Text Mining Group Project

OPIM-5671- Data Mining and Business Intelligence

Professor Sudip Bhattacharjee

Evaluating Eateries: A Comprehensive Analysis of Yelp Reviews

Group 3

Nimish Pastaria

Pranjal Jain

Pratik More

Priyanshu Kumar

Rabab Abbasi

Introduction:

This project dives into Yelp reviews to understand the customers' sentiments towards restaurants. By analyzing the language and expressions in these reviews, we aim to uncover what drives positive or negative sentiments. Whether it's about the food, the ambiance, or the service, we'll look into how these aspects influence a customer's overall impression. Our analysis will help identify patterns that explain why customers feel a certain way about their dining experiences. Understanding these sentiments is crucial for restaurants aiming to improve their service and for customers looking for the best dining experiences. Before we proceed, we'd like to highlight that we did all of our analysis in Python because of two reasons:

1) Most of us had macs and SAS runs slow on mac.
2) It was easier to work together using GCollab

We did use SAS and created some supervised models and achieved an accuracy of around 60% which can be found in the later part of the report. We moved to Python for most of our analysis implementing unsupervised learning by leveraging pre-trained python models and libraries to give us better actionable insights from our dataset.

Problem Statement:

Online reviews are a critical aspect of consumer decision-making in today's digital age, especially in the service industry. Understanding the sentiment of these reviews and predicting the ratings can provide invaluable insights for restaurants to improve their services and for consumers to make informed decisions. This project aims to analyze Yelp reviews to identify sentiment trends, and key aspects that influence customer satisfaction, and develop a model to predict the star ratings based on the review text.

Dataset:

1. Introduction To Dataset

DATA SOURCES

Yelp public dataset: A comprehensive dataset made open source by Yelp for academic and research purposes, containing reviews, user information, business metadata, and more.

https://www.yelp.com/dataset/download

The Yelp dataset contains the following key columns relevant to our project:

Column Name	Description
Review ID	A unique identifier for each review.
User ID	The ID of the user who wrote the review.
Business ID	The ID of the restaurant being reviewed.
Stars	The rating given by the user, ranging from 1 to 5.
Text	The full text of the review.
Cool	Engagement metrics indicating how other users perceived the review.
Funny	Engagement metrics indicating how other users perceived the review.
Useful	Engagement metrics indicating how other users perceived the review.
Date	The date when the review was posted.

2. Data Exploration and Modification

Raw Data

Our primary dataset is a subset of the Yelp Dataset Challenge, consisting of ~14,000 reviews. This dataset included information such as the review text, user ratings, restaurants, and more as described above. The reviews covered 5 restaurants, offering a comprehensive overview of customer feedback along with star ratings, which served as a proxy for customer sentiment.

The revised Yelp dataset contains the following key columns relevant to our project:

Column Name	Description
Business ID	The ID of the restaurant being reviewed.
Stars	The rating given by the user, ranging from 1 to 5.
Text	The full text of the review.
Date	The date when the review was posted.

Data Preprocessing

Before we began any type of data processing, we sliced data to find the top restaurants which had the most number of text reviews. Subsequently, the sliced file underwent extensive preprocessing, including tokenization, stop words removal, lowercasing, and optional steps like stemming and lemmatization, resulting in four distinct versions of the processed data. Following are the 4 variations:

i) Tokenization, lowercasing, and removal of stop words.
ii) Tokenization, lowercasing, removal of stop words, and stemming.
iii) Tokenization, lowercasing, removal of stop words, and lemmatization.
iv) Tokenization, lowercasing, removal of stop words, lemmatization, and stemming.

We decided to proceed with only lemmatization as it provided more meaningful word reductions for sentiment analysis and topic modeling. As part of pre-processing, we followed 6 steps as shown below to prepare our data for analysis:

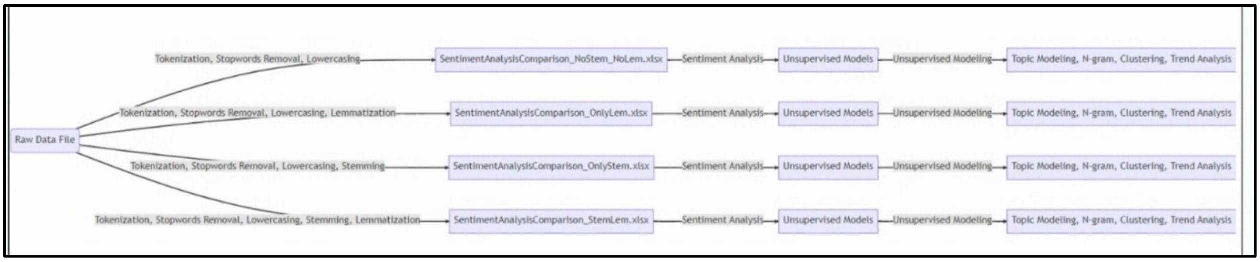
Data Loading: We start by importing the necessary libraries such as pandas for data manipulation, NLTK for natural language processing, and Spacy for advanced linguistic features.

Data Cleaning: The raw review text often contains noise in the form of punctuation, special characters, and irrelevant words like stop words. Our first task is to clean this text, preserving only the meaningful content. Using regular expressions, we strip away any non-alphabetical characters and convert the text to lowercase to maintain uniformity. Tokenization: The cleaned text is then tokenized, which means breaking it down into individual terms or words. This is crucial for the subsequent steps where we need to analyze these terms separately.

Stop Words Removal: Common words that do not contribute to the sentiment of the review are filtered out. To this end, we have compiled a custom list of stop words that includes typical stop words from the NLTK library along with domain-specific terms that are frequent in our dataset but do not hold significant sentiment value.

Stemming and Lemmatization: In our code, we have the functionality to perform stemming and lemmatization which are techniques to reduce words to their root form. For instance, "running" would be reduced to "run". However, based on our preliminary analysis, we are currently opting for lemmatization over stemming, as it provides more meaningful word reductions.

Applying Preprocessing: After cleaning and preprocessing functions were applied to the text column of our dataset, we created new column called 'Preprocessed_Text' with the preprocessed text which we then used for our sentiment classification and theme detection.
The pre-processing ensured that our analysis would be based on the specific information drawn from the Yelp reviews and the variations aimed to evaluate their impact on the accuracy of sentiment analysis and topic modeling. Refer below for a visual depiction of the same:



File Naming Convention: Files were named according to the preprocessing techniques applied, e.g., SentimentAnalysisComparison_NoStem_NoLem.xlsx for data without stemming or lemmatization.

```
Dataset shape: (33064, 11)
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 33064 entries, 0 to 33063
Data columns (total 11 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Restaurant Name     33064 non-null  object
 1   Location            33064 non-null  object
 2   Text                33064 non-null  object
 3   stars               33064 non-null  int64
 4   Sentiment           33064 non-null  object
 5   Date                33064 non-null  datetime64[ns]
 6   Preprocessed_Text   33064 non-null  object
 7   Day                 33064 non-null  int64
 8   Month               33064 non-null  int64
 9   Year                33064 non-null  int64
 10  MonthYear           33064 non-null  object
dtypes: datetime64[ns](1), int64(4), object(6)
memory usage: 2.8+ MB
None
              stars           Day         Month          Year
count  33064.000000  33064.000000  33064.000000  33064.000000
mean       4.009920     15.814693      6.443897   2016.026131
std        1.220121      8.863494      3.409282      3.026041
min        1.000000      1.000000      1.000000   2005.000000
25%        3.000000      8.000000      3.000000   2014.000000
50%        4.000000     16.000000      6.000000   2016.000000
75%        5.000000     24.000000      9.000000   2018.000000
max        5.000000     31.000000     12.000000   2022.000000
Positive    24543
Negative     4596
Neutral      3925
Name: Sentiment, dtype: int64
```

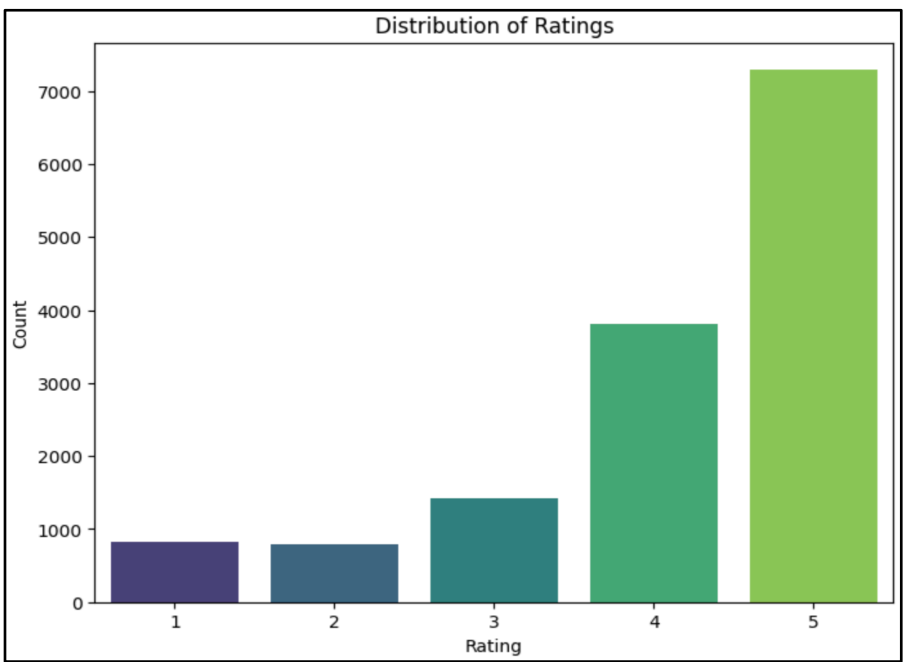
Please note that we did not find any missing values or outliers in our dataset.
Please also note that we created new columns for 'Day', 'Month', and 'Year' by extracting these values from the 'Date' column to help in our trend analysis later.

Exploratory Data Analysis (EDA)

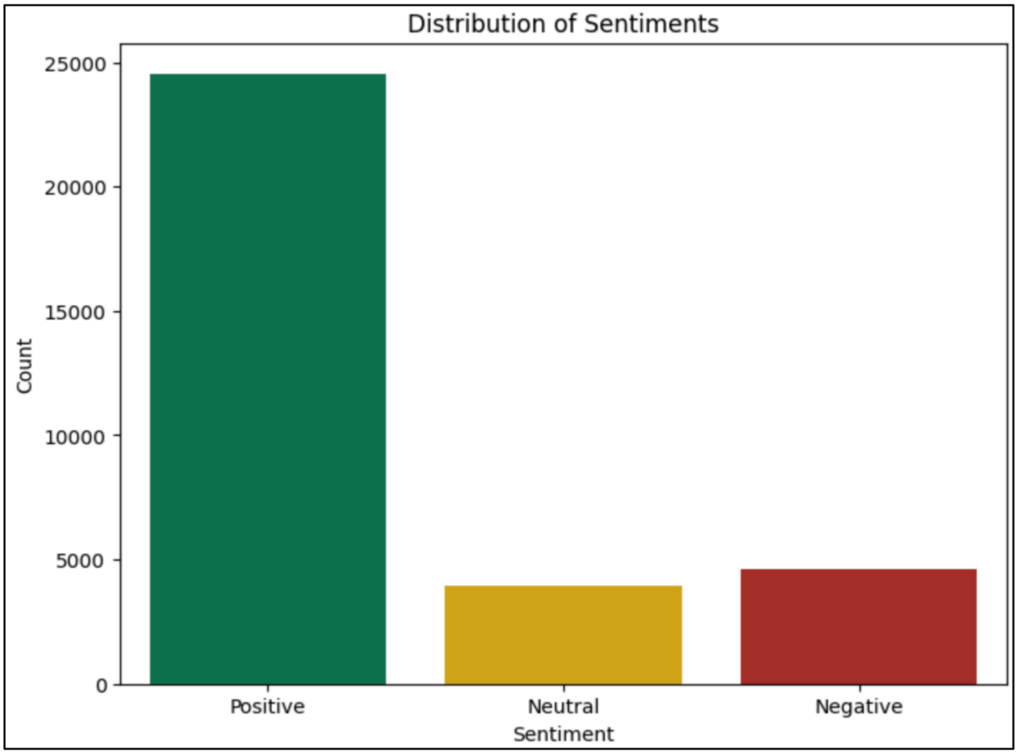
We began our EDA by looking at the descriptive and summary statistics of the processed dataset to gauge the data structure, distribution of variables and correlations between variables. This included,
1) Basic Information and Statistics to get a basic understanding of the dataset's structure and content
2) Shape of the dataset
3) Descriptive statistics for numeric columns
4) Descriptive statistics for the 'sentiment' column
5) Distribution of ratings
6) Visualizing the distribution of sentiments
8) Calculating the average ratings by restaurant
9) Counting the number of reviews per restaurant
10) Analyzing the distribution of sentiment per restaurant
11) Identifying top restaurants by positive sentiment
12) Identifying top restaurants by negative sentiment
13) Graph of top restaurants by sentiment percentage
14) Distribution of Reviews by Location
15) Calculating average ratings by location
16) Sentiment analysis by location

Some of the results are as shown below:

- Overall distribution of Ratings:

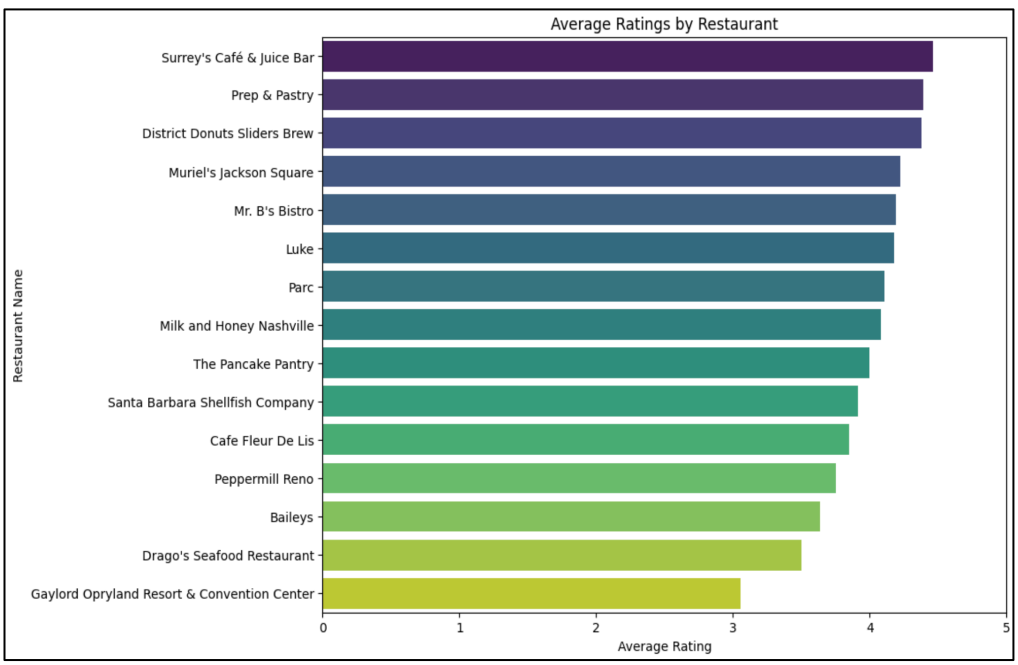


- Overall Distribution of Sentiments:

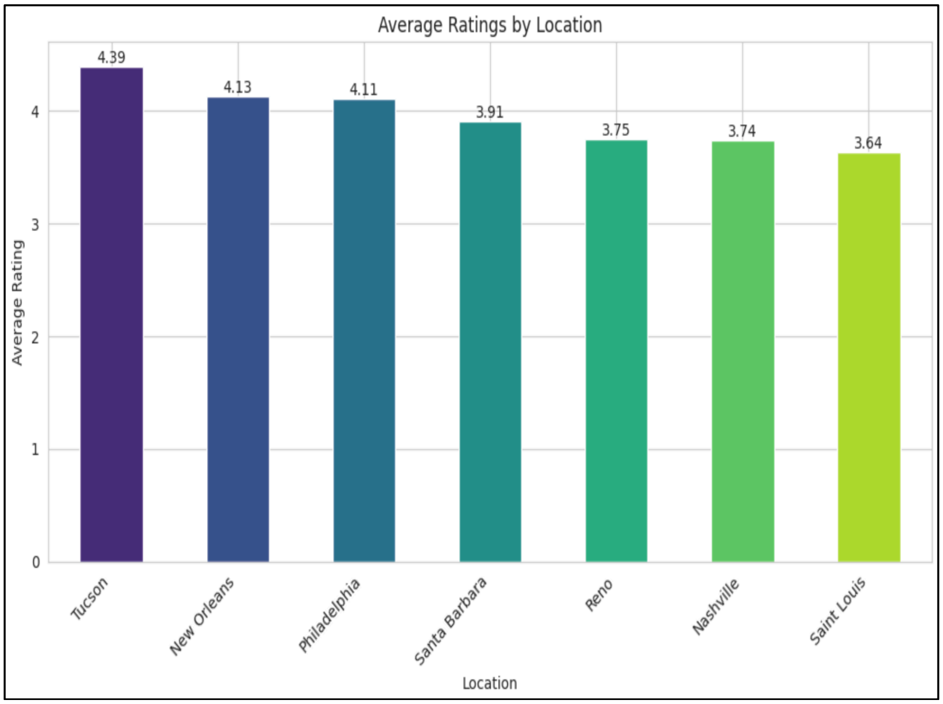


- Average Ratings by Restaurants:

Restaurants with high average ratings are likely meeting or exceeding customer expectations, while those with lower averages may struggle with consistency, quality, or service.



