

INTRODUCTION

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.

Many visualizations are presented using Qlik to demonstrate various aspects of airline operations such as flight schedules, passenger demographics, and performance metrics, etc.

The main purpose of this project is to leverage Qlik's analytical capabilities to uncover patterns, trends, and correlations within this data and aid the decision-making process for airlines, airports, and related stakeholders.

PROBLEM STATEMENT

An airport authority wants to enhance operational efficiency by analyzing flight schedules, passenger flows, countries covered under the authority, etc. By integrating Qlik with synthetic airline data, they can identify bottlenecks in airport operations, predict peak traffic periods, and allocate resources effectively to streamline processes and improve overall efficiency.

DATA COLLECTION

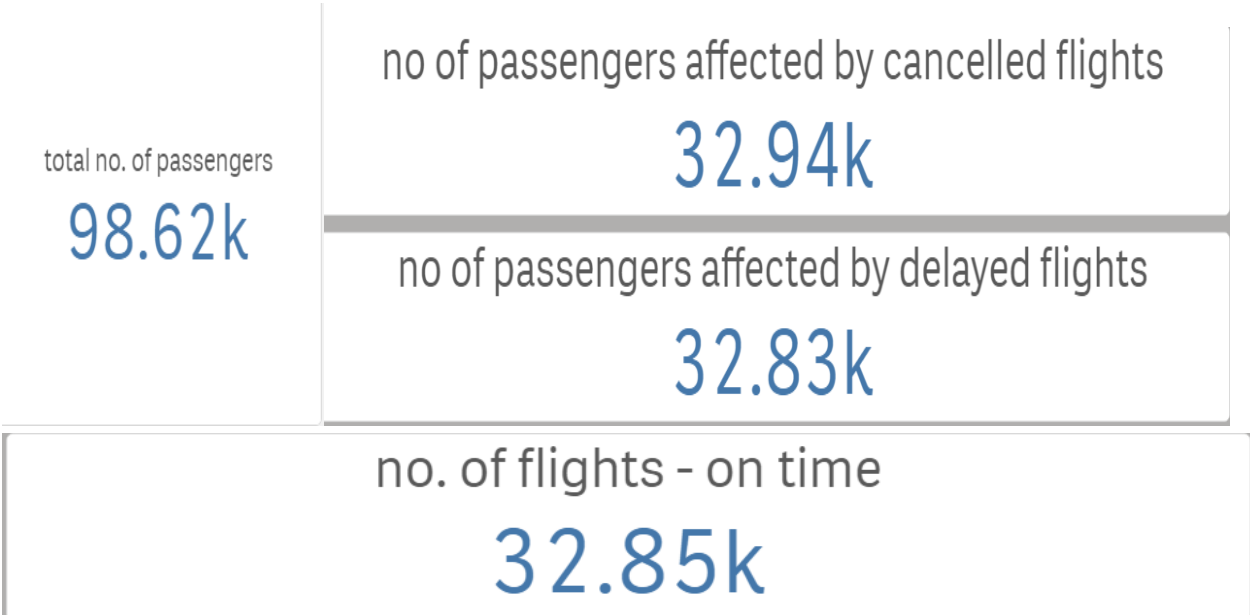
(https://drive.google.com/file/d/1kdi2Tw0C5YRtWCfTw9hWlrzwrdd4YgT7/view?usp=drive_link)

The dataset for this project is taken from kaggle. 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. Researchers and industry experts can leverage this dataset to analyze trends in passenger behavior, optimize travel experiences, evaluate pilot performance, and enhance overall flight operations.

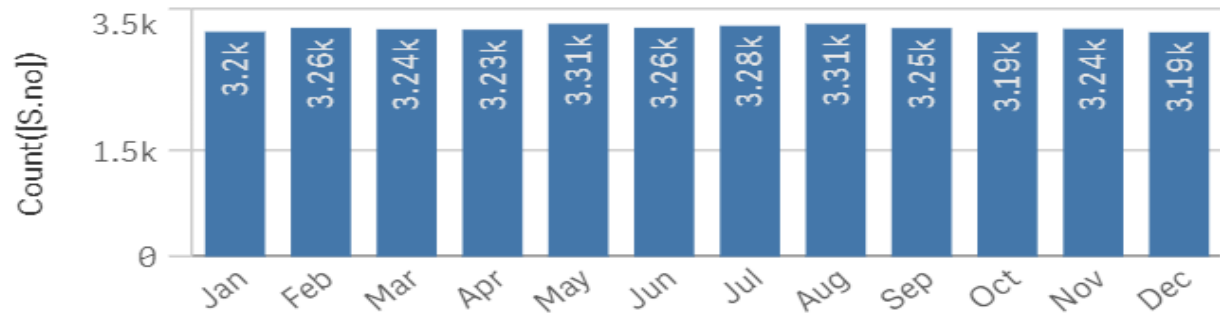
DATA PREPARATION

In the given dataset, a column of serial number is added and a column of pilot name is removed for better visualizations. A new column is also generated titled "month" derived from the existing column "Departure Date". This modified data is then uploaded in Qlik Cloud for visualization.

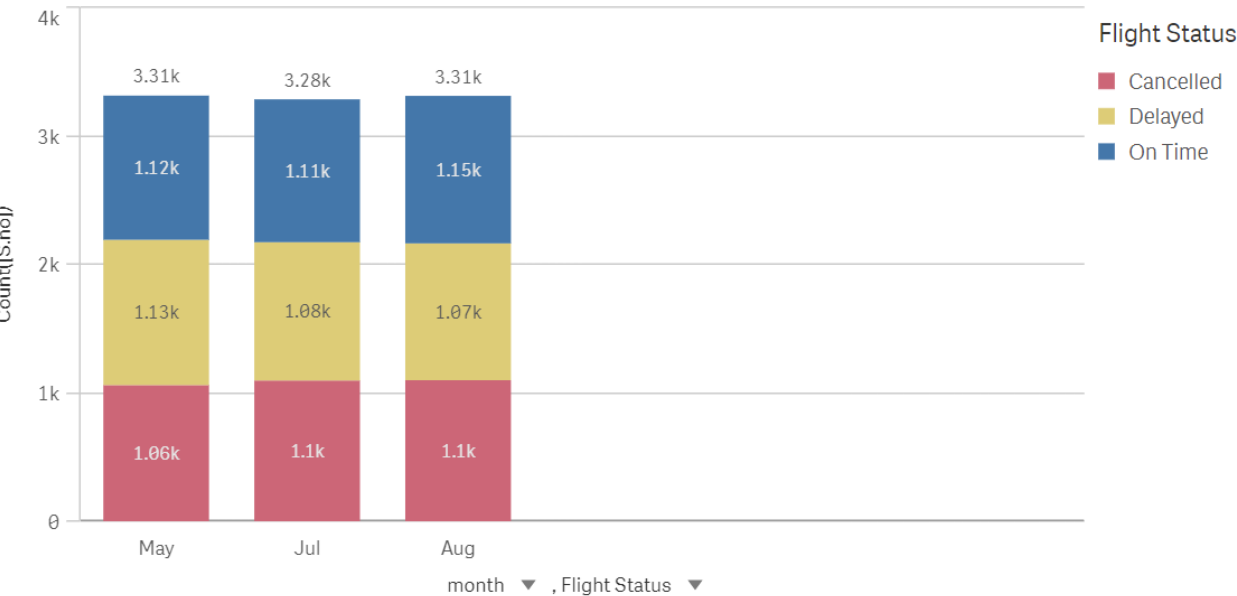
DATA VISUALISATION



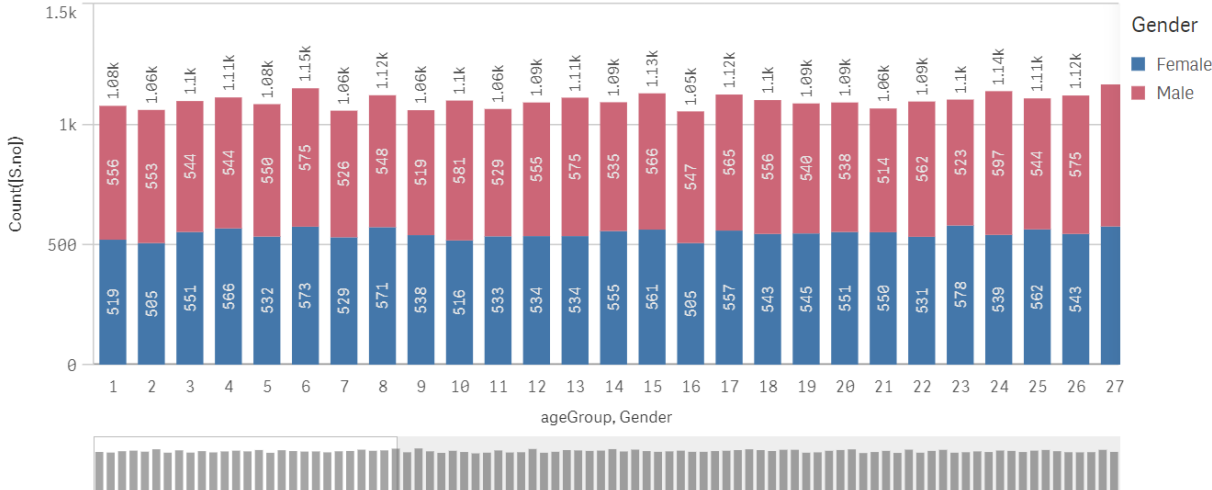
No. of passengers travelled - Month wise



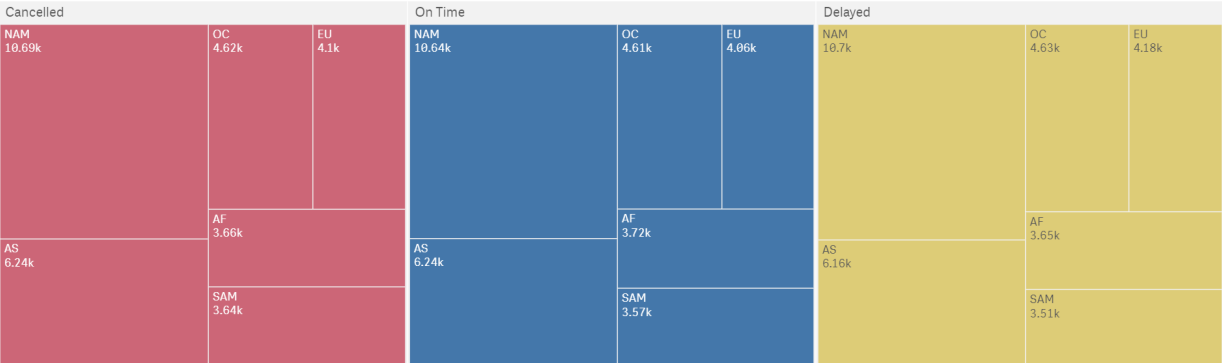
Top 3 month flight status wise



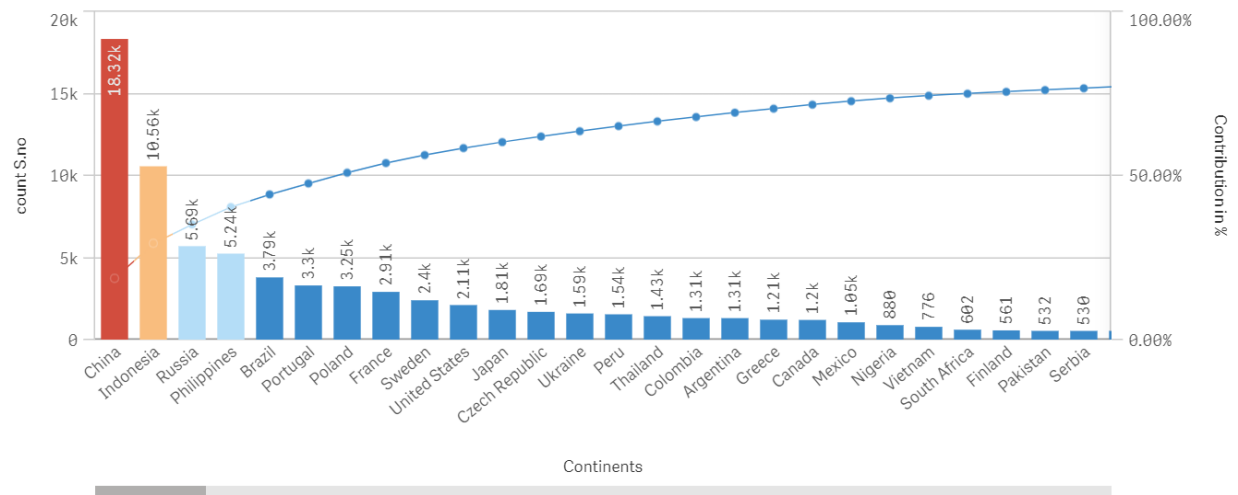
Age group as per gender



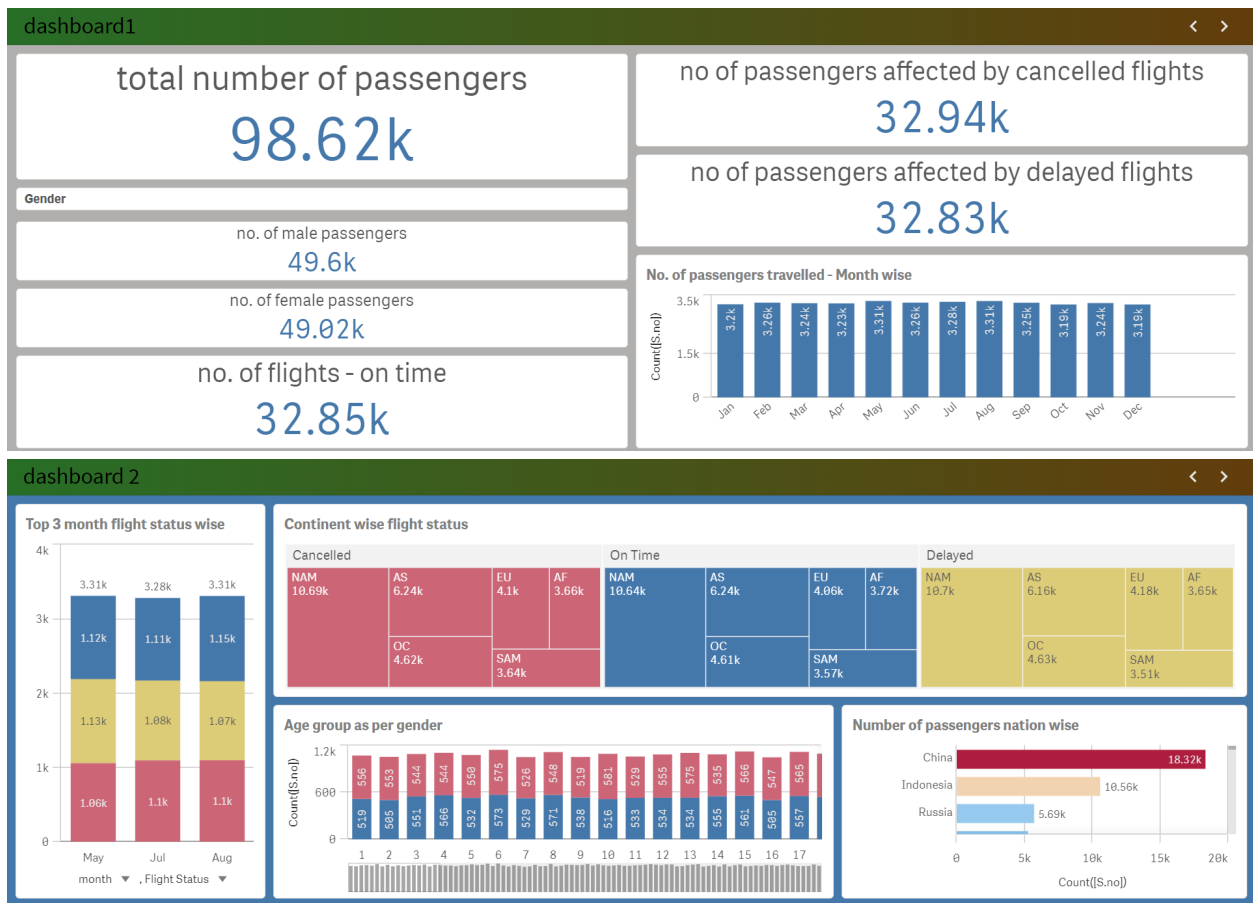
Continent wise flight status



No. of passengers - Nation wise analysis



DASHBOARD

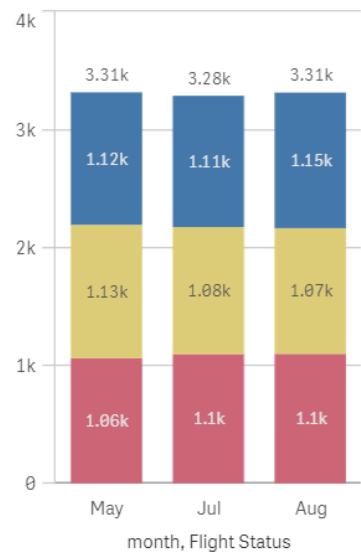


DESIGN OF STORY /REPORT

Exploring Insights from Synthetic Airline Data Analysis with Qlik

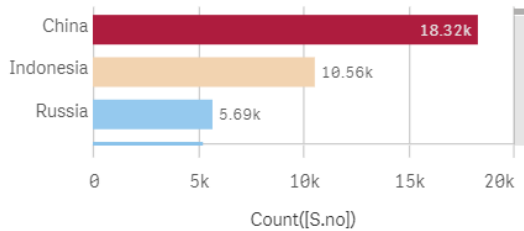
total number of passengers
98.62k
no. of flights - on time
32.85k
no of passengers affected by delayed flights
32.83k
no of passengers affected by cancelled flights
32.94k

Top 3 month flight status wise

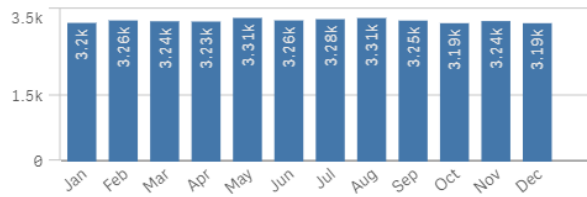


Exploring Insights from Synthetic Airline Data Analysis with Qlik

Number of passengers nation wise



No. of passengers travelled - Month wise



Continent wise flight status

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

PERFORMANCE TESTING

amount of data rendered:

https://drive.google.com/file/d/1junKarcZISh78RAj0NG_n-azEXoF4bO/view?usp=drive_link

Utilization of data filters:

