Documentation in

Qlik Sense project

In Airline Data

For SmartInternz virutual internship

By,

Aashish Minson

SICSR

Table of Content

1 Introduction

1..1 Overview: A brief description about your project

This project is to analyze the data in Airlines using Qlik Sense, a powerful business intelligence and visual analytics tool. The main objective is to understand the workings of the dataset, how each individual data is related to each other and then using these relations, deriving conclusions or insights with trial-and-error method and visualize them in dashboard which helps all levels of businesses in airline industry to deduct and to plan and create strategies for the betterment of airlines and passengers. This involves identifying trends, patterns in flight status, analysing passenger behaviour and demographics, insights on flight routes and more.

1.2 Purpose: The use of this project. What can be achieved using this.

By utilizing powerful capabilities of business intelligence and visual analytics of Qlik Sense, dataset of airlines, and understanding of dataset, we can transform data in dataset into actionable insights. This helps in enhancement in decision making, improved operational efficiency, strategic planning, customer experience management, and risk management. This will help in planning for airline strategic goals of airlines, increasing the comfort of passengers and planning for optimum operational efficiency which can result in driving business performance and customer satisfaction.

1.3 Technical Architecture

Here, this project revolves around Qlik Sense, airline dataset and the user who uses Qlik Sense with trial-and-error method derives insights from the given dataset. For this project, we used 30-day free trial Qlik Sense Cloud with the dataset Airlines.csv downloaded from Kaggle.

First, we downloaded the dataset from Kaggle, uploaded the dataset into Qlik Cloud Sense and then pre-processed the dataset. Then we derived new column names for easy handling of data and then created insights in dashboard and then created stories from this. Please find below for same dataset screenshot.

Passenger	First Name	Last Name	Gender	Age	Nationality	Airport Na	Airport Co	Country N	Airport Co	Continent	Departure	Arrival Air	Pilot Name	Flight Status
10856	Edithe	Leggis	Female	62	Japan	Coldfoot A	US	United Sta	NAM	North Am	6/28/2022	CXF	Edithe Leg	On Time
43872	Elwood	Catt	Male	62	Nicaragua	Kugluktuk	CA	Canada	NAM	North Am	12/26/202	YCO	Elwood Ca	On Time
42633	Darby	Felgate	Male	67	Russia	Grenoble-	FR	France	EU	Europe	1/18/2022	GNB	Darby Felg	On Time
78493	Dominica	Pyle	Female	71	China	Ottawa / (	CA	Canada	NAM	North Am	9/16/2022	YND	Dominica	Delayed
82072	Bay	Pencost	Male	21	China	Gillespie F	US	United Sta	NAM	North Am	2/25/2022	SEE	Bay Penco	On Time
39630	Lora	Durbann	Female	55	Brazil	Coronel H	BR	Brazil	SAM	South Am	########	LEC	Lora Durb	On Time
11940	Rand	Bram	Male	73	Ivory Coas	Duxford A	GB	United Kin	EU	Europe	10/30/202	QFO	Rand Bran	Cancelled

2 Problem Understanding

2.1 Specify the business problem

The airline industry is a complex and dynamic market where political, economical and social issues can affect the normal flight operations. Therefore, it is crucial to learn from past history, learn new insights and predict for new issues from past and present data.

Given dataset contains tons of data 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

It is a difficult task to extract meaningful insights from given dataset and the business problem is to identify and understand the dataset to analyze the data to get insights that can help improve airline operations and customer satisfaction.

2.2 Business requirements

In a complex and dynamic airline domain, business requirements keep on changing with market changes. To improve operational efficiency, enhance customer satisfaction, and support strategic decision-making in the airline industry, we need insights from analysis of dataset in visual format which can help all business levels of the domain to take quick and meaningful decision to stay ahead of competitors improve customer satisfaction and smooth operational efficiency.

2.3 Literature Survey

As airline domain is a complex and dynamic industry which can affect its operational efficiency with external interferences, there are many studies done for the same. Some studies are:

- A comprehensive survey titled “Data Analytics for Air Travel Data: A Survey and New Perspectives” by Haiman Tian et al.

- A study titled “Topic Modelling and Sentiment Analysis of Online Review for Airlines”

and more. These studies dedicate their time and effort to learn and highlight the need and scope of data analytics in transforming airline data into meaningful insights.

3 Data collection

3.1 Collect the dataset

For data to be analysed, we downloaded the data Airlines.csv from Kaggle which has the comprehensive dataset of airline passengers and their flights. The dataset includes details 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.

Passenger	First Name	Last Name	Gender	Age	Nationality	Airport Na	Airport Co	Country N	Airport Co	Continent	Departure	Arrival Air	Pilot Nam	Flight Status
10856	Edithe	Leggis	Female	62	Japan	Coldfoot A	US	United Sta	NAM	North Am	6/28/2022	CXF	Edithe Leg	On Time
43872	Elwood	Catt	Male	62	Nicaragua	Kugluktuk	CA	Canada	NAM	North Am	12/26/202	YCO	Elwood Ca	On Time
42633	Darby	Felgate	Male	67	Russia	Grenoble-	FR	France	EU	Europe	1/18/2022	GNB	Darby Felg	On Time
78493	Dominica	Pyle	Female	71	China	Ottawa / (	CA	Canada	NAM	North Am	9/16/2022	YND	Dominica	Delayed
82072	Bay	Pencost	Male	21	China	Gillespie F	US	United Sta	NAM	North Am	2/25/2022	SEE	Bay Penco	On Time
39630	Lora	Durbann	Female	55	Brazil	Coronel H	BR	Brazil	SAM	South Am	########	LEC	Lora Durb	On Time
11940	Rand	Bram	Male	73	Ivory Coas	Duxford A	GB	United Kin	EU	Europe	10/30/202	QFO	Rand Bran	Cancelled

The data is collected from various sources, including airline databases and passenger management systems. The data is provided in a structured format (such as CSV or Excel), which makes it easy to upload into Qlik Sense Cloud for analysis. Data collection is a crucial step in any data analytics project, as the quality and relevance of the data directly impact the accuracy and usefulness of the analysis results.

3.2 Connect Data with Qlik Sense

Once data is downloaded from Kaggle, it is then analysed to understand the data in the dataset and then uploaded to Qlik Sense Cloud where we can preprocess the data and perform necessary actions to get insights.

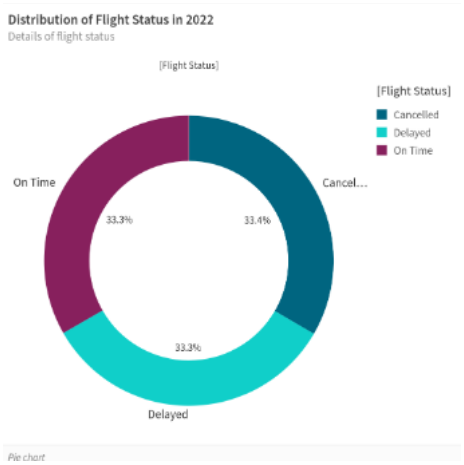
4 Data Preparation

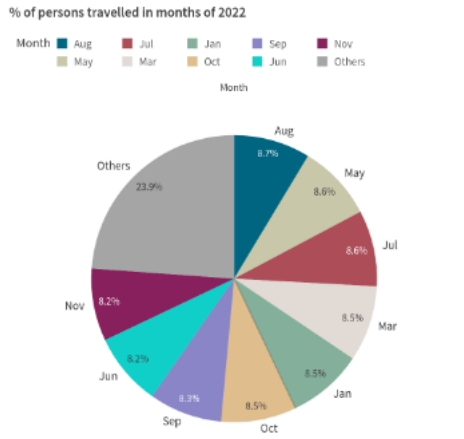
Once dataset is uploaded into Qlik Sense Cloud, we then analyze the data for null values. After removing null values, we then determine if there is any data which has no meaning which we can drop. Then we create new columns to extract meaningful data into it which helps in creating meaningful visual insights.

5 Data Visualizations

Here, after performing all data pre-processing and deletions, we then determine meaningful insights which we then convert into visual analytics using powerful tools of Qlik Sense Cloud which will help the viewer to understand the insights in first understanding and then take necessary actions or plans for smooth operation and customer satisfaction.

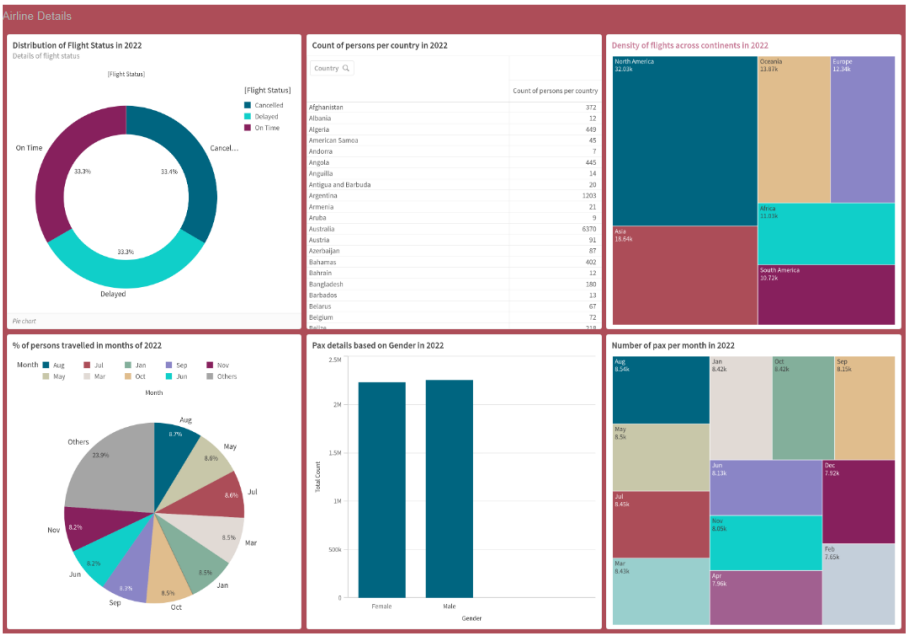
Please find below samples of related visual insights:

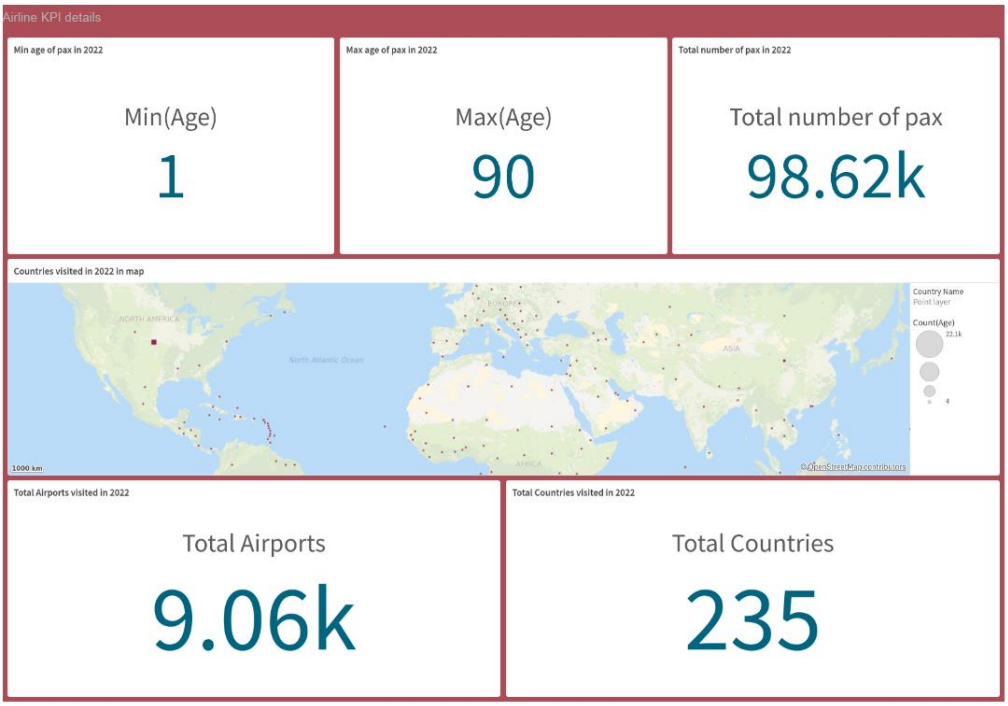




6 Dashboard

Please find below for related dashboards:





7 Report

The report creation involves identifying the insights, analysing them and then visualizing the insights in structured way in dashboard. Then we convey the insights to non-technical users and let them derive insights and make data-driven decisions.

8 Performance Testing

8.1 Amount of Data Rendered

Performance testing is a crucial aspect of a data analytics project as from this, we can determine whether the system can handle given amount of data and show the expected performance. Here, we processed 98K records from the dataset.

8.2 Utilization of Data filters

Data filters are a critical part of any business intelligence tool as they enable users to focus on specific subsets of data. However, the utilization of data filters can significantly impact the performance of the system, especially when dealing with large datasets.

9 Conclusion

From this project, we get to learn on how to load data into Qlik Sense Cloud, create student account, how to upload dataset, how to pre-process dataset, how to understand the data, how to create visual insights with data and measures which will be helpful for business users who can streamline operations and enhance customer satisfaction in airline industry.