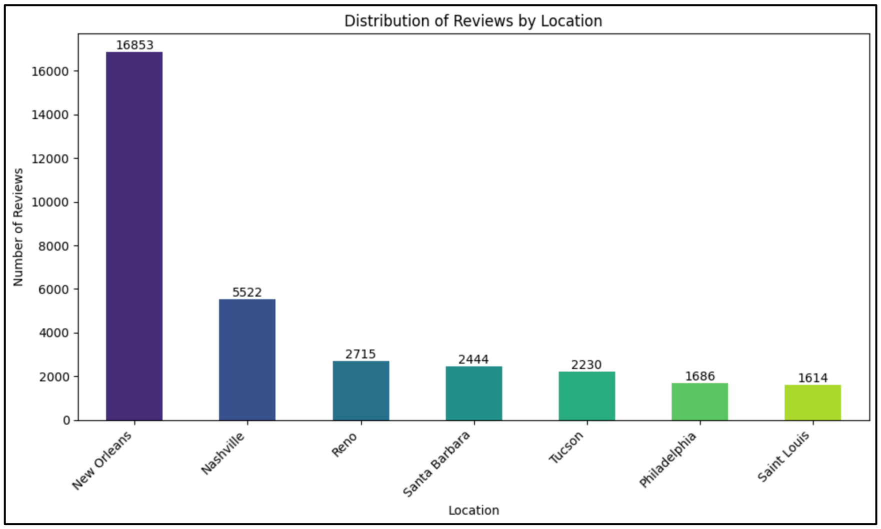
- Average Ratings by Location:



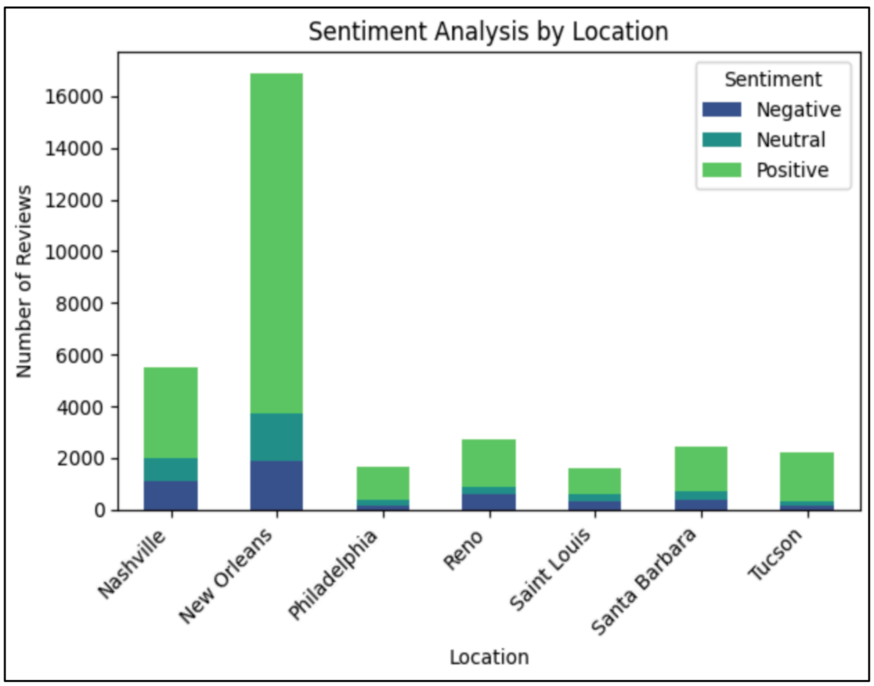
- Number of Review per Restaurant

```
Luke                                            4661
Peppermill Reno                                 2715
Santa Barbara Shellfish Company                 2444
Prep & Pastry                                   2230
Surrey's Café & Juice Bar                       2120
Mr. B's Bistro                                  2116
District Donuts Sliders Brew                    2110
The Pancake Pantry                              2091
Muriel's Jackson Square                         2008
Drago's Seafood Restaurant                      1947
Cafe Fleur De Lis                               1891
Milk and Honey Nashville                        1761
Parc                                            1686
Gaylord Opryland Resort & Convention Center     1670
Baileys                                         1614
Name: Restaurant Name, dtype: int64
```

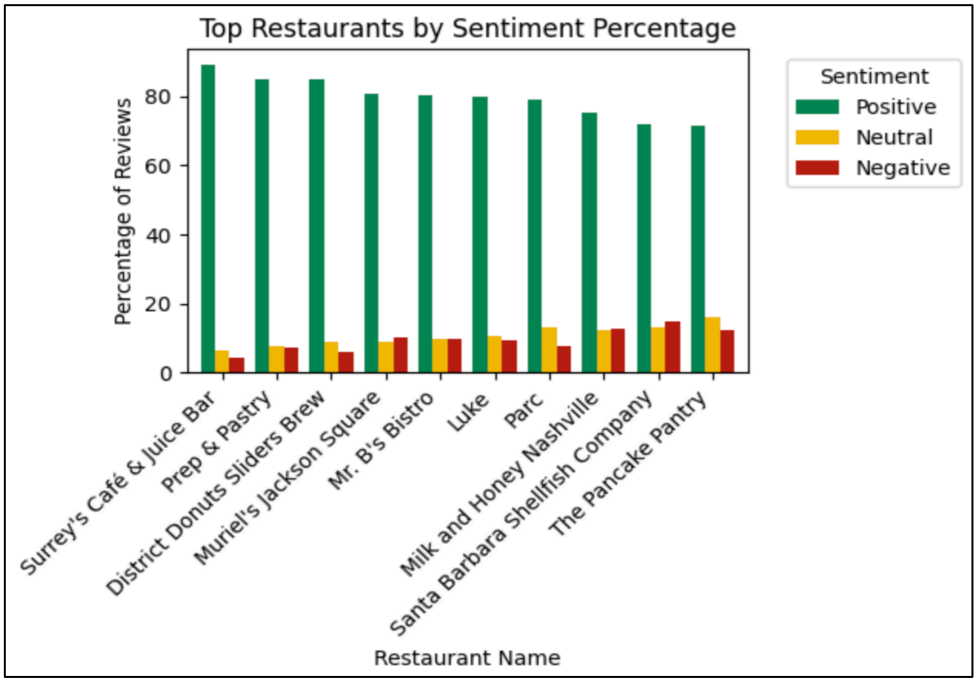
- Distribution of Reviews by Location



- Sentiment Analysis by Location:



- Sentiment Analysis by Restaurants:



Analysis and Approach:

<u>Unsupervised Learning</u>:

1) The sentiments derived from ratings would be compared to the sentiments generated using inbuilt NLTK and TextBlob Python libraries that come with pre-trained sentiment analysis models. These pre-trained models are trained on large datasets, making them capable of providing reasonable baseline sentiment results for various types of text in our processed dataset. Please note these libraries created additional sentiment columns to our original sentiment column.

2) We'll then compare the actual sentiments from the original file to the results of NLTK and TextBlob. Star Ratings as Sentiment Indicators where Star ratings were categorized into Negative (1-2 stars), Neutral (3 stars), and Positive (4-5 stars) sentiments.

Preliminary Analysis to get a sense of direction:

<u>Word Clouds</u>: Generated to visualize frequently mentioned words across different sentiments and restaurants, highlighting common themes and concerns.
<u>N-Gram Analysis</u>: Explored common phrases that were referred together to further understand customer feedback patterns.

Modeling:

Text Similarity and Clustering: Grouped reviews into clusters using Topic Modeling and KMeans clustering to discover underlying patterns and similarities among them.
Trend Analysis: Examined how certain topics, words, or sentiments trended over time, offering insights into changing customer preferences
We'll then be using the outputs to derive insights into what drove the customers to feel a certain way towards the restaurant.

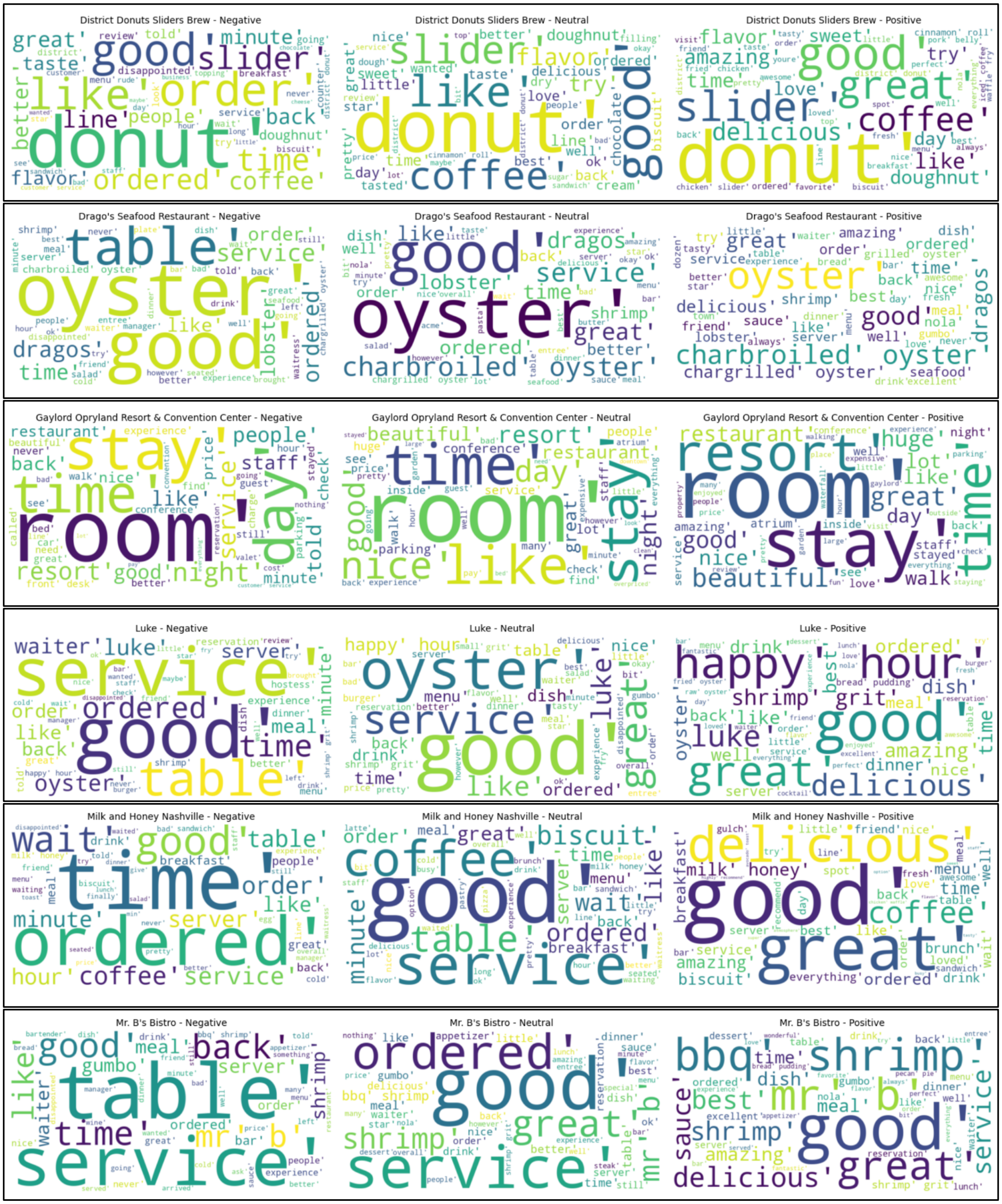
Supervised Learning:

Running the preprocessed data in SAS Text Miner to generate text topics and supervised models that predicted the sentiment of reviews with a certain accuracy.

Unsupervised Learning:

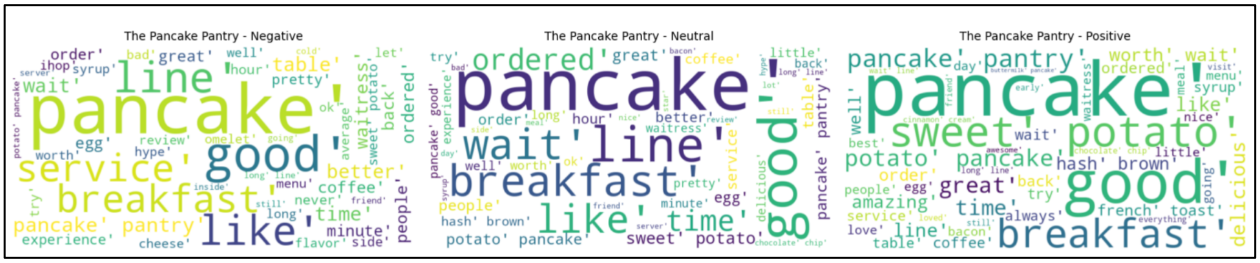
WordCloud Results (Per Restaurant aggregated by sentiment):

Having our data explored and preprocessed, we started our initial analysis using word clouds to deduce significant words from the text reviews. Using graphical representation to determine the prominence of words in the reviews, the bigger sized words in the cloud meant that the term appeared more frequently whereas the smaller sized terms words were not used as frequently. We generated word clouds for each of our restaurants segregated by sentiment some of which are shown on the slides here. We cap the word count at 50 to focus on the most relevant terms.
By visualizing the data in this way, we gained quick, intuitive insights. We can easily get to know which words are most frequently associated with positive or negative experiences. Also, this analysis based on word frequency gave us some information on the common themes and concerns raised by customers, guiding us in understanding the driving factors behind customer satisfaction or dissatisfaction.



