratings

One way to help restaurants elevate stars is to predict star ratings using machine learning models. We mean to predict stars and observe essential features that affect lead.

The graph shows list of input features for all machine learning models.

```
f1 <- as.formula(stars ~ review_count + RestaurantsTakeOut +
BusinessAcceptsCreditCards + NoiseLevel + RestaurantsPriceRange2 + Alcohol + trendy
+ classy + lunch + dinner + brunch + RestaurantsAttire + GoodForKids +
RestaurantsReservations + RestaurantsGoodForGroups + HasTV + BikeParking +
RestaurantsDelivery + OutdoorSeating + WiFi + Caters)
```

Linear Regression

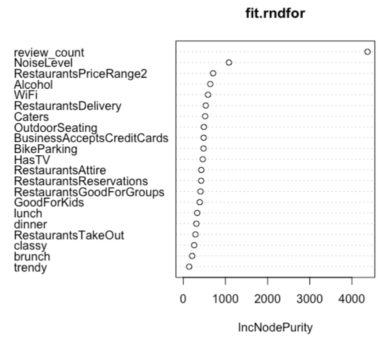
First, we considered the linear regression model to capture essential features with resulting coefficients and p-values. However, the resulting r-square of 0.12 is too low to proceed with further analysis. Therefore, we decide to move on to more flexible models to capture nonlinearity.

Boosting

	MSE	R square
Linear Regression	0.8	0.12
Boosting Forest	0.46	0.51
Random Forest	0.21	0.638

We then used Boosting model and resulted in much better predictions. The performance is much better with an MSE of 0.46, a standard error of 0.7. However, the difference of 0.7 instars could cause a real star difference.

Random Forest



Finally, Random Forest has shown the best results with an MSE of 0.21 and r square of 0.638. With random forest, it will result in 0.46 variance in predicting ratings, but with this model, the restaurant can better understand where they stand in terms of ratings.

To further understand what affects star ratings, we have conducted Mean Decrease Gini (IncNodePurity) to measure feature importance. As shown in the plot on the right, "review_count" displays the highest score, which reflects its importance in predicting stars, followed by "NoiseLevel" and "RestaurnatPriceRange2".

The dataset feature might not be the most predictive when predicting stars, as MSE computed is desired to be smaller. However, we can gain insight into the importance of review counts in restaurant star ratings. Therefore, restaurants might be incentivized to have proportions targeting more customer reviews.

Clustering without star ratings

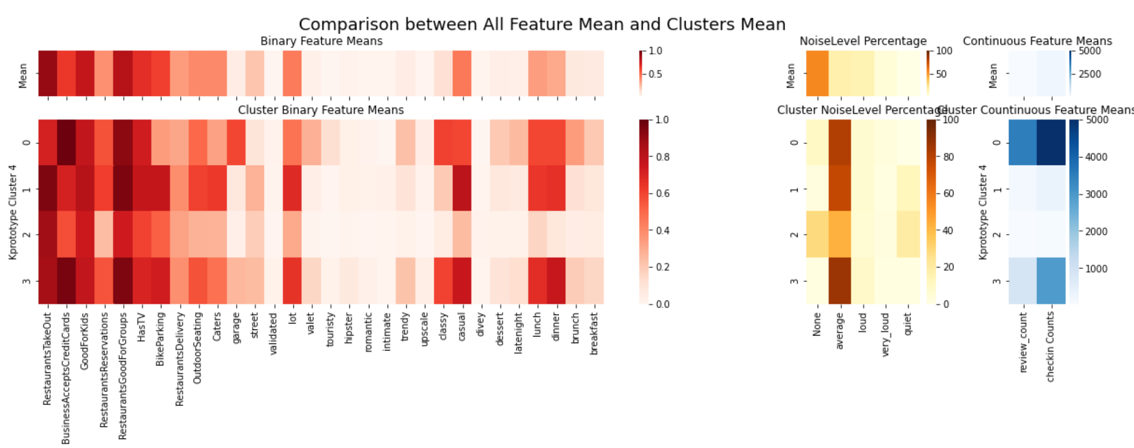
The exploratory analysis observed that stars rated from user reviews are highly correlated with restaurant stars. We suspect review stars are potentially subjective and inconsistent. Therefore, we decided to cluster restaurants without star ratings and purely based on their infrastructures, containing only restaurant features, review counts, and check-in counts. First, we applied umap to roughly observe clustering in two dimensions (shown on the right). There are around 13 clusters observed. Several methods with different cluster groups are conducted, including Hierarchical clustering with Jaccard distances, Kmodes and [Kprototype](). We have applied standard scaling to numerical variables, review count, and check-in counts.

