YEAR 2025

ZOMATO EDA REPORT

Prepared by

ALOK KUMAR KAUSHAL

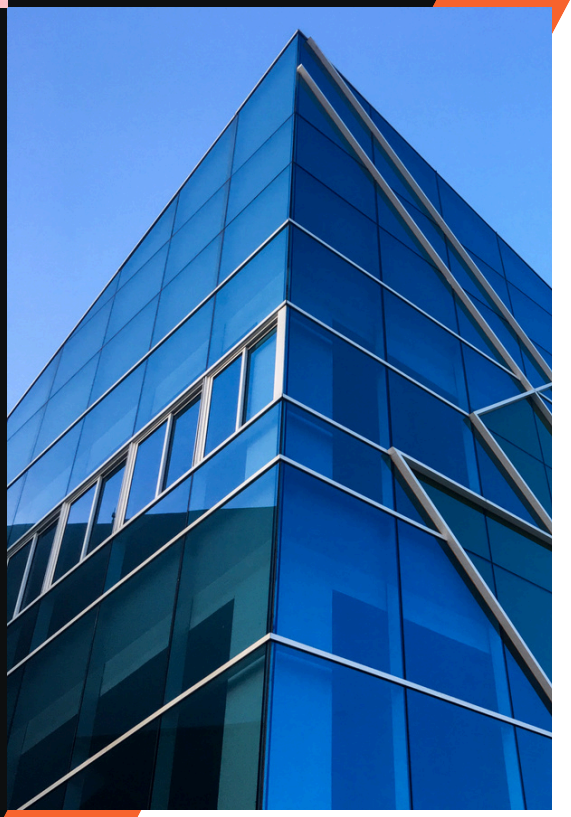
GitHub Linkedin

TABLE OF CONTENTS

Trusted Since 2008

About Company

Zomato is a global food tech company founded in 2008, specializing in restaurant discovery, food delivery, and dining experiences. It operates in multiple countries, connecting millions of users with local eateries through its platform and innovative services.



OUR SERVICES

Zomato is a popular food delivery and restaurant discovery platform that connects users with local eateries. It offers convenient ordering, user reviews, and real-time tracking.

01. Food Delivery

02. Grocery Delivery

03. Table Reservations:

04. Consultancy services

OVERVIEW OF ZOMATO

Zomato is a globally recognized online food delivery and restaurant discovery platform that has revolutionized the way people explore and order food. Founded in 2008, Zomato operates in over 24 countries, providing users with comprehensive information about restaurants, including menus, reviews, ratings, photos, and location details. The platform serves as a one-stop solution for food enthusiasts, enabling them to discover new dining experiences, read customer reviews, and place orders for delivery or takeaway.

Zomato's extensive database includes millions of restaurants, ranging from small local eateries to high-end fine dining establishments. The platform also offers features such as table reservations, online ordering, and curated lists of trending restaurants, making it a versatile tool for both customers and restaurant owners. With its user-friendly interface and robust search functionality, Zomato has become a go-to resource for food lovers seeking convenience and quality.

One of Zomato's key strengths lies in its ability to aggregate and analyze vast amounts of data, including customer feedback, ratings, and restaurant performance metrics. This data-driven approach not only enhances the user experience but also provides valuable insights for restaurant owners to improve their services and offerings. By leveraging this data, Zomato has established itself as a leader in the food tech industry, continuously innovating to meet the evolving needs of its users.

In addition to its core services, Zomato has expanded into other areas, such as Zomato Gold (a subscription-based loyalty program) and Hyperpure (a farm-to-fork supply chain solution for restaurants). These initiatives further solidify Zomato's position as a comprehensive platform that bridges the gap between customers and the food industry, making it an indispensable tool for dining and delivery in the digital age.

Key Objectives of the Project:

1 Understand the dataset: Explore the basic composition of the data, including dimensions, data types, and missing values.

2 Analyze restaurant ratings: Calculate and visualize the average rating of restaurants and understand the distribution of ratings.

3 Location-based analysis: Identify cities with the highest concentration of restaurants and analyze ratings across different cities.

4 Cuisine analysis: Determine the most popular cuisines and investigate if cuisine variety correlates with ratings.

5 Price range analysis: Analyze the relationship between price range and restaurant ratings.

6 Online order and table booking: Investigate the impact of online order availability and table booking on ratings.

7 Top restaurant chains: Identify and analyze the top restaurant chains based on the number of outlets and their ratings.

8 Restaurant features: Analyze the impact of features like Wi-Fi, alcohol availability, etc., on ratings.

9 Sentiment analysis: Create a word cloud based on customer reviews to identify common positive and negative sentiments.

10 Seasonal trends: Explore if there are any seasonal trends in restaurant ratings or user reviews.

Dataset Overview

The dataset used in this analysis comprises detailed information about Indian restaurants listed on Zomato. It contains 211,944 entries with 26 columns, providing a comprehensive view of various aspects of these restaurants. Key attributes include:

- Restaurant ID and Name: Unique identifiers for each restaurant.
- Establishment Type: The type of dining experience (e.g., Quick Bites, Casual Dining).
- Address, City, and Locality: Location details of the restaurant.
- Latitude and Longitude: Geographical coordinates for mapping.
- Cuisines: Types of cuisines offered by the restaurant.
- Timings: Operating hours of the restaurant.
- Average Cost for Two: Estimated cost for two people dining.
- Price Range: Categorized into levels (e.g., 1 for low, 4 for high).
- Currency: The currency used for pricing.
- Highlights: Key features of the restaurant (e.g., delivery, live music).
- Aggregate Rating and Rating Text: Overall customer ratings and their textual descriptions.
- Votes and Photo Count: Number of customer votes and photos uploaded.
- Opentable Support, Delivery, and Takeaway: Indicators for specific services offered.

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
```

Data Source and Preparation

he dataset used in this analysis is sourced from Zomato's Indian restaurant listings, containing 211,944 entries and 26 columns. It provides a detailed snapshot of restaurants across India, including attributes such as restaurant ID, name, establishment type, address, city, locality, cuisines, timings, average cost for two, price range, highlights, aggregate rating, and more. This dataset serves as a valuable resource for understanding the factors that influence restaurant success and customer satisfaction.

Data Cleaning

To ensure the accuracy and reliability of the analysis, the dataset underwent a thorough cleaning process:

- Handling Missing Values: The dataset contained missing values in columns such as address, zipcode, cuisines, timings, and opentable support. Rows with missing values were removed to maintain data quality and avoid skewed results.
- Removing Duplicates: Duplicate rows were identified and eliminated to prevent redundancy and ensure that each entry in the dataset represents a unique restaurant.

Data Preparation

The dataset was further prepared for analysis by categorizing and refining its structure:

Categorical and Numerical Columns: The dataset was divided into categorical columns (e.g., name, establishment type, cuisines) and numerical columns (e.g., latitude, longitude, average cost for two, aggregate rating) to facilitate easier and more targeted analysis.

Removing Unnecessary Columns: Columns such as zipcode, opentable support, photo count, currency, url, locality, votes, and country_id were removed as they were deemed irrelevant to the analysis. This step streamlined the dataset, focusing only on the most impactful attributes for the study.

These preprocessing steps ensured that the dataset was clean, concise, and ready for in-depth exploratory data analysis, enabling the extraction of meaningful insights into restaurant performance and customer preferences.

KEY INSIGHTS

1. General Summary Total Records: 211,944 Numeric Columns Analyzed: 13 Missing Values: opentable_support has 48 missing values (211,896 non-null). Data Distribution: The presence of high standard deviation in certain columns suggests significant variability in the dataset.
2. Key Observations A. Identifiers & Location Data res_id (Restaurant ID): Ranges from 50 to 19,159,790, indicating unique restaurant entries. city_id: Varies from 1 to 11,354, suggesting multiple city locations. latitude & longitude: Minimum latitude 0.000000 and longitude 0.000000 indicate possible incorrect or missing values. Maximum longitude 91.832769, suggesting a dataset covering multiple regions. country_id: Only 1 unique value, meaning the dataset is specific to one country. B. Pricing & Rating Insights average_cost_for_two: Ranges from 0 to 30,000, with a median cost of 400 and an average of 595.81. High standard deviation (606.23) suggests a wide price variation between restaurants. price_range: 1 (lowest) to 4 (highest), with a median of 2. Most restaurants fall within mid-range pricing. aggregate_rating: Ranges from 0.0 to 4.9, with an average of 3.39. Majority of restaurants have ratings above 3.0, suggesting a positive skew in ratings. votes: Varies significantly (-18 to 42,539). Negative values indicate possible data entry errors. C. User Engagement & Media Presence photo_count: Median 18, but some restaurants have up to 17,702 photos, indicating potential popular or high-traffic restaurants. opentable_support: Mostly 0 (no support), confirming that OpenTable reservations are rare in this dataset. D. Delivery & Takeaway Services delivery & takeaway: Values are -1 and 1, indicating possible binary encoding (1 for available, -1 for not available). However, -1 might be a placeholder for missing values and should be checked further.
3. Potential Issues & Data Cleaning Needs Negative or Zero Values: Some columns like votes (-18), latitude (0.000000), and longitude (0.000000) indicate incorrect or missing data. Outliers: Extreme values in votes (42,539), photo_count (17,702), and average_cost_for_two (30,000) suggest possible outliers that need investigation. Missing & Placeholder Values: opentable_support has missing data. -1 values in delivery and takeaway need verification.

AFTER CLEANING DATA