District Donuts Sliders Brew - Negative
great' good' slider' minute' taste' like' order' line' people' back' better' donut' time' flavor' ordered' coffee' doughnut' disappointed' breakfast' told' review' service' wait' try'
District Donuts Sliders Brew - Neutral
nice' slider' better' doughnut' flavor' ordered' like' donut' good' coffee' sweet' little' taste' delicious' try' love' order' line' well' time' day' tasted' best' ok' back' cream' chocolate' wanted' star' pretty'
District Donuts Sliders Brew - Positive
flavor' sweet' good' amazing' try' time' great' slider' love' coffee' delicious' donut' day' like' doughnut' cinnamon' roll'
Drago's Seafood Restaurant - Negative
table' order' service' oyster' good' like' dragos' time' lobster' ordered' charbroiled' oyster' shrimp' told' back' manager'
Drago's Seafood Restaurant - Neutral
dish' like' good' dragos' well' service' back' lobster' time' order' oyster' shrimp' great' ordered' better' charbroiled' oyster' chargrilled' oyster'
Drago's Seafood Restaurant - Positive
try' great' amazing' dish' order' ordered' service' oyster' time' back' nice' shrimp' best' delicious' good' sauce' friend' like' lobster' server' well' nola' charbroiled' oyster' chargrilled' oyster' seafood' dragos'
Gaylord Opryland Resort & Convention Center - Negative
restaurant' stay' people' back' staff' time' price' nice' like' room' day' service' resort' good' night' minute' told' check' parking'
Gaylord Opryland Resort & Convention Center - Neutral
beautiful' resort' time' day' restaurant' conference' room' good' stay' great' nice' like' night' walk' parking' check' find'
Gaylord Opryland Resort & Convention Center - Positive
restaurant' resort' huge' room' lot' like' great' time' day' stay' good' nice' amazing' atrium' inside' staff' stayed' beautiful' walk' night'
Luke - Negative
waiter' luke' server' service' ordered' good' meal' minute' time' order' like' back' table' oyster' dish' experience'
Luke - Neutral
happy' hour' table' nice' oyster' luke' service' menu' dish' good' great' back' drink' time' like' ordered'
Luke - Positive
drink' ordered' happy' hour' dish' shrimp' grit' best' good' oyster' back' like' luke' amazing' great' dinner' nice' delicious' time'
Milk and Honey Nashville - Negative
wait' good' table' drink' breakfast' time' order' like' minute' server' ordered' great' hour' coffee' service' back'
Milk and Honey Nashville - Neutral
order' meal' great' biscuit' coffee' time' good' menu' wait' like' minute' table' ordered' breakfast' service' back'
Milk and Honey Nashville - Positive
delicious' little' friend' nice' milk' honey' good' menu' breakfast' coffee' time' great' best' like' brunch' amazing' biscuit' service' ordered' drink' wait'
Mr. B's Bistro - Negative
good' drink' meal' back' table' people' server' gumbo' like' waiter' time' ordered' mr' b' service' shrimp' order' experience'
Mr. B's Bistro - Neutral
ordered' appetizer' little' dinner' like' good' sauce' gumbo' bbq' shrimp' meal' great' shrimp' service' mr' b' table' dish' time'
Mr. B's Bistro - Positive
bbq' time' shrimp' dessert' drink' back' dish' mr' b' best' meal' good' sauce' shrimp' amazing' delicious' great' service' excellent'

Muriel's Jackson Square - Negative
Muriel's Jackson Square - Neutral
Muriel's Jackson Square - Positive
Parc - Negative
Parc - Neutral
Parc - Positive
Peppermill Reno - Negative
Peppermill Reno - Neutral
Peppermill Reno - Positive
Prep & Pastry - Negative
Prep & Pastry - Neutral
Prep & Pastry - Positive
Santa Barbara Shellfish Company - Negative
Santa Barbara Shellfish Company - Neutral
Santa Barbara Shellfish Company - Positive
Surrey's Café & Juice Bar - Negative
Surrey's Café & Juice Bar - Neutral
Surrey's Café & Juice Bar - Positive



Unsupervised Learning: N-Gram Analysis

In the case of our project, a "bi-gram" analysis means we're looking at every two words that come one after another in restaurant reviews and a “tri-gram” means looking at three words in consequence.
From the previous word cloud, looking at just one word based on it's frequency doesn't tell us much. Like, if the word "cold" is mentioned you wouldn't know if we’re talking about the weather, the food, or the service. But if we heard "cold pancake," that'll provide a much better context.
In our analysis, we generated and inspected the top 10 bi-grams for each restaurant and sentiment category. By doing this, we could pinpoint recurring themes or specific issues that customers talk about. For example, if we see "not clean" a lot in different reviews for a restaurant, it told us that cleanliness is a problem. It helped us spot patterns—like if "tasty donut" pops up often, that might be what the restaurant is famous for.
By doing this with lots of reviews, we got a better picture of what's good or bad at each place.
The outcome of this analysis informed us not only of the most mentioned topics but also how certain attributes were perceived together, painting a more nuanced picture of the customer experience. This level of detail can potentially provide actionable insights for restaurant management, like which areas to focus on for improvement or to highlight in marketing materials.

- Top 10 bi-grams for each restaurant and sentiment:

- Top 10 tri-grams for each restaurant and sentiment:

Unsupervised Learning: Clustering (Using TfidfVectorizer and K-Means)

Clustering_Restr_Sentiment.docx

Sample restaurant output shown below:

Restaurant: Peppermill Reno, Sentiment: Positive

```
Top terms per cluster:
Cluster 0: room | tower | tuscany | bed | nice | bathroom | suite | pool | stay | stayed
Cluster 1: spa | love | pool | amazing | room | always | stay | great | best | day
Cluster 2: great | peppermill | time | service | room | stay | experience | back | thank | staff
Cluster 3: casino | good | best | peppermill | like | great | room | lot | game | time
Cluster 4: nice | clean | room | great | staff | stay | friendly | good | pool | casino
```

Restaurant: Peppermill Reno, Sentiment: Neutral

```
Top terms per cluster:
Cluster 0: buffet | casino | great | good | restaurant | nice | pool | clean | staff | room
Cluster 1: room | nice | bed | suite | like | good | casino | great | peppermill | tub
Cluster 2: room | time | peppermill | stay | back | like | tower | great | night | year
Cluster 3: casino | room | peppermill | walk | floor | machine | slot | sign | nice | staff
Cluster 4: room | housekeeping | service | ready | night | stay | day | give | stayed | kid
```

Restaurant: Peppermill Reno, Sentiment: Negative

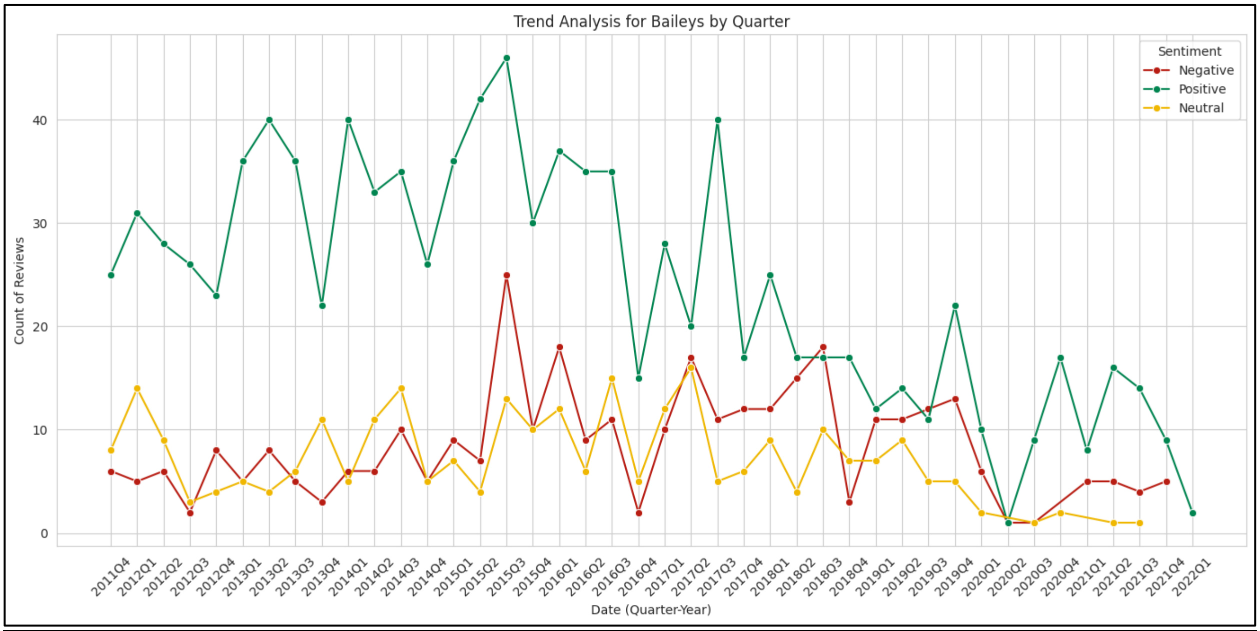
```
Top terms per cluster:
Cluster 0: casino | time | slot | machine | peppermill | like | dealer | drink | better | table
Cluster 1: room | day | never | stay | peppermill | service | casino | great | time | people
Cluster 2: fee | charge | room | service | customer | resort | spa | day | like | back
Cluster 3: room | smoke | casino | smell | like | wing | north | tower | elevator | peppermill
Cluster 4: room | desk | front | told | stay | night | back | check | time | suite
```

<u>Unsupervised Learning: TOPIC MODELING</u> (Top 5 topics for each restaurant and each sentiment)
Sample output below:

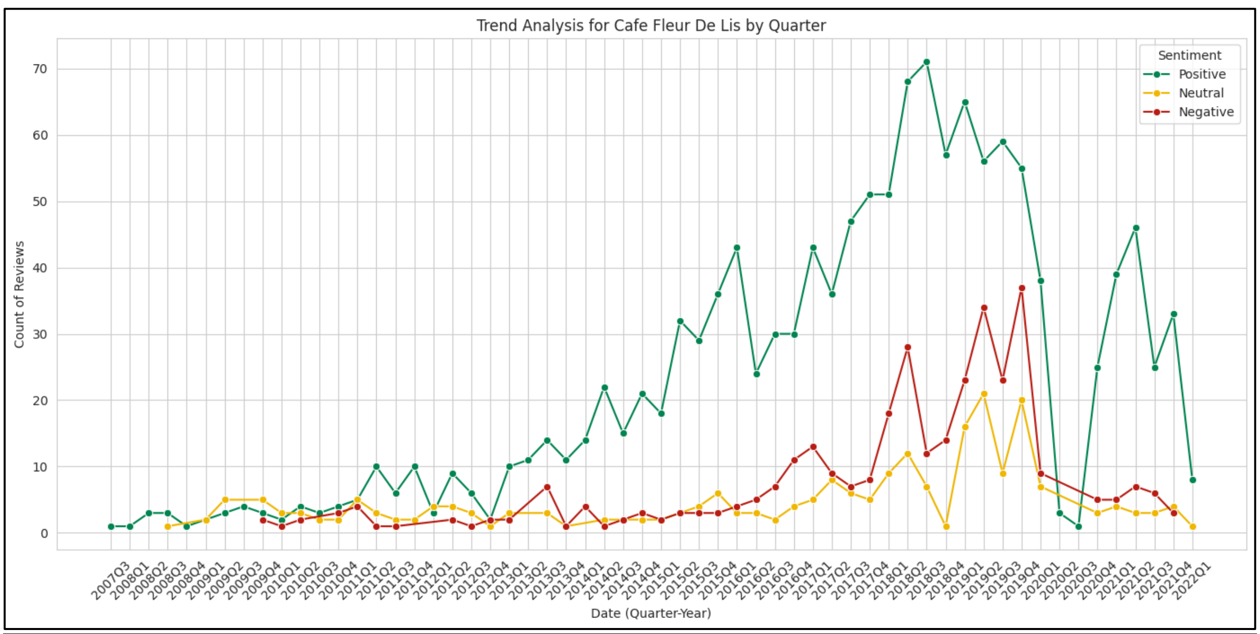
```
Restaurant: Peppermill Reno, Sentiment: Positive
Topic 0: 0.022*"peppermill" + 0.019*"room" + 0.012*"stay" + 0.010*"pool" + 0.010*"time" + 0.009*"love" + 0.009*"casino" + 0.009*"great" + 0.007*"nice" + 0.006*"service"
Topic 1: 0.028*"room" + 0.013*"spa" + 0.012*"peppermill" + 0.011*"time" + 0.008*"stay" + 0.008*"great" + 0.006*"nice" + 0.006*"service" + 0.006*"amazing" + 0.006*"casino"
Topic 2: 0.009*"peppermill" + 0.008*"room" + 0.007*"time" + 0.006*"casino" + 0.005*"review" + 0.005*"pool" + 0.005*"stayed" + 0.004*"service" + 0.004*"back" + 0.004*"experience"
Topic 3: 0.029*"room" + 0.013*"great" + 0.012*"nice" + 0.012*"casino" + 0.010*"peppermill" + 0.010*"time" + 0.009*"pool" + 0.009*"good" + 0.009*"stay" + 0.008*"service"
Topic 4: 0.020*"room" + 0.014*"peppermill" + 0.009*"stay" + 0.009*"like" + 0.008*"casino" + 0.007*"great" + 0.007*"service" + 0.006*"nice" + 0.005*"time" + 0.005*"clean"

Restaurant: Peppermill Reno, Sentiment: Neutral
Topic 0: 0.017*"room" + 0.009*"great" + 0.008*"casino" + 0.008*"time" + 0.006*"like" + 0.005*"clean" + 0.005*"good" + 0.004*"peppermill" + 0.004*"nice" + 0.004*"service"
Topic 1: 0.008*"stay" + 0.008*"room" + 0.008*"nice" + 0.006*"told" + 0.005*"check" + 0.005*"found" + 0.005*"end" + 0.005*"great" + 0.004*"time" + 0.004*"staff"
Topic 2: 0.011*"room" + 0.010*"good" + 0.010*"casino" + 0.009*"service" + 0.008*"time" + 0.008*"peppermill" + 0.007*"like" + 0.006*"great" + 0.006*"nice" + 0.006*"customer"
Topic 3: 0.032*"room" + 0.008*"time" + 0.008*"nice" + 0.007*"casino" + 0.007*"stay" + 0.006*"peppermill" + 0.006*"like" + 0.006*"bed" + 0.006*"good" + 0.006*"night"
Topic 4: 0.021*"room" + 0.014*"peppermill" + 0.008*"casino" + 0.007*"nice" + 0.007*"night" + 0.007*"like" + 0.006*"time" + 0.006*"pool" + 0.006*"stay" + 0.005*"good"

Restaurant: Peppermill Reno, Sentiment: Negative
Topic 0: 0.024*"room" + 0.009*"night" + 0.009*"peppermill" + 0.009*"service" + 0.008*"back" + 0.008*"time" + 0.007*"casino" + 0.007*"stay" + 0.006*"call" + 0.006*"told"
Topic 1: 0.012*"peppermill" + 0.011*"room" + 0.010*"casino" + 0.010*"like" + 0.008*"time" + 0.007*"service" + 0.006*"stay" + 0.005*"back" + 0.005*"table" + 0.005*"people"
Topic 2: 0.014*"room" + 0.008*"peppermill" + 0.007*"time" + 0.006*"told" + 0.005*"day" + 0.005*"like" + 0.005*"nice" + 0.004*"back" + 0.004*"night" + 0.004*"floor"
Topic 3: 0.033*"room" + 0.008*"like" + 0.006*"stay" + 0.006*"peppermill" + 0.005*"time" + 0.005*"fee" + 0.005*"casino" + 0.005*"resort" + 0.005*"service" + 0.005*"back"
Topic 4: 0.034*"room" + 0.010*"desk" + 0.009*"front" + 0.009*"told" + 0.009*"time" + 0.008*"stay" + 0.007*"back" + 0.007*"check" + 0.006*"like" + 0.005*"night"
```

