Import Libraries

Load Dataset

	Booking_ID	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	type_of_meal_plan	required
0	INN00001	2	0	1	2	Meal Plan 1	
1	INN00002	2	0	2	3	Not Selected	
2	INN00003	1	0	2	1	Meal Plan 1	
3	INN00004	2	0	0	2	Meal Plan 1	
4	INN00005	2	0	1	1	Not Selected	

	Booking_ID	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	type_of_meal_plan	req
36270	INN36271	3	0	2	6	Meal Plan 1	
36271	INN36272	2	0	1	3	Meal Plan 1	
36272	INN36273	2	0	2	6	Meal Plan 1	
36273	INN36274	2	0	0	3	Not Selected	
36274	INN36275	2	0	1	2	Meal Plan 1	

```
(36275, 19)
```

```
Index(['Booking_ID', 'no_of_adults', 'no_of_children', 'no_of_weekend_nights',
       'no_of_week_nights', 'type_of_meal_plan', 'required_car_parking_space',
       'room_type_reserved', 'lead_time', 'arrival_year', 'arrival_month',
       'arrival_date', 'market_segment_type', 'repeated_guest',
       'no_of_previous_cancellations', 'no_of_previous_bookings_not_canceled',
       'avg_price_per_room', 'no_of_special_requests', 'booking_status'],
      dtype='object')
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 36275 entries, 0 to 36274
Data columns (total 19 columns):
 #   Column                                Non-Null Count  Dtype  
---  ------                                --------------  -----  
 0   Booking_ID                            36275 non-null  object 
 1   no_of_adults                          36275 non-null  int64  
 2   no_of_children                        36275 non-null  int64  
 3   no_of_weekend_nights                  36275 non-null  int64  
 4   no_of_week_nights                     36275 non-null  int64  
 5   type_of_meal_plan                     36275 non-null  object 
 6   required_car_parking_space            36275 non-null  int64  
 7   room_type_reserved                    36275 non-null  object 
 8   lead_time                             36275 non-null  int64  
 9   arrival_year                          36275 non-null  int64  
 10  arrival_month                         36275 non-null  int64  
 11  arrival_date                          36275 non-null  int64  
 12  market_segment_type                   36275 non-null  object 
 13  repeated_guest                        36275 non-null  int64  
 14  no_of_previous_cancellations          36275 non-null  int64  
 15  no_of_previous_bookings_not_canceled  36275 non-null  int64  
 16  avg_price_per_room                    36275 non-null  float64
 17  no_of_special_requests                36275 non-null  int64  
 18  booking_status                        36275 non-null  object 
dtypes: float64(1), int64(13), object(5)
memory usage: 5.3+ MB
```

	Booking_ID	type_of_meal_plan	room_type_reserved	market_segment_type	booking_status
count	36275	36275	36275	36275	36275
unique	36275	4	7	5	2
top	INN00001	Meal Plan 1	Room_Type 1	Online	Not_Canceled
freq	1	27835	28130	23214	24390

