

ZOMATO

EXPLORATORY DATA ANALYSIS



PROJECT BY :
NAINIKA THAPA

ABOUT ZOMATO

Zomato, an Indian multinational restaurant aggregator and food delivery company, has significantly transformed the food industry since its inception in 2008 by Deepinder Goyal and Pankaj Chaddah. Originally launched as "Foodiebay," it was rebranded to Zomato in 2010 and has since expanded its operations to over 24 countries and more than 10,000 cities. Zomato's platform provides users with comprehensive information on restaurants, including menus, reviews, and ratings, making it easier for people to discover new dining experiences. In addition to restaurant discovery, Zomato offers online food delivery and table reservations, enhancing the convenience and accessibility of dining. The company's acquisition of Blinkit, a fast grocery delivery service, has further broadened its scope, allowing users to quickly access groceries in addition to restaurant meals. Through initiatives like Zomato Gold/Zomato Pro, the company also offers exclusive deals and discounts, adding value to its services. Overall, Zomato continues to innovate and adapt, cementing its position as a leading player in the global food and dining industry.

PURPOSE OF THIS PROJECT

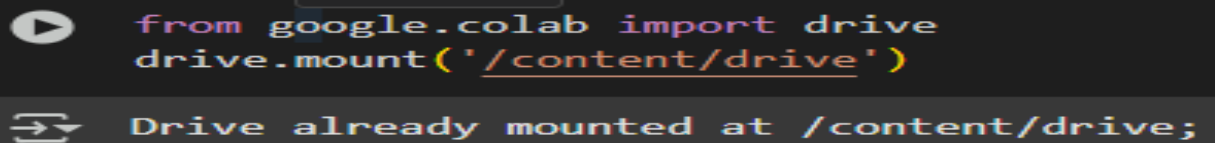
- The primary purpose of this project appears to be **exploratory data analysis(EDA)** and visualization of data related to Indian restaurants .
- Analyzing the distribution of restaurant ratings, average ratings, and understanding the overall rating landscape.
- Identifying popular restaurant locations, exploring the distribution of restaurant ratings across different cities, and potentially finding areas with high restaurant density.
- Determining the most popular cuisines, investigating the relationship between cuisine variety and restaurant ratings, and gaining insights into cuisine preferences .
- Examining the relationship between price range and restaurant ratings, visualizing average costs for different price categories, and understanding the price-quality dynamics in the market.
- Analyzing the distribution of restaurants based on features like Wi-Fi and alcohol availability, investigating if the presence of certain features correlates with higher ratings, and gaining insights into restaurant characteristics .

DATASET LOADING

```
# IMPORTING REQUIRED LIBRARIES
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from wordcloud import WordCloud , STOPWORDS
import warnings
warnings.filterwarnings('ignore')
from collections import Counter
import re
import plotly.express as px
from textblob import TextBlob
```

•Importing required libraries for performing data analysis and creating visualizations in Python. **PANDAS** is used for data manipulation ,**NUMPY** is used for numerical operations, **SEABORN AND MATPLOTLIB** for creating static plots , **Wordclouds** are primarily used for visualizing the frequency of words in a text. The more frequent a word, the larger and bolder it appears in the word cloud. **Warnings** module is used to control and manage warning messages that may arise during code execution. **Counter** is a subclass of dict that is used for counting hashable objects. It provides a convenient way to count the frequency of items in a list or other iterable. **Re** module provides regular expression operations for pattern matching and text manipulation. Regular expressions are powerful tools for searching, extracting, and manipulating text based on patterns. **Plotly.express** is a high-level interface for creating interactive visualizations using the Plotly library. It provides a simple and concise way to generate a variety of charts, including bar charts, line charts, scatter plots, and more.

TextBlob library provides tools for natural language processing (NLP) tasks, such as sentiment analysis, part-of-speech tagging, and noun phrase extraction. It is built on top of the NLTK library and provides a user-friendly interface for common NLP operations.

A screenshot of a Google Colab code cell. The top part shows the code: `from google.colab import drive` and `drive.mount('/content/drive')`. The bottom part shows the output: a play button icon followed by the text "Drive already mounted at /content/drive;".

```
from google.colab import drive
drive.mount('/content/drive')

Drive already mounted at /content/drive;
```

Then connecting the Google Drive to the google collab so you can access your files. It then sets the path to CSV file with loan information and reads this file into a dataframe.

DATASET LOADING

df.head()

	res_id	name	establishment	url	address	city	city_id	locality	latitude	longitude	...
0	3400299	Bikanervala	['Quick Bites']	https://www.zomato.com/agra/bikanervala-khanda...	Kalyani Point, Near Tulsi Cinema, Bypass Road,...	Agra	34	Khandari	27.211450	78.002381	...
1	3400005	Mama Chicken Mama Franky House	['Quick Bites']	https://www.zomato.com/agra/mama-chicken-mama-...	Main Market, Sadar Bazaar, Agra Cantt, Agra	Agra	34	Agra Cantt	27.160569	78.011583	...
2	3401013	Bhagat Halwai	['Quick Bites']	https://www.zomato.com/agra/bhagat-halwai-2-sh...	62/1, Near Easy Day, West Shivaji Nagar, Goalp...	Agra	34	Shahganj	27.182938	77.979684	...
3	3400290	Bhagat Halwai	['Quick Bites']	https://www.zomato.com/agra/bhagat-halwai-civi...	Near Anjana Cinema, Nehru Nagar, Civil Lines, ...	Agra	34	Civil Lines	27.205668	78.004799	...
4	3401744	The Salt Cafe Kitchen & Bar	['Casual Dining']	https://www.zomato.com/agra/the-salt-cafe-kitc...	1C,3rd Floor, Fatehabad Road, Tajganj, Agra	Agra	34	Tajganj	27.157709	78.052421	...

I've named the dataset as 'df'. Then i use df.head() method . This is a method that displays the first few rows of the DataFrame. By default, it shows the first 5 rows. You can specify a different number of rows to display by passing an integer as an argument to the head() method .

df.shape

(211944, 26)

There are 211944 rows and 26 columns present in the dataset .

DESCRIPTION


```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 211944 entries, 0 to 211943
Data columns (total 26 columns):
#   Column                Non-Null Count  Dtype  
---  -
0   res_id                211944 non-null int64  
1   name                  211944 non-null object 
2   establishment          211944 non-null object 
3   url                   211944 non-null object 
4   address               211810 non-null object 
5   city                  211944 non-null object 
6   city_id               211944 non-null int64  
7   locality              211944 non-null object 
8   latitude              211944 non-null float64 
9   longitude              211944 non-null float64 
10  zipcode                48757 non-null  object 
11  country_id            211944 non-null int64  
12  locality_verbose      211944 non-null object 
13  cuisines               210553 non-null object 
14  timings                208070 non-null object 
15  average_cost_for_two  211944 non-null int64  
16  price_range            211944 non-null int64  
17  currency               211944 non-null object 
18  highlights             211944 non-null object 
19  aggregate_rating      211944 non-null float64 
20  rating_text            211944 non-null object 
21  votes                  211944 non-null int64  
22  photo_count            211944 non-null int64  
23  opentable_support      211896 non-null float64 
24  delivery               211944 non-null int64  
25  takeaway               211944 non-null int64  
dtypes: float64(4), int64(9), object(13)
memory usage: 42.0+ MB
```

- **res_id**: Unique identifier for each restaurant.
- **name**: Name of the restaurant.
- **establishment_types**: Type of establishment (e.g., restaurant, cafe, bar).
- **url**: URL link to the restaurant's page.
- **address**: Address of the restaurant.
- **city**: City where the restaurant is located.
- **city_id**: Unique identifier for the city.
- **locality**: Specific locality within the city.
- **country_id**: Unique identifier for the country.
- **latitude**: Latitude coordinates of the restaurant's location.
- **longitude**: Longitude coordinates of the restaurant's location.
- **cuisines**: List of cuisines offered by the restaurant.

- timings:** Operating hours of the restaurant.
- average_cost_for_two:** Average cost for a meal for two people.
- price_range:** Price range category (1-4, with 4 being the most expensive).
- currency:** Currency used for pricing.
- highlights:** List of features or highlights of the restaurant.
- aggregate_rating:** Overall rating of the restaurant.
- rating_text:** Textual representation of the rating (e.g., Excellent, Very Good).
- votes:** Number of votes or reviews received by the restaurant.
- photo_count:** Number of photos associated with the restaurant.
- opentable_support:** Indicates if the restaurant accepts reservations through OpenTable (1 for yes, 0 for no).
- delivery:** Indicates if the restaurant offers delivery (1 for yes, 0 for no).
- take_away:** Indicates if the restaurant offers takeaway (1 for yes, 0 for no).

Dropping the duplicate values

```
 df.drop_duplicates(inplace=True)
```

`df.drop_duplicates()` this function identifies and removes duplicate rows from the dataframe .

`inplace=True` this argument modifies the DataFrame directly, avoiding the need to assign it to a new variable.

DESCRIPTION

```
df.describe()
```

	res_id	city_id	latitude	longitude	country_id	average_cost_for_two
count	2.119440e+05	211944.000000	211944.000000	211944.000000	211944.0	211944.000000
mean	1.349411e+07	4746.785434	21.499758	77.615276	1.0	595.812229
std	7.883722e+06	5568.766386	22.781331	7.500104	0.0	606.239363
min	5.000000e+01	1.000000	0.000000	0.000000	1.0	0.000000
25%	3.301027e+06	11.000000	15.496071	74.877961	1.0	250.000000
50%	1.869573e+07	34.000000	22.514494	77.425971	1.0	400.000000
75%	1.881297e+07	11306.000000	26.841667	80.219323	1.0	700.000000
max	1.915979e+07	11354.000000	10000.000000	91.832769	1.0	30000.000000

There are 13 numerical columns i.e. res_id , city_id ,latitude,longitude,country_id , average_cost_for_two, price_range ,aggregate_rating , votes ,photo_count, opentable_support , delivery , takeaway and rest of the columns are categorical .

CHECKING NULL VALUES

```
[22] df.isnull().sum().sum()
```

```
1577
```

There are 1577 null values present in the dataset .

DATA CLEANING

Now , checking the null values

```
[22] df.isnull().sum().sum()
```

```
⇒ 1577
```

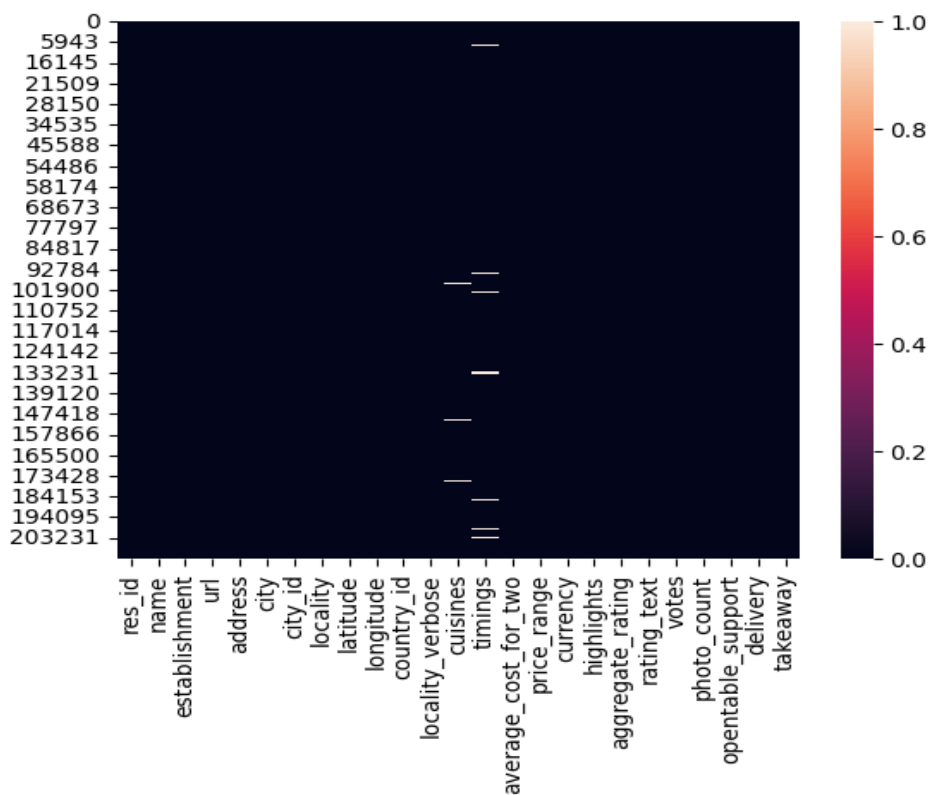
There are 1577 null values present in the dataset .

```
# Filling missing values
df['cuisines'].fillna('Unknown',inplace=True)
df['timings'].fillna('Not available',inplace=True)
df['opentable_support'].fillna(0,inplace=True)
df['address'].fillna('Not available',inplace=True)
```

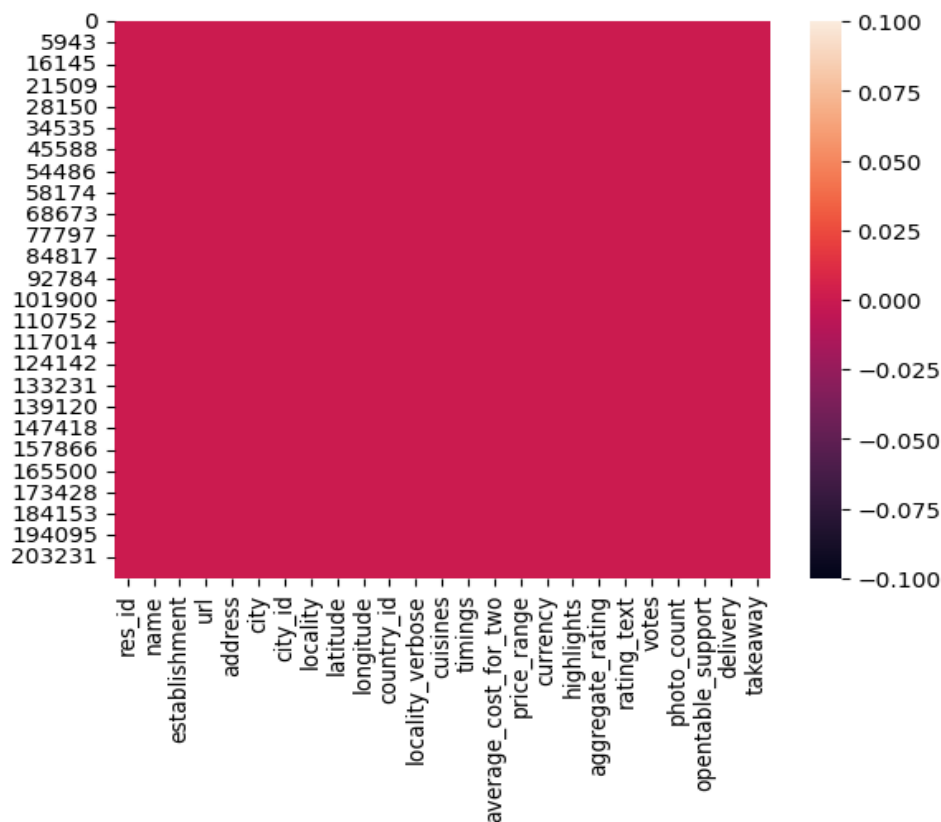
NOW , filling the null values present in the columns 'cuisines' , 'timings' , 'opentable_support' , 'address' using fillna function is crucial for data cleaning and preprocessing. It allows you to handle missing data in a way that is appropriate for your analysis or model.

DATA CLEANING

Heatmap before cleaning



Heatmap after cleaning



DATA CLEANING

Handling Outliers : Here i used IQR method for outliers detection and box plot for visualization of outliers .

Outliers in 'average_cost_for_two' : Outliers present in the column and it was imputed by median

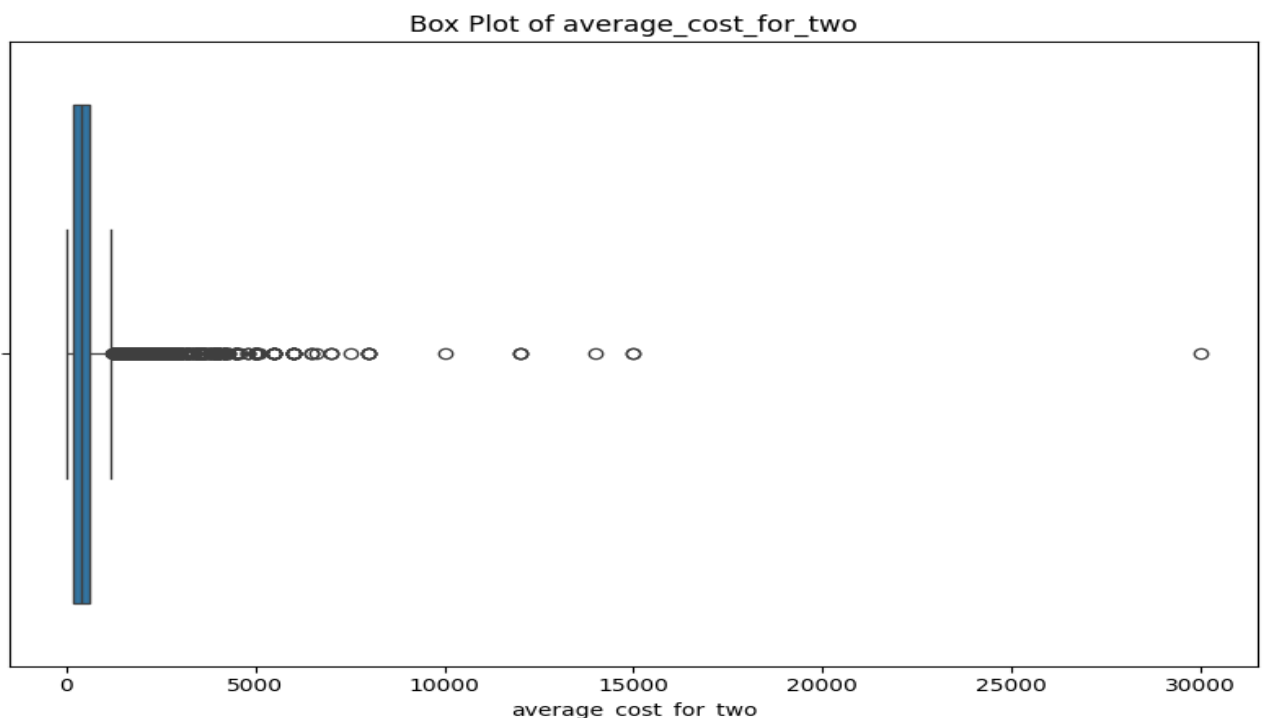
```
Q1 = df['average_cost_for_two'].quantile(0.25)
Q3 = df['average_cost_for_two'].quantile(0.75)
IQR = Q3 - Q1
lower_bound = Q1 - 1.5 * IQR
upper_bound = Q3 + 1.5 * IQR
df_1 = df[((df['average_cost_for_two'] < lower_bound) | (df['average_cost_for_two'] > upper_bound))]

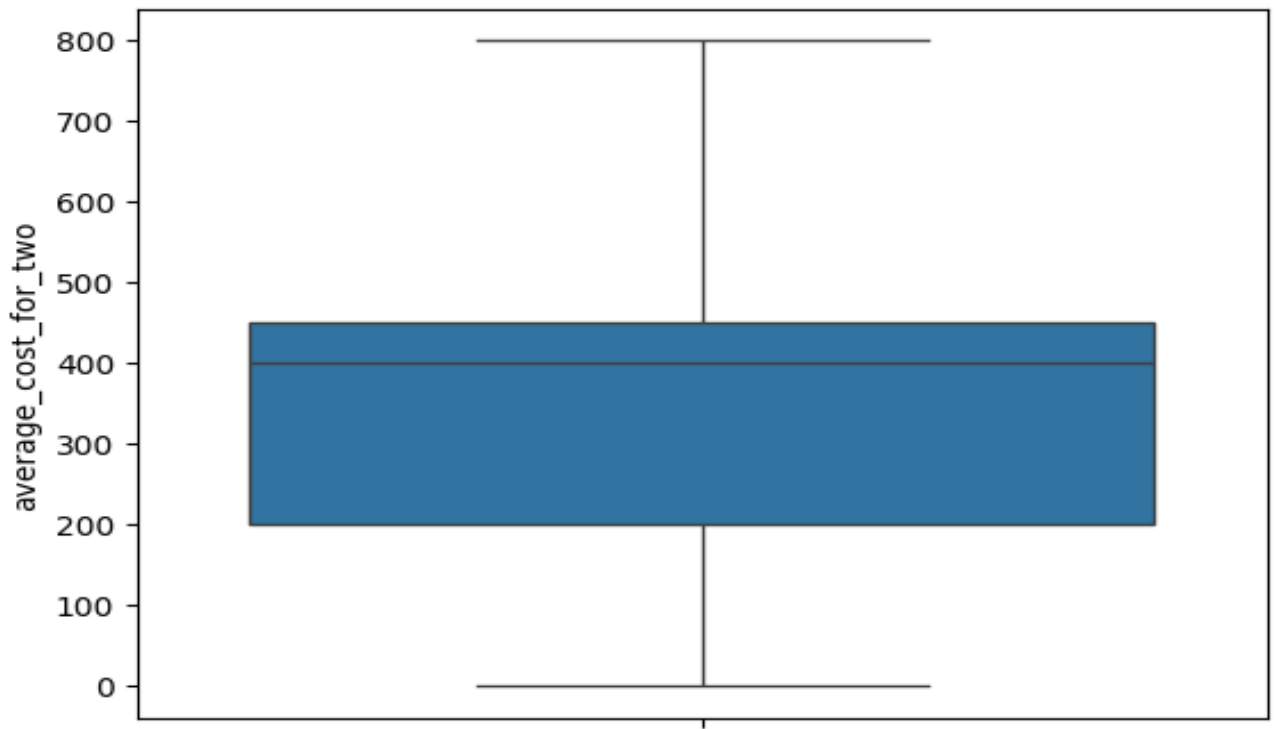
# Filling the outliers with median

median_avg = df['average_cost_for_two'].median()
df['average_cost_for_two'] = np.where(df['average_cost_for_two'] > upper_bound, median_avg, df['average_cost_for_two'])
df['average_cost_for_two'] = np.where(df['average_cost_for_two'] < lower_bound, median_avg, df['average_cost_for_two'])

sns.boxplot(df['average_cost_for_two'])
plt.show()
```

The outliers present in the 'average_cost_for_two' attribute are imputed by median , the box plot shows the zero outliers present in average_cost_for_two term .
Boxplot of average_cost_for_two with outliers .





