# zometo data analysis project

## importing libraries

In [3]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [4]:
```
dataframe = pd.read_csv("zomato_data.csv")
print(dataframe)
```

```
                    name online_order book_table   rate  votes  \
0                  Jalsa          Yes        Yes  4.1/5    775
1         Spice Elephant          Yes         No  4.1/5    787
2        San Churro Cafe          Yes         No  3.8/5    918
3    Addhuri Udupi Bhojana           No         No  3.7/5     88
4           Grand Village            No         No  3.8/5    166
..                    ...          ...        ...    ...    ...
143      Melting Melodies            No         No  3.3/5      0
144       New Indraprasta            No         No  3.3/5      0
145          Anna Kuteera          Yes         No  4.0/5    771
146                Darbar            No         No  3.0/5     98
147         Vijayalakshmi          Yes         No  3.9/5     47

     approx_cost(for two people) listed_in(type)
0                            800          Buffet
1                            800          Buffet
2                            800          Buffet
3                            300          Buffet
4                            600          Buffet
..                           ...             ...
143                          100          Dining
144                          150          Dining
145                          450          Dining
146                          800          Dining
147                          200          Dining

[148 rows x 7 columns]
```

In [5]:
```
dataframe
```

Out[5]:

| | name | online_order | book_table | rate | votes | approx_cost(for two people) | listed_in(type) |
|---|---|---|---|---|---|---|---|
| **0** | Jalsa | Yes | Yes | 4.1/5 | 775 | 800 | Buffet |
| **1** | Spice Elephant | Yes | No | 4.1/5 | 787 | 800 | Buffet |
| **2** | San Churro Cafe | Yes | No | 3.8/5 | 918 | 800 | Buffet |
| **3** | Addhuri Udupi Bhojana | No | No | 3.7/5 | 88 | 300 | Buffet |
| **4** | Grand Village | No | No | 3.8/5 | 166 | 600 | Buffet |
| **...** | ... | ... | ... | ... | ... | ... | ... |
| **143** | Melting Melodies | No | No | 3.3/5 | 0 | 100 | Dining |
| **144** | New Indraprasta | No | No | 3.3/5 | 0 | 150 | Dining |
| **145** | Anna Kuteera | Yes | No | 4.0/5 | 771 | 450 | Dining |
| **146** | Darbar | No | No | 3.0/5 | 98 | 800 | Dining |
| **147** | Vijayalakshmi | Yes | No | 3.9/5 | 47 | 200 | Dining |

148 rows × 7 columns

In [8]:
```
def handelRate(value):
    value = str(value).split('/')
    value = value[0];
    return float(value)

dataframe['rate'] = dataframe['rate'].apply(handelRate)
print(dataframe.head())
```

```
                  name online_order book_table  rate  votes  \
0                Jalsa          Yes        Yes   4.1    775
1         Spice Elephant          Yes         No   4.1    787
2        San Churro Cafe          Yes         No   3.8    918
3  Addhuri Udupi Bhojana           No         No   3.7     88
4          Grand Village           No         No   3.8    166

   approx_cost(for two people) listed_in(type)
0                           800          Buffet
1                           800          Buffet
2                           800          Buffet
3                           300          Buffet
4                           600          Buffet
```

In [6]:
```
dataframe.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 148 entries, 0 to 147
Data columns (total 7 columns):
 #   Column                       Non-Null Count  Dtype 
---  ------                       --------------  ----- 
 0   name                         148 non-null    object
 1   online_order                 148 non-null    object
 2   book_table                   148 non-null    object
 3   rate                         148 non-null    object
 4   votes                        148 non-null    int64 
 5   approx_cost(for two people)  148 non-null    int64 
 6   listed_in(type)              148 non-null    object
dtypes: int64(2), object(5)
memory usage: 8.2+ KB
```

