

In [24]:

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

In [25]:

```
df = pd.read_excel(r'C:\Users\NITU\Downloads\zomato.xlsx')
```

In [26]:

```
df.head()
```

Out[26]:

	url	address	name	online_order	book_table
0	https://www.zomato.com/bangalore/jalsa-banasha...	942, 21st Main Road, 2nd Stage, Banashankari, ...	Jalsa	Yes	Yes
1	https://www.zomato.com/bangalore/spice-elephan...	2nd Floor, 80 Feet Road, Near Big Bazaar, 6th ...	Spice Elephant	Yes	No
2	https://www.zomato.com/SanchurroBangalore?cont...	1112, Next to KIMS Medical College, 17th Cross...	San Churro Cafe	Yes	No
3	https://www.zomato.com/bangalore/addhuri-udupi...	1st Floor, Annakuteera, 3rd Stage, Banashankar...	Addhuri Udupi Bhojana	No	No
4	https://www.zomato.com/bangalore/grand-village...	10, 3rd Floor, Lakshmi Associates, Gandhi Baza...	Grand Village	No	No

In [27]:

```
df.shape
```

Out[27]:

```
(51717, 17)
```

In [28]:

df.columns

Out[28]:

```
Index(['url', 'address', 'name', 'online_order', 'book_table', 'rate', 'votes',
      'phone', 'location', 'rest_type', 'dish_liked', 'cuisines',
      'approx_cost(for two people)', 'reviews_list', 'menu_item',
      'listed_in(type)', 'listed_in(city)'],
      dtype='object')
```

In [29]:

```
df= df.drop(['url','address','phone','menu_item','dish_liked','reviews_list'],axis=1)
df.head()
```

Out[29]:

	name	online_order	book_table	rate	votes	location	rest_type	cuisines	approx_two
0	Jalsa	Yes	Yes	4.1/5	775	Banashankari	Casual Dining	North Indian, Mughlai, Chinese	
1	Spice Elephant	Yes	No	4.1/5	787	Banashankari	Casual Dining	Chinese, North Indian, Thai	
2	San Churro Cafe	Yes	No	3.8/5	918	Banashankari	Cafe, Casual Dining	Cafe, Mexican, Italian	
3	Addhuri Udupi Bhojana	No	No	3.7/5	88	Banashankari	Quick Bites	South Indian, North Indian	
4	Grand Village	No	No	3.8/5	166	Basavanagudi	Casual Dining	North Indian, Rajasthani	

In [30]:

df.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 51717 entries, 0 to 51716
Data columns (total 11 columns):
 #   Column                                Non-Null Count  Dtype
---  -
 0   name                                51717 non-null  object
 1   online_order                        51717 non-null  object
 2   book_table                          51717 non-null  object
 3   rate                                43942 non-null  object
 4   votes                              51717 non-null  int64
 5   location                           51696 non-null  object
 6   rest_type                          51490 non-null  object
 7   cuisines                           51672 non-null  object
 8   approx_cost(for two people)        51371 non-null  float64
 9   listed_in(type)                    51717 non-null  object
10   listed_in(city)                    51717 non-null  object
dtypes: float64(1), int64(1), object(9)
memory usage: 4.3+ MB
```

In [31]:

df['rate'].unique()

Out[31]:

```
array(['4.1/5', '3.8/5', '3.7/5', '3.6/5', '4.6/5', '4.0/5', '4.2/5',
       '3.9/5', '3.1/5', '3.0/5', '3.2/5', '3.3/5', '2.8/5', '4.4/5',
       '4.3/5', 'NEW', '2.9/5', '3.5/5', nan, '2.6/5', '3.8 /5', '3.4/5',
       '4.5/5', '2.5/5', '2.7/5', '4.7/5', '2.4/5', '2.2/5', '2.3/5',
       '3.4 /5', '-', '3.6 /5', '4.8/5', '3.9 /5', '4.2 /5', '4.0 /5',
       '4.1 /5', '3.7 /5', '3.1 /5', '2.9 /5', '3.3 /5', '2.8 /5',
       '3.5 /5', '2.7 /5', '2.5 /5', '3.2 /5', '2.6 /5', '4.5 /5',
       '4.3 /5', '4.4 /5', '4.9/5', '2.1/5', '2.0/5', '1.8/5', '4.6 /5',
       '4.9 /5', '3.0 /5', '4.8 /5', '2.3 /5', '4.7 /5', '2.4 /5',
       '2.1 /5', '2.2 /5', '2.0 /5', '1.8 /5'], dtype=object)
```

In [32]:

```
def handlerate(value):
    if(value=='NEW' or value=='_'):
        return np.nan
    else:
        value = str(value).split('/')
        value = value[0]
        return float(value)
df['rate'] =df['rate'].apply(handlerate)
```

In [33]:

```
df['rate'].head()
```

Out[33]:

```
0    4.1/5
1    4.1/5
2    3.8/5
3    3.7/5
4    3.8/5
Name: rate, dtype: object
```

In [34]:

```
df .rate.isnull().sum()
```

Out[34]:

7775

In [35]:

```
rename(columns={'approx_cost(for two people)': 'cost2plates', 'listed_in(type)': 'Type'}, inplace=True)
df.head()
```

Out[35]:

	name	online_order	book_table	rate	votes	location	rest_type	cuisines	cost2plates
0	Jalsa	Yes	Yes	4.1/5	775	Banashankari	Casual Dining	North Indian, Mughlai, Chinese	81
1	Spice Elephant	Yes	No	4.1/5	787	Banashankari	Casual Dining	Chinese, North Indian, Thai	81
2	San Churro Cafe	Yes	No	3.8/5	918	Banashankari	Cafe, Casual Dining	Cafe, Mexican, Italian	81
3	Addhuri Udupi Bhojana	No	No	3.7/5	88	Banashankari	Quick Bites	South Indian, North Indian	31
4	Grand Village	No	No	3.8/5	166	Basavanagudi	Casual Dining	North Indian, Rajasthani	61

In [36]:

```
df['location'].unique()
```

Out[36]:

```
array(['Banashankari', 'Basavanagudi', 'Mysore Road', 'Jayanagar',  
      'Kumaraswamy Layout', 'Rajarajeshwari Nagar', 'Vijay Nagar',  
      'Uttarahalli', 'JP Nagar', 'South Bangalore', 'City Market',  
      'Nagarbhavi', 'Bannerghatta Road', 'BTM', 'Kanakapura Road',  
      'Bommanahalli', nan, 'CV Raman Nagar', 'Electronic City', 'HSR',  
      'Marathahalli', 'Sarjapur Road', 'Wilson Garden', 'Shanti Nagar',  
      'Koramangala 5th Block', 'Koramangala 8th Block', 'Richmond Road',  
      'Koramangala 7th Block', 'Jalahalli', 'Koramangala 4th Block',  
      'Bellandur', 'Whitefield', 'East Bangalore', 'Old Airport Road',  
      'Indiranagar', 'Koramangala 1st Block', 'Frazer Town', 'RT Nagar',  
      'MG Road', 'Brigade Road', 'Lavelle Road', 'Church Street',  
      'Ulsoor', 'Residency Road', 'Shivajinagar', 'Infantry Road',  
      'St. Marks Road', 'Cunningham Road', 'Race Course Road',  
      'Commercial Street', 'Vasanth Nagar', 'HBR Layout', 'Domlur',  
      'Ejipura', 'Jeevan Bhima Nagar', 'Old Madras Road', 'Malleshwaram',  
      'Seshadripuram', 'Kammanahalli', 'Koramangala 6th Block',  
      'Majestic', 'Langford Town', 'Central Bangalore', 'Sanjay Nagar',  
      'Brookefield', 'ITPL Main Road, Whitefield',  
      'Varthur Main Road, Whitefield', 'KR Puram',  
      'Koramangala 2nd Block', 'Koramangala 3rd Block', 'Koramangala',  
      'Hosur Road', 'Rajajinagar', 'Banaswadi', 'North Bangalore',  
      'Nagawara', 'Hennur', 'Kalyan Nagar', 'New BEL Road', 'Jakkur',  
      'Rammurthy Nagar', 'Thippasandra', 'Kaggadasapura', 'Hebbal',  
      'Kengeri', 'Sankey Road', 'Sadashiv Nagar', 'Basaveshwara Nagar',  
      'Yeshwantpur', 'West Bangalore', 'Magadi Road', 'Yelahanka',  
      'Sahakara Nagar', 'Peenya'], dtype=object)
```

In [37]:

```
df['listed_in(city)'].unique()
```

Out[37]:

```
array(['Banashankari', 'Bannerghatta Road', 'Basavanagudi', 'Bellandur',  
      'Brigade Road', 'Brookefield', 'BTM', 'Church Street',  
      'Electronic City', 'Frazer Town', 'HSR', 'Indiranagar',  
      'Jayanagar', 'JP Nagar', 'Kalyan Nagar', 'Kammanahalli',  
      'Koramangala 4th Block', 'Koramangala 5th Block',  
      'Koramangala 6th Block', 'Koramangala 7th Block', 'Lavelle Road',  
      'Malleshwaram', 'Marathahalli', 'MG Road', 'New BEL Road',  
      'Old Airport Road', 'Rajajinagar', 'Residency Road',  
      'Sarjapur Road', 'Whitefield'], dtype=object)
```

In [38]:

```
df = df.drop(['listed_in(city)'],axis = 1)
```

In [39]:

```
df.head()
```

Out[39]:

	name	online_order	book_table	rate	votes	location	rest_type	cuisines	cost2plates
0	Jalsa	Yes	Yes	4.1/5	775	Banashankari	Casual Dining	North Indian, Mughlai, Chinese	81
1	Spice Elephant	Yes	No	4.1/5	787	Banashankari	Casual Dining	Chinese, North Indian, Thai	81
2	San Churro Cafe	Yes	No	3.8/5	918	Banashankari	Cafe, Casual Dining	Cafe, Mexican, Italian	81
3	Addhuri Udupi Bhojana	No	No	3.7/5	88	Banashankari	Quick Bites	South Indian, North Indian	31
4	Grand Village	No	No	3.8/5	166	Basavanagudi	Casual Dining	North Indian, Rajasthani	61

In [40]:

```
df['cost2plates'].unique()
```

Out[40]:

```
array([ 800.,  300.,  600.,  700.,  550.,  500.,  450.,  650.,  400.,
        900.,  200.,  750.,  150.,  850.,  100., 1200.,  350.,  250.,
        950., 1000., 1500., 1300.,  199.,   80., 1100.,  160., 1600.,
        230.,  130.,   50.,  190., 1700.,   nan, 1400.,  180., 1350.,
       2200., 2000., 1800., 1900.,  330., 2500., 2100., 3000., 2800.,
       3400.,   40., 1250., 3500., 4000., 2400., 2600.,  120., 1450.,
        469.,   70., 3200.,   60.,  560.,  240.,  360., 6000., 1050.,
       2300., 4100., 5000., 3700., 1650., 2700., 4500.,  140.] )
```

In [41]:

```
plt.figure(figsize = (16,10))
ax = sns.countplot(df['location'])
plt.xticks(rotation=90)

