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

```
In [ ]: df=pd.read_excel("Hospitality.xlsx")
```

```
In [10]: df.shape
Out[10]: (50000, 14)
```

```
In [9]: df.head()
```

Out[9]:

	Avg Room Rate	reservation_id	check_in_date	stay_duration	adults	children	room_type	special_requests_flag	booking_channel	reservation_status	advanced_booking	Property	Date	Rate Type
0	71.10	779087-Y5-9824-SA	4/15/2020	13	3	4	Single	Yes	Call Center	Completed	Yes	The Chord	4/15/2020	Weekday
1	71.10	984023-QO-5015-YG	2020-12-06 00:00:00	3	3	3	Single	No	Travel Agent	Completed	Yes	The Chord	2020-12-06 00:00:00	Weekday
2	172.38	518066-UQ-2315-FK	2/25/2020	11	4	2	Queen	No	Call Center	Completed	Yes	The Sankey	2/25/2020	Weekday
3	172.38	130339-H9-2116-KE	9/15/2020	11	4	1	Queen	No	Website	No-Show	Yes	The Sankey	9/15/2020	Weekday
4	199.00	961051-40-0956-EO	2020-05-01 00:00:00	14	2	4	Double	Yes	Walk-in	Completed	No	The Sankey	2020-05-01 00:00:00	Weekend

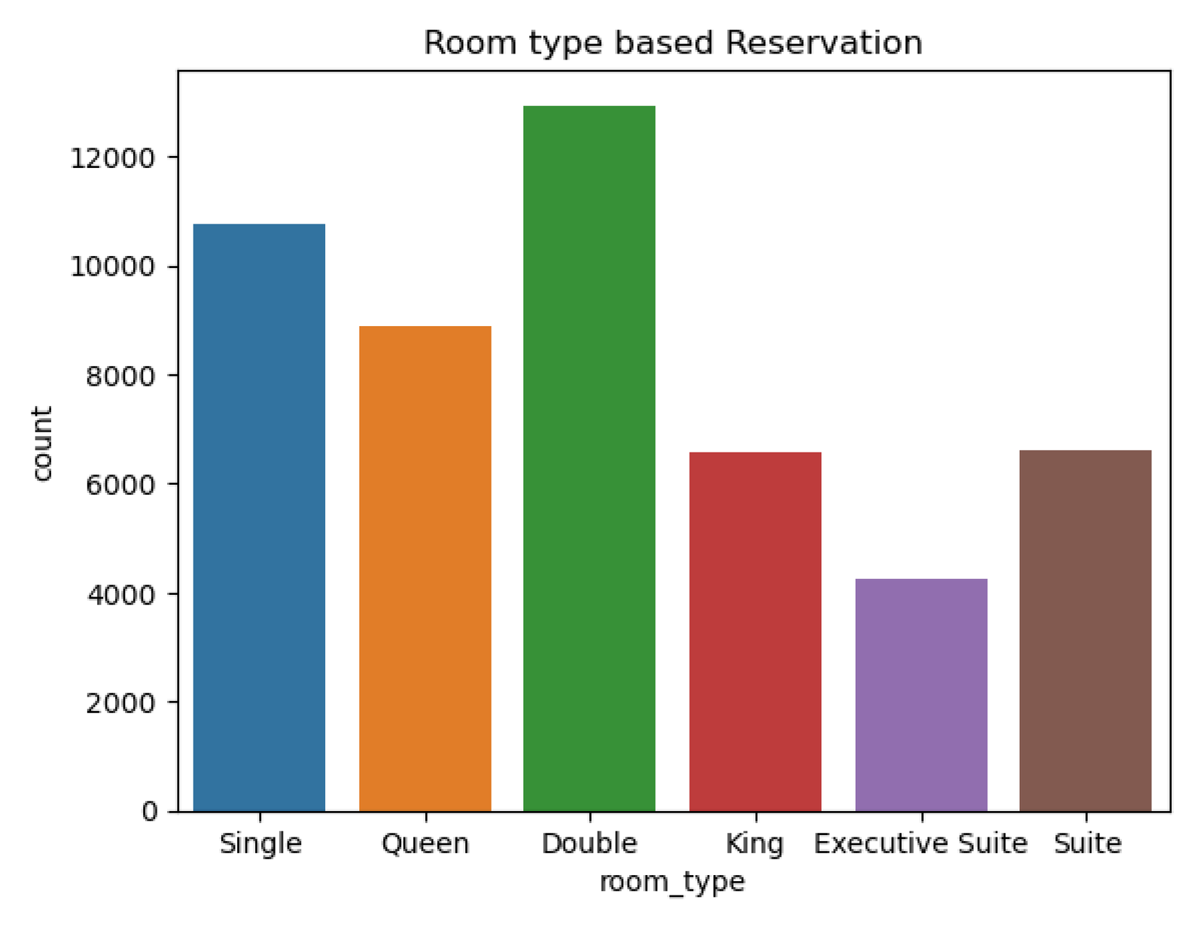
```
In [11]: df["Avg Room Rate"].describe()
Out[11]: count    50000.000000
         mean       147.147144
         std         48.316584
         min         71.100000
         25%        103.950000
         50%        152.100000
         75%        177.450000
         max        288.550000
         Name: Avg Room Rate, dtype: float64
```

```
In [12]: df.isnull().sum()
Out[12]: Avg Room Rate            0
         reservation_id           0
         check_in_date            0
         stay_duration            0
         adults                   0
         children                 0
         room_type                0
         special_requests_flag    0
         booking_channel          0
         reservation_status       0
         advanced_booking         0
         Property                 0
         Date                     0
         Rate Type                0
         dtype: int64
```

Room Type Analysis

```
In [13]: df["room_type"].value_counts()
Out[13]: Double            12936
         Single            10764
         Queen              8872
         Suite              6623
         King               6565
         Executive Suite    4240
         Name: room_type, dtype: int64
```

```
In [35]: sns.countplot(x="room_type",y=None,data=df)
         plt.title("Room type based Reservation")
Out[35]: Text(0.5, 1.0, 'Room type based Reservation')
```



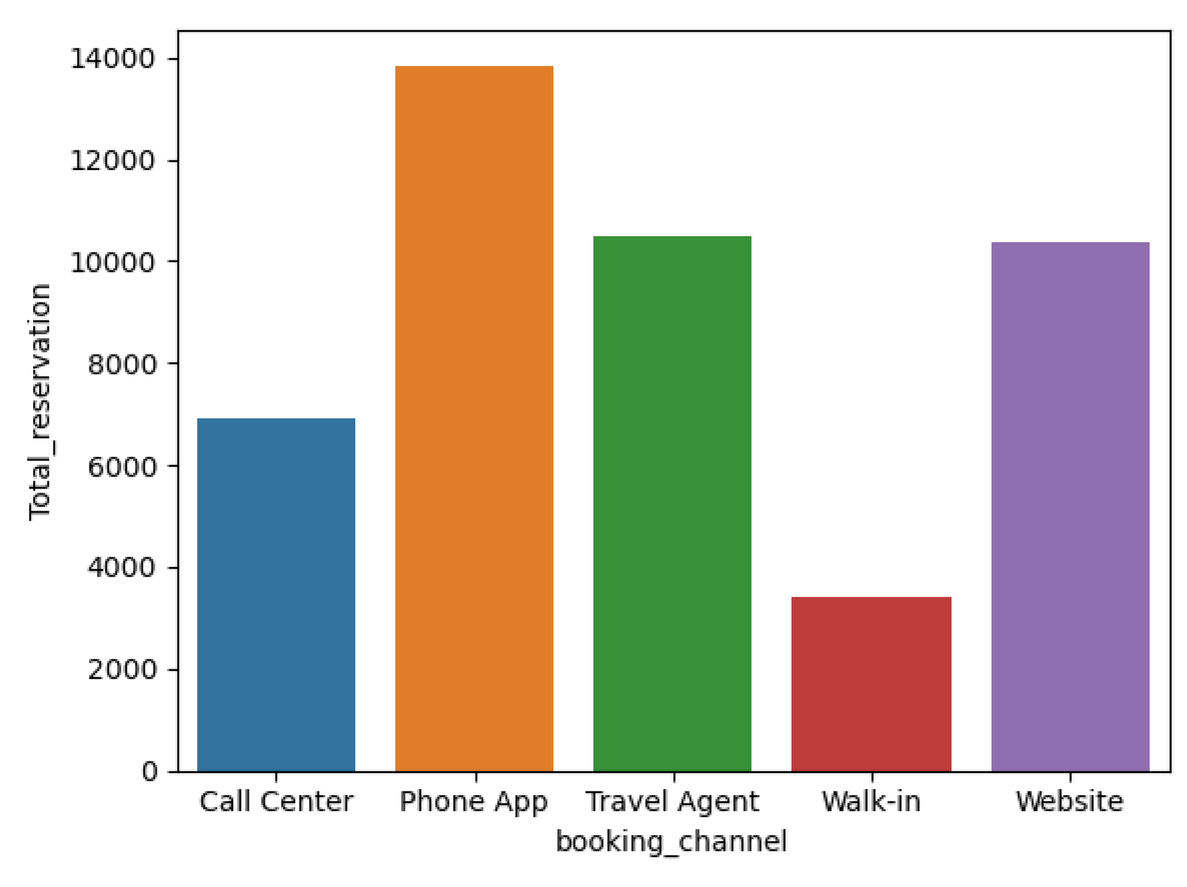
Booking Channel Insights

```
In [16]: df["booking_channel"].value_counts()
Out[16]: Phone App       15391
         Travel Agent    11574
         Website         11532
         Call Center      7700
         Walk-in          3803
         Name: booking_channel, dtype: int64
```

```
In [17]: booking_status=df[df["reservation_status"]=="Completed"].groupby("booking_channel").size().reset_index().rename(columns={0:"Total_reservation"})
```

```
In [36]: print(booking_status)
           booking_channel  Total_reservation
         0     Call Center               6920
         1       Phone App              13848
         2    Travel Agent              10488
         3         Walk-in               3400
         4         Website              10379
```

```
In [18]: #Based on 'Completed' reservation
         sns.barplot(x="booking_channel",y="Total_reservation",data=booking_status)
Out[18]: <Axes: xlabel='booking_channel', ylabel='Total_reservation'>
```



