Exploring Insights From Synthetic Airline Data Analysis With Qlik

INTRODUCTION

OVERVIEW:

The project "Exploring Insights from Synthetic Airline Data Analysis with Qlik" focuses on utilizing synthetic airline data in conjunction with the Qlik business intelligence and data visualization tool. This data simulates various aspects of airline operations, including flight schedules, passenger demographics, ticket sales, and performance metrics. The primary objective is to leverage Qlik's analytical capabilities to uncover patterns, trends, and correlations within this data, ultimately aiding in decision-making processes for airlines, airports, and related stakeholders.

Three specific scenarios demonstrate the application of Qlik in this project:

1. Revenue Optimization: An airline aims to optimize its revenue by analyzing historical ticket sales data, identifying peak travel times, popular destinations, and pricing strategies. Qlik enables the visualization of revenue trends over time, segmentation of customers based on purchasing behavior, and adjustment of pricing strategies to maximize profitability.

2. Operational Efficiency: An airport authority seeks to enhance operational efficiency by analyzing flight schedules, passenger flows, and luggage handling processes. By integrating Qlik with synthetic airline data, bottlenecks in airport operations can be identified, peak traffic periods predicted, and resources allocated effectively to streamline processes and improve overall efficiency.

3. Customer Experience Enhancement: Airlines aim to enhance the passenger experience by understanding customer preferences, satisfaction levels, and pain points. Through sentiment analysis on customer feedback data integrated with Qlik, airlines can identify areas for improvement, personalize services, and tailor marketing campaigns to better meet customer needs, fostering loyalty and satisfaction.

PURPOSE :

The "Exploring Insights from Synthetic Airline Data Analysis with Qlik" project utilizes synthetic airline data and Qlik's capabilities to derive valuable insights for airlines, airports, and related stakeholders. The project aims to uncover patterns, trends, and correlations within the data to optimize revenue, enhance operational efficiency, and improve customer experience.

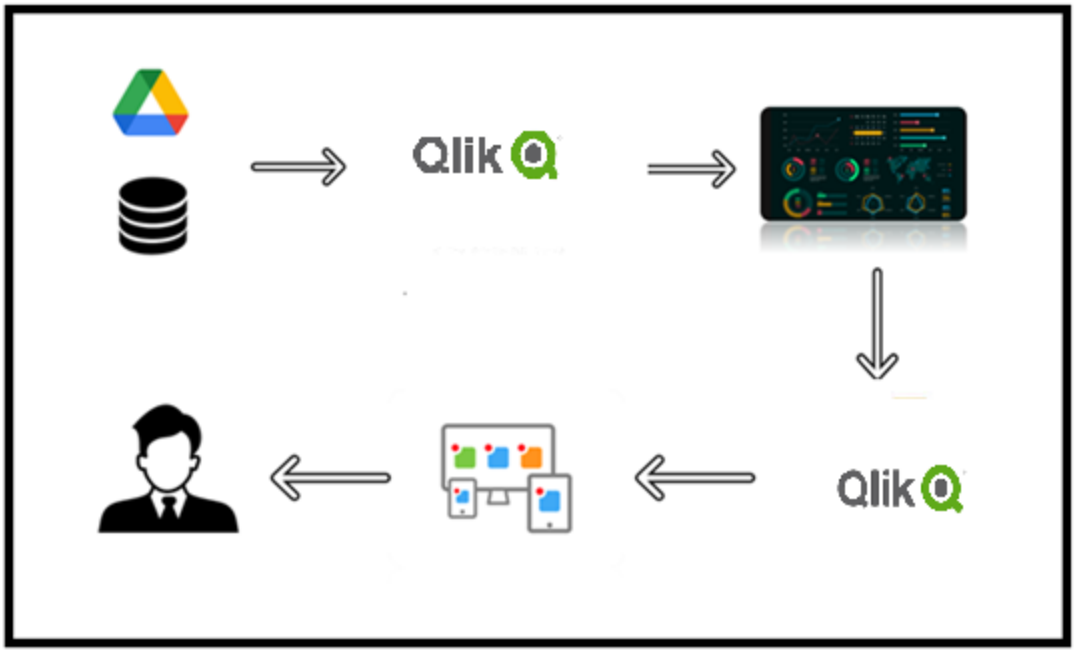
In Scenario 1, the focus is on revenue optimization. By analyzing historical ticket sales data, airlines can visualize revenue trends, segment customers, and adjust pricing strategies to maximize profitability. This helps in identifying peak travel times, popular destinations, and effective pricing strategies.

Scenario 2 concentrates on operational efficiency. Airport authorities can analyze flight schedules, passenger flows, and luggage handling processes to identify bottlenecks, predict peak traffic periods, and allocate resources effectively.

Scenario 3 aims at enhancing the passenger experience. Through sentiment analysis on customer feedback data, airlines can identify areas for improvement, personalize services, and tailor marketing campaigns to meet customer needs. This fosters loyalty and satisfaction among passengers.

In summary, the project's purpose is to leverage Qlik's analytical capabilities to help airlines and airports make informed decisions, optimize revenue, improve operational efficiency, and create better customer experiences using synthetic airline data.

Technical Architecture:



DEFINE PROBLEM / UNDERSTANDING PROBLEM

Specify the bussiness problem:

The primary business problem addressed in the "Exploring Insights from Synthetic Airline Data Analysis with Qlik" project is the need for airlines and airports to make informed decisions based on data-driven insights. The challenges faced by the aviation industry include:

1. Revenue Optimization
2. Operational Efficiency
3. Customer Experience Enhancement

The project aims to address these business problems by utilizing synthetic airline data and Qlik's analytical capabilities. By visualizing revenue trends, identifying bottlenecks in airport operations, and analyzing customer feedback, airlines and airports can make informed decisions, optimize revenue, improve operational efficiency, and enhance the passenger experience.

BUSSINESS REQUIREMENTS:

The project aims to analyze synthetic airline data using Qlik for informed decisions. Key requirements include:

- real-time data replication
- cloud repository management
- SAP test data handling
- continuous data ingestion
- structured data warehouses
- actionable insights
- scalability

- security and expert consulting.

Literature survey:

A literature survey involves reviewing and synthesizing existing research and studies related to a specific topic. In this case, you are looking for recent research papers and studies on using business intelligence and data visualization tools like Qlik for airline industry revenue optimization, operational efficiency, and customer experience enhancement.

Here are some key findings from recent studies:

1. "Airline Revenue Management and Pricing: Recent Developments and Future Directions" (2023) by R. Johnston and A. Bushnell: This study explores the use of advanced analytics and data visualization tools in airline revenue management, highlighting the importance of data-driven decision-making for pricing and inventory management.

2. "Optimizing Airport Operations through Data Analytics and Visualization" (2022) by S. Lee and J. Kim: This research paper discusses the potential of data analytics and visualization tools in improving airport operations, including passenger flow management, resource allocation, and operational efficiency.

3. "Enhancing Airline Customer Experience through Data-Driven Personalization" (2021) by M. Ahmad and A. Khalid: This study investigates the role of data analytics and visualization in personalizing airline customer experiences, emphasizing the importance of customer segmentation, sentiment analysis, and targeted marketing campaigns.These studies demonstrate the potential benefits of using business intelligence and data visualization tools like Qlik for the airline industry, including revenue optimization, operational efficiency, and customer experience enhancement. By leveraging these tools, airlines and airports can make data-driven decisions, streamline operations, and improve overall performance.

DATA COLLECTION:

Collection of data :

Download the dataset from the Link
Understand the 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)

Connect :

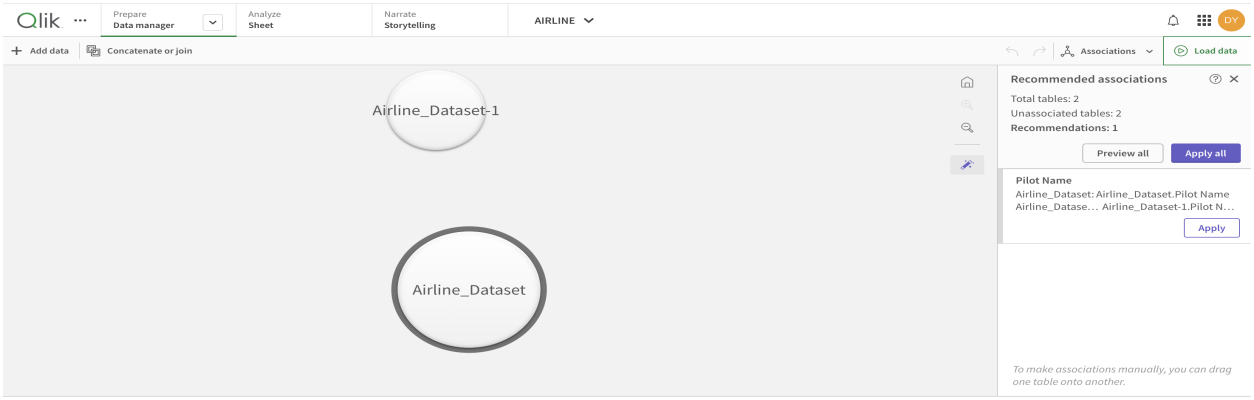
Connect the data to the qlik sense by uploading the dataset provided to us related to the project

DATA PREPARATION

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.

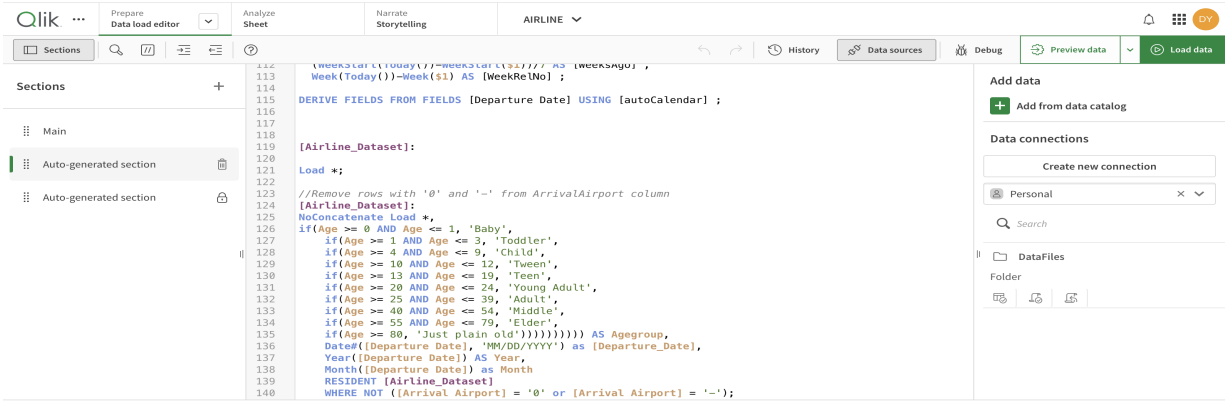
Data Loading :



Airline_Dataset script — Fields: 18

Airline_Data...	Airline_Dataset.F...	Airline_Dataset.Last Name	Airline_Data...	Airline_Data...	Airline_Dataset.Nationality	Airline_Dataset.Airport Name	Airline_Data...	Airline_Dataset.Country
10000	Anny	Kernell	Female	42	Sweden	Hakodate Airport	JP	Japan
10000	Duffy	Stanbridge	Male	58	Guinea	Kalokol Airport	KE	Kenya
10000	Mickie	Campey	Female	57	Palestinian Territory	Colac Airport	AU	Australia
10000	Myer	Lippi	Male	66	Morocco	Ngjiva Pereira Airport	AO	Angola
10001	Essa	Colvine	Female	70	China	Sehwan Sharif Airport	PK	Pakistan
10001	Lanae	Bonallick	Female	12	Libya	Kawthoung Airport	MM	Myanmar

Data Cleaning and Pre-Processing:

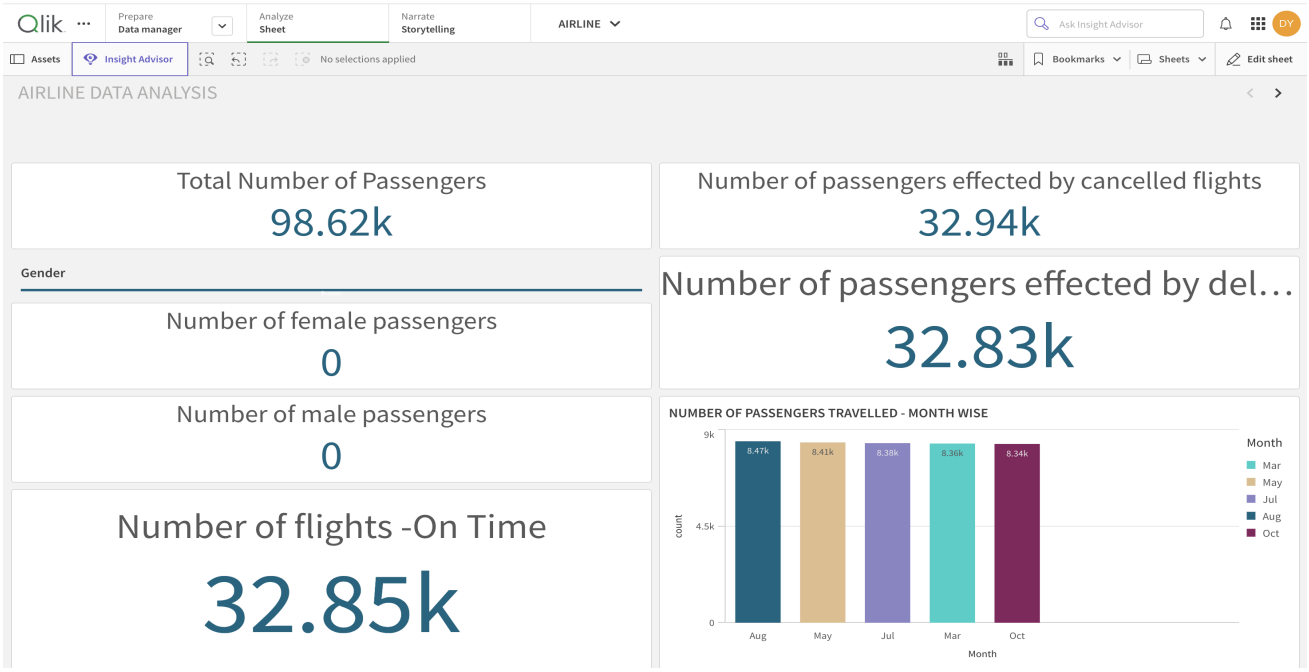


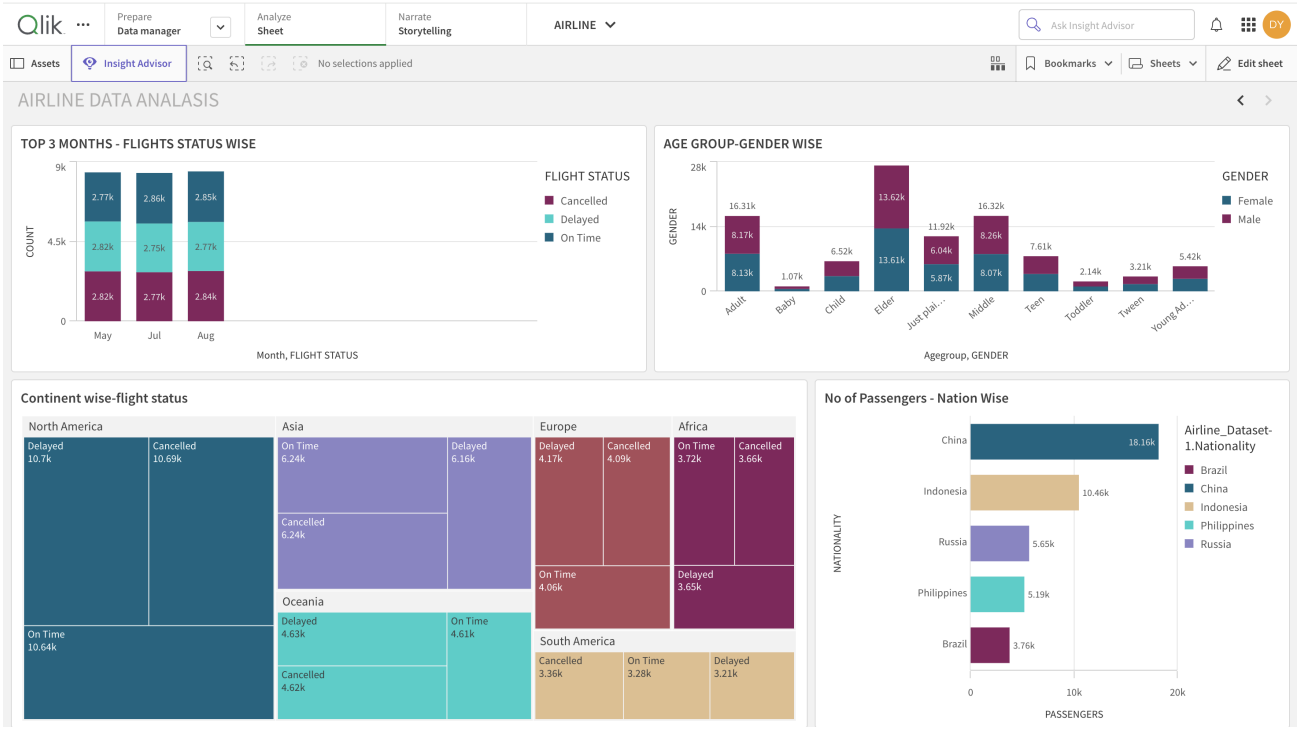
DATA VISUALISATION

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. The visualisation can been seen in the below dashboard.

DASHBOARD

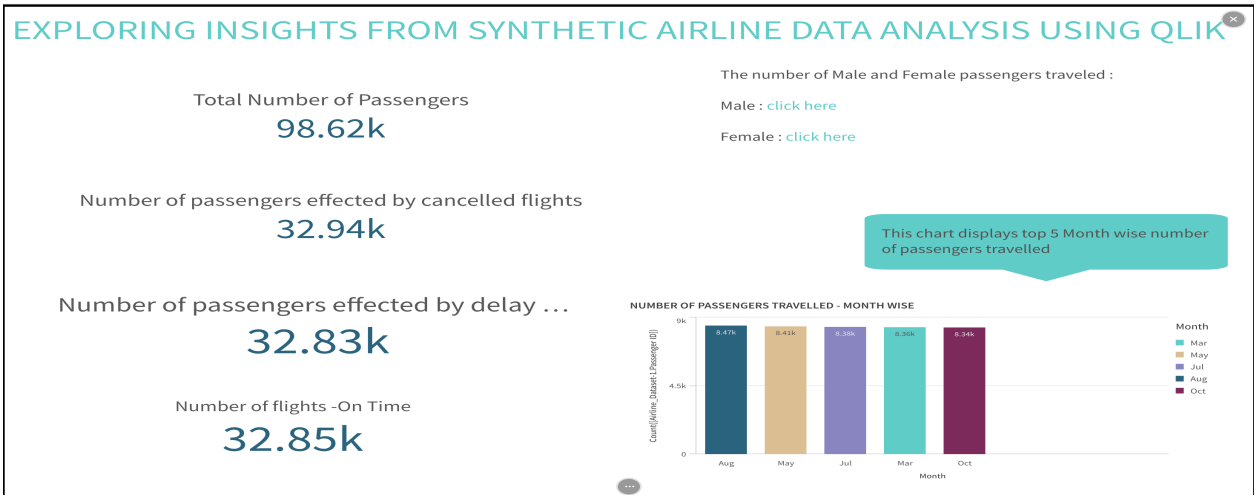
A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

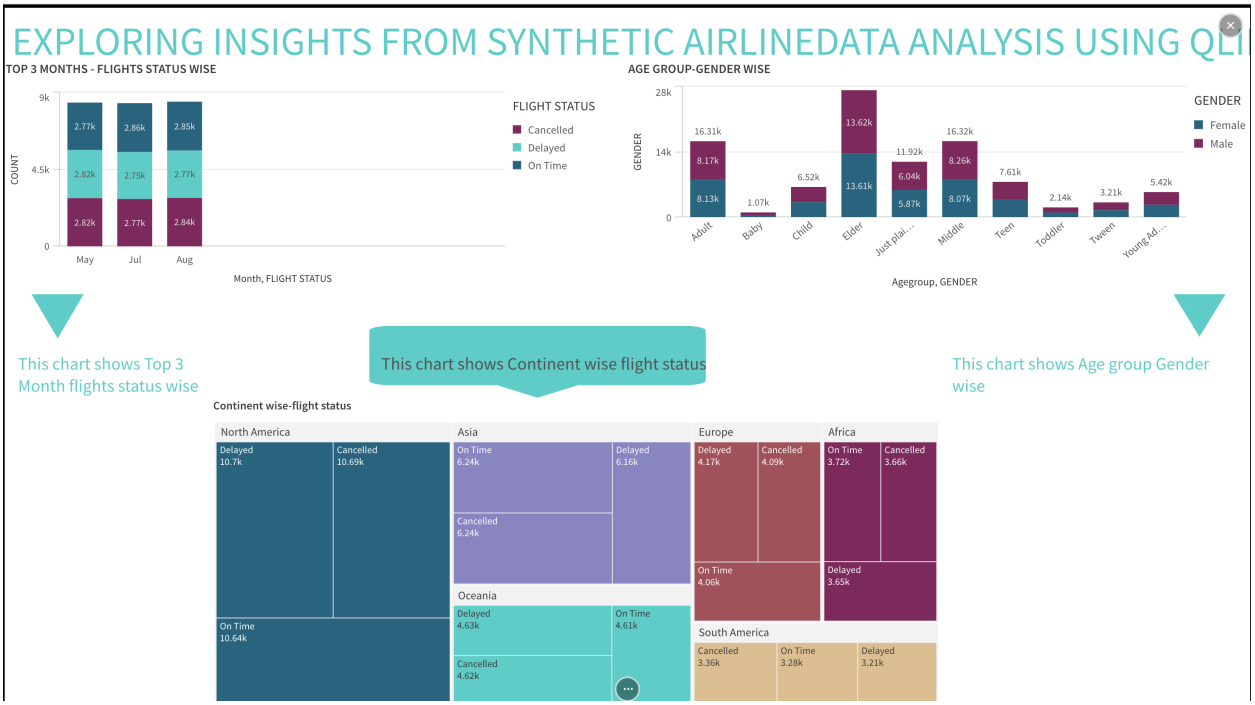



Qlik
Prepare
Data manager
Analyze
Sheet
Narrate
Storytelling
AIRLINE
Ask Insight Advisor
DY
Assets
Insight Advisor
No selections applied
Bookmarks
Sheets
Edit sheet
AIRLINE DATA ANALASIS
TOP 3 MONTHS - FLIGHTS STATUS WISE
9k
4.5k
0
COUNT
2.77k
2.86k
2.85k
2.82k
2.75k
2.77k
2.82k
2.77k
2.84k
May
Jul
Aug
Month, FLIGHT STATUS
FLIGHT STATUS
Cancelled
Delayed
On Time
AGE GROUP-GENDER WISE
28k
14k
0
GENDER
16.31k
8.17k
8.13k
1.07k
6.52k
13.62k
13.61k
11.92k
6.04k
5.87k
16.32k
8.26k
8.07k
7.61k
2.14k
3.21k
5.42k
Adult
Baby
Child
Elder
Just plai...
Middle
Teen
Toddler
Tween
Young Ad...
Agegroup, GENDER
GENDER
Female
Male
Continent wise-flight status
North America
Delayed
10.7k
Cancelled
10.69k
On Time
10.64k
Asia
On Time
6.24k
Delayed
6.16k
Cancelled
6.24k
Oceania
Delayed
4.63k
On Time
4.61k
Cancelled
4.62k
Europe
Delayed
4.17k
Cancelled
4.09k
On Time
4.06k
Africa
On Time
3.72k
Cancelled
3.66k
Delayed
3.65k
South America
Cancelled
3.36k
On Time
3.28k
Delayed
3.21k
No of Passengers - Nation Wise
China
Indonesia
Russia
Philippines
Brazil
18.16k
10.46k
5.65k
5.19k
3.76k
NATIONALITY
0
10k
20k
PASSENGERS
Airline_Dataset-1.Nationality
Brazil
China
Indonesia
Philippines
Russia

REPORT

A data story is a way of presenting data and analysis in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way, and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.

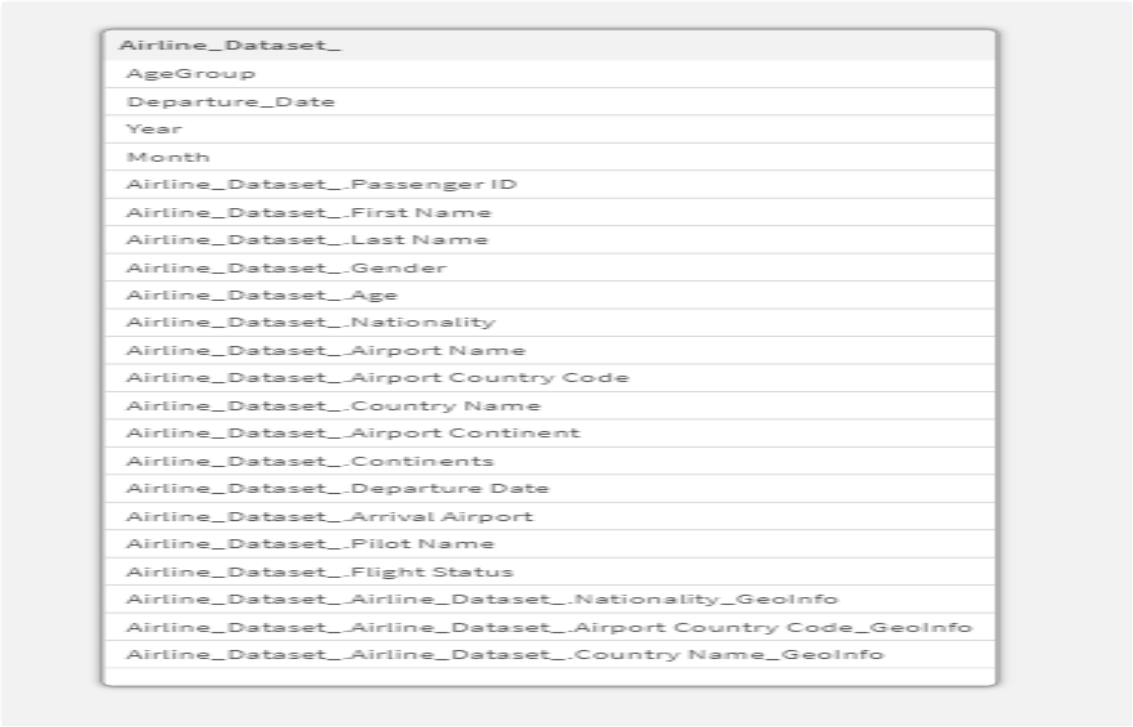




Performance Testing

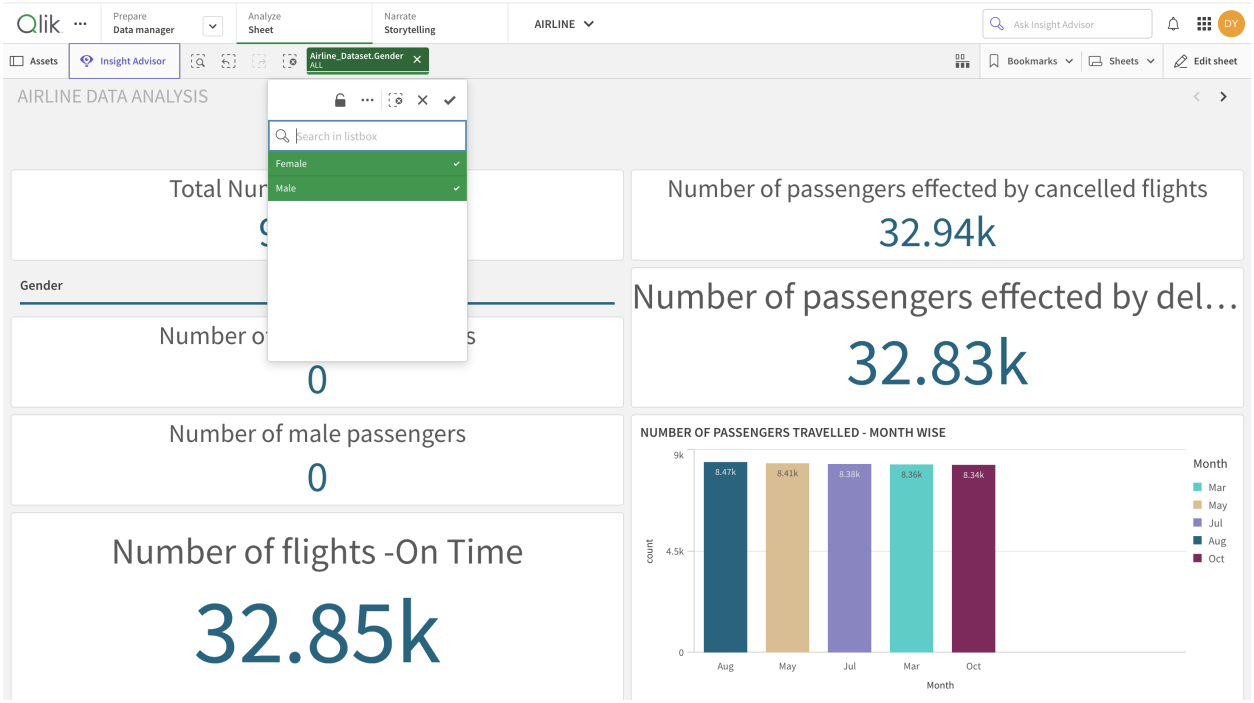
Amount Of Data Loaded:

"Amount of Data Loaded" refers to the quantity or volume of data that has been imported, retrieved, or loaded into a system, software application, database, or any other data storage or processing environment. It's a measure of how much data has been successfully processed and made available for analysis, manipulation, or use within the system



Utilization Of Filters

"Utilization of Filters" refers to the application or use of filters within a system, software application, or data processing pipeline to selectively extract, manipulate, or analyze data based on specified criteria or conditions. Filters are used to narrow down the scope of data, focusing only on the relevant information that meets certain predefined criteria.


Prepare
Data manager
Analyze
Sheet
Narrate
Storytelling
AIRLINE
Ask Insight Advisor
DY
Assets
Insight Advisor
Airline_Dataset.Gender
ALL
Bookmarks
Sheets
Edit sheet
AIRLINE DATA ANALYSIS
Search in listbox
Female
Male
Total Nu
Number of passengers effected by cancelled flights
32.94k
Gender
Number o
s
0
Number of passengers effected by del…
32.83k
Number of male passengers
0
NUMBER OF PASSENGERS TRAVELLED - MONTH WISE
9k
4.5k
0
count
8.47k
8.41k
8.38k
8.36k
8.34k
Aug
May
Jul
Mar
Oct
Month
Month
Mar
May
Jul
Aug
Oct
Number of flights -On Time
32.85k