Zomanto Data Analysis

February 23, 2024

[9]:
```
# Importing the necessary libraries.
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

[10]:
```
# Let's load the data and store it in a variable.
df = pd.read_csv(r'C:\Users\91702\Documents\zomato.csv')
```

[54]:
```
df.shape
```

[54]:
```
(51717, 17)
```

[55]:
```
df.columns
```

[55]:
```
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')
```

[412]:
```
#Dropping few columns since its not needed for our analysis
df.drop(['url',
         'address',
         'phone',
         'menu_item',
         'dish_liked',
         'reviews_list'],
        axis = 1,
        inplace = True)
```

[413]:
```
df.head(4)
```

[413]:
```
                    name online_order book_table  rate  votes      location  \
0                  Jalsa          Yes        Yes  4.1/5    775  Banashankari
1         Spice Elephant          Yes         No  4.1/5    787  Banashankari
2        San Churro Cafe          Yes         No  3.8/5    918  Banashankari
3  Addhuri Udupi Bhojana           No         No  3.7/5     88  Banashankari
```

```
             rest_type                        cuisines  \
0        Casual Dining  North Indian, Mughlai, Chinese
1        Casual Dining     Chinese, North Indian, Thai
2  Cafe, Casual Dining          Cafe, Mexican, Italian
3          Quick Bites      South Indian, North Indian

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

[58]:

```
#Checking the data types
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  object
 9   listed_in(type)              51717 non-null  object
 10  listed_in(city)              51717 non-null  object
dtypes: int64(1), object(10)
memory usage: 4.3+ MB
```

1 Dropping Duplicates

[59]:

```
df.drop_duplicates(inplace = True)
df.shape
```

[59]:

```
(51609, 11)
```

[61]:

```
# Cleaning Rate Column.
df['rate'].unique()
```

[61]:

```
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)
```

[62]:

```
# Removing the 'NEW', '-' and '/5' from Rate Column since its not needed.
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)
```

[63]:

```
df.head(4)
```

[63]:

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

             rest_type                        cuisines  \
0        Casual Dining  North Indian, Mughlai, Chinese   
1        Casual Dining     Chinese, North Indian, Thai   
2  Cafe, Casual Dining          Cafe, Mexican, Italian   
3          Quick Bites      South Indian, North Indian   

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

1.1 Filling Null values in Rate Column with Mean

[64]:

```
df['rate'].fillna(df['rate'].mean(),
                  inplace = True)
df['rate'].isnull().sum()
```

[64]:

```
0
```

[65]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 51609 entries, 0 to 51716
Data columns (total 11 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   name                         51609 non-null  object
 1   online_order                 51609 non-null  object
 2   book_table                   51609 non-null  object
 3   rate                         51609 non-null  float64
 4   votes                        51609 non-null  int64
 5   location                     51588 non-null  object
 6   rest_type                    51382 non-null  object
 7   cuisines                     51564 non-null  object
 8   approx_cost(for two people)  51265 non-null  object
 9   listed_in(type)              51609 non-null  object
 10  listed_in(city)              51609 non-null  object
dtypes: float64(1), int64(1), object(9)
memory usage: 4.7+ MB
```

1.2 Dropping Null values

[66]:
```
df.dropna(inplace = True)
```

[67]:
```
df.rename(columns = {'approx_cost(for two people)' : 'Cost2plates',
                     'listed_in(type)' : 'Type' },
          inplace = True)
```

[68]:
```
df.head()
```

[68]:
```
                    name online_order book_table  rate  votes      location  \
0                  Jalsa          Yes        Yes   4.1    775  Banashankari
1         Spice Elephant          Yes         No   4.1    787  Banashankari
2        San Churro Cafe          Yes         No   3.8    918  Banashankari
3  Addhuri Udupi Bhojana           No         No   3.7     88  Banashankari
4          Grand Village           No         No   3.8    166  Basavanagudi

             rest_type                        cuisines Cost2plates    Type  \
0        Casual Dining  North Indian, Mughlai, Chinese         800  Buffet
1        Casual Dining     Chinese, North Indian, Thai         800  Buffet
2  Cafe, Casual Dining          Cafe, Mexican, Italian         800  Buffet
3          Quick Bites      South Indian, North Indian         300  Buffet
4        Casual Dining        North Indian, Rajasthani         600  Buffet

  listed_in(city)
0    Banashankari
1    Banashankari
```

```
2    Banashankari
3    Banashankari
4    Banashankari
```

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

```
[69]: 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', 'CV Raman Nagar', 'Electronic City', 'HSR',
             'Marathahalli', 'Wilson Garden', 'Shanti Nagar',
             'Koramangala 5th Block', 'Koramangala 8th Block', 'Richmond Road',
             'Koramangala 7th Block', 'Jalahalli', 'Koramangala 4th Block',
             'Bellandur', 'Sarjapur Road', '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)
```

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

```
[70]: 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)
```

1.2.1 Since we have Listed_in(City) and location both, So I'm keeping only one

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

```
[76]: df['Cost2plates'].unique()
```

```
[76]: array(['800', '300', '600', '700', '550', '500', '450', '650', '400',
             '900', '200', '750', '150', '850', '100', '1,200', '350', '250',
             '950', '1,000', '1,500', '1,300', '199', '80', '1,100', '160',
             '1,600', '230', '130', '50', '190', '1,700', '1,400', '180',
             '1,350', '2,200', '2,000', '1,800', '1,900', '330', '2,500',
             '2,100', '3,000', '2,800', '3,400', '40', '1,250', '3,500',
             '4,000', '2,400', '2,600', '120', '1,450', '469', '70', '3,200',
             '60', '560', '240', '360', '6,000', '1,050', '2,300', '4,100',
             '5,000', '3,700', '1,650', '2,700', '4,500', '140'], dtype=object)
```

1.2.2 Removing the commas(,) from Cost2plates column values

```
[414]: def handlecomma(value):
           if ',' in value:
               value = value.replace(',', '')
               return float(value)
           else:
               return float(value)


       # df['Cost2plates'] = df['Cost2plates'].apply(handlecomma)
       # df['Cost2plates'].unique()
```

1.2.3 Cleaning rest type column

```
[108]: rest_types = df['rest_type'].value_counts(ascending = False)
       rest_types
```

```
[108]: rest_type
       Quick Bites           19010
       Casual Dining         10253
       others                 9003
       Cafe                   3682
       Delivery               2574
       Dessert Parlor         2242
       Takeaway, Delivery     2008
       Bakery                 1140
       Casual Dining, Bar     1130
       Name: count, dtype: int64
```

[415]:

```
rest_types_lessthan1000 = rest_types[rest_types<1000]

def handle_rest_type(value):
    if(value in rest_types_lessthan1000):
        return 'others'
    else:
        return value

df['rest_type'] = df['rest_type'].apply(handle_rest_type)
df['rest_type'].value_counts()
```

[415]:

```
rest_type
Quick Bites                      19132
Casual Dining                    10330
Cafe                              3732
Delivery                          2604
Dessert Parlor                    2263
                                 …
Dessert Parlor, Kiosk                2
Food Court, Beverage Shop            2
Dessert Parlor, Food Court           2
Sweet Shop, Dessert Parlor           1
Quick Bites, Kiosk                   1
Name: count, Length: 93, dtype: int64
```

1.2.4 Cleaning location column

[109]:

```
location = df['location'].value_counts(ascending = False)

location_lessthan300 = location[location<300]

def handle_location(value):
    if (value in location_lessthan300):
        return 'others'
    else:
        return value

df['location'] = df['location'].apply(handle_location)
df['location'].value_counts()
```

[109]:

```
location
BTM                         5056
others                      4954
HSR                         2494
Koramangala 5th Block       2479
JP Nagar                    2218
Whitefield                  2105
```

```
Indiranagar              2026
Jayanagar                1916
Marathahalli             1805
Bannerghatta Road        1609
Bellandur                1268
Electronic City          1246
Koramangala 1st Block    1236
Brigade Road             1210
Koramangala 7th Block    1174
Koramangala 6th Block    1127
Sarjapur Road            1047
Koramangala 4th Block    1017
Ulsoor                   1011
Banashankari              902
MG Road                   893
Kalyan Nagar              841
Richmond Road             803
Malleshwaram              721
Frazer Town               714
Basavanagudi              684
Residency Road            671
Brookefield               656
New BEL Road              644
Banaswadi                 640
Kammanahalli              639
Rajajinagar               591
Church Street             566
Lavelle Road              518
Shanti Nagar              508
Shivajinagar              498
Cunningham Road           490
Domlur                    482
Old Airport Road          437
Ejipura                   433
Commercial Street         370
St. Marks Road            343
Name: count, dtype: int64
```

1.2.5 Cleaning Cuisines Columns

[156]:

```python
cuisines = df['cuisines'].value_counts(ascending = False)

cuisines_lessthan100 = cuisines[cuisines<100]

def handle_cuisines(value):
    if (value in cuisines_lessthan100):
        return 'others'
    else:
        return value

df['cuisines'] = df['cuisines'].apply(handle_cuisines)
df['cuisines'].value_counts()
```

[156]:

```
cuisines
others                               26159
North Indian                          2852
North Indian, Chinese                 2351
South Indian                          1820
Biryani                                903
                                     …
South Indian, Chinese, North Indian    105
North Indian, Mughlai, Chinese         104
South Indian, Fast Food                104
Italian, Pizza                         102
North Indian, Chinese, Seafood         102
Name: count, Length: 70, dtype: int64
```

[119]:

```
df.head(10)
```

[119]:

```
                                              name online_order book_table  \
0                                            Jalsa          Yes        Yes
1                                   Spice Elephant          Yes         No
2                                  San Churro Cafe          Yes         No
3                            Addhuri Udupi Bhojana           No         No
4                                    Grand Village           No         No
5                                  Timepass Dinner          Yes         No
6  Rosewood International Hotel - Bar & Restaurant           No         No
7                                           Onesta          Yes        Yes
8                                   Penthouse Cafe          Yes         No
9                                        Smacznego          Yes         No

   rate  votes      location      rest_type                        cuisines  \
0   4.1    775  Banashankari  Casual Dining  North Indian, Mughlai, Chinese
1   4.1    787  Banashankari  Casual Dining                          others
2   3.8    918  Banashankari         others                          others
3   3.7     88  Banashankari    Quick Bites      South Indian, North Indian
4   3.8    166  Basavanagudi  Casual Dining                          others
5   3.8    286  Basavanagudi  Casual Dining                    North Indian
6   3.6      8        others  Casual Dining                          others
7   4.6   2556  Banashankari         others                          others
8   4.0    324  Banashankari           Cafe                          others
9   4.2    504  Banashankari           Cafe                          others

   Cost2plates    Type
```

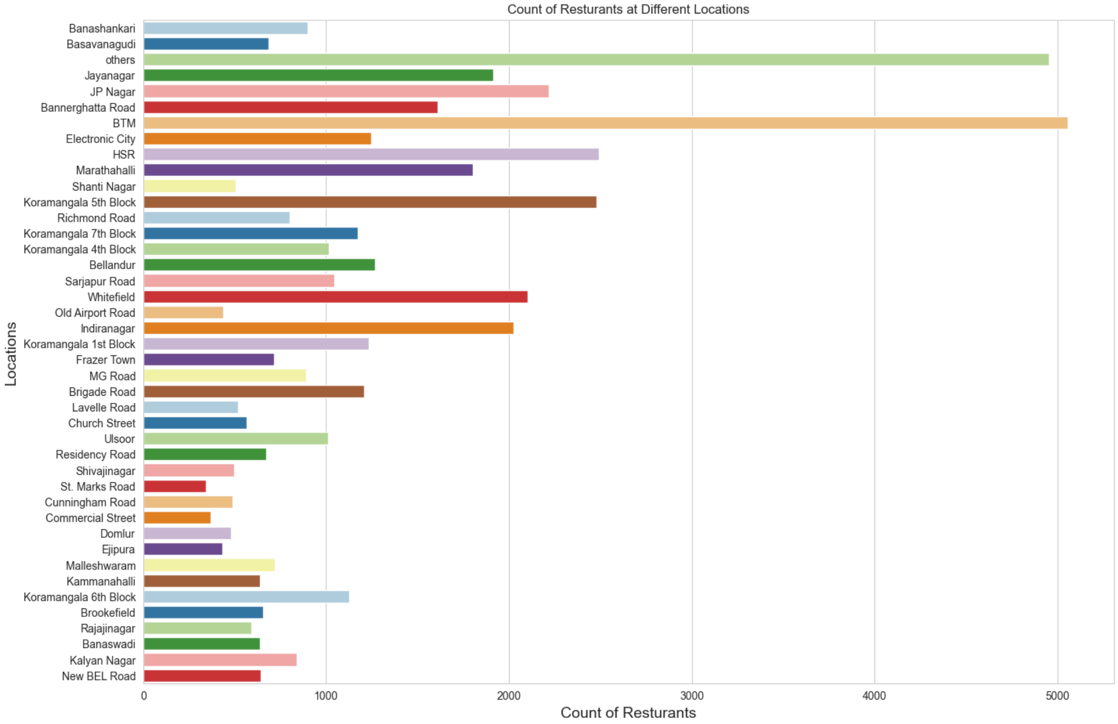
```
0        800.0  Buffet
1        800.0  Buffet
2        800.0  Buffet
3        300.0  Buffet
4        600.0  Buffet
5        600.0  Buffet
6        800.0  Buffet
7        600.0   Cafes
8        700.0   Cafes
9        550.0   Cafes
```

1.3 Completed with the Data Cleaning. Now let's visualize the dataset!!!

Count plot on location

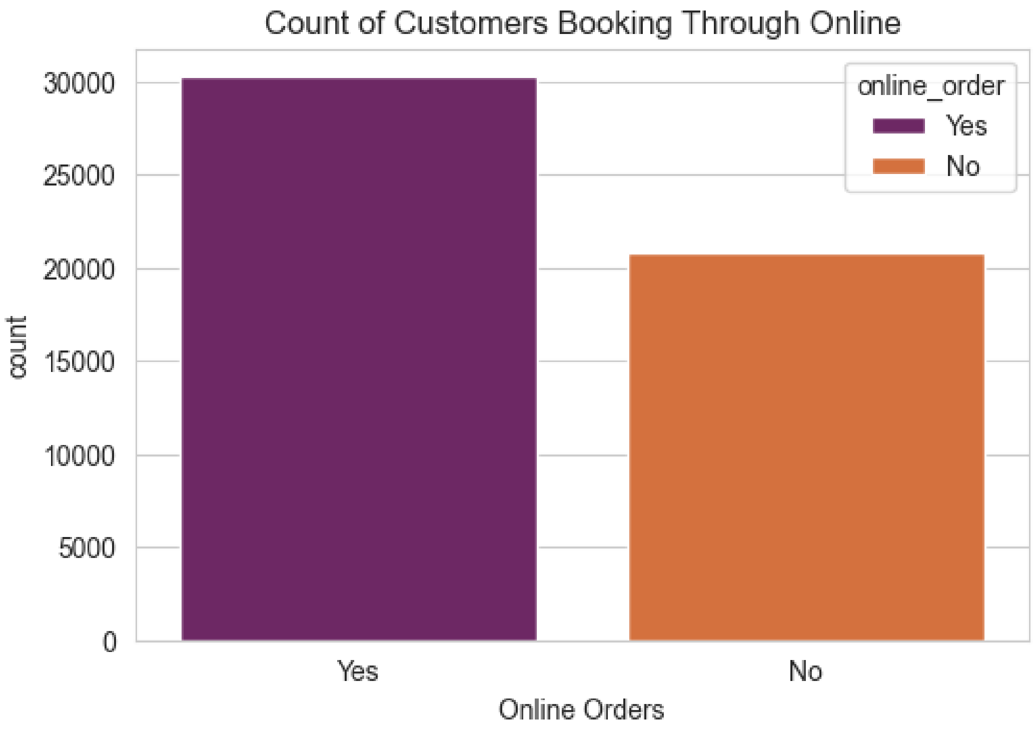
[204]:

```
plt.figure(figsize = (16,11))

ax = sns.countplot(y = df['location'],
                   palette = 'Paired',
                   hue = df['location'])
plt.title('Count of Resturants at Different Locations')
plt.xlabel('Count of Resturants', fontsize = 14)
plt.ylabel('Locations',  fontsize = 14)

plt.show()
```



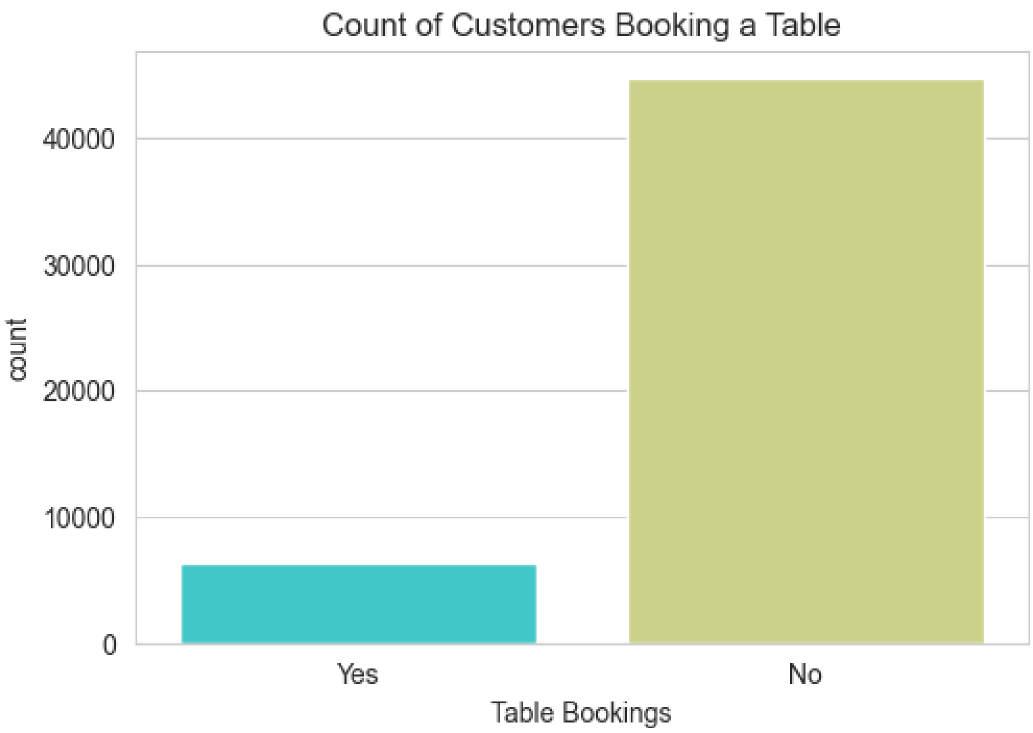
Visualizing the Online Orders

```
[205]: plt.figure(figsize = (6,4))

sns.countplot(x = df['online_order'],
              palette = 'inferno',
              hue = df['online_order'], legend = True)

plt.title('Count of Customers Booking Through Online')
plt.xlabel('Online Orders')

plt.show()
```



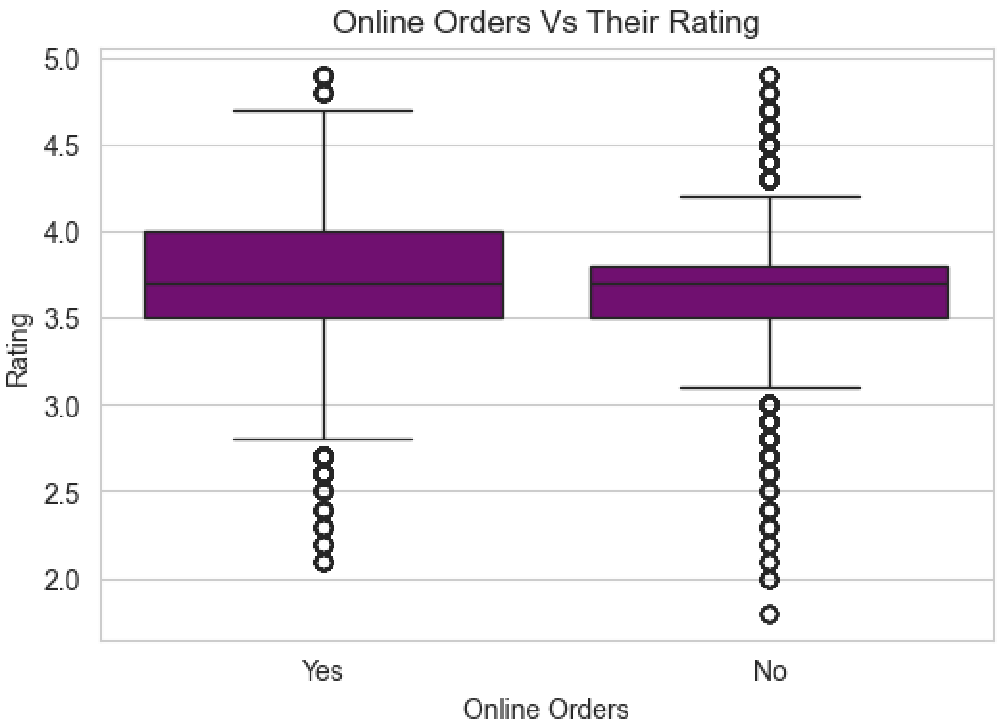
Visualizing Book Table

```
[206]: plt.figure(figsize = (6,4))

sns.countplot(x = df['book_table'],
              palette = 'rainbow',
              hue = df['book_table'])

plt.title('Count of Customers Booking a Table')
plt.xlabel('Table Bookings')

plt.show()
```



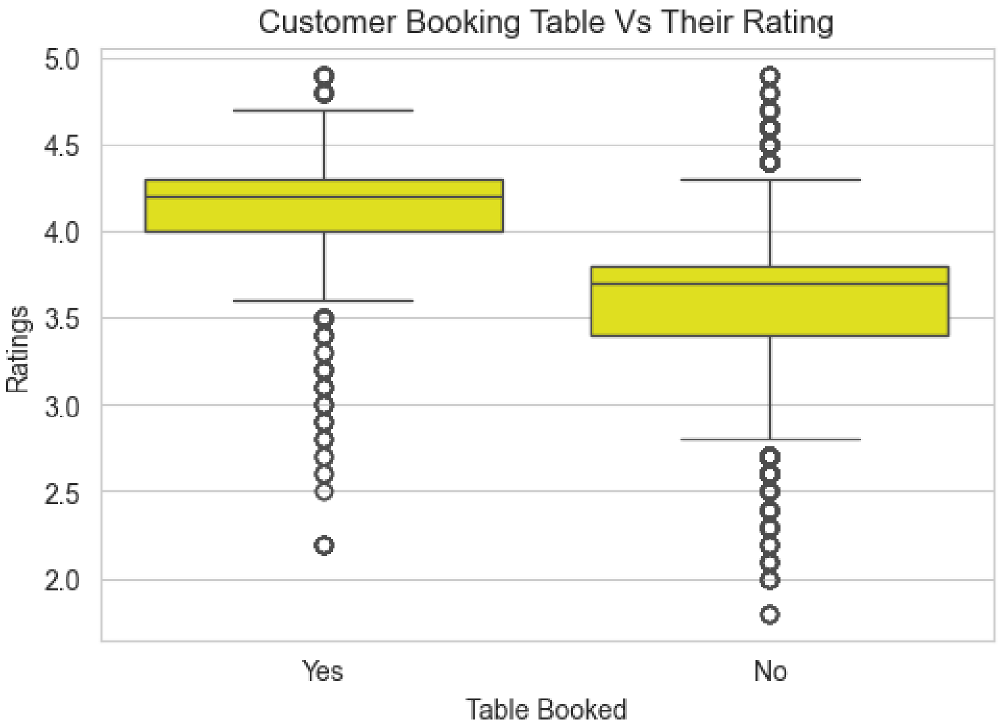
Visualizing Online Order Vs Rate

[242]:
```
plt.figure(figsize = (6,4))

sns.boxplot(data = df,
            x = 'online_order',
            y = 'rate',
           color = 'purple')

plt.title('Online Orders Vs Their Rating')
plt.xlabel('Online Orders')
plt.ylabel('Rating')

plt.show()
```



Visualizing Book Table Vs Rate

[240]:
```
plt.figure(figsize = (6,4))

sns.boxplot(data = df,
            x = 'book_table',
            y = 'rate',
           color = 'yellow')

plt.title('Customer Booking Table Vs Their Rating')
plt.xlabel('Table Booked')
plt.ylabel('Ratings')

