**Report**

**Exploring Insights from Synthetic Airline Dataset with Qlik**

**Table of Contents**

1. Abstract

2. Introduction

2.1 Technical Architecture:

3. App Creation and Data Upload

3.1. Creating a New App

3.2. Upload the Data File

4. Data Preparation

4.1. Removing Duplicates and Null Values

4.2. Sample Code

5. Visualization Creation

5.1. Key Performance Indicators

5.2. Continent-wise Flight Status

6. Dashboard and Storytelling

6.1. Dashboard Creation

6.2. Storytelling

7. Project Analysis and Scope

7.1. Analysis

7.2. Scope

8. Conclusion

\

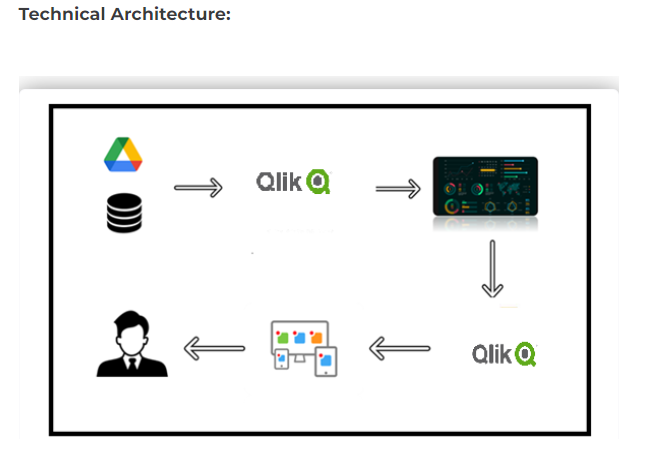
# Abstract

This report presents a comprehensive analysis of airline flight performance using Qlik Cloud, a robust data visualization and analytics tool. The project encompasses setting up Qlik Cloud, creating a new application, uploading and preparing data, and generating insightful visualizations. Our primary objective is to measure key performance indicators (KPIs) and analyse flight data across different continents to derive actionable insights. These insights aim to enhance operational efficiency, improve customer satisfaction, and inform strategic decision-making within the airline industry.

# Introduction

Qlik Cloud is a powerful tool designed for data analytics and visualization, enabling users to create interactive and informative dashboards. In this project, we leverage Qlik Cloud to analyse a comprehensive airline dataset. This analysis focuses on key performance metrics such as the total number of passengers, the number of on-time flights, delayed flights, and cancelled flights. By visualizing this data, we aim to uncover patterns and insights that can support strategic decisions and operational improvements in the airline industry. This project demonstrates how Qlik Cloud can transform raw data into meaningful visual representations that facilitate deeper understanding and actionable insights.

