In [87]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from matplotlib import style
```

In [19]:
```
dim_date= pd.read_csv("dim_date.csv")
dim_hotels= pd.read_csv("dim_hotels.csv")
dim_room = pd.read_csv("dim_rooms.csv")
fact_aggregated_bookings= pd.read_csv("fact_aggregated_bookings.csv")
fact_bookings = pd.read_csv("fact_bookings.csv")
```

In [4]:
```
dim_date.head()
```

Out[4]:

	date	mmm yy	week no	day_type
0	01-May-22	May 22	W 19	weekend
1	02-May-22	May 22	W 19	weekeday
2	03-May-22	May 22	W 19	weekeday
3	04-May-22	May 22	W 19	weekeday
4	05-May-22	May 22	W 19	weekeday

In [5]:
```
dim_date.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 92 entries, 0 to 91
Data columns (total 4 columns):
 #   Column    Non-Null Count  Dtype 
---  ------    --------------  ----- 
 0   date      92 non-null     object
 1   mmm yy    92 non-null     object
 2   week no   92 non-null     object
 3   day_type  92 non-null     object
dtypes: object(4)
memory usage: 3.0+ KB
```

In [7]:
```
dim_date.shape
```

Out[7]: (92, 4)

In [8]:
```
dim_date.describe()
```

Out[8]:

	date	mmm yy	week no	day_type
count	92	92	92	92
unique	92	3	14	2
top	01-May-22	May 22	W 19	weekeday
freq	1	31	7	65

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

Out[9]:

	property_id	property_name	category	city
0	16558	Atliq Grands	Luxury	Delhi
1	16559	Atliq Exotica	Luxury	Mumbai
2	16560	Atliq City	Business	Delhi
3	16561	Atliq Blu	Luxury	Delhi
4	16562	Atliq Bay	Luxury	Delhi

In [10]:
```
dim_hotels.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 25 entries, 0 to 24
Data columns (total 4 columns):
 #   Column         Non-Null Count  Dtype 
---  ------         --------------  ----- 
 0   property_id    25 non-null     int64 
 1   property_name  25 non-null     object
 2   category       25 non-null     object
 3   city           25 non-null     object
dtypes: int64(1), object(3)
memory usage: 932.0+ bytes
```

In [11]:
```
dim_hotels.shape
```

Out[11]: (25, 4)

In [20]:
```
dim_room.head()
```

Out[20]:

	room_id	room_class
0	RT1	Standard
1	RT2	Elite
2	RT3	Premium
3	RT4	Presidential

In [21]:
```
dim_room.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4 entries, 0 to 3
Data columns (total 2 columns):
 #   Column      Non-Null Count  Dtype 
---  ------      --------------  ----- 
 0   room_id     4 non-null      object
 1   room_class  4 non-null      object
dtypes: object(2)
memory usage: 196.0+ bytes
```

In [22]:
```
dim_room.shape
```

Out[22]: (4, 2)

In [23]:
```
fact_aggregated_bookings.head()
```

Out[23]:

	property_id	check_in_date	room_category	successful_bookings	capacity
0	16559	01-May-22	RT1	25	30
1	19562	01-May-22	RT1	28	30
2	19563	01-May-22	RT1	23	30
3	17558	01-May-22	RT1	13	19
4	16558	01-May-22	RT1	18	19

In [24]:
```
fact_aggregated_bookings.shape
```

Out[24]: (9200, 5)

In [25]:
```
fact_bookings.head()
```

Out[25]:

	booking_id	property_id	booking_date	check_in_date	checkout_date	no_guests
0	May012216558RT11	16558	2022-04-27	2022-05-01	2022-05-02	3
1	May012216558RT12	16558	2022-04-30	2022-05-01	2022-05-02	2
2	May012216558RT13	16558	2022-04-28	2022-05-01	2022-05-04	2
3	May012216558RT14	16558	2022-04-28	2022-05-01	2022-05-02	2
4	May012216558RT15	16558	2022-04-27	2022-05-01	2022-05-02	4

In [26]:
```
fact_bookings.shape
```

Out[26]: (134590, 12)

In [27]:
```
fact_bookings.describe()
```

Out[27]:

	property_id	no_guests	ratings_given	revenue_generated	revenue_realized
count	134590.000000	134590.000000	56683.000000	134590.000000	134590.000000
mean	18061.113493	2.036808	3.619004	14916.013188	12696.123256
std	1093.055847	1.031766	1.235009	6452.868072	6928.108124
min	16558.000000	1.000000	1.000000	6500.000000	2600.000000
25%	17558.000000	1.000000	3.000000	9900.000000	7600.000000
50%	17564.000000	2.000000	4.000000	13500.000000	11700.000000
75%	18563.000000	2.000000	5.000000	18000.000000	15300.000000
max	19563.000000	6.000000	5.000000	45220.000000	45220.000000

In [28]:
```
fact_bookings.isnull().sum()
```

