AIRLINE DATA MANAGEMENT AND ANALYSIS

POWER BI PROJECT

ABSTRACT

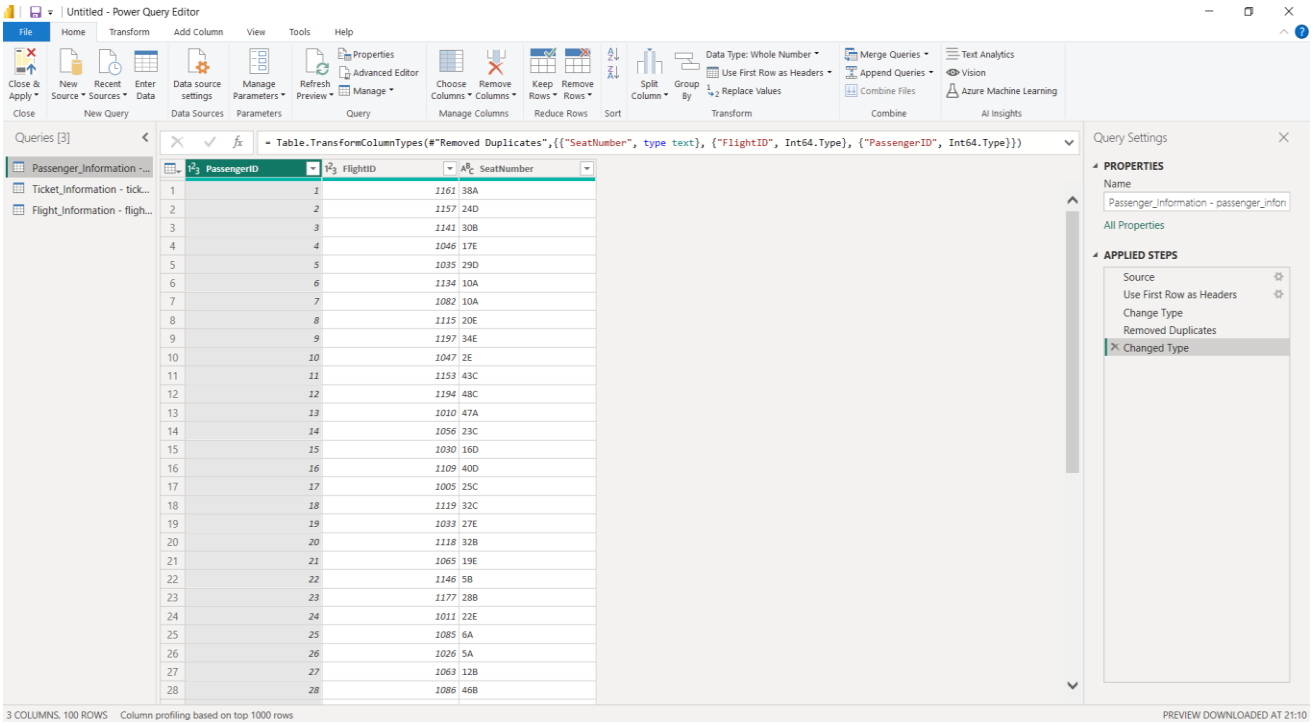
To analyse and visualize airline data for operational insights, passenger management, and ticket booking trends using Power BI.

