Documentation for "Exploring Insights from Synthetic Airline Data Analysis with Qlik"

Introduction

The project "Exploring Insights from Synthetic Airline Data Analysis with Qlik" leverages synthetic airline data to uncover valuable insights using Qlik, a powerful business intelligence and data visualization tool. This synthetic data simulates various aspects of airline operations, including flight schedules, passenger demographics, ticket sales, and performance metrics. The goal is to utilize Qlik's analytical capabilities to identify patterns, trends, and correlations within this data, aiding decision-making processes for airlines, airports, and related stakeholders.

Procedure

1. Data Collection

-Objective: Gather synthetic airline data that simulates real-world operations.
 - Actions Taken: I collected the synthetic dataset, which includes fields such as flight schedules, passenger demographics, ticket sales, and performance metrics.

2. Data Loading

 - Objective: Import the collected data into Qlik for further analysis.
 - Steps Taken:
  - I opened Qlik Sense.
  - I used the data load editor to import the dataset into Qlik.
  - I ensured that all relevant fields were loaded properly.

3. Data Preparation

 - Objective: Prepare the data for analysis by cleaning and sorting it.
- Steps Taken:
  - I identified and handled any missing or inconsistent data.
  - I normalized data formats (e.g., date formats, text case).
  - I sorted and categorized data based on relevant fields such as departure date, passenger age, and flight status.

4. Data Cleaning and Sorting

- Objective: Ensure data quality and consistency.
- Steps Taken:
  - I removed duplicate records.
  - I corrected any erroneous data entries.
  - I validated data integrity by cross-checking with predefined rules or constraints.

5. Initial Data Visualization

- Objective: Create initial visualizations to understand data distribution and key metrics.

-Visualizations Created:

1. Total Passenger Count:

- I created a bar chart showing the total count of passengers.

2. Number of Passengers - Nation Wise:

- I used a pie chart to display the distribution of passengers by nationality.

3. Passengers by Flight Status:

- I created a stacked bar chart showing the number of passengers based on flight status (On Time, Delayed, Cancelled).

4. Count of Age by Departure Date and First Name:

- I used a line chart to depict the count of passengers' ages over time, grouped by first names.

5. Flight Travel - Age Groups:

- I created a histogram to show the distribution of different age groups among travelers.

6. Scenario-Based Analysis

-Objective: Use visualizations to address specific scenarios for actionable insights.

Scenarios Addressed:

1. Revenue Optimization:

- I analyzed historical ticket sales data.
- I visualized revenue trends over time.
- I segmented customers based on purchasing behavior.
- I adjusted pricing strategies based on the insights to maximize profitability.

2. Operational Efficiency:

- I analyzed flight schedules, passenger flows, and luggage handling processes.
- I identified bottlenecks in airport operations.
- I predicted peak traffic periods.
- I allocated resources effectively to streamline processes and improve efficiency.

**3. Customer Experience Enhancement:**

- I performed sentiment analysis on customer feedback data.
- I identified areas for improvement.
- I personalized services and tailored marketing campaigns.
- I fostered loyalty and satisfaction by addressing customer needs and pain points.
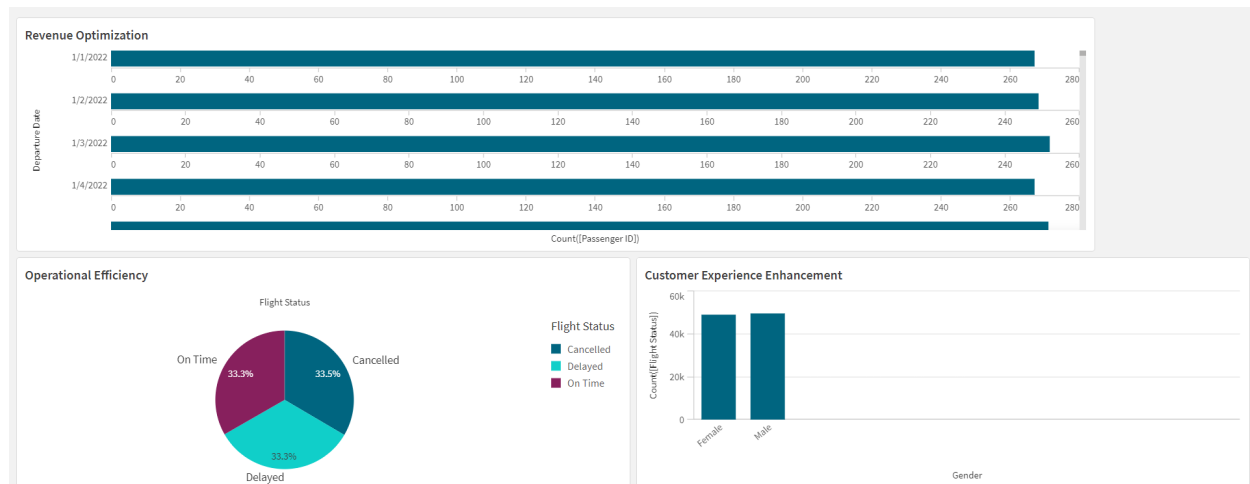
7. Final Visualization and Reporting

- Objective: Compile the findings into comprehensive visual reports.
- Steps Taken:
  - I used Qlik's dashboarding capabilities to create interactive reports.
  - I included key visualizations that address the scenarios.
 - I summarized insights and provided actionable recommendations.

8. Dashboard

-Dashboard 1:
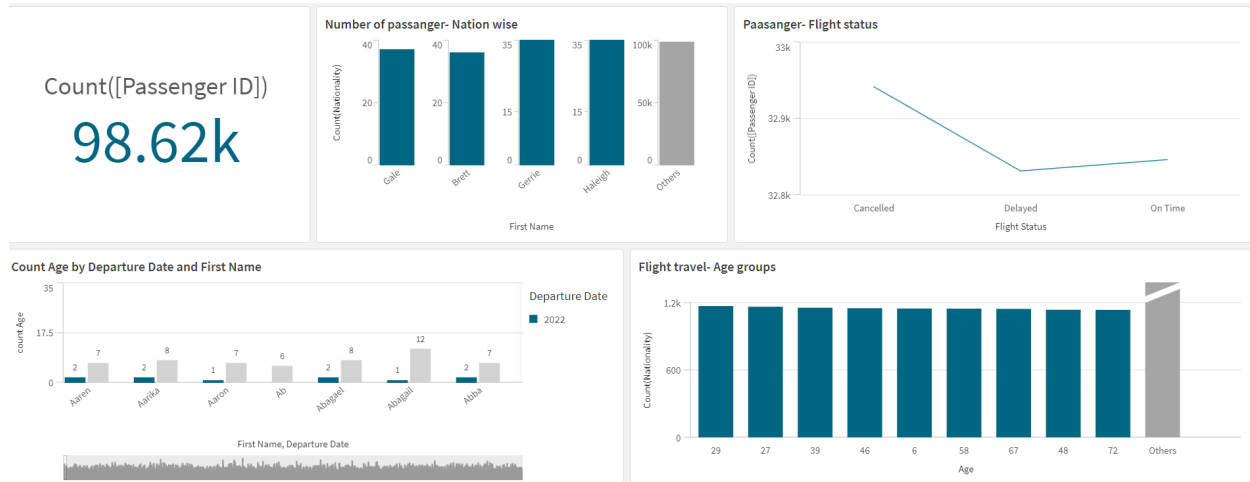
-Dashboard 2:

9. Amount Of Data Loaded

Age

Airport Continent

🌐 Airport Country ...

Airport Name

Arrival Airport

Continents

🌐 Country Name

📅 Departure Date ▶

First Name

Flight Status

Gender

Last Name

🌐 Nationality

Airport Name

Arrival Airport

Continents

🌐 Country Name

📅 Departure Date ▶

First Name

Flight Status

Gender

Last Name

🌐 Nationality

Passenger ID

Pilot Name

10.Data Pre-Processing – Qlik Sense Script

```
Set dataManagerTables = '','Airline Dataset Updated - v2';
//This block renames script tables from non generated section which conflict with the names of managed tables

For each name in $(dataManagerTables)
    Let index = 0;
    Let currentName = name;
    Let tableNumber = TableNumber(name);
    Let matches = 0;
    Do while not IsNull(tableNumber) or (index > 0 and matches > 0)
        index = index + 1;
        currentName = name & '-' & index;
        tableNumber = TableNumber(currentName)
        matches = Match('$(currentName)', $(dataManagerTables));
    Loop
    If index > 0 then
            Rename Table '$(name)' to '$(currentName)';
    EndIf;
Next;
Set dataManagerTables = ;


Unqualify *;
```

```
__countryAliasesBase:
LOAD
    Alias AS [__Country],
    ISO3Code AS [__ISO3Code]
FROM [lib://DataFiles/countryAliases.qvd]
(qvd);

__countryGeoBase:
LOAD
    ISO3Code AS [__ISO3Code],
    ISO2Code AS [__ISO2Code],
    Polygon AS [__Polygon]
FROM [lib://DataFiles/countryGeo.qvd]
(qvd);

__countryName2IsoThree:
MAPPING LOAD
    __Country,
    __ISO3Code
RESIDENT __countryAliasesBase;

__countryCodeIsoThree2Polygon:
MAPPING LOAD
```

```
[Airline Dataset Updated - v2]:
LOAD
    [Passenger ID],
    [First Name],
    [Last Name],
    [Gender],
    [Age],
    [Nationality],
    [Airport Name],
    [Airport Country Code],
    [Country Name],
    [Airport Continent],
    [Continents],
    Date(Date#([Departure Date], 'MM-DD-YYYY') ) AS [Departure Date],
    [Arrival Airport],
    [Pilot Name],
    [Flight Status],
    APPLYMAP( '__countryCodeIsoThree2Polygon', APPLYMAP( '__countryName2IsoThree', LOWER([Nationality])), '-'
    APPLYMAP( '__countryCodeIsoTwo2Polygon', UPPER([Airport Country Code]), '-') AS [Airline Dataset Updated
    APPLYMAP( '__countryCodeIsoThree2Polygon', APPLYMAP( '__countryName2IsoThree', LOWER([Country Name])), '-
 FROM [lib://DataFiles/Airline Dataset Updated - v2.csv]
(txt, utf8, embedded labels, delimiter is ',', msq);
```

```
    Dual(Year($1), YearStart($1)) AS [Year] Tagged ('$axis', '$year'),
    Dual('Q'&Num(Ceil(Num(Month($1))/3)),Num(Ceil(NUM(Month($1))/3),00)) AS [Quarter] Tagged ('$quarter', '$cyc
    Dual(Year($1)&'-Q'&Num(Ceil(Num(Month($1))/3)),QuarterStart($1)) AS [YearQuarter] Tagged ('$yearquarter', '
    Dual('Q'&Num(Ceil(Num(Month($1))/3)),QuarterStart($1)) AS [_YearQuarter] Tagged ('$yearquarter', '$hidden',
    Month($1) AS [Month] Tagged ('$month', '$cyclic'),
    Dual(Year($1)&'-'&Month($1), monthstart($1)) AS [YearMonth] Tagged ('$axis', '$yearmonth', '$qualified'),
    Dual(Month($1), monthstart($1)) AS [_YearMonth] Tagged ('$axis', '$yearmonth', '$simplified', '$hidden'),
    Dual('W'&Num(Week($1),00), Num(Week($1),00)) AS [Week] Tagged ('$weeknumber', '$cyclic'),
    Date(Floor($1)) AS [Date] Tagged ('$axis', '$date', '$qualified'),
    Date(Floor($1), 'D') AS [_Date] Tagged ('$axis', '$date', '$hidden', '$simplified'),
    If (DayNumberOfYear($1) <= DayNumberOfYear(Today()), 1, 0) AS [InYTD] ,
    Year(Today())-Year($1) AS [YearsAgo] ,
    If (DayNumberOfQuarter($1) <= DayNumberOfQuarter(Today()),1,0) AS [InQTD] ,
    4*Year(Today())+Ceil(Month(Today())/3)-4*Year($1)-Ceil(Month($1)/3) AS [QuartersAgo] ,
    Ceil(Month(Today())/3)-Ceil(Month($1)/3) AS [QuarterRelNo] ,
    If(Day($1)<=Day(Today()),1,0) AS [InMTD] ,
    12*Year(Today())+Month(Today())-12*Year($1)-Month($1) AS [MonthsAgo] ,
    Month(Today())-Month($1) AS [MonthRelNo] ,
    If(WeekDay($1)<=WeekDay(Today()),1,0) AS [InWTD] ,
    (WeekStart(Today())-WeekStart($1))/7 AS [WeeksAgo] ,
    Week(Today())-Week($1) AS [WeekRelNo] ;

DERIVE FIELDS FROM FIELDS [Departure Date] USING [autoCalendar] ;
```

11.Utilization Of Filters
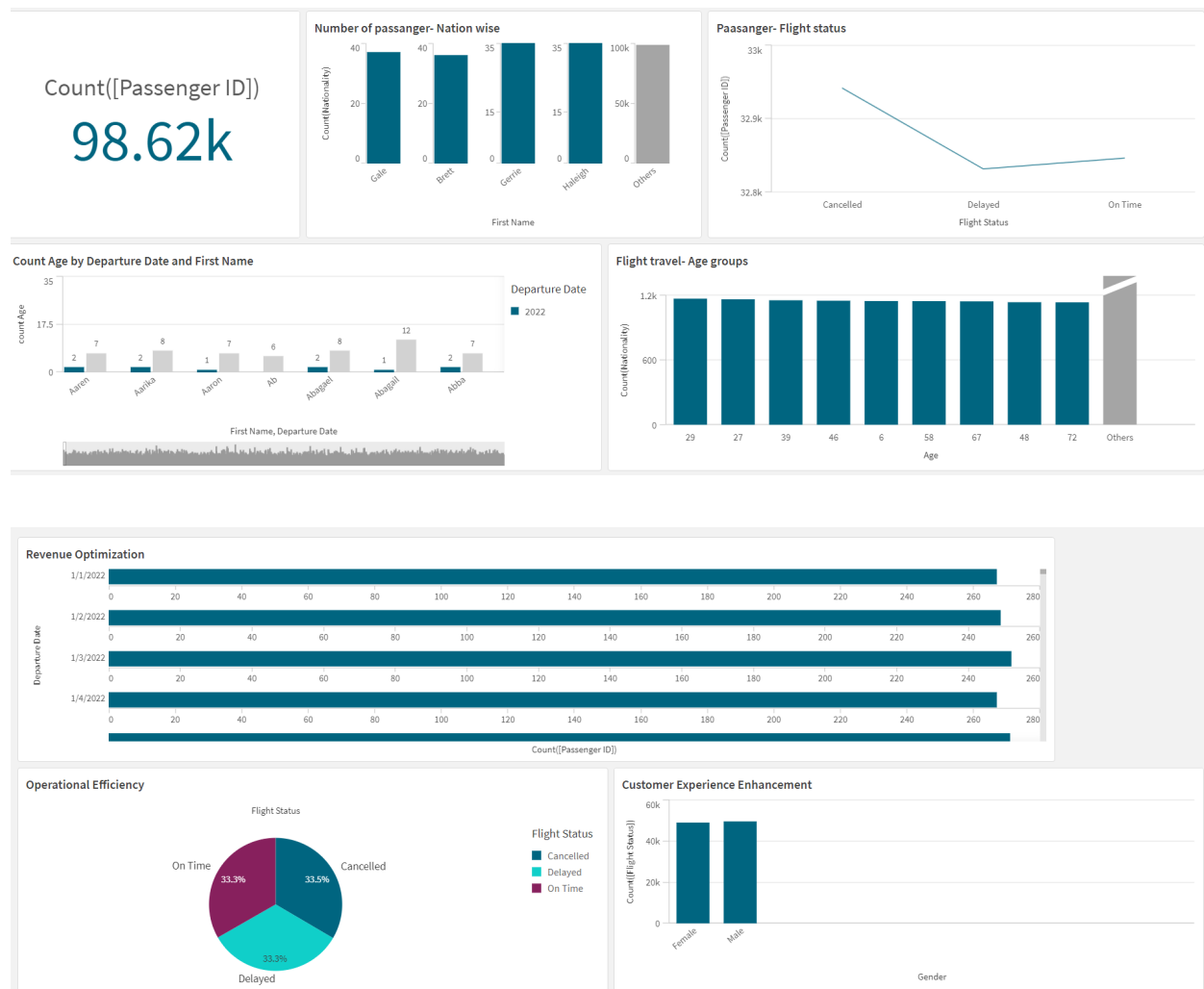
12.No Of Visualizations/ Graphs

- passanger count
- Number of passanger- Nation wise
- Passanger- Flight status
- Count Age by Departure Date and First Name
- Flight travel- Age groups
- Revenue Optimization
- Operational Efficiency
- Customer Experience Enhancement

Conclusion

The project "Exploring Insights from Synthetic Airline Data Analysis with Qlik" demonstrates the transformative power of data analytics in the aviation industry. By leveraging Qlik, I was able to derive meaningful insights from synthetic airline data, supporting revenue optimization, operational efficiency, and customer experience enhancement. This documentation outlines the systematic approach I took to achieve these insights, showcasing the value of data-driven decision-making in the airline industry.