# Project Report

## ***Exploring insights from Airline data analysis with Qlik***

Submitted by: Chinmmay D. Dudhediya

## 1. Introduction:

### 1.1 Overview:

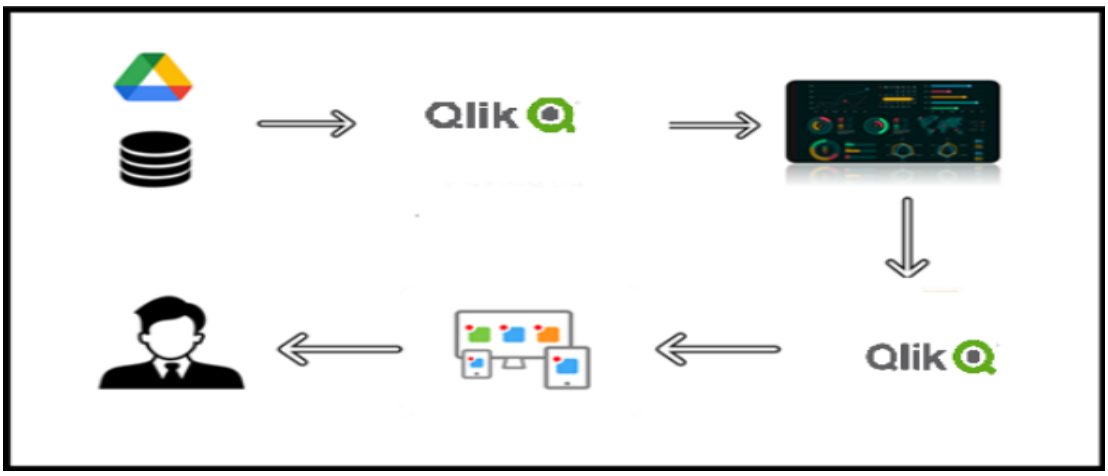
- The project "Exploring Insights from Synthetic Airline Data Analysis with Qlik" involves utilizing synthetic airline data to derive valuable insights using Qlik, a business intelligence and data visualization tool.
- In this project, the synthetic airline data simulates various aspects of airline operations, including flight schedules, passenger demographics, ticket sales, and performance metrics.

### 1.2 Purpose:

- The objective is to leverage Qlik's analytical capabilities to uncover patterns, trends, and correlations within this data, aiding in decision-making processes for airlines, airports, and related stakeholders.

1.3 Technical Architecture:

- **Collect the dataset**
- **Connect Data with Qlik Sense**
- **Prepare the Data for Visualization**
- **Visualizations**
- **Responsive and Design of Dashboard**
- **Story Creation**



# 2. Define Problem / Problem Understanding:

## 2.1 Specify the business problem:

- With narrow operating margins and intense competition, the aviation industry is exceedingly competitive. Among them include the airlines airports and their related service providers who must optimize their operations to improve customer satisfaction and engage in data driven decision making. However, currently, there is a problem of being unable to examine large amounts of information from varied sources for valuable insights.

Some of these problems include:

- Operational inefficiency: This includes challenges with flight delays or cancellations as well as resource allocation at airports and airlines.
- Customer experience: It is critical but difficult to understand and improve passenger satisfaction due to disparate data sources.
- Revenue management: Proper pricing strategies need to be optimized by airlines as well as effective revenue streams control.
- Maintenance and safety: Predictive maintenance requires analyzing complex data sets for avoiding any kind of issues before they arise.
- Market competitiveness: Continuous monitoring and analysis is required to keep pace with market trends and rival strategies.

## 2.2 Business requirements:

Functional Requirements:

- Data Integration: The software should be able to link data from different sources, such as flight schedules, customer opinions, weather information and financial records.
- Real-Time Analysis: The system must allow for analysis of data that is taking place now for making quick decisions.

- Interactive Dashboards: Provide easy-to-use dashboards through which stakeholders would perceive and interact with the data.

- Predictive Analytics: These are tools that forecast future developments and potential problems based on past patterns.
- Pattern and Trend Analysis: Find out what the numbers say about trends, correlations and patterns so you can act accordingly.
- Custom Reporting: Reports have to be tailored to meet specific needs of various players in the airline industry.

Non-Functional Requirements:

- Scalability: Massive amount of data should be accommodated by a system while it grows with increasing loads of data.
- Security: Ensuring privacy of information regardless of its sensitivity when it comes to passenger details along with operations.
- Performance: Quick processing time with real-time analytics as well as reports.
- User Training: Availability of training materials and support for enhancing use of these analytical tools by users.
- Integration Capability: Be seamlessly incorporated into the systems being used by airlines and airports.

## 2.3 Literature Survey:

Qlik's Role in Aviation Data Analysis:

- The software is important for analyzing complicated aviation data such as the ones found in airlines.
- According to research, it allows easier exploration and discovery of relationships among data than traditional BI tools .

Case Studies and Industry Applications:

- Airlines: it has been shown that Qlik is used by major airlines in optimizing flight schedules, improving revenue management, and on-time performance with the use of predictive analytics .
- Airports: Airports have utilized Qlik as a means of managing passenger flow, which enhances security screening procedures and improves operational

efficiency .

Predictive Maintenance:

- This study also indicated that these tools can be used to reduce downtime and costs through equipment failure prediction hence enabling timely maintenance scheduling using Qlik's tools .

Customer Satisfaction and Experience:

- Use of this approach has also been studied in relation to customer feedback analysis expressed through Qlik satisfaction surveys. This helps airlines adjust their services more effectively towards meeting customer requirements during travel.

Revenue and Pricing Optimization:

- In particular, literature stresses the use of advanced analytics related to dynamic pricing and revenue management. In real time data processing abilities including predictive modeling capabilities.

Operational Efficiency:

- Studies show that airports and airlines use Qlik's dashboards and real-time analytics to allow them monitor the performance of their operational activities, so as to cut costs and make their operations more efficient.

References:


- [1] "The Power of Associative Analytics in the Airline Industry," Journal of Aviation Management.
- [2] "Data-Driven Decision Making in Aviation: A Qlik Perspective," Aviation Analytics Journal.
- [3] "Case Study: Optimizing Flight Schedules with Qlik," Airline Business Review.
- [4] "Improving Airport Operations with Qlik Analytics," Airport Management Today.
- [5] "Predictive Maintenance in Aviation: Leveraging Qlik," Maintenance & Engineering.
- [6] "Enhancing Passenger Experience through Data Analytics," Customer Experience Journal.
- [7] "Revenue Management Strategies in Airlines using Qlik," Revenue Optimization Quarterly.
- [8] "Operational Efficiency in Airports: The Role of Qlik," Operations Research in Aviation.