plt.show()
```



Visualizing Online Orders Facility, Location Wise

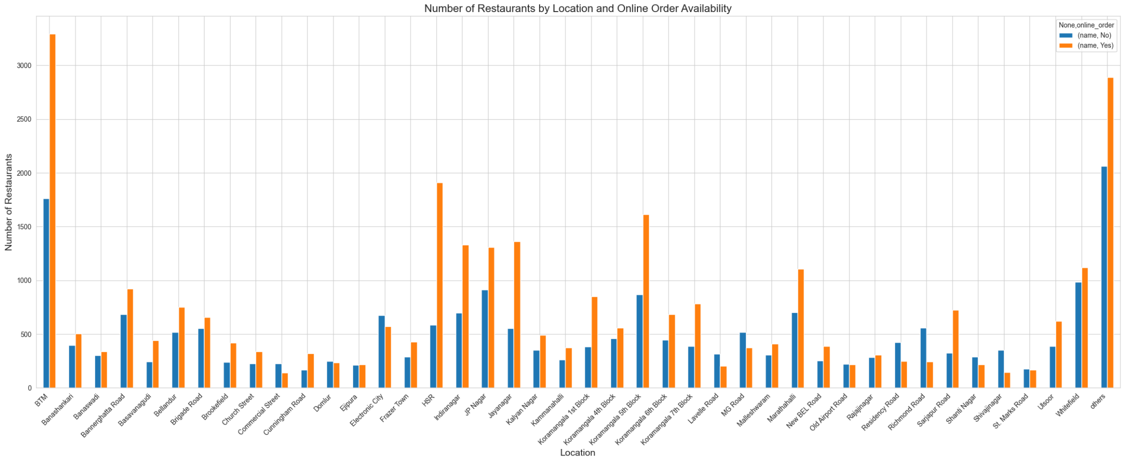
[296]:
```
df1 = df.groupby(['location', 'online_order'])['name'].count()
df1.to_csv('location_online.csv')

df1 = pd.read_csv('location_online.csv')
df1 = pd.pivot_table(df1,
                     values=None,
                     index=['location'],
                     columns=['online_order'],
                     fill_value=0,
                     aggfunc='sum')  # Use 'sum' as a string

df1
```

[296]:
```
                        name
online_order              No   Yes
location
BTM                     1763  3293
Banashankari             397   505
Banaswadi                302   338
Bannerghatta Road        685   924
Basavanagudi             243   441
Bellandur                517   751
```

```
Brigade Road            552   658
Brookefield             239   417
Church Street           226   340
Commercial Street       228   142
Cunningham Road         168   322
Domlur                  247   235
Ejipura                 214   219
Electronic City         676   570
Frazer Town             287   427
HSR                     584  1910
Indiranagar             697  1329
JP Nagar                911  1307
Jayanagar               552  1364
Kalyan Nagar            350   491
Kammanahalli            264   375
Koramangala 1st Block   384   852
Koramangala 4th Block   459   558
Koramangala 5th Block   866  1613
Koramangala 6th Block   445   682
Koramangala 7th Block   389   785
Lavelle Road            315   203
MG Road                 520   373
Malleshwaram            309   412
Marathahalli            701  1104
New BEL Road            255   389
Old Airport Road        221   216
Rajajinagar             286   305
Residency Road          424   247
Richmond Road           557   246
Sarjapur Road           323   724
Shanti Nagar            289   219
Shivajinagar            354   144
St. Marks Road          176   167
Ulsoor                  389   622
Whitefield              986  1119
others                 2064  2890
```

[297]:

```
df1.plot(kind='bar',
         figsize=(29,10),
         color=['#1f77b4', '#ff7f0e'])

plt.title('Number of Restaurants by Location and Online Order Availability',␣
  ↪fontsize=16)
plt.xlabel('Location', fontsize=14)
plt.ylabel('Number of Restaurants', fontsize=14)
plt.xticks(rotation=45, ha='right')
plt.show()
```



Visualizing Book Table Facility, Location Wise

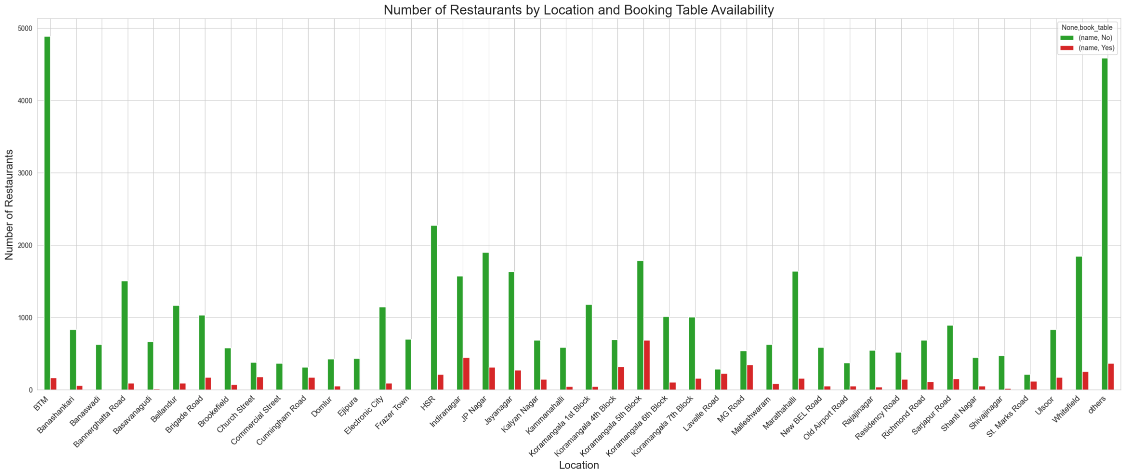
[283]:
```
df2 = df.groupby(['location', 'book_table'])['name'].count()
df2.to_csv('location_booktable.csv')

df2 = pd.read_csv('location_booktable.csv')
df2 = pd.pivot_table(df2,
                     values=None,
                     index=['location'],
                     columns=['book_table'],
                     fill_value=0,
                     aggfunc='sum')  # Use 'sum' as a string

df2
```

[283]:
```
                       name
book_table               No  Yes
location
BTM                    4889  167
Banashankari            839   63
Banaswadi               632    8
Bannerghatta Road      1510   99
Basavanagudi            668   16
Bellandur              1170   98
Brigade Road           1034  176
Brookefield             582   74
Church Street           385  181
Commercial Street       370    0
Cunningham Road         315  175
Domlur                  427   55
```

```
Ejipura                   433    0
Electronic City          1148   98
Frazer Town               706    8
HSR                      2277  217
Indiranagar              1578  448
JP Nagar                 1903  315
Jayanagar                1637  279
Kalyan Nagar              692  149
Kammanahalli              590   49
Koramangala 1st Block    1186   50
Koramangala 4th Block     695  322
Koramangala 5th Block    1787  692
Koramangala 6th Block    1015  112
Koramangala 7th Block    1012  162
Lavelle Road              290  228
MG Road                   546  347
Malleshwaram              632   89
Marathahalli             1642  163
New BEL Road              588   56
Old Airport Road          378   59
Rajajinagar               550   41
Residency Road            522  149
Richmond Road             687  116
Sarjapur Road             893  154
Shanti Nagar              451   57
Shivajinagar              475   23
St. Marks Road            219  124
Ulsoor                    834  177
Whitefield               1852  253
others                   4587  367
```

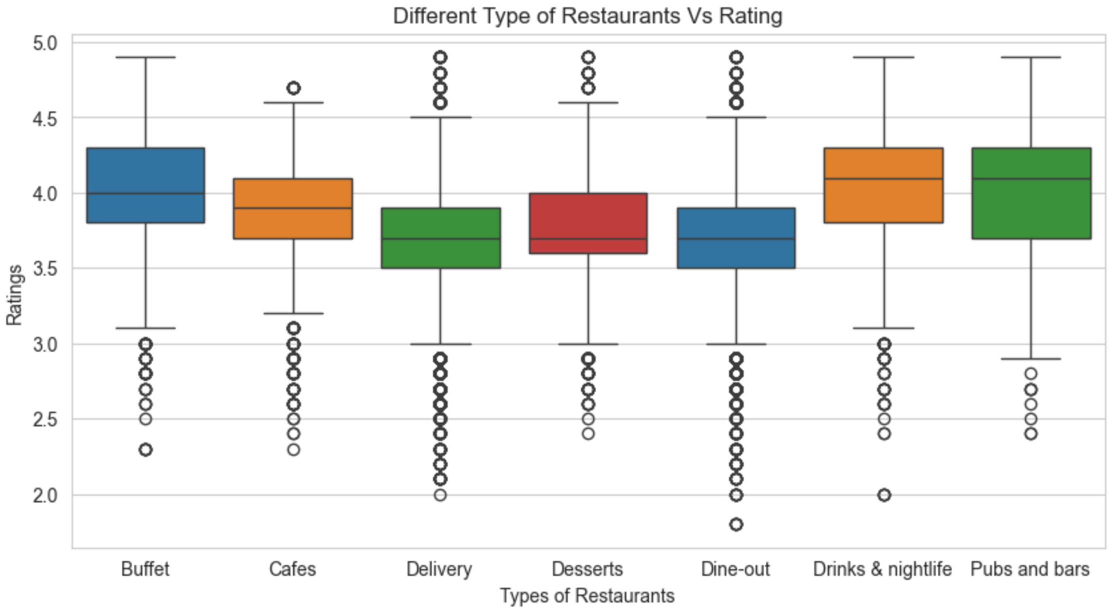
[321]:

```
df2.plot(kind='bar',
         figsize=(29, 10),
         color=['#2ca02c', '#d62728'])

plt.title('Number of Restaurants by Location and Booking Table Availability',␣
 ↪fontsize=20)
plt.xlabel('Location', fontsize=16)
plt.ylabel('Number of Restaurants', fontsize=16)
plt.xticks(rotation=45, ha='right', fontsize=12)

plt.show()
```



Visualizing Types of Resturants Vs Rate

[320]:
```
plt.figure(figsize = (10,5))

custom_colors = {"Type1": "#1f77b4",
                 "Type2": "#ff7f0e",
                 "Type3": "#2ca02c",
                 "Type4": "#d62728"}
sns.boxplot(data = df,
            x = 'Type',
            y = 'rate',
            palette = custom_colors.values(),
            hue = 'Type')

plt.title('Different Type of Restaurants Vs Rating')
plt.xlabel('Types of Restaurants')
plt.ylabel('Ratings')

plt.show()
```



Grouping Types of Resturants, Location Wise

[332]:

```
df3 = df.groupby(['location', 'Type'])['name'].count()

df3.to_csv('location_type.csv')
df3 = pd.read_csv('location_type.csv')

df3 = pd.pivot_table(df3,
                    values = None,
                    index = ['location'],
                    columns = ['Type'],
                    fill_value = 0,
                    aggfunc = 'sum')
df3
```

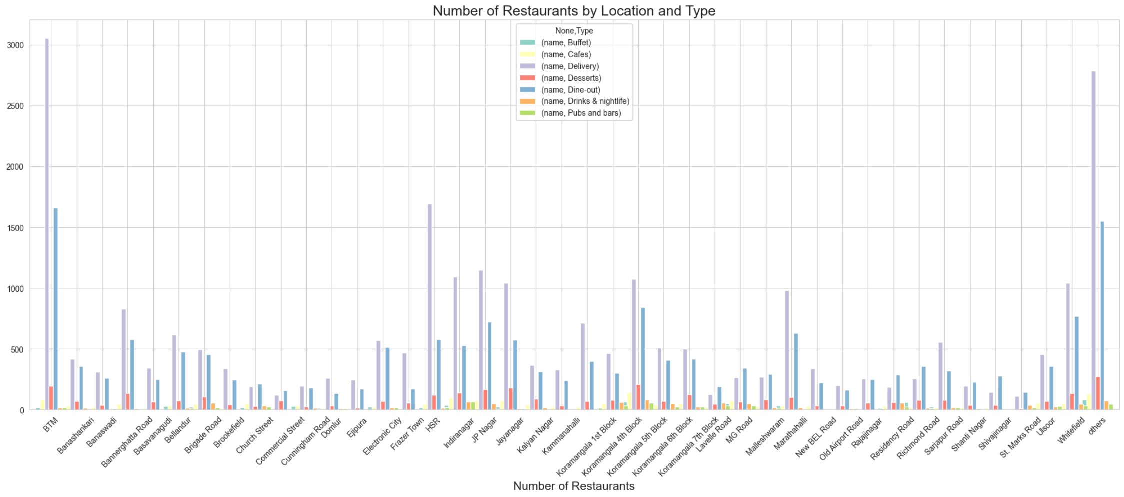
[332]:

```
                        name                                    \
Type                  Buffet Cafes Delivery Desserts Dine-out
location
BTM                       21    83     3053      198     1660
Banashankari               7    36      418       71      356
Banaswadi                  0    24      310       37      262
Bannerghatta Road          9    46      828      137      578
Basavanagudi               7    11      344       66      251
Bellandur                 28    36      617       75      479
Brigade Road              25    46      497      108      455
Brookefield                6    17      339       45      245
Church Street             19    51      193       29      215
```

