```
]: df_date = pd.read_csv("C:\\Users\\91990\\anlsys_pro_main_1\\dim_date.csv")
   df_hotels = pd.read_csv("C:\\Users\\91990\\anlsys_pro_main_1\\dim_hotels.csv")
   df_rooms = pd.read_csv("C:\\Users\\91990\\anlsys_pro_main_1\\dim_rooms.csv")
   df_agg_bookings = pd.read_csv("C:\\Users\\91990\\anlsys_pro_main_1\\fact_aggregated_bookings.csv")
```

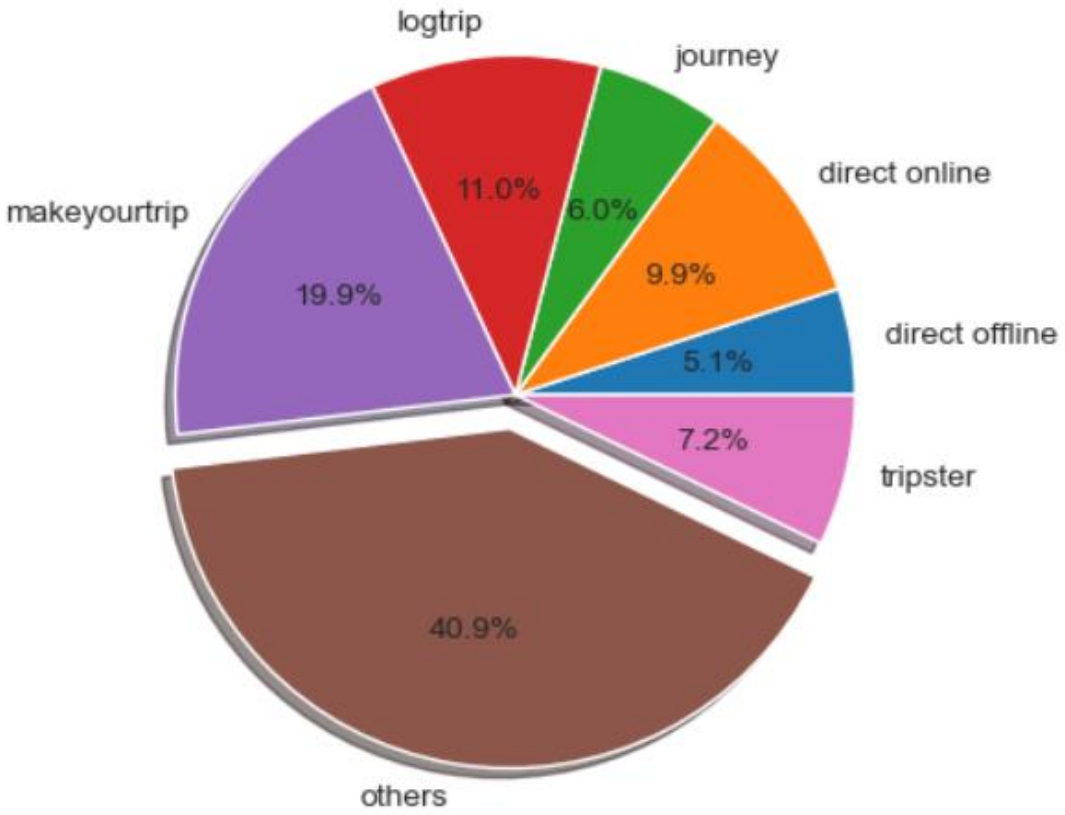
```
In [24]: avg = df_bookings[df_bookings['room_category'] == "RT4"]['revenue_realized'].mean()
         std =  df_bookings[df_bookings['room_category'] == "RT4"]['revenue_realized'].std()
         avg,std
```

