Text Analysis for Variance in Restaurant Chain Reviews

BYGY/ISGY Web Analytics  
Professor Yilu Zhou

December 10, 2020

Section 1 - Team 5

Vanessa Asaro  
Annemarie Donohue  
Urvesh Bhadani  
Nicolas Clarke de Dromantin

**Executive Summary**

Yelp, a highly popular crowd-sourced local business review and social networking site, has more than 178 million unique visitors every month. Yelp is engrained in consumer behavior patterns, with 45% of customers likely to check reviews on Yelp before visiting a business (*Yelp Fact Sheet: Stats Your Business Needs to Know* 2020) . The type of business that gets the most reviews and tractions on Yelp are restaurants (*Fast Facts*). For restaurants, each location has their own page. This includes a 0-5 star rating, and the option to leave a written comment, video and/or pictures. An increased rating of one star will increase revenue 5-9% (Luca, ["*Reviews, Reputation, and Revenue: The Case of Yelp.com*](http://www.hbs.edu/faculty/Pages/download.aspx?name=12-016.pdf)”). Smaller, local businesses are able to thrive by having their main source of new customers, “word of mouth” reviews, now be public. This has left chain restaurants in a revenue deficit as Yelp has increased in popularity due to the accessibility of local reviews. Yelp granted larger businesses the competition of local restaurants by assuring the meal would be pleasurable. People go to chains to know what they are getting, especially if they are out of state. Now, local places are validated as deserving of attention to travelers due to Yelp. Ratings are indicative of a restaurants’ performance in revenue (Luca, ["Reviews, Reputation, and Revenue: The Case of Yelp.com](http://www.hbs.edu/faculty/Pages/download.aspx?name=12-016.pdf)”) .

This report will focus on the Yelp ratings and reviews of a popular Mexican Restaurant chain, Tacombi. Starting their company in 2005 from a van in Yucatan and Playa del Carmen, Mexico, Tacombi captured their dream of serving authentic Meixcan food into a reality (*Company: Story*). Now a successful chain of restaurants in New York City, there are 8 total locations spanning Manhattan and Brooklyn. Coming to New York in 2008 with their street taco van, Tacombi opened their first flagship restaurant, Nolita, in 2010. Our analysis will specifically focus on the flagship due to Nolita having the highest amount of reviews and shares, yet the lowest rating amongst the chain. This project will analyze the ratings and reviews of Tacombi locations on Yelp to improve ratings, leading to increasing net revenue. We will explain how the Nolita location can improve with web content analytics by analyzing ratings and reviews. Ratings will show a distribution of the star ratings, along with a comparison to a highly rated location, Bleecker St. Reviews will be evaluated through crawling Yelp reviews and using text analysis to analyze and compare terminology of poor reviews. With recommendations to increase their rating based on our analysis, Tacombi will be more informed on what their customer is thinking and where to implement change based on our project.

# 

# Business Goal Analysis

# 

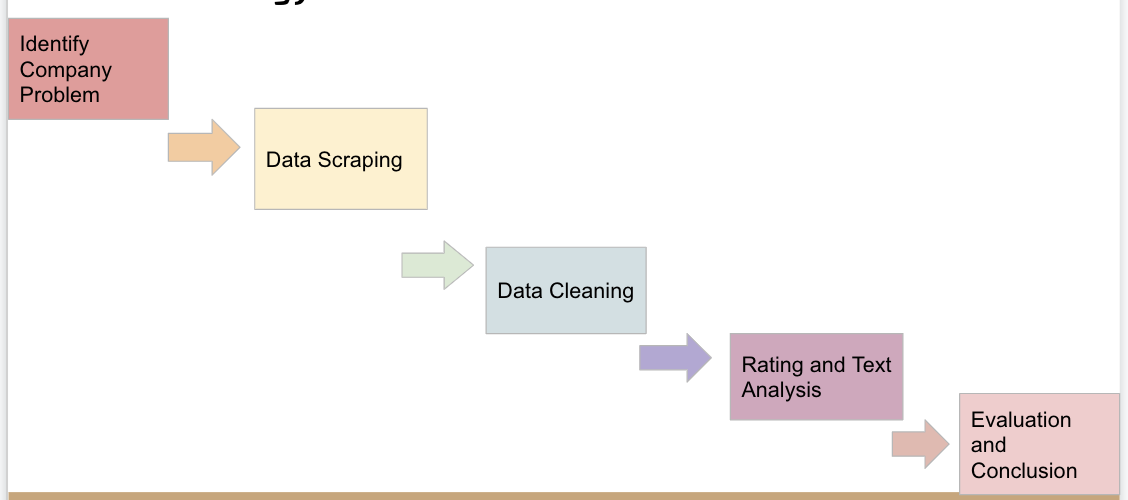
# The goal of this study is to understand how Tacombi can increase their rating and reviews for their flagship, Nolita, location. The difference in ratings is unusual because each location offers the same menu and prices with similar atmospheres. By comparing the locations of Nolita and Bleeker, the result will show why there is a variance in average rating, more reviews and/or lower scores. On Yelp the average locations, including Bleeker, have ~200 reviews with a 4 star rating. Nolita has 1100+ reviews with a 3.5 star rating. While exact net revenue behind each location is not publicly disclosed, we can draw a conclusion based on reviews that Nolita is currently not as profitable as the other locations.

