SQL Capstone Project

Please answer the following questions using Airline DB database.

Instruction to attempt questions:

- Students need to write queries for the questions mentioned in the using Airline DB database
- Read the questions carefully before writing the query in **Airline Playground** (in the Playground chapter of SQL)
- Airline DB: https://www.skillovilla.com/playground/sql?exerciseId=0181e251-6ea8-4595-ae2b-0c690119f8db

How to submit the capstone:

- Copy the SQL query code and paste it in the answer section in this file.
- Once the assignment is done, submit the file over LMS.

Invalid Submissions:

- Pasting pictures of the code as answer is **NOT** acceptable.
- Uploading output data (CSVs) of the SQL queries is **NOT** acceptable.

Write your answers(query) in the answer and submit it. To write the answer in the assignment, please follow the below example in yellow

Example:

Questions*: Extract all the columns of the flights table*

Answer: *SELECT * FROM flights*

SQL Capstone Project

Attempt the following Questions-

1. ***Represent the "book_date" column in "yyyy-mmm-dd" format using Bookings table***

   Expected output: book_ref, book_date (in "yyyy-mmm-dd" format) , total amount

Answer:

```
SELECT
book_ref,
TO_CHAR (book_date, 'yyyy-mon-dd') as book_date, total_amount
FROM bookings
```

2. **Get the following columns in the exact same sequence.**
   Expected columns in the output: ticket_no, boarding_no, seat_number, passenger_id, passenger_name.

Answer:

```
SELECT
 bp.ticket_no, boarding_no, seat_no, passenger_id, passenger_name
 from boarding_passes bp
 JOIN tickets t
 ON bp.ticket_no = t.ticket_no
```

3. **Write a query to find the seat number which is least allocated among all the seats?**

Answer:

```
WITH T1 AS (
SELECT s.seat_no, COUNT(ticket_no) as ticket_no
FROM boarding_passes bp
LEFT JOIN seats s
ON s.seat_no = bp.seat_no
GROUP BY 1)

SELECT
seat_no
FROM T1
WHERE ticket_no = (SELECT MIN(ticket_no)FROM T1)
```

4. ***In the database, identify the month wise highest paying passenger name and passenger id.***

Expected output: Month_name("mmm-yy" format), passenger_id, passenger_name and total amount

Answer:

```
WITH T1 AS (
SELECT TO_CHAR(book_date, 'mon-yy') as Month_name, passenger_id,
passenger_name, total_amount
from bookings b
join tickets t
on b.book_ref = t.book_ref
group by 1,2,3,4), T2 AS (

SELECT
Month_name, passenger_id,passenger_name,total_amount,
DENSE_RANK() OVER(PARTITION BY Month_name ORDER BY total_amount
DESC) AS RANK_1
FROM T1 )
SELECT
Month_name, passenger_id, passenger_name,total_amount
FROM T2
WHERE RANK_1 = 1
```

5. ***In the database, identify the month wise least paying passenger name and passenger id?***

Expected output: Month_name("mmm-yy" format), passenger_id, passenger_name and total amount

Answer:

```
WITH T1 AS (
SELECT TO_CHAR(book_date, 'mon-yy') as Month_name, passenger_id,
passenger_name, total_amount
from bookings b
join tickets t
on b.book_ref = t.book_ref
group by 1,2,3,4), T2 AS (
```

```
SELECT
Month_name, passenger_id,passenger_name,total_amount,
DENSE_RANK() OVER(PARTITION BY Month_name ORDER BY total_amount ASC)
AS RANK_1
FROM T1 )
SELECT
Month_name, passenger_id, passenger_name,total_amount
FROM T2
WHERE RANK_1 = 1
```

6. **Identify the travel details of non stop journeys or return journeys (having more than 1 flight).**

Expected Output: Passenger_id, passenger_name, ticket_number and flight count.

Answer:

```
SELECT
Passenger_id,
passenger_name,
t.ticket_no,
COUNT(tf.flight_id)flight_count
FROM tickets t
JOIN BOARDING_PASSES BP
ON T.TICKET_NO = BP.TICKET_NO
HAVING COUNT(tf.flight_id)>1
ORDER BY 1 ASC
```

7. **How many tickets are there without boarding passes?**

Expected Output: just one number is required.

Answer

```
SELECT
 COUNT(*)
 FROM tickets t
 LEFT JOIN boarding_passes bp
 ON t.ticket_no = bp.ticket_no
 WHERE boarding_no IS NULL
```

8. **Identify details of the longest flight (using flights table)?**
Expected Output: Flight number, departure airport, arrival airport, aircraft code and durations.

