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

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
    T.ticket_no,
    BP.boarding_no,
    BP.SEAT_NO,
    T.passenger_id,
    T.passenger_name
FROM TICKETS T
JOIN BOARDING_PASSES BP
ON T.ticket_no = BP.ticket_no
```

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

Answer:

```
WITH RNK AS (SELECT
    DISTINCT SEAT_NO,
    COUNT(SEAT_NO),
    ROW_NUMBER() OVER(ORDER BY COUNT(SEAT_NO) ASC) AS RN
FROM SEATS
GROUP BY 1)
SELECT
    SEAT_NO
FROM RNK
WHERE RN = 1
```

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 RNK AS (
    SELECT
```

```
        TO_CHAR(B.BOOK_DATE,'mon-YY') AS MONTH_NAME,

        T.PASSENGER_ID,

        T.PASSENGER_NAME,

        B.TOTAL_AMOUNT,

        ROW_NUMBER() OVER(PARTITION BY
TO_CHAR(B.BOOK_DATE,'mon-YY') ORDER BY B.TOTAL_AMOUNT DESC)
AS RK

    FROM BOOKINGS B

    JOIN TICKETS T

    ON T.BOOK_REF = B.BOOK_REF
)
SELECT

    MONTH_NAME,

    PASSENGER_ID,

    PASSENGER_NAME,

    TOTAL_AMOUNT

FROM RNK

WHERE RK = 1

ORDER  BY 4 DESC
```

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 RNK AS (

    SELECT

        TO_CHAR(B.BOOK_DATE,'mon-YY') AS MONTH_NAME,

        T.PASSENGER_ID,

        T.PASSENGER_NAME,

        B.TOTAL_AMOUNT,
```

```
        ROW_NUMBER() OVER(PARTITION BY
TO_CHAR(B.BOOK_DATE,'mon-YY') ORDER BY B.TOTAL_AMOUNT ASC)
AS RK
    FROM BOOKINGS B
    JOIN TICKETS T
    ON T.BOOK_REF = B.BOOK_REF
)
SELECT
    MONTH_NAME,
    PASSENGER_ID,
    PASSENGER_NAME,
    TOTAL_AMOUNT
FROM RNK
WHERE RK = 1
ORDER  BY 4 ASC
```

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
    T.PASSENGER_ID,
    T.PASSENGER_NAME,
    T.TICKET_NO,
    COUNT(TF.FLIGHT_ID) AS FLIGHT_COUNT
FROM TICKETS T
JOIN TICKET_FLIGHTS TF
ON T.TICKET_NO= TF.TICKET_NO
GROUP BY 1,2,3
HAVING COUNT(TF.FLIGHT_ID)>1
```

7. **How many tickets are there without boarding passes?**
Expected Output: just one number is required.

Answer: SELECT

   COUNT(T.TICKET_NO)

FROM TICKETS T

LEFT JOIN BOARDING_PASSES BP

ON T.TICKET_NO = BP.TICKET_NO

WHERE BP.TICKET_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: SELECT

   FLIGHT_NO,

   DEPARTURE_AIRPORT,

   ARRIVAL_AIRPORT,

   AIRCRAFT_CODE,

   TO_CHAR(SCHEDULED_ARRIVAL-

SCHEDULED_DEPARTURE,'HH24:MI') AS DURATIONS

FROM FLIGHTS

ORDER BY 5 DESC

LIMIT 1

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,
    CONCAT(scheduled_departure, ' - ', scheduled_arrival) AS timings
FROM flights
WHERE EXTRACT(HOUR FROM scheduled_departure) BETWEEN 6 AND
10 OR EXTRACT(HOUR FROM scheduled_departure) = 11 AND
EXTRACT(MINUTE FROM scheduled_departure) = 0
```

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  EARLYMORNFLIGHT AS ( SELECT
    FLIGHT_ID,
    FLIGHT_NO,
    SCHEDULED_DEPARTURE,
    SCHEDULED_ARRIVAL,
    DEPARTURE_AIRPORT,
    ROW_NUMBER() OVER (PARTITION BY DEPARTURE_AIRPORT ORDER
BY  SCHEDULED_DEPARTURE) AS RN
    FROM FLIGHTS
    WHERE SCHEDULED_DEPARTURE ::TIME BETWEEN '00:00:00' AND '11:59:59')
SELECT
    EM.FLIGHT_ID,
    EM.FLIGHT_NO,
    EM.SCHEDULED_DEPARTURE,
    EM.SCHEDULED_ARRIVAL,
    EM.DEPARTURE_AIRPORT,
```

