

# AIRLINE ANALYSIS IN MySQL

WE COVER BASIC SQL , INTERMEDIATE SQL, ADVANCE SQL.

## BASIC SQL QUESTIONS

### 1. LIST ALL COLUMNS AND THE FIRST 5 ROWS FROM THE DATASET

--- Basic SQL Questions

-- 1. List all columns and the first 5 rows from the dataset.

```
SELECT * FROM airline_quality  
LIMIT 5;
```

	ID	Gender	Age	Customer Type	Type of Travel	Class	Flight Distance	Departure Delay	Arrival Delay	Departure and Arrival Time Convenience	Ease of Online Booking	Check Service
▶	1	Male	48	First-time	Business	Business	821	2	5	3	3	4
	2	Female	35	Returning	Business	Business	821	26	39	2	2	3
	3	Male	41	Returning	Business	Business	853	0	0	4	4	4
	4	Male	50	Returning	Business	Business	1905	0	0	2	2	3
	5	Female	49	Returning	Business	Business	3470	0	1	3	3	3

### 2. COUNT THE TOTAL NUMBER OF RECORDS.

--- 2. Count the total number of records.

- `SELECT COUNT(*) AS total_records FROM airline_quality;`

total_records
25309

### 3. FIND THE NUMBER OF MALE AND FEMALE PASSENGERS.

--- 3. Find the number of male and female passengers.

```
SELECT Gender, COUNT(*) AS count  
FROM airline_quality  
GROUP BY Gender;
```

Result Grid		
	Gender	count
▶	Male	12470
	Female	12839

### 4. GET DISTINCT TRAVEL CLASSES AVAILABLE

--- 4. Get distinct travel classes available

```
SELECT DISTINCT Class FROM airline_quality;
```

Result Grid	
	Class
▶	Business
	Economy
	Economy Plus

### 5. SHOW ALL UNIQUE VALUES IN THE CUSTOMER TYPE COLUMN.

```
5     --- 5. Show all unique values in the Customer Type column.  
6  
7 •   SELECT DISTINCT "Customer Type" FROM airline_quality;  
8
```

Result Grid    F	
	Customer Type
▶	Customer Type

## 6. . COUNT HOW MANY FLIGHTS HAD A RECORDED ARRIVAL DELAY.

```
--- 6. Count how many flights had a recorded Arrival Delay.  
  
•   SELECT COUNT(*) AS arrival_delay_count  
    FROM airline_quality  
    WHERE "Arrival Delay" IS NOT NULL;
```

Result Grid    Filter Rows:	
	arrival_delay_count
▶	25309

## 7. . FIND THE MINIMUM AND MAXIMUM AGE OF PASSENGERS.

```
--- 7. . Find the minimum and maximum age of passengers.  
  
•   SELECT MIN(Age) AS min_age, MAX(Age) AS max_age  
    FROM airline_quality;
```

	min_age	max_age
▶	7	85

## 8. LIST THE NUMBER OF PASSENGERS BY TYPE OF TRAVEL.

--- 8. List the number of passengers by Type of Travel.

```
• SELECT "Type of Travel", COUNT(*) AS total
  FROM airline_quality
  GROUP BY "Type of Travel";
```

	Type of Travel	total
▶	Type of Travel	25309

## INTERMEDIATE SQL QUESTIONS

### 1.CALCULATE THE AVERAGE AGE OF PASSENGERS.

```
-- ○ Intermediate SQL Questions
--- 5. Calculate the average age of passengers.

• SELECT AVG(Age) AS average_age FROM airline_quality;
```

Result Grid	
	average_age
▶	39.3787

## 1. FIND THE AVERAGE FLIGHT DISTANCE BY TRAVEL CLASS

--- 6. Find the average flight distance by travel class

```
SELECT Class, AVG("Flight Distance") AS avg_distance  
FROM airline_quality  
GROUP BY Class;
```

Result Grid		
	Class	avg_distance
▶	Business	0
	Economy	0
	Economy Plus	0

## 2. Get the number of satisfied vs. dissatisfied customers

--- 7. Get the number of satisfied vs. dissatisfied customers.

```
SELECT Satisfaction, COUNT(*) AS count  
FROM airline_quality  
GROUP BY Satisfaction;
```

Satisfaction	count
Neutral or Dissatisfied	14640
Satisfied	10669

### 3. List the top 5 flights with the highest departure delays.

--- 8. List the top 5 flights with the highest departure delays.

```
SELECT ID, "Departure Delay"
FROM airline_quality
ORDER BY "Departure Delay" DESC
LIMIT 5;
```

