Hotel Reservations

By

Tasneem Alnaasan

The dataset is about Hotel Reservation. There are 17 columns; BookingID, Hotel, BookingDate, ArrivalDate, LeadTime, Nights, Guests, DistributionChannel, CustomerType, Country, DepositeType, AvgDailyRate, Status, StatusUpdate, Cancelled, Revenue, RevenueLoss.

The dataset for the years 2015, 2016, 2017. The analysis will focus on the year 2016.

The analysis focuses on:

1- Total Revenue and Average daily rate

2-Cancellation Rate

3-Room Status

4- Booking Channels

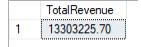5-Hotels and Customer types.

1. Showing Dataset

```sql
select *
from HotelReservation..Hotel
```

	BookingID	Hotel	BookingDate	ArrivalDate	LeadTime	Nights	Guests	DistributionChannel	CustomerType	Country	DepositType	AvgDailyRate	Status	StatusUpdate	Cancel...	Revenue	RevenueLoss
1	1	Resort	2014-07-2...	2015-07-01 0...	342	0	2	Direct	Transient	Portugal	No Deposit	0.00	Check_Out	2015-07-01 00:...	0	0.00	0.00
2	2	Resort	2013-06-2...	2015-07-01 0...	737	0	2	Direct	Transient	Portugal	No Deposit	0.00	Check_Out	2015-07-01 00:...	0	0.00	0.00
3	3	Resort	2015-06-2...	2015-07-01 0...	7	1	1	Direct	Transient	United Kingdom	No Deposit	75.00	Check_Out	2015-07-02 00:...	0	75.00	0.00
4	4	Resort	2015-06-1...	2015-07-01 0...	13	1	1	Corporate	Transient	United Kingdom	No Deposit	75.00	Check_Out	2015-07-02 00:...	0	75.00	0.00
5	5	Resort	2015-06-1...	2015-07-01 0...	14	2	2	Online Travel Agent	Transient	United Kingdom	No Deposit	98.00	Check_Out	2015-07-03 00:...	0	196.00	0.00
6	6	Resort	2015-06-1...	2015-07-01 0...	14	2	2	Online Travel Agent	Transient	United Kingdom	No Deposit	98.00	Check_Out	2015-07-03 00:...	0	196.00	0.00
7	7	Resort	2015-07-0...	2015-07-01 0...	0	2	2	Direct	Transient	Portugal	No Deposit	107.00	Check_Out	2015-07-03 00:...	0	214.00	0.00
8	8	Resort	2015-06-2...	2015-07-01 0...	9	2	2	Direct	Transient	Portugal	No Deposit	103.00	Check_Out	2015-07-03 00:...	0	206.00	0.00
9	9	Resort	2015-04-0...	2015-07-01 0...	85	3	2	Online Travel Agent	Transient	Portugal	No Deposit	82.00	Canceled	2015-05-06 00:...	1	0.00	-246.00
10	10	Resort	2015-04-1...	2015-07-01 0...	75	3	2	Offline Travel Agent	Transient	Portugal	No Deposit	105.50	Canceled	2015-04-22 00:...	1	0.00	-316.50

2.Total Revenue for the year 2016

```sql
Select SUM(Revenue) TotalRevenue

From HotelReservation..Hotel  Where DATEPART(yy,ArrivalDate) = 2016
```

	TotalRevenue
1	13303225.70

3.Monthly Total Revenue

```sql
select DATEPART(MM,ArrivalDate) ArrivalMonth, Sum(Revenue) as Total_Revenue

From HotelReservation..Hotel where DATEPART(yy,ArrivalDate)=2016

Group by DATEPART(MM,ArrivalDate)  Order by Total_Revenue  Desc
```

	ArrivalMonth	Total_Revenue
1	8	1844836.15
2	7	1590291.05
3	9	1480759.14
4	6	1439152.88
5	10	1247601.84
6	5	1231884.13
7	4	1126280.09
8	3	919801.53
9	11	850531.67
10	12	701367.12
11	2	579198.72
12	1	291521.38

The highest 4 values for Revenue were in August, July, September and June.

4.Total Revenue by Country

```sql
select Top 5 Country, SUM(Revenue) as Total_Revenue

from HotelReservation..Hotel  where DATEPART(yy,ArrivalDate)=2016 Group by Country order by Total_Revenue Desc
```

	Country	Total_Revenue
1	Portugal	3928589.33
2	United Kingdom	1834079.59
3	France	1499968.26
4	Germany	1035933.43
5	Spain	957749.35

The most revenue is from customers who came from Portugal, United Kingdom, France, Germany and Spain.

5.Total Revenue by DistributionChannel

```sql
select DistributionChannel, SUM(Revenue) as Total_Revenue

from HotelReservation..Hotel  Where DATEPART(yy,ArrivalDate)=2016  Group by
DistributionChannel order by Total_Revenue Desc
```

	DistributionChannel	Total_Revenue
1	Online Travel Agent	7562215.22
2	Offline Travel Agent	3440153.25
3	Direct	1849005.04
4	Corporate	451852.19

Online Travel Agent channels contribute a big portion of the total revenue (i.e, most customers use online travel agents for their reservations).

6.Average Daily Rate by month

```sql
select  DATEPART(MM,ArrivalDate) ArrivalMonth, avg(AvgDailyRate) as AverageDailyRate

 from HotelReservation..Hotel   Where DATEPART(yy,ArrivalDate) =2016

 group by  DATEPART(MM,ArrivalDate)

 Order by  AverageDailyRate Desc
```

	ArrivalMonth	AverageDailyRate
1	8	142.8935
2	7	125.4845
3	9	114.7495
4	6	106.9773
5	5	96.3992
6	10	95.1117
7	4	88.9189
8	12	86.3702
9	11	80.7375
10	3	79.0693
11	2	70.1022
12	1	64.7676

The highest Average Daily Rate is for the months 8, 7, 9, and 6.

