Exploring Insights From Synthetic

Airline Data Analysis With Qlik 1

# INTRODUCTION

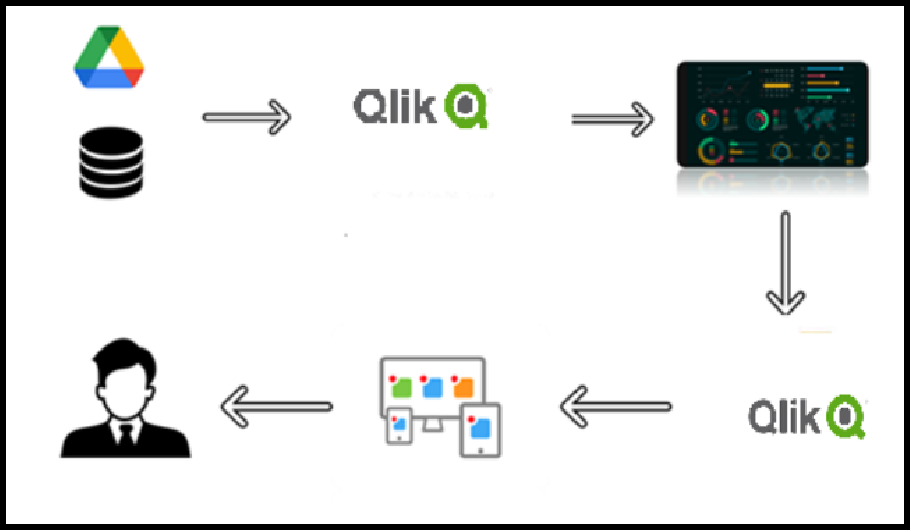
### OVERVIEW:

This document provides a detailed guide on "Exploring insights from synthetic airline data using Qlik Sense". The project involves analyzing various aspects of airline operations, performance, and customer satisfaction through interactive visualizations.

### PURPOSE:

The goal of this project is to create a comprehensive dashboard in Qlik Sense that provides insights into the airline industry's performance using synthetic data. In this project, the synthetic airline data simulates various aspects of airline operations, including ﬂight schedules, passenger demographics, ticket sales, and performance metrics. 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. **TECHNICAL ARCHITECHTURE**: A diagram showing the major components and their interactions.Components are
     + Data Sources
     + Data Ingestion and Storage
     + Qlik Sense Environment
     + User Interface



# PROBLEM UNDERSTANDING

### SPECIFY THE BUSINESS PROBLEM:

The key business problems are

* + - **Flight Performance Optimization**:Frequent ﬂight delays and cancellations negatively impact customer satisfaction and airline reputation.
    - **Passenger Analysis:**Lack of detailed understanding of passenger travel trends and preferences.
* **Revenue Management**:Ineﬃcient revenue management leading to potential revenue losses.
* **Operational Eﬃciency:**Ineﬃciencies in ﬂight operations, such as prolonged turnaround times and resource allocation issues.
* **Customer Satisfaction and Experience** : Insuﬃcient insights into factors affecting customer satisfaction and loyalty.

#### BUSINESS REQUIREMENTS:

* Data Integration and Preparation
* Performance Analysis
* Operational Eﬃciency
* Customer Experience
* Interactive Dashboards and Reports
* Predictive Analysis

By focusing on these requirements, the project aims to provide comprehensive insights into airline operation, performance, and customer experience.

#### LITERATURE SURVEY:

The use of synthetic airline data for analysis offers a versatile and powerful tool for exploring various aspects of airline operations, passenger behavior, and ﬁnancial performance. Continued advancements in data generation and analysis techniques will further enhance the ability to derive actionable insights from synthetic data, beneﬁting both researchers and industry practitioners.

# DATA COLLECTION

#### COLLECT THE DATA:

Source Data: Obtain a synthetic airline dataset. Ensure it includes relevant ﬁelds such as Date, Airline, PassengerCount, FlightDelay,Destination, etc.

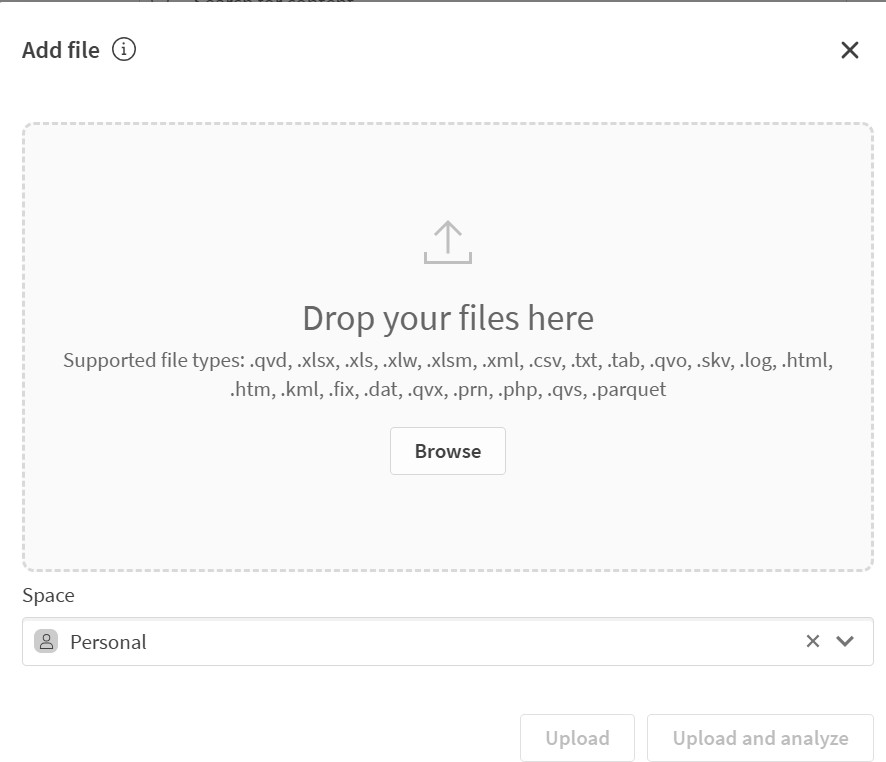
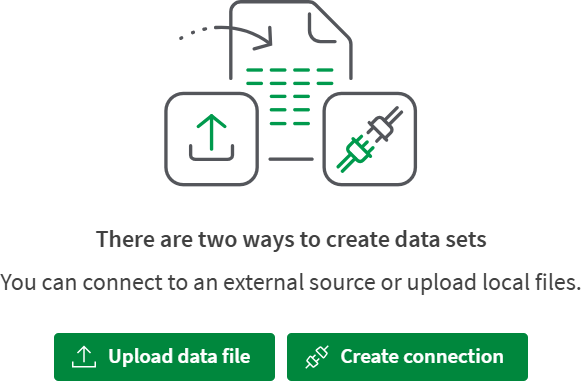
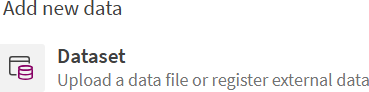
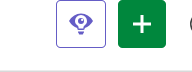
Format: The dataset can be in formats like CSV, Excel, or a database table.



The dataset prominently incorporates ﬁelds such as Passenger ID, First Name, Last Name, Gender, Age, Nationality, Airport Name, Airport Country Code, Country Name, Airport Continent, Continents, Departure Date, Arrival Airport, Pilot Name, and Flight Status.

#### CONNECT DATA WITH QLIK SENSE:

After collecting the dataset ,Open Qlik Sense and create a new app.Click + symbol ont he top right corner.Click "add dataset".Click "Upload data ﬁle".click on "Browse" and add the required dataseet from your device.Then click on "upload and analyze".Then the new app will be created to create your visualisations.



