Project Report

On Business Analytics

INDEX

INTRODUCTION

1.1 Overview : A brief description of Project:

Airline Data Analysis With Qlik

The project "Exploring Insights from Synthetic Airline Data Analysis with Qlik" aims to utilize synthetic airline data to uncover valuable insights using the powerful business intelligence and data visualization tool, Qlik. The synthetic airline data replicates key aspects of airline operations such as flight schedules, passenger demographics, ticket sales, and performance metrics.

By leveraging Qlik's advanced analytical capabilities, the project seeks to identify patterns, trends, and correlations within the synthetic airline data. These insights are intended to assist airlines, airports, and other stakeholders in making informed decisions related to their operations and strategies.

Ultimately, through the analysis of synthetic airline data with Qlik, this project aims to provide actionable insights that can enhance operational efficiency, improve customer experiences, and drive strategic decision-making within the aviation industry

1.2Purpose: The use of the Project:

- **Decision-Making Support:** The project aims to provide valuable insights derived from the analysis of synthetic airline data to support decision-making processes for airlines, airports, and other stakeholders in the aviation industry. These insights can aid in optimizing operations, improving efficiency, and enhancing overall performance.

- **Understanding Market Dynamics:** The project seeks to uncover trends and patterns in flight schedules, passenger demographics, ticket sales, and performance metrics, offering a deeper understanding of market dynamics.

- **Enhancing Operational Efficiency:** Through the analysis of synthetic airline data, the project aims to highlight opportunities for enhancing operational efficiency within airline operations. This could include optimizing flight schedules, improving resource allocation, and streamlining processes to reduce costs and improve overall performance.

- **Improving Customer Experiences:** By gaining insights from passenger demographics and behavior analysis, the project can assist in improving customer experiences and tailoring services to meet the needs and preferences of different passenger segments. This could lead to increased customer satisfaction and loyalty.

- **Strategic Decision-Making:** The insights derived from the analysis of synthetic airline data can also be valuable in strategic decision-making processes. By understanding market trends, competitor analysis, and performance metrics, airlines and related stakeholders can make informed decisions to stay competitive and drive growth.

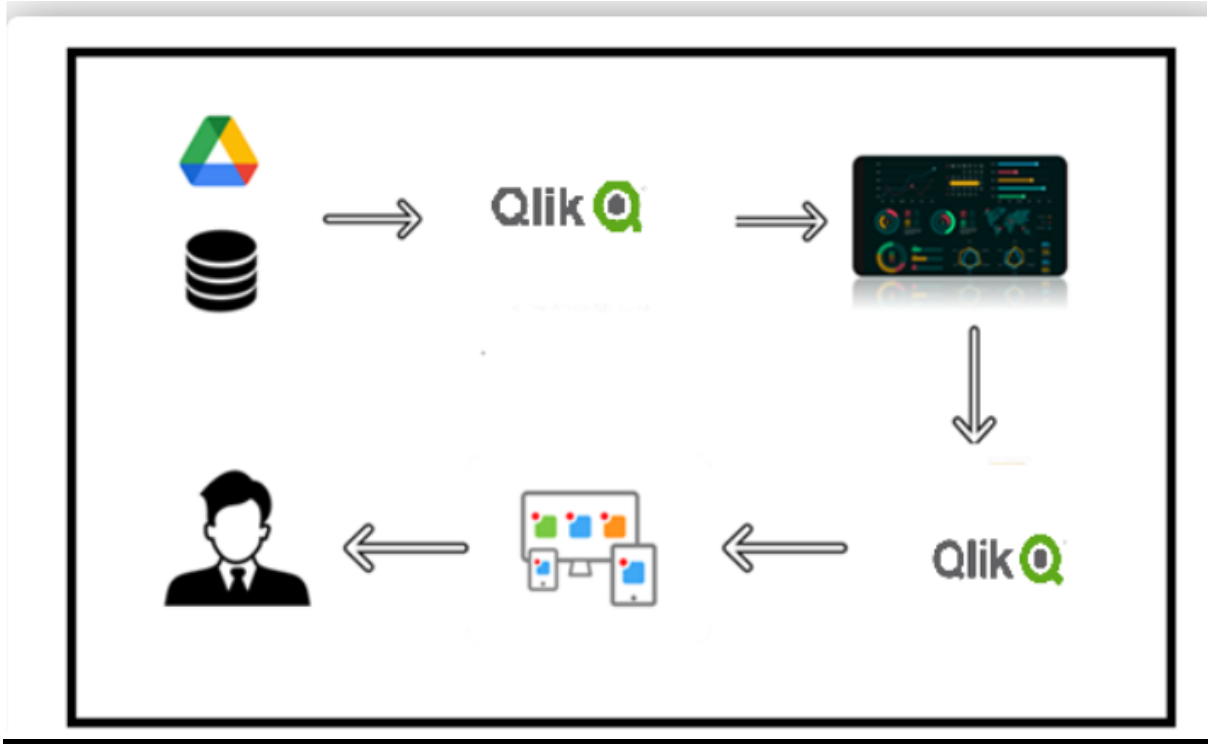
1.3 Technical Architecture:

- **Qlik Sense:** Qlik Sense is the business intelligence and data visualization tool used for analyzing and visualizing the synthetic airline data.

