Executive Summary

Zomato Data Analysis

This project aims to derive actionable insights from a real-world Zomato dataset by applying data cleaning techniques, exploratory data analysis (EDA), and effective data visualization. The primary focus is to understand user preferences, restaurant performance, cost dynamics, and digital behavior patterns within the food and hospitality sector. The insights drawn have implications for business strategy, customer engagement, and digital transformation within the restaurant industry.

---

1. Data Cleaning and Preprocessing

- The dataset underwent rigorous cleaning, especially the transformation of the **'rate'** column, originally in the format "X.X/5". It was parsed and converted into a **numeric float** for quantitative analysis.

- Missing and inconsistent values in the `rate` and `approx_cost(for two people)` columns were addressed to ensure clean, structured, and analysis-ready data.

- This step laid the foundation for accurate and meaningful visualization and interpretation.

---

2. Restaurant Type Distribution

- Restaurants were categorized by their type (e.g., Dine-in, Café, Quick Bites).

- **Key Insight**: **Dining-type restaurants dominated the dataset**, accounting for the **largest share**, followed by Cafés and Quick Bites.

- **Data Insight**: Dining establishments formed approximately **45–50%** of all restaurants analyzed.

- **Business Implication**: The popularity of dining restaurants suggests a focus on experience-based dining, which could be leveraged by service-enhancing strategies like ambiance curation or loyalty programs.

---

3. Votes and Engagement

- Total customer votes were aggregated by restaurant type to assess user engagement.
- **Observation**: **Dining restaurants received the highest cumulative number of votes**, indicating strong customer interaction and possibly higher footfall.
- **Percentage Insight**: Dining restaurants received nearly **50%–55% of all customer votes**, reinforcing their dominance.
- **Interpretation**: High voting could be a reflection of better service, more visibility, or popularity on the platform.

---

4. Rating Distribution

- The analysis of customer ratings revealed:
    - The majority of ratings fall between **3.5 and 4.0**.
    - Very few restaurants received ratings below 2.5 or above 4.5.
- **Percentage Breakdown**:
    - **~35–40%** of restaurants had ratings between **3.5 to 4.0**.
    - **~25–30%** were rated between **4.0 and 4.5**.
- **Conclusion**: The rating distribution is **narrow and competitive**, emphasizing the importance for restaurants to maintain service consistency to remain relevant.

---

5. Customer Spending Patterns

- The `approx_cost(for two people)` column was analyzed to understand typical spending behavior.
- **Finding**: A substantial portion of customers prefer mid-range budget options.
- **Cost Insight**:

    - **~40–45%** of restaurants had an average cost for two between **₹200 and ₹300**.
    - Higher-end dining experiences (₹800 and above) formed a much smaller share, less than **10%**.
- **Business Insight**: Most users prefer affordable options. Restaurants targeting the budget-conscious crowd may see higher traffic and engagement.

---

6. Online vs Offline Ordering Behavior

- Ratings were compared based on whether restaurants offered **online orders**.
- **Insight**: Restaurants that supported **online ordering received higher average ratings**.
    - **Online Order Rating Average**: **~4.0**
    - **Offline Order Rating Average**: **~3.6**
- **Conclusion**: Enhanced customer experience, convenience, and faster service in online delivery could contribute to this trend.
- **Actionable Strategy**: Restaurants can improve their ratings by integrating efficient online delivery platforms and offering promotions through them.

---

7. Heatmap Analysis: Restaurant Type vs Order Mode

- A heatmap was generated to show the relationship between restaurant types and preferred order modes (online vs offline).
- **Findings**:
    - **Dining establishments** predominantly receive **offline orders**.
    - **Cafés and Quick Bites** tend to get **more online orders**.
- **Implication**:
    - Casual food outlets are adapting to digital trends better and cater to a younger, mobile-savvy audience.

- Traditional dining setups still rely on walk-in customers and could benefit from digitization.

Conclusion & Recommendations

This Zomato dataset analysis has uncovered distinct behavioral patterns:

1. **Dining restaurants** are the most preferred category and receive the highest customer engagement.

2. **Mid-range price points** (₹200–₹300) dominate customer preferences, indicating a value-driven audience.

3. **Customer ratings are heavily concentrated** around the 3.5–4.0 mark, leaving little room for differentiation unless restaurants go beyond expectations.

4. **Online ordering significantly correlates with higher customer satisfaction**, emphasizing the need for restaurants to invest in digital platforms.

5. There's a **clear segmentation of restaurant types** by order preferences—dining for offline, cafés for online.

Future Scope

- Integrating sentiment analysis from user reviews could provide even deeper insights.

- Predictive modeling (e.g., predicting ratings or churn) using machine learning would add a predictive layer to this analysis.

- Further segmentation based on location or cuisine could offer hyper-targeted business strategies.
