Zomato data analysis project

Introduction

This project analyzes Bangalore's Zomato restaurant data to uncover insights about customer preferences, restaurant popularity, online ordering trends, and spending behaviors. The dataset includes structured details per restaurant: name, online order availability, table booking, rating, votes, estimated cost, and category.Zomato-data.xlsx

Objectives:

- Understand customer preferences and behavior.

- Compare online vs. offline restaurant engagement.

- Help Zomato and restaurateurs make informed decisions.

Data Description

The provided dataset ("Zomato-data.xlsx") consists of:

Column	Description
name	Restaurant name
online_order	Availability of online ordering ('Yes'/'No')
book_table	Table booking availability ('Yes'/'No')
rate	Restaurant rating (string, e.g., '4.1/5')
votes	Number of customer votes
approx_cost(for two people)	Estimated cost for two people (numeric)
listed_in(type)	Restaurant category (Dining, Cafes, Buffet, other)

Data was reviewed for duplicate entries, incorrect types, and missing values, followed by necessary cleaning.

Methodology

1. Environment Setup

- Installed Python (3.x), Jupyter Notebook.

- Installed necessary libraries: Pandas, NumPy, Matplotlib, Seaborn.

2. Data Loading

- Imported dataset into a Pandas DataFrame.

- Displayed the first rows to verify loading and columns.

3. Data Cleaning

- **Converted Ratings**: Split the 'rate' column to extract the numeric value and convert to float.

- **Checked Data Types**: Used DataFrame methods to confirm correct typing for each column.

- **Null/Missing Values**: Confirmed no missing values remain after data type adjustments.

4. Exploratory Data Analysis & Visualization

Analysis 1: Restaurant Type Distribution

- Used a count plot to visualize the frequency of each restaurant type.

- **Finding**: Dining was notably the most common restaurant category.

Analysis 2: Customer Votes by Type

- Grouped by 'listed_in(type)', summed up votes, and presented results as a bar chart.

- **Finding**: Dining restaurants attracted the highest vote count, indicating popularity.

Analysis 3: Ratings Distribution

- Plotted histogram of the ratings column to visually assess the spread.

- **Finding**: Most restaurants received ratings in the range of 3.5 to 4.0.

Analysis 4: Spending Patterns for Couples

- Used count plot (bar graph) for 'approx_cost(for two people)' to see typical couple's spending.

- **Finding**: The most frequent spending range for couples was between 300 and 800 (currency not specified).

Analysis 5: Online vs Offline Ratings

- Created box plot comparing 'rate' between restaurants with and without online ordering.

- **Finding**: Restaurants offering online orders had slightly higher median ratings.

Analysis 6: Online Orders by Restaurant Type

- Created a pivot table showing counts of online/offline order availability per type.

- Visualized with a heatmap.

- **Finding**: Cafes and "other" types strongly favored online orders, Dining skewed towards offline.

Key Analytical Questions Addressed

1. **What type of restaurant do the majority of customers order from?**

   o Dining restaurants are the most common and most ordered from.

2. **How many votes has each type of restaurant received from customers?**

   o 'Dining' leads in cumulative votes, followed by Cafes and Buffet.

3. **What ratings do the majority of restaurants receive?**

   o Ratings cluster from 3.5 to 4.0, denoting moderate customer satisfaction.

4. **Zomato observed most couples order food online. What is their average spending?**

   o Majority spend between 300 and 800 (currency units) for two people.

5. **Which mode (online or offline) receives the maximum rating?**

   o Online-ordering restaurants tend to receive slightly higher ratings.

6. **Which type received more offline orders for special offers targeting?**

   o Dining type received more offline orders, suggesting this segment for targeted offers.

Insights & Business Recommendations

- **Dining popularity**: Dining restaurants are most frequented. More resources and offers can be targeted here.

- **Online convenience**: Online ordering correlates with slightly better customer ratings. Improving online capabilities is beneficial.

- **Spending range focus**: Offers tailored for 300-800 currency range may be attractive for couples.

- **Votes illustrate engagement**: Tracking votes and reviews alongside type helps measure engagement and satisfaction.

- **Offline Dining offers**: Offline-oriented Dining could benefit from exclusive promotions to incentivize digital engagement.

Visualizations Summary

- **Countplot**: Restaurant types distribution with distinct colors.

- **Barplot**: Total votes per restaurant type.

- **Histogram**: Ratings distribution among restaurants.

- **Countplot**: Spending patterns among couples ordering.

- **Boxplot**: Ratings comparison for online vs offline ordering.

- **Heatmap**: Online ordering spread by restaurant type.

Conclusion

This analysis presents a comprehensive view of restaurant dynamics in Bangalore's Zomato ecosystem. Dining restaurants dominate both in presence and customer engagement. Online ordering is growing in popularity, offering a route for enhanced customer satisfaction. Spending data guides targeted promotional efforts, especially for couples and offline Dining restaurants.

Business stakeholders and Zomato can leverage these findings for strategic campaigns, operational improvements, and improved product offerings. The methodology provides a framework for ongoing restaurant analysis using Python and open-source libraries.