Out[28]:
```
booking_id              0
property_id             0
booking_date            0
check_in_date           0
checkout_date           0
no_guests               0
room_category           0
booking_platform        0
ratings_given       77907
booking_status          0
revenue_generated       0
revenue_realized        0
dtype: int64
```

In [29]:
```
dim_date['date']=pd.to_datetime(dim_date['date'])
fact_aggregated_bookings['check_in_date']=pd.to_datetime(fact_aggregated_bookings['
fact_bookings['booking_date']=pd.to_datetime(fact_bookings['booking_date'])
fact_bookings['check_in_date']=pd.to_datetime(fact_bookings['check_in_date'])
fact_bookings['checkout_date']=pd.to_datetime(fact_bookings['checkout_date'])
```

```
C:\Users\nagpa\AppData\Local\Temp\ipykernel_23132\3222500642.py:1: UserWarning: Coul
d not infer format, so each element will be parsed individually, falling back to `da
teutil`. To ensure parsing is consistent and as-expected, please specify a format.
  dim_date['date']=pd.to_datetime(dim_date['date'])
C:\Users\nagpa\AppData\Local\Temp\ipykernel_23132\3222500642.py:2: UserWarning: Coul
d not infer format, so each element will be parsed individually, falling back to `da
teutil`. To ensure parsing is consistent and as-expected, please specify a format.
  fact_aggregated_bookings['check_in_date']=pd.to_datetime(fact_aggregated_bookings
['check_in_date'])
```

In [30]:
```
fact_bookings['ratings_given']=fact_bookings['ratings_given'].fillna(0)
fact_bookings.isna().sum()
```

Out[30]:
```
booking_id           0
property_id          0
booking_date         0
check_in_date        0
checkout_date        0
no_guests            0
room_category        0
booking_platform     0
ratings_given        0
booking_status       0
revenue_generated    0
revenue_realized     0
dtype: int64
```

Revenue Analysis

In [31]:
```
revenue_df=pd.merge(dim_hotels,fact_bookings,how='left',on='property_id')
revenue_df.head()
```

Out[31]:

	property_id	property_name	category	city	booking_id	booking_date	check_i
0	16558	Atliq Grands	Luxury	Delhi	May012216558RT11	2022-04-27	2022
1	16558	Atliq Grands	Luxury	Delhi	May012216558RT12	2022-04-30	2022
2	16558	Atliq Grands	Luxury	Delhi	May012216558RT13	2022-04-28	2022
3	16558	Atliq Grands	Luxury	Delhi	May012216558RT14	2022-04-28	2022
4	16558	Atliq Grands	Luxury	Delhi	May012216558RT15	2022-04-27	2022

In [32]:
```
# Hotelwise Revenue
hotel_revenue=revenue_df.groupby(['property_name','city']).agg({'revenue_realized':
hotel_revenue.reset_index(inplace=True)
hotel_revenue['Revenue in Millions']=hotel_revenue['Revenue in Millions']/1000000
hotel_revenue['Revenue in Millions']=hotel_revenue['Revenue in Millions'].round(2)
hotel_revenue
```

Out[32]:

	property_name	city	Revenue in Millions
0	Atliq Bay	Bangalore	82.44
1	Atliq Bay	Delhi	56.44
2	Atliq Bay	Hyderabad	69.26
3	Atliq Bay	Mumbai	51.91
4	Atliq Blu	Bangalore	72.96
5	Atliq Blu	Delhi	57.93
6	Atliq Blu	Hyderabad	56.04
7	Atliq Blu	Mumbai	73.92
8	Atliq City	Bangalore	81.88
9	Atliq City	Delhi	54.93
10	Atliq City	Hyderabad	61.01
11	Atliq City	Mumbai	88.00
12	Atliq Exotica	Bangalore	60.02
13	Atliq Exotica	Hyderabad	47.84
14	Atliq Exotica	Mumbai	212.44
15	Atliq Grands	Bangalore	54.49
16	Atliq Grands	Delhi	36.06
17	Atliq Grands	Hyderabad	46.25
18	Atliq Grands	Mumbai	74.73
19	Atliq Palace	Bangalore	68.60
20	Atliq Palace	Delhi	89.14
21	Atliq Palace	Hyderabad	44.84
22	Atliq Palace	Mumbai	101.51
23	Atliq Seasons	Mumbai	66.13

In [34]:
```
revenue_pivot=hotel_revenue.pivot(index='city',columns='property_name',values='Reve
revenue_pivot
```

Out[34]:

property_name city	Atliq Bay	Atliq Blu	Atliq City	Atliq Exotica	Atliq Grands	Atliq Palace	Atliq Seasons
Bangalore	82.44	72.96	81.88	60.02	54.49	68.60	NaN
Delhi	56.44	57.93	54.93	NaN	36.06	89.14	NaN
Hyderabad	69.26	56.04	61.01	47.84	46.25	44.84	NaN
Mumbai	51.91	73.92	88.00	212.44	74.73	101.51	66.13

In [35]:
```
revenue_pivot.plot(kind='bar')
plt.ylabel('Revenue In Millions')
