Title of report: “aviation events/accidents analysis”

Contents

Figure Table

Execution Summary

This dataset includes many different factors associated with aviation events, such as injuries, accidents, damage to aircraft, weather, and more. The primary objective of the project is to derive significant insights from this dataset in order to support the continuous improvement of aviation safety protocols and practices.

Methods and Approach:

Data Collection and Preprocessing: The Aviation Department dataset is collected and

preprocessed at the start of the project. To prepare the data for subsequent analysis, this entails cleaning it up, dealing with missing values, and guaranteeing data consistency.
Exploratory Data Analysis (EDA) - The report explores exploratory analysis using statistical methods and Power BI visualizations. It looks for patterns, trends, and correlations in the dataset. Understanding the links between many data, such as injury severity, aircraft damage, weather, flying phases, and more, is made easier with the help of visual representations. **Comprehensive Analysis -** To find undiscovered insights, the study employs a variety of statistical and machine learning techniques to do a comprehensive analysis. Predictive modeling may be used to detect trends that lead to particular events or to foresee possible risk factors.

Principal Results:

Perceptive Results - Supported by data-driven analysis, this dataset offers perceptive conclusions regarding the events in the aviation sector.
Recommendations – This report makes suggestions for strengthening safety procedures, reducing risks, and raising general aviation safety requirements in light of the findings.
Visual Representations - This report makes use of Power BI visuals to provide aesthetically pleasing representations that facilitate understanding of intricate data patterns and relationships. **Practical Utility –** In this report project's goal is to give aviation industry stakeholders a useful tool to support safety procedures and policies.
Data from year 1980 to 2022 is used here like this –

Introduction

International air travel safety standards are being improved in large part through aviation safety and the investigation of aviation occurrences. This report on the **Aviation** Department dataset is a thorough examination of a large dataset that includes numerous aspects of aviation occurrences.

Dataset provides a thorough overview of all occurrences that take place in the aviation domain. These incidents include everything from accidents to different kinds of injuries, damage severity to aircraft details, and even environmental conditions and flight phases. This report's goal is to better comprehend aviation mishaps by examining, analyzing, and extrapolating insightful information from this large dataset. The goal of this project is to provide a thorough resource with insightful information that may aid in the ongoing improvement of aviation safety procedures and practices.

Data Source Description

https://www.kaggle.com/datasets/khsamaha/aviation-accident-database-synopses

This dataset provides a thorough overview of aviation occurrences, including crashes, injuries of various degrees, damage assessments of aircraft, and specific aircraft data, such as weather and various flying phases. The main goal of this report is to examine these incidents, perform analysis, and draw important conclusions from this large dataset. The ultimate objective is to produce a tool that advances knowledge about aircraft accidents and provides data that may be useful in improving aviation safety protocols and practices. Through a thorough examination and analysis of this dataset with the help of power bi visualizations, the report seeks to offer a thorough grasp of aviation events, covering everything from environmental conditions to accidents, in order to support the continuous enhancement of safety protocols in the aviation industry.

The data sheet has around 47870 rows and 31 columns. This information shows the details about the Aviation Department.

Aviation dataset explanation table

ID	COLUMN NAME	DESCRIPTION	DATA TYPE
1	Event.Id	Shows the event unique id	INT
2	Investigation.Type	Shows the investigation type	VAR/CHAR
3	Accident.Number	Shows the particular number of accident	INT
4	Event.Date	Shows the date of event	DATE/TIME
5	Location	Shows the exact location	VAR/CHAR
6	Country	Shows the name of country	VAR/CHAR
7	Latitude	Shows the latitude	INT
8	Longitude	Shows the longitude	INT
9	Airport.Code	Shows the particular airport code	VAR/CHAR
10	Airport.Name	Shows the name of airport	VAR/CHAR
11	Injury.Severity	Shows the injury type	VAR/CHAR

12	Aircraft.damage	Shows the damage of aircraft	VAR/CHAR
13	Aircraft.Category	Shows the category of aircraft	DATE/TIME
14	Registration.Number	Shows the number of registration	INT
15	Make	Shows the make of aircraft	VAR/CHAR
16	Model	Shows the model of aircraft	VAR/CHAR
17	Amateur.Built	Shows the amateur built	VAR/CHAR
18	Number.of.Engines	Number of engines in aircraft	INT
19	Engine.Type	Type of engine working	VAR/CHAR
20	FAR.Description	FAR description of aircraft	VAR/CHAR
21	Schedule	Shows the schedule	VAR/CHAR
22	Purpose.of.flight	Explain the purpose of flight	VAR/CHAR
23	Air.carrier	Shows the name of air carrier	VAR/CHAR
24	Total.Fatal.Injuries	Shows the total fatal injuries	INT
25	Total.Serious.Injuries	Shows the count of total serious injuries	INT
26	Total.Minor.Injuries	Shows the count of total minor injuries	INT
27	Total.Uninjured	Shows the total number of uninjured	INT
28	Weather.Condition	Shows the weather condition	VAR/CHAR
29	Broad.phase.of.flight	Shows the broad phase of flight	VAR/CHAR
30	Report.Status	Shows the status of report	VAR/CHAR
31	Publication.Date	Shows the date of publication	DATE/TIME

Related Business Questions

A project on an online shopping portal requires the creation of a theory or framework, which requires an awareness of a variety of business, user experience, technology, and market trends (Zhang, J., Le, T. and Chen, J., 2023; Becker, L.T. and Gould, E.M., 2019).

Some of the business questions are as:

1. In which phase of the accident did the highest number of injuries occur and what is the count of accidents in such phase?
2. What is the name of the country where least number of accidents were recorded?
3. Which company's aircraft caused the highest number of accidents and injuries?
4. Which model of aircraft caused the most accidents?
5. How many such accidents occurred in which minor aircraft damage was recorded?
6. According to flight purpose, which flights had the highest number of accidents and injuries?
7. In which year the maximum number of air accidents occurred?

Report Objectives

Here are five report objectives for the project report on aviation dataset:-

- **Extensive Examination of Aviation Events:** Examine and evaluate the dataset that includes aviation incidents, accidents, injuries, aircraft damage, and related elements in great detail. With regard to the events, their underlying causes, and the connections found in the dataset, this goal seeks to present a comprehensive picture.
- **Assessment of Safety Practices and Procedures:** By establishing correlations between injury severity, weather, flight phases, and other variables, evaluate the efficacy of current aviation safety protocols and practices. In order to lower the risk of accidents and injuries, this goal seeks to pinpoint possible areas for safety measure enhancement.
- **Creation of Visual Tools for Data Representation:** Make use of Power BI visualizations to produce interactive tools and perceptive graphical representations that help users grasp the dataset in an easy-to-understand manner. The goal is to make difficult-to-understand aviation data visually clear for stakeholders to understand.
- **Finding Recurring Patterns, Trends, and Risk variables:** Look for patterns, trends, and risk variables related to aviation mishaps. These may include relationships between weather, flight objectives, engine types, and aircraft damage. This goal is to draw attention to possible trouble spots or high-risk elements that could lead to mishaps.
- **Suggestions for Improved Aviation Safety Procedures:** Produce definitive findings and suggestions derived from the dataset analysis. The aim of this study is to strengthen industry safety standards by suggesting practical changes and enhancements to current aviation safety practices.

These goals seek to take advantage of the comprehensiveness of the Aviation dataset in order to examine past events and get practical insights for enhancing aviation safety protocols.

Key Findings

The work done using Power Bi and dataset visualizations to address business problems will be reviewed in this part, along with pertinent findings and suggestions. (Answers of business questions)

1. Country Wise Number of Accidents and Injuries

The following dashboard shows the different countries in which most of aviation accident happened. The world map shown below has the complete locus of the countries with major accidents. This page shows that 41 countries was included in this analysis, in which total 47870 accidents took place.

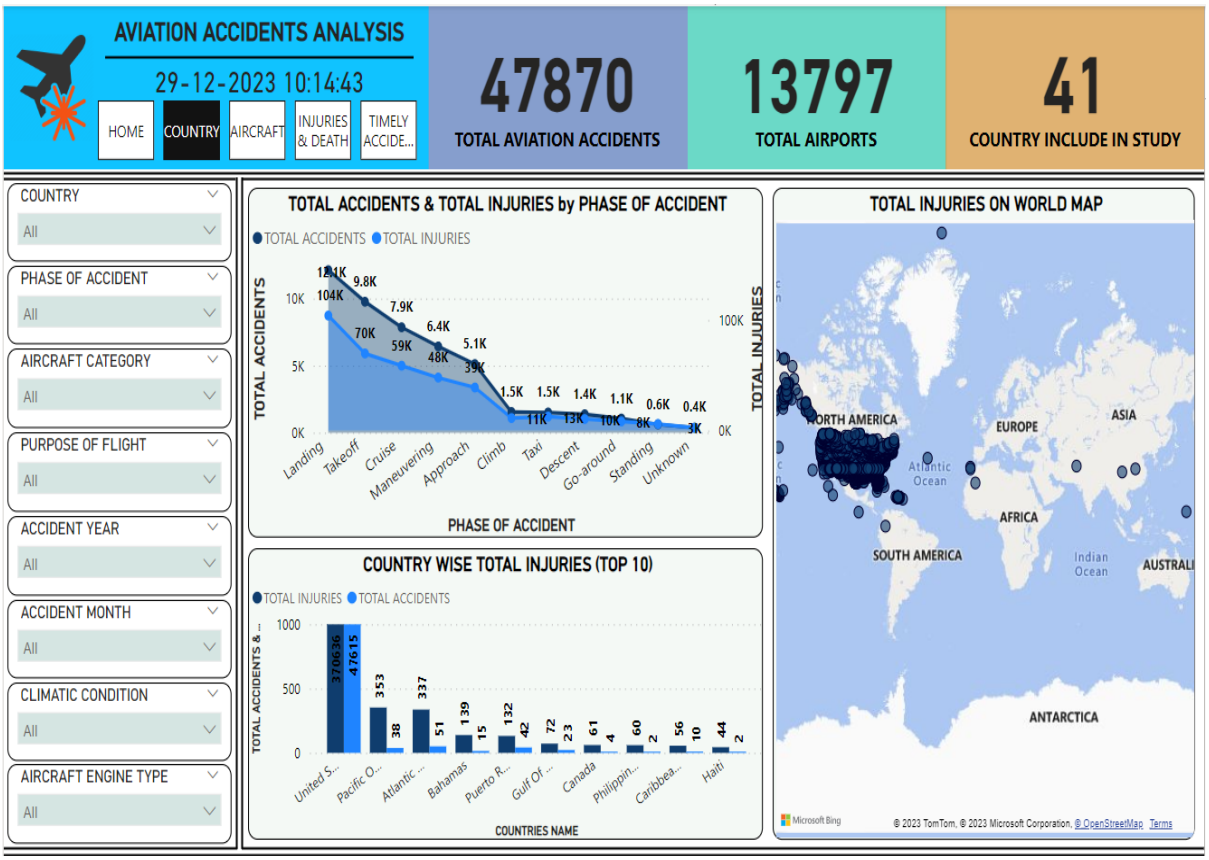


Figure-1: Country Wise Number of Accidents and Injuries dashboard

Most Injured and Accidental Countries:

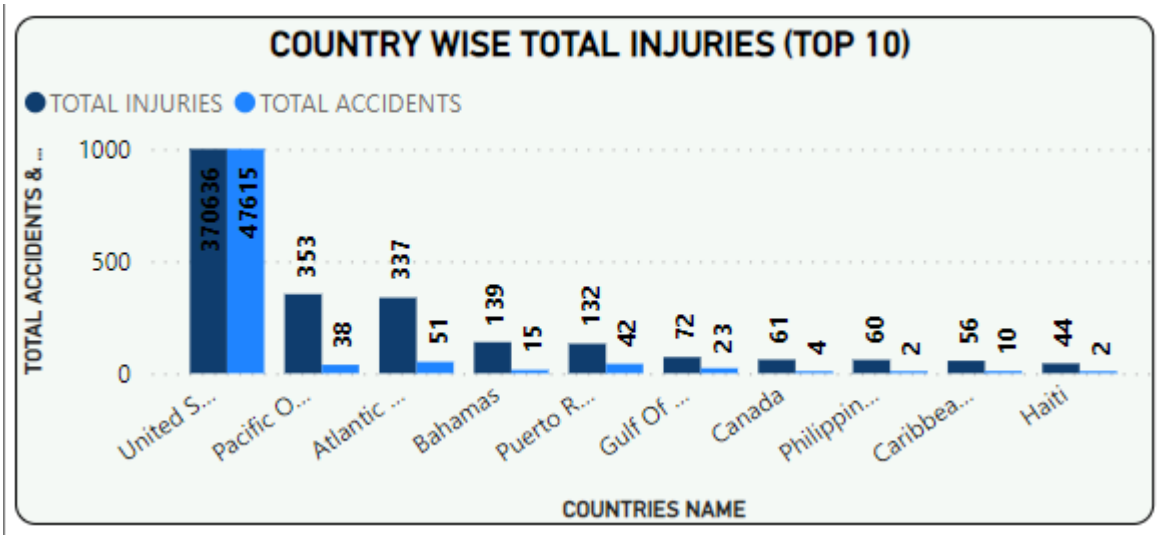


Figure-2: Graph-Most Injured and Accidental Countries

The Clustered column chart shows the number of total injuries and total accidents in countries around the world. The countries are listed on the right side of the graph, with the United States having the most 'total injuries', followed by the Pacific Ocean, the Atlantic Ocean, the Bahamas. The numbers on the left side of the graph show the number of total injuries and total accidents for each country. For example, the United States has 370636 total injuries and 47615 total accidents.

Most Injuries and Accident on Phase of Landing:

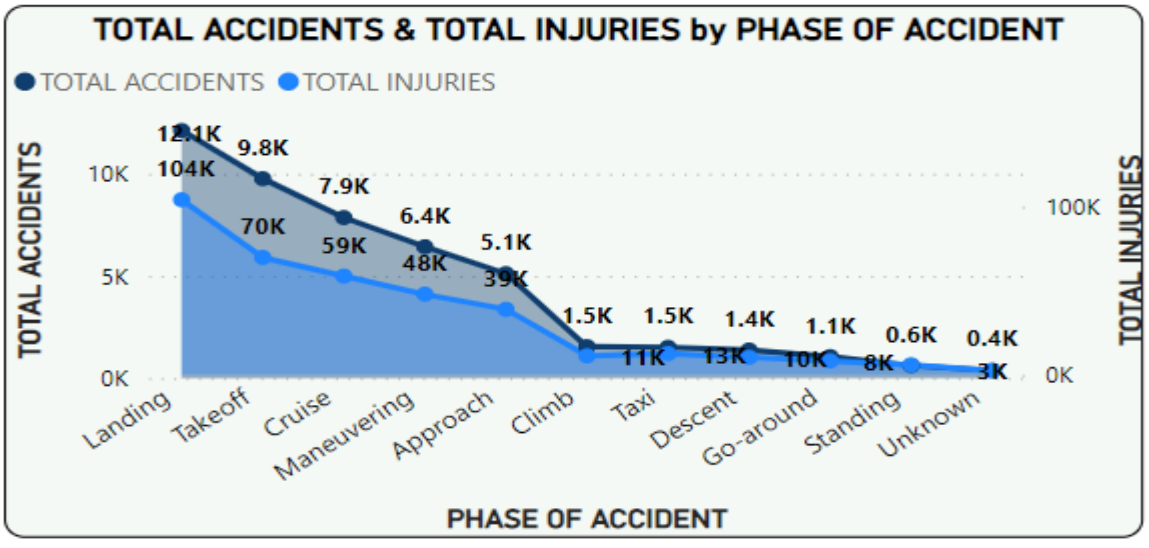


Figure-3: Graph- Most Injuries and Accident on Phase of Landing

The area line chart shows the number of 'total accidents and total injuries by phase of accident'. The phases of accidents are listed on the bottom of the graph, from left to right: Landing, Takeoff, Cruise, Maneuvering, Approach, Climb, Taxi, Descent, Go-around, Standing, and Unknown. The numbers on the left side of the graph show the number of total accidents and total injuries for each phase of accident. For example, the Landing phase has 104K total accidents and 12.1K total injuries.

2. Aircraft Company and Model Wise Accident History

The below page of the dashboard shows the Aircraft Company and Model Wise Accident History through which most of aviation accident happened. There were total 9 aircraft categories having total 6880 aircraft models. There were total 4035 aircraft making companies, which aircraft accident analysis done.

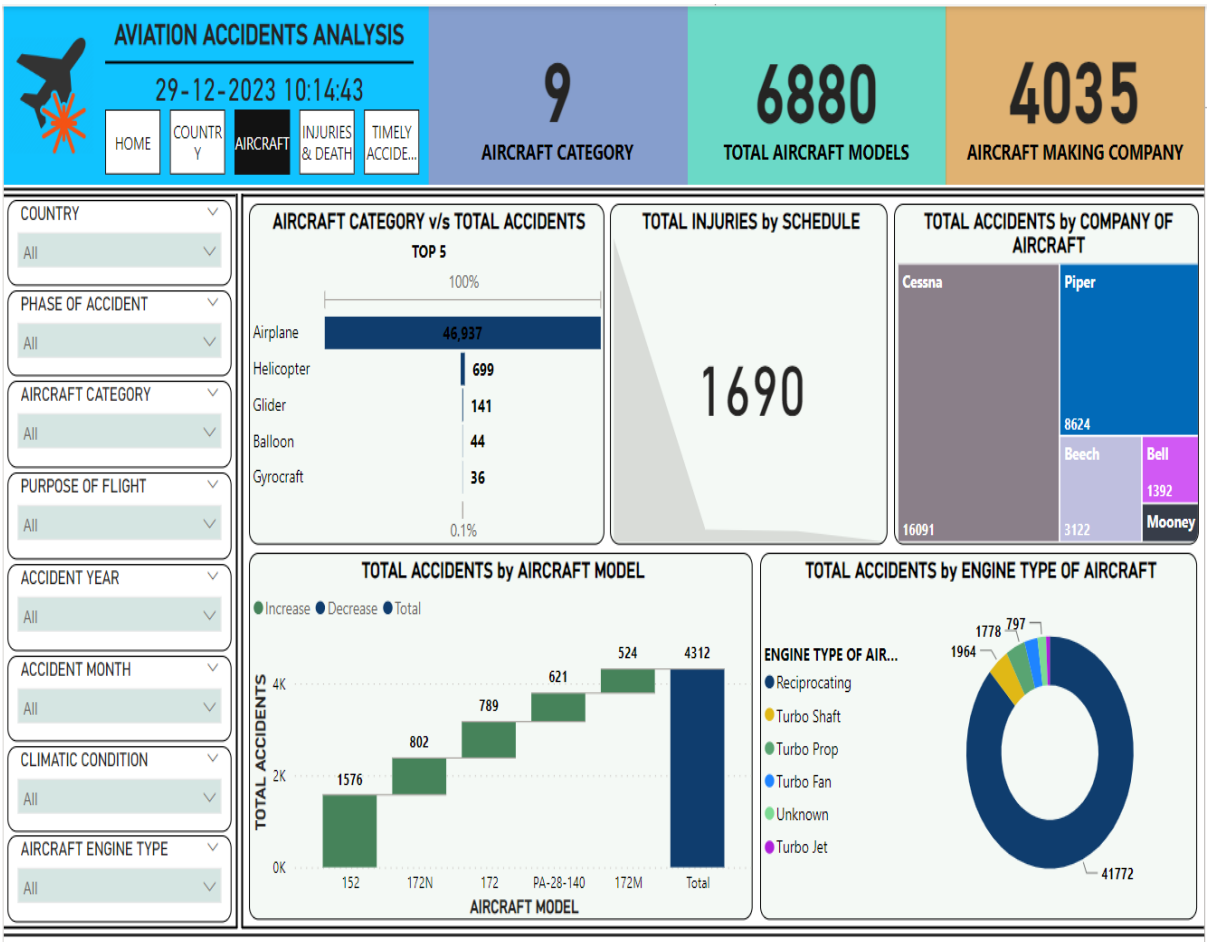


Figure-4: Aircraft Company and Model Wise Accident History dashboard

Total Accidents According to Category of Aircraft:

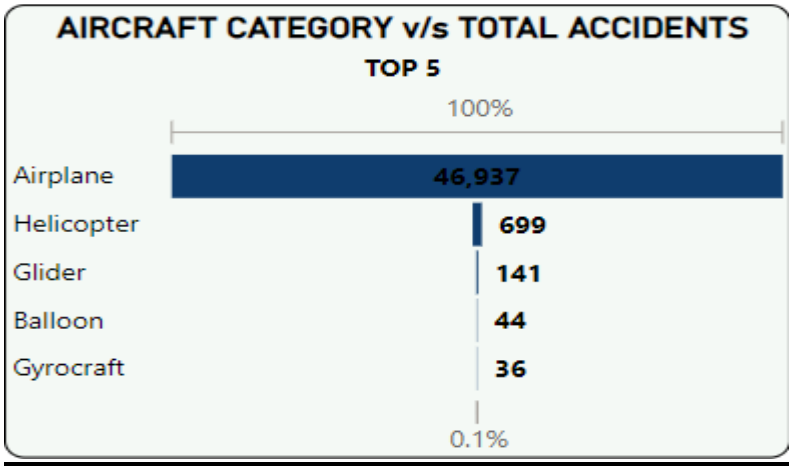


Figure-5: Graph- Total Accidents According to Category of Aircraft

The graph shows the percentage of accidents by type of aircraft, where the aircraft types are listed on the left side of the graph and the percentage of accidents is shown on the right. The airplane has most accident took place, which are 46937 followed by Helicopter with total 699 aviation accidents

Total Accidents According to Model of Aircraft:

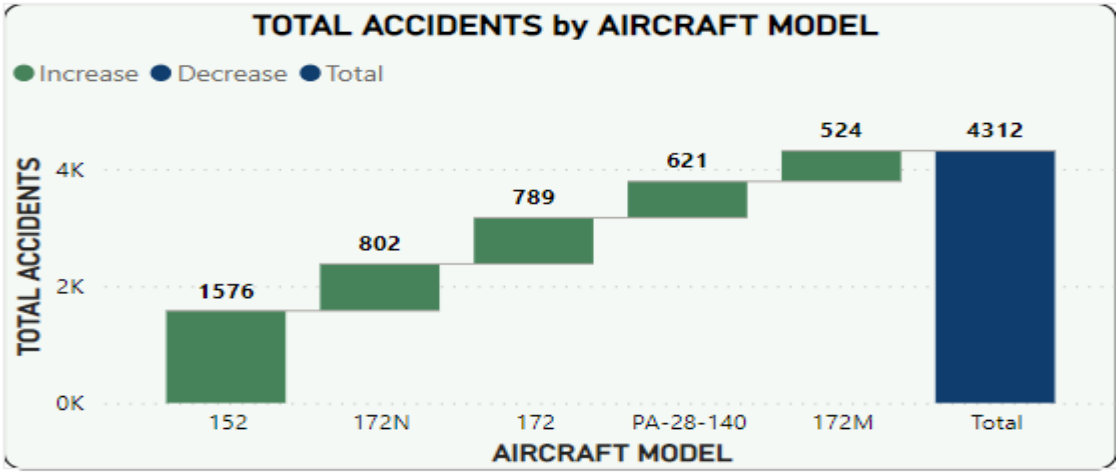


Figure-6: Graph- Total Accidents According to Model of Aircraft

The waterfall chart shows the total number of accidents for different aircraft models. The X-axis lists the aircraft models, while the Y-axis shows the number of accidents. The model '152' has most number of the total accident, followed by model 172N and 172. They have 1576, 802, 789 accidents respectfully.

KPI of Total Injuries of Accidents:

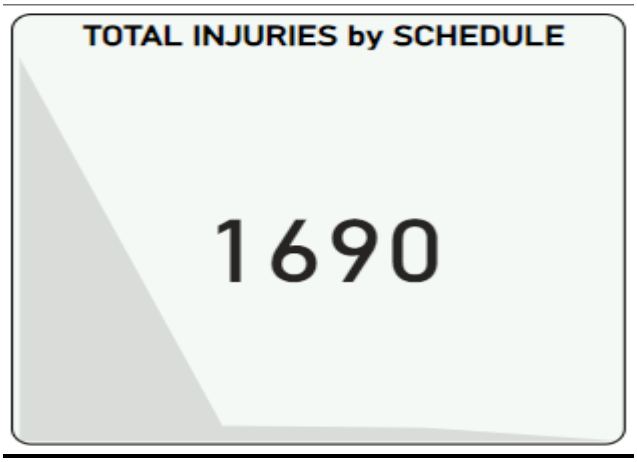


Figure-7: KPI-Total Injuries of Accidents

The above KPI shows the total number of injuries according to aircraft schedule, which is 1690.

Total Accidents According to Aircraft Making Company:

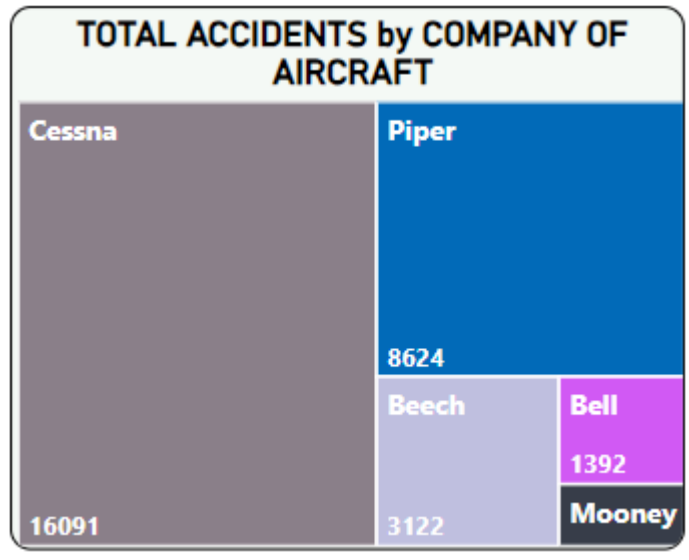


Figure-8: Graph- Total Accidents According to Aircraft Making Company

The Tree-map (tile map) shows a graph of the total number of accidents for different aircraft manufacturers, Aviation Accident and Incident Database. The data in the graph covers accidents that occurred between 1982 and 2020. The top 3 aircraft making companies are 'Cessna, Piper and Beech', having total accidents 16091, 8624, and 3122 respectfully.

3. Finding of the Total Injuries and Death of Accidents

The below page of the dashboard shows the Total Injuries and Death of Accidents in aviation accident. There were total 372278 injuries in all aviation accidents, of which 254175 injuries are serious and 98579 injuries are minor injuries.

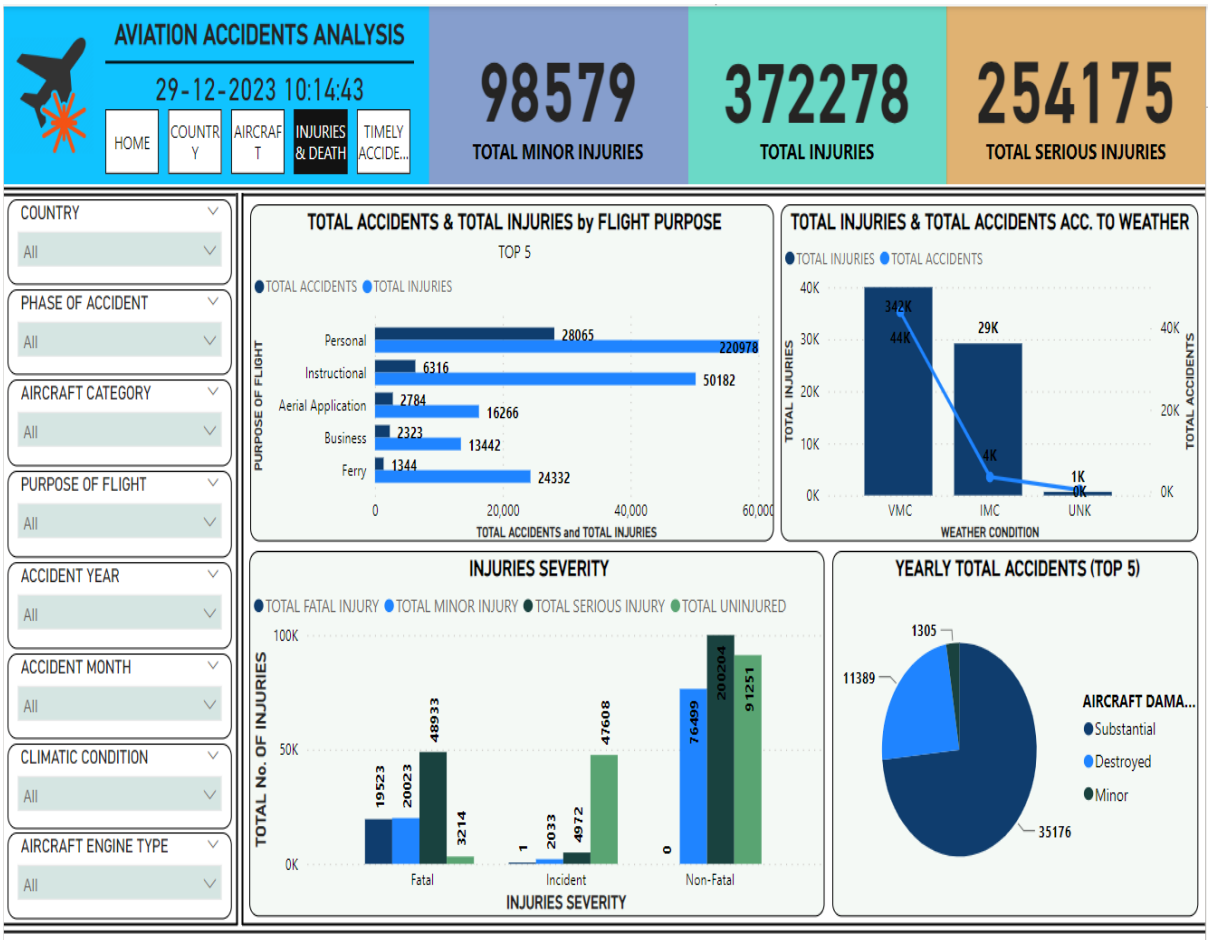


Figure-9: Total Injuries and Death of Accidents dashboard

Total Accidents and Total Accidents:

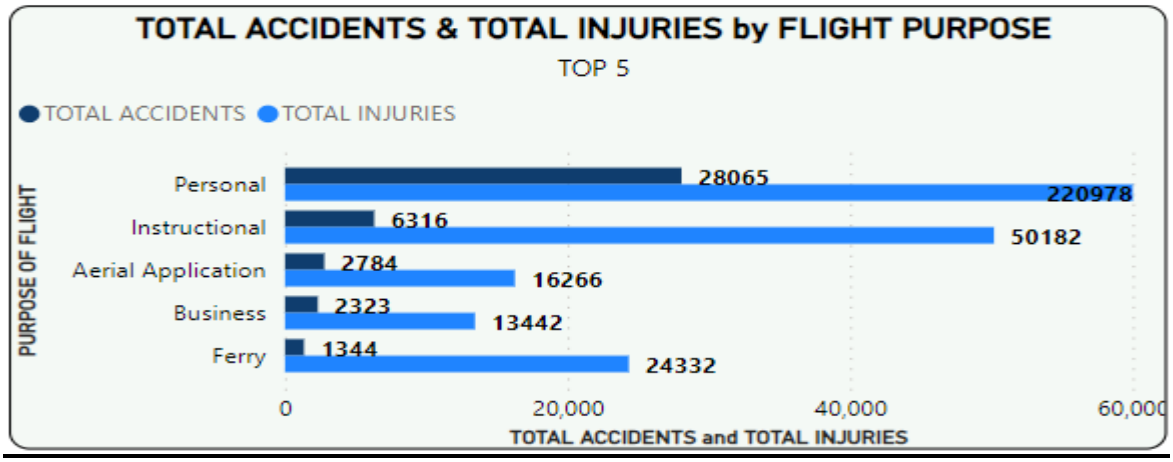


Figure-10: Graph- Total Accidents and Total Accidents

The clustered bar chart shows the total number of accidents and injuries by flight purpose based on data the General Aviation Accident and Incident Database.

The top five flight purposes with the most accidents according to the graph are Personal flying - 28,065 accidents, Instructional flying - 6,316 accidents, Aerial application - 2,784 accidents, Business flying - 2,323 accidents, Ferry flying - 1,344 accidents.

Whereas the top five flight purposes with the most injuries are Personal flying - 220,978 injuries, Instructional flying - 50,182 injuries, Aerial application - 16,266 injuries, Business flying - 13,442 injuries, Ferry flying - 24,332 injuries respectfully.

Total Injuries According to Weather Conditions:

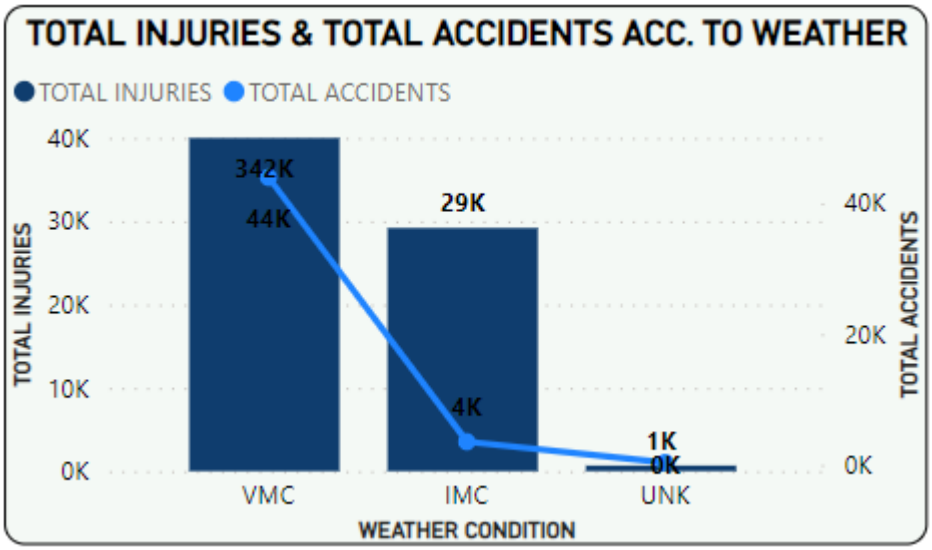


Figure-11: Graph- Total Injuries According to Weather Conditions

The line-Clustered column chart showing the total number of injuries and total accidents according to weather condition. The weather conditions are listed on the left side of the graph, and the number of injuries and accidents are shown on the right side.

The graph shows that the most common weather condition for both accidents and injuries is VMC (visual meteorological conditions), followed by IMC (instrument meteorological conditions) and OK (other known conditions).

Aircraft Damage Conditions:

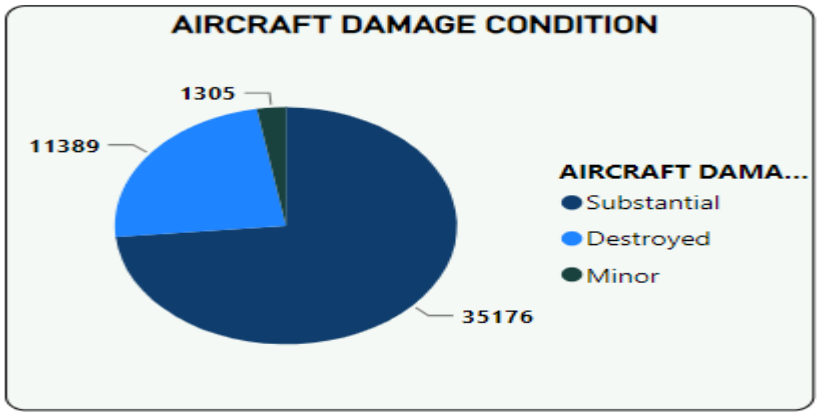


Figure-12: Graph- Aircraft Damage Conditions

The pie chart shows the aircraft damage condition of various aircraft. Total 35176 aircrafts are substantial damaged whereas 11389 are destroyed completely.

Injuries Severity:

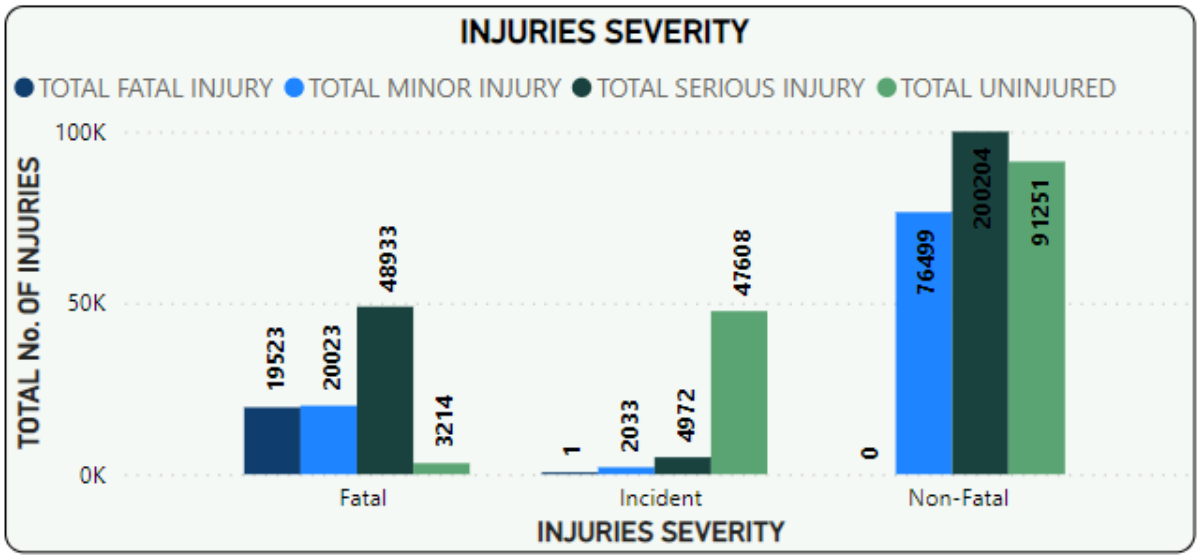


Figure-13: Graph-Injuries Severity

The clustered column chart shows the injuries severity of aviation accidents. The 'Non-Fatal type injuries' are higher than other type of injuries. There were total 200204 injuries of serious category in Non-Fatal Severity.

However, Incident type injury severity are much lower than other injuries severity.

4. Accidents and Injuries of Aviation Accidents by Time-Duration

The below page of the dashboard shows the Accidents and Injuries of Aviation Accidents by Time-Duration. There were total 47870 aviation accidents. The number of airlines are surveyed were 1664 in 41 countries worldwide.

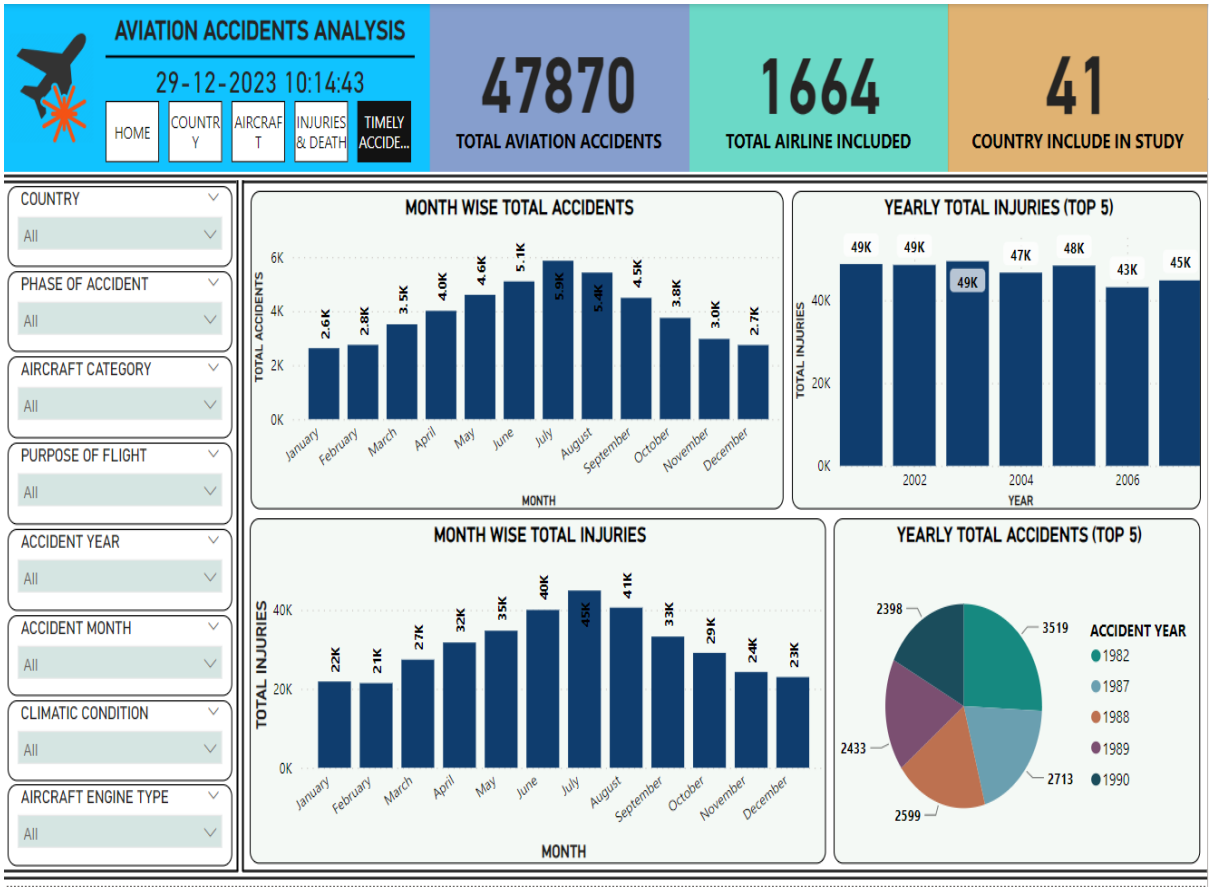


Figure-14: Accidents and Injuries of Aviation Accidents by Time-Duration dashboard

Month wise total Accidents:

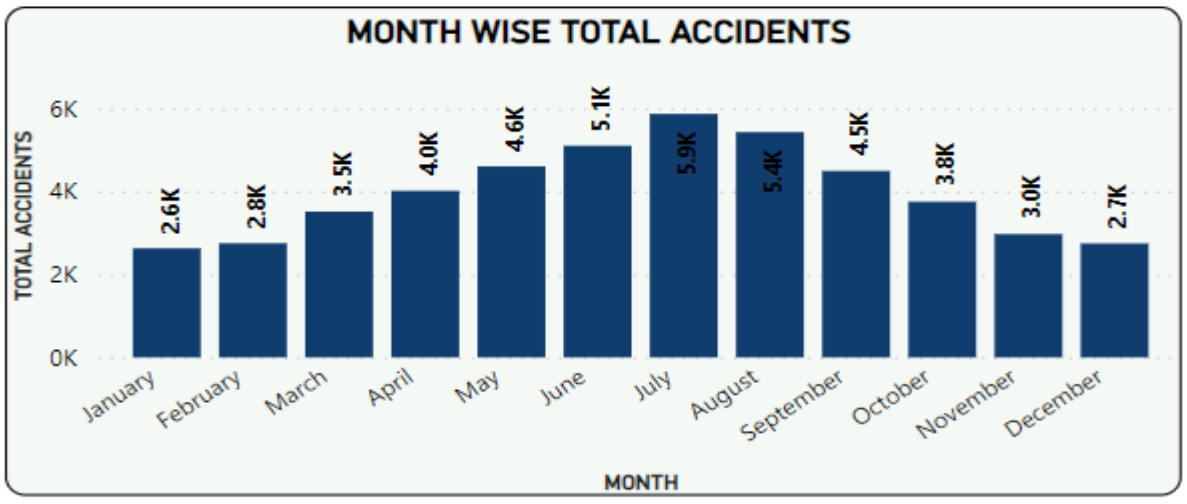


Figure-15: Graph- Month wise total Accidents

The clustered column chart shows the number of accidents per month for commercial aviation, general aviation, and air taxis/charter operations. The y-axis appears to show the number of accidents whereas the x-axis appears to show the months of the year.

The July, August and June months has most number of aviation accidents, which are 5.9K, 5.4K and 5.1K

It looks like the number of accidents for all three categories is highest in the summer months and lowest in the winter months.

Month wise total Injuries:

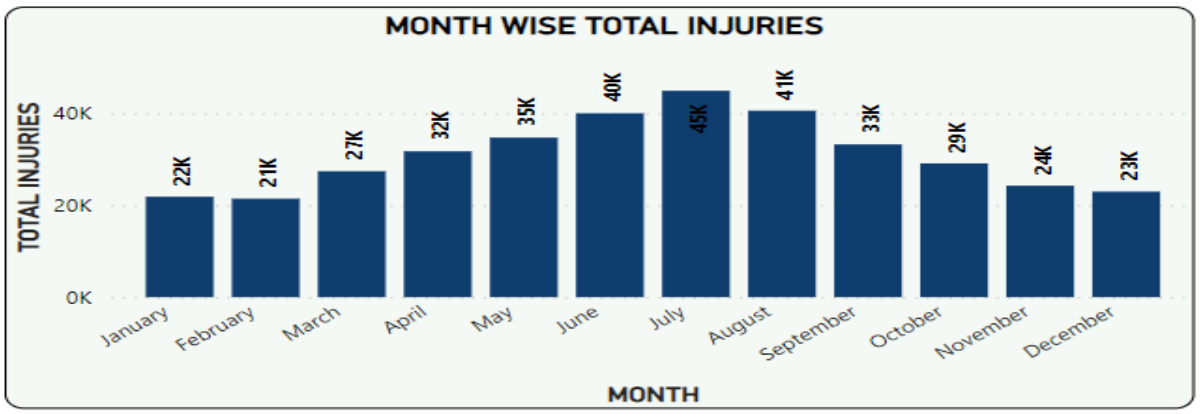


Figure-16: Graph- Month wise total Injuries

The clustered column chart visualize the "Monthly Wise Total Injuries. The X-axis labelled "Month," it likely shows the months of the year whereas the Y-axis labelled "Total Injuries".

The month July, August and June has the most number of injuries which are 45K, 41K and 40K respectfully. However, the month February has the lowest number of injuries i.e. 21K only.

Year wise total Injuries:

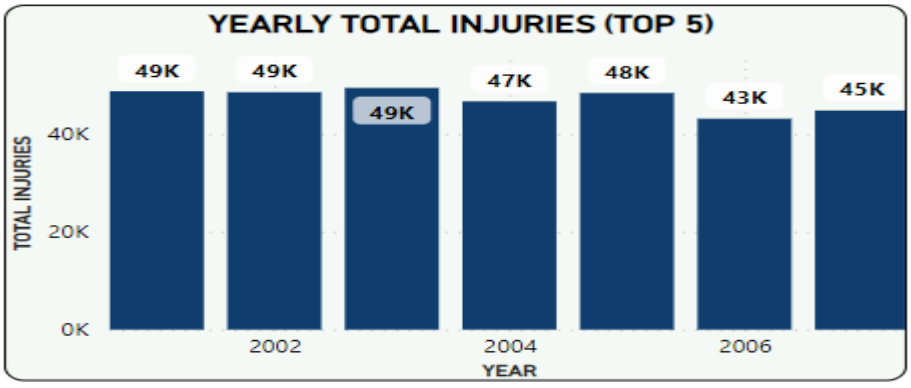


Figure-17: Graph- Year wise total Injuries

The clustered column chart titled as "Yearly Total Injuries (Top 5)" with data apparently from the year 2002 to 2006.

The X-axis labelled "YEAR," it shows the top 5 years from the dataset which are from 2002 to 2006. The Y-axis labelled "TOTAL INJURIES," it likely represents the total number of injuries recorded within each year.

Year wise total Accidents:

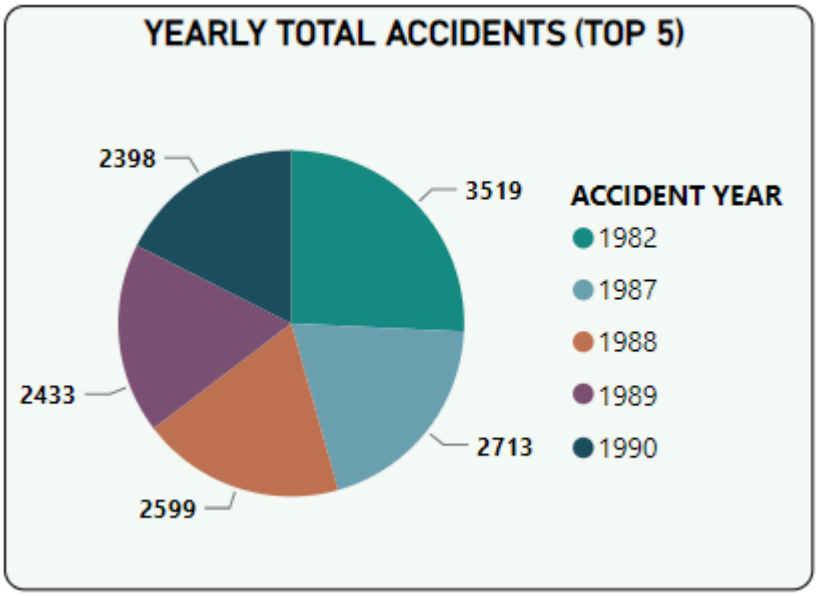


Figure-18: Graph- Year wise total Accidents

The above pie chart shows the top 5 year in which most numbers of accidents took placed. The top 5 years are 1982, 1987, 1988, 1989 and 1990. The year 1982 has most number of accidents which are 3519 followed by year 1987 having 2713 accidents.

Conclusions

To sum up, the thorough examination of the **Aviation dataset** has yielded insightful information about a range of aviation-related topics. This report provides a comprehensive summary of incidents ranging from accidents to various injury types and damage severity by exploring characteristics like Event ID, Investigation Type, and Aircraft Category. The principal objective of the research was to augment comprehension of aircraft accidents and facilitate continuous enhancements in safety protocols. This report has effectively provided a thorough picture of aviation occurrences, spanning everything from environmental conditions to accidents, through the usage of Power BI visualizations. In the end, the project is a useful tool that supports the ongoing improvement of aviation safety procedures and practices.

While studying all the points, we have reached some general conclusions which are as follows:

- The number of accidents and injuries occurring at the time of landing was highest.
- The United States has the highest number of accidents.
- Aircraft model number 152 caused the most accidents.
- Highest number of air accidents were recorded in the year 1982

Recommendations

- Cessna company should be honored for having the least number of accidents.
- Aircraft company should work to improve the quality of its model number 152.
- The aircrafts which are having reciprocating type engine should either replace it or improve it.

References

1. Fultz, A.J. and Ashley, W.S., 2016. Fatal weather-related general aviation accidents in the United States. *Physical Geography*, *37*(5), pp.291-312.
2. Johnson, C. and Holloway, C., 2007. A longitudinal analysis of the causal factors in major aviation accidents in the USA from 1976 to 2006.
3. Patriarca, R., Di Gravio, G., Cioponea, R. and Licu, A., 2019. Safety intelligence: Incremental proactive risk management for holistic aviation safety performance. *Safety science*, *118*, pp.551-567.
4. Amato, A., Cozzolino, G., Maisto, A. and Pelosi, S., 2021. Monitoring Airplanes Faults Through Business Intelligence Tools. In *Advances on P2P, Parallel, Grid, Cloud and Internet Computing: Proceedings of the 15th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC-2020) 15* (pp. 224-234). Springer International Publishing.
5. Bhargava, M.G., Phani, K.T., Kiran, S. and Rao, D.R., 2018. Analysis and design of visualization of educational institution database using power bi tool. *Global Journal of Computer Science and Technology*, *18*, pp.1-8.
6. Aspin, A., 2014. *High Impact Data Visualization with Power View, Power Map, and Power BI*. Apress.

Appendix

Load the data sheet in power-bi

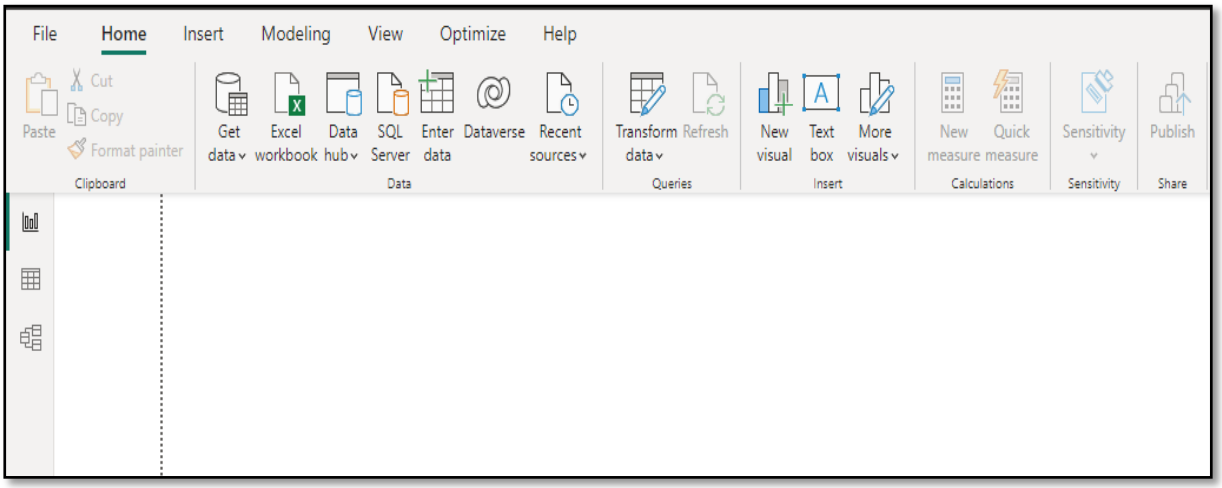
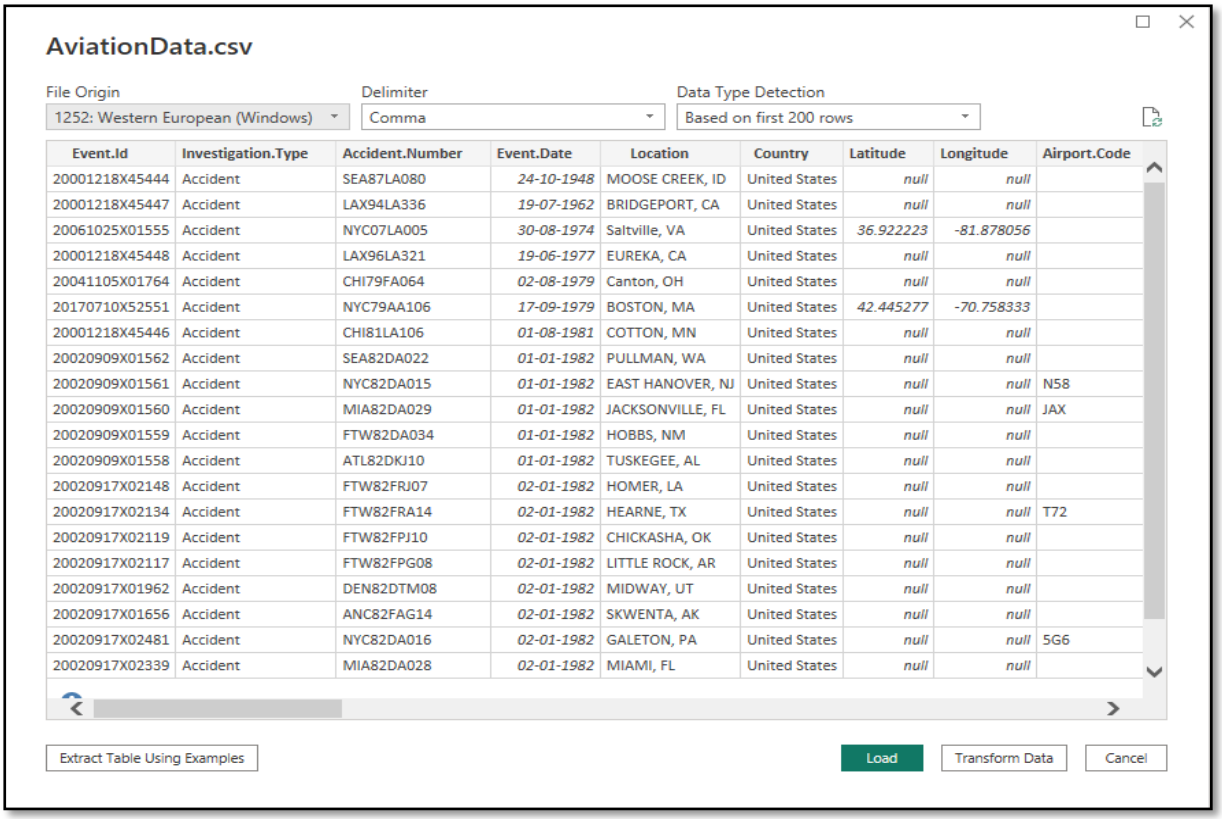


Figure-19: Data Importing Option

- After that, select the file from Excel. And load them.



Event.Id	Investigation.Type	Accident.Number	Event.Date	Location	Country	Latitude	Longitude	Airport.Code
20001218X45444	Accident	SEA87LA080	24-10-1948	MOOSE CREEK, ID	United States	null	null	
20001218X45447	Accident	LAX94LA336	19-07-1962	BRIDGEPORT, CA	United States	null	null	
20061025X01555	Accident	NYC07LA005	30-08-1974	Saltville, VA	United States	36.922223	-81.878056	
20001218X45448	Accident	LAX96LA321	19-06-1977	EUREKA, CA	United States	null	null	
20041105X01764	Accident	CHI79FA064	02-08-1979	Canton, OH	United States	null	null	
20170710X52551	Accident	NYC79AA106	17-09-1979	BOSTON, MA	United States	42.445277	-70.758333	
20001218X45446	Accident	CHI81LA106	01-08-1981	COTTON, MN	United States	null	null	
20020909X01562	Accident	SEA82DA022	01-01-1982	PULLMAN, WA	United States	null	null	
20020909X01561	Accident	NYC82DA015	01-01-1982	EAST HANOVER, NJ	United States	null	null	N58
20020909X01560	Accident	MIA82DA029	01-01-1982	JACKSONVILLE, FL	United States	null	null	JAX
20020909X01559	Accident	FTW82DA034	01-01-1982	HOBBS, NM	United States	null	null	
20020909X01558	Accident	ATL82DKJ10	01-01-1982	TUSKEGEE, AL	United States	null	null	
20020917X02148	Accident	FTW82FRJ07	02-01-1982	HOMER, LA	United States	null	null	
20020917X02134	Accident	FTW82FRA14	02-01-1982	HEARNE, TX	United States	null	null	T72
20020917X02119	Accident	FTW82FPJ10	02-01-1982	CHICKASHA, OK	United States	null	null	
20020917X02117	Accident	FTW82FPG08	02-01-1982	LITTLE ROCK, AR	United States	null	null	
20020917X01962	Accident	DEN82DTM08	02-01-1982	MIDWAY, UT	United States	null	null	
20020917X01656	Accident	ANC82FAG14	02-01-1982	SKWENTA, AK	United States	null	null	
20020917X02481	Accident	NYC82DA016	02-01-1982	GALETON, PA	United States	null	null	5G6
20020917X02339	Accident	MIA82DA028	02-01-1982	MIAMI, FL	United States	null	null	

Figure-20: Data Sheet loading

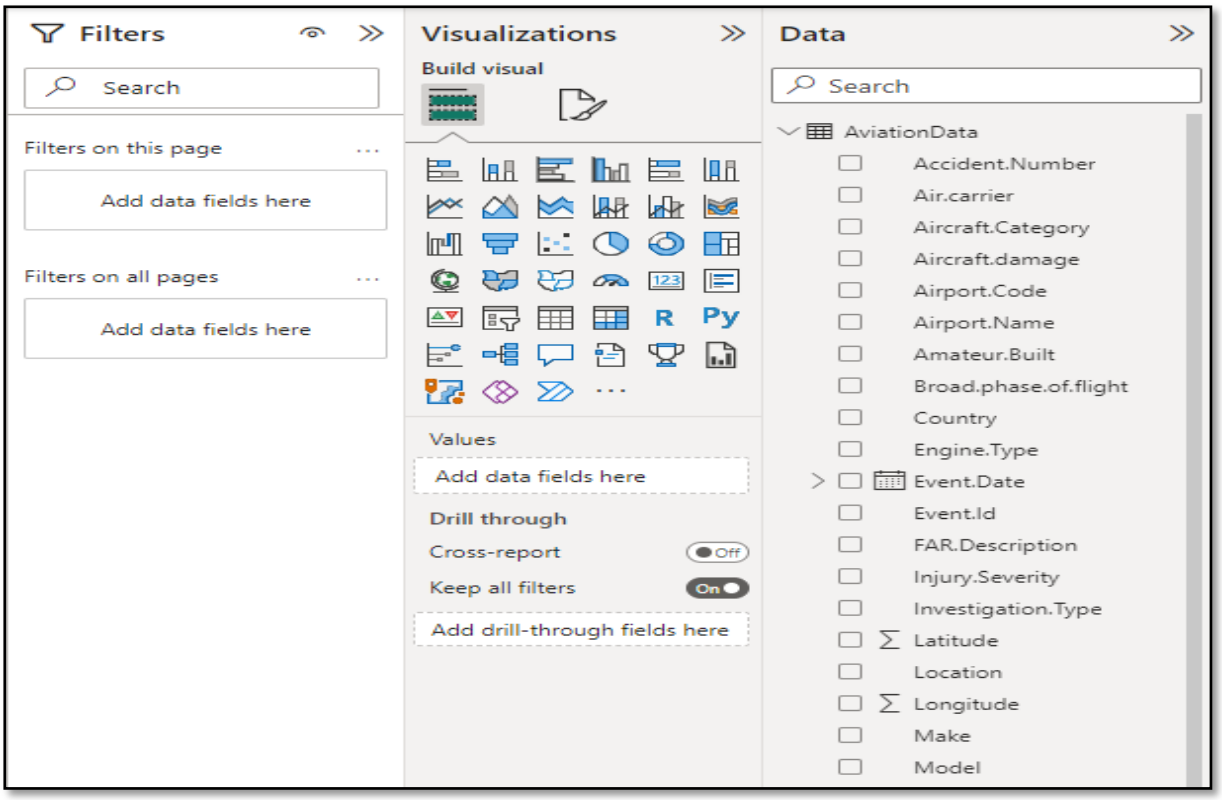


Figure-21: Data View In power BI

Data Cleansing & Transformation Steps

Data transformation in Power BI Desktop empowers to shape raw data into meaningful insights. It acts as sculpting tool, transforming messy, disorganized data into a polished masterpiece ready for analysis. Someone can cleanse and restructure data, calculate new measures, merge datasets, and even pivot tables to gain different perspectives. Power BI offers a visual interface with intuitive tools like M Query and the Query Editor, making data transformation accessible even for non-programmers.

