**SQL Capstone Project**

**Airline DB Database**

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

b.ticket\_no,

b.boarding\_no, b.seat\_no, t.passenger\_id, t.passenger\_name

from BOARDING\_PASSES as b

join tickets as t on b.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

from seats

group by 1

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 month\_totals as

(select

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

passenger\_id, passenger\_name,

sum(amount) as total\_amount,

ROW\_NUMBER() OVER (PARTITION BY TO\_CHAR(book\_date, 'mon-yy') ORDER BY SUM(amount) DESC) AS rn

from tickets as t

join bookings as b

on b.book\_ref = t.book\_ref

join ticket\_flights tf

on tf.ticket\_no = t.ticket\_no

group by 1,2,3)

select

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

from month\_totals

where rn = 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 month\_totals as

(select

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

passenger\_id, passenger\_name,

sum(amount) as total\_amount,

ROW\_NUMBER() OVER (PARTITION BY TO\_CHAR(book\_date, 'mon-yy') ORDER BY SUM(amount) ASC) AS rn

from tickets as t

join bookings as b

on b.book\_ref = t.book\_ref

join ticket\_flights tf

on tf.ticket\_no = t.ticket\_no

group by 1,2,3)

select

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

from month\_totals

where rn = 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

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

count(flight\_id) as flight\_count

from tickets t

join ticket\_flights tf on tf.ticket\_no = t.ticket\_no

group by 1,2,3

having count(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) AS tickets\_without\_boarding\_passes

FROM tickets as t

LEFT JOIN boarding\_passes as 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,

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

from flights

group by 1,2,3,4

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,

CASE WHEN EXTRACT(HOUR FROM scheduled\_departure) >= 6 AND EXTRACT(HOUR FROM scheduled\_departure) < 12 THEN 'Morning'

ELSE 'Afternoon/Evening/Night'

END AS timings

FROM flights

WHERE EXTRACT(HOUR FROM scheduled\_departure) >= 6 AND EXTRACT(HOUR FROM scheduled\_departure) < 12

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:**

SELECT

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

CASE WHEN EXTRACT(HOUR FROM scheduled\_departure) >= 2 AND EXTRACT(HOUR FROM scheduled\_departure) < 6 THEN 'Morning'

ELSE 'Afternoon/Evening/Night'

END AS timings

FROM flights

WHERE EXTRACT(HOUR FROM scheduled\_departure) >= 2 AND EXTRACT(HOUR FROM scheduled\_departure) < 6

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

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

Expected Output : Count of aircraft codes

**Answer:**

select

Count(aircraft\_code)

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 = (

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 = (

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,

a.aircraft\_code,

range

from aircrafts a

join flights f

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

group by 1,2

having 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(\*) 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 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

flight\_no,

f.aircraft\_code,

range,

departure\_airport

from flights f

join aircrafts a

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

where range 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 = '%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 = '%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 airports a

join flights f

on a.airport\_code = f.arrival\_airport

where status = 'Cancelled'

group by 1

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 a.aircraft\_code = f.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 Last\_Flights AS (

SELECT

flight\_id,

flight\_no,

scheduled\_departure,

departure\_airport,

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

FROM flights

)

SELECT

flight\_id,

flight\_no,

scheduled\_departure,

departure\_airport

FROM Last\_Flights

WHERE flight\_rank = 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 AS t

JOIN ticket\_flights AS tf

ON tf.ticket\_no = t.ticket\_no

JOIN flights AS f

ON f.flight\_id = tf.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:**

WITH RankedFlights AS (

SELECT

flight\_id,

flight\_no,

scheduled\_departure,

departure\_airport,

ROW\_NUMBER() OVER (PARTITION BY departure\_airport ORDER BY scheduled\_departure) AS row\_num

FROM flights

WHERE status = 'Cancelled'

)

SELECT

flight\_id,

flight\_no,

scheduled\_departure,

departure\_airport

FROM RankedFlights

WHERE row\_num = 1

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

*Expected Output : Flight\_id*

**Answer:**

with Airbus\_flight as (

SELECT

f.flight\_id,

a.model

FROM

flights as f

JOIN aircrafts as a

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

WHERE

status = 'Cancelled'

)

SELECT

flight\_id

FROM Airbus\_flight

WHERE model LIKE '%Airbus%' --

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

*Expected Output : Flight\_no, range*

**Answer:**

select

f.flight\_id,

max(a.range) as range

from flights f

join aircrafts a

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

group by 1