plt.title('City & Hotel Wise Revenue')
```

Out[35]: Text(0.5, 1.0, 'City & Hotel Wise Revenue')

Highest Revenue Generated city is Mumbai , property name = Atliq Exotica

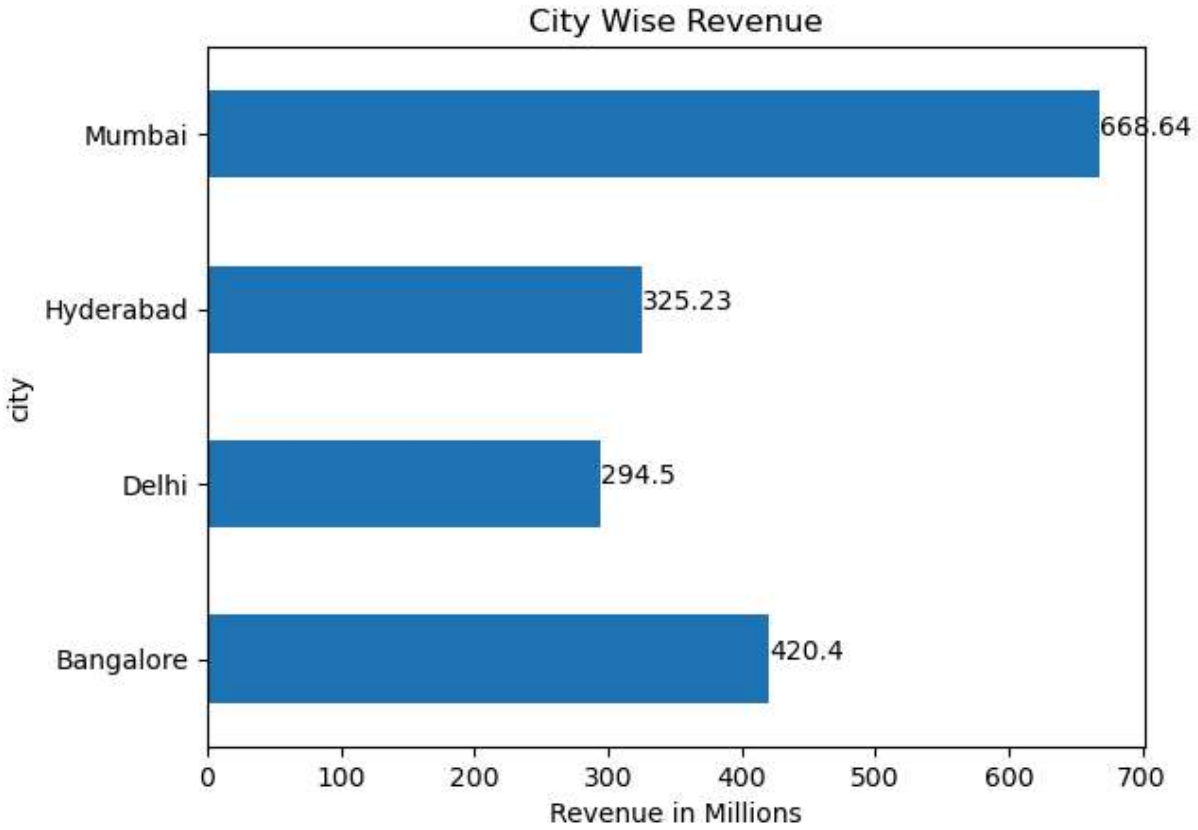
In [36]:
```
# Citywise Revenue In Millions
city_revenue=revenue_df.groupby(['city']).agg({'revenue_realized':'sum'}).rename(co
```

```
city_revenue['Revenue in Millions']=city_revenue['Revenue in Millions']/1000000
city_revenue['Revenue in Millions']=city_revenue['Revenue in Millions'].round(2)
city_revenue.sort_values(by='Revenue in Millions',ascending=False)
```

Out[36]:

city	Revenue in Millions
Mumbai	668.64
Bangalore	420.40
Hyderabad	325.23
Delhi	294.50

In [37]:
```
ax=city_revenue.plot(kind='barh',legend=False)
plt.xlabel('Revenue in Millions')
plt.title('City Wise Revenue')
for index, value in enumerate(city_revenue['Revenue in Millions']):
    ax.text(value, index, str(value))
plt.show()
```



In [38]:
```
revenue_tr=pd.merge(dim_date,revenue_df,how='left',left_on='date',right_on='check_i
revenue_trend=revenue_tr.groupby(['week no','property_name']).agg({'revenue_realize
revenue_trend['Revenue in Millions']=revenue_trend['Revenue in Millions']/1000000
revenue_trend['Revenue in Millions']=revenue_trend['Revenue in Millions'].round(2)
revenue_trend.reset_index(inplace=True)
revenue_trend
```

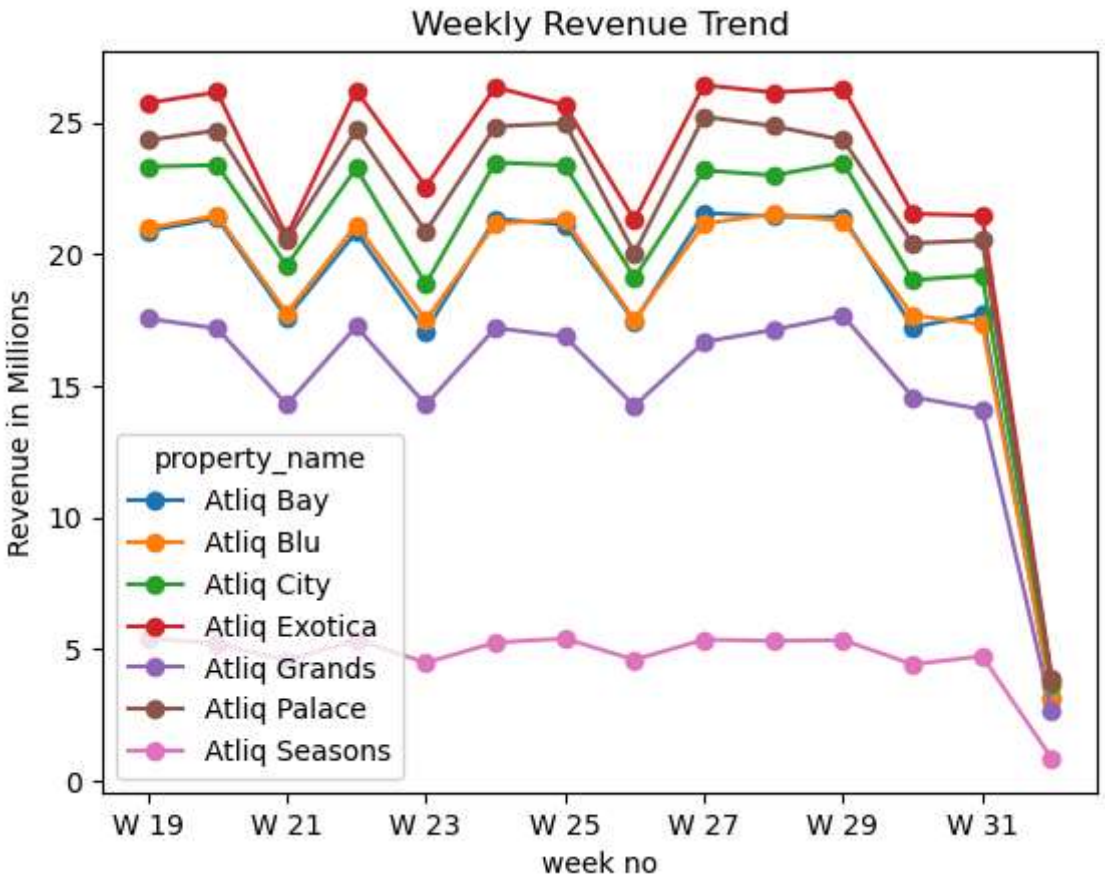
Out[38]:

	week no	property_name	Revenue in Millions
0	W 19	Atliq Bay	20.87
1	W 19	Atliq Blu	20.98
2	W 19	Atliq City	23.32
3	W 19	Atliq Exotica	25.74
4	W 19	Atliq Grands	17.55
...	...	...	...
93	W 32	Atliq City	3.66
94	W 32	Atliq Exotica	3.85
95	W 32	Atliq Grands	2.63
96	W 32	Atliq Palace	3.79
97	W 32	Atliq Seasons	0.83

98 rows × 3 columns

In [39]:
```
# Weekly Hotelwise Revenue Trend
pivot_data=revenue_trend.pivot(index='week no',columns='property_name',values='Reve
pivot_data.plot(kind='line',marker='o')
plt.ylabel('Revenue in Millions')
plt.title('Weekly Revenue Trend')
```

Out[39]:
```
Text(0.5, 1.0, 'Weekly Revenue Trend')
```



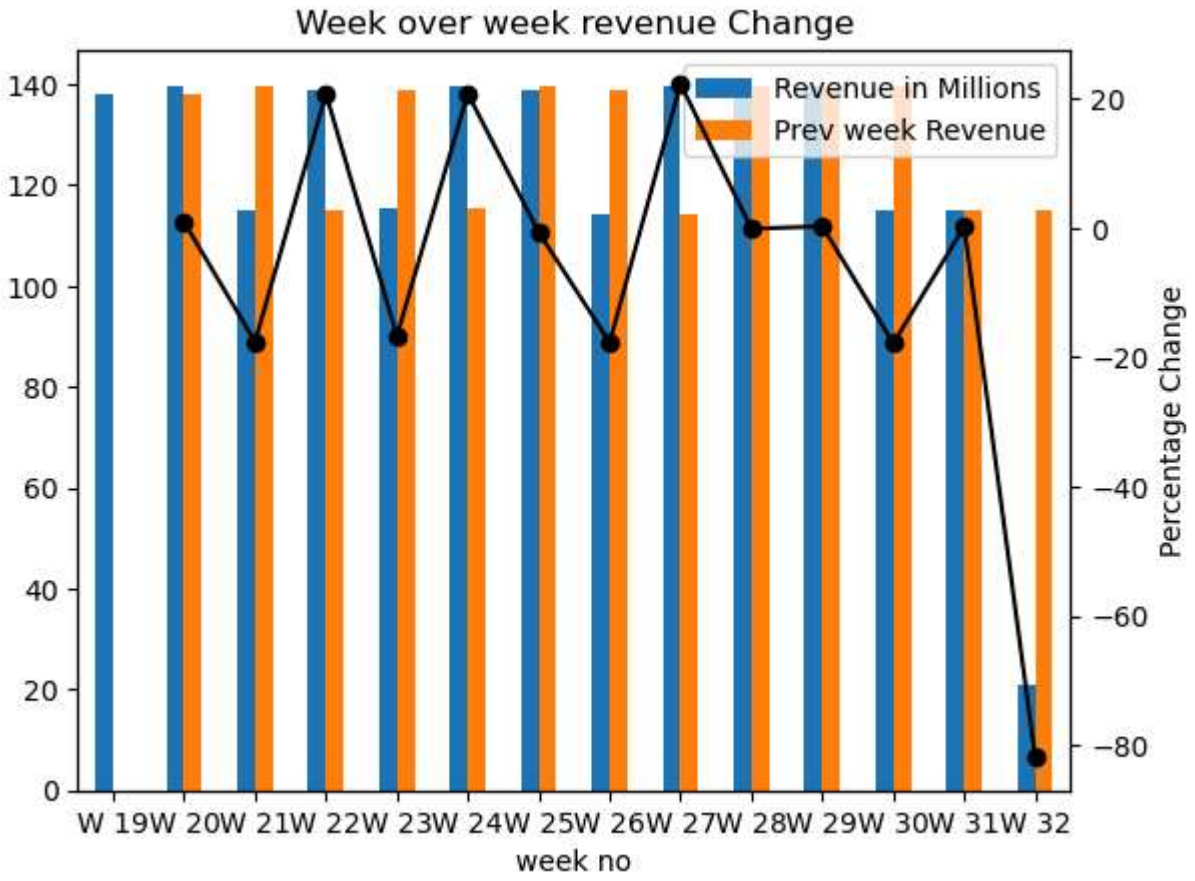
In [40]:
```
# Week over Week Revenue Trend
atliq_revenue_trend=revenue_tr.groupby(['week no']).agg({'revenue_realized':'sum'})
atliq_revenue_trend['Revenue in Millions']=atliq_revenue_trend['Revenue in Millions
atliq_revenue_trend['Revenue in Millions']=atliq_revenue_trend['Revenue in Millions
atliq_revenue_trend['Prev week Revenue']=atliq_revenue_trend['Revenue in Millions']
atliq_revenue_trend['Change Percentage']=((atliq_revenue_trend['Revenue in Millions
atliq_revenue_trend
```

Out[40]:

week no	Revenue in Millions	Prev week Revenue	Change Percentage
W 19	138.18	NaN	NaN
W 20	139.44	138.18	0.911854
W 21	114.92	139.44	-17.584624
W 22	138.72	114.92	20.710059
W 23	115.57	138.72	-16.688293
W 24	139.58	115.57	20.775288
W 25	138.67	139.58	-0.651956
W 26	114.15	138.67	-17.682267
W 27	139.56	114.15	22.260184
W 28	139.38	139.56	-0.128977
W 29	139.73	139.38	0.251112
W 30	114.81	139.73	-17.834395
W 31	115.04	114.81	0.200331
W 32	21.01	115.04	-81.736787

In [41]:
```
atliq_revenue_trend[['Revenue in Millions','Prev week Revenue']].plot(kind='bar')
atliq_revenue_trend['Change Percentage'].plot(secondary_y=True,color='black',marker
plt.ylabel('Percentage Change')
plt.title('Week over week revenue Change')
```

Out[41]:
```
Text(0.5, 1.0, 'Week over week revenue Change')
```



Booking Analysis

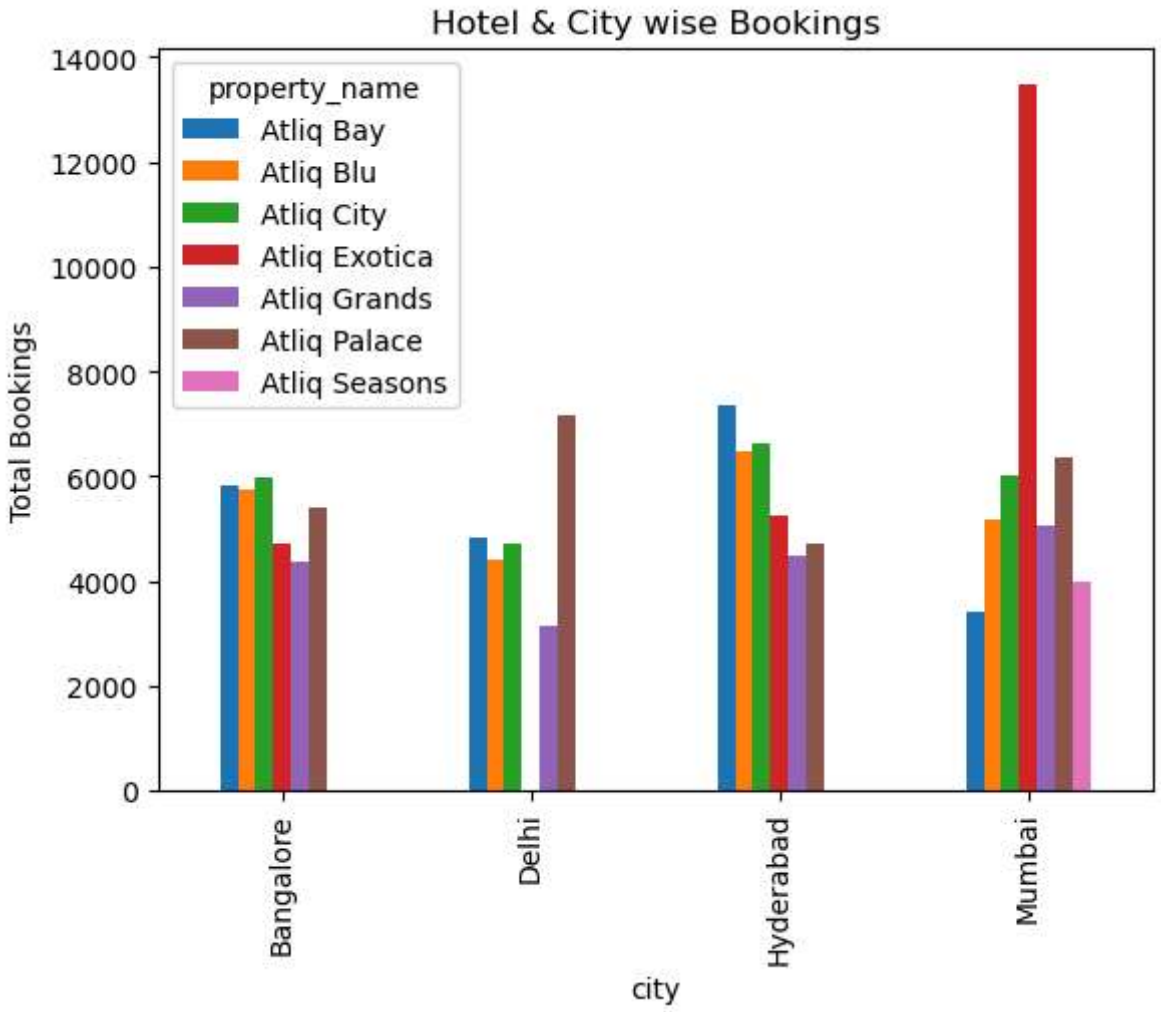
In [42]:
```
# Hotelwise Bookings
hotel_bookings=revenue_df.groupby(['property_name','city']).agg({'booking_id':'nuni
hotel_bookings.reset_index(inplace=True)
hotel_bookings
```

Out[42]:

	property_name	city	Total bookings
0	Atliq Bay	Bangalore	5812
1	Atliq Bay	Delhi	4820
2	Atliq Bay	Hyderabad	7333
3	Atliq Bay	Mumbai	3424
4	Atliq Blu	Bangalore	5736
5	Atliq Blu	Delhi	4418
6	Atliq Blu	Hyderabad	6458
7	Atliq Blu	Mumbai	5183
8	Atliq City	Bangalore	5979
9	Atliq City	Delhi	4693
10	Atliq City	Hyderabad	6638
11	Atliq City	Mumbai	6013
12	Atliq Exotica	Bangalore	4705
13	Atliq Exotica	Hyderabad	5256