# 3. Data Collection:

3.1 Collect the dataset and understanding:

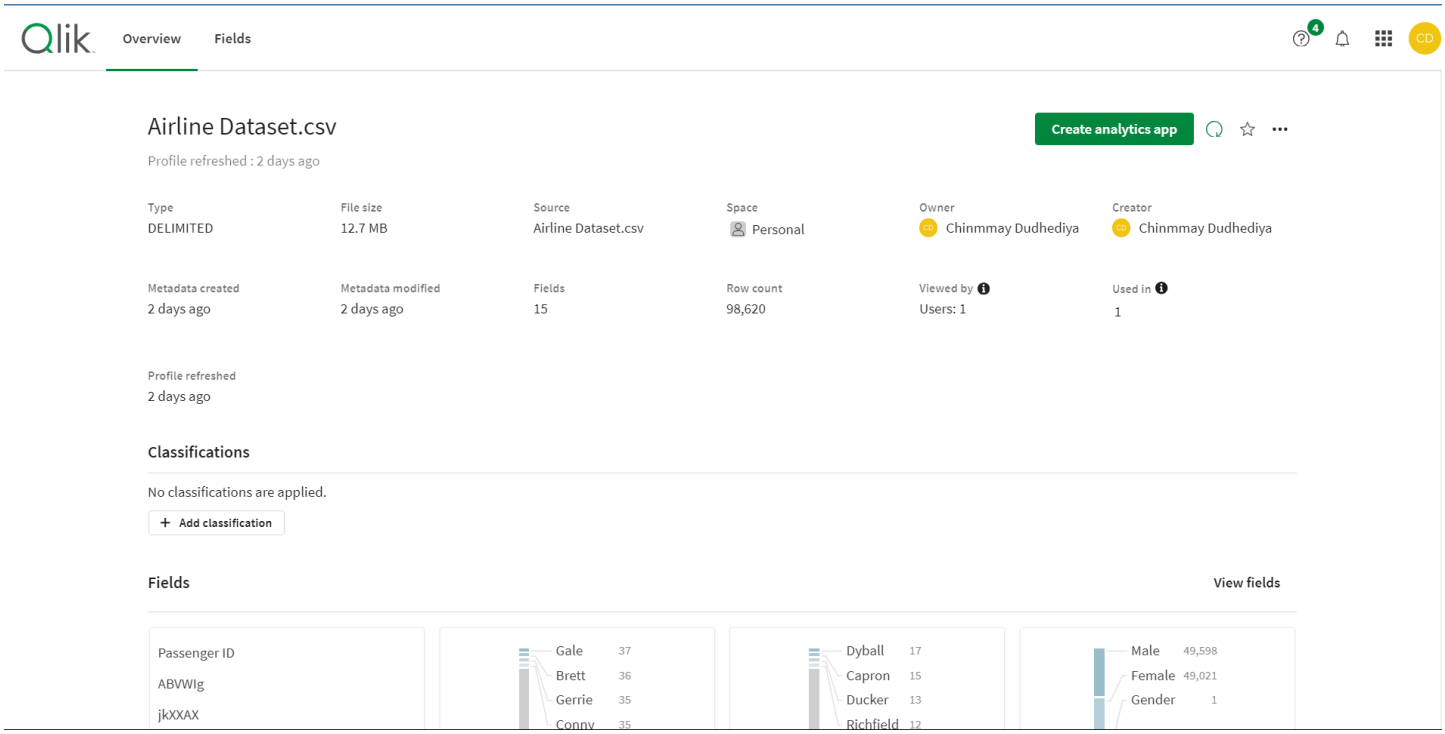
- Collecting a dataset, here from "https://www.kaggle.com/datasets/iamsouravbanerjee/airline-dataset/data "
- Data contains all the meta information regarding the columns described in the CSV files

Column Description of the Dataset:

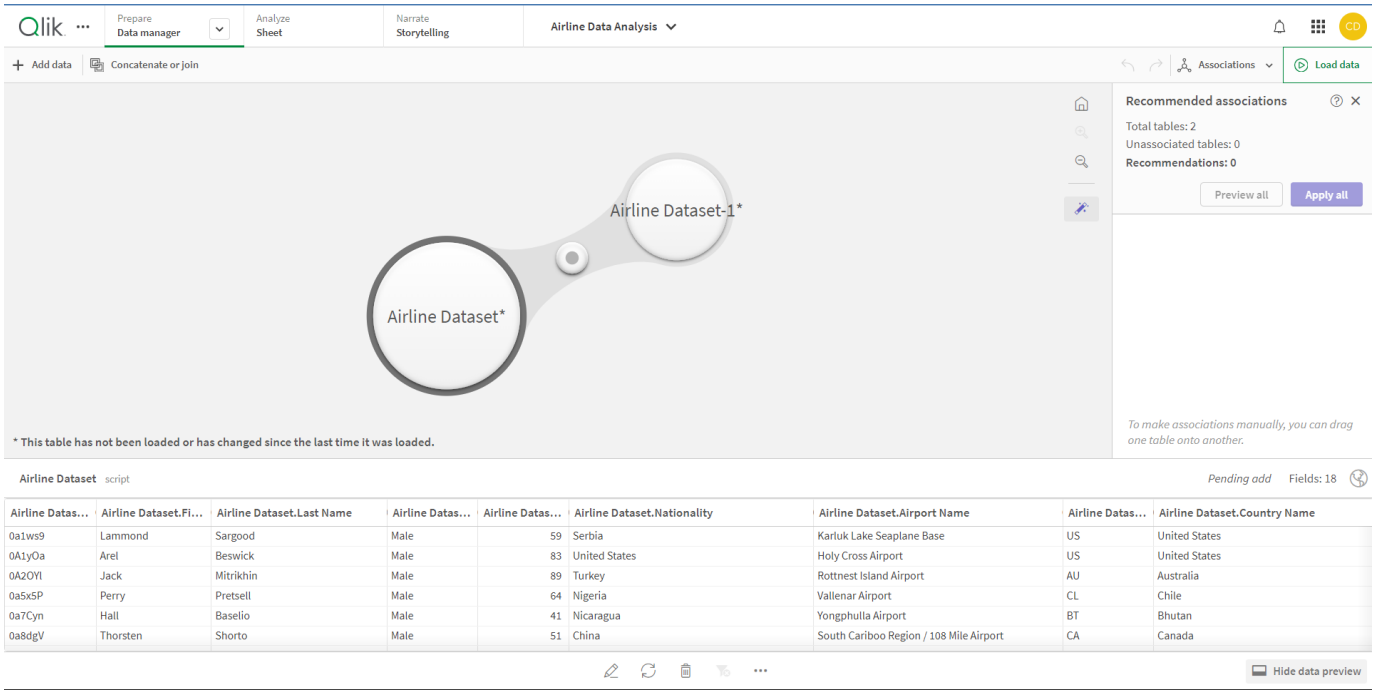
- Passenger ID - Unique identifier for each passenger
- First Name - First name of the passenger
- Last Name - Last name of the passenger
- Gender - Gender of the passenger
- Age - Age of the passenger
- Nationality - Nationality of the passenger
- Airport Name - Name of the airport where the passenger boarded
- Airport Country Code - Country code of the airport's location
- Country Name - Name of the country the airport is located in
- Airport Continent - Continent where the airport is situated
- Continents - Continents involved in the flight route
- Departure Date - Date when the flight departed
- Arrival Airport - Destination airport of the flight
- Pilot Name - Name of the pilot operating the flight
- Flight Status - Current status of the flight (e.g., on-time, delayed, canceled)

## 3.2 Connect Data with Qlik Sense:

- Create a new analytics app and then load the dataset to connect with the qlik.




