DATA SCIENCE

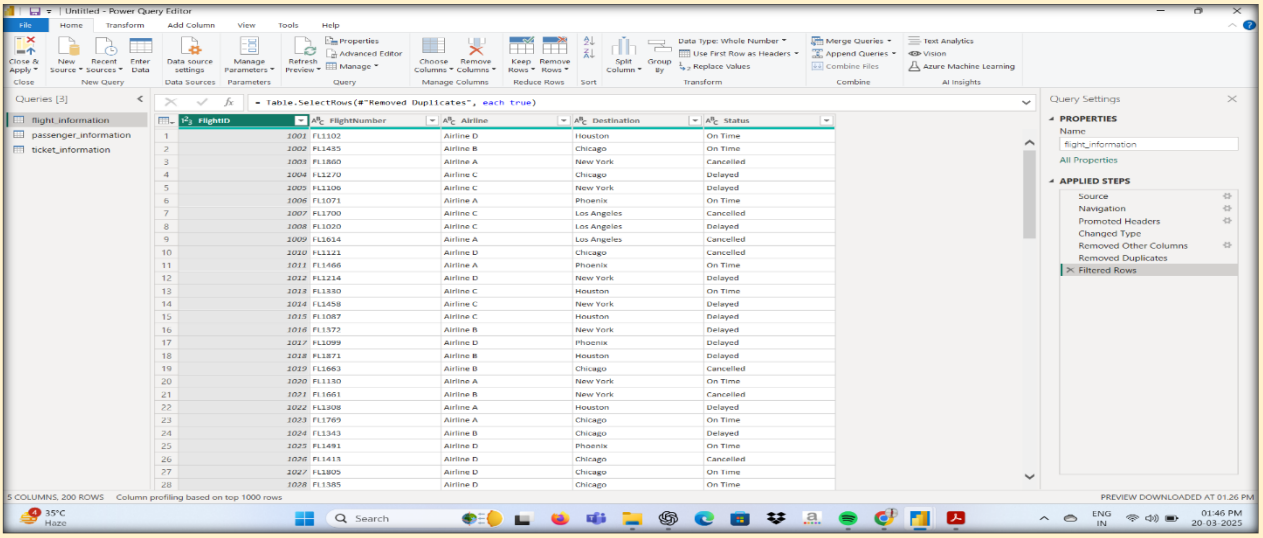
POWER BI PROJECT

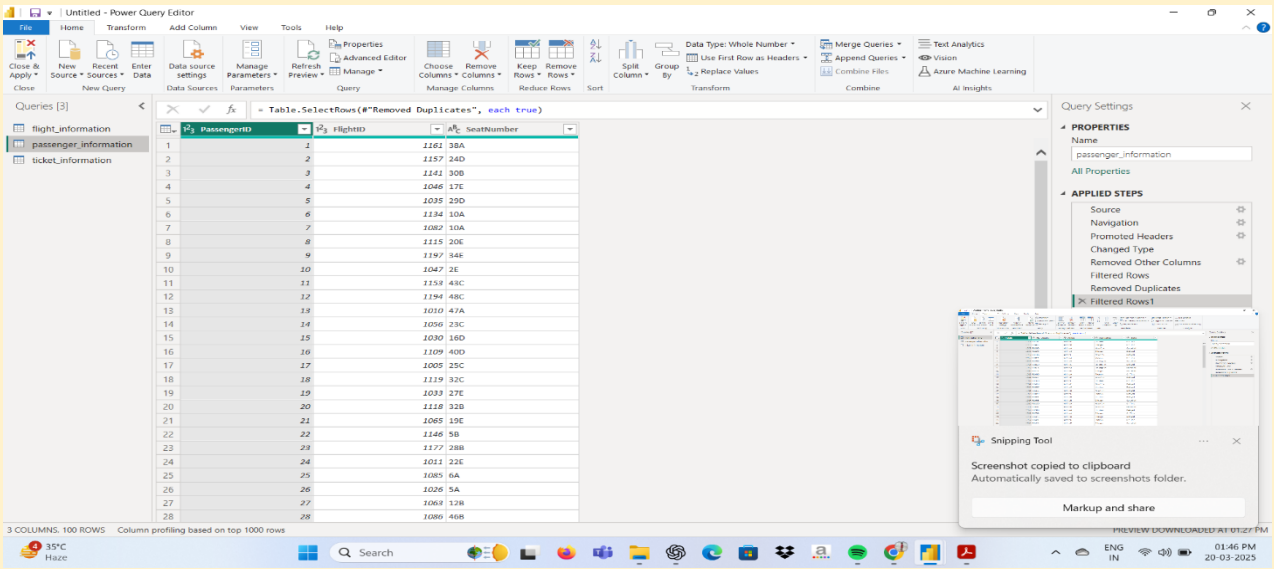
Task 1. Data Preparation and Cleaning

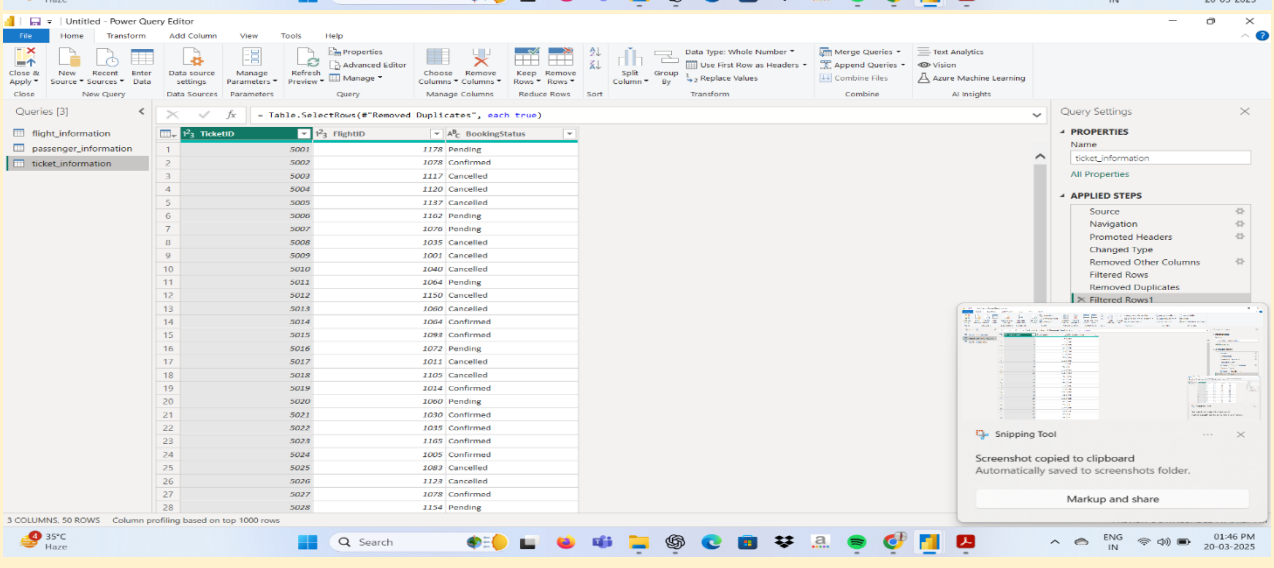
Steps taken :

