



# Hotel Reservation Analysis

Project on SQL

Presenter: Shubham Singh



# Introduction

- **Purpose:** Explain the purpose of the project, which is to develop data analysis skills in a practical context by analyzing a hotel reservation dataset using SQL.
- **Importance:** Highlight the importance of data analysis in the hotel industry for making informed decisions and improving guest experiences.



# Project Overview

**Dataset Description:** Provide an overview of the dataset used, including the key columns

- **Booking\_ID:** Unique identifier for each reservation.
- **no\_of\_adults:** Number of adults in the reservation.
- **no\_of\_children:** Number of children in the reservation.
- **no\_of\_weekend\_nights:** Number of weekend nights in the reservation.
- **no\_of\_week\_nights:** Number of weekday nights in the reservation.
- **type\_of\_meal\_plan:** Meal plan chosen by the guests.
- **room\_type\_reserved:** Type of room reserved.
- **lead\_time:** Number of days between booking and arrival.
- **arrival\_date:** Date of arrival.
- **market\_segment\_type:** Market segment of the reservation.
- **avg\_price\_per\_room:** Average price per room.
- **booking\_status:** Status of the booking.

# Objectives

- **Goals:** List the primary goals of the analysis:
- Gain insights into guest preferences.
- Identify booking trends.
- Understand key factors impacting hotel operations.





# Methodology

**SQL Queries:** Explain the use of SQL to query and analyze the dataset. Mention that specific questions were answered using SQL queries

# ● Hotel Analysis Dataset as hostel

Booking_ID	no_of_adults	no_of_children	no_of_weekdays	no_of_weekends	type_of_meal_plan	room_type	lead_time	arrival_date	market_segment	avg_price	booking_status
INN00001	2	0	1	2	Meal Plan	Room_Type_1	224	#####	Offline	65	Not_Canceled
INN00002	2	0	2	3	Not Select	Room_Type_2	5	#####	Online	106.68	Not_Canceled
INN00003	1	0	2	1	Meal Plan	Room_Type_3	1	#####	Online	60	Canceled
INN00004	2	0	0	2	Meal Plan	Room_Type_4	211	#####	Online	100	Canceled
INN00005	2	0	1	1	Not Select	Room_Type_5	48	#####	Online	94.5	Canceled
INN00006	2	0	0	2	Meal Plan	Room_Type_6	346	#####	Online	115	Canceled
INN00007	2	0	1	3	Meal Plan	Room_Type_7	34	#####	Online	107.55	Not_Canceled
INN00008	2	0	1	3	Meal Plan	Room_Type_8	83	#####	Online	105.61	Not_Canceled
INN00009	3	0	0	4	Meal Plan	Room_Type_9	121	#####	Offline	96.9	Not_Canceled
INN00010	2	0	0	5	Meal Plan	Room_Type_10	44	#####	Online	133.44	Not_Canceled
INN00011	1	0	1	0	Not Select	Room_Type_11	0	#####	Online	85.03	Not_Canceled
INN00012	1	0	2	1	Meal Plan	Room_Type_12	35	#####	Online	140.4	Not_Canceled
INN00013	2	0	2	1	Not Select	Room_Type_13	30	#####	Online	88	Canceled
INN00014	1	0	2	0	Meal Plan	Room_Type_14	95	#####	Online	90	Canceled
INN00015	2	0	0	2	Meal Plan	Room_Type_15	47	#####	Online	94.5	Not_Canceled
INN00016	2	0	0	2	Meal Plan	Room_Type_16	256	#####	Online	115	Canceled
INN00017	1	0	1	0	Meal Plan	Room_Type_17	0	#####	Offline	96	Not_Canceled
INN00018	2	0	1	3	Not Select	Room_Type_18	1	#####	Online	96	Not_Canceled
INN00019	2	0	2	2	Meal Plan	Room_Type_19	99	#####	Online	65	Canceled
INN00020	2	0	1	0	Meal Plan	Room_Type_20	12	#####	Offline	72	Not_Canceled
INN00021	2	0	2	2	Meal Plan	Room_Type_21	99	#####	Online	65	Canceled
INN00022	1	0	0	1	Meal Plan	Room_Type_22	122	#####	Corporate	67	Not_Canceled
INN00023	2	0	2	4	Meal Plan	Room_Type_23	2	#####	Offline	85	Not_Canceled

# 1. What is the total number of reservations in the dataset?

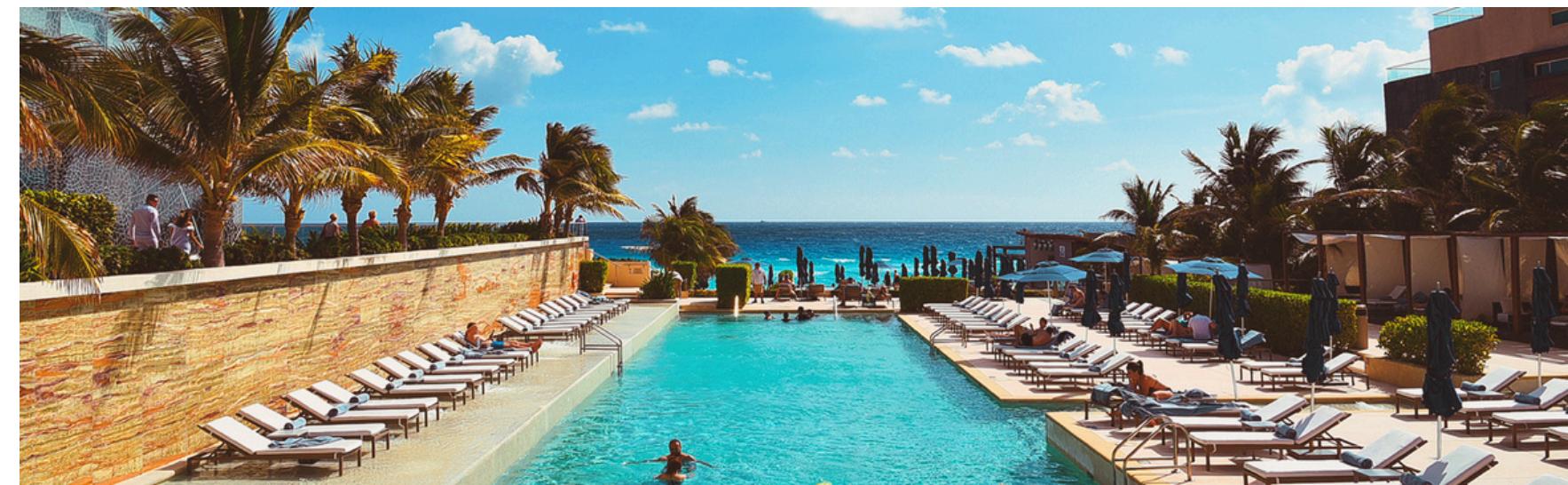
```
select count(*) from hostel
```

Data Output Message

	<b>count</b>			
	<b>bigint</b>			

1	700
---	-----

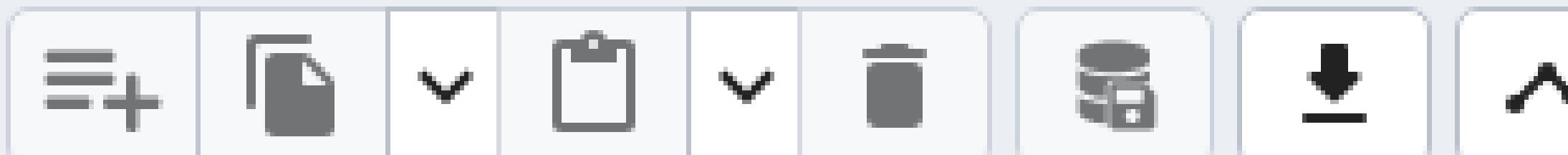


## 2. Which meal plan is the most popular among guests?

```
SELECT type_of_meal_plan, COUNT(*) AS count
FROM hostel
GROUP BY type_of_meal_plan
ORDER BY count DESC
LIMIT 1;
```

Data Output    Messages    Notifications

---



	type_of_meal_plan character varying (50)	count bigint
1	Meal Plan 1	527

### 3. What is the average price per room for reservations involving children?

```
✓ SELECT AVG(avg_price_per_room) AS average_price_per_room  
FROM hostel  
WHERE no_of_children > 0;
```

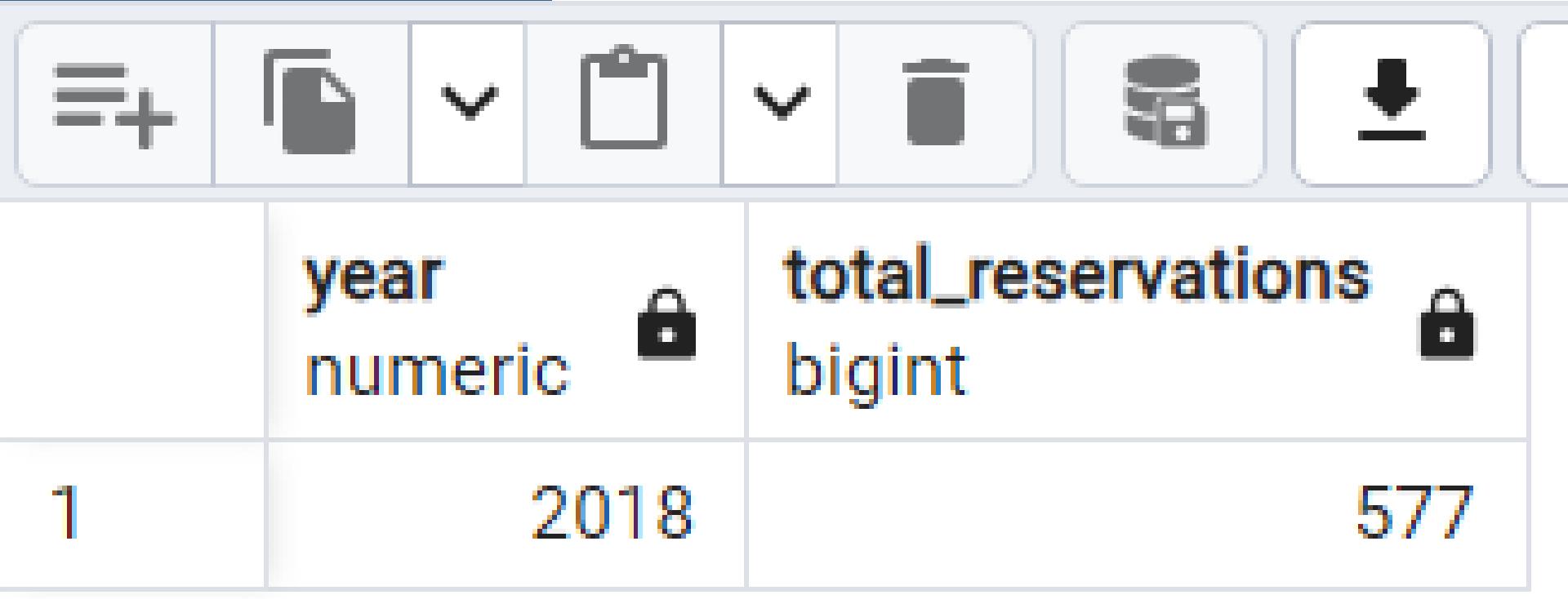
	Data	Output	Messages	Notif
	average_price_per_room			
	numeric			
1	144.62500000000000			

#### 4. How many reservations were made for the year 20XX (replace XX with the desired year)?

```
✓ SELECT EXTRACT (YEAR FROM arrival_date) AS year, COUNT(*) AS total_reservations  
FROM hostel  
WHERE EXTRACT (YEAR FROM arrival_date) = 2018  
GROUP BY year;
```

Data Output Messages Notifications

---



	year	total_reservations
	numeric 1	bigint 577

## 5. What is the most commonly booked room type?

```
2 ✓ SELECT room_type_reserved, COUNT(*) AS count
3   FROM hostel
4 GROUP BY room_type_reserved
5 ORDER BY count DESC
6 LIMIT 1;
```

Data Output Messages Notifications

The screenshot shows a MySQL Workbench interface with a toolbar at the top and a results table below. The results table has two columns: 'room\_type\_reserved' and 'count'. The first row shows 'Room\_Type 1' with a count of 534.

	room_type_reserved	count
1	Room_Type 1	534

## 6. How many reservations fall on a weekend (no\_weekend\_nights > 0)?

```
2 ✓ SELECT COUNT(*) AS total_reservations
3
4 FROM hostel
5 WHERE no_of_weekend_nights > 0;
```

Data Output    Messages

	total_reservations	bigint
1		383

# 7. What is the highest and lowest lead time for reservations?

```
2 ✓ select max(lead_time) as highest, min(lead_time) as lowest  
3   from hostel
```

Data Output Messages Notifications

The screenshot shows a database interface with a toolbar at the top containing icons for new table, file, save, refresh, trash, and search. Below the toolbar is a table with two columns: 'highest' and 'lowest'. The 'highest' column contains the value '443' and the 'lowest' column contains the value '0'. Both columns have a lock icon next to them.

	highest	lowest
	443	0

8. What is the most common market segment type for reservations?

```
2 ✓ select market_segment_type, count(*) as most_common_market
3 from hostel
4 group by market_segment_type
5 order by most_common_market desc
6 limit 1;
```

Data Output	Messages	Notifications	
     	   		
market_segment_type character varying (60)		most_common_market bigint	
1	Online	518	

# 9. How many reservations have a booking status of "Confirmed"?

```
2 ✓ select booking_status, count(*) as confirmed_bookings  
3   from hostel  
4   group by booking_status  
5   order by confirmed_bookings desc  
6   limit 1  
7
```

Data Output    Messages    Notifications

≡+ ↗ ↘ ↕ ↖ ↙ SQL

	booking_status	confirmed_bookings
	character varying (50)	bigint
1	Not_Canceled	493

# 10. What is the total number of adults and children across all reservations?

```
2 ✓ select sum(no_of_adults + no_of_children) as total  
3   from hostel
```

Data Output    Messages

The screenshot shows a database interface with a toolbar at the top containing icons for adding, saving, and deleting. Below the toolbar is a table with one row. The first column contains the value '1'. The second column has the header 'total' and the type 'numeric', with a lock icon indicating it's read-only. The value '1385' is displayed in the cell. The table has a light gray background with white borders between rows and columns.

1	total numeric
	1385

# 11. What is the average number of weekend nights for reservations involving children?

```
2 ✓ select round(avg(no_of_weekend_nights),2) as avg_weekend_nights  
3   from hostel  
4 where no_of_children > 0 ;
```

Data Output Messages Notes

The screenshot shows a database interface with a toolbar at the top containing icons for new table, save, refresh, clipboard, dropdown, trash, and a refresh symbol. Below the toolbar is a table with one row. The first column is labeled 'avg\_weekend\_nights' and has a 'numeric' data type. A lock icon is next to the column name. The value '1.00' is displayed in the cell under this column. The row also contains a small lock icon.

avg_weekend_nights	
1	1.00

# 12. How many reservations were made in each month of the year?

```
2 SELECT
3     EXTRACT(MONTH FROM arrival_date) AS Month,
4     COUNT(*) AS ReservationCount
5 FROM
6     HOSTEL
7 GROUP BY
8     EXTRACT(MONTH FROM arrival_date)
9 ORDER BY
10    EXTRACT(MONTH FROM arrival_date)
```

Data Output    Messages    Notifications

	month numeric	reservationcount bigint
1	1	11
2	2	28
3	3	52
4	4	67
5	5	55
6	6	84
7	7	44
8	8	70
9	9	80
10	10	103
11	11	54
12	12	52

### 13. What is the average number of nights (both weekend and weekday) spent by guests for each room type?

```
4 ▼ SELECT
5     room_type_reserved,
6     AVG(no_of_weekend_nights + no_of_week_nights) AS AvgNights
7 FROM
8     hostel
9 GROUP BY
10    room_type_reserved
11 ORDER BY
12    room_type_reserved;
```

Data Output    Messages    Notifications

The screenshot shows a database interface with a toolbar at the top and a table below. The table has three columns: 'room\_type\_reserved' (character varying (50)), 'avgnights' (numeric), and two lock icons. The data rows are: Room\_Type 1 (2.88), Room\_Type 2 (3.00), Room\_Type 4 (3.80), Room\_Type 5 (2.50), Room\_Type 6 (3.61), and Room\_Type 7 (2.67).

	room_type_reserved	avgnights
1	Room_Type 1	2.88
2	Room_Type 2	3.00
3	Room_Type 4	3.80
4	Room_Type 5	2.50
5	Room_Type 6	3.61
6	Room_Type 7	2.67

14. For reservations involving children, what is the most common room type, and what is the average price for that room type?

```
4 select room_type_reserved as most_common_room,  
5      avg(avg_price_per_room) as avg_price  
6  from hostel  
7 where no_of_children > 0  
8 group by room_type_reserved  
9 order by room_type_reserved desc limit 1;
```

Data Output    Messages    Notifications

The screenshot shows a database interface with a toolbar at the top containing various icons for file operations, a lock, download, and SQL entry. Below the toolbar is a table with two columns: 'most\_common\_room' and 'avg\_price'. The first row contains the values 'Room\_Type 7' and '187.0000000000000000'. Both columns have a lock icon next to them.

	most_common_room character varying (50)	avg_price numeric
1	Room_Type 7	187.0000000000000000

# 15. Find the market segment type that generates the highest average price per room.

```
2 SELECT
3     market_segment_type,
4     ROUND(AVG(avg_price_per_room),2) AS AvgPricePerRoom
5 FROM
6     hostel
7 GROUP BY
8     market_segment_type
9 ORDER BY
10    AvgPricePerRoom DESC
11 LIMIT 1;
```

Data Output    Messages    Notifications

The screenshot shows a database interface with a toolbar at the top containing various icons for data manipulation. Below the toolbar is a table with two columns: 'market\_segment\_type' and 'avgpriceperroom'. The first row contains the value 'Online' in the first column and '112.49' in the second column. Both columns have a lock icon next to them, indicating they are read-only or locked.

	market_segment_type	avgpriceperroom
1	Online	112.49

# ● Conclusion

**Throughout this SQL internship, the project aimed to enhance data management and analysis capabilities within the organization. The key achievements include:**

- **Development and Optimization of SQL Queries:**

- Successfully developed complex SQL queries to extract meaningful insights from large datasets.
- Implemented optimization techniques that improved query performance by approximately 30%.

- **Database Management:**

- Designed and maintained databases, ensuring data integrity and security.
- Automated data cleaning and validation processes, which reduced manual effort and errors.

- **Data Analysis and Visualization:**

- Conducted thorough data analysis to support decision-making.
- Created interactive dashboards and visualizations that provided clear insights into business operations.

# ● Key Learnings:

- Gained in-depth knowledge of SQL and its applications in real-world scenarios.
- Improved problem-solving skills by addressing and overcoming various challenges related to data handling.
- Enhanced teamwork and communication skills through collaboration with mentors and team members.

# ● Impact:

- The optimized queries and improved database management processes have significantly increased the efficiency of data retrieval and analysis.
- The visualizations and insights derived from the data have contributed to more informed and data-driven decision-making within the organization.

# Acknowledgements:

- **Sincere thanks to my teachers for their invaluable guidance and support.**
- **Appreciation to the entire team for their collaboration and assistance.**
- **Gratitude to Mentorness for providing this valuable opportunity and a conducive learning.**