Decoding Dining Delight: Correlating Operational and Service Metrics with Restaurant Ratings

Sherry Yang yyang35@uw.edu

Xinlin He xlhe@uw.edu

Rayan Alghamdi rayan3gh@uw.edu

Masaru Chida mchida@uw.edu

Andrew Ng ngand000@uw.edu

Abstract

In an era where online reviews significantly impact consumer choice, understanding the determinants of restaurant ratings has become vital for the hospitality industry. Addressing this challenge is difficult due to the multifaceted nature of factors influencing customer satisfaction. Recognizing the limitations of current research which often relies heavily on singular datasets, our study aims to cast a wider net for a more comprehensive analysis. This study presents a extensive examination of factors influencing customer ratings of restaurants, spanning both operational aspects and online review characteristics. We investigated a range of hypotheses centered on operational factors such as service options, food preparation time, cost, and operating hours, alongside static attributes like location, cuisine type, parking availability, and restaurant size. Utilizing a robust dataset from a prominent online review platform, our analysis diverges from previous research that typically relied on single-source data or surveys. We employed advanced statistical methods to comprehensively assess how a variety of factors collectively influence customer ratings. The findings highlight the complex interplay between a restaurant's operational strategies and its online reputation. Key insights reveal that while traditional factors like food quality and service remain critical, other operational factors such as pricing strategy, opening hours, and even the frequency of customer reviews notably influence ratings. This research sheds light on the multifaceted nature of customer satisfaction in the restaurant industry and provides empirical guidance for restaurateurs aiming to enhance their online presence and customer appeal.

1 Introduction

According to Luca [1], a one-star increase in Yelp rating leads to a 5-9% increase in revenue. However, only a few restaurant owners know which elements (e.g. quality food, and services) of their business lead to higher customer ratings and reviews. Answering this question is crucial when operating a restaurant business since restaurant owners have a finite amount of resources to improve their restaurants. It is important to know which element to focus both resources and effort.

In our project, we utilized Linear Regression to examine the Yelp, Uber Eats, and Google Maps datasets. These datasets contain information about service options, food preparation time, cost, and operating hours, as well as static attributes, including location, cuisine type, parking availability, and restaurant size. These datasets held a huge amount of categorical data that using one-hot-encoding did not provide a high correlation between the data given and the rating. However, by using the combination of these datasets and turning the categorical data into empirical data, we were able to provide a higher correlation between the data and the rating we were trying to predict. What we discovered was that price matters more relative to poor reviews, but for good reviews, the environment matters more. We also found that opening hours, brand size, and number of restaurants in a category correlate the most with the rating of the restaurant. This can be interpreted that as a customer, if one can't choose a good restaurant, a restaurant that is open later is usually better. For restaurant owners, it means that choosing a bigger brand may bring greater challenges for good reviews on their restaurant, The restaurant owners should also avoid popular categories if they want to improve their chances of getting a good review on their restaurant.

2 Dataset

In our study, we analyzed factors impacting restaurant ratings using diverse datasets, focusing on operational characteristics, customer sentiment, and industry competition. We primarily used Yelp's dataset, rich in textual reviews and ratings from over 50,000 restaurants and 1.25 million reviews, providing data on location, cuisine, and service options, and concentrated on restaurant-specific data, excluding other business types. We also incorporated Uber Eats USA data, covering menu diversity and pricing for 60,000 restaurants, to examine how these factors influence customer satisfaction and ratings. Additionally, we used a smaller dataset from Google Maps, featuring around 1,100 restaurant reviews, to cross-validate Yelp findings and broaden our understanding of customer perspectives. Extensive data preprocessing was essential due to the large and varied datasets; we parsed each dataset, creating custom tables with relevant information to improve analysis efficiency and merged Yelp tables to combine user attributes with their reviews.

3 Analytical Approach

3.1 Hypothesis 1: Category Competitiveness

The competitiveness within a restaurant category, indicated by the total number of restaurants, is negatively associated with the average rating of the restaurants in that category.

This hypothesis aims to verify our idea of how higher competitiveness in a restaurant category results in customers being more critical of those restaurants. The intuition comes from a few reasons, but most prominently from the fact that customers have access to many other alternatives. The key variable of interest, the average rating, was calculated for each category and was plotted against the number of restaurants in said category. It is important to note that this analysis focused on the top 30 categories for restaurants because we do not have enough data points on very niche categories. The preliminary visual analysis through a scatter plot with a fitted regression line suggested a negative correlation, as depicted in the accompanying graph (Figure 1).