Consultant agencies, especially ones specializing in data analysis, will have producers turn to them for better understanding of their consumer. In order to optimize the feedback of consumers, web content analytics is utilized, and therefore in this project. Discovering information that is useful from documents, data and web contents is later used for collection and optimization of results. Here the feedback for analysis is given to us through Yelp. When crawled, this website will exemplify trends, like common phrases in reviews. Since the reviews go back years, a time series analysis will be done to see if there are seasonal trends. By understanding the descriptive analysis, you will be able to implement methods for improvement in Tacombi’s ratings and revenue through predictive analysis. The only risks would depend on what actions are taken in accordance with the outcome of trends, but knowledge of what the customer is thinking, granting the producer more information, will be a significant advantage. Due to the fact all the locations have the same menu and ambiance, the results of this study can assist with location specific service improvement and marketing planning. Part of the analysis will include analyzing the differences in ratings and reviews of the Bleeker Street location as a reference. Since the Tacombi name highly associates itself with locally sourced ingredients, authenticity and hospitality, those variables will be highly valued. This analysis will grant Tacombi more information on their highly valued morals in order to make sure they are seen and appreciated by the customer, therefore increasing revenue.

**Dataset Description**

This report consists of 2 dataset categories, one with the Nolita Tacombi Yelp Reviews and one with the Bleecker Street Tacombi Yelp Reviews. To get these datasets, the data was extracted through web scraping using Python. This HTML was extracted through importing and running modules of selenium through webdriver and beautiful soup. Selenium webdriver was chosen for its ability to collect APIs of open source used for testing of a web application. Beautifulsoup was used to parse or extract data from the HTML, and can also be used on XML. The scrape of the Nolita Location consisted of 56 pages of Yelp data and the scrape of The Bleecker Street Location consisted of 20 pages of Yelp data. To get all these pages of HTML, multiple loops were used to place URLs into a list, extract drivers, and form drivers to get soups. This soups list is the HTML, used in the following code to extract necessary elements in our study.

# System Design



Our project will revolve around analyzing the Yelp reviews and ratings of Tacombi’s Nolita location. The Bleecker Street, West Village location will also be used for analysis, as explained above. The crawled Yelp data includes reviews, users, ratings, and date of reviews from both locations.

*Scraping the Data*

The necessary elements that were extracted during the first round of crawling were user, date, rating, and review. A couple rounds of scraping and modifying code was required to get the same number of users, date, rating, and reviews. Some revising was done to match up previous comments and business owner comments, which are stated within one review. Using the pandas module, the data in these lists were appended to columns within a data frame. Pandas was chosen for the data structure operation manipulation and analysis. The Nolita data consists of 1135 rows with 4 columns and the Bleecker Street data consisted of 389 rows and 4 columns.

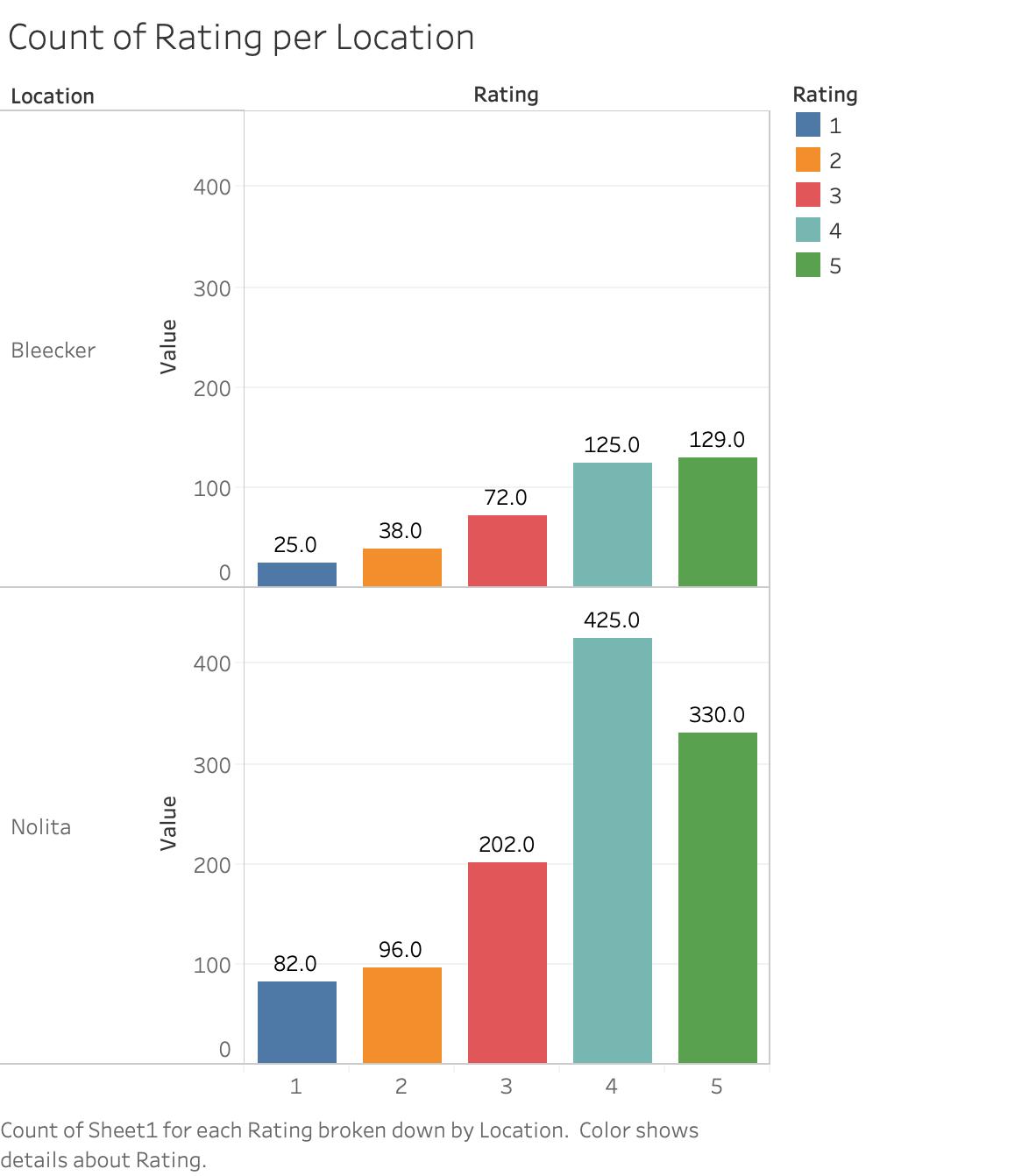
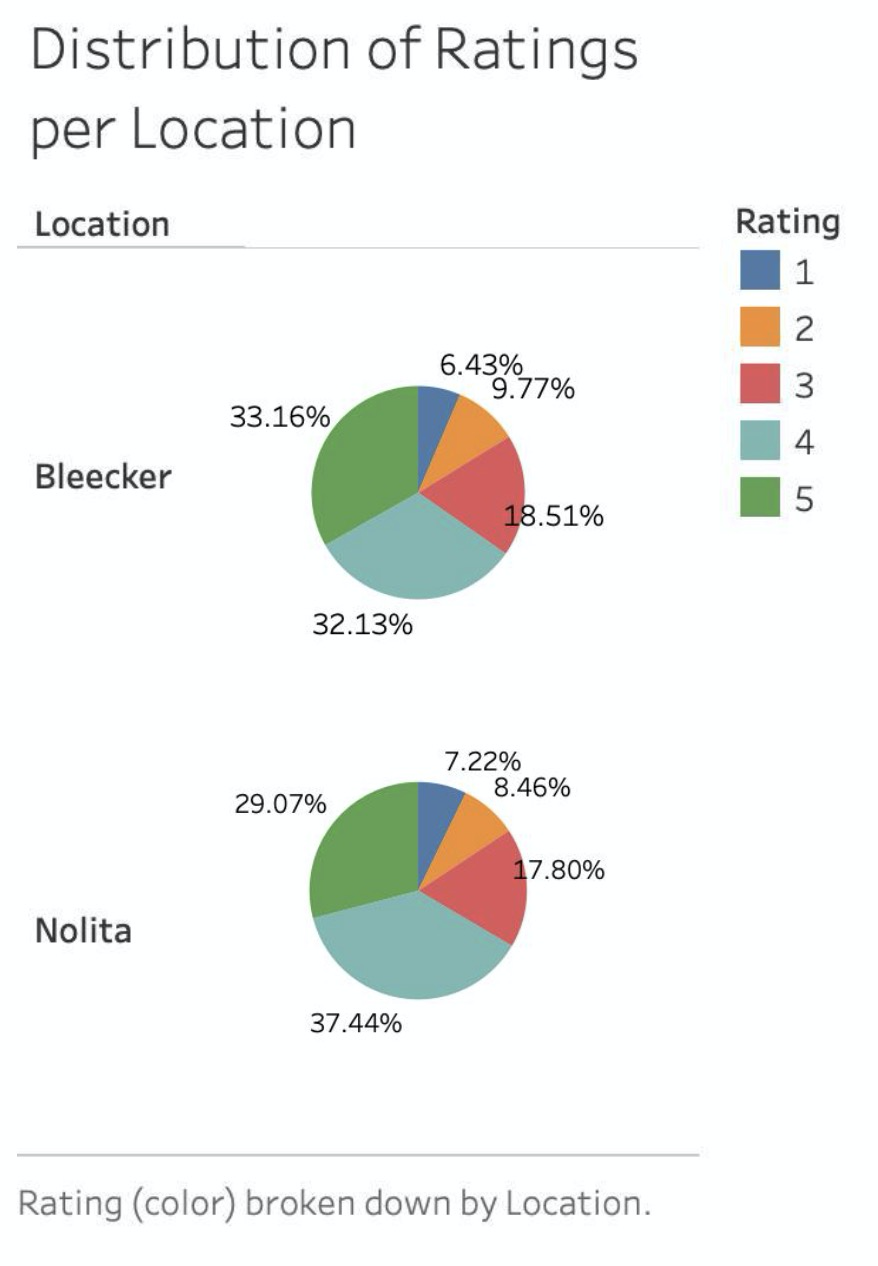
*Cleaning the Data*

The first step in cleaning the data was cleaning the date column. This column needs to have date values only to be used in time-series analysis. But, attached to the date on some reviews there was ‘Updated review’ and ‘Previous review”. With pandas in python, the date was extracted from the dataframe, ran through a loop, and removed through splicing the indexes with those letters. This data was exported to a file. Further cleaning and organizing of the data from the original dataset involved new columns based on users who had previous reviews. This was done through extracracting previous date from dates, previous rating from rating and the previous review content from review. These were put in a new column that aligned with the user.

