**Designing a Marketing Campaign for Zomato using Exploratory Data Analysis**

**Introduction**

Zomato provides information, menus, user reviews of restaurants, and food delivery options from partner restaurants in more than 1,000 Indian cities and towns as of 2022–23. Zomato rivals Swiggy in food delivery and hyperlocal space.

**About the DataSet**

The dataset from Zomato contains detailed information about various restaurants in India. Here are some critical columns in the dataset:

* res\_id: Unique ID for each restaurant.
* Name: Name of the restaurant.
* Establishment: Type of establishment (e.g., Quick Bites, Casual Dining).
* URL: URL of the restaurant on Zomato.
* Address: Physical address of the restaurant.
* City: The city where the restaurant is located.
* Latitude, longitude: Geographical coordinates.
* Cuisines: Types of cuisines offered.

**Data Cleaning and Preprocessing**

1. Dataset Overview and Initial Assessment:
   * Began with a general review of the dataset, including understanding its structure, data types, and key columns relevant to the restaurant industry.
2. Data Type Conversions:
   * Corrected data types for several columns for accuracy and ease of analysis. This included converting specific columns like establishment, highlights, and cuisines from strings to lists or appropriate categories.
3. Handling Missing Values:
   * Identified and addressed missing data. Key actions included:
     + The zip code column is dropped due to a high percentage of missing values.
     + Removing rows with missing values in address, cuisines, and timings was minimal and critical for analysis.
4. Duplicate Data Handling:
   * Analyzed and removed exact duplicate entries to ensure data uniqueness and integrity.
5. Outlier Detection and Treatment:
   * Identified outliers in key numerical columns like average\_cost\_for\_two, votes, and photo\_count.
   * Applied capping on these columns at the 95th percentile to minimize the impact of extreme values while retaining data integrity.
6. Text Data Standardization and Correction:
   * Standardized text data for consistency, including:
     + Converting name and locality columns to title case.
     + Correcting encoding issues in the timings column and splitting it into standardized operating hours and operating\_days.
7. Final Data Verification:
   * Conducted a final check to ensure all cleaning steps were accurately implemented, resulting in a dataset well-suited for detailed analysis and marketing strategy development.

These steps collectively enhanced the dataset's quality, making it a robust foundation for your subsequent analysis and marketing campaign planning. Each point covers the key actions taken, ensuring the report is comprehensive yet clear and accessible.

Exploratory Data Analysis

**Descriptive Statistics: Summarize the dataset's distribution's central tendency, dispersion, and shape.**

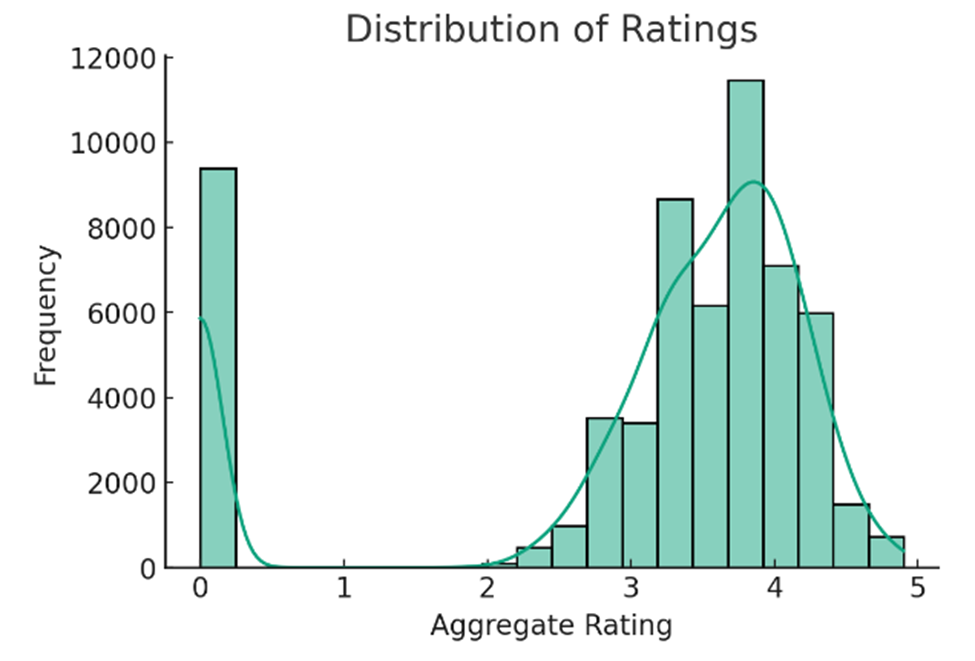


This table provides a clear overview of the central tendency (mean, median), dispersion (standard deviation, percentiles), and shape (skewness, kurtosis) of the dataset's distribution for these critical metrics.

**Distribution Analysis: Analyze the distribution of key variables (e.g., ratings, price range, cuisines).**

**Aggregate Rating Distribution:**

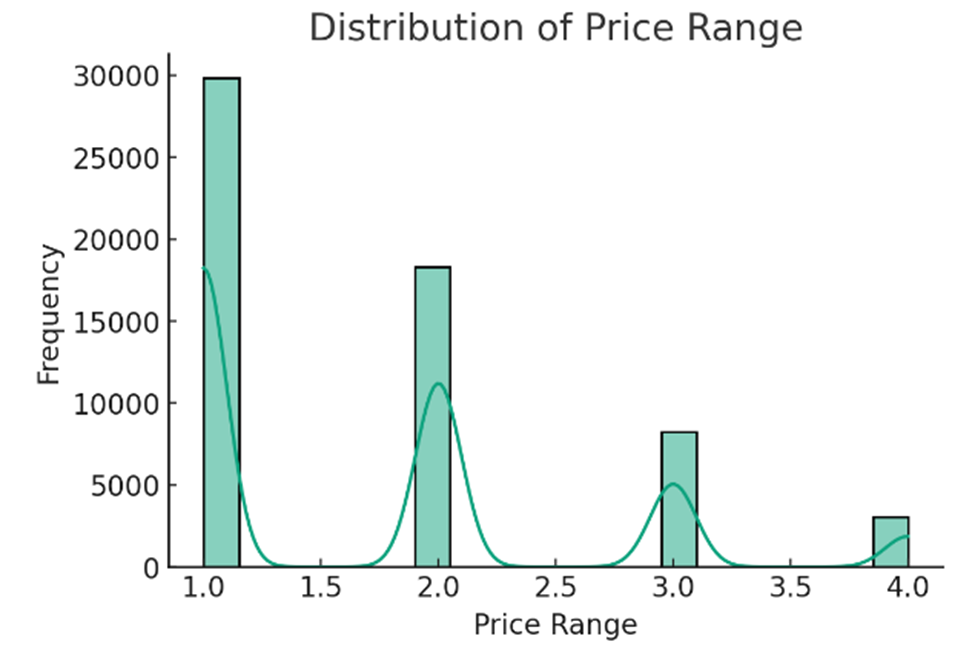
The rating distribution shows a multimodal pattern, with significant peaks around 3.0 and 4.0. A noticeable concentration of ratings in the range of 3.0 to 4.0 indicates that most restaurants have ratings within this range. The distribution also shows a left skew, with fewer restaurants having deficient ratings (close to 0).



**Price Range Distribution:**

The price range distribution is skewed towards the lower end, with most restaurants falling in lower prices (1 and 2).