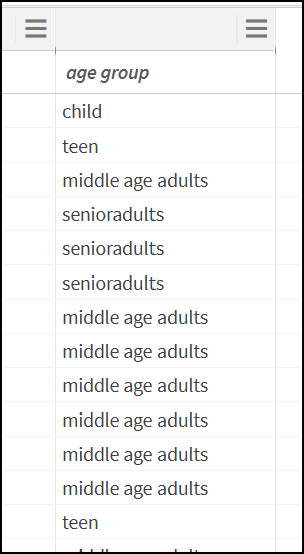
1. **DATA PREPARATION**

#### PREPARING 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.

In the app,go to data manager and then select data load editor and embed the code required to create new columns in the dataset (or) go to the edit option in data manager and select "Add ﬁeld" and add "calculated ﬁeld" and select the name for the coloumn and give the expression on which you want to calculate the data of the column. From the dataset,I created "age group" column with the expression : if(Age<4,'toodler',

if(Age<12,'child', if(Age<19,'teen', if(Age<30,'adults', if(Age<=59,'middle age adults', if(Age>60,'senioradults'))))))

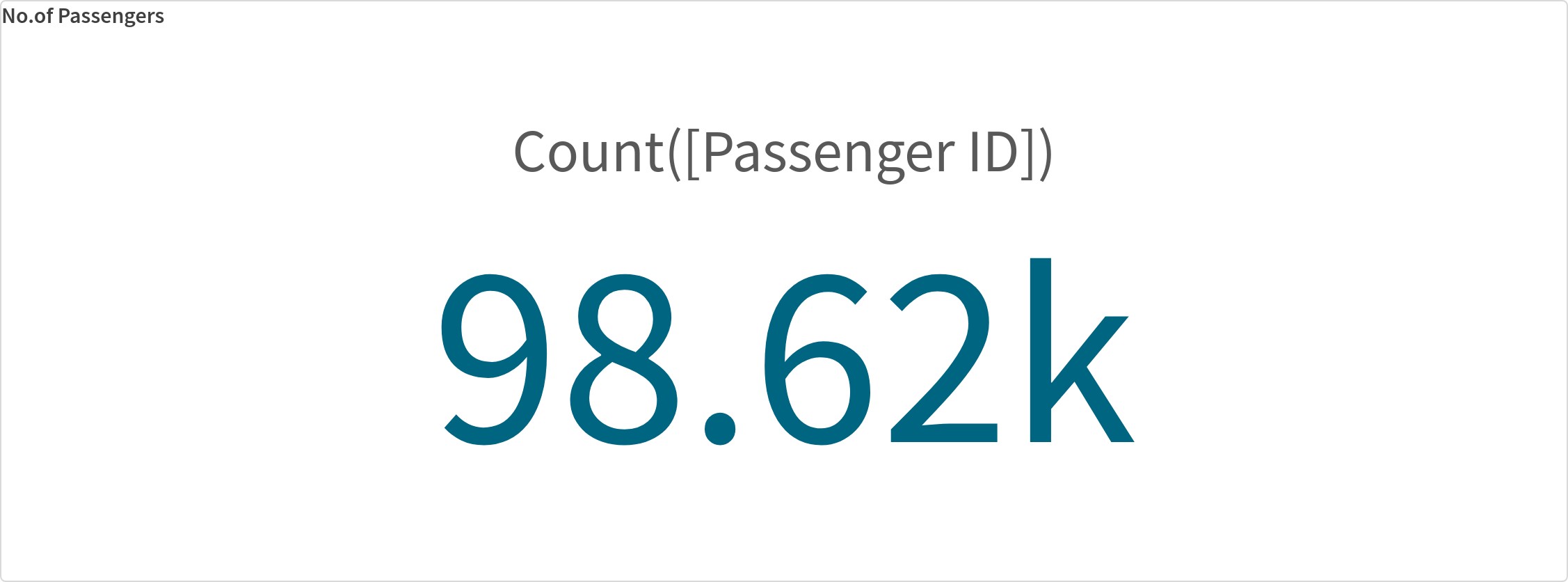
This created a new column in the dataset.

**4.1 Creation of age group column for data preparation**

## DATA VISUALISATIONS

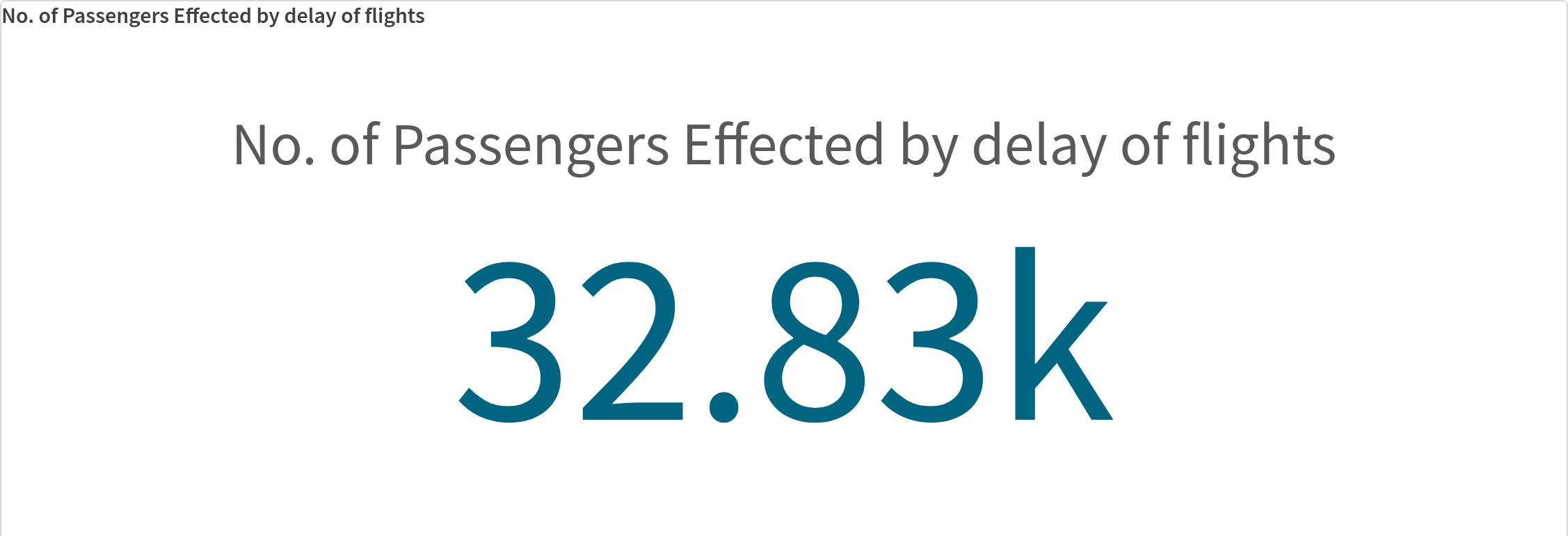
#### VISUALISATIONS:

Data visualization refers to the representation of data in graphical formats such as charts, graphs, dashboards, or infographics. This makes it easier for you to see trends, recognize relationships, and uncover data-driven insights from large, complex data sets. These insights can increase eﬃciency, revenue, and proﬁts for your organization.