Text(72, 0, 'Banaswadi'),
Text(73, 0, 'North Bangalore'),
Text(74, 0, 'Nagawara'),
Text(75, 0, 'Hennur'),
Text(76, 0, 'Kalyan Nagar'),
Text(77, 0, 'New BEL Road'),
Text(78, 0, 'Jakkur'),
Text(79, 0, 'Rammurthy Nagar'),
Text(80, 0, 'Thippasandra'),
Text(81, 0, 'Kaggadasapura'),
Text(82, 0, 'Hebbal'),
Text(83, 0, 'Kengeri'),
Text(84, 0, 'Sankey Road'),
Text(85, 0, 'Sadashiv Nagar'),
Text(86, 0, 'Basaveshwara Nagar'),
Text(87, 0, 'Yeshwantpur'),
Text(88, 0, 'West Bangalore'),
Text(89, 0, 'Magadi Road'),
Text(90, 0, 'Yelahanka'),
Text(91, 0, 'Sahakara Nagar').
```

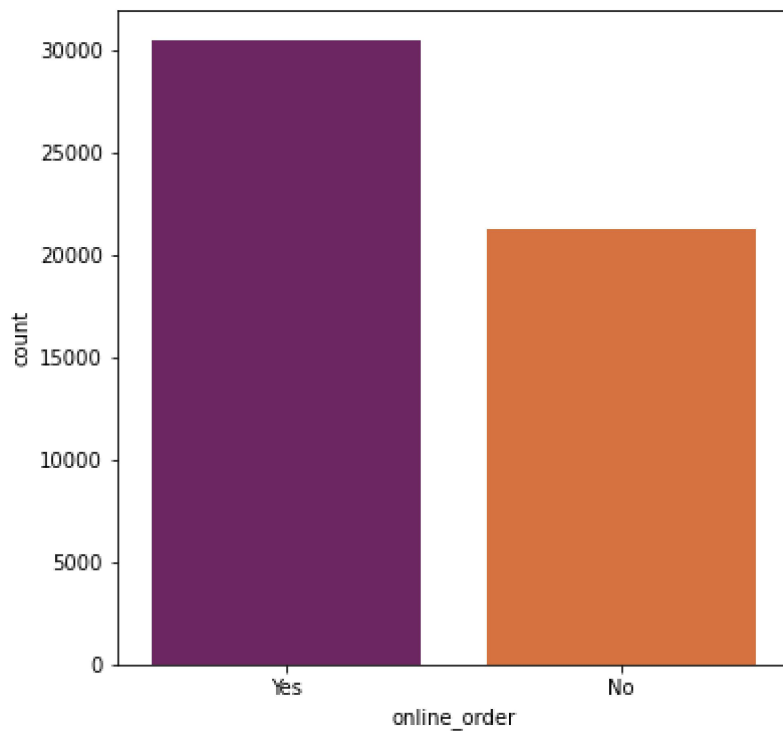
In [42]:

```
plt.figure(figsize = (6,6))  
sns.countplot(df['online_order'],palette = 'inferno')
```

C:\Users\NITU\anaconda3\lib\site-packages\seaborn\\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.  
warnings.warn(

Out[42]:

<AxesSubplot:xlabel='online\_order', ylabel='count'>





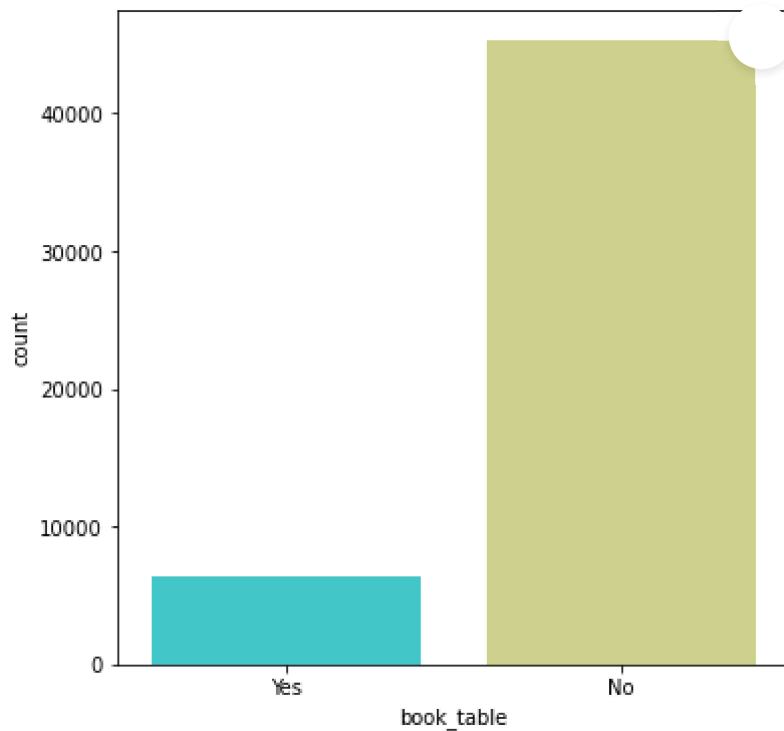
In [43]:

```
plt.figure(figsize = (6,6))  
sns.countplot(df['book_table'],palette = 'rainbow')
```

C:\Users\NITU\anaconda3\lib\site-packages\seaborn\\_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.  
warnings.warn(

Out[43]:

<AxesSubplot:xlabel='book\_table', ylabel='count'>



In [ ]:

In [ ]:

In [ ]:

In [ ]: