HOTEL RESERVATION ANALYSIS

TALAL MANSHOOR

Batch – MIP-DA-10

PROBLEM STATEMENT

Tourism is one of constantly growing and highly profitable industry. Thus hotels are more keen about their customers, their services, and their value in industry.
The problem statement is to retrieve insights from the dataset provided by hotel and help them in making data driven decisions yielding more better customer experience and profit.

DATASET

Guest Information	Reservation Details	Booking Context
Booking_ID: Unique identifier for each reservation.	no_of_weekend_nights: Number of nights booked on weekends. no_of_week_nights: Number of nights booked on weekdays.	market_segment_type: Category of customer the reservation belongs to.
no_of_adults: Number of adults in the reservation.	arrival_date: Date the guests are scheduled to arrive.	avg_price_per_room: Average price per room for the reservation.
no_of_children: Number of children in the reservation.	type_of_meal_plan: Meal plan chosen by the guests	booking_status: Current status of the reservation
	room_type_reserved: The type of room reserved by the guests.	

- Dataset contains following columns which are represented in group form for better understanding

ARRIVAL_DATE COLUMN INCONSISTENCY

- Arrival_date had date in two different formats in string data type which is first changed to same format string then changed data type to date

```
23 •  SELECT
24        arrival_date,
25        CASE
26            WHEN arrival_date LIKE '%/%/%' THEN DATE_FORMAT(STR_TO_DATE(arrival_date, '%d/%m/%Y'), '%d-%m-%Y')
27            WHEN arrival_date LIKE '%-%-%' THEN DATE_FORMAT(STR_TO_DATE(arrival_date, '%d-%m-%Y'), '%d-%m-%Y')
28            ELSE arrival_date
29        END AS formatted_date
30    FROM hotel_reservation_dataset;
31
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

arrival_date	formatted_date
02/10/2017	02-10-2017
06/11/2018	06-11-2018
28-02-2018	28-02-2018
20-05-2018	20-05-2018
11/04/2018	11-04-2018
13-09-2018	13-09-2018
15-10-2017	15-10-2017

Result 10

```
3      -- Q1 : What is the total number of reservations in
4 •    SELECT COUNT(*) AS total_reservations
5      FROM hotel_reservation_dataset;
```

	total_reservations
▶	700

WHAT IS THE TOTAL NUMBER OF RESERVATIONS IN THE DATASET?

```sql
-- Q2 : Which meal plan is most popular among guests?
SELECT type_of_meal_plan, COUNT(*) AS plan_count
FROM hotel_reservation_dataset
GROUP BY type_of_meal_plan
LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

type_of_meal_plan	plan_count
Meal Plan 1	527

WHICH MEAL PLAN IS THE MOST POPULAR AMONG GUESTS?

```
13      -- Q3 : What is the average price per room for reservations invo
14      SELECT AVG(avg_price_per_room) AS avg_price_with_children
15      FROM hotel_reservation_dataset
16      WHERE no_of_children > 0;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

avg_price_with_children
144.56833333333336

WHAT IS THE AVERAGE PRICE PER ROOM FOR RESERVATIONS INVOLVING CHILDREN?

```sql
-- Q4 : How many reservations were made for the year 20XX (replace XX wit
SELECT EXTRACT(YEAR FROM arrival_date) AS reservation_year, COUNT(*) AS r
FROM hotel_reservation_dataset
GROUP BY reservation_year;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

reservation_year	reservation_count
2017	123
2018	577

HOW MANY RESERVATIONS WERE MADE FOR THE YEAR 20XX?

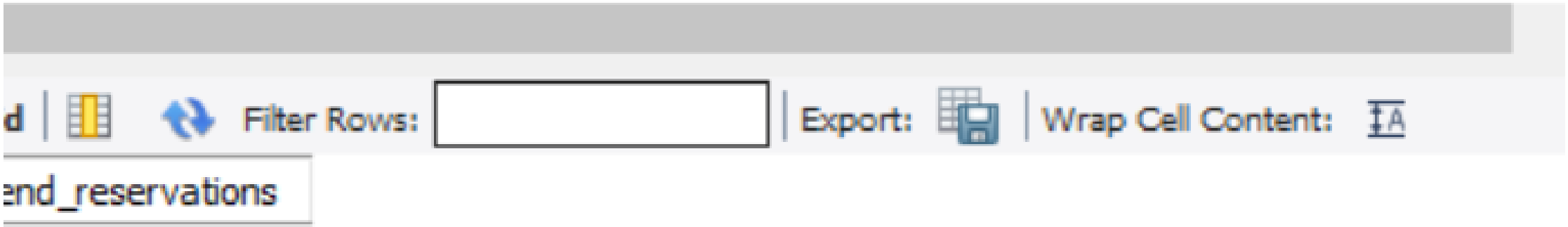
```
24 •    SELECT room_type_reserved, count(*) AS room_count
25      FROM hotel_reservation_dataset
26      GROUP BY room_type_reserved
```

Result Grid | Filter Rows: | Export: | Wrap Cell C

room_type_reserved	room_count
Room_Type 1	534
Room_Type 4	130
Room_Type 2	8
Room_Type 6	18
Room_Type 5	4
Room_Type 7	6

WHAT IS THE MOST COMMONLY BOOKED ROOM TYPE?

```sql
-- Q6 : How many reservations fall on a weekend (no_of_weekend_ni
SELECT COUNT(*) AS weekend_reservations FROM hotel_reservation_da
WHERE no_of_weekend_nights > 0;
```



HOW MANY RESERVATIONS FALL ON A WEEKEND?

```
33 •    SELECT
34        MAX(lead_time) AS highest_lead_time,
35        MIN(lead_time) AS lowest_lead_time
36      FROM hotel_reservation_dataset;
37
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

highest_lead_time	lowest_lead_time
443	0

WHAT IS THE HIGHEST AND LOWEST LEAD TIME FOR RESERVATIONS?