ACTIVITY 1: TOTAL NO.OF PASSENGERS USING KPI

#### ACTIVITY 2: NUMBER OF PASSENGERS EFFECTED BY CANCELLED FLIGHTS USING KPI

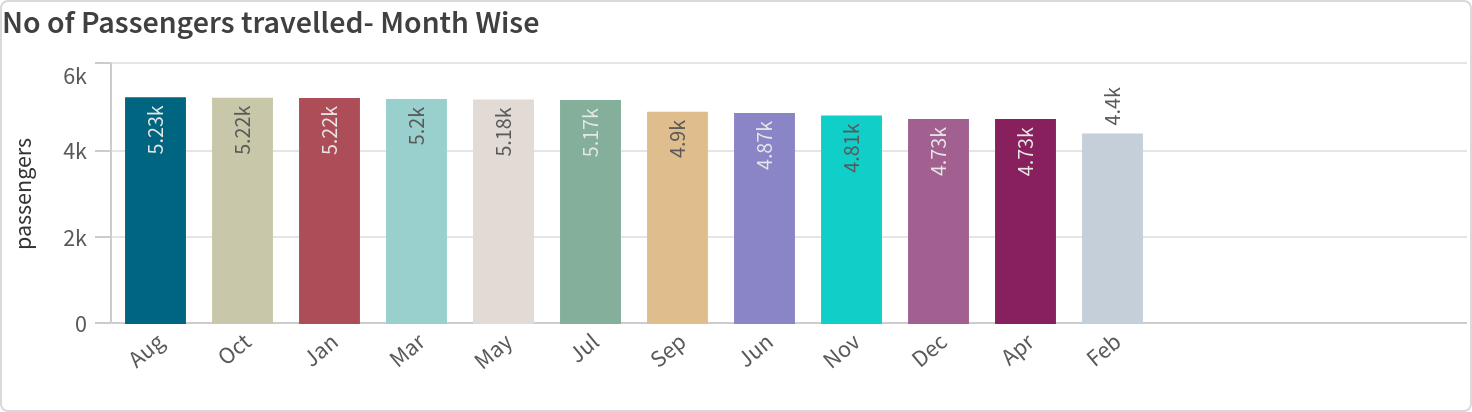


**ACTIVITY 3: NUMBER OF PASSENGERS EFFECTED BY DELAY OF FLIGHT USING KPI**

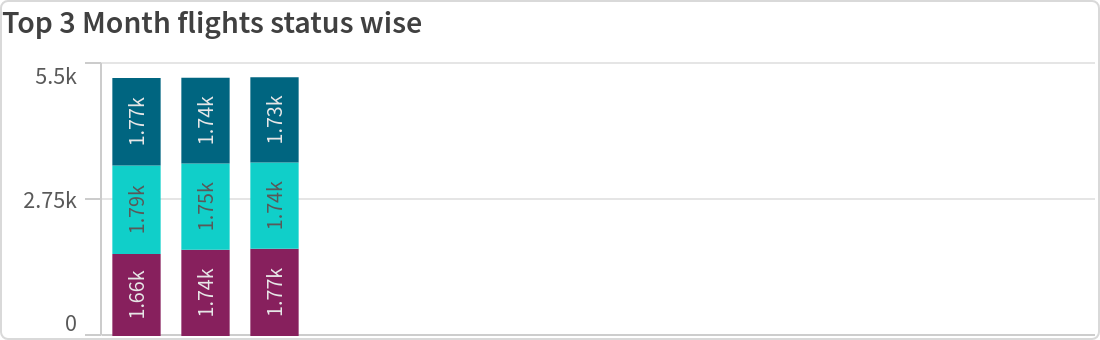
#### ACTIVITY 4: NUMBER OF FLIGHTS ON TIME USING KPI



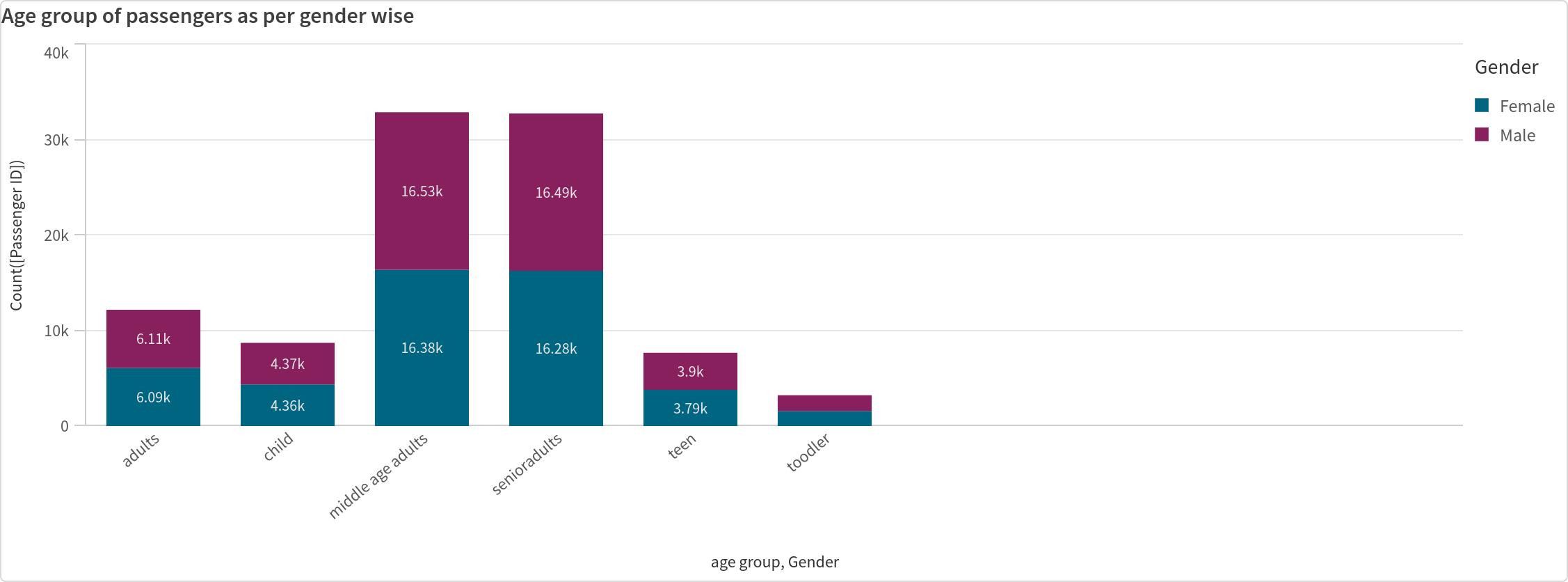
**ACTIVITY 5: NUMBER OF PASSENGERS TRAVELLED MONTH WISE USING BAR CHART**



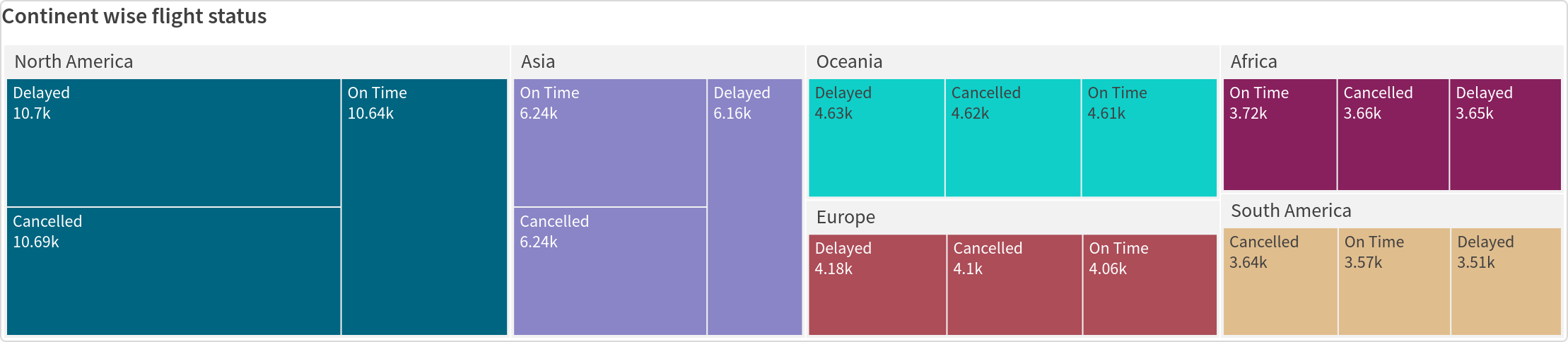
#### ACTVITY 6: TOP 3 MONTHS FLIGHT STATUS WISE USING STACK BAR CHART



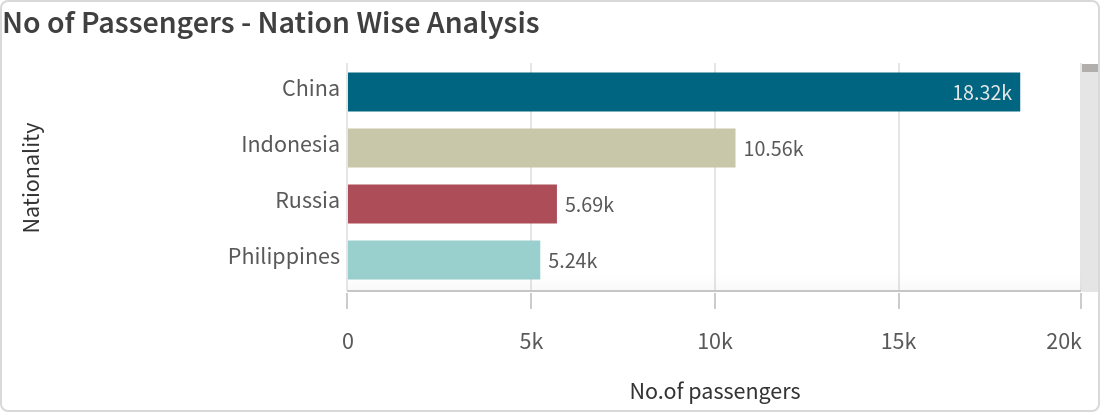
**ACTIVITY 7: AGE GROUP OF PASSENGERS AS PER GENDER WISE USING BAR CHART STACK**



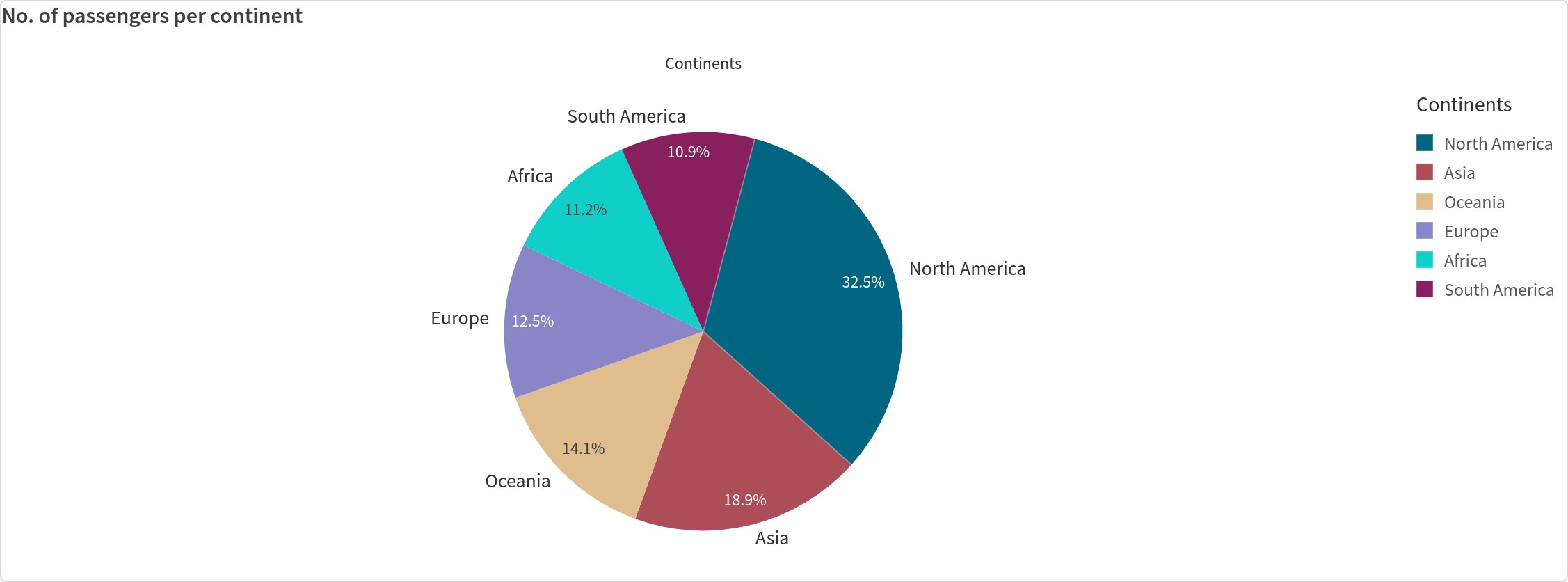
#### ACTIVITY 8: CONTINENT WISE FLIGHT STATUS USING TREE MAP



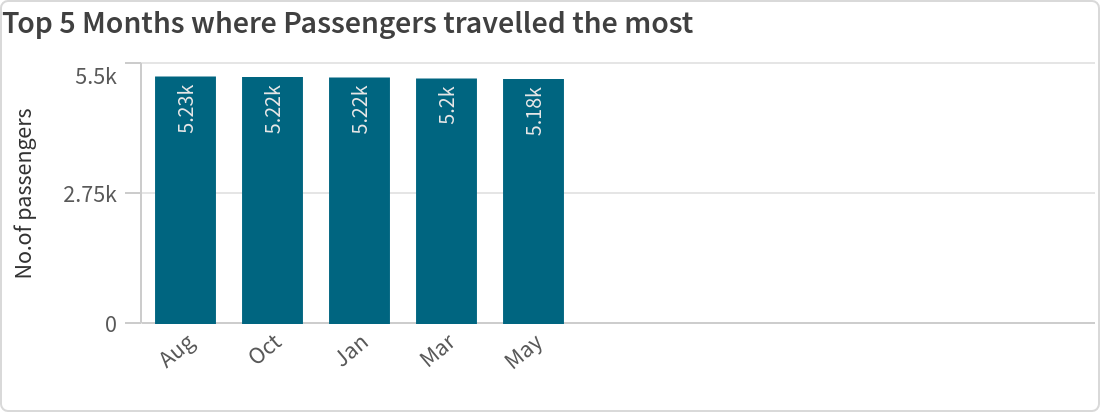
**ACTIVITY 9: NUMBER OF PASSENGERS \_ NATION WISE ANALYSIS USING HORIZANTAL BAR CHART**



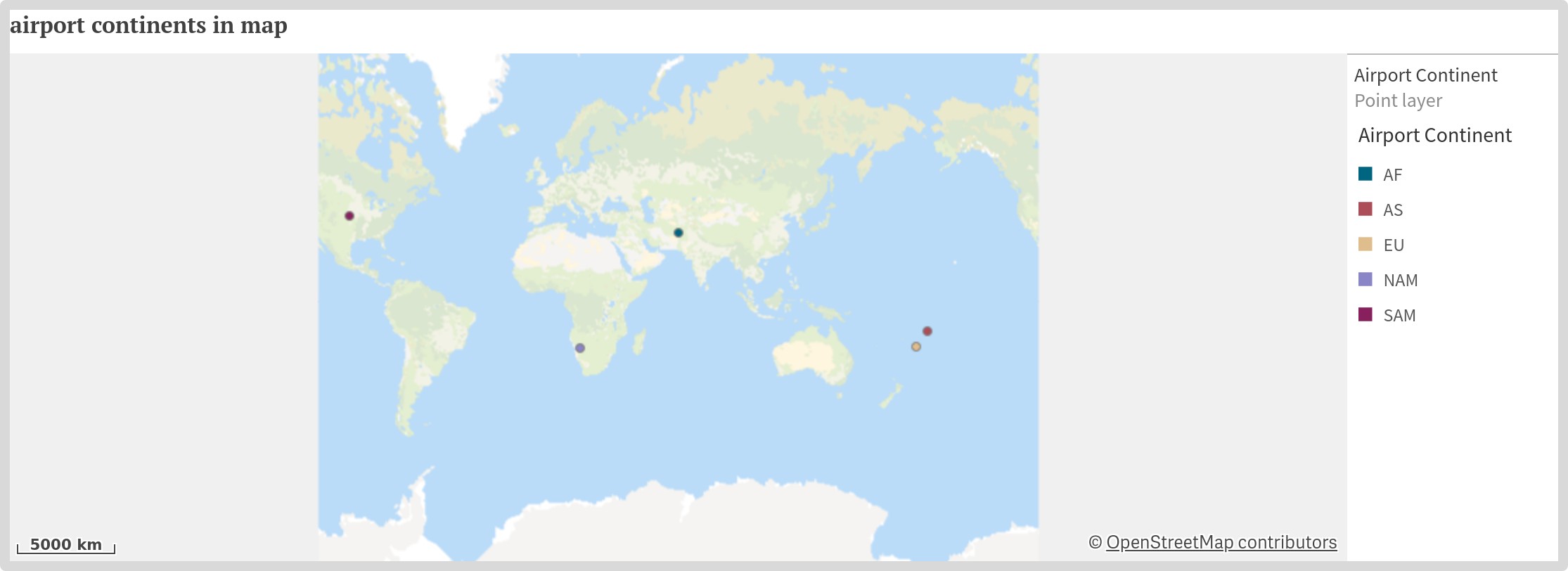
#### ACTIVITY 10: NUMBER OF PASSENGERS PER CONTINENT USING PIE CHART



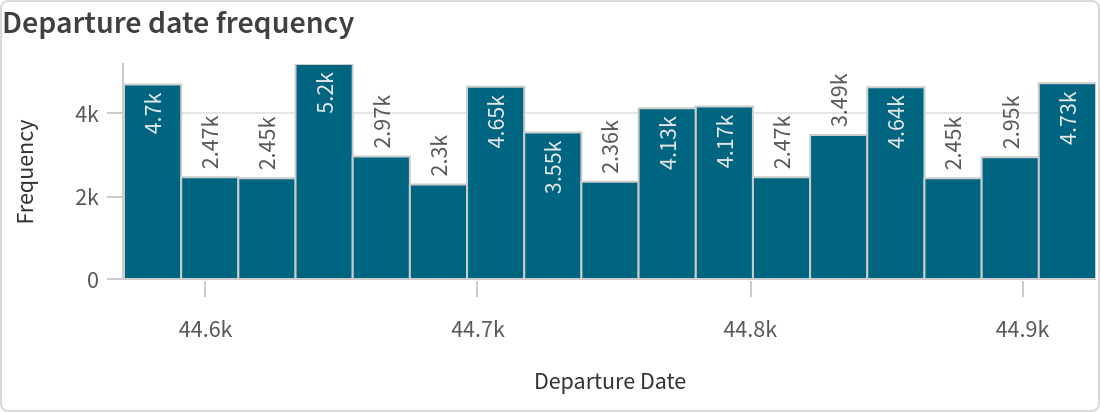
**ACTIVITY 11: TOP 5 MONTHS WHERE PASSENGERS TRAVELLED THE MOST USING BAR CHART**



#### ACTIVITY 12: AIRPORT CONTINENTS IN MAP

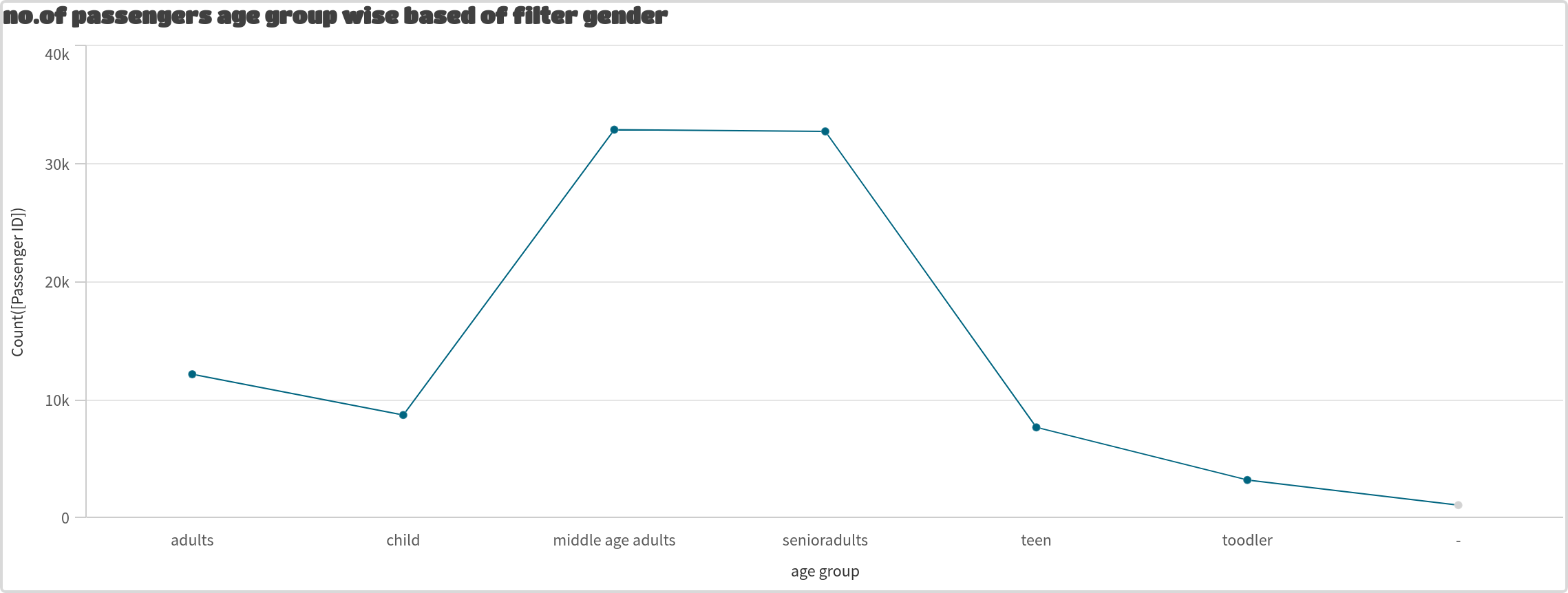


**ACTIVITY 13: DEPARTURE DATE FREQUENCY USING HISTOGRAM**

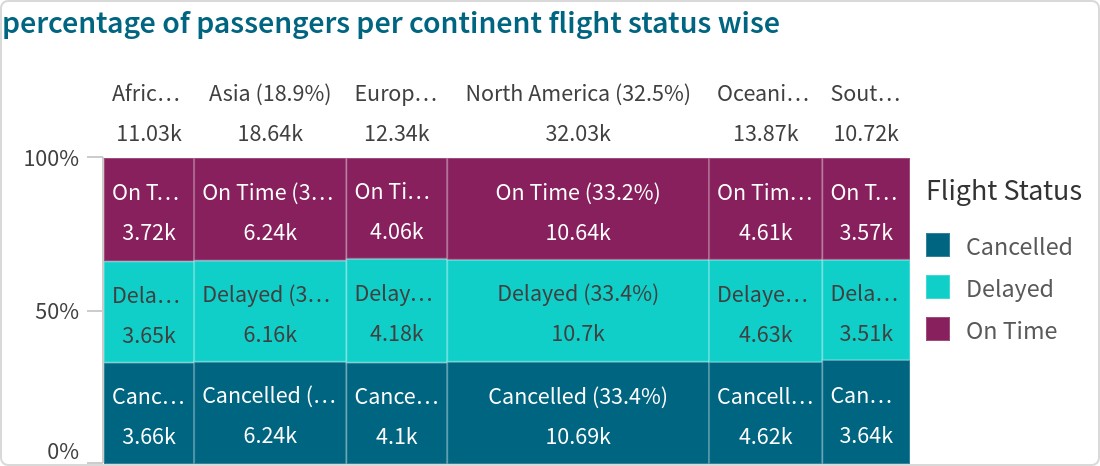


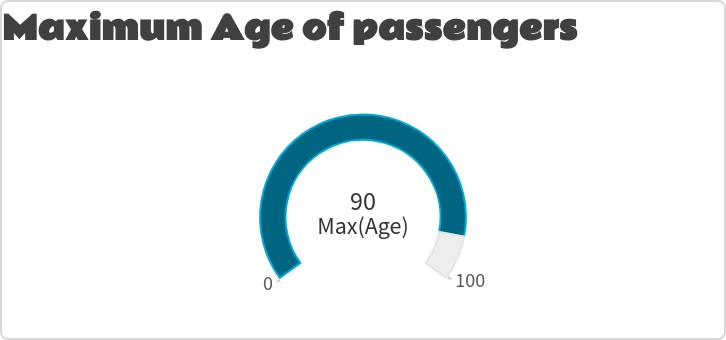
#### ACTIVITY 14: NUMBER OF PASSENGERS AS PER AGE GROUP USING PIE CHART

**ACTIVITY 15: NO.OF PASSENGERS AS PER GENDER USING LINE CHART**

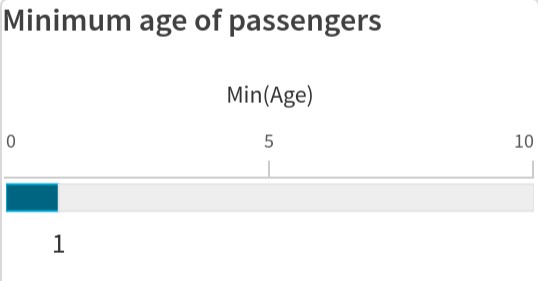


#### ACTIVITY 16: PERCENTAGE OS PASSENGERS PER CONTINENT-FLIGHT STATUS WISE USING TREE MAP

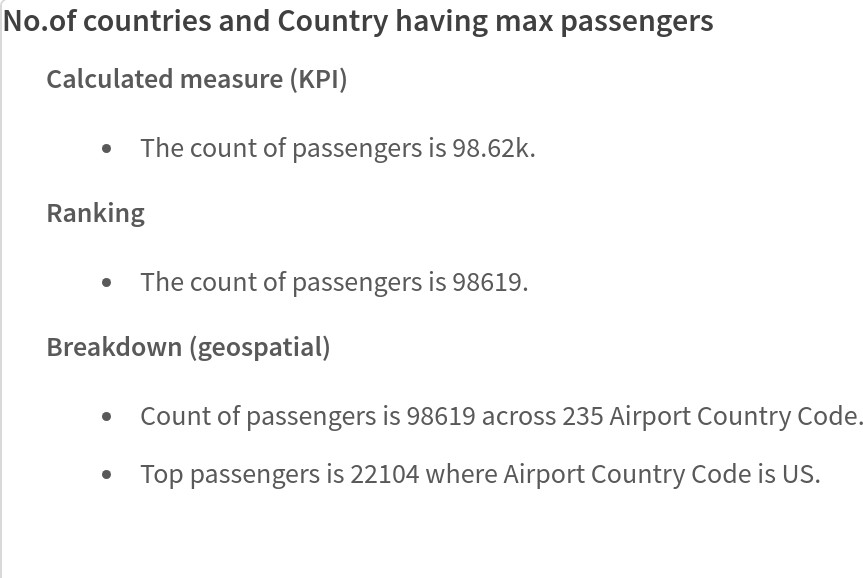


**ACTIVITY 17: MAXIMUM AGE AMONG PASSENGERS USING GAUGE RADIAL CHART**

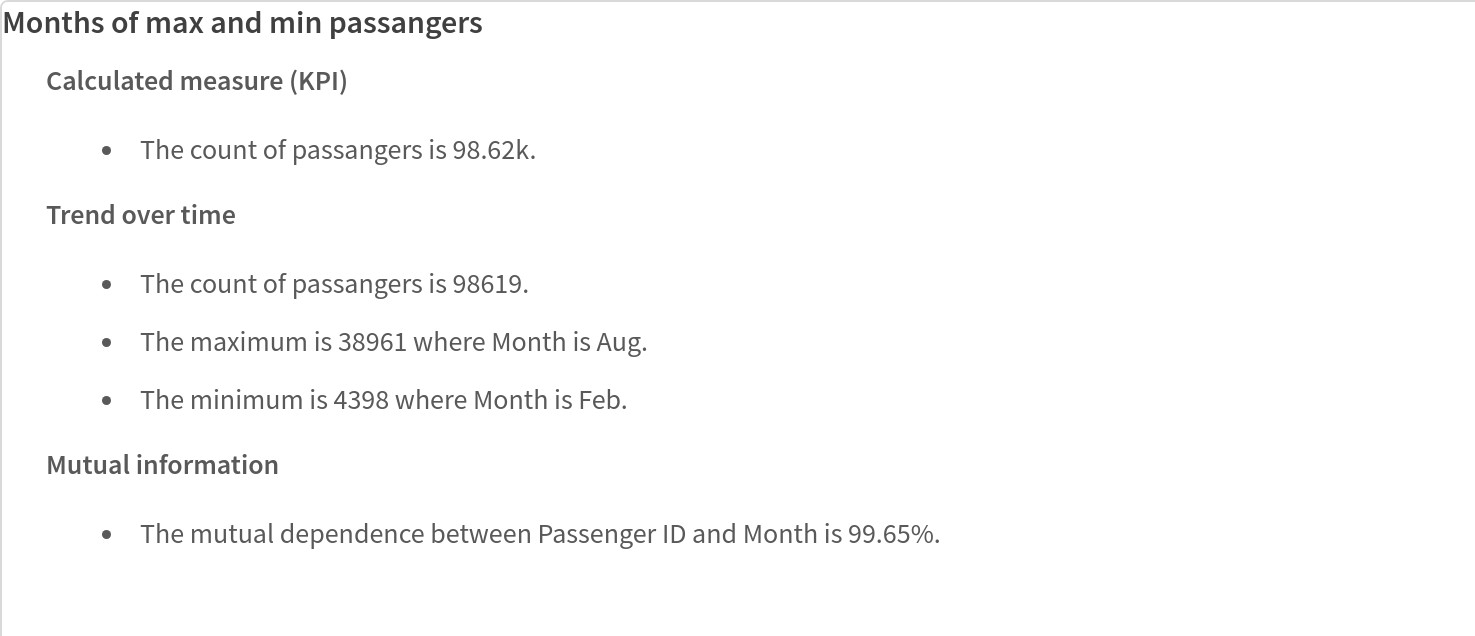
#### ACTIVITY 18: MINIMUM AGE AMONG PASSANGERS USING GAUGE BAR CHART



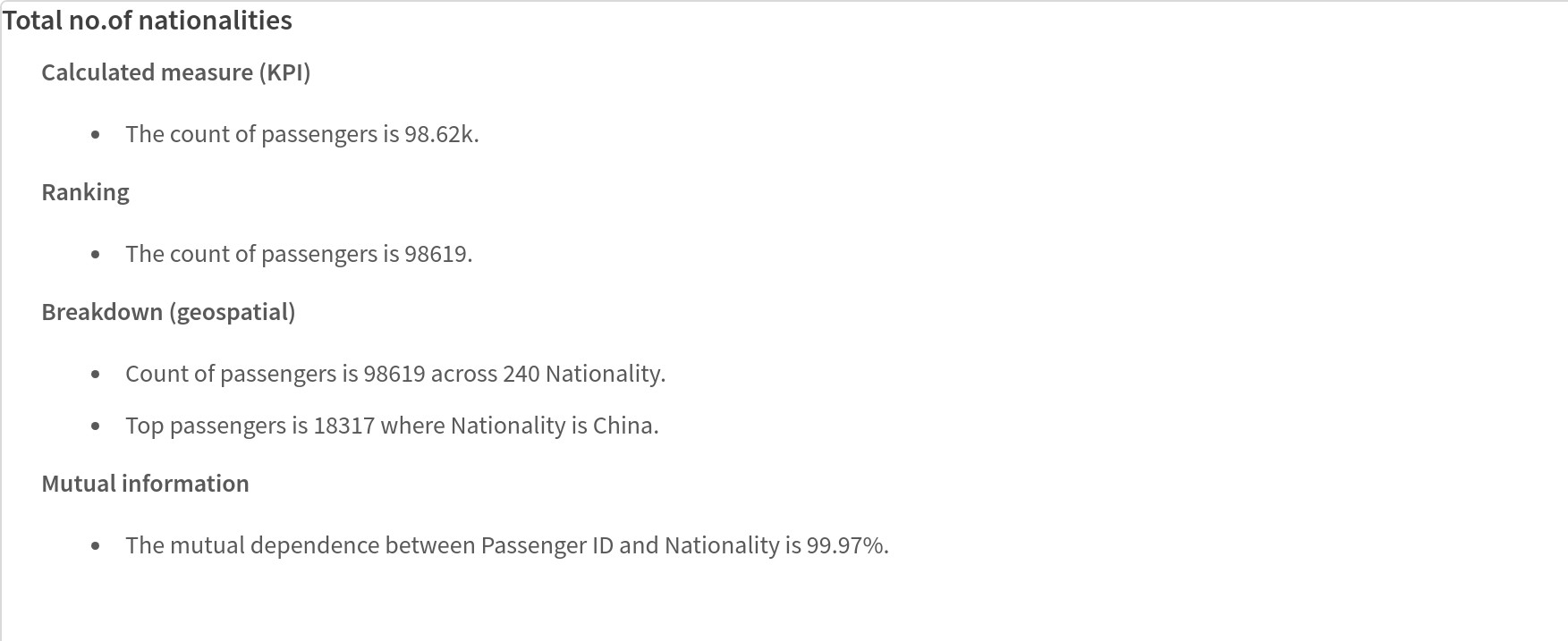
**ACTIVITY 19: TOTAL NUMBER OF COUNTIRES AND COUNTRY HAVING MAXIMUM PASSENGERS USING NL INSIGHTS**



#### ACTIVITY 20: MONTHS HAVING MAXIMUM AND MINIMUM PASSENGERS USING NL INSIGHTS



**ACTIVITY 21: TOTAL NUMBER OF NATIONALITIES AND NATIONALITY HAVING MAXIMUM PASSENGERS USING NL INSIGHTS**



#### INFORMATION ABOUT THE VISUALISATIONS I USED:

**KPI:** In a KPI visualization, you can have one or two measures and no dimensions. With two measures, the second value automatically becomes a complementary value and is shown with a smaller font size.



☆

**BARCHART:** Grouping and stacking bars makes it easy to visualize grouped data. The bar chart is also useful when you want to compare values side by side.



☆

**TREE MAP:** Treemaps display hierarchical data by using nested rectangles, that is, smaller rectangles within a larger rectangle.



☆

**PIE CHARTS**:The pie chart displays the relation between values as well as the relation of a single value to the total.



☆

**HISTROGRAM:**The histogram is suitable for visualizing distribution of numerical data over a continuous interval, or a certain time period.



☆

**MAP CHART**:Maps have many ways to present your data. You can add multiple layers to your map to display different types of information on the same map.



☆

**GAUGE**:The gauge is designed to show a single measure value and visualize how to interpret that value.



☆

**NL INSIGHTS**:The NL insights visualization displays natural language insights and information about selected dimensions and measures.



☆

**LINE CHART:**The line chart is used to show trends over time. The dimension is always on the x-axis, and the measures are always on the y-axis



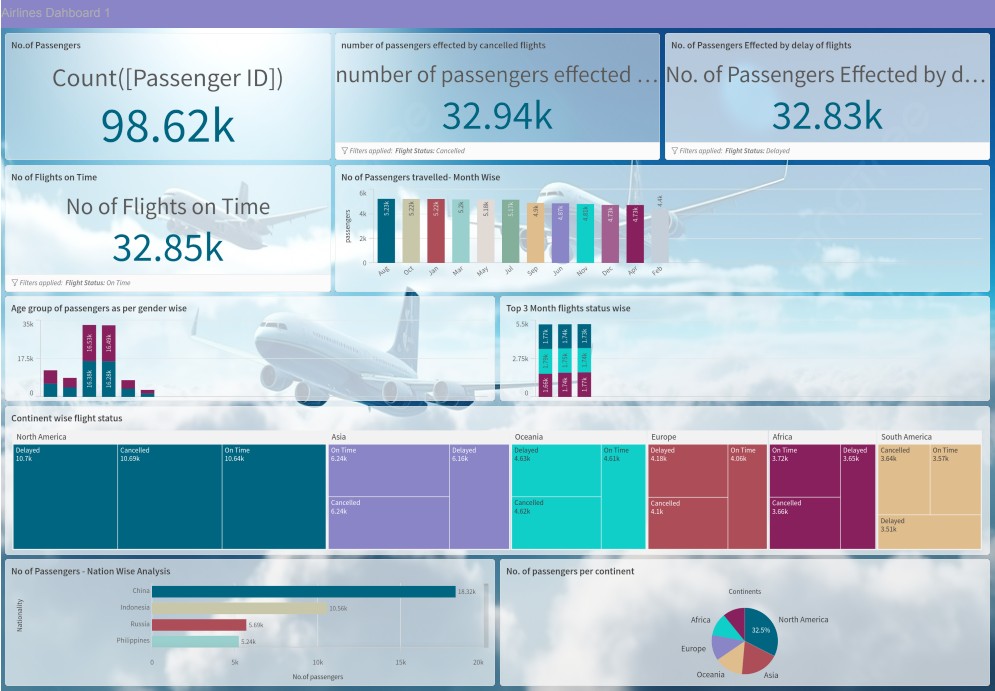
☆

## DASHBOARD

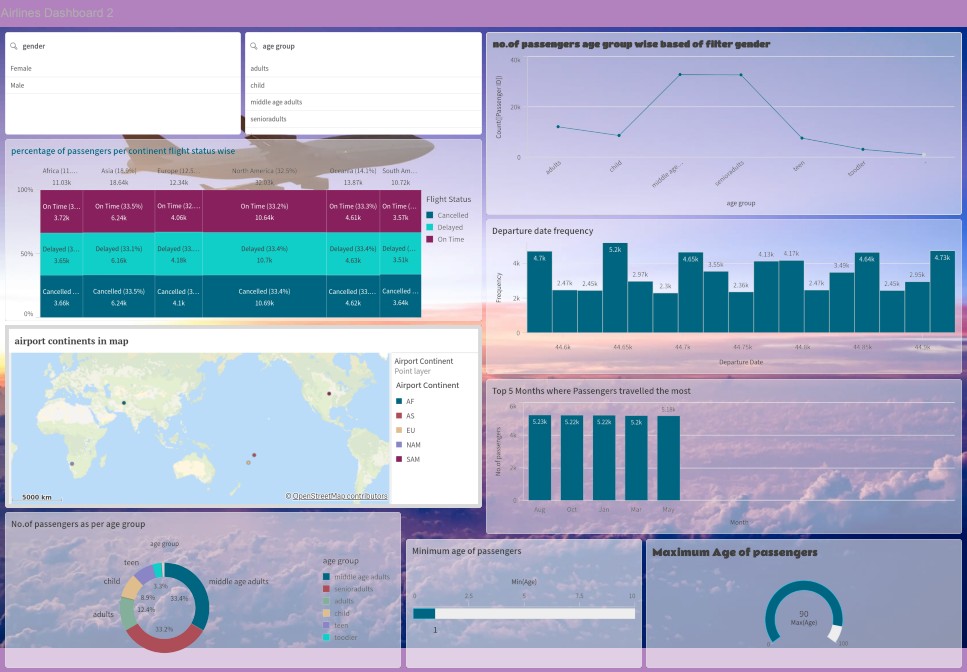
A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data and are typically designed for a speciﬁc purpose or use case. Dashboards can be used in a variety of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

* 1. ***RESPONSIVE AND DESIGN OF DASHBOARD:***

#### DASHBOARD 1:



**DASHBOARD 2:**



**DASHBOARD 3:**

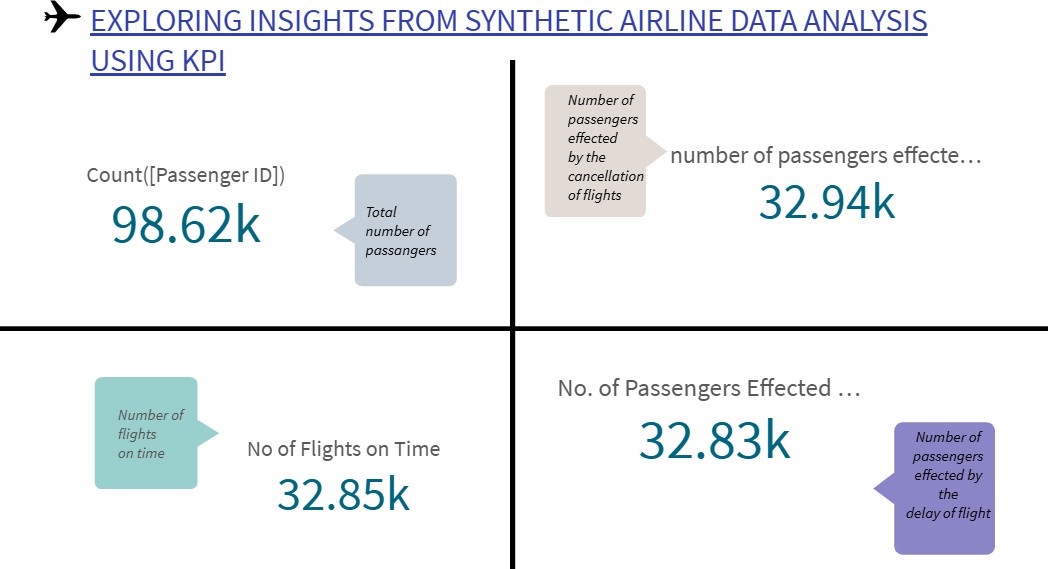


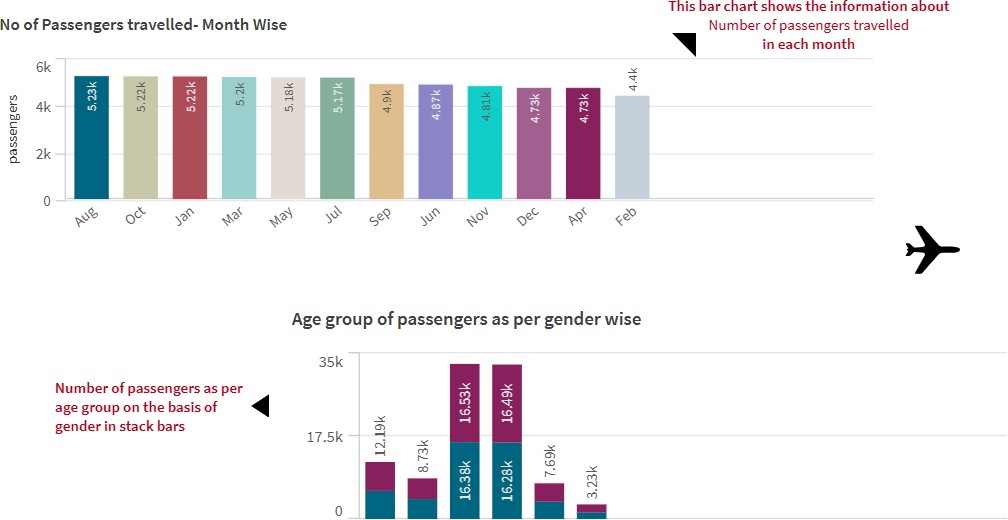
## STORY

#### STORY CREATION:

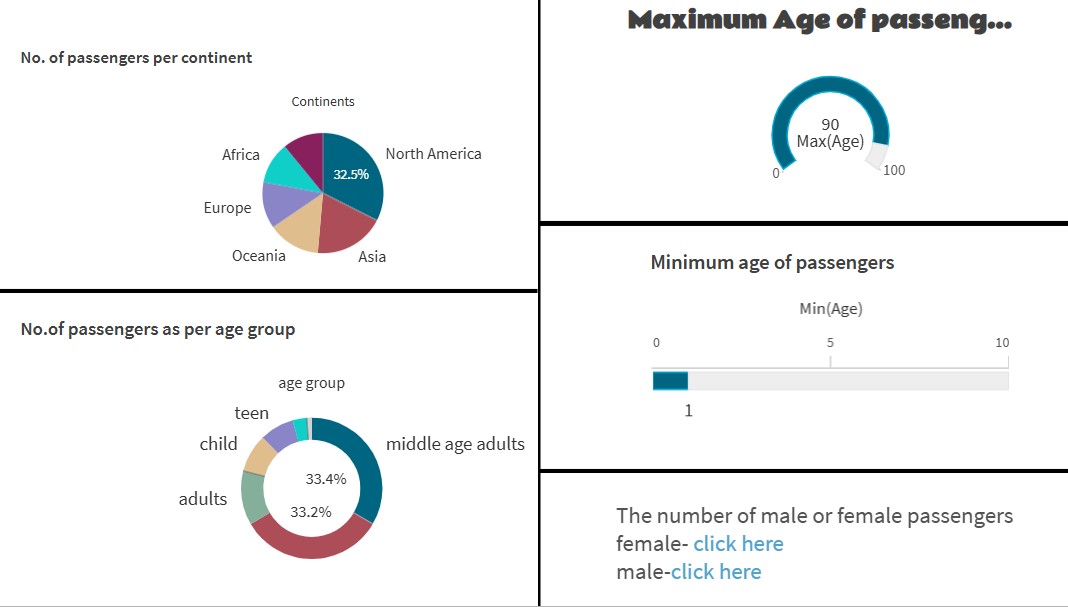
A data story is a way of presenting data and analysis in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and 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.





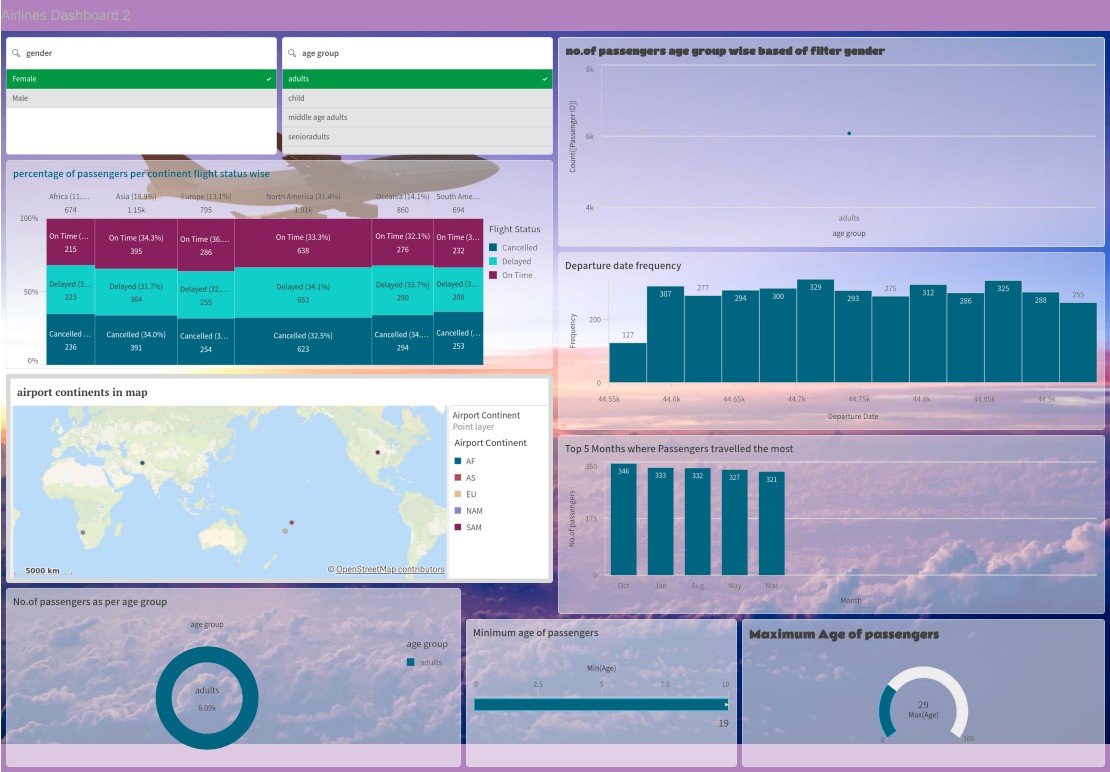






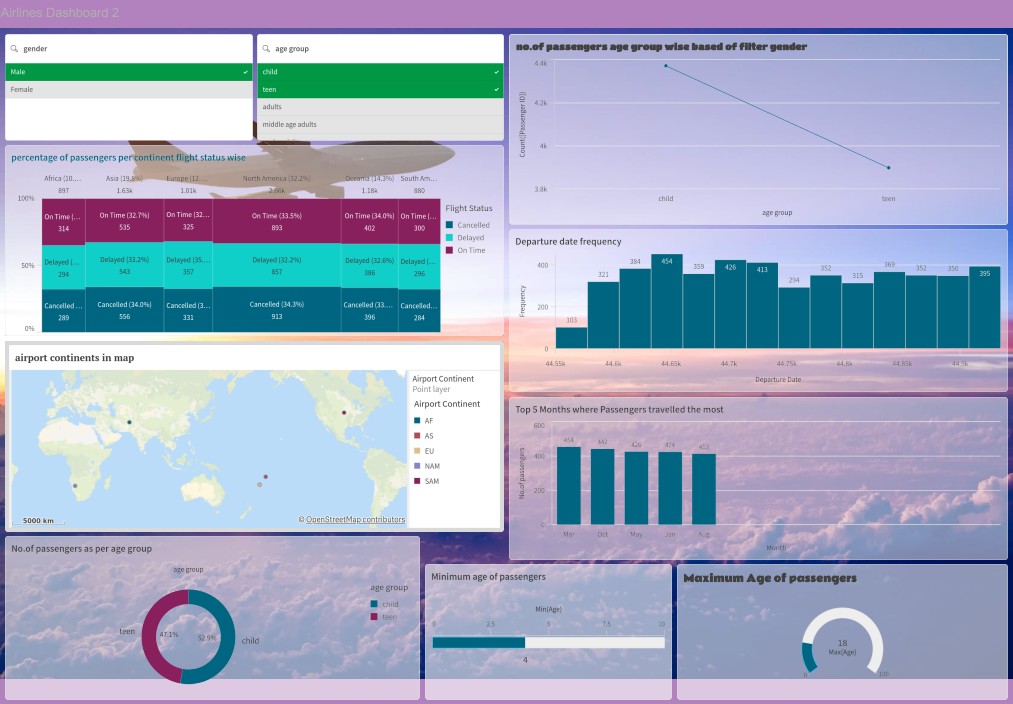
## PERFORMANCE TESTING

#### UTILISATION OF DATA FILTERS



**-**In this dashboard, I used 2 ﬁlters on the basis of gender and age group

-According to female gender and age group adult, the visualisations give the above data.



**-**The dashboard having ﬁlters gender and age group.

-The gender is "male" and taken two age groups "child" and "teen" gives the following information in the view as visualizations.