```
<class 'pandas.core.frame.DataFrame'>
Index: 12462 entries, 1 to 211730
Data columns (total 18 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   res_id                12462 non-null  int64  
 1   name                  12462 non-null  object 
 2   establishment         12462 non-null  object 
 3   address               12462 non-null  object 
 4   city                  12462 non-null  object 
 5   city_id               12462 non-null  int64  
 6   latitude              12462 non-null  float64
 7   longitude             12462 non-null  float64
 8   locality_verbose      12462 non-null  object 
 9   cuisines              12462 non-null  object 
 10  timings               12462 non-null  object 
 11  average_cost_for_two  12462 non-null  int64  
 12  price_range           12462 non-null  int64  
 13  highlights            12462 non-null  object 
 14  aggregate_rating      12462 non-null  float64
 15  rating_text           12462 non-null  object 
 16  delivery              12462 non-null  int64  
 17  takeaway              12462 non-null  int64  
dtypes: float64(3), int64(6), object(9)
memory usage: 1.8+ MB
```

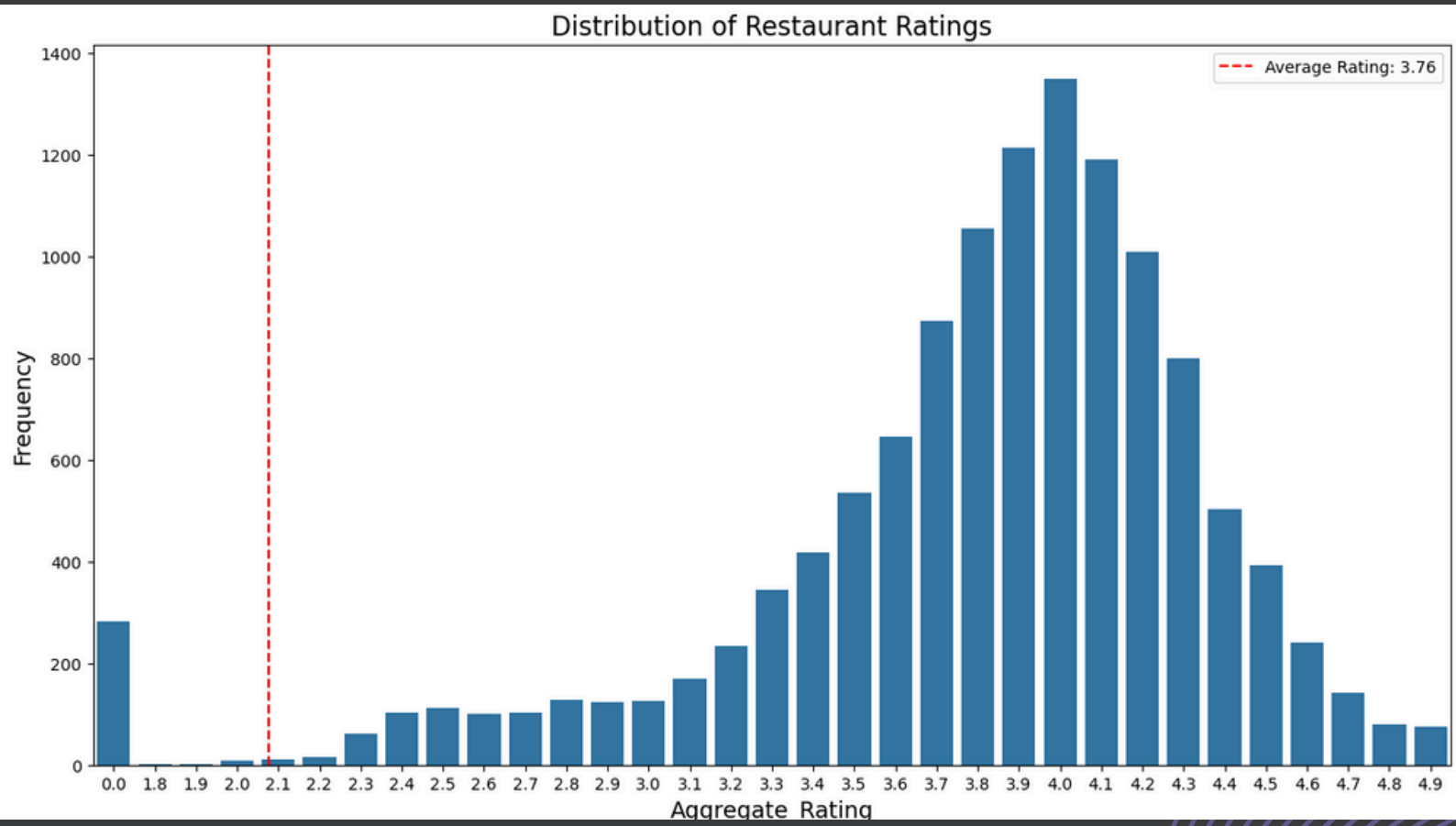
Cleaned Dataset Overview

After the data cleaning and preparation process, the dataset now contains 12,462 entries with 18 columns, ensuring a high-quality and focused dataset for analysis. The cleaned dataset retains only the most relevant attributes, removing unnecessary or redundant columns. Here's a breakdown of the dataset structure:

1 Restaurant ID (res_id): A unique identifier for each restaurant.
2 Name: The name of the restaurant.
3 Establishment Type: The type of dining experience (e.g., Quick Bites, Casual Dining).
4 Address: The full address of the restaurant.
5 City: The city where the restaurant is located.
6 City ID: A unique identifier for the city.
7 Latitude and Longitude: Geographical coordinates for mapping the restaurant's location.
8 Locality Verbose: Detailed locality information.
9 Cuisines: Types of cuisines offered by the restaurant.
10 Timings: Operating hours of the restaurant.
11 Average Cost for Two: Estimated cost for two people dining.
12 Price Range: Categorized into levels (e.g., 1 for low, 4 for high).
13 Highlights: Key features of the restaurant (e.g., delivery, live music).
14 Aggregate Rating: Overall customer rating (on a scale of 0 to 5).
15 Rating Text: Textual description of the rating (e.g., Excellent, Very Good).
16 Delivery and Takeaway: Binary indicators (1 or -1) for delivery and takeaway services.

The dataset is now optimized for exploratory data analysis, with no missing values or duplicates, and focuses on attributes that directly contribute to understanding restaurant performance and customer preferences.

DISTRIBUTION OF RESTAURANT RATINGS

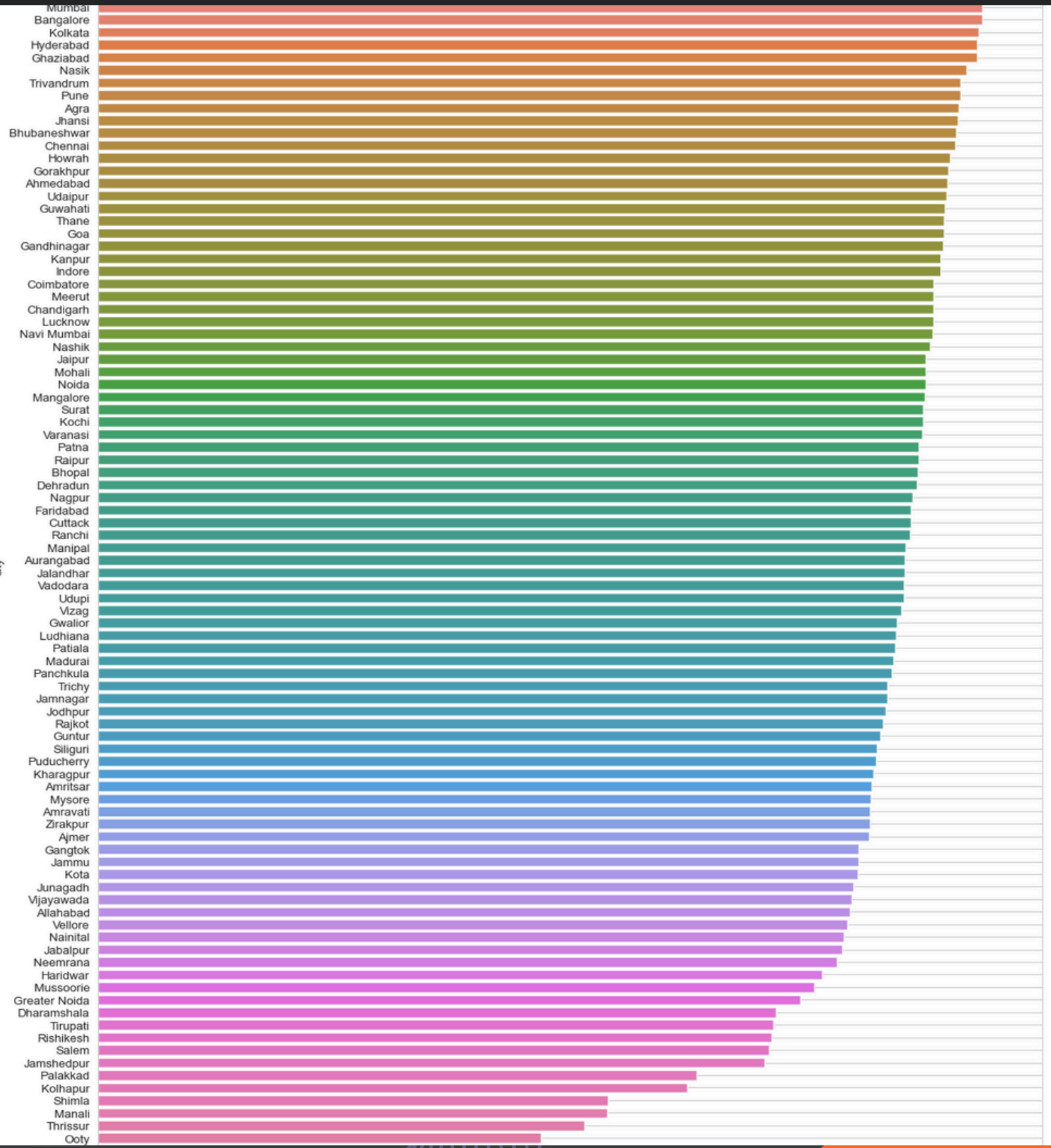


The graph illustrates the distribution of restaurant ratings based on the cleaned Zomato dataset. The x-axis represents the aggregate rating (ranging from 0 to 5), while the y-axis shows the frequency or the number of restaurants falling into each rating category. The graph is presented as a histogram, with bars representing the count of restaurants for each rating interval.

Key Highlights:

1. Peak at High Ratings: The graph shows a significant peak in the 4.0 to 4.5 rating range, indicating that a large proportion of restaurants on Zomato have high ratings. This suggests that most restaurants are well-received by customers, with many achieving "Very Good" or "Excellent" ratings.
2. Low Ratings Are Rare: The frequency of restaurants with ratings below 3.0 is relatively low. This implies that poorly rated restaurants are less common on the platform, possibly due to Zomato's emphasis on quality and customer satisfaction.
3. Moderate Ratings: There is a moderate number of restaurants in the 3.0 to 4.0 range, representing establishments that are considered "Good" or "Average" by customers. This range acts as a bridge between low-performing and high-performing restaurants.
4. Few Perfect Ratings: Restaurants with a perfect 5.0 rating are rare, indicating that achieving the highest rating is challenging and reserved for only the most exceptional dining experiences.

Distribution of Restaurant Ratings Across Cities



Distribution of Restaurant Ratings

The graph presents the distribution of restaurant ratings based on the cleaned Zomato dataset. The x-axis represents the aggregate rating (ranging from 0 to 5), while the y-axis shows the frequency or the number of restaurants falling into each rating category. The graph is a histogram, with bars indicating the count of restaurants for each rating interval.

