AIRLINE ANALYSIS USING SQL

Business Problem:

The aviation company faces a challenge in enhancing its overall business performance by optimizing various aspects of its operations.

The main challenges include:

1. Optimizing the fleet composition based on passenger preferences and operational efficiency,
2. Implementing targeted marketing strategies to enhance customer engagement and loyalty,
3. Planning and executing operational expansions, particularly in identifying potential hub airports for improved connectivity.
4. Strategically identifying peak revenue seasons for effective promotional planning.

The airline aims to leverage data-driven insights to make informed decisions in these areas, ultimately improving customer satisfaction, maximizing revenue, and positioning itself for sustained growth in the competitive aviation market.

Objectives for the Project:

1. **Customer Segmentation and Preferences:**
    - Understand customer demographics and travel preferences.
    - Identify patterns in gender and travel class preferences.
2. **Airport Optimization:**
    - Determine the busiest departure airports.
    - Identify potential hub airports for operational planning.
    - Analyze destination airports for arrival patterns.
3. **Route Optimization:**
    - Identify the longest and shortest flight routes for strategic planning.
    - Categorize routes into short, intermediate, and long distances.
4. **Revenue Management:**
    - Optimize pricing strategies based on historical ticket sales and customer preferences.
    - Identify the most profitable travel classes.
    - Analyze yearly revenue trends and monthly variations.
5. **Customer Booking Insights:**
    - Identify customers with the highest number of ticket bookings.
    - Determine passengers traveling by specific classes for targeted services.
    - Implement a function to specify if complimentary services are provided for specific travel classes.

The analysis and insights provided through SQL queries form the basis for informed decision-making in addressing the identified challenges and achieving the set objectives.

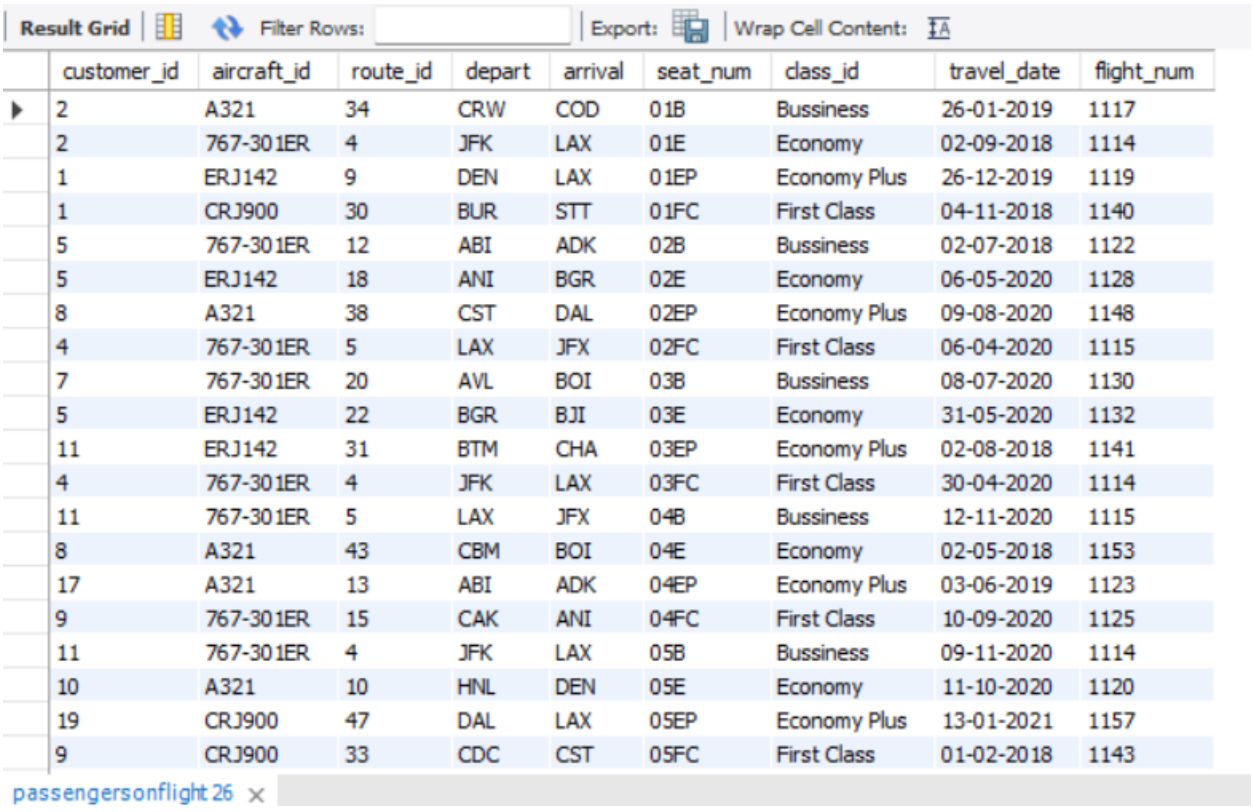
About the Dataset:

There are four tables in the "aviation" database namely: `passengersonflight`, `customer`, `routes`, and `ticketdetails`.

Below is a summary of each table:

1. passengersonflight Table:

- Columns: passenger_id, flight_id, customer_id, route_id, travel_class, seat_number.
- Foreign Key Constraints:
  - `fk` on `customer_id` referencing `customer(customer_id)`.
  - `fk2` on `route_id` referencing `routes(route_id)`.



Result Grid | Filter Rows: | Export: | Wrap Cell Content:

customer_id	aircraft_id	route_id	depart	arrival	seat_num	class_id	travel_date	flight_num
2	A321	34	CRW	COD	01B	Bussiness	26-01-2019	1117
2	767-301ER	4	JFK	LAX	01E	Economy	02-09-2018	1114
1	ERJ142	9	DEN	LAX	01EP	Economy Plus	26-12-2019	1119
1	CRJ900	30	BUR	STT	01FC	First Class	04-11-2018	1140
5	767-301ER	12	ABI	ADK	02B	Bussiness	02-07-2018	1122
5	ERJ142	18	ANI	BGR	02E	Economy	06-05-2020	1128
8	A321	38	CST	DAL	02EP	Economy Plus	09-08-2020	1148
4	767-301ER	5	LAX	JFX	02FC	First Class	06-04-2020	1115
7	767-301ER	20	AVL	BOI	03B	Bussiness	08-07-2020	1130
5	ERJ142	22	BGR	BJI	03E	Economy	31-05-2020	1132
11	ERJ142	31	BTM	CHA	03EP	Economy Plus	02-08-2018	1141
4	767-301ER	4	JFK	LAX	03FC	First Class	30-04-2020	1114
11	767-301ER	5	LAX	JFX	04B	Bussiness	12-11-2020	1115
8	A321	43	CBM	BOI	04E	Economy	02-05-2018	1153
17	A321	13	ABI	ADK	04EP	Economy Plus	03-06-2019	1123
9	767-301ER	15	CAK	ANI	04FC	First Class	10-09-2020	1125
11	767-301ER	4	JFK	LAX	05B	Bussiness	09-11-2020	1114
10	A321	10	HNL	DEN	05E	Economy	11-10-2020	1120
19	CRJ900	47	DAL	LAX	05EP	Economy Plus	13-01-2021	1157
9	CRJ900	33	CDC	CST	05FC	First Class	01-02-2018	1143

passengersonflight 26

2. customer Table:

- Columns: customer_id, first_name, last_name, email, phone_number.
- Primary Key Constraint: `pk` on `customer_id`.

Result Grid | Filter Rows: | Edit:

customer_id	first_name	last_name	date_of_birth	gender
1	Julie	Sam	12-01-1989	F
2	Steve	Ryan	03-04-1983	M
3	Morris	Lois	09-12-1993	M
4	Cathenna	Emily	14-09-1977	F
5	Aaron	Kim	18-02-1991	M
6	Alexander	Scot	12-02-1985	M
7	Anderson	Stewart	11-01-1992	M
8	Floyd	Ted	21-02-1993	M
9	Leo	Travis	22-03-1994	M
10	Melvin	Tracy	23-04-1995	M
11	Roger	Walson	24-05-1996	M
12	Shirley	Wally	25-06-1997	F
13	Solomon	Walter	26-07-1998	M
14	Carol	Vernon	27-08-1999	F
15	Linda	William	28-09-1986	F
16	Chirstine	Willis	06-10-1987	F
17	Catherine	Shad	09-11-1988	F
18	Gloria	Richie	04-12-1989	F
19	Joyce	Paul	02-06-1990	F
20	Sara	Oliver	01-01-1991	F

customer 27

3. routes Table:

- Columns: route_id, departure_airport, arrival_airport, distance, duration.
- Primary Key Constraint: `pk1` on `route_id`.

Result Grid | Filter Rows: | Edit: | Export/Import:

route_id	flight_num	origin_airport	destination_airport	aircraft_id	distance_miles
1	1111	EWR	HNL	767-301ER	4962
2	1112	HNL	EWR	767-301ER	4962
3	1113	EWR	LHR	A321	3466
4	1114	JFK	LAX	767-301ER	2475
5	1115	LAX	JFK	767-301ER	2475
6	1116	HNL	LAX	767-301ER	2556
7	1117	LAX	ORD	A321	1745
8	1118	ORD	EWR	A321	719
9	1119	DEN	LAX	ERJ142	862
10	1120	HNL	DEN	A321	3365
12	1122	ABI	ADK	767-301ER	4300
13	1123	ADK	BQN	A321	2232
14	1124	BQN	CAK	A321	2445
15	1125	CAK	ANI	767-301ER	2000
16	1126	ALB	APN	A321	1700
17	1127	APN	BLV	767-301ER	1900
18	1128	ANI	BGR	ERJ142	2450
19	1129	ATW	AVL	A321	2222
20	1130	AVL	BOI	767-301ER	3134
21	1131	BFL	BET	A321	2425

routes 28

4. ticketdetails Table:

- Columns: ticket_id, customer_id, flight_id, ticket_price, purchase_date.
- Foreign Key Constraint: `fk1` on `customer_id` referencing `customer(customer_id)`.

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

p_date	customer_id	aircraft_id	class_id	no_of_tickets	a_code	Price_per_ticket	brand
26-12-2018	27	767-301ER	Economy	1	DAL	130	Emirates
02-02-2020	22	ERJ142	Economy Plus	1	AGB	220	Jet Airways
03-03-2020	21	CRJ900	Bussiness	1	BOH	490	Bristish Airways
04-04-2020	4	767-301ER	First Class	1	AGB	390	Emirates
05-05-2020	5	ERJ142	Economy	1	CTM	120	Jet Airways
07-07-2020	7	767-301ER	Bussiness	1	BFS	430	Emirates
08-08-2020	8	A321	Economy Plus	1	DAL	275	Qatar Airways
09-09-2020	9	767-301ER	First Class	1	BOH	380	Emirates
10-10-2020	10	A321	Economy	1	MCO	135	Qatar Airways
11-11-2020	11	767-301ER	Bussiness	1	AGB	465	Emirates
12-12-2020	19	CRJ900	Economy Plus	1	DEN	225	Bristish Airways
01-01-2019	13	A321	First Class	1	YVR	395	Qatar Airways
02-02-2019	14	ERJ142	Economy	1	CTM	120	Jet Airways
03-03-2019	25	767-301ER	Bussiness	1	BHX	499	Emirates
04-04-2019	16	CRJ900	First Class	1	YVR	395	Bristish Airways
03-05-2019	17	A321	Economy Plus	1	BFS	250	Qatar Airways
06-06-2019	18	767-301ER	Economy	1	YVR	190	Emirates
07-07-2019	24	A321	Bussiness	1	CTM	480	Qatar Airways
09-08-2019	20	CRJ900	First Class	1	MCO	365	Bristish Airways
21-09-2019	25	767-301ER	Economy	1	BOH	150	Emirates

