

Make My Trip

# Data Analysis Case Study

## DATA UNDERSTANDING

#### There are two tables

- 1) booking\_table: contains information about bookings made by users.
- 2) user\_table: contains information about users.

	booking_table				
Booking_id	Booking_date	User_id	Line_of_business		
b1	23-03-2022	u1	Flight		
b2	27-03-2022	u2	Flight		
b3	28-03-2022	u1	Hotel		
b4	31-03-2022	u4	Flight		
b5	02-04-2022	u1	Hotel		
b6	02-04-2022	u2	Flight		
b7	06-04-2022	u5	Flight		
b8	06-04-2022	u6	Hotel		
b9	06-04-2022	u2	Flight		
b10	10-04-2022	u1	Flight		
b11	12-04-2022	u4	Flight		
b12	16-04-2022	u1	Flight		
b13	19-04-2022	u2	Flight		
b14	20-04-2022	u5	Hotel		
b15	22-04-2022	u6	Flight		
b16	26-04-2022	u4	Hotel		
b17	28-04-2022	u2	Hotel		
b18	30-04-2022	u1	Hotel		
b19	04-05-2022	u4	Hotel		
b20	06-05-2022	u1	Flight		

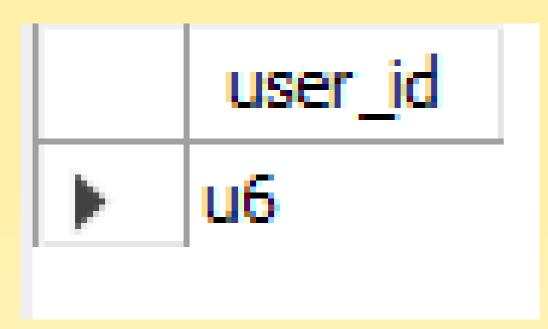
user_table		
User_id	Segment	
<b>u1</b>	<b>s1</b>	
u2	<b>s1</b>	
u3	<b>s1</b>	
u4	<b>s</b> 2	
u5	s2	
u6	<b>s</b> 3	
u7	s3	
u8	s3	
u9	s3	
u10	s3	

 Write a query to get the following output.
 Segment, total user count, total users who booked flight in April 2022.

	segment	total_user_count	user_who_booked_flight_in_apr2022
•	s1	3	2
	s2	2	2
	s3	5	1

 Write a query to find the users whose first booking was a hotel booking.

```
WITH cte AS (
SELECT
    *,
    RANK() OVER(PARTITION BY user_id ORDER BY booking_date)
    AS booking_rank
FROM booking_table
)
SELECT
    DISTINCT(user_id)
FROM cte
WHERE booking_rank = 1 AND line_of_business = 'Hotel';
```



 Write a query to find the days between first boking and last booking for each user.

```
SELECT

user_id,

DATEDIFF(MAX(booking_date),MIN(booking_date))

AS days_diff

FROM booking_table

GROUP BY user_id;
```

	user_id	days_diff
<b>&gt;</b>	u1	44
	u2	32
	u4	34
	u5	14
	u6	16

 Write a query to count the number of flight and hotel bookings for each user segment in 2022.

```
SELECT
    segment,
    SUM(CASE WHEN line_of_business = 'Flight' THEN 1 ELSE 0 END)
    AS flight_bookings,
    SUM(CASE WHEN line_of_business = 'Hotel' THEN 1 ELSE 0 END)
    AS hotel_bookings
FROM user_table ut
INNER JOIN booking_table bt
    ON ut.user_id = bt.user_id
WHERE YEAR(booking_date) = 2022
GROUP BY segment;
```

	segment	flight_bookings	hotel_bookings
•	s1	8	4
	s2	3	3
	s3	1	1