7. Cancellation Rate by month

```sql
with T1 as (select  DATEPART(MM,ArrivalDate) ArrivalMonth, count(Cancelled) As
CancelledCount, (select count (*) from HotelReservation..Hotel where
DATEPART(yy,ArrivalDate) =2016) as All_records

            from HotelReservation..Hotel   where Cancelled =0  and
DATEPART(yy,ArrivalDate) =2016 group by  DATEPART(MM,ArrivalDate))

Select  ArrivalMonth, CancelledCount, All_records,
Cast(((((CancelledCount)/(All_records*1.0))*100) As Decimal (10,2)) as CancellationRate

From T1  Order By  CancellationRate Desc
```

	ArrivalMonth	CancelledCount	All_records	CancellationRate
1	10	3689	56707	6.51
2	5	3563	56707	6.28
3	9	3372	56707	5.95
4	4	3367	56707	5.94
5	3	3347	56707	5.90
6	8	3238	56707	5.71
7	6	3196	56707	5.64
8	7	3073	56707	5.42
9	11	2818	56707	4.97
10	2	2554	56707	4.50
11	12	2462	56707	4.34
12	1	1691	56707	2.98

October showed the highest cancellation rate followed by May.

8.Cancelation Rate based on LeadTime

```sql
with T1 as (select   Cancelled, case when LeadTime <= 30 Then '0-30 Days' Else '> 30
Days' End As LeadTimeCategories

                        from HotelReservation..Hotel   where Cancelled =0 and
DATEPART(yy,ArrivalDate)=2016),

    T2 as (select LeadTimeCategories, Cast(((count(Cancelled)*1.0)/(select count (*)
from HotelReservation..Hotel where DATEPART(yy,ArrivalDate)=2016)*100) AS Decimal (10,2))
as CancellationRate

        from T1   group by LeadTimeCategories )


select LeadTimeCategories, CancellationRate

from T2 order by LeadTimeCategories desc
```

	LeadTimeCategories	CancellationRate
1	0-30 Days	26.89
2	> 30 Days	37.25

More than 30 days as lead time (no. of days between a guest books a room and the time the guest scheduled to arrive at the hotel) scored the highest cancellation rate.

9.Room Status By month

```
 Update HotelReservation..Hotel

set Status= 'No_Show'

    Where Status= 'No-Show'



Update HotelReservation..Hotel

set Status= 'Check_Out'

    Where Status= 'Check-Out'

  SELECT [ArrivalMonth], Canceled,No_Show, Check_Out   FROM (select
DATEPART(MM,ArrivalDate) ArrivalMonth,

                    Status, COUNT(Status) As Count_Status

 From HotelReservation..Hotel  Where DATEPART(yy,ArrivalDate)=2016 Group by
DATEPART(MM,ArrivalDate), Status) as T1

 PIVOT (sum   (Count_Status)  for Status in (Canceled,No_Show, Check_Out)) As PivotTable
```

	ArrivalMonth	Canceled	No_Show	Check_Out
1	1	507	50	1691
2	2	1192	145	2554
3	3	1406	71	3347
4	4	2007	54	3367
5	5	1850	65	3563
6	6	2046	50	3196
7	7	1461	38	3073
8	8	1786	39	3238
9	9	1993	29	3372
10	10	2459	55	3689
11	11	1597	39	2818
12	12	1365	33	2462

1-Most rooms are checked out while a large proportion of rooms are being canceled.

2-The number of rooms that were booked but the customers did not show up (No_Show) was inconsiderable.

3-Most check out rooms were during October and May.

10. The most effective booking channels

select DistributionChannel, count(DistributionChannel) as ChannelCount

from HotelReservation..Hotel Where DATEPART(yy,ArrivalDate)=2016 Group by DistributionChannel order by ChannelCount Desc

	DistributionChannel	ChannelCount
1	Online Travel Agent	34358
2	Offline Travel Agent	12258
3	Direct	6799
4	Corporate	3292

Online Travel Agent channels play a huge role in hotel reservation. The least effective channel in hotel reservation is the Corporate channel.

11. The most popular hotels

select Hotel, Number_of_Booking From(Select Hotel, count(ArrivalDate) AS Number_of_Booking

From HotelReservation..Hotel where DATEPART(yy,ArrivalDate)=2016 Group by Hotel) AS T1 order by Number_of_Booking Desc

	Hotel	Number_of_Booking
1	City	38140
2	Resort	18567

City hotels got much more customers compared to Resort hotels.

12.Comparing hotels based on Customer Type

Select CustomerType, City, Resort

From (Select Hotel, CustomerType, count(CustomerType) AS Number_Customer

```
                                        From HotelReservation..Hotel where
DATEPART(yy,ArrivalDate)=2016 Group by Hotel, CustomerType) AS T1

PIVOT (sum (Number_Customer)  for Hotel in (City, Resort )) As PivotTable  order by City
desc , Resort desc
```

	CustomerType	City	Resort
1	Transient	30298	14093
2	Transient-Party	7600	3793
3	Contract	180	576
4	Group	62	105

1.Transient is the main customer type, followed by Transient-Party and contract customer type.

2.City hotels are the most popular hotels for Transient and Transient-Party customers.

3.However, Resort Hotel data show a higher portion of Contract customer type at a total of 576, while the number for City Hotel only at 180.

4.Group customer type is inconsiderable.