

Zomato Sales Data Analysis Project

Summary

Project Overview & Methodology

The project is a comprehensive Exploratory Data Analysis (EDA) of a Zomato dataset consisting of **148 restaurant listings** across 7 key parameters: name, online order status, table booking availability, rating, votes, average cost, and service type.

Data Processing Steps:

- **Cleaning:** The most critical transformation was converting the string-based `rate` column (e.g., "4.1/5") into a numerical float format (e.g., 4.1) using a custom `handleRate` function.
 - **Validation:** The notebook performed a null-value check confirming **no missing data** was present, ensuring the reliability of the statistical findings.
-

2. In-Depth Question Analysis & Visualizations

A. Market Composition: Restaurant Types

- **Question:** What type of restaurant dominates the market?
- **Analysis:** The project uses a `sns.countplot` on the `listed_in(type)` column.
- **Visualization Explanation:** A bar chart identifies the frequency of each category (e.g., Buffet, Cafes, Dining). This reveals which business models are most prevalent in the current food ecosystem.

B. Customer Engagement: Popularity by Service

- **Question:** Which restaurant types receive the most consumer interaction (votes)?
- **Analysis:** The data is grouped by restaurant type, and the `votes` column is summed.
- **Visualization Explanation:** A bar plot shows that while some types might be fewer in number, they might attract a significantly higher volume of votes, indicating higher customer engagement or "foodie" interest.

C. Financial Landscape: Cost Distribution

- **Question:** What is the average pricing for a meal for two?

- **Analysis:** The `describe()` function and likely distribution plots were used to analyze `approx_cost(for two people)`.
- **Key Finding:** The average cost is approximately **418 units**, with the majority of entries falling between 200 and 600.

D. Strategic Correlation: Heatmap Analysis

- **Question:** How does the service type relate to online ordering habits?
 - **Analysis:** A **Pivot Table** was created cross-referencing `listed_in(type)` with `online_order` status.
 - **Visualization Explanation:**
 - **The Heatmap:** Using `sns.heatmap` with the '`YlGnBu`' color map, the code visually highlights "hot zones".
 - **In-Depth Insight:** This helps identify market gaps—for instance, if "Dining" restaurants have low "Online Order" numbers, it signifies an untapped delivery potential or a strictly traditional dining model.
-

3. Findings & Observations

- **Rating Health:** The ratings are generally healthy, though the range varies significantly across different price points.
 - **Cost vs. Type:** There is a clear link between the "Type" of restaurant and its "Approximate Cost," with Buffet and Fine Dining naturally occupying the higher cost brackets.
 - **Online Presence:** A significant portion of the dataset indicates "Yes" for online ordering, reflecting a shift toward digital food delivery.
-

4. Strategic Recommendations

1. **Optimization for Delivery:** For restaurant types currently lacking "Online Order" options (as identified in the heatmap), there is a significant opportunity to increase revenue by integrating with the Zomato delivery platform.
2. **Price-Point Targeting:** Since the mean cost is ~418, new entrants should target the **200–500 price range** to appeal to the largest segment of the customer base.
3. **Customer Retention via Engagement:** Restaurants with low "Votes" but high "Ratings" should implement "Review & Vote" campaigns. Increasing the vote count improves visibility on the Zomato algorithm, even if the rating stays the same.
4. **Service Diversification:** "Dining" type restaurants should consider adding "Cafe" or "Buffet" elements if the data shows those types have higher engagement (votes) in their specific price bracket.