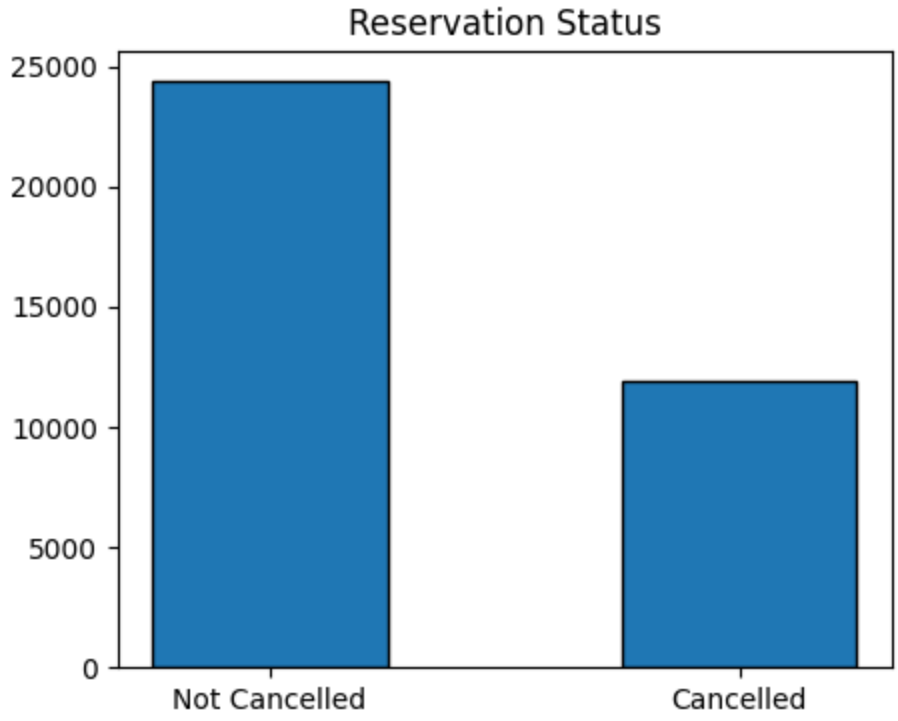
```
Booking_ID
['INN00001' 'INN00002' 'INN00003' ... 'INN36273' 'INN36274' 'INN36275']
__________________________________________________________
type_of_meal_plan
['Meal Plan 1' 'Not Selected' 'Meal Plan 2' 'Meal Plan 3']
__________________________________________________________
room_type_reserved
['Room_Type 1' 'Room_Type 4' 'Room_Type 2' 'Room_Type 6' 'Room_Type 5'
 'Room_Type 7' 'Room_Type 3']
__________________________________________________________
market_segment_type
['Offline' 'Online' 'Corporate' 'Aviation' 'Complementary']
__________________________________________________________
booking_status
['Not_Canceled' 'Canceled']
__________________________________________________________
```

```
Booking_ID                            0
no_of_adults                          0
no_of_children                        0
no_of_weekend_nights                  0
no_of_week_nights                     0
type_of_meal_plan                     0
required_car_parking_space            0
room_type_reserved                    0
lead_time                             0
arrival_year                          0
arrival_month                         0
arrival_date                          0
market_segment_type                   0
repeated_guest                        0
no_of_previous_cancellations          0
no_of_previous_bookings_not_canceled  0
avg_price_per_room                    0
no_of_special_requests                0
booking_status                        0
dtype: int64
```

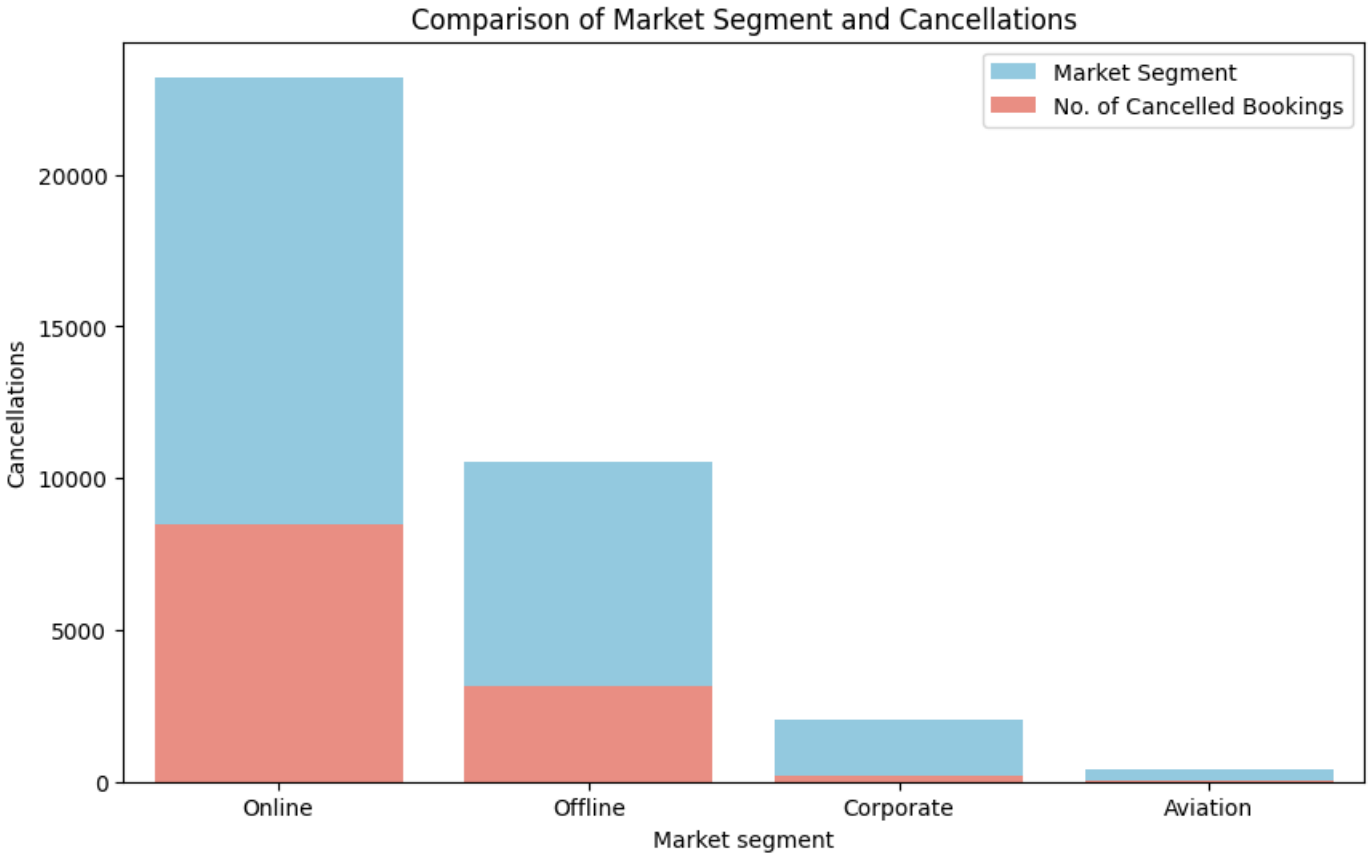
	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	required_car_parking_space	lead
count	36275.000000	36275.000000	36275.000000	36275.000000	36275.000000	36275.0
mean	1.844962	0.105279	0.810724	2.204300	0.030986	85.2
std	0.518715	0.402648	0.870644	1.410905	0.173281	85.9
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.0
25%	2.000000	0.000000	0.000000	1.000000	0.000000	17.0
50%	2.000000	0.000000	1.000000	2.000000	0.000000	57.0
75%	2.000000	0.000000	2.000000	3.000000	0.000000	126.0
max	4.000000	10.000000	7.000000	17.000000	1.000000	443.0

```
Not_Canceled    0.672364
Canceled        0.327636
Name: booking_status, dtype: float64
```

