Flight Data Analysis and Reporting

Objective:

Analyze airline operations data—including aircraft specifications, flight schedules, bookings, and passenger information—to gain insights into:

- Flight performance and punctuality
- Aircraft utilization and seating efficiency
- Revenue distribution across flights and booking channels
- Passenger behavior and boarding patterns

Key Stakeholders:

Airline management, operations team, revenue management, customer service, and data analysts.

Schema Overview

Your schema includes the following core entities:

- AIRCRAFTS_DATA: Details about aircraft (e.g., model, range).
- AIRPORTS_DATA: Information on airports with spatial coordinates and localized names.
- FLIGHTS: Flight schedules, statuses, and references to departure/arrival airports and aircraft.
- SEATS: Seat details per aircraft including fare conditions.
- BOOKINGS & TICKETS: Reservation details and associated ticket information.
- TICKET_FLIGHTS & BOARDING_PASSES: Junction tables linking tickets with flights, along with fare and boarding details.

Project Steps

1. Data Import

1. Download the Airlines_schema_scripts.zip file from vLearn
2. Unzip the Airlines_schema_scripts.zip file
3. Open MySQL Workbench
4. Connect to your Database
5. Open and Execute the following SQL Scripts:
    a. airlines_data_tickets.sql
    b. airlines_data_aircrafts.sql
    c. airlines_data_airports.sql
    d. airlines_data_boarding_passes.sql
    e. airlines_data_bookings.sql
    f. airlines_data_flights.sql
    g. airlines_data_seats.sql
    h. airlines_data_ticket_flights.sql
6. After successful execution of all scripts, your schema is ready for performing analytics

Hint: Reverse Engineer ER Diagram for visual view of schema

2. Analysis Queries

Write SQL Select Queries for the following:

1. Calculate the Average Delay in Departures for Delayed Flight (3 Marks)

Determine, for flights that actually departed later than scheduled, what the average departure delay is (in minutes).

2. List Flights Grouped by Aircraft to See Usage Frequency (3 Marks)

Show how frequently each aircraft is used by counting the number of flights operated by each aircraft.

3. Calculate Total Revenue per Flight from the Ticket_Flights Table (3 Marks)

   Aggregate the total revenue generated by each flight by summing up the fare amounts from the ticket_flights table.

4. Analyze Boarding Numbers per Flight to Detect Patterns in Boarding Pass Allocation (3 Marks)

   Examine boarding pass data to identify how passengers are being boarded by calculating the average boarding number per flight.

5. Determine Occupancy per Aircraft (3 Marks)

   Identify seat occupancy by comparing the number of boarding passes (occupied seats) to the total seats available on an aircraft.

6. Top 3 Flights by Revenue (3 Marks)

   Identify the three flights that generated the highest revenue based on ticket sales.

7. Average Flight Duration by Aircraft Model (3 Marks)

   Determine the average flight duration for each aircraft model, allowing you to see how flight performance might vary between different models.

8. Flight Count per Airport (Departure) (3 Marks)

   Count how many flights depart from and arrive at each airport to assess airport activity levels.

9. Flight Count per Airport (Arrival) (3 Marks)

   Count how many flights depart from and arrive at each airport to assess airport activity levels.

10. Daily Booking Trends (3 Marks)

Analyze trends in bookings over time by grouping bookings by date, counting total bookings, and summing up the revenue.

11. Frequent Routes Analysis (3 Marks)

Identify the most common routes by counting the number of flights for each departure–arrival airport pair.

12. Passenger Boarding Summary per Flight (3 Marks)

Determine how many passengers boarded each flight by counting the boarding passes issued.

13. Average Boarding Number per Flight (3 Marks)

Assess the boarding process by calculating the average boarding order number for each flight, which may indicate the boarding strategy or sequence.

14. Total Spend per Passenger (3 Marks)

Aggregate the total amount spent by each passenger across all their tickets by joining the tickets and ticket_flights tables.

15. Fare-Condition Distribution (3 Marks)

Analyze the distribution of ticket volume and revenue across different fare conditions.

16. Revenue by Route (3 Marks)

Aggregate total revenue generated by each departure→arrival airport route.

17. Peak Booking Hours (2 Marks)

Identify the top booking hours of the day by volume of bookings.

Note: Kindly write solution for all queries in a single SQL script file

Artifacts to be generated (For Learners):

- SQL Script File(.sql)
- Artifacts generated need to be submitted in vLearn on or before the deadline.
- SQL Script File Name:
    - File name: firstname_lastname_CPDA_Batch.sql
    - E.g., Kartik_Mudaliar_CPDA_B1.sql
- Zip your script file and submit