- **Data Storage:** The synthetic airline data is stored in a data repository or database that can be easily accessed by Qlik for analysis.

- **Dashboard and Report Generation:** The insights and visualizations created within Qlik Sense can be used to build interactive dashboards and reports that can be shared with stakeholders for decision-making purposes

- **User Interaction and Analysis:** End users access the Qlik dashboards through various devices. They can interact with the visualizations, apply filters, drill down into details, and generate custom reports..



PROBLEM UNDERSTANDING

2.1 Specify the business problem:

The project "Exploring Insights from Synthetic Airline Data Analysis with Qlik" associated with the number of flights and passengers in different continents and respective countries airports. In our project we will focus on which continent has maximum passengers and flights. At the end of project we come to know about continents airline need to focus on for following :

- Identify continents and specific countries with growing passenger numbers and flight volumes.
- Focus on expanding market presence and increasing flight frequencies to high-demand destinations.
- Analyze fare trends and passenger spending patterns in high-traffic regions.
- Implement dynamic pricing strategies and promotional campaigns in high-demand areas to maximize revenue.
- Optimize flight schedules and routes to improve load factors and profitability.
- Identification of key markets for strategic focus.

2.2 Business Requirements:

1. Data Analysis Objectives:

- Identify Continents with Highest Traffic: Determine the top continents based on:
    - Total passenger volume
    - Total number of flights
- Analyze Country-Level Trends (Optional): Within the top continents, identify countries with significant passenger and flight traffic.

2. Data Requirements:

- Synthetic Airline Data:

    Passenger data (e.g., origin, destination, date of travel)

    Flight data (e.g., flight number, origin airport, destination airport, departure date/time, arrival date/time)

    Geographic Data (if not included in the above):

3. Reporting and Visualization:

- Qlik Sense Dashboard: Develop an interactive dashboard that:

    Displays key metrics (passenger volume, flight frequency) by continent.

4. Success Criteria:

- Accurate and Reliable Insights: Ensure the analysis is based on accurate data and sound methodologies.
- Clear and Actionable Recommendations: Provide specific recommendations for airline focus, supported by data-driven insights.

.

2.3 Literature Survey:

- **Airline Industry Trends:** Look for recent industry reports that discuss global or regional trends in airline traffic, passenger volumes, and emerging markets. Industry associations, consulting firms, and government agencies often publish reports on these topics.

- **Geographical Analysis:** Search for literature that discusses the geographical distribution of airline traffic, including passenger numbers and flight frequencies across continents and countries.

- **Data Analysis and Visualization Techniques:** Find research articles or best practice guides on analyzing and visualizing airline data. Look for techniques and methodologies used to identify trends, patterns, and outliers in large datasets.

- **Factors Affecting Airline Focus:** Explore literature on the factors that influence airline focus, such as economic indicators, tourism trends, business travel patterns, and competition in the airline industry.

- **Synthetic Data Analysis:** Search for any literature or best practices related to analysis using synthetic datasets, including considerations for data accuracy, representation, and potential biases.

DATA COLLECTION

3.1 Collect the Dataset:

Firstly , we download the dataset from given website. Airline data holds immense importance as it offers insights into the functioning and efficiency of the aviation industry. It provides valuable information about flight routes, schedules, passenger demographics, and preferences, which airlines can leverage to optimize their operations and enhance customer experiences.

Dataset Glossary (Column-wise)

- 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)

