zomato

DECODING DINING DYNAMICS

CONTENTS

ABOUT THE DATASET

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

1. **Res_id**: Unique ID for each restaurant.
2. **Name**: Name of the restaurant.
3. **Establishment**: Type of establishment (e.g., Quick Bites, Casual Dining).
4. **URL**: URL of the restaurant on Zomato.
5. **Address**: Physical address of the restaurant.
6. **City**: The city where the restaurant is located.
7. **Latitude**, **longitude**: Geographical coordinates.
8. **Cuisines**: Types of cuisines offered.
9. **Average_cost_for_two**: Average cost for two people.
10. **Price_range**: Price range category.
11. **Currency**: Currency used for prices.
12. **Highlights**: Key features or services offered.
13. **Aggregate_rating**: Average rating on Zomato.
14. **Rating_text**: Rating category (e.g., Excellent, Very Good).
15. **Votes**: Number of votes the restaurant has received.
16. **Photo_count**: Number of photos uploaded for the restaurant.
17. **Delivery**, **takeaway**: Information about delivery and takeaway services.

DATA CLEANING PROCESS

The data cleaning process was crucial in preparing the Zomato dataset for thorough analysis. This process ensured the accuracy, consistency, and usability of the data. Here's how we approached it:

Handling Missing Values:

Zipcode	**163,187**	**76.99%**
Cuisines	**1391**	**0.66%**
Timings	**3874**	**1.83%**
Opentable_support	**48**	**0.02%**

Cleaning

- The zip code column is dropped due to a high percentage of missing values.
- Rows with missing values in address, cuisines, and timings were removed, and they were minimal and critical for analysis.

Duplicate Data Handling:

- Analyzed and removed exact duplicate entries to ensure data uniqueness and integrity.

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.

Text Data Standardization and Correction:

Standardized text data for consistency, including:

1. ·Converting **name** and **locality** columns to title case.
2. ·Correcting encoding issues in the **timings** column and splitting it into standardized **operating hours** and **operating**.

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.

EXPLORATORY DATA ANALYSIS

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

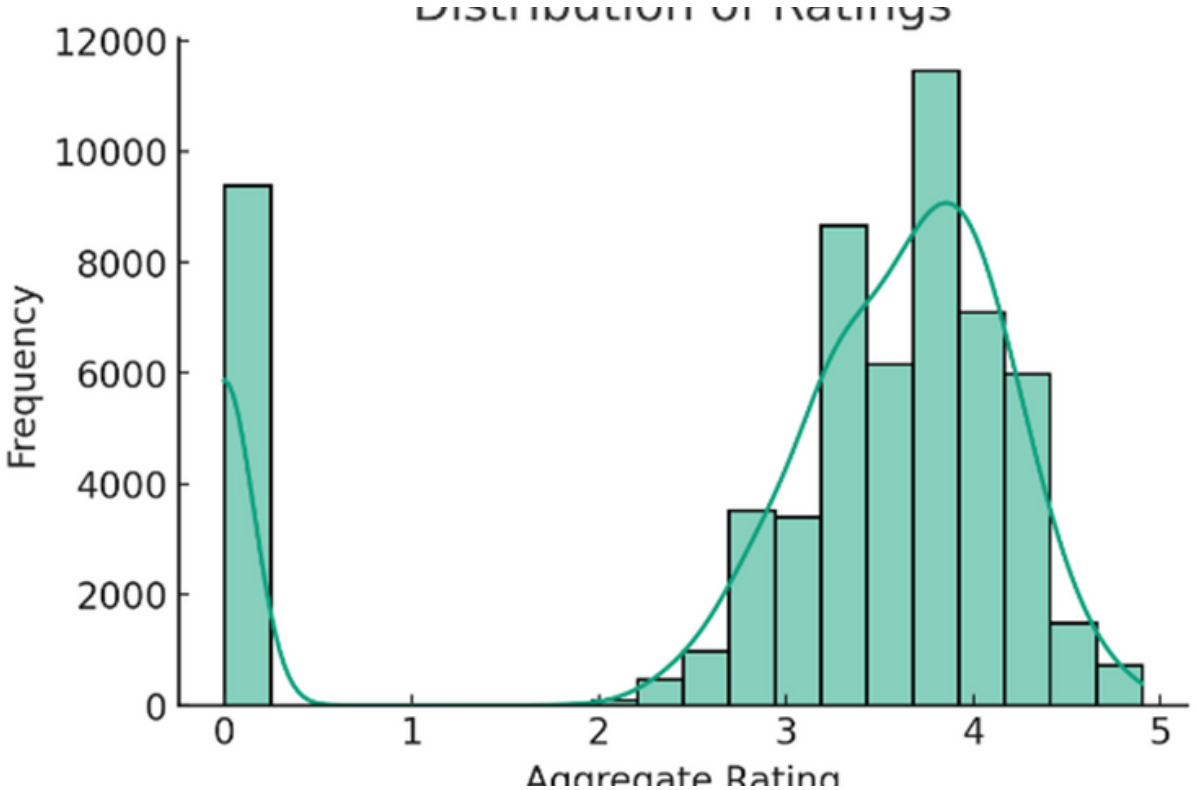
Statistical Measure	Average Cost for Two	Price Range	Aggregate Rating	Votes	Photo Count
Count	59464.00	59464.00	59464.00	59464.00	59464.00
Mean	503.16	1.74	3.07	200.07	121.31
Std Dev	409.77	0.88	1.41	329.96	248.77
Min	0.00	1.00	0.00	-18.00	0.00
25%	200.00	1.00	3.00	7.00	2.00
50% (Median)	400.00	1.00	3.60	45.00	12.00
75%	600.00	2.00	4.00	216.00	86.00
Max	1600.00	4.00	4.90	1230.00	967.85
Skewness	1.43	0.98	-1.45	2.08	2.52
Kurtosis	1.22	0.04	0.71	3.32	5.28

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).

