Airline Data Management and Analysis Using Power BI

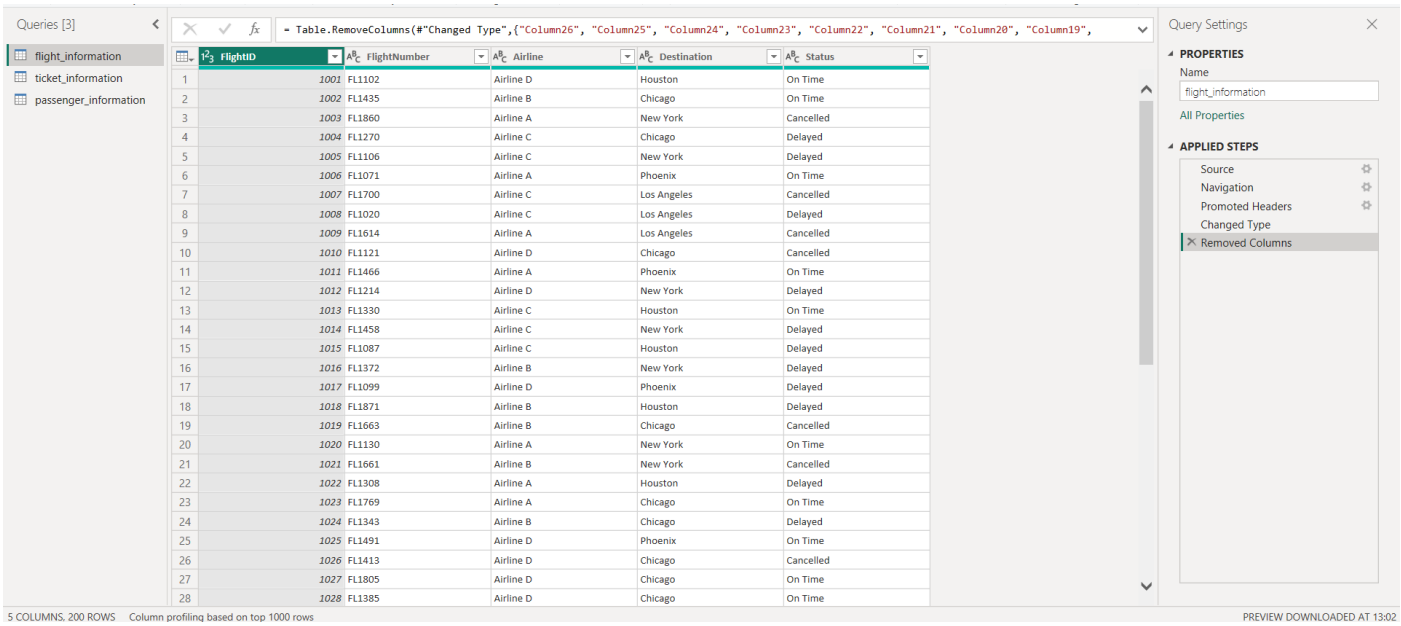
Objective: - To analyze and visualize airline data for operational insights, passenger management, and ticket booking trends using Power BI.

TASK 1: - DATA PREPARATION AND CLEANING.

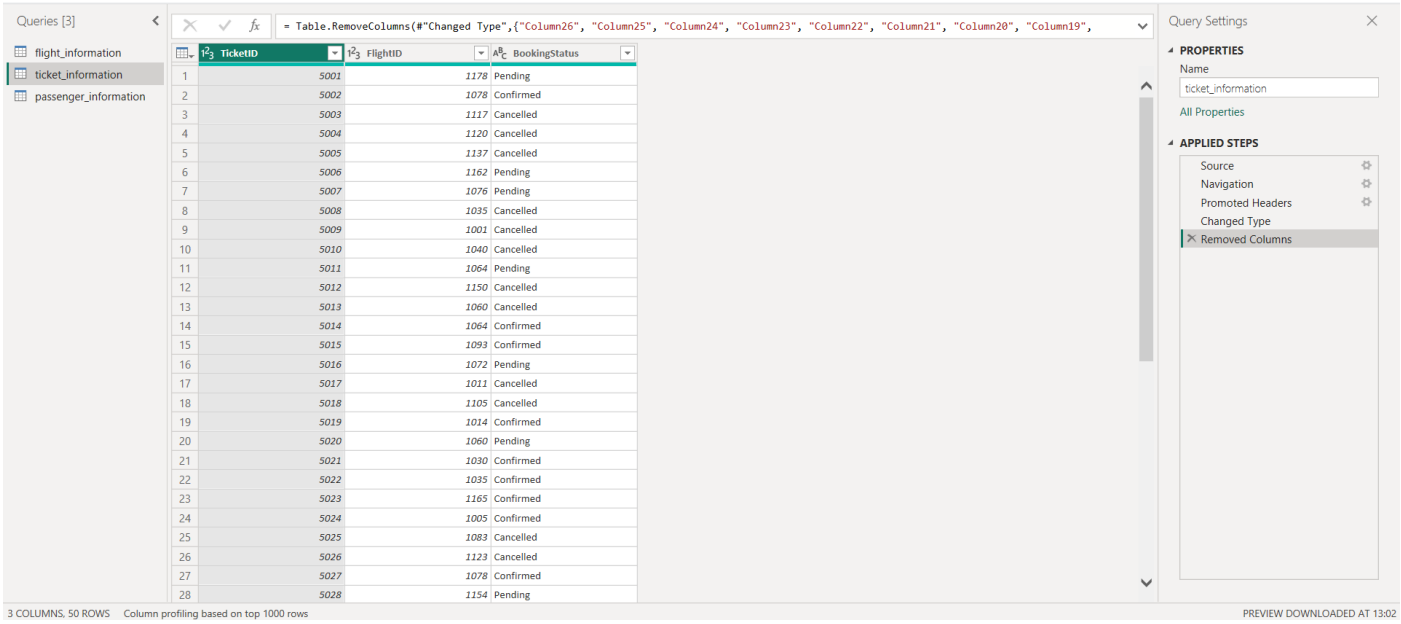
STEP 1: First, I take datasets individually and put them into power query editor.

STEP 2: Then I cleaned each data table by cleaning unnecessary columns and formatting data type.

