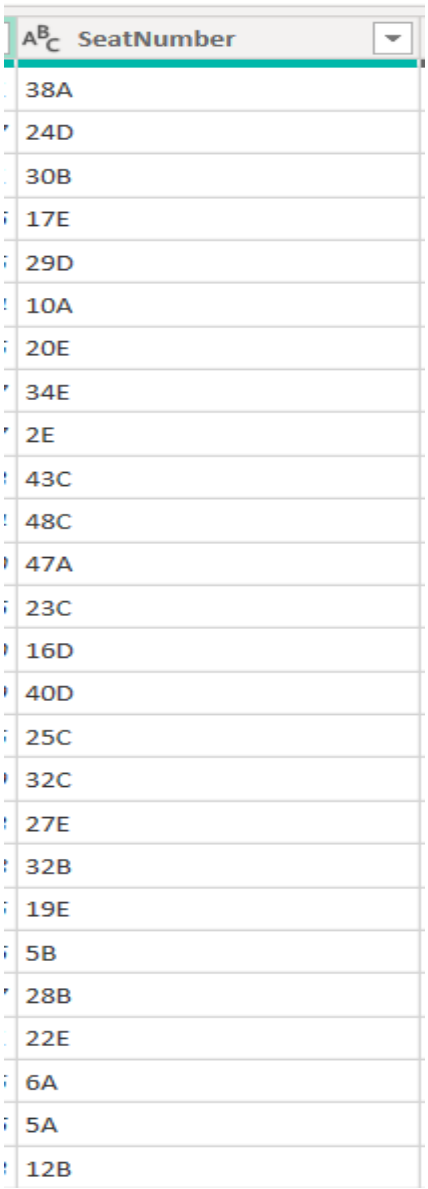
Airline Data Management and Analysis Using Power BI

1. Data Preparation and Cleaning

1. Remove Duplicates

- Step 1: Open Power BI Desktop.
- Step 2: Go to the Home tab, and click on Transform Data. This opens the Power Query Editor.
- Step 3: In the Power Query Editor, select the table you want to clean.
- Step 4: Select the columns where you want to check for duplicates.
- Step 5: On the ribbon, under the Transform tab, click Remove Duplicates. This will remove rows that are identical based on the selected columns.



SeatNumber
38A
24D
30B
17E
29D
10A
20E
34E
2E
43C
48C
47A
23C
16D
40D
25C
32C
27E
32B
19E
5B
28B
22E
6A
5A
12B

Handle Missing Values

- Step 1: In the Power Query Editor, select the column with missing values.
- Step 2: To remove rows with missing values:
    - Go to the Home tab, click Remove Rows and then select Remove Blank Rows.
- Step 3: To replace missing values with a specific value
    - Right-click on the column header.
    - Select Replace Values.

FlightID	FlightNumber	Airline	Destination	Status
1001	FL1102	Airline D	Houston	On Time
1002	FL1435	Airline B	Chicago	On Time
1003	FL1860	Airline A	New York	Cancelled
1004	FL1270	Airline C	Chicago	Delayed
1005	FL1106	Airline C	New York	Delayed
1006	FL1071	Airline A	Phoenix	On Time
1007	FL1700	Airline C	Los Angeles	Cancelled
1008	FL1020	Airline C	Los Angeles	Delayed
1009	FL1614	Airline A	Los Angeles	Cancelled
1010	FL1121	Airline D	Chicago	Cancelled
1011	FL1466	Airline A	Phoenix	On Time
1012	FL1214	Airline D	New York	Delayed
1013	FL1330	Airline C	Houston	On Time
1014	FL1458	Airline C	New York	Delayed
1015	FL1087	Airline C	Houston	Delayed
1016	FL1372	Airline B	New York	Delayed
1017	FL1099	Airline D	Phoenix	Delayed
1018	FL1871	Airline B	Houston	Delayed
1019	FL1663	Airline B	Chicago	Cancelled
1020	FL1130	Airline A	New York	On Time
1021	FL1661	Airline B	New York	Cancelled
1022	FL1308	Airline A	Houston	Delayed
1023	FL1769	Airline A	Chicago	On Time
1024	FL1343	Airline B	Chicago	Delayed
1025	FL1491	Airline D	Phoenix	On Time
1026	FL1413	Airline D	Chicago	Cancelled
1027	FL1805	Airline D	Chicago	On Time
1028	FL1385	Airline D	Chicago	On Time
1029	FL1191	Airline D	Los Angeles	On Time
1030	FL1955	Airline B	Phoenix	On Time
1031	FL1276	Airline B	New York	On Time

PassengerID	FlightID	SeatNumber
1	1161	38A
2	1157	24D
3	1141	30B
4	1046	17E
5	1035	29D
6	1134	10A
8	1115	20E
9	1197	34E
10	1047	2E
11	1153	43C
12	1194	48C
13	1010	47A
14	1056	23C
15	1030	16D
16	1109	40D
17	1005	25C
18	1119	32C
19	1033	27E
20	1118	32B
21	1065	19E
22	1146	5B
23	1177	28B
24	1011	22E
25	1085	6A
26	1026	5A
27	1063	12B

TicketID	FlightID	BookingStatus
5001	1178	Pending
5002	1078	Confirmed
5003	1117	Cancelled
5004	1120	Cancelled
5005	1137	Cancelled
5006	1162	Pending
5007	1076	Pending
5008	1035	Cancelled
5009	1001	Cancelled
5010	1040	Cancelled
5011	1064	Pending
5012	1150	Cancelled
5013	1060	Cancelled
5014	1064	Confirmed
5015	1093	Confirmed
5016	1072	Pending
5017	1011	Cancelled
5018	1105	Cancelled
5019	1014	Confirmed
5020	1060	Pending
5021	1030	Confirmed
5022	1035	Confirmed
5023	1165	Confirmed
5024	1005	Confirmed
5025	1083	Cancelled

Format Columns

- Step 1: In the Power Query Editor, select the column you want to format.
- Step 2: To change the data type:
    - Right-click the column header and select Change Type.
    - Choose the appropriate data type (ex: →, text, number, date, etc.).
- Step 3: For text formatting
    - Use the Transform tab options like:
        - Format > Uppercase or Lowercase.
        - Trim to remove spaces.
        - Clean to remove non-printing characters.

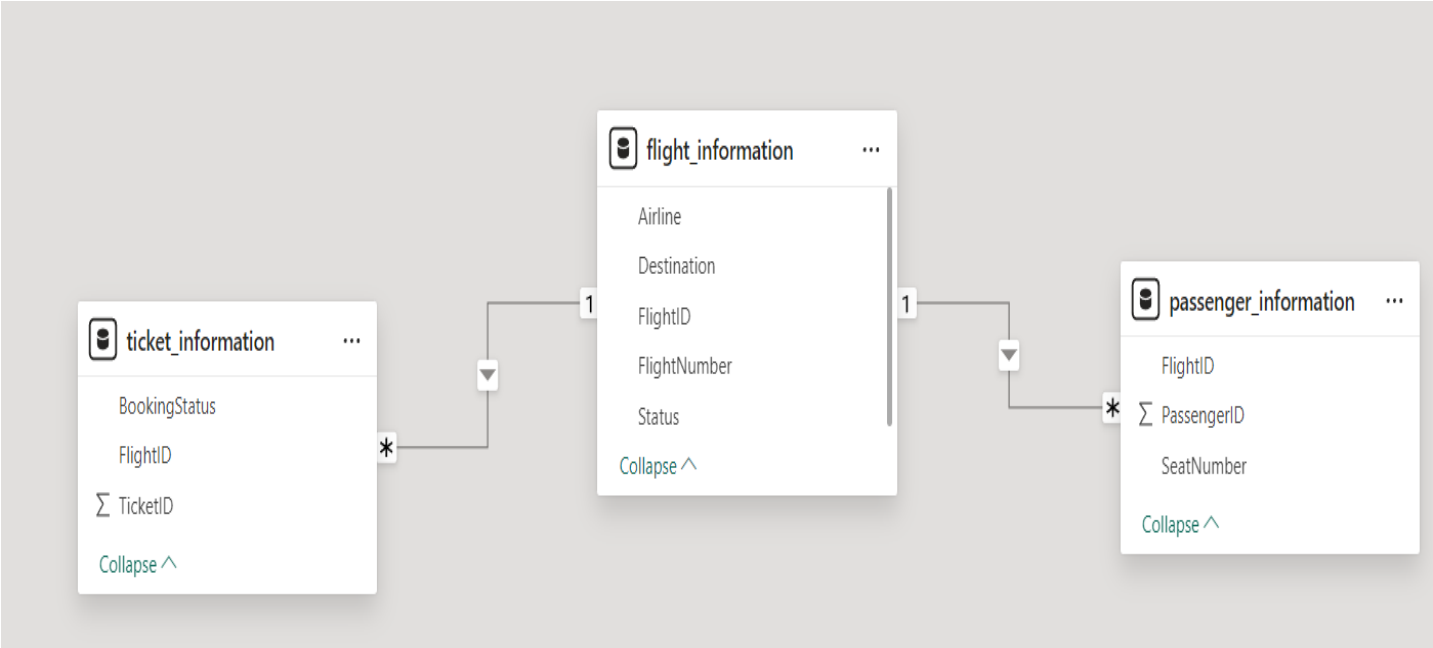
TicketID	FlightID	BookingStatus
5001	1178	Pending
5002	1078	Confirmed
5003	1117	Cancelled
5004	1120	Cancelled
5005	1137	Cancelled
5006	1162	Pending
5007	1076	Pending
5008	1035	Cancelled
5009	1001	Cancelled
5010	1040	Cancelled
5011	1064	Pending

TicketID	FlightID	BookingStatus
5001	1178	Pending
5002	1078	Confirmed
5003	1117	Cancelled
5004	1120	Cancelled
5005	1137	Cancelled
5006	1162	Pending
5007	1076	Pending
5008	1035	Cancelled
5009	1001	Cancelled

PassengerID	FlightID	SeatNumber
1	1161	38A
2	1157	24D
3	1141	30B
4	1046	17E
5	1035	29D
6	1134	10A
8	1115	20E
9	1197	34E
10	1047	2E
11	1153	43C

2. Create Relationships Between Tables

- Step 1: In the Relationships View, you'll need to connect the tables that have a FlightID column.
- Step 2: Drag the FlightID column from the first table and drop it onto the FlightID column in the second table. This will automatically create a relationship.
- Step 3: Do the same for the third table. Drag the FlightID column from the Flights table and drop it onto the FlightID column of the third table.



Power BI should automatically detect the relationship type (One-to-Many), assuming that the FlightID is unique in the Flights table.

3. Enhanced Data Insights

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

Step 1: In the Power Query Editor, go to the "Add Column" tab and click "Conditional Column".