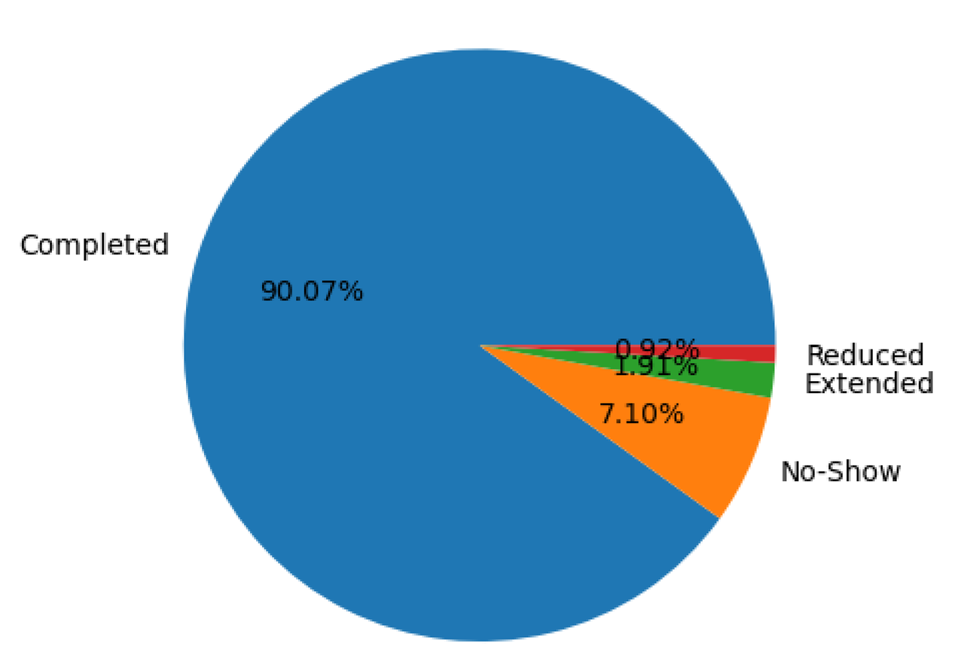
Observation:The most effective booking channel is Phone App & The least effective booking channel is Walk-in

Reservation Status Analysis

```
In [19]: df["reservation_status"].value_counts()
Out[19]: Completed    45035
         No-Show       3550
         Extended       953
         Reduced        462
         Name: reservation_status, dtype: int64
```

```
In [20]: values=df["reservation_status"].value_counts().values
         name=df["reservation_status"].value_counts().index
```

```
In [38]: plt.pie(values,labels=name,autopct="%1.2f%%")
Out[38]: ([<matplotlib.patches.Wedge at 0x1ec8226f950>,
           <matplotlib.patches.Wedge at 0x1ec82279350>,
           <matplotlib.patches.Wedge at 0x1ec8227b010>,
           <matplotlib.patches.Wedge at 0x1ec82280c10>],
          [Text(-1.0469071293725583, 0.33761733141074063, 'Completed'),
           Text(1.0127951594386826, -0.42923882049224504, 'No-Show'),
           Text(1.0923590271578676, -0.12942857407356972, 'Extended'),
           Text(1.0995365740026752, -0.03192682932048767, 'Reduced')],
          [Text(-0.5710402523850318, 0.18415490804222215, '90.07%'),
           Text(0.5524337233301905, -0.2341302657230427, '7.10%'),
           Text(0.5958321966315641, -0.07059740404012893, '1.91%'),
           Text(0.5997472221832774, -0.017414634174811452, '0.92%')])
```



Rate Type Comparison

```
In [58]: df[df["Rate Type"]=="Weekday"].groupby("Rate Type")["Avg Room Rate"].mean()
Out[58]: Rate Type
         Weekday    142.031001
         Name: Avg Room Rate, dtype: float64
```

```
In [59]: df[df["Rate Type"]=="Weekend"].groupby("Rate Type")["Avg Room Rate"].mean()
Out[59]: Rate Type
         Weekend    159.810204
         Name: Avg Room Rate, dtype: float64
```

Observation:There is a significant disparity between room rates on weekdays and weekends, with the average room rate being notably higher during weekends compared to weekdays.

Property Performance

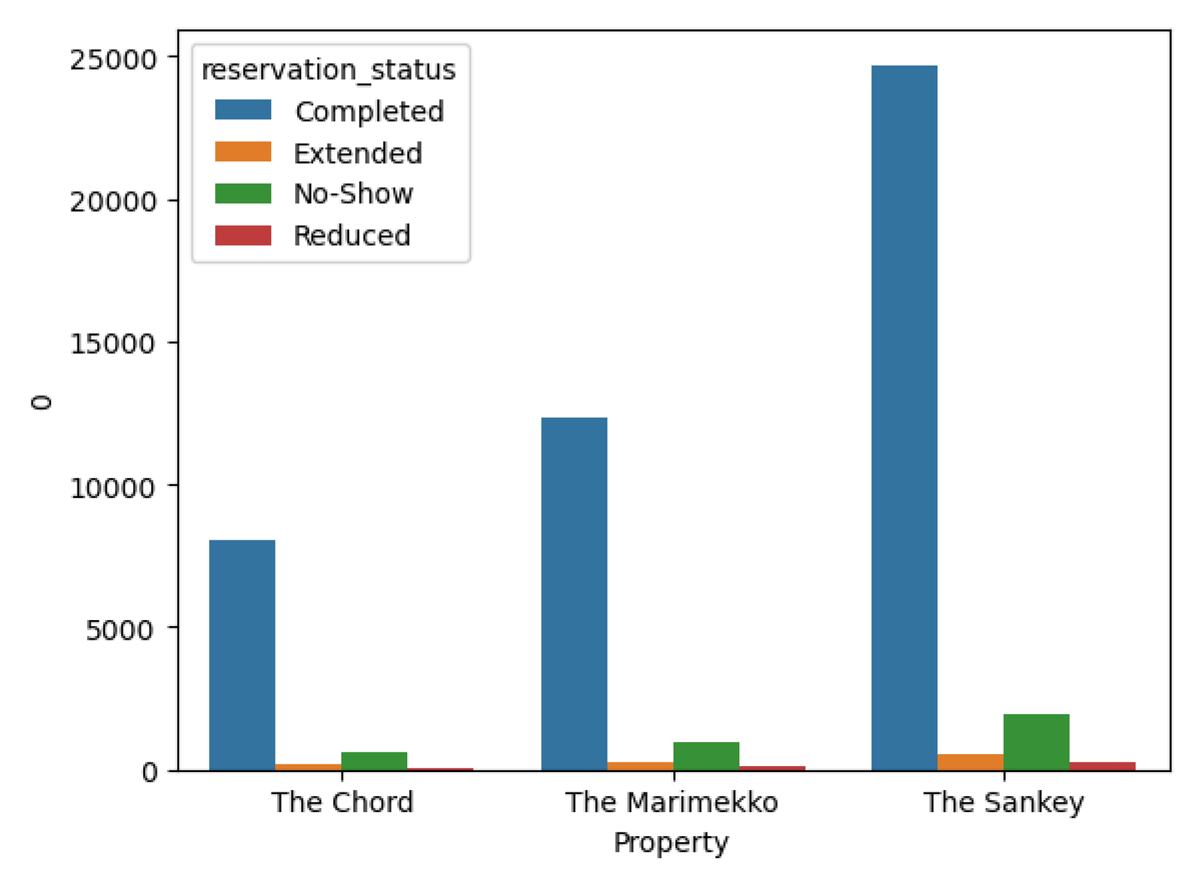
```
In [22]: df[["Property","reservation_status"]].groupby(["Property","reservation_status"]).size().reset_index()
```