- This shows the collected and loaded " .csv " file




- This shows the dataset connected with the Qlik Sense

# 4. Data Preparation:

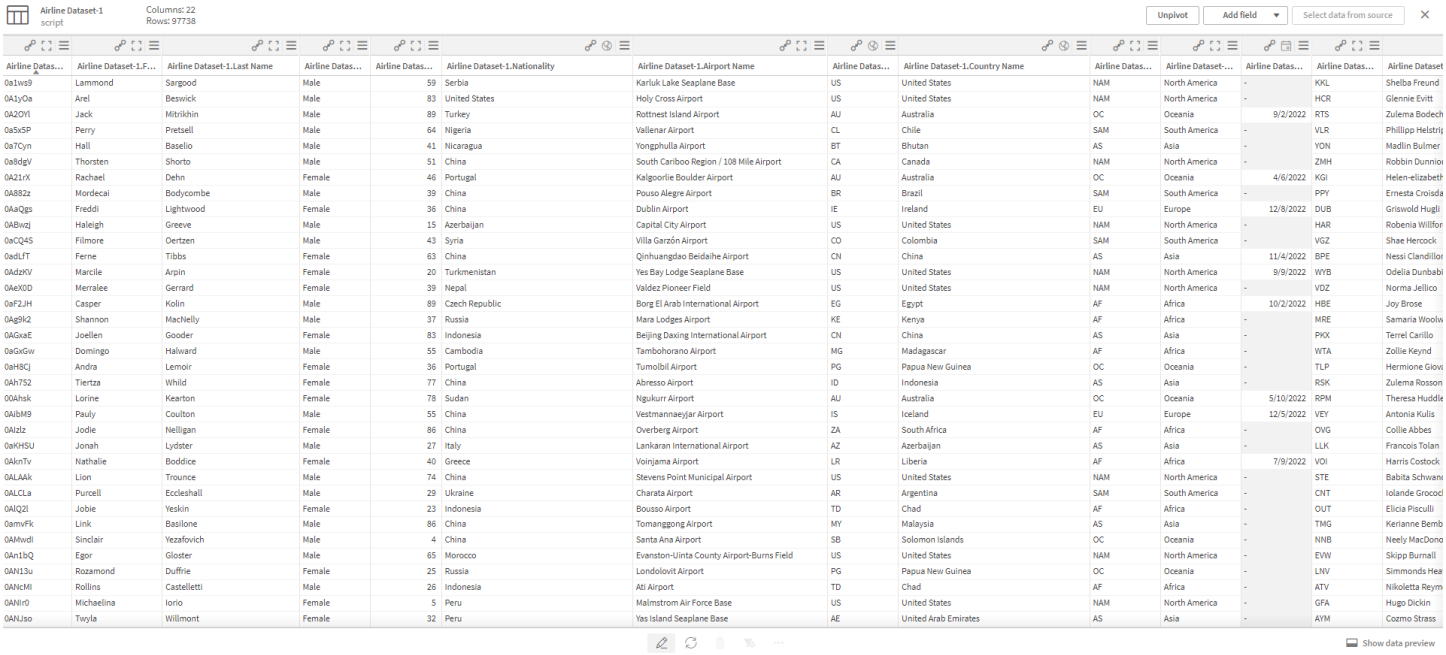
## 4.1 Prepare the Data for Visualization:

- Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for visualization software, and ensuring the data is accurate and complete.

- This process helps to make the data easily understandable and ready for creating visualizations to gain insights into performance and efficiency. Since the data is already cleaned, we can move to visualization.

In our project we used:

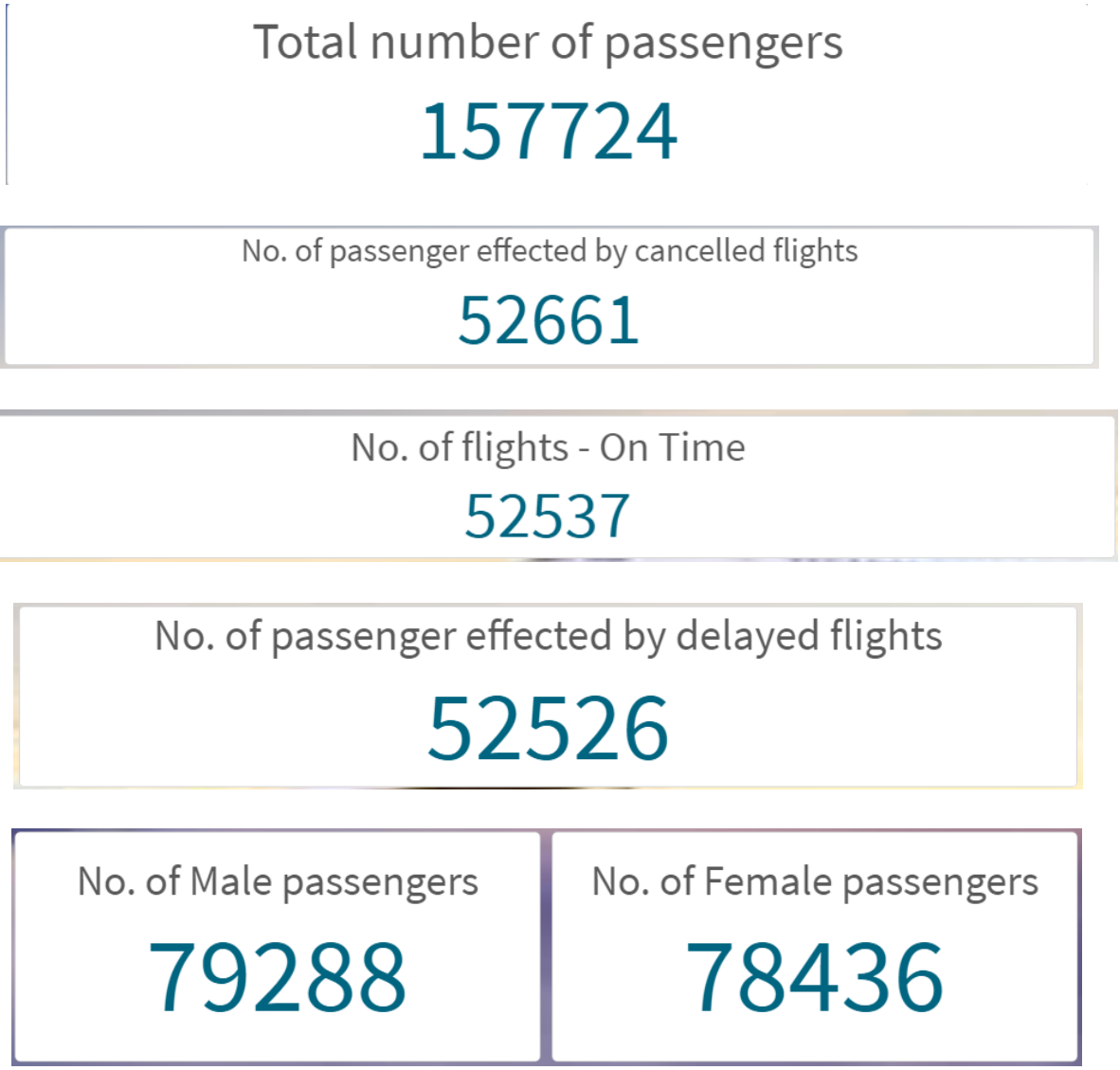
```
[Airline Dataset]:
Load *;
//Remove rows with 'B' and from ArrivalAirport column
[Airline Dataset]:
NOCONCATENATE LOAD *,
if(Age >=0 AND Age <= 1, 'Baby',
if(Age >= 1 AND Age <= 3, 'Toddler',
if(Age >= 4 AND Age <= 9, 'Child',
if(Age >=10 AND Age <= 12, 'Tween',
if(Age >=13 AND Age <= 19, 'Teen',
if(Age >=20 AND Age <=24, 'Young Adult',
if (Age >=25 AND Age <=39, 'Adult',
if (Age >=40 AND Age <= 54, 'Middle',
if(Age >=55 AND Age <= 79, 'Elder',
if(Age >=80, 'Old'))))))))) AS AgeGroup,
Date#([Departure Date], 'MM/DD/YYYY') as [Departure_Date],
Year([Departure Date]) AS Year,
Month([Departure Date]) as Month
RESIDENT [Airline Dataset]
Where NOT ([Arrival Airport] = '0' OR [Arrival Airport] = '-');
```