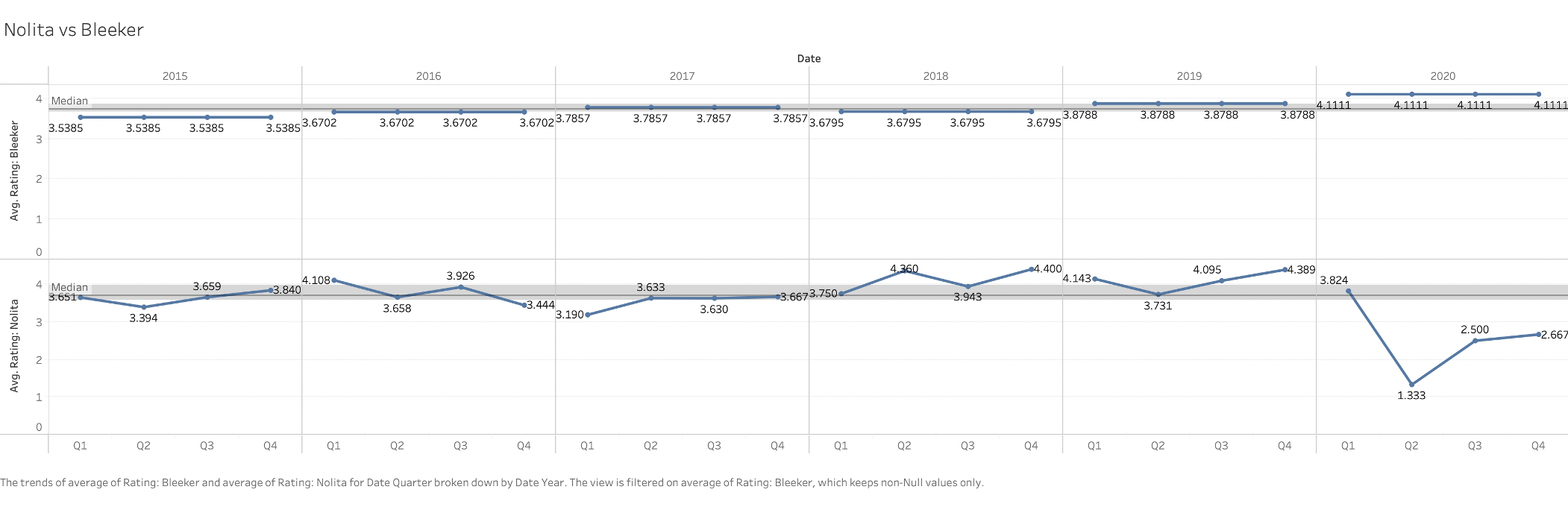
**Data Analysis: System implementation**

*Ratings*

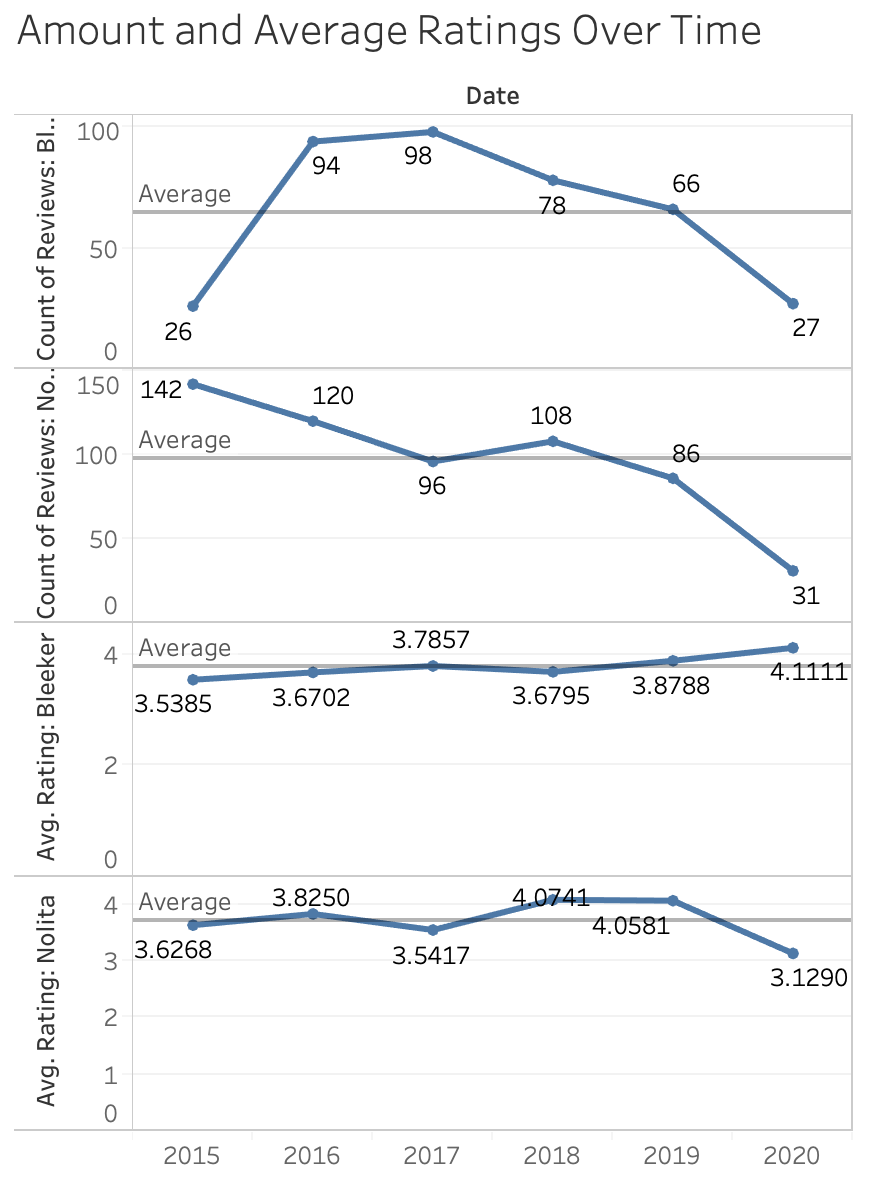
To get a baseline understanding of where the flagship location stands on Yelp, the ratings were analyzed. The data visualization tool Tableau utilized the datasets for Nolita and Bleeker, exported from the cleaning above, to display the distribution of ratings within each location. Tableau is an interactive data visualization software, used here to show results in a manner that is easily digestible. The 0-5 star rating is explained through the legend, while a pie chart was used to understand the deviations of ratings out of a whole. Unexpectedly, the distribution between the two locations is extremely similar. The count bar graph is another way of showing the distribution. Using the python code, we were able to get an average mean for Nolita as 3.73, and Bleeker as 3.76. This leads us to believe that the cut off for 3 ½ stars is .75, and anything above is rounded to the nearest whole number. Even though the ratings differ by 0.03, having the visual of the yellow star 3 ½ rating is still a visual difference from the red 4 star rating. As stated previously, every star counts, even part of stars. The median, however, was 4.0 for both Nolita and Bleeker. This continues to show how similar the ratings between the two are.



Now having a general understanding of the ratings, further analysis in Tableau was done through a comparison of average ratings by quarter of Bleeker and Nolita over time. A time series analysis with a linear regression overlay, showing the rating over time, is shown on the graph below. A linear regression is shown in grey, emphasizing the median and quartiles. The Nolita location opened in 2010, and Bleeker was opened 5 years later (2015). The Bleeker location shows consistency, with a slight increase from year to year. Nolita shows an increase then decreases from quarter to quarter. Generally, there is a decrease in the second quarter and increase in the 3rd, while the 4th and 1st are more stagnant. This means that the fall is less profitable, and winter has a raise. The last section of Nolita is probably skewed due to the pandemic, having a dramatic decrease with little comeback, which started in the second quarter of 2020. However, Bleeker continues to stay consistent, even raising their ratings through the closings of restaurants due to Covid-19. As a result of the pandemic, many restaurants have closed. To be able to raise ratings and constantly keep them high through the past pandemic year is an achievement.



Another analysis was performed through Tableau to show when the most ratings were submitted, and how the ratings were affected. The count of reviews and average rating are calculated per year (not cumulative) for the past 5 years. The labels from top to bottom are Count of Reviews: Bleeker, Count of Reviews: Nolita, Avg Rating: Bleeker, and Avg Rating: Nolita. Generally, the more reviews submitted the better the average rating was. Both locations have this trend, but is very slight. The last year is very skewed due to the pandemic, so the number of ratings is bound to be lower with less people allowed to be served (besides take out/delivery). Even still, Bleeker continues to increase in rates with significantly less reviews, while Nolita, also with a much lower number of reviews than average, takes a hard hit on their average rating.



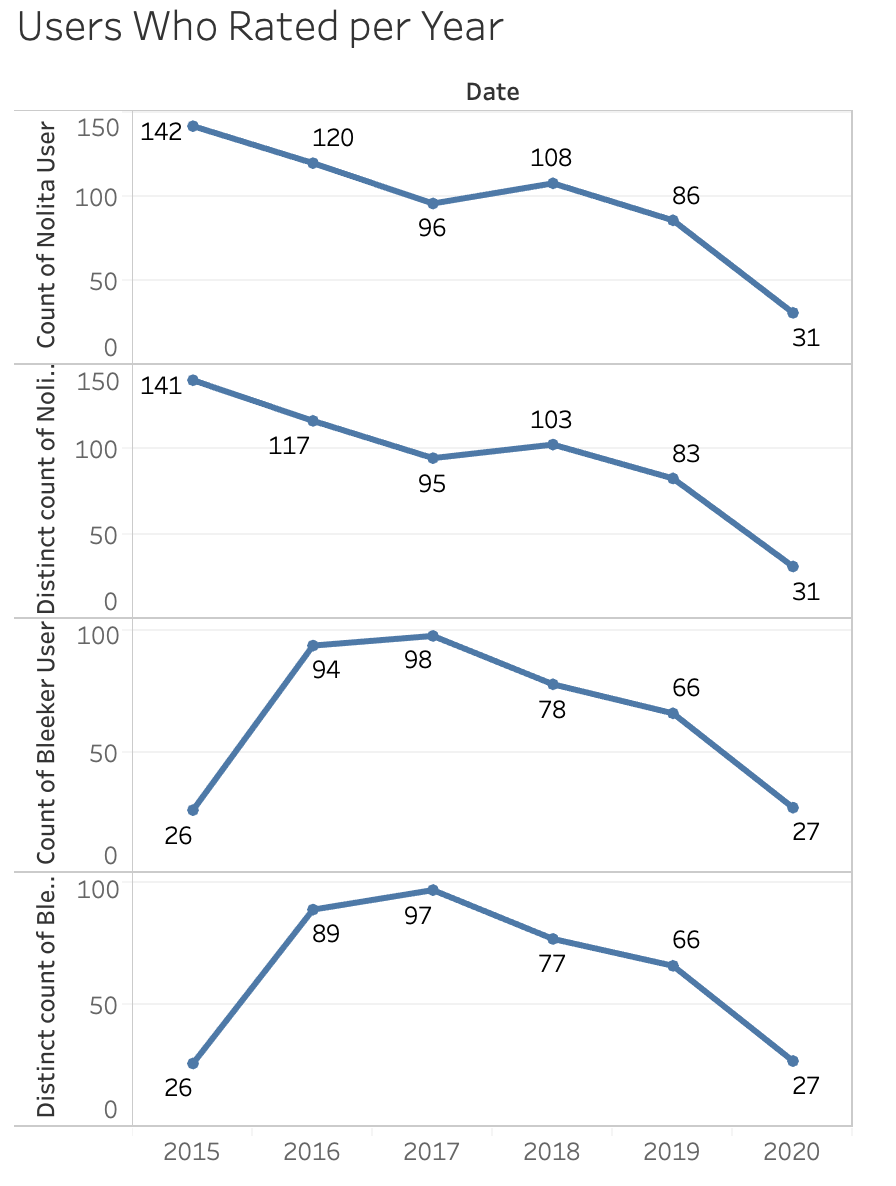
*Reviews*

Once the data was scraped and cleaned, the reviews were analyzed to comprehend the text behind the ratings. The reviews will give insight in terminology of what was liked and disliked, which is important to understand what the restaurant should enhance or change.

After looking at frequent words for all reviews, the 178 reviews for only the ratings under 3 stars for the Nolita Location were extracted using Python. Then, each review was split up into word strings. Using the collections module, a counter was imported to count the most common words in that set. While the top words were stopwords or other broad words, such as ‘the,’ ‘and,’ and ‘food.’ The module asks for input for the number of words to look up, so we chose the top 200 relevant words. Significant words found, followed by their frequency, were: [('good', 60) ('back', 42) ('service', 40) ('time', 39) ('better', 36) ('never', 35) ('small', 30) ('wait', 29) ('minutes', 28) ('tasted', 27) ('bad', 26) ('prices', 18) ('waitress', 17) ('staff', 17) ('atmosphere', 16) ('expensive', 16) ('back.',16) , ('waiter',15), ('service.', 13)]. These words were investigated, being an insight into opinions prominently exemplified in bad reviews. Since ‘price’ and ‘food’ are supposedly consistently used throughout restaurants, service was looked at. Service was a topic that was different between locations, while ‘food’ and ‘price’ were the same. Out of the 45 reviews with the word ‘service’/’service.,’ 33 had a negative service comment talking about slow service or how unsatisfactory their service experience was. Out of the comments that mention the service explicitly, 78% of text associated service with negative service. Diving into this further, on the topic of customer service, the reviews with the frequent words of ‘waitress’, ‘waiter’, and ‘staff’ were looked at. An overwhelming amount of those reviews mentioned members of wait staff being rude, inattentive, and inadequate. Specific words include being homophobic, racist and rude to someone about being in a wheel chair. Another point of service, service time, had reviews with the words ‘time’, ‘wait’, and ‘minutes’ inspected. The reviews with these words mention a long waiting time on three different points of customer experience at the restaurant, which include long waits for food to come to table after ordering, wait for waiters to attend to them and the wait to be seated. The long wait time for food coming to the table potentially has to do with kitchen efficiency or possibly waiter picking up the food. The long wait time for waiters to attend to the tables has to do with customer service, either between the servers or the hostess. If people are waiting at their tables for time longer than they would’ve normally stayed due to the service, revenue will be affected as table turnover slows down. This could go on to affect reviews as people who are given a long wait time on busy days waiting even longer to be seated. Going through the text analysis, ‘food poisoning’ was also found a few times. Reviews mentioning food poisoning were examined and there 8 reviews mentioning it. Hearing someone got food poisoning is definitely a strong convincer for others not to eat at a restaurant. However, this is an extremely low count, 8 being less than 1% of the reviews. However, this is especially concerning for a restaurant who prides itself on local and sustainability sourced ingredients.