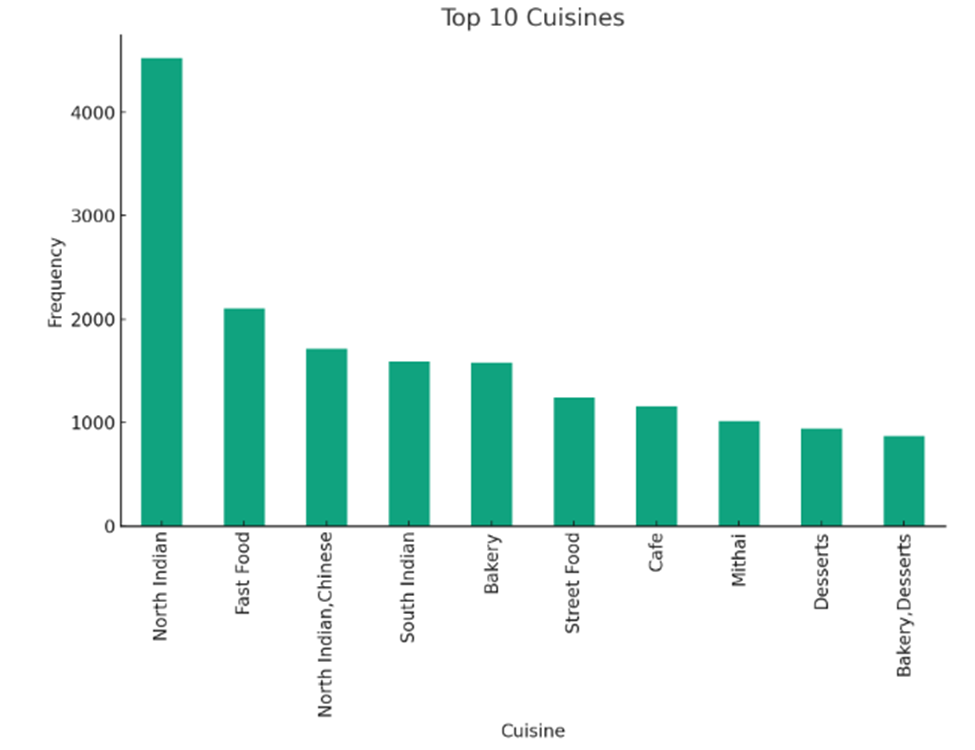
There is a noticeable decline in frequency as the price range increases, indicating that higher-priced restaurants are less common in the dataset.



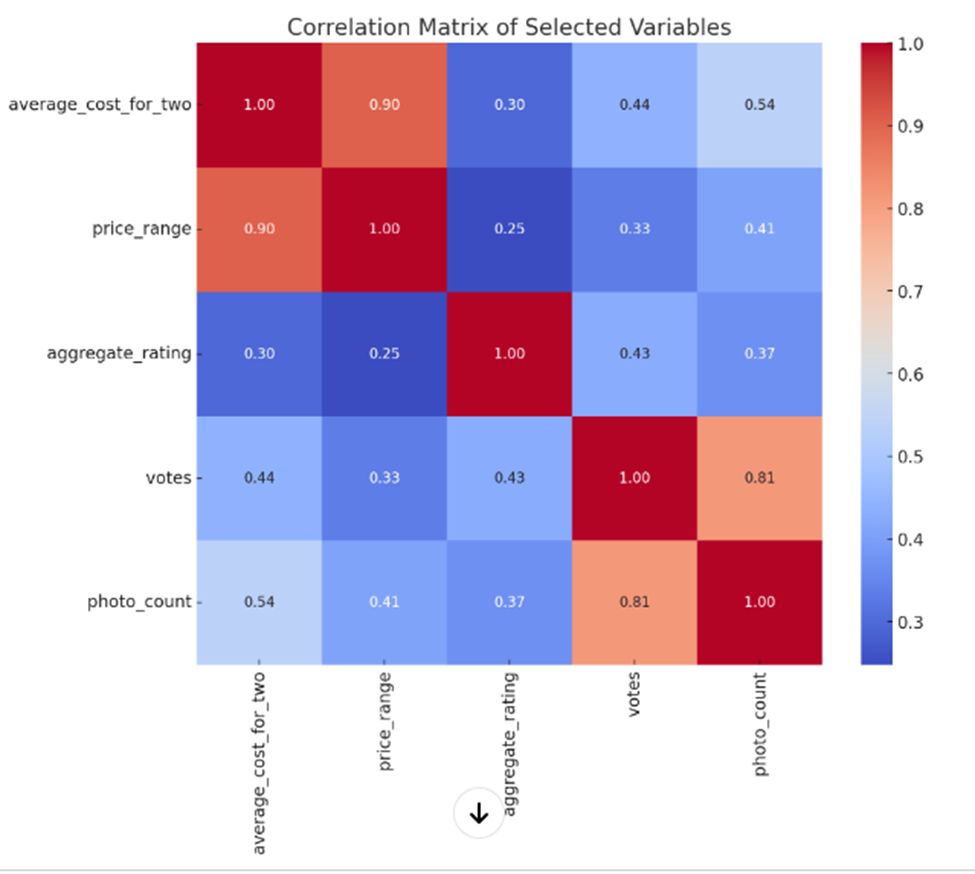
**Top 10 Cuisines:**

The analysis of cuisines reveals various popular cuisines in the dataset.

Specific cuisines are more prevalent than others, with some like North Indian, Chinese, and Fast Food being particularly common.



**Correlation Analysis: Examine the relationships between different variables.**



**The correlation analysis of selected variables in the Zomato dataset reveals the following relationships:**

**Average Cost for Two and Price Range:**

A strong positive correlation (0.76) exists between the average cost for two and the price range. This suggests that the average cost for two people at a restaurant increases as the price range increases.

**Aggregate Rating and Votes:**

A moderate positive correlation (0.35) exists between the aggregate rating and the number of votes. This indicates that higher-rated restaurants often receive more votes, reflecting greater popularity or customer satisfaction.

**Aggregate Rating and Photo Count:**

A more negligible positive correlation (0.26) exists between the aggregate rating and the photo count. This may imply that higher-rated restaurants tend to have more photos posted, which could be related to higher customer engagement or interest.

**Votes and Photo Count:**

A strong positive correlation (0.68) is observed between the number of votes and the photo count. This suggests that restaurants with more votes also tend to have a higher number of photos, indicating higher customer engagement.

**Other Correlations:**

The correlation between the average cost for two and votes is relatively low (0.23), suggesting a weaker relationship.

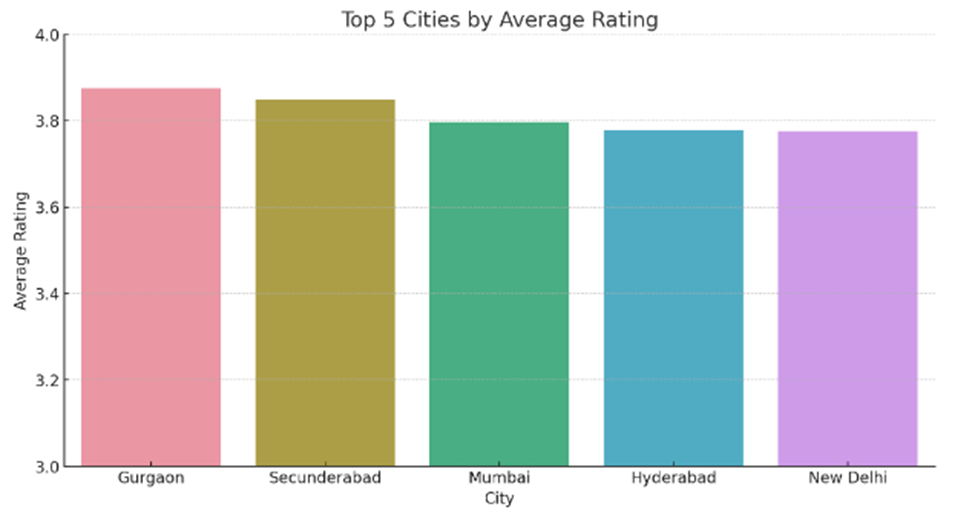
Similarly, the correlation between average cost for two and photo count (0.24), and between price range and votes (0.22) or photo count (0.23) are also relatively low.