Kprototype Cluster 4	stars	Count
0	3.851562	64
1	3.622035	14754
2	3.385310	26393
3	3.933197	1220

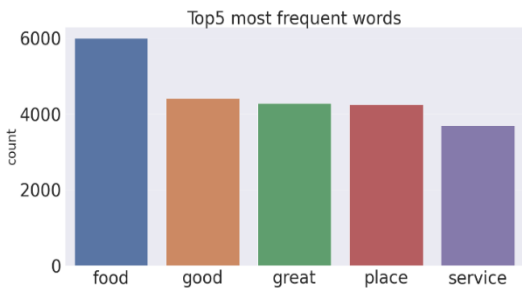
KPrototype with four clusters is chosen as it is the most interpretable and disguisable between restaurant clusters. (Most other clustering methods either produce clusters with one restaurant or displays similar attributes.) The table records average star ratings and restaurant counts for each cluster. Different cluster groups with various attributes do not exhibit significant rating differences. To find what features distinguish clusters, heatmaps are plotted below, showing the overall feature means (upper row) and 4 cluster means (bottom four rows).

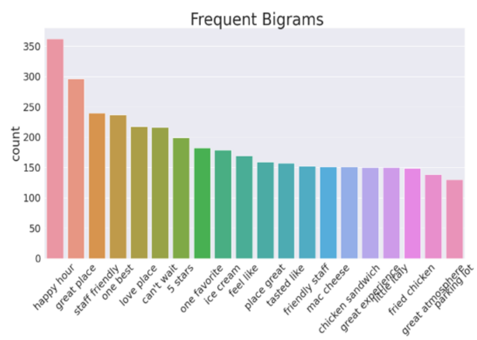
Observing binary variables (red), classiness and casualness varied in four clusters. Whether or not restaurants offer car and bike parking seems to be relevant for clustering. What restaurants serve in terms of lunch and dinner weight high as well in KPrototype clustering. Noise level, which showed an essential feature in Random Forest analysis, did not matter as expected. The percentage distribution between each noise level seemed uniform across clusters. The most apparent clustering features are review count and check-in counts (blue), as cluster 0 obtains higher review and check-in counts, followed by cluster 3. Cluster 3, which brings the highest star ratings, seemed to be casual and classy restaurants offering dinner with moderate review counts. Restaurants in cluster 3 are observed to obtain better services as they provide take-outs, accept credit cards, are better for kids, and are better for groups than other clusters. On the other hand, cluster 2 restaurants with the lowest average stars are much weaker than services. The point regarding restaurant services also resonates with the text analysis explained in later sections.

Clustering without user input has pointed that services offered by restaurants and review counts have differentiated from one restaurant to another. Therefore, we paid particular attention to users' reviews on restaurants for our text analysis, trying to observe what exactly users are looking for.



Text Analysis
PART1: 2019 reviews for all of the 825
Cleveland restaurants





Customers care about food, service, and happy hour

Food: The frequency of tokens analysis reveals that customers most care about food and service. The most common word is "food" which mentioned more than 6000 times in 2019. Also, the term "delicious" has ranked among the Top20 most common words. These seem to indicate that food taste and food quality might still be the most critical factor for evaluating a restaurant.

Service: Customers also seem to care a lot about the quality of services. Other than the word "service", related words including "staff" and "friendly" are also among the hottest words.

Happy Hour and others: Besides, restaurants may want to consider providing a "happy hour" menu as it was brought up by the customers nearly 400 times in reviews. Additionally, people also talked about experience, atmosphere, and parking lot frequently.

Hottest Food trend in Cleveland

From the frequency plot and word clouds analysis, we can observe that the classic fast-food options appeared the most frequently in all the Cleveland 2019 reviews, such as pizza, ice cream, chicken, fried chicken, chicken sandwich, and mac cheese. In terms of drinks, yelp users tend to comment a lot about beer and bars.

Specific food trend:

Healthy Food: Eating healthy is the trend now. More and more people are starting to pay attention to their diet health, so meals such as salad and soup appear highly frequent in reviews. Moreover, when analyzing the most popular restaurant in 2019 – the Townhall, we noticed that healthy meals might be gaining more popularity than we expected, as evidenced by the fact that gluten-free, veggie burger, grass feed, and dietary restriction are among the hottest words in Townhall restaurant' reviews.