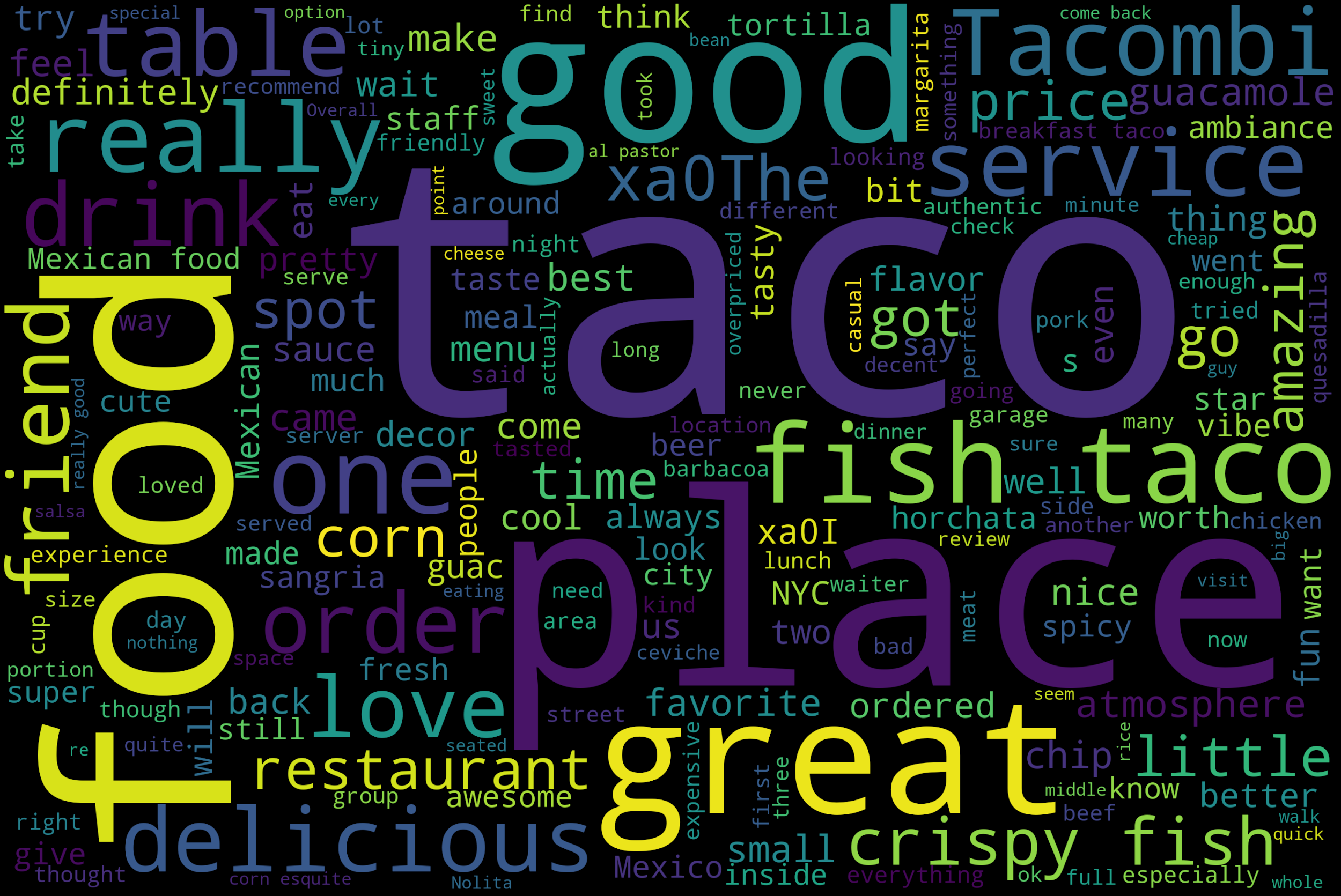
Nolita was then compared to the 63 reviews with 1 and 2 star ratings for Bleecker Street Location, using the same method as above to find the top 200 frequent words. Some significant words found, followed by their frequency, were: [('good', 16) ('time', 13) ('better', 11) ('waiter', 8) ('tasted', 8) ('small', 8) ('bad', 7) ("wouldn't", 7) ('worth', 7) ('decent', 7) ('expensive', 7) ('dry', 6) ('minutes', 6) ('waiters', 6) ('service', 6) ('horrible', 4) ('dry', 6) ('minutes', 6) ('waiters', 6) ('horrible', 4)]. When looking at some of the same words we can derive insights that the lower rating terminology for Bleeker is not as harsh, when words like bad and never were replaced with decent or small. For Bleecker Street, ‘Service’ was in 9.5% of the bad ratings and for Nolita, ‘Service’ was in 22.5% of the bad ratings. Looking at time in service, slow wait time was only in 3 reviews. That is only 4% of all bad reviews for this location.

Another attribute of analysis looked at was the number of users that sent in a review (distinct count), and the total number of reviews. As shown in the Tableau chart below, Nolita had 583 ratings over the past 5 years, with 13 repeated users. This makes up for 2.3% of the ratings for this location. Bleeker had a total of 389, with 7 repeated users, making up 1.8% of ratings. Even though these are both small percentages, every rating does count. Both of these are about 2%, which is a very low number, meaning that users almost never leave more than one rating. Instead, this data should be viewed in the sense that a first impression is very important, and normally sticks with the customer.