Devina M Vasudevan

DATA SCIENCE PGC JUNE 1ST BATCH

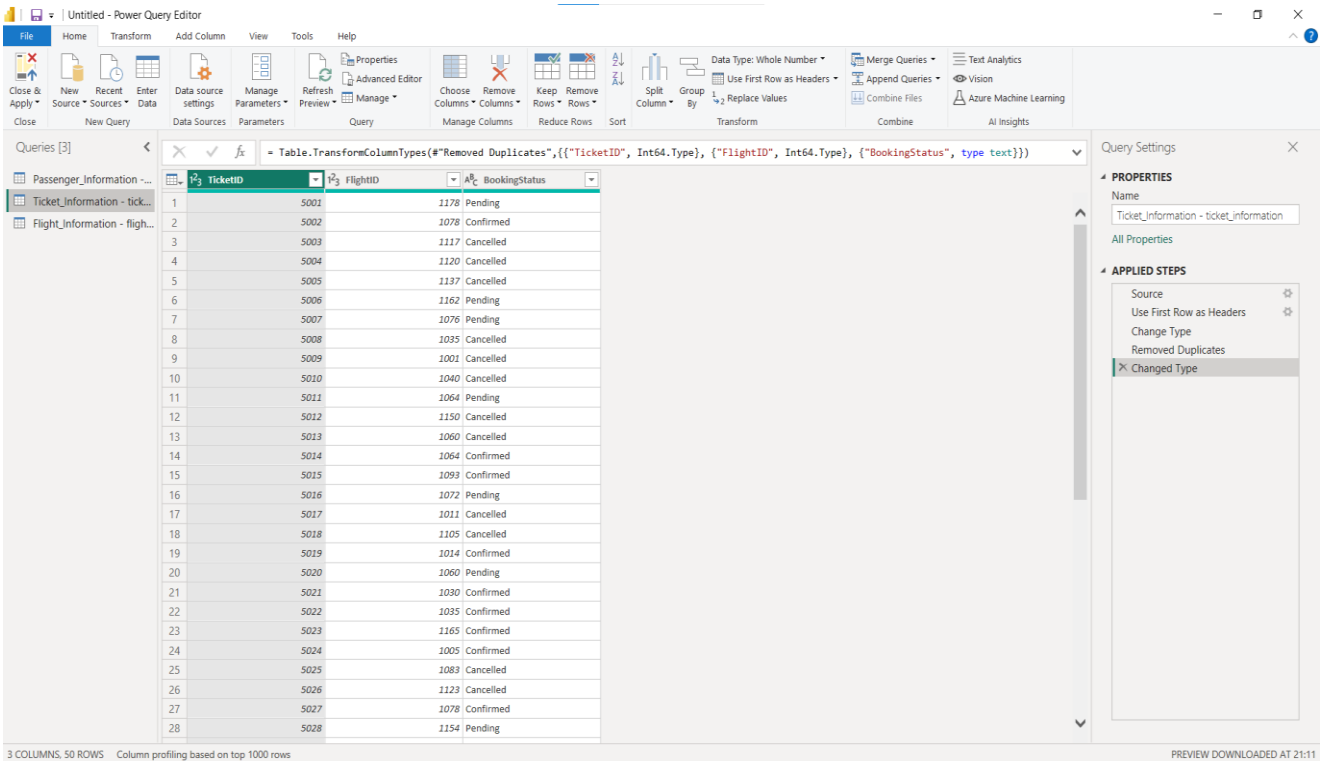
TASK 1: DATA PREPARATION AND CLEANING

PASSENGER INFORMATION



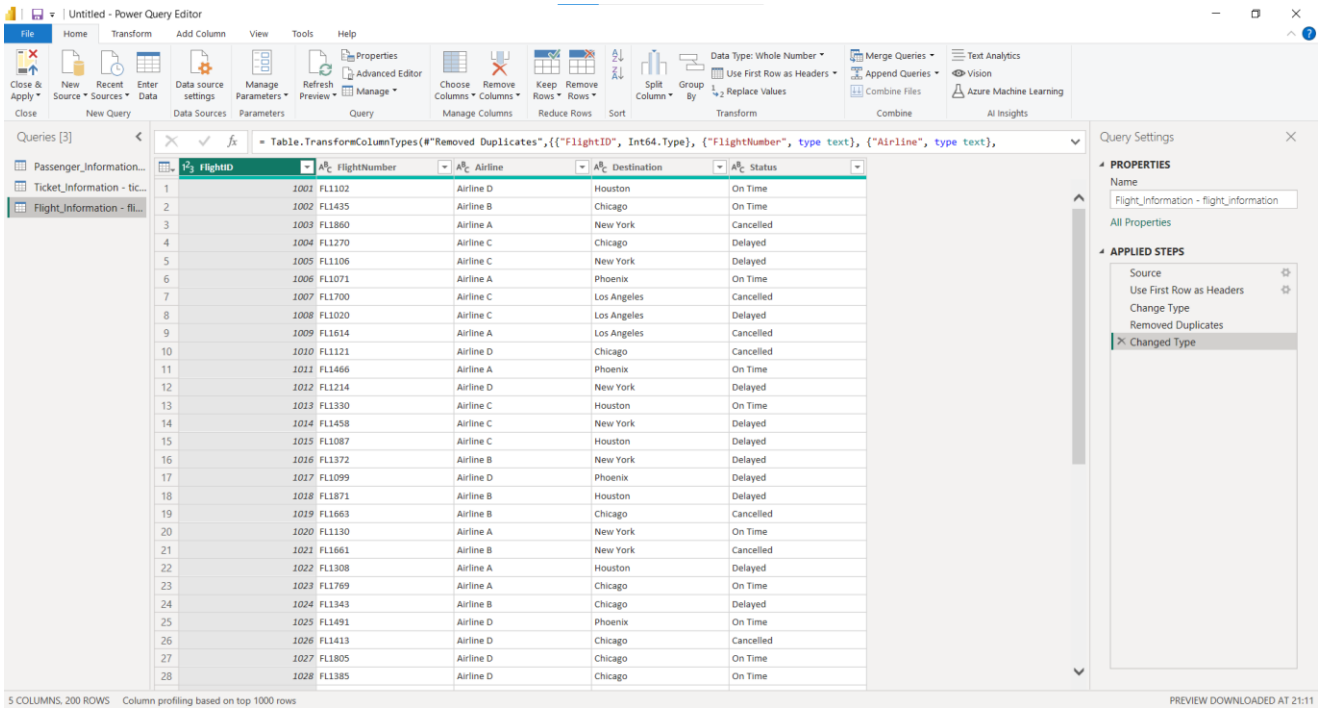
	PassengerID	FlightID	SeatNumber
1	1	1161	38A
2	2	1157	24D
3	3	1141	30B
4	4	1046	17E
5	5	1035	29D
6	6	1134	10A
7	7	1082	10A
8	8	1115	20E
9	9	1197	34E
10	10	1047	2E
11	11	1153	43C
12	12	1194	48C
13	13	1010	47A
14	14	1056	23C
15	15	1030	16D
16	16	1109	40D
17	17	1005	25C
18	18	1119	32C
19	19	1033	27E
20	20	1118	32B
21	21	1065	19E
22	22	1146	5B
23	23	1177	28B
24	24	1011	22E
25	25	1085	6A
26	26	1026	5A
27	27	1063	12B
28	28	1086	46B

TICKET INFORMATION



	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
10	5010	1040	Cancelled
11	5011	1064	Pending
12	5012	1150	Cancelled
13	5013	1060	Cancelled
14	5014	1064	Confirmed
15	5015	1093	Confirmed
16	5016	1072	Pending
17	5017	1011	Cancelled
18	5018	1105	Cancelled
19	5019	1014	Confirmed
20	5020	1060	Pending
21	5021	1030	Confirmed
22	5022	1035	Confirmed
23	5023	1165	Confirmed
24	5024	1005	Confirmed
25	5025	1083	Cancelled
26	5026	1123	Cancelled
27	5027	1078	Confirmed
28	5028	1154	Pending

