Marielle Kristine de la Viña

Homework! Work with: *tutorial.flights.*

Task 1: Categorize Arrival Delays

Question: Use a CTE to create a new column called `arrival_delay_category` using a `CASE WHEN` statement with the following rules:

- `'Early'` if `arrival_delay` is less than 0
- `'On Time'` if `arrival_delay` is exactly 0
- `'Late'` if `arrival_delay` is greater than 0

Then, using the CTE, return the count of flights in each delay category.

```
WITH with_delay_category as (
select arrival_delay,
(CASE
  WHEN arrival_delay > 0 then 'Late'
  WHEN arrival_delay < 0 then 'Early'
  WHEN arrival_delay = 0 then 'On Time'
  WHEN arrival_delay is NULL then 'No Data'
END) AS arrival_delay_category
from tutorial.flights
)

select arrival_delay_category, count(*) as flights
from with_delay_category
group by arrival_delay_category
order by flights desc
```

Task 2: Classify Flights by Distance

Question: Using a `CASE WHEN` statement inside a CTE, classify each flight by distance into a column called `distance_category`:

- `'Short-haul'` if `distance` is less than 800
- `'Medium-haul'` if `distance` is between 800 and 2200
- `'Long-haul'` if `distance` is greater than 2200

From the CTE, return the number of flights and average air time per `distance_category`.

```
WITH with_distance_category as (
select distance,air_time,
(CASE
  WHEN distance is NULL then 'No Data'
  WHEN distance < 800 then 'short-haul'
  WHEN distance > 2200 then 'long-haul'
  WHEN distance between 800 and 2200 then 'medium-haul'
END) AS distance_category
from tutorial.flights
)

select distance_category, count(*) as flights, avg(air_time) as avg_air_time
from with_distance_category
group by distance_category
```

Task 3: Cancellations by Reason

Question: Create a CTE that uses a `CASE WHEN` statement to categorize the `cancellation_reason` column into:

- `'Weather'`
- `'Carrier'`
- `'Security'`
- `'Other'` for any other values or NULLs

Filter the dataset to only include cancelled flights. From the CTE, return the number of cancellations by reason category.

```
WITH with_cancellation_category as (
select cancellation_reason,
(CASE
  when upper(cancellation_reason) like '%WEATHER%' then 'WEATHER'
  when upper(cancellation_reason) like '%CARRIER%' then 'CARRIER'
  when upper(cancellation_reason) like '%SECURITY%' then 'SECURITY'
  else 'Other'
END) as cancellation_category
from tutorial.flights
where was_cancelled is true
)

select cancellation_category, count(*) as flights
from with_cancellation_category
group by cancellation_category
order by flights desc
```

Task 4: Departure Delay Severity

Create a CTE that adds a column called `departure_delay_severity` using a `CASE WHEN` statement:

- `'None'` for 0 minutes
- `'Minor'` for 1–15 minutes
- `'Moderate'` for 16–60 minutes
- `'Severe'` for more than 60 minutes

Then return the total number of flights in each severity category.

```
WITH with_delay_category as (
select departure_delay,
(CASE
  WHEN departure_delay is NULL THEN 'No Data'
  WHEN departure_delay <= 0 THEN 'None'
  WHEN departure_delay BETWEEN 1 AND 15 THEN 'Minor'
  WHEN departure_delay BETWEEN 16 AND 60 THEN 'Moderate'
  WHEN departure_delay > 60 THEN 'Severe'
END) as delay_category
from tutorial.flights
)

select delay_category, count(*) as flights
from with_delay_category
group by delay_category
order by flights desc
```

Task 5: Rate Airline Delay Performance

Question: First, create a CTE that calculates the **average arrival delay per airline**.
 Then, using a second CTE or inside the same one, use a `CASE WHEN` statement to create a new column called `delay_rating`:

- `'Excellent'` if the average delay is less than 0
- `'Good'` if it's between 0 and 10
- `'Poor'` if it's greater than 10

Finally, return each airline's name, average arrival delay, and delay rating.

```
with with_avg_delay as (
  select airline_name, avg(arrival_delay) as avg_arrival_delay,
  (CASE
  WHEN avg(arrival_delay) < 0 THEN 'Excellent'
  WHEN avg(arrival_delay)  BETWEEN 0 and 10 THEN 'Good'
  WHEN avg(arrival_delay) >10 THEN 'Poor'
  END) as delay_rating
  from tutorial.flights
  group by airline_name
)

select *
from with_avg_delay
```

Task 6: Departure Delay Severity

Use a CTE to create a new column called `multi_delay_flag` using a `CASE WHEN` statement. This column should be:

- `'Yes'` if the flight has more than one of the following delay types greater than 0: `weather_delay`, `carrier_delay`, `late_aircraft_delay`, `nas_delay`, or `security_delay`
- `'No'` otherwise

Then return the count of flights that fall into each category.

```
WITH with_delay_flags as (
select flight_number,
(CASE
  WHEN weather_delay > 0 then 1
  else 0
END) as weather_delay_flag,
(CASE
  WHEN carrier_delay > 0 then 1
  else 0
END) as carrier_delay_flag,
(CASE
  WHEN late_aircraft_delay > 0 then 1
  else 0
END) as late_aircraft_delay_flag,
(CASE
  WHEN security_delay > 0 then 1
  else 0
END) as security_delay_flag
from tutorial.flights
)

select
(CASE
  WHEN
(weather_delay_flag+carrier_delay_flag+late_aircraft_delay_flag+security_delay_flag) >=
2 then 'Yes'
  ELSE 'No'
END) as multi_delay_flag,
count(*) as flights
from with_delay_flags
group by multi_delay_flag
```