# type of resturant

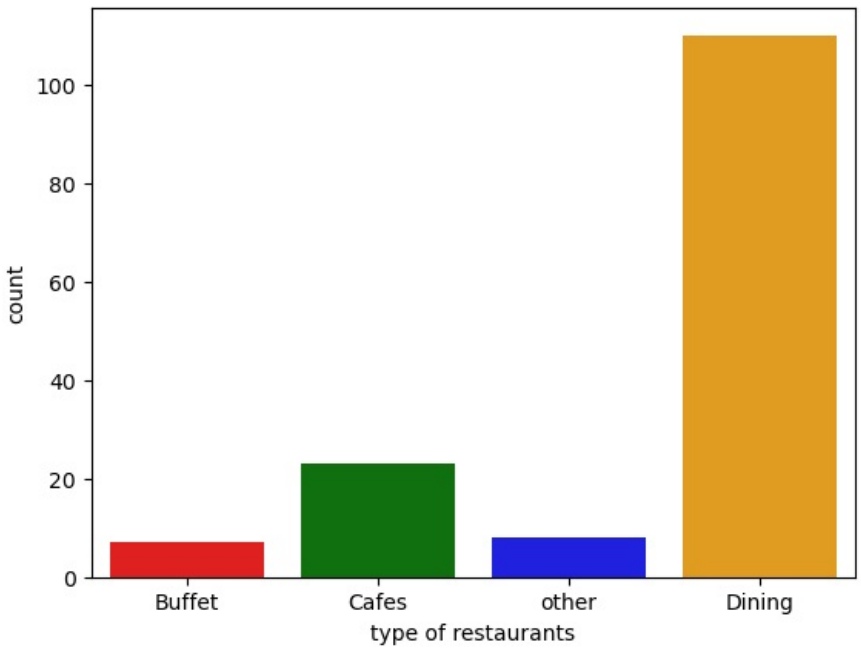
In [7]:
```
dataframe.head()
```

Out[7]:

| | name | online_order | book_table | rate | votes | approx_cost(for two people) | listed_in(type) |
|---|---|---|---|---|---|---|---|
| **0** | Jalsa | Yes | Yes | 4.1/5 | 775 | 800 | Buffet |
| **1** | Spice Elephant | Yes | No | 4.1/5 | 787 | 800 | Buffet |
| **2** | San Churro Cafe | Yes | No | 3.8/5 | 918 | 800 | Buffet |
| **3** | Addhuri Udupi Bhojana | No | No | 3.7/5 | 88 | 300 | Buffet |
| **4** | Grand Village | No | No | 3.8/5 | 166 | 600 | Buffet |

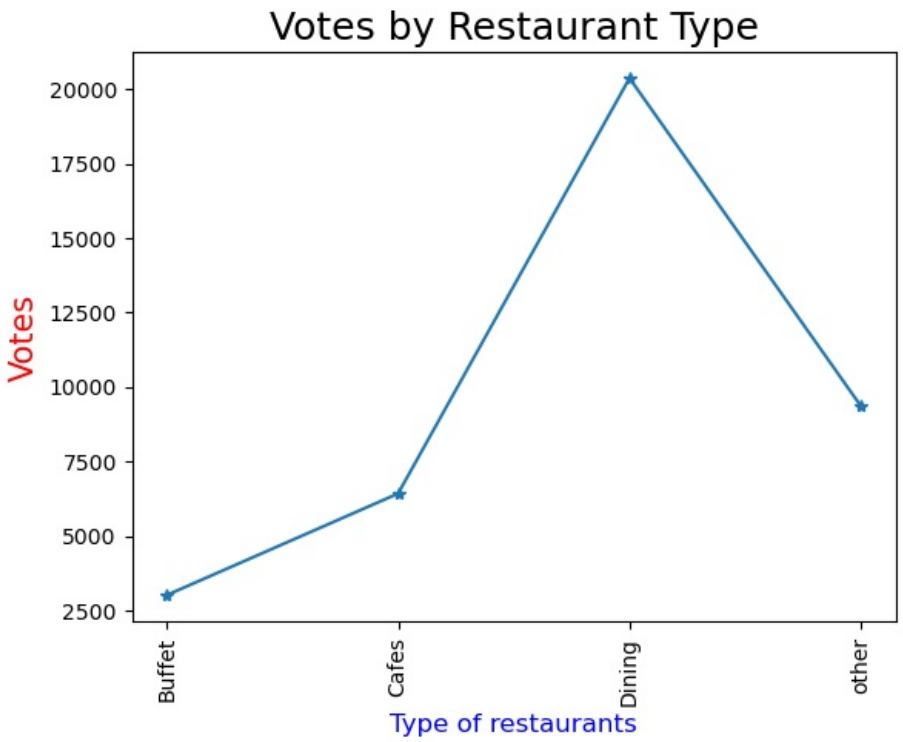
In [8]:
```
sns.countplot(
    x=dataframe['listed_in(type)'],
    palette=['red', 'green', 'blue', 'orange'],legend = False
)

plt.xlabel("type of restaurants")
plt.show()
```

```
C:\Users\ABC\AppData\Local\Temp\ipykernel_14224\1330362676.py:1: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable
to `hue` and set `legend=False` for the same effect.

  sns.countplot(
```

