Qlik

Project Title:

Exploring Insights From Synthetic Airline Data Analysis With Qlik

PROBLEM STATEMENT :-

Exploring Insights From Synthetic Airline Data Analysis With Qlik

INTRODUCTION

- In the fast-paced and ever-evolving airline industry, the ability to glean actionable insights from data is paramount for maintaining operational efficiency, enhancing customer satisfaction, and achieving financial success. This project, "Exploring Insights From Synthetic Airline Data Analysis With Qlik," is designed to demonstrate the power of data analytics in the aviation sector using Qlik, a leading data visualization and business intelligence.

- This project leverages synthetic airline data to extract valuable insights using Qlik, a business intelligence and data visualization tool. The data simulates various aspects of airline operations, including flight schedules, passenger demographics, ticket sales, and performance metrics. The objective is to use Qlik's capabilities to identify patterns, trends, and correlations within the data, aiding decision-making for airlines, airports, and related stakeholders.

PROJECT FLOW :-

The project flow for analyzing synthetic airline data using Qlik involves several structured phases. Each phase builds upon the previous one to ensure a comprehensive analysis, from data preparation to actionable insights.

Social Impact:
• By analyzing passenger demographics and preferences, airlines can failor services to better meet customer needs, enhancing overall satisfaction and loyalty

Business Impact:
Airlines can optimize flight routes, schedules, and resources based on data insights, reducing costs and improving overall efficiency.

DATA COLLECTION & EXTRACTION

1. **Import from Kaggle:** You likely used Kaggle's interface to download your desired dataset, possibly as a CSV, Excel, or other supported file format.
2. **Upload to Qlik Sense:** In Qlik Sense, you might have used the Data Load Editor or a similar functionality to upload the downloaded file from your local machine.

Next Steps:

With the data loaded into Qlik Sense, you can now leverage its capabilities for data exploration and analysis. Here are some potential next steps:

- **Data Cleaning and Transformation:** Use Qlik Sense's data manipulation tools to clean the data (e.g., handle missing values, format inconsistencies) and transform it into the desired format for analysis.
- **Data Exploration:** Build visualizations like charts, graphs, and tables to explore the data and uncover patterns and insights. Qlik Sense's associative engine allows you to dynamically filter and analyze the data based on your selections.
- **Building Applications:** Create interactive dashboards or data applications within Qlik Sense to share your insights with colleagues or stakeholders.

Data Collection and Extraction:-Link

DATA PREPARATION:

Data Loading

Step-by-Step Guide to Load Data into Qlik Sense Cloud

1. **Log in to Qlik Sense Cloud:**
   - Open your web browser and navigate to the Qlik Sense Cloud login page.
   - Log in with your credentials.
2. **Create a New App:**
   - From the Qlik Sense Cloud hub, click on “Create new app”.
   - Give your app a meaningful name and click “Create”.
3. **Open the App:**
   - After creating the app, click on “Open app”.
4. **Add Data:**
   - In the app, click on “Add data” to begin the process of loading your dataset.
5. **Upload Your Data:**
   - **Data Files:** If your data is in a file (e.g., CSV, Excel):
     - Click on “Data files”.
     - Drag and drop your file into the designated area or click “Select file” to browse and upload your file.

Data Pre-Processing:

- **Renaming Columns:**
  - In the Data Manager, you can easily rename columns.
  - Click on a table, and then click on the column you want to rename.
  - Enter the new name for the column.
- **Adding New Fields (Calculated Fields):**
  - In the Data Load Editor, you can add new fields using expressions.
  - For example, to add a new field for total no of passengers travelled.

Setting Null Values:

- To handle null values, you can use the set`Null()` function or conditional statements in your script.
- For example, to replace null values in a column with a default value

Splitting Fields:

- You can use the `splitfunction()` function to split fields.
- For example, if you have a field with values in the format "Origin-Destination", you can split it into two fields:

Applying Filters:

- You can filter data during the load process using the `WHERE` clause.

Data collection and pre-processing-Link

Data Visualization:

Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, maps, and plots, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

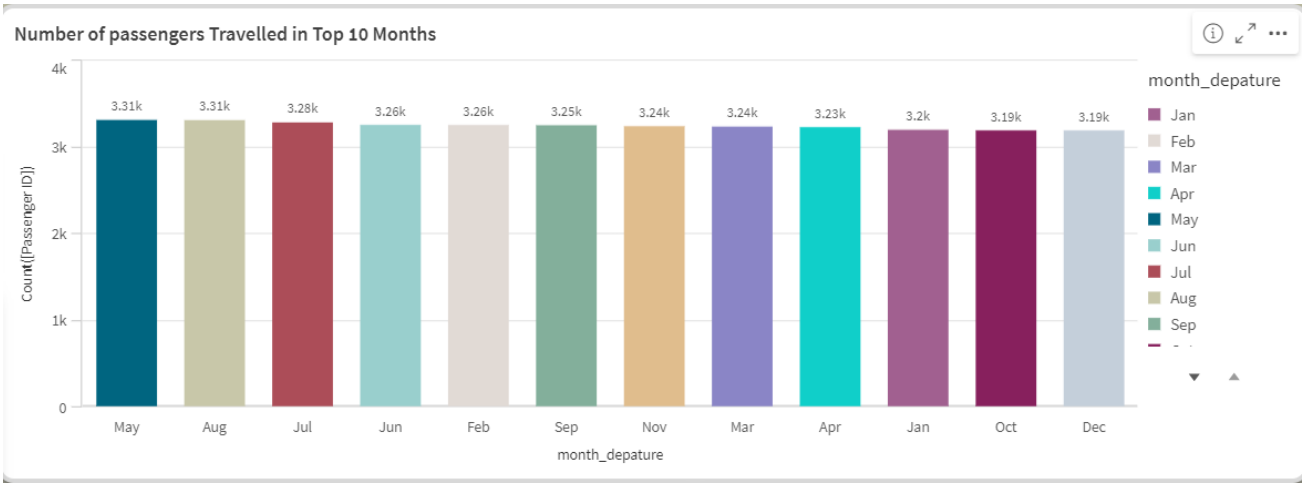
Examples:

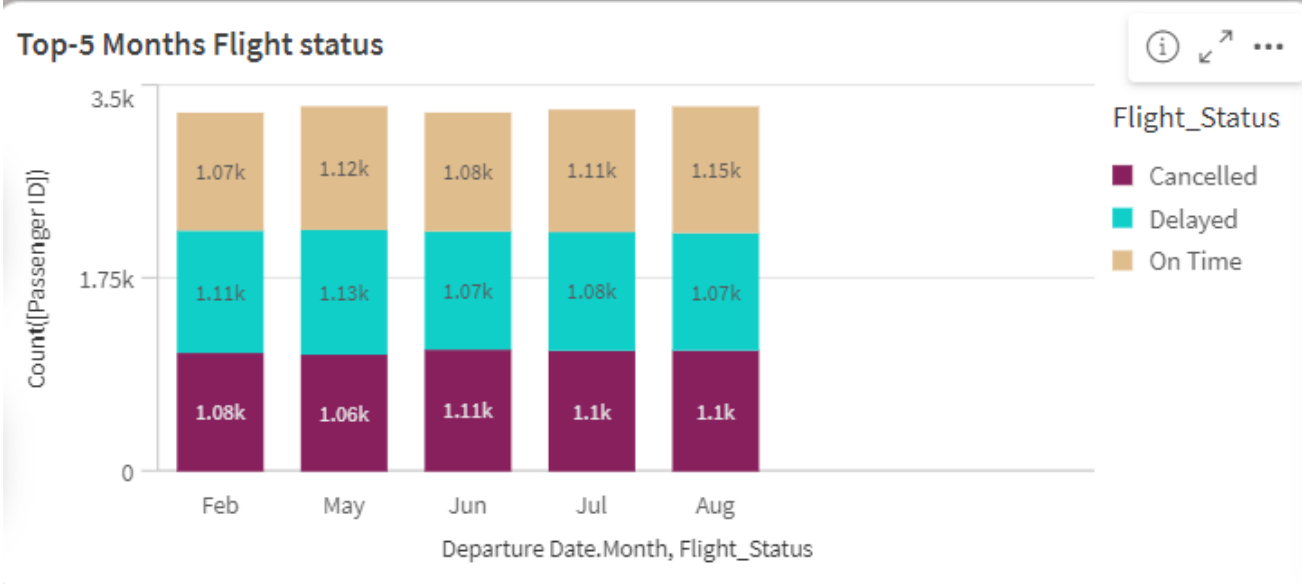
Kpi(Key Performance Indicator):

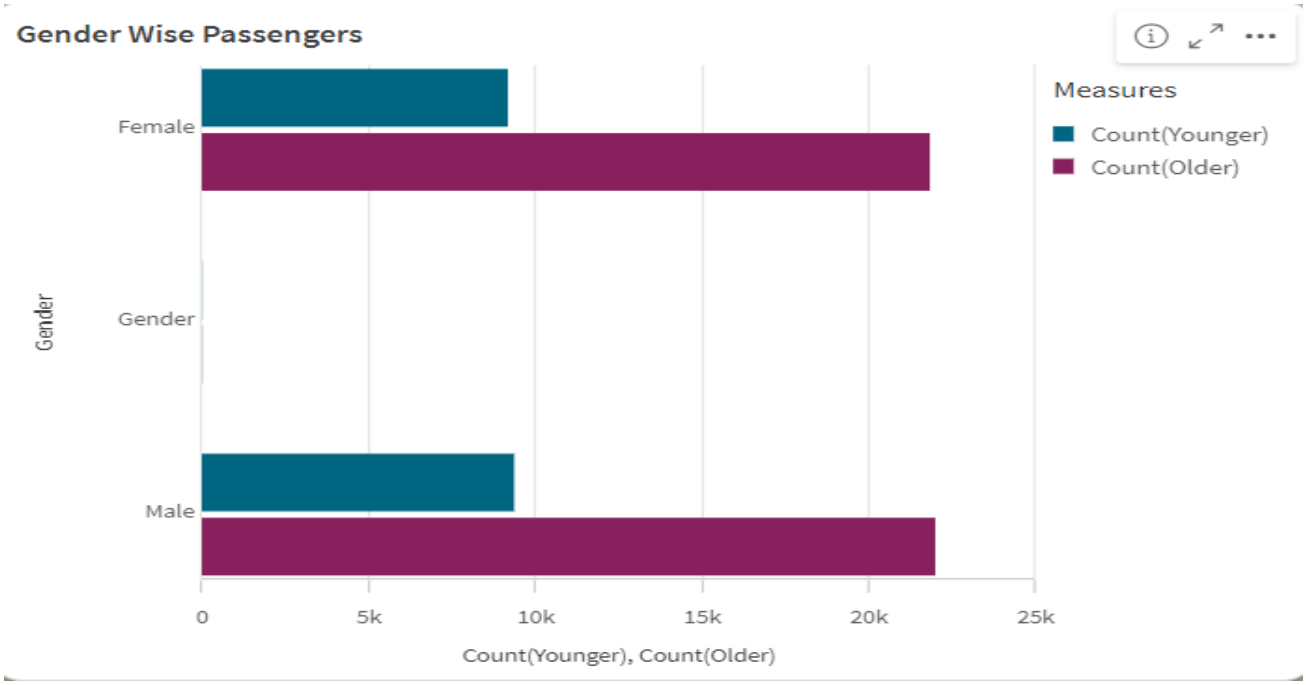
Key Performance Indicators (KPIs) are measurable values that demonstrate how effectively an organization is achieving its key business objectives. Organizations use KPIs to evaluate their success at reaching targets. They can be applied at multiple levels to gauge overall performance or specific processes.

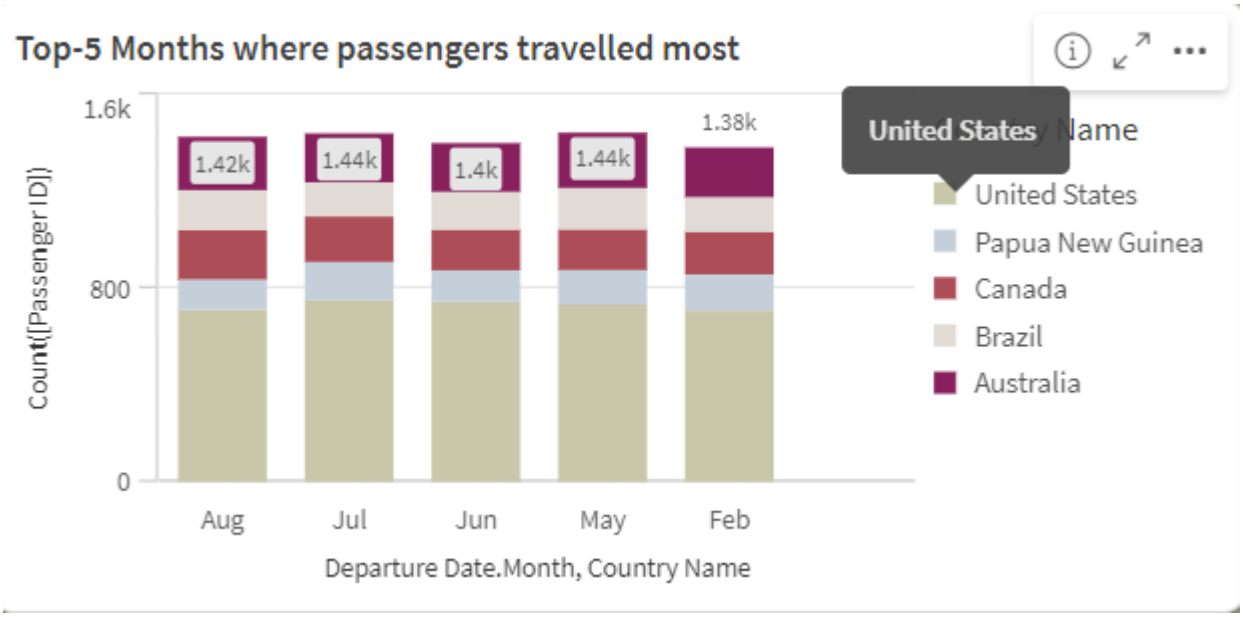
Bar charts:

Bar charts are one of the most common types of data visualization used to compare different categories or groups of data. They are particularly effective for displaying categorical data with numerical values. In a bar chart, each category is represented by a bar, and the length or height of the bar corresponds to the value of the category.



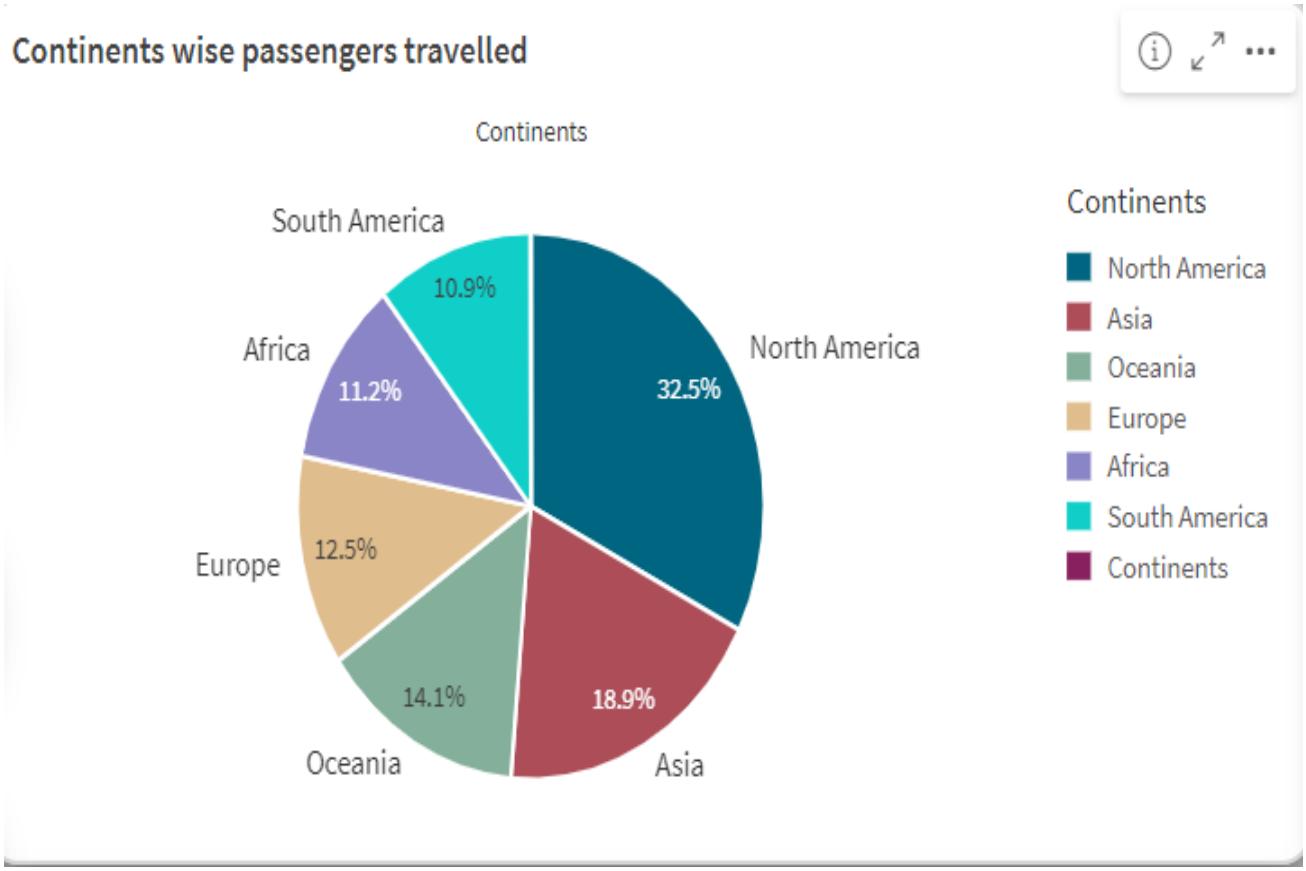

Top-5 Months Flight status
3.5k
1.75k
0
Count([Passenger ID])
1.07k
1.12k
1.08k
1.11k
1.15k
1.11k
1.13k
1.07k
1.08k
1.07k
1.08k
1.06k
1.11k
1.1k
1.1k
Feb
May
Jun
Jul
Aug
Departure Date.Month, Flight_Status
Flight_Status
Cancelled
Delayed
On Time