ticketdetails 29

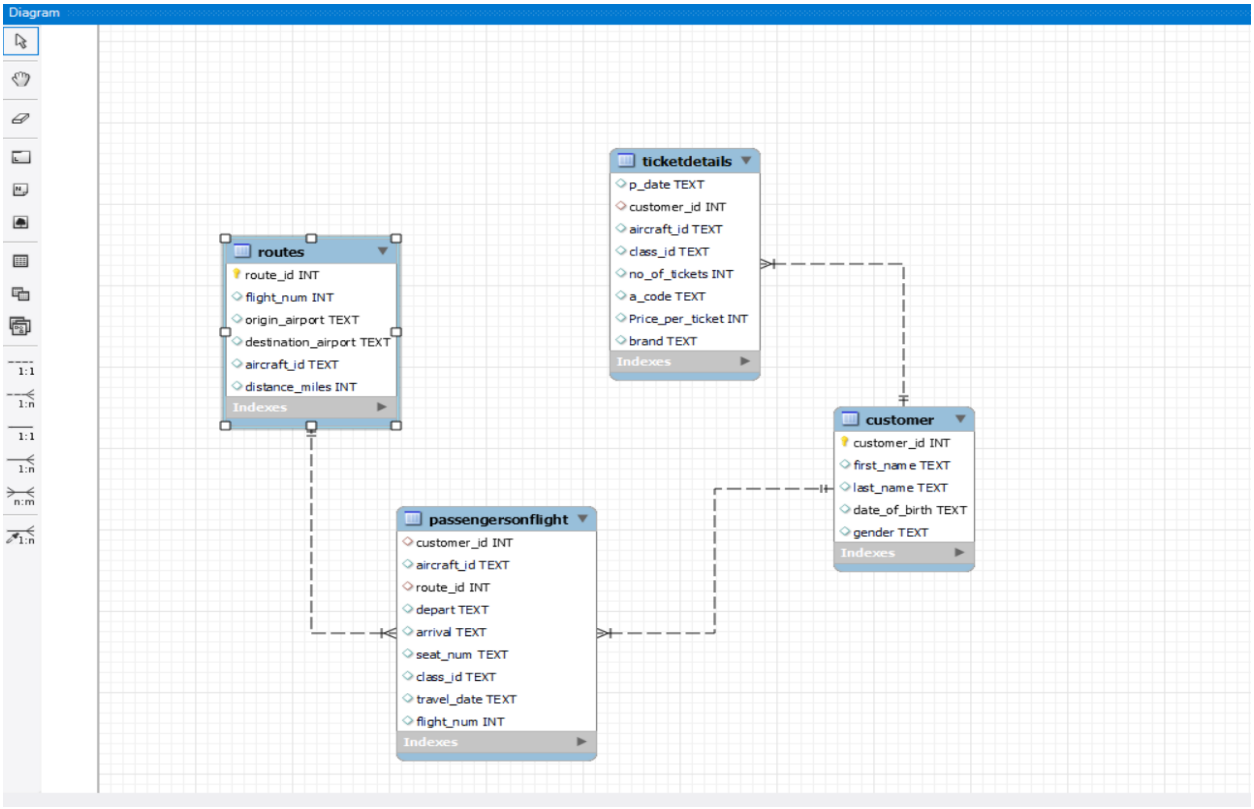
These tables are interconnected through foreign key relationships, allowing for the analysis of passenger data, customer details, flight routes, and ticket information for the aviation company's business problem and objectives.

Queries along with screenshot of the output and inferences are as follows:

```
create database aviation;

use aviation;

alter table customer add constraint pk primary key(customer_id);

alter table routes add constraint pk1 primary key(route_id);

alter table passengersonflight add constraint fk foreign key(customer_id) references
customer(customer_id);

alter table ticketdetails add constraint fk1 foreign key(customer_id) references
customer(customer_id);
```

alter table passengersonflight add constraint fk2 foreign key(route_id) references routes(route_id);

The ER Diagram:



 # Lets first identify the types of aircraft we have and how many passengers are travelling.

```
SELECT aircraft_id, sum(no_of_tickets) as 'total ticket'
FROM aviation.ticketdetails
group by aircraft_id
order by 'total ticket' DESC;
```

Result Grid | Filter Rows:

aircraft_id	total ticket
767-301ER	18
ERJ142	9
CRJ900	10
A321	13

From the above we can see that there are 4 types of aircraft:

1. Airbus A321- a commercial aircraft
2. Boeing 767-301ER- a commercial aircraft
3. Embraer ERJ142 - a regional jet (capacity 42)
4. Bombardier CRJ900 -a regional jet (capactity-80-90)

- **And Maximum passengers are travelling with Boeing 767-301ER i.e out of 50 passengers, 18 passengers travelled through Boeing.**

❖ **Identifying the Customer Segmentation and Preferences:**
  - Understanding the demographics of customers and their travel preferences.
  - Identifying patterns in gender and travel class preferences.

```
SELECT c.gender, t.class_id, COUNT(*) as customer_count,
        AVG(t.Price_per_ticket) as     avg_ticket_price
FROM customer c
JOIN ticketdetails t
ON c.customer_id = t.customer_id
GROUP BY c.gender,t.class_id
order by gender DESC;
```

Result Grid | Filter Rows: | Export:

gender	class_id	customer_count	avg_ticket_price
M	Bussiness	12	467.0000
M	Economy	9	142.2222
M	Economy Plus	4	252.5000
M	First Class	5	388.0000
F	Bussiness	1	430.0000
F	Economy	5	142.0000
F	Economy Plus	6	241.6667
F	First Class	8	368.1250

Inferences Drawn:

- Out of 50 passengers, 30 are male(60%) and 20 female(40%).
- The count is highest for Male customers in the Business class, suggesting that more males booked tickets in the Business class compared to other classes.