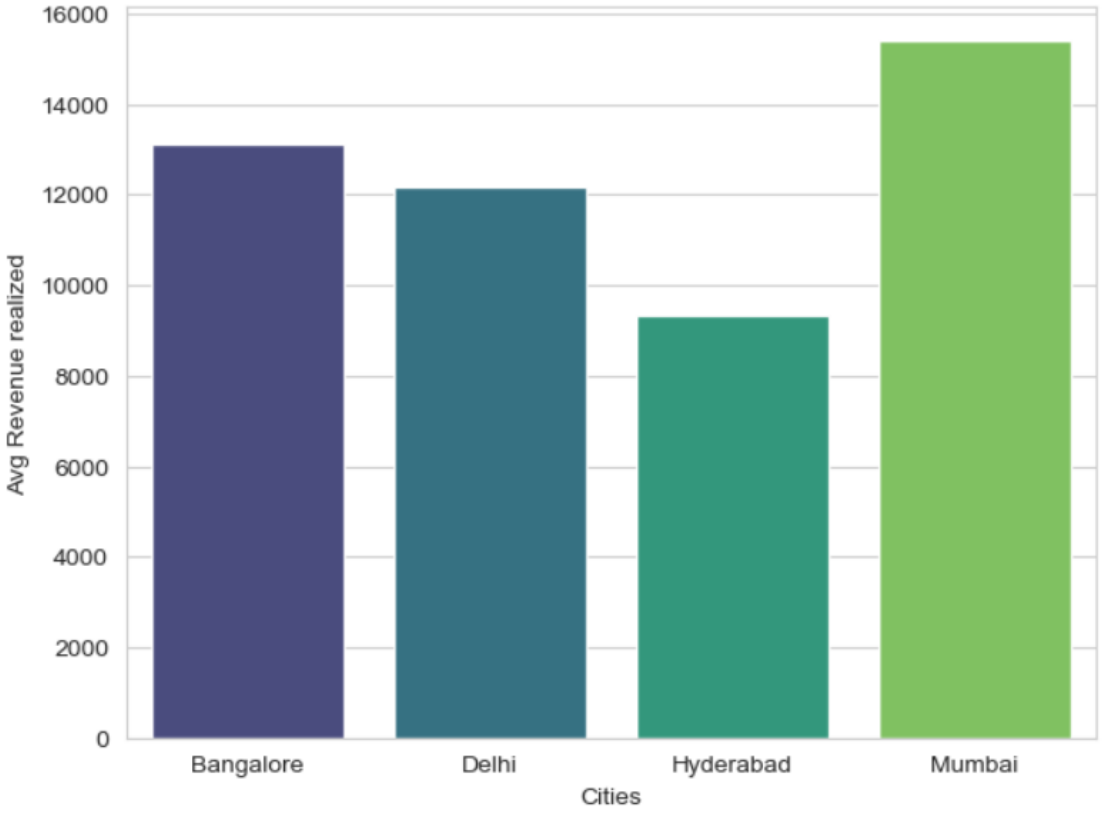
```
Out[24]: (23439.308443780723, 9048.599075739625)
```