Topic_Modeling.docx

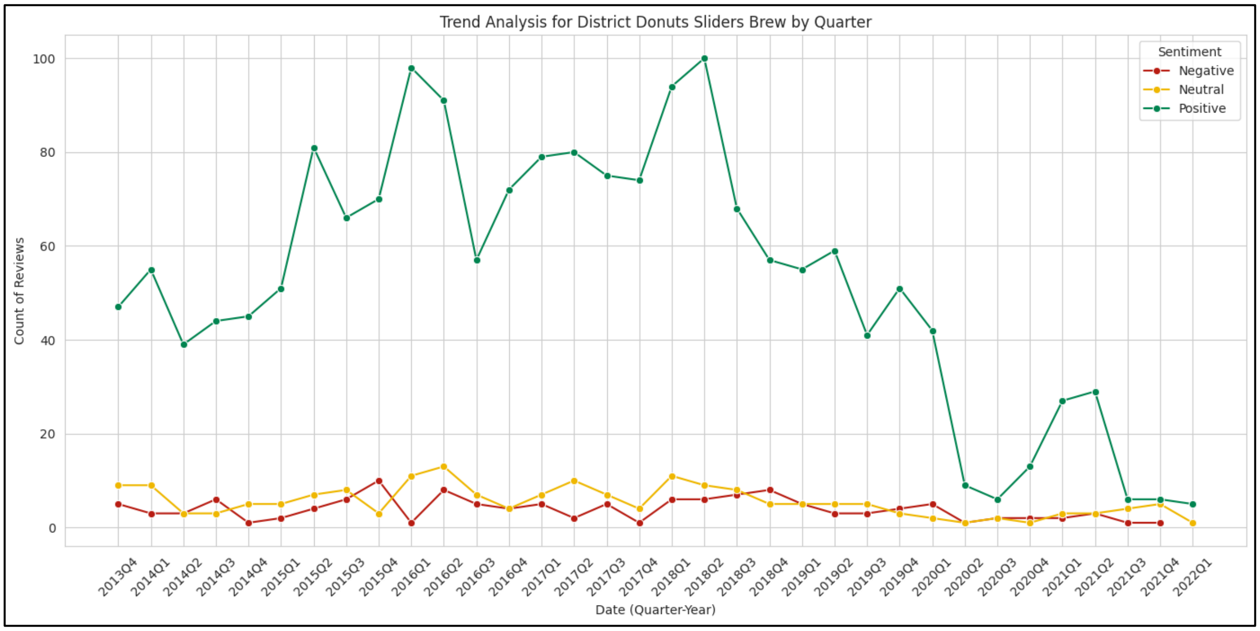
<u>Trend Analysis Over Time:</u>

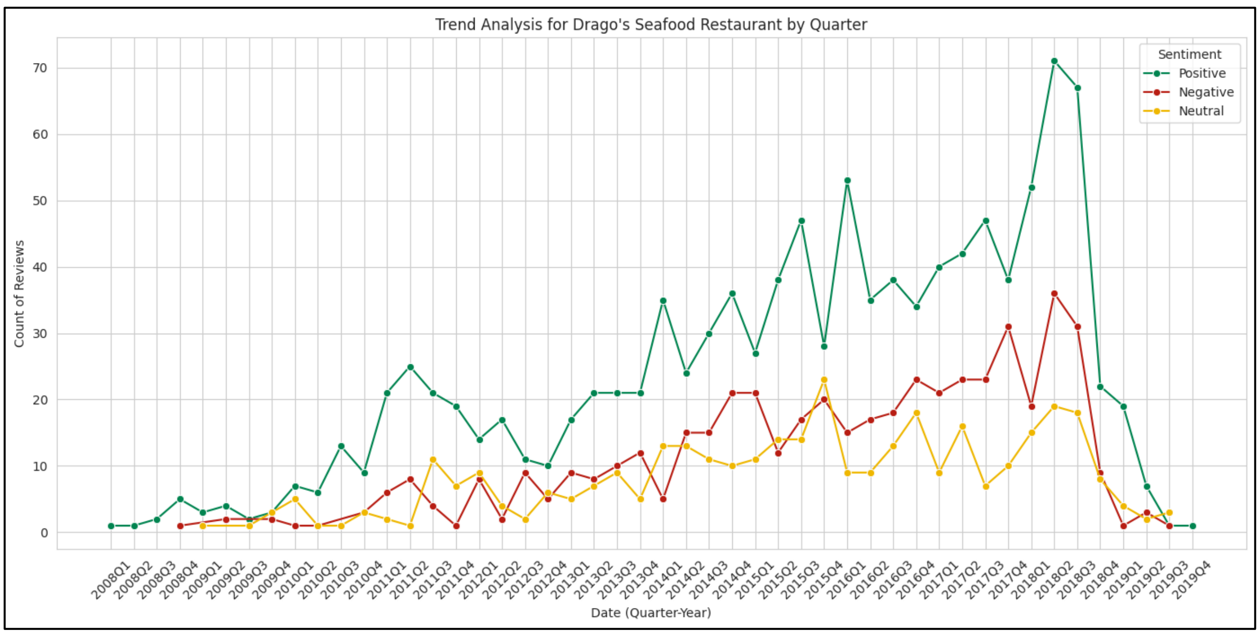


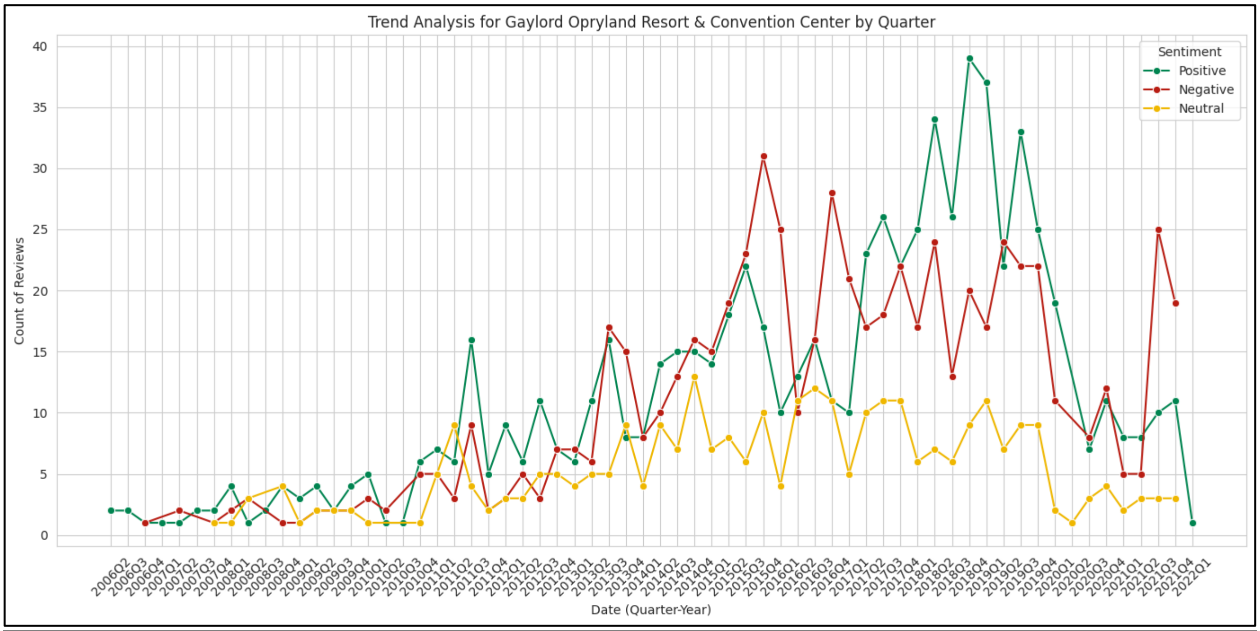
Trend Analysis for Baileys by Quarter
Count of Reviews
Date (Quarter-Year)
Sentiment
Negative
Positive
Neutral



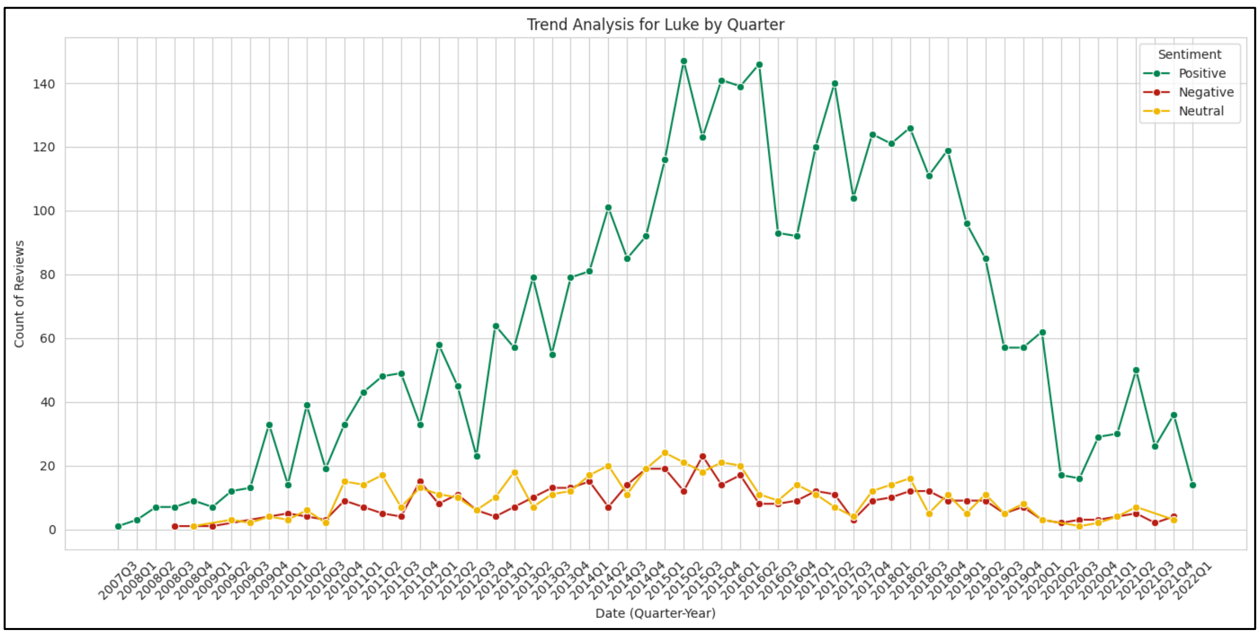
Trend Analysis for Cafe Fleur De Lis by Quarter
Count of Reviews
Date (Quarter-Year)
Sentiment
Positive
Neutral
Negative


Trend Analysis for District Donuts Sliders Brew by Quarter
Sentiment
Negative
Neutral
Positive
Count of Reviews
Date (Quarter-Year)

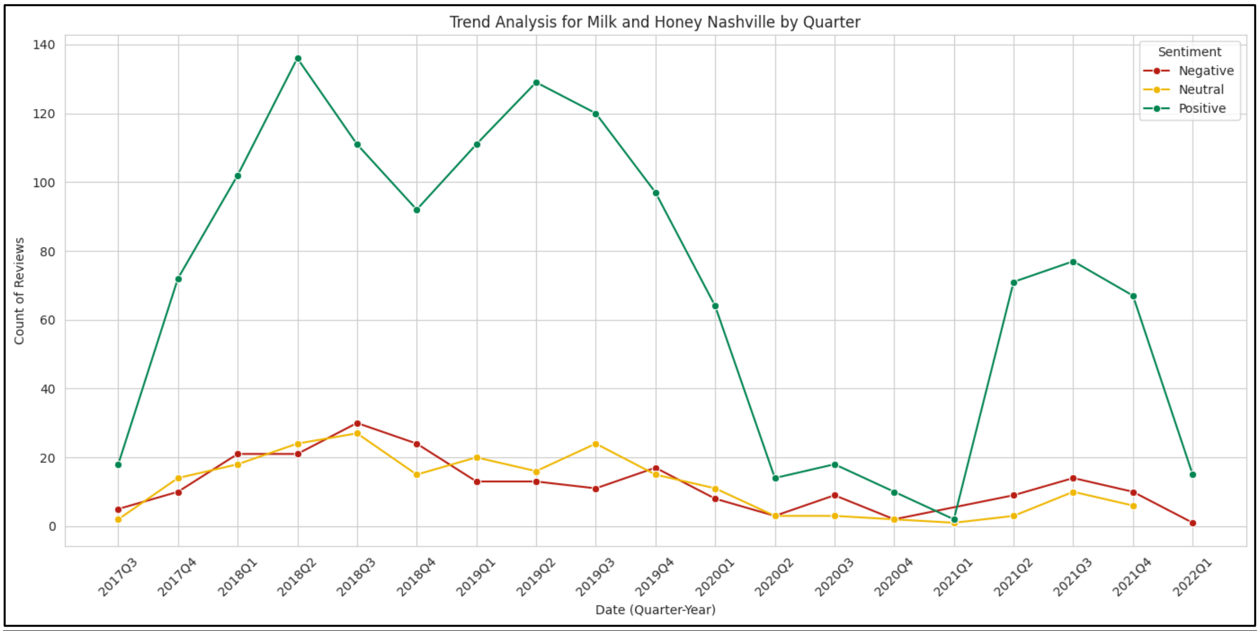

Trend Analysis for Drago's Seafood Restaurant by Quarter
Sentiment
Positive
Negative
Neutral
Count of Reviews
Date (Quarter-Year)