A variety of fast-food topics: Based on the topic model, some word that are related to fast food, such as "fried chicken," "sweet potato," "chicken sandwich" appear in the topics generated. A Mexican food stands out along with the keyword "taco".

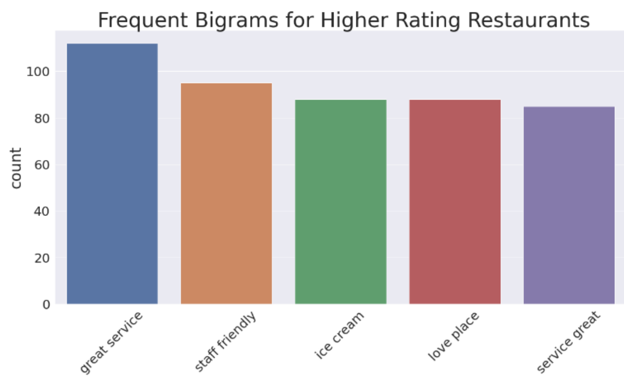
Asian cuisine topic: We can see items such as soup dumplings, ramen/noodles, spicy food, and Korean food appear in the Asian food topic, which indicating that these might be the popular items in Asian cuisine.

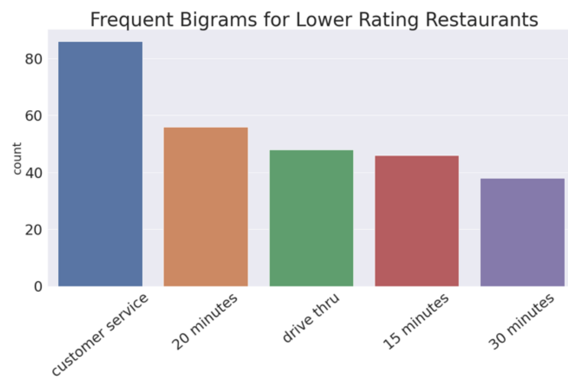
PART2: Analysis on higher rating restaurants V.S. lower rating restaurants

We classified the Cleveland restaurants whose general ratings are higher than four as higher rating restaurants, and those ratings are lower 3.5 as lower rating restaurants. By comparing the two groups, we aim to gain some business insights on differences between higher and lower restaurants and help them operate better.

A. Bigram Frequency Analysis (excluding "food" related bigrams)

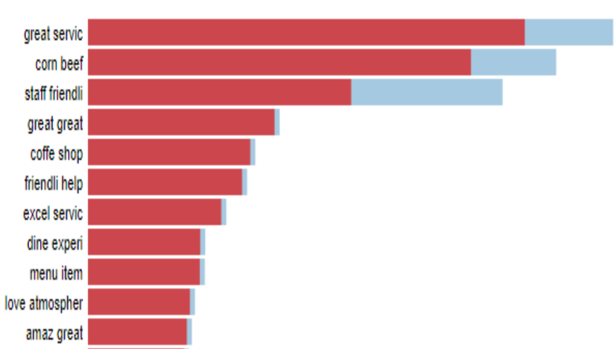
As we've discussed food topics in *PART1*, now we focus on analyzing what other factors are useful for differentiating higher rating restaurants from lower rating restaurants. We can see the quality of customer service is the central aspect. Besides, long waiting times seem to associate with low ratings.



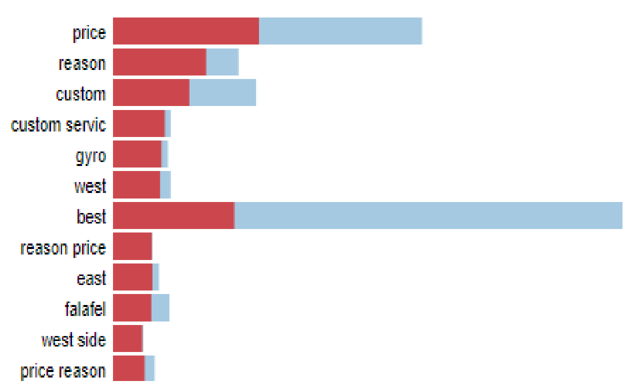


B. What other factors could contribute to high ratings?

"High quality of service, reasonable price, happy hour menu, great atmosphere, and great location tend to attract a lot of positive comments and good ratings, according to our tokens frequency analysis and LDA topic modeling results. It's not hard to imagine friendly staff and reasonable price contribute a lot to customer satisfaction. But there are other insights we can gain from the reviews. For example, live music and outdoor seating space can create a better atmosphere. Westside Market in Cleveland is very popular and seems to be a great place to open a restaurant.



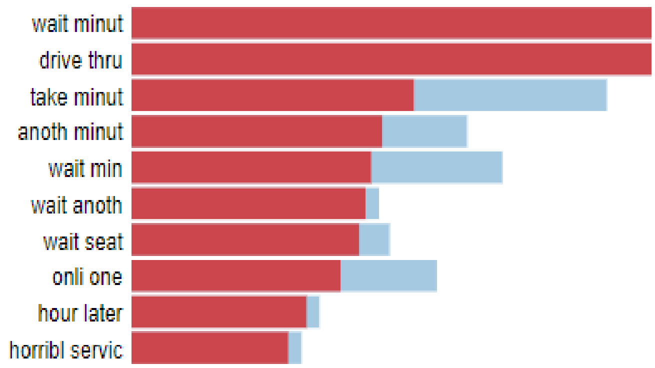
Service-related topic



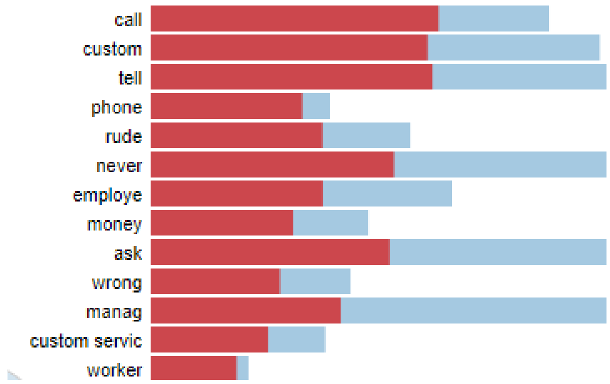
Price-related topic

C. What other factor could lead to low ratings?

Wait Too Long



Staff's Poor communication & Poor Attitude



Gain more context through searching for keywords using the concordance function
Service as keyword:
"Seriously horrible service! There is a lack of restaurant management,"
"Some of the absolute worst service I 've ever had. We waited 50 min"
"The service was extremely slow and our waitress kept ignoring… "
"it ruined my afternoon because he was so rude"
Manager as keyword:
"I was put on hold and told a manager was not available for several hours"
"The manager didn't seem to care about."
"The manager was passive aggressive and made the experience terrible"
"The only positive is the manager was very nice."

Recommendations
Through all the text analyses we performed for the lower rating restaurants, we found out that, in addition to food quality, "long waiting time" and "staff's poor attitude" might the issues that bring down the ratings.

Reduce waiting time
Patrons of those restaurants complained about waiting too long during the process, including staying in lines to get seats, waiting for the meals to be ready, and waiting in the drive-through line. We found out long waiting time has led to many poor service perceptions. Thus, the restaurants must reduce wait times actively. There are several ways that a restaurant may want to consider. For example, 1) pre-schedule staff and make sure there are enough employees during rush hours, 2) accept reservations, 3) create fast-paced workflow and improve table bussing efficiency, and 4) prepare snacks for customers who are waiting.

Improve service skills of the restaurant's managers and its employees
"Rude employees" and "a lack of response" tend to result in many negative comments and lower ratings. And managers' attitude is the key to delivering better services. From the reviews, we noticed that many customers would try to communicate with restaurant managers when they didn't feel good about their dining experiences. If managers can listen to the customers and solve problems immediately, it would help ease customers' anger and disappointments and thus reduce the number of negative reviews. To improve customer service, it's essential for restaurant owners to train employees in effective service techniques, such as explaining "friendly" to the team, training in problem solving skills, and helping the team maintain a positive attitude by valuing the employees.

Conclusion

Eventually, reviewers are looking for Good service, a good atmosphere, and good food. To achieve these “Good”s,

Restaurants can predict star ratings with upgrades in restaurant attributes or increase review counts with a random forest model. Clustering inferred the importance of having restaurant infrastructures. Overall, more useful reviews result in higher restaurant ratings. Most importantly, restaurants should pay special attention to reducing wait time and providing better customer experiences.

While the dataset provided a wide coverage range of topics, it can produce more profound results by analyzing more specific areas like “Vegan food.” We also suggested Yelp add location variables in terms of distance from the city for location analysis. Yelp can also provide star ratings for these three crucial features for restaurants to better understand their performance in each category.