14	Atliq Exotica	Mumbai	13480
15	Atliq Grands	Bangalore	4371
16	Atliq Grands	Delhi	3153
17	Atliq Grands	Hyderabad	4475
18	Atliq Grands	Mumbai	5036
19	Atliq Palace	Bangalore	5413
20	Atliq Palace	Delhi	7147
21	Atliq Palace	Hyderabad	4728
22	Atliq Palace	Mumbai	6337
23	Atliq Seasons	Mumbai	3982

In [43]:
```
booking_pivot=hotel_bookings.pivot(index='city',columns='property_name',values='Tot
booking_pivot.plot(kind='bar')
plt.ylabel('Total Bookings')
plt.title('Hotel & City wise Bookings')
```

Out[43]:
```
Text(0.5, 1.0, 'Hotel & City wise Bookings')
```



In [44]:
```
# Hotelwise cancellation percentage, No Show Percentage, Check out percentage
df1=revenue_df.groupby(['property_name','booking_status']).agg({'booking_id':'nuniq
df2=revenue_df.groupby(['property_name']).agg({'booking_id':'nunique'}).rename(colu
df3=pd.merge(df1,df2,how='inner',on='property_name')
df3
```

Out[44]:

property_name	Cancelled	Checked Out	No Show	Toatal Bookings
Atliq Bay	5314	14965	1110	21389
Atliq Blu	5373	15267	1155	21795
Atliq City	5811	16365	1147	23323
Atliq Exotica	5713	16557	1171	23441
Atliq Grands	4273	11914	848	17035
Atliq Palace	5949	16532	1144	23625
Atliq Seasons	987	2811	184	3982

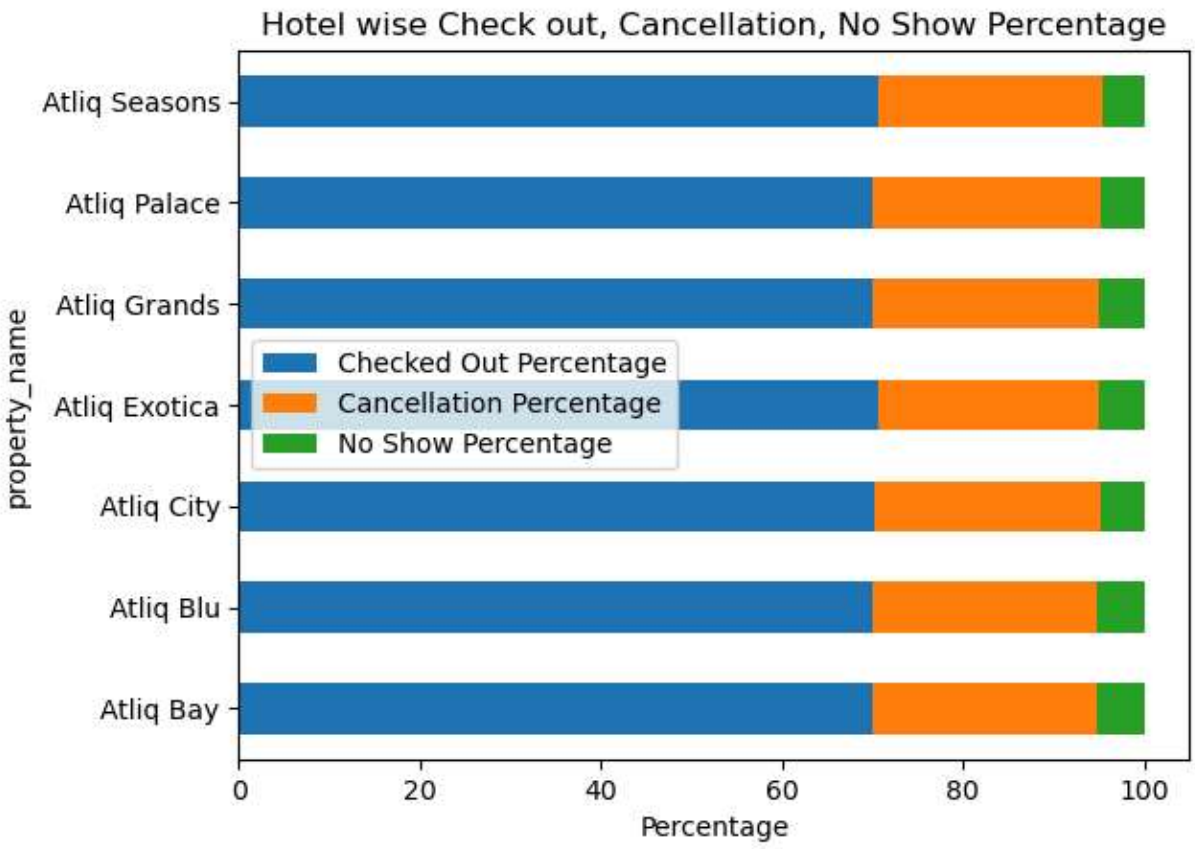
In [45]:
```
df3['Cancellation Percentage']=df3['Cancelled']*100/df3['Toatal Bookings']
df3['Checked Out Percentage']=df3['Checked Out']*100/df3['Toatal Bookings']
df3['No Show Percentage']=df3['No Show']*100/df3['Toatal Bookings']
df4=df3[['Checked Out Percentage','Cancellation Percentage','No Show Percentage']]
df4
```

Out[45]:

property_name	Checked Out Percentage	Cancellation Percentage	No Show Percentage
Atliq Bay	69.965870	24.844546	5.189583
Atliq Blu	70.048176	24.652443	5.299381
Atliq City	70.166788	24.915320	4.917892
Atliq Exotica	70.632652	24.371827	4.995521
Atliq Grands	69.938362	25.083651	4.977986
Atliq Palace	69.976720	25.180952	4.842328
Atliq Seasons	70.592667	24.786539	4.620794

In [88]:
```
df4.plot(kind='barh',stacked=True)
plt.xlabel('Percentage')
plt.title('Hotel wise Check out, Cancellation, No Show Percentage')
```

Out[88]:
```
Text(0.5, 1.0, 'Hotel wise Check out, Cancellation, No Show Percentage')
```

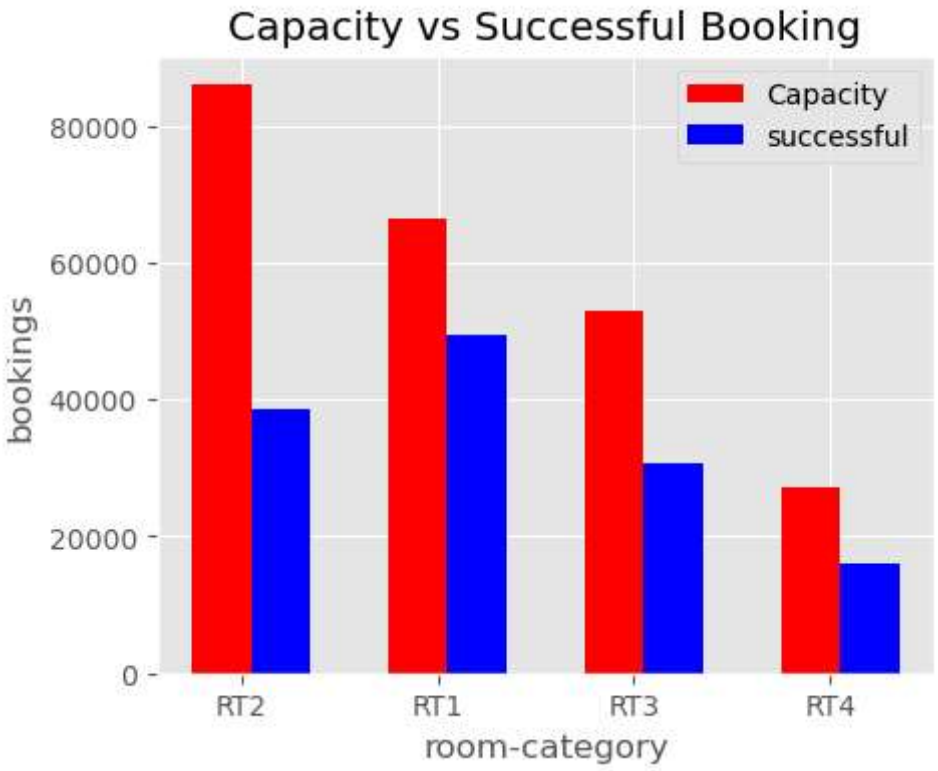


In [92]:
```
success=fact_aggregated_bookings.groupby("room_category").successful_bookings.sum()
capacity=fact_aggregated_bookings.groupby("room_category").capacity.sum()
capacity=fact_aggregated_bookings.groupby("room_category").capacity.sum().sort_valu
xplot=np.arange(4)
xplot
```

Out[92]:
```
array([0, 1, 2, 3])
```

In [93]:
```
style.use("ggplot")
plt.figure(figsize=(5,4))
plt.bar(xplot,capacity,width=0.3,color="red",label="Capacity")
plt.bar(xplot+0.3,success,width=0.3,color="blue",label="successful")
plt.xticks(xplot+0.1,["RT2","RT1","RT3","RT4"])
plt.xlabel("room-category")
plt.ylabel("bookings")
plt.title("Capacity vs Successful Booking")
plt.legend()
```

Out[93]:
```
<matplotlib.legend.Legend at 0x206d2efc590>
```



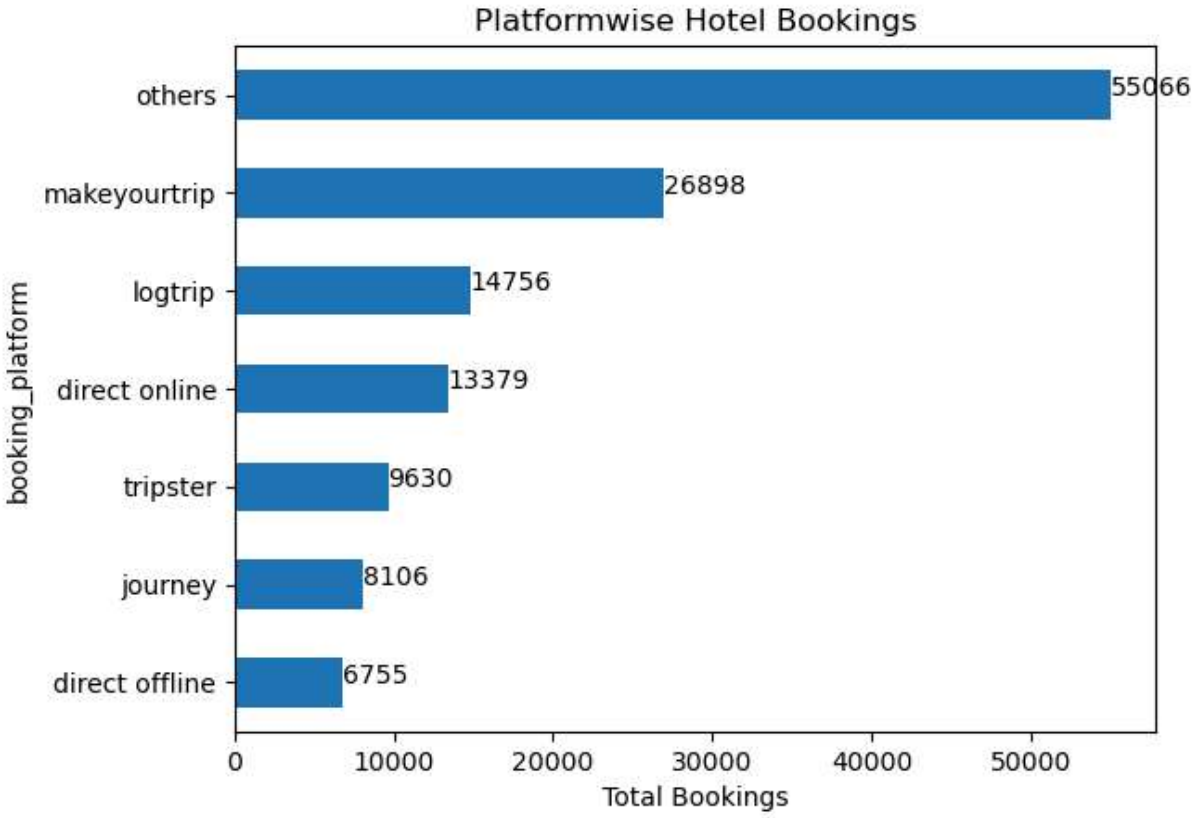
In [47]:
```
# booking platform wise
platform_df=revenue_df.groupby(['booking_platform']).agg({'booking_id':'nunique'}).
platform_df.sort_values(by='Total Bookings',ascending=False)
```

Out[47]:

booking_platform	Total Bookings
others	55066
makeyourtrip	26898
logtrip	14756
direct online	13379
tripster	9630
journey	8106
direct offline	6755

In [48]:
```
platform_df=platform_df.sort_values(by='Total Bookings')
```

In [49]:
```
ax=platform_df.plot(kind='barh',legend=False)
plt.xlabel('Total Bookings')
plt.title('Platformwise Hotel Bookings')
for index, value in enumerate(platform_df['Total Bookings']):
    ax.text(value, index, str(value))
plt.show()
```



In [51]:
```
rooms_df=pd.merge(dim_room,fact_bookings,how='left',left_on='room_id',right_on='roo
rooms_df.head()
```

Out[51]:

	room_id	room_class	booking_id	property_id	booking_date	check_in_date	che
0	RT1	Standard	May012216558RT11	16558	2022-04-27	2022-05-01	
1	RT1	Standard	May012216558RT12	16558	2022-04-30	2022-05-01	
