**SQL Capstone Project**

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,**

**boarding\_no,**

**seat\_no,**

**passenger\_id,**

**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:** **select**

**seat\_no**

**from boarding\_passes**

**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 booking\_details as (select**

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

**passenger\_id,**

**passenger\_name,**

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

**from bookings b**

**join tickets t**

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

**group by 1,2,3),**

**rank\_data as (select \*,**

**row\_number() over(partition by month\_name order by total\_amount desc) as rnk**

**from booking\_details)**

**select**

**month\_name,**

**passenger\_id,**

**passenger\_name,**

**total\_amount**

**from rank\_data**

**where rnk=1**

**order by total\_amount desc**

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 booking\_details as (select**

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

**passenger\_id,**

**passenger\_name,**

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

**from bookings b**

**join tickets t**

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

**group by 1,2,3),**

**rank\_data as (select \*,**

**row\_number() over(partition by month\_name order by total\_amount asc) as rnk**

**from booking\_details)**

**select**

**month\_name,**

**passenger\_id,**

**passenger\_name,**

**total\_amount**

**from rank\_data**

**where rnk=1**

**order by total\_amount asc**

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 f**

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

**from tickets t**

**left join boarding\_passes b**

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

**where boarding\_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 duration\_details as(select**

**flight\_no,**

**departure\_airport,**

**arrival\_airport,**

**aircraft\_code,**

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

**from flights**

**where actual\_arrival is not null and actual\_departure is not null**

**),**

**rank\_data as (select \*,**

**dense\_rank() over(order by duration desc) as rnk**

**from duration\_details)**

**select**

**flight\_no,**

**departure\_airport,**

**arrival\_airport,**

**aircraft\_code,**

**duration**

**from rank\_data**

**where rnk=1**

**order by duration desc**

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:**  **with morning\_flights as(select**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**extract(hour from scheduled\_departure) as timing**

**from flights )**

**select**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**timing**

**from morning\_flights**

**where timing >=6 and timing<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 tb1 as(select**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**departure\_airport,**

**to\_char(scheduled\_departure,'HH24:MI:SS') as timings**

**from flights),**

**rank\_data as (select \*,**

**dense\_rank() over(partition by departure\_airport order by scheduled\_departure asc) as rnk**

**from tb1**

**where timings between '02:00:00' and '06:00:00')**

**select**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**scheduled\_arrival,**

**departure\_airport,**

**timings**

**from rank\_data**

**where rnk=1**

**order by timings asc**

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'**

**order by 1 asc**

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**

**a.aircraft\_code,**

**s.fare\_conditions,**

**count(seat\_no) as seat\_no**

**from aircrafts a**

**join seats s**

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

**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(distinct aircraft\_code) as count\_of\_aircraft\_codes**

**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: with departure\_count as (select**

**airport\_name,**

**count(departure\_airport) as departure\_count**

**from flights f**

**join airports a**

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

**group by 1**

**order by 2 desc**

**limit 1)**

**select**

**airport\_name**

**from departure\_count**

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

Expected Output : Airport\_name

**Answer:**  **with name as (select**

**airport\_name,**

**count(departure\_airport) as departure\_count**

**from flights f**

**join airports a**

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

**group by 1**

**order by 2 asc**

**limit 1 )**

**select**

**airport\_name**

**from name**

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

Expected Output : Flight Count

**Answer:** **select**

**count(flight\_id)**

**from flights**

**where actual\_departure is null**

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

1. **Identify flight ids having range between 3000 to 6000**

Expected Output : Flight\_Number , aircraft\_code, ranges

**Answer:** **select**

**f.flight\_no,**

**f.aircraft\_code,**

**range**

**from flights f**

**join aircrafts a**

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

**count(flight\_id) 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(flight\_id) 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(flight\_id) 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**

**f.flight\_no,**

**f.aircraft\_code,**

**a.range,**

**f.departure\_airport**

**from aircrafts a**

**join flights f**

**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**

**flight\_id,**

**model as aircraft\_model**

**from flights f**

**join aircrafts a**

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

**where model like '%Airbus%'**

**and (status ='Cancelled' or 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**

**flight\_id,**

**model as aircraft\_model**

**from flights f**

**join aircrafts a**

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

**where model like '%Boeing%'**

**and (status='Cancelled' or status='Delayed')**

1. **Which airport(name) has most cancelled flights (arriving)?**

Expected Output : Airport\_name

**Answer**  **with tb1 as (select**

**airport\_name,**

**arrival\_airport,**

**status,**

**count(flight\_id) as flight\_count**

**from flights f**

**join airports a**

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

**where status='Cancelled'**

**group by 1,2,3),**

**rank\_data as (select \*,**

**dense\_rank() over(order by flight\_count desc) as rnk**

**from tb1)**

**select**

**airport\_name**

**from rank\_data**

**where rnk=1**

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

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

**Answer:** **select**

**flight\_id,**

**model as aircraft\_model**

**from flights f**

**join aircrafts a**

**on f.aircraft\_code=a.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 tb1 as (select**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport,**

**date(scheduled\_departure) as departure\_date**

**from flights),**

**rank\_data as (select \*,**

**row\_number() over(partition by departure\_airport, departure\_date order by scheduled\_departure desc) as rnk**

**from tb1)**

**select**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**from rank\_data**

**where rnk=1**

**order by departure\_airport,scheduled\_departure**

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 t**

**join ticket\_flights tf**

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

**join flights f**

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

**where status ='Cancelled'**

**group by 1**

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

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

**Answer:** **with cancelled\_flight as (select**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport,**

**status,**

**date(scheduled\_departure) as departure\_date**

**from flights),**

**rank\_data as(select \*,**

**row\_number() over(partition by departure\_airport,departure\_date order by scheduled\_departure asc) as rnk**

**from cancelled\_flight**

**where status='Cancelled')**

**select**

**flight\_id,**

**flight\_no,**

**scheduled\_departure,**

**departure\_airport**

**from rank\_data**

**where rnk=1**

**order by departure\_airport,scheduled\_departure**

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

*Expected Output : Flight\_id*

**Answer:** **select**

**flight\_id**

**from flights f**

**join aircrafts a**

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

**where model like '%Airbus%'**

**and status='Cancelled'**

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

*Expected Output : Flight\_no, range*

**Answer:** **with tb1 as (select**

**flight\_id,**

**range**

**from flights f**

**join aircrafts a**

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

**rank\_data as(select \*,**

**dense\_rank() over(order by range desc) as rnk**

**from tb1)**

**select**

**flight\_id,**

**range**

**from rank\_data**

**where rnk=1**

**order by range desc**