EM.SCHEDULED_DEPARTURE AS TIMINGS

from EARLYMORNFLIGHT AS EM

INNER JOIN AIRPORTS AS A

ON EM.DEPARTURE_AIRPORT = A.AIRPORT_CODE

WHERE EM.RN =1

ORDER BY 5 ASC

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

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

Answer: SELECT

   COUNT(AIRCRAFT_CODE)

FROM SEATS

WHERE FARE_CONDITIONS = 'Business'

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

Answer:

WITH CTE AS (SELECT

   DEPARTURE_AIRPORT,

   COUNT(DEPARTURE_AIRPORT) AS CT

FROM FLIGHTS

GROUP BY 1

ORDER BY 2 DESC)

SELECT

   DEPARTURE_AIRPORT

FROM CTE

LIMIT 1

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

Answer:

WITH CTE AS (SELECT

   DEPARTURE_AIRPORT,

   COUNT(DEPARTURE_AIRPORT) AS CT

FROM FLIGHTS

GROUP BY 1

ORDER BY 2 ASC)

SELECT

   DEPARTURE_AIRPORT

FROM CTE

LIMIT 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'
```

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

Answer:

```
SELECT
    F.FLIGHT_NO,
    F.AIRCRAFT_CODE,
    A.RANGE
FROM FLIGHTS F
JOIN AIRCRAFTS A
ON F.AIRCRAFT_CODE = A.AIRCRAFT_CODE
WHERE A.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_ID) AS FLIGHT_COUNT
FROM FLIGHTS
WHERE DEPARTURE_AIRPORT IN ('URS','KUF') AND ARRIVAL_AIRPORT IN('URS','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_ID) AS FLIGHT_COUNT
```

FROM FLIGHTS

WHERE DEPARTURE_AIRPORT IN ('NOZ','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) AS 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

   DISTINCT F.FLIGHT_NO,

   A.AIRCRAFT_CODE,

   A.RANGE,

   F.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
    F.FLIGHT_ID,
    A.MODEL AS AIRCRAFT_MODEL
FROM FLIGHTS F
JOIN AIRCRAFTS A
ON F.AIRCRAFT_CODE = A.AIRCRAFT_CODE
WHERE F.STATUS IN ('Cancelled','Delayed') AND A.MODEL LIKE '%Airbus%'
```

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 F.STATUS IN ('Cancelled','Delayed') AND A.MODEL LIKE '%Boeing%'
```

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

Answer:
```
WITH CTE AS (SELECT
    A.AIRPORT_NAME,
    COUNT(F.STATUS),
    RANK() OVER(ORDER BY COUNT(F.STATUS) DESC) AS RNK
FROM FLIGHTS F
JOIN AIRPORTS A
ON A.AIRPORT_CODE = F.ARRIVAL_AIRPORT
WHERE F.STATUS = 'Cancelled'
GROUP BY 1)
```

```
SELECT
    AIRPORT_NAME
FROM CTE
WHERE RNK = 1
```

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

Answer:
```
SELECT
    F.FLIGHT_ID,
    A.MODEL
FROM FLIGHTS F
JOIN AIRCRAFTS A
ON A.AIRCRAFT_CODE = F.AIRCRAFT_CODE
WHERE A.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 CTE AS(SELECT
    FLIGHT_ID,
    FLIGHT_NO,
    SCHEDULED_DEPARTURE,
    DEPARTURE_AIRPORT,
    RANK() OVER(PARTITION BY DEPARTURE_AIRPORT ORDER BY SCHEDULED_DEPARTURE
DESC) AS RNK
FROM FLIGHTS)
SELECT
    FLIGHT_ID,
    FLIGHT_NO,
    SCHEDULED_DEPARTURE,
    DEPARTURE_AIRPORT
```

FROM CTE

WHERE RNK = 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

   T.PASSENGER_NAME,

   TF.AMOUNT AS 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'

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

Answer:

WITH CTE AS(SELECT

   FLIGHT_ID,

   FLIGHT_NO,

   SCHEDULED_DEPARTURE,

   DEPARTURE_AIRPORT,

   RANK() OVER(PARTITION BY DEPARTURE_AIRPORT ORDER BY SCHEDULED_DEPARTURE

ASC) AS RNK

FROM FLIGHTS)

SELECT

   FLIGHT_ID,

   FLIGHT_NO,

   SCHEDULED_DEPARTURE,

```
    DEPARTURE_AIRPORT
FROM CTE
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 A.AIRCRAFT_CODE = F.AIRCRAFT_CODE
WHERE F.STATUS = 'Canacelled' AND A.MODEL LIKE '%Airbus%'
```

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

Answer:
```
SELECT
    F.FLIGHT_NO,
    MAX(A.RANGE) AS RANGE
FROM FLIGHTS F
JOIN AIRCRAFTS A
ON A.AIRCRAFT_CODE = F.AIRCRAFT_CODE
GROUP BY 1
```