Sushi Restaurants – Yelp Reviews Analysis

Jack Bressett, Rachel Studer, Ran Zhao

University of Wisconsin – Madison

Introduction

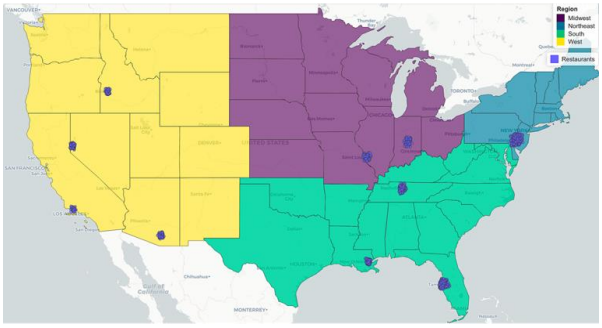
Yelp is a widely used platform with 178 million distinct visitors monthly. Research shows 45% of people utilize Yelp before visiting a business, and revenue has shown to increase 9% for each 1-star increase in Yelp rating (Elad, 2022). Thus, increasing one's rating is essential to businesses, specifically restaurants since these make up 17% of the total Yelp reviews (Yelp, 2022). To simplify the Yelp dataset, we focused on analyzing sushi restaurants throughout the United States, as available. Our team proposes the following questions:

Does region play a role in ratings? What topics get brought up within each rating level? What elements of service quality lead to higher stars? Which types of sushi rolls should restaurants serve? What reviews or tips for a business most embody that business?

These questions help determine how sushi businesses can specifically improve their ratings and avoid unsatisfactory reviews, thus earning higher revenue in the end.

Data Pre-Processing

The Yelp data consisted of 3 JSON files (business, reviews, tips) covering 1052 sushi restaurants - clustered near ten large cities involving 13 U.S. states. There are 168,293 reviews encompassing all restaurants.



Business Data

This dataset provides information on location, business type, and aspects of the business itself which were omitted from our analysis. To gather only restaurants with a focus of sushi, the business type variable needed to be cleaned first by filtering on 'Restaurant' and 'Sushi,' then hand filtering to eliminate those with other main focuses. A small portion of cities required renaming due to various spellings (e.g.: Saint, St.).

Reviews & Tips Data

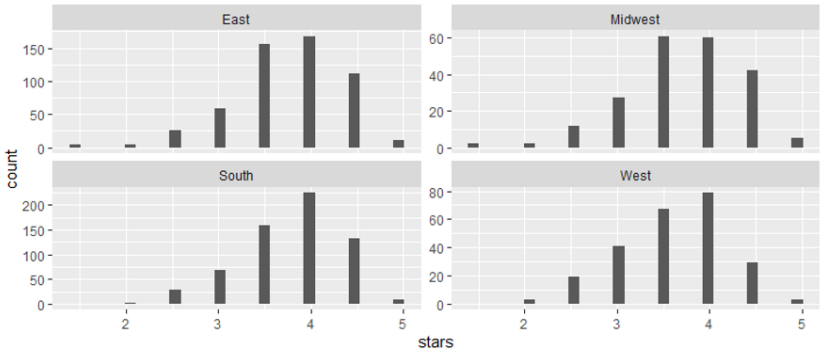
Each review on Yelp for all businesses is given within this dataset. Information on the star ratings, review text, and reviewer identification is available alongside the business identification to link the datasets together. To process this file, we had to parse through the review text for the selected sushi restaurants while throwing out words (stop words) unnecessary to our analysis (e.g. 'sushi,' 'the,' 'and,' 'I'). Reviews are also classified as positive if the star is greater than 3 or negative if less than 3. The tips dataset provides suggestions left by reviewers for a specific restaurant or those interested, like the structure of the reviews data. Ultimately, we are interested

in what topics are brought up by those who leave tips and need to perform text cleaning like the reviews dataset.