❖ Airport Optimization:

- Determing the busiest Airport

1. Departure Airport

```
SELECT origin_airport, COUNT(*) as departure_count
FROM routes
GROUP BY origin_airport
ORDER BY departure_count DESC;
```

Result Grid | Filter Rows:

origin_airport	departure_count
HNL	3
EWR	2
LAX	2
JFK	1
ORD	1
DEN	1
ABI	1
ADK	1
BQN	1
CAK	1
ALB	1
APN	1
ANI	1

Result 35

Inferences Drawn:

1. **Busiest Origin Airports:**

- HNL (Honolulu International Airport): It appears to be the busiest origin airport with 3 departures, suggesting a high demand for flights departing from Honolulu.
- EWR (Newark Liberty International Airport): It has 2 departures, indicating a relatively busy airport for flight departures.

- LAX (Los Angeles International Airport): Similarly, LAX has 2 departures, showing significant departure activity.

2.**Limited Departures from Some Airports:**

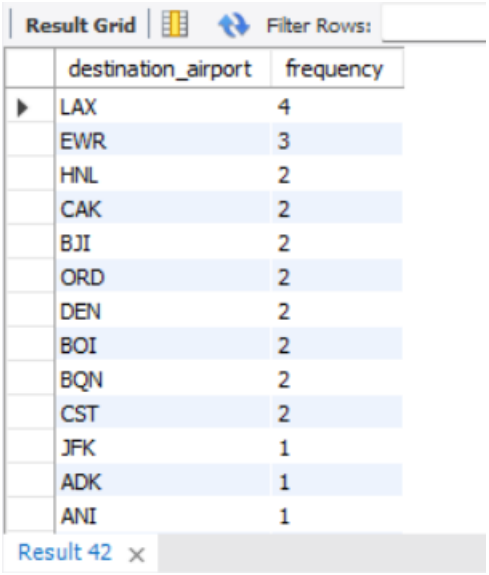
- Several airports, including JFK, ORD, DEN, ABI, ADK, BQN, CAK, ALB, APN, ANI, ATW, AVL, BFL, BGR, BLV, BJI, RDM, BET, and others, have only 1 departure each.
- This may indicate that these airports have relatively lower departure frequencies in the given dataset.

3.Potential Hub Airports:

- HNL, with the highest departure count, might be a significant hub or a popular starting point for flights in the context of this dataset.

2. Destination Airport

```
select destination_airport, count(*) as frequency
from routes
group by destination_airport
order by frequency DESC;
```



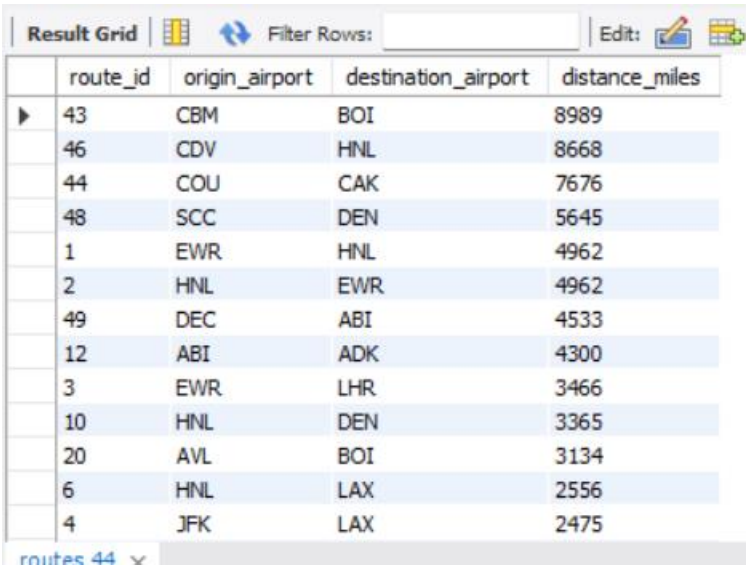
Inferences Drawn:

1.Most Common Destination Airports:

- LAX (Los Angeles International Airport): It is the most frequently occurring destination airport with 4 flights in the dataset, indicating a high demand for flights arriving at Los Angeles.
- EWR (Newark Liberty International Airport): EWR follows closely with 3 flights as the destination, suggesting significant arrival activity.

❖ **# Route Optimization - Longest and Shortest Routes**

```
SELECT route_id, origin_airport, destination_airport,distance_miles
FROM routes
GROUP BY route_id, origin_airport, destination_airport, distance_miles
ORDER BY distance_miles DESC;
```



Result Grid | Filter Rows: | Edit:

route_id	origin_airport	destination_airport	distance_miles
43	CBM	BOI	8989
46	CDV	HNL	8668
44	COU	CAK	7676
48	SCC	DEN	5645
1	EWR	HNL	4962
2	HNL	EWR	4962
49	DEC	ABI	4533
12	ABI	ADK	4300
3	EWR	LHR	3466
10	HNL	DEN	3365
20	AVL	BOI	3134
6	HNL	LAX	2556
4	JFK	LAX	2475

routes 44

Inferences Drawn:

1. **Longest Routes:**
    - The routes with the highest distances are at the top, such as route_id 43 from CBM to BOI with 8989 miles and route_id 46 from CDV to HNL with 8668 miles.
    - These long-distance routes might involve intercontinental flights.
2. **Shorter Routes:**
    - Some routes have shorter distances, like route_id 32 from CLD to CHI (246 miles), route_id 28 from BOS to CDC (246 miles), and route_id 27 from BOI to CLD (578 miles).
    - Shorter routes are typically domestic flights or flights within a continent.