Step 2: Name the Column: Give it a meaningful name, e.g., Flight Status Classification.

Step 3: Condition: Set up your conditions based on the status of the flight. For example:

- If the Status is "On Time," then the result will be "Best".
- If the Status is "Delayed," then the result will be "To Be Improved".

Step 4: Your setup might look like this:

- If: Status
- equals: On Time
- then: Best
- else: To Be Improved

Click "OK" to apply the conditional column

Status	Flight_status_classification
On Time	Best
On Time	Best
Cancelled	To be improved
Delayed	To be improved
Delayed	To be improved
On Time	Best
Cancelled	To be improved
Delayed	To be improved
Cancelled	To be improved
Cancelled	To be improved
On Time	Best
Delayed	To be improved
On Time	Best
Delayed	To be improved
Delayed	To be improved
Delayed	To be improved
Delayed	To be improved
Delayed	To be improved
Cancelled	To be improved
On Time	Best
Cancelled	To be improved
Delayed	To be improved
On Time	Best
Delayed	To be improved
On Time	Best
Cancelled	To be improved
On Time	Best
On Time	Best
On Time	Best
On Time	Best

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

Step 1: In the "Add Column" tab, click "Column from Examples" and select "From All Columns".

This allows you to provide an example of the desired output, and Power BI will try to extract it automatically.

Step 2: In the first row of the new column that appears, type the Flight Number manually based on an example from the FlightNumber column. For example:

- If the FlightNumber is "FL1102" in the original column, type 1102 in the new column.

Step 3: Power BI will try to extract a pattern and auto-complete the rest of the column based on your example.

Step 4: If the correct flight number is extracted, you can confirm by clicking "OK".

FlightNumber	Airline	Destination	Status	Flight_status_classification	flight number
FL1102	Airline D	Houston	On Time	Best	1102
FL1435	Airline B	Chicago	On Time	Best	1435
FL1860	Airline A	New York	Cancelled	To be improved	1860
FL1270	Airline C	Chicago	Delayed	To be improved	1270
FL1106	Airline C	New York	Delayed	To be improved	1106
FL1071	Airline A	Phoenix	On Time	Best	1071
FL1700	Airline C	Los Angeles	Cancelled	To be improved	1700
FL1020	Airline C	Los Angeles	Delayed	To be improved	1020
FL1614	Airline A	Los Angeles	Cancelled	To be improved	1614
FL1121	Airline D	Chicago	Cancelled	To be improved	1121
FL1466	Airline A	Phoenix	On Time	Best	1466
FL1214	Airline D	New York	Delayed	To be improved	1214
FL1330	Airline C	Houston	On Time	Best	1330
FL1458	Airline C	New York	Delayed	To be improved	1458
FL1087	Airline C	Houston	Delayed	To be improved	1087
FL1372	Airline B	New York	Delayed	To be improved	1372
FL1099	Airline D	Phoenix	Delayed	To be improved	1099
FL1871	Airline B	Houston	Delayed	To be improved	1871
FL1663	Airline B	Chicago	Cancelled	To be improved	1663
FL1130	Airline A	New York	On Time	Best	1130
FL1661	Airline B	New York	Cancelled	To be improved	1661
FL1308	Airline A	Houston	Delayed	To be improved	1308

4. Calculations Using DAX

1)Total passengers for a specific flight.

Step1: Go to Table View Select the passenger_information Dataset

Step2: Click on new measure

In DAX formula bar

```
Total Passengers for Specific Flight =
DISTINCTCOUNT(passenger_information[PassengerID])
```

Step3: Go to Report View

In Report view go to Visualization Pane and click on table visuals and drag the Measure to column & FlightID.

FlightID	Total Passengers for Specific Flight
1001	1
1003	2
1005	1
1006	1
1010	1
1011	1
1012	1
1026	1
1027	1
1030	1
1032	1
1033	1
1034	1
1035	1
1038	1
1039	1
1046	1
1047	1
1050	2
1051	1
1052	1
1053	1
1055	1
1056	1
1057	1
1059	1
1063	1
1065	1
1066	1
1072	1
Total	**79**

2)Total tickets booked

Step1: Go to Table View Select the ticket_information Dataset

Step2: Click on new measure

In DAX formula bar

```
Total Tickets Booked =
CALCULATE(
    COUNTROWS(ticket_information),
    ticket_information[BookingStatus] = "Confirmed"
)
```

Step3: Go to Report View

In Report view go to Visualization Pane and click on table visuals and drag the Measure to column & FlightID.

Total Tickets Booked	FlightID
1	1005
1	1014
1	1023
1	1030
1	1035
1	1040
1	1042
1	1064
2	1078
1	1093
1	1097
1	1104
1	1143
1	1165
1	1166
1	1179
17	

3)Filtered table showing "Best" flights only

Step1: Go to Table View Select the flight_information Dataset

Step2: Click on new measure

In DAX formula bar

```
1  Best Flight = FILTER(flight_information,flight_information[Status]="On Time")
```

Step3: Go to Report View

In Report view go to Visualization Pane and click on table visuals and drag the Measure to column & FlightID.

Best	1001
Best	1002
Best	1006
Best	1011
Best	1013
Best	1020
Best	1023
Best	1025
Best	1027
Best	1028
Best	1029
Best	1030
Best	1031
Best	1033
Best	1034
Best	1036
Best	1039
Best	1043
Best	1044
Best	1046
Best	1048
Best	1050
Best	1052
Best	1053
Best	1055
Best	1057
Best	1060
Best	1061
Best	1062
Best	1064

5. Visualization and Interactive Features

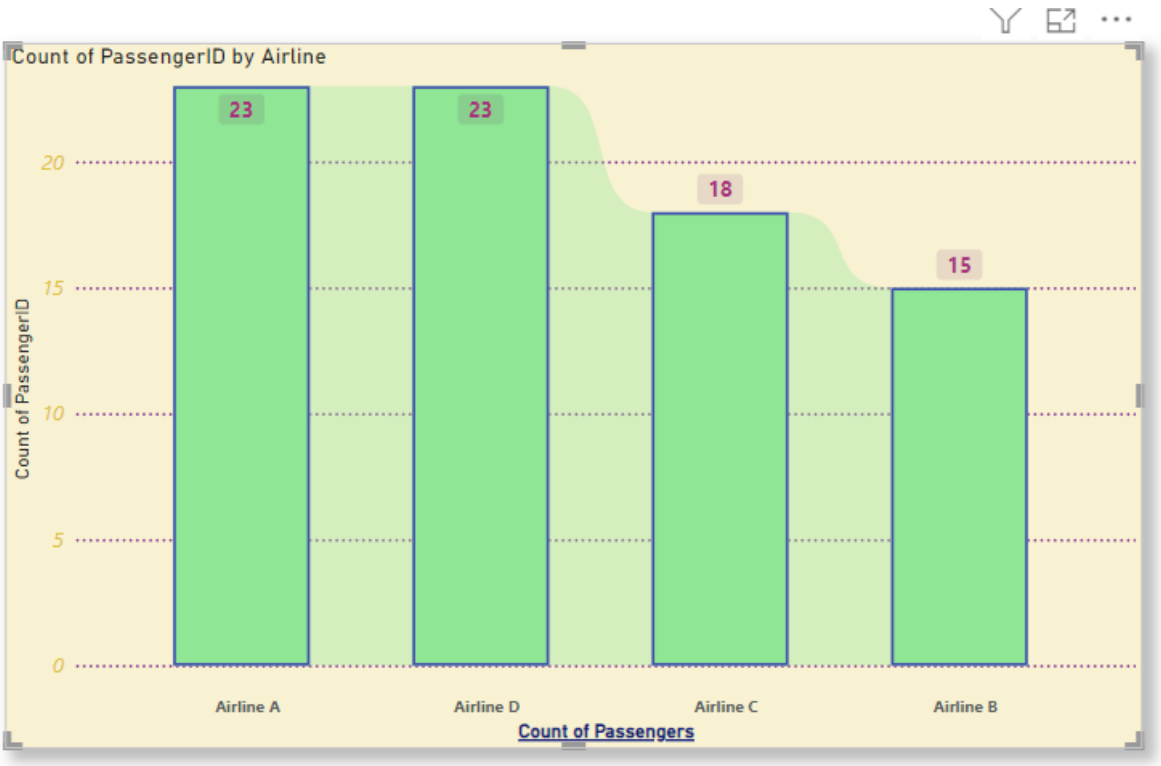
● Create visuals for:

1) Passenger count by airline.

Step 1: In report view go to visualization pane

Step 2: In visualization pane select Stacked Column Chart in x – axis Airline and in y – axis count of PassengerID