2	RT1	Standard	May012216558RT13	16558	2022-04-28	2022-05-01	
3	RT1	Standard	May012216558RT14	16558	2022-04-28	2022-05-01	
4	RT1	Standard	May012216558RT15	16558	2022-04-27	2022-05-01	

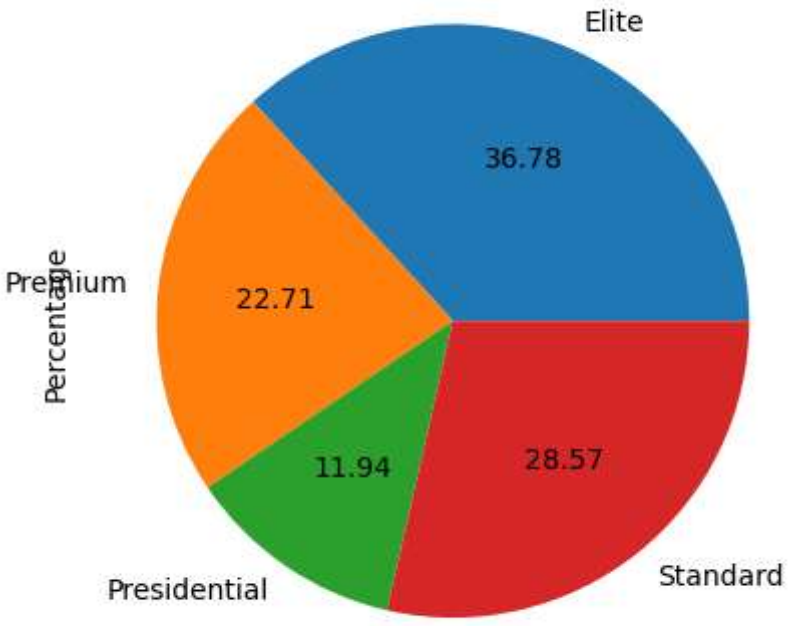
In [53]:
```
# Room Categorywise Bookings
room_bookings=rooms_df.groupby(['room_class']).agg({'booking_id':'nunique'})
room_bookings['Percentage']=room_bookings['booking_id']*100/room_bookings['booking_
room_bookings['Percentage']=room_bookings['Percentage'].round(2)
room_bookings
```

Out[53]:

room_class	booking_id	Percentage
Elite	49505	36.78
Premium	30566	22.71
Presidential	16073	11.94
Standard	38446	28.57

In [54]:
```
# Room Caegory wise Booking Percentage
room_bookings['Percentage'].plot(kind='pie',subplots=True,autopct='%0.2f')
```

Out[54]:
```
array([<Axes: ylabel='Percentage'>], dtype=object)
```



In [55]:
```
# Booking Trends weekly
weekly_bookings=revenue_tr.groupby(['week no','property_name']).agg({'booking_id':'
weekly_bookings.reset_index(inplace=True)
weekly_bookings_pivot=weekly_bookings.pivot(index='week no',columns='property_name'
weekly_bookings_pivot
```

Out[55]:

property_name	Atliq Bay	Atliq Blu	Atliq City	Atliq Exotica	Atliq Grands	Atliq Palace	Atliq Seasons
week no							
W 19	1745	1771	1897	1907	1411	1913	321
W 20	1746	1789	1906	1906	1384	1908	319
W 21	1441	1465	1586	1546	1134	1600	270
W 22	1726	1754	1896	1913	1389	1933	323
W 23	1421	1475	1559	1624	1149	1591	270
W 24	1756	1784	1916	1937	1394	1929	325
W 25	1734	1779	1897	1895	1381	1930	327
W 26	1424	1459	1568	1574	1141	1581	270
W 27	1758	1777	1883	1911	1379	1933	321
W 28	1753	1781	1867	1902	1380	1924	322
W 29	1750	1784	1922	1921	1408	1909	324
W 30	1427	1457	1569	1569	1145	1585	268
W 31	1451	1450	1566	1557	1130	1593	274
W 32	257	270	291	279	210	296	48

In [56]:
```
weekly_bookings_pivot.plot(kind='line',marker='o')
plt.ylabel('Total Bookings')
plt.title('Weekly Hotel Bookings')
```

Out[56]:
```
Text(0.5, 1.0, 'Weekly Hotel Bookings')
```

In [57]:
```
x=revenue_tr.groupby(['week no','booking_status']).agg({'booking_id':'nunique'})
x.reset_index(inplace=True)
x
```

Out[57]:

	week no	booking_status	booking_id
0	W 19	Cancelled	2779
1	W 19	Checked Out	7628
2	W 19	No Show	558
3	W 20	Cancelled	2711
4	W 20	Checked Out	7699
5	W 20	No Show	548
6	W 21	Cancelled	2226
7	W 21	Checked Out	6331
8	W 21	No Show	485
9	W 22	Cancelled	2772
10	W 22	Checked Out	7585
11	W 22	No Show	577
12	W 23	Cancelled	2222
13	W 23	Checked Out	6399
14	W 23	No Show	468
15	W 24	Cancelled	2797
16	W 24	Checked Out	7688
17	W 24	No Show	556
18	W 25	Cancelled	2748
19	W 25	Checked Out	7657
20	W 25	No Show	538
21	W 26	Cancelled	2299
22	W 26	Checked Out	6292
23	W 26	No Show	426
24	W 27	Cancelled	2691
25	W 27	Checked Out	7734
26	W 27	No Show	537
27	W 28	Cancelled	2643
28	W 28	Checked Out	7757
29	W 28	No Show	529

	week no	booking_status	booking_id
30	W 29	Cancelled	2733
31	W 29	Checked Out	7777
32	W 29	No Show	508
33	W 30	Cancelled	2210
34	W 30	Checked Out	6348
35	W 30	No Show	462
36	W 31	Cancelled	2193
37	W 31	Checked Out	6347
38	W 31	No Show	481
39	W 32	Cancelled	396
40	W 32	Checked Out	1169
41	W 32	No Show	86

In [58]:
```
y=revenue_tr.groupby(['week no']).agg({'booking_id':'nunique'}).rename(columns={'bo
y.reset_index(inplace=True)
y
```

Out[58]:

	week no	Total Bookings
0	W 19	10965
1	W 20	10958
2	W 21	9042
3	W 22	10934
4	W 23	9089
5	W 24	11041
6	W 25	10943
7	W 26	9017
8	W 27	10962
9	W 28	10929
10	W 29	11018
11	W 30	9020
12	W 31	9021
13	W 32	1651

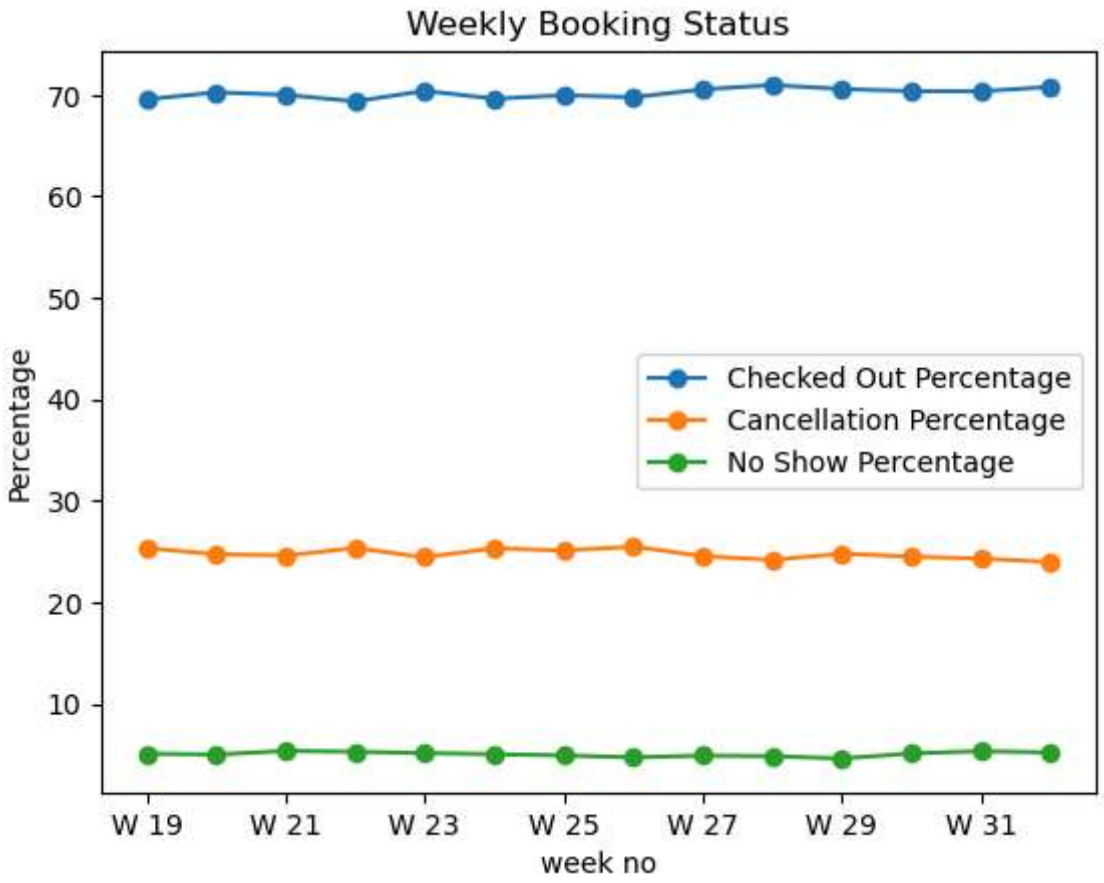
In [59]:
```
x1=x.pivot(index='week no',columns='booking_status',values='booking_id')
hotel_booking=pd.merge(x1,y,on='week no')
hotel_booking['Cancellation Percentage']=hotel_booking['Cancelled']*100/hotel_booki
hotel_booking['Checked Out Percentage']=hotel_booking['Checked Out']*100/hotel_book
hotel_booking['No Show Percentage']=hotel_booking['No Show']*100/hotel_booking['Tot
booking_status_trend=hotel_booking[['week no','Checked Out Percentage','Cancellatio