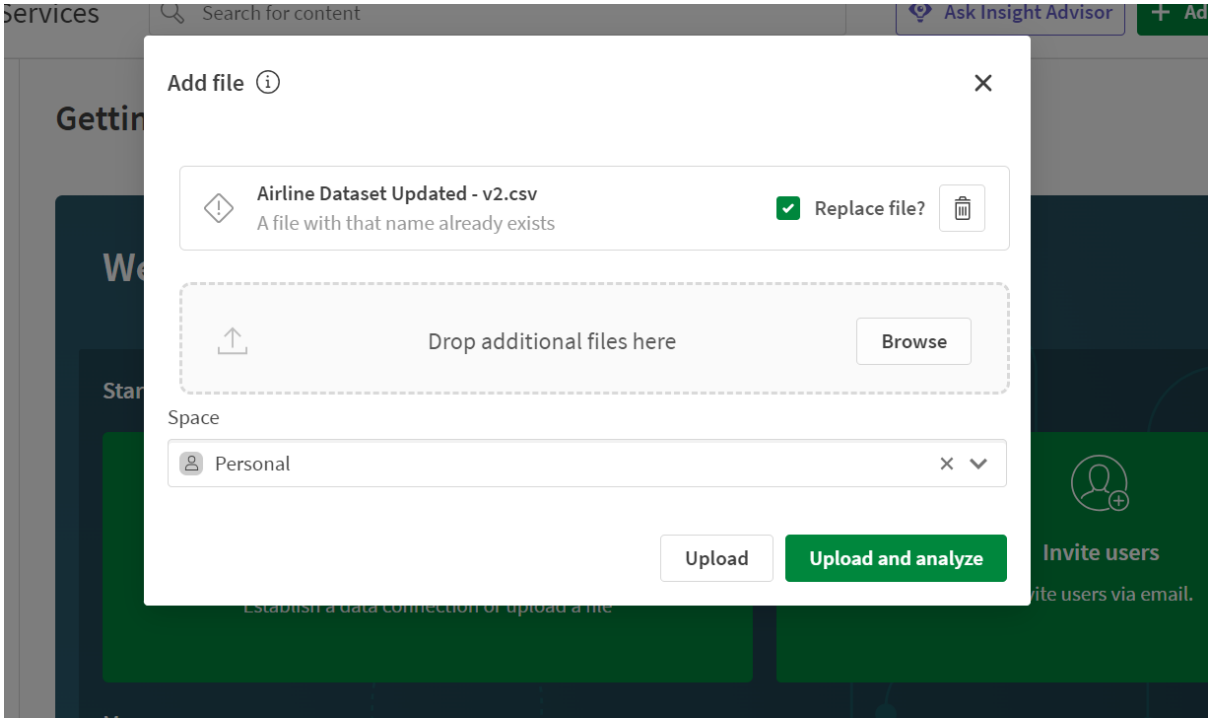
Structure of the Dataset

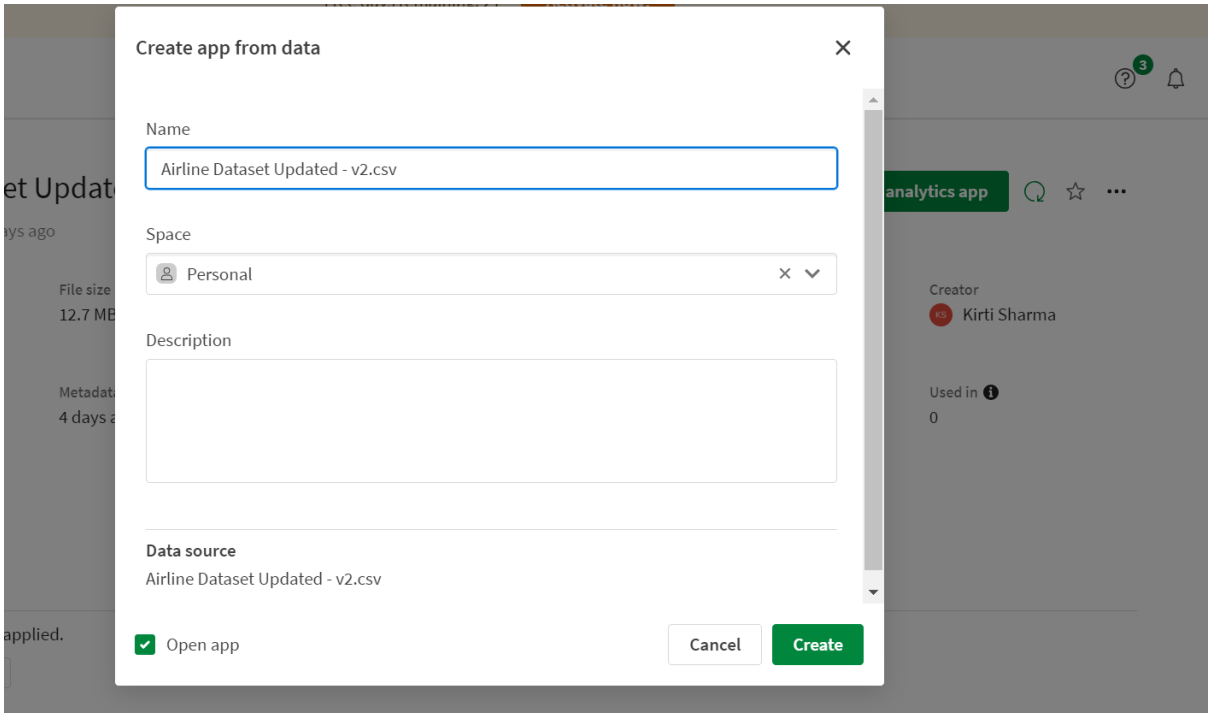
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Passenger	First Nam	Last Name	Gender	Age	Nationalit	Airport Na	Airport Co	Country N	Airport Co	Continent	Departure	Arrival Air	Pilot Nam	Flight Status	
2	ABVWIg	Edithe	Leggis	Female	62	Japan	Coldfoot A	US	United Sta	NAM	North Am	6/28/2022	CXF	Fransisco	On Time	
3	jkXXAX	Elwood	Catt	Male	62	Nicaragua	Kugluktuk	CA	Canada	NAM	North Am	12/26/202	YCO	Marla Par	On Time	
4	CdUz2g	Darby	Felgate	Male	67	Russia	Grenoble-	FR	France	EU	Europe	1/18/2022	GNB	Rhonda Ai	On Time	
5	BRS38V	Dominica	Pyle	Female	71	China	Ottawa / (	CA	Canada	NAM	North Am	9/16/2022	YND	Kacie Com	Delayed	
6	9kvTLo	Bay	Pencost	Male	21	China	Gillespie F	US	United Sta	NAM	North Am	2/25/2022	SEE	Ebonee Tr	On Time	
7	nMJKVh	Lora	Durbann	Female	55	Brazil	Coronel H	BR	Brazil	SAM	South Am	########	LEC	Inglis Doll	On Time	
8	8IPFPE	Rand	Bram	Male	73	Ivory Coas	Duxford A	GB	United Kin	EU	Europe	10/30/202	QFO	Stanislas T	Cancelled	
9	pqixbY	Perceval	Dallosso	Male	36	Vietnam	Maestro V	BR	Brazil	SAM	South Am	########	STM	Sharyl Eas	Cancelled	
10	QNAs2R	Aleda	Pigram	Female	35	Palestinia	Venice Ma	IT	Italy	EU	Europe	8/20/2022	VCE	Daryn Bar	On Time	
11	3jmudz	Burlie	Schustl	Male	13	Thailand	Vermilion	CA	Canada	NAM	North Am	########	YVG	Alameda (	On Time	