Answer:

```
with t1 as (
SELECT
 flight_no, departure_airport, arrival_airport, aircraft_code,
 (scheduled_arrival - scheduled_departure) as duration
 FROM flights
 )
, t2 as (
 SELECT
  flight_no, departure_airport, arrival_airport, aircraft_code,
duration,
  DENSE_RANK() OVER(ORDER BY duration DESC) as rank_1
  from t1)

SELECT
 flight_no, departure_airport, arrival_airport, aircraft_code,
duration
 FROM t2
 WHERE rank_1 = 1
 ORDER BY 5 DESC
```

9. **Identify details of all the morning flights (morning means between 6AM to 11 AM, using flights table)?**
Expected output: flight_id, flight_number, scheduled_departure, scheduled_arrival and timings.

Answer:

```
SELECT
flight_id,
flight_no,
scheduled_departure,
scheduled_arrival,
cast(scheduled_departure as time) timings
FROM flights
WHERE CAST(scheduled_departure as time) BETWEEN '06:00:00' AND
'10:59:59'
```

10. Identify the earliest morning flight available from every airport.
Expected output: flight_id, flight_number, scheduled_departure, scheduled_arrival, departure airport and timings.

Answer:

```
 WITH  T1 AS (
    SELECT
    flight_id, flight_no, scheduled_departure, scheduled_arrival,
departure_airport,
    CAST(scheduled_departure as time) timings,
    DENSE_RANK () OVER (PARTITION BY DEPARTURE_AIRPORT ORDER BY
CAST(scheduled_departure as time) ASC) RANK_1
    FROM FLIGHTS

 )
 SELECT
  flight_id, flight_no, scheduled_departure, scheduled_arrival,
departure_airport, timings
  FROM T1
  WHERE CAST(scheduled_departure as time) between '02:00:00' and
'05:59:59' and RANK_1 = 1
Order by 6
```

11. Questions: Find list of airport codes in Europe/Moscow timezone
Expected Output: Airport_code.

Answer:

```
SELECT
  airport_code
FROM airports
WHERE timezone = 'Europe/Moscow'
```

12. Write a query to get the count of seats in various fare condition for every aircraft code?
Expected Outputs: Aircraft_code, fare_conditions ,seat count

Answer:

```
SELECT
 aircraft_code, fare_conditions, COUNT(seat_no) as seat_count
 FROM seats
GROUP BY 1,2
ORDER BY 1,2
```

13. How many aircrafts codes have at least one Business class seats?
Expected Output: Count of aircraft codes

Answer:

```
SELECT
COUNT (DISTINCT a.aircraft_code)
 FROM aircrafts a
 JOIN  seats s
 ON a.aircraft_code = s.aircraft_code
 WHERE fare_conditions = 'Business'
```

14. Find out the name of the airport having maximum number of departure flight
Expected Output : Airport_name

Answer:

```
WITH T1 AS (
SELECT airport_name, COUNT(departure_AIRPORT) AS total_departures,
RANK()OVER(ORDER BY COUNT(DEPARTURE_AIRPORT)DESC) R1
FROM AIRPORTS a
join FLIGHTS f
on a.airport_code=f.departure_AIRPORT
GROUP BY airport_name)

SELECT
airport_name
FROM T1
WHERE R1=1
```

15. Find out the name of the airport having least number of scheduled departure flights
Expected Output : Airport_name

Answer:

```
WITH T1 AS (
SELECT airport_name, COUNT(departure_AIRPORT) AS total_departures,
RANK()OVER(ORDER BY COUNT(DEPARTURE_AIRPORT)ASC) R1
FROM AIRPORTS a
join FLIGHTS f
on a.airport_code=f.departure_AIRPORT
GROUP BY airport_name)

SELECT
airport_name
FROM T1
WHERE R1=1
```

16. How many flights from 'DME' airport don't have actual departure?
Expected Output : Flight Count

Answer:

```
SELECT
 COUNT(flight_id) as flight_count
FROM flights
WHERE departure_airport = 'DME' AND actual_departure IS NULL
```

17. Identify flight ids having range between 3000 to 6000
Expected Output : Flight_Number , aircraft_code, ranges

Answer:

```
SELECT
 flight_no, f.aircraft_code, range
 FROM flights f
 JOIN aircrafts a
 ON f.aircraft_code = a.aircraft_code
 WHERE range BETWEEN 3000 AND 6000
```

18. Write a query to get the count of flights flying between URS and KUF?
Expected Output : Flight_count

Answer:

```
SELECT count(*) flight_count
FROM flights
WHERE (departure_airport='KUF' AND arrival_airport='URS') OR
 (departure_airport= 'URS' AND arrival_airport='KUF')
```

19. Write a query to get the count of flights flying from either from NOZ or KRR?
Expected Output : Flight count

Answer:

```
 SELECT
COUNT(flight_no) FLIGHT_COUNT
FROM flights
WHERE departure_airport = 'NOZ' OR departure_airport = 'KRR'
```

20. Write a query to get the count of flights flying from KZN,DME,NBC,NJC,GDX,SGC,VKO,ROV
Expected Output : Departure airport ,count of flights flying from these airports.

