BONDA NAVEEN KUMAR

Flight Delay and Cancellation Analysis

1. Introduction

Flight delays and cancellations pose serious challenges to the U.S. aviation industry, affecting passenger satisfaction, airline operations, and economic efficiency. The objective of this project is to analyze the 2015 U.S. domestic flights dataset to uncover the key factors behind delays and cancellations, benchmark airline and airport performance, explore temporal trends, and provide actionable recommendations to improve overall flight reliability.

2. Methodology

2.1 Data Sources:

- **flights.csv**: Core dataset containing detailed records for each flight.
- **airlines.csv**: Airline code-to-name mapping.
- **airports.csv**: Airport information including city, state, and location.

2.2 Tools Used:

- **SQL**: PostgreSQL for data cleaning, transformation, and exploratory analysis.
- **Power BI**: Dashboard creation and data visualization.
- **Excel**: Used for initial inspection and column data-type validation.

2.3 Approach:

- **Phase 1**: Data ingestion, schema creation, and raw data load using CSV imports.
- **Phase 2**: SQL-based cleaning (datetime formatting, null handling), enrichment (cancellation reason descriptions), and creation of derived columns like flight_date, hour.
- **Phase 3**: KPI definitions and exploratory SQL queries to examine performance by airline, airport, time, and delay types.
- **Phase 4**: Power BI dashboard design covering KPIs, airline and airport performance, and temporal delay trends.

3. Key Findings & Analysis

3.1 Primary Causes of Delays & Cancellations:

- **Airline delay** and **Late aircraft delay** were leading contributors to overall delay time.
- **Weather** and **National Airspace System (NAS)** were significant contributors to systemic delays.
- **Cancellations** were most commonly attributed to weather and airline internal decisions.

3.2 Airline Performance:

- Airlines varied widely in terms of **On-Time Performance (OTP)** and **Cancellation Rates**.
- Example:
    - **Alaska Airlines** showed the highest OTP.
    - **Frontier Airlines** had higher cancellation and delay rates.

3.3 Airport Performance:

- Major hubs like **ORD (Chicago O'Hare)** and **ATL (Atlanta)** experienced higher delays.
- Smaller airports had lower delay percentages but also fewer total flights.

3.4 Temporal Trends:

- Delays increased gradually throughout the day, peaking in the late evening (5 PM to 8 PM).
- **Summer months** (June–August) experienced the highest average delay durations.
- **Weekends** showed slightly fewer cancellations but no significant improvement in OTP.

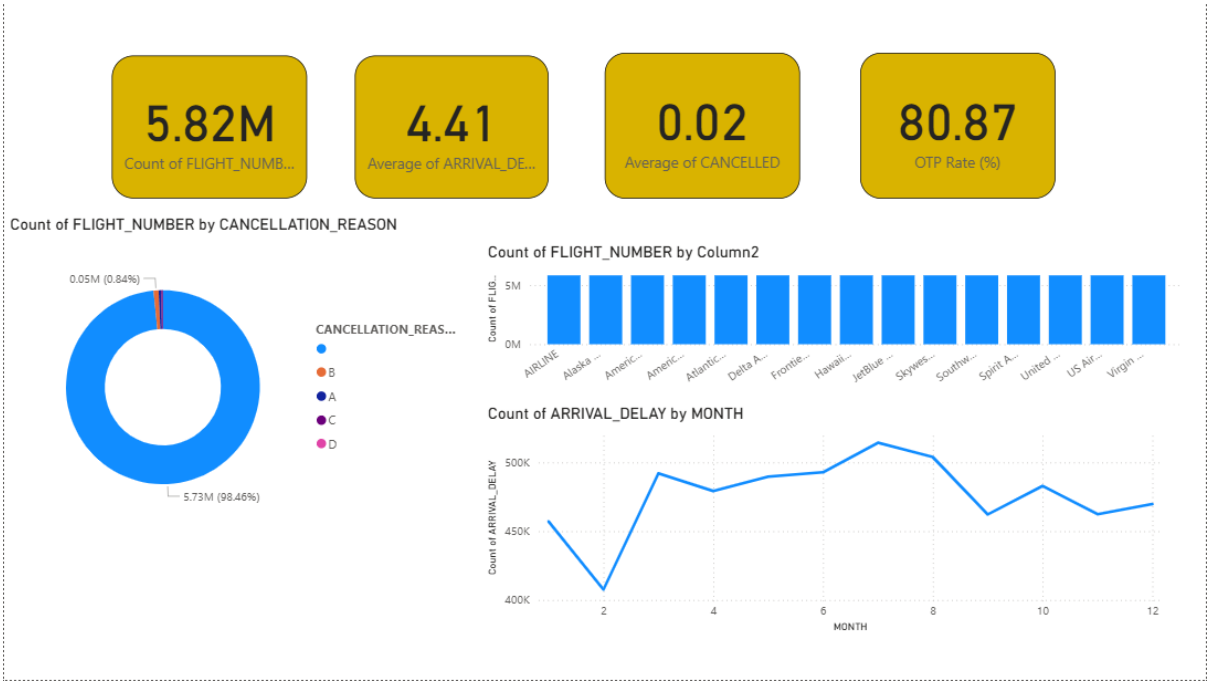
3.5 Routes & Distance:

- Long-haul routes had higher probability of cumulative delays, particularly for connecting flights.

