Zomato Data Analysis Project

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

```python
data=pd.read_csv('Zomato data .csv',encoding='latin1')
```

```python
data
```

	name	online_order	book_table	rate	votes	approx_cost(for two people)	listed_in(type)
0	Jalsa	Yes	Yes	4.1	775	800	Buffet
1	Spice Elephant	Yes	No	4.1	787	800	Buffet
2	San Churro Cafe	Yes	No	3.8	918	800	Buffet
3	Addhuri Udupi Bhojana	No	No	3.7	88	300	Buffet
4	Grand Village	No	No	3.8	166	600	Buffet
...	...	...	...	...	...	...	...
143	Melting Melodies	No	No	3.3	0	100	Dining
144	New Indraprasta	No	No	3.3	0	150	Dining
145	Anna Kuteera	Yes	No	4.0	771	450	Dining
146	Darbar	No	No	3.0	98	800	Dining
147	Vijayalakshmi	Yes	No	3.9	47	200	Dining

148 rows × 7 columns

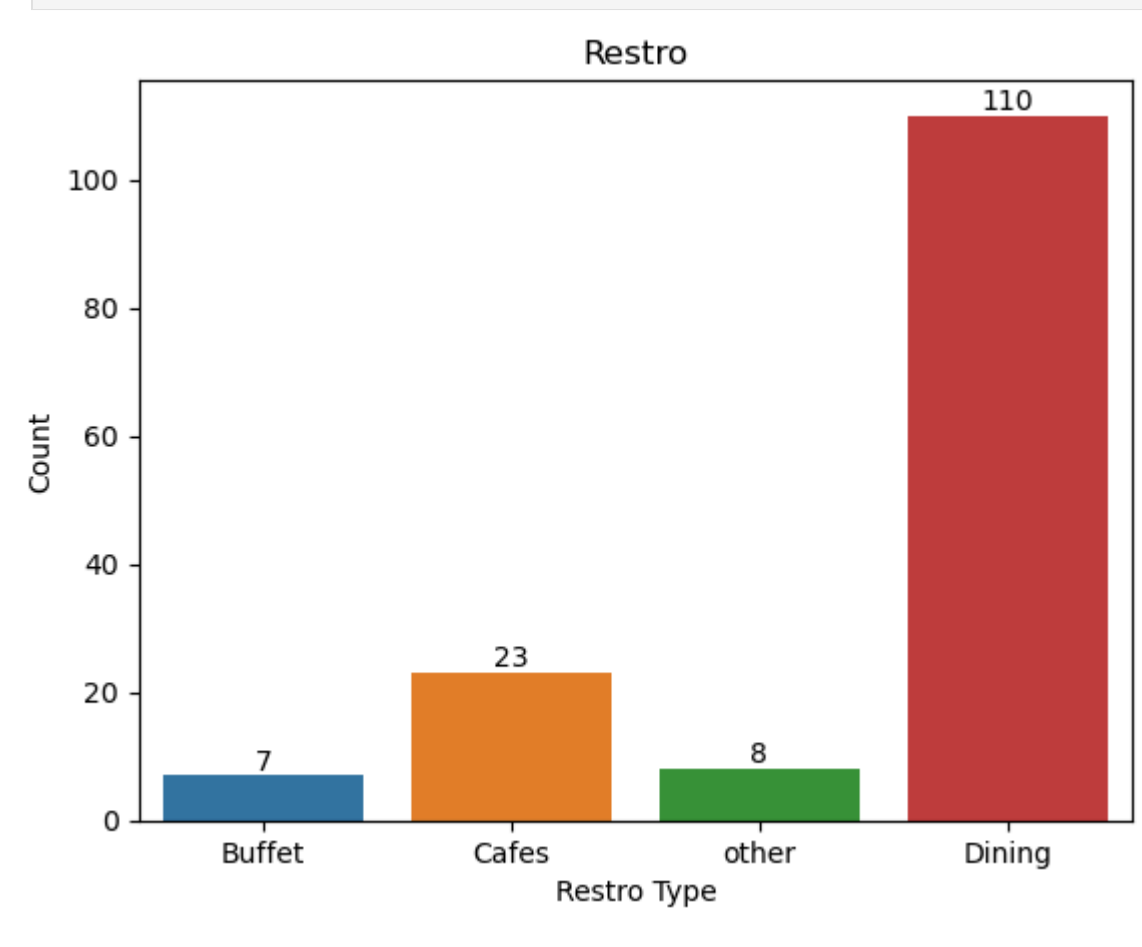
data['rate']=data['rate'].str.split('/').str[0].astype(float)

```python
data.info()
```

```python
data.isnull().sum()
```

```python
data['name'].unique()
```

```python
ax=sns.countplot(x='listed_in(type)',data=data,hue='listed_in(type)')
ax
plt.title('Restro')
plt.ylabel('Count')
plt.xlabel('Restro Type')
for i in ax.containers:
    ax.bar_label(i)
plt.show()
```

