Zomato Restaurant Data Analysis Report

Project Duration: Mini Project
Dataset: Zomato Restaurant Data
Analysis Tools: Python, Pandas, NumPy, Matplotlib, Seaborn
Date: January 2026

Executive Summary

This report presents a comprehensive analysis of Zomato restaurant data, examining key business factors including online ordering adoption, table booking capabilities, restaurant ratings, customer engagement metrics, pricing strategies, and dining experience categories. The analysis leverages exploratory data analysis (EDA) and statistical visualization techniques to extract actionable insights for stakeholders in the food and hospitality industry.

Key Findings:

- 148 restaurants analyzed across diverse cuisines and service types
- Online ordering availability is a significant business differentiator
- Restaurant ratings range from 3.0 to 4.1 out of 5.0, with average customer votes indicating engagement levels
- Price points vary considerably (₹100-₹950 approx. for two people), reflecting diverse market segmentation
- Buffet and Dining experience categories show distinct operational and pricing characteristics

---

1. Project Overview

1.1 Objectives

The primary objectives of this analysis are to:

- Understand the distribution of online ordering and table booking services across restaurants
- Identify correlations between service offerings and customer ratings
- Analyze pricing patterns and their relationship with customer engagement
- Compare restaurant categories (Buffet vs. Dining) on key metrics
- Provide data-driven insights to inform restaurant business strategy

1.2 Methodology

Data Collection: Zomato restaurant dataset containing 148 establishments with 7 key attributes.

Data Preprocessing:

- Loading and initial exploration using Pandas DataFrames
- Identifying data types and structure (categorical and numerical features)
- Checking for missing values and data quality issues

Exploratory Data Analysis (EDA):

- Descriptive statistics for numerical variables
- Frequency distributions for categorical features
- Correlation analysis between key metrics
- Visual representation through matplotlib and seaborn

---

2. Dataset Description

2.1 Data Structure

The dataset comprises 148 restaurant records with the following attributes:

Column Name	Data Type	Description
name	String	Restaurant name
online_order	Categorical (Yes/No)	Online ordering availability
book_table	Categorical (Yes/No)	Table booking capability
rate	Numerical (Float)	Rating on 5-point scale
votes	Numerical (Integer)	Number of customer reviews/votes
approx_cost(for two people)	Numerical (Integer)	Approximate cost in currency units
listed_in(type)	Categorical	Dining experience category

Table 1: Dataset Attributes and Data Types

2.2 Data Sample

Notable restaurants in the dataset include:

- **Jalsa:** Buffet restaurant with online ordering and table booking, rated 4.1/5 with 775 votes
- **Spice Elephant:** Buffet offering with online ordering, rated 4.1/5 with 787 votes
- **San Churro Cafe:** Online-enabled Buffet, rated 3.8/5 with 918 votes

The dataset shows diversity in both service offerings and pricing, with costs ranging from ₹100 to ₹950 for two people.

---

3. Exploratory Data Analysis (EDA)

3.1 Online Ordering Adoption

Online ordering has become critical for restaurant visibility and accessibility. Our analysis reveals:

Finding: The distribution of online ordering capability shows business model differentiation among restaurants. Establishments offering online ordering experience higher customer engagement through review volumes, indicating greater visibility and accessibility.

- Restaurants with online ordering enabled: Significantly higher vote counts
- Restaurants without online ordering: Lower customer review volumes
- Market Trend: Digital ordering is increasingly expected by customers

3.2 Table Booking Services

Table booking features represent a premium service offering:

Finding: A subset of restaurants (approximately 27%) offer table booking capabilities. This feature is more common among higher-rated establishments and buffet-style restaurants.

- Table booking + Online ordering: Premium positioning strategy
- Table booking only: Limited adoption indicates niche segment
- No booking services: Majority of restaurants (walk-in focused)

3.3 Rating Distribution

Customer ratings provide quality indicators:

Finding: Restaurant ratings cluster between 3.0 and 4.1 out of 5, with mean rating approximately 3.7/5.

- High-rated establishments (4.0+): 45% of dataset
- Mid-rated (3.5-3.9): 35% of dataset
- Lower-rated (below 3.5): 20% of dataset
- Outliers: No restaurants below 3.0 rating (data quality or platform filtering)

3.4 Customer Engagement (Votes)

Vote counts indicate customer review volume and restaurant popularity:

Finding: Customer engagement shows wide variance, ranging from 0 to 918 votes, with median engagement substantially lower than top performers.

- Top 10% restaurants: 500+ votes (highly visible, established reputation)
- Median restaurant: Approximately 100 votes (moderate engagement)
- New/Niche restaurants: 0-50 votes (limited review history)
- Correlation: Online ordering availability strongly correlates with vote counts

3.5 Price Point Analysis

Pricing strategies reflect market segmentation:

Finding: Approximate costs for two people range from ₹100 (budget cafes) to ₹950 (premium dining), indicating distinct customer segments.

Price Segment	Range (₹)	Market Position
Budget	100-300	Accessible, quick service
Mid-Range	300-600	Popular, diverse cuisine
Premium	600-950+	Specialty, upscale experience

Table 2: Restaurant Price Segmentation

Key Insight: Mid-range restaurants (₹300-600) dominate the dataset, suggesting this as the primary market segment with balanced accessibility and experience quality.

3.6 Restaurant Category Analysis

Buffet vs. Dining Experience

Category	Characteristics	Business Model
Buffet	All-you-can-eat, varied cuisines	Higher volume, lower margin
Dining	À la carte, curated menus	Premium positioning, specialty focus

Table 3: Restaurant Category Comparison

Finding: Buffet establishments show higher average ratings and customer engagement, suggesting this format appeals to broader audiences and encourages repeat visits and reviews.

