

**K.L.N College of Information Technology, Pottapalayam**

**Department of (computer science and engineering)**

**HX 8001- Professional Readiness for Innovation, Employability and Entrepreneurship**

**“Project Report”**

**“Airlines Data Analytics for Aviation Industry ”**

**Team ID: PNT2022TMID52743**

BALA SUNDER S	910719104004
HARI KRISHNAN R B	910719104011
MOHAMED NEFEES A	910719104017
RUBIN RAJ A	910719104022

# **INTRODUCTION:**

## **Project Overview :**

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

## **Purpose :**

To provide better Airline and AirPort services and to avoid delays in Air Travel across different locations at Municipality level. 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.

# **LITERATURE SURVEY :**

## **Existing problem :**

The airport codes may refer to either the IATA airport code, a three-letter code that 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.

## **References :**

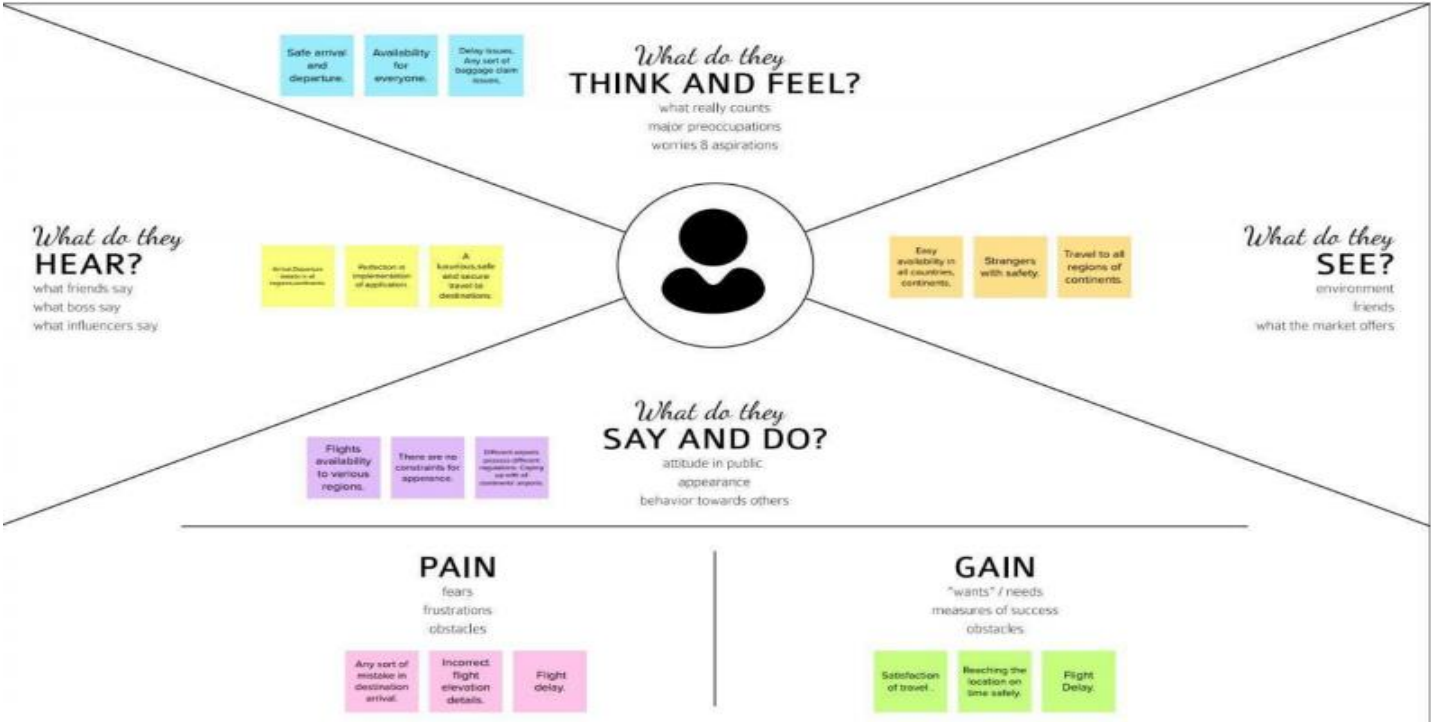
- 1.Data Science And Analytics In Aviation(2020): Authors:Sai-Ho-Chung,Hoi-Lam-ma
- 2.Data Analytics for Air Travel Data(2021): Authors:Haiman Tian,Yudong Tao
- 3.Topological Data Analysis For Aviation Applications(2018): Authors: Max Z. Li,Megan S. Ryerson and Hamsa Balakrishnan
- 4.Operational Efficiency Versus Financial Mobility In The Global Airline Industry(2015): Author:Hoi-Lam-ma
- 5.An Evaluation Of The Operational Performance And Profitability Of The U.S. Airlines(2021): Author:Emillio Collar

## **Problem Statement Definition :**

To identify and manage many people traveling this summer, they are noticing first –hand that airlines are facing major challenges, including numerous flight cancellations and delays.

# IDEATION & PROPOSED SOLUTION :

## Empathy Map Canvas:



# Ideation & Brainstorming :

## Bala Sunder

Airports need best visualisation to prevent delays in the air travel

Airports need to prepare organisational data on more than 8700 companies including airports and MRO's.

Air taxi services helps to accomodate between 4 to 18 people,used particularly by business travellers.

AI is the optimal solution for flight optimization kind of problem.

## Rubin Raj

Passengers need the shortest airline route for cost efficiency

MNC's need real time KPI's and visualisations for maintaining their trade.

Airports can prepared for advanced air mobility.

Telegram,flight radar 24 is useful app for airline or aviation news.

## Hari Krishnan

It enables the stockmarket buyers to invest suitable airline stocks.

To build relationships with customers that go beyond frequent flyer programmes

The project should give real time notification for users about changing dates.

To develop preliminary guidelines and standard for aircraft alerting system.

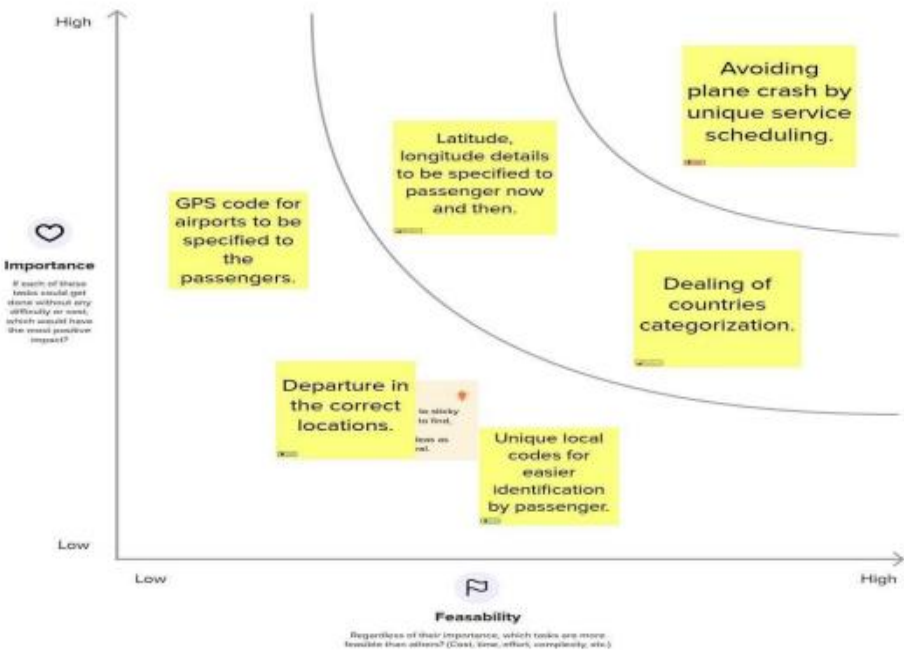
## Mohamed Nafees

It helps the travel and tourist agencies to create mobile apps for services

Convenience to buy is nothing but how and where does the customer wants to purchase.

Passengers need to access latest realtime news and insights in aviation frms.

On time flight environment,snacks and more legroom might be obvious contributors to a good experience and more loyalty



**Proposed Solution :**

S.no	Parameter	Description
1.	Problem Statement(problem to be solved)	The airport codes may refer to either the IATA airport code, a threeletter code that is used in passenger reservation, ticketing and baggagehandling 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.
2.	Idea/Solution Description	Machine learning and analytics have touched almost all the fields around the globe including the aviation industry. With the growth of data, the use of analytics in the airline industry is the next big wave. The purpose of data analytics in aviation is to examine the vast amount of data generated daily and provide useful information to airlines,

		<p>airports and other aviation stakeholders so that they can improve their operational planning and execution, as well as any related products and services. Airlines use AI systems with built-in machine learning algorithms to collect and analyze flight data regarding each route distance and altitudes, aircraft type and weight, weather, etc. Based on findings from data, systems estimate the optimal amount of fuel needed for a flight</p>
3.	Novelty/Uniqueness	<p><b>1. Cost Reduction</b> Airlines are very concerned about baggage handling metrics like lost-bag tally, SLAs. They rely on real-time baggage tracking data to avoid losing, damaging or delaying</p>

		<p>bags and face compliance issues.</p> <p><b>2.Fuel</b></p> <p><b>Management-</b></p> <p>Airlines track real-time fuel consumption data on Dashboards from take-off to landing. This monitoring is crucial to be ultraefficient in reducing fuel costs and airline emissions.</p> <p><b>3.Revenue</b></p> <p><b>Maximization-</b>Airlines segment customers, target with personalized offers, optimize pricing in realtime using predictive analytics techniques such as modelling and forecasting.</p>
4.	Social Impact/Customer Satisfaction	<p>Trajectory Optimization</p> <ul style="list-style-type: none"> <li>• Predictive Maintenance</li> <li>• Delay Estimation</li> <li>• Targeted Advertising</li> <li>• Crew Performance Assessment</li> <li>• Sentiment Analysis</li> <li>• Prediction of Customer Behaviour.</li> </ul>



5.	Business Model(Revenue Model)	<p>The 4 Most Important Business Models for Airlines</p> <ol style="list-style-type: none"> <li>1. Full-Service Carriers. Full-service carriers are airlines that operate with a business model that includes offering a range of preflight and onboard services with the price of the ticket.</li> <li>2. Low-Cost Carriers</li> <li>3. Charter Airlines.</li> <li>4. Cargo Airlines.</li> </ol>
----	-------------------------------	---

6.	Scalability of the solution	<p>Data analytics has revolved around every industry, including aviation. Technology has changed how business is conducted and helps to make better decisions. As a result, data analytics plays a vital role in the aviation industry. It assists in collecting data and planning a powerful strategy that helps to grow business overall. According to a report, after adopting</p>
----	-----------------------------	---

# Problem Solution fit :



# REQUIREMENT ANALYSIS :

## Functional Requirements:

FR No	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Gmail.
FR-2	User Confirmation	Confirmation via Email
FR-3	Search for flights	The registered user can search one way,round trip and multiple destination flights by choosing specific dates and destination.
FR-4	Specify passenger	Customer select the number of passengers and their category either adults,infant or child.
FR-5	Sorting flight	Customer will sort the flight either by price or duration of the flight and will register.
FR-6	Better airline service	Provide better airline service by analysing time consuming,comfort of passenger.

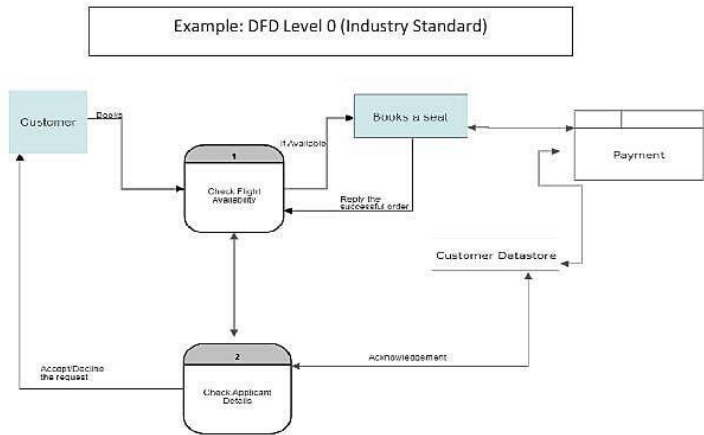
## Non Functional Requirements:

FR No	Non-Functional Requirement	Description
NFR-1	Usability	It defines how difficult it will be for a user to learn and operate the system and it can be assessed from different points of view.
NFR-2	Security	Software is protected from unauthorized access to the system and its stored data. There will be more security to the passenger.
NFR-3	Reliability	To ensure that the aircraft maintenance program tasks are effective and their periodicity is adequate.
NFR-4	Performance	Revenue is often looked at on a passenger revenue per available seat mile basis.
NFR-5	Availability	Where all required maintenance is accomplished and the aircraft is airworthy, as defined by the regulations and is considered available for flight.
NFR-6	Scalability	The capability of a system, network, or process to handle a growing amount of work.

# PROJECT DESIGN:

## Data Flow Diagrams:

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

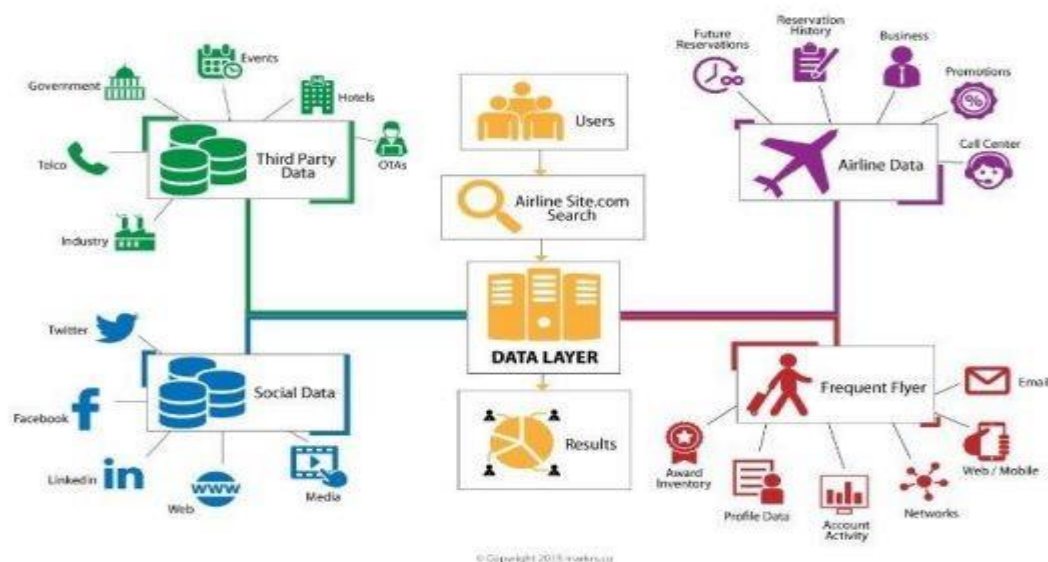


## Solution & Technical Architecture :

The Deliverable shall include the architectural diagram as below and the information as per the table1 & table2.

**Example:**

### Airline Data Analytics For Aviation Industry



## User Stories :

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Web user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the application through Gmail.		Medium	Sprint-1
	Login	USN-4	As a user, I can log into the application by entering email & password.	I can get to access my web portal	High	Sprint-1
	Dashboard	USN-5	As a user, I can get to know what my dashboard consists of.	I can my details of my registration.	Low	Sprint-2
Customer Care Executive	Organization	USN-6	The organization which owns this airplane analysis system will enable the option to customers to reach out the organization if <ul style="list-style-type: none"><li>they have any problem with the organization's system of customer interaction or</li><li>airplane issues- delay, landing in a different location</li></ul>	The customer care workers will help out the customers in trouble.	High	Sprint-1
Administrator	Administration	USN-7	The organization takes in-charge of the administrative policies of different departments like: <ul style="list-style-type: none"><li>registration</li><li>flight booking</li><li>delay visualization</li><li>generation of delay report</li></ul>	As an administrator, confirmation of user while registration is done.	High	Sprint-1



# PROJECT PLANNING AND SCHEDULING

## Sprint Planning & Estimation

Use the below template to create product backlog and sprint schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority
Sprint-1	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming that.	2	Low
Sprint-1	Registration	USN-2	As a user, I will receive confirmation email once I have registered for the application	3	High
Sprint-1	Login	USN-3	As a user, I adapt to logging into the system with credentials.	2	Low
Sprint-1	Designation of Region	USN-4	As a user, I can collect the dataset and select the region of interest to be monitored and analysed	5	Medium
Sprint-2	Exploration Of The Data	USN-5	As a developer, I will explore the given dataset through cognos.	6	High
Sprint-2	Visualization Of The Dataset	USN-6	As a developer, I will visualize the given dataset into a dashboard using cognos.	6	High
Sprint-3	Customization Of The Dashboard	USN-7	As a user, I can customize the visualized dashboard.	6	Medium
Sprint-3	Ease of Access	USN-8	As a user, I can easily access and manipulate the dashboard.	6	Medium
Sprint-4	Report Generation	USN-9	As a user, I can view the detailed report of my visualization.	6	High
Sprint-4	Establishment of the Dashboard	USN-10	As a developer, I established the dashboard into a website and submit the website.	6	High

### Project Tracker, Velocity & Burndown Chart: (4 Marks)

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	12	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	12	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	12	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	12	19 Nov 2022

### Velocity:

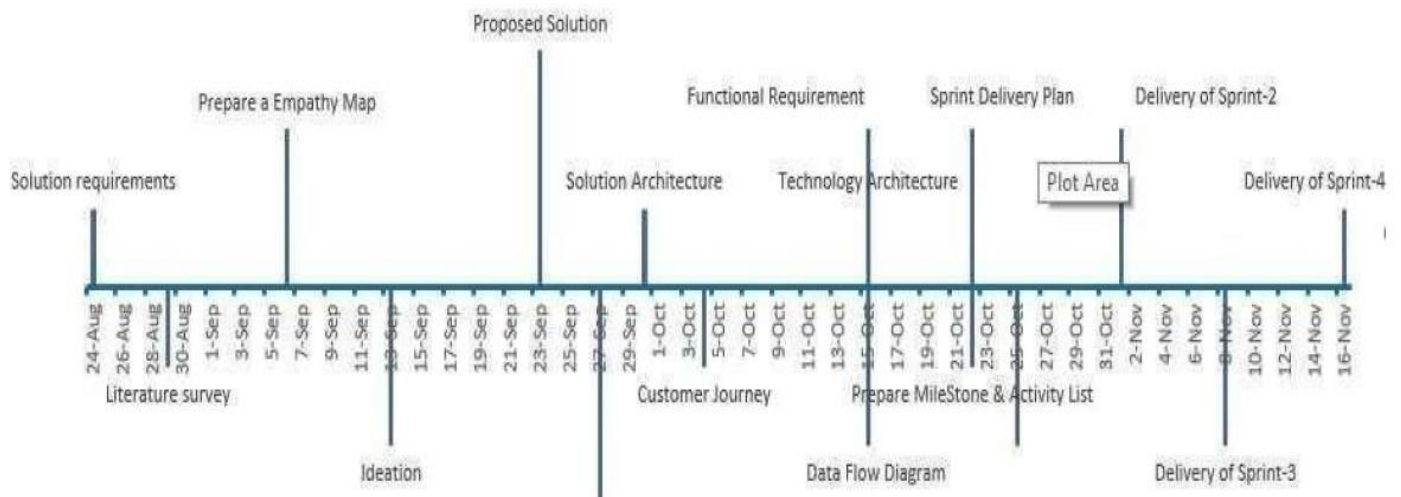
Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

$$\text{Average velocity} = \text{Sprint duration} / \text{velocity} = 12/6 = 2$$

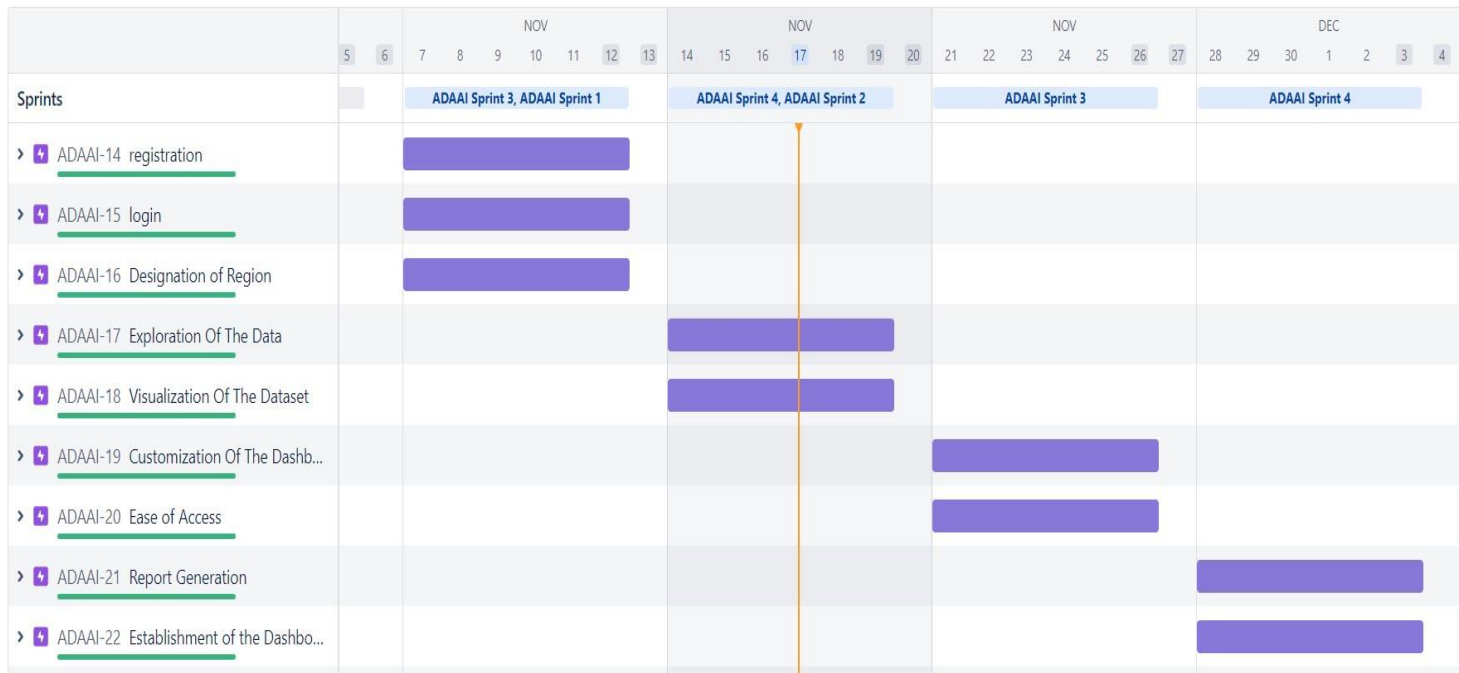


## Sprint Delivery Schedule :

A milestone schedule, or milestone chart, is a timeline that uses milestones to divide a project schedule into major phases. A milestone chart is a way to visualize the most important steps of our project. Each milestone the team achieves brings us closer to completing the project. As a result, milestones provide a sense of accomplishment and show the team how the work they're doing contributes to the overarching project objective.



# REPORT FROM JIRA



# WORKING WITH THE DATASETS AND DATA VISUALISATION :

## Working With The Dataset

- Understand the Dataset
- Load the Dataset
- Perform Joins of the Dataset tables

## Understanding The Dataset :

The data can be downloaded from the Links :  
AirStats data on airports around the world

- Airports.csv

#	Field Name	Data Type
1	id	Int
2	ident	Text
3	type	Text
4	name	Text
5	latitude_deg	Geo
6	longitude_deg	Geo
7	elevation_ft	int
8	continent	Text
9	iso_country	Text
10	iso_region	Text
11	municipality	Text
12	scheduled_service	Boolean
13	gps_code	Text
14	iata_code	Text
15	local_code	Text
16	home_link	Text
17	wikipedia_link	Text
18	keywords	Text

- Countries.csv

#	Field Name	Type
1	id	Int
2	code	Text
3	name	Text
4	continent	Text
5	wikipedia_link	Text
6	keywords	Text

- Regions.csv

#	Field Name	Type
1	id	Int
2	code	Text
3	name	Text
4	continent	Text
5	wikipedia_link	Text
6	keywords	Text

## DATASET LINK:

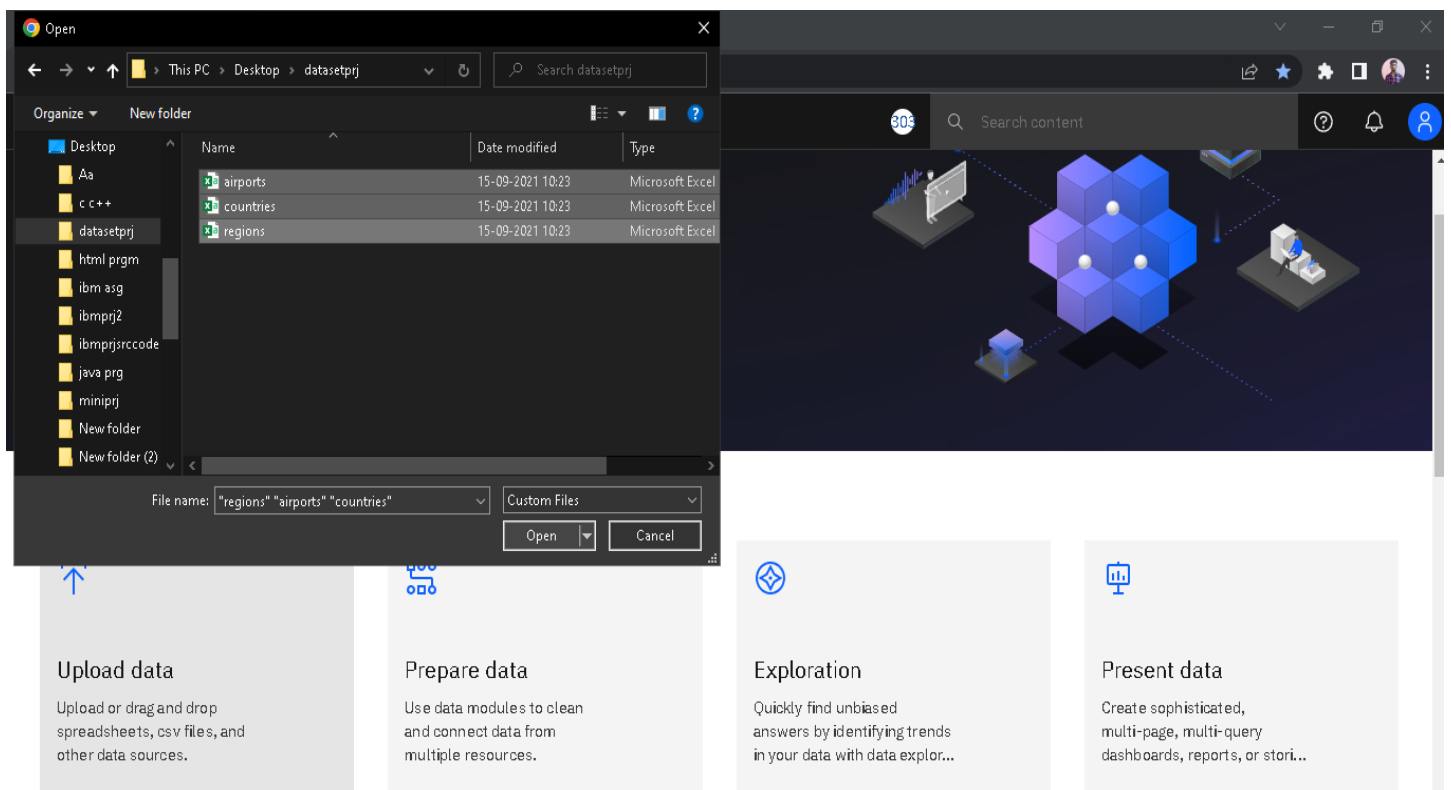
<https://www.kaggle.com/patrasaurabh/airstats-data-on-airports-around-the-world>

## Loading Of Dataset

Before you build a view and analyze your data, you must first connect the data to IBM Cognos. Cognos supports connecting to a wide variety of data, stored in a variety of places.

The data might be stored on your computer in a spreadsheet or a text file, or in a big data, relational, or cube (multidimensional) database on a server in your enterprise.

In our case, we will be using a spreadsheet or text file for making our analysis.



# Data Preparation :

- Validate all the tables - airports, countries, regions
- Create calculated field - Continent Name using the codes

IBM Cognos Analytics with Watson

AirStrats-Dataprep

Search content

Properties

Grid

Relationships

Custom tables

	Continent Name	Row Id	id	code	name	continent
Europe	1	302672	AD	Andorra	EU	
Asia	2	302618	AE	United Arab Emirates	AS	
Asia	3	302619	AF	Afghanistan	AS	
North America	4	302722	AG	Antigua and Barbuda	NA	
North America	5	302723	AI	Anguilla	NA	
Europe	6	302673	AL	Albania	EU	
Asia	7	302620	AM	Armenia	AS	
Africa	8	302556	AO	Angola	AF	
Antarctica	9	302615	AQ	Antarctica	AN	
South America	10	302789	AR	Argentina	SA	

Search

AirStrats-Dataprep

Navigation paths

airports.csv

countries.csv

Continent Name

Row Id

id

code

name

continent

wikipedia\_link

keywords

regions.csv

# Joining Of Tables :

Joining of Tables Airports, Countries and Regions with the related columns.

IBM Cognos Analytics with Watson

Create relationship

Table 1

countries.csv

code

AD

AE

AF

AG

Search

# Row Id

# id

abc code

name

continent

Table 2

regions.csv

iso\_country

AD

AD

AD

AD

Search

name

continent

iso\_country

wikipedia\_link

keywords

1

N

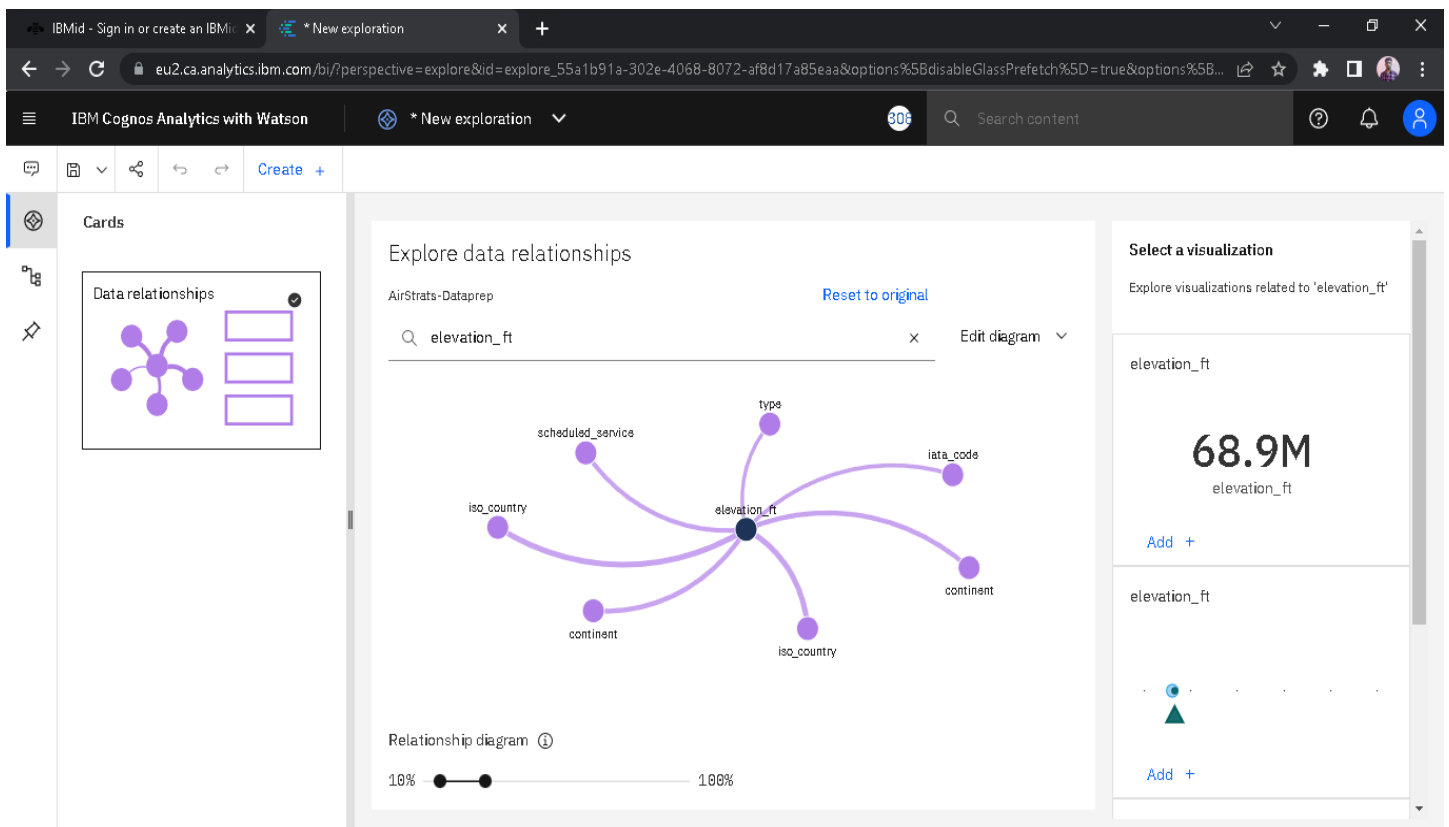
Match selected columns

Row Id	id	code	name
1	302672	AD	Andorra
2	302618	AE	United Arab Em
3	302619	AF	Afghanistan

Row Id	id	code	local_code
1	302811	AD-02	02
2	302812	AD-03	03
3	302813	AD-04	04
4	302814	AD-05	05

## Exploration Of Data :

- Explore from data directly or via an existing asset in a Dashboard or Story.
- Leverage advanced analytics in an accessible way, opening the door for any user to surface compelling new insights.
- Interact with contextual recommendations that guide users to greater understanding of their data.
- Start exploring immediately with an intuitive, natural language tool that lowers the barriers to entry for the world of analytics.





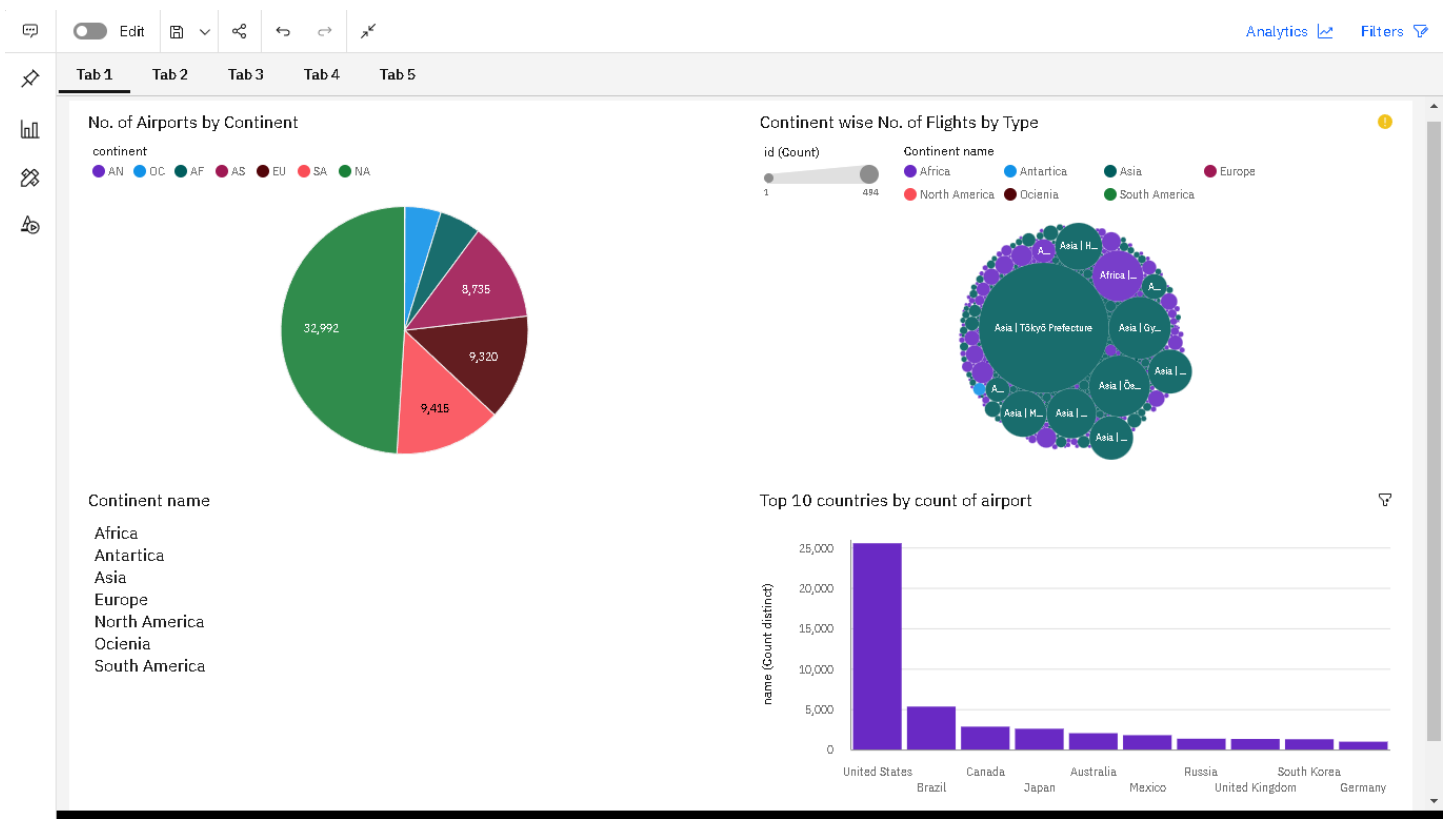
# Data Visualization :

Using the given dataset, we plan to create various graphs and charts to highlight the insights and visualizations.

## Representation Of Flight Count By Categories :

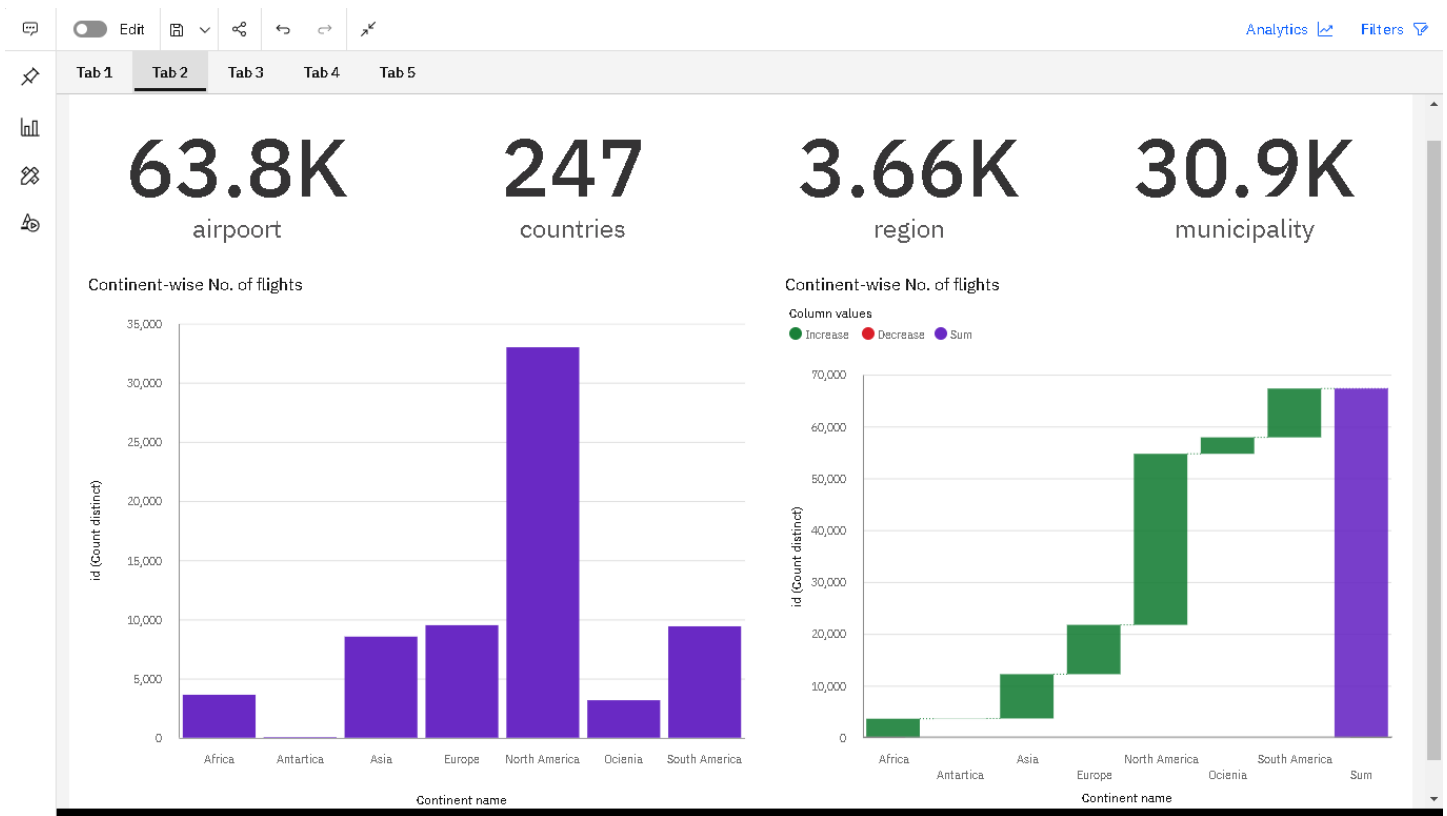
Representation of Flight Count by Categories.

- 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.



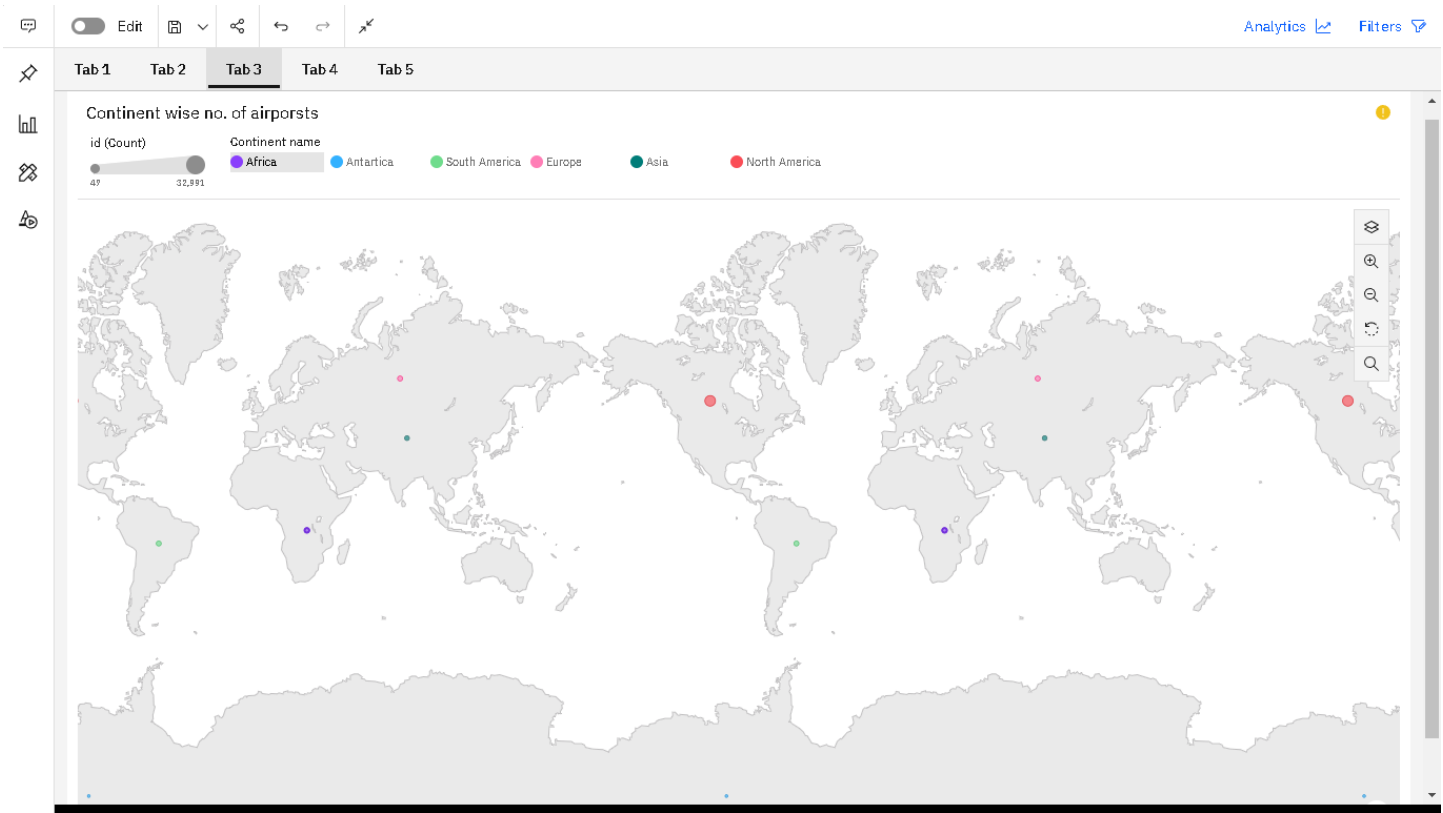
# No Of Flights By Countries , Regions And Airports :

- 1) Build the Summary Cards showing the
  - Number of Countries, Number of distinct Regions, Number of Airports and Number of Municipalities
- 2) Build the number of Airports by Countries using a Column Chart
- 3) Build a Waterfall-Chart showing the number of Airports by Continents.



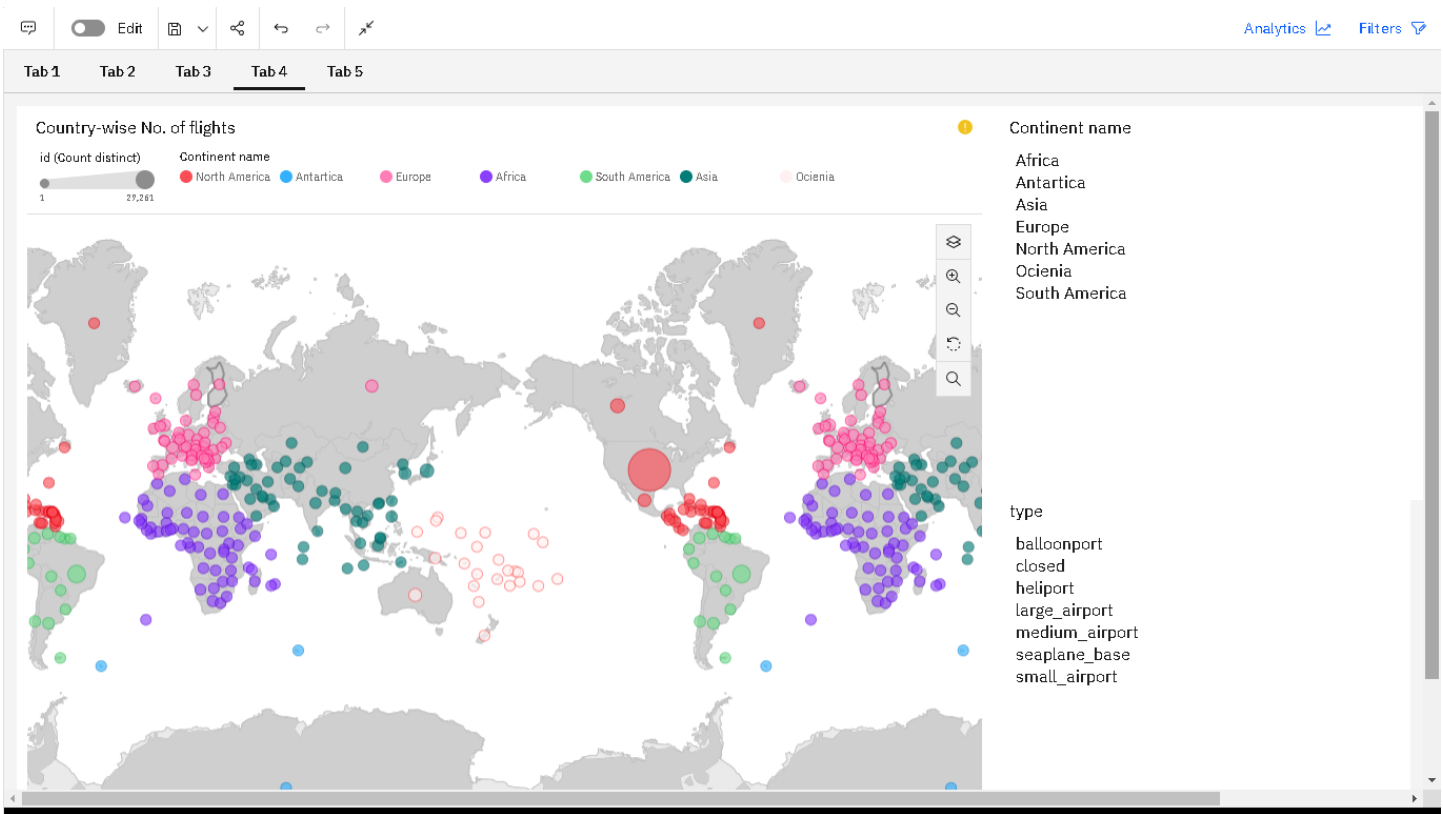
# Continent Wise Count Of Airports Using Geo Map :

Geo-Map - Continent-wise No. of flights.



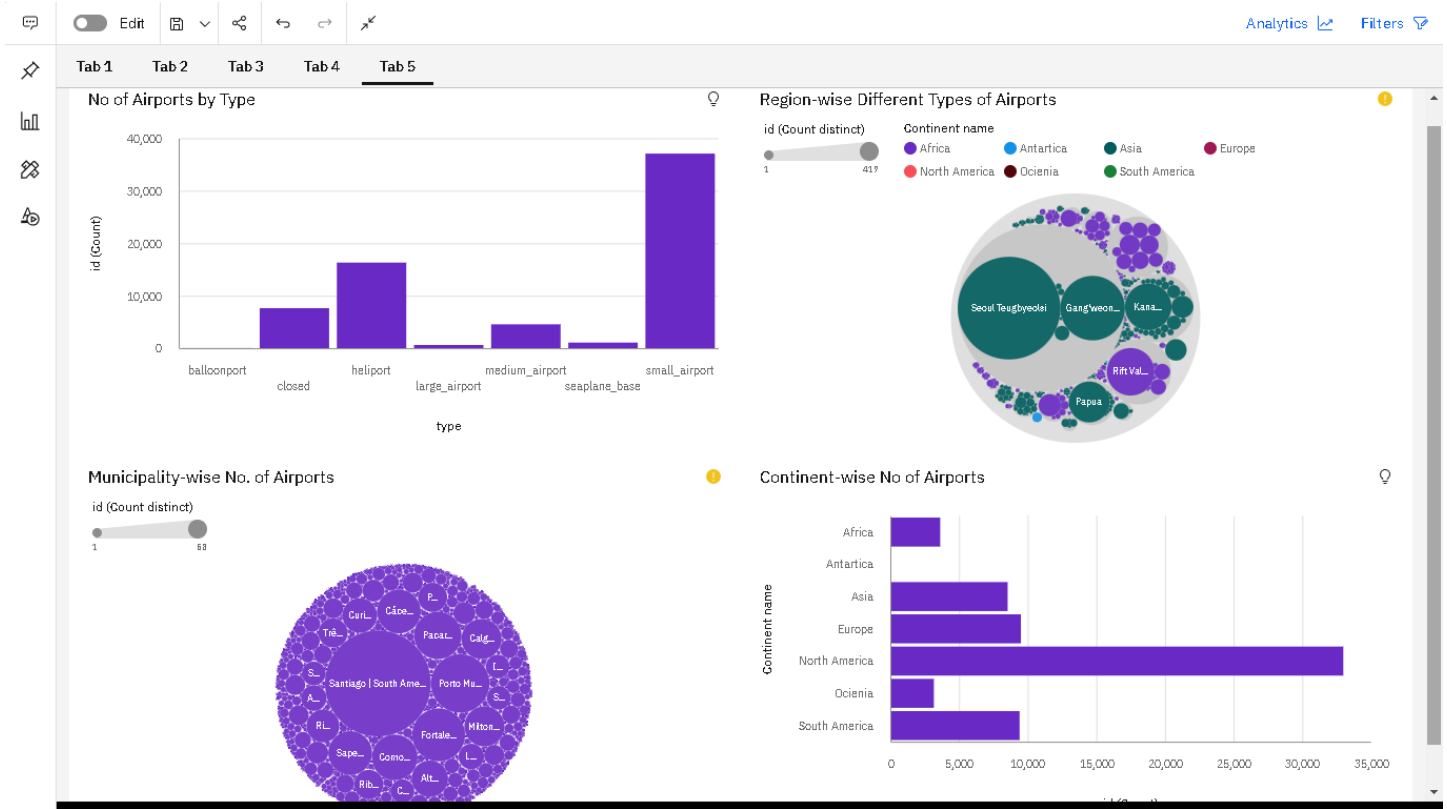
# Country Wise Airports With Types :

- 1.Geo-Map - Country-wise No. of flights
- 2.Continent Filter
- 3.Flight-Type filter



## **Dashboard showing count of flights by Types,Countries and Continents:**

1. Column-Chart - No of Airports by Type
2. Hierarchy Bubble Chart - Region-wise Different Types of Airports
3. Packed bubble Chart - Municipality-wise No. of Airports
4. Bar Chart - Continent-wise No of Airports



TESTING :

Test Cases :

Test case ID	Feature Type	Component	Test Scenario	Steps To Execute	Expected Result	Actual Result	Status
LoginPage_TC_OO1	Functional	Home Page	Verify user is able to see the Login/Sign up popup when user clicked on My account button	1.Enter URL and click go 2.Click on My Account dropdown button 3.Verify login/Sing up popup displayed or not	Login/Sig nup popup should display	Working as expected	Pass
LoginPage_TC_OO2	UI	dashboa rd page	verify user is able to see airport report in dashboa rd page	1.Airstat dashboard will be displayed. 2.Check if each tab can able to access. 3.Click on the required dataset. 4.OBTain the report	required visualisat ion wi ll be display ed on the dashboa rd	working as expected	pass

# User Acceptance Testing :

## Defect Analysis

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved.

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	10	4	2	3	20
Duplicate	1	0	3	0	4
External	2	3	0	1	6
Fixed	11	2	4	20	37
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won't Fix	0	5	2	1	8
Totals	24	14	13	26	77



# Test Case Analysis

This report shows the number of test cases that have passed, failed, and untested

Section	Total Cases	Not Tested	Fa il	Pa ss
Print Engine	7	0	0	7
Client Application	51	0	0	51
Security	2	0	0	2
Outsource Shipping	3	0	0	3
Exception Reporting	9	0	0	9
Final Report Output	4	0	0	4
Version Control	2	0	0	2

## RESULTS :

### Performance Metrics :

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visulizations / Graphs - 18
2.	Data Responsiveness	It shows the output when any of the dataset is selected.
3.	Utilization of Data Filters	Various filter methods were used to filter the dataset values like sort,top or bottom,format data etc.,
4.	Effective User Story	No of tabs Added - 5
5.	Descriptive Reports	No of Visulizations / Graphs -18

# **ADVANTAGES & DISADVANTAGES :**

## **Advantages :**

- It improves the average turnaround time needed to cater to market trends
- Properly implemented data modules help flight operators bag more customers and profits
- Predictive analytics is the key to preparing for future crises and put a mitigation plan in place
- It helps businesses make data-backed and more informed policy decisions
- Not just sales and customer service, data analytics play a vital role in flight operations and maintenance too

## **Disadvantages :**

- Air transport is a costly service. Its operational costs are too high. Middle class and poor people can not afford its cash.
- Air transport is prone to accidents. A small mistake can be very dangerous for passengers. Hijacking of planes is easily possible.
- For creating aviation facilities, huge investments are required. The cost of aero planes, construction and maintenance of aerodromes and control mechanism needs a capital expenditure.

## **CONCLUSION :**

Flight delays are a major problem in civil aviation. They incur direct and indirect costs, such as maintenance at the gate, extra fees for crew, food service, and lodging. They also affect passenger satisfaction. Flight delay is inevitable and it plays an important role in both profits and losses of the airlines. An accurate estimation of flight delay is critical for airlines because the results can be applied to increase customer satisfaction and the incomes of airline agencies. So, the prediction and analysis of flight delays are of great significance to airlines, passengers, and airports. Predicting delays will help an airport to adjust resource allocations, quickly analyse the causes, and take measures to reduce or eliminate delays. Therefore, it delivers a well-friendly graphical UI and gives a proper delay rate to the users.

## **FUTURE SCOPE :**

To illustrate, airlines bear high costs due to delays and cancellations that include expenses on maintenance and compensations to travellers stuck in airports. With nearly 30 % of the total delay time caused by unplanned maintenance, predictive analytics applied to fleet technical support is a reasonable solution.

## **APPENDIX :**

### **GitHub & Project Demo Link :**

#### **Github repositories :**

<https://github.com/IBM-EPBL/IBM-Project-27010-1660043641>

#### **Project Demo Link :**

<https://www.youtube.com/watch?v=QBZ29xqj6sM&feature=youtu.be>