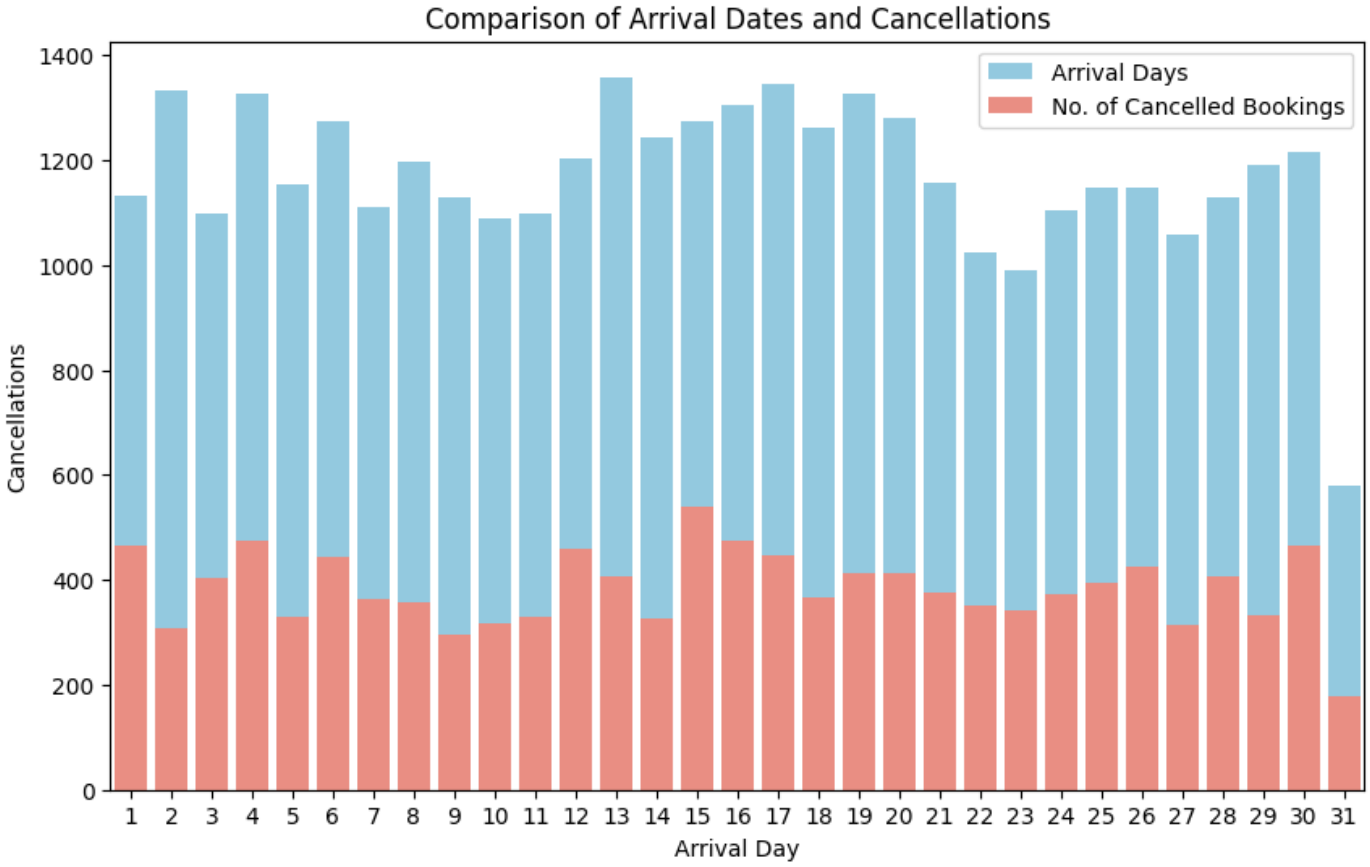
Analysis of Cancelled and Not_Cancelled Bookings



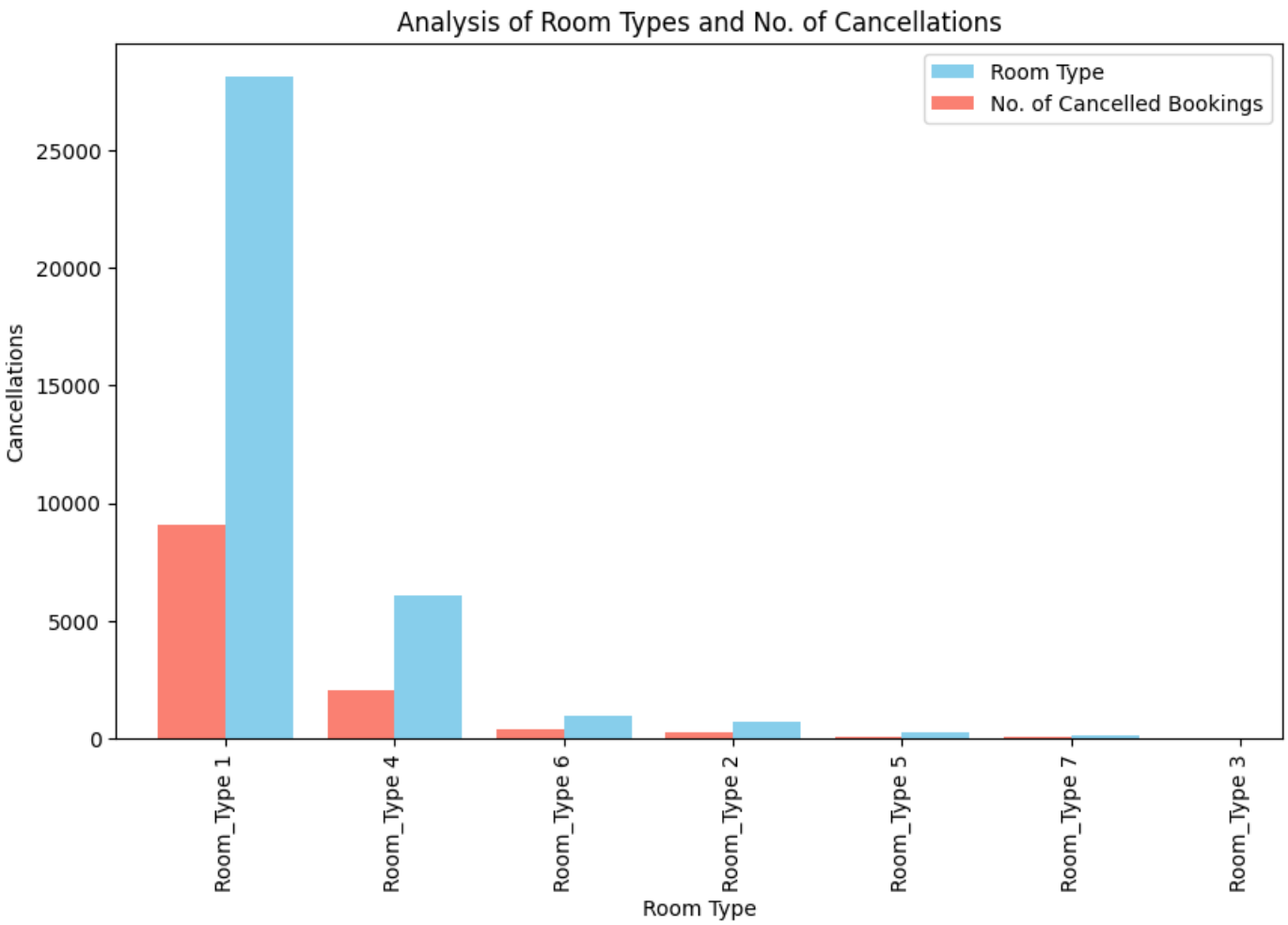
Analysis of Market Segment Type and Cancelled Bookings



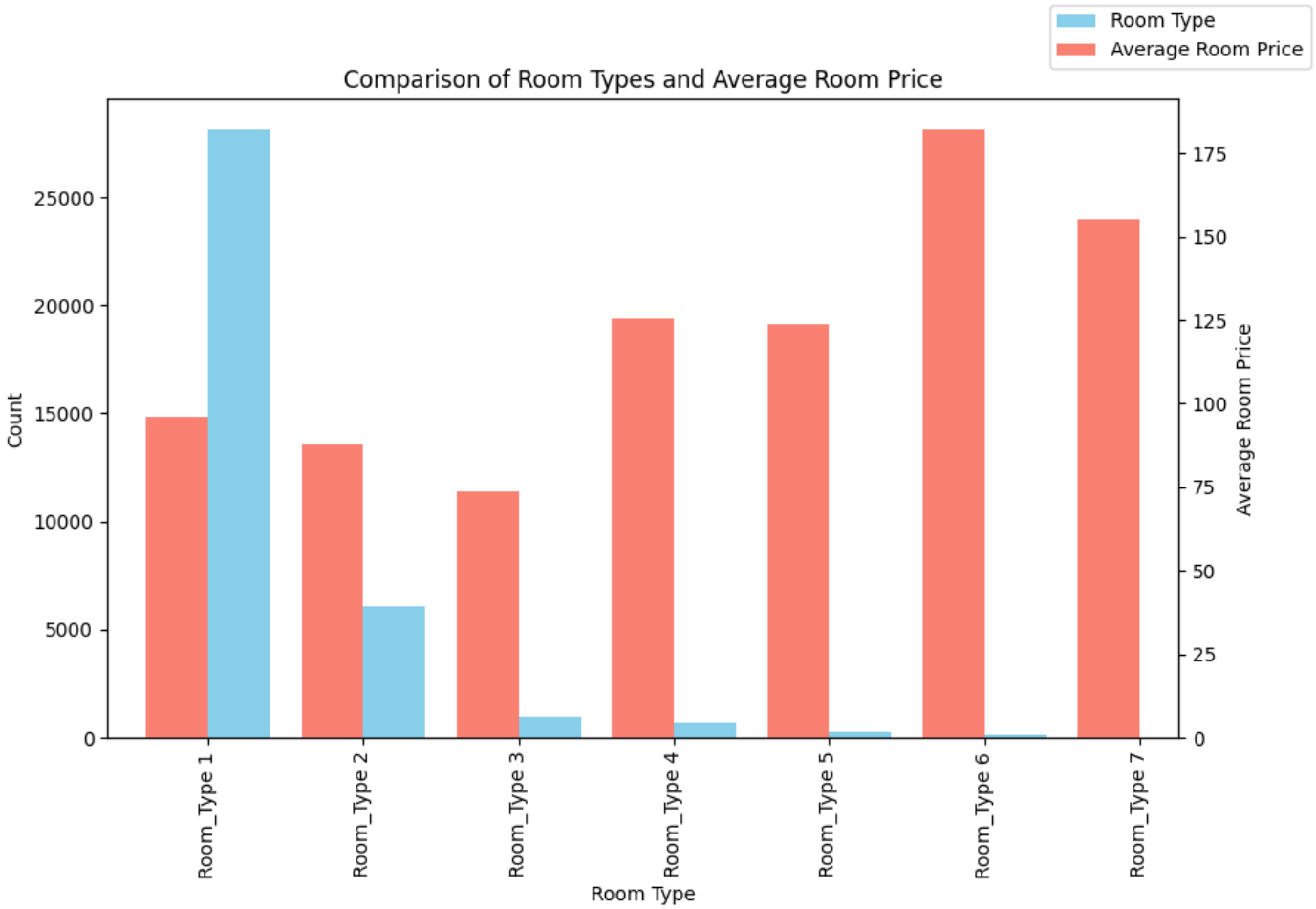
Analysis of Arrival Date and Booking Cancellations



Analysis of Room Type and No. of Cancellations

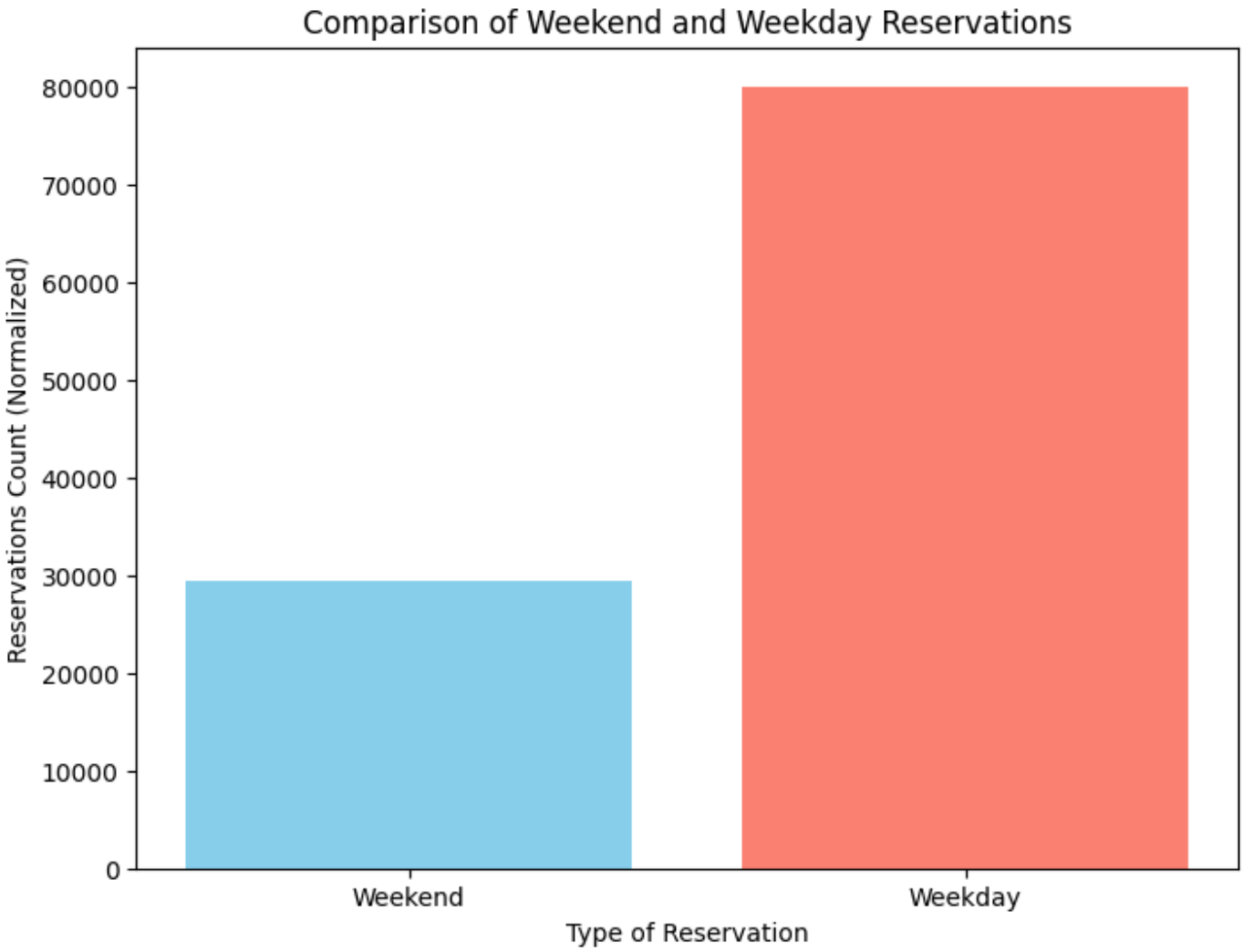


Analysis of Room Type and Average Room Price

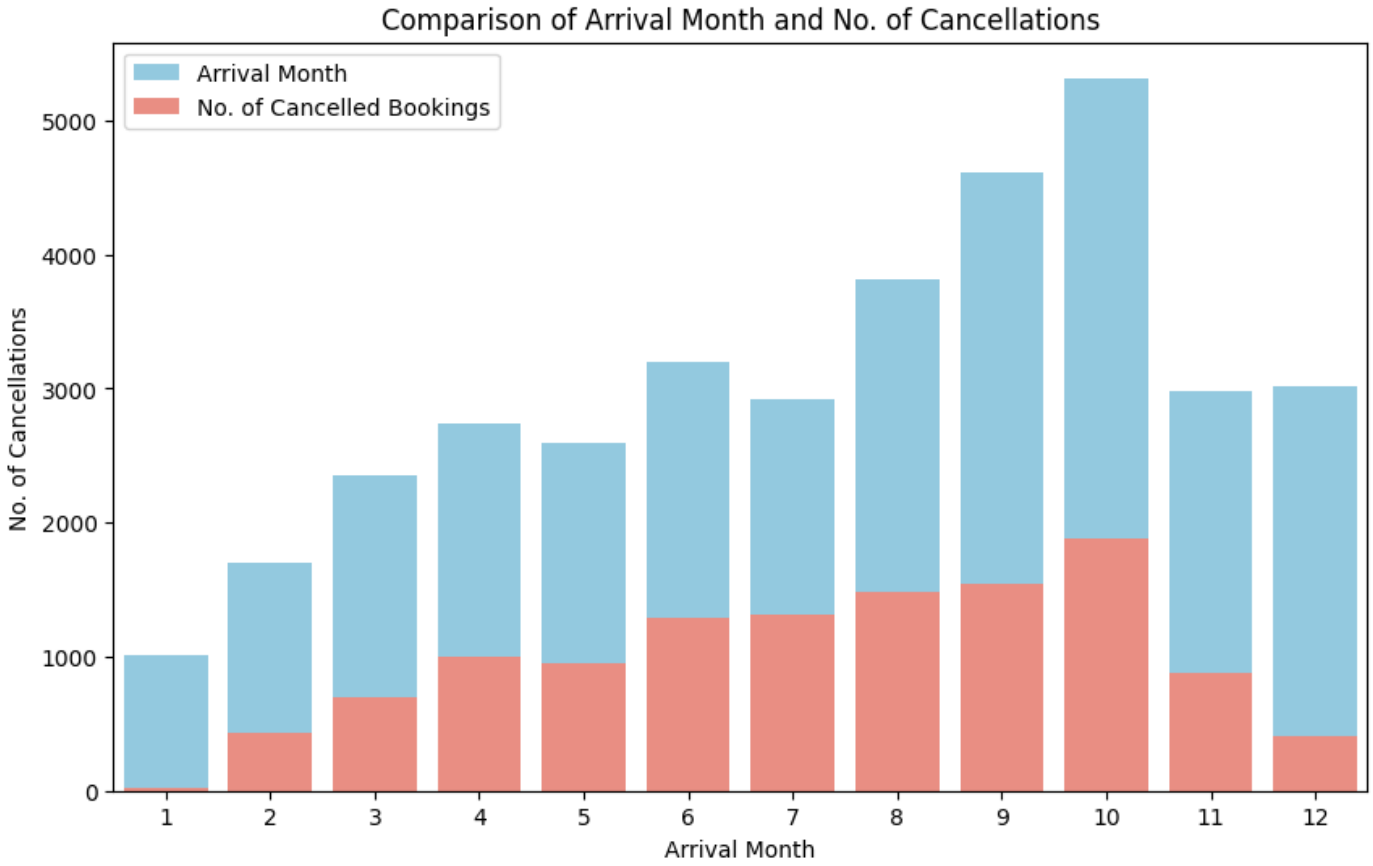


Analyzing No. of Weekend Night Reservations to No. of Weekday Night Reservations

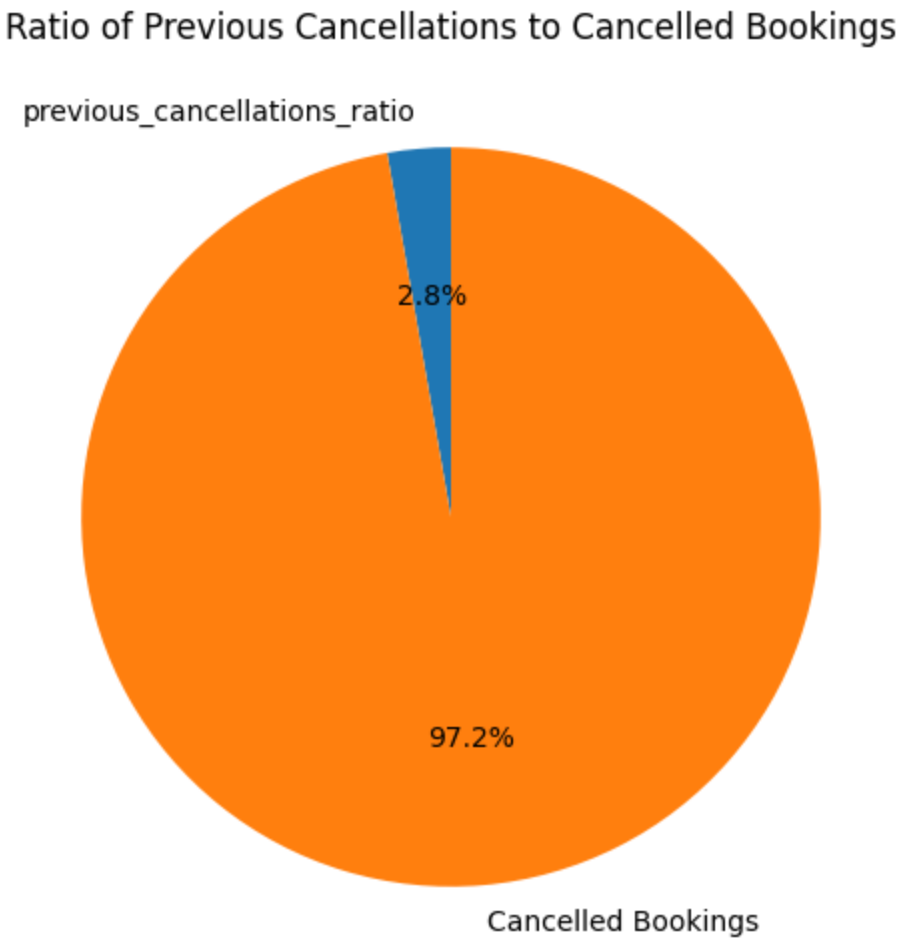
```
Weekend Reservations: 29409
Weekday Reservations: 79961
```



Analysis of Arrival Months and Cancellations



Analysis of Previous Cancellations Ratio to Overall Cancellations



Regression Analysis of 'Cancelled Bookings' and 'Average Price per Room'

```
T-Statistic: 27.433239020374042
P-Value: 5.2303189247282316e-164
```

T-tests and P-test analysis carried out suggest a very large T-statistic and an extremely small p-value, indicating that the average room price has a statistically significant impact on the cancellation rate. With such a small p-value (close to 0), you can confidently reject the null hypothesis and conclude that there is indeed a significant relationship between the average room price and the cancellation rate. We investigate further..

```
                            OLS Regression Results
