**Capstone Project - Airline db**

100 marks

**Resources**

**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](https://www.skillovilla.com/assignment/%E2%80%A2%09https:/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**
* Uploading output data (CSVs) of the SQL queries is **NOT**

**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

1. **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

    b.ticket\_no,

    b.boarding\_no,

    b.seat\_no,

    t.passenger\_id,

    t.passenger\_name

 FROM

    boarding\_passes b

    inner join tickets t

    on t.ticket\_no=b.ticket\_no

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

**Answer:**

 SELECT

   seat\_no,

   count(seat\_no)

 FROM boarding\_passes

 GROUP BY seat\_no

 ORDER BY 2 ASC

 LIMIT 1

1. ***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:**

SELECT \*

FROM (

    SELECT

        TO\_CHAR(book\_date, 'Mon-YY') as Month,

        t.passenger\_id,

        t.passenger\_name,

        b.total\_amount,

        DENSE\_RANK() OVER (PARTITION BY TO\_CHAR(book\_date, 'Mon-YY') ORDER BY b.total\_amount DESC) AS RNK

    FROM bookings b

    JOIN tickets t ON b.book\_ref = t.book\_ref

) AS sub

WHERE RNK = 1

1. ***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:**

SELECT \*

FROM (

    SELECT

        TO\_CHAR(book\_date, 'Mon-YY') as Month,

        t.passenger\_id,

        t.passenger\_name,

        b.total\_amount,

        DENSE\_RANK() OVER (PARTITION BY TO\_CHAR(book\_date, 'Mon-YY') ORDER BY b.total\_amount ASC) AS RNK

    FROM bookings b

    JOIN tickets t ON b.book\_ref = t.book\_ref

) AS sub

WHERE RNK = 1

1. **Identify the travel details of the flights having return journey (more than 1 flight).**

Expected Output: Passenger\_id, passenger\_name, ticket\_number and flight count.

**Answer:**

SELECT

    t.passenger\_id,

    t.passenger\_name,

    f.ticket\_no,

    COUNT(f.flight\_id) flight\_count

FROM ticket\_flights f

JOIN tickets t

ON f.ticket\_no = t.ticket\_no

GROUP BY 1, 2, 3

HAVING COUNT(f.flight\_id) > 1

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

Expected Output: just one number is required.

**Answer:**

SELECT COUNT(DISTINCT ticket\_no) AS tickets\_without\_boarding\_passes

FROM tickets

WHERE ticket\_no NOT IN (SELECT DISTINCT ticket\_no FROM boarding\_passes)

1. **Identify details of the longest flight (using flights table)?**

Expected Output: Flight number, departure airport, arrival airport, aircraft code and durations.

**Answer:**

WITH CTE AS (

    SELECT

        flight\_no as flight\_number,

        departure\_airport,

        arrival\_airport,

        aircraft\_code,

        MAX(scheduled\_arrival - scheduled\_departure) as duration,

        DENSE\_RANK() OVER (ORDER BY MAX(scheduled\_arrival - scheduled\_departure) DESC) AS RNK

    FROM flights

    GROUP BY 1, 2, 3, 4

)

SELECT

    flight\_number,

    departure\_airport,

    arrival\_airport,

    aircraft\_code,

    duration

FROM CTE

WHERE RNK = 1

1. **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 as flight\_number,

    scheduled\_departure,

    scheduled\_arrival,

    EXTRACT(hour FROM scheduled\_departure) as timing

FROM flights

WHERE EXTRACT(hour FROM scheduled\_departure) BETWEEN 6 AND 11

1. **Identify the earliest morning flight available from every airport.Early morning: 2:00 am to 6:00 am.**

Expected output: flight\_id, flight\_number, scheduled\_departure, scheduled\_arrival, departure airport and timings.

**Answer:**

SELECT

    \*

FROM(

    SELECT

         flight\_id,

         flight\_no as flight\_number,

         scheduled\_departure,

         arrival\_airport,

         aircraft\_code,

        scheduled\_arrival,

         max (scheduled\_arrival-scheduled\_departure) as duration,

         DENSE\_RANK() over(partition by departure\_airport order by  max (scheduled\_arrival-scheduled\_departure) DESC) as RNK

FROM flights

GROUP BY 1,2,3,4

) as Table\_A

WHERE RNK=1

1. **Questions:** **Find list of airport codes in Europe/Moscow timezone**

Expected Output:  Airport\_code.