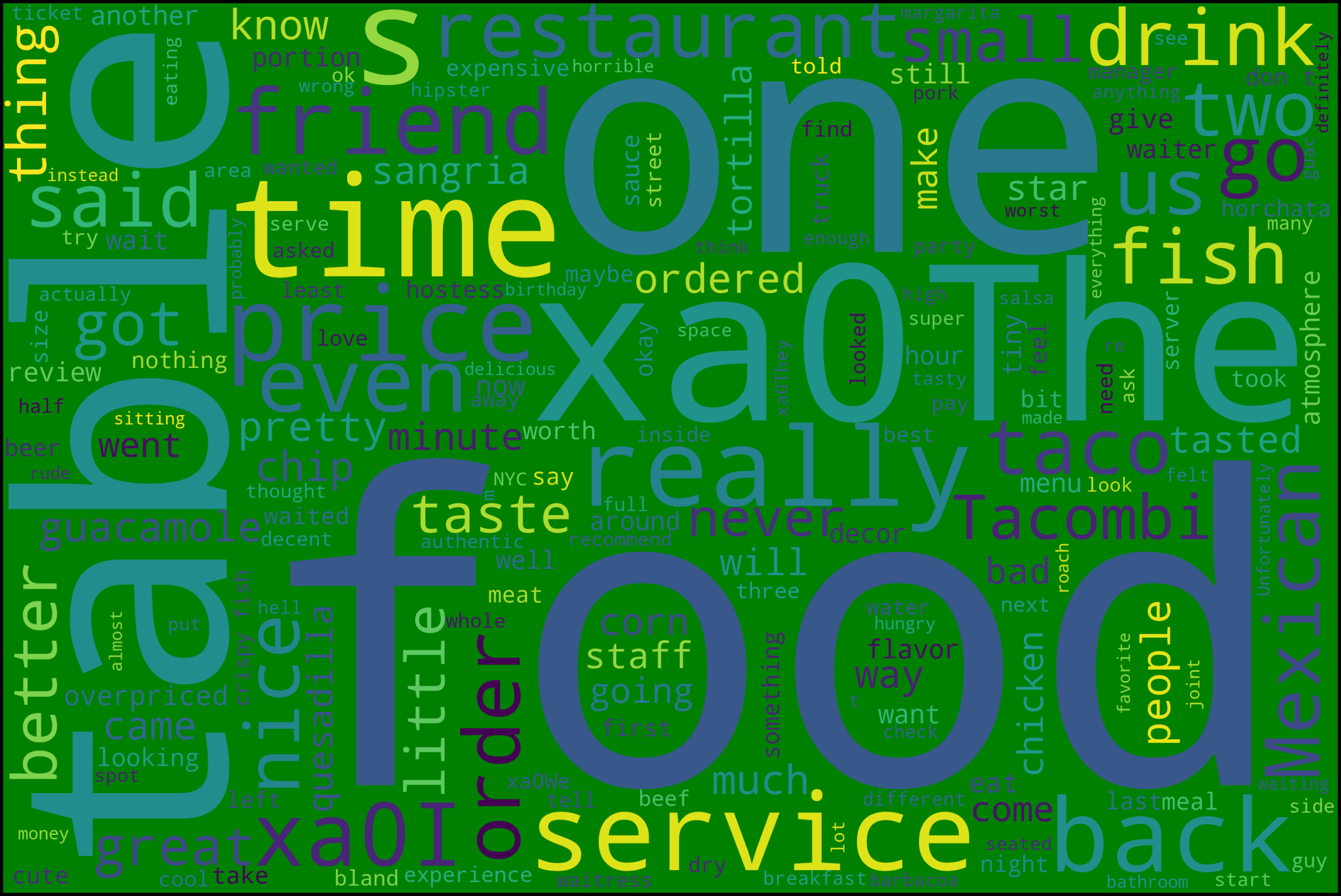


Lastly, reviews were coded through Python into a set of Word Clouds for visual representation for what was found in the text analysis. We created Word Clouds for both locations and made different ones based on 1 and 2 star reviews and overall reviews. Data visualization is a way to engage the client in a clear way, and WordClouds served that purpose here for the presentation section of our project.

Nolita: all reviews



Nolita: 1 and 2 star reviews:

****

These negative reviews were consistent with two recurring issues of the flagship location, price and service. Customers at Tacombi Nolita were dissatisfied with the service of the restaurant and the price they were paying for the value they were getting. Again, although the means were close and median rating was the same, the terminology within the reviews differed.

**Evaluation**

Web content analytics is discovering functional information from documents, data and/or web contents. This Tacombi analysis brings to attention to how much every review counts. With the ratings being very similar, both having the same median, Nolita had a yellow 3 ½ star review with Bleeker having a red 4 star review. Visually, this shows how even locations varying in average ratings by 0.03 can affect stars, which affects revenue. Extracting words with high term frequencies (TF) gave us a better insight of how to enhance Nolita. Without text analytics, the location would not understand the severity of the need to improve their service and lower wait time. Another example would be having more reviews normally means a higher rating. This would not have been proven without having a side by side comparison of the average rating distribution and amount of ratings. Without web analytics, a business cannot properly understand their niche audience. There is an importance that in order to receive optimization, actions must be taken to change behaviors that are lacking, and enhance what is working.

**Conclusion and Future Direction**

Tacombi has opportunities to incorporate conclusions that we have formed from the analysis. The ratings are really similar between Nolita and Bleeker, but the reviews held differences. ‘Service,’ ‘wait time’ and ‘food poisoning’ were the top phrases for poor reviews for Nolita, which were not as significant for Bleeker. Suggestions would include increased training and efficiently for waiters and chefs in order to work on these complaints. Workers in customer service must implement patience and hearing the customer out, not racist. Nolita can increase net revenue by looking into these three phrases that were complained about.

In the future, there are insights that would only be accessible that are disclosed from the public. Analyzing the net revenue in a time series and comparing it to the trends seen in the analysis would solidify that the net revenue has a direct correlation with reviews. Although the correlation between positive ratings and increase in revenue is normally a true statement, we could analyze the exact impact every rating has by month. Another aspect, like having a high turnover rate for servers or management, would contribute to low service and morale. There are aspects of a company that are not open to public knowledge, and understanding the struggles or strides for the company behind closed doors would help put the rating and review analysis into context, therefore being more helpful. Public knowledge not stated on Yelp, like the demographics of the people who left positive reviews, could be collected in order to have a better understanding of who the target audience is. This leads into targeted ads, which we did not touch upon but would be in Tacombi’s interest to do. These demographic details are not given on Yelp, which is why they were not included in the analysis.

As for marketing ideas, New York University has multiple dorm buildings practically across the street, so this could be a great audience. The mexican food is already moderately priced, and having a student discount for the students would make the restaurant even more accessible. Another idea would be to target non-locals, considering that Nolita is a commercial area. There are a lot of large retail stores near the Nolita location, attracting a lot of tourists. The Bleeker St. location is in the West Village, where there is a more community feel. Marketing ads towards these non-locals through Instagram or Yelp would be sure to increase revenue. By taking into account the analysis and suggestions drawn through this project, Tacombi will be able to increase their net revenue in their flagship location.

**References**

Company: Story. (n.d.). Retrieved December 10, 2020, from https://tacombi.com/company/story

Fast Facts. (n.d.). Retrieved December 10, 2020, from https://www.yelp-press.com/company/fast-facts/default.aspx

Luca, Michael. ["Reviews, Reputation, and Revenue: The Case of Yelp.com."](http://www.hbs.edu/faculty/Pages/download.aspx?name=12-016.pdf) Harvard Business School Working Paper, No. 12-016, September 2011. (Revised March 2016. Revise and resubmit at the *American Economic Journal - Applied Economics*.)

“Yelp Fact Sheet: Stats Your Business Needs to Know.” *ReviewTrackers*, 15 Sept. 2020, www.reviewtrackers.com/blog/yelp-factsheet/.