Problem Statement

Exploring Insights From Synthetic Airline Data Analysis With Qlik

INTRODUCTION

This project leverages synthetic airline data to extract valuable insights using Qlik, a business intelligence and data visualization tool. The data simulates various aspects of airline operations, including flight schedules, passenger demographics, ticket sales, and performance metrics. The objective is to use Qlik's capabilities to identify patterns, trends, and correlations within the data, aiding decision-making for airlines, airports, and related stakeholders.

PROJECT FLOW

DATA COLLECTION AND EXTRACTION

- **Collect Dataset from Kaggle**:
  - **Identify Relevant Dataset**: Search for airline-related datasets on Kaggle that include flight schedules, passenger demographics, ticket sales, and performance metrics.
  - **Download Data**: Download the chosen dataset(s) from Kaggle.
- **Extract Data Files**:
  - **Unzip Files**: If the dataset is provided as a compressed file (e.g., ZIP), extract all the files to a designated folder on your PC.
  - **Organize Files**: Ensure all data files are organized and accessible, typically in formats such as CSV, Excel, or JSON.
  - Demonstration video :-Link

DATA PREPARATION

DATA LOADING INTO THE QLIK CLOUD ENVIRONMENT

- **Access Qlik Cloud Platform**:
  - Open your web browser and navigate to the Qlik Cloud platform.
- **Login with Valid Credentials**:

    - Enter your username and password to log in to your Qlik Cloud account.
- **Navigate to Data Load Editor**:
    - Once logged in, you'll be directed to the Qlik Cloud interface.
    - Locate and access the Data Load Editor, which is typically found in the app creation or data management section.
- **Import Dataset**:
    - Look for the option to import dataset files into the Qlik Cloud environment.
    - Click on the import button and browse to select the dataset file(s) that you want to load into Qlik Cloud.

App Creation:

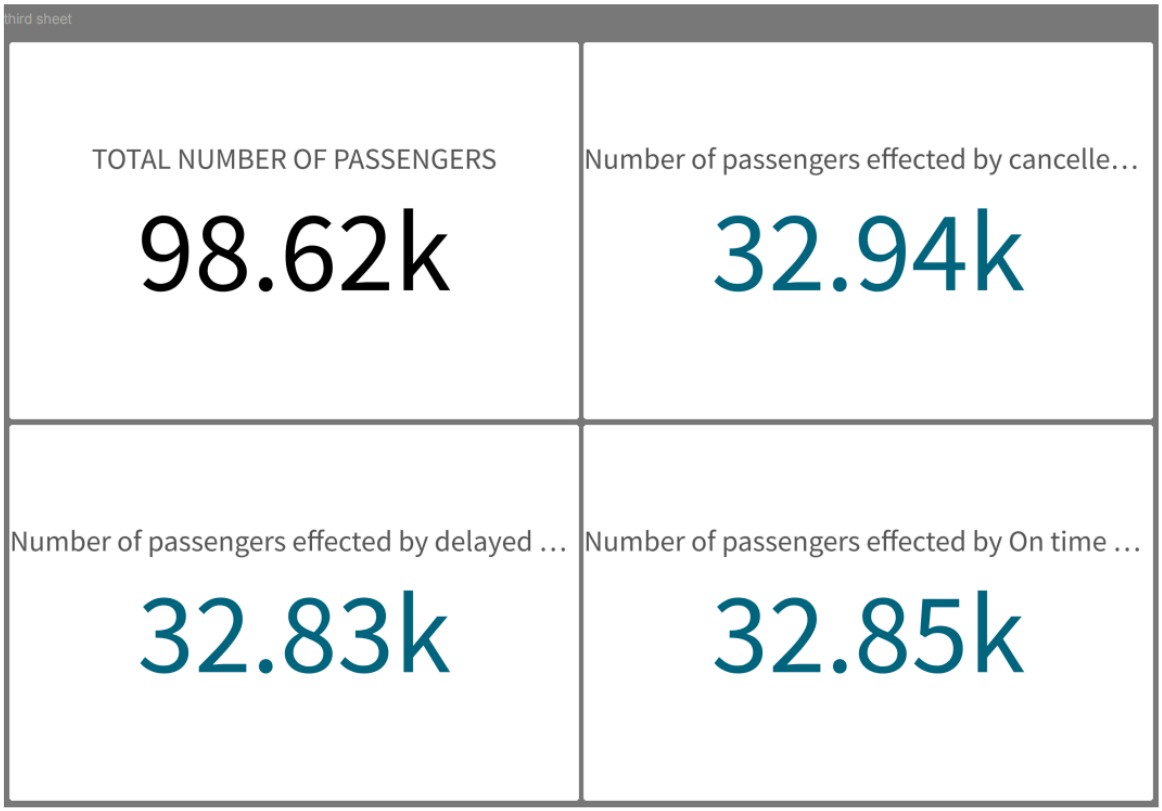
- Upon importing the dataset, Qlik Cloud automatically creates a new app containing the loaded data.
- You can now start analyzing and visualizing the data within the newly created app in the Qlik Cloud environment.