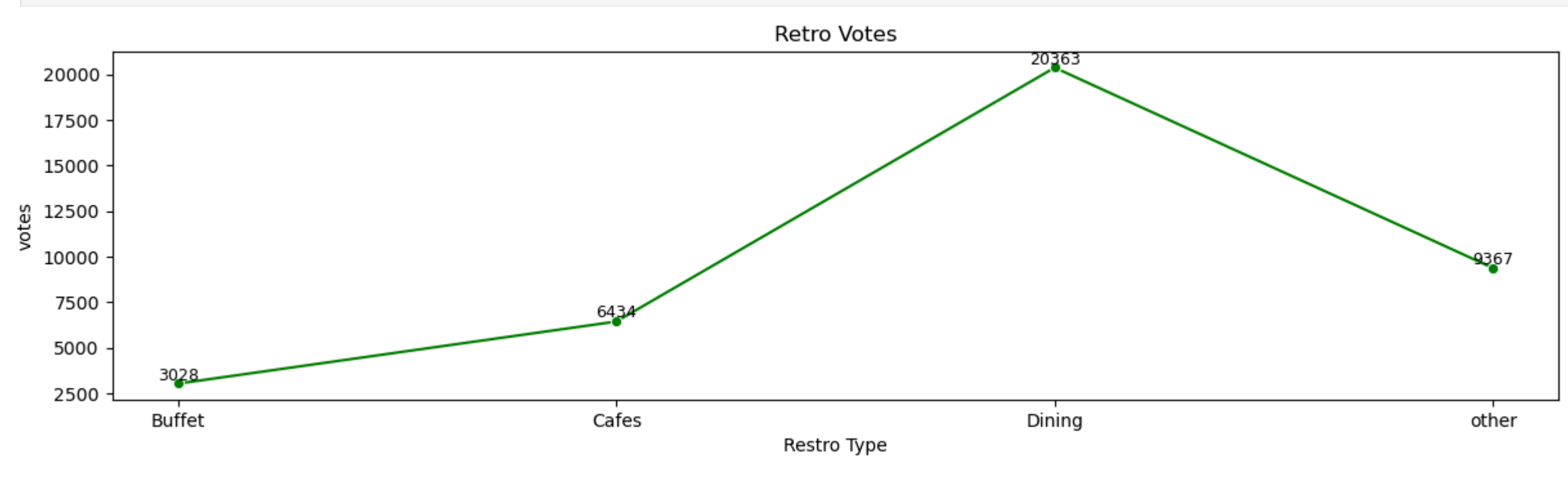


conclusion - majority of customer are falls in Dinning category

```python
restro_votes=data.groupby('listed_in(type)')['votes'].sum().reset_index()
restro_votes
```