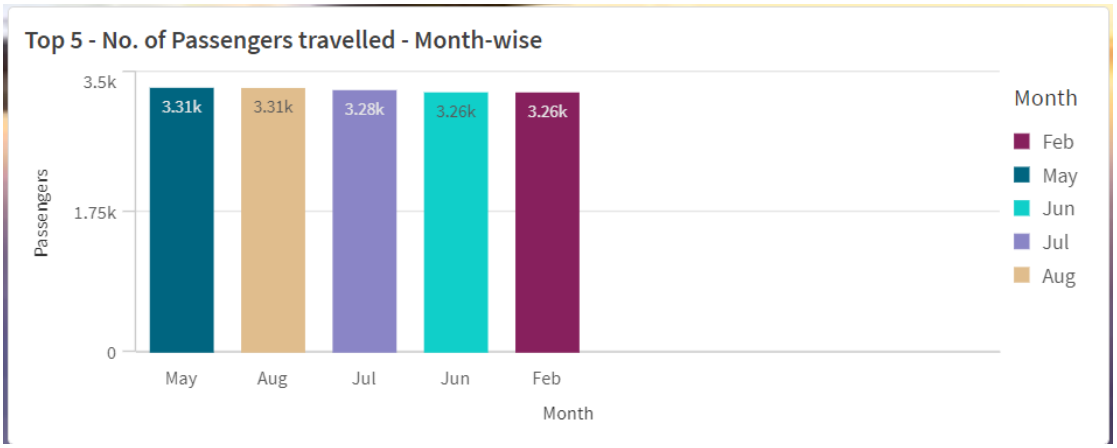

Airline Dataset-1
script
Columns: 22
Rows: 97738
Unpivot
Add field
Select data from source
Airline Datas...
Airline Dataset-1.F...
Airline Dataset-1.Last Name
Airline Datas...
Airline Datas...
Airline Dataset-1.Nationality
Airline Dataset-1.Airport Name
Airline Datas...
Airline Dataset-1.Country Name
Airline Datas...
Airline Dataset-...
Airline Datas...
Airline Datas...
Airline Dataset
0a1ws9 Lammond Sargood Male 59 Serbia Karluk Lake Seaplane Base US United States NAM North America - KKL Shelba Freund
0A1yOa Arel Beswick Male 83 United States Holy Cross Airport US United States NAM North America - HCR Glennie Evitt
0A2OYl Jack Mitrikhin Male 89 Turkey Rottnest Island Airport AU Australia OC Oceania 9/2/2022 RTS Zulema Bodech
0aSx5P Perry Pretsell Male 64 Nigeria Vallenar Airport CL Chile SAM South America - VLR Phillipp Helstrij
0a7Cyn Hall Baselio Male 41 Nicaragua Yongphulla Airport BT Bhutan AS Asia - YON Madlin Bulmer
0a8dgV Thorsten Shorto Male 51 China South Cariboo Region / 108 Mile Airport CA Canada NAM North America - ZMH Robbin Dunnio
0A21rX Rachael Dehn Female 46 Portugal Kalgoorlie Boulder Airport AU Australia OC Oceania 4/6/2022 KGI Helen-elizabeth
0A882z Mordecai Bodycombe Male 39 China Pouso Alegre Airport BR Brazil SAM South America - PPY Ernesta Croisda
0AaQgs Freddi Lightwood Female 36 China Dublin Airport IE Ireland EU Europe 12/8/2022 DUB Griswold Hugli
0ABwzj Haleigh Greeve Male 15 Azerbaijan Capital City Airport US United States NAM North America - HAR Robenia Willfor
0aCQ4S Filmore Oertzen Male 43 Syria Villa Garzón Airport CO Colombia SAM South America - VGZ Shae Hercock
0adLfT Ferne Tibbs Female 63 China Qinhuangdao Beidaihe Airport CN China AS Asia 11/4/2022 BPE Nessi Clandillor
0AdzKV Marcile Arpin Female 20 Turkmenistan Yes Bay Lodge Seaplane Base US United States NAM North America 9/9/2022 WYB Odelia Dunbabi
0AeX0D Merralee Gerrard Female 39 Nepal Valdez Pioneer Field US United States NAM North America - VDZ Norma Jellico
0aF2JH Casper Kolin Male 89 Czech Republic Borg El Arab International Airport EG Egypt AF Africa 10/2/2022 HBE Joy Brose
0Ag9k2 Shannon MacNelly Male 37 Russia Mara Lodges Airport KE Kenya AF Africa - MRE Samaria Woolw
0AGxaE Joellen Gooder Female 83 Indonesia Beijing Daxing International Airport CN China AS Asia - PKX Terrel Carillo
0aGxGw Domingo Halward Male 55 Cambodia Tambohorano Airport MG Madagascar AF Africa - WTA Zollie Keynd
0aH8CJ Andra Lemoir Female 36 Portugal Tumolbil Airport PG Papua New Guinea OC Oceania - TLP Hermione Giov
0Ah7S2 Tiertza Whild Female 77 China Abresso Airport ID Indonesia AS Asia - RSK Zulema Rosson
00Ahsk Lorine Kearton Female 78 Sudan Ngukurr Airport AU Australia OC Oceania 5/10/2022 RPM Theresa Huddle
0AibM9 Pauly Coulton Male 55 China Vestmannaeyjar Airport IS Iceland EU Europe 12/5/2022 VEY Antonia Kulis
0AIzIz Jodie Nelligan Female 86 China Overberg Airport ZA South Africa AF Africa - OVG Collie Abbes
0aKHSU Jonah Lydster Male 27 Italy Lankaran International Airport AZ Azerbaijan AS Asia - LLK Francois Tolan
0AknTv Nathalie Boddice Female 40 Greece Voinjama Airport LR Liberia AF Africa 7/9/2022 VOI Harris Costock
0ALAAk Lion Trounce Male 74 China Stevens Point Municipal Airport US United States NAM North America - STE Babita Schwan
0ALCLa Purcell Eccleshall Male 29 Ukraine Charata Airport AR Argentina SAM South America - CNT Iolande Grococ
0AlQ2l Jobie Yeskin Female 23 Indonesia Bousso Airport TD Chad AF Africa - OUT Elicia Pisculli
0amvFk Link Basilone Male 86 China Tomanggong Airport MY Malaysia AS Asia - TMG Kerianne Bemb
0AMwdl Sinclair Yezafovich Male 4 China Santa Ana Airport SB Solomon Islands OC Oceania - NNB Neely MacDono
0An1bQ Egor Gloster Male 65 Morocco Evanston-Uinta County Airport-Burns Field US United States NAM North America - EVW Skipp Burnall
0AN13u Rozamond Duffrie Female 25 Russia Londolovit Airport PG Papua New Guinea OC Oceania - LNV Simmonds Hea
0ANcMl Rollins Castelletti Male 26 Indonesia Ati Airport TD Chad AF Africa - ATV Nikoletta Reym
0ANlr0 Michaelina Iorio Female 5 Peru Malmstrom Air Force Base US United States NAM North America - GFA Hugo Dickin
0ANJso Twyla Willmont Female 32 Peru Yas Island Seaplane Base AE United Arab Emirates AS Asia - AYM Cozmo Strass
Show data preview

# 5. Data Visualizations:

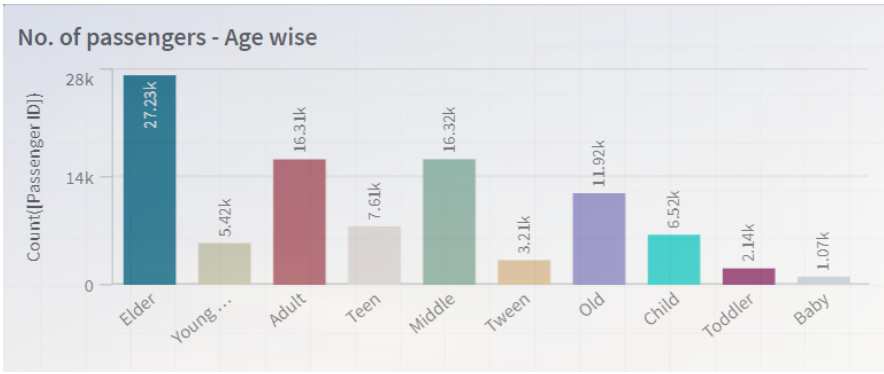
5.1 Visualizations:

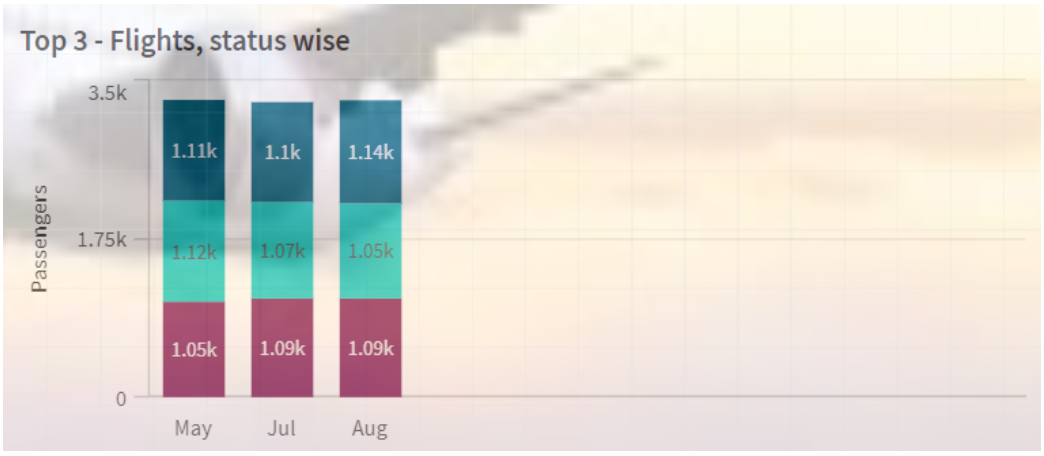
- Data visualization is the process of creating graphical representations of data to help people understand and explore the information.
- The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret.
- By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

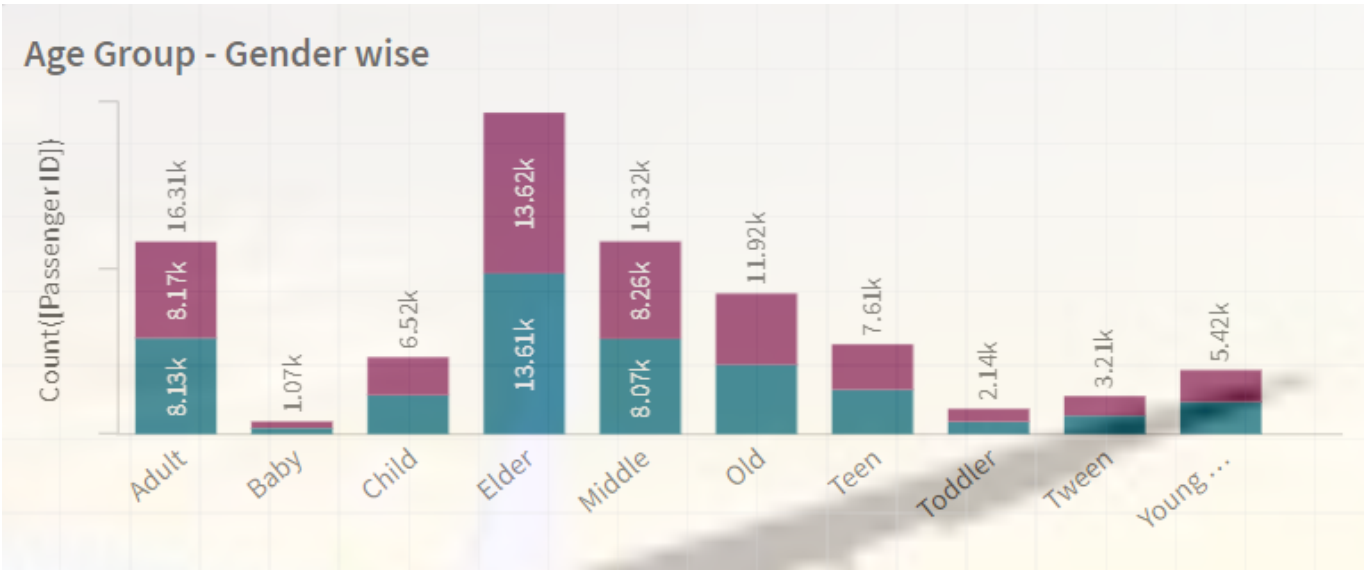


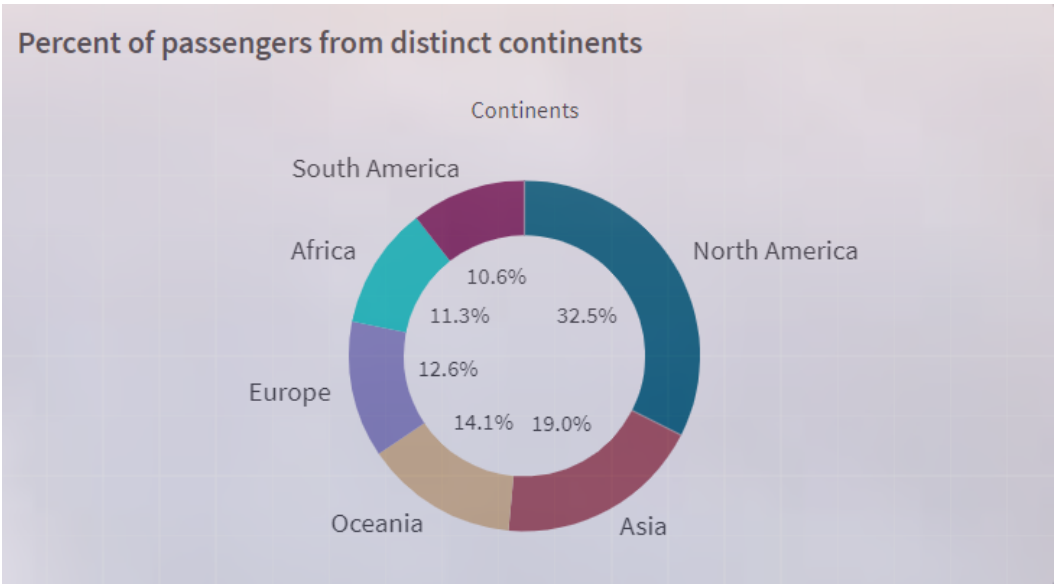

Top 5 - No. of Passengers travelled - Month-wise
3.5k
1.75k
0
Passengers
3.31k
3.31k
3.28k
3.26k
3.26k
May
Aug
Jul
Jun
Feb
Month
Month
Feb
May
Jun
Jul
Aug


Continent wise - Flight Status
North America
Delayed
17.09k
Cancelled
17.07k
On Time
17.05k
Asia
On Time
10.03k
Cancelled
10k
Delayed
9.91k
Oceania
Delayed
7.44k
On Time
7.43k
Cancelled
7.39k
Africa
On Time
5.98k
Cancelled
5.91k
Delayed
5.89k
South Ame...
Cancelled
5.68k
On Time
5.51k
Delayed
5.47k
Europe
Delayed
6.72k
Cancelled
6.61k
On Time
6.54k


No. of passengers - Age wise
28k
14k
0
Count({Passenger ID})
27.23k
5.42k
16.31k
7.61k
16.32k
3.21k
11.92k
6.52k
2.14k
1.07k
Elder
Young...
Adult
Teen
Middle
Tween
Old
Child
Toddler
Baby


Top 3 - Flights, status wise
Passengers
3.5k
1.75k
0
1.11k
1.12k
1.05k
1.1k
1.07k
1.09k
1.14k
1.05k
1.09k
May
Jul
Aug

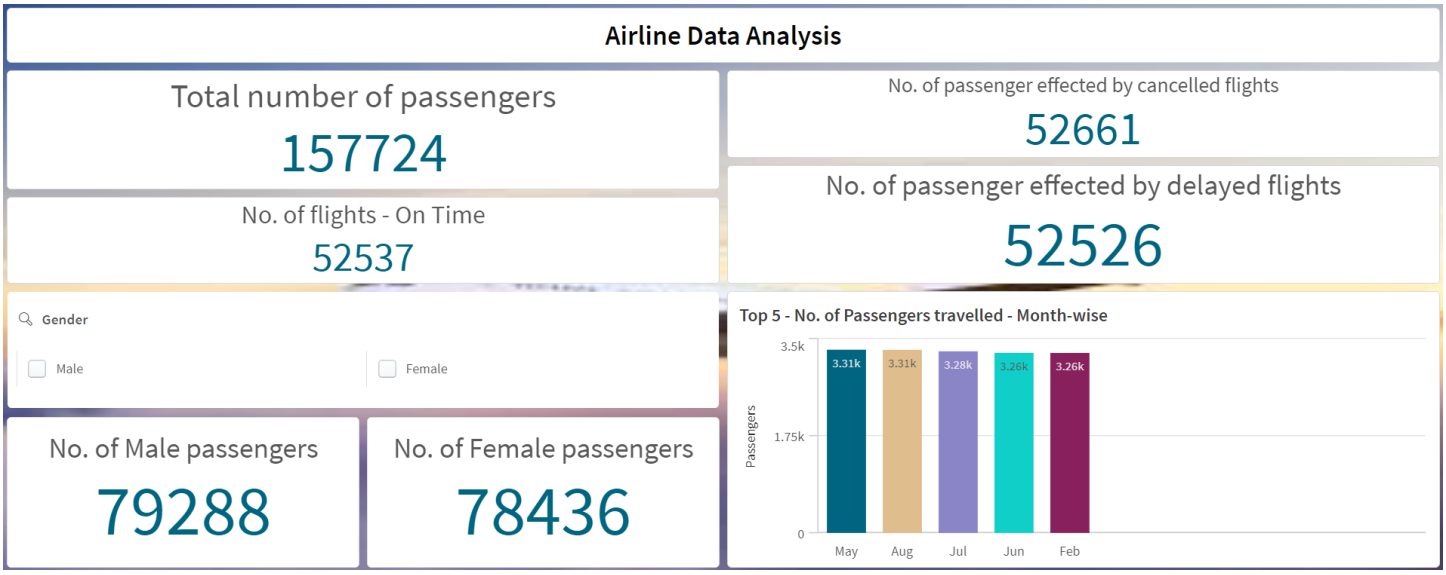

Age Group - Gender wise
Count([Passenger ID])
16.31k
8.17k
8.13k
1.07k
6.52k
13.62k
13.61k
16.32k
8.26k
8.07k
11.92k
7.61k
2.14k
3.21k
5.42k
Adult
Baby
Child
Elder
Middle
Old
Teen
Toddler
Tween
Young…


Percent of passengers from distinct continents
Continents
South America
Africa
North America
10.6%
11.3%
32.5%
12.6%
Europe
14.1%
19.0%
Oceania
Asia

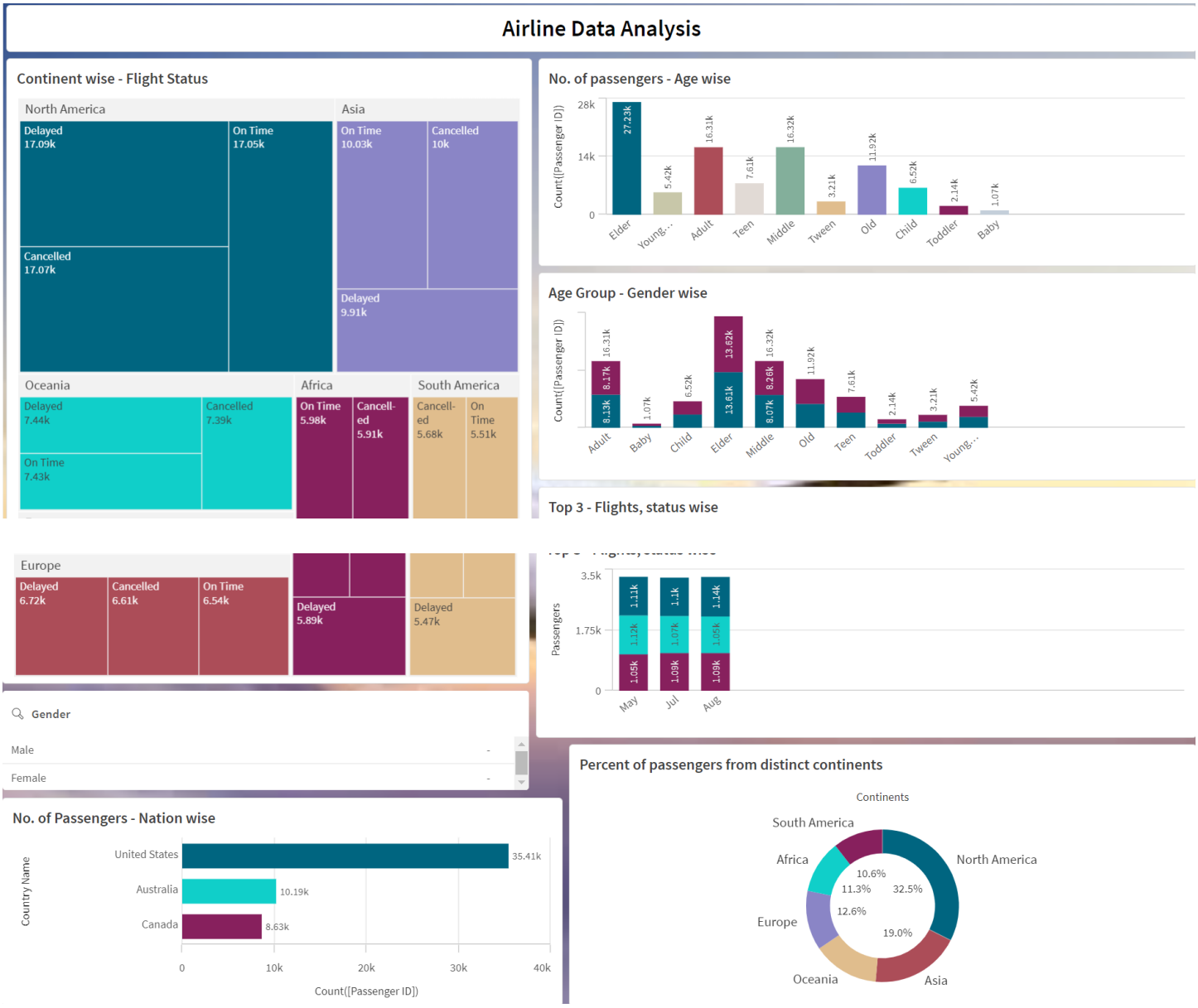
# 6. Dashboard:

## 6.1 Responsive and Design of Dashboard:

- A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case.