FLIGHT INFORMATION

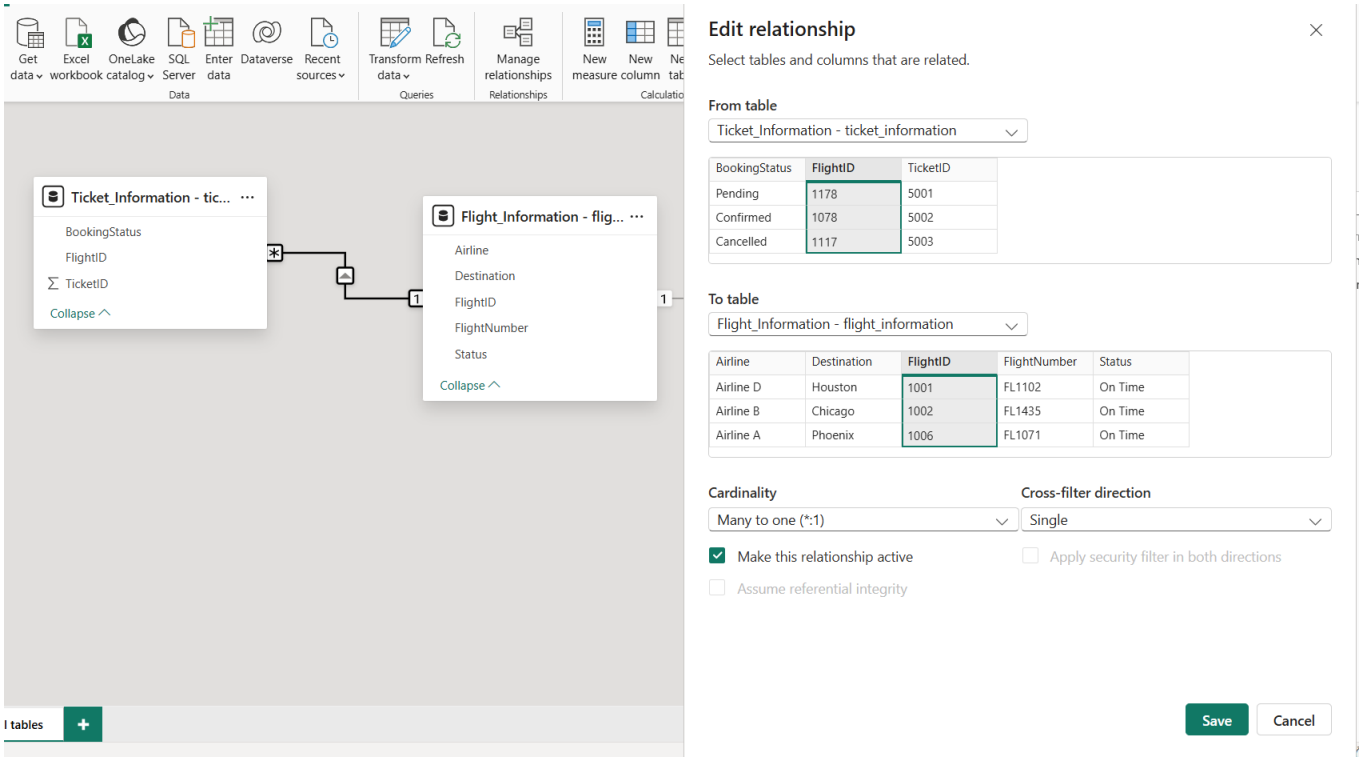


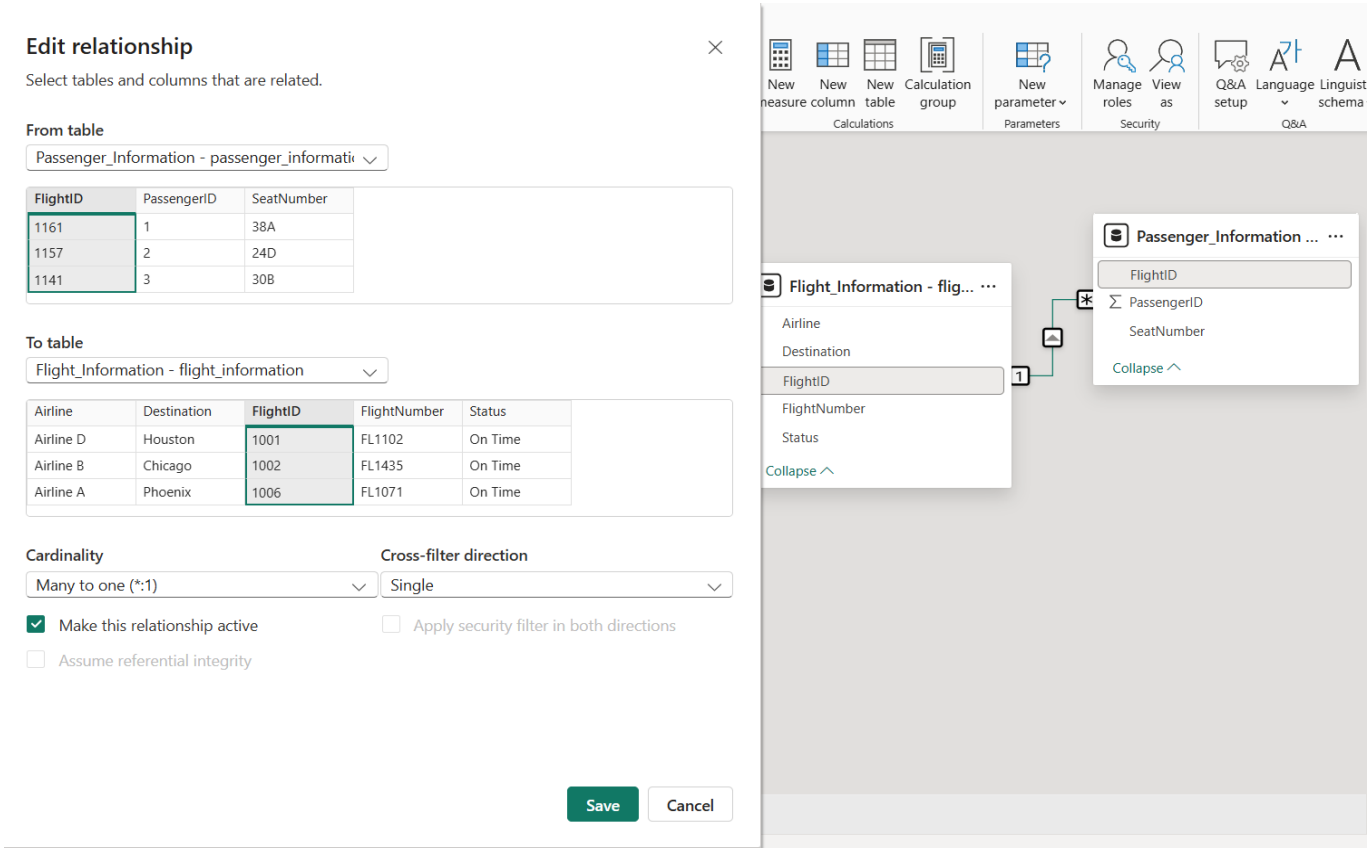
	FlightID	FlightNumber	Airline	Destination	Status
1	1001	FL1102	Airline D	Houston	On Time
2	1002	FL1435	Airline B	Chicago	On Time
3	1003	FL1860	Airline A	New York	Cancelled
4	1004	FL1270	Airline C	Chicago	Delayed
5	1005	FL1106	Airline C	New York	Delayed
6	1006	FL1071	Airline A	Phoenix	On Time
7	1007	FL1700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

TASK 2: DATA MODELING

RELATIONSHIP BETWEEN DATASETS:

MANY TO ONE RELATION

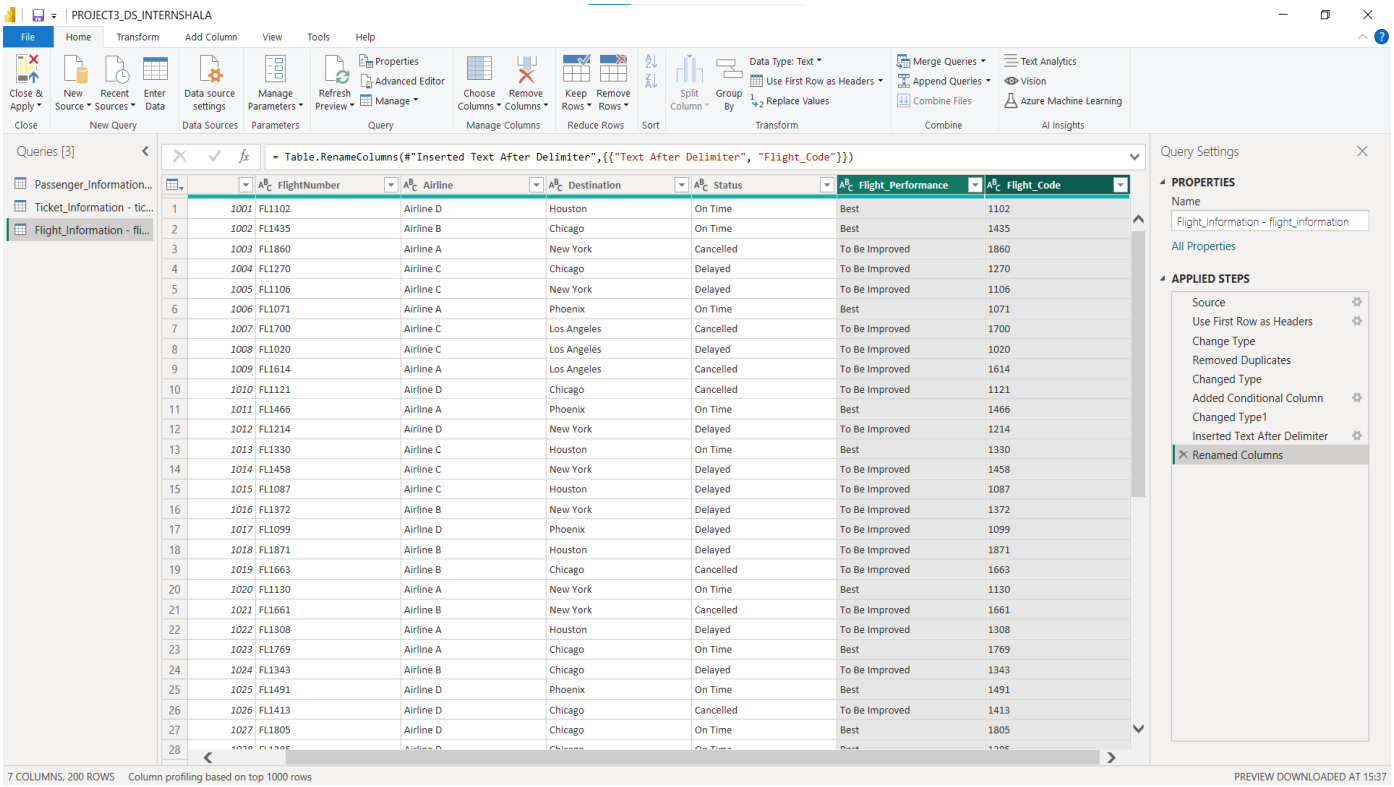




TASK 3: ENHANCED DATA INSIGHTS

Two columns added:

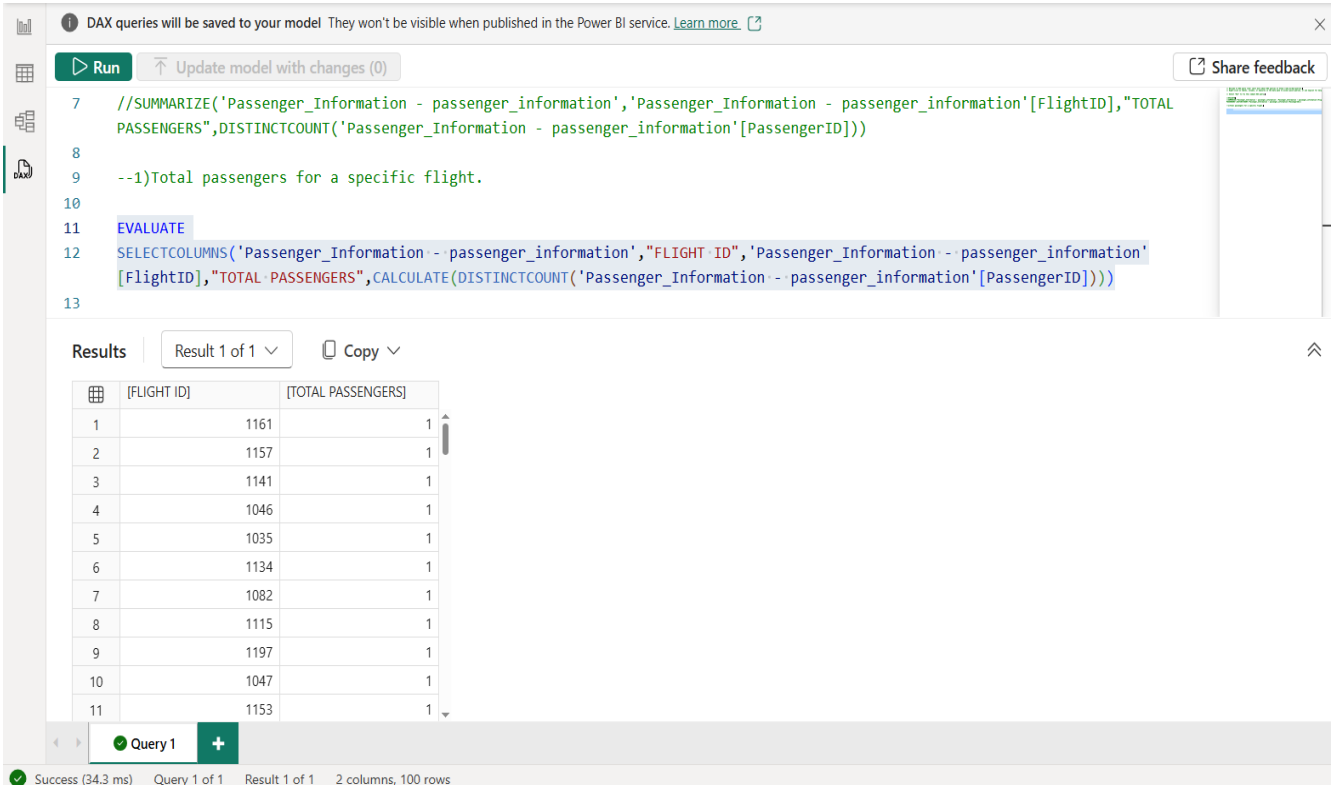
Flight_Performance with Conditional Column and Flight code with Column From Example



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

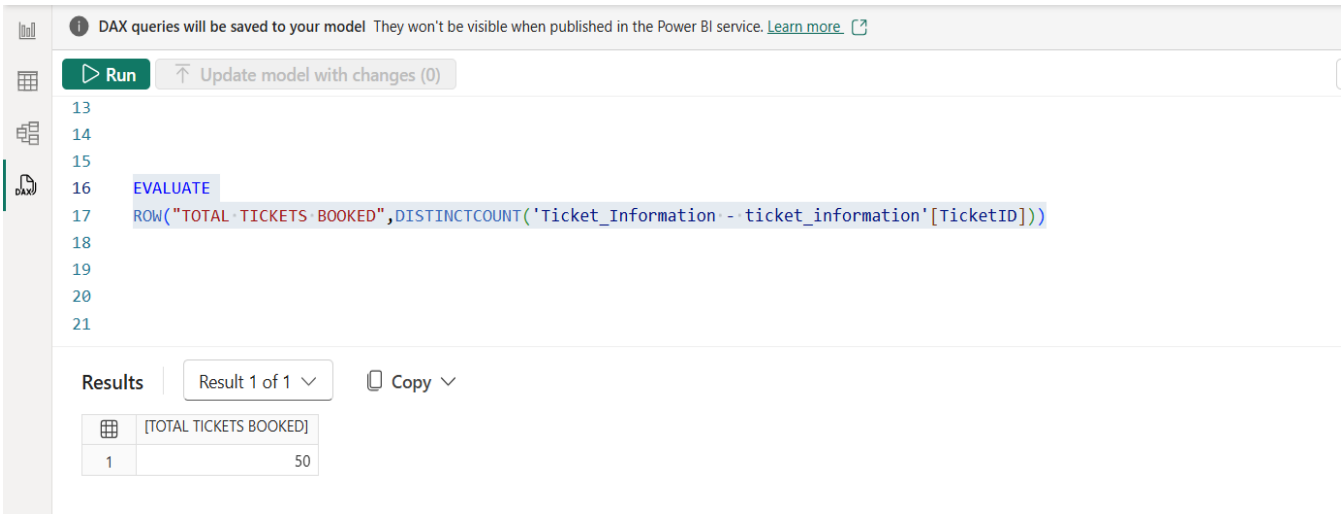
TASK 4: CALCULATIONS USING DAX

TOTAL PASSENGERS FOR A SPECIFIC FLIGHT



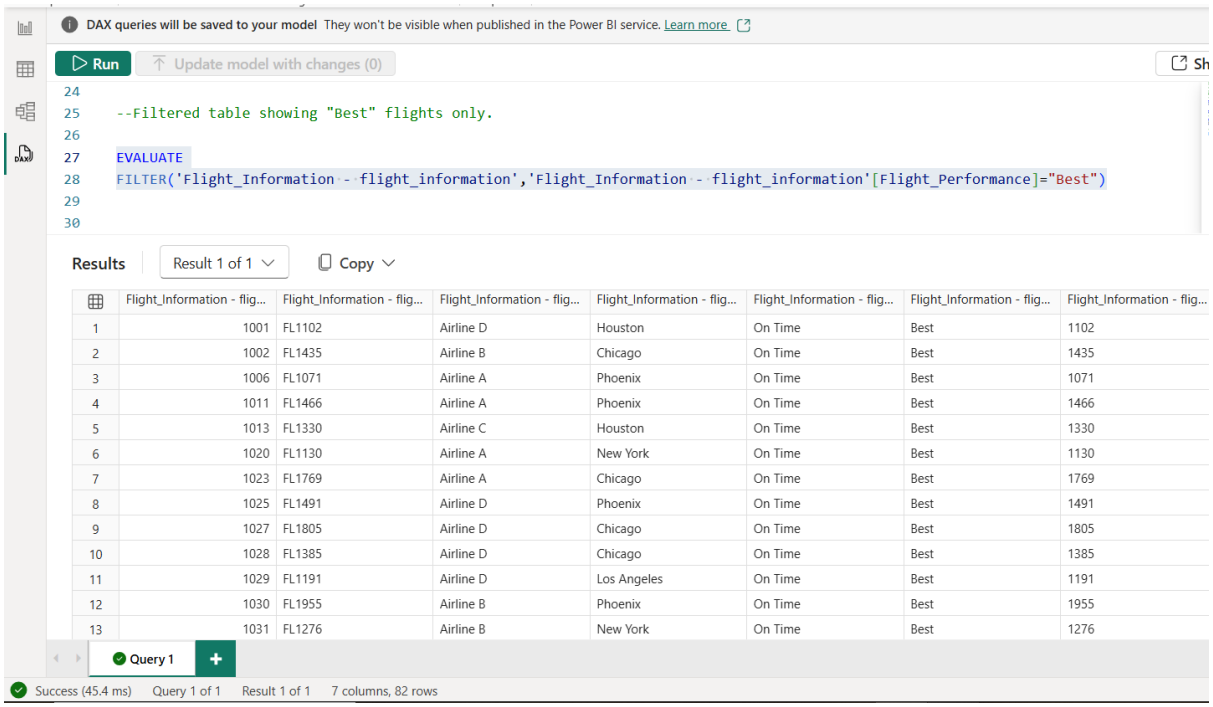
	[FLIGHT ID]	[TOTAL PASSENGERS]
1	1161	1
2	1157	1
3	1141	1
4	1046	1
5	1035	1
6	1134	1
7	1082	1
8	1115	1
9	1197	1
10	1047	1
11	1153	1

TOTAL TICKETS BOOKED



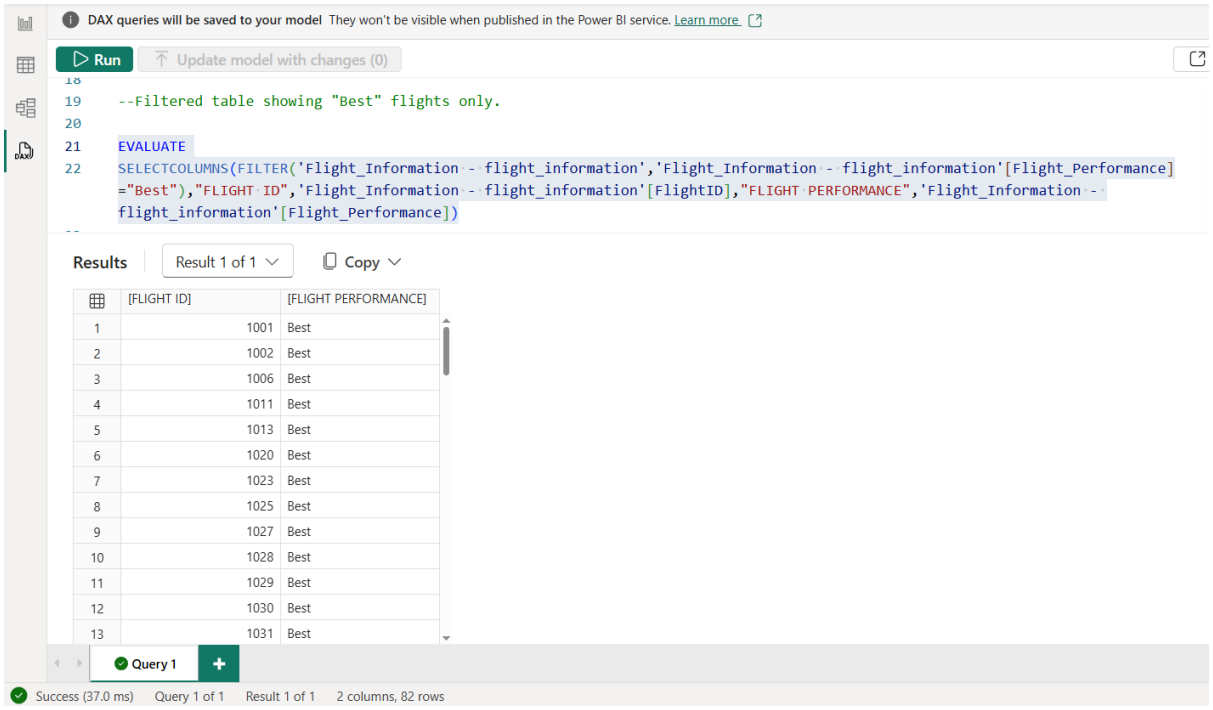
	[TOTAL TICKETS BOOKED]
1	50

FILTERED TABLE SHOWING ALL INFORMATION OF BEST FLIGHTS ONLY



	Flight_Information - flig...	Flight_Information - flig...	Flight_Information - flig...	Flight_Information - flig...	Flight_Information - flig...	Flight_Information - flig...	Flight_Information - flig...
1	1001	FL1102	Airline D	Houston	On Time	Best	1102
2	1002	FL1435	Airline B	Chicago	On Time	Best	1435
3	1006	FL1071	Airline A	Phoenix	On Time	Best	1071
4	1011	FL1466	Airline A	Phoenix	On Time	Best	1466
5	1013	FL1330	Airline C	Houston	On Time	Best	1330
6	1020	FL1130	Airline A	New York	On Time	Best	1130
7	1023	FL1769	Airline A	Chicago	On Time	Best	1769
8	1025	FL1491	Airline D	Phoenix	On Time	Best	1491
9	1027	FL1805	Airline D	Chicago	On Time	Best	1805
10	1028	FL1385	Airline D	Chicago	On Time	Best	1385
11	1029	FL1191	Airline D	Los Angeles	On Time	Best	1191
12	1030	FL1955	Airline B	Phoenix	On Time	Best	1955
13	1031	FL1276	Airline B	New York	On Time	Best	1276

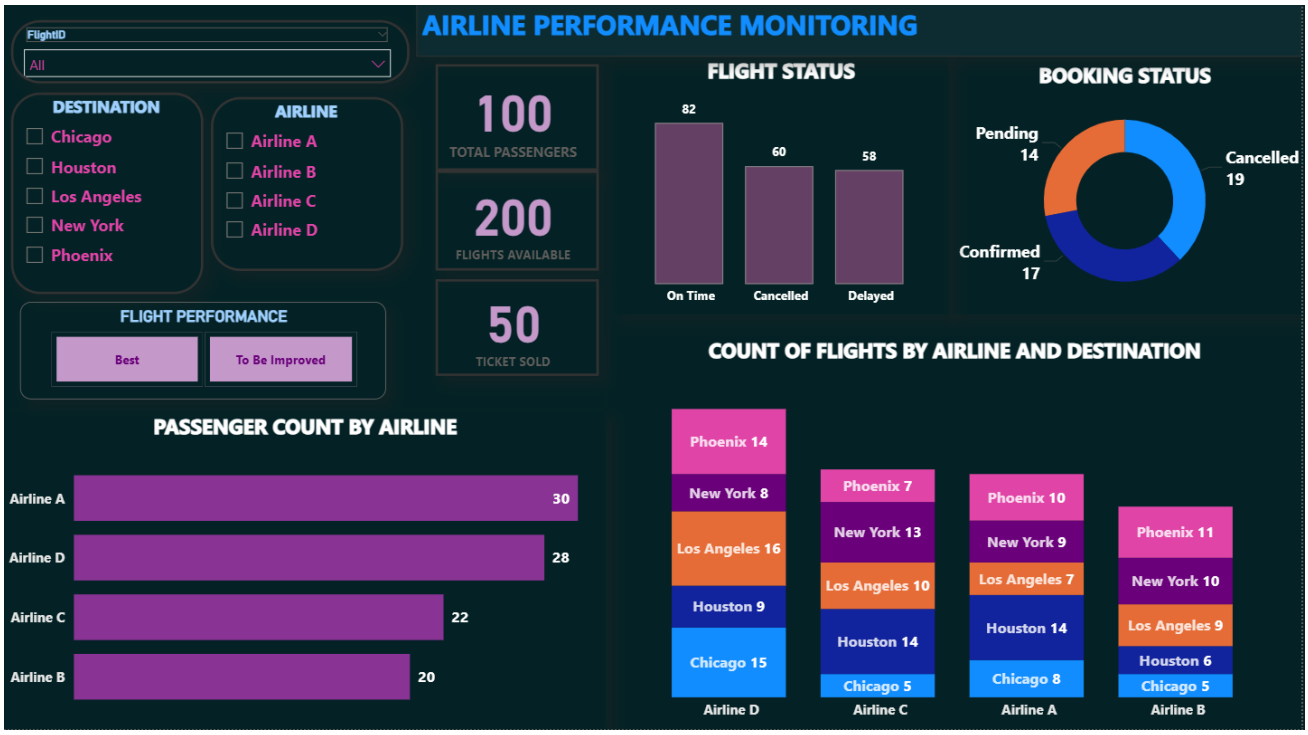
FILTERED TABLE SHOWING BEST FLIGHTS AND FLIGHT ID'S ONLY



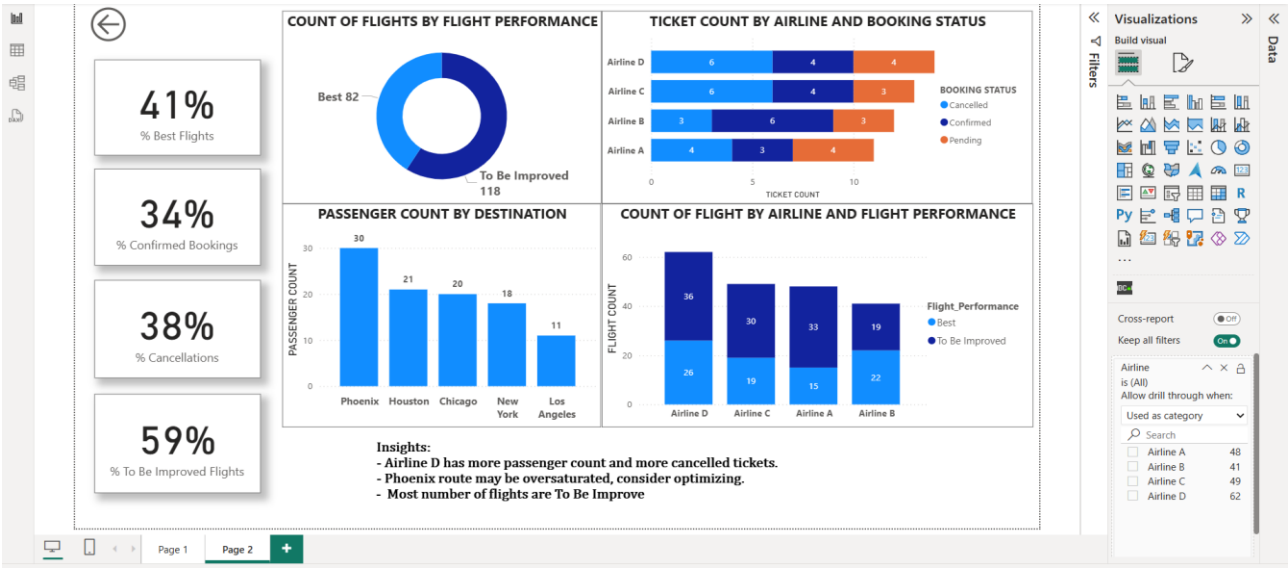
	[FLIGHT ID]	[FLIGHT PERFORMANCE]
1	1001	Best
2	1002	Best
3	1006	Best
4	1011	Best
5	1013	Best
6	1020	Best
7	1023	Best
8	1025	Best
9	1027	Best
10	1028	Best
11	1029	Best
12	1030	Best
13	1031	Best