booking_status_trend
```

Out[59]:

	week no	Checked Out Percentage	Cancellation Percentage	No Show Percentage
0	W 19	69.566803	25.344277	5.088919
1	W 20	70.259171	24.739916	5.000913
2	W 21	70.017695	24.618447	5.363858
3	W 22	69.370770	25.352113	5.277117
4	W 23	70.403785	24.447134	5.149081
5	W 24	69.631374	25.332850	5.035776
6	W 25	69.971671	25.111944	4.916385
7	W 26	69.779306	25.496285	4.724409
8	W 27	70.552819	24.548440	4.898741
9	W 28	70.976302	24.183365	4.840333
10	W 29	70.584498	24.804865	4.610637
11	W 30	70.376940	24.501109	5.121951
12	W 31	70.358053	24.309943	5.332003
13	W 32	70.805572	23.985463	5.208964

In [60]:
```
x1=booking_status_trend.set_index('week no')
x1.plot(kind='line',marker='o')
plt.ylabel('Percentage')
plt.title('Weekly Booking Status')
```

Out[60]:
```
Text(0.5, 1.0, 'Weekly Booking Status')
```



In [99]:
```
#Average rating by City
fact_booking_hotel=pd.merge(fact_bookings,dim_hotels,on="property_id")
fact_booking_hotel.head()
```

Out[99]:

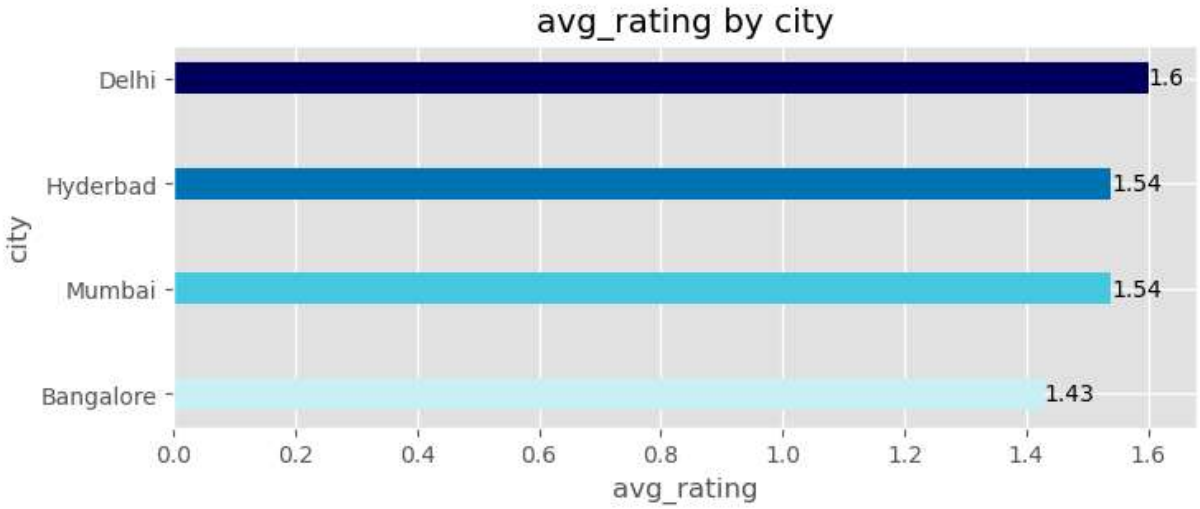
	booking_id	property_id	booking_date	check_in_date	checkout_date	no_guests
0	May012216558RT11	16558	2022-04-27	2022-05-01	2022-05-02	3
1	May012216558RT12	16558	2022-04-30	2022-05-01	2022-05-02	2
2	May012216558RT13	16558	2022-04-28	2022-05-01	2022-05-04	2
3	May012216558RT14	16558	2022-04-28	2022-05-01	2022-05-02	2
4	May012216558RT15	16558	2022-04-27	2022-05-01	2022-05-02	4

In [100…
```
avg_rating=fact_booking_hotel.groupby("city").ratings_given.mean().round(2).sort_va
avg_rating
```

Out[100…
```
city
Bangalore    1.43
Hyderabad    1.54
Mumbai       1.54
Delhi        1.60
Name: ratings_given, dtype: float64
```

In [101…

```
style.use("ggplot")
plt.figure(figsize=(8,3))
c=["#CAF0F8","#48CAE4","#0077B6","#03045E"]
plt.barh(xplot,avg_rating,height=0.3,color=c)
plt.yticks(xplot,["Bangalore","Mumbai","Hyderbad","Delhi"])
plt.xlabel("avg_rating")
plt.ylabel("city")
plt.title("avg_rating by city")
for index,value in enumerate(avg_rating):
 plt.text(value,index,str(value),va="center")
```



Occupancy Analysis

In [62]:

```
oc=pd.merge(dim_hotels,fact_aggregated_bookings,how='left',on='property_id')
occ_df=pd.merge(dim_date,oc,how='left',left_on='date',right_on='check_in_date')
occ_df.head()
```

Out[62]:

	date	mmm yy	week no	day_type	property_id	property_name	category	city	check_i
0	2022-05-01	May 22	W 19	weekend	16558	Atliq Grands	Luxury	Delhi	2022
1	2022-05-01	May 22	W 19	weekend	16558	Atliq Grands	Luxury	Delhi	2022
2	2022-05-01	May 22	W 19	weekend	16558	Atliq Grands	Luxury	Delhi	2022
3	2022-05-01	May 22	W 19	weekend	16558	Atliq Grands	Luxury	Delhi	2022
4	2022-05-01	May 22	W 19	weekend	16559	Atliq Exotica	Luxury	Mumbai	2022

In [63]:
```
# Hotelwise Occupancy
hotel_occ=occ_df.groupby(['city','property_name']).agg({'successful_bookings':'sum'
hotel_occ['Occupancy']=hotel_occ['successful_bookings']*100/hotel_occ['capacity']
hotel_occ.reset_index(inplace=True)
hotel_occ
```

Out[63]:

	city	property_name	successful_bookings	capacity	Occupancy
0	Bangalore	Atliq Bay	5812	8832	65.806159
1	Bangalore	Atliq Blu	5736	10764	53.288740
2	Bangalore	Atliq City	5979	9108	65.645586
3	Bangalore	Atliq Exotica	4705	8740	53.832952
4	Bangalore	Atliq Grands	4371	9844	44.402682
5	Bangalore	Atliq Palace	5413	10120	53.488142
6	Delhi	Atliq Bay	4820	9016	53.460515
7	Delhi	Atliq Blu	4418	6716	65.783204
8	Delhi	Atliq City	4693	8740	53.695652
9	Delhi	Atliq Grands	3153	4784	65.907191
10	Delhi	Atliq Palace	7147	10764	66.397250
11	Hyderabad	Atliq Bay	7333	11132	65.873158
12	Hyderabad	Atliq Blu	6458	9844	65.603413
13	Hyderabad	Atliq City	6638	10028	66.194655
14	Hyderabad	Atliq Exotica	5256	11776	44.633152
15	Hyderabad	Atliq Grands	4475	8372	53.451983
16	Hyderabad	Atliq Palace	4728	8924	52.980726
17	Mumbai	Atliq Bay	3424	7636	44.840230
18	Mumbai	Atliq Blu	5183	7820	66.278772
19	Mumbai	Atliq City	6013	11316	53.137151
20	Mumbai	Atliq Exotica	13480	20424	66.000783
21	Mumbai	Atliq Grands	5036	9384	53.665814