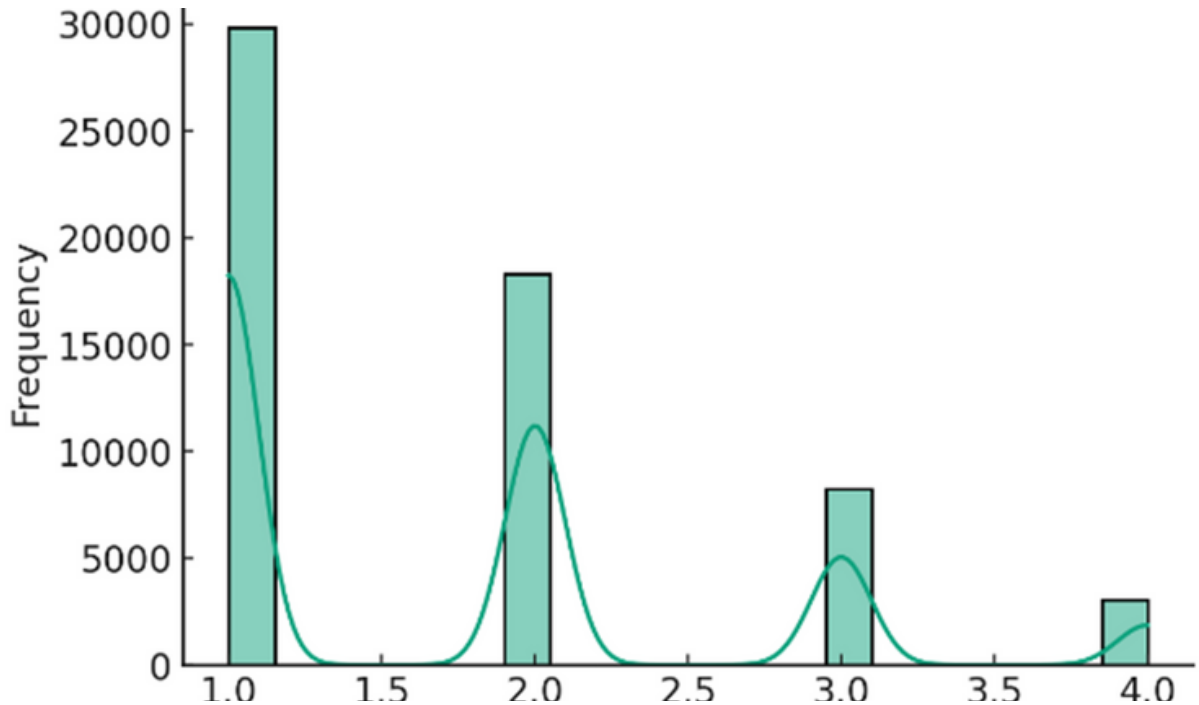
Aggregate Rating Distribution:

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).



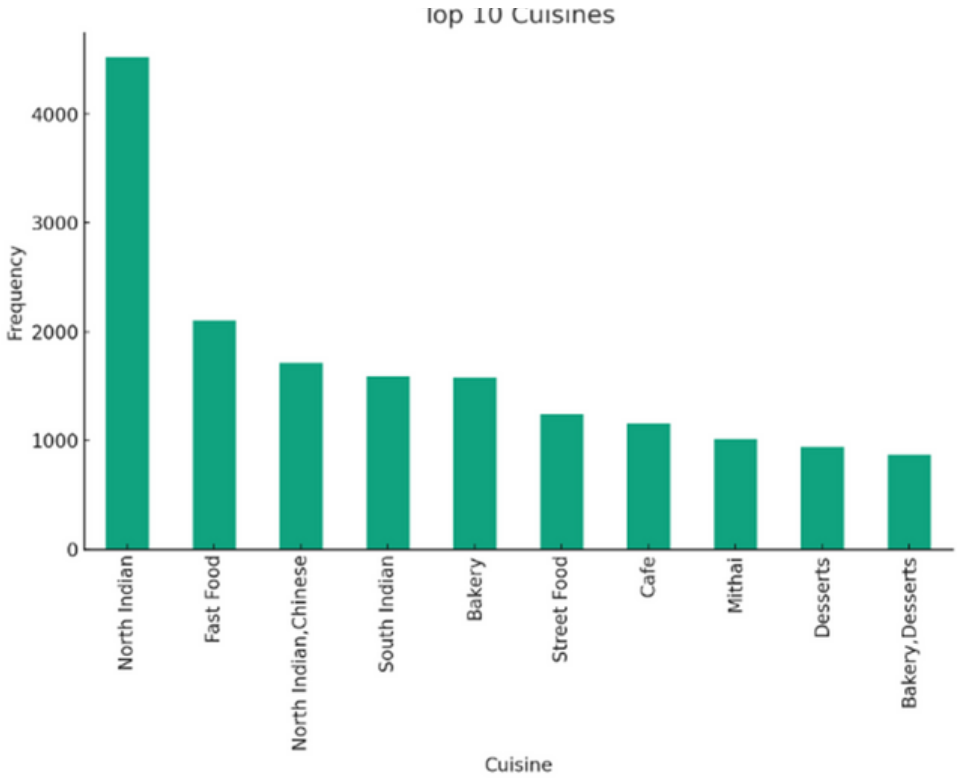
Price Range Distribution:

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.

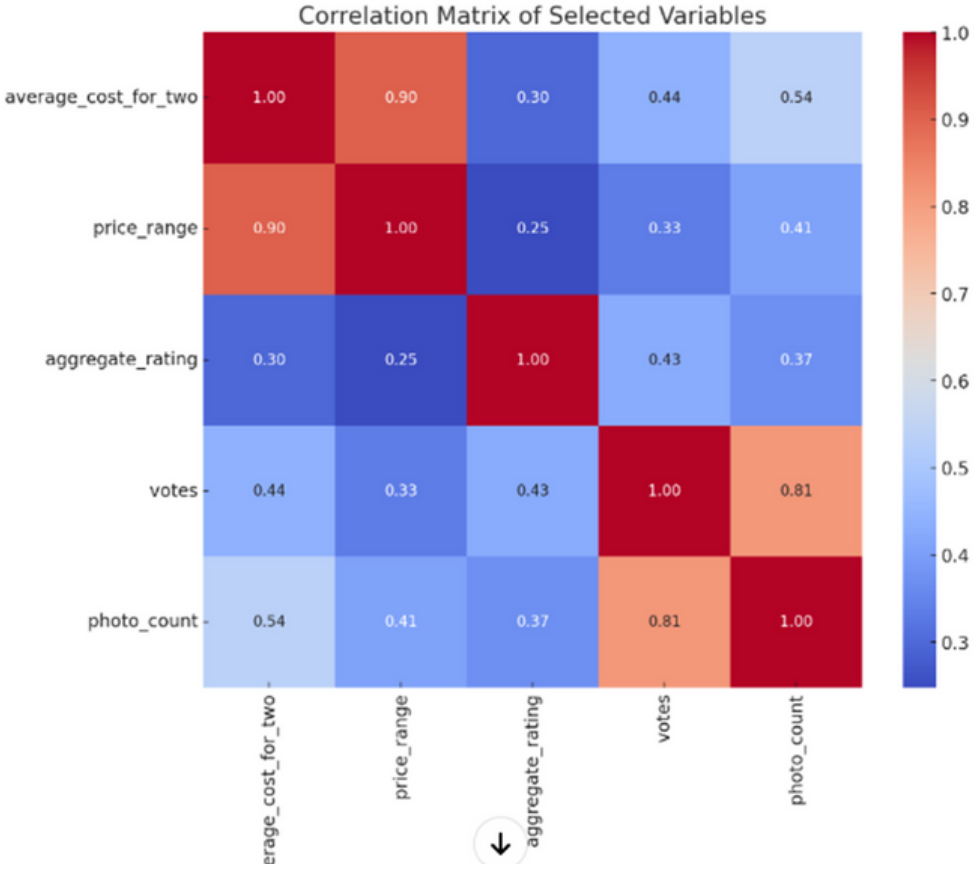


Top 10 Cuisines:

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.