==============================================================================
Dep. Variable:         booking_status   R-squared:                       0.020
Model:                            OLS   Adj. R-squared:                  0.020
Method:                 Least Squares   F-statistic:                     752.6
Date:                Mon, 23 Oct 2023   Prob (F-statistic):          5.23e-164
Time:                        11:11:23   Log-Likelihood:                -23661.
No. Observations:               36275   AIC:                          4.733e+04
Df Residuals:                   36273   BIC:                          4.734e+04
Df Model:                           1
Covariance Type:            nonrobust
==========================================================================================
                             coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------------------
const                     0.1304      0.008     17.177      0.000       0.116       0.145
avg_price_per_room        0.0019   6.95e-05     27.433      0.000       0.002       0.002
==============================================================================
Omnibus:                   312722.067   Durbin-Watson:                   2.010
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             6013.016
Skew:                           0.704   Prob(JB):                         0.00
Kurtosis:                       1.588   Cond. No.                         340.
==============================================================================

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
```

0.020 is observed to be the R-Squared value, indicating a 2% variance in cancellation rate. This means there is no real relationship between Average Room Prices and Cancellation rate. Other Factors should be considered.

Summary

Observations

- High Price does NOT Affect Cancellation rates
- Most Bookings and Cancellations occur from ONLINE bookings type
- October always has most Bookings and has most Cancellations also
- Room 1 has the Highest number of Booking Cancellations

Suggestions

- Prices for Online Bookings should be reduced
- Set Policy to ensure Customers that make bookings are not encouraged to Cancel
- Promotions can be carried out to improve online bookings
- Promotions and discounts can be given for rooms (except room 1)