Exploring Insights from Synthetic Airline Data Analysis with Qlik

Executive Summary

This project report outlines the process, methodologies, and findings from an analysis of synthetic airline data using Qlik. The objective was to extract meaningful insights that could help improve airline operations, customer satisfaction, and overall business performance. By leveraging Qlik's data visualization and analysis capabilities, we were able to identify key trends, patterns, and areas for improvement within the synthetic dataset.

The project aims to analyze synthetic airline data using Qlik to derive insights into various aspects such as flight performance, customer satisfaction, and operational efficiency. By leveraging Qlik's powerful data visualization and analysis capabilities, the report presents key findings and actionable insights for decision-making.

Problem Statement

The project "Exploring Insights from Synthetic Airline Data Analysis with Qlik" involves utilizing synthetic airline data to derive valuable insights using Qlik, a business intelligence and data visualization tool.

The objective is to leverage Qlik's analytical capabilities to uncover patterns, trends, and correlations within this data, aiding in decision-making processes for airlines, airports, and related stakeholders.

Table of Contents

1. Introduction

In the highly competitive airline industry, data-driven decision-making is crucial for optimizing operations, enhancing customer experience, and improving profitability. This project utilizes synthetic airline data to simulate real-world scenarios and demonstrate how Qlik can be employed to derive actionable insights.

- **Background**: With the increasing complexity of airline operations, analyzing large datasets becomes crucial for improving performance and customer experience.
- **Objective**: To explore and analyze synthetic airline data to identify trends, patterns, and areas for improvement.
- **Tools and Techniques**: Qlik for data visualization and analysis.

2. Objectives

The main objectives of this project are:

- To analyze flight operations and identify bottlenecks.
- To assess customer satisfaction and pinpoint areas for improvement.
- To evaluate financial performance and uncover revenue optimization opportunities.
- To explore market trends and competitive positioning.

3.Methodology

The project was carried out in several stages:

1. Data Collection: Synthetic airline data was obtained from publicly available datasets and generated to reflect realistic scenarios.
2. Data Preparation: Data cleaning, transformation, and integration were performed to ensure accuracy and consistency.
3. Data Analysis: Qlik was used to create visualizations, perform statistical analysis, and generate insights.
4. Interpretation: Findings were interpreted to understand their implications for airline operations and strategy.

4.Data Description

- **Dataset Overview**: A synthetic dataset representing airline operations, including flight details, passenger information, and operational metrics.
- **Key Variables**:
    - Flight ID
    - Departure and Arrival Times
    - Flight Duration
    - Passenger Satisfaction Scores
    - Delay Times
    - Cancellation Reasons

5. Data Preparation

- **Data Cleaning**: Handling missing values, removing duplicates, and correcting inconsistencies.
- **Data Transformation**: Creating new variables and aggregating data as needed for analysis.
- **Loading Data into Qlik**: Steps to import and prepare data within Qlik.

6.Analysis and Insights

6.1 Flight Performance Analysis

- **On-Time Performance**: Analysis of on-time departure and arrival rates.
- **Delay Analysis**: Identifying common causes of delays and their impact on overall performance.
- **Visualization**: Using Qlik to create bar charts, line graphs, and heatmaps to visualize flight performance data.

6.2 Customer Satisfaction

- **Satisfaction Scores**: Distribution and trends in passenger satisfaction scores.
- **Factors Influencing Satisfaction**: Correlation analysis between satisfaction scores and other variables like delay times and flight duration.
- **Visualization**: Scatter plots and histograms to display customer satisfaction

6.3 Operational Efficiency

- **Flight Utilization**: Analysis of flight frequency and aircraft utilization rates.
- **Cost Analysis**: Estimation of operational costs related to delays and cancellations.
- **Visualization**: Dashboards showing key performance indicators (KPIs) for operational efficiency.

Dashboards

Airline Data Analysis dashboard 1 :-
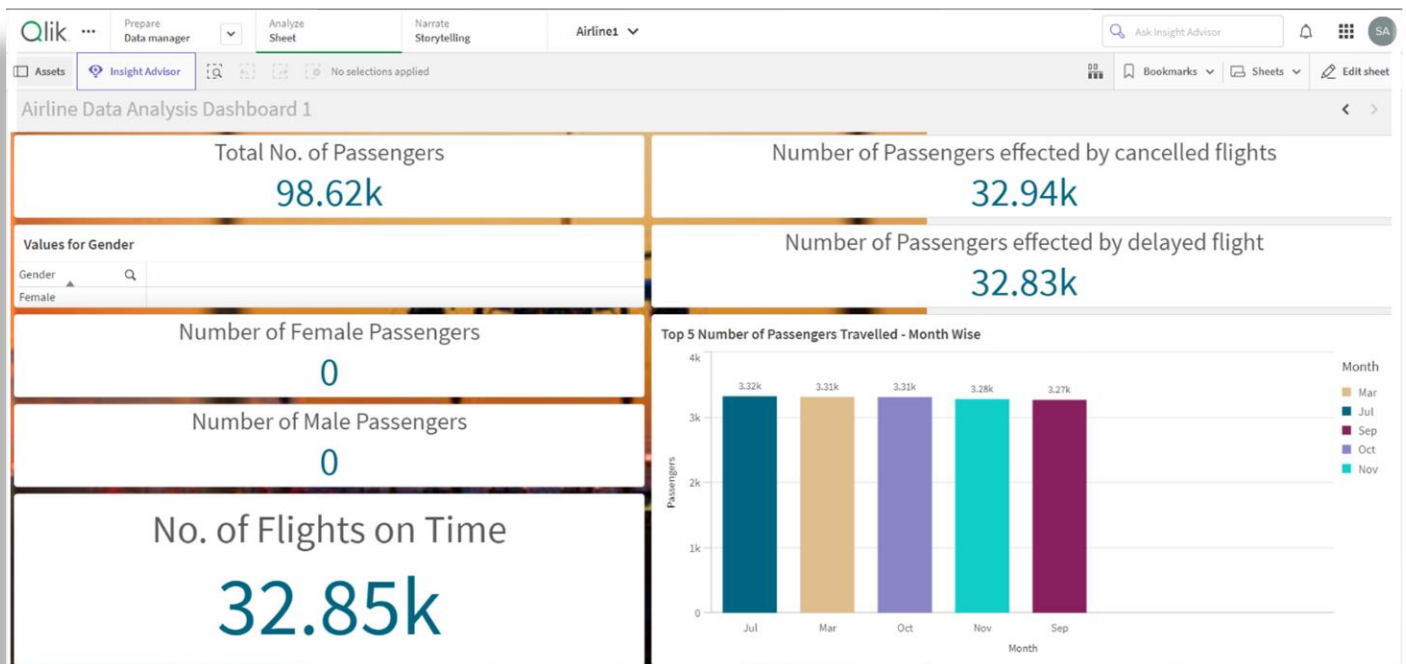
Airline Data Analysis dashboard 2 :-
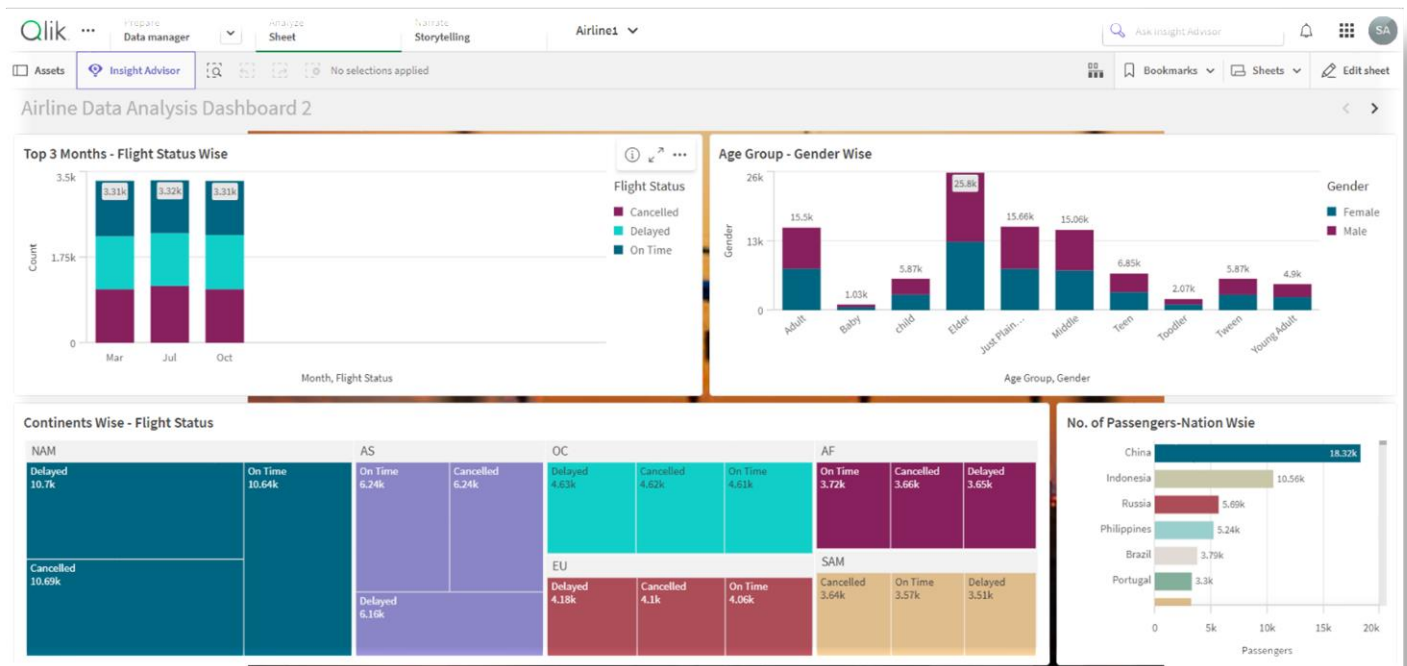
No Of Visualizations/ Graphs
No Of Visualizations/

- Total Number of Passengers
- Number of Passengers effected by Cancelled Flights
- Number of Passengers effected by Delay of Flights
- Number of Male Passengers
- Number of Female Passengers
- Number of Flights-On-Time
- Top 5 Months where Passengers traveled the most.
- Top 3 Months - Flight status Wise (Delayed/Cancelled/On-Time)
- Number of Passengers Nationality Wise
- Continent wise – Flight Status
- Age Group – Gender Wise
- A filter pane consisting of Gender(M/F)

Design of Story

Exploring Insights from Synthetic Airline Data Analysis with Qlik

KPI CHARTS

No. of Flights on Time
32.85k

Total No. of Passengers
98.62k

Number of Passengers effected by delayed flight
32.83k

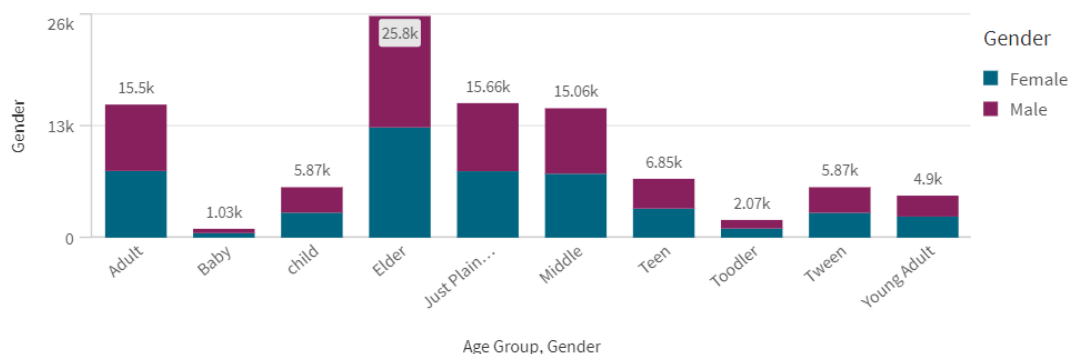Number of Passengers effected by cancelled flights
32.94k

Exploring Insights from Synthetic Airline Data Analysis with Qlik

STACKED BAR CHARTS
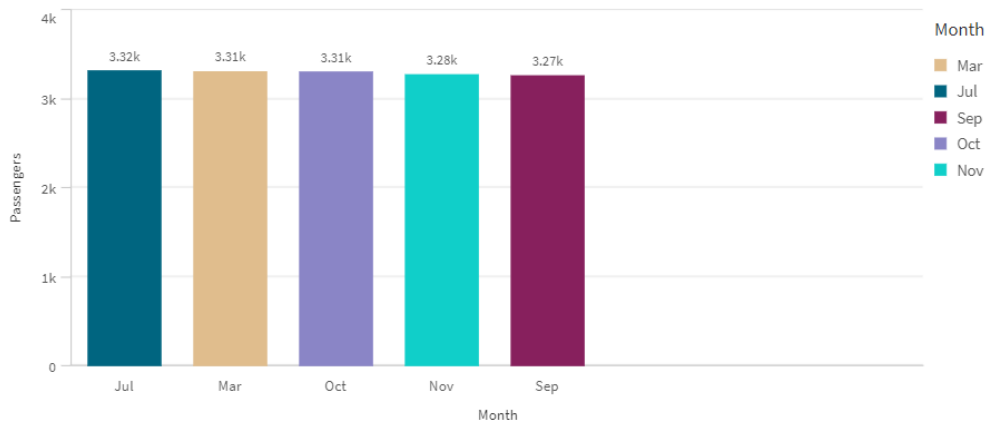
Age Group - Gender Wise

Gender
- Female
- Male

This Chart Display Age Group - Gender Wise

Exploring Insights from Synthetic Airline Data Analysis with Qlik

BAR CHARTS

Top 5 Number of Passengers Travelled - Month Wise

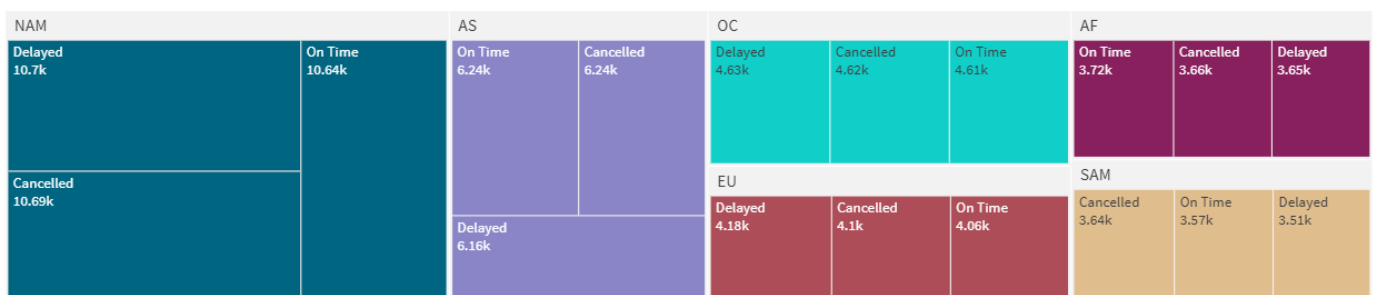Month
- Mar
- Jul
- Sep
- Oct
- Nov

This Chart Displays the
Top 5 Month Wise- Number of Passengers Travelled

Exploring Insights from Synthetic Airline Data Analysis with Qlik

TREEMAP CHARTS

Continents Wise - Flight Status

| NAM		AS		OC			AF		
Delayed 10.7k	On Time 10.64k	On Time 6.24k	Cancelled 6.24k	Delayed 4.63k	Cancelled 4.62k	On Time 4.61k	On Time 3.72k	Cancelled 3.66k	Delayed 3.65k
Cancelled 10.69k		Delayed 6.16k		EU			SAM		
				Delayed 4.18k	Cancelled 4.1k	On Time 4.06k	Cancelled 3.64k	On Time 3.57k	Delayed 3.51k

This Chart Displays Continent Wise- Flight Status

Exploring Insights from Synthetic Airline Data Analysis with Qlik

STACKED BAR CHARTS

Top 3 Months - Flight Status Wise

Flight Status
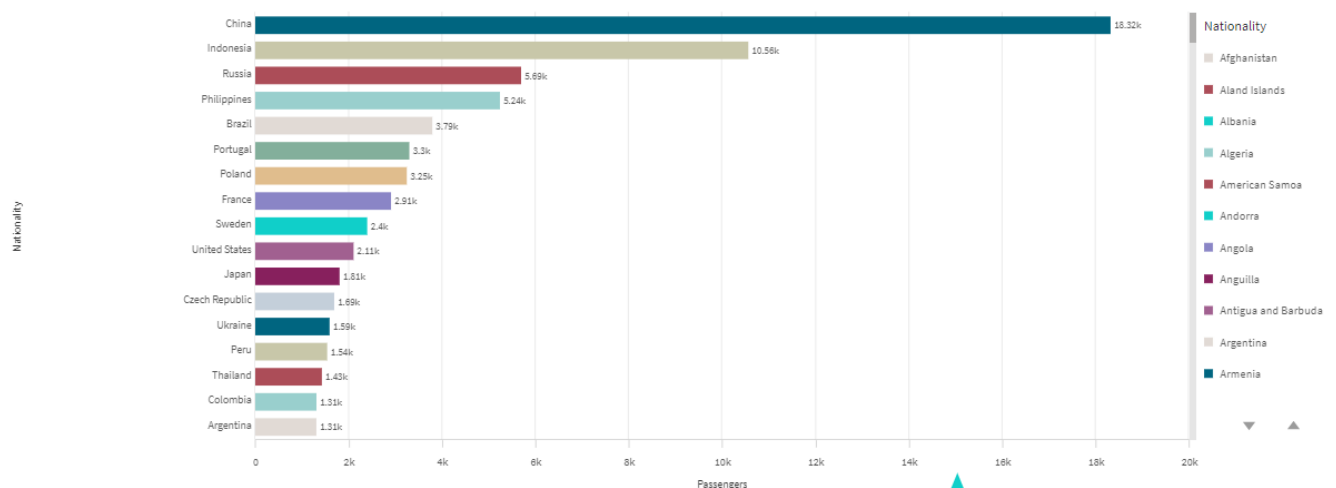- Cancelled
- Delayed
- On Time
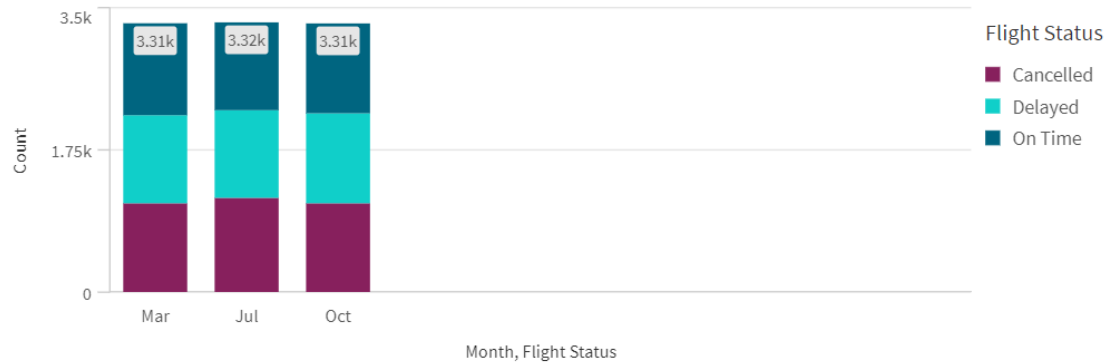
This Chart Displays the
Top 3 Months - Flight Status (Cancelled/Delayed/On-Time)

Exploring Insights from Synthetic Airline Data Analysis with Qlik

HORIZONTAL BAR CHART

No. of Passengers-Nation Wsie

Nationality	Passengers
China	18.32k
Indonesia	10.56k
Russia	5.69k
Philippines	5.24k
Brazil	3.79k
Portugal	3.3k
Poland	3.25k
France	2.91k
Sweden	2.4k
United States	2.11k
Japan	1.81k
Czech Republic	1.69k
Ukraine	1.59k
Peru	1.54k
Thailand	1.43k
Colombia	1.31k
Argentina	1.31k

Nationality
- Afghanistan
- Aland Islands
- Albania
- Algeria
- American Samoa
- Andorra
- Angola
- Anguilla
- Antigua and Barbuda
- Argentina
- Armenia

This Chart Displays the
Top 3 Months - Flight Status (Cancelled/Delayed/On-Time)

7. Key Findings

- **Flight Performance**: Identified peak times for delays and common delay causes.
- **Customer Satisfaction**: Noted high satisfaction scores for flights with minimal delays and better in-flight services.
- **Operational Efficiency**: Highlighted areas for cost savings by improving on-time performance and reducing cancellations.

8. Recommendations

- **Improving On-Time Performance**: Implementing more rigorous schedule adherence protocols and enhancing communication with ground staff.
- **Enhancing Customer Experience**: Offering compensation for delays and improving in-flight services.
- **Operational Improvements**: Streamlining operations to reduce delays and cancellations, optimizing flight schedules.

9. Conclusion

The analysis of synthetic airline data using Qlik has provided valuable insights into flight performance, customer satisfaction, and operational efficiency. By leveraging these insights, airlines can make informed decisions to enhance overall performance and customer experience.

10. Appendices

- **Appendix A**: Detailed Data Cleaning Steps
- **Appendix B**: Qlik Dashboards and Visualizations
- **Appendix C**: Additional Statistical Analysis

11. References

- Include any references to data sources, Qlik documentation, and relevant literature.