**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](file:///C:\Users\Pravallika\Downloads\•%09https:\www.skillovilla.com\playground\sql%3fexerciseId=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**

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,**

**B.boarding\_no,**

**B.seat\_no,**

**T.passenger\_id,**

**T.passenger\_name**

**FROM Tickets T**

**JOIN Boarding\_passes B 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**

**FROM Boarding\_Passes**

**GROUP BY seat\_no**

**ORDER BY COUNT (\*) 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:**

**WITH TotalAmount AS (**

**SELECT**

**TO\_CHAR(b.book\_date, 'Mon-YY') AS month\_name,**

**T.passenger\_id,**

**T.passenger\_name,**

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

**FROM BOOKINGS B**

**JOIN TICKETS T ON B.book\_ref = T.book\_ref**

**GROUP BY TO\_CHAR(b.book\_date, 'Mon-YY'),T.passenger\_id,T.passenger\_name**

**),**

**RankedTotals AS (**

**SELECT**

**month\_name,**

**passenger\_id,**

**passenger\_name,**

**total\_amount,**

**Row\_number() OVER (PARTITION BY month\_name ORDER BY total\_amount DESC) AS rank**

**FROM Totalamount**

**)**

**SELECT**

**month\_name,**

**passenger\_id,**

**passenger\_name,**

**total\_amount**

**FROM RankedTotals**

**WHERE rank =1**

**ORDER BY month\_name**

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 MonthlyTotals AS (**

**SELECT**

**TO\_CHAR(b.book\_date, 'Mon-YY') AS month\_name,**

**T.passenger\_id,**

**T.passenger\_name,**

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

**FROM BOOKINGS B**

**JOIN TICKETS T ON B.book\_ref = T.book\_ref**

**GROUP BY TO\_CHAR(b.book\_date, 'Mon-YY'),T.passenger\_id,T.passenger\_name**

**),**

**RankedMonthlyTotals AS (**

**SELECT**

**month\_name,**

**passenger\_id,**

**passenger\_name,**

**total\_amount,**

**Row\_number() OVER (PARTITION BY month\_name ORDER BY total\_amount ASC) AS rank**

**FROM MonthlyTotals**

**)**

**SELECT**

**month\_name,**

**passenger\_id,**

**passenger\_name,**

**total\_amount**

**FROM RankedMonthlyTotals**

**WHERE rank =1**

**ORDER BY month\_name**

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:**

**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 T.Passenger\_id,T.passenger\_name,T.ticket\_no**

**HAVING COUNT(TF.flight\_id) > 1**

1. **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 B ON T.ticket\_no = b.ticket\_no**

**WHERE B.ticket\_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:**

**SELECT**

**Flight\_no,**

**departure\_airport,**

**arrival\_airport,**

**aircraft\_code,**

**(actual\_arrival - actual\_departure) as durations**

**FROM flights**

**WHERE actual\_arrival IS NOT NULL AND actual\_departure IS NOT NULL**

**ORDER BY durations DESC**

**LIMIT 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,**

**scheduled\_departure,**

**scheduled\_arrival,**

**scheduled\_departure:: time as timing**

**FROM FLIGHTS**

**WHERE scheduled\_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 FLIGHTIME AS (**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**departure\_airport,**

**scheduled\_departure:: time as timing**

**FROM Flights),**

**Timeranking AS (**

**SELECT \*,**

**row\_number()over(partition by departure\_airport order by timing) AS rnk**

**FROM FLIGHTIME)**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**departure\_airport,**

**timing**

**FROM Timeranking**

**WHERE rnk =1 AND timing BETWEEN '02:00:00' AND '06:00:00'**

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

Expected Output: Airport\_code.

**Answer:**

**Note: Distinct is used to remove duplicates**

**SELECT**

**DISTINCT 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)**

**FROM SEATS**

**GROUP BY Aircraft\_code,fare\_conditions**

**ORDER BY Aircraft\_code,fare\_conditions**

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

Expected Output : Count of aircraft codes

**Answer:**

**SELECT**

**COUNT(DISTINCT Aircraft\_code)**

**FROM SEATS**

**WHERE fare\_conditions = 'Business' AND SEAT\_NO IS NOT NULL**

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

Expected Output : Airport\_name

**Answer:**

**SELECT**

**airport\_name**

**FROM airports A**

**JOIN flights F on A.airport\_code = F.departure\_airport**

**GROUP BY airport\_name**

**ORDER BY COUNT(F.scheduled\_departure) desc**

**LIMIT 1**

**Note:In the question it is mentioned as departure, where it was not clearly mentioned as scheduled\_departure or actual\_departure, Hence I have considered as scheduled\_departure considering the below question. However even if we consider the actual\_departure the output is same.**

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

Expected Output : Airport\_name

**Answer:**

**SELECT**

**airport\_name**

**FROM airports A**

**JOIN flights F on A.airport\_code = F.departure\_airport**

**GROUP BY airport\_name**

**ORDER BY COUNT(F.scheduled\_departure) asc**

**LIMIT 1**

1. **How many flights from ‘DME’ airport don’t have actual departure?**

Expected Output : Flight Count

**Answer:**

**SELECT**

**COUNT(\*) 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,**

**F.aircraft\_code,**

**A.range**

**FROM FLIGHTS F**

**JOIN AIRCRAFTS A on A.aircraft\_code = F.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) 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 departure\_airport**

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 F.Flight\_no,F.aircraft\_code,A.range,F.departure\_airport**

**FROM FLIGHTS F**

**JOIN AIRCRAFTS A ON A.aircraft\_code = F.aircraft\_code**

**WHERE A.range BETWEEN 3000 AND 6000**

**AND F.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 as aircraft\_model**

**FROM Flights F**

**JOIN Aircrafts A ON A.aircraft\_code = F.aircraft\_code**

**WHERE A.model LIKE '%Airbus%'**

**AND (F.status = 'Cancelled' OR F.status = '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**

**F.Flight\_id,**

**A.model AS aircraft\_model**

**FROM Flights F**

**JOIN Aircrafts A ON A.aircraft\_code = F.aircraft\_code**

**WHERE A.model LIKE '%Boeing%'**

**AND (F.status = 'Cancelled' OR F.status = 'Delayed')**

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 A.airport\_code = F.arrival\_airport**

**WHERE f.status = 'Cancelled'**

**GROUP BY A.airport\_name**

**ORDER BY COUNT(F.flight\_id) 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**

**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:**

**WITH RankedFlights AS (**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport,**

**row\_number() OVER (PARTITION BY departure\_airport ORDER BY scheduled\_departure DESC) AS rn**

**FROM flights**

**)**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**FROM RankedFlights**

**WHERE rn = 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 TICKETS T**

**JOIN TICKET\_FLIGHTS TF on TF.ticket\_no = T.ticket\_no**

**JOIN FLIGHTS F on F.flight\_id = TF.flight\_id**

**WHERE status = 'Cancelled'**

**GROUP BY T.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 CancelledFlights AS (**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**FROM flights**

**WHERE status = 'Cancelled'**

**),**

**RankedFlights AS (**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport,**

**ROW\_NUMBER() OVER (PARTITION BY departure\_airport ORDER BY scheduled\_departure) AS rn**

**FROM CancelledFlights**

**)**

**SELECT**

**Flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**FROM RankedFlights**

**WHERE rn = 1**

1. ***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\_codec**

**WHERE model like '%Airbus%'**

**AND status = 'cancelled'**

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

*Expected Output : Flight\_no, range*

**Answer:**

**Note: We can use Distinct if we any duplicate flight\_id, However in this we see there are no duplicate values and we are getting same output if we use Distinct or not**

**SELECT**

**F.flight\_id,**

**A.range**

**FROM**

**FLIGHTS F**

**JOIN**

**AIRCRAFTS A ON F.aircraft\_code = A.aircraft\_code**

**WHERE (A.range) = (SELECT MAX(A.range) FROM AIRCRAFTS A)**