Out[22]:

	Property	reservation_status	0
0	The Chord	Completed	8039
1	The Chord	Extended	165
2	The Chord	No-Show	646
3	The Chord	Reduced	84
4	The Marimekko	Completed	12298
5	The Marimekko	Extended	246
6	The Marimekko	No-Show	983
7	The Marimekko	Reduced	139
8	The Sankey	Completed	24698
9	The Sankey	Extended	542
10	The Sankey	No-Show	1921
11	The Sankey	Reduced	239

```
In [23]: prop=df[["Property","reservation_status"]].groupby(["Property","reservation_status"]).size().reset_index()
```

```
In [24]: sns.barplot(y=0,x="Property",data=prop,hue="reservation_status")
Out[24]: <Axes: xlabel='Property', ylabel='0'>
```




```
In [60]: df[df["Property"]=="The Chord"].groupby("Property")["Avg Room Rate"].mean()
Out[60]: Property
         The Chord    88.133999
         Name: Avg Room Rate, dtype: float64
```

```
In [61]: df[df["Property"]=="The Marimekko"].groupby("Property")["Avg Room Rate"].mean()
Out[61]: Property
         The Marimekko    110.153924
         Name: Avg Room Rate, dtype: float64
```

```
In [62]: df[df["Property"]=="The Sankey"].groupby("Property")["Avg Room Rate"].mean()
Out[62]: Property
         The Sankey    184.839581
         Name: Avg Room Rate, dtype: float64
```

Observation:The Sankey Property has the highest Avg Room rate

Advanced Booking Analysis

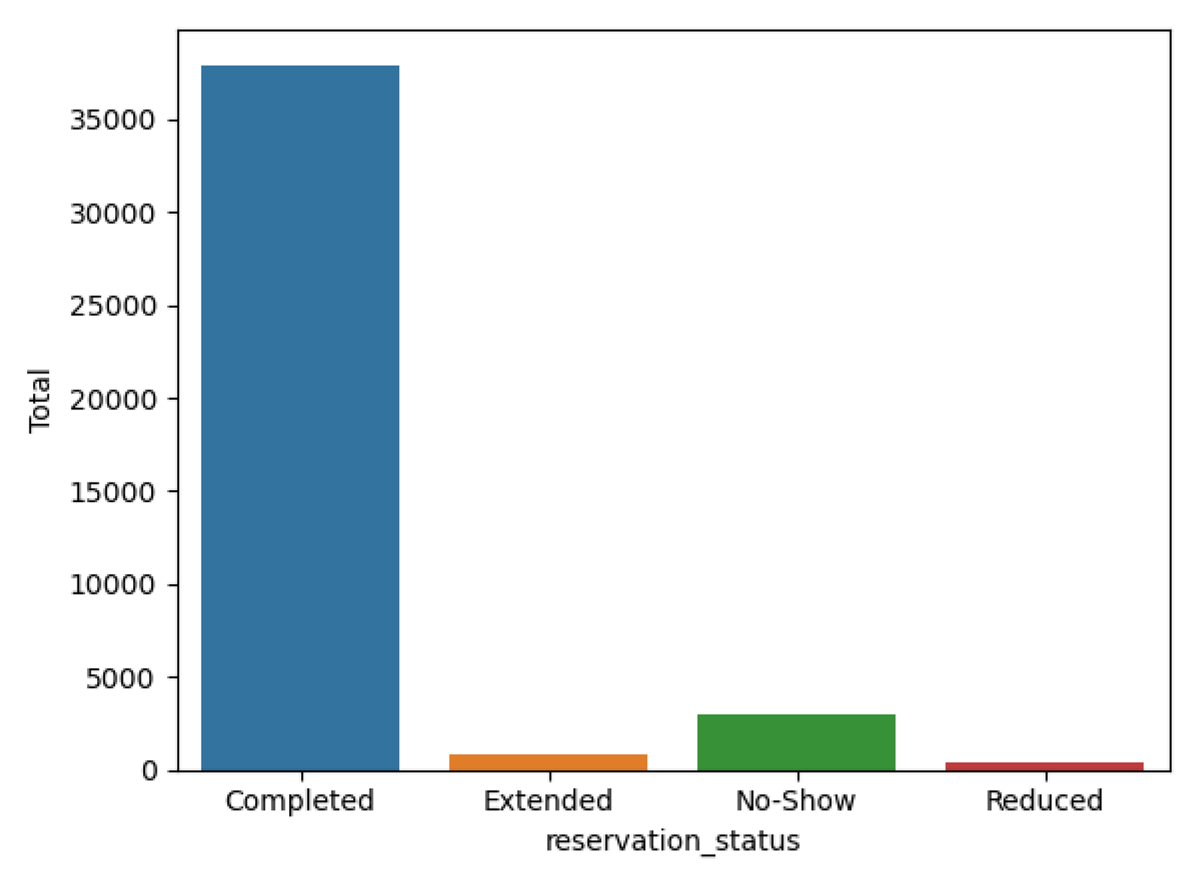
```
In [26]: df[df["advanced_booking"]=="Yes"].groupby("reservation_status").size().reset_index().rename(columns={0:"Total"})
```