22	Mumbai	Atliq Palace	6337	9568	66.231187
23	Mumbai	Atliq Seasons	3982	8924	44.621246

In [64]:
```
hotel_occ_pivot=hotel_occ.pivot(index='city',columns='property_name',values='Occupa
hotel_occ_pivot.plot(kind='barh')
```

```
plt.xlabel('Occupancy Percentage')
plt.title('City & Hotel wise Occupancy')
```

Out[64]: Text(0.5, 1.0, 'City & Hotel wise Occupancy')

In [66]:
```
# Room Class wise Occupancy
room_occ=pd.merge(dim_room,occ_df,how='left',left_on='room_id',right_on='room_categ
room_occ_df=room_occ.groupby(['room_class']).agg({'successful_bookings':'sum','capa
room_occ_df['Occupancy']=room_occ_df['successful_bookings']*100/room_occ_df['capaci
room_occ_df['Occupancy']=room_occ_df['Occupancy'].round(2)
room_occ_df
```

Out[66]:

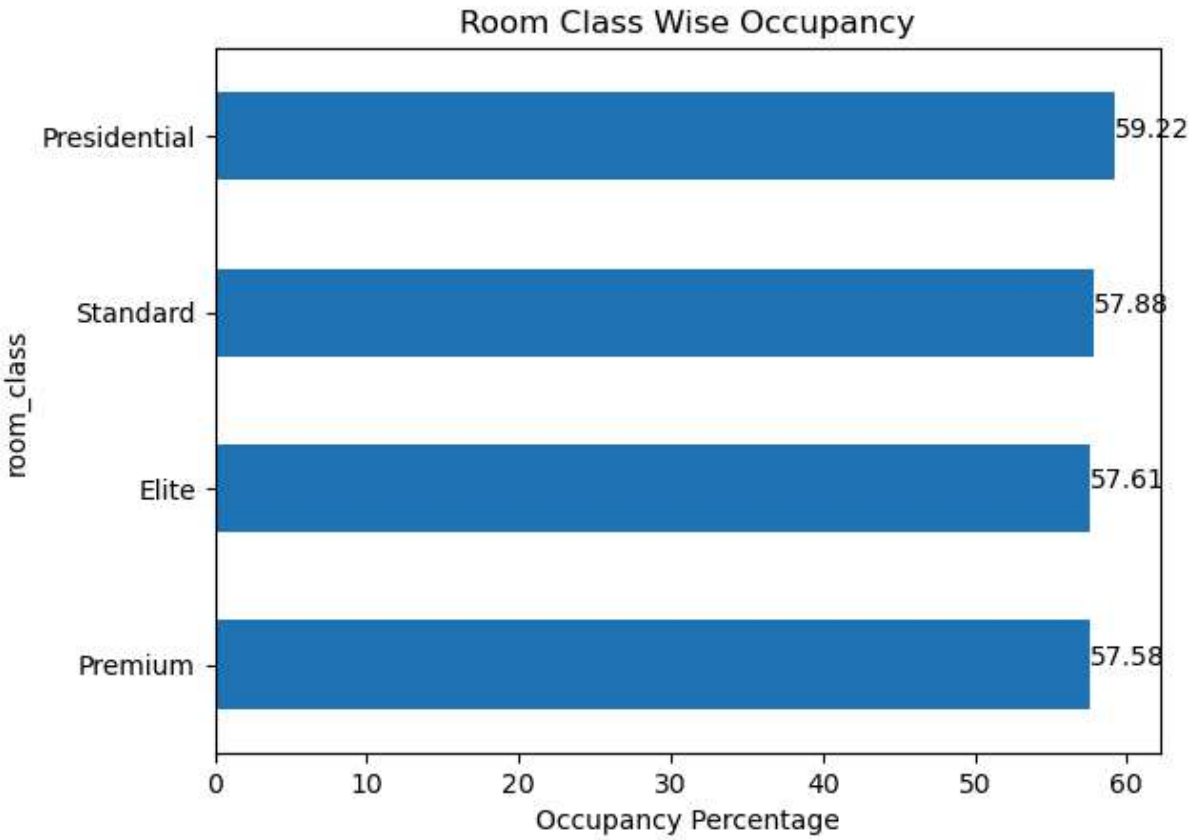
room_class	successful_bookings	capacity	Occupancy
Elite	49505	85928	57.61
Premium	30566	53084	57.58
Presidential	16073	27140	59.22
Standard	38446	66424	57.88

In [67]:
```
room_df=room_occ_df[['Occupancy']]
room_df=room_df.sort_values(by='Occupancy')
room_df
```

Out[67]:

room_class	Occupancy
Premium	57.58
Elite	57.61
Standard	57.88
Presidential	59.22

In [68]:
```
ax=room_df.plot(kind='barh',legend=False)
plt.xlabel('Occupancy Percentage')
plt.title('Room Class Wise Occupancy')
for index, value in enumerate(room_df['Occupancy']):
    ax.text(value, index, str(value))
plt.show()
```



In [69]:
```
# Daywise Occupancy
occ_df['Day Name']=occ_df['date'].dt.day_name()
occ_df['Day No']=occ_df['date'].dt.dayofweek
day_occ=occ_df.groupby(['Day Name','Day No','property_name']).agg({'successful_book
day_occ['Occupancy']=day_occ['successful_bookings']*100/day_occ['capacity']
day_occx=day_occ[['Occupancy']]
z=day_occx.reset_index()
z.sort_values(by='Day No')
```

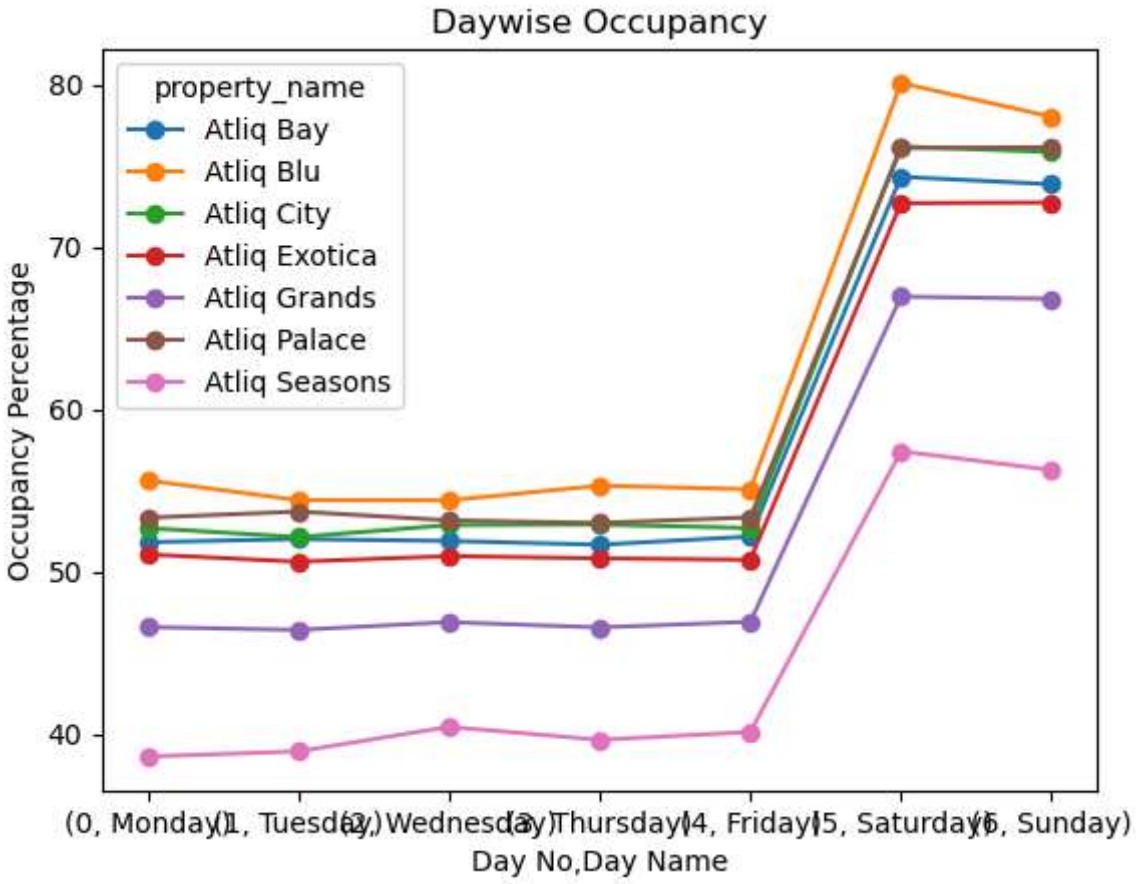
Out[69]:

	Day Name	Day No	property_name	Occupancy
13	Monday	0	Atliq Seasons	38.620143
7	Monday	0	Atliq Bay	51.816776
8	Monday	0	Atliq Blu	55.618204
9	Monday	0	Atliq City	52.690502
10	Monday	0	Atliq Exotica	51.063094
11	Monday	0	Atliq Grands	46.590909
12	Monday	0	Atliq Palace	53.324946
35	Tuesday	1	Atliq Bay	52.010050
36	Tuesday	1	Atliq Blu	54.409988
37	Tuesday	1	Atliq City	52.130733
38	Tuesday	1	Atliq Exotica	50.613656
41	Tuesday	1	Atliq Seasons	38.937351
39	Tuesday	1	Atliq Grands	46.416084
40	Tuesday	1	Atliq Palace	53.702372
47	Wednesday	2	Atliq Palace	53.163192
48	Wednesday	2	Atliq Seasons	40.444092
42	Wednesday	2	Atliq Bay	51.894086
43	Wednesday	2	Atliq Blu	54.389851
44	Wednesday	2	Atliq City	52.871073
45	Wednesday	2	Atliq Exotica	50.959378
46	Wednesday	2	Atliq Grands	46.896853
34	Thursday	3	Atliq Seasons	39.651071
33	Thursday	3	Atliq Palace	53.001438
32	Thursday	3	Atliq Grands	46.569056
31	Thursday	3	Atliq Exotica	50.821089
30	Thursday	3	Atliq City	52.907187
29	Thursday	3	Atliq Blu	55.296013
28	Thursday	3	Atliq Bay	51.662157
0	Friday	4	Atliq Bay	52.164670
1	Friday	4	Atliq Blu	55.074507

	Day Name	Day No	property_name	Occupancy
2	Friday	4	Atliq City	52.690502
3	Friday	4	Atliq Exotica	50.734659
4	Friday	4	Atliq Grands	46.918706
5	Friday	4	Atliq Palace	53.342919
6	Friday	4	Atliq Seasons	40.126883
16	Saturday	5	Atliq City	76.182737
20	Saturday	5	Atliq Seasons	57.414750
19	Saturday	5	Atliq Palace	76.078361
18	Saturday	5	Atliq Grands	66.936189
17	Saturday	5	Atliq Exotica	72.687986
15	Saturday	5	Atliq Blu	80.084575
14	Saturday	5	Atliq Bay	74.313877
26	Sunday	6	Atliq Palace	76.134846
21	Sunday	6	Atliq Bay	73.851400
22	Sunday	6	Atliq Blu	78.010471
23	Sunday	6	Atliq City	75.838364
25	Sunday	6	Atliq Grands	66.801948
27	Sunday	6	Atliq Seasons	56.259205
24	Sunday	6	Atliq Exotica	72.728732

In [70]:
```
z_pivot=z.pivot(index=['Day No','Day Name'],columns='property_name',values='Occupan
z_pivot.plot(kind='line',marker='o')
plt.ylabel('Occupancy Percentage')
plt.title('Daywise Occupancy')
```

Out[70]:
```
Text(0.5, 1.0, 'Daywise Occupancy')
```



In [71]:
```
weekly_occ=occ_df.groupby(['week no','property_name']).agg({'successful_bookings':'
weekly_occx=weekly_occ.reset_index()