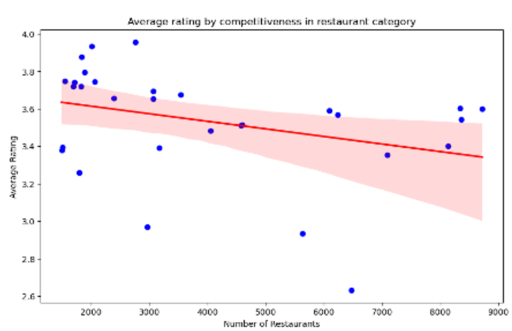


Figure 1: Scatter Plot Illustrating the Relationship Between the Number of Restaurants in a Category and Their Average Rating, With a Fitted Regression Line Indicating a Negative Correlation.

To quantitatively test Hypothesis 1, we employed statistical models that account for potential confounding variables such as restaurant price range, location, and cuisine type. We anticipated that these variables might also influence the average ratings independently of category competitiveness. Using a regression framework, we isolated the effect of category competitiveness on average ratings while controlling for these factors. The initial regression analysis yielded a statistically significant negative coefficient for the number of restaurants in a category, suggesting that as the number of competitors increases, the average rating of a restaurant in that category tends to decrease.

We also explored the potential for nonlinear effects, considering that the impact of competition might escalate or diminish as the number of restaurants grows. To this end, we incorporated polynomial terms of the competitiveness measure into our regression models. In parallel, we conducted a subgroup analysis to investigate whether the observed effect of competitiveness on average ratings differed across various price levels or types of cuisine. Finally, to ensure the robustness of our findings, we validated our analytical approach using a subset of the data and compared the results with the entire dataset. The consistency of our findings across these checks would lend credence to our hypothesis, suggesting that the density of competition within a restaurant category is one of the determinants of its average rating.

3.2 Hypothesis 2: Service Types

Restaurants with certain service types are more likely to receive higher ratings.

Each business on Yelp is labeled with the types of services it offers. This hypothesis investigates the relationship between restaurant ratings and the services provided. We aim to investigate the data to identify service types that correlate with higher-rated restaurants. Initially, we processed the Yelp dataset by filtering out restaurants that lacked specific service type attributes, thus focusing on those with well-defined services. We then calculated the frequency of each service type, identifying the most common attributes such as credit card acceptance, parking availability, and takeout services, among others. Subsequently, we categorized restaurants into groups with higher ratings (four stars and above) and lower ratings (below four stars). We visualized the top nine attributes across all restaurants and attempted to observe the differences in ratings between those that offer a specific service type and those that do not. These service types include credit card acceptance, takeout options, bike parking, child-friendliness, group accommodation, TV availability, delivery, outdoor seating, and catering. We visualized the presence (true) or absence (false) of these services across ratings using a violin plot, as shown in **??** in the appendix.

3.3 Hypothesis 3: Opening Hours

Restaurants with higher opening hours tend to receive more reviews but tend to have bad ratings.

This hypothesis aims to verify the impact of the restaurant's opening hours per day on the average ratings and number of reviews. The long opening hours usually lead to large foot traffic, which leads to more non-curious customers, thus leading to lower ratings. Therefore, a negative relationship between opening hours and average rating should be expected, but a positive correlation between opening hours and number of reviews. From testing this hypothesis, this research is able to provide suggestions to restaurants on how to adjust their opening hours in order to balance the traffic and performance. The opening time of each restaurant is indicated by the Yelp dataset with one opening window in 24:00 format on each day of the week. We filter out the day with 0 opening hours and handle overnight cases by adding up 24 to close time, then average each restaurant's opening hour per day. Restaurants with 0 opening hours on all seven days a week are filtered out. Finally, a scatter plot of the average rating against the opening hour is generated to check the relationship. Correlation and slope are calculated to help to verify the significance of the correlation between them.

3.4 Hypothesis 4: Brand Size

Restaurants that are larger brands tend to receive more reviews but tend to have bad ratings.