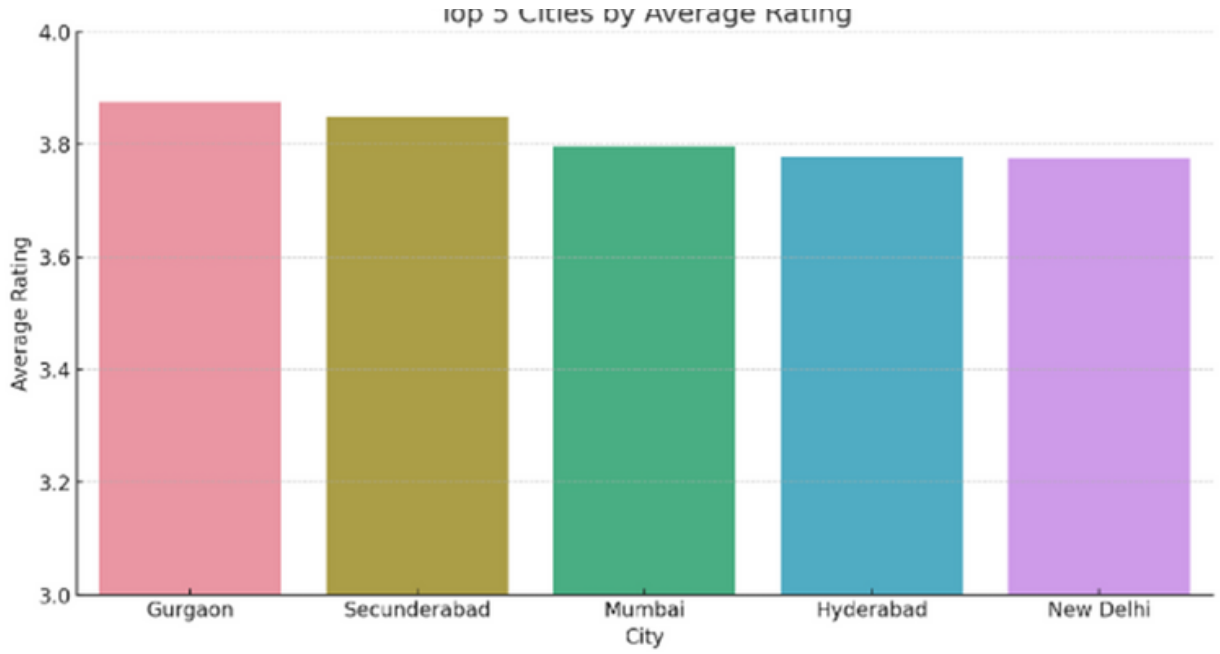
Votes and Photo Count:
A strong positive correlation (0.68) is observed between the number of votes and the photo count. This suggests that restaurants with more votes also tend to have a higher number of photos, indicating higher customer engagement.

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:

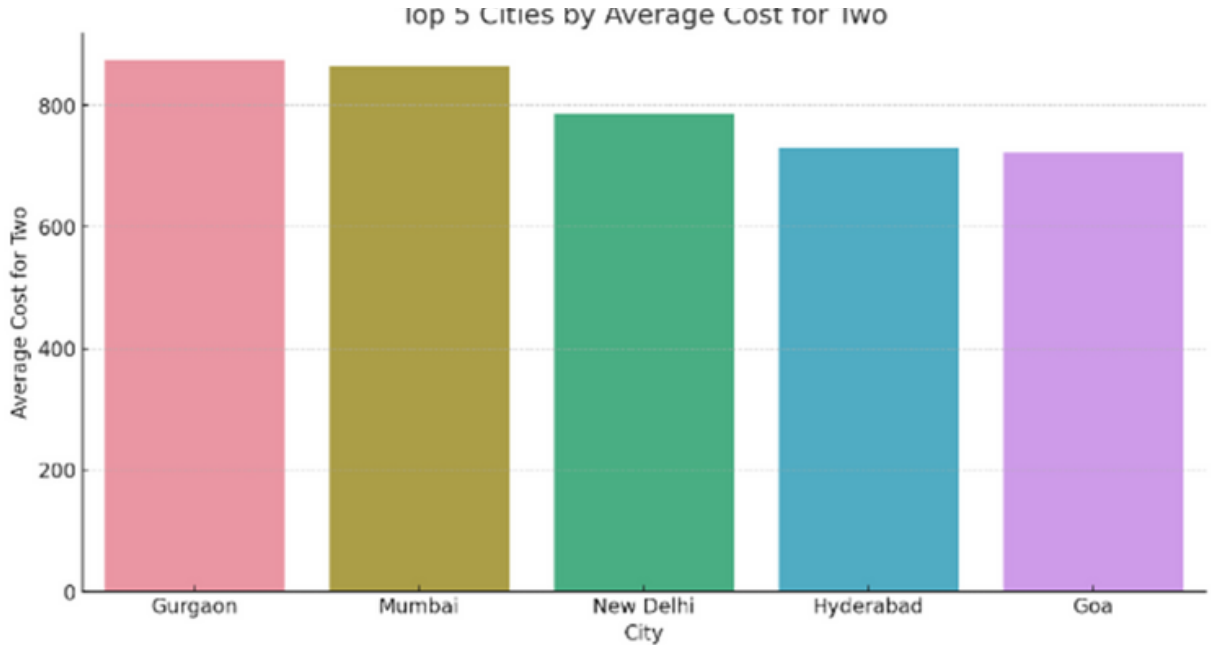
Compare the restaurant trends and customer preferences across different cities or regions in India.

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.



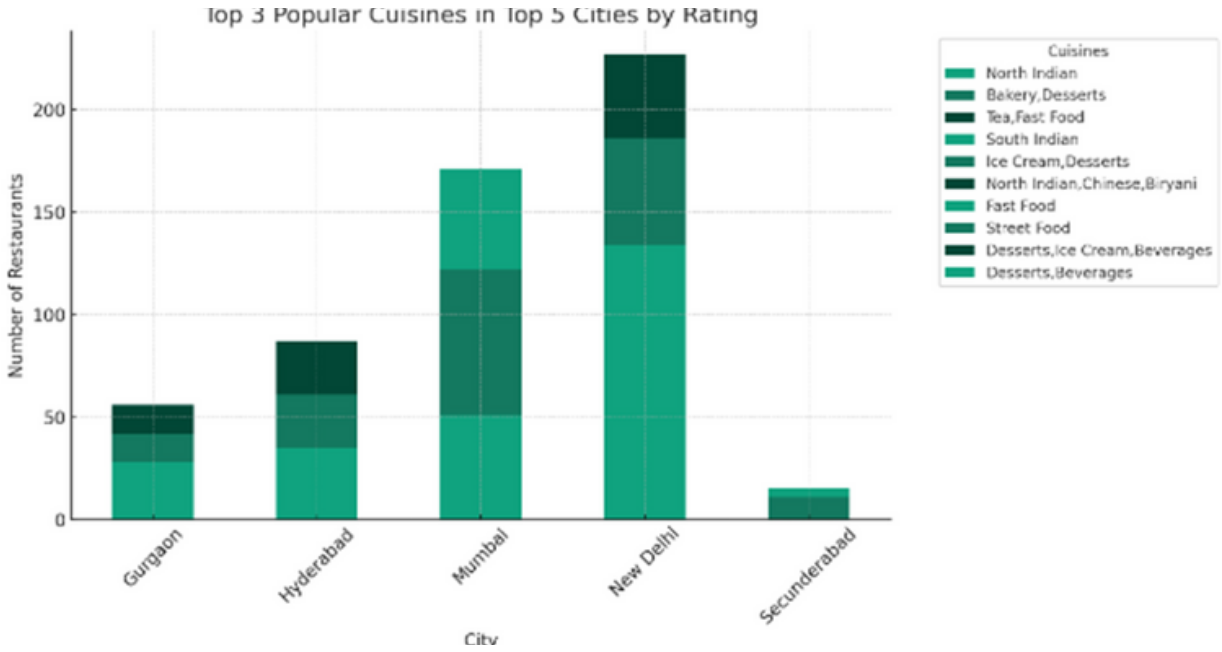
Top 5 Cities by Average Rating:

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 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.

