



AIR TRAVEL



AIR TRAVEL – SQL PROJECT

THIS PROJECT, AIR TRAVEL DATA ANALYSIS, IS CREATED BY MOHIT KUMAR
IT USES SQL TO ANALYZE AIRLINE DATA SUCH AS AIRPORTS, FLIGHTS,
PASSENGERS, AND DISTANCES.

THE MAIN GOAL IS TO FIND USEFUL INSIGHTS LIKE BUSIEST ROUTES,
PASSENGER TRENDS, AND SEAT UTILIZATION.

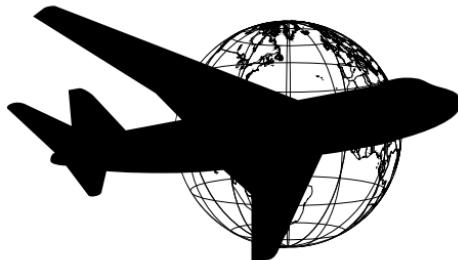


1.FIND THE TOTAL NUMBER OF PASSENGERS BETWEEN EACH ORIGIN AND DESTINATION AIRPORT IN DESCENDING ORDER.

```
create database airport_db;
use airport_db;
show tables;
select * from airports2 limit 5;
select
Origin_airport, Destination_airport,
sum(Passenger) as Total_passe
from airports2
group by
Origin_airport,
Destination_airport
order by Total_passe desc;
```



Result Grid			
	Origin_airport	Destination_airport	Total_passe
▶	PDX	RDM	429036
	SFO	RDM	47513
	SEA	RDM	40803
	LMT	RDM	13269
	SJC	RDM	5993



2.WHICH ORIGIN-DESTINATION AIRPORT PAIR HAS THE HIGHEST NUMBER OF PASSENGERS?

```
select
Origin_airport,Destination_airport,
avg(cast(Passengers as float)/nullif(Seats,0))*100 as avg_seats_utl
from airports2
group by
Origin_airport,
Destination_airport
order by avg_seats_utl desc;
```



	Origin_airport	Destination_airport	avg_seats_utl
▶	EKO	RDM	78.3172946597994
	LWS	RDM	72.2222222222221
	MHK	AMW	70
	CLM	RDM	69.4444444444444
	PDX	RDM	58.06128883196268



Q3. WHICH ORIGIN AIRPORT HAS THE HIGHEST NUMBER OF PASSENGERS AND TOTAL FLIGHTS?

```
select
Origin_airport,
sum(Passenger) as Total_passe,
count(Flights) as total_flights
from airports2
group by
Origin_airport
order by Total_passe, total_flights desc;
```



Origin_airport	Total_passe	total_flights
PDT	2	1
ACV	4	1
SLE	5	1
AST	6	1
PUW	9	1



Q4. WHAT IS THE TOTAL DISTANCE COVERED FROM EACH ORIGIN AIRPORT?

- ```
select
Origin_airport,
sum(Distance)
from airports2
group by
Origin_airport
order by Origin_airport desc;
```



|   | Origin_airport | sum(Distance) |
|---|----------------|---------------|
| ▶ | YKM            | 324           |
|   | SLE            | 102           |
|   | SJC            | 2862          |
|   | SFO            | 18942         |
|   | SEA            | 15504         |



#### Q5. WHAT IS THE MONTHLY AVERAGE DISTANCE, TOTAL FLIGHTS, AND PASSENGERS FOR EACH YEAR?

```
40 • select avg(Distance) as avg_distance,
41 year(Fly_date) as year,
42 month(Fly_date) as month,
43 count(flights) as total_flights,
44 sum(Passengers) as total_passenger
45 from airports2
46 group by year(Fly_date),month(Fly_date)
47 order by year,month;
```



|   | Origin_airport | sum(Distance) |
|---|----------------|---------------|
| ▶ | YKM            | 324           |
|   | SLE            | 102           |
|   | SJC            | 2862          |
|   | SFO            | 18942         |
|   | SEA            | 15504         |



#### Q6. FIND THE MINIMUM NUMBER OF PASSENGERS BETWEEN EACH ORIGIN AND DESTINATION AIRPORT.

```
49 • select origin_airport as o,
50 destination_airport as d,
51 min(passengers) as p
52 from
53 airports2
54 group by o,d
55 order by p asc;
```



Result grid | Filter Rows

|  | o   | d   | p |
|--|-----|-----|---|
|  | ACV | RDM | 4 |
|  | SLE | RDM | 5 |
|  | AST | RDM | 6 |
|  | EAT | RDM | 8 |
|  | YKM | RDM | 8 |



#### Q7. FIND THE TOP 3 ORIGIN AIRPORTS WITH THE HIGHEST NUMBER OF FLIGHTS.

```
58 • select origin_airport as oa,
59 count(flights) as fl
60 from airports2
61 group by oa
62 order by fl desc
63 limit 3;
```

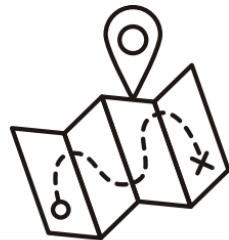


|   | oa  | fl  |
|---|-----|-----|
| ▶ | PDX | 279 |
|   | LMT | 96  |
|   | SEA | 68  |



#### Q8. FIND THE ORIGIN AND DESTINATION AIRPORT PAIR WITH THE MAXIMUM TRAVEL DISTANCE.

- ```
select origin_airport as er,destination_airport as dr  
      ,max(distance) as di  
   from airports2  
group by er,dr  
order by di desc;
```



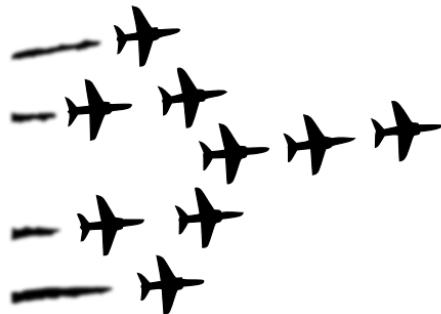
	er	dr	di
▶	LAX	RDM	726
	SJC	RDM	477
	SFO	RDM	462
	EKO	RDM	362
	GEG	RDM	291



Q9. IDENTIFY THE MONTH WITH THE MOST BUSY AND LEAST BUSY FLIGHT ACTIVITY.

with df as |

```
(select  
month(fly_date) as df,  
count(flights) as cf  
from airports2  
group by df )  
select  
df,  
cf,
```



	df	cf	status
▶	1	61	most busy
	7	41	least busy