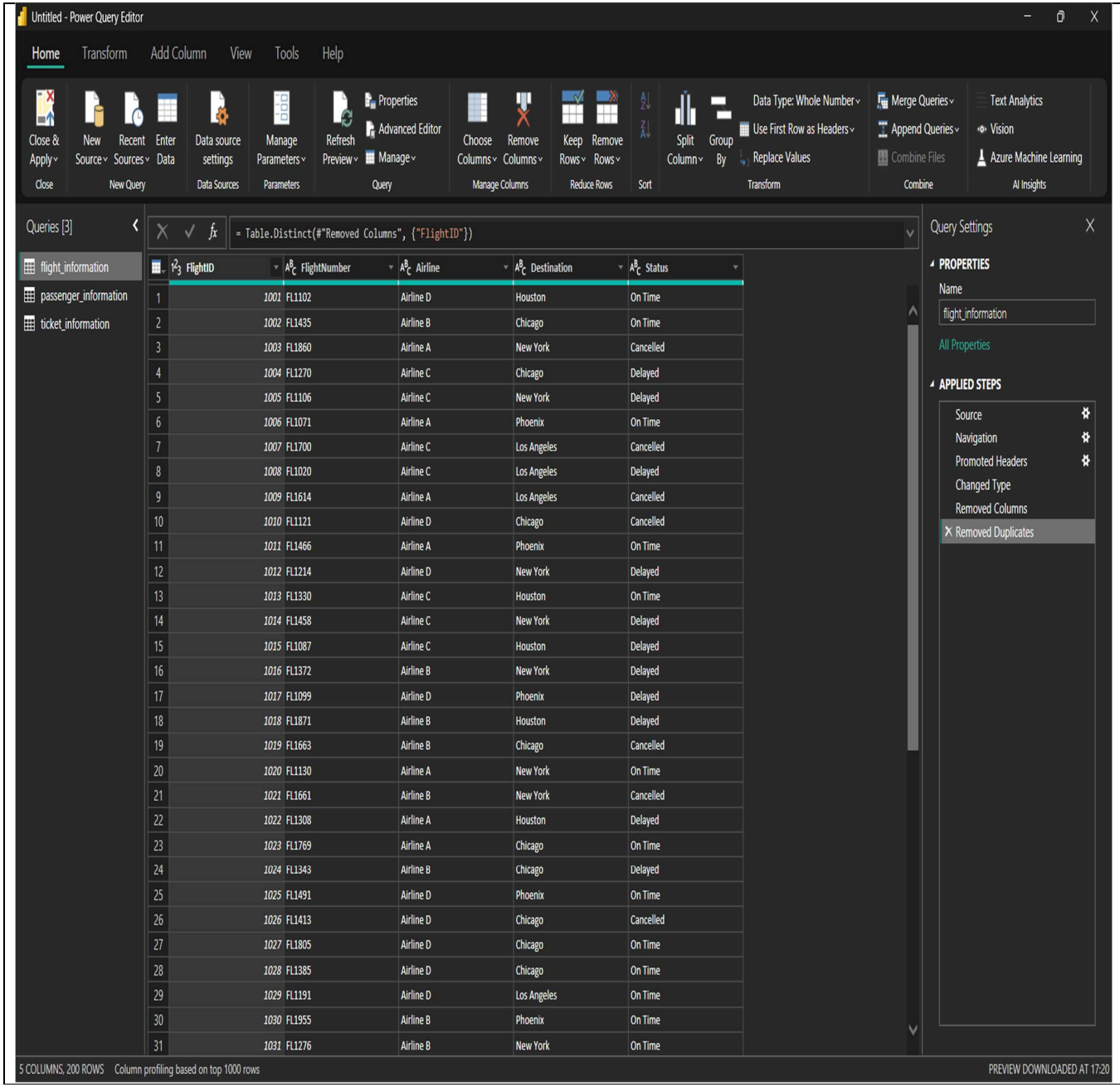


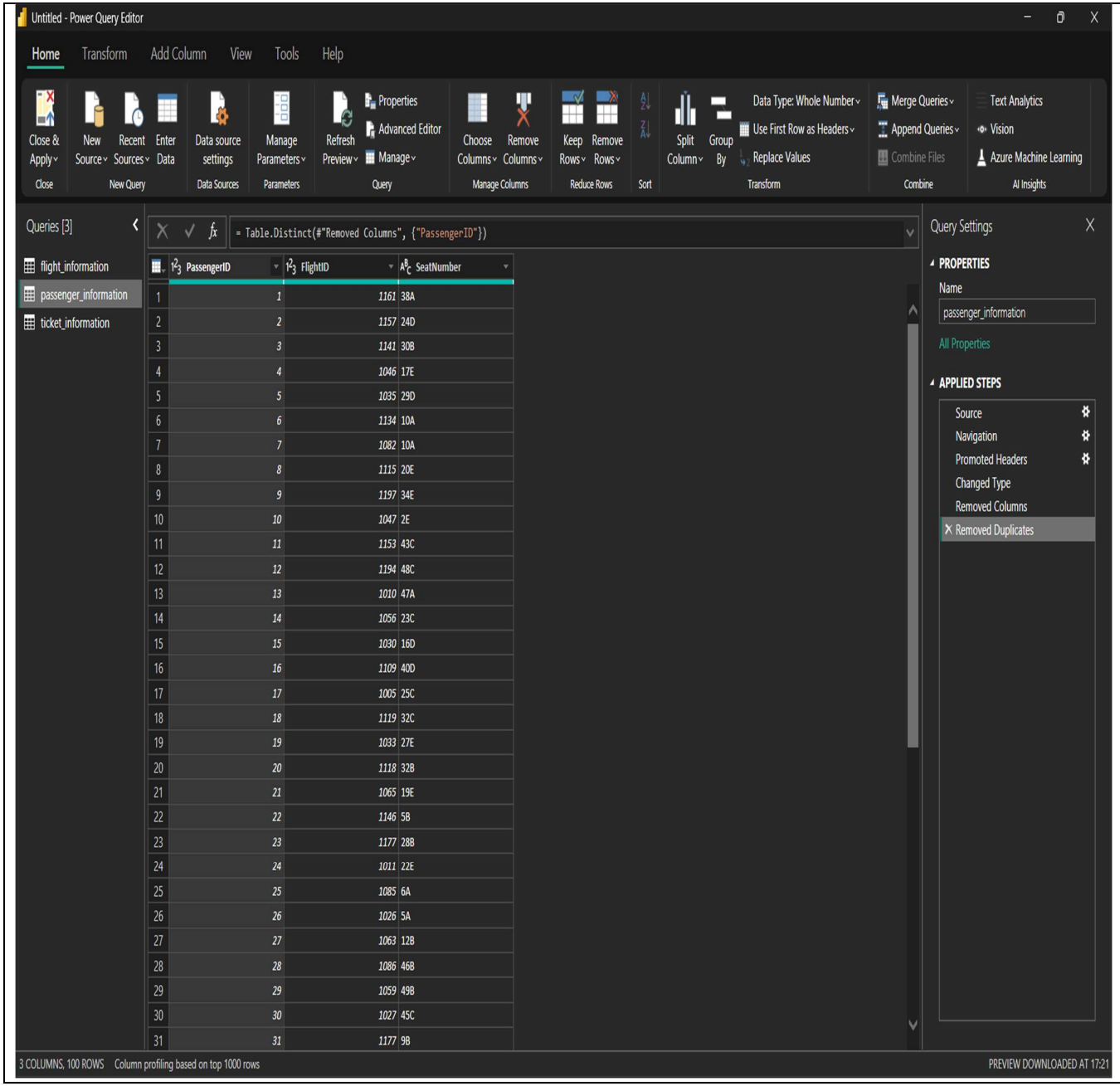
Task 1: Data Preparation and Cleaning

Extract and transform data in Power Query. Clean data: remove duplicates, handle missing values, and format columns.

Steps Taken:

<u>Flight information dataset:</u>

- *From the home tab in power query opted the remove rows option with duplicated as the dataset did not have any duplicates so none were removed.*
- *From the transform tab selected the option use first row as header to promote the headers by 1 row.*
- *From the drop down in column headers assigned the data types to each column.*
- *The data set had a lot of extra columns so from the home tab in power query choose the remove column option to remove those columns.*



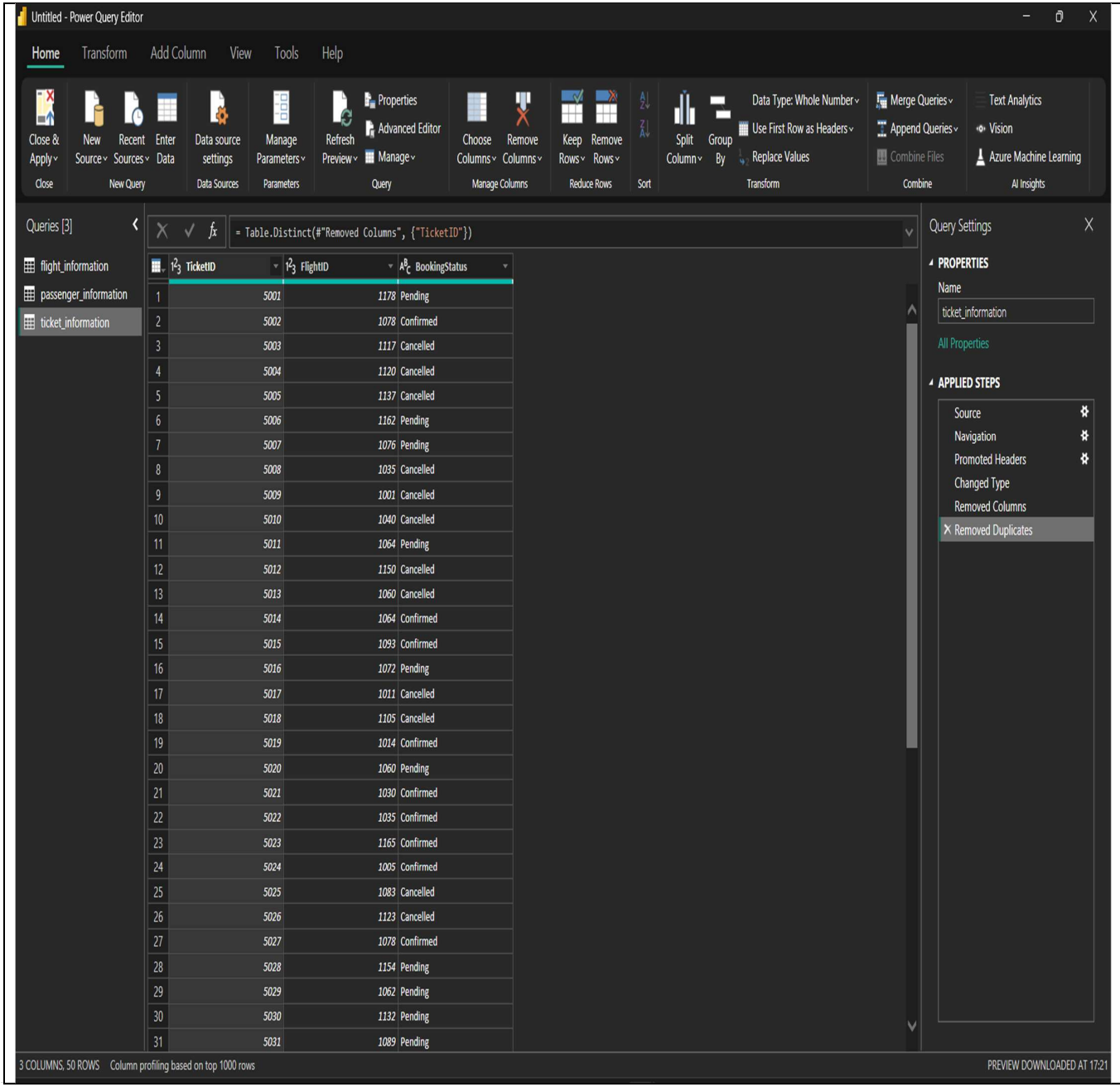
Task 1: Data Preparation and Cleaning

Extract and transform data in Power Query. Clean data: remove duplicates, handle missing values, and format columns.

Steps Taken:

<u>Passenger information dataset:</u>

- *From the home tab in power query opted the remove rows option with duplicated as the dataset did not have any duplicates so none were removed.*
- *From the drop down in column headers assigned the data types to each column.*
- *The data set had a lot of extra columns so from the home tab in power query choose the remove column option to remove those columns.*



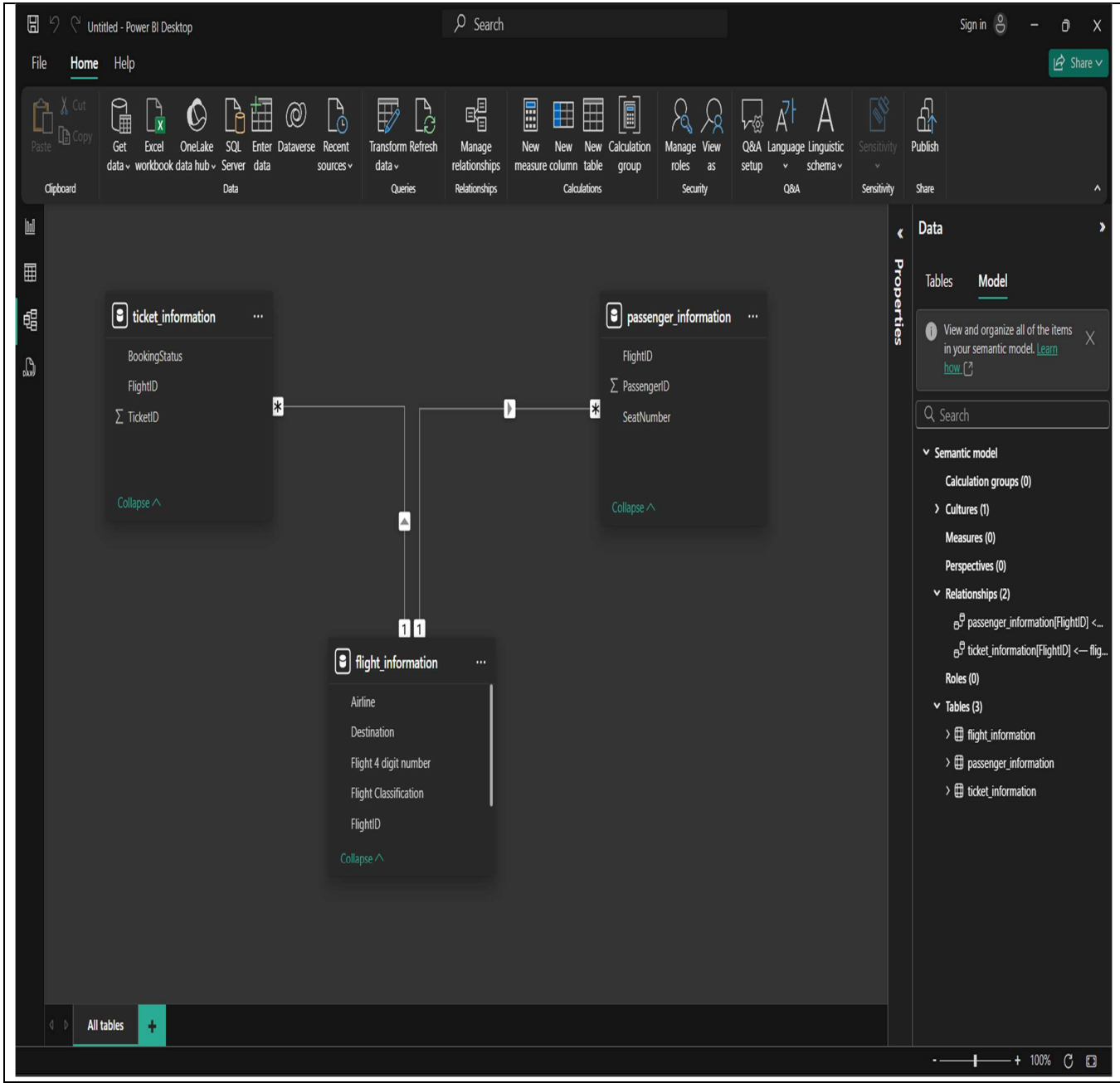
Task 1: Data Preparation and Cleaning

Extract and transform data in Power Query. Clean data: remove duplicates, handle missing values, and format columns.

Steps Taken:

<u>Ticket_information dataset:</u>

- *From the home tab in power query opted the remove rows option with duplicated as the dataset did not have any duplicates so none were removed.*
- *From the drop down in column headers assigned the data types to each column.*
- *The data set had a lot of extra columns so from the home tab in power query choose the remove column option to remove those columns.*

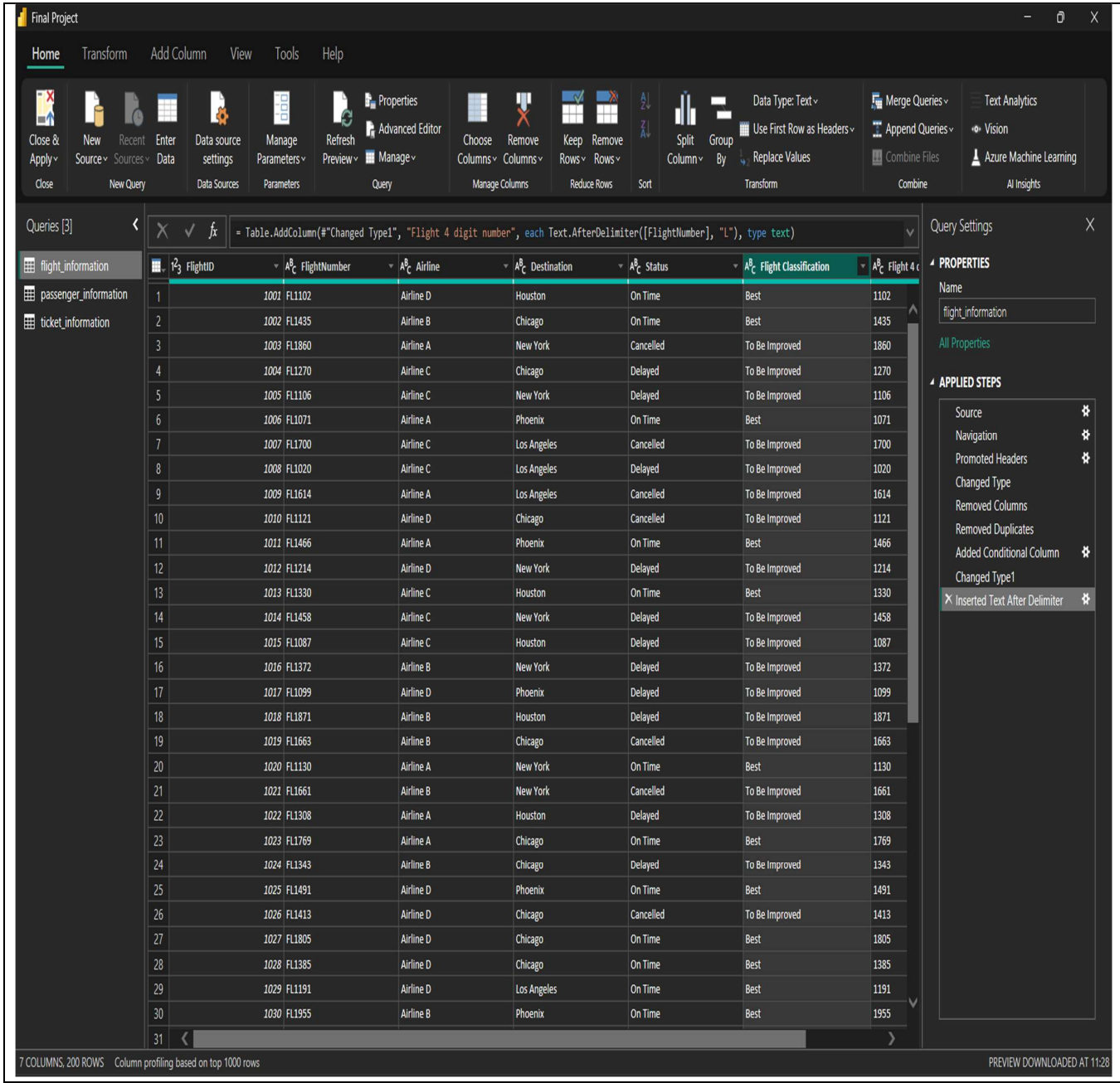


Task 2: Data Modelling

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

Steps Taken:

- *On the power BI desktop switched to model view where all the tables are listed.*
- *From the **ticket_information** table selected the **flightID** column and dragged it over similar column of **flight_information** table to create a many-one relationship.*
- *From the **passenger_information** table selected the **flightID** column and dragged it over similar column of **flight_information** table to create a many-one relationship.*
- *Relationship **cannot** be formed in ticket_information to passenger_information as it would be many to many relationship.*
- *In the image we can see the **star-schema model** with flight_infomation table in the middle.*



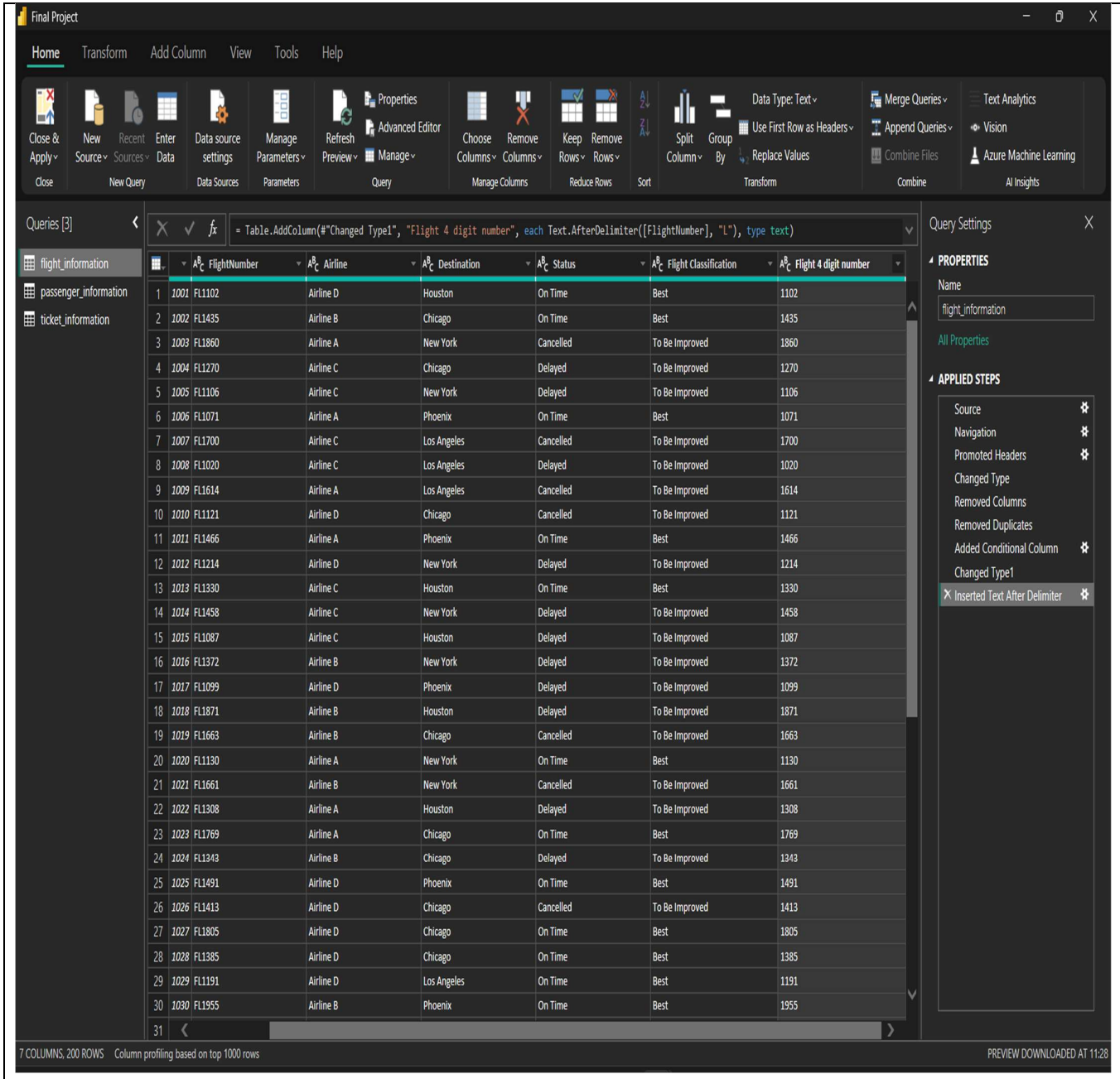
	FlightID	FlightNumber	Airline	Destination	Status	Flight Classification	Flight 4 c
1	1001	FL1102	Airline D	Houston	On Time	Best	1102
2	1002	FL1435	Airline B	Chicago	On Time	Best	1435
3	1003	FL1860	Airline A	New York	Cancelled	To Be Improved	1860
4	1004	FL1270	Airline C	Chicago	Delayed	To Be Improved	1270
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
28	1028	FL1385	Airline D	Chicago	On Time	Best	1385
29	1029	FL1191	Airline D	Los Angeles	On Time	Best	1191
30	1030	FL1955	Airline B	Phoenix	On Time	Best	1955

Task 3: Enhanced Data Insights

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

Steps Taken:

- *On Power Query- Add column tab selected the **conditional column option**.*
- *In the dialog box set the conditions for "**Best**" when "**On Time**".*
- *And "**To Be Improved**" when "**Cancelled**" and "**Delayed**".*
- *And then renamed the column as "**Flight Classification**".*
- *As shown in the image a new column has been created with the set conditions as classifications.*



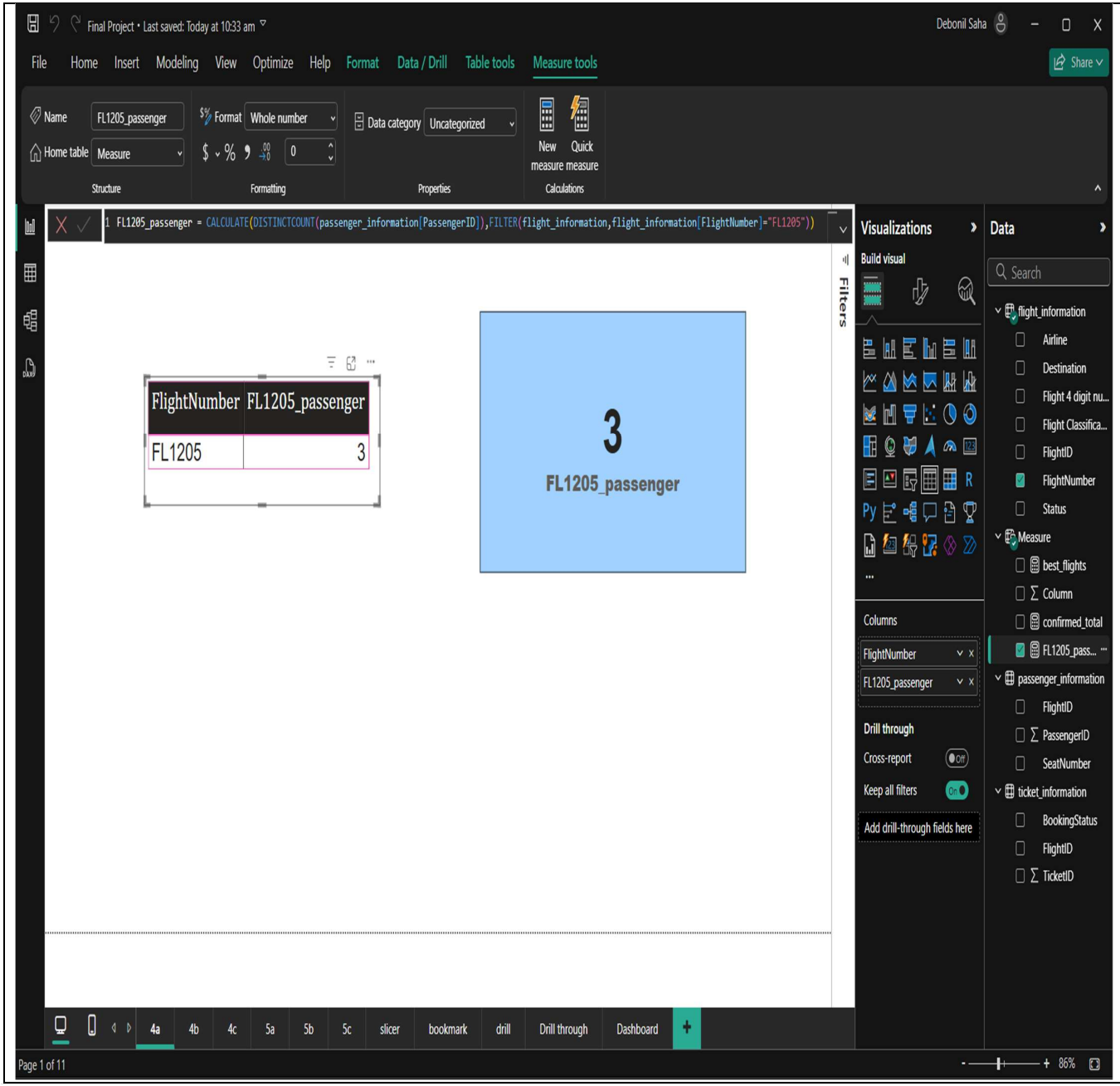
	FlightNumber	Airline	Destination	Status	Flight Classification	Flight 4 digit number
1	1001 FL1102	Airline D	Houston	On Time	Best	1102
2	1002 FL1435	Airline B	Chicago	On Time	Best	1435
3	1003 FL1860	Airline A	New York	Cancelled	To Be Improved	1860
4	1004 FL1270	Airline C	Chicago	Delayed	To Be Improved	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
28	1028 FL1385	Airline D	Chicago	On Time	Best	1385
29	1029 FL1191	Airline D	Los Angeles	On Time	Best	1191
30	1030 FL1955	Airline B	Phoenix	On Time	Best	1955
31						