```sql
38      -- Q8 : What is the most common market type for reservations?
39      SELECT market_segment_type, COUNT(*) AS segment_count
40      FROM hotel_reservation_dataset
41      GROUP BY market_segment_type
42      ORDER BY segment_count DESC
43
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

market_segment_type	segment_count
Online	518
Offline	140
Corporate	27
Complementary	14
Aviation	1

WHAT IS THE MOST COMMON MARKET SEGMENT TYPE FOR RESERVATIONS?

```
44      -- Q9 : How many reservations have a booking status of "Confirme
45 •    SELECT COUNT(*) AS confirmed_reservations
46      FROM hotel_reservation_dataset
47      WHERE booking_status = 'Not_Canceled';
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

	confirmed_reservations
▶	493

HOW MANY RESERVATIONS HAVE A BOOKING STATUS OF "CONFIRMED"?

```
49      -- Q10 : What is the total number of adults and children across all reser
50 •    SELECT
51        SUM(no_of_adults) AS total_adults,
52        SUM(no_of_children) AS total_children
53      FROM hotel_reservation_dataset;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

	total_adults	total_children
▶	1316	69

WHAT IS THE TOTAL NUMBER OF ADULTS AND CHILDREN ACROSS ALL RESERVATIONS?

```
55      -- Q11 : What is the average number of weekend nights for reser
56 •    SELECT AVG(no_of_weekend_nights) AS avg_weekend_nights_with_chi
57      FROM hotel_reservation_dataset
58      WHERE no_of_children > 0;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

avg_weekend_nights_with_children
1.0000

WHAT IS THE AVERAGE NUMBER OF WEEKEND NIGHTS FOR RESERVATIONS INVOLVING CHILDREN?

```sql
SELECT
  MONTH(arrival_date) AS Month,
  MONTHNAME(arrival_date) AS MonthName,
  SUM(CASE WHEN YEAR(arrival_date) = 2017 THEN 1 ELSE 0 END) AS 2017_reservations,
  SUM(CASE WHEN YEAR(arrival_date) = 2018 THEN 1 ELSE 0 END) AS 2018_reservations
FROM hotel_reservation_dataset
GROUP BY MonthName, Month
ORDER BY Month;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Month	MonthName	2017_reservations	2018_reservations
1	January	0	11
2	February	0	28
3	March	0	52
4	April	0	67
5	May	0	55
6	June	0	84
7	July	8	36
8	August	14	56
9	September	35	45
10	October	40	63
11	November	13	41
12	December	13	39

HOW MANY RESERVATIONS WERE MADE IN EACH MONTH OF THE YEAR?

```sql
-- Q13 : What is the average number of nights (both weekend and weekday) spent by guests for each room type?
SELECT room_type_reserved,
AVG(no_of_week_nights + no_of_weekend_nights) AS Avg_Total_Nights
FROM hotel_reservation_dataset
GROUP BY room_type_reserved
ORDER BY room_type_reserved;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

room_type_reserved	Avg_Total_Nights
Room_Type 1	2.8783
Room_Type 2	3.0000
Room_Type 4	3.8000
Room_Type 5	2.5000
Room_Type 6	3.6111
Room_Type 7	2.6667

WHAT IS THE AVERAGE NUMBER OF NIGHTS SPENT BY GUESTS FOR EACH ROOM TYPE?

```sql
-- Q14 : For reservations involving children, what is the most common room type, and what is the
SELECT room_type_reserved, COUNT(*) AS room_count,
ROUND(AVG(avg_price_per_room),2) AS Avg_Price
FROM hotel_reservation_dataset
WHERE no_of_children > 0
GROUP BY room_type_reserved
ORDER BY room_count DESC
LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

room_type_reserved	room_count	Avg_Price
Room_Type 1	24	123.12

FOR RESERVATIONS INVOLVING CHILDREN, WHAT IS THE MOST COMMON ROOM TYPE, AND WHAT IS THE AVERAGE PRICE FOR THAT ROOM TYPE?

```sql
-- Q15 : Find the market segment type that generates the highest average price per ro
SELECT market_segment_type AS Market_Segment,
ROUND(AVG(avg_price_per_room),2) AS Avg_Price
FROM hotel_reservation_dataset
GROUP BY Market_Segment
ORDER BY Avg_Price DESC
LIMIT 1;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Market_Segment	Avg_Price
Online	112.46

FIND THE MARKET SEGMENT TYPE THAT GENERATES THE HIGHEST AVERAGE PRICE PER ROOM.

CONCLUSION

ALL THE QUESTIONS ASKED BY HOTEL MANAGEMENT ARE ANSWERED USING SQL QUERY AND PROVIDED TO STAKE HOLDERS FOR IMPROVEMENTS IN BUSINESS PROCESS