Airline Data Management and Analysis Using Power BI

1. Introduction

Project Title

Airline Data Management and Analysis Using Power BI

Problem Statement

The airline industry operates with numerous complexities, requiring effective data management to optimize flight schedules, passenger handling, and ticket booking systems. This project leverages **Power BI** to analyze airline operations, identify inefficiencies, and improve decision-making for enhanced customer satisfaction.

Objective

To transform raw airline data into actionable insights by:

- Cleaning and modeling datasets.
- Performing calculations using **DAX**.
- Creating interactive dashboards for real-time analysis.

2. Datasets Used

1. **Flight_Information**:
    - Columns: `FlightID`, `FlightNumber`, `Airline`, `Destination`, `Status`.
2. **Passenger_Information**:

- Columns: `PassengerID`, `FlightID`, `SeatNumber`.

3. **Ticket_Information**:
   - Columns: `TicketID`, `FlightID`, `BookingStatus`.

---

3. Methodology & Implementation

3.1 Data Preparation and Cleaning

- **Steps Taken**:
  - Removed null columns and duplicate rows in Power Query.
  - Validated data quality using **Column Distribution** and **Column Profile**.
  - Formatted columns (e.g., `FlightID` as Whole Number, `Status` as Text).
- **Deliverable**: *(Replace with actual screenshot)*

*Suggestion: Use **"Replace Values"** for consistent formatting (e.g., "On-Time" → "On Time").*

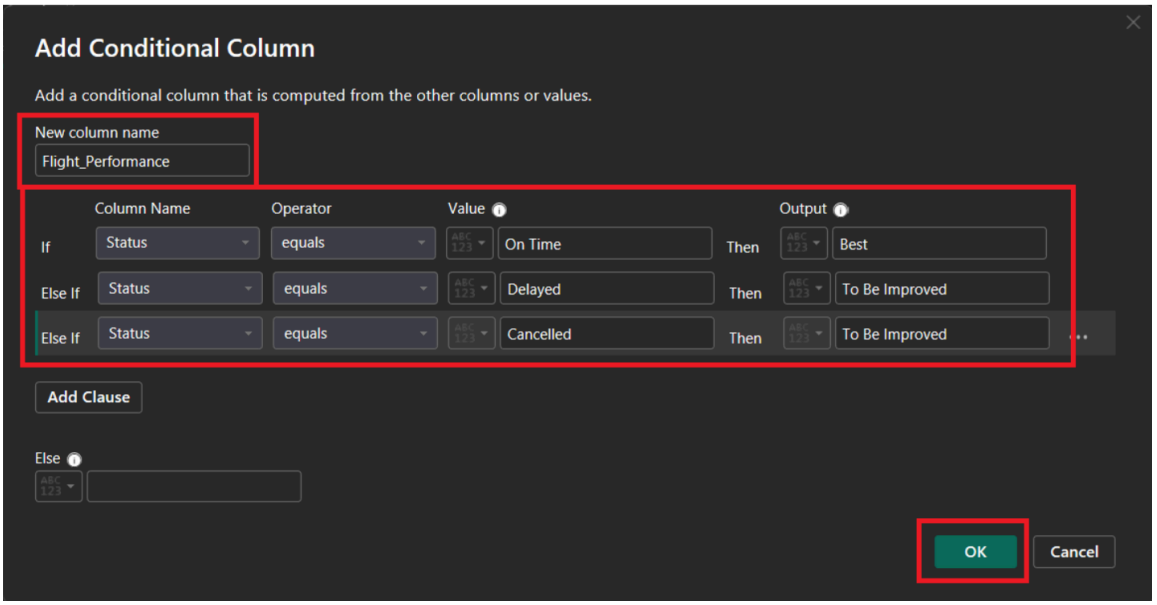
3.2 Data Modeling

- **Relationships**:
  - Established **one-to-many** relationships between `Flight_Information` (primary) and `Passenger_Information` / `Ticket_Information` using `FlightID`.
  - Cross-filter direction: **Single**.
- **Deliverable**: *(Show relationships diagram)*

*Suggestion: Use **"Manage Relationships"** to verify referential integrity.*

3.3 Enhanced Data Insights

- **Conditional Column**:
  - Created `Flight_Performance` to classify flights:
    - **"Best"** (Status = "On Time").
    - **"To Be Improved"** (Status = "Delayed" or "Cancelled").
- **Column from Examples**:
  - Extracted flight numbers (e.g., "FL1102" → "1102") into `ExtractedFlightNumber`.
- **Deliverable**: *(Show DAX logic)*



*Suggestion: Add a **tooltip** to explain performance categories.*

3.4 Calculations Using DAX

- **Key Measures**:

```
Total_Passengers = COUNT(Passenger_Information[PassengerID])
Total_Tickets_Booked = COUNT(Ticket_Information[TicketID])
Best_Flights = FILTER(Flight_Information, Flight_Information[Flight_Performance] = "Best")
```

- **Deliverable**: *(Show card visuals with totals)*

*Suggestion: Use **ALLEXCEPT()** to preserve filters (e.g., for airline-specific analysis).*

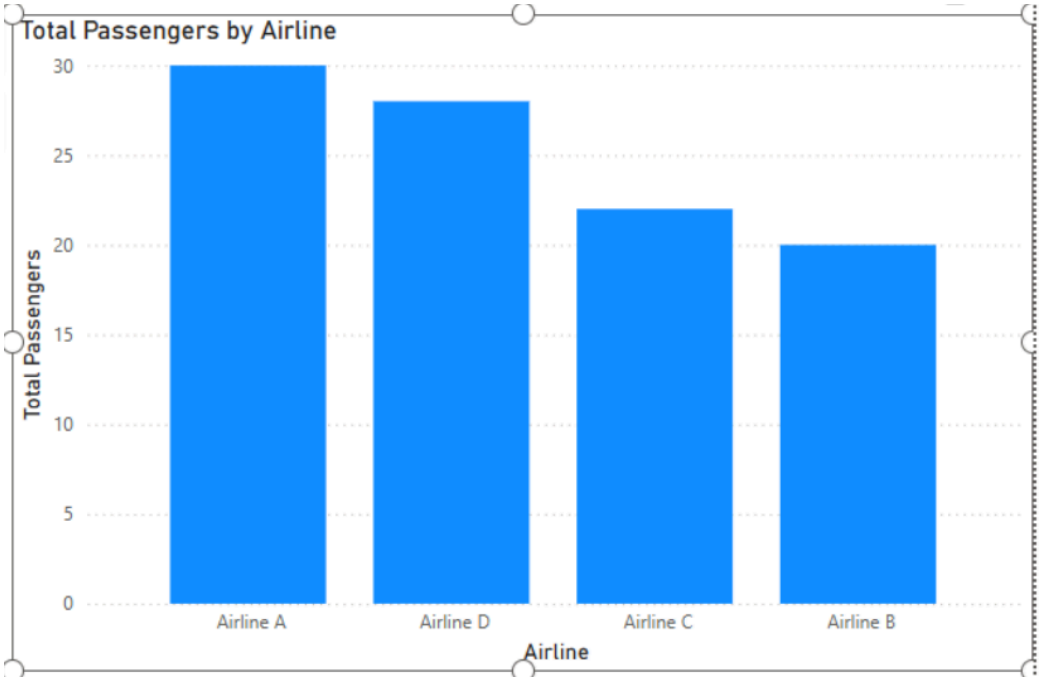
---

4. Visualization & Dashboard

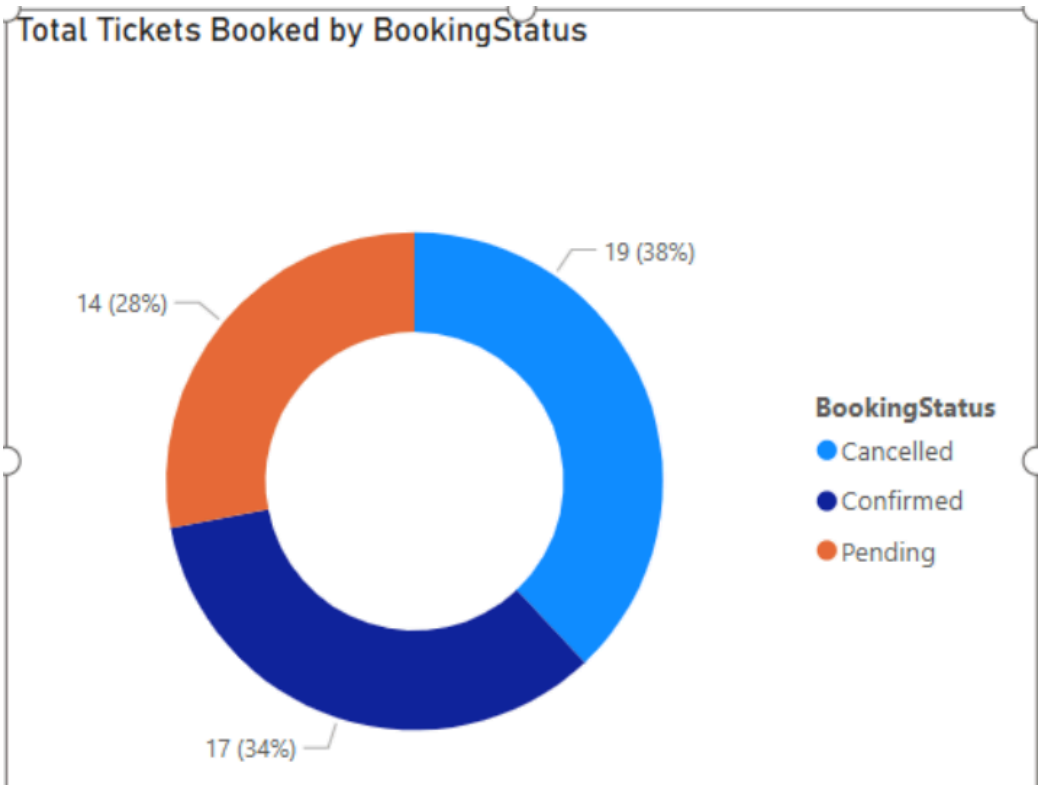
4.1 Recommended Charts

1. **Passenger Count by Airline**
    - **Chart Type**: Stacked Bar Chart
    - **Fields**:
        - X-axis: Airline

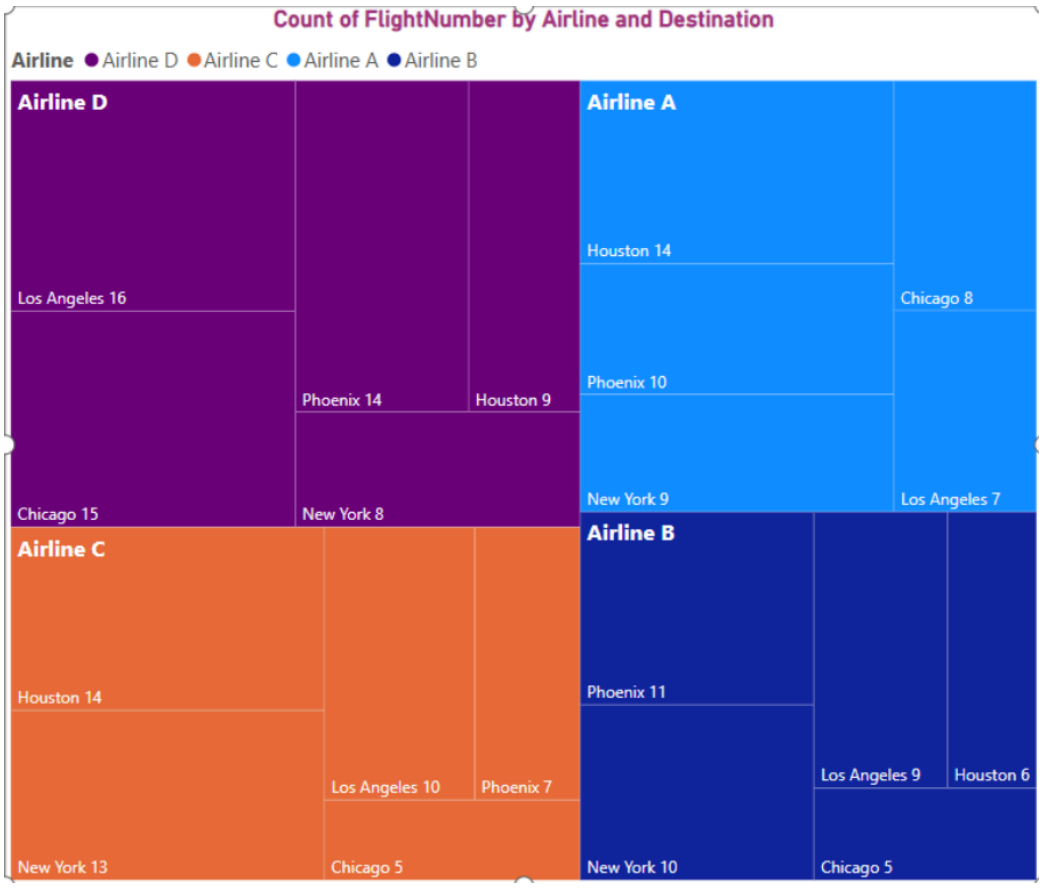
- Y-axis: `Total_Passengers`



2. **Ticket Booking Statuses**
   - **Chart Type**: Donut Chart
   - **Fields**:
     - Legend: `BookingStatus`
     - Values: `Total_Tickets_Booked`



3. **Flights by Airline & Destination**
    - **Chart Type**: Heatmap
    - **Fields**:
        - Rows: `Airline`
        - Columns: `Destination`
        - Values: `FlightNumber` (Count)



*Suggestion: Use **slicers** for `Airline` and `Destination` to enhance interactivity.*

4.2 Interactive Features

- **Slicers**: Added for `Airline` and `Destination` to filter visuals dynamically.
- **Tooltips**: Configured to show `Passenger Count` and `BookingStatus` details on hover.
- **Navigation Buttons**: Created for airline-specific pages (e.g., "Airline A Overview").

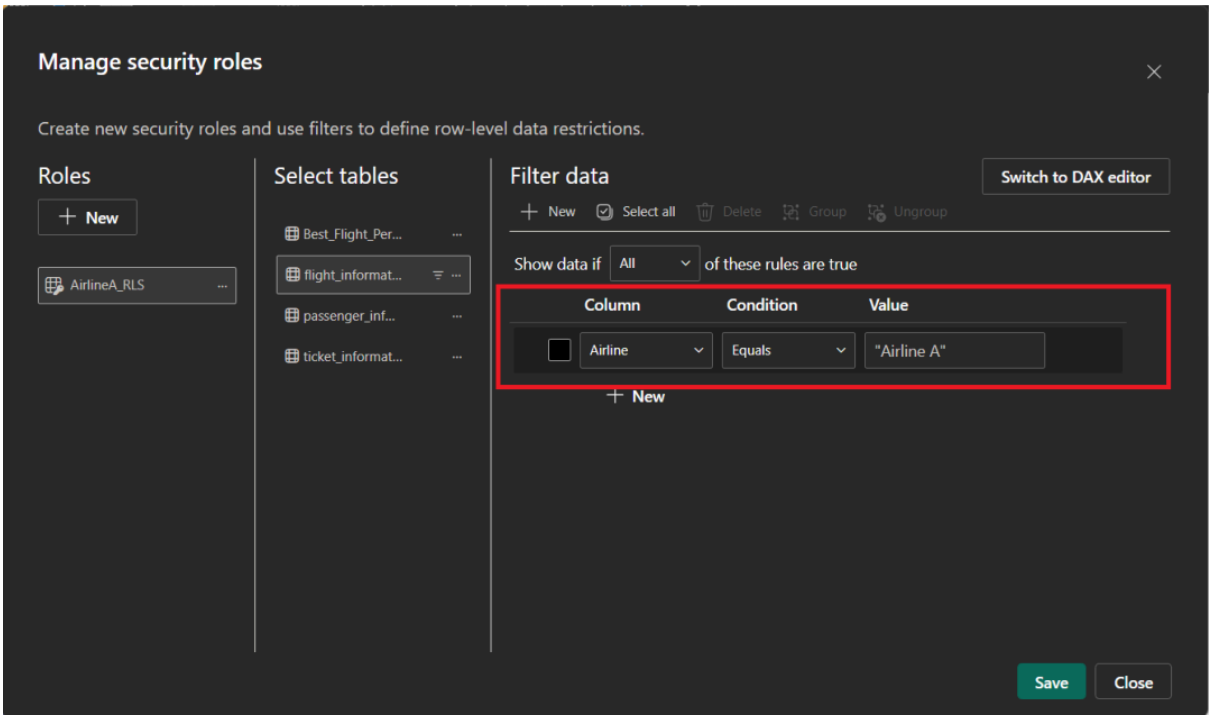
Final Dashboard:

(Include a full-view screenshot)

5. Power BI Service Deployment

- **Published** to workspace **"AirDMA"**.
- **Row-Level Security (RLS)**:
  - Configured for `Airline A` to restrict data access.
- **Scheduled Refresh**: Set to run daily at **5 PM**.

Deliverable:



(Show security settings)

---

6. Conclusion & Recommendations

- **Key Insights**:
  - **Airline D** had the highest passenger count (28).
  - **Los Angeles** was the most popular destination (42 passengers).
  - **22% of tickets** were in "Pending" status, indicating potential revenue leakage.
- **Improvements**:
  - *Investigate delays for "To Be Improved" flights.*
  - *Optimize ticket confirmation processes to reduce "Pending" statuses.*

Appendices:

- Power Query M Code snippets.
- DAX formulas used.
- Dataset sample (Excel).