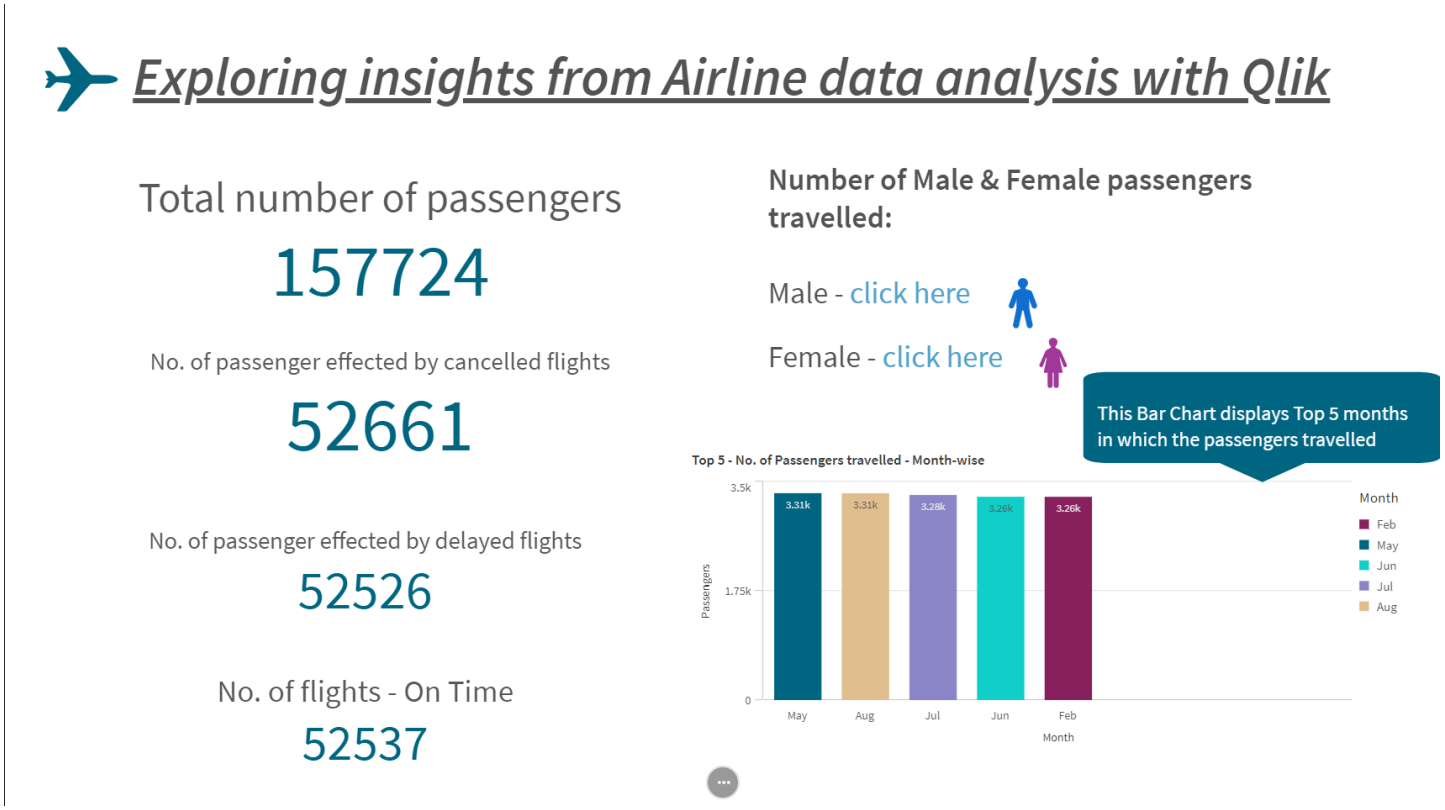

- Dashboard - 1

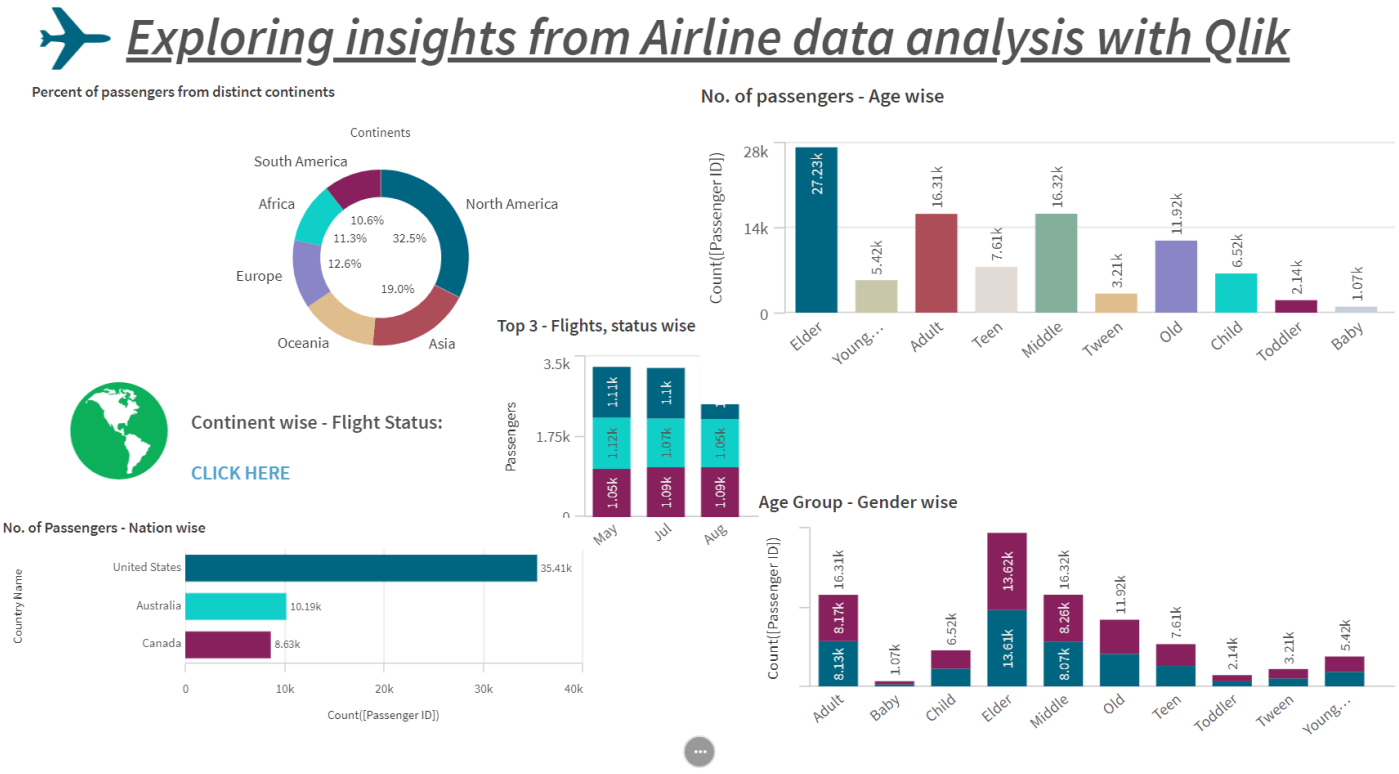

Airline Data Analysis
Continent wise - Flight Status
North America
Delayed 17.09k
On Time 17.05k
Cancelled 17.07k
Asia
On Time 10.03k
Cancelled 10k
Delayed 9.91k
Oceania
Delayed 7.44k
Cancelled 7.39k
On Time 7.43k
Africa
On Time 5.98k
Cancell-ed 5.91k
South America
Cancell-ed 5.68k
On Time 5.51k
Europe
Delayed 6.72k
Cancelled 6.61k
On Time 6.54k
Delayed 5.89k
Delayed 5.47k
No. of passengers - Age wise
Count([Passenger ID])
28k
14k
0
27.23k
5.42k
16.31k
7.61k
16.32k
3.21k
11.92k
6.52k
2.14k
1.07k
Elder
Young...
Adult
Teen
Middle
Tween
Old
Child
Toddler
Baby
Age Group - Gender wise
Count([Passenger ID])
16.31k
8.17k
8.13k
1.07k
6.52k
13.62k
13.61k
16.32k
8.26k
8.07k
11.92k
7.61k
2.14k
3.21k
5.42k
Adult
Baby
Child
Elder
Middle
Old
Teen
Toddler
Tween
Young...
Top 3 - Flights, status wise
Passengers
3.5k
1.75k
0
1.11k
1.12k
1.05k
1.1k
1.07k
1.09k
1.14k
1.05k
1.09k
May
Jul
Aug
Gender
Male
Female
Percent of passengers from distinct continents
Continents
South America
Africa
North America
Europe
Oceania
Asia
10.6%
11.3%
32.5%
12.6%
19.0%
No. of Passengers - Nation wise
Country Name
United States
Australia
Canada
35.41k
10.19k
8.63k
0
10k
20k
30k
40k
Count([Passenger ID])

- Dashboard - 2

# 7. Report:

## 7.1 Report Creation:



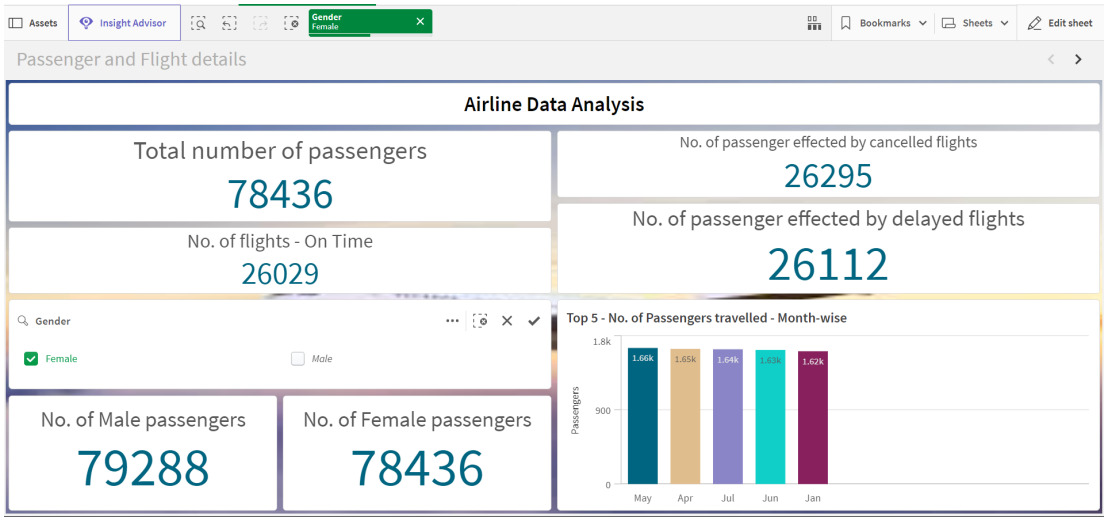
# 8. Performance Training:

8.1 Amount of Data Rendered:

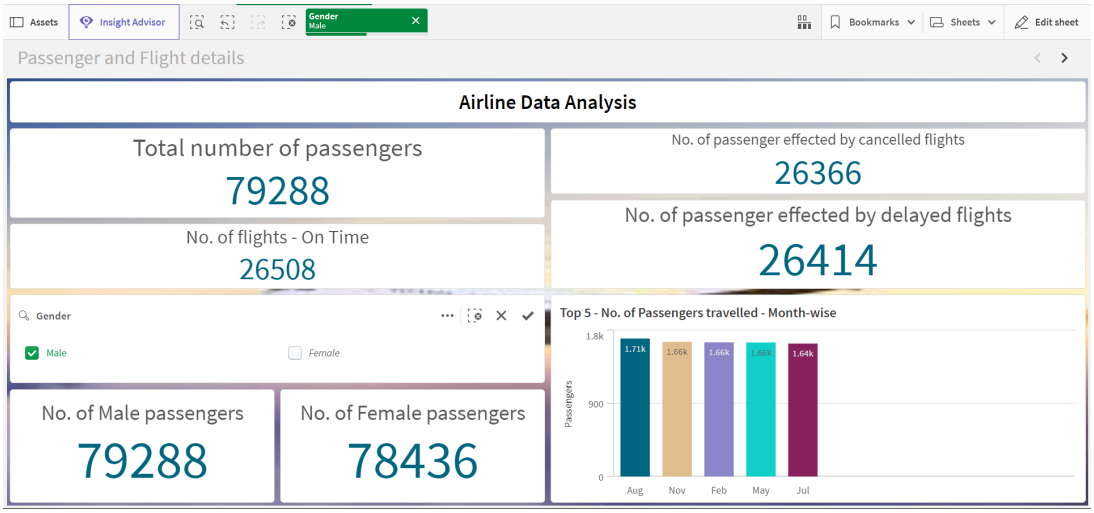
- Airline Dataset-1.Passenger ID
- Airline Dataset-1.First Name
- Airline Dataset-1.Last Name
- Airline Dataset-1.Gender
- Airline Dataset-1.Age
- Airline Dataset-1.Nationality
- Airline Dataset-1.Airport Name
- Airline Dataset-1.Airport Country Code
- Airline Dataset-1.Country Name
- Airline Dataset-1.Airport Continent
- Airline Dataset-1.Continents
- Airline Dataset-1.Departure Date
- Airline Dataset-1.Arrival Airport
- Airline Dataset-1.Pilot Name
- Airline Dataset-1.Flight Status
- Airline Dataset-1.Airline Dataset..Nationality_GeoInfo
- Airline Dataset-1.Airline Dataset..Airport Country Code_GeoInfo
- Airline Dataset-1.Airline Dataset..Country Name_GeoInfo
- AgeGroup
- Departure_Date
- Year
- Month

8.2 Utilization of Data Filters:

- Utilizing data filters in Qlik is a crucial aspect of interactive data analysis and visualization. Filters enable users to narrow down data sets to specific subsets, making it easier to focus on relevant information, identify trends, and gain insights.
- Filters are used to narrow down the scope of data, focusing only on the relevant information that meets certain predefined criteria.
- Filters can be applied in various ways, such as through selection boxes, sliders, buttons, and dynamic filtering expressions.




- This shows the responsive dashboard which on applying the Gender filter as 'Female' shows the analysis on Senior Female passengers.




- This shows the responsive dashboard which on applying the Gender filter as 'Male' shows the analysis on Senior Male passengers.

# Conclusion:

- Creating reports in Qlik involves designing and developing interactive dashboards along with the use of bookmark and visualizations that provide meaningful insights from the data.

- The project underscored Qlik's powerful role in transforming airline data into actionable insights. By uncovering patterns, trends, and correlations, Qlik empowered stakeholders to enhance operational efficiency, improve customer satisfaction, ensure safety, optimize revenue, and maintain a competitive edge in the aviation industry.

- This data-driven approach positions airlines and airports to navigate industry complexities and drive sustainable growth.