* **Project Title:** -

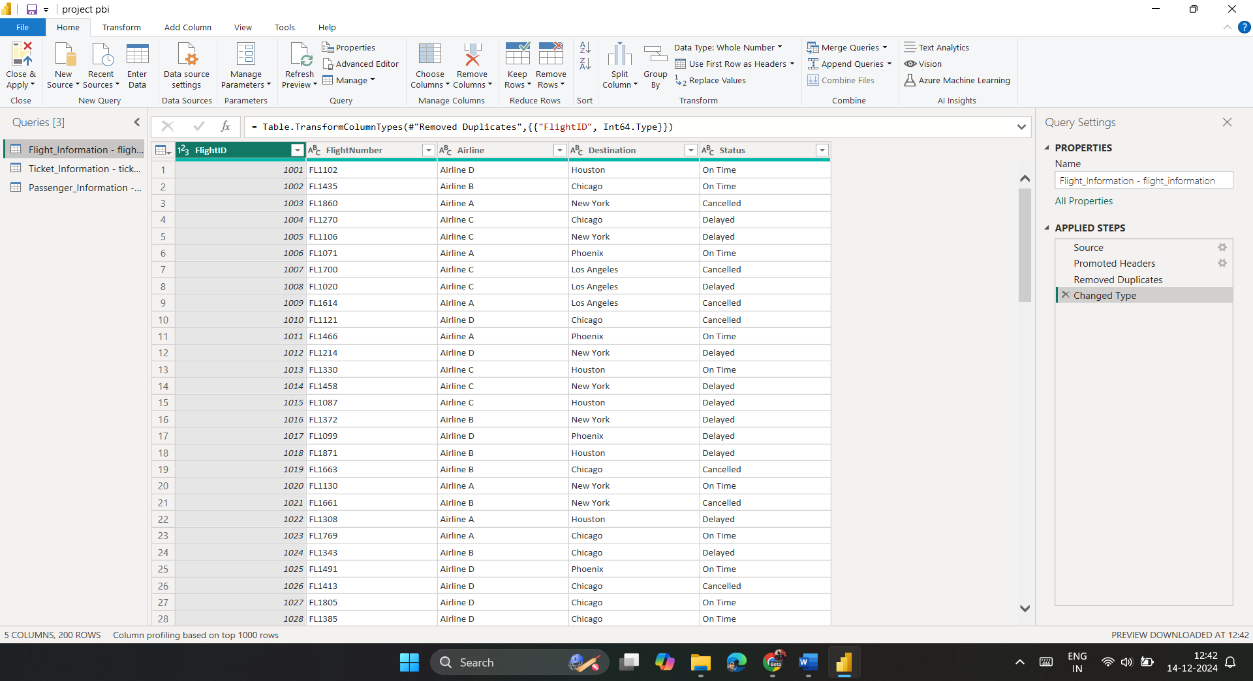
Airline Data Management and Analysis Using Power BI

* **Problem Statement: -**

The airline industry operates with numerous complexities, requiring effective data management and insights into flight schedules, passenger details, and ticketing systems. This project aims to analyze airline operations for improving efficiency and customer satisfaction.

* **Datasets Used: -**
* **Flight\_Information:** FlightID, FlightNumber, Airline, Destination, Status
* **Passenger\_Information:** PassengerID, FlightID, SeatNumber
* **Ticket\_Information:** TicketID, FlightID, BookingStatus

1. **Data Preparation and Cleaning**

* **Extract and transform data in Power Query.**
* **Clean data: remove duplicates, handle missing values, & format columns.**
*  **Deliverables: Screenshot of Power Query Editor showing cleaned data.**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

In the Data Preparation and Cleaning stage, the data was loaded and transformed using Power Query Editor:

* **Loading Data to Power Query**:
  + First, the data was loaded into Power BI using the Get Data option.
  + Once the data was loaded, the Power Query Editor was used to transform and clean the dataset.
* **Removing Duplicates**:
  + Duplicate entries were identified and removed to ensure the dataset had only unique records.
  + missing values were handled, and columns like FlightID, PassengerID, and TicketID
* **Handling Missing Values**:
  + Missing or null values were handled using various techniques
  + Remove the blank rows this help remove the null cells and filling the required data like n/a without keeping it empty

1. **Data Modelling: -**

* A screenshot of a computer

  Description automatically generated**Create relationships between datasets (FlightID as the key).**
* **Understand cardinality and configure** A screenshot of a computer

  Description automatically generated**the model appropriately.**
* **Deliverables: Screenshot of the data model with relationships.**

In Data Modeling, relationships were created between the different datasets to enable seamless analysis and reporting.

* **Building Relationships**:

The Flight\_Information, Passenger\_Information, and Ticket\_Information tables were connected using FlightID as the primary key.

* **One-to-Many Relationships**:

One-to-Many relationships were set to ensure that each flight can have multiple passengers and tickets, but each passenger and ticket can be associated with only one flight.

* **Cardinality:**

The Cardinality was set to control how data is filtered between tables. This step ensures that the filters work correctly and that no data is missed or misrepresented.

1. **Enhanced Data Insights: -**

* **Add a conditional column to classify flights as "Best" or "To Be Improved" based on status.**
* **Use "Column from Examples" to extract the flight number from FlightNumber.**
* **Deliverables: Screenshot of the transformed data**

Enhanced Data Insights were added by creating conditional columns and calculated fields to add business context to the data.

* **Creating Conditional Columns:**

Flights were classified as either "Best" or "To Be Improved" based on their status, using the Conditional Column feature in Power Query.

* **Column With Examples:**

The FlightNumber was extracted from the Flight\_Information table using the Column from Examples feature. This added a new field for analysis and visualization.

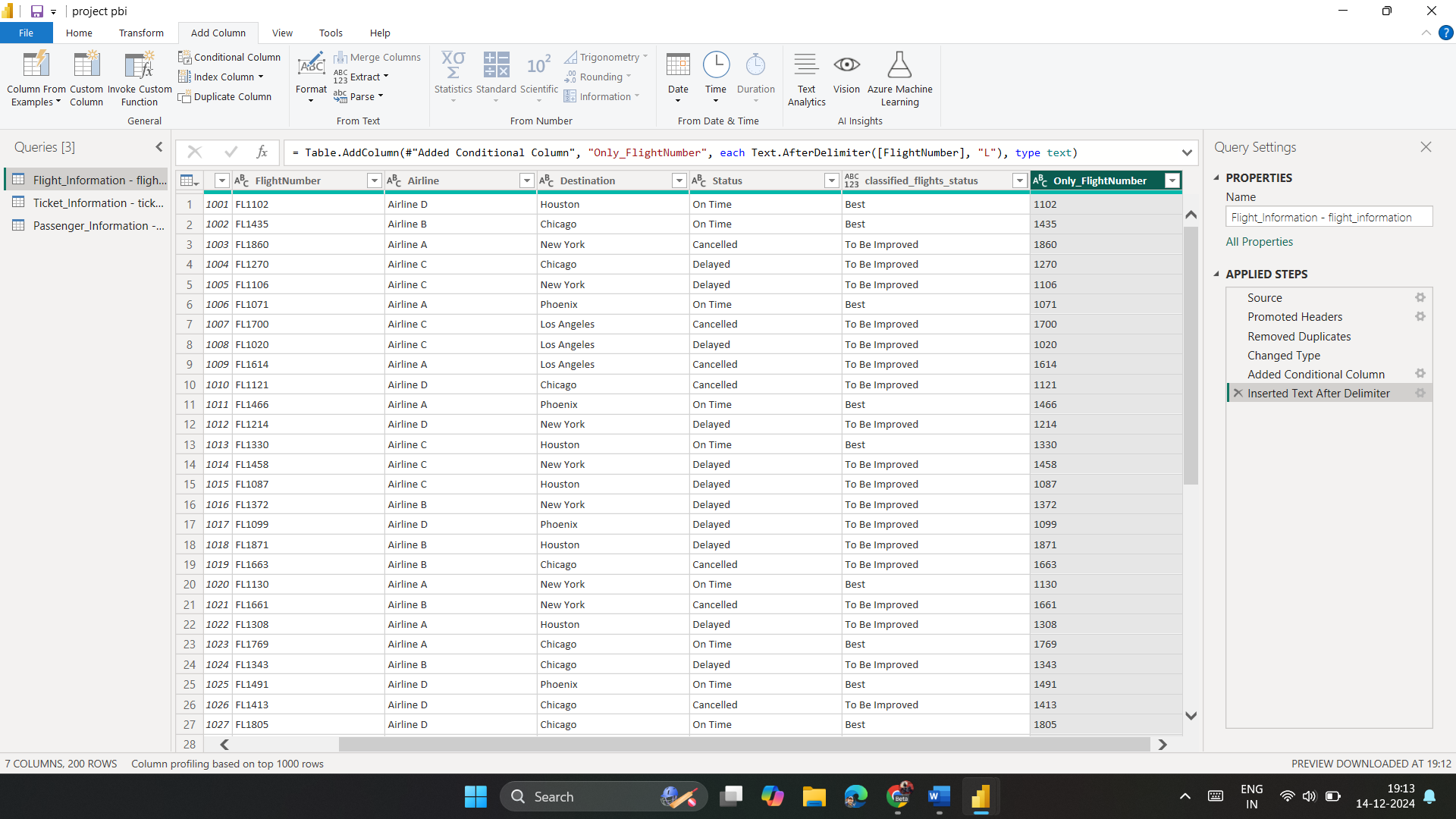
A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generatedA screenshot of a computer

Description automatically generated



1. **Calculations Using DAX: -**

* **Calculate:**
* **Total passengers for a specific flight.**

In DAX Calculations, various metrics were calculated to provide deeper insights into the data.

* Total Passengers:

A measure was created using the DAX function to count the number of passengers per flight.

* Dax Formula:

**TotalPassengers = COUNTROWS(Passenger\_Information)**

* Total Tickets:

Another measure was created to count the total number of tickets booked.

* Dax Formula:

**TotalTickets = COUNTROWS(Ticket\_Information)**

* Best Flights Table:

A new table was filtered to display only the flights classified as "Best".

* Dax Formula:

**BestFlights = FILTER(Flight\_Information, Flight\_Information[Status] = "Best")**

A screenshot of a computer

Description automatically generated

* **Total tickets booked.**

A screenshot of a computer

Description automatically generated

* **Filtered table showing "Best" flights only.**
* **Deliverables: Screenshot of DAX calculations and results.**

A screenshot of a computer

Description automatically generated

1. **Visualization and Interactive Features**

* **Create visuals for:**
* **Passenger count by airline.**
* Donut Chart: Passenger Counts Grouped by Airline
* Data Field:Category: Airlines (e.g., Airline A, Airline B).
* Values: Total Passenger Count (calculated using DAX measure: TotalPassengers).
* Visualization Details:
  + Chart Type: Donut Chart.
  + Data Colors: Distinct colors for each airline for better clarity.
  + A computer screen shot of a pie chart

    Description automatically generatedPercentage Display: Enabled to show each airline's share of total passengers.
* **Ticket booking statuses.**
* Bar Chart: Ticket Booking Statuses
* X-Axis: Ticket Status (Booked, Cancelled, Pending).
* Y-Axis: Count of Tickets (using DAX measure: TotalTickets).
* Visualization Details:
  + Chart Type: Clustered Bar Chart.
  + Data Labels: Enabled for clear visibility of ticket counts.
  + Colors: Each bar represents a ticket status with a distinct color for better differentiation.

A screenshot of a computer

Description automatically generated

* **Flights by airline and destination.**

A screenshot of a computer

Description automatically generated

* Map Visual: Flights by Destination and Airline
* Location Field: Destination City (e.g., New York, London, Tokyo).
* Legend: Airlines (e.g., Airline A, Airline B).
* Visualization Details:
  + Chart Type: Map Visual.
  + Size: Proportional to the number of flights to each destination.
  + Tooltips: Display total flights and airline name on hover for detailed insights.
* **Add interactive features for:**
* **Destination and Airline.**

A screenshot of a computer

Description automatically generated

* Slicers
* Destination Slicers:
  + - Field: Destination City.
    - Selection: Multi-select enabled for filtering multiple cities simultaneously.
* Airline Slicers:
  + - Field: Airline Name.
    - Selection: Dropdown style for compact display.
* **Quick views.**

A screenshot of a computer

Description automatically generated

* Predefined Bookmarks:
* Created bookmarks for common insights, such as:
  + Top Destinations: Displays cities with the highest passenger counts.
  + Airline Performance: Highlights metrics for all airlines.
* Buttons for Navigation:
* Buttons linked to bookmarks were added for a user-friendly experience.
* "View Top Destinations" button applies filters and navigates to the Top Destinations view.
* "Reset View" button clears all filters and restores the default dashboard layout.
* **Airline-specific pages.**
* **Deliverables: Screenshots of all visuals and interactive features.**

A screenshot of a computer

Description automatically generated

* Drill-Through Pages:
* Separate pages were created for each airline to provide detailed, airline-specific insights.
* Fields Used: Airline Name as the drill-through field.
* Details Displayed on Drill-Through Pages:
* Passenger counts for that airline,
* Top destinations served by the airline,
* Ticket statuses and flight performance metrics.

1. **Final Dashboard and Power BI Service**

* **Design a comprehensive dashboard with key visuals and insights.**

A computer screen shot of a computer screen

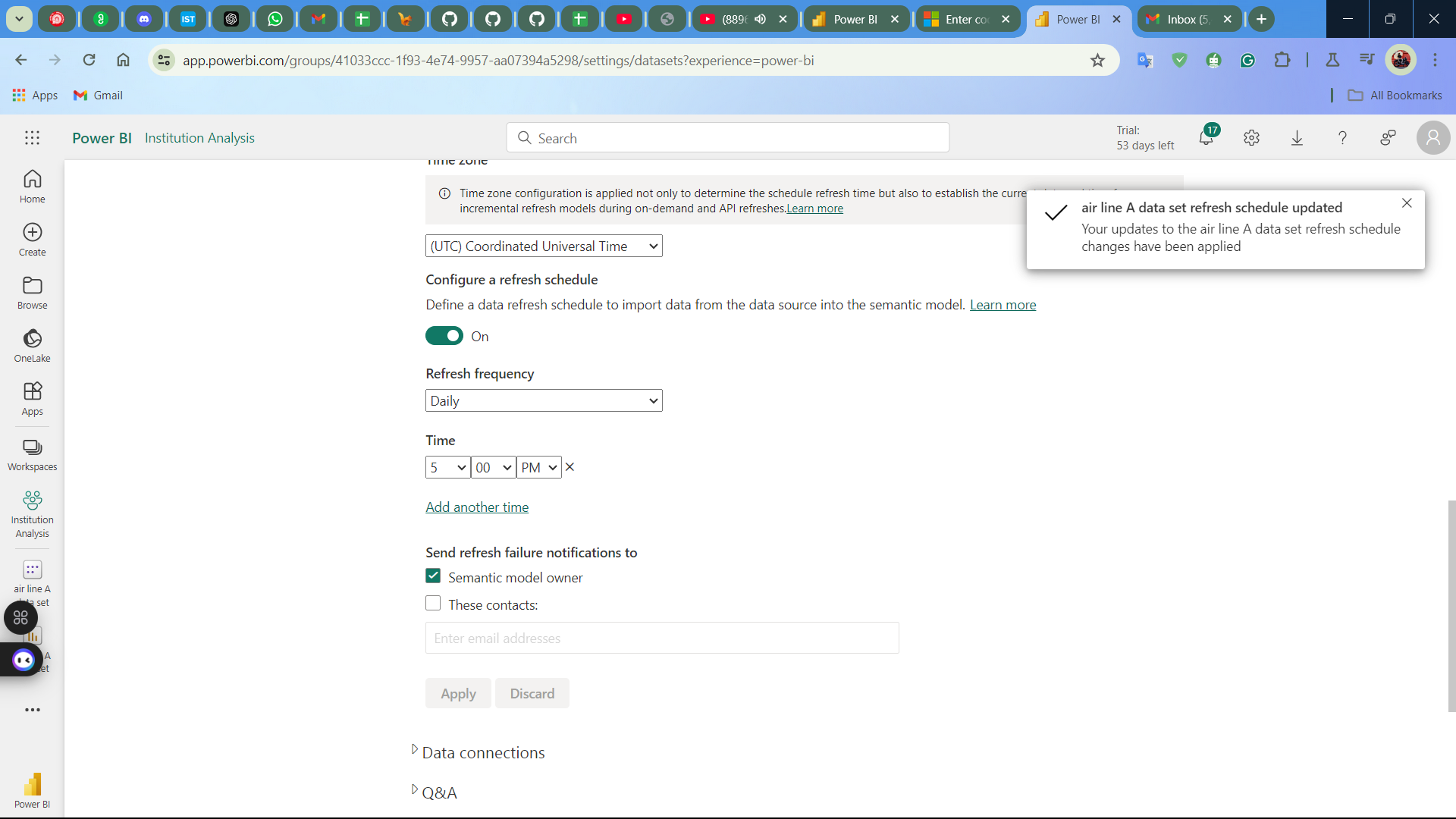
Description automatically generated

* Visuals Added:
* **Passenger Counts by Airline** (Donut Chart).
* **Ticket Statuses** (Clustered Bar Chart).
* **Flights by Destination and Airline** (Map Visual).
* Layout Design: Arranged visuals in a grid pattern for clarity:
  + Top Section: Slicers for Destination and Airline.
  + Middle Section: Donut Chart and Bar Chart side by side for comparison.
  + Bottom Section: Map Visual for geographic insights.
* Formatting:
* Used consistent fonts, colors, and borders for professional appearance.
* Added a title: **"Airline Performance Dashboard"** for clear identification.
* By clicking in power bi service where we can create the final dashboard and publish as well.
* **Configure Row-Level Security (RLS) for Airline A data, assign it to a user.**
  + Define Roles:
    - Created a role called "data analyst".
    - Added a filter to the Airline Name column: [Airline] = "Airline A".
  + Assign Users:
    - Assigned specific email IDs of internshala staff to the role in Power BI Service.
  + Testing RLS:
    - Verified RLS functionality using the View as Role feature in Power BI Desktop.
* Users from Airline A can only see their airline's data, ensuring security and privacy.

A screenshot of a computer

Description automatically generated

* Set up a schedule refresh at 5 PM daily.



* + Upload to Power BI Service: Published the dashboard to Power BI Service.
  + Configure Scheduled Refresh:
    - Set up a refresh schedule from the settings in Power BI Service:
      * Frequency: Daily.
      * Time: 5 PM.
    - Ensured proper data source credentials were provided for seamless refresh.
  + Test the Refresh:
    - Triggered a manual refresh to confirm the setup works as expected.
* Deliverables: Screenshot of the published dashboard and RLS configuration.
* **Note:- The video file size is large so I used this video link:** <https://drive.google.com/file/d/1l1o9N2Ok7XwnTPB4wHNCvCjUQhzg-qOh/view?usp=sharing>

**UMA MAHESWAR NETHI**

**DATA SCIENCE LEARNER**