This hypothesis examines the influence of brand size on average ratings and the number of reviews for restaurants. We posit that larger, well-known brands should attract more foot traffic, resulting in a higher volume of reviews. Furthermore, due to their widespread recognition and appeal to a broad customer base, these big-brand restaurants may receive a diverse range of reviews, which could impact their overall rating. To quantify brand size, we will analyze the Yelp dataset, focusing on the top 20 most frequently mentioned restaurants. This approach ensures that the brands identified are recognizable in everyday life. We created a scatter plot to visualize the relationship between average rating and brand size. Additionally, we calculated correlation coefficients and regression slopes to assess the statistical significance of the relationship between brand size and restaurant ratings.

3.5 Hypothesis 5: Factors in Text Reviews

Factors that influence ratings stay constant between good ratings and bad ratings.

This hypothesis aims to explore all popular factors appearing in text reviews and test their relationship to the rating. The first stage works are done on the Google map dataset since it has fewer reviews. We extract all nouns in all review text in Google Maps, sort them by frequency, go over the most frequent one, and summarize them into five types of content: Waiting times, Price fairness, Environment, Service, and Food quality. There are also some frequently discussed topics, like parking and toilets, that detail the facilities not related to restaurant quality, thus they are not included. We manually tag all nouns with frequently > 1 if it's related to one of these 5 types. Finally, for each record of review text, we count a binary value of whether this text includes a certain type or not. The final result for each record data, is a length = 5 vector indicates whether this content talked about certain aspects or not, and the corresponding rating of this review, which is an integer between 1 and 5.

Since the rating is a discrete value between 1 and 5, and the input factors are also a binary value of either mentioned this topic or not. This makes it hard to make any numerical analysis depending on it, since all values are discrete, and the result would lack certainties. Therefore, we expand the rating to sentiment by graph. The box plot verifies the sentiment could represent the rating of a review text. Finally, a master graph about the probability of mentioning certain aspects in review against the sentiments they produce.

Lastly, in order to analyze the specific relationship between waiting time and rating, a larger dataset, Yelp's review text dataset was used. We filtered the text mentioning 'waiting' and 'hour' or 'minute', located to a certain phase including these words, and abstracted the number near hour and minute in text, finally producing a box graph between waiting time distribution to the rating.

3.6 Construct Validity

In the part of our research paper about how well our study measures what it aims to measure, we pay close attention to ensure the things we're looking at, such as service options and operational hours, accurately represent the important aspects of how restaurants operate and provide service.

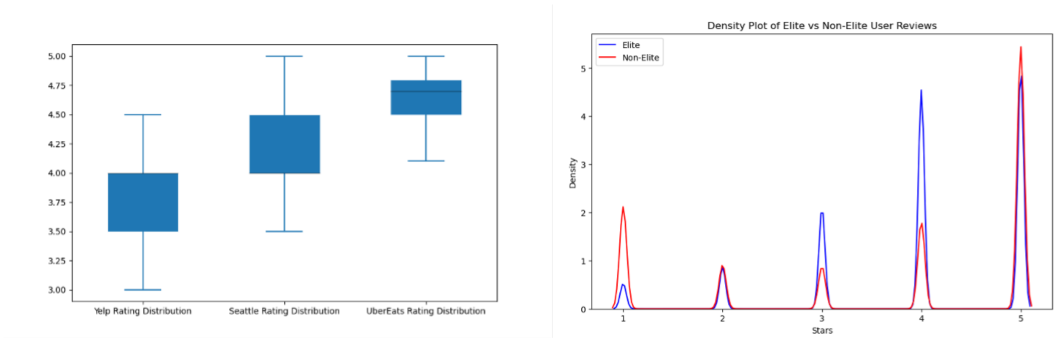
Brand Size: The size of a restaurant brand was calculated based on the number of branches it has. This basically means counting the total number of restaurants under a specific brand name in our datasets.

Waiting Time: To accurately measure waiting times in our analysis, we refined a multi-technique approach throughout the project. We extracted waiting time data solely from customer reviews. This involved filtering reviews for phrases indicating waiting (e.g., "waiting," "long time"), and identifying adjacent words or numerical values indicating duration, such as "2 hours." We further enhanced accuracy by including terms like "whole," "half," or "full," often used to describe time spans. Our dataset is regularly updated to ensure real-time relevance, guaranteeing that our analysis reflects the most current customer experiences up to 2023, thus avoiding outdated data issues.

3.7 Internal Validity

For our research to be reliable, it's important that our study design truly reflects how things work in the real restaurant industry. The accuracy of our investigation, especially when looking at factors like service options and hours, depends on how well we capture the essential elements that affect restaurant ratings in the actual industry context. To make sure our findings are trustworthy, we analyzed possible confounding variables and determined if it could have an effect on our data and try to account for it. What we found was that there was a bias to have more 1 star reviews among non-elite members when compared to elite members which follows a smooth distribution as shown in Figure 2a. We also analyzed to see if there was any bias in our dataset so we compared the Yelp dataset distribution of reviews and compared it to the Seattle Rating Distribution and UberEats dataset Distribution. What we found was that Seattle Dataset and Yelp Dataset have the same median however the Seattle Dataset is skewed to the right and the UberEats Dataset has a lot higher median and distribution seem to be biased towards higher ratings.

By looking at the distribution in Figure 2b, we found that Non-Elite Members give a lot more lower reviews than those of elite members as Elite-Members give a stable curve that drops off at one star while Non-Elite Members give significantly more one star reviews and instead drop off at two stars. This implies that can there is a possible bias where restaurants which gets more Elite Members would on average get higher reviews than those that don't.



(a) Box plot Verify the Rating Distribution Across Dataset (b) Rating Distribution of Elite/Non-Elite Users

Figure 2: Validity of Average Rating

3.8 External Validity

To understand the broader impact of our research, we want our findings to be useful beyond our specific study on restaurant ratings. We aim to make our insights applicable to different types of restaurants. Our research, primarily grounded in data from Yelp, UberEats USA, and Google Maps, offers a comprehensive analysis of factors influencing restaurant ratings. The Yelp dataset encompasses a broad spectrum of restaurants across over 500 cities, offering varied cuisine types, price ranges, and service options. This diversity provides a robust foundation for generalizing our findings to a wide array of restaurant types and business models. By integrating data from multiple platforms like Uber Eats and Google Maps, we capture a multifaceted view of the restaurant industry. This approach enables our findings to be more reflective of the industry as a whole rather than being limited to a single source of customer feedback or operational perspective. The inclusion of restaurants from a large number of cities enhances the potential for our findings to be relevant in various geographical contexts. We can generalize our findings to any restaurant in the U.S. this way.

4 Results and Findings

4.1 Category Competitiveness

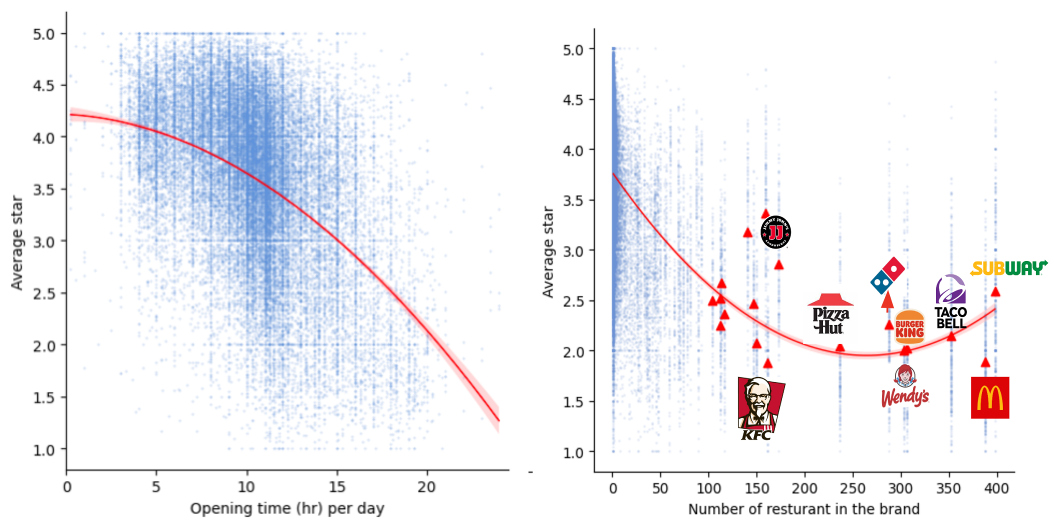
Our investigation into Hypothesis 1 revealed a distinct negative correlation between the competitiveness of a restaurant category and its average customer rating. This was visually evident in the scatter plot shown in Figure 1, where an increase in the number of restaurants within a category corresponded with a decrease in their average rating. This trend suggests that as restaurant categories become more competitive, with a higher density of establishments, their average customer ratings tend to decline. The statistical analysis reinforced these observations. Our regression models, focusing solely on the relationship between category competitiveness and average ratings, indicated a significant negative correlation. This result provides insight into customer behavior and preferences in the context of the restaurant industry, highlighting the impact of competition on customer satisfaction and perceived quality. It suggests that in highly competitive restaurant categories, establishments might face greater challenges in maintaining higher average ratings, possibly due to heightened customer expectations and comparisons.