TASK 5: VISUALIZATION AND INTERACTIVE FEATURES

DASHBOARD

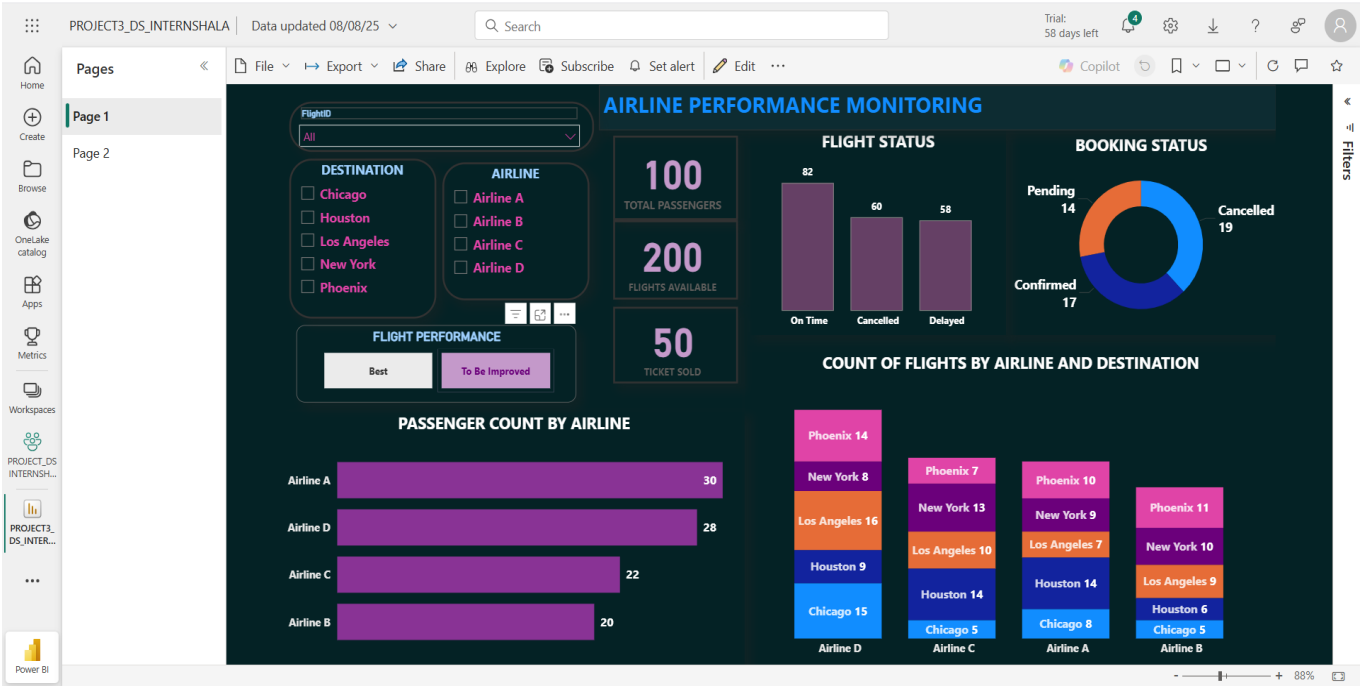


- Created Bar Chart for analysing the count of passengers by Airline.
- Created Donut Chart for analysing the Booking Status.
- Created Stacked column chart for Number of Flights by airline and destination.
- Created Stacked column chart for Flight Status.
- Created 3 Cards for ‘Total Passengers’, ‘Total Flights Available’, ’Ticket Sold’.
- Created 4 Slicers for Airline, Destination, Flight id and Flight Performance