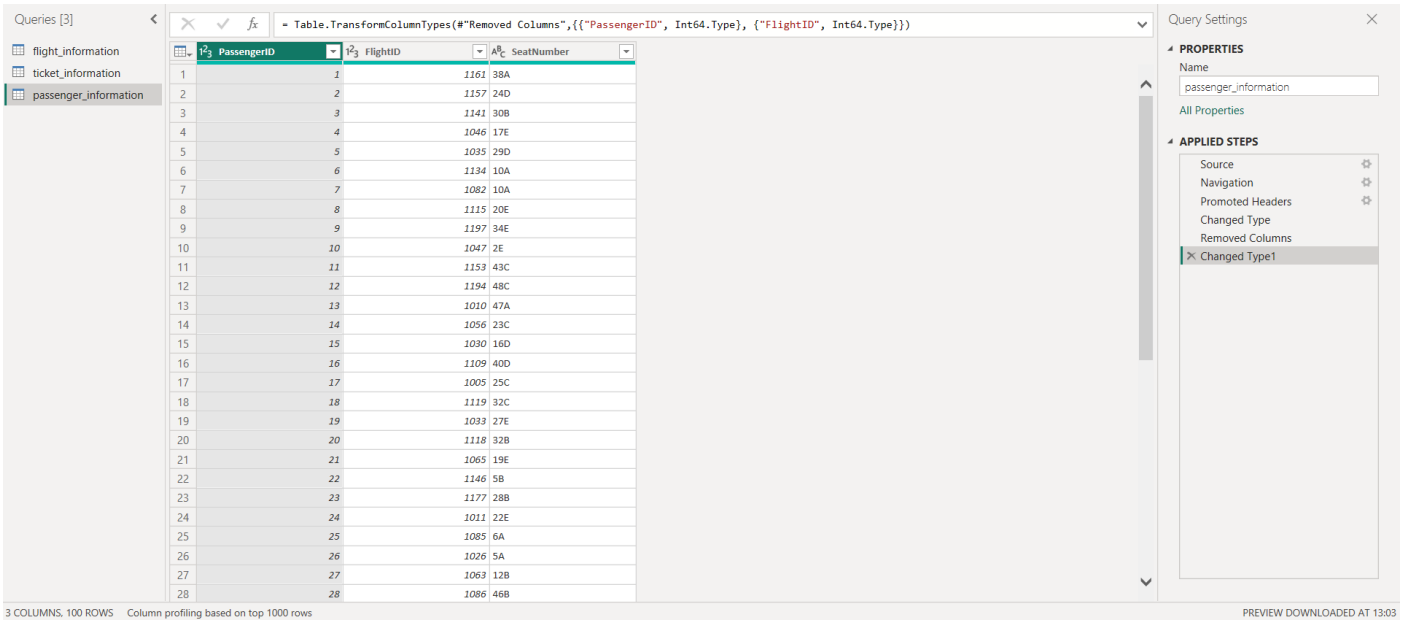
1. Flight Information



2. Ticket Information



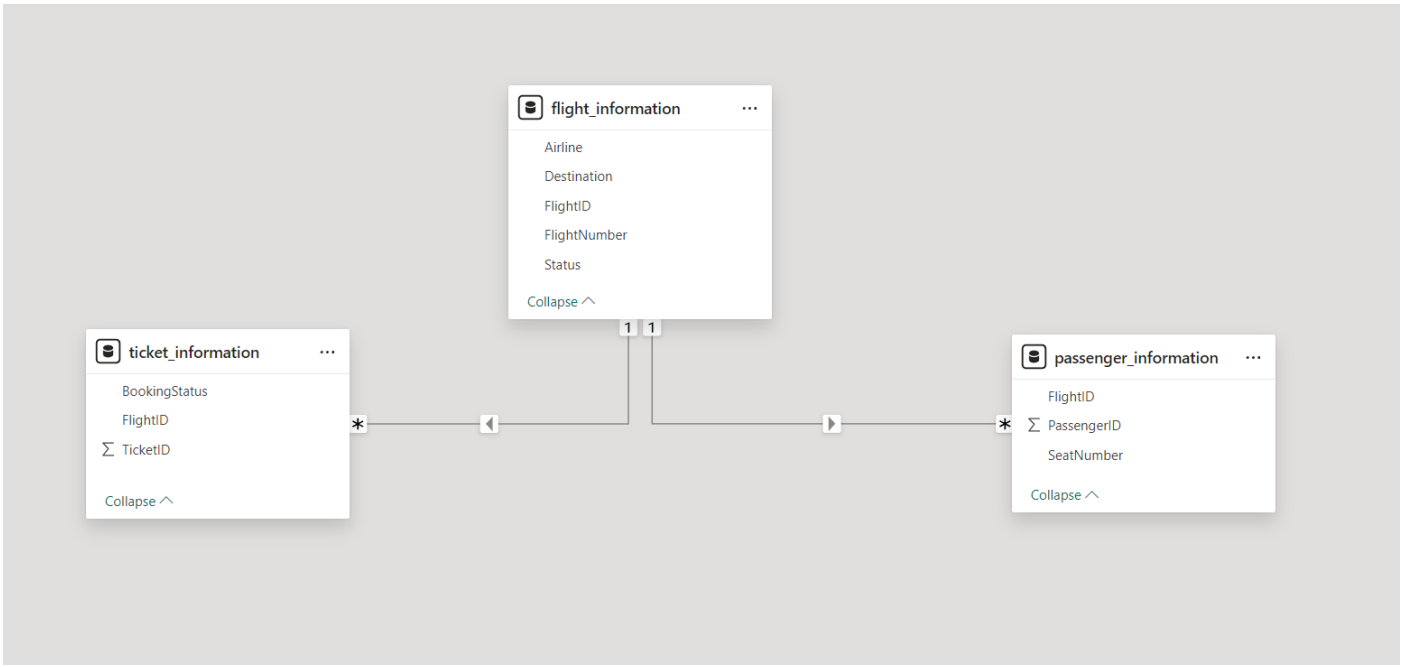
3. Passenger Information



TASK 2: - DATA MODELING.

STEP 1: After cleaning the data, I click on ‘close & apply’ button and load the datasets in POWER BI DESKTOP.

STEP 2: Then I go to ‘Model View’ to create relationships between these tables and choose cardinality “Many to One”.

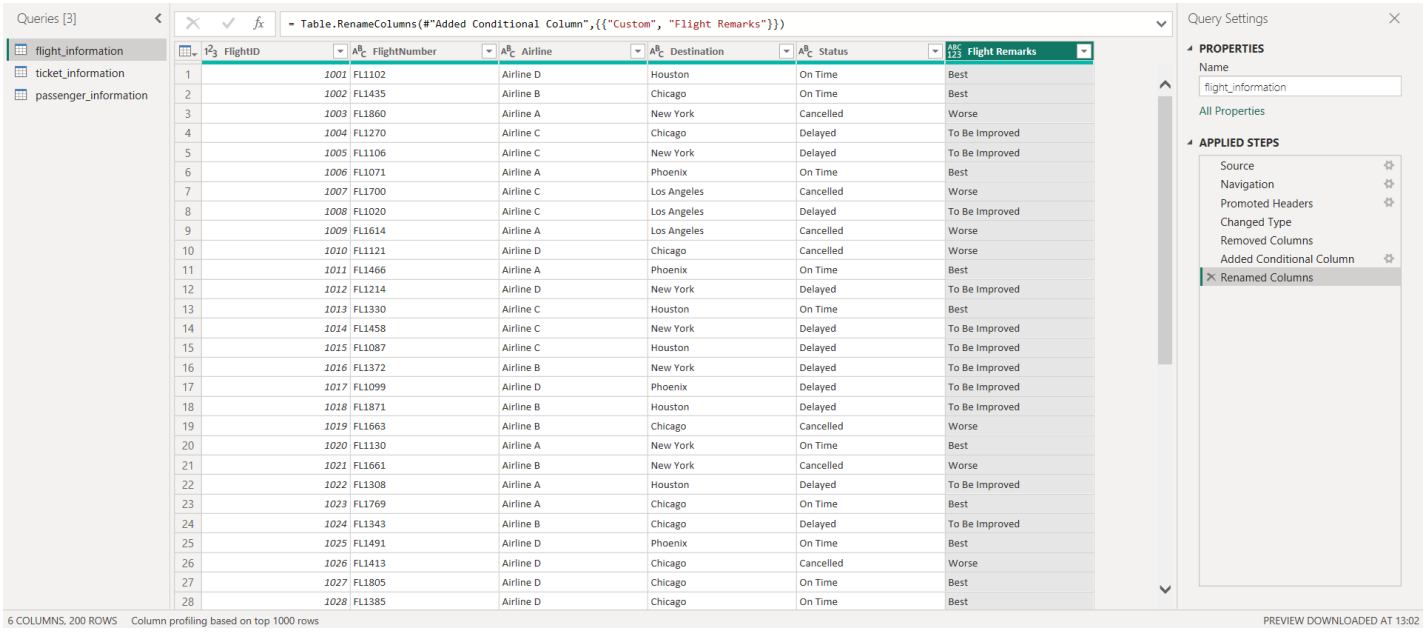


TASK 3: - ENHANCED DATA INSIGHTS.

