Airlines Data Analytics for Aviation Industry

Smart bridge Data Analytics Final Project by **Team – 405**

TEAM MEMBERS:

STUDENT NAME	REG NUMBER	E-MAIL	CAMPUS
N ROHIT	20BCD7097	rohit.20bcd7097@vitap.ac.in	VIT-AP
V.AASRITHA SAI	20BCE7074	aasritha.20bce7074@vitap.ac.in	VIT-AP
MRK SAKETH	20BCD7091	saketh.20bcd7091@vitap.ac.in	VIT-AP
SHIVAM KUMAR	20BBS0040	shivam.kumar2020b@vitstudent.ac.in	VIT-Vellore

Contents

Introduction

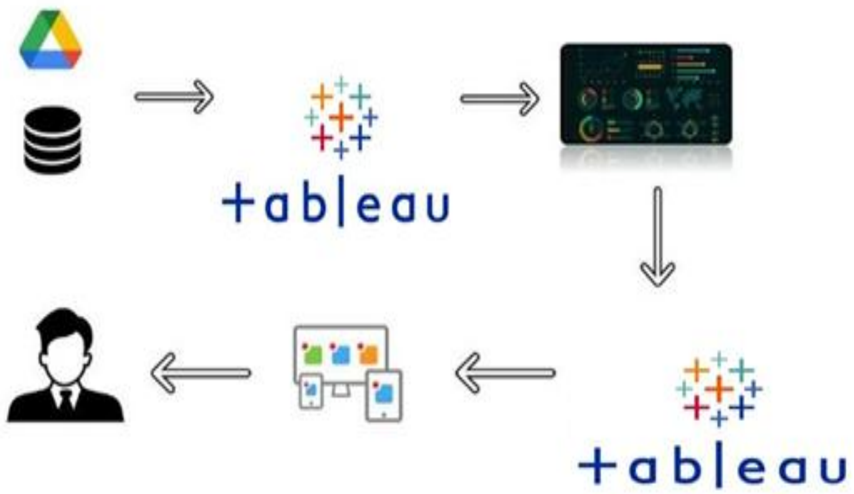
The aviation industry heavily relies on efficient airline and airport services to provide seamless travel experiences for passengers. Delays in air travel can lead to inconvenience, increased costs, and customer dissatisfaction. This project aims to utilize data analytics techniques to analyze airline data and improve services, specifically focusing on avoiding delays and enhancing the overall travel experience at the municipality level. The airport codes may refer to either IATA airport code, a three-letter code which is used in passenger reservation, ticketing and baggage-handling systems, or the ICAO airport code which is a four-letter code used by ATC systems and for airports that do not have an IATA airport code.

Purpose and Objectives

The purpose of this project is to provide airports, airlines, and the travelling public with a comprehensive analysis of airline data to identify patterns and factors influencing delays. The objectives of the project are as follows:

- Analyze airline data to understand the causes of delays at the municipality level.
- Provide insights into the performance of airlines in terms of on-time arrivals and departures.
- Identify potential areas for improvement in airline and airport services.
- Offer recommendations for optimizing operations and reducing delays in air travel.

Architecture:



Project Flow:

- Users create multiple analytical graphs/charts/Visualizations.
- Using the Analytical Visualizations, build required Dashboard(s).
- Saving and visualizing the final dashboard.

DATA VISUALIZATION:

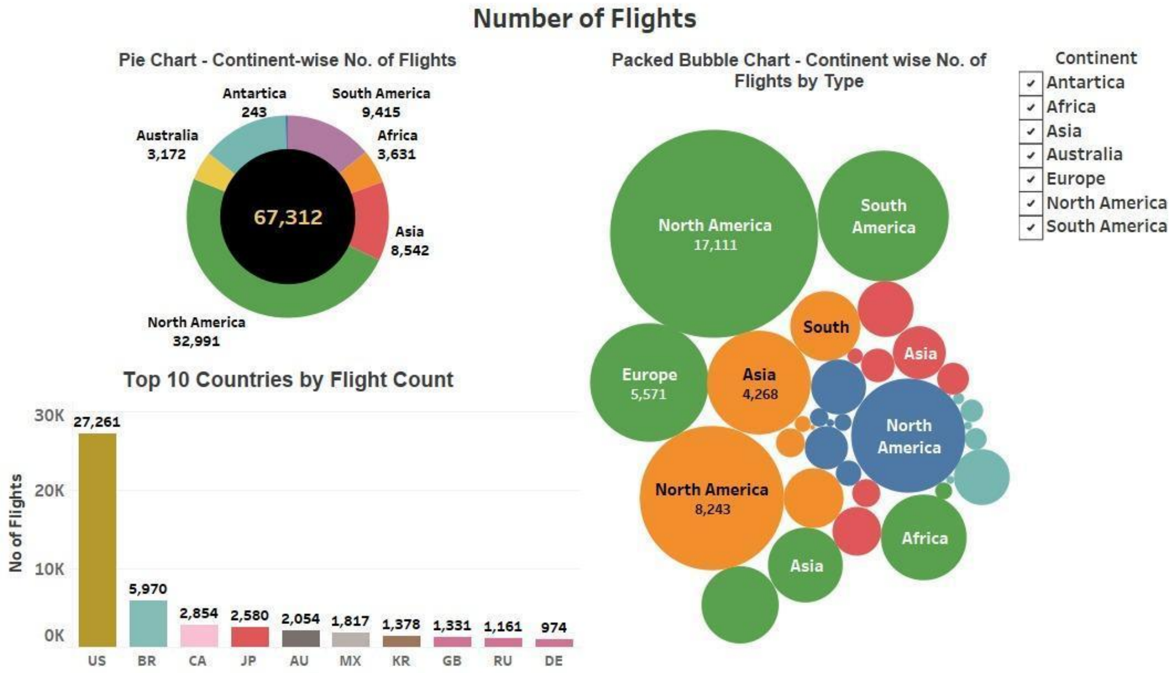
To accomplish this, we have to complete all the activities and tasks listed below:

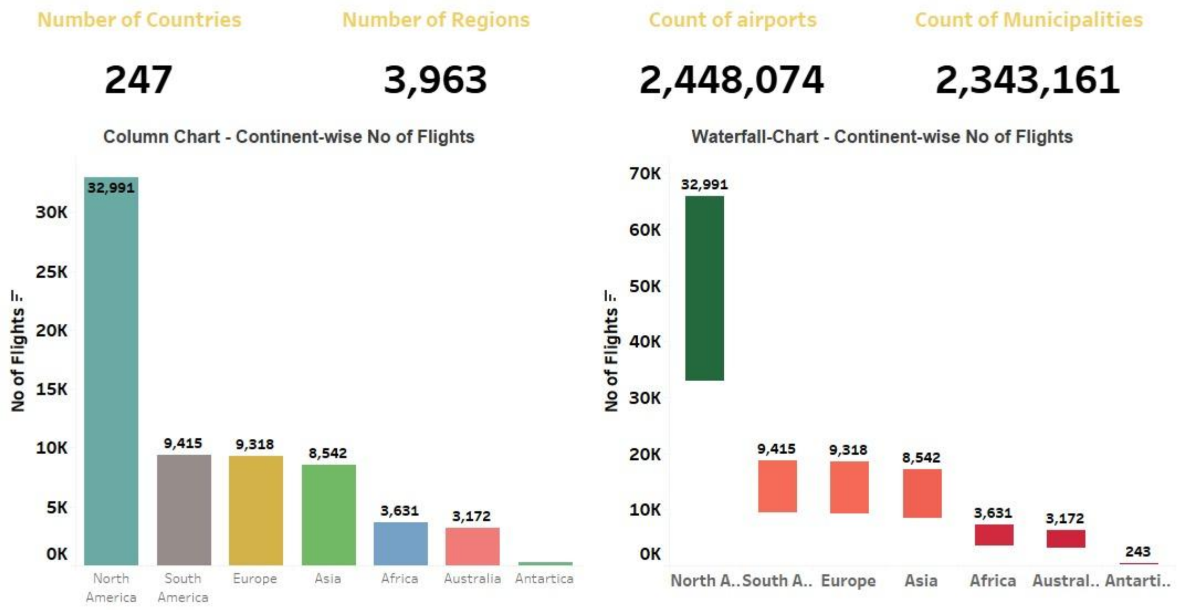
- IBM Cloud Account
- Login to Congas Analytics
- Working with the Dataset
    - Understanding the Dataset
    - Loading the Dataset
- Data Visualization Charts
    - Build the following visualizations

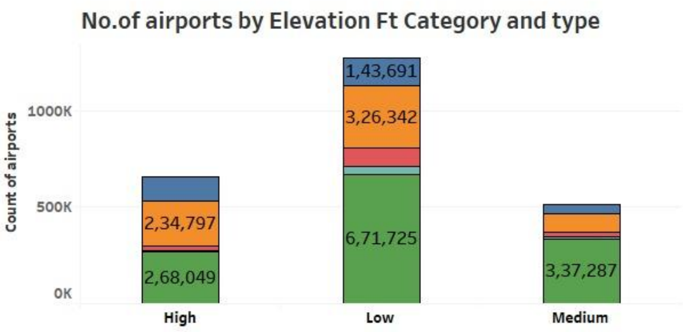
1. Pie Chart - Continent-wise No. of Flights
2. Packed Bubble Chart - Continent wise No. of Flights by Type - Colored with Type
3. Continent List - Filter
4. Top 10 Countries by Flights
5. Countries - Summary Card
6. Regions - Summary Card
7. Airports - Summary Card
8. Municipalities - Summary Card
9. Column Chart - Continent-wise No of Flights
10. Waterfall-Chart - Continent-wise No of Flights
11. Geo-Map - Continent-wise No. of flights
12. Geo-Map - Country-wise No. of flights
13. Continent Filter
14. Flight-Type filter
15. Column-Chart - No of Airports by Type
16. Hierarchy Bubble Chart - Region-wise Different Types of Airports
17. Packed bubble Chart - Municipality-wise No. of Airports
18. Bar Chart - Continent-wise No of Airports

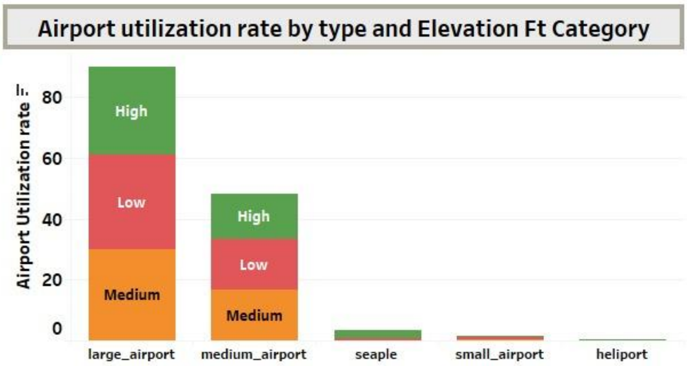
Visualizations:

Below dashboard we can see Pie Chart - Continent-wise No. of Flights, Packed Bubble Chart - Continent wise, No. of Flights by Type - Colored with Type, Continent List - Filter and Top 10 Countries by Flights



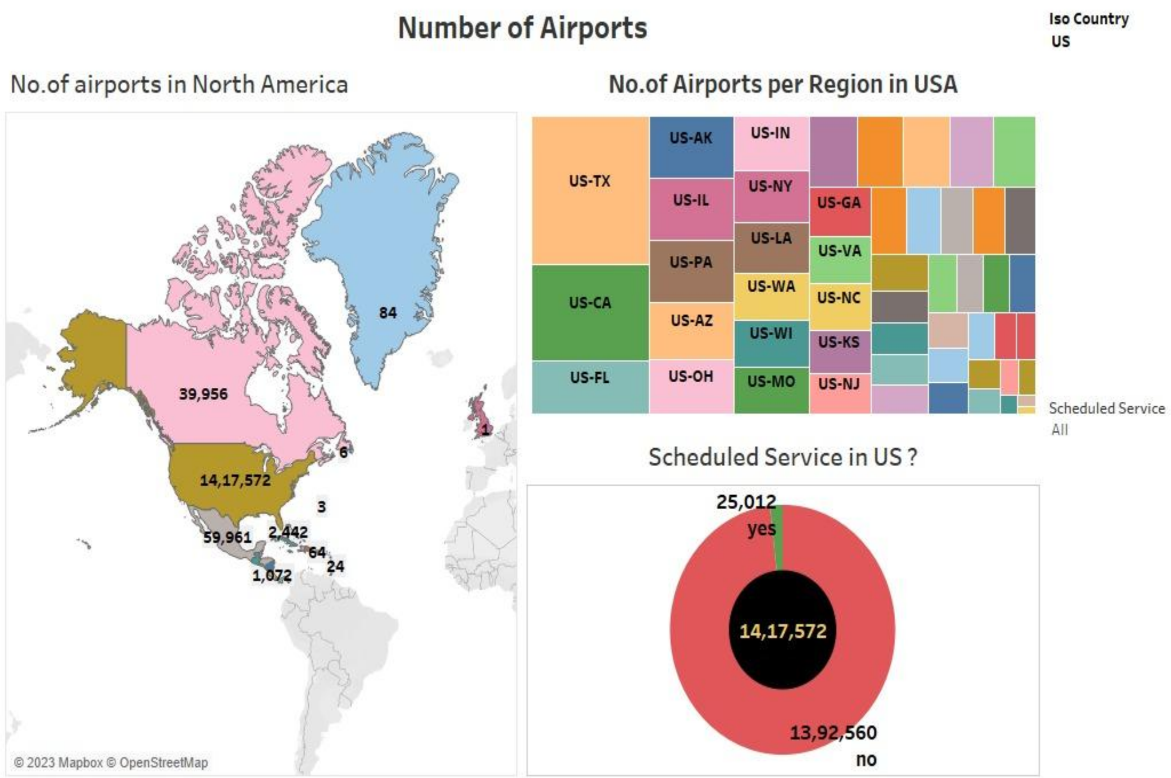




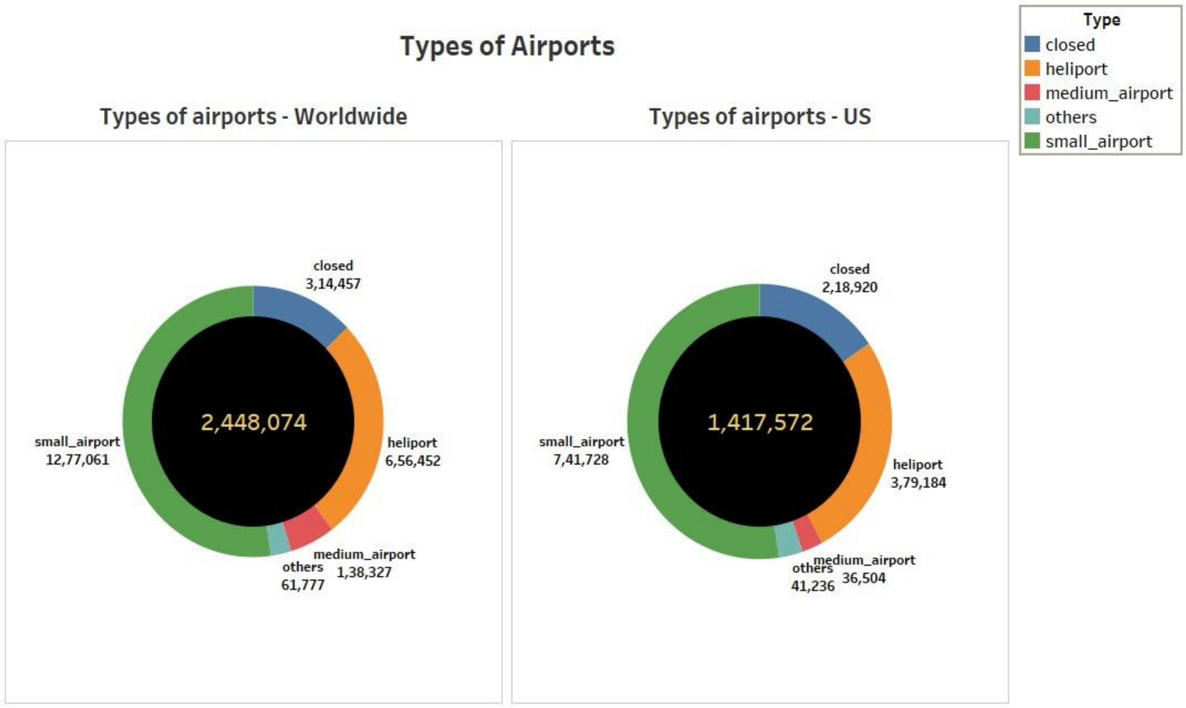


Other Visualizations:

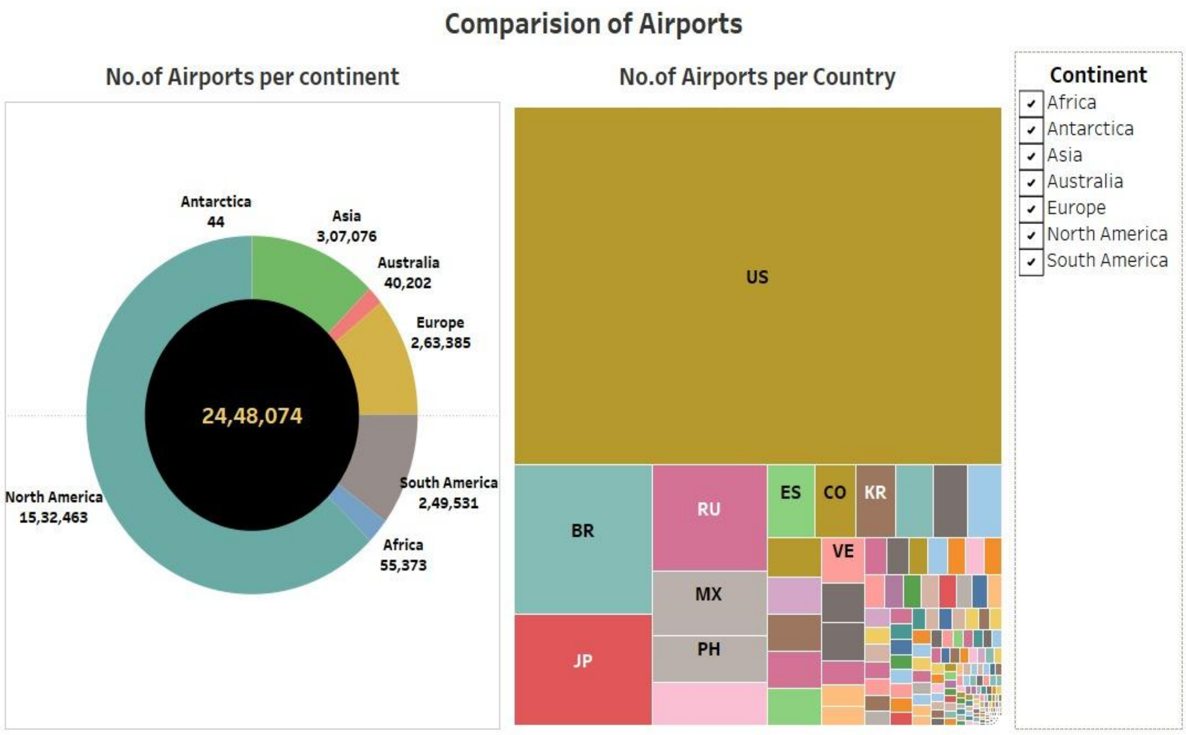
Visualization of No. of airports in North America and No. of Airports per Region in USA



Comparison of Airports in North America and in USA



Comparison of Airports per continent and per Country



Literature Survey

The aviation industry is a complex and dynamic sector that heavily relies on efficient airline operations and airport services to ensure a seamless travel experience for passengers. The use of data analytics in the aviation industry has gained significant attention in recent years, as it enables airlines and airports to analyze vast amounts of data to improve operational efficiency, optimize scheduling, and mitigate delays. This literature survey aims to explore the existing body of research on airlines data analytics in the aviation industry, focusing on its applications, challenges, and benefits.

Data Analytics Applications in the Aviation Industry:

Numerous studies have highlighted the applications of data analytics in the aviation industry. One key area is predicting and managing flight delays. Researchers have developed predictive models that utilize historical data, weather conditions, air traffic information, and other relevant factors to forecast potential delays. By leveraging data analytics techniques such as machine learning and regression analysis, airlines and airports can proactively manage their operations and allocate resources effectively.

Furthermore, data analytics has proven valuable in enhancing customer satisfaction. By analyzing customer feedback, sentiment analysis, and social media data, airlines can identify trends, preferences, and areas for improvement. This information can then be utilized to personalize services, improve in-flight experiences, and tailor customer interactions.

Challenges in Airlines Data Analytics:

While data analytics presents numerous opportunities, it also poses several challenges in the aviation industry. One significant challenge is the complexity and variety of data sources. Airlines collect data from various systems, including reservation systems, operational databases, and customer feedback platforms. Integrating and consolidating these diverse datasets require robust data management and integration techniques.

Another challenge lies in data quality and accuracy. Incomplete or erroneous data can hinder the effectiveness of data analytics models and lead to incorrect insights. Ensuring data accuracy and consistency through data cleansing, validation, and normalization processes is crucial to obtain reliable results.

Benefits and Impacts of Airlines Data Analytics:

The use of data analytics in the aviation industry has shown several benefits and impacts. First and foremost, it enables airlines and airports to optimize their operations and reduce delays. By identifying patterns and causal factors behind delays, airlines can implement proactive measures to mitigate their impact. This, in turn, leads to increased operational efficiency, improved on-time performance, and enhanced customer satisfaction.

Data analytics also facilitates evidence-based decision-making. Airlines can utilize insights gained from data analysis to make informed decisions regarding fleet management, route optimization, pricing strategies, and resource allocation. Additionally, data-driven decision-making enhances cost-effectiveness and enables airlines to respond more effectively to market changes and customer demands.

Methodology

Data Collection:

Data will be collected from various sources, including airline databases, airport records, flight schedules, and historical on-time performance data. The data will encompass variables such as flight details (e.g., flight numbers, origin, and destination), scheduled departure and arrival times, actual departure and arrival times, and delays.

Data Pre-processing:

The collected data will undergo pre-processing to ensure its quality and consistency. This involves cleaning the data, handling missing values, and standardizing formats. The airport codes will be converted to a unified format, considering both IATA airport codes and ICAO airport codes.

Exploratory Data Analysis (EDA):

EDA techniques will be applied to explore the dataset and identify patterns, trends, and correlations related to delays. Descriptive statistics, data visualization, and correlation analysis will be conducted to gain insights into the relationships between variables and their impact on delays.

Performance Evaluation Metrics:

Performance evaluation metrics will be defined to assess airlines' on-time performance. These metrics may include on-time departure rate, on-time arrival rate, average delay duration, and overall airline performance scores. The metrics will be calculated based on the municipality level analysis.

Data Analysis and Findings

Overview of the Airline Industry Dataset:

The airline industry dataset will include information on flight details, scheduled and actual departure/arrival times, delays, and other relevant variables. The dataset will cover flights across different municipalities and airports, enabling a comprehensive analysis of delays at the municipality level.

Descriptive Statistics:

Descriptive statistics will be computed to summarize the dataset, including measures such as mean, median, standard deviation, and percentiles. These statistics will provide insights into the average delay durations, distributions of delays, and overall performance of airlines in terms of on-time departures and arrivals.

Data Visualization:

Data visualization techniques, including line charts, bar graphs, and heat maps, will be used to visually represent patterns and trends related to delays. Visualizations will help identify peak delay periods, compare performance across airlines, and highlight potential areas for improvement.

Correlation Analysis:

Correlation analysis will be conducted to identify relationships between variables and their impact on delays. Correlation coefficients will be calculated, and visual representations such as scatter plots or heat maps will be used to illustrate significant correlations between factors like weather conditions, airline performance, and delay durations.

Key Insights and Interpretation:

Factors Affecting Delay Occurrences:

The analysis will provide insights into the factors that contribute to delays at the municipality level. This may include weather conditions, airport congestion, air traffic control issues, or operational inefficiencies. Understanding these factors will help identify areas for improvement and potential solutions to reduce delays.

Performance Comparison across Airlines:

By evaluating the on-time performance metrics, a comparison of airlines' performance in terms of delays will be made. This will identify airlines that consistently deliver on-time services and those that need improvement. Airlines with exemplary performance can serve as benchmarks for others to enhance their operations.

Recommendations:

Based on the findings, recommendations will be provided to airports and airlines to optimize operations and reduce delays. These recommendations may include improving communication and coordination between airlines and airports, implementing efficient scheduling strategies, enhancing maintenance and operational procedures, and leveraging real-time data for proactive decision-making.

Limitations and Challenges:

Data Limitations:

The analysis may face challenges related to the availability and quality of data. Data inconsistencies, missing values, and biases may impact the accuracy and reliability of the findings. Efforts will be made to address these limitations and ensure data integrity.

External Factors:

The analysis may be influenced by external factors that are beyond the control of airlines and airports, such as extreme weather events, air traffic control disruptions, or security concerns. These factors should be considered when interpreting the results and making recommendations.

TABLEAU PUBLIC:

Dashboards link:

https://public.tableau.com/views/Dashboard_28_16879724216690/1?:language=en-US&:display_count=n&:origin=viz_share_link

Story link:
https://public.tableau.com/app/profile/rohit.n5087/viz/Story_16881365226800/Story1?publish=yes

Source Code:

Project_Files/Web Integration/templates/index.html

SPSGP-524797-Airlines-Data-Analytics-for-Avaition-Industry/Project_Files/Web Integration/templates/index.html at main · smartinternz02/SPSGP-524797-Airlines-Data-Analytics-for-Avaition-Industry · GitHub

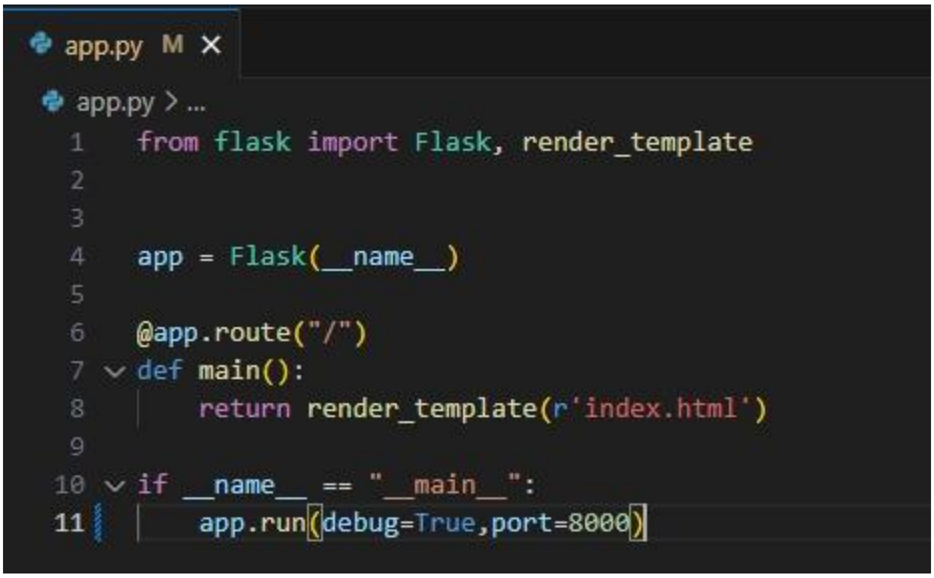
Basic HTML SYNTAX:

```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Document</title>
</head>
<body>

</body>
</html>
```

Embedded Code:

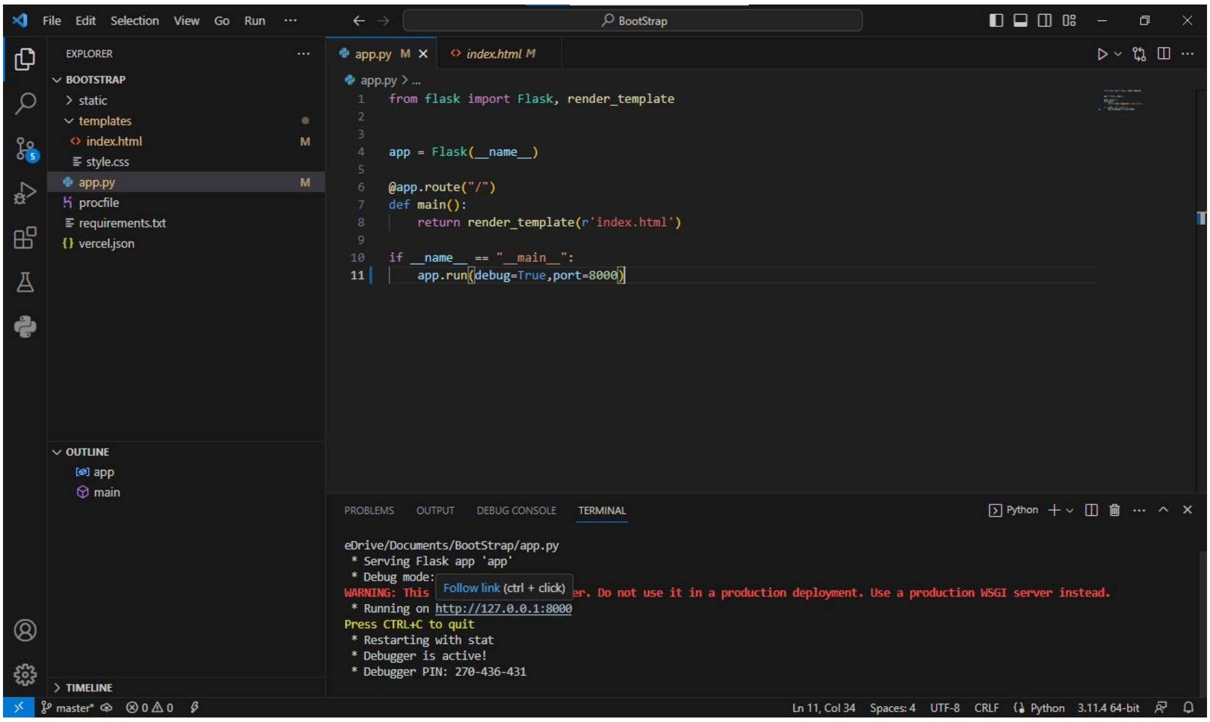
FLASK: (App.py)



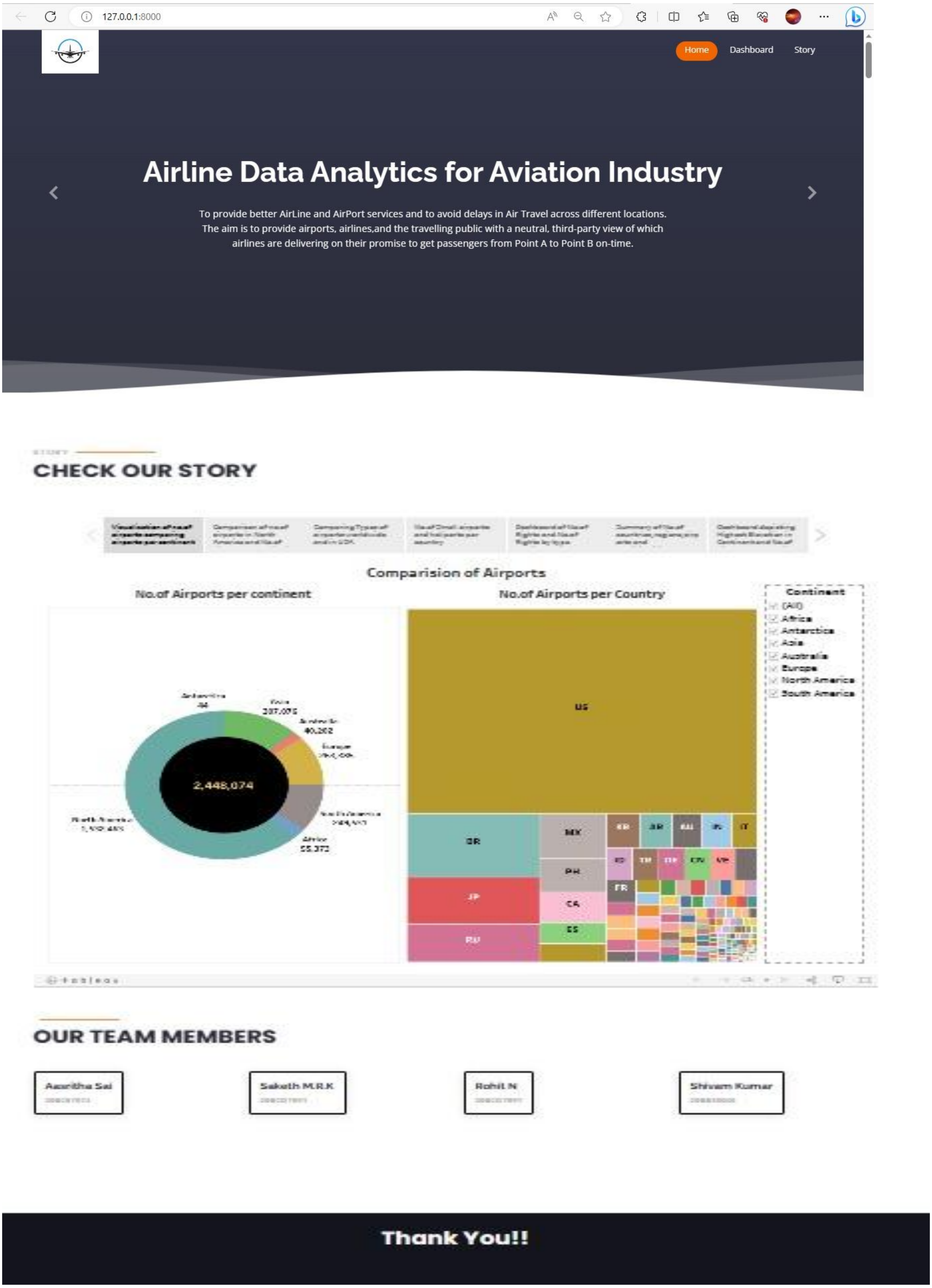
Website link :

https://github.com/smartinternz02/SPSGP-524797-Airlines-Data-Analytics-for-Avaition-Industry.git

Download this file run python file app.py then the web link of our project html will open.



Web page:


127.0.0.1:8000
Home
Dashboard
Story
Airline Data Analytics for Aviation Industry
To provide better AirLine and AirPort services and to avoid delays in Air Travel across different locations.
The aim is to provide airports, airlines,and the travelling public with a neutral, third-party view of which
airlines are delivering on their promise to get passengers from Point A to Point B on-time.
CHECK OUR STORY
Comparision of Airports
No.of Airports per continent
No.of Airports per Country
2,448,074
US
BR
JP
RU
MX
PH
CA
ES
FR
Continent
(All)
Africa
Antarctica
Asia
Australia
Europe
North America
South America
OUR TEAM MEMBERS
Aasritha Sai
Saketh M.R.K
Rohit N
Shivam Kumar
Thank You!!

Conclusion:

Summary of Project Objectives and Findings:

This project aimed to analyze airline data to avoid delays in air travel and enhance airline and airport services at the municipality level. The analysis provided insights into the causes of delays, performance comparisons across airlines, and recommendations for improvement.

Significance for the Aviation Industry:

The findings of this project are significant for airports, airlines, and the travelling public as they provide actionable insights to optimize operations and enhance the travel experience. By addressing delays and improving on-time performance, the project contributes to the overall efficiency and customer satisfaction in the aviation industry.

Future Directions and Further Research:

Future research can explore the integration of real-time data sources, such as live flight data and weather information, to enhance the accuracy of delay predictions and proactive measures. Additionally, the analysis can be expanded to cover other factors impacting the travel experience, such as baggage handling, customer service, and security procedures.