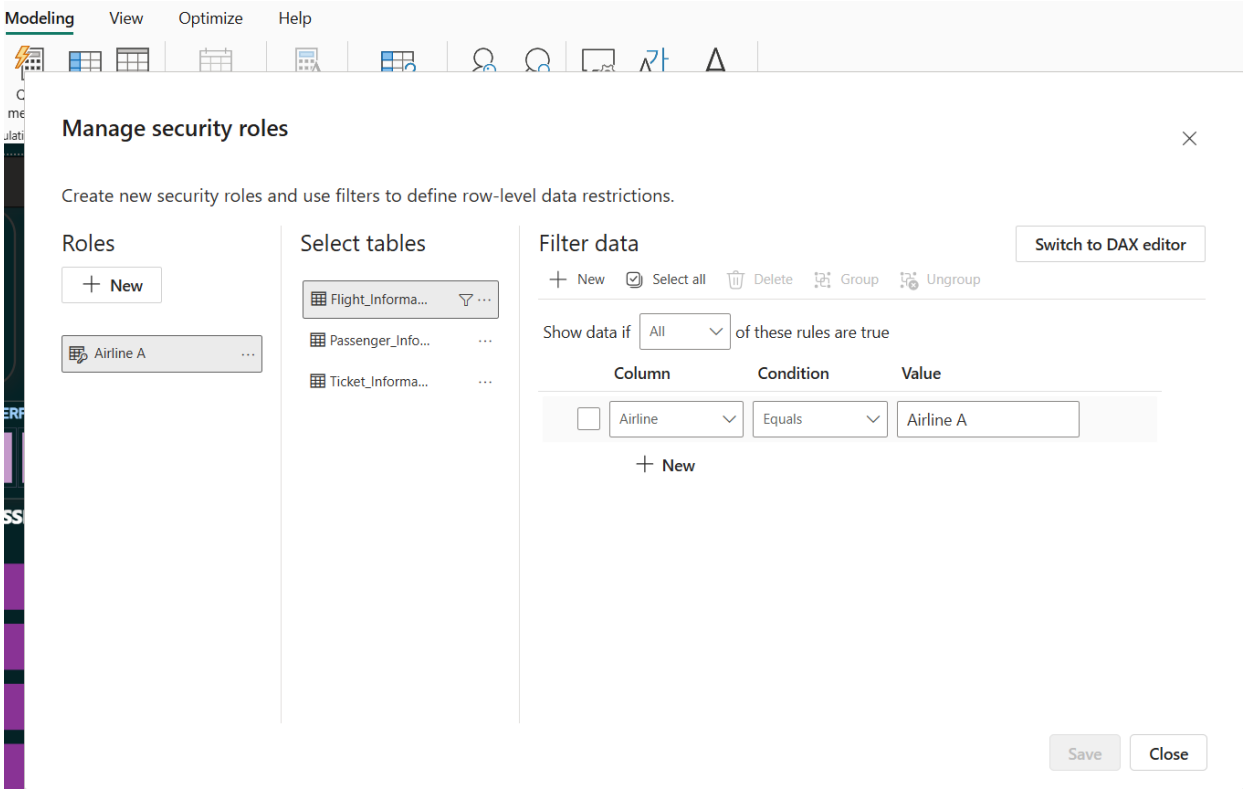


TASK 6: FINAL DASHBOARD AND POWER BI SERVICE

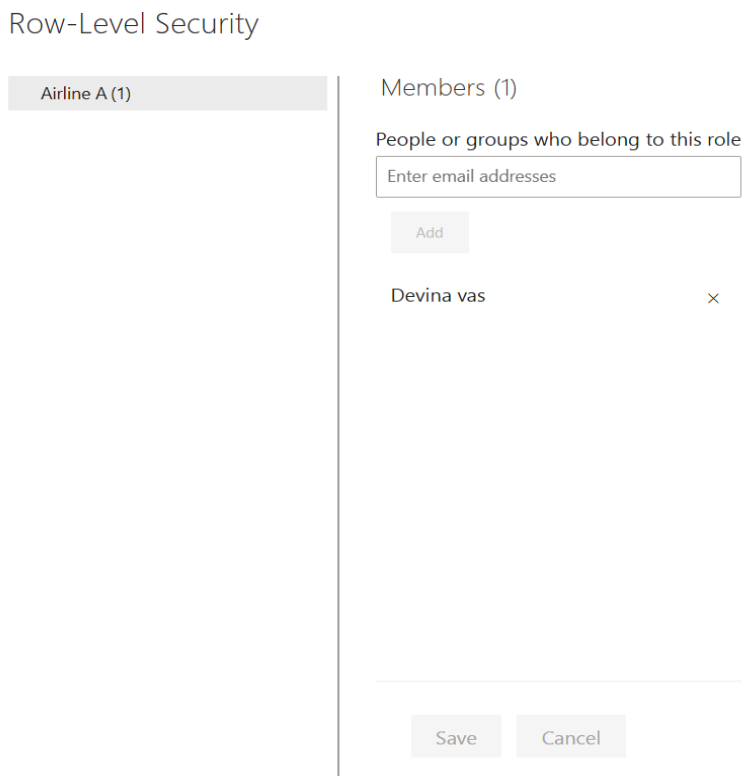
DESIGNED A COMPREHENSIVE DASHBOARD WITH KEY VISUALS AND INSIGHTS IN POWER BI SERVICE



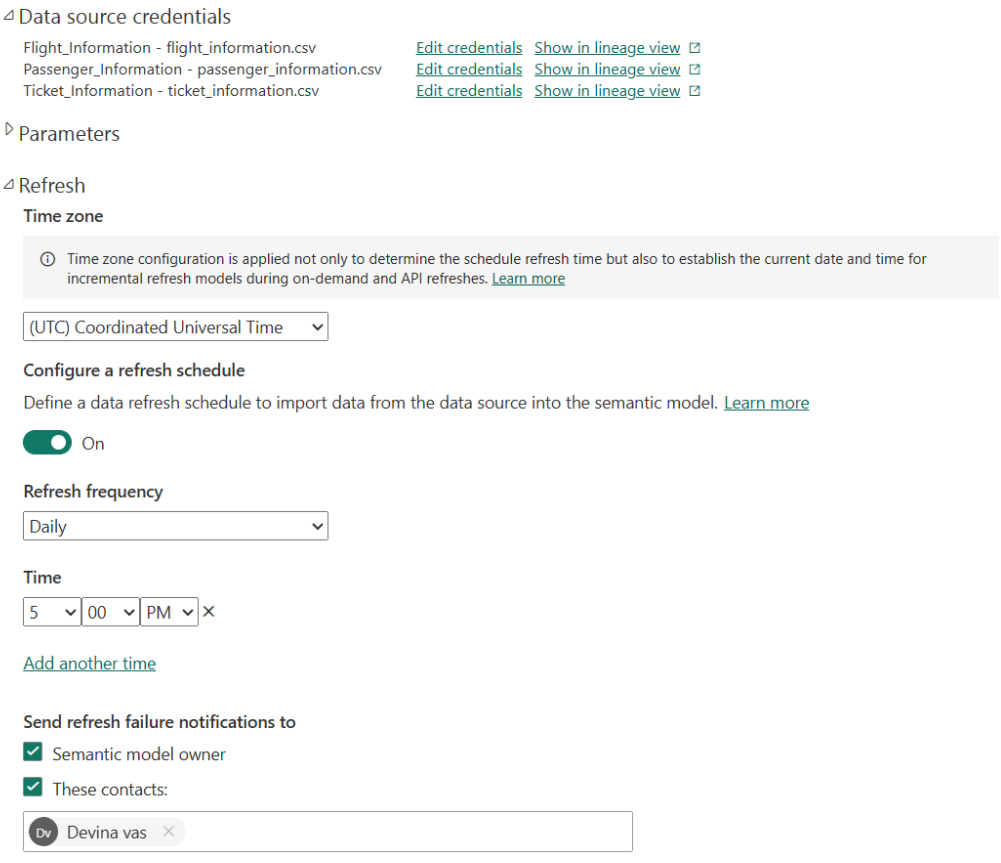
CONFIGURED **ROW-LEVEL SECURITY** FOR **AIRLINE A** DATA AND ASSIGN IT TO A USER:



SET UP ROW LEVEL SECURITY IN POWER BI SERVICE:



SET UP A SCHEDULE **REFRESH AT 5 PM DAILY**:



CONCLUSION

- **Flight Performance**: Only 41% of flights are marked as ‘Best’, with the majority needing improvement.
- **Booking Status**: A significant portion of tickets (38%) are cancelled, and only 34% are confirmed, highlighting inefficiencies in the booking process. The ticket booking process needs enhancement to reduce cancellations.
- **Passenger Trends**: The route to Phoenix shows maximum passenger traffic and flight frequency, requiring better traffic distribution.
- **Airline D** has a greater number of passengers and flights but also records more cancellations and underperformance, requiring urgent system review.
- The data reveals an irregularity, where **ticket purchase** figures are either higher or lower than the actual number of passengers traveling.

- THANK YOU -