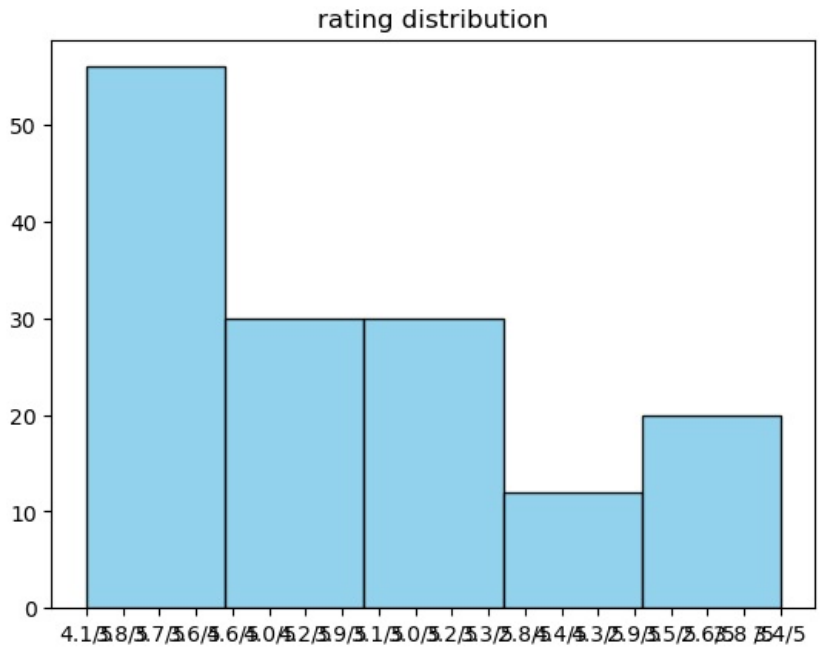



In [12]:

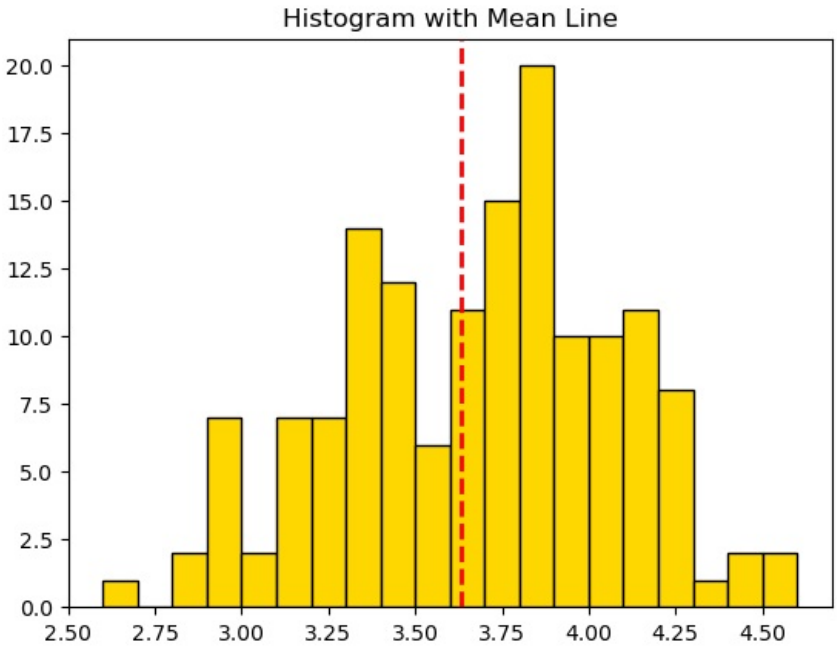
```
grouped_data = dataframe.groupby('listed_in(type)')['votes'].sum().reset_index()

plt.plot(grouped_data['listed_in(type)'], grouped_data['votes'],
         marker='*')

plt.xlabel("Type of restaurants", color="blue", fontsize=12)
plt.ylabel("Votes", color="red", fontsize=15)
plt.title("Votes by Restaurant Type", fontsize=18)

plt.xticks(rotation=90)
plt.grid(False)
plt.show()
```




