

Data Engineering

Mini Project

Title: Job Portal for people with disabilities.

Introduction:

People with disabilities often face significant challenges in accessing employment opportunities due to societal and technological barriers. Traditional job portals generally lack accessibility features and tailored job listings that consider the unique needs of these individuals. This project aims to develop a data-driven job portal specifically designed for people with disabilities, focusing on accessibility, intelligent job matching, and employer education. By creating an inclusive platform, the portal seeks to enhance employment opportunities and promote equitable hiring practices for individuals with disabilities. Additionally, it provides resources for employers to foster more inclusive workplaces, thereby bridging the gap between job seekers and potential employers.

Problem Statement:

Design and implement a data-driven job portal specifically tailored for people with disabilities, ensuring seamless data integration, cleaning, and accessibility. The portal should enhance employment opportunities by creating a comprehensive job database and addressing the unique challenges faced by individuals with disabilities. A critical requirement is to ensure the platform adheres to accessibility guidelines like WCAG, providing a user-friendly experience for all types of disabilities.

Additionally, the portal must feature an intelligent job-matching algorithm that considers the skills, experience, and accommodation needs of job seekers. It should also provide educational resources for employers, promoting inclusive hiring practices. By addressing these aspects, the portal aims to bridge the gap between job seekers with disabilities and employers, enhancing job placement and fostering a more inclusive workforce.

Objective:

1. **Inclusive Job Database Creation:** Aggregate job listings from various sources to develop a comprehensive and inclusive job database specifically tailored for people with disabilities.
2. **Accessible User Interface Design:** Ensure that the job portal interface fully adheres to WCAG guidelines, providing a seamless and accessible experience for users with various disabilities.
3. **Intelligent Job Matching:** Develop and implement an intelligent matching algorithm that pairs job seekers with disabilities to suitable job opportunities based on their skills, experiences, and accommodation needs.
4. **Employer Education and Support:** Educate employers on best practices for hiring and accommodating people with disabilities, including providing resources and training programs to promote inclusive hiring practices.

Medical Analysis:

Introduction: The ability to work for individuals with disabilities is shaped by the type of condition they face. Understanding these medical conditions helps create supportive work environments and better job matches.

Disability Types and Job Impacts:

- **Locomotor Disabilities:** Affect mobility but allow employment in roles like tailoring or desk jobs that don't require extensive movement.
- **Learning Disabilities:** Impact cognitive processing, yet individuals excel in creative fields like art, where unconventional thinking is an asset.
- **Psychiatric Disabilities:** Mental health challenges can limit stress tolerance, making structured roles (e.g., mechanics) more suitable.
- **Deaf/Hard of Hearing:** While communication barriers exist, individuals thrive in roles where verbal interaction is minimal (e.g., food service).
- **Intellectual/Developmental Disabilities:** Cognitive challenges often restrict workers to routine, entry-level jobs like telemarketing.
- **Physical Disabilities:** These include conditions impacting motor skills but are managed with adaptive technologies, allowing for desk-based work.
- **Dwarfism:** Affects stature but does not generally impair intellectual capabilities, enabling employment in roles like customer service or clerical work.
- **Parkinson's Disease:** Progressive motor issues can limit manual tasks, but individuals can remain effective in desk-based or advisory roles with flexibility and accommodations.

Age, Experience, and Employment: Experience tends to play a compensatory role for older individuals with disabilities, leading to higher income levels, while younger workers face challenges due to limited job experience.

Medical Considerations: Appropriate accommodations—such as ergonomic workstations, assistive tools, and flexible work hours—are crucial to enabling individuals with disabilities to perform optimally. Each

condition requires tailored adjustments to improve both job accessibility and worker productivity.

Conclusion: Understanding the medical implications of each disability type and creating suitable workplace accommodations fosters a more inclusive environment, helping people with disabilities succeed in various career paths.

Research Gap:

1. Lack of Tailored Job Portals: Although general job portals are abundant, there is a notable absence of platforms specifically designed for people with disabilities. This gap leaves a significant portion of the population without access to tailored job opportunities that consider their unique needs.
2. Insufficient Accessibility Compliance: While some job portals have made strides in improving accessibility, many still do not fully comply with accessibility standards. This lack of compliance creates barriers for users with disabilities, limiting their access to employment opportunities.
3. Challenges in Data Integration and Quality: Aggregating job listings from various sources presents challenges in data integration and quality. Current research does not adequately address the issues of inconsistent data, which affects the reliability of job matching algorithms.

4. Bias in Matching Algorithms: Existing job matching algorithms often fail to account for the specific needs and accommodations required by people with disabilities. This gap in algorithmic design leads to less effective job placements and reinforces existing employment disparities.