3.1 – ADD CONDITIONAL COLUMN.

STEP 1: To Add Conditional Column, I select ‘Status’ column from “flight information” table in power query editor.

STEP 2: After selecting the column, I go to transform tab and click on ‘conditional column’, then I put the values in it for On time flights I choose ‘Best’, for delayed flights I choose ‘To be improved’ and for cancelled flights I choose ‘Worse’.



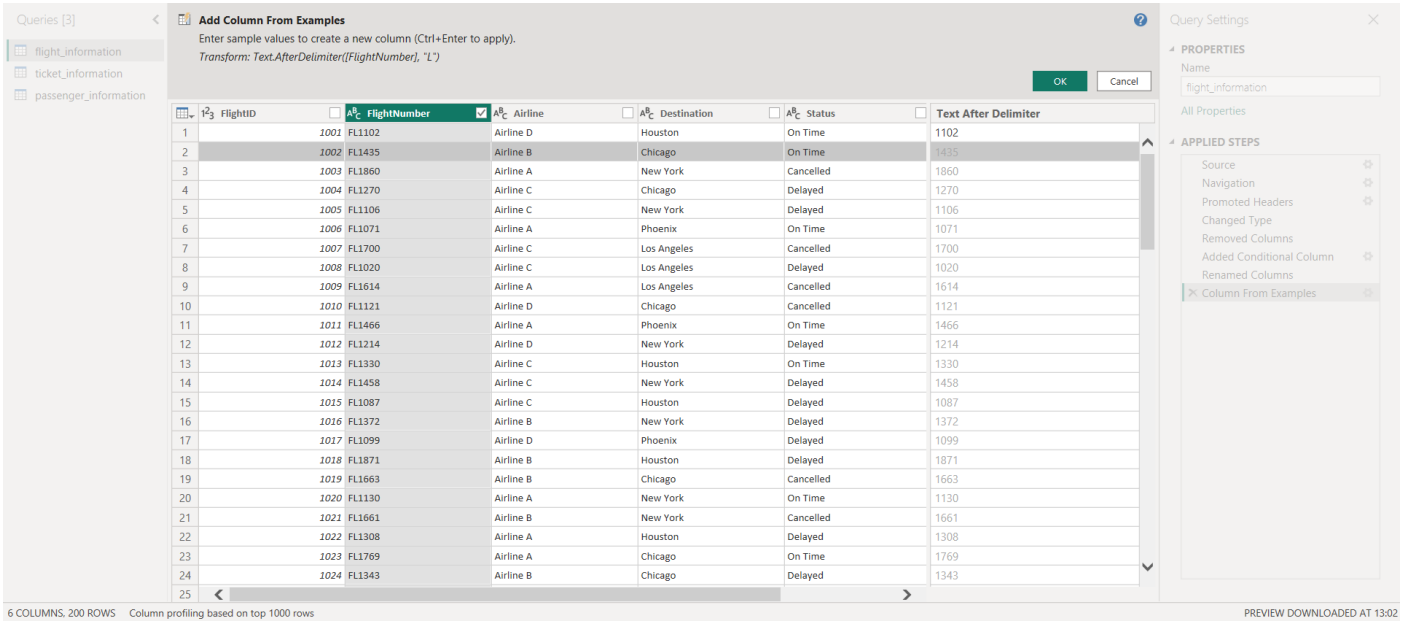
	FlightID	FlightNumber	Airline	Destination	Status	Flight Remarks
1	1001	FL1102	Airline D	Houston	On Time	Best
2	1002	FL1435	Airline B	Chicago	On Time	Best
3	1003	FL1860	Airline A	New York	Cancelled	Worse
4	1004	FL1270	Airline C	Chicago	Delayed	To Be Improved
5	1005	FL1106	Airline C	New York	Delayed	To Be Improved
6	1006	FL1071	Airline A	Phoenix	On Time	Best
7	1007	FL1700	Airline C	Los Angeles	Cancelled	Worse
8	1008	FL1020	Airline C	Los Angeles	Delayed	To Be Improved
9	1009	FL1614	Airline A	Los Angeles	Cancelled	Worse
10	1010	FL1121	Airline D	Chicago	Cancelled	Worse
11	1011	FL1466	Airline A	Phoenix	On Time	Best
12	1012	FL1214	Airline D	New York	Delayed	To Be Improved
13	1013	FL1330	Airline C	Houston	On Time	Best
14	1014	FL1458	Airline C	New York	Delayed	To Be Improved
15	1015	FL1087	Airline C	Houston	Delayed	To Be Improved
16	1016	FL1372	Airline B	New York	Delayed	To Be Improved
17	1017	FL1099	Airline D	Phoenix	Delayed	To Be Improved
18	1018	FL1871	Airline B	Houston	Delayed	To Be Improved
19	1019	FL1663	Airline B	Chicago	Cancelled	Worse
20	1020	FL1130	Airline A	New York	On Time	Best
21	1021	FL1661	Airline B	New York	Cancelled	Worse
22	1022	FL1308	Airline A	Houston	Delayed	To Be Improved
23	1023	FL1769	Airline A	Chicago	On Time	Best
24	1024	FL1343	Airline B	Chicago	Delayed	To Be Improved
25	1025	FL1491	Airline D	Phoenix	On Time	Best
26	1026	FL1413	Airline D	Chicago	Cancelled	Worse
27	1027	FL1805	Airline D	Chicago	On Time	Best
28	1028	FL1385	Airline D	Chicago	On Time	Best

3.2 – COLUMN FROM EXAMPLE.

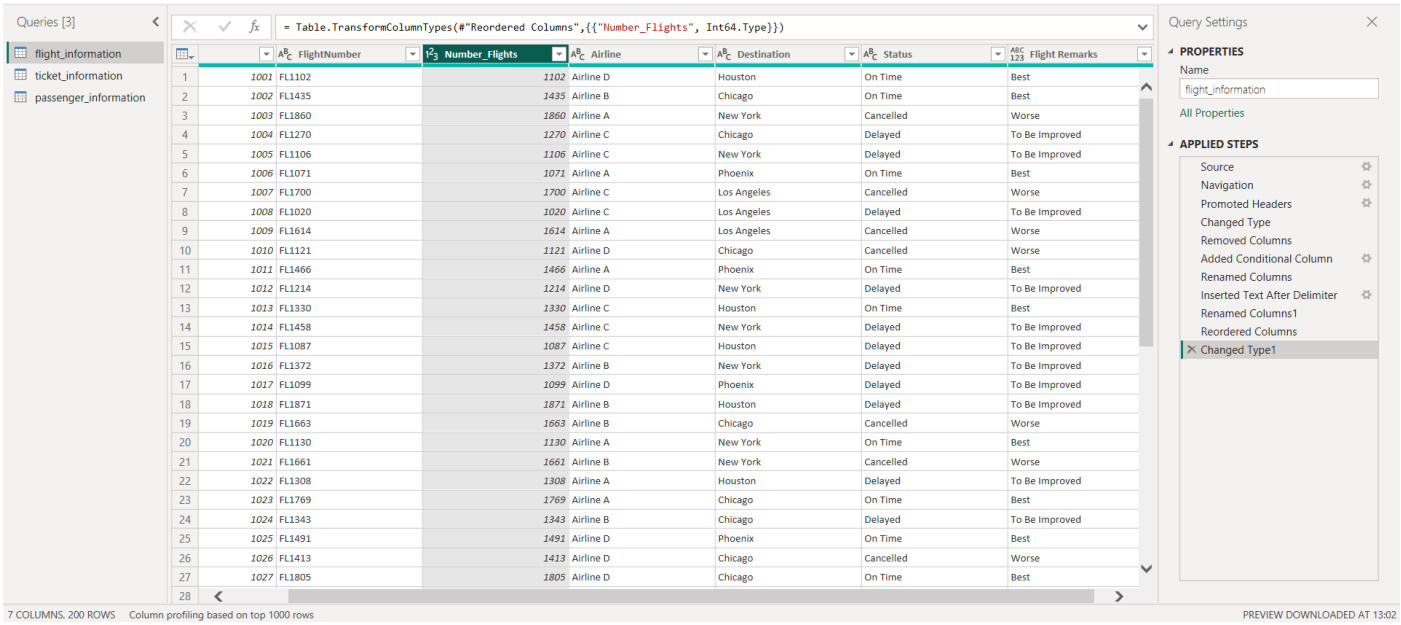
STEP 1: Fist, I go to “flight information” table and select ‘FlightNumber’ column.

STEP 2: Then I go to transform tab and go to column from examples and click on down arrow, then I click on ‘from selection’.

STEP 3: Then I extract the number from ‘FlightNumber’ column and name it as ‘Number_Flights’ column.



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



		FlightNumber	Number_Flights	Airline	Destination	Status	Flight Remarks
1	1001	FL1102	1102	Airline D	Houston	On Time	Best
2	1002	FL1435	1435	Airline B	Chicago	On Time	Best
3	1003	FL1860	1860	Airline A	New York	Cancelled	Worse
4	1004	FL1270	1270	Airline C	Chicago	Delayed	To Be Improved
5	1005	FL1106	1106	Airline C	New York	Delayed	To Be Improved
6	1006	FL1071	1071	Airline A	Phoenix	On Time	Best
7	1007	FL1700	1700	Airline C	Los Angeles	Cancelled	Worse
8	1008	FL1020	1020	Airline C	Los Angeles	Delayed	To Be Improved
9	1009	FL1614	1614	Airline A	Los Angeles	Cancelled	Worse
10	1010	FL1121	1121	Airline D	Chicago	Cancelled	Worse
11	1011	FL1466	1466	Airline A	Phoenix	On Time	Best
12	1012	FL1214	1214	Airline D	New York	Delayed	To Be Improved
13	1013	FL1330	1330	Airline C	Houston	On Time	Best
14	1014	FL1458	1458	Airline C	New York	Delayed	To Be Improved
15	1015	FL1087	1087	Airline C	Houston	Delayed	To Be Improved
16	1016	FL1372	1372	Airline B	New York	Delayed	To Be Improved
17	1017	FL1099	1099	Airline D	Phoenix	Delayed	To Be Improved
18	1018	FL1871	1871	Airline B	Houston	Delayed	To Be Improved
19	1019	FL1663	1663	Airline B	Chicago	Cancelled	Worse
20	1020	FL1130	1130	Airline A	New York	On Time	Best
21	1021	FL1661	1661	Airline B	New York	Cancelled	Worse
22	1022	FL1308	1308	Airline A	Houston	Delayed	To Be Improved
23	1023	FL1769	1769	Airline A	Chicago	On Time	Best
24	1024	FL1343	1343	Airline B	Chicago	Delayed	To Be Improved
25	1025	FL1491	1491	Airline D	Phoenix	On Time	Best
26	1026	FL1413	1413	Airline D	Chicago	Cancelled	Worse
27	1027	FL1805	1805	Airline D	Chicago	On Time	Best
28							

TASK 4: - CALCULATION USING DAX.

4.1: - CALCULATE TOTAL PASSENGERS FOR SPECIFIC FLIGHT.

STEP 1: First, I go to Home Tab and click on 'New Measure' and create a DAX formula.

STEP 2: The DAX formula which I create is specific flight for passengers = CALCULATE(COUNT(passenger_information[PassengerID]),flight_information[FlightNumber]="FL1686")

STEP 3: After creating DAX formula, I used single-row card visual from visualization pane and put the column 'Specific flight for passengers' in field pane.

```
1 Specific flight for passengers = CALCULATE(COUNT(passenger_information[PassengerID]),flight_information[FlightNumber]="FL1686")
```

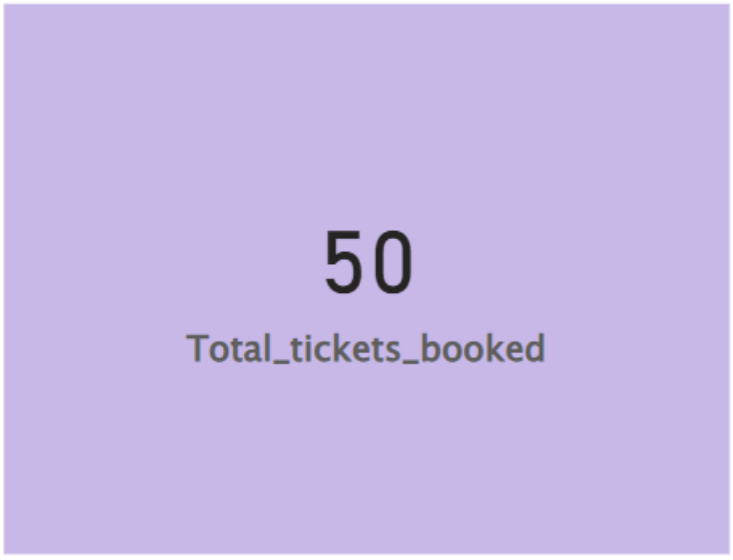
4.2: - CALCULATE TOTAL TICKETS BOOKED.

STEP 1: I add another page and go to Home tab and click on 'New Measure' and create a DAX formula.

