**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 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**

**bp.ticket\_no,**

**bp.boarding\_no,**

**bp.seat\_no,**

**t.passenger\_id,**

**t.passenger\_name**

**from**

**BOARDING\_PASSES as bp  join tickets as t**

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

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

**Answer:** **select**

**seat\_no,**

**count(\*) as total\_count**

**from**

**seats**

**group by**

**seat\_no**

**order by**

**total\_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 ranking as (**

**select**

**to\_char(book\_date, 'mmm-yy') as month\_name,**

**passenger\_id,**

**passenger\_name,**

**sum(total\_amount) as amount,**

**dense\_rank() over(partition by to\_char(book\_date, 'mmm-yy') order by sum(total\_amount) desc) as rankk**

**from tickets as t join bookings as b on t.book\_ref = b.book\_ref**

**group by to\_char(book\_date, 'mmm-yy'), passenger\_id, passenger\_name**

**)**

**select**

**month\_name,**

**passenger\_id,**

**passenger\_name,**

**amount**

**from**

**ranking**

**where**

**rankk = 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 ranking as (**

**select**

**to\_char(book\_date, 'mmm-yy') as month\_name,**

**passenger\_id,**

**passenger\_name,**

**sum(total\_amount) as amount,**

**dense\_rank() over(partition by to\_char(book\_date, 'mmm-yy') order by sum(total\_amount) asc) as rankk**

**from tickets as t join bookings as b on t.book\_ref = b.book\_ref**

**group by to\_char(book\_date, 'mmm-yy'), passenger\_id, passenger\_name**

**)**

**select**

**month\_name,**

**passenger\_id,**

**passenger\_name,**

**amount**

**from**

**ranking**

**where**

**rankk = 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(f.flight\_id) as flight\_count**

**from tickets as t left join ticket\_flights as f**

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

**group by 1,2,3**

**having count(f.flight\_id)>1**

**order by 4**

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

Expected Output: just one number is required.

**Answer:** **SELECT COUNT(\*)**

**FROM tickets t**

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

**WHERE b.ticket\_no IS NULL;**

**Output: 251**

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,**

**max(actual\_arrival-actual\_departure) as duration**

**from flights**

**group  by 1,2,3,4**

**order by 5 desc**

**limit 2**

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,**

**EXTRACT(HOUR FROM scheduled\_departure) as departuretime ,**

**EXTRACT(HOUR FROM scheduled\_arrival) arrivaltime**

**FROM**

**flights**

**WHERE**

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

**AND EXTRACT(HOUR FROM scheduled\_departure) < 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 ranking AS (**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**departure\_airport,**

**EXTRACT(HOUR FROM scheduled\_departure)  AS timings,**

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

**FROM**

**flights**

**WHERE**

**EXTRACT(HOUR FROM scheduled\_departure) >= 2**

**AND EXTRACT(HOUR FROM scheduled\_departure) < 12**

**)**

**SELECT**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**departure\_airport,**

**timings**

**FROM**

**ranking**

**WHERE**

**rank = 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) as seat\_count**

**from seats**

**group by 1,2**

**order by 3**

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

Expected Output : Count of aircraft codes

**Answer:** **SELECT**

**COUNT(a.aircraft\_code) as count\_of\_ac**

**from aircrafts as a left join seats as s**

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

**where s.fare\_conditions = 'Business'**

**having count(s.fare\_conditions) >=1**

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

Expected Output : Airport\_name

**Answer:** **SELECT**

**a.airport\_name**

**FROM**

**airports AS a**

**JOIN**

**flights AS f**

**ON**

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

**GROUP BY**

**a.airport\_name**

**ORDER BY**

**COUNT(f.flight\_id) 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**

**airports AS a**

**JOIN**

**flights AS f**

**ON**

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

**GROUP BY**

**a.airport\_name**

**ORDER BY**

**COUNT(f.flight\_id) ASC**

**LIMIT 1;**

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

Expected Output : Flight Count

**Answer:** **select**

**count(flight\_id)**

**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 f.Flight\_no,**

**c.aircraft\_code,**

**range**

**from flights as f join aircrafts as c**

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

**where c.range between 3000 and 6000**

**order by Flight\_no,range**

1. **Write a query to get the count of flights flying between URS and KUF?**

Expected Output : Flight\_count

**Answer:** **select**

**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(\*)**

**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(\*)**

**from flights**

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

**group by 1**

**order by 2**

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,**

**c.aircraft\_code,**

**range,**

**departure\_airport**

**from flights as f join aircrafts as c**

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

**where c.range between 3000 and 6000**

**and (departure\_airport= 'DME')**

**order by Flight\_no,range**

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, ac.model**

**from flights f**

**join  aircrafts ac on f.aircraft\_code = ac.aircraft\_code**

**where**

**ac.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, ac.model**

**from flights f join aircrafts ac**

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

**where**

**ac.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 f.arrival\_airport = a.airport\_code**

**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 Flight\_id, model**

**from flights as f join aircrafts ac on f.aircraft\_code = ac.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 table1 as (**

**select**

**f.flight\_id,**

**f.flight\_no,**

**f.scheduled\_departure,**

**f.departure\_airport,**

**dense\_rank() over (**

**partition by f.departure\_airport, to\_char(scheduled\_departure,'DD-MM-YYYY')**

**order by f.scheduled\_departure desc**

**) as rank**

**from**

**flights f**

**)**

**select**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**from**

**table1**

**where**

**rank = 1**

**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**

**passenger\_name,**

**sum(amount) as total\_refund**

**from tickets as t  join ticket\_flights as tf**

**on t.ticket\_no  = tf.ticket\_no**

**join flights as f**

**on f.flight\_id = tf.flight\_id**

**where f.status = 'Cancelled'**

**group by passenger\_name**

**order by 2**

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

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

**Answer:** **with table1 as (**

**select**

**f.flight\_id,**

**f.flight\_no,**

**f.scheduled\_departure,**

**f.departure\_airport,**

**dense\_rank() over (**

**partition by f.departure\_airport, to\_char(scheduled\_departure,'DD-MM-YYYY')**

**order by f.scheduled\_departure**

**) as rank**

**from**

**flights f**

**where**

**f.status = 'Cancelled'**

**)**

**select**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**from**

**table1**

**where**

**rank = 1**

**order by**

**scheduled\_departure;**

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

*Expected Output : Flight\_id*

**Answer:** **select f.flight\_id**

**from flights f join aircrafts ac on f.aircraft\_code = ac.aircraft\_code**

**where**

**f.status = 'Cancelled'**

**and ac.model like '%Airbus%'**

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

*Expected Output : Flight\_no, range*

**Answer:** **select**

**f.flight\_no,**

**max(ac.range) as maxrange**

**from  aircrafts ac join flights as f**

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

**group by flight\_no**

**order by maxrange desc**