**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](•%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** 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-mmm-dd') AS formatted\_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 t.ticket\_no ,**

**boarding\_no , seat\_no as seat\_number , passenger\_id , passenger\_name**

**FROM tickets as t**

**join boarding\_passes as bp on t.ticket\_no =bp.ticket\_no**

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

**Answer:**

**with t1 as (SELECT seat\_no, COUNT(seat\_no) AS seat\_count**

**FROM boarding\_passes**

**GROUP BY seat\_no**

**ORDER BY seat\_count ASC)**

**SELECT seat\_no**

**from t1**

**where seat\_count<=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:**

**WITH cte1 AS (**

**SELECT**

**TO\_CHAR(Book\_date, 'mmm-yy') AS month\_name, passenger\_id,passenger\_name,**

**SUM(total\_amount) AS total\_amount**

**FROM TICKETS AS T**

**JOIN Bookings AS B ON B.book\_ref = T.book\_ref**

**GROUP BY 1, passenger\_id, passenger\_name**

**),**

**t2 AS (**

**SELECT**

**\*,**

**RANK() OVER (PARTITION BY month\_name ORDER BY total\_amount DESC) AS rnk**

**FROM cte1**

**)**

**SELECT**

**month\_name,**

**passenger\_id,**

**passenger\_name,**

**total\_amount**

**FROM t2**

**WHERE rnk = 1**

**ORDER BY 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:**

**WITH cte1 AS (**

**SELECT**

**TO\_CHAR(Book\_date, 'MON-YYYY') AS month\_name, passenger\_id,passenger\_name,**

**SUM(total\_amount) AS total\_amount**

**FROM TICKETS AS T**

**JOIN Bookings AS B ON B.book\_ref = T.book\_ref**

**GROUP BY 1, passenger\_id, passenger\_name**

**),**

**t2 AS (**

**SELECT**

**\*,**

**RANK() OVER (PARTITION BY month\_name ORDER BY total\_amount ) AS rnk**

**FROM cte1**

**)**

**SELECT**

**month\_name,**

**passenger\_id,**

**passenger\_name,**

**total\_amount**

**FROM t2**

**WHERE rnk = 1**

**ORDER BY 1**

1. **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:**

**with t1 as(select ticket\_no,count(flight\_id) as flight\_count**

**from ticket\_flights**

**group by ticket\_no)**

**select passenger\_id,passenger\_name,t1.ticket\_no as ticket\_number,flight\_count**

**from t1**

**join tickets t**

**on t1.ticket\_no = t.ticket\_no**

**where flight\_count > 1**

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

Expected Output: just one number is required.

**Answer:**

**SELECT count(t.ticket\_no) FROM ticket\_flights t left join BOARDING\_PASSES b on t.ticket\_no = b.ticket\_no**

**where b.boarding\_no is NULL**

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

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

**Answer:**

**with b1 as (select distinct flight\_no , departure\_airport , arrival\_airport , aircraft\_code ,(scheduled\_arrival - scheduled\_departure) as duration,**

**dense\_rank() over(order by (scheduled\_arrival - scheduled\_departure) desc) as rnk\_**

**from flights**

**order by rnk\_ asc)**

**select flight\_no , departure\_airport , arrival\_airport , aircraft\_code ,duration**

**from b1**

**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:**

**with B as (select \*,TO\_CHAR(scheduled\_departure, 'HH24:MI:SS')as departure\_time , TO\_CHAR(scheduled\_arrival, 'HH24:MI:SS')as arrival\_time,CASE**

**WHEN EXTRACT(HOUR FROM scheduled\_departure) >= 6 AND**

**EXTRACT(HOUR FROM scheduled\_departure) <= 11 THEN 'Morning'**

**ELSE 'Not Morning' END AS timings**

**from flights)**

**select flight\_id,flight\_no,scheduled\_departure,scheduled\_arrival,timings**

**from B**

**where departure\_time between '06:00:00' and '11:00:00'**

1. **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 cte1 AS (**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**departure\_airport,**

