# Exploring Insights from Airlines Data Analysis using Qlik

### INTRODUCTION

#### Overview:

This project, "Exploring Insights from Synthetic Airline Data Analysis with Qlik,"

utilizes synthetic airline data to derive valuable insights and support decision-making for airlines, airports, and related stakeholders. The data encompasses various aspects of airline operations, including ﬂight schedules, passenger demographics,

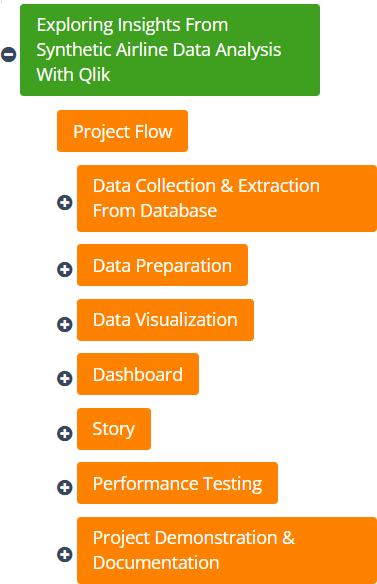
ticket sales, and performance metrics. By leveraging Qlik's powerful analytical and visualization capabilities, we aim to uncover patterns, trends, and correlations within this data to address key business challenges.

#### Purpose:

The primary purpose of this project is to demonstrate how Qlik can be used to analyze and visualize synthetic airline data to achieve speciﬁc business objectives:

* + - **Revenue Optimization**: Analyzing historical ticket sales to identify peak travel times, popular destinations, and effective pricing strategies.
    - **Operational Eﬃciency**: Enhancing airport operational eﬃciency by identifying bottlenecks and predicting peak traﬃc periods.
    - **Customer Experience Enhancement**: Improving passenger experience by understanding customer preferences and pain points through sentiment analysis of feedback data.

#### Technical Architecture:



1

# Exploring Insights from Airlines Data Analysis using Qlik

### Deﬁne Problem / Problem Understanding

#### Specify the Business Problem:

The airline industry faces several challenges that can be addressed through data analysis:

1. **Revenue Optimization**: Airlines need to maximize proﬁtability by identifying optimal pricing strategies and understanding sales trends.
2. **Operational Eﬃciency**: Airports must streamline operations to handle passenger ﬂows and luggage handling effectively.
3. **Customer Experience**: Airlines aim to enhance customer satisfaction and loyalty by addressing service quality issues and personalizing experiences.

#### Business Requirements:

To address these problems, the following business requirements are identiﬁed:

* + - Detailed analysis of ticket sales data to identify revenue opportunities.
    - Assessment of ﬂight schedules and passenger ﬂows to improve operational eﬃciency.
    - Sentiment analysis of customer feedback to enhance service quality and customer experience.

#### Literature Survey:

A review of existing literature highlights the importance of data analytics in the airline industry. Previous studies have shown how data-driven decision-making can lead to signiﬁcant improvements in revenue management, operational eﬃciency, and customer satisfaction. Here are a few relevant studies:

* + - **"Airline Data Analytics: An Overview" (Journal of Air Transport Management)**: This study provides a comprehensive overview of how airlines use data analytics to improve various aspects of their operations. It discusses the role of data in optimizing ﬂight schedules, enhancing customer experience, and managing revenue. The ﬁndings emphasize that airlines leveraging data analytics can achieve signiﬁcant competitive advantages.
    - **"Predictive Analytics in Airline Operations: A Case Study" (Journal of Airline and Airport Management)**: This article explores the use of predictive analytics in airline operations. It highlights how airlines can use historical data to

forecast demand, predict maintenance needs, and manage crew scheduling. The study demonstrates that predictive analytics can lead to more eﬃcient operations and cost savings.

2

# Exploring Insights from Airlines Data Analysis using Qlik

* + - **"Improving Airline Customer Loyalty Through Data Analysis" (Journal of Travel Research)**: This research focuses on how airlines can use customer data to enhance loyalty programs. By analyzing customer preferences and travel patterns, airlines can tailor their loyalty programs to better meet customer needs. The study shows that data-driven loyalty programs can increase customer satisfaction and retention.