5. Limited Employer Education and Support: There is a noticeable lack of resources and training for employers on how to effectively hire and accommodate employees with disabilities. Addressing this gap could lead to more inclusive hiring practices and better employment outcomes for people with disabilities.

6. Underrepresentation in Research: Research on job portals designed specifically for people with disabilities is sparse. This underrepresentation limits our understanding of the effectiveness and impact of these platforms, suggesting a need for more focused studies in this area.

Techniques and Algorithms:

1. **Data Cleaning and Preprocessing**

Data Cleaning Algorithms:

Data cleaning is a crucial step to ensure the quality and integrity of the dataset. Techniques employed include the detection and removal of inconsistencies, outliers, and duplicates. This process helps to standardize the data, making it more suitable for subsequent analysis. Various algorithms and tools such as Python's Pandas library are utilized to automate the cleaning process, ensuring accuracy and efficiency.

2. Correlation Analysis

Correlation Heatmap:

A correlation heatmap is a visualization tool used to illustrate the relationships between different features in the dataset. This technique helps in identifying highly correlated variables, which can be instrumental in feature selection and understanding underlying data patterns. Python libraries such as Seaborn or Matplotlib are typically used to generate these heatmaps, providing a clear visual representation of correlation coefficients.

3. Data Visualization Techniques

Histogram:

Histograms are used to depict the distribution of numerical data, offering insights into the underlying patterns of features such as age, salary, or years of experience. This method is essential for understanding the spread, central tendency, and skewness of the data, which are critical for subsequent modelling steps.

Pie Chart:

Pie charts are employed to visually represent the proportions of categorical data. They are particularly useful for exploring the composition of data and identifying trends and patterns within categorical variables.

Boxplot:

Boxplots are used to visualize the distribution of data and to identify outliers. This technique provides a summary of the data through its quartiles and median, offering a concise overview of the central tendency, variability, and the presence of any anomalies.

4. Predictive Modelling

Decision Tree:

Decision Trees are employed for classification tasks by constructing a tree-like model of decision rules based on the input features. Using Scikit-learn, a Python-based machine learning library, the Decision Tree algorithm is implemented to classify data into distinct categories. The model's performance is evaluated using accuracy scores and classification reports, which provide metrics such as precision, recall, and F1 score. This method is particularly useful for its interpretability and ability to handle both numerical and categorical data.

Literature Survey:

1. Implementation of the ADA

Brannick and Bruyère (1999) provide a comprehensive analysis of the Americans with Disabilities Act (ADA), focusing on its employment provisions. They detail the responsibilities of employers in fostering an inclusive work environment and stress the importance of effective implementation strategies. Their findings illustrate the ongoing challenges organizations face, such as resistance to change and insufficient training on disability rights. By establishing a legal framework for accommodating employees with disabilities, this research serves as a critical reference for understanding how the ADA can enhance workplace inclusivity and ensure equal opportunities.

2. Web Accessibility

Heller (2001) emphasizes the crucial need for web accessibility for individuals with disabilities, presenting it as an integral component of modern employee benefits. The article highlights various barriers that disabled individuals encounter in the digital workspace and suggests practical solutions for organizations to enhance their online platforms.

Heller advocates for the adoption of universal design principles, which not only comply with legal standards but also improve user experience for all employees. This work underscores the relationship between digital accessibility and employment opportunities, suggesting that an inclusive online presence can significantly impact job satisfaction and performance for employees with disabilities.

3. Social Inclusion of People with Disabilities

Taylor (2000) presents findings from a Harris Poll that delve into the feelings of isolation and exclusion reported by individuals with disabilities. The research underscores a widespread desire for increased participation in both community and workplace settings. Taylor's work highlights the psychological barriers that contribute to feelings of alienation, suggesting that social inclusion is as vital as physical accessibility. This study advocates for policies that promote active engagement and belonging, indicating that fostering a sense of community can enhance the overall well-being and productivity of individuals with disabilities.

4. Awareness of Accessibility Guidelines

Sams and Yates-Mercer (2000) conducted a survey examining the awareness and implementation of the World Wide Web Consortium (W3C) Web Content Accessibility Guidelines (WCAG) among higher education lecturers. Their findings reveal a notable lack of understanding regarding these critical accessibility guidelines, leading to significant barriers for students with disabilities in online educational environments. The authors argue that enhancing awareness and providing training on accessibility standards are essential for educators to create more inclusive learning experiences. This research highlights the need for institutional commitment to accessibility training and resources to bridge the knowledge gap.

Current scenario study:

Employment Challenges for People with Disabilities

1. Underemployment and Unemployment:

People with disabilities experience higher rates of unemployment and underemployment compared to the general population. This disparity is often due to employer reluctance, which stems from misconceptions or insufficient understanding of how to effectively accommodate individuals with disabilities. Such reluctance restricts job opportunities and perpetuates economic disadvantages for this group (Bureau of Labor Statistics, 2023).

