Zomato Data Analysis Project Report

1. Business Problem Statement

The aim of this project is to analyze the Zomato dataset to derive insights into customer preferences regarding restaurant types, order methods, spending habits, and ratings. By understanding these trends, Zomato can enhance its offerings, optimize marketing strategies, and improve customer satisfaction.

2. Data Preparation

2.1. Installing Necessary Libraries

The following Python libraries were used for data manipulation and visualization:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

2.2. Loading the Dataset

The dataset was loaded from a CSV file into a pandas DataFrame:

```
dataframe = pd.read_csv("Zomato data.csv")
```

2.3. Data Cleaning

The rate column, originally in the format "X/Y", was converted to a float representing only the rating:

```
def correctRate(value):
    value = str(value).split("/")
    return float(value[0])
dataframe['rate'] = dataframe['rate'].apply(correctRate)
```

After cleaning, the dataset contained no missing values, ensuring reliable analysis.

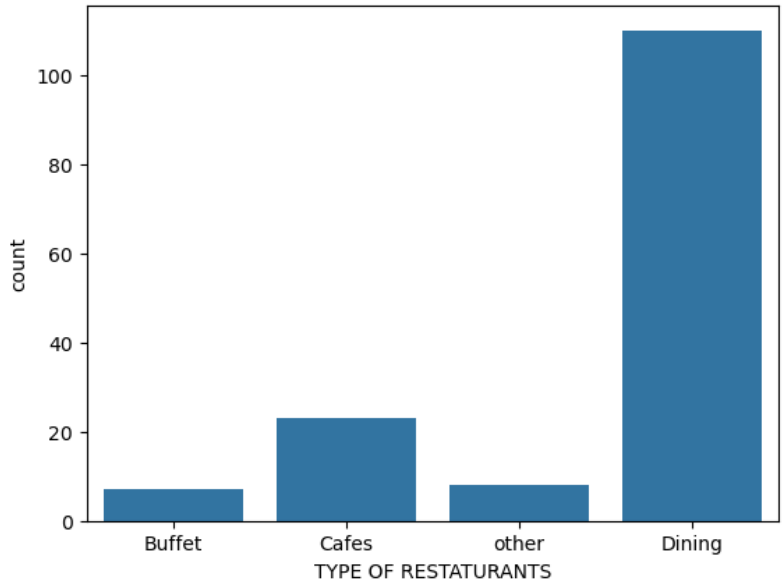
3. Exploratory Data Analysis

3.1. Restaurant Types

Question: What type of restaurants do the majority of customers order from?

Visualization:

```
sns.countplot(x=dataframe['listed_in(type)'])
plt.xlabel("TYPE OF RESTAURANTS")
plt.show()
```



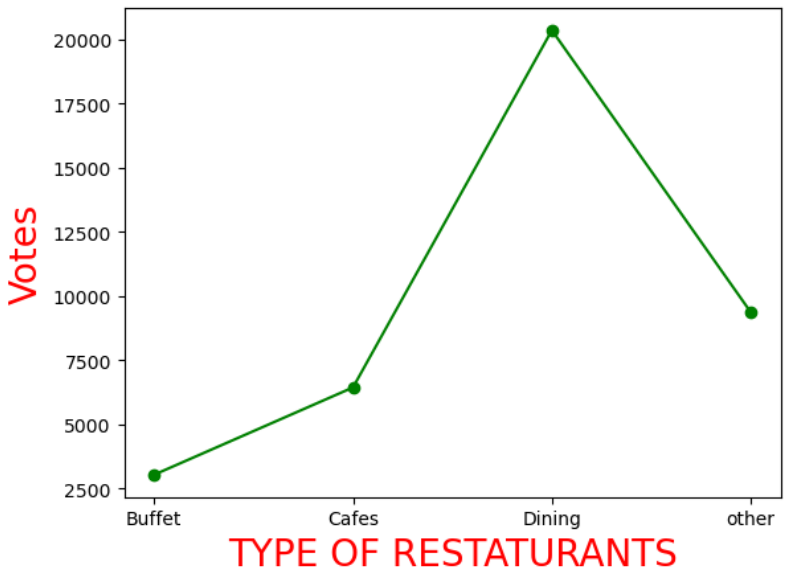
Findings: The analysis revealed that the majority of customers order from **Dining** type restaurants. This indicates a significant preference for dining experiences over other categories.

3.2. Customer Votes

Question: How many votes has each type of restaurant received from customers?

