**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**

**T.ticket\_no,**

**BP.boarding\_no,**

**BP.seat\_no as seat\_number,**

**T.passenger\_id,**

**T.passenger\_name**

**from**

**tickets as T**

**join boarding\_passes as BP**

**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: with t1 as (**

**Select bp.seat\_no, dense\_rank() over(order**

**by count(bp.seat\_no) asc) as rnk**

**From boarding\_passes as bp**

**Group by 1**

**)**

**select seat\_no from t1**

**where rnk = 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 t1 AS (**

**SELECT**

**To\_char(B.book\_date, 'mmm-yy') AS month\_name,**

**T.passenger\_id,**

**T.passenger\_name,**

**B.total\_amount AS total\_amount,**

**row\_number() OVER (PARTITION BY To\_char(B.book\_date, 'mmm-yy') ORDER BY B.total\_amount DESC) AS rnk**

**FROM**

**Bookings AS B**

**JOIN**

**Tickets AS T ON B.book\_ref = T.book\_ref**

**GROUP BY 1,2,3,4**

**)**

**SELECT month\_name, passenger\_id, passenger\_name, total\_amount**

**FROM t1**

**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 t1 AS (**

**SELECT**

**To\_char(B.book\_date, 'mmm-yy') AS month\_name,**

**T.passenger\_id,**

**T.passenger\_name,**

**B.total\_amount AS total\_amount,**

**row\_number() OVER (PARTITION BY To\_char(B.book\_date, 'mmm-yy') ORDER BY B.total\_amount asc) AS rnk**

**FROM**

**Bookings AS B**

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

**GROUP BY 1,2,3,4**

**)**

**SELECT \***

**FROM t1**

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

**From tickets t**

**Join ticket\_flights tf on tf.ticket\_no = t.ticket\_no**

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

**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(\*)**

**From tickets as T**

**Left join boarding\_passes as BP**

**On BP.ticket\_no = T.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: with t1 as (**

**Select**

**Flight\_no,**

**Departure\_airport,**

**Arrival\_airport,**

**Aircraft\_code,**

**(scheduled\_arrival-scheduled\_departure) as durations,**

**rank() over(order by (scheduled\_arrival-scheduled\_departure) desc) as rnk**

**From Flights**

**Group by 1,2,3,4,5**

**)**

**select \***

**from t1**

**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, scheduled\_departure, scheduled\_arrival,**

**extract(hour from scheduled\_departure) ||':' || LPAD(extract(minute from scheduled\_departure)::TEXT,2, '0') as timings**

**from flights**

**where extract(hour from scheduled\_departure) >= 6 and extract(hour from scheduled\_departure) < 11**

**Order by 3**

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 flight\_id, flight\_no, scheduled\_departure, scheduled\_arrival, departure\_airport,**

**Timings**

**From (select flight\_id, flight\_no, scheduled\_departure, scheduled\_arrival, departure\_airport, row\_number() over (Partition by departure\_airport order by scheduled\_departure) as rn, Extract(Hour from scheduled\_departure) || ':' || LPAD (Extract(minute from scheduled\_departure)::TEXT, 2, '0') as timings**

**From flights**

**Where extract(hour from scheduled\_departure) >= 2 and extract(hour from scheduled\_departure) < 6) as subquery**

**Where rn = 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(\*) as 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 distinct count(A.aircraft\_code)**

**From aircrafts as A**

**Join seats as S**

**On A.aircraft\_code = S.aircraft\_code**

**Where S.fare\_conditions = 'Business'**

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

Expected Output : Airport\_name

**Answer: WITH t1 AS (**

**SELECT**

**cast(A.airport\_name AS JSON)->>'en' as airport\_name,**

**ROW\_NUMBER() OVER (ORDER BY COUNT(F.departure\_airport) DESC) AS rnk**

**FROM**

**flights AS F**

**JOIN airports AS A**

**ON F.departure\_airport = A.airport\_code**

**GROUP BY A.airport\_name**

**)**

**SELECT airport\_name**

**FROM t1**

**WHERE rnk = 1**

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

Expected Output : Airport\_name

**Answer: WITH t1 AS (**

**SELECT**

**A.airport\_name,**

**ROW\_NUMBER() OVER (ORDER BY COUNT(F.departure\_airport) ASC) AS rnk**

**FROM**

**flights AS F**

**JOIN**

**airports AS A ON F.departure\_airport = A.airport\_code**

**GROUP BY A.airport\_name**

**)**

**SELECT airport\_name**

**FROM t1**

**WHERE rnk = 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 as ranges**

**From Flights as F**

**Join Aircrafts as 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 distinct count(flight\_id) as flight\_count**

**from flights**

**where departure\_airport = 'URS' and arrival\_airport = 'KUF' or departure\_airport = 'KUF' and arrival\_airport = 'URS'**

1. **Write a query to get the count of flights flying from either from NOZ or KRR?**

Expected Output : Flight count

**Answer: Select distinct 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(departure\_airport)**

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

**from flights as F**

**left join aircrafts as A**

**on A.aircraft\_code = F.aircraft\_code**

**where F.departure\_airport = 'DME' and A.range between 3000 and 6000**

**group by 1,2,3,4**

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 as F**

**join aircrafts as 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**

**from flights as F**

**join aircrafts as 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: with t1 as (**

**Select A.airport\_name, count(A.airport\_name), row\_number() over(order by count(airport\_name) desc) as rnk**

**From Airports as A**

**Join Flights as F**

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

**Where f.status = 'Cancelled'**

**Group by 1**

**Order by 2 desc**

**)**

**select airport\_name**

**from t1**

**where rnk = 1**

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

*Expected Output : Flight\_id,aircraft\_model*

**Answer: select distinct F.flight\_id, A.model**

**from flights as F**

**join aircrafts as 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 T1 AS (**

**SELECT**

**FLIGHT\_ID,**

**FLIGHT\_NO,**

**SCHEDULED\_DEPARTURE,**

**TO\_CHAR(SCHEDULED\_DEPARTURE, 'YYYY-MM-DD') AS DATE,**

**DEPARTURE\_AIRPORT,**

**ROW\_NUMBER() OVER(PARTITION BY DEPARTURE\_AIRPORT ORDER BY SCHEDULED\_DEPARTURE DESC) AS RNK**

**FROM**

**FLIGHTS**

**ORDER BY 4 ASC**

**)**

**SELECT FLIGHT\_ID,**

**FLIGHT\_NO,**

**SCHEDULED\_DEPARTURE,**

**TO\_CHAR(SCHEDULED\_DEPARTURE, 'YYYY-MM-DD') AS DATE,**

**DEPARTURE\_AIRPORT**

**FROM T1**

**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, TF.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 TF.flight\_id = F.flight\_id**

**where F.status = 'Cancelled'**

**group by 1,2**

**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: Select flight\_id, flight\_no, scheduled\_departure, departure\_airport**

**From ( select flight\_id, flight\_no, scheduled\_departure, departure\_airport,**

**Row\_number() over(partition by departure\_airport, date(scheduled\_departure)**

**order by scheduled\_departure) as rn**

**From flights**

**Where status = 'Cancelled') as subquery**

**Where rn = 1**

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

*Expected Output : Flight\_id*

**Answer: select flight\_id**

**from flights as F**

**join aircrafts as 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\_no, range*

**Answer: Select**

**f.flight\_no, a.range**

**from**

**flights as f**

**join**

**aircrafts as a**

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

**order by 2 desc**