Name: Bhushan Lilhare

Exploring Insights From Synthetic Airline Data Analysis With Qlik

1. INTRODUCTION:

1.1 OVERVIEW:

The project, titled "Exploring Insights from Synthetic Airline Data Analysis with Qlik," aims to extract relevant insights from synthetic airline data using Qlik, a powerful business intelligence and data visualisation platform. This synthetic dataset mimics several aspects of airline operations, including aircraft schedules, passenger demographics, ticket sales, and performance measures, giving a solid base for study.

Scenario 1: Revenue Optimization

For airlines aiming to boost revenue, analyzing historical ticket sales data is crucial. Qlik can be used to visualize revenue trends over time, identify peak travel periods, and popular destinations. Additionally, customer segmentation based on purchasing behavior can inform dynamic pricing strategies, ultimately maximizing profitability.

Scenario 2: Operational Efficiency

Airport authorities can leverage Qlik to improve operational efficiency by examining flight schedules, passenger flows, and luggage handling processes. This analysis helps identify operational bottlenecks, predict peak traffic periods, and optimize resource allocation, thereby enhancing overall efficiency.

Scenario 3: Customer Experience Enhancement

Improving passenger experience is a priority for airlines. By integrating customer feedback and performing sentiment analysis using Qlik, airlines can pinpoint areas needing improvement, personalize services, and tailor marketing efforts. This leads to increased customer satisfaction and loyalty.

1.2 PURPOSE

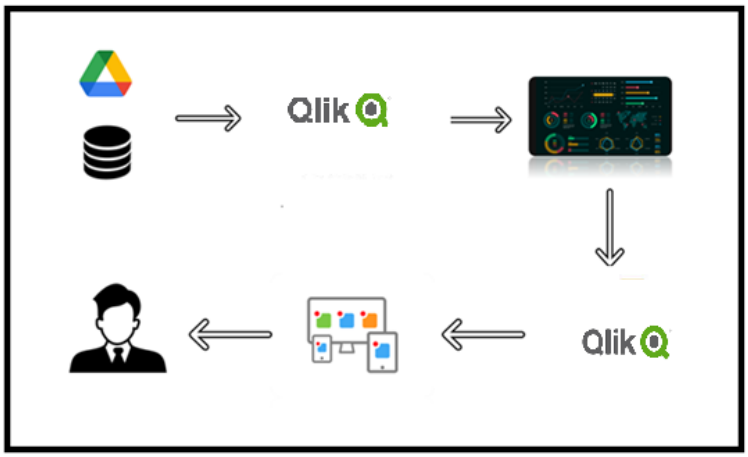
Project Goal: Using Qlik to Investigate Insights from Synthetic Airline Data Analysis

Name: Bhushan Lilhare

The goal of this project is to analyse synthetic airline data using Qlik, a potent business intelligence and data visualisation tool. By doing this, the team hopes to find important patterns and insights in the data that will help airports and airlines make decisions. In particular, the project aims to:

1. Maximize Revenue: Examine past ticket sales data to determine the busiest travel periods, well-liked locations, and sensible price policies.
2. Improve Operational Efficiency: Identify bottlenecks and enhance resource allocation by looking at flight schedules, passenger flows, and luggage handling procedures.
3. Improve Customer Experience: Airlines can customise services and improve customer experiences by using sentiment analysis on customer input to learn preferences and pain areas

1.3 TECHNICAL ARCHITECTURE



Technical Architecture Overview

1. Data Sources:

The architecture starts with a variety of data sources. In this scenario, the data is obtained from both a cloud storage service (Google Drive) and a regular database.These data sources contain raw data that must be analysed and visualised.

2. Data Integration with Qlik:

The raw data from the cloud storage and database is loaded into Qlik. Qlik serves as a central platform for data integration.
During this phase, Qlik extracts, transforms, and loads (ETL) the data to prepare it for analysis.

3. Data processing and visualisation:

Once connected into Qlik, the data is processed to produce useful insights. Qlik creates data visualisations and dashboards that are interactive and easy to understand. This is represented by the dashboard icon.

4. User Access and Interactions:

The visualized data is accessible to end consumers through various platforms, such as PCs, tablets, and smartphones, ensuring flexibility and accessibility for data interaction from any location.Users have the convenience of accessing their Qlik interface from any device to analyze and view data dashboards.

5. Feedback Loops:

Through interactions with visualizations, users make data-driven decisions, with their insights and interactions sent back into the system.This feedback loop aids in refining and enhancing data processing and visualization techniques.

2. PROBLEM UNDERSTANDING

2.1 BUSINESS PROBLEM

The aviation sector is a complex one, with airlines and airports facing several issues in optimising revenue, increasing operating efficiency, and improving customer experience. Traditional data analysis approaches frequently fail to provide actionable insights due to the sheer volume and complexity of airline operations data.

Exploring Insights from Synthetic Airline Data Analysis with Qlik" addresses the business problem of effective data utilisation to support educated decision-making processes. Airlines and airports must sift through massive datasets containing flight schedules, passenger demographics, ticket sales, and performance indicators in order to discover important patterns, trends, and relationships.

This project will use Qlik's analytical skills to address this difficulty by analysing synthetic airline data. By visualising and interpreting the data, stakeholders can get knowledge about revenue optimisation techniques, operational bottlenecks, and customer preferences. Finally, the goal is to provide airlines, airports, and other aviation stakeholders with actionable data that enhance efficiency, profitability, and customer pleasure in the industry.

2.2 BUSINESS REQUIREMENTS

Name: Bhushan Lilhare

1. Synthetic Airline Data Generation: Create synthetic data mimicking various facets of airline operations, including flight schedules, passenger demographics, ticket sales, and performance indicators.

2. Qlik Integration and Utilization: Incorporate Qlik, a robust business intelligence and data visualization tool, to analyze and interpret the synthetic airline data effectively.

3. Analytical Exploration: Leverage Qlik's analytical capabilities to identify crucial patterns, trends, and relationships within the airline data, aiding decision-making processes.

4. Specific Business Scenario Analysis: Address key business scenarios such as revenue optimization, operational bottleneck identification, and customer preference analysis through tailored analysis using Qlik.

5. Accessibility and Flexibility: Ensure that the visualized data is accessible across multiple platforms (PCs, tablets, smartphones) to allow stakeholders to interact with the data from any location.

6. Feedback Integration: Establish mechanisms to incorporate user insights and interactions from the visualizations back into the system, facilitating continuous refinement and improvement of data analysis techniques.

2.3 LITERATURE SURVEY

In the pursuit of leveraging data analytics to enhance decision-making processes within the aviation industry, relevant literature provides valuable insights across various domains:

1. **Business Intelligence and Analytics (Chen et al., 2012):** This foundational work underscores the transformative potential of business intelligence and analytics, highlighting their role in extracting valuable insights from vast datasets.
2. **Qlik's Analytical Capabilities (QlikTech International AB, 2019):** Qlik Sense is identified as a potent data analytics platform, offering powerful data visualisation and analytical capabilities essential for uncovering insights from synthetic airline data.

3. **Airline Networks and Revenue Optimization (Cook & Goodwin, 2008; Phillips, 2005):** Insights from these sources shed light on strategic methodologies for optimising revenue within airline networks through dynamic pricing strategies, aligning with the project's objective of revenue optimization.

4. **Airport Operations and Efficiency (Bazargan, 2016; Santos & Robin, 2010):** Studies on airline operations and resource allocation offer valuable

frameworks for enhancing operational efficiency, aligning with the project's aim to streamline airport operations using Qlik analytics.

5. **Customer Experience Enhancement (Pang & Lee, 2008; Smith & Wheeler, 2002):** Literature on sentiment analysis and customer experience management provides insights into leveraging customer feedback data to enhance passenger satisfaction and loyalty, aligning with the project's focus on customer experience enhancement.

6. **Synthetic Data Generation (Rubin, 1993; Gupta & Verma, 2019):** Exploration of synthetic data generation techniques offers insights into creating representative datasets for analysis, facilitating the project's utilization of synthetic airline data for uncovering actionable insights.

3. DATA COLLECTION

3.1 DATASET COLLECTION

Airline data is critical because it provides insights into the operation and effectiveness of the aviation industry. It provides vital information on aircraft routes, schedules, passenger demographics, and preferences, allowing airlines to optimise operations and improve customer experiences. Airlines can recognise patterns in data on delays, cancellations, and performance and apply strategies to improve punctuality and reduce disruptions. Furthermore, regulatory organisations and legislators rely on this data to maintain safety standards, enforce rules, and make sound choices about aviation policies. Researchers and analysts use airline data to investigate market trends, assess environmental implications, and develop plans for long-term industry success. To summarise, data is essential for informed decision-making, operational excellence, and overall advancement in the aviation business.

CONTENT

This dataset comprises diverse parameters relating to airline operations on a global scale. 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. These columns collectively provide comprehensive insights into passenger demographics, travel details, flight routes, crew information, and flight statuses.This dataset can be used to analyse trends in passenger behaviour, optimize travel experiences, evaluate pilot performance, and enhance overall flight operations.

Name: Bhushan Lilhare

DATASET GLOSSARY

- **Passenger ID** - Unique identifier for each passenger
- **First Name** - First name of the passenger
- **Last Name** - Last name of the passenger
- **Gender** - Gender of the passenger
- **Age** - Age of the passenger
- **Nationality** - Nationality of the passenger
- **Airport Name** - Name of the airport where the passenger boarded
- **Airport Country Code** - Country code of the airport's location
- **Country Name** - Name of the country the airport is located in
- **Airport Continent** - Continent where the airport is situated
- **Continents** - Continents involved in the flight route
- **Departure Date** - Date when the flight departed
- **Arrival Airport** - Destination airport of the flight
- **Pilot Name** - Name of the pilot operating the flight
- **Flight Status** - Current status of the flight (e.g., on-time, delayed, canceled)

3.2 DATA CONNECTION WITH QLIK SENSE

1. **Integration:**
   - Connect the dataset to Qlik Sense by importing the CSV or database file containing the airline data.
   - Ensure the data fields align correctly with the dataset glossary to maintain data integrity and accuracy.
2. **Data Loading:**
   - Use Qlik Sense's data load editor to load the dataset into the application.
   - Verify that each column, such as Passenger ID, First Name, Last Name, and Flight Status, is correctly mapped to its respective field in Qlik Sense.
3. **Data Modeling:**
   - Create associations between relevant data fields to enable comprehensive analysis.
   - For example, link Passenger ID with First Name, Last Name, Gender, Age, and Nationality to analyze passenger demographics.
4. **Dashboard Creation:**
   - Utilize Qlik Sense’s visualization tools to create interactive dashboards that display key insights from the dataset.
   - Include charts and graphs that show trends in passenger demographics, flight routes, and flight statuses.
5. **Analysis and Insights:**
   - Leverage Qlik Sense’s analytical capabilities to uncover patterns and trends within the data.

- Analyze flight delays, cancellations, and on-time performance to identify areas for operational improvement.

6. Decision Support:
   - Use the insights derived from Qlik Sense to support informed decision-making processes.
   - Optimize flight schedules, enhance customer experiences, and improve overall operational efficiency based on data-driven insights.

By connecting this rich dataset to Qlik Sense, stakeholders can gain valuable insights into airline operations, leading to enhanced decision-making and overall industry advancement.

4. DATA PREPARATION

4.1 DATA PREPARATION FOR VISUALIZATION

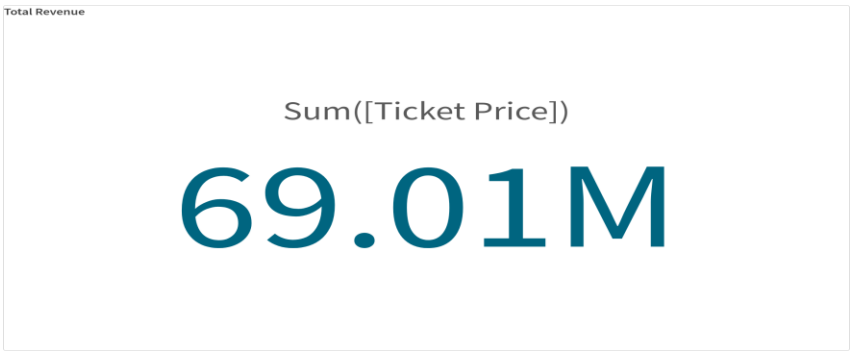
1. Data Cleaning:
   - Although the data is already cleaned, typically, this step involves removing any irrelevant or missing data to ensure accuracy and completeness. Clean data forms the foundation for meaningful visualizations.
2. Data Transformation:
   - Transform the data into formats suitable for visualization. This may include converting categorical data into numerical values, standardizing date formats, and aggregating data to highlight significant metrics.
3. Data Exploration:
   - Explore the data to identify patterns, trends, and anomalies. Understanding these elements can help in deciding which aspects of the data to highlight in visualizations.
4. Data Filtering:
   - Focus on specific subsets of data that are relevant to the analysis objectives. This helps in narrowing down the data to the most impactful information.
5. Preparing Data for Visualization Software:
   - Ensure the data is formatted and structured to be compatible with visualization tools like Qlik. This includes organizing data fields and setting up data hierarchies.
6. Ensuring Data Accuracy and Completeness:
   - Double-check the data to ensure it is accurate and complete, which is crucial for generating reliable insights from visualizations.

By following these steps, the data becomes easily understandable and ready for creating visualizations. This process aids in gaining insights into performance and efficiency, ultimately supporting informed decision-making in the aviation industry.

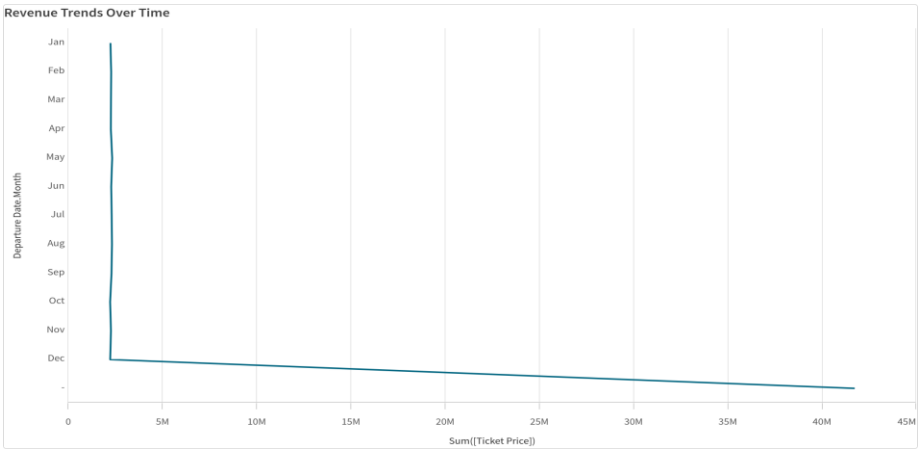
5.1 DATA VISUALIZATION

5.1.1. Total Revenue of Tickets sold

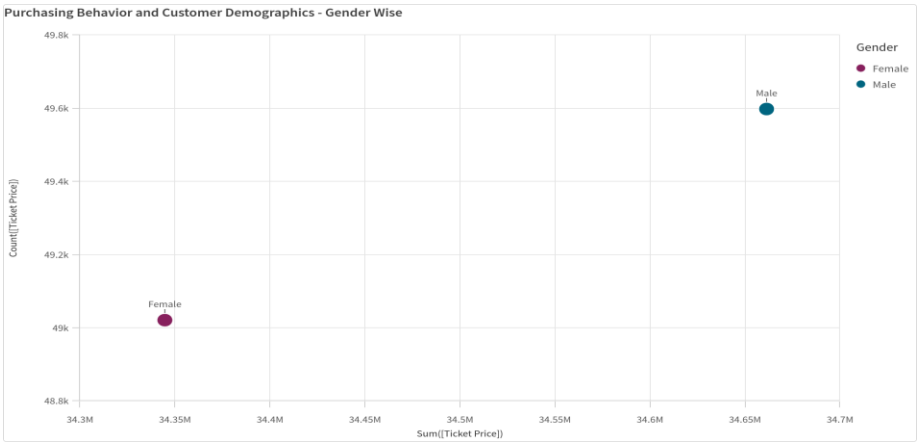
Name: Bhushan Lilhare



5.1.2. Revenue Trends Over Time

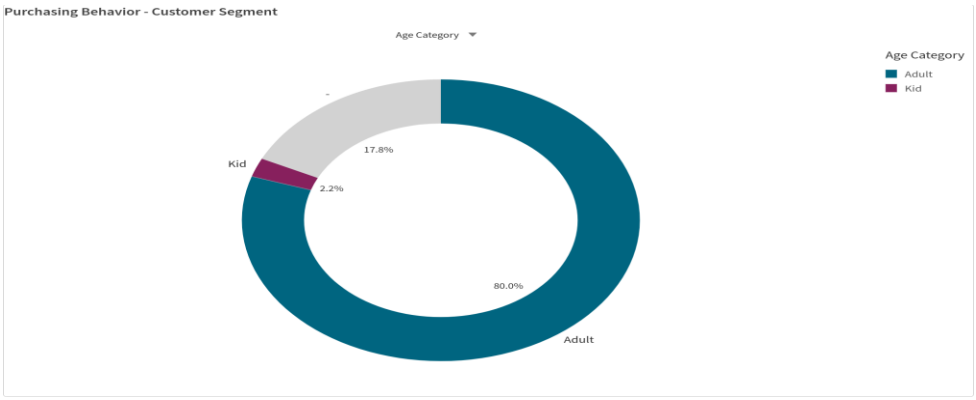


5.1.3. Purchasing Behavior and Demographics

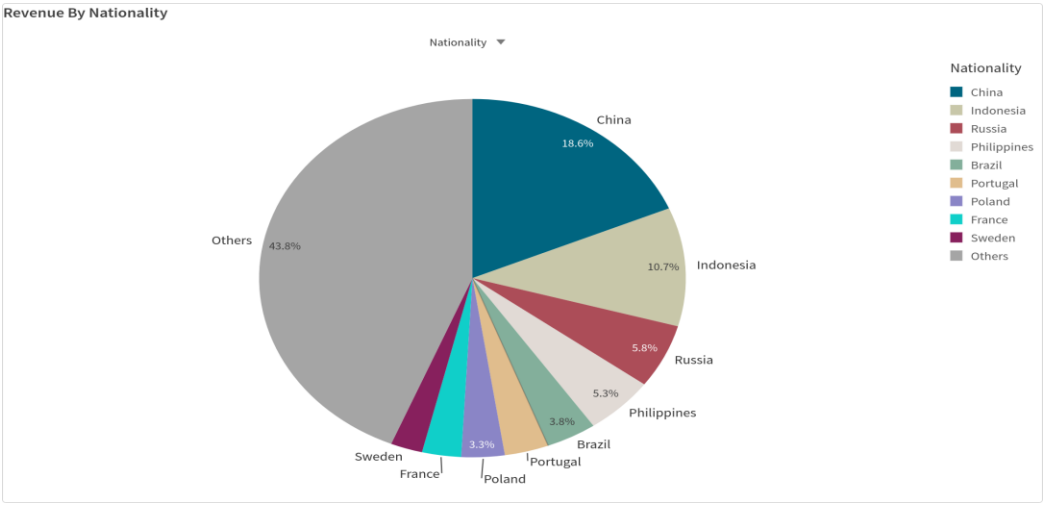


5.1.4. Customer Segmentation

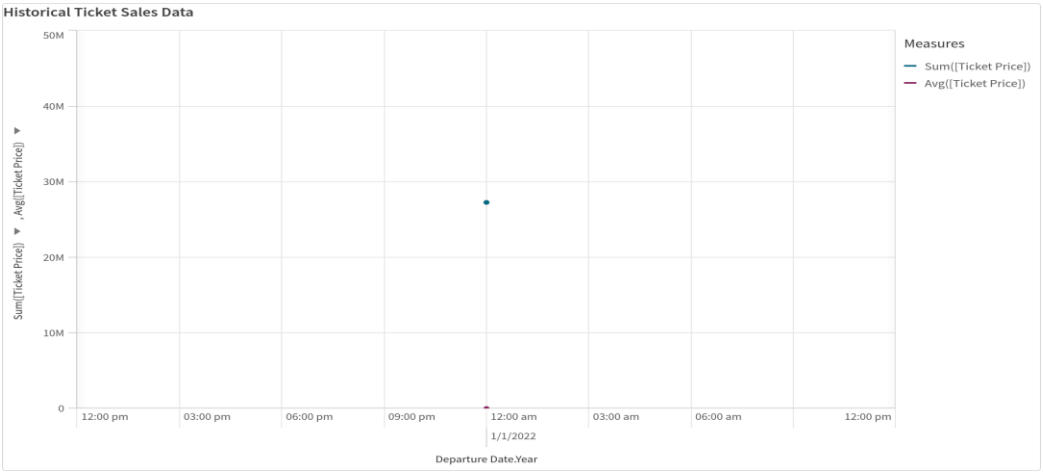
Name: Bhushan Lilhare



5.1.5. Revenue By Nationality



5.1.6. Ticket Sales Data History



5.2.1. Highest Average Tickets

Name: Bhushan Lilhare

Highest Average Ticket by Continents

Calculated measure (KPI)

- The Avg([Ticket Price]) is 699.7.

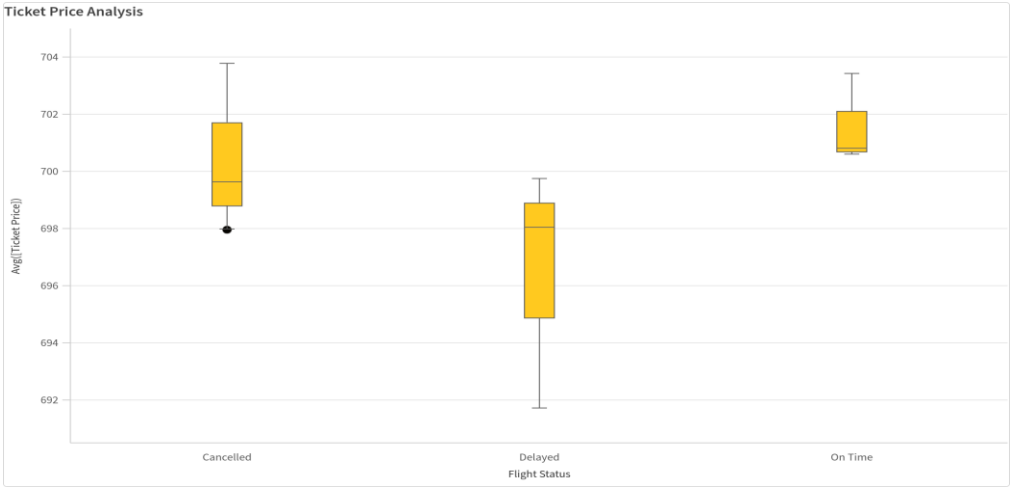
Ranking

- The Avg([Ticket Price]) is 699.7234703251949.
- The top Continents is South America with Avg([Ticket Price]) average equal to 700.8368520246302.

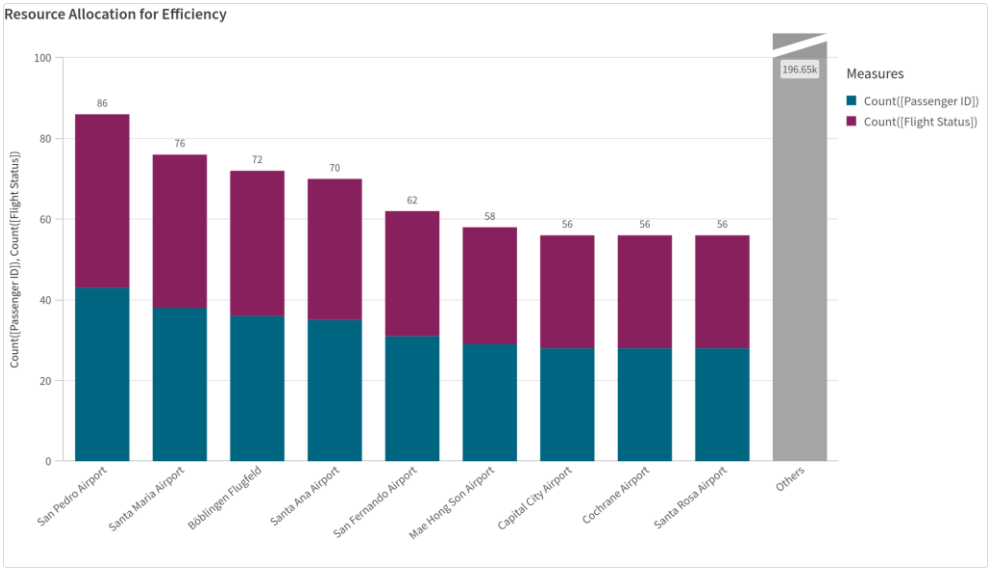
Mutual information

- The mutual dependence between Ticket Price and Continents is 0%.

5.2.2. Ticket Price Analysis

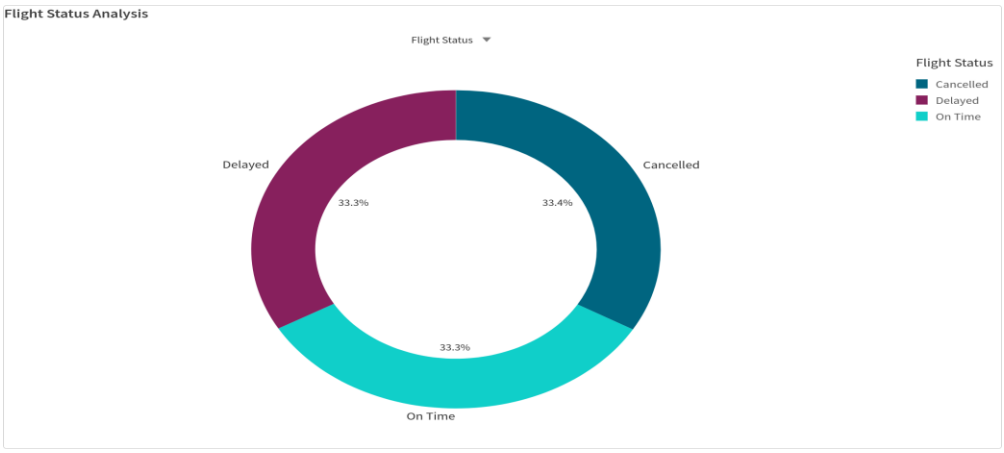


5.2.3. Resource Allocation for Efficiency

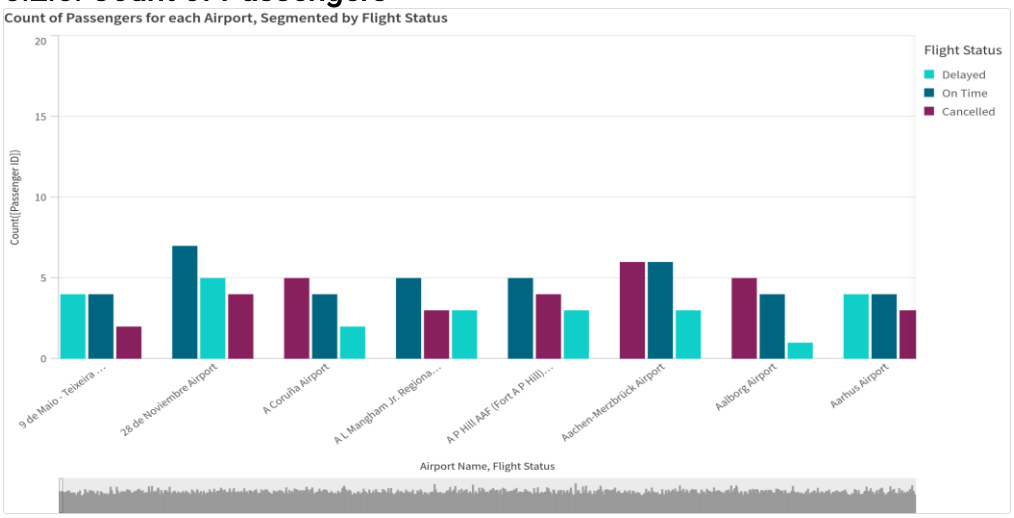


5.2.4. Flight Status Analysis of Airline

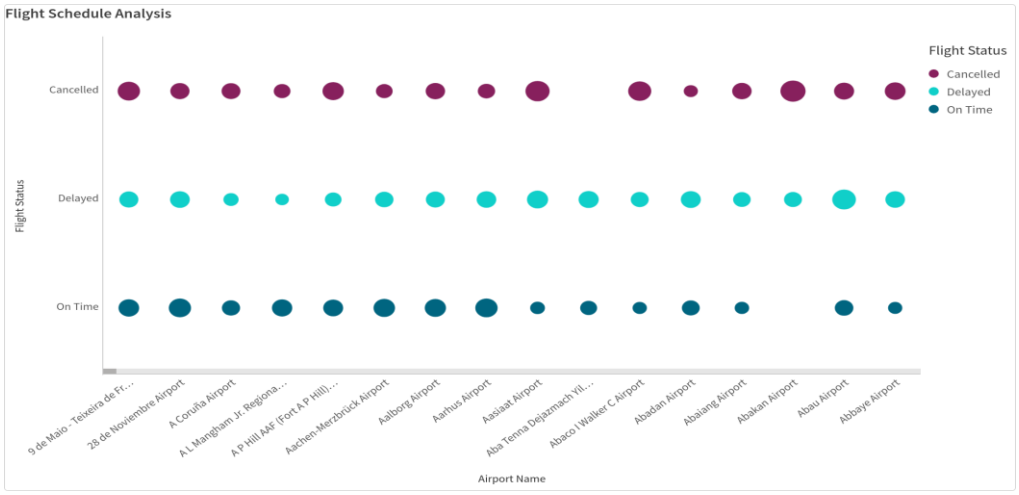
Name: Bhushan Lilhare



5.2.5. Count of Passengers

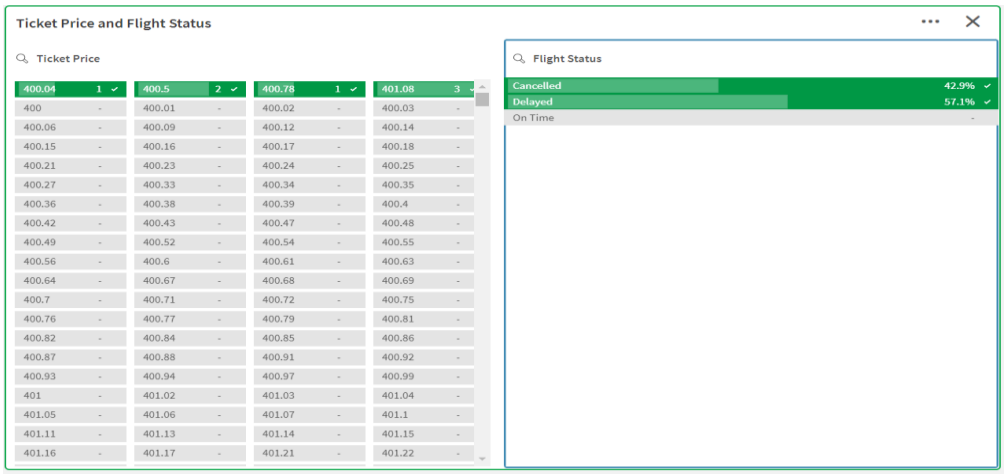


5.2.6. Flight Schedule Analysis

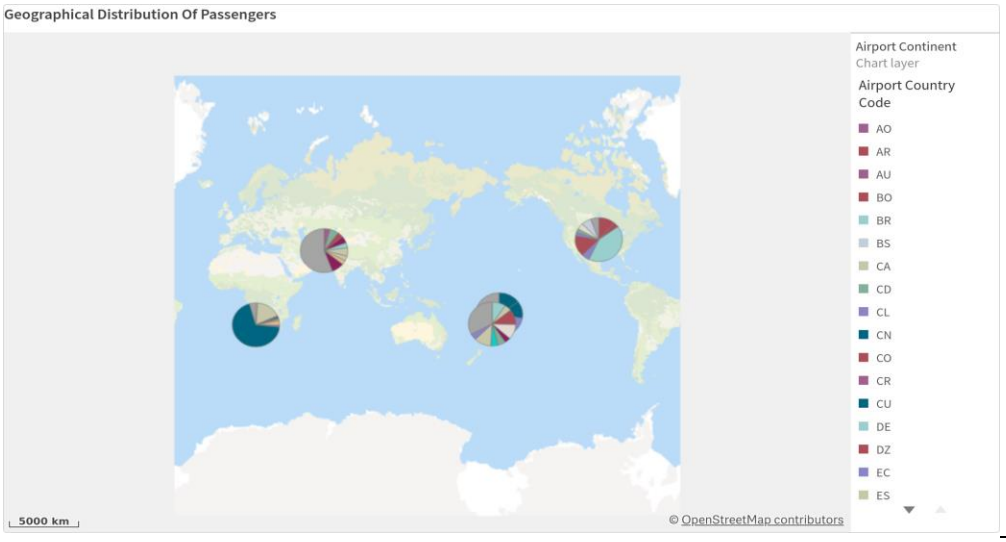


5.2.7. Filter of Ticket Price and Flight Status

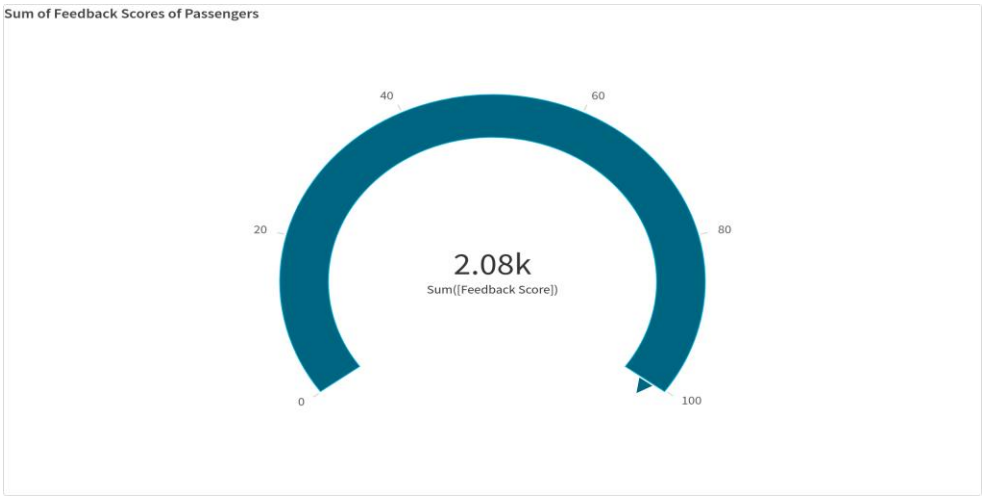
Name: Bhushan Lilhare



5.3.1. Geographical Distribution of Passengers



5.3.2. Sum of Feedback Scores



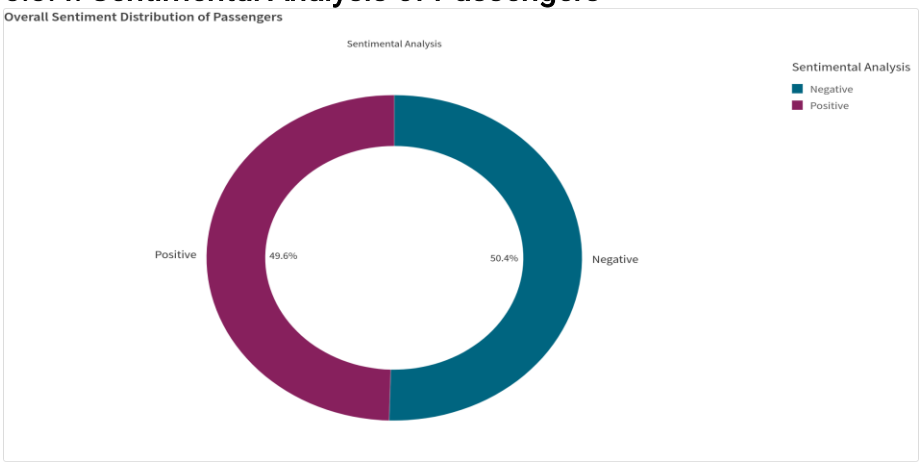
5.3.3. Negative Feedback Analysis

Name: Bhushan Lilhare

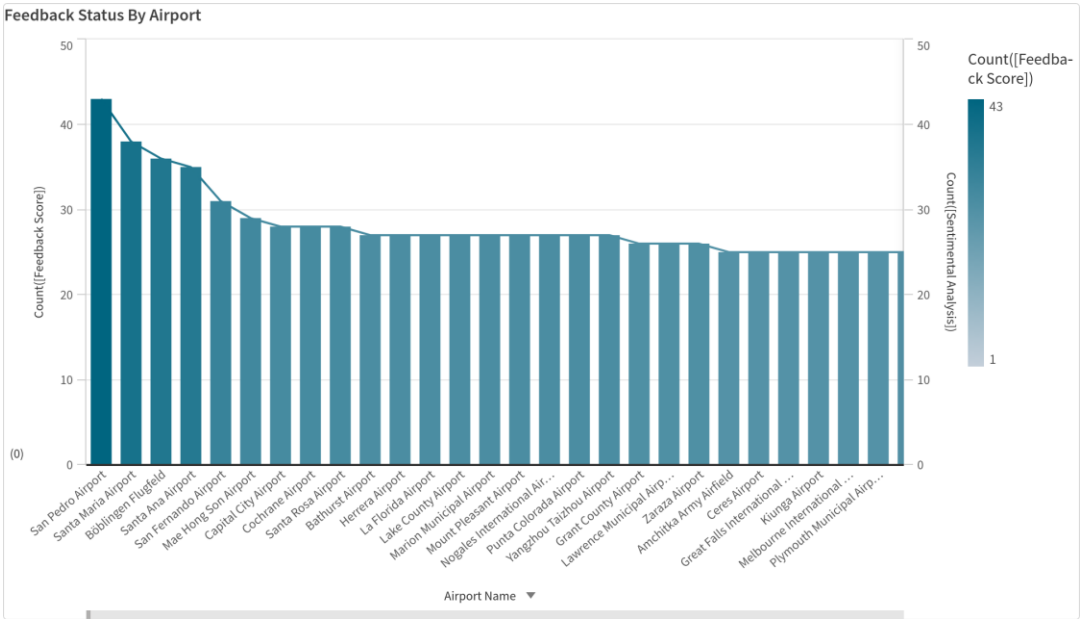
Pain Points and Improvement Areas

Sentimental Analysis	AW		AD		GW		TT		GM		CX		IO		
	Avg...	Cou...	Avg...	Cou...	Avg...	Cou...	Avg...	Cou...	Avg...	Cou...	Avg...	Cou...	Avg...	Cou...	Avg...
Negative	0	3	-0.125	4	-0.13...	3	-0.22...	11	-0.23...	9	-0.23...	6	-0.275	8	-0.28
Positive	0.56...	6	0.76...	3	0.7	7	0.54...	12	0.41...	6	0.56	5	0.625	8	0.675

5.3.4. Sentimental Analysis of Passengers



5.3.5. Feedback Scores of Passengers by Airport

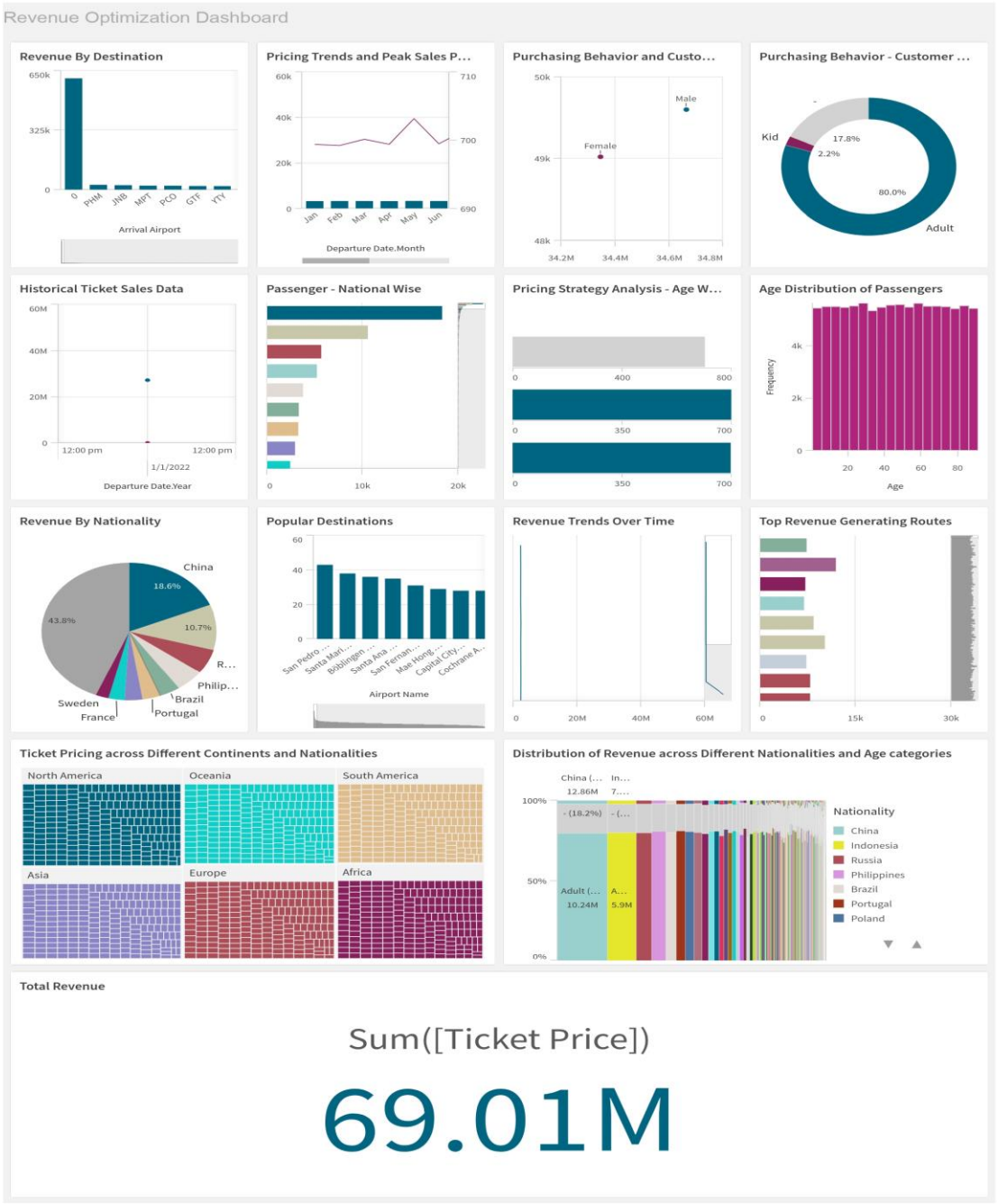


6. Dashboard

6.1 Design of Dashboard

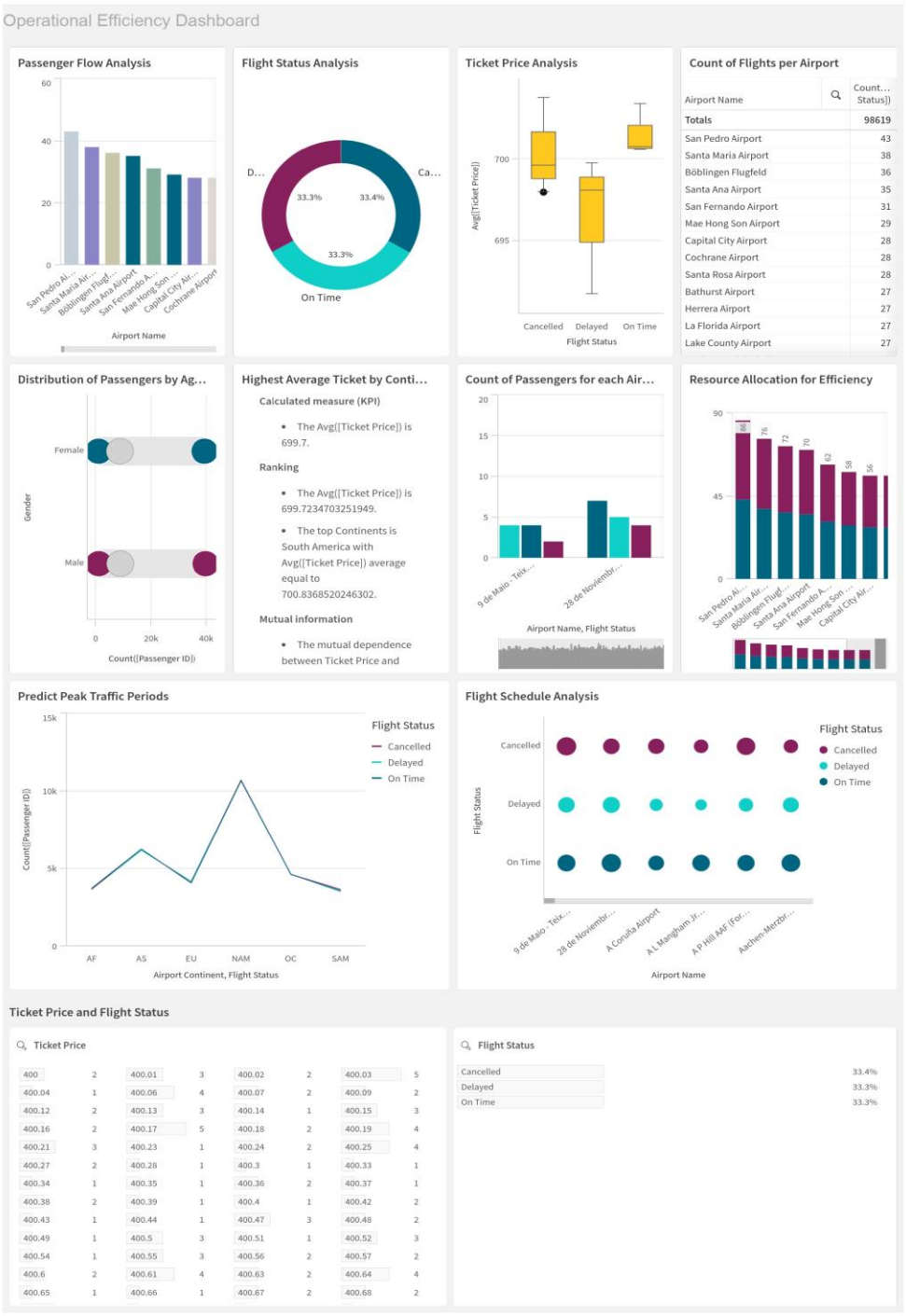
Name: Bhushan Lilhare

6.1.1 Revenue Optimization Dashboard



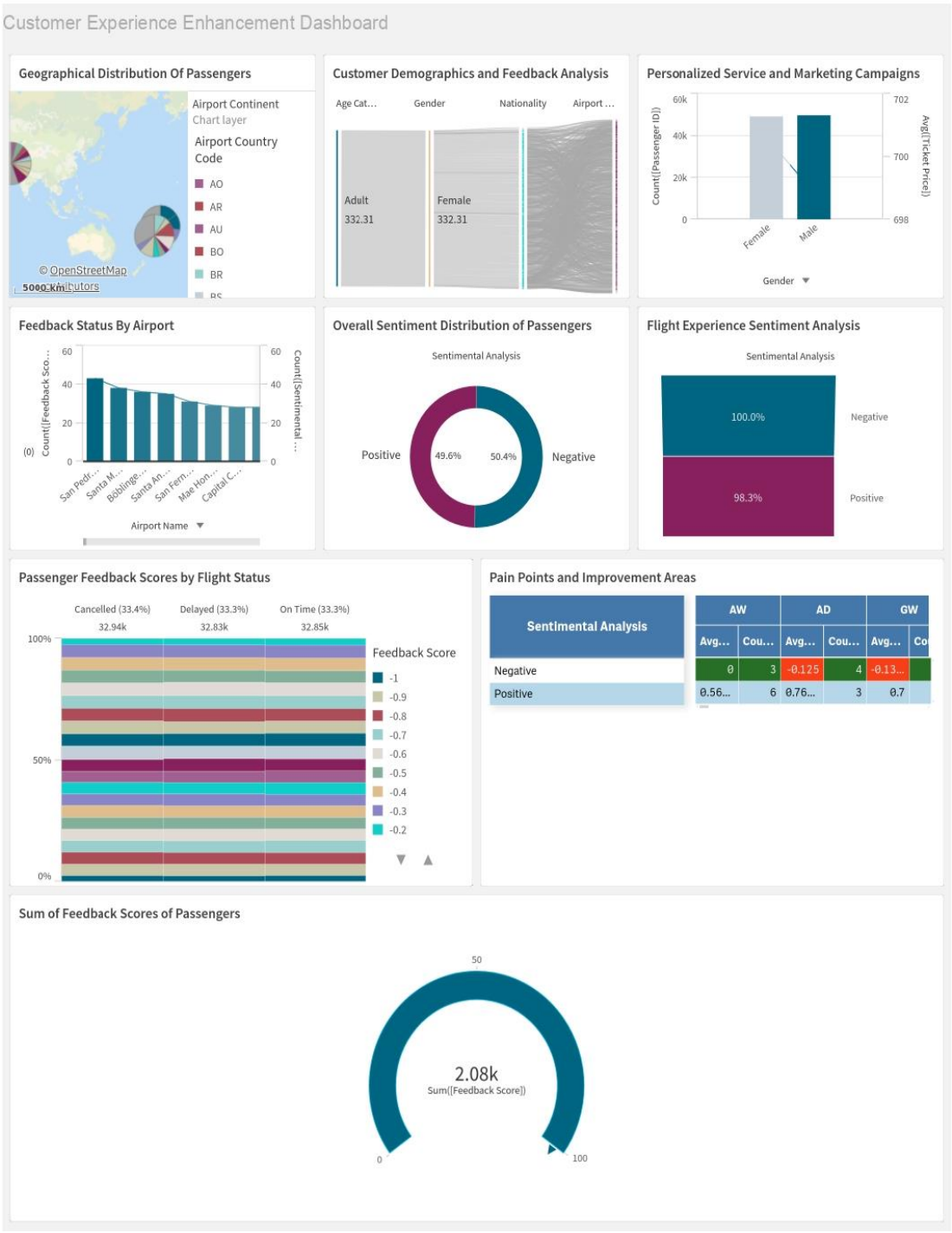
6.1.2 Operational Efficiency

Name: Bhushan Lilhare



6.1.3. Customer Experience Enhancement Dashboard

Name: Bhushan Lilhare

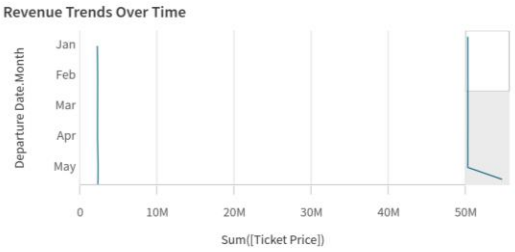


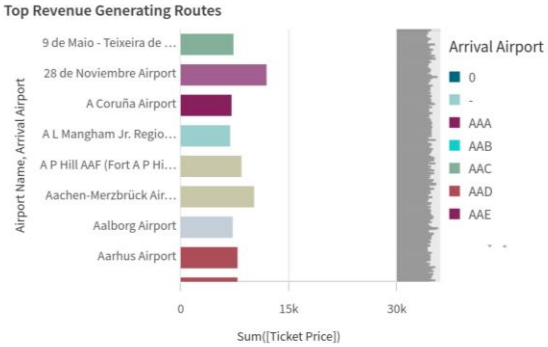
7. REPORT

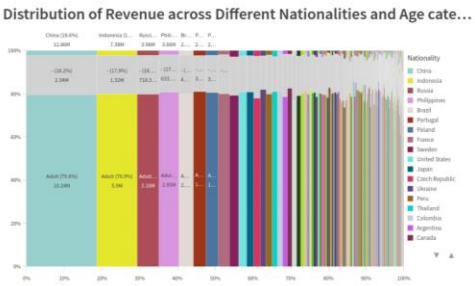
Name: Bhushan Lilhare

7.1. CREATION OF REPORT - STORYTELLING

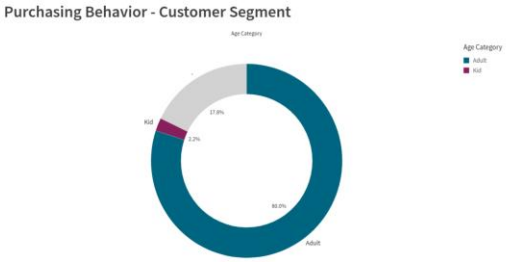
REVENUE TREND ANALYSIS OF AIRLINE DATA

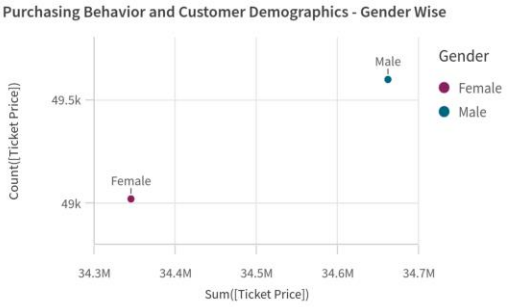


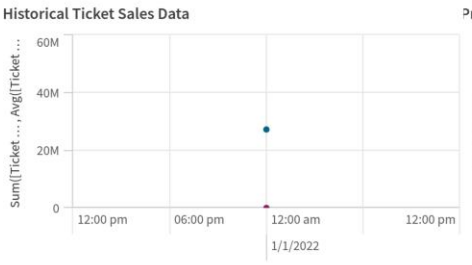




PRICING STRATEGY AND PURCHASING BEHAVIOR

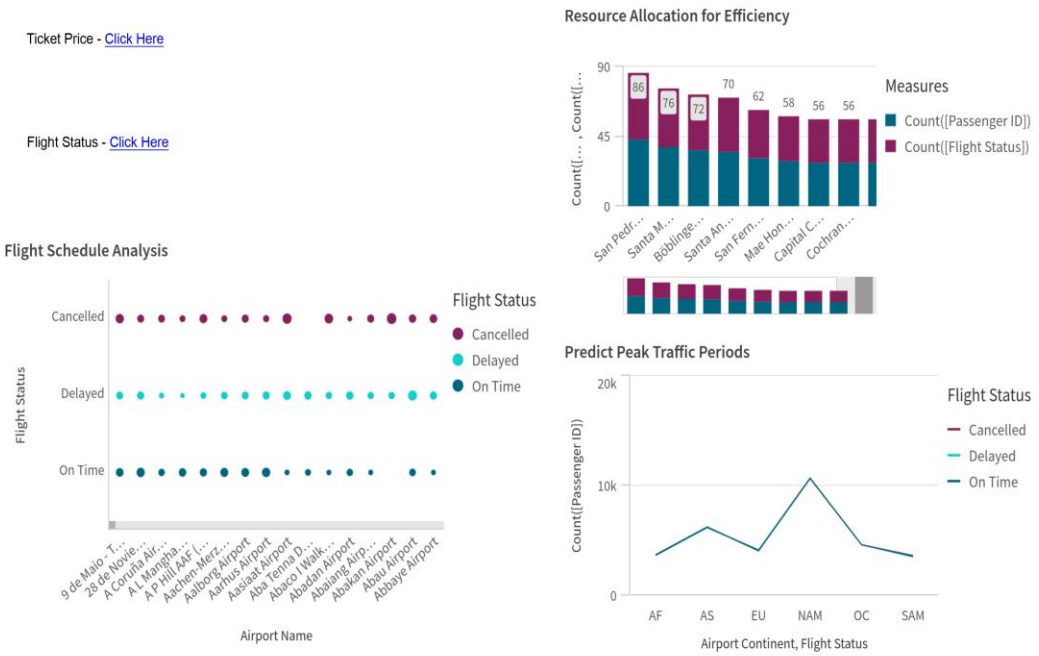




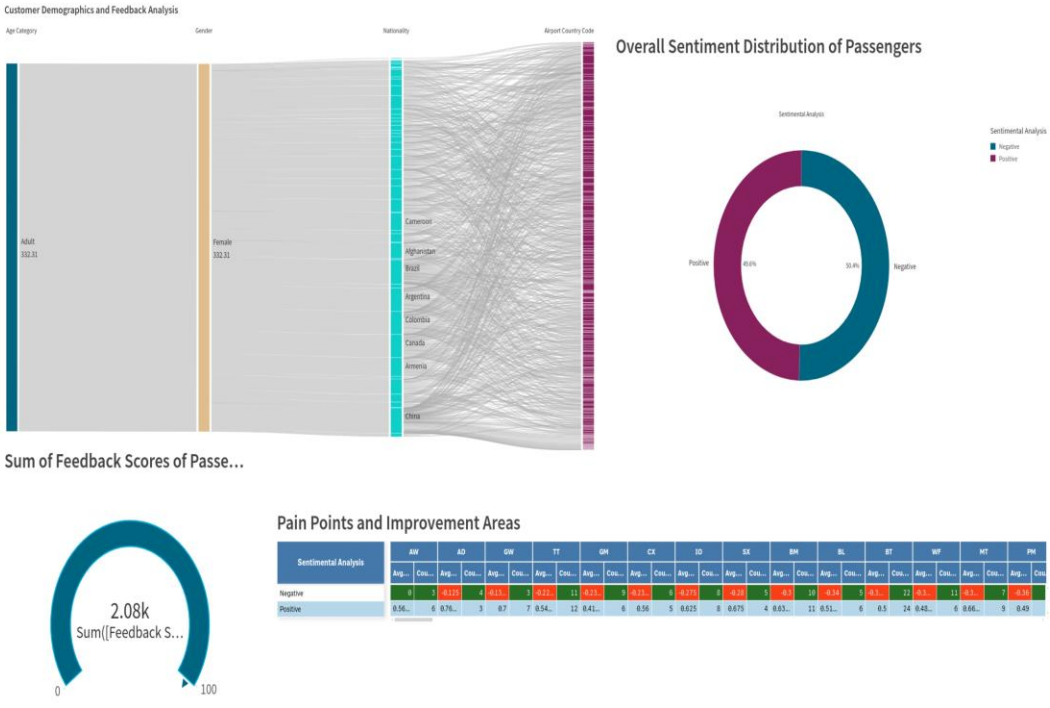


OPERATIONAL EFFICIENCY OF AIRLINE DATA

Name: Bhushan Lilhare



CUSTOMER DEMOGRAPHICS AND SENTIMENTAL ANALYSIS

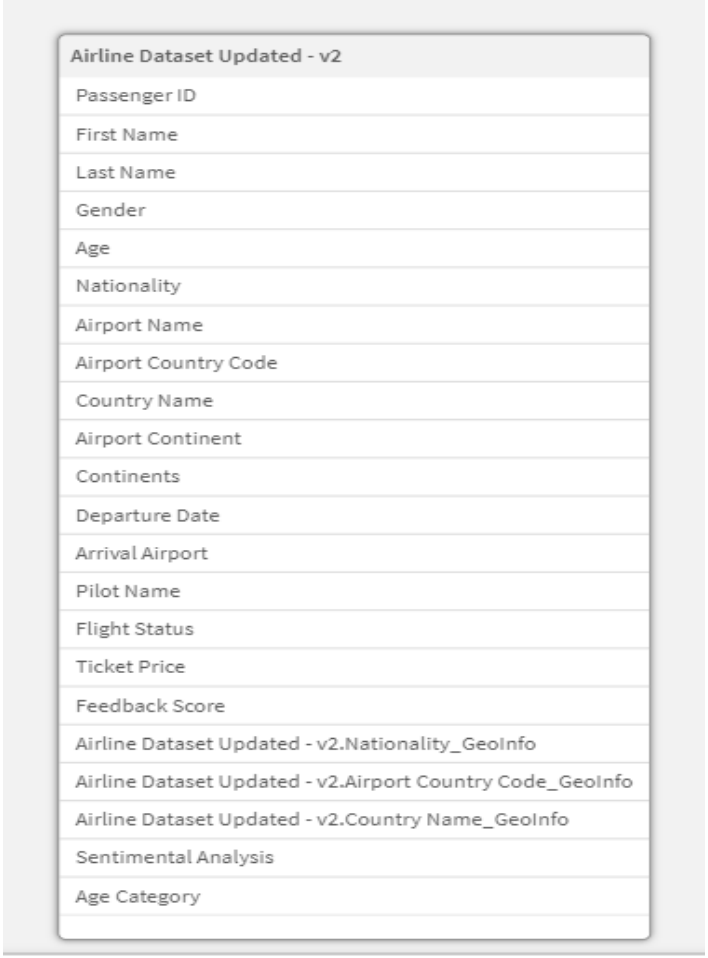


8. PERFORMANCE TESTING

8.1 AMOUNT OF DATA LOADED

Name: Bhushan Lilhare

The dataset loaded into the Qlik Sense app, titled "Airline Dataset Updated - v2," comprises various fields capturing detailed information about airline passengers and their travel details. The dataset includes identifiers and personal information such as Passenger ID, First Name, Last Name, Gender, Age, and Nationality. It also contains extensive travel-related data including Airport Name, Airport Country Code, Country Name, Airport Continent, Departure Date, Arrival Airport, and Pilot Name. Additional fields cover Flight Status, Ticket Price, Feedback Score, Sentimental Analysis, and Age Category. The dataset is enriched with geoinformation fields like Nationality_GeoInfo, Airport Country Code_GeoInfo, and Country Name_GeoInfo, providing a comprehensive set of data for analysis.



8.2 DATA FILTERS USED

8.2.1. FLIGHT STATUS FILTER

The Flight Status data includes three key statuses: "**Delayed," "On Time," and "Cancelled."** These statuses are used as filters in the Operational Efficiency dashboard to easily identify and analyze the performance of flights.

Name: Bhushan Lilhare

8.2.2. TICKET PRICE FILTER

The dataset includes various ticket prices for different passengers. This data is used as a filter in the Operational Efficiency dashboard to easily view and analyze ticket prices.

Name: Bhushan Lilhare