```
WITH session_based AS (
                SELECT s.session_id
        ,s.user_id
        ,s.trip_id
        ,s.session_start
        ,s.session_end
        ,s.page_clicks
        ,s.flight_discount
        ,s.flight_discount_amount
        ,s.hotel_discount
        ,s.hotel_discount_amount
        ,s.flight_booked
        ,s.hotel_booked
        ,s.cancellation
        ,EXTRACT( EPOCH FROM (s.session_end-s.session_start)) AS
session_duration_per_sec
        ,f.origin_airport
        ,f.destination
        ,f.destination_airport
        ,f.seats
        ,f.return_flight_booked
        ,f.departure_time
        ,f.return_time
        ,f.checked_bags
                                        ,f.trip_airline
        ,f.destination_airport_lat
        ,f.destination_airport_lon
        ,f.base_fare_usd
        ,h.hotel_name
        ,CASE WHEN h.nights < 0 THEN ABS(h.nights)
                                                WHEN h.nights = 0 THEN 1
                                                ELSE h.nights END AS num_nights
        ,h.rooms
        ,h.check_in_time
        ,h.check_out_time
        ,h.hotel_per_room_usd AS hotel_price_per_room_night_usd
                        ,u.home_airport_lat
                        ,u.home_airport_lon
                        ,EXTRACT(MONTH FROM departure_time) AS departure_month
                FROM sessions s
                LEFT JOIN users u
                ON s.user_id = u.user_id
                LEFT JOIN flights f
                ON s.trip_id = f.trip_id
```

```
                    LEFT JOIN hotels h
                    ON s.trip_id = h.trip_id
                    WHERE s.user_id IN (SELECT user_id
                FROM sessions
                WHERE session_start > '2023-01-04'
                GROUP BY user_id
                HAVING COUNT(*) > 7)

)
, session_user_based AS(
                    SELECT        user_id
                                                    ,SUM(page_clicks)     AS
num_clicks
                    ,COUNT(distinct session_id)  AS num_sessions
                    ,ROUND(AVG(session_duration_per_sec)::numeric,2) AS
avg_session_duration
                            FROM session_based
                            GROUP BY user_id
),
trip_based AS(
                            SELECT user_id
                                    ,COUNT(trip_id) as num_trips
                                    ,SUM(CASE WHEN flight_booked AND
return_flight_booked THEN 2
                                                                    WHEN
flight_booked THEN 1
                                                                    ELSE 0
                                                            END) AS num_flights
                                    ,SUM(base_fare_usd) AS original_flight_fare
                                    ,SUM((base_fare_usd) * (1 -
COALESCE(flight_discount_amount,0))) AS discounted_flight_fare
                                    ,COUNT(hotel_booked) FILTER (WHERE
hotel_booked = 'true') AS num_booked_hotel
                                    ,SUM(hotel_price_per_room_night_usd *
num_nights*rooms) AS original_hotel_fare
                                    ,SUM((hotel_price_per_room_night_usd *
num_nights*rooms) * (1 - COALESCE(hotel_discount_amount,0))) AS discounted_hotel_fare
                                    ,AVG(EXTRACT(DAY FROM departure_time -
session_end)) AS avg_time_before_trip
                                    ,SUM(haversine_distance(home_airport_lat,
home_airport_lon, destination_airport_lat, destination_airport_lon)) AS distance_flown
                                                    ,SUM((hotel_price_per_room_night_usd *
num_nights*rooms) * (1 - COALESCE(hotel_discount_amount,0)))
```

```
                                        + SUM((base_fare_usd) * (1 -
COALESCE(flight_discount_amount,0))) AS total_booking_price
                                        ,COUNT(flight_discount) FILTER (WHERE
flight_discount = 'true') AS flight_discount
                                        ,COUNT(hotel_discount) FILTER (WHERE
hotel_discount = 'true') AS hotel_discount
                                    ,ROUND(AVG(flight_discount_amount),2) AS
avg_flight_discount
                                    ,ROUND(AVG(hotel_discount_amount),2) AS
avg_hotel_discount
                                        ,SUM(num_nights) AS total_nights
                                        ,SUM(rooms) AS total_rooms
                                FROM session_based
                            WHERE trip_id IS NOT NULL
                                        AND trip_id NOT IN (SELECT distinct trip_id
                                                FROM session_based
                                                WHERE cancellation)
                            GROUP BY user_id
),
main_query AS (
                                SELECT sub.*
                                                ,EXTRACT(YEAR FROM age(now(),
u.birthdate)) AS age
                    ,u.gender
                    ,u.married
                    ,u.has_children
                    ,u.home_country
                    ,u.home_city
                    ,EXTRACT(YEAR FROM age(now(), u.sign_up_date)) AS time_spent
                    ,t.num_trips
                    ,t.num_flights
                                        ,t.original_flight_fare
                                        ,t.discounted_flight_fare
                                        ,t.flight_discount
                                        ,t.num_booked_hotel
                                        ,t.total_nights
                                        ,t.total_rooms
                                        ,t.original_hotel_fare
                    ,t.discounted_hotel_fare
                    ,t.distance_flown
                                        ,t.total_booking_price
                                        ,t.hotel_discount
                                        ,t.avg_flight_discount
                                        ,t.avg_hotel_discount
```

```
                                ,t.avg_time_before_trip
                        FROM users u
                        LEFT JOIN session_user_based AS sub
                        ON sub.user_id = u.user_id
                        JOIN trip_based t
                        ON sub.user_id = t.user_id)


SELECT      *
                        , CASE      WHEN age >= 60 THEN 'Senior Traveller'
                        WHEN has_children THEN 'Family Traveller'
          WHEN num_trips <= 2 THEN 'Dreamer'
          WHEN age < 30 AND num_trips > 2 THEN 'Young frequent traveller'
          WHEN age >= 30 AND num_trips BETWEEN 2 AND 5 THEN 'Middle-aged frequent
traveller'
          WHEN age >= 30 AND num_trips > 5 THEN 'Business traveller'
          ELSE 'others'
       END AS groups
     ,CASE WHEN age BETWEEN 17 AND 29 THEN '17-29'
                     WHEN age BETWEEN 30 AND 59 THEN '30-59'
        ELSE '60+'
       END AS age_bucket
     ,CASE WHEN (flight_discount + hotel_discount) > 5 THEN 'Bargain Wizard'
                     WHEN (flight_discount + hotel_discount) BETWEEN 1 AND 5 THEN
'Medium bargainer'
        ELSE 'Other'
       END AS discount_bucket
FROM main_query
;
```