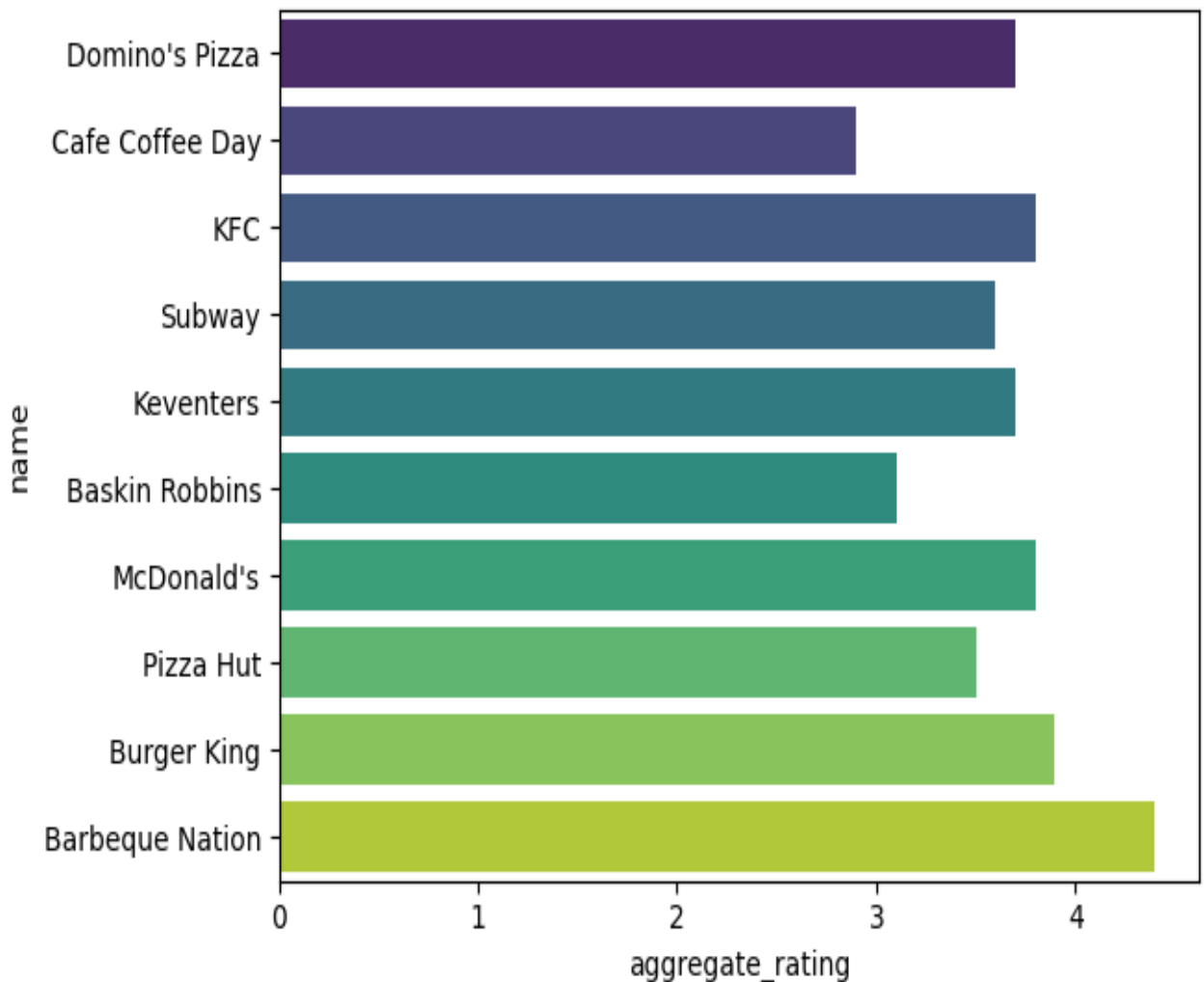
Boxplot of average_cost_for_two after outlier handling.

Outliers were present in the all numerical columns, other than average_cost_for_two were also addressed using the IQR method and capping to boundary values.

DATA VISUALIZATION

Basic Statistics :

Calculating and visualizing the average rating of restaurants .

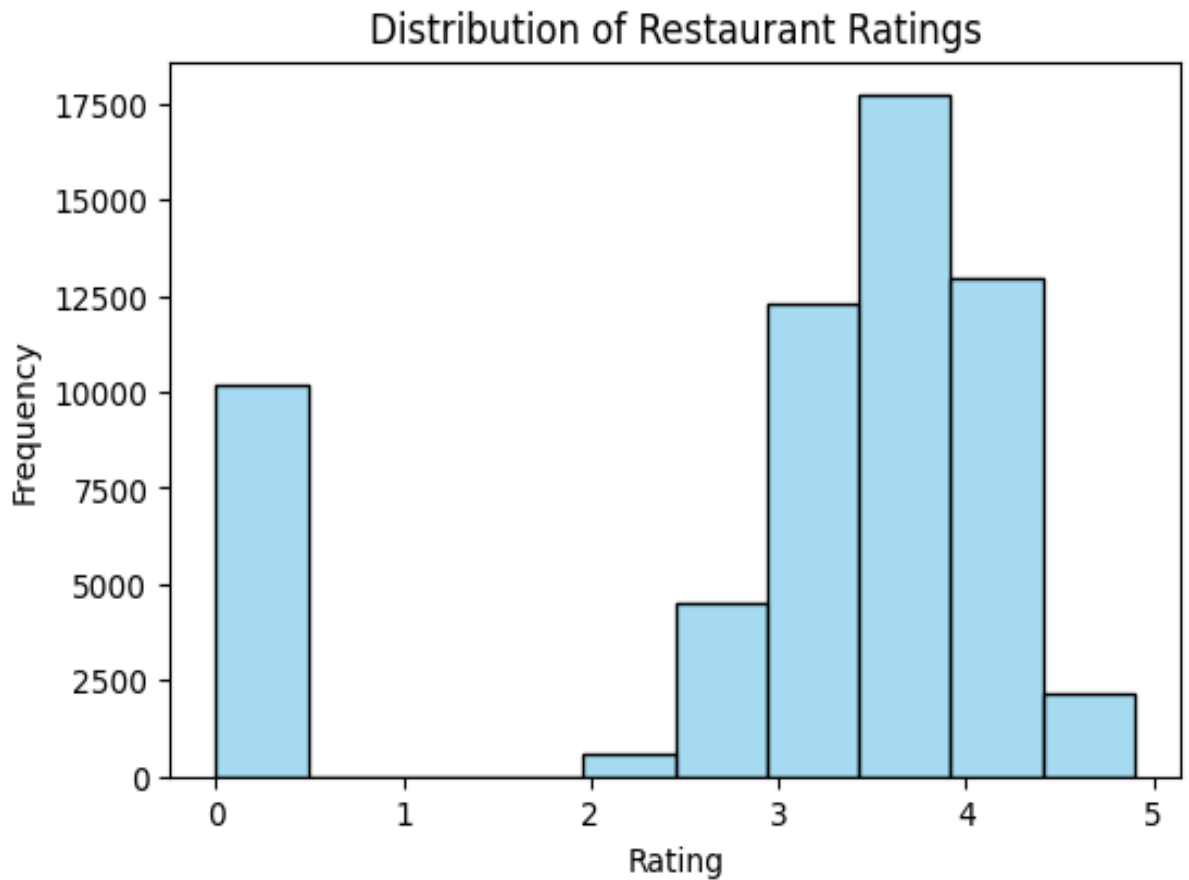


KEY INSIGHTS

- Barbeque Nation stands out with the highest aggregate rating , indicating it is the most popular or well-regarded among the listed food chains .
- Burger King and Pizza Hut follow closely behind , suggesting they also enjoy high customer satisfaction .
- Domino's Pizza has the lowest aggregate rating , which might indicate room for improvement in customer satisfaction or popularity .
- The ratings for McDonald's , Baskin Robbins , Keventers, Subway , and KFC are relatively close , showing a competitive landscape among these food chains .

DATA VISUALIZATION

Analyze the distribution of restaurants ratings to understand the overall rating landscape