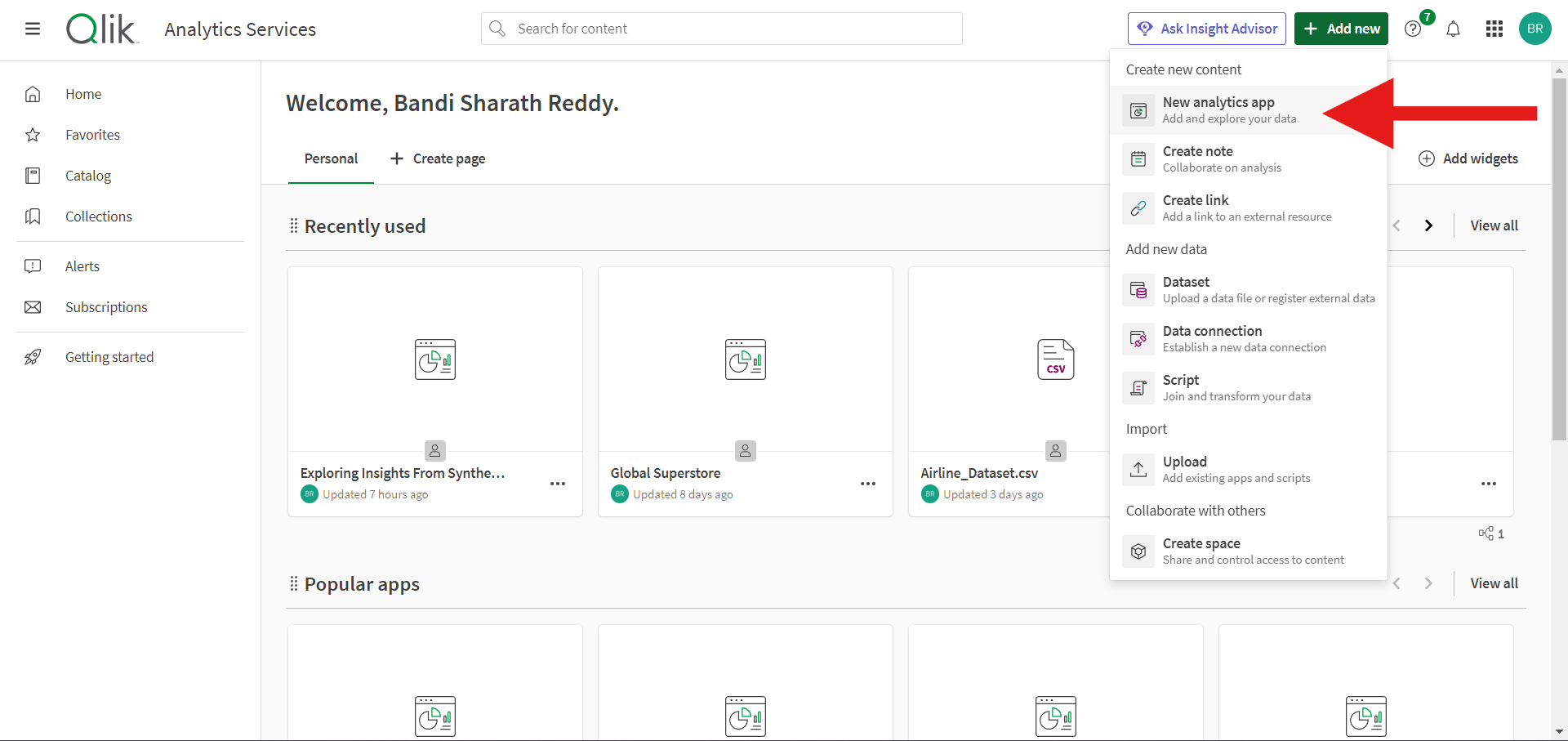
**Technical Architecture:**



**Create a New App and Upload Data:**

**Create a New App:**

1. Click on the "Create App" button to start a new project.

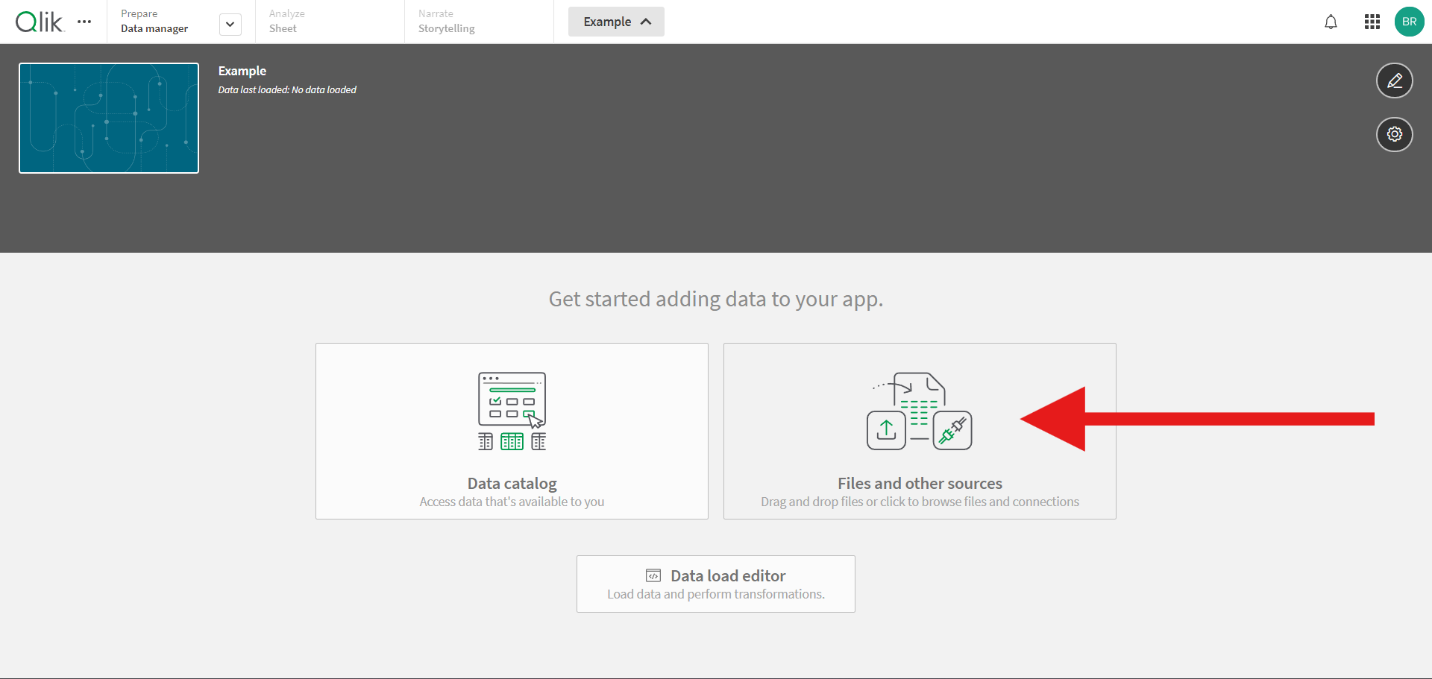


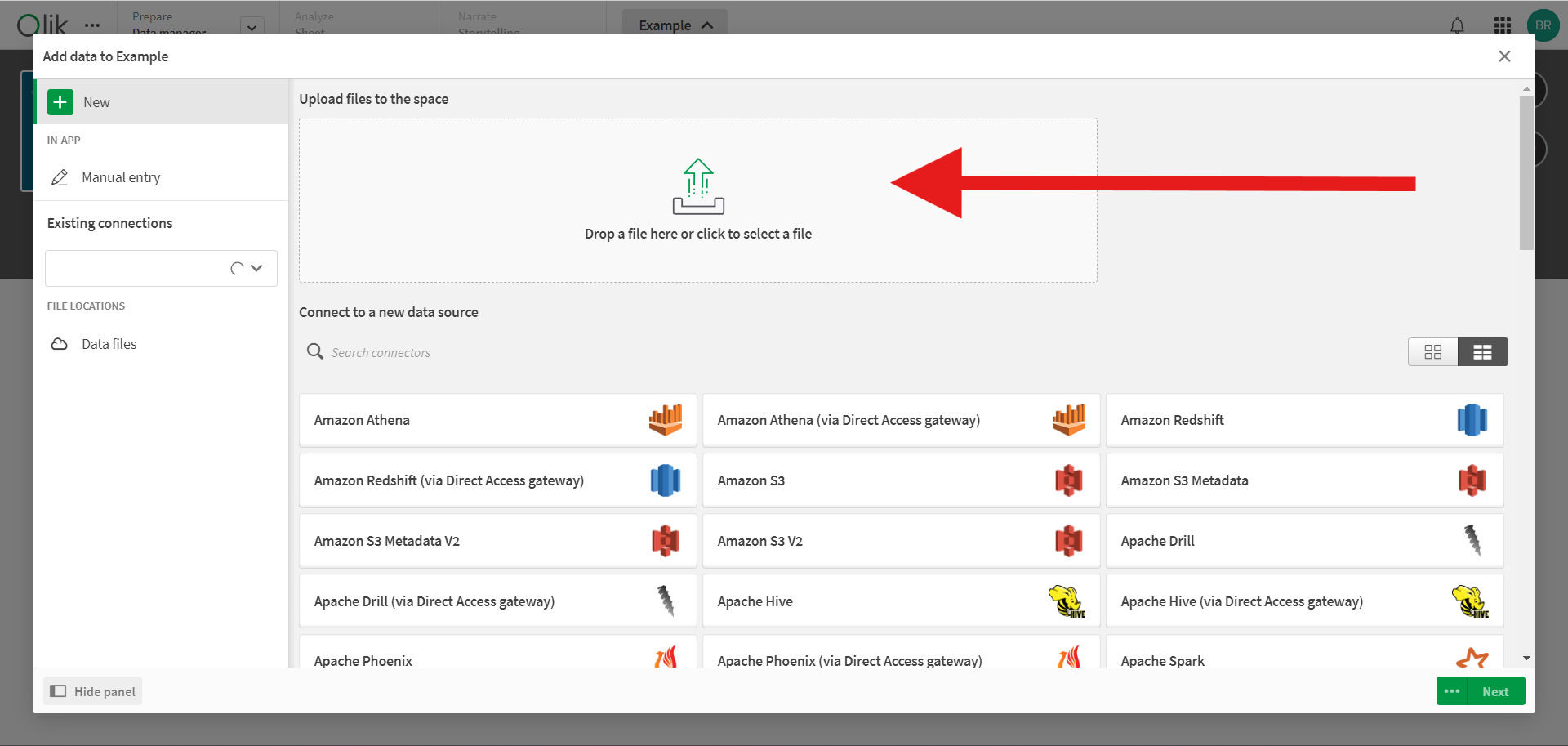
**Upload the Data File:**

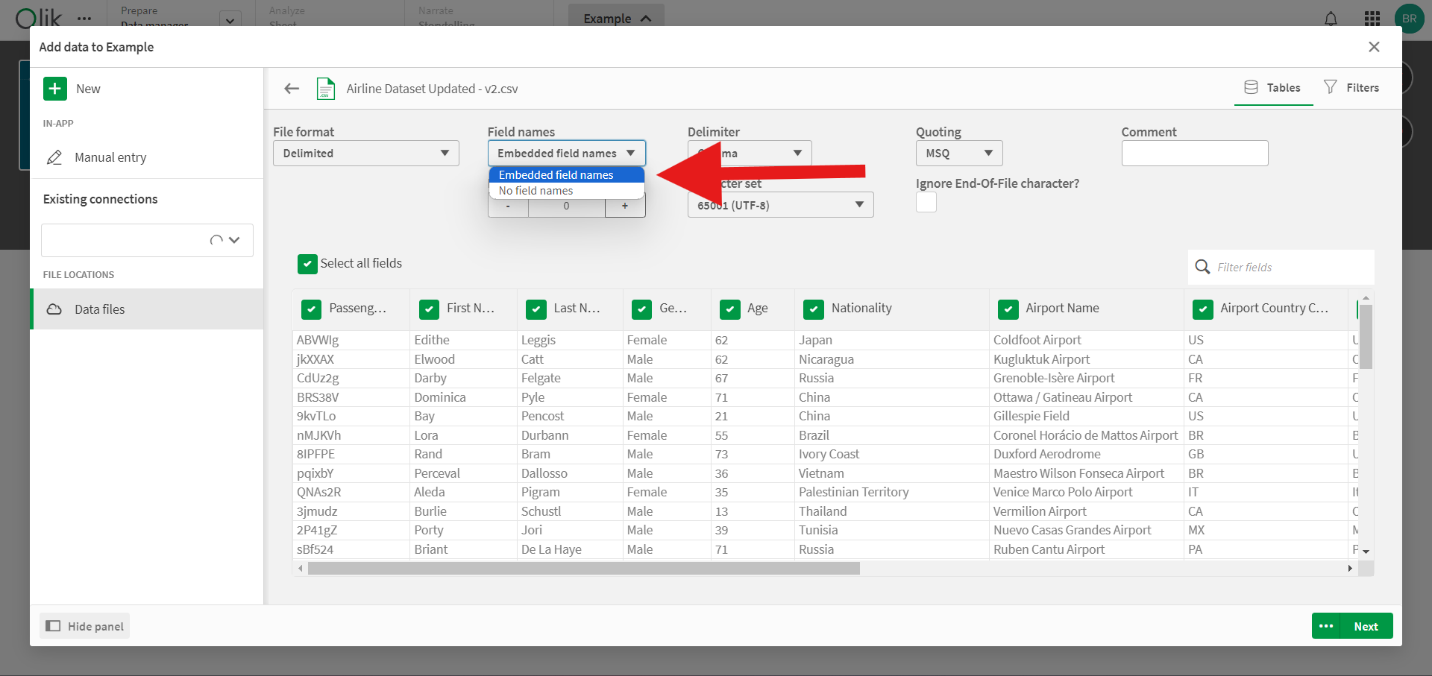
1. Go to Skill Wallet and download the project flow data set.

Dataset Link : <https://www.kaggle.com/datasets/iamsouravbanerjee/airline-dataset/data?select=Airline+Dataset+Updated+-+v2.csv>

1. In Qlik Sense, upload this data set into the new app. Ensure the dataset is embedded correctly to use the first row as headers if it is not done automatically.







# Data Preparation

**Removing Duplicates and Null Values**

**Data Load Editor:**

1. Go to the 'Prepare' tab and open the Data Load Editor.
2. Modify the default Qlik script to handle duplicates and null values as per the requirements.

**Sample Code:**

Use the following Qlik Script to pre-process the airline dataset:

```

// Load data from the source

LOAD

[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],

[Flight Status]

FROM [lib://YourDataFolder/Airline\_dataset.csv]

(txt, codepage is 1252, embedded labels, delimiter is ',', msq);

// Remove duplicates

NoConcatenate

LOAD Distinct \*

Resident Airline\_dataset;

// Handle null values by replacing them with appropriate values

LOAD

[Passenger ID],

If(Len(Trim([First Name]))=0, 'Unknown', [First Name]) as [First Name],

If(Len(Trim([Last Name]))=0, 'Unknown', [Last Name]) as [Last Name],

If(IsNull(Gender), 'Unknown', Gender) as Gender,

If(IsNull(Age), 0, Age) as Age,

If(Len(Trim(Nationality))=0, 'Unknown', Nationality) as Nationality,

If(Len(Trim([Airport Name]))=0, 'Unknown', [Airport Name]) as [Airport Name],

If(Len(Trim([Airport Country Code]))=0, 'Unknown', [Airport Country Code]) as [Airport Country Code],

If(Len(Trim([Country Name]))=0, 'Unknown', [Country Name]) as [Country Name],

If(Len(Trim([Airport Continent]))=0, 'Unknown', [Airport Continent]) as [Airport Continent],

If(Len(Trim(Continents))=0, 'Unknown', Continents) as Continents,

If(IsNull([Departure\_Date]), '1900-01-01', [Departure\_Date]) as [Departure\_Date],

If(Len(Trim([Arrival Airport]))=0, 'Unknown', [Arrival Airport]) as [Arrival Airport],

If(Len(Trim([Pilot Name]))=0, 'Unknown', [Pilot Name]) as [Pilot Name],

If(Len(Trim([Flight Status]))=0, 'Unknown', [Flight Status]) as [Flight Status],

MonthName([Departure\_Date]) AS MonthYear

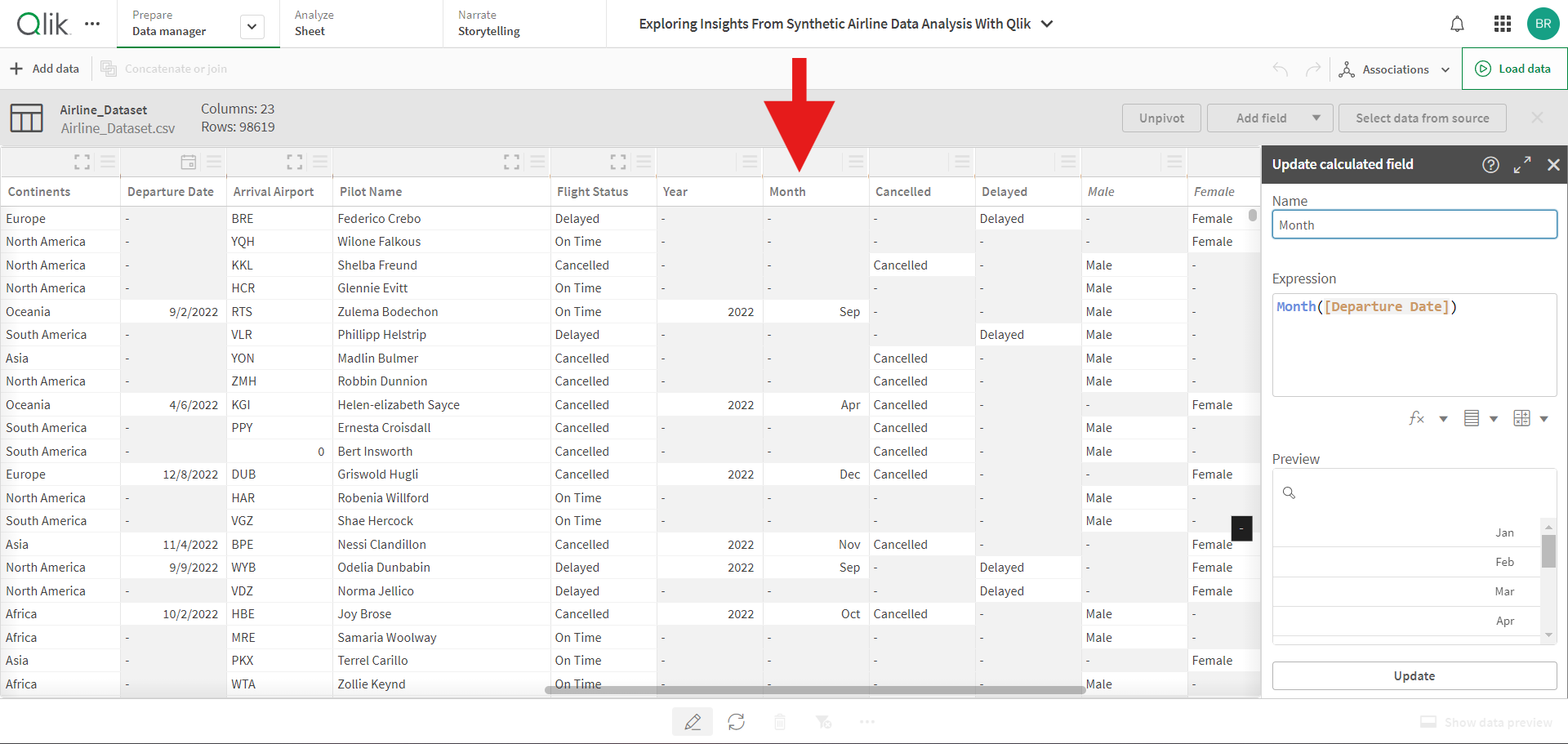
Resident Airline\_dataset;

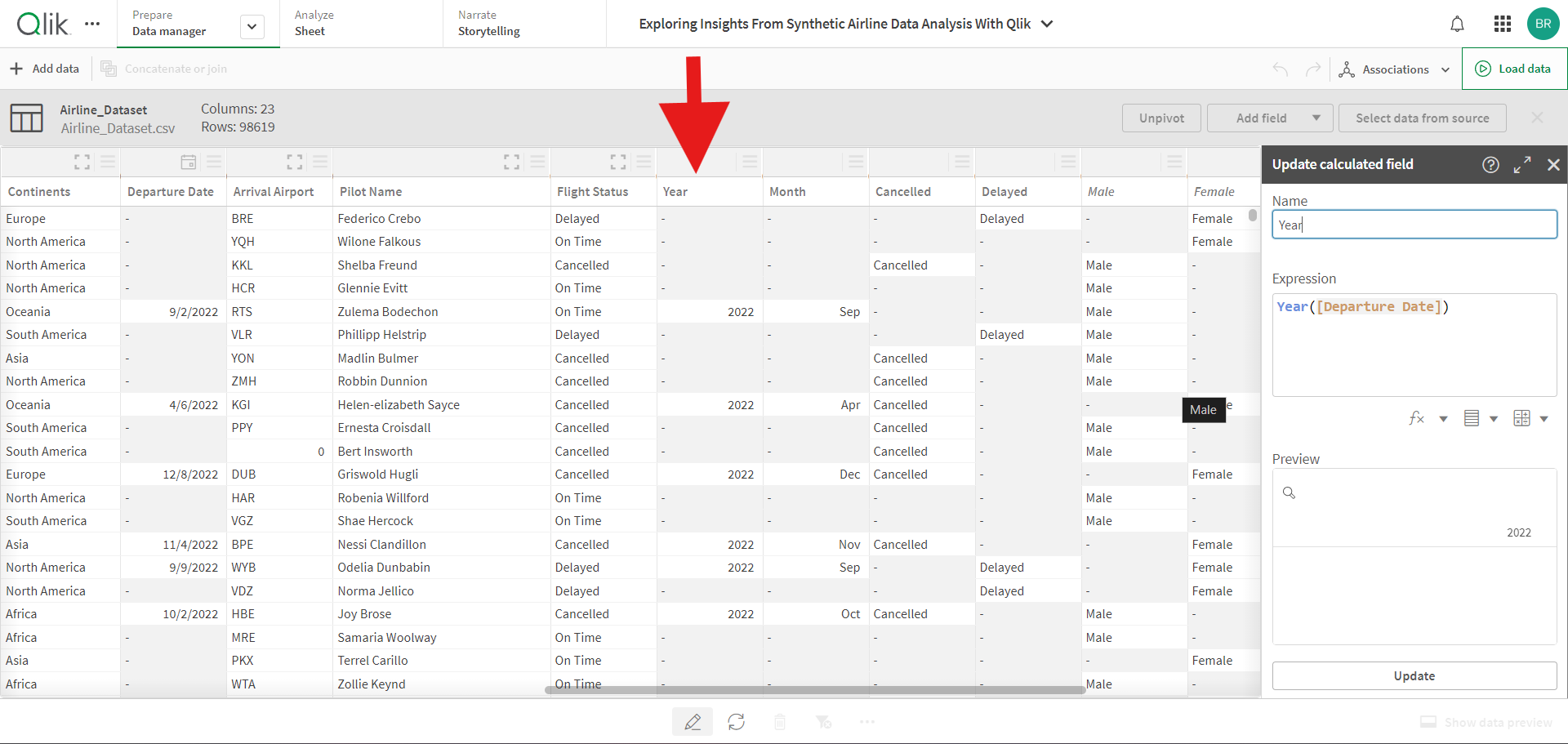
// Drop the temporary table

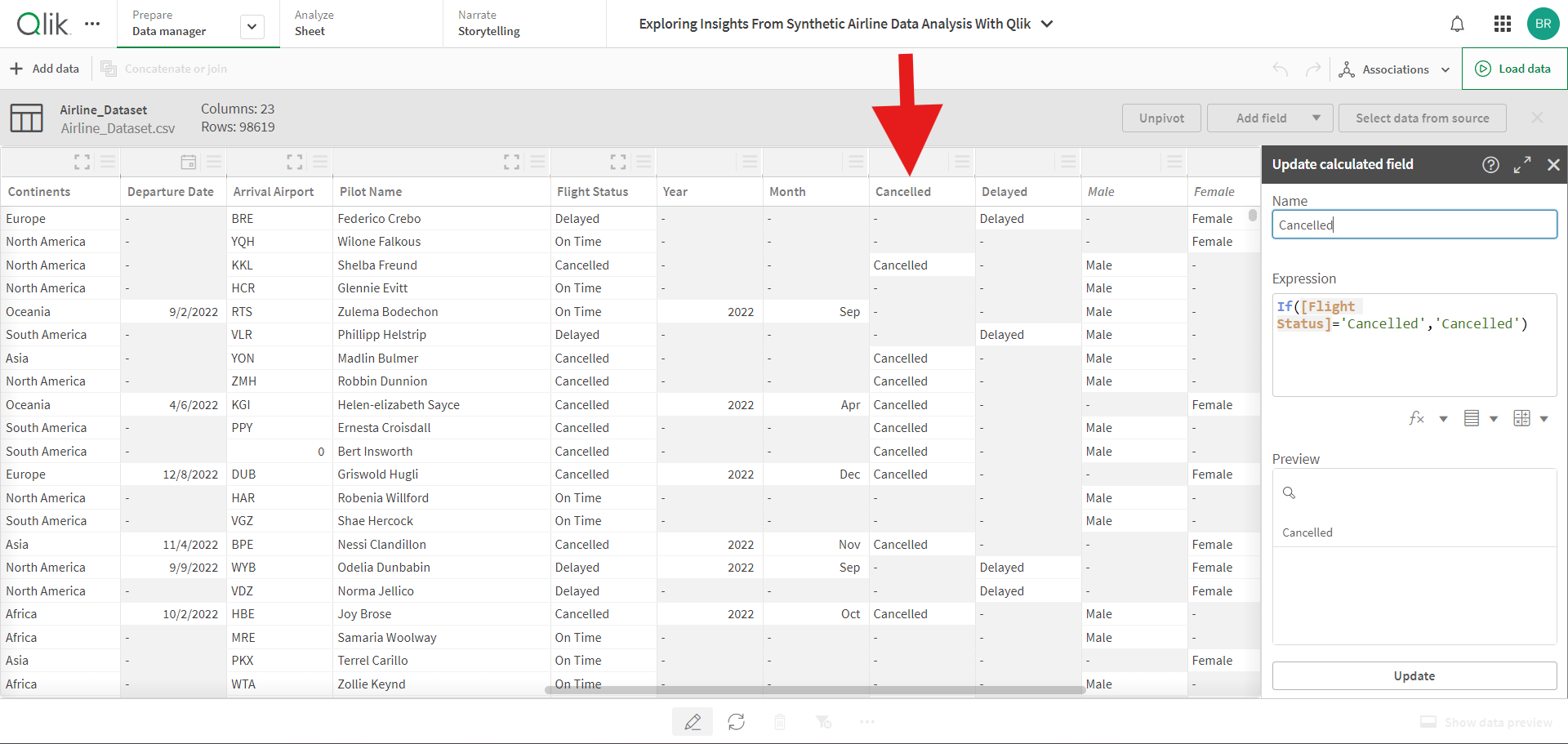
Drop Table Airline\_dataset;

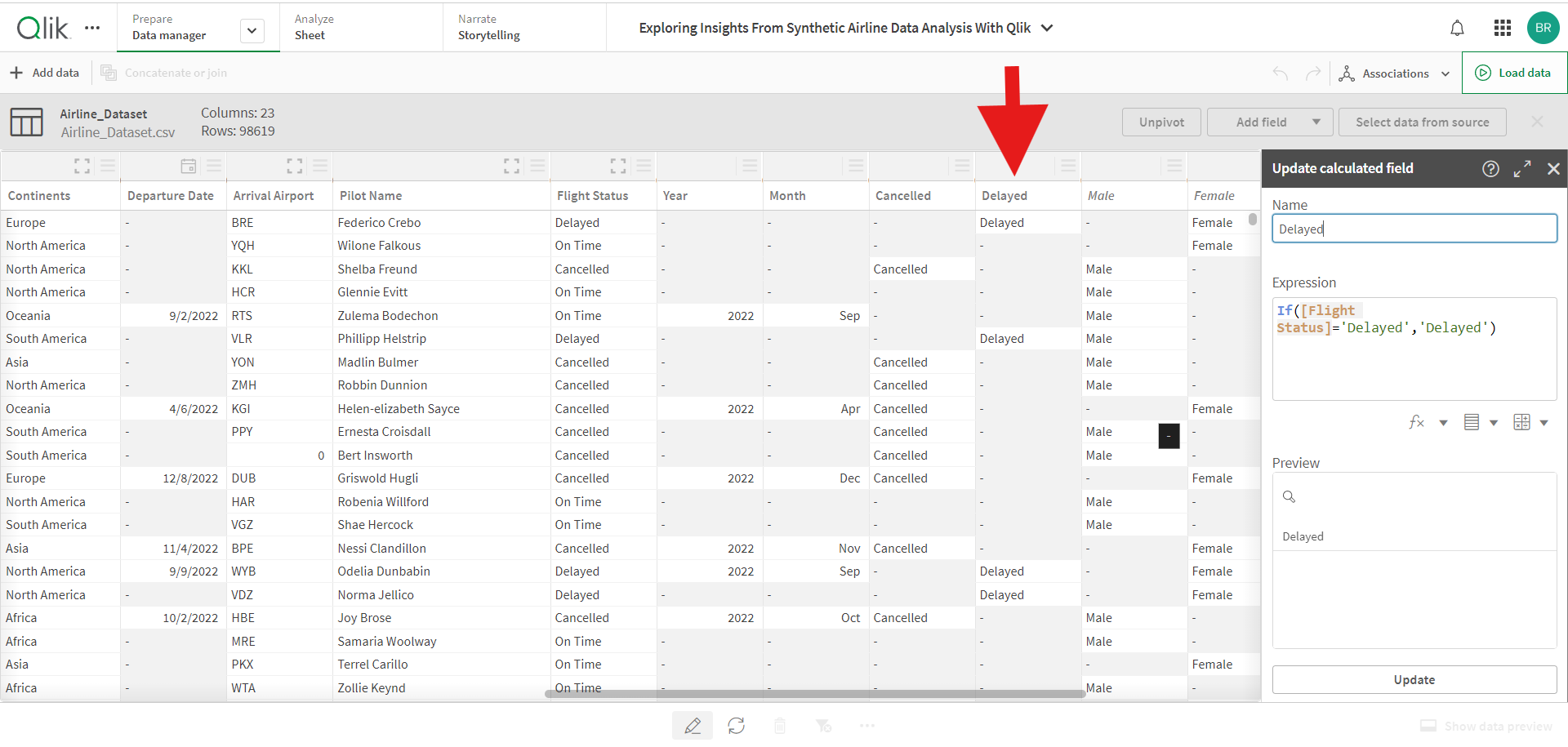
```

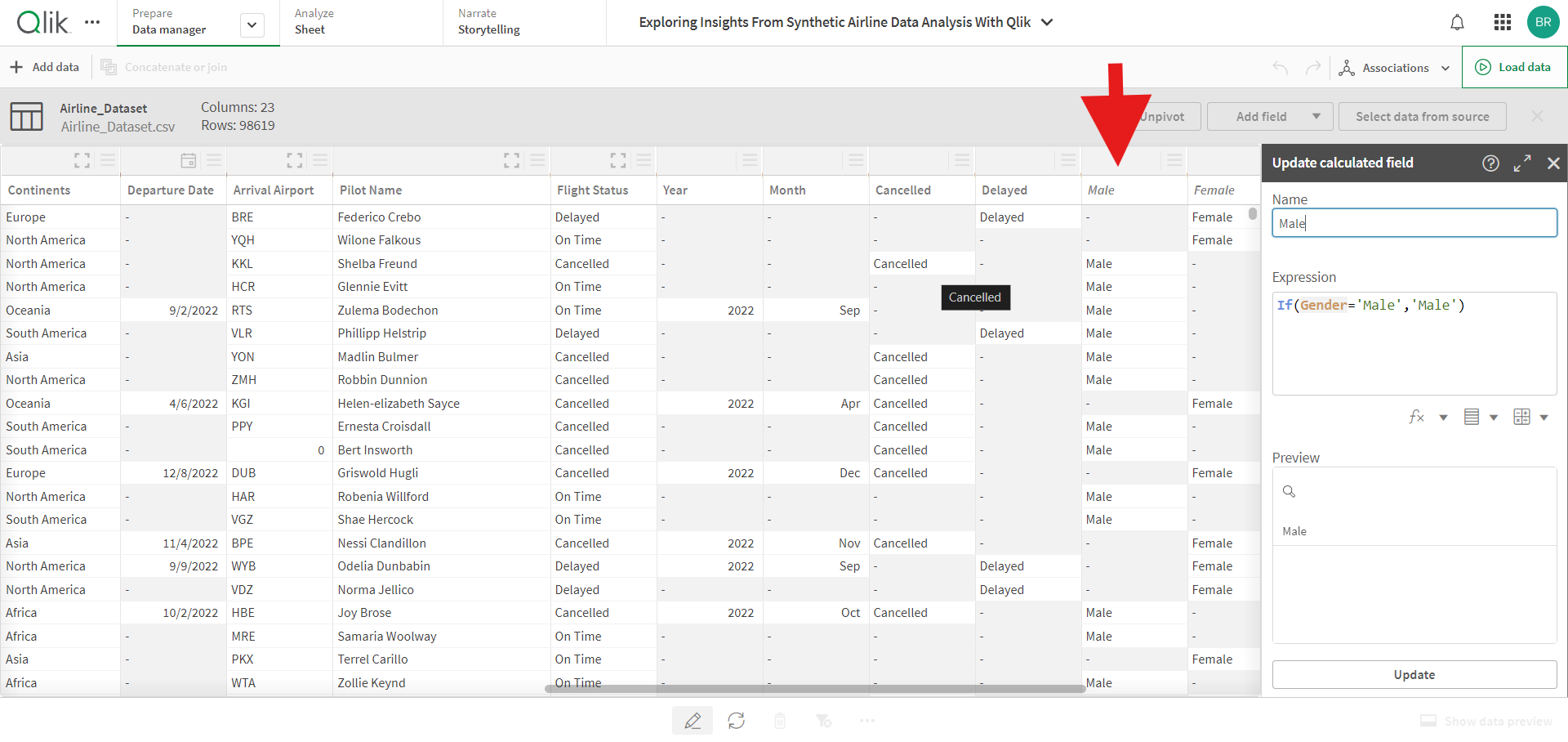
**The dataset is already clean [no need to use above script], so you can start preprocessing it directly without the need for additional data cleaning.**

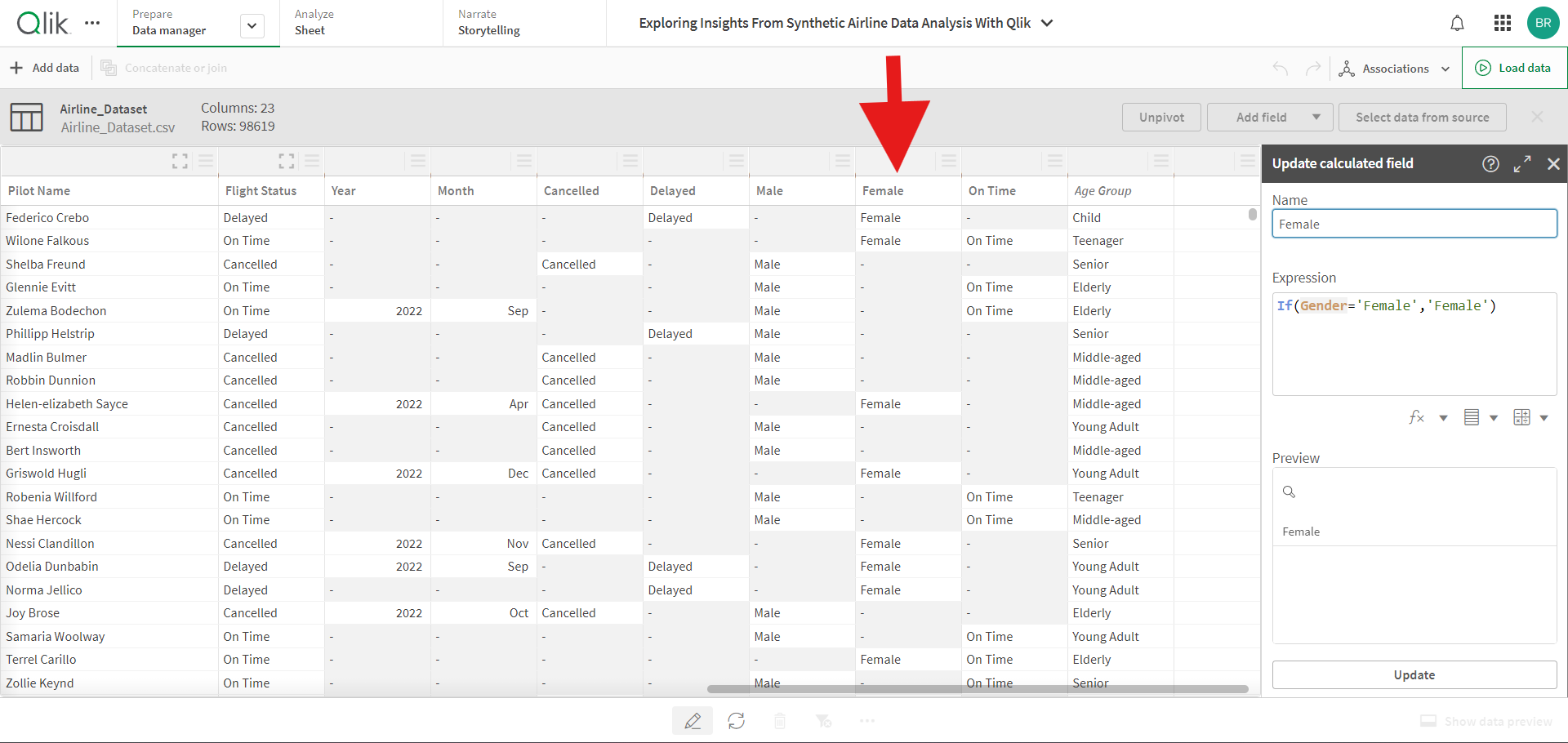
****

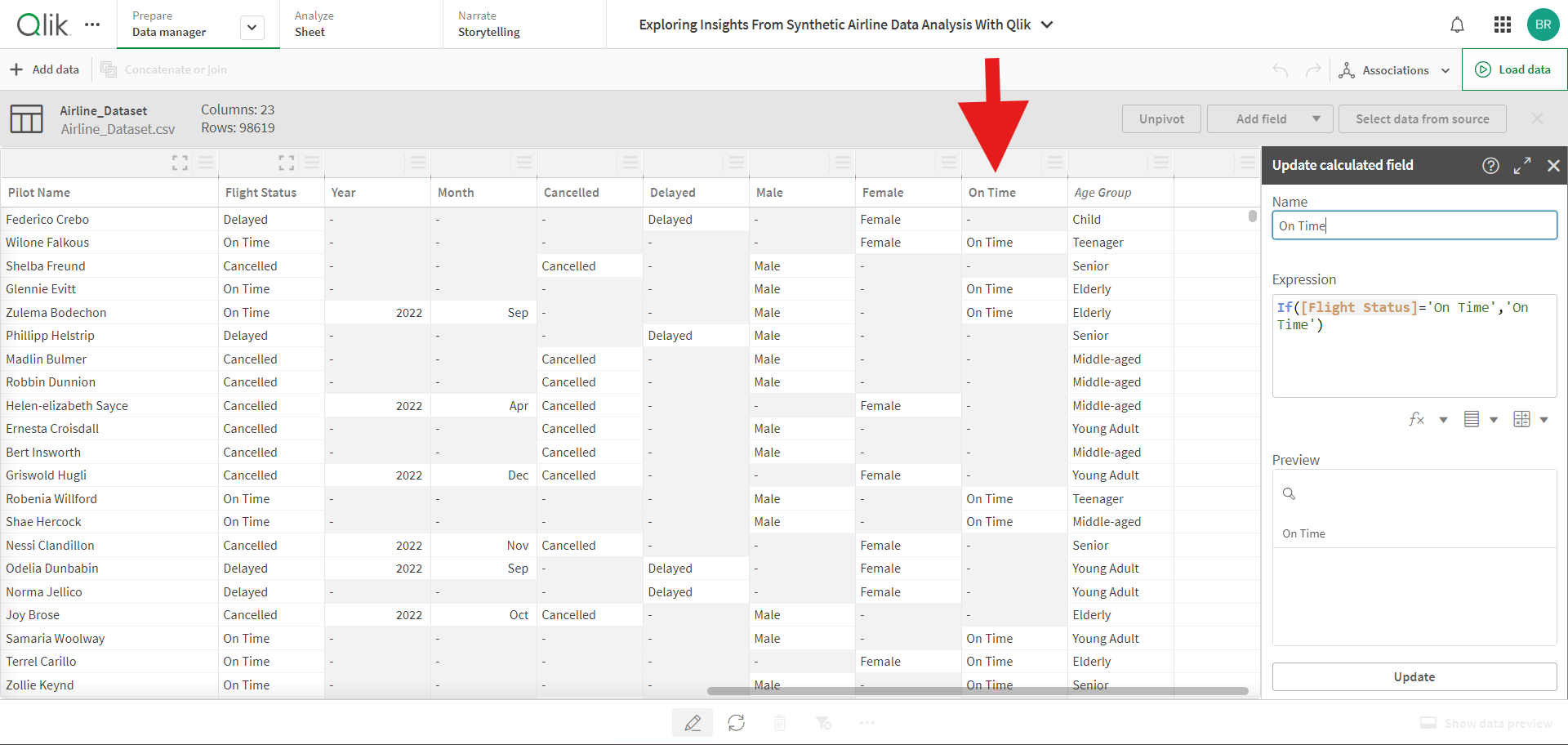
****

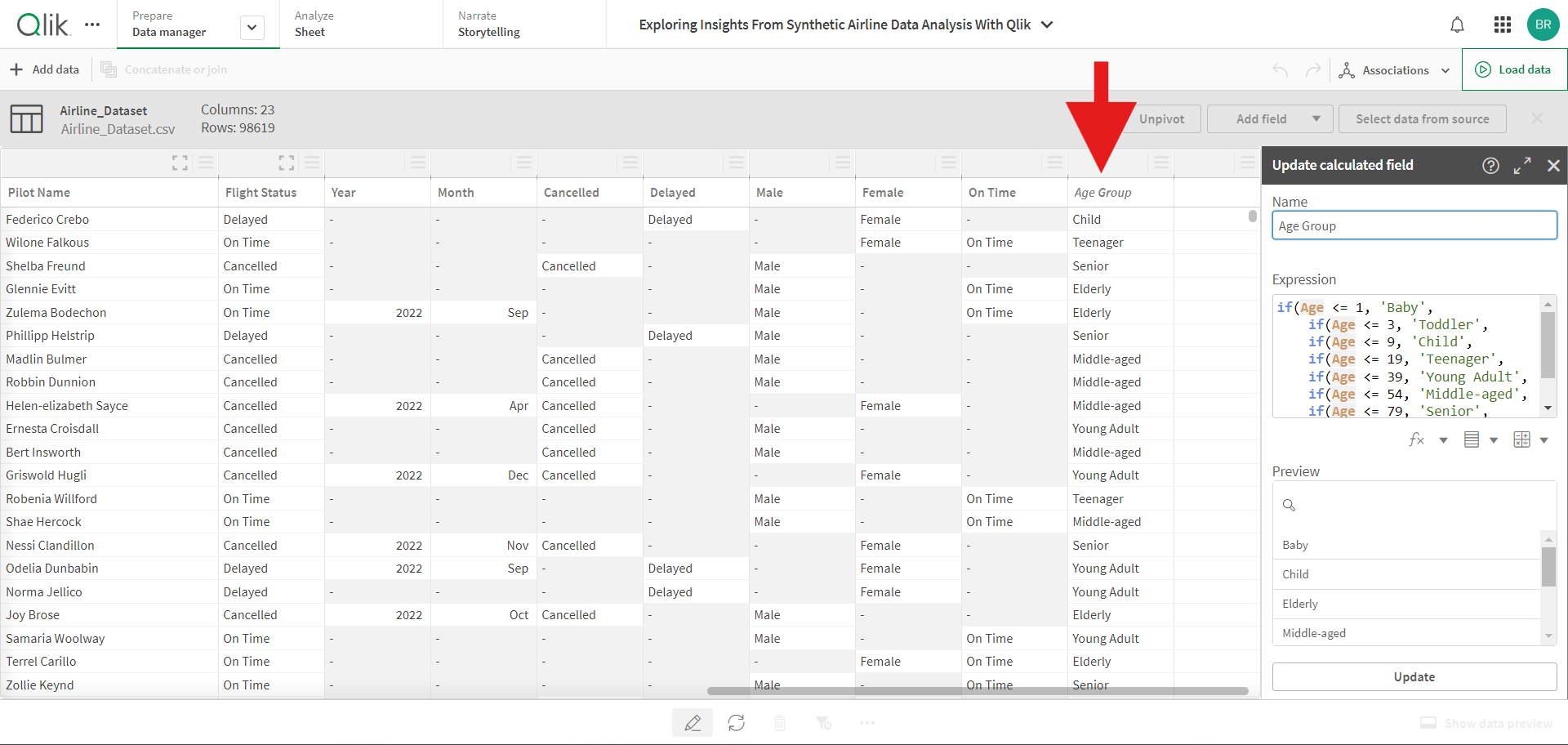
****

****









# Visualization Creation

**Key Performance Indicators**

**Total Passengers:**

Measure: `Count (distinct [Passenger ID])`

**On-time Flights:**

Measure: `Count ({< [Flight Status] = {'On Time'} >} DISTINCT [Passenger ID])`

**Delayed Flights**:

Measure: `Count ({< [Flight Status] = {'Delayed'} >} DISTINCT [Passenger ID])`

**Cancelled Flights:**

Measure: `Count ({< [Flight Status] = {'Cancelled'} >} DISTINCT [Passenger ID])`

**Continent-wise Flight Status** Visualization: Tree Map

Dimension: `Continent`

Rectangle Size: `Flight Status`

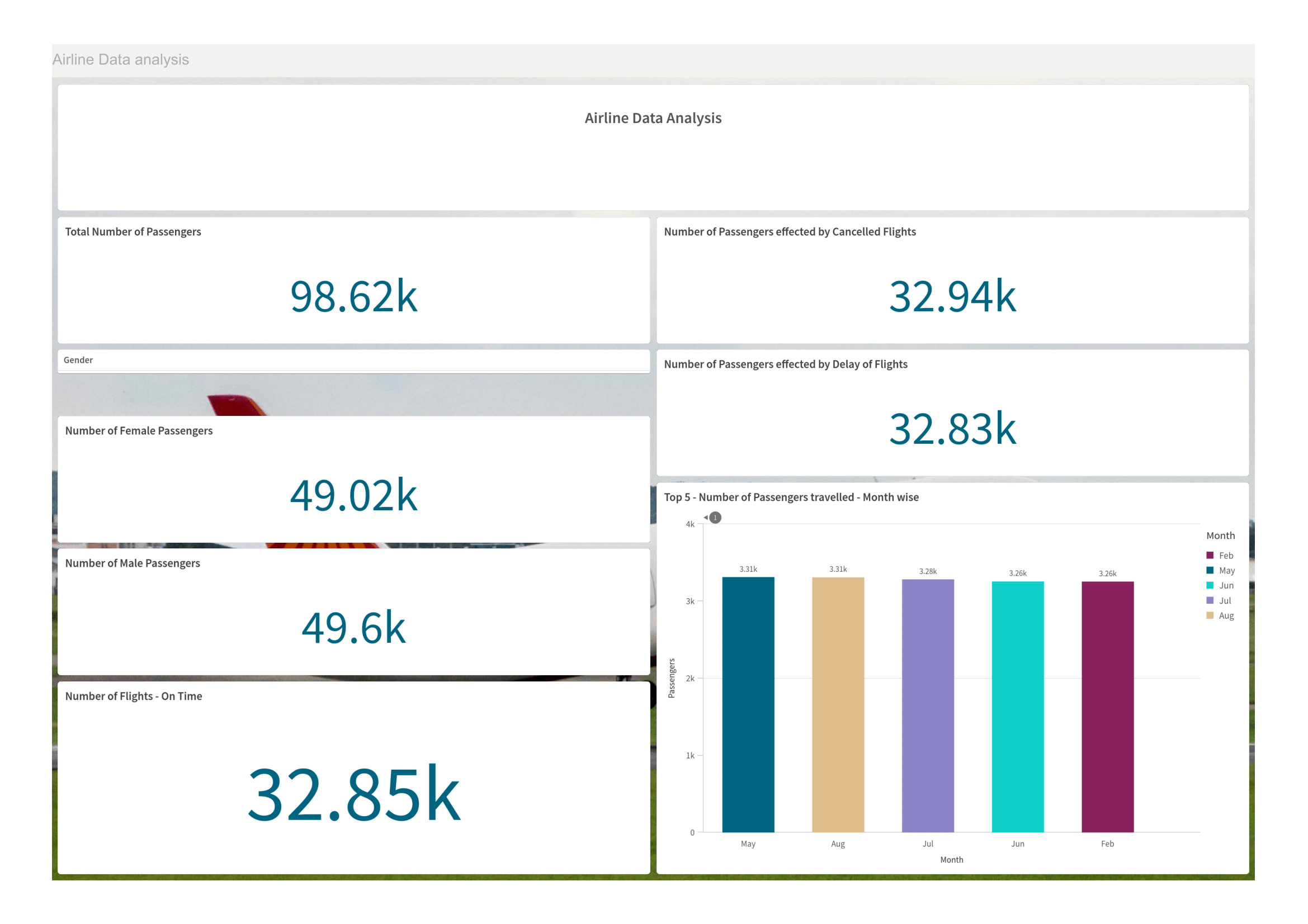
Measure: `Count (Flight Status)`

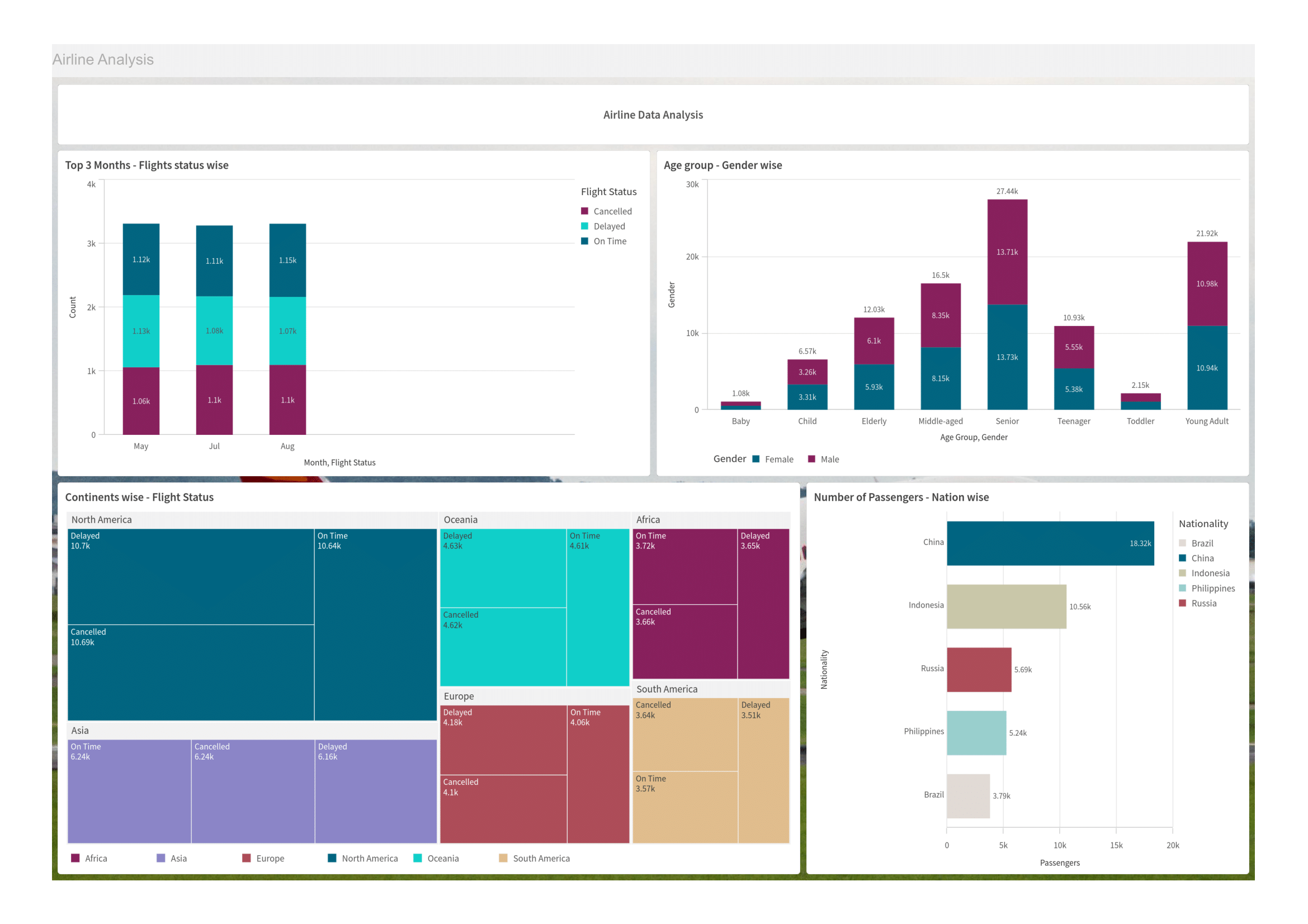
# Dashboard and Storytelling

**Dashboard Creation**

**Design Dashboards:**

1. Design and create two dashboards, strategically aligning the visualizations and key performance indicators to provide a comprehensive overview of the airline’s performance.
2. Ensure the dashboards are interactive and user-friendly, enabling stakeholders to easily derive insights.

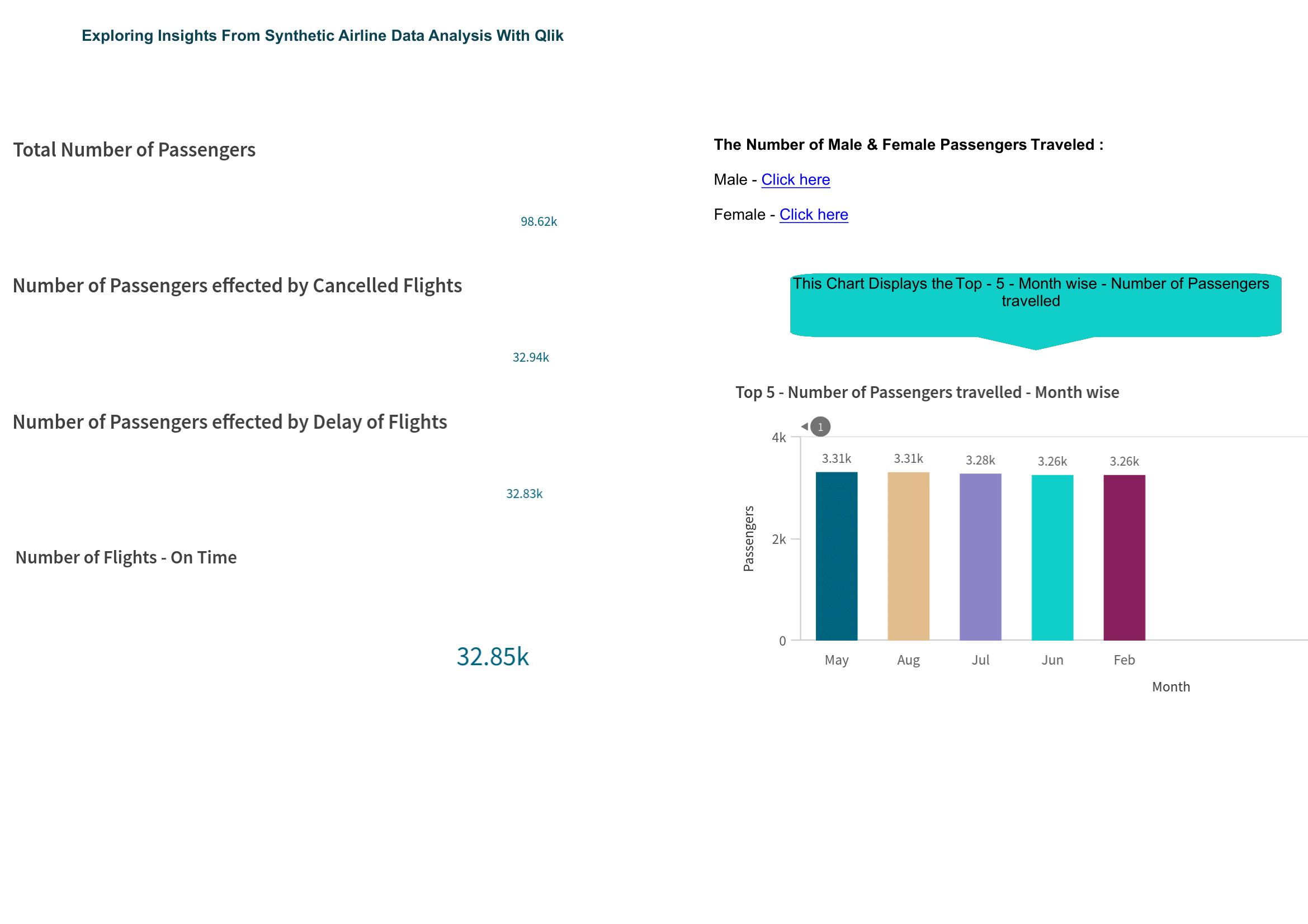


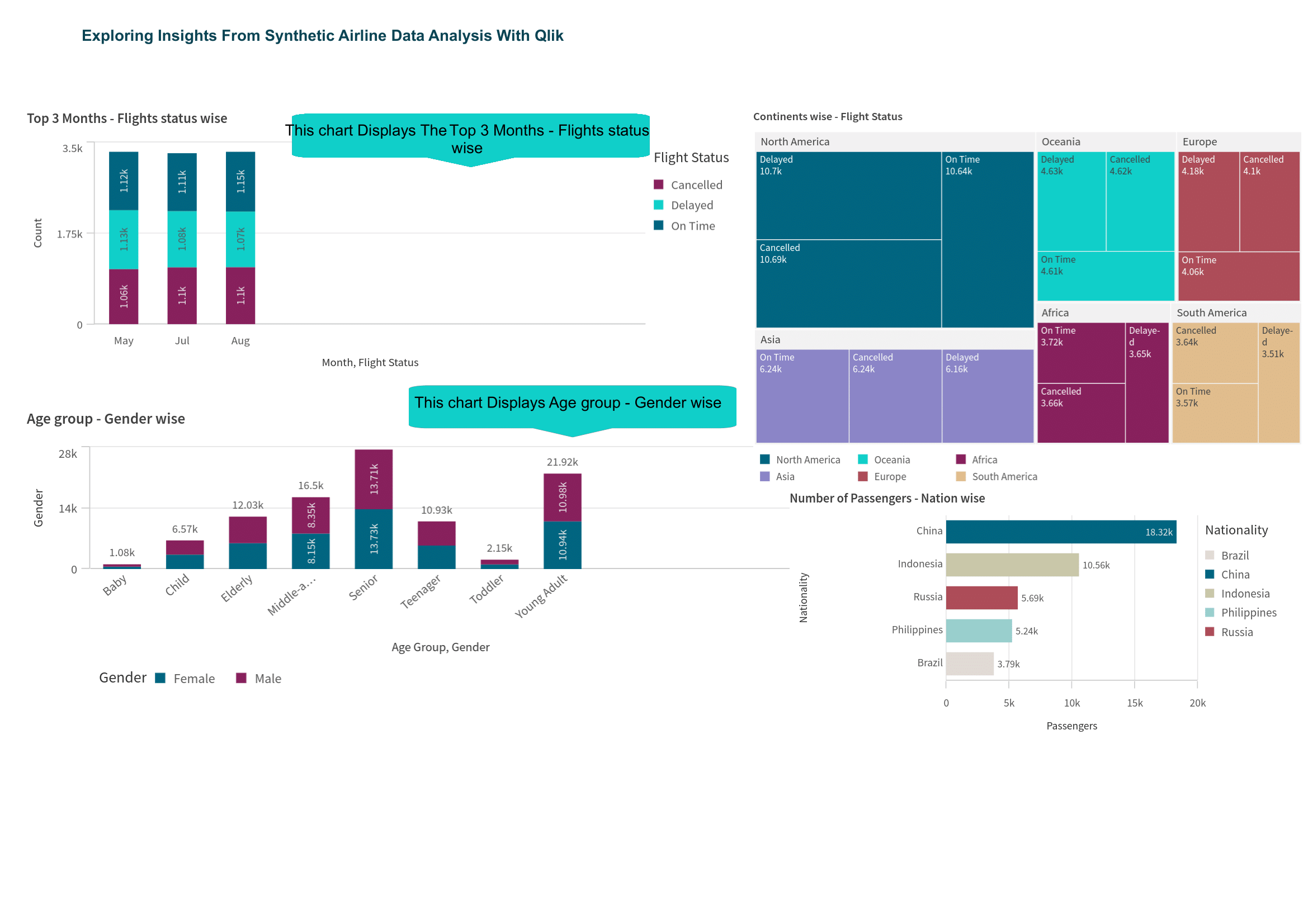


**Storytelling**

**Create a Story:**

1. Use the storytelling feature in Qlik Cloud to create a compelling presentation (PPT).
2. Include visualizations and narrative explanations to communicate the insights derived from the data effectively. The storytelling approach helps in engaging the audience and making the data more accessible and understandable.





# Project Analysis and Scope

**Analysis:**

The analysis of the project data involves several steps to ensure data quality and derive meaningful insights. Initially, the raw data is pre-processed to handle duplicates and null values. However, since the provided dataset is already clean, we can proceed directly to structuring the data for analysis. This involves organizing and transforming the data to create key performance indicators (KPIs) such as the total number of passengers, on-time flights, delayed flights, and cancelled flights. We start by structuring passenger data to calculate the total number of passengers and categorizing flight statuses to distinguish between on-time, delayed, and cancelled flights. We then develop KPIs to measure these metrics, using visualizations like bar charts and line graphs to represent them. For continent-wise analysis, we create tree maps and geographic maps to examine the distribution of flight statuses across different continents, helping identify regional performance patterns and potential issues. Further analysis focuses on operational performance, examining factors contributing to on-time performance, identifying common causes of flight delays, and assessing their impact on overall performance. These visualizations and analyses facilitate an in-depth understanding of operational efficiency and customer satisfaction, providing actionable insights for strategic decision-making.

**Scope:**

The scope of this project encompasses the following areas:

1. Data Quality Improvement: The raw data is pre-processed to handle duplicates and null values, ensuring the accuracy and completeness of the dataset.

1. Performance Measurement: Key performance indicators (KPIs) such as the total number of passengers, on-time flights, delayed flights, and cancelled flights are developed to monitor flight performance.

1. Geographical Analysis: Visualizations, such as tree maps, are used to analyze flight status distribution across different continents, providing insights into regional performance variations.

1. Business Insights: Deriving actionable insights to improve operational efficiency, customer satisfaction, and strategic decision-making.

1. Dashboard Creation: Creating interactive dashboards that provide a comprehensive view of the airline's performance.

1. Storytelling and Presentation: Using Qlik Cloud’s storytelling feature to create presentations that effectively communicate the insights derived from the data.

# Conclusion

This report outlines the comprehensive process of setting up Qlik Cloud, preparing data, creating visualizations, and compiling them into a cohesive story. The provided script and steps ensure a systematic approach to analyzing and presenting data effectively using Qlik Cloud. The project analysis emphasizes the importance of data quality, performance measurement, and geographical analysis in deriving business insights and making informed decisions. The scope of the project demonstrates the potential of Qlik Cloud in transforming raw data into valuable insights that can drive operational improvements and strategic planning.

**GitHub link:** [**https://github.com/Bandi-Sharath-Reddy-SRB/Exploring-Insights-From-Synthetic-Airline-Data-Analysis-With-Qlik**](https://github.com/Bandi-Sharath-Reddy-SRB/Exploring-Insights-From-Synthetic-Airline-Data-Analysis-With-Qlik%20)