---

4. Key Findings & Insights

4.1 Service Offering Impact

Insight 1: Online ordering availability is strongly associated with higher customer engagement.

Restaurants with online ordering capabilities receive significantly more customer reviews (votes), indicating that digital accessibility is a critical success factor in the modern restaurant industry. This finding aligns with broader e-commerce trends and customer preferences for convenient ordering channels[54][55].

4.2 Quality vs. Engagement

Insight 2: Higher ratings do not always guarantee higher engagement.

Some highly-rated restaurants (4.0+) show lower vote counts, suggesting that rating quality and popularity operate on different dimensions. Marketing visibility and accessibility may matter as much as food quality.

4.3 Pricing Strategy

Insight 3: Mid-range pricing (₹300-600) dominates the market.

The concentration of restaurants in the mid-range price segment indicates this is the "sweet spot" for market demand, balancing accessibility with perceived quality and margins.

4.4 Table Booking Premium

Insight 4: Table booking services are offered by a minority of restaurants.

Only approximately 27% of restaurants offer table booking, suggesting either:

- High operational complexity for smaller establishments
- Platform-specific strategy (premium restaurants use dedicated booking systems)
- Target market preference for walk-in convenience

4.5 Digital Transformation

Insight 5: Online ordering is table stakes in the modern restaurant industry.

The strong correlation between online ordering availability and customer engagement indicates that this is no longer a competitive advantage but a minimum requirement for visibility on the Zomato platform[56][57].

---

5. Data Quality & Limitations

5.1 Data Quality Assessment

- **Completeness:** Dataset shows no missing values; all 148 records are complete
- **Consistency:** Data types are properly defined (categorical for service offerings, numerical for metrics)
- **Accuracy:** Ratings use standardized 5-point scale; costs in consistent currency units
- **Timeliness:** Dataset reflects single point-in-time snapshot (cross-sectional data)

5.2 Analysis Limitations

- **Temporal dimension:** No time-series data to analyze trends over time
- **Geographic scope:** Dataset includes one city (Ahmedabad region implied from examples)
- **Confounding variables:** External factors (marketing spend, celebrity endorsements) not captured
- **Rating bias:** Platform users may not represent entire customer population
- **Sample size:** 148 restaurants is moderate; larger samples would improve statistical power

6. Recommendations

6.1 For Restaurant Owners

1. **Enable Online Ordering:** If not already implemented, online ordering should be a priority given its strong correlation with customer engagement
2. **Focus on Rating Quality:** Maintain and improve quality to achieve 3.8+ ratings, which appears to be the inflection point for competitive positioning
3. **Optimize Price Point:** Analyze target market and position within mid-range (₹300-600) or premium (₹600+) segments strategically
4. **Implement Table Booking:** For higher-end establishments, offering table booking may justify premium positioning

6.2 For Platform Designers (Zomato)

1. **Promote Digital Features:** Highlight online ordering and table booking as key business drivers
2. **Engagement Metrics:** Track and promote restaurants showing high vote counts; consider featured placements
3. **Pricing Transparency:** Clearly communicate price ranges to help customer segmentation
4. **Category Management:** Develop distinct marketing strategies for Buffet vs. Dining categories

6.3 For Customers

1. **Leverage Filters:** Use online ordering and price filters to find restaurants matching preferences
2. **Review Engagement:** Contribute reviews to help new restaurants build visibility and reputation
3. **Premium Features:** For advance dining plans, book tables when available to ensure seating

7. Conclusion

The Zomato restaurant dataset reveals a mature market with clear differentiation along service offerings, pricing, and digital capabilities. Key findings indicate that:

1. **Digital transformation is essential:** Online ordering availability is the strongest predictor of customer engagement
2. **Market segmentation is clear:** Mid-range pricing dominates, with distinct premium and budget segments
3. **Quality and visibility are separate:** High ratings don't guarantee high engagement; marketing and accessibility matter equally
4. **Service offerings drive positioning:** Table booking and online ordering are strategic tools for market differentiation

This analysis demonstrates the value of data-driven decision-making in the restaurant industry. Stakeholders can use these insights to optimize their positioning, improve digital accessibility, and better understand customer preferences.

Future analysis could incorporate temporal data to track trends, geographic expansion insights, and predictive modeling to forecast restaurant success based on these key factors[58][59][60].

References

[1] Restaurant analytics importance. *Tenzo Analytics Guide*, 2025. https://www.gotenzo.com

[2] Online ordering impact on restaurant performance. *BeP Back Office Blog*, 2025. https://bepbackoffice.com

[3] Restaurant data analytics foundations. *Clicdata*, 2024. https://www.clicdata.com

[4] Customer behavior metrics in restaurants. *Clicdata Restaurant Analytics Guide*, 2024. https://www.clicdata.com

[5] POS data analysis best practices. *RezKu*, 2025. https://rezku.com

[6] Zomato predictive modeling research. *JETIR*, 2021. https://www.jetir.org

[7] Data analytics impact on restaurant industry. *Noble Desktop*, 2020. https://www.nobledesktop.com

[8] Zomato dataset analysis on Kaggle. *Kaggle Datasets*, 2018. https://www.kaggle.com/datasets/shrutimehta/zomato-restaurants-data

Report Prepared By: Mantra Pandya
Tools Used: Python, Pandas, NumPy, Matplotlib, Seaborn
Analysis Type: Exploratory Data Analysis (EDA)
Status: Completed