Task 3: Enhanced Data Insights

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

Steps Taken:

- *On Power Query- Add column tab selected the **Column from example** option.*
- *With the new column created entered **2 rows** with the **last 4 digits** from the **FlightNumber** column.*
- *This leads to the entire column being created on its own.*
- *Then renamed the column to "**Flight 4 digit number**".*

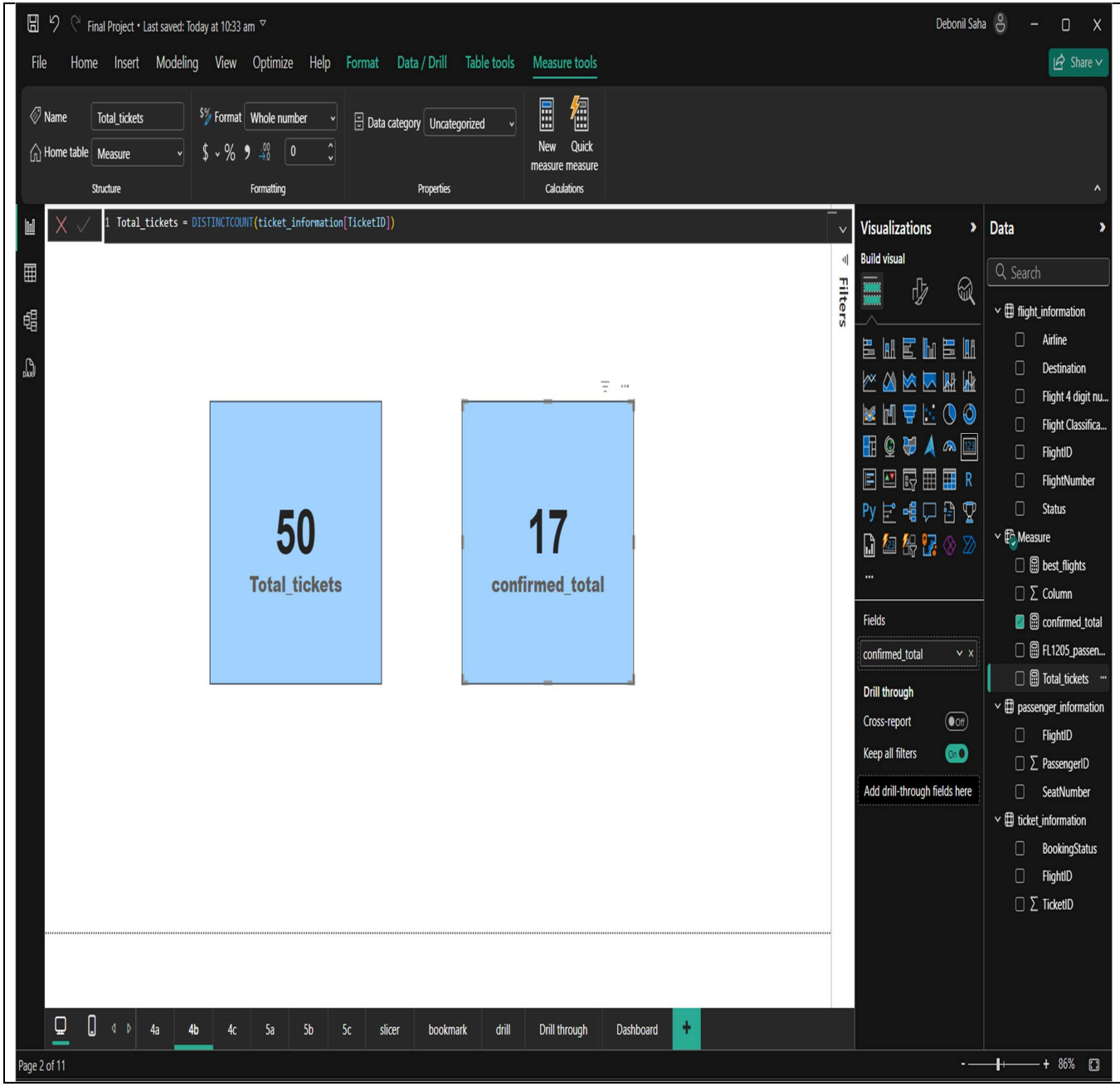


Task 4: Calculations Using DAX

Calculate: Total passengers for a specific flight.

Steps Taken:

- *Created a new measure with the DAX formula:* ***FL1205_passenger = CALCULATE(DISTINCTCOUNT(passenger_information[PassengerID]),FILTER(flight_information,flight_information[FlightNumber]="FL1205"))***
- *This measure counts the number of passengers in the flight* ***FL1205****.*
- *From the visualization pane created a new visual using single card and applied the created measure as the field.*
- *From the visualization pane created another table visual and added the column FlightNumber and the measure FL1205_passenger.*
- *In both the visuals we can see the number of passengers in the flight FL1205.*

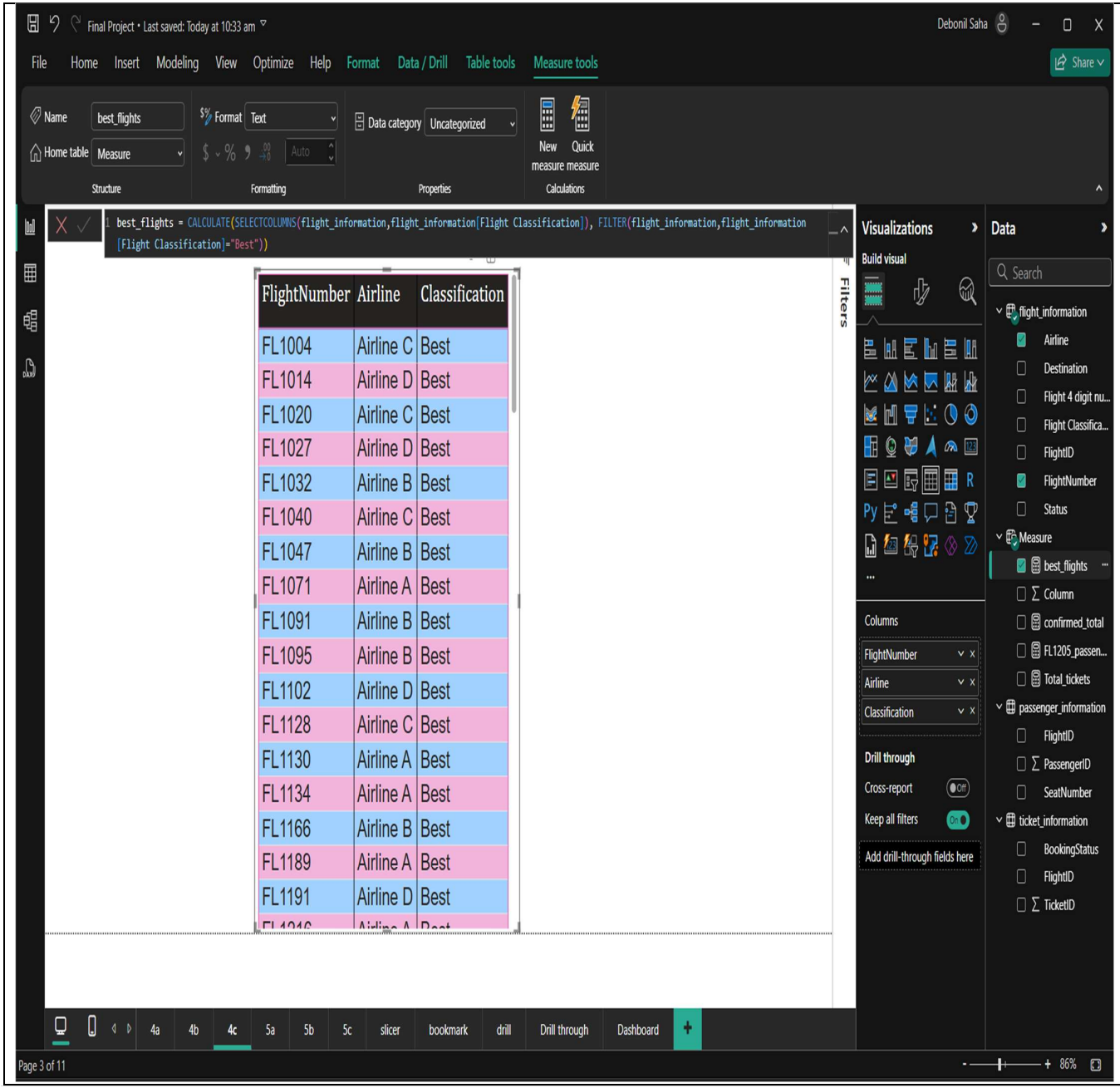


Task 4: Calculations Using DAX

Calculate: Total tickets booked.