STEP 2: The DAX formula is "Total_tickets_booked = COUNT(ticket_information[TicketID])"

STEP 3: After creating the DAX formula, I take single-row card visual from visualization pane and put 'Total_tickets_booked' column in field pane.

```
1 Total_tickets_booked = COUNT(ticket_information[TicketID])
```



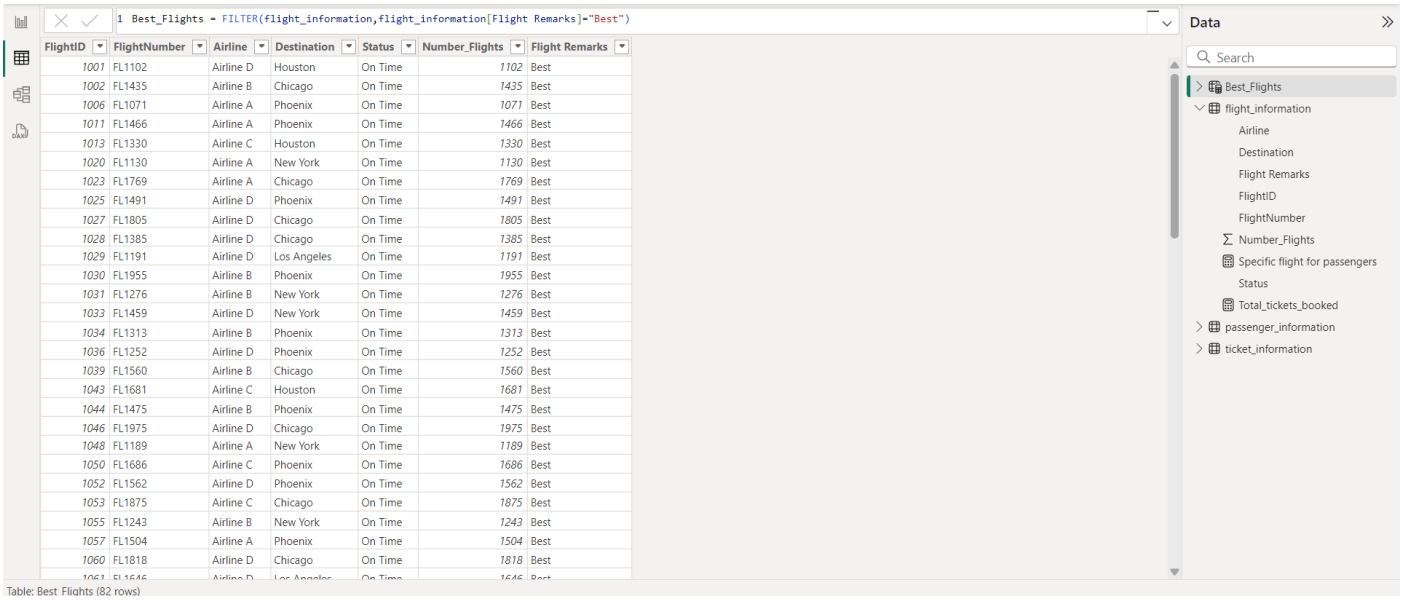
4.3: - FILTER TABLE SHOWING 'BEST' FLIGHTS ONLY.

STEP 1: For creating filter table, I go to Table view and then I go to Home Tab > click on 'New Table'.

STEP 2: After creating new table I have to create the DAX formula which is Best_Flights = FILTER(flight_information,flight_information[Flight Remarks]="Best")

STEP 3: And after creating the DAX formula I click on ENTER and my new filter table created which has data only according to 'Best' flights.

```
1  Best_Flights = FILTER(flight_information,flight_information[Flight Remarks]="Best")
```



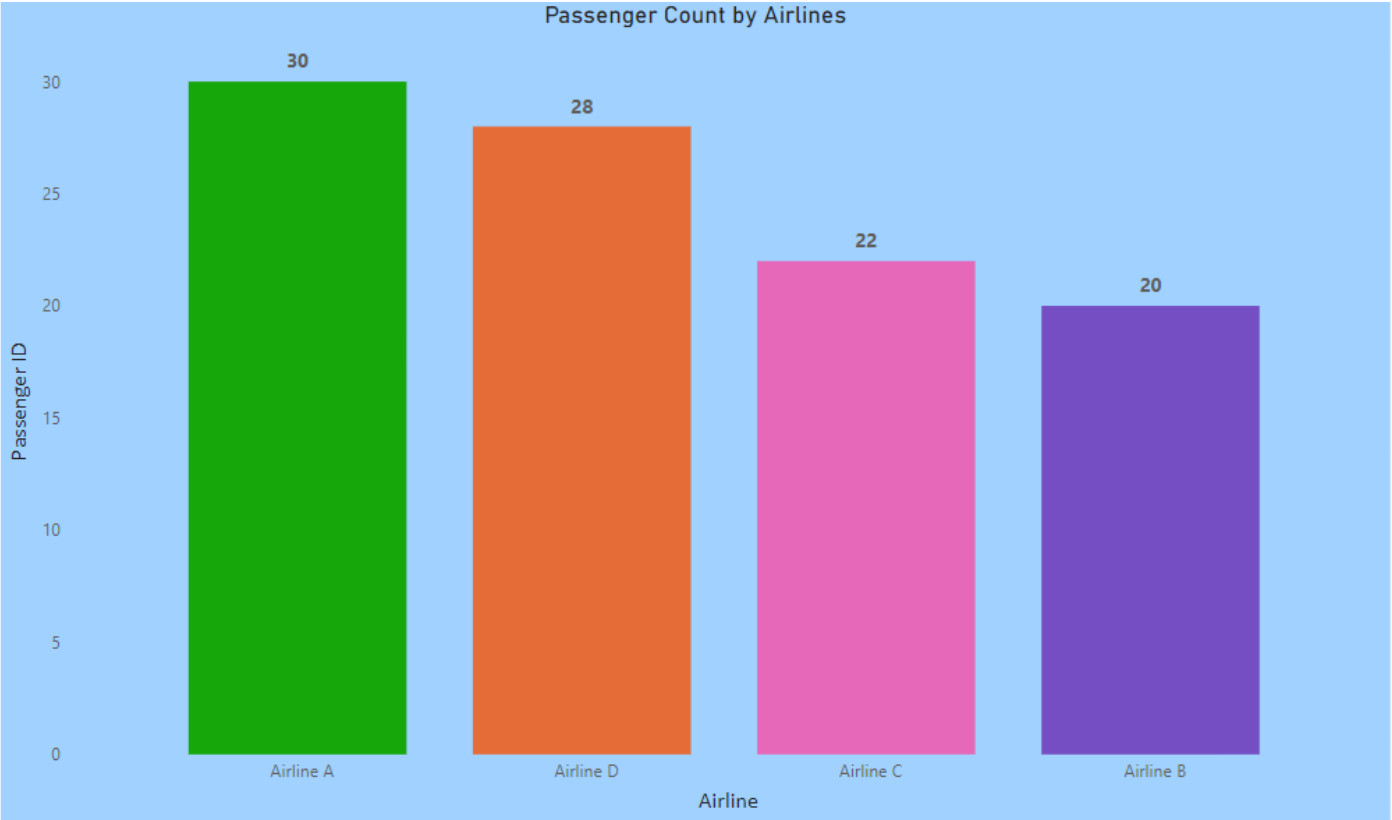
TASK 5: - VISUALIZATION AND INTERACTIVE FEATURES.

5.1.1: - CREATE VISUAL FOR PASSENGER COUNT BY AIRLINE.

STEP 1: First, I take a column chart visual from visualization pane.

STEP 2: Then I put 'Airline' column in X-axis section and 'PassengerID' column in Y-axis section, then I used count function in 'PassengerID' column.