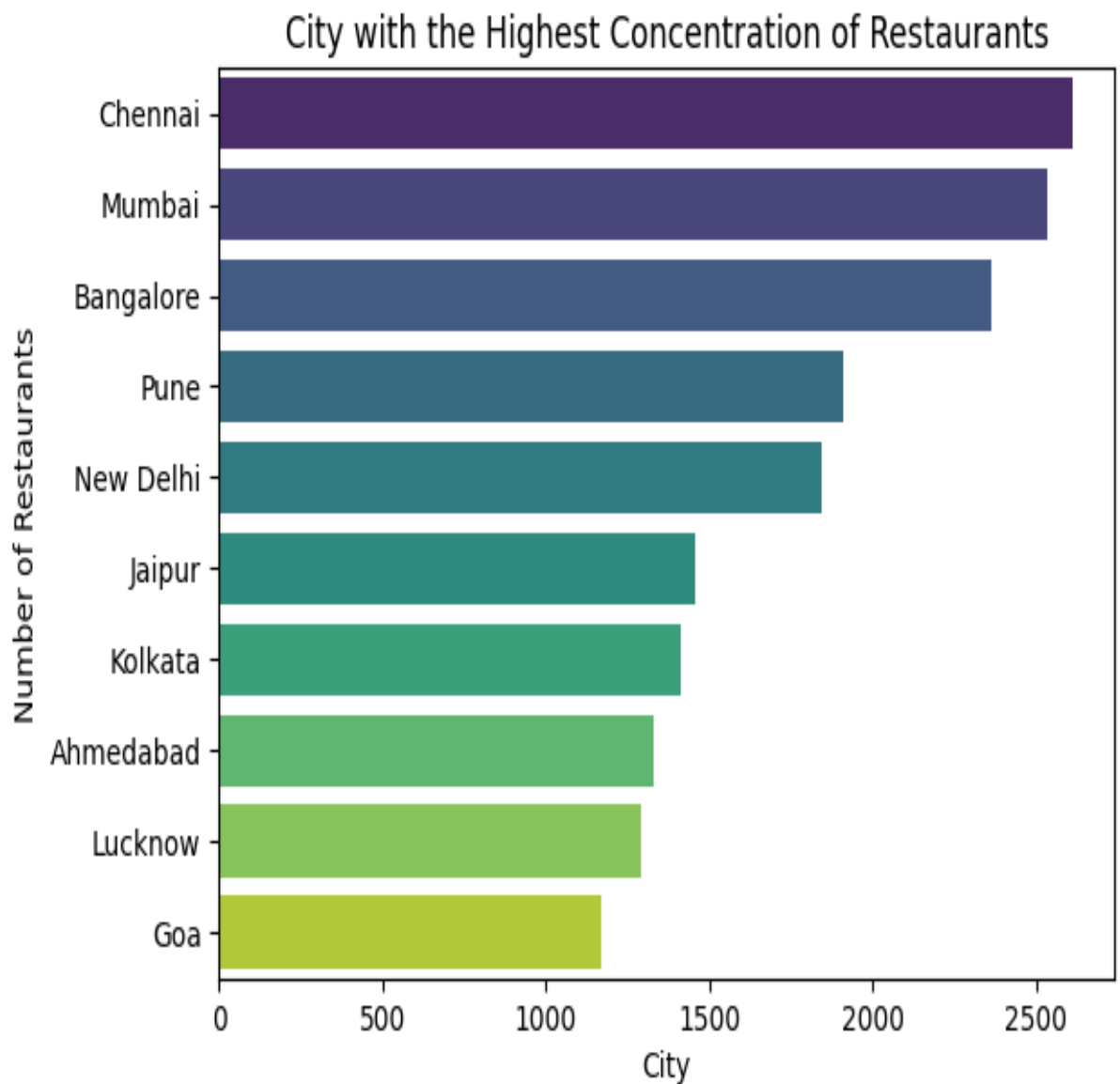
KEY INSIGHTS

- **Rating Distribution:** The most common rating is 4, with approximately 13,000 ratings. This is followed by a rating of 0, with around 10,500 ratings. Ratings of 3 and 5 have approximately 5,000 and 2,500 ratings, respectively. The least common rating is 2, with around 500 ratings.
- **Skewed Distribution:** The distribution is skewed towards higher ratings, with the majority of ratings being 4. This suggests that most customers are generally satisfied with their dining experiences.
- **Zero Ratings:** The significant number of 0 ratings might indicate a large number of customers who were extremely dissatisfied or possibly a data entry issue.
- **Low Ratings:** Ratings of 2 and 3 are relatively low, indicating that fewer customers had a mediocre experience.
- **High Ratings:** Ratings of 5 are also relatively low compared to 4, suggesting that while many customers are satisfied, fewer are extremely satisfied.

DATA VISUALIZATION

Location Analysis

Identify the city with the highest concentration of restaurants.

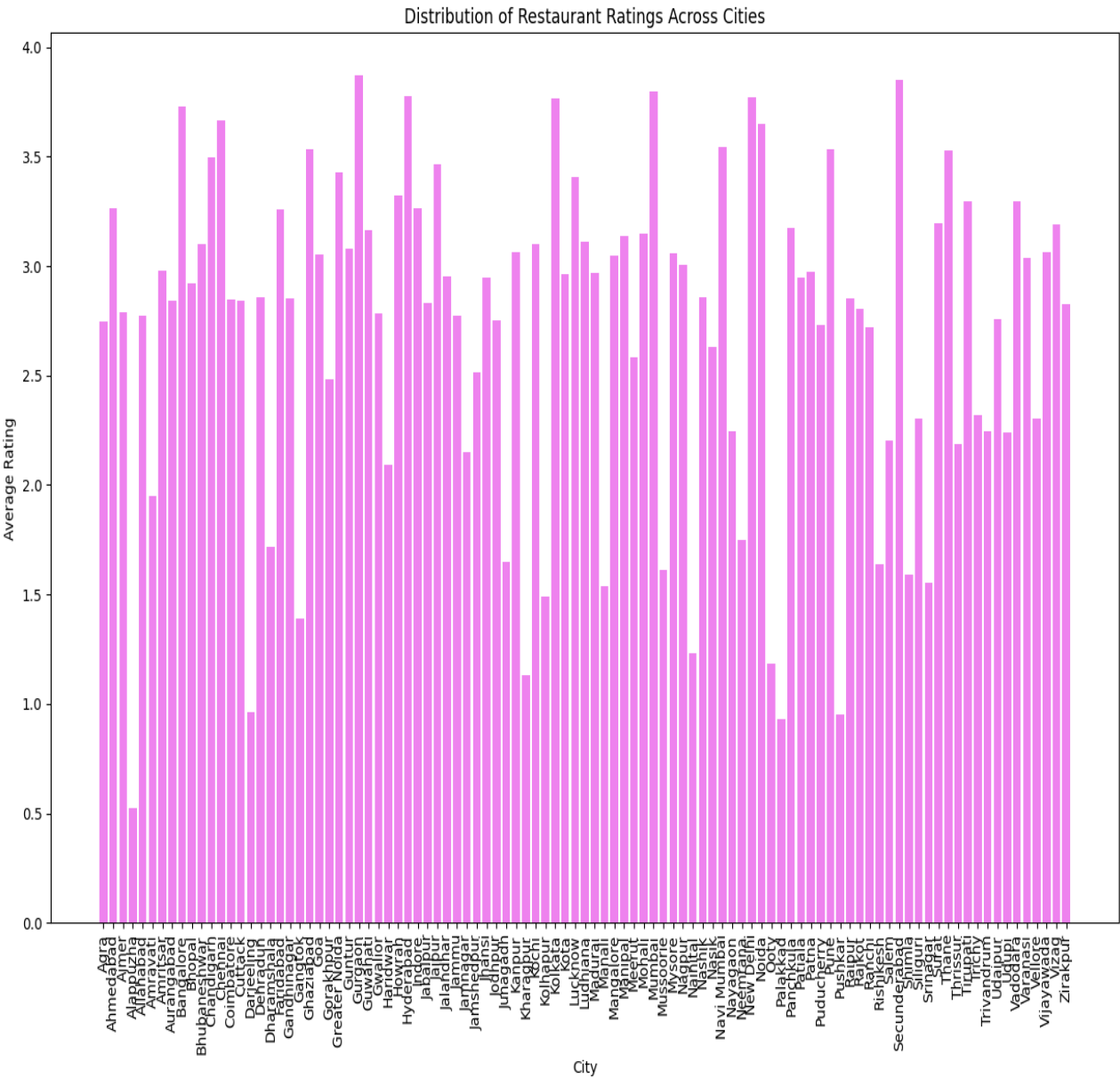


KEY INSIGHTS

- Chennai has the highest concentration of restaurants among the listed cities, followed by Mumbai and Bangalore.
- Pune, New Delhi, and Jaipur also have a significant number of restaurants, indicating a vibrant food scene.
- Kolkata, Ahmedabad, Lucknow, and Goa have relatively fewer restaurants compared to the top three cities but still offer a variety of dining options.

DATA VISUALIZATION

Visualize the distribution of restaurant ratings across different cities.



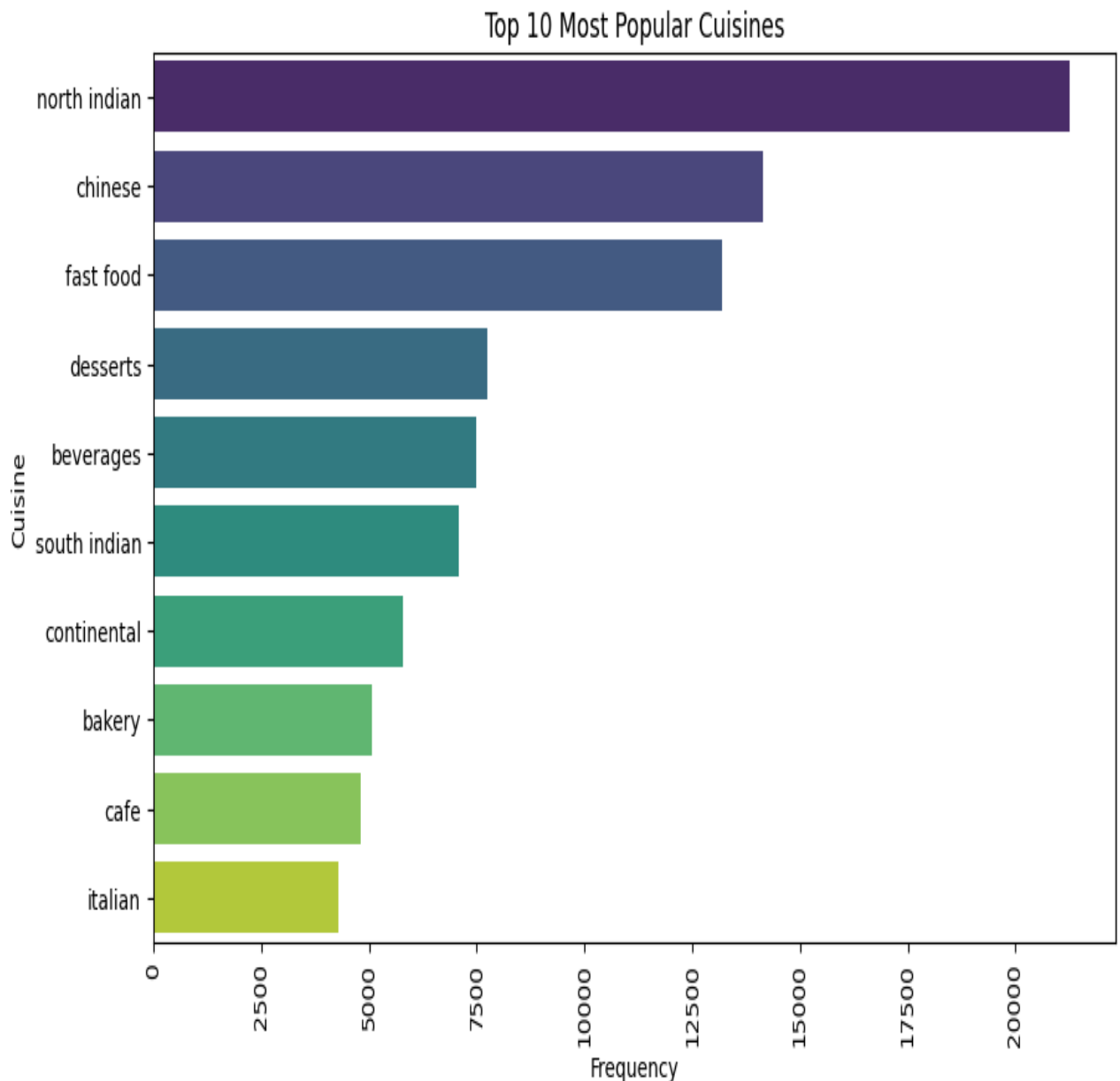
KEY INSIGHTS

- **Rating Variation:** The average restaurant ratings across cities generally range between 2.5 and 4.0. Cities like Amritsar, Chandigarh, and Jaipur have higher average ratings close to 4.0, while cities like Bhubaneswar, Guwahati, and Patna have lower average ratings closer to 2.5.
- **City Performance:** Cities with the highest average restaurant ratings include Amritsar, Chandigarh, and Jaipur. Cities with the lowest average restaurant ratings include Bhubaneswar, Guwahati, and Patna.
- **City Comparison:** Comparing cities, it is evident that there is a significant variation in restaurant ratings. Cities like Amritsar and Chandigarh outperform cities like Bhubaneswar and Guwahati in terms of average restaurant ratings. The performance of cities can be compared by looking at the height of the bars, with taller bars indicating higher average ratings and shorter bars indicating lower average ratings.

DATA VISUALIZATION

Cuisine Analysis:

Determine the most popular cuisines among the listed restaurants.