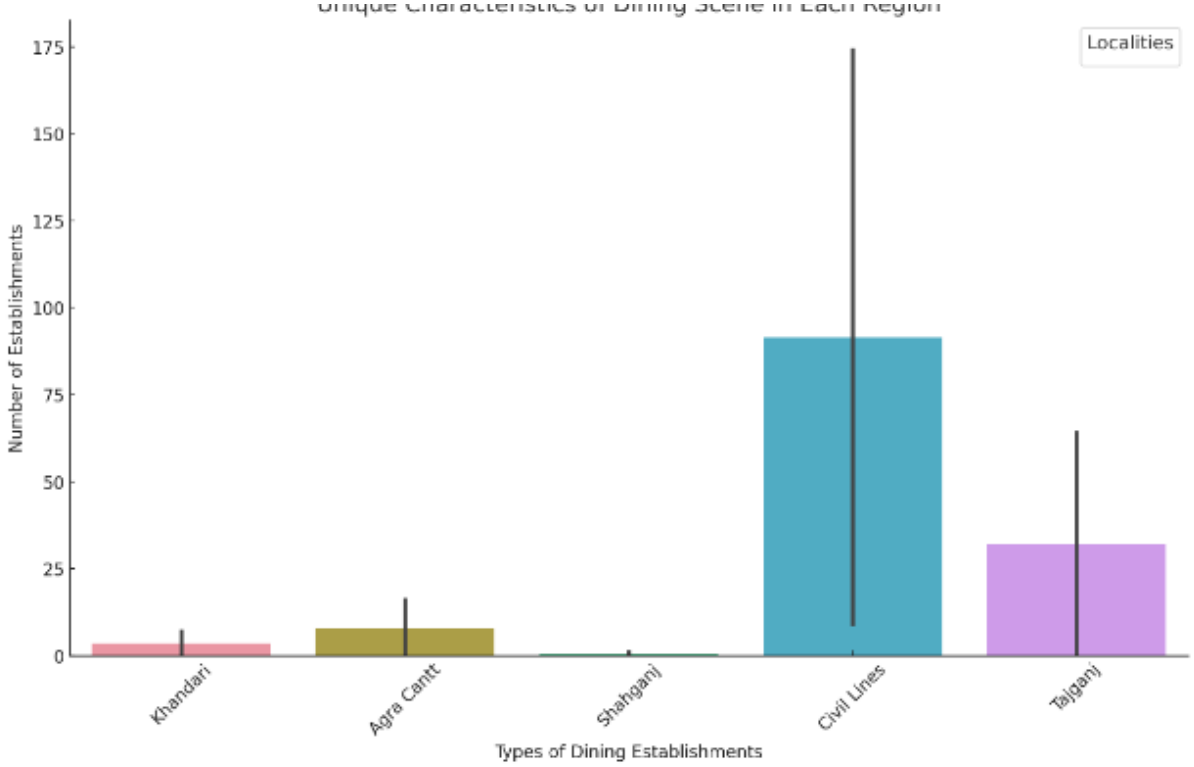


Identify unique characteristics of the dining scene in each region.

Top 5 cuisines across different days of the week.

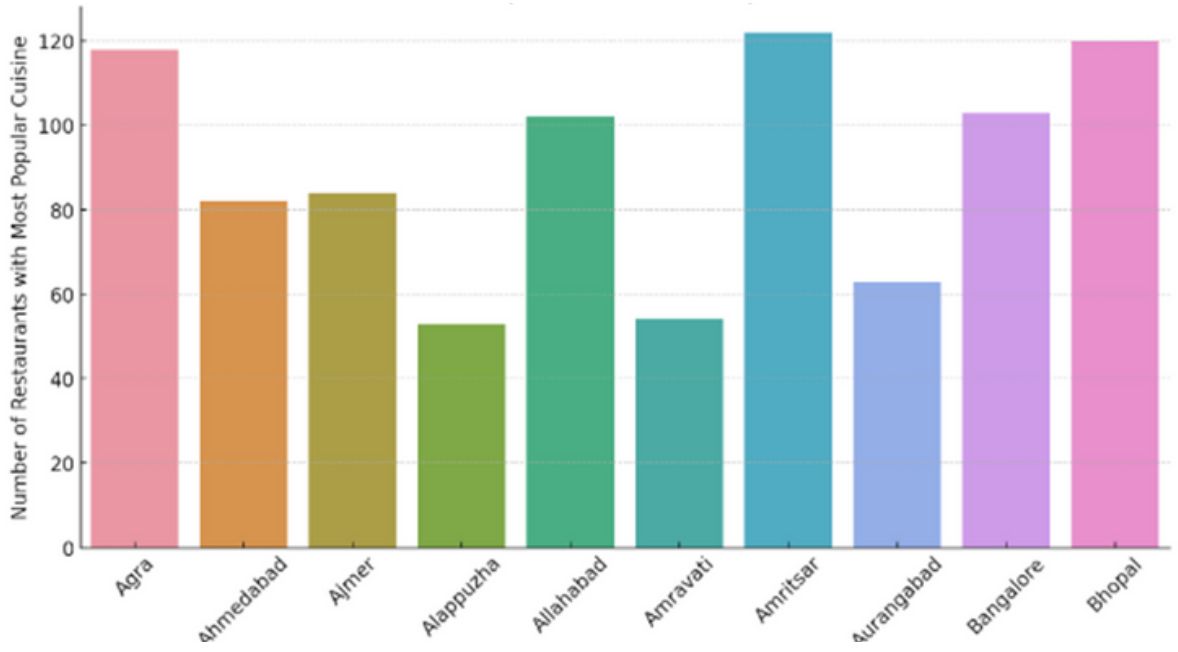
The graph above illustrates the unique characteristics of the dining scene in a selection of five regions (localities) from the dataset. Each bar represents a different type of dining establishment (such as restaurants, bistros, etc.), and the height of the bar indicates the number of establishments of that type within each locality.

This visualization helps us understand the diversity and distribution of dining options in different regions. It highlights how certain dining establishments may be more prevalent in one region than others, offering insights into regional dining preferences and trends.