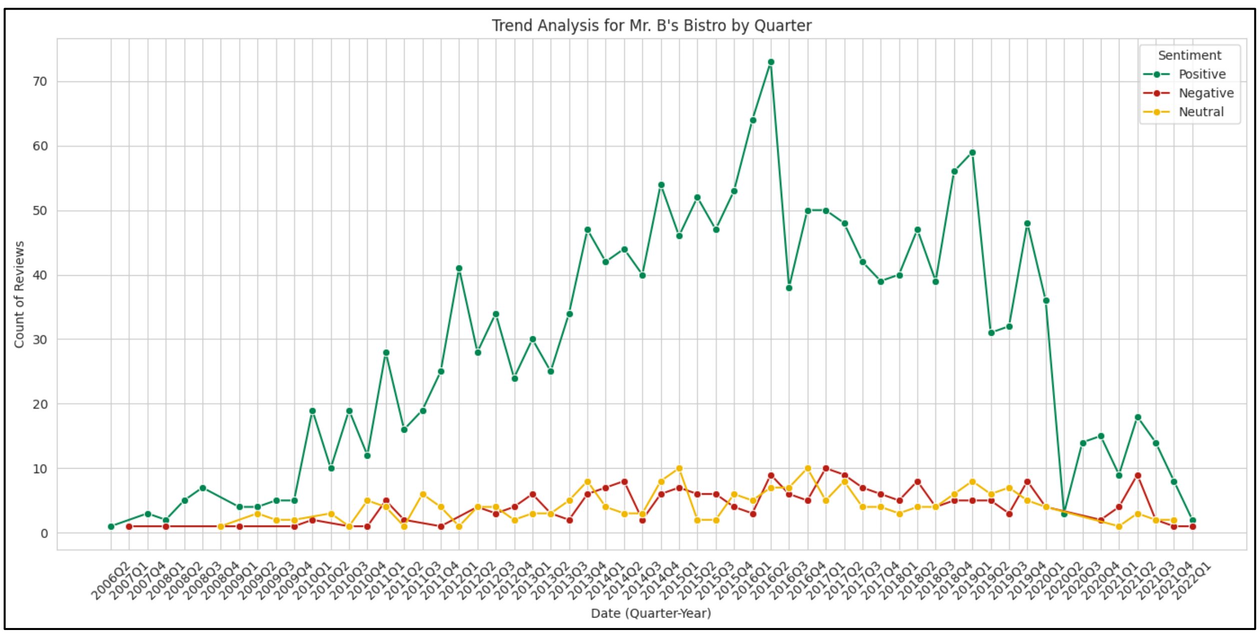


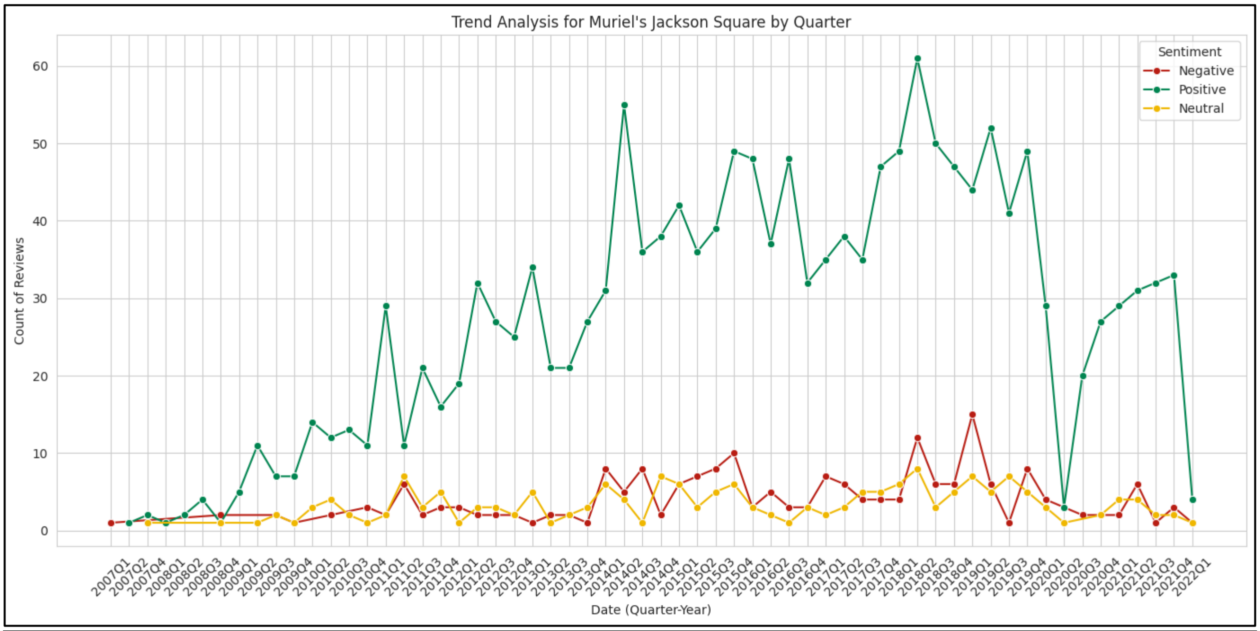
Trend Analysis for Gaylord Opryland Resort & Convention Center by Quarter
Sentiment
Positive
Negative
Neutral
Count of Reviews
Date (Quarter-Year)

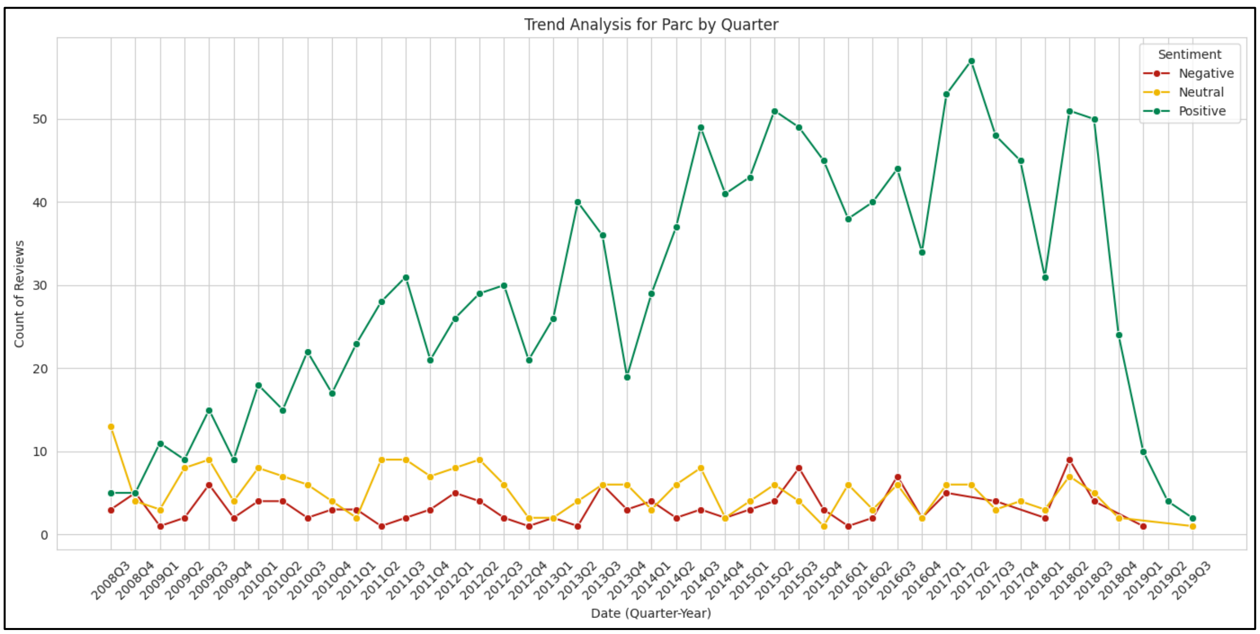


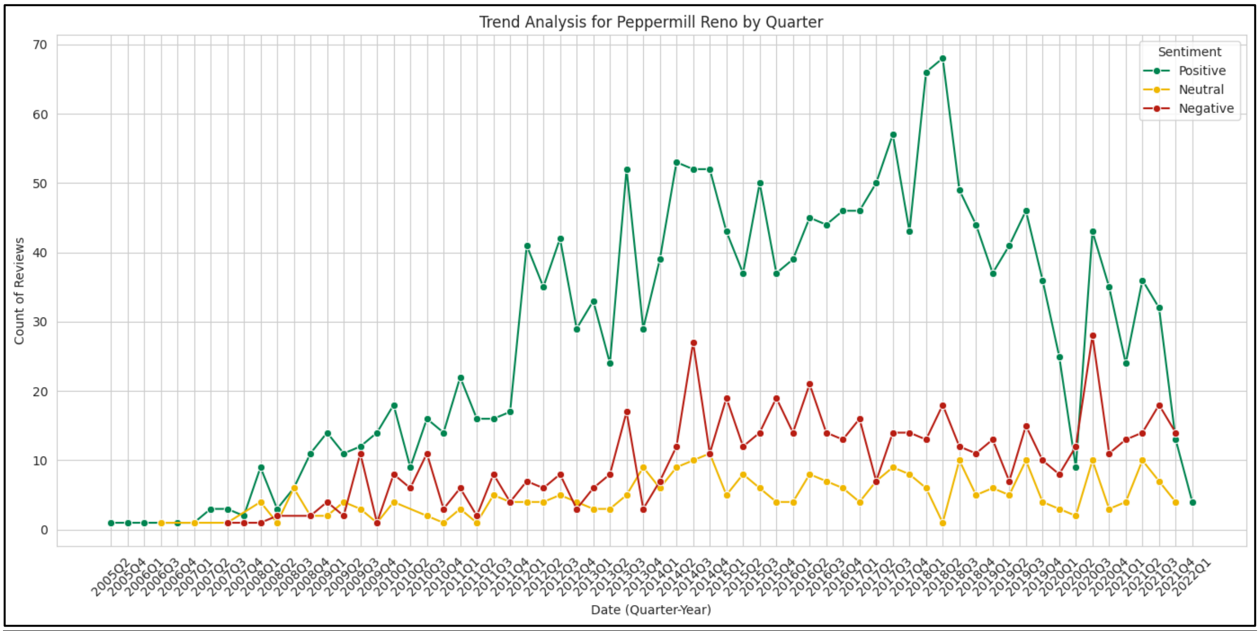
Trend Analysis for Luke by Quarter
Sentiment
Positive
Negative
Neutral
Count of Reviews
Date (Quarter-Year)


Trend Analysis for Milk and Honey Nashville by Quarter
Sentiment
Negative
Neutral
Positive
Count of Reviews
Date (Quarter-Year)

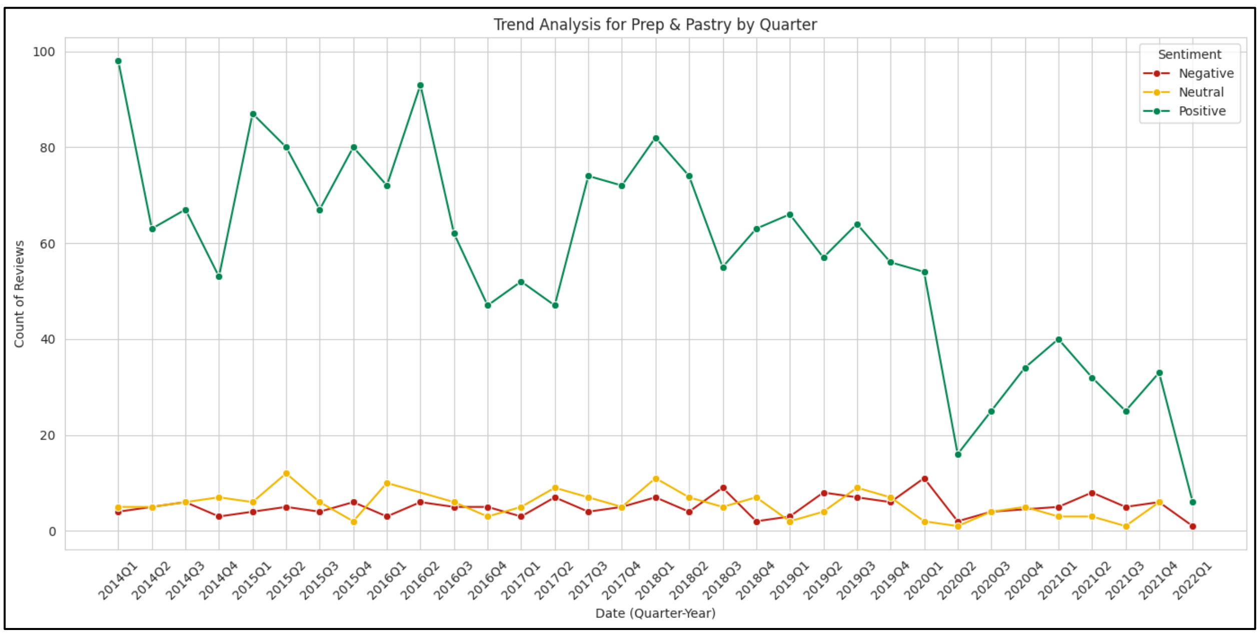

Trend Analysis for Mr. B's Bistro by Quarter
Sentiment
Positive
Negative
Neutral
Count of Reviews
Date (Quarter-Year)


Trend Analysis for Muriel's Jackson Square by Quarter
Sentiment
Negative
Positive
Neutral
Count of Reviews
Date (Quarter-Year)

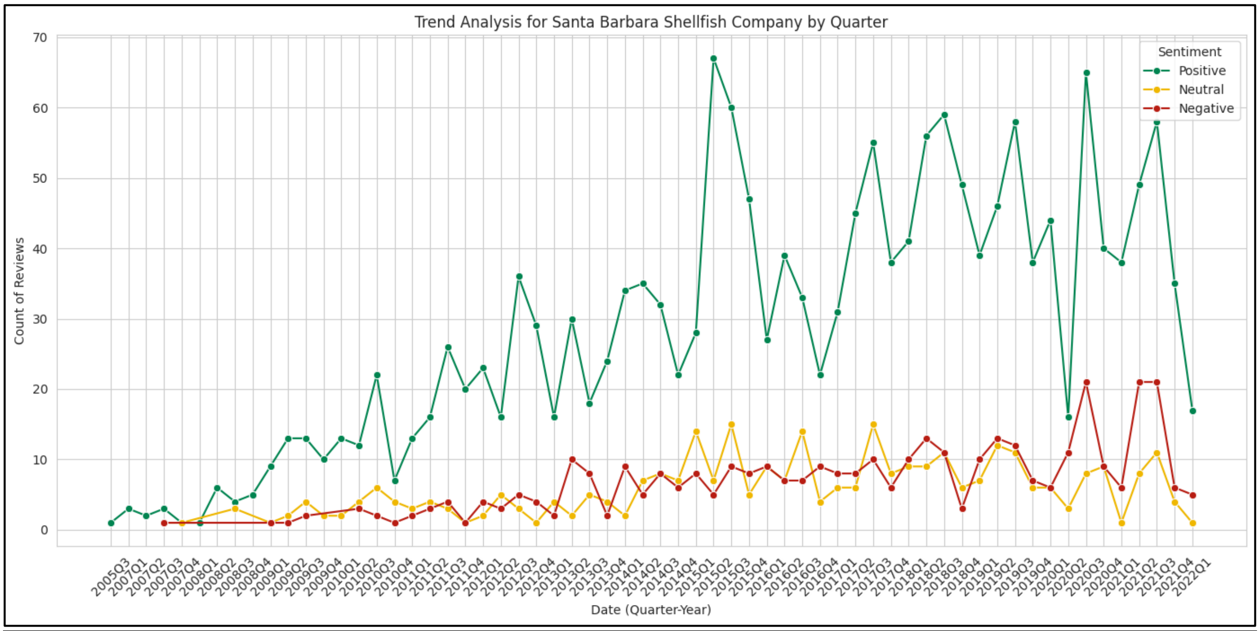

Trend Analysis for Parc by Quarter
Sentiment
Negative
Neutral
Positive
Count of Reviews
Date (Quarter-Year)