DATA PRE-PROCESSING

- **Data Cleaning**:
    - Handle missing values: Use functions like `IsNull()` or `Null()` to identify and handle missing values by either replacing them or removing them.
    - Remove duplicates: Utilize the `DropDuplicates` function to eliminate duplicate records from the dataset.
    - Standardize data: Convert data into consistent formats (e.g., date formats, currency formats) using Qlik functions like `Date()`, `Num()`, `Money()`.
- **Data Transformation**:
    - Create calculated fields: Use Qlik script to create new fields based on existing data, such as calculating total sales or deriving new metrics.
    - Rename fields: Use the `Rename Fields` statement to rename fields for clarity and consistency.
    - Concatenate tables: Combine multiple tables into a single table using the `Concatenate` function, ensuring consistency in field names and data types.
- **Data Filtering**:
    - Apply filters: Use the `Where` clause to filter data based on specific conditions, such as date ranges or categorical variables.
    - Remove outliers: Use functions like `RangeSum()` or `Above()` and `Below()` to identify and filter out outliers from the dataset.
- **Data Integration**:
    - Merge tables: Combine multiple tables by common fields using Qlik's data model viewer or the `Join` statement.
    - Establish relationships: Create associations between tables based on common fields to enable seamless data exploration and analysis.
- **Data Validation**:
    - Verify data integrity: Check for data consistency and accuracy after

preprocessing steps by reviewing sample records and running validation tests.

- Debug and troubleshoot: Identify and resolve any issues or errors encountered during the preprocessing phase using Qlik's debugging tools and error logs.



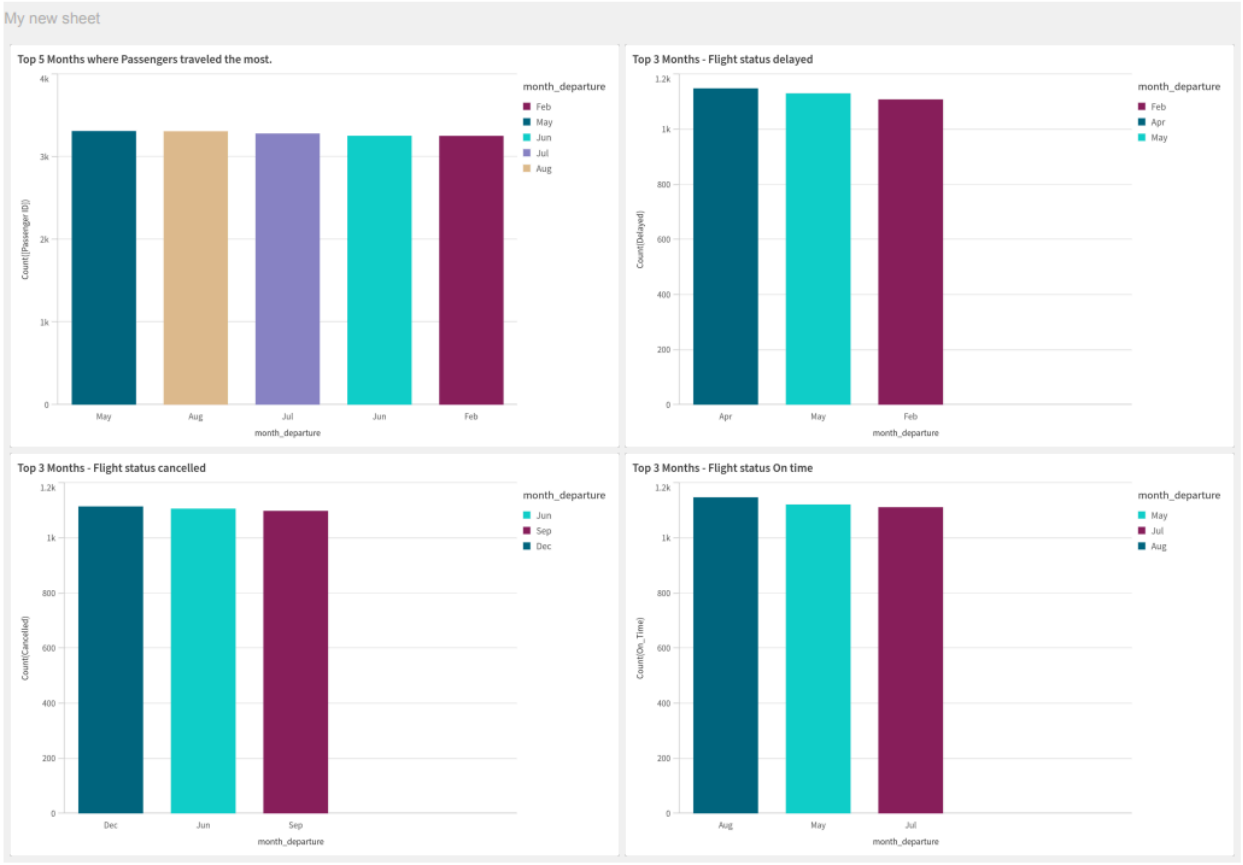
BarChart

bar chart is a visual representation of data that uses rectangular bars to display the frequency, count, or value of different categories. It's a fundamental visualization tool that helps users analyze and compare data across different dimensions or categories.

- **Drag and Drop the Bar Chart Object:**
  - From the visualization panel, drag the bar chart object onto the sheet.
- **Configure the Bar Chart:**
  - In the properties panel, configure the bar chart by selecting dimensions and measures:
    - **Dimension (X-Axis):** Choose the field that represents the categories or groups you want

to compare.

Measure (Y-Axis): Select the field that represents the values you want to measure.

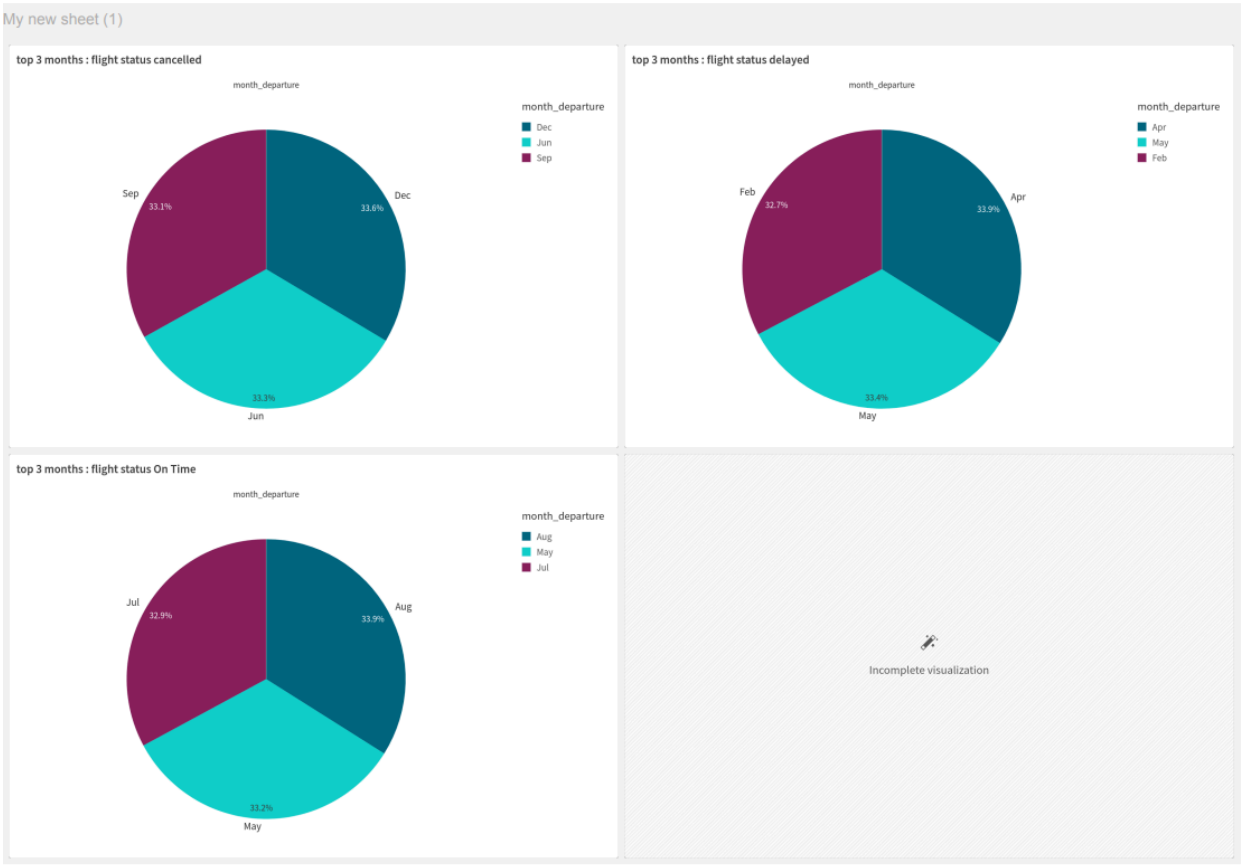


Pie chart

a pie chart is a visualization that displays data as a circular chart divided into slices to illustrate numerical proportions. Each slice represents a proportion of the whole, and the size of each slice is proportional to the value it represents relative to the total.

- **Drag and Drop the Pie Chart Object:**
  - From the visualization panel, drag the pie chart object onto the sheet.
- **Configure the Pie Chart:**

- In the properties panel, configure the pie chart by selecting dimensions and measures:
    - **Dimension:** Choose the field that represents the categories or groups you want to visualize.
    - **Measure:** Select the field that represents the values you want to measure or analyze.
    - You can also add additional dimensions or measures for further analysis, such as grouping by time periods or adding calculated expressions.

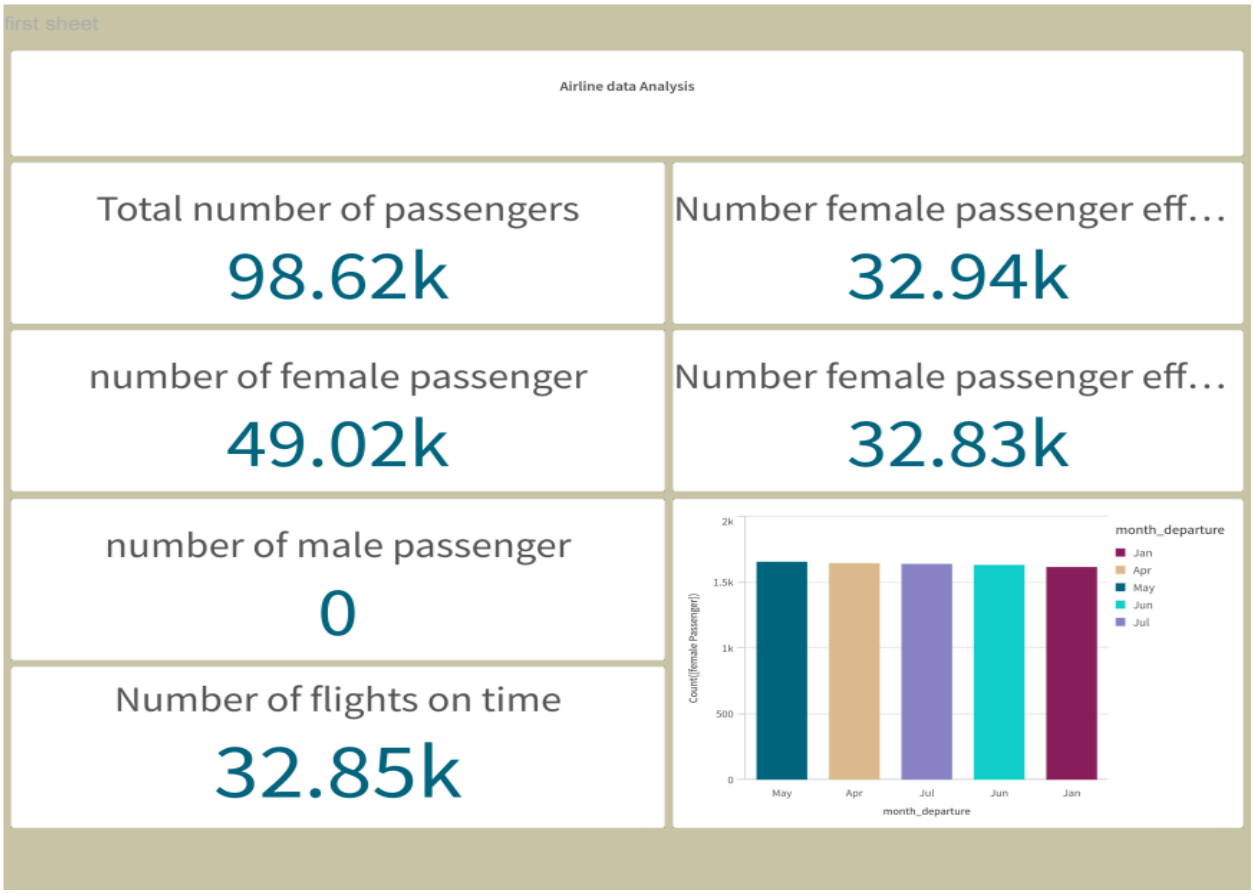


Dashboard Creation

Steps to Create a Dashboard in Qlik

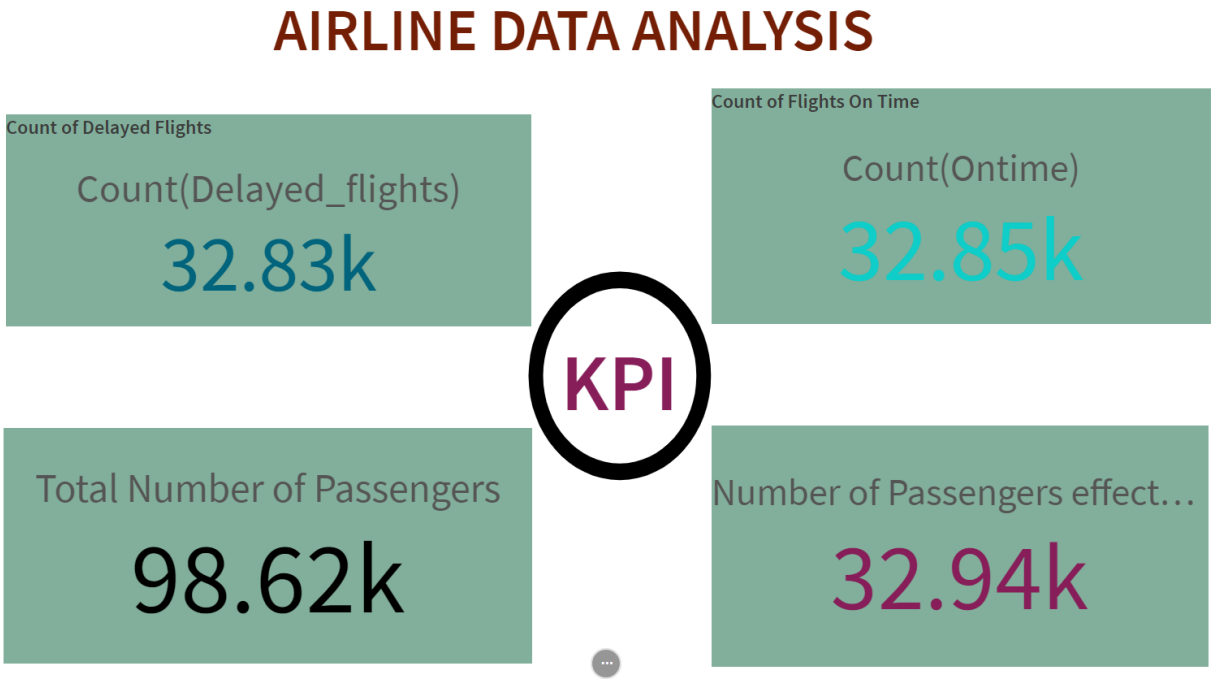
1. **Define Objectives**:
    - Identify the purpose and key performance indicators (KPIs) that the dashboard should track.
1. **Prepare Data**:
    - Connect to relevant data sources and clean the data to ensure accuracy and consistency.
    - Create a logical data model to define relationships between different data tables.

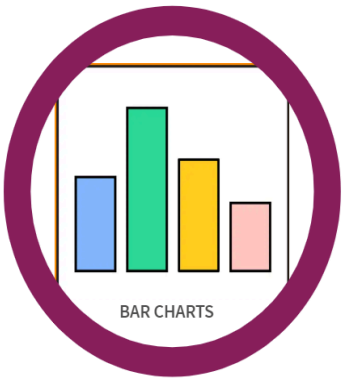
1. **Set Up the Qlik Environment**:
    - Create a new application in Qlik Sense or QlikView.
    - Load data into the application using the Data Load Editor.
1. **Design the Dashboard**:
    - Plan the layout, deciding on the placement of charts, tables, and filters.
    - Add visualizations such as bar charts, line charts, pie charts, and tables to represent data.
    - Include interactive filters to allow users to explore the data dynamically.
1. **Enhance User Experience**:
    - Apply custom styling and themes to make the dashboard visually appealing.
    - Add tooltips and labels for additional context.
    - Ensure the dashboard is responsive and works well on various devices.
1. **Add Interactivity**:
    - Enable drill-down capabilities for detailed data exploration.
    - Use Qlik's scripting language for dynamic calculations and expressions.

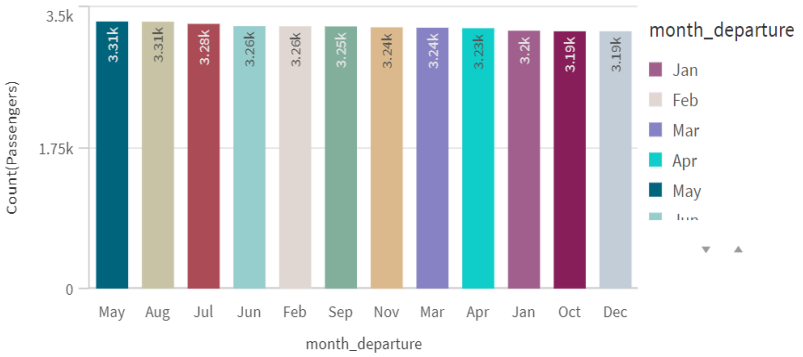


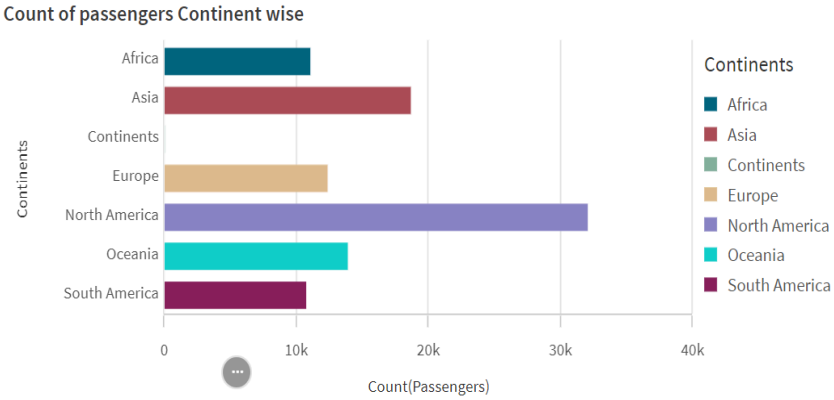
Story Creation

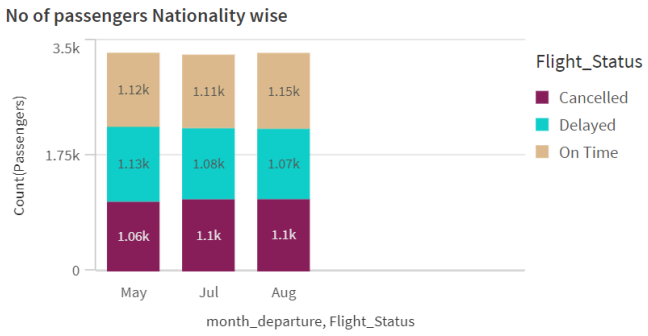
A data story is a way of presenting data and analysis in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way, and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.

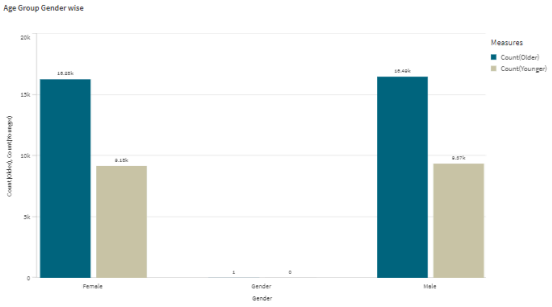


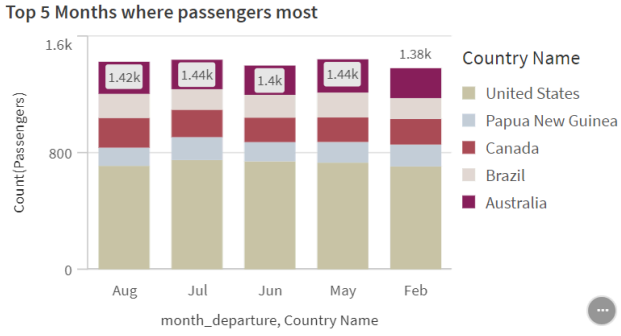


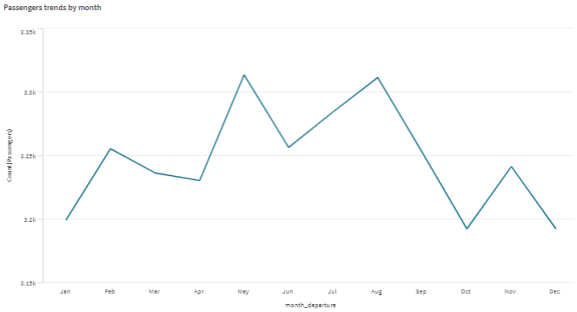




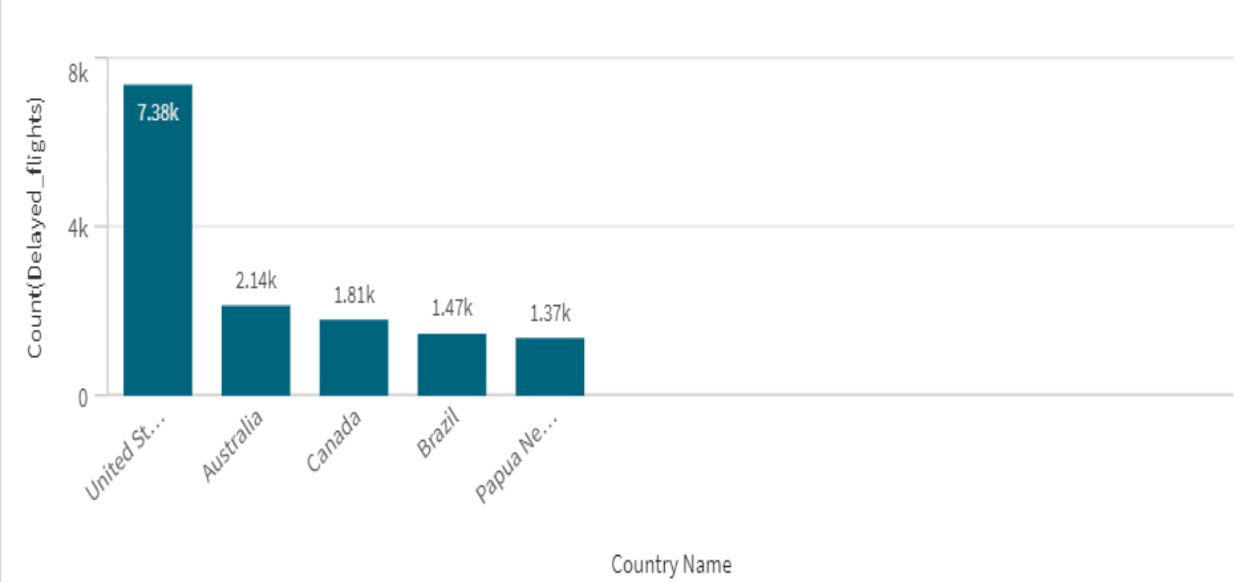








Conclusions and Insights

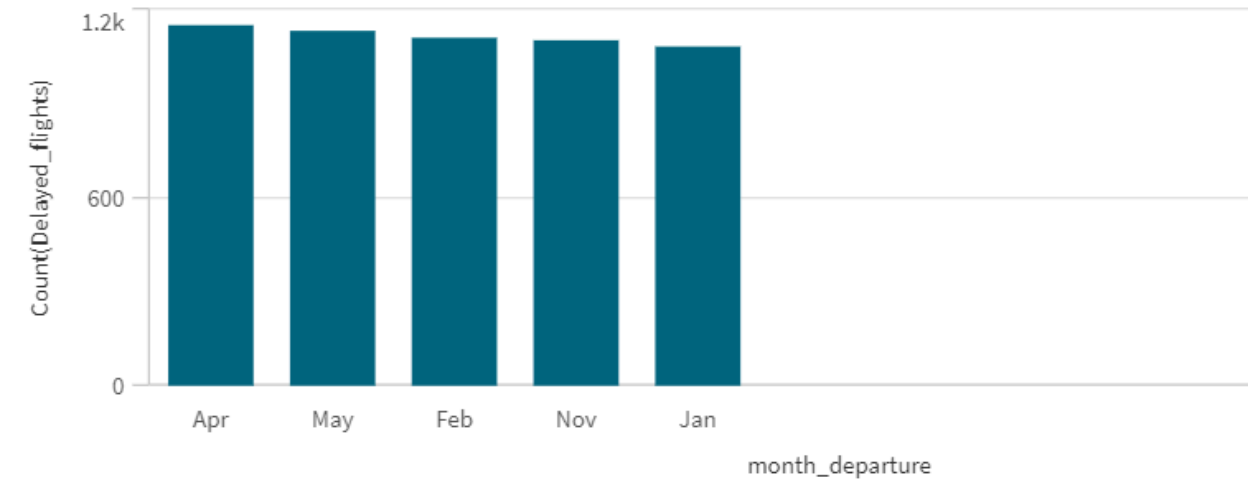


Reasons for the flight delays

Weather-Related Delays:

- Extreme weather conditions (such as storms, snow, or fog) can disrupt flight schedules.

Airports may close runways or limit operations during adverse weather



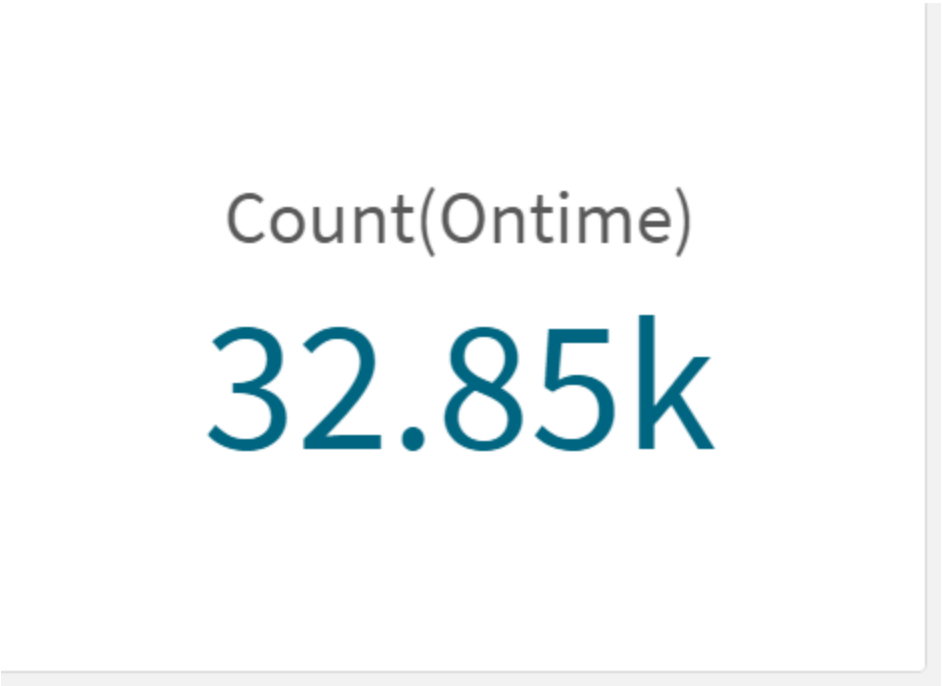
Reasons for the delay of flights in the top five months

- **February**: In the USA, February varies significantly by region. For example:
  - **Northern States**: Cold temperatures with snow in places like New

York and Chicago.

- **Southern States**: Milder weather in Florida and Texas.

- **April**: Springtime! Expect milder temperatures across most states, with blossoming flowers and longer daylight hours.
- **November**: Fall season. Leaves change color in the Northeast, while Southern states remain relatively warm.
- **January**: Winter in full swing, especially in the Northern states. Snow and cold temperatures prevail.



Reasons

Passengers:

- **Peace of mind:** Knowing a flight is on time reduces pre-flight stress and allows passengers to relax and plan accordingly. This is especially helpful for connecting flights or those with tight schedules.
- **Better time management:** Passengers can plan their arrival at the airport accordingly, avoiding unnecessary layovers or missed connections.
- **Proactive adjustments:** If a flight is delayed, passengers can be informed and make changes to their travel plans, like rebooking connecting flights or arranging transportation.

Airline Management:

- **Improved efficiency:** On-time flights mean smoother operations for the airline. This allows for better gate and crew utilization, reducing delays for subsequent flights.
- **Reduced costs:** On-time departures and arrivals help airlines avoid costs associated with delays, such as additional staff time, compensation for stranded passengers, and rescheduling flights.
- **Better customer satisfaction:** On-time flights lead to happier passengers, which can improve customer loyalty and brand reputation.