

```
In [1]: import warnings
warnings.filterwarnings("ignore")
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
from sklearn.model_selection import train_test_split
import statsmodels.stats.api as sms
from statsmodels.stats.outliers_influence import variance_inflation_factor
import statsmodels.api as sm
from statsmodels.tools.tools import add_constant
from sklearn import metrics
from sklearn.metrics import f1_score, accuracy_score, recall_score, precision_score, c
from sklearn import metrics
from sklearn.model_selection import GridSearchCV
from sklearn.tree import DecisionTreeClassifier
from sklearn import tree
import warnings
from statsmodels.tools.sm_exceptions import ConvergenceWarning
warnings.simplefilter('ignore', ConvergenceWarning)
```

Here im importing all needed packages

```
In [2]: df = pd.read_csv('/Users/conne/Downloads/INNHotelsGroup.csv')
```

Here im loading my data

The next few lines im getting an overview of the code

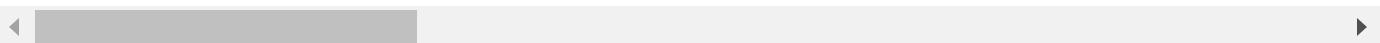
```
In [3]: df.head()
```

	Booking_ID	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	type_of_meal_p
0	INN00001	2	0	1	2	Meal Pla
1	INN00002	2	0	2	3	Not Sele
2	INN00003	1	0	2	1	Meal Pla
3	INN00004	2	0	0	2	Meal Pla
4	INN00005	2	0	1	1	Not Sele

```
In [4]: df.tail()
```

Out[4]:

	Booking_ID	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	type_of_m
36270	INN36271	3	0	2	6	Me
36271	INN36272	2	0	1	3	Me
36272	INN36273	2	0	2	6	Me
36273	INN36274	2	0	0	3	Not
36274	INN36275	2	0	1	2	Me



In [5]:

```
df.info()
df.shape
```

```
<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
(36275, 19)
```

Out[5]:

```
In [6]: print(df.value_counts(subset='type_of_meal_plan'))
print(df.value_counts(subset='room_type_reserved'))
print(df.value_counts(subset='market_segment_type'))
print(df.value_counts(subset='booking_status'))
print(df.value_counts(subset='repeated_guest'))
```

```
type_of_meal_plan
Meal Plan 1      27835
Not Selected     5130
Meal Plan 2      3305
Meal Plan 3          5
dtype: int64
room_type_reserved
Room_Type 1      28130
Room_Type 4      6057
Room_Type 6      966
Room_Type 2      692
Room_Type 5      265
Room_Type 7      158
Room_Type 3          7
dtype: int64
market_segment_type
Online           23214
Offline          10528
Corporate        2017
Complementary    391
Aviation          125
dtype: int64
booking_status
Not_Canceled    24390
Canceled         11885
dtype: int64
repeated_guest
0            35345
1            930
dtype: int64
```

```
In [7]: print(df[ "type_of_meal_plan" ].value_counts(1))
print(df[ "room_type_reserved" ].value_counts(1))
print(df[ "market_segment_type" ].value_counts(1))
print(df[ "booking_status" ].value_counts(1))
```

```
Meal Plan 1      0.767333
Not Selected     0.141420
Meal Plan 2      0.091110
Meal Plan 3          0.000138
Name: type_of_meal_plan, dtype: float64
Room_Type 1      0.775465
Room_Type 4      0.166975
Room_Type 6      0.026630
Room_Type 2      0.019076
Room_Type 5      0.007305
Room_Type 7      0.004356
Room_Type 3          0.000193
Name: room_type_reserved, dtype: float64
Online           0.639945
Offline          0.290227
Corporate        0.055603
Complementary    0.010779
Aviation          0.003446
Name: market_segment_type, dtype: float64
Not_Canceled    0.672364
Canceled         0.327636
Name: booking_status, dtype: float64
```

77.55% of people get room type 1 only 14.14% of people dont get a meal plan

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

	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	required_car_parking_sr
count	36275.000000	36275.000000	36275.000000	36275.000000	36275.000000
mean	1.844962	0.105279	0.810724	2.204300	0.030000
std	0.518715	0.402648	0.870644	1.410905	0.173000
min	0.000000	0.000000	0.000000	0.000000	0.000000
25%	2.000000	0.000000	0.000000	1.000000	0.000000
50%	2.000000	0.000000	1.000000	2.000000	0.000000
75%	2.000000	0.000000	2.000000	3.000000	0.000000
max	4.000000	10.000000	7.000000	17.000000	1.000000

◀ ▶

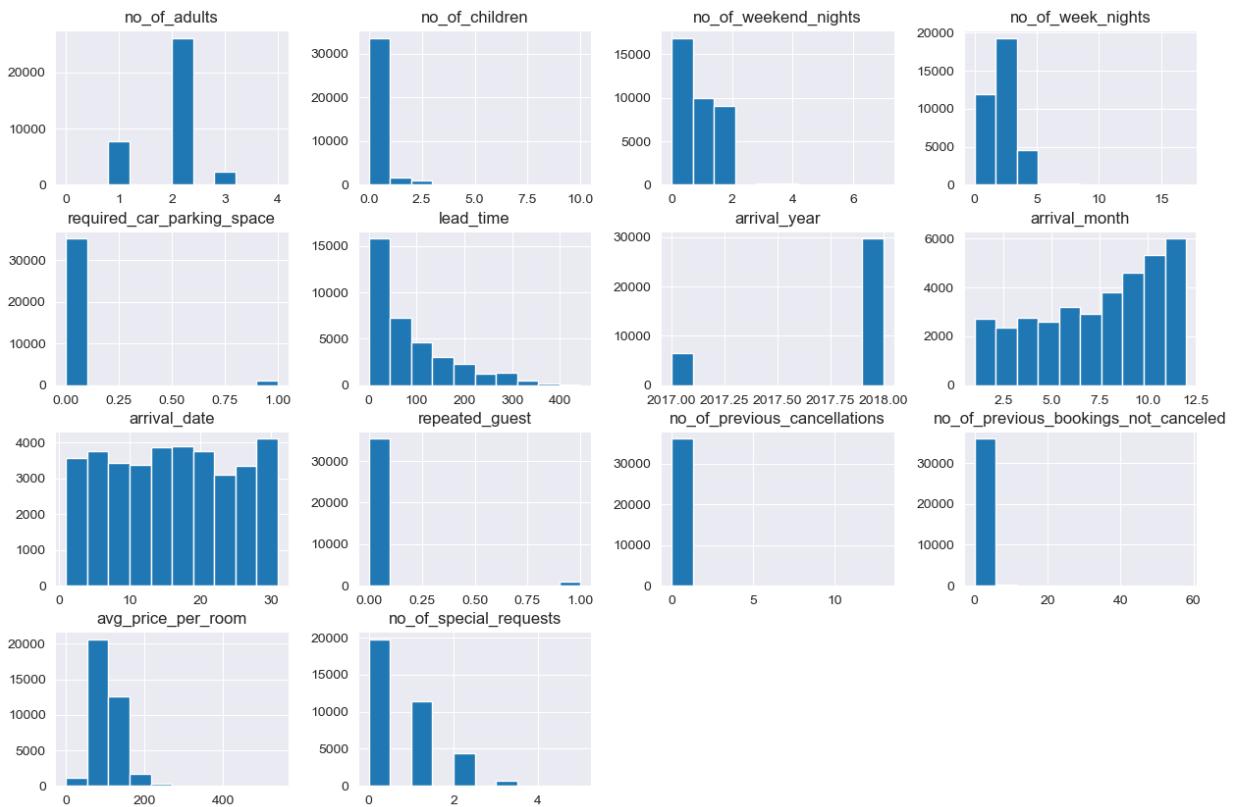
Most people dont require a place to park so this probably has little impact on cancellations

In [9]: `df.isna().values.any()`

Out[9]: `False`

i have no missing values

In [10]: `sns.set_style("darkgrid")
df.hist(figsize=(15, 10))
plt.show()`



Most people are getting rooms for 2 adults Most people came in 2018 Bussiness increases in the last 4 months of the year Almost everyone who booked a room didn't need a parking spot

```
In [11]: cols = ["required_car_parking_space", "repeated_guest"]
df[cols] = df[cols].replace(0, "no")
df[cols] = df[cols].replace(1, "yes")
```

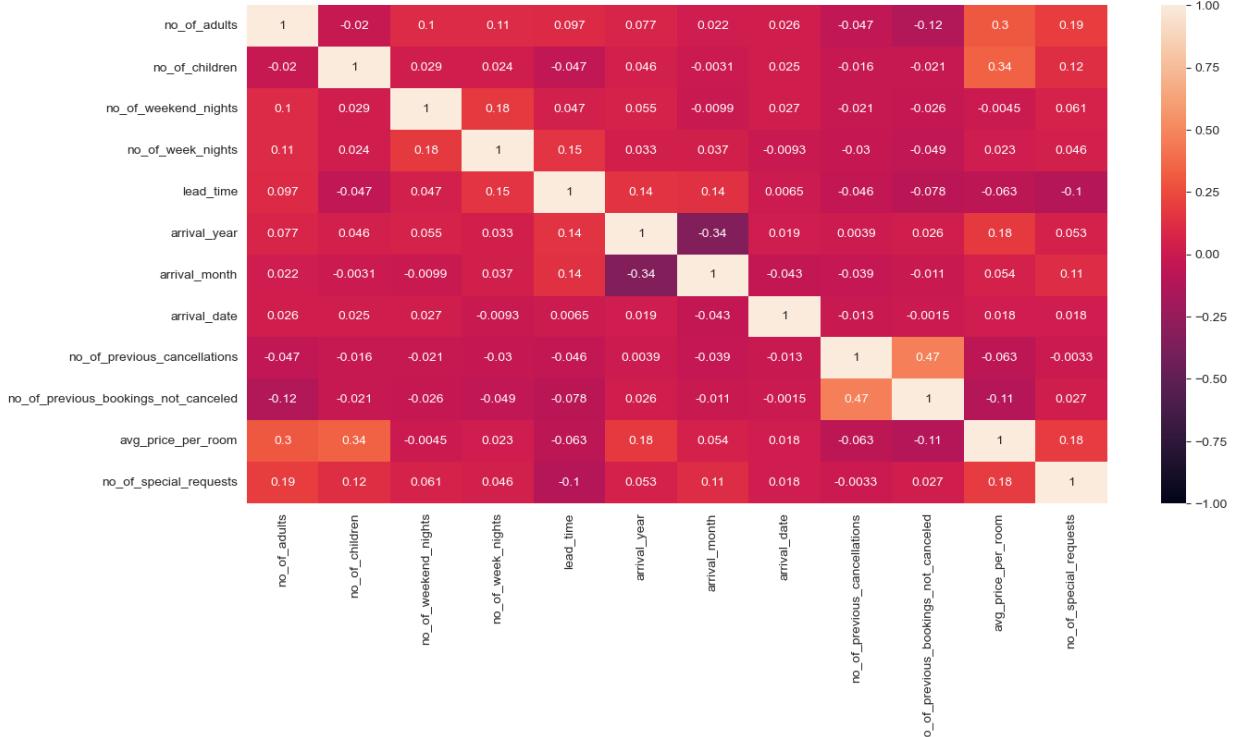
```
In [12]: cols = ["required_car_parking_space", "repeated_guest"]
df[cols] = df[cols].replace(0, "no")
df[cols] = df[cols].replace(1, "yes")
df['required_car_parking_space'].value_counts()
```

```
Out[12]: no      35151
yes     1124
Name: required_car_parking_space, dtype: int64
```

```
In [13]: df['repeated_guest'].value_counts()
```

```
Out[13]: no      35345
yes      930
Name: repeated_guest, dtype: int64
```

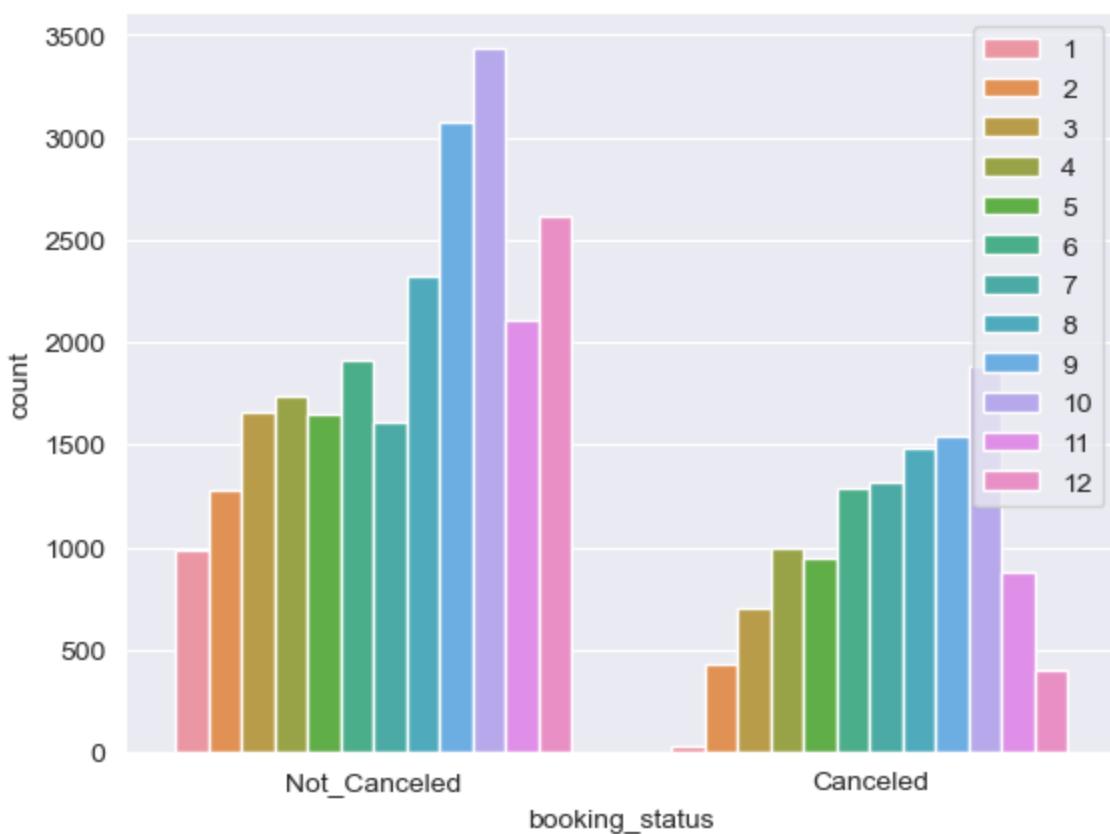
```
In [14]: plt.figure(figsize=(15, 7))
sns.heatmap(df.corr(), annot=True, vmin=-1, vmax=1)
plt.show()
```



There are not highly correlated values

```
In [15]: sns.countplot(data=df, x='booking_status', hue='arrival_month')
plt.legend(loc='upper right')
```

```
Out[15]: <matplotlib.legend.Legend at 0x24138a25ad0>
```



Question 1: the ratio of not canceled to canceled increase as you get later in to the year

In [16]: `df.columns`

Out[16]: `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')`

In [17]: `sns.pairplot(df[['no_of_adults', 'no_of_children', 'no_of_weekend_nights', 'no_of_week_nights', 'arrival_month', 'no_of_previous_cancellations', 'no_of_previous_bookings_not_canceled', 'avg_price_per_room', 'no_of_special_requests']])
plt.show()`



The pair plot shows us are data columns aren't very correlated with eachother

In [18]: `df["arrival_month"].value_counts(1)`

```
Out[18]: 10    0.146575
         9    0.127112
         8    0.105114
         6    0.088298
         12   0.083280
        11   0.082150
         7    0.080496
         4    0.075424
         5    0.071620
         3    0.065003
         2    0.046975
         1    0.027953
Name: arrival_month, dtype: float64
```

October is the busiest month of the year as 14% of people visit in that month.

```
In [19]: df["market_segment_type"].value_counts()
```

```
Out[19]: Online      23214
          Offline     10528
          Corporate    2017
          Complementary 391
          Aviation      125
Name: market_segment_type, dtype: int64
```

```
In [20]: df["market_segment_type"].value_counts(1)
```

```
Out[20]: Online      0.639945
          Offline     0.290227
          Corporate    0.055603
          Complementary 0.010779
          Aviation      0.003446
Name: market_segment_type, dtype: float64
```

Question 2: most guests come from online

```
In [21]: df_online = df.copy()
df_online['market_segment_type'] = df_online['market_segment_type'].replace(['Offline'])
df_online = df_online.dropna()
df_offline = df.copy()
df_offline['market_segment_type'] = df_offline['market_segment_type'].replace(['Online'])
df_offline = df_offline.dropna()
df_corporate = df.copy()
df_corporate['market_segment_type'] = df_corporate['market_segment_type'].replace(['Corporate'])
df_corporate = df_corporate.dropna()
df_complementary = df.copy()
df_complementary['market_segment_type'] = df_complementary['market_segment_type'].replace(['Complementary'])
df_complementary = df_complementary.dropna()
df_aviation = df.copy()
df_aviation['market_segment_type'] = df_aviation['market_segment_type'].replace(['Aviation'])
df_aviation = df_aviation.dropna()
```

```
In [22]: df_online.isna().sum()
```

```
Out[22]: 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
```

```
In [23]: df_online.shape
```

```
Out[23]: (23214, 19)
```

```
In [24]: df_online.mean()
```

```
Out[24]: no_of_adults          1.939476  
no_of_children           0.151977  
no_of_weekend_nights     0.886577  
no_of_week_nights         2.289911  
lead_time                 75.334238  
arrival_year               2017.872878  
arrival_month              7.380417  
arrival_date                15.690618  
no_of_previous_cancellations 0.013225  
no_of_previous_bookings_not_canceled 0.012105  
avg_price_per_room        112.256855  
no_of_special_requests    0.842250  
dtype: float64
```

```
In [25]: df_offline.shape
```

```
Out[25]: (10528, 19)
```

```
In [26]: df_offline.mean()
```

```
Out[26]: no_of_adults           1.777641
          no_of_children        0.021087
          no_of_weekend_nights   0.730528
          no_of_week_nights      2.180661
          lead_time              122.872625
          arrival_year            2017.722074
          arrival_month            7.572758
          arrival_date             15.396087
          no_of_previous_cancellations 0.011113
          no_of_previous_bookings_not_canceled 0.010828
          avg_price_per_room       91.632679
          no_of_special_requests   0.202603
          dtype: float64
```

In [27]: `df_corporate.shape`

Out[27]: (2017, 19)

In [28]: `df_corporate.mean()`

```
Out[28]: no_of_adults           1.230045
          no_of_children        0.009916
          no_of_weekend_nights   0.427863
          no_of_week_nights      1.488845
          lead_time              21.818047
          arrival_year            2017.753099
          arrival_month            7.103619
          arrival_date             15.695092
          no_of_previous_cancellations 0.166584
          no_of_previous_bookings_not_canceled 2.070402
          avg_price_per_room       82.911740
          no_of_special_requests   0.222112
          dtype: float64
```

In [29]: `df_complementary.shape`

Out[29]: (391, 19)

In [30]: `df_complementary.mean()`

```
Out[30]: no_of_adults           1.483376
          no_of_children        0.125320
          no_of_weekend_nights   0.329923
          no_of_week_nights      1.240409
          lead_time              12.035806
          arrival_year            2017.644501
          arrival_month            7.723785
          arrival_date             15.017903
          no_of_previous_cancellations 0.209719
          no_of_previous_bookings_not_canceled 2.475703
          avg_price_per_room       3.141765
          no_of_special_requests   0.882353
          dtype: float64
```

In [31]: `df_aviation.shape`

Out[31]: (125, 19)

In [32]: `df_aviation.mean()`

```
Out[32]: no_of_adults      1.016
          no_of_children    0.000
          no_of_weekend_nights 1.160
          no_of_week_nights    2.856
          lead_time            5.488
          arrival_year         2018.000
          arrival_month         7.120
          arrival_date          15.360
          no_of_previous_cancellations 0.040
          no_of_previous_bookings_not_canceled 0.208
          avg_price_per_room    100.704
          no_of_special_requests 0.000
          dtype: float64
```

Question 3: The average price for an online room is 112.25 The average price for an offline room is 91.63 The average price for an corporate room is 82.91 The average price for an complementary room is 3.14 The average price for an aviation room is 100.7

```
In [33]: df["booking_status"].value_counts()
```

```
Out[33]: Not_Canceled    24390
          Canceled        11885
          Name: booking_status, dtype: int64
```

```
In [34]: df["booking_status"].value_counts(1)
```

```
Out[34]: Not_Canceled    0.672364
          Canceled        0.327636
          Name: booking_status, dtype: float64
```

Question 4: 32.76% of bookings are canceled

```
In [35]: df_repeat = df.copy()
          df_repeat['repeated_guest'] = df_repeat['repeated_guest'].replace(['no'], np.nan)
          df_repeat = df_repeat.dropna()
          df_repeat.shape
```

```
Out[35]: (930, 19)
```

```
In [116...]: df_repeat["booking_status"].value_counts(1)
```

```
Out[116]: Not_Canceled    0.982796
          Canceled        0.017204
          Name: booking_status, dtype: float64
```

Question 5: 1.72% of repeat customers canceled.

```
In [37]: df["no_of_special_requests"].value_counts()
```

```
Out[37]: 0    19777
          1    11373
          2     4364
          3     675
          4      78
          5      8
          Name: no_of_special_requests, dtype: int64
```

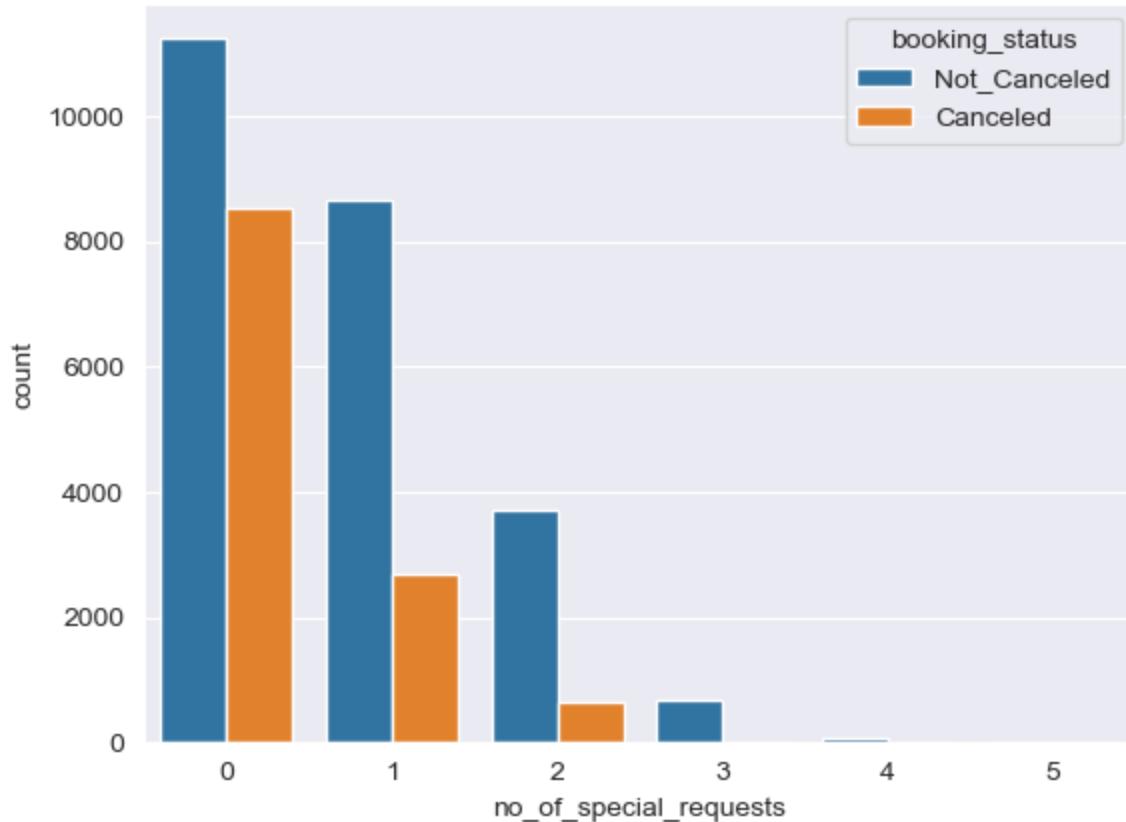
```
In [38]: print("There have been", 36275-19777,"people to have a special request.")
```

There have been 16498 people to have a special request.

Question 6: Yes, when people make special requests the cancelation rate goes down. There have been 16498 people to have at least one special request.

```
In [39]: sns.countplot(data=df, x='no_of_special_requests', hue='booking_status')
```

```
Out[39]: <Axes: xlabel='no_of_special_requests', ylabel='count'>
```



```
In [40]: df = df.drop(columns=['Booking_ID'])
```

```
In [41]: df['booking_status'] = df['booking_status'].replace(['Not_Canceled'], 0)
df['booking_status'] = df['booking_status'].replace(['Canceled'], 1)
df["booking_status"].value_counts(1)
```

```
Out[41]: 0    0.672364
1    0.327636
Name: booking_status, dtype: float64
```

```
In [42]: X = df.drop(["booking_status"], axis=1)
Y = df["booking_status"]
X = add_constant(X)
X = pd.get_dummies(X, drop_first=True)
X_train, X_test, y_train, y_test = train_test_split(
    X, Y, test_size=0.30, random_state=1, stratify=Y)
```

Here im creating my vars and splitting my data

```
In [43]: logit = sm.Logit(y_train, X_train.astype(float))
#lg = Logit.fit(max_iter=500)
```

```
In [44]: lg = logit.fit(max_iter=500)
```

```
Warning: Maximum number of iterations has been exceeded.  
Current function value: 0.422351  
Iterations: 35
```

```
In [45]: print(lg.summary())
```

Logit Regression Results

Dep. Variable:	booking_status	No. Observations:	25392
Model:	Logit	Df Residuals:	25364
Method:	MLE	Df Model:	27
Date:	Sat, 30 Sep 2023	Pseudo R-squ.:	0.3322
Time:	01:57:00	Log-Likelihood:	-10724.
converged:	False	LL-Null:	-16060.
Covariance Type:	nonrobust	LLR p-value:	0.000

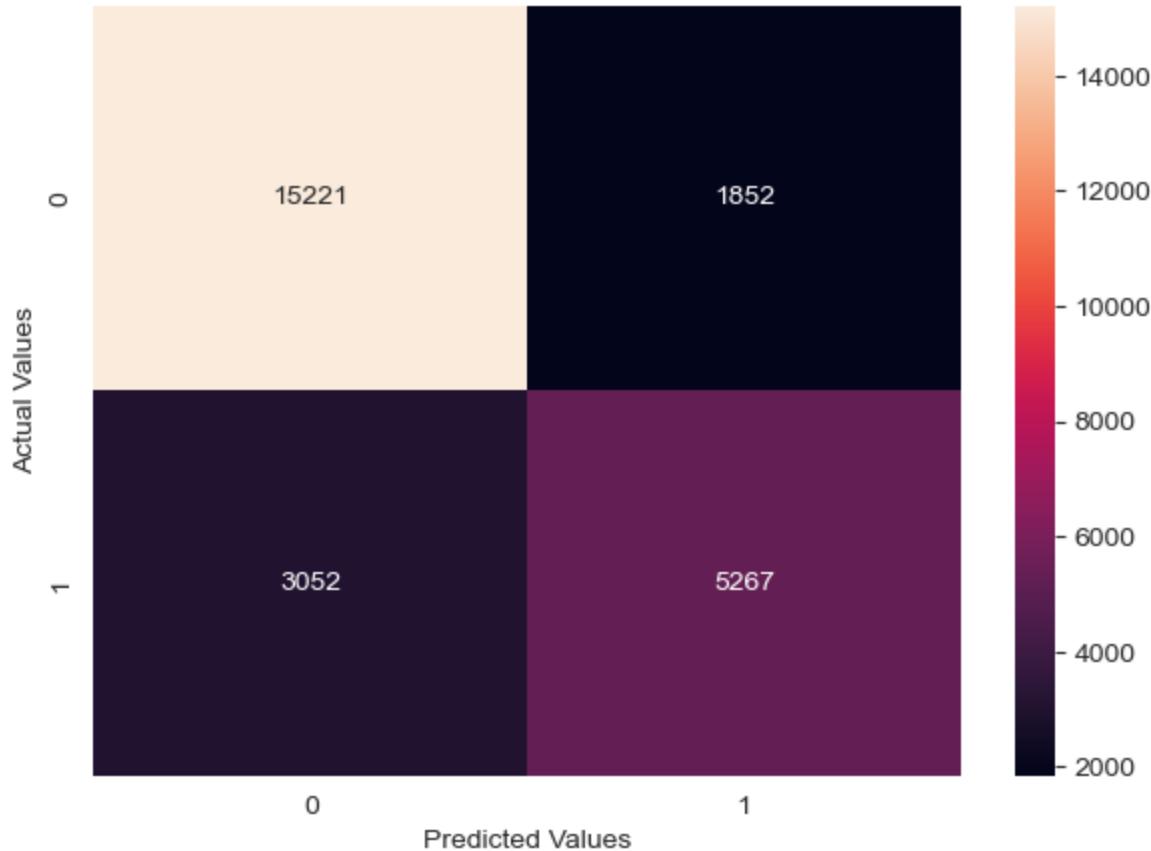
		coef	std err	z	P> z
[0.025	0.975]				
const		-886.9563	121.332	-7.310	0.000
124.762	-649.150				-1
no_of_adults		0.0333	0.038	0.883	0.377
-0.041	0.107				
no_of_children		0.0833	0.061	1.371	0.170
-0.036	0.202				
no_of_weekend_nights		0.1461	0.020	7.368	0.000
0.107	0.185				
no_of_week_nights		0.0354	0.012	2.881	0.004
0.011	0.059				
lead_time		0.0158	0.000	58.946	0.000
0.015	0.016				
arrival_year		0.4382	0.060	7.288	0.000
0.320	0.556				
arrival_month		-0.0475	0.006	-7.315	0.000
-0.060	-0.035				
arrival_date		0.0030	0.002	1.528	0.127
-0.001	0.007				
no_of_previous_cancellations		0.3476	0.102	3.413	0.001
0.148	0.547				
no_of_previous_bookings_not_canceled		-1.3821	0.906	-1.526	0.127
-3.157	0.393				
avg_price_per_room		0.0184	0.001	24.903	0.000
0.017	0.020				
no_of_special_requests		-1.4904	0.030	-48.963	0.000
-1.550	-1.431				
type_of_meal_plan_Meal Plan 2		0.1746	0.067	2.607	0.009
0.043	0.306				
type_of_meal_plan_Meal Plan 3		10.1526	107.971	0.094	0.925
201.466	221.771				-
type_of_meal_plan_Not Selected		0.1995	0.053	3.741	0.000
0.095	0.304				
required_car_parking_space_yes		-1.6146	0.137	-11.769	0.000
-1.883	-1.346				
room_type_reserved_Room_Type 2		-0.4172	0.133	-3.129	0.002
-0.679	-0.156				
room_type_reserved_Room_Type 3		1.1881	1.891	0.628	0.530
-2.518	4.895				
room_type_reserved_Room_Type 4		-0.2679	0.053	-5.019	0.000
-0.373	-0.163				
room_type_reserved_Room_Type 5		-0.6822	0.215	-3.173	0.002
-1.104	-0.261				
room_type_reserved_Room_Type 6		-0.8460	0.153	-5.534	0.000
-1.146	-0.546				
room_type_reserved_Room_Type 7		-1.3601	0.298	-4.566	0.000

```

-1.944      -0.776
market_segment_type_Complementary    -24.5570   2.24e+04   -0.001   0.999   -
4.39e+04    4.39e+04
market_segment_type_Corporate       -0.8514     0.276     -3.087   0.002
-1.392      -0.311
market_segment_type_Offline        -1.7645     0.264     -6.688   0.000
-2.282      -1.247
market_segment_type_Online         0.0070     0.261     0.027   0.979
-0.504      0.518
repeated_guest_yes                -1.9186     0.767     -2.502   0.012
-3.421      -0.416
=====
=====
```

```
In [46]: pred_train = lg.predict(X_train) > 0.5
pred_train = np.round(pred_train)
```

```
In [47]: cm = confusion_matrix(y_train, pred_train)
plt.figure(figsize=(7, 5))
sns.heatmap(cm, annot=True, fmt="g")
plt.xlabel("Predicted Values")
plt.ylabel("Actual Values")
plt.show()
```



```
In [48]: print("Accuracy on training set : ", accuracy_score(y_train, pred_train))
```

Accuracy on training set : 0.8068683049779458

The model is preforming fairly well

```
In [49]: vif_series = pd.Series ([variance_inflation_factor(X_train.values, i) for i in range(>)
print("VIF values: \n\n{} \n".format(vif_series))]
```

VIF values:

const	3.959545e+07
no_of_adults	1.345059e+00
no_of_children	2.007213e+00
no_of_weekend_nights	1.067255e+00
no_of_week_nights	1.094373e+00
lead_time	1.401920e+00
arrival_year	1.433261e+00
arrival_month	1.277398e+00
arrival_date	1.007629e+00
no_of_previous_cancellations	1.322011e+00
no_of_previous_bookings_not_canceled	1.570863e+00
avg_price_per_room	2.032631e+00
no_of_special_requests	1.247233e+00
type_of_meal_plan_Meal Plan 2	1.261818e+00
type_of_meal_plan_Meal Plan 3	1.007964e+00
type_of_meal_plan_Not Selected	1.279209e+00
required_car_parking_space_yes	1.034943e+00
room_type_reserved_Room_Type 2	1.094581e+00
room_type_reserved_Room_Type 3	1.003898e+00
room_type_reserved_Room_Type 4	1.355784e+00
room_type_reserved_Room_Type 5	1.030903e+00
room_type_reserved_Room_Type 6	1.990106e+00
room_type_reserved_Room_Type 7	1.092074e+00
market_segment_type_Complementary	4.350047e+00
market_segment_type_Corporate	1.663442e+01
market_segment_type_Offline	6.251301e+01
market_segment_type_Online	6.947183e+01
repeated_guest_yes	1.750194e+00
	dtype: float64

I will now drop all p-values greater than .05

```
In [50]: X_train1 = X_train.drop("market_segment_type_Complementary", axis=1)
logit1 = sm.Logit(y_train, X_train1.astype(float))
lg1 = logit1.fit()
pred_train1 = lg1.predict(X_train1)
pred_train1 = np.round(pred_train1)
print("Accuracy on training set : ", accuracy_score(y_train, pred_train1))
```

Optimization terminated successfully.

Current function value: 0.422630

Iterations 16

Accuracy on training set : 0.8063563327032136

```
In [51]: print(lg1.summary())
```

Logit Regression Results

Dep. Variable:	booking_status	No. Observations:	25392
Model:	Logit	Df Residuals:	25365
Method:	MLE	Df Model:	26
Date:	Sat, 30 Sep 2023	Pseudo R-squ.:	0.3318
Time:	01:57:01	Log-Likelihood:	-10731.
converged:	True	LL-Null:	-16060.
Covariance Type:	nonrobust	LLR p-value:	0.000

		coef	std err	z	P> z
[0.025	0.975]				
const		-889.3698	121.210	-7.337	0.000
126.936	-651.803				-1
no_of_adults		0.0274	0.038	0.729	0.466
-0.046	0.101				
no_of_children		0.0790	0.061	1.300	0.194
-0.040	0.198				
no_of_weekend_nights		0.1483	0.020	7.480	0.000
0.109	0.187				
no_of_week_nights		0.0368	0.012	3.003	0.003
0.013	0.061				
lead_time		0.0158	0.000	59.008	0.000
0.015	0.016				
arrival_year		0.4392	0.060	7.312	0.000
0.321	0.557				
arrival_month		-0.0478	0.006	-7.356	0.000
-0.060	-0.035				
arrival_date		0.0029	0.002	1.510	0.131
-0.001	0.007				
no_of_previous_cancellations		0.3469	0.102	3.406	0.001
0.147	0.547				
no_of_previous_bookings_not_canceled		-1.3680	0.898	-1.523	0.128
-3.128	0.392				
avg_price_per_room		0.0187	0.001	25.388	0.000
0.017	0.020				
no_of_special_requests		-1.4921	0.030	-49.021	0.000
-1.552	-1.432				
type_of_meal_plan_Meal Plan 2		0.1717	0.067	2.563	0.010
0.040	0.303				
type_of_meal_plan_Meal Plan 3		3.4639	2.902	1.194	0.233
-2.224	9.152				
type_of_meal_plan_Not Selected		0.2030	0.053	3.809	0.000
0.099	0.308				
required_car_parking_space_yes		-1.6165	0.137	-11.778	0.000
-1.885	-1.348				
room_type_reserved_Room_Type 2		-0.4154	0.133	-3.115	0.002
-0.677	-0.154				
room_type_reserved_Room_Type 3		0.8265	1.601	0.516	0.606
-2.312	3.965				
room_type_reserved_Room_Type 4		-0.2672	0.053	-5.005	0.000
-0.372	-0.163				
room_type_reserved_Room_Type 5		-0.6930	0.215	-3.227	0.001
-1.114	-0.272				
room_type_reserved_Room_Type 6		-0.8536	0.153	-5.585	0.000
-1.153	-0.554				
room_type_reserved_Room_Type 7		-1.3740	0.298	-4.615	0.000

-1.958	-0.790				
market_segment_type_Corporate		-0.4569	0.259	-1.766	0.077
-0.964	0.050				
market_segment_type_Offline		-1.3692	0.246	-5.573	0.000
-1.851	-0.888				
market_segment_type_Online		0.4001	0.243	1.646	0.100
-0.076	0.876				
repeated_guest_yes		-1.9144	0.765	-2.501	0.012
-3.415	-0.414				
=====					
=====					

```
In [52]: X_train2 = X_train1.drop("room_type_reserved_Room_Type 3", axis=1)
logit2 = sm.Logit(y_train, X_train2.astype(float))
lg2 = logit2.fit()

pred_train2 = lg2.predict(X_train2)
pred_train2 = np.round(pred_train2)

print("Accuracy on training set : ", accuracy_score(y_train, pred_train2))
```

Optimization terminated successfully.
 Current function value: 0.422635
 Iterations 16
 Accuracy on training set : 0.8063957151858853

```
In [53]: print(lg2.summary())
```

Logit Regression Results

Dep. Variable:	booking_status	No. Observations:	25392
Model:	Logit	Df Residuals:	25366
Method:	MLE	Df Model:	25
Date:	Sat, 30 Sep 2023	Pseudo R-squ.:	0.3318
Time:	01:57:01	Log-Likelihood:	-10732.
converged:	True	LL-Null:	-16060.
Covariance Type:	nonrobust	LLR p-value:	0.000

		coef	std err	z	P> z
[0.025	0.975]				
const		-889.0510	121.207	-7.335	0.000
126.612	-651.490				-1
no_of_adults		0.0274	0.038	0.729	0.466
-0.046	0.101				
no_of_children		0.0790	0.061	1.299	0.194
-0.040	0.198				
no_of_weekend_nights		0.1483	0.020	7.480	0.000
0.109	0.187				
no_of_week_nights		0.0368	0.012	2.999	0.003
0.013	0.061				
lead_time		0.0158	0.000	59.009	0.000
0.015	0.016				
arrival_year		0.4390	0.060	7.310	0.000
0.321	0.557				
arrival_month		-0.0477	0.006	-7.353	0.000
-0.060	-0.035				
arrival_date		0.0029	0.002	1.510	0.131
-0.001	0.007				
no_of_previous_cancellations		0.3469	0.102	3.406	0.001
0.147	0.547				
no_of_previous_bookings_not_canceled		-1.3681	0.898	-1.523	0.128
-3.129	0.392				
avg_price_per_room		0.0187	0.001	25.386	0.000
0.017	0.020				
no_of_special_requests		-1.4922	0.030	-49.025	0.000
-1.552	-1.433				
type_of_meal_plan_Meal Plan 2		0.1716	0.067	2.562	0.010
0.040	0.303				
type_of_meal_plan_Meal Plan 3		3.4623	2.900	1.194	0.233
-2.222	9.146				
type_of_meal_plan_Not Selected		0.2030	0.053	3.807	0.000
0.098	0.307				
required_car_parking_space_yes		-1.6165	0.137	-11.779	0.000
-1.886	-1.348				
room_type_reserved_Room_Type 2		-0.4155	0.133	-3.116	0.002
-0.677	-0.154				
room_type_reserved_Room_Type 4		-0.2672	0.053	-5.006	0.000
-0.372	-0.163				
room_type_reserved_Room_Type 5		-0.6929	0.215	-3.227	0.001
-1.114	-0.272				
room_type_reserved_Room_Type 6		-0.8534	0.153	-5.584	0.000
-1.153	-0.554				
room_type_reserved_Room_Type 7		-1.3738	0.298	-4.614	0.000
-1.957	-0.790				
market_segment_type_Corporate		-0.4602	0.259	-1.779	0.075

-0.967	0.047				
market_segment_type_Offline		-1.3724	0.246	-5.587	0.000
-1.854	-0.891				
market_segment_type_Online		0.3970	0.243	1.634	0.102
-0.079	0.873				
repeated_guest_yes		-1.9148	0.765	-2.501	0.012
-3.415	-0.414				
<hr/>					
<hr/>					

```
In [54]: X_train3 = X_train2.drop("no_of_adults", axis=1)
logit3 = sm.Logit(y_train, X_train3.astype(float))
lg3 = logit3.fit()

pred_train3 = lg3.predict(X_train3)
pred_train3 = np.round(pred_train3)

print("Accuracy on training set : ", accuracy_score(y_train, pred_train3))
```

Optimization terminated successfully.
Current function value: 0.422646
Iterations 16
Accuracy on training set : 0.8065926275992439

```
In [55]: print(lg3.summary())
```

Logit Regression Results

Dep. Variable:	booking_status	No. Observations:	25392
Model:	Logit	Df Residuals:	25367
Method:	MLE	Df Model:	24
Date:	Sat, 30 Sep 2023	Pseudo R-squ.:	0.3318
Time:	01:57:02	Log-Likelihood:	-10732.
converged:	True	LL-Null:	-16060.
Covariance Type:	nonrobust	LLR p-value:	0.000

		coef	std err	z	P> z
[0.025	0.975]				
const		-885.9115	121.142	-7.313	0.000
123.346	-648.477				-1
no_of_children		0.0737	0.060	1.220	0.222
-0.045	0.192				
no_of_weekend_nights		0.1489	0.020	7.520	0.000
0.110	0.188				
no_of_week_nights		0.0369	0.012	3.007	0.003
0.013	0.061				
lead_time		0.0158	0.000	59.388	0.000
0.015	0.016				
arrival_year		0.4375	0.060	7.288	0.000
0.320	0.555				
arrival_month		-0.0478	0.006	-7.369	0.000
-0.061	-0.035				
arrival_date		0.0030	0.002	1.534	0.125
-0.001	0.007				
no_of_previous_cancellations		0.3479	0.102	3.415	0.001
0.148	0.548				
no_of_previous_bookings_not_canceled		-1.3704	0.899	-1.525	0.127
-3.132	0.391				
avg_price_per_room		0.0187	0.001	25.747	0.000
0.017	0.020				
no_of_special_requests		-1.4898	0.030	-49.240	0.000
-1.549	-1.431				
type_of_meal_plan_Meal Plan 2		0.1723	0.067	2.573	0.010
0.041	0.304				
type_of_meal_plan_Meal Plan 3		3.4727	2.912	1.193	0.233
-2.235	9.180				
type_of_meal_plan_Not Selected		0.2056	0.053	3.865	0.000
0.101	0.310				
required_car_parking_space_yes		-1.6143	0.137	-11.767	0.000
-1.883	-1.345				
room_type_reserved_Room_Type 2		-0.4193	0.133	-3.146	0.002
-0.681	-0.158				
room_type_reserved_Room_Type 4		-0.2586	0.052	-4.969	0.000
-0.361	-0.157				
room_type_reserved_Room_Type 5		-0.6914	0.215	-3.219	0.001
-1.112	-0.270				
room_type_reserved_Room_Type 6		-0.8466	0.153	-5.546	0.000
-1.146	-0.547				
room_type_reserved_Room_Type 7		-1.3589	0.297	-4.576	0.000
-1.941	-0.777				
market_segment_type_Corporate		-0.4549	0.259	-1.758	0.079
-0.962	0.052				
market_segment_type_Offline		-1.3574	0.245	-5.541	0.000

-1.837	-0.877				
market_segment_type_Online		0.4148	0.242	1.715	0.086
-0.059	0.889				
repeated_guest_yes		-1.9179	0.768	-2.498	0.012
-3.423	-0.413				
=====					
=====					

```
In [56]: X_train4 = X_train3.drop("no_of_children", axis=1)
logit4 = sm.Logit(y_train, X_train4.astype(float))
lg4 = logit4.fit()

pred_train4 = lg4.predict(X_train4)
pred_train4 = np.round(pred_train4)

print("Accuracy on training set : ", accuracy_score(y_train, pred_train4))
```

Optimization terminated successfully.
Current function value: 0.422675
Iterations 16
Accuracy on training set : 0.806710775047259

```
In [57]: print(lg4.summary())
```

Logit Regression Results

Dep. Variable:	booking_status	No. Observations:	25392
Model:	Logit	Df Residuals:	25368
Method:	MLE	Df Model:	23
Date:	Sat, 30 Sep 2023	Pseudo R-squ.:	0.3317
Time:	01:57:02	Log-Likelihood:	-10733.
converged:	True	LL-Null:	-16060.
Covariance Type:	nonrobust	LLR p-value:	0.000

		coef	std err	z	P> z
[0.025	0.975]				
const		-889.7590	121.092	-7.348	0.000
127.095	-652.423				-1
no_of_weekend_nights		0.1492	0.020	7.532	0.000
0.110	0.188				
no_of_week_nights		0.0370	0.012	3.017	0.003
0.013	0.061				
lead_time		0.0158	0.000	59.374	0.000
0.015	0.016				
arrival_year		0.4394	0.060	7.322	0.000
0.322	0.557				
arrival_month		-0.0479	0.006	-7.369	0.000
-0.061	-0.035				
arrival_date		0.0030	0.002	1.551	0.121
-0.001	0.007				
no_of_previous_cancellations		0.3481	0.102	3.416	0.001
0.148	0.548				
no_of_previous_bookings_not_canceled		-1.3706	0.899	-1.525	0.127
-3.132	0.391				
avg_price_per_room		0.0188	0.001	26.059	0.000
0.017	0.020				
no_of_special_requests		-1.4882	0.030	-49.242	0.000
-1.547	-1.429				
type_of_meal_plan_Meal Plan 2		0.1712	0.067	2.556	0.011
0.040	0.303				
type_of_meal_plan_Meal Plan 3		3.4722	2.925	1.187	0.235
-2.261	9.205				
type_of_meal_plan_Not Selected		0.2001	0.053	3.776	0.000
0.096	0.304				
required_car_parking_space_yes		-1.6149	0.137	-11.773	0.000
-1.884	-1.346				
room_type_reserved_Room_Type 2		-0.3797	0.129	-2.939	0.003
-0.633	-0.126				
room_type_reserved_Room_Type 4		-0.2660	0.052	-5.145	0.000
-0.367	-0.165				
room_type_reserved_Room_Type 5		-0.6838	0.215	-3.185	0.001
-1.105	-0.263				
room_type_reserved_Room_Type 6		-0.7323	0.120	-6.097	0.000
-0.968	-0.497				
room_type_reserved_Room_Type 7		-1.3033	0.292	-4.465	0.000
-1.875	-0.731				
market_segment_type_Corporate		-0.4605	0.259	-1.779	0.075
-0.968	0.047				
market_segment_type_Offline		-1.3615	0.245	-5.556	0.000
-1.842	-0.881				
market_segment_type_Online		0.4150	0.242	1.714	0.086

```
-0.059      0.889
repeated_guest_yes           -1.9184      0.768      -2.498      0.012
-3.424      -0.413
=====
=====
```

```
In [58]: X_train5 = X_train4.drop("type_of_meal_plan_Meal Plan 3", axis=1)
logit5 = sm.Logit(y_train, X_train5.astype(float))
lg5 = logit5.fit()

pred_train5 = lg5.predict(X_train5)
pred_train5 = np.round(pred_train5)

print("Accuracy on training set : ", accuracy_score(y_train, pred_train5))
```

```
Optimization terminated successfully.
    Current function value: 0.422715
    Iterations 16
Accuracy on training set :  0.8066713925645873
```

```
In [59]: print(lg5.summary())
```

Logit Regression Results

Dep. Variable:	booking_status	No. Observations:	25392
Model:	Logit	Df Residuals:	25369
Method:	MLE	Df Model:	22
Date:	Sat, 30 Sep 2023	Pseudo R-squ.:	0.3317
Time:	01:57:02	Log-Likelihood:	-10734.
converged:	True	LL-Null:	-16060.
Covariance Type:	nonrobust	LLR p-value:	0.000

		coef	std err	z	P> z
[0.025	0.975]				
const		-885.5409	121.030	-7.317	0.000
122.755	-648.326				-1
no_of_weekend_nights		0.1491	0.020	7.529	0.000
0.110	0.188				
no_of_week_nights		0.0370	0.012	3.014	0.003
0.013	0.061				
lead_time		0.0158	0.000	59.376	0.000
0.015	0.016				
arrival_year		0.4373	0.060	7.291	0.000
0.320	0.555				
arrival_month		-0.0479	0.006	-7.381	0.000
-0.061	-0.035				
arrival_date		0.0030	0.002	1.561	0.119
-0.001	0.007				
no_of_previous_cancellations		0.3481	0.102	3.416	0.001
0.148	0.548				
no_of_previous_bookings_not_canceled		-1.3695	0.899	-1.524	0.128
-3.131	0.392				
avg_price_per_room		0.0189	0.001	26.112	0.000
0.017	0.020				
no_of_special_requests		-1.4882	0.030	-49.243	0.000
-1.547	-1.429				
type_of_meal_plan_Meal Plan 2		0.1695	0.067	2.531	0.011
0.038	0.301				
type_of_meal_plan_Not Selected		0.2005	0.053	3.785	0.000
0.097	0.304				
required_car_parking_space_yes		-1.6142	0.137	-11.769	0.000
-1.883	-1.345				
room_type_reserved_Room_Type 2		-0.3794	0.129	-2.937	0.003
-0.633	-0.126				
room_type_reserved_Room_Type 4		-0.2666	0.052	-5.156	0.000
-0.368	-0.165				
room_type_reserved_Room_Type 5		-0.6848	0.215	-3.190	0.001
-1.106	-0.264				
room_type_reserved_Room_Type 6		-0.7349	0.120	-6.119	0.000
-0.970	-0.500				
room_type_reserved_Room_Type 7		-1.3066	0.292	-4.476	0.000
-1.879	-0.734				
market_segment_type_Corporate		-0.4650	0.259	-1.797	0.072
-0.972	0.042				
market_segment_type_Offline		-1.3650	0.245	-5.571	0.000
-1.845	-0.885				
market_segment_type_Online		0.4103	0.242	1.696	0.090
-0.064	0.885				
repeated_guest_yes		-1.9187	0.768	-2.499	0.012

```
-3.424      -0.414
```

```
=====
```

```
=====
```

```
In [60]: X_train6 = X_train5.drop("no_of_previous_bookings_not_canceled", axis=1)
logit6 = sm.Logit(y_train, X_train6.astype(float))
lg6 = logit6.fit()

pred_train6 = lg6.predict(X_train6)
pred_train6 = np.round(pred_train6)

print("Accuracy on training set : ", accuracy_score(y_train, pred_train6))
```

Optimization terminated successfully.

Current function value: 0.422882

Iterations 11

Accuracy on training set : 0.806710775047259

```
In [61]: print(lg6.summary())
```

Logit Regression Results

Dep. Variable:	booking_status	No. Observations:	25392		
Model:	Logit	Df Residuals:	25370		
Method:	MLE	Df Model:	21		
Date:	Sat, 30 Sep 2023	Pseudo R-squ.:	0.3314		
Time:	01:57:02	Log-Likelihood:	-10738.		
converged:	True	LL-Null:	-16060.		
Covariance Type:	nonrobust	LLR p-value:	0.000		
5	0.975]				
const	-881.3654	121.049	-7.281	0.000	-1118.61
7	-644.114				
no_of_weekend_nights	0.1491	0.020	7.529	0.000	0.11
0	0.188				
no_of_week_nights	0.0371	0.012	3.023	0.003	0.01
3	0.061				
lead_time	0.0158	0.000	59.463	0.000	0.01
5	0.016				
arrival_year	0.4352	0.060	7.256	0.000	0.31
8	0.553				
arrival_month	-0.0479	0.006	-7.377	0.000	-0.06
1	-0.035				
arrival_date	0.0030	0.002	1.528	0.126	-0.00
1	0.007				
no_of_previous_cancellations	0.2868	0.078	3.695	0.000	0.13
5	0.439				
avg_price_per_room	0.0189	0.001	26.138	0.000	0.01
7	0.020				
no_of_special_requests	-1.4896	0.030	-49.287	0.000	-1.54
9	-1.430				
type_of_meal_plan_Meal Plan 2	0.1687	0.067	2.518	0.012	0.03
7	0.300				
type_of_meal_plan_Not Selected	0.2021	0.053	3.814	0.000	0.09
8	0.306				
required_car_parking_space_yes	-1.6106	0.137	-11.745	0.000	-1.87
9	-1.342				
room_type_reserved_Room_Type 2	-0.3798	0.129	-2.938	0.003	-0.63
3	-0.126				
room_type_reserved_Room_Type 4	-0.2670	0.052	-5.163	0.000	-0.36
8	-0.166				
room_type_reserved_Room_Type 5	-0.6844	0.215	-3.190	0.001	-1.10
5	-0.264				
room_type_reserved_Room_Type 6	-0.7362	0.120	-6.129	0.000	-0.97
2	-0.501				
room_type_reserved_Room_Type 7	-1.3078	0.292	-4.479	0.000	-1.88
0	-0.735				
market_segment_type_Corporate	-0.4727	0.259	-1.828	0.068	-0.98
0	0.034				
market_segment_type_Offline	-1.3630	0.245	-5.566	0.000	-1.84
3	-0.883				
market_segment_type_Online	0.4127	0.242	1.707	0.088	-0.06
1	0.887				
repeated_guest_yes	-3.0382	0.597	-5.086	0.000	-4.20
9	-1.867				

```
=====
=====
In [62]: X_train7 = X_train6.drop("arrival_date", axis=1)
logit7 = sm.Logit(y_train, X_train7.astype(float))
lg7 = logit7.fit()

pred_train7 = lg7.predict(X_train7)
pred_train7 = np.round(pred_train7)

print("Accuracy on training set : ", accuracy_score(y_train, pred_train7))
```

Optimization terminated successfully.
Current function value: 0.422928
Iterations 11
Accuracy on training set : 0.8061988027725268

```
In [63]: print(lg7.summary())
```

Logit Regression Results

Dep. Variable:	booking_status	No. Observations:	25392
Model:	Logit	Df Residuals:	25371
Method:	MLE	Df Model:	20
Date:	Sat, 30 Sep 2023	Pseudo R-squ.:	0.3313
Time:	01:57:02	Log-Likelihood:	-10739.
converged:	True	LL-Null:	-16060.
Covariance Type:	nonrobust	LLR p-value:	0.000
5	0.975]		
const	-881.8023	121.089	-7.282
3	-644.472		0.000
no_of_weekend_nights	0.1499	0.020	7.568
1	0.189		0.000
no_of_week_nights	0.0367	0.012	2.990
3	0.061		0.003
lead_time	0.0158	0.000	59.469
5	0.016		0.000
arrival_year	0.4355	0.060	7.257
8	0.553		0.000
arrival_month	-0.0485	0.006	-7.476
1	-0.036		0.000
no_of_previous_cancellations	0.2866	0.078	3.683
4	0.439		0.000
avg_price_per_room	0.0189	0.001	26.128
7	0.020		0.000
no_of_special_requests	-1.4881	0.030	-49.284
7	-1.429		0.000
type_of_meal_plan_Meal Plan 2	0.1712	0.067	2.557
0	0.302		0.011
type_of_meal_plan_Not Selected	0.2023	0.053	3.819
8	0.306		0.000
required_car_parking_space_yes	-1.6121	0.137	-11.759
1	-1.343		0.000
room_type_reserved_Room_Type 2	-0.3771	0.129	-2.919
0	-0.124		0.004
room_type_reserved_Room_Type 4	-0.2646	0.052	-5.121
6	-0.163		0.000
room_type_reserved_Room_Type 5	-0.6848	0.214	-3.195
5	-0.265		0.001
room_type_reserved_Room_Type 6	-0.7340	0.120	-6.112
9	-0.499		0.000
room_type_reserved_Room_Type 7	-1.3033	0.292	-4.462
6	-0.731		0.000
market_segment_type_Corporate	-0.4674	0.259	-1.807
4	0.040		0.071
market_segment_type_Offline	-1.3644	0.245	-5.569
5	-0.884		0.000
market_segment_type_Online	0.4128	0.242	1.706
1	0.887		0.088
repeated_guest_yes	-3.0488	0.600	-5.085
4	-1.874		0.000

```
In [64]: X_train8 = X_train7.drop("market_segment_type_Online", axis=1)
logit8 = sm.Logit(y_train, X_train8.astype(float))
lg8 = logit8.fit()

pred_train8 = lg8.predict(X_train8)
pred_train8 = np.round(pred_train8)

print("Accuracy on training set : ", accuracy_score(y_train, pred_train8))
```

Optimization terminated successfully.
Current function value: 0.422990
Iterations 11
Accuracy on training set : 0.8065926275992439

```
In [65]: print(lg8.summary())
```

Logit Regression Results

Dep. Variable:	booking_status	No. Observations:	25392
Model:	Logit	Df Residuals:	25372
Method:	MLE	Df Model:	19
Date:	Sat, 30 Sep 2023	Pseudo R-squ.:	0.3312
Time:	01:57:02	Log-Likelihood:	-10741.
converged:	True	LL-Null:	-16060.
Covariance Type:	nonrobust	LLR p-value:	0.000
5	0.975]		
const	-869.8889	120.911	-7.194
0	-632.908		0.000
no_of_weekend_nights	0.1498	0.020	7.568
1	0.189		0.000
no_of_week_nights	0.0362	0.012	2.953
2	0.060		0.003
lead_time	0.0159	0.000	59.908
5	0.016		0.000
arrival_year	0.4298	0.060	7.172
2	0.547		0.000
arrival_month	-0.0492	0.006	-7.595
2	-0.036		0.000
no_of_previous_cancellations	0.2892	0.078	3.722
7	0.441		0.000
avg_price_per_room	0.0190	0.001	26.645
8	0.020		0.000
no_of_special_requests	-1.4855	0.030	-49.272
5	-1.426		0.000
type_of_meal_plan_Meal Plan 2	0.1669	0.067	2.492
6	0.298		0.013
type_of_meal_plan_Not Selected	0.2098	0.053	3.974
6	0.313		0.000
required_car_parking_space_yes	-1.6148	0.137	-11.780
3	-1.346		0.000
room_type_reserved_Room_Type 2	-0.3734	0.129	-2.891
7	-0.120		0.004
room_type_reserved_Room_Type 4	-0.2678	0.052	-5.189
9	-0.167		0.000
room_type_reserved_Room_Type 5	-0.6879	0.214	-3.211
8	-0.268		0.001
room_type_reserved_Room_Type 6	-0.7419	0.120	-6.186
7	-0.507		0.000
room_type_reserved_Room_Type 7	-1.3150	0.292	-4.504
7	-0.743		0.000
market_segment_type_Corporate	-0.8710	0.103	-8.445
3	-0.669		0.000
market_segment_type_Offline	-1.7725	0.052	-34.288
4	-1.671		0.000
repeated_guest_yes	-3.0747	0.597	-5.152
4	-1.905		0.000

```
In [66]: X_train9 = X_train8.drop("type_of_meal_plan_Meal Plan 2", axis=1)
logit9 = sm.Logit(y_train, X_train9.astype(float))
```

```
lg9 = logit9.fit()  
  
pred_train9 = lg9.predict(X_train9)  
pred_train9 = np.round(pred_train9)  
  
print("Accuracy on training set : ", accuracy_score(y_train, pred_train9))
```

Optimization terminated successfully.
Current function value: 0.423112
Iterations 11
Accuracy on training set : 0.806710775047259

In [67]: print(lg9.summary())

Logit Regression Results

Dep. Variable:	booking_status	No. Observations:	25392
Model:	Logit	Df Residuals:	25373
Method:	MLE	Df Model:	18
Date:	Sat, 30 Sep 2023	Pseudo R-squ.:	0.3310
Time:	01:57:03	Log-Likelihood:	-10744.
converged:	True	LL-Null:	-16060.
Covariance Type:	nonrobust	LLR p-value:	0.000
5	0.975]		
const	-796.7090	117.085	-6.805
1	-567.227		0.000
no_of_weekend_nights	0.1504	0.020	7.598
2	0.189		0.000
no_of_week_nights	0.0350	0.012	2.862
1	0.059		0.004
lead_time	0.0160	0.000	61.153
5	0.017		0.000
arrival_year	0.3935	0.058	6.781
0	0.507		0.000
arrival_month	-0.0511	0.006	-7.941
4	-0.038		0.000
no_of_previous_cancellations	0.2876	0.078	3.688
5	0.440		0.000
avg_price_per_room	0.0195	0.001	28.194
8	0.021		0.000
no_of_special_requests	-1.4823	0.030	-49.287
1	-1.423		0.000
type_of_meal_plan_Not Selected	0.2113	0.053	4.001
8	0.315		0.000
required_car_parking_space_yes	-1.6199	0.137	-11.826
8	-1.351		0.000
room_type_reserved_Room_Type 2	-0.3772	0.129	-2.918
1	-0.124		0.004
room_type_reserved_Room_Type 4	-0.2764	0.051	-5.367
7	-0.175		0.000
room_type_reserved_Room_Type 5	-0.7014	0.214	-3.275
1	-0.282		0.001
room_type_reserved_Room_Type 6	-0.7726	0.119	-6.480
6	-0.539		0.000
room_type_reserved_Room_Type 7	-1.3635	0.292	-4.673
5	-0.792		0.000
market_segment_type_Corporate	-0.8694	0.103	-8.424
2	-0.667		0.000
market_segment_type_Offline	-1.7440	0.050	-34.665
3	-1.645		0.000
repeated_guest_yes	-3.0657	0.599	-5.120
9	-1.892		0.000

In [68]: odds = np.exp(lg9.params)

```
# adding the odds to a dataframe
pd.DataFrame(odds, X_train9.columns, columns=["odds"]).T
```

Out[68]:

	const	no_of_weekend_nights	no_of_week_nights	lead_time	arrival_year	arrival_month	no_of_previous_cancellations
odds	0.0	1.162288		1.035666	1.016125	1.482135	0.950189

In [69]:

```
# finding the percentage change
perc_change_odds = (np.exp(lg9.params) - 1) * 100

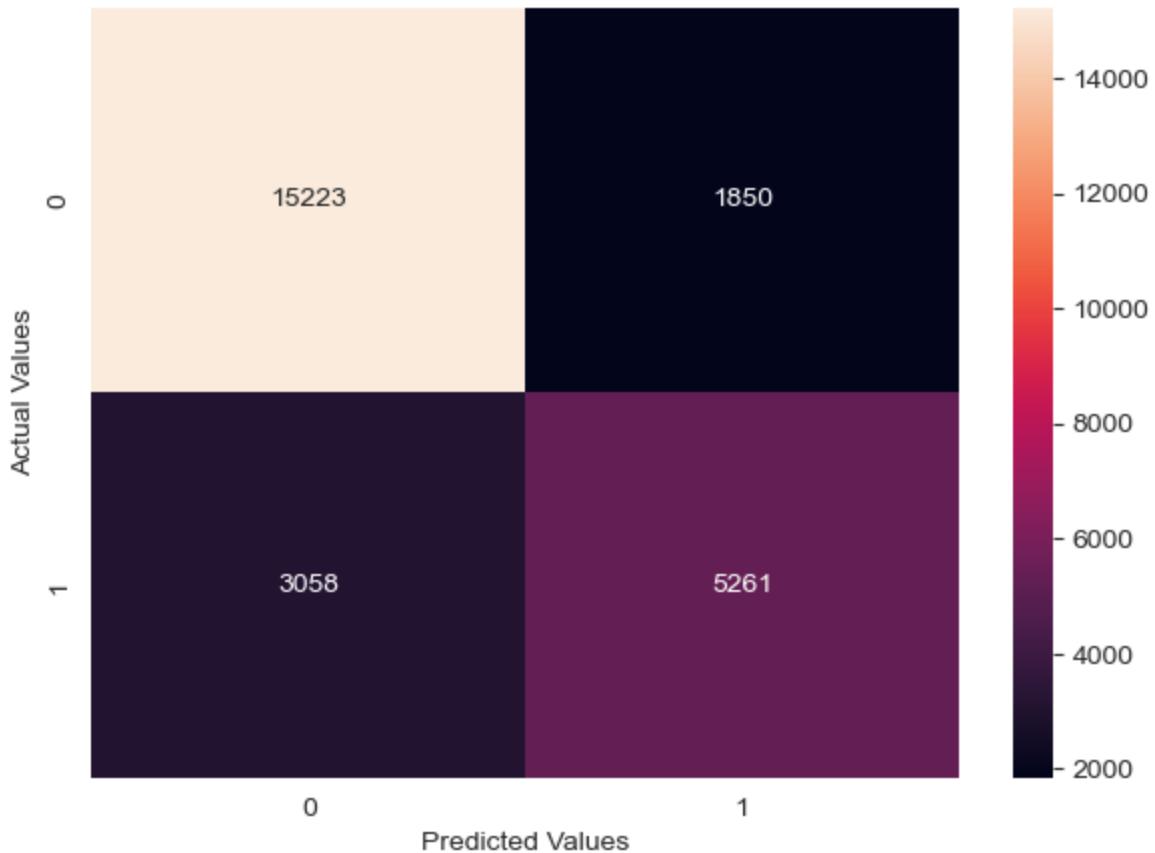
# adding the change_odds% to a dataframe
pd.DataFrame(perc_change_odds, X_train9.columns, columns=["change_odds%"])
```

Out[69]:

	change_odds%
const	-100.000000
no_of_weekend_nights	16.228751
no_of_week_nights	3.566640
lead_time	1.612464
arrival_year	48.213533
arrival_month	-4.981121
no_of_previous_cancellations	33.320739
avg_price_per_room	1.969566
no_of_special_requests	-77.288387
type_of_meal_plan_Not Selected	23.532481
required_car_parking_space_yes	-80.208023
room_type_reserved_Room_Type 2	-31.422024
room_type_reserved_Room_Type 4	-24.147562
room_type_reserved_Room_Type 5	-50.408765
room_type_reserved_Room_Type 6	-53.817498
room_type_reserved_Room_Type 7	-74.423216
market_segment_type_Corporate	-58.081635
market_segment_type_Offline	-82.517173
repeated_guest_yes	-95.337927

In [70]:

```
cm = confusion_matrix(y_train, pred_train9)
plt.figure(figsize=(7, 5))
sns.heatmap(cm, annot=True, fmt="g")
plt.xlabel("Predicted Values")
plt.ylabel("Actual Values")
plt.show()
```



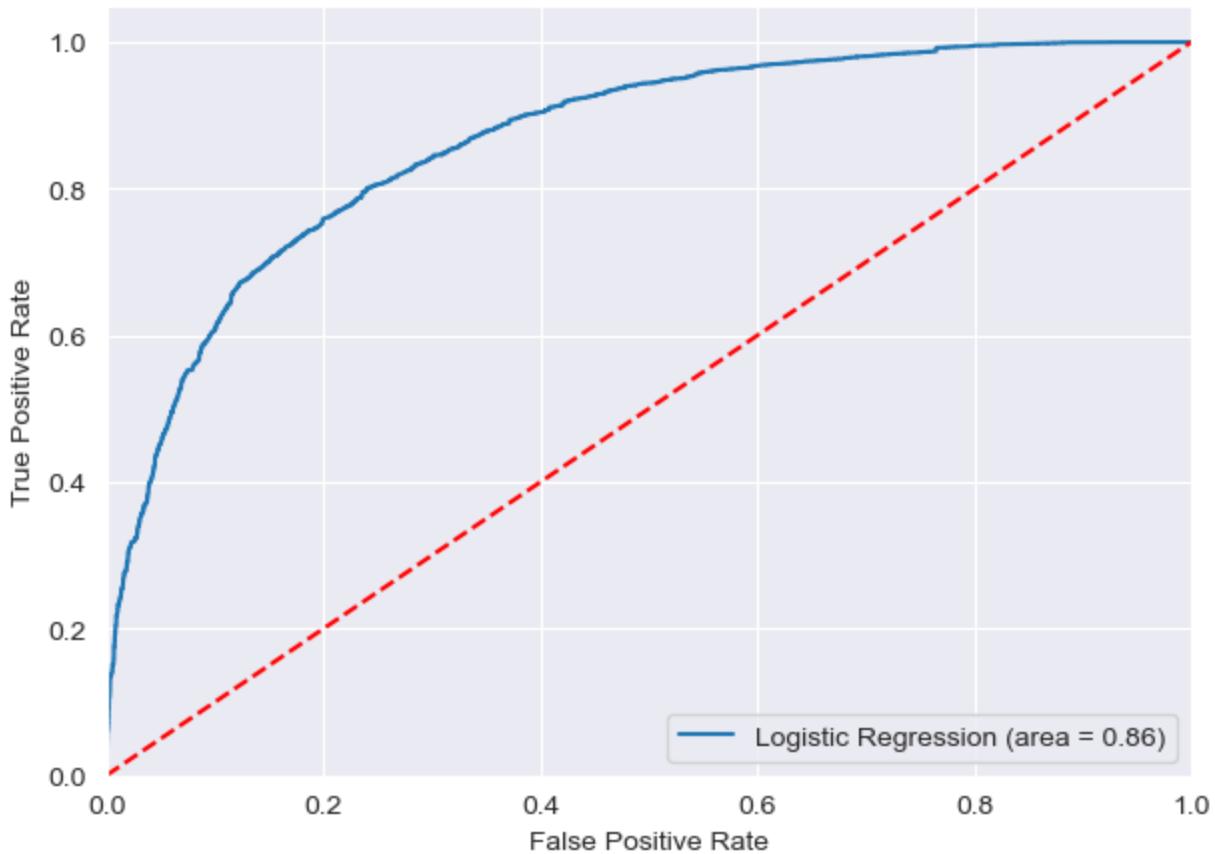
It seems the Predicted value seems twice as likely to incorrectly guess people wont be there when they actually are

```
In [71]: print("Accuracy on training set : ", accuracy_score(y_train, pred_train9))
```

```
Accuracy on training set :  0.806710775047259
```

```
In [72]: logit_roc_auc_train = roc_auc_score(y_train, lg9.predict(X_train9))
fpr, tpr, thresholds = roc_curve(y_train, lg9.predict(X_train9))
plt.figure(figsize=(7, 5))
plt.plot(fpr, tpr, label="Logistic Regression (area = %0.2f)" % logit_roc_auc_train)
plt.plot([0, 1], [0, 1], "r--")
plt.xlim([0.0, 1.0])
plt.ylim([0.0, 1.05])
plt.xlabel("False Positive Rate")
plt.ylabel("True Positive Rate")
plt.title("Receiver operating characteristic")
plt.legend(loc="lower right")
plt.show()
```

Receiver operating characteristic



```
In [73]: X_test = X_test.drop(["market_segment_type_Complementary", "room_type_reserved_Room_Type", "no_of_adults", "no_of_children", "type_of_meal_plan_Meal Plan 3", "no_of_previous_bookings_not_canceled", "arrival_date", "market_segment_type", "type_of_meal_plan_Meal Plan 2"], axis=1)
```

```
In [74]: pred_test = lg9.predict(X_test) > 0.5
pred_test = np.round(pred_test)
print("Accuracy on training set : ", accuracy_score(y_train, pred_train4))
print("Accuracy on test set : ", accuracy_score(y_test, pred_test))
```

Accuracy on training set : 0.806710775047259
 Accuracy on test set : 0.8020766332812643

The final model preformed well with an 80% accuracy

```
In [75]: odds = np.exp(lg9.params)
perc_change_odds = (np.exp(lg9.params) - 1) * 100
pd.set_option("display.max_columns", None)
pd.DataFrame({"Odds": odds, "Change_odd%": perc_change_odds}, index=X_train3.columns).
```

	const	no_of_children	no_of_weekend_nights	no_of_week_nights	lead_time	arrival_year
Odds	0.0	NaN	1.162288	1.035666	1.016125	1.48213
Change_odd%	-100.0	NaN	16.228751	3.566640	1.612464	48.21353

```
In [76]: df['repeated_guest'].value_counts()
```

```
Out[76]: no      35345
yes      930
Name: repeated_guest, dtype: int64
```

```
In [77]: col = ["required_car_parking_space"]
df[col] = df[col].replace("no", 0)
df[col] = df[col].replace("yes", 1)
col = ["type_of_meal_plan"]
df[col] = df[col].replace("Not Selected", 0)
df[col] = df[col].replace("Meal Plan 1", 1)
df[col] = df[col].replace("Meal Plan 2", 2)
df[col] = df[col].replace("Meal Plan 3", 3)
col = ["room_type_reserved"]
df[col] = df[col].replace("Room_Type 1", 1)
df[col] = df[col].replace("Room_Type 2", 2)
df[col] = df[col].replace("Room_Type 3", 3)
df[col] = df[col].replace("Room_Type 4", 4)
df[col] = df[col].replace("Room_Type 5", 5)
df[col] = df[col].replace("Room_Type 6", 6)
df[col] = df[col].replace("Room_Type 7", 7)
col = ["market_segment_type"]
df[col] = df[col].replace("Online", 1)
df[col] = df[col].replace("Offline", 2)
df[col] = df[col].replace("Corporate", 3)
df[col] = df[col].replace("Complementary", 4)
df[col] = df[col].replace("Aviation", 5)
col = ["repeated_guest"]
df[col] = df[col].replace("no", 0)
df[col] = df[col].replace("yes", 2)
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 36275 entries, 0 to 36274
Data columns (total 18 columns):
 #   Column           Non-Null Count Dtype
 ---  -- 
 0   no_of_adults     36275 non-null  int64
 1   no_of_children   36275 non-null  int64
 2   no_of_weekend_nights  36275 non-null  int64
 3   no_of_week_nights 36275 non-null  int64
 4   type_of_meal_plan 36275 non-null  int64
 5   required_car_parking_space 36275 non-null  int64
 6   room_type_reserved 36275 non-null  int64
 7   lead_time         36275 non-null  int64
 8   arrival_year      36275 non-null  int64
 9   arrival_month     36275 non-null  int64
 10  arrival_date      36275 non-null  int64
 11  market_segment_type 36275 non-null  int64
 12  repeated_guest    36275 non-null  int64
 13  no_of_previous_cancellations 36275 non-null  int64
 14  no_of_previous_bookings_not_canceled 36275 non-null  int64
 15  avg_price_per_room 36275 non-null  float64
 16  no_of_special_requests 36275 non-null  int64
 17  booking_status    36275 non-null  int64
dtypes: float64(1), int64(17)
memory usage: 5.0 MB
```

```
In [78]: X = df.drop("booking_status" , axis=1)
y = df["booking_status"]
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=.30, random_state=
```

```
In [79]: X = pd.get_dummies(X, drop_first=True)
X.head()
```

Out[79]:

	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	type_of_meal_plan	requirec
0	2	0	1	2	1	
1	2	0	2	3	0	
2	1	0	2	1	1	
3	2	0	0	2	1	
4	2	0	1	1	0	

```
In [80]: print("Percentage of classes in training set:")
print(y_train.value_counts(normalize=True))
print("Percentage of classes in test set:")
print(y_test.value_counts(normalize=True))
```

Percentage of classes in training set:
0 0.670644
1 0.329356
Name: booking_status, dtype: float64
Percentage of classes in test set:
0 0.676376
1 0.323624
Name: booking_status, dtype: float64

```
In [81]: model = DecisionTreeClassifier(criterion="gini", random_state=1)
model.fit(X_train, y_train)
```

Out[81]:

DecisionTreeClassifier(random_state=1)

```
In [82]: print("Accuracy on training set : ",model.score(X_train, y_train))
print("Accuracy on test set : ",model.score(X_test, y_test))
```

Accuracy on training set : 0.994210775047259
Accuracy on test set : 0.8712671138472847

```
In [83]: y.sum(axis = 0)
```

Out[83]: 11885

```
In [84]: def model_performance_classification_sklearn(model, predictors, target):
    """
        Function to compute different metrics to check classification model performance
        model: classifier
        predictors: independent variables
        target: dependent variable
    
```

```

    """
# predicting using the independent variables
pred = model.predict(predictors)

acc = accuracy_score(target, pred) # to compute Accuracy
recall = recall_score(target, pred) # to compute Recall
precision = precision_score(target, pred) # to compute Precision
f1 = f1_score(target, pred) # to compute F1-score

# creating a dataframe of metrics
df_perf = pd.DataFrame(
    {"Accuracy": acc, "Recall": recall, "Precision": precision, "F1": f1},
    index=[0],
)

return df_perf

```

In [85]: `def confusion_matrix_sklearn(model, predictors, target):`

```

    """
    To plot the confusion_matrix with percentages

    model: classifier
    predictors: independent variables
    target: dependent variable
    """

    y_pred = model.predict(predictors)
    cm = confusion_matrix(target, y_pred)
    labels = np.asarray(
        [
            ["{0:0.0f}".format(item) + "\n{0:.2%}".format(item / cm.flatten().sum())]
            for item in cm.flatten()
        ]
    ).reshape(2, 2)

    plt.figure(figsize=(6, 4))
    sns.heatmap(cm, annot=labels, fmt="")
    plt.ylabel("True label")
    plt.xlabel("Predicted label")

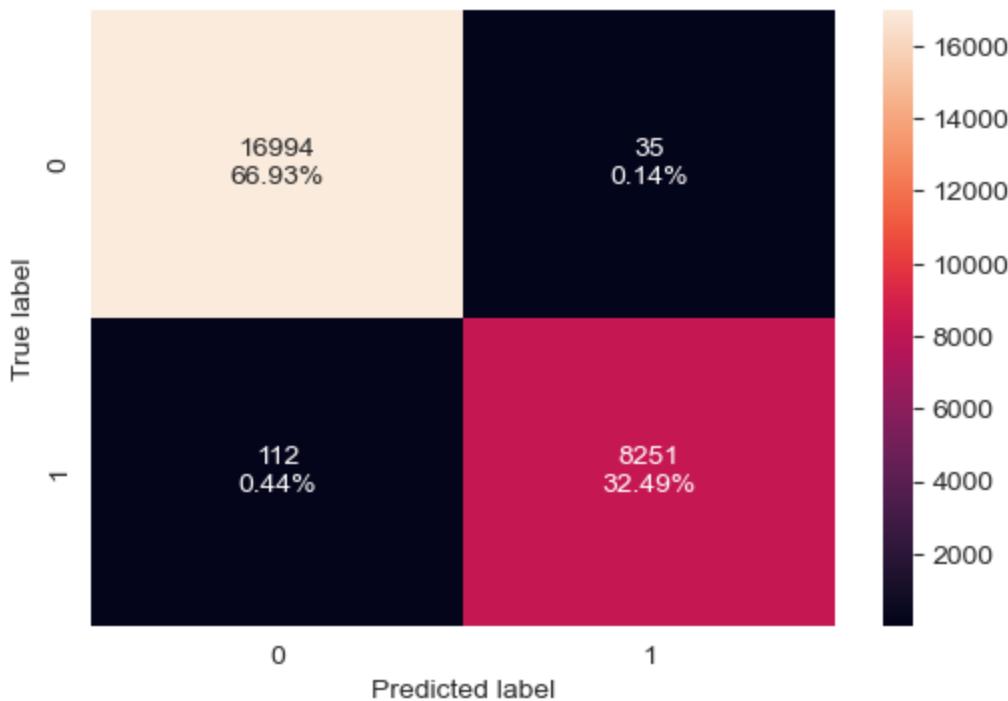
```

In [86]: `decision_tree_perf_train = model_performance_classification_sklearn(model, X_train, y_train)`
`decision_tree_perf_train`

Out[86]:

	Accuracy	Recall	Precision	F1
0	0.994211	0.986608	0.995776	0.991171

In [87]: `confusion_matrix_sklearn(model, X_train, y_train)`

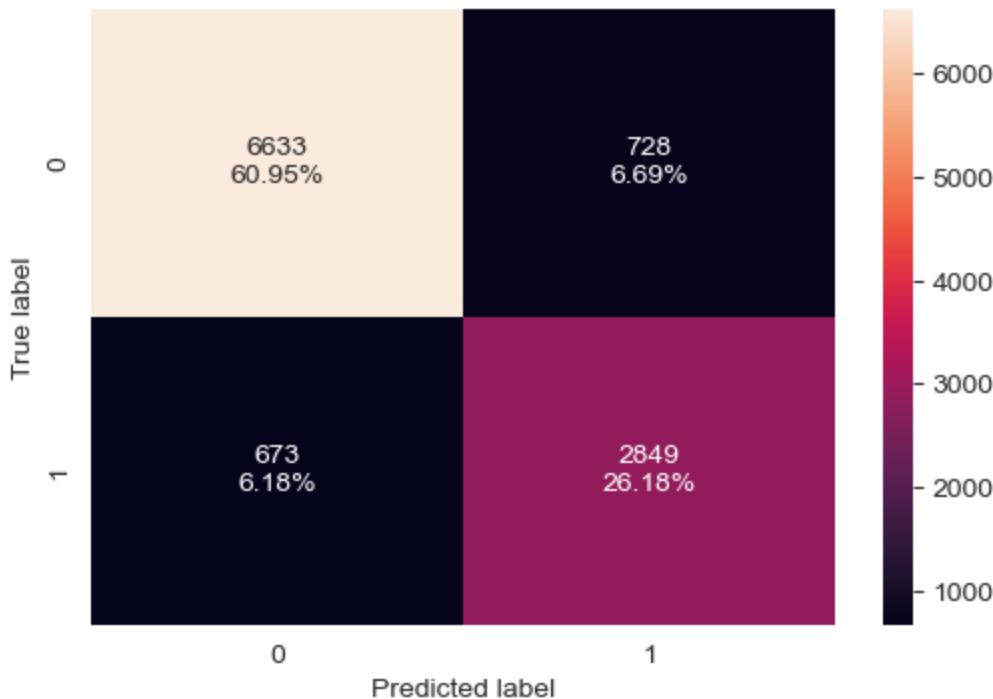


```
In [88]: decision_tree_perf_test = model_performance_classification_sklearn(  
    model, X_test, y_test  
)  
decision_tree_perf_test
```

Out[88]:

	Accuracy	Recall	Precision	F1
0	0.871267	0.808915	0.796477	0.802648

```
In [89]: confusion_matrix_sklearn(model, X_test, y_test)
```

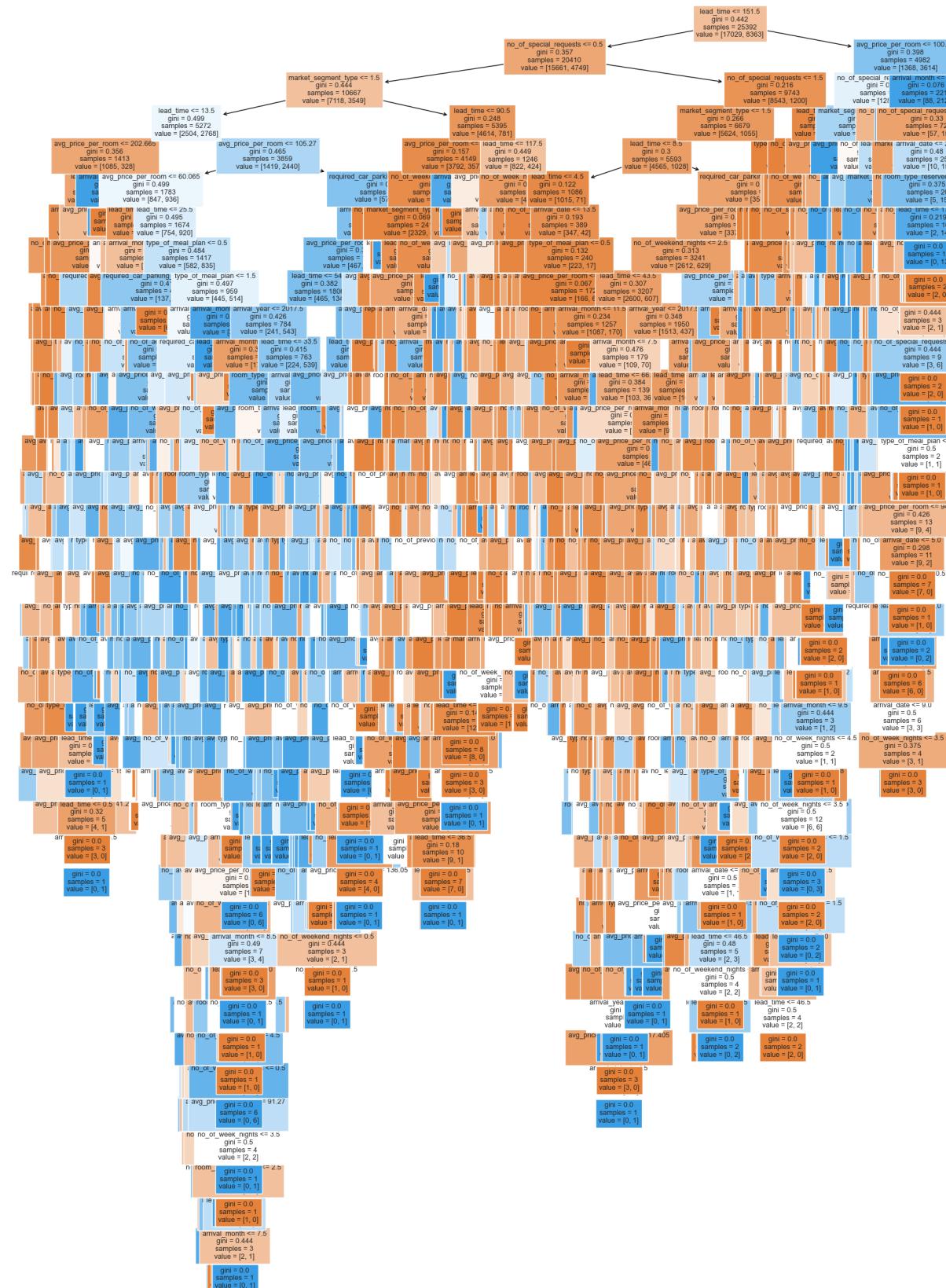


```
In [90]: column_names = list(X.columns)
feature_names = column_names
print(feature_names)

['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', 'arr
ival_year', 'arrival_month', 'arrival_date', 'market_segment_type', 'repeated_guest',
'no_of_previous_cancellations', 'no_of_previous_bookings_not_canceled', 'avg_price_pe
r_room', 'no_of_special_requests']
```

```
In [91]: names = list(X.columns)
plt.figure(figsize=(20,30))
out = tree.plot_tree(model, feature_names=feature_names, filled=True, fontsize=9, node_ids
for o in out:
    arrow = o.arrow_patch
    if arrow is not None:
        arrow.set_edgecolor('black')
        arrow.set_linewidth(1)
plt.show()
```

Project 4



```
In [92]: print(tree.export_text(model, feature_names=feature_names, show_weights=True))
```

```
--- lead_time <= 151.50
    --- no_of_special_requests <= 0.50
        --- market_segment_type <= 1.50
            --- lead_time <= 13.50
                --- avg_price_per_room <= 202.67
                    --- lead_time <= 3.50
                        --- arrival_month <= 5.50
                            --- no_of_weekend_nights <= 1.50
                                --- arrival_month <= 1.50
                                    --- weights: [56.00, 0.00] class: 0
                                --- arrival_month > 1.50
                                    --- avg_price_per_room <= 77.50
                                        --- weights: [24.00, 0.00] class: 0
                                    --- avg_price_per_room > 77.50
                                        --- arrival_date <= 26.50
                                            --- truncated branch of depth 14
                                        --- arrival_date > 26.50
                                            --- truncated branch of depth 4
                                --- no_of_weekend_nights > 1.50
                                    --- arrival_date <= 25.50
                                        --- avg_price_per_room <= 134.22
                                            --- lead_time <= 2.50
                                                --- weights: [17.00, 0.00] class: 0
                                            --- lead_time > 2.50
                                                --- truncated branch of depth 3
                                        --- avg_price_per_room > 134.22
                                            --- weights: [0.00, 2.00] class: 1
                                    --- arrival_date > 25.50
                                        --- arrival_month <= 1.50
                                            --- weights: [1.00, 0.00] class: 0
                                        --- arrival_month > 1.50
                                            --- weights: [0.00, 14.00] class: 1
                --- arrival_month > 5.50
                    --- no_of_week_nights <= 8.50
                        --- arrival_month <= 9.50
                            --- avg_price_per_room <= 76.35
                                --- avg_price_per_room <= 74.40
                                    --- truncated branch of depth 4
                                --- avg_price_per_room > 74.40
                                    --- truncated branch of depth 2
                            --- avg_price_per_room > 76.35
                                --- avg_price_per_room <= 118.04
                                    --- truncated branch of depth 9
                                --- avg_price_per_room > 118.04
                                    --- truncated branch of depth 16
                        --- arrival_month > 9.50
                            --- avg_price_per_room <= 178.00
                                --- lead_time <= 1.50
                                    --- truncated branch of depth 4
                                --- lead_time > 1.50
                                    --- truncated branch of depth 5
                            --- avg_price_per_room > 178.00
                                --- lead_time <= 0.50
                                    --- weights: [0.00, 1.00] class: 1
                                --- lead_time > 0.50
                                    --- weights: [2.00, 0.00] class: 0
                            --- no_of_week_nights > 8.50
                                --- weights: [0.00, 1.00] class: 1
                    --- lead_time > 3.50
                        --- avg_price_per_room <= 99.38
```

Project 4

```

    |--- avg_price_per_room <= 78.90
    |   |--- no_of_weekend_nights <= 4.50
    |   |   |--- no_of_week_nights <= 5.50
    |   |   |   |--- arrival_date <= 23.50
    |   |   |   |   |--- weights: [100.00, 0.00] class: 0
    |   |   |   |   |--- arrival_date > 23.50
    |   |   |   |   |   |--- truncated branch of depth 3
    |   |   |   |--- no_of_week_nights > 5.50
    |   |   |   |   |--- type_of_meal_plan <= 0.50
    |   |   |   |   |   |--- weights: [0.00, 1.00] class: 1
    |   |   |   |   |--- type_of_meal_plan > 0.50
    |   |   |   |   |   |--- weights: [4.00, 0.00] class: 0
    |   |   |--- no_of_weekend_nights > 4.50
    |   |   |   |--- weights: [0.00, 1.00] class: 1
    |--- avg_price_per_room > 78.90
    |--- arrival_month <= 11.50
    |   |--- arrival_month <= 1.50
    |   |   |--- weights: [23.00, 0.00] class: 0
    |   |--- arrival_month > 1.50
    |   |   |--- no_of_week_nights <= 4.50
    |   |   |   |--- truncated branch of depth 11
    |   |   |--- no_of_week_nights > 4.50
    |   |   |   |   |--- truncated branch of depth 2
    |   |--- arrival_month > 11.50
    |   |   |--- weights: [42.00, 0.00] class: 0
--- avg_price_per_room > 99.38
    |--- arrival_month <= 8.50
    |   |--- required_car_parking_space <= 0.50
    |   |   |--- avg_price_per_room <= 119.25
    |   |   |   |--- avg_price_per_room <= 117.25
    |   |   |   |   |--- truncated branch of depth 8
    |   |   |   |   |--- avg_price_per_room > 117.25
    |   |   |   |   |   |--- truncated branch of depth 3
    |   |--- avg_price_per_room > 119.25
    |   |   |--- no_of_weekend_nights <= 1.50
    |   |   |   |--- truncated branch of depth 12
    |   |   |--- no_of_weekend_nights > 1.50
    |   |   |   |   |--- truncated branch of depth 6
    |--- required_car_parking_space > 0.50
    |   |--- weights: [5.00, 0.00] class: 0
--- arrival_month > 8.50
    |--- lead_time <= 9.50
    |   |--- arrival_date <= 6.50
    |   |   |--- arrival_date <= 4.50
    |   |   |   |--- truncated branch of depth 3
    |   |   |--- arrival_date > 4.50
    |   |   |   |   |--- truncated branch of depth 2
    |   |--- arrival_date > 6.50
    |   |   |--- arrival_date <= 18.50
    |   |   |   |--- truncated branch of depth 5
    |   |   |--- arrival_date > 18.50
    |   |   |   |   |--- weights: [34.00, 0.00] class: 0
--- lead_time > 9.50
    |--- arrival_month <= 11.50
    |   |--- arrival_date <= 26.00
    |   |   |--- truncated branch of depth 8
    |   |   |--- arrival_date > 26.00
    |   |   |   |--- weights: [5.00, 0.00] class: 0
--- arrival_month > 11.50
    |--- weights: [10.00, 0.00] class: 0

```

```
    |--- avg_price_per_room > 202.67
    |--- arrival_month <= 11.50
    |   |--- weights: [0.00, 32.00] class: 1
    |--- arrival_month > 11.50
    |   |--- weights: [1.00, 0.00] class: 0
--- lead_time > 13.50
    --- avg_price_per_room <= 105.27
        --- avg_price_per_room <= 60.07
            --- lead_time <= 84.50
                --- lead_time <= 51.50
                    |--- lead_time <= 50.50
                        |--- avg_price_per_room <= 21.67
                            |--- weights: [19.00, 0.00] class: 0
                        |--- avg_price_per_room > 21.67
                            |--- avg_price_per_room <= 49.84
                                |--- truncated branch of depth 2
                            |--- avg_price_per_room > 49.84
                                |--- truncated branch of depth 4
                        |--- lead_time > 50.50
                            |--- weights: [0.00, 1.00] class: 1
                    |--- lead_time > 51.50
                        |--- weights: [32.00, 0.00] class: 0
                |--- lead_time > 84.50
                    --- arrival_year <= 2017.50
                        --- arrival_date <= 19.00
                            |--- lead_time <= 139.00
                                |--- weights: [0.00, 8.00] class: 1
                            |--- lead_time > 139.00
                                |--- weights: [1.00, 0.00] class: 0
                        |--- arrival_date > 19.00
                            |--- lead_time <= 87.50
                                |--- weights: [0.00, 1.00] class: 1
                            |--- lead_time > 87.50
                                |--- no_of_weekend_nights <= 0.50
                                    |--- truncated branch of depth 2
                                |--- no_of_weekend_nights > 0.50
                                    |--- weights: [6.00, 0.00] class: 0
                    |--- arrival_year > 2017.50
                        |--- avg_price_per_room <= 59.43
                            |--- weights: [14.00, 0.00] class: 0
                        |--- avg_price_per_room > 59.43
                            |--- weights: [0.00, 1.00] class: 1
--- avg_price_per_room > 60.07
    --- lead_time <= 25.50
        --- arrival_month <= 11.50
            |--- arrival_month <= 1.50
                |--- weights: [29.00, 0.00] class: 0
            |--- arrival_month > 1.50
                |--- arrival_year <= 2017.50
                    |--- lead_time <= 14.50
                        |--- weights: [0.00, 1.00] class: 1
                    |--- lead_time > 14.50
                        |--- truncated branch of depth 3
                |--- arrival_year > 2017.50
                    |--- no_of_week_nights <= 3.50
                        |--- truncated branch of depth 14
                    |--- no_of_week_nights > 3.50
                        |--- truncated branch of depth 4
            |--- arrival_month > 11.50
                |--- weights: [54.00, 0.00] class: 0
```

```

    |--- lead_time >  25.50
    |--- type_of_meal_plan <= 0.50
    |   |--- required_car_parking_space <= 0.50
    |   |   |--- arrival_year <= 2017.50
    |   |   |   |--- lead_time <= 57.00
    |   |   |   |   |--- weights: [6.00, 0.00] class: 0
    |   |   |   |--- lead_time > 57.00
    |   |   |   |   |--- weights: [0.00, 1.00] class: 1
    |   |--- arrival_year > 2017.50
    |   |   |--- no_of_adults <= 1.50
    |   |   |   |--- truncated branch of depth 5
    |   |   |   |--- no_of_adults > 1.50
    |   |   |   |   |--- truncated branch of depth 15
    |--- required_car_parking_space > 0.50
    |   |--- weights: [6.00, 0.00] class: 0
    |--- type_of_meal_plan > 0.50
    |   |--- type_of_meal_plan <= 1.50
    |   |   |--- arrival_year <= 2017.50
    |   |   |   |--- lead_time <= 60.50
    |   |   |   |   |--- truncated branch of depth 6
    |   |   |   |--- lead_time > 60.50
    |   |   |   |   |--- truncated branch of depth 10
    |   |--- arrival_year > 2017.50
    |   |   |--- required_car_parking_space <= 0.50
    |   |   |   |--- truncated branch of depth 28
    |   |   |   |--- required_car_parking_space > 0.50
    |   |   |   |   |--- weights: [12.00, 0.00] class: 0
    |--- type_of_meal_plan > 1.50
    |   |--- arrival_month <= 5.00
    |   |   |--- weights: [2.00, 0.00] class: 0
    |   |--- arrival_month > 5.00
    |   |   |--- lead_time <= 57.50
    |   |   |   |--- truncated branch of depth 2
    |   |   |   |--- lead_time > 57.50
    |   |   |   |   |--- weights: [0.00, 35.00] class: 1
--- avg_price_per_room > 105.27
|--- required_car_parking_space <= 0.50
|--- arrival_month <= 10.50
|   |--- avg_price_per_room <= 195.30
|   |--- lead_time <= 54.50
|   |   |--- arrival_year <= 2017.50
|   |   |   |--- arrival_month <= 9.50
|   |   |   |   |--- truncated branch of depth 4
|   |   |   |--- arrival_month > 9.50
|   |   |   |   |--- weights: [11.00, 0.00] class: 0
|--- arrival_year > 2017.50
|   |--- lead_time <= 33.50
|   |   |--- truncated branch of depth 16
|   |--- lead_time > 33.50
|   |   |--- truncated branch of depth 15
|--- lead_time > 54.50
|--- arrival_month <= 8.50
|   |--- lead_time <= 135.50
|   |   |--- truncated branch of depth 20
|   |--- lead_time > 135.50
|   |   |--- truncated branch of depth 9
|--- arrival_month > 8.50
|   |--- lead_time <= 59.50
|   |   |--- truncated branch of depth 6
|   |--- lead_time > 59.50

```

Project 4

```

| | | | --- truncated branch of depth 12
|--- avg_price_per_room > 195.30
|--- type_of_meal_plan <= 0.50
|   |--- weights: [1.00, 0.00] class: 0
|--- type_of_meal_plan > 0.50
|   |--- no_of_adults <= 1.50
|   |   |--- lead_time <= 59.50
|   |   |   |--- weights: [0.00, 6.00] class: 1
|   |   |   |--- lead_time > 59.50
|   |   |   |   |--- weights: [1.00, 0.00] class: 0
|   |   |--- no_of_adults > 1.50
|   |   |   |--- weights: [0.00, 92.00] class: 1
--- arrival_month > 10.50
|--- lead_time <= 22.50
|   |--- no_of_adults <= 1.50
|   |   |--- arrival_month <= 11.50
|   |   |   |--- weights: [0.00, 4.00] class: 1
|   |   |   |--- arrival_month > 11.50
|   |   |   |   |--- weights: [1.00, 0.00] class: 0
|   |   |--- no_of_adults > 1.50
|   |   |   |--- weights: [22.00, 0.00] class: 0
--- lead_time > 22.50
|--- avg_price_per_room <= 168.06
|   |--- avg_price_per_room <= 147.75
|   |   |--- no_of_week_nights <= 3.50
|   |   |   |--- truncated branch of depth 8
|   |   |   |--- no_of_week_nights > 3.50
|   |   |   |   |--- truncated branch of depth 4
|   |   |   |--- avg_price_per_room > 147.75
|   |   |   |   |--- weights: [0.00, 15.00] class: 1
|--- avg_price_per_room > 168.06
|   |--- lead_time <= 80.00
|   |   |--- no_of_week_nights <= 6.50
|   |   |   |--- truncated branch of depth 3
|   |   |   |--- no_of_week_nights > 6.50
|   |   |   |   |--- weights: [0.00, 1.00] class: 1
|   |--- lead_time > 80.00
|   |   |--- weights: [0.00, 1.00] class: 1
--- required_car_parking_space > 0.50
|--- no_of_week_nights <= 8.00
|   |--- weights: [39.00, 0.00] class: 0
|--- no_of_week_nights > 8.00
|   |--- weights: [0.00, 1.00] class: 1
segment_type > 1.50
d_time <= 90.50
avg_price_per_room <= 201.50
|--- no_of_weekend_nights <= 0.50
|   |--- market_segment_type <= 2.50
|   |   |--- weights: [1609.00, 0.00] class: 0
|--- market_segment_type > 2.50
|   |--- lead_time <= 16.50
|   |   |--- avg_price_per_room <= 68.50
|   |   |   |--- repeated_guest <= 1.00
|   |   |   |   |--- avg_price_per_room <= 48.50
|   |   |   |   |   |--- weights: [63.00, 0.00] class: 0
|   |   |   |   |   |--- avg_price_per_room > 48.50
|   |   |   |   |   |--- truncated branch of depth 8
|   |   |   |   |--- repeated_guest > 1.00
|   |   |   |   |   |--- weights: [109.00, 0.00] class: 0
|--- avg_price_per_room > 68.50

```

Project 4

```

--- arrival_date <= 29.50
    |--- no_of_adults <= 1.50
        |--- truncated branch of depth 15
    |--- no_of_adults > 1.50
        |--- truncated branch of depth 6
--- arrival_date > 29.50
    |--- arrival_date <= 30.50
        |--- weights: [0.00, 4.00] class: 1
    |--- arrival_date > 30.50
        |--- truncated branch of depth 2
--- lead_time > 16.50
    |--- avg_price_per_room <= 135.00
        |--- arrival_month <= 11.50
            |--- lead_time <= 17.50
                |--- truncated branch of depth 3
            |--- lead_time > 17.50
                |--- truncated branch of depth 12
        |--- arrival_month > 11.50
            |--- weights: [29.00, 0.00] class: 0
    |--- avg_price_per_room > 135.00
        |--- weights: [0.00, 8.00] class: 1
--- no_of_weekend_nights > 0.50
    |--- lead_time <= 68.50
        |--- no_of_weekend_nights <= 4.50
            |--- lead_time <= 1.50
                |--- arrival_date <= 27.50
                    |--- no_of_week_nights <= 5.50
                        |--- truncated branch of depth 7
                    |--- no_of_week_nights > 5.50
                        |--- truncated branch of depth 2
                |--- arrival_date > 27.50
                    |--- arrival_month <= 2.50
                        |--- truncated branch of depth 2
                    |--- arrival_month > 2.50
                        |--- truncated branch of depth 5
            |--- lead_time > 1.50
                |--- arrival_month <= 9.50
                    |--- lead_time <= 59.50
                        |--- truncated branch of depth 17
                    |--- lead_time > 59.50
                        |--- truncated branch of depth 5
                |--- arrival_month > 9.50
                    |--- market_segment_type <= 4.50
                        |--- truncated branch of depth 14
                    |--- market_segment_type > 4.50
                        |--- truncated branch of depth 3
        |--- no_of_weekend_nights > 4.50
            |--- weights: [0.00, 8.00] class: 1
--- lead_time > 68.50
    |--- avg_price_per_room <= 99.98
        |--- arrival_month <= 3.50
            |--- avg_price_per_room <= 62.50
                |--- weights: [21.00, 0.00] class: 0
            |--- avg_price_per_room > 62.50
                |--- lead_time <= 77.00
                    |--- truncated branch of depth 2
                |--- lead_time > 77.00
                    |--- truncated branch of depth 5
        |--- arrival_month > 3.50
            |--- lead_time <= 71.50

```

Project 4

```

|--- no_of_week_nights <= 3.50
|   |--- weights: [3.00, 0.00] class: 0
|--- no_of_week_nights > 3.50
|   |--- weights: [0.00, 2.00] class: 1
|--- lead_time > 71.50
|   |--- no_of_week_nights <= 2.50
|       |--- truncated branch of depth 6
|       |--- no_of_week_nights > 2.50
|           |--- truncated branch of depth 7
|--- avg_price_per_room > 99.98
|   |--- no_of_adults <= 1.50
|       |--- avg_price_per_room <= 123.25
|           |--- weights: [0.00, 52.00] class: 1
|       |--- avg_price_per_room > 123.25
|           |--- weights: [1.00, 0.00] class: 0
|   |--- no_of_adults > 1.50
|       |--- arrival_date <= 23.50
|           |--- arrival_date <= 3.00
|               |--- weights: [0.00, 2.00] class: 1
|           |--- arrival_date > 3.00
|               |--- truncated branch of depth 6
|       |--- arrival_date > 23.50
|           |--- avg_price_per_room <= 131.67
|               |--- truncated branch of depth 3
|           |--- avg_price_per_room > 131.67
|               |--- weights: [2.00, 0.00] class: 0
|--- avg_price_per_room > 201.50
|   |--- arrival_date <= 28.00
|       |--- weights: [0.00, 17.00] class: 1
|   |--- arrival_date > 28.00
|       |--- weights: [2.00, 0.00] class: 0
--- lead_time > 90.50
--- lead_time <= 117.50
    |--- avg_price_per_room <= 93.58
        |--- arrival_date <= 6.50
            |--- no_of_week_nights <= 2.50
                |--- no_of_weekend_nights <= 1.50
                    |--- avg_price_per_room <= 80.38
                        |--- type_of_meal_plan <= 0.50
                            |--- weights: [1.00, 0.00] class: 0
                        |--- type_of_meal_plan > 0.50
                            |--- truncated branch of depth 4
                    |--- avg_price_per_room > 80.38
                        |--- room_type_reserved <= 2.50
                            |--- weights: [3.00, 0.00] class: 0
                        |--- room_type_reserved > 2.50
                            |--- weights: [0.00, 1.00] class: 1
                |--- no_of_weekend_nights > 1.50
                    |--- arrival_date <= 5.00
                        |--- weights: [5.00, 0.00] class: 0
                    |--- arrival_date > 5.00
                        |--- weights: [0.00, 1.00] class: 1
            |--- no_of_week_nights > 2.50
                |--- arrival_date <= 5.50
                    |--- weights: [35.00, 0.00] class: 0
            |--- arrival_date > 5.50
                |--- arrival_month <= 9.00
                    |--- weights: [3.00, 0.00] class: 0
            |--- arrival_month > 9.00
                |--- no_of_week_nights <= 3.50

```

Project 4

```

|--- weights: [1.00, 0.00] class: 0
|--- no_of_week_nights > 3.50
|--- weights: [0.00, 2.00] class: 1
--- arrival_date > 6.50
|--- avg_price_per_room <= 66.50
|--- no_of_weekend_nights <= 1.50
|--- arrival_date <= 9.50
|   |--- no_of_week_nights <= 3.00
|   |--- weights: [0.00, 1.00] class: 1
|   |--- no_of_week_nights > 3.00
|   |--- weights: [2.00, 0.00] class: 0
|--- arrival_date > 9.50
|   |--- weights: [24.00, 0.00] class: 0
|--- no_of_weekend_nights > 1.50
|   |--- avg_price_per_room <= 58.75
|   |--- weights: [6.00, 0.00] class: 0
|--- avg_price_per_room > 58.75
|   |--- lead_time <= 97.50
|       |--- truncated branch of depth 2
|   |--- lead_time > 97.50
|       |--- weights: [0.00, 39.00] class: 1
|--- avg_price_per_room > 66.50
|--- type_of_meal_plan <= 1.50
|--- arrival_date <= 29.50
|   |--- type_of_meal_plan <= 0.50
|   |--- weights: [0.00, 1.00] class: 1
|   |--- type_of_meal_plan > 0.50
|       |--- truncated branch of depth 11
|--- arrival_date > 29.50
|   |--- lead_time <= 96.00
|   |--- weights: [8.00, 0.00] class: 0
|   |--- lead_time > 96.00
|       |--- truncated branch of depth 4
|--- type_of_meal_plan > 1.50
|   |--- avg_price_per_room <= 82.50
|   |--- avg_price_per_room <= 76.25
|   |--- weights: [1.00, 0.00] class: 0
|   |--- avg_price_per_room > 76.25
|   |--- weights: [0.00, 7.00] class: 1
|--- avg_price_per_room > 82.50
|   |--- no_of_adults <= 1.50
|   |--- weights: [11.00, 2.00] class: 0
|   |--- no_of_adults > 1.50
|       |--- weights: [1.00, 0.00] class: 0
--- avg_price_per_room > 93.58
|--- arrival_date <= 16.50
|--- arrival_month <= 7.50
|--- lead_time <= 108.50
|   |--- avg_price_per_room <= 125.00
|   |--- no_of_week_nights <= 0.50
|   |--- weights: [0.00, 1.00] class: 1
|   |--- no_of_week_nights > 0.50
|       |--- truncated branch of depth 4
|   |--- avg_price_per_room > 125.00
|       |--- weights: [0.00, 1.00] class: 1
|--- lead_time > 108.50
|--- lead_time <= 111.50
|   |--- weights: [5.00, 0.00] class: 0
|--- lead_time > 111.50
|   |--- arrival_date <= 3.50

```

Project 4

```

|--- weights: [12.00, 1.00] class: 0
|--- arrival_date > 3.50
|--- weights: [1.00, 0.00] class: 0
--- arrival_month > 7.50
    --- avg_price_per_room <= 108.50
        --- arrival_date <= 14.50
            --- lead_time <= 113.50
                --- weights: [4.00, 0.00] class: 0
                --- lead_time > 113.50
                    --- truncated branch of depth 2
                --- arrival_date > 14.50
                    --- weights: [0.00, 47.00] class: 1
            --- avg_price_per_room > 108.50
                --- no_of_weekend_nights <= 0.50
                    --- weights: [42.00, 0.00] class: 0
                --- no_of_weekend_nights > 0.50
                    --- arrival_date <= 9.50
                        --- truncated branch of depth 4
                    --- arrival_date > 9.50
                        --- truncated branch of depth 3
            --- arrival_date > 16.50
                --- arrival_month <= 8.50
                    --- avg_price_per_room <= 127.39
                        --- no_of_adults <= 1.50
                            --- weights: [0.00, 50.00] class: 1
                        --- no_of_adults > 1.50
                            --- no_of_week_nights <= 2.50
                                --- truncated branch of depth 4
                            --- no_of_week_nights > 2.50
                                --- truncated branch of depth 4
                    --- avg_price_per_room > 127.39
                        --- weights: [2.00, 0.00] class: 0
            --- arrival_month > 8.50
                --- arrival_year <= 2017.50
                    --- weights: [0.00, 3.00] class: 1
                --- arrival_year > 2017.50
                    --- avg_price_per_room <= 101.34
                        --- weights: [0.00, 3.00] class: 1
                    --- avg_price_per_room > 101.34
                        --- avg_price_per_room <= 177.83
                            --- weights: [8.00, 0.00] class: 0
                        --- avg_price_per_room > 177.83
                            --- weights: [0.00, 1.00] class: 1
            --- lead_time > 117.50
                --- no_of_week_nights <= 1.50
                    --- arrival_date <= 7.50
                        --- weights: [51.00, 0.00] class: 0
                --- arrival_date > 7.50
                    --- avg_price_per_room <= 93.58
                        --- avg_price_per_room <= 65.38
                            --- weights: [0.00, 3.00] class: 1
                        --- avg_price_per_room > 65.38
                            --- avg_price_per_room <= 89.88
                                --- weights: [24.00, 0.00] class: 0
                            --- avg_price_per_room > 89.88
                                --- avg_price_per_room <= 91.33
                                    --- weights: [8.00, 2.00] class: 0
                                --- avg_price_per_room > 91.33
                                    --- weights: [1.00, 0.00] class: 0
                    --- avg_price_per_room > 93.58

```

Project 4

```

--- arrival_date <= 28.00
    --- type_of_meal_plan <= 1.50
        --- no_of_adults <= 2.50
            --- weights: [0.00, 17.00] class: 1
            --- no_of_adults > 2.50
                --- weights: [1.00, 1.00] class: 0
        --- type_of_meal_plan > 1.50
            --- arrival_year <= 2017.50
                --- truncated branch of depth 3
            --- arrival_year > 2017.50
                --- weights: [1.00, 0.00] class: 0
    --- arrival_date > 28.00
        --- weights: [13.00, 1.00] class: 0
--- no_of_week_nights > 1.50
    --- no_of_adults <= 1.50
        --- weights: [113.00, 0.00] class: 0
    --- no_of_adults > 1.50
        --- lead_time <= 125.50
            --- avg_price_per_room <= 90.85
                --- avg_price_per_room <= 87.50
                    --- no_of_week_nights <= 3.50
                        --- truncated branch of depth 4
                    --- no_of_week_nights > 3.50
                        --- truncated branch of depth 2
                --- avg_price_per_room > 87.50
                    --- weights: [0.00, 10.00] class: 1
            --- avg_price_per_room > 90.85
                --- weights: [14.00, 0.00] class: 0
        --- lead_time > 125.50
            --- avg_price_per_room <= 216.00
                --- arrival_date <= 19.50
                    --- arrival_date <= 10.50
                        --- truncated branch of depth 5
                    --- arrival_date > 10.50
                        --- truncated branch of depth 9
                --- arrival_date > 19.50
                    --- lead_time <= 128.00
                        --- truncated branch of depth 2
                    --- lead_time > 128.00
                        --- weights: [75.00, 0.00] class: 0
            --- avg_price_per_room > 216.00
                --- weights: [0.00, 1.00] class: 1
--- no_of_special_requests > 0.50
    --- no_of_special_requests <= 1.50
        --- market_segment_type <= 1.50
            --- lead_time <= 8.50
                --- lead_time <= 4.50
                    --- no_of_week_nights <= 10.00
                        --- avg_price_per_room <= 219.86
                            --- avg_price_per_room <= 157.64
                                --- arrival_date <= 4.50
                                    --- weights: [82.00, 0.00] class: 0
                                --- arrival_date > 4.50
                                    --- arrival_date <= 27.50
                                        --- truncated branch of depth 11
                                    --- arrival_date > 27.50
                                        --- weights: [69.00, 0.00] class: 0
                            --- avg_price_per_room > 157.64
                                --- avg_price_per_room <= 158.50
                                    --- weights: [0.00, 1.00] class: 1

```

Project 4

```

--- avg_price_per_room > 158.50
|--- no_of_week_nights <= 3.50
|   |--- truncated branch of depth 6
|--- no_of_week_nights > 3.50
|   |--- truncated branch of depth 3
--- avg_price_per_room > 219.86
|--- avg_price_per_room <= 223.58
|   |--- weights: [0.00, 1.00] class: 1
--- avg_price_per_room > 223.58
|--- avg_price_per_room <= 237.25
|   |--- weights: [5.00, 0.00] class: 0
--- avg_price_per_room > 237.25
|--- arrival_date <= 11.50
|   |--- weights: [3.00, 0.00] class: 0
|--- arrival_date > 11.50
|   |--- truncated branch of depth 2
--- no_of_week_nights > 10.00
|--- weights: [0.00, 2.00] class: 1
lead_time > 4.50
|--- arrival_date <= 13.50
|   |--- arrival_month <= 9.50
|   |   |--- arrival_year <= 2017.50
|   |   |--- arrival_date <= 11.00
|   |       |--- avg_price_per_room <= 119.50
|   |       |--- weights: [4.00, 0.00] class: 0
|   |       |--- avg_price_per_room > 119.50
|   |       |   |--- truncated branch of depth 2
|   |       |--- arrival_date > 11.00
|   |       |   |--- weights: [0.00, 3.00] class: 1
|   |--- arrival_year > 2017.50
|   |   |--- avg_price_per_room <= 185.50
|   |   |   |--- no_of_adults <= 0.50
|   |   |       |--- weights: [0.00, 1.00] class: 1
|   |   |   |--- no_of_adults > 0.50
|   |   |       |--- truncated branch of depth 8
|   |   |--- avg_price_per_room > 185.50
|   |   |   |--- arrival_month <= 8.50
|   |   |       |--- weights: [0.00, 2.00] class: 1
|   |   |   |--- arrival_month > 8.50
|   |   |       |--- weights: [1.00, 0.00] class: 0
|--- arrival_month > 9.50
|   |--- avg_price_per_room <= 157.12
|   |   |--- weights: [43.00, 0.00] class: 0
|--- avg_price_per_room > 157.12
|   |--- room_type_reserved <= 2.50
|   |   |--- weights: [0.00, 1.00] class: 1
|   |   |--- room_type_reserved > 2.50
|   |       |--- weights: [2.00, 0.00] class: 0
--- arrival_date > 13.50
|--- type_of_meal_plan <= 0.50
|   |--- avg_price_per_room <= 126.33
|   |   |--- arrival_date <= 21.50
|   |       |--- weights: [24.00, 0.00] class: 0
|   |   |--- arrival_date > 21.50
|   |       |--- arrival_date <= 23.50
|   |           |--- truncated branch of depth 3
|   |           |--- arrival_date > 23.50
|   |               |--- weights: [17.00, 0.00] class: 0
|--- avg_price_per_room > 126.33
|   |--- arrival_month <= 8.50

```

Project 4

```

|--- arrival_date <= 26.00
|   |--- weights: [7.00, 0.00] class: 0
|--- arrival_date > 26.00
|   |--- truncated branch of depth 2
|--- arrival_month > 8.50
|   |--- lead_time <= 6.50
|       |--- truncated branch of depth 2
|   |--- lead_time > 6.50
|       |--- truncated branch of depth 3
|--- type_of_meal_plan > 0.50
|   |--- avg_price_per_room <= 139.57
|       |--- no_of_week_nights <= 3.50
|           |--- arrival_month <= 2.50
|               |--- truncated branch of depth 3
|           |--- arrival_month > 2.50
|               |--- weights: [97.00, 0.00] class: 0
|   |--- no_of_week_nights > 3.50
|       |--- arrival_date <= 23.50
|           |--- weights: [10.00, 0.00] class: 0
|       |--- arrival_date > 23.50
|           |--- truncated branch of depth 2
|   |--- avg_price_per_room > 139.57
|       |--- arrival_date <= 15.50
|           |--- arrival_date <= 14.50
|               |--- weights: [2.00, 0.00] class: 0
|           |--- arrival_date > 14.50
|               |--- truncated branch of depth 3
|       |--- arrival_date > 15.50
|           |--- avg_price_per_room <= 140.50
|               |--- truncated branch of depth 2
|           |--- avg_price_per_room > 140.50
|               |--- truncated branch of depth 3
|--- lead_time > 8.50
|   |--- required_car_parking_space <= 0.50
|       |--- avg_price_per_room <= 127.62
|           |--- no_of_weekend_nights <= 2.50
|               |--- lead_time <= 43.50
|                   |--- arrival_month <= 11.50
|                       |--- arrival_month <= 1.50
|                           |--- weights: [87.00, 0.00] class: 0
|                       |--- arrival_month > 1.50
|                           |--- truncated branch of depth 23
|                   |--- arrival_month > 11.50
|                       |--- weights: [128.00, 0.00] class: 0
|   |--- lead_time > 43.50
|       |--- arrival_year <= 2017.50
|           |--- arrival_month <= 7.50
|               |--- truncated branch of depth 4
|           |--- arrival_month > 7.50
|               |--- truncated branch of depth 10
|       |--- arrival_year > 2017.50
|           |--- arrival_month <= 8.50
|               |--- truncated branch of depth 20
|           |--- arrival_month > 8.50
|               |--- truncated branch of depth 21
|   |--- no_of_weekend_nights > 2.50
|       |--- avg_price_per_room <= 119.12
|           |--- arrival_month <= 1.50
|               |--- weights: [3.00, 0.00] class: 0
|           |--- arrival_month > 1.50

```

Project 4

```
|--- no_of_week_nights <= 8.50
|   |--- truncated branch of depth 4
|--- no_of_week_nights > 8.50
|   |--- weights: [0.00, 12.00] class: 1
|--- avg_price_per_room > 119.12
|   |--- weights: [3.00, 0.00] class: 0
--- avg_price_per_room > 127.62
|--- lead_time <= 142.50
|   |--- arrival_month <= 8.50
|       |--- arrival_date <= 19.50
|           |--- avg_price_per_room <= 177.15
|               |--- truncated branch of depth 15
|           |--- avg_price_per_room > 177.15
|               |--- truncated branch of depth 6
|           |--- arrival_date > 19.50
|               |--- arrival_date <= 27.50
|                   |--- truncated branch of depth 11
|               |--- arrival_date > 27.50
|                   |--- truncated branch of depth 10
|           |--- arrival_month > 8.50
|               |--- arrival_month <= 11.50
|                   |--- arrival_year <= 2017.50
|                       |--- truncated branch of depth 7
|                   |--- arrival_year > 2017.50
|                       |--- truncated branch of depth 21
|               |--- arrival_month > 11.50
|                   |--- lead_time <= 100.50
|                       |--- weights: [49.00, 0.00] class: 0
|                   |--- lead_time > 100.50
|                       |--- truncated branch of depth 2
|--- lead_time > 142.50
|   |--- avg_price_per_room <= 142.65
|       |--- no_of_week_nights <= 2.50
|           |--- weights: [5.00, 0.00] class: 0
|       |--- no_of_week_nights > 2.50
|           |--- no_of_week_nights <= 3.50
|               |--- weights: [0.00, 2.00] class: 1
|           |--- no_of_week_nights > 3.50
|               |--- weights: [1.00, 0.00] class: 0
|       |--- avg_price_per_room > 142.65
|           |--- avg_price_per_room <= 182.49
|               |--- weights: [0.00, 11.00] class: 1
|           |--- avg_price_per_room > 182.49
|               |--- arrival_month <= 7.50
|                   |--- weights: [0.00, 1.00] class: 1
|               |--- arrival_month > 7.50
|                   |--- weights: [1.00, 0.00] class: 0
--- required_car_parking_space > 0.50
|--- no_of_week_nights <= 7.50
|   |--- weights: [180.00, 0.00] class: 0
|--- no_of_week_nights > 7.50
|   |--- weights: [0.00, 1.00] class: 1
--- market_segment_type > 1.50
|--- type_of_meal_plan <= 0.50
|   |--- lead_time <= 63.00
|       |--- market_segment_type <= 2.50
|           |--- weights: [16.00, 0.00] class: 0
|       |--- market_segment_type > 2.50
|           |--- arrival_date <= 14.50
|               |--- weights: [3.00, 0.00] class: 0
```

```
    |--- arrival_date > 14.50
    |   |--- weights: [2.00, 1.00] class: 0
    --- lead_time > 63.00
    |--- weights: [0.00, 6.00] class: 1
    --- type_of_meal_plan > 0.50
    --- lead_time <= 102.50
        --- no_of_week_nights <= 11.00
            --- avg_price_per_room <= 141.50
                --- lead_time <= 91.50
                    --- avg_price_per_room <= 129.50
                        |--- weights: [856.00, 0.00] class: 0
                        --- avg_price_per_room > 129.50
                            --- avg_price_per_room <= 131.50
                                |--- truncated branch of depth 2
                                --- avg_price_per_room > 131.50
                                    |--- weights: [15.00, 0.00] class: 0
        --- lead_time > 91.50
            --- no_of_children <= 0.50
                |--- weights: [43.00, 0.00] class: 0
            --- no_of_children > 0.50
                --- arrival_date <= 16.50
                    |--- weights: [0.00, 2.00] class: 1
                --- arrival_date > 16.50
                    |--- weights: [2.00, 0.00] class: 0
        --- avg_price_per_room > 141.50
            --- avg_price_per_room <= 144.00
                |--- weights: [0.00, 1.00] class: 1
            --- avg_price_per_room > 144.00
                --- market_segment_type <= 2.50
                    |--- weights: [15.00, 0.00] class: 0
                --- market_segment_type > 2.50
                    --- arrival_date <= 15.50
                        |--- weights: [2.00, 0.00] class: 0
                    --- arrival_date > 15.50
                        |--- weights: [0.00, 1.00] class: 1
    --- no_of_week_nights > 11.00
        |--- weights: [0.00, 1.00] class: 1
    lead_time > 102.50
        --- lead_time <= 104.50
            --- no_of_week_nights <= 2.50
                --- avg_price_per_room <= 67.65
                    |--- weights: [1.00, 0.00] class: 0
                --- avg_price_per_room > 67.65
                    |--- weights: [0.00, 4.00] class: 1
            --- no_of_week_nights > 2.50
                |--- weights: [4.00, 0.00] class: 0
    lead_time > 104.50
        --- avg_price_per_room <= 141.75
            --- no_of_week_nights <= 2.50
                --- avg_price_per_room <= 83.39
                    --- arrival_month <= 3.50
                        |--- weights: [6.00, 0.00] class: 0
                    --- arrival_month > 3.50
                        |--- truncated branch of depth 4
                --- avg_price_per_room > 83.39
                    --- lead_time <= 143.50
                        |--- truncated branch of depth 3
                    --- lead_time > 143.50
                        |--- truncated branch of depth 2
            --- no_of_week_nights > 2.50
```

Project 4

```

    |--- avg_price_per_room <= 122.00
    |   |--- weights: [54.00, 0.00] class: 0
    |--- avg_price_per_room > 122.00
    |   |--- avg_price_per_room <= 131.75
    |       |--- weights: [0.00, 1.00] class: 1
    |       |--- avg_price_per_room > 131.75
    |           |--- weights: [1.00, 0.00] class: 0
    |--- avg_price_per_room > 141.75
    |   |--- type_of_meal_plan <= 1.50
    |       |--- weights: [0.00, 2.00] class: 1
    |       |--- type_of_meal_plan > 1.50
    |           |--- weights: [1.00, 0.00] class: 0
--- no_of_special_requests > 1.50
|--- lead_time <= 90.50
|   |--- no_of_week_nights <= 3.50
|       |--- weights: [2126.00, 0.00] class: 0
|--- no_of_week_nights > 3.50
|   |--- no_of_week_nights <= 9.50
|       |--- no_of_special_requests <= 2.50
|           |--- lead_time <= 6.50
|               |--- weights: [43.00, 0.00] class: 0
|           |--- lead_time > 6.50
|               |--- arrival_month <= 11.50
|                   |--- avg_price_per_room <= 93.09
|                       |--- lead_time <= 54.00
|                           |--- truncated branch of depth 8
|                           |--- lead_time > 54.00
|                               |--- truncated branch of depth 5
|                               |--- avg_price_per_room > 93.09
|                                   |--- avg_price_per_room <= 107.29
|                                       |--- truncated branch of depth 3
|                                       |--- avg_price_per_room > 107.29
|                                           |--- truncated branch of depth 13
|                                           |--- arrival_month > 11.50
|                                               |--- weights: [26.00, 0.00] class: 0
|--- no_of_special_requests > 2.50
|   |--- weights: [70.00, 0.00] class: 0
|--- no_of_week_nights > 9.50
|   |--- weights: [0.00, 2.00] class: 1
--- lead_time > 90.50
|--- avg_price_per_room <= 202.95
|--- arrival_month <= 8.50
|   |--- arrival_year <= 2017.50
|       |--- arrival_month <= 7.50
|           |--- arrival_date <= 4.50
|               |--- weights: [1.00, 0.00] class: 0
|           |--- arrival_date > 4.50
|               |--- arrival_date <= 26.00
|                   |--- weights: [0.00, 5.00] class: 1
|               |--- arrival_date > 26.00
|                   |--- avg_price_per_room <= 91.62
|                       |--- weights: [1.00, 0.00] class: 0
|                       |--- avg_price_per_room > 91.62
|                           |--- weights: [0.00, 1.00] class: 1
|--- arrival_month > 7.50
|   |--- arrival_date <= 24.50
|       |--- lead_time <= 98.50
|           |--- arrival_date <= 10.50
|               |--- weights: [1.00, 0.00] class: 0
|               |--- arrival_date > 10.50

```

Project 4

Project 4

```
    |--- lead_time <= 341.00
    |   |--- lead_time <= 173.00
    |   |   |--- arrival_date <= 3.50
    |   |   |   |--- arrival_year <= 2017.50
    |   |   |   |   |--- no_of_week_nights <= 1.50
    |   |   |   |   |   |--- weights: [1.00, 0.00] class: 0
    |   |   |   |   |   |--- no_of_week_nights > 1.50
    |   |   |   |   |   |--- weights: [61.00, 6.00] class: 0
    |   |   |   |   |--- arrival_year > 2017.50
    |   |   |   |   |   |--- weights: [1.00, 0.00] class: 0
    |   |   |   |--- arrival_date > 3.50
    |   |   |   |--- arrival_month <= 5.00
    |   |   |   |   |--- weights: [3.00, 0.00] class: 0
    |   |   |   |   |--- arrival_month > 5.00
    |   |   |   |   |   |--- weights: [0.00, 9.00] class: 1
    |--- lead_time > 173.00
    |   |--- arrival_month <= 5.50
    |   |   |--- avg_price_per_room <= 88.00
    |   |   |   |--- weights: [9.00, 0.00] class: 0
    |   |   |   |--- avg_price_per_room > 88.00
    |   |   |   |   |--- weights: [0.00, 3.00] class: 1
    |--- arrival_month > 5.50
    |   |--- avg_price_per_room <= 98.00
    |   |   |--- avg_price_per_room <= 55.21
    |   |   |   |--- truncated branch of depth 2
    |   |   |   |--- avg_price_per_room > 55.21
    |   |   |   |   |--- truncated branch of depth 6
    |   |   |   |--- avg_price_per_room > 98.00
    |   |   |   |   |--- no_of_weekend_nights <= 0.50
    |   |   |   |   |   |--- weights: [1.00, 0.00] class: 0
    |   |   |   |   |   |--- no_of_weekend_nights > 0.50
    |   |   |   |   |   |--- weights: [0.00, 1.00] class: 1
    --- lead_time > 341.00
    |--- arrival_date <= 8.50
    |   |--- avg_price_per_room <= 88.33
    |   |   |--- weights: [0.00, 7.00] class: 1
    |--- avg_price_per_room > 88.33
    |   |--- weights: [1.00, 1.00] class: 0
    --- arrival_date > 8.50
    |--- arrival_date <= 24.50
    |   |--- avg_price_per_room <= 80.00
    |   |   |--- weights: [5.00, 0.00] class: 0
    |--- avg_price_per_room > 80.00
    |   |--- lead_time <= 381.50
    |   |   |--- truncated branch of depth 2
    |   |   |--- lead_time > 381.50
    |   |   |   |--- weights: [3.00, 2.00] class: 0
    |--- arrival_date > 24.50
    |   |--- weights: [0.00, 3.00] class: 1
    --- no_of_adults > 1.50
    |--- avg_price_per_room <= 84.58
    |--- lead_time <= 244.00
    |   |--- no_of_week_nights <= 1.50
    |   |   |--- no_of_weekend_nights <= 1.50
    |   |   |   |--- arrival_date <= 19.00
    |   |   |   |   |--- lead_time <= 166.50
    |   |   |   |   |   |--- weights: [2.00, 0.00] class: 0
    |   |   |   |   |   |--- lead_time > 166.50
    |   |   |   |   |   |--- truncated branch of depth 2
    |--- arrival_date > 19.00
```

Project 4

```
|   |   |--- weights: [3.00, 0.00] class: 0
|--- no_of_weekend_nights > 1.50
|   |--- weights: [24.00, 0.00] class: 0
|--- no_of_week_nights > 1.50
|   |--- avg_price_per_room <= 66.50
|       |--- no_of_weekend_nights <= 0.50
|           |--- avg_price_per_room <= 64.80
|               |--- weights: [2.00, 0.00] class: 0
|           |--- avg_price_per_room > 64.80
|               |--- weights: [0.00, 7.00] class: 1
|       |--- no_of_weekend_nights > 0.50
|           |--- no_of_week_nights <= 5.50
|               |--- truncated branch of depth 2
|           |--- no_of_week_nights > 5.50
|               |--- weights: [0.00, 1.00] class: 1
|       |--- avg_price_per_room > 66.50
|           |--- type_of_meal_plan <= 1.50
|               |--- avg_price_per_room <= 75.75
|                   |--- truncated branch of depth 4
|               |--- avg_price_per_room > 75.75
|                   |--- truncated branch of depth 6
|           |--- type_of_meal_plan > 1.50
|               |--- weights: [0.00, 1.00] class: 1
|--- lead_time > 244.00
|--- arrival_month <= 11.50
|   |--- arrival_year <= 2017.50
|       |--- weights: [34.00, 0.00] class: 0
|   |--- arrival_year > 2017.50
|       |--- avg_price_per_room <= 80.38
|           |--- no_of_week_nights <= 3.50
|               |--- truncated branch of depth 6
|           |--- no_of_week_nights > 3.50
|               |--- truncated branch of depth 4
|       |--- avg_price_per_room > 80.38
|           |--- weights: [11.00, 0.00] class: 0
|--- arrival_month > 11.50
|   |--- weights: [37.00, 0.00] class: 0
|--- avg_price_per_room > 84.58
|--- arrival_month <= 11.50
|   |--- no_of_weekend_nights <= 1.50
|       |--- room_type_reserved <= 2.50
|           |--- no_of_adults <= 2.50
|               |--- arrival_date <= 2.50
|                   |--- weights: [1.00, 0.00] class: 0
|               |--- arrival_date > 2.50
|                   |--- truncated branch of depth 5
|           |--- no_of_adults > 2.50
|               |--- weights: [1.00, 0.00] class: 0
|       |--- room_type_reserved > 2.50
|           |--- weights: [4.00, 0.00] class: 0
|--- no_of_weekend_nights > 1.50
|--- arrival_month <= 6.50
|   |--- weights: [0.00, 13.00] class: 1
|--- arrival_month > 6.50
|   |--- weights: [14.00, 0.00] class: 0
|--- arrival_month > 11.50
|   |--- weights: [9.00, 0.00] class: 0
--- no_of_special_requests > 0.50
--- no_of_weekend_nights <= 0.50
|--- lead_time <= 180.50
```

```

    |--- lead_time <= 159.50
    |--- arrival_month <= 8.50
    |   |--- weights: [8.00, 0.00] class: 0
    |--- arrival_month >  8.50
    |   |--- no_of_week_nights <= 2.50
    |   |   |--- weights: [0.00, 4.00] class: 1
    |   |--- no_of_week_nights >  2.50
    |   |   |--- lead_time <= 156.50
    |   |   |   |--- weights: [0.00, 1.00] class: 1
    |   |--- lead_time >  156.50
    |   |   |   |--- weights: [2.00, 0.00] class: 0
    |--- lead_time >  159.50
    |--- no_of_adults <= 0.50
    |   |--- weights: [0.00, 1.00] class: 1
    |--- no_of_adults >  0.50
    |   |--- arrival_date <= 1.50
    |   |   |--- no_of_week_nights <= 1.50
    |   |   |   |--- weights: [0.00, 2.00] class: 1
    |   |   |--- no_of_week_nights >  1.50
    |   |   |   |--- weights: [2.00, 0.00] class: 0
    |--- arrival_date >  1.50
    |   |--- weights: [48.00, 0.00] class: 0
    --- lead_time >  180.50
    |--- market_segment_type <= 1.50
    |--- no_of_special_requests <= 2.50
    |   |--- arrival_month <= 11.50
    |   |   |--- no_of_week_nights <= 0.50
    |   |   |   |--- weights: [2.00, 0.00] class: 0
    |   |   |--- no_of_week_nights >  0.50
    |   |   |   |--- weights: [0.00, 125.00] class: 1
    |--- arrival_month >  11.50
    |   |--- lead_time <= 300.50
    |   |   |--- lead_time <= 226.50
    |   |   |   |--- weights: [0.00, 3.00] class: 1
    |   |   |--- lead_time >  226.50
    |   |   |   |--- lead_time <= 272.00
    |   |   |   |   |--- weights: [6.00, 0.00] class: 0
    |   |   |   |--- lead_time >  272.00
    |   |   |   |   |--- truncated branch of depth 3
    |   |   |--- lead_time >  300.50
    |   |   |   |--- weights: [0.00, 5.00] class: 1
    |--- no_of_special_requests >  2.50
    |   |--- weights: [12.00, 0.00] class: 0
    --- market_segment_type >  1.50
    |--- no_of_adults <= 2.50
    |   |--- lead_time <= 356.00
    |   |   |--- lead_time <= 302.50
    |   |   |   |--- weights: [15.00, 0.00] class: 0
    |   |   |--- lead_time >  302.50
    |   |   |   |--- weights: [2.00, 1.00] class: 0
    |--- lead_time >  356.00
    |   |--- weights: [0.00, 1.00] class: 1
    |--- no_of_adults >  2.50
    |   |--- weights: [0.00, 2.00] class: 1
    --- no_of_weekend_nights >  0.50
    |--- market_segment_type <= 1.50
    |--- no_of_week_nights <= 9.50
    |   |--- arrival_month <= 11.50
    |   |   |--- arrival_date <= 27.50
    |   |   |   |--- avg_price_per_room <= 81.12

```

Project 4

```
    |--- lead_time <= 153.50
    |    |--- no_of_week_nights <= 2.50
    |    |    |--- weights: [1.00, 0.00] class: 0
    |    |--- no_of_week_nights > 2.50
    |    |    |--- weights: [0.00, 2.00] class: 1
    |--- lead_time > 153.50
    |    |--- lead_time <= 157.50
    |    |    |--- truncated branch of depth 2
    |    |--- lead_time > 157.50
    |    |    |--- truncated branch of depth 3
    --- avg_price_per_room > 81.12
    |--- no_of_week_nights <= 6.50
    |    |--- lead_time <= 233.00
    |    |    |--- truncated branch of depth 13
    |    |--- lead_time > 233.00
    |    |    |--- truncated branch of depth 9
    |--- no_of_week_nights > 6.50
    |    |--- lead_time <= 204.50
    |    |    |--- truncated branch of depth 2
    |    |--- lead_time > 204.50
    |    |    |--- weights: [0.00, 3.00] class: 1
    --- arrival_date > 27.50
    |--- no_of_week_nights <= 1.50
    |    |--- lead_time <= 224.50
    |    |    |--- lead_time <= 175.50
    |    |    |    |--- weights: [1.00, 0.00] class: 0
    |    |    |--- lead_time > 175.50
    |    |    |    |--- weights: [0.00, 10.00] class: 1
    |--- lead_time > 224.50
    |    |--- weights: [4.00, 0.00] class: 0
    |--- no_of_week_nights > 1.50
    |    |--- lead_time <= 269.00
    |    |    |--- lead_time <= 176.00
    |    |    |    |--- truncated branch of depth 4
    |    |    |    |--- lead_time > 176.00
    |    |    |    |--- truncated branch of depth 4
    |    |    |--- lead_time > 269.00
    |    |    |--- weights: [0.00, 3.00] class: 1
    --- arrival_month > 11.50
    |--- arrival_date <= 14.50
    |    |--- arrival_date <= 3.00
    |    |    |--- weights: [0.00, 1.00] class: 1
    |--- arrival_date > 3.00
    |    |--- lead_time <= 217.50
    |    |    |--- weights: [8.00, 0.00] class: 0
    |--- lead_time > 217.50
    |    |--- no_of_weekend_nights <= 1.50
    |    |    |--- weights: [3.00, 0.00] class: 0
    |--- no_of_weekend_nights > 1.50
    |    |    |--- weights: [0.00, 1.00] class: 1
    --- arrival_date > 14.50
    |--- no_of_week_nights <= 2.50
    |    |--- no_of_special_requests <= 1.50
    |    |    |--- weights: [0.00, 8.00] class: 1
    |--- no_of_special_requests > 1.50
    |    |--- arrival_date <= 19.50
    |    |    |--- weights: [0.00, 1.00] class: 1
    |--- arrival_date > 19.50
    |    |    |--- weights: [2.00, 0.00] class: 0
    --- no_of_week_nights > 2.50
```

Project 4

```

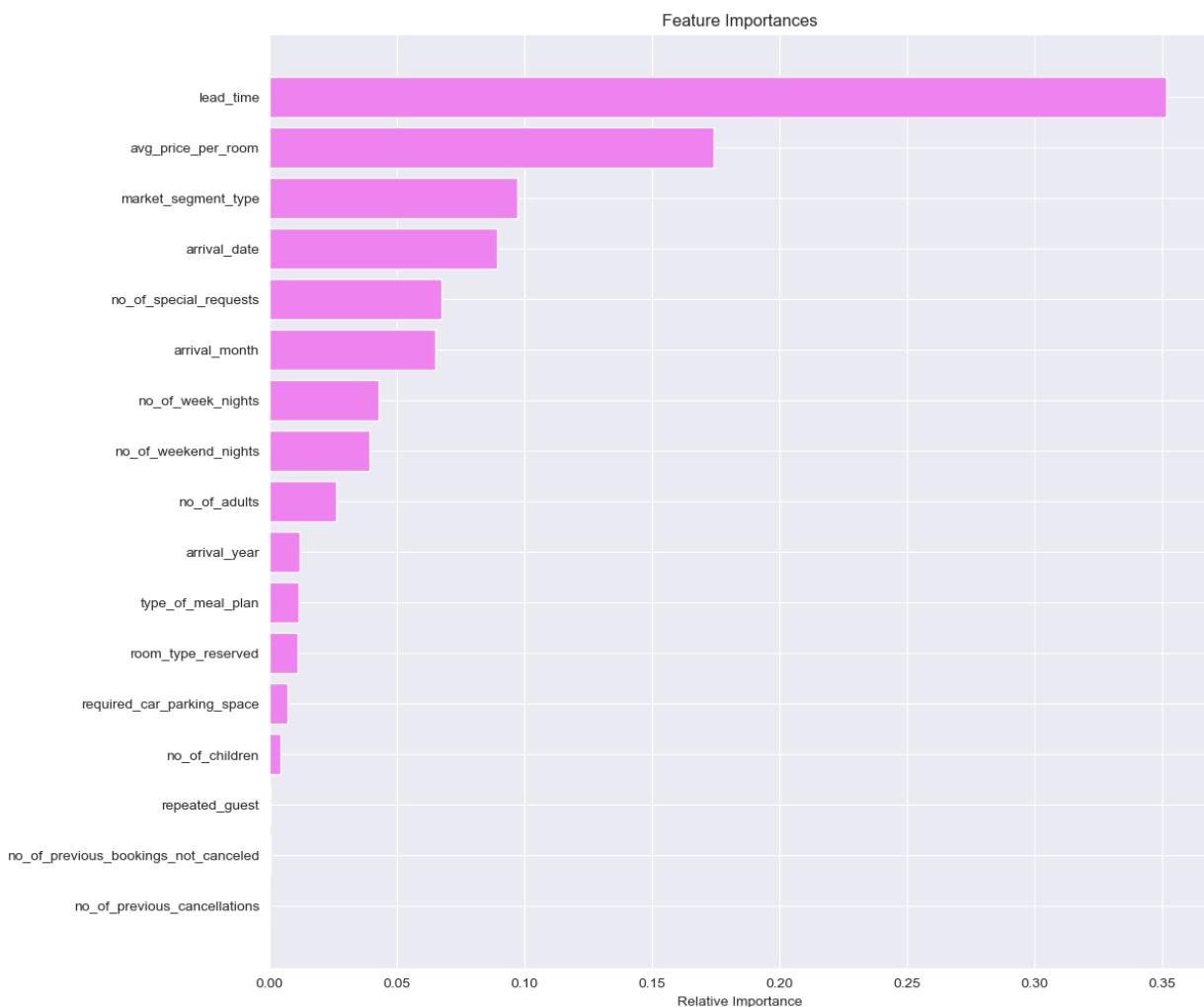
    |--- no_of_week_nights <= 4.50
    |--- lead_time <= 281.50
    |--- weights: [8.00, 0.00] class: 0
    |--- lead_time > 281.50
    |--- truncated branch of depth 4
    |--- no_of_week_nights > 4.50
    |--- no_of_special_requests <= 2.50
    |--- truncated branch of depth 2
    |--- no_of_special_requests > 2.50
    |--- weights: [2.00, 0.00] class: 0
    |--- no_of_week_nights > 9.50
    |--- lead_time <= 198.00
    |--- weights: [1.00, 0.00] class: 0
    |--- lead_time > 198.00
    |--- weights: [0.00, 7.00] class: 1
--- market_segment_type > 1.50
    |--- lead_time <= 348.50
    |--- no_of_week_nights <= 5.50
    |--- arrival_date <= 30.00
    |--- weights: [137.00, 0.00] class: 0
    |--- arrival_date > 30.00
    |--- no_of_week_nights <= 3.00
    |--- weights: [3.00, 0.00] class: 0
    |--- no_of_week_nights > 3.00
    |--- weights: [2.00, 1.00] class: 0
    |--- no_of_week_nights > 5.50
    |--- avg_price_per_room <= 76.75
    |--- weights: [1.00, 0.00] class: 0
    |--- avg_price_per_room > 76.75
    |--- weights: [0.00, 1.00] class: 1
    |--- lead_time > 348.50
    |--- avg_price_per_room <= 84.00
    |--- avg_price_per_room <= 58.50
    |--- weights: [1.00, 0.00] class: 0
    |--- avg_price_per_room > 58.50
    |--- weights: [6.00, 2.00] class: 0
    |--- avg_price_per_room > 84.00
    |--- weights: [1.00, 1.00] class: 0
--- avg_price_per_room > 100.04
    |--- arrival_month <= 11.50
    |--- no_of_special_requests <= 2.50
    |--- weights: [0.00, 2108.00] class: 1
    |--- no_of_special_requests > 2.50
    |--- weights: [31.00, 0.00] class: 0
    |--- arrival_month > 11.50
    |--- no_of_special_requests <= 0.50
    |--- weights: [47.00, 0.00] class: 0
    |--- no_of_special_requests > 0.50
    |--- arrival_date <= 24.50
    |--- weights: [5.00, 0.00] class: 0
    |--- arrival_date > 24.50
    |--- room_type_reserved <= 2.50
    |--- no_of_week_nights <= 3.00
    |--- weights: [0.00, 1.00] class: 1
    |--- no_of_week_nights > 3.00
    |--- weights: [3.00, 0.00] class: 0
    |--- room_type_reserved > 2.50
    |--- lead_time <= 172.50
    |--- no_of_adults <= 2.50
    |--- weights: [0.00, 1.00] class: 1

```

```
| | | | | |--- no_of_adults > 2.50
| | | | | |--- weights: [2.00, 0.00] class: 0
| | | | |--- lead_time > 172.50
| | | | | |--- weights: [0.00, 13.00] class: 1
```

```
In [93]: importances = model.feature_importances_
indices = np.argsort(importances)

plt.figure(figsize=(12, 12))
plt.title("Feature Importances")
plt.barh(range(len(indices)), importances[indices], color="violet", align="center")
plt.yticks(range(len(indices)), [feature_names[i] for i in indices])
plt.xlabel("Relative Importance")
plt.show()
```



```
In [94]: estimator = DecisionTreeClassifier(random_state=1)

# Grid of parameters to choose from

parameters = {
    "max_depth": [np.arange(2, 50, 5), None],
    "criterion": ["entropy", "gini"],
    "splitter": ["best", "random"],
    "min_impurity_decrease": [0.000001, 0.00001, 0.0001],
}
```

```
# Type of scoring used to compare parameter combinations
acc_scoring = make_scorer(recall_score)

# Run the grid search
grid_obj = GridSearchCV(estimator, parameters, scoring=acc_scoring, cv=5)
grid_obj = grid_obj.fit(X_train, y_train)

# Set the clf to the best combination of parameters
estimator = grid_obj.best_estimator_

# Fit the best algorithm to the data.
estimator.fit(X_train, y_train)
```

Out[94]:

```
▼ DecisionTreeClassifier
DecisionTreeClassifier(criterion='entropy', min_impurity_decrease=1e-05,
random_state=1)
```

In [95]:

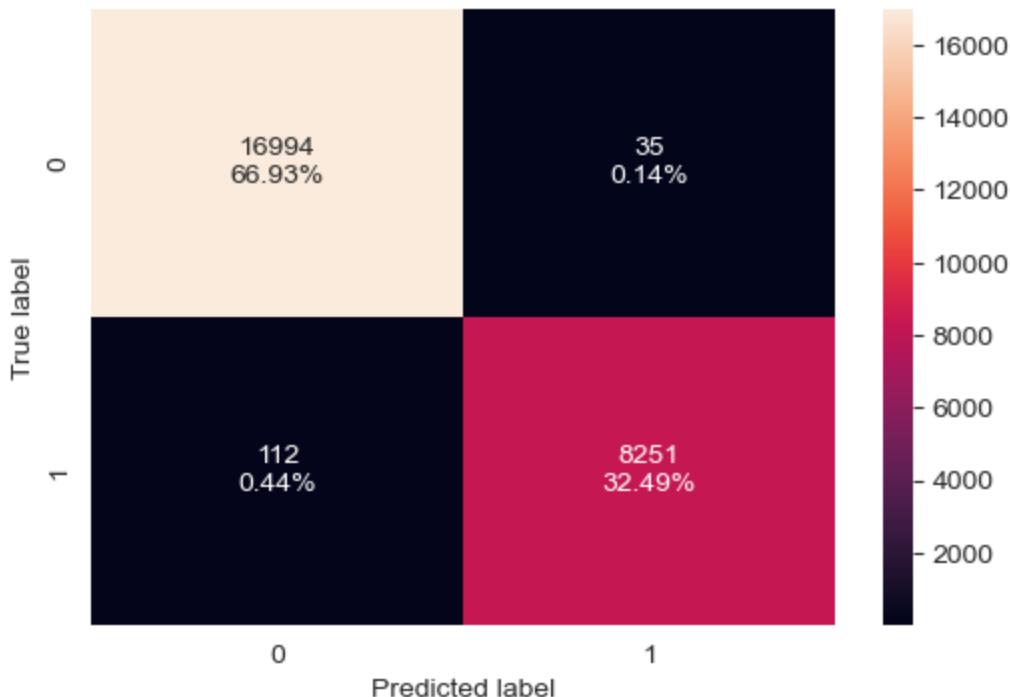
```
decision_tree_tune_perf_train = model_performance_classification_sklearn(
    estimator, X_train, y_train
)
decision_tree_tune_perf_train
```

Out[95]:

	Accuracy	Recall	Precision	F1
0	0.994211	0.986608	0.995776	0.991171

In [96]:

```
confusion_matrix_sklearn(estimator, X_train, y_train)
```



In [97]:

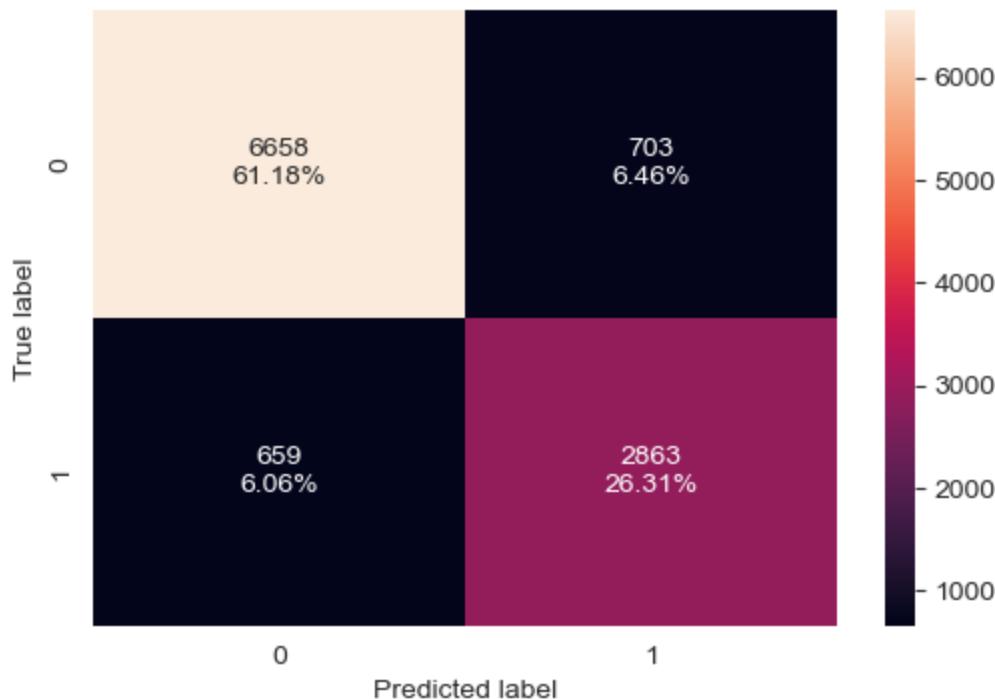
```
decision_tree_tune_perf_test = model_performance_classification_sklearn(
    estimator, X_test, y_test
)
decision_tree_tune_perf_test
```

Out[97]:

	Accuracy	Recall	Precision	F1
0	0.874851	0.81289	0.80286	0.807844

In [98]:

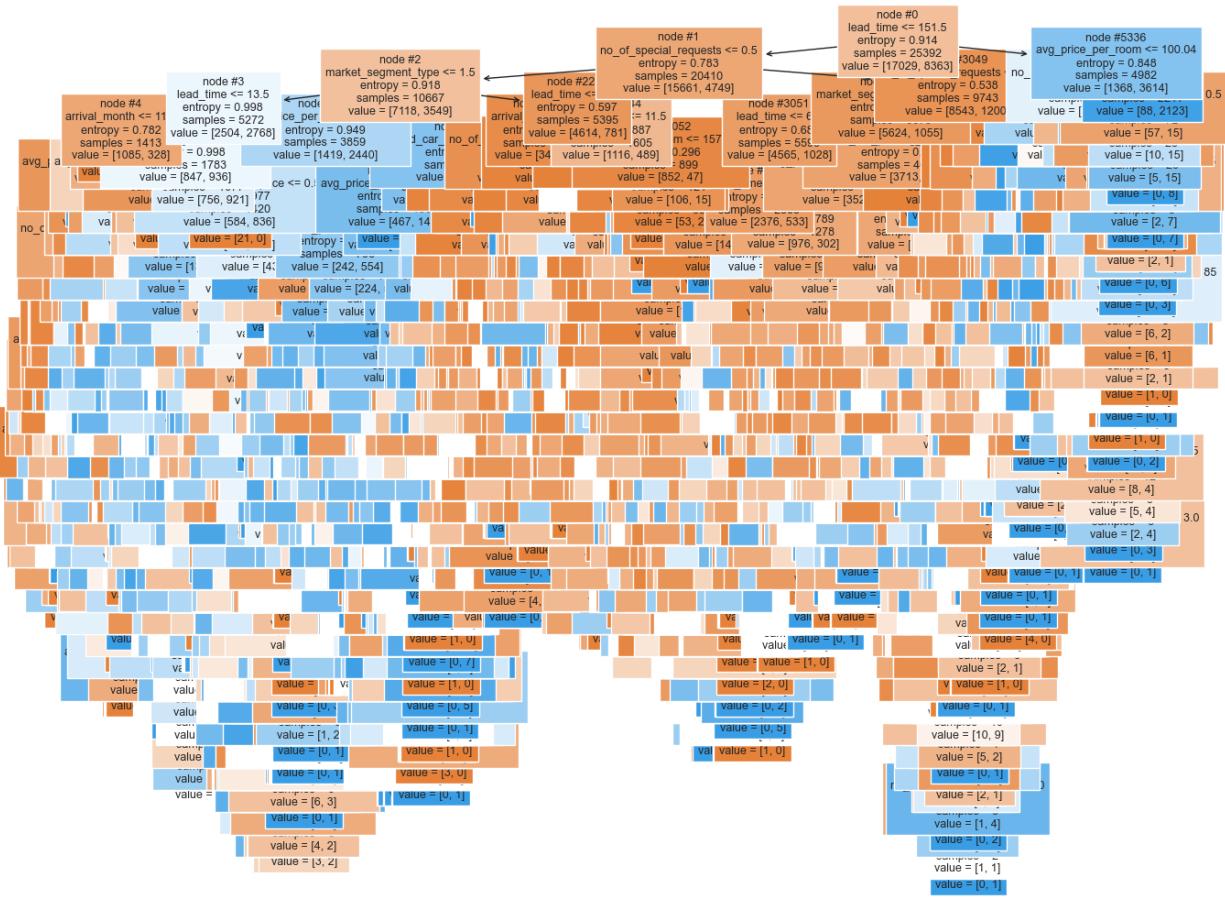
```
confusion_matrix_sklearn(estimator, X_test, y_test)
```



In [99]:

```
plt.figure(figsize=(15, 12))

tree.plot_tree(
    estimator,
    feature_names=feature_names,
    filled=True,
    fontsize=9,
    node_ids=True,
    class_names=None,
)
plt.show()
```



I will now print the design tree

```
In [100]: clf = DecisionTreeClassifier(random_state=1)
path = clf.cost_complexity_pruning_path(X_train, y_train)
ccp_alphas, impurities = path ccp_alphas, path impurities
pd.DataFrame(path)
```

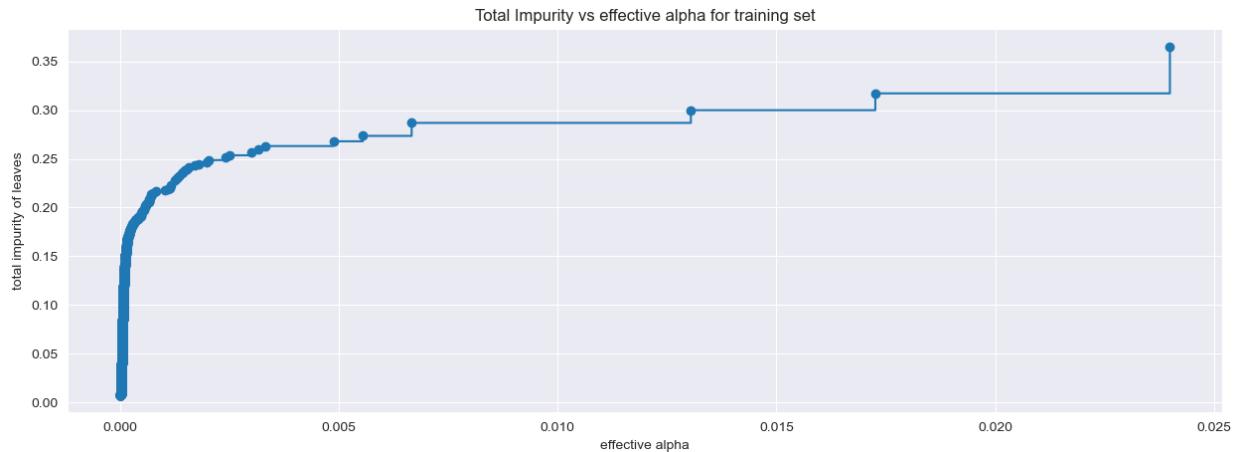
Out[100]:

	ccp_alphas	impurities
0	0.000000e+00	0.007572
1	4.327745e-07	0.007573
2	4.688391e-07	0.007573
3	5.329960e-07	0.007574
4	6.133547e-07	0.007575
...
1337	6.665684e-03	0.286897
1338	1.304480e-02	0.299942
1339	1.725993e-02	0.317202
1340	2.399048e-02	0.365183
1341	7.657789e-02	0.441761

1342 rows × 2 columns

In [101...]

```
fig, ax = plt.subplots(figsize=(15, 5))
ax.plot(ccp_alphas[:-1], impurities[:-1], marker="o", drawstyle="steps-post")
ax.set_xlabel("effective alpha")
ax.set_ylabel("total impurity of leaves")
ax.set_title("Total Impurity vs effective alpha for training set")
plt.show()
```



In [102...]

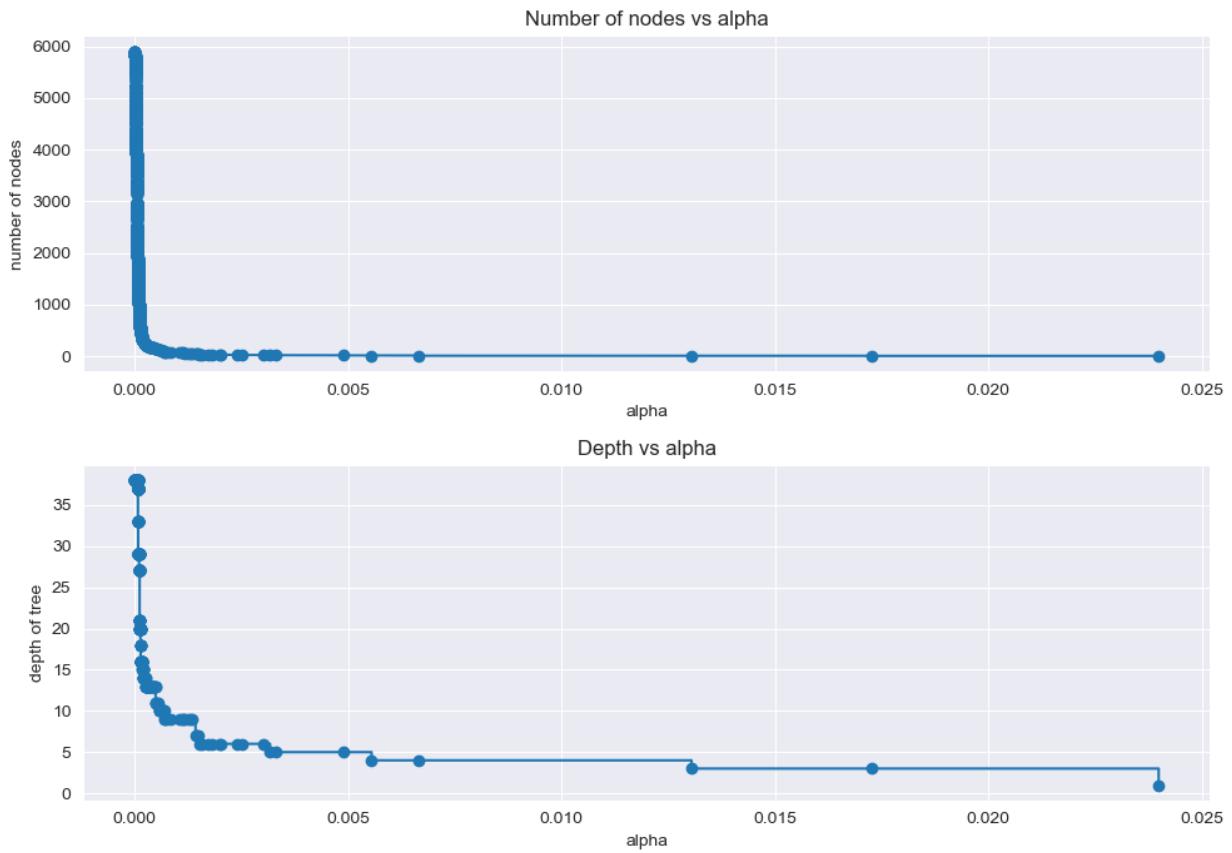
```
clfs = []
for ccp_alpha in ccp_alphas:
    clf = DecisionTreeClassifier(random_state=1, ccp_alpha=ccp_alpha)
    clf.fit(X_train, y_train)
    clfs.append(clf)
print(
    "Number of nodes in the last tree is: {} with ccp_alpha: {}".format(
        clfs[-1].tree_.node_count, ccp_alphas[-1]))
```

Number of nodes in the last tree is: 1 with ccp_alpha: 0.07657789477371357

In [103...]

```
clfs = clfs[:-1]
ccp_alphas = ccp_alphas[:-1]

node_counts = [clf.tree_.node_count for clf in clfs]
depth = [clf.tree_.max_depth for clf in clfs]
fig, ax = plt.subplots(2, 1, figsize=(10, 7))
ax[0].plot(ccp_alphas, node_counts, marker="o", drawstyle="steps-post")
ax[0].set_xlabel("alpha")
ax[0].set_ylabel("number of nodes")
ax[0].set_title("Number of nodes vs alpha")
ax[1].plot(ccp_alphas, depth, marker="o", drawstyle="steps-post")
ax[1].set_xlabel("alpha")
ax[1].set_ylabel("depth of tree")
ax[1].set_title("Depth vs alpha")
fig.tight_layout()
```



```
In [104...]: recall_train = []
for clf in clfs:
    pred_train = clf.predict(X_train)
    values_train = recall_score(y_train, pred_train)
    recall_train.append(values_train)
recall_test = []
for clf in clfs:
    pred_test = clf.predict(X_test)
    values_test = recall_score(y_test, pred_test)
    recall_test.append(values_test)
fig, ax = plt.subplots(figsize=(15, 5))
ax.set_xlabel("alpha")
ax.set_ylabel("Recall")
ax.set_title("Recall vs alpha for training and testing sets")
ax.plot(ccp_alphas, recall_train, marker="o", label="train", drawstyle="steps-post")
ax.plot(ccp_alphas, recall_test, marker="o", label="test", drawstyle="steps-post")
ax.legend()
plt.show()
```



```
In [105]: index_best_model = np.argmax(recall_test)
best_model = clfs[index_best_model]
print(best_model)
```

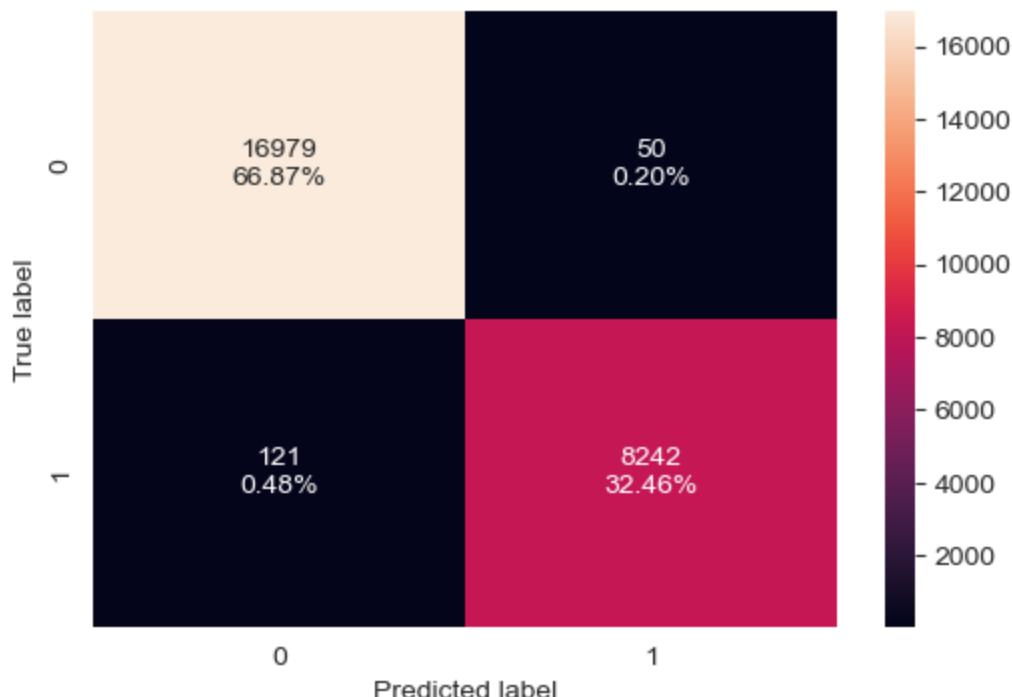
```
DecisionTreeClassifier(ccp_alpha=2.4891092944023008e-05, random_state=1)
```

```
In [106]: decision_tree_postpruned_perf_train = model_performance_classification_sklearn(
    best_model, X_train, y_train
)
decision_tree_postpruned_perf_train
```

```
Out[106]:
```

	Accuracy	Recall	Precision	F1
0	0.993266	0.985532	0.99397	0.989733

```
In [107]: confusion_matrix_sklearn(best_model, X_train, y_train)
```



```
In [108]: decision_tree_postpruned_perf_test = model_performance_classification_sklearn(
    best_model, X_test, y_test
```

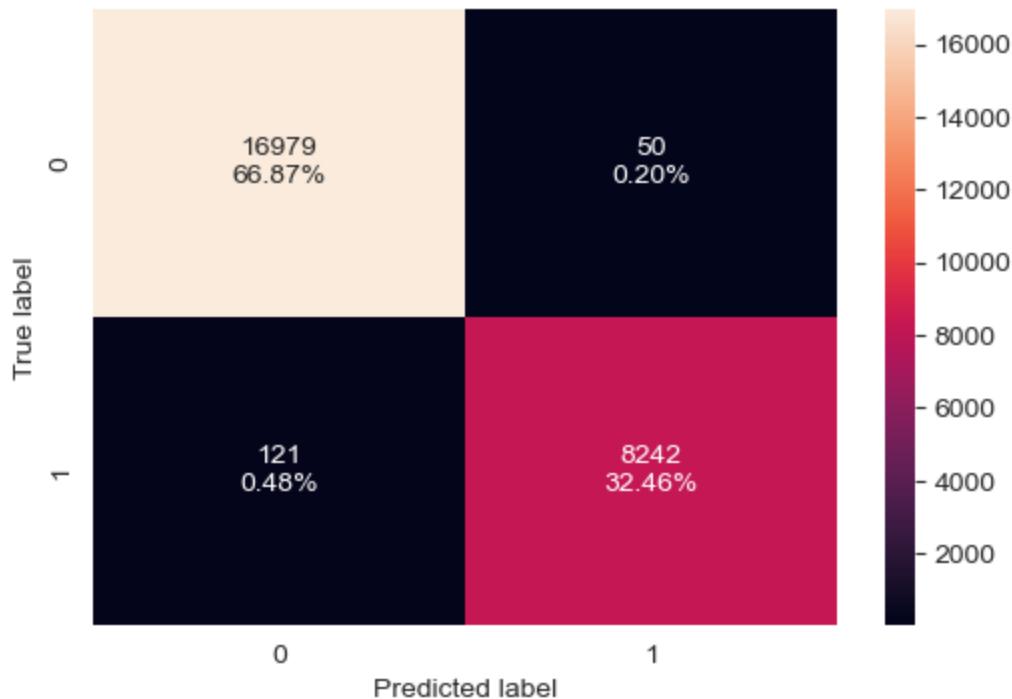
```
)  
decision_tree_postpruned_perf_test
```

Out[108]:

	Accuracy	Recall	Precision	F1
0	0.871451	0.809483	0.796591	0.802985

In [109...]

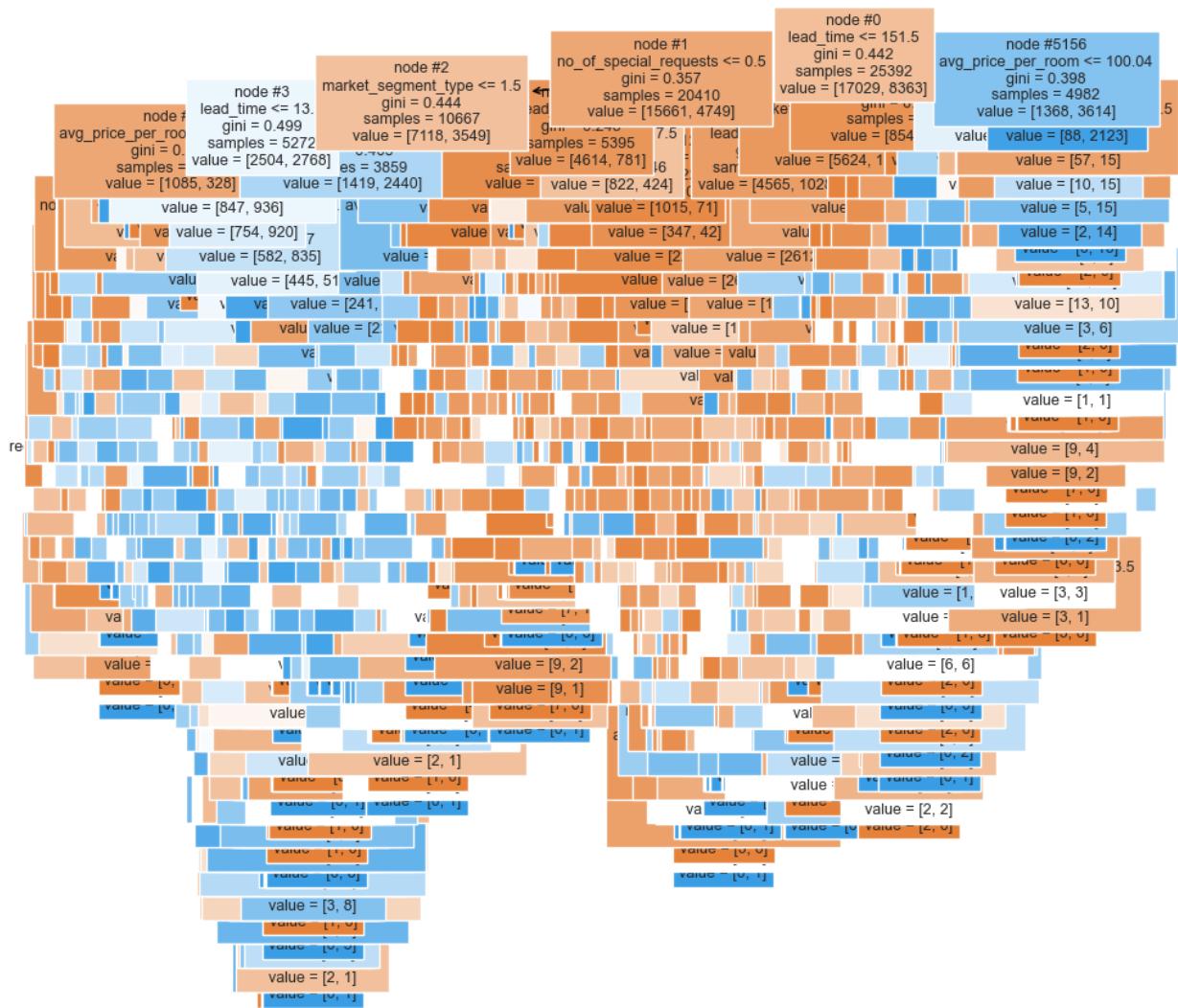
```
confusion_matrix_sklearn(best_model, X_train, y_train)
```



In [110...]

```
plt.figure(figsize=(10, 10))

out = tree.plot_tree(
    best_model,
    feature_names=feature_names,
    filled=True,
    fontsize=9,
    node_ids=True,
    class_names=None,
)
for o in out:
    arrow = o.arrow_patch
    if arrow is not None:
        arrow.set_edgecolor("black")
        arrow.set linewidth(1)
plt.show()
```



```
In [111]: print(tree.export_text(best_model, feature_names=feature_names, show_weights=True))
```

```
--- lead_time <= 151.50
    --- no_of_special_requests <= 0.50
        --- market_segment_type <= 1.50
            --- lead_time <= 13.50
                --- avg_price_per_room <= 202.67
                    --- lead_time <= 3.50
                        --- arrival_month <= 5.50
                            --- no_of_weekend_nights <= 1.50
                                --- arrival_month <= 1.50
                                    --- weights: [56.00, 0.00] class: 0
                                --- arrival_month > 1.50
                                    --- avg_price_per_room <= 77.50
                                        --- weights: [24.00, 0.00] class: 0
                                    --- avg_price_per_room > 77.50
                                        --- arrival_date <= 26.50
                                            --- truncated branch of depth 14
                                        --- arrival_date > 26.50
                                            --- truncated branch of depth 4
                                --- no_of_weekend_nights > 1.50
                                    --- arrival_date <= 25.50
                                        --- avg_price_per_room <= 134.22
                                            --- lead_time <= 2.50
                                                --- weights: [17.00, 0.00] class: 0
                                            --- lead_time > 2.50
                                                --- truncated branch of depth 3
                                        --- avg_price_per_room > 134.22
                                            --- weights: [0.00, 2.00] class: 1
                                    --- arrival_date > 25.50
                                        --- arrival_month <= 1.50
                                            --- weights: [1.00, 0.00] class: 0
                                        --- arrival_month > 1.50
                                            --- weights: [0.00, 14.00] class: 1
                --- arrival_month > 5.50
                    --- no_of_week_nights <= 8.50
                        --- arrival_month <= 9.50
                            --- avg_price_per_room <= 76.35
                                --- avg_price_per_room <= 74.40
                                    --- truncated branch of depth 4
                                --- avg_price_per_room > 74.40
                                    --- truncated branch of depth 2
                            --- avg_price_per_room > 76.35
                                --- avg_price_per_room <= 118.04
                                    --- weights: [103.00, 2.00] class: 0
                                --- avg_price_per_room > 118.04
                                    --- truncated branch of depth 16
                        --- arrival_month > 9.50
                            --- avg_price_per_room <= 178.00
                                --- lead_time <= 1.50
                                    --- truncated branch of depth 4
                                --- lead_time > 1.50
                                    --- truncated branch of depth 5
                            --- avg_price_per_room > 178.00
                                --- lead_time <= 0.50
                                    --- weights: [0.00, 1.00] class: 1
                                --- lead_time > 0.50
                                    --- weights: [2.00, 0.00] class: 0
                            --- no_of_week_nights > 8.50
                                --- weights: [0.00, 1.00] class: 1
                    --- lead_time > 3.50
                        --- avg_price_per_room <= 99.38
```

Project 4

```

    |--- avg_price_per_room <= 78.90
    |   |--- no_of_weekend_nights <= 4.50
    |   |   |--- no_of_week_nights <= 5.50
    |   |   |   |--- arrival_date <= 23.50
    |   |   |   |   |--- weights: [100.00, 0.00] class: 0
    |   |   |   |   |--- arrival_date > 23.50
    |   |   |   |   |   |--- truncated branch of depth 3
    |   |   |   |--- no_of_week_nights > 5.50
    |   |   |   |   |--- type_of_meal_plan <= 0.50
    |   |   |   |   |   |--- weights: [0.00, 1.00] class: 1
    |   |   |   |   |--- type_of_meal_plan > 0.50
    |   |   |   |   |   |--- weights: [4.00, 0.00] class: 0
    |   |   |--- no_of_weekend_nights > 4.50
    |   |   |   |--- weights: [0.00, 1.00] class: 1
    |--- avg_price_per_room > 78.90
    |--- arrival_month <= 11.50
    |   |--- arrival_month <= 1.50
    |   |   |--- weights: [23.00, 0.00] class: 0
    |   |--- arrival_month > 1.50
    |   |   |--- no_of_week_nights <= 4.50
    |   |   |   |--- truncated branch of depth 11
    |   |   |--- no_of_week_nights > 4.50
    |   |   |   |   |--- truncated branch of depth 2
    |   |--- arrival_month > 11.50
    |   |   |--- weights: [42.00, 0.00] class: 0
--- avg_price_per_room > 99.38
    |--- arrival_month <= 8.50
    |   |--- required_car_parking_space <= 0.50
    |   |   |--- avg_price_per_room <= 119.25
    |   |   |   |--- avg_price_per_room <= 117.25
    |   |   |   |   |--- truncated branch of depth 8
    |   |   |   |   |--- avg_price_per_room > 117.25
    |   |   |   |   |   |--- truncated branch of depth 3
    |   |--- avg_price_per_room > 119.25
    |   |   |--- no_of_weekend_nights <= 1.50
    |   |   |   |--- truncated branch of depth 12
    |   |   |--- no_of_weekend_nights > 1.50
    |   |   |   |   |--- truncated branch of depth 6
    |--- required_car_parking_space > 0.50
    |   |--- weights: [5.00, 0.00] class: 0
--- arrival_month > 8.50
    |--- lead_time <= 9.50
    |   |--- arrival_date <= 6.50
    |   |   |--- arrival_date <= 4.50
    |   |   |   |--- truncated branch of depth 3
    |   |   |--- arrival_date > 4.50
    |   |   |   |   |--- truncated branch of depth 2
    |--- arrival_date > 6.50
    |   |--- arrival_date <= 18.50
    |   |   |--- truncated branch of depth 5
    |   |   |--- arrival_date > 18.50
    |   |   |   |--- weights: [34.00, 0.00] class: 0
--- lead_time > 9.50
    |--- arrival_month <= 11.50
    |   |--- arrival_date <= 26.00
    |   |   |--- truncated branch of depth 8
    |   |   |--- arrival_date > 26.00
    |   |   |   |--- weights: [5.00, 0.00] class: 0
--- arrival_month > 11.50
    |--- weights: [10.00, 0.00] class: 0

```

```
    |--- avg_price_per_room > 202.67
    |--- arrival_month <= 11.50
    |   |--- weights: [0.00, 32.00] class: 1
    |--- arrival_month > 11.50
    |   |--- weights: [1.00, 0.00] class: 0
--- lead_time > 13.50
    --- avg_price_per_room <= 105.27
        --- avg_price_per_room <= 60.07
            --- lead_time <= 84.50
                --- lead_time <= 51.50
                    |--- lead_time <= 50.50
                        |--- avg_price_per_room <= 21.67
                            |--- weights: [19.00, 0.00] class: 0
                        |--- avg_price_per_room > 21.67
                            |--- avg_price_per_room <= 49.84
                                |--- truncated branch of depth 2
                            |--- avg_price_per_room > 49.84
                                |--- weights: [18.00, 1.00] class: 0
                        |--- lead_time > 50.50
                            |--- weights: [0.00, 1.00] class: 1
                    |--- lead_time > 51.50
                        |--- weights: [32.00, 0.00] class: 0
                |--- lead_time > 84.50
                    --- arrival_year <= 2017.50
                        --- arrival_date <= 19.00
                            |--- lead_time <= 139.00
                                |--- weights: [0.00, 8.00] class: 1
                            |--- lead_time > 139.00
                                |--- weights: [1.00, 0.00] class: 0
                        |--- arrival_date > 19.00
                            |--- lead_time <= 87.50
                                |--- weights: [0.00, 1.00] class: 1
                            |--- lead_time > 87.50
                                |--- no_of_weekend_nights <= 0.50
                                    |--- truncated branch of depth 2
                                |--- no_of_weekend_nights > 0.50
                                    |--- weights: [6.00, 0.00] class: 0
                    |--- arrival_year > 2017.50
                        |--- avg_price_per_room <= 59.43
                            |--- weights: [14.00, 0.00] class: 0
                        |--- avg_price_per_room > 59.43
                            |--- weights: [0.00, 1.00] class: 1
--- avg_price_per_room > 60.07
    --- lead_time <= 25.50
        --- arrival_month <= 11.50
            |--- arrival_month <= 1.50
                |--- weights: [29.00, 0.00] class: 0
            |--- arrival_month > 1.50
                |--- arrival_year <= 2017.50
                    |--- lead_time <= 14.50
                        |--- weights: [0.00, 1.00] class: 1
                    |--- lead_time > 14.50
                        |--- truncated branch of depth 3
                |--- arrival_year > 2017.50
                    |--- no_of_week_nights <= 3.50
                        |--- truncated branch of depth 14
                    |--- no_of_week_nights > 3.50
                        |--- truncated branch of depth 4
            |--- arrival_month > 11.50
                |--- weights: [54.00, 0.00] class: 0
```

```

    |--- lead_time >  25.50
    |--- type_of_meal_plan <= 0.50
    |   |--- required_car_parking_space <= 0.50
    |   |   |--- arrival_year <= 2017.50
    |   |   |   |--- lead_time <= 57.00
    |   |   |   |   |--- weights: [6.00, 0.00] class: 0
    |   |   |   |--- lead_time > 57.00
    |   |   |   |   |--- weights: [0.00, 1.00] class: 1
    |   |--- arrival_year > 2017.50
    |   |   |--- no_of_adults <= 1.50
    |   |   |   |--- truncated branch of depth 5
    |   |   |   |--- no_of_adults > 1.50
    |   |   |   |   |--- truncated branch of depth 15
    |--- required_car_parking_space > 0.50
    |   |--- weights: [6.00, 0.00] class: 0
    |--- type_of_meal_plan > 0.50
    |   |--- type_of_meal_plan <= 1.50
    |   |   |--- arrival_year <= 2017.50
    |   |   |   |--- lead_time <= 60.50
    |   |   |   |   |--- truncated branch of depth 6
    |   |   |   |--- lead_time > 60.50
    |   |   |   |   |--- truncated branch of depth 10
    |   |--- arrival_year > 2017.50
    |   |   |--- required_car_parking_space <= 0.50
    |   |   |   |--- truncated branch of depth 28
    |   |   |   |--- required_car_parking_space > 0.50
    |   |   |   |   |--- weights: [12.00, 0.00] class: 0
    |--- type_of_meal_plan > 1.50
    |   |--- arrival_month <= 5.00
    |   |   |--- weights: [2.00, 0.00] class: 0
    |   |--- arrival_month > 5.00
    |   |   |--- lead_time <= 57.50
    |   |   |   |--- truncated branch of depth 2
    |   |   |   |--- lead_time > 57.50
    |   |   |   |   |--- weights: [0.00, 35.00] class: 1
--- avg_price_per_room > 105.27
|--- required_car_parking_space <= 0.50
|--- arrival_month <= 10.50
|   |--- avg_price_per_room <= 195.30
|   |--- lead_time <= 54.50
|   |   |--- arrival_year <= 2017.50
|   |   |   |--- arrival_month <= 9.50
|   |   |   |   |--- truncated branch of depth 4
|   |   |   |--- arrival_month > 9.50
|   |   |   |   |--- weights: [11.00, 0.00] class: 0
|--- arrival_year > 2017.50
|   |--- lead_time <= 33.50
|   |   |--- truncated branch of depth 16
|   |--- lead_time > 33.50
|   |   |--- truncated branch of depth 15
|--- lead_time > 54.50
|--- arrival_month <= 8.50
|   |--- lead_time <= 135.50
|   |   |--- truncated branch of depth 20
|   |--- lead_time > 135.50
|   |   |--- truncated branch of depth 9
|--- arrival_month > 8.50
|   |--- lead_time <= 59.50
|   |   |--- truncated branch of depth 6
|   |--- lead_time > 59.50

```

Project 4

```

| | | | --- truncated branch of depth 12
| --- avg_price_per_room > 195.30
| --- type_of_meal_plan <= 0.50
| | --- weights: [1.00, 0.00] class: 0
| --- type_of_meal_plan > 0.50
| | --- no_of_adults <= 1.50
| | | --- lead_time <= 59.50
| | | | --- weights: [0.00, 6.00] class: 1
| | | --- lead_time > 59.50
| | | | --- weights: [1.00, 0.00] class: 0
| | --- no_of_adults > 1.50
| | | --- weights: [0.00, 92.00] class: 1
--- arrival_month > 10.50
--- lead_time <= 22.50
| --- no_of_adults <= 1.50
| | --- arrival_month <= 11.50
| | | --- weights: [0.00, 4.00] class: 1
| | --- arrival_month > 11.50
| | | --- weights: [1.00, 0.00] class: 0
| --- no_of_adults > 1.50
| | --- weights: [22.00, 0.00] class: 0
--- lead_time > 22.50
| --- avg_price_per_room <= 168.06
| | --- avg_price_per_room <= 147.75
| | | --- no_of_week_nights <= 3.50
| | | | --- truncated branch of depth 8
| | | --- no_of_week_nights > 3.50
| | | | --- truncated branch of depth 4
| | --- avg_price_per_room > 147.75
| | | --- weights: [0.00, 15.00] class: 1
| --- avg_price_per_room > 168.06
| | --- lead_time <= 80.00
| | | --- no_of_week_nights <= 6.50
| | | | --- truncated branch of depth 3
| | | --- no_of_week_nights > 6.50
| | | | --- weights: [0.00, 1.00] class: 1
| | --- lead_time > 80.00
| | | --- weights: [0.00, 1.00] class: 1
--- required_car_parking_space > 0.50
| --- no_of_week_nights <= 8.00
| | --- weights: [39.00, 0.00] class: 0
| --- no_of_week_nights > 8.00
| | | --- weights: [0.00, 1.00] class: 1
_segment_type > 1.50
lead_time <= 90.50
avg_price_per_room <= 201.50
| --- no_of_weekend_nights <= 0.50
| | --- market_segment_type <= 2.50
| | | --- weights: [1609.00, 0.00] class: 0
| --- market_segment_type > 2.50
| | --- lead_time <= 16.50
| | | --- avg_price_per_room <= 68.50
| | | | --- repeated_guest <= 1.00
| | | | | --- avg_price_per_room <= 48.50
| | | | | | --- weights: [63.00, 0.00] class: 0
| | | | --- avg_price_per_room > 48.50
| | | | | --- truncated branch of depth 8
| | | --- repeated_guest > 1.00
| | | | | --- weights: [109.00, 0.00] class: 0
| --- avg_price_per_room > 68.50

```

Project 4

```

    |--- arrival_date <= 29.50
    |   |--- no_of_adults <= 1.50
    |   |   |--- truncated branch of depth 11
    |   |--- no_of_adults >  1.50
    |   |   |--- truncated branch of depth 6
    |--- arrival_date >  29.50
    |   |--- arrival_date <= 30.50
    |   |   |--- weights: [0.00, 4.00] class: 1
    |   |--- arrival_date >  30.50
    |   |   |--- truncated branch of depth 2
    --- lead_time >  16.50
    |--- avg_price_per_room <= 135.00
    |   |--- arrival_month <= 11.50
    |   |   |--- lead_time <= 17.50
    |   |   |   |--- truncated branch of depth 3
    |   |--- lead_time >  17.50
    |   |   |--- truncated branch of depth 12
    |--- arrival_month >  11.50
    |   |--- weights: [29.00, 0.00] class: 0
    |--- avg_price_per_room >  135.00
    |   |--- weights: [0.00, 8.00] class: 1
--- no_of_weekend_nights >  0.50
--- lead_time <= 68.50
    |--- no_of_weekend_nights <= 4.50
    |   |--- lead_time <= 1.50
    |   |   |--- arrival_date <= 27.50
    |   |   |   |--- no_of_week_nights <= 5.50
    |   |   |   |   |--- truncated branch of depth 7
    |   |   |   |--- no_of_week_nights >  5.50
    |   |   |   |--- truncated branch of depth 2
    |--- arrival_date >  27.50
    |   |--- arrival_month <= 2.50
    |   |   |--- truncated branch of depth 2
    |--- arrival_month >  2.50
    |   |--- truncated branch of depth 2
    --- lead_time >  1.50
    |--- arrival_month <= 9.50
    |   |--- lead_time <= 59.50
    |   |   |--- truncated branch of depth 17
    |--- lead_time >  59.50
    |   |--- truncated branch of depth 5
    |--- arrival_month >  9.50
    |   |--- market_segment_type <= 4.50
    |   |   |--- truncated branch of depth 11
    |   |--- market_segment_type >  4.50
    |   |   |--- truncated branch of depth 3
    |--- no_of_weekend_nights >  4.50
    |   |--- weights: [0.00, 8.00] class: 1
--- lead_time >  68.50
    |--- avg_price_per_room <= 99.98
    |   |--- arrival_month <= 3.50
    |   |   |--- avg_price_per_room <= 62.50
    |   |   |   |--- weights: [21.00, 0.00] class: 0
    |   |--- avg_price_per_room >  62.50
    |   |   |--- lead_time <= 77.00
    |   |   |   |--- truncated branch of depth 2
    |   |--- lead_time >  77.00
    |   |   |   |--- truncated branch of depth 5
    |--- arrival_month >  3.50
    |   |--- lead_time <= 71.50

```

Project 4

```

|--- no_of_week_nights <= 3.50
|   |--- weights: [3.00, 0.00] class: 0
|--- no_of_week_nights > 3.50
|   |--- weights: [0.00, 2.00] class: 1
|--- lead_time > 71.50
|   |--- no_of_week_nights <= 2.50
|       |--- truncated branch of depth 4
|       |--- no_of_week_nights > 2.50
|           |--- truncated branch of depth 7
|--- avg_price_per_room > 99.98
|   |--- no_of_adults <= 1.50
|       |--- avg_price_per_room <= 123.25
|           |--- weights: [0.00, 52.00] class: 1
|       |--- avg_price_per_room > 123.25
|           |--- weights: [1.00, 0.00] class: 0
|   |--- no_of_adults > 1.50
|       |--- arrival_date <= 23.50
|           |--- arrival_date <= 3.00
|               |--- weights: [0.00, 2.00] class: 1
|           |--- arrival_date > 3.00
|               |--- truncated branch of depth 6
|       |--- arrival_date > 23.50
|           |--- avg_price_per_room <= 131.67
|               |--- truncated branch of depth 3
|           |--- avg_price_per_room > 131.67
|               |--- weights: [2.00, 0.00] class: 0
|--- avg_price_per_room > 201.50
|   |--- arrival_date <= 28.00
|       |--- weights: [0.00, 17.00] class: 1
|   |--- arrival_date > 28.00
|       |--- weights: [2.00, 0.00] class: 0
--- lead_time > 90.50
--- lead_time <= 117.50
    |--- avg_price_per_room <= 93.58
        |--- arrival_date <= 6.50
            |--- no_of_week_nights <= 2.50
                |--- no_of_weekend_nights <= 1.50
                    |--- avg_price_per_room <= 80.38
                        |--- type_of_meal_plan <= 0.50
                            |--- weights: [1.00, 0.00] class: 0
                        |--- type_of_meal_plan > 0.50
                            |--- truncated branch of depth 4
                    |--- avg_price_per_room > 80.38
                        |--- room_type_reserved <= 2.50
                            |--- weights: [3.00, 0.00] class: 0
                        |--- room_type_reserved > 2.50
                            |--- weights: [0.00, 1.00] class: 1
                |--- no_of_weekend_nights > 1.50
                    |--- arrival_date <= 5.00
                        |--- weights: [5.00, 0.00] class: 0
                    |--- arrival_date > 5.00
                        |--- weights: [0.00, 1.00] class: 1
            |--- no_of_week_nights > 2.50
                |--- arrival_date <= 5.50
                    |--- weights: [35.00, 0.00] class: 0
            |--- arrival_date > 5.50
                |--- arrival_month <= 9.00
                    |--- weights: [3.00, 0.00] class: 0
            |--- arrival_month > 9.00
                |--- no_of_week_nights <= 3.50

```

Project 4

```
|--- weights: [1.00, 0.00] class: 0
|--- no_of_week_nights > 3.50
|--- weights: [0.00, 2.00] class: 1
--- arrival_date > 6.50
|--- avg_price_per_room <= 66.50
|--- no_of_weekend_nights <= 1.50
|--- arrival_date <= 9.50
|--- no_of_week_nights <= 3.00
|--- weights: [0.00, 1.00] class: 1
|--- no_of_week_nights > 3.00
|--- weights: [2.00, 0.00] class: 0
|--- arrival_date > 9.50
|--- weights: [24.00, 0.00] class: 0
|--- no_of_weekend_nights > 1.50
|--- avg_price_per_room <= 58.75
|--- weights: [6.00, 0.00] class: 0
|--- avg_price_per_room > 58.75
|--- lead_time <= 97.50
|--- truncated branch of depth 2
|--- lead_time > 97.50
|--- weights: [0.00, 39.00] class: 1
--- avg_price_per_room > 66.50
|--- type_of_meal_plan <= 1.50
|--- arrival_date <= 29.50
|--- type_of_meal_plan <= 0.50
|--- weights: [0.00, 1.00] class: 1
|--- type_of_meal_plan > 0.50
|--- truncated branch of depth 10
|--- arrival_date > 29.50
|--- lead_time <= 96.00
|--- weights: [8.00, 0.00] class: 0
|--- lead_time > 96.00
|--- truncated branch of depth 4
|--- type_of_meal_plan > 1.50
|--- avg_price_per_room <= 82.50
|--- avg_price_per_room <= 76.25
|--- weights: [1.00, 0.00] class: 0
|--- avg_price_per_room > 76.25
|--- weights: [0.00, 7.00] class: 1
|--- avg_price_per_room > 82.50
|--- weights: [12.00, 2.00] class: 0
--- avg_price_per_room > 93.58
|--- arrival_date <= 16.50
|--- arrival_month <= 7.50
|--- lead_time <= 108.50
|--- avg_price_per_room <= 125.00
|--- no_of_week_nights <= 0.50
|--- weights: [0.00, 1.00] class: 1
|--- no_of_week_nights > 0.50
|--- truncated branch of depth 4
|--- avg_price_per_room > 125.00
|--- weights: [0.00, 1.00] class: 1
|--- lead_time > 108.50
|--- weights: [18.00, 1.00] class: 0
|--- arrival_month > 7.50
|--- avg_price_per_room <= 108.50
|--- arrival_date <= 14.50
|--- lead_time <= 113.50
|--- weights: [4.00, 0.00] class: 0
|--- lead_time > 113.50
```

Project 4

```
| | | --- weights: [6.00, 9.00] class: 1
| | | --- arrival_date > 14.50
| | | | --- weights: [0.00, 47.00] class: 1
| | | --- avg_price_per_room > 108.50
| | | | --- no_of_weekend_nights <= 0.50
| | | | | --- weights: [42.00, 0.00] class: 0
| | | | --- no_of_weekend_nights > 0.50
| | | | | --- arrival_date <= 9.50
| | | | | | --- truncated branch of depth 3
| | | | | --- arrival_date > 9.50
| | | | | | --- truncated branch of depth 3
| | | | | | --- truncated branch of depth 3
--- arrival_date > 16.50
| --- arrival_month <= 8.50
| | --- avg_price_per_room <= 127.39
| | | --- no_of_adults <= 1.50
| | | | --- weights: [0.00, 50.00] class: 1
| | | --- no_of_adults > 1.50
| | | | --- no_of_week_nights <= 2.50
| | | | | --- truncated branch of depth 2
| | | | --- no_of_week_nights > 2.50
| | | | | | --- truncated branch of depth 4
| | | --- avg_price_per_room > 127.39
| | | | --- weights: [2.00, 0.00] class: 0
--- arrival_month > 8.50
| --- arrival_year <= 2017.50
| | --- weights: [0.00, 3.00] class: 1
| --- arrival_year > 2017.50
| | --- avg_price_per_room <= 101.34
| | | --- weights: [0.00, 3.00] class: 1
| | --- avg_price_per_room > 101.34
| | | --- avg_price_per_room <= 177.83
| | | | --- weights: [8.00, 0.00] class: 0
| | | --- avg_price_per_room > 177.83
| | | | --- weights: [0.00, 1.00] class: 1
--- lead_time > 117.50
| --- no_of_week_nights <= 1.50
| | --- arrival_date <= 7.50
| | | --- weights: [51.00, 0.00] class: 0
| | --- arrival_date > 7.50
| | | --- avg_price_per_room <= 93.58
| | | | --- avg_price_per_room <= 65.38
| | | | | --- weights: [0.00, 3.00] class: 1
| | | | --- avg_price_per_room > 65.38
| | | | | --- weights: [33.00, 2.00] class: 0
| | | --- avg_price_per_room > 93.58
| | | --- arrival_date <= 28.00
| | | | --- type_of_meal_plan <= 1.50
| | | | | --- no_of_adults <= 2.50
| | | | | | --- weights: [0.00, 17.00] class: 1
| | | | | --- no_of_adults > 2.50
| | | | | | --- weights: [1.00, 1.00] class: 0
| | | | --- type_of_meal_plan > 1.50
| | | | | --- arrival_year <= 2017.50
| | | | | | --- weights: [18.00, 30.00] class: 1
| | | | | --- arrival_year > 2017.50
| | | | | | --- weights: [1.00, 0.00] class: 0
| | | | --- arrival_date > 28.00
| | | | | | --- weights: [13.00, 1.00] class: 0
--- no_of_week_nights > 1.50
| --- no_of_adults <= 1.50
```

```

|--- weights: [113.00, 0.00] class: 0
|--- no_of_adults > 1.50
|--- lead_time <= 125.50
|   |--- avg_price_per_room <= 90.85
|   |   |--- avg_price_per_room <= 87.50
|   |   |   |--- no_of_week_nights <= 3.50
|   |   |   |   |--- truncated branch of depth 3
|   |   |   |   |--- no_of_week_nights > 3.50
|   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |--- avg_price_per_room > 87.50
|   |   |   |   |   |--- weights: [0.00, 10.00] class: 1
|   |   |--- avg_price_per_room > 90.85
|   |   |   |--- weights: [14.00, 0.00] class: 0
|--- lead_time > 125.50
|   |--- avg_price_per_room <= 216.00
|   |   |--- arrival_date <= 19.50
|   |   |   |--- arrival_date <= 10.50
|   |   |   |   |--- truncated branch of depth 5
|   |   |   |   |--- arrival_date > 10.50
|   |   |   |   |   |--- truncated branch of depth 9
|   |   |   |   |--- arrival_date > 19.50
|   |   |   |   |   |--- weights: [83.00, 1.00] class: 0
|   |   |   |   |--- avg_price_per_room > 216.00
|   |   |   |   |   |--- weights: [0.00, 1.00] class: 1
--- no_of_special_requests > 0.50
--- no_of_special_requests <= 1.50
|--- market_segment_type <= 1.50
|--- lead_time <= 8.50
|--- lead_time <= 4.50
|   |--- no_of_week_nights <= 10.00
|   |   |--- avg_price_per_room <= 219.86
|   |   |   |--- avg_price_per_room <= 157.64
|   |   |   |   |--- arrival_date <= 4.50
|   |   |   |   |   |--- weights: [82.00, 0.00] class: 0
|   |   |   |   |--- arrival_date > 4.50
|   |   |   |   |   |--- arrival_date <= 27.50
|   |   |   |   |   |   |--- truncated branch of depth 10
|   |   |   |   |   |--- arrival_date > 27.50
|   |   |   |   |   |   |--- weights: [69.00, 0.00] class: 0
|--- avg_price_per_room > 157.64
|   |--- avg_price_per_room <= 158.50
|   |   |--- weights: [0.00, 1.00] class: 1
|--- avg_price_per_room > 158.50
|   |--- no_of_week_nights <= 3.50
|   |   |--- truncated branch of depth 6
|   |   |--- no_of_week_nights > 3.50
|   |   |   |--- truncated branch of depth 3
|--- avg_price_per_room > 219.86
|--- avg_price_per_room <= 223.58
|   |--- weights: [0.00, 1.00] class: 1
|--- avg_price_per_room > 223.58
|   |--- avg_price_per_room <= 237.25
|   |   |--- weights: [5.00, 0.00] class: 0
|   |--- avg_price_per_room > 237.25
|   |   |--- arrival_date <= 11.50
|   |   |   |--- weights: [3.00, 0.00] class: 0
|   |   |   |--- arrival_date > 11.50
|   |   |   |   |--- truncated branch of depth 2
|--- no_of_week_nights > 10.00
|   |--- weights: [0.00, 2.00] class: 1

```

```

    |--- lead_time > 4.50
    |--- arrival_date <= 13.50
    |--- arrival_month <= 9.50
    |--- arrival_year <= 2017.50
    |--- arrival_date <= 11.00
    |--- avg_price_per_room <= 119.50
    |--- weights: [4.00, 0.00] class: 0
    |--- avg_price_per_room > 119.50
    |--- truncated branch of depth 2
    |--- arrival_date > 11.00
    |--- weights: [0.00, 3.00] class: 1
    --- arrival_year > 2017.50
    |--- avg_price_per_room <= 185.50
    |--- no_of_adults <= 0.50
    |--- weights: [0.00, 1.00] class: 1
    |--- no_of_adults > 0.50
    |--- truncated branch of depth 8
    |--- avg_price_per_room > 185.50
    |--- arrival_month <= 8.50
    |--- weights: [0.00, 2.00] class: 1
    |--- arrival_month > 8.50
    |--- weights: [1.00, 0.00] class: 0
    --- arrival_month > 9.50
    |--- avg_price_per_room <= 157.12
    |--- weights: [43.00, 0.00] class: 0
    --- avg_price_per_room > 157.12
    |--- room_type_reserved <= 2.50
    |--- weights: [0.00, 1.00] class: 1
    |--- room_type_reserved > 2.50
    |--- weights: [2.00, 0.00] class: 0
    --- arrival_date > 13.50
    |--- type_of_meal_plan <= 0.50
    |--- avg_price_per_room <= 126.33
    |--- arrival_date <= 21.50
    |--- weights: [24.00, 0.00] class: 0
    --- arrival_date > 21.50
    |--- arrival_date <= 23.50
    |--- truncated branch of depth 3
    |--- arrival_date > 23.50
    |--- weights: [17.00, 0.00] class: 0
    --- avg_price_per_room > 126.33
    |--- arrival_month <= 8.50
    |--- arrival_date <= 26.00
    |--- weights: [7.00, 0.00] class: 0
    --- arrival_date > 26.00
    |--- truncated branch of depth 2
    |--- arrival_month > 8.50
    |--- lead_time <= 6.50
    |--- truncated branch of depth 2
    |--- lead_time > 6.50
    |--- weights: [1.00, 6.00] class: 1
    --- type_of_meal_plan > 0.50
    |--- avg_price_per_room <= 139.57
    |--- no_of_week_nights <= 3.50
    |--- arrival_month <= 2.50
    |--- truncated branch of depth 3
    |--- arrival_month > 2.50
    |--- weights: [97.00, 0.00] class: 0
    |--- no_of_week_nights > 3.50
    |--- arrival_date <= 23.50

```

Project 4

```

|--- weights: [10.00, 0.00] class: 0
|--- arrival_date > 23.50
|--- truncated branch of depth 2
--- avg_price_per_room > 139.57
|--- arrival_date <= 15.50
|--- arrival_date <= 14.50
|--- weights: [2.00, 0.00] class: 0
|--- arrival_date > 14.50
|--- truncated branch of depth 3
|--- arrival_date > 15.50
|--- avg_price_per_room <= 140.50
|--- truncated branch of depth 2
|--- avg_price_per_room > 140.50
|--- truncated branch of depth 3
--- lead_time > 8.50
--- required_car_parking_space <= 0.50
|--- avg_price_per_room <= 127.62
|--- no_of_weekend_nights <= 2.50
|--- lead_time <= 43.50
|--- arrival_month <= 11.50
|--- arrival_month <= 1.50
|--- weights: [87.00, 0.00] class: 0
|--- arrival_month > 1.50
|--- truncated branch of depth 23
|--- arrival_month > 11.50
|--- weights: [128.00, 0.00] class: 0
|--- lead_time > 43.50
|--- arrival_year <= 2017.50
|--- arrival_month <= 7.50
|--- truncated branch of depth 4
|--- arrival_month > 7.50
|--- truncated branch of depth 10
|--- arrival_year > 2017.50
|--- arrival_month <= 8.50
|--- truncated branch of depth 20
|--- arrival_month > 8.50
|--- truncated branch of depth 21
|--- no_of_weekend_nights > 2.50
|--- avg_price_per_room <= 119.12
|--- arrival_month <= 1.50
|--- weights: [3.00, 0.00] class: 0
|--- arrival_month > 1.50
|--- no_of_week_nights <= 8.50
|--- truncated branch of depth 4
|--- no_of_week_nights > 8.50
|--- weights: [0.00, 12.00] class: 1
|--- avg_price_per_room > 119.12
|--- weights: [3.00, 0.00] class: 0
--- avg_price_per_room > 127.62
|--- lead_time <= 142.50
|--- arrival_month <= 8.50
|--- arrival_date <= 19.50
|--- avg_price_per_room <= 177.15
|--- truncated branch of depth 15
|--- avg_price_per_room > 177.15
|--- truncated branch of depth 6
|--- arrival_date > 19.50
|--- arrival_date <= 27.50
|--- truncated branch of depth 11
|--- arrival_date > 27.50

```

Project 4

```

| | | | | --- truncated branch of depth 10
| | | | | --- arrival_month > 8.50
| | | | | --- arrival_month <= 11.50
| | | | | --- arrival_year <= 2017.50
| | | | | --- truncated branch of depth 7
| | | | | --- arrival_year > 2017.50
| | | | | --- truncated branch of depth 21
| | | | | --- arrival_month > 11.50
| | | | | --- lead_time <= 100.50
| | | | | --- weights: [49.00, 0.00] class: 0
| | | | | --- lead_time > 100.50
| | | | | --- truncated branch of depth 2
| | | | | --- lead_time > 142.50
| | | | | --- avg_price_per_room <= 142.65
| | | | | --- no_of_week_nights <= 2.50
| | | | | --- weights: [5.00, 0.00] class: 0
| | | | | --- no_of_week_nights > 2.50
| | | | | --- no_of_week_nights <= 3.50
| | | | | --- weights: [0.00, 2.00] class: 1
| | | | | --- no_of_week_nights > 3.50
| | | | | --- weights: [1.00, 0.00] class: 0
| | | | | --- avg_price_per_room > 142.65
| | | | | --- avg_price_per_room <= 182.49
| | | | | --- weights: [0.00, 11.00] class: 1
| | | | | --- avg_price_per_room > 182.49
| | | | | --- arrival_month <= 7.50
| | | | | --- weights: [0.00, 1.00] class: 1
| | | | | --- arrival_month > 7.50
| | | | | --- weights: [1.00, 0.00] class: 0
| | | | | --- required_car_parking_space > 0.50
| | | | | --- no_of_week_nights <= 7.50
| | | | | --- weights: [180.00, 0.00] class: 0
| | | | | --- no_of_week_nights > 7.50
| | | | | --- weights: [0.00, 1.00] class: 1
| | | | | --- market_segment_type > 1.50
| | | | | --- type_of_meal_plan <= 0.50
| | | | | --- lead_time <= 63.00
| | | | | --- weights: [21.00, 1.00] class: 0
| | | | | --- lead_time > 63.00
| | | | | --- weights: [0.00, 6.00] class: 1
| | | | | --- type_of_meal_plan > 0.50
| | | | | --- lead_time <= 102.50
| | | | | --- no_of_week_nights <= 11.00
| | | | | --- avg_price_per_room <= 141.50
| | | | | --- lead_time <= 91.50
| | | | | | --- avg_price_per_room <= 129.50
| | | | | | --- weights: [856.00, 0.00] class: 0
| | | | | | --- avg_price_per_room > 129.50
| | | | | | --- avg_price_per_room <= 131.50
| | | | | | --- truncated branch of depth 2
| | | | | | --- avg_price_per_room > 131.50
| | | | | | --- weights: [15.00, 0.00] class: 0
| | | | | --- lead_time > 91.50
| | | | | --- no_of_children <= 0.50
| | | | | --- weights: [43.00, 0.00] class: 0
| | | | | --- no_of_children > 0.50
| | | | | --- arrival_date <= 16.50
| | | | | | --- weights: [0.00, 2.00] class: 1
| | | | | --- arrival_date > 16.50
| | | | | | --- weights: [2.00, 0.00] class: 0

```

```

|--- avg_price_per_room > 141.50
|   |--- avg_price_per_room <= 144.00
|   |   |--- weights: [0.00, 1.00] class: 1
|   |--- avg_price_per_room > 144.00
|   |   |--- market_segment_type <= 2.50
|   |   |   |--- weights: [15.00, 0.00] class: 0
|   |   |--- market_segment_type > 2.50
|   |   |   |--- arrival_date <= 15.50
|   |   |   |   |--- weights: [2.00, 0.00] class: 0
|   |   |   |--- arrival_date > 15.50
|   |   |   |   |--- weights: [0.00, 1.00] class: 1
|--- no_of_week_nights > 11.00
|   |--- weights: [0.00, 1.00] class: 1
--- lead_time > 102.50
|--- lead_time <= 104.50
|   |--- no_of_week_nights <= 2.50
|   |   |--- avg_price_per_room <= 67.65
|   |   |   |--- weights: [1.00, 0.00] class: 0
|   |   |--- avg_price_per_room > 67.65
|   |   |   |--- weights: [0.00, 4.00] class: 1
|--- no_of_week_nights > 2.50
|   |--- weights: [4.00, 0.00] class: 0
--- lead_time > 104.50
|--- avg_price_per_room <= 141.75
|   |--- no_of_week_nights <= 2.50
|   |   |--- avg_price_per_room <= 83.39
|   |   |   |--- arrival_month <= 3.50
|   |   |   |   |--- weights: [6.00, 0.00] class: 0
|   |   |   |--- arrival_month > 3.50
|   |   |   |   |--- truncated branch of depth 4
|   |   |--- avg_price_per_room > 83.39
|   |   |   |--- lead_time <= 143.50
|   |   |   |   |--- weights: [29.00, 1.00] class: 0
|   |   |   |--- lead_time > 143.50
|   |   |   |   |--- truncated branch of depth 2
|--- no_of_week_nights > 2.50
|   |--- avg_price_per_room <= 122.00
|   |   |--- weights: [54.00, 0.00] class: 0
|   |--- avg_price_per_room > 122.00
|   |   |--- avg_price_per_room <= 131.75
|   |   |   |--- weights: [0.00, 1.00] class: 1
|   |   |--- avg_price_per_room > 131.75
|   |   |   |--- weights: [1.00, 0.00] class: 0
|--- avg_price_per_room > 141.75
|   |--- type_of_meal_plan <= 1.50
|   |   |--- weights: [0.00, 2.00] class: 1
|   |--- type_of_meal_plan > 1.50
|   |   |--- weights: [1.00, 0.00] class: 0
--- no_of_special_requests > 1.50
|--- lead_time <= 90.50
|   |--- no_of_week_nights <= 3.50
|   |   |--- weights: [2126.00, 0.00] class: 0
|--- no_of_week_nights > 3.50
|   |--- no_of_week_nights <= 9.50
|   |   |--- no_of_special_requests <= 2.50
|   |   |   |--- lead_time <= 6.50
|   |   |   |   |--- weights: [43.00, 0.00] class: 0
|   |   |   |--- lead_time > 6.50
|   |   |   |   |--- arrival_month <= 11.50
|   |   |   |   |   |--- avg_price_per_room <= 93.09

```

Project 4

```

|--- lead_time <= 54.00
|   |--- truncated branch of depth 8
|--- lead_time > 54.00
|   |--- truncated branch of depth 5
|--- avg_price_per_room > 93.09
|   |--- avg_price_per_room <= 107.29
|       |--- truncated branch of depth 3
|       |--- avg_price_per_room > 107.29
|           |--- truncated branch of depth 13
|           |--- arrival_month > 11.50
|               |--- weights: [26.00, 0.00] class: 0
|--- no_of_special_requests > 2.50
|   |--- weights: [70.00, 0.00] class: 0
|--- no_of_week_nights > 9.50
|   |--- weights: [0.00, 2.00] class: 1
|--- lead_time > 90.50
|--- avg_price_per_room <= 202.95
|   |--- arrival_month <= 8.50
|       |--- arrival_year <= 2017.50
|           |--- arrival_month <= 7.50
|               |--- arrival_date <= 4.50
|                   |--- weights: [1.00, 0.00] class: 0
|               |--- arrival_date > 4.50
|                   |--- arrival_date <= 26.00
|                       |--- weights: [0.00, 5.00] class: 1
|                   |--- arrival_date > 26.00
|                       |--- avg_price_per_room <= 91.62
|                           |--- weights: [1.00, 0.00] class: 0
|                       |--- avg_price_per_room > 91.62
|                           |--- weights: [0.00, 1.00] class: 1
|               |--- arrival_month > 7.50
|                   |--- arrival_date <= 24.50
|                       |--- lead_time <= 98.50
|                           |--- arrival_date <= 10.50
|                               |--- weights: [1.00, 0.00] class: 0
|                           |--- arrival_date > 10.50
|                               |--- weights: [0.00, 1.00] class: 1
|                       |--- lead_time > 98.50
|                           |--- weights: [11.00, 0.00] class: 0
|                   |--- arrival_date > 24.50
|                       |--- weights: [0.00, 1.00] class: 1
|--- arrival_year > 2017.50
|--- lead_time <= 150.50
|   |--- no_of_children <= 0.50
|       |--- arrival_date <= 29.50
|           |--- avg_price_per_room <= 157.65
|               |--- truncated branch of depth 9
|               |--- avg_price_per_room > 157.65
|                   |--- truncated branch of depth 3
|           |--- arrival_date > 29.50
|               |--- no_of_week_nights <= 2.50
|                   |--- weights: [2.00, 0.00] class: 0
|               |--- no_of_week_nights > 2.50
|                   |--- truncated branch of depth 2
|   |--- no_of_children > 0.50
|       |--- arrival_month <= 4.50
|           |--- no_of_weekend_nights <= 1.00
|               |--- weights: [2.00, 0.00] class: 0
|           |--- no_of_weekend_nights > 1.00
|               |--- weights: [0.00, 2.00] class: 1

```

Project 4

```

    |--- arrival_month > 4.50
    |   |--- no_of_week_nights <= 4.50
    |   |   |--- truncated branch of depth 4
    |   |   |--- no_of_week_nights > 4.50
    |   |   |   |--- truncated branch of depth 3
    |   |--- lead_time > 150.50
    |   |   |--- weights: [0.00, 3.00] class: 1
--- arrival_month > 8.50
    |--- no_of_special_requests <= 2.50
    |--- avg_price_per_room <= 90.42
    |   |--- arrival_month <= 11.50
    |   |   |--- lead_time <= 107.00
    |   |   |   |--- room_type_reserved <= 1.50
    |   |   |   |   |--- truncated branch of depth 4
    |   |   |   |--- room_type_reserved > 1.50
    |   |   |   |   |--- weights: [4.00, 0.00] class: 0
    |--- lead_time > 107.00
    |   |--- arrival_date <= 17.00
    |   |   |--- truncated branch of depth 3
    |   |--- arrival_date > 17.00
    |   |   |   |--- weights: [7.00, 0.00] class: 0
    |--- arrival_month > 11.50
    |   |--- lead_time <= 101.00
    |   |   |--- weights: [11.00, 0.00] class: 0
    |--- lead_time > 101.00
    |   |--- arrival_date <= 7.50
    |   |   |--- weights: [0.00, 2.00] class: 1
    |   |--- arrival_date > 7.50
    |   |   |   |--- truncated branch of depth 5
    |--- avg_price_per_room > 90.42
    |   |--- no_of_adults <= 1.50
    |   |   |--- weights: [11.00, 0.00] class: 0
    |--- no_of_adults > 1.50
    |   |--- avg_price_per_room <= 153.15
    |   |   |--- no_of_weekend_nights <= 1.50
    |   |   |   |--- truncated branch of depth 8
    |   |   |   |--- no_of_weekend_nights > 1.50
    |   |   |   |   |--- truncated branch of depth 6
    |   |   |--- avg_price_per_room > 153.15
    |   |   |   |--- arrival_date <= 26.50
    |   |   |   |   |--- truncated branch of depth 3
    |   |   |   |--- arrival_date > 26.50
    |   |   |   |   |--- truncated branch of depth 2
    |--- no_of_special_requests > 2.50
    |   |--- weights: [52.00, 0.00] class: 0
--- avg_price_per_room > 202.95
    |--- weights: [0.00, 7.00] class: 1
--- lead_time > 151.50
    |--- avg_price_per_room <= 100.04
    |   |--- no_of_special_requests <= 0.50
    |   |--- market_segment_type <= 1.50
    |   |   |--- avg_price_per_room <= 2.50
    |   |   |   |--- no_of_weekend_nights <= 0.50
    |   |   |   |--- lead_time <= 205.00
    |   |   |   |   |--- weights: [2.00, 0.00] class: 0
    |--- lead_time > 205.00
    |   |--- arrival_date <= 19.00
    |   |   |--- weights: [0.00, 4.00] class: 1
    |--- arrival_date > 19.00
    |   |--- weights: [1.00, 0.00] class: 0

```

```

--- no_of_weekend_nights > 0.50
    |--- weights: [9.00, 0.00] class: 0
--- avg_price_per_room > 2.50
    |--- arrival_month <= 11.50
        |--- weights: [0.00, 525.00] class: 1
    --- arrival_month > 11.50
        |--- no_of_weekend_nights <= 0.50
            |--- lead_time <= 263.50
                |--- avg_price_per_room <= 76.87
                    |--- no_of_week_nights <= 1.50
                        |--- weights: [1.00, 0.00] class: 0
                    |--- no_of_week_nights > 1.50
                        |--- weights: [0.00, 5.00] class: 1
                |--- avg_price_per_room > 76.87
                    |--- weights: [3.00, 0.00] class: 0
            |--- lead_time > 263.50
                |--- weights: [0.00, 7.00] class: 1
        --- no_of_weekend_nights > 0.50
            |--- no_of_week_nights <= 1.50
                |--- arrival_date <= 3.50
                    |--- weights: [1.00, 0.00] class: 0
                |--- arrival_date > 3.50
                    |--- weights: [0.00, 6.00] class: 1
            |--- no_of_week_nights > 1.50
                |--- weights: [0.00, 58.00] class: 1
--- market_segment_type > 1.50
    |--- no_of_adults <= 1.50
        |--- lead_time <= 163.50
            |--- arrival_date <= 7.00
                |--- weights: [0.00, 15.00] class: 1
        |--- arrival_date > 7.00
            |--- lead_time <= 160.50
                |--- weights: [4.00, 0.00] class: 0
            |--- lead_time > 160.50
                |--- weights: [1.00, 1.00] class: 0
    |--- lead_time > 163.50
        |--- lead_time <= 341.00
            |--- lead_time <= 173.00
                |--- arrival_date <= 3.50
                    |--- weights: [63.00, 6.00] class: 0
                |--- arrival_date > 3.50
                    |--- arrival_month <= 5.00
                        |--- weights: [3.00, 0.00] class: 0
                    |--- arrival_month > 5.00
                        |--- weights: [0.00, 9.00] class: 1
            |--- lead_time > 173.00
                |--- arrival_month <= 5.50
                    |--- avg_price_per_room <= 88.00
                        |--- weights: [9.00, 0.00] class: 0
                    |--- avg_price_per_room > 88.00
                        |--- weights: [0.00, 3.00] class: 1
            |--- arrival_month > 5.50
                |--- avg_price_per_room <= 98.00
                    |--- avg_price_per_room <= 55.21
                        |--- truncated branch of depth 2
                    |--- avg_price_per_room > 55.21
                        |--- truncated branch of depth 6
                |--- avg_price_per_room > 98.00
                    |--- no_of_weekend_nights <= 0.50
                        |--- weights: [1.00, 0.00] class: 0

```

```

|--- no_of_weekend_nights > 0.50
|   |--- weights: [0.00, 1.00] class: 1
--- lead_time > 341.00
|--- arrival_date <= 8.50
|   |--- avg_price_per_room <= 88.33
|   |   |--- weights: [0.00, 7.00] class: 1
|   |--- avg_price_per_room > 88.33
|   |   |--- weights: [1.00, 1.00] class: 0
|--- arrival_date > 8.50
|   |--- arrival_date <= 24.50
|       |--- avg_price_per_room <= 80.00
|       |   |--- weights: [5.00, 0.00] class: 0
|       |--- avg_price_per_room > 80.00
|       |   |--- weights: [12.00, 7.00] class: 0
|       |--- arrival_date > 24.50
|           |--- weights: [0.00, 3.00] class: 1
--- no_of_adults > 1.50
|--- avg_price_per_room <= 84.58
|--- lead_time <= 244.00
|   |--- no_of_week_nights <= 1.50
|       |--- no_of_weekend_nights <= 1.50
|           |--- arrival_date <= 19.00
|               |--- lead_time <= 166.50
|                   |--- weights: [2.00, 0.00] class: 0
|               |--- lead_time > 166.50
|                   |--- truncated branch of depth 2
|               |--- arrival_date > 19.00
|                   |--- weights: [3.00, 0.00] class: 0
|   |--- no_of_weekend_nights > 1.50
|       |--- weights: [24.00, 0.00] class: 0
|--- no_of_week_nights > 1.50
|   |--- avg_price_per_room <= 66.50
|       |--- no_of_weekend_nights <= 0.50
|           |--- avg_price_per_room <= 64.80
|               |--- weights: [2.00, 0.00] class: 0
|           |--- avg_price_per_room > 64.80
|               |--- weights: [0.00, 7.00] class: 1
|       |--- no_of_weekend_nights > 0.50
|           |--- no_of_week_nights <= 5.50
|               |--- truncated branch of depth 2
|           |--- no_of_week_nights > 5.50
|               |--- weights: [0.00, 1.00] class: 1
|   |--- avg_price_per_room > 66.50
|       |--- type_of_meal_plan <= 1.50
|           |--- avg_price_per_room <= 75.75
|               |--- weights: [85.00, 1.00] class: 0
|           |--- avg_price_per_room > 75.75
|               |--- truncated branch of depth 6
|       |--- type_of_meal_plan > 1.50
|           |--- weights: [0.00, 1.00] class: 1
--- lead_time > 244.00
|--- arrival_month <= 11.50
|   |--- arrival_year <= 2017.50
|       |--- weights: [34.00, 0.00] class: 0
|   |--- arrival_year > 2017.50
|       |--- avg_price_per_room <= 80.38
|           |--- no_of_week_nights <= 3.50
|               |--- truncated branch of depth 6
|           |--- no_of_week_nights > 3.50
|               |--- truncated branch of depth 4

```

Project 4

Project 4

```
|--- weights: [0.00, 3.00] class: 1
|--- lead_time > 226.50
|--- lead_time <= 272.00
|--- weights: [6.00, 0.00] class: 0
|--- lead_time > 272.00
|--- truncated branch of depth 3
|--- lead_time > 300.50
|--- weights: [0.00, 5.00] class: 1
--- no_of_special_requests > 2.50
|--- weights: [12.00, 0.00] class: 0
--- market_segment_type > 1.50
|--- no_of_adults <= 2.50
|--- lead_time <= 356.00
|--- weights: [17.00, 1.00] class: 0
|--- lead_time > 356.00
|--- weights: [0.00, 1.00] class: 1
|--- no_of_adults > 2.50
|--- weights: [0.00, 2.00] class: 1
--- no_of_weekend_nights > 0.50
|--- market_segment_type <= 1.50
|--- no_of_week_nights <= 9.50
|--- arrival_month <= 11.50
|--- arrival_date <= 27.50
|--- avg_price_per_room <= 81.12
|--- lead_time <= 153.50
|--- no_of_week_nights <= 2.50
|--- weights: [1.00, 0.00] class: 0
|--- no_of_week_nights > 2.50
|--- weights: [0.00, 2.00] class: 1
|--- lead_time > 153.50
|--- lead_time <= 157.50
|--- truncated branch of depth 2
|--- lead_time > 157.50
|--- truncated branch of depth 3
|--- avg_price_per_room > 81.12
|--- no_of_week_nights <= 6.50
|--- lead_time <= 233.00
|--- truncated branch of depth 13
|--- lead_time > 233.00
|--- truncated branch of depth 9
|--- no_of_week_nights > 6.50
|--- lead_time <= 204.50
|--- truncated branch of depth 2
|--- lead_time > 204.50
|--- weights: [0.00, 3.00] class: 1
--- arrival_date > 27.50
|--- no_of_week_nights <= 1.50
|--- lead_time <= 224.50
|--- lead_time <= 175.50
|--- weights: [1.00, 0.00] class: 0
|--- lead_time > 175.50
|--- weights: [0.00, 10.00] class: 1
|--- lead_time > 224.50
|--- weights: [4.00, 0.00] class: 0
|--- no_of_week_nights > 1.50
|--- lead_time <= 269.00
|--- lead_time <= 176.00
|--- truncated branch of depth 4
|--- lead_time > 176.00
|--- truncated branch of depth 4
```

```

|   |   |   |   |   |--- lead_time > 269.00
|   |   |   |   |   |--- weights: [0.00, 3.00] class: 1
|   |   |   |--- arrival_month > 11.50
|   |   |   |   |--- arrival_date <= 14.50
|   |   |   |   |   |--- arrival_date <= 3.00
|   |   |   |   |   |--- weights: [0.00, 1.00] class: 1
|   |   |   |--- arrival_date > 3.00
|   |   |   |   |--- lead_time <= 217.50
|   |   |   |   |   |--- weights: [8.00, 0.00] class: 0
|   |   |   |--- lead_time > 217.50
|   |   |   |   |--- no_of_weekend_nights <= 1.50
|   |   |   |   |   |--- weights: [3.00, 0.00] class: 0
|   |   |   |   |--- no_of_weekend_nights > 1.50
|   |   |   |   |   |--- weights: [0.00, 1.00] class: 1
|   |   |--- arrival_date > 14.50
|   |   |   |--- no_of_week_nights <= 2.50
|   |   |   |   |--- no_of_special_requests <= 1.50
|   |   |   |   |   |--- weights: [0.00, 8.00] class: 1
|   |   |   |   |--- no_of_special_requests > 1.50
|   |   |   |   |   |--- arrival_date <= 19.50
|   |   |   |   |   |   |--- weights: [0.00, 1.00] class: 1
|   |   |   |   |   |--- arrival_date > 19.50
|   |   |   |   |   |   |--- weights: [2.00, 0.00] class: 0
|   |   |--- no_of_week_nights > 2.50
|   |   |   |--- no_of_week_nights <= 4.50
|   |   |   |   |--- lead_time <= 281.50
|   |   |   |   |   |--- weights: [8.00, 0.00] class: 0
|   |   |   |   |--- lead_time > 281.50
|   |   |   |   |   |--- truncated branch of depth 4
|   |   |--- no_of_week_nights > 4.50
|   |   |   |   |   |--- no_of_special_requests <= 2.50
|   |   |   |   |   |   |--- truncated branch of depth 2
|   |   |   |   |   |--- no_of_special_requests > 2.50
|   |   |   |   |   |   |--- weights: [2.00, 0.00] class: 0
|--- no_of_week_nights > 9.50
|--- lead_time <= 198.00
|   |--- weights: [1.00, 0.00] class: 0
|--- lead_time > 198.00
|   |--- weights: [0.00, 7.00] class: 1
|--- market_segment_type > 1.50
|--- lead_time <= 348.50
|   |--- no_of_week_nights <= 5.50
|   |   |--- weights: [142.00, 1.00] class: 0
|--- no_of_week_nights > 5.50
|   |--- avg_price_per_room <= 76.75
|   |   |--- weights: [1.00, 0.00] class: 0
|   |--- avg_price_per_room > 76.75
|   |   |--- weights: [0.00, 1.00] class: 1
|--- lead_time > 348.50
|   |--- weights: [8.00, 3.00] class: 0
|--- avg_price_per_room > 100.04
|--- arrival_month <= 11.50
|   |--- no_of_special_requests <= 2.50
|   |   |--- weights: [0.00, 2108.00] class: 1
|--- no_of_special_requests > 2.50
|   |--- weights: [31.00, 0.00] class: 0
|--- arrival_month > 11.50
|--- no_of_special_requests <= 0.50
|   |--- weights: [47.00, 0.00] class: 0
|--- no_of_special_requests > 0.50

```

```

    |   |   |   |   |   --- arrival_date <= 24.50
    |   |   |   |   |   |--- weights: [5.00, 0.00] class: 0
    |   |   |   |   --- arrival_date > 24.50
    |   |   |   |   |--- room_type_reserved <= 2.50
    |   |   |   |   |   |--- no_of_week_nights <= 3.00
    |   |   |   |   |   |--- weights: [0.00, 1.00] class: 1
    |   |   |   |   |   --- no_of_week_nights > 3.00
    |   |   |   |   |   |--- weights: [3.00, 0.00] class: 0
    |   |   |   |   --- room_type_reserved > 2.50
    |   |   |   |   |--- lead_time <= 172.50
    |   |   |   |   |   |--- no_of_adults <= 2.50
    |   |   |   |   |   |   |--- weights: [0.00, 1.00] class: 1
    |   |   |   |   |   |--- no_of_adults > 2.50
    |   |   |   |   |   |   |--- weights: [2.00, 0.00] class: 0
    |   |   |   |   |--- lead_time > 172.50
    |   |   |   |   |   |--- weights: [0.00, 13.00] class: 1

```

In [112...]

```
print(pd.DataFrame(best_model.feature_importances_, columns=["Imp"], index=X_train.columns))
```

	Imp
lead_time	0.352079
avg_price_per_room	0.174046
market_segment_type	0.097631
arrival_date	0.088920
no_of_special_requests	0.068086
arrival_month	0.064782
no_of_week_nights	0.042594
no_of_weekend_nights	0.039044
no_of_adults	0.026149
arrival_year	0.011808
type_of_meal_plan	0.011506
room_type_reserved	0.011142
required_car_parking_space	0.007128
no_of_children	0.004282
repeated_guest	0.000591
no_of_previous_bookings_not_canceled	0.000213
no_of_previous_cancellations	0.000000

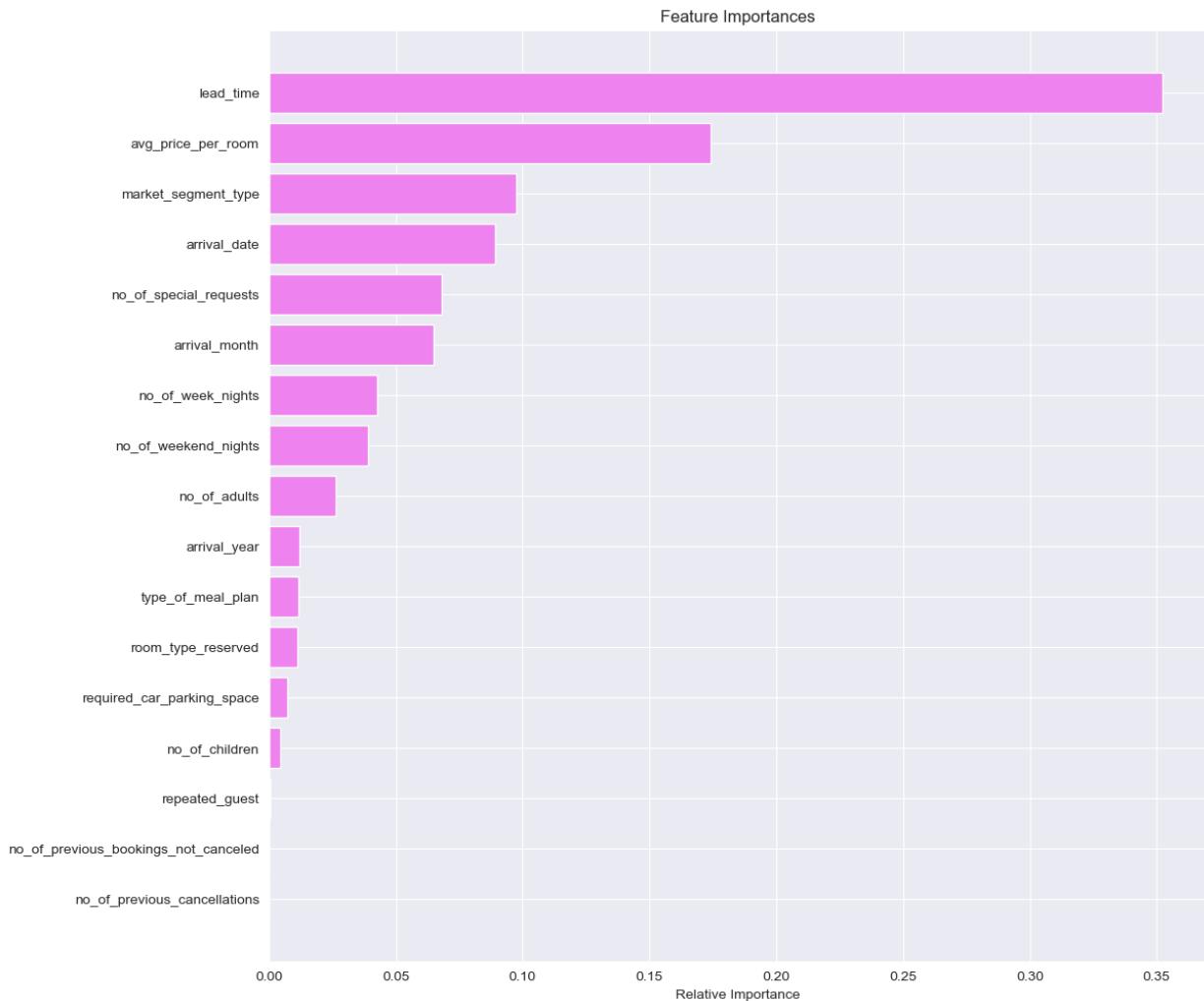
In [113...]

```

importances = best_model.feature_importances_
indices = np.argsort(importances)

plt.figure(figsize=(12, 12))
plt.title("Feature Importances")
plt.barh(range(len(indices)), importances[indices], color="violet", align="center")
plt.yticks(range(len(indices)), [feature_names[i] for i in indices])
plt.xlabel("Relative Importance")
plt.show()

```



We see the top three final factors are lead time, average price of room, and the market segment type.

In [114...]

```
models_train_comp_df = pd.concat(
    [
        decision_tree_perf_train.T,
        decision_tree_tune_perf_train.T,
        decision_tree_postpruned_perf_train.T,
    ],
    axis=1,
)
models_train_comp_df.columns = [
    "Decision Tree sklearn",
    "Decision Tree (Pre-Pruning)",
    "Decision Tree (Post-Pruning)",
]
print("Training performance comparison:")
models_train_comp_df
```

Training performance comparison:

Out[114]:

	Decision Tree sklearn	Decision Tree (Pre-Pruning)	Decision Tree (Post-Pruning)
Accuracy	0.994211	0.994211	0.993266
Recall	0.986608	0.986608	0.985532
Precision	0.995776	0.995776	0.993970
F1	0.991171	0.991171	0.989733

In [115...]

```
models_train_comp_df = pd.concat(
    [
        decision_tree_perf_test.T,
        decision_tree_tune_perf_test.T,
        decision_tree_postpruned_perf_test.T,
    ],
    axis=1,
)
models_train_comp_df.columns = [
    "Decision Tree sklearn",
    "Decision Tree (Pre-Pruning)",
    "Decision Tree (Post-Pruning)",
]
print("Test set performance comparison:")
models_train_comp_df
```

Test set performance comparison:

Out[115]:

	Decision Tree sklearn	Decision Tree (Pre-Pruning)	Decision Tree (Post-Pruning)
Accuracy	0.871267	0.874851	0.871451
Recall	0.808915	0.812890	0.809483
Precision	0.796477	0.802860	0.796591
F1	0.802648	0.807844	0.802985

The final model true is displaying high accuracy

Both the Logistical regression model and the desion tree were very accurate in there predictions. Lead time is by far the biggest driving factor in cancellations. My reccomendation would be to try and reduce the average number of days between booking and arriving. You could also slightly over sell rooms far in advanced and then wait for cancellations leaving you with less empty rooms.