#### "Revenue Management in Airlines: Data-Driven Approaches" (Annals of

**Operations Research)**: This paper discusses various data-driven approaches to revenue management in the airline industry. It examines how airlines use data to set dynamic pricing, manage seat inventory, and forecast demand. The ﬁndings suggest that data-driven revenue management strategies can signiﬁcantly increase proﬁtability.

These studies collectively illustrate the transformative potential of data analytics in

the airline industry. They provide valuable insights into how airlines can leverage data to optimize operations, enhance customer experience, and drive revenue growth.

1. **Data Collection**

#### Collect the Dataset:

The dataset used for this project is sourced from Kaggle, speciﬁcally the "Airlines " dataset. This dataset includes detailed information on Airports and passengers across various states and Union Territories in India. The key features of the dataset include:

* + 1. **Passenger ID**: A unique identiﬁer for each passenger.
    2. **First Name**: The ﬁrst name of the passenger.
    3. **Last Name**: The last name of the passenger.
    4. **Gender**: The gender of the passenger (e.g., Male, Female, Other).
    5. **Age**: The age of the passenger.
    6. **Nationality**: The country of citizenship of the passenger.
    7. **Airport Name**: The name of the airport from which the passenger departs or arrives.
    8. **Airport Country Code**: The country code where the airport is located.
    9. **Country Name**: The name of the country where the airport is located. 10.**Airport Continent**: The continent where the airport is located.

1. **Continents**: Continent of departure or arrival.
2. **Departure Date**: The date when the ﬂight departs. 13.**Arrival Airport**: The airport where the ﬂight arrives.
3. **Pilot Name**: The name of the pilot operating the ﬂight.
4. **Flight Status**: The current status of the ﬂight (e.g., On Time, Delayed, Cancelled).

3

# Exploring Insights from Airlines Data Analysis using Qlik

### Connect Data with Qlik Sense

To analyze the dataset using Qlik Sense, follow these steps:

#### Extract the Dataset:

* + - * After downloading the dataset, extract the ﬁles to a speciﬁc location on your device.

#### Create a New Qlik Sense App:

* + - * Open Qlik Sense and create a new app named "Exploring Insights from Synthetic Airline Data Analysis."
      * Open the newly created app.

#### Add Data to Qlik Sense:

* + - * Click on "Data Manager."
      * Click on "Add data" and select the dataset ﬁle from the location where it was extracted.

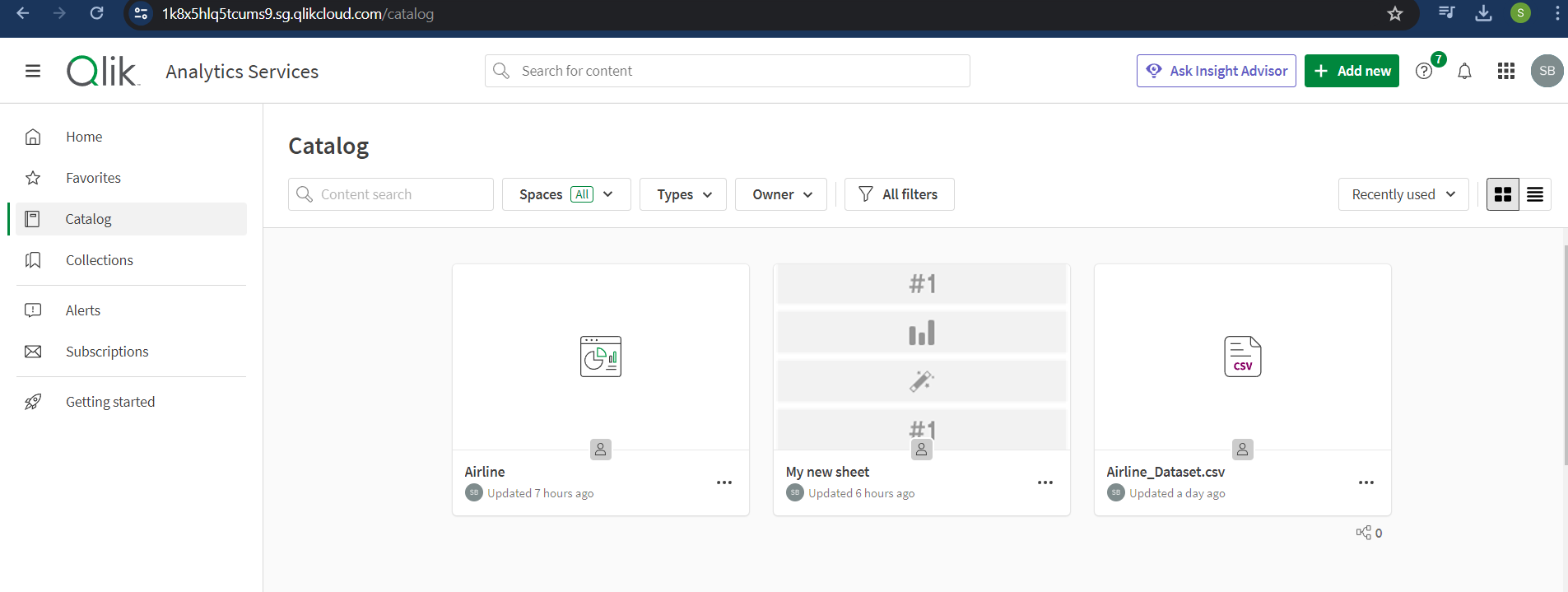
#### Data Integration:

* + - * Ensure that all relevant ﬁelds from the dataset are correctly mapped in Qlik Sense.
      * Check for any inconsistencies or missing values in the dataset and clean the data if necessary.

#### Data Mapping:

* + - * Map ﬁelds such as Passenger ID, First Name, Last Name, Gender, Age, Nationality, Airport Name, Airport Country Code, Country Name, Airport Continent, Departure Date, Arrival Airport, Pilot Name, Flight Status, and Age Group to ensure they are correctly recognized by Qlik Sense for analysis.

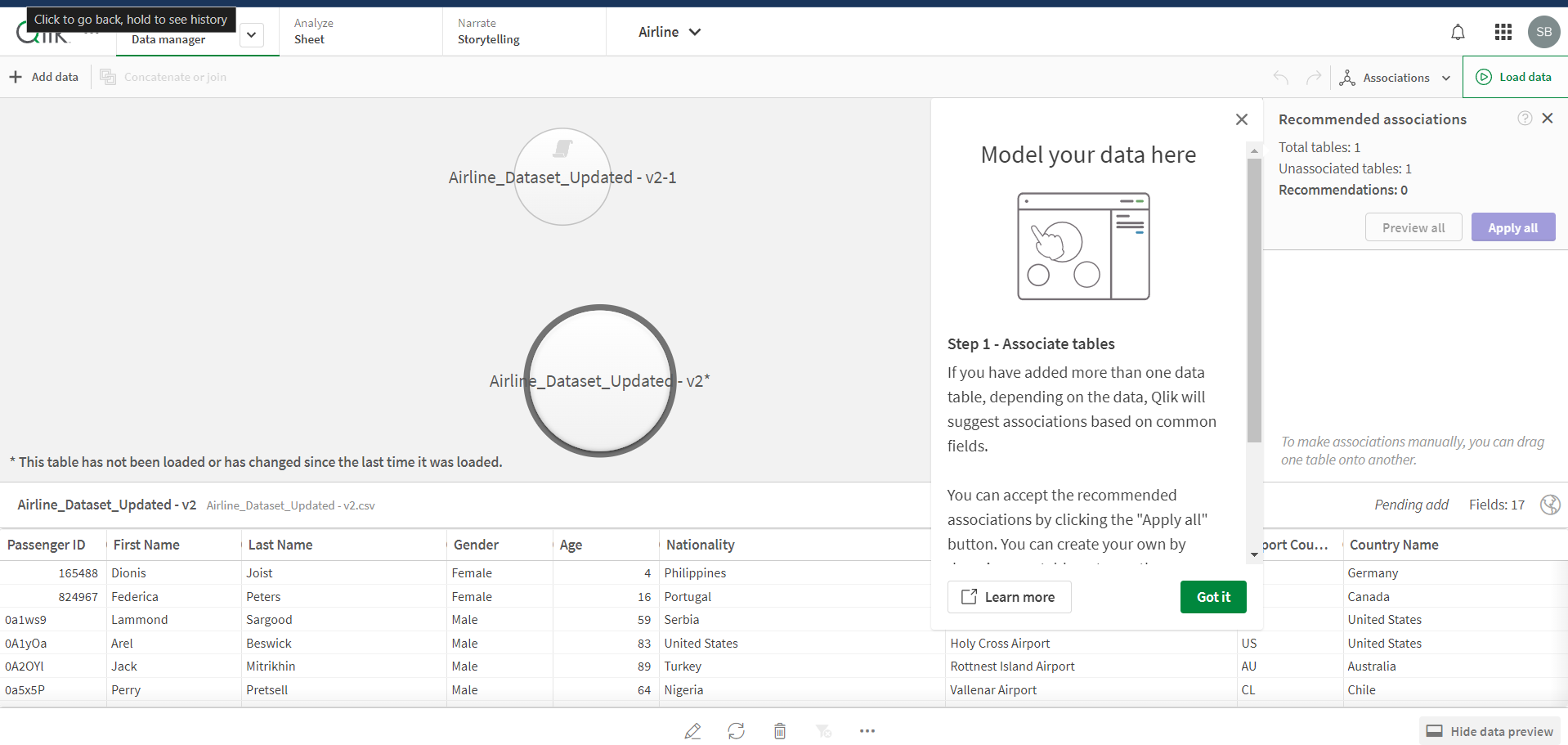
By following these steps, you can successfully integrate and prepare your synthetic airline dataset in Qlik Sense for comprehensive analysis and visualization.



4

# Exploring Insights from Airlines Data Analysis using Qlik

> we need to click on add data on left side top corner .



## Data Preparation

### Prepare the Data for Visualization:

#### Clean the Data

* + - **Remove Inconsistencies:**
      * After downloading the dataset and converting it from CSV to Excel format, inspect the data for inconsistencies and anomalies.
      * Rectify any discrepancies in the data entries to ensure uniformity.

#### Handle Missing Values:

* + - * Identify and address any missing values in the dataset.
      * Fill in missing data points with appropriate values or remove records with substantial missing information.

#### Transform the Data

* + - **Format for Analysis:**
      * Ensure the data is in a suitable format for analysis and visualization.
      * Check that dates, times, and numerical values are correctly formatted.
      * Ensure all ﬁelds are appropriately labeled.

#### Aggregate and Categorize Data

* + - **Remove Extra Columns:**
      * Identify and eliminate unnecessary columns such as 'others' and 'average' that are not relevant to the project's analysis.
      * During the data addition process in Qlik Sense, select only the columns

5

# Exploring Insights from Airlines Data Analysis using Qlik

required for analysis and discard extraneous ones.

#### Remove 'Total' Rows:

* + - * Identify rows that contain "total" values, which are direct additions of each column.
      * Remove these rows from the dataset to ensure that aggregate data does not skew the analysis.

#### Re-upload Cleaned Data:

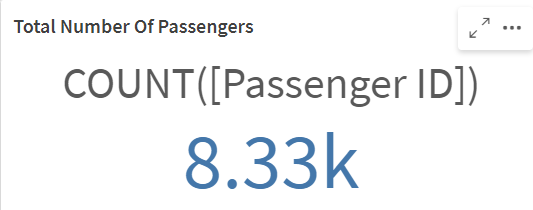
* + - * After cleaning the dataset by removing unwanted columns and rows, re-upload the cleaned ﬁles to Qlik Sense.

#### Data Association

* + - **Qlik Sense Recommendations:**
      * Utilize Qlik Sense's recommendations for data associations to link related data ﬁelds across different tables.
      * Ensure that the data is properly connected and ready for comprehensive analysis.

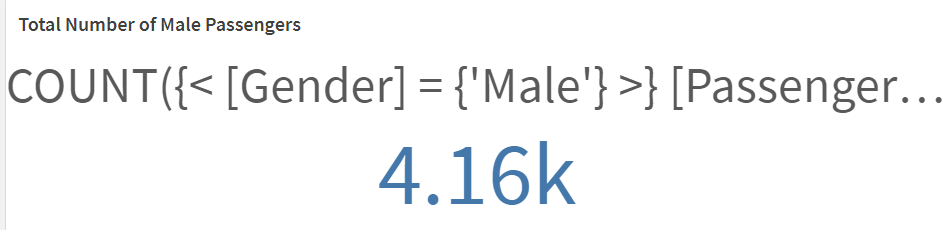
1. **Data Visualizations:**

### Key Performance Indicators (KPIs):



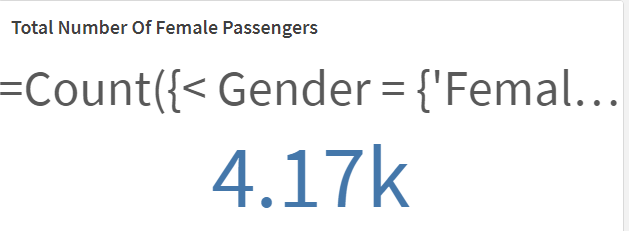
#### Total Number of Male Passengers

* + - * Key Performance Indicator (KPI) visualizations were created to display the total number of Airports.
      * This KPI provides a clear snapshot of the overall male passengers, helping to understand the scale number of airports.



#### Total Number of Female Passengers

* + - * Key Performance Indicator (KPI) visualizations were created to display the total number of Airports.
      * This KPI provides a clear snapshot of the overall female passengers, helping to understand the scale number of airports.



6

# Exploring Insights from Airlines Data Analysis using Qlik

# Line Chart : AgeGroup > 40

# Number of Passengers by Age Group (>40):

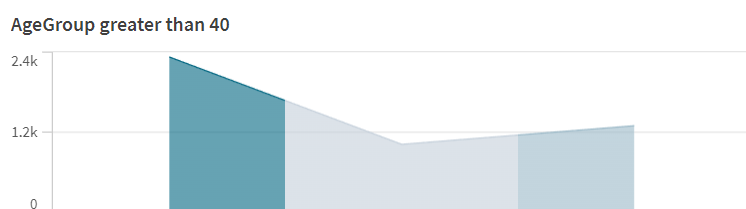
# The area chart displays the count of passengers by different age groups over 40. The age groups listed are Elder, Just plain old, and Midlife.

# Elder has the highest number of passengers (approximately 29k).

# This is followed by the Just plain old group with around 11k passengers.

# The Midlife group has approximately 15k passengers.

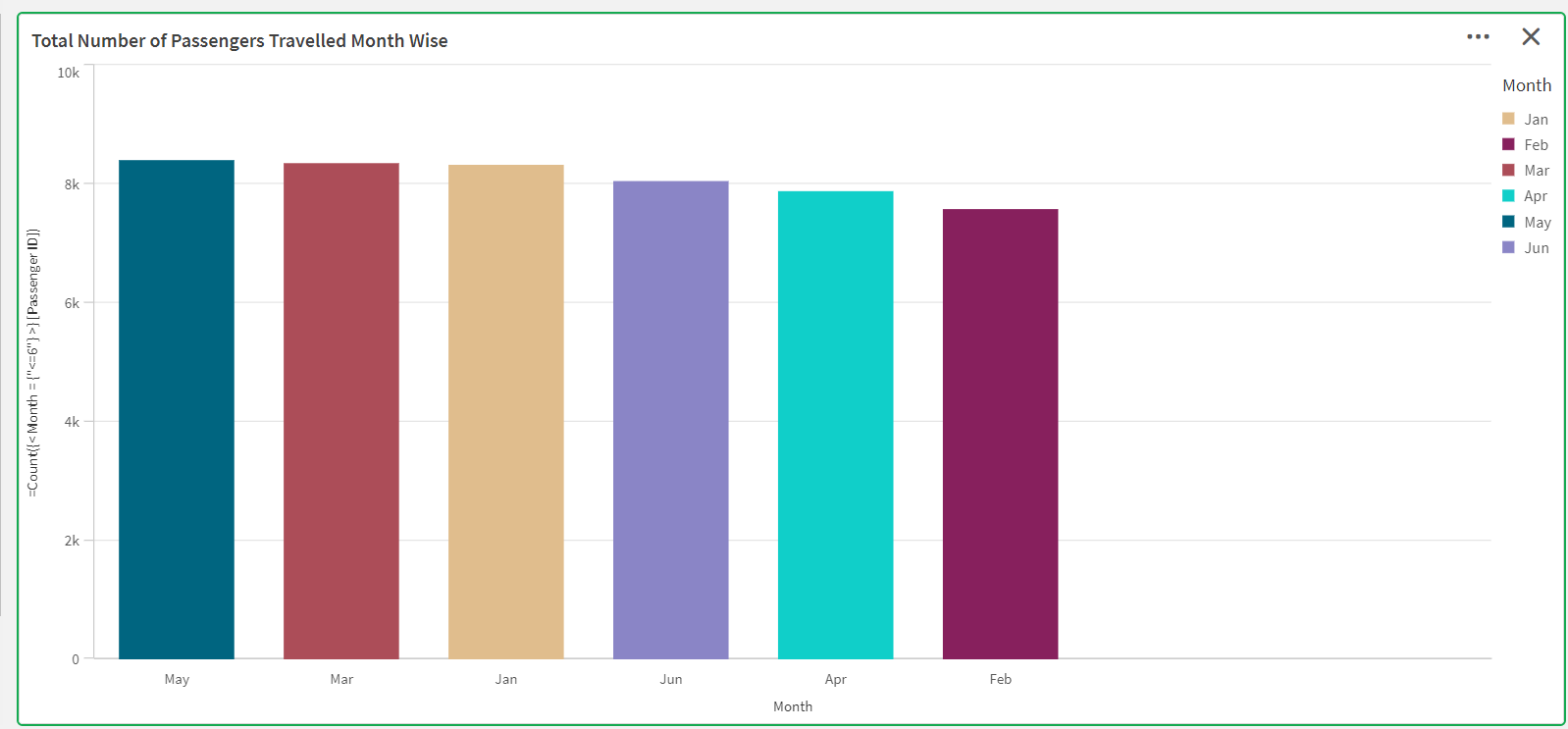
# This visualization is useful for understanding the distribution of passengers in various age groups over 40 and analyzing travel patterns based on age demographics.



#### 5.2 Bar Graph:

●

* **Total Number of Passengers:**Another KPI visualization was created to highlight the total number of passengers travelled in aeroplane.
* This KPI underscores the provides a clear overview of the total number of passengers.



* **Number of Passengers Travelled - Month-wise**
  + The bar chart displays the count of passengers who have traveled each month. Months range from August to February. The y-axis shows the number of passengers, with values ranging from 0 to 10K. Some months have more than 8K passengers, while others have fewer. This visualization is relevant for analyzing seasonal travel patterns or assessing business performance in transportation services over these months.
  + **Insight**: Identiﬁes seasonal travel patterns and business performance trends in transportation services across the speciﬁed months.
  + This bar chart helps stakeholders understand ﬂuctuations in passenger travel volumes, enabling them to make informed decisions regarding capacity planning, marketing strategies, and resource allocation based on observed trends.7

Exploring Insights from Airlines Data Analysis using Qlik

#### Number of Passengers Travelled - Nation wise:

* + The horizontal bar chart displays the count of passengers by nationality.
  + The countries listed are China, Indonesia, Russia, Philippines, Brazil, and Portugal. China has the highest number of passengers

(approximately 13.2k), followed by Indonesia (10.5k), Russia (5.7k), Philippines (5.2k), Brazil, and Portugal (both around 3.3k).

* + This visualization is useful for understanding passenger demographics and travel patterns.
    - **Age Group - Gender wise**:
      * The horizontal bar chart compares the count of females and males across different age groups.
      * The age groups listed are Adult, Baby, Child, Elderly, Infant, Middle Age, Teen, Toddler, and Young Adult.
      * The highest count is in the Adult category for females (approximately 27.4k), followed by males (around 16.4k). This visualization is relevant for demographic studies or marketing analysis.

8

# Exploring Insights from Airlines Data Analysis using Qlik

## Dashboard:

## 

**Exploring Insights from Airlines Data Analysis using Qlik**

travel activity captured in the dataset.

#### Number of Passengers Travelled - Month-wise:

* + Passenger travel peaks during the months of August, May, and July, with each month seeing more than 8.4k passengers.
  + Other months have slightly lower passenger counts, ranging from 7.65k in February to 8.15k in September.

#### Number of Passengers - Nation-wise:

* + China leads signiﬁcantly in passenger numbers with 18.32k passengers.
  + Indonesia follows with 10.56k passengers, and other notable nations include Russia (5.69k), Philippines (5.24k), Brazil (3.79k), and Portugal (3.3k).

#### Age Group - Gender-wise:

* + The elder age group (27.44k) has the highest travel frequency.
  + Adults (16.43k) and the middle-aged group (16.5k) also show high travel activity.
  + Noticeable gender differences exist within age groups, particularly in the elder and middle-aged categories where males are signiﬁcantly more frequent travelers.

#### Recommendations

1. **Resource Allocation**:
   * Allocate additional resources during peak travel months (August, May, and July) to handle the increased passenger ﬂow eﬃciently.
   * Ensure suﬃcient staﬃng and operational readiness at airports to accommodate high passenger volumes.

#### Marketing Strategies:

* + Develop targeted marketing campaigns for nations with lower

passenger numbers (e.g., Brazil, Portugal) to boost travel interest and diversify passenger sources.

* + Explore partnerships or promotions to attract more passengers from these countries.

#### Targeted Campaigns:

* + Develop age-speciﬁc marketing campaigns, especially targeting the elder, adult, and middle-aged groups, considering their high travel

frequency.

* + Tailor marketing messages and services to address the preferences and needs of different gender demographics within these age groups.

#### Actionable Insights

1. **Staﬃng and Scheduling**:
   * Prepare for increased demand in peak months by adjusting staﬃng

13

**Exploring Insights from Airlines Data Analysis using Qlik**

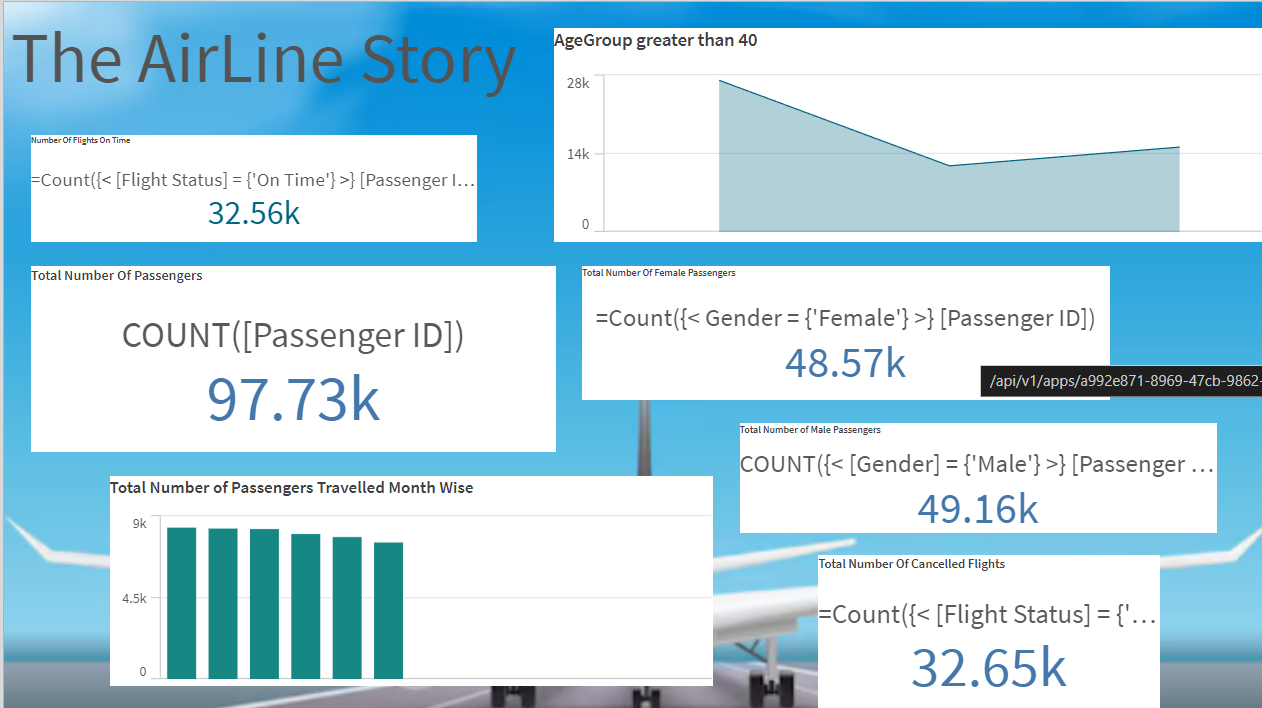
#### Number of Passengers Travelled - Month wise:

* + A vertical bar graph shows passenger counts across different months.
  + August (Aug) and January (Jan) have the highest passenger numbers, while other months exhibit ﬂuctuations.

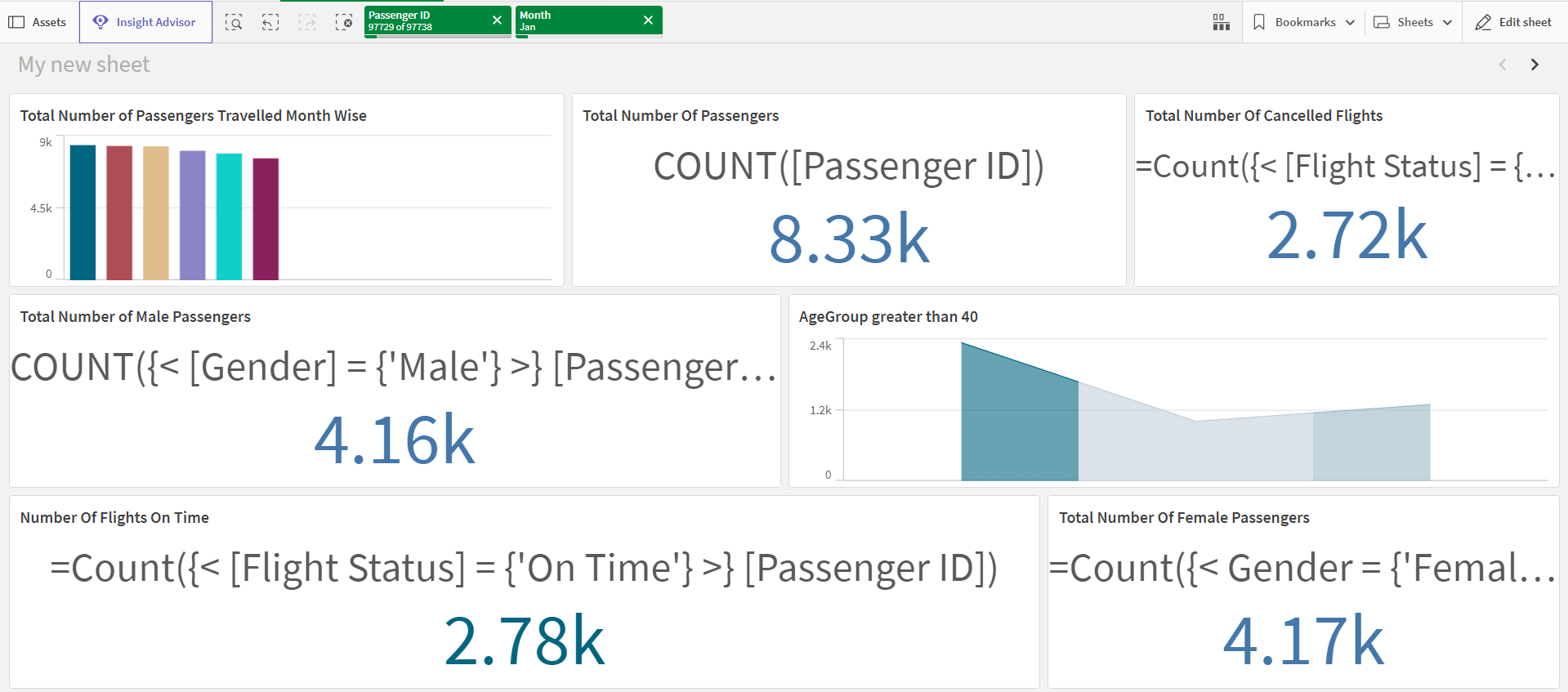
#### Age Group - Gender wise:

* + This stacked bar chart categorizes passengers by age group and further breaks down the distribution by gender.
  + The Adult category has the highest count for both females (approximately 27.4k) and males (around 16.4k).

Story Telling:



DashBoard :



The dashboard provides several key metrics related to passengers and flight statuses for a specific month (January). Here's a summary of the information presented:

A bar chart displays the monthly distribution of the total number of passengers, with values around 9,000.

Total Number of Passengers:

The total count of passengers is 8.33k (8,330).

Total Number of Cancelled Flights:

The total number of canceled flights is 2.72k (2,720).

Total Number of Male Passengers:

The total count of male passengers is 4.16k (4,160).

Total Number of Female Passengers:

The total count of female passengers is 4.17k (4,170).

Number of Flights On Time:

The total number of on-time flights is 2.78k (2,780).

Age Group Greater than 40:

A chart showing the distribution of passengers aged over 40,

Though specific numbers are not provided in the summary.

These metrics give insights into passenger demographics, flight punctuality, and cancellations, helping to understand the performance and customer profile for the airline in January.

By:

SUMAN BHANDARI

14