ID	Departure Delay
1	Departure Delay
2	Departure Delay
3	Departure Delay
4	Departure Delay
5	Departure Delay

### 4. Find the minimum, maximum, and average arrival delay.

--- Find the minimum, maximum, and average arrival delay.

```
SELECT
    MIN("Arrival Delay") AS min_arrival_delay,
    MAX("Arrival Delay") AS max_arrival_delay,
    AVG("Arrival Delay") AS avg_arrival_delay
FROM airline_quality;
```

	min_arrival_delay	max_arrival_delay	avg_arrival_delay
▶ Arrival Delay	Arrival Delay	0	

**5. Count how many flights had no delays at all (both arrival and departure = 0).**

```
--- Count how many flights had no delays at all (both arrival and departure = 0).
```

```
SELECT COUNT(*) AS no_delay_flights
FROM airline_quality
WHERE "Arrival Delay" = 0 AND "Departure Delay" = 0;
```

	no_delay_flights
▶	25309

**6. Group passengers by age brackets (e.g., 18–30, 31–45, etc.).**

```
--- Group passengers by age brackets (e.g., 18–30, 31–45, etc.).
```

```
SELECT
CASE
    WHEN Age BETWEEN 18 AND 30 THEN '18-30'
    WHEN Age BETWEEN 31 AND 45 THEN '31-45'
    WHEN Age BETWEEN 46 AND 60 THEN '46-60'
    ELSE '60+'
END AS age_group,
COUNT(*) AS count
FROM airline_quality
GROUP BY age_group;
```

	age_group	count
▶	46-60	7058
	31-45	8281
	18-30	5987
	60+	3983

## 7. Find the average rating for each customer type across all service categories.

```
--- Find the average rating for each customer type across all service categories.
SELECT "Customer Type",
       AVG("Check-in Service") AS avg_checkin,
       AVG("Online Boarding") AS avg_boarding,
       AVG("Seat Comfort") AS avg_seat_comfort
FROM airline_quality
GROUP BY "Customer Type";
```

	Customer Type	avg_checkin	avg_boarding	avg_seat_comfort
▶	Customer Type	0	0	0

## ● Advanced SQL Questions

### 1. Find the average in-flight service rating for each customer type, grouped by travel type

- 9. Find the average in-flight service rating for each customer type, grouped by travel type
- ```
SELECT "Customer Type", "Type of Travel", AVG("In-flight Service") AS avg_service
FROM airline_quality
GROUP BY "Customer Type", "Type of Travel";
```

Result Grid | Filter Rows: Export:

|   | Customer Type | Type of Travel | avg_service |
|---|---------------|----------------|-------------|
| ▶ | Customer Type | Type of Travel | 0           |

## 2. Identify if there's a correlation between satisfaction and in-flight entertainment.

```
--- 10. Identify if there's a correlation between satisfaction and in-flight entertainment.

• SELECT "In-flight Entertainment", Satisfaction, COUNT(*) AS count
  FROM airline_quality
  GROUP BY "In-flight Entertainment", Satisfaction
  ORDER BY "In-flight Entertainment";
```

Result Grid | Filter Rows: Export:

|   | In-flight Entertainment | Satisfaction            | count |
|---|-------------------------|-------------------------|-------|
| ▶ | In-flight Entertainment | Neutral or Dissatisfied | 14640 |
|   | In-flight Entertainment | Satisfied               | 10669 |

## 3. Calculate average ratings across all service categories for each travel class.

```
--- 11. Calculate average ratings across all service categories for each travel class.

• SELECT Class,
      AVG("Seat Comfort") AS avg_seat,
      AVG("Leg Room Service") AS avg_leg_room,
      AVG("Cleanliness") AS avg_clean,
      AVG("Food and Drink") AS avg_food,
      AVG("In-flight Service") AS avg_service,
      AVG("In-flight Wifi Service") AS avg_wifi,
      AVG("In-flight Entertainment") AS avg_entertainment
    FROM airline_quality
   GROUP BY Class;
```

|   | Class        | avg_seat | avg_leg_room | avg_clean | avg_food | avg_service | avg_wifi | avg_entertainment |
|---|--------------|----------|--------------|-----------|----------|-------------|----------|-------------------|
| ▶ | Business     | 0        | 0            | 0         | 0        | 0           | 0        | 0                 |
|   | Economy      | 0        | 0            | 0         | 0        | 0           | 0        | 0                 |
|   | Economy Plus | 0        | 0            | 0         | 0        | 0           | 0        | 0                 |