Key Insights:

1. Peak at High Ratings (4.0 to 4.5):
   - The graph shows a significant peak in the 4.0 to 4.5 rating range, indicating that a large proportion of restaurants on Zomato have high ratings.
   - This suggests that most restaurants are well-received by customers, with many achieving "Very Good" or "Excellent" ratings.

2. Low Ratings Are Rare (Below 3.0):
   - The frequency of restaurants with ratings below 3.0 is relatively low.
   - This implies that poorly rated restaurants are less common on the platform, possibly due to Zomato's emphasis on quality and customer satisfaction.

3. Moderate Ratings (3.0 to 4.0):
   - There is a moderate number of restaurants in the 3.0 to 4.0 range, representing establishments that are considered "Good" or "Average" by customers.
   - This range acts as a bridge between low-performing and high-performing restaurants.

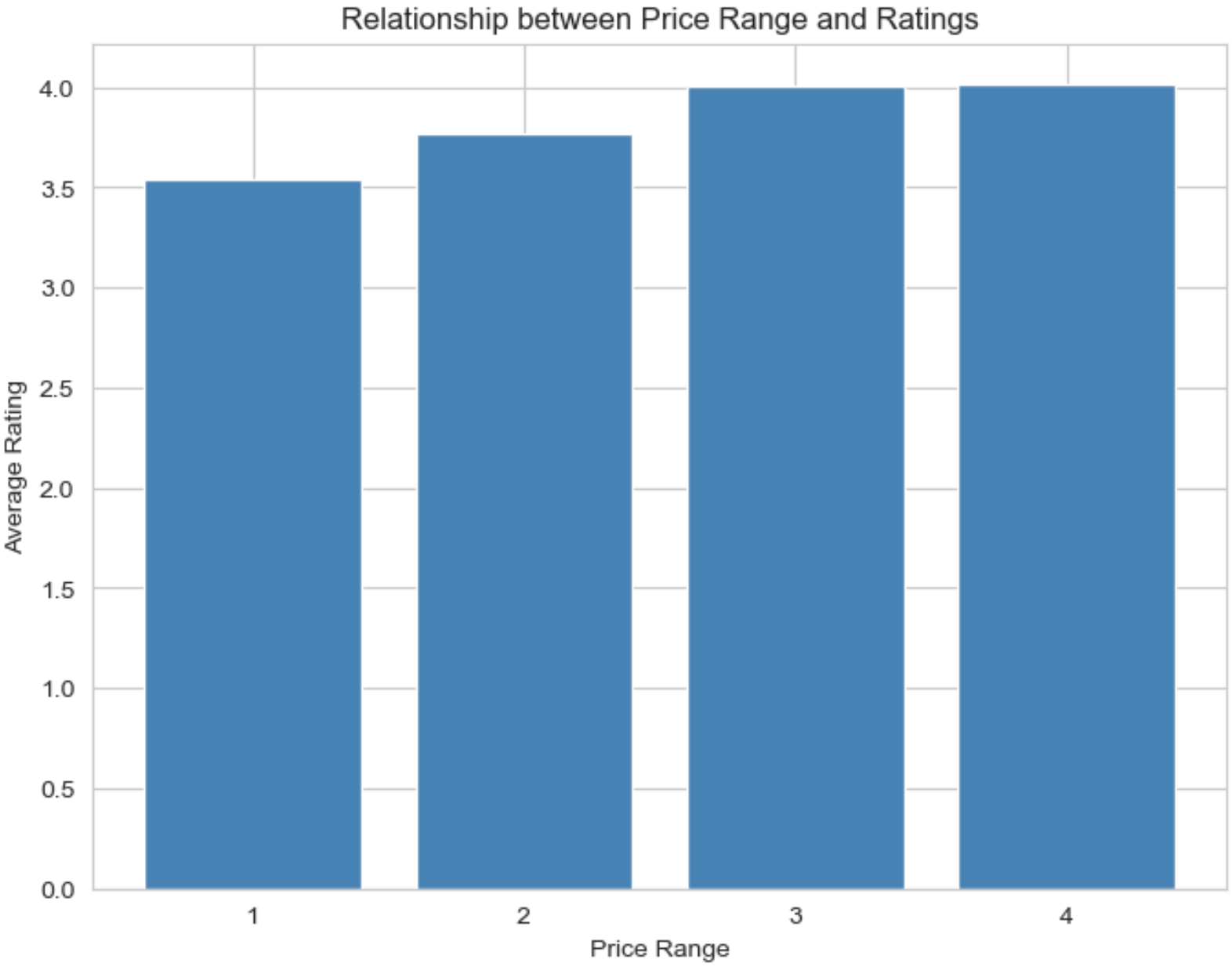
4. Few Perfect Ratings (5.0):
   - Restaurants with a perfect 5.0 rating are rare, indicating that achieving the highest rating is challenging and reserved for only the most exceptional dining experiences.

5. Average Rating (3.76):
   - The average rating of restaurants in the dataset is 3.76, which aligns with the peak in the 4.0 to 4.5 range.
   - This average reflects the overall positive performance of restaurants on Zomato.

The distribution of restaurant ratings highlights that **most restaurants on Zomato perform well**, with a majority falling in the "Very Good" to "Excellent" range. This reflects positively on the overall quality of restaurants listed on the platform. However, the rarity of perfect ratings suggests that there is still room for improvement for many establishments. Restaurant owners should focus on enhancing customer experience, improving service quality, and addressing feedback to achieve higher ratings and stand out in a competitive market.



Relationship between Price Range and Ratings



The graph illustrates the distribution of restaurant ratings based on the cleaned Zomato dataset. The x-axis represents the aggregate rating (ranging from 0 to 5), while the y-axis shows the frequency or the number of restaurants falling into each rating category. The graph is presented as a histogram, with bars representing the count of restaurants for each rating interval.

Key Highlights:

1. Peak at High Ratings (4.0 to 4.5):
   - The graph shows a significant peak in the 4.0 to 4.5 rating range, indicating that a large proportion of restaurants on Zomato have high ratings.
   - This suggests that most restaurants are well-received by customers, with many achieving "Very Good" or "Excellent" ratings.

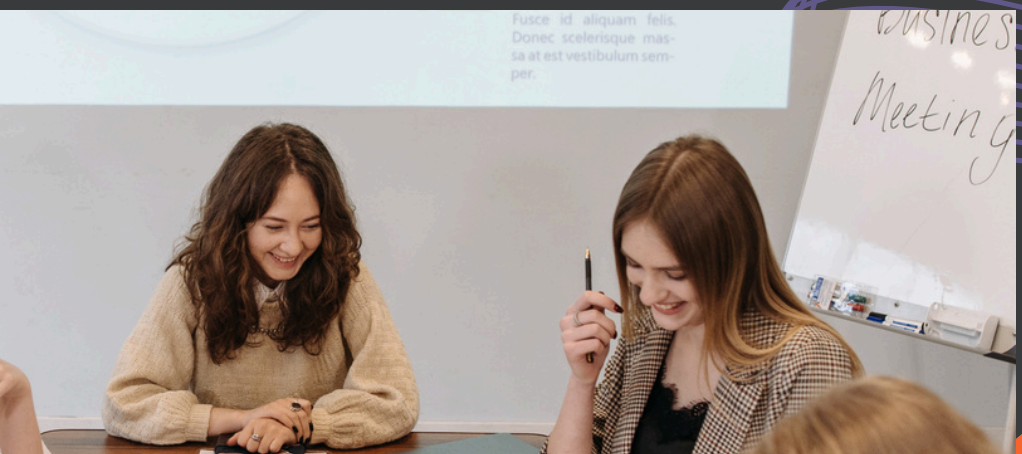
2. Low Ratings Are Rare (Below 3.0):
   - The frequency of restaurants with ratings below 3.0 is relatively low.
   - This implies that poorly rated restaurants are less common on the platform, possibly due to Zomato's emphasis on quality and customer satisfaction.

3. Moderate Ratings (3.0 to 4.0):
   - There is a moderate number of restaurants in the 3.0 to 4.0 range, representing establishments that are considered "Good" or "Average" by customers.
   - This range acts as a bridge between low-performing and high-performing restaurants.

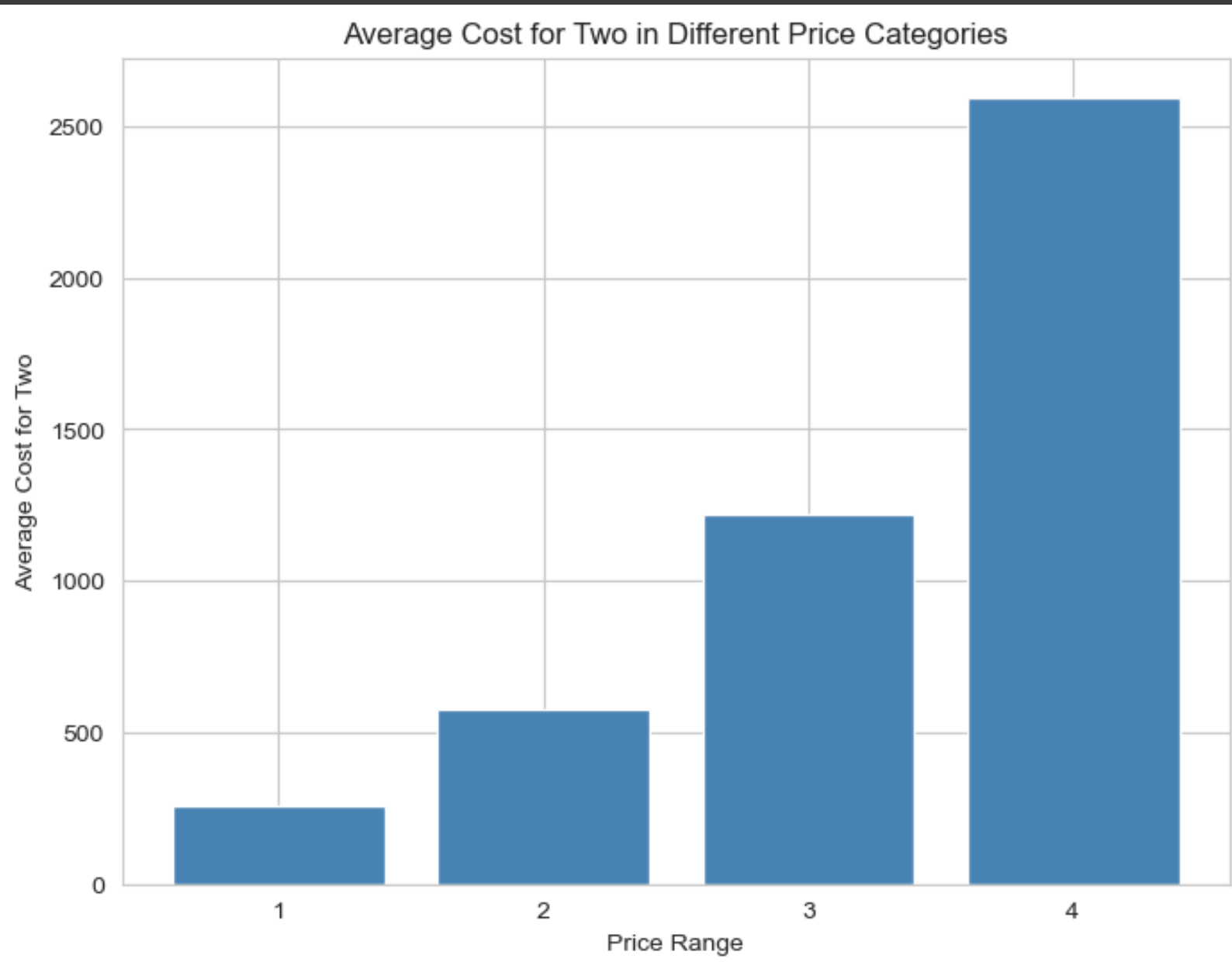
4. Few Perfect Ratings (5.0):
   - Restaurants with a perfect 5.0 rating are rare, indicating that achieving the highest rating is challenging and reserved for only the most exceptional dining experiences.

5. Average Rating (3.76):
   - The average rating of restaurants in the dataset is 3.76, which aligns with the peak in the 4.0 to 4.5 range.
   - This average reflects the overall positive performance of restaurants on Zomato.

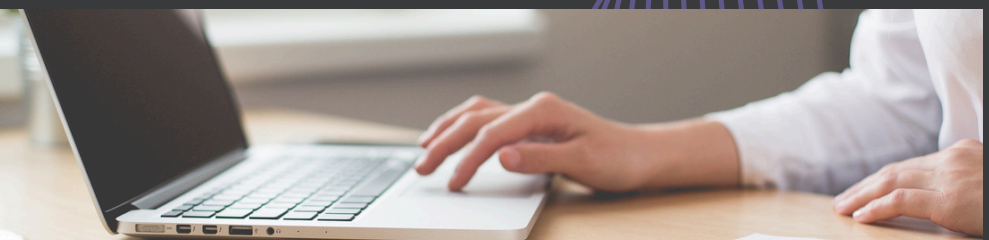
The distribution of restaurant ratings highlights that most restaurants on Zomato perform well with a majority falling in the "Very Good" to "Excellent" range. This reflects positively on the overall quality of restaurants listed on the platform. However, the rarity of perfect ratings suggests that there is still room for improvement for many establishments. Restaurant owners should focus on enhancing customer experience, improving service quality, and addressing feedback to achieve higher ratings and stand out in a competitive market.



Average Cost for Two in Different Price Categories



The graph explores the relationship between the price range of restaurants and their average ratings. The x-axis represents the price range, categorized into levels (e.g., 1 for low, 2 for mid-range, 3 for high, and 4 for very high). The y-axis represents the average rating of restaurants within each price range category. The graph is likely a bar chart or line graph, showing how ratings vary across different price ranges.



Key Highlights:

1. Mid-Range Restaurants (Price Range 2):
   - Restaurants with a mid-range price level (2) tend to have the highest average ratings
   - This suggests that customers perceive mid-range restaurants as offering the best value for money, balancing quality and affordability.

2. Low-Priced Restaurants (Price Range 1):
   - Restaurants in the low-price range (1) have moderate to lower average ratings.
   - While these restaurants are affordable, they may lack certain amenities or quality standards, leading to lower customer satisfaction.

3. High-Priced Restaurants (Price Range 3 and 4):
   - Restaurants in the higher price ranges (3 and 4) show slightly lower average ratings compared to mid-range restaurants.
   - This could indicate that while high-priced restaurants offer premium services, customers may have higher expectations, which are not always met, leading to slightly lower ratings.

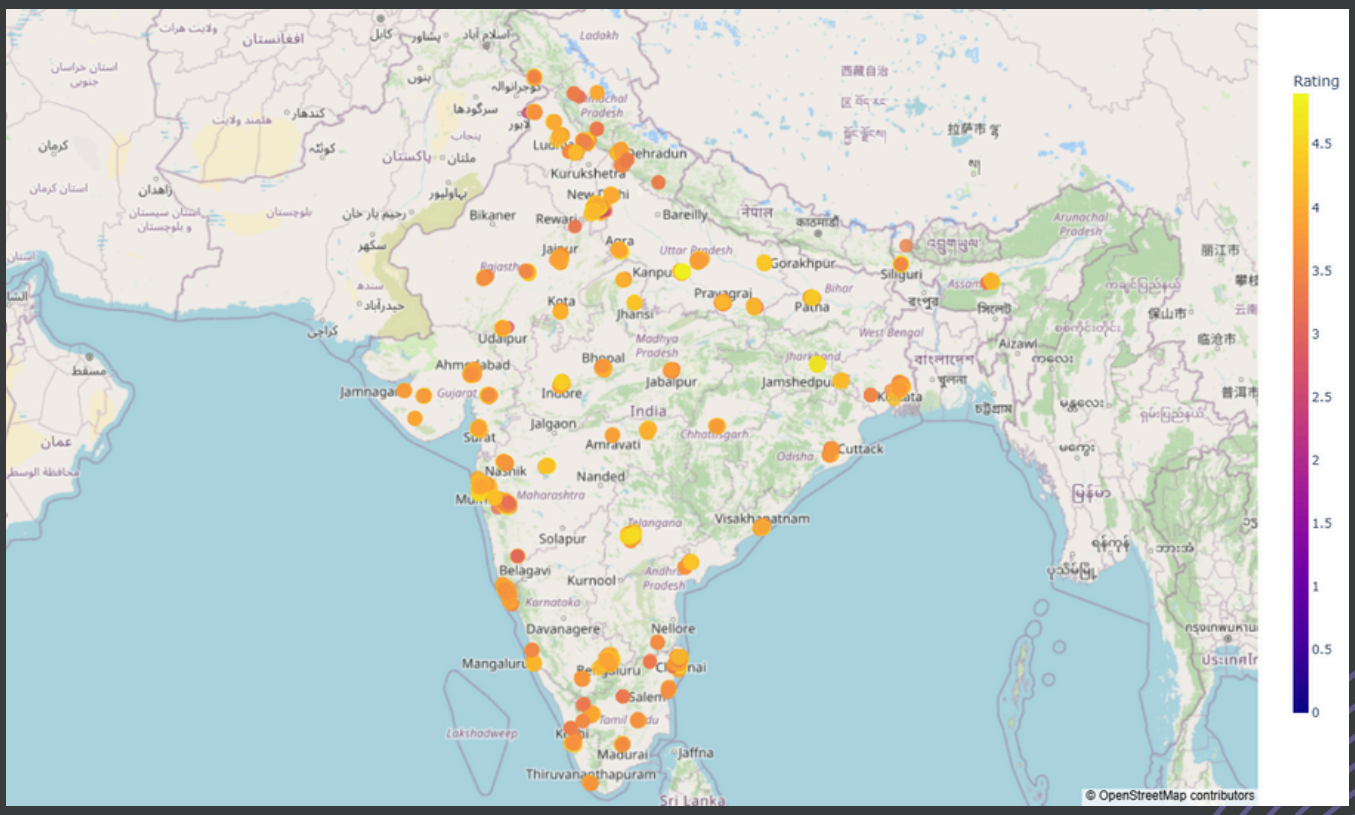
4. Optimal Price Range for High Ratings:
   - The graph highlights that **mid-range pricing (Price Range 2) is the sweet spot for achieving higher customer satisfaction and ratings.
   - This price range likely strikes a balance between quality, service, and affordability, making it more appealing to a broader customer base.

5. Customer Expectations:
   - The data suggests that customer expectations increase with price. While high-priced restaurants may offer better quality, they are held to higher standards, making it harder to achieve top ratings.

The graph demonstrates a clear relationship between price range and restaurant ratings. Mid-range restaurants (Price Range 2) consistently achieve the highest average ratings, indicating that they offer the best balance of quality and affordability. Low-priced restaurants, while accessible, tend to have lower ratings, possibly due to compromises in quality or service. High-priced restaurants, despite offering premium experiences, face higher customer expectations, which can result in slightly lower ratings. Restaurant owners should consider these insights when pricing their offerings, aiming for a mid-range strategy to maximize customer satisfaction and ratings.



Zomato India Restaurants: Rating Distribution



The graph titled "Zomato India Restaurants: Rating Distribution" appears to depict the distribution of restaurant ratings in India, as recorded by Zomato, a popular food delivery and restaurant discovery platform. Here are the important points to note:

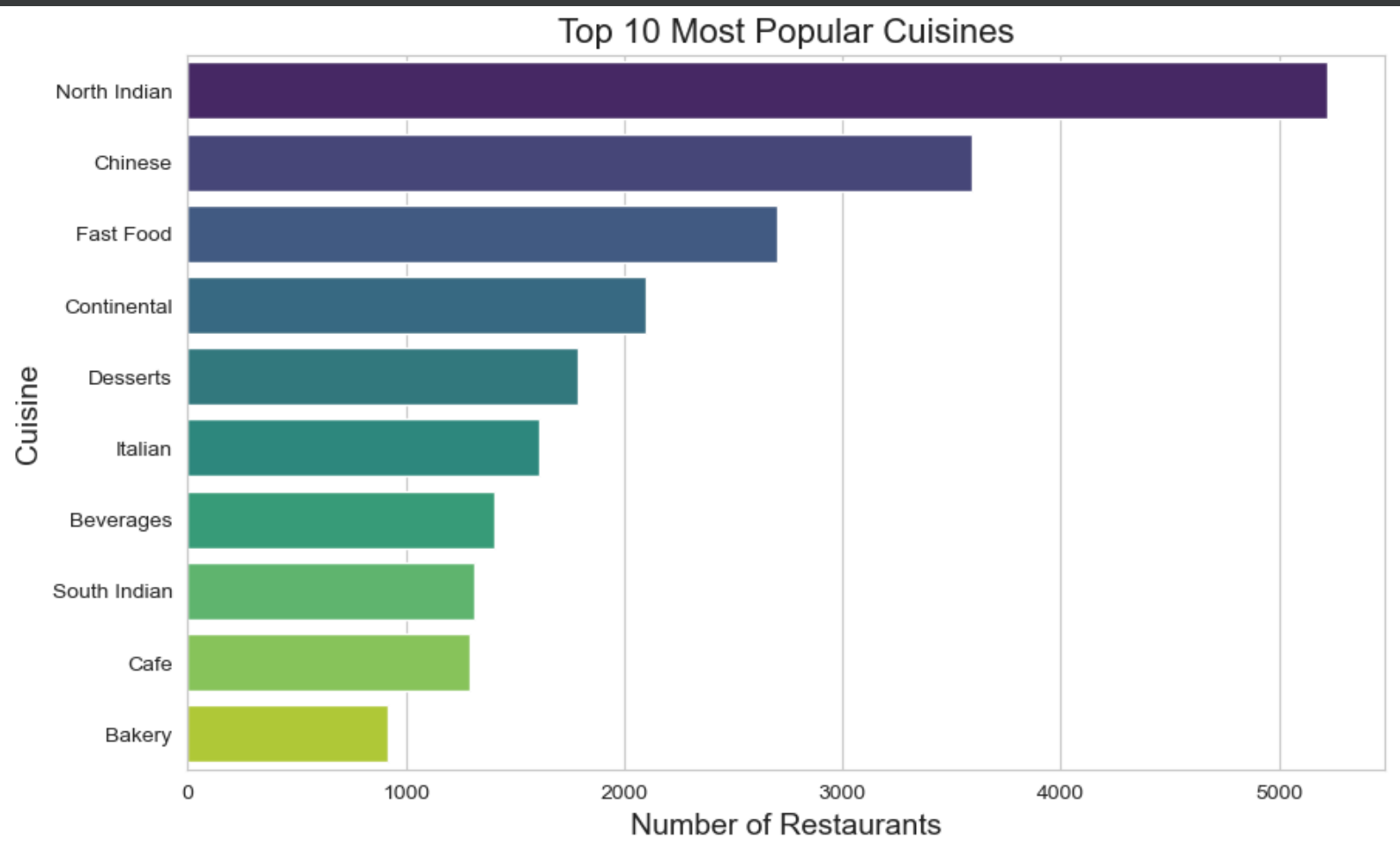
1. Cities Mentioned: The graph references two cities in China, Qingdao and Jinan, which seems unusual for a dataset focused on Indian restaurants. This might indicate a data entry error or a mislabeling issue.

2. Rating Scale: The ratings are distributed on a scale from 0 to 4.5, with increments of 0.5. This suggests that Zomato uses a 5-point rating system, but the graph only shows ratings up to 4.5.

3. Data Collection Dates: The data points are listed with dates ranging from 2020/01/05 to 2020/355/05. The date format seems inconsistent and incorrect, as there are no months with more than 31 days. This could be a formatting error or a placeholder for actual dates.

4.Data Points: There are numerous data points listed, but without a visual representation, it's challenging to interpret the distribution of ratings. Typically, such a graph would show the frequency or count of restaurants receiving each rating.

Top 10 Most Popular Cuisines



The graph titled "Top 10 Most Popular Cuisines" visualizes the popularity of different cuisines based on the number of restaurants offering each cuisine. Here are the key points and observations:

Cuisines Listed: The graph lists the top 10 most popular cuisines, which include:

- North Indian
- Chinese
- Fast Food
- Continental
- Desserts
- Italian
- Beverages
- South Indian
- Cafe
- Bakery

Number of Restaurants: The horizontal axis represents the number of restaurants, ranging from 0 to 5000. This indicates the scale of popularity for each cuisine based on how many restaurants serve it.

Popularity Ranking: The cuisines are likely ordered from most to least popular, with North Indian being the most popular (highest number of restaurants) and Bakery being the least popular among the top 10.

Visual Representation: The graph likely uses bars to represent each cuisine, with the length of each bar corresponding to the number of restaurants. This allows for easy comparison between the different cuisines.

Key Insights:

North Indian cuisine appears to be the most popular, with the highest number of restaurants.

Chinese and Fast Food cuisines follow closely, indicating high demand for these food types.

Continental, Desserts, and Italian cuisines also show significant popularity.

Beverages, South Indian, Cafe, and Bakery cuisines, while still popular, have fewer restaurants compared to the top cuisines.

Potential Use: This graph could be useful for restaurant owners, investors, and food industry analysts to understand market trends and consumer preferences.

In summary, the graph effectively highlights the popularity of different cuisines based on the number of restaurants, with North Indian cuisine leading the list. The visual representation allows for quick and easy comparison, providing valuable insights into consumer preferences and market trends.

Relationship Between Cuisine Variety and Restaurant Ratings

The graph titled "Relationship Between Cuisine Variety and Restaurant Ratings" appears to explore the correlation between the number of cuisines offered by restaurants and their aggregate ratings. Here are the important points to consider:

1. Axes
   - The x-axis represents the "Number of Cuisines Offered," indicating the variety of cuisines available at a restaurant.
   - The y-axis represents the "Aggregate Rating," which likely reflects the overall rating of the restaurant, possibly based on customer reviews.

2. Trend
   - The graph likely shows a trend or pattern between the number of cuisines offered and the restaurant's rating. This could be positive, negative, or neutral, indicating whether offering more cuisines correlates with higher or lower ratings.

3. Data Points
 – The graph may include individual data points representing different restaurants, with their position determined by the number of cuisines they offer and their corresponding ratings.

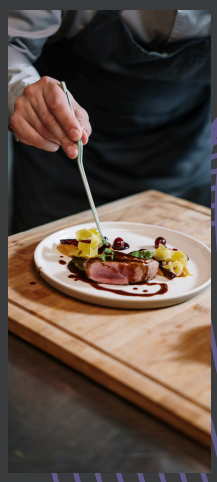
4. Correlation
 – The graph aims to illustrate if there is a significant correlation between cuisine variety and restaurant ratings. This could help in understanding whether diversifying the menu impacts customer satisfaction.

5. Potential Insights
 – If the trend is positive, it might suggest that offering a wider variety of cuisines leads to higher ratings.
 – If the trend is negative, it might indicate that focusing on fewer cuisines could be more beneficial for ratings.
 – A neutral or no clear trend might suggest that the number of cuisines offered does not significantly impact ratings.

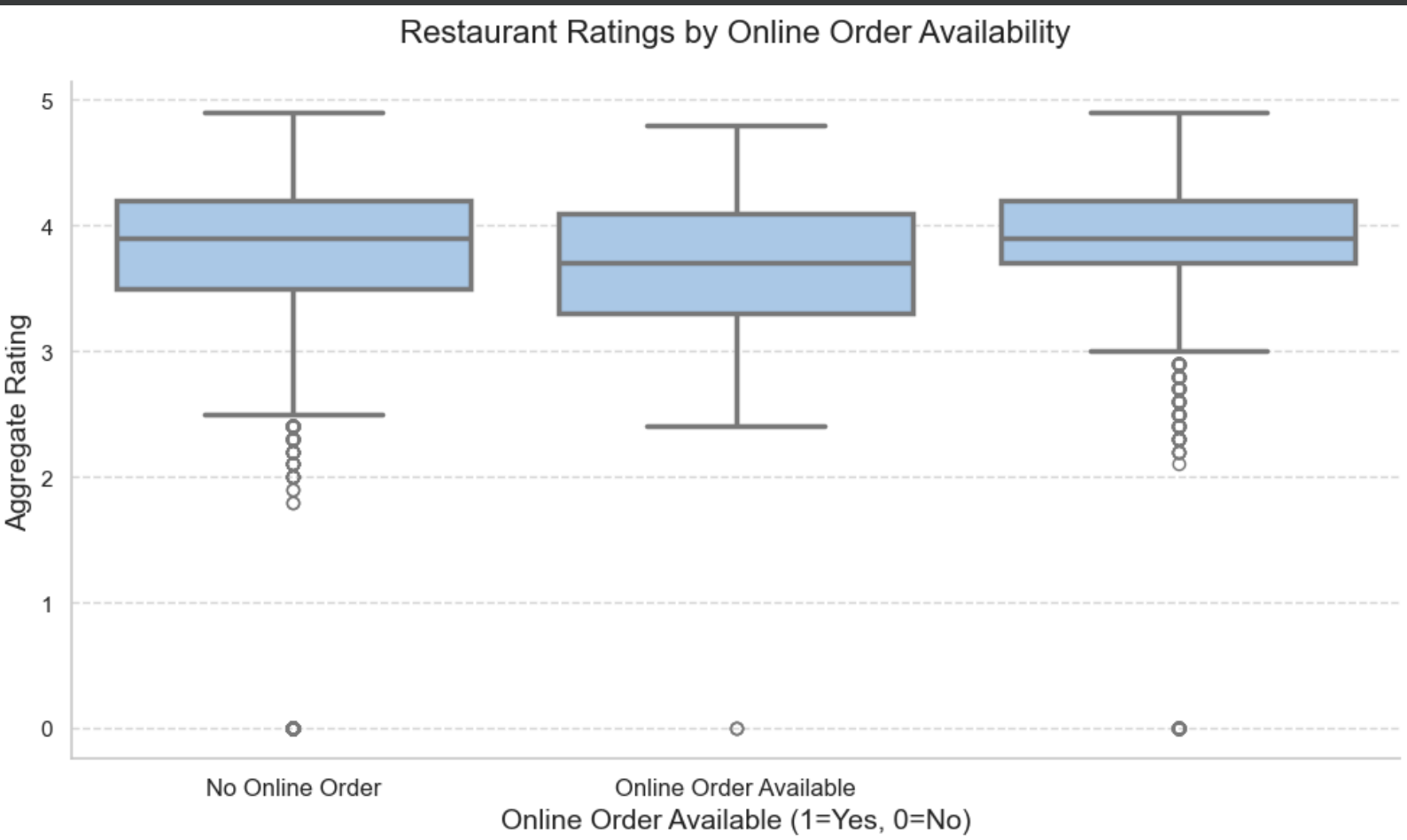
6. Implications
 – The findings could have practical implications for restaurant owners in terms of menu planning and customer satisfaction strategies.

Overall, the graph is a valuable tool for analyzing how the diversity of cuisines offered by restaurants might influence their overall rati





Restaurant Ratings by Online Order Availability



1. X-Axis (Horizontal Axis): Represents the availability of online ordering, categorized into two groups:
   - "No Online Order" (0)
   - "Online Order Available" (1)

2. Y-Axis (Vertical Axis): Represents the average rating of restaurants, likely on a scale from 0 to 5.

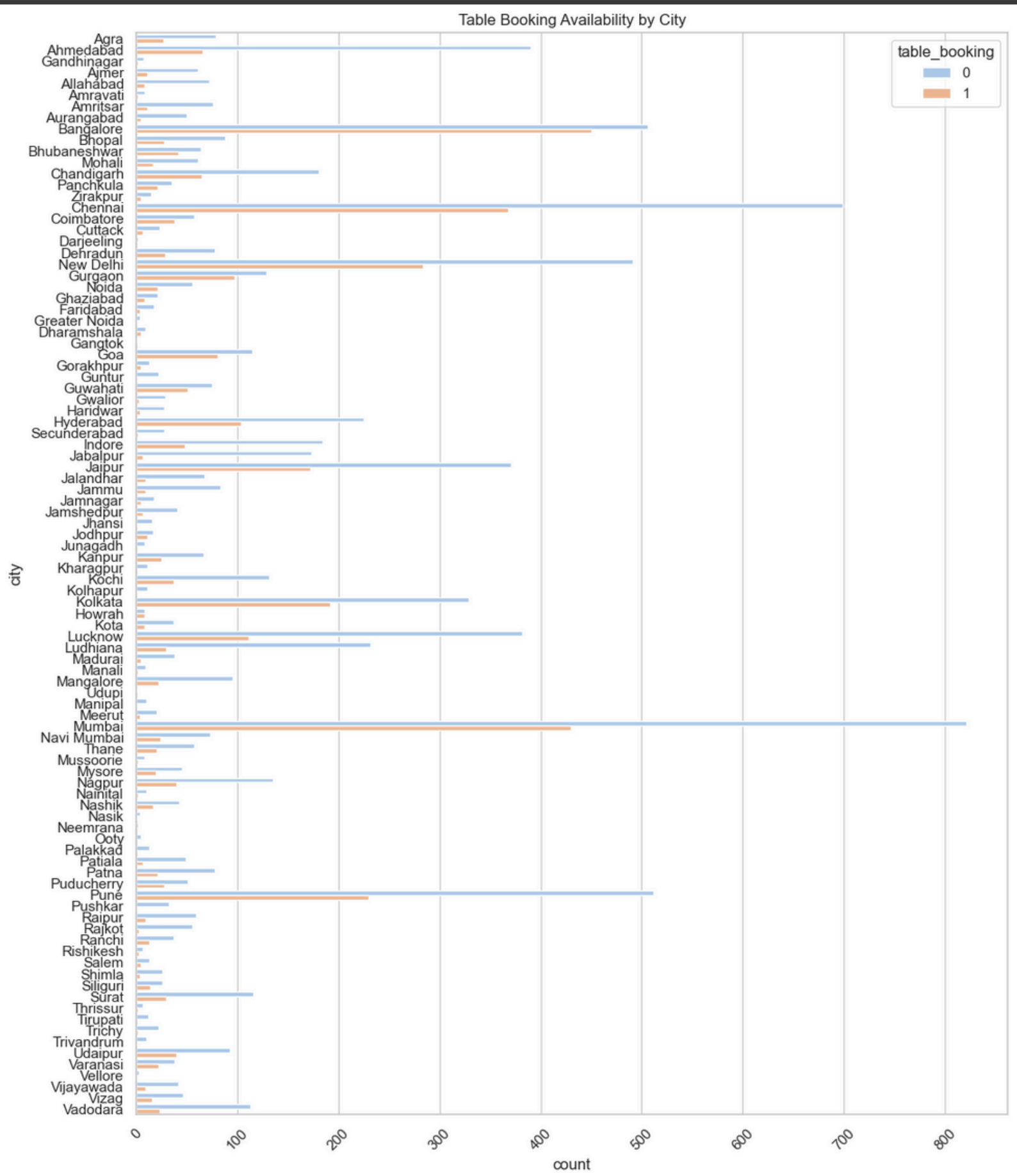
3. Data Points:
   - For restaurants without online ordering (0), the average rating is approximately 1.
   - For restaurants with online ordering available (1), the average rating is around 4.

4. Trend: The graph indicates that restaurants offering online ordering tend to have significantly higher average ratings compared to those that do not.

5. Implications: This suggests a positive correlation between the availability of online ordering services and customer satisfaction, as reflected in higher ratings.

6. Visual Representation: The graph likely uses bars or points to represent the average ratings for each category, making it easy to compare the two groups.

Table Booking Availability by City



The content provided appears to be a table rather than a graph, titled "Table Booking Availability by City." Here are the key points based on the table:

1. Table Structure: The table lists various cities and indicates whether table booking is available (1) or not (0).

2. Cities and Availability:
   - Agra, Ahmedabad, Gandhinagar: Table booking is not available (0).
   - All other listed cities:** Table booking is available (1).

3. Geographical Coverage: The table includes a wide range of cities from different regions, indicating a broad geographical scope.

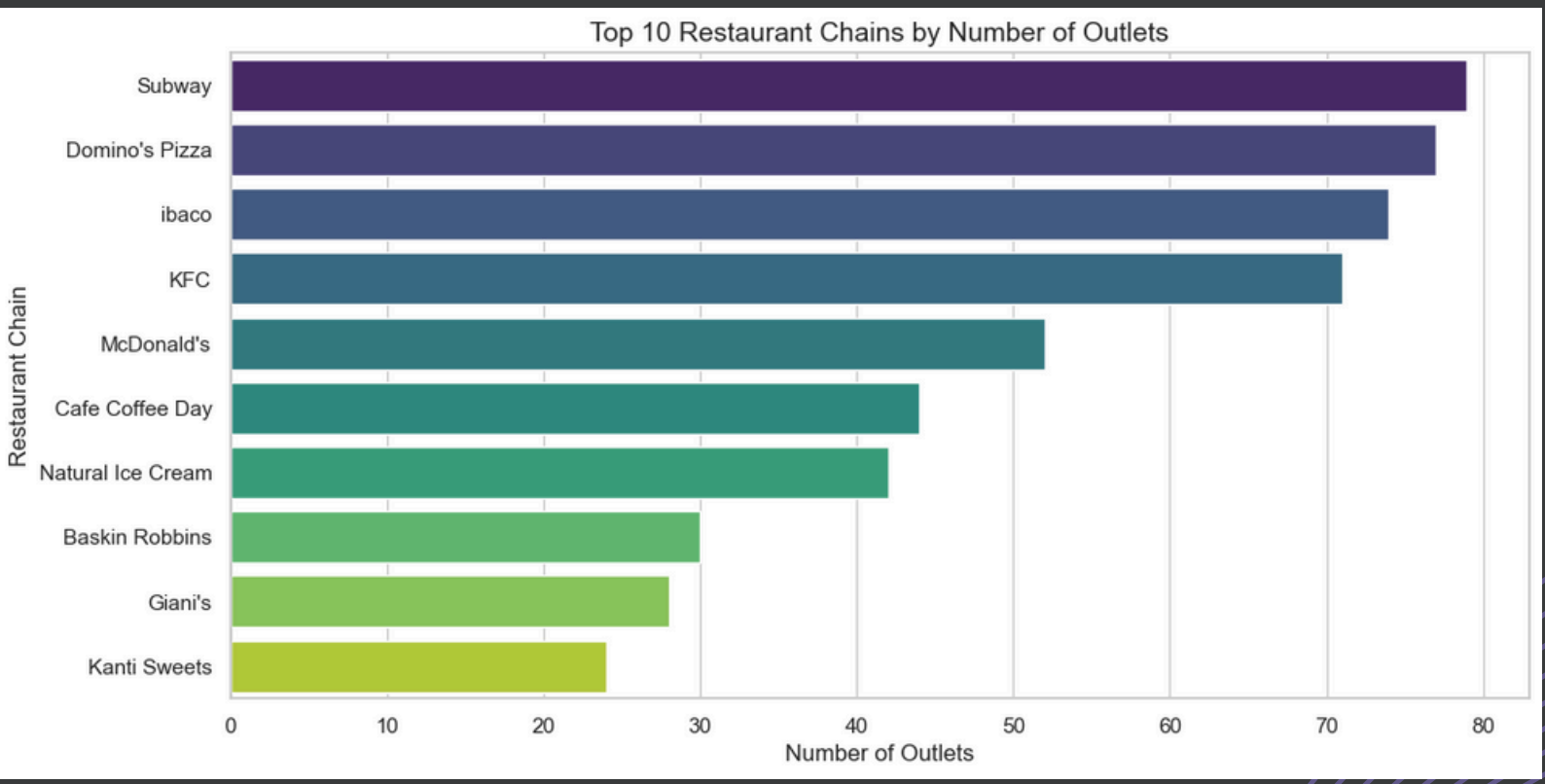
4. Implications: The availability of table booking services is widespread across most cities, with only a few exceptions.

5. Data Representation: The table uses a binary system (0 and 1) to represent the availability of table booking, making it straightforward to interpret.

6. Potential Use: This information could be useful for customers looking to book tables in specific cities or for businesses analyzing the prevalence of table booking services.

Overall, the table provides a clear and concise overview of table booking availability across various cities, highlighting that the service is widely available with only a few exceptions.

Top 10 Restaurant Chains by Number of Outlets



The graph titled "Top 10 Restaurant Chains by Number of Outlets" visually represents the number of outlets for the top 10 restaurant chains. Here are the key points:

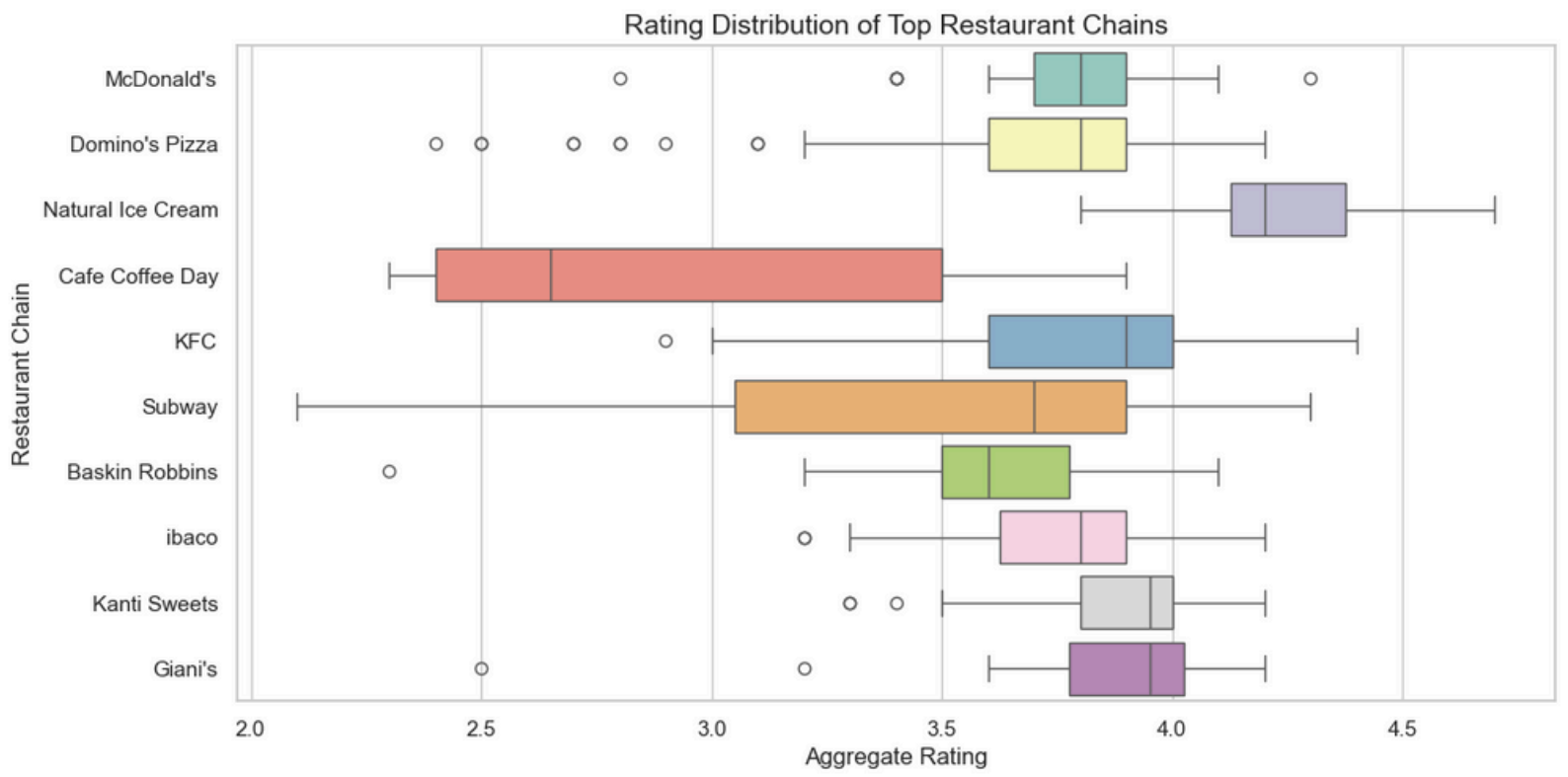
1. X-Axis (Horizontal Axis): Represents the "Number of Outlets," ranging from 0 to 80.

2. Y-Axis (Vertical Axis): Lists the "Restaurant Chains," which include:
   - Subway
   - Domino's Pizza
   - Ibaco
   - KFC
   - McDonald's
   - Cafe Coffee Day
   - Natural Ice Cream
   - Baskin Robbins
   - Giani's
   - Kanti Sweets

3. Data Representation: The graph likely uses horizontal bars to represent the number of outlets each restaurant chain has, making it easy to compare the chains.
4. Key Observations:
 – Subway appears to have the highest number of outlets, likely close to 80.
 – Domino's Pizza follows with a significant number of outlets.
 – The number of outlets decreases progressively for the subsequent chains, with Kanti Sweets having the fewest.
5. Implications: The graph highlights the market presence and expansion of these restaurant chains, with Subway and Domino's Pizza leading in terms of outlet numbers.
6. Visual Clarity: The use of a bar graph provides a clear and straightforward comparison of the number of outlets across different restaurant chains.

Overall, the graph effectively communicates the scale of each restaurant chain's operations, emphasizing the dominance of Subway and Domino's Pizza in terms of the number of outlets.

Rating Distribution of Top Restaurant Chains



The graph titled "Rating Distribution of Top Restaurant Chains" illustrates the aggregate ratings of various top restaurant chains. Here are the key points:

1. X-Axis (Horizontal Axis): Represents the "Aggregate Rating," with a scale ranging from 2.0 to 4.5.

2. Y-Axis (Vertical Axis): Lists the "Restaurant Chains," which include:
   - McDonald's
   - Domino's Pizza
   - Natural Ice Cream
   - Cafe Coffee Day
   - KFC
   - Subway
   - Baskin Robbins
   - Ibaco
   - Kanti Sweets
   - Giani's

3. Data Representation: The graph likely uses horizontal bars or points to represent the aggregate rating of each restaurant chain, allowing for easy comparison.

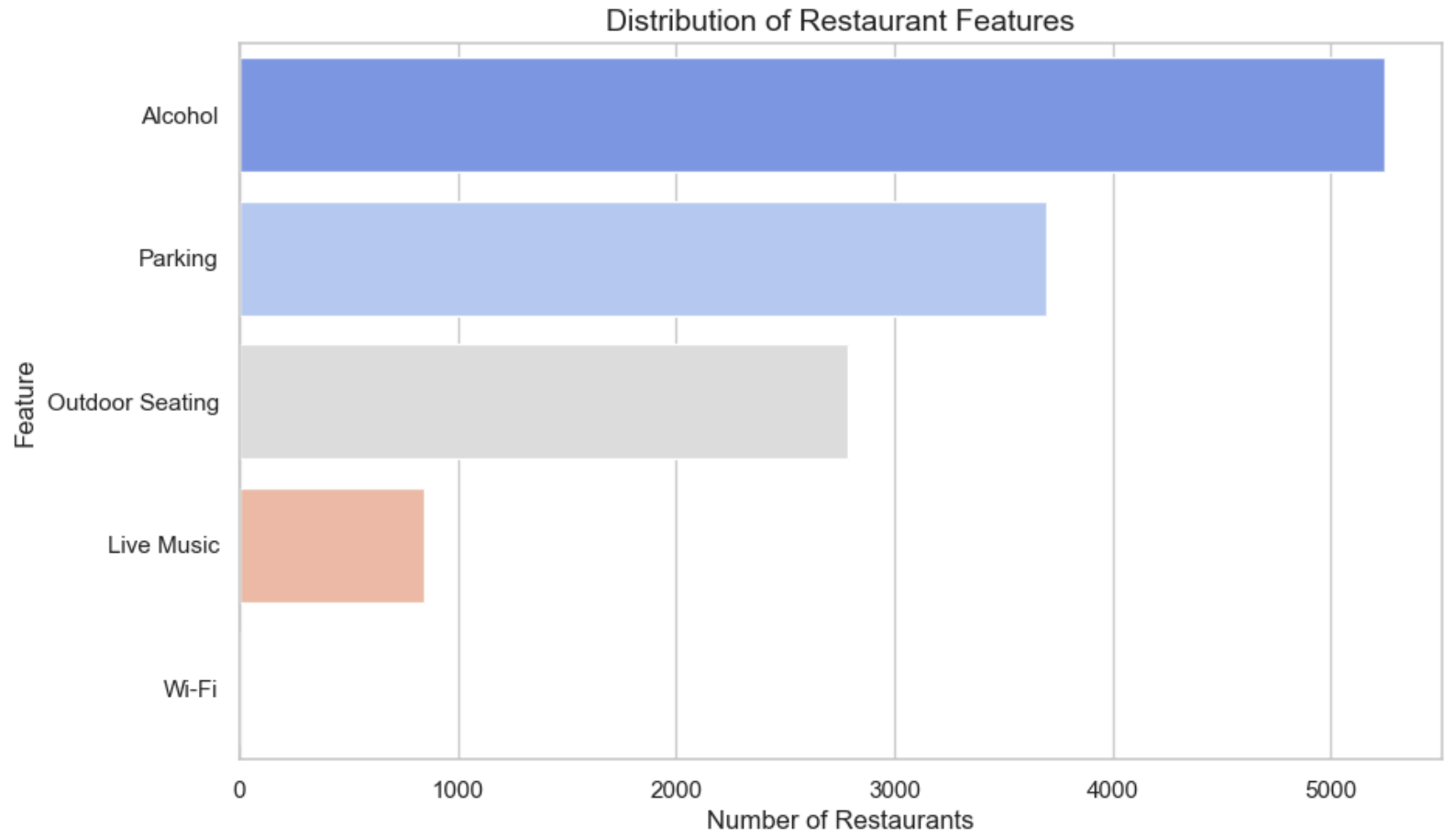
4. Key Observations:
 - The ratings are distributed between 2.0 and 4.5, indicating a range of customer satisfaction levels.
 - Natural Ice Cream and Cafe Coffee Day appear to have higher aggregate ratings, possibly close to 4.5.
 - McDonald's and Domino's Pizza also have relatively high ratings.
 - Kanti Sweets and Giani's seem to have lower ratings, closer to 2.0.

5. Implications: The graph highlights the varying levels of customer satisfaction across different restaurant chains, with some chains like Natural Ice Cream and Cafe Coffee Day performing exceptionally well.

6. Visual Clarity: The use of a bar or point graph provides a clear and straightforward comparison of the aggregate ratings, making it easy to identify which chains are rated higher or lower.

Overall, the graph effectively communicates the customer satisfaction levels of top restaurant chains, emphasizing the high performance of Natural Ice Cream and Cafe Coffee Day, and the relatively lower performance of Kanti Sweets and Giani's.

Distribution of Restaurant Features



The graph titled "Distribution of Restaurant Features" illustrates the prevalence of various features across a number of restaurants. Here are the key points:

X-Axis (Horizontal Axis): Represents the "Number of Restaurants," with a scale ranging from 0 to 5000.

Y-Axis (Vertical Axis): Lists the "Restaurant Features," which include:

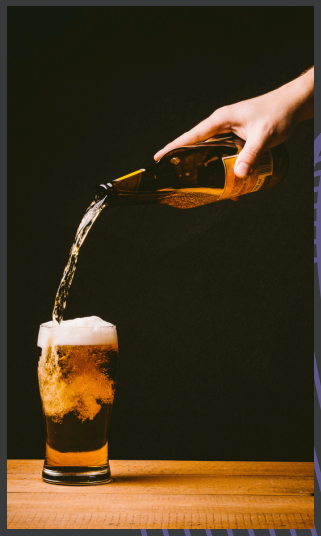
Alcohol

Parking

Outdoor Seating

Live Music

Wi-Fi



- Data Representation: The graph likely uses horizontal bars to represent the number of restaurants that offer each feature, making it easy to compare the prevalence of each feature.

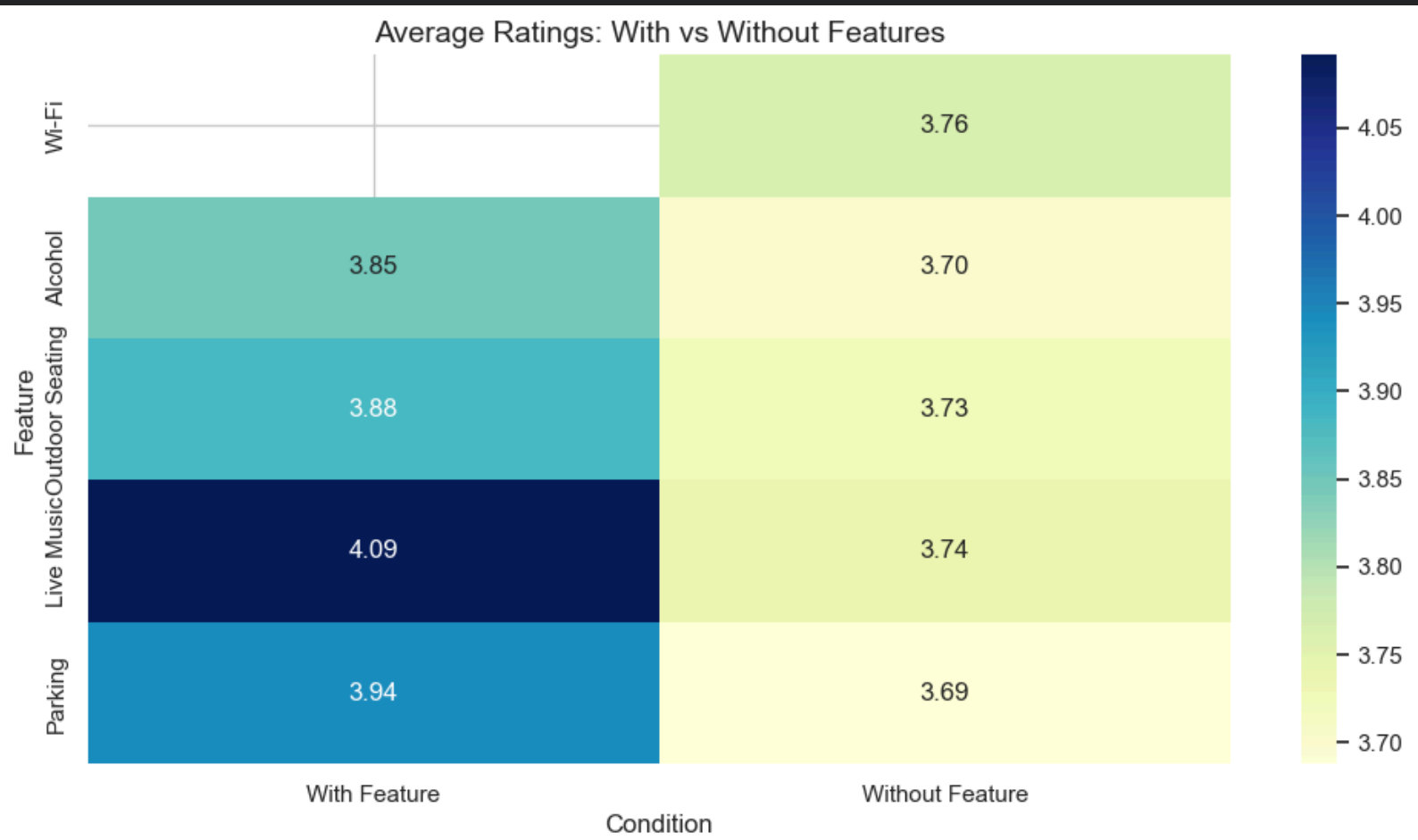
- Key Observations:
  - Parking appears to be the most common feature, with the highest number of restaurants offering it, possibly close to 5000.
  - Wi-Fi and Outdoor Seating also seem to be widely available, with a significant number of restaurants providing these features.
  - Alcohol and Live Music are less common, with fewer restaurants offering these features.

- Implications: The graph highlights the varying availability of different restaurant features, with parking being the most prevalent and live music being the least common.

- Visual Clarity: The use of a bar graph provides a clear and straightforward comparison of the number of restaurants offering each feature, making it easy to identify which features are more or less common.

- Overall, the graph effectively communicates the distribution of various restaurant features, emphasizing the widespread availability of parking and the relative rarity of live music and alcohol services.

Average Ratings: With vs Without Features



The graph titled "Average Ratings: With vs Without Features" compares the average ratings of restaurants based on whether they offer certain features or not. Here are the key points:

1. X-Axis (Horizontal Axis): Likely represents different conditions or features being compared, such as "With Feature" and "Without Feature."

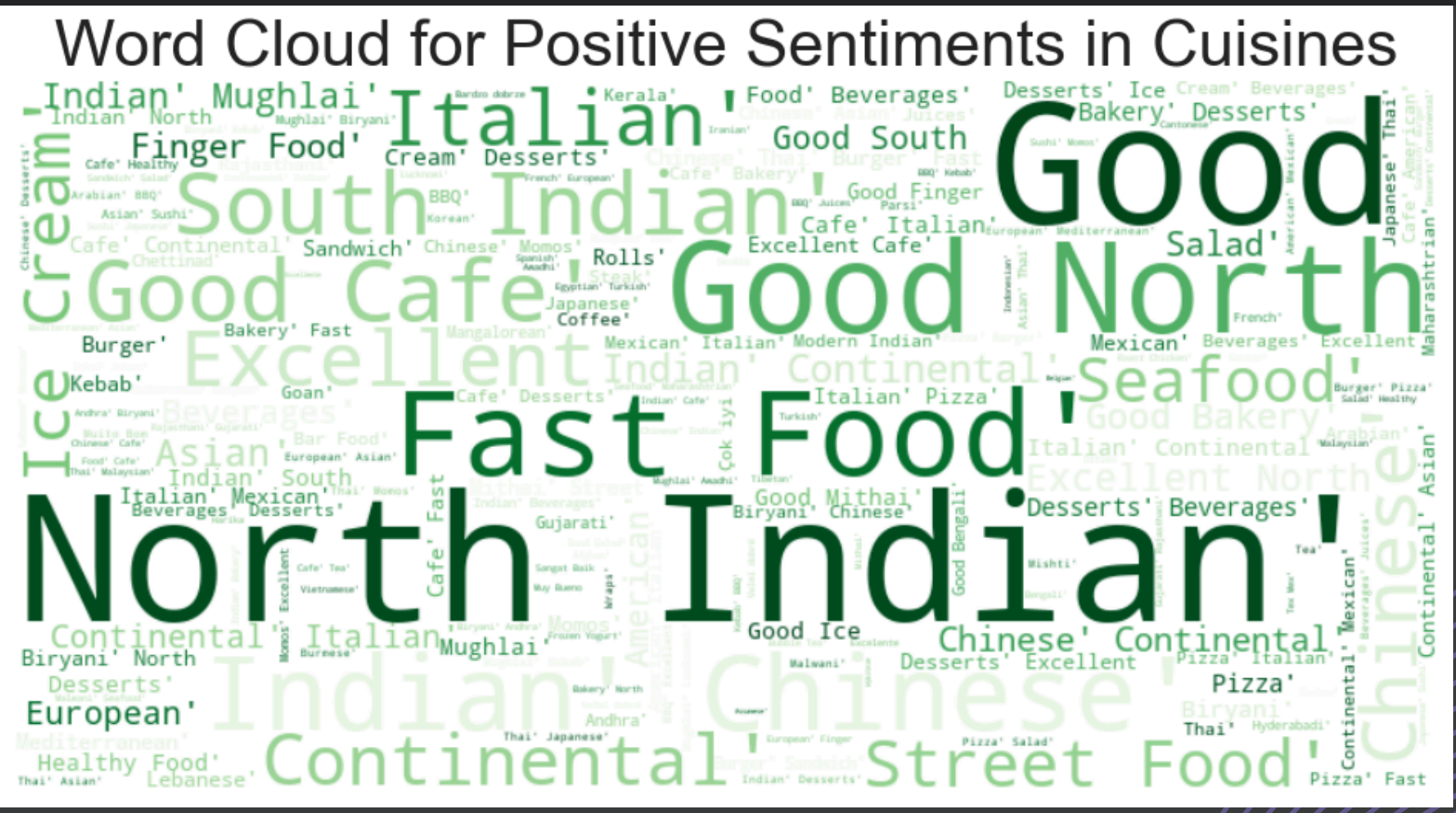
2. Y-Axis (Vertical Axis): Represents the "Average Rating," with values ranging approximately from 3.70 to 4.10.

3. Data Representation: The graph likely uses bars or points to represent the average ratings for restaurants with and without specific features.

4. Key Observations:
   - With Feature: The average ratings for restaurants with the feature are generally higher, with values around 3.85 to 4.09.
   - Without Feature: The average ratings for restaurants without the feature are lower, ranging from approximately 3.69 to 3.74.

5. Implications: The graph suggests that restaurants offering certain features tend to have higher average ratings compared to those that do not. This indicates that these features positively impact customer satisfaction.

Word Cloud for Positive Sentiments in Cuisines

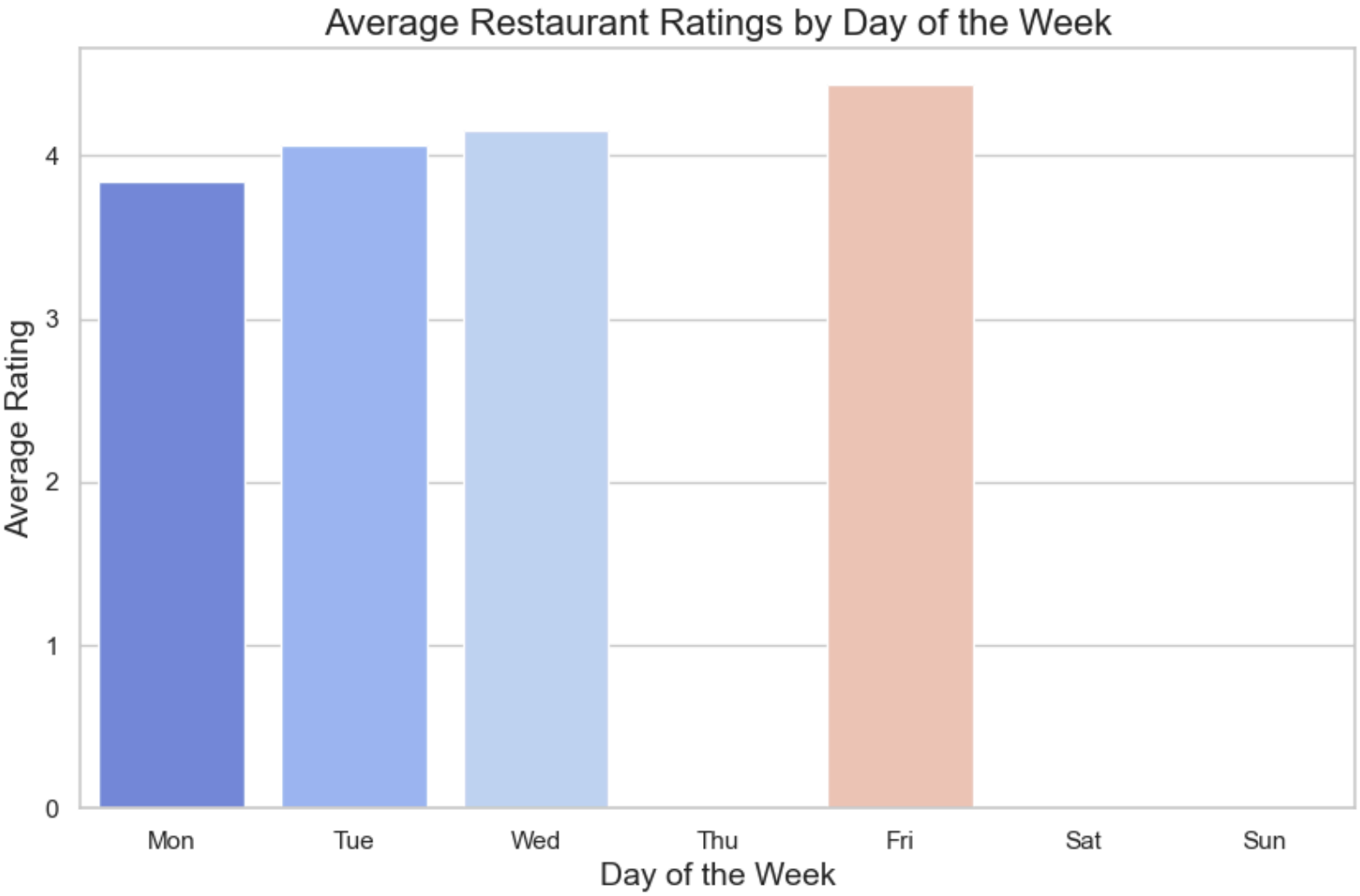


The image is a word cloud titled "Word Cloud for Positive Sentiments in Cuisines." It visually represents frequently mentioned words associated with positive sentiments in food reviews. Larger words indicate higher frequency, while smaller ones appear less often.

Prominent words include "North Indian," "Good," "Fast Food," "Italian," and "Cafe," suggesting these cuisines and food types receive the most positive feedback. Other notable words include "South Indian," "Chinese," "Continental," "Seafood," "Ice Cream," and "Bakery." The presence of words like "Excellent," "Healthy," "Finger Food," and "Desserts" highlights aspects that contribute to positive customer experiences.

The word cloud uses shades of green to display the data, making it visually appealing and easy to interpret. This visualization helps identify popular cuisines and trends based on customer sentiments, providing valuable insights into food preferences.

Average Restaurant Ratings by Day of the Week

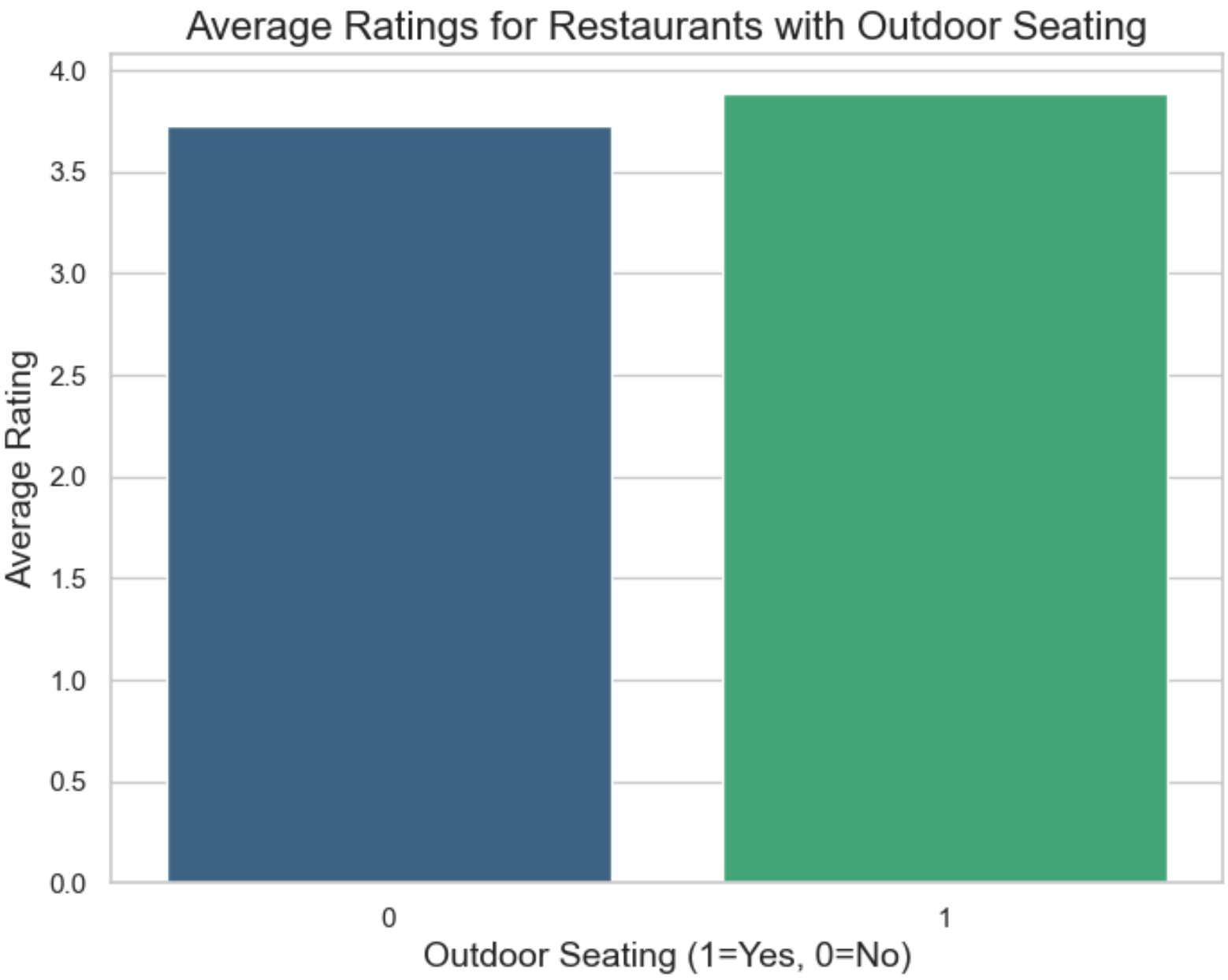


The bar chart titled "Average Restaurant Ratings by Day of the Week" displays the average ratings of restaurants across different days. The x-axis represents the days of the week, while the y-axis represents the average rating on a scale from 0 to 5.

The chart shows ratings from Monday to Friday, with no data for Saturday and Sunday. The ratings appear to increase gradually from Monday to Friday. Monday has the lowest rating, while Friday has the highest, suggesting that restaurants receive better ratings towards the end of the workweek.

The bars are color-coded, transitioning from darker blue on Monday to lighter shades on Wednesday and then to a peach color on Friday. This visualization helps in identifying patterns in customer satisfaction, showing that people might rate restaurants higher as the week progresses, possibly due to better service, mood, or different dining experiences.

Average Ratings for Restaurants with Outdoor Seating

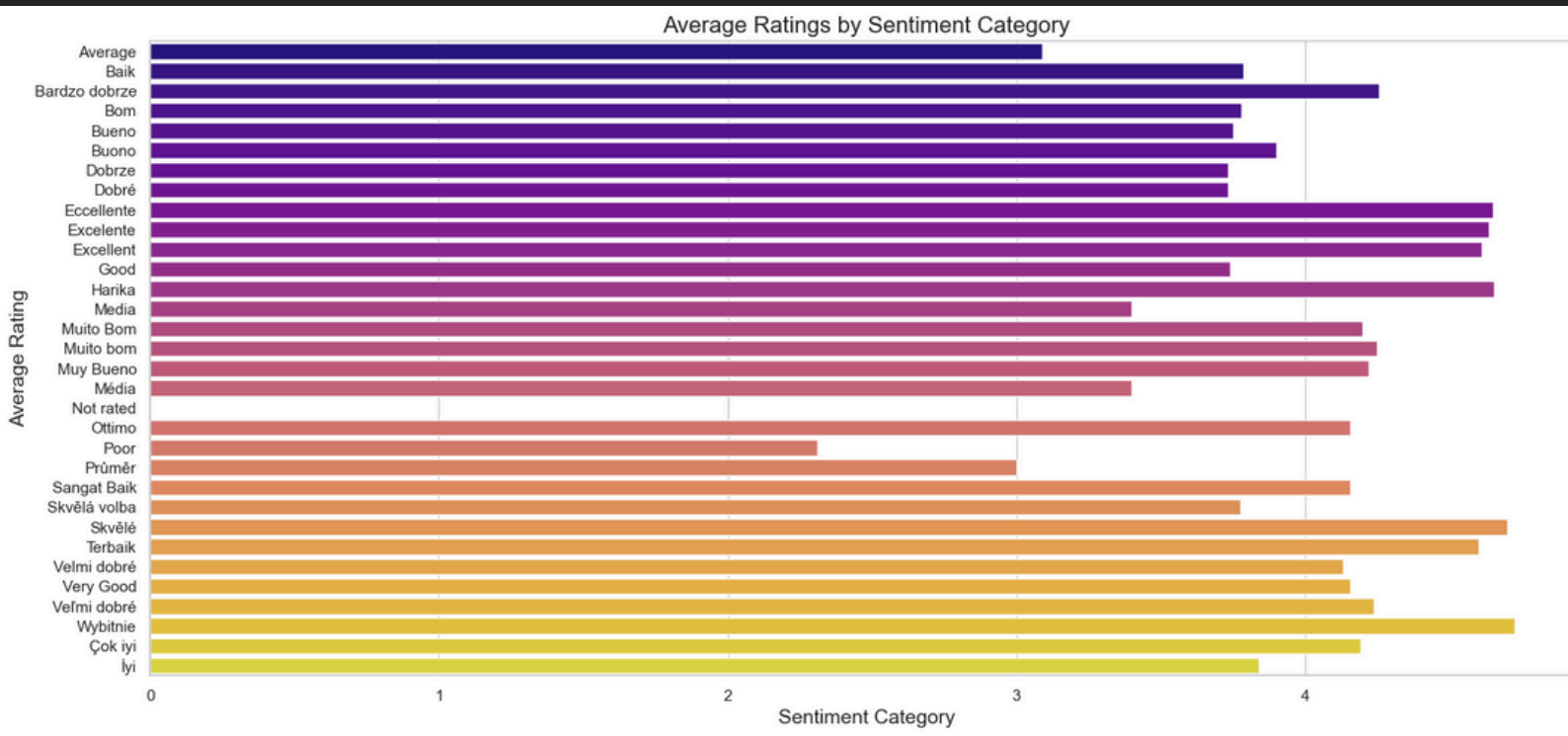


The bar chart titled "Average Ratings for Restaurants with Outdoor Seating" compares the average ratings of restaurants with and without outdoor seating. The x-axis represents outdoor seating availability, where 0 indicates no outdoor seating and 1 indicates the presence of outdoor seating. The y-axis represents the average rating on a scale from 0 to 5.

The chart shows that restaurants with outdoor seating (1) have a slightly higher average rating compared to those without it (0). This suggests that customers may prefer dining experiences that include outdoor seating, potentially due to ambiance, open-air environment, or better overall experience.

The bars are color-coded differently, with a dark blue bar for restaurants without outdoor seating and a green bar for those with it. This visualization provides insights into customer preferences and could be useful for restaurant owners considering adding outdoor seating to enhance customer satisfaction.

Average Ratings by Sentiment Category



The bar chart titled "Average Ratings by Sentiment Category" visualizes the relationship between sentiment categories and their corresponding average ratings. The x-axis represents the sentiment category, while the y-axis lists different sentiment labels in multiple languages (such as "Excellent," "Muy Bueno," "Velmi dobré," etc.).

The color gradient transitions from dark purple (lower ratings) to yellow (higher ratings), indicating how ratings vary across sentiments. Higher-rated sentiments, such as "Skvélé" and "Çok iyi," appear towards the bottom in yellow, while lower-rated sentiments like "Poor" and "Media" appear in darker shades towards the top.

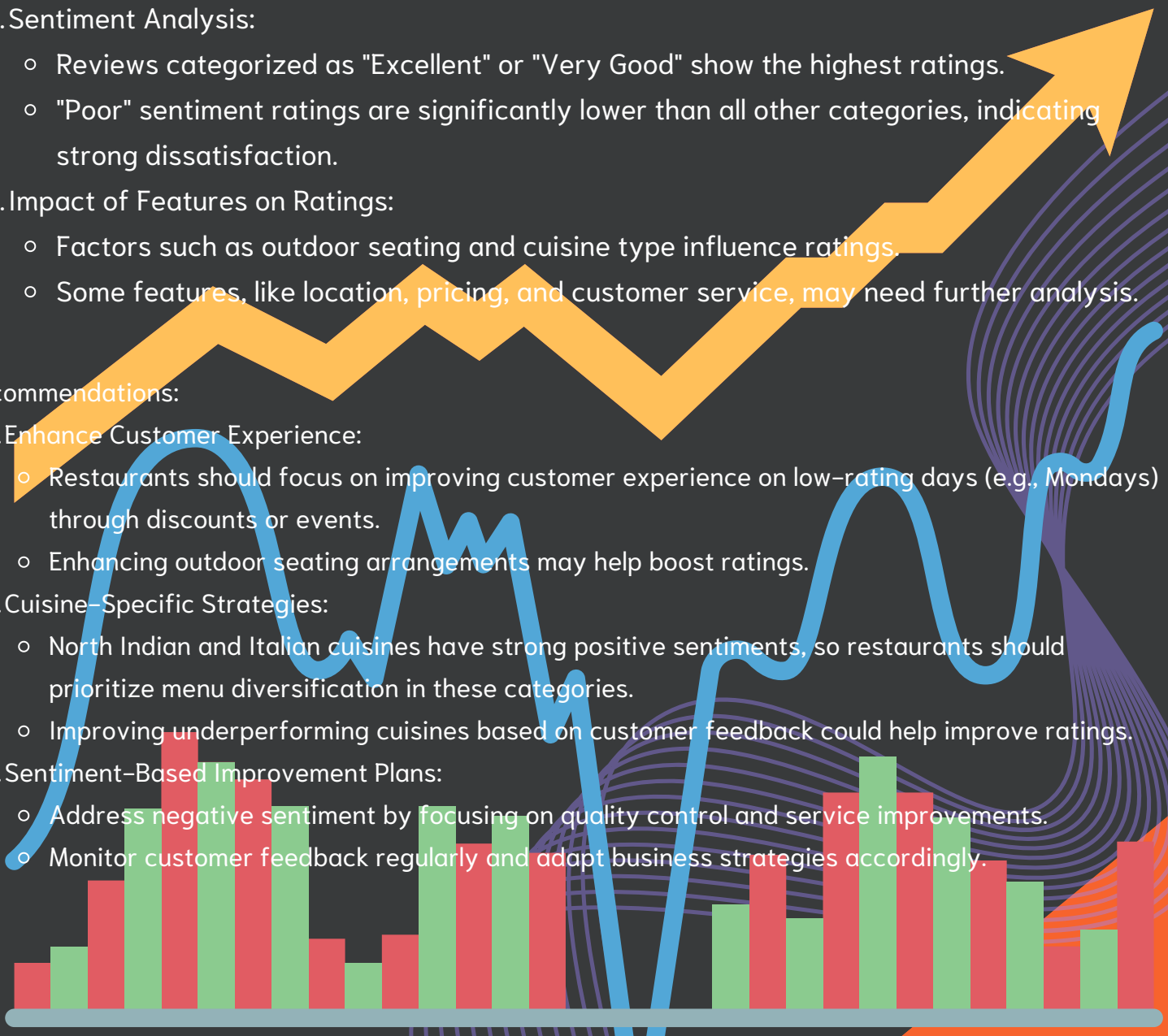
This graph highlights that positive sentiments tend to be associated with higher ratings, while neutral or negative sentiments receive lower ratings. The multilingual representation suggests an international dataset, making it useful for analyzing customer feedback trends across different languages and cultures.

Key Insights and Recommendation

1. Restaurant Ratings Trends:
   - The highest average ratings are observed on Fridays, while Mondays show relatively lower ratings.
   - Restaurants with outdoor seating tend to have slightly higher average ratings than those without.
2. Cuisine Sentiments:
   - North Indian, Italian, and Fast Food cuisines receive the most positive sentiments.
   - Words like "Good," "Excellent," and "Cafe" dominate the positive sentiment word cloud, indicating customer preferences.
3. Sentiment Analysis:
   - Reviews categorized as "Excellent" or "Very Good" show the highest ratings.
   - "Poor" sentiment ratings are significantly lower than all other categories, indicating strong dissatisfaction.
4. Impact of Features on Ratings:
   - Factors such as outdoor seating and cuisine type influence ratings.
   - Some features, like location, pricing, and customer service, may need further analysis.

Recommendations:

1. Enhance Customer Experience:
   - Restaurants should focus on improving customer experience on low-rating days (e.g., Mondays) through discounts or events.
   - Enhancing outdoor seating arrangements may help boost ratings.
2. Cuisine-Specific Strategies:
   - North Indian and Italian cuisines have strong positive sentiments, so restaurants should prioritize menu diversification in these categories.
   - Improving underperforming cuisines based on customer feedback could help improve ratings.
3. Sentiment-Based Improvement Plans:
   - Address negative sentiment by focusing on quality control and service improvements.
   - Monitor customer feedback regularly and adapt business strategies accordingly.



Conclusion and Final Thoughts

Conclusion: The analysis reveals that customer sentiment and ratings are influenced by multiple factors, including cuisine type, seating arrangements, and day of the week. Restaurants can leverage these insights to enhance customer satisfaction and optimize business strategies. Positive sentiments align with specific cuisines, and negative reviews highlight areas that need improvement. Implementing targeted strategies can help businesses maximize ratings and customer retention.

Final Thoughts: Data-driven decision-making is crucial in the restaurant industry. Understanding customer sentiment helps restaurants refine their offerings and improve the overall dining experience. Sentiment analysis and rating trends provide actionable insights for business growth.

Future Work:

- Expanding the analysis to include pricing, service speed, and location-based variations.
- Utilizing machine learning models to predict customer ratings based on reviews.
- Conducting a deeper analysis of customer demographics to personalize restaurant services.
- Enhancing sentiment analysis with Natural Language Processing (NLP) for more accurate customer insights.

THANKYOU