STEP 3: After that I used format pane and change columns color differently, also I show data labels on each column. And I changed the background color too.

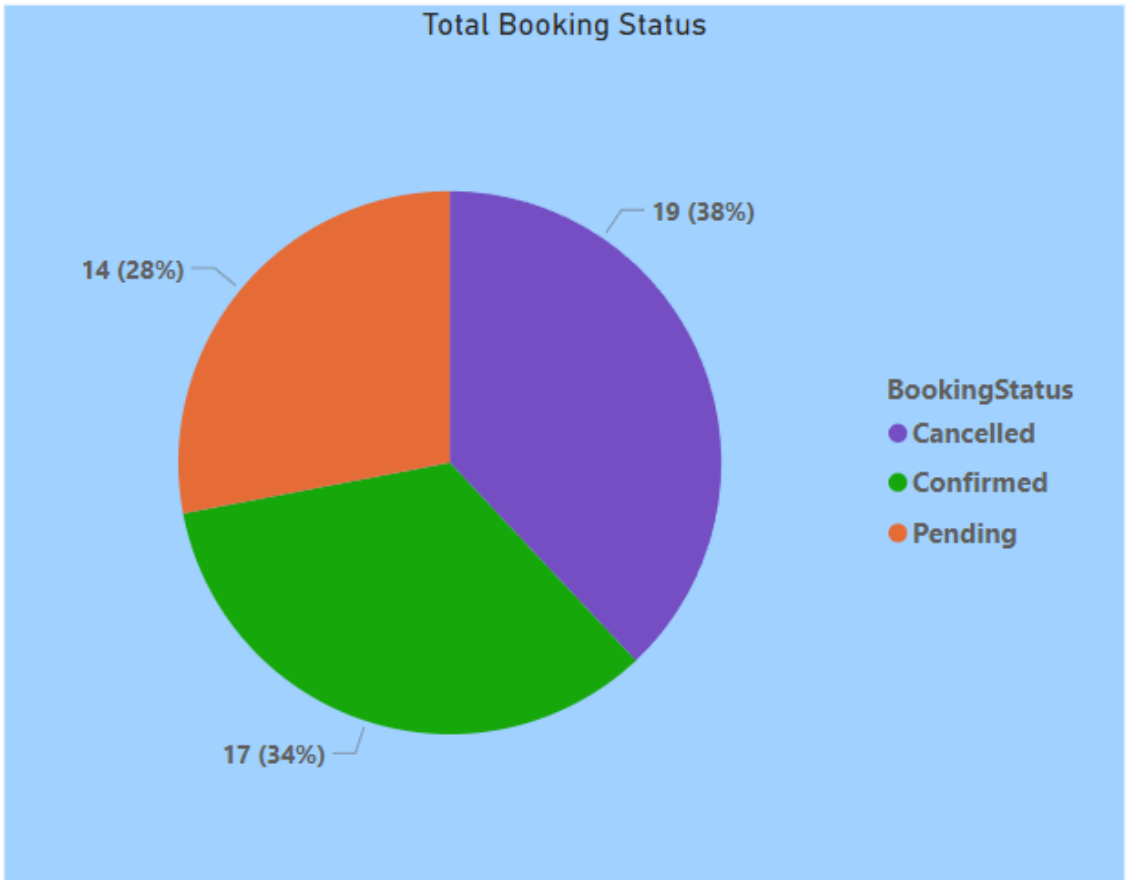


5.1.2: - CREATE VISUAL FOR TICKET BOOKING STATUS.

STEP 1: First, I take pie chart visual from visualization pane.

STEP 2: Then I put 'BookingStatus' column in Legend section and 'TicketID' in Values section, then I used the count function in 'TicketID' column.

STEP 3: Then I go to format pane and changes the background colour of the visual and changes the slices colour to make it attractive.



5.1.3: - CREATE VISUAL FOR FLIGHTS BY AIRLINE AND DESTINATION.

STEP 1: First, I take Matrix visual from visualization pane.

STEP 2: Then I put 'Destination' column in Rows section, 'Airline' column in Columns section and 'FlightID' column in Values section, then I used count function in 'FlightID' column.

STEP 3: After that I changed the background colour of the visual, changed the layout and style to minimal and make the column header to 'bold'.

Flights by Destination and Airlines

Destination	Airline A	Airline B	Airline C	Airline D	Total
Chicago	8	5	5	15	**33**
Houston	14	6	14	9	**43**
Los Angeles	7	9	10	16	**42**
New York	9	10	13	8	**40**
Phoenix	10	11	7	14	**42**
Total	**48**	**41**	**49**	**62**	**200**

5.2.1: - ADD INTERACTIVE FEATURES FOR DESTINATION AND AIRLINE & QUICK VIEWS.

STEP 1: First, I added all the visuals in a same page.

STEP 2: Then I added slicer from visualization pane for both Destination and Airline.

STEP 3: For quick view, I added q&a feature from visualization pane to search anything related to the visuals or given dataset.

Flights by Destination and Airlines

Destination	Airline A	Airline B	Airline C	Airline D	Total
Chicago	8	5	5	15	33
Houston	14	6	14	9	43
Los Angeles	7	9	10	16	42
New York	9	10	13	8	40
Phoenix	10	11	7	14	42
Total	48	41	49	62	200

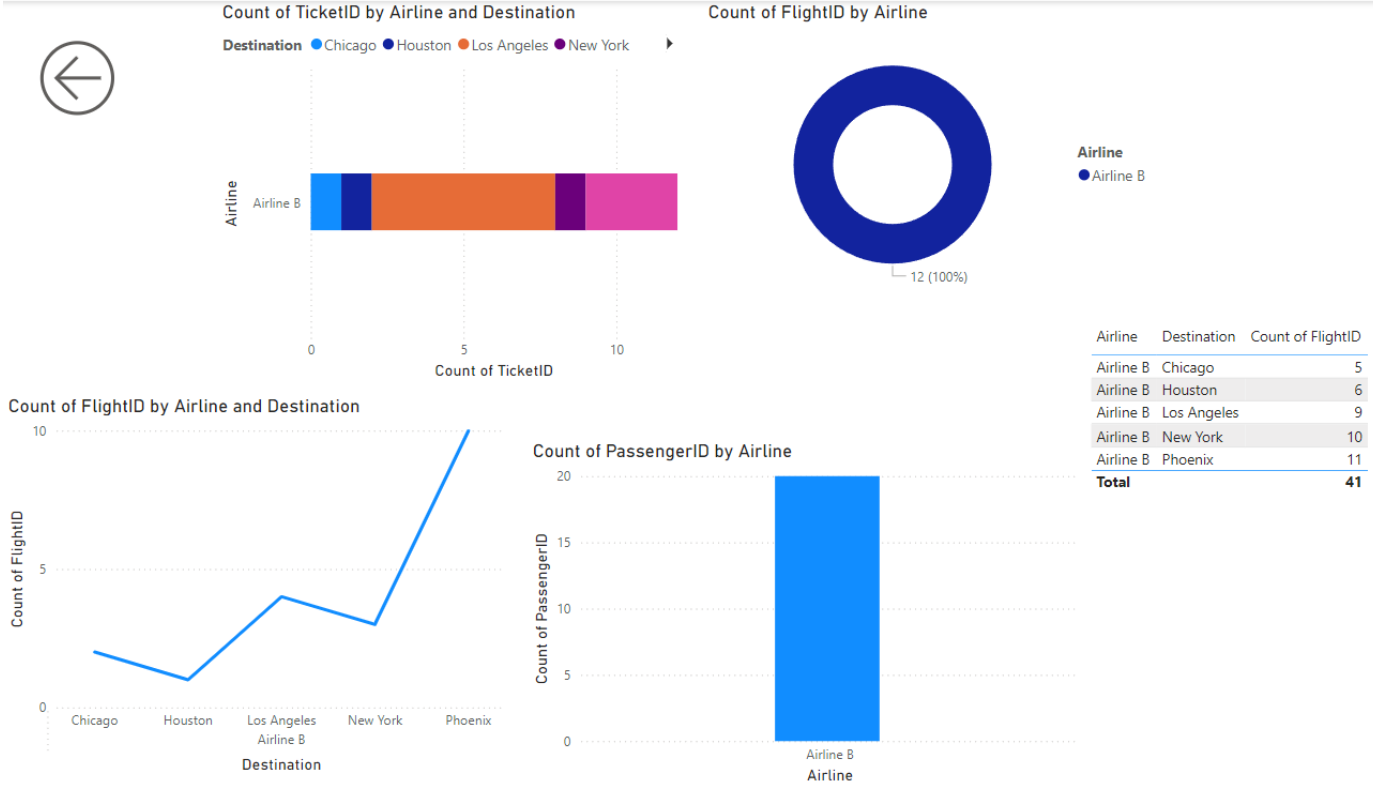
5.2.2: - ADD INTERACTIVE FEATURES FOR AIRLINE-SPECIFIC PAGE.