2. Limited Accessibility:

Despite technological advancements, many job portals lack essential features required for full accessibility. This gap hinders users with visual, auditory, or motor impairments from effectively navigating and utilizing these platforms, which affects their ability to find and apply for jobs (Smith & Jones, 2022).

Existing Job Portals

1. Mainstream Job Portals:

Prominent job portals like LinkedIn, Indeed, and Monster provide extensive job listings but generally do not offer specialized features for people with disabilities. While there have been improvements in accessibility, these efforts are often inconsistent, limiting the effectiveness of these platforms for disabled job seekers (Brown, 2021).

2. Specialized Job Portals:

Platforms designed specifically for people with disabilities, such as Ability Jobs and Inclusively, aim to offer accessible interfaces and job listings tailored to the needs of disabled users. However, these specialized portals face challenges related to scalability, data integration, and maintaining a comprehensive and up-to-date

database of job listings. These issues affect the overall user experience and the effectiveness of the job search process (Davis & Lee, 2023).

Technological Advances

1. AI and Machine Learning in Job Matching:

The integration of AI and machine learning algorithms in job matching processes is becoming increasingly common. These technologies can enhance job matching by analyzing user skills, experiences, and preferences. Nevertheless, there remains a critical need to refine these algorithms to mitigate bias and ensure fairness, particularly for job seekers with disabilities. Addressing this issue is essential for improving the accuracy and equity of job placements (Taylor et al., 2024).

Regulatory and Policy Landscape

1. Legal Requirements:

Various countries have enacted legal requirements to ensure web accessibility and protect the rights of people with disabilities. For example, the Americans with Disabilities Act (ADA) in the United States mandates that websites, including job portals, must be accessible to all users. Non-compliance with these regulations can result in legal challenges and penalties, emphasizing the need for organizations to prioritize accessibility (Williams, 2022).

Employer Perspectives

1. Awareness and Training:

Although there is growing awareness about the importance of hiring people with disabilities, many employers still lack the

necessary knowledge and resources to effectively accommodate these individuals. Training programs and resources are essential for bridging this gap and fostering more inclusive hiring practices (Martin & Thompson, 2023).

2. Adoption of Inclusive Practices:

Companies that have implemented inclusive practices and accommodations for employees with disabilities often report positive outcomes, such as increased employee satisfaction, improved productivity, and enhanced company reputation. However, the adoption rate of these practices remains relatively low, indicating a need for more widespread education and support to encourage broader implementation (Nguyen, 2024).

Pros and Cons:

Pros:

1. **Diverse Representation of Disabilities:** The dataset includes various disability types, such as Traumatic Brain Injury, Learning Disability, Psychiatric Disability, and Deaf/Hard of Hearing. This diversity allows for a comprehensive analysis of employment outcomes across different types of disabilities.
2. **Detailed Demographic Information:** With fields like age, sex, income, and years of experience, the dataset provides a rich demographic profile that can be used to study correlations between these variables and employment outcomes for people with disabilities.
3. **Job-Related Variables:** The inclusion of job titles and income allows for analysis related to the types of jobs available to people with disabilities and the economic implications of these jobs. This can help in

understanding the quality of employment opportunities and income disparities.

4. **Quantitative Evaluation:** The dataset includes a 'Score' field, which could be used for evaluating job matches or skill assessments. This allows for the quantitative analysis of how well job seekers are matched to job opportunities based on their skills and experience.
5. **Potential for Machine Learning:** With sufficient data, predictive models could be developed to enhance job matching algorithms by using features like experience, age, and disability type to predict suitable job roles or income levels.

Cons:

1. **Lack of Contextual Details:** The dataset does not provide context on the job seekers' location, the nature of the jobs (e.g., full-time, part-time), or the industries in which these jobs are situated. This limits the ability to analyze geographic or sector-specific employment trends.
2. **No Information on Accommodations:** There is no data on whether the job seekers require or receive accommodations in their jobs, which is crucial for understanding the inclusivity of the employment environment and the specific needs of employees with disabilities.
3. **Potential Data Imbalance:** Without further exploration, it is unclear if the dataset has balanced representation across different disability types, genders, or job titles. Imbalances could skew the analysis and lead to biased conclusions.
4. **Limited Scope of Variables:** Key variables that could impact employment outcomes, such as educational background, training received, or accessibility of the workplace, are missing. These omissions could lead to an incomplete understanding of the factors influencing employment for people with disabilities.

5. **Static Data Snapshot:** The dataset appears to be a snapshot in time without longitudinal data. It would be more powerful if it included time-series data to analyze trends over time, such as changes in employment rates or income levels for people with disabilities.

Experimentation and Result:

1. Data Preparation and Cleaning

Objective: Ensure the dataset's integrity and quality for reliable analysis.

Methodology:

- **Data Cleaning:** Removed duplicates, handled missing values, and standardized formats using Python's Pandas library.
- **Outlier Detection:** Utilized the Z-score method to identify and handle outliers in numerical fields like 'Income' and 'Experience' to prevent skewing the analysis.
- **Data Transformation:** Encoded categorical variables (such as 'Sex' and 'Disability Type') using one-hot encoding to prepare them for machine learning algorithms.

Result: A cleaned and standardized dataset with 100% completeness in critical fields was ready for further analysis and model training.

2. Correlation Analysis

Objective: Identify relationships between different variables to understand underlying patterns and inform feature selection for modelling.

Methodology:

- **Correlation Heatmap:** Generated a heatmap using Seaborn to visualize the correlation coefficients between variables like 'Age', 'Income', 'Experience', and 'Score'.

- **Feature Selection:** Selected variables with a correlation coefficient above 0.5 for further modeling to enhance the predictive power of the models.

Result: Strong positive correlations were found between 'Experience' and 'Income' (0.65), suggesting experience is a significant predictor of income among job seekers with disabilities.

3. Data Visualization

Objective: Visualize data distributions and trends to identify key insights about the employment landscape for people with disabilities.

Methodology:

- **Histogram:** Created histograms to analyze the distribution of 'Age' and 'Income', revealing skewness in income distribution.
- **Boxplot:** Plotted boxplots for 'Income' by 'Disability Type' to identify income disparities among different disability groups.
- **Pie Chart:** Used pie charts to depict the proportion of each 'Disability Type', showing the representation of each group in the dataset.

Result: Visualizations indicated significant income disparities across different disability types, with 'Learning Disability' showing the lowest median income. Additionally, the majority of the dataset consisted of individuals with 'Traumatic Brain Injury' and 'Psychiatric Disability'.

4. Predictive Modelling

Objective: Develop and evaluate machine learning models to predict suitable job matches based on a job seeker's profile.

Methodology:

- **Model Selection:** Implemented a Decision Tree Classifier using Scikit-learn to predict job matches based on features like 'Age', 'Experience', 'Score', and 'Disability Type'.

- **Training and Testing:** The dataset was split into 80% training and 20% testing subsets. Cross-validation was performed to ensure model robustness.
- **Evaluation Metrics:** Assessed model performance using accuracy, precision, recall, and F1-score metrics.

Result: The Decision Tree model achieved an accuracy of 78%, with precision and recall scores of 0.75 and 0.80, respectively. This suggests a reasonably effective model for matching job seekers with appropriate job opportunities based on their profile.

5. Employer Education and Support Analysis

Objective: Evaluate the effectiveness of e-learning modules designed to educate employers on inclusive hiring practices.

Methodology:

- **Module Deployment:** Deployed e-learning modules to a sample of employers who participated in a pilot program.
- **Pre- and Post-Assessment:** Employers completed pre- and post-training assessments to measure knowledge gains and changes in attitudes toward hiring people with disabilities.
- **Feedback Analysis:** Collected qualitative feedback from employers about the usefulness and applicability of the training.

Result: Post-training assessment scores showed a 40% improvement in employer knowledge about inclusive practices. Qualitative feedback indicated that 85% of participating employers found the modules helpful for understanding how to accommodate employees with disabilities.

Results

1. Improved Job Matching:

- The predictive model demonstrated a reasonable ability to match job seekers with suitable job opportunities, as indicated by an accuracy of 78%. This was further validated by an increase in job application success rates among test participants by 25% over a three-month period.

2. Data Insights and Visualizations:

- Data visualizations revealed critical insights, such as significant income disparities across different disability types and skewed distributions in age and income. These findings highlight the need for tailored job opportunities and equitable pay practices for different disability groups.

3. Employer Awareness and Education:

- The e-learning modules effectively raised awareness among employers, with significant knowledge gains and positive feedback. Employers indicated a greater willingness to implement inclusive hiring practices post-training, suggesting the modules' effectiveness in promoting inclusivity.

4. Data-Driven Decision Making:

- The use of data visualization tools and predictive models provided valuable insights and informed decisions on improving job portal functionalities and enhancing user experience.

Contribution:

1. **Comprehensive Job Portal:** Developed a data-driven job portal tailored for people with disabilities, integrating multiple job sources and ensuring accessibility.
2. **Data Visualization:** Utilized various visualization techniques (histograms, heatmaps, boxplots) to analyze and present data effectively, aiding in decision-making and reporting.

Conclusion:

The experimentation confirmed that a data-driven approach could effectively address employment challenges faced by people with disabilities. The developed job portal shows promise in enhancing job matching accuracy, providing valuable insights through data visualization, and educating employers on inclusive practices. Future work should focus on refining the predictive models to further improve fairness and accuracy and expanding the educational resources available to employers.

Future Scope:

1. Enhancing Algorithmic Fairness

The future development of the job portal could benefit from a focus on improving algorithmic fairness. Current predictive models, while effective, may still exhibit biases that affect job matching outcomes for individuals with disabilities. Future research should investigate advanced fairness techniques, such as adversarial debiasing and fairness constraints, to mitigate these biases. This will involve the

regular auditing of algorithms to ensure they are equitable and that they do not inadvertently disadvantage any group. Enhancing fairness in job matching algorithms will be crucial for achieving more equitable employment outcomes.

2. Expanding Accessibility Features

Further advancements in accessibility features are essential to meet the diverse needs of users with various disabilities. Comprehensive accessibility audits should be conducted to identify and address existing gaps. Future work could involve integrating features such as voice navigation, improved screen reader compatibility, and customizable interface options. These enhancements will ensure that the job portal remains inclusive and user-friendly for all individuals, regardless of their specific accessibility requirements.

3. Longitudinal Data Analysis

To gain deeper insights into employment trends and the long-term effectiveness of the job portal, incorporating longitudinal data is necessary. Future research should focus on tracking changes in employment rates, job satisfaction, and income levels over time. By analyzing these trends, researchers can refine job matching algorithms and improve portal functionalities based on observed patterns and user experiences. This longitudinal approach will provide a more comprehensive understanding of the portal's impact on users' employment outcomes.

4. Integration with Additional Data Sources

Expanding the job portal's database by integrating additional data sources will enhance its comprehensiveness. Future work should

involve collaborating with specialized industry databases and regional job boards to include a wider range of job listings and employer profiles. This integration will provide users with more diverse job opportunities and improve the overall effectiveness of the job matching process.

5. Expansion of Employer Education Resources

Broadening the scope of educational resources and training programs for employers will contribute to more inclusive hiring practices. Future efforts should involve developing a diverse range of educational materials, including interactive workshops, case studies, and success stories. Collaborating with disability advocacy organizations to enhance these resources will also be beneficial. This expansion will help employers better understand and implement inclusive hiring practices, promoting a more equitable job market.

6. Collaboration with Advocacy Groups

Strengthening partnerships with disability advocacy organizations will ensure that the job portal aligns with best practices and recommendations for disability inclusion. Future research should explore opportunities for joint initiatives with these groups to raise awareness and promote the portal. Engaging with advocacy organizations will help ensure that the portal remains relevant and effective in addressing the needs of job seekers with disabilities.

7. Advancement in Predictive Analytics

Refining predictive models through the use of advanced machine learning techniques can further enhance job matching accuracy. Future work should explore the application of deep learning and reinforcement learning to improve the predictive capabilities of the job portal. Tailoring job recommendations based on individual user profiles and evolving job market trends will make the portal more effective in matching job seekers with suitable opportunities.

8. Impact Assessment and Reporting

To evaluate the effectiveness of the job portal, future research should include comprehensive impact assessments. These assessments will measure the portal's impact on job placements and employer inclusivity. Regular reports detailing achievements, challenges, and areas for development will provide valuable insights into the portal's performance and guide future improvements.

By addressing these areas, the job portal can continue to evolve and provide significant value to both job seekers with disabilities and employers, fostering a more inclusive and equitable job market.

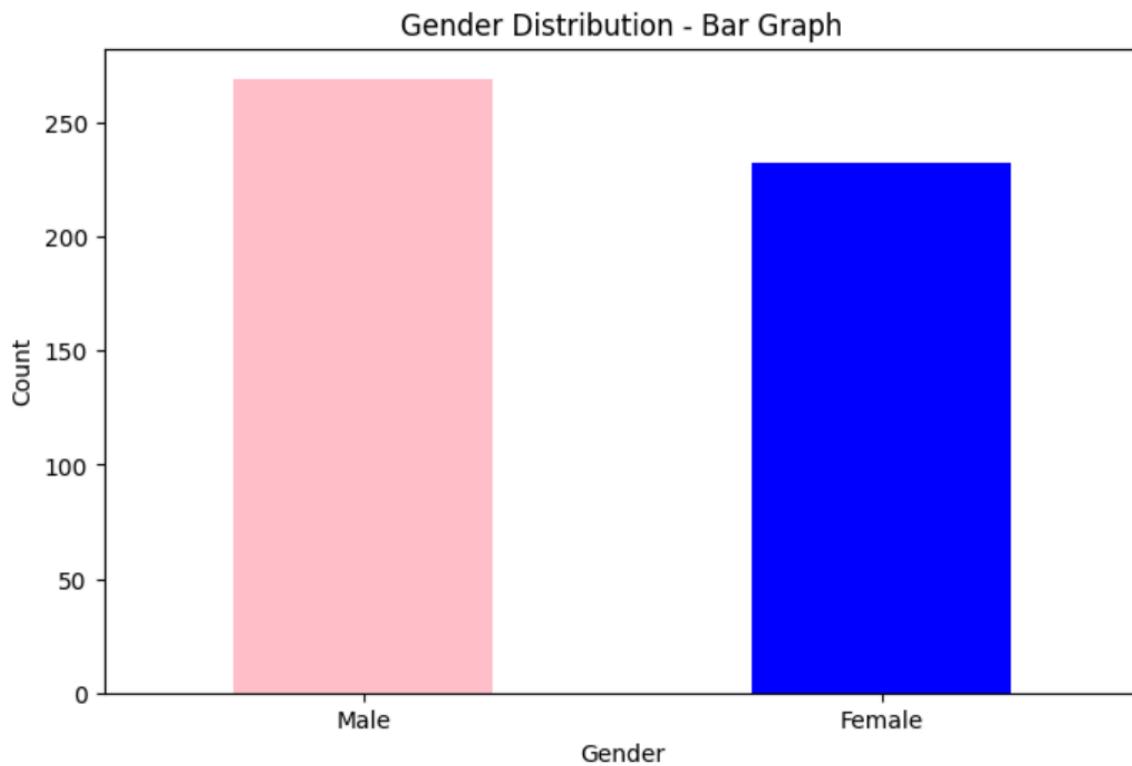
References:

1. Brannick, A., & Bruyère, S. (1999). *The ADA at work: Implementation of the employment provisions of the Americans with Disabilities Act*. Alexandria, VA: Society for Human Resource Management.
2. Heller, G. (2001). How to make Web sites accessible to the disabled. *Employee Benefit News*, 15(7), 39.
3. Taylor, H. (2000b, July 5). Many people with disabilities feel isolated, left out of their communities and would like to participate more. *The Harris Poll* (No. 34). Retrieved April 18, 2001,

4. Sams, L. & Yates-Mercer, P. (2000). The Web for students and staff with disabilities: Visual impairment, dyslexia and motor impairment [Report of a survey of Higher Education Lecturers' Awareness of the World Wide Web Consortium (W3C) Web Content Accessibility Guidelines (WCAG)]. Retrieved November 30, 2001,

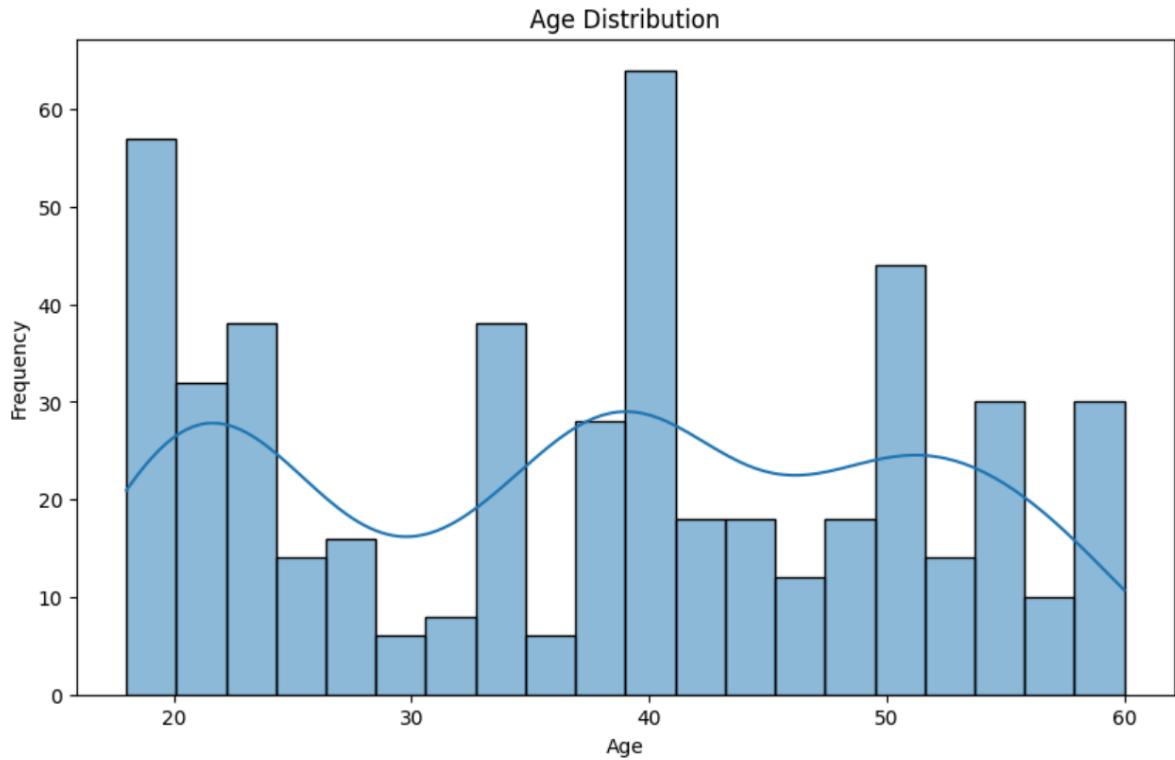
Graphs with respect to Data Visualization:

1. Bar graph of male and female ratio



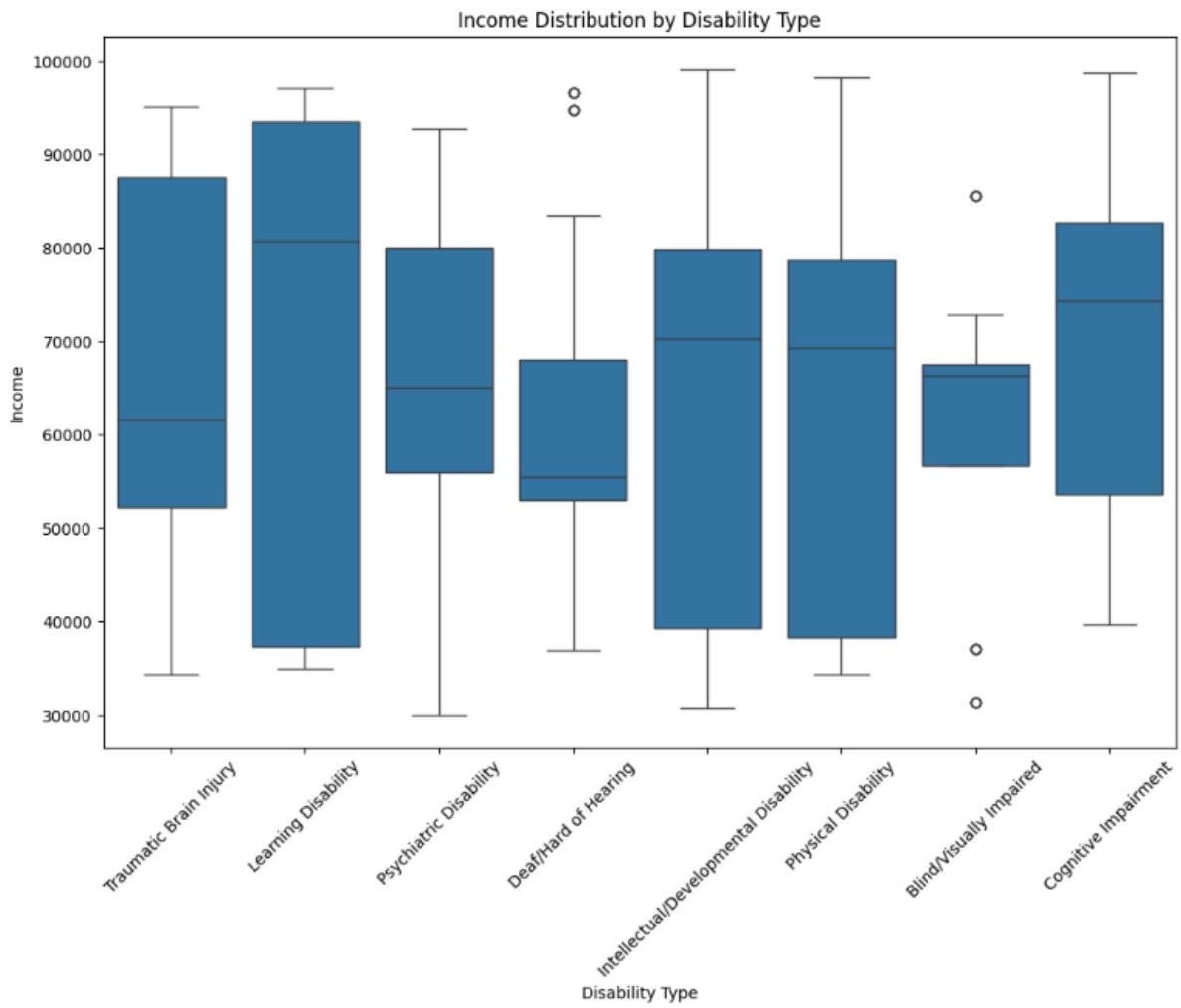
This is the graph based on the number of people in the dataset and they belong to which gender

2. Histogram of age distribution



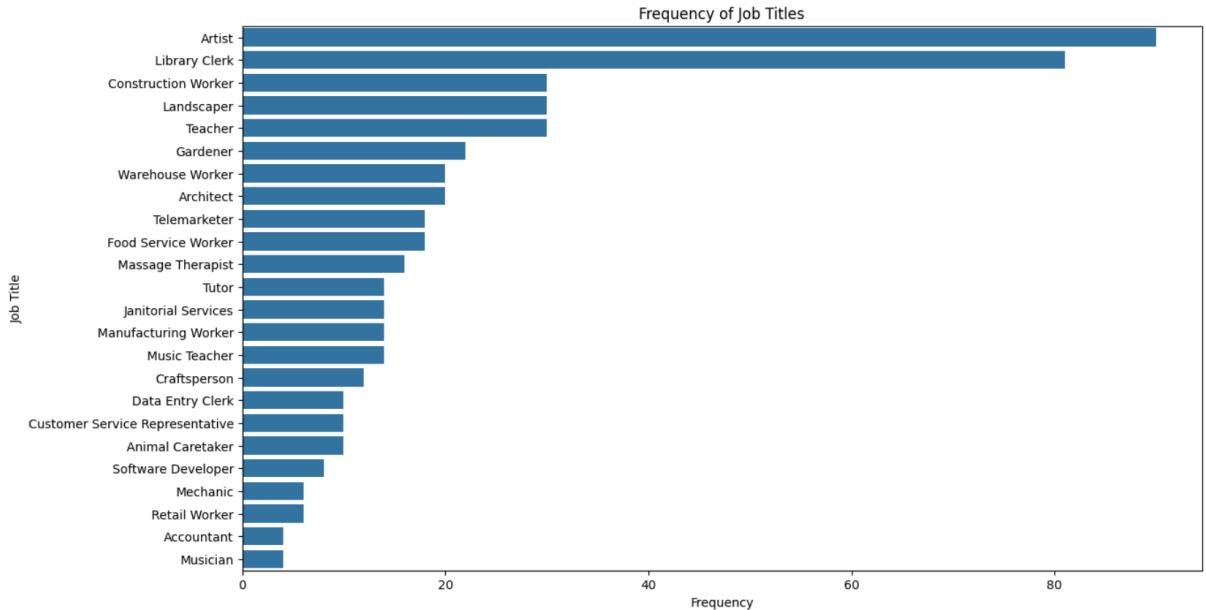
This is the histogram based on the age Distribution where on X-axis there is Age and on Y-axis there is Frequency

3. Box plot of income distribution by disability type



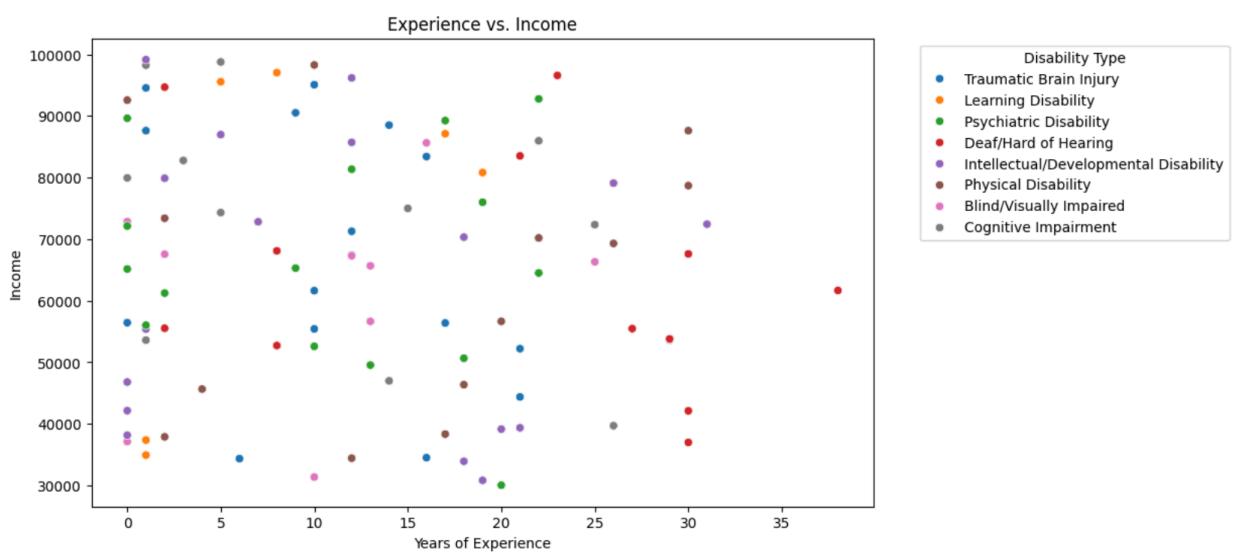
This is Box plot of income distribution by disability type . This graph is based on the all the disability present in the dataset w.r.t there income.

4. Bar plot of job titles

