Flight Performance & Aviation Analytics for AeroStat Airlines

You are a Data Analyst at AeroStat Airlines, a global airline company that operates flights across multiple airports. The company is facing challenges in flight punctuality, cancellations, route efficiency, and passenger experience. Your role is to analyze the dataset using SQL queries, Power BI/Tableau for visualization, and machine learning for predictive analytics.

Dataset
Link: https://drive.google.com/file/d/1M2d_D_h1_t_TIcrPlWGKNqGwb86ofAZ8/view

PART 1: SQL QUERIES (10 Questions)

To upload data to SQL use the following code snippet and edit according to the requirement.
https://github.com/Ayushi0214/SQL-Python-Ecommerce-Project/blob/main/csv_to_sql.py

Perform data analysis using SQL to help AeroStat Airlines make strategic operational decisions.

1 Flight Delay Analysis

- Calculate the average departure and arrival delay for all flights in the last 6 months.

2 Most Frequent Routes

- Identify the top 10 most popular flight routes based on the total number of flights.

3 Airline Performance Ranking

- Rank airlines by their on-time performance (lowest average delay).

4 Cancellation Trends

- Find the top reasons for flight cancellations and their frequency.

5 Airport Congestion Analysis

- Identify airports with the highest number of flight departures and arrivals.

6 Weather-Related Disruptions

- Analyze how weather impacts delays and cancellations.

7 Flight Distance & Duration Trends

- Calculate the average flight duration for different distance ranges.

8️⃣ Day-of-Week Flight Performance

- Find which day of the week has the highest on-time flight performance.

9️⃣ Seasonality in Air Traffic

- Identify which months have the highest number of flights and delays.

🔟 Passenger Connectivity & Hub Efficiency

- Identify which airports act as the largest hubs based on the number of connecting flights.

PART 2: DASHBOARD VISUALIZATION (Power BI/Tableau)

Create a multi-page interactive dashboard to analyze airline performance and operational efficiency.

1️⃣ Flight Operations Overview

- KPIs to Include:
  - Total Flights Operated
  - On-Time Performance Rate
  - Average Flight Duration
  - Average Departure Delay
  - Average Arrival Delay
- Reports to Generate:
  - Flight activity trends over time
  - Airline-wise on-time performance
  - Flight punctuality distribution

2️⃣ Airline Performance Comparison

- KPIs to Include:
  - Airline-wise Punctuality Score
  - Cancellation Rate by Airline
  - Average Delay per Airline
- Reports to Generate:
  - Ranking of airlines based on punctuality
  - Airline-specific cancellation patterns
  - Delay trends across airlines

3️⃣ Route & Airport Analytics

- KPIs to Include:
  - Top 10 Busiest Airports
  - Most Frequent Flight Routes

    - Airport Congestion Index
- Reports to Generate:
    - Heatmap of flight movements across airports
    - Delay trends by airport
    - Traffic flow between major hubs

4️⃣ Flight Cancellation & Delay Patterns

- KPIs to Include:
    - Cancellation Rate by Reason
    - Weather Delay Impact
    - Late Aircraft Delay Contribution
- Reports to Generate:
    - Flight delay reasons analysis
    - Cancellation breakdown by external factors
    - Seasonal disruption trends

5️⃣ Passenger Experience & Service Quality

- KPIs to Include:
    - Customer Satisfaction Proxy Score
    - Missed Connections Rate
    - Delay Recovery Effectiveness
- Reports to Generate:
    - Customer impact of flight delays
    - Passenger connectivity efficiency
    - Factors influencing customer satisfaction

6️⃣ Financial & Operational Efficiency

- KPIs to Include:
    - Flight Cost per Mile
    - Fuel Efficiency by Route
    - Revenue Loss Due to Delays
- Reports to Generate:
    - Financial impact of cancellations and delays
    - Cost efficiency comparison across different routes
    - Airline fuel and operational efficiency insights

PART 3: MACHINE LEARNING (Predictive Flight Delay Model) (OPTIONAL)

1️⃣ Flight Delay Prediction

- Train a classification model to predict whether a flight will be delayed or not based on departure airport, airline, weather, and time of day.

2️⃣ Optimal Flight Scheduling Model

- Build a machine learning model to predict optimal departure times for minimizing delays.

Additional Requirements

- Each dashboard should include at least 5-6 charts per page
- Reports should allow filters, drill-downs, and segmentation
- Analysis should be actionable and relevant for airline decision-making

Expected Deliverables

1. SQL Queries:

- A complete set of queries in a well presented PPT addressing all the tasks in Phase 1.

2. Interactive Dashboard:

- A visually compelling and interactive dashboard in Power BI/Tableau showcasing airline performance insights. Attach multiple screenshots of your dashboard.

3. Machine Learning Model:

- Python code for preprocessing and training the predictive model.
- A detailed evaluation report of the model's performance.

4. Insights Report:

- Comprehensive summary of findings from SQL analysis and visualizations.
- Actionable recommendations for airlines to reduce delays, optimize operations, and enhance customer experience.

Note - Merge all the files into one pdf file and submit it.