In order to transform the data set, click on “Transform data” option available in the ribbon menu.

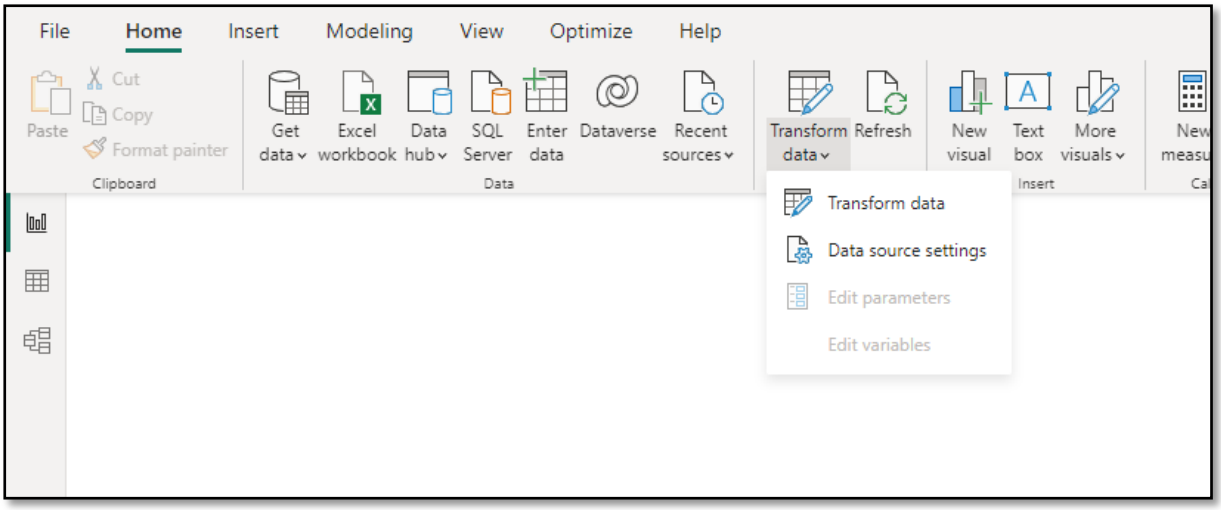
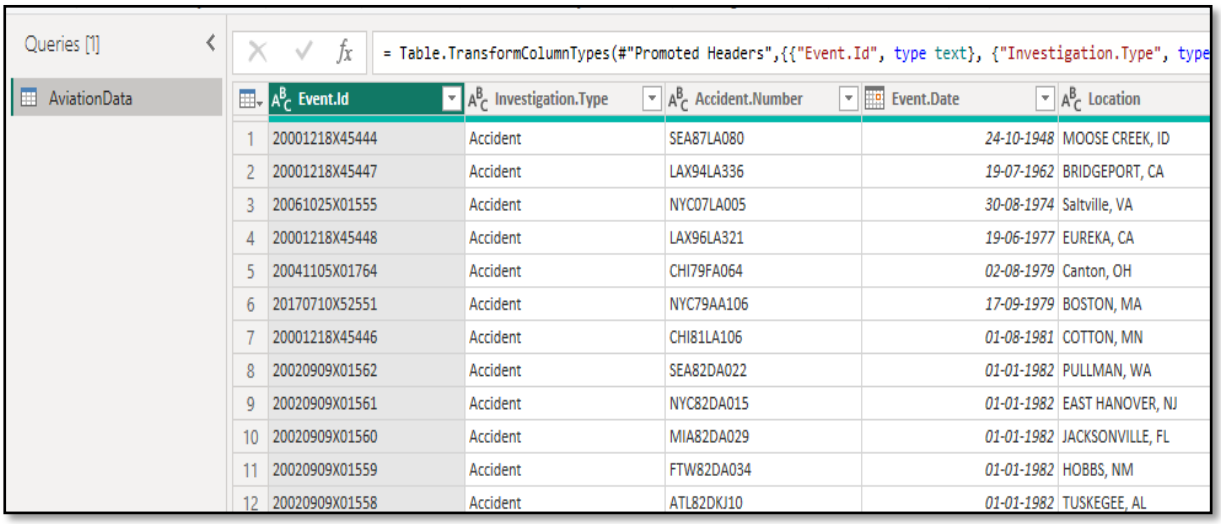


Figure-22: Data Transformation Option



	Event.Id	Investigation.Type	Accident.Number	Event.Date	Location
1	20001218X45444	Accident	SEA87LA080	24-10-1948	MOOSE CREEK, ID
2	20001218X45447	Accident	LAX94LA336	19-07-1962	BRIDGEPORT, CA
3	20061025X01555	Accident	NYC07LA005	30-08-1974	Saltville, VA
4	20001218X45448	Accident	LAX96LA321	19-06-1977	EUREKA, CA
5	20041105X01764	Accident	CHI79FA064	02-08-1979	Canton, OH
6	20170710X52551	Accident	NYC79AA106	17-09-1979	BOSTON, MA
7	20001218X45446	Accident	CHI81LA106	01-08-1981	COTTON, MN
8	20020909X01562	Accident	SEA82DA022	01-01-1982	PULLMAN, WA
9	20020909X01561	Accident	NYC82DA015	01-01-1982	EAST HANOVER, NJ
10	20020909X01560	Accident	MIA82DA029	01-01-1982	JACKSONVILLE, FL
11	20020909X01559	Accident	FTW82DA034	01-01-1982	HOBBS, NM
12	20020909X01558	Accident	ATL82DKJ10	01-01-1982	TUSKEGEE, AL

Figure-23: View of Data Transformation

Removing Null Columns from Data Sheet

- If the datasheet has any column completely null then it can be removed from the data sheet.

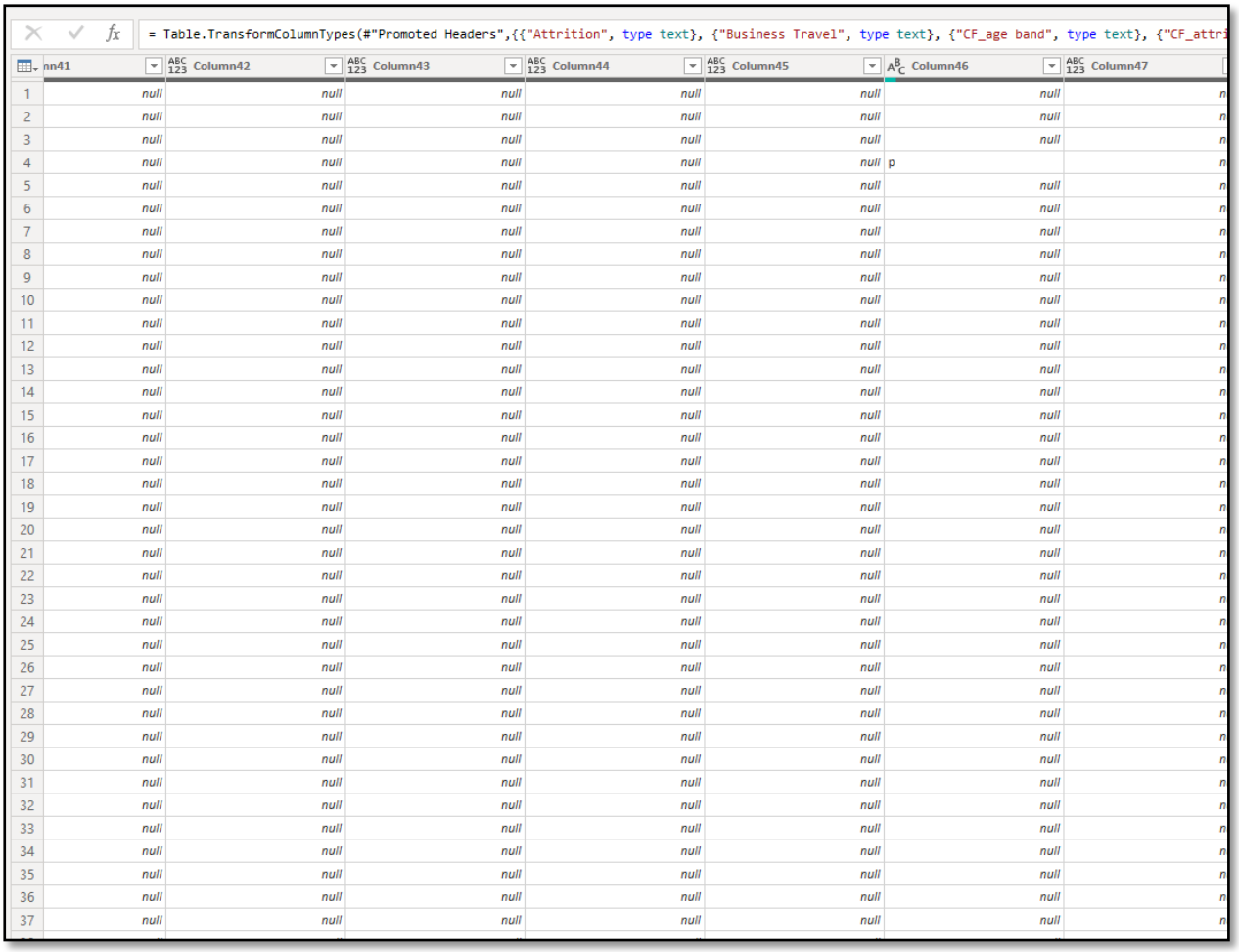


Figure-24: Null Columns in Dataset

- The null column can removed by two ways,

  (a) Either by options “remove Columns” in Ribbon tab or

  (b) By clicking right column on null column header.

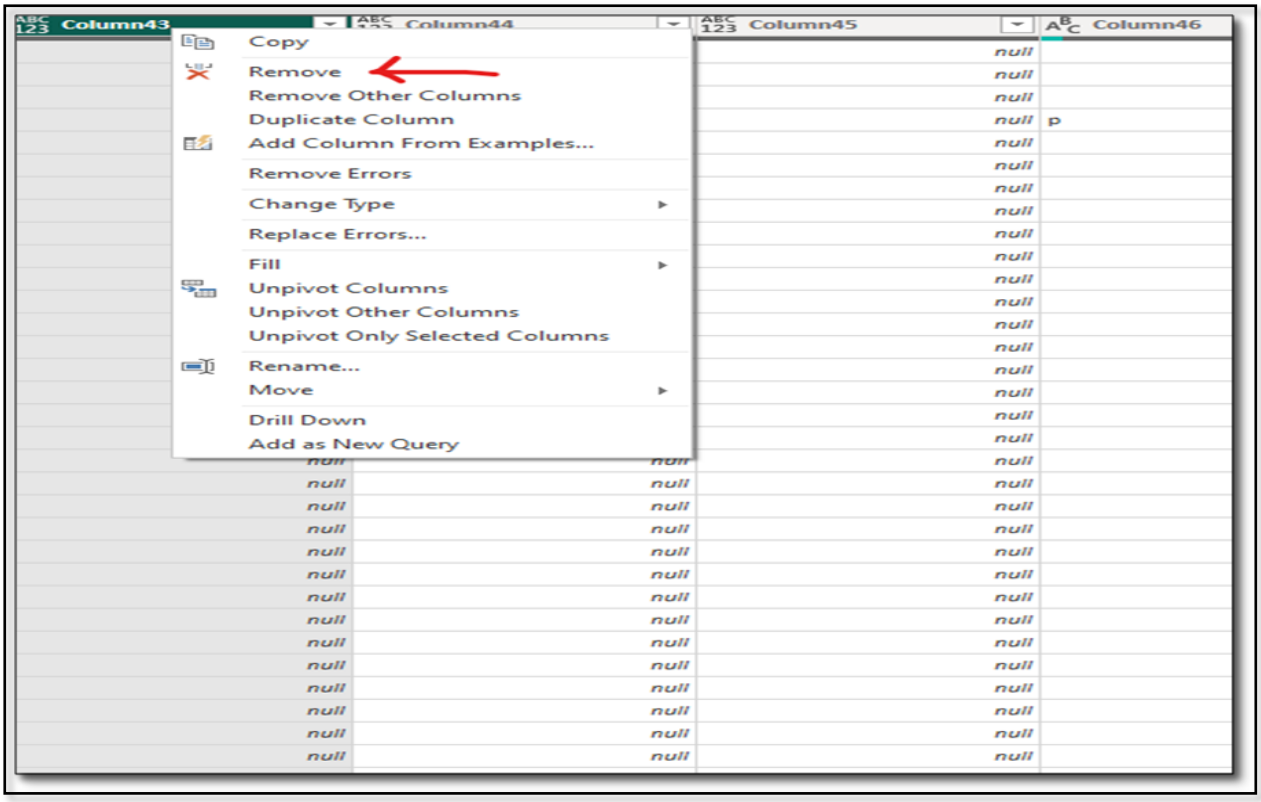


Figure-25: Removing Null Columns

Removing Null values in Column

- Delete null values from Power BI data sets helps in preventing the problems. We may make sure that your analyses and visualizations are based on valid and full data by eliminating null values.
- Removing null values can also speed up the processing of your Power BI dashboards and reports because it will require less data to analyze.
- If any column is having null then it can be replaced using replace values.

i. From column aircraft.Category

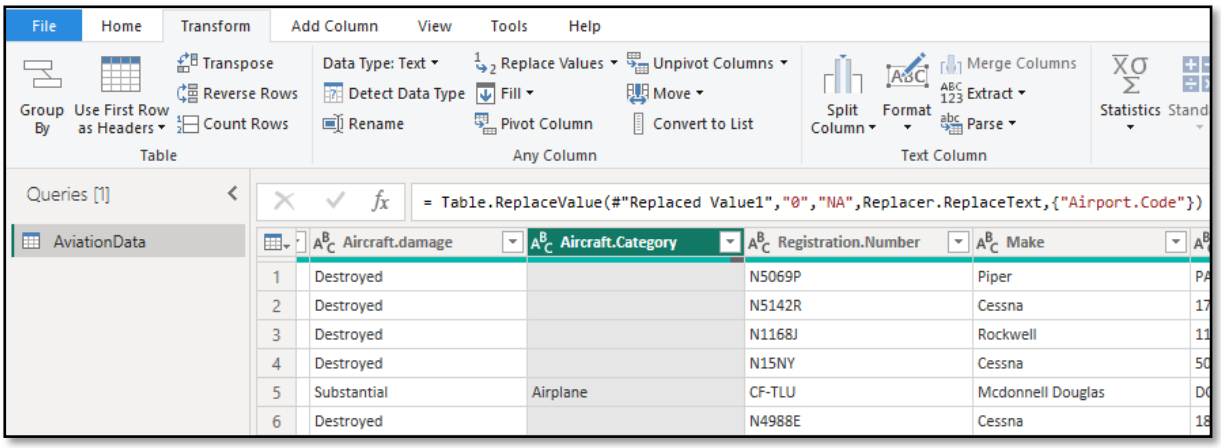


Figure-26: Replacing Null values from column aircraft.Category

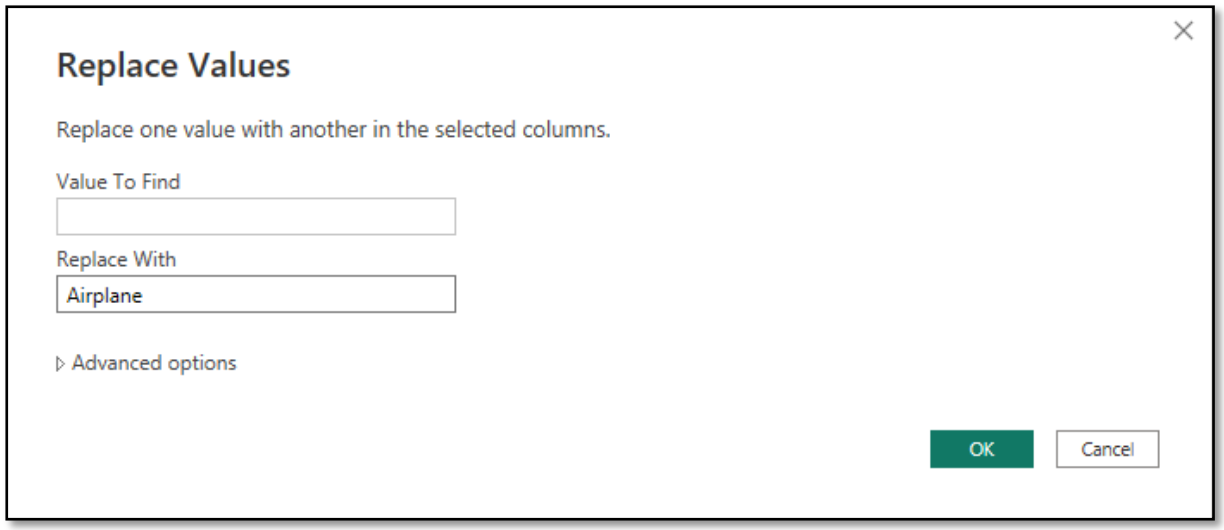


Figure-27: Replacing Null values

Changing Data Types of Columns

i. For the column Number.of.Engines
Number.of.Engines Decimal Number → Whole Number

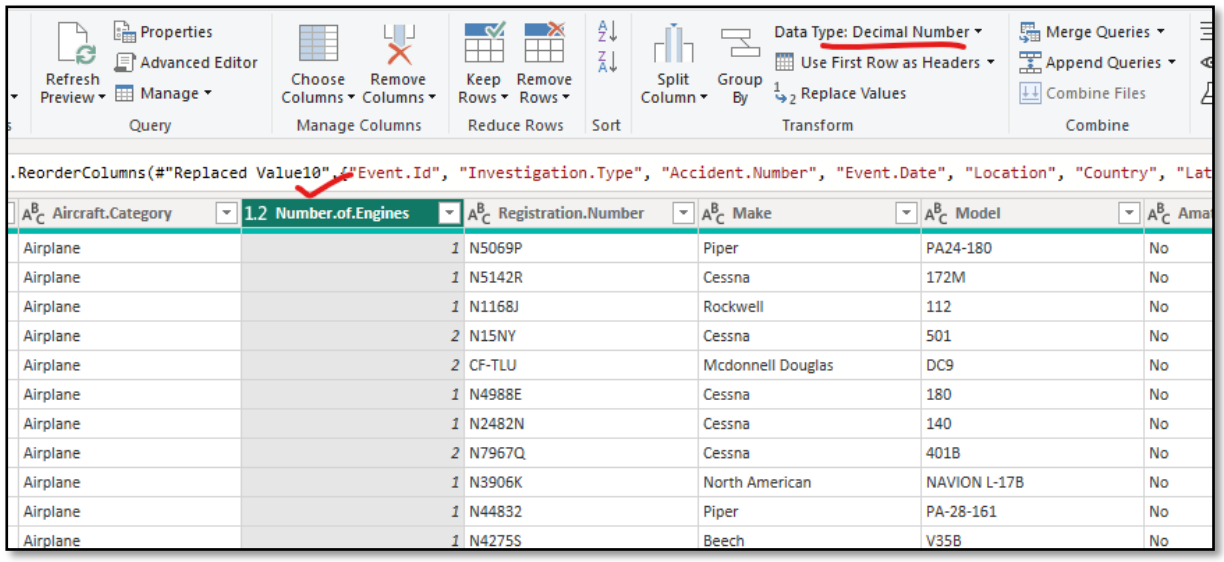


Figure-28: Data Type of 'Number.of.Engines'