STEP 1: For this I take a new page and add some visuals regarding to Airline data like a donut chart showing 'FlightID by Airline', a column chart showing 'PassengerID by Airline', a table and line chart visual showing 'Airline, Destination and count of FlightID' and a bar chart showing 'TicketID by Airline and Destination'.

STEP 2: After taking this, I go to visualization pane and below there is an option "Drill Through" in which I put the 'Airline' column and then a 'back' button is created automatically.

STEP 3: Then I go to previous page and right click on 'Airline B' in bar chart visual and go to Drill Through.

STEP 4: After clicking the Drill Through option, I get the data related to only 'Airline B'.



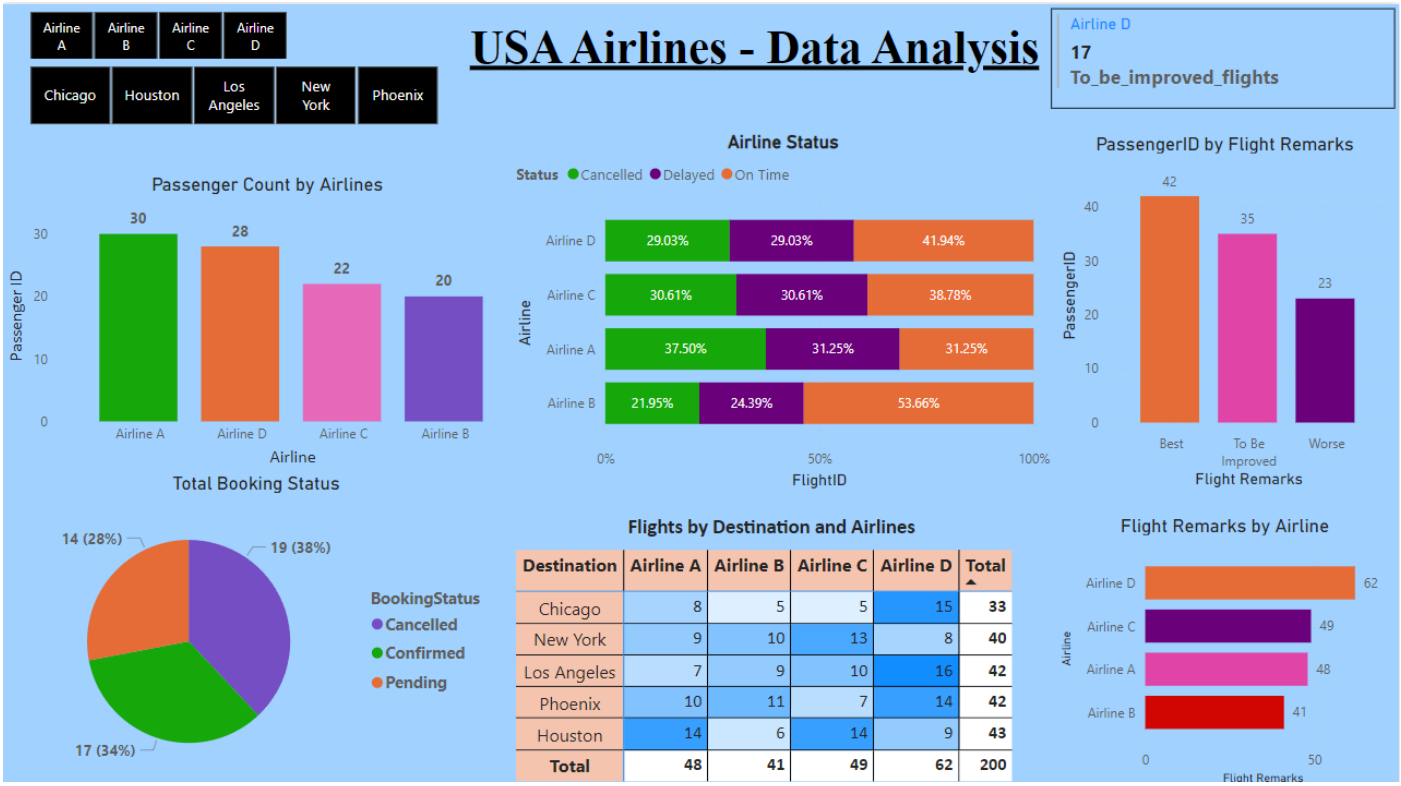
TASK 6: - FINAL DASHBOARD AND POWER BI SERVICE.

6.1: - DESIGN A COMPREHENSIVE DASHBOARD WITH KEY VISUALS AND INSIGHTS.

STEP 1: First, I take all the visuals from different pages into one page and added a few more like I take 100% stacked bar chart visual from visualization pane and put 'Airline' column in Y-axis, 'FlightID' column in X-axis and 'Status' column in Legend section.

STEP 2: Also, I take a column chart and put 'Flight Remarks' column in X-axis and 'PassengerID' in Y-axis section. Last, I added bar chart from visualization pane and put 'Airline' column in Y-axis and 'Flight Remarks' column in X-axis section.

STEP 3: Then I added a multi-row card but for that first I created a DAX formula which is To_be_improved_flights = CALCULATE(DISTINCTCOUNT(flight_information[FlightNumber]),FILTER(flight_information,flight_information[Airline]="Airline D"&&flight_information[Flight Remarks]= "To be improved")) , then I take this created column named 'To_be_improved_flights' in fields section and 'Airline' column also in fields section.



