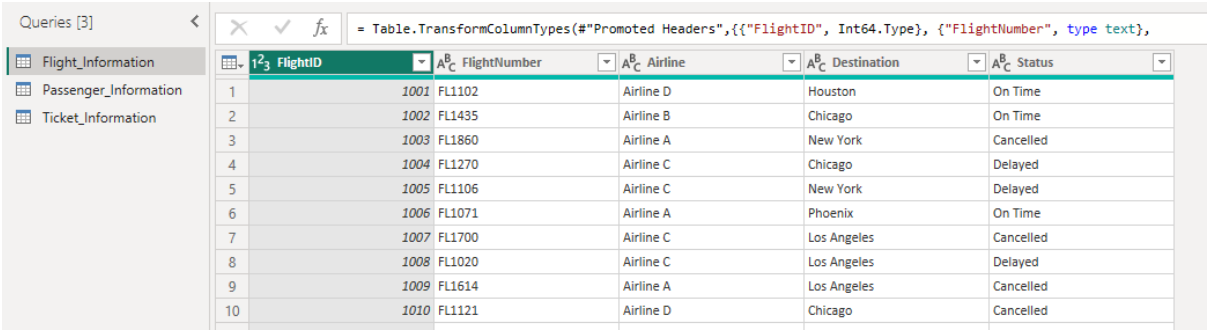
Final Project

Project Title: Airline Data Management and Analysis Using Power BI

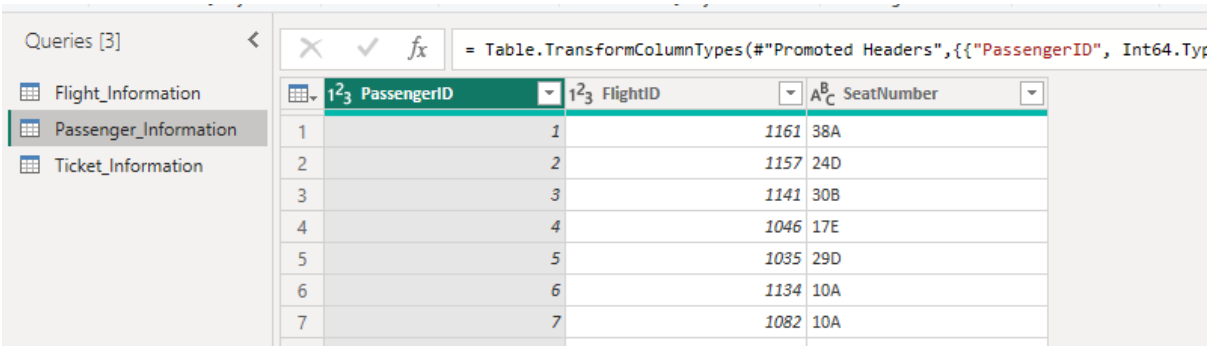
1. Data Preparation and Cleaning

- **Extract and transform data in Power Query.**

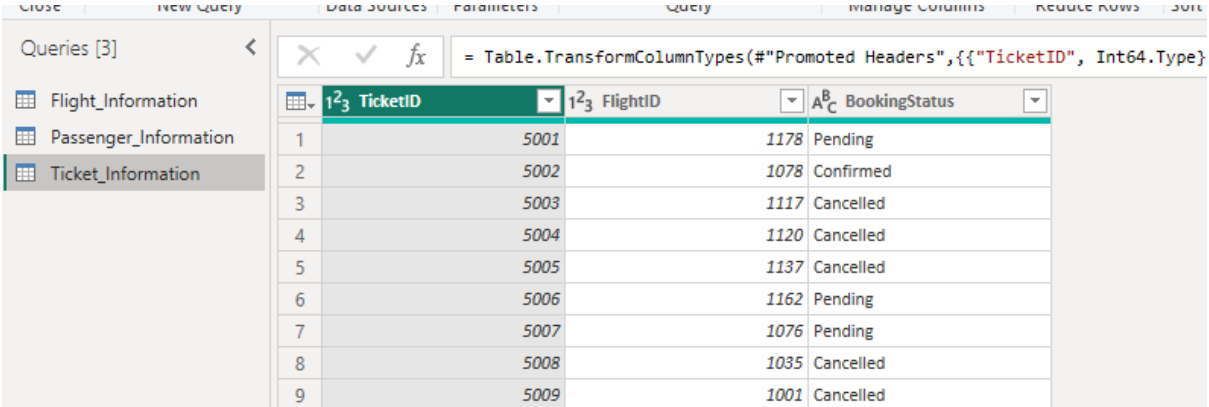
Ans: Open Power BI, then navigate to the Data tab and select Get Data > From File, depending on where the data is stored. Import the three datasets—Flight_Information, Passenger_Information, and Ticket_Information—and open each dataset in the Power Query Editor.



	FlightID	FlightNumber	Airline	Destination	Status
1	1001	FL1102	Airline D	Houston	On Time
2	1002	FL1435	Airline B	Chicago	On Time
3	1003	FL1860	Airline A	New York	Cancelled
4	1004	FL1270	Airline C	Chicago	Delayed
5	1005	FL1106	Airline C	New York	Delayed
6	1006	FL1071	Airline A	Phoenix	On Time
7	1007	FL1700	Airline C	Los Angeles	Cancelled
8	1008	FL1020	Airline C	Los Angeles	Delayed
9	1009	FL1614	Airline A	Los Angeles	Cancelled
10	1010	FL1121	Airline D	Chicago	Cancelled



	PassengerID	FlightID	SeatNumber
1	1	1161	38A
2	2	1157	24D
3	3	1141	30B
4	4	1046	17E
5	5	1035	29D
6	6	1134	10A
7	7	1082	10A

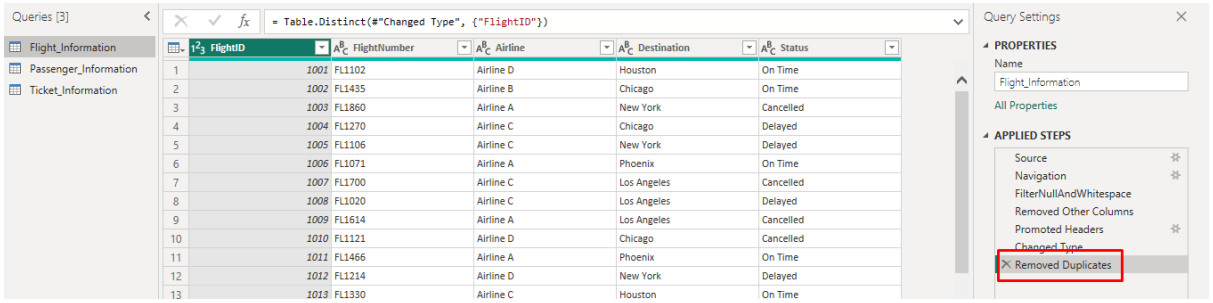


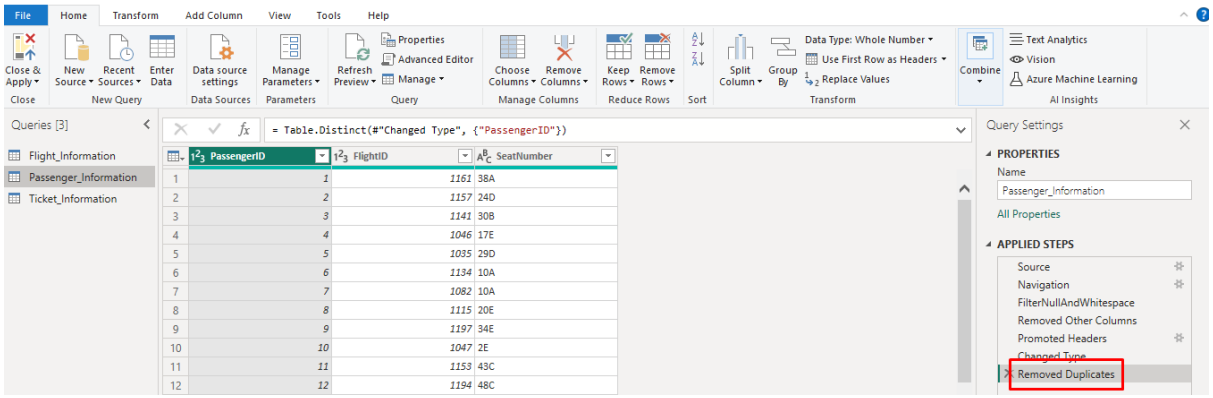
	TicketID	FlightID	BookingStatus
1	5001	1178	Pending
2	5002	1078	Confirmed
3	5003	1117	Cancelled
4	5004	1120	Cancelled
5	5005	1137	Cancelled
6	5006	1162	Pending
7	5007	1076	Pending
8	5008	1035	Cancelled
9	5009	1001	Cancelled

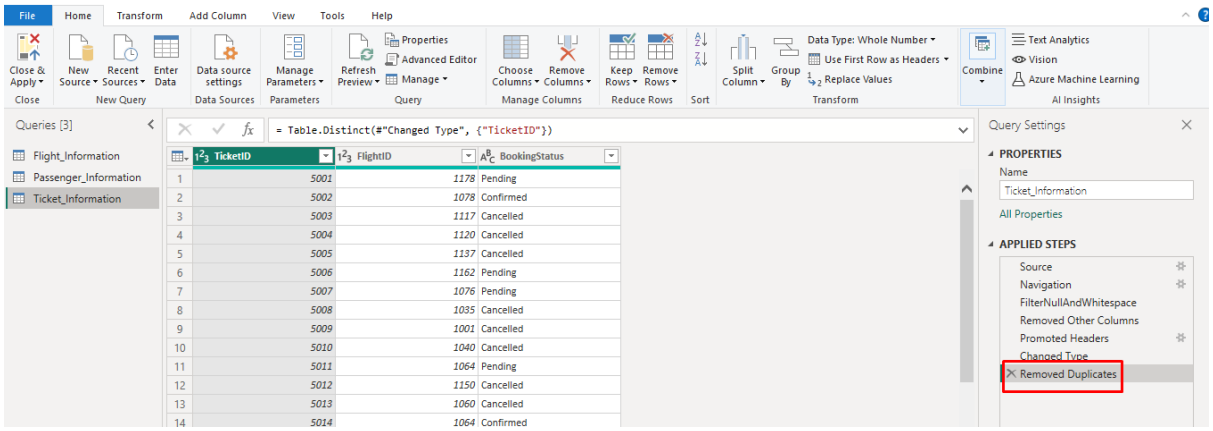
● **Clean data: remove duplicates, handle missing values, and format columns.**

Ans: In the Power Query Editor, for the “Flight_Information” table, select the “Flight_ID” column that defines unique rows. Similarly, for the “Passenger_Information” table, select the “PassengerID” column, and for the “Ticket_Information” table, select the “TicketID” column.

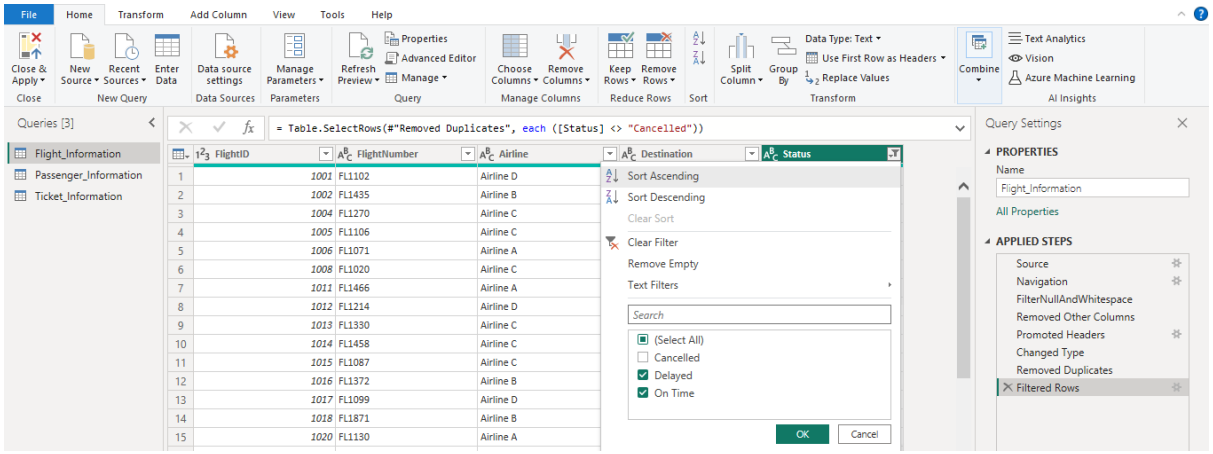
Next, navigate to the “Home” tab, click the “Remove Rows” dropdown, and select “Remove Duplicates”.

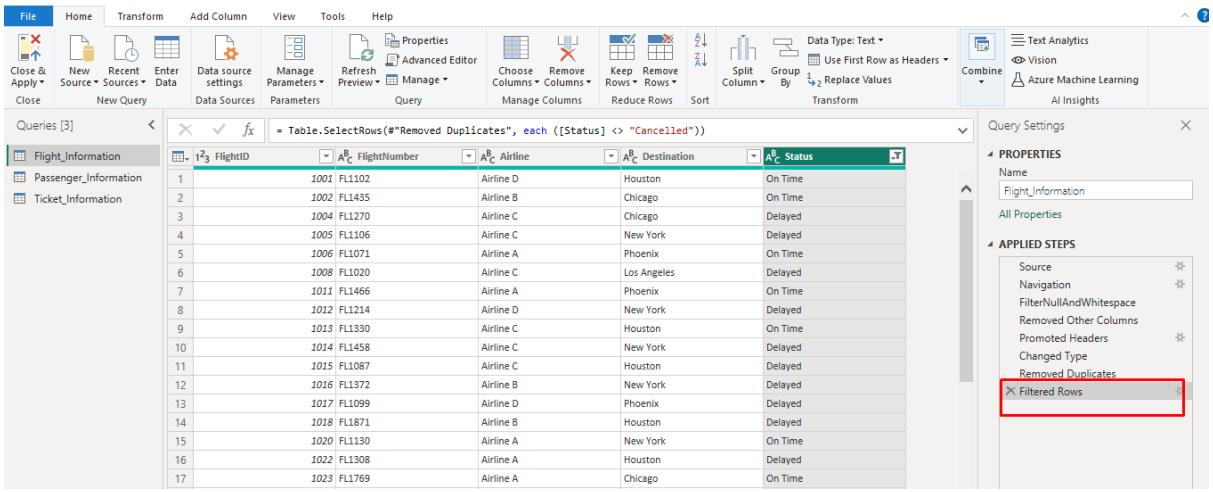


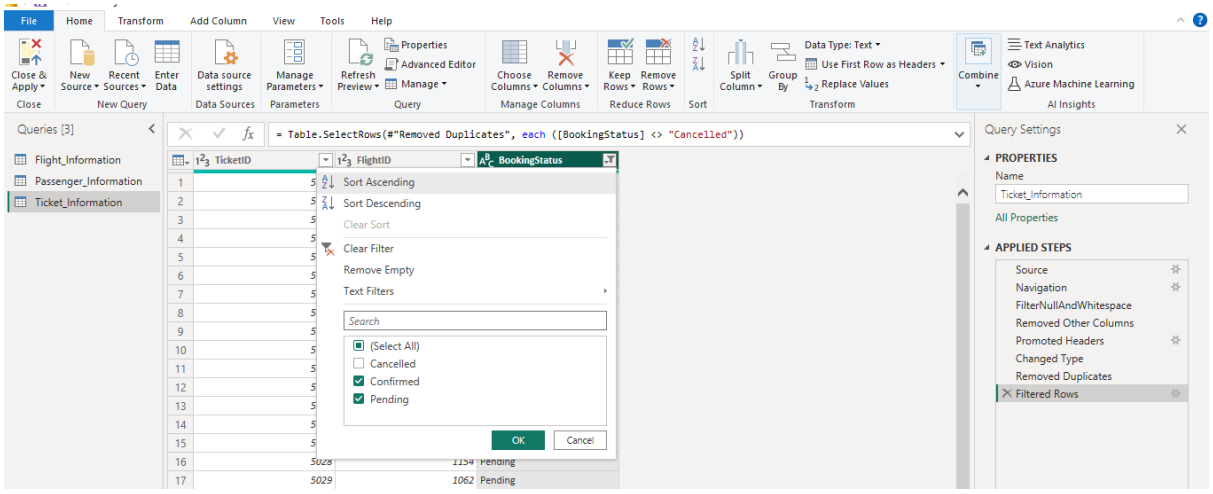


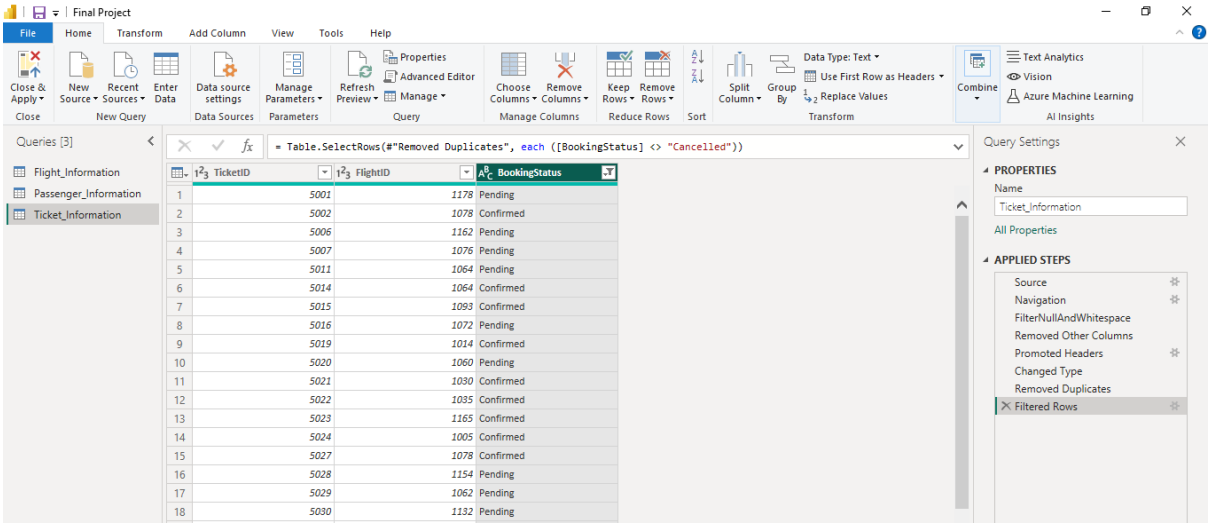


In Power query editor, for “Flight Information” table, filtered the “Status” column and unchecked “Cancelled”, cause this is the unnecessary rows. Same task doing for “Ticket_Information” table, filtered “BookingStatus” column and unchecked “Cancelled”.


File
Home
Transform
Add Column
View
Tools
Help
Close & Apply
New Source
Recent Sources
Enter Data
Data source settings
Manage Parameters
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Text
Use First Row as Headers
Replace Values
Combine
Text Analytics
Vision
Azure Machine Learning
Close
New Query
Data Sources
Parameters
Query
Manage Columns
Reduce Rows
Sort
Transform
AI Insights
Queries [3]
Flight_Information
Passenger_Information
Ticket_Information
= Table.SelectRows(#"Removed Duplicates", each ([Status] <> "Cancelled"))
FlightID
FlightNumber
Airline
Destination
Status
1001 FL1102 Airline D
1002 FL1435 Airline B
1003 FL1270 Airline C
1004 FL1106 Airline C
1005 FL1071 Airline A
1006 FL1020 Airline C
1007 FL1466 Airline A
1008 FL1214 Airline D
1009 FL1330 Airline C
1010 FL1458 Airline C
1011 FL1087 Airline C
1012 FL1372 Airline B
1013 FL1099 Airline D
1014 FL1871 Airline B
1015 FL1130 Airline A
Sort Ascending
Sort Descending
Clear Sort
Clear Filter
Remove Empty
Text Filters
Search
(Select All)
Cancelled
Delayed
On Time
OK
Cancel
Query Settings
PROPERTIES
Name
Flight_Information
All Properties
APPLIED STEPS
Source
Navigation
FilterNullAndWhitespace
Removed Other Columns
Promoted Headers
Changed Type
Removed Duplicates
Filtered Rows


File
Home
Transform
Add Column
View
Tools
Help
Close & Apply
New Source
Recent Sources
Enter Data
Data source settings
Manage Parameters
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Text
Use First Row as Headers
Replace Values
Combine
Text Analytics
Vision
Azure Machine Learning
Close
New Query
Data Sources
Parameters
Query
Manage Columns
Reduce Rows
Sort
Transform
AI Insights
Queries [3]
Flight_Information
Passenger_Information
Ticket_Information
= Table.SelectRows(#"Removed Duplicates", each ([Status] <> "Cancelled"))
FlightID
FlightNumber
Airline
Destination
Status
1001 FL1102 Airline D Houston On Time
1002 FL1435 Airline B Chicago On Time
1003 FL1270 Airline C Chicago Delayed
1004 FL1106 Airline C New York Delayed
1005 FL1071 Airline A Phoenix On Time
1006 FL1020 Airline C Los Angeles Delayed
1007 FL1466 Airline A Phoenix On Time
1008 FL1214 Airline D New York Delayed
1009 FL1330 Airline C Houston On Time
1010 FL1458 Airline C New York Delayed
1011 FL1087 Airline C Houston Delayed
1012 FL1372 Airline B New York Delayed
1013 FL1099 Airline D Phoenix Delayed
1014 FL1871 Airline B Houston Delayed
1015 FL1130 Airline A New York On Time
1016 FL1308 Airline A Houston Delayed
1017 FL1769 Airline A Chicago On Time
Query Settings
PROPERTIES
Name
Flight_Information
All Properties
APPLIED STEPS
Source
Navigation
FilterNullAndWhitespace
Removed Other Columns
Promoted Headers
Changed Type
Removed Duplicates
Filtered Rows


File
Home
Transform
Add Column
View
Tools
Help
Close & Apply
New Source
Recent Sources
Enter Data
Data source settings
Manage Parameters
Refresh Preview
Properties
Advanced Editor
Manage
Choose Columns
Remove Columns
Keep Rows
Remove Rows
Split Column
Group By
Data Type: Text
Use First Row as Headers
Replace Values
Combine
Text Analytics
Vision
Azure Machine Learning
Close
New Query
Data Sources
Parameters
Query
Manage Columns
Reduce Rows
Sort
Transform
AI Insights
Queries [3]
Flight_Information
Passenger_Information
Ticket_Information
= Table.SelectRows(#"Removed Duplicates", each ([BookingStatus] <> "Cancelled"))
TicketID
FlightID
BookingStatus
Sort Ascending
Sort Descending
Clear Sort
Clear Filter
Remove Empty
Text Filters
Search
(Select All)
Cancelled
Confirmed
Pending
OK
Cancel
Query Settings
PROPERTIES
Name
Ticket_Information
All Properties
APPLIED STEPS
Source
Navigation
FilterNullAndWhitespace
Removed Other Columns
Promoted Headers
Changed Type
Removed Duplicates
Filtered Rows



----------X----------

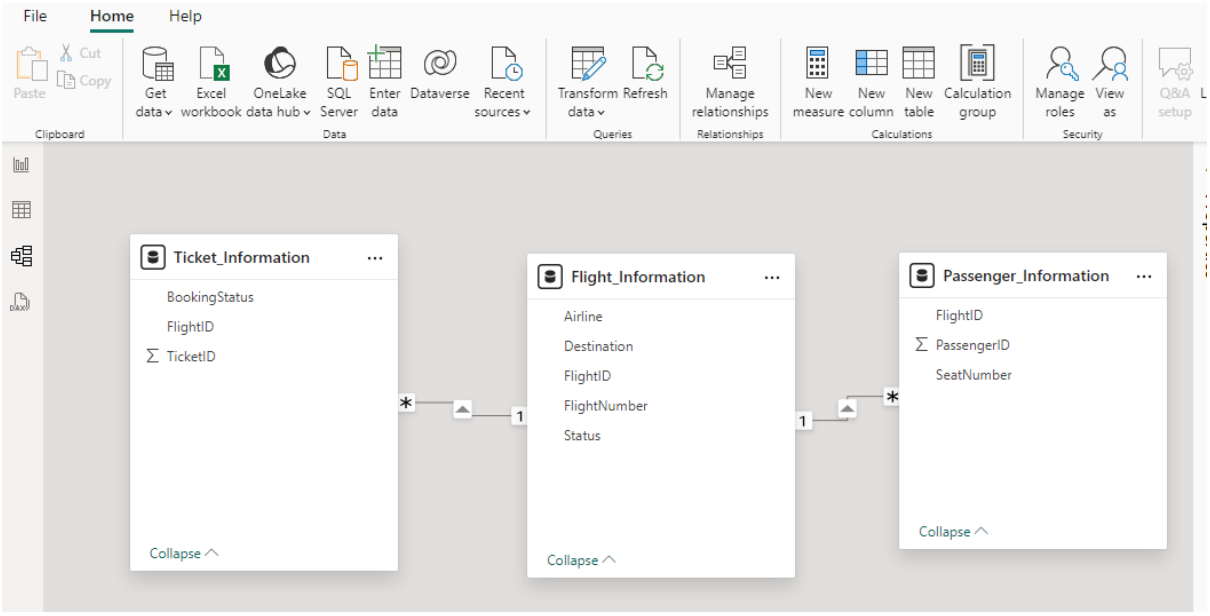
2. Data Modeling

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

Ans: In the “Model View”, Drag the “Flight_ID” column from the “Flight_Information” table and drop it onto the corresponding “Flight_ID” column in the “Passenger_Information” table.

Similarly, link the “Flight_ID” column from the “Flight_Information” table to the “Flight_ID” column in the “Ticket_Information” table.

The relationships are set to “One-to-Many”, with the *Flight_Information* table as the One side, and the *Passenger_Information* and *Ticket_Information* tables as the Many sides.



----------X----------

3. Enhanced Data Insights

● Add a conditional column to classify flights as "Best" or "To Be Improved" based on status.

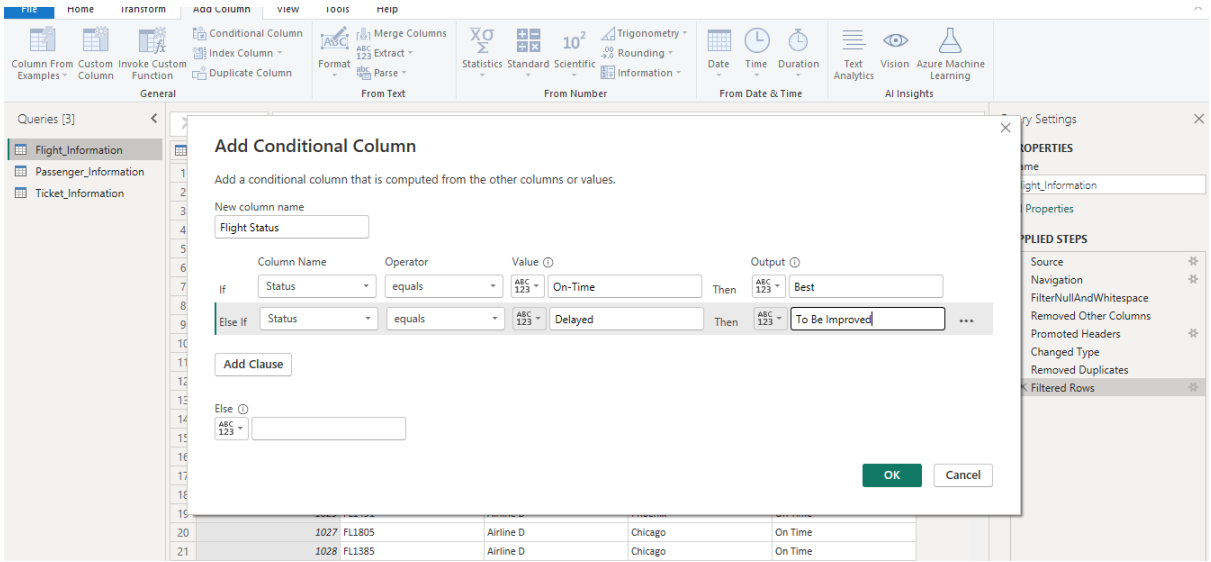
● Use "Column from Examples" to extract the flight number from FlightNumber.

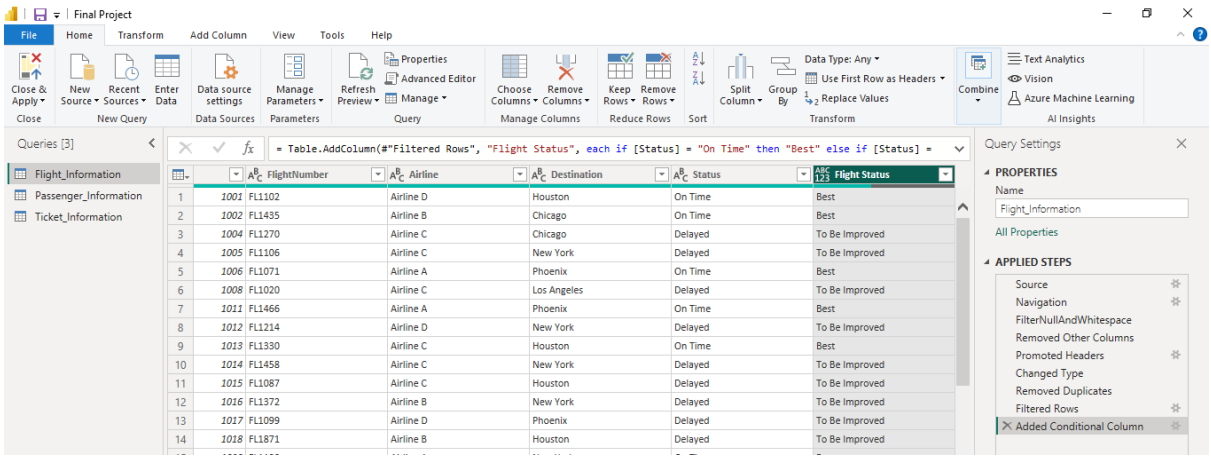
Ans: In the Power Query editor, Select the “Flight_Information” table, then go to the “Add Column” tab and click “Conditional Column”. Set the column name as “Flight Status”.

Then set the condition

- If Status equals 'On-Time' then 'Best'.
- If Status equals 'Delayed' then 'To Be Improved'.

and click “OK”.





For, extract the flight number from “FlightNumber” column, in the “Flight_Information” table, choose the “FlightNumber” column.

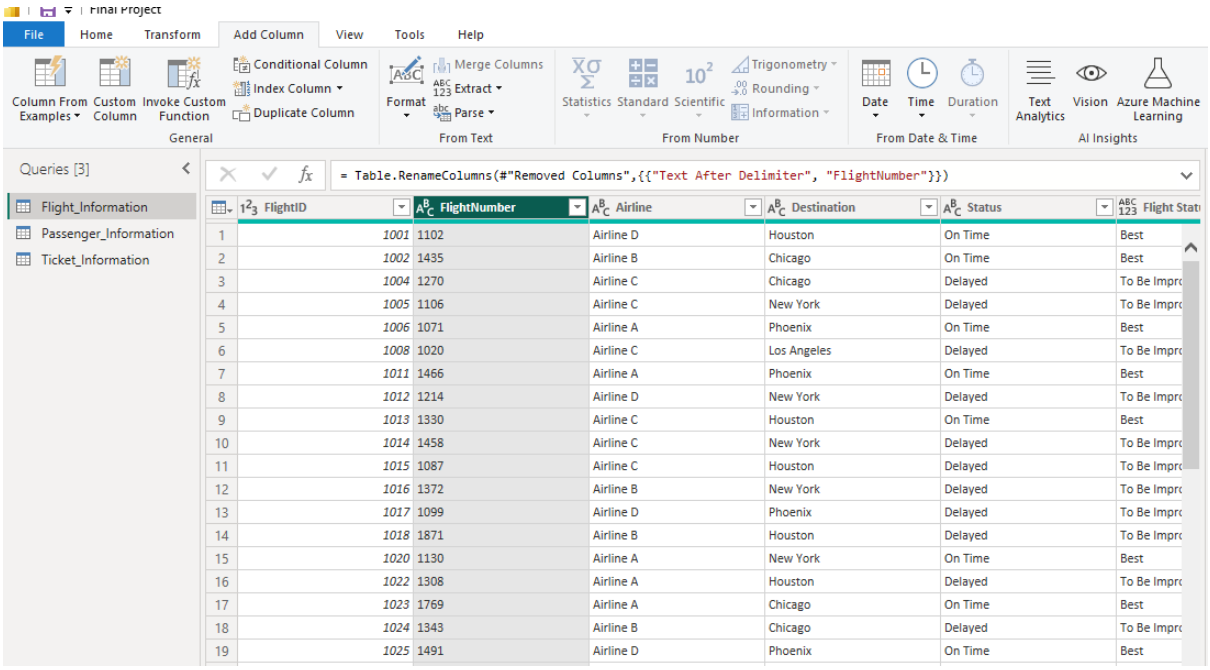
Then go to the “Add Column” tab and click “Column from Examples**”** > “From Selection**”**.

Then in the data preview section, start typing the **desired output** in the new column. For example,

If FlightNumber is "FL1102" and I want to extract "1102" then type "1102" in the new column.

After that, Power Query will automatically detect the pattern and suggest a formula to extract the flight number.

Then click “OK”.



----------X----------

4. Calculations Using DAX

- **Calculate:**

  - **Total tickets booked.**

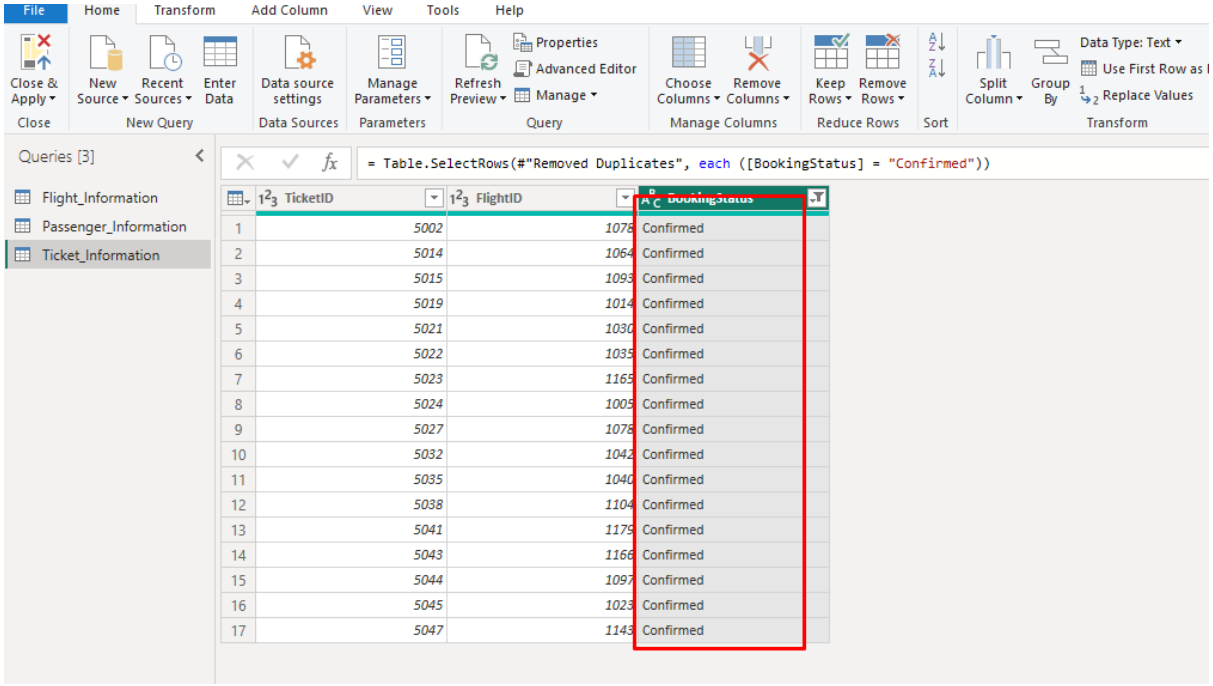
  - **Filtered table showing "Best" flights only.**

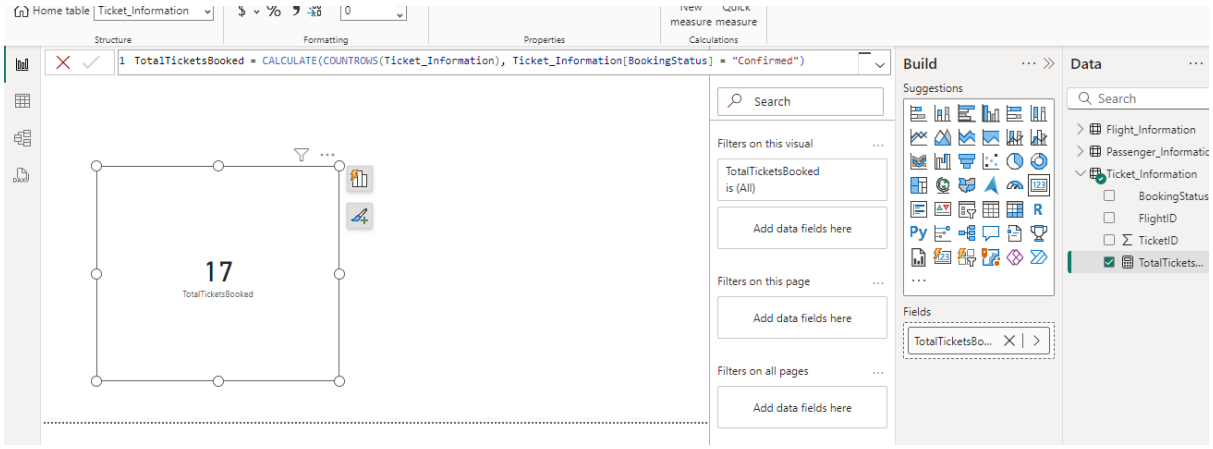
Ans: To calculate the “Total tickets booked” using DAX,

At first creating a measure, that counts the rows in the Ticket_Information table where the BookingStatus indicates that the ticket has been confirmed:

TotalTicketsBooked = CALCULATE(COUNTROWS(Ticket_Information),
Ticket_Information[BookingStatus] = "Confirmed")

Then Use a card visual to display the booked tickets.





So, total 17 tickets are booked after calculation.

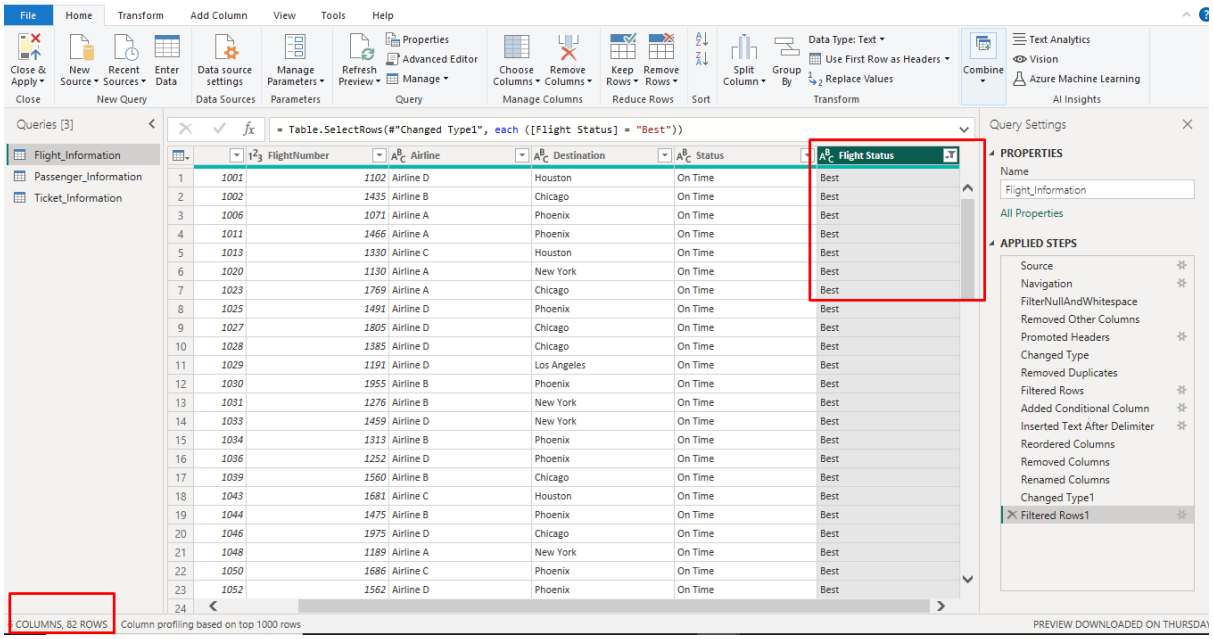
Calculate filtered table showing "Best" flights only:

Create a filtered table in Power BI that shows only the "Best" flights using DAX, firstly create a calculated table.

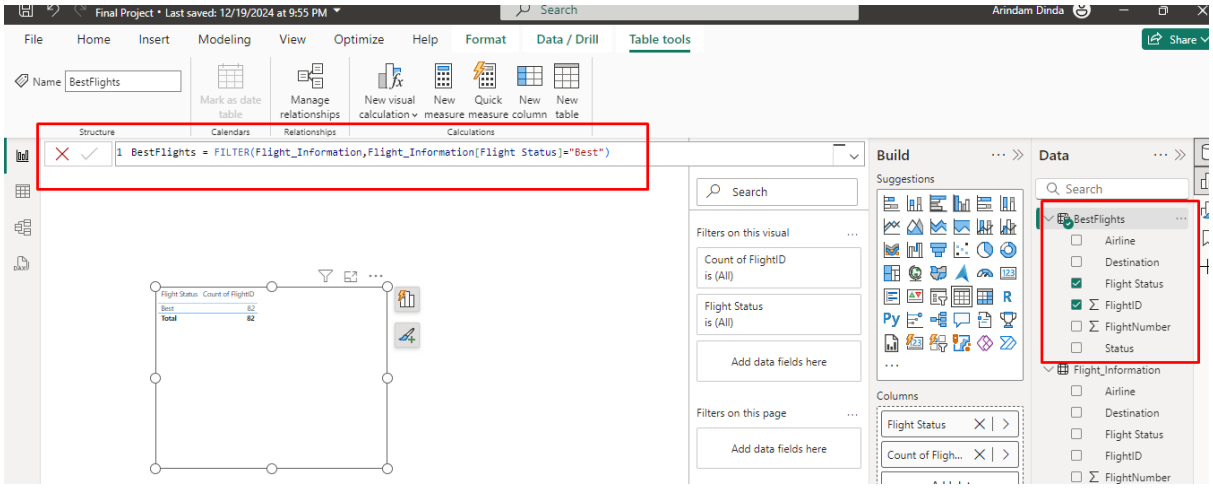
In “Modeling” tab, Select “New Table”. Then enter the DAX expression for “Best Flights”.

BestFlights = FILTER(Flight_Information,Flight_Information[Flight Status]="Best")

Then create a table visual, then drag the “Flight Status” and “Flight Number” column.



In Power Query Editor", only 82 best flight are there, after filtered,



----------X----------

5. Visualization and Interactive Features

● **Create visuals for: ○ Passenger count by airline.**

○ **Ticket booking statuses.**

○ **Flights by airline and destination.**

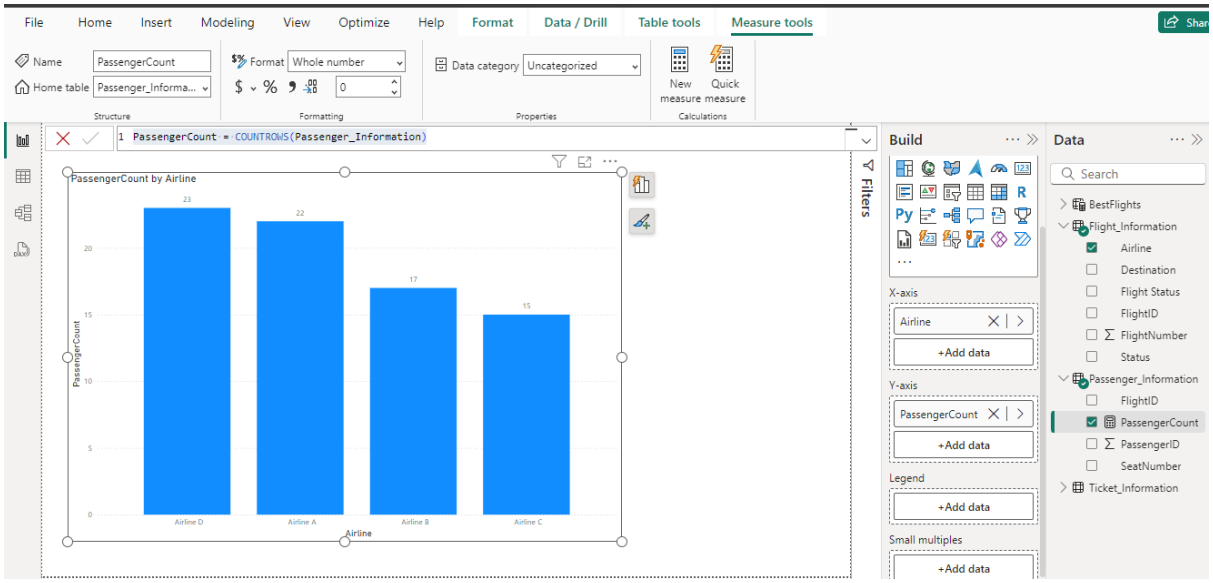
Ans: For Passenger count by airline-

Create a DAX measure in "Passenger_Information" to calculate the total number of passengers:

PassengerCount = COUNTROWS(Passenger_Information)

This measure counts the rows in the Passenger_Information table, where each row represents a passenger.

Then using “Stacked Column Chart”, this visual perfectly analyse.

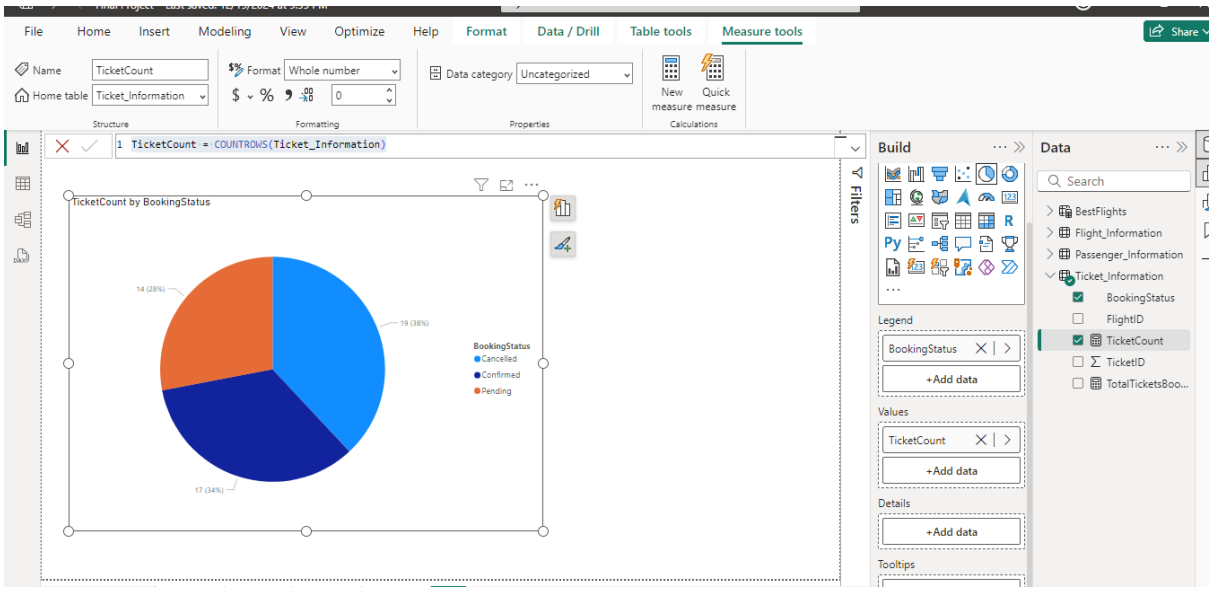


For ticket booking statuses:

Create a DAX formula in “Ticket_Information” to find the ticket count:

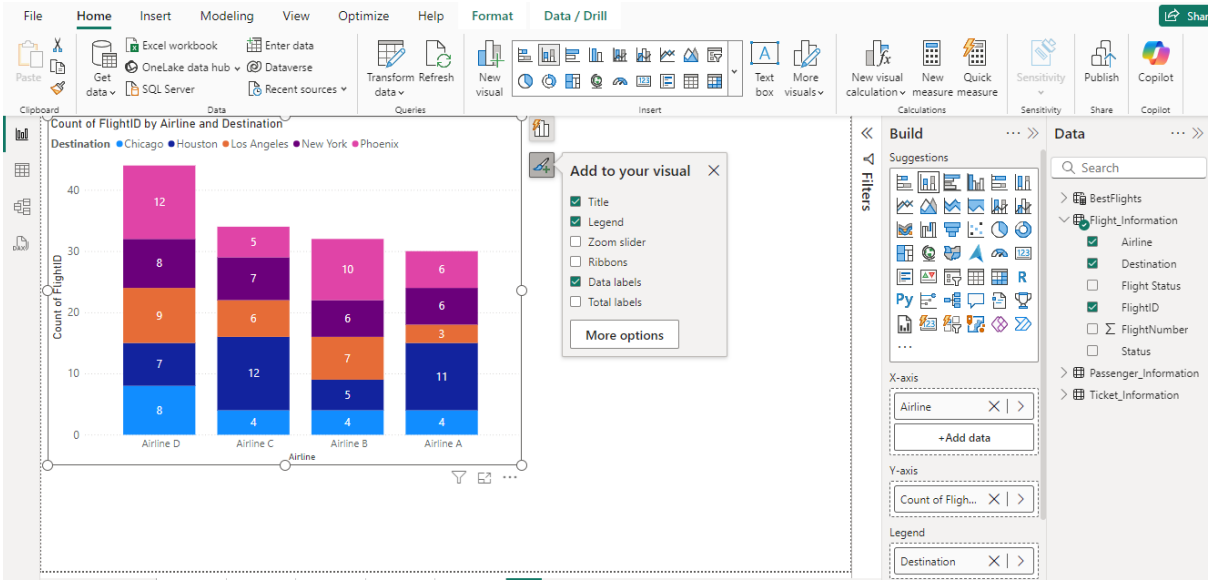
TicketCount = COUNTROWS(Ticket_Information)

Then using a “Pie Chart”, drag “BookingStatus” in “Legend” field and “TicketCount” in “Values” field.



For Flights by airline and destination:

For this, “Stacked column chart” is the best option.



Add interactive features for:

○ Destination and Airline.

○ Quick views.

○ Airline-specific pages.

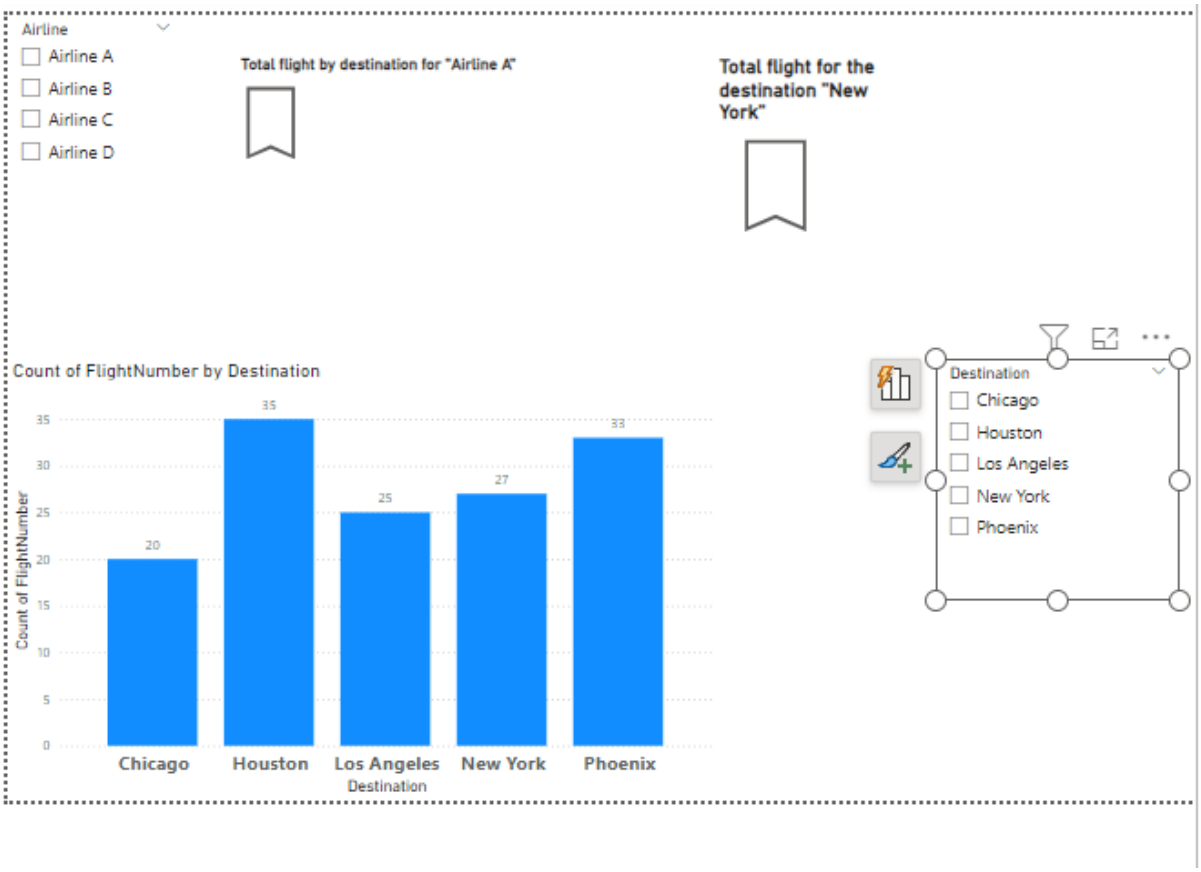
Ans:

<u>For Destination and Airline</u>:

To add interactive features for “Destination**”** and “Airline**”**, we can incorporate Slicers, Tooltips, Drill-through, and Interactive Filtering. These features allow users to filter data dynamically, view detailed information, and interact with the visuals on the report.

Slicers for Destination and Airline*:* Create two slicer, for “Airline” & “Destination.

Quick Views Using Bookmarks and Buttons: Quick views allow users to quickly navigate between different views of the data, such as overviews or detailed breakdowns for specific airlines.

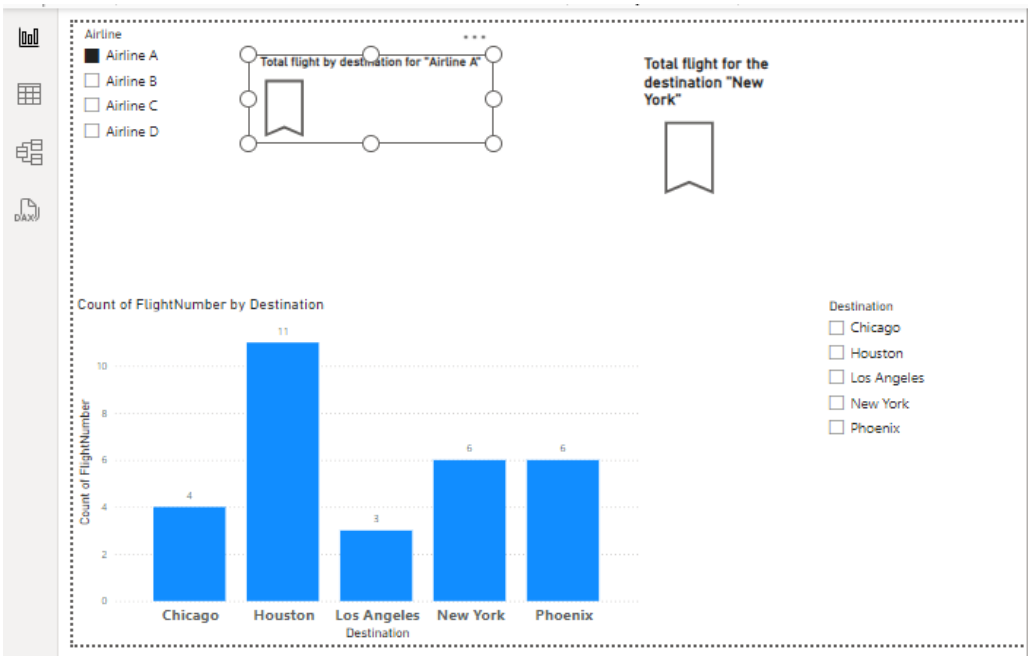


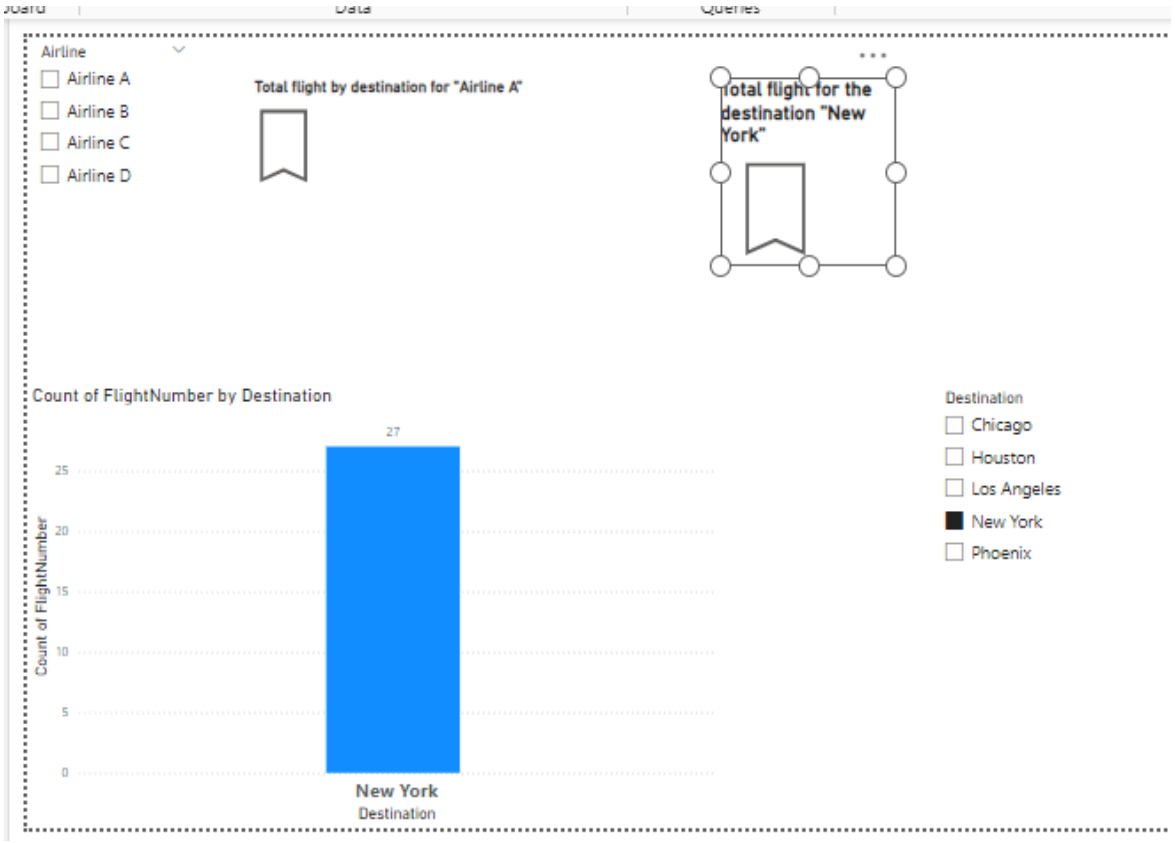
After that, go to the “View” tab and select “Bookmarks”. Click add to create bookmarks and making two individuals view:

- Total flight by destination for "Airline A".
- Total flight for the destination "New York"

Then add two individual “Bookmark” button from “Insert” ribbon and assign those two views in action field:

This view for “Airline A”

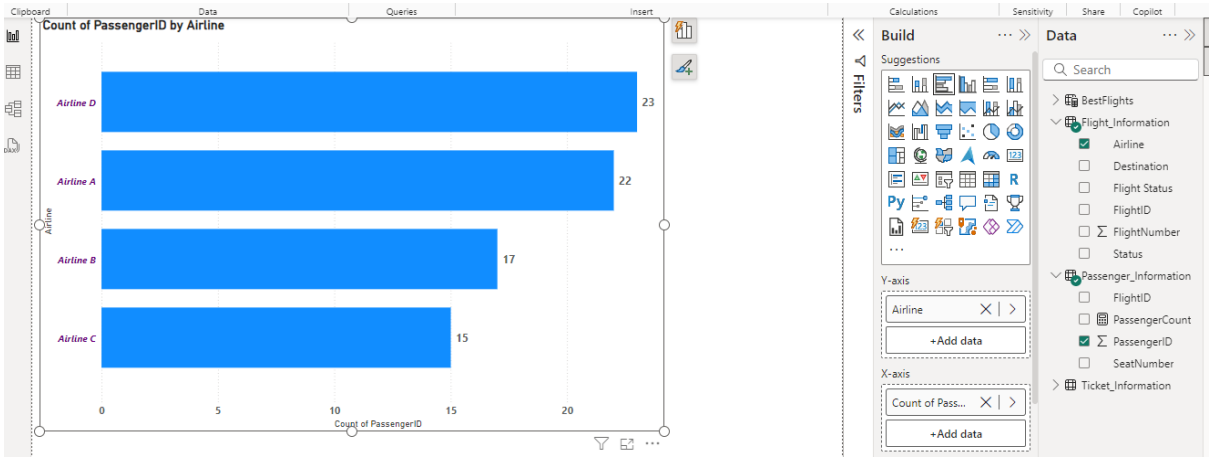




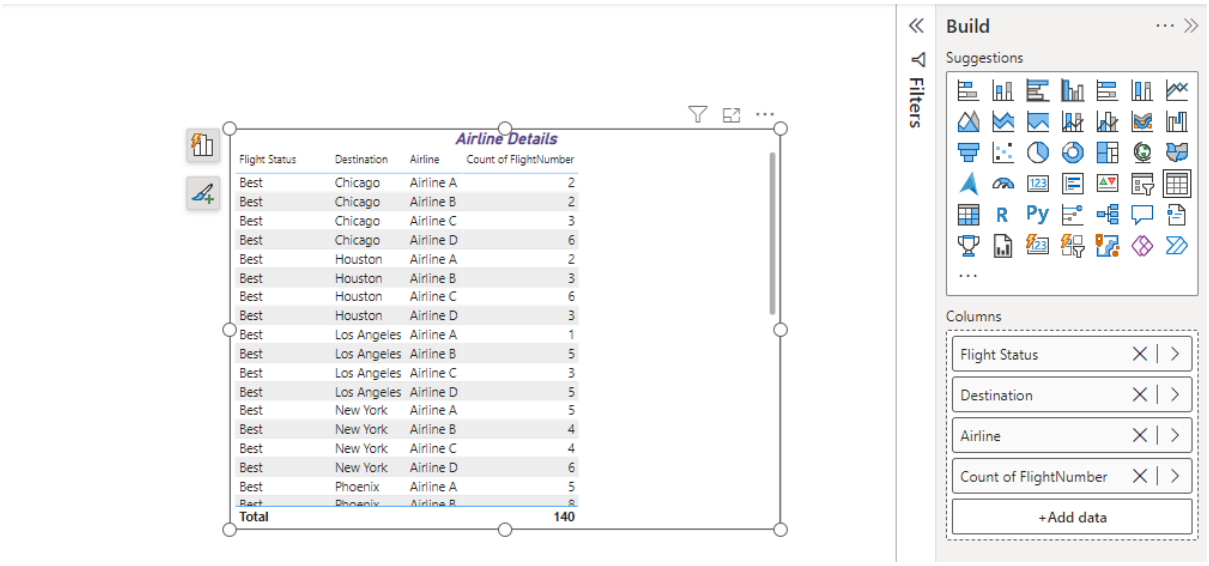
This view for “Total flight for the destination "New York"

For Airline-specific pages : Using “Drillthrough” pages, we can view detailed information related to that airline by right-click on an Airline.

At first create a new report page using “Clustered Bar Chart” and make a visual “Count of Passenger ID by Airline”.

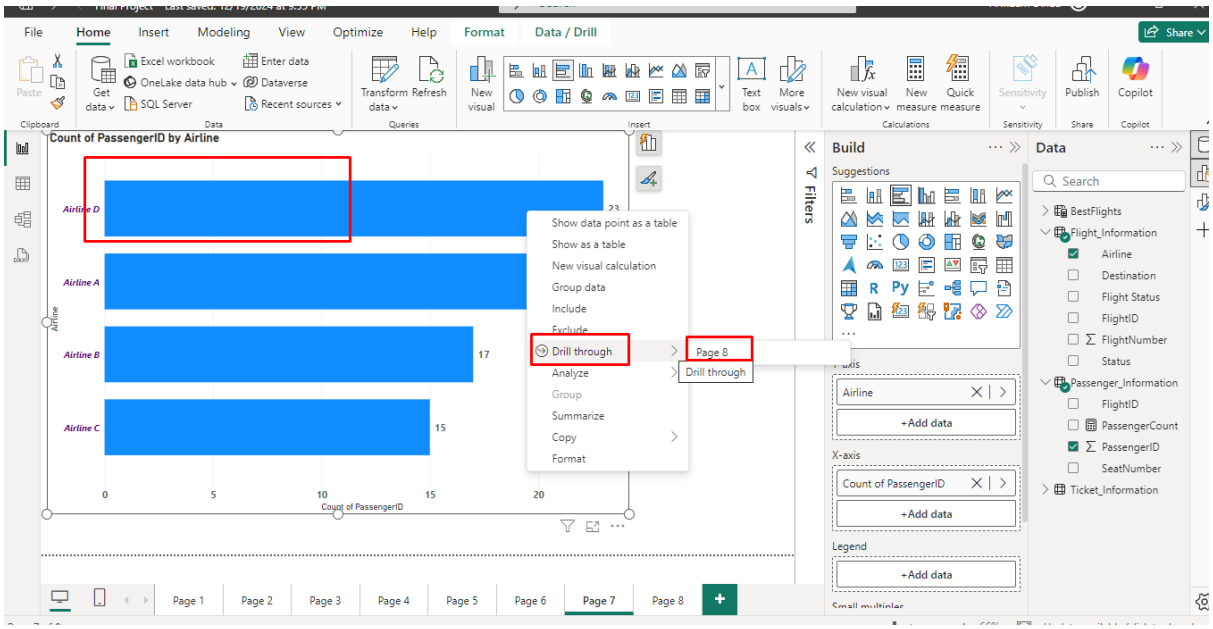


Then in another page, create a new visual table “Airline Details”

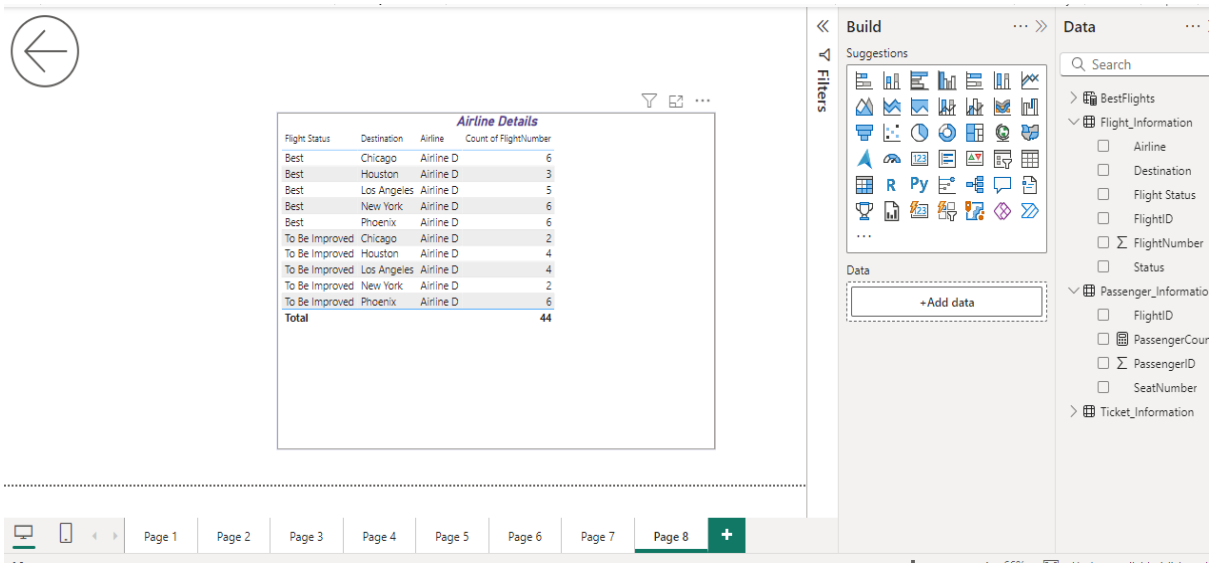


Then, Drag the “Airline” field from the “Flight_Information” table to the Drill-through section in the Fields pane. This will allow the drill-through feature to pass the selected Airline context from one visual to the drill-through page.

After that, in “Clustered Bar Chart”, right-click on “Airline D” chart, click “Drillthrough”> ”Page 8”



Now, the visuals will show data specific to the selected Airline.



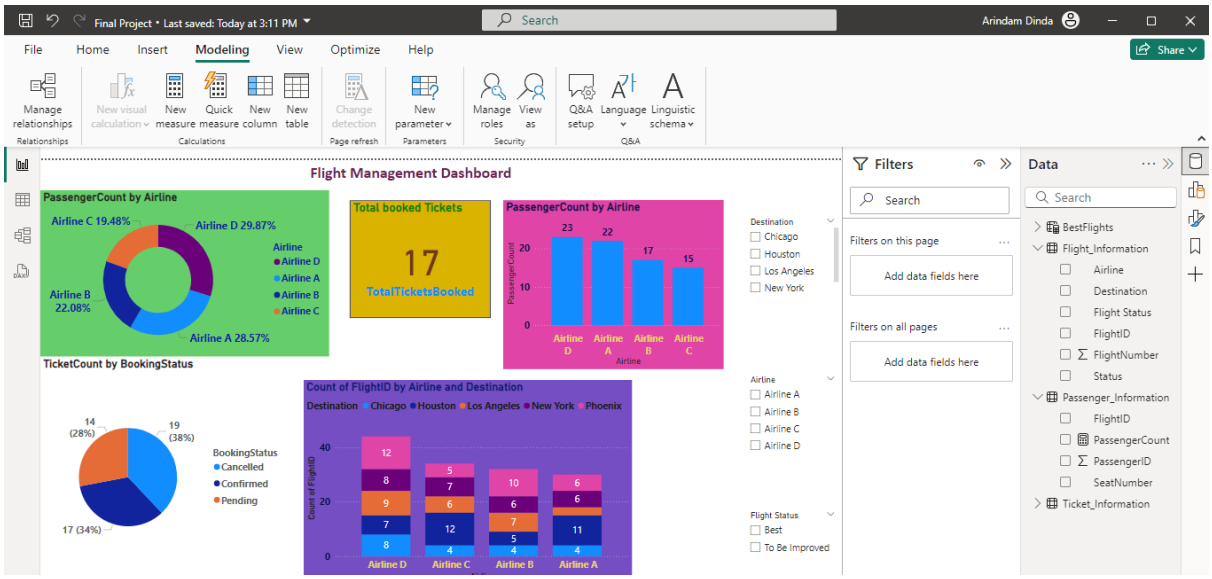
----------X----------

6. Final Dashboard and Power BI Service

- **Design a comprehensive dashboard with key visuals and insights.**
- **Configure Row-Level Security (RLS) for Airline A data and assign it to a user.**
- **Set up a schedule refresh at 5 PM daily.**

Ans:

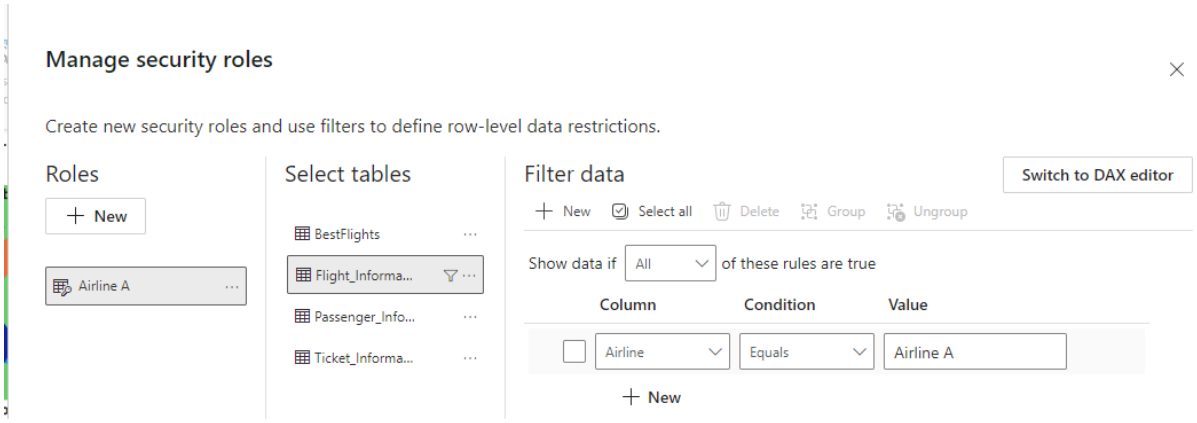
<u>Design a comprehensive dashboard with key visuals and insights</u>

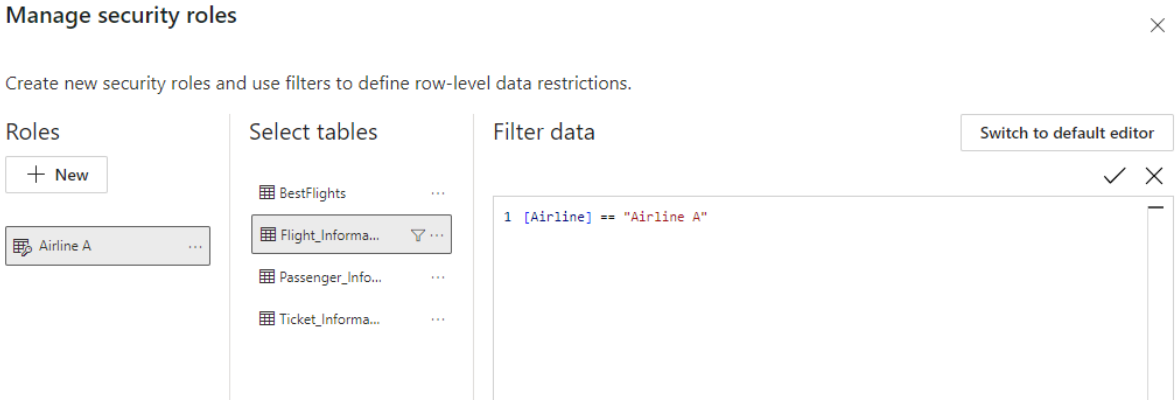


Configure Row-Level Security (RLS) for Airline A data and assign it to a user

For configuration, go to the “Modeling” tab in the ribbon, click on “Manage Roles”.

In the “Manage Roles” window, click Create and name the role as “Airline A”

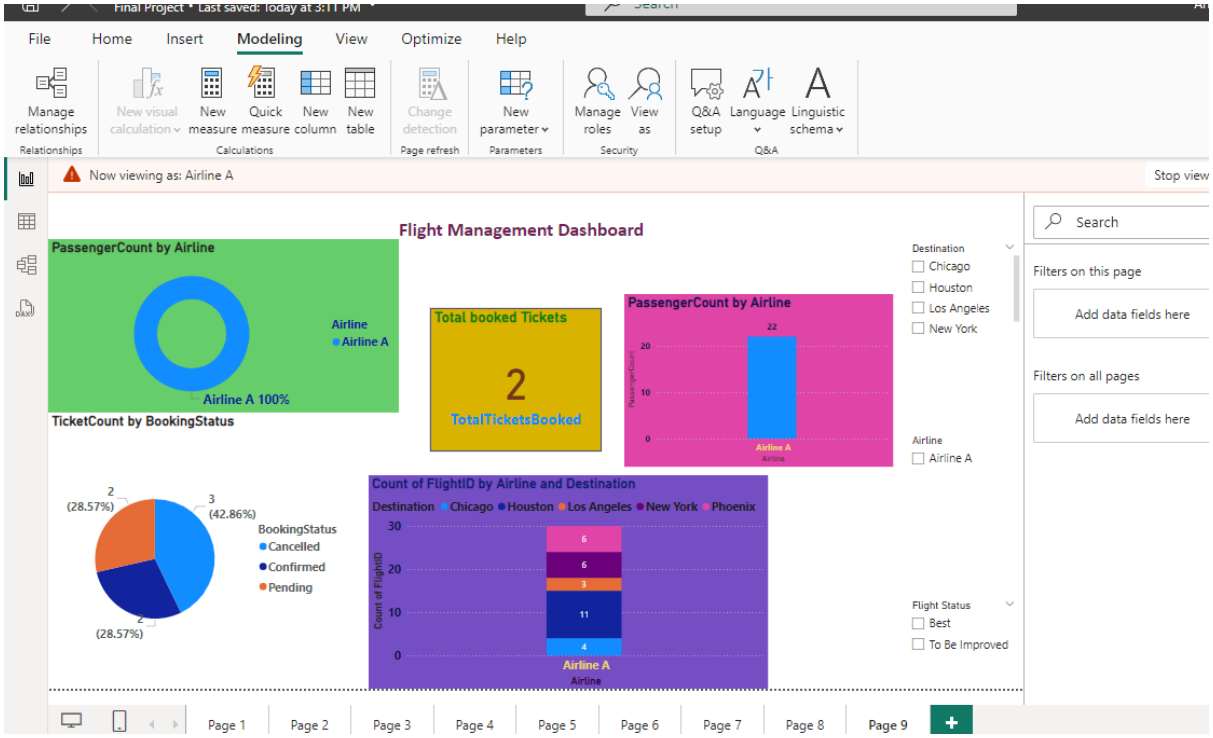




Then click “Save”.

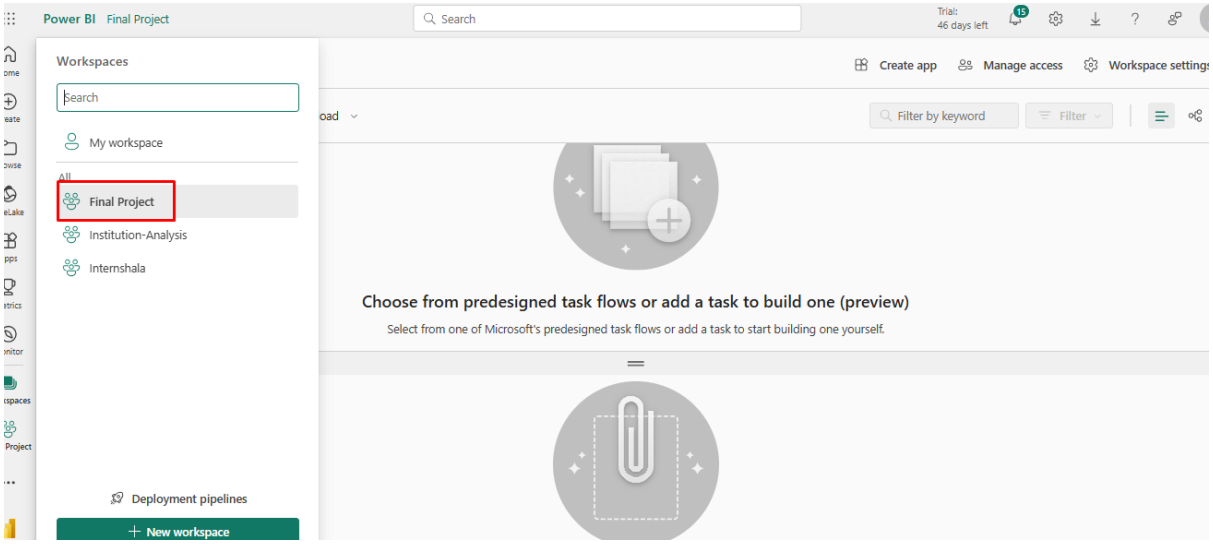
After that need to validating this “RLS Configuration”.

So, go to the “Modeling” tab and click “View as Roles” and select the role “Airline A”.



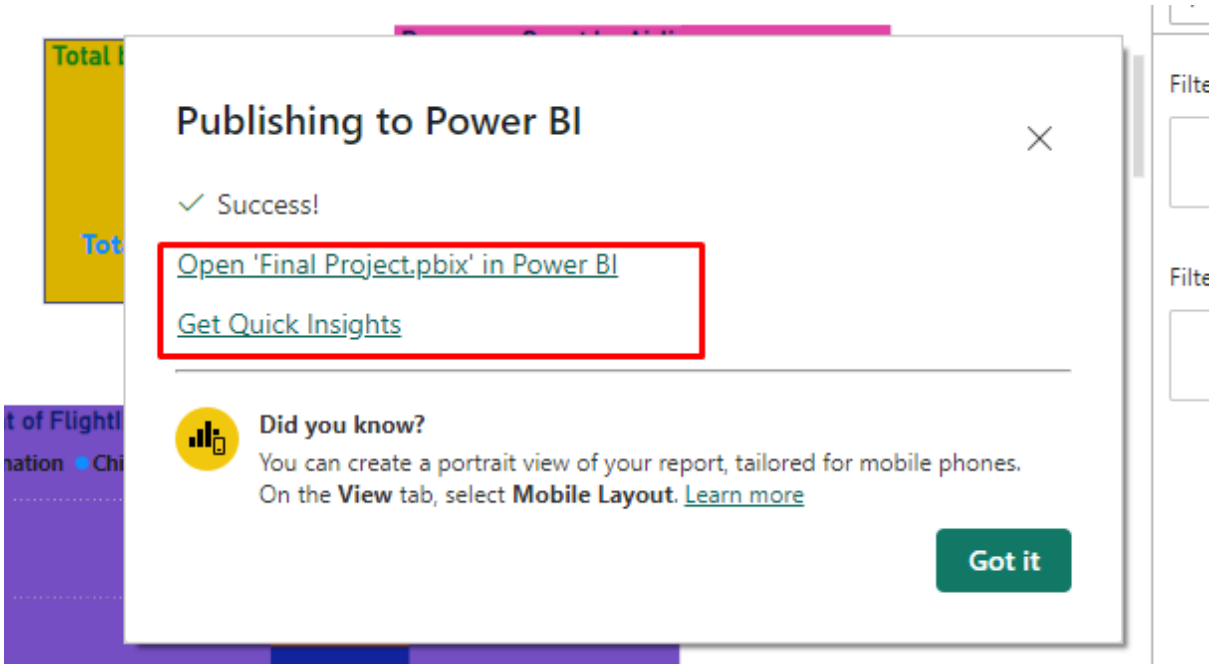
Now, all values showing as “Airline A” data.

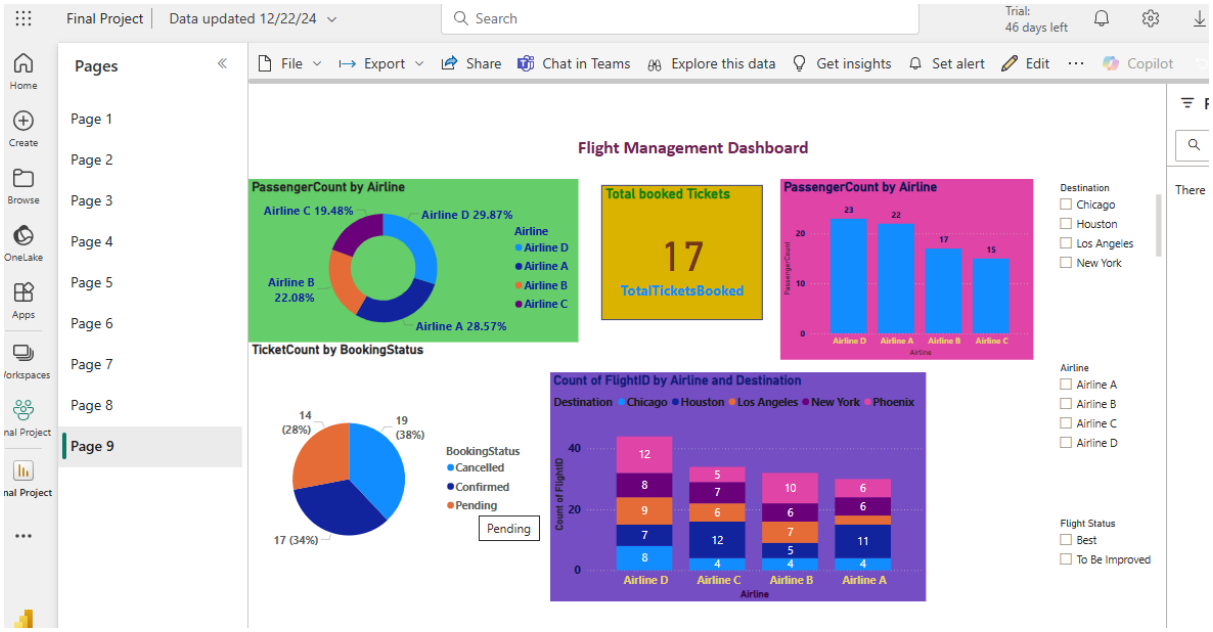
For assigning, at first create a workspace in “Power BI Service” and give the name as “Final Project”.

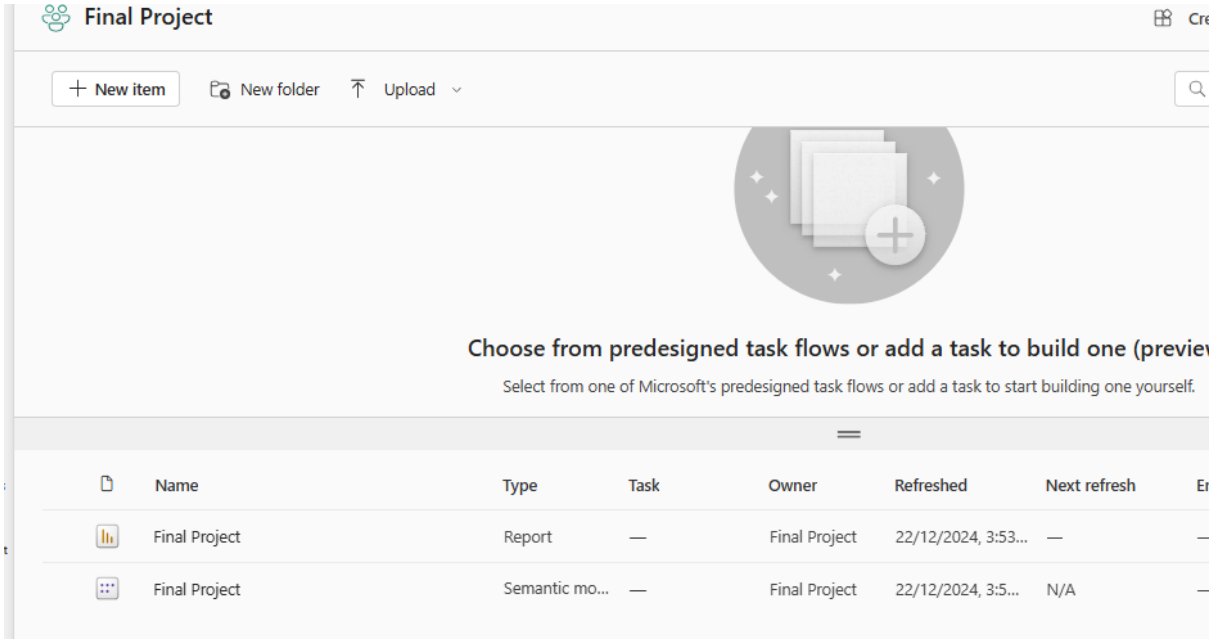


Then coming to the power BI desktop, in the “Home” ribbon, click on “Publish”, the select “Final Project” and click “Ok”.

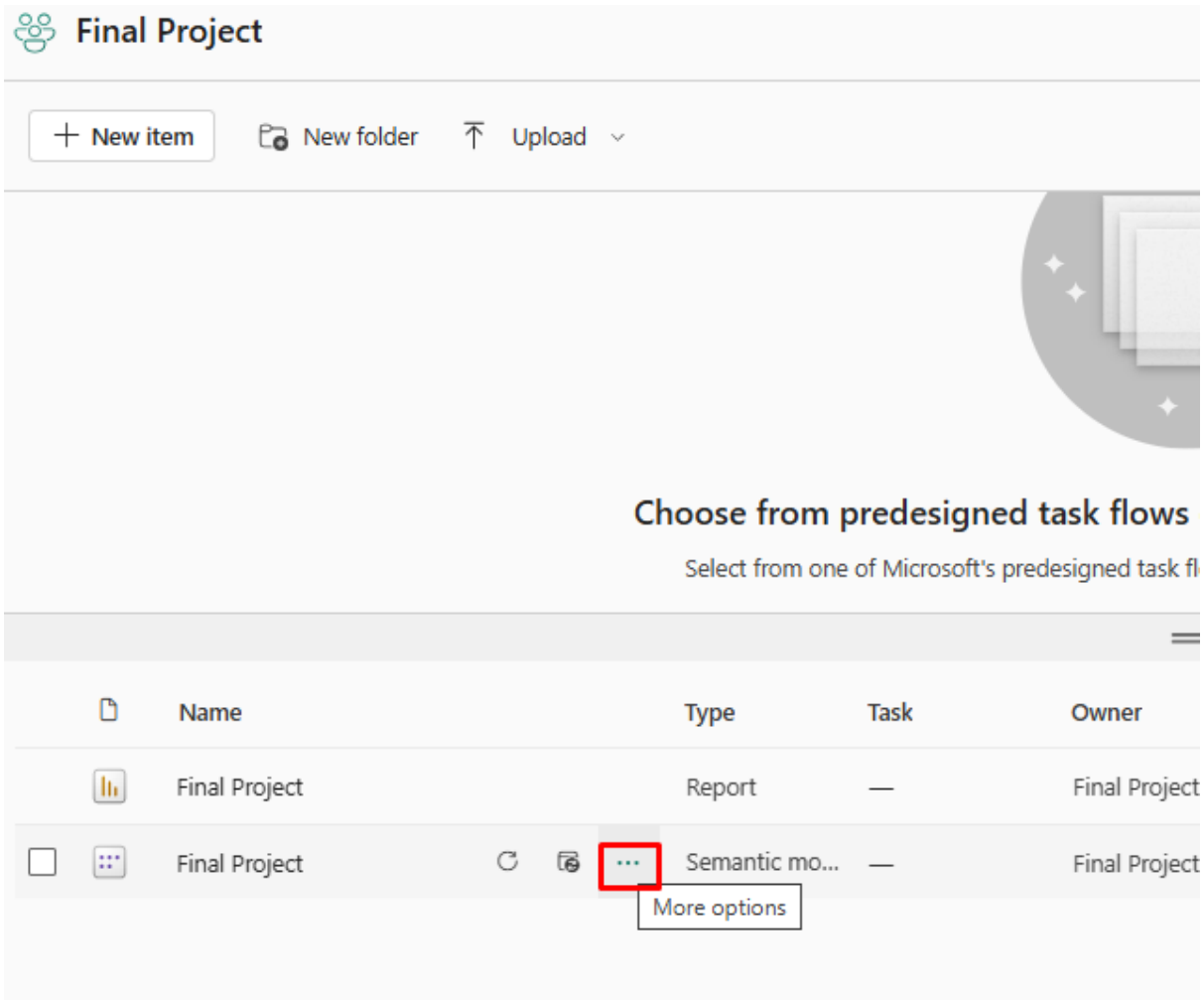
Then click on this marking link.



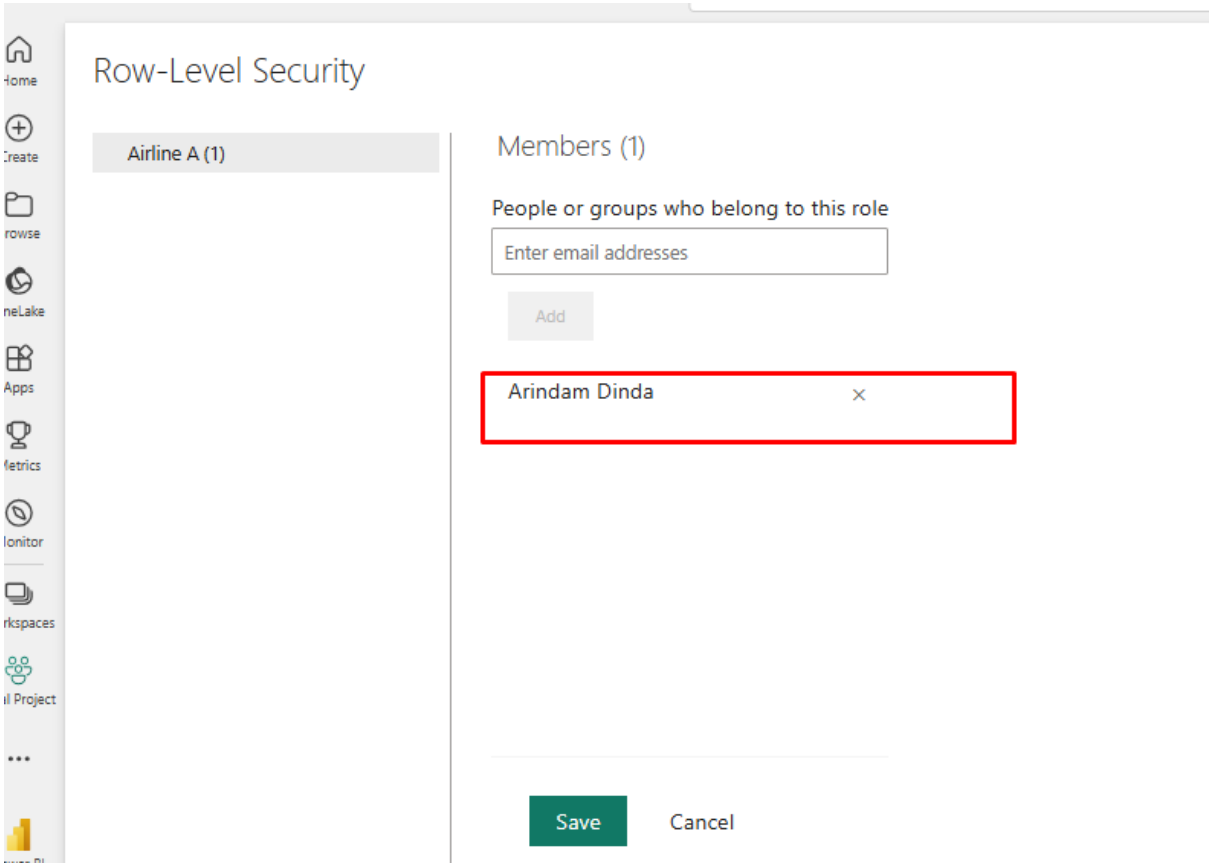




After that, click on three dot of the “Final Project” data set.

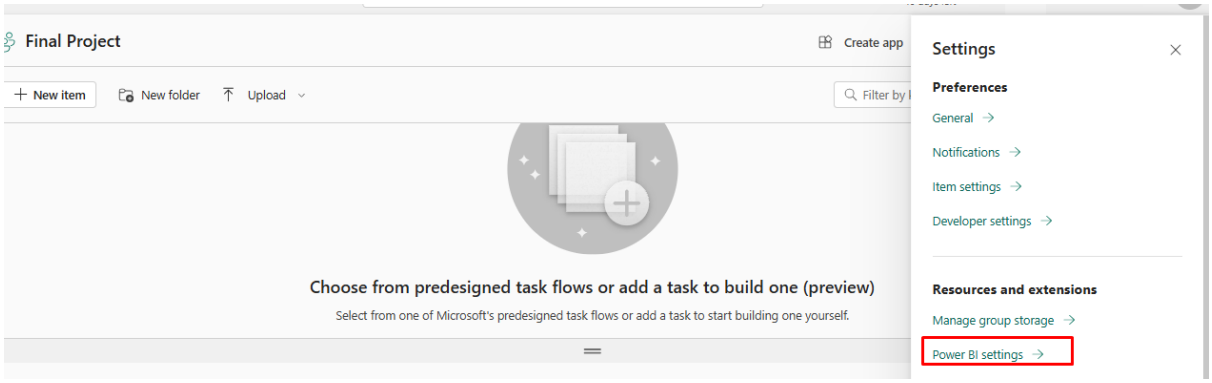


“row level security” dashboard will open. After that in the blank field “Member” enter any email id, which I want to assign. Then click “Add”. This is the way to assign any people.

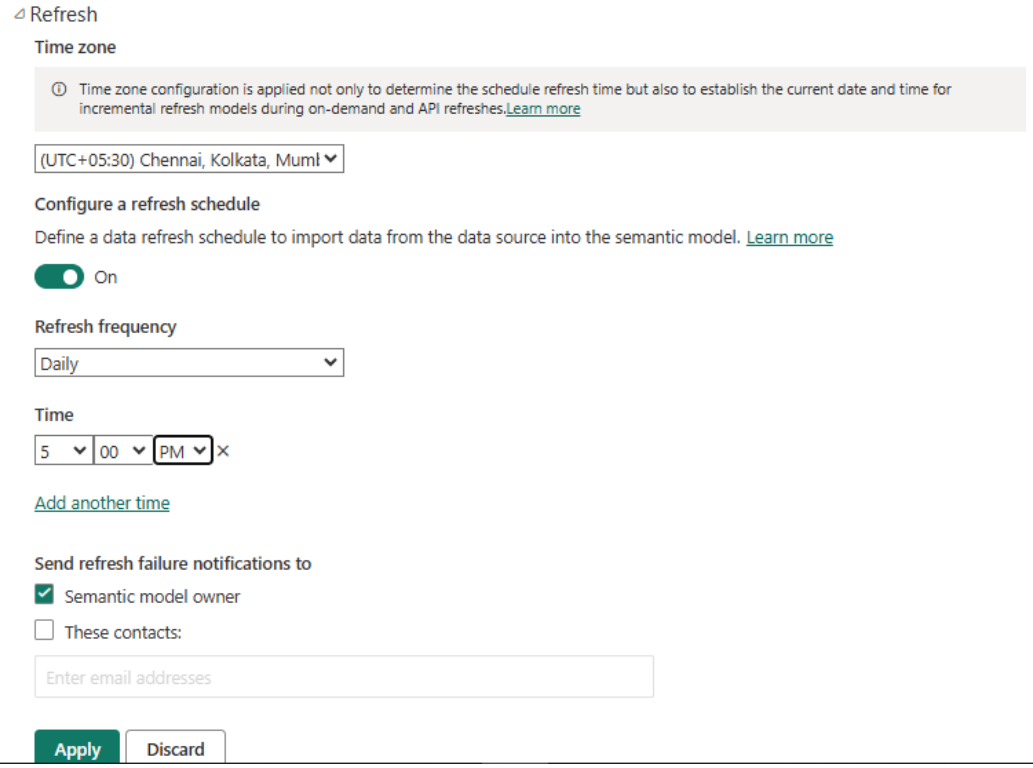


Set up a schedule refresh at 5 PM daily:

Go to the settings of “Power BI Service”, then click “Power BI setting”



Then click on “Semantic models” and scroll down, we can see “Refresh” option. After that set the time as per requirement and click on apply



---**End of the Project**---

-