❖ **Revenue Management:**

- Optimizing pricing strategies based on historical ticket sales and customer preferences.
- Identifying the most profitable travel classes.

```
SELECT class_id, COUNT(*) as tickets_sold, AVG(Price_per_ticket) as avg_ticket_price,
                SUM(Price_per_ticket) as total_revenue
FROM ticketdetails
GROUP BY class_id
order by total_revenue DESC;
```

Result Grid | Filter Rows: | Export:

class_id	tickets_sold	avg_ticket_price	total_revenue
Bussiness	13	464.1538	6034
First Class	13	375.7692	4885
Economy Plus	10	246.0000	2460
Economy	14	142.1429	1990

Inferences Drawn :

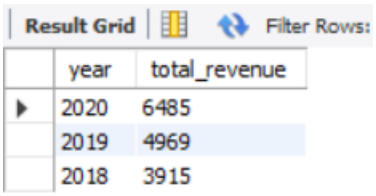
1. Ticket Sales by Class:
   - "Economy" class has the highest number of ticket sales (14), indicating that it is a popular choice among passengers.
   - "Economy Plus","First Class" and "Bussiness" classes have 10 and 13 ticket sales, respectively, suggesting a moderate level of demand.

2. Total Revenue:
   - "Bussiness" class generates the highest total revenue ($6034), primarily due to its higher average ticket prices.
   - "First Class" follows with a total revenue of $4885.
   - "Economy Plus" generates $2460 in total revenue, while "Economy" generates $1990.

3. Pricing Strategy Considerations:
   - The higher average ticket prices in premium classes ("Bussiness" and "First Class") contribute significantly to the overall revenue. Consider maintaining or optimizing pricing strategies for these classes to maximize revenue.
   - "Economy" class, while having a lower average ticket price, contributes to a substantial number of ticket sales. Adjusting pricing or promotions in this class may attract more budget-conscious passengers.

 # Yearly Revenue Trend:

```
SELECT DATE_FORMAT(STR_TO_DATE(p_date, '%m/%d/%Y'), '%Y') AS year,
        SUM(Price_per_ticket * no_of_tickets) AS total_revenue
FROM ticketdetails
GROUP BY year
ORDER BY total_revenue desc;
```



Result Grid | Filter Rows:

year	total_revenue
2020	6485
2019	4969
2018	3915

Inferences Drawn:

1. Revenue Growth:
    - There is an apparent increase in revenue from 2018 to 2019, and further growth from 2019 to 2020.
    - This indicates positive business performance and increasing ticket sales over the years.

 # Year-wise Monthy Revenue

```
SELECT YEAR(STR_TO_DATE(p_date, '%m/%d/%Y')) AS year,
        MONTHNAME(STR_TO_DATE(p_date, '%m/%d/%Y')) AS month,
        SUM(Price_per_ticket * no_of_tickets) AS monthly_revenue
FROM ticketdetails
GROUP BY year, month
ORDER BY year DESC, monthly_revenue DESC;
```

Result Grid | Filter Rows:

year	month	monthly_revenue
2020	March	980
2020	November	930
2020	July	860
2020	April	780
2020	September	760
2020	December	675
2020	August	550
2020	February	440
2020	October	270
2020	May	240
2019	January	1690
2019	March	499
2019	July	480
2019	October	410
2019	April	395
2019	August	365
2019	May	250
2019	November	250
2019	June	190
2019	December	170
2019	September	150
2019	February	120
2018	April	510
2018	December	430
2018	July	430

Result 23

Inferences Drawn:

1. **2020 Monthly Revenue Analysis:**
    - The highest monthly revenue in 2020 is observed in March (980), followed by November (930) and July (860).
    - December has the lowest monthly revenue in 2020 (675).

2. **2019 Monthly Revenue Analysis:**
    - January stands out as the month with the highest revenue in 2019 (1690), followed by March (499) and July (480).
    - September and February have the lowest monthly revenues in 2019 (150 and 120, respectively).

3. 2018 Monthly Revenue Analysis:
    - April has the highest monthly revenue in 2018 (510), followed by December, July, and November (all with 430).
    - February and September have the lowest monthly revenues in 2018 (100 and 130, respectively).

4. Seasonal Trends:
    - March and November seem to be strong months across all years.

5. Business Planning:
    - We can plan marketing and promotional activities during high-revenue months. For instance,
    - March and November may be months to focus on promotions or events.

6. Identifying Challenges:
    - Low-revenue months can be explored to identify challenges or opportunities for improvement. For example, if certain months consistently have lower revenue, it might be worth investigating potential causes.

 # Now, lets extract few information for further analysis using concepts of Window function, Roll up function, Stored Procedure etc.

 # Number of tickets booked by Passengers

```
select c2.customer_id, CONCAT(c2.first_name,' ',c2.last_name) AS 'NAME',
COUNT(t2.no_of_tickets) AS "TicketsBooked"
from customer c2
JOIN ticketdetails t2 USING(customer_id)
GROUP BY c2.customer_id, NAME
order by TicketsBooked desc;
```

Result Grid | Filter Rows:

customer_id	NAME	TicketsBooked
11	Roger Walson	3
19	Joyce Paul	3
5	Aaron Kim	3
18	Gloria Richie	2
4	Cathenna Emily	2
25	Moss Morris	2
8	Floyd Ted	2
9	Leo Travis	2
14	Carol Vernon	2
20	Sara Oliver	2
29	Watson Ronald	2
1	Julie Sam	2
2	Steve Ryan	2
46	Louis Douglas	2
16	Chirstine Willis	1

Result 40

Inferences Drawn:

- Roger Walson, Joyce Paul, and Aaron Kim have booked the maximum number of tickets (3 each).
- Businesses could consider identifying customers with consistent high-ticket bookings for potential VIP programs or special offers to encourage continued engagement.

 # Lets identify the customers who have travelled by Economy Plus class using Group By and Having clause on the passengers_on_flights table.

```
select c4.customer_id, CONCAT(c4.first_name,' ',c4.last_name) as Passenger_name
from customer c4
JOIN passengersonflight p4  USING(customer_id)
GROUP BY c4.customer_id, Passenger_name
HAVING SUM(p4.class_id="Bussiness")>0;
```

Result Grid | Filter Rows:

customer_id	Passenger_name
2	Steve Ryan
5	Aaron Kim
7	Anderson Stewart
11	Roger Walson
15	Linda William
21	Chirsty Josh
24	Calvin Willis
25	Moss Morris
29	Watson Ronald
33	Mark Ethan
49	Russell Peter

Inference Drawn:

- 11 passengers were travelling by Business class. So, the airline could consider providing complementary services in order to retain the passengers.

 # Lets find the maximum ticket price for each class using window functions on the ticket_details table.

```
select *,
MAX(Price_per_ticket) OVER (partition by class_id) AS MAX_TICKET_PRICE
from ticketdetails;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

p_date	customer_id	aircraft_id	class_id	no_of_tickets	a_code	Price_per_ticket	brand	MAX_TICKET_PRICE
3/12/2020	33	CRJ900	Bussiness	1	BOH	490	Bristish Airways	510
4/1/2018	29	ERJ142	Bussiness	1	EME	510	Jet Airways	510
7/7/2020	7	767-301ER	Bussiness	1	BFS	430	Emirates	510
7/7/2019	24	A321	Bussiness	1	CTM	480	Qatar Airways	510
11/1/2018	15	A321	Bussiness	1	BFS	430	Qatar Airways	510
1/25/2019	2	A321	Bussiness	1	YVR	505	Qatar Airways	510
11/11/2020	11	767-301ER	Bussiness	1	AGB	465	Emirates	510
10/22/2019	29	A321	Bussiness	1	PEK	410	Qatar Airways	510
7/1/2018	5	767-301ER	Bussiness	1	BFS	430	Emirates	510
11/8/2020	11	767-301ER	Bussiness	1	AGB	465	Emirates	510
12/26/2018	27	767-301ER	Economy	1	DAL	130	Emirates	190
9/1/2018	2	767-301ER	Economy	1	DAL	130	Emirates	190
12/1/2018	28	ERJ142	Economy	1	BHX	170	Jet Airways	190
5/5/2020	5	ERJ142	Economy	1	CTM	120	Jet Airways	190
2/2/2019	14	ERJ142	Economy	1	CTM	120	Jet Airways	190
12/19/2018	31	767-301ER	Economy	1	DAL	130	Emirates	190
10/10/2020	10	A321	Economy	1	MCO	135	Qatar Airways	190
5/30/2020	5	ERJ142	Economy	1	CTM	120	Jet Airways	190
10/7/2020	46	A321	Economy	1	MCO	135	Qatar Airways	190
6/6/2019	18	767-301ER	Economy	1	YVR	190	Emirates	190
9/21/2019	25	767-301ER	Economy	1	BOH	150	Emirates	190
12/24/2019	14	767-301ER	Economy	1	BHX	170	Emirates	190
2/1/2018	19	767-301ER	Economy	1	AGB	100	Emirates	190
5/1/2018	8	A321	Economy	1	YVR	190	Qatar Airways	190
2/2/2020	22	ERJ142	Economy...	1	AGB	220	Jet Airways	295
12/12/2020	19	CRJ900	Economy...	1	DEN	225	Bristish Airways	295

Result 50

Inference Drawn:

1. **Price Variation Across Classes:**
   - Different travel classes ("Economy," "Economy Plus,", "First Class" and "Bussiness") have distinct maximum ticket prices.
   - Bussiness Class tickets have the highest maximum price (510), followed by, First Class, Economy Plus and Economy classes.*/

 # Lets calculate the total price of all tickets booked by a customer across different aircraft IDs using rollup function.