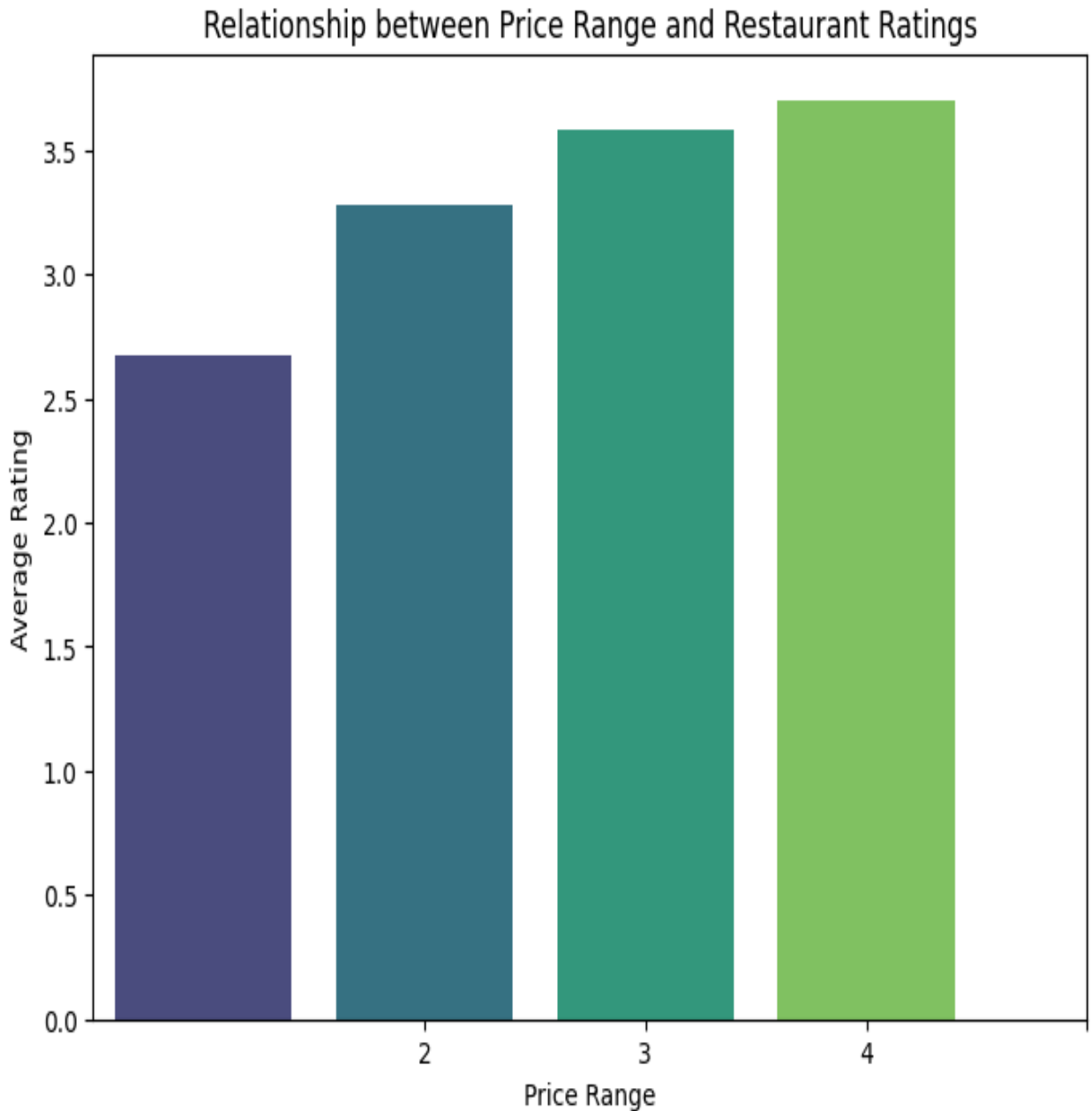
KEY INSIGHTS

- North Indian cuisine is the most popular, followed by Chinese and Fast Food.
- Desserts and Beverages also have a significant presence, indicating a strong preference for sweet treats and drinks.
- South Indian and Continental cuisines are also popular, showcasing a diverse range of food preferences.
- Bakery and Cafe options are well-liked, suggesting a trend towards casual dining and quick bites.
- Italian cuisine rounds out the top 10, highlighting its global appeal.

DATA VISUALIZATION

Price Range and Rating

Analyze the relationship between price range and restaurant ratings.

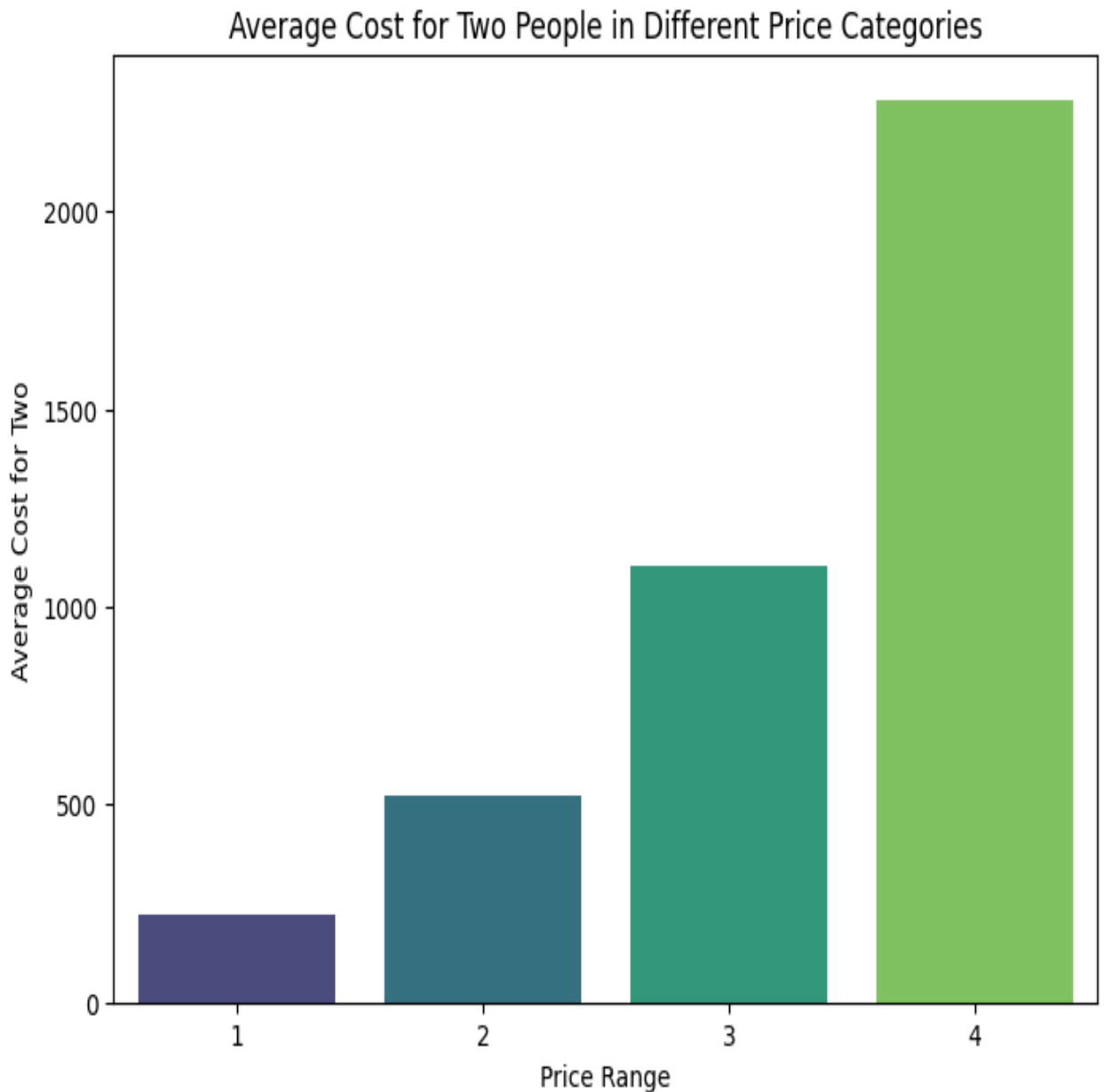


KEY INSIGHTS

- **Positive Correlation:** There is a positive correlation between the price range and the average restaurant rating. As the price range increases, the average rating also increases.
- **Higher Ratings for Higher Prices:** Restaurants in the highest price range (4) have the highest average rating, while those in the lowest price range (1) have the lowest average rating.
- **Customer Expectations:** This suggests that higher-priced restaurants tend to receive better ratings from customers, possibly due to higher expectations being met.

DATA VISUALIZATION

Visualize the average cost for two people in different price categories.



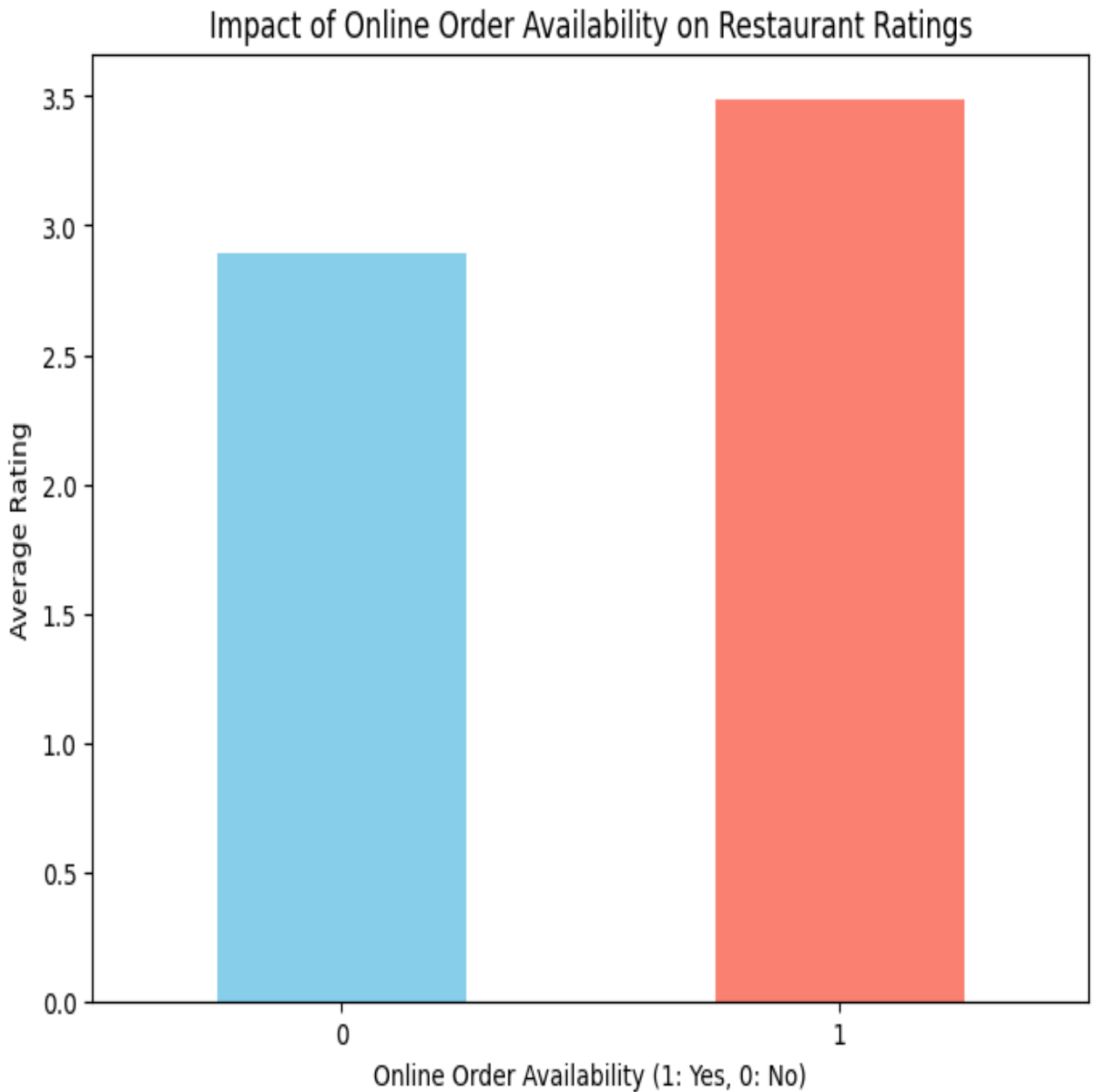
KEY INSIGHTS

- Price Range 1: The average cost for two people is approximately ₹200.
- Price Range 2: The average cost for two people is approximately ₹500.
- Price Range 3: The average cost for two people is approximately ₹450.
- Price Range 4: The average cost for two people is approximately ₹400.