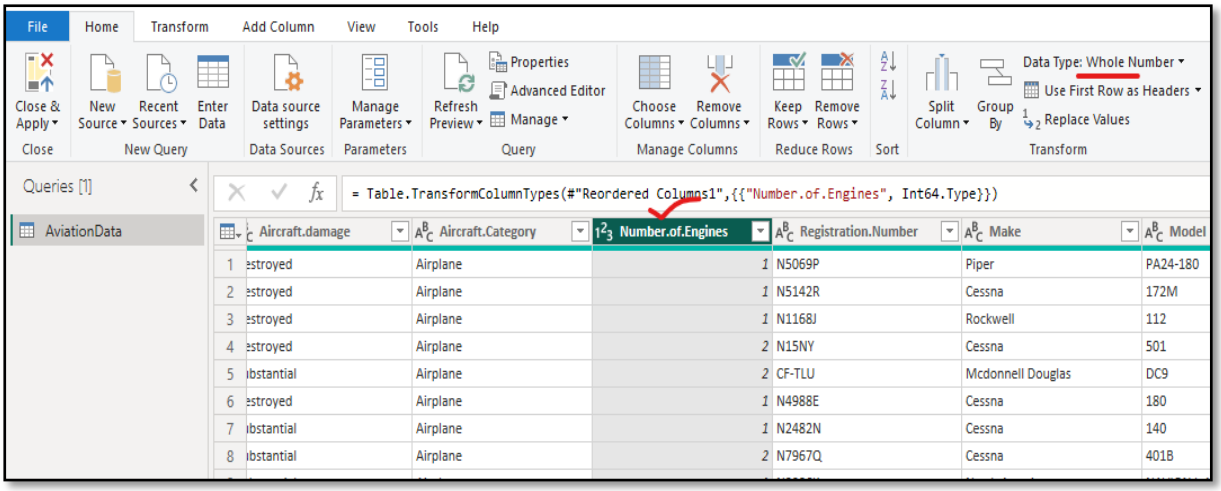


Figure-29: Changed Data Type of 'Number.of.Engines'

= Table.TransformColumnTypes(#"Reordered Columns1",{{"Number.of.Engines", Int64.Type}})

ii. For the column Total.Fatal.Injuries

Total.Fatal.Injuries Decimal Number → Whole Number

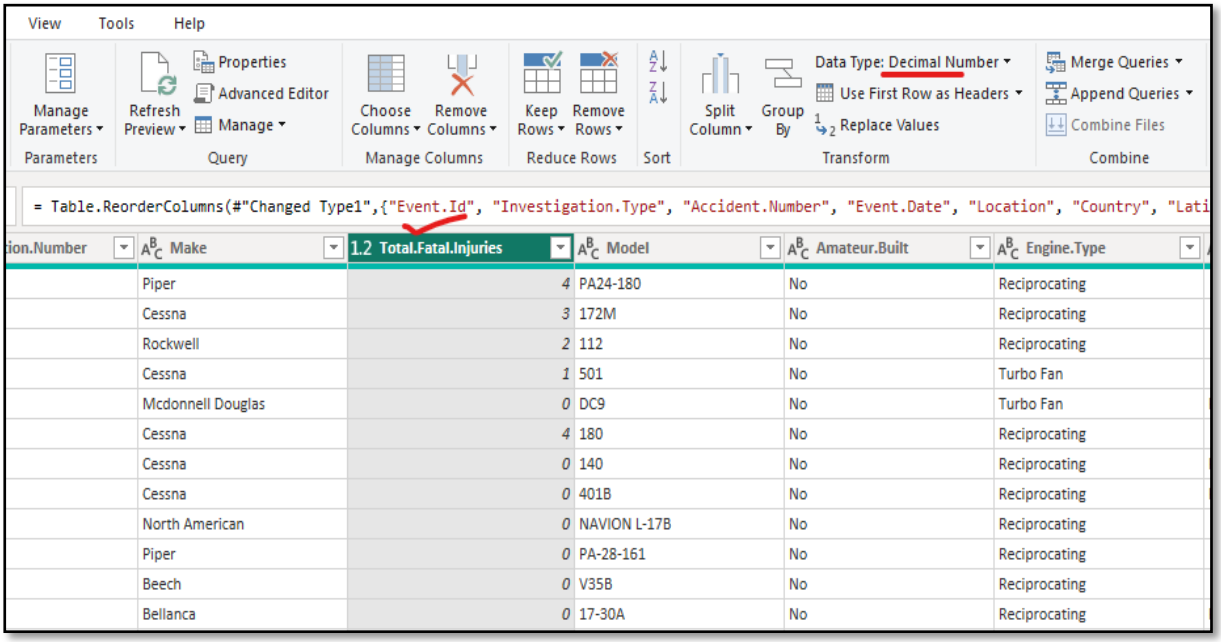


Figure-30: Data Type of 'Total.Fatal.Injuries'

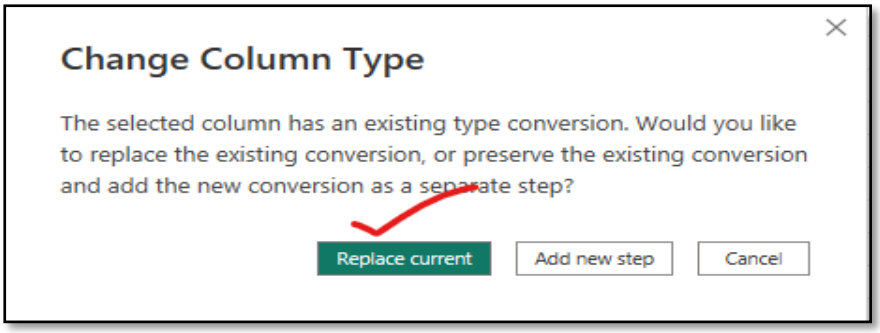


Figure-31: Data Type Changing Pop-up

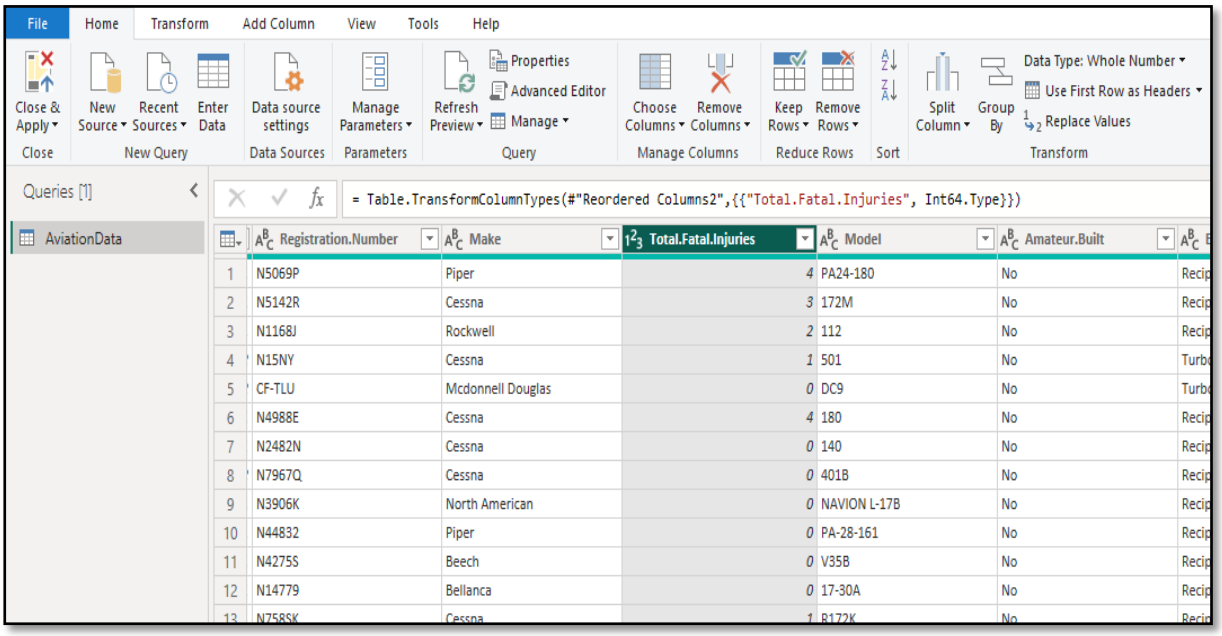


Figure-32: Changed Data Type of 'Total.Fatal.Injuries'

= Table.TransformColumnTypes(#"Reordered Columns2",{{"Total.Fatal.Injuries", Int64.Type}})

Adding an Index Column

- Power BI Desktop lets us easily add an index column to our data tables, offering several ways to customize its numbering. The "Add Column" tab provides a simple "Index Column" option, automatically assigning sequential numbers starting from 0. Alternatively, we can configure the starting number, increment, and even choose "Custom" to define your own unique numbering logic. These index columns can prove

valuable for sorting visuals, tracking order, or supporting other transformation patterns within your Power BI reports.

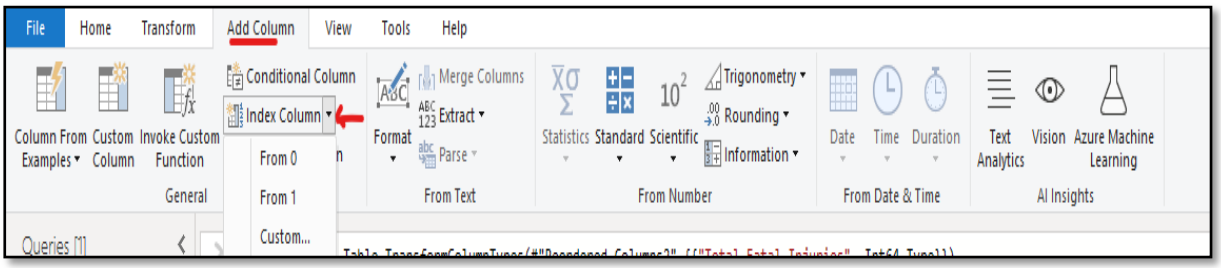


Figure-33: Index Column Option

Broad.phase.of.flight	Report.Status	Publication.Date	Index
Unknown	Probable Cause	19-09-1996	1
Cruise	Probable Cause	26-02-2007	2
Cruise	Probable Cause	12-09-2000	3
Approach	Probable Cause	16-04-1980	4
Climb	Probable Cause	19-09-2017	5
Unknown	Probable Cause	06-11-2001	6
Takeoff	Probable Cause	01-01-1982	7
Landing	Probable Cause	01-01-1982	8
Cruise	Probable Cause	01-01-1982	9
Approach	Probable Cause	01-01-1982	10
Landing	Probable Cause	01-01-1982	11
Cruise	Probable Cause	02-01-1983	12
Takeoff	Probable Cause	02-01-1983	13
Cruise	Probable Cause	02-01-1983	14
Cruise	Probable Cause	02-01-1983	15
Taxi	Probable Cause	02-01-1983	16
Unknown	Probable Cause	02-01-1983	17
Taxi	Probable Cause	02-01-1983	18
Cruise	Probable Cause	02-01-1983	19

Figure-34: Index Column added

= Table.AddIndexColumn(#"Removed Columns", "Index", 1, 1, Int64.Type)

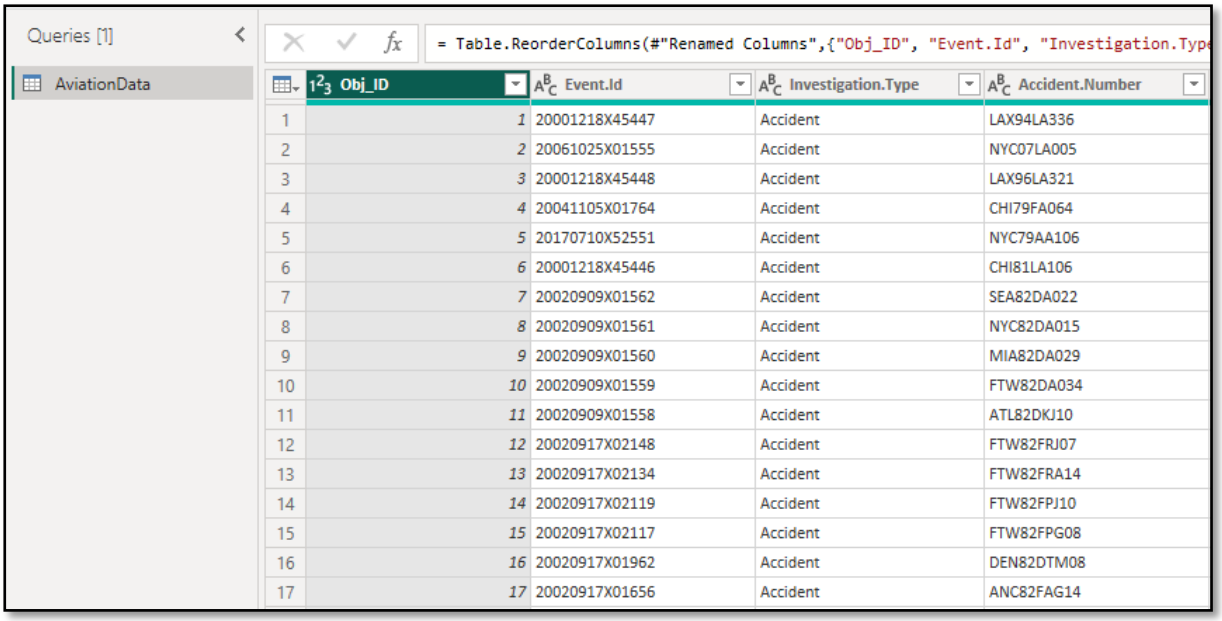


Figure-35: Index Column name changed

= Table.RenameColumns(#"Added Index1",{{"Index", "Obj_ID"}})

Bi Data Modeling

Fact and Dimension in Datasheet

Dimension Table:

- In a data warehouse, a dimension table is a table that contains dimensions—the descriptive information about the data/information. Time, location, product, customer, and employee are a few examples of dimensions.
- The dimension table is a crucial part of the Star Schema. The dimensions along which the values of the attributes are taken in the fact table are included in a dimension table.

Fact Table:

- A fact table is a table in a data warehouse that stores quantitative data or metrics, also known as facts. These facts are typically numerical measures, such as sales revenue, quantity sold, or profit.
- Fact tables also contain foreign keys that link to the primary keys of related dimension tables. These relationships enable analysts to combine quantitative data with descriptive attributes for comprehensive analysis.

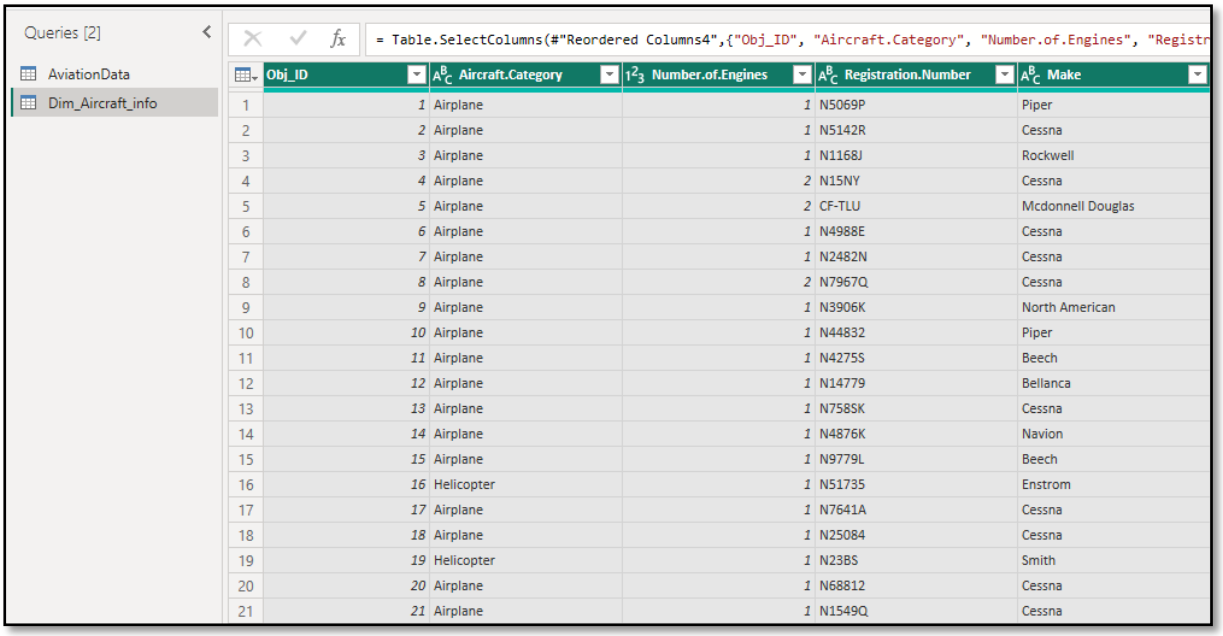
Difference between Fact and Dimension Table

Feature	Fact Tables	Dimension Tables
Purpose	Stores quantitative, measurable data	Stores descriptive, contextual data
Granularity	Contains the most granular level of detail	Contains aggregated or summarized information
Structure	Typically contains many rows	Typically contains fewer rows
Primary key	Often a composite key, combining multiple dimensions	Usually has a single primary key, uniquely identifying each entity
Relationships	Connected to multiple dimension tables via foreign keys	Typically not directly connected to other dimension tables
Data types	Primarily numeric values	Primarily text or descriptive values, but can include dates or codes

Dimension Dim_Aircraft_info

The "Dim_Aircraft_info" table contains the following columns-

i. Obj_ID
ii. Aircraft.Category
iii. Number.of.Engines
iv. Registration.Number
v. Make
vi. Model
vii. Amateur.Built
viii. Engine.Type
ix. FAR.Description



	Obj_ID	Aircraft.Category	Number.of.Engines	Registration.Number	Make
1	1	Airplane	1	N5069P	Piper
2	2	Airplane	1	N5142R	Cessna
3	3	Airplane	1	N1168J	Rockwell
4	4	Airplane	2	N15NY	Cessna
5	5	Airplane	2	CF-TLU	Mcdonnell Douglas
6	6	Airplane	1	N4988E	Cessna
7	7	Airplane	1	N2482N	Cessna
8	8	Airplane	2	N7967Q	Cessna
9	9	Airplane	1	N3906K	North American
10	10	Airplane	1	N44832	Piper
11	11	Airplane	1	N4275S	Beech
12	12	Airplane	1	N14779	Bellanca
13	13	Airplane	1	N758SK	Cessna
14	14	Airplane	1	N4876K	Navion
15	15	Airplane	1	N9779L	Beech
16	16	Helicopter	1	N51735	Enstrom
17	17	Airplane	1	N7641A	Cessna
18	18	Airplane	1	N25084	Cessna
19	19	Helicopter	1	N23BS	Smith
20	20	Airplane	1	N68812	Cessna
21	21	Airplane	1	N1549Q	Cessna