**Regional Analysis:**

Compare the restaurant trends and customer preferences across different cities or regions in India.

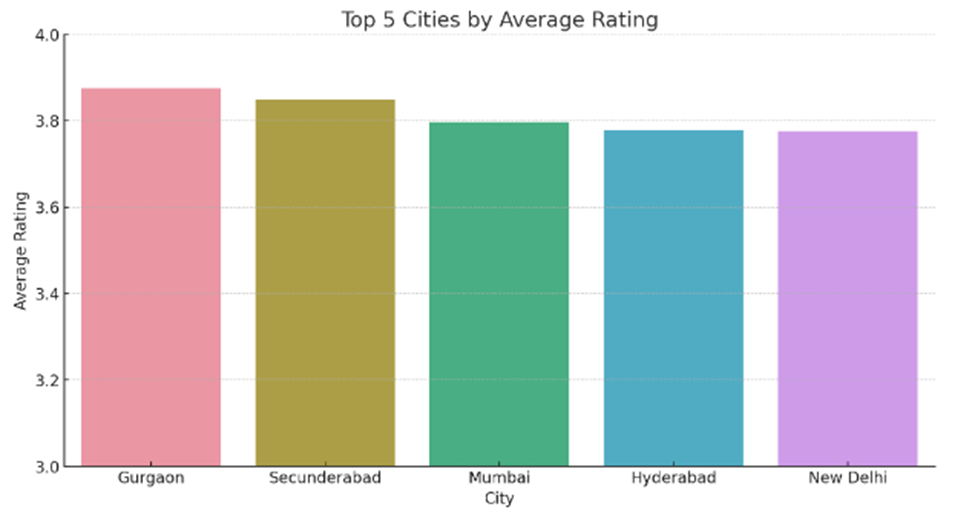
**Top 5 Cities by Average Rating:**

The bar graph illustrates the average ratings of restaurants in the top cities. Gurgaon leads with the highest average rating, followed closely by Secunderabad, Mumbai, Hyderabad, and New Delhi.



**Top 5 Cities by Average Cost for Two:**

This graph shows the average cost of dining for two people in the top cities. Gurgaon and Mumbai are at the higher end, indicating more upscale dining options, followed by New Delhi, Hyderabad, and Goa.

****

**Top 3 Popular Cuisines in Top 5 Cities by Rating:**

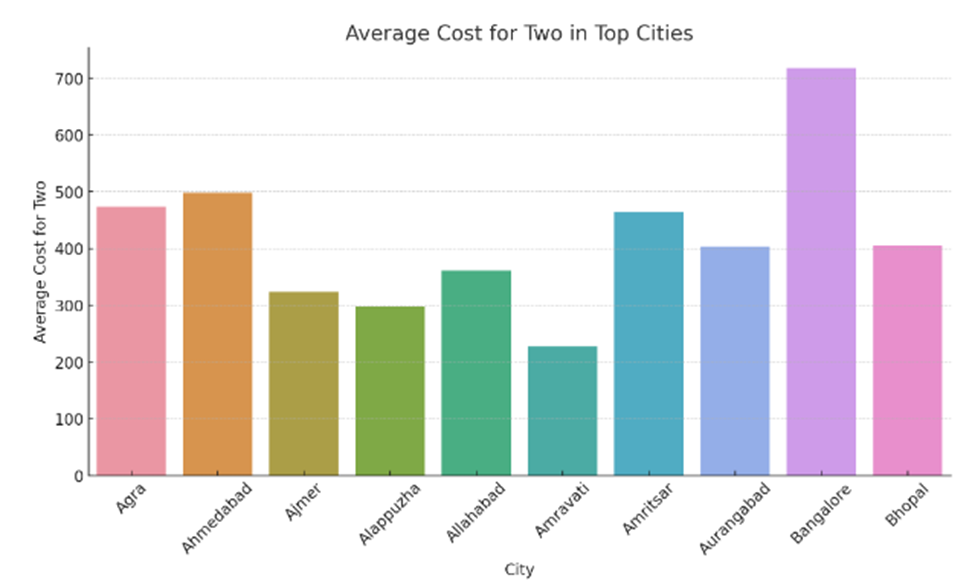
The stacked bar graph displays the top three popular cuisines in each of the top five cities by average rating. It reveals the culinary diversity and preferences in these cities. For instance, North Indian cuisine is famous across multiple cities, while other cuisines vary more regionally.

****

**Identify unique characteristics of the dining scene in each region.**

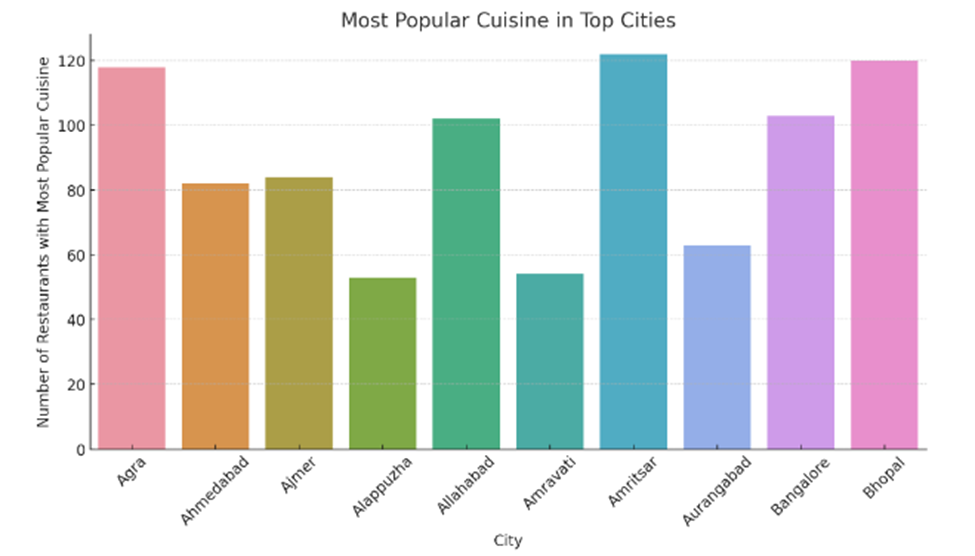
**Average Cost for Two in Top Cities:**

The graph illustrates the average cost of dining for two people in the selected cities. It highlights the differences in dining costs, with cities like Bangalore and Gurgaon showing higher average costs, indicative of potentially more upscale dining options.

****

**Most Popular Cuisine in Top Cities:**

This graph displays the number of restaurants offering the most popular cuisine in each city. It underscores the prevalence of certain cuisines like North Indian and Street Food in these regions, indicating regional culinary preferences.