**Answer:**

SELECT

    airport\_code

FROM airports

WHERE timezone = 'Europe/Moscow'

1. **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) seat\_count

FROM seats

GROUP BY 1, 2

1. **How many aircrafts codes have at least one Business class seats?**

Expected Output : Count of aircraft codes

**Answer:**

SELECT

    COUNT(DISTINCT aircraft\_code) AS Count\_of\_aircraft\_codes

FROM seats

WHERE fare\_conditions = 'Business'

1. **Find out the name of the airport having maximum number of departure flight**

Expected Output : Airport\_name

**Answer:**

SELECT

    a.airport\_name

FROM flights f

JOIN airports a

ON a.airport\_code = f.departure\_airport

GROUP BY 1

ORDER BY COUNT(f.departure\_airport) DESC

LIMIT 1

1. **Find out the name of the airport having least number of scheduled departure flights**

Expected Output : Airport\_name

**Answer:**

SELECT

    a.airport\_name

FROM flights f

JOIN airports a

ON a.airport\_code = f.departure\_airport

GROUP BY 1

ORDER BY COUNT(f.departure\_airport) ASC

LIMIT 1

1. **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

1. **Identify flight ids having range between 3000 to 6000**

Expected Output : Flight\_Number , aircraft\_code, ranges

**Answer:**

SELECT

    f.flight\_no,

    a.aircraft\_code,

    a.range

FROM flights f

JOIN aircrafts a

ON f.aircraft\_code = a.aircraft\_code

WHERE a.range BETWEEN 3000 AND 6000

1. **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 = 'URS' AND arrival\_airport = 'KUF'

1. **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) FLight\_count

FROM flights

WHERE departure\_airport IN('NOZ','KRR')

1. **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) count\_of\_flights

FROM flights

WHERE departure\_airport IN('KZN','DME','NBC','NJC','GDX','SGC','VKO','ROV')

GROUP BY 1

1. **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

    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 a.range BETWEEN 3000 AND 6000 AND departure\_airport = 'DME'

1. **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 aircraft\_model

FROM flights f

JOIN aircrafts a

ON f.aircraft\_code=a.aircraft\_code

WHERE model like '%Airbus%' AND status IN ('Delayed','Cancelled')

1. **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 aircraft\_model

FROM flights f

JOIN aircrafts a

ON f.aircraft\_code=a.aircraft\_code

WHERE model like '%Boeing%' AND status IN ('Delayed','Cancelled')

1. **Which airport(name) has most cancelled flights (arriving)?**

Expected Output : Airport\_name.

**Answer:**

SELECT

    a.airport\_name

FROM flights f

JOIN airports a

ON f.departure\_airport=a.airport\_code

GROUP BY 1

ORDER BY COUNT(f.status = 'Cancelled') DESC

LIMIT 1

1. ***Identify flight ids which are using “Airbus aircrafts”***

*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 '%Airbus%'

1. ***Identify date-wise last flight id flying from every airport?***

*Expected Output: Flight\_id,flight\_number,schedule\_departure,departure\_airport*

**Answer:**

SELECT

    \*

FROM (

    SELECT

        flight\_id,

        flight\_no,

        scheduled\_departure,

        departure\_airport,

        DENSE\_RANK() OVER (PARTITION BY departure\_airport ORDER BY scheduled\_departure DESC) AS RNK

    FROM flights

) AS Table\_1

WHERE RNK = 1

1. ***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,

    SUM(tf.amount) AS total\_refund

FROM flights f

JOIN ticket\_flights tf

ON tf.flight\_id = f.flight\_id

JOIN tickets t

ON t.ticket\_no = tf.ticket\_no

WHERE f.status = 'Cancelled'

GROUP BY 1

1. ***Identify date wise first cancelled flight id flying for every airport?***

*Expected Output : Flight\_id,flight\_number,schedule\_departure,departure\_airport*

**Answer:**

SELECT

    \*

FROM (

    SELECT

        flight\_id,

        flight\_no,

        scheduled\_departure,

        departure\_airport,

        DENSE\_RANK() OVER (PARTITION BY departure\_airport ORDER BY scheduled\_departure ASC) AS RNK

    FROM flights

    WHERE status = 'Cancelled'

) AS Table\_1

WHERE RNK = 1

1. ***Identify list of Airbus flight ids which got cancelled.***

*Expected Output : Flight\_id*

**Answer:**

SELECT

    f.flight\_id

FROM flights f

JOIN aircrafts a

ON a.aircraft\_code = f.aircraft\_code

WHERE a.model LIKE '%Airbus%' AND f.status = 'Cancelled'

1. ***Identify list of flight ids having highest range.***

*Expected Output : Flight\_id, range*

SELECT

    f.flight\_no,

    MAX(a.range) AS range

FROM flights f

JOIN aircrafts a

ON f.aircraft\_code = a.aircraft\_code

GROUP BY 1