Steps Taken:

- *In order to calculate the total number of tickets we need to create a new measure with the DAX formula:* ***Total_tickets = DISTINCTCOUNT(ticket_information[TicketID])***
- *From the visualization pane created a single card visual with the created measure as the field to view the result- "50" as shown in the image.*
- *We can also get the number of tickets booked out of this with a DAX measure formula:* ***confirmed_total = CALCULATE(DISTINCTCOUNT(ticket_information[TicketID]), FILTER(ticket_information,ticket_information[BookingStatus]="Confirmed"))***
- *Plotting this measure into single card we can see the result as "17".*

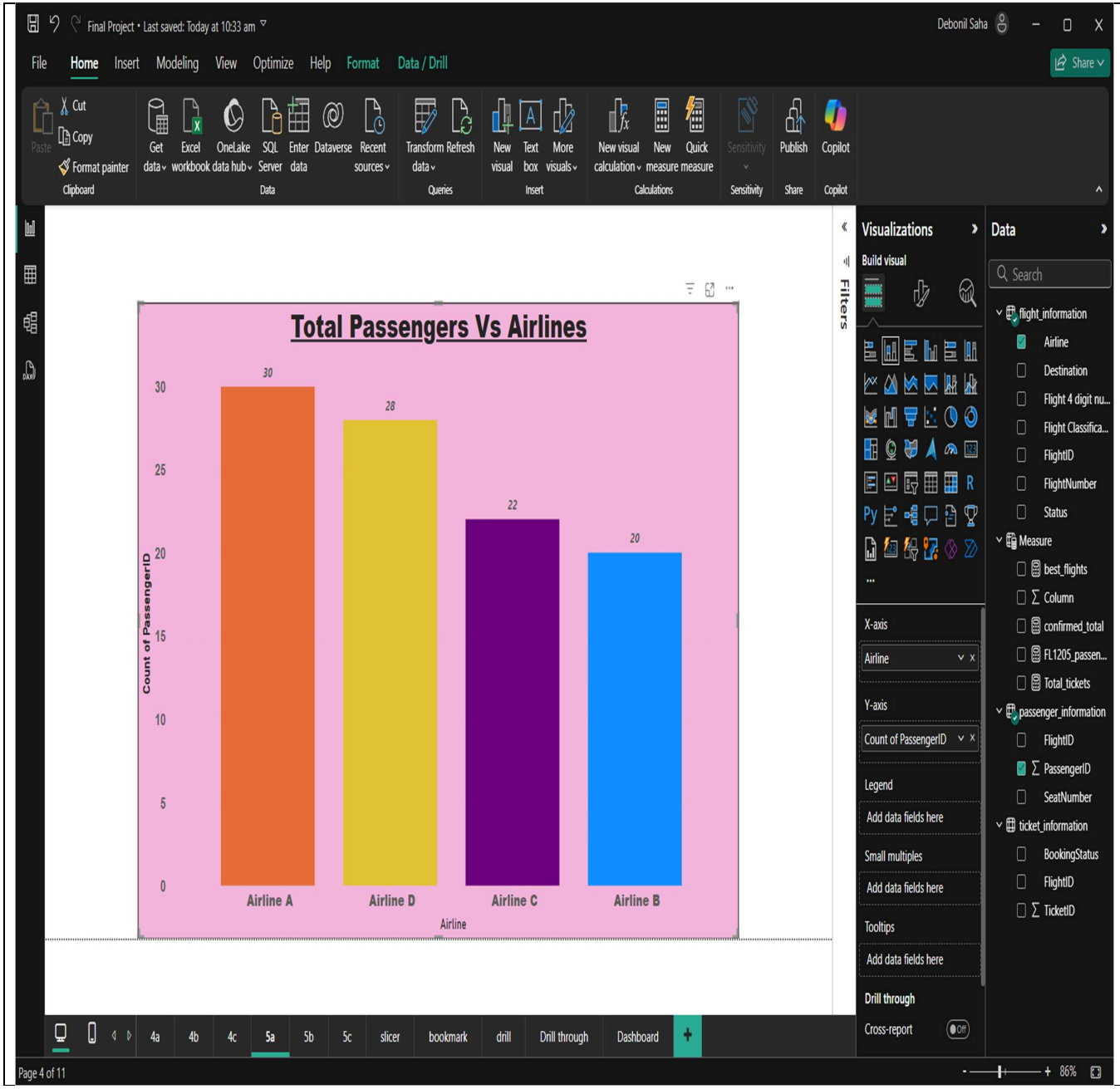


Task 4: Calculations Using DAX

Calculate: Filtered table showing "Best" flights only.

Steps Taken:

- *To get list of flights with only "Best" classification we can use DAX formula:* **best_flights = CALCULATE(SELECTCOLUMNS(flight_information,flight_information[Flight Classification]), FILTER(flight_information,flight_information[Flight Classification]="Best"))**
- *Creating a table visual and adding the fields FlightNumber, Airline and the created measure we get the list of all the flights classified as "Best".*

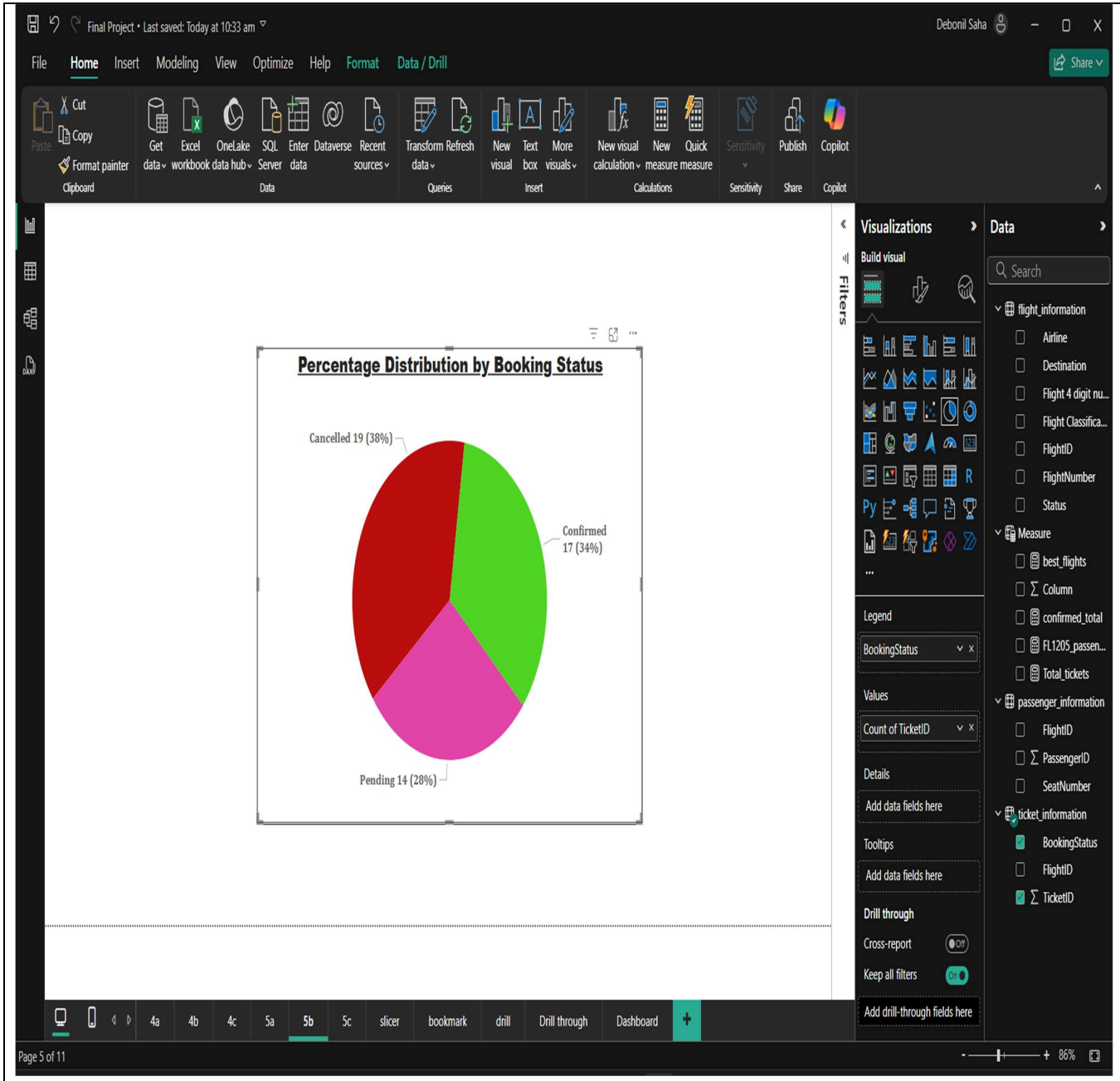


Task 5: Visualization and Interactive Features

Create visuals for: Passenger count by airline.