Data Analysis

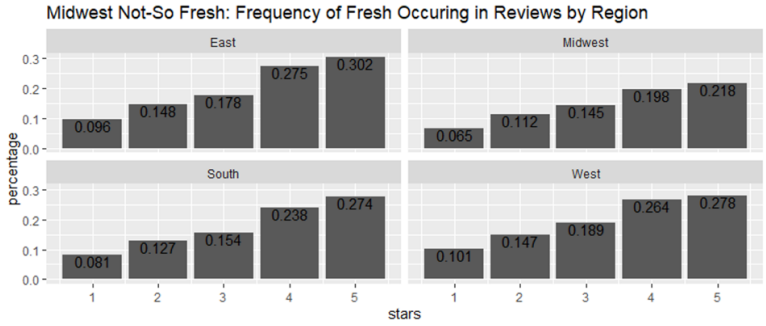
Exploratory Data Analysis

Sushi restaurants' average Yelp ratings increase in frequency from 1 to 4-stars, but there are fewer 4.5-stars and minimal 5-star ratings. This finding holds between all regions of the US (Midwest, East, South, West).



N-Gram Results – Popular Topics

The n-gram analysis was split into two tasks, a broad look at all tips and sushi reviews by star level, and a focused look per individual business's reviews organized by review rating and tips. The broad analysis demonstrated trending topics amongst the sample of restaurants, notably the word 'fresh' appears very often prompting a further look by region. Upon conclusion of a Chi-Squared test for difference in proportions, the Midwest does show a significantly lower percentage of positive ratings referencing 'fresh', $\chi^2(4,\ 168{,}293) = 15.707,\ p - value\ <\ 0.01$.



Overall, the broad analysis revealed a division of opinion on cuisine such as the spicy tuna roll, a celebration for options like green tea ice cream and miso soup, and a strong fixation on prompt service. Additionally, patrons tended to also prefer all-you-can-eat options and alcohol.

In the app, business owners can explore the n-gram tables for positive and negative reviews and tips for their business to discover new insights to achieve better ratings. To make this efficient for the shiny app, thousands of n-gram analysis were run outside the app, then stored efficiently for fast use.

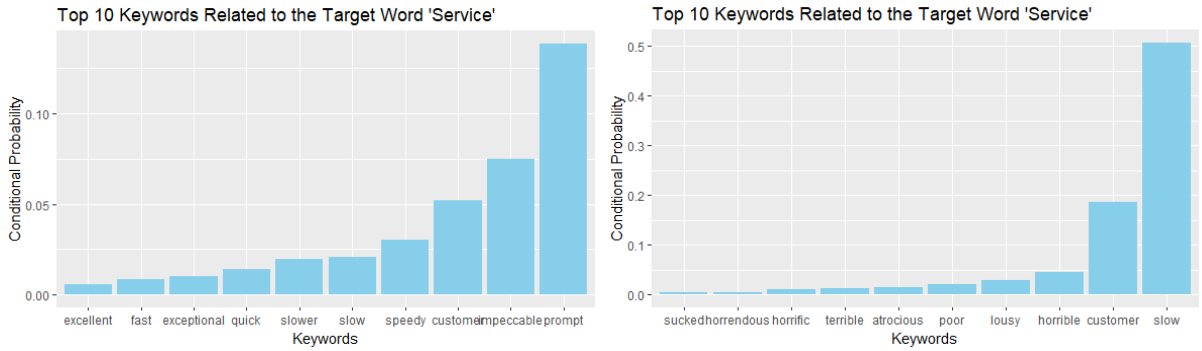
Latent Dirichlet Allocation Results – Unique Topics and Important Tips/Reviews

LDA identifies abstract topics and finds popular tokens and unique tokens that are used less but still carry strong importance. Since we use an n-gram analysis for popular tokens, LDA is used for finding tokens unique to businesses using a discriminative scoring method. We find rare topics not found in the n-gram analysis which we highlight in the shiny app per business and per review/tip dataset. The algorithm is then used to find which specific review or tip most embodies the abstract topics the algorithm finds. Thus, we have an efficient tool for showing business owners the statistically most relevant tips and reviews to their restaurant. To make the app efficient, thousands of pre-processed files are organized for the app to read a specific file.

Word2Vec Results

Service is a principal factor of customer satisfaction in all restaurants and sushi rolls are classic for any sushi restaurant. Our Word2Vec models are built to answer questions about why people love or hate the service or rolls in sushi restaurants.

