1. Introduction

1.1 **OverView**:

The project "Exploring Insights from Synthetic Airline Data Analysis with Qlik" focuses on using Qlik's powerful capabilities and data visualisation to analyse a synthetic airline dataset. The Vata considers a variety of airline operations factors, including flight schedules, flight status, performance metrics, and so on.

1.2 **Purpose:**

The project's goal is to use airline data to derive actionable insights that can help improve decision-making processes in the domain. The purpose includes identifying revenue trends, travel periods, peak traffic periods, and customer feedback.

1.3 **Technical Architecture:**

Some Technical Architecture for the Project Are:

1.Excel
2.Qlik Sense
3.Custom Visualisation
4.Story Telling
5.Access Controls

2. Define Problem

2.1 **Specify Business Problem:**

The problem definition is to use Data Analysis Tools like Qlik to optimise airline revenue, increase efficiency, and improve customer experience.

2.2 **Business Requirements:**

1. Data Integration: Add synthetic airline data to Qlik, including flight schedules, passenger demographics, ticket sales, and performance metrics.
2. Data Visualisation: Create interactive dashboards and visualisations to show revenue trends, operational efficiency metrics, and customer satisfaction levels.
3. Predictive Analytics: Use Qlik's predictive analytics tools to forecast demand, identify potential operational bottlenecks, and anticipate customer behaviour.
4. Customer Segmentation Analysis: Use purchasing behaviour, travel frequency, and feedback to tailor marketing and operational strategies.

2.3 **Literature Survey:**

1. Historical Data Analysis: According to research, analysing past ticket sales can reveal peak travel times and popular destinations.1

2. Resource Allocation: According to the literature, predictive analytics can help improve resource allocation by forecasting peak traffic periods.

3. Customer Segmentation: Research shows that segmenting customers based on demographics and purchasing behaviour can help personalise services and improve satisfaction.

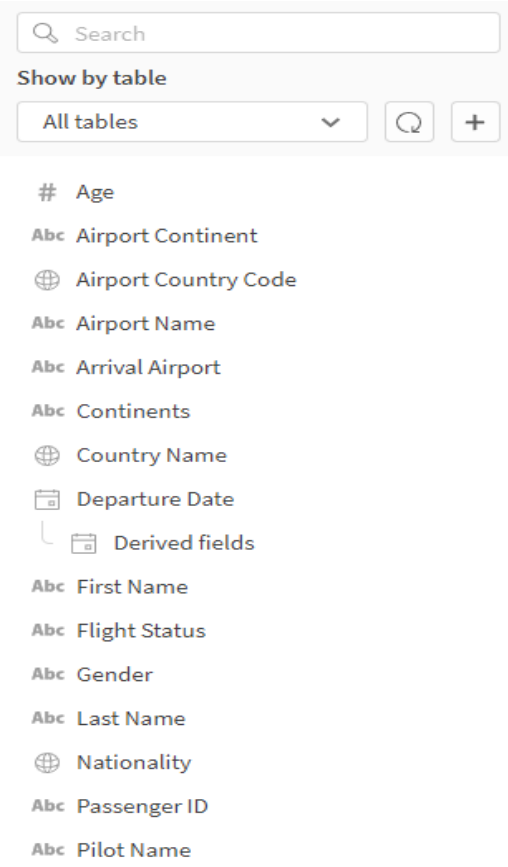
3. Data Collection:

3.1 Collect the Dataset:

Data contains all the meta information regarding the columns described in the CSV files

3.2 Connect Data with Qlik Sense:

effectively connect and visualize synthetic airline data in Qlik Sense, uncovering valuable insights into revenue optimization, operational efficiency, and customer



4. Prepare Data for Visualisation:

4.1 Prepare Data for Visualization:

```
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://DataFiles/Airline Dataset Updated - v2.csv]
(txt, utf8, embedded labels, delimiter is ',', msq);


[Airline Dataset Updated - v2]:

Load *;
```

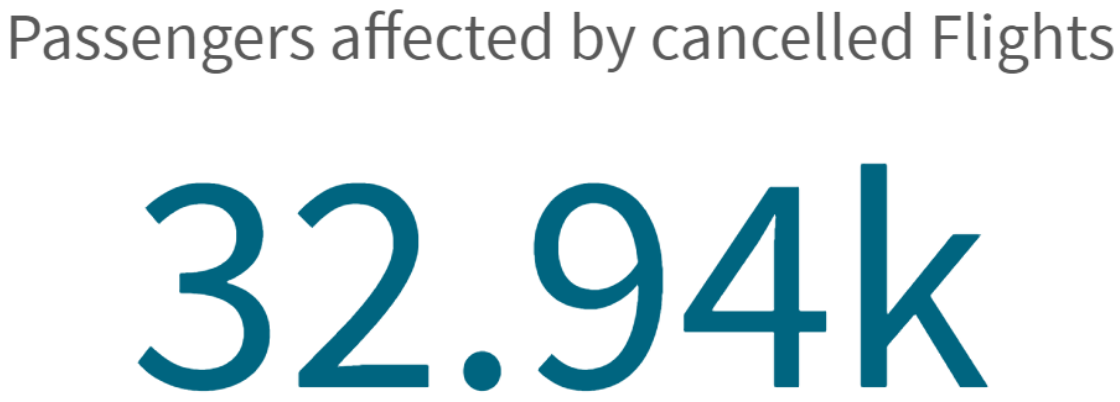
```
[Airline Dataset Updated - v2]:
NOCONCATENATE LOAD *,

if(Age >=0 AND Age <= 1, 'Baby',
    if(Age >= 1 AND Age <= 3, 'Toddler',
    if(Age >= 4 AND Age <= 9, 'Child',
    if(Age >= 10 AND Age <= 12, 'Tween',
    if(Age >= 13 AND Age <= 19,'Teen',
    if(Age >= 20 AND Age <= 24, 'Young Adult',
    if(Age >= 25 AND Age <= 39, 'Adult',
    if(Age >= 40 AND Age <= 54, 'Middle',
    if(Age >= 55 AND Age <= 79, 'Elder',
    if(Age >= 80, 'Just plain old'))))))))) AS AgeGroup,

Date#([Departure Date],'MM/DD/YYYY') as [Departure_Date],
Year([Departure Date]) As Year,
Month([Departure Date]) as Month
RESIDENT [Airline Dataset Updated - v2]
WHERE NOT ([Arrival Airport] = '0' OR [Arrival Airport] = '-');
```

5. Data Visualization

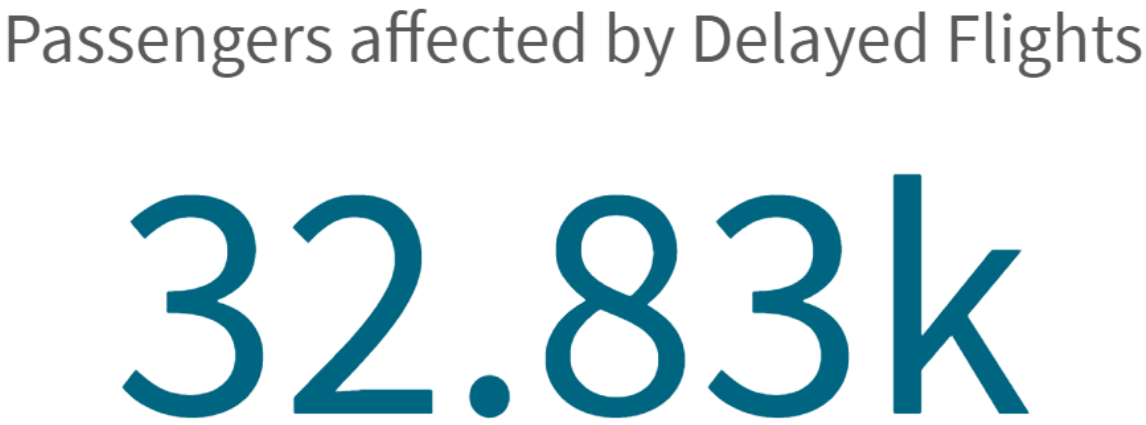
5.1 Visualization

1.Total Number of Passengers:

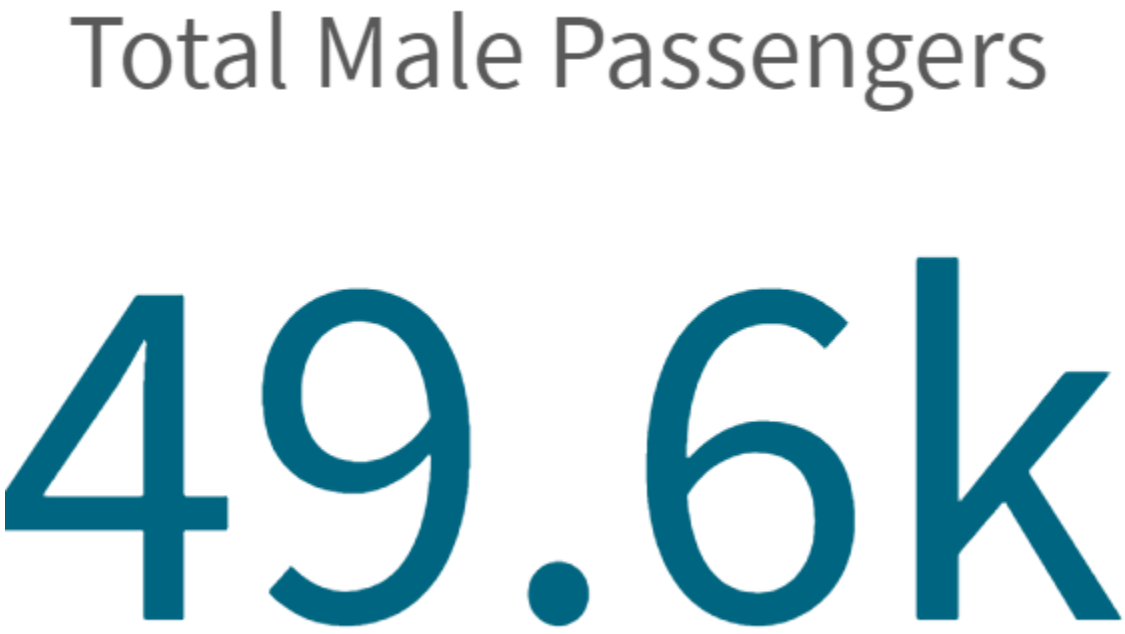
2. Number of Passengers affected by cancellled Flights:



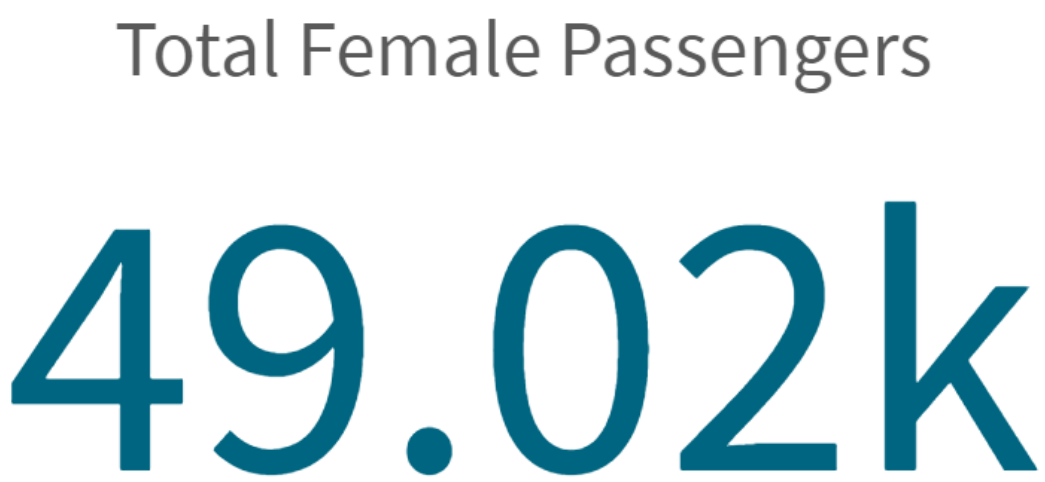
3. Number of Passengers affected by Delayed Flights:



4.Number of Male Passengers:

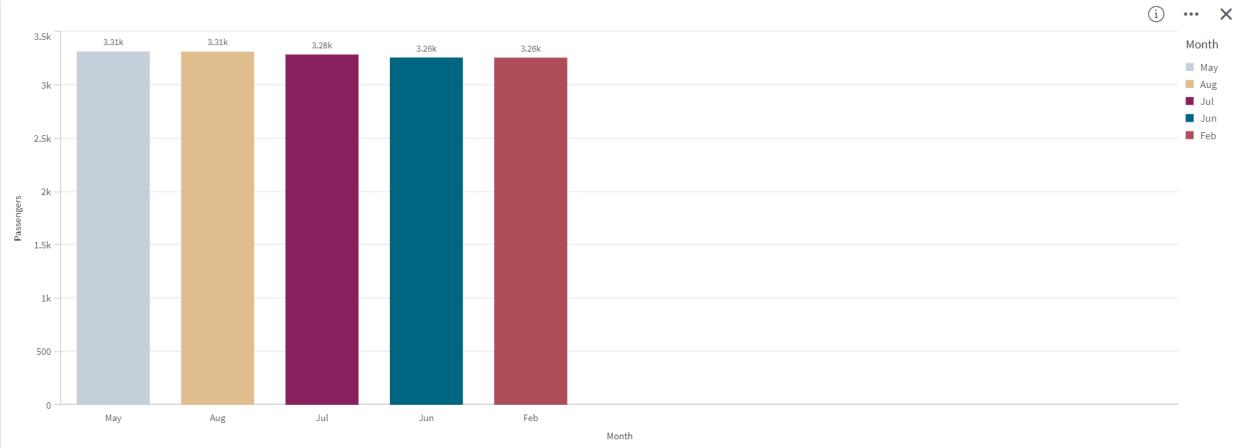


5.Number of Female Passengers:

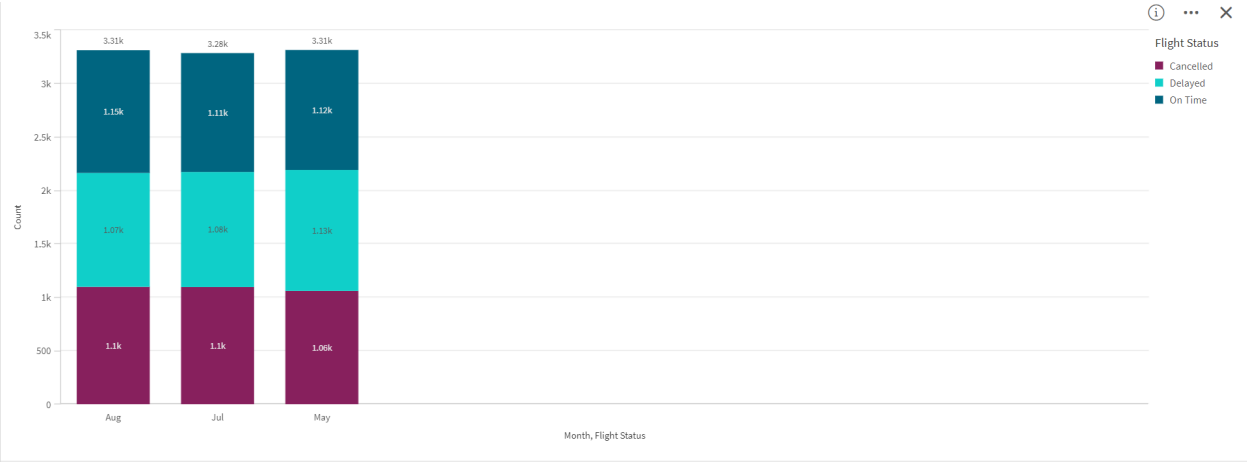


6.Number of Flights on Time:

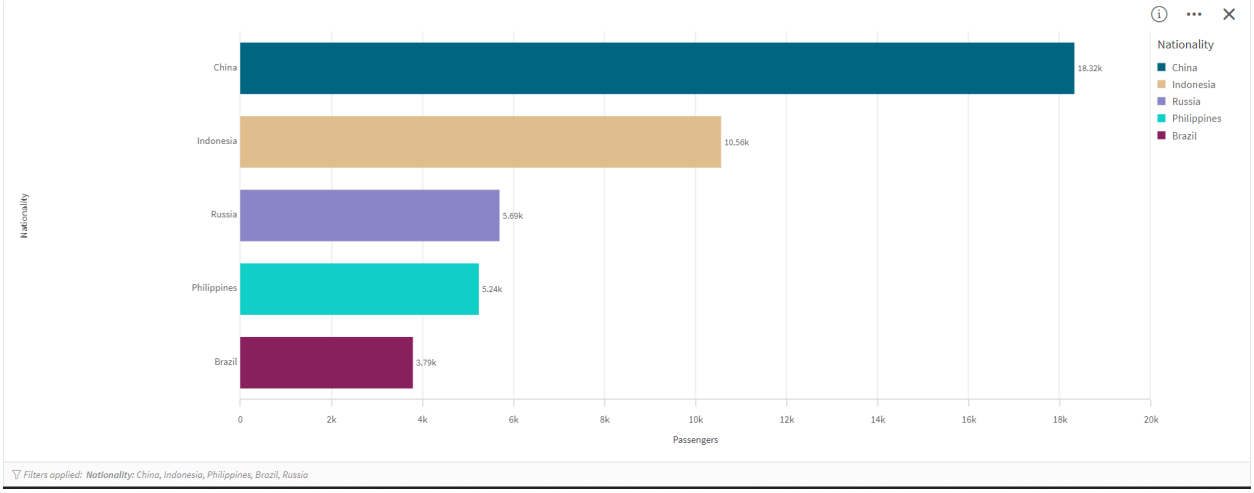
7.Top 5 Months where Passengers Travle the Most:



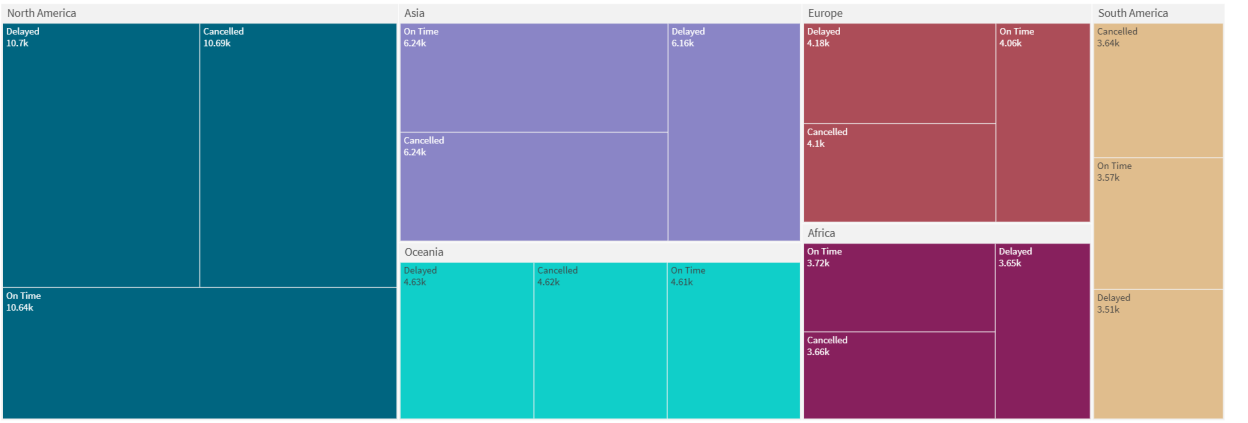
8.Top 3 Months Flight Status Wise:



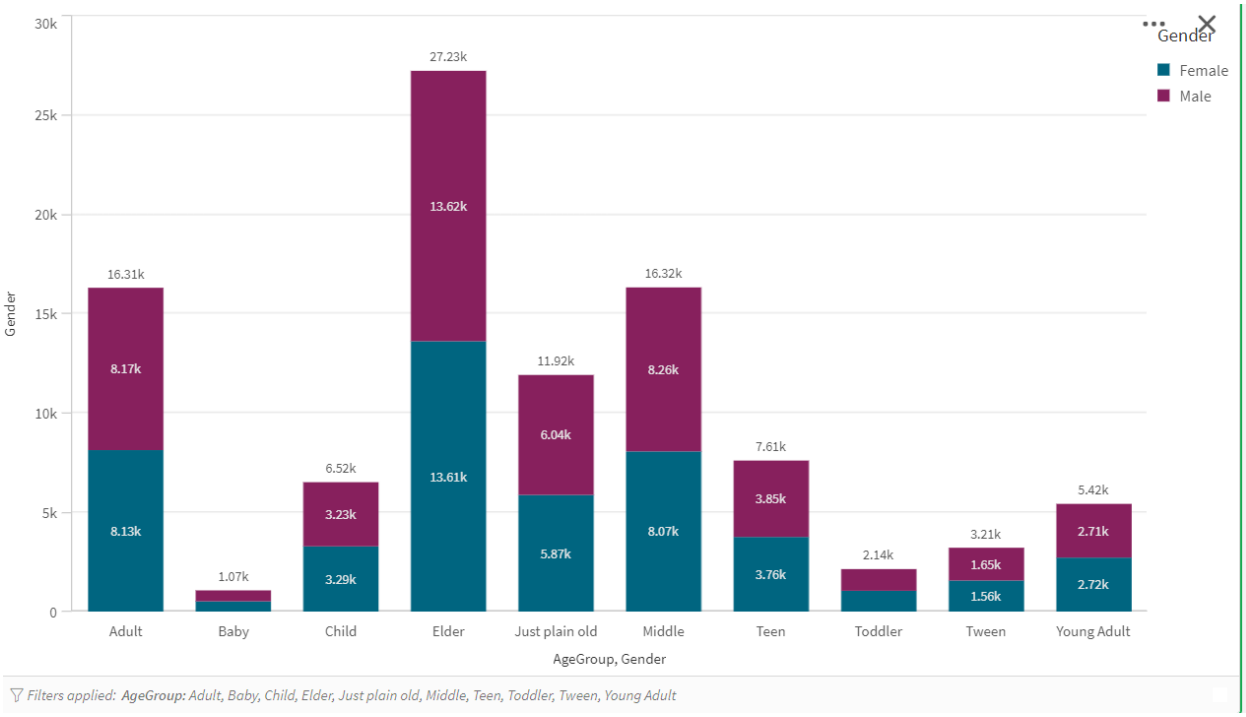
9. Number of Passengers nationality wise:



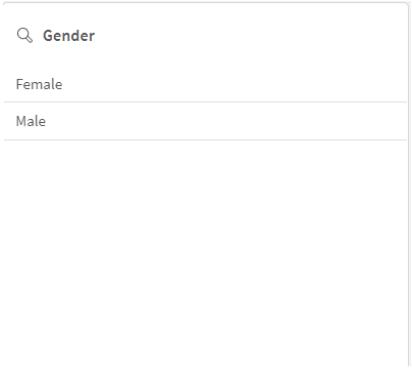
10. Continent Wise Flight Status:



11.Age Group Gender Wise:

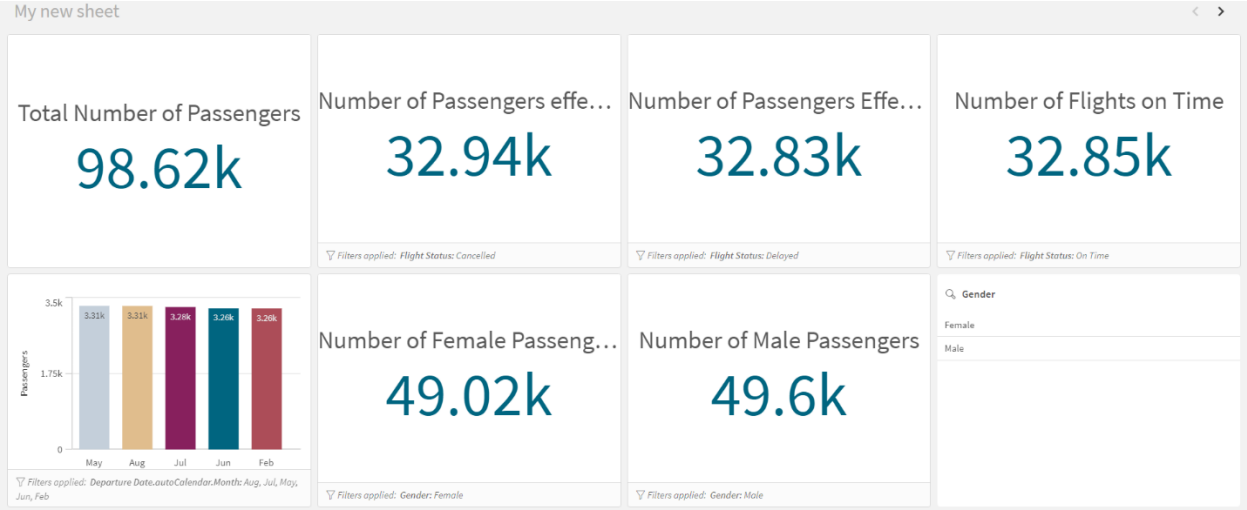


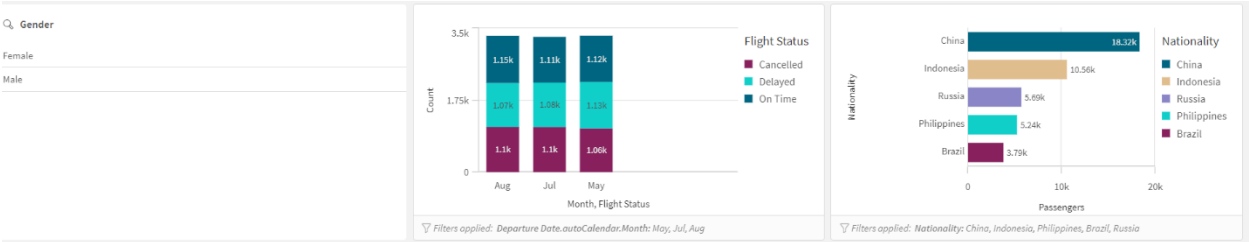
12. Filter Pane Consisting of Gender(M/F):

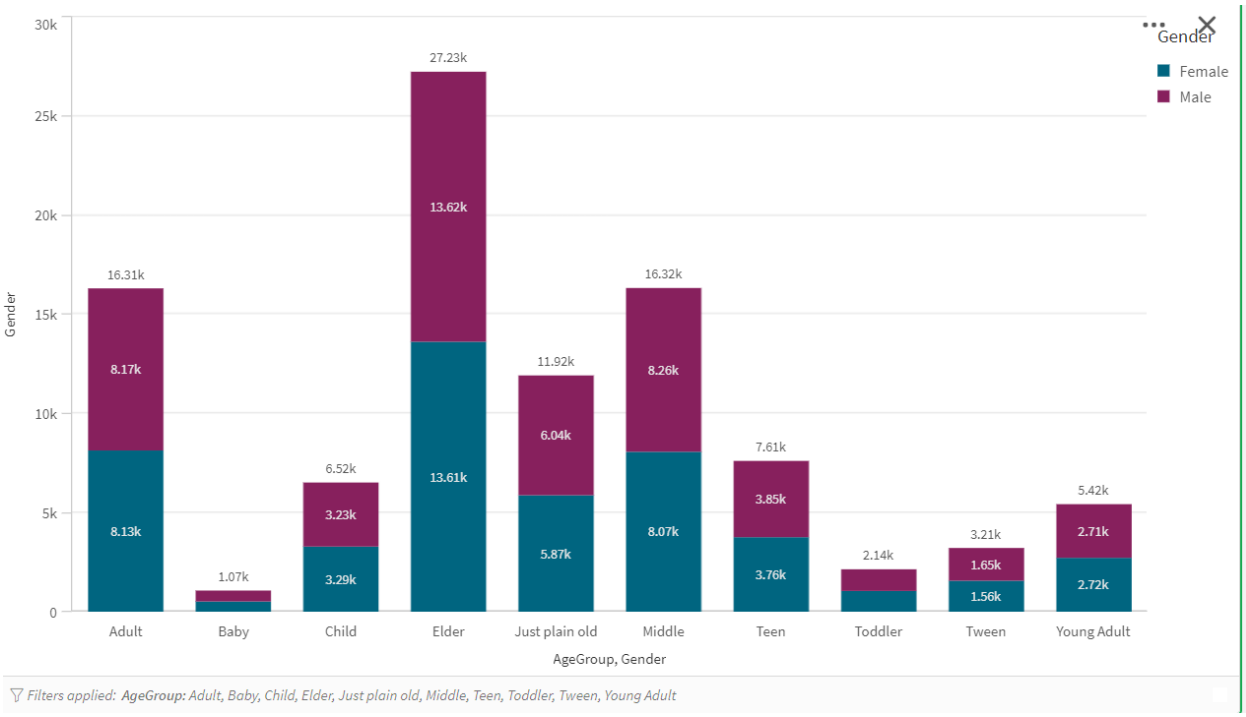


6.Dashboards/Sheets:

6.1 **Responsive and Design of Dashboards**







7.Report:

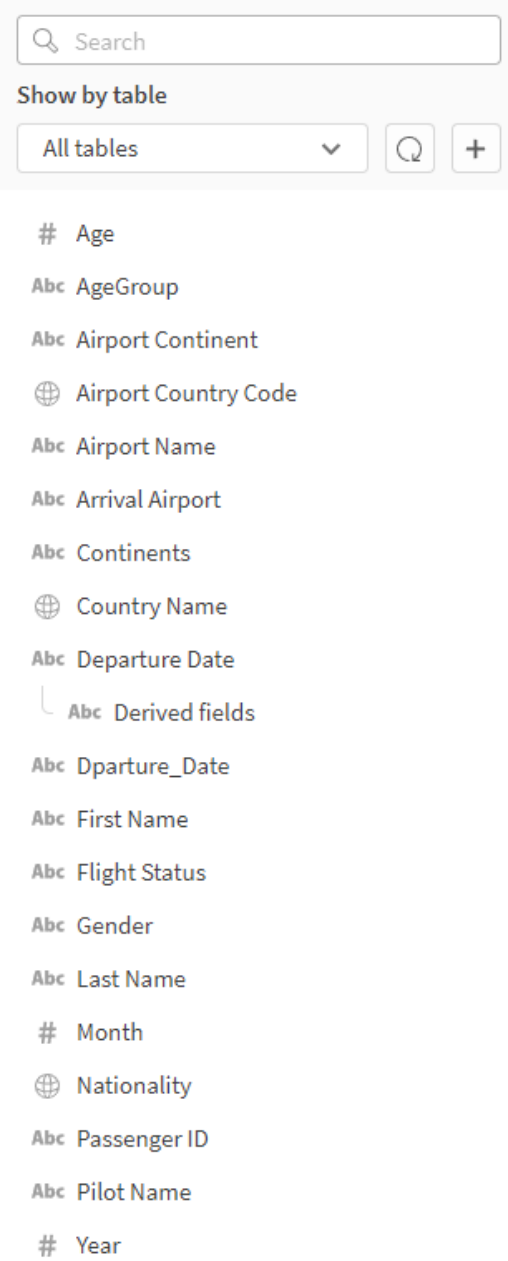
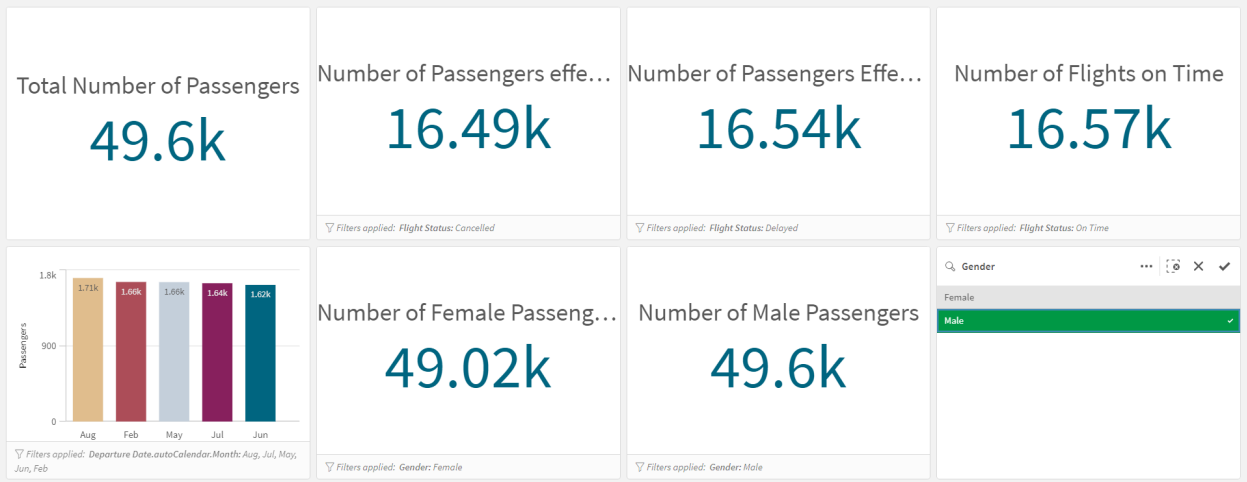
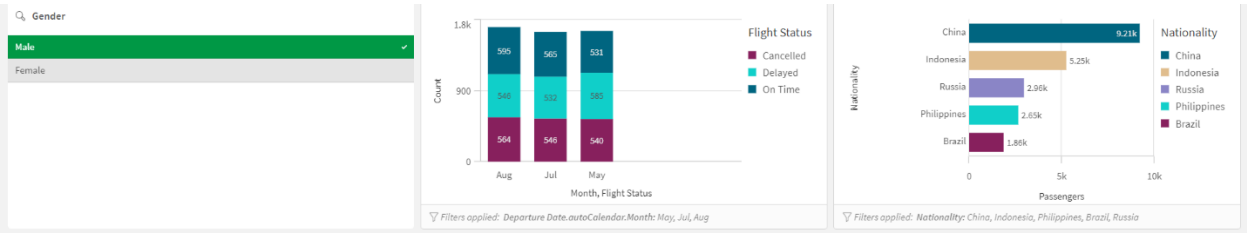
7.1 **Report Creation**

The report uses synthetic data to provide insight into various aspects of airline operations.The visualisations provide various insights into the dataset.Some of the insights are provided below.1. There are 98.62K passengers travelling, with 49.6K males and 49.02K females.2. A total of 32.94K passengers were affected by cancelled flights, 32.83K by delayed flights, and 32.85K by flights that arrived on time.3. The highest number of passengers travelled during the month of May, totaling 3.31k.4. China has the highest number of passengers (18.32K).5. Age Group Elder has the highest number of passengers (27.23K), and Baby

Agegroup is the Minimum Number of Passengers.

8.Performance Testing:

8.1 **Amount of Data Rendered:**



8.2 **Utilization of Data Filters:**