Steps Taken:

- *From the visualization pane inserted a new visual for **column chart**.*
- *Added the fields **Airline in X-axis** and **count of PassengerID as Y-axis**.*
- *Formatting the chart with **title, effects** and adding **data labels** we get the desired result.*

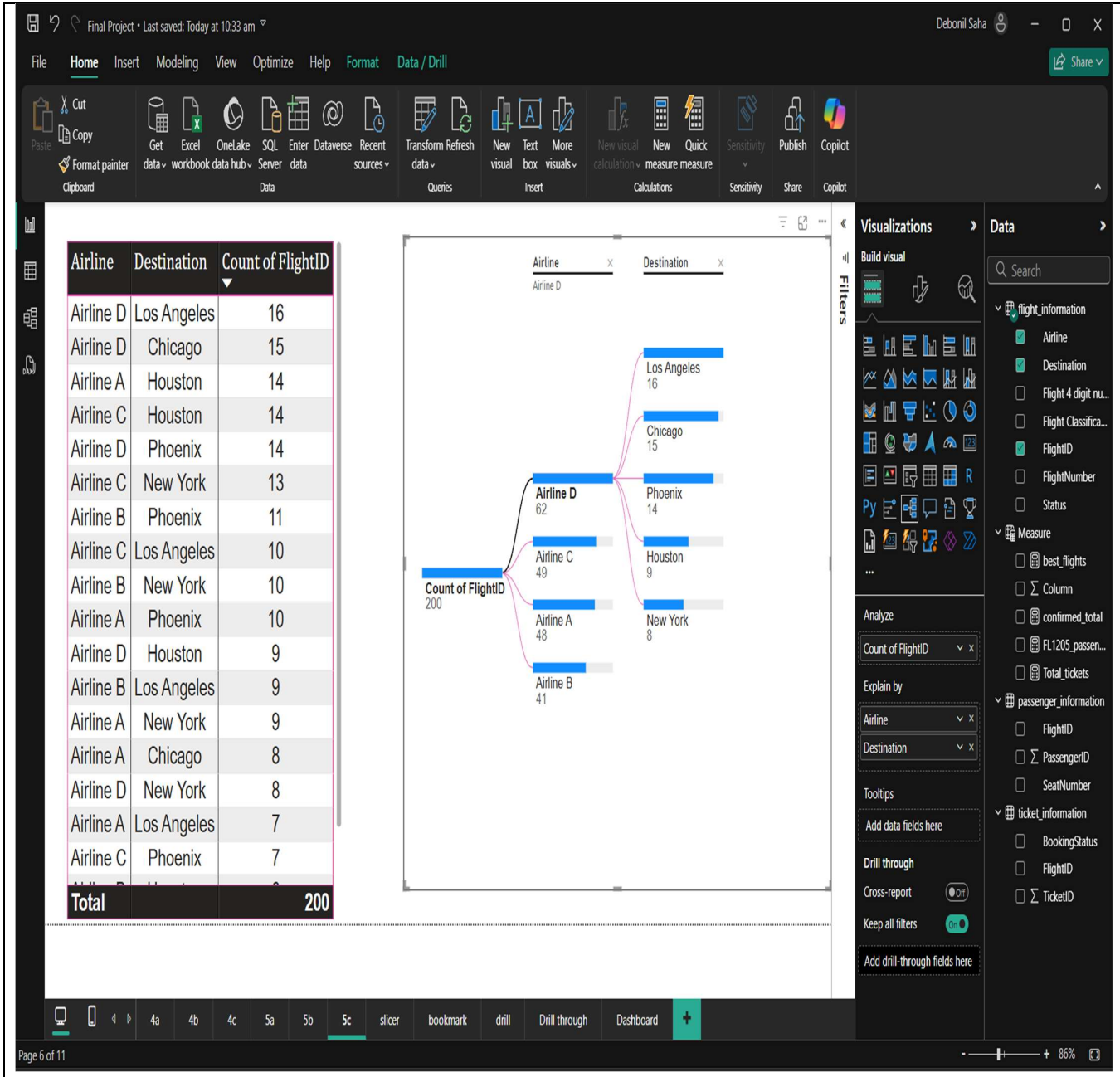


Task 5: Visualization and Interactive Features

Create visuals for: Ticket booking statuses.

Steps Taken:

- *From the visualization pane inserted the **pie chart**.*
- *Added the column **BookingStatus as legend and Count of TicketID as Values**.*
- *Formatting the chart with removing with different slice colors, adding a title we get the given pie chart.*

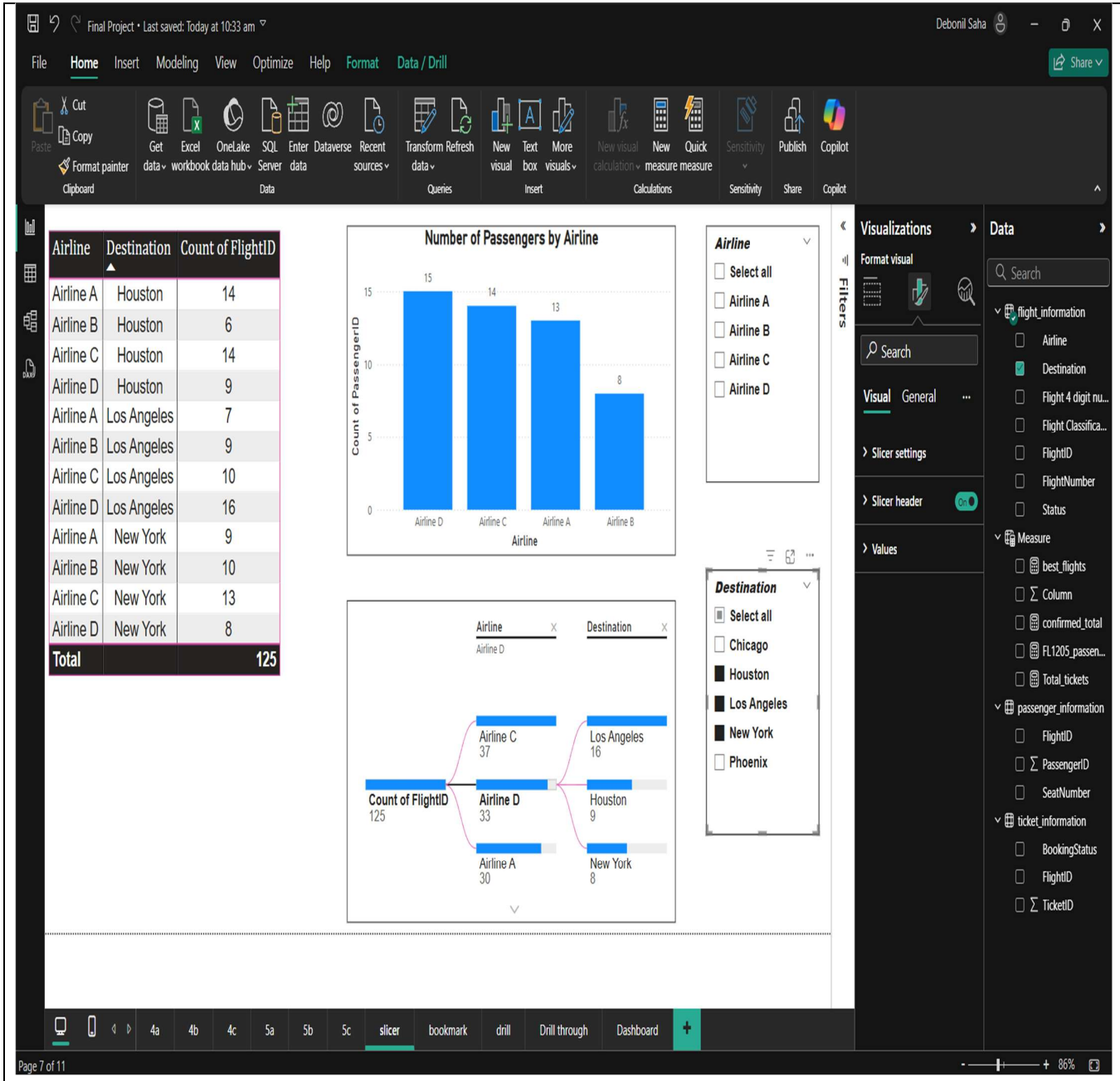


Task 5: Visualization and Interactive Features

Create visuals for: Flights by airline and destination.

Steps Taken:

- *From the visualization pane inserted the **table visual** added the columns Airline, Destination and count of FlightID, using this visual we can see the number of flights from a particular airline and destination.*
- *Added another visual **decomposition tree** by analzing the count of FlightID by Airline and Destination. Using this visual we can see the number of flights first integrating it on airline basis and then by destination.*

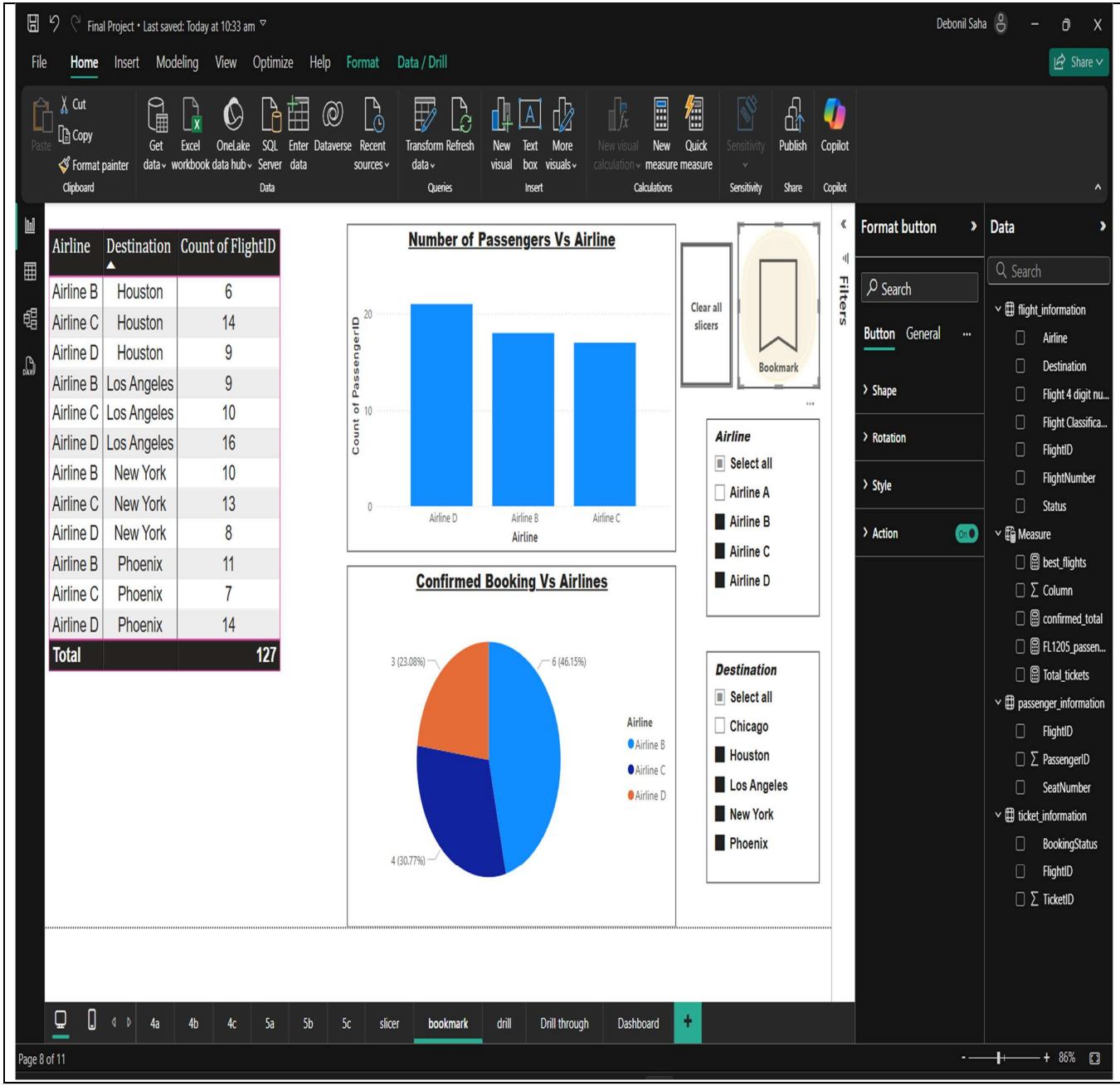


Task 5: Visualization and Interactive Features

Add interactive features for: Destination and Airline.

