

Data Analysis Hotel Reservation Project

**By
Sahar Sayed Elasersawi**

Introduction

- ❑ This project is talking about the analysis of Hotel Reservation.
- ❑ I used SQL analysis methodologies and techniques to analyze the dataset to gain different insights and customers preferences.
- ❑ I used MySQL coding to answer 15 questions according to the stakeholders requirements.
- ❑ I answered the 15 questions to make good analysis to be shown clearly to the stakeholders through this report.

Data Preparation

- ❑ Open MySQL Workbench.
- ❑ Upload the Hotel reservation dataset as CSV file to the database.
- ❑ Start analys and answer questions.

Understanding the Dataset

- ❑ The dataset includes the following columns:
- ❑ `Booking_ID`: A unique identifier for each hotel reservation.
- ❑ `no_of_adults`: The number of adults in the reservation.
- ❑ `no_of_children`: The number of children in the reservation.
- ❑ `no_of_weekend_nights`: The number of nights in the reservation that fall on weekends.
- ❑ `no_of_week_nights`: The number of nights in the reservation that fall on weekdays.
- ❑ `type_of_meal_plan`: The meal plan chosen by the guests.
- ❑ `room_type_reserved`: The type of room reserved by the guests.
- ❑ `lead_time`: The number of days between booking and arrival.
- ❑ `arrival_date`: The date of arrival.
- ❑ `market_segment_type`: The market segment to which the reservation belongs.
- ❑ `avg_price_per_room`: The average price per room in the reservation.
- ❑ `booking_status`: The status of the booking.



```
1 • SELECT * FROM hotels.`hotel reservation dataset`  
2 LIMIT 5
```

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

	Booking_ID	no_of_adults	no_of_children	no_of_weekend_nights	no_of_week_nights	type_of_meal_plan	room_type_reserved	lead_time	arrival_date	market_segmen
▶	INN00001	2	0	1	2	Meal Plan 1	Room_Type 1	224	02-10-2017	Offline
	INN00002	2	0	2	3	Not Selected	Room_Type 1	5	06-11-2018	Online
	INN00003	1	0	2	1	Meal Plan 1	Room_Type 1	1	28-02-2018	Online
	INN00004	2	0	0	2	Meal Plan 1	Room_Type 1	211	20-05-2018	Online
	INN00005	2	0	1	1	Not Selected	Room_Type 1	48	11-04-2018	Online



```
1  -- 1. What is the total number of reservations in the dataset?
2  -- SELECT * FROM hotels.`hotel reservation dataset`;
3  •  SELECT COUNT(*) AS total_reservations
4     FROM hotels.`hotel reservation dataset`
5
```

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

	total_reservations
▶	700



```

1  -- 2. Which meal plan is the most popular among guests?
2  • SELECT type_of_meal_plan , COUNT(*) AS count
3     FROM hotels.`hotel reservation dataset`
4     GROUP BY type_of_meal_plan
5     ORDER BY count DESC;
6

```

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

	type_of_meal_plan	count
▶	Meal Plan 1	527
	Not Selected	109
	Meal Plan 2	64



```
1  -- 3. What is the average price per room for reservations involving children?
2  -- SELECT no_of_children , avg_price_per_room
3  -- FROM hotels.`hotel reservation dataset`;
4  • SELECT round (AVG(avg_price_per_room),2) AS avg_price_per_room
5  FROM hotels.`hotel reservation dataset`
```

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

	avg_price_per_room
▶	144.57



```

1  -- 4. How many reservations were made for the year 20XX (replace XX with the desired year)?
2  Save the script to a file. 4 hotels.`hotel reservation dataset`;
3  • SELECT
4  year(str_to_date( arrival_date, '%d-%m-%Y')) as year_of_arrival,
5  COUNT(*) AS reservations_per_year
6  FROM hotels.`hotel reservation dataset`
7  group by year_of_arrival
8  order by year_of_arrival DESC
9

```

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

	year_of_arrival	reservations_per_year
▶	2018	577
	2017	123



```
1  -- 5- What is the most commonly booked room type?
2  •  SELECT room_type_reserved, COUNT(*) AS room_count
3     FROM hotels.`hotel reservation dataset`
4     GROUP BY room_type_reserved
5     ORDER BY room_count DESC
6     LIMIT 1
```

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

	room_type_reserved	room_count
▶	Room_Type 1	534

6_No. of weekend reservations

7_highest & lowest lead time

8_common market segment

9_No. of confirmed reservations*



```
1 -- 6. How many reservations fall on a weekend (no_of_weekend_nights > 0)?
2 -- SELECT * FROM hotels.`hotel reservation dataset`;
3 • SELECT COUNT(*) AS weekend_reservations
4 FROM hotels.`hotel reservation dataset`
5 WHERE no_of_weekend_nights > 0;
```

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

	weekend_reservations
▶	383



```
1  -- 7. What is the highest and lowest lead time for reservations?
2  -- SELECT * FROM hotels.`hotel reservation dataset`;
3  • SELECT
4      MAX(lead_time) AS highest_lead_time,
5      MIN(lead_time) AS lowest_lead_time
6  FROM hotels.`hotel reservation dataset`
7  WHERE lead_time <> 0
8
```

Result Grid Filter Rows: Export: Wrap Cell Content:

	highest_lead_time	lowest_lead_time
▶	443	1

```

1  -- 8. What is the most common market segment type for reservations?
2  -- Execute the selected portion of the script or everything, if there is no selection
3  • SELECT market_segment_type, COUNT(*) AS segment_count
4     FROM hotels.`hotel reservation dataset`
5     GROUP BY market_segment_type
6     ORDER BY segment_count DESC
    
```

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

Fetch rows:

market_segment_type	segment_count
Online	518

9_No. of confirmed reservations x

10_No. of adults & no. of childre...

11_no. of weekends with children

12_reservations per month & year



```
1  -- 9. How many reservations have a booking status of "Confirmed"?
2  -- SELECT * FROM hotels.`hotel reservation dataset`;
3  • SELECT booking_status , COUNT(*) AS confirmed_reservations
4  FROM hotels.`hotel reservation dataset`
5  where booking_status = 'Not_canceled';
6
```

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

	booking_status	confirmed_reservations
▶	Not_Canceled	493

on market segment

9_No. of confirmed reservations*

10_No. of adults & no. of childre... x

11_no. of weekends with childre...



Limit to 1000 rows



```
1  -- 10. What is the total number of adults and children across all reservations?
2  -- SELECT * FROM hotels.`hotel reservation dataset`;
3  •  SELECT
4      SUM(no_of_adults) AS total_adults,
5      SUM(no_of_children) AS total_children
6  FROM hotels.`hotel reservation dataset`
7
```

Result Grid



Filter Rows:

Export:



Wrap Cell Content:



	total_adults	total_children
▶	1316	69



```

1  -- 11. What is the average number of weekend nights for reservations involving children?
2  -- SELECT * FROM hotels.`hotel reservation dataset`;
3
4  ● SELECT AVG(no_of_weekend_nights) AS avg_weekend_nights_with_children
5     FROM hotels.`hotel reservation dataset`
6     WHERE no_of_children > 0
7

```




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

	avg_weekend_nights_with_children
▶	1.0000


```

1  -- 12. How many reservations were made in each month of the year?
2  • SELECT
3      YEAR(STR_TO_DATE(arrival_date, '%d-%m-%Y')) AS years,
4      MONTH(STR_TO_DATE(arrival_date, '%d-%m-%Y')) AS months,
5      COUNT(*) AS reservation_count
6  FROM
7      hotels.`hotel reservation dataset`
8  GROUP BY
9      years, months
10 ORDER BY
11     years, months
12

```

Result Grid
 

 Filter Rows:
 Export: 
 Wrap Cell Content: 

	years	months	reservation_count
	2017	7	8
	2017	8	14
	2017	9	35
	2017	10	40
	2017	11	13
	2017	12	13
	2018	1	11
	2018	2	28
	2018	3	52



```

1  -- 13. What is the average number of nights (both weekend and weekday) spent by guests for each room type?
2  -- SELECT * FROM hotels.`hotel reservation dataset`;
3  • SELECT room_type_reserved , round(avg( no_of_weekend_nights + no_of_week_nights )) as avg_nights
4  FROM hotels.`hotel reservation dataset`
5  GROUP BY room_type_reserved
6
    
```

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

room_type_reserved	avg_nights
Room_Type 1	3
Room_Type 4	4
Room_Type 2	3
Room_Type 6	4
Room_Type 5	3
Room_Type 7	3



```

1  -- 14. For reservations involving children, what is the most common room type, and what is the average price for that room type?
2  -- SELECT * FROM hotels.`hotel reservation dataset`;
3  • SELECT room_type_reserved , count(*) AS most_common_room ,
4  ROUND (avg(avg_price_per_room), 2) as average_price
5  FROM hotels.`hotel reservation dataset`
6  WHERE no_of_children > 0
7  group by room_type_reserved
8  order by most_common_room DESC
9  limit 1
10

```

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

	room_type_reserved	most_common_room	average_price
▶	Room_Type 1	24	123.12

11_no. of weekends with children

12_reservations per month &year

13_Average nights

14_ common room & average pr...

15_market segment with the hig...



```
1  -- 15. Find the market segment type that generates the highest average price per room.
2  -- SELECT * FROM hotels.`hotel reservation dataset`;
3  • SELECT avg_price_per_room , market_segment_type
4  from hotels.`hotel reservation dataset`
5  order by avg_price_per_room DESC
6  limit 1
-
```

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

	avg_price_per_room	market_segment_type
▶	258	Online

Conclusion

Data Analytics insights and recommendations

- ❑ We have 700 reservations on 2017 and 2018.
- ❑ Meal plan1 is the most requested by 75%.
- ❑ Average price per room involving children is 144.57.
- ❑ Total reservation in 2017 is 123 and in 2018 is 577 i.e gross rate by more than 300% and this is very good indication for the hotel strategy in general including marketing activities and pricing as well.
- ❑ Room type 1 is the most requested type by about 76% than others room types, its average price is 123.12.

- ❑ Weekend reservation is almost same like weekdays which is good.
- ❑ Online reservations is preferred than other types by more than 75%.
- ❑ Most reservations are confirmed , good indication of serious customers.
- ❑ Adults are more than children which is reasonable.

Thank You