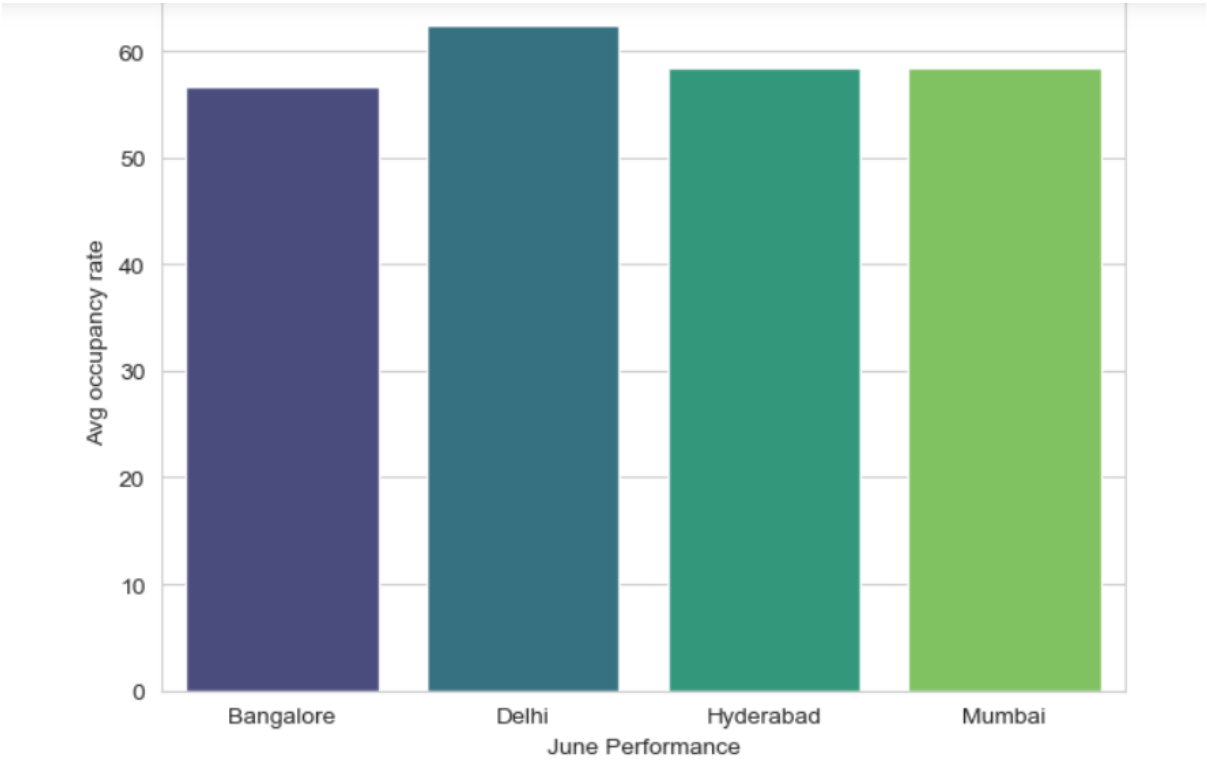
```
In [25]: high_rev_rls = avg + 3*std
         low_rev_rls = avg - 3*std
```

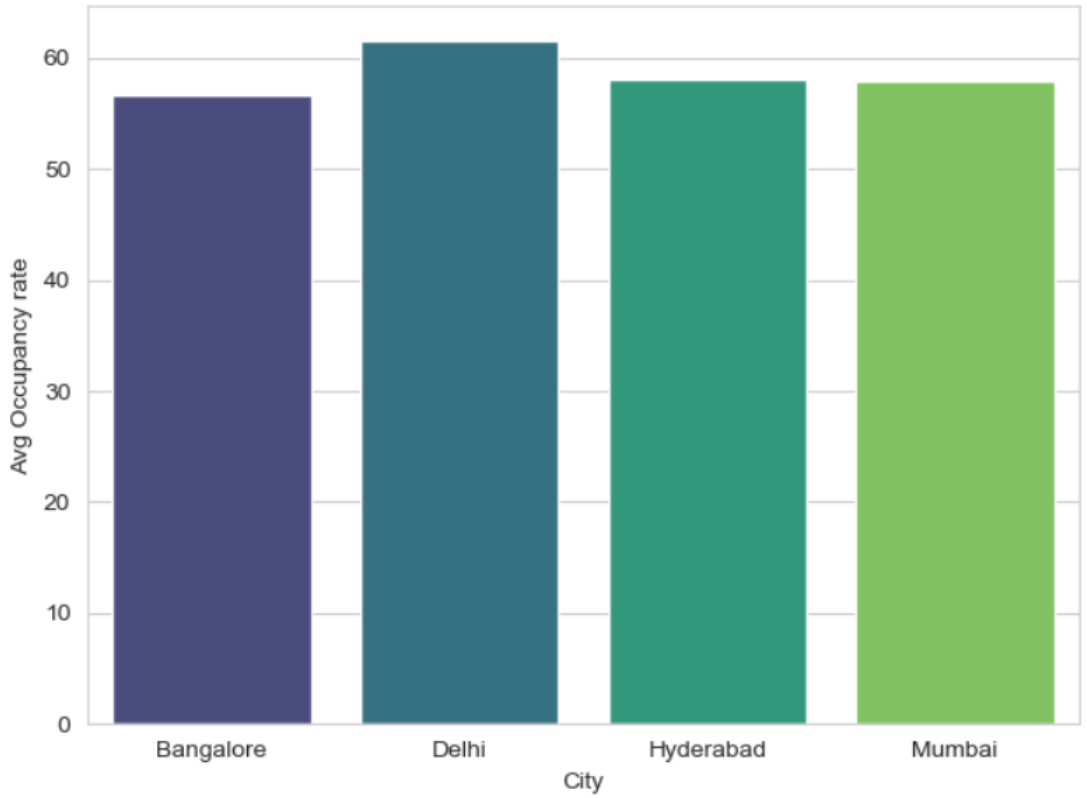
```
In [26]: df_bookings[df_bookings['revenue_realized'] > high_rev_rls]
```

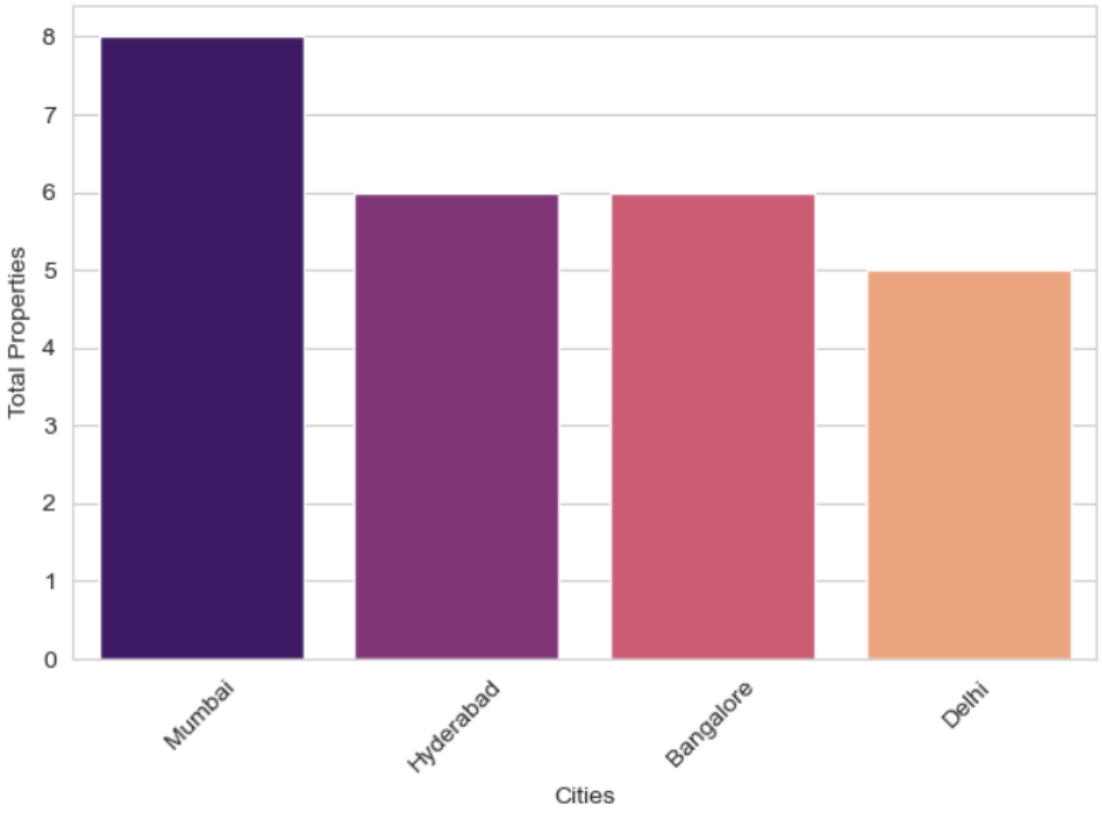
```
In [5]: import matplotlib.pyplot as  plt
        import seaborn as sns
        sns.set_style("whitegrid")
        plt.figure(figsize=(10, 6))
        a = df_bookings['booking_platform'].value_counts().sort_values(ascending=True)
        sns.barplot(x=a.index, y=a.values, palette='viridis')
        plt.xlabel('Booking Platform')
        plt.ylabel('Total Bookings')
        plt.xticks(rotation=45)
        plt.tight_layout()
        plt.show()
```

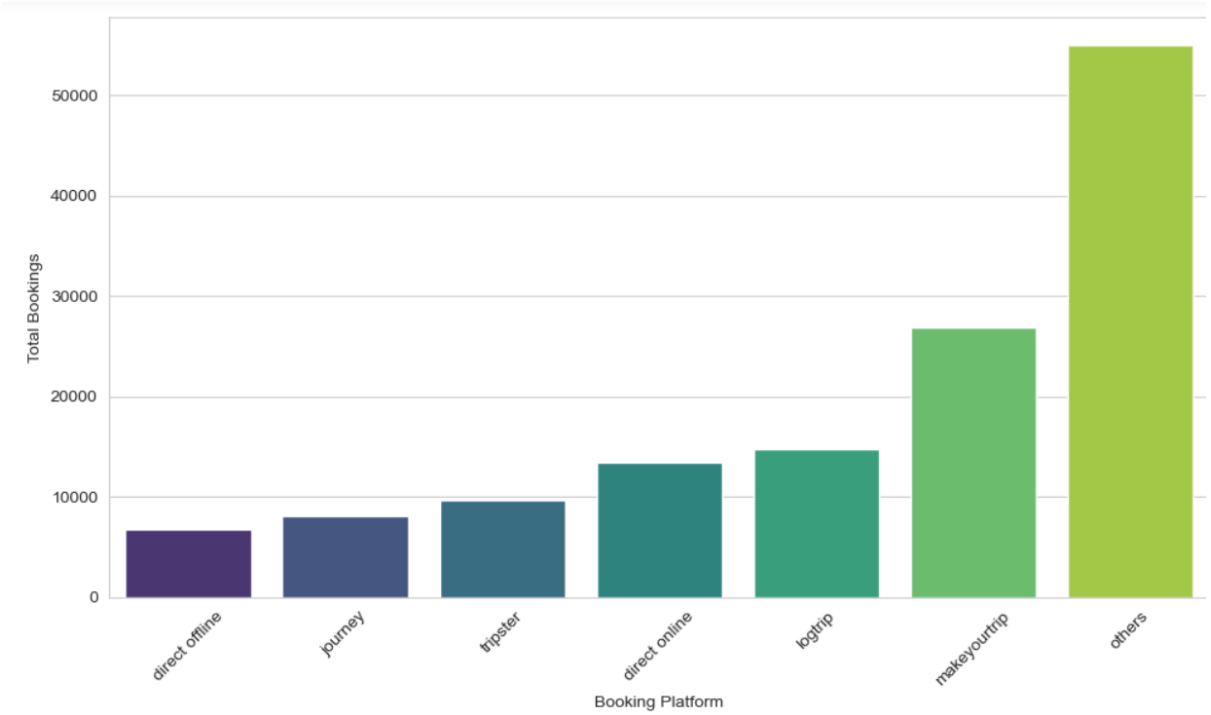



16000
14000
12000
10000
8000
6000
4000
2000
0
Avg Revenue realized
Bangalore
Delhi
Hyderabad
Mumbai
Cities


60
50
40
30
20
10
0
Avg occupancy rate
Bangalore
Delhi
Hyderabad
Mumbai
June Performance


Avg Occupancy rate
60
50
40
30
20
10
0
Bangalore
Delhi
Hyderabad
Mumbai
City


Total Properties
8
7
6
5
4
3
2
1
0
Mumbai
Hyderabad
Bangalore
Delhi
Cities


50000
40000
30000
20000
10000
0
Total Bookings
direct offline
journey
tripster
direct online
ibgtrip
makeyourtrip
others
Booking Platform