4.2 Service Types

The data suggests that for certain service attributes, the presence of that service is associated with a higher median Yelp rating. Specifically, restaurants that offer bike parking and outdoor seating tend to have a median rating of 4.0, compared to a median of 3.5 for those that do not offer these services. This could indicate that amenities such as bike parking and outdoor seating are valued by customers and may contribute to higher overall satisfaction. On the other hand, services like accepting credit cards, providing takeout, being good for kids, suitable for groups, having TVs, offering delivery, and catering show the opposite trend; restaurants not offering these services have a median rating of 4.0, whereas those that do have a median of 3.5. These findings challenge the initial assumption that the availability of certain services would always correspond to higher ratings. This complexity highlights the importance of considering multiple factors when evaluating what influences customer ratings on Yelp.

4.3 Opening hours

As shown in Figure 3a and the Table 1, the correlation coefficient between brand size and average rating was found to be negative (r = -525), suggesting a substantial inverse relationship. In essence, as the size of the brand increases, the average rating tends to decrease. This unexpected finding implies that customers' expectations of larger brands may not align with their actual dining experiences. Larger brands, which are often fast-food chains, might be meeting basic expectations of convenience and consistency, but they are not exceeding customer expectations in a way that motivates positive feedback. The correlation's slope, at -500 restaurants per star, quantifies the rate at which the average rating declines as brand size increases. Specifically, for every increase of 500 restaurants in a brand, there is a decrease of one star in the rating, signifying a moderate decline in relation to our brand size metric. It's important to note that the count of restaurants in a brand is based solely on the Yelp dataset, which represents just a fraction of the entire Yelp dataset. Additionally, this decrease is not entirely linear; for extremely large brands, the trend of decline tends to plateau.



(a) Negative relationship between opening hour and average rate of restaurants.

(b) Negative relationship between brand size and average rate of restaurants.

Figure 3: Relationship Between Brand Size, Opening Hour and Average Rating

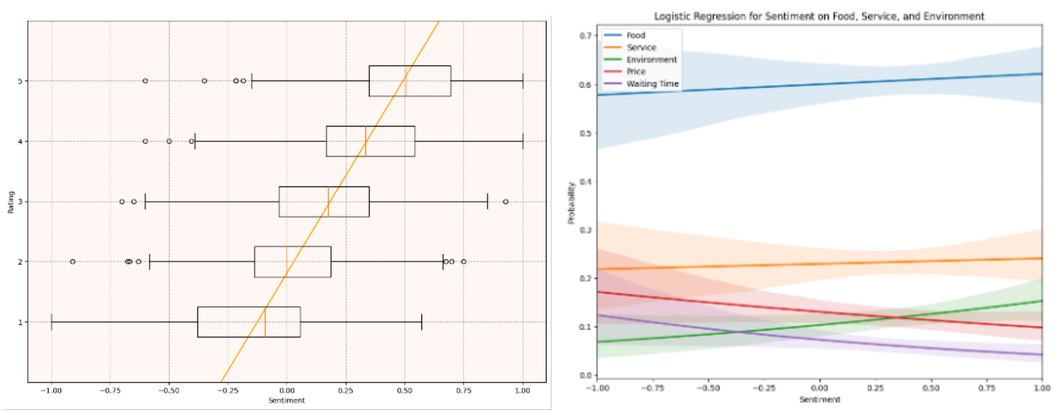
4.4 Brand size

With Figure 3b and Table 1, opening hours exhibited a negative correlation with average ratings (r = -0.449), albeit to a lesser extent than brand size. The correlation suggests that restaurants with more extended hours do not necessarily garner higher average ratings. This could reflect a market dynamic where restaurants with longer hours are possibly overextending their operations, leading to a dilution of quality or service that impacts customer satisfaction. The slope (-7.79hr / star) of this relationship indicates that it is open 7.79 hr more which leads to 1 star in average rating decreases for each additional hour that a restaurant remains open.

These findings underscore a counterintuitive aspect of consumer behavior: more is not always better. In the case of brand size, the ubiquity and predictability of large brands seem to be at odds with the factors that drive higher customer ratings. For opening hours, the apparent convenience of extended availability does not translate into better customer experiences as measured by average ratings.

4.5 Waiting times, Price fairness, Environment, Service, and Food quality

Our probability graph Figure 4b, which cross-references the likelihood of mentioning each factor against the sentiment scores, provides a visual representation of their relationship. It is evident that certain aspects, such as waiting times and price fairness, have a more pronounced effect on negative sentiments, whereas positive sentiments are more likely to be associated with the environment and food quality.



(a) Box plot verifies relation of sentiment and rating

(b) Probabilities of Mention Different Aspects

Figure 4: The Probabilities of Mention Certain Aspects to Changes of with Sentiment on Rating

Correlation Analysis				
	Average Rating		Number of Reviews	
	Correlation	Slope	Correlation	Slope
Cost	+0.059	+0.0036 $^{USD}/_{Star}$	-0.004	-0.061 $^{USD}/_{Review}$
Num. in Category	-0.277	-4.80·10^{-5} $/_{Star}$	+0.003	+6.34·10^{4} $/_{Review}$
Brand Size	-0.525	-500 $/_{Star}$	-0.128	-0.34 $/_{Review}$
Distance	-0.011	-2.64·10^{-5} $^{Km}/_{Star}$	-0.026	-0.017 $^{Km}/_{Review}$
Opening Hours	-0.449	7.79$^{Hr}/_{Star}$	-0.117	-7.8 $^{Hr}/_{Review}$
Population	+0.045	+5.71·10^{-10} $/_{Star}$	+0.093	+3.15·10^{7} $/_{Review}$

Table 1: Correlation Between Different Factors to Average Review and Number of Rating

5 Discussion of results/insights

5.1 Category competitiveness

We believe that the negative correlation between categorical competitiveness and average restaurant ratings arises due to multiple factors. Our interpretation starts by considering the effects on the customers as the number of restaurants in one category increases. We believe that customers become more critical of restaurants with a lot of competitors in their category due to higher expectations and a wider array of options. In a more competitive market, customers are exposed to various standards of service, quality, and overall culinary experience. This elevates their standards and expectations to other restaurants in the same category. Additionally, this abundance of alternatives makes them less tolerant of below-par experiences. If they have a closely similar option that gives them even a slightly better experience, they are more likely to not care for one or two other similar options; and therefore be more critical in rating it; which explains the negative correlation we observed.

5.2 Service types

Initial analysis of the correlation between service types and restaurant ratings has yielded results for certain services. However, it is challenging to determine whether a definitive correlation exists, primarily due to a significant amount of missing data. Approximately 10% of the restaurants in the Yelp dataset lack service type attributes. Given that service types are reported by customers rather than provided directly by businesses, this data missingness complicates the analysis, making it difficult to establish a reliable correlation. Nonetheless, it can be suggested that restaurants benefit from a complete profile of services provided, as this can be advantageous for customers. Restaurants should endeavor to complete their service type information across various platforms to ensure users have comprehensive information. Should one factor matter more to users, it is likely that the restaurant will attract more traffic, reviews, or even higher ratings.

5.3 Opening hours

The negative correlation between the opening hours and the restaurant ratings indicates that customers are more satisfied with restaurants that are open for limited hours than those that open for extended hours. One possibility is the fact that a restaurant stays open for only limited hours and manages to stay in business is a sign that the restaurant has either a quality menu, quality service, or both, which allows them to attract enough foot traffic to make a profit. On the other hand, a restaurant that stays for extended hours could be a sign that the restaurant cannot attract enough foot traffic in a short amount of time. Another possibility is that the restaurants that stay open for limited hours are more likely to attract regular customers or customers who plan their visits in advance and already have a favorable opinion of the restaurant, which results in higher ratings. In contrast, restaurants that stay open for extended hours are more likely to attract non-regular or less interested customers, which results in lower ratings.

5.4 Brand size

Our analysis reveals a paradoxical trend where larger brands, despite their market presence and customer reach, tend to have lower average ratings. This phenomenon can be dissected by considering consumer behavior and expectations. Customers often approach large brands with a set of established expectations. These expectations are shaped by widespread marketing, social proof, and the brand's historical performance. When a restaurant is part of a prominent brand, customers expect a certain level of quality and service that is consistent across the brand's locations. However, this level of expectation also leaves little room for exceeding customer satisfaction, as the expectation is already high. Therefore, unless the brand consistently offers exceptional experiences or novelty, customers may feel their experiences are just meeting the baseline, which does not incentivize positive reviews. For business owners considering joining a franchise, the data indicates that while being part of a large brand may guarantee a certain level of foot traffic and recognition, it may be challenging to achieve high performance or receive positive feedback. This could result in a lack of a sense of achievement or fulfillment from customer interactions.

