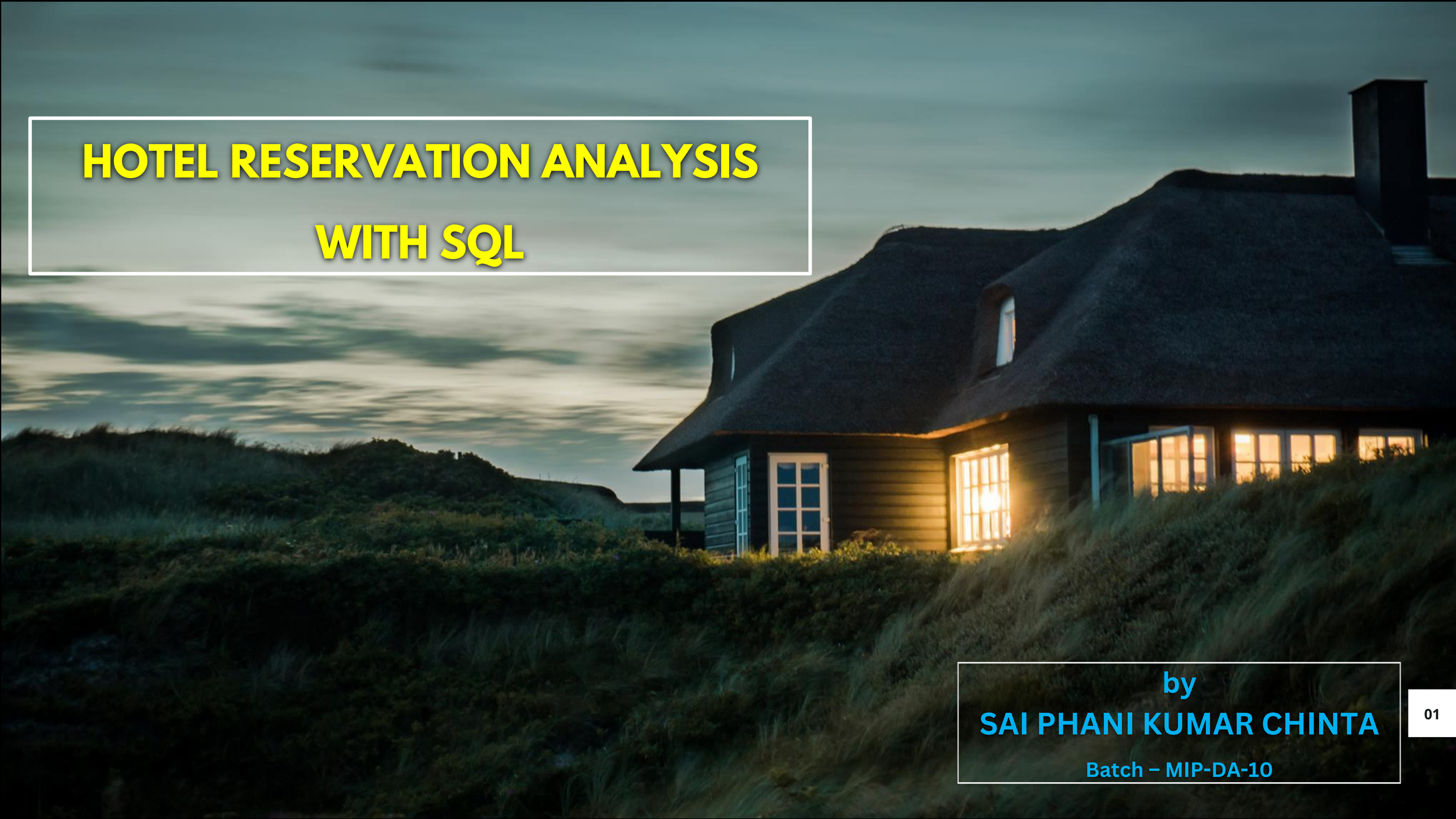
HOTEL RESERVATION ANALYSIS WITH SQL



by
SAI PHANI KUMAR CHINTA
Batch – MIP-DA-10

PROJECT OVERVIEW

- The hotel industry relies on data to make informed decisions and provide a better guest experience.

- The main objective is to perform Exploratory Data Analysis on the Hotel dataset to find the insights about the guest preferences, booking trends & other key factors influence hotel operations

- In this task i will be using MySQL Database Management system tool to analyze the data and answer the questions about dataset

DATASET DETAILS

The dataset includes the following 12 features:

- **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 reservatio belongs.
- **avg_price_per_room**: The average price per room in the reservation.
- **booking_status**: The status of the booking.

EDA QUESTIONS

1. What is the total number of reservations in the dataset?
2. Which meal plan is the most popular among guests?
3. What is the average price per room for reservations involving children?
4. How many reservations were made for the year 2018?
5. What is the most commonly booked room type?
6. How many reservations fall on a weekend (no_of_weekend_nights > 0)?
7. What is the highest and lowest lead time for reservations?
8. What is the most common market segment type for reservations?
9. How many reservations have a booking status of "Confirmed"?
10. What is the total number of adults and children across all reservations?
11. What is the average number of weekend nights for reservations involving children?
12. How many reservations were made in each month of the year?
13. What is the average number of nights (both weekend and weekday) spent by guests for each room type?
14. For reservations involving children, what is the most common room type, and what is the average price for that room type?
15. Find the market segment type that generates the highest average price per room.

DATABASE & TABLE CREATION

QUERY

Hotel_reservation solution*

Don't Limit

```sql
CREATE DATABASE Hotel_reservation;

use hotel_reservation;

create table if not exists hotel_reservation_data
(
Booking_id varchar(20) PRIMARY KEY,
no_of_adults INT,
no_of_children INT,
no_of_weekend_nights INT,
no_of_week_nights INT,
type_of_meal_plan VARCHAR(50),
room_type_reserved VARCHAR(50),
lead_time INT,
arrival_date text,
market_segment_type varchar(20),
average_price_per_room DECIMAL,
booking_status varchar(20)
);
```

SCHEMA STRUCTURE

QUERY

```
desc hotel_reservation_data;
```

RESULT

Field	Type	Null	Key	Default	Extra
Booking_id	varchar(20)	NO	PRI	NULL	
no_of_adults	int	YES		NULL	
no_of_children	int	YES		NULL	
no_of_weekend_nights	int	YES		NULL	
no_of_week_nights	int	YES		NULL	
type_of_meal_plan	varchar(50)	YES		NULL	
room_type_reserved	varchar(50)	YES		NULL	
lead_time	int	YES		NULL	
arrival_date	text	YES		NULL	
market_segment_type	varchar(20)	YES		NULL	
average_price_per_ro...	decimal(10...	YES		NULL	
booking_status	varchar(20)	YES		NULL	

IMPORTED DATASET INTO MYSQL

QUERY

```
SELECT
    *
FROM
    hotel_reservation_data;
```

RESULT

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_segment_type	average_price_per_room	booking_status
INN00001	2	0	1	2	Meal Plan 1	Room_Type 1	224	2017-10-02	Offline	65	Not_Canceled
INN00002	2	0	2	3	Not Selected	Room_Type 1	5	2018-11-06	Online	107	Not_Canceled
INN00003	1	0	2	1	Meal Plan 1	Room_Type 1	1	2018-02-28	Online	60	Canceled
INN00004	2	0	0	2	Meal Plan 1	Room_Type 1	211	2018-05-20	Online	100	Canceled
INN00005	2	0	1	1	Not Selected	Room_Type 1	48	2018-04-11	Online	95	Canceled
INN00006	2	0	0	2	Meal Plan 2	Room_Type 1	346	2018-09-13	Online	115	Canceled
INN00007	2	0	1	3	Meal Plan 1	Room_Type 1	34	2017-10-15	Online	108	Not_Canceled
INN00008	2	0	1	3	Meal Plan 1	Room_Type 4	83	2018-12-26	Online	106	Not_Canceled
INN00009	3	0	0	4	Meal Plan 1	Room_Type 1	121	2018-07-06	Offline	97	Not_Canceled
INN00010	2	0	0	5	Meal Plan 1	Room_Type 4	44	2018-10-18	Online	133	Not_Canceled
INN00011	1	0	1	0	Not Selected	Room_Type 1	0	2018-09-11	Online	85	Not_Canceled
INN00012	1	0	2	1	Meal Plan 1	Room_Type 4	35	2018-04-30	Online	140	Not_Canceled
INN00013	2	0	2	1	Not Selected	Room_Type 1	30	2018-11-26	Online	88	Canceled
INN00014	1	0	2	0	Meal Plan 1	Room_Type 1	95	2018-11-20	Online	90	Canceled
INN00015	2	0	0	2	Meal Plan 1	Room_Type 1	47	2017-10-20	Online	95	Not_Canceled
INN00016	2	0	0	2	Meal Plan 2	Room_Type 1	256	2018-06-15	Online	115	Canceled
INN00017	1	0	1	0	Meal Plan 1	Room_Type 1	0	2017-10-05	Offline	96	Not_Canceled

DATA QUERYING

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

QUERY

```
SELECT
    COUNT(booking_id) AS number_of_reservations
FROM
    hotel_reservation_data;
```

RESULT

	number_of_reservations
▶	700

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

QUERY

```
SELECT
    type_of_meal_plan,
    COUNT(booking_id) AS number_of_meal_bookings
FROM
    hotel_reservation_data
GROUP BY type_of_meal_plan
ORDER BY number_of_meal_bookings DESC
LIMIT 1;
```

RESULT

type_of_meal_plan	number_of_meal_bookings
Meal Plan 1	527

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

QUERY

```
SELECT
    ROUND(AVG(average_price_per_room), 2) AS avg_price_per_room
FROM
    hotel_reservation_data
WHERE
    no_of_children > 0;
```

RESULT

	avg_price_per_room
▶	144.63

4. How many reservations were made for the year 2018?

QUERY

```
SELECT
    YEAR(arrival_date) AS year_of_arrival,
    COUNT(booking_id) AS number_of_bookings
FROM
    hotel_reservation_data
WHERE
    YEAR(arrival_date) = 2018
GROUP BY YEAR(arrival_date);
```

RESULT

	year_of_arrival	number_of_bookings
▶	2018	577

5. What is the most commonly booked room type?

QUERY

```
SELECT
    room_type_reserved,
    COUNT(booking_id) AS number_of_bookings
FROM
    hotel_reservation_data
GROUP BY room_type_reserved
ORDER BY COUNT(booking_id) DESC
LIMIT 1;
```

RESULT

	room_type_reserved	number_of_bookings
▶	Room_Type 1	534

6. How many reservations fall on a weekend (no_of_weekend_nights > 0)?

QUERY

```
SELECT
    COUNT(no_of_weekend_nights) as no_of_reservations_on_weekend
FROM
    hotel_reservation_data
WHERE
    no_of_weekend_nights > 0;
```

RESULT

	no_of_reservations_on_weekend
▶	383

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

QUERY

```
SELECT
    MAX(lead_time) AS maximum_lead_time,
    MIN(lead_time) AS minimum_lead_time
FROM
    hotel_reservation_data;
```

RESULT

	maximum_lead_time	minimum_lead_time
▶	443	0

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

QUERY

```
SELECT
    market_segment_type,
    COUNT(booking_id) as no_of_reservations
FROM
    hotel_reservation_data
GROUP BY market_segment_type
ORDER BY COUNT(booking_id) DESC
limit 1;
```

RESULT

	market_segment_type	no_of_reservations
▶	Online	518

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

QUERY

```
SELECT
    COUNT(Booking_id) AS Confirmed_reservations
FROM
    hotel_reservation_data
WHERE
    booking_status = 'Not_canceled';
```

RESULT

Confirmed_reservations
493

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

QUERY

```
SELECT
    SUM(no_of_adults) as number_of_adults,
    SUM(no_of_children) AS number_of_children,
    SUM(no_of_adults) + SUM(no_of_children) as total_members
FROM
    hotel_reservation_data;
```

RESULT

	number_of_adults	number_of_children	total_members
▶	1316	69	1385

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

QUERY

```
SELECT
    Round(AVG(no_of_weekend_nights))as no_of_weekend_nights_children
FROM
    hotel_reservation_data
WHERE
    no_of_children > 0;
```

RESULT

	no_of_weekend_nights_children
▶	1

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

QUERY

```
SELECT
    EXTRACT(MONTH FROM arrival_date) as Month_no,
    DATE_FORMAT(arrival_date, '%M') AS Month_name,
    COUNT(booking_id) AS no_of_reservations
FROM hotel_reservation_data
GROUP BY month_name,Month_no
ORDER BY Month_no;
```

RESULT

Month_no	Month_name	no_of_reservations
1	January	11
2	February	28
3	March	52
4	April	67
5	May	55
6	June	84
7	July	44
8	August	70
9	September	80
10	October	103
11	November	54
12	December	52

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

QUERY

```
SELECT
    room_type_reserved AS room_type,
    ROUND(AVG(no_of_week_nights) + AVG(no_of_weekend_nights),2) AS avg_no_of_nights
FROM
    hotel_reservation_data
GROUP BY room_type_reserved
ORDER BY room_type;
```

RESULT

room_type	avg_no_of_nights
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
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?

QUERY

```
SELECT
    room_type_reserved as room_type,
    COUNT(Booking_id) AS no_of_reservations,
    Round(AVG(average_price_per_room),2) AS avg_price
FROM
    hotel_reservation_data
WHERE
    no_of_children > 0
GROUP BY room_type_reserved
ORDER BY no_of_reservations desc
limit 1;
```

RESULT

room_type	no_of_reservations	avg_price
Room_Type 1	24	123.21

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

QUERY

```
SELECT
    market_segment_type,
    ROUND(AVG(average_price_per_room),2) as avg_room_price
FROM
    hotel_reservation_data
GROUP BY market_segment_type
ORDER BY avg_room_price desc
LIMIT 1;
```

RESULT

	market_segment_type	avg_room_price
▶	Online	112.49

THANK YOU