12	2P41gZ	Porty	Jori	Male	39	Tunisia	Nuevo Cas	MX	Mexico	NAM	North Am	5/27/2022	NCG	Rasia Fide	Cancelled	
13	sBf524	Briant	De La Hay	Male	71	Russia	Ruben Car	PA	Panama	NAM	North Am	########	SYP	Alina Floo	Delayed	
14	PlwJZT	Kalie	Scoble	Female	47	Sweden	Loralai Air	PK	Pakistan	AS	Asia	3/19/2022	LRG	Madelena	Delayed	
15	iU75x3	Catriona	Beaument	Female	77	Russia	Cudal Airp	AU	Australia	OC	Oceania	3/24/2022	CUG	Margie Be	Delayed	
16	GUta6R	Amberly	Handling	Female	32	China	Farmingto	US	United Sta	NAM	North Am	########	FAM	Lothaire E	Delayed	
17	8qA80a	Dyna	De'Vere - I	Female	22	China	Oudtshoo	ZA	South Afri	AF	Africa	7/18/2022	OUH	Neila Gier	Cancelled	
18	2haCDu	Janella	Hardaker	Female	28	Colombia	Zaraza Air	VE	Venezuela	SAM	South Am	9/23/2022	ZRZ	Shaylynn /	On Time	
19	WiVl8o	Alvin	Wenzel	Male	12	Greece	Enshi Airp	CN	China	AS	Asia	3/29/2022	ENH	Alfie Mach	Delayed	
20	e0H5LI	Jerrine	Peeters	Female	87	Philippine	Thompsor	CA	Canada	NAM	North Am	########	YTH	Chandra D	Cancelled	
21	nL8kyD	Warner	Driutti	Male	62	China	Guilin Lian	CN	China	AS	Asia	########	KWL	Marita Ho	Cancelled	
22	9iT79e	Paige	Hayhow	Male	24	Sweden	Crested Bu	US	United Sta	NAM	North Am	2/19/2022	CSE	Clyde Win	Delayed	
23	kEARqP	Dorisa	Skill	Female	19	Ukraine	St Augusti	CA	Canada	NAM	North Am	########	YIF	Ilyse Bartl	Cancelled	
24	dx3NWh	Bobbye	Patmore	Female	45	China	Port Berg	MG	Madagasc	AF	Africa	8/24/2022	WPB	Stella Pitt	On Time	

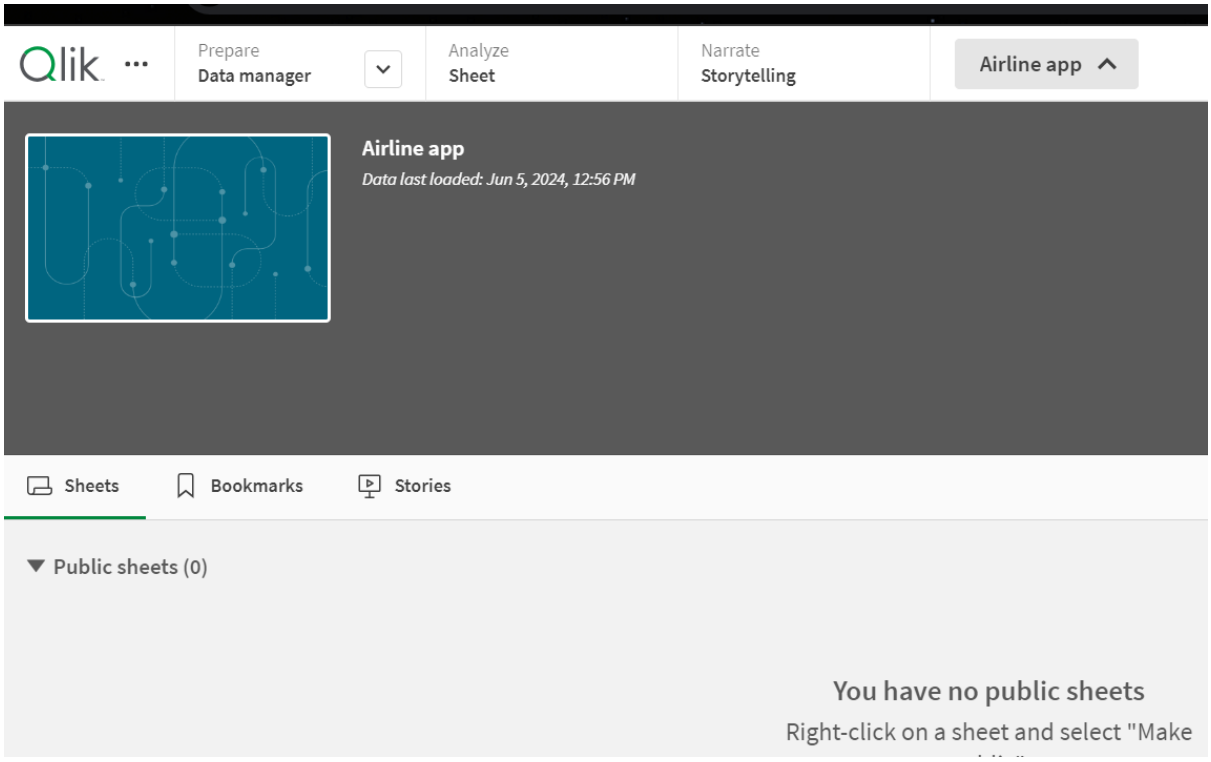
Airline Dataset Updated - v2

3.2 Connect the Data with Qlik Sense:

- **Open Qlik Sense:** Launch Qlik Sense Desktop or access Qlik Sense Enterprise through your web browser.
- **Create a New App:** Click on "Create a New App" and give it a meaningful name related to your project.
- **Add Data:** Open your new app and click on "Add data" to load your synthetic airline data. You can load data from different sources like files, databases, or web data connectors.
  - Select the data source type (e.g., file, database) and follow the prompts to connect. For instance, if you're using a CSV file:
  - Click on "Data files" and drag your CSV file into the upload area.
  - Click "Next" after the file is uploaded.
  - Qlik Sense will automatically recognize the data structure. Confirm or adjust the field mappings as needed.




Create app from data
Name
Airline Dataset Updated - v2.csv
Space
Personal
Description
Data source
Airline Dataset Updated - v2.csv
Open app
Cancel
Create
analytics app
Creator
Kirti Sharma
Used in
0


Prepare
Data manager
Analyze
Sheet
Narrate
Storytelling
Airline app
Airline app
Data last loaded: Jun 5, 2024, 12:56 PM
Sheets
Bookmarks
Stories
Public sheets (0)
You have no public sheets
Right-click on a sheet and select "Make

DATA PREPARATION

4.1 Prepare the Data for Visualization:

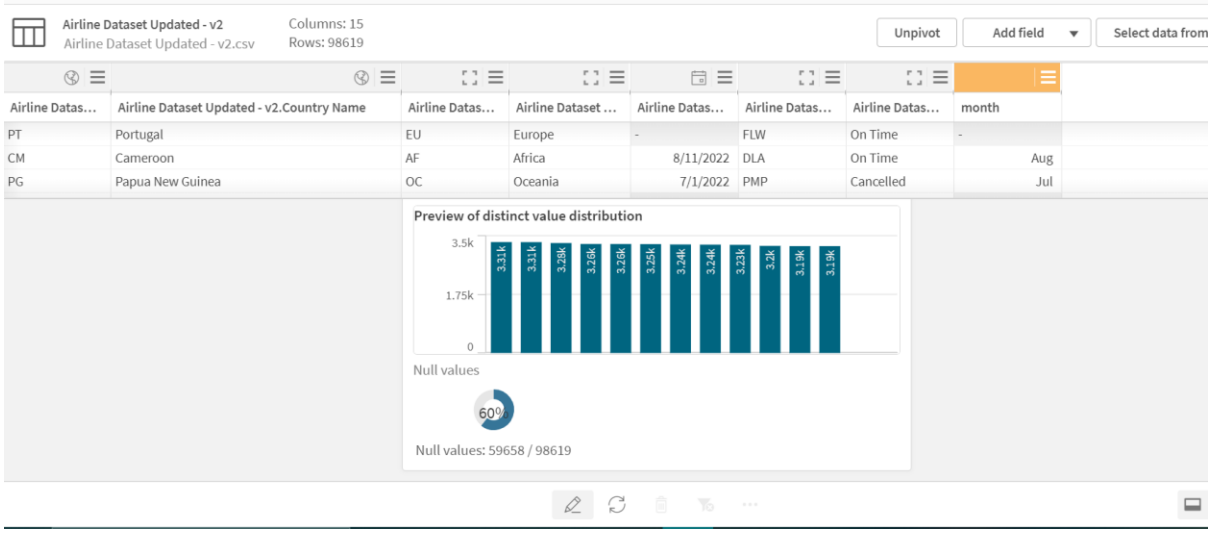
Preparing the data for visualization involves cleaning the data to remove irrelevant or missing data, transforming the data into a format that can be easily visualized, exploring the data to identify patterns and trends, filtering the data to focus on specific subsets of data, preparing the data for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

Data Cleaning & Transformation (in Qlik Sense or Before):

- Missing Values: Decide how to handle them (e.g., imputation, removal, using a default value).

Data Reduction (If Necessary):

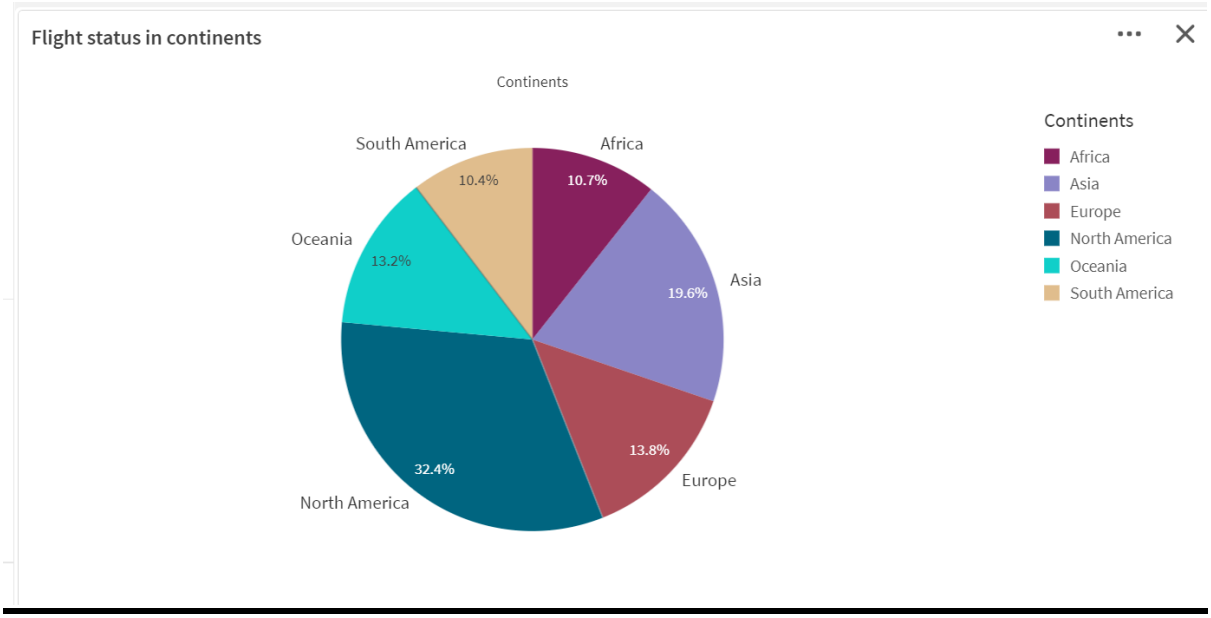
- Large Datasets: Qlik Sense handles large volumes well, but extremely large datasets might benefit from aggregation or sampling before loading.
- .



DATA VISUALIZATION

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.

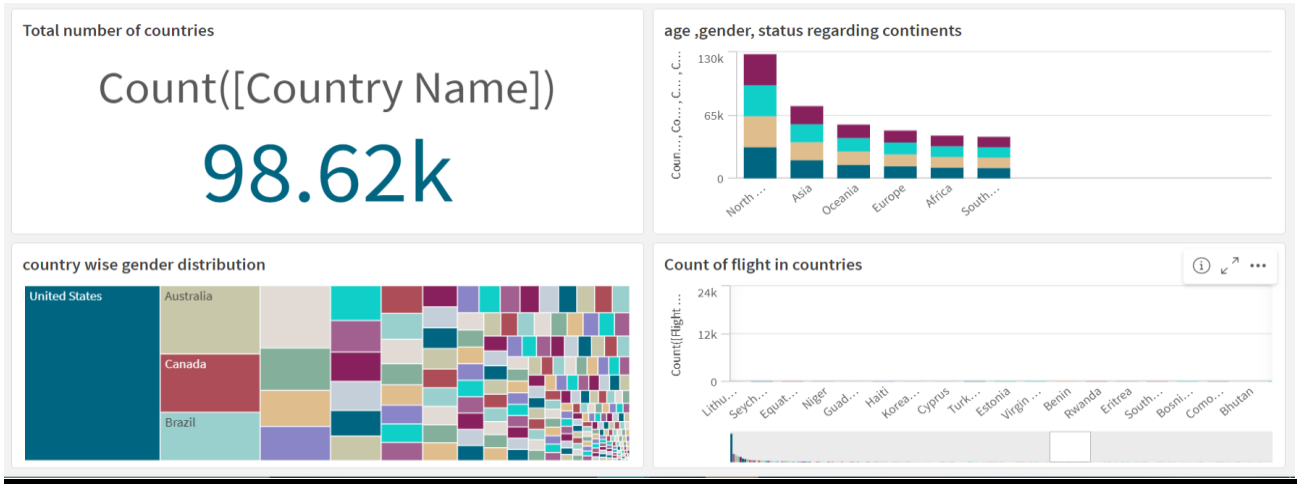


DASHBOARDS

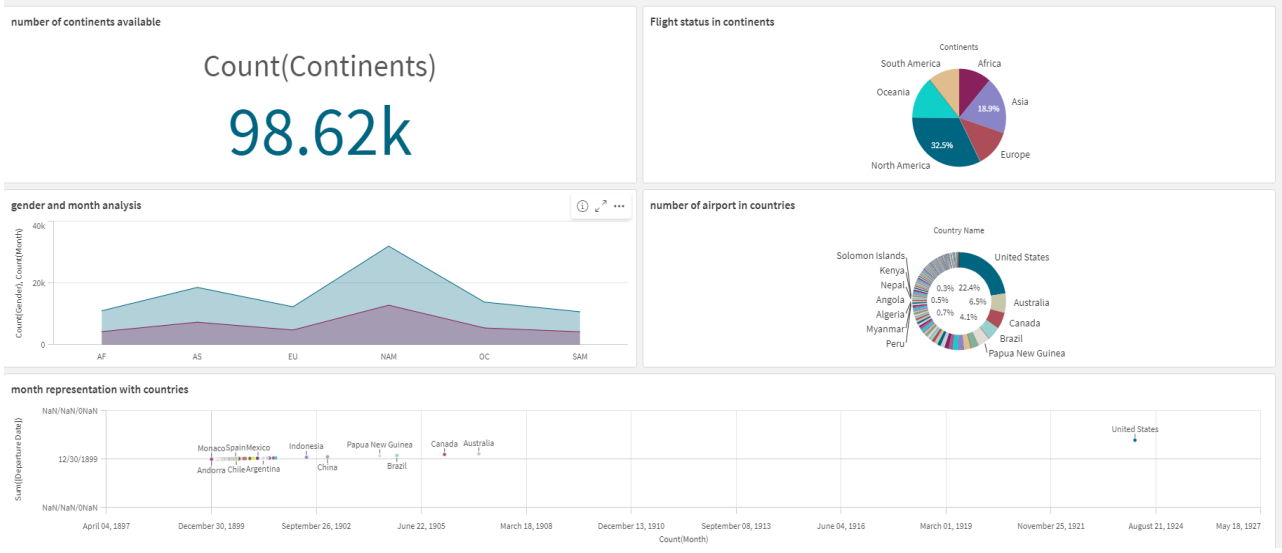
6.1 Responsive and Design of Dashboard:

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.

Dashboard1:



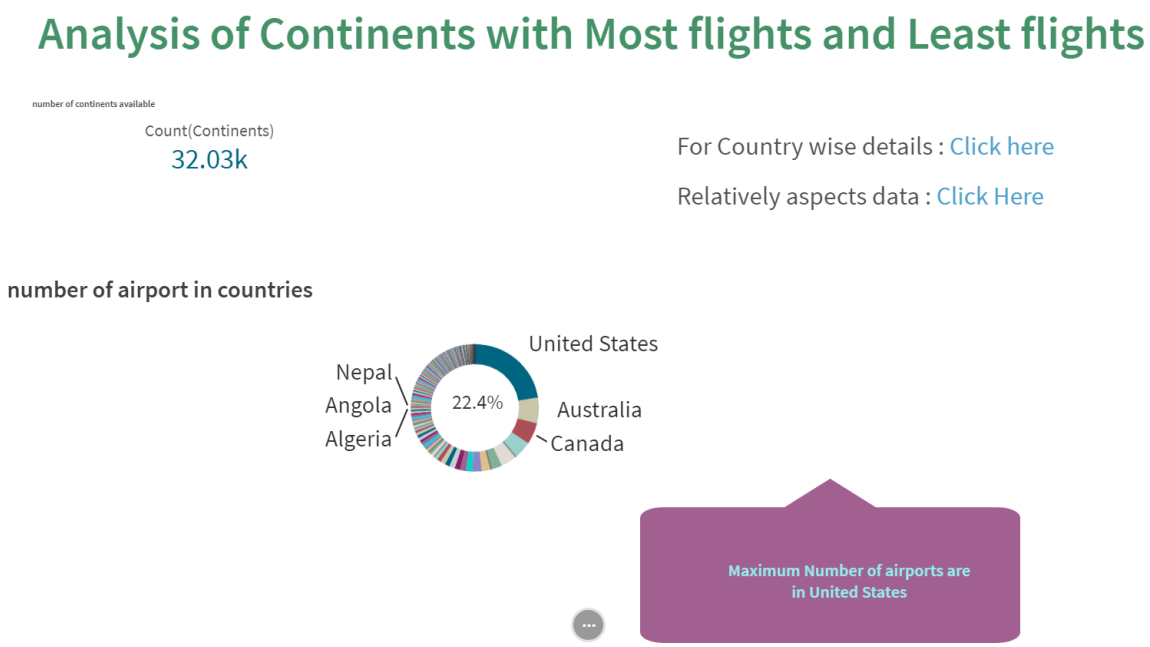
Dashboard2:



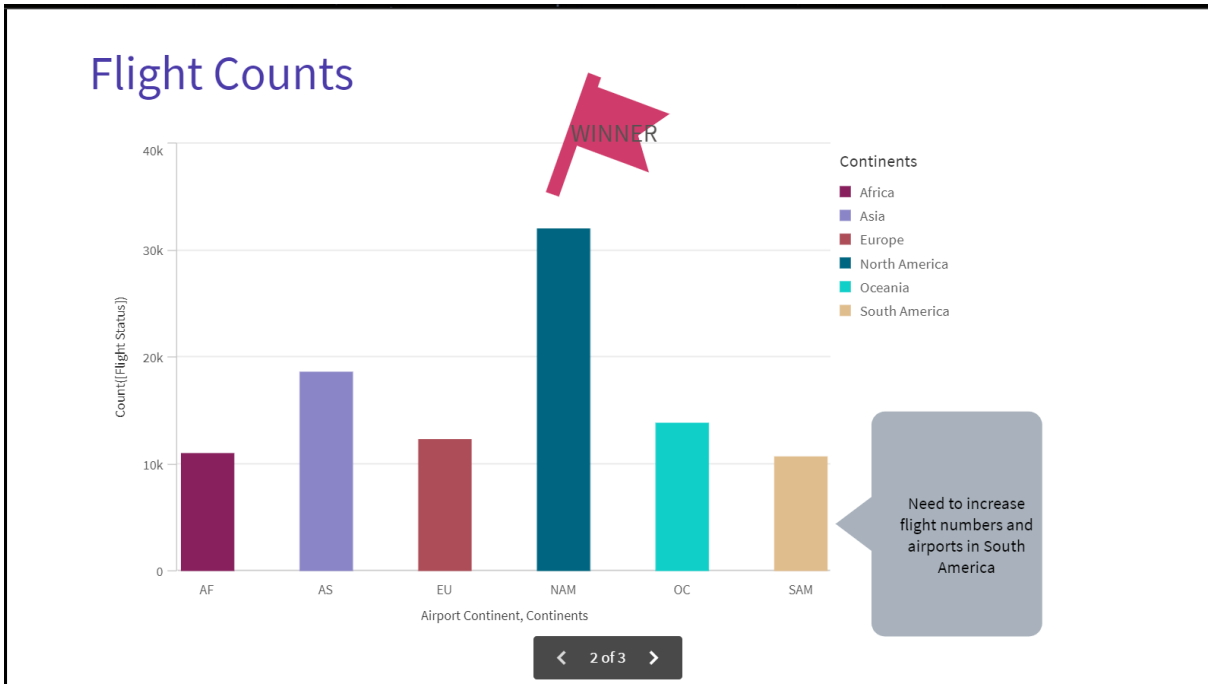
REPORT

7.1 Report Creation:

Page1:



Page2:



Page3:


Comparison
number of continents available
Count(Continents)
10.72k
SA count
age ,gender, status regarding continents
Count([Airport Continent]), Count(Nationality), Count(Gender), Count(Age)
150k
100k
50k
0
North America
Asia
Oceania
Europe
Africa
South America
Measures
Count([Airport Continent])
Count(Nationality)
Count(Gender)
Count(Age)
Reasons of less number of flights in South America
country wise gender distribution
United States
Canada
Indonesia
Russian Federation
Brazil
Colombia
India
Papua New Guinea
France
Australia
United Kingdom
China

PERFORMANCE TESTING

8.1 Amount of Data Rendered:

"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.

Number of Visualization:

- Total number of countries
- age ,gender, status regarding continents
- country wise gender distribution
- Count of flight in countries
- Flights in Continents
- number of continents available
- month representation with countries
- number of airport in countries
- gender and month analysis
- total number of passengers

8.2 Utilization of Data Filters:

It 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.