Steps Taken:

- *From the visualization pane added the **slicer visual** with the **field as Airlines**.*
- *Added another **slicer visual** with **field as Destinations.***
- *Using these 2 slicers we can filter the data into the shown visual as per our requirements.*

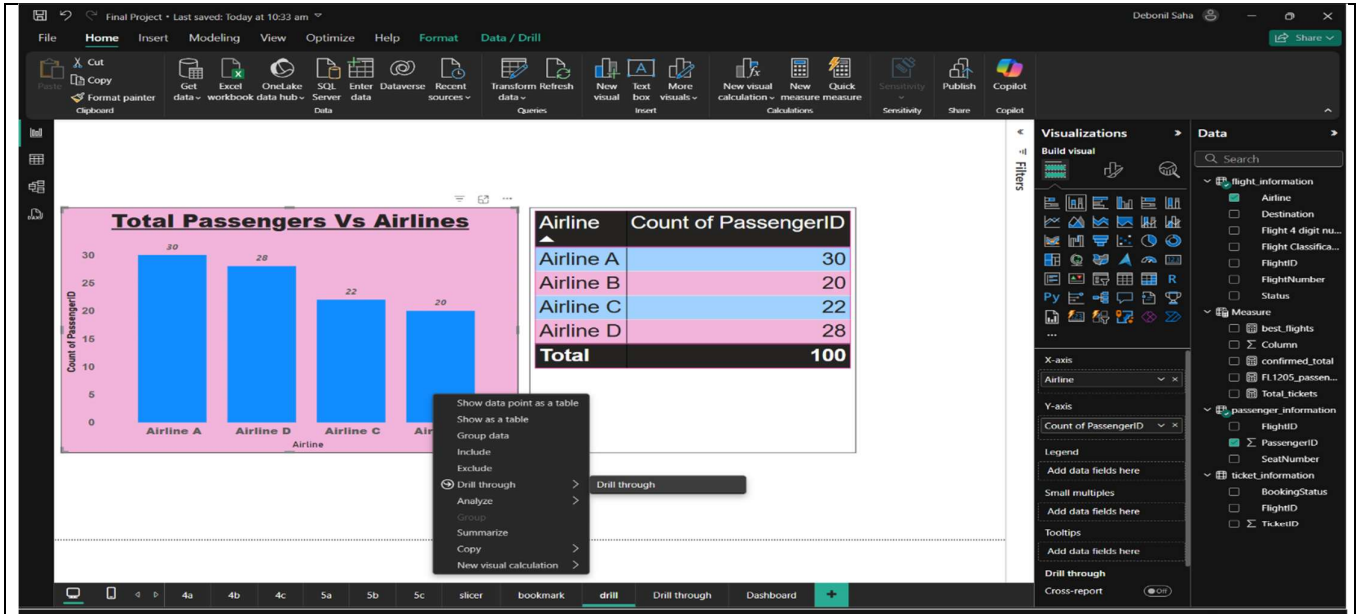


Task 5: Visualization and Interactive Features

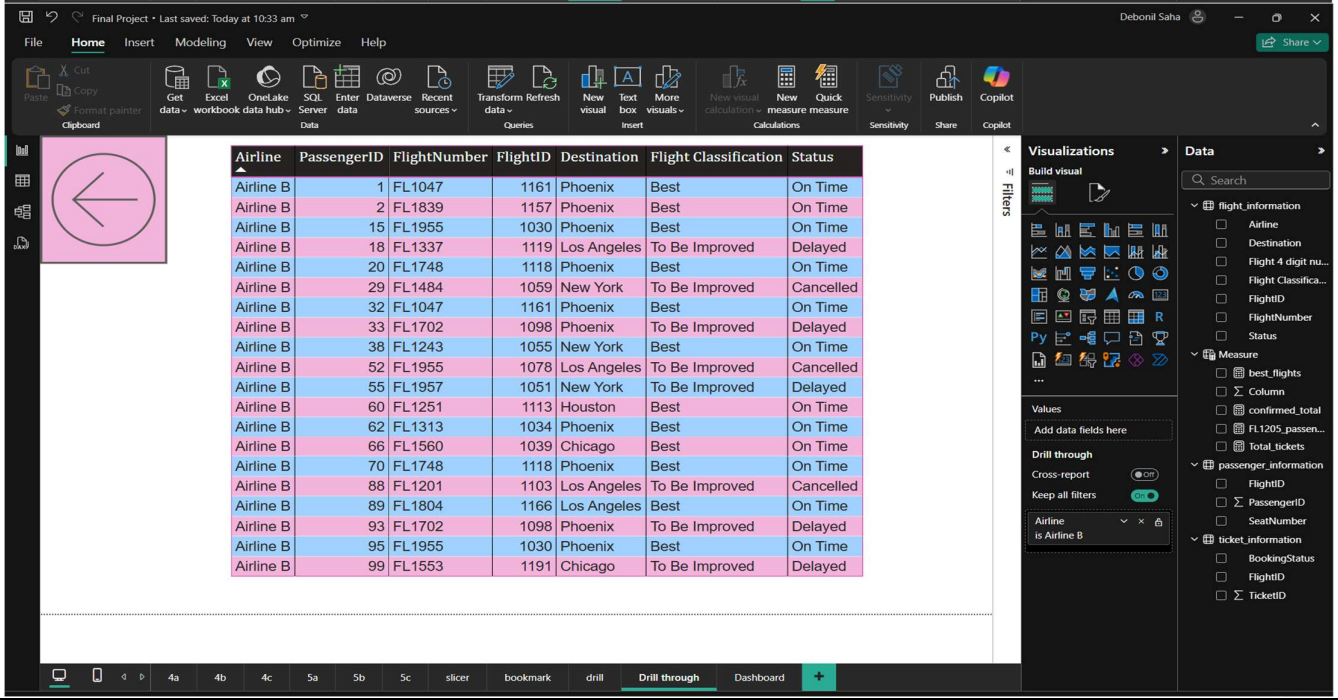
Add interactive features for: Quick views.

Steps Taken:

- *From the **insert tab-button option** selected the **bookmark button** and added it to the current report.*
- *From the view tab added the **bookmarks pane**.*
- *Then selected my required filter condition using slicers.*
- *Then in the bookmarks pane added this scenario as a bookmark.*
- *Then by formatting the bookmarks button added the **action** as the bookmark created.*
- *So now every-time we click the bookmarks button we will be taken the same scenario that was added as bookmark.*



Airline	Count of PassengerID
Airline A	30
Airline B	20
Airline C	22
Airline D	28
Total	**100**



Airline	PassengerID	FlightNumber	FlightID	Destination	Flight Classification	Status
Airline B	1	FL1047	1161	Phoenix	Best	On Time
Airline B	2	FL1839	1157	Phoenix	Best	On Time
Airline B	15	FL1955	1030	Phoenix	Best	On Time
Airline B	18	FL1337	1119	Los Angeles	To Be Improved	Delayed
Airline B	20	FL1748	1118	Phoenix	Best	On Time
Airline B	29	FL1484	1059	New York	To Be Improved	Cancelled
Airline B	32	FL1047	1161	Phoenix	Best	On Time
Airline B	33	FL1702	1098	Phoenix	To Be Improved	Delayed
Airline B	38	FL1243	1055	New York	Best	On Time
Airline B	52	FL1955	1078	Los Angeles	To Be Improved	Cancelled
Airline B	55	FL1957	1051	New York	To Be Improved	Delayed
Airline B	60	FL1251	1113	Houston	Best	On Time
Airline B	62	FL1313	1034	Phoenix	Best	On Time
Airline B	66	FL1560	1039	Chicago	Best	On Time
Airline B	70	FL1748	1118	Phoenix	Best	On Time
Airline B	88	FL1201	1103	Los Angeles	To Be Improved	Cancelled
Airline B	89	FL1804	1166	Los Angeles	Best	On Time
Airline B	93	FL1702	1098	Phoenix	To Be Improved	Delayed
Airline B	95	FL1955	1030	Phoenix	Best	On Time
Airline B	99	FL1553	1191	Chicago	To Be Improved	Delayed

Task 5: Visualization and Interactive Features

Add interactive features for: Airline-specific pages.