```
select p5.customer_id,p5.aircraft_id, sum(t5.no_of_tickets * t5.Price_per_ticket) AS "Total Price"
from passengersonflight p5
JOIN ticketdetails t5 using(customer_id)
GROUP BY p5.customer_id,p5.aircraft_id WITH ROLLUP;
```

Result Grid | Filter Rows:

customer_id	aircraft_id	Total Price
1	CRJ900	570
1	ERJ142	570
1	NULL	1140
2	767-301ER	635
2	A321	635
2	NULL	1270
4	767-301ER	1560
4	NULL	1560
5	767-301ER	670
5	ERJ142	1340
5	NULL	2010
7	767-301ER	430
7	NULL	430
8	A321	930
8	NULL	930
9	767-301ER	770
9	CRJ900	770
9	NULL	1540
10	A321	135
10	NULL	135
11	767-301ER	2450
11	ERJ142	1225
11	NULL	3675
13	A321	395
13	NULL	395

Result 51

Lets create a stored procedure to get the details of all passengers flying between a range of routes defined in run time.

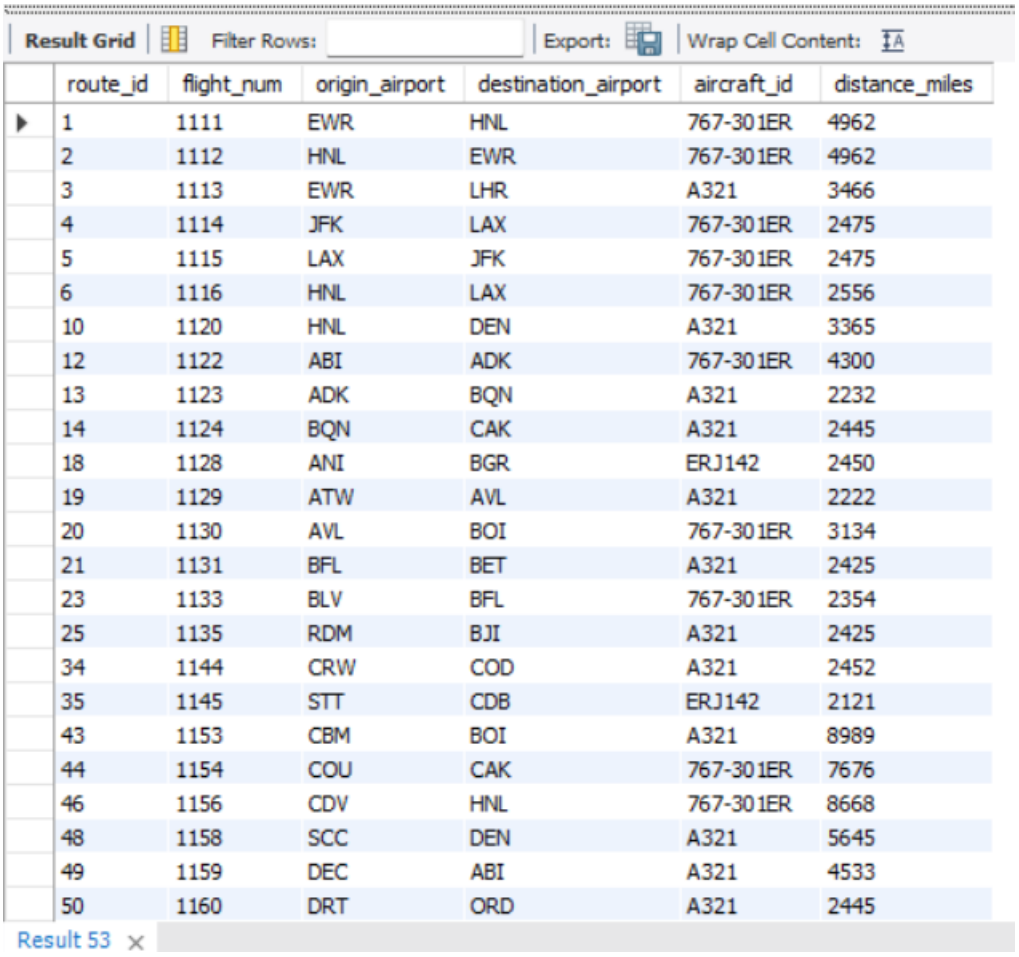
```
DELIMITER &&
CREATE PROCEDURE getAllDetails(IN start_route int, IN end_route int)
BEGIN
        select * from passengersonflight
   WHERE route_id BETWEEN start_route AND end_route;
END &&
DELIMITER ;
CALL getAllDetails(1,14);
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

customer_id	aircraft_id	route_id	depart	arrival	seat_num	class_id	travel_date	flight_num
18	767-301ER	1	EWR	HNL	13FC	First Class	01-04-2018	1111
2	767-301ER	4	JFK	LAX	01E	Economy	02-09-2018	1114
4	767-301ER	4	JFK	LAX	03FC	First Class	30-04-2020	1114
11	767-301ER	4	JFK	LAX	05B	Bussiness	09-11-2020	1114
4	767-301ER	5	LAX	JFX	02FC	First Class	06-04-2020	1115
11	767-301ER	5	LAX	JFX	04B	Bussiness	12-11-2020	1115
46	A321	8	ORD	EWR	12FC	First Class	08-07-2011	1118
1	ERJ142	9	DEN	LAX	01EP	Economy Plus	26-12-2019	1119
29	ERJ142	9	DEN	LAX	11B	Bussiness	03-05-2018	1119
10	A321	10	HNL	DEN	05E	Economy	11-10-2020	1120
5	767-301ER	12	ABI	ADK	02B	Bussiness	02-07-2018	1122
17	A321	13	ABI	ADK	04EP	Economy Plus	03-06-2019	1123
13	A321	13	ADK	BQN	06FC	First Class	05-01-2019	1123
15	A321	14	BQN	CAK	06B	Bussiness	02-11-2018	1124
24	A321	14	BQN	CAK	08B	Bussiness	22-07-2019	1124

Lets create a stored procedure that extracts all the details from the routes table where the travelled distance is more than 2000 miles.

```
DELIMITER &&
CREATE PROCEDURE getAllRoutes()
BEGIN
        select * from routes
   WHERE distance_miles >2000;
END &&
DELIMITER ;

CALL getAllRoutes();
```



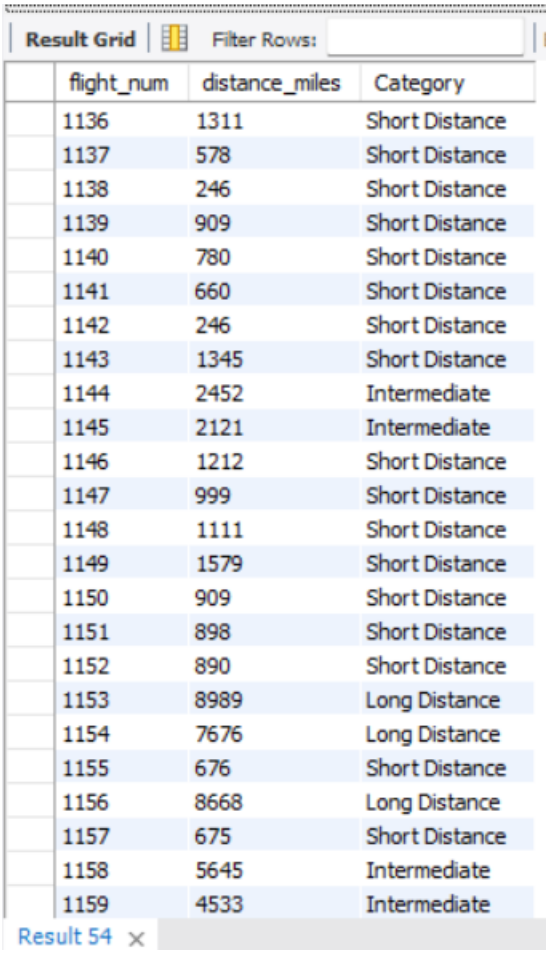
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