The data indicates a direct correlation between the price range and the average cost for two people. As the price range increases, the average cost also increases proportionally. However, Price Range 2 has the highest average cost for two people, around ₹500, while Price Range 1 has the lowest average cost for two people, around ₹200. The average cost for two people in Price Range 3 and Price Range 4 is relatively close, around ₹450 and ₹400, respectively.

DATA VISUALIZATION

Investigate the impact of online order availability on restaurant ratings.



KEY INSIGHTS

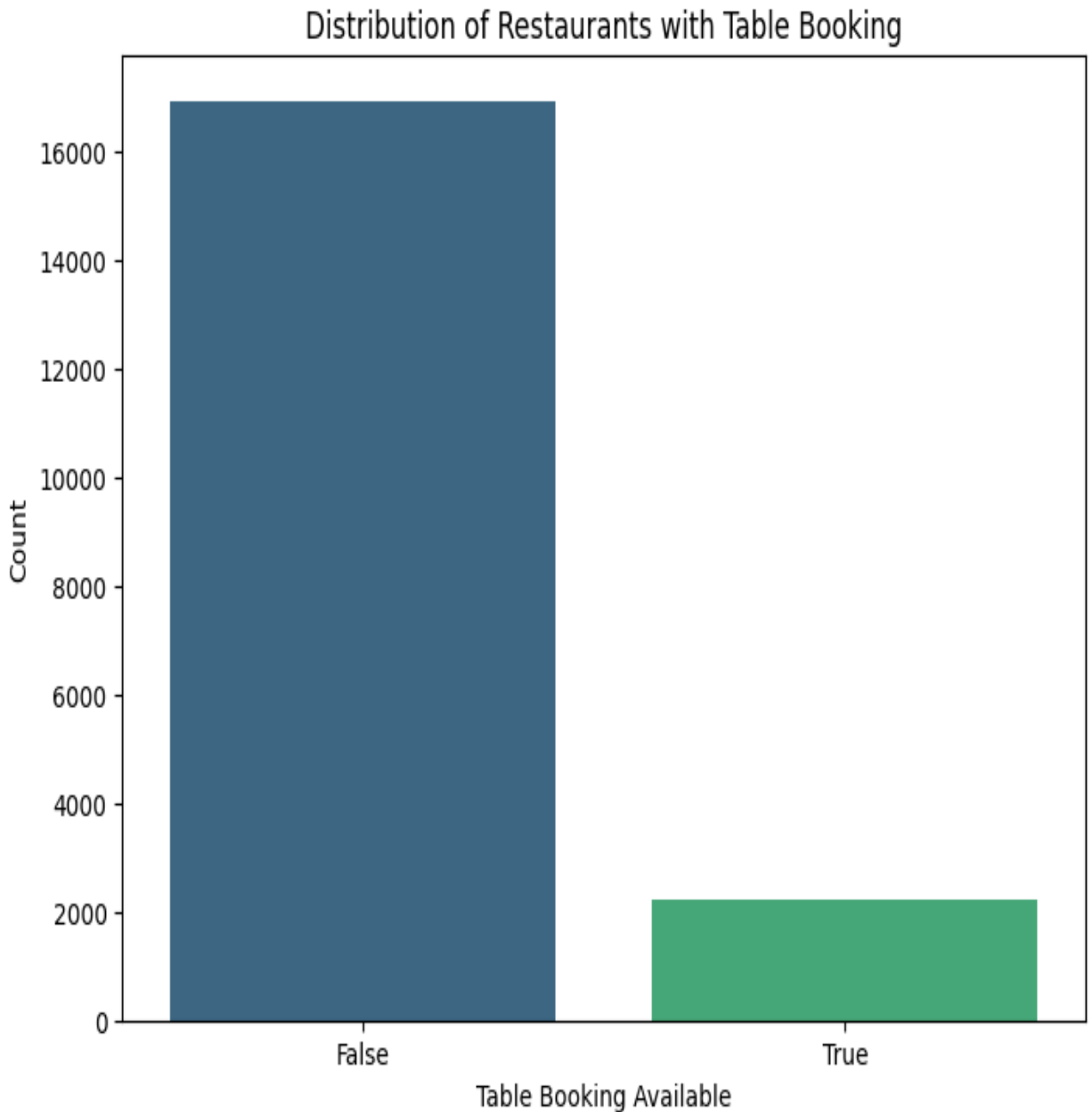
The bar chart is titled "Impact of Online Order Availability on Restaurant Ratings." It compares the average ratings of restaurants based on whether they offer online ordering or not. Here's a summary of the insights:

- Restaurants without Online Ordering: The average rating is approximately 3.0.
- Restaurants with Online Ordering: The average rating is approximately 3.5.

This suggests that restaurants offering online ordering tend to have higher average ratings compared to those that do not. This could be due to the convenience and accessibility that online ordering provides to customers.

DATA VISUALIZATION

Analyze the distribution of restaurants that offer table booking.



KEY INSIGHTS

The bar chart is titled "Distribution of Restaurants with Table Booking." Here's a summary of the insights:

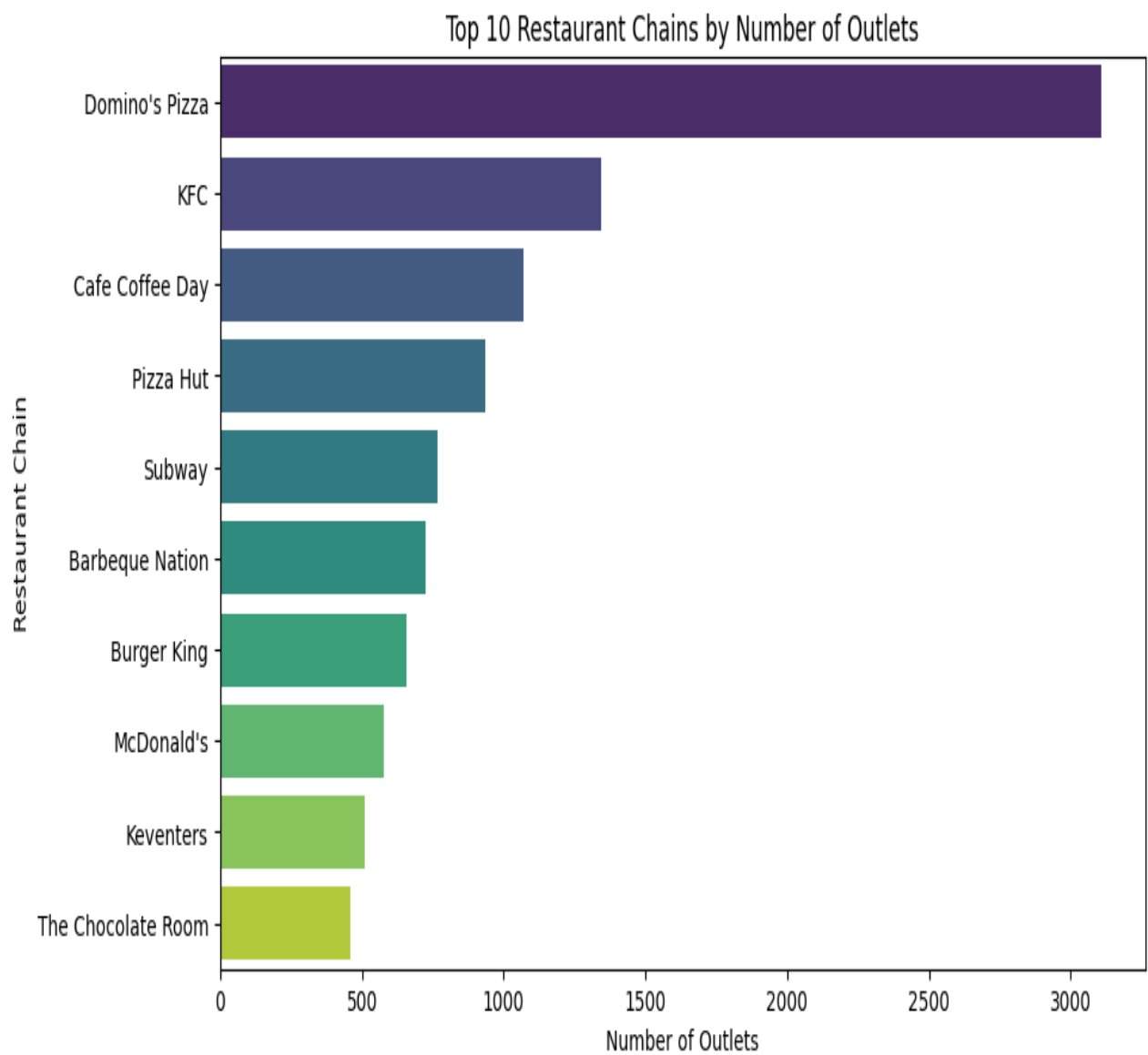
- **Table Booking Not Available:** A significantly higher number of restaurants do not offer table booking, with over 16,000 restaurants falling into this category.
- **Table Booking Available:** Only around 2,000 restaurants offer table booking services.

This information highlights the prevalence of restaurants that do not provide table booking services, which could be relevant for understanding customer service trends or planning for restaurant reservations

DATA VISUALIZATION

Top Restaurant Chains

Identify and visualize the top restaurant chains based on the number of outlets.