Steps Taken:

- *Drill through feature is applied by assigning a drill page which contains the values you want to see while doing a deeper analysis.*
- *On the drill page in the visualization pane added **the Airline as field in drill through option.***
- *Now when we select any other visual from any other page in the report and choose airline option as drill though we are directed to the drill page.*
- *As shown in the image we drill the airlines in the column chart where we get directed to drill page where information about the passenger and other flight details are available.*

Task 6: Final Dashboard and Power BI Service

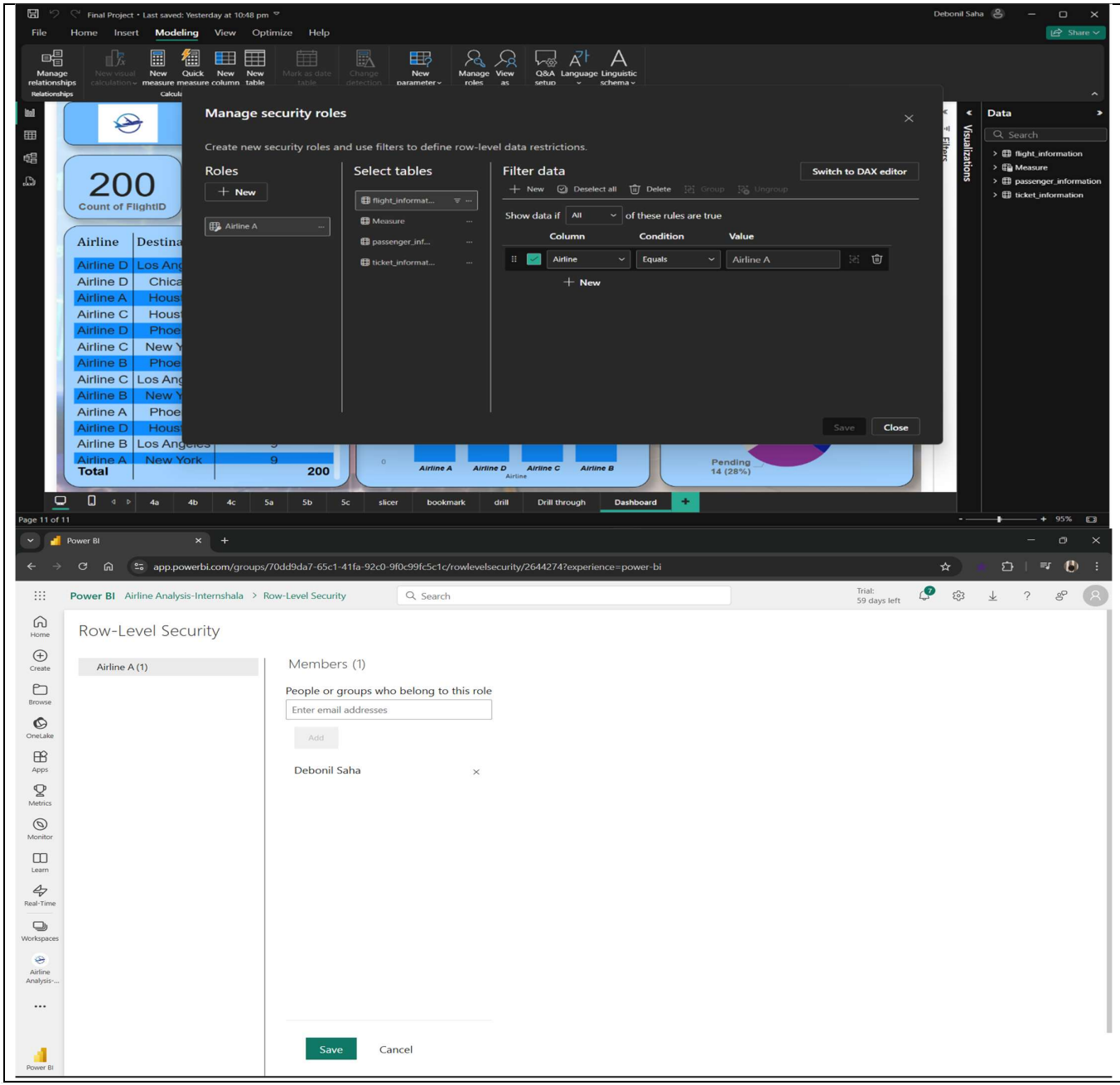
*Design a **comprehensive dashboard** with key visuals and insights.*

Key Insights:

- *Total flights: 200*
- *Total passengers: 100*
- *Number of unique destinations: 5*
- *Tickets analysed: 50 in which only 17 are confirmed*
- *Airline D operates the highest number of flights to a single destination, specifically Los Angeles (16 flights).*
- *Other destinations, such as Chicago (15 flights) and Houston (14 flights), are also popular among different airlines.*
- *Phoenix is serviced by multiple airlines but has significant coverage from Airline C and Airline A.*
- *By Airline: Airline A has the highest passenger count (30), followed closely by Airline D (28), indicating their strong market share.*
- *By Destination: Phoenix leads with 30% of passengers, followed by Houston (21%) and Chicago (20%). Los Angeles and New York have smaller passenger shares (11% and 18%, respectively).*

Contd:
Key Insights:

- *"To Be Improved" flights account for 59% (118 flights), while "Best" flights make up 41% (82 flights). This indicates a significant opportunity for quality improvement across services.*
- *Confirmed bookings constitute 34%, while 38% of flights are cancelled, and 28% remain pending. This indicates potential inefficiencies in booking or operational processes.*

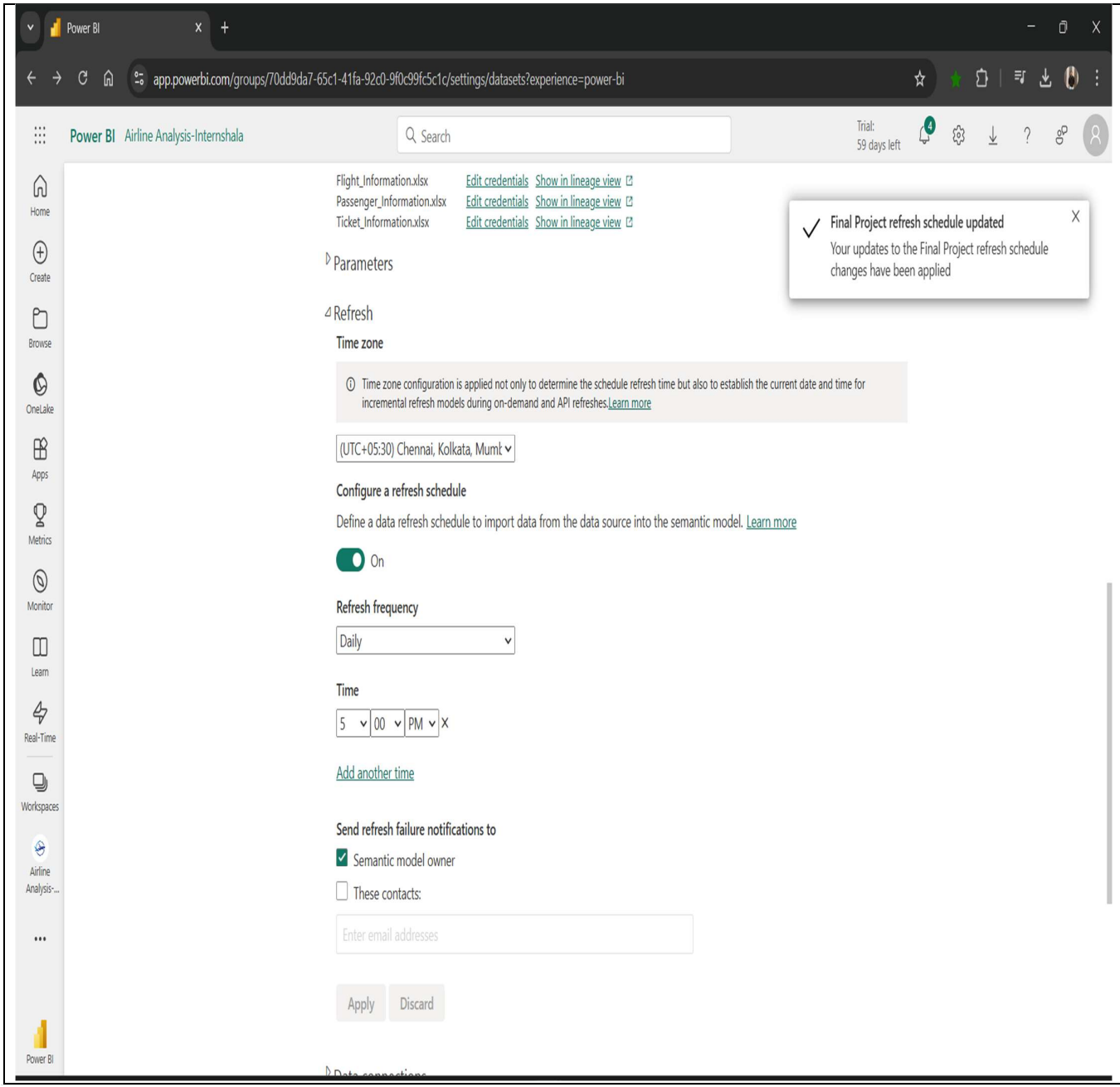


Task 6: Final Dashboard and Power BI Service

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

Steps Taken:

- *On the **modelling tab** selected the option **Manage roles**.*
- *Created a new role named "Airline A" with the **condition** from the flight_information table as **value equals "Airline A".***
- ***Publish** the report again on the cloud service website of power BI.*
- *Now on the power BI service website- my workspace of Airline Analysis, selected the **security option** from file type **semantics model**.*
- *Assigned the newly created role to a email-id or user as shown in the image.*



Task 6: Final Dashboard and Power BI Service

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

Steps Taken:

- *After **publishing** the report on a workspace named "Airline Analysis-Internshala" and installing personal **Data Gateway**.*
- *In the **Power BI settings** selected the **semantics model** which had an option for **refresh**.*
- *In the refresh option entered the time **5PM daily** and saved the settings.*
- *So as shown in the image the entire workspace will refresh the linked dataset at 5PM IST.*

Video Submission:

Link: https://drive.google.com/file/d/1gMs6bVTKTKRqi-9G1frZLHSs3WmMweeD/view?usp=sharing