**A**

**PROJECT REPORT**

**ON**

**AIRLINE METRIX : DELAYS AND CANCELLATIONS**

**SUBMITTED BY**

**Borekar Aditi Pramod**

**SUBMITTED TO**

**SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE**

**IN FULFILLMENT OF DEGREE**

**MASTER OF COMPUTER APPLICATION (SEM-II)**

**UNDER THE GUIDANCE OF**

**Dr. Deeplai Gavhane**

**Through,**



**Sadhu Vaswani Institute of Management Studies for Girls,**

**Koregaon Park, Pune 411001**

**2024-2026**

**DECLARATION BY STUDENT**

To,

The Director,

SVIMS, Koregaon Park, Pune

I, undersigned hereby declare that this project titled, **" AIRLINE METRIX : DELAYS AND CANCELATIONS "** written and submitted by me to SPPU, Pune, in partial fulfilment of the requirement of the award of the degree of **MASTER OF COMPUTER APPLICATION (MCA-II)** under the guidance of **Dr. Shveti Chandan, Dr. Deepali Gavhane** this is my original work.

I further declare that to the best of my knowledge and belief, this project has not been submitted to this or any other University or Institution for the award of any Degree.

**Place: Pune**

**Date: (Aditi Pramod Borekar)**

**ACKNOWLEDGEMENT**

I extend my sincere gratitude to Dr. B. H. Nanwani, Dr. Neeta Raskar-Lokhande Dr. Shevti Chandan, Dr. Deeplai Gavhane for allowing me to carry out the study and for their constant encouragement, valuable suggestions, and guidance during the research work.

I extend my special gratitude to my dearest family members and friends who encouraged and motivated me to complete the project report.

**Place: Pune**

**Date: Aditi Pramod Borekar**

**Airport Analysis and Delay in Power BI**

**1. Project Overview**

**Introduction:**

Flight delays and cancellations are critical operational challenges for the aviation industry. These disruptions not only affect customer satisfaction but also lead to increased operational costs, missed connections, and logistical headaches for both passengers and airport authorities. Understanding the root causes and patterns behind these delays can enable more informed decisions and proactive strategies.

This Power BI project, titled **"Airport Analysis and Delay"**, aims to explore domestic U.S. flight data to analyze delay durations, cancellation frequencies, and how these metrics vary by destination, distance, and date. The interactive dashboards provide a visual platform to uncover trends, correlations, and anomalies in flight performance.

By transforming raw flight data into actionable insights, this report seeks to empower airline and airport management with a comprehensive overview of operational efficiency. The analysis supports smarter scheduling, better resource allocation, and ultimately, a smoother travel experience for passengers.

**Key Objectives**

* Analyze flight delay patterns with respect to distance, date, and destination.
* Identify destinations with the highest cancellation rates.
* Compare expected vs actual departure and arrival times to measure operational efficiency.
* Provide interactive dashboards that facilitate dynamic filtering and exploration of delay trends.

**Target Audience**

* **Airline Operations Teams**: For performance monitoring and delay reduction.
* **Airport Management**: To identify problematic routes and improve scheduling.
* **Aviation Analysts**: For data-backed decision-making and reporting.
* **Academic or Training Institutions**: For students and researchers analyzing transportation data using BI tools.

**2.Challenges Faced**

During the execution of the project, several challenges were encountered:

* **Data Quality Issues**: Many rows contained missing or incomplete values, especially in time columns.
* **Time Conversion**: Scheduled and actual times were in string formats, requiring complex parsing to calculate delays accurately.
* **Visual Design Balance**: Maintaining simplicity in the dashboard while ensuring all necessary information is accessible was a key design challenge.
* **Performance Optimization**: Large datasets initially caused sluggish performance until filtering and aggregation techniques were applied.
* **DAX Complexity**: Writing DAX expressions for measures like average delay and cancellations took time and iterative testing.

**3 Process**

The project followed a structured process to ensure data was accurately transformed into meaningful insights:

1. **Data Collection**:
   * Imported raw flight data from CSV files into Power BI.
2. **Data Cleaning and Preparation**:
   * Removed null or missing values.
   * Standardized time formats for better comparison.
   * Renamed columns for clarity and usability.
3. **Data Transformation**:
   * Calculated actual delays by comparing scheduled and actual times using DAX.
   * Grouped distances into bins to analyze trends over different flight ranges.
4. **Data Modeling**:
   * Created calculated columns and measures.
   * Maintained a single flat table for simplicity, as all fields were in one dataset.
5. **Visualization**:
   * Built interactive visuals including pie charts, line charts, tables, and KPIs.
   * Implemented slicers for user filtering (e.g., by origin, destination, date).
6. **Dashboard Design**:
   * Structured the report into intuitive sections.
   * Applied consistent color themes and layout formatting.
7. **Testing and Validation**:
   * Verified calculated fields.
   * Cross-checked filtered views for consistency and accuracy.

**4. Objectives**

The main objectives of this project are:

* To provide insights into flight delays and cancellations using visual analytics.
* To help airline operators understand patterns based on flight distance, destination, and time.
* To calculate and visualize metrics such as average departure delay and total cancellations.
* To create a user-friendly and interactive dashboard for easy data exploration.
* To enhance operational efficiency and support data-driven decision-making in the airline industry.

**4. Significance**

This project holds significant value for stakeholders within the aviation sector. With increasing passenger numbers and tightening operational margins, identifying and minimizing delays is crucial. By leveraging Power BI to visualize delay and cancellation trends, decision-makers are empowered with a data-driven understanding of performance issues.

The insights gained from this analysis can help streamline operations, improve on-time performance, and enhance overall customer experience. Additionally, the use of visual analytics makes the information more accessible to non-technical users, promoting a culture of data-informed decisions across all levels of airline and airport operations.

Ultimately, the report serves as a decision support tool to drive efficiency, reduce operational disruptions, and support continuous improvement in air travel services.

**5. Industry Background**

Air travel has become one of the most relied upon modes of transportation. However, with increased traffic comes operational complexity. Delay analysis is essential to optimize scheduling, reduce cost, and enhance the customer experience.

The U.S. Bureau of Transportation Statistics provides detailed records of flights, including delays and cancellations. This data provides a unique opportunity to explore performance trends in real-world scenarios using BI tools like Power BI.

**5. Data Sources**

**a. Data Sources**

* **Dataset Name**: U.S. Domestic Flights Dataset
* **Source Format**: CSV file
* **Imported Using**: Power BI Desktop
* **Fields Included**:
  + Flight Date (FL\_DATE)
  + Origin & Destination
  + Expected and Actual Departure/Arrival Times
  + Cancellation Status
  + Distance (in miles)

**b. Data Transformations**

* Renamed columns for clarity and consistency.
* Removed rows with null or missing values.
* Standardized date and time formats.
* Created calculated columns such as Departure\_Delay.
* Grouped distances into bins for trend analysis.

**c. Data Modelling**

* **Data Model Type**: Single-table (flat model)
* **Calculated Columns and Measures**:
  + Departure\_Delay = DATEDIFF(EXPECT\_DEP\_TIME, ACTUAL\_DEP\_TIME, MINUTE)
  + Average Departure Delay, Total Cancellations, etc.
* **Data Dictionary**:
  + FL\_DATE: Date of the flight
  + ORIGIN, DESTINATION: Cities involved in the flight
  + EXPECT\_DEP\_TIME, ACTUAL\_DEP\_TIME: Scheduled vs. actual departure times
  + CANCELLED: Binary indicator of cancellation (0 or 1)
  + DISTANCE: Flight distance in miles

**6. MECE Breakdown**

**Delay Analysis**

* Average delay by flight distance and city
* Time-based delay trends (daily/monthly)

**Cancellation Analysis**

* Cancellations by destination and time period
* Percentage of cancelled flights vs completed flights

**Distance-Based Analysis**

* Short, Medium, Long distance categories
* Relationship between distance and delays or cancellations

**Time-Based Trends**

* Delay trends by time of day (morning, evening, night)
* Weekday vs weekend performance

**7. Calculations & Formulas**

**a. Custom Calculations**

* Departure\_Delay = DATEDIFF(EXPECT\_DEP\_TIME, ACTUAL\_DEP\_TIME, MINUTE)
* Distance\_Category = SWITCH(TRUE(), DISTANCE < 300, "Short", DISTANCE < 1000, "Medium", "Long")
* Total\_Cancellations = COUNTROWS(FILTER(Flights, CANCELLED = 1))

**b. User Interactivity**

* Slicers: Origin, Destination, Date
* Drill-through: Destination-level detail
* Filters: Distance range, cancellation status

**8. Exploratory Data Analysis**

**Example 1:**

* **Problem**: Which destination has the most cancelled flights?
* **Visualization**: Pie chart of cancellations by destination
* **Insight**: Jackson, Mississippi had the highest share of cancellations.