5.5 Waiting times, Price fairness, Environment, Service, and Food quality

The results illuminate the importance of operational and experiential factors in shaping customer satisfaction. Waiting time, a significant driver of customer dissatisfaction, demonstrated a clear trend: as wait times increase, ratings decline, with a substantial number of 1-star ratings associated with longer waits. This suggests that customers have a low tolerance for waiting, and their likelihood of leaving a negative review increases dramatically with the wait time.

Price fairness also emerged as a critical factor in customer reviews. Instances of perceived unfair pricing were commonly associated with negative sentiment. Conversely, positive reviews often highlighted environmental aspects, suggesting that ambiance and setting are influential in creating delightful dining experiences.

Our analysis further suggests that businesses could significantly reduce negative reviews by ensuring service efficiency, particularly in minimizing wait times. For example, serving customers within 10 minutes could potentially eliminate 75% of bad reviews.

Moreover, while the service and food quality are foundational to customer satisfaction, the nuances of the dining experience, such as ambiance and perceived value, are equally important. Restaurants should not only focus on the tangibles but also on the intangibles that contribute to the overall dining experience. The environment, which could be a differentiator in a saturated market, should be leveraged to enhance customer satisfaction and, consequently, ratings.

In conclusion, our analysis underscores the need for a holistic approach to restaurant management that prioritizes efficiency, fairness, and the dining environment to meet and exceed customer expectations.

6 Limitations

Our study, while comprehensive, encompasses several limitations that should be acknowledged. These limitations stem from the dataset characteristics, the scope of the analysis, and inherent challenges in the methodology. Below, we outline the key limitations:

Dataset Specificity: The primary dataset used in our study was the Yelp dataset. While extensive, it represents a specific segment of online review platforms and may not fully encapsulate the dynamics present in other platforms or geographical locations.

Platform Review Manipulation: We are aware that some large restaurant businesses have private relationships with restaurant rating platforms like Yelp and that they may use that to remove some of their extremely negative reviews. However, since we do not have direct access to that information, we cannot account for its potential impact.

Subjectivity of Reviews: Customer reviews are inherently subjective. Factors such as personal bias, expectations, and individual experiences can significantly influence the ratings, making it challenging to draw objective conclusions.

Causal Inference: Our study primarily relies on correlational analysis. The nature of this analysis does not allow for strong causal inferences to be drawn about the relationship between restaurant competitiveness and average ratings.

Temporal Dynamics: Our study does not account for changes over time, such as evolving customer preferences or shifts in restaurant management, which could impact the relevance and applicability of our findings in a longitudinal context.

Data Quality and Completeness: As with any analysis reliant on large datasets, the quality and completeness of the data are pivotal. Inconsistencies, missing data, or errors within the datasets can potentially bias the results.

Recognizing these limitations is essential for a nuanced understanding of our study's findings and for guiding future research in this area.

7 Related Work

Hengyun et al (2023) [2]. studied the effect of competition on online reviews. What they found was competition influences both positive and negative review manipulations. But we studied which factors effected the average review. Also, reviews are more influenced by competitors of the same cuisine than by those of different cuisines. While we looked at how competitors effected reviews in general. Also, high-priced restaurants are more influenced by competition in terms of positive review manipulation than low-priced ones. However, such differences were observed for negative review manipulations.

Hu et al (2022) [3]. focused on how food photo types in restaurant reviews affect consumers' purchase intention. What they found was that reviews with process-focused food photos led to stronger purchase intentions than those with outcome-focused photos like pictures of the food itself. Also, the type of food photo had a significant effect on purchase intentions. Meanwhile we studied what caused bad reviews versus what caused good reviews

Benkhe et al (2015) [4]. researched why people use Yelp. They surveyed Yelp-users and found that they used Yelp to efficiently find information. Their main goal was to reduce search time and avoid poor purchase decisions. Users trusted Yelp because of the strong community feeling generated by using Yelp and perceiving the high-quality reviews they saw as trustable. Furthermore, users are more likely to disregard extremely positive or negative reviews in their search for good restaurants. Finally, Users give reviews because they believe it would help others make informed decisions. Meanwhile we used this analysis to help provide reasons why restaurant owners should care about reviews.