KEY INSIGHTS

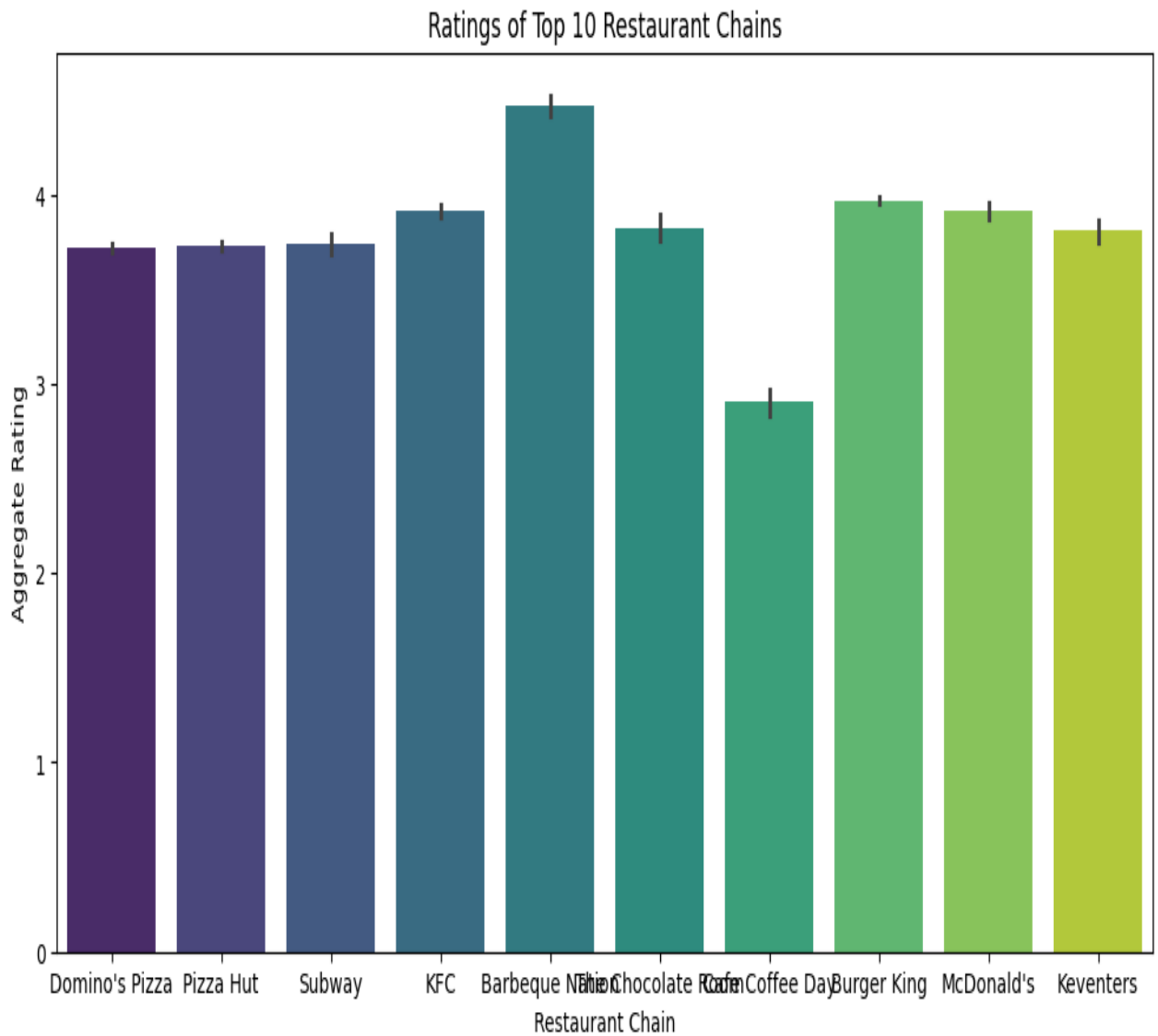
The bar chart is titled "Top 10 Restaurant Chains by Number of Outlets." It lists the top 10 restaurant chains and the number of their outlets. Here's a summary of the insights:

- Domino's Pizza has the highest number of outlets, indicating its widespread presence and popularity.
- KFC and Cafe Coffee Day follow closely behind, also having a significant number of outlets, suggesting their strong market penetration.
- Pizza Hut, Subway, and Barbeque Nation have a moderate number of outlets, indicating their presence in various locations.
- Burger King, McDonald's, Keventers, and The Chocolate Room have relatively fewer outlets compared to the top three chains but still maintain a considerable presence.

This chart provides a clear comparison of the scale of different restaurant chains based on their number of outlets. By comparing the heights of the bars, we can identify the chains with the most extensive network of outlets and infer their popularity and market dominance. Chains with taller bars have a greater number of outlets and a wider reach, while those with shorter bars have a smaller presence.

DATA VISUALIZATION

Explore the ratings of these top chains.



KEY INSIGHTS

The bar chart is titled "Ratings of Top 10 Restaurant Chains." It compares the aggregate ratings of different restaurant chains on a scale from 0 to 5. Here's a summary of the insights:

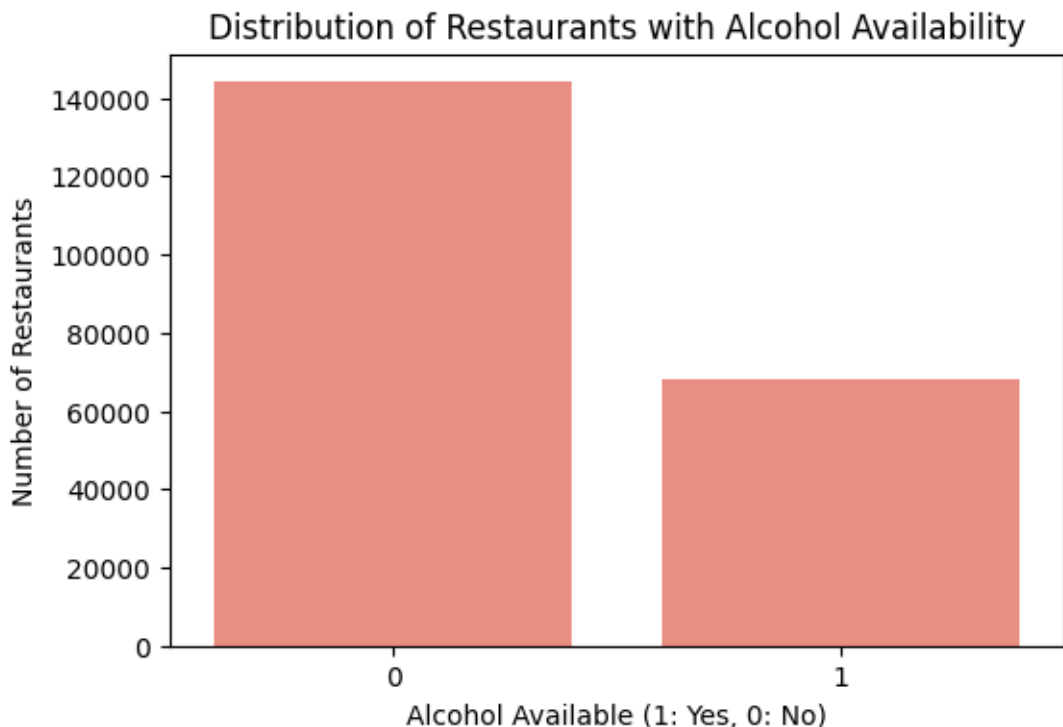
- Barbeque Nation has the highest rating, indicating it is the most popular or well-regarded among the listed chains.
- Burger King and Pizza Hut follow closely behind, suggesting they also enjoy high customer satisfaction.
- Domino's Pizza has the lowest rating, which might indicate room for improvement in customer satisfaction or popularity.
- The ratings for McDonald's, Baskin Robbins, Keventers, Subway, and KFC are relatively close, showing a competitive landscape among these food chains.

This chart provides a comparative overview of customer satisfaction or quality perception across these popular restaurant chains .

DATA VISUALIZATION

Restaurant Features

Analyze the distribution of restaurants based on features like Wi-Fi, Alcohol availability, etc.



KEY INSIGHTS

The first bar chart is titled "Distribution of Restaurants with Wi-Fi." It shows the number of restaurants that offer Wi-Fi services compared to those that do not. Here's a summary of the insights:

- **Restaurants without Wi-Fi:** There are over 200,000 restaurants that do not offer Wi-Fi services. This indicates that a significant majority of restaurants do not provide Wi-Fi to their customers.
- **Restaurants with Wi-Fi:** The chart does not show any bars for restaurants with Wi-Fi, suggesting that the number of restaurants offering Wi-Fi is negligible or not represented in the data.

This chart highlights the prevalence of restaurants that do not offer Wi-Fi services, which could be relevant for understanding customer preferences or planning for future amenities.

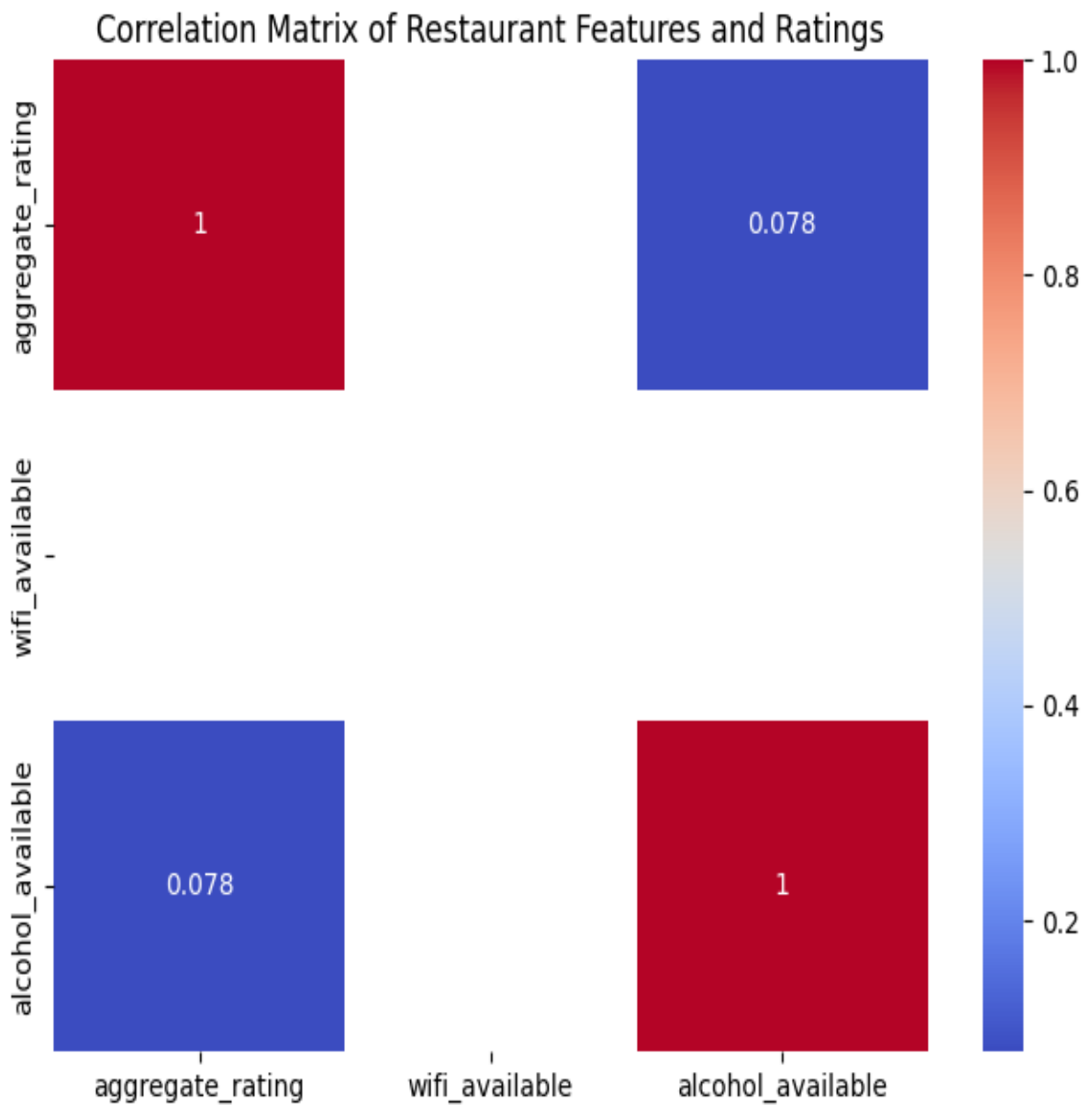
The second bar chart is titled "Distribution of Restaurants with Alcohol Availability." It shows the number of restaurants that offer alcohol services compared to those that do not. Here's a summary of the insights:

- Restaurants without Alcohol: There are around 140,000 restaurants that do not serve alcohol. This indicates that a significant majority of restaurants do not provide alcohol to their customers .
- Restaurants with Alcohol: There are around 80,000 restaurants that do serve alcohol. This shows that while a considerable number of restaurants offer alcohol, they are still outnumbered by those that do not.