In our case, good reviews have greater than three stars and bad reviews have less than three. First, we look at "service" as a context word and get the following results. The left plot presents the keywords for service in good reviews and the right plot presents bad reviews.

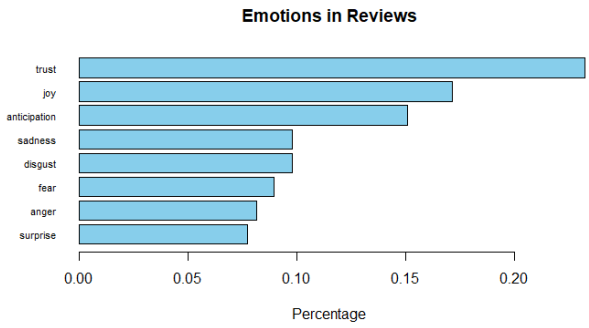


The horizontal axis is the top 10 frequently occurring words along with the target word. The vertical axis can be interpreted as conditional probabilities, that is, probability distributions of the keywords given the target word. For example, in bad reviews, once a review contains 'service,' 'slow' will occur near it with a probability greater than 0.5.

In the Word2Vec results for 'roll', we can notice most of the keywords are the names of various kinds of rolls. These words are informative, from which we can know what kind of rolls people love most and what kinds of rolls are underperforming in all sushi restaurants analyzed.

Sentiment Analysis Results

Sentiment analysis is the process of understanding the opinions of people about a subject. We use package 'tidytext' in R to calculate sentiment score for each individual sushi restaurant. The score is the average difference between the numbers of positive words and negative words in each review. A business with sentiment score of 2 means there are two more positive words than negative words on average in each of its reviews. The higher the sentiment score, the more satisfied the customers are with the restaurant.



We also call the National Research Council sentiment dictionary to calculate the presence of eight different emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, trust. Here is the percentage of the emotions in one business's reviews. Given these emotions for reviews, business owners can better improve their restaurant. For this business, the percentage of surprise emotion is significantly smaller than the other three positive emotions, so the business owner may consider doing something to make customers more surprised, like changing the decoration, etc.

Business Recommendations

Our analysis on frequently used words and phrases shows that freshness and timeliness is driving higher ratings. Additionally, references to all-you-can-eat sushi options appear frequently in both good and bad reviews, meaning restaurants should investigate their prices and options for this offering. Our app allows owners to explore topics specific to their business for decision making.

Freshness is a very frequent topic amongst reviewers. There is a difference between regions in positive ratings referencing freshness. The Midwest shows 5% fewer good ratings referencing freshness compared to all restaurants. We recommend investing in the sourcing of very fresh fish and ingredients to gain positive reviews, especially in inland states.

From our Word2Vec results, we conclude that patrons notice when things move slowly more often than quickly. Avoiding slowdowns will lessen the volume of 1 and 2-star reviews. Promptness in delivery or customer service was shown to increase high ratings, so acknowledging quick service will aid an establishment.

Through the same process, we found Rainbow rolls to be high performing and appear more often in positive reviews. Allocating menu space to detail this roll, either in words or pictures, will improve sales and, consequently, solidify high ratings. We also noticed that crab-based rolls, like Spider and California, tend to appear in negative reviews frequently. This tells us that restaurants should also allocate time and money into improving ingredients other than just fresh fish.

References & Contributions

Elad, B. (2022, May 18) *Yelp Statistics 2022 Demographics, Users and Facts*. Enterprise Apps Today. https://www.enterpriseappstoday.com/stats/yelp-statistics.html

Yelp (2022, Sept 30) *Fast Facts: Yelp Metrics as of September 30, 2022.* Yelp Inc. https://www.yelp-press.com/company/fast-facts/default.aspx

JB: Code: Text Cleaning, N-Grams – Popular Topics, LDA – Unique Topics + Most Relevant Quotes, Shiny App – Data Engineering, Topics Playground. Report: N-gram, LDA, Overall Edits

RS: Code: Data cleaning, Map visualization, Shiny App. GitHub organization. Report: Intro, Business Recommendations.

RZ: Code: Word2Vec, sentiment analysis. Report: Word2Vec, Sentiment Analysis.