

# **Airlines Data Analytics for Aviation Industry**

## **Team members**

737819ITR034 - KARAMCHAND SUBASH K

737819ITR035-KARTHI M

737819ITR033-KALAYARASI R

737819ITR036-KARTHIKEYAN S

# **TABLE OF CONTENTS**

## **1. INTRODUCTION**

1.1 Project Overview

1.2 Purpose

## **2. LITERATURE SURVEY**

2.1 Existing problem

2.2 References

2.3 Problem Statement Definition

## **3. IDEATION & PROPOSED SOLUTION**

3.1 Empathy Map Canvas

3.2 Ideation & Brainstorming

3.3 Proposed Solution

3.4 Problem Solution fit

## **4. REQUIREMENT ANALYSIS**

4.1 Functional requirement

4.2 Non-Functional requirements

## **5. PROJECT DESIGN**

5.1 Data Flow Diagrams

5.2 Solution & Technical Architecture

5.3 User Stories

## **6. PROJECT PLANNING & SCHEDULING**

6.1 Sprint Planning & Estimation

6.2 Sprint Delivery Schedule

6.3 Reports from JIRA

## **7. CODING & SOLUTIONING**

7.1 Feature 1

7.2 Feature 2

7.3 Database Schema (if Applicable)

## **8. TESTING**

8.1 Test Cases

8.2 User Acceptance Testing

## **9. RESULTS**

9.1 Performance Metrics

## **10. ADVANTAGES & DISADVANTAGES**

## **11. CONCLUSION**

## **12. FUTURE SCOPE**

## **13. APPENDIX**

Source Code

GitHub & Project Demo Link

## 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-Lamma

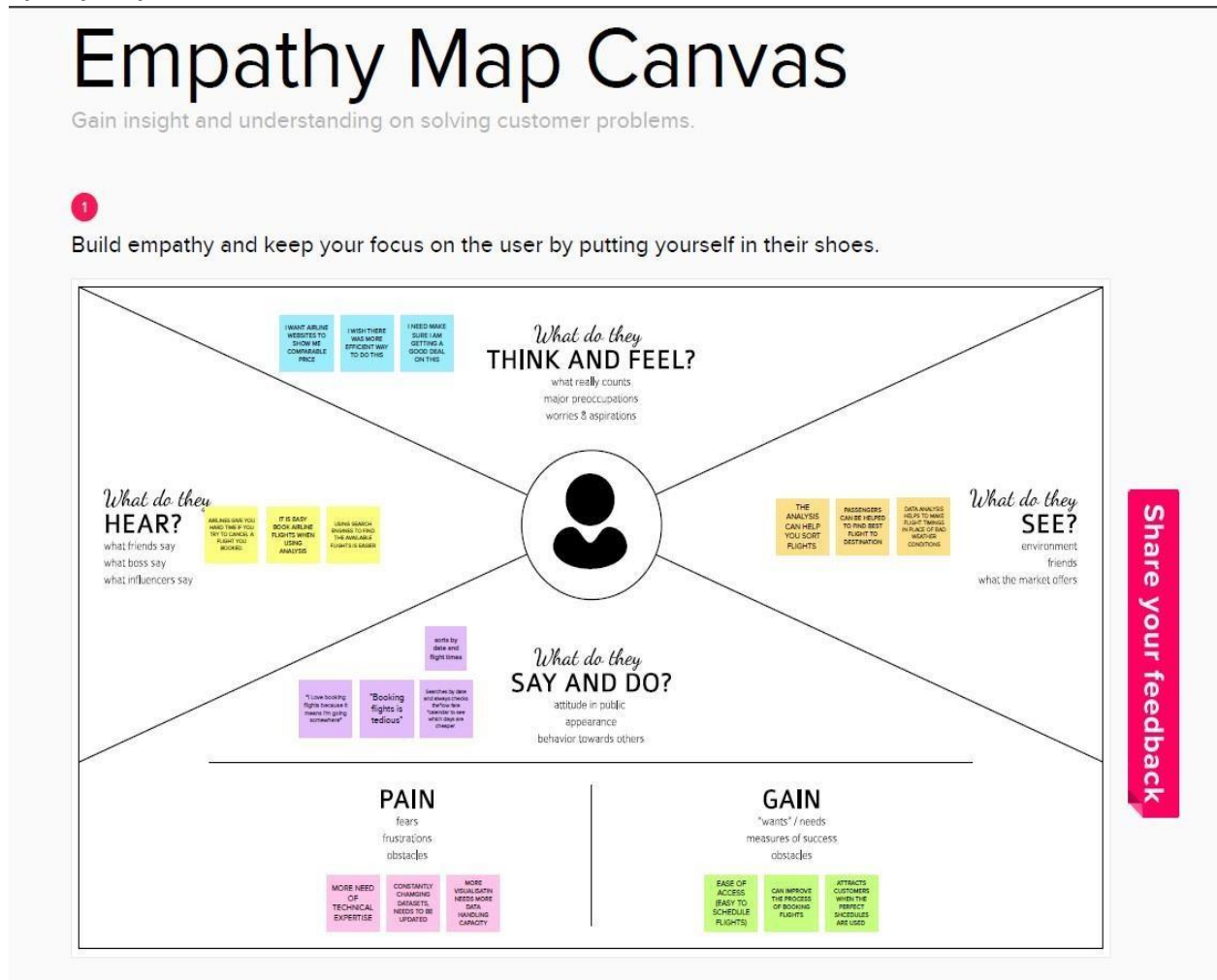
5.An Evaluation Of The Operational Performance And Profitability Of The U.S. Airlines(2021): Author:Emillio Collar

### Problem Statement Definition :

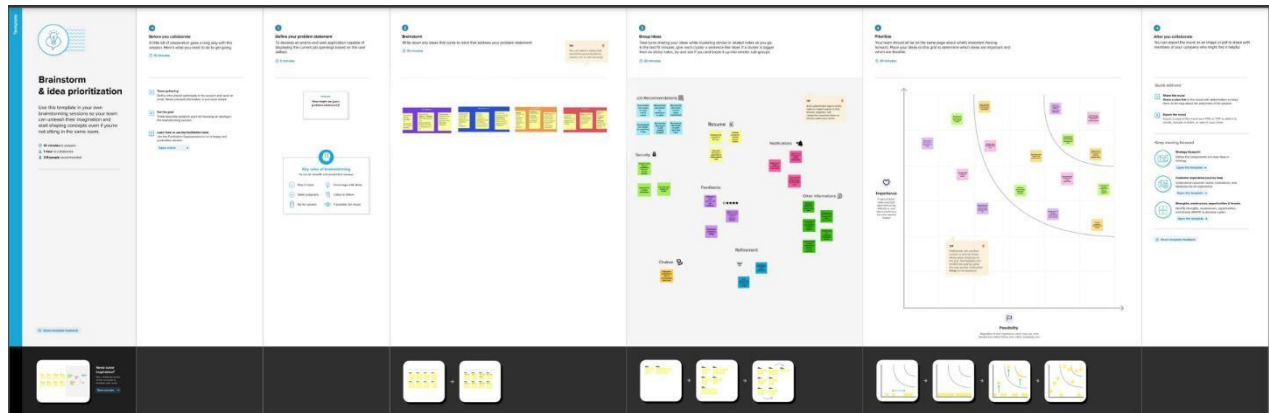
To identify and manage many people traveling this summer, they are noticing first –hand that airlines are facing major challenges, including numerous flight cancellations and delays.

### ► IDEATION & PROPOSED SOLUTION :

### Empathy Map Canvas:



## Ideation & Brainstorming :



## Proposed Solution :

S.No.	Parameter	Description
-------	-----------	-------------

1.	Problem Statement (Problem to be solved)	<p>❖ With the growing demand for air transportation and the limited ability to increase capacity at some key points in the air transportation system, there are concerns that in the future the system will not scale to meet demand. This situation will result in the generation and the propagation of delays throughout the system, impacting passengers' quality of travel and more broadly the economy.</p>
2.	Idea / Solution description	<p>❖ Understanding traveler demand for specific city pairs and pricing flights can be done using data analytics project.</p> <p>❖ Airlines use this biometric technology as a boarding option. The equipment scans travelers' faces and matches them with photos stored in border control agency databases. These can be handled with the aforementioned project.</p>
3.	Novelty / Uniqueness	<p>❖ The ultimate benefits of big data analytics include timely responses to current and future market demands, improved planning and strategically aligned decision making, as</p>

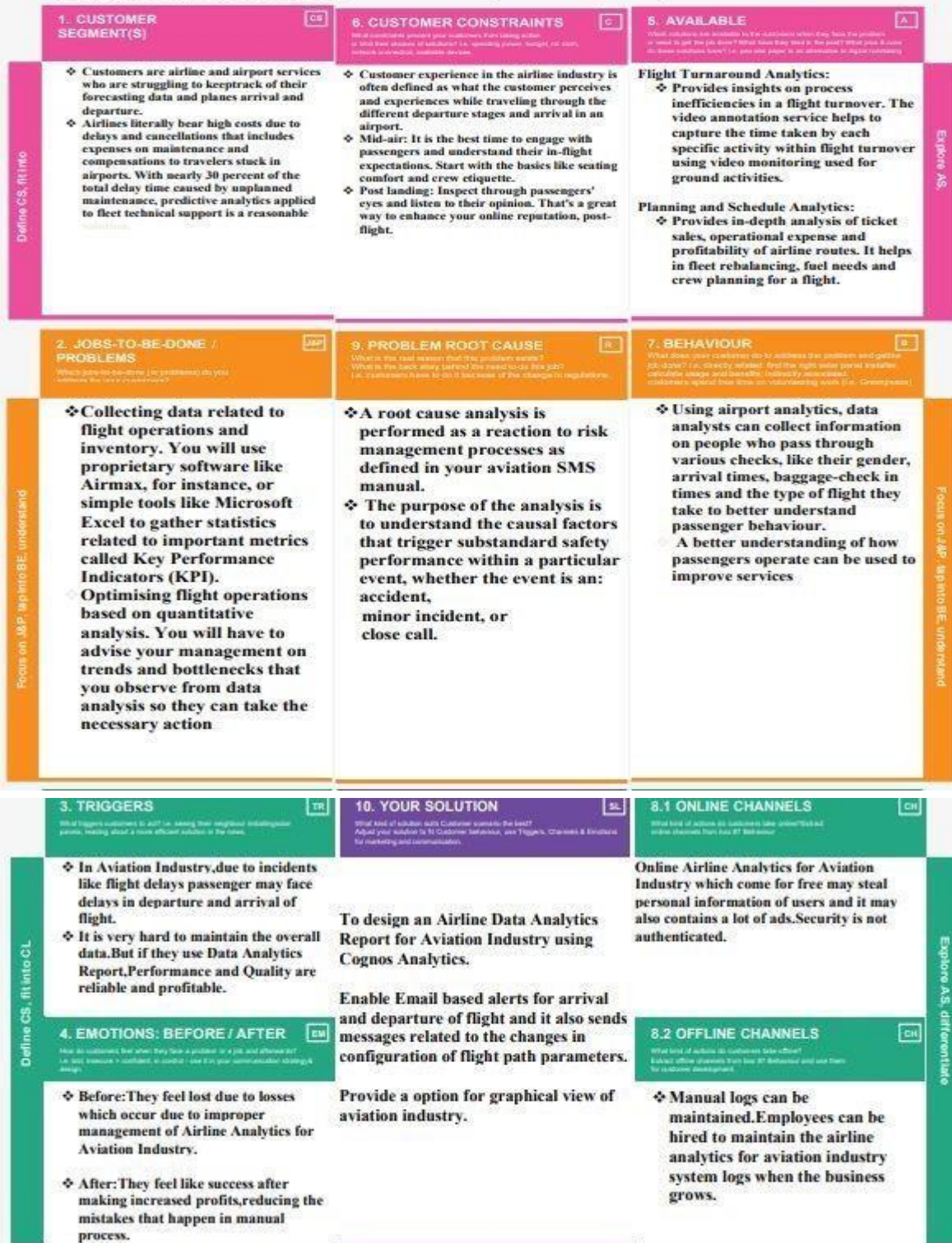
		<p>well as crystal clear comprehension and monitoring of all main performance drivers relevant to the airline industry.</p> <p>❖ Due to the use of smart data analytics, passengers will avoid many issues with baggage tracking. While radio-frequency identification prevents mishandling the baggage, predictive analysis assists in improving the predictability of fleet reliability.</p>
--	--	--

4.	Social Impact / Customer Satisfaction	<p>❖ Data analytics helps the industry to understand customers' preferences and other maintenance issues. ❖ For instance, analysis of ticket booking helps the industry to target the customers with personalized offers while optimizing the price in realtime using predictive analysis techniques. As a result, by gathering meaningful data, airlines can fetch more bookings in the given timeframe.</p>
----	---------------------------------------	---



5.	Business Model (Revenue Model)	<p>❖ Business models innovation in airlines can contribute to the creation of value, competitive advantage and profitability with new possibilities of action.</p> <p>❖ A revenue model is a</p>
		<p>blueprint that shows how a startup business will earn revenue or gross income from its standard business operations, and how it will pay for operating costs and expenses.</p>
6.	Scalability of the Solution	<p>❖ The Cloud Cognos Analytics is not only for particular organization/governments. ❖ Aviation industry acting under international, domestic or private are also getting satisfied with the aviation data analyzing process provided as per their needs.</p>

**Problem Solution fit :**



## ► REQUIREMENT ANALYSIS :

**Functional Requirements:**

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	customer Registration	customer can make Registration through Gmail
FR-2	User Confirmation	After the Registration the customer will get confirmation through mail.
FR-3	Visualizing data	User can visualize the Regular trends of delay of flights Using IBM cognos Analytics
FR-4	Generating Report	User can view the flight delay report

---

**Non-functional Requirements:**

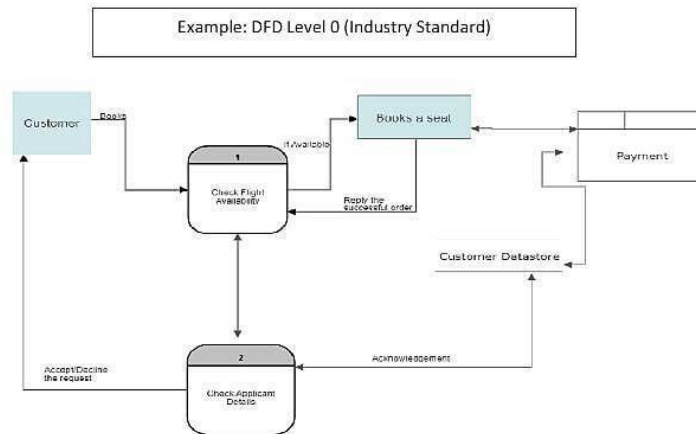
Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	<b>Usability</b>	The application will have a simple and user-friendly graphical interface. Users will be able to understand and use all the features of the application easily. Any action has to be performed with just a few clicks
NFR-2	<b>Security</b>	The main security concern is for users account hence proper login mechanism should be used to avoid hacking. The organization system should not disclose personal information of users and other organization details to public.
NFR-3	<b>Reliability</b>	When the system is disconnected or frozen due to over access at the same time, it should save all the process of the users made up to the point of abnormal happenings.
NFR-4	<b>Performance</b>	The system should require a fair amount of speed especially while browsing through the catalogue.
NFR-5	<b>Availability</b>	The system shall be available 24 hours a day 7 days a week. User can access at anytime.
NFR-6	<b>Scalability</b>	Large Number of users can access the website

**► 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.

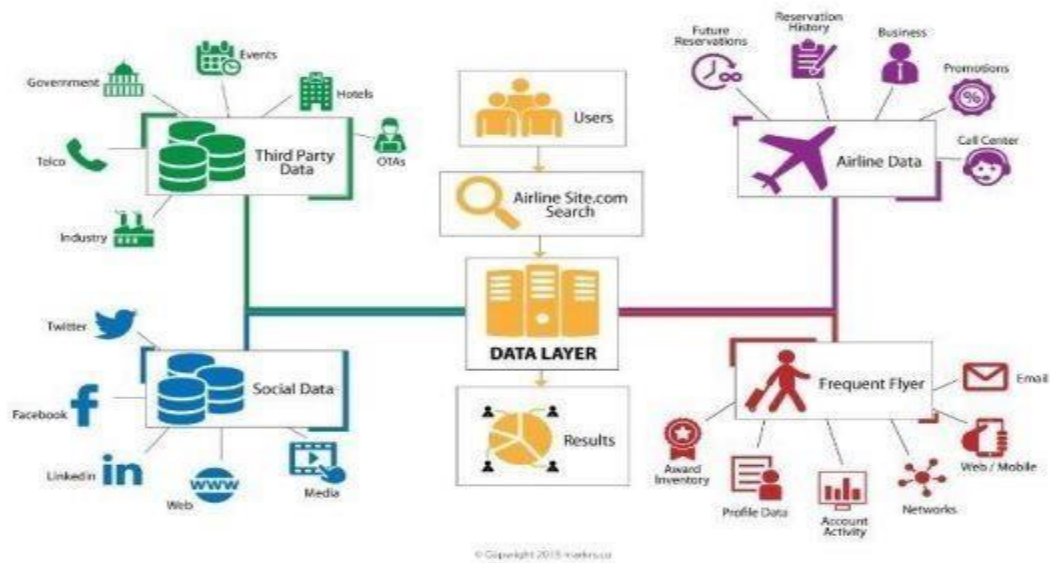


## Solution & Technical Architecture :

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 & SCHEDULING

### Sprint Planning & Estimation:

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.	5	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	10	Medium
Sprint-2	Exploration Of The Data	USN-5	As a developer, I will explore the given dataset through cognos.	8	High
Sprint-2	Visualization Of The Dataset	USN-6	As a developer, I will visualize the given dataset into a dashboard using cognos.	12	High
Sprint-3	Customization Of The Dashboard	USN-7	As a user, I can customize the visualized dashboard.	12	Medium
Sprint-3	Ease of Access	USN-8	As a user, I can easily access and manipulate the dashboard.	8	Medium



#### 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)

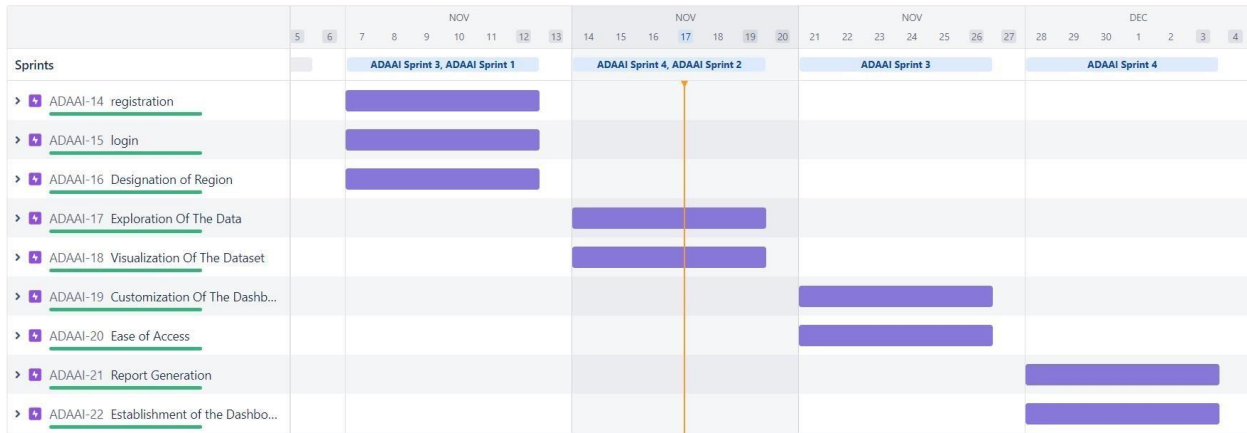
$$\text{Average velocity} = \text{Sprint duration} / \text{velocity} = 12/6 = 2$$

#### 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.



## Reports 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 :

1. AirStats data on airports around the world
  2. Circum - Airport Performance Reports
  3. Resources Coverage data
- 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	local_code	Text
4	name	Text
5	continent	Text
6	iso_country	Text
7	wikipedia_link	Text
8	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 :**

Data Preparation.

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

#### **Joining Of Tables :**

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

#### **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 :**

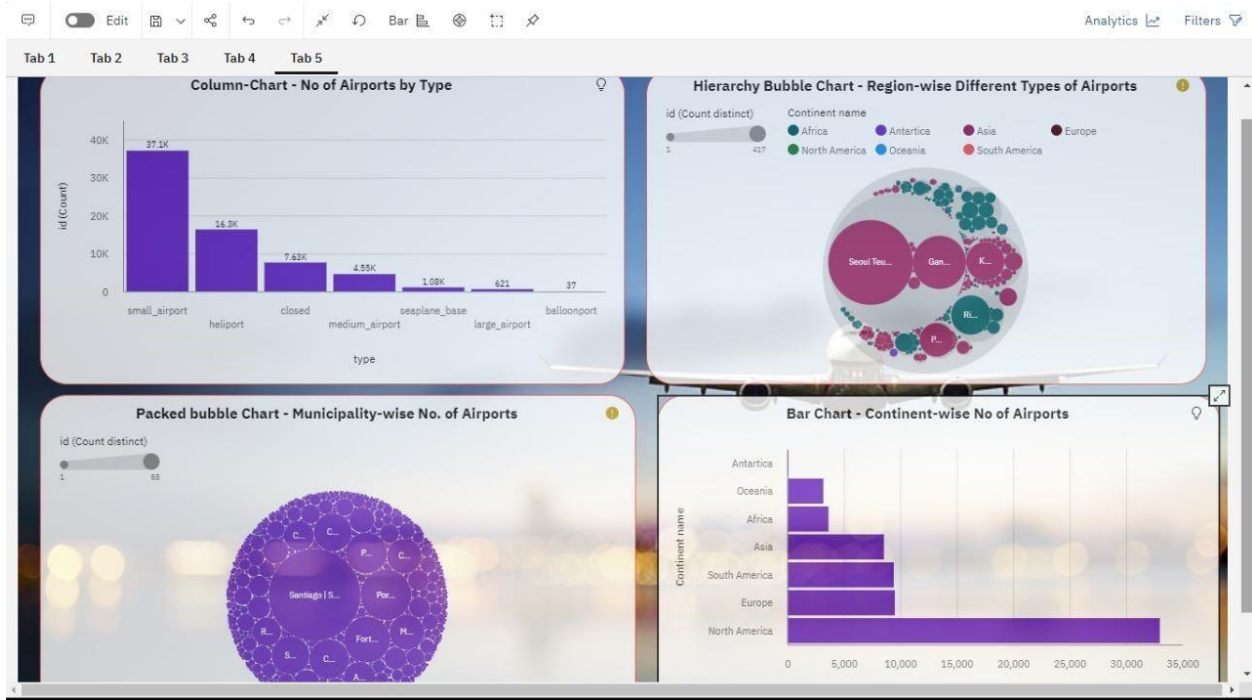
Geo-Map - Continent-wise No. of flights.

### **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 popu p 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					Subtotal
	Severity 1	Severity 2	Severity 3	Severity 4	
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
	2	0	0	2

Version

Control

► **RESULTS :**

**Performance Metrics :**

**Model Performance Testing:**

Project team shall fill the following information in model performance testing template.

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visualizations / 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 Visualizations / 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 :

### Source Code:

#### Source code for Login Page:

```
<!DOCTYPE html>
<html lang="en">

<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>Sign In...</title>
  <link rel="stylesheet" href="style.css">
  <link rel="stylesheet"
href="https://cdnjs.cloudflare.com/ajax/libs/fontawesome/5.15.3/css/all.min.css" />
</head>

<body>
  <div class="wrapper">
    <h1>Book a flight!</h1>
```

```

<form action="#">
<div class="field email">
  <div class="input-area">
    <input type="text" placeholder="Email">
    <i class="icon fas fa-envelope"></i>
    <i class="error error-icon fas fa-exclamation-circle"></i>
  </div>
  <div class="error error-txt">Email is required!!</div>
</div>
<div class="field password">
  <div class="input-area">
    <input type="password" placeholder="Passcode">
    <i class="icon fas fa-lock"></i>
    <i class="error error-icon fas fa-exclamation-circle"></i>
  </div>
  <div class="error error-txt">Password is required!!</div>
</div>
<div class="pass-txt"><a href="#">Forgot password?</a></div>
<input type="submit" value="Login">
</form>
<div class="sign-txt"> Not registered <a href="#">Signup </a></div>
</div>

<script src="script.js"></script>

</body>

</html>

```

### Source code for Dashboard page:

```

<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <meta http-equiv="X-UA-Compatible" content="IE=edge">

```

```

    <meta name="viewport" content="width=device-width, initial-scale=1.0">
<title>Flight analysis</title>
</head>
<body>
    <h1 style="text-align: center;">Airline Data Analytics in Aviation Industry</h1>

    <button><a
href="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&id=iFA7770C28D4D4F0BA3
9260F152E6A887&objRef=iFA7770C28D4D4F0BA39260F152E6A887&options%5BdisableGl
assPrefetch%5D=true&options%5Bcollections%5D%5BcanvasExtension%5D%5Bid%5D=com.i
bm.bi.dashboard.canvasExtension&options%5Bcollections%5D%5BfeatureExtension%5D%5Bid
%5D=com.ibm.bi.dashboard.core-
features&options%5Bcollections%5D%5Bbuttons%5D%5Bid%5D=com.ibm.bi.dashboard.button
s&options%5Bcollections%5D%5Bwidget%5D%5Bid%5D=com.ibm.bi.dashboard.widgets&opti
ons%5Bcollections%5D%5BcontentFeatureExtension%5D%5Bid%5D=com.ibm.bi.dashboard.co
ntentfeatures&options%5Bcollections%5D%5BsaveServices%5D%5Bid%5D=com.ibm.bi.dashboar
d.saveServices&options%5Bcollections%5D%5Btemplates%5D%5Bid%5D=com.ibm.bi.dashboar
d.templates&options%5Bcollections%5D%5BvisualizationExtension%5D%5Bid%5D=com.ibm.
bi.dashboard.visualizationExtensionCA&options%5Bcollections%5D%5BboardModel%5D%5Bi
d%5D=com.ibm.bi.dashboard.boardModelExtension&options%5Bcollections%5D%5BcontentTy
pes%5D%5Bid%5D=com.ibm.bi.dashboard.contentTypes&options%5Bcollections%5D%5Bservi
ceExtension%5D%5Bid%5D=com.ibm.bi.dashboard.serviceExtension&options%5Bcollections%
5D%5BblayoutExtension%5D%5Bid%5D=com.ibm.bi.dashboard.layoutExtension&options%5Bc
ollections%5D%5BcolorSetExtensions%5D%5Bid%5D=com.ibm.bi.dashboard.colorSetExtensio
ns&options%5Bconfig%5D%5Bproduct%5D=CA&options%5Bconfig%5D%5BeditPropertiesLa
bel%5D=true&options%5Bconfig%5D%5BenableCustomVisualizations%5D=true&options%5B
config%5D%5BassetTags%5D%5B%5D=dashboard&options%5Bconfig%5D%5BfilterDock%5
D=true&options%5Bconfig%5D%5BshowMembers%5D=true&options%5Bconfig%5D%5Bupg
rades%5D=dashboardcore%2Fjs%2Fdashboard%2Fupgrades&options%5Bconfig%5D%5BassetTyp
e%5D=exploration
&options%5Bconfig%5D%5BgeoService%5D=CA&options%5Bconfig%5D%5BsmartTitle%5D
=true&options%5Bconfig%5D%5BnavigationGroupAction%5D=true&options%5Bconfig%5D%
5BenableDataQuality%5D=false&options%5Bconfig%5D%5BmemberCalculation%5D=false&is
AuthoringMode=true&boardId=iFA7770C28D4D4F0BA39260F152E6A887" >View
Dashboard</a></button>
</body>

<style>
body{
margin: 0;
padding: 0;
background: url("https://wallpaperaccess.com/full/1470792.jpg");
    }  button{      margin-
left: 600px;      border-
radius: 10px;      padding:

```

```
20px;    background-  
color: blue;  
  
    }  
    button:hover{  
        background-color: yellow;  
    }  
</style>  
</html>
```

**GitHub & Project Demo Link :**

**Github repositories :**

<https://github.com/IBM-EPBL/IBM-Project-29325-1660124127>

**Project Demo Link :**

[https://drive.google.com/file/d/1O45J8EOhiwpgbCCQBP9euFLiGCVOsxVk/view?usp=share link](https://drive.google.com/file/d/1O45J8EOhiwpgbCCQBP9euFLiGCVOsxVk/view?usp=share_link)

