Project Report Format

1. INTRODUCTION

1.1 Project Overview

Job analysis is a systematic procedure to analyse the requirements for the job role and job profile.
Glassdoor is a website and online platform that provides information about jobs, salaries, and Companies.

Job analysis is a systematic approach to defining the job role, description, requirements, responsibilities, evaluation, etc. It helps in finding out the required level of education, skills, knowledge, training, etc for the job position.

It also depicts the job worth i.e. measurable ffectiveness of the job and contribution of job to the organization. Thus, it effectively contributes to setting up the compensation package for the job Position.

Lack of analysis of Glassdoor jobs can result in limited understanding of job market trends, difficulty in finding relevant job opportunities, inability to attract and retain top talent, and lack of insight into company branding and reputation.

The purpose of this project is to conduct an analysis of Glassdoor job postings to gain insights into current and emerging job market trends, identify in-demand skills and experience, and understand how employers can improve their employer branding and reputation to attract and retain top talent

1.2 Purpose

The purpose of this project is to use data analysis skills and tools to explore the job market for various positions in India based on data from glassdoor.com, a popular website for job seekers and employers. The project aims to answer questions such as:

- What are the most common and lucrative roles in India?

- How does the salary vary by location, industry, etc?
- How does the rating of the company affect the salary and satisfaction of the employees?

- What are the best practices and tips for applying and interviewing for jobs in India?

The project also intends to showcase the ability to use web scraping techniques to collect data from online sources, data manipulation and visualization libraries to process and display data.

The project will also provide a web application that allows users to interact with the data and the models through dashboards and charts. The project will also produce a comprehensive report that documents the methodology, findings, and limitations of the data analysis process.

2. LITERATURE SURVEY

2.1 EXISTING PROBLEM

How to scrape the data from Glassdoor.com efficiently and reliably, without violating the terms of service or getting blocked by the website?

How to visualize the data to explore the patterns and trends in the job market for data analysts, such as the distribution of salaries, skills, ratings, and locations?

How to present the findings and insights from the data analysis project in a clear and compelling way, using charts, graphs, tables, and reports?

2.2 REFERENCES

- [Glassdoor Jobs Data Analysis | Kaggle]: This is a dataset of jobs data scraped from Glassdoor.com for a self-learning project. It contains information such as job title, company name, location, salary, rating, industry, sector, etc. You can use this data to perform exploratory data analysis and gain insights into the job market.

(1) Glassdoor Jobs Data Analysis | Kaggle. https://www.kaggle.com/datasets/defrinogionaldo/glassdoor-jobs-data-analysis.

2.3 PROBLEM STATEMENT DEFINITION

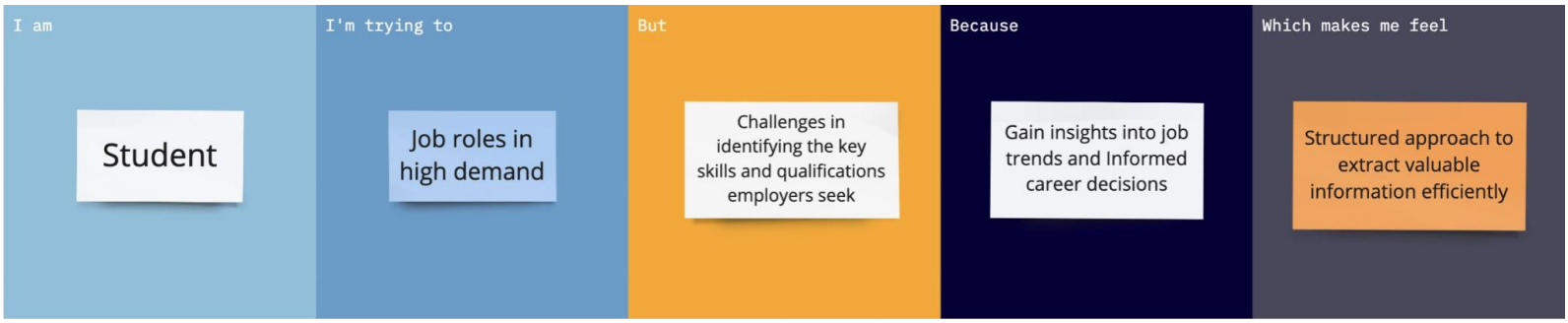
The purpose of this project is to use data analysis skills and tools to explore the job market for various positions in India based on data from glassdoor.com, a popular website for job seekers and employers. The project aims to answer questions such as:

- What are the most common and lucrative roles in India?

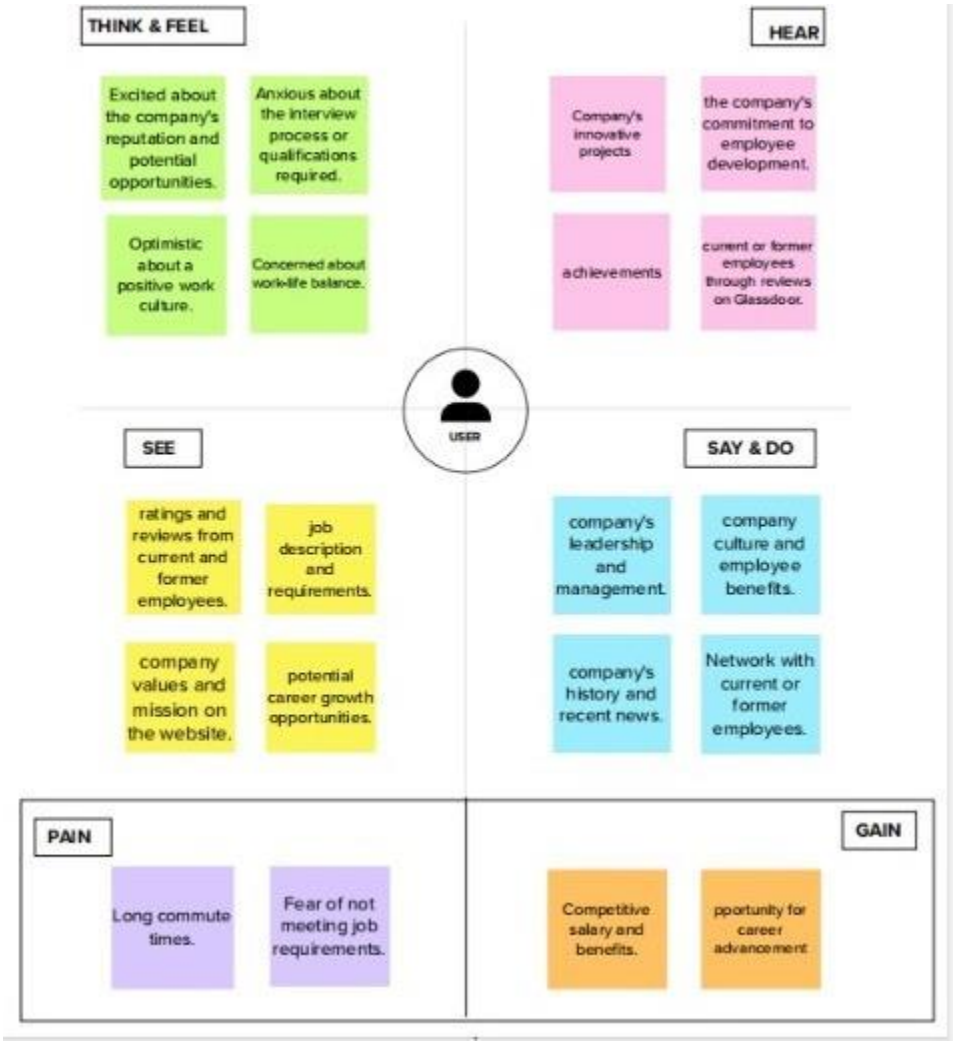
The project also intends to showcase the ability to use web scraping techniques to collect data from online sources, data manipulation and visualization libraries to process and display data, and machine learning algorithms to build and evaluate predictive models. The project will also demonstrate the proficiency in using Python as the main programming language and various frameworks and platforms such as Plotly, Dash, scikit-learn, etc. The project will also provide a web application that allows users to interact with the data and the models through dashboards and charts. The project will also produce a comprehensive report that documents the methodology, findings, and limitations of the data analysis process.

3. IDEATION AND PROPOSED SOLUTION

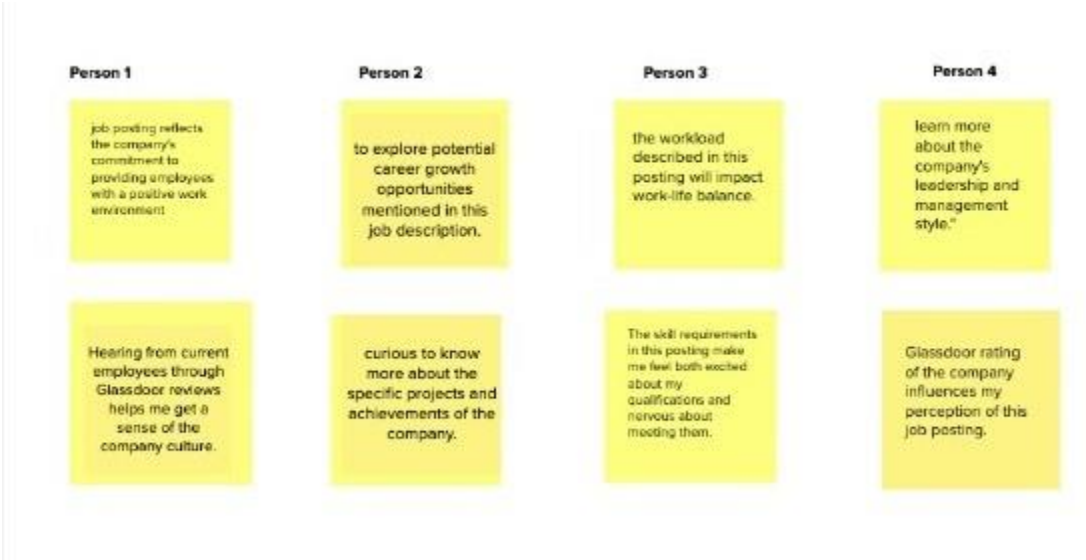
Problem Statement Definition:



3.1 EMPATHY MAP CANVAS

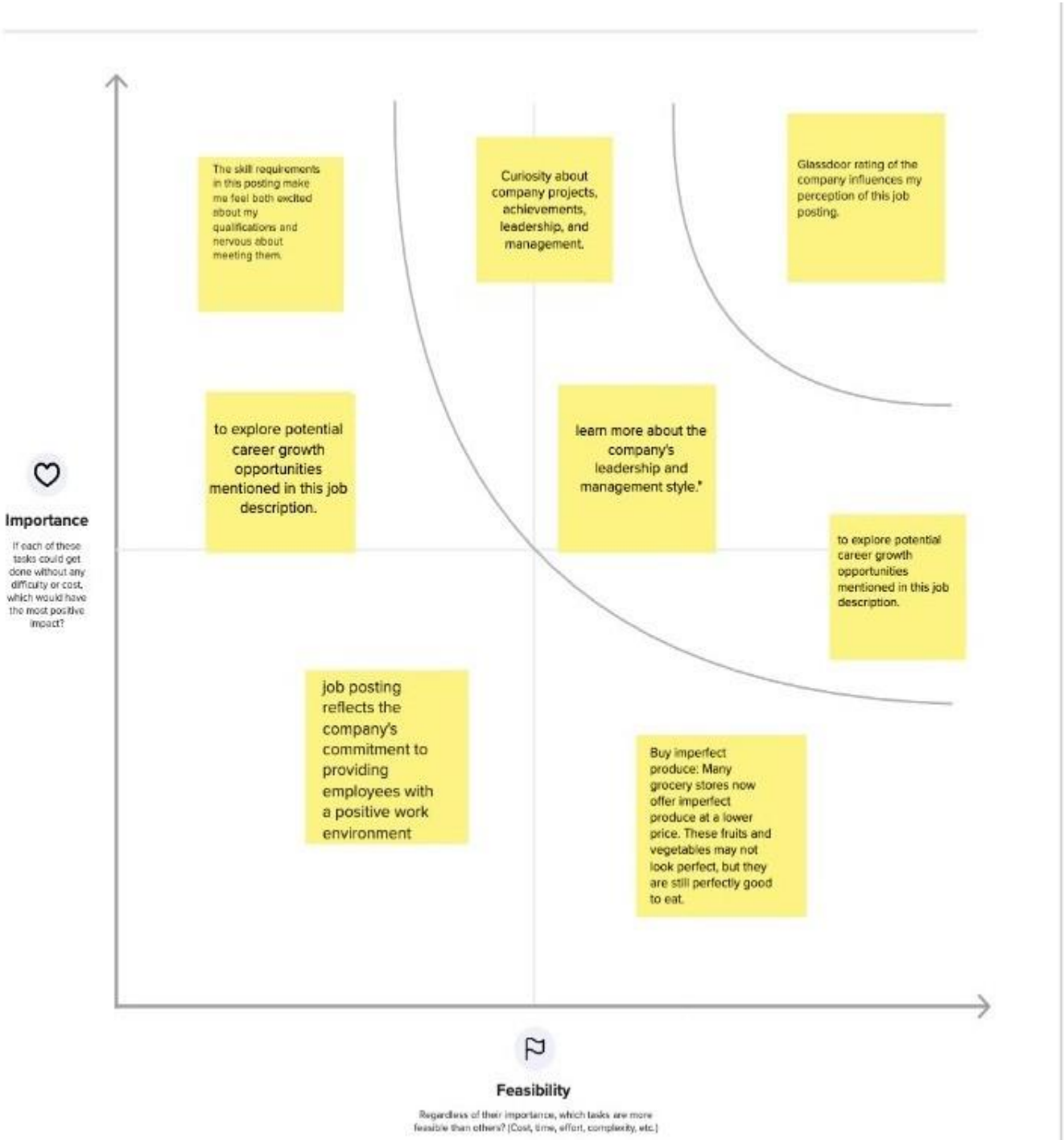


3.2 Ideation & Brainstorming BRAINSTORM:



GROUP IDEAS:

IDEA PRIORITIZATION:



4. REQUIREMENT ANALYSIS:

4.1 Functional requirement

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	**User Registration and profiles**	User registration and profile creation to track job seekers and employers.
FR-2	**Job Listing and Search**	Ability for employers to post job listings with details. Job search functionality with filters for job seekers.
FR-3	**Company Profile**	Detailed company profiles with information like company size, location, reviews, ratings, and benefits.
FR-4	**Job Application**	Application submission and tracking for job seekers. Resume and cover letter upload features.
FR-5	**Review and Rating system**	System for users (current or former employees) to write reviews and rate companies
FR-6	**Salary Information**	Salary and compensation data for job positions. Integration with salary data sources and tools**.**

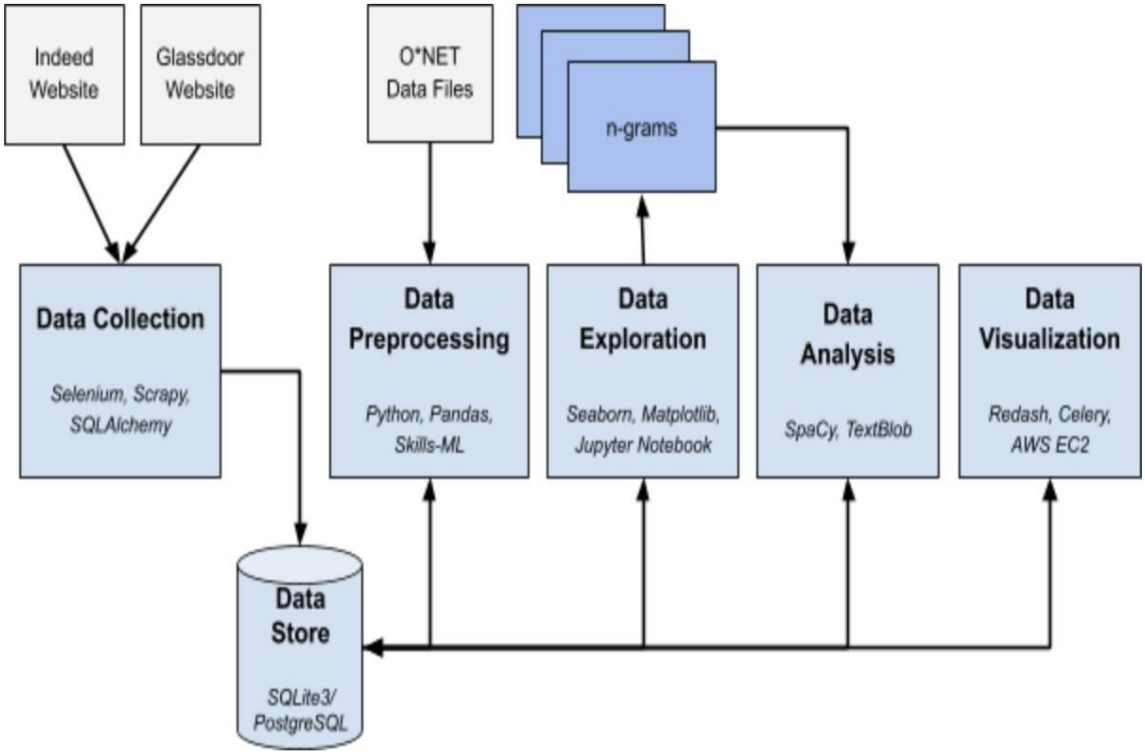
4.2 Non-Functional requirements

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The user interface should be user-friendly and intuitive, catering to a wide range of users.
NFR-2	**Security**	User data, including personal information, should be stored securely and protected against unauthorized

NFR-3	**Reliability**	The system should be available and reliable, with minimal downtime.
NFR-4	**Performance**	The system should be responsive, providing quick search results and analysis of job data.
NFR-5	**Compatibility**	The system should be compatible with various web browsers and mobile devices to reach a broad audience
NFR-6	**Scalability**	The system should be able to handle a growing database of job listings and user profiles.

5. PROJECT DESIGN

5.1 Data Flow Diagrams


Indeed Website
Glassdoor Website
O*NET Data Files
n-grams
Data Collection
Selenium, Scrapy, SQLAlchemy
Data Preprocessing
Python, Pandas, Skills-ML
Data Exploration
Seaborn, Matplotlib, Jupyter Notebook
Data Analysis
SpaCy, TextBlob
Data Visualization
Redash, Celery, AWS EC2
Data Store
SQLite3/ PostgreSQL

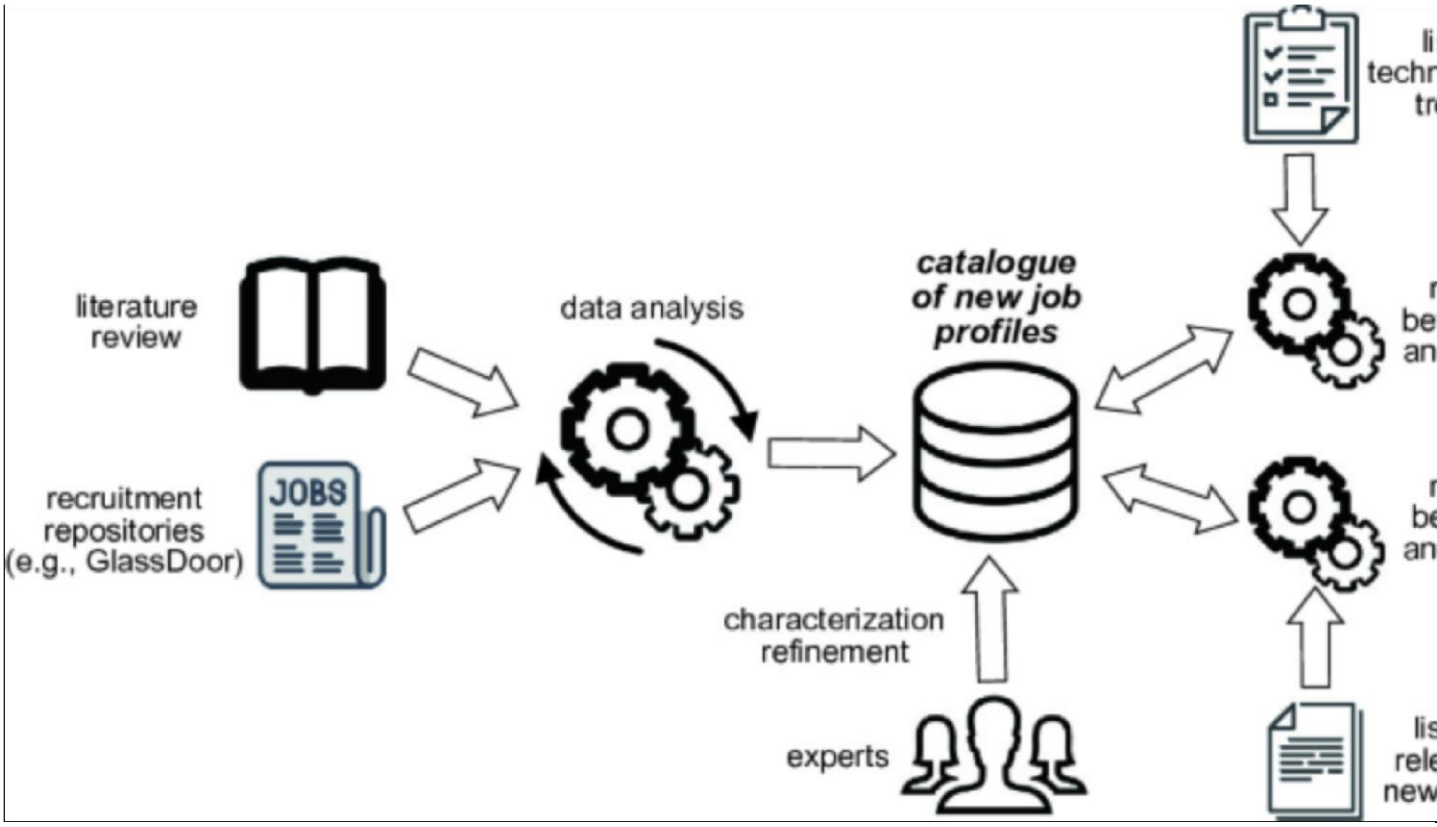
5.1 STORIES:

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
Data Analyst	Data retrieval and processing	USN-1	As a Data Analyst, I want to retrieve job data from Glassdoor so that I can analyze it for trends and insights.	The data should include job descriptions, salaries, ratings, and company information.	High	Sprint-1
		USN-2	As a Data Analyst, I want to analyze job market gaps by comparing job supply and demand to help bridge the imbalance.	Recommendations are provided for bridging the gap through training, education, or talent acquisition	Medium	Sprint-1
		USN-3	As a Data Analyst, I want to clean and preprocess Glassdoor job data to ensure accuracy and consistency in my analysis.	The cleaned data is stored in a separate file or database.	High	Sprint-1
		USN-4	As a Data Analyst, I want to analyze Glassdoor job data to identify trends, high-demand roles, and salary insights.	The analysis results are presented in a clear and visually appealing format (e.g., charts, graphs, reports).	High	Sprint-2
Job Seeker	Job Search and Evaluation	USN-5	As a Job Seeker, I want to search for jobs on Glassdoor and evaluate them based on company culture, benefits, and employee reviews.	Company culture information includes values, mission, and employee testimonials.	High	Sprint-1
HR Manager	Competitive Analysis	USN-6	As an HR Manager, I want to perform competitive analysis by comparing our company's Glassdoor ratings and reviews with industry competitors.	The analysis is presented in a format that allows easy interpretation.	Medium	Sprint-2

5.2 SOLUTION ARCHITECTURE:



6.PROJECT PLANNING AND SCHEDULING:

6.1 TECHNICAL ARCHITECTURE:

6.2 SPRINT PLANNING AND ESTIMATION:

6.3 SPRINT DELIVERY SCHEDULE:

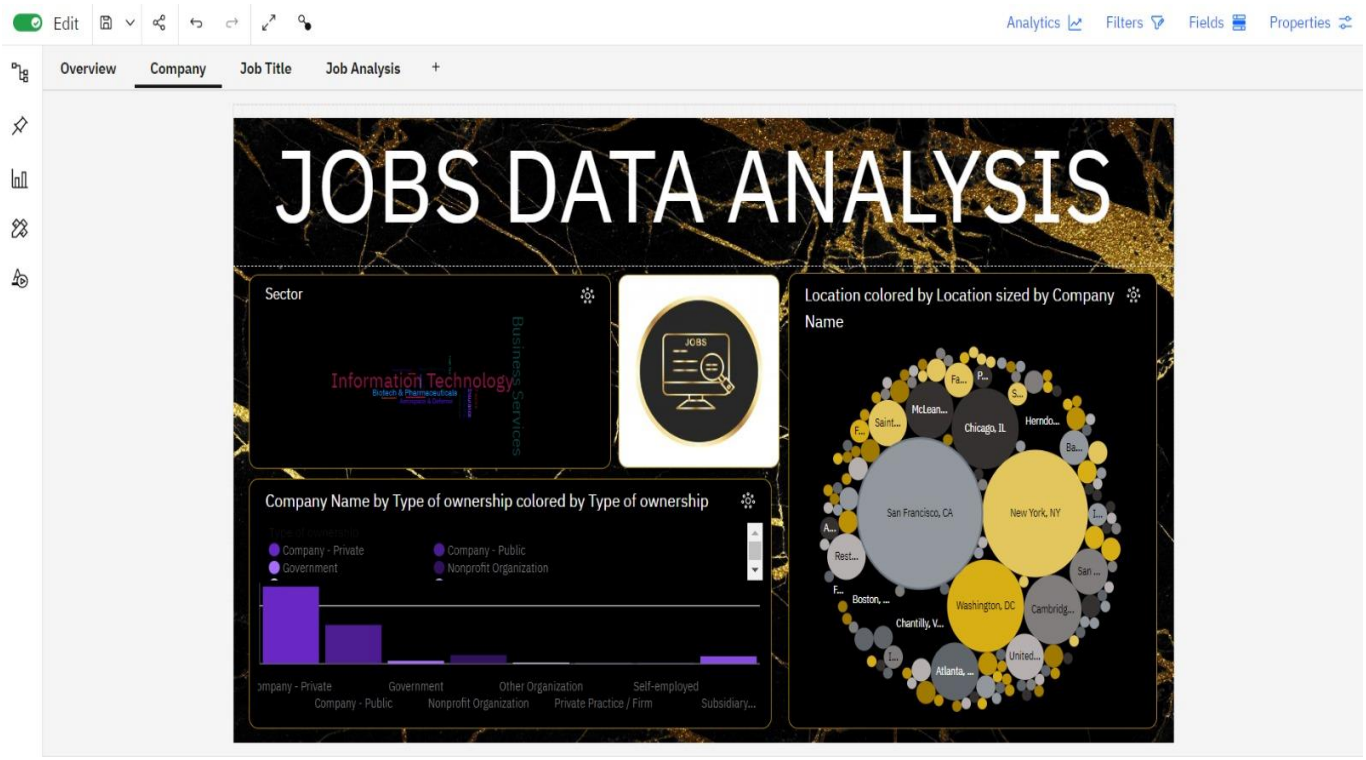
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Data retrieval and processing	USN-1	As a Data Analyst, I want to retrieve job data from Glassdoor so that I can analyze it for trends and insights.	2	High	4
Sprint-1		USN-2	Analyst, I want to analyze job market gaps by comparing job supply and demand to help bridge the imbalance.	1	High	4
Sprint-2		USN-3	As a Data Analyst, I want to clean and preprocess Glassdoor job data to ensure accuracy and consistency in my analysis.	3	High	4
Sprint-2		USN-4	As a Data Analyst, I want to analyze Glassdoor job data to identify trends, high-demand roles, and salary insights.	3	Medium	4
Sprint-3	Job search and evaluation	USN-5	I want to search for jobs on Glass door and evaluate them based on company culture, benefits, and employee reviews.	3	High	4
Sprint-4	HR Manager	USN-6	I want to perform competitive analysis by comparing our company's Glassdoor ratings and reviews with industry competitors.	2	Medium	4

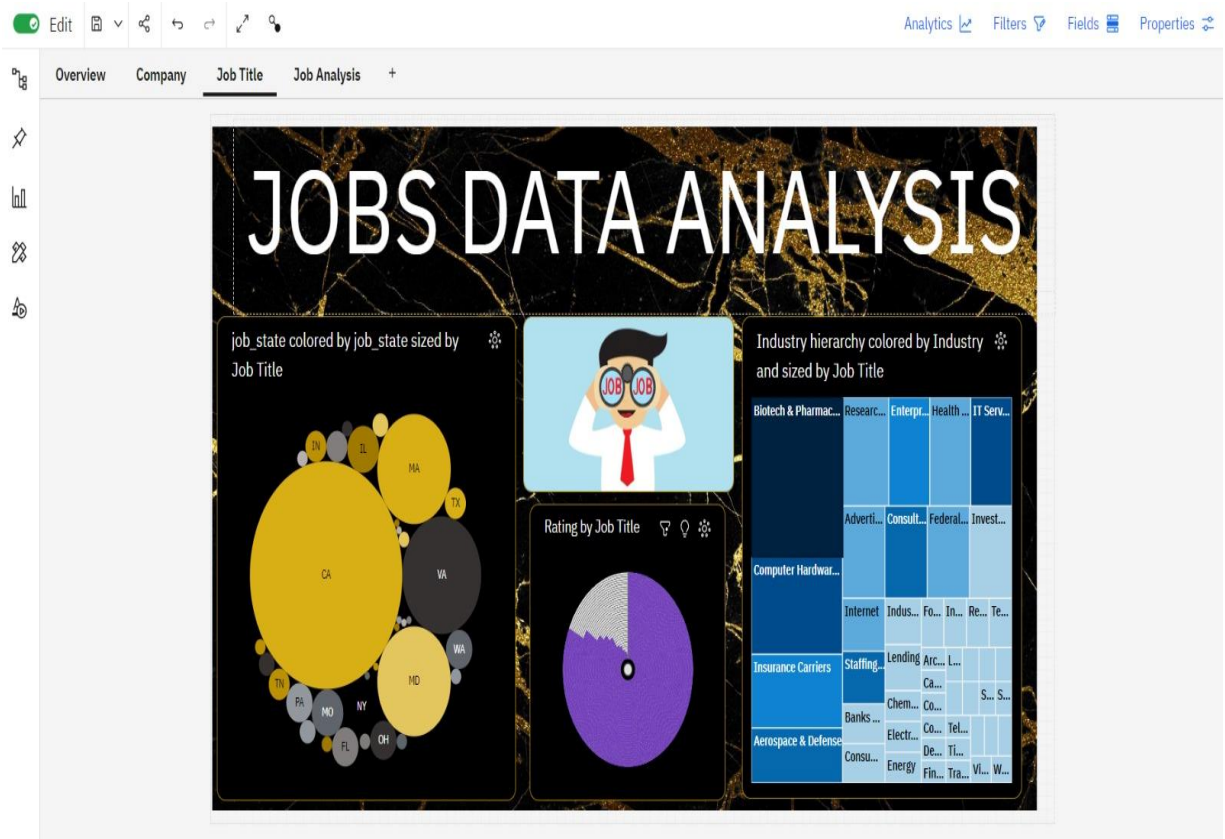
8.PERFORMANCE TESTING:

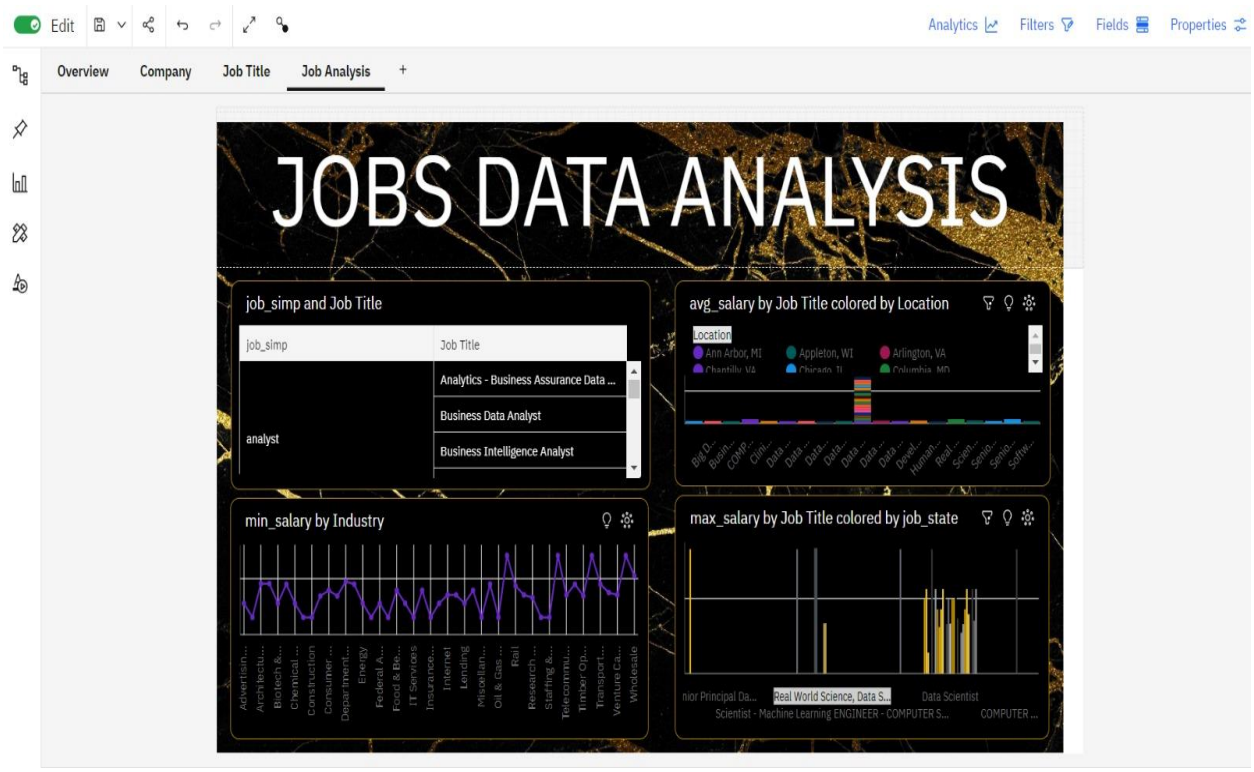
8.1 PERFORMANCE METRICS:

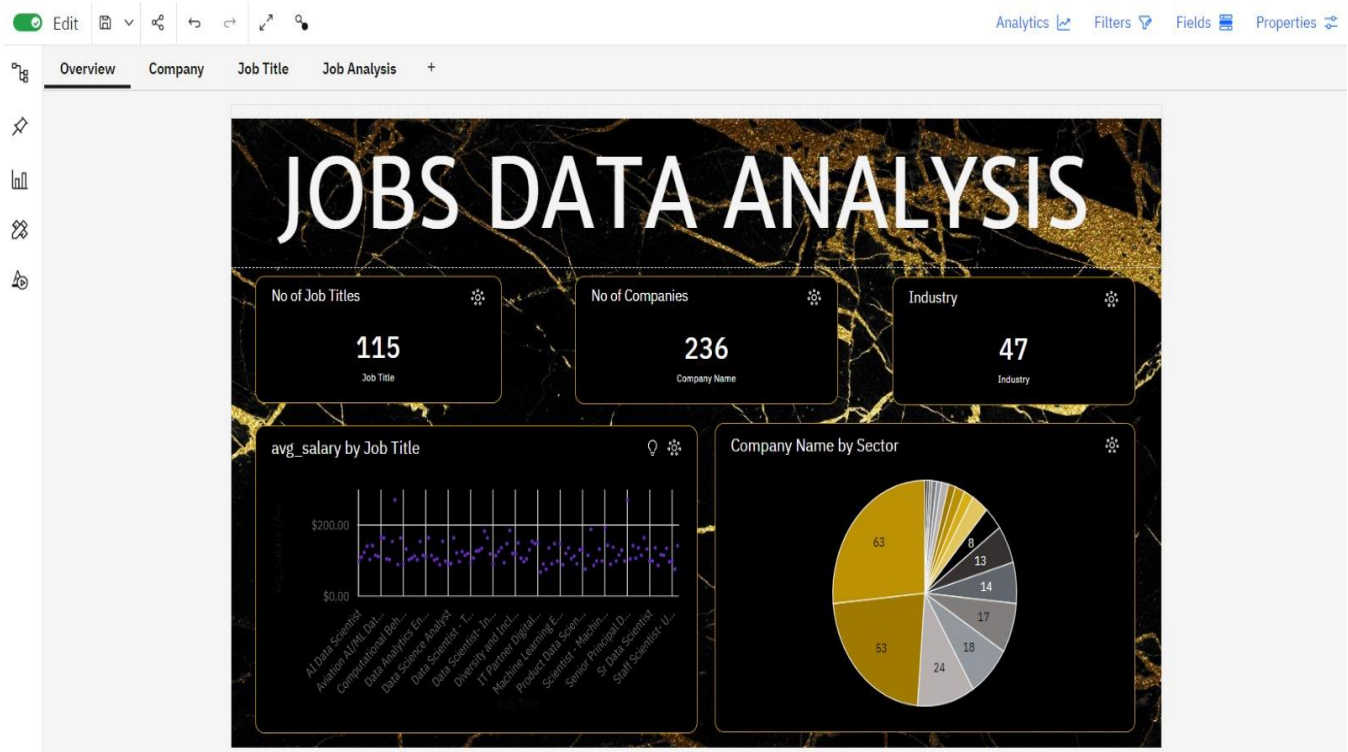
S.No.	Parameter	Screenshot / Values
1.	Dashboard design	No of Visualizations / Graphs - 20
2.	Data Responsiveness	20
3.	Amount Data to Rendered (DB2 Metrics)	NA
4.	Utilization of Data Filters	6
5.	Effective User Story	No of Scene Added -
6.	Descriptive Reports	No of Visulizations / Graphs -

9.RESULTS: 9.1 OUTPUT SCREENSHOTS:






Edit
Analytics
Filters
Fields
Properties
Overview
Company
Job Title
Job Analysis
JOBS DATA ANALYSIS
job_simp and Job Title
job_simp
Job Title
analyst
Analytics - Business Assurance Data ...
Business Data Analyst
Business Intelligence Analyst
avg_salary by Job Title colored by Location
Location
Ann Arbor, MI
Appleton, WI
Arlington, VA
min_salary by Industry
max_salary by Job Title colored by job_state
Real World Science, Data S...
Data Scientist
Scientist - Machine Learning ENGINEER - COMPUTER S...
COMPUTER ...


Edit
Analytics
Filters
Fields
Properties
Overview
Company
Job Title
Job Analysis
JOBS DATA ANALYSIS
No of Job Titles
115
Job Title
No of Companies
236
Company Name
Industry
47
Industry
avg_salary by Job Title
$200.00
$0.00
Company Name by Sector
63
8
13
14
17
18
24
53

10.ADVANTAGE & DISADVANTAGES:

ADVANTAGES:

- The project could help job seekers find and apply for their ideal jobs, by providing them with information about the job market trends, the skills and qualifications required, the expectations and satisfaction of the employees, and the reputation and culture of the employers.

- The project could also help employers attract and retain qualified and engaged talent, by providing them with feedback and suggestions on how to improve their company profile, employer branding, and employee engagement.

- The project could also contribute to the academic and social research on the topics of labor economics, human resources, organizational behavior, and career development, by providing a rich and diverse source of data and analysis.

DISADVANTAGES:

Glassdoor relies on user-generated content that may not always be accurate, reliable, or updated. Some users may provide false or biased information, either intentionally or unintentionally, that can affect the quality and validity of the data. You may need to verify and cross-check the data from other sources before using it for your analysis or decision making.

- Glassdoor also faces some privacy and ethical challenges, as it collects and displays personal and sensitive information from its users and the companies. You may need to respect the privacy and confidentiality of the data and comply with the terms and conditions of Glassdoor. You may also need to avoid plagiarism and bias when using or presenting the data.

- Data analysis is a fast-changing and competitive field that requires constant learning and adaptation. You may need to keep up with the latest developments and trends in data science, such as artificial intelligence, machine learning, big data, etc. You may also need to deal with high pressure, dull tasks, and indefinite job roles**.**

11.CONCLUSIONS:

Data scientists job has highest salary compared to other sectors.

The most common sectors that hire data analysts in India are IT Services, Financial Services, Health Care, E-commerce and Renewable Energy

The Financial Services sector is the second most highly performed sector for data analysts in India, as it offers a competitive salary, a diverse range of projects, and a high demand for analytical skills

The Health Care sector is the third most highly performed sector for data analysts in India, as it offers a stable and rewarding career, a social impact, and a growing need for data-driven solutions[13].

The E-commerce sector is the fourth most highly performed sector for data analysts in India, as it offers a dynamic and innovative environment, a large amount of data, and a high potential for growth

12.FUTURE SCOPE:

Developing a web application or a dashboard that can display the analysis results in an interactive and user-friendly way. This can help users to explore different aspects of the job market, such as the most popular skills, industries, locations, salaries, etc. Users can also filter and customize the results according to their preferences and needs.

Applying more advanced data analysis techniques, such as natural language processing, sentiment analysis, topic modeling, etc. This can help to extract more meaningful and nuanced information from the job descriptions, such as the tone, mood, style, keywords, etc. This can also help to identify the hidden patterns and trends in the job market that are not easily observable from the numerical data.

Building a machine learning model that can predict the salary range or the rating of a job posting based on its features, such as the job title, company name, industry, location, skills, etc. This can help job seekers to estimate their worth and negotiate their salary expectations. It can also help employers to set competitive and fair salaries for their employees.

Creating a recommender system that can suggest relevant and suitable job opportunities to job seekers based on their profile, preferences, skills, experience, etc. This can help job seekers to find their ideal jobs faster and easier.

13. APPENDIX

GitHub Project

https://github.com/Mukesh-Veer/Naan-Mudhalvan

Video Demo Link

https://drive.google.com/file/d/1sLpEQfgSjaZ_Svt7ZzOdsa9zZlmr97nE/view?usp=share_link