Most Popular Cuisine in Top Cities:

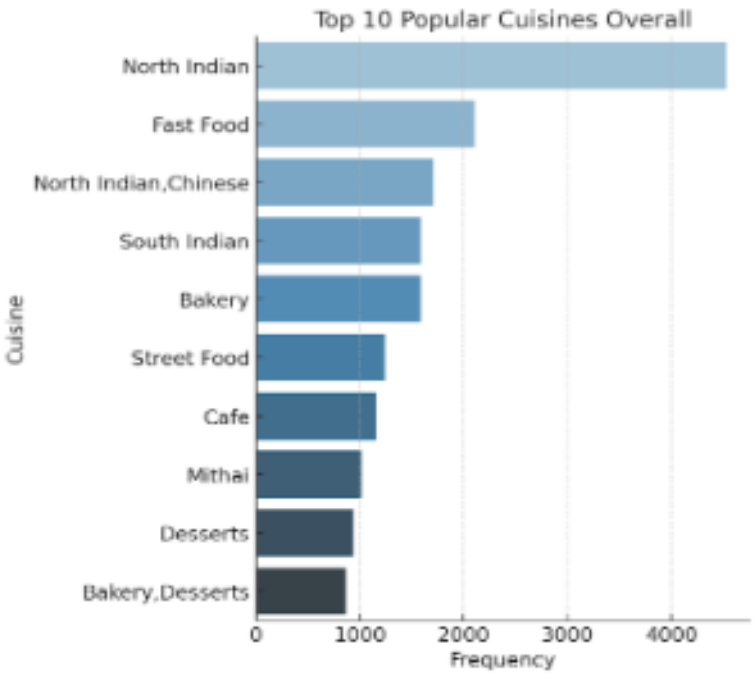
This graph displays the number of restaurants offering the most popular cuisine in each city. It underscores the prevalence of specific cuisines like North Indian and Street Food in these regions, indicating regional culinary preferences.



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.

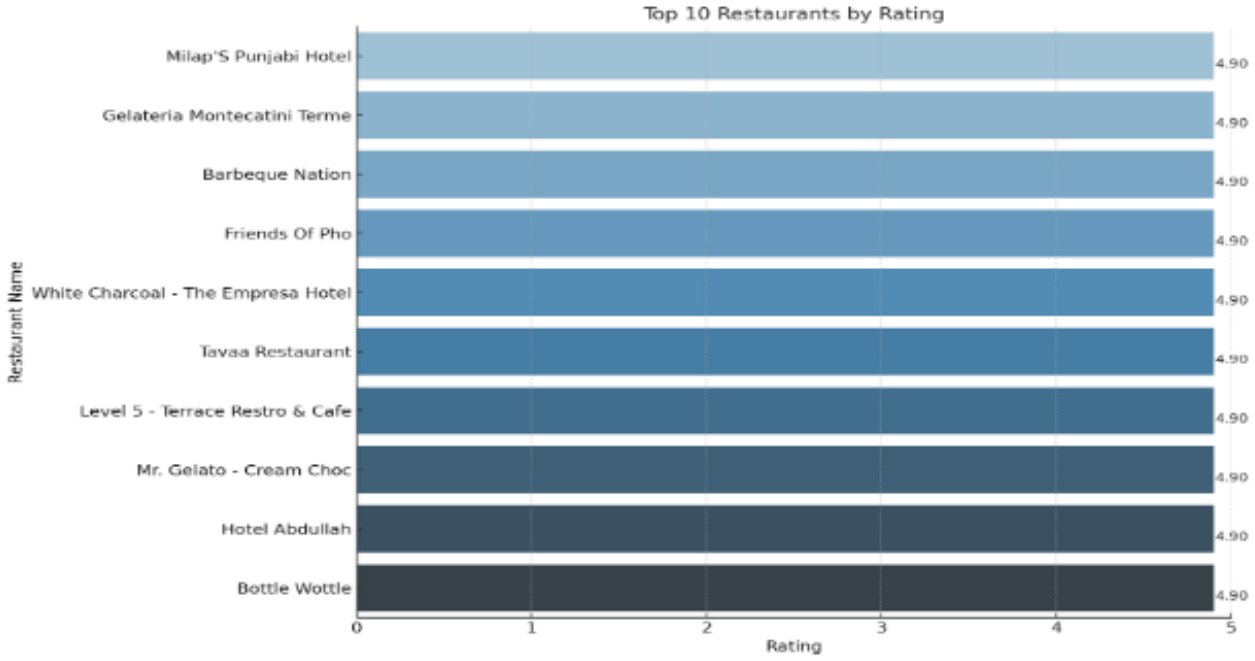
Top Ten Cuisines



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.

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

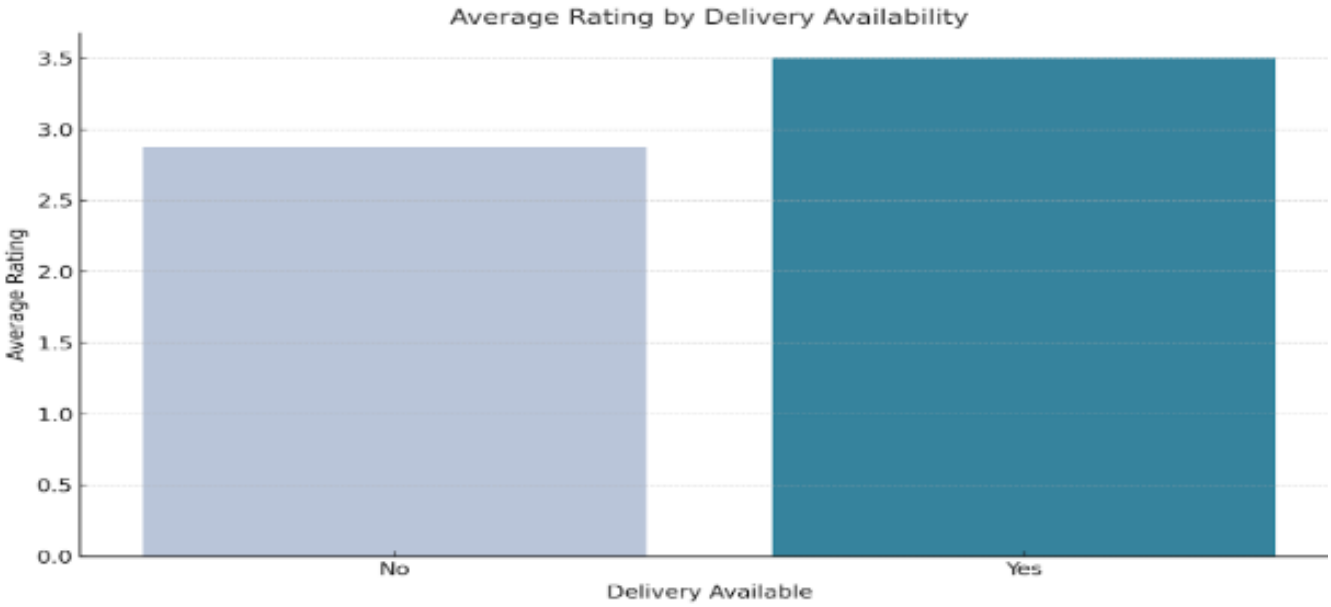
This chart provides a clear and concise overview of the highest-rated restaurants in the data.