Visualization:

```
grouped_data = dataframe.groupby('listed_in(type)')['votes'].sum()
result = pd.DataFrame({'votes': grouped_data})
plt.plot(result, c="green", marker="o")
plt.xlabel("TYPE OF RESTAURANTS", c="red", size=20)
plt.ylabel("Votes", c="red", size=20)
plt.show()
```



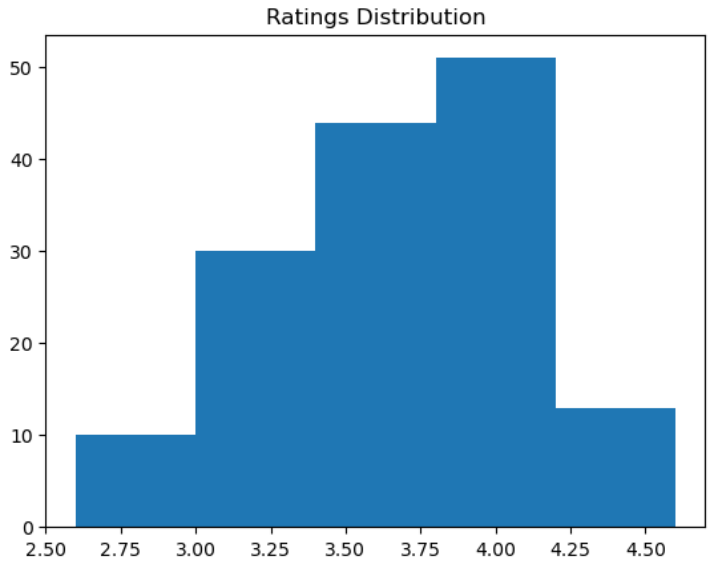
Findings: The **Dining** restaurants received the maximum number of votes, suggesting that they are more popular among customers and may benefit from enhanced marketing and promotions.

3.3. Ratings Distribution

Question: What are the ratings majority of restaurants have received?

Visualization:

```
plt.hist(dataframe['rate'], bins=100)
plt.title("Ratings Distribution")
plt.show()
```



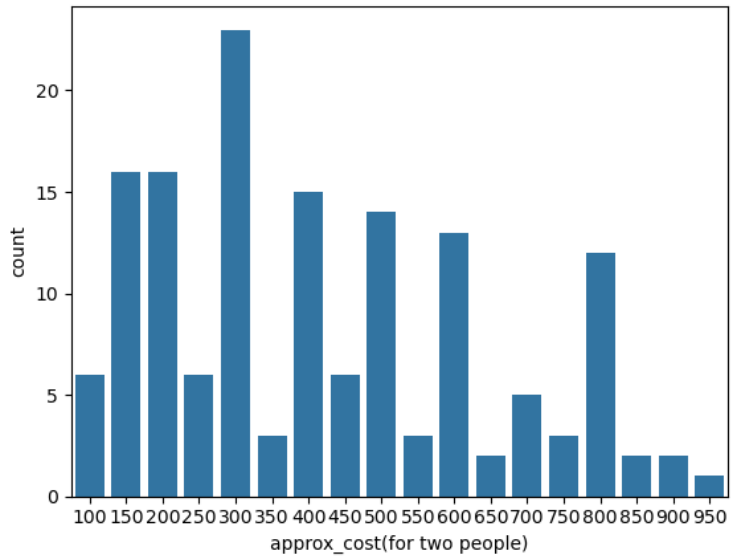
Findings: The histogram indicates that ratings are predominantly concentrated between **3.5** and **4.0**, with a peak around **4.1**. This suggests that most customers are satisfied with their dining experiences, but there is still room for improvement.

3.4. Average Spending

Question: What is the average spending on each order?

Visualization:

```
cdata = dataframe['approx_cost(for two people)']
sns.countplot(x=cdata)
plt.show()
```



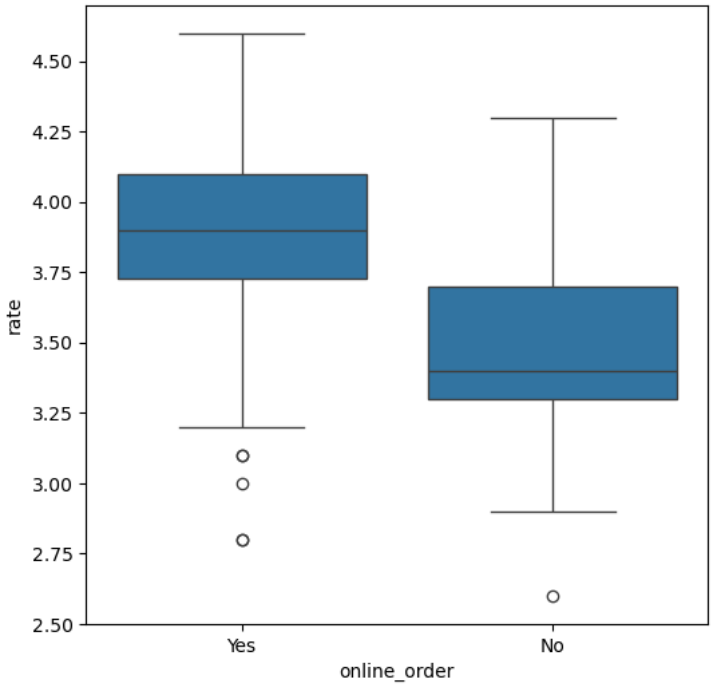
Findings: The analysis shows that customers generally prefer to dine at restaurants with an approximate budget of around **300** for two people. This information can help Zomato tailor its marketing strategies to attract budget-conscious customers.