Answer:

```
SELECT
Departure_airport,
COUNT(flight_id) flight_count
FROM flights
WHERE departure_airport in
('KZN','DME','NBC','NJC','GDX','SGC','VKO','ROV')
GROUP BY 1
```

21. **Write a query to extract flight details having range between 3000 and 6000 and flying from DME**

Expected Output :Flight_no,aircraft_code,range,departure_airport

Answer:

```
SELECT
flight_no, f.aircraft_code, range, departure_airport
FROM flights f
JOIN aircrafts a
ON f.aircraft_code=a.aircraft_code
where range between 3000 AND 6000 AND departure_airport='DME'
```

22. **Find the list of flight ids which are using aircrafts from "Airbus" company and got cancelled or delayed**

Expected Output : Flight_id,aircraft_model

Answer:

```
SELECT
flight_id, model as aircraft_model
FROM flights f
JOIN aircrafts a
ON f.aircraft_code=a.aircraft_code
WHERE model like '%Airbus%' and (F.status = 'Cancelled' or F.status
= 'Delayed')
```

23. **Find the list of flight ids which are using aircrafts from "Boeing" company and got cancelled or delayed**

Expected Output : Flight_id,aircraft_model

Answer:

```
SELECT
 F.FLIGHT_ID, A.MODEL AS AIRCRAFT_MODEL
 FROM FLIGHTS F
 JOIN AIRCRAFTS A
 ON F.AIRCRAFT_CODE = A.AIRCRAFT_CODE
where A.model like '%Boeing%' and (F.status = 'Cancelled' or
F.status = 'Delayed')
```

24. Which airport(name) has most cancelled flights (arriving)?
Expected Output : Airport_name

Answer:

```
with t1 as (
    SELECT airport_name, rank() over(partition by airport_name order
by count(flight_id) desc)as r1
from flights f
JOIN airports a
ON f.arrival_airport=a.airport_code
WHERE status = 'Cancelled'
group by 1
)
select airport_name
from t1
where r1=1
```

25. Identify flight ids which are using "Airbus aircrafts"
Expected Output : Flight_id,aircraft_model

Answer:

```

SELECT
flight_id, model AS AIRCRAFT_MODEL
from flights f
join AIRCRAFTS a
ON f.aircraft_code=a.aircraft_code
WHERE model like '%Airbus%'
```

26. ***Identify date-wise last flight id flying from every airport?***
Expected Output: Flight_id,flight_number,schedule_departure,departure_airport

Answer:

```
WITH T1 AS (
     SELECT
      FLIGHT_ID, FLIGHT_NO, SCHEDULED_DEPARTURE , DEPARTURE_AIRPORT,
      ROW_NUMBER() OVER(PARTITION BY DEPARTURE_AIRPORT ORDER BY
SCHEDULED_DEPARTURE DESC) AS RANK_1
    FROM FLIGHTS)

SELECT
 FLIGHT_ID, FLIGHT_NO, SCHEDULED_DEPARTURE, DEPARTURE_AIRPORT
 FROM T1
 WHERE RANK_1 = 1
```

27. ***Identify list of customers who will get the refund due to cancellation of the flights and how much amount they will get?***
Expected Output : Passenger_name,total_refund.

Answer:

```
SELECT passenger_name, SUM(amount) total_refund
FROM tickets t
JOIN ticket_flights tf
ON t.ticket_no=tf.ticket_no
JOIN flights f
ON f.flight_id=tf.flight_id
WHERE f.STATUS = 'Cancelled'
GROUP BY 1
```

28. *Identify date wise first cancelled flight id flying for every airport?*
Expected Output : Flight_id,flight_number,schedule_departure,departure_airport

Answer:

```
WITH FC AS(
SELECT flight_id,flight_no,scheduled_departure,departure_airport,
ROW_NUMBER() OVER(PARTITION BY departure_airport ORDER BY
SCHEDULED_DEPARTURE ASC) AS RNK
from flights
WHERE status='Cancelled'
)
SELECT
flight_id,flight_no,scheduled_departure,departure_airport
FROM FC
WHERE RNK=1
```

29. *Identify list of Airbus flight ids which got cancelled.*
Expected Output : Flight_id

Answer:

```
SELECT flight_id
FROM FLIGHTS F
JOIN AIRCRAFTS A
ON F.AIRCRAFT_CODE=A.AIRCRAFT_CODE

WHERE MODEL LIKE '%Airbus%' and Status = 'Cancelled'
```

30. ***Identify list of flight ids having highest range.***
Expected Output : Flight_no, range

Answer:

```
WITH T1 AS (
SELECT FLIGHT_NO, range,
RANK()OVER(PARTITION BY FLIGHT_ID ORDER BY RANGE DESC)AS R1
FROM FLIGHTS F
JOIN AIRCRAFTS A
ON F.AIRCRAFT_CODE=A.AIRCRAFT_CODE
)

SELECT FLIGHT_NO, range
FROM T1
WHERE R1=1
ORDER BY 1
```