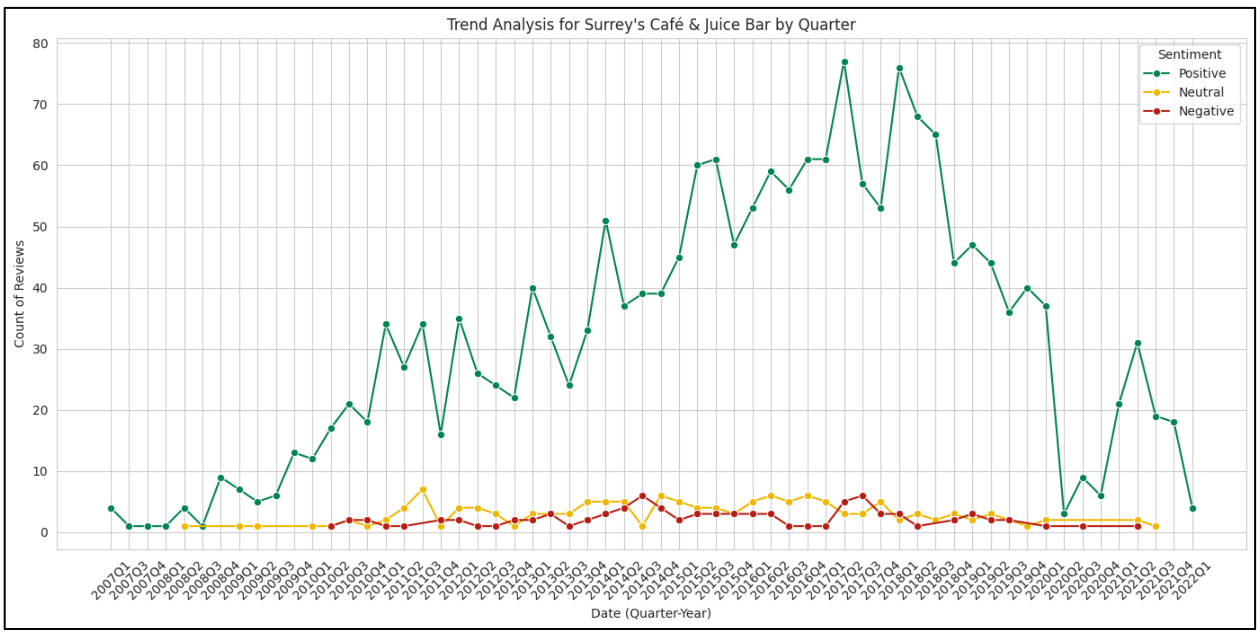


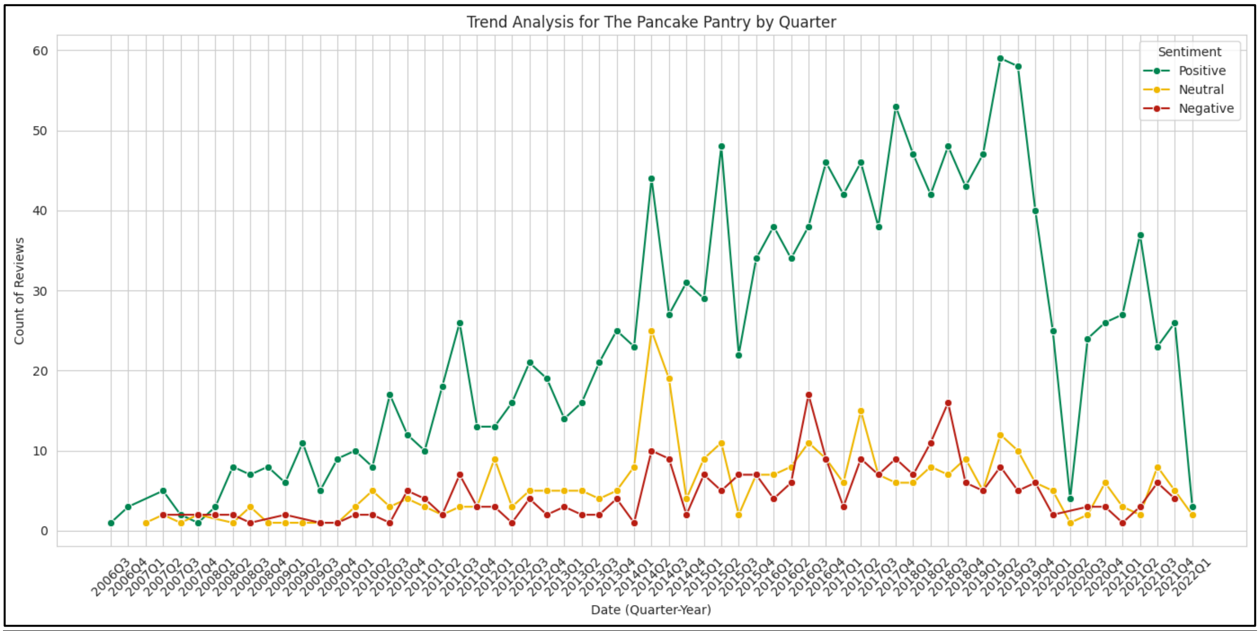
Trend Analysis for Peppermill Reno by Quarter
Sentiment
Positive
Neutral
Negative
Count of Reviews
Date (Quarter-Year)



Trend Analysis for Prep & Pastry by Quarter
Sentiment
Negative
Neutral
Positive
Count of Reviews
Date (Quarter-Year)


Trend Analysis for Santa Barbara Shellfish Company by Quarter
Count of Reviews
Date (Quarter-Year)
Sentiment
Positive
Neutral
Negative


Trend Analysis for Surrey's Café & Juice Bar by Quarter
Count of Reviews
Date (Quarter-Year)
Sentiment
Positive
Neutral
Negative

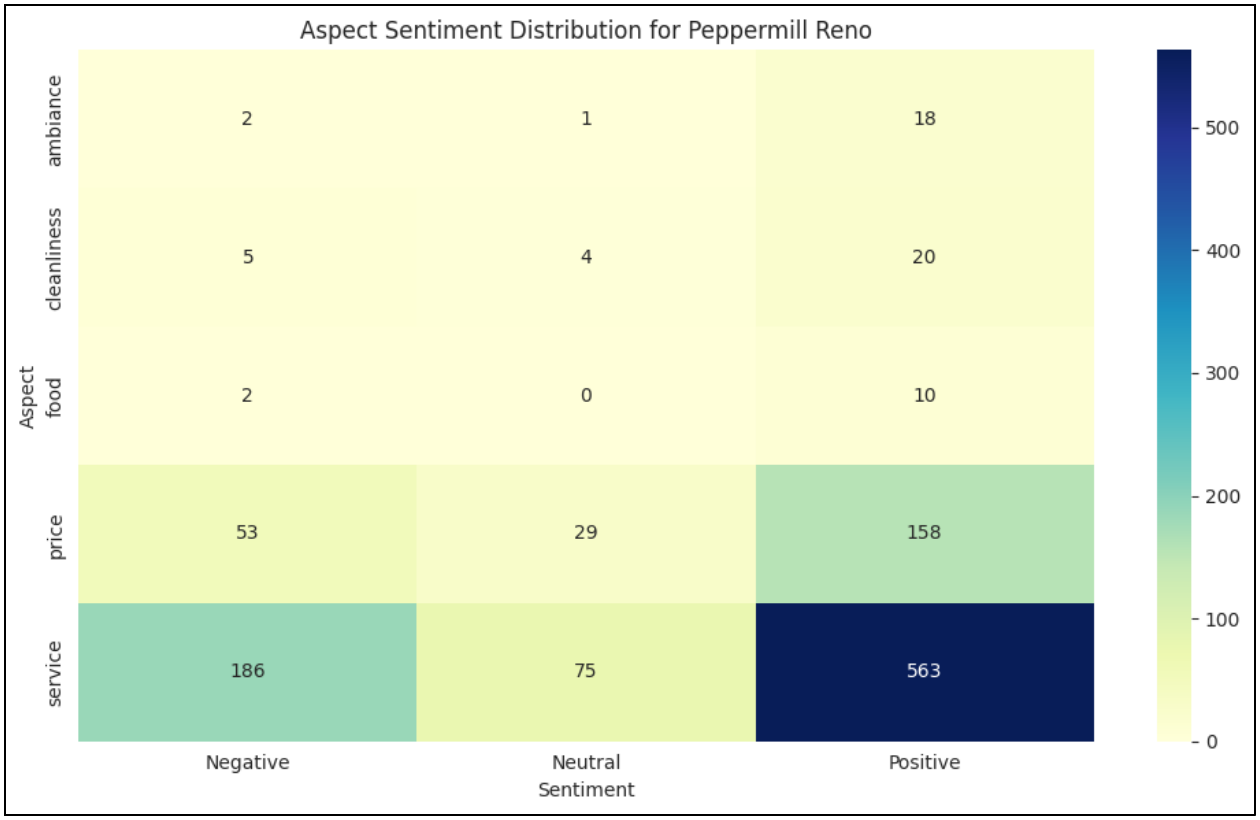


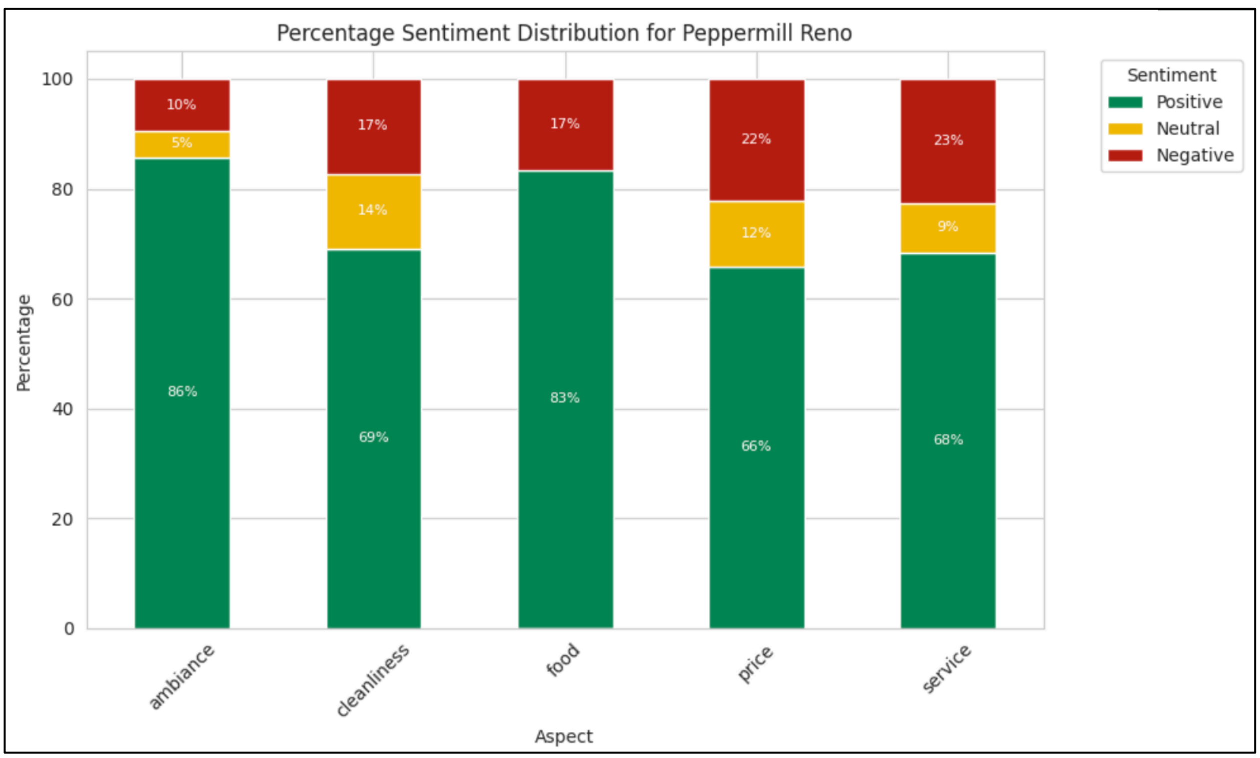
Aspect Based Sentiment Analysis

Aspect Sentiment Distribution

By examining specific aspects like service quality, food quality, ambiance, and price, this analysis can pinpoint what aspects contribute most to positive and negative sentiments. It offers a nuanced view of customer feedback, beyond simple satisfaction or dissatisfaction.
We try to determine which aspects of the service (e.g., food quality, service speed, ambiance, price) are most frequently associated with negative sentiments so that the business can prioritize these aspects for improvement.
For restaurants with predominantly negative feedback, implement targeted improvements in the identified areas. This may include staff training, menu adjustments, or ambiance enhancements.
Celebrate and promote the strengths identified through positive feedback to attract more customers.

Sample Output Below:


Aspect Sentiment Distribution for Peppermill Reno
Aspect
ambiance
cleanliness
food
price
service
2
1
18
5
4
20
2
0
10
53
29
158
186
75
563
Negative
Neutral
Positive
Sentiment
500
400
300
200
100
0


Percentage Sentiment Distribution for Peppermill Reno
Percentage
100
80
60
40
20
0
10%
5%
86%
17%
14%
69%
17%
83%
22%
12%
66%
23%
9%
68%
ambiance
cleanliness
food
price
service
Aspect
Sentiment
Positive
Neutral
Negative

Findings and Business Implications by Restaurant

1. Baileys

Findings	Business Implications
Positive reviews peak above 35 at times but can dip below 20 in other quarters.	Assess factors contributing to high peaks of satisfaction and address causes of significant dips to stabilize positive sentiment.
Negative sentiment spikes to nearly 15 reviews in some quarters while being as low as 5 in others.	Investigate operational or external factors that may correlate with spikes in negative sentiment to implement corrective actions.
Neutral reviews consistently range between 5 and 10 per quarter.	Engage with customers providing neutral feedback to understand the barriers to a higher rating and to enhance their experience.
Review counts show a downward trend, with the most recent quarters below the initial counts.	Analyze market trends, competition, and internal changes that might have contributed to the decline in review frequency.

2. Cafe fleur de Lis

Findings	Business Implications
Positive reviews have increased significantly, peaking at over 60 reviews in some quarters before a sharp decline.	Investigate the factors that led to the peak in positive reviews to replicate success and understand the reasons for the subsequent decline to avoid future downturns.
Negative reviews have peaks that coincide with declines in positive reviews, reaching up to approximately 20 reviews.	Examine events or changes in the business that may have caused dissatisfaction during the periods of increased negative reviews.
Neutral reviews remain low, generally under 10 reviews per quarter, indicating a polarized customer perception.	Develop strategies to engage customers who left neutral reviews to encourage more positive experiences and feedback.

3. District donut sliders brew

Findings	Business Implications

Positive reviews saw a peak of over 100 in Q2 of an unspecified year, followed by a steep decline in subsequent quarters.	Analyze the factors contributing to the high volume of positive feedback during the peak and address the causes behind the dramatic decline to regain customer satisfaction.
Negative reviews have minor fluctuations but a noticeable peak around Q3 of the following year, reaching close to 20 reviews.	Investigate potential operational issues, menu changes, or service disruptions that correlate with the increase in negative reviews during this period.
Neutral reviews maintain a fairly low profile with a slight uptick in the same quarter as the negative review peak.	Understand the experiences of customers leaving neutral feedback to identify missed opportunities for creating highly positive experiences.

4. Drago's Seafood Restaurant

Findings	Business Implications
A sharp increase in positive reviews reaching over 70 reviews in recent quarters.	Explore what changes or initiatives led to this surge to continue the positive momentum.
Negative reviews fluctuate but have notable peaks, with one reaching around 30 reviews.	Identify any recurring issues or events that align with these peaks to prevent future dissatisfaction.
Neutral reviews show an upward trend in the earlier quarters, peaking at around 20 reviews, followed by a decline.	Investigate the reasons behind the initial increase and subsequent decline in neutral reviews to better target the middle-ground customer experiences.

5. Gaylord Opryland Resort & Convention Center

Findings	Business Implications
The positive reviews show an overall increasing trend, peaking at around 35 reviews.	Capitalize on what is working well during peak times to sustain and further improve customer satisfaction.
Negative reviews exhibit a peak that coincides with a dip in positive reviews, reaching about 20 reviews.	Investigate service disruptions or event-related issues that might correlate with negative review spikes.
Neutral reviews are comparatively low, but they show an increase parallel to the negative reviews at certain points.	Engage with customers providing neutral feedback to understand the lack of enthusiasm and improve their experiences.

6. Luke

Findings	Business Implications
Positive reviews showed a strong increasing trend, peaking at around 120 reviews, before a significant downturn.	Understand the factors behind the peak periods of positive reviews and address the causes of the recent downturn to recover lost ground.
Negative reviews remain relatively constant with slight fluctuations, indicating a persistent undercurrent of dissatisfaction.	Investigate the reasons behind the consistent negative feedback to address underlying issues that may be affecting overall customer sentiment.
Neutral reviews stay low and stable, suggesting that customers are more likely to have a clear positive or negative impression.	Use the clarity of customer impressions to focus on enhancing aspects that are already well-received and improving or eliminating those that are not.

7. Milk and Honey Nashville

Findings	Business Implications
Positive reviews experience a significant peak, reaching over 130, followed by a steep decline.	Examine the reasons behind the peak to understand what drove the positive sentiment and address the factors that led to the decline.
Negative reviews have a consistent presence, with a slight upward trend, peaking around 30 reviews.	Investigate persistent issues that may be leading to negative reviews and develop strategies to resolve them.
Neutral reviews are stable but begin to show a slight increase at the end of the period, suggesting the potential for increased customer dissatisfaction.	Engage with customers leaving neutral reviews to encourage more positive experiences and to identify any areas of unmet expectations.