Average Rating by Price Range

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.

Average Rating by Delivery Availability



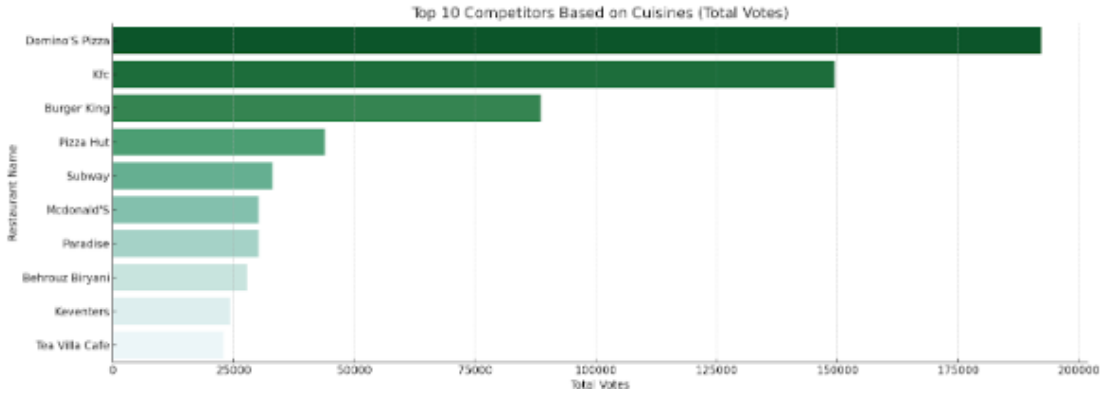
The visualization above shows restaurants' average ratings based on delivery service availability. It further explains that the availability of delivery service does not significantly influence the average ratings of restaurants. This suggests that factors such as food quality, customer service, and overall dining experience play a more crucial role in determining a restaurant's rating than the availability of delivery services.

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

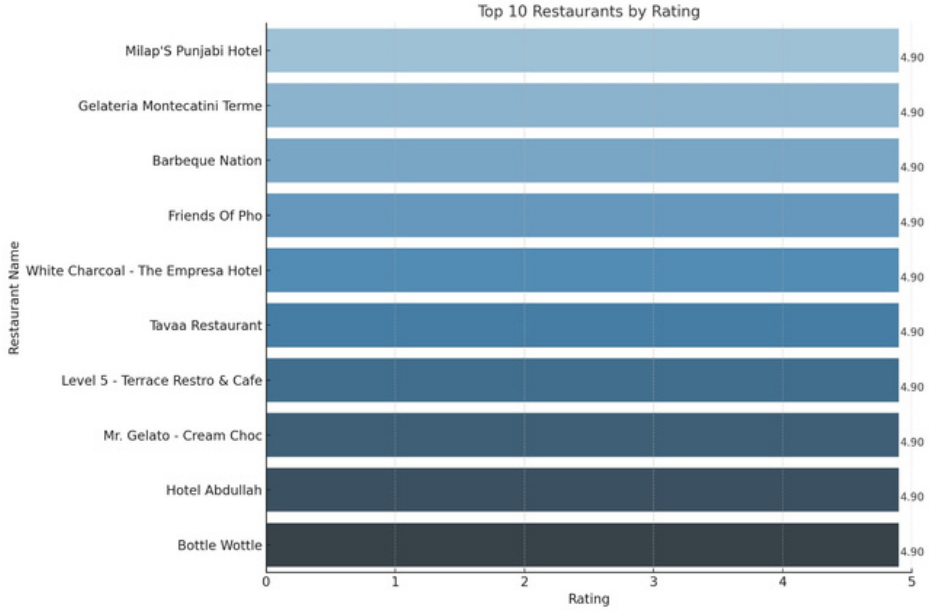
Top ten Competitors based on Cuisines



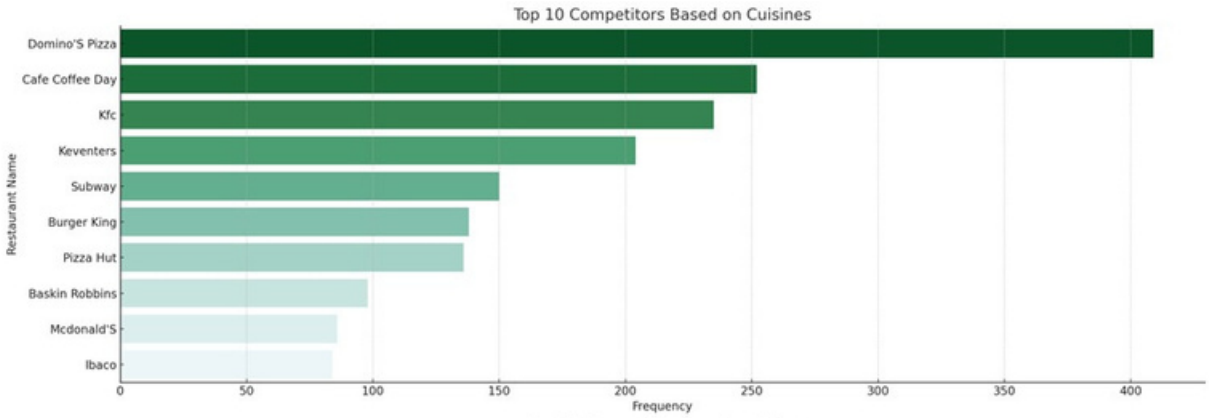
The chart shows the restaurants receiving the most votes in specific cuisine categories, providing a clearer picture of customer engagement and popularity.

Top ten Competitors based on Cuisines

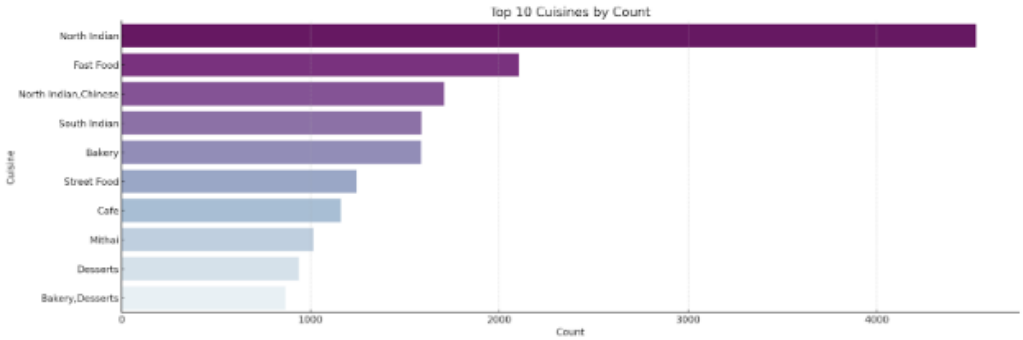
This chart provides a clear and concise overview of the highest-rated restaurants in the data.



