TEAM ID	PNT2022TMID28566
PROJECT NAME	AIRLINE DATA ANALYTICS FOR AVIATION INDUSTRY

TEAM MEMBERS

LOKESH V
SOLOMON SAGAYAM D J
ASHWIN P
ALBIN FELIX A

Airlines Data Analytics for Aviation Industry

INTRODUCTION:

Airlines Data Analytics for Aviation Industry

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.

Goal of the Project:

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.

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.

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.

Data Science and Analytics in Aviation (2020):

Authors: Sai-Ho-Chung, Hoi-Lam-ma

The researcher in this article cited that, Due to the rapid development of advanced technologies nowadays, a massive amount of real time data regarding flight information, flight performance, airport conditions, air traffic conditions, weather, ticket prices, passengers comments, crew comments, etc., are all available from a diverse set of sources, including flight performance monitoring systems, operational systems of airlines and airports, and social media platforms.

Development of data analytics in aviation and related applications is also growing rapidly. This paper concisely examines data science and analytics in aviation studies in several critical areas, namely big data analysis, air transport network management, forecasting, and machine learning. The papers featured in this special issue are also introduced and reviewed, and future directions for data science and analytics in aviation are discussed.

Data Analytics for Air Travel Data (2021):

Authors: Haiman Tian, Yudong Tao

The researcher in this article cited that, From the start, the airline industry has remarkably connected countries all over the world through rapid long-distance transportation, helping people overcome geographic barriers. Consequently, this has ushered in substantial economic growth, both nationally and internationally. The airline industry produces vast amounts of data, capturing a diverse set of information about their operations, including data related to passengers, freight, flights, and much more. Analyzing air travel data can advance the understanding of airline market dynamics, allowing companies to provide customized, efficient, and safe transportation services. Due to big data challenges in such a complex environment, the benefits of drawing insights from the air travel data in the airline industry have not yet been fully explored. They introduce existing data sources commonly used in the papers surveyed and summarize their availability. Finally, we discuss several potential research directions to better harness airline data in the future. They anticipate this study to be used as a comprehensive reference for both members of the airline industry and academic scholars with an interest in airlineresearch.

Topological Data Analysis for Aviation Applications (2018):

Authors: Max Z. Li,Megan S. Ryerson and HamsaBalakrishnan

Aviation data sets are increasingly high-dimensional and sparse. Consequently, the underlying features and interactions are not easily uncovered by traditional data analysis methods. Recent advancements in applied mathematics introduce topological methods, offering a newapproach to obtain these features. This paper applies the fundamental notions underlying topological data analysis and persistent homology (TDA/PH) to aviation data analytics. We review past aviation research that leverage topological methods, and present a new computational case study exploring the topology of airport surface connectivity. In each case, we connect abstract topological features with real-world processes in aviation, and highlight potential operational and managerial insights.

Operational Efficiency Versus Financial Mobility in the Global Airline Industry (2015):

Author: Hoi-Lam-ma

The researcher in this article cited that, the events of September 11th,2001 precipitated an almost unprecedented financial crisis for the world airline industry. However, it is not clear that these events represent a discrete, industry disruption or whether, in fact, airlines were already entering a period of economic challenges that would demand new strategic orientations on their part. This study investigates the structural drivers of operational efficiency as well as the financial posture of airlines on the eve of September 11th.

A sample of 38 airlines from North America, Europe, Asia and the Middle East was utilized to investigate whether relative operational efficiency

implied superior financial mobility (as defined by Donaldson). Data envelopment analysis was utilized to derive efficiency scores for individual airlines. The underlying structural drivers of efficiency were then investigated. It was found that the traditional framework developed in the literature still provided reasonable explanatory power for realized relative operational efficiency. However, the second stage of the analysis found that relative operational efficiency did not inherently imply superior financial mobility. As such, airlines that had chosen relatively efficient operational strategies found themselves in positions of vulnerability with regard to financial mobility and thus suffered the consequences in the post-September 11th environment.

An Evaluation of the Operational Performance and Profitability of the U.S.Airlines (2021):

Author: Emillio Collar

The researcher in this article cited that, since 2008, a series of mega-mergers has dramatically changed the U.S. airline industry. Despite the presence of fewer airlines in the market, the competition remains intense, which forces airlines to continually search for ways to increase their efficiency to maintain survival and financial sustainability. To evaluate airline performance and disentangle the causes of inefficiency, this paper applied a two-stage network data envelopment analysis approach and a truncated regression to investigate the performance of nine U.S.-based airlines from 2015 to 2019. Our empirical results reveal that during the sample period, airlines' operating efficiency steadily improved, but the efficiency in the profitability stage stagnated. Therefore, strategic resource allocations are needed for airlines to see further advances in their overall efficiency. On average, airlines operating in the low-cost business model yielded higher

efficiency scores than their peers operating in the full-service framework. While an airline's size, measured in terms of total assets, has a positive influence on operating efficiency, a larger number of full-time employee equivalents hinders efficiency outcomes, which indicates the importance of enhancing labor efficiency among carriers.

The Relationship Between On-Time Performance And Airline Market Share(2020):

Author:Yoshinori Suzuki

The researcher in this article cited that, we propose a new method of modeling the relationship between on-time performance and market share in the airline industry. The idea behind the method is that the passengers' decision to remain (use same airline) or switch (use otherairlines) at time *t* depends on whether they have experienced flight delays at time *t*−1 or not. More specifically, we posit that the passengers who experienced flight delays are more likely to switch airlines for the subsequent flight than those passengers who did not experience delays. To capture such effect, we develop an aggregate- level Markovian type model that estimates the transition probability matrices separately for the passengers who experienced flight delays at time *t*−1 and for those who did not experience delays. The model was calibrated with the US DOT data. The study results imply that, once experiencing flight delays, passengers are more likely to

switch airlines. The results also imply that on-time performance affects a carrier's market share primarily through the passengers' experience, and not though the "advertisement" of performance.

Airline Finance (2021):

Author:Peter.S.Morel

The researcher in this article cited that, It is supported at each stage by practical airline examples and recent data, *Airline Finance* examines the financial trends and longer term prospects for the airline industry as a whole, contrasting the developments for the major regions and airlines together with critical discussion of key issues that affect the industry as a whole. Important techniques in financial analysis are applied to the airlines as well as their investors such as banks and other financial institutions. This book is written for employees of airlines, airports and their suppliers, and investment bank and other analysts. It is also popular for use by universities and in-house courses on air transport management, within both academia and industry.

Airline Route Profitability Analysis And Optimization Using Big Data Analytics On Aviation Data Sets Under Heuristic Techniques (2016):

Authors: Kasturi E, Prasanna Devi Sb, Vinu Kiran Sb,Manivannan Sc

Researchers in this article cited that, applying vital decisions for newairline routes and aircraft utilization are important factors for airline decision

making. For data driven analysis key points such as airliners route distance, availability on seats/freight/mails and fuel are considered. The airline route profitability optimization model is proposed based on performing Bigdata analytics over large scale aviation data under multiple heuristic methods, based on which practical problems are analyzed. Analysis should be done based on key criteria, identified by operational needs and load revenues from operational systems e.g. passenger, cargo, freights, airport, country, aircraft, seat class etc. The result shows that the analysis is simple andconvenient with concrete decision.

Analysis Of Flight Data Using Clustering Techniques For Detecting Abnormal Operations(2015):

Author:Lishaui Li,Santanu Das

The researcher in this article cited that,the airline industry is movingtoward proactive risk management, which aims to identify and mitigate risks before accidents occur. However, existing methods forsuch efforts are limited. They rely on predefined criteria to identify risks, leaving emergent issues undetected. This paper presents a newmethod, cluster-based anomaly detection to detect abnormal flights, which can support domain experts in detecting anomalies and associated risks from routine airline operations. The new method, enabled by data from the flight data recorder, applies

clustering techniques to detect abnormal flights of unique data patterns. Compared with existing methods, the new method no longer requires predefined criteria or domain knowledge. Tests were conducted usingtwo sets of operational data consisting of 365 B777 flights and 25,519A320 flights. The performance of cluster-based anomaly detection to detect abnormal flights was compared with those of multiple kernel anomaly detection, which is another data-driven anomaly detection algorithm in recent years, as well as with exceedance detection, whichis the current method employed by the airline industry. Results showed that both cluster-based anomaly detection to detect abnormal flights and multiple kernel anomaly detection were able to identify operationally significant anomalies, surpassing the capability of exceedance detection. Cluster-based anomaly detection to detect abnormal flights performed better with continuous parameters, whereas multiple kernel anomaly detection was more sensitive towarddiscrete parameters.

Data Analytics Of Skytrax's Airport ReviewAnd Ratings(2015):

Author:KrityaBunchongchit

The researcher in this article cited that, this study investigates the perception of passengers of airport service attributes, using data from the Skytrax Airport Review websites. Overall, a total of 7358 reviews were collected from the website, together with other related passenger data, namely review headers, passenger types, rating scores of airport attributes and the overall rating. This study focused on investigating each group of passenger types to identify underlying differences amongst airport's

passenger segmentation, particularly on the leisure travelers. The study performed different techniques of data analysis including sentiment analysis, lemmatization and partial least square – structural equation modelling (PLS-SEM) to reveal key patterns derived from the available data, which the normal survey data or the interview data may not have revealed. The research contributes to airport passenger segmentation by highlighting the differences found in the travelers segmented by Skytrax. The study also provides practical implications to airport managers.

Post Pandemic Aviation Market Recovery: Experience and lessons from China(2021)

Author:Achim.I.Czemy

The researcher in this article cited that, China was the first aviation market in the world hit hard by COVID-19 and has been recovering gradually as the pandemic became largely under control within mainland China. This study reviews the recovery pattern influenced by the Chinese government's aviation policy choices, in the hope that our discussions and findings will help improve aviation policy responses elsewhere. While the domestic market in mainland China has enjoyed a quick recovery to about 80% of the pre-crisis level by July 2020, the recovery of international services has been much slower, due to the bilateral route and flight frequency/capacity control and strict requirements for health check and quarantine. China's domestic aviation market was recovered by about 80% in two months after the pandemic became under good control. Most other countries with a “curve flattening” strategy, instead of full pandemic control, may not expect the fast recovery path China has achieved. A British “travel corridor” approach may be more practical for Western countries to follow, albeit more likely to be subject to serious

setbacks and disruptions. The aviation fee reductions and cost support China and many other countries have been using are helpful by reducing airlines' marginal costs, but not sufficient for carriers to return to profitability or sustainable operations. Capital injection and/or credit guarantee may be needed for many airlines to survive. With various, often uncoordinated, regulations imposed in international markets, airlines based in open economies that have small domestic markets will face particularly serious challenges during the recovery process.

Sustainability Reporting In The Airline Industry: Current Literature And Future Research Avenues(2022):

Authors: Malgorzata Zieba and Eljas Johansson

Researchers in this article cited that, sustainability reporting (SR) allows organizations to communicate their non-financial impacts to stakeholders. It has also become a widespread business practice in aviation, a transport sector that contributes significantly to global warming. Academia has begun to examine SR in the context of airlines surprisingly late, and no comprehensive reviews of its respective developments have been made so far. Consequently, a systematic literature review was performed with an exclusive focus on airline SR to synthesize its associated scholarly research and distinguish the common concerns and gaps that have emerged from it. The analyzed publications indicate that the industry has lacked a unified policy and common understanding of how to define and measure sustainability, which has led to inconsistent SR practices.

This causes ambiguity between the real actions and promotional communication through which airlines may legitimize their operations. Academia and various airline stakeholders would benefit from more in-depth studies examining the stakeholder views and quality of disclosures,

helping the industry improve its SR.

Changes In Air Passenger Demand As A Result Of The COVID-19 Crisis: Using Big Data To Inform Tourism Policy(2021):

Author:Galle Immaculadago

The researcher in this article cited that, this paper develops a methodology for the early detection of reactivation of tourist markets to help mitigate the effects of the COVID-19 crisis, using Skyscanner data on air passenger searches (>5,000 million) and picks (>600 million), for flights between November 2018 and December 2020, through Forward Keys. For future travel during the May to September 2020 period, the desire to travel (based on the number of flight searches) has dropped by about 30% in Europe and the Americas, and by about 50% in Asia, while intention to travel (the number of flight picks, the final selections amongst flight searches) has dropped a further 10–20%. Most source markets remain optimistic about air travel during the last quarter of 2020, suggesting a U shape recovery. However, optimism has dwindled as time passes, suggesting a flatline L shape. A traffic light dashboard for domestic and inbound air travel demand to Spain shows how destination managers might use Big Datarelating to the early recovery of key source markets to develop targeted marketing strategies. We show how Big Data provides timelygranular data essential in highly volatile situations, and we argue that destination management organizations must improve their Big Data analytical and evidence-based, decision-making skills.

The Impact Of Social Media And Offline Influences On Consumer Behaviour: An Analysis Of The Low-Cost Airline Industry(2018):

Author:Carla Ruiz

The researcher in this article cited that, this study analyses the impact of social media as well as offline environments upon tourist online purchase and recommendation behaviour of low-cost airline services.Drawing on the Theory of Reasoned Action (TRA), this research considers the effect of offline social influences (interpersonal and external influences) and analyses online Consumer-to-Consumer (C2C) information exchanges as a driver of customer attitude towardsonline purchases. We propose that these factors improve online repurchase intentions and positive word-of-mouth communication (WOM and e-WOM) in low-cost settings. Using structural equation modelling, the conceptual model is tested with a sample of 441 Spanish Internet buyers of low-cost airline services. Interpersonal offline influences (e.g. friends, relatives, and family) have a significant effect on online repurchase intentions and WOM but do not affect e-WOM. External offline influences (e.g. media and experts), however, only affect consumer intentions to recommend future purchases of low-cost airline services on social networking travel sites and have no effect on online repurchase intentions or WOM.

Assessing Quality Of Air Transport Service: A Comparative Analysis Of Two Evaluation Models (2021)

Author:Denise Dumiko De Medeiros

The researcher in this article cited that, this paper aims to analyze the opinion of tourists about airlines' service in a developing country. For this, the study proposes to make a comparative analysis of two evaluation models (SERVQUAL and SERVPERF) to investigate the factors that influence the formation of perceived quality in airline services, using statistical techniques such as Cluster Analysis and Structural Equation Modeling. Although the results were not the same, the result of both analyzes indicated two common dimensions (tangibles and empathy) that influence the customer's perception of the airline service quality. The main conclusion of this study is that the two analyzes are convergent for the study sample. The SERVQUAL and cluster analysis allow airline managers to identify and prioritize gaps in service delivery according to criticality, aiming at the allocation of efficient resources by the airline. The SERVPERF and SEM provide statistical evidence of the impact of different dimensions of service quality on customer satisfaction, highlighting the direct relationship between satisfaction and dimensions.

Considering how customers evaluate the service provided by airlines, particularly regarding the service they receive from airport employees, this study has relevance for decisions taken by airline managers to develop quality services, and provide guidelines for improvements in airline services.

Domestic Code Sharing, Alliances, And Airfares In The U.S. Airline Industry(2017):

Author:Harumi Ito,Darmin Lee

The researcher in this article cited that,this paper examines the impact of domestic code-sharing alliances on airfares. Our analysis yieldstwo novel and somewhat surprising findings that have yet to be documented in the literature. First, unlike with international code sharing, we find that the overwhelming majority of domestic code- share itineraries involve a single operating carrier, a phenomenon that we refer to as virtual code sharing. Second, we find that these virtual code-sharing itineraries are priced lower than itineraries operated and marketed by a single carrier in the same market. We suggest that carriers may be using virtual code sharing—in large part—as a generic product to compete for the most price-sensitive passengers.

Customer segmentation revisited: The caseof the airline industry (2018):

Author:Iwanvon Wartburg

The researcher in this article cited that,although the application of segmentation is a topic of central importance in marketing literature and practice, managers tend to rely on intuition and on traditional segmentation techniques based on socio-demographic variables. In the airline industry, it is regarded as common sense to separate between business and economy passengers. However, the simplicity of this segmentation logic no longer matches the ever more complex and heterogeneous choices made by customers. Airline companies relying solely on flight class as the

segmentation criterion may not be able to customize their product offerings and marketing policies to an appropriate degree in order to respond to the shifting importance and growing complexity of customer choice drivers, e.g. flexibility and price as a result of liberalization in the airline industry. Thus, there is a need to re-evaluate the traditional market segmentation criterion.

By analyzing the stated preference data of more than 5800 airline passengers, we show that segmenting into business and leisure (a) does not sufficiently capture the preference heterogeneity among customers and (b) leads to a misunderstanding of consumer preferences. We apply latent class modeling to our data and propose an alternative segmentation approach: we profile the identified segments along behavioral and socio-demographic variables. We combine our findings with observable consumer characteristics to derive pronounced fencing mechanisms for isolating and addressing customer segments receptive for tailored product packages.

A Hierarchical Model Of Service Quality In The Airline Industry(2015):

Author:Ching Chang-Cheng

The researcher in this article cited that,the purpose of this study is to enhance understanding of service quality in the airline industry by developing a conceptual framework and measurement scale. Based on an extensive literature review, qualitative and empirical research, a hierarchical model of service quality for the airline industry is proposed. Analysis of data from 544 passengers indicates that the proposed model fits the data well. Reliability and validity of the measurement scale are established using a

pilot test and the substantive survey. This study extends the literature on service quality in the fields of transportation management by providing a comprehensive framework and measurement scale. Theoretical and managerial implications are discussed.

Applied Cognitive Task Analysis in Aviation(2017):

Authors: Thomas L. Seamster and Richard E. Redding

Researchers in this article cited that, due to the requirements of automatic system design, and new needs for the training of complex tasks, Cognitive Task Analysis (CTA) has been used with increasing frequency in recent years by the airline industry and air traffic control community. Its power is reflected in the literature on professional training and systems design, where CTA is often cited as one of the most promising new technologies, especially for the complex cognitive tasks now confronting those working in aviation. The objective of this book is to bridge the gap between research and practice, to make what we know about CTA available to practitioners in the field. The book focuses on cognitive psychology and artificial intelligence analyses of aviation tasks. It is designed to help readers identify and solve specific design and training problems, in the flight deck, air traffic control and operations contexts. Distilling experience and guidelines from the best aviation cognitive analyses in accessible form, it is the first comprehensive volume on CTA, and is written for practitioners of cognitive analysis in aviation. It provides an overview of analyses to date; methods of data collection; and recommendations for designing and conducting CTA for use in instructional design, systems development, and

evaluation. The first part of the book provides the principles and foundations of CTA, describing traditional approaches to task analysis and ways that cognitive analyses can be integrated with the analysis and development processes. The next part details how to: select the appropriate method or methods; determine job tasks that can be trained for automatic performance; extract knowledge structures; analyse mental models; and identify the decision-making and problem-solving strategies associated with experienced job

.performance. The authors also describe when to use and how to design and conduct a cognitive task analysis; how to use CTA along with traditional task analysis and ISD; and how to use CTA in training program development and systems design, as well as in personnel selection and evaluation. The current demand for cognitive analyses makes this a timely volume for those in aviation and, more generally, the industrial development and training communities.

Readers will find this a thorough presentation of cognitive analyses in aviation and a highly usable guide in the design, implementation and interpretation of CTA. The book will be useful to instructional developers, aviation equipment and systems designers, researchers, government regulatory personnel, human resource managers, instructors, pilots, air traffic controllers, and operations staff.

References:

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

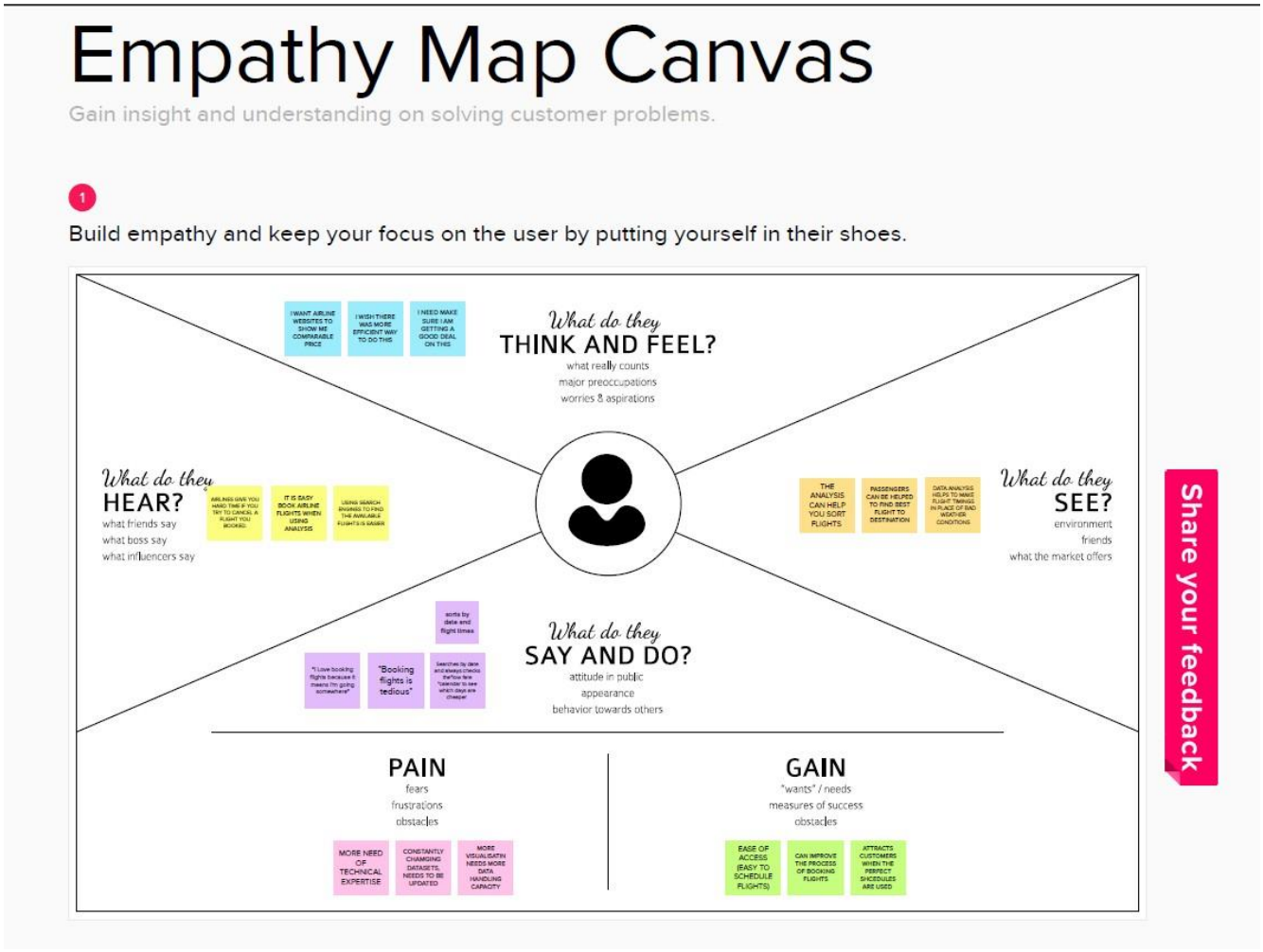
Problem Statement Definition:

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.

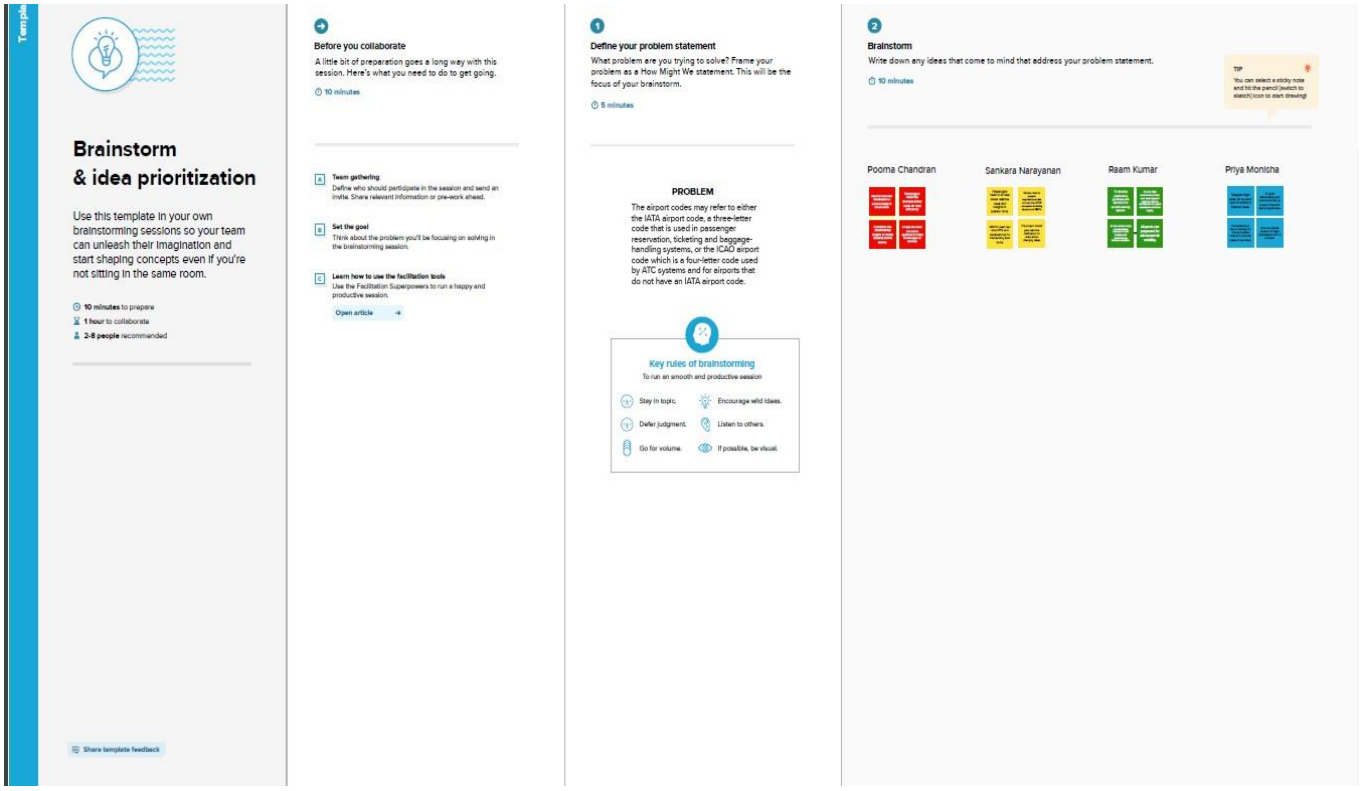
Who does the problem affect?	Many people will get frustrated to sit idlein airport. Passengers who need to handle international meeting and also for some emergency purpose.
What are the boundaries of the problem?	Adverse weather conditions, low ceilings, and low visibility conditionsstrongly influence flight delays.
What is the issue?	Inclement weather, such as thunderstorm, hurricane, or blizzard. Late arrival of theaircraft to be used for the flight from a previous flight. Maintenance problems with the aircraft.
When does the issue occurs?	A flight delay is when an airline flight takes off and /or lands later than its scheduled time. A cancellation occurs when the airline does not operate the flightat all for a certain reason.
Where is the issue occurring?	The least punctual airline of all U.S. carriers was Frontier which had an average on-time arrival of 73.14 percent. Coming injust above that was JetBlue with an on-timearrival of 73.5 percent.

Why is it important that we fix theproblem?	Passengers are increasingly affected bydelays and cancellations. Thus these problems has to be fixed

IDEATION



Ideation & Brainstorming :

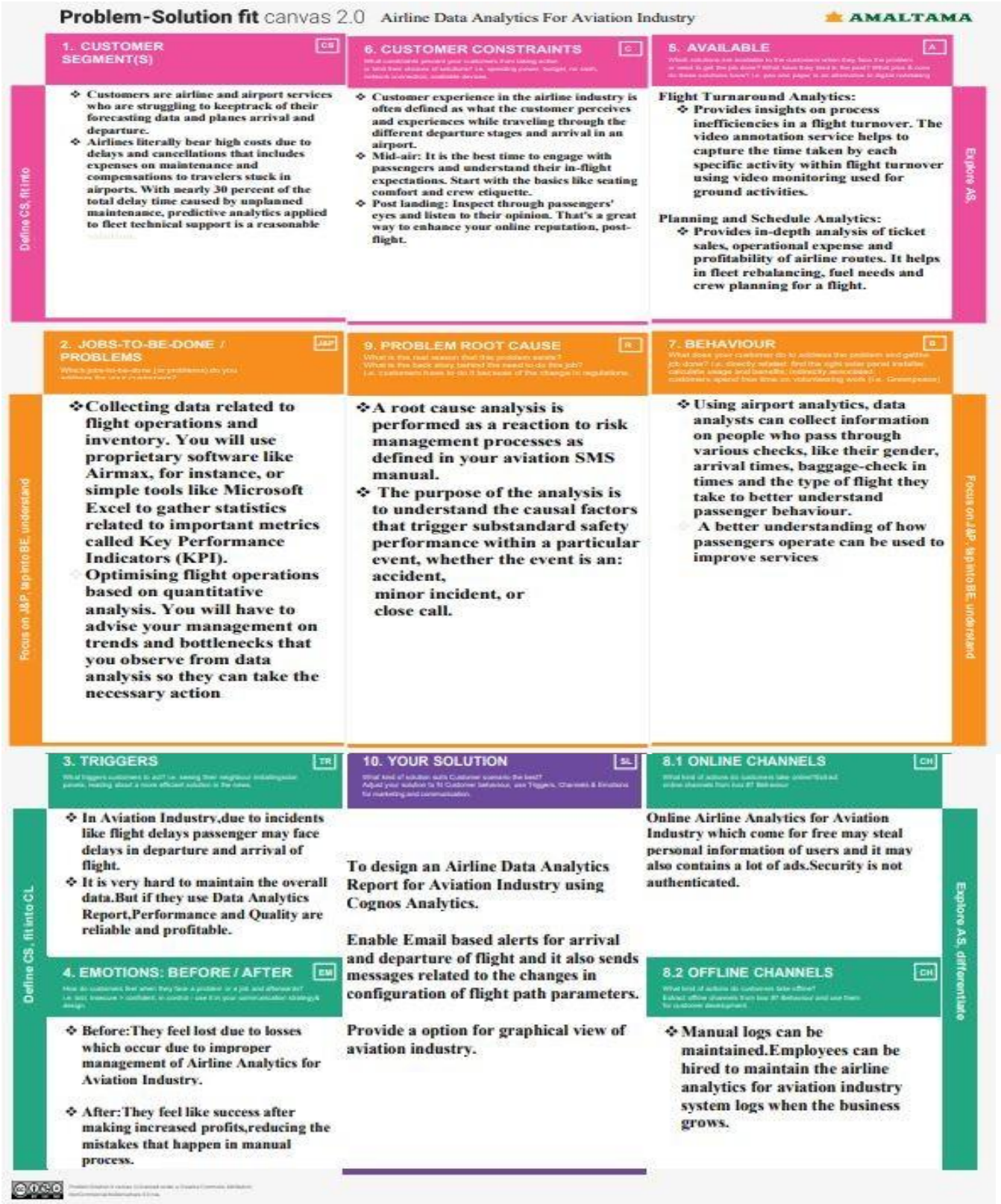


S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	❖ With the growing demand for air transportation and the limited ability to increase capacity at some key points in the air transportation system, there are concerns that in the future the system will not scale to meet demand. This situation will result in the generation and the propagation of delays throughout the system, impacting passengers' quality of travel and more broadly the economy.
2.	Idea / Solution description	❖ Understanding traveler demand for specific city pairs and pricing flights can be done using data analytics project. ❖ 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.
3.	Novelty / Uniqueness	❖ The ultimate benefits of big data analytics include timely responses to current and future market demands, improved planning and strategically aligned decision making, as

		well as crystal clear comprehension and monitoring of all main performance drivers relevant to the airline industry. ❖ 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.
4.	Social Impact / Customer Satisfaction	❖ 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 real-time using predictive analysis techniques. As a result, by gathering meaningful data, airlines can fetch more bookings in the given timeframe.
5.	Business Model (Revenue Model)	❖ Business models innovation in airlines can contribute to the creation of value, competitive advantage and profitability with new possibilities of action. ❖ A revenue model is a

		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.
6.	Scalability of the Solution	❖ 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.

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 Registeration through Gmail
FR-2	User Confirmation	After the Registeration 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:

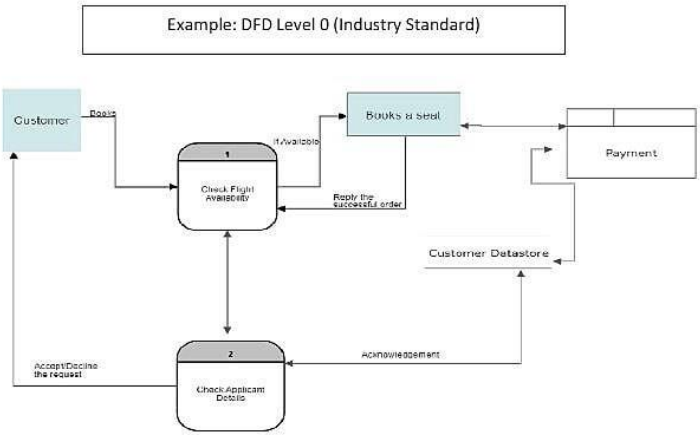
Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	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	**Security**	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	**Reliability**	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	**Performance**	The system should require a fair amount of speed especially while browsing through the catalogue.
NFR-5	**Availability**	The system shall be available 24 hours a day 7 days a week. User can access at anytime.
NFR-6	**Scalability**	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.

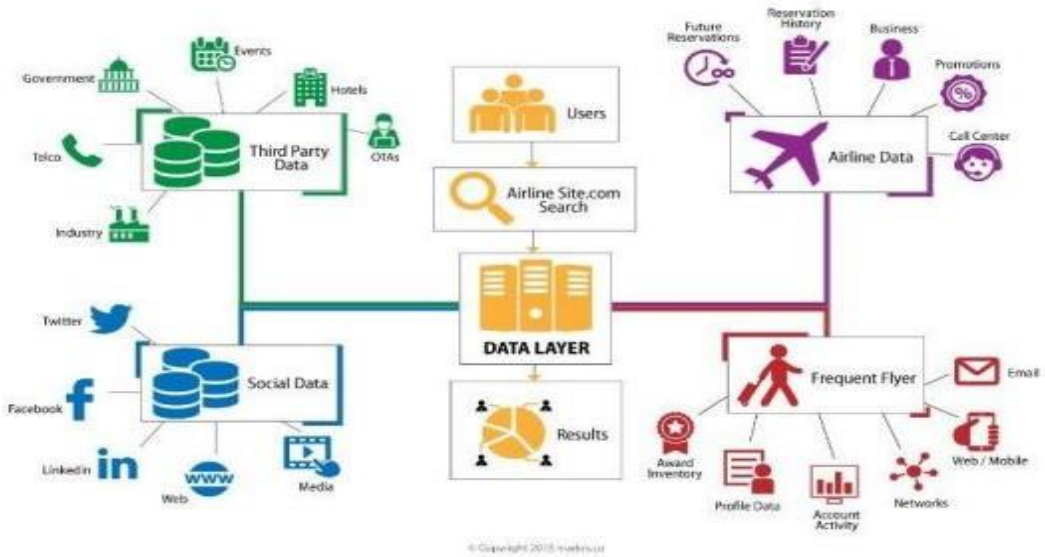


SOLUTION & TECHNICAL ARCHITECTURE:

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 • they have any problem with the organization's system of customer interaction or • airplane issues- delay, landing in a different location	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: • registration • flight booking • delay visualization • generation of delay report	As an administrator, confirmation of user while registration is done.	High	Sprint-1

PROJECT PLANNING & 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)

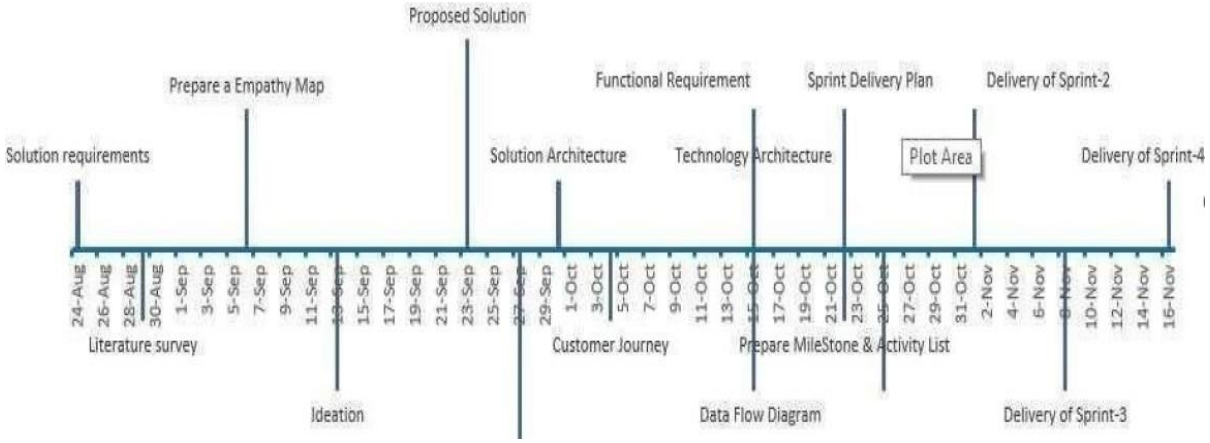
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	12	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	12	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022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

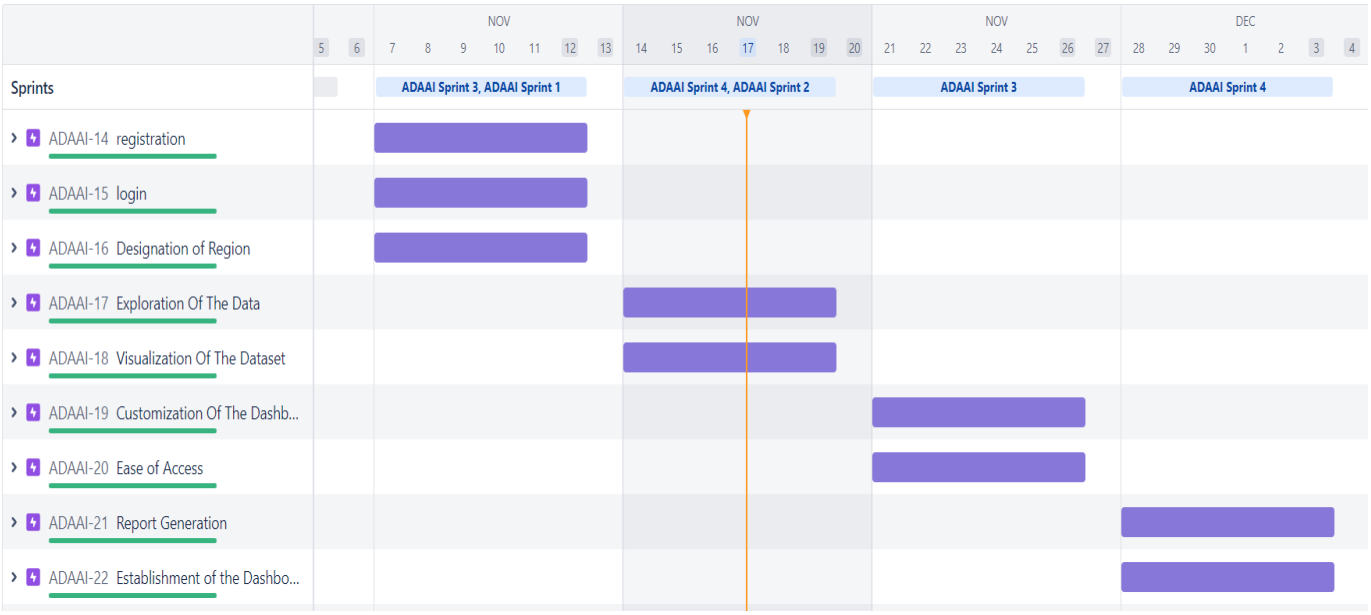
Average velocity=Sprint duration / velocity=12/6=2

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. Eachmilestone 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:

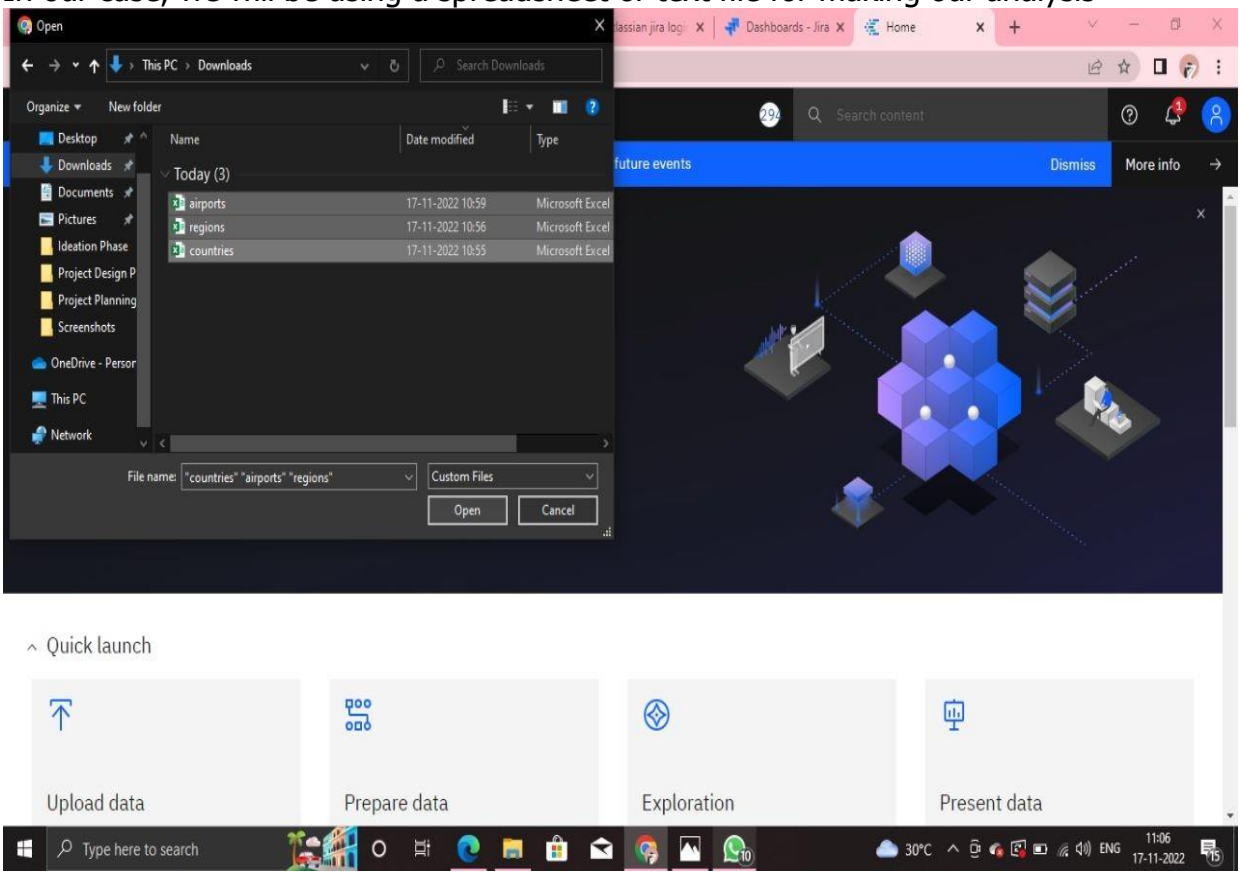
https://www.kaggle.com/patrasaurabh/airstats-data-on-airports-around-the-world

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.

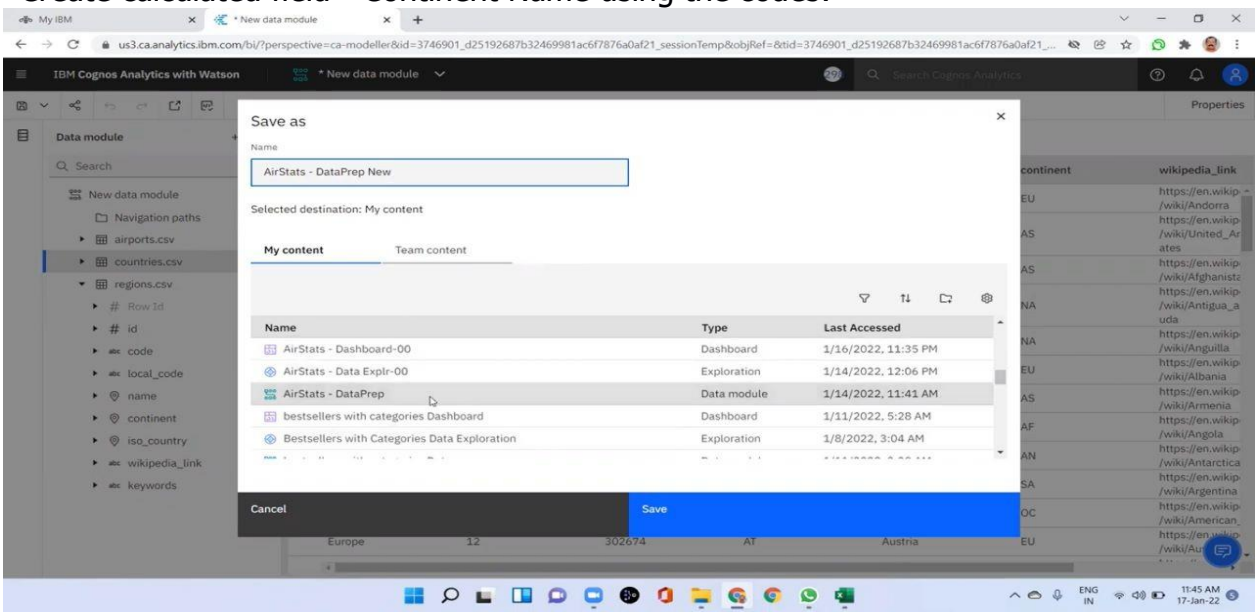
In our case, we will be using a spreadsheet or text file for making our analysis



DATA PREPARATION:

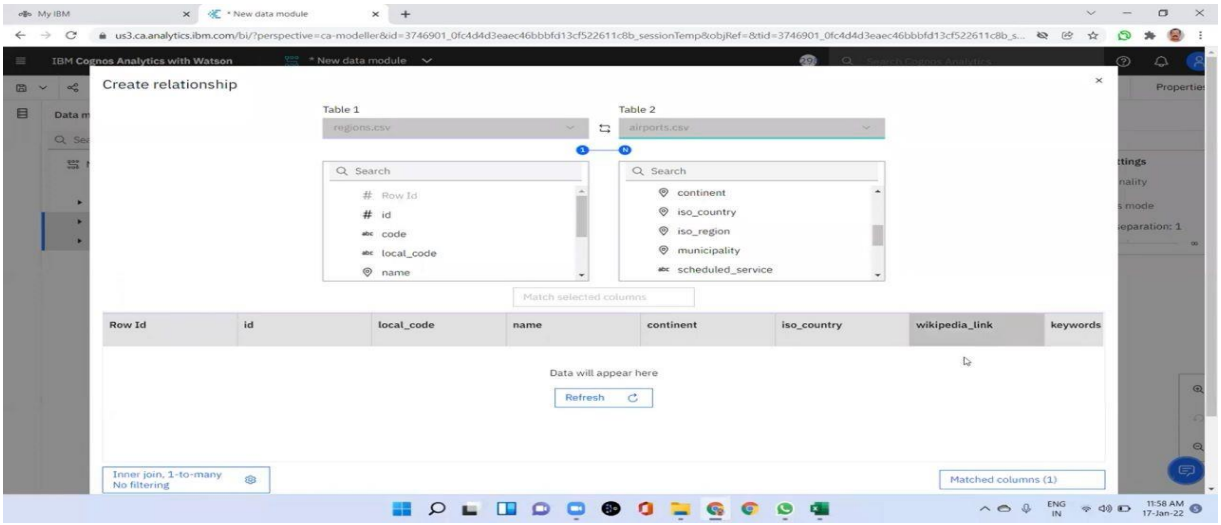
Data Preparation.

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



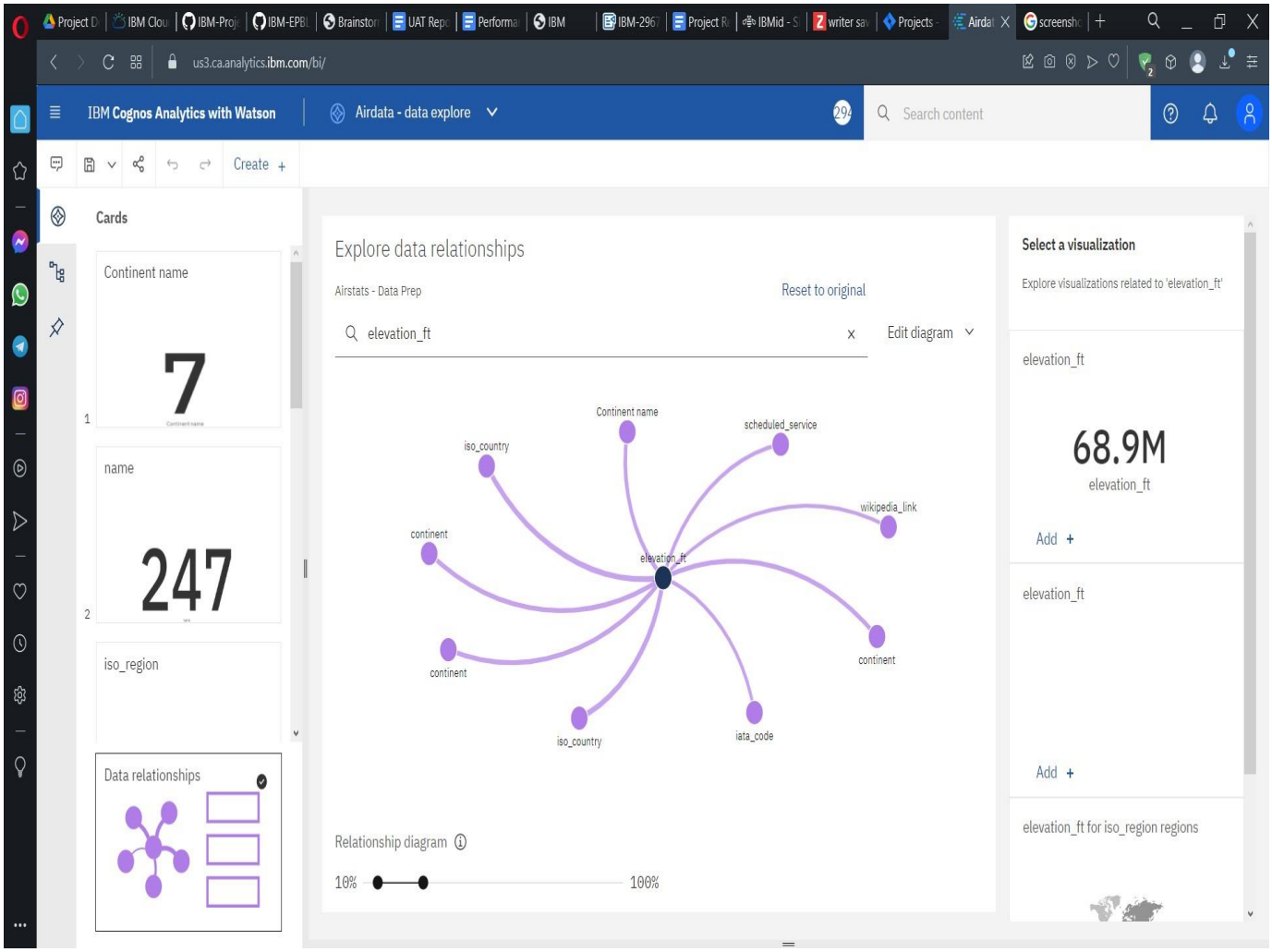
Joining Of Tables :

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 tosurface 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.



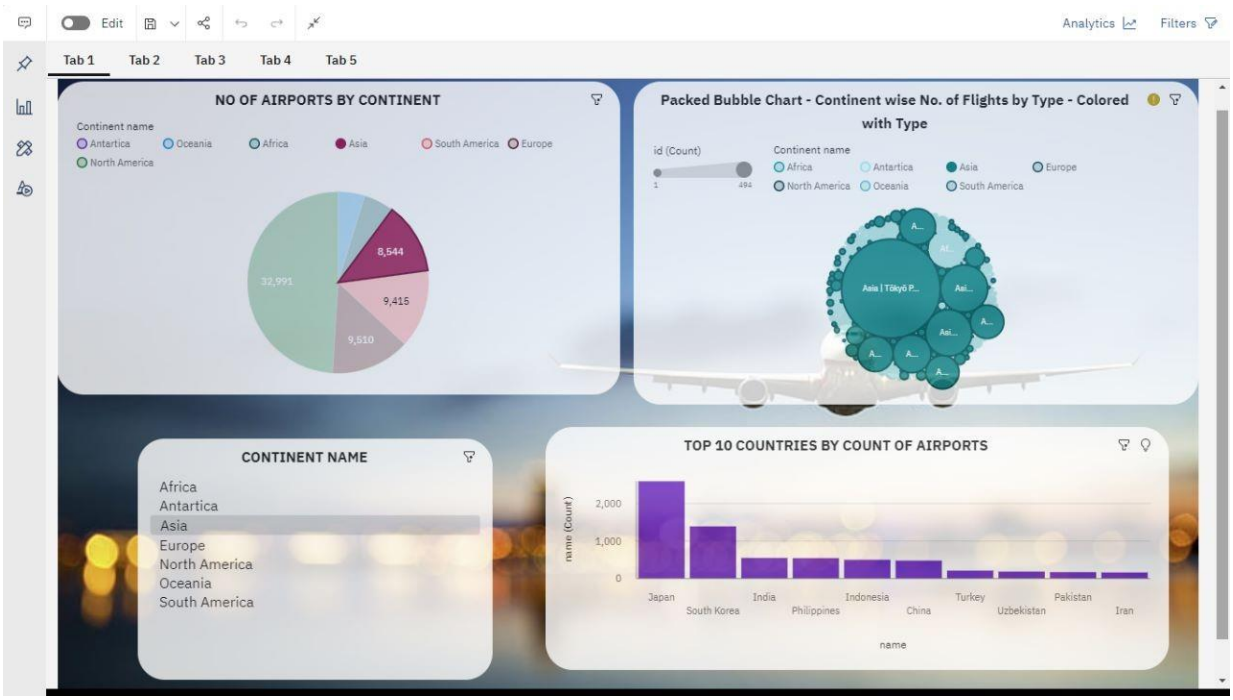
DATA VISUALIZATION :

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

Representation Of Flight Count By Categories :

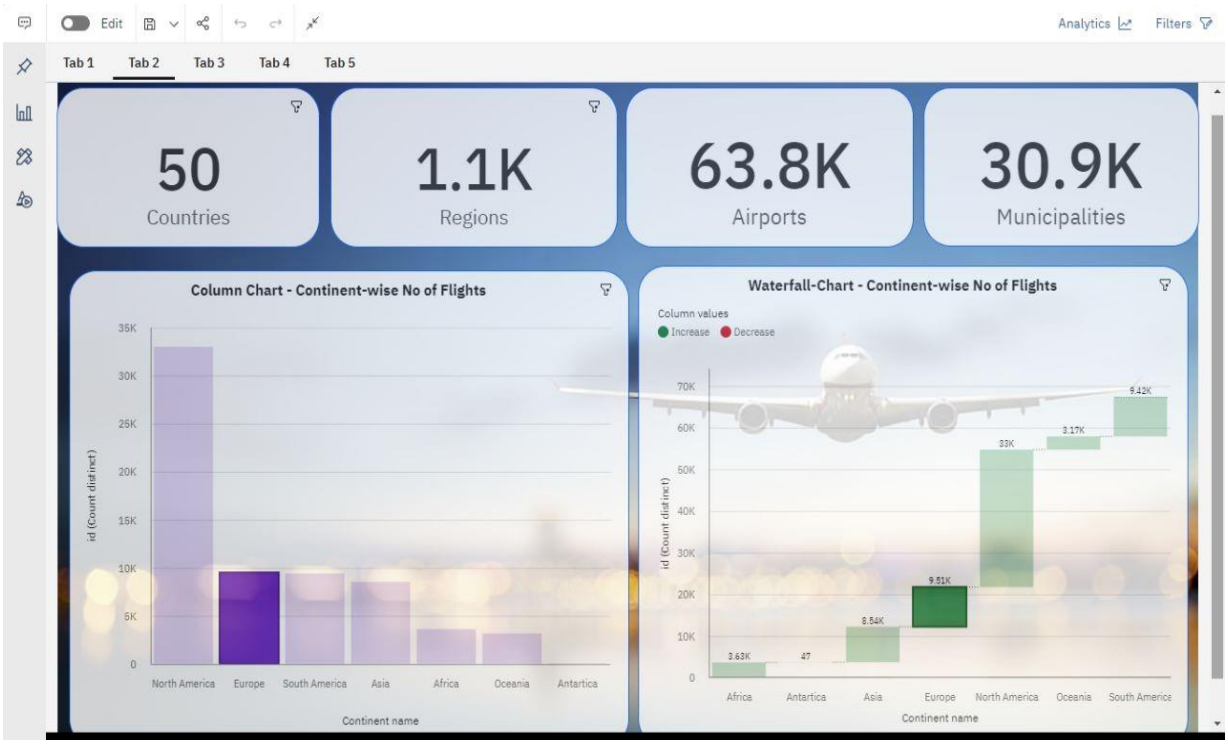
Representation of Flight Count by Categories.

- Pie Chart - Continent-wise No. of Flights.
- Packed Bubble Chart - Continent wise No. of Flights by Type - Colored with Type.
- Continent List - Filter.
- Top 10 Countries by Flights.



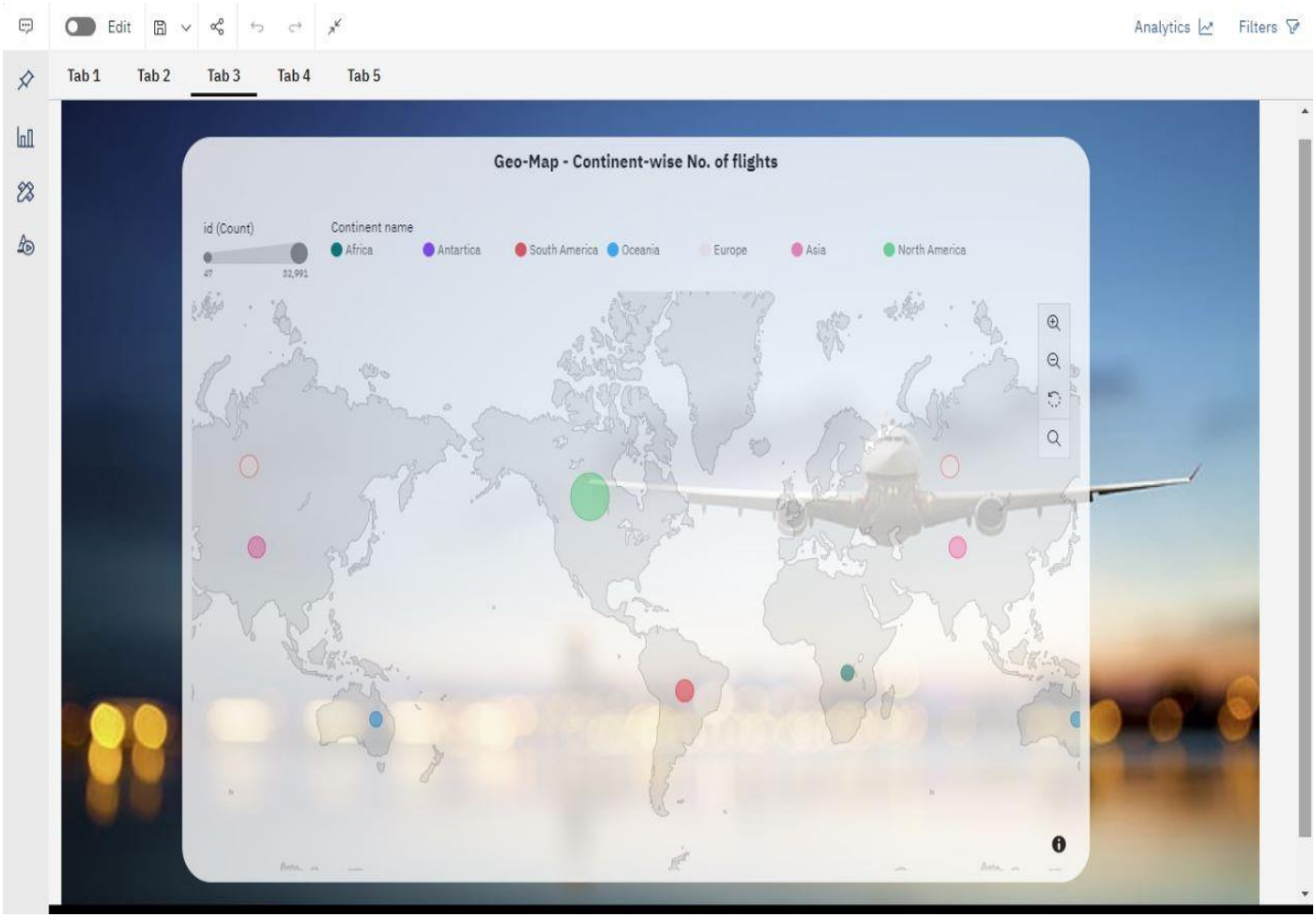
No Of Flights by Countries, Regions And Airports:

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.



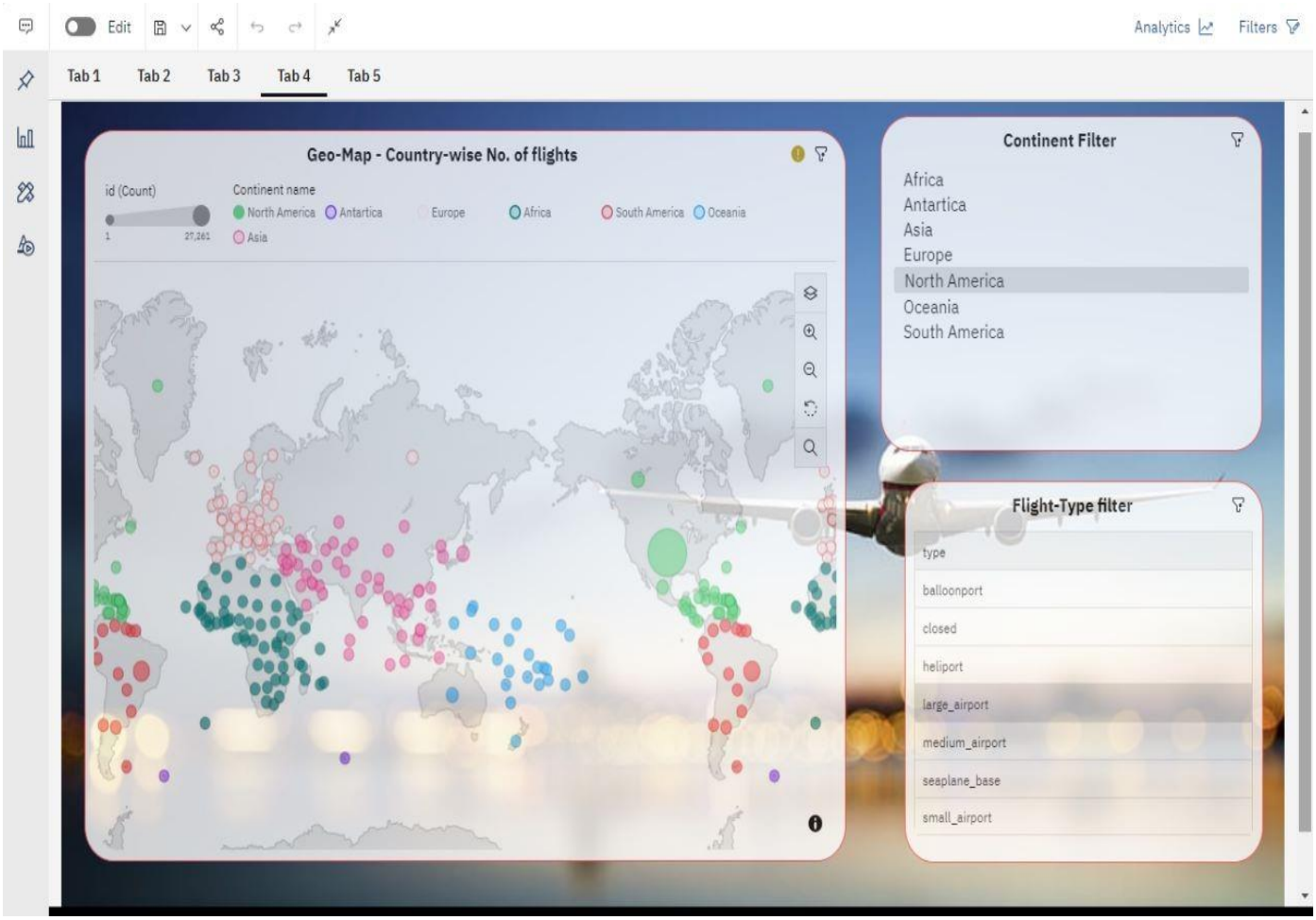
Continent Wise Count Of Airports Using Geo Map:

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



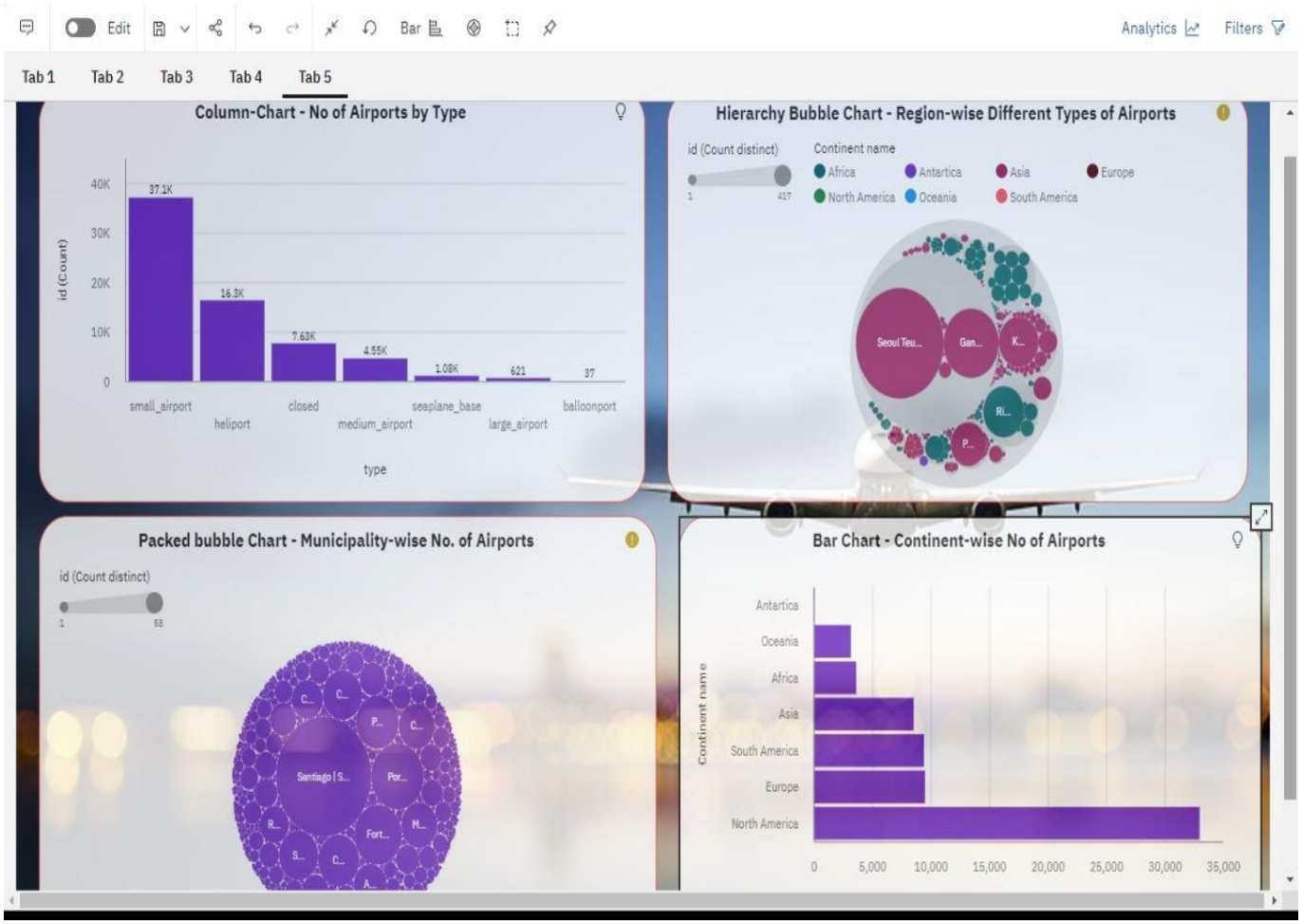
Country Wise Airports With Types :

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/Signup popup should display	Working as expected	Pass
LoginPage_TC_OO2	UI	dashboard page	verify user is able to see airport report in dashboard 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 visualisation will be displayed on the dashboard	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	Fail	Pass
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:

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:

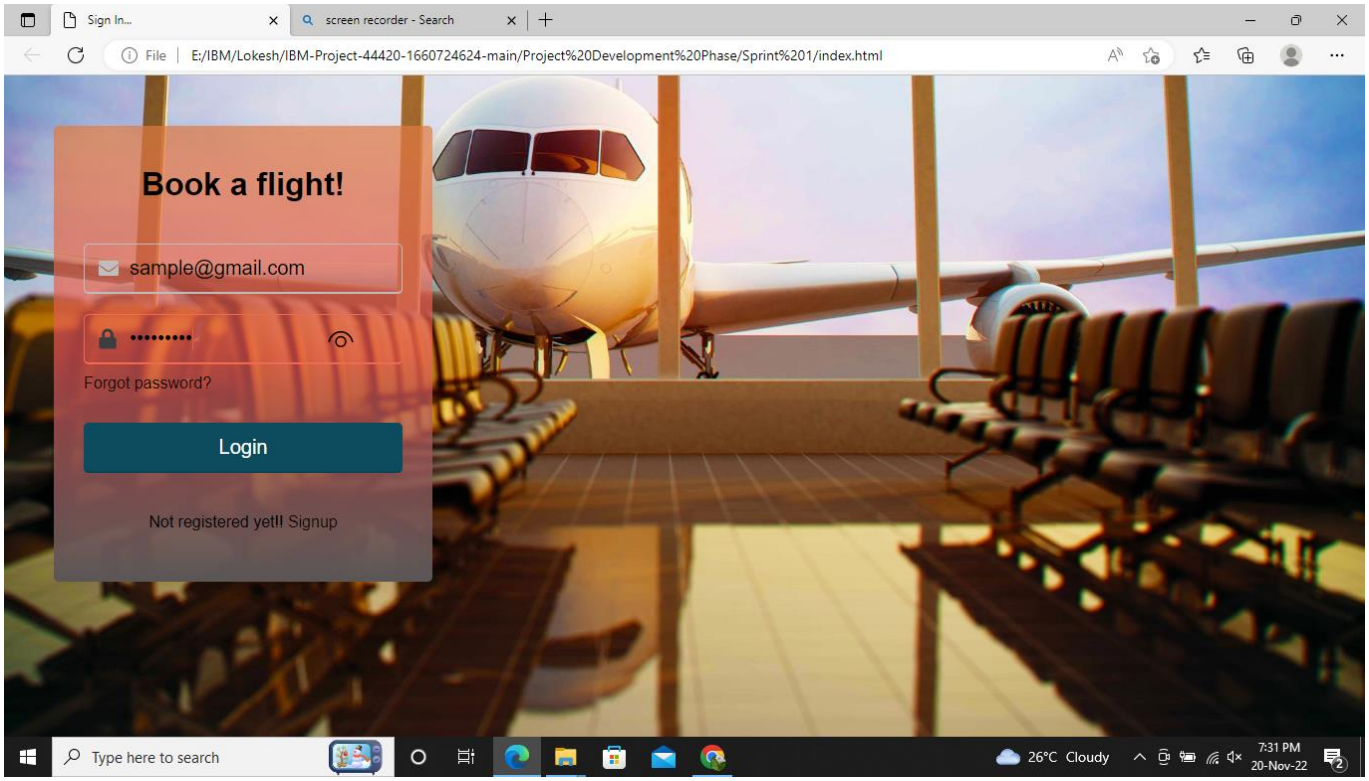
- It improves the average turnaround time needed to cater to market trends
- Properly implemented data modules help flight operators bag more customers and profits
- Predictive analytics is the key to preparing for future crises and put a mitigation plan in place
- It helps businesses make data-backed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

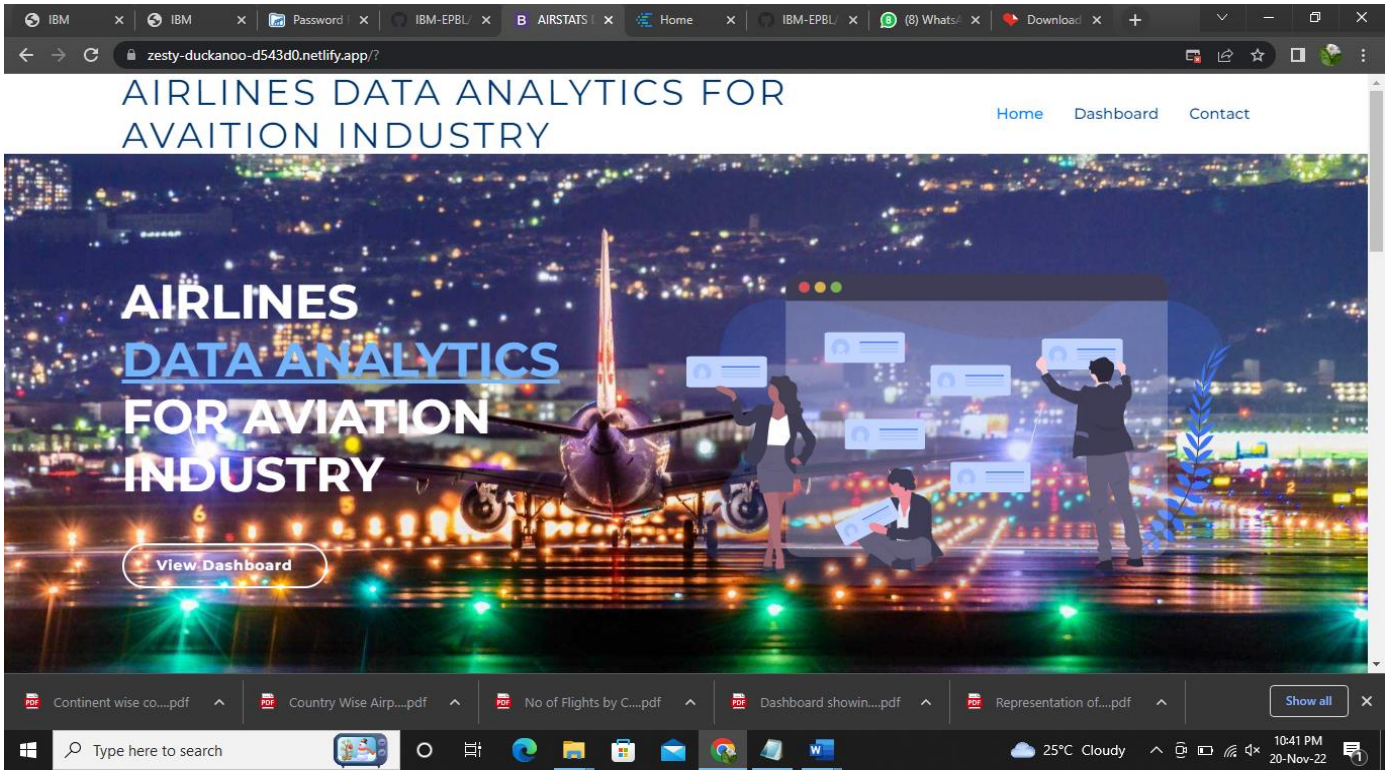
Disadvantages:

- Air transport is a costly service. Its operational costs are too high. Middle class and poor people cannot affect 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.

PROJECT DEMONSTRATION:

Click the Link: https://capnpeace.github.io/Airlines-Data-Analytics-for-Avaition-Industry/
Then require to login with mail id and password to visit our dashboard representation.





Click Dashboard to view the data representation of Continent wise airports with help of id, flight code.

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>Login Form</title>
    <link rel="stylesheet" href="style.css">
    <link rel="stylesheet" href="C:\Users\PC\OneDrive\Desktop\style.css"
    />
</head>

<body>
    <div class="wrapper">
```

```
        <header>Login Form</header>
        <form   action="https://zesty-duckanoo-d543d0.netlify.app/">
            <div class="field email">
                <div class="input-area">
                    <input type="text" placeholder="Email Address">
                    <i class="icon fas fa-envelope"></i>
                    <i class="error error-icon fas fa-exclamation-
circle"></i
>
                </div>
                <div class="error error-txt">Email can't be blank</div>
            </div>
            <div class="field password">
                <div class="input-area">
                    <input type="password" placeholder="Password">
                    <i class="icon fas fa-lock"></i>
                    <i class="error error-icon fas fa-exclamation-
circle"></i
>
                </div>
                <div class="error error-txt">Password can't be blank</div>
            </div>
            <div class="pass-txt"><a href="#">Forgot password?</a></div>

            <input type="submit" value="Login">
        </form>
        <div class="sign-txt">Not yet member? <a href="#">Signup
now</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 content="width=device-width, initial-scale=1.0" name="viewport">

  <title>AIRSTATS DASHBOARD</title>
  <meta content="" name="description">
  <meta content="" name="keywords">

  <!-- Favicons -->
  <link href="assets/img/favicon.png" rel="icon">
  <link href="assets/img/apple-touch-icon.png" rel="apple-touch-icon">

  <!-- Google Fonts -->
  <link
href="https://fonts.googleapis.com/css?family=Open+Sans:300,300i,400,400i,600
,600i,700,700i|Montserrat:300,400,500,700"   rel="stylesheet">

  <!-- Vendor CSS Files -->
  <link href="assets/vendor/aos/aos.css" rel="stylesheet">
  <link href="assets/vendor/bootstrap/css/bootstrap.min.css"
rel="stylesheet">
  <link   href="assets/vendor/bootstrap-icons/bootstrap-icons.css"
rel="stylesheet">
  <link   href="assets/vendor/glightbox/css/glightbox.min.css"
```

```
rel="stylesheet">
  <link href="assets/vendor/swiper/swiper-bundle.min.css" rel="stylesheet">

  <!-- Template Main CSS File -->
  <link href="assets/css/style.css" rel="stylesheet">

  <!--   =======================================================
  * Template Name: NewBiz - v4.9.1
  * Template URL: https://bootstrapmade.com/newbiz-bootstrap-business-
template/
  * Author: BootstrapMade.com
  * License: https://bootstrapmade.com/license/
  =======================================================   -->
</head>

<body>

  <!-- ======= Header ======= -->
  <header id="header" class="fixed-top d-flex align-items-center">
    <div class="container d-flex justify-content-between">

      <div class="logo">
        <!-- Uncomment below if you prefer to use an text logo -->
        <h1><a href="index.html">Airlines Data Analytics for Aviation
Industry</a></h1>

      </div>

      <nav id="navbar" class="navbar">
        <ul>
          <li><a class="nav-link scrollto active" href="#hero">Home</a></li>

          <li><a class="nav-link scrollto"
href="#services">Dashboard</a></li>

          <li><a class="nav-link scrollto" href="#contact">Contact</a></li>
        </ul>
        <i class="bi bi-list mobile-nav-toggle"></i>
      </nav><!-- .navbar -->

    </div>
```

```
  </header><!-- #header -->

 <!-- ======= Hero Section ======= -->
 <section id="hero" class="clearfix">
  <div class="container" data-aos="fade-up">

    <div class="hero-img" data-aos="zoom-out" data-aos-delay="200">
      <img src="assets/img/hero-img.svg" alt="" class="img-fluid">
    </div>

    <div class="hero-info" data-aos="zoom-in" data-aos-delay="100">
      <h2>AIRLINES<br><span>DATA ANALYTICS</span><br>FOR AVIATION
INDUSTRY</h2>
      <div>
        <a href="#services" class="btn-services scrollto">View Dashboard</a>
      </div>
    </div>

  </div>
</section><!-- End Hero Section -->

  <main id="main">



    <!-- ======= Services Section ======= -->
    <section id="services" class="section-bg">
      <div class="container" data-aos="fade-up">
        <header class="section-header">
          <h3>AIRSTATS ANALYSIS DASHBOARD</h3>
          <iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;pathRef=.
my_folders%2FAIR%2BSTATS%2BDASHBOARD&amp;closeWindowOnLastView=true&amp;ui_ap
pbar=false&amp;ui_navbar=false&amp;shareMode=embedded&amp;action=view&amp;mod
e=dashboard&amp;subView=model0000018447f5966e_00000002" width="1300"
height="1000" frameborder="0" gesture="media" allow="encrypted-media"
allowfullscreen="">
          </iframe>
        </header>
      </div>
     </section><!-- End Services Section -->
```

```
    <!-- ======= Contact Section ======= -->
    <section id="contact">
      <div class="container-fluid" data-aos="fade-up">
        <div class="section-header">
          <h3>Contact Us</h3>
        </div>
        <div class="row">
          <div class="col-lg-6">
            <div class="row">
              <div class="col-md-5 info">
                <i class="bi bi-geo-alt"></i>
                <p>GCE TLY</p>
              </div>
              <div class="col-md-4 info">
                <i class="bi bi-envelope"></i>
                <p>https://github.com/capnpeace.com</p>
              </div>
          </div>

        </div>

      </div>
    </section><!-- End Contact Section -->

  </main>
  <!-- End #main -->



  <a href="#" class="back-to-top d-flex align-items-center justify-content-
center"><i class="bi bi-arrow-up-short"></i></a>

  <!-- Vendor JS Files -->
  <script   src="assets/vendor/purecounter/purecounter_vanilla.js"></script>
  <script src="assets/vendor/aos/aos.js"></script>
  <script   src="assets/vendor/bootstrap/js/bootstrap.bundle.min.js"></script>
  <script   src="assets/vendor/glightbox/js/glightbox.min.js"></script>
  <script   src="assets/vendor/isotope-layout/isotope.pkgd.min.js"></script>
  <script   src="assets/vendor/swiper/swiper-bundle.min.js"></script>
```

```
  <script   src="assets/vendor/php-email-form/validate.js"></script>

  <!-- Template Main JS File -->
  <script src="assets/js/main.js"></script>

</body>

</html>
```

Software Requirement & Specification

- IBM Cognos
- GitHub
- JIRA
- Google Collaborator
- Google Drive
- Notepad
- MS Excel
- JavaScript
- Python
- CSS

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 analyze 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 travelers 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.