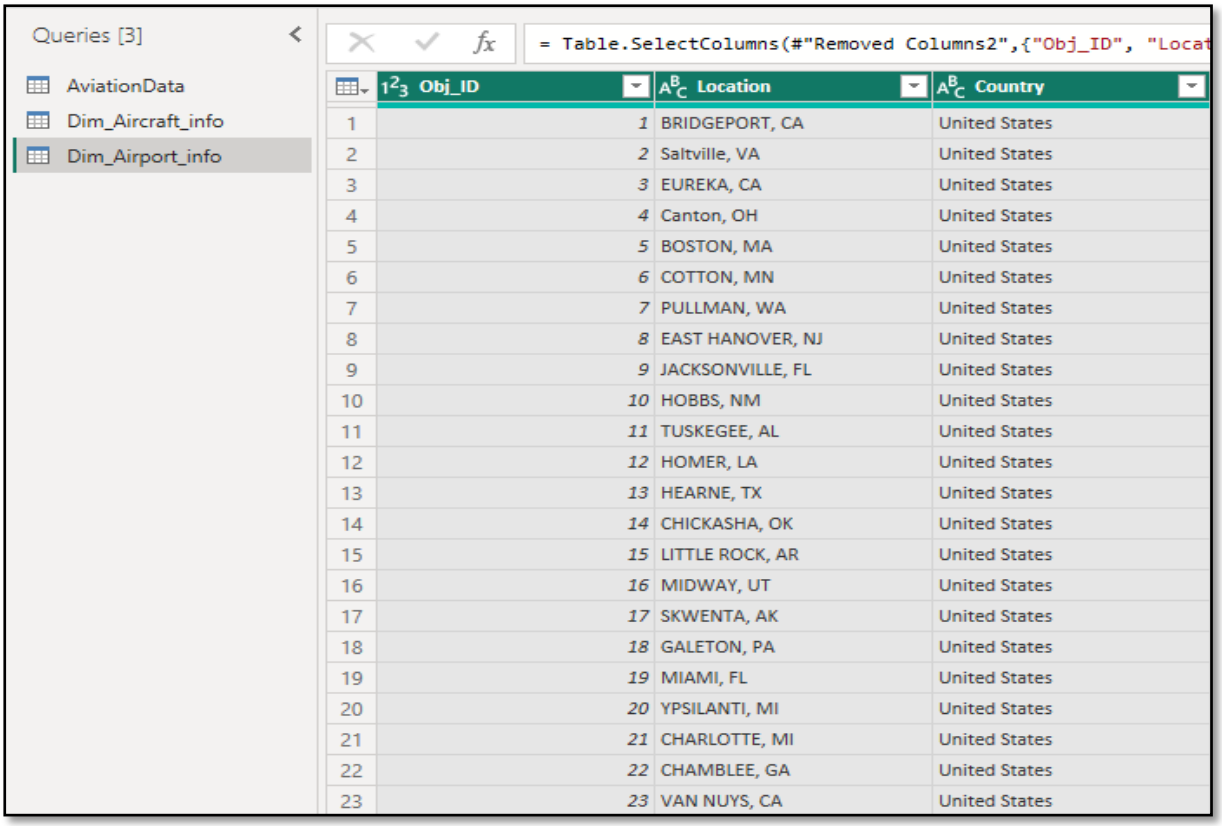
Figure-36: creation of Dimension Dim_Aircraft_info Table

= Table.SelectColumns(#"Reordered Columns4",{"Obj_ID", "Aircraft.Category", "Number.of.Engines", "Registration.Number", "Make", "Model", "Amateur.Built", "Engine.Type", "FAR.Description"})

Dimension Dim_Airport_info

The “Dim_Airport_info” table contains the following columns-

i. Obj_ID
ii. Location
iii. Country
iv. Latitude
v. Longitude
vi. Airport.Code
vii. Airport.Name



	Obj_ID	Location	Country
1	1	BRIDGEPORT, CA	United States
2	2	Saltville, VA	United States
3	3	EUREKA, CA	United States
4	4	Canton, OH	United States
5	5	BOSTON, MA	United States
6	6	COTTON, MN	United States
7	7	PULLMAN, WA	United States
8	8	EAST HANOVER, NJ	United States
9	9	JACKSONVILLE, FL	United States
10	10	HOBBS, NM	United States
11	11	TUSKEGEE, AL	United States
12	12	HOMER, LA	United States
13	13	HEARNE, TX	United States
14	14	CHICKASHA, OK	United States
15	15	LITTLE ROCK, AR	United States
16	16	MIDWAY, UT	United States
17	17	SKWENTA, AK	United States
18	18	GALETON, PA	United States
19	19	MIAMI, FL	United States
20	20	YPSILANTI, MI	United States
21	21	CHARLOTTE, MI	United States
22	22	CHAMBLEE, GA	United States
23	23	VAN NUYS, CA	United States

Figure-37: creation of Dimension Dim_Airport_info Table

= Table.SelectColumns(#"Removed Columns2",{"Obj_ID", "Location", "Country", "Latitude", "Longitude", "Airport.Code", "Airport.Name"})

Dimension Dim_Flight_info

The “Dim_Flight_info” table contains the following columns-

i. Obj_ID
ii. Schedule
iii. Purpose.of.flight
iv. Air.carrier

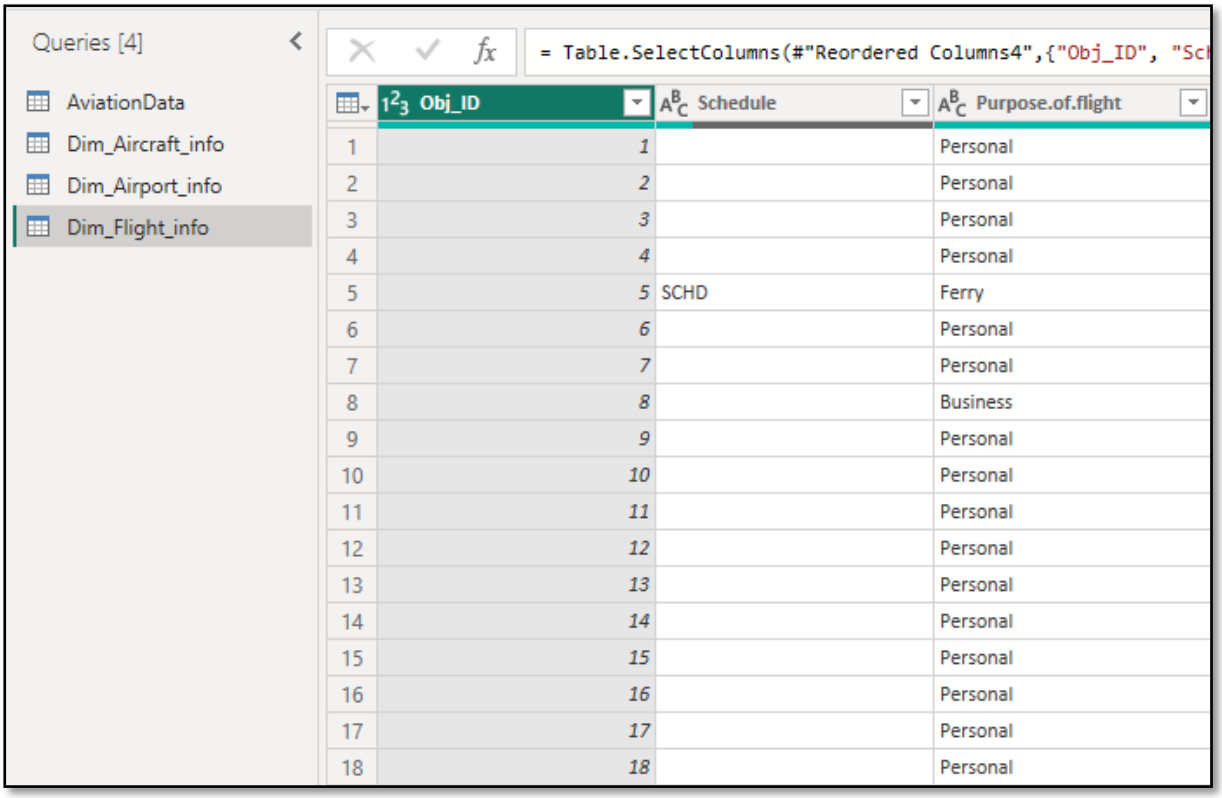


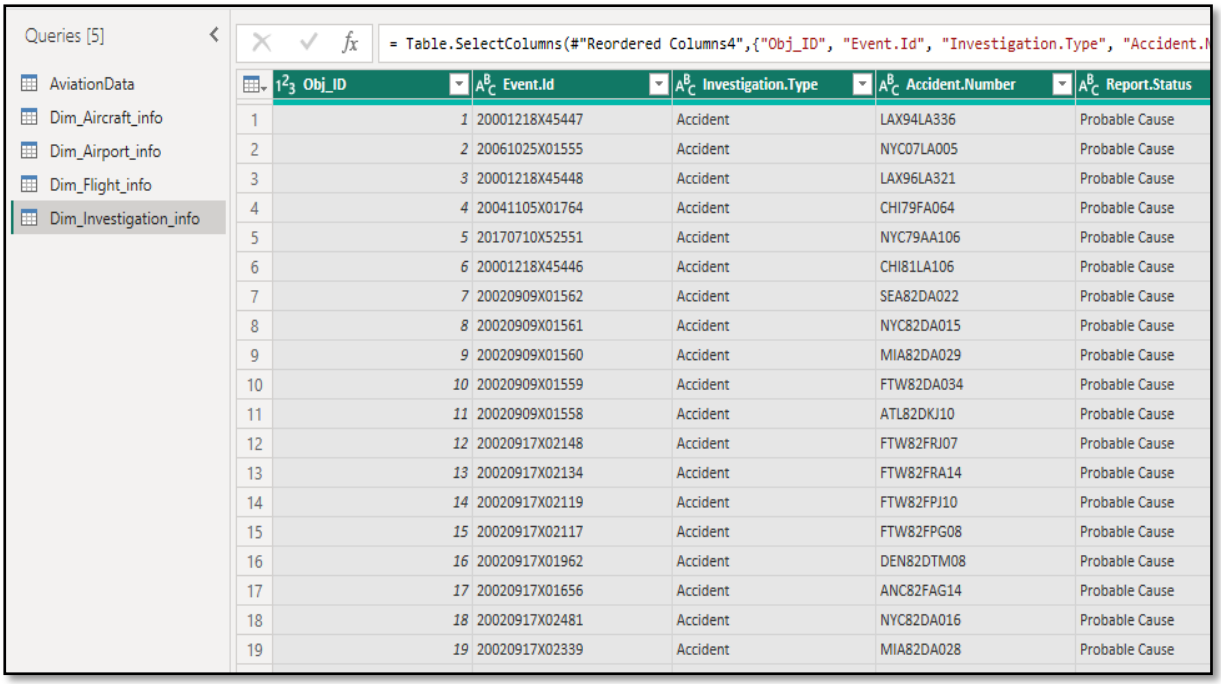
Figure-38: creation of Dimension Dim_Flight_info Table

= Table.SelectColumns(#"Reordered Columns4",{"Obj_ID", "Schedule", "Purpose.of.flight", "Air.carrier"})

Dimension Dim_Investigation_info

The "Dim_Investigation_info" table contains the following columns-

i. Obj_ID
ii. Event.Id
iii. Investigation.Type
iv. Report.Status
v. Publication.Date



	Obj_ID	Event.Id	Investigation.Type	Accident.Number	Report.Status
1	1	20001218X45447	Accident	LAX94LA336	Probable Cause
2	2	20061025X01555	Accident	NYC07LA005	Probable Cause
3	3	20001218X45448	Accident	LAX96LA321	Probable Cause
4	4	20041105X01764	Accident	CHI79FA064	Probable Cause
5	5	20170710X52551	Accident	NYC79AA106	Probable Cause
6	6	20001218X45446	Accident	CHI81LA106	Probable Cause
7	7	20020909X01562	Accident	SEA82DA022	Probable Cause
8	8	20020909X01561	Accident	NYC82DA015	Probable Cause
9	9	20020909X01560	Accident	MIA82DA029	Probable Cause
10	10	20020909X01559	Accident	FTW82DA034	Probable Cause
11	11	20020909X01558	Accident	ATL82DKJ10	Probable Cause
12	12	20020917X02148	Accident	FTW82FRJ07	Probable Cause
13	13	20020917X02134	Accident	FTW82FRA14	Probable Cause
14	14	20020917X02119	Accident	FTW82FPJ10	Probable Cause
15	15	20020917X02117	Accident	FTW82FPG08	Probable Cause
16	16	20020917X01962	Accident	DEN82DTM08	Probable Cause
17	17	20020917X01656	Accident	ANC82FAG14	Probable Cause
18	18	20020917X02481	Accident	NYC82DA016	Probable Cause
19	19	20020917X02339	Accident	MIA82DA028	Probable Cause

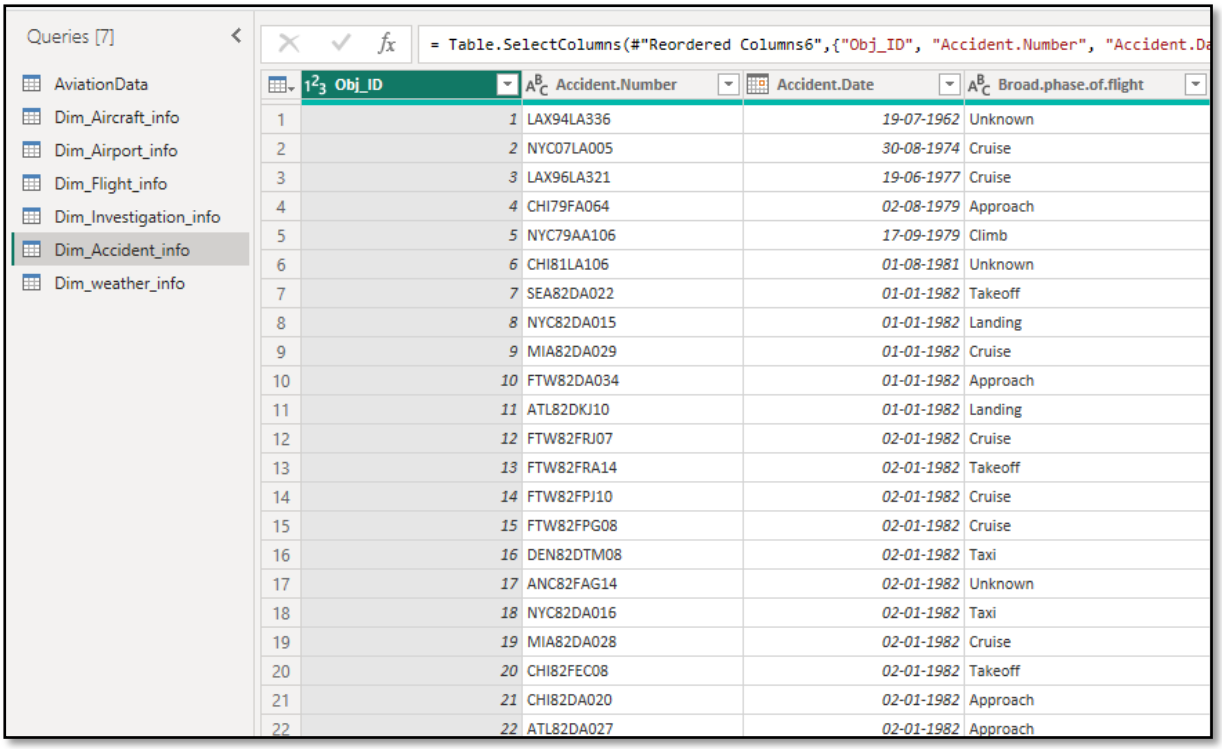
Figure-39: creation of Dimension Dim_Investigation_info Table

= Table.SelectColumns(#"Reordered Columns4",{"Obj_ID", "Event.Id", "Investigation.Type", "Accident.Number", "Report.Status", "Publication.Date"})

Dimension Dim_Accident_info

The “Dim_Accident_info” table contains the following columns-

i. Obj_ID
ii. Accident.Number
iii. Accident.Date
iv. Broad.phase.of.flight
v. Aircraft.damage
vi. State
vii. Country



	Obj_ID	Accident.Number	Accident.Date	Broad.phase.of.flight
1	1	LAX94LA336	19-07-1962	Unknown
2	2	NYC07LA005	30-08-1974	Cruise
3	3	LAX96LA321	19-06-1977	Cruise
4	4	CHI79FA064	02-08-1979	Approach
5	5	NYC79AA106	17-09-1979	Climb
6	6	CHI81LA106	01-08-1981	Unknown
7	7	SEA82DA022	01-01-1982	Takeoff
8	8	NYC82DA015	01-01-1982	Landing
9	9	MIA82DA029	01-01-1982	Cruise
10	10	FTW82DA034	01-01-1982	Approach
11	11	ATL82DKJ10	01-01-1982	Landing
12	12	FTW82FRJ07	02-01-1982	Cruise
13	13	FTW82FRA14	02-01-1982	Takeoff
14	14	FTW82FPJ10	02-01-1982	Cruise
15	15	FTW82FPG08	02-01-1982	Cruise
16	16	DEN82DTM08	02-01-1982	Taxi
17	17	ANC82FAG14	02-01-1982	Unknown
18	18	NYC82DA016	02-01-1982	Taxi
19	19	MIA82DA028	02-01-1982	Cruise
20	20	CHI82FEC08	02-01-1982	Takeoff
21	21	CHI82DA020	02-01-1982	Approach
22	22	ATL82DA027	02-01-1982	Approach

Figure-40: creation of Dimension Dim_Accident_info Table

= Table.SelectColumns(#"Reordered Columns6",{"Obj_ID", "Accident.Number", "Accident.Date", "Broad.phase.of.flight", "Aircraft.damage", "State", "Country"})

Dimension Dim_weather_info

The “Dim_weather_info” table contains the following columns-

i. Obj_ID
ii. Weather.Condition

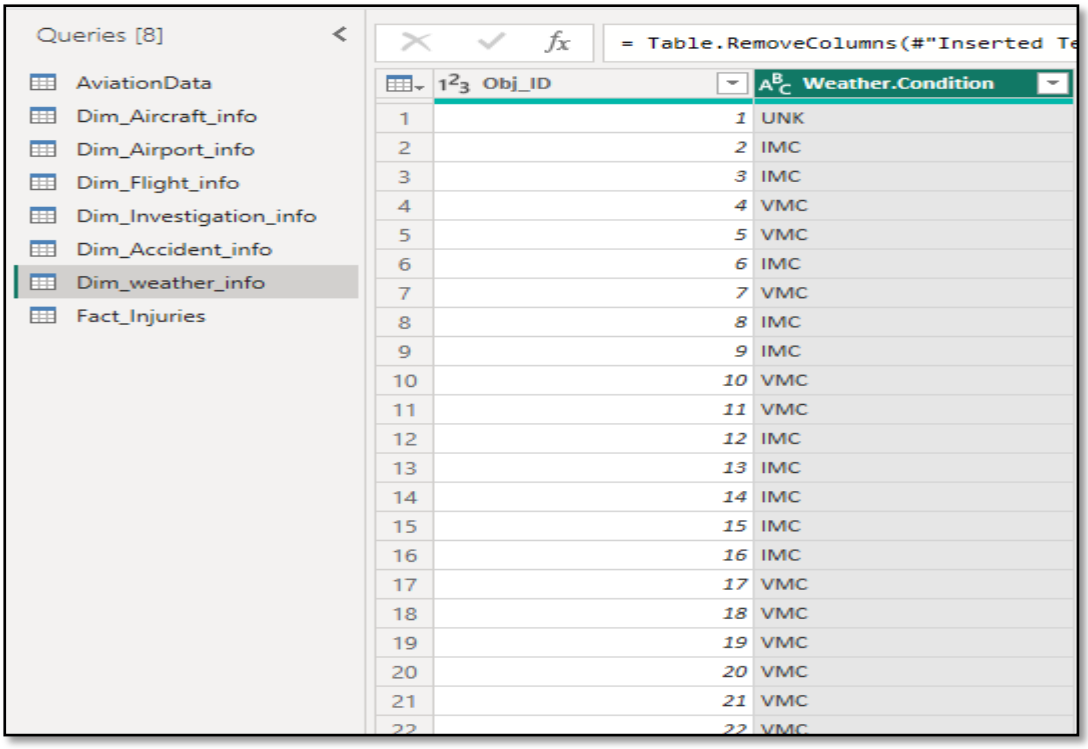


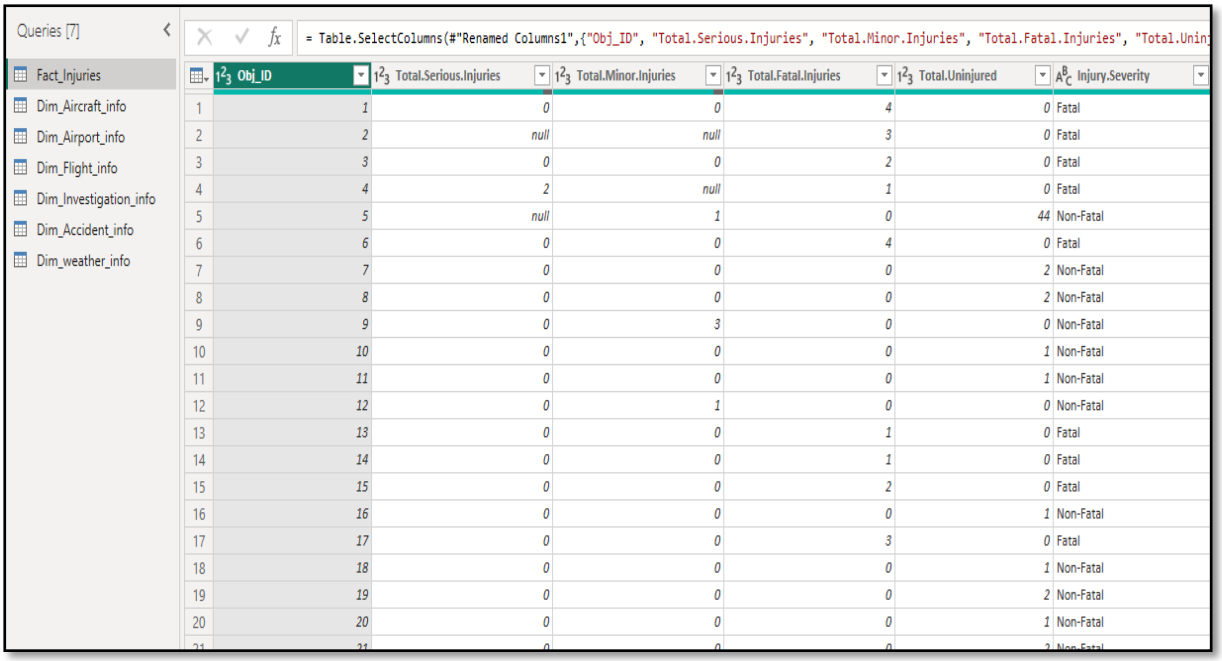
Figure-41: creation of Dimension Dim_weather_info Table