```
Commercial Street          0    13      121       77      159
Cunningham Road           29    34      194       26      184
Domlur                    15    13      261       35      135
Ejipura                    0     0      245       16      172
Electronic City           23    24      570       71      516
Frazer Town                1    11      470       56      172
HSR                       19    49     1694      120      580
Indiranagar               38    97     1091      140      529
JP Nagar                  45    76     1151      166      722
Jayanagar                 27    77     1043      182      575
Kalyan Nagar               9    45      366       88      315
Kammanahalli               2    27      329       35      240
Koramangala 1st Block      3    26      716       70      398
Koramangala 4th Block     21    53      464       81      302
Koramangala 5th Block     65   146     1075      209      842
Koramangala 6th Block     18    43      511       70      411
Koramangala 7th Block     25    52      503      127      417
Lavelle Road              30    27      127       50      191
MG Road                   51    76      266       68      343
Malleshwaram              11    31      269       85      291
Marathahalli              34    32      980      105      630
New BEL Road               4    29      338       33      224
Old Airport Road          12     5      200       35      164
Rajajinagar               10     4      258       55      251
Residency Road            20    31      187       63      289
Richmond Road             63    21      257       78      356
Sarjapur Road             25    22      558       82      319
Shanti Nagar               9    22      198       39      229
Shivajinagar               6    17      143       37      280
St. Marks Road             5    10      111       10      145
Ulsoor                    16    56      456       71      359
Whitefield                28    51     1041      137      768
others                    83   133     2787      276     1553


Type                  Drinks & nightlife Pubs and bars
location
BTM                                   22            19
Banashankari                          14             0
Banaswadi                              6             1
Bannerghatta Road                      9             2
Basavanagudi                           5             0
Bellandur                             17            16
Brigade Road                          57            22
Brookefield                            4             0
Church Street                         36            23
Commercial Street                      0             0
```

```
Cunningham Road                       16             7
Domlur                                12            11
Ejipura                                0             0
Electronic City                       21            21
Frazer Town                            2             2
HSR                                   14            18
Indiranagar                           65            66
JP Nagar                              51             7
Jayanagar                             12             0
Kalyan Nagar                          18             0
Kammanahalli                           6             0
Koramangala 1st Block                  7            16
Koramangala 4th Block                 62            34
Koramangala 5th Block                 84            58
Koramangala 6th Block                 51            23
Koramangala 7th Block                 25            25
Lavelle Road                          59            34
MG Road                               53            36
Malleshwaram                          20            14
Marathahalli                          22             2
New BEL Road                           8             8
Old Airport Road                      12             9
Rajajinagar                            3            10
Residency Road                        55            26
Richmond Road                         16            12
Sarjapur Road                         19            22
Shanti Nagar                           9             2
Shivajinagar                           7             8
St. Marks Road                        40            22
Ulsoor                                23            30
Whitefield                            47            33
others                                75            47
```

```
[351]: df3.plot(kind = 'bar',
                figsize = (25,9),
                width = 1.2)

       plt.title('Number of Restaurants by Location and Type', fontsize = 18)
       plt.xlabel('Location', fontsize = 15)
       plt.xlabel('Number of Restaurants', fontsize = 15)

       plt.xticks(rotation = 45)

       plt.show()
```

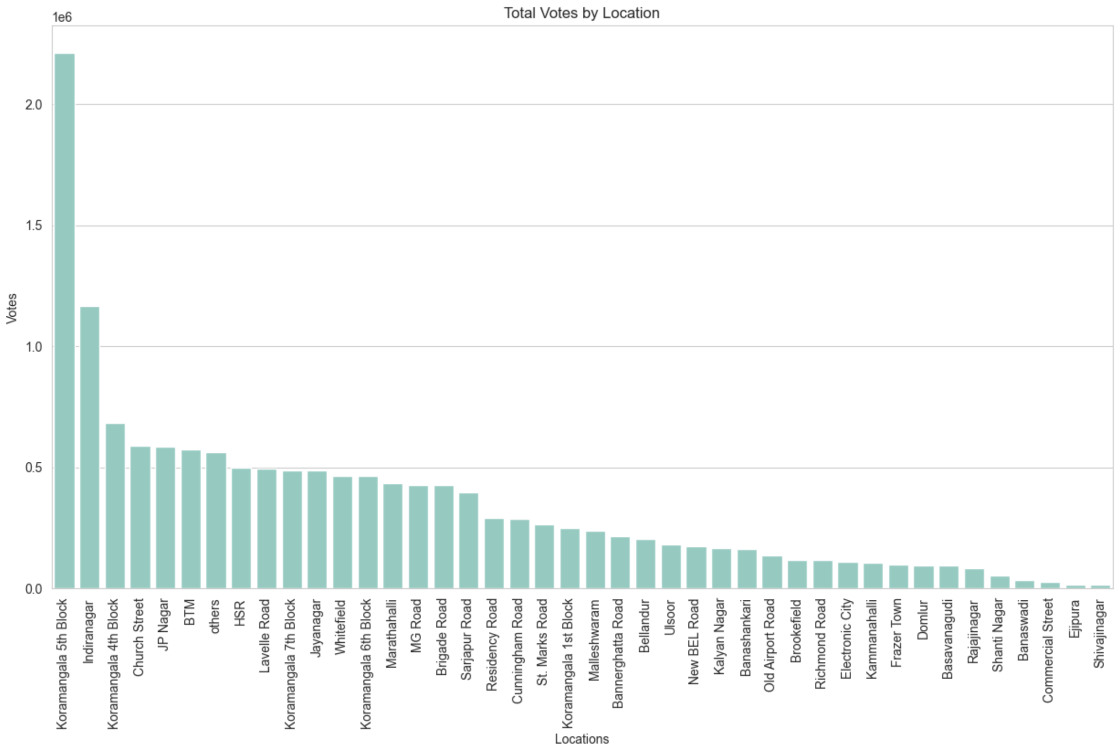


Number of Votes, Location Wise

[359]:
```
df4 = df[['location', 'votes']]
df4.drop_duplicates()

df5 = df4.groupby(['location'])['votes'].sum()
df5 = df5.to_frame()
df5 = df5.sort_values('votes', ascending = False)
df5.head(6)
```

[359]:
```
                           votes
location
Koramangala 5th Block    2214083
Indiranagar              1165909
Koramangala 4th Block     685156
Church Street             590306
JP Nagar                  586522
BTM                       573668
```

[381]:
```
plt.figure(figsize = (15,8))
sns.barplot(x=df5.index,
            y='votes',
            data=df5)

plt.title('Total Votes by Location')
plt.xlabel('Locations')
plt.ylabel('Votes')
plt.xticks(rotation = 90)

plt.show()
```



1.3.1 Visualizing Top Cuisines

[12]:

```
df6 = df[['cuisines', 'votes']]
df6.drop_duplicates()

df7 = df6.groupby(['cuisines'])['votes'].sum()
df7 = df7.to_frame()
df7 = df7.sort_values('votes', ascending = False)
df7
```

[12]:

```
                                              votes
cuisines
North Indian                                 558654
Pizza, Cafe, Italian                         347520
Cafe, American, Burger, Steak                301059
North Indian, Chinese                        259864
North Indian, Mughlai, South Indian, Chinese 227930
…                                              …
Italian, Pizza, Fast Food                         0
Salad, Beverages, Juices                          0
Bengali, Oriya, South Indian                      0
Salad, Asian, Continental, Healthy Food           0
Lucknowi, Kebab, Rolls                            0
```

```
[2723 rows x 1 columns]
```

[13]:

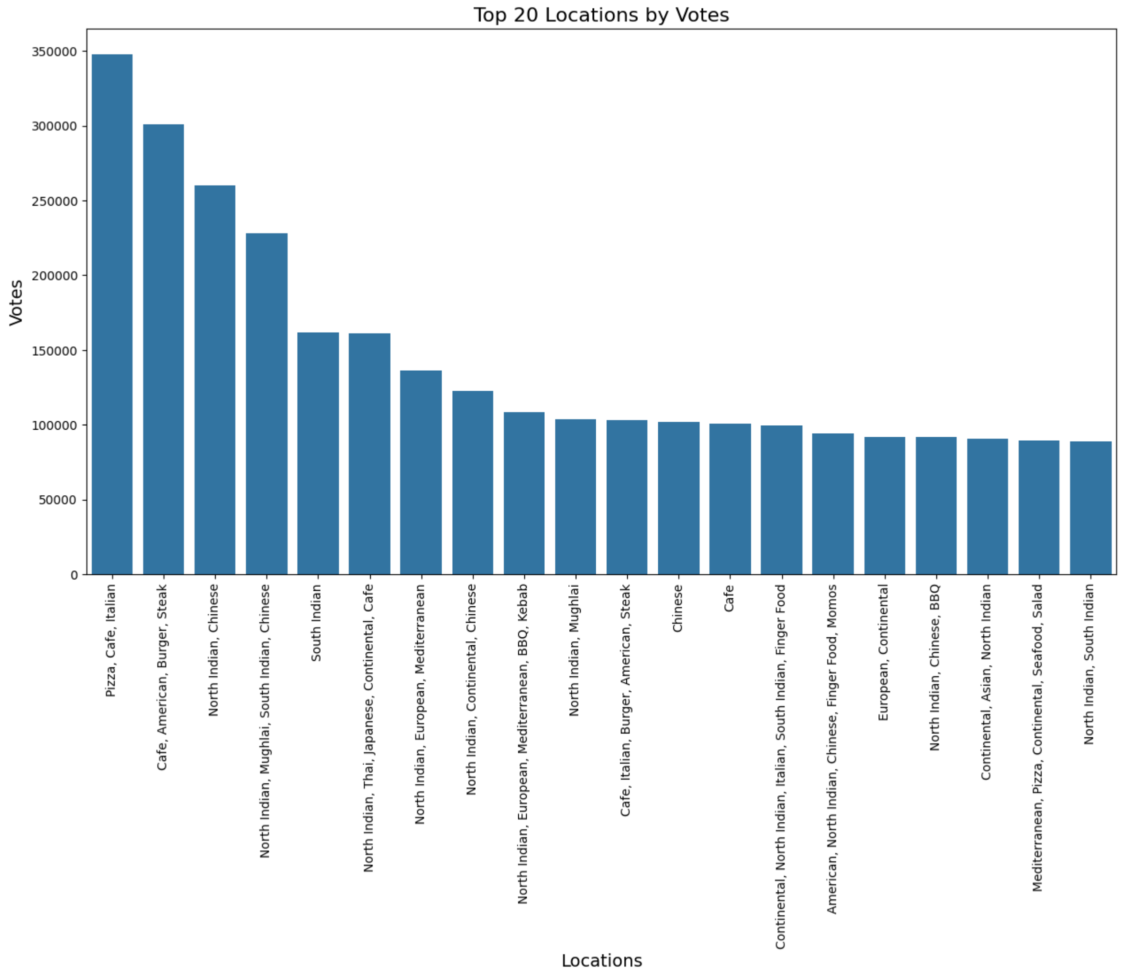
```
df7 = df7.iloc[1:, :]  # Remove the first row to exclude the row containing␣
 ↪"other cuisines" and retain only the other rows.
df7
```

[13]:

```
                                              votes
cuisines
Pizza, Cafe, Italian                         347520
Cafe, American, Burger, Steak                301059
North Indian, Chinese                        259864
North Indian, Mughlai, South Indian, Chinese 227930
South Indian                                 161991
…                                               …
Italian, Pizza, Fast Food                         0
Salad, Beverages, Juices                          0
Bengali, Oriya, South Indian                      0
Salad, Asian, Continental, Healthy Food           0
Lucknowi, Kebab, Rolls                            0

[2722 rows x 1 columns]
```

[15]:

```
# Selecting the top 10 entries from df7
top_20_df7 = df7.head(20)

plt.figure(figsize=(15, 8))

sns.barplot(x=top_20_df7.index, y='votes', data=top_20_df7)

plt.title('Top 20 Locations by Votes', fontsize=16)
plt.xlabel('Locations', fontsize=14)
plt.ylabel('Votes', fontsize=14)
plt.xticks(rotation=90)

plt.show()
```


Top 20 Locations by Votes
Votes
350000
300000
250000
200000
150000
100000
50000
0
Pizza, Cafe, Italian
Cafe, American, Burger, Steak
North Indian, Chinese
North Indian, Mughlai, South Indian, Chinese
South Indian
North Indian, Thai, Japanese, Continental, Cafe
North Indian, European, Mediterranean
North Indian, Continental, Chinese
North Indian, European, Mediterranean, BBQ, Kebab
North Indian, Mughlai
Cafe, Italian, Burger, American, Steak
Chinese
Cafe
Continental, North Indian, Italian, South Indian, Finger Food
American, North Indian, Chinese, Finger Food, Momos
European, Continental
North Indian, Chinese, BBQ
Continental, Asian, North Indian
Mediterranean, Pizza, Continental, Seafood, Salad
North Indian, South Indian
Locations

[ ]: