

APPENDIX 1

AIRLINE DATA ANALYTICS FOR AVIATION INDUSTRY

IBM-DOCUMENTATION

Industry Mentor(s) Name : MOHAMMED AZHAR UDDIN

Faculty Mentos(s) Name : KEERTHIKA K

TEAM ID : PNT2022TMID12425

Submitted by

JANARTHANAN S

KARTHICK RAJ S

ABDUL THAVAB A

HARISH S D

in partial fulfillment for the award of the degree

Of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

KARPAGAM COLLEGE OF ENGINEERING

ANNA UNIVERSITY: CHENNAI 600 025

NOVEMBER 2022

S.NO	Table of Content	Pg. No
1	Introduction	1
1.1	Project Overview	1
1.2	Purpose	1
2	Literature Survey	2
2.1	Existing Problem	2
2.2	References	2
2.3	Problem Statement Definition	2
3	Ideation & Proposed Solution	3
3.1	Empathy Map Canvas	3
3.2	Ideation & Brainstorming	3
3.3	Proposed Solution	4
3.4	Problem Solution Fit	4
4	Requirement Analysis	6
4.1	Functional Requirement	6
4.2	Non-Functional Requirement	7
5	Project Design	8
5.1	Data Flow Diagram	8
5.2	Solution & Technical Architecture	8
6	Project Planning & Scheduling	10
6.1	Sprint Planning & Estimation	10
6.2	Sprint Delivery Schedule	11
7	Testing	14
7.1	Test Cases	14
7.2	User Acceptance Testing	14
8	Results	15
8.1	Performance Metrics	15
9	Advantages & Disadvantages	18
10	Conclusion	18
11	Future Scope	18
12	Appendix	19
12.1	GitHub & Project Demo Link	19

1. INTRODUCTION

1.1 PROJECT OVERVIEW

Flight delays not only cause inconvenience to passengers, but also cost the carriers billions of dollars. The Federal Aviation Administration (FAA) considers a flight to be **delayed** when it is 15 minutes later than its scheduled time, while a **cancellation** occurs when the airline does not operate the flight at all for a certain reason.

Carriers attribute flight delays to several causes such as bad weather conditions, airport congestion, airspace congestion, and use of smaller aircraft by airlines. These delays and cancellations tarnish the airlines' reputation, often resulting in loss of demand by passengers. Delays and cancellations affect both passengers and air carriers. By resulting in increased travel time and increased expenses on food and lodging, they cause stress among passengers.

1.2 PURPOSE

Big data refers to the huge amounts of information in the structured and unstructured form that cannot be processed using traditional data systems. Big data technology facilitates the utilization of high volumes of external and internal data to create new products, services and improve business operations. In the era of big data, airlines can provide services that are more satisfying to customers and to stay competitive in their fierce marketplace. Airlines can reap many benefits from big data, but many challenges still remain.

Big data can also help airlines to enhance customer service. For starters, airlines, just like any other retail or e-commerce company, use big data to provide customers with personalized experiences and offers.

2. LITERATURE SURVEY

2.1 EXISTING PROBLEM

Airlines literally bear high costs due to delays and cancellations that includes expenses on maintenance and compensations to travelers stuck in airports. With nearly 30 percent of the total delay time caused by unplanned maintenance, predictive analytics applied to fleet technical support is a reasonable solution.

