

# **PROJECT REPORT TITLES**

## **1.1 Overview**

The project "Exploring Insights from Synthetic Airline Data Analysis with Qlik" makes use of Qlik, a potent business intelligence and data visualization platform, to extract relevant insights from synthetic data representing many aspects of airline operations. Flight schedules, passenger demographics, and performance indicators are all included in the synthetic data, which offers a comprehensive dataset to mimic actual airline operations. The main goal is to enhance airline, airport, and related stakeholder decision-making processes by using Qlik's analytical capabilities to find patterns, trends, and correlations within this data.

## **1.2 Purpose**

The primary purpose of this project is to demonstrate how Qlik can be utilized to analyze and visualize synthetic airline data, leading to actionable insights across multiple scenarios:

**Revenue Optimization:** Airlines can determine popular locations, peak travel periods, and efficient pricing methods by examining past data on ticket sales. By using Qlik to visualize these information, pricing and

marketing tactics can be changed to optimize profitability.

**Operational Efficiency:** By looking at flight schedules, passenger flows, and luggage handling procedures, airport authorities can improve operational efficiency. Qlik may be used to predict peak traffic times, find operational bottlenecks, and allocate resources optimally, which can optimize workflows and boost productivity.

**Consumer Experience Enhancement:** By using sentiment analysis on feedback data, airlines can better understand consumer preferences and pain spots, leading to increased passenger happiness. Airlines may improve customer happiness and loyalty by personalizing services, identifying areas for improvement, and adjusting marketing campaigns with the use of Qlik's visualization capabilities.

## **1.3 Technical Architecture**

**Data Source:** The synthetic airline data, encompassing multiple datasets such as aircraft schedules, passenger demographics, ticket sales, and performance indicators, is the source of the data. These datasets might be kept in databases or in different file formats like CSV or JSON.

**Data Ingestion:** Third-party or built-in connectors can be used to link Qlik Sense or QlikView to the data sources. Data extraction from many sources and input into Qlik's associative data model constitute data ingestion.

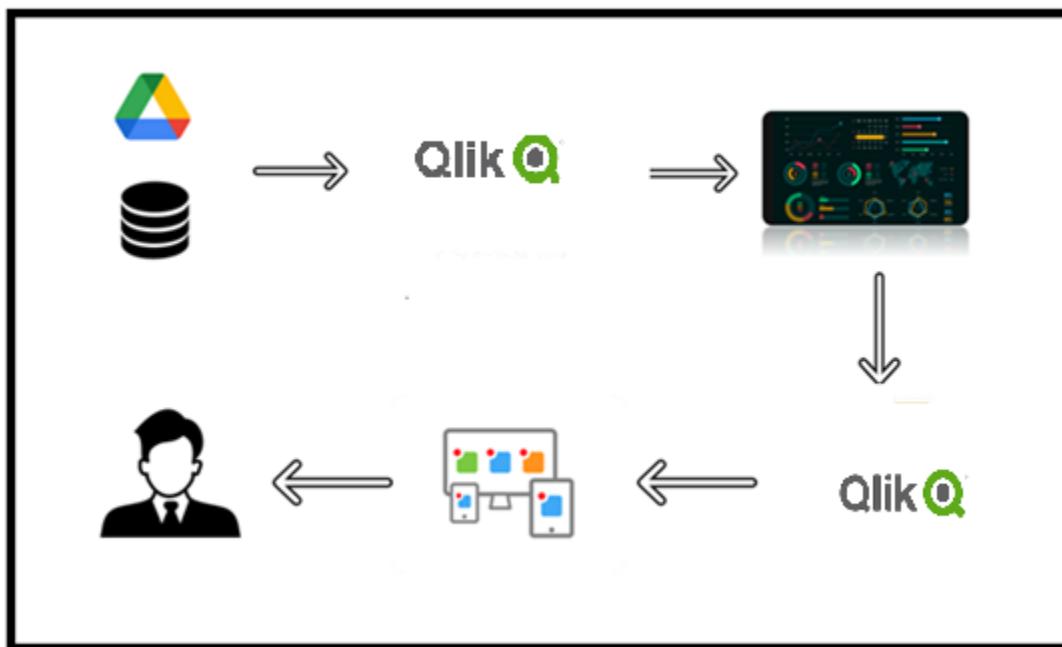
**Data Transformation:** To guarantee that the data is in the proper format for analysis, Qlik transforms and cleans the data. This process involves applying business rules, filtering, aggregating, and merging data.

**Data modeling:** Relationships between various datasets can be created using Qlik's associative model. Through dynamic data exploration and analysis made possible by this paradigm, users can engage with the data to uncover insights.

**Visualization Layer:** Qlik uses the modified data to build a variety of visuals, including scatter plots, bar charts, line charts, and dashboards. Users are better able to investigate the data, spot trends, and gain new insights thanks to these visualizations.

**Analytical Capabilities:** Qlik's robust analytics engine facilitates set analysis, sophisticated data modeling, and intricate computations, providing deep insights into the data. Actionable insights can be obtained through interactive dashboards and reports that users can generate.

**User Interface:** Using the Qlik interface, the visuals and insights are shown in the last stage.



## **2.1 Specifying the Business Problem**

The "Exploring Insights from Synthetic Airline Data Analysis with Qlik" project tackles a variety of complex business issues that relate to many facets of airline operations:

**Revenue Optimization:** In an environment of competitive pricing and varying demand, airlines must maximize revenue. To increase profitability, they must determine the busiest travel periods, well-liked travel locations, and the best pricing schemes.

**Operational Efficiency:** By efficiently controlling flight schedules, passenger flows, and baggage handling procedures, airport authorities aim to improve operational efficiency. They must identify bottlenecks, forecast times of high traffic, and effectively distribute resources.

**Enhancing the Customer Experience:** By learning about the preferences, levels of happiness, and pain spots of their customers, airlines hope to increase passenger satisfaction. To promote loyalty and happiness, they require information to address areas for improvement and personalize offerings.

### **MAIN BUSINESS PROBLEM USED IN THIS PROJECT:**

**Challenges Faced by Senior Citizens in Flight Delay and Cancellation:**

- Physical strain from prolonged waits.
- Disrupted medical needs.
- Inadequate facilities.
- Stress and anxiety.
- Financial strain.
- Safety concerns.
- Difficulty in coordination with family.
- Need for better support and clear communication.

## **2.2 Business Requirements**

### **Maximizing Revenue**

**Data analysis:** Look for trends and patterns in the past ticket sales data.

**Visualization:** Show historical revenue trends, peak travel periods, and well-liked locations.

**Consumer Segmentation:** To customize pricing tactics, divide up your consumer base according to their purchasing habits.

**Pricing Strategy:** Using data-driven insights, modify pricing strategies to optimize profitability.

### **Efficiency of Operations**

**Analysis of Flight Schedules:** Examine flight schedules to find inefficiencies.

**Analyze passenger flow:** To forecast peak travel times and enhance the distribution of resources.

**Processes for Handling Baggage:** Examine luggage handling procedures to find and fix bottlenecks.

**Resource Allocation:** To improve productivity and simplify processes, allocate resources wisely.

### **Enhancement of Customer Experience**

**Customer Feedback Analysis:** To identify pain spots and satisfaction levels, use sentiment analysis on customer feedback data.

**Personalization:** Tailor services and marketing initiatives with the help of client data insights.

**Service Enhancement:** To improve the overall traveler experience, pinpoint areas that need attention.

## **2.3 Literature Survey**

A literature survey involves reviewing existing research, studies, and articles related to the project topics. For this project, the literature survey includes:

### **Maximizing Revenue:**

Research on dynamic pricing models in the airline sector, with an emphasis on how airlines can instantly modify prices in response to market conditions and rivalry.

Research on demand forecasting methods for estimating peak travel periods and well-liked locations.

Articles on revenue management systems and how airlines are using them to increase profitability.

### **Efficiency of Operations:**

**Airport Operations Management Literature:** This section covers topics such as baggage handling, flight scheduling, and passenger flow management.

Bottleneck analysis is the study of finding and fixing operational bottlenecks at airports to increase productivity.

Research on efficient methods for allocating resources to optimize airport operations.

### **Enhancement of Customer Experience**

**Customer Sentiment Analysis:** Studying how to use feedback data to discover pain spots and measure customer satisfaction through sentiment analysis techniques.

**Personalization in Marketing:** Articles about marketing tactics that emphasize the use of consumer data by airlines to customize campaigns and offerings.

**Customer Loyalty Programs:** Research on methods to increase customer loyalty as well as the efficacy of such programs in the airline sector.

## 3.1 Collect the Dataset

The screenshot shows the Qlik Sense dataset overview for 'AIRLINE DATASET PROJECT WORK.csv'. The dataset was last refreshed 3 days ago. It contains 15 fields and 98,619 rows. The owner and creator are OJASVI RAJ. The file is 12.7 MB and is located in the AIRLINE DATASET PROJECT WORK space.

Type	File size	Source	Space	Owner	Creator
DELIMITED	12.7 MB	AIRLINE DATASET PROJECT WO...	Personal	OJASVI RAJ	OJASVI RAJ

Metadata details:

Created	Modified	Fields	Row count	Viewed by	Used in
3 days ago	3 days ago	15	98,619	1 User	0

Profile refreshed 3 days ago.

This shows the dataset collected as .csv file

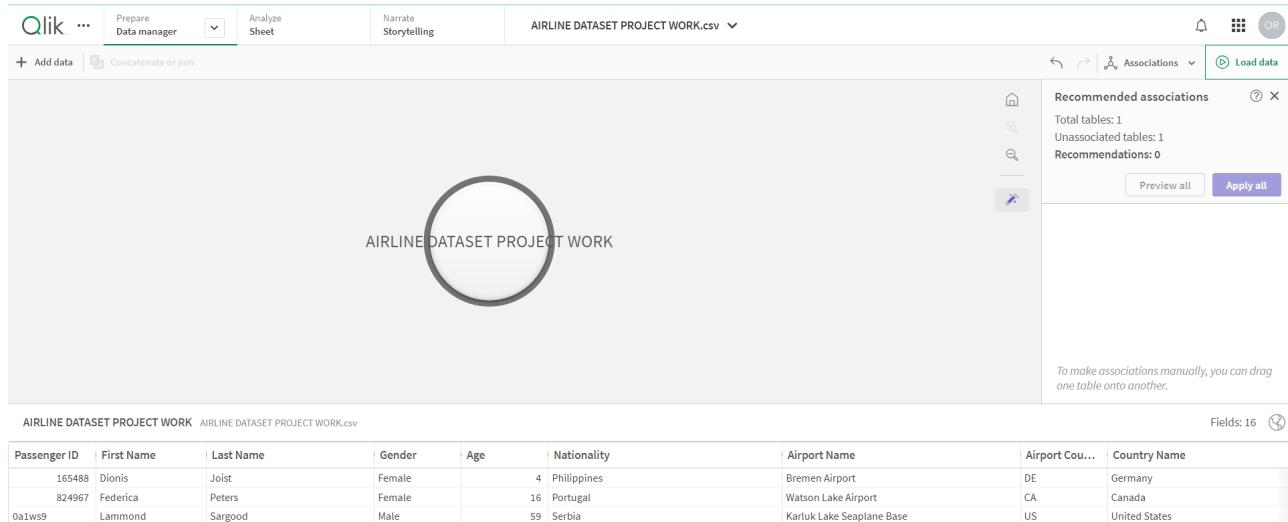
## 3.2 Connect Data with Qlick Sense

The screenshot shows the Qlik Sense data manager interface with the dataset 'AIRLINE DATASET PROJECT WORK' selected. A large circular placeholder is visible in the center of the workspace. The bottom half of the screen displays a preview of the dataset table.

Passenger ID	First Name	Last Name	Gender	Age	Nationality	Airport Name	Airport Cou...	Country Name
165488	Dionis	Joist	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
0A20YI	Jack	Mitrikhin	Male	89	Turkey	Rottnest Island Airport	AU	Australia
0a5x5P	Perry	Pretsell	Male	64	Nigeria	Vallenar Airport	CL	Chile

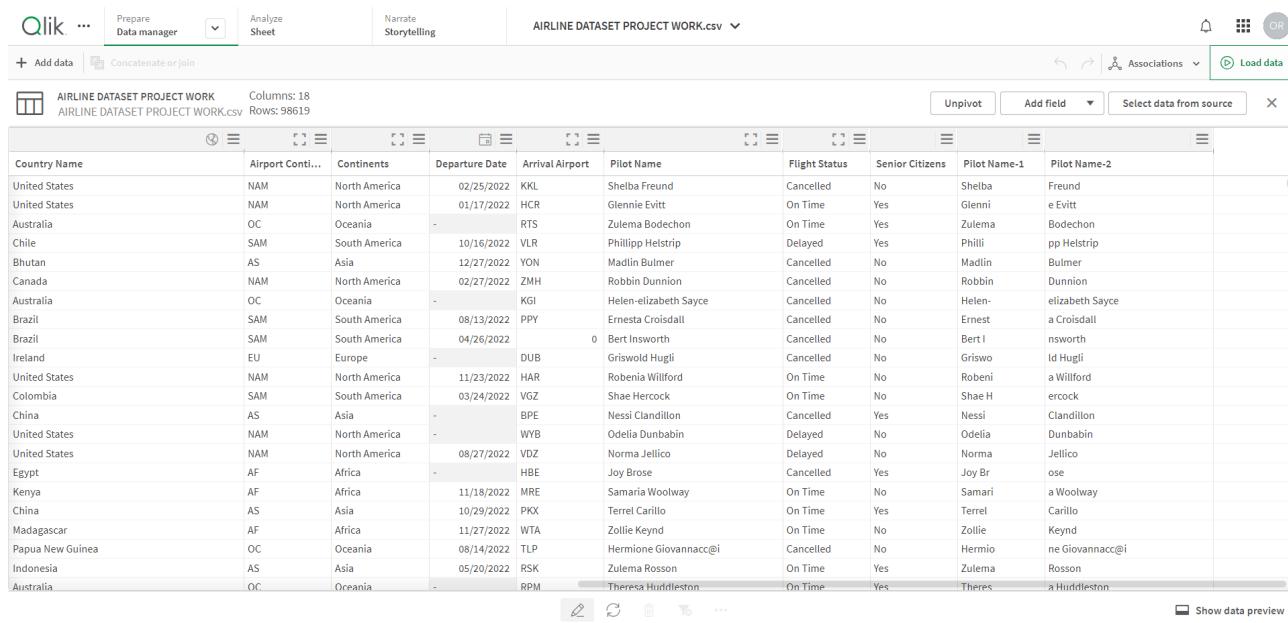
This shows the dataset connected with Qlik Sense.

## 4.1 Prepare the Data for Visualization



AIRLINE DATASET PROJECT WORK AIRLINE DATASET PROJECT WORK.csv Fields: 16

Passenger ID	First Name	Last Name	Gender	Age	Nationality	Airport Name	Airport Cou...	Country Name
165488	Dionis	Joist	Female	4	Philippines	Bremen Airport	DE	Germany
824967	Federica	Peters	Female	16	Portugal	Watson Lake Airport	CA	Canada
0a1ws9	Lammond	Sargood	Male	59	Serbia	Karuk Lake Seaplane Base	US	United States
0AiyOa	Arel	Beswick	Male	83	United States	Holy Cross Airport	US	United States
0A2OYI	Jack	Mitrikhin	Male	89	Turkey	Rottnest Island Airport	AU	Australia
0a5xSP	Perry	Pretsell	Male	64	Nigeria	Vallenar Airport	CL	Chile

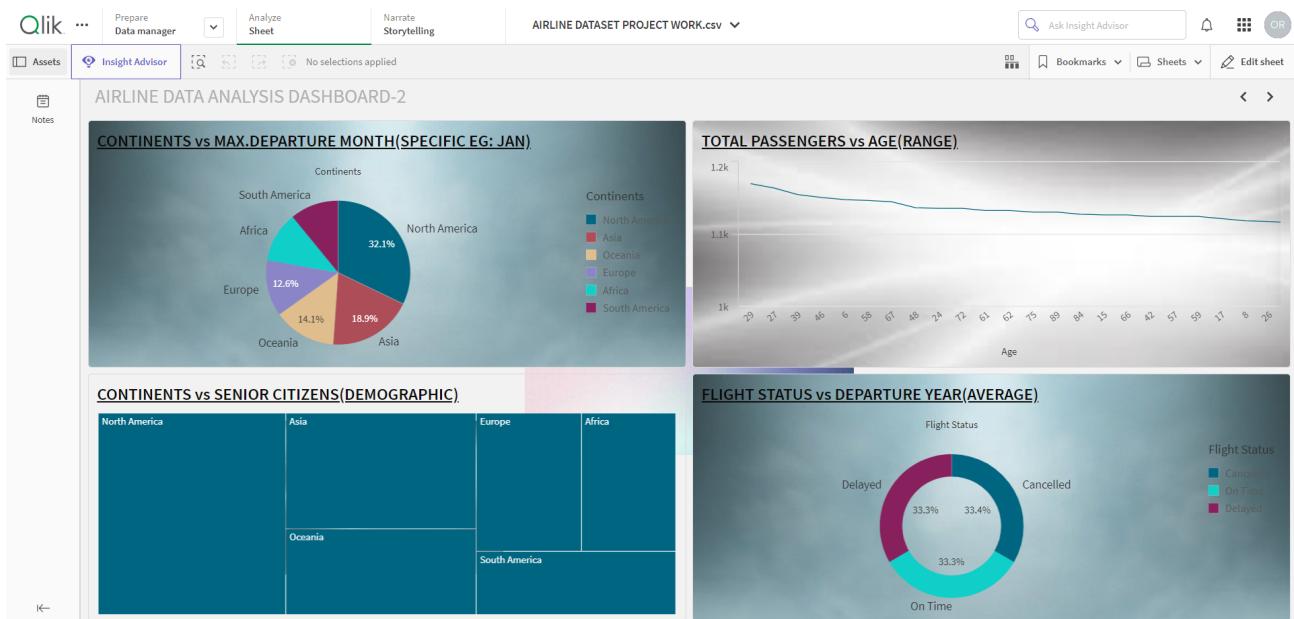
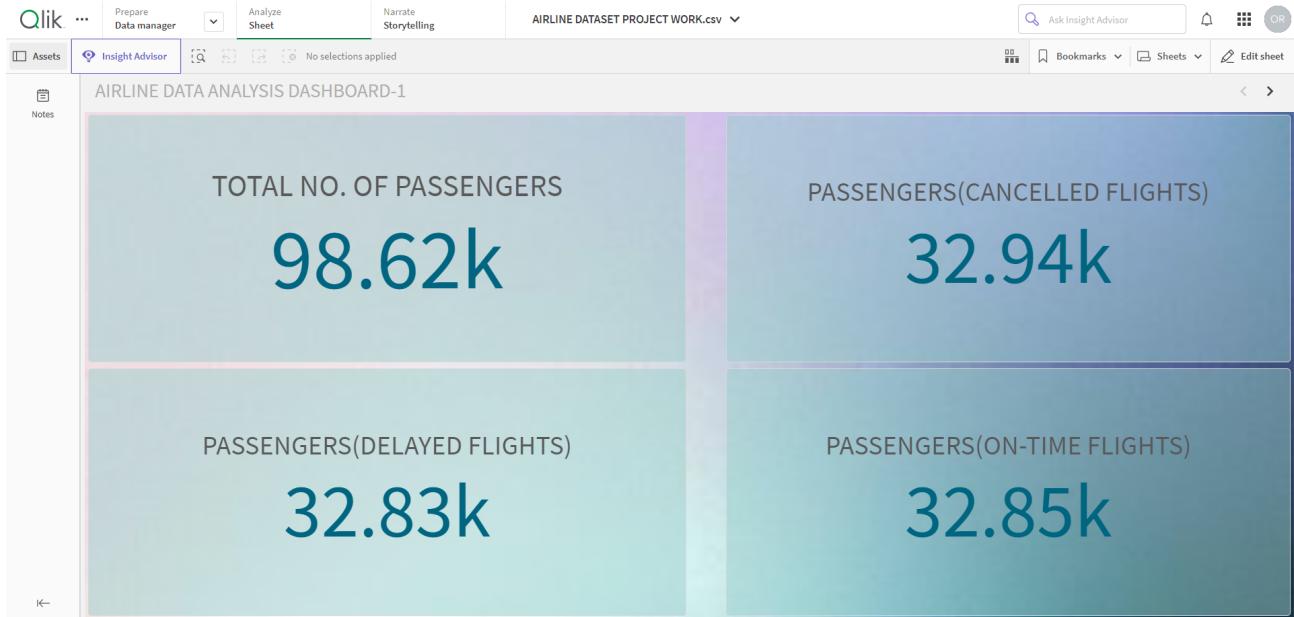


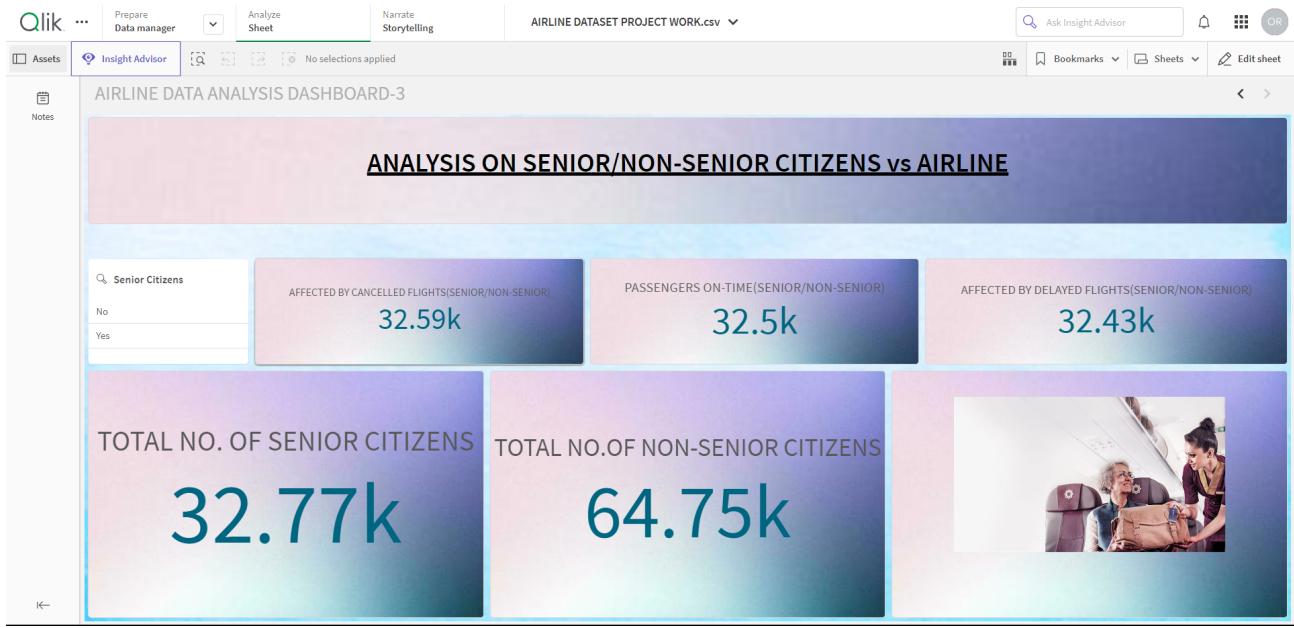
AIRLINE DATASET PROJECT WORK Columns: 18 AIRLINE DATASET PROJECT WORK.csv Rows: 98619

Country Name	Airport Conti...	Continents	Departure Date	Arrival Airport	Pilot Name	Flight Status	Senior Citizens	Pilot Name-1	Pilot Name-2
United States	NAM	North America	02/25/2022	KKL	Shelba Freund	Cancelled	No	Shelba	Freund
United States	NAM	North America	01/17/2022	HCR	Glenne Evitt	On Time	Yes	Glenne	Evitt
Australia	OC	Oceania	-	RTS	Zulema Bodechon	On Time	Yes	Zulema	Bodechon
Chile	SAM	South America	10/16/2022	VLR	Phillip Helstrip	Delayed	Yes	Phillip	Helstrip
Bhutan	AS	Asia	12/27/2022	YON	Madin Bulmer	Cancelled	No	Madin	Bulmer
Canada	NAM	North America	02/27/2022	ZMH	Robbin Dunnion	Cancelled	No	Robbin	Dunnion
Australia	OC	Oceania	-	KGI	Helen-elizabeth Sayce	Cancelled	No	Helen-elizabeth	Sayce
Brazil	SAM	South America	08/13/2022	PPY	Ernesta Croisdall	Cancelled	No	Ernest	Croisdall
Brazil	SAM	South America	04/26/2022	0	Bert Insworth	Cancelled	No	Bert I	nsworth
Ireland	EU	Europe	-	DUB	Griswold Hugli	Cancelled	No	Griswo	ld Hugli
United States	NAM	North America	11/23/2022	HAR	Robenia Willford	On Time	No	Robeni	a Willford
Colombia	SAM	South America	03/24/2022	VGZ	Shae Hercock	On Time	No	Shae H	ercock
China	AS	Asia	-	BPE	Nessi Clandillon	Cancelled	Yes	Nessi	Clandillon
United States	NAM	North America	-	WVB	Odelia Dunbabin	Delayed	No	Odelia	Dunbabin
United States	NAM	North America	08/27/2022	VDZ	Norma Jellico	Delayed	No	Norma	Jellico
Egypt	AF	Africa	-	HBE	Joy Brose	Cancelled	Yes	Joy Br	ose
Kenya	AF	Africa	11/18/2022	MRE	Samaria Woolway	On Time	No	Samari	a Woolway
China	AS	Asia	10/29/2022	PKX	Terrel Carillo	On Time	Yes	Terrel	Carillo
Madagascar	AF	Africa	11/27/2022	WTA	Zollie Keynd	On Time	No	Zollie	Keynd
Papua New Guinea	OC	Oceania	09/14/2022	TLP	Hermione Giovannacci	Cancelled	No	Hermio	ne Giovannacci
Indonesia	AS	Asia	05/20/2022	RSK	Zulema Rosson	On Time	Yes	Zulema	Rosson
Australia	OC	Oceania	-	RPM	Theresa Hurdleston	On Time	Yes	Theres	a Hurdleston

These show the prepared format for data visualizations.

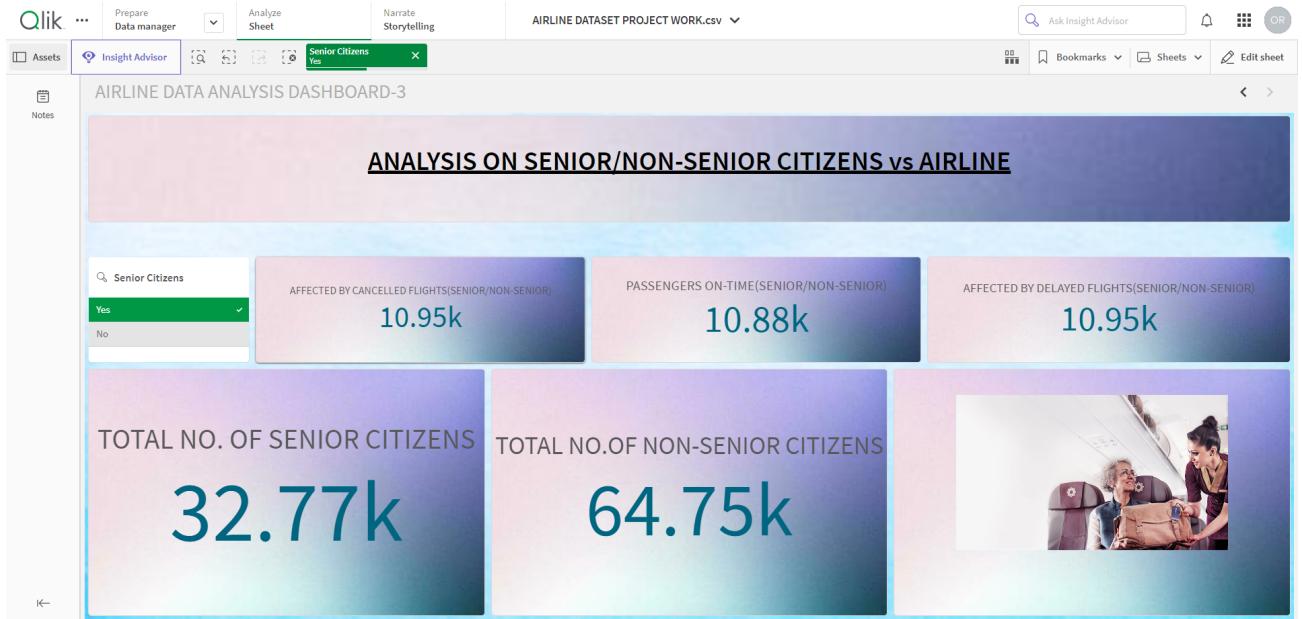
# 5.1 Visualization



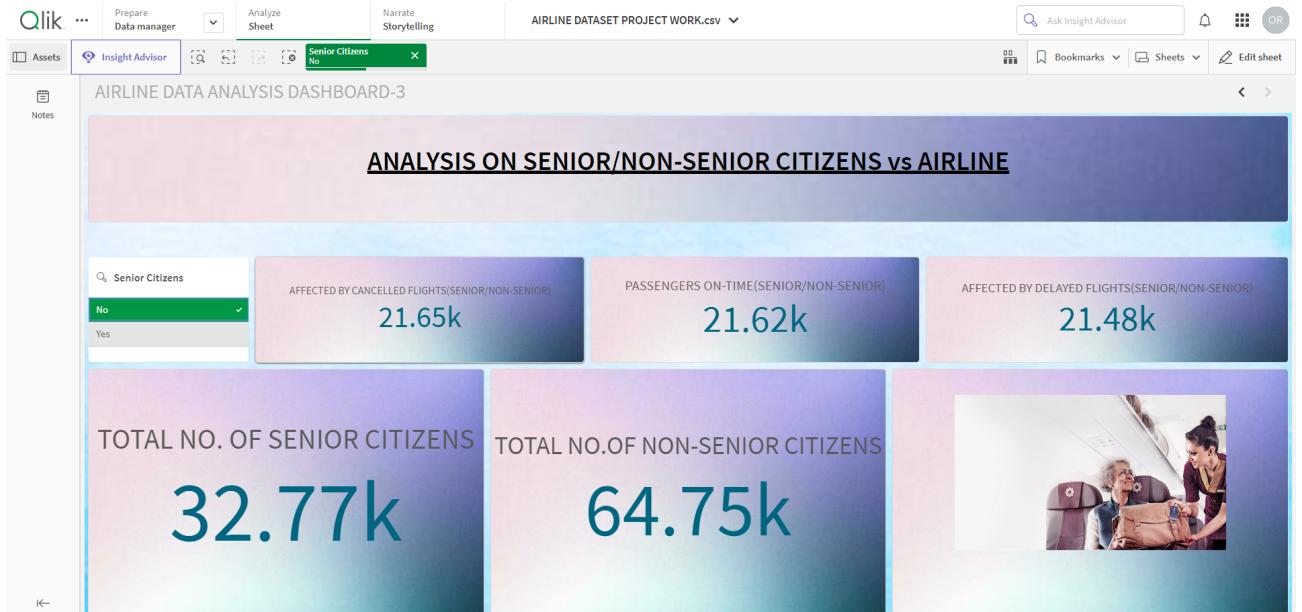


**These slides show the simplified visualizations which balance the detail of information with rendering performance.**

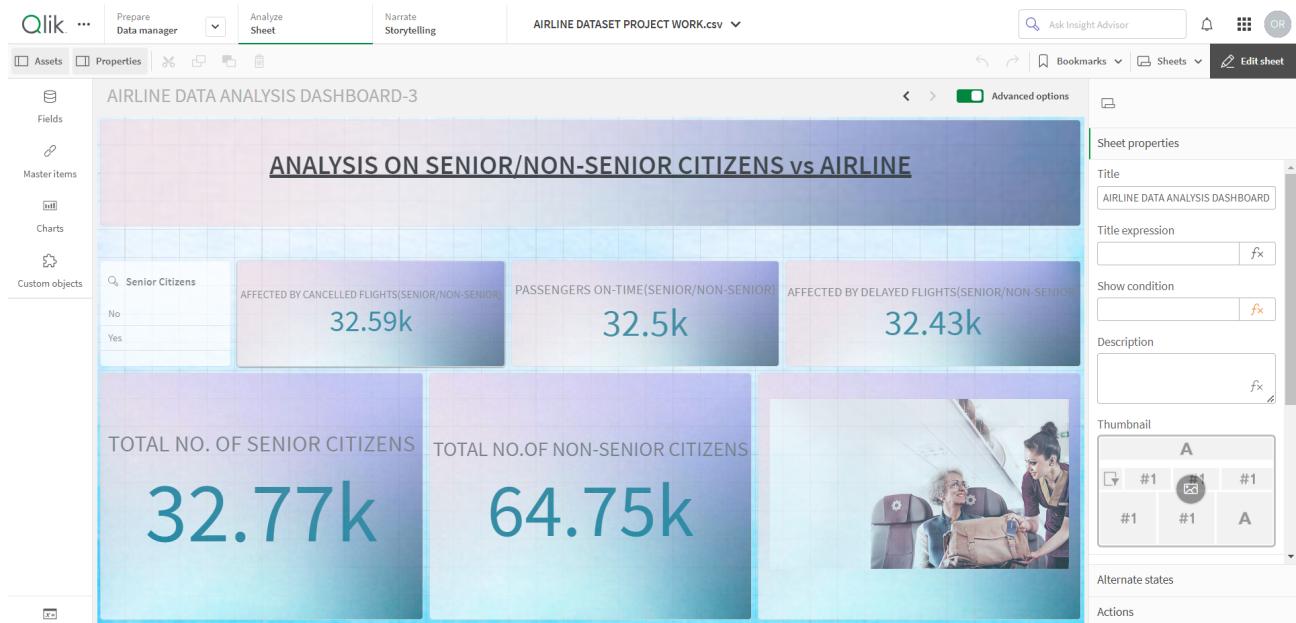
## 6.1 Responsive and Design of Dashboard



This shows the responsive dashboard which on applying the SENIOR CITIZENS filter as 'YES' shows the analysis on Senior Citizens vs Airline.



This shows the responsive dashboard which on applying the SENIOR CITIZENS filter as 'NO' shows the analysis on Non-Senior Citizens vs Airline.



This slide shows the designed dashboard (AIRLINE DATA ANALYSIS-3)

## 7.1 Report Creation



## No Of Visualizations/ Graphs

- 1) Total Number of Passengers
- 2) Number of Passengers effected by Cancelled Flights
- 3) Number of Passengers effected by Delay of Flights
- 4) Number of Flights-On-Time
- 5) Continents vs Max. Departure Month(Specific e.g. : Jan)
- 6) Total Passengers vs Age(Range)
- 7) Responsive Dashboard which on applying the SENIOR CITIZENS filter as 'YES' shows the analysis on Senior Citizens vs Airline.
- 8) Responsive Dashboard which on applying the SENIOR CITIZENS filter as 'NO' shows the analysis on Non-Senior Citizens vs Airline.

## EXPLORING INSIGHTS FROM SYNTHETIC AIRLINE DATA ANALYSIS

TOTAL NO. OF PASSENGERS

**98.62k**

PASSENGERS(CANCELLED FLIGHTS)

**32.94k**

PASSENGERS(DELAYED FLIGHTS)

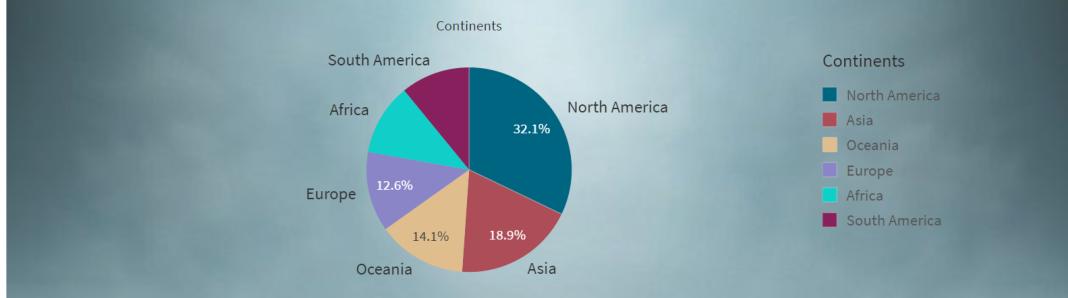
**32.83k**

PASSENGERS(ON-TIME FLIGHTS)

**32.85k**

## EXPLORING INSIGHTS FROM SYNTHETIC AIRLINE DATA ANALYSIS

CONTINENTS vs MAX.DEPARTURE MONTH(SPECIFIC EG: JAN).

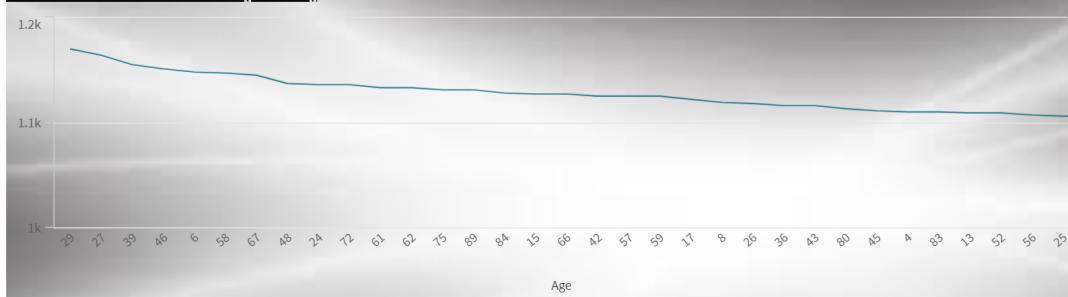


This chart displays the Continents with the data of a specific max Departure Month (for e.g: January)

...

## EXPLORING INSIGHTS FROM SYNTHETIC AIRLINE DATA ANALYSIS

TOTAL PASSENGERS vs AGE(RANGE).



This chart displays the total no. of passengers within a range of age.

...

## EXPLORING INSIGHTS FROM SYNTHETIC AIRLINE DATA ANALYSIS

TOTAL NO. OF SENIOR CITIZENS

**32.77k**

**SENIOR CITIZENS**  
Click here

FFECTED BY CANCELLED FLIGHTS(SENIOR/NON-SENIOR)

**10.95k**

PASSENGERS ON-TIME(SENIOR/NON-SENIOR)

**10.88k**

FFECTED BY DELAYED FLIGHTS(SENIOR/NON-SENIOR)

**10.95k**

## EXPLORING INSIGHTS FROM SYNTHETIC AIRLINE DATA ANALYSIS

TOTAL NO.OF NON-SENIOR CITIZENS

**64.75k**

**NON SENIOR CITIZENS**  
Click here

FFECTED BY CANCELLED FLIGHTS(SENIOR/NON-SENIOR)

**21.65k**

PASSENGERS ON-TIME(SENIOR/NON-SENIOR)

**21.62k**

FFECTED BY DELAYED FLIGHTS(SENIOR/NON-SENIOR)

**21.48k**

**Creating reports in Qlik involves designing and developing interactive dashboards along with the use of bookmark and visualizations that provide meaningful insights from the data.**

## 8.1 Amount of Data Rendered

The screenshot displays the Qlik Sense interface with two main sections: Overview and Fields.

**Overview Page:**

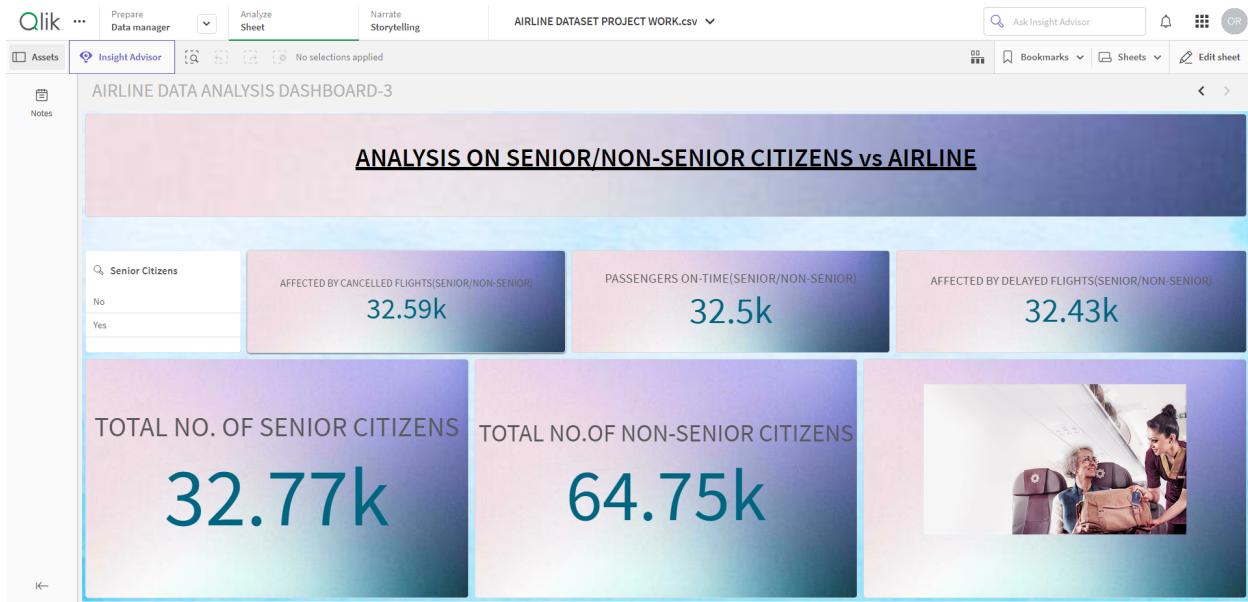
- Dataset Information:** AIRLINE DATASET PROJECT WORK.csv (refreshed 3 days ago)
- Type:** DELIMITED
- File size:** 12.7 MB
- Source:** AIRLINE DATASET PROJECT WO...
- Space:** Personal
- Owner:** OJASVI RAJ
- Creator:** OJASVI RAJ
- Metadata created:** 3 days ago
- Metadata modified:** 3 days ago
- Fields:** 15
- Row count:** 98,619
- Viewed by:** 1 User: 1
- Used in:** 0

**Fields Page:**

- Classifications:** No classifications are applied. (+ Add classification)
- Fields:** ABVWlg, jkXXAX, CdUz2g, +98616 additional unique values; Passenger ID (Distinct values: 98,619); First Name (Distinct values: 8,437), showing a bar chart for top names: Gale (37), Brett (36), Gerrie (35), Conny (35), Other (98,476); Last Name (Distinct values: 41,658), showing a bar chart for top names: Dyball (17), Capron (15), Ducker (13), Richfield (12), Other (98,562); Gender (Distinct values: 2), showing a bar chart for Male (49,598) and Female (49,021); Age (Distinct values: 90), showing a histogram with an average of 45.5; Nationality (Distinct values: 240), showing a bar chart for China (18,317), Indonesia (10,559), Russia (5,693), Philippines (5,239), Other (58,811); Airport Name (Distinct values: 9,062), showing a bar chart for San Pedro Airport (43), Santa Maria Airport (38), Böblingen Flugfeld (36), Santa Ana Airport (35), Other (98,467); Airport Country Code (Distinct values: 235), showing a bar chart for US (22,104), AU (6,370), CA (5,424), BR (4,504), Other (60,217).
- View fields:** Buttons for viewing specific field details.

**Rendering data involves loading, processing, and displaying data in charts, tables, and other visualization objects within the Qlik environment.**

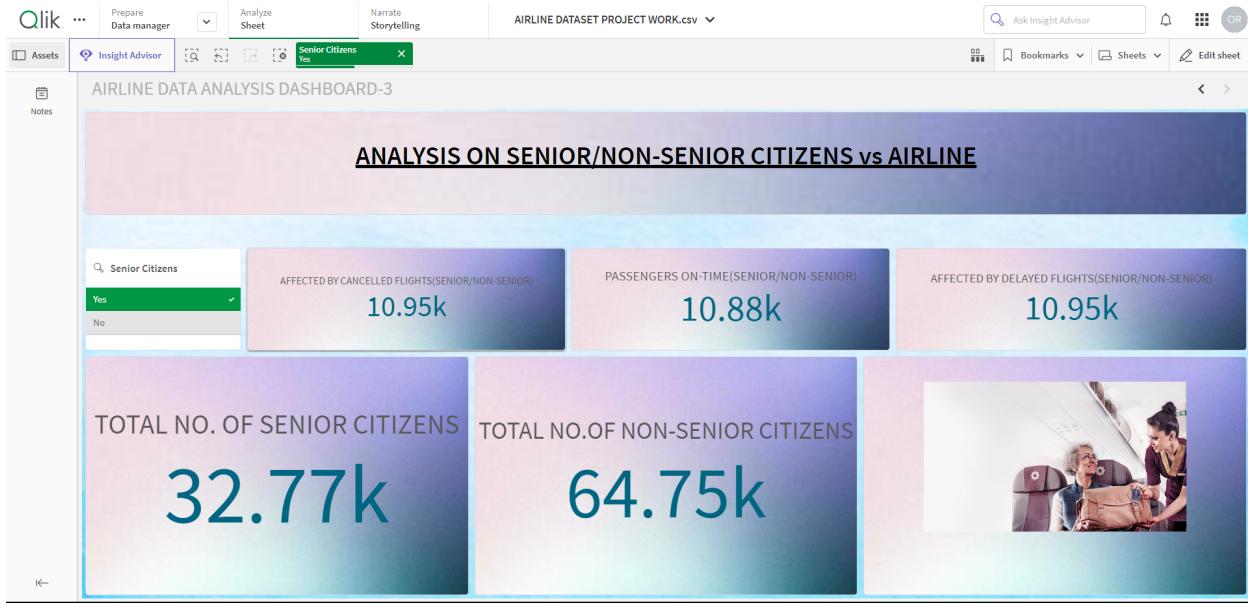
## 8.2 Utilization of Data Filters



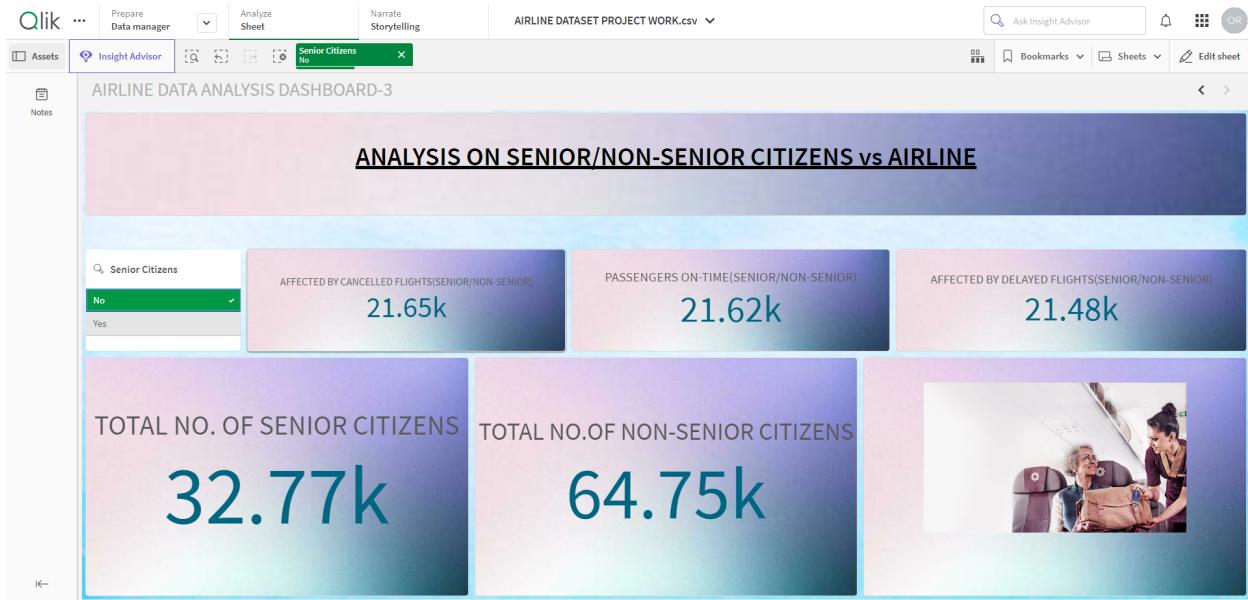
**Utilizing data filters in Qlik is a crucial aspect of interactive data analysis and visualization. Filters enable users to narrow down data sets to specific subsets, making it easier to focus on relevant information, identify trends, and gain insights.**

**Filters can be applied in various ways, such as through selection boxes, sliders, buttons, and dynamic filtering expressions.**

**Below are some detailed insights into how data filters are utilized in Qlik, along with practical examples.**



This shows the responsive dashboard which on applying the SENIOR CITIZENS filter as 'YES' shows the analysis on Senior Citizens vs Airline.



This shows the responsive dashboard which on applying the SENIOR CITIZENS filter as 'NO' shows the analysis on Non-Senior Citizens vs Airline.