**Example 2:**

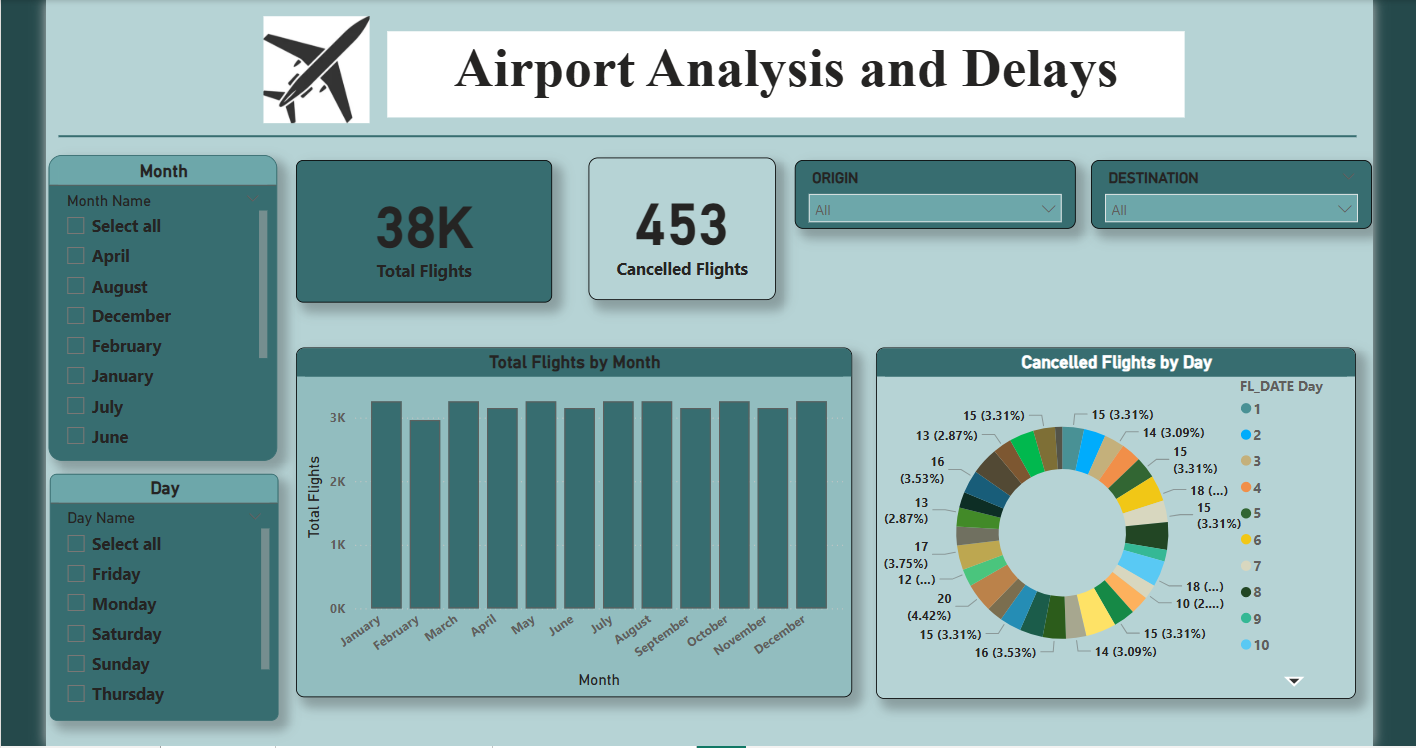
* **Problem**: Do longer flights have more delays?
* **Visualization**: Scatter plot of distance vs average delay
* **Insight**: Shorter flights experienced more delays than longer ones.