**CAST(scheduled\_departure AS TIME) AS timings**

**FROM flights**

**),**

**Rank\_table AS (**

**SELECT**

**\*,**

**RANK() OVER (PARTITION BY departure\_airport ORDER BY timings ASC) AS rnk1**

**FROM cte1**

**)**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**departure\_airport,**

**CASE**

**WHEN timings BETWEEN '02:00:00' AND '06:00:00' THEN 'Early Morning'**

**ELSE 'Not Early Morning'**

**END AS timings\_**

**FROM Rank\_table**

**WHERE rnk1 = 1 AND timings BETWEEN '02:00:00' AND '06:00:00';**

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) as seat\_count

from seats

group by 1, 2

order by 2 asc

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

Expected Output : Count of aircraft codes

**Answer:**

**select count(aircraft\_code) 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:**

with b1 as (select airport\_name, count(scheduled\_departure)

from flights f

left join airports a

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

group by 1

order by 2 desc

limit 1)

select airport\_name

from b1

limit 1

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

Expected Output : Airport\_name

**Answer:**

with b1 as (select airport\_name, count(scheduled\_departure)

from flights f

left join airports a

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

group by 1

order by 2 asc)

select airport\_name

from b1

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 distinct flight\_no , f.aircraft\_code , range as ranges

from flights f

left join aircrafts a

on f.aircraft\_code = a.aircraft\_code

where 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 in ('URS','KUF') and arrival\_airport in ('URS', '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(\*) as 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) as count\_of\_flights**

**from flights**

**where departure\_airport in ('KZN','DME','NBC','NJC','GDX','SGC','VKO','ROV')**

**group by 1**

**order by 2 desc**

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 distinct flight\_no , f.aircraft\_code , range ,departure\_airport**

**from flights f**

**left join aircrafts a**

**on f.aircraft\_code = a.aircraft\_code**

**where departure\_airport in ('DME') and range between 3000 and 6000**

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 flight\_id ,**

**model as aircraft\_model**

**from flights f**

**left join aircrafts a**

**on f.aircraft\_code = a.aircraft\_code**

**where model like '%Airbus%' and status in ('Cancelled' , 'delayed')**

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 flight\_id , model as aircraft\_model**

**from flights f**

**left join aircrafts a**

**on f.aircraft\_code = a.aircraft\_code**

**where model like '%Boeing%' and status in ('Cancelled' , 'delayed')**

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

Expected Output : Airport\_name

**Answer:WITH AIRPORT AS (**

**SELECT airport\_name , count(\*) as cancelled\_flights**

**FROM flights as f left join airports as a**

**on a.airport\_code = f.ARRIVAL\_airport**

**WHERE status = 'Cancelled'**

**GROUP BY 1 )**

**SELECT Airport\_name**

**FROM AIRPORT**

**WHERE cancelled\_flights = (SELECT MAX(cancelled\_flights) From AIRPORT)**

1. ***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

left join aircrafts a

on f.aircraft\_code = a.aircraft\_code

where 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:**

**WITH ranked\_flights AS (**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport,**

**dense\_rank() OVER (PARTITION BY departure\_airport ORDER BY scheduled\_departure DESC) AS RANK\_**

**FROM**

**flights**

**)**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**FROM**

**ranked\_flights**

**WHERE**

**RANK\_ = 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:**

**WITH CanceledFlights AS (**

**SELECT**

**t.passenger\_name,**

**tf.amount AS ticket\_amount**

**FROM**

**TICKETS t**

**JOIN TICKET\_FLIGHTS tf ON t.ticket\_no = tf.ticket\_no**

**JOIN FLIGHTS f ON tf.flight\_id = f.flight\_id**

**WHERE**

**f.status = 'Cancelled'**

**)**

**SELECT**

**cf.passenger\_name,**

**SUM(cf.ticket\_amount) AS total\_refund**

**FROM**

**CanceledFlights cf**

**GROUP BY**

**cf.passenger\_name;**

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

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

**Answer:**

**WITH ranked\_cancelled\_flights AS (**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport,**

**Dense\_rank() OVER (PARTITION BY departure\_airport ORDER BY scheduled\_departure) AS Rank\_**

**FROM**

**flights**

**WHERE**

**status = 'Cancelled'**

**)**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**FROM**

**ranked\_cancelled\_flights**

**WHERE**

**Rank\_ = 1;**

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

*Expected Output : Flight\_id*

**Answer:**

select Flight\_id

from flights f

left join aircrafts a

on f.aircraft\_code = a.aircraft\_code

where model like '%Airbus%' AND status in ('Cancelled')

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

*Expected Output : Flight\_no, range*

**Answer:**

**with flightrangecte as (SELECT**

**f.flight\_id,**

**a.range,**

**dense\_rank() OVER (ORDER BY a.range DESC) AS row\_num**

**FROM**

**flights f**

**left JOIN aircrafts a**

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

**SELECT**

**flight\_id,range**

**FROM**

**FlightRangeCTE**

**WHERE**

**row\_num = 1**

**order by 1 asc;**