In [9]:
```
plt.hist(dataframe['rate'],bins= 5,alpha= 0.9,color = 'skyblue',edgecolor ='black')
plt.title("rating distribution")
plt.show()
```

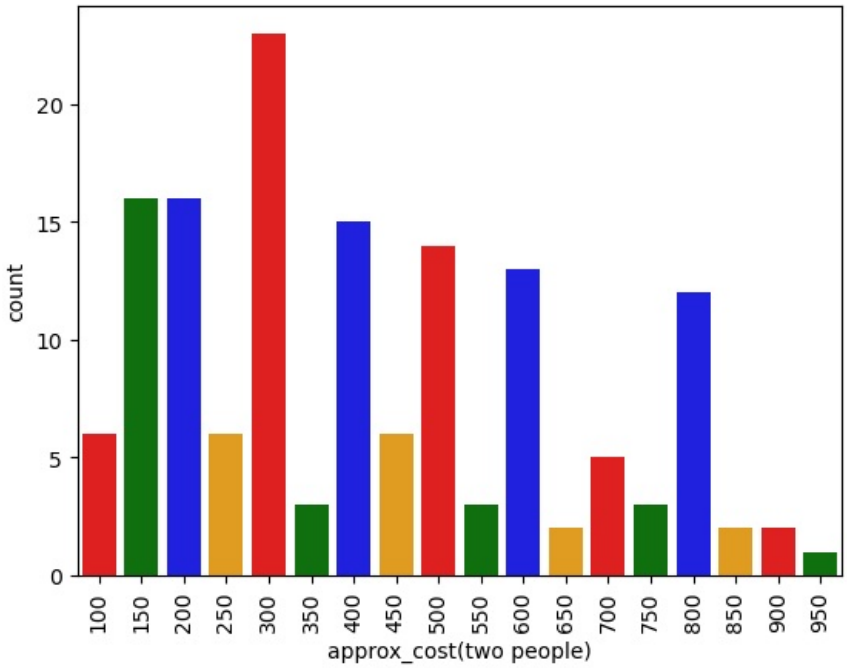



In [14]:
```
import numpy as np

plt.hist(dataframe['rate'], bins=20, color='gold', edgecolor='black')

plt.axvline(dataframe['rate'].mean(), color='red', linestyle='--', linewidth=2)
plt.title("Histogram with Mean Line")

plt.show()
```