Step 3: Go to format pane and format our chart

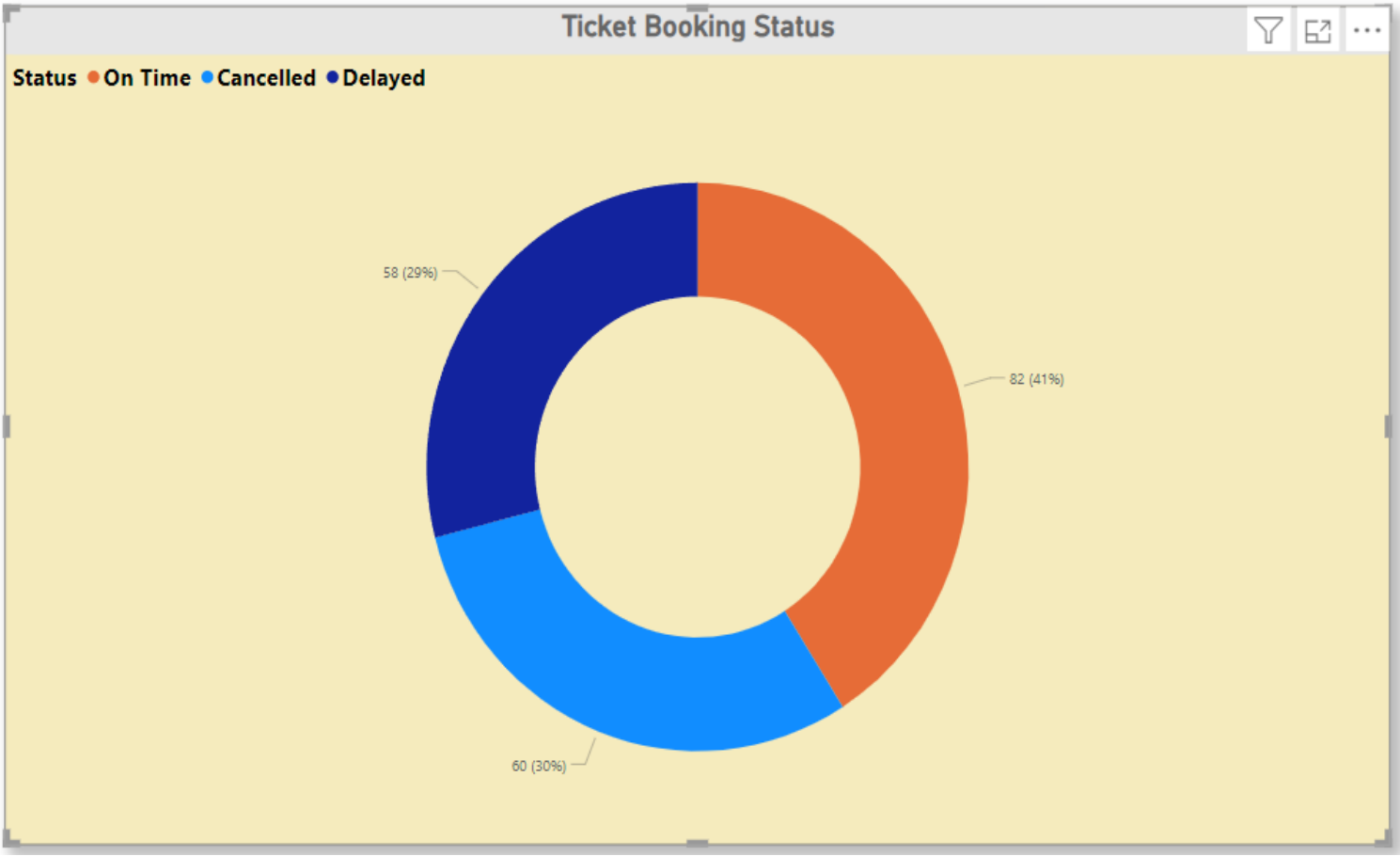


2)Ticket booking statuses.

Step 1: In report view go to visualization pane

Step 2: In visualization pane select Donut Chart in legend area Status and Values are Count Of Status.

Step 3: Go to format pane and format our chart

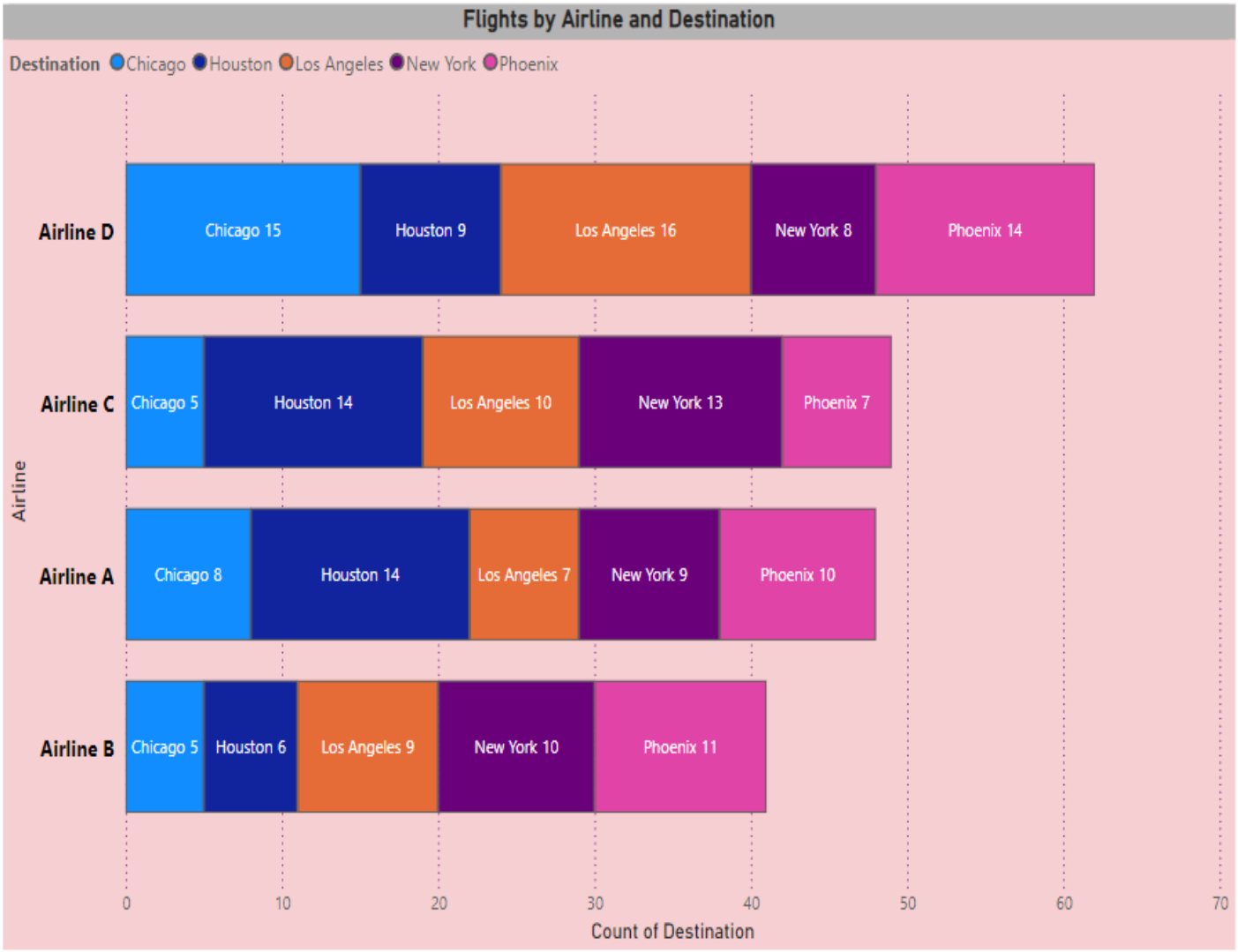


3)Flights by airline and destination.

Step 1: In report view go to visualization pane

Step 2: In visualization pane select Stacked Bar Chart in in Y –axis Airline in X – axis Count Of Destination in Legend area Destination.

Step 3: Go to format pane and format our chart



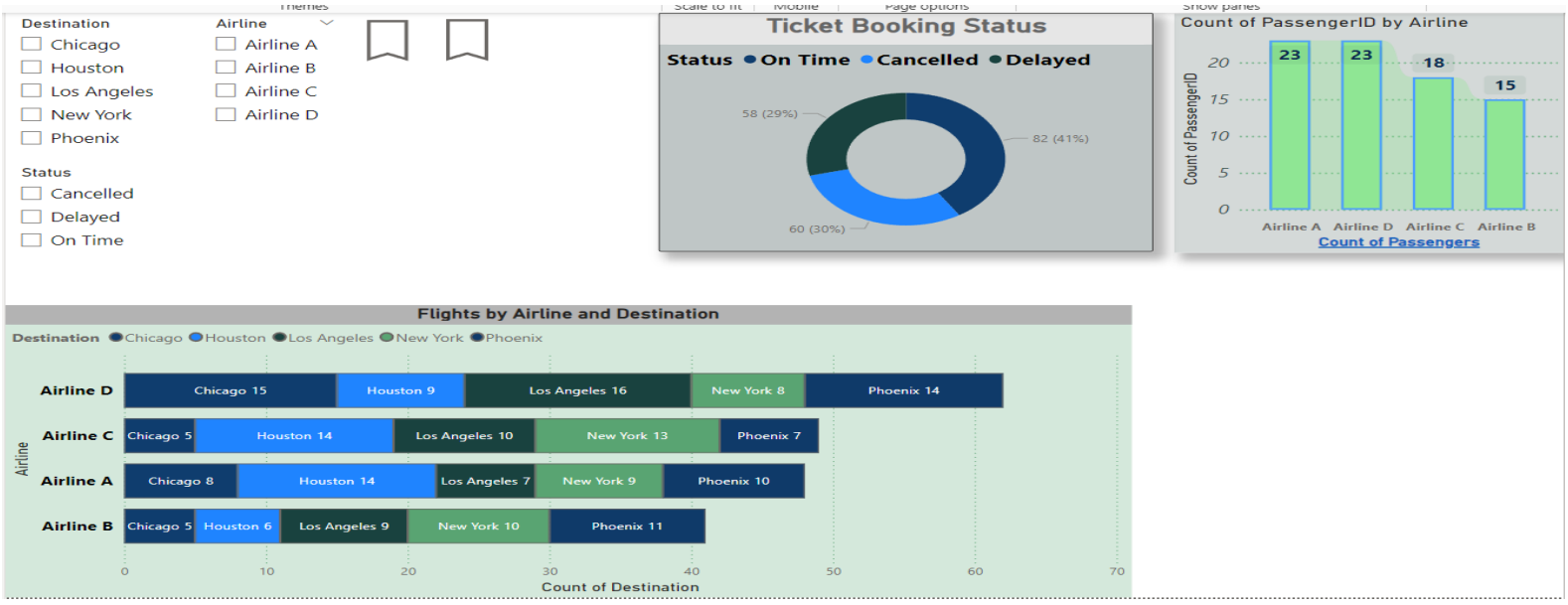
● Add interactive features for:

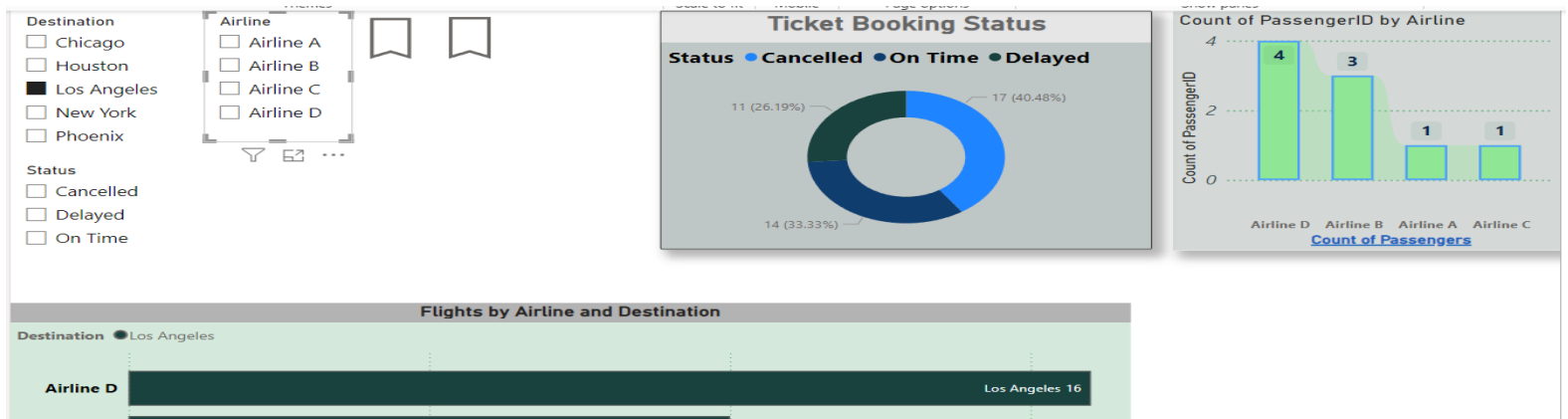
○ Destination and Airline.