= Table.SelectColumns(#"Reordered Columns2",{"Serial Number", ","Weather.Condition "})

Fact Fact_Injuries

The "Fact_Injuries" table contains the following columns-

i. Obj_ID
ii. Total all type Injured
iii. Injury.Severity
iv. Total.Uninjured
v. Total.Minor.Injuries
vi. Total.Serious.Injuries
vii. Total.Fatal.Injuries



	Obj_ID	Total.Serious.Injuries	Total.Minor.Injuries	Total.Fatal.Injuries	Total.Uninjured	Injury.Severity
1	1	0	0	4	0	Fatal
2	2	null	null	3	0	Fatal
3	3	0	0	2	0	Fatal
4	4	2	null	1	0	Fatal
5	5	null	1	0	44	Non-Fatal
6	6	0	0	4	0	Fatal
7	7	0	0	0	2	Non-Fatal
8	8	0	0	0	2	Non-Fatal
9	9	0	3	0	0	Non-Fatal
10	10	0	0	0	1	Non-Fatal
11	11	0	0	0	1	Non-Fatal
12	12	0	1	0	0	Non-Fatal
13	13	0	0	1	0	Fatal
14	14	0	0	1	0	Fatal
15	15	0	0	2	0	Fatal
16	16	0	0	0	1	Non-Fatal
17	17	0	0	3	0	Fatal
18	18	0	0	0	1	Non-Fatal
19	19	0	0	0	2	Non-Fatal
20	20	0	0	0	1	Non-Fatal

Figure-42: creation of Fact Fact_Injuries Table

= Table.SelectColumns(#"Renamed Columns1",{"Obj_ID", "Total.Serious.Injuries", "Total.Minor.Injuries", "Total.Fatal.Injuries", "Total.Uninjured", "Injury.Severity"})

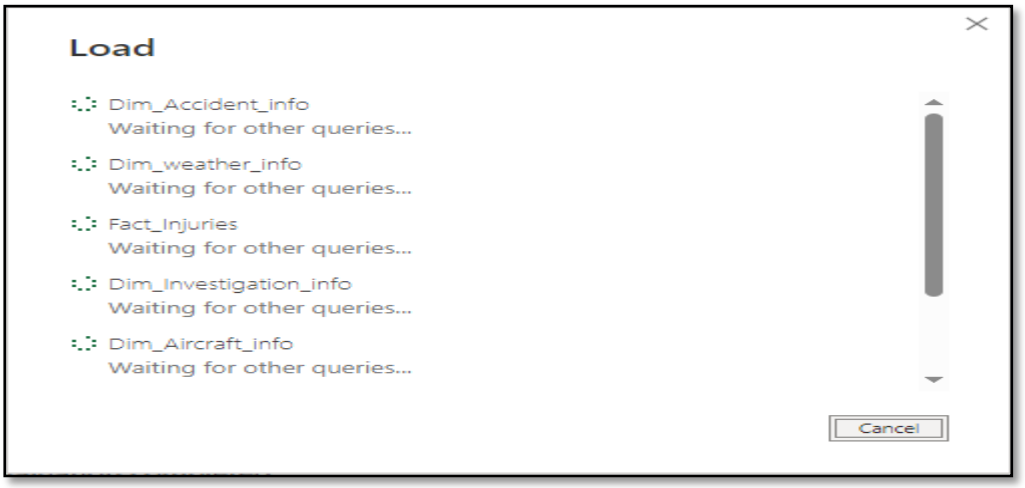


Figure-43: Loading of Dimension and Fact Table

List of Fact and Dimensions Tables

- The list of all fact and dimension tables can be seen in right-top side of the power-bi screen.
- Here in the following image, the list of fact and dimension can be seen.

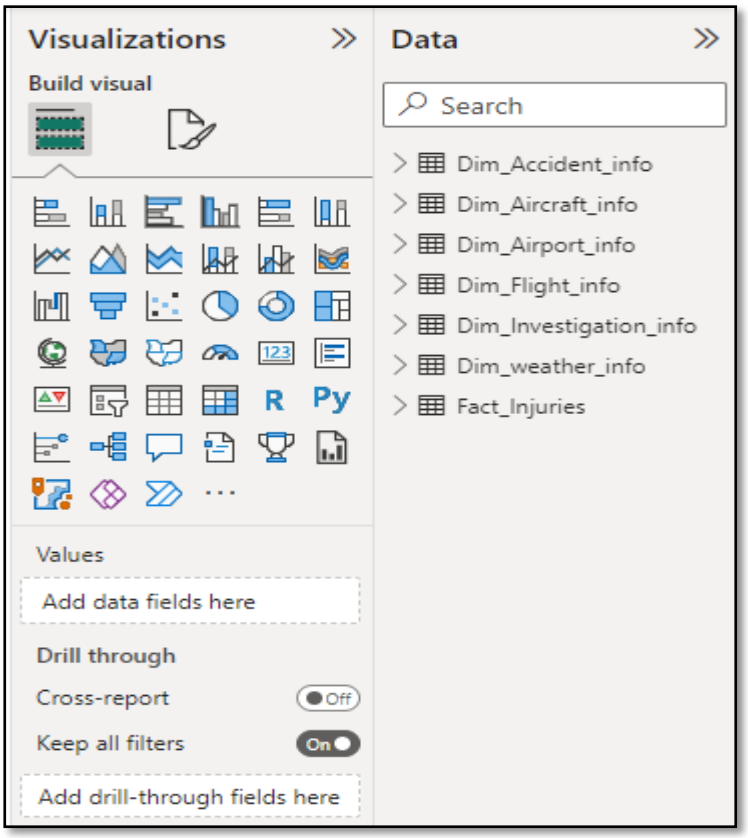


Figure-44: List of fact and dimensions in power-Bi

Data models and relationships

A star schema is a type of data model that resembles a star, with a central fact table surrounded by dimension tables. The fact table contains the quantitative data, such as sales figures, website visits, or product inventory. Each dimension table provides descriptive information about a specific aspect of the data, such as customers, products, time-periods, or locations

Star Schema [one-to-one relationship]

- The Star Schema is a popular data modeling technique used in the design of data.
- It organizes data into a central fact table surrounded by dimension tables.
- The central concept resembles a star, with the fact table at the center and dimension tables radiating out like points on the star.
- This design simplifies querying and reporting, making it easier for analysts to retrieve and analyze information.

Benefits of the Star Schema [one-to-one relationship]

- Simplicity: The structure is straightforward, making it easy to understand and use.
- Query Performance: Queries involving the fact table and one or more dimension tables can be efficient, as the relationships are well-defined.
- Flexibility: It allows for easy expansion by adding new dimensions or modifying existing ones without affecting the entire schema.