Destination	Airline A	Airline B	Airline C	Airline D	Total
Chicago	8	5	5	15	33
New York	9	10	13	8	40
Los Angeles	7	9	10	16	42
Phoenix	10	11	7	14	42
Houston	14	6	14	9	43
Total	48	41	49	62	200

6.2: - CONFIGURE ROW-LEVEL SECURITY(RLS) FOR AIRLINE A DATA AND ASSIGN IT TO A USER.

STEP 1: To configure RLS, first I go to Modelling tab > Go to manage roles.

STEP 2: After clicking on manage roles, select the role first and rename it to ‘Airline A’, then select table “Flight_information” in which the ‘Airline’ column is and then filter data by selecting the column, give the condition “equals” and value “Airline A” and save it.

STEP 3: Then to check the role, I go to ‘modelling tab’ again and go to ‘view as’ and choose the option ‘Airline A’, then my whole report changes to ‘Airline A’ data.

STEP 4: Now I have to assign a role in Power BI service, before that I created a workspace in Power BI service and name it as “Airline Data”.

STEP 5: Then I published my report in my workspace and then I go to ‘semantic model’ and right click on it and select ‘Security’. After that I assign a role in RLS and put email id of that user which I assigned it to.

Manage security roles

Create new security roles and use filters to define row-level data restrictions.

Successfully applied role changes.

Roles

+ New

Airline A

Select tables

Best_Flights

flight_informat...

passenger_inf...

ticket_informat...

Filter data

Switch to DAX editor

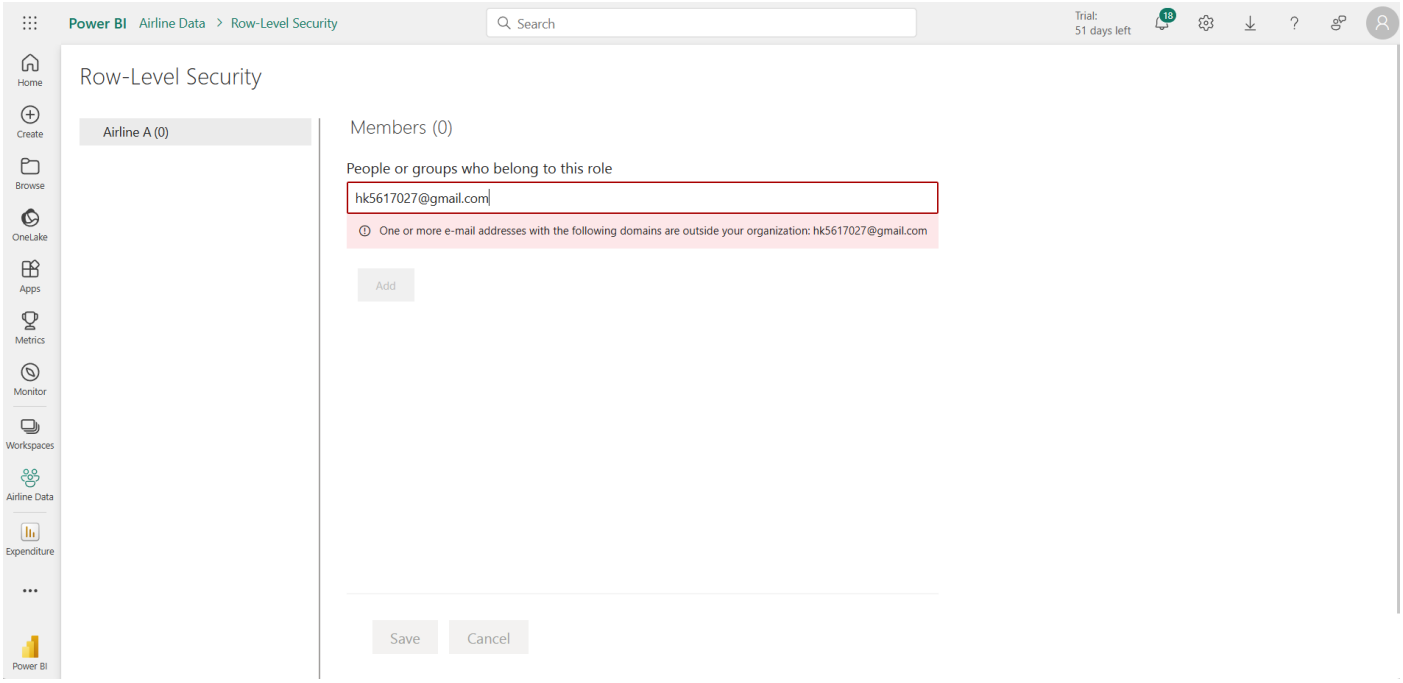
+ New | Select all | Delete | Group | Ungroup

Show data if All of these rules are true

Column	Condition	Value
Airline	Equals	Airline A

+ New

Save | Close

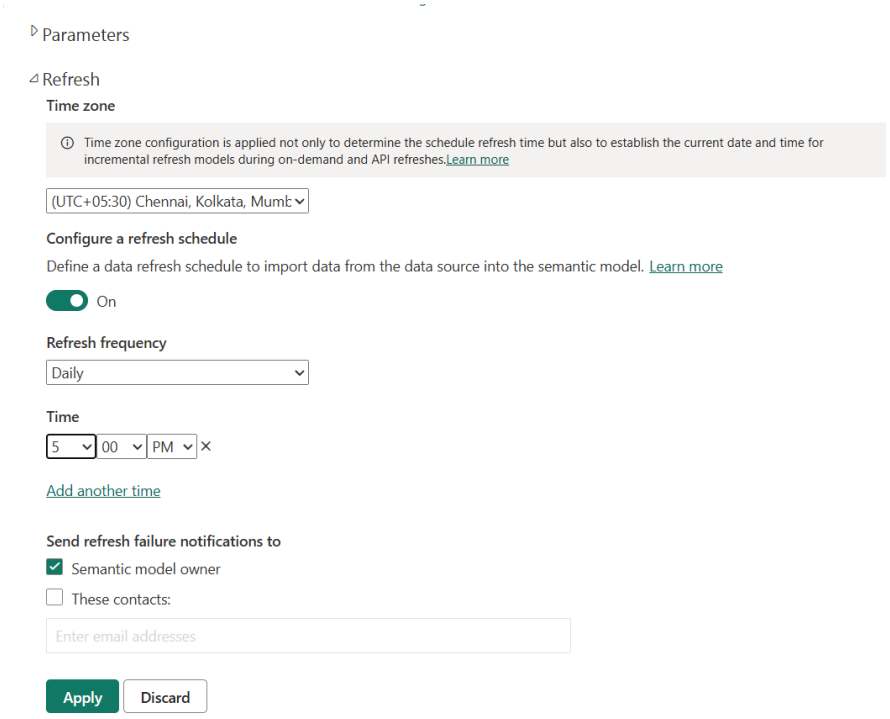


6.3: - SET UP A SCHEDULE REFRESH AT 5PM DAILY.

STEP 1: First, I open workspace in Power Bi service and go to Power BI setting.

STEP 2: After opening the power Bi setting > go to semantic model and go to “Refresh”.

STEP 3: There I change the time zone to New Delhi, turn on the refresh schedule button, change the refresh frequency to “Daily” and change the time to 5P.M and click on apply.



Video Explanation:
https://drive.google.com/file/d/1szSEKDMB5rXL0HVZujRWcFkILVZDQ5K4/view?usp=sharing