

Exploring Insights From Synthetic Airline Data Analysis With Qlik

1. INTRODUCTION

1.1 Overview

This project explores insights derived from synthetic airline data using Qlik Sense, a leading data analytics and visualization tool. The project aims to demonstrate how synthetic data, which mimics real-world data, can be used to perform meaningful analysis and derive insights in the context of the airline industry.

1.2 Purpose

The purpose of this project is to showcase the capabilities of Qlik Sense in handling and analyzing large datasets, even when they are synthetic. The project aims to:

- Identify patterns and trends in airline operations.
- Provide actionable insights for decision-making.
- Demonstrate the use of Qlik Sense for creating interactive and dynamic visualizations.

1.3 Technical Architecture

The technical architecture involves:

- Data source: Synthetic airline data.
- Data processing: Data preparation using Qlik Sense.
- Data visualization: Interactive dashboards and reports created in Qlik Sense.
- Deployment: Qlik Sense server for rendering and sharing visualizations.

2. DEFINE PROBLEM / PROBLEM UNDERSTANDING

2.1 Specify the Business Problem

The business problem addressed is optimizing airline operations to enhance efficiency, reduce costs, and improve customer satisfaction. This involves analyzing flight schedules, delays, passenger demographics, and operational costs.

Exploring Insights From Synthetic Airline Data Analysis With Qlik

2.2 Business Requirements

- Ability to handle large datasets.
- Interactive and user-friendly visualizations.
- Real-time data processing and analysis.
- Insights into key performance metrics such as on-time performance, load factors, and revenue.

2.3 Literature Survey

A review of existing literature on airline data analysis reveals common challenges such as data volume, complexity, and the need for real-time insights. Studies have shown that data analytics can significantly improve operational efficiency and customer satisfaction in the airline industry.

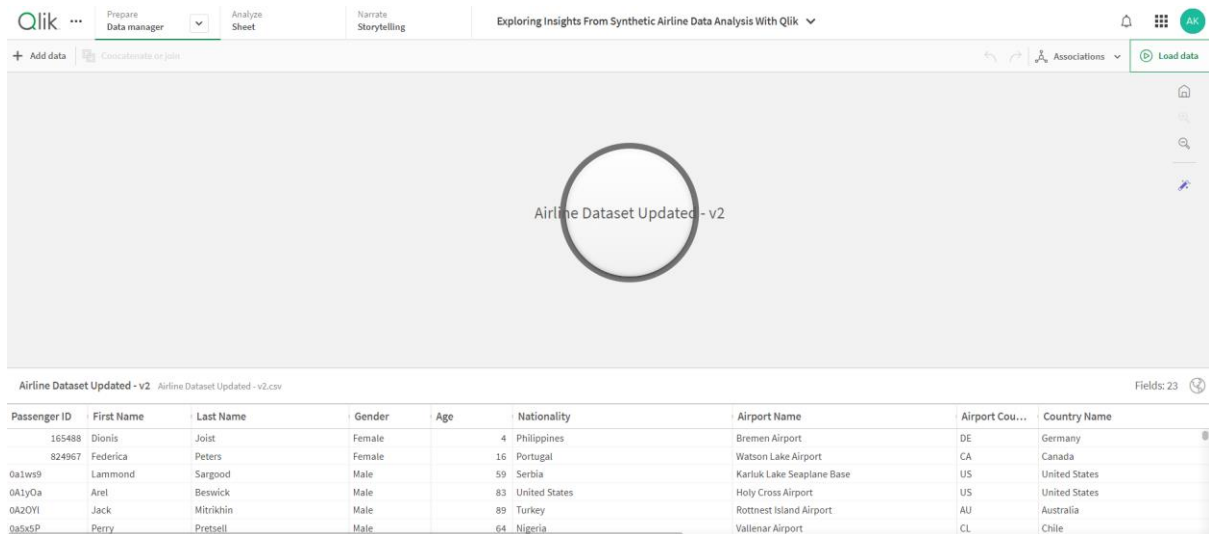
Additionally, studies emphasize the role of data analytics in managing disruptions caused by weather, economic shifts, and geopolitical events. The comprehensive application of data analytics thus enables airlines to make informed decisions, improve performance, and maintain a competitive edge in the market.

3. DATA COLLECTION

3.1 Collect the Dataset

The dataset comprises synthetic data generated to reflect real-world airline operations, including flight schedules, passenger information, and operational metrics.

Exploring Insights From Synthetic Airline Data Analysis With Qlik



Passenger ID	First Name	Last Name	Gender	Age	Nationality	Airport Name	Airport Cou...	Country Name
165488	Dionis	Jolist	Female	4	Philippines	Bremen Airport	DE	Germany
824967	Federica	Peters	Female	16	Portugal	Watson Lake Airport	CA	Canada
0a1ws9	Lammond	Sargood	Male	59	Serbia	Karluk Lake Seaplane Base	US	United States
0A1y0a	Arel	Beswick	Male	83	United States	Holy Cross Airport	US	United States
0A20VI	Jack	Mitrikin	Male	89	Turkey	Rottneest Island Airport	AU	Australia
0a5x5P	Perry	Pretsell	Male	64	Nigeria	Vallenar Airport	CL	Chile

3.2 Connect Data with Qlik Sense

The dataset is uploaded to Qlik Sense, where it is connected using Qlik's data load editor. The data is then structured and organized for analysis.

4. DATA PREPARATION

4.1 Prepare the Data for Visualization

Data preparation involves cleaning the dataset, handling missing values, and transforming data into a suitable format for visualization. This step ensures the accuracy and reliability of the data used for analysis.

5. DATA VISUALIZATIONS

5.1 Visualizations

Various visualizations are created to explore different aspects of the data:

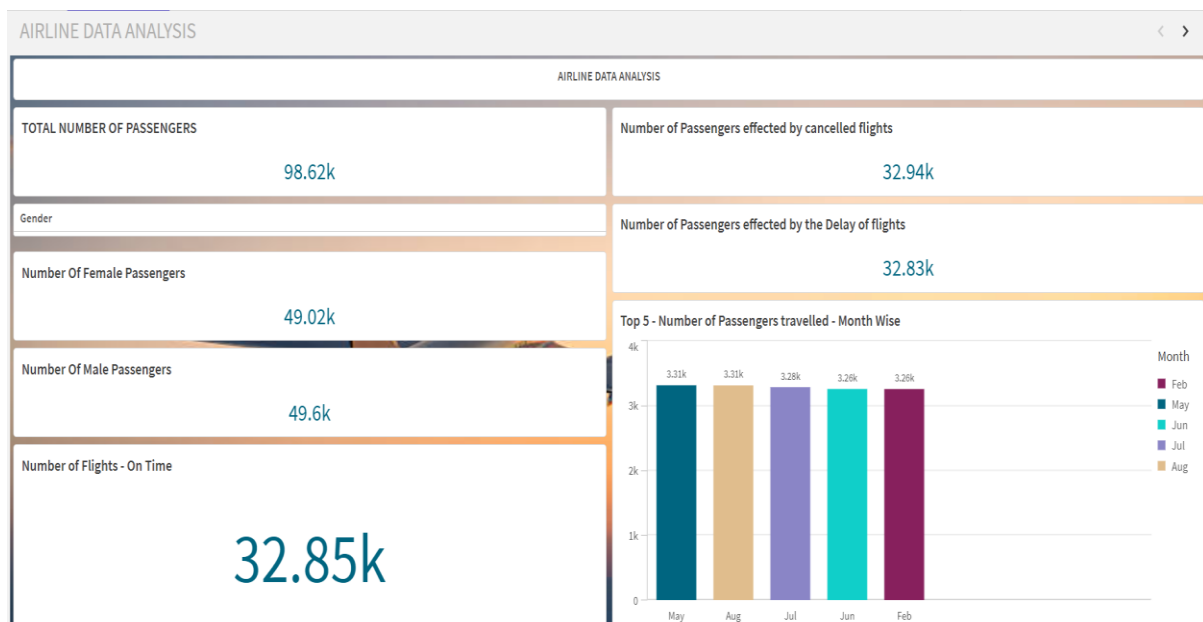
Exploring Insights From Synthetic Airline Data Analysis With Qlik

- Bar charts to compare flight delays across different airlines.
- Line graphs to show trends in passenger numbers over time.
- Heat maps to visualize flight density and popular routes.
- Pie charts to represent the distribution of operational costs.

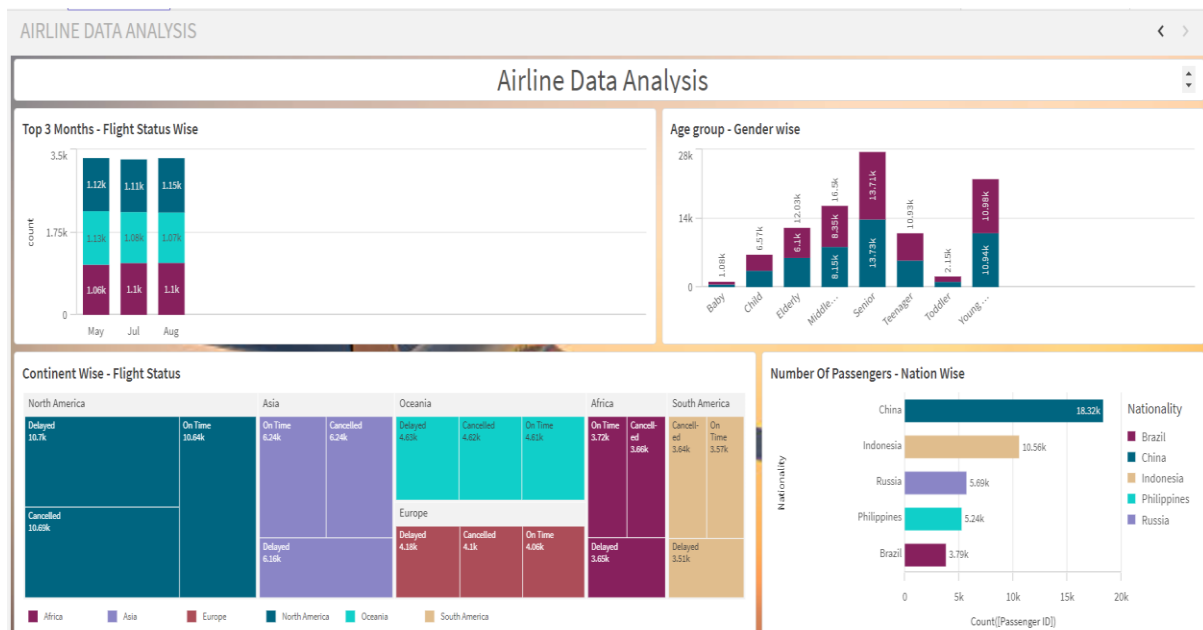
6. DASHBOARD

6.1 Responsive and Design of Dashboard

The dashboard is designed to be responsive and user-friendly, allowing users to interact with the data through filters and drill-down capabilities. It includes multiple views and perspectives to provide a comprehensive overview of the airline operations.



Exploring Insights From Synthetic Airline Data Analysis With Qlik



7. REPORT

7.1 Report Creation

Reports are generated to summarize the findings from the data analysis. These reports include key insights, visualizations, and recommendations for improving airline operations. They include key insights that highlight trends, patterns, and anomalies in the data, offering a deeper understanding of various operational aspects. Visualizations such as charts, graphs, and dashboards are employed to present complex data in an accessible and intuitive manner, making it easier for decision-makers to grasp the implications. Recommendations based on the analysis are also included, aimed at improving airline operations. These may cover strategies for optimizing flight schedules, enhancing customer service, reducing operational costs, and increasing overall efficiency. By leveraging the power of data analytics, these reports serve as a valuable tool for driving informed decision-making and fostering continuous improvement within the airline industry.

Exploring Insights From Synthetic Airline Data Analysis With Qlik

Exploring Insights From Synthetic Airline Data Analysis With Qlik

TOTAL NUMBER OF PASSENGERS

98.62k

Number of Passengers effected by cancelled flights

32.94k

Number of Passengers effected by the Delay of flights

32.83k

Number of Flights - On Time

32.85k

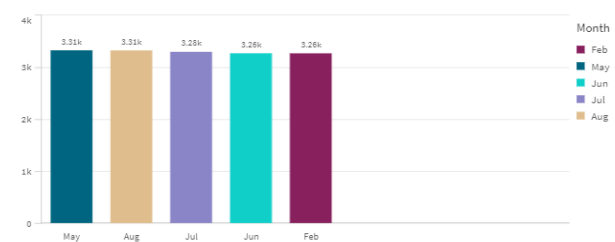
THE NUMBER OF MALE AND FEMALE PASSENGERS TRAVELLED

MALE - [Click here](#)

FEMALE - [Click here](#)

This chart displays the Top 5 - Month Wise - Number Of Passengers Travelled

Top 5 - Number of Passengers travelled - Month Wise

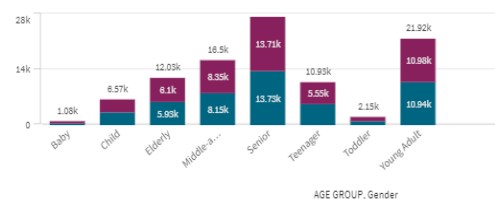


Exploring Insights From Synthetic Airline Data Analysis With Qlik

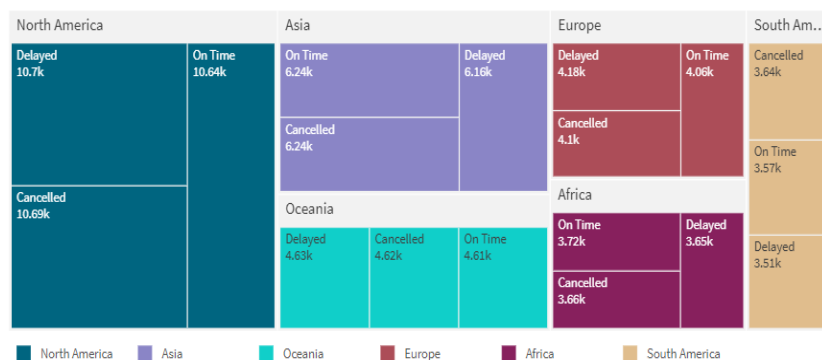
Top 3 Months - Flight Status Wise



Age group - Gender wise

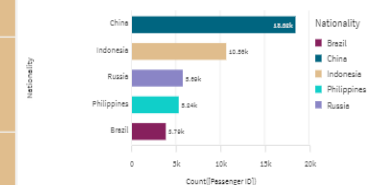


Continent Wise - Flight Status



This chart shows Age Group - Gender Wise

Number Of Passengers - Nation Wise



Exploring Insights From Synthetic Airline Data Analysis With Qlik

8. PERFORMANCE TESTING

8.1 Amount of Data Rendered

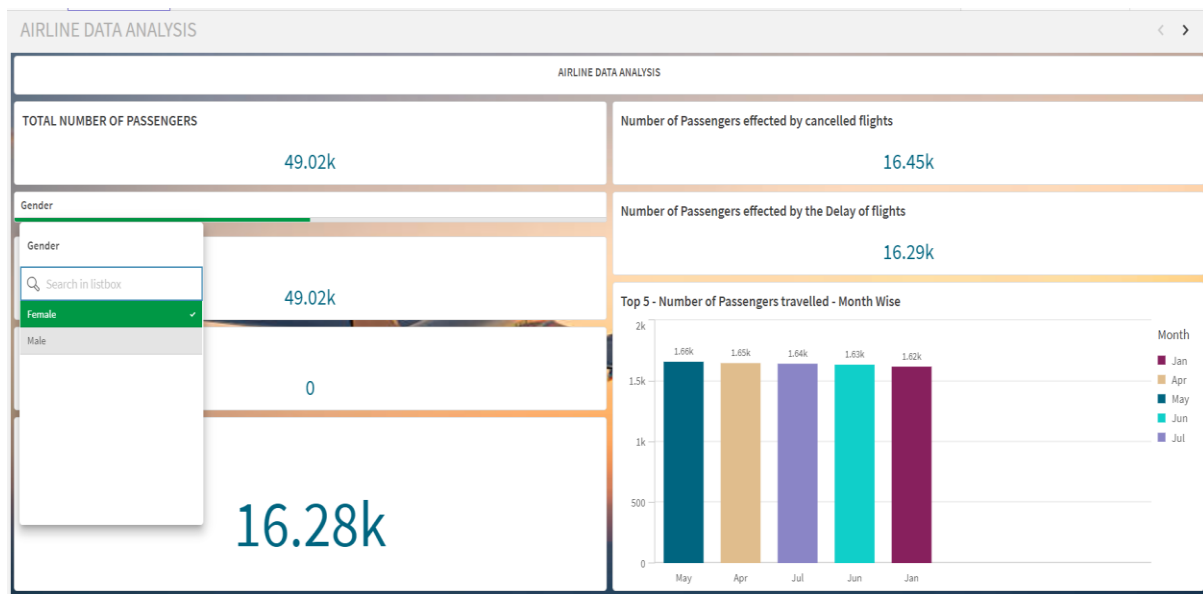
Performance testing is conducted to evaluate how Qlik Sense handles large volumes of data. This involves testing the rendering speed and responsiveness of visualizations when dealing with substantial datasets.

Airline Dataset Updated - v2	
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	
Airline Dataset Updated - v2.Nationality_GeoInfo	
Airline Dataset Updated - v2.Airport Country Code_GeoInfo	
Airline Dataset Updated - v2.Country Name_GeoInfo	
Year	
Month	
Cancelled	
MALE	
FEMALE	
Delayed	
On Time	
AGE GROUP	

8.2 Utilization of Data Filters

The effectiveness of data filters is assessed to ensure users can efficiently navigate and analyze specific subsets of data. This includes testing the performance impact of applying multiple filters simultaneously.

Exploring Insights From Synthetic Airline Data Analysis With Qlik



By following this structured approach, the project demonstrates the powerful capabilities of Qlik Sense in transforming synthetic airline data into actionable insights, ultimately aiding in the optimization of airline operations.

Exploring Insights From Synthetic Airline Data Analysis With Qlik

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. The data is then structured to create key performance indicators (KPIs) such as the total number of passengers, on-time flights, delayed flights, and canceled flights. Visualizations, such as tree maps, help in understanding the distribution of flight status across different continents.

Scope:

The scope of this project encompasses the following areas:

Data Quality Improvement: Ensuring the accuracy and completeness of data by handling duplicates and null values.

Geographical Analysis: Using visualizations to analyze flight performance across different continents, providing insights into regional performance variations.

Performance Measurement: Developing KPIs to monitor flight performance, including on-time performance, delays, and cancellations.

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

Conclusion:

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

GitHub link: <https://github.com/am4426/AIRLINE-DATA-ANALYSIS>