weekly_occx['Occupancy']=weekly_occx['successful_bookings']*100/weekly_occx['capaci
weekly_occx
```

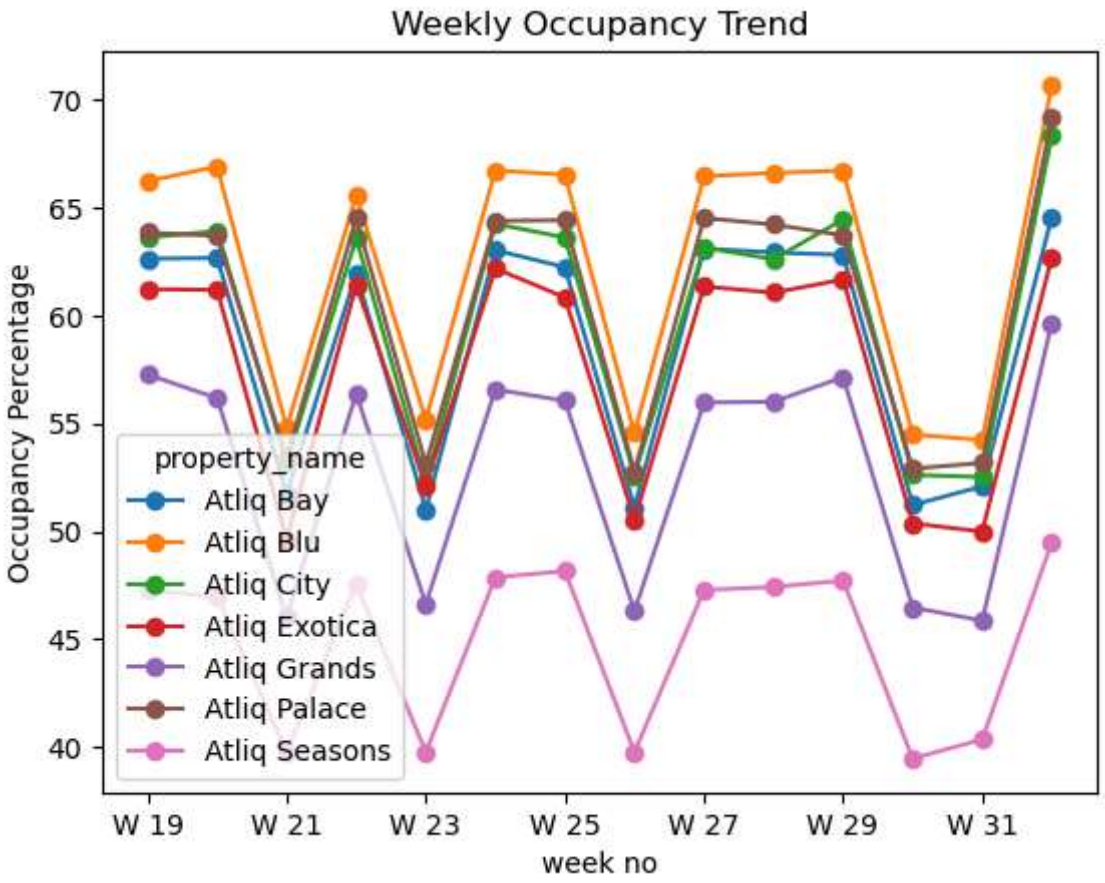
Out[71]:

	week no	property_name	successful_bookings	capacity	Occupancy
0	W 19	Atliq Bay	1745	2786	62.634602
1	W 19	Atliq Blu	1771	2674	66.230366
2	W 19	Atliq City	1897	2982	63.615023
3	W 19	Atliq Exotica	1907	3115	61.219904
4	W 19	Atliq Grands	1411	2464	57.264610
...	...	...	...	...	...
93	W 32	Atliq City	291	426	68.309859
94	W 32	Atliq Exotica	279	445	62.696629
95	W 32	Atliq Grands	210	352	59.659091
96	W 32	Atliq Palace	296	428	69.158879
97	W 32	Atliq Seasons	48	97	49.484536

98 rows × 5 columns

In [72]:
```
weekly_occx_pivot=weekly_occx.pivot(index='week no',columns='property_name',values=
weekly_occx_pivot.plot(kind='line',marker='o')
plt.ylabel('Occupancy Percentage')
plt.title('Weekly Occupancy Trend')
```

Out[72]:
```
Text(0.5, 1.0, 'Weekly Occupancy Trend')
```



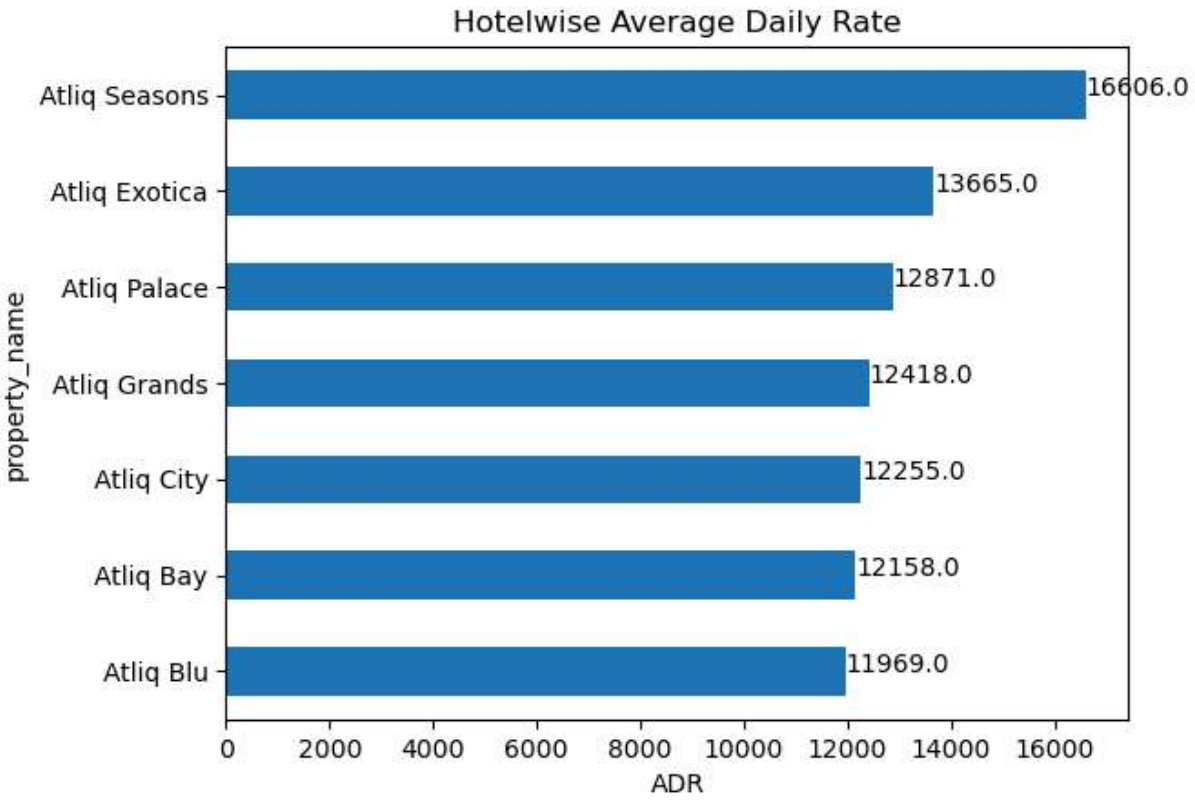
ADR (Average Daily Rate) Analysis

In [73]:
```
adr=revenue_tr.groupby(['property_name']).agg({'booking_id':'nunique','revenue_real
adr['ADR']=adr['revenue_realized']/adr['booking_id']
adr['ADR']=adr['ADR'].round(0)
adrx=adr[['ADR']]
adry=adrx.sort_values(by='ADR')
adrx
```

Out[73]:

property_name	ADR
Atliq Bay	12158.0
Atliq Blu	11969.0
Atliq City	12255.0
Atliq Exotica	13665.0
Atliq Grands	12418.0
Atliq Palace	12871.0
Atliq Seasons	16606.0

In [74]:
```
ax=adry.plot(kind='barh',legend=False)
plt.xlabel('ADR')
plt.title('Hotelwise Average Daily Rate')
for index, value in enumerate(adry['ADR']):
    ax.text(value, index, str(value))
plt.show()
```



Daily Booked Room Nights (DBRN), Daily Sellable Room Nights (DSRN), Daily Utilized Room Nights (DURN) Analysis

In [75]:
```
min_date=dim_date['date'].min()
max_date=dim_date['date'].max()
no_of_days=(max_date-min_date)/np.timedelta64(1,'D')
no_of_days
```

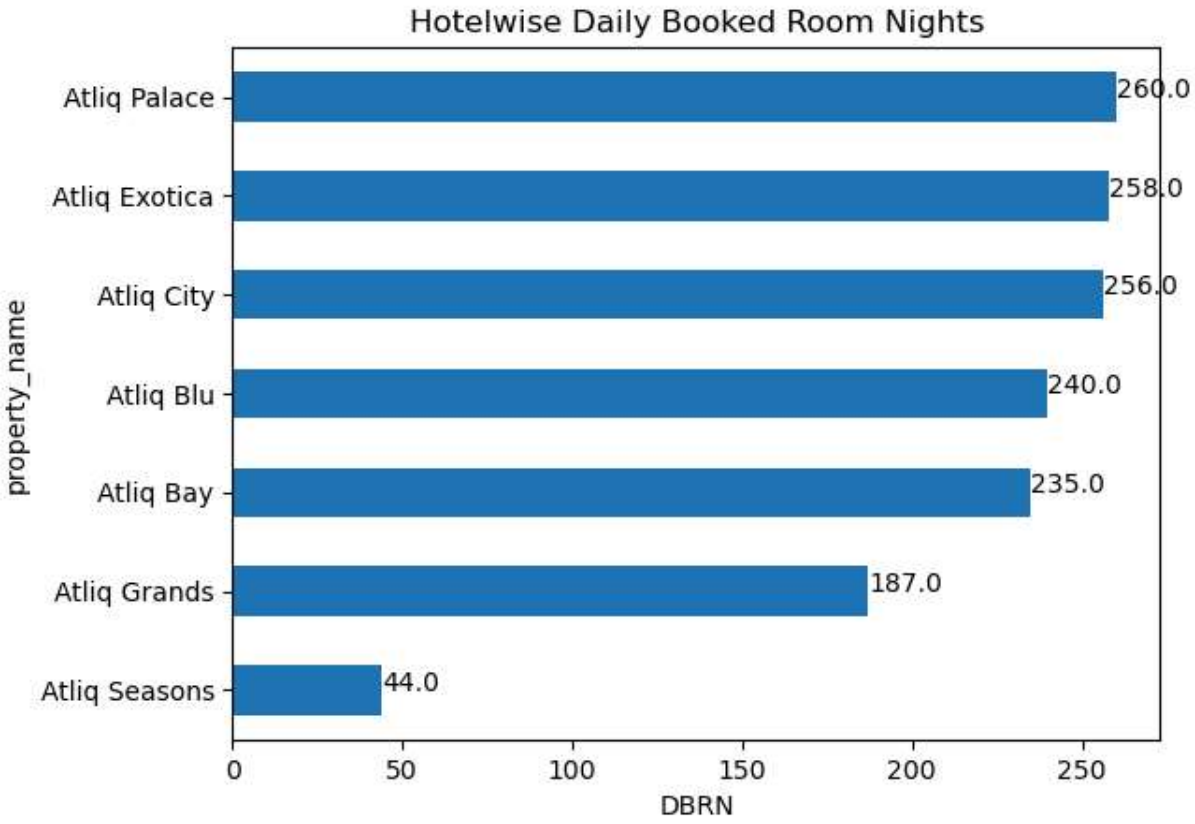
Out[75]: 91.0

In [82]:
```
#DBRN
DBRN=adr[['booking_id']]/no_of_days
DBRN['booking_id']=DBRN['booking_id'].round(0)
DBRNx=DBRN.sort_values(by='booking_id')
DBRNx
```

Out[82]:

property_name	booking_id
Atliq Seasons	44.0
Atliq Grands	187.0
Atliq Bay	235.0
Atliq Blu	240.0
Atliq City	256.0
Atliq Exotica	258.0
Atliq Palace	260.0

In [77]:
```
ax=DBRNx.plot(kind='barh',legend=False)
plt.xlabel('DBRN')
plt.title('Hotelwise Daily Booked Room Nights')
for index, value in enumerate(DBRNx['booking_id']):
    ax.text(value, index, str(value))
plt.show()
```



In [79]:
```
#DSRN
hotel_capacity=pd.merge(dim_hotels,fact_aggregated_bookings,how='left',on='property
hotel_capacity['DSRN']=hotel_capacity/no_of_days
hotel_capacity['DSRN']=hotel_capacity['DSRN'].round(0)
hotel_capacity
hotel_capacityx=hotel_capacity.sort_values(by='DSRN')
```

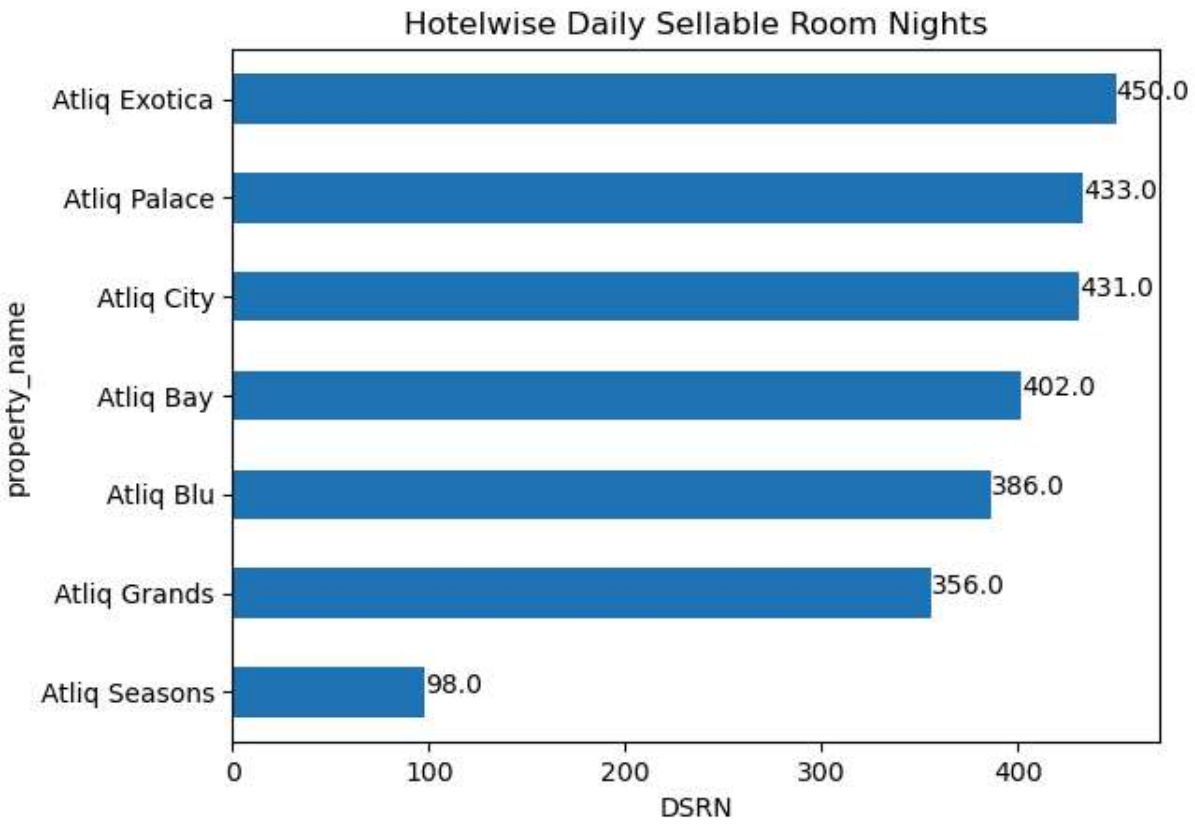
```
hotel_capacityy=hotel_capacityx[['DSRN']]
hotel_capacityy
```

Out[79]:

property_name	DSRN
Atliq Seasons	98.0
Atliq Grands	356.0
Atliq Blu	386.0
Atliq Bay	402.0
Atliq City	431.0
Atliq Palace	433.0
Atliq Exotica	450.0

In [81]:

```
ax=hotel_capacityy.plot(kind='barh',legend=False)
plt.xlabel('DSRN')
plt.title('Hotelwise Daily Sellable Room Nights')
for index, value in enumerate(hotel_capacityy['DSRN']):
    ax.text(value, index, str(value))
plt.show()
```



In [83]:

```
# DURN
m=revenue_tr[revenue_tr['booking_status']=='Checked Out']
n=m.groupby(['property_name']).agg({'booking_id':'nunique'})
```

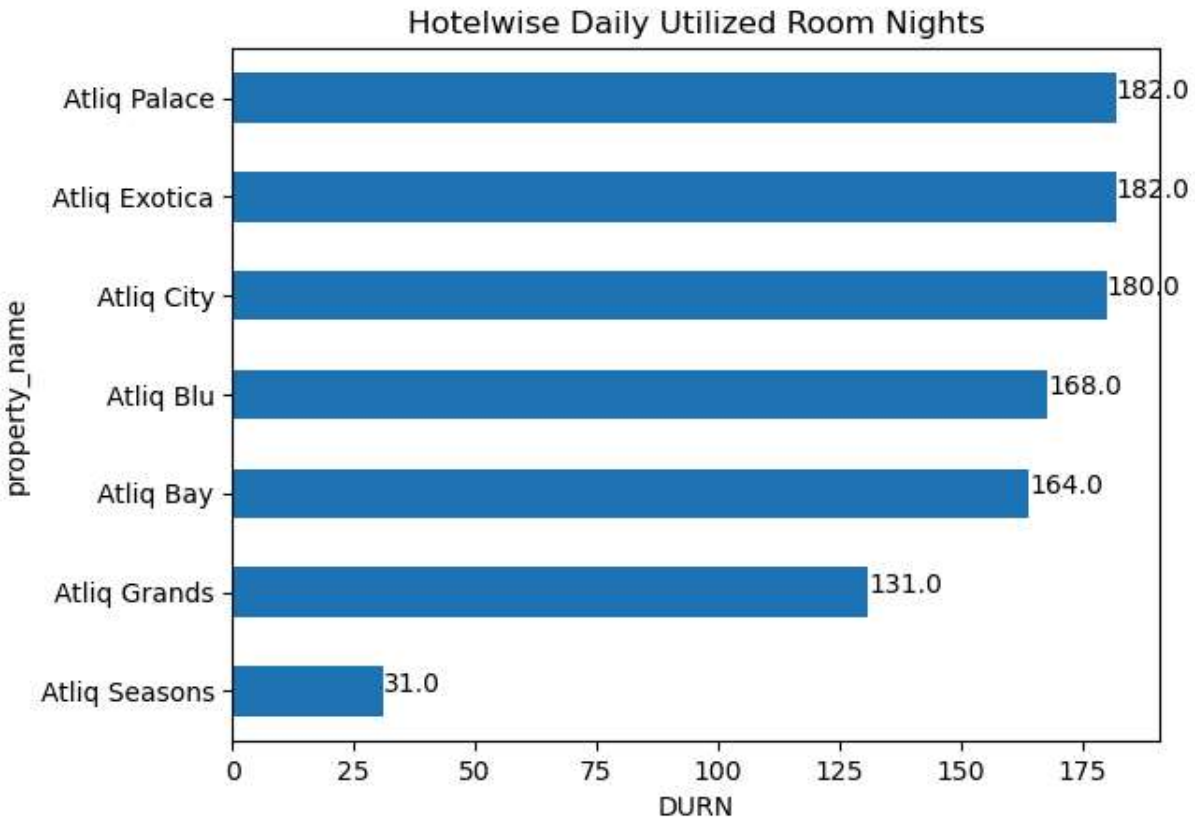
```
n['DURN']=n['booking_id']/no_of_days
n['DURN']=n['DURN'].round(0)
n1=n.sort_values(by='DURN')
n2=n1[['DURN']]
n2
```

Out[83]:

property_name	DURN
Atliq Seasons	31.0
Atliq Grands	131.0
Atliq Bay	164.0
Atliq Blu	168.0
Atliq City	180.0
Atliq Exotica	182.0
Atliq Palace	182.0

In [85]:

```
ax=n2.plot(kind='barh',legend=False)
plt.xlabel('DURN')
plt.title('Hotelwise Daily Utilized Room Nights')
for index, value in enumerate(n2['DURN']):
    ax.text(value, index, str(value))
plt.show()
```



In [ ]: