SQL Capstone Project

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

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;
```

2. **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
bs.ticket_no,
bs.boarding_no,
bs.seat_no,
t.passenger_id,
t.passenger_name
from boarding_passes bs
join tickets t
on t.ticket_no=bs.ticket_no;
```

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

Answer:

```
select
min(seat_no) as least_allocated_seat
from seats;
```

4. ***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,
t.passenger_id,
t.passenger_name,
b.total_amount
from bookings b
join tickets t
```

```
   on t.book_ref=b.book_ref
),
t2 as (
   select *,
   rank() over(partition by month order by total_amount desc) as rank
   from t1
)
select
month,
passenger_id,
passenger_name,
total_amount
from t2
where rank=1
order by month;
```

5. ***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,
   t.passenger_id,
   t.passenger_name,
   b.total_amount
   from bookings b
   join tickets t
   on t.book_ref=b.book_ref
```

```
),
t2 as (
   select *,
   rank() over(partition by month order by total_amount asc) as rank
   from t1
)
select
month,
passenger_id,
passenger_name,
total_amount
from t2
where rank=1
order by month;
```

6. **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
```

7. **How many tickets are there without boarding passes?**
Expected Output: just one number is required.

Answer:

```
select
count(ticket_no)
from boarding_passes
where boarding_no is null
```

8. **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 flight_duration
from flights
group by 1,2,3,4
```

9. **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,
to_char(scheduled_departure,'HH24:MI') as timings
from flights
```

```
where extract(hour from scheduled_departure) between 6 and 10
order by scheduled_departure;
```

10. 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 t1 as(
   select
flight_id,
flight_no,
scheduled_departure,
scheduled_arrival,
departure_airport,
to_char(scheduled_departure,'HH24:MI') as timings ,
row_number() over(partition by departure_airport order by scheduled_departure) as rnk
from flights
where extract(hour from scheduled_departure) between 6 and 10
order by scheduled_departure)

select
flight_id,
flight_no,
scheduled_departure,
scheduled_arrival,
departure_airport,
timings
from t1
where rnk =1
order by departure_airport;
```

11. Questions: Find list of airport codes in Europe/Moscow timezone
Expected Output: Airport_code.

Answer:

```
select
airport_code
from airports
where timezone='Europe/Moscow'
```

12. 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
```

13. How many aircrafts codes have at least one Business class seats?
Expected Output : Count of aircraft codes

Answer:

```
select
count(distinct aircraft_code) as count_of_aircraft_codes
from seats
where fare_conditions='Business';
```

14. Find out the name of the airport having maximum number of departure flight
Expected Output : Airport_name

Answer:

```
select
```

```
a.airport_name
from flights f
join airports a
on f.departure_airport=a.airport_code
group by a.airport_name
order by count(*) desc
limit 1;
```

15. Find out the name of the airport having least number of scheduled departure flights
Expected Output : Airport_name

Answer:

```
select
a.airport_name
from flights f
join airports a
on f.departure_airport=a.airport_code
group by a.airport_name
order by count(*)
limit 1;
```

16. 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;
```

17. 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
from aircrafts a
join flights f
on a.aircraft_code=f.aircraft_code
where range between 3000 and 6000;
```

18. 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';
```

19. 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');
```

20. 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) count_of_flights
```

```
from flights
where departure_airport in ('KZN','DME','NBC','NJC','GDX','SGC','VKO','ROV')
group by 1
order by departure_airport;
```

21. **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,
a.aircraft_code,
a.range,
f.departure_airport
from aircrafts a
join flights f
on f.aircraft_code=a.aircraft_code
where range between 3000 and 6000 and departure_airport='DME'
```

22. **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
from flights f
join aircrafts a
on a.aircraft_code=f.aircraft_code
where model like '%Airbus%' and status in ('Cancelled','Delayed');
```

23. **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 f
join aircrafts a
on a.aircraft_code=f.aircraft_code
where model like '%Boeing%' and status in ('Cancelled','Delayed');
```

24. **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;
```

25. ***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 a.aircraft_code=f.aircraft_code
```

```
where model like '%Airbus%' ;
```

26. *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,
departure_airport,
row_number() over(partition by departure_airport, DATE(scheduled_departure) order by
scheduled_departure desc) as rnk
from flights)

select
flight_id,
flight_no,
scheduled_departure,
departure_airport
from t1
where rnk = 1
order by departure_airport,
scheduled_departure;
```

27. *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 tickets t
join bookings b
on b.book_ref=t.book_ref
join ticket_flights tf
on tf.ticket_no=t.ticket_no
join flights f
on tf.flight_id=f.flight_id
where f.status='Cancelled'
group by 1
order by 2 desc ;
```

28. *Identify date wise first cancelled flight id flying for every airport?*
Expected Output : Flight_id,flight_number,schedule_departure,departure_airport

Answer:

```
with t1 as(
select
flight_id,
flight_no,
scheduled_departure,
departure_airport,
row_number() over(partition by departure_airport,DATE(scheduled_departure) order by
scheduled_departure) as rnk
from flights
where status='Cancelled'
)
select
flight_id,
flight_no,
scheduled_departure,
```

```
departure_airport
from t1
where rnk = 1
order by scheduled_departure,
departure_airport;
```

29. *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 a.aircraft_code=f.aircraft_code
where a.model like '%Airbus%' and f.status='Cancelled';
```

30. *Identify list of flight ids having highest range.*
Expected Output : Flight_no, range

Answer:

```
select
f.flight_no,
max(a.range) as range
from flights f
join aircrafts a
on f.aircraft_code=a.aircraft_code
group by 1
order by 2 desc
```