2.2 REFERENCES

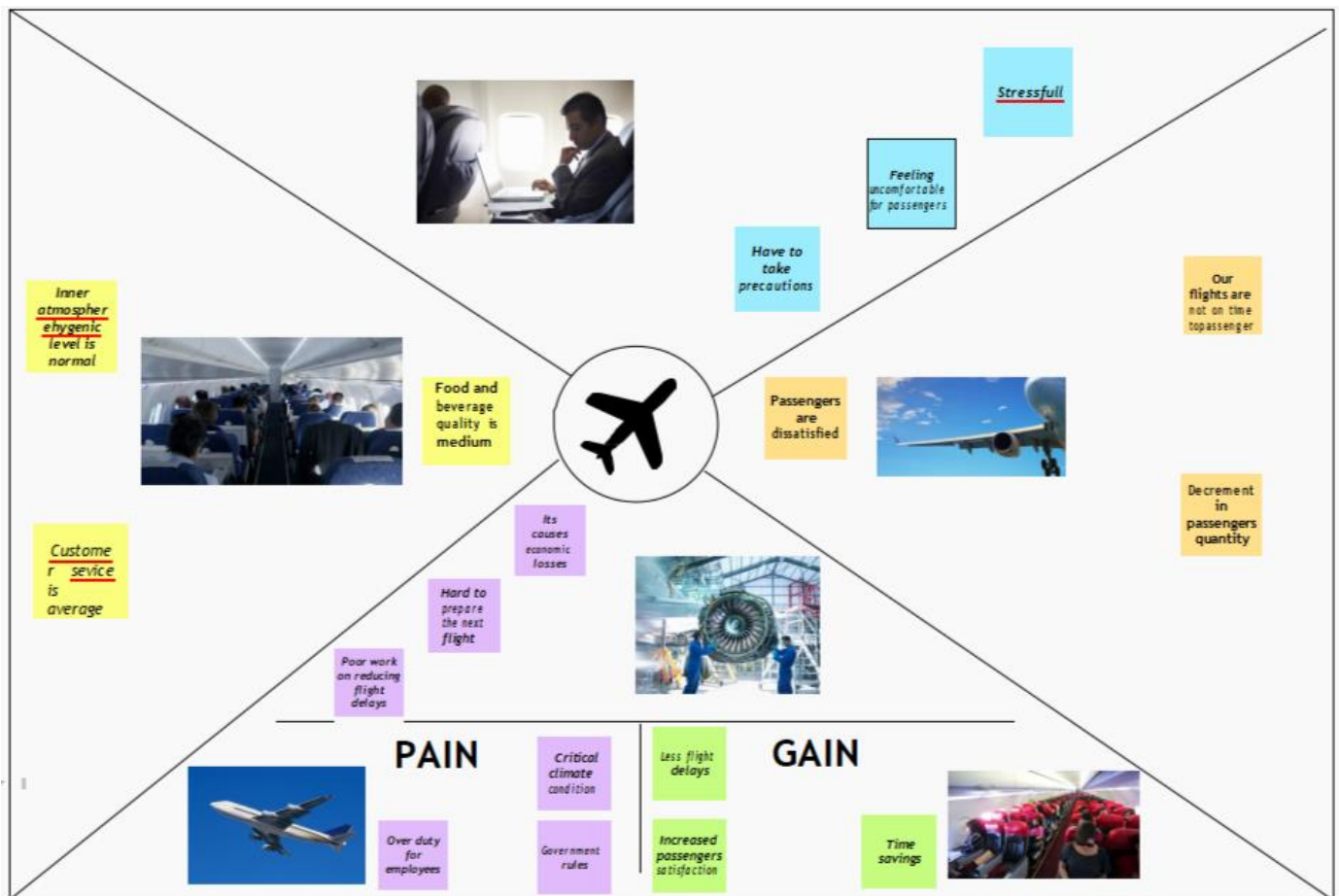
- <https://www.trefis.com/stock/dal/articles/375013/what-is-the-impact-of-flight-delays>.
- <https://blog.datumize.com/5-relevant-examples-of-a-big-data-case-study-from-the-airline-industry>
- <https://www.sciencedirect.com/science/article/abs/pii/S0739885912002053>
- <https://www.hindawi.com/journals/jat/2019/3525912/>
- https://catsr.vse.gmu.edu/SYST660/BigDataAnalysisInAviation_AOTP
- https://www.researchgate.net/publication/301589272_Flight_scheduling_through_optimizing_flight_block_and_ground_times_A_case_study
- <http://www.c-jump.com/bcc/c157c/Week10hw/Week10hw.html>
- <https://scholarworks.calstate.edu/downloads/qr46r081g>
- https://news.berkeley.edu/2010/10/18/flight_delays/3/
- <https://towardsdatascience.com/data-cleaning-101-948d22a92e4>

2.3 PROBLEM STATEMENT DEFINITION

- Collecting data set.
- Upload the dataset into cognos.
- Open the properties->data module.
- If null value is present in character field use mode method.
- If null value is present in continuous field use average or median.
- Display the data in respective charts.
- Create conclusion using summary.

3. IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS



3.2 IDEATION & BRAINSTORMING



Before you collaborate

A little bit of preparation goes a long way with this session. Here's what you need to do to get going.

10 minutes



Team gathering
Define who should participate in the session and send an invite. Share relevant information or pre-work ahead.



Set the goal
Think about the problem you'll be focusing on solving in the brainstorming session.



Learn how to use the facilitation tools
Use the Facilitation Superpowers to run a happy and productive session.

[Open article](#)



1

Define your problem statement

What problem are you trying to solve? Frame your problem as a *How Might We* statement. This will be the focus of your brainstorm.

5 minutes

PROBLEM

Using Cognos display the data



Key rules of brainstorming

To run an smooth and productive session



Stay in topic.



Defer judgment.



Go for volume.



Encourage wild ideas.



Listen to others.



If possible, be visual.

3.3 PROPOSED SOLUTION

S.No	Parameter	Description
1.	Problem Statement (Problem to be solved)	To analyze Weather data and flight departure time to calculate possibility of flight arrival and departure time and make it more convenient.
2.	Idea / Solution description	To use the Cognos analytics and Tableau, create an interactive dashboard that will answer both the customer and airline queries.
3.	Novelty / Uniqueness	One of the Efficient ways to analyze the problem is Visualization technique. The dashboard will easily provide the insight with the help of customer data.
4.	Social Impact / Customer Satisfaction	It clearly makes a huge impact on Society, Even common people can easily understand the dataset by just simply looking into it.
5.	Business Model (Revenue Model)	Business model can make revenue to the particular airline by answering the simple queries from the customer's side on how to reduce the risk of flight delay and arrival time
6.	Scalability of the Solution	The provided Solution is highly Scalable and Feasible as compared to the other common Techniques Such as Data Mining, ML, DL and soon.

3.3 PROBLEM SOLUTION FIT

Problem-Solution fit canvas 2.0 Airline Data Analytics For Aviation Industry



1. CUSTOMER SEGMENT(S)

CS

- ❖ Customers are airline and airport services who are struggling to keep track of their forecasting data and planes arrival and departure.
- ❖ Airlines literally bear high costs due to delays and cancellations that includes expenses on maintenance and compensations to travelers stuck in airports. With nearly 30 percent of the total delay time caused by unplanned maintenance, predictive analytics applied to fleet technical support is a reasonable solution.

6. CUSTOMER CONSTRAINTS

C

What constraints prevent your customers from taking action or limit their choices of solutions? (i.e. spending power, budget, no cash, network connection, available devices)

- ❖ Customer experience in the airline industry is often defined as what the customer perceives and experiences while traveling through the different departure stages and arrival in an airport.
- ❖ Mid-air: It is the best time to engage with passengers and understand their in-flight expectations. Start with the basics like seating comfort and crew etiquette.
- ❖ Post landing: Inspect through passengers' eyes and listen to their opinion. That's a great way to enhance your online reputation, post-flight.

5. AVAILABLE

A

Which solutions are available to the customers when they face the problem or need to get the job done? What have they tried in the past? What pros & cons do these solutions have? (i.e. pen and paper is an alternative to digital notetaking)

Flight Turnaround Analytics:

- ❖ Provides insights on process inefficiencies in a flight turnover. The video annotation service helps to capture the time taken by each specific activity within flight turnover using video monitoring used for ground activities.

Planning and Schedule Analytics:

- ❖ Provides in-depth analysis of ticket sales, operational expense and profitability of airline routes. It helps in fleet rebalancing, fuel needs and crew planning for a flight.

2. JOBS-TO-BE-DONE / PROBLEMS

J&P

Which jobs-to-be-done (or problems) do you

Collecting data related to flight operations and inventory. You will use proprietary software like Airmax, for instance, or simple tools like Microsoft Excel to gather statistics related to important metrics called Key Performance Indicators (KPI).

Optimising flight operations based on quantitative analysis. You will have to advise your management on trends and bottlenecks that you observe from data analysis so they can take the

9. PROBLEM ROOT CAUSE

R

What is the real reason that this problem exists? What is the back story behind the need to do this job? (i.e. customers have to do it because of the change in regulations)

- ❖ A root cause analysis is performed as a reaction to risk management processes as defined in your aviation SMS manual.
- ❖ The purpose of the analysis is to understand the causal factors that trigger substandard safety performance within a particular event, whether the event is an: accident, minor incident, or close call.

7. BEHAVIOUR

B

What does your customer do to address the problem and get the job done? (i.e. directly related: find the right solar panel installer, calculate usage and benefits; indirectly associated: customers spend free time on volunteering work (i.e. Greenpeace))

Using airport analytics, data analysts can collect information on people who pass through various checks, like their gender, arrival times, baggage-check in times and the type of flight they take to better understand passenger behaviour. A better understanding of how passengers operate can be used to improve services.

Focus on J&P, tap into BE, understand

Focus on J&P, tap into BE, understand

Define CS, fit into CL	3. TRIGGERS TR What triggers customers to act? i.e. seeing their neighbour installing solar panels, reading about a more efficient solution in the news.	10. YOUR SOLUTION SL What kind of solution suits Customer scenario the best? Adjust your solution to fit Customer behaviour, use Triggers, Channels & Emotions for marketing and communication.	8.1 ONLINE CHANNELS CH What kind of actions do customers take online? Extract online channels from box #7 Behaviour.	Explore AS, differentiate
	❖ In Aviation Industry, due to incidents like flight delays passenger may face delays in departure and arrival of flight. ❖ It is very hard to maintain the overall data. But if they use Data Analytics Report, Performance and Quality are reliable and profitable.	To design an Airline Data Analytics Report for Aviation Industry using Cognos Analytics. Enable Email based alerts for arrival and departure of flight and it also sends messages related to the changes in configuration of flight path parameters.	Online Airline Analytics for Aviation Industry which come for free may steal personal information of users and it may also contain a lot of ads. Security is not authenticated.	
	4. EMOTIONS: BEFORE / AFTER EM How do customers feel when they face a problem or a job and afterwards? i.e. lost, insecure > confident, in control - use it in your communication strategy & design.	Provide an option for graphical view of aviation industry.	8.2 OFFLINE CHANNELS CH What kind of actions do customers take offline? Extract offline channels from box #7 Behaviour and use them for customer development.	
	❖ Before: They feel lost due to losses which occur due to improper management of Airline Analytics for Aviation Industry. ❖ After: They feel like success after making increased profits, reducing the mistakes that happen in manual process.		❖ Manual logs can be maintained. Employees can be hired to maintain the airline analytics for aviation industry system logs when the business grows.	

4. IDEATION & PROPOSED SOLUTION

4.1 FUNCTIONAL REQUIREMENTS

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through website
FR-2	User Confirmation	Confirmation via Email Confirmation via Message
FR-3	Data Cleaning	We clean the data because there are many potential for data to be duplicated or incorrectly labeled when merging multiple data sources
FR-4	Reliability	Users may utilize this dashboard in an effective, efficient, and reliable manner since it is consistent and reliable for them.
FR-5	Accuracy	Dashboard accurately predicts the patient's health risks based on the length of their stay.

4.2 NON-FUNCTIONAL REQUIREMENTS

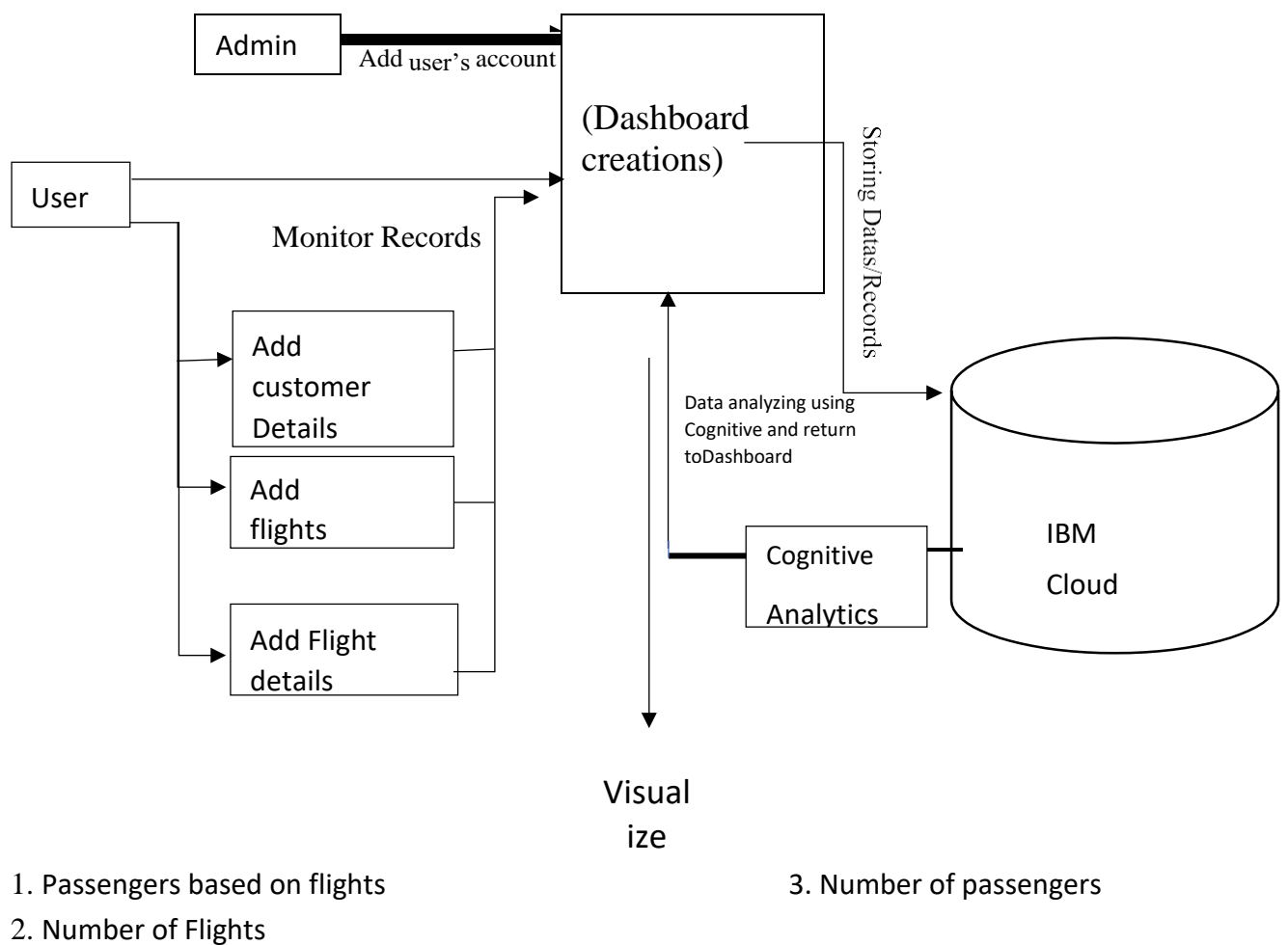
Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	In order to provide a clear understanding of the patient's Length Of Stay, this dashboard makes use of data visualization techniques including charts and graphs.
NFR-2	Security	Only users who have the password can access the website. High degrees of security are provided through the use of encryption techniques to secure the database.
NFR-3	Reliability	Users will find this dashboard to be constant and dependable, assisting them in using it effectively, efficiently, and dependably.
NFR-4	Performance	The project must respond quickly to the user's actions or even if the user has to wait the waiting period must be short.
NFR-5	Availability	The project is independent of platforms. On practically every platform, it functions flawlessly.
NFR-6	Scalability	The project enables concurrent usage of the data by several people. Because adding features and improving the website is simple, it is very scalable.

5. PROJECT DESIGN

5.1 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.



5.2 SOLUTION & TECHNICAL ARCHITECTURE

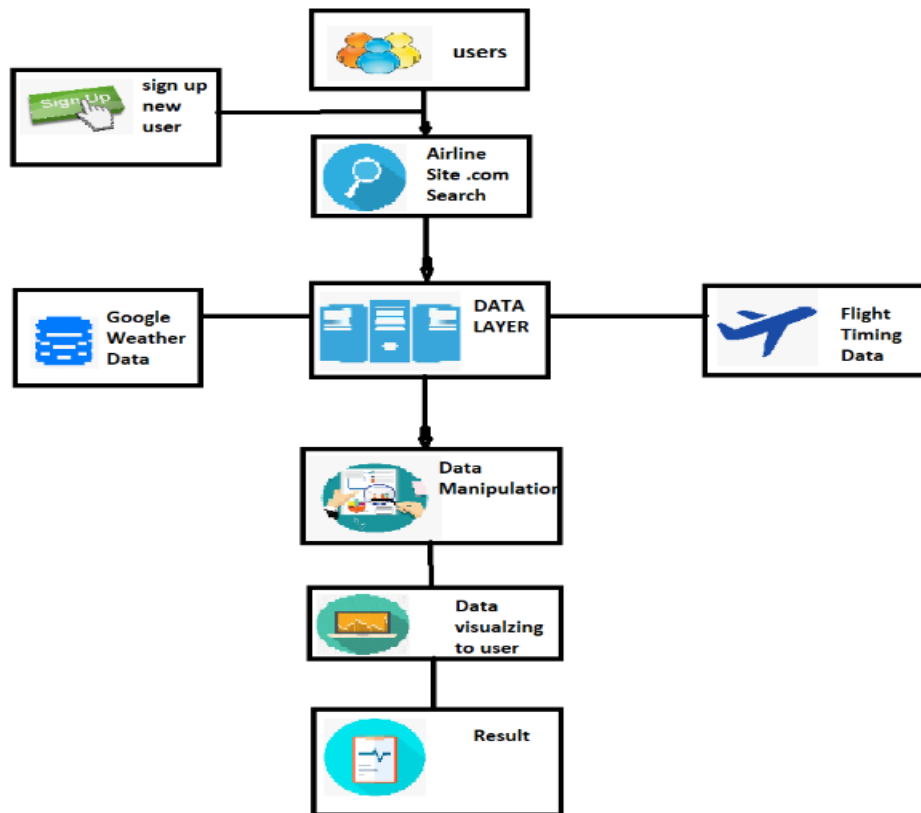
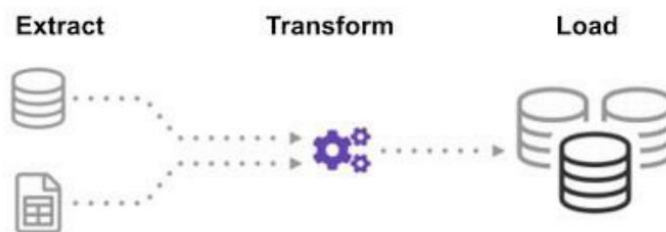


Table-1 : Components & Technologies:

S. No	Component	Description	Technology
1.	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular Js /React Js etc.
2.	Data Processing	The data from the dataset is pre-processed	IBM Cognos Analytics
3.	Database	Database Service on IBM Cloud	IBM Cloud
4.	File Storage	File storage requirements	IBM Block Storage or Other StorageService or Local Filesystem
5.	Uploading Dataset	Uploading dataset to the IBM Cognos	IBM Cognos Analytics
6.	Data Visualization	The data is visualized into different forms	IBM Cognos Analytics and python
7.	Data Prediction	Prediction of the Flight departure and Arrival Time	IBM Cognos Analytics and python

Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
1.	Open-Source Frameworks	Dashboard frameworks in the form of charts,graphs and more.	IBM Cognos
2.	Scalable Architecture	Scalability 3-tier => Database, application, server	Web server – HTML, CSS Application server- Python
3.	Availability	This dashboard is available for Aviation providers which helps them to managing the Airline resources	IBM Cognos
4.	Performance	This dashboard can contain numerous data and provides interactive visualizing performance whichhelps in easy usage	IBM Cognos



6.PROJECT PLANNING & SHEDULING

6.1 SPRINT PLANNING & ESTIMATION

MILESTONES	TASKS
MILESTONE – 1	Data Collecting process (Datasets)
MILESTONE – 2	Required Datasets are uploaded on the IBM Cognitive Platform.
MILESTONE – 3	Data Exploration and Data Visualization
MILESTONE - 4	To Create a Interactive Dashboard.
MILESTONE - 5	Display the Insights in the Dashboard

MILESTONE - 6	Construct a Standardized Data Set and use the needed data with the Assistance of a Python Program
MILESTONE - 7	Use of different algorithm with Google Colab to achieve the desired result with more accuracy.
MILESTONE - 8	Making the output simpler to understand and more efficient.
MILESTONE - 9	Deployed in the Github

6.2 SPRINT PLANNING & ESTIMATION

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Register	USN-1	Data Collecting process (Datasets) Required Datasets are uploaded on the IBM Cognitive Platform	10	Medium	Janarthanan,Karthick raj
Sprint-1	Login	USN-1	Data Exploration and Data Visualization	20	High	Abdul Thavab,Harish
Sprint-2	Dashboard	USN-2	To Create a Interactive Dashboard. Display the Insights in the Dashboard	10	High	Janarthanan,Harish
Sprint-2	Dashboard	USN-3	Construct a Standardized Data Set and use the needed data with the Assistance of a Python Program	10	High	Abdul Thavab,Harish
Sprint-3	Dashboard	USN-4	Use of different algorithm with Google Colab to achieve the desired result with more accuracy.	20	High	Karthickraj,Jaanarthanan

Sprint-4	Virtualizes	USN-5	Making the output simpler to understand and more efficient. Deployed in the Github.	20	High	Harish,Abdul Thavab
----------	-------------	-------	--	----	------	---------------------

6.3 Sprint Delivery Schedule

TITLE	DESCRIPTION	DATE
Literature Survey & Information Gathering	Literature survey on the selected project & gathering information by referring the technical papers, research publications , journals etc.	1 SEPTEMBER 2022
Prepare Empathy Map	Prepare Empathy Map Canvas to capture the user Pains & Gains, Prepare list of problem Statements that are to be solved by this project.	7 SEPTEMBER 2022
Ideation	List the ideas by organizing a brainstorming session and prioritize the top 3 ideas based on the feasibility & importance.	14 SEPTEMBER 2022
Proposed Solution	Prepare the proposed solution document, which includes novelty, feasibility of idea, revenue model, social impact, scalability of solution, etc.	21 SEPTEMBER 2022
Problem Solution Fit	Prepare problem - solution fit document.	27 SEPTEMBER 2022
Solution Architecture	Prepare solution architecture document.	29 SEPTEMBER 2022

7.TESTING

7.1 TEST CASES

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

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	0	0	0	0
Client Application	0	0	0	0
Security	0	0	0	0
Outsource Shipping	0	0	0	0
Exception Reporting	0	0	0	0
Final Report Output	0	0	0	0
Version Control	0	0	0	0

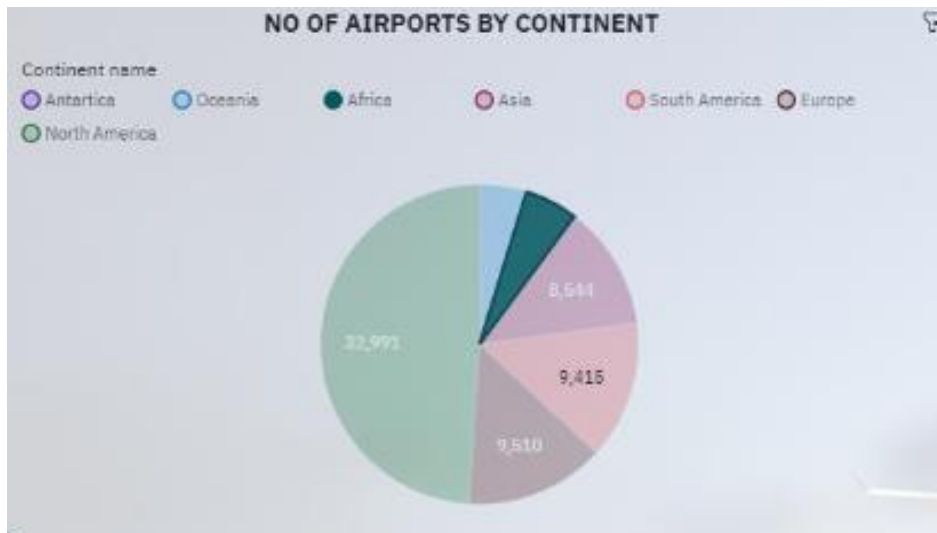
7.2 USER ACCEPTANCE TESTING

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

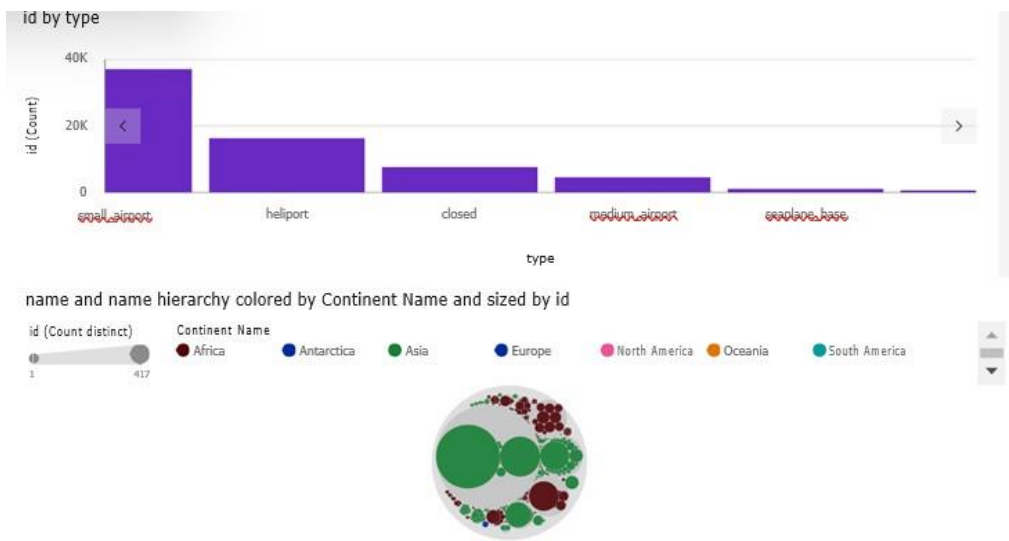
Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	5	5	3	0	13
Duplicate	0	0	0	0	0
External	7	5	1	0	13
Fixed	11	8	7	5	31
Not Reproduced	1	0	0	0	1
Skipped	0	0	0	0	0
Won't Fix	0	0	0	0	0
Totals	24	18	11	5	58

8. RESULT

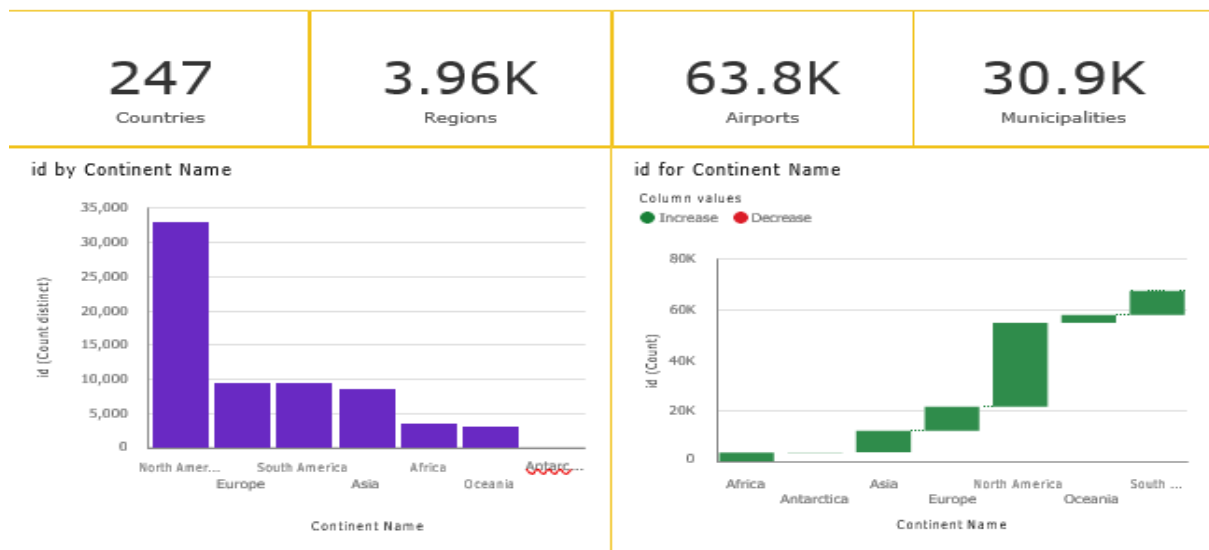
- Dashboard for Number of Airports by continent



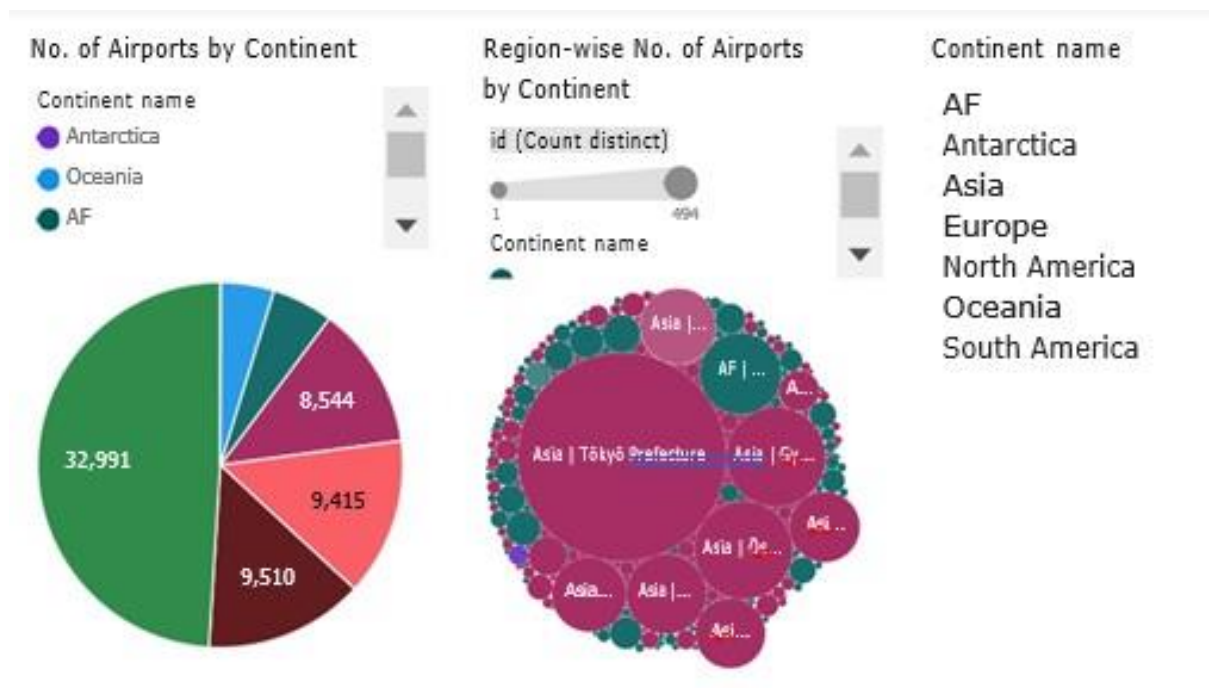
- Count of flights by types,countries and continents



● No Of Flights By Countries,Regions and Airports



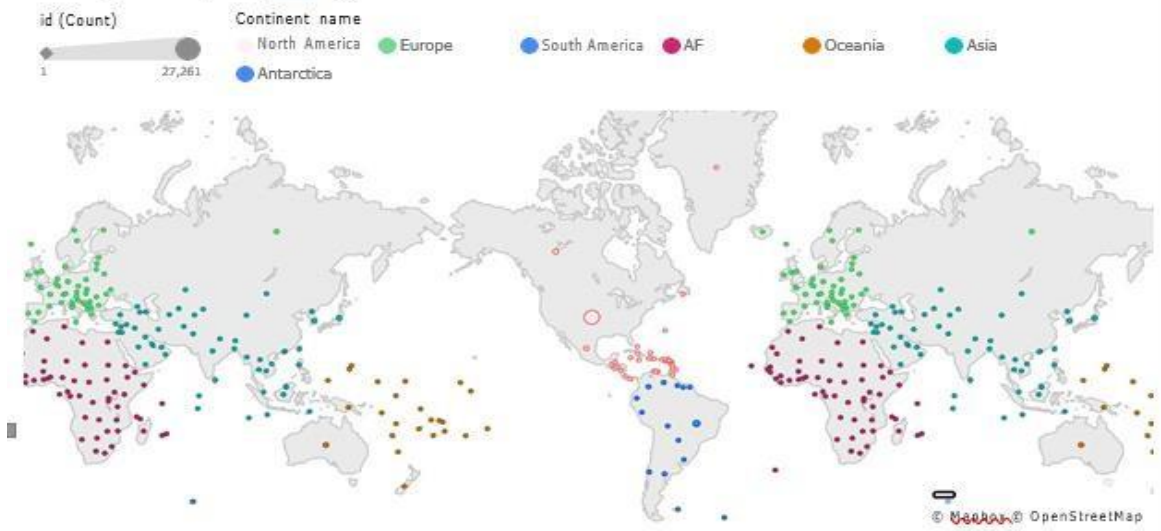
● Representation Of Flight Count By Categories



● Country wise Airports With Types

Tab 1

Country Wise Airports With Types



● Continent wise Count of Airports Using Geo Map



9. ADVANTAGES & DISADVANTAGES

ADVANTAGES:

- Helps an organization and customer to make a better decision
- Increase the efficiency of the work
- The analytics keeps you updated of your customer behavioral changes.
- Personalization of Flight details.
- Improving quality of service and Airlines.

DISADVANTAGES:

- Lack of alignment within teams
- Lack of commitment and patience
- Low quality of data
- Privacy Concerns
- Complexity and Bias

10.CONCLUSION

The project is completed successfully which accurately predict the Arrival and departure time for each flight on case by case basis so that the Aviation Industry and customers can use this information for optimal resource allocation and better functioning.

11.FUTURE SCOPE

Airline data analytics helps in analyzing the Flight details via hospital that the availability of Flights and number of Flights to the Customer and hence, it reduce the manpower and time of the respective Airport

12. FUTURE SCOPE

Github

<https://github.com/IBM-EPBL/IBM-Project-13113-1659511266>