3.5. Ratings by Order Mode

Question: Which mode (online or offline) received the most ratings?

Visualization:

```
plt.figure(figsize=(6, 6))
sns.boxplot(x='online_order', y='rate', data=dataframe)
plt.show()
```



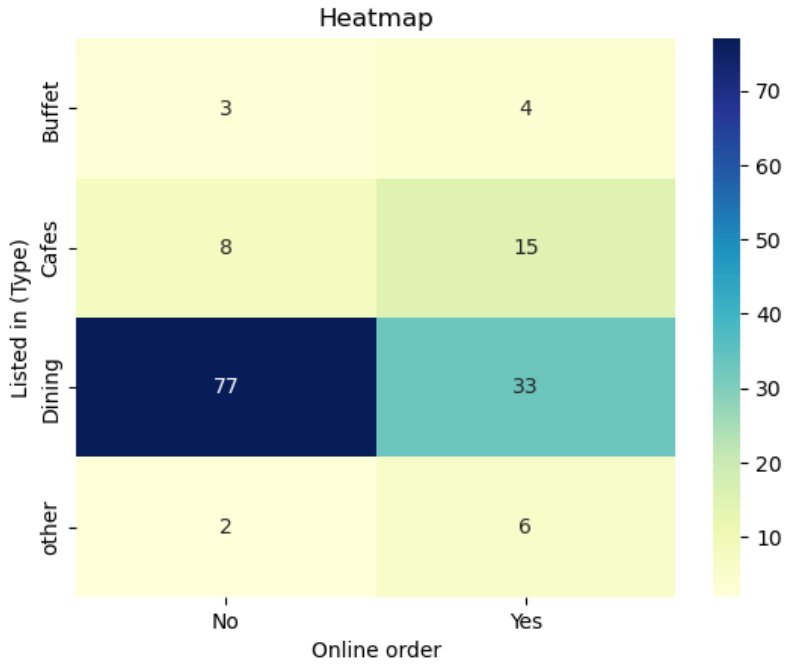
Findings: The box plot indicates that online orders tend to receive better ratings compared to offline orders. This insight can guide Zomato to focus on enhancing the online ordering experience, potentially leading to higher customer satisfaction and retention.

3.6. Offline Orders by Restaurant Type

Question: Which type of restaurant received more offline orders?

Visualization:

```
pivot_table = dataframe.pivot_table(index='listed_in(type)', columns='online_order', aggfunc='size', fill_value=0)
sns.heatmap(pivot_table, annot=True, cmap="YlGnBu", fmt='d')
plt.title("Heatmap")
plt.xlabel("Online order")
plt.ylabel("Listed in (Type)")
plt.show()
```



Findings: The heatmap shows that **Dining** restaurants receive a significant number of offline orders, suggesting that Zomato should enhance its online offerings for this category to capture the offline customer base.

4. Conclusions and Suggestions

Conclusions

1. The majority of customers prefer **Dining** restaurants, which also receive the highest number of votes.
2. Ratings are generally positive, with most restaurants scoring between **3.5** and **4.1**.
3. Customers prefer to spend around **300** for two people, indicating a budget-friendly market segment.
4. Online orders tend to receive better ratings compared to offline orders, highlighting an opportunity for improving the online experience.
5. **Dining** restaurants dominate offline orders, indicating a potential market for online enhancements.

Suggestions

1. **Marketing Strategies:** Focus on promoting Dining restaurants through targeted campaigns, highlighting customer favorites and high-rated options.
2. **Online Experience:** Invest in enhancing the online ordering experience, as it correlates with higher ratings and customer satisfaction.
3. **Budget-Friendly Offers:** Introduce special offers or discounts for restaurants within the budget range of **300** to attract more customers.
4. **Feedback Mechanisms:** Implement feedback collection for both online and offline orders to understand customer preferences better and improve service quality.