Step 1 : Home

Step 2: Remove duplicates/Blank rows/Format

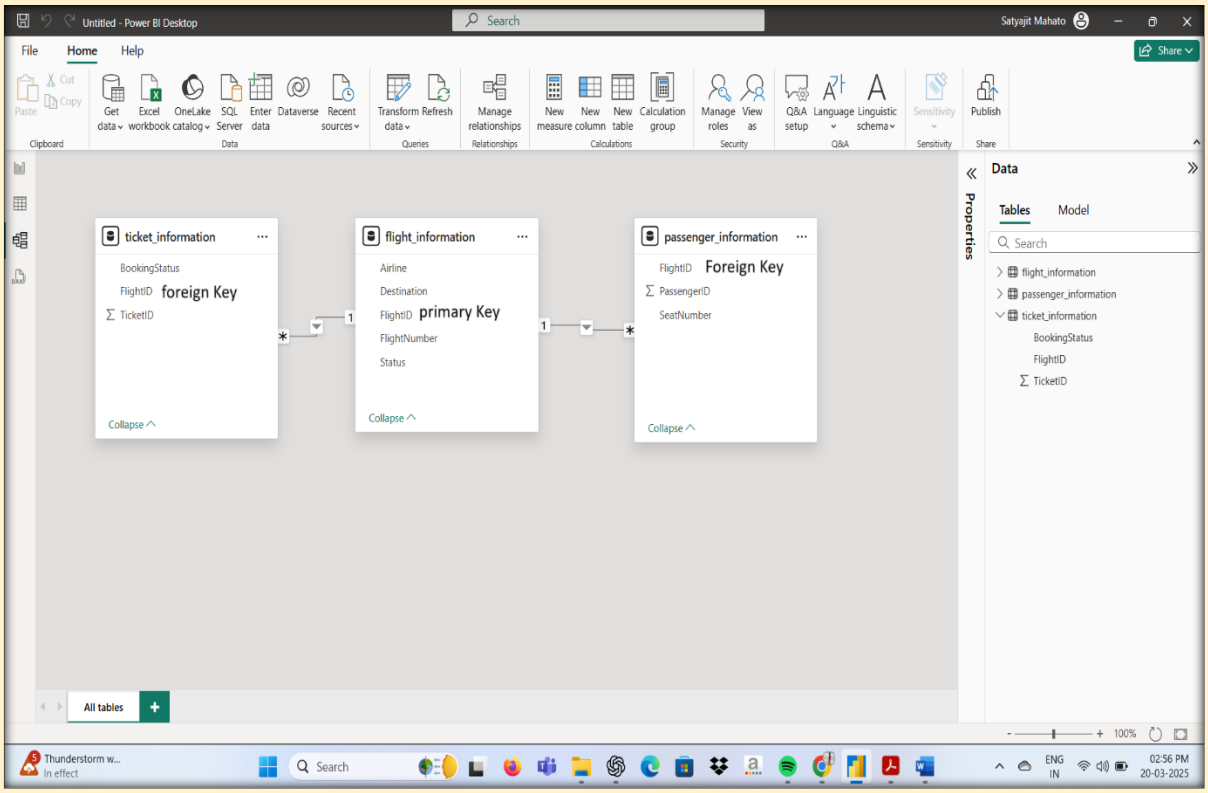


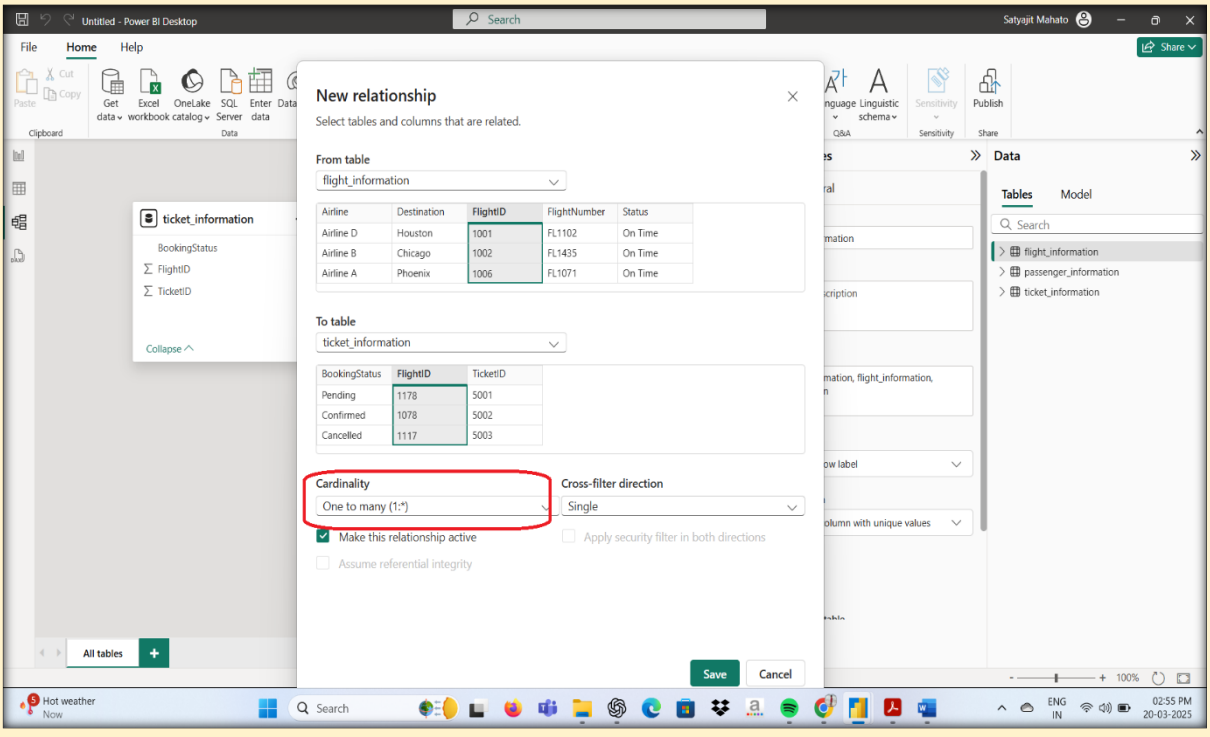




Task 2. Data Modeling

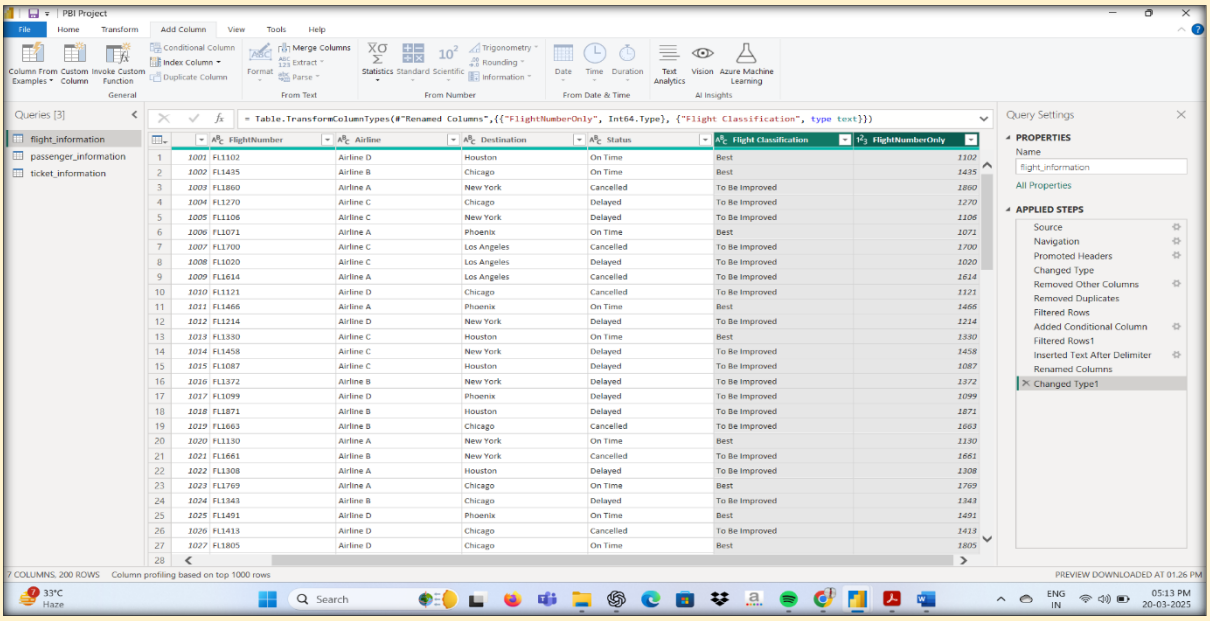
Flight_ID column is unique in flight_information which means all IDs are unique whereas flight_ID column in ticket_information and passenger_information query has repeated flight_ID values.





Task 3: Enhanced Data Insights

- Conditional Column displaying classification of flights as “Best” or “To be improved”
- Used “column from example” option to extract flight number from flight number

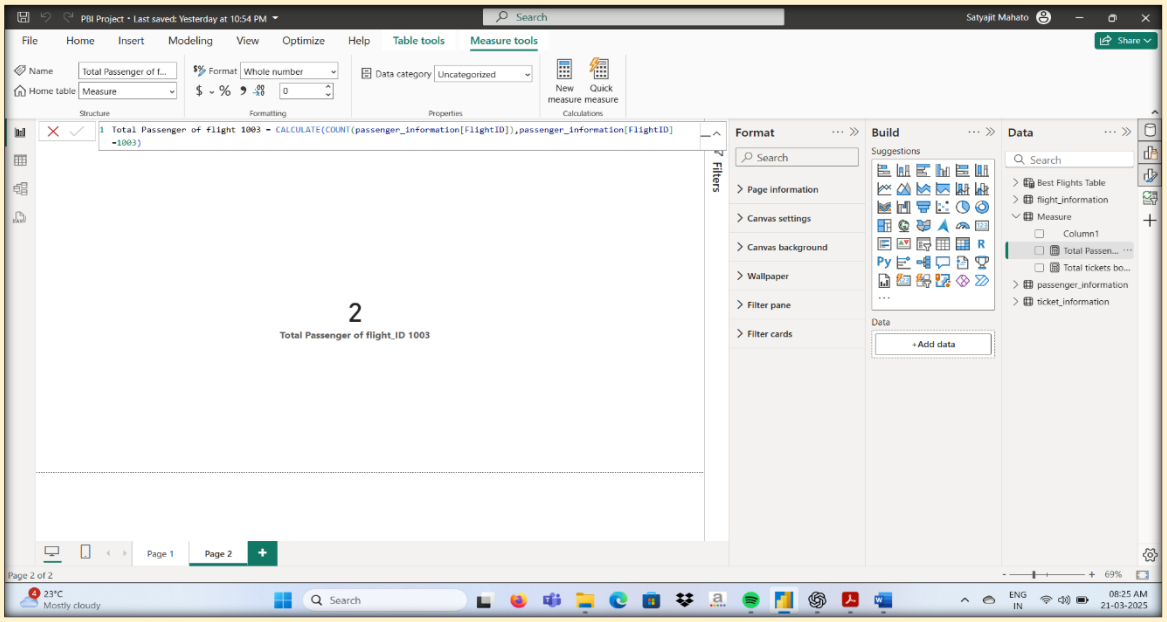


Task 4 :Calculations Using DAX

- Total passengers for a specific flight.

DAX function

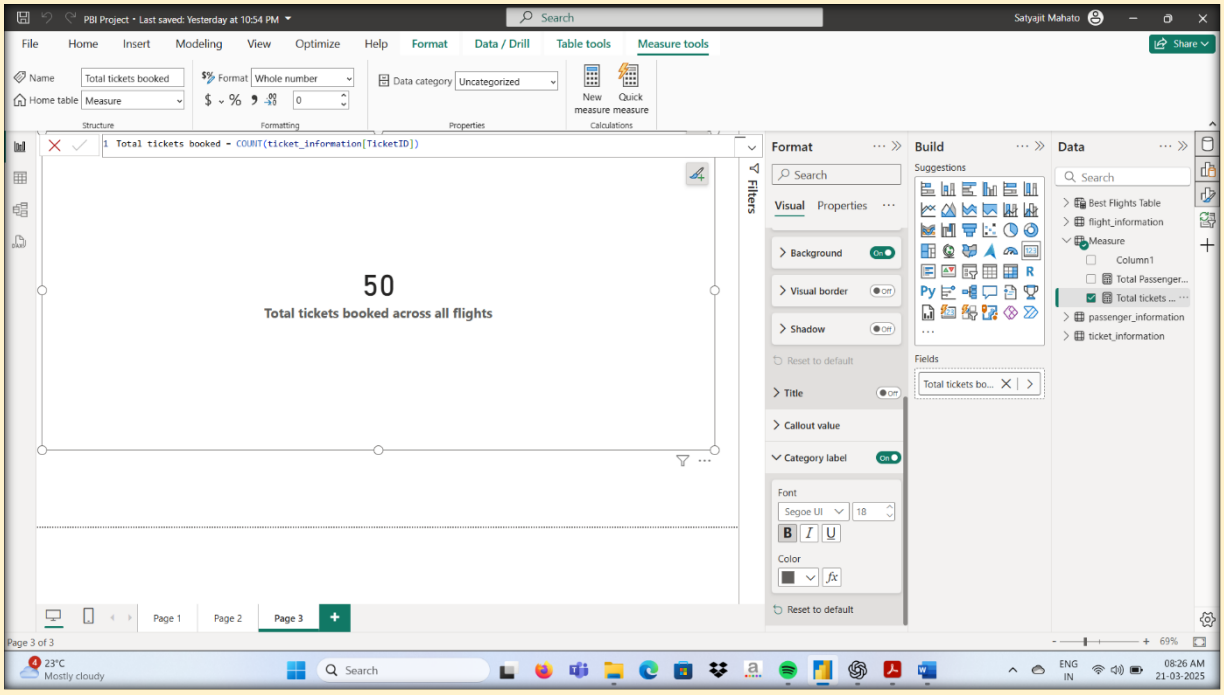
Total Passenger of flight 1003 =
CALCULATE(COUNT(passenger_information[FlightID]),passenger_information[FlightID]=1003)



- Total tickets booked across all flights.

DAX function

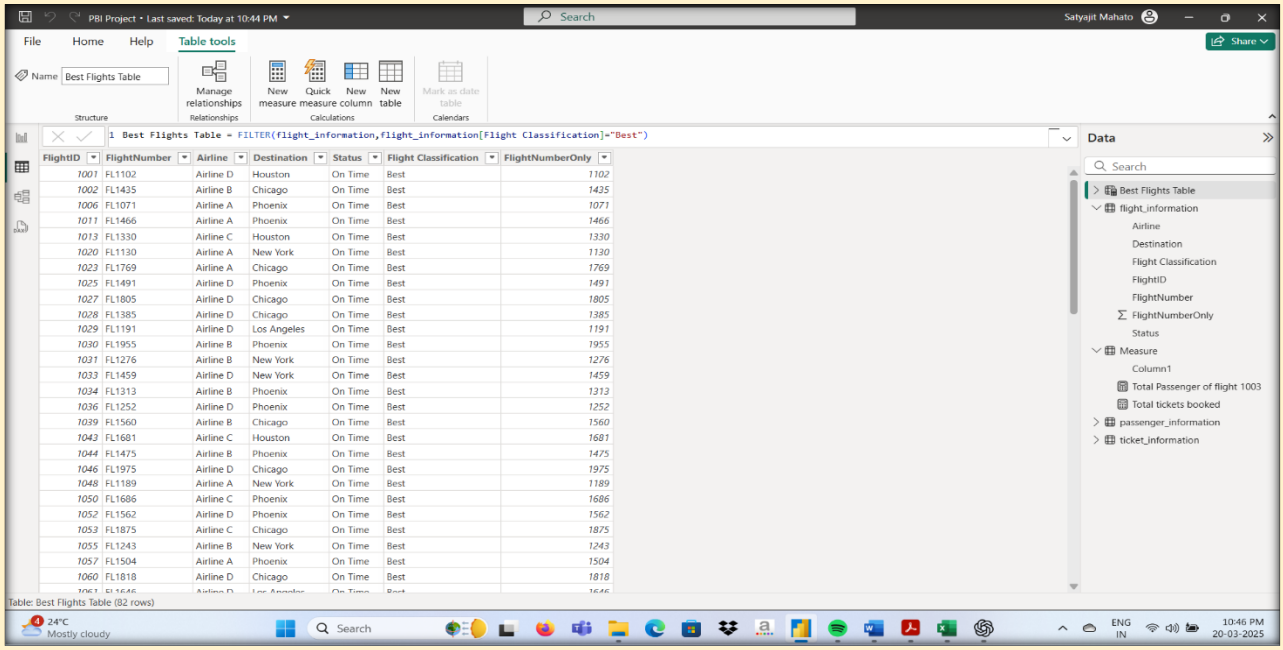
Total tickets booked = COUNT(ticket_information[TicketID])



- Filtered table showing "Best" flights only.

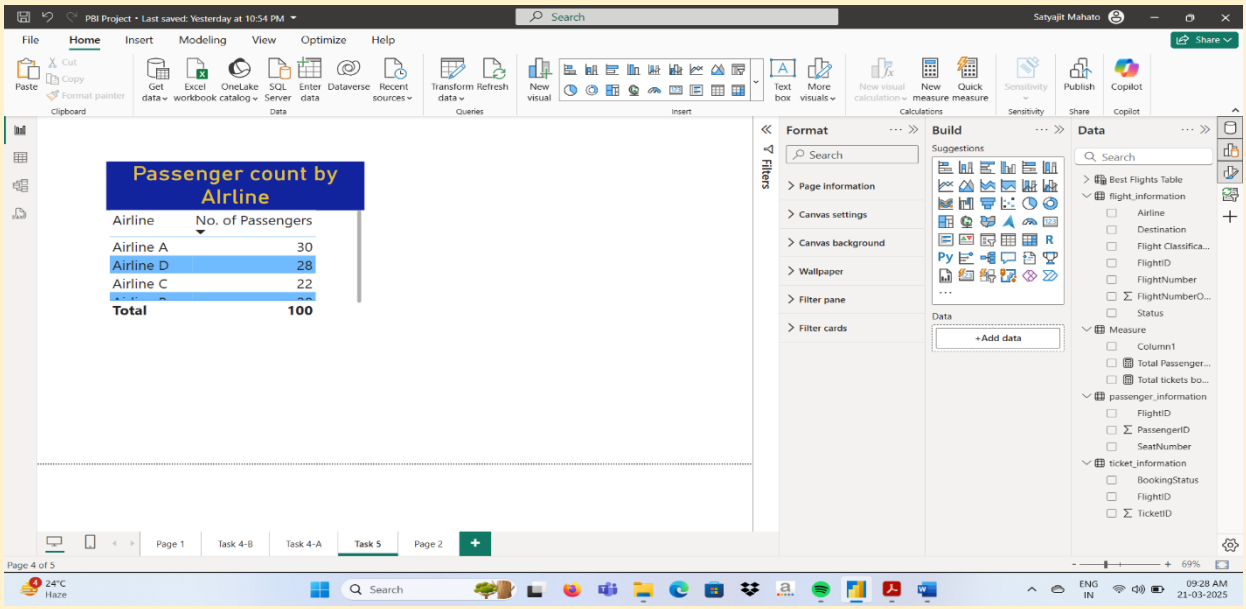
DAX Function

Best Flights Table = FILTER(flight_information,flight_information[Flight Classification]="Best")

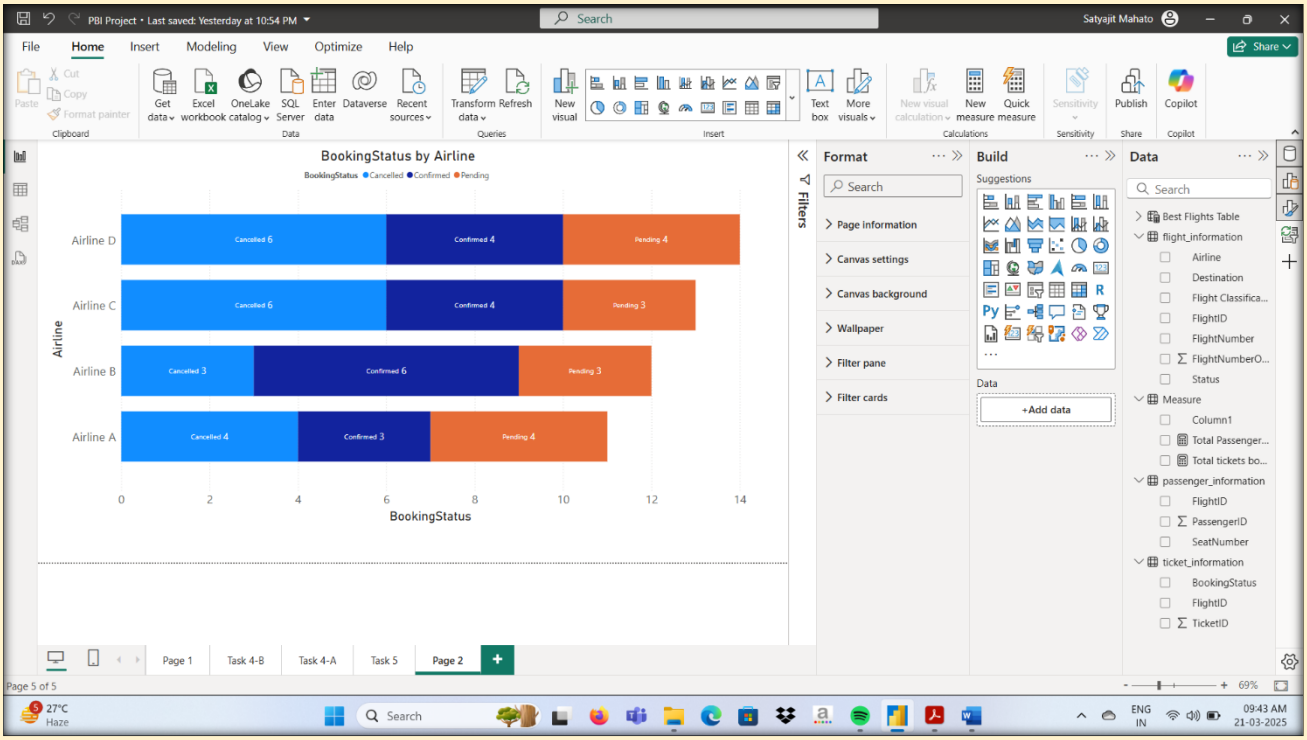


Task 5. Visualization and Interactive Features

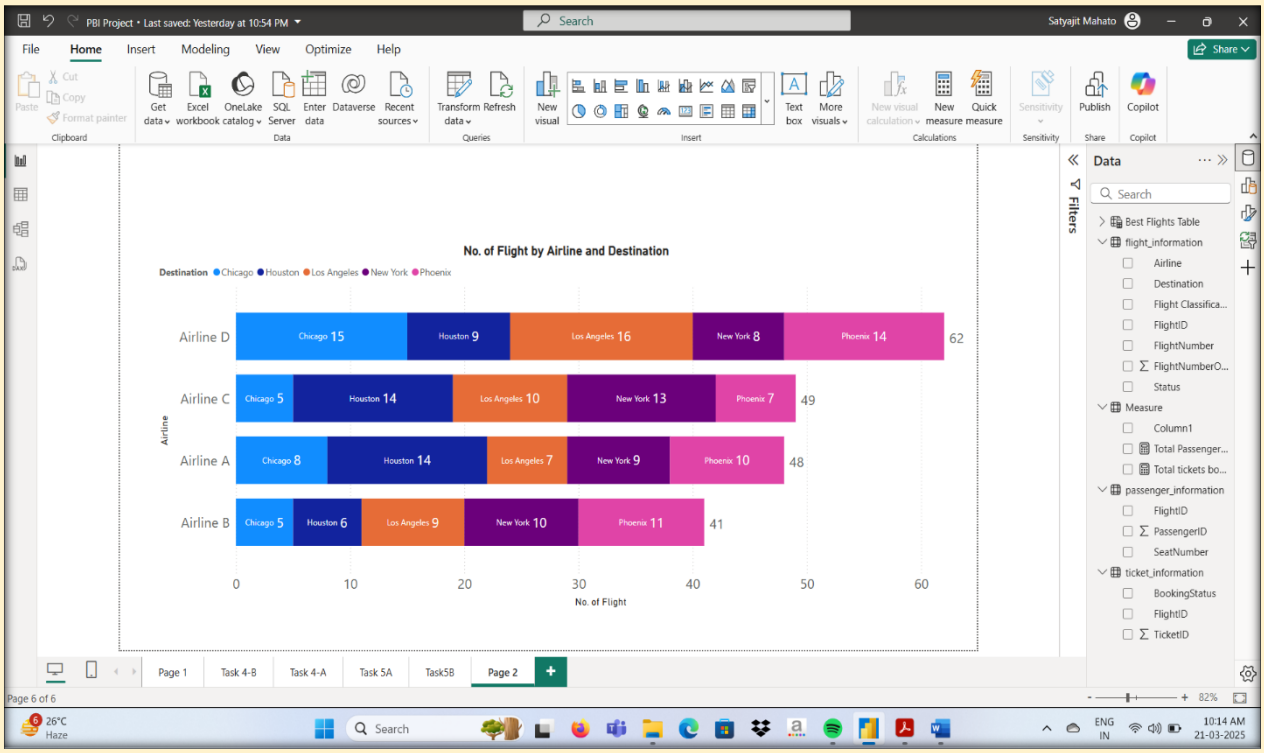
○ Passenger count by airline.



○ Ticket booking statuses.

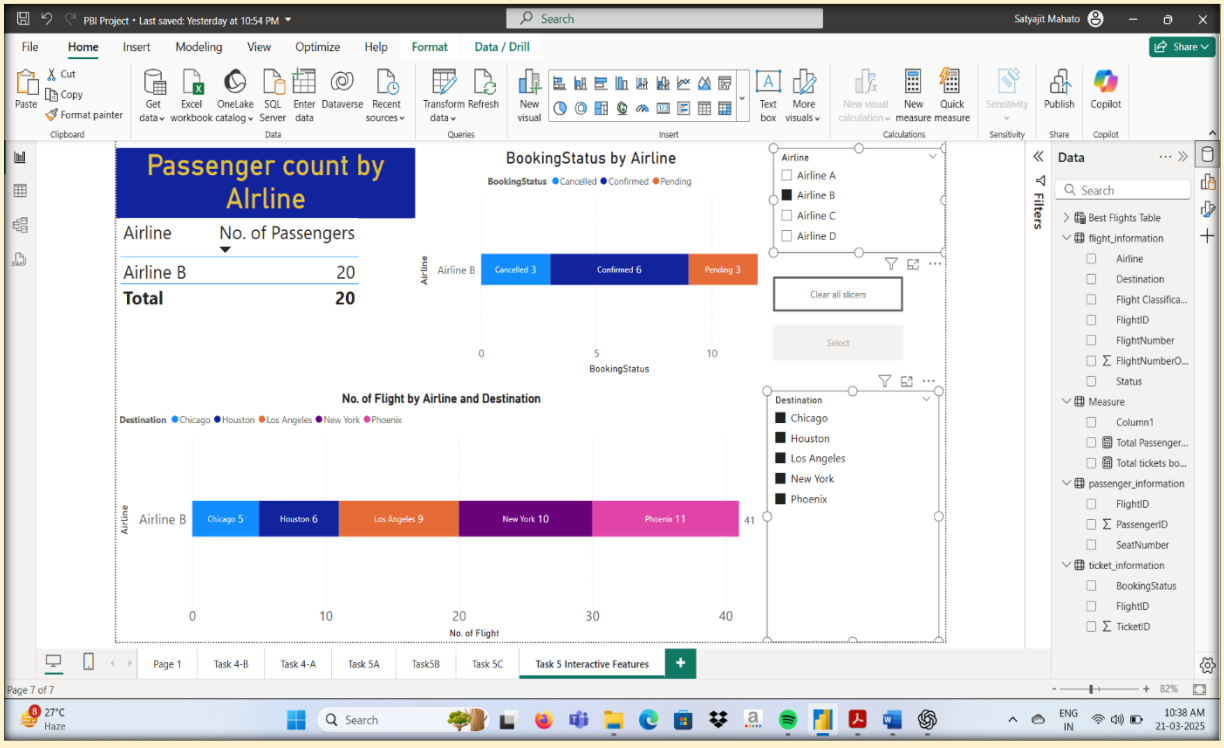


○ Flights by airline and destination.

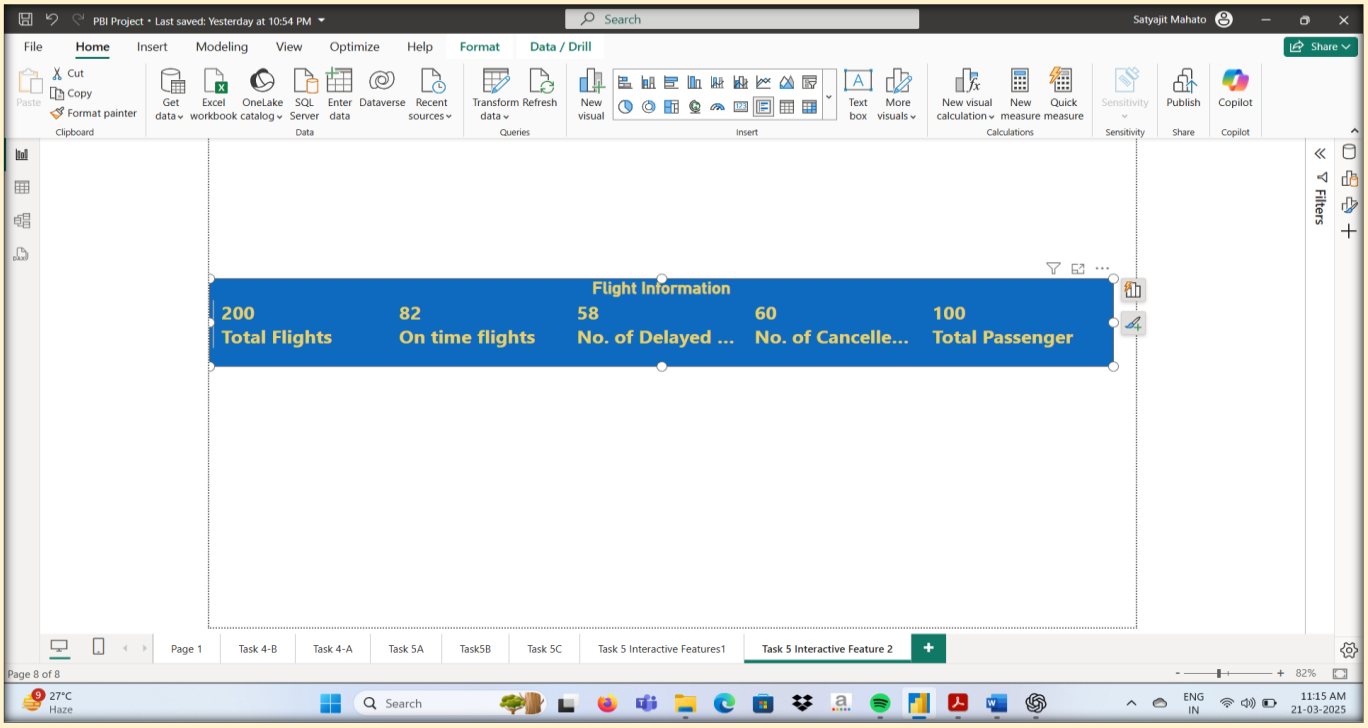


Interactive Feature

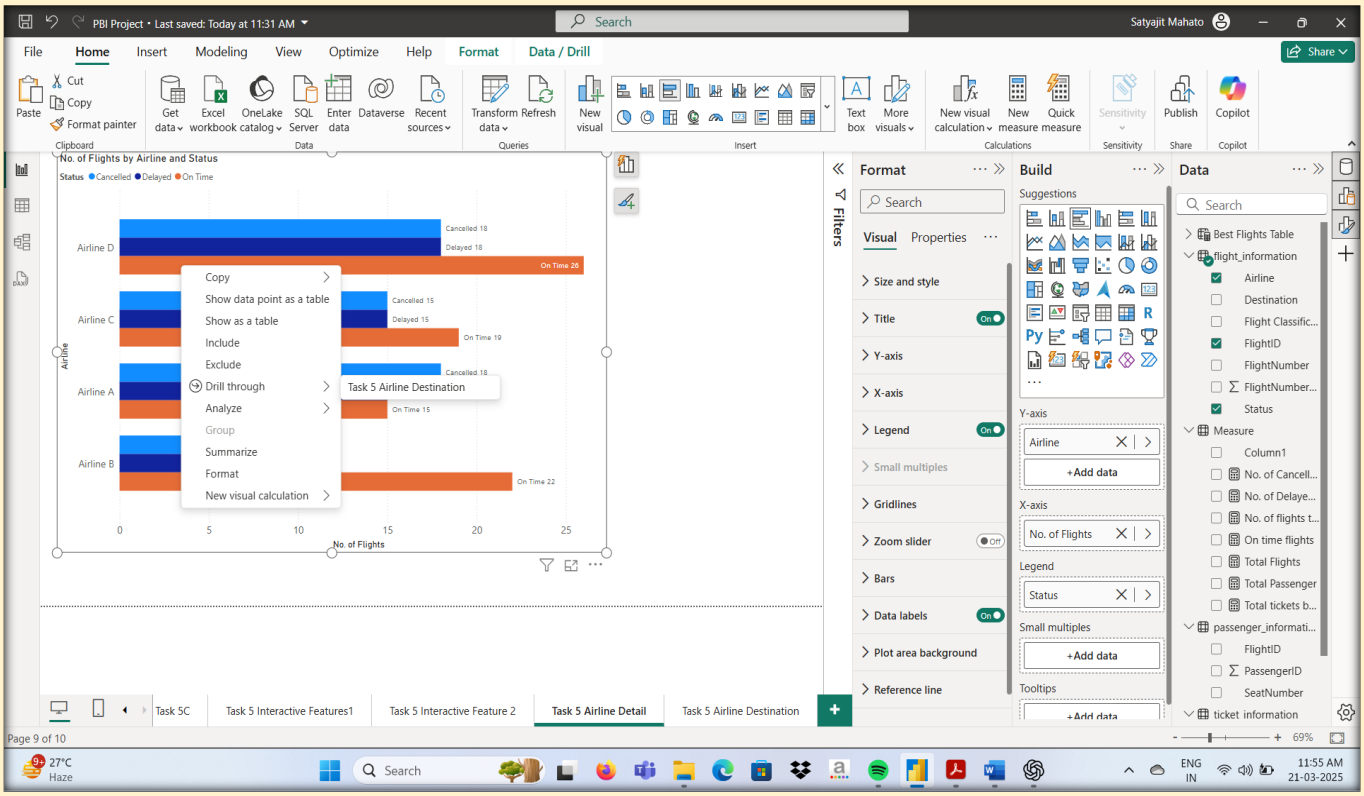
Destination and Airline



Quick Views

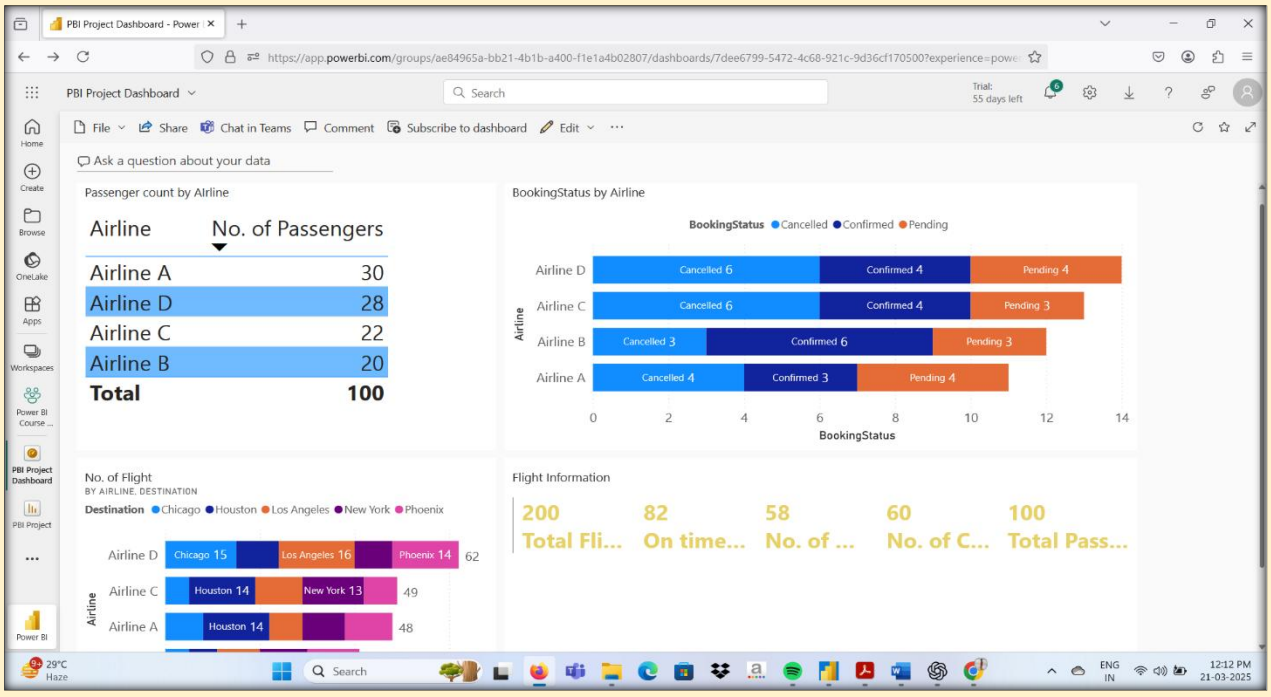


Airline Specific View using Drillthrough

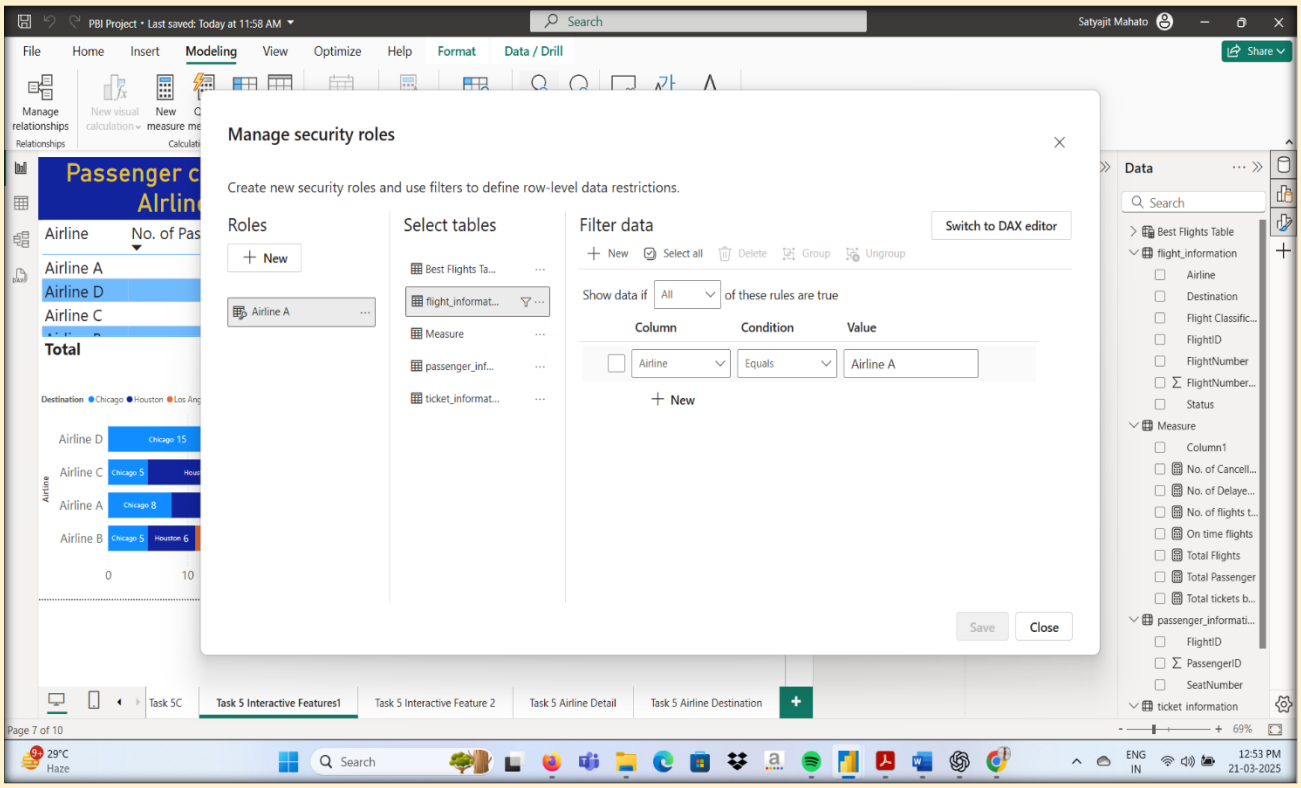


Task 6. Final Dashboard and Power BI Service

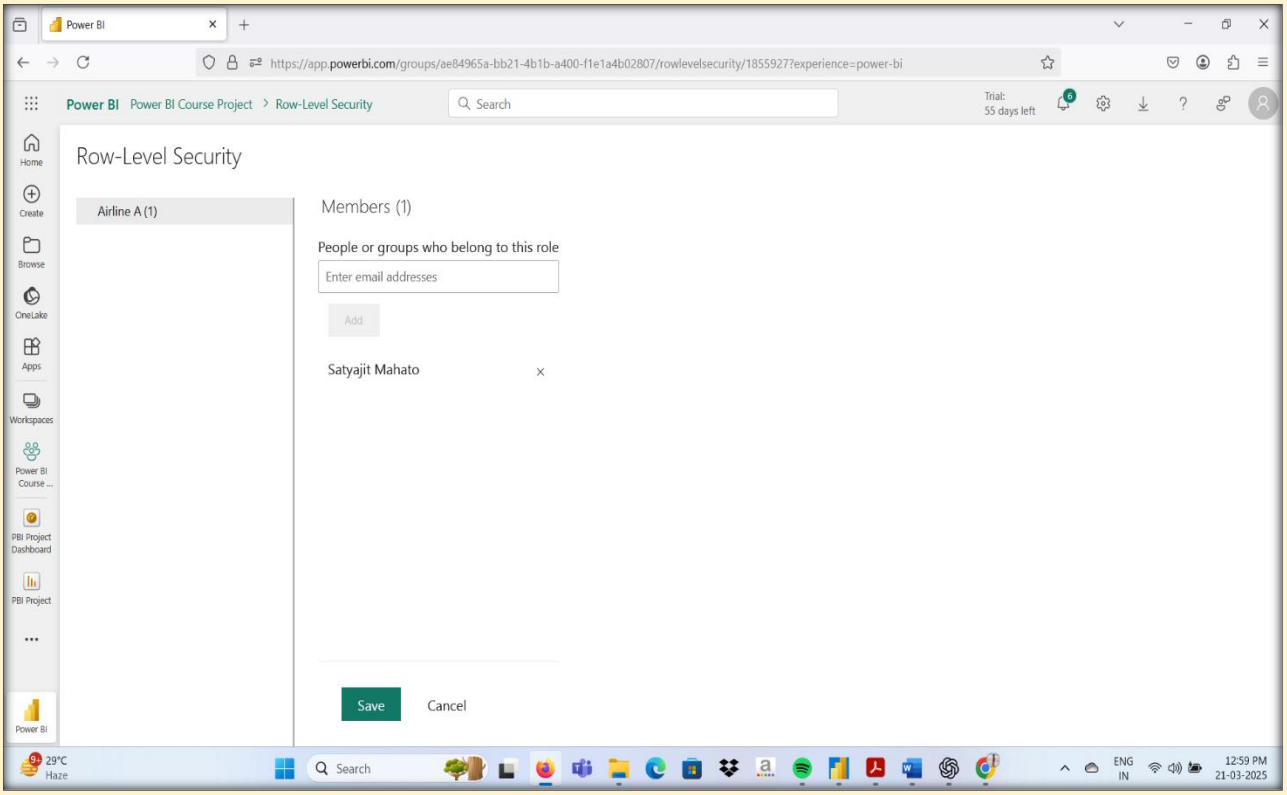
● **Dashboard** with key visuals and insights.



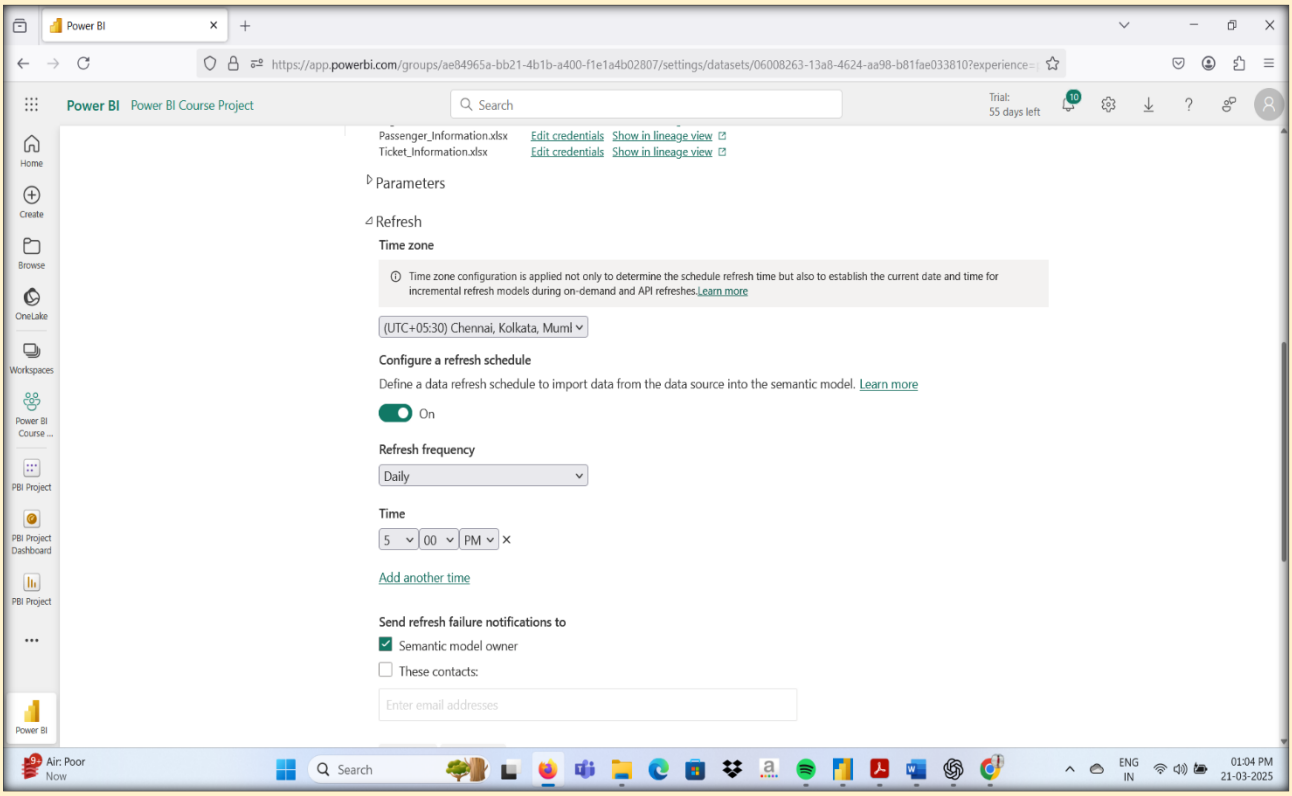
● Configuration of **Row-Level Security (RLS)** for Airline A data



● Assigned to a user



● Set up a **schedule refresh** at 5 PM daily.



Key Findings:

1. Airline D operates highest no. of flights i.e. 62
2. 40 % flights to Los Angeles and 39 % flights to Chicago were cancelled