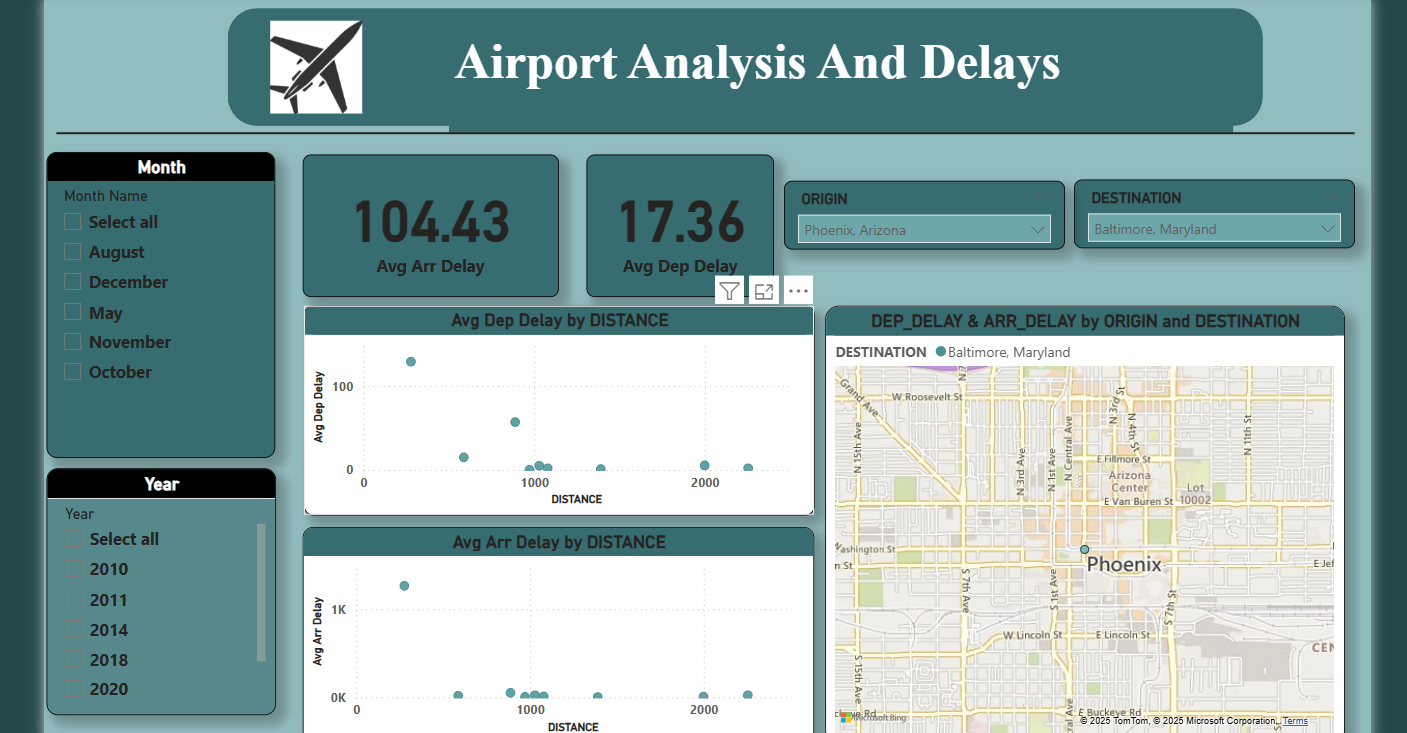
**Example 3:**

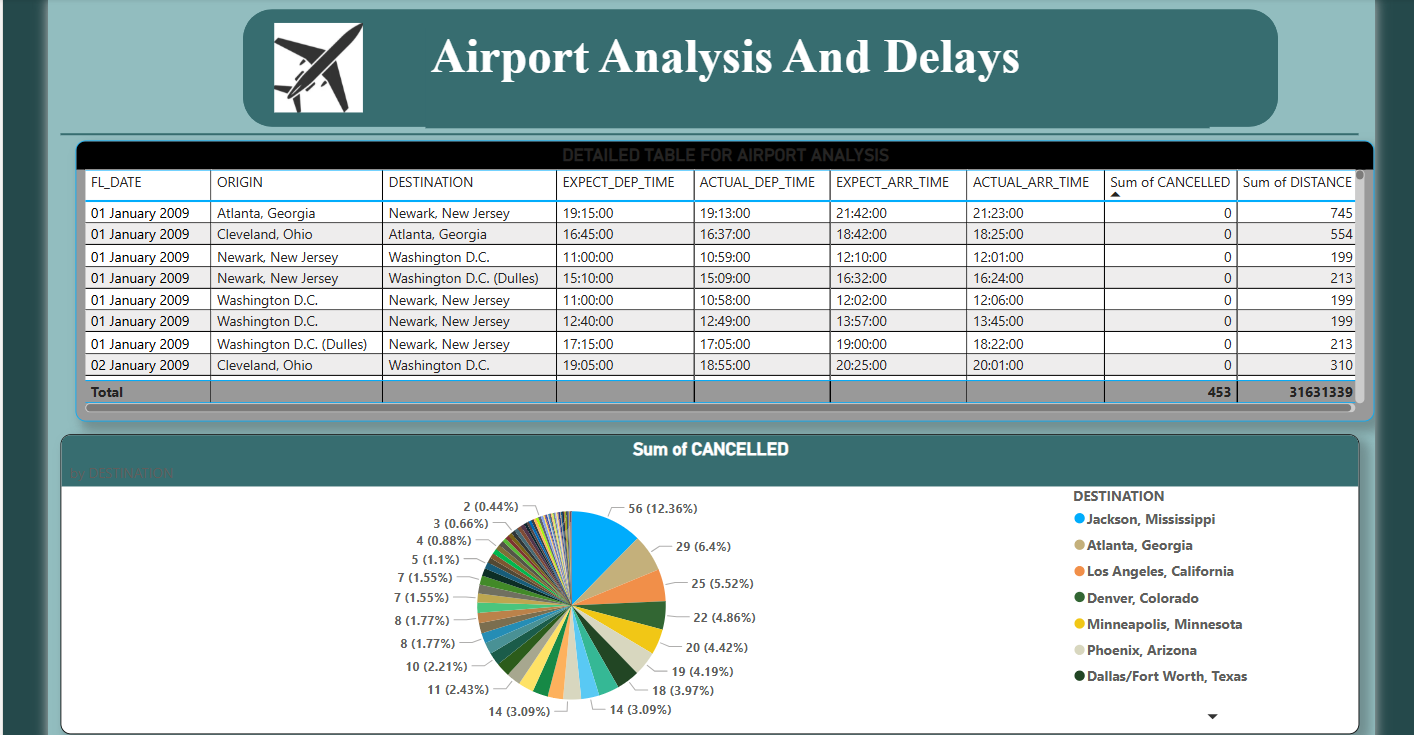
* **Problem**: Are delays worse at certain times of the day?
* **Visualization**: Line chart of average delay by hour
* **Insight**: Evening and night-time flights saw more delays on average.

**9 Power BI Dashboard**

* **Components**:
  + Flight data table
  + KPI cards (total distance, cancellations, delays)
  + Pie chart (cancellations by city)
  + Scatter plot (delay vs distance)
  + Line chart (delay over time)
* **Interactive features**:
  + Date and destination slicers
  + Drill-through for route details







**10. Testing Plan**

**a. Plan**

* Validate all DAX measures manually
* Ensure visual filters and slicers work as expected

**b. Test Cases**

|  |  |
| --- | --- |
| **Test** | **Outcome** |
| Filter on "Houston" | Only Houston flights display |
| Select Cancelled = 1 | Only cancelled flights shown |
| Time filter = January | Shows January data only |

**c. Validation**

* Cross-check delay values
* Confirm summary totals match detail records

**11. Conclusion**

* The "Airport Analysis and Delay" project successfully transforms raw flight data into actionable insights using the analytical and visual capabilities of Power BI. By focusing on delay trends, cancellation patterns, and the relationship between distance and performance, the dashboard provides an accessible and powerful tool for operational analysis.
* The insights gained from this project can support airport and airline management in improving flight scheduling, resource allocation, and customer experience. The use of interactive visualizations ensures that stakeholders can explore the data dynamically, make informed decisions, and address inefficiencies in real time.
* Overall, this project showcases the value of business intelligence in aviation operations and demonstrates how Power BI can serve as a comprehensive platform for data exploration and strategic planning in the transportation sector.

**12. References :**

1. U.S. Department of Transportation – Bureau of Transportation Statistics (BTS)  
   https://www.transtats.bts.gov
2. Microsoft Power BI Documentation  
   <https://learn.microsoft.com/en-us/power-bi/>
3. DAX Guide by SQLBI  
   <https://dax.guide/>
4. Power BI Community and Forums  
   https://community.powerbi.com/
5. Kaggle – U.S. Domestic Flights Dataset (if applicable) <https://www.kaggle.com>*.*