4. Dashboard Overview (Power BI)

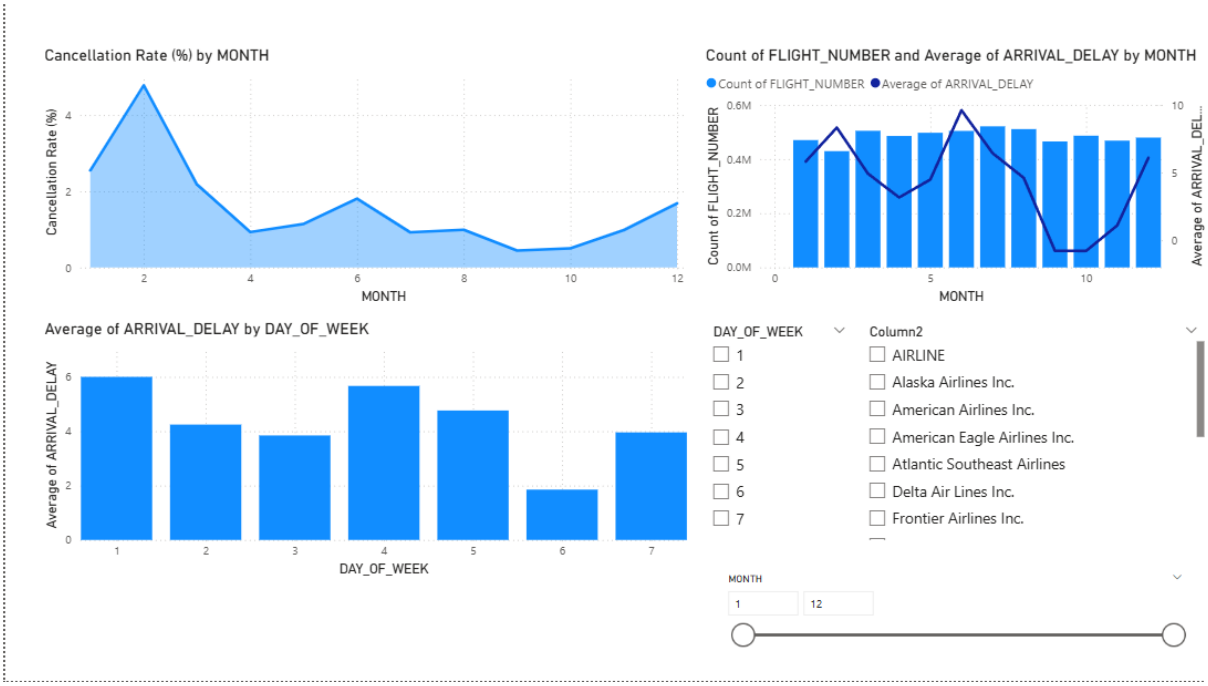
Page 1: Overview Dashboard

- **KPIs:** OTP Rate, Avg Departure Delay, Avg Arrival Delay, Cancellation Rate.
- **Visuals:**
    - KPI cards
    - Pie chart of cancellation reasons
        - Stacked column chart showing delay type contributions



Page 2: Airline Performance Dashboard

- **X-Axis:** Airline Name
- **Y-Axis:** OTP Rate, Avg Delay, Cancellation Rate
- **Visuals:**
    - Clustered bar chart (Comparing delay metrics)
    - Line chart showing temporal OTP trend per airline
    - Matrix visual of airline vs average delay type contribution



5. Recommendations

1. **Improve Morning Operations:** Since delay accumulates over the day, airlines and airports should focus on timely departures in the morning to reduce cascading delays.
2. **Invest in Weather Resilience:** Weather-related cancellations could be reduced by improving ground handling and predictive forecasting systems, particularly in weather-sensitive regions.
3. **Optimize Airline Scheduling:** Airlines should re-evaluate buffer times and aircraft turnaround schedules to reduce late aircraft delays.

6. Conclusion

Through data-driven analysis using SQL and Power BI, we have identified actionable insights into the performance of airlines and airports in the U.S. during 2015. The project highlights critical factors behind delays and cancellations and provides practical recommendations that stakeholders can implement to improve operational efficiency and passenger satisfaction. Further enhancements could involve integrating live weather feeds and incorporating aircraft maintenance data for more robust forecasting.