Steps:

- Using Slicers for Destination and Airline:

1. Add Slicers:
   - From the Visualizations pane, select the Slicer icon.
   - Drag the Destination field into the Field section of the slicer. This will create a slicer for destinations.
   - Add another slicer and drag the Airline Name field into the Field section of this slicer. This will create a slicer for airlines.
2. Position Slicers:
   - Arrange these slicers in a logical place on your report canvas.
3. Interactivity:
   - When we select a destination or an airline from the slicer, all the other visuals on the report page will update dynamically to reflect the selected values.



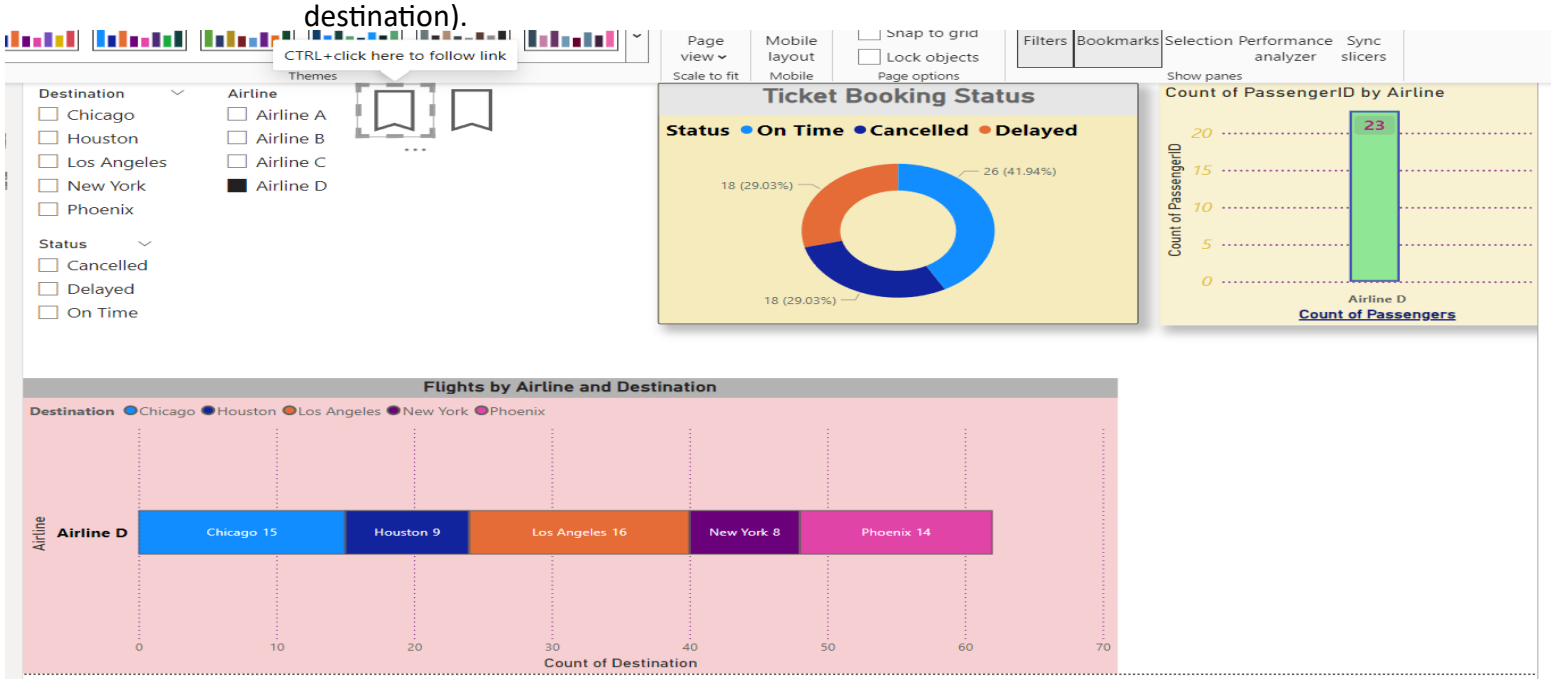


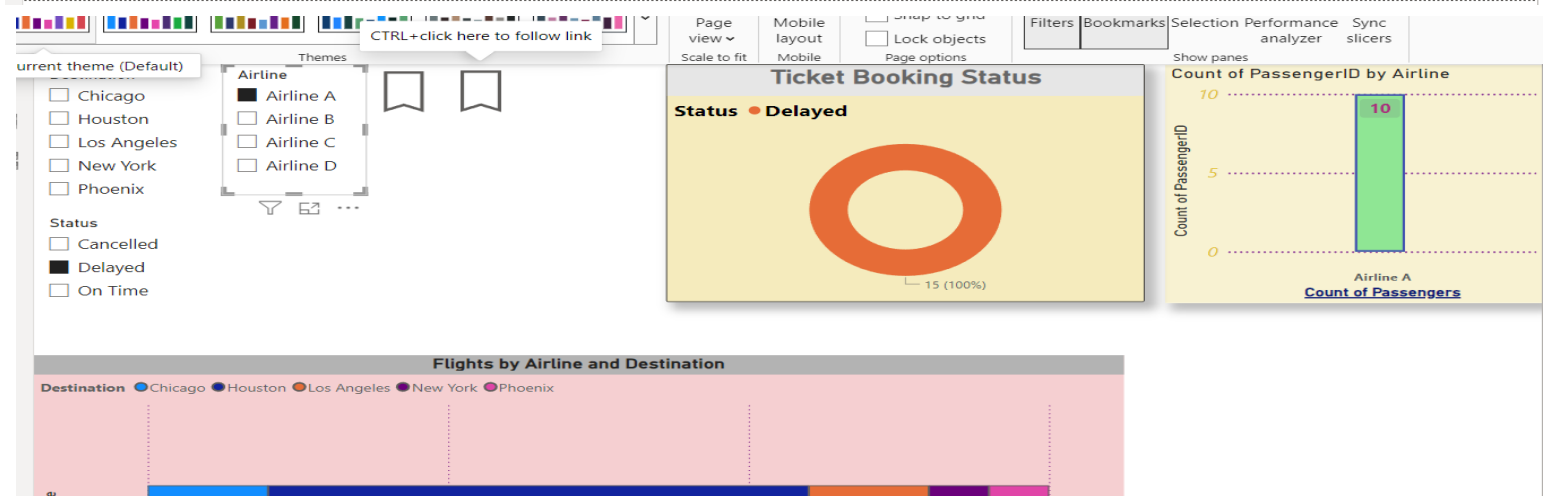
o Quick views.

Steps:

- Creating Bookmarks for Quick Views:

1. Set Up Views:
   - Arrange your visuals on the report canvas to create different views (e.g., one with all data, another showing only flight count by airline, etc.).
2. Create Bookmarks:
   - In the View tab on the top ribbon, select Bookmarks Pane.
   - Click on Add to create a new bookmark.
   - Name the bookmark.
   - For each view you want to create, click on the Add button and name it accordingly.
   - After selecting a visual state (for example, filtering to show only one airline or destination).





○ Airline-specific pages

Add Airline-Specific Filters:

- On each new page, add visuals that focus on the specific airline.
- Add a Slicer for Airline Name and set it to filter by the airline corresponding to that page (e.g., on the "Airline A" page, set the slicer to filter only Airline A).
- Alternatively, you can use Visual Level Filters to limit the data to a specific airline on each page.

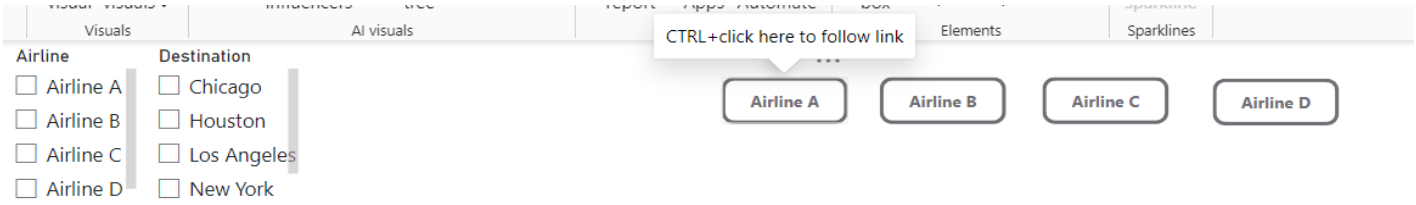
Create Navigation Buttons:

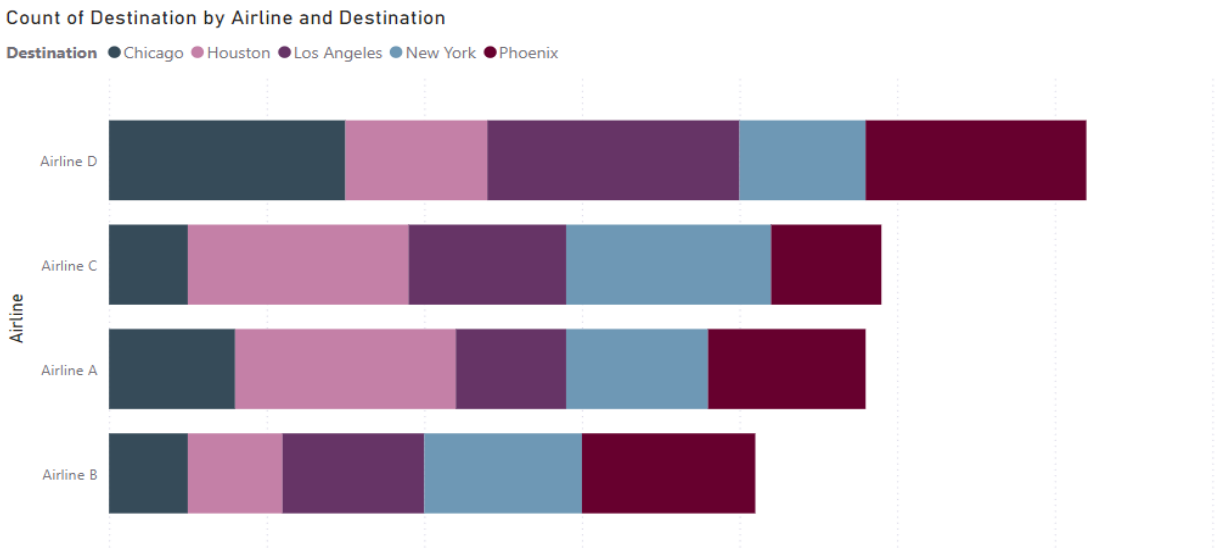
- On each page including the main page, add Buttons from the Insert tab.
- In the Action settings of the button, set the Type to Page Navigation.
- Assign the button to navigate to the specific airline page, on the main page, add a button for each airline that will take to that airline's page.

Add Home Button for Navigation:

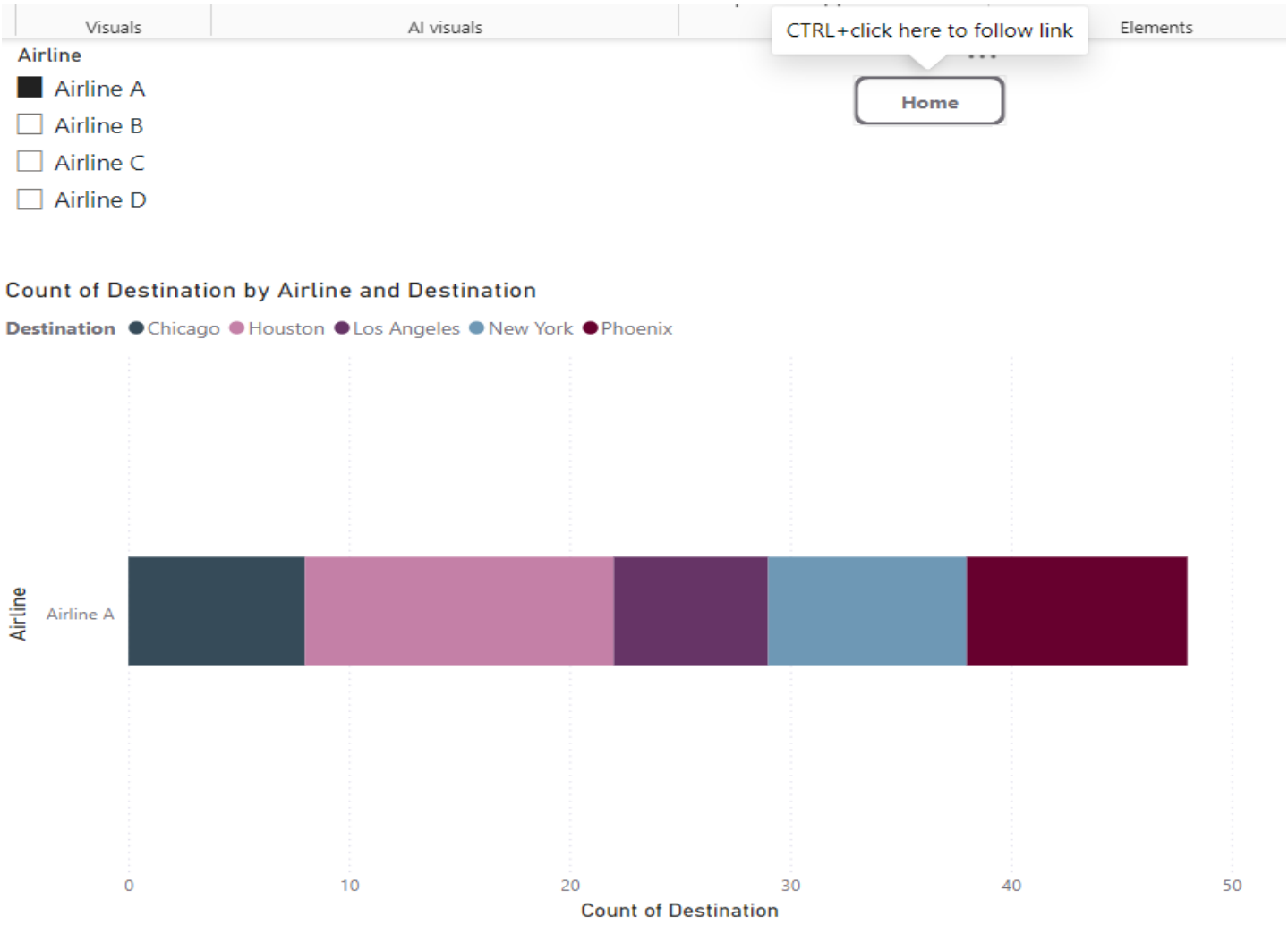
- To allow users to return to the main page, add a Home Button that navigates back to the main dashboard.
- This can be done using a Button with an Action type of Page Navigation, pointing to the main page.

Home Page

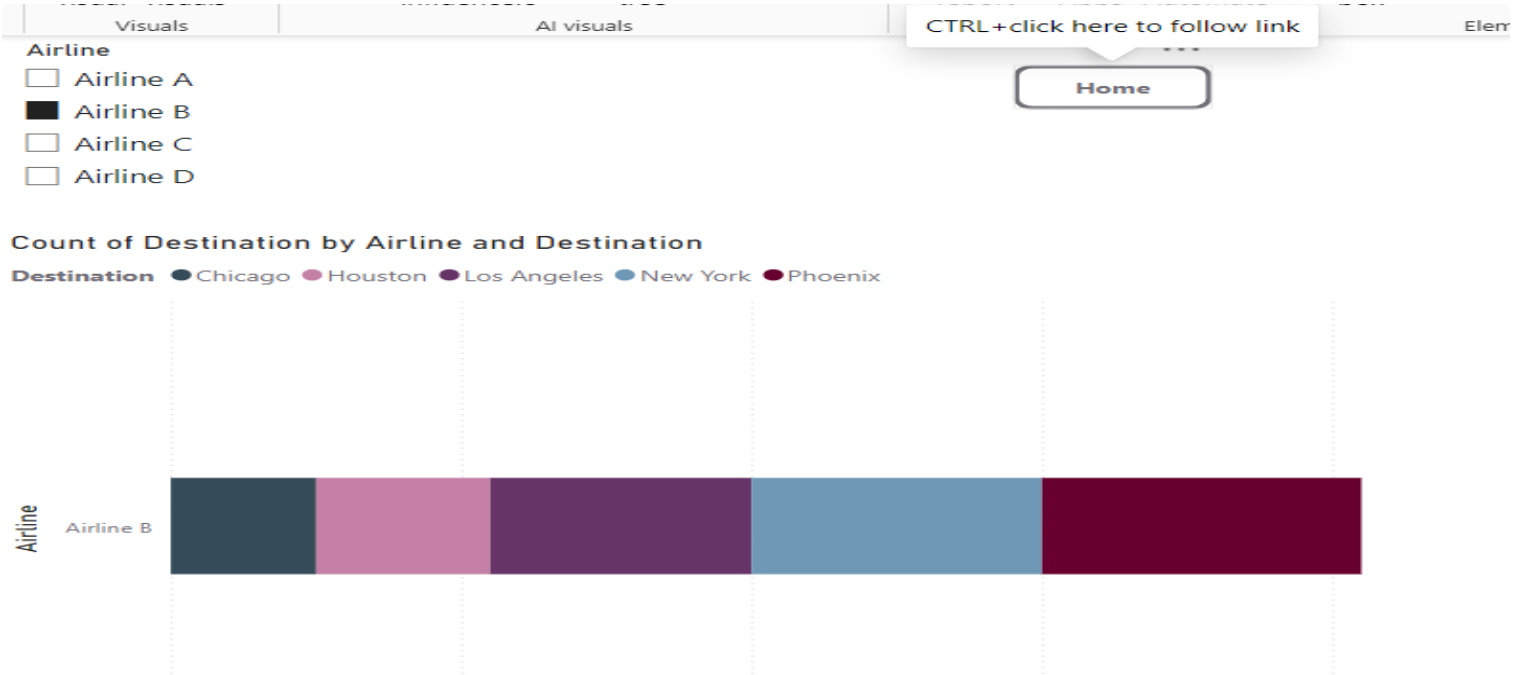




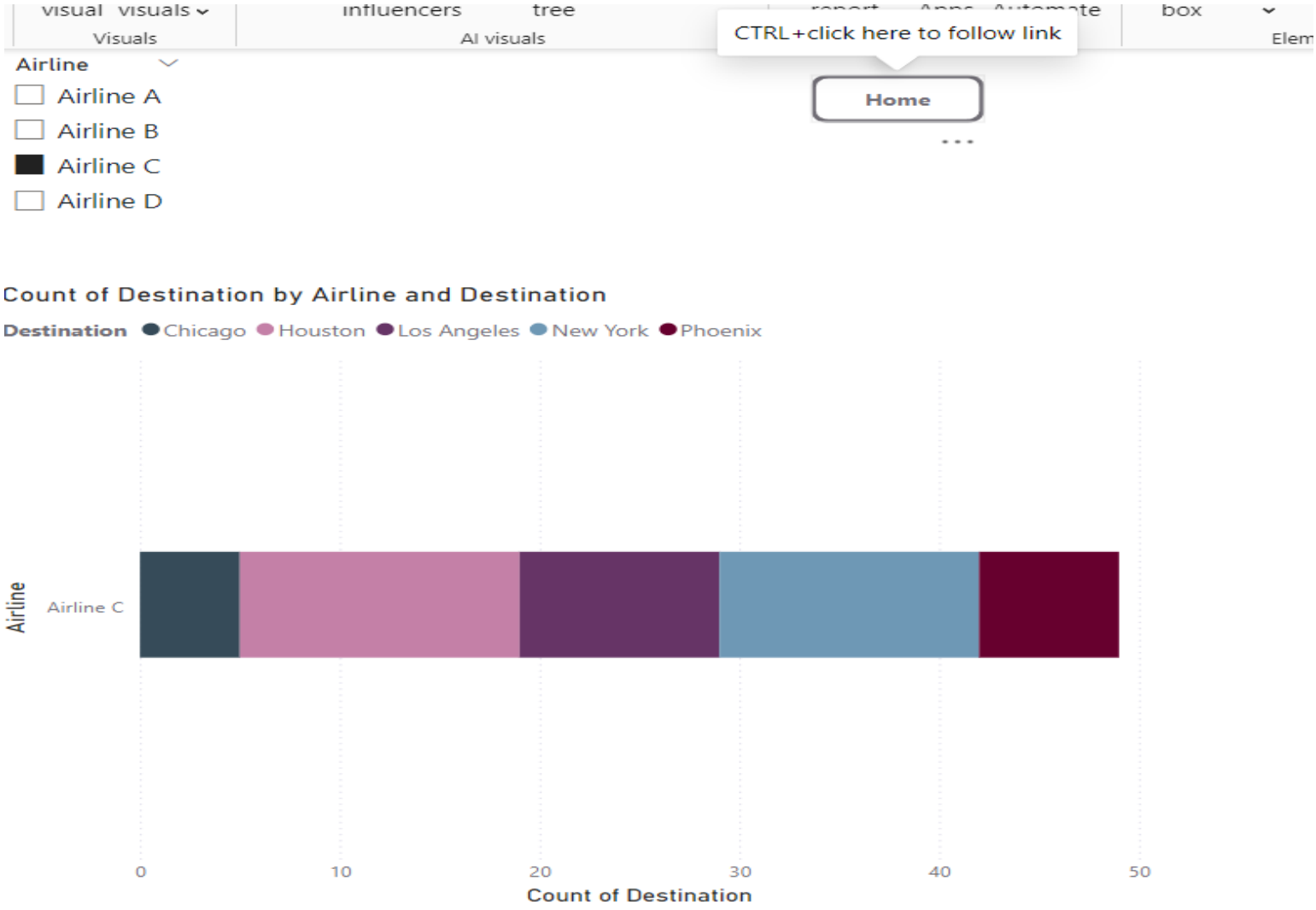
Airline A



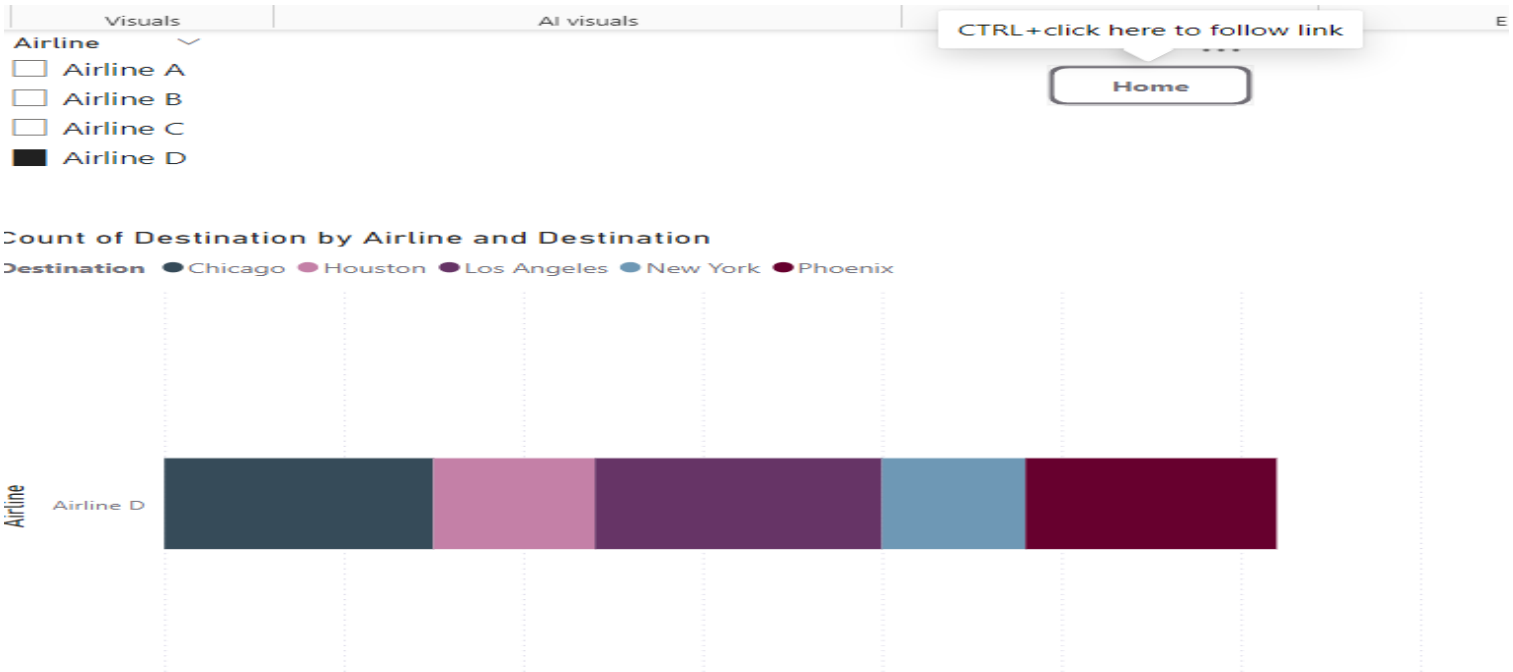
Airline B



Airline C



Airline D

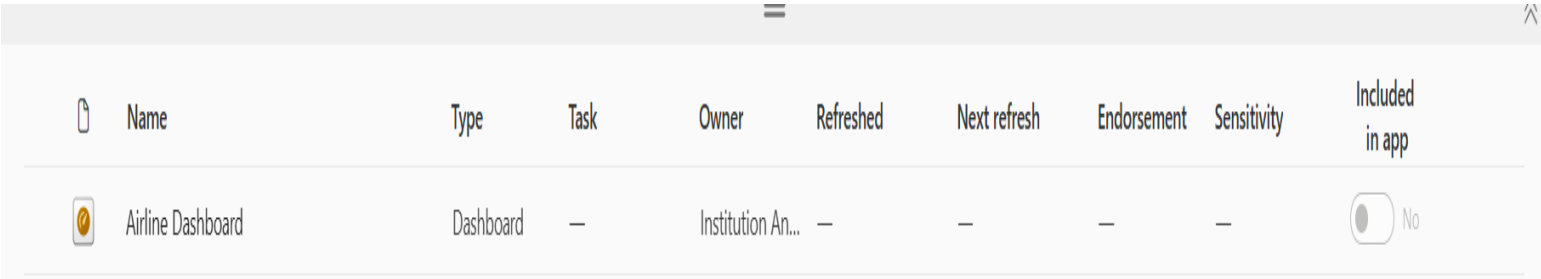


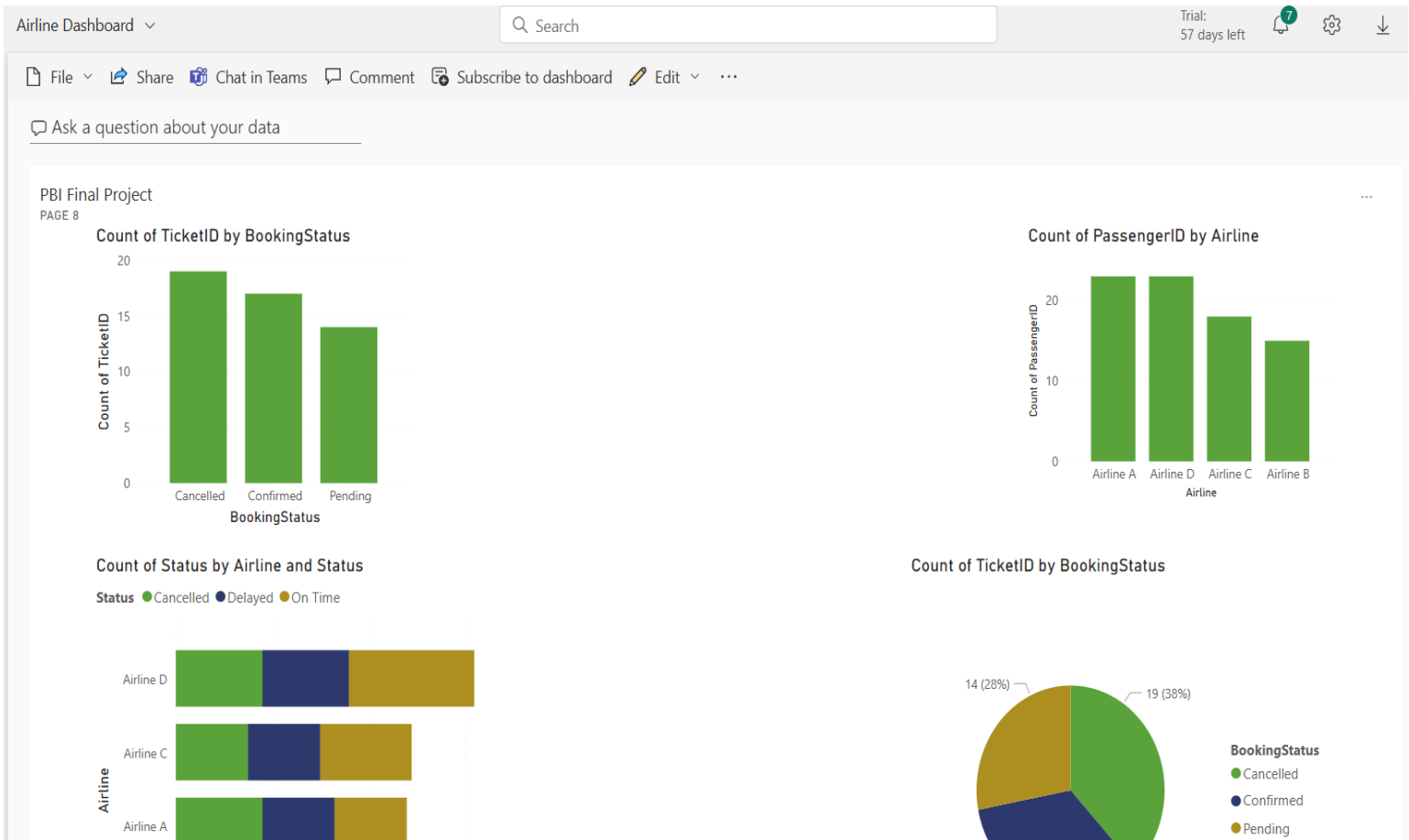
6. Final Dashboard and Power BI Service

● Design a comprehensive dashboard with key visuals and insights

Start by determining the key insights:

- Passenger Count by Airline: Shows the total number of passengers for each airline.
- Ticket Booking Status: Displays the number of bookings based on their status (Booked, Pending, Cancel).
- Flights by Airline and Destination: Displays the number of flights per airline and their destinations.
- Airline Performance: Shows key performance metrics (e.g., total flights, cancellations, on-time performance).
- Booking Trends (Time-based): Shows trends over time for bookings and cancellations.





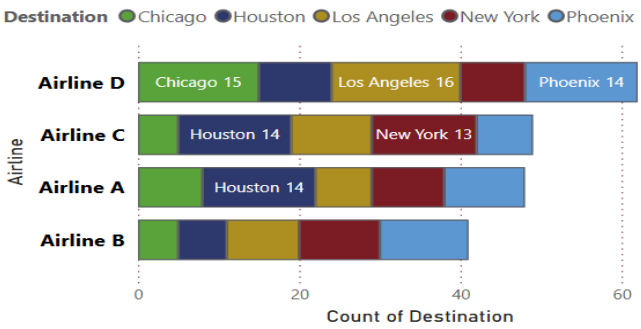
Airline Dashboard

Search

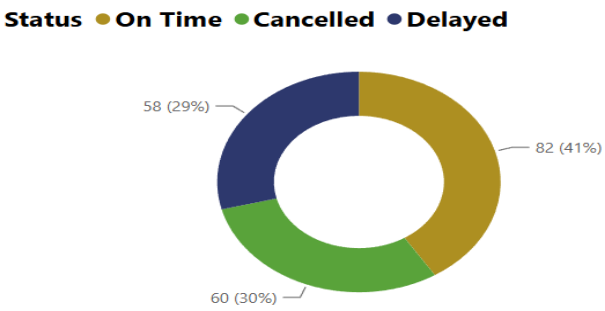
File | Share | Chat in Teams | Comment | Subscribe to dashboard | Edit | ...

Ask a question about your data

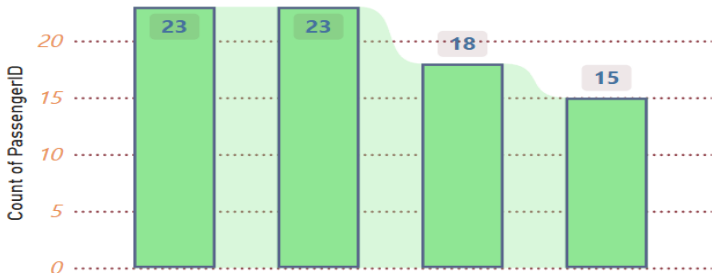
Flights by Airline and Destination



Ticket Booking Status



Count of PassengerID
BY AIRLINE



Total Tickets Booked
BY FLIGHTID

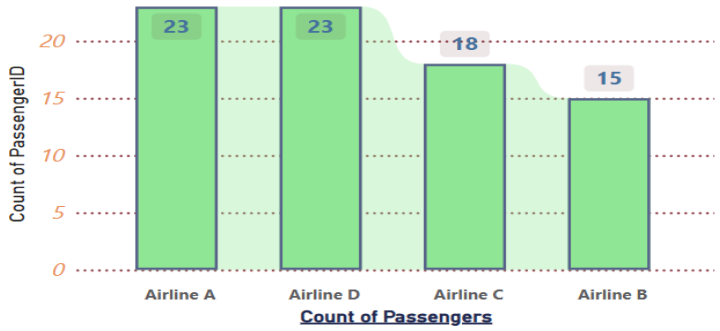
Total Tickets Booked	FlightID
1	1005
1	1014
1	1023
1	1030
1	1035
1	1040
1	1042
1	1064
2	1078
1	1093

Airline Dashboard

Search

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

Ask a question about your data



Total Tickets Booked	FlightID
1	1005
1	1014
1	1023
1	1030
1	1035
1	1040
1	1042
1	1064
2	1078
1	1093
1	1097
17	

Total Passengers for Specific Flight
BY FLIGHTID

FlightID	Total Passengers for Specific Flight
1001	1
1003	2
1005	1
1006	1
1010	1
1011	1
1012	1
1026	1
1027	1
1030	1
1032	1
1033	1
Total	**79**

Flight_status_classification, FlightID

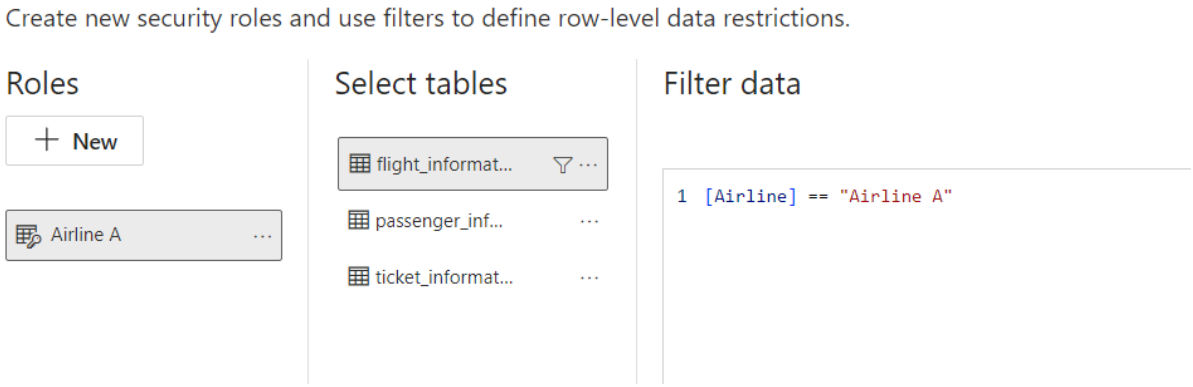
Flight_status_classification	FlightID
Best	1001
Best	1002
Best	1006
Best	1011
Best	1013
Best	1020
Best	1023
Best	1025
Best	1027
Best	1028
Best	1029
Best	1030

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

Step 1: To Setup row level security Go to Modeling Tab and click on Manage roles

Step 2: Name the row level security Airline A

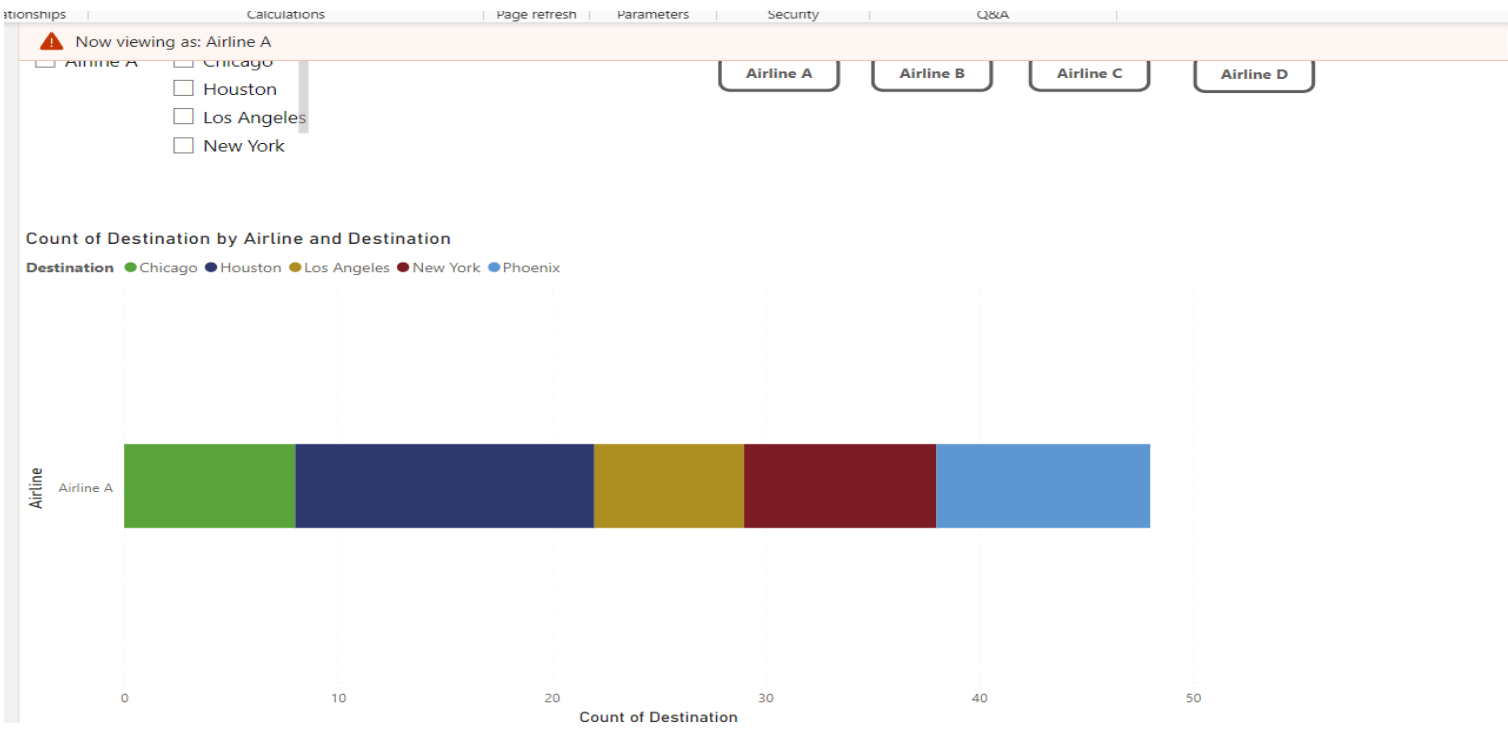
Step 3: Select the flight_information data set and click on new Column Airline, Conditions Equals, Value Airline A

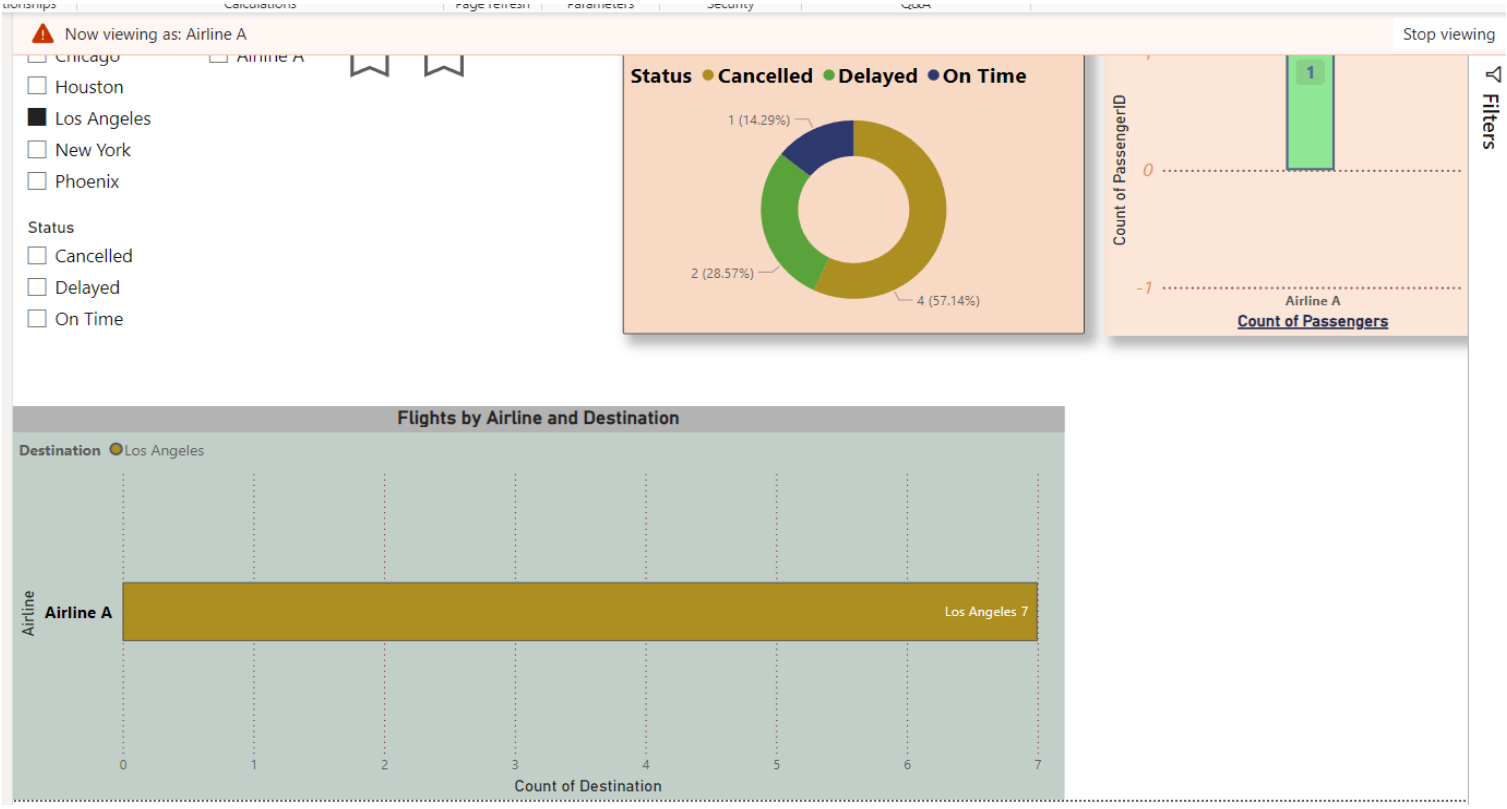


Step 4: Move to DAX editor and correct the value “”” Airline A”””” to “Airline A” and Click on save after that close.

Step 5: To assign this RLS go to the Modeling Tab and click on View as and Click on Airline A after that ok.

You will see: --

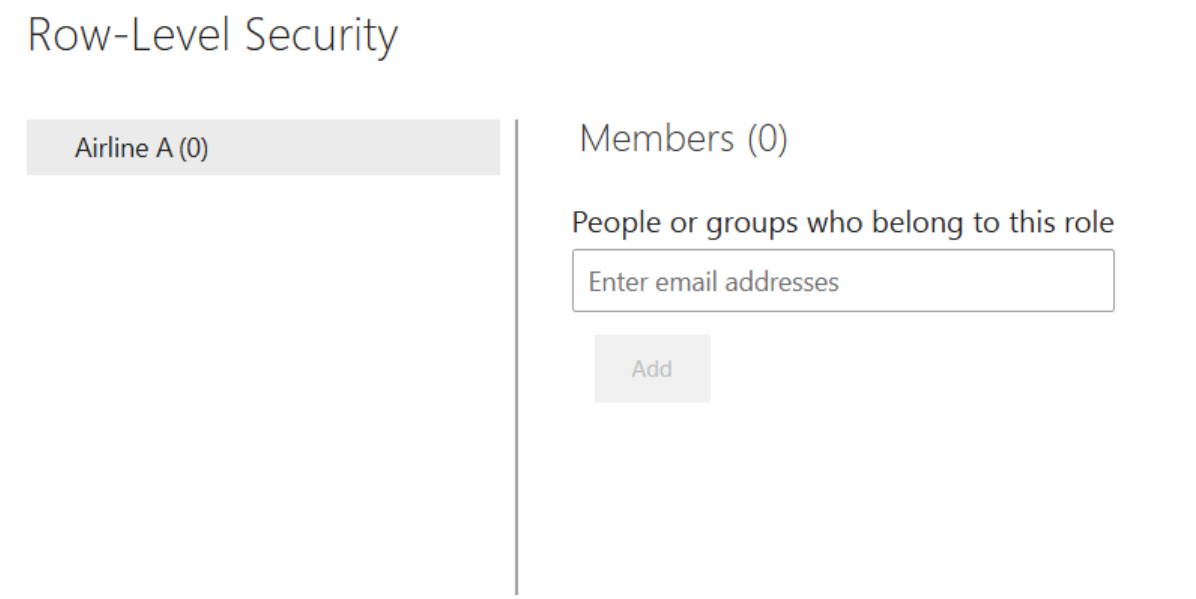



Now viewing as: Airline A
Stop viewing
Houston
Los Angeles
New York
Phoenix
Status
Cancelled
Delayed
On Time
Status Cancelled Delayed On Time
1 (14.29%)
2 (28.57%)
4 (57.14%)
Count of PassengerID
1
0
-1
Airline A
Count of Passengers
Filters
Flights by Airline and Destination
Destination Los Angeles
Airline
Airline A
Los Angeles 7
0
1
2
3
4
5
6
7
Count of Destination

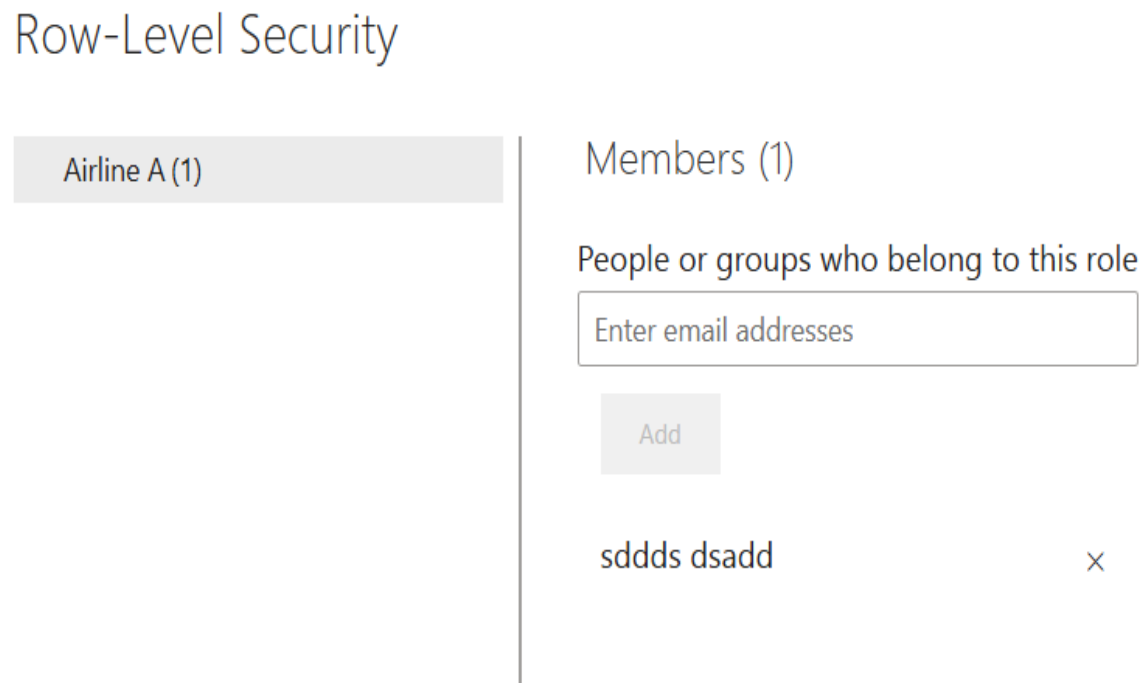

Now viewing as: Airline A
Flights by Airline and Destination
Destination Chicago Houston Los Angeles New York Phoenix
Airline
Airline A
Chicago 8
Houston 14
Los Angeles 7
New York 9
Phoenix 10
0
10
20
30
40
50
Count of Destination

Step 6: You have to publish the Report in Workspace you have created (Institution analysis)

Step 7: To Apply the RLS on the Cloud You have to go to Your workspace you have created and after that you will select the Dataset and click on 3 dots and go to security you will see RLS you will created.



Step 8: In the People or Groups who belong to this role type the email ID you want to assign The role and add them after that save.



Set up a schedule refresh at 5 PM daily

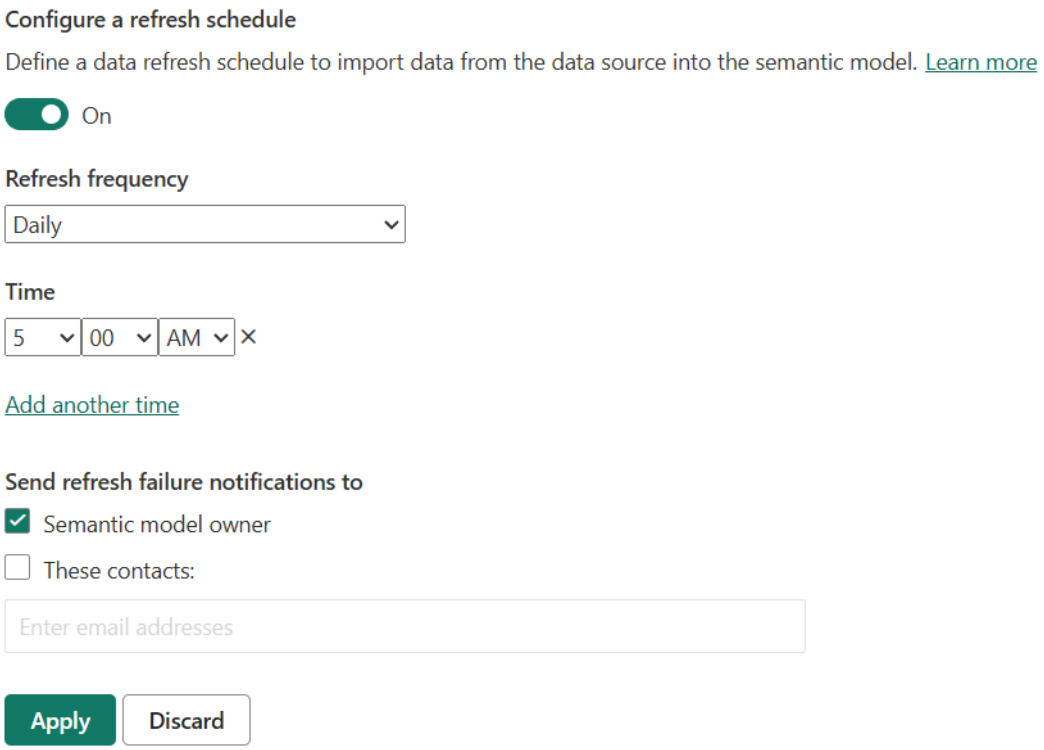
Step 3: To Schedule Refresh open “Institution Analysis” workspace and go to setting and go to power BI Settings.

Step 4: In setting you will see Semantic Models click on this and scroll down you will see the

Refresh option it is of turn it on choose [Refresh Frequency (Daily)] after that choose

The [Time Zone India] after that set the [Time 05 – 00—AM Daily] Refresh failure

Notification sends to Semantic Model Owner and click on (Apply)



Video Introduction of Project:→
https://drive.google.com/file/d/1u4gS9m5vRYiC7_fV1zIUb5Td9qq0XxT4/view?usp=drivesdk