Manage Relationship

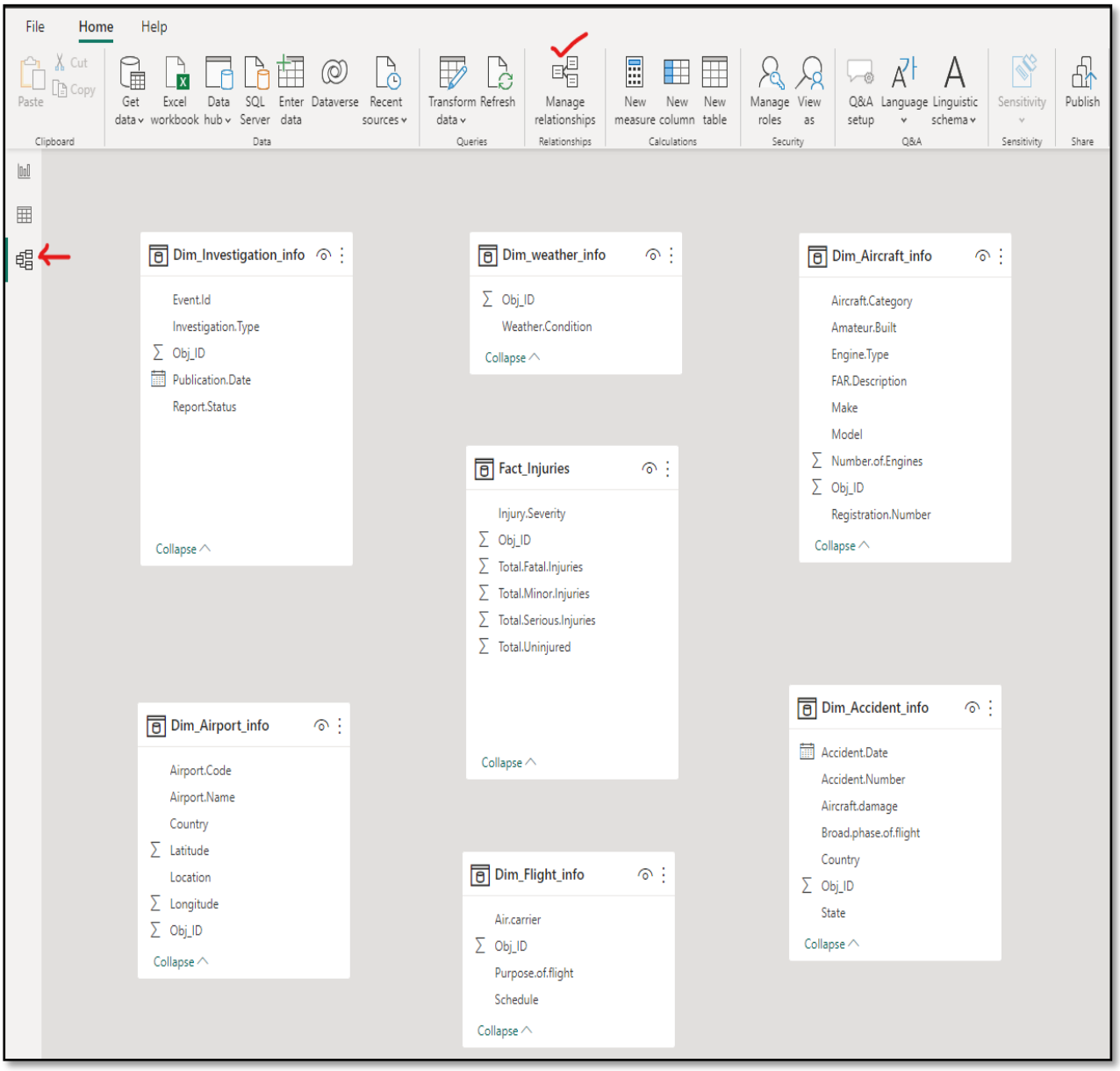


Figure-45: Manage Relationship Option

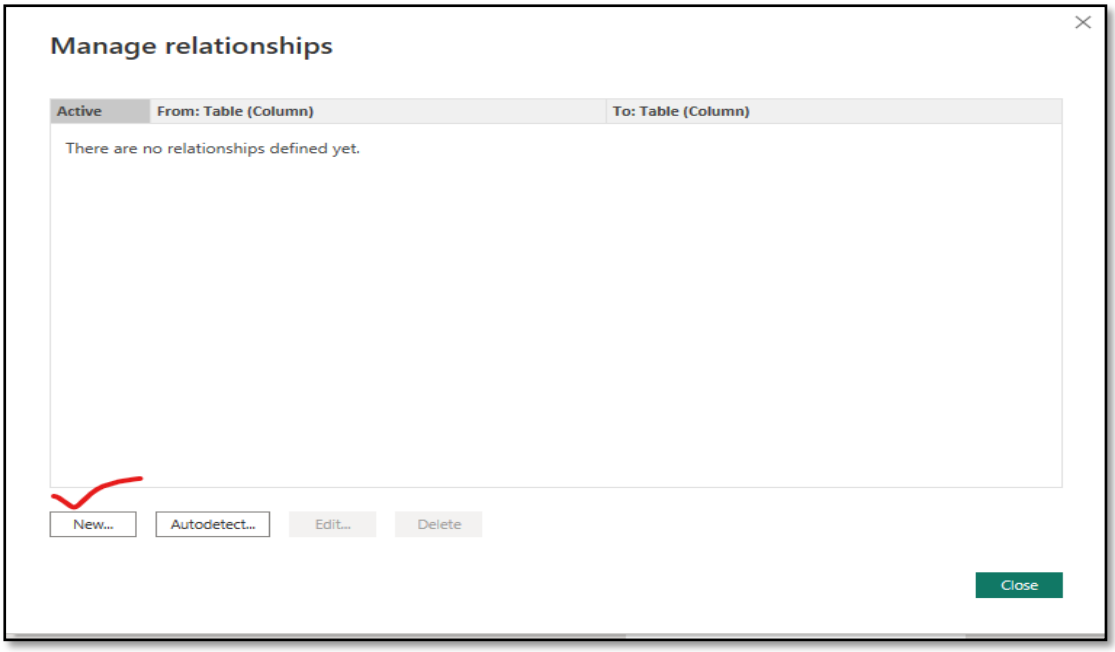


Figure-46: Manage Relationship pop up

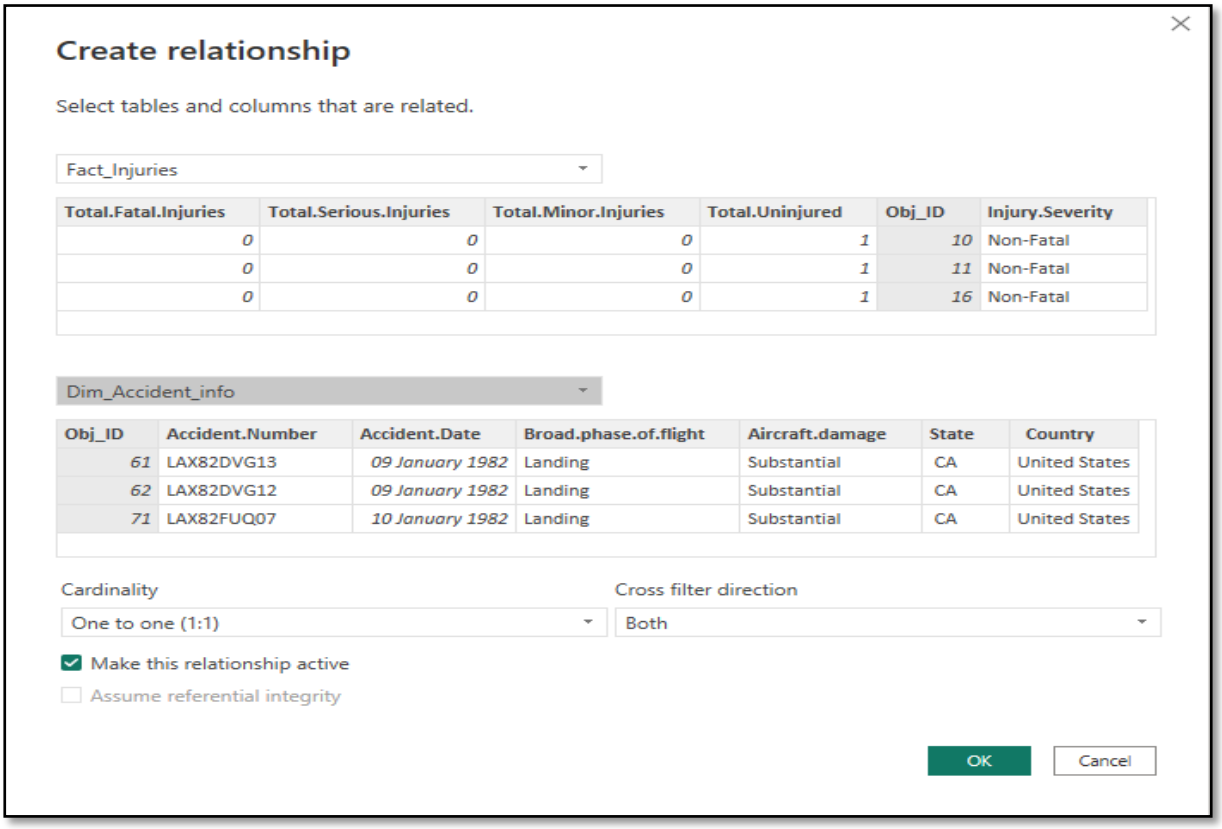


Figure-47: Create Relationship b/w Fact_Injuries & Dim_Accident_info

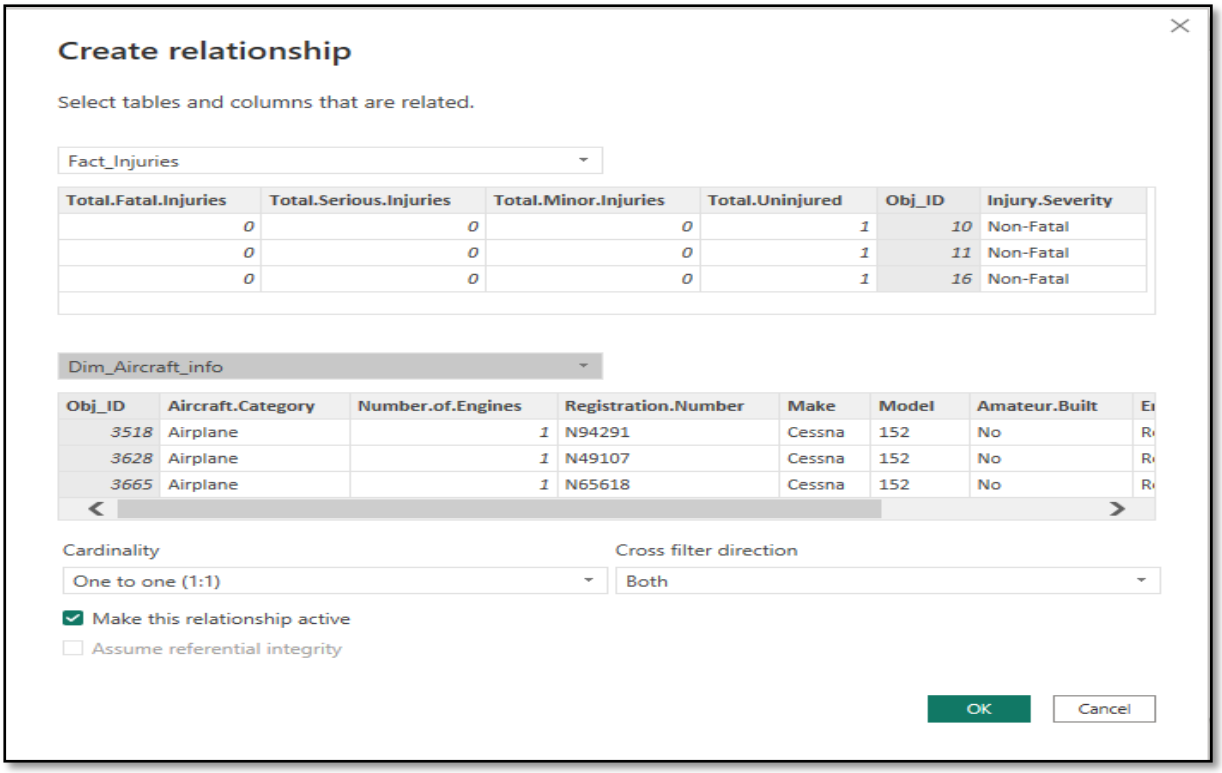


Figure-48: Create Relationship b/w Fact_Injuries & Dim_Aircraft_info

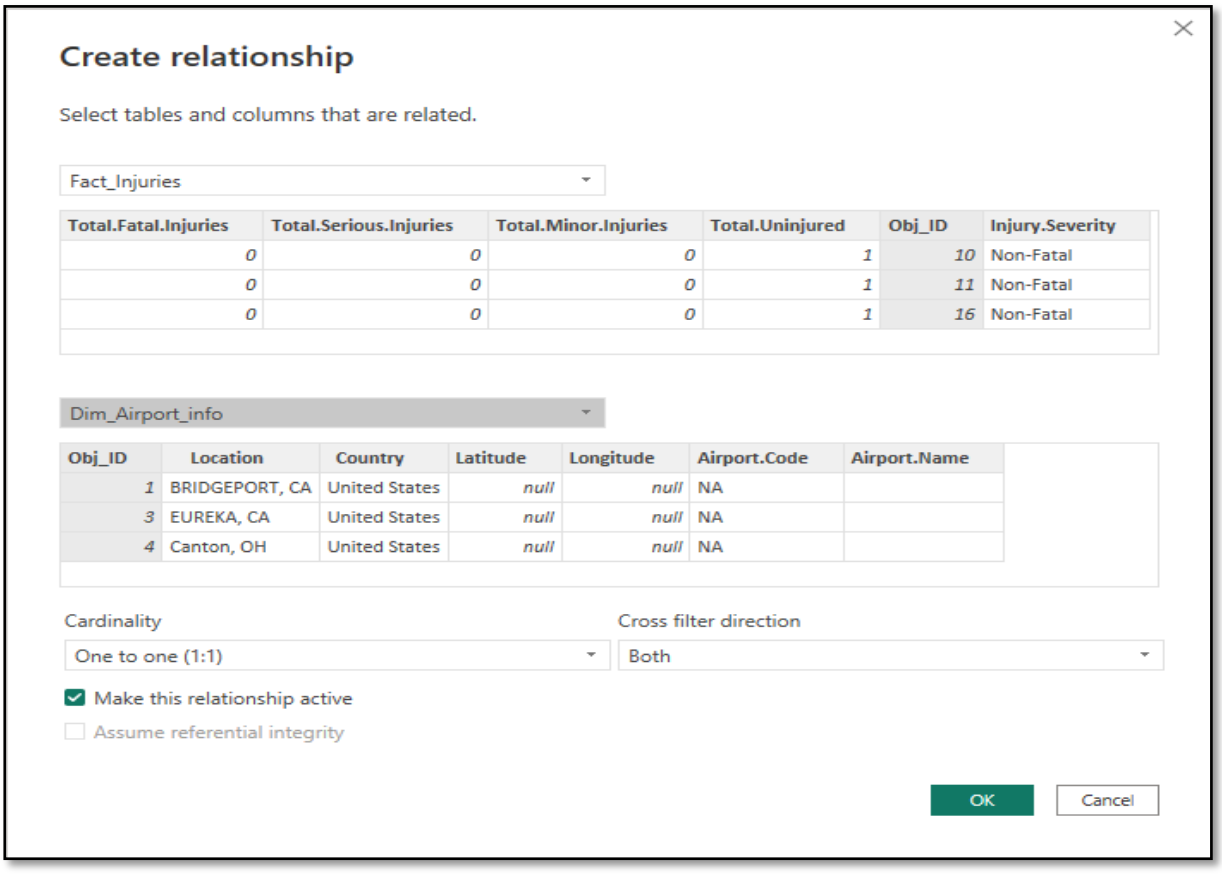


Figure-49: Create Relationship b/w Fact_Injuries & Dim_Airport_info

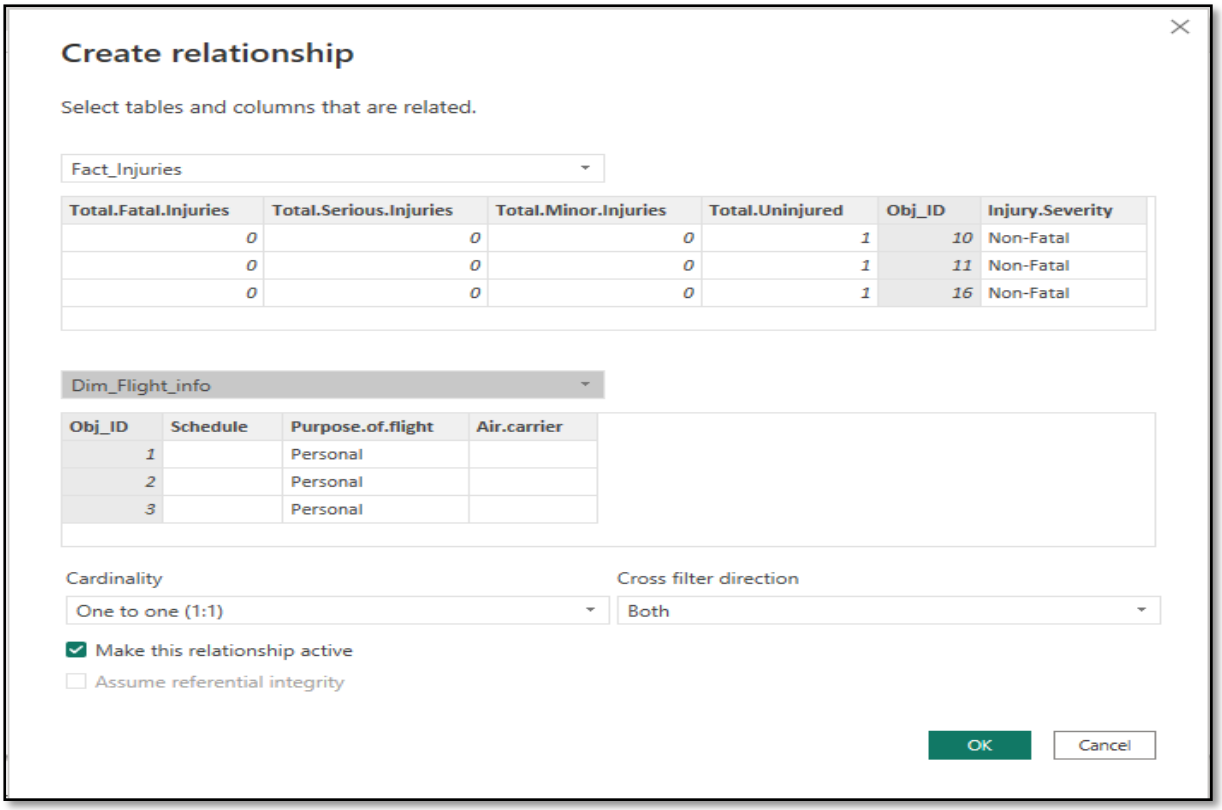


Figure-50: Create Relationship b/w Fact_Injuries & "Dim_Flight_info

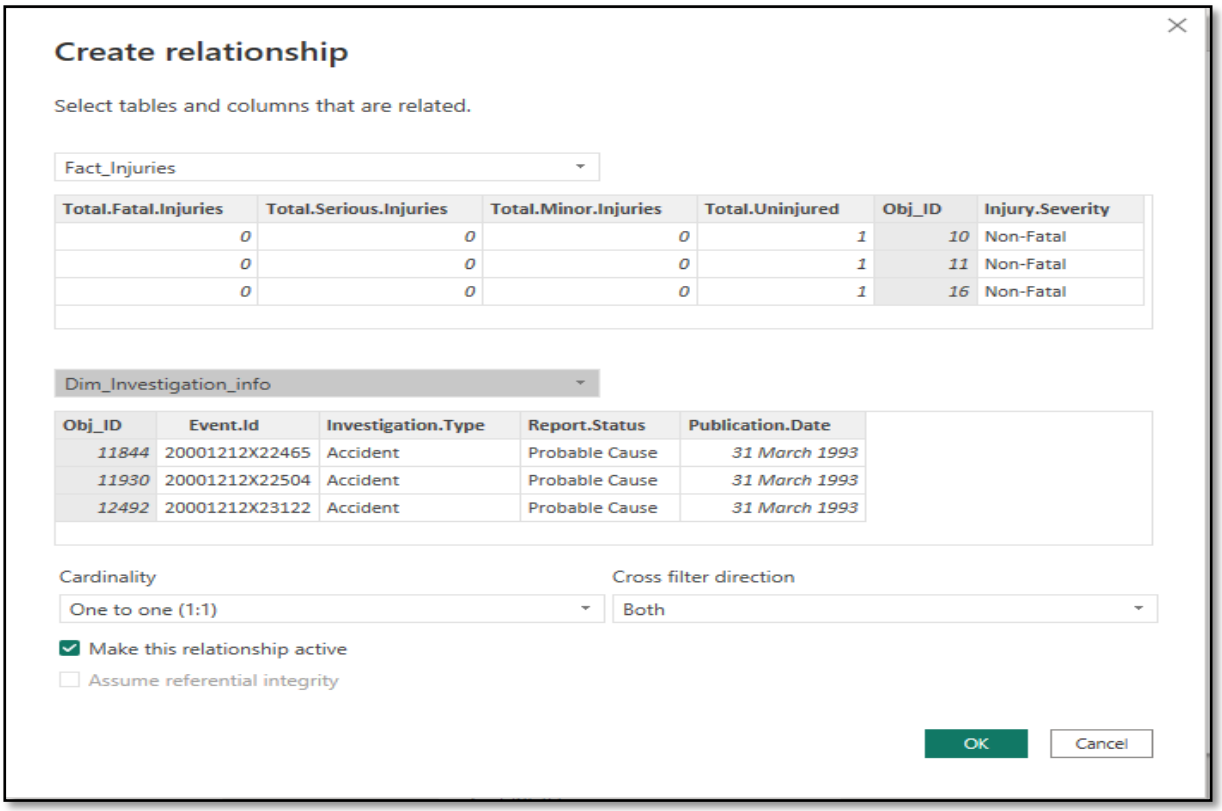


Figure-51: Create Relationship b/w Fact_Injuries & Dim_Investigation_info

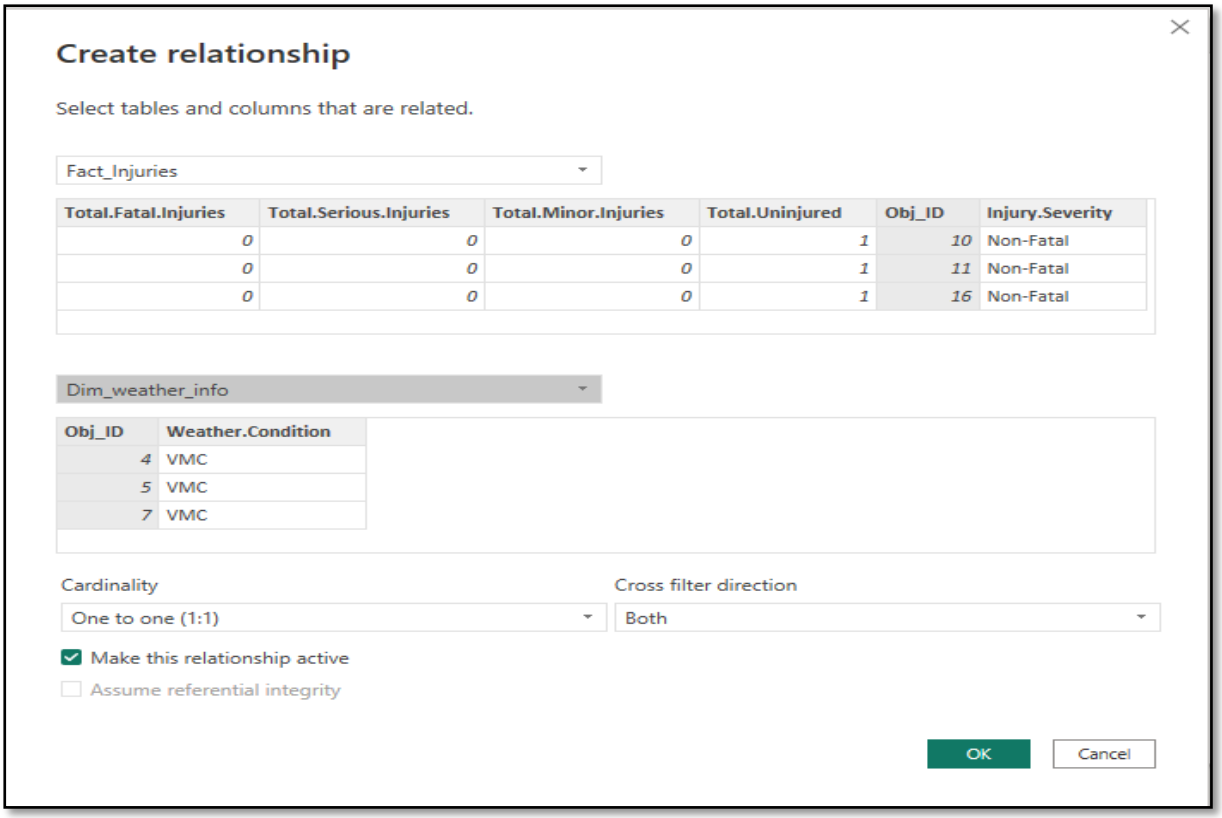


Figure-52: Create Relationship b/w Fact_Injuries & Dim_weather_info

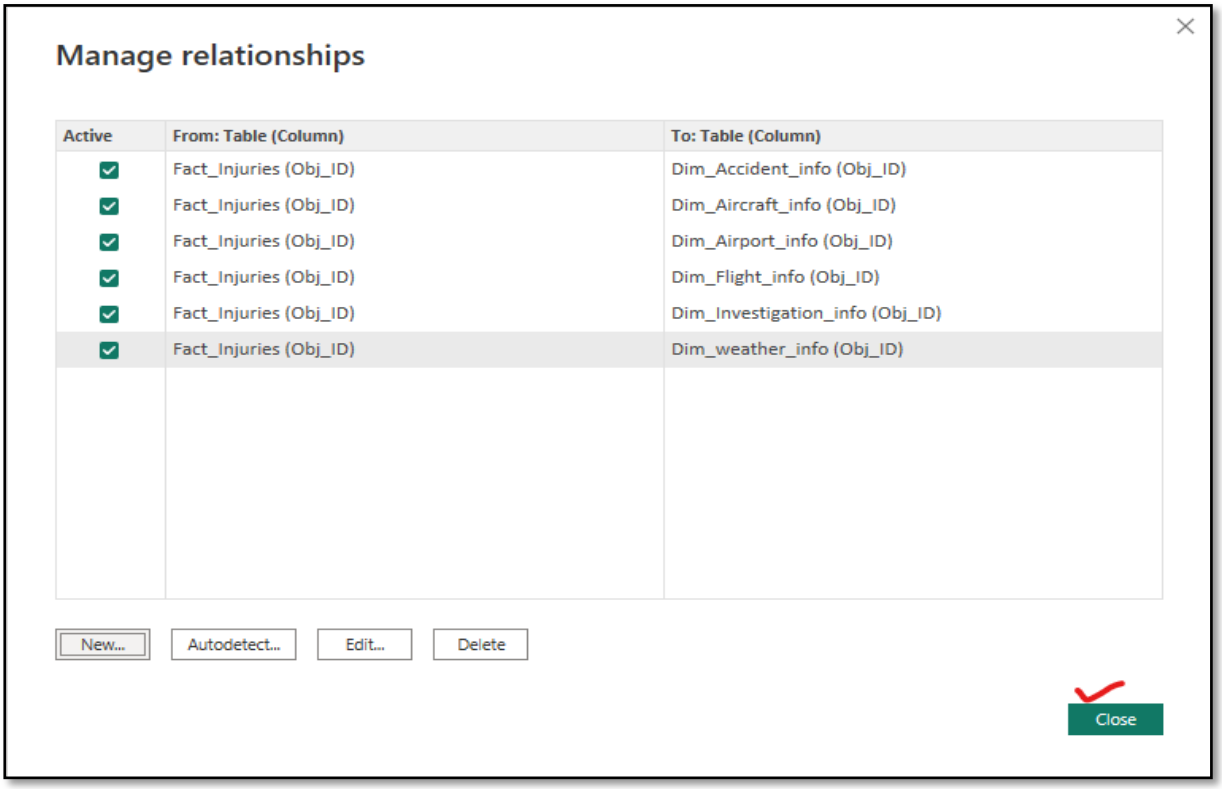


Figure-53: List of all created relationships

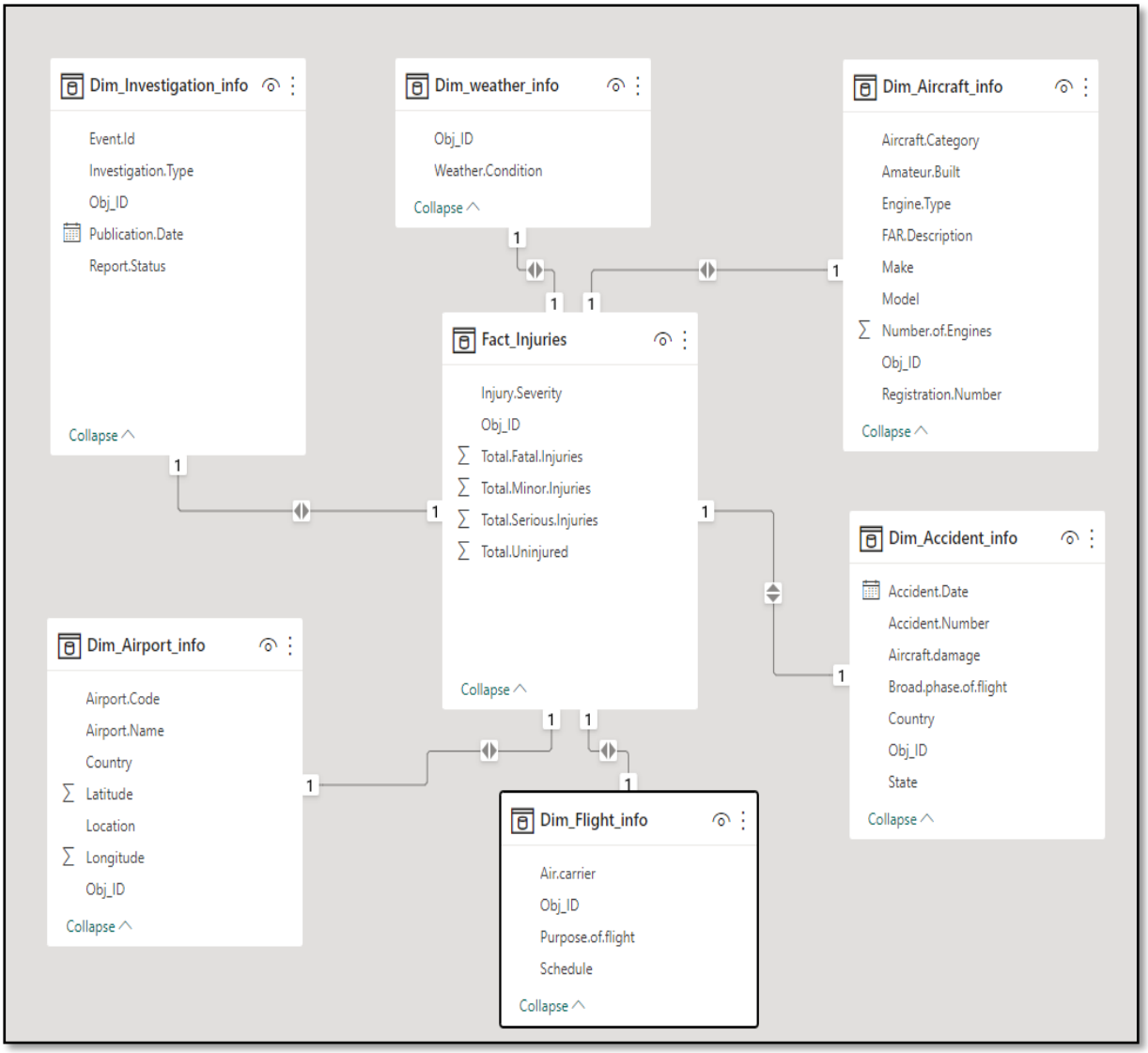


Figure-54: Star Schema of Relationships

Data Visualization

Figure-55: Home page of Dashboard

The buttons, which are used on the home page, are:

i. Home
ii. Country
iii. Aircraft
iv. Injuries and Death
v. Timely Accident

These buttons can be used to navigate on the dashboard. For jumping on the required page, use “ctrl + click”.