This chart highlights the prevalence of restaurants that do not offer alcohol services, which could be relevant for understanding customer preferences or planning for future amenities.

DATA VISUALIZATION

Investigate if the presence of certain features correlates with higher ratings.



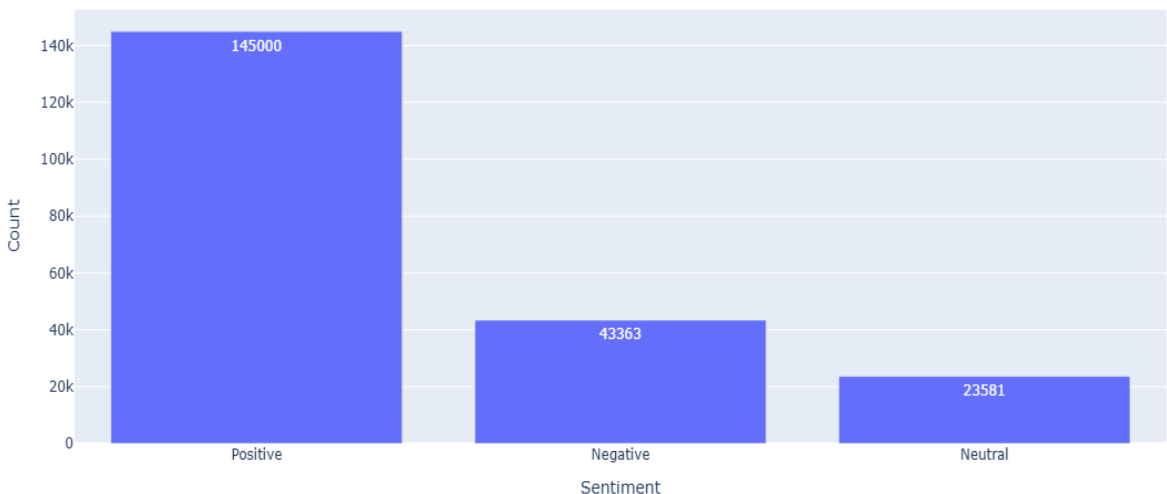
KEY INSIGHTS

The correlation matrix you uploaded shows the relationship between three variables: `aggregate_rating`, `wifi_available`, and `alcohol_available`. Here are the insights:

- **Aggregate Rating:** This is the overall rating of the restaurants.
- **Wi-Fi Availability:** Indicates whether the restaurant offers Wi-Fi services.
- **Alcohol Availability:** Indicates whether the restaurant serves alcohol.

The correlation values are represented by colors, with a color bar on the right indicating the range from -1 to 1. The diagonal elements show a perfect correlation of 1, as expected. The off-diagonal elements show the correlation between different features.

These values suggest a weak positive correlation between the availability of Wi-Fi and alcohol with the overall rating of the restaurants. This means that while there is a slight tendency for restaurants with Wi-Fi and alcohol to have higher ratings, the relationship is not very strong.



KEY INSIGHTS

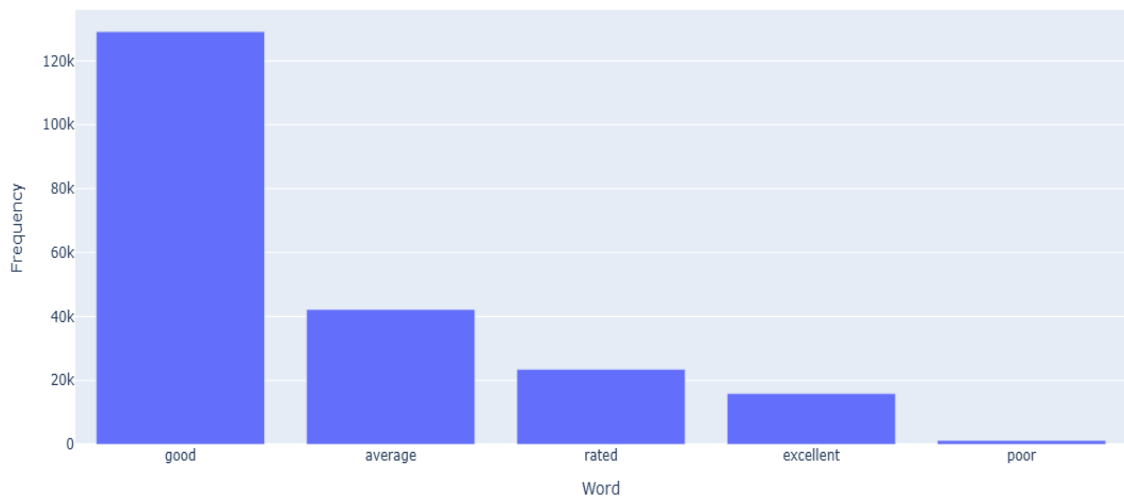
The word cloud is titled "Word Cloud for Customer Reviews." It visually represents the frequency of words used in customer reviews, with larger words indicating higher frequency. The most prominent words in the word cloud are "Good," "Average," and "Excellent," suggesting these are the most common terms used in the reviews. Other words like "rated" and "Poor" also appear but are less prominent. This visualization provides a quick and clear overview of customer sentiment and the most frequently mentioned terms in reviews.

The bar chart is titled "Sentiment Analysis of Customer Reviews." It categorizes customer reviews into three sentiments: Positive, Negative, and Neutral. Here are the key insights:

- **Positive Reviews:** The majority of customer reviews are positive, with a count of 145,000. This indicates a high level of customer satisfaction.
- **Negative Reviews:** There are 43,363 negative reviews, which is significantly lower than the positive reviews but still noteworthy.
- **Neutral Reviews:** The count of neutral reviews is 23,581, suggesting that a smaller portion of customers had an average experience.

Overall, the sentiment analysis shows that most customers are happy with their experiences, but there is still room for improvement to address the negative feedback.

Most Common Words in Reviews



The bar chart is titled "Most Common Words in Reviews." It displays the frequency of six words: "good," "average," "rated," "excellent," and "poor." Here are the key insights:

- **Good:** The word "good" has the highest frequency, with over 120,000 occurrences. This indicates that many reviewers found their experience to be positive.
- **Average:** The word "average" follows with around 50,000 occurrences, suggesting that a significant number of reviewers had a mediocre experience.
- **Rated:** The word "rated" appears approximately 30,000 times, indicating that many reviews included a rating.

- Excellent: The word "excellent" has about 20,000 occurrences, showing that some reviewers had an outstanding experience.
- Poor: The word "poor" has a very low frequency, close to zero, indicating that very few reviewers had a negative experience.

This chart highlights the most frequently used words in reviews, with "good" being the most common term used by reviewers.

FINAL REPORT

Summarizing the key findings and insights obtained from the analysis. Providing recommendations for restaurant owners and Zomato users based on the identified success factors :

Key Findings and Insights

Restaurant Ratings and Popularity:

- Barbeque Nation consistently receives high ratings, indicating strong customer satisfaction and popularity.
- Burger King and Pizza Hut also enjoy high customer satisfaction. Domino's Pizza, despite having the most outlets, has the lowest aggregate rating among the top chains, suggesting a potential need for improvement.
- Chennai, Mumbai, and Bangalore have the highest concentration of restaurants, indicating a vibrant and competitive food scene.
- Restaurants in higher price ranges tend to receive better ratings, potentially due to higher customer expectations.

Cuisine and Price:

- Average cost for two people varies across price ranges, with Price Range 2 having the highest average cost.

- North Indian, Chinese, and Fast Food are the most popular cuisines. There's a positive correlation between the variety of cuisines offered and restaurant ratings, suggesting that offering diverse options can enhance customer satisfaction.

Online Ordering and Table Booking:

- Restaurants offering online ordering tend to have higher average ratings compared to those that don't, highlighting the importance of convenience and accessibility.
- Table booking is not widely available, indicating a potential opportunity for restaurants to implement this feature to improve customer experience.

Restaurant Features:

- Most restaurants do not offer Wi-Fi, indicating a potential area for improvement.
- Alcohol availability is more common than Wi-Fi, although a majority of restaurants still do not serve alcohol.
- There's a weak positive correlation between Wi-Fi and alcohol availability with restaurant ratings.

Customer Sentiment:

- Sentiment analysis reveals that the majority of customer reviews are positive, indicating a generally positive dining experience.
- "Good," "average," and "excellent" are the most frequently used words in reviews.

RECOMMENDATIONS

For Restaurant Owners:

- **Focus on Quality and Customer Satisfaction:** Prioritize delivering high-quality food and excellent customer service to achieve high ratings and positive reviews. Consider strategies to address negative feedback and continuously improve customer experience.
- **Offer Online Ordering:** Implement online ordering services to enhance convenience and accessibility for customers, which can lead to higher ratings.
- **Diversify Cuisine Options:** Consider offering a variety of cuisines to cater to a broader customer base and potentially improve ratings.
- **Consider Table Booking:** Explore the possibility of implementing table booking services to enhance customer experience and convenience.
- **Improve Amenities:** Evaluate the potential benefits of offering Wi-Fi and other amenities to attract more customers and improve ratings.
- **Monitor Customer Reviews:** Actively track and analyze customer reviews to understand sentiments and identify areas for improvement.

For Zomato Users:

- Explore Top-Rated Restaurants: Use ratings and reviews to discover popular and highly-rated restaurants like Barbeque Nation, Burger King, and Pizza Hut.
- Filter by Cuisine and Price: Utilize Zomato's filters to find restaurants that offer preferred cuisines and fit within budget constraints.
- Consider Online Ordering: Opt for restaurants with online ordering for convenience and accessibility.
- Check for Table Booking: If planning a visit to a popular restaurant, check if table booking is available to avoid waiting times.
- Read Reviews: Before trying a new restaurant, read customer reviews to get an understanding of the overall experience and potential issues.
- Provide Feedback: Share your dining experiences through reviews to help other users and contribute to improving restaurant quality.



THANK YOU FOR READING

FOR CODING PART , KINDLY VISIT THE LINK BELOW

[Zomato Analysis.ipynb - Colab](#)