8. Mr. B's Bistro

Findings	Business Implications

Positive reviews show several peaks, with the highest around 70 reviews, indicating periods of high customer satisfaction.	Investigate what drove the high levels of satisfaction during peak periods and strive to replicate those conditions consistently.
Negative reviews exhibit a gradual increase over time, with peaks that roughly coincide with drops in positive reviews.	Examine possible operational or service issues that correspond with the increase in negative reviews to address underlying problems.
Neutral reviews show minor fluctuations but generally remain low, suggesting that experiences at Mr. B's Bistro tend to be polarizing.	Engage with customers leaving neutral reviews to understand their lack of strong sentiment and find ways to enhance their dining experience.

9. Muriel’s Jackson Square

Findings	Business Implications
Positive reviews show an upward trend, peaking at around 50 reviews, indicating periods when the restaurant excelled in customer satisfaction.	Analyze operational strategies, menu offerings, and service protocols during peak times to identify key factors contributing to success.
Negative reviews have a few significant spikes but tend to stay below 10 reviews, pointing to isolated incidents rather than widespread issues.	Investigate the root causes of the spikes in negative reviews to understand specific customer grievances and address them.
Neutral reviews have minimal presence and fluctuate slightly, implying that most customers have definitive opinions about their experience.	Further explore the reasons behind the neutral feedback to identify potential areas for improvement that could turn these into positive reviews.

10. Parc

Findings	Business Implications
Positive reviews peak at over 40 but show volatility, indicating inconsistency in customer satisfaction.	Investigate the reasons behind the fluctuations to maintain and improve consistency in the positive dining experience.

Negative reviews show a gradual increase, with a few notable peaks around 10 reviews, suggesting specific areas of concern.	Identify and address any operational, service, or menu issues corresponding with negative feedback peaks.
Neutral reviews are consistent but remain low, which might indicate that customer experiences tend to be decisive.	Look into converting neutral feedback into positive by enhancing certain aspects of the dining experience.

11. Peppermill Reno

Findings	Business Implications
Positive reviews show an overall upward trend with recent peaks nearing 60 reviews, indicating improved customer satisfaction.	Investigate and reinforce the factors contributing to increased satisfaction to maintain and continue the positive trend.
Negative reviews fluctuate over time, with peaks reaching around 20 reviews, signifying areas of potential concern.	Analyze specific instances or operational changes that correlate with peaks in negative reviews to understand and rectify the issues.
Neutral reviews are less common but show slight increases at times, suggesting occasional customer ambivalence.	Look into experiences leading to neutral reviews to identify and enhance aspects that could improve customer perception.

12. Prep & Pastry

Findings	Business Implications
Positive reviews show considerable volatility, with peaks reaching up to 80 reviews, indicative of periods of high customer satisfaction.	Determine the drivers behind periods of high satisfaction and implement strategies to maintain consistency in positive experiences.
Negative reviews remain relatively low but have small spikes, suggesting occasional customer dissatisfaction.	Delve into the specific instances that led to dissatisfaction to address any service or product quality issues.
Neutral reviews are consistently low, which may indicate that customers often have clear opinions about their experiences.	Understand the factors that prevent neutral reviewers from becoming positive

	promoters of the brand and address these areas.

13. Santa Barbara Shellfish Company

Findings	Business Implications
Positive reviews have high peaks, indicating moments when the restaurant particularly excelled, but there are notable dips that need attention.	Identify factors that contribute to peak experiences and understand the causes of the dips to maintain more consistent satisfaction levels.
Negative reviews demonstrate a gradual increase over time, with some peaks coinciding with dips in positive reviews.	Investigate any changes in service, product quality, or customer expectations that align with increased negative reviews.
Neutral reviews are the least frequent but show a slight upward trend, indicating potential for significant improvement.	Explore what is preventing neutral reviewers from being more enthusiastic to identify improvement areas.

14. District donut sliders brew

Findings	Business Implications
Positive reviews show an upward trend with some variability, peaking at around 70 reviews.	Investigate the driving factors behind high-satisfaction periods and strive for consistency to maintain high levels of positive reviews.
Negative reviews have occasional spikes but are generally lower than positive reviews, indicating isolated instances of dissatisfaction.	Pinpoint and address the specific causes of dissatisfaction that lead to negative review spikes.
Neutral reviews are present but relatively low, suggesting that most experiences elicit a stronger sentiment, either positive or negative.	Explore ways to elevate neutral experiences into positive ones and understand the barriers to achieving a higher satisfaction level.

15. The Pancake Pantry

Findings	Business Implications

Positive reviews show an upward trend with peaks of up to 50 reviews, reflecting moments of high customer satisfaction.	Understand what factors contribute to these peaks to maintain and further enhance customer satisfaction.
Negative reviews are relatively few but show some peaks, suggesting instances of dissatisfaction that need addressing.	Analyze what occurred during times of increased negative feedback to prevent future occurrences.
Neutral reviews are consistent but lower in number, indicating a potential opportunity to convert neutral experiences to positive ones.	Investigate the causes behind neutral reviews to transform these experiences into more positive encounters.

Conclusion and Recommendations

1. Baileys

Conclusion	Recommendations
Sentiment analysis reveals that positive sentiment can fluctuate significantly quarter over quarter.	Establish a customer feedback system to consistently monitor and react to customer sentiment, aiming to reduce the variability in positive review counts.
The highest negative feedback peak nearly triples the count of some of the lowest points.	Implement a robust customer service recovery plan that activates when negative feedback reaches a defined threshold, such as more than 10 negative reviews in a quarter.
Neutral sentiment shows less variability but persists, suggesting missed opportunities for delighting customers.	Initiate a program to follow up with customers who leave neutral feedback to uncover actionable insights that can turn a satisfactory experience into an exceptional one.
Declining engagement in recent periods is noticeable, which may impact the business's reputation and growth.	Develop and launch a targeted outreach campaign, possibly involving incentives for leaving reviews, to re-engage customers and encourage more frequent feedback.

2. Cafe fleur de Lis

Conclusion	Recommendations

The significant rise and fall in positive sentiment indicate periods of high customer satisfaction followed by a potential issue that needs addressing.	Implement consistent quality checks and service standards to maintain the high level of customer satisfaction reflected in peak positive reviews.
The correlation between the increase in negative reviews and the decrease in positive ones suggests an inverse relationship.	Conduct a thorough review of negative feedback to identify and address systemic issues that lead to customer dissatisfaction.
The low count of neutral reviews implies that customers have strong opinions about their experiences, leaning toward positive or negative.	Leverage the strong reactions captured in reviews to refine the customer experience and drive more consistently positive outcomes.
The sharp decline in review counts across all sentiments may impact the café's online presence and attractiveness to new customers.	Engage in active social media and online reputation management to encourage more reviews and to address any concerns promptly.

3. District donut sliders brew

Conclusion	Recommendations
The restaurant experiences a volatile customer satisfaction pattern, with sharp increases and decreases in positive feedback.	Develop strategies to maintain quality and service to prevent future downturns in positive feedback.
A correlation between the peaks of negative reviews and downturns of positive reviews suggests an inverse relationship.	Establish a rapid response system to mitigate the factors contributing to negative reviews, especially during identified peak periods.
Despite the overall lower counts, the pattern of neutral reviews potentially indicates consistency in the average customer experience.	Introduce incentives for customers to leave detailed feedback to better understand the middle-ground experiences and how to enhance them.
The sharp decline in reviews after the peak period may indicate reduced customer engagement or operational challenges.	Initiate a customer re-engagement campaign using social media, loyalty programs, and targeted marketing to encourage repeat visits and reviews.

4. District donut sliders brew

Conclusion	Recommendations

The restaurant has experienced a substantial increase in positive feedback, indicating successful periods or offerings that resonate well with customers.	Maintain and replicate the successful aspects of the restaurant's offerings, and ensure consistent delivery of the quality that led to the positive reviews.
The peaks in negative sentiment correspond with drops in positive reviews, suggesting that certain factors adversely affect customer experiences.	Conduct an in-depth analysis of operational, staffing, or menu changes during these periods to determine and rectify the causes of negative feedback.
The neutral reviews' trend may signify that some customers' expectations are not fully met, despite not being entirely dissatisfied.	Implement a customer feedback loop that specifically targets customers leaving neutral reviews to understand their needs and turn their experiences into positive ones.
The overall trend of reviews, particularly the decrease in negative and neutral reviews, suggests improving customer sentiment.	Leverage the recent increase in positive sentiment in marketing strategies to attract new customers and incentivize repeat business.

5. Gaylord Opryland Resort & Convention Center

Conclusion	Recommendations
There's an observable growth in customer satisfaction, but attention is needed on the coinciding peaks of negative reviews.	Monitor customer feedback closely, especially during peak negative periods, and address the concerns raised to prevent a recurrence.
Negative and neutral feedback may indicate specific issues that are not yet fully understood or addressed.	Implement a comprehensive feedback system that can trace the source of dissatisfaction to specific services or operational areas for improvement.
The consistency in low neutral reviews suggests customers often have a decisive opinion about their experience.	Use detailed surveys or feedback mechanisms to delve into why experiences are not consistently positive and to uncover specific areas for enhancement.
Peaks in reviews could be associated with specific events or seasons, impacting the sentiment.	Perform a seasonal analysis of operational performance and customer traffic to align services and staffing with expected demand.

6. Luke

Conclusion	Recommendations
Luke's increasing trend in positive feedback suggests successful operations, which have recently seen a decline that needs to be investigated.	Conduct a thorough review of changes in operations, staff, menu, or service that coincided with the decline in positive reviews to inform recovery strategies.
The stability of negative feedback, even with the rise in positive reviews, indicates potential areas of consistent customer dissatisfaction.	Introduce a detailed customer feedback system that targets areas highlighted in negative reviews to implement specific improvements.
The low level of neutral feedback implies that experiences at Luke are memorable, prompting customers to provide strong feedback.	Leverage the strong responses as a marketing tool to showcase what Luke does best, while internally addressing the negative aspects to prevent them from overshadowing the positives.
The recent decline in reviews after a peak suggests an event or change that has had a significant impact on customer perception.	Analyze recent events, changes in customer service, or environmental factors that could have led to the downturn, and engage with customers to reassure and regain their trust.

7. Milk and Honey Nashville

Conclusion	Recommendations
The peak in positive reviews indicates a period of high customer satisfaction, which is not sustained.	Capitalize on what was working well during the peak period and identify changes that may have caused the downturn to reverse the trend.
The increase in negative reviews, even as positive reviews decline, suggests underlying issues that need addressing.	Implement targeted improvements in areas frequently mentioned in negative reviews, such as customer service or food quality.
The rise in neutral reviews towards the end indicates that some	Offer follow-up incentives for customers who leave reviews to gather more detailed feedback and to show a commitment to improvement.

customers' needs may not be fully met.	

8. Mr. B's Bistro

Conclusion	Recommendations
The restaurant has experienced significant fluctuations in positive sentiment, with notable peaks that suggest periods of excellence.	Establish best practices based on the peak periods to maintain high standards that lead to positive customer experiences.
An upward trend in negative sentiment alongside positive review peaks suggests specific areas of the business may be overlooked.	Introduce a customer service initiative focused on continuous improvement, particularly addressing areas commonly cited in negative reviews.
The stability in neutral reviews indicates a consistent customer base whose expectations are met but not exceeded.	Develop a customer engagement program to convert neutral experiences into positive ones, potentially through personalized service or loyalty rewards.

9. Muriel’s Jackson Square

Conclusion	Recommendations
The overall trend indicates that Muriel’s Jackson Square often meets or exceeds customer expectations, reflected by the number of positive reviews.	Maintain and reinforce the qualities and practices that lead to high customer satisfaction to ensure consistent positive experiences.
The spikes in negative feedback, while not frequent, suggest that certain aspects occasionally fail to meet customer expectations.	Establish a mechanism for immediate response and resolution to negative reviews to mitigate their impact and prevent future occurrences.
The relative rarity of neutral feedback suggests a strong customer reaction, either positive or negative, to the dining experience.	Engage customers post-visit with surveys to gain more nuanced insights into their experiences, aiming to convert any neutrality into positivity.

10. Parc

Conclusion	Recommendations
The pattern of positive reviews suggests that when Parc performs well, it significantly exceeds customer expectations.	Capitalize on successful practices from peak periods to create more consistent positive experiences.
The presence and growth of negative feedback highlight the importance of addressing any potential recurring issues.	Establish a quality assurance process to mitigate factors that lead to negative reviews, ensuring they are addressed proactively.
The low number of neutral reviews may offer an opportunity to create a memorable experience for every guest.	Introduce initiatives to engage diners more fully, such as personalized service or follow-up on their dining experience.

11. Peppermill Reno

Conclusion	Recommendations
The increasing positive feedback trend suggests that Peppermill Reno is on a trajectory of growth in terms of customer satisfaction.	Maintain the service or product quality that has led to the upsurge in positive reviews, ensuring consistent customer experiences.
The peaks in negative sentiment indicate specific periods or events where customer experience may be lacking.	Set up a feedback loop to quickly identify and address any customer issues, particularly during peak periods of negative feedback.
The presence of neutral feedback, while minimal, presents an opportunity to convert satisfied customers into enthusiastic advocates.	Implement a follow-up system to engage customers after their visit, aiming to elevate their experiences from neutral to positive.

12. Prep & Pastry

Conclusion	Recommendations

The fluctuating pattern in positive reviews suggests that when Prep & Pastry excels, it can significantly surpass customer expectations.	Maintain the standards and offerings that correlate with peak positive feedback and aim to replicate these conditions to minimize fluctuations.
The presence of negative feedback, while not extensive, points to specific areas of potential improvement.	Implement a detailed review of negative feedback to identify and rectify operational challenges or customer service issues.
The low number of neutral reviews presents an opportunity to focus on converting every customer experience into a positive one.	Develop initiatives to actively engage with all customers post-visit, particularly those with neutral feedback, to elevate their experience.

13. Santa Barbara Shellfish Company

Conclusion	Recommendations
The restaurant experiences significant fluctuations in customer satisfaction, suggesting that when the experience is positive, it's very well received.	Leverage the successful elements from peak periods to create a more consistent experience that minimizes negative feedback.
The presence and pattern of negative reviews suggest specific intervals or reasons for customer dissatisfaction that need to be addressed.	Implement a real-time feedback system to quickly capture and address negative experiences before they escalate.
The low frequency of neutral reviews points to opportunities to impress every customer and move the needle towards a positive response.	Introduce a program or service feature aimed at turning satisfactory experiences into exceptional ones, such as personalized touches or loyalty incentives.

14. Surrey's Café & Juice Bar

Conclusion	Recommendations
The presence of negative reviews at certain points suggests that there are areas for potential service improvement.	Implement a system to closely monitor and respond to negative feedback promptly, aiming to prevent recurring issues.