Yu and Margolin (2021) [5] studied reviews and how status-chasing affected reviews. They found it only took around 100 reviews for a restaurant to achieve a stable rating. They highlighted how elite members will review differently and we found the same thing in our study revealing elite members giving higher ratings on average than those of non-elite members.

8 Ethics

In this project, we handle sensitive customer data, necessitating a thorough consideration of potential risks associated with its misuse. A primary concern is the potential for inadvertently revealing socio-economic statuses of highly active users. For instance, frequent ratings of high-end restaurants might expose an individual's economic standing or even their habitual locations, posing privacy risks. Moreover, the nature of comments left by users warrants attention. The tone and content of reviews could inadvertently become a target for internet-based negativity, such as hate speech or racism. It's crucial for users to be aware of these implications when posting comments. To ensure ethical handling of data, our project adheres strictly to the terms of use for three key datasets: Yelp Academic Dataset, UberEats, and Google Maps. We guarantee that no information or insights derived from these sources will be utilized for commercial purposes. Our focus remains exclusively academic, aiming to understand the impact of various factors on restaurant ratings and to derive actionable insights for both consumers and businesses.

9 Future Work

Our current project integrates diverse datasets, including those with ratings from online orders as well as dine-in experiences. For future research, a promising avenue would be to segregate these datasets based on the nature of the dining experience. Analyzing the differences in customer behavior and preferences between online and dine-in scenarios could yield fascinating insights. Another area of interest lies in addressing discrepancies across the datasets used. Future projects could focus on harmonizing these datasets by identifying and linking common attributes. This approach would facilitate more comprehensive analyses and potentially unveil new patterns and trends. Additionally, Yelp's API offers a wealth of detailed information about businesses, which was not fully utilized in our current scope. Leveraging this API in future work could significantly enhance the predictive accuracy of our models and unearth deeper insights into consumer behavior and business performance.

10 Conclusion

In conclusion, this comprehensive study has uncovered the intricate nature of factors influencing restaurant ratings on online platforms. We found that customer satisfaction and ratings are influenced not only by traditional factors like food quality and service but also by a range of other operational aspects such as pricing, competitiveness, and the different types of services offered. These insights support some conventional wisdom in the restaurant industry but also reveal new insights and suggest that a more nuanced, holistic approach is crucial for enhancing customer satisfaction and improving online ratings.

Our findings provide actionable guidance for both customers and restaurant owners. For customers, this research allows them to make more informed choices when choosing what restaurants to go to and ensure they get a high-quality experience. For restaurant owners, the research guides them and emphasizes the importance of considering all significant aspects of the dining experience in their operational strategies. Moreover, this research paves the way for future studies to further explore these complex dynamics, particularly in the context of evolving consumer behaviors and the digitalization of the hospitality industry.

Overall, the study makes a significant contribution to our understanding of what drives customer ratings in the restaurant sector, offering a rich perspective for industry professionals aiming to adapt and thrive in a competitive market.

References

[1] Luca, Michael, Reviews, Reputation, and Revenue: The Case of Yelp.Com (March 15, 2016). *Harvard Business School NOM Unit Working Paper No. 12-016*, Available at SSRN: https://ssrn.com/abstract=1928601 or http://dx.doi.org/10.2139/ssrn.1928601

[2] Li, H. et al. (2023) Competition and restaurant online review manipulations: A dynamic panel data analysis. *International Journal of Hospitality Management 115*. United Kingdom: Elsevier Ltd.

[3] Liu, H. et al. (2022) Process vs. outcome: Effects of food photo types in online restaurant reviews on consumers' purchase intention *International Journal of Hospitality Management 102*. United Kingdom: Elsevier Ltd.

[4] Parikh, A. et al. (2015) A Qualitative Assessment of Yelp.Com Users' Motivations to Submit and Read Restaurant Reviews*Journal of Culinary Science & Technology 13*. pp. 1-18 United Kingdom: Taylor & Francis Group.

[5] Yu C Margolin D (2021) (2021) The disproportion of crowd wisdom: The impact of status seeking on Yelp reviews, *Public Library of Science ONE 16(6)*. San Fransico.