Top 10 Competitors by Restaurant Count



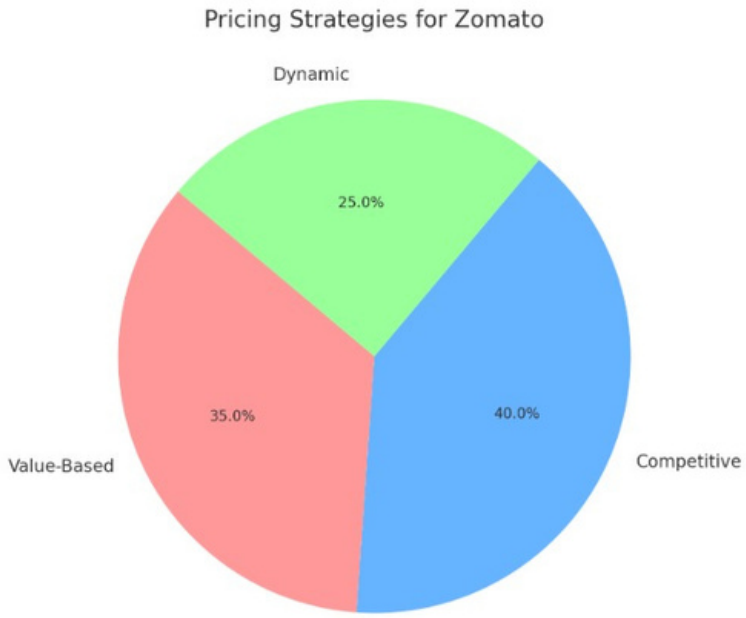
Top ten cuisines by count



Marketing Campaign for Zomato

"Embark on a journey through Zomato's vibrant marketplace, where every click opens doors to diverse culinary experiences. Here, we'll explore how Zomato's strategic marketing moves have carved its path in the bustling world of food delivery."

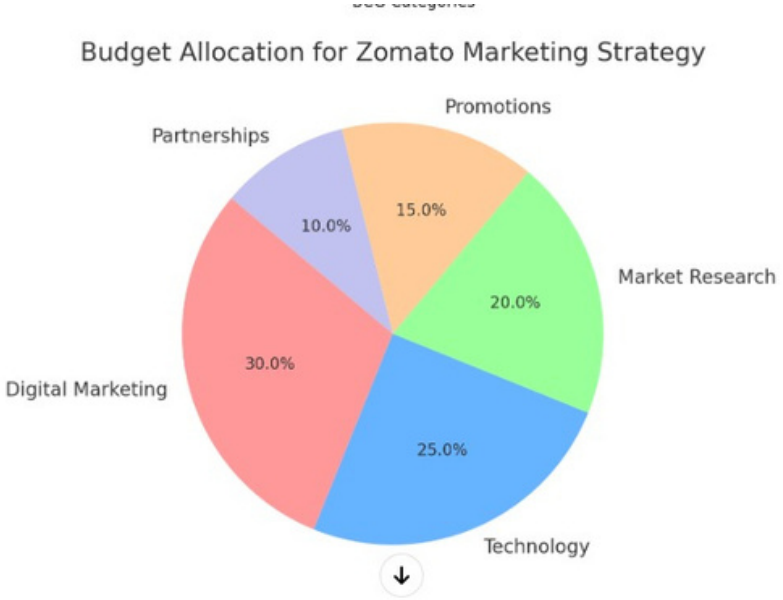
Pricing Strategies



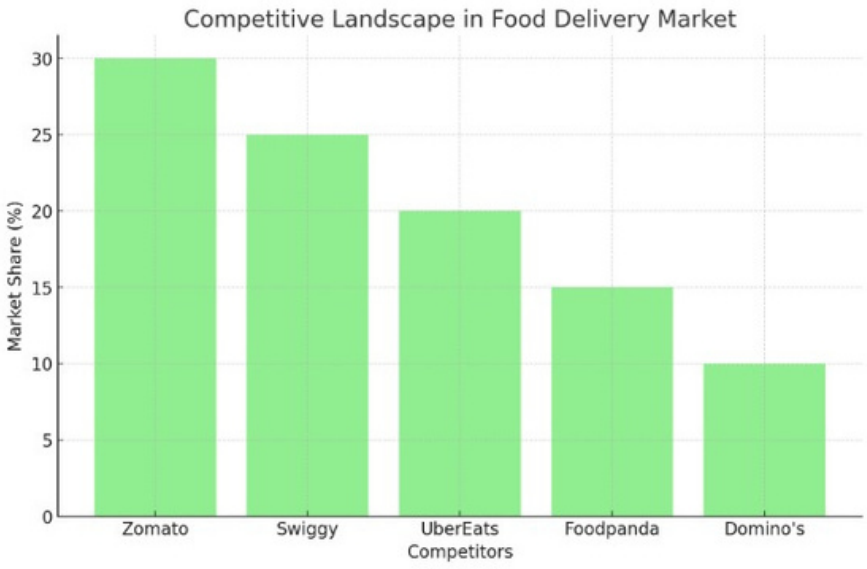
The pie chart above represents the preferences for different pricing strategies for Zomato:
Value-Based Pricing: Comprising 35% of the strategy preference, this approach focuses on setting prices based on the perceived value to customers.
Competitive Pricing: The largest segment at 40%, this strategy involves aligning prices with competitors to stay competitive in the market.
Dynamic Pricing: Making up 25% of the preference, this strategy varies prices based on demand, time, or other factors, offering flexibility in pricing.
This visualization provides an overview of how Zomato might focus on different pricing strategies to effectively cater to market demands and competitive dynamics.

Budget Allocation

Budget Allocation Chart: The pie chart shows the proposed distribution of Zomato's marketing budget across various channels and initiatives.



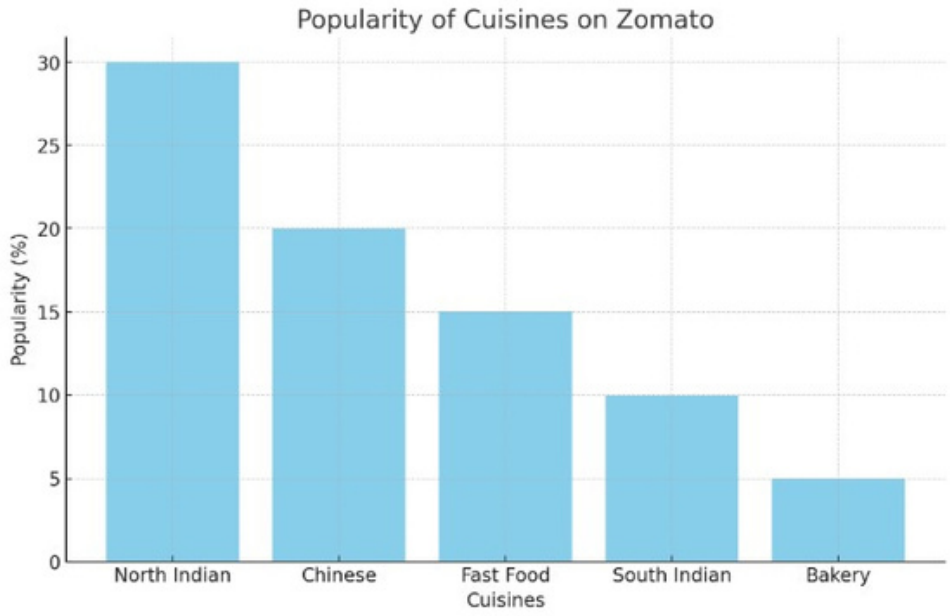
Competitor analysis



Competitive Landscape Chart: This chart displays Zomato's market share in comparison to its main competitors, highlighting its position in the food delivery market.