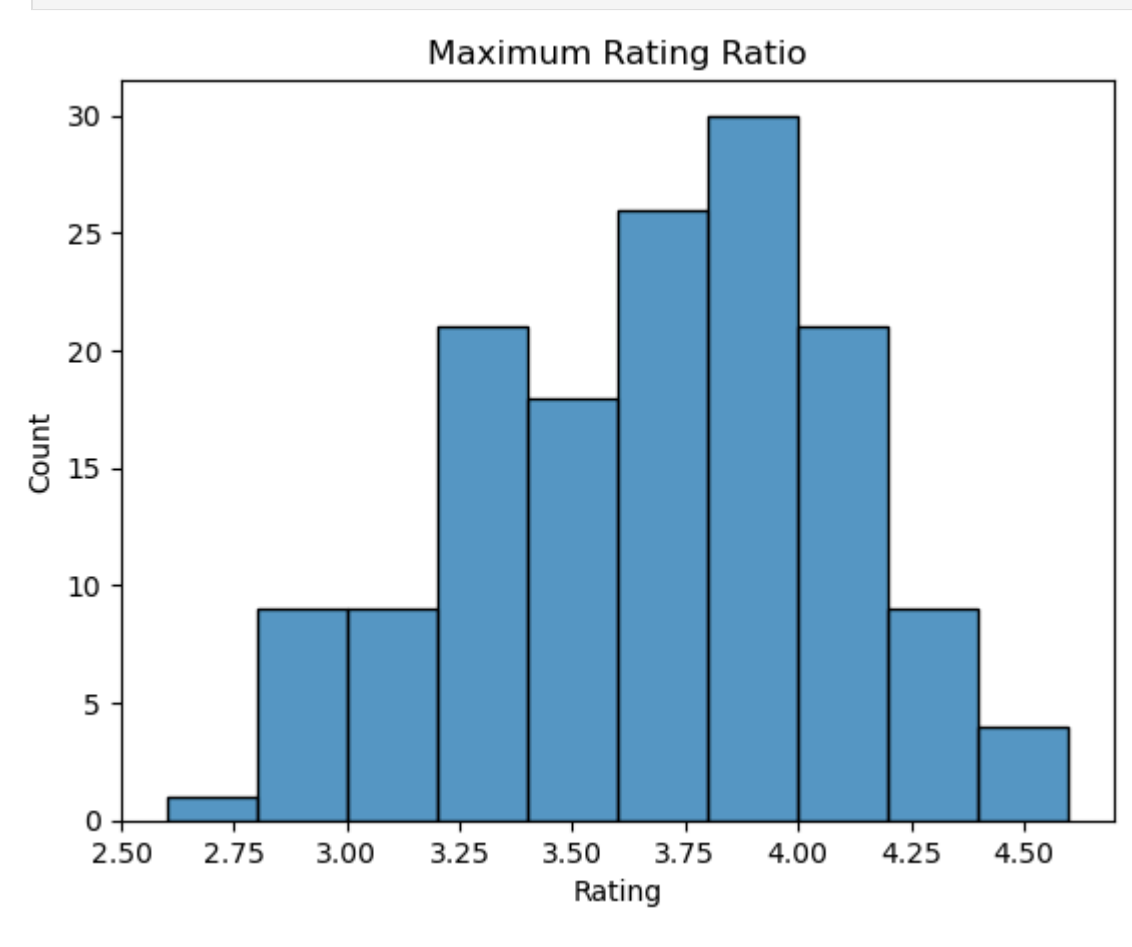
	listed_in(type)	votes
0	Buffet	3028
1	Cafes	6434
2	Dining	20363
3	other	9367

```python
plt.figure(figsize=(14.5, 3.5))

sns.lineplot(x='listed_in(type)',y='votes',data=restro_votes,marker='o',color='g')
plt.xlabel('Restro Type')
plt.title('Retro Votes')

for x, y in zip(restro_votes['listed_in(type)'], restro_votes['votes']):
    plt.text(x, y, str(y), ha='center', va='bottom', fontsize=9)
plt.show()
```



