

PROJECT POWER BI

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

- 1. Flight Information:** Includes FlightID, FlightNumber, Airline, Destination, and Status.
- 2. Passenger Information:** Includes PassengerID, FlightID, and SeatNumber.
- 3. Ticket Information:** Includes TicketID, FlightID, and BookingStatus.

1. Data Preparation and Cleaning (10 Marks)

- Extract and transform data in Power Query.
- Clean data: remove duplicates, handle missing values, and format columns.

1. Data Preparation and Cleaning – 10 Marks

Steps:

- Load all three datasets into Power BI using Power Query.
- Remove duplicate rows (e.g., duplicate flight IDs or ticket records).
- Ensure consistent data types (e.g., Date Time for schedule columns, Text for BookingStatus).
- Rename columns for clarity and consistency (e.g., “FlightID” instead of “flight_id”).
- Format columns (e.g., convert status to proper case, trim unnecessary spaces).
- Promote headers also
- Format titles by using _ in between

Home Transform Add Column View Tools Help

CLOSE & APPLY CLOSE NEW QUERY

Queries [4]

Flight_information

Ticket_information

Passenger_information

Measure Table

DATA SOURCE SETTINGS DATA SOURCES MANAGE PARAMETERS REFRESH PREVIEW PROPERTIES ADVANCED EDITOR

CHOOSE COLUMNS REMOVE COLUMNS KEEP ROWS REMOVE ROWS REDUCE ROWS SORT GROUP BY TRANSFORM

MERGE QUERIES APPEND QUERIES COMBINE FILES COMBINE AI INSIGHTS

Text Analytics Vision Azure Machine Learning

Flight_ID Airline Destination Status FlightNumber Text After Delimiter Status_ty

	Flight_ID	Airline	Destination	Status	FlightNumber	Text After Delimiter	Status_ty
1	1001	Airline D	Houston	On Time	FL1102	1102	Best
2	1002	Airline B	Chicago	On Time	FL1435	1435	Best
3	1003	Airline A	New York	Cancelled	FL1860	1860	Needs Attn
4	1004	Airline C	Chicago	Delayed	FL1270	1270	To be impr
5	1005	Airline C	New York	Delayed	FL1106	1106	To be impr
6	1006	Airline A	Phoenix	On Time	FL1071	1071	Best
7	1007	Airline C	Los Angeles	Cancelled	FL1700	1700	Needs Attn
8	1008	Airline C	Los Angeles	Delayed	FL1020	1020	To be impr
9	1009	Airline A	Los Angeles	Cancelled	FL1614	1614	Needs Attn
10	1010	Airline D	Chicago	Cancelled	FL1121	1121	Needs Attn
11	1011	Airline A	Phoenix	On Time	FL1466	1466	Best
12	1012	Airline D	New York	Delayed	FL1214	1214	To be impr
13	1013	Airline C	Houston	On Time	FL1330	1330	Best
14	1014	Airline C	New York	Delayed	FL1458	1458	To be impr
15	1015	Airline C	Houston	Delayed	FL1087	1087	To be impr
16	1016	Airline B	New York	Delayed	FL1372	1372	To be impr
17	1017	Airline D	Phoenix	Delayed	FL1099	1099	To be impr
18	1018	Airline B	Houston	Delayed	FL1871	1871	To be impr
19	1019	Airline B	Chicago	Cancelled	FL1663	1663	Needs Attn
20	1020	Airline A	New York	On Time	FL1130	1130	Best
21	1021	Airline B	New York	Cancelled	FL1661	1661	Needs Attn
22	1022	Airline A	Houston	Delayed	FL1308	1308	To be impr
23	1023	Airline A	Chicago	On Time	FL1769	1769	Best
24	1024	Airline B	Chicago	Delayed	FL1343	1343	To be impr
25	1025	Airline D	Phoenix	On Time	FL1491	1491	Best
26	1026	Airline D	Chicago	Cancelled	FL1419	1419	Needs Attn

7 COLUMNS, 200 ROWS Column profiling based on top 1000 rows

PREVIEW DOWNLOADED ON 14 JUNE 2023

Query Settings

PROPERTIES

Name: Flight_information
All Properties

APPLIED STEPS

Source Navigation Promoted Headers Changed Type Removed Other Columns Filtered Rows Removed Duplicates Added Conditional Column Inserted Text After Delimiter Reordered Columns Renamed Columns

The screenshot shows the Power BI desktop interface with the following details:

- Home ribbon:** Home, Transform, Add Column, View, Tools, Help.
- Left sidebar:** Close & Apply, New Source, Recent Sources, Enter Data, Data source settings, Manage Parameters, Refresh Preview, Advanced Editor, Properties, Choose Columns, Remove Columns, Keep Rows, Remove Rows, Split Column, Group By, Replace Values, Data Type: Whole Number, Merge Queries, Append Queries, Combine Files, Combine, Text Analytics, Vision, Azure Machine Learning, AI Insights.
- Queries [4]:** Flight_information, Ticket_information (selected), Passenger_information, Measure Table.
- Table View:** A table titled "Ticket_information" with columns: Ticket_ID, Flight_ID, Booking_Status. The data shows various flight bookings with their status (Pending, Confirmed, Cancelled).
- Properties pane:** Shows the selected item is "Ticket_information".
 - PROPERTIES:** Name: Ticket_information, All Properties.
 - APPLIED STEPS:** Shows the steps taken:
 - Source
 - Navigation
 - Promoted Headers
 - Changed Type
 - Removed Other Columns
 - Removed Duplicates
 - Renamed Columns (highlighted)

The screenshot shows the Microsoft Power BI Data Editor interface. The top navigation bar includes Home, Transform, Add Column, View, Tools, and Help. The ribbon has sections for Close & Apply, New Source, Recent Sources, Enter Data, Data source settings, Manage Parameters, Refresh Preview, Properties, Advanced Editor, Query, Choose Columns, Remove Columns, Keep Rows, Remove Rows, Sort, Split Column, Group By, Data Type: Whole Number, Merge Queries, Append Queries, Combine Files, and Combine. On the right, there are sections for Text Analytics, Vision, Azure Machine Learning, and AI Insights. The main area shows a query titled 'Passenger_information' with the formula: `= Table.RenameColumns("Removed Duplicates", {"PassengerID", "Passenger_ID"}, {"FlightID", "Flight_ID"}, {"SeatNumber", "Seat_Number"})`. Below it is a preview of a table with three columns: Passenger_ID, Flight_ID, and Seat_Number. The table has 26 rows of data. The bottom right corner of the preview area says 'PREVIEW DOWNLOADED AT 19:39'. The 'APPLIED STEPS' pane on the right lists the steps taken: Source, Navigation, Promoted Headers, Changed Type, Removed Other Columns, Removed Duplicates, and Renamed Columns.

2. Data Modeling (10 Marks)

- Create relationships between datasets (FlightID as the key).
- Understand cardinality and configure the model appropriately

STEP 1: Establish Relationships:

- Use FlightID as the primary key to create the following relationships:
 - Flight Information[FlightID] → Passenger Information[FlightID]
 - Flight Information[FlightID] → Ticket Information[FlightID]

STEP 2: Set Cardinality:

- Understand and configure relationships as:
 - One-to-Many (1:*):
 - One flight can have many passengers (1:*)
 - One flight can have many tickets (1:*)
- Ensure "Cross filter direction" is set correctly (usually “Single”).

The screenshot shows the Power BI Home screen with the following elements:

- Top Bar:** Home, Help, Cut, Copy, Get data, Excel, OneLake, SQL Server, Enter Data, Data, Transform Refresh data, Manage relationships, New measure column, New table, Calculation group, New parameter, Manage roles, View as Security, Q&A, Language setup, Sensitivity, Publish.
- Table Preview Area:**
 - Passenger_information:** Flight_ID, Passenger_ID, Seat_Number.
 - Flight_information:** Airline, Destination, Flight_ID, FlightNumber, Status, Status_type, Text After Delimiter.
 - Ticket_information:** Booking_Status, Flight_ID, Ticket_ID.
 - Best Flights:** Airline, Destination, Flight_ID, Status, Column1.
- Bottom Navigation:** All tables, +.

The screenshot shows the Power BI Manage relationships dialog with the following details:

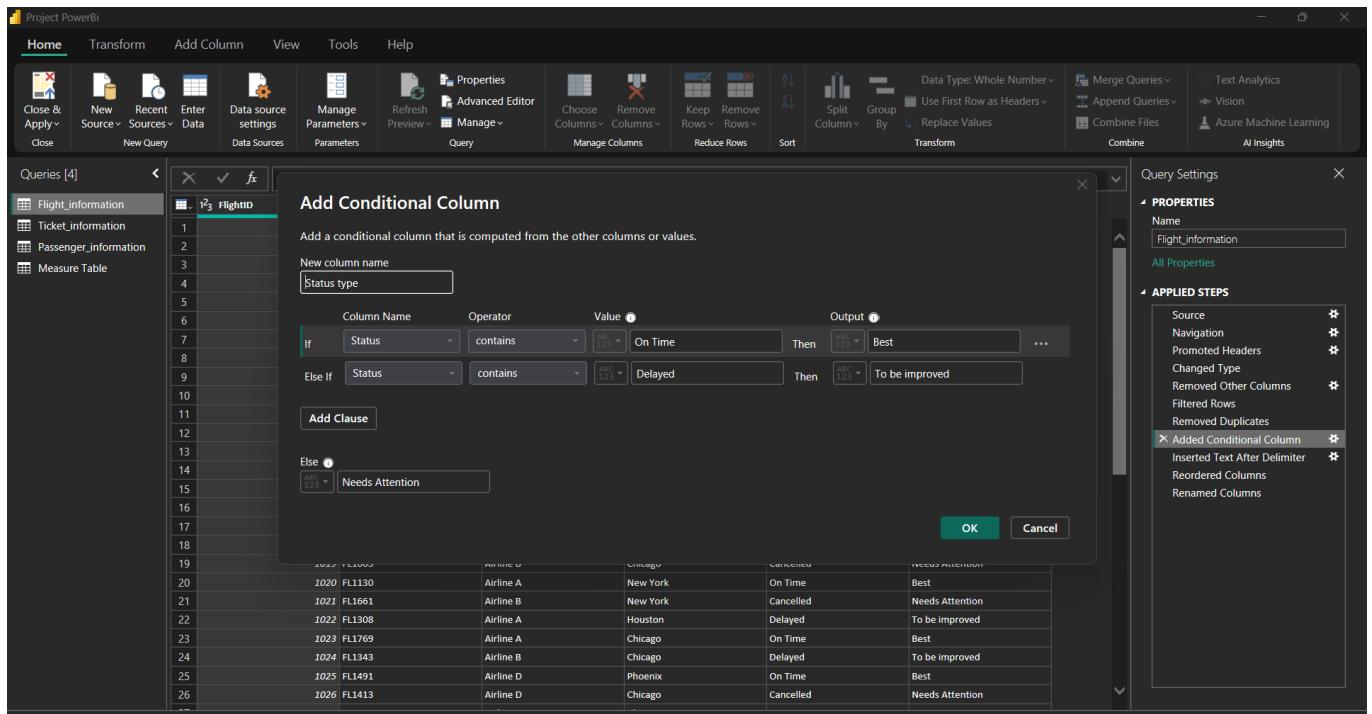
- From: table (column)**: Passenger_information (Flight_ID), Ticket_information (Flight_ID).
- Relationship**: *—(4)—1.
- To: table (column)**: Flight_information (Flight_ID).
- Status**: Active.
- Buttons:** + New relationship, Autodetect, Edit, Delete, Filter.
- Bottom Navigation:** All tables, +.

3. Enhanced Data Insights (10 Marks) ● 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.

Steps:

- 1. Create Conditional Column for Flight Classification:**

- **Use Power Query's "Add Column > Conditional Column" option.**
- **Logic example:**
 - **If Status = "On Time", then "Best"**
 - **If delayed, then "to be improved"**
 - **Else needs attention**
- **Rename column as status type**



Use Column From Examples:

- **In Power Query, select "Add Column > Column from Examples"**
- **Extract a simplified Flight Code (e.g., just the numeric part) from FlightNumber.**

Example: From FL1102 extract 1102.

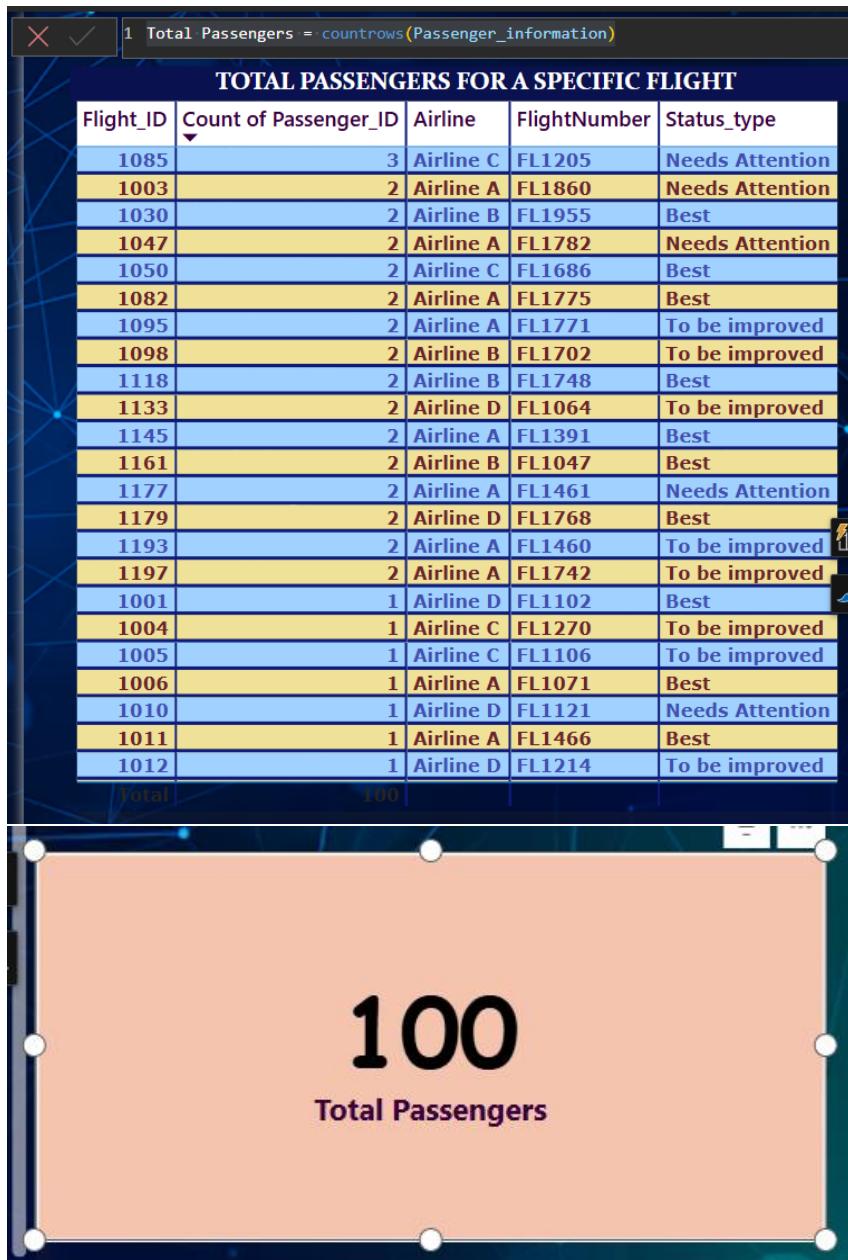
- **This helps in grouping or filtering flights based on numeric codes.**
- **Rename the column as FlightCode**

The screenshot shows the Power BI desktop interface with the 'Add Column' tab selected. The main area displays a table titled 'Flight_information' with 200 rows and 6 columns. The columns are: Destination, Status, FlightNumber, Text After Delimiter, Status_type, and FLIGHT_CODE. A context menu is open over the 'Text After Delimiter' column, with the option 'Text After Delimiter - Copy' highlighted. The 'APPLIED STEPS' pane on the right lists several steps taken to process the data, including 'Renamed Columns', 'Added Conditional Column', and 'Duplicated Column'.

4. Calculations Using DAX (10 Marks)

● Calculate:

- Total passengers for a specific flight.
- Total tickets booked.
- Filtered table showing "Best" flights only.



Steps4(a)

1. In the report view ,select enter data and create measure table.
2. Create measure
 $\text{Total Passengers} = \text{countrows}(\text{Passenger_information})$
3. Choose card visual for total no. of passengers
4. Create table for number of costumers for each flight.

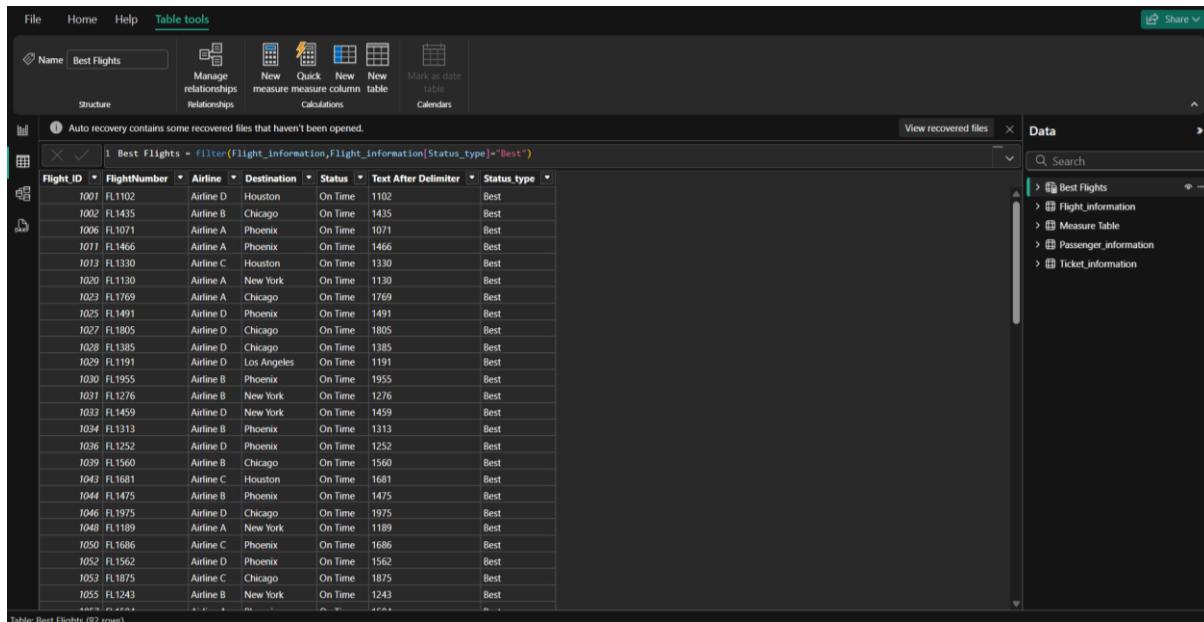
Step4(b)

1.Create measure In measure table

2.Total tickets =countrows(ticket_information)

Steps4(c)

1. Go to table view ,click on new table
2. In formula bar, write measure
3. Best Flights =
`filter(Flight_information,Flight_information[Status_type]="Best")`
4. You will get the new table as filter table



Flight ID	FlightNumber	Airline	Destination	Status	Text After Delimiter	Status type
1001	FL1102	Airline D	Houston	On Time	1102	Best
1002	FL1435	Airline B	Chicago	On Time	1435	Best
1006	FL1071	Airline A	Phoenix	On Time	1071	Best
1011	FL1466	Airline A	Phoenix	On Time	1466	Best
1013	FL1330	Airline C	Houston	On Time	1330	Best
1020	FL1130	Airline A	New York	On Time	1130	Best
1023	FL1769	Airline A	Chicago	On Time	1769	Best
1025	FL1491	Airline D	Phoenix	On Time	1491	Best
1027	FL1805	Airline D	Chicago	On Time	1805	Best
1028	FL1385	Airline D	Chicago	On Time	1385	Best
1029	FL1191	Airline D	Los Angeles	On Time	1191	Best
1030	FL1955	Airline B	Phoenix	On Time	1955	Best
1031	FL1276	Airline B	New York	On Time	1276	Best
1033	FL1459	Airline D	New York	On Time	1459	Best
1034	FL1313	Airline B	Phoenix	On Time	1313	Best
1036	FL1252	Airline D	Phoenix	On Time	1252	Best
1039	FL1560	Airline B	Chicago	On Time	1560	Best
1043	FL1681	Airline C	Houston	On Time	1681	Best
1044	FL1475	Airline B	Phoenix	On Time	1475	Best
1046	FL1975	Airline D	Chicago	On Time	1975	Best
1048	FL1189	Airline A	New York	On Time	1189	Best
1050	FL1686	Airline C	Phoenix	On Time	1686	Best
1052	FL1562	Airline D	Phoenix	On Time	1562	Best
1053	FL1875	Airline C	Chicago	On Time	1875	Best
1055	FL1243	Airline B	New York	On Time	1243	Best

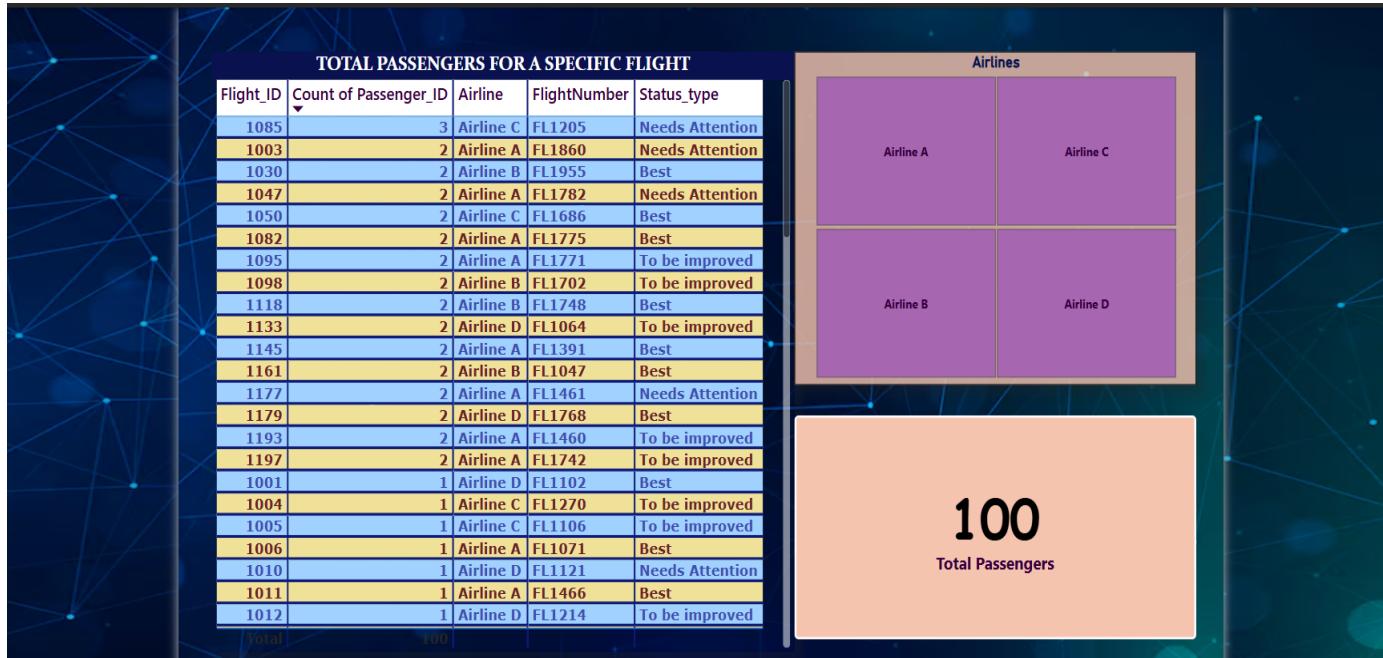
5. Visualization and Interactive Features (20 Marks)

- Create visuals for:
 - Passenger count by airline.
 - Ticket booking statuses.
 - Flights by airline and destination

Step 5(a)

1. From visualisation pane, select table view, add columns flight_id, count of passenger_id,airline, flight number,status_type.
 - From visualisation pane, add one card view , add measure of total passenger in it. Total Passengers:
 $\text{Total Passengers} = \text{DISTINCTCOUNT}(\text{Passenger_information})$

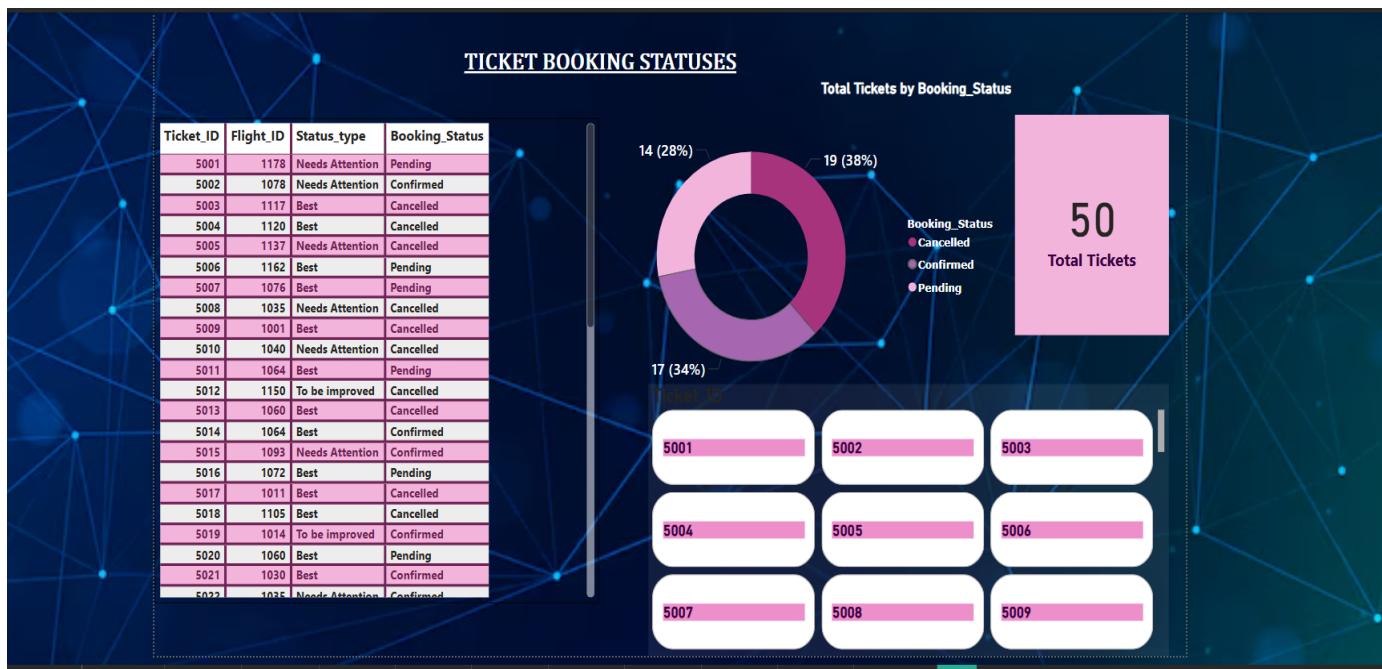
2. From visualisation pane add slicer in tile format and add airline in it.
3. Format all visuals using different backgrounds, font colour and styles.



TICKET BOOKING STATUS

1. Add table from visualisation pane , add columns ticket_id, flight_id, status_type, booking_status
2. Add a donut chart, with category booking_status, and value is total tickets from measure table
3. Add a one card visual of total number of tickets
4. Total Tickets = COUNTROWS(Ticket_information)

5. Add slicer with ticket id
6. Add different colours and styles to data, by changing colours and and increasing the size of font



Flights by airline and destination.

Steps

1. Add bar chart , For x axis= destination, for y axis= count of flight_id
2. Format it more by cell element, add different colours for minimum, maximum as well as middle value.
3. Add legends
4. Add slicer for different airlines.



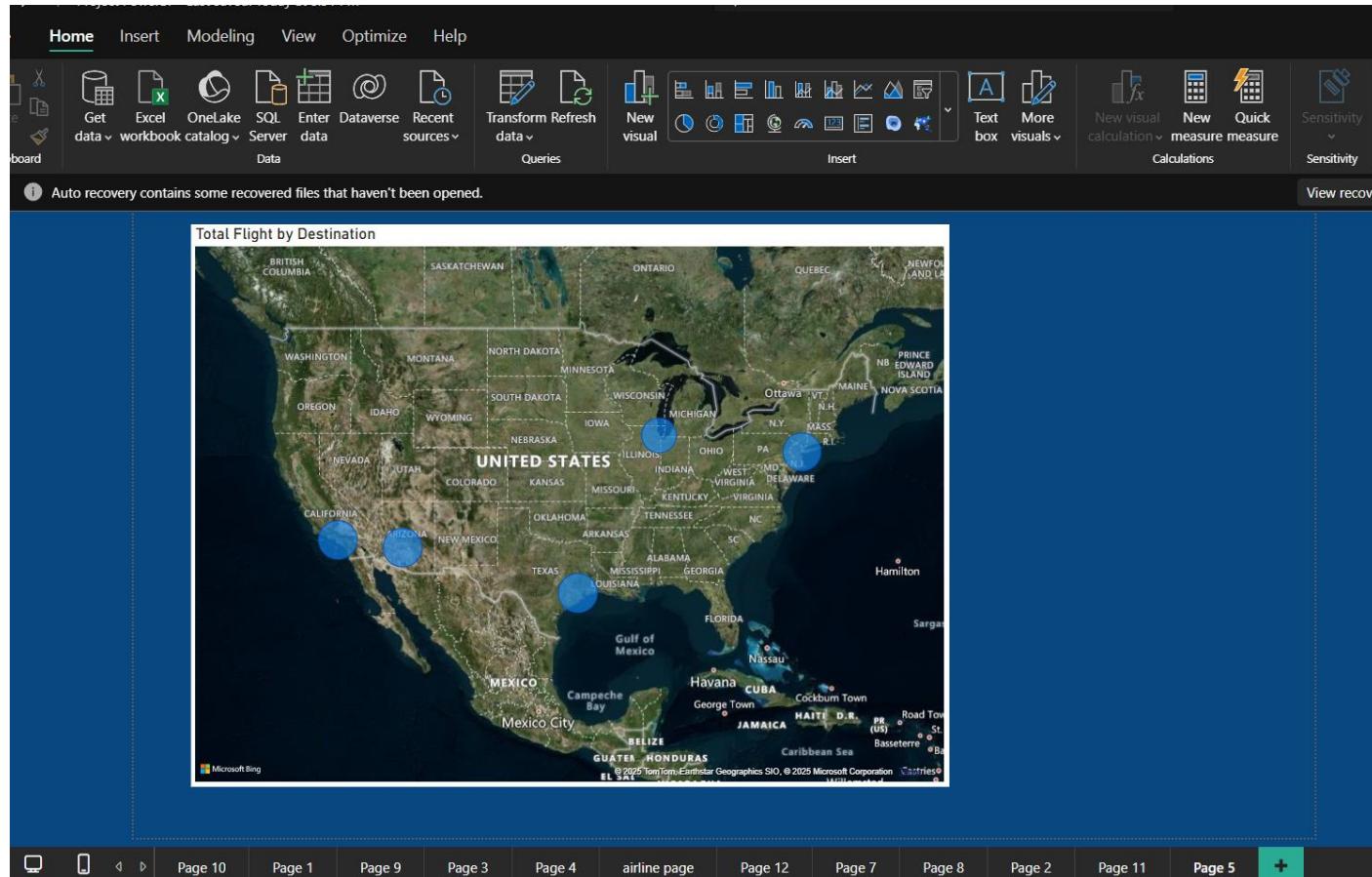
● Add interactive features for:

- Destination and Airline.
- Quickviews.
- Airline-specific pages

Steps for destination and airlines,

Steps

1. Create a map visual with destination as location and value as total flights
Total Flight = countrows(Flight_information)
2. Map settings as Aerial view



Quickviews

Steps

1. Step 1: Create a New Drillthrough Page. Add a new page in your Power BI report.
2: Add a Drillthrough Field that is flight_id

3.Drag the field you want to drill through (e.g., FlightID) into the "Drillthrough" filter well on the right pane.

3. Add visuals relevant to the context of the drillthrough.

4. For a FlightID drillthrough page, add:

5. Total Passengers (card) and total tickets visual.

6. Add table add flight id ,airline, status type,destination

7. Add a back button to return to the main report page:

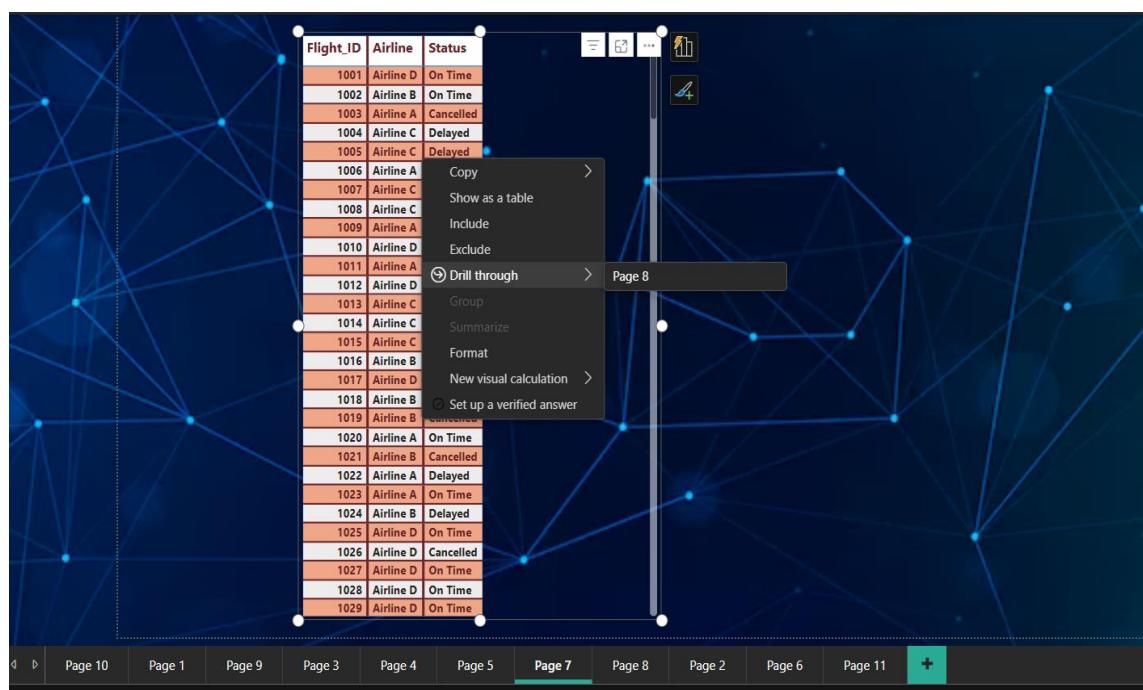
8. Insert → Buttons → Back

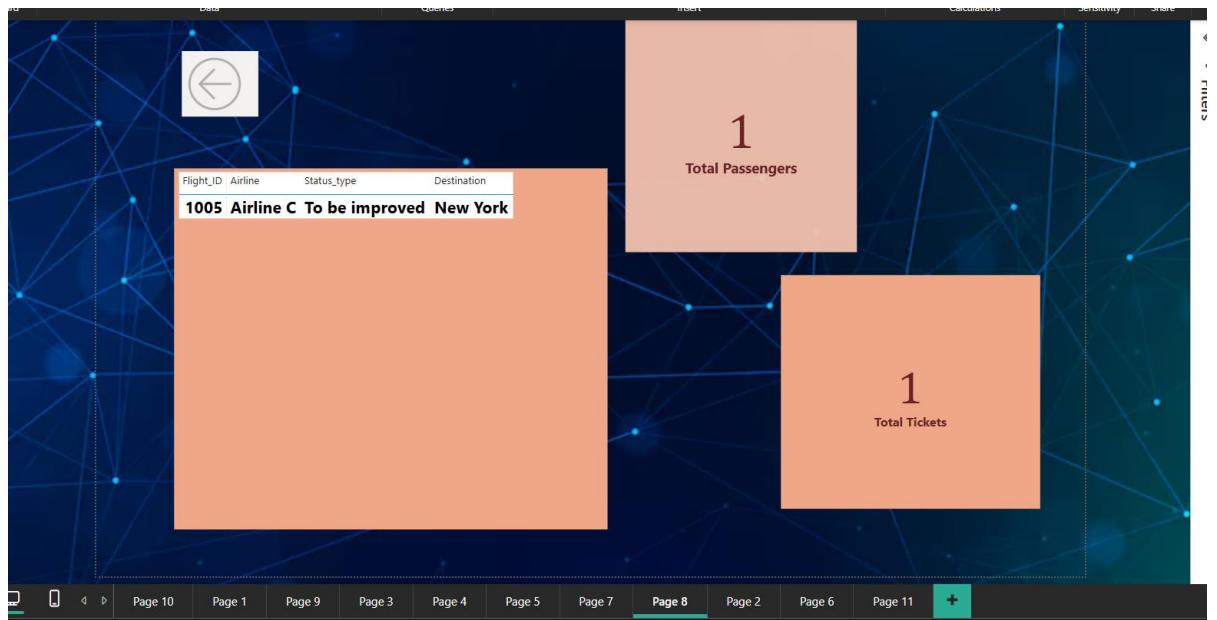
9. Set action type to Back in the Format pane.

10. Step 4: Use Drillthrough from Main Page that is page 8

11. Go to your main report/dashboard page.

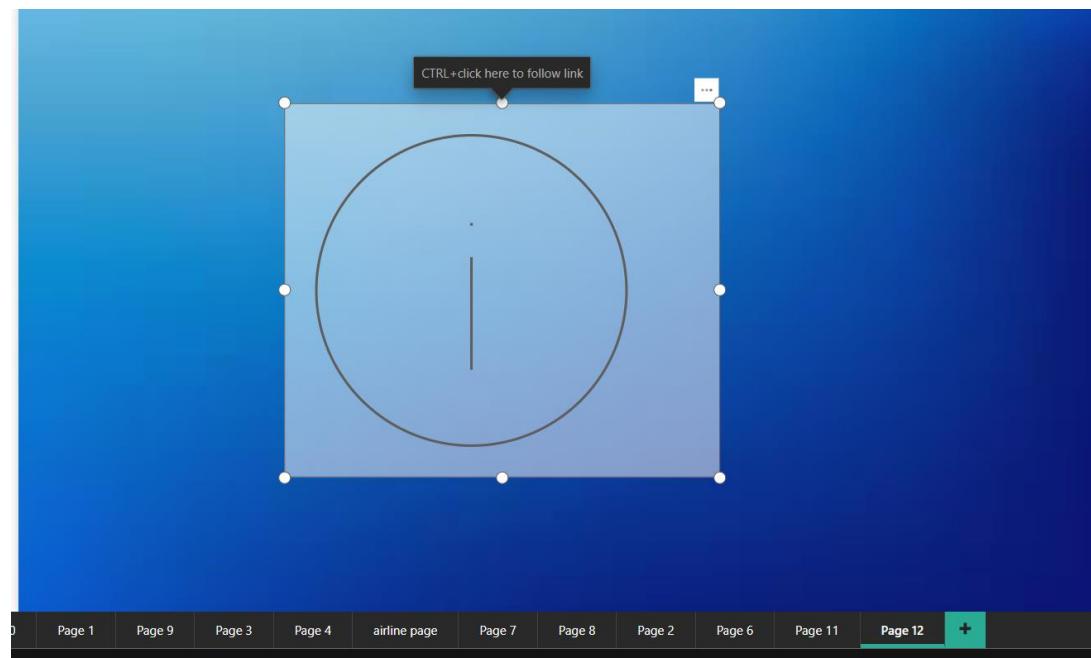
12. Right-click any visual element





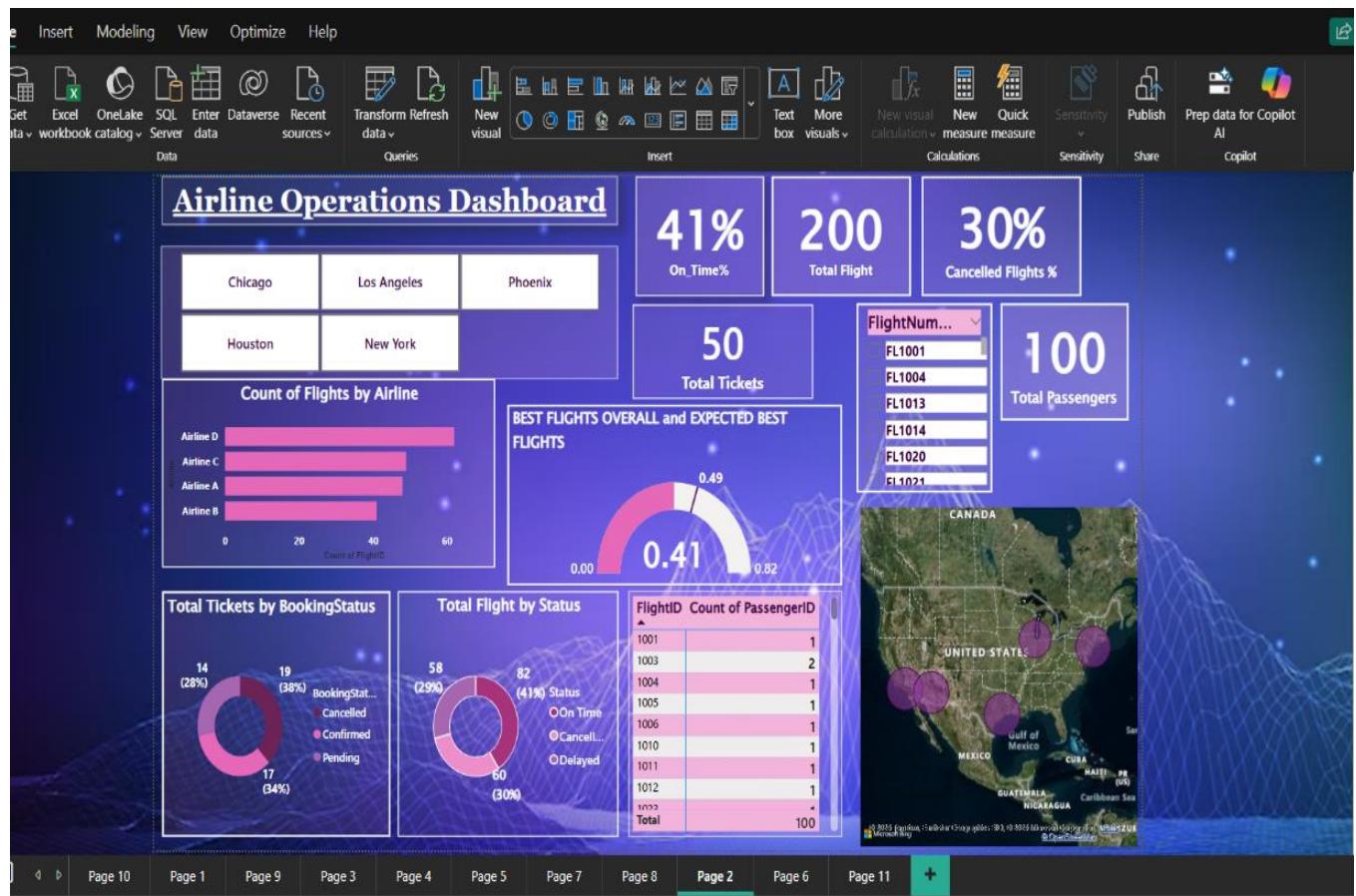
AIRLINE SPECIFIC PAGES

Create a button for page navigation and add action as airline page by clicking on ctrl+click the button will work





- 5. Final Dashboard and Power BI Service (20 Marks)** ● Design a comprehensive dashboard with key visuals and insights. ● Configure Row-Level Security (RLS) for Airline A data and assign it to a user. ● Setup a schedule refresh at 5 PM daily.



2. Create Key Visual Components

◆ KPIs (Top Row)

- Use Card visuals:

- On Time %:

On_Time% = DIVIDE([On_Time],[Total Flight],0)

**On_Time = CALCULATE(COUNTROWS(Flight_information),
Flight_information[Status] = "On Time")**

- Total Flight:

Total Flight = COUNTROWS(Flights_information)

- Cancelled Flights %:

Cancelled% Cancelled Flights % = divide([Cancelled flights],[Total Flight],0)

Cancelled flights =

calculate(countrows(Flight_information),Flight_information[Status] = "Cancelled")

- Total Tickets:

Total Tickets = COUNTROWS(Tickets_information)

- Total Passengers:

Total Passengers = DISTINCTCOUNT(Passenger_information)

- Slicer for Flight Number

Add a Slicer visual with FlightNumber to allow filtering across the report.

- Location Filter (Destination)

Use a slicer for destination : Chicago, Los Angeles, Houston, etc.

Field: destination from flight table

- Count of Flights by Airline (Bar Chart)

Visual: Clustered Bar Chart

X-axis: Count of FlightID

Y-axis: Airline

- Donut Charts

Total Tickets by Booking Status:

Legend: BookingStatus

Values: total tickets from measure table

- **Total Flights by Status:**

Legend: Flight Status

Values: total flights from measure table

- **Gauge Visual**

Best Flights Overall (Custom KPI Metric or Normalized Score),

BEST FLIGHTS OVERALL = DIVIDE([best flights%],[Total Flight],0)

EXPECTED BEST FLIGHTS = [BEST FLIGHTS OVERALL]*1.20 (I took it more the calculation which is 2%higher)

best flights% = countrows('Best Flights')

Best flight table is filter table that we have created

Use a Gauge visual.

Value: Best flights overall

Target/Maxexpected best flights

- **Table: FlightID vs. Count of PassengerID**

Visual: Table

Rows: FlightID

Values: Count of PassengerID

- **Map: Flight or City Visualization**

Visual: Map or Filled Map

Legends: destination

Bubble Size: Count of Flights from measure table.

We can also create dashboard in power bi service by pinning different visuals.

AIRLINE DATA

File Share Chat in Teams Comment Subscribe to dashboard Edit ...

55 days left

Flight ID	Airline	Status
1001	Airline D	On Time
1002	Airline B	On Time
1003	Airline A	Cancelled
1004	Airline C	Delayed
1005	Airline C	Delayed
1006	Airline A	On Time
1007	Airline C	Cancelled
1008	Airline C	Delayed
1009	Airline A	Cancelled
1010	Airline D	Cancelled
1011	Airline A	On Time
1012	Airline D	Delayed

Count of Flights by Airline

Airline	Count
Airline D	60
Airline C	45
Airline A	45
Airline B	40

Ticket_ID, Flight_ID, Status_type, Booking_Status

Ticket_ID	Flight_ID	Status_type	Booking_Status
5001	1178	Needs Attention	Pending
5002	1078	Needs Attention	Confirmed
5003	1117	Best	Cancelled
5004	1120	Best	Cancelled
5005	1137	Needs Attention	Cancelled
5006	1162	Best	Pending
5007	1076	Best	Pending
5008	1035	Needs Attention	Cancelled
5009	1001	Best	Cancelled
5010	1040	Needs Attention	Cancelled

AIRLINE DATA - Proper analysis of data of different airline...

AIRLINE DATA DASHBOARD

TOTAL PASSENGERS FOR A SPECIFIC FLIGHT

Flight_ID	Count of Passenger_ID	Airline	FlightNum
1085	3	Airline C	FL1205
1003	2	Airline A	FL1860
1030	2	Airline B	FL1955
1047	2	Airline A	FL1782
1050	2	Airline C	FL1686
1082	2	Airline A	FL1775
1095	2	Airline A	FL1771
1098	2	Airline B	FL1702
1118	2	Airline B	FL1748
1122	2	Airline D	FL1064
Total	100		

Ticket_ID, Flight_ID, Status_type, Booking_Status

Ticket_ID	Flight_ID	Status_type	Booking_Status
5001	1178	Needs Attention	Pending
5002	1078	Needs Attention	Confirmed
5003	1117	Best	Cancelled

Count of Flight_ID BY DESTINATION

Destination	Count of Flight_ID
Houston	43
Los Angeles	38
Phoenix	40
New York	40
Chicago	33

Total Tickets

On_Time%

41%

Total Flight

200

Total Passengers



● Configure Row-Level Security (RLS) for Airline A data and assign it to a user.

Step 1: Create RLS Role in Power BI Desktop

1. Open your .pbix report in Power BI Desktop.
2. Go to the Modeling tab in the top ribbon.
3. Click on Manage Roles.
4. Click Create to make a new role.
5. Rename the role to:Airline_A
6. Select the table that contains the Airline column (e.g., FlightInformation).
7. In the DAX filter box, enter: [Airline] = "Airline A"

Step 2: Validate the Role

Click View as Roles.

1. Select AirlineA_Role.
2. Confirm that the report now only shows data for Airline_A.
3. Click File > Publish > Publish to Power BI Service.

- Open the workspace where your dataset is published

1.Datasets + dataflows, click the More options (:) next to your dataset

2.Choose Security

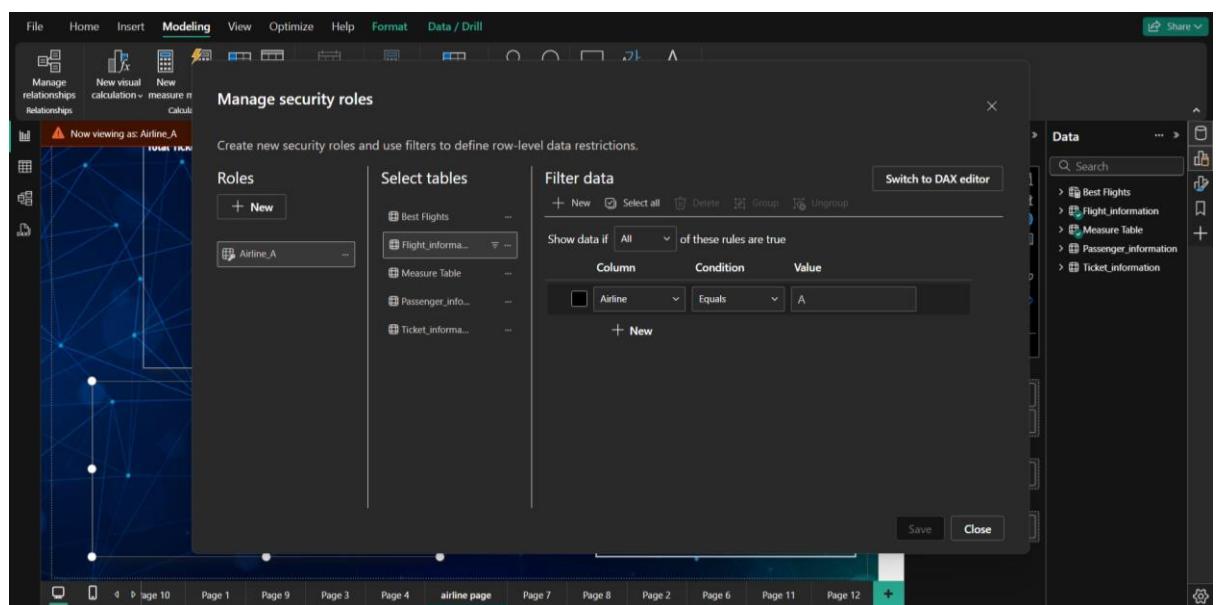
3.You'll see the role you created (e.g., AirlineA_Role)

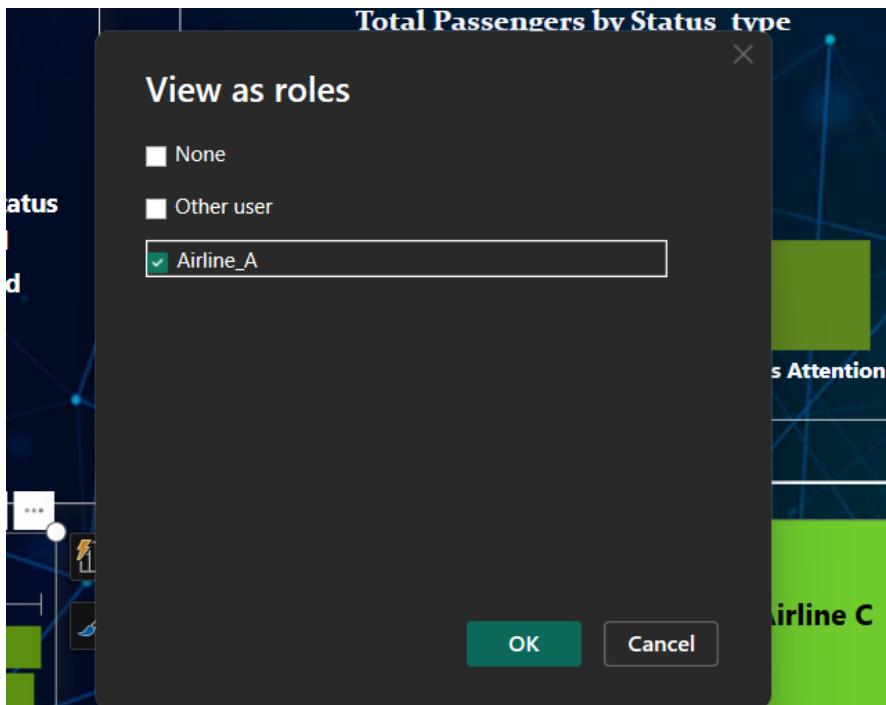
4.Click on the role name

5.In the Members box, enter the user's email ID, such as:

xegil74150@calorp.com

6. Add and save





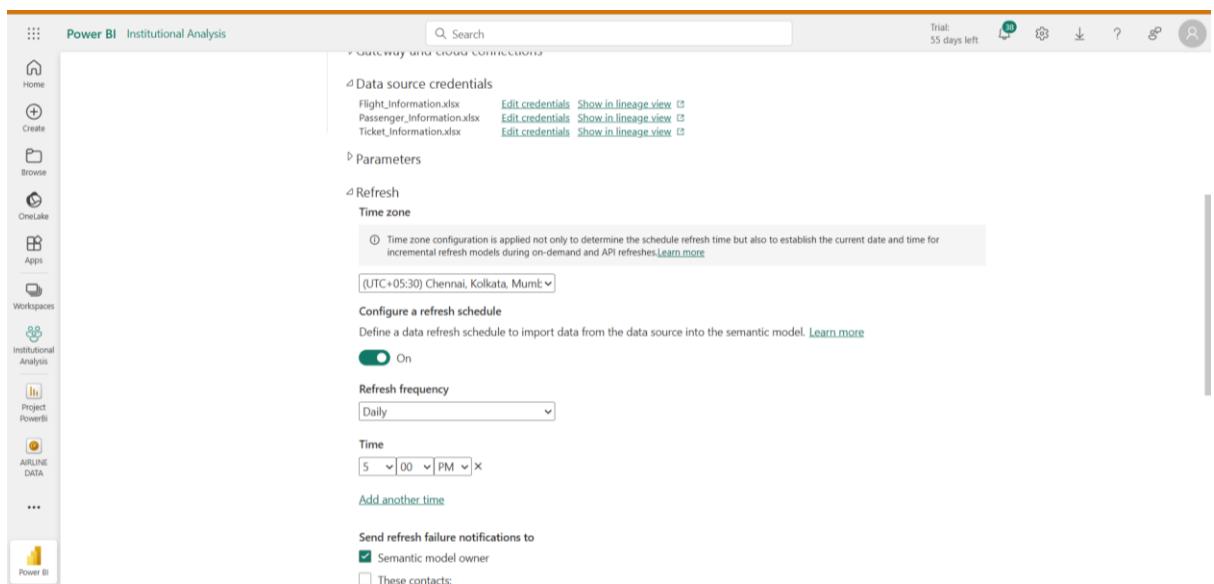
A screenshot of the Power BI Row-Level Security settings page. The left sidebar shows navigation options like Home, Create, Browse, OneLake, Apps, Metrics, Workspaces, and the current workspace "Airline data analysis". The main area is titled "Row-Level Security" and shows a role named "Airline_A (1)". Under "Members (1)", it lists "Simran Soni" with an "x" button to remove. There is a search bar at the top and "Save" and "Cancel" buttons at the bottom.

- **Setup a schedule refresh at 5 PM daily**

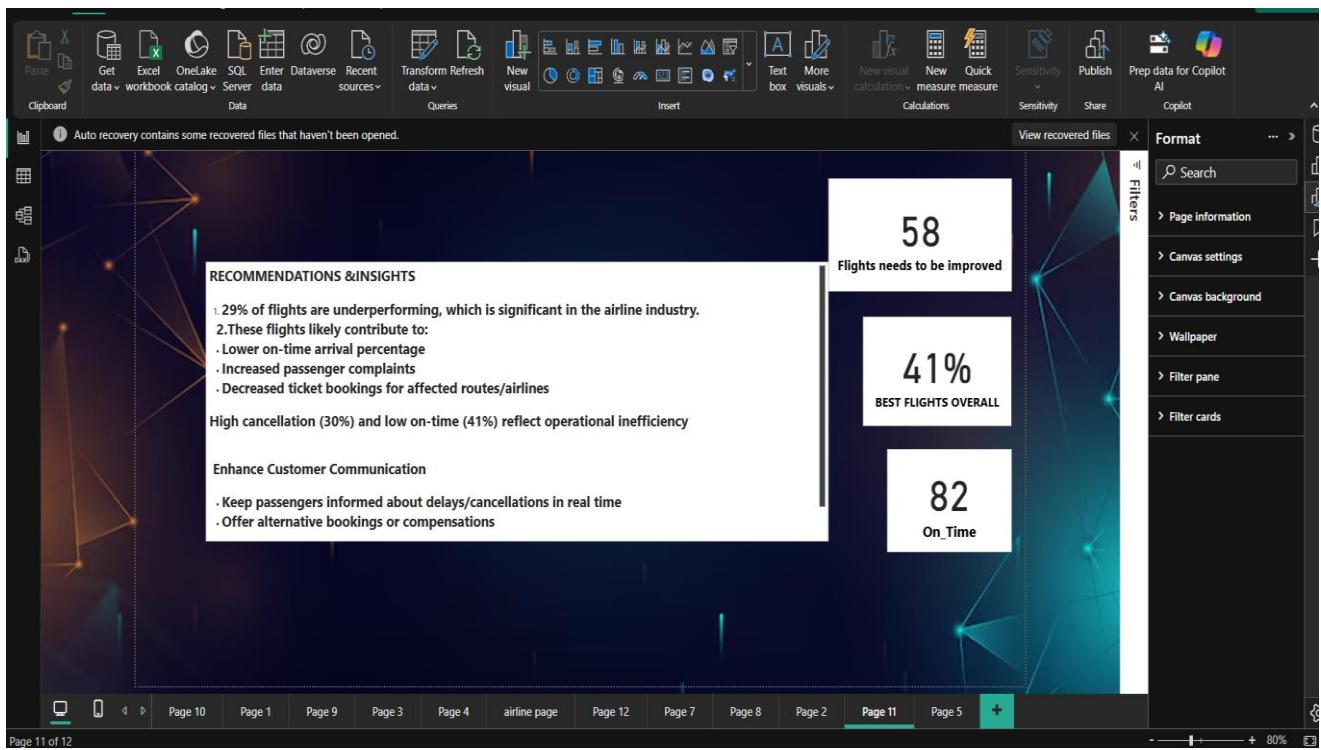
Step 1: Download gateway

1. In the top left panel download gateway of Institutional analysis workspace
2. Download personal gateway
3. Set up the gateway
4. In the institutional analysis of airline data go to settings> power bi setting> semantic models> schedule refresh
5. Under **Refresh frequency:** Select **Daily**.

Under **Time**, select **5:00 pm**



Some Insights and suggestions:



Data Insights and Recommendations

Flights needs to be improved =

`calculate(countrows(Flight_information),Flight_information[Status_type]="To be improved")`

1. Highest Number of Flights from Houston

- **Houston emerges as the most active destination, indicating it's a major hub or a highly demanded location in the current flight network.**

2. Maximum Flights by Airline D – 62 Flights

- **Airline D is operating the highest number of flights, showing strong operational capacity or customer preference.**

3. Minimum Flights by Airline B – 41 Flights

- **Airline B has the lowest number of flights, which may point to limited market coverage, operational restrictions, or low demand.**

4. Recommendations:

1. **Strengthen Operations in Houston**
2. **Analyze Airline D's Success Factors**
3. **Expanding to high-demand destinations like Houston**

4. Running pilot routes or promotional campaigns to boost demand

Video part 1 : https://drive.google.com/file/d/1ojdJxVpgOsKb8pHJ6_bSE4lk9PYDu-KV/view?usp=sharing

Video part 2:

https://drive.google.com/file/d/1hPG4Z39a0a6L1ZKhb1tcgDnL2r8KA1_r/view?usp=sharing