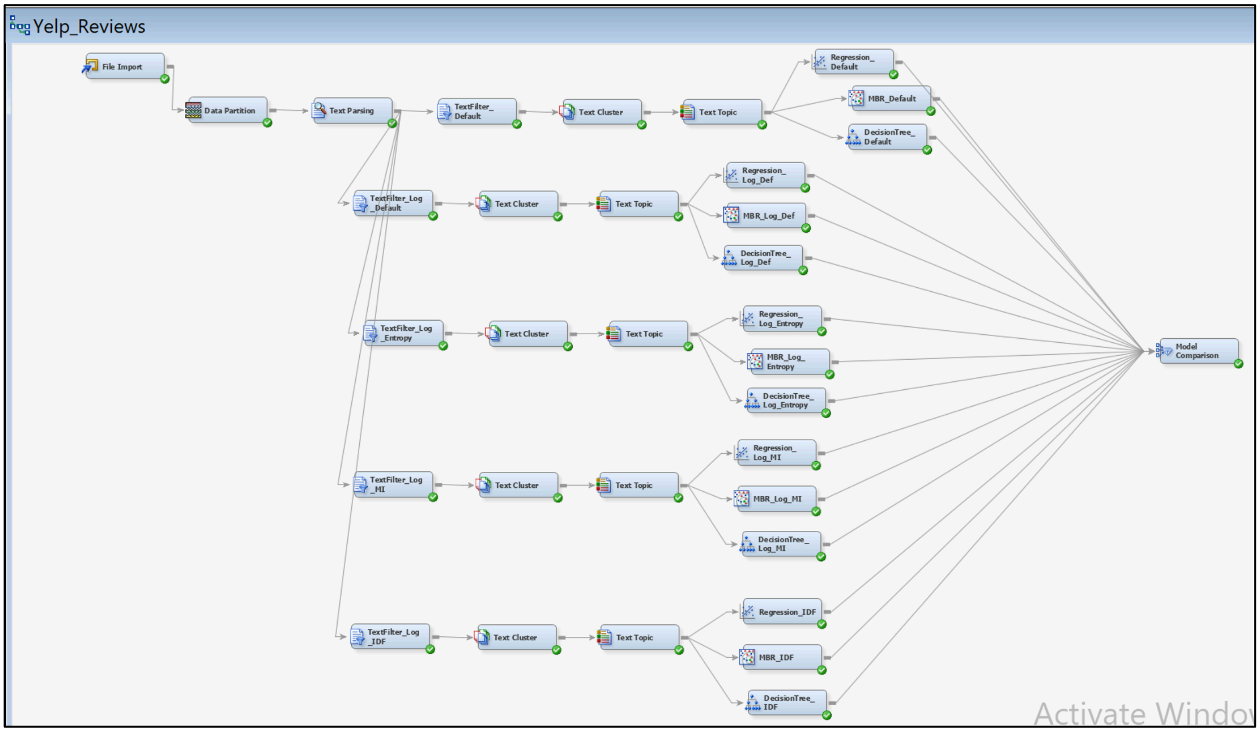
The low frequency of neutral feedback indicates an opportunity to turn good experiences into great ones.	Introduce an initiative to proactively seek customer feedback to gain actionable insights, particularly from those who had a neutral experience.

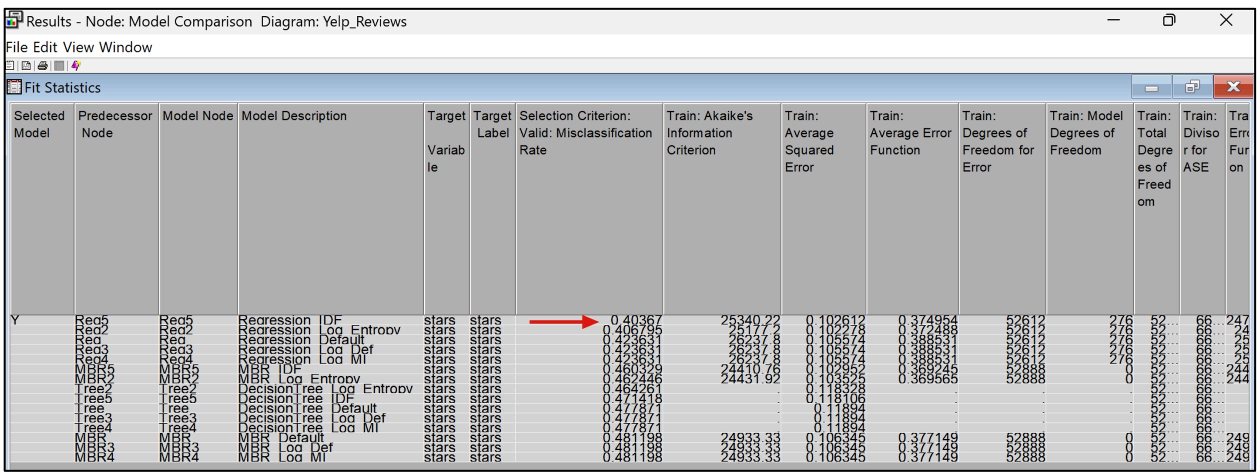
15. The Pancake Pantry

Conclusion	Recommendations
The Pancake Pantry generally achieves positive feedback from customers, with room to improve consistency in delivering high-quality experiences.	Focus on consistently delivering the products and services that correlate with positive feedback to increase overall satisfaction.
The occurrence of negative reviews at certain intervals suggests that there are specific areas that could be improved.	Create a system for collecting and analyzing feedback in real time to quickly address any negative experiences.
The relatively low but steady number of neutral reviews presents an opportunity for enhancement.	Engage with customers leaving neutral feedback to gain insights into how to elevate their experience to a positive one.

Supervised Learning:

Following is the SAS diagram for our Yelp dataset where we partitioned the data in 40-30-30 split, followed by text parsing and filtering using different options including default, log, Entropy, Mutual Information and so on. This parsed data was then fed to a text cluster followed by text topics to generate a cluster of words that directed towards a particular sentiment. These clusters were then fed to different models like Logistic Regression, MBR and Decision Trees using different configurations of log, entropy, mutual information and inverse document frequency.
We observed that our best performing model turned out to be Logistic Regression with IDF with a misclassification rate of around 40%.





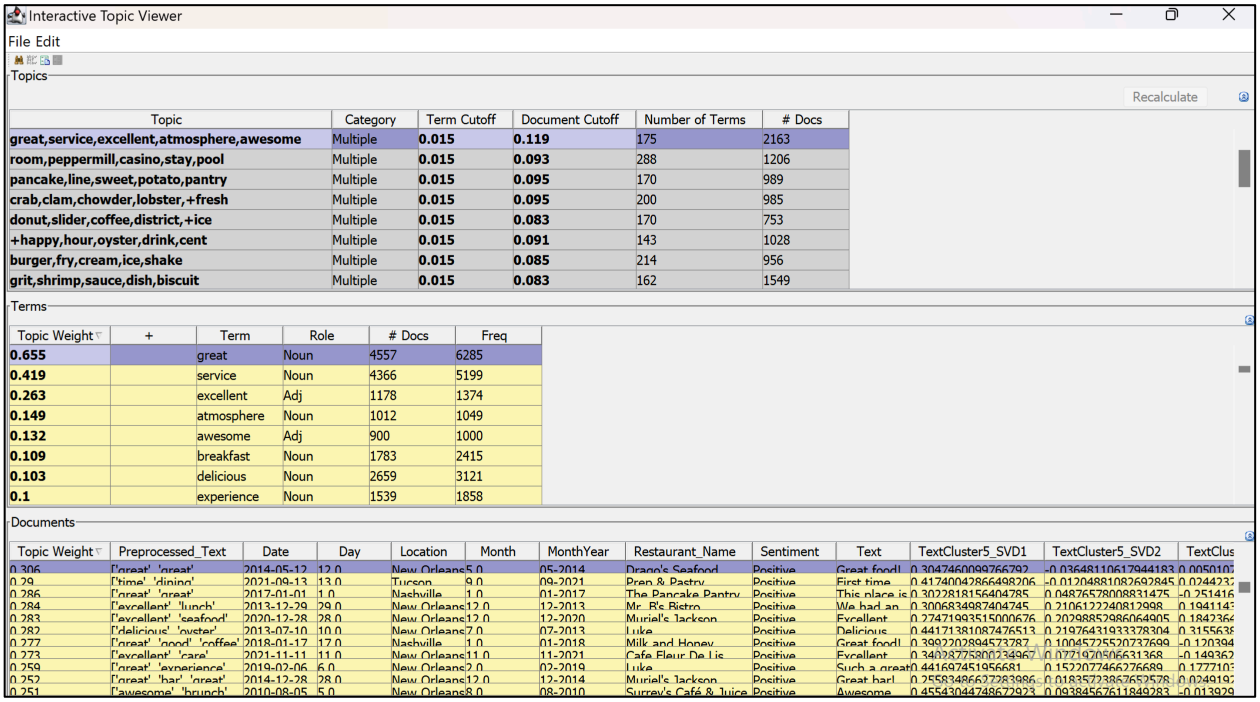
Results - Node: Model Comparison Diagram: Yelp_Reviews

File Edit View Window

Fit Statistics

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion: Valid: Misclassification Rate	Train: Akaike's Information Criterion	Train: Average Squared Error	Train: Average Error Function	Train: Degrees of Freedom for Error	Train: Model Degrees of Freedom	Train: Total Degrees of Freedom	Train: Divisor for ASE	Train: Error Function
Y	Reg5	Reg5	Regression_IDF	stars	stars	0.40367	25340.22	0.102612	0.374954	52612	276	52...	66...	247
	Reg2	Reg2	Regression_Log_Entropy	stars	stars	0.406795	25177.2	0.102278	0.372488	52612	276	52...	66...	24
	Reg	Reg	Regression_Default	stars	stars	0.423631	26237.8	0.105574	0.388531	52612	276	52...	66...	25
	Reg3	Reg3	Regression_Log_Def	stars	stars	0.423631	26237.8	0.105574	0.388531	52612	276	52...	66...	25
	Reg4	Reg4	Regression_Log_MI	stars	stars	0.423631	26237.8	0.105574	0.388531	52612	276	52...	66...	25
	MBR5	MBR5	MBR_IDF	stars	stars	0.460329	24410.76	0.102952	0.369245	52888	0	52...	66...	244
	MBR2	MBR2	MBR_Log_Entropy	stars	stars	0.462446	24431.92	0.103525	0.369565	52888	0	52...	66...	244
	Tree2	Tree2	DecisionTree_Log_Entropy	stars	stars	0.464261	.	0.118328	.	.	.	52...	66...	
	Tree5	Tree5	DecisionTree_IDF	stars	stars	0.471418	.	0.118106	.	.	.	52...	66...	
	Tree	Tree	DecisionTree_Default	stars	stars	0.477871	.	0.11894	.	.	.	52...	66...	
	Tree3	Tree3	DecisionTree_Log_Def	stars	stars	0.477871	.	0.11894	.	.	.	52...	66...	
	Tree4	Tree4	DecisionTree_Log_MI	stars	stars	0.477871	.	0.11894	.	.	.	52...	66...	
	MBR	MBR	MBR_Default	stars	stars	0.481198	24933.33	0.106345	0.377149	52888	0	52...	66...	249
	MBR3	MBR3	MBR_Log_Def	stars	stars	0.481198	24933.33	0.106345	0.377149	52888	0	52...	66...	249
	MBR4	MBR4	MBR_Log_MI	stars	stars	0.481198	24933.33	0.106345	0.377149	52888	0	52...	66...	249

Topic Viewer results:



Interactive Topic Viewer

File Edit

Topics

Recalculate

Topic	Category	Term Cutoff	Document Cutoff	Number of Terms	# Docs
great,service,excellent,atmosphere,awesome	Multiple	**0.015**	**0.119**	175	2163
room,peppermill,casino,stay,pool	Multiple	**0.015**	**0.093**	288	1206
pancake,line,sweet,potato,pantry	Multiple	**0.015**	**0.095**	170	989
crab,clam,chowder,lobster,+fresh	Multiple	**0.015**	**0.095**	200	985
donut,slider,coffee,district,+ice	Multiple	**0.015**	**0.083**	170	753
+happy,hour,oyster,drink,cent	Multiple	**0.015**	**0.091**	143	1028
burger,fry,cream,ice,shake	Multiple	**0.015**	**0.085**	214	956
grit,shrimp,sauce,dish,biscuit	Multiple	**0.015**	**0.083**	162	1549

Terms

Topic Weight	+	Term	Role	# Docs	Freq
0.655		great	Noun	4557	6285
0.419		service	Noun	4366	5199
0.263		excellent	Adj	1178	1374
0.149		atmosphere	Noun	1012	1049
0.132		awesome	Adj	900	1000
0.109		breakfast	Noun	1783	2415
0.103		delicious	Noun	2659	3121
0.1		experience	Noun	1539	1858

Documents

Topic Weight	Preprocessed_Text	Date	Day	Location	Month	MonthYear	Restaurant_Name	Sentiment	Text	TextCluster5_SVD1	TextCluster5_SVD2	TextClus
0.306	['great', 'great'	2014-05-12	12.0	New Orleans	5.0	05-2014	Drago's Seafood	Positive	Great food!	0.3047460099766792	-0.0364811061794418З	0.005010
0.29	['time', 'dining'	2021-09-13	13.0	Tucson	9.0	09-2021	Prep & Pastry	Positive	First time	0.41740042866498206	-0.01204881082692845	0.024423
0.286	['great', 'great'	2017-01-01	1.0	Nashville	1.0	01-2017	The Pancake Pantry	Positive	This place is	0.3022818156404785	0.04876578008831475	-0.251416
0.284	['excellent', 'lunch'	2013-12-29	29.0	New Orleans	12.0	12-2013	Mr. B's Bistro	Positive	We had an	0.3006834987404745	0.2106122240812998	0.194114
0.283	['excellent', 'seafood'	2020-12-28	28.0	New Orleans	12.0	12-2020	Muriel's Jackson	Positive	Excellent	0.27471993515000676	0.20298852986064905	0.184236
0.282	['delicious', 'oyster'	2013-07-10	10.0	New Orleans	7.0	07-2013	Luke	Positive	Delicious	0.44171381087476513	0.21976431933378304	0.315563
0.277	['great', 'good', 'coffee'	2018-01-17	17.0	Nashville	1.0	01-2018	Milk and Honey	Positive	Great food!	0.3992202894573787	0.10045725520737699	-0.120394
0.273	['excellent', 'care'	2021-11-11	11.0	New Orleans	11.0	11-2021	Cafe Fleur De Lis	Positive	Excellent	0.34028775800234967	0.07719705066313б8	-0.149362
0.259	['great', 'experience'	2019-02-06	6.0	New Orleans	2.0	02-2019	Luke	Positive	Such a great	0.441697451956681	0.1522077466276689	0.177710
0.252	['great', 'bar', 'great'	2014-12-28	28.0	New Orleans	12.0	12-2014	Muriel's Jackson	Positive	Great bar!	0.2558348662728З986	0.01835723867652578	0.024919
0.251	['awesome', 'brunch'	2010-08-05	5.0	New Orleans	8.0	08-2010	Surrey's Café & Juice	Positive	Awesome	0.45543044748672923	0.09384567611849283	-0.013929

REFERENCES

- Fabio Nelli. (2023). Python Data Analytics: With Pandas, NumPy, and Matplotlib. Retrieved from Semantic Scholar.
- Nezar Abdennur, G. Fudenberg, Ilya M. Flyamer, Aleksandra Galitsyna, A. Goloborodko, Maxim Imakaev, Sergey V. Venev. (2022). Bioframe: operations on genomic intervals in Pandas dataframes. Bioinformatics, 40. Retrieved from Semantic Scholar.
- John Hunter, Darren Dale. (2020). Matplotlib. In Learning Scientific Programming with Python. Retrieved from Semantic Scholar.
- TextBlob: Emmy Danny Ajik, Aminu Bashir Suleiman, & Muhammad Ibrahim. (2023). ENHANCING USER EXPERIENCE THROUGH SENTIMENT ANALYSIS FOR KATSINA STATE TRANSPORT AGENCY: A TEXTBLOB APPROACH. FUDMA Journal of Sciences. Link
- NLTK: Nihaarika Agarwal & Vijayalaxmi. (2023). Covid-19 Impact on Mental Health: Sentiment Analysis Using NLTK. 2023 6th International Conference on Contemporary Computing and Informatics (IC3I). Link
- TFIDFVectorizer: Vipin Kumar & Basant Subba. (2020). A TfidfVectorizer and SVM based sentiment analysis framework for text data corpus. 2020 National Conference on Communications (NCC). Link