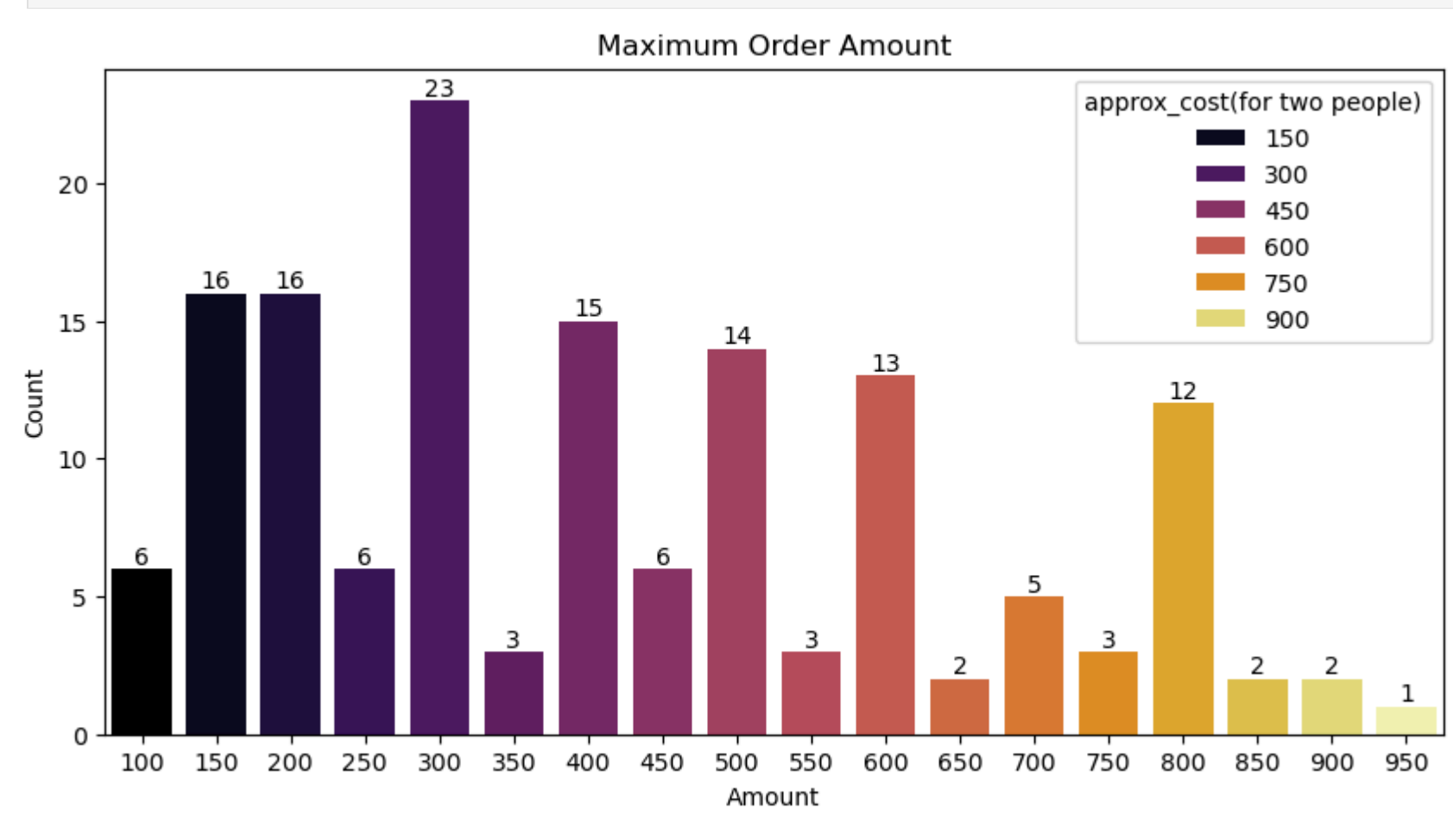
conclusion - Dining restro has maximum votes

```python
sns.histplot(data['rate'],bins=10)
plt.title("Maximum Rating Ratio")
plt.xlabel('Rating')
plt.ylabel('Count')
plt.show()
```



conclusion - maximum rating falls in between 3.5 to 4

```python
plt.figure(figsize=(10,5))
ax=sns.countplot(x='approx_cost(for two people)',data=data,hue='approx_cost(for two people)',palette='inferno')
ax
plt.title('Maximum Order Amount')
plt.ylabel('Count')
plt.xlabel('Amount')
for i in ax.containers:
    ax.bar_label(i)
plt.show()
```



conclusion - couples prefer order to most of 300 amount

```python
sns.boxplot(x='online_order',y='rate',data=data)
plt.title('Online order & Rating')
plt.show()
```

conclusion- Online order has maximum rating than offline order

```python
heatmap_data = data.pivot_table(index='listed_in(type)', columns='online_order', aggfunc='size', fill_value=0)
sns.heatmap(heatmap_data, annot=True, fmt="d", cmap="YlGnBu")
plt.title("Heatmap of Online Order and Restro Type")
plt.xlabel("Online Order")
plt.ylabel("Restro Type")
plt.show()
```