SQL TIPS AND TRICKS

PART 30

Case Study

On Travel Company

MAYURI DANDEKAR

Example--

booking

booking_id	booking_date	user_id	line_of_business
b1	2022-03-23	u1	flight
b2	2022-03-27	u2	flight
b3	2022-03-28	u1	hotel
b4	2022-03-31	u4	flight
b5	2022-04-02	u1	hotel
b6	2022-04-02	u2	flight
b7	2022-04-06	u5	flight
b8	2022-04-06	u6	hotel
b9	2022-04-06	u2	flight
b10	2022-04-10	u1	flight
b11	2022-04-12	u4	flight
b12	2022-04-16	u1	flight
b13	2022-04-19	u2	flight
b14	2022-04-20	u5	hotel
b15	2022-04-22	u6	flight
b16	2022-04-26	u4	hotel
b17	2022-04-28	u2	hotel
b18	2022-04-30	u1	hotel
b19	2022-05-04	u1	hotel
b20	2022-05-06	u1	flight

user

user_id	segment
u1	s1
u2	s1
u3	s1
u4	s2
u5	s2
u6	s3
u7	s3
u8	s3
u9	s3
u10	s3

Write a query that gives below output

segment	no_of_users	user_who_booked_flight_in_apr2022
s1	3	2
s2	2	2
s3	5	1

```sql
SELECT u.segment, COUNT(DISTINCT u.user_id) AS no_of_users,
COUNT(DISTINCT CASE WHEN b.line_of_business='flight' AND
        b.booking_date BETWEEN '2022-04-01' AND '2022-04-30'
        THEN b.user_id END) AS user_who_booked_flight_in_apr2022
FROM user u
LEFT JOIN booking b ON u.user_id = b.user_id
GROUP BY u.segment;
```

Write a query to identify users whose first booking was hotel booking

```
36      -- method 1
37  •   SELECT * FROM (
38          SELECT *,
39          RANK() OVER(PARTITION BY user_id ORDER BY booking_date) AS rn
40          FROM booking) rank_users
41      WHERE rn=1 AND line_of_business='hotel';
42
```

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

booking_id	booking_date	user_id	line_of_business	rn
b8	2022-04-06	u6	hotel	1

```sql
-- method 2
SELECT DISTINCT user_id FROM (
    SELECT *,
    FIRST_VALUE(line_of_business) OVER(PARTITION BY user_id ORDER BY booking_date) AS first_booking
    FROM booking) book_users
WHERE first_booking='hotel';
```

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

user_id
u6

Write a query to calculate days between first and last booking of each user

```sql
SELECT user_id, MIN(booking_date) AS first_booking,
    MAX(booking_date) AS last_booking,
    DATEDIFF(MAX(booking_date), MIN(booking_date)) AS no_of_days
FROM booking
GROUP BY user_id;
```

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

user_id	first_booking	last_booking	no_of_days
u1	2022-03-23	2022-05-06	44
u2	2022-03-27	2022-04-28	32
u4	2022-03-31	2022-04-26	26
u5	2022-04-06	2022-04-20	14
u6	2022-04-06	2022-04-22	16

Write a query to find the No. of flight and hotel bookings in each of the user segments for the year 2022

```
58      -- by segment
59  •   SELECT segment,
60      SUM(CASE WHEN line_of_business='flight' THEN 1 ELSE 0 END) AS flight_booking,
61      SUM(CASE WHEN line_of_business='hotel' THEN 1 ELSE 0 END) AS hotel_booking
62      FROM booking b
63      INNER JOIN user u ON b.user_id=u.user_id
64      GROUP BY segment;
```

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

segment	flight_booking	hotel_booking
s1	8	5
s2	3	2
s3	1	1

```sql
-- by user_id
SELECT user_id,
SUM(CASE WHEN line_of_business='flight' THEN 1 ELSE 0 END) AS flight_booking,
SUM(CASE WHEN line_of_business='hotel' THEN 1 ELSE 0 END) AS hotel_booking
FROM booking b
-- INNER JOIN user u ON b.user_id=u.user_id
GROUP BY user_id;
```

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

user_id	flight_booking	hotel_booking
u1	4	4
u2	4	1
u4	2	1
u5	1	1
u6	1	1

THANK YOU

MAYURI DANDEKAR