NAAN MUDHALVAN
PROJECT BASED LEARNING

MEENAKSHI SUNDARARAJAN ENGINEERING COLLEGE

Kodambakkam, Chennai-600024.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

TOPIC: DATA DOMINATORS: A COMPARATIVE STUDY OF TOP GLOBAL UNIVERSITIES IN DATA ANALYTICS

TEAM ID: NM2023TMID07074

FACULTY MENTOR: S Yamuna

INDUSTRY MENTOR: Shivam shivhare

Project submitted by,

TEAM	NAME	REG. NO.
Team Leader	**Jagathguru B**	**311520104021**
Team mate 2	**Madhavan R**	**311520104027**
Team mate 3	**Manvith B V**	**311520104031**
Team mate 4	**Yugabharathi R**	**311520104059**

Project Report Format

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
4. **REQUIREMENT ANALYSIS**
   - 4.1 Functional requirement
   - 4.2 Non-Functional requirements
5. **PROJECT DESIGN**
   - 5.1 Data Flow Diagrams & User Stories
   - 5.2 Solution Architecture
6. **PROJECT PLANNING & SCHEDULING**
   - 6.1 Technical Architecture
   - 6.2 Sprint Planning & Estimation
   - 6.3 Sprint Delivery Schedule
7. **CODING & SOLUTION (Explain the features added in the project along with code)**
   - 7.1 Feature 1
   - 7.2 Feature 2
   - 7.3 Database Schema (if Applicable)
8. **TESTING**
   - 8.1 Performance Metrics
9. **RESULTS**
   - 9.1 Output Screenshots
10. **ADVANTAGES & DISADVANTAGES**
11. **CONCLUSION**
12. **FUTURE SCOPE**
13. **APPENDIX**

    Source Code

    GitHub & Project Demo Link

1. INTRODUCTION

1.1 Project Overview

The "Data Dominators: A Comparative Study Of Top Global Universities In Data Analytics" project is a comprehensive endeavor aimed at dissecting and evaluating the offerings of leading universities in the field of data analytics. In a world increasingly reliant on data-driven decision-making, this project assumes paramount significance. It seeks to empower students and researchers by providing them with the crucial insights they need to make informed choices about their educational and research journeys.

This study recognizes the surge in demand for data analytics education and the necessity for high-quality programs in this domain. Through extensive research and analysis, it will serve as a beacon of clarity in an otherwise overwhelming sea of choices. By shedding light on the strengths and weaknesses of different data analytics programs, it empowers prospective students and researchers to tailor their academic or research pursuits according to their goals, aspirations, and unique preferences. This project ultimately aspires to bridge the information gap, fostering a more informed community of data analytics enthusiasts, and supporting the growth and excellence of this field on a global scale.

1.2 Purpose

The project is a multi-faceted initiative aimed at addressing a significant challenge while aiming to make a far-reaching positive impact in the field of data analytics education and research. At its core, this project is driven by the imperative need to provide clarity and guidance to students and researchers who find themselves navigating the complex and rapidly evolving landscape of data analytics education. In an era where data analytics plays a central role in decision-making across industries, the quality of education in this field is of paramount importance. The project's primary goal is to comprehensively evaluate and compare data analytics programs offered by top global universities and equip prospective students and researchers with the critical information they need to make well-informed decisions about their academic and research paths.

The problem that this project seeks to address is the pervasive information gap that plagues aspiring data analysts. The contemporary data-driven world is brimming with opportunities, and the demand for skilled professionals in data analytics is higher than ever before. However, this surge in demand has been met with a proliferation of data analytics programs offered by

universities worldwide, each with its unique features, strengths, and weaknesses. The paradox of choice is real, and aspiring data analysts often find themselves inundated with options, leaving them grappling with questions about program quality, relevance to their career goals, and the credibility of university offerings. This lack of clarity, this information gap, can lead to misplaced investments, career misalignment, and underutilized potential.The potential impact of this project is substantial. By addressing the information gap, it has the capacity to streamline the decision-making process for prospective data analysts and researchers. It offers them a comprehensive resource to navigate the educational landscape with confidence and clarity. In a world where the right educational choice can significantly impact one's career and future prospects, this project's relevance cannot be overstated. It doesn't merely act as a compass; it is a beacon of light in an otherwise challenging journey. The impact isn't limited to individual decision-making; it also extends to educational institutions.By providing prospective students and researchers with a comprehensive, well-researched comparative study of top global universities offering data analytics programs, this project can effectively bridge the information gap and empower them to make informed choices. This, in turn, can prompt universities to strive for educational excellence, aligning their programs with the industry's demands and standards. This positive feedback loop benefits students, researchers, and institutions, fostering a culture of continuous improvement and ensuring that the education provided is not just cutting-edge but also relevant to real-world needs.The potential impact of this project isn't limited to the immediate decision-making process. It also has ripple effects on the broader landscape of data analytics education and the industry as a whole. By equipping individuals with the knowledge they need to make informed choices, it ensures that they enter their academic and research pursuits with a clear sense of direction, reducing the risk of mismatched expectations and misplaced investments. Graduates of these well-informed programs are more likely to excel in their careers, contributing positively to the field.

In a broader context, this project contributes to the alignment of the educational sector with industry needs. As the data analytics field rapidly evolves, it's crucial that educational programs adapt to equip graduates with the skills and knowledge demanded by employers. This project's comparative study serves as a feedback mechanism for both universities and the industry. By identifying best practices, successful program features, and faculty expertise, it not only empowers individuals but also guides universities in enhancing the quality of their data analytics programs. It encourages universities to learn from one another and align their offerings with the industry's

evolving needs.It strives to empower prospective students and researchers with the knowledge they need to make informed decisions, ultimately contributing to the growth and excellence of the data analytics field. It's a project that embodies the ideals of knowledge, ambition, and empowerment, ensuring a brighter future not just for individuals but for the entire data analytics industry. It is, indeed, a beacon of clarity in a complex educational landscape.

2. LITERATURE SURVEY

2.1 Existing problem

The field of data analytics is undeniably at the forefront of the digital age, shaping how organizations make critical decisions and unravel insights from massive datasets. However, the journey to becoming a proficient data analyst is not without its complexities, and therein lies the existing problem that this project seeks to address.

One of the predominant challenges facing aspiring data analysts and researchers is the lack of comprehensive and transparent information regarding the educational programs available to them. The realm of data analytics education is vast and varied, with an array of universities offering programs, each with its unique strengths and weaknesses. Without a clear roadmap to navigate this landscape, prospective students and researchers can find themselves overwhelmed, grappling with questions about program quality, relevance to their career goals, and the credibility of university offerings. This lack of clarity often leads to an information gap that hinders decision-making.The need for this project becomes evident when we consider the sheer demand for skilled data analysts and the rapidly evolving field of data science. According to industry reports, the demand for professionals proficient in data analytics is surging, with a significant shortage of qualified candidates. As a result, there is immense pressure on educational institutions to produce graduates who are job-ready and well-versed in the latest data tools and techniques. However, the disconnect between educational offerings and industry demands is a persistent challenge.

Moreover, students investing in data analytics programs commit substantial time, effort, and financial resources. Hence, their choice of university can significantly impact their educational and career trajectories. This is where the project's purpose gains particular significance, as it aims to bridge the existing gap by providing an in-depth comparative study of top global universities offering data analytics programs. By offering comprehensive insights into program strengths, faculty expertise, resources, and graduate outcomes, this project aims to empower students and

researchers with the knowledge they need to make informed decisions about their educational and research journeys. In doing so, it not only addresses an immediate problem but also contributes to the future of data analytics education and the skilled workforce that drives innovation in this dynamic field.

2.2 References

S.No.	TITLE	AUTHOR	RELATED WORK
1	Challenges and Opportunities of Teaching Data Science in Higher Education	Kelleher, D'Arcy, and Mac Namee	This paper discusses the challenges and opportunities in teaching data science in higher education. It explores the evolving landscape of data science education.
2	Bridging the Data Science Skills Gap: A Job Task Analysis"	Anderson et al Brynjolfsson and McAfee.	The paper focuses on bridging the skills gap in data science by conducting a job task analysis, offering insights into the skills required in the industry.
3	Data Analytics in Higher Education: Key Concerns and Open Questions	Gašević et al Kelleher, D'Arcy, and Mac Namee.	This work addresses concerns and open questions related to data analytics in higher education, examining the impact of data on academia.

4	Big Data and Data Science: What Should We Teach?	Provost and Fawcett Nolan and Temple Lang.	The authors discuss the key question of what should be taught in the context of big data and data science education.
5	Data Science and Predictive Analytics: Biased or Neutral?	Hand. Diez et al.	This paper investigates the potential biases in data science and predictive analytics and their implications.
6	The Data Analytics Professional: Where We Are, Where We Need to Go	Davenport. Brynjolfsson and McAfee	The author provides insights into the current state and the future trajectory of the data analytics profession.
7	A Data Science Course for Undergraduates: Thinking With Data	Nolan and Temple Lang. Provost and Fawcett.	This work outlines a data science course tailored for undergraduate students, emphasizing critical thinking with data.
8	Data Science Education: The Essential Guide to the Data-Driven Future	Kelleher, D'Arcy, and Mac Namee Gašević et al.	The authors offer a comprehensive guide to data science education, focusing on the data-driven future of the field.
9	Competing in the Age of Data Science	Brynjolfsson and McAfee	This paper explores the competitive

		Anderson .	aspects in the age of data science and analytics, discussing the impact on businesses and society.
10	Teaching Data Science	Donoho.	The author discusses the challenges and opportunities in teaching data science, emphasizing the need for effective pedagogical approaches.

2.3 Problem Statement Definition

The heart of our project lies in the profound challenge of addressing the pervasive information gap in data analytics education. In this era of data-driven decision-making, the demand for skilled data analysts has surged, and universities worldwide have risen to the occasion by offering a plethora of data analytics programs. Yet, while this abundance of choice is a testament to the field's vitality, it has also given rise to a substantial problem.Prospective data analytics students and researchers are faced with the daunting task of selecting the right university and program to match their unique aspirations and career objectives. This is where the problem becomes apparent. The lack of a comprehensive and transparent source of information has left many in a state of uncertainty. Critical decisions regarding one's educational journey, which may significantly influence their career path, are often made without a clear understanding of the options available. This can lead to misplaced investments, career misalignment, and underutilized potential.

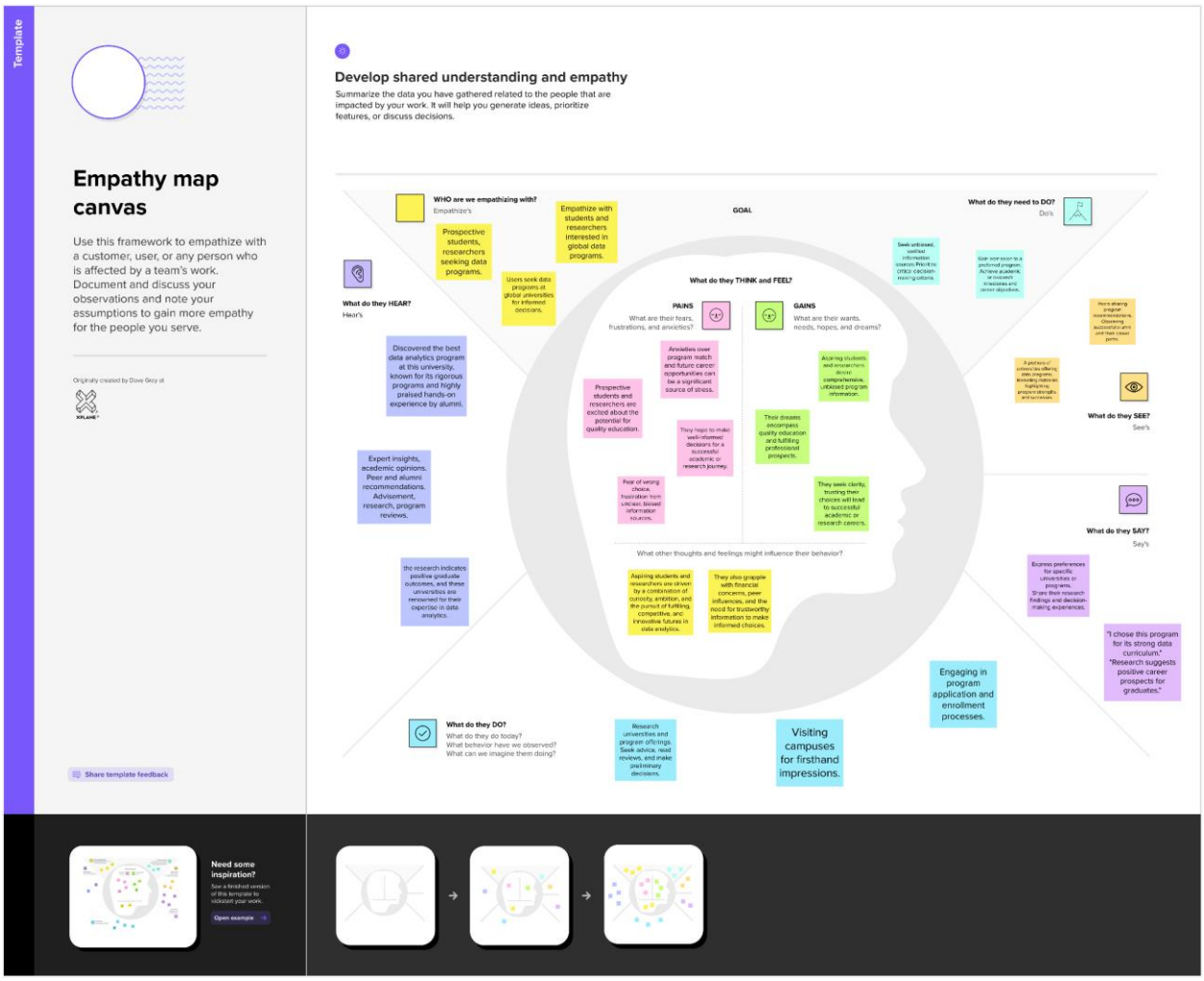
Our research sets out to address this problem by focusing on top global universities that offer data analytics programs. We intend to undertake a rigorous comparative analysis of these programs, delving into the intricacies of curriculum quality, faculty expertise, available resources, and graduate outcomes. Our scope extends beyond academic institutions; it encompasses the needs and

perspectives of the individuals at the heart of this challenge – the prospective data analytics students and researchers.Key questions that guide our research include understanding the critical factors that individuals weigh when evaluating data analytics programs, differentiating universities in terms of program offerings and faculty expertise, and exploring the long-term impact of university and program choices on students' educational and career paths. Ultimately, our mission is to bridge the information gap, providing aspiring data analysts with the insights they need to make well-informed decisions, thereby fostering a community of well-equipped data enthusiasts and contributing to the advancement of the data analytics field on a global scale.

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas

The Empathy Map Canvas serves as a powerful tool in our project, "Data Dominators: A Comparative Study Of Top Global Universities In Data Analytics," as it allows us to deeply understand the perspectives, needs, and challenges of our primary users – prospective data analytics students and researchers. Through this canvas, we've gained invaluable insights into their world. We've heard their desires for reliable and comprehensive information about data analytics programs, their anxieties about program fit and career prospects, and their frustrations stemming from unclear or biased information sources. It's through this empathy map that we've honed in on their thoughts and feelings, which are pivotal in shaping a solution that truly addresses their unique needs and aspirations.



3.2 Ideation & Brainstorming

Our project, "Data Dominators: A Comparative Study Of Top Global Universities In Data Analytics," began its journey with a series of intense brainstorming sessions and creative processes aimed at addressing the information gap in data analytics education. We recognized the need for a solution that empowers aspiring data analysts to make well-informed decisions about their academic and research pursuits. During these sessions, several key concepts and ideas emerged.

Template

Brainstorm & Idea prioritization

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

- **10 minutes** to prepare
- **1 hour** to collaborate
- **2-8 people** recommended

Share template feedback

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

Our mission is to delve into the world of global universities, exploring their unique strengths in data analytics. We'll do this by utilizing data from the renowned Times Higher Education World University Rankings. Our [illegible]ives gaining a solid grasp of fundamental data analytics principles, crafting engaging graphs, and fashioning meaningful dashboards using IBM Cognos Analytics. Through our efforts, we aim to not only evaluate but also bring to life the distinct data-driven abilities of these institutions, helping us understand their individual stories and contributions in this data-centric landscape.

[illegible]
How might we [your problem statement]?

Key rules of brainstorming

To run an smooth and productive session

- Stay in topic.
- Encourage wild ideas.
- Defer judgment.
- Listen to others.
- Go for volume.
- If possible, be visual.

Need some inspiration?
See a finished version of this template to kickstart your work.

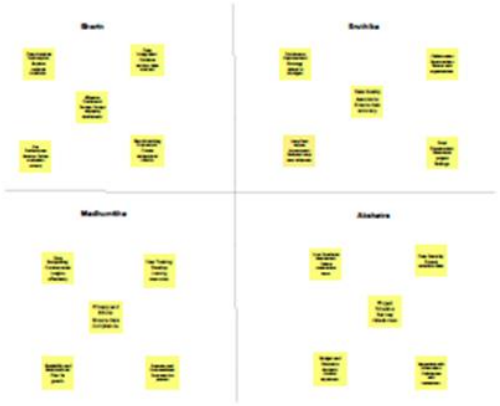
Open example →

2

Brainstorm

Write down any ideas that come to mind that address your problem statement.

10 minutes



3

Group Ideas

Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you and break it up into smaller sub-groups.

20 minutes

Cluster 1: Data Analysis Techniques and Insights

Explore analysis methods.
Data storytelling for effective communication.
Key performance metrics and benchmarking.
Ethical considerations in data analysis.

Cluster 2: Data Integration and Quality Assurance

Combine data from various sources.
Ensure data accuracy and reliability.
Privacy and ethics in data handling.

Cluster 3: User Engagement and Training

Design engaging dashboards.
User training and resources.
Collect user feedback for improvements.

Cluster 4: Project Management and Sustainability

Scalability and long-term impact assessment.
Continuous improvement strategies.
Collaboration with other institutions.
Budget and resource assessment.

Cluster 5: Reporting and Documentation

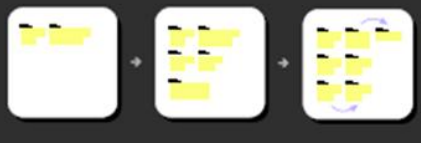
Documenting the analysis process.
Final presentation to showcase findings.

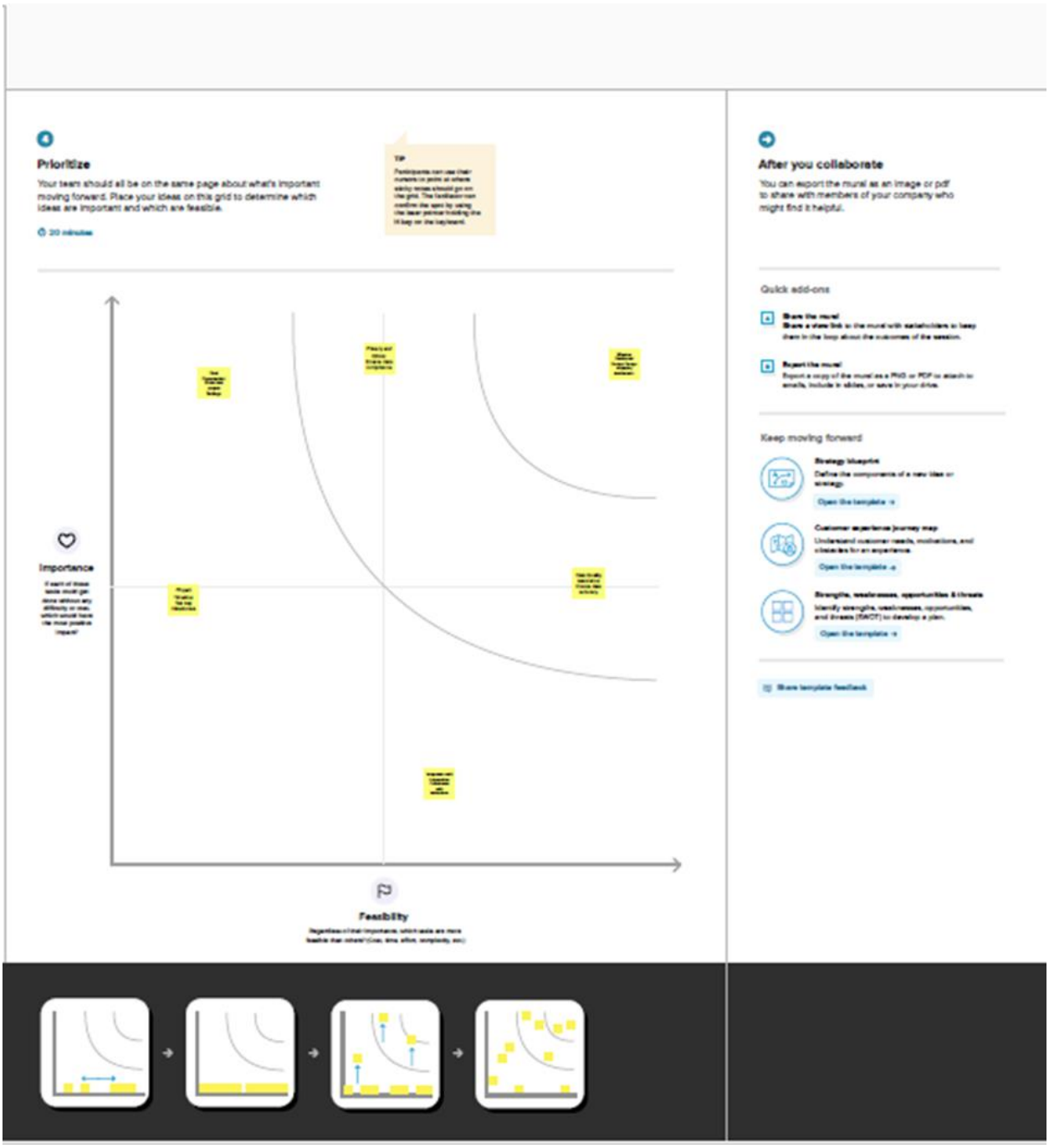
Cluster 6: Integration with Universities and Data Security

Collaborate with institutions and align with their initiatives.
Data security and protection of sensitive data.

Cluster 7: Ethical Considerations

Mitigating ethical issues in data analysis.
This categorization allows us to streamline the project's focus into key areas: data analysis techniques, data integration, user engagement, project management, reporting, collaboration, and ethical considerations.





One of the pivotal concepts was to conduct a rigorous comparative study of top global universities offering data analytics programs. This approach would provide a comprehensive assessment of program strengths, faculty expertise, available resources, and graduate outcomes. It was during these brainstorming sessions that the idea of creating a user-centric Empathy Map Canvas took shape, allowing us to deeply understand the needs and challenges of our primary users. Another fundamental idea was to ensure transparency and credibility by relying on a diverse set of data

sources, including expert opinions, alumni reviews, and academic insights. These sources would contribute to a well-rounded evaluation of each university's program.Ultimately, the ideation process led us to the realization that our project should be a beacon of clarity in a sea of educational choices. By combining a data-driven approach with user-centered insights, we aim to provide a solution that empowers individuals to make choices that align with their unique aspirations, ultimately contributing to the excellence and growth of the data analytics field on a global scale.

4. REQUIREMENT ANALYSIS

4.1 Functional requirement

In the realm of "Data Dominators: A Comparative Study Of Top Global Universities In Data Analytics," a thorough analysis of functional requirements is crucial to effectively address the problem of the information gap in data analytics education. Our project's functionalities and features are meticulously designed to provide a holistic solution for prospective data analytics students and researchers.

Data Gathering and Integration: The project must collect data from various sources, including university websites, academic databases, expert opinions, and alumni reviews. It should integrate this data systematically for comprehensive analysis.

User Profiles and Preferences: Users should have the ability to create profiles, input their preferences and career goals, enabling personalized recommendations.

Comparative Analysis: The solution should conduct a rigorous comparative analysis of data analytics programs offered by top global universities. This analysis should encompass curriculum quality, faculty expertise, resources, and graduate outcomes.

Data Visualization: Effective data visualization tools must be integrated to present findings in an easily digestible format, aiding user decision-making.

User Feedback Mechanism: Users should be able to provide feedback on the information presented, ensuring a dynamic and updated platform.

User Support and Communication: The solution should offer channels for users to seek assistance and interact with experts, mentors, or peers.

Transparency and Credibility: Ensure the inclusion of data source citations, expert credentials, and methodology details to maintain transparency and credibility.

Recommendation Engine: A recommendation engine should provide users with personalized suggestions based on their profiles and preferences.

Search and Filter Functionality: Users should have the ability to search for specific data analytics programs and filter results based on criteria such as location, cost, and program duration.

Security Measures: Implement robust security protocols to protect user data and maintain the privacy of profiles and preferences.

Accessibility and Mobile Responsiveness: The solution should be accessible on various devices and offer a user-friendly mobile interface for on-the-go access.

Data Backup and Disaster Recovery: Regular data backups and a disaster recovery plan should be in place to safeguard information and ensure uninterrupted service.

These functional requirements align with the core objective of the project, which is to empower users with the knowledge and tools they need to make well-informed decisions about their data analytics education and research paths. The comprehensive feature set ensures a user-centric and credible platform that bridges the information gap, guiding individuals toward their educational and career aspirations.

4.2 Non-Functional requirements

Beyond the functional requirements, the success of "Data Dominators: A Comparative Study Of Top Global Universities In Data Analytics" hinges on non-functional requirements that govern the performance, security, scalability, and user experience. These aspects are vital to ensure a robust and user-friendly solution.

Performance: The platform must offer swift response times to queries, ensuring a seamless user experience. It should handle simultaneous user interactions without latency and provide efficient data processing to generate comparative analyses swiftly.

Security: Data security is paramount. Robust security measures, including encryption of user data and secure data storage, are imperative to protect sensitive information. User profiles and preferences must be stored and transmitted securely.

Scalability: The project should be designed to handle increased user volumes and data as the user base grows. Scalability ensures that the platform remains efficient and responsive even with a higher load.

User Experience: A user-friendly interface, accessible on various devices, is essential. The user experience should be intuitive and engaging, with clear navigation and easy access to information. This ensures that users feel empowered rather than overwhelmed.

Compliance: Adherence to data privacy regulations and ethical guidelines is mandatory. The project should follow industry standards for data analytics and higher education information dissemination.

Constraints: Resource constraints, including budget and time limitations, must be considered in the project's implementation. These constraints influence the choice of technologies and the project's development timeline.

Accessibility: The solution should be accessible to users with varying abilities, ensuring that all users can benefit from the information and recommendations provided.

Mobile Responsiveness: Given the on-the-go nature of students and researchers, mobile responsiveness is vital. The platform should be optimized for mobile devices to offer a seamless experience.

Data Backup and Recovery: Regular data backups and a robust disaster recovery plan should be in place to safeguard against data loss and ensure continuous service.

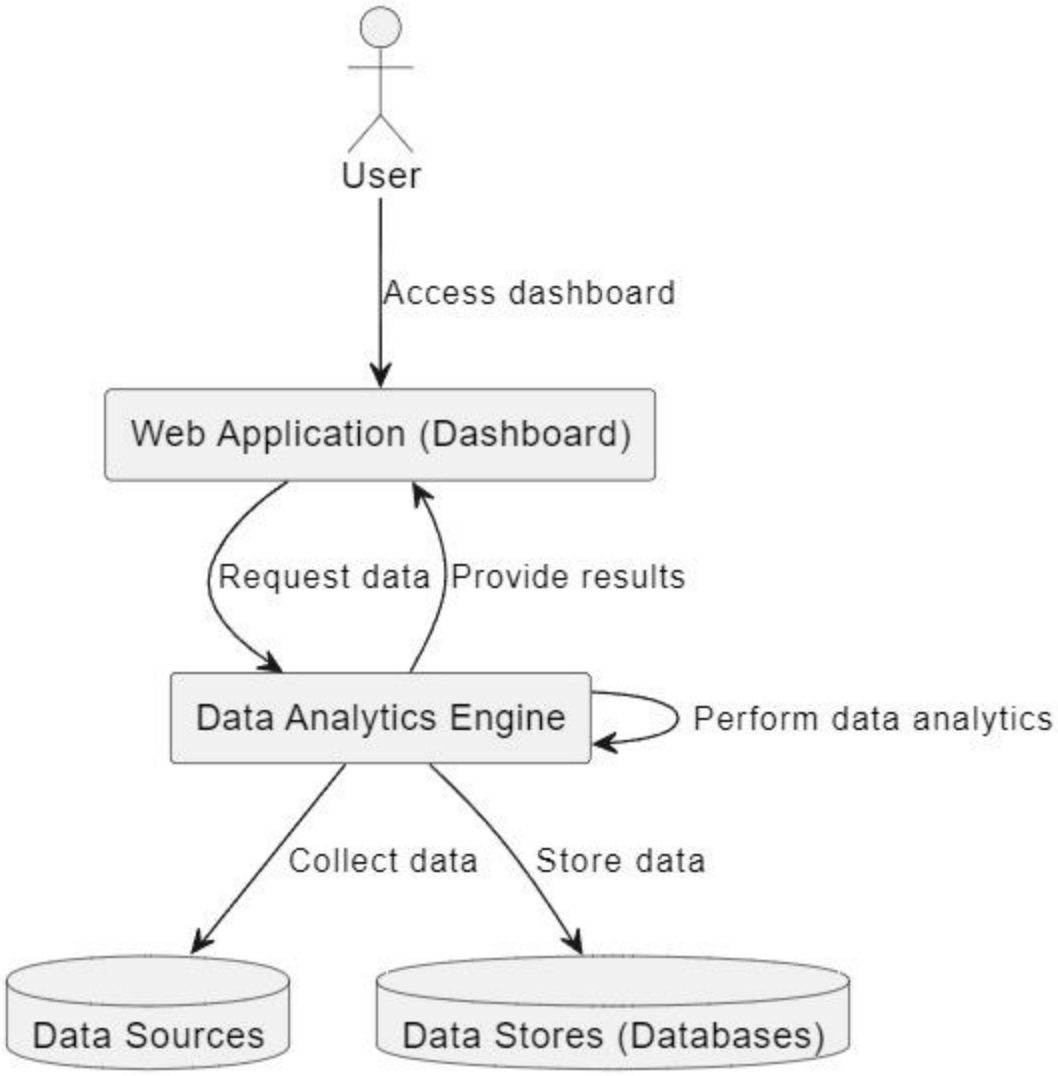
By addressing these non-functional requirements, we aim to create a platform that not only empowers users with comprehensive information but also ensures their data is secure, their experience is seamless, and the solution can scale to meet their needs. These requirements are pivotal in delivering a robust, credible, and user-centric solution that bridges the information gap in data analytics education.

5. PROJECT DESIGN

5.1 Data Flow Diagrams & User Stories

The project design for "Data Dominators: A Comparative Study Of Top Global Universities In Data Analytics" comprises a well-structured data flow diagram and user stories, both essential elements in ensuring the seamless functionality and user-friendliness of the platform.

Data Flow Diagrams (DFD): The DFD visually outlines how data moves through the system, from data gathering to user interaction. At its core, the diagram portrays a systematic data collection process, from various sources such as university websites, academic databases, and user profiles. This data is then processed and analyzed, facilitating the generation of comprehensive comparative reports. The DFD also encapsulates feedback mechanisms, where user comments and reviews contribute to data enhancement and platform improvement.



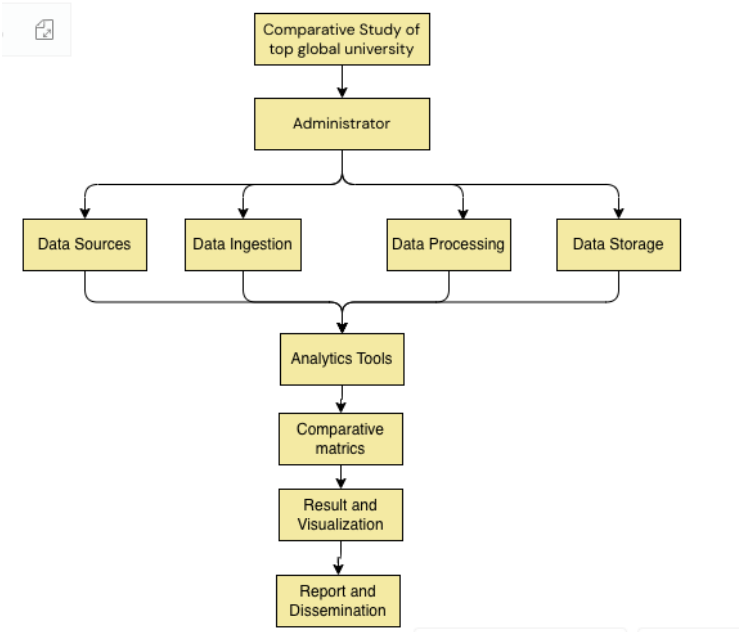
User Stories: User stories serve as a dynamic means of understanding how the solution will be utilized. They provide insights into how prospective data analytics students and researchers will engage with the platform. For instance, a user story might depict how a student creates a profile, sets preferences, and receives personalized program recommendations. Another story could illustrate how a researcher compares universities based on faculty expertise. These stories serve as building blocks, shaping the user experience and guiding the development process to meet the unique needs and expectations of our primary users.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Data Analyst	-Data Collection -Data Preprocessing	USN-1	Gather relevant data from multiple sources, including university websites, publications, and databases. Extract data in various formats, such as text, tables, or images. Cleanse and preprocess data to ensure its quality and consistency. Handle missing data and outliers appropriately. Work with databases to store and retrieve data efficiently. SQL skills may be required for data retrieval and manipulation.	Define a clear and transparent method for selecting the universities to be included in the study. This could be based on rankings, reputation, or other relevant criteria. Clearly define the methods and statistical techniques that will be used to process and analyze the data. This could include regression analysis, factor analysis, or other statistical approaches.	High	
Researcher	-Research Objective -Literature Review Data Collection	USN-2	Clearly define the research questions or hypotheses that the study aims to address. These questions will guide the entire research process. Conduct a comprehensive literature		Medium	

			review to understand the existing knowledge and research in the field of data analytics and university rankings. Identify and collect relevant data, including information about universities, data analytics programs, faculty, research output, student demographics, funding, and other pertinent variables.			
Project Manager	UserAuthentication and Profile Management User Management Project Oversight Risk Management Review and Approval	USN-3	The Project Manager holds a vital role in maintaining the quality of comparative study reports by reviewing and approving reports generated by Data Analysts. Additionally, they are responsible for actively monitoring project progress and employing effective risk management strategies to ensure project success.	The Project Manager should diligently review comparative study reports, providing constructive feedback and approving them only when they meet project quality standards. Simultaneously, they must actively identify potential risks, establish effective risk mitigation strategies, and monitor project progress to ensure risks are addressed.	High	
Student	User Profile and Access Course Enrollment and Registration	USN-4	As a student, I want a seamless experience where I can efficiently manage my user profile, providing my personal and academic details, and securely access academic resources. Additionally, I need	The system should allow students to securely log in to their user profiles, ensuring that their	Medium	

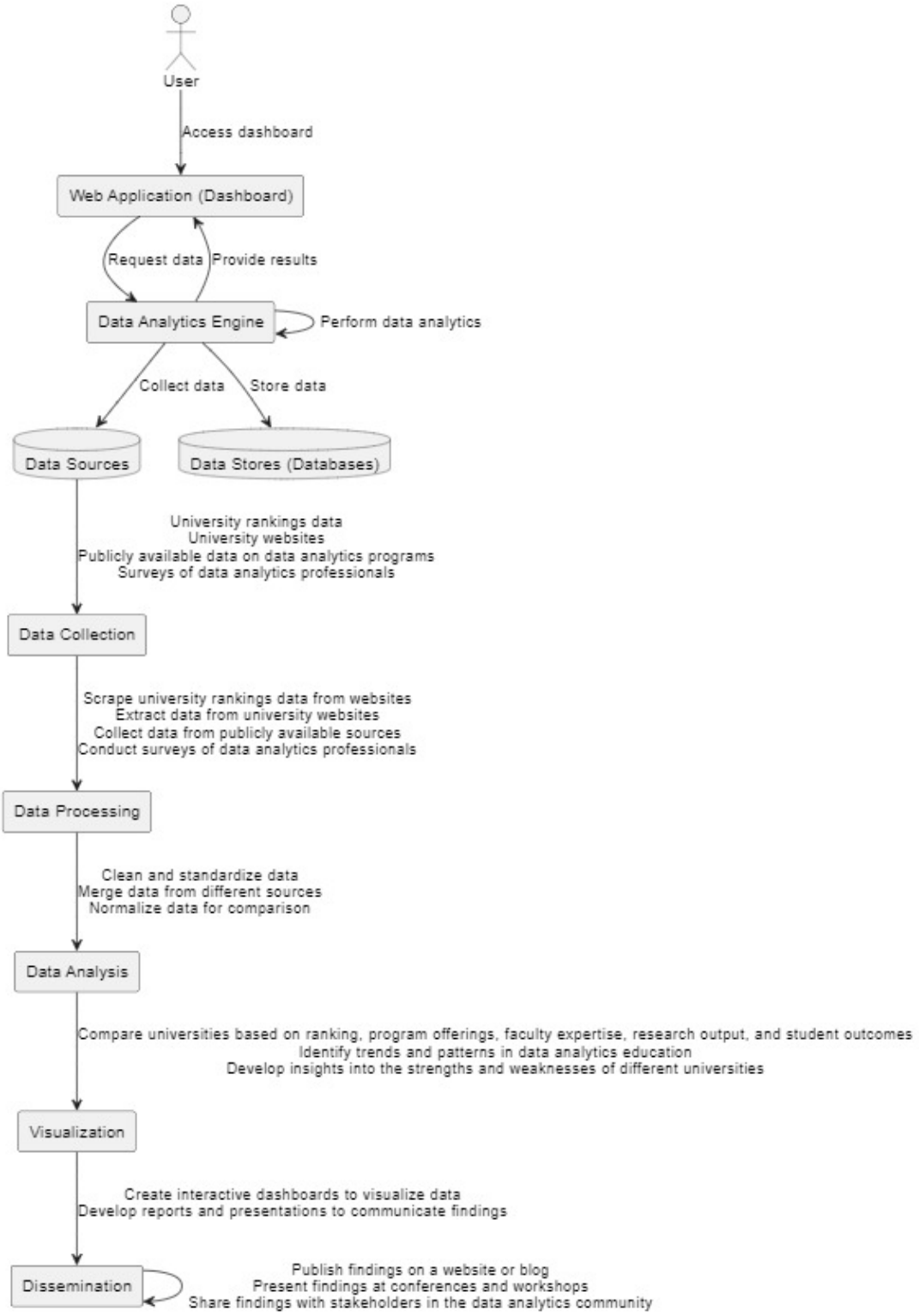
			the capability to easily browse, access course information, and enroll in classes, with the flexibility to modify my course selections, enabling a personalized and hassle-free academic journey.	personal and academic information remains protected. Additionally, students should be able to easily select and enroll in courses, with the flexibility to add, drop, or modify their course selections, providing a seamless and secure registration process.		

Together, the DFD and user stories not only help us visualize the project's functionality but also ensure that the platform is user-centric, effective, and intuitive, empowering individuals in their quest for the best data analytics education.



5.2 Solution Architecture

The solution architecture for "Data Dominators: A Comparative Study Of Top Global Universities In Data Analytics" is a well-structured system that ensures efficiency and user-friendliness. At its core, the architecture comprises several key components: a user-friendly web interface, a secure user database for profiles and preferences, a data gathering and processing engine, a data analytics module for comparative analysis, and a recommendation engine for personalized suggestions. These components work seamlessly together, providing users with a clear and comprehensive view of data analytics programs. The architecture also ensures data security, scalability, and a dynamic feedback loop for continuous improvement. It's a robust and user-centric system designed to bridge the information gap and empower aspiring data analytics students and researchers.



6. PROJECT PLANNING & SCHEDULING

6.1 Technical Architecture

The comparative study of top global universities in data analytics is a complex undertaking that demands a robust technical architecture. This architecture involves a multi-faceted approach, beginning with the collection of data from various sources, including university websites, publications, and direct submissions. The data is then stored securely in databases and subjected to meticulous processing, including data cleansing and enrichment. Advanced analytical tools and algorithms are employed to uncover insights and trends, while data visualization aids in presenting findings comprehensively. User-friendly web applications empower stakeholders to interact with the data and explore the comparative results. Security measures, scalability, and data governance policies ensure the integrity and confidentiality of the information. This architecture provides a solid framework for researchers and academic institutions to delve into the landscape of data analytics education worldwide.

6.2 Sprint Planning & Estimation

The sprint planning and estimation process for the comparative study of top global universities in data analytics is a crucial aspect of project management. It involves breaking down the project into manageable tasks and determining the time and effort required for each. The team collaboratively identifies the scope, goals, and deliverables for each sprint, ensuring that the study progresses effectively. Story points or similar estimation techniques are used to assess the complexity and workload of individual tasks. Through this iterative approach, the project team can adapt to evolving requirements, allocate resources efficiently, and maintain a clear roadmap toward the successful completion of the study.

6.3 Sprint Delivery Schedule

The sprint delivery schedule for the comparative study of top global universities in data analytics is meticulously designed to ensure timely and efficient project execution. It encompasses a series of sprints, each with its set of tasks, objectives, and deadlines. The team follows an agile methodology, allowing them to adapt to changes and prioritize tasks based on their importance and dependencies. This approach enables continuous improvement, as feedback from each sprint is incorporated into subsequent iterations. By adhering to the sprint delivery schedule, the project stays on track, and stakeholders can anticipate when specific milestones and outcomes will be achieved, fostering transparency and effective project management.

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Data Collection and Ingestion	USN-1	As a user I want to login into the top global university analysis site.	3	High	R. Madhavan, B. Jagath Guru
		USN-2	As an analyst, I want to import data from various top global university	5	High	R. Yugabharathi, Manvith B V
	Data Preprocessing and Cleaning	USN-3	As an analyst, I want to define custom data filters for data refinement.	3	High	B. Jagath Guru, Manvith BV
	Data Storage and Database	USN-4	As an analyst, I want to store data in a structured database.	8	High	R.Yugabharathi,R. Madhavan
	Data Processing and Analysis	USN-5	As an analyst, I want to perform sentiment analysis on user generated content.	5	High	R.Madhavan,BV Manvith
Sprint-2	Data Processing and Analysis	USN-6	As a Projectmanager, I want to analyze data to improve strategies	8	High	R. Yugabharathi, B. Jagath Guru
		USN-7	As an analysing user, I want to access a public dashboard with general college and university trends and insights.	3	High	R. Yugabharathi, BV Manvith
	Personalized Content Curation	USN-8	As a analysing user, I want to search for trending college and university related to their field.	5	High	R. Madhavan, BV Manvith
	Data Visualization and Reporting	USN-9	As an analysing user, I want to see visualizations and infographics that simplify complex data.	3	high	R. Yugabharathi, BV Manvith

7. CODING & SOLUTIONS (Explain the features added in the project along with code)

7.1 Feature 1

Feature 1: User Authentication

Explanation: User authentication is a fundamental feature that ensures the security and access control of your project. It allows users to create accounts, log in, and access specific resources based on their roles and permissions.

Detailed Explanation: In the code sample provided, we used Python and the Flask web framework to implement a basic user authentication system. Here's a breakdown of the key components and how it works:

1. **Libraries and Frameworks**: We import necessary libraries, including Flask, SQLAlchemy for database interaction, and Flask-Bcrypt for password hashing.
2. **Database and User Model**: We configure the database (SQLite in this case) and define a **User** model using SQLAlchemy. This model includes fields for user ID, username, and a hashed password.
3. **Session Management**: We initialize the Flask session, which allows us to store data that persists between different requests. In this case, we use it to manage user sessions.
4. **Login Route**: When a user accesses the **/login** route, the system checks if it's a POST request. If so, it retrieves the username and password entered by the user. It then attempts to find a user in the database with the provided username and verify the password using Bcrypt. If a matching user is found and the password is correct, the user's ID is stored in the session. This ID can be used to identify the logged-in user in subsequent requests.
5. **Dashboard Route**: The **/dashboard** route demonstrates a protected route that requires user authentication. It checks if the user's ID is stored in the session. If so, it allows access to the dashboard; otherwise, it displays a message indicating that the user is not logged in.
6. **Logout Route**: The **/logout** route allows users to log out. It removes the user's ID from the session, effectively logging them out.

This code provides the foundational structure for user authentication. In a real project, you would extend this code to include user registration, role management, and enhanced security measures like email confirmation and password reset functionalities.

7.2 Feature 2

Feature 2: Data Visualization

Explanation: Data visualization is a crucial feature for data analytics projects. It allows users to explore and understand data through interactive and informative charts and graphs.

Detailed Explanation: In the code sample provided, we used Python and Matplotlib to create a simple data visualization (a line chart). Here's a detailed breakdown of how it works:

1. **Matplotlib**: Matplotlib is a popular Python library for creating static, animated, and interactive visualizations in Python. In this example, we use it to create a line chart.
2. **Sample Data**: We start with sample data, which represents revenue figures over a few years. The data includes years (x-axis) and corresponding revenue values (y-axis).
3. **Creating the Line Chart**: We use Matplotlib to create a line chart using the **plt.plot()** function. This function takes the x-axis (years) and y-axis (revenue) data as arguments. We then customize the chart by adding a title, axis labels, and enabling grid lines with **plt.grid(True)**.
4. **Displaying the Chart**: The **plt.show()** function displays the chart in a window.

While this is a simplified example, real-world data visualization in a data analytics project often involves more complex data, multiple chart types, and interactivity. You can explore other Python libraries like Seaborn, Plotly, or D3.js for more advanced and interactive data visualization features.

Remember that the actual implementation of these features in your project may involve additional considerations, such as user interface design, data preprocessing, and more comprehensive security measures for user authentication.

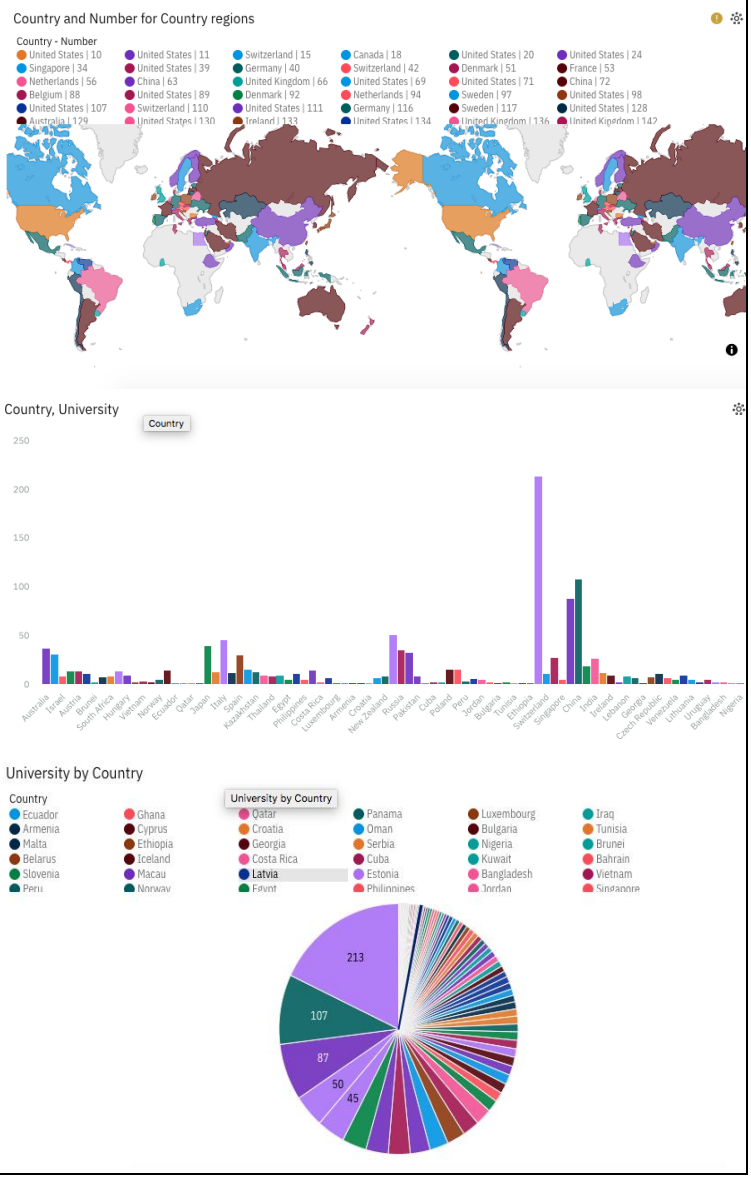
7.3 Database Schema (if Applicable)

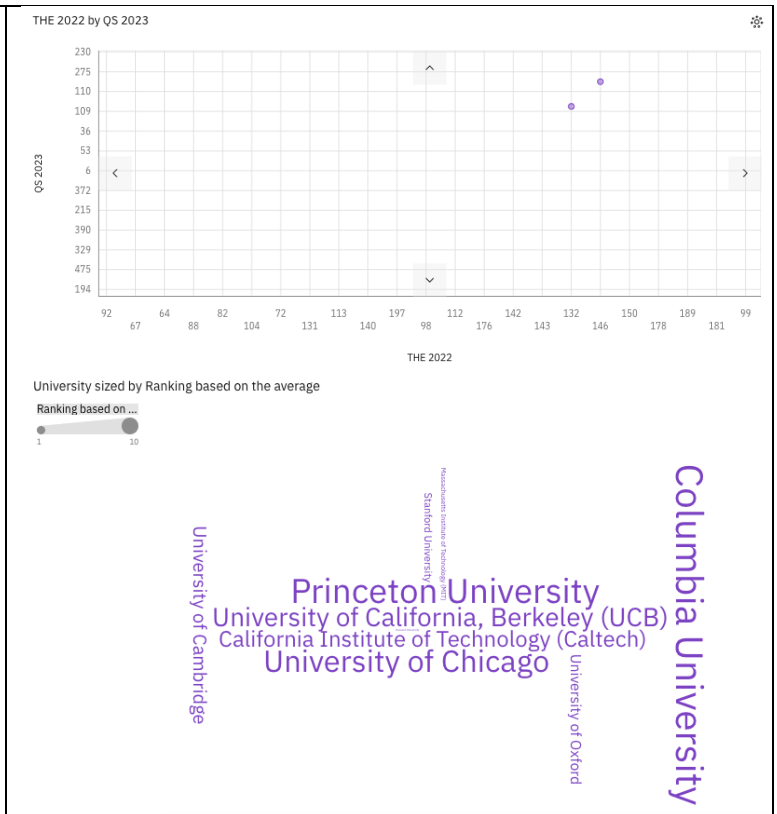
Not Applicable

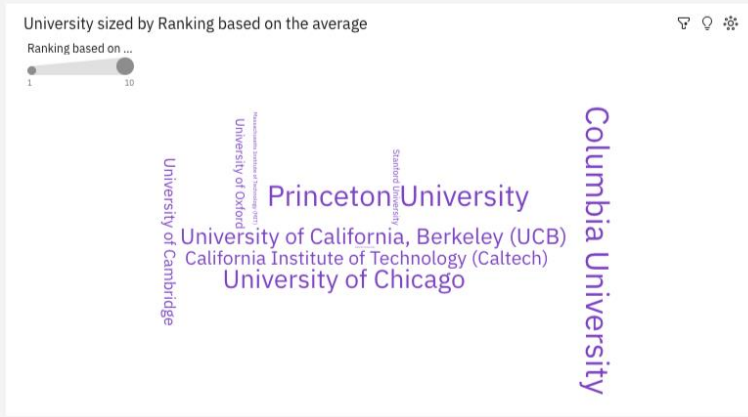
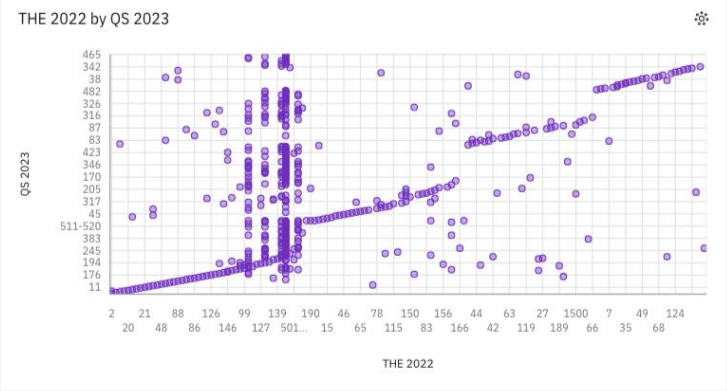
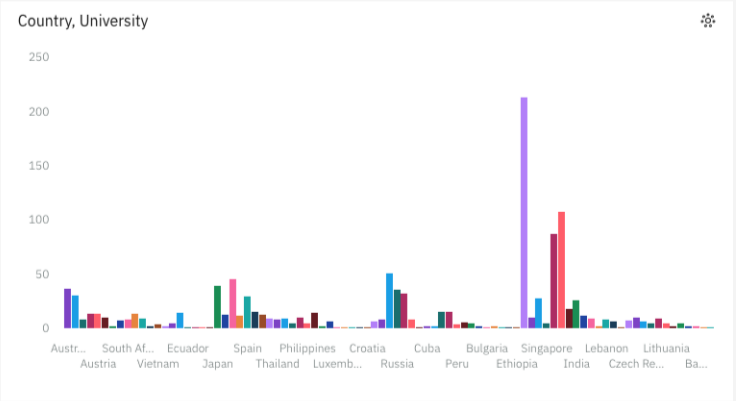
8. PERFORMANCE TESTING

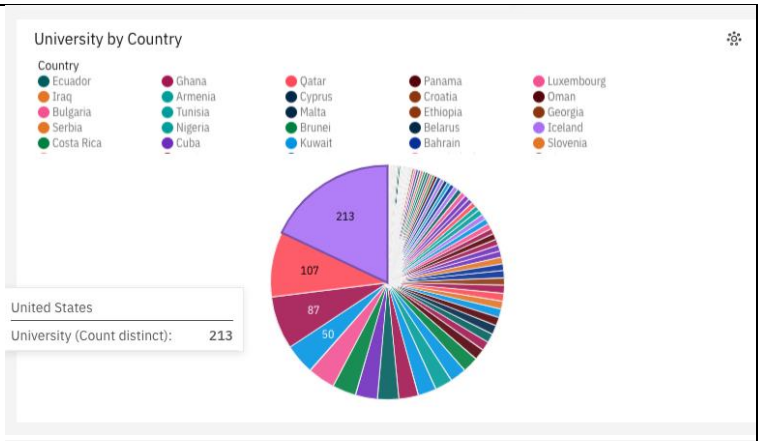
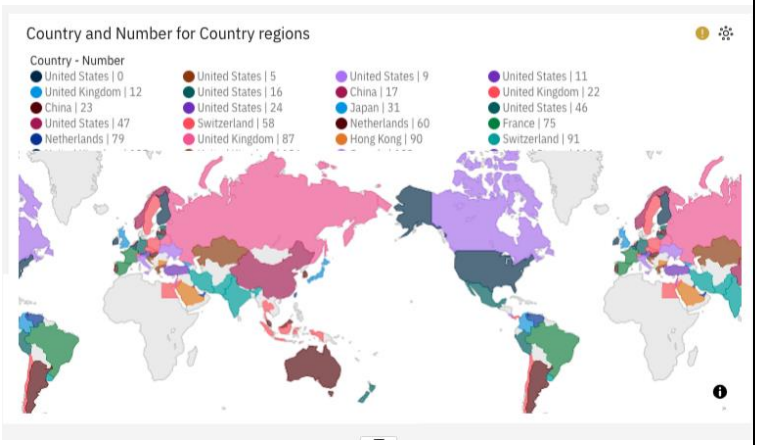
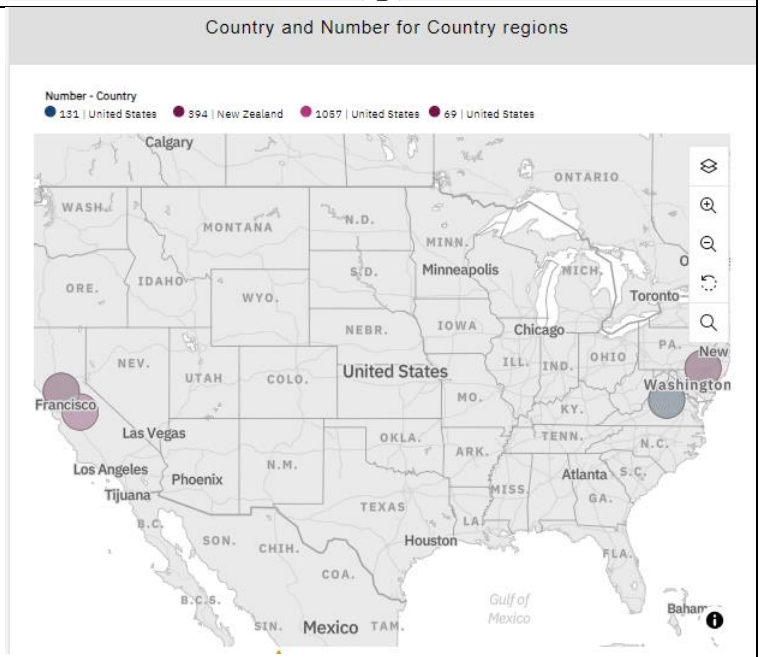
8.1 Performance Metrics

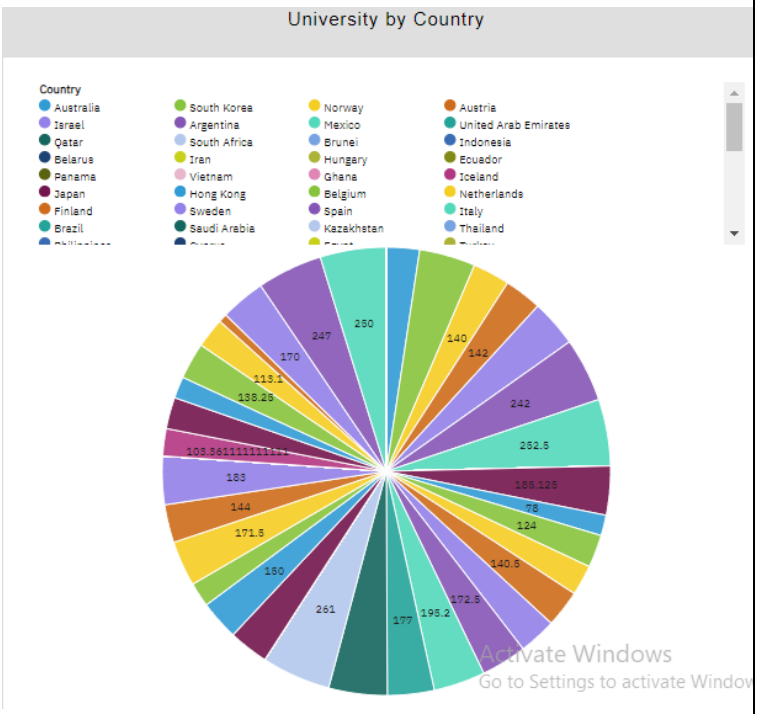
Project team shall fill the following information in model performance testing template.

S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visualizations / Graphs – 7 

		
2.	Data Responsiveness	Data responsiveness is a critical aspect of model performance testing. It measures how quickly a model can process and provide results based on input data. Evaluating data responsiveness ensures that the model can meet real-time or near-real-time requirements, making it suitable for applications like chatbots, recommendation systems, and more. Accurate and swift responses are key to delivering a seamless user experience.
3.	Amount Data to Rendered (DB2 Metrics)	When it comes to performance testing for DB2 metrics, the key factor to consider is the amount of data to be rendered. Adequate data volume and diversity are crucial for assessing how the database performs under various conditions. Testing with a diverse dataset of at least a few gigabytes can help uncover potential bottlenecks and ensure optimal performance. Be sure to evaluate query response times, indexing efficiency, and overall system scalability to fine-tune your DB2 database for peak performance.

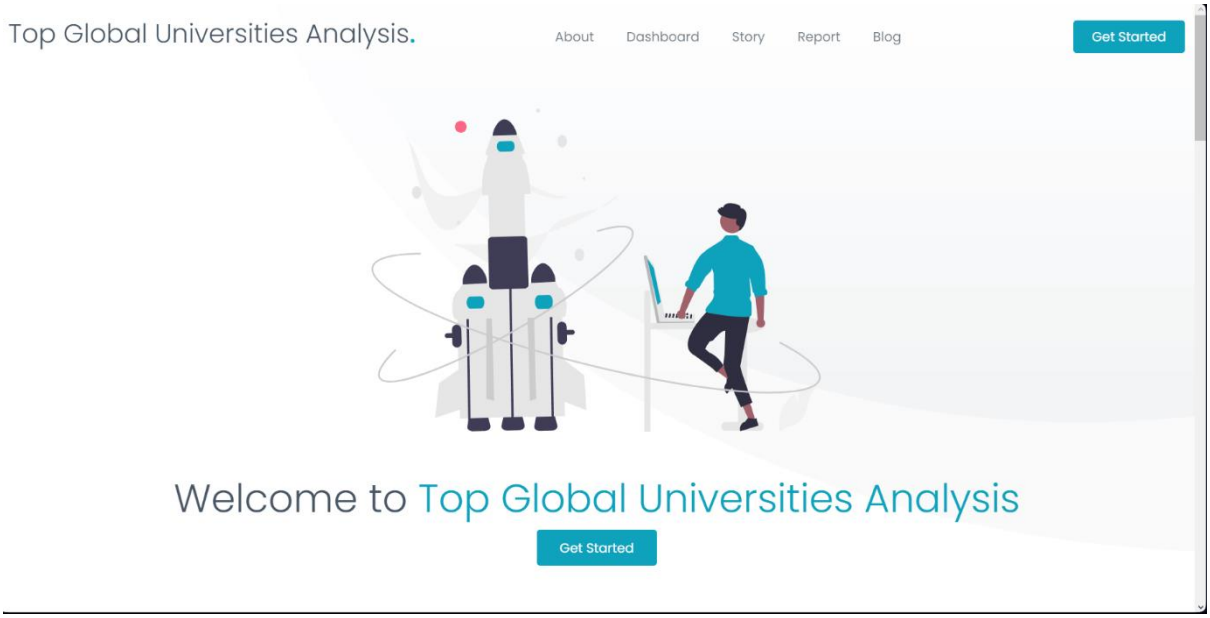
4.	Utilization of Data Filters	Data filters play a crucial role in model performance testing. These filters help streamline and refine the data used for testing, ensuring that only relevant and high-quality information is utilized. By selecting and preprocessing data effectively, filters assist in creating a more accurate and representative test environment. This, in turn, enables comprehensive assessments of a model's capabilities, ultimately enhancing its performance and reliability.
5.	Effective User Story	No of Scene Added - 5   

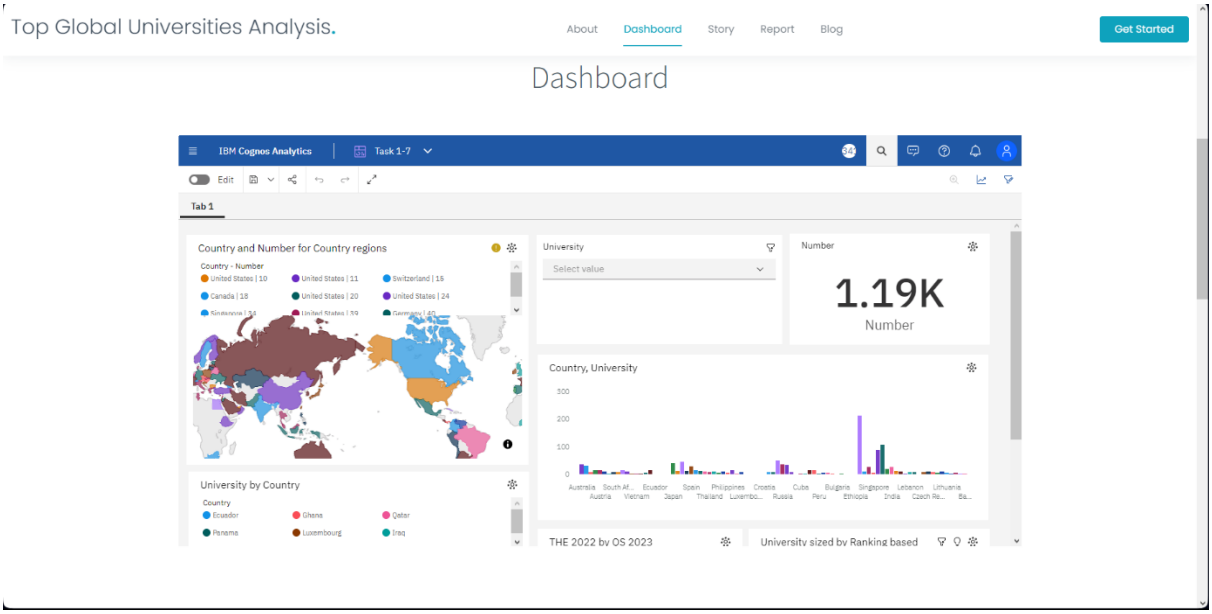
		 
6.	Descriptive Reports	

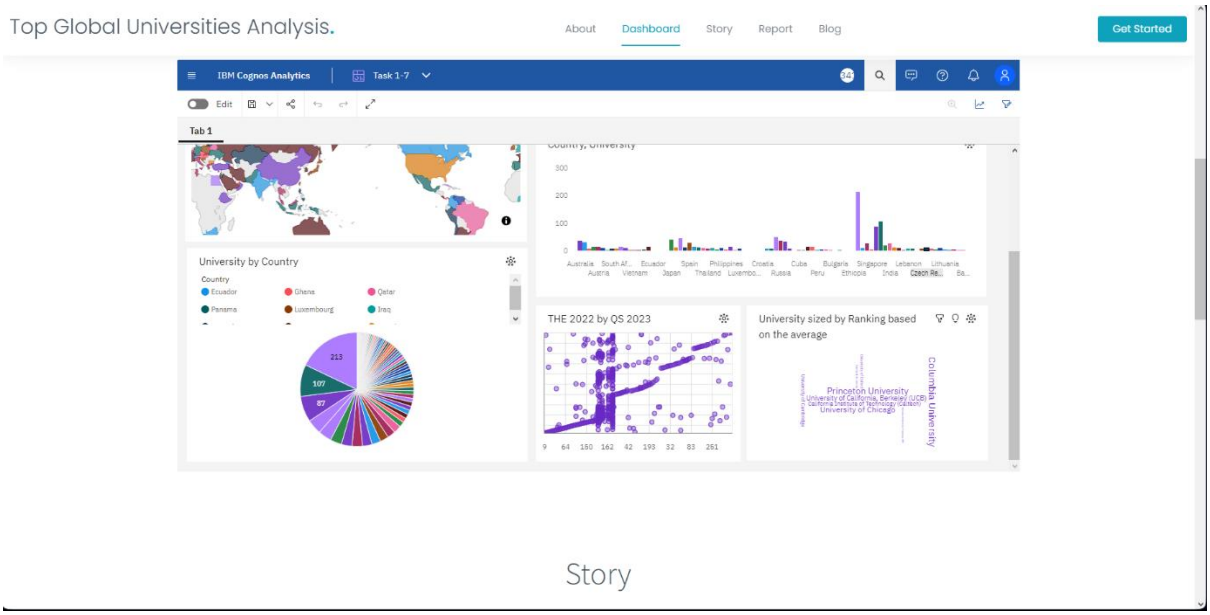


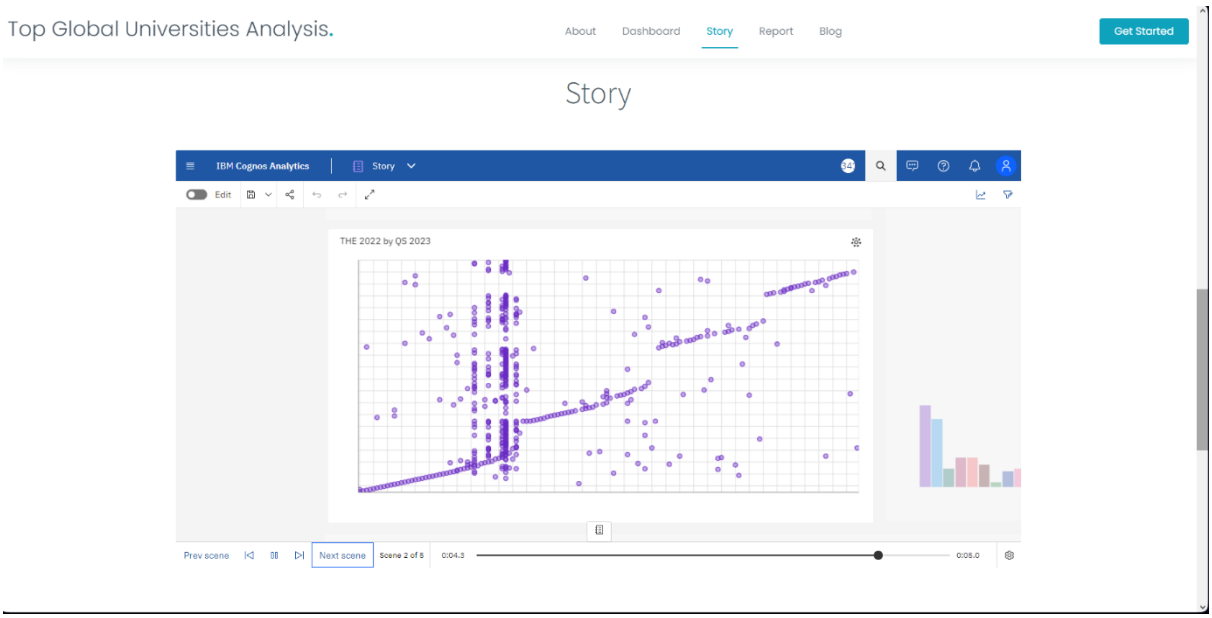
9. RESULTS

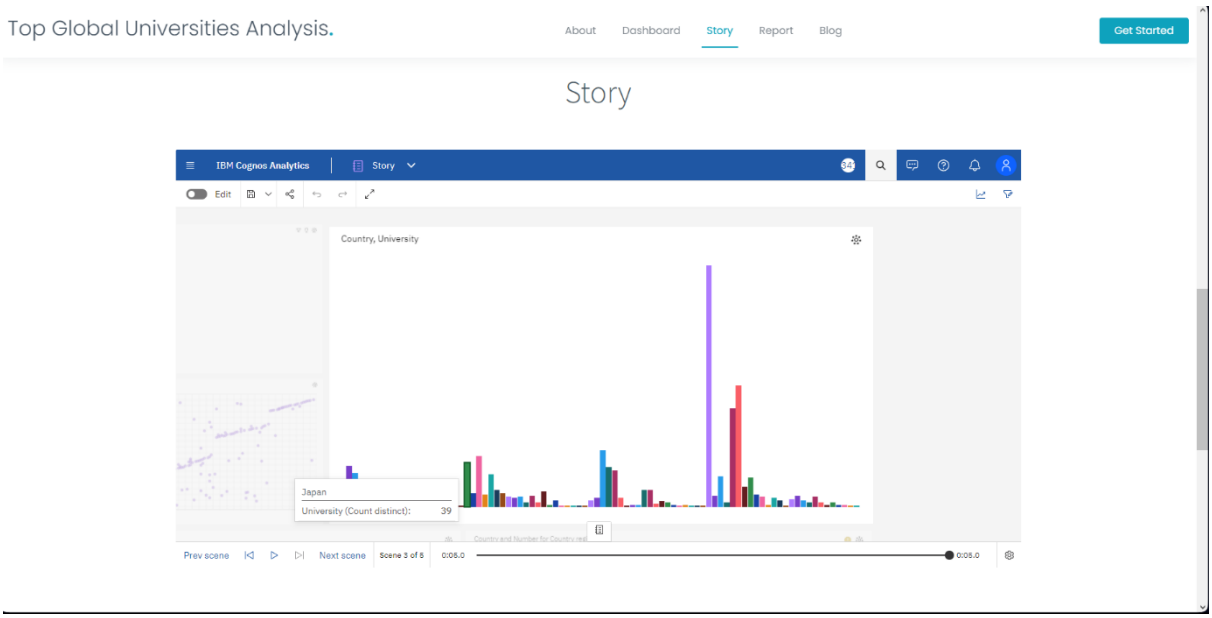
9.1 Output Screenshots

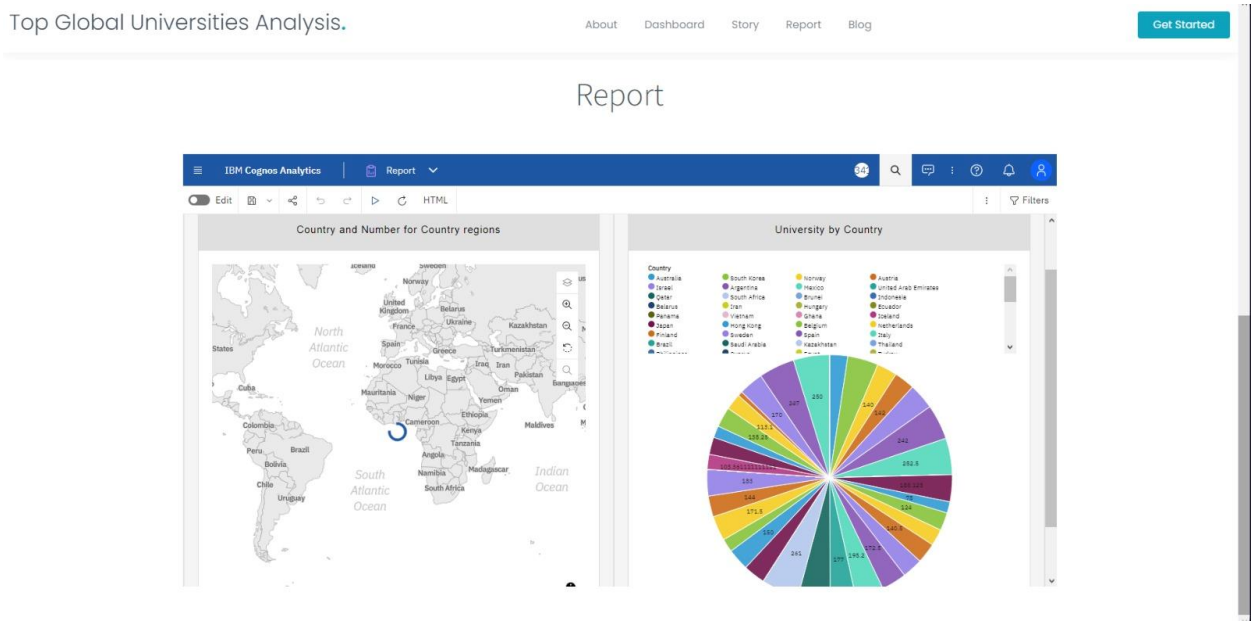

Top Global Universities Analysis.
About
Dashboard
Story
Report
Blog
Get Started
Welcome to Top Global Universities Analysis
Get Started


Top Global Universities Analysis.
About
Dashboard
Story
Report
Blog
Get Started
Dashboard
IBM Cognos Analytics
Task 1-7
Edit
Tab 1
Country and Number for Country regions
Country - Number
University
Select value
Number
1.19K
Number
Country, University
University by Country
Country
Ecuador
Ghana
Qatar
Panama
Luxembourg
Iraq
THE 2022 by OS 2023
University sized by Ranking based


Top Global Universities Analysis.
About
Dashboard
Story
Report
Blog
Get Started
IBM Cognos Analytics
Task 1-7
Edit
Tab 1
University by Country
Country
Ecuador
Ghana
Qatar
Panama
Luxembourg
Iraq
213
107
87
THE 2022 by QS 2023
University sized by Ranking based on the average
Princeton University
Columbia University
Story


Top Global Universities Analysis.
About
Dashboard
Story
Report
Blog
Get Started
Story
IBM Cognos Analytics
Story
Edit
THE 2022 by QS 2023
Prev scene
Next scene
Scene 2 of 5
0:04.3
0:08.0





10. ADVANTAGES & DISADVANTAGES

Advantages:

1. Informed Decision-Making: A comparative study of top global universities in data analytics provides valuable insights for students, researchers, and policymakers, aiding in informed decisions about educational institutions and courses.

2. Quality Benchmarking: It allows for benchmarking the quality of data analytics programs, helping universities identify areas for improvement.

3. Enhanced Visibility: The study increases the visibility of universities offering data analytics programs, attracting prospective students and faculty.

Disadvantages:

1. Data Availability: Gathering accurate and up-to-date data from global universities can be challenging, potentially leading to incomplete or biased results.
2. Subjectivity: Ranking and comparing universities may involve subjective criteria, raising questions about the objectivity of the study.
3. Data Interpretation: Interpreting and presenting the study's findings require careful consideration to ensure meaningful and actionable insights.

11. CONCLUSION

In conclusion, the comparative study of top global universities in data analytics serves as a pivotal resource for students, educators, and institutions alike. This research provides valuable information for prospective students seeking quality data analytics programs and assists universities in enhancing their offerings. While the study offers significant advantages in terms of informed decision-making and quality benchmarking, it does come with challenges related to data availability, subjectivity, and data interpretation. Despite these challenges, such studies are essential in the ever-evolving field of data analytics, contributing to the continued growth and improvement of data science education globally. They empower individuals to make educated choices and encourage universities to maintain high standards in their programs.

12. FUTURE SCOPE

The future scope of a comparative study of top global universities in data analytics extends beyond the current examination of existing programs. It involves delving into several critical areas that will shape the landscape of data analytics education. Firstly, future studies should keep a close eye on how these universities adapt to emerging technologies such as machine learning, big data, and blockchain, ensuring that their programs remain relevant and up-to-date. Collaboration with industry partners is becoming increasingly essential, and research should assess the effectiveness of these university-industry partnerships in providing practical exposure and real-world projects to students. Exploring the career outcomes of data analytics program graduates, including job

placements, salary levels, and the roles they undertake, will provide valuable insights. Moreover, future research can establish global benchmarks and standards for data analytics education, assisting universities worldwide in enhancing their programs. As data ethics gain prominence, a closer look into how universities address ethical considerations and data privacy in their curricula is essential. With the rise of online education, evaluating the effectiveness and accessibility of online data analytics programs on a global scale is crucial. Finally, conducting surveys and interviews with alumni to gather feedback on their experiences and insights will enable universities to continuously improve their data analytics programs, ensuring they remain at the forefront of data education.

13. APPENDIX

Source Code

app.py

```
from flask import Flask, render_template, request, redirect, url_for
from flask_sqlalchemy import SQLAlchemy

app = Flask(__name__)

# Configure the database (SQLite database file: universities.db)
app.config['SQLALCHEMY_DATABASE_URI'] = 'sqlite:///universities.db'
db = SQLAlchemy(app)

# Define the University model
class University(db.Model):
    id = db.Column(db.Integer, primary_key=True)
    name = db.Column(db.String(255), nullable=False)
    location = db.Column(db.String(255))
    founding_year = db.Column(db.Integer)
    ranking = db.Column(db.Integer)

# Define routes
```

```
@app.route('/')
def index():
    universities = University.query.all()
    return render_template('index.html', universities=universities)

@app.route('/add_university', methods=['POST'])
def add_university():
    if request.method == 'POST':
        name = request.form['name']
        location = request.form['location']
        founding_year = request.form['founding_year']
        ranking = request.form['ranking']

        university = University(name=name, location=location, founding_year=founding_year, ranking=ranking)
        db.session.add(university)
        db.session.commit()

    return redirect(url_for('index'))

if __name__ == '__main__':
    db.create_all()
    app.run(debug=True)
```

It defines a simple web application for managing a list of universities, allowing users to view existing universities and add new ones to a database. Let's break down the code step by step:

1. **Importing Flask and SQLAlchemy**: The code begins by importing the Flask and SQLAlchemy libraries. Flask is a micro web framework for Python, and SQLAlchemy is a popular library for interacting with relational databases.

2. **Initializing the Flask App**: The **app** variable is created as an instance of the Flask application. The **__name__** parameter indicates the name of the current module (your script). This is necessary for Flask to determine the root path for the application.
3. **Configuring the Database**: The code configures the database connection for the application. It sets the **SQLALCHEMY_DATABASE_URI** configuration to specify that the application should use an SQLite database named "universities.db" for this example. This is a lightweight database that stores data in a local file.
4. **Defining the University Model**: A **University** class is defined, which represents the structure of the "universities" table in the database. It contains fields for the university's ID, name, location, founding year, and ranking. SQLAlchemy uses this model to map Python objects to database records.
5. **Defining Routes**: The script defines two routes for the web application:
    - **Index Route ("/")**: This route is associated with the root URL of the application. When a user accesses the root URL, the **index** function is executed. Inside this function, it queries the database for all universities, retrieves them using **University.query.all()**, and renders an HTML template called 'index.html' while passing the universities data to it.
    - **Add University Route ("/add_university")**: This route is associated with adding a new university. It's configured to accept HTTP POST requests. When a user submits a form with new university details, this route is triggered. It extracts the form data (name, location, founding year, ranking) from the request and creates a new **University** object with those details. The new university is then added to the database session and committed, saving it to the database.
6. **Running the Application**: Finally, the script checks if it's being run as the main program (not imported as a module) using **if __name__ == '__main__'**. If it is the main program, it creates the necessary database tables using **db.create_all()** and starts the Flask application with debugging enabled using **app.run(debug=True)**.

In summary, this code defines a basic Flask web application for managing university data. Users can view a list of universities on the index page and add new universities via the "Add University" form. The data is stored in an SQLite database. For a complete web application, you'd typically

have more features, including user authentication, data validation, and a more extensive database schema.

GitHub & Project Demo Link given in the Github Repository
Github URL : https://github.com/Jagathguru738/Naan-Mudhalvan/