route_id	flight_num	origin_airport	destination_airport	aircraft_id	distance_miles
1	1111	EWR	HNL	767-301ER	4962
2	1112	HNL	EWR	767-301ER	4962
3	1113	EWR	LHR	A321	3466
4	1114	JFK	LAX	767-301ER	2475
5	1115	LAX	JFK	767-301ER	2475
6	1116	HNL	LAX	767-301ER	2556
10	1120	HNL	DEN	A321	3365
12	1122	ABI	ADK	767-301ER	4300
13	1123	ADK	BQN	A321	2232
14	1124	BQN	CAK	A321	2445
18	1128	ANI	BGR	ERJ142	2450
19	1129	ATW	AVL	A321	2222
20	1130	AVL	BOI	767-301ER	3134
21	1131	BFL	BET	A321	2425
23	1133	BLV	BFL	767-301ER	2354
25	1135	RDM	BJI	A321	2425
34	1144	CRW	COD	A321	2452
35	1145	STT	CDB	ERJ142	2121
43	1153	CBM	BOI	A321	8989
44	1154	COU	CAK	767-301ER	7676
46	1156	CDV	HNL	767-301ER	8668
48	1158	SCC	DEN	A321	5645
49	1159	DEC	ABI	A321	4533
50	1160	DRT	ORD	A321	2445

Result 53

 # Lets create a stored procedure that groups the distance travelled by each flight into three categories. The categories are, short distance travel (SDT) for >=0 AND <= 2000 miles, intermediate distance travel (IDT) for >2000 AND <=6500, and long-distance travel (LDT) for >6500.

```
DELIMITER &&
CREATE PROCEDURE getCategories()
BEGIN
        select flight_num,distance_miles,
   CASE
                WHEN distance_miles BETWEEN 0 AND 2000 THEN  "Short Distance"
                WHEN distance_miles BETWEEN 2001 AND 6500 THEN  "Intermediate"
                ELSE "Long Distance"
   END AS "Category"
   from routes;
END &&
DELIMITER ;

CALL getCategories();
```



Result Grid | Filter Rows:

flight_num	distance_miles	Category
1136	1311	Short Distance
1137	578	Short Distance
1138	246	Short Distance
1139	909	Short Distance
1140	780	Short Distance
1141	660	Short Distance
1142	246	Short Distance
1143	1345	Short Distance
1144	2452	Intermediate
1145	2121	Intermediate
1146	1212	Short Distance
1147	999	Short Distance
1148	1111	Short Distance
1149	1579	Short Distance
1150	909	Short Distance
1151	898	Short Distance
1152	890	Short Distance
1153	8989	Long Distance
1154	7676	Long Distance
1155	676	Short Distance
1156	8668	Long Distance
1157	675	Short Distance
1158	5645	Intermediate
1159	4533	Intermediate

Result 54

 # Lets extract ticket purchase date, customer ID, class ID and specify if the complimentary services are provided for the specific class using a stored function in stored procedure on the ticket_details table.

 # Condition: If the class is Business and Economy Plus, then complimentary services are given as Yes, else it is No

```
DELIMITER &&
CREATE PROCEDURE getComplementaryServiceDetails()
BEGIN
select p_date,customer_id,class_id,
CASE
        WHEN class_id = "Bussiness"  THEN "YES"
   WHEN class_id = "Economy Plus"  THEN "YES"
   ELSE "NO"
END AS "Complimentary Services"
from ticketdetails;
END &&
DELIMITER ;
```

CALL getComplementaryServiceDetails();

Result Grid | Filter Rows: | Export:

p_date	customer_id	class_id	Complimentary Services
12/26/2018	27	Economy	NO
2/2/2020	22	Economy Plus	YES
3/3/2020	21	Bussiness	YES
4/4/2020	4	First Class	NO
5/5/2020	5	Economy	NO
7/7/2020	7	Bussiness	YES
8/8/2020	8	Economy Plus	YES
9/9/2020	9	First Class	NO
10/10/2020	10	Economy	NO
11/11/2020	11	Bussiness	YES
12/12/2020	19	Economy Plus	YES
1/1/2019	13	First Class	NO
2/2/2019	14	Economy	NO
3/3/2019	25	Bussiness	YES
4/4/2019	16	First Class	NO
5/3/2019	17	Economy Plus	YES
6/6/2019	18	Economy	NO
7/7/2019	24	Bussiness	YES
8/9/2019	20	First Class	NO
9/21/2019	25	Economy	NO
10/22/2019	29	Bussiness	YES

Result 55

CONCLUSION:

In conclusion, the SQL analysis of the aviation company's dataset yielded actionable insights for strategic decision-making.

- For instance, the identification of **Boeing 767-301ER as the preferred aircraft** suggests potential fleet optimization strategies.
- The observation that **male customers predominantly choose Business class** highlights an opportunity for targeted marketing and services tailored to this demographic.
- The exploration of departure and destination airports, such as **HNL**(Honolulu International Airport) **and LAX** (Los Angeles International Airport), provides a basis for operational planning and **potential hub** expansion.
- The examination of revenue by class underscores the significance of premium classes like **Business and First Class in generating higher total revenue**.
- The identification of **peak revenue months like March and November** enables strategic promotional planning.
- The use of stored procedures and functions further streamlines data extraction for ongoing analysis, supporting the aviation company's goal of sustained growth and enhanced customer satisfaction.