3. 48 % flights to Houston were delayed.
3. Airline A and D has highest number of cancellations
4. Total number of passengers were 100
5. Total tickets booked across all flights are 50
6. Total 19 tickets were cancelled
7. Total flights operated were 200
8. 82 flights were on time across all flights
9. 58 flights were delayed across all flights
10. 60 flights were cancelled across all flights
11. 53% flights of airline B are on time.
12. 37.5% flights of Airline A are cancelled
13. More than 25% flights are delayed on an average across all flights

Recommendations:

1. Several flights were delayed and cancelled adding to customer dissatisfaction. Investigate the reasons and Reduce flight delays and cancellation overall to improve customer satisfaction.
2. Investigate the reason behind ticket cancellation
3. Find out why Airline A and D has highest cancellations.
4. Find out the reasons behind delayed flights to Houston
5. Enhance customer satisfaction by improving on-time performance
6. Optimize flight scheduling based popular routes and occupancy rates

In conclusion, this project highlights how Power BI can revolutionize airline data management.
By leveraging data analytics, airlines can:

- ✓ Reduce delays and cancellations
- ✓ Improve ticketing efficiency
- ✓ Enhance passenger satisfaction

Video Explanation Link

https://drive.google.com/file/d/1D6LzT3i9e3lfmMdPR22TGWnAbXx35oqS/view?usp=drive_link