DATA ANALYTICS

THE FUTURE OF WORK: DATA ANALYSIS OF GLASSDOOR JOBS

Team leader:

SINGANA KEERTHI

Team members:

V.SASI

KANNEMADUGU NIKITHA

ADALAM VINUSHA

CHENNAKAYALA SRAVANI

Contents

INTRODUCTION

* 1. Overview

A brief description about your project

* 1. Purpose

The use of this project. What can be achieved using this.

2 LITERATURE SURVEY

2.1 Existing problem Existing approaches or method to solve this problem

2.2 Proposed solution What is the method or solution suggested by you?

3 THEORITICAL ANALYSIS

3.1 Block diagram Diagrammatic overview of the project.

3.2 Hardware / Software designing Hardware and software requirements of the project

4 RESULT Final findings (Output) of the project along with screenshots.

5 ADVANTAGES & DISADVANTAGES

List of advantages and disadvantages of the proposed solution

6 APPLICATIONS The areas where this solution can be applied

7 CONCLUSION Summarizing the entire work and findings.

8 FUTURE SCOPE Enhancements that can be made in the future.

**PEOJECT REPORT**

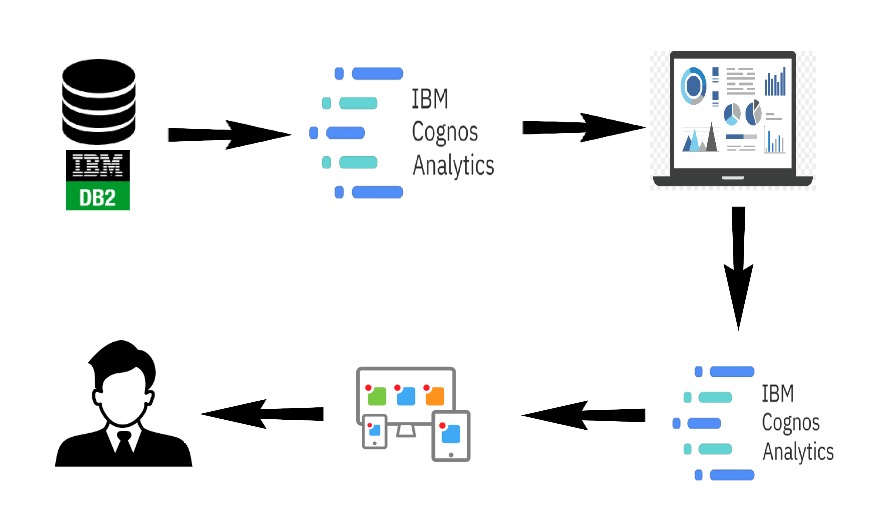
**1.Introduction**

**1.1 Overview**

**THE FUTURE OF WORK: DATA ANALYSIS OF GLASSDOOR JOBS**

The project aims to conduct a comprehensive analysis of Glassdoor job postings, utilizing data mining and natural language processing techniques. It seeks to provide valuable insights into current and emerging job market trends, in-demand skills and qualifications, and employer branding and reputation. By understanding the job market dynamics and employer requirements, the analysis will benefit both job seekers and employers, enabling employers to refine their talent acquisition strategies and attract top talent, while job seekers can enhance their employability and focus on developing sought-after skills. Ultimately, the project's insights aim to contribute to better decision-making and increased efficiency in the job market.

**Technical Architecture:**

****

**1.2 Purpose**

* The analysis helps employers tailor job postings and recruitment efforts to attract candidates with the right qualifications and experience, leading to more effective hiring processes.
* Companies can understand how they are perceived by job seekers and employees, identifying areas of improvement to enhance employer branding and attract top talent.
* Job seekers gain insights into in-demand skills, focusing their efforts on developing relevant expertise, increasing employability.
* HR departments use insights to make informed decisions on talent management, training programs, and compensation packages.
* Organizations gain competitive insights on competitor hiring practices and employer branding efforts.
* Analysis predicts future job market trends, helping with workforce planning and business alignment.
* Understanding factors that retain talent helps implement effective retention strategies, reducing turnover costs and maintaining a stable workforce.

**2.LITERATURE SURVEY**

**2.1 Existing Problem**

The project aims to conduct a comprehensive analysis of Glassdoor job postings, utilizing data mining and natural language processing techniques. It seeks to provide valuable insights into current and emerging job market trends, in-demand skills and qualifications, and employer branding and reputation. By understanding the job market dynamics and employer requirements, the analysis will benefit both job seekers and employers, enabling employers to refine their talent acquisition strategies and attract top talent, while job seekers can enhance their employability and focus on developing sought-after skills. Ultimately, the project's insights aim to contribute to better decision-making and increased efficiency in the job market.

1. Data Quality and Completeness:

The analysis heavily relies on data available on Glassdoor, and any inaccuracies, inconsistencies, or missing data in job postings can impact the reliability and validity of the results

1. Bias in User-Generated Content:

Glassdoor reviews and ratings are user-generated, which may introduce bias or subjective opinions about companies and job roles, affecting the objectivity of the analysis.

1. Limited Data Scope:

The analysis is limited to the data available on Glassdoor, which may not fully represent the entire job market, potentially leading to incomplete or skewed insights.

1. Changing Job Market Dynamics:

The job market is dynamic, and trends can change rapidly. The analysis may provide insights for a specific time frame, but these trends may not remain consistent over an extended period.

**2.2 Proposed Solution**

Proposed Solutions for the Project:

1. Data Validation and Cleaning: Implement rigorous data validation and cleaning processes to ensure the quality and completeness of the data extracted from Glassdoor. This involves identifying and handling any inaccuracies, inconsistencies, or missing data points to enhance the reliability of the analysis.

2. Sentiment Analysis and Bias Mitigation: Apply sentiment analysis techniques to identify and mitigate potential bias in user-generated content on Glassdoor. By quantifying sentiment and subjective opinions, the analysis can be adjusted for any undue influence on the insights.

3. Incorporate Multiple Data Sources: To overcome the limitation of the data scope, supplement Glassdoor data with other relevant job market datasets. Integrating data from various sources can provide a more comprehensive and representative view of the job market trends.

4. Real-Time Data Monitoring: Recognizing the dynamic nature of the job market, establish a system for real-time data monitoring to continuously update the analysis with the latest job postings and market trends, providing timely insights.

5. Privacy Compliance and Ethical Considerations: Ensure compliance with data privacy regulations and address any legal considerations related to data usage from Glassdoor. This includes obtaining appropriate permissions and using the data responsibly and ethically.

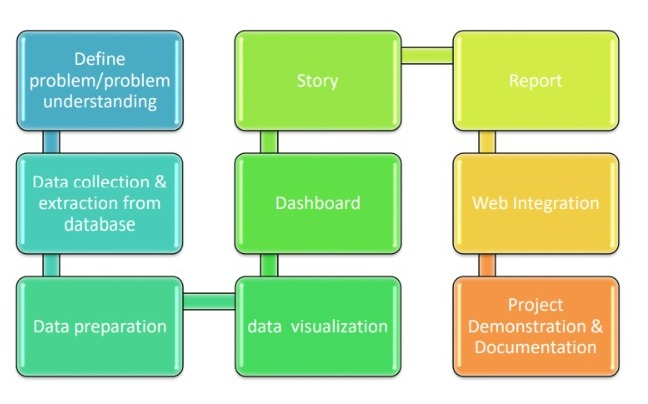
6. Longitudinal Analysis: Conduct a longitudinal analysis to track job market trends over time. This approach can identify patterns and changes in demand for skills and qualifications, offering more insightful predictions for the future.

7. Diversified Job Role Categorization: Develop a flexible system for categorizing non-standard or emerging job roles that may not fit conventional classifications. This enables a more accurate representation and analysis of diverse job opportunities.

8. Expert Review and Validation: Seek expert review and validation of the analysis results to ensure objectivity and accuracy. Involving domain experts can provide valuable insights and enhance the credibility of the findings.

**3.THEORITICAL ANALYSIS**

**3.1 Block diagram**

****

**Diagrammatic overview of the project.**

Define problem/problem understanding: -

1. Specify the business problem.
2. Business requirements.
3. Literature survey.
4. Social or business impact.

Data collection & extraction from database: -

Data collection is the process of gathering and measuring information on variables of interest, in an established systematic fashion that enables one to answer stated research questions, test hypothesis, and Evaluate outcomes and generate insights from the data.

Data preparation: -

In this milestone, we will see how to prepare the data for building visualizations.

Data visualization: -

Data visualization is the process of creating graphical representations of data in order to help peoples understand and explore the information. The goal of data visualization is to make complex data sets more accessible, intuitive, and easier to interpret. By using visual elements such as charts, graphs, and maps, data visualizations can help people quickly identify patterns, trends, and outliers in the data.

Dashboard: -

A dashboard is a graphical user interface (GUI) that displays information and data in an organized, easy-to-read format. Dashboards are often used to provide real-time monitoring and analysis of data, and are typically designed for a specific purpose or use case. Dashboards can be used in a variety of settings, such as business, finance, manufacturing, healthcare, and many other industries. They can be used to track key performance indicators (KPIs), monitor performance metrics, and display data in the form of charts, graphs, and tables.

Story: -

A data story is a way of presenting data and analysis in a narrative format, with the goal of making the information more engaging and easier to understand. A data story typically includes a clear introduction that sets the stage and explains the context for the data, a body that presents the data and analysis in a logical and systematic way, and a conclusion that summarizes the key findings and highlights their implications. Data stories can be told using a variety of mediums, such as reports, presentations, interactive visualizations, and videos.

Report: -

A report is a document that presents information in a specific format and layout, usually based on data from a database or other data source. A report in IBM Cognos can contain various elements, such as tables, charts, graphs, and images, as well as text and data elements, and it is designed to be used by business users to help them better understand their data and make informed decisions. There are several different types of reports available in IBM Cognos, including list reports, crosstab reports, chart reports, and report studio reports, among others. The type of report that you choose will depend on the specific needs and requirements of your organization, as well as the data that you need to present

Web integration: -

Publishing helps we to track and monitor key performance metrics, to communicate results and progress. Help a publisher stay informed, make better decisions, and communicate their performance to others.

Project demonstration & documentation: -

Activity 1: - Record explanation Video for project end to end solution.

Activity 2: - Project Documentation-Step by step project development procedure

**3.2 Hardware / Software designing**

The hardware and software requirements for a project involving the estimation and prediction of hospitalization and medical care costs using data analytics can vary depending on the scale of the project, the size of the dataset, and the complexity of the analysis. Here are some general hardware and software requirements for such a project:

**Hardware Requirements:**

1. Computer: A relatively powerful computer with sufficient processing power, memory (RAM), and storage capacity is essential for handling large datasets and running computationally intensive data analytics tasks.

2. Storage: Ample storage space is needed to store the dataset, intermediate results, and model files. Consider using SSDs or high-capacity HDDs for faster data access.

**Software Requirements:**

**1. Data Management and Preprocessing:**

A. **Python or R**: Both Python and R are popular programming languages for data analysis. They have a wide range of libraries and packages for data manipulation, cleaning, and preprocessing.

B. **Pandas (Python) or Data table (R):** These libraries offer efficient data structures and functions for data manipulation and cleaning.

C. **NumPy (Python) or Matrix (R):** For numerical operations and handling large arrays or matrices.

D. **Data Visualization:** Libraries such as Matplotlib (Python) or ggplot2 (R) for creating insightful data visualizations.

E. **IBM Cognos:** IBM Cognos is a web based integrated business intelligence suite by IBM. It provides a toolset for reporting, analytics, score carding, and monitoring of events and metrics.

**2. Integrated Development Environment (IDE):**

A**. Jupyter Notebook:** For interactive data analysis and exploration.

B. **R Studio (R) or Visual Studio Code (Python):** For a comprehensive code development and debugging environment.

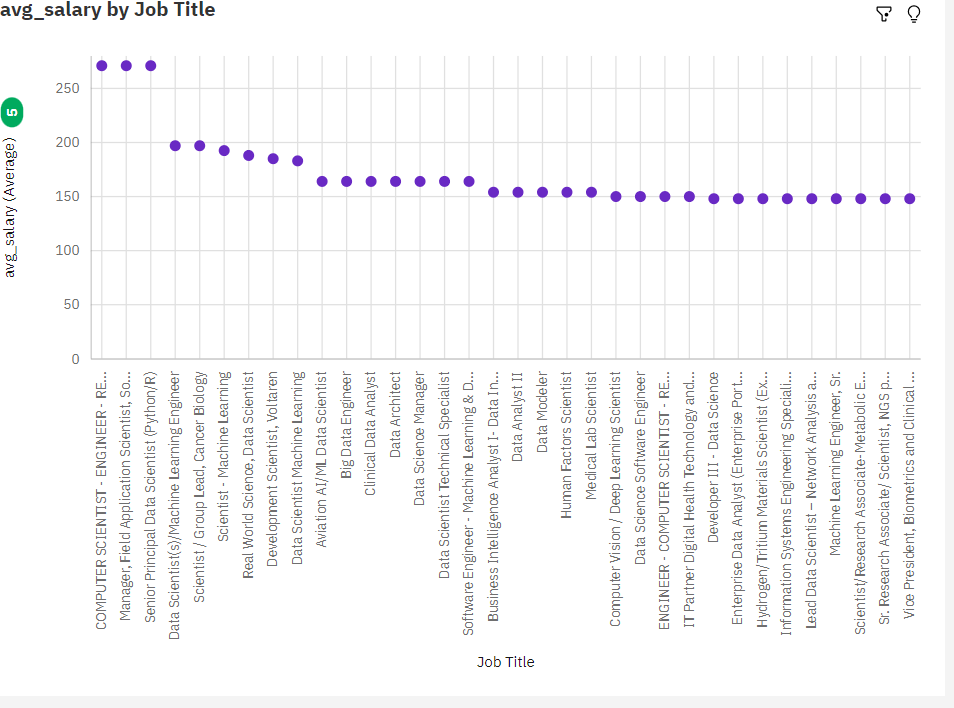
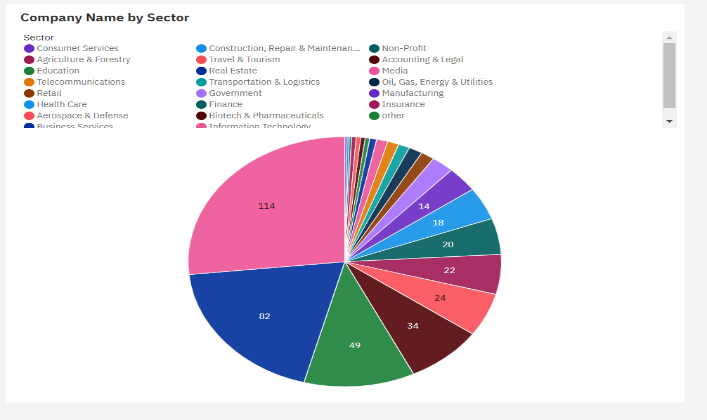
C. **Anaconda:** Anaconda software helps you create an environment for many different versions of python and packages version.

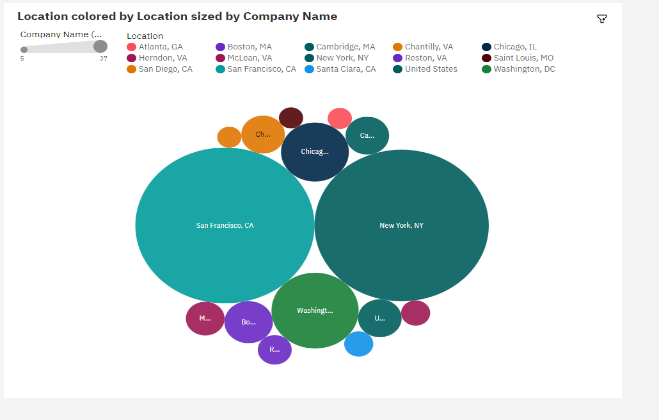
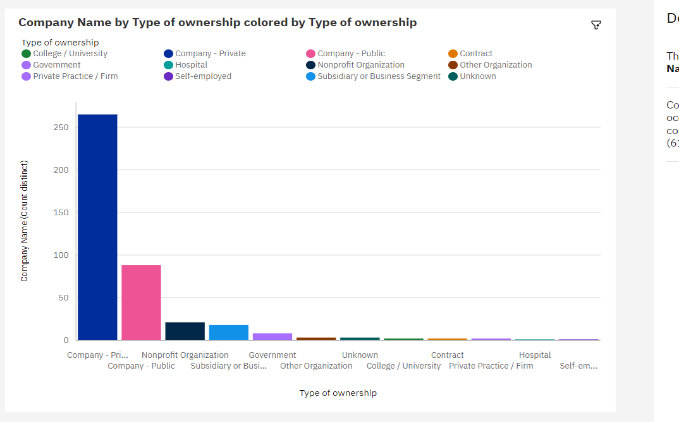
**3. Project Management and Collaboration Tools:**

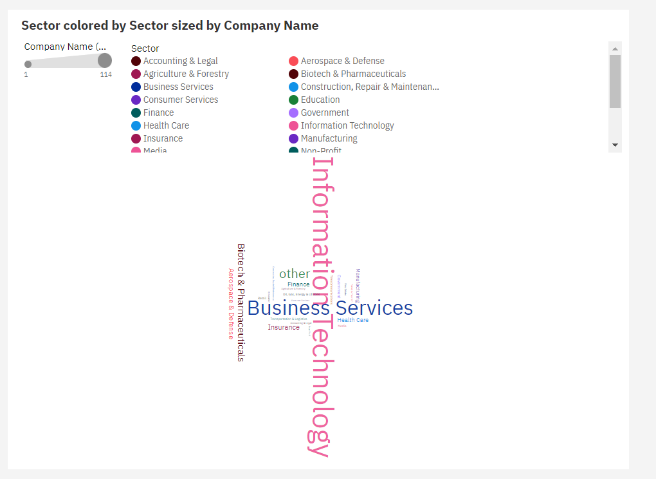
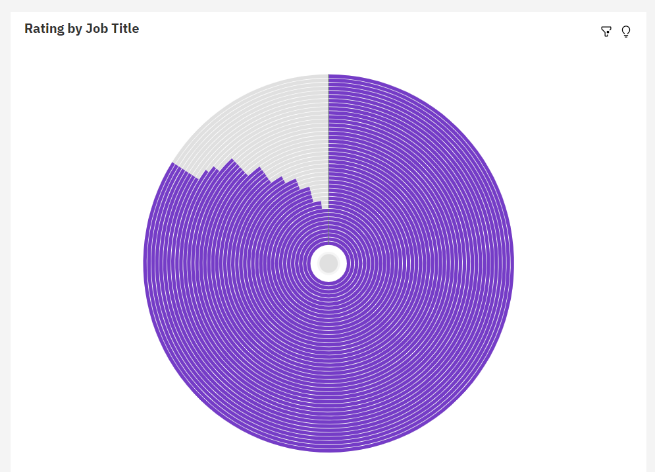
• Communication Tools: Microsoft Teams, SmartInternz, APLMS, zoom, or other communication platforms for team collaboration. These hardware and software requirements provide a foundation for building and executing a data analytics project for the estimation and prediction of hospitalization and medical care costs. Depending on the specific needs and complexity of the project, additional tools and resources may be necessary. It's essential to consider scalability, performance, and data security while selecting the hardware and software components for the project.

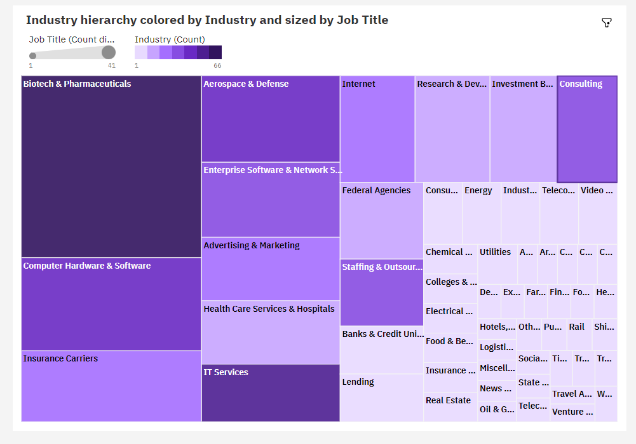
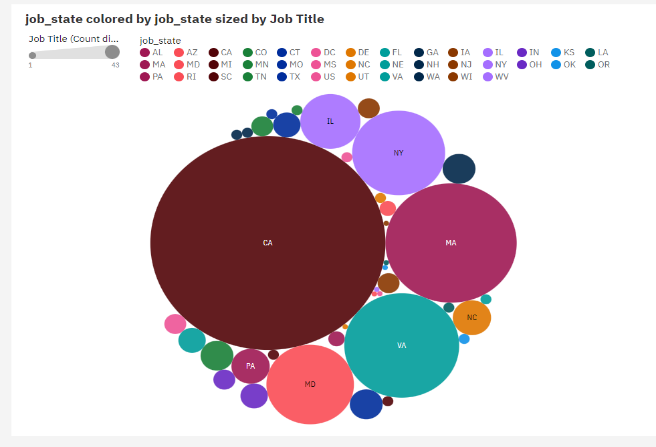
**VISUALIZATION**

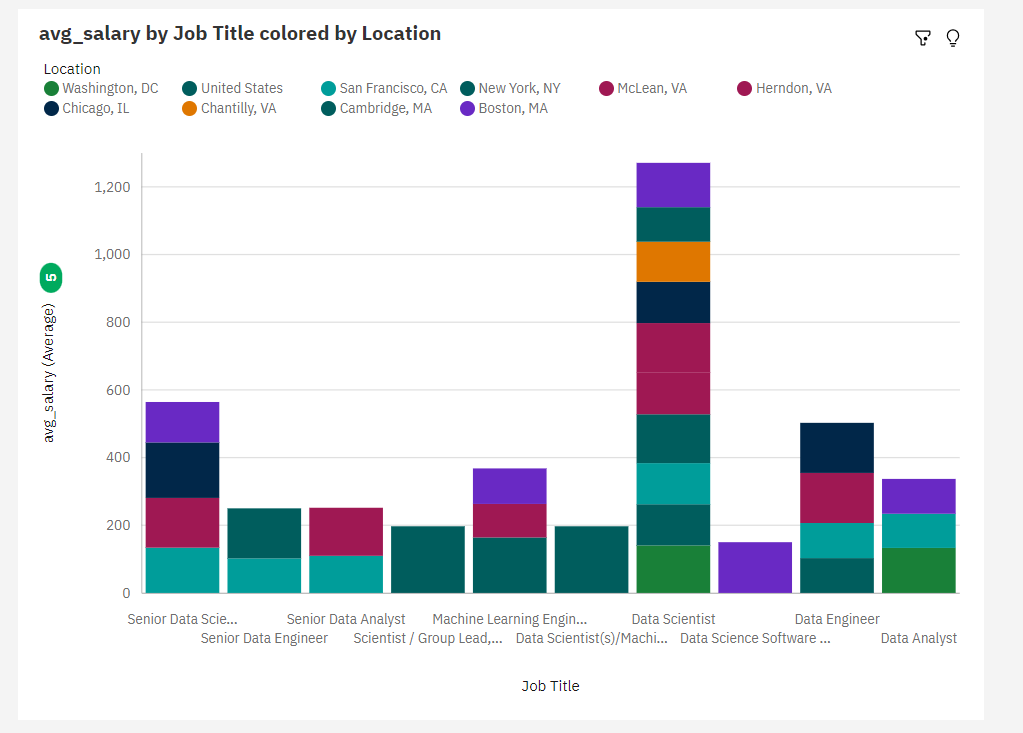
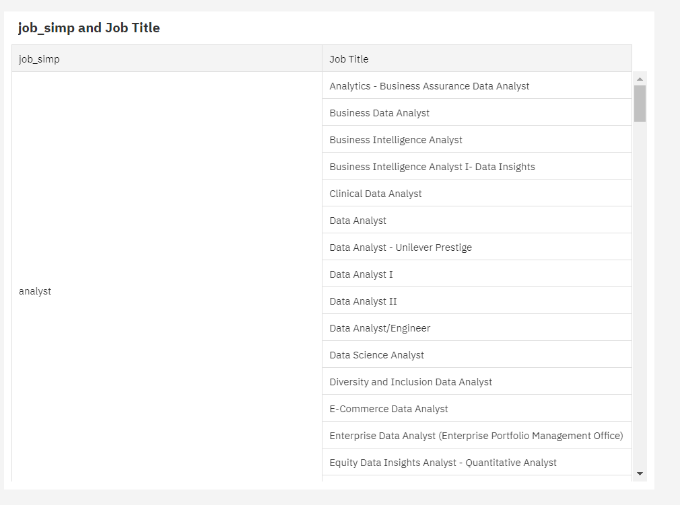
**RESULTS**

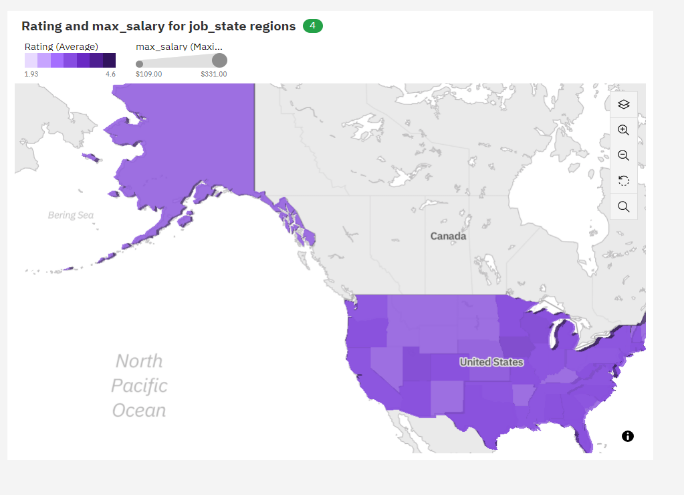
 

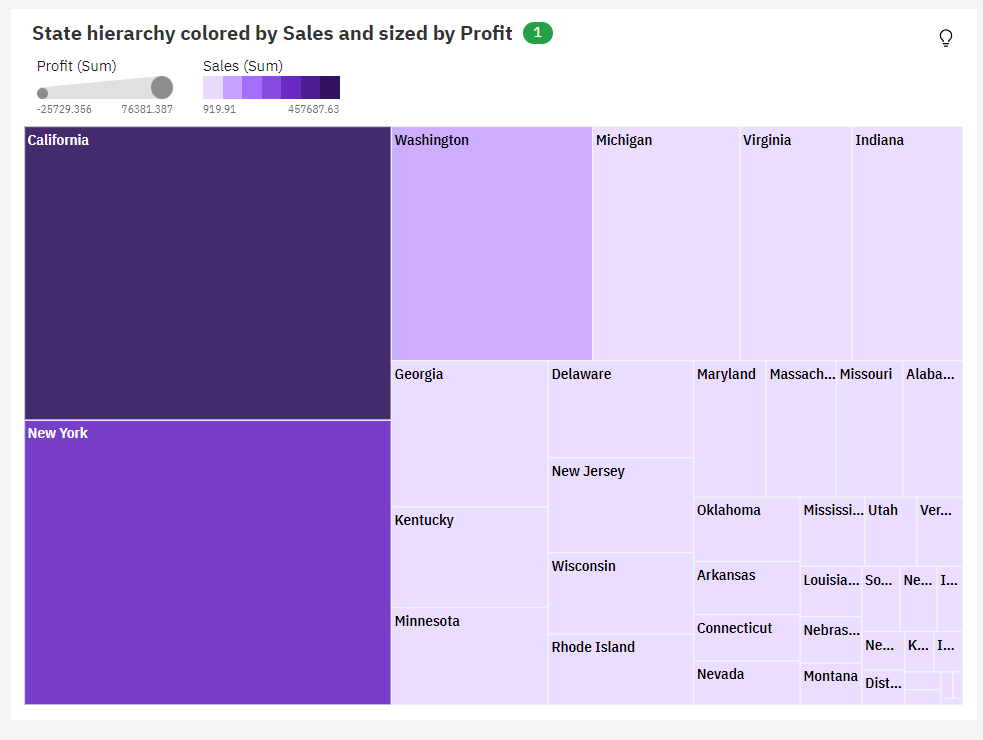
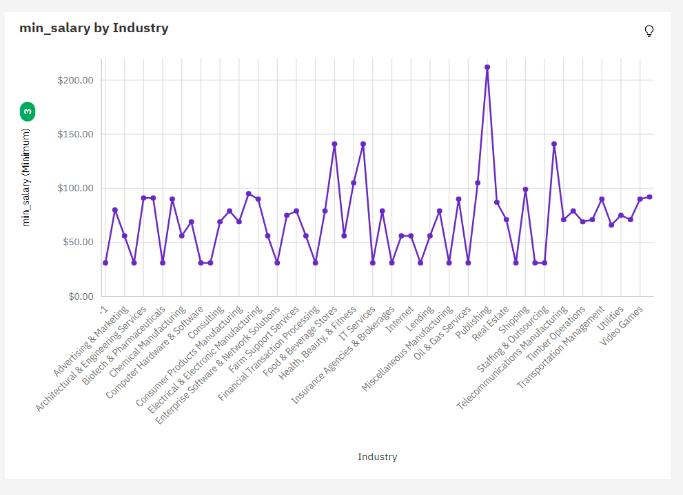
 

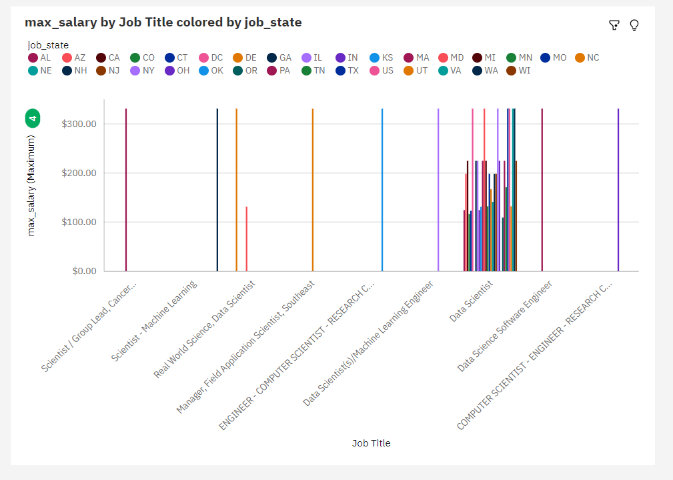
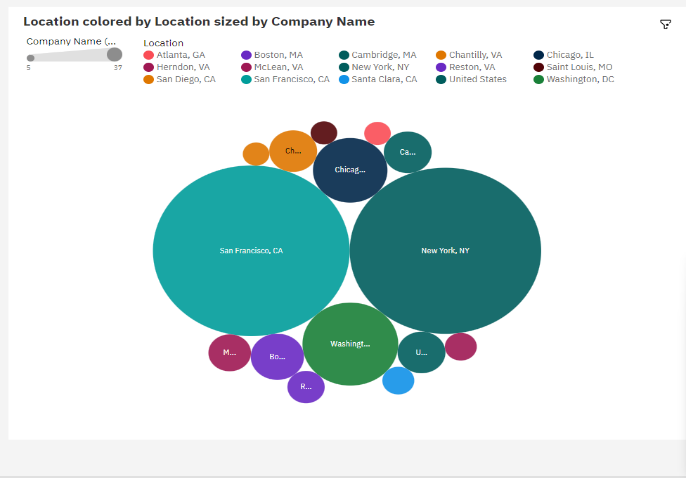
 

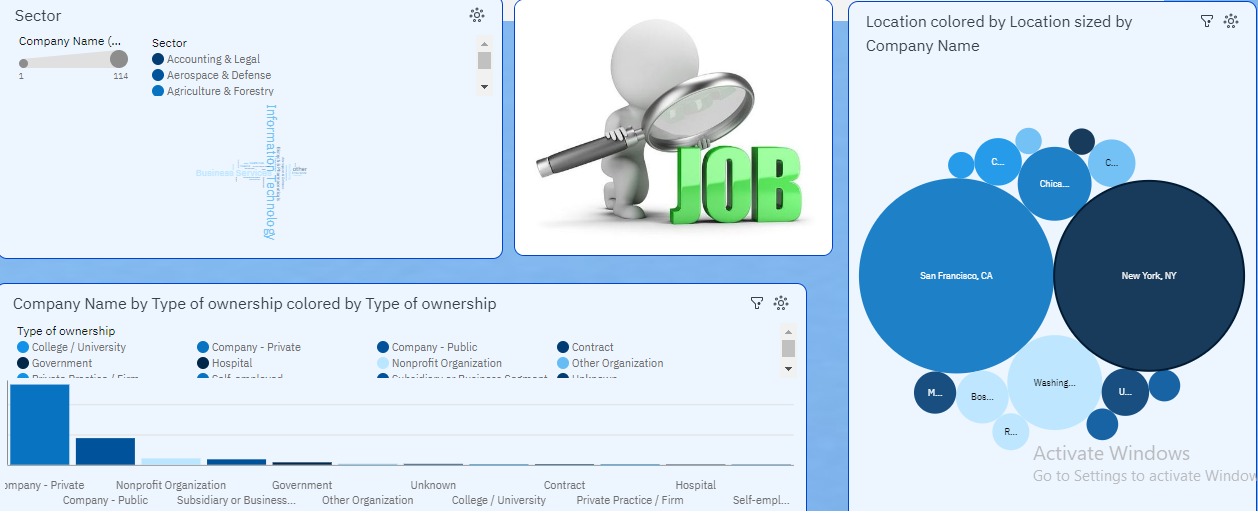
 

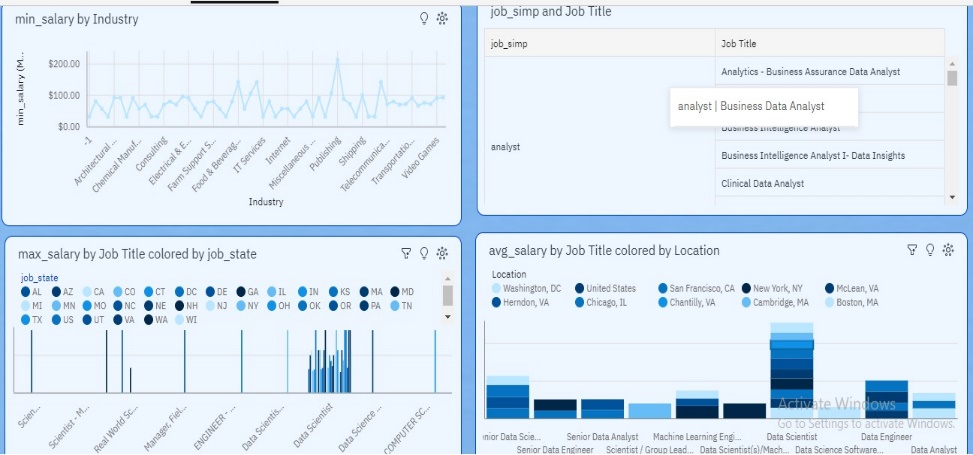
 

**DASH BOARD**

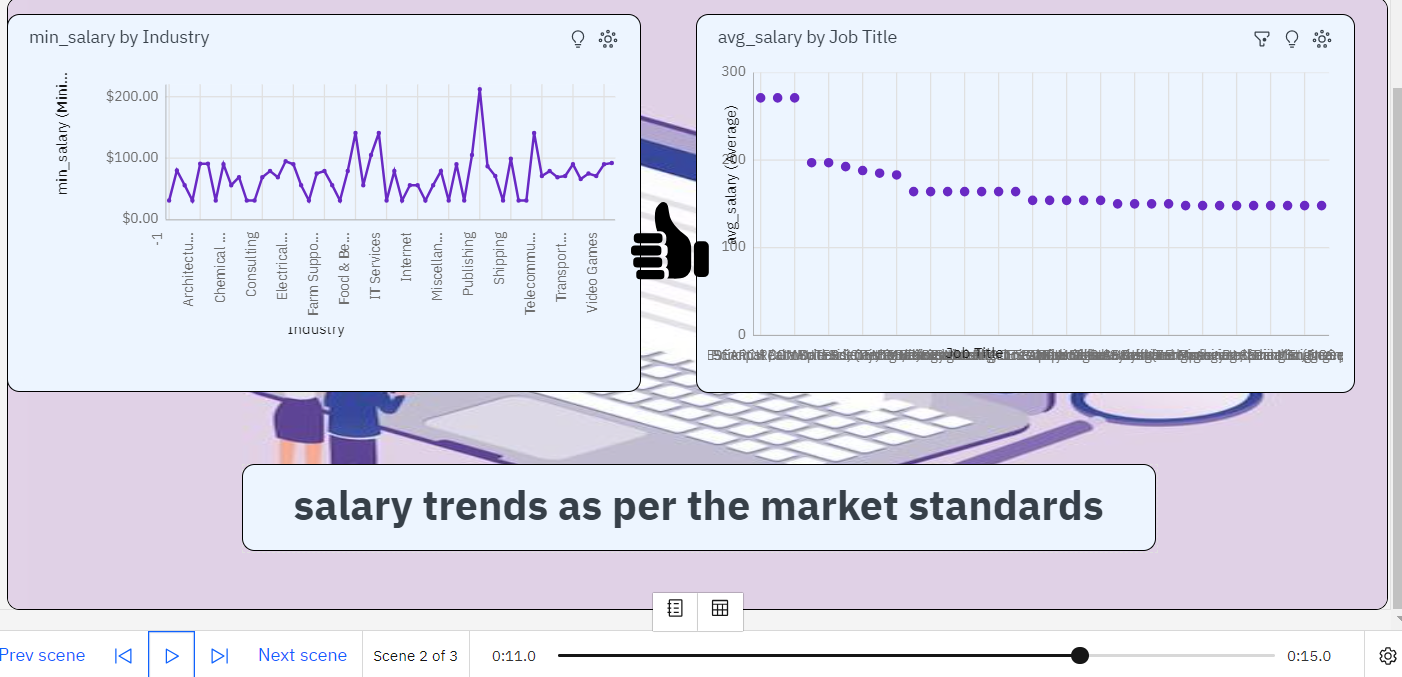


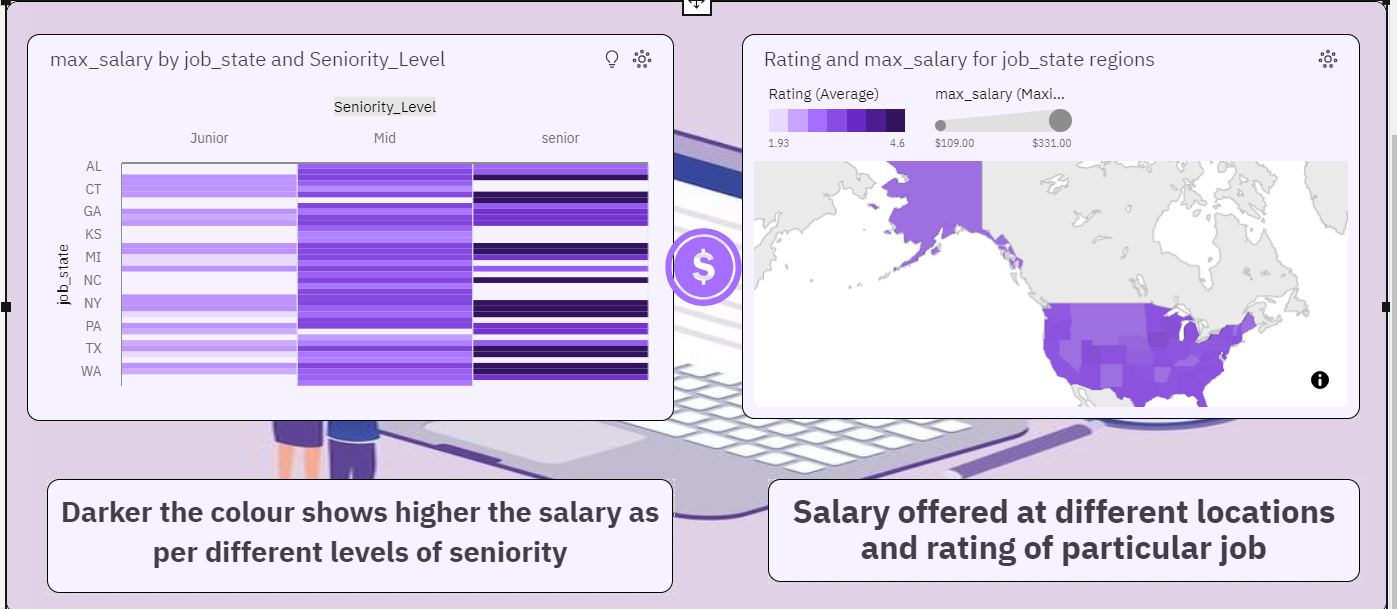




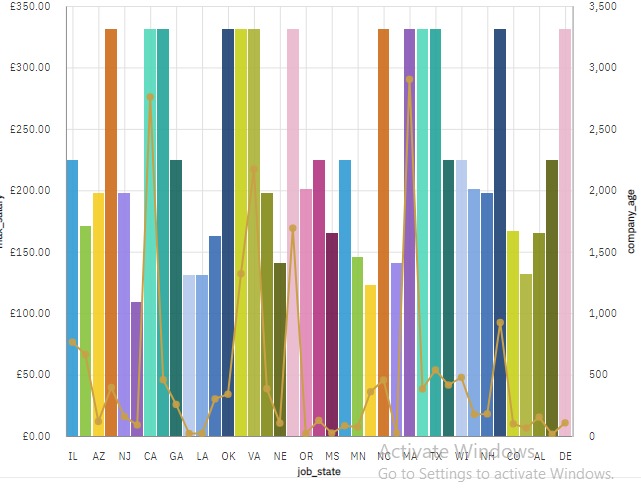
**STORY**

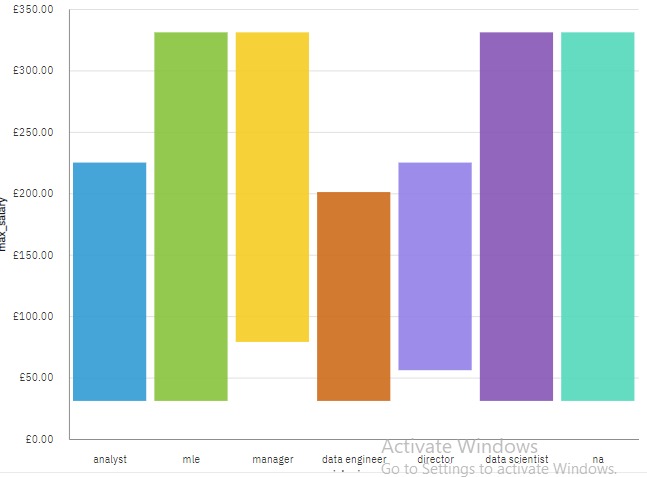






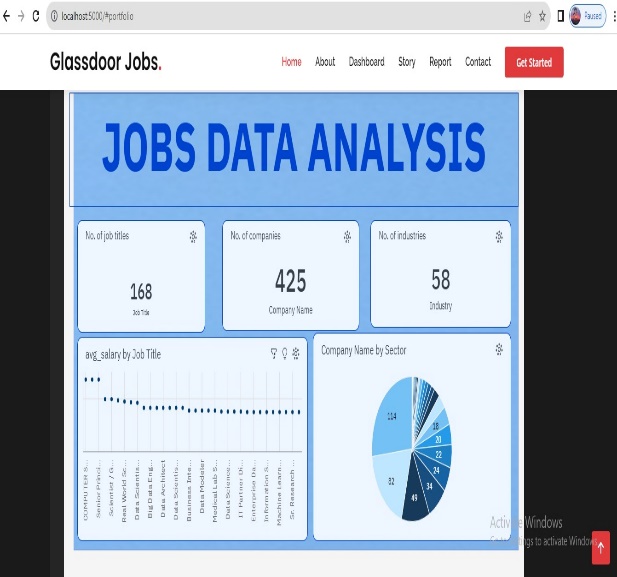
**REPORT**

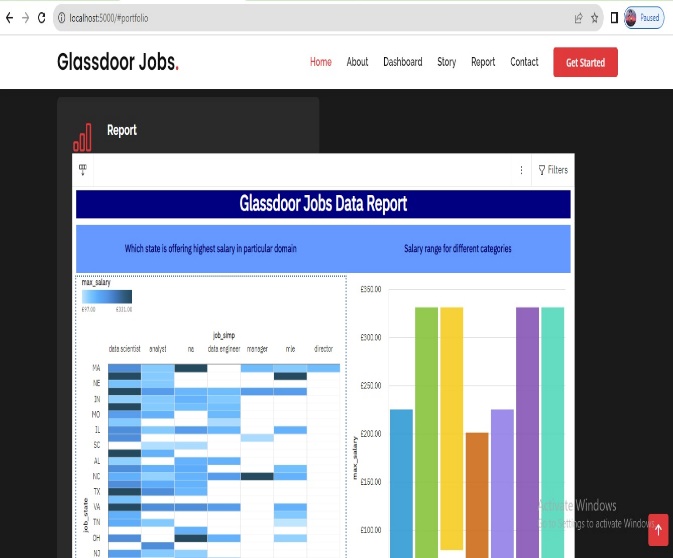
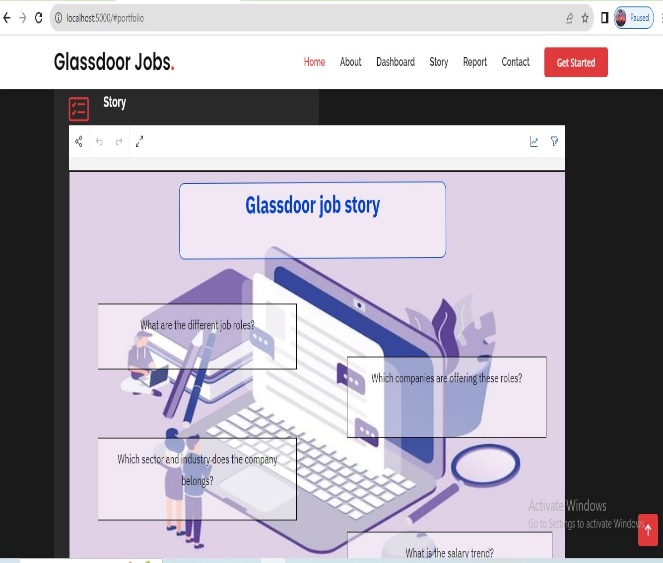






WEB INTEGRATION

**5.ADVANTAGES & DISADVANTAGES**

**Advantages of the Project:**

1. **Data-Driven Insights:** The project's analysis relies on data from Glassdoor, providing objective and data-driven insights into job market trends, in-demand skills, and employer branding, enabling informed decision-making.

2. **Talent Acquisition Optimization:** Employers can leverage the analysis findings to optimize their talent acquisition strategies, tailoring job postings and recruitment efforts to attract the most suitable candidates for specific job roles.

3. **Enhanced Employer Branding:** Understanding employer branding perceptions allows companies to improve their image and reputation, making them more appealing to potential candidates and increasing their ability to attract top talent.

4. **Improved Employability for Job Seekers:** Job seekers gain valuable insights into the skills and qualifications sought by employers, enabling them to focus on developing relevant expertise, increasing their employability and competitiveness in the job market.

5. **Strategic HR Decision-Making:** HR departments can utilize the analysis results for strategic decision-making related to talent management, training programs, compensation packages, and overall human resource planning.

6. **Competitive Advantage for Companies:** Access to market intelligence on competitor hiring practices and branding efforts provides companies with a competitive advantage in attracting and retaining talent.

7. **Forecasting Future Job Trends:** The project's longitudinal analysis can help predict future job market trends, allowing companies to align their business strategies with emerging market demands.

8. **Efficient Employee Retention Strategies:** Understanding factors that retain talent helps companies implement effective retention strategies, reducing turnover costs and maintaining a stable and satisfied workforce.

9. **Real-Time Monitoring Capability:** Implementing real-time data monitoring ensures that the analysis remains up-to-date, providing the latest job market insights for immediate use.

**Disadvantages of the Project:**

1. **Data Limitations and Bias:** The analysis heavily relies on data available on Glassdoor, which may not fully represent the entire job market. Additionally, user-generated content may introduce bias and subjective opinions, potentially affecting the objectivity of the analysis.

2. **Data Privacy Concerns:** Accessing and using data from Glassdoor requires adherence to data privacy regulations and ethical considerations. There might be legal restrictions on the use of such data for commercial purposes, leading to potential challenges in data collection and usage.

3. **Inaccuracies in User-Generated Content:** Glassdoor reviews and ratings are based on user-generated content, which may contain inaccuracies or false information, impacting the reliability of the insights derived from the analysis.

4. **Challenges in Data Cleaning and Validation:** Preprocessing the data to ensure its quality and completeness can be time-consuming and require expertise, especially when dealing with large datasets.

5. **Dynamic Nature of the Job Market:** The job market is continuously evolving, and trends may change rapidly. The analysis may provide insights for a specific time frame, but those trends may not remain consistent over an extended period.

6. **Subjectivity in Job Analysis:** Job analysis itself involves some degree of subjectivity, as job roles and requirements may be interpreted differently by individuals or organizations, potentially influencing the analysis outcomes.

7. **Resource Intensive:** Implementing data mining and natural language processing techniques can be computationally intensive, requiring significant computational resources and expertise.

8. **Non-Standard Job Roles:** Some job roles may be non-standard or emerging, making it challenging to categorize and analyse them effectively within conventional frameworks.

9. **Limitations in Real-Time Data:** Glassdoor data might not be real-time, and delays in updating the analysis with the latest information could impact the accuracy and relevance of immediate job market insights.

10. **Dependence on Glassdoor's Platform:** The analysis is contingent on the availability and stability of Glassdoor's platform. Any changes or disruptions to the platform could affect data collection and subsequent analysis.

**6. APPLICATIONS**

**Applications of the Project:**

1. **Talent Acquisition and Recruitment:** The analysis findings can be directly applied by employers to optimize their talent acquisition and recruitment strategies. Tailoring job postings and identifying in-demand skills enables them to attract the right candidates more effectively.

2. **Skill Development and Training Programs:** Educational institutions and training providers can utilize the analysis insights to design skill development programs that align with the current and future job market demands, improving graduates' employability.

3. **HR Strategy and Workforce Planning:** HR departments can leverage the analysis results to make informed decisions on workforce planning, talent management, and compensation strategies, leading to more efficient HR practices.

4. **Employer Branding and Employee Retention:** The understanding of employer branding perceptions allows companies to make necessary improvements and enhance their reputation to attract and retain top talent.

5. **Market Intelligence and Competitor Analysis:** Companies can gain market intelligence on competitor hiring practices and branding efforts, enabling them to benchmark their performance and develop competitive strategies.

6. **Job Seeker Empowerment:** Job seekers can utilize the analysis insights to tailor their job search and skill development efforts, improving their chances of finding relevant job opportunities.

7. **Consulting and Recruitment Services:** HR consulting firms and recruitment agencies can use the analysis findings to offer specialized services to companies seeking talent acquisition and talent management solutions.

8. **Industry Research and Trend Analysis:** The project's findings can contribute to broader industry research and trend analysis, helping policymakers and industry leaders make informed decisions about workforce development and economic planning.

9. **Employee Satisfaction Surveys and Feedback:** Companies can use the analysis to gain insights into employee perceptions and sentiments, leading to targeted employee satisfaction surveys and feedback mechanisms.

10. **Start-up and Small Business Support:** Start-ups and small businesses can benefit from the analysis by understanding the job market demands and employer branding strategies, helping them attract the right talent in a competitive landscape.

**7. CONCLUSION**

The project aimed to conduct a comprehensive analysis of Glassdoor job postings to gain valuable insights into the current and emerging job market trends, identify in-demand skills and experience sought by employers, and offer recommendations to enhance employer branding and reputation. Through the systematic application of data mining and natural language processing techniques, the analysis provided data-driven and objective insights for both job seekers and employers.

The project's advantages included the ability to optimize talent acquisition strategies, enhance employer branding, and improve employability for job seekers by focusing on relevant skills. Strategic HR decision-making, competitive advantage, and improved employee retention were also among the project's benefits.

However, certain limitations, such as data quality and privacy concerns, bias in user-generated content, and the dynamic nature of the job market, were recognized. Nonetheless, addressing these challenges and leveraging the proposed solutions can enhance the validity and reliability of the analysis.

The applications of the project findings extend to talent acquisition, skill development, HR strategy, and market intelligence for companies, while also providing valuable insights for industry research and policymaking. Moreover, start-ups and small businesses can benefit from the analysis to strengthen their talent acquisition and branding efforts.

The project's analysis of Glassdoor job postings provides a valuable resource for job seekers and employers, enabling better decision-making and increasing efficiency in the job market. By understanding job market trends, in-demand skills, and employer branding perceptions, stakeholders can navigate the dynamic job market landscape with greater confidence and success. The project's findings contribute to bridging the gap between job seekers and employers, ultimately fostering a more thriving and competitive job market.

**8. FUTURE SCOPE**

The future scope of the project presents exciting opportunities for enhancing the analysis of Glassdoor job postings. Firstly, integrating additional data sources from various job platforms and professional networks can provide a more comprehensive view of the job market. By analyzing data from multiple sources, the project can offer a richer understanding of job trends and employer preferences, leading to more robust and accurate insights for job seekers and employers.

Secondly, advancements in natural language processing techniques can significantly improve the analysis's quality and objectivity. Leveraging advanced sentiment analysis models and NLP algorithms can help mitigate bias in user-generated content, ensuring a more balanced and unbiased perspective of employer branding and company reputation. This enhancement would lead to more reliable recommendations for employers to enhance their branding strategies.

Thirdly, predictive analytics can be incorporated into the analysis to forecast future job market trends. By applying machine learning models, the project can offer valuable foresight into the demand for specific skills and industries, empowering both job seekers and employers to make informed decisions aligned with the future job market demands.

Lastly, geographical and industry-specific analyses hold great potential for tailoring insights to localized job markets and specific sectors. By focusing on regional or industry-specific trends, the project can provide targeted guidance for job seekers seeking opportunities in specific areas or sectors and employers looking to tap into niche talent pools.

By embracing these future enhancements, the project can remain dynamic and adaptable to the ever-changing job market landscape. The utilization of real-time data, machine learning, and advanced analytics will transform the analysis of Glassdoor job postings into a powerful tool for job seekers and employers alike, contributing to a more efficient and inclusive job market ecosystem. The project's future scope represents a promising journey towards empowering individuals and organizations to make informed decisions that align with the evolving demands of the job market.

**THANKING YOU….**