Out[26]:

	reservation_status	Total
0	Completed	37945
1	Extended	773
2	No-Show	2985
3	Reduced	371

```
In [27]: df1=df[df["advanced_booking"]=="Yes"].groupby("reservation_status").size().reset_index().rename(columns={0:"Total"})
```

```
In [28]: sns.barplot(x="reservation_status",y="Total",data=df1)
Out[28]: <Axes: xlabel='reservation_status', ylabel='Total'>
```




```
In [39]: df[["advanced_booking","reservation_status"]].groupby(["advanced_booking","reservation_status"]).size().reset_index().rename(columns={0:"Total"})
```

Out[39]:

	advanced_booking	reservation_status	Total
0	No	Completed	7090
1	No	Extended	180
2	No	No-Show	565
3	No	Reduced	91
4	Yes	Completed	37945
5	Yes	Extended	773
6	Yes	No-Show	2985
7	Yes	Reduced	371

```
In [53]: df[df["advanced_booking"]=="Yes"].groupby("advanced_booking")["Avg Room Rate"].mean()
Out[53]: advanced_booking
         Yes    147.170475
         Name: Avg Room Rate, dtype: float64
```

```
In [54]: df[df["advanced_booking"]=="No"].groupby("advanced_booking")["Avg Room Rate"].mean()
Out[54]: advanced_booking
         No    147.023298
         Name: Avg Room Rate, dtype: float64
```

Observation:There is slight difference in advance booking rate & regular booking rate

Special Requests Impact

```
In [29]: df[df["special_requests_flag"]=="Yes"].groupby("reservation_status").size().reset_index().rename(columns={0:"Total"})
```

Out[29]:

	reservation_status	Total
0	Completed	22424
1	Extended	515
2	No-Show	1768
3	Reduced	233

```
In [30]: df[["special_requests_flag","reservation_status"]].groupby(["special_requests_flag","reservation_status"]).size().reset_index().rename(columns={0:"Total"})
```

Out[30]:

	special_requests_flag	reservation_status	Total
0	No	Completed	22611
1	No	Extended	438
2	No	No-Show	1782
3	No	Reduced	229
4	Yes	Completed	22424
5	Yes	Extended	515
6	Yes	No-Show	1768
7	Yes	Reduced	233

```
In [56]: df[df["special_requests_flag"]=="Yes"].groupby("special_requests_flag")["Avg Room Rate"].mean()
Out[56]: special_requests_flag
         Yes    147.056233
         Name: Avg Room Rate, dtype: float64
```

```
In [57]: df[df["special_requests_flag"]=="No"].groupby("special_requests_flag")["Avg Room Rate"].mean()
Out[57]: special_requests_flag
         No    147.237621
         Name: Avg Room Rate, dtype: float64
```

Observation:There is no such difference in Avg Room Rate with and without Special Request

```
In [ ]:
```