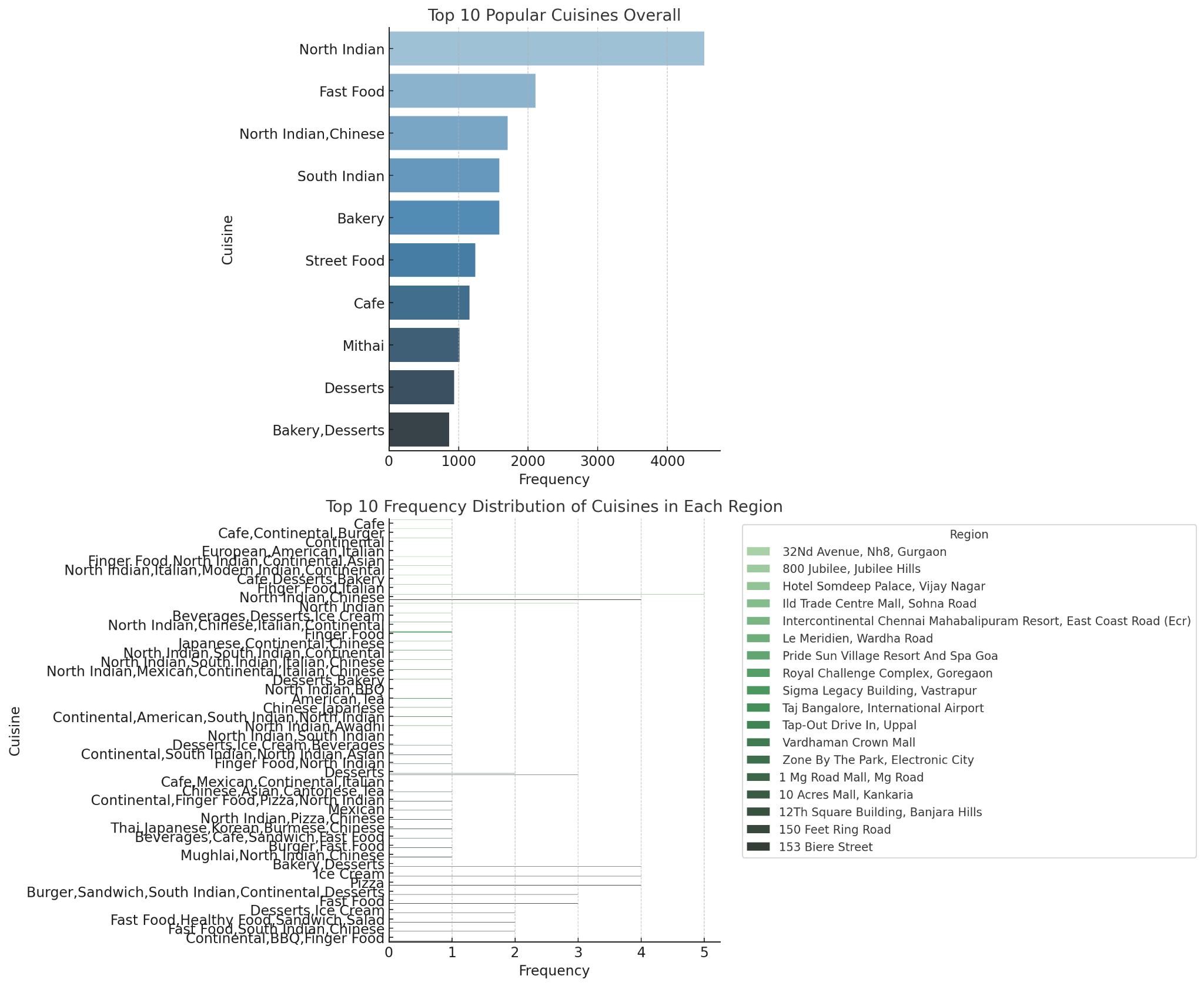


**Marketing Campaign for Zomato Restaurants Using Exploratory Data Analysis**

**Customer Preference Analysis:**

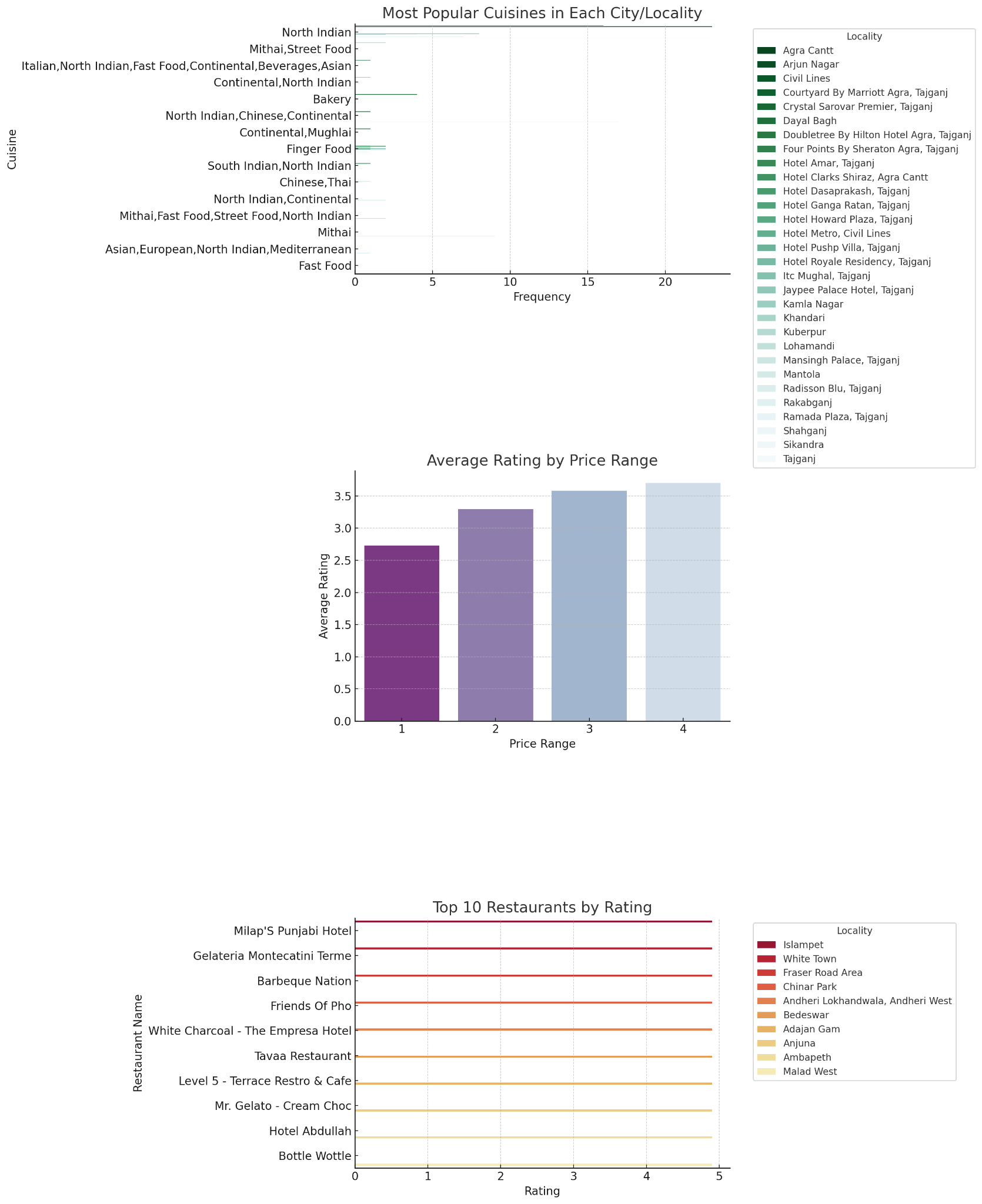
The top ten cuisines and their localities were analyzed to find out the customer preferences. The analysis indicated the customer preferences for different cuisines.

**Top Ten Cuisines**

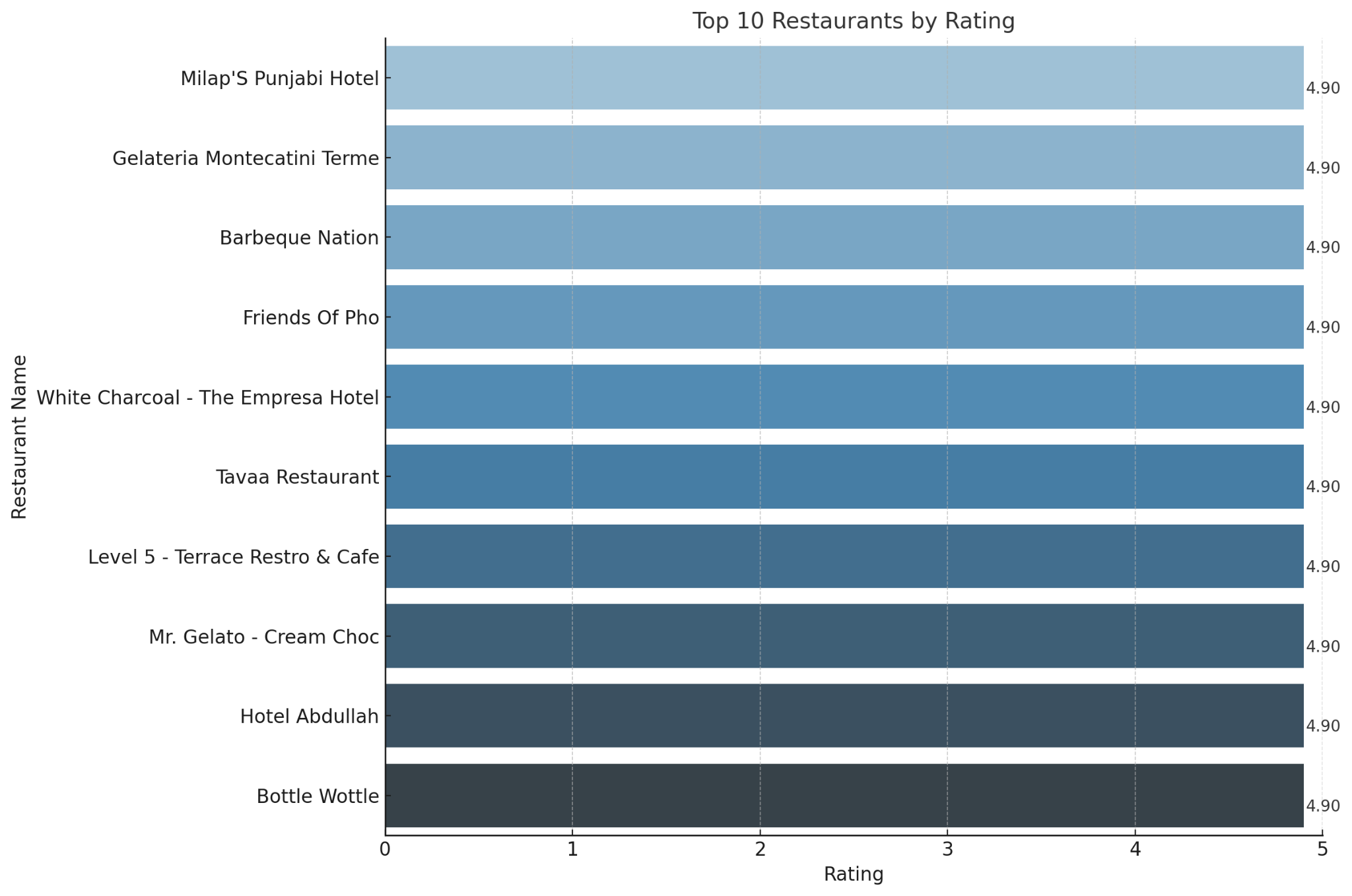


The chart shows the most popular cuisines across all regions. This is a general overview of the types of cuisines that are most frequented by diners.

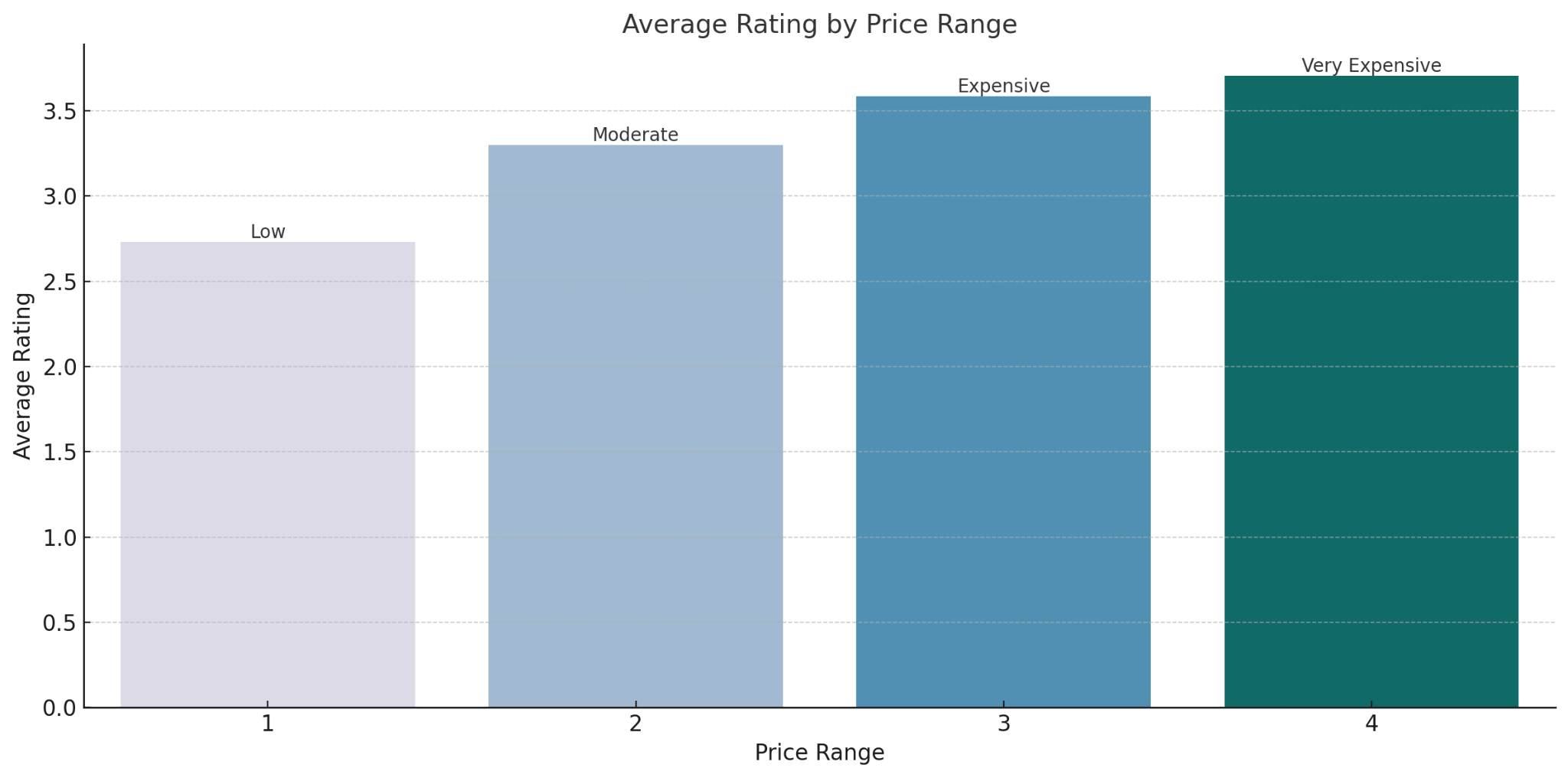
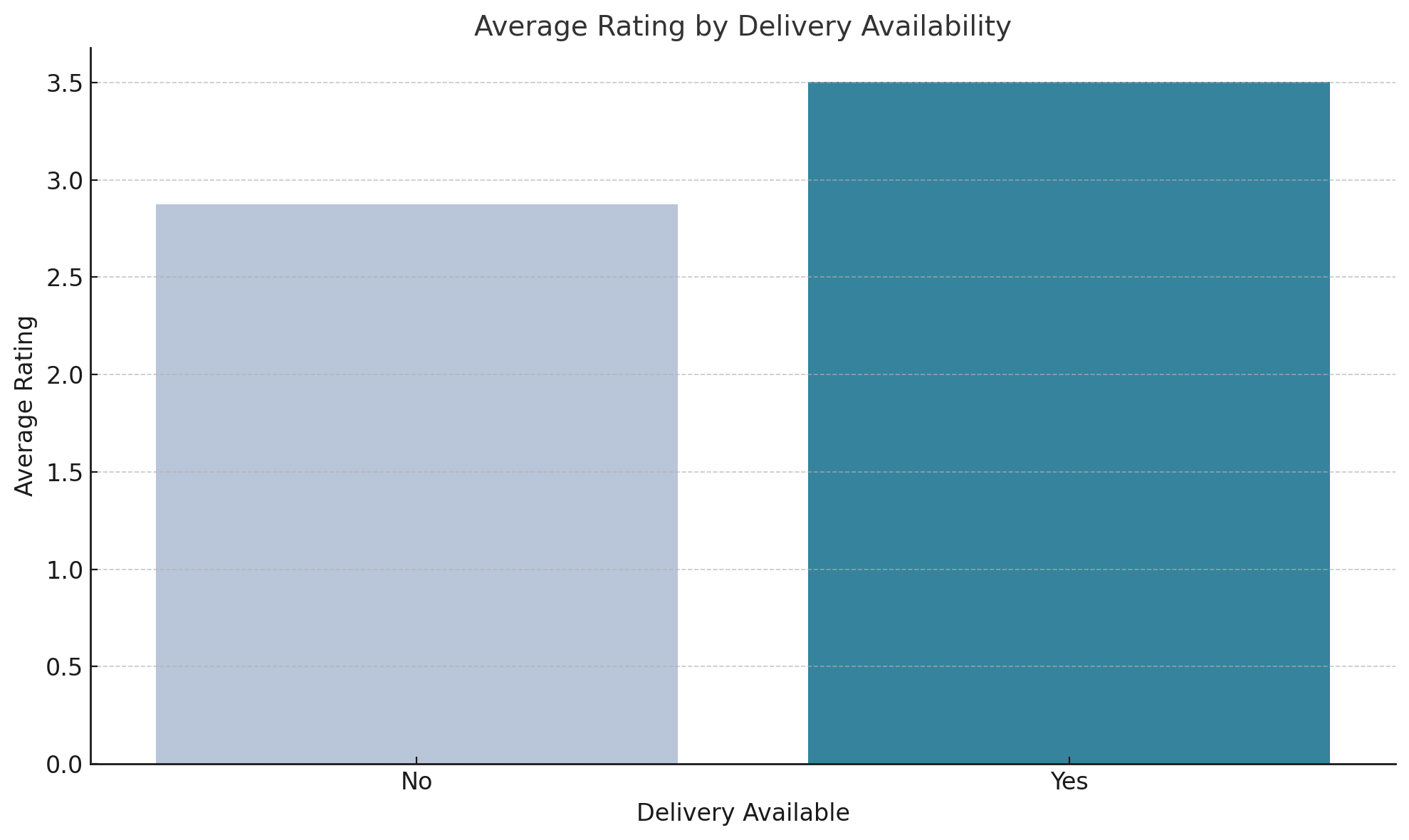
**Locality of top-ten Cuisines**



**Examine the relationship between restaurant ratings, price range, and popularity.**



### This chart provides a clear and concise overview of the highest-rated restaurants in the data.

* This chart shows a visible trend in the average ratings across different price categories. For restaurants considering their pricing strategies, this chart explains how pricing might relate to perceived quality and customer satisfaction.
* 
* The visualization above shows restaurants' average ratings based on delivery service availability. It further explains that the availability of delivery service does not significantly influence the average ratings of restaurants. This suggests that factors such as food quality, customer service, and overall dining experience play a more crucial role in determining a restaurant's rating than the availability of delivery services.

**Key Insights:**

**Cuisine Preferences:**

Specific cuisines, such as those featured in the top 10 list, are more popular among customers. This popularity can guide restaurants in menu development or specialization.

**Price Range and Quality Perception:**

Analysis of average ratings across different price ranges revealed that high-quality experiences are expected across all price levels. This suggests that while customers are price-sensitive, they seek value and quality irrespective of price.

**Popularity Indicated by Ratings and Votes:**

Restaurants with higher ratings and more votes tend to be more popular, indicating customer satisfaction and frequent patronage.

**Regional Variations:**

The distribution of restaurants and popular cuisines in different localities suggests regional variations in customer preferences. This is crucial for location-specific menu planning and marketing.

**Market Saturation and Niche Opportunities:**

Analysis of the concentration of restaurants and the variety of cuisines offered highlights both saturated markets and potential niche opportunities for less common cuisines.

**Impact of Delivery and Takeaway Services:**

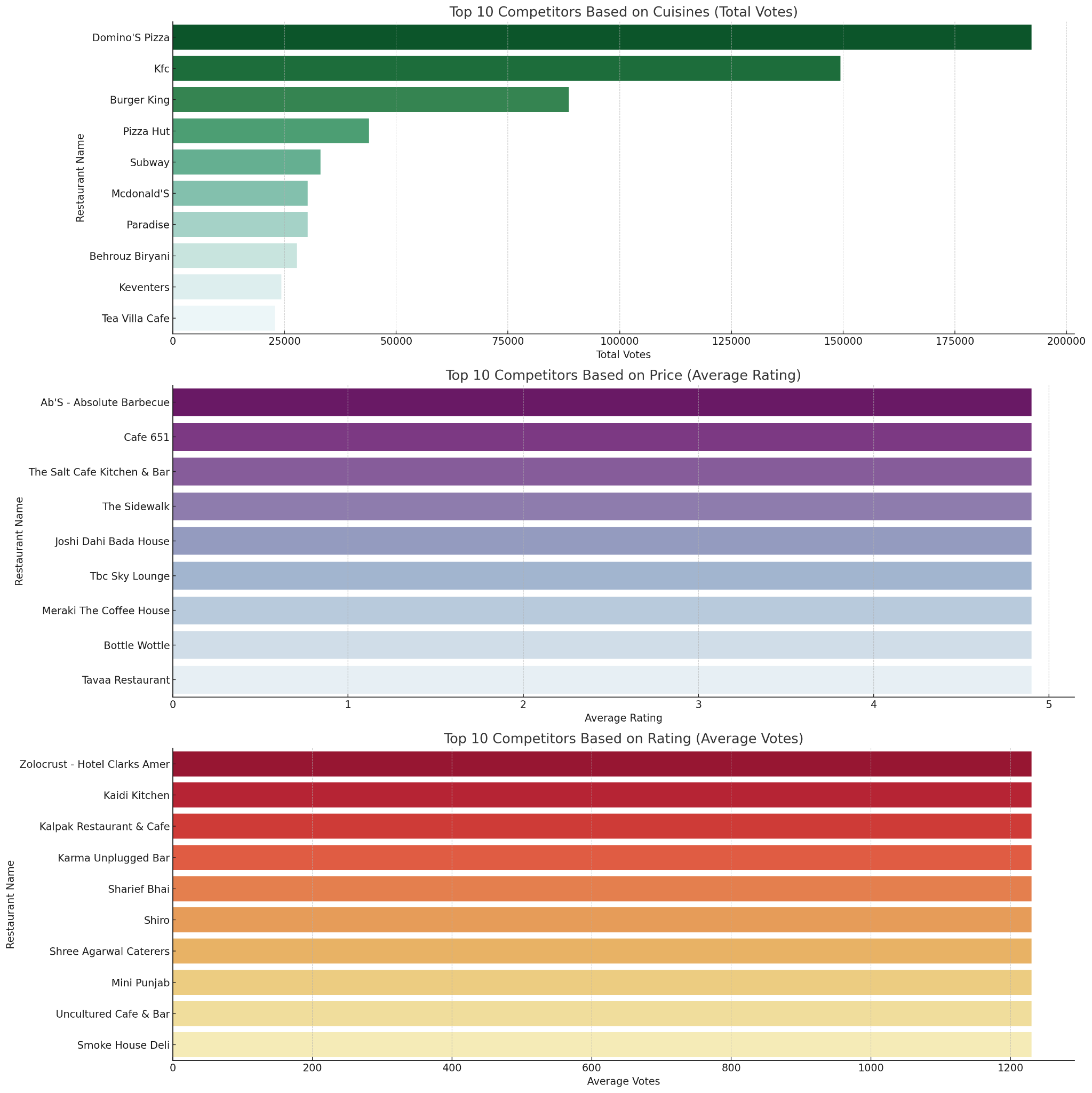
The availability of delivery and takeaway services did not significantly influence the average ratings, implying that food quality and service are the primary drivers of customer satisfaction.

**Competitive Analysis:**

To perform the competitor analysis on the given restaurant data, the following steps were undertaken:

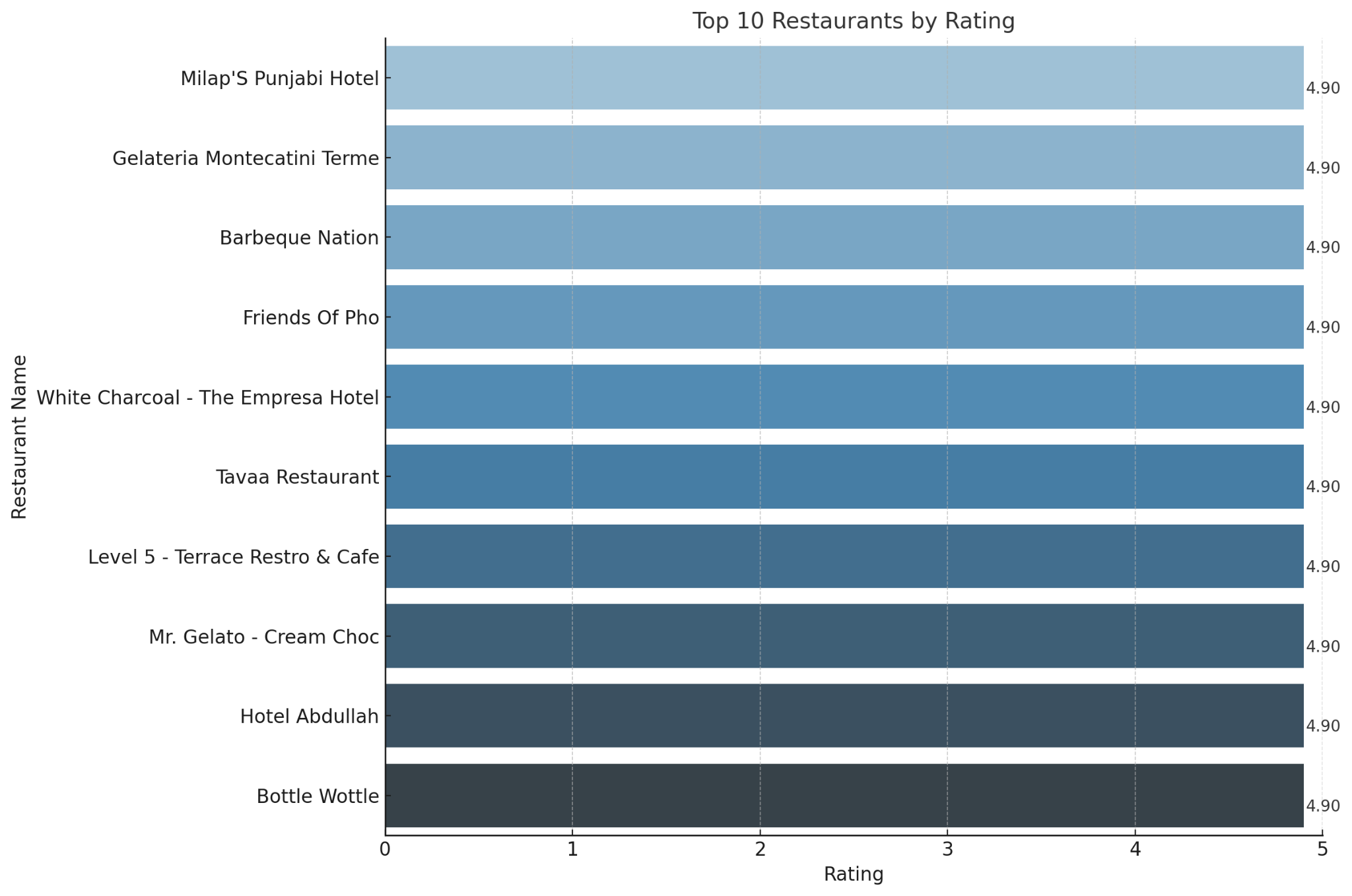
* Identified major Competitors:
* Identified significant competitors based on ratings
* Analyzed the strengths and weaknesses of these competitors
* Analyzed the market presence of major competitors
* Identified most popular cuisines among customers

**Top ten Competitors based on Cuisines**



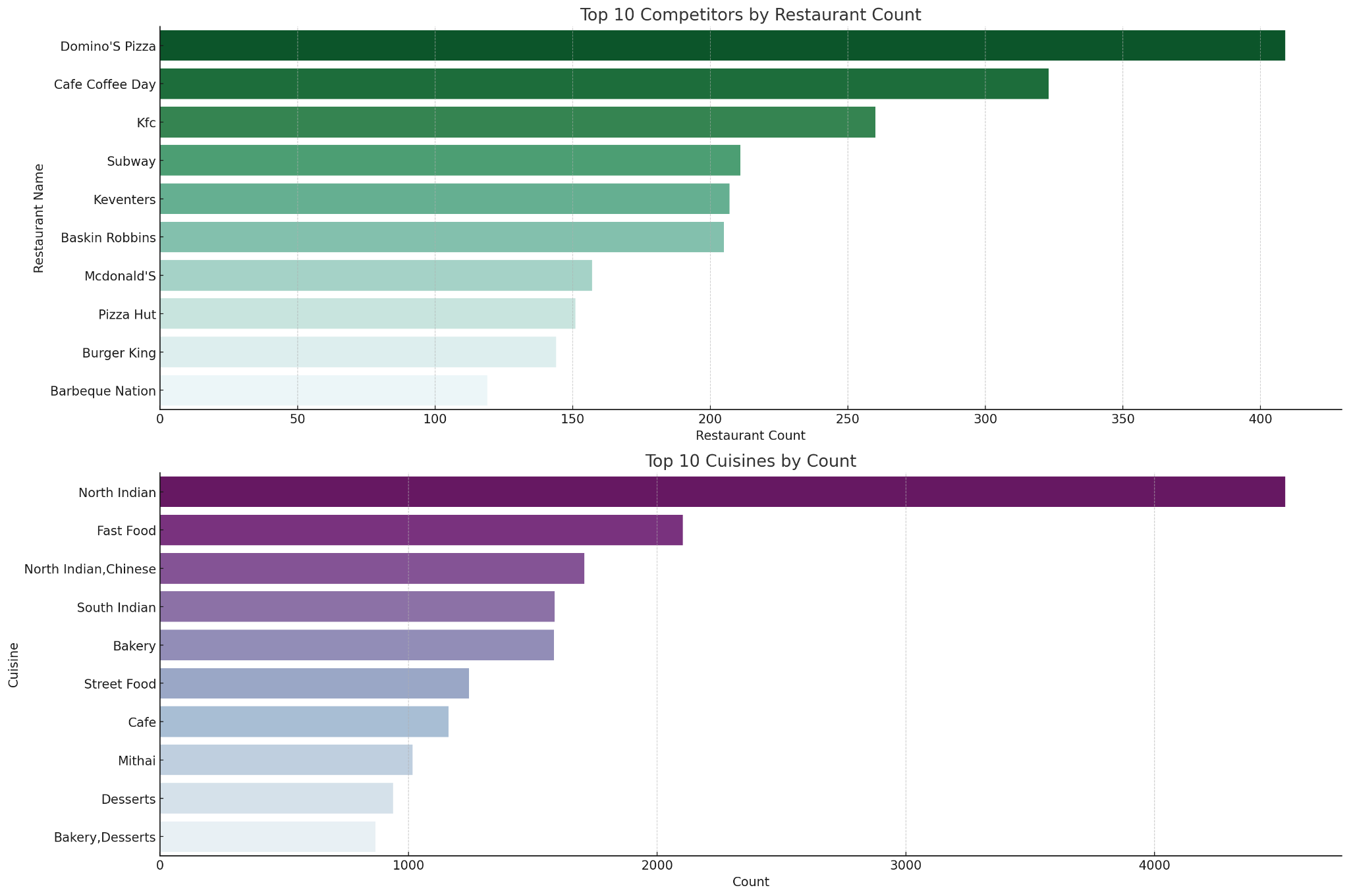
The chart shows the restaurants receiving the most votes in specific cuisine categories, providing a clearer picture of customer engagement and popularity.

**Top ten Competitors based on Ratings:**



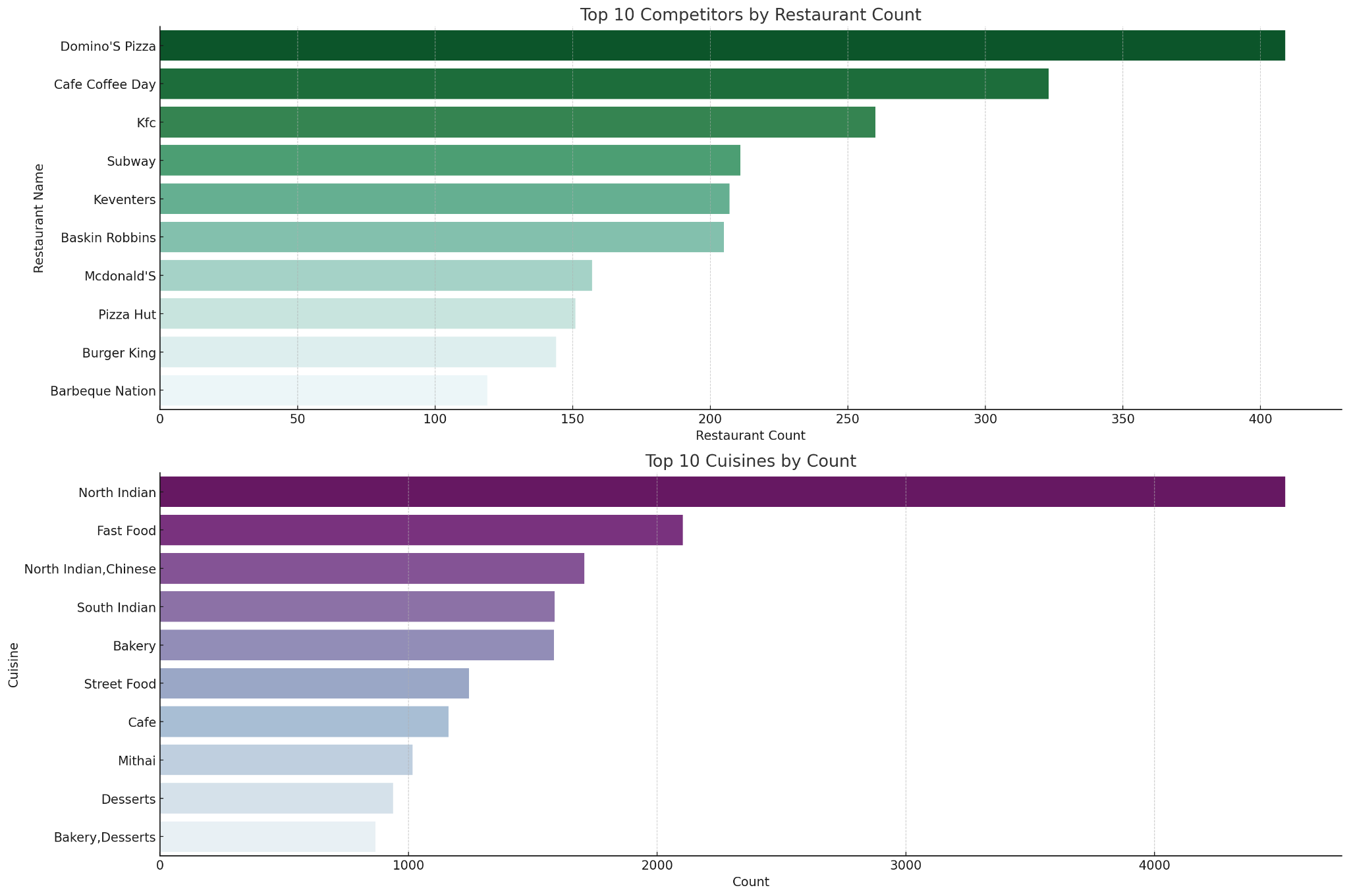
### This chart provides a clear and concise overview of the highest-rated restaurants in the data.

**Top 10 Competitors by Restaurant Count**



The chart shows the names of restaurants with the highest count, indicating the significant players in terms of market presence. These are likely well-established brands with multiple outlets.

**Top ten cuisines by count**



This chart ranks the most common cuisines based on their prevalence in the market. It reflects consumer preferences and the potential saturation levels of different cuisines.

**Key Insights:**

* **Established Competitors**: Identifying the restaurants with the most outlets can guide strategies for competing against or learning from established players.
* **Consumer Preferences and Opportunities:** Understanding the most common cuisines can help tap into existing demand or explore niche markets with less common cuisines for differentiation.
* Regional Targeting: Analyze the popular cuisines and dining preferences in various cities. For example, if North Indian cuisine is popular in a specific city, focus marketing efforts on restaurants offering these cuisines in that area.
* Customer Segmentation: Focus on customer preferences revealed in the data. For instance, if certain cuisines are highly popular, tailor marketing to highlight these options. Also, consider price sensitivity and quality expectations across different segments.
* Differentiation from Competitors: Identify the strengths and weaknesses of competitors through the analysis. Highlight unique offerings or superior aspects of Zomato-partnered restaurants.
* Promotional Tactics:
  + Discounts: Offer discounts on popular cuisines or during specific times to drive traffic.
  + Loyalty Programs: Develop a loyalty program that rewards frequent customers, encouraging repeat business.
  + Special Events: Host or sponsor food festivals or culinary events showcasing popular or unique cuisines.
* Leveraging Data Insights: Utilize insights on customer ratings, price ranges, and cuisine popularity to inform marketing messages and promotions.