**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-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:**

**WITH SeatCounts AS (**

**SELECT**

**seat\_no,**

**COUNT(\*) AS total\_count,**

**RANK() OVER (ORDER BY COUNT(\*) ASC) AS rank**

**FROM**

**boarding\_passes**

**GROUP BY**

**seat\_no**

**)**

**SELECT**

**seat\_no**

**FROM**

**SeatCounts**

**WHERE**

**rank = 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 MonthlyPayments AS (**

**select \*,**

**RANK() OVER (PARTITION BY month, year ORDER BY total\_amount DESC) AS rnk FROM**

**(**

**SELECT**

**EXTRACT(MONTH FROM b.book\_date) AS month,**

**EXTRACT(YEAR FROM b.book\_date) AS year,**

**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**

**month, year, t.passenger\_id, t.passenger\_name ) t1**

**)**

**SELECT**

**TO\_CHAR(TO\_DATE(CONCAT(month, '-', year), 'MM-YYYY'), 'Mon-YY') AS month\_name,**

**passenger\_id,**

**passenger\_name,**

**total\_amount**

**FROM**

**MonthlyPayments**

**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:**

**WITH MonthlyPayments AS (**

**select \*,**

**RANK() OVER (PARTITION BY month, year ORDER BY total\_amount asc) AS rnk FROM**

**(**

**SELECT**

**EXTRACT(MONTH FROM b.book\_date) AS month,**

**EXTRACT(YEAR FROM b.book\_date) AS year,**

**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**

**month, year, t.passenger\_id, t.passenger\_name ) t1**

**)**

**SELECT**

**TO\_CHAR(TO\_DATE(CONCAT(month, '-', year), 'MM-YYYY'), 'Mon-YY') AS month\_name,**

**passenger\_id,**

**passenger\_name,**

**total\_amount**

**FROM**

**MonthlyPayments**

**WHERE**

**rnk = 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:**

**SELECT**

**t.passenger\_id,**

**t.passenger\_name,**

**t.ticket\_no,**

**COUNT(\*) AS flight\_count**

**FROM**

**Tickets t**

**JOIN**

**boarding\_passes bp ON t.ticket\_no = bp.ticket\_no**

**GROUP BY**

**t.passenger\_id, t.passenger\_name, t.ticket\_no**

**HAVING**

**COUNT(\*) > 1**

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

Expected Output: just one number is required.

**Answer:**

**SELECT**

**COUNT(\*) AS ticket\_count\_without\_boarding\_passes**

**FROM Tickets t**

**LEFT JOIN boarding\_passes bp ON t.ticket\_no = bp.ticket\_no**

**WHERE bp.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 duration**

**FROM**

**flights**

**ORDER BY**

**duration desc nulls last**

**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,**

**concat(TO\_CHAR(scheduled\_departure, 'HH:MI AM'),'-',TO\_CHAR(scheduled\_arrival, 'HH:MI AM'))as timings**

**FROM**

**flights**

**WHERE**

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

**AND EXTRACT(HOUR FROM scheduled\_departure) < 11**

**and EXTRACT(HOUR FROM scheduled\_arrival) >= 6**

**AND EXTRACT(HOUR FROM scheduled\_arrival) < 11**

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 RankedFlights AS (**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**departure\_airport,**

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

**FROM**

**flights**

**WHERE**

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

**AND EXTRACT(HOUR FROM scheduled\_departure) < 11**

**)**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**departure\_airport,**

**TO\_CHAR(scheduled\_departure, 'HH:MI AM') AS timings**

**FROM**

**RankedFlights**

**WHERE**

**flight\_rank = 1**

**order by TO\_CHAR(scheduled\_departure, 'HH:MI AM')**

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(\*) AS seat\_count**

**FROM**

**seats**

**GROUP 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) AS aircraftcodes\_count**

**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**

**airport\_name**

**FROM**

**airports**

**WHERE**

**airport\_code IN (**

**SELECT**

**departure\_airport**

**FROM**

**flights**

**GROUP BY**

**departure\_airport**

**ORDER BY**

**COUNT(\*) DESC**

**LIMIT 1)**

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**

**WHERE**

**airport\_code IN (**

**SELECT**

**departure\_airport**

**FROM**

**flights**

**GROUP BY**

**departure\_airport**

**ORDER BY**

**COUNT(\*) 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**

**flight\_no,**

**aircraft\_code,**

**flight\_id AS ranges**

**FROM**

**flights**

**WHERE**

**flight\_id 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(\*) 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(\*) 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(\*) AS flight\_count**

**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**

**flight\_no,**

**aircraft\_code,**

**flight\_id AS range,**

**departure\_airport**

**FROM**

**flights**

**WHERE**

**flight\_id 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 as aircraft\_model**

**FROM**

**flights f**

**JOIN**

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

**WHERE**

**a.model like '%Airbus%'**

**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**

**airports a**

**JOIN**

**flights f ON a.airport\_code = f.arrival\_airport**

**WHERE**

**f.status = 'Cancelled'**

**GROUP BY**

**a.airport\_name**

**ORDER BY**

**COUNT(\*) 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**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**FROM**

**flights**

**WHERE (departure\_airport,scheduled\_departure) in**

**(**

**select departure\_airport , last\_flight from**

**(**

**select**

**to\_char(scheduled\_departure,'yyyy-mm-dd') as Date , departure\_airport , max(scheduled\_departure) as last\_flight**

**from flights**

**group by to\_char(scheduled\_departure,'yyyy-mm-dd') , departure\_airport**

**order by departure\_airport , to\_char(scheduled\_departure,'yyyy-mm-dd') desc**

**)t1**

**)**

**order by scheduled\_departure desc**

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(b.total\_amount) AS total\_refund**

**FROM**

**bookings b**

**JOIN**

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

**Join**

**Boarding\_Passes bp**

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

**JOIN**

**flights f ON f.flight\_id = bp.flight\_id**

**WHERE**

**f.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:**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**FROM**

**flights**

**WHERE (departure\_airport,scheduled\_departure) in**

**(**

**select departure\_airport , first\_cancelled\_flight from**

**(**

**select**

**to\_char(scheduled\_departure,'yyyy-mm-dd') as Date , departure\_airport , min(scheduled\_departure) as first\_cancelled\_flight**

**from flights**

**where status = 'Cancelled'**

**group by to\_char(scheduled\_departure,'yyyy-mm-dd') , departure\_airport**

**order by departure\_airport , to\_char(scheduled\_departure,'yyyy-mm-dd') desc**

**)t1**

**)**

**order by scheduled\_departure desc ;**

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

**WHERE**

**a.model = 'Airbus'**

**AND f.status = 'Cancelled'**

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

*Expected Output : Flight\_no, range*

**Answer:**

**SELECT**

**flight\_no,**

**flight\_id AS Range**

**FROM**

**flights**

**WHERE**

**flight\_id = (**

**SELECT**

**MAX(flight\_id)**

**FROM**

**flights**

**)**