#### 4. Find customers who experienced both arrival and departure delays greater than 30 minutes

--- 12. Find customers who experienced both arrival and departure delays greater than 30 minutes

```
SELECT *
FROM airline_quality
WHERE "Departure Delay" > 30 AND "Arrival Delay" > 30;
```

| ID | Gender | Age | Customer Type | Type of Travel | Class | Flight Distance | Departure Delay | Arrival Delay | Departure and Arrival Time Convenience | Ease of Online Booking | Check-in Service |
|----|--------|-----|---------------|----------------|-------|-----------------|-----------------|---------------|----------------------------------------|------------------------|------------------|
|    |        |     |               |                |       |                 |                 |               |                                        |                        |                  |

#### 5. Identify customer segments (based on gender and travel type) with highest satisfaction rates.

--- Identify customer segments (based on gender and travel type) with highest satisfaction rates.

```
SELECT Gender, "Type of Travel",
       COUNT(*) AS total,
       SUM(CASE WHEN Satisfaction = 'Satisfied' THEN 1 ELSE 0 END) AS satisfied_count,
       ROUND(
           100.0 * SUM(CASE WHEN Satisfaction = 'Satisfied' THEN 1 ELSE 0 END) / COUNT(*), 2
       ) AS satisfaction_rate
  FROM airline_quality
 GROUP BY Gender, "Type of Travel"
 ORDER BY satisfaction_rate DESC;
```

|   | Gender | Type of Travel | total | satisfied_count | satisfaction_rate |
|---|--------|----------------|-------|-----------------|-------------------|
| ▶ | Male   | Type of Travel | 12470 | 5262            | 42.20             |
|   | Female | Type of Travel | 12839 | 5407            | 42.11             |

#### 6. Find the top 3 age groups that fly business class the most.

```

--- Find the top 3 age groups that fly business class the most.

SELECT
CASE
    WHEN Age BETWEEN 18 AND 30 THEN '18-30'
    WHEN Age BETWEEN 31 AND 45 THEN '31-45'
    WHEN Age BETWEEN 46 AND 60 THEN '46-60'
    ELSE '60+'
END AS age_group,
COUNT(*) AS count
FROM airline_quality
WHERE Class = 'Business'
GROUP BY age_group
ORDER BY count DESC
LIMIT 3;

```

Result Grid | Filter Rows:

|   | age_group | count |
|---|-----------|-------|
| ▶ | 31-45     | 4259  |
|   | 46-60     | 3923  |
|   | 18-30     | 2114  |

## 7. Compare satisfaction between first-time and returning customers.

```

--- Compare satisfaction between first-time and returning customers.

SELECT "Customer Type",
Satisfaction,
COUNT(*) AS count
FROM airline_quality
GROUP BY "Customer Type", Satisfaction
ORDER BY "Customer Type", Satisfaction;

```

Result Grid | Filter Rows: Export:

|   | Customer Type | Satisfaction            | count |
|---|---------------|-------------------------|-------|
| ▶ | Customer Type | Neutral or Dissatisfied | 14640 |
|   | Customer Type | Satisfied               | 10669 |

## 8. Get the correlation between delay time and satisfaction.

```
--- Get the correlation between delay time and satisfaction.

|>

SELECT
CASE
    WHEN "Arrival Delay" + "Departure Delay" > 60 THEN 'High Delay'
    WHEN "Arrival Delay" + "Departure Delay" > 30 THEN 'Moderate Delay'
    ELSE 'Low/No Delay'
END AS delay_level,
Satisfaction,
COUNT(*) AS count
FROM airline_quality
GROUP BY delay_level, Satisfaction
ORDER BY delay_level;
```

Result Grid | Filter Rows: Export:

|   | delay_level  | Satisfaction            | count |
|---|--------------|-------------------------|-------|
| ▶ | Low/No Delay | Neutral or Dissatisfied | 14640 |
|   | Low/No Delay | Satisfied               | 10669 |

## 9. Determine if online services (boarding, booking) affect satisfaction.

```
--- Determine if online services (boarding, booking) affect satisfaction.
```

```
SELECT
    ROUND(AVG("Online Boarding"), 2) AS avg_boarding,
    ROUND(AVG("Ease of Online Booking"), 2) AS avg_booking,
    Satisfaction
FROM airline_quality
GROUP BY Satisfaction;
```

| Result Grid |              |             | Filter Rows:            | Export: |
|-------------|--------------|-------------|-------------------------|---------|
|             | avg_boarding | avg_booking | Satisfaction            |         |
| ▶           | 0            | 0           | Neutral or Dissatisfied |         |
|             | 0            | 0           | Satisfied               |         |

## Optional: Data Cleaning Queries (before analysis)

### 1. Check for any null values in Arrival Delay.

```
--- Optional: Data Cleaning Queries (before analysis)
--- Check for any null values in Arrival Delay.
```

```
SELECT COUNT(*) AS null_arrival_delay
FROM airline_quality
WHERE "Arrival Delay" IS NULL;
```

| Result Grid |                    | Filter |
|-------------|--------------------|--------|
|             | null_arrival_delay |        |
| ▶           | 0                  |        |

### 2. Remove records with null Arrival Delay (if needed):

```
--- Remove records with null Arrival Delay (if needed):
```

```
DELETE FROM airline_quality  
WHERE "Arrival Delay" IS NULL;
```

### 3. Convert satisfaction into binary values for ML (Satisfied = 1, Neutral/Dissatisfied = 0).

```
--- Convert satisfaction into binary values for ML (Satisfied = 1, Neutral/Dissatisfied = 0).
```

```
SELECT *,  
CASE  
    WHEN Satisfaction = 'Satisfied' THEN 1  
    ELSE 0  
END AS satisfaction_binary  
FROM airline_quality;
```

| ID | Gender | Age | Customer Type | Type of Travel | Class    | Flight Distance | Departure Delay | Arrival Delay | Departure and Arrival Time Convenience | Ease of Online Booking |
|----|--------|-----|---------------|----------------|----------|-----------------|-----------------|---------------|----------------------------------------|------------------------|
| 1  | Male   | 48  | First-time    | Business       | Business | 821             | 2               | 5             | 3                                      | 3                      |
| 2  | Female | 35  | Returning     | Business       | Business | 821             | 26              | 39            | 2                                      | 2                      |
| 3  | Male   | 35  | Returning     | Business       | Business | 853             | 0               | 0             | 4                                      | 4                      |
| 4  | Male   | 50  | Returning     | Business       | Business | 1905            | 0               | 0             | 2                                      | 2                      |
| 5  | Female | 49  | Returning     | Business       | Business | 3470            | 0               | 1             | 3                                      | 3                      |
| 6  | Male   | 43  | Returning     | Business       | Business | 3788            | 0               | 0             | 4                                      | 4                      |
| 7  | Male   | 43  | Returning     | Business       | Business | 1963            | 0               | 0             | 3                                      | 3                      |
| 8  | Female | 60  | Returning     | Business       | Business | 853             | 0               | 3             | 3                                      | 4                      |
| 9  | Male   | 50  | Returning     | Business       | Business | 2607            | 0               | 0             | 1                                      | 1                      |

| it | Leg Room Service | Cleanliness | Food and Drink | In-flight Service | In-flight Wifi Service | In-flight Entertainment | Baggage Handling | Satisfaction            | satisfaction_binary |
|----|------------------|-------------|----------------|-------------------|------------------------|-------------------------|------------------|-------------------------|---------------------|
| 2  | 5                | 5           | 5              | 3                 | 5                      | 5                       | 5                | Neutral or Dissatisfied | 0                   |
| 5  | 5                | 3           | 5              | 2                 | 5                      | 5                       | 5                | Satisfied               | 1                   |
| 3  | 5                | 5           | 3              | 4                 | 3                      | 3                       | 3                | Satisfied               | 1                   |
| 5  | 4                | 4           | 5              | 2                 | 5                      | 5                       | 5                | Satisfied               | 1                   |
| 4  | 5                | 4           | 3              | 3                 | 3                      | 3                       | 3                | Satisfied               | 1                   |
| 4  | 3                | 3           | 4              | 4                 | 4                      | 4                       | 4                | Satisfied               | 1                   |
| 5  | 4                | 5           | 5              | 3                 | 5                      | 5                       | 5                | Satisfied               | 1                   |
| 4  | 4                | 4           | 3              | 4                 | 3                      | 3                       | 3                | Satisfied               | 1                   |
| 4  | 3                | 3           | 4              | 4                 | 4                      | 4                       | 4                | Neutral or Dissatisfied | 0                   |