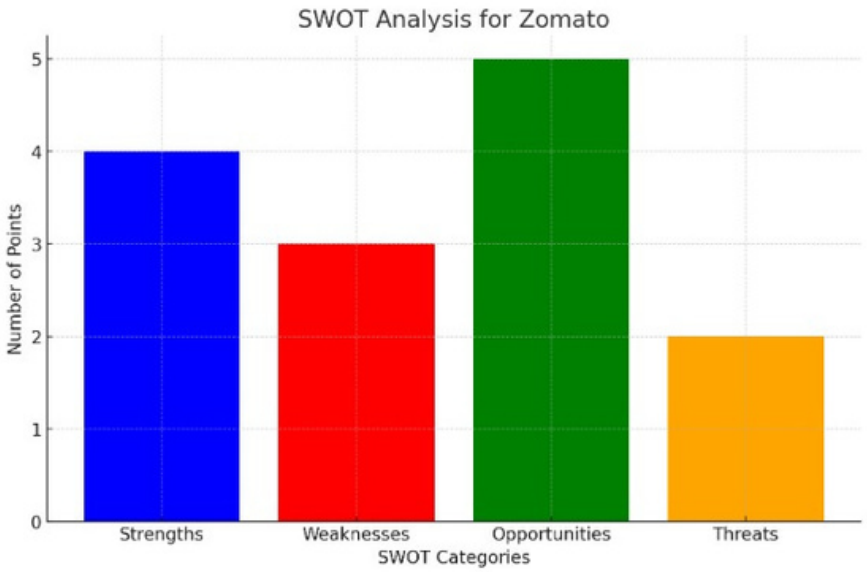
Popularity of cuisines on Zomato

This bar graph shows the popularity of different cuisines on Zomato, indicating where to focus menu development and marketing efforts.



SWOT Analysis Zomato

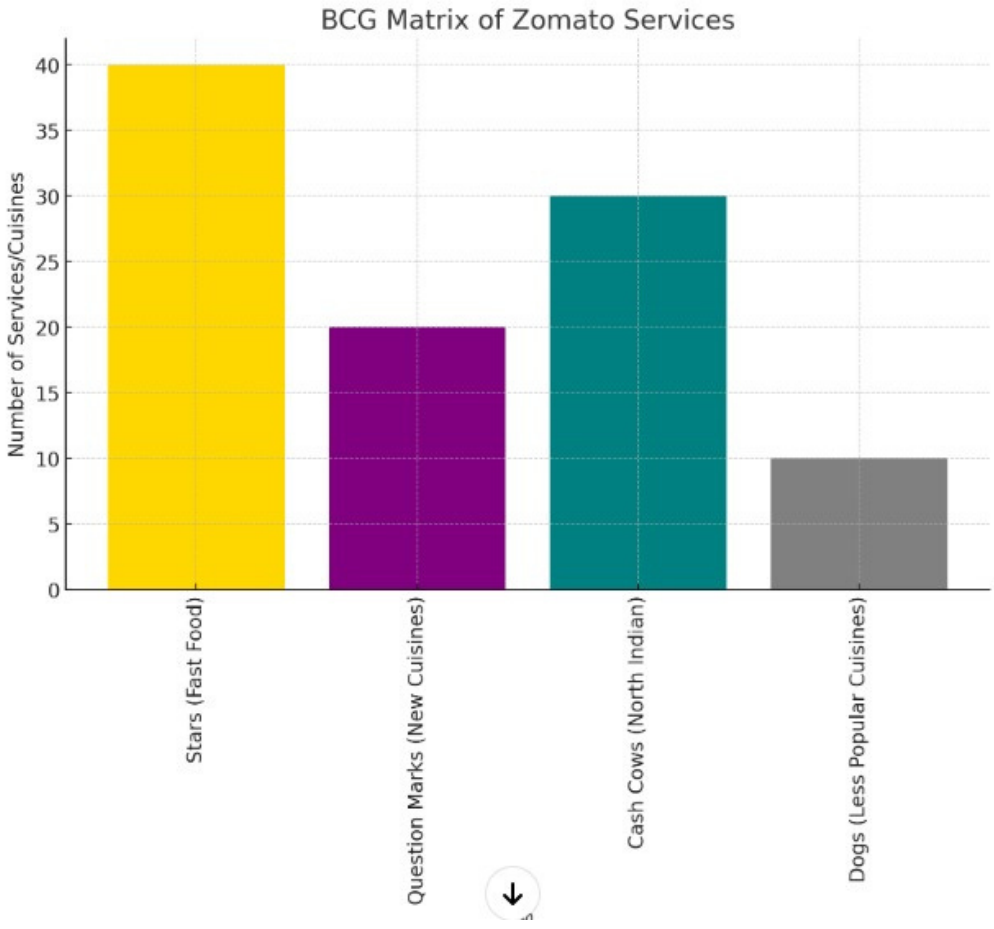
SWOT Analysis Diagram: The bar chart represents Zomato's strengths, weaknesses, opportunities, and threats, providing a quick visual summary of the SWOT analysis.



BCG Matrix for Zomato

The revised graph above illustrates the BCG Matrix for Zomato's services, with each category labeled to indicate a specific type of service or cuisine:

Stars (Fast Food): Representing popular and fast-growing segments like fast food.
Question Marks (New Cuisines): New or experimental cuisines with potential for growth but currently uncertain market share.
Cash Cows (North Indian): Established and popular cuisines like North Indian have a significant and steady market share.
Dogs (Less Popular Cuisines): Cuisines that have lesser popularity and market growth.
This visualization aids in understanding the strategic positioning of different types of services or cuisines offered by Zomato and can guide resource allocation and strategic focus.



Conclusion

The conclusion of the emphasizes several key strategic recommendations for Zomato:

Discounts: It suggests offering discounts on popular cuisines or during specific times to drive traffic, which can help in attracting more customers and increasing sales.

Loyalty Programs: The development of a loyalty program is recommended to reward frequent customers, thereby encouraging repeat business. This approach can enhance customer loyalty and long-term engagement.

Special Events: Hosting or sponsoring food festivals or culinary events showcasing popular or unique cuisines is recommended. This strategy can help build brand awareness and attract diverse customer segments.

Leveraging Data Insights: This stresses the importance of utilizing insights on customer ratings, price ranges, and other key metrics to inform strategic decisions.

Overall, this concludes with a focus on using targeted marketing strategies and data-driven insights to enhance Zomato's market presence, customer engagement, and overall business performance. These strategies are aimed at reinforcing Zomato's position in the competitive food delivery and restaurant industry.