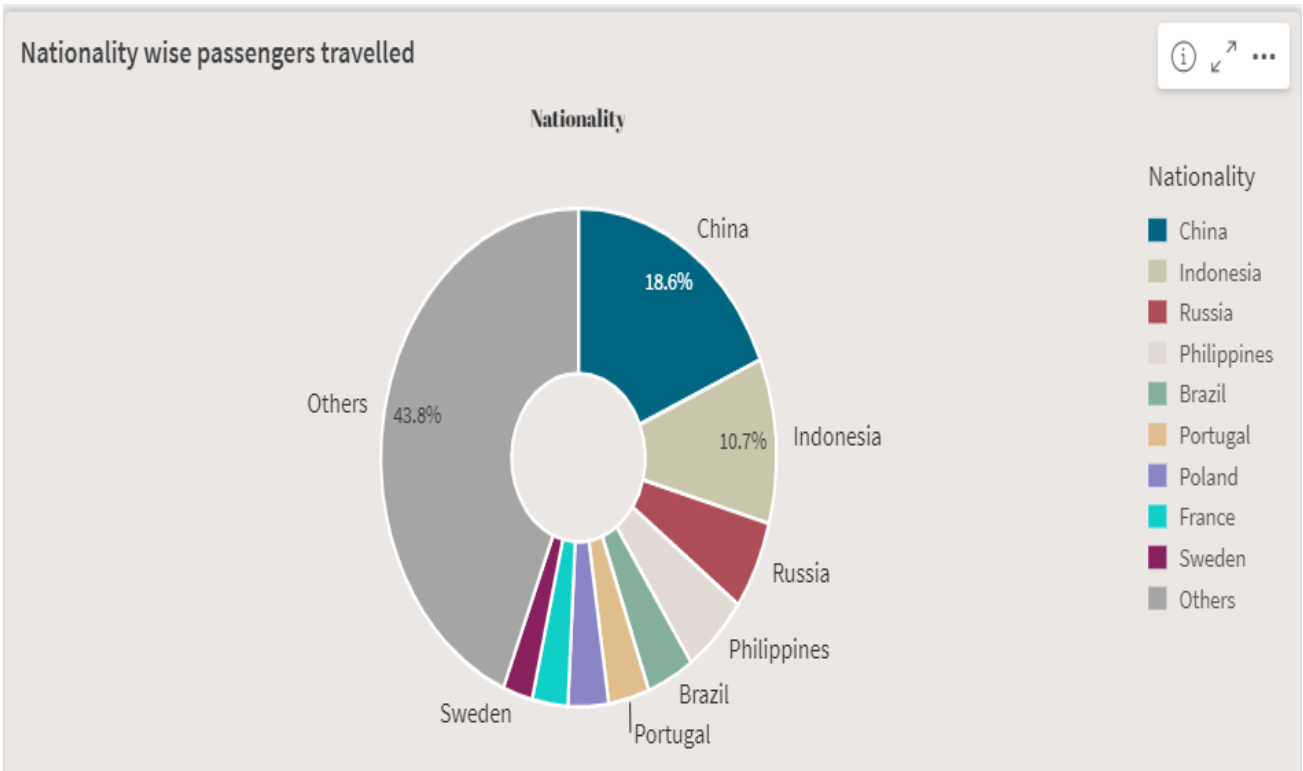

Gender Wise Passengers
Female
Gender
Male
Gender
0
5k
10k
15k
20k
25k
Count(Younger), Count(Older)
Measures
Count(Younger)
Count(Older)


Top-5 Months where passengers travelled most
1.6k
800
0
Count([Passenger ID])
1.42k
1.44k
1.4k
1.44k
1.38k
Aug
Jul
Jun
May
Feb
Departure Date.Month, Country Name
United States
Country Name
United States
Papua New Guinea
Canada
Brazil
Australia

Pie charts:

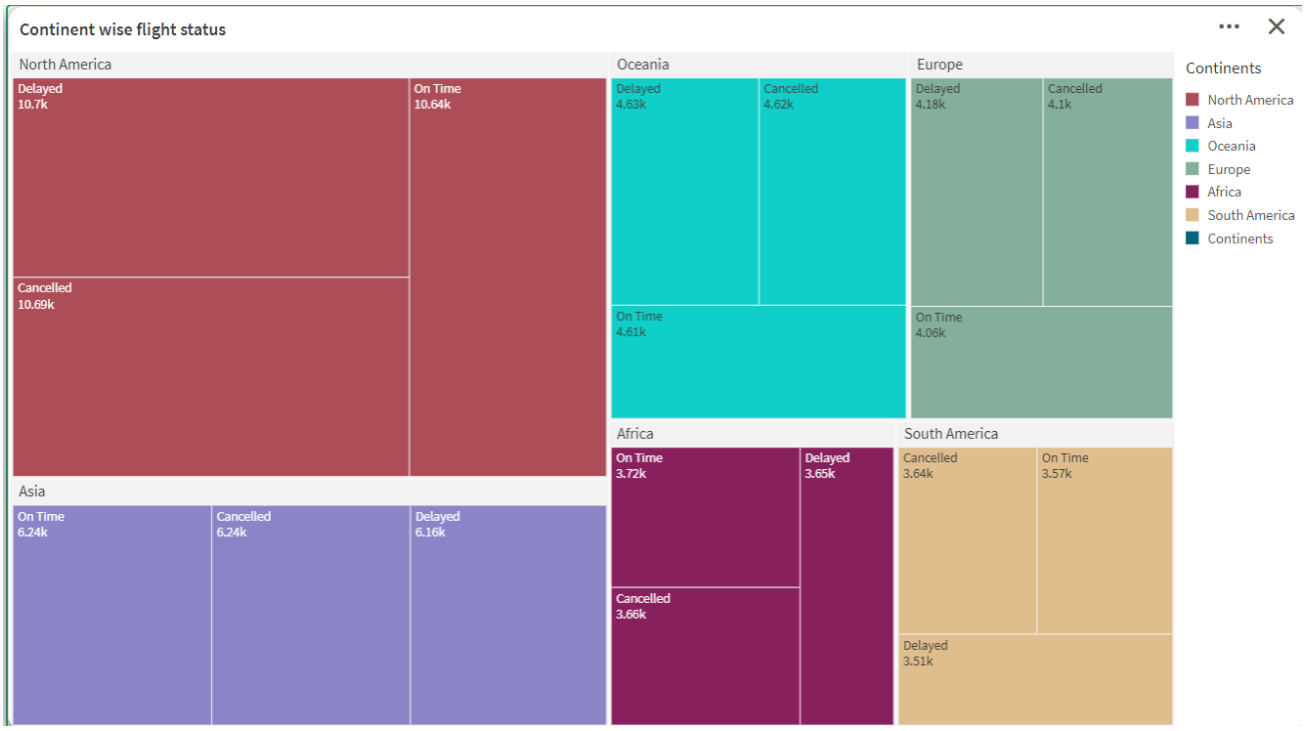
A pie chart is a circular statistical graphic divided into slices to illustrate numerical proportions. Each slice of the pie represents a category, and the size of each slice is proportional to the category's value.





Treemap:

A treemap is a visualization that displays hierarchical data using nested rectangles. Each branch of the hierarchy is represented by a rectangle, which is then tiled with smaller rectangles representing sub-branches. The size and color of each rectangle can represent different dimensions of the data.



Line Charts:

Line charts are a powerful tool for visualizing data points over a period of time. They are particularly effective for showing trends, patterns, and fluctuations, making them ideal for time series data. In a line chart, data points are connected by straight lines, which makes it easy to see changes over time.

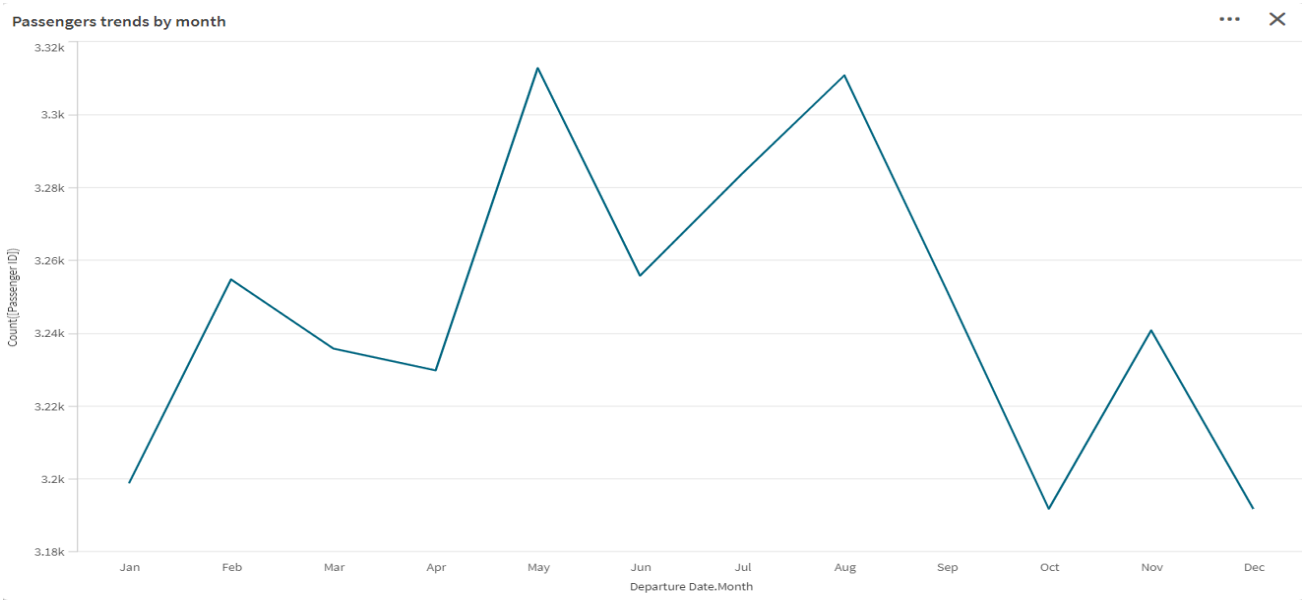


Table:

Table refers to a structured arrangement of data organized in rows and columns. It's a fundamental concept used to store, manipulate, and analyze data. Tables provide a systematic way to present and understand information, making it easier to identify patterns, relationships, and insights within the data.

country wise flight status

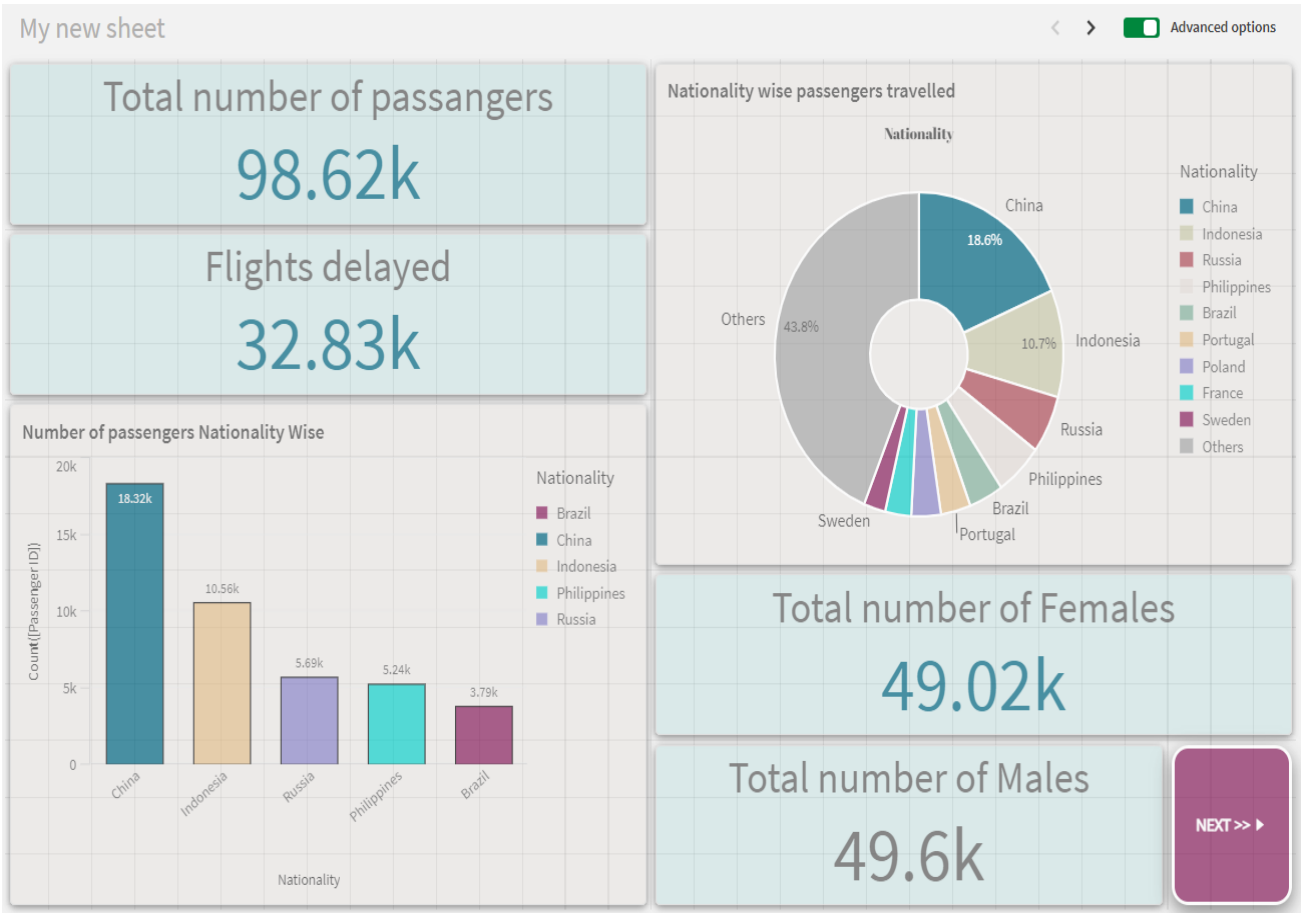
Country Name	count(cancellled flights)	Count(Delayed)
Totals	32942	32831
Afghanistan	137	121
Albania	5	5
Algeria	151	151
American Samoa	21	10
Andorra	4	1
Angola	142	154
Anguilla	7	4
Antigua and Barbuda	7	7
Argentina	394	427
Armenia	7	6
Aruba	4	2
Australia	2131	2142
Austria	37	27
Azerbaijan	30	25
Bahamas	134	134
Bahrain	5	3
Bangladesh	68	55
Barbados	7	3
Belarus	20	26
Belgium	29	20
Belize	78	68
Benin	16	11
Bermuda	9	7
Bhutan	16	19
Bolivia, Plurinational State of	128	144

Data Visualization-Link

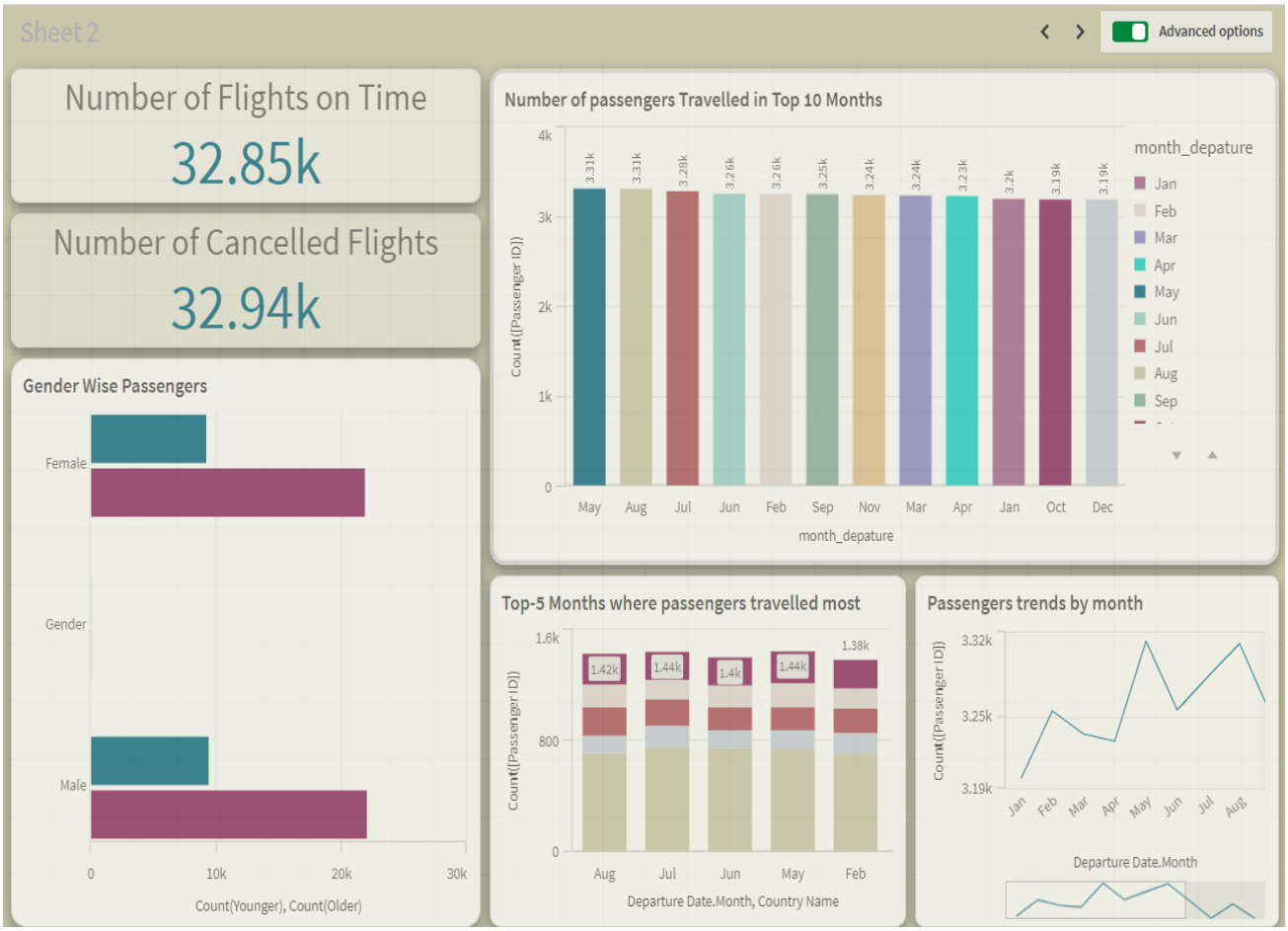
Dashboards:

Dashboard is a visual representation of data insights and analysis, designed to provide users with a comprehensive view of key performance indicators (KPIs), trends, and metrics. Qlik dashboards typically consist of multiple interactive visualizations and charts that allow users to explore and analyze data dynamically.

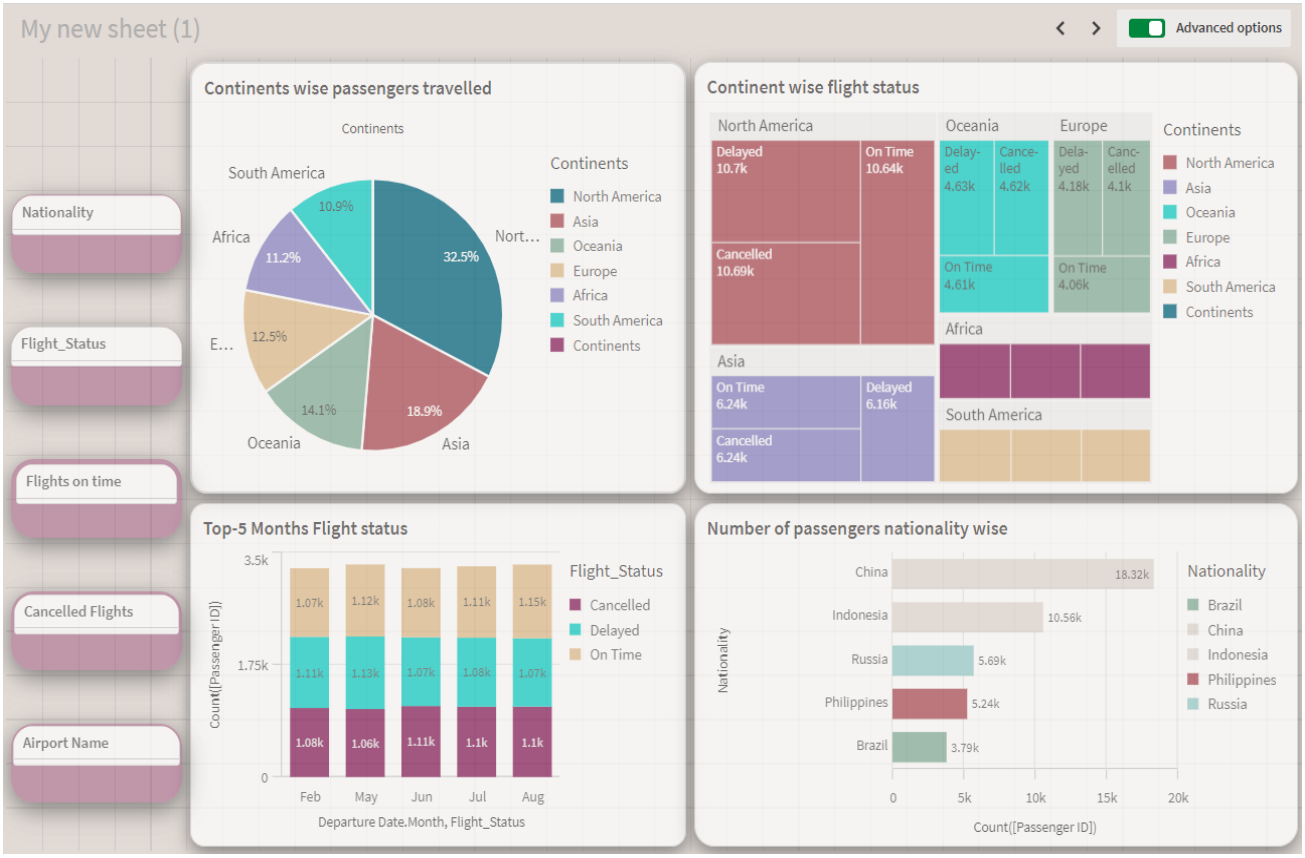
Dashboard-1:



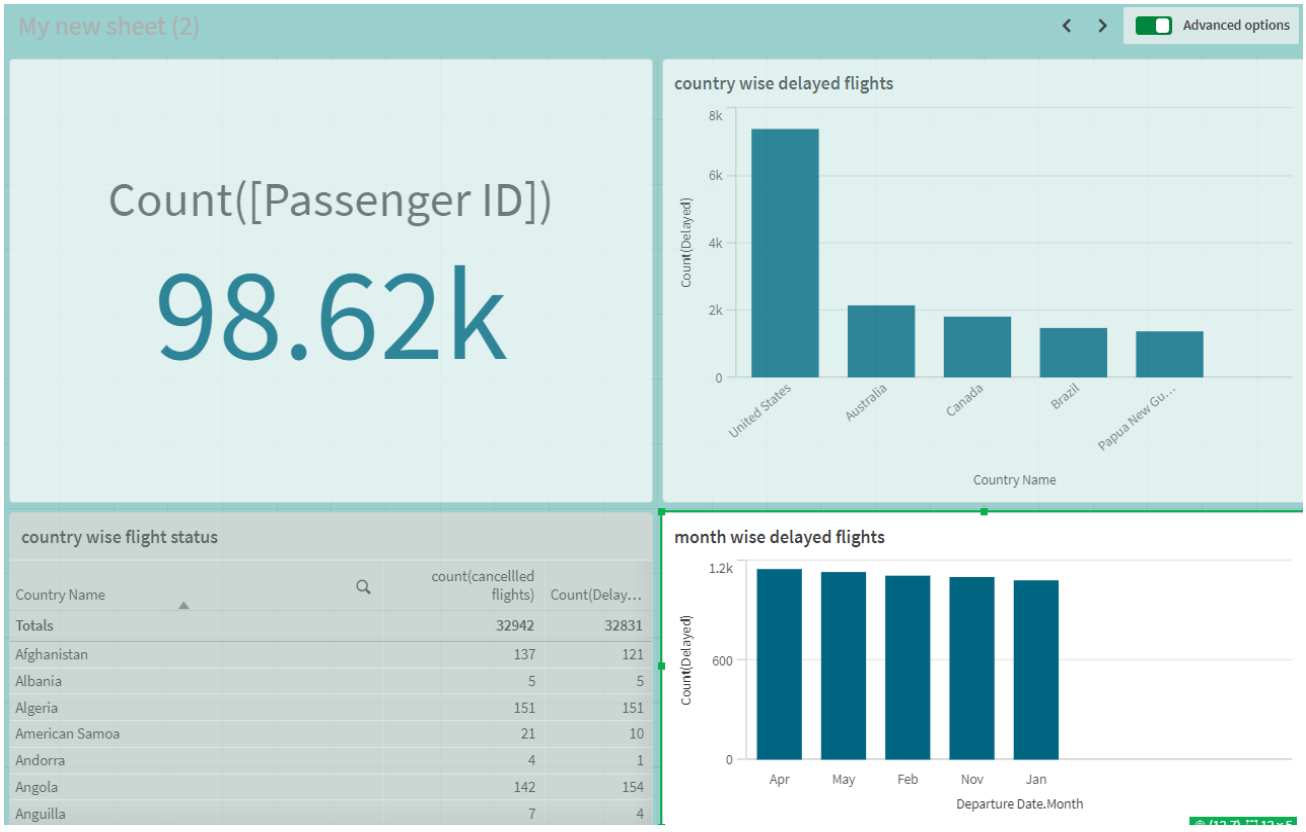
Dashboard-2:



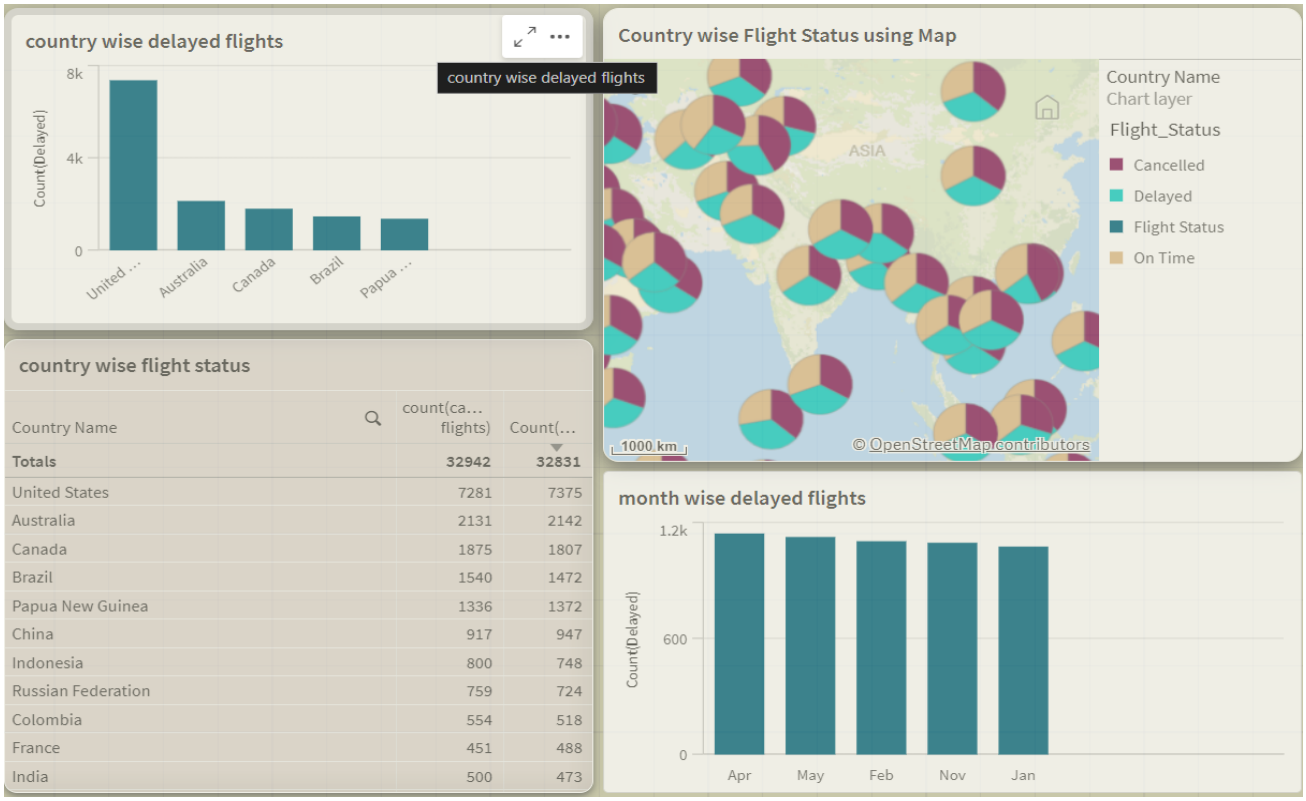
Dashboard-3:



Dashboard-4:



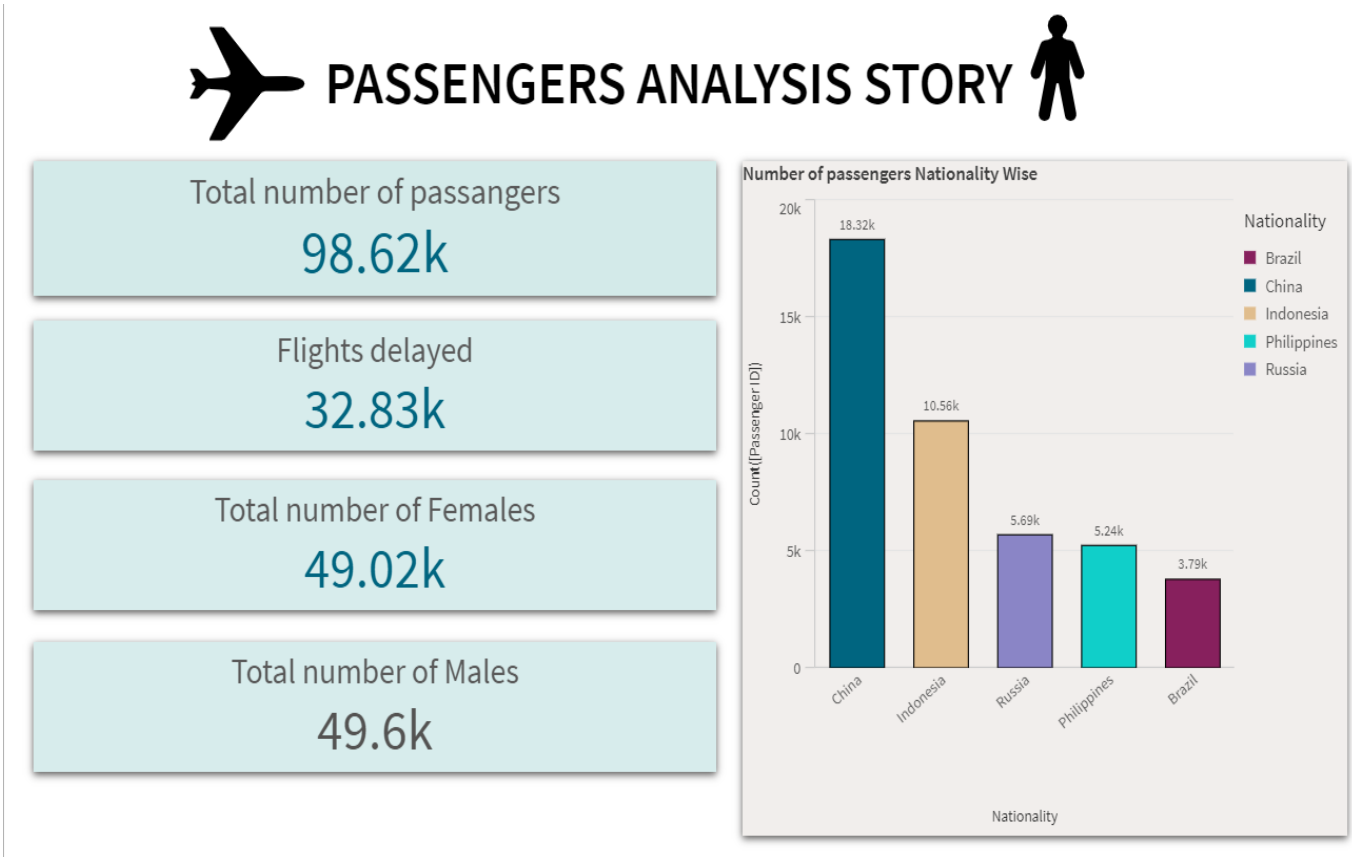
Dashboard-5:

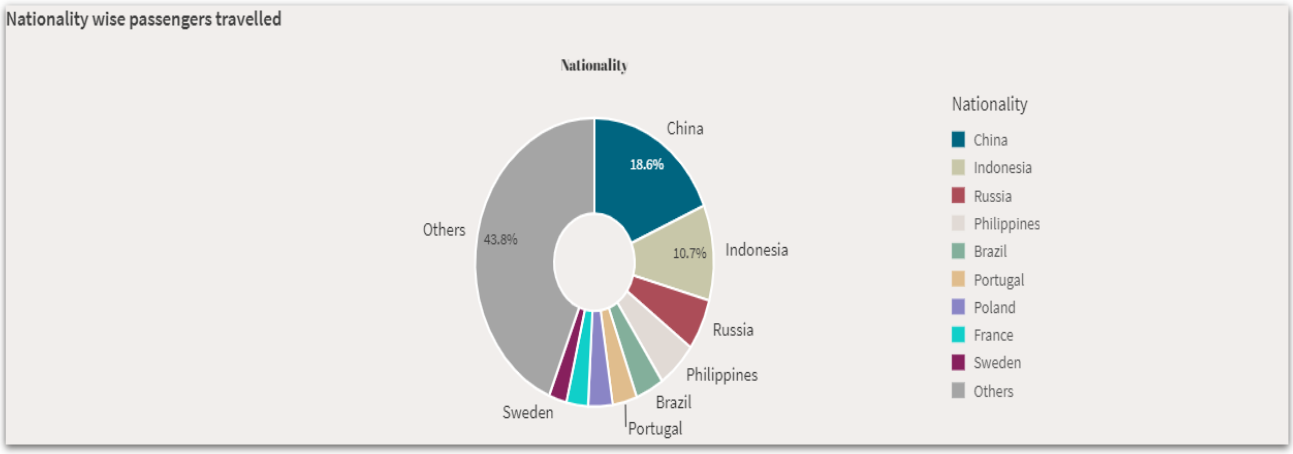


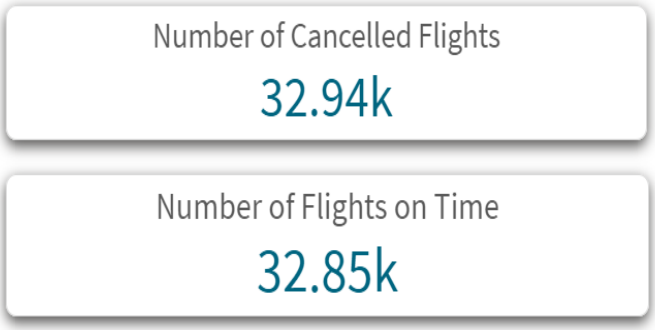
Dashboard Creation:-[Link]

Story:

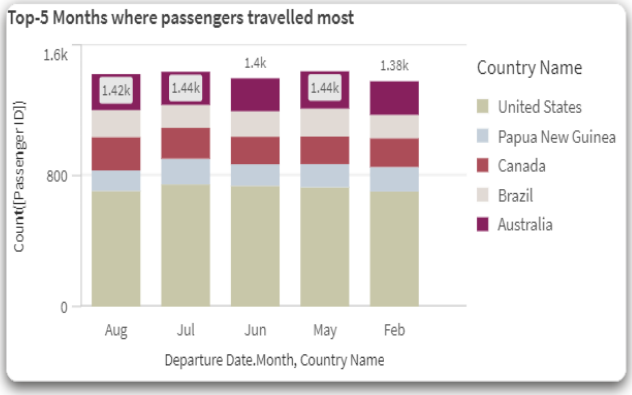
In Qlik, a **story** is a feature that allows users to create interactive presentations or narratives based on data visualizations and insights. Qlik Stories enable users to weave together multiple visualizations, text, images, and other elements into a cohesive storyline, guiding viewers through the data analysis process and communicating key findings and insights effectively.

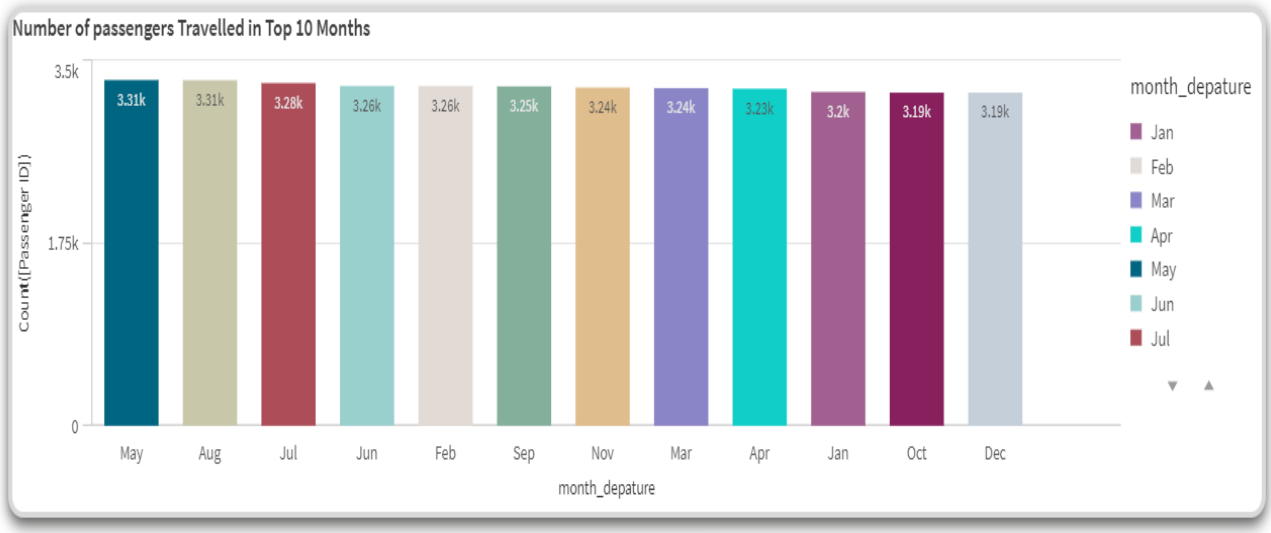


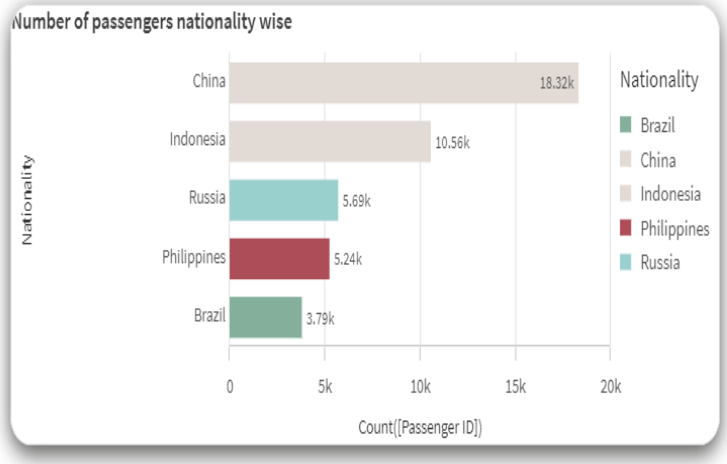

Nationality wise passengers travelled
Nationality
China
18.6%
Indonesia
10.7%
Russia
Philippines
Brazil
Portugal
Sweden
Others
43.8%
Nationality
China
Indonesia
Russia
Philippines
Brazil
Portugal
Poland
France
Sweden
Others

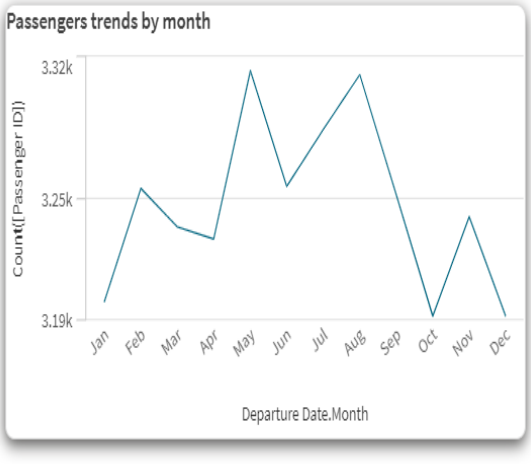

Number of Cancelled Flights
32.94k

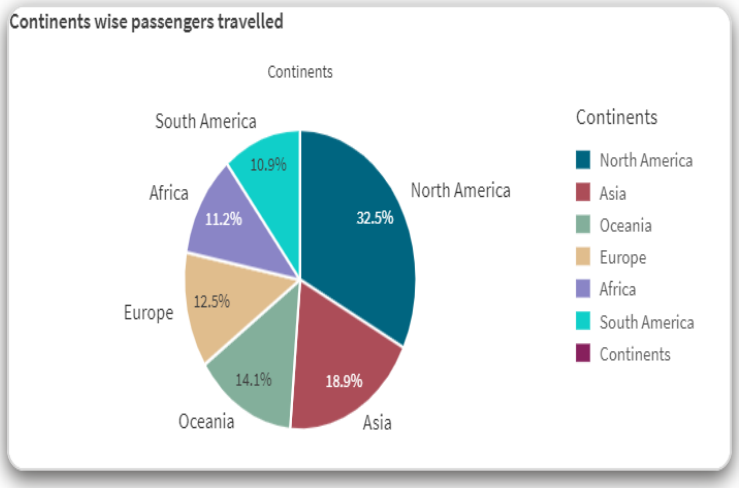
Number of Flights on Time
32.85k

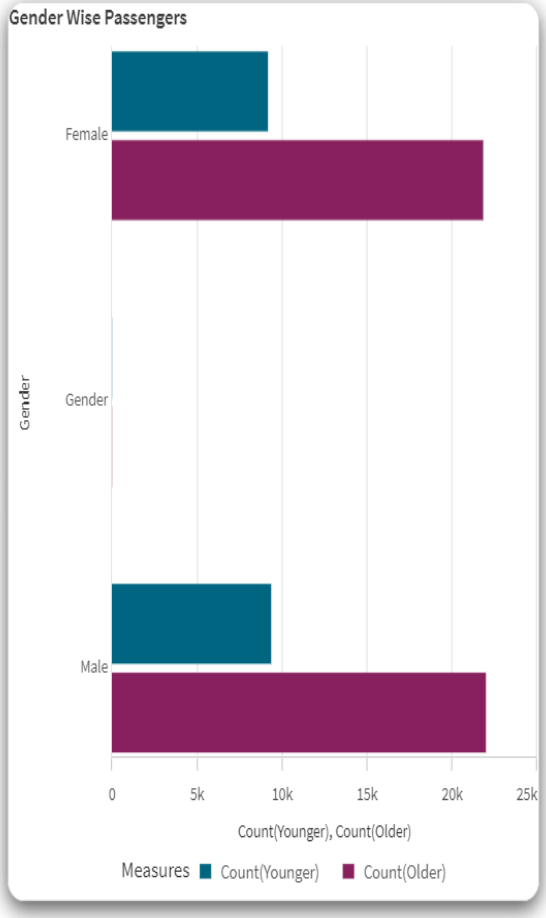

Top-5 Months where passengers travelled most
1.6k
800
0
Count([Passenger ID])
1.42k
1.44k
1.4k
1.44k
1.38k
Aug
Jul
Jun
May
Feb
Departure Date.Month, Country Name
Country Name
United States
Papua New Guinea
Canada
Brazil
Australia

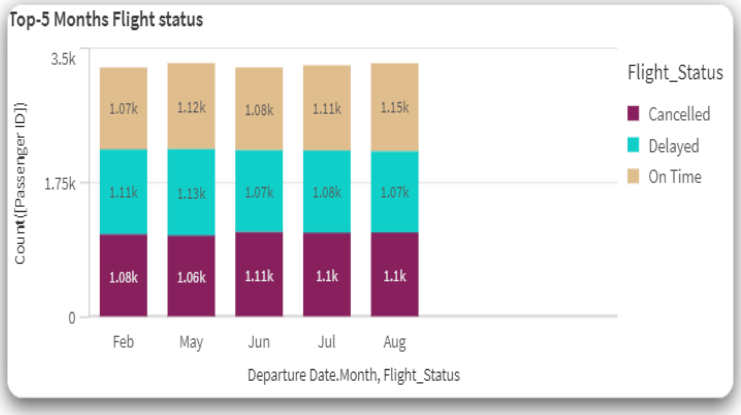

Number of passengers Travelled in Top 10 Months
3.5k
1.75k
0
Count([Passenger ID])
3.31k
3.31k
3.28k
3.26k
3.26k
3.25k
3.24k
3.24k
3.23k
3.2k
3.19k
3.19k
May
Aug
Jul
Jun
Feb
Sep
Nov
Mar
Apr
Jan
Oct
Dec
month_depature
month_depature
Jan
Feb
Mar
Apr
May
Jun
Jul


Number of passengers nationality wise
Nationality
China
18.32k
Indonesia
10.56k
Russia
5.69k
Philippines
5.24k
Brazil
3.79k
0
5k
10k
15k
20k
Count([Passenger ID])
Nationality
Brazil
China
Indonesia
Philippines
Russia


Passengers trends by month
3.32k
3.25k
3.19k
Count([Passenger ID])
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Departure Date.Month


Continents wise passengers travelled
Continents
South America
Africa
North America
Europe
Oceania
Asia
10.9%
11.2%
32.5%
12.5%
14.1%
18.9%
Continents
North America
Asia
Oceania
Europe
Africa
South America
Continents


Gender Wise Passengers
Female
Gender
Male
Gender
0
5k
10k
15k
20k
25k
Count(Younger), Count(Older)
Measures
Count(Younger)
Count(Older)


Top-5 Months Flight status
3.5k
1.75k
0
Count([Passenger ID])
1.07k
1.12k
1.08k
1.11k
1.15k
1.11k
1.13k
1.07k
1.08k
1.07k
1.08k
1.06k
1.11k
1.1k
1.1k
Feb
May
Jun
Jul
Aug
Departure Date.Month, Flight_Status
Flight_Status
Cancelled
Delayed
On Time


Continent wise flight status
North America
Delayed
10.7k
On Time
10.64k
Cancelled
10.69k
Asia
On Time
6.24k
Cancelled
6.24k
Delayed
6.16k
Oceania
Delayed
4.63k
Cancelled
4.62k
On Time
4.61k
Europe
Delayed
4.18k
Cancelled
4.1k
On Time
4.06k
Africa
On Time
3.72k
Delayed
3.65k
Cancelled
3.66k
South America
Cancelled
3.64k
Delayed
3.51k
On Time
3.57k
Continents
North America
Asia
Oceania
Europe
Africa
South America
Continents

Map:

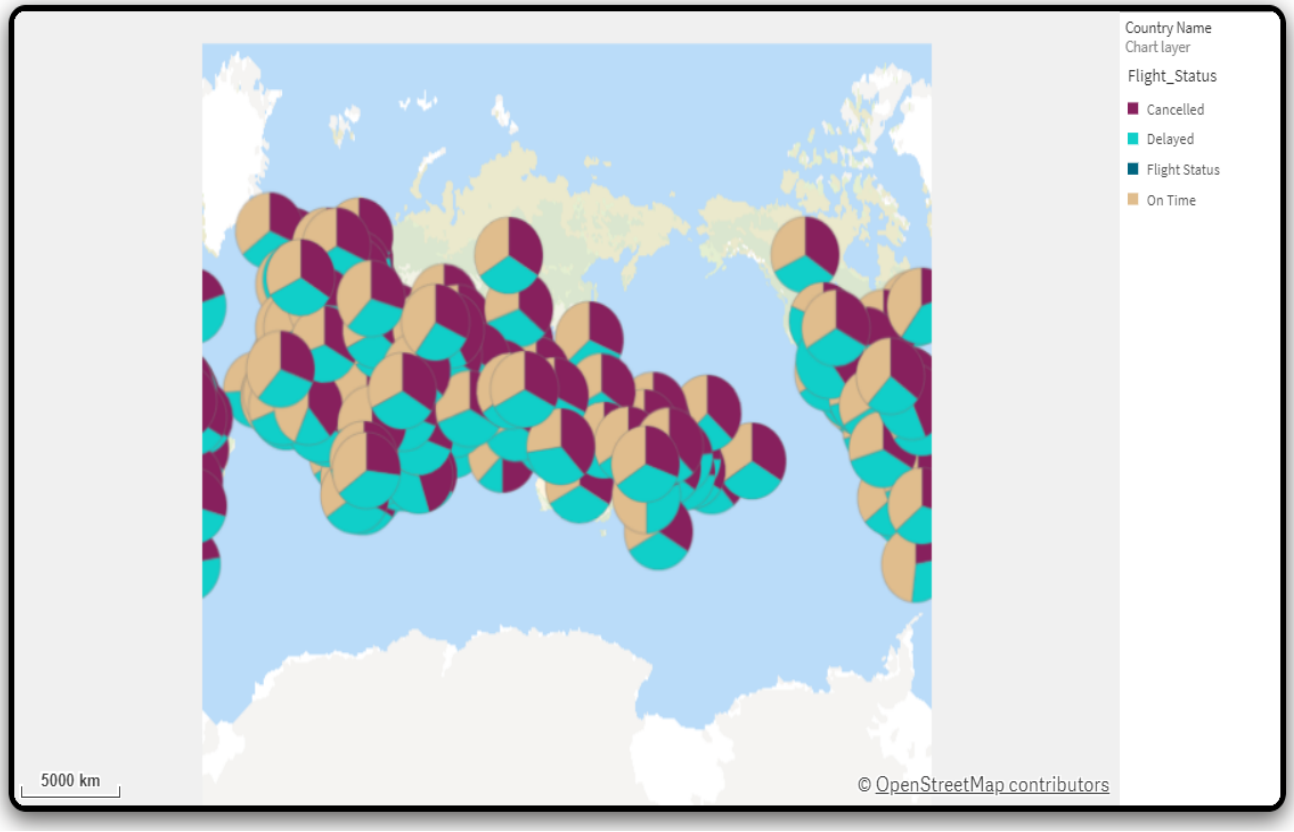


Table:

country wise flight status

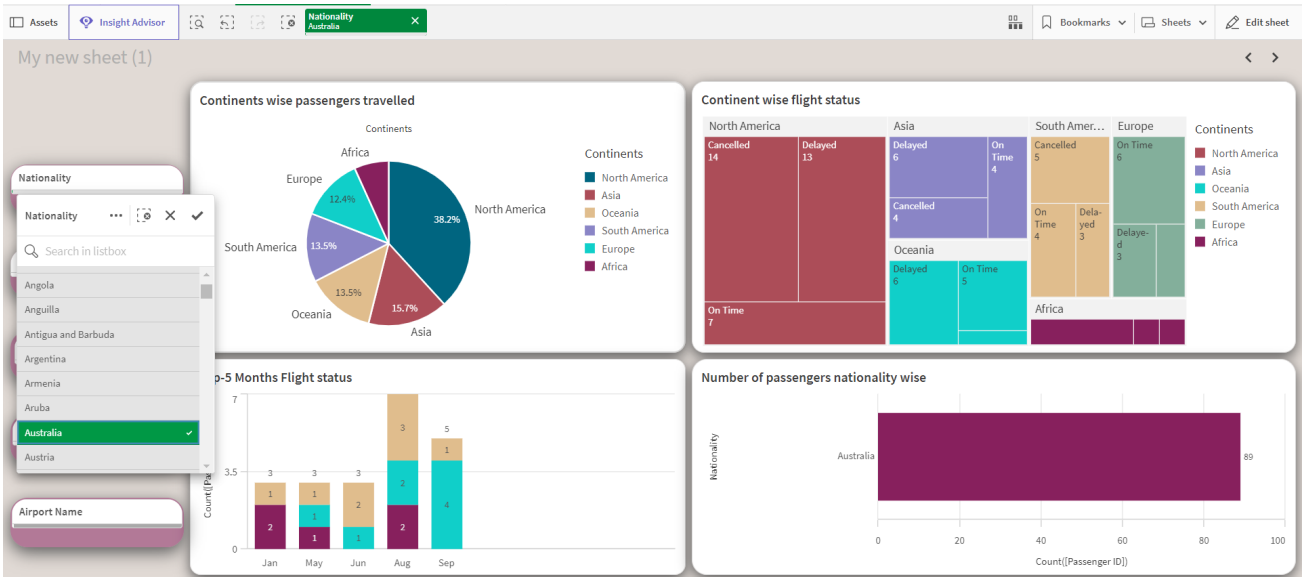
Country Name	count(cancellled flights)	Count(Delayed)
Totals	**32942**	**32831**
United States	7281	7375
Australia	2131	2142
Canada	1875	1807
Brazil	1540	1472
Papua New Guinea	1336	1372
China	917	947
Indonesia	800	748
Russian Federation	759	724
Colombia	554	518
France	451	488
India	500	473
United Kingdom	466	460
Argentina	394	427
Germany	411	385
Mexico	359	353
Japan	316	321
South Africa	305	313
Iran, Islamic Republic of	267	280
Philippines	250	272
Pakistan	242	252
Italy	206	250
Congo, The Democratic Republic of the	210	238
Malaysia	206	235
Turkey	254	221
Venezuela, Bolivarian Republic of	258	221

Performance Testing

Utilization of Data Filters:

Utilizing filters in data visualization involves implementing and applying various types of filters to improve the clarity, interactivity, and analytical power of visual representations.

1.country wise flight status



2.Total number of flights on time