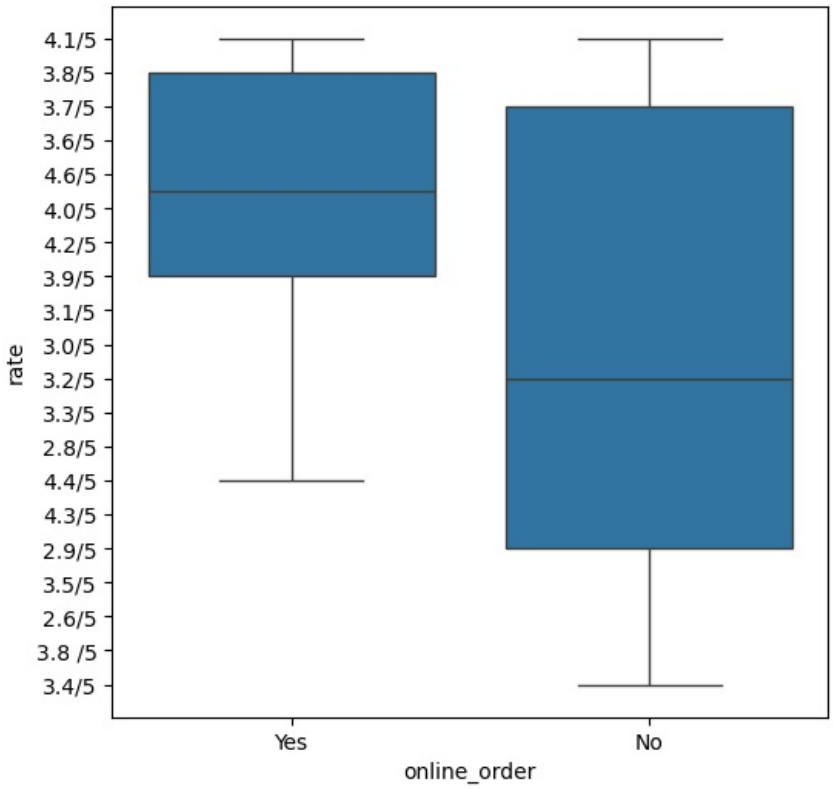



In [10]:
```
coupal_data = dataframe["approx_cost(for two people)"]
sns.countplot(x=coupal_data,
palette=['red', 'green', 'blue', 'orange'],legend = False)


plt.xlabel("approx_cost(two people)" )
plt.xticks(rotation=90)

plt.show()
```

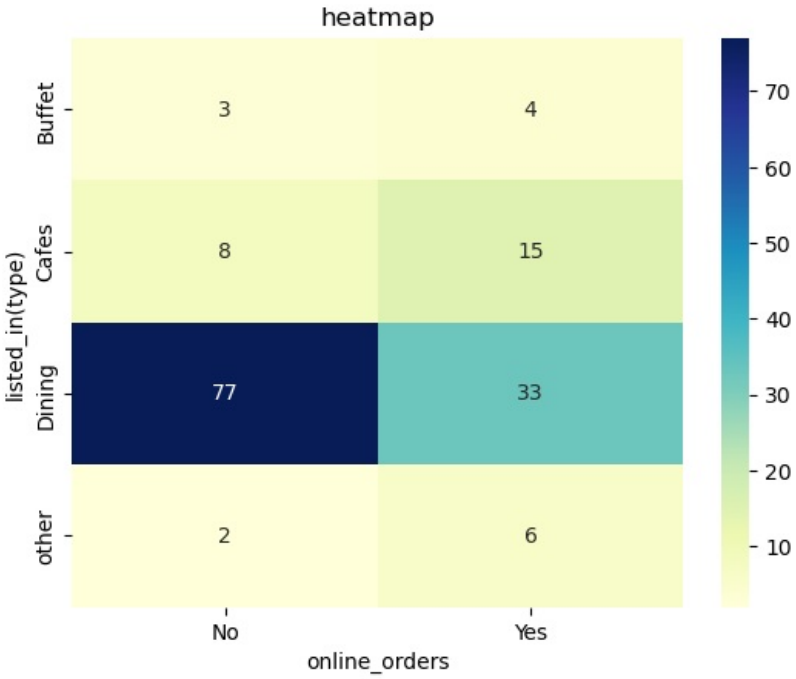
```
C:\Users\ABC\AppData\Local\Temp\ipykernel_14224\1147537393.py:2: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `x` variable
to `hue` and set `legend=False` for the same effect.

  sns.countplot(x=coupal_data,
C:\Users\ABC\AppData\Local\Temp\ipykernel_14224\1147537393.py:2: UserWarning:
The palette list has fewer values (4) than needed (18) and will cycle, which may produce an uninterpretable plot
.
  sns.countplot(x=coupal_data,
```




In [12]:
```
import matplotlib.pyplot as plt
import seaborn as sns
plt.figure(figsize =(6,6))
sns.boxplot(data = dataframe,x = "online_order", y= "rate")
plt.show()
```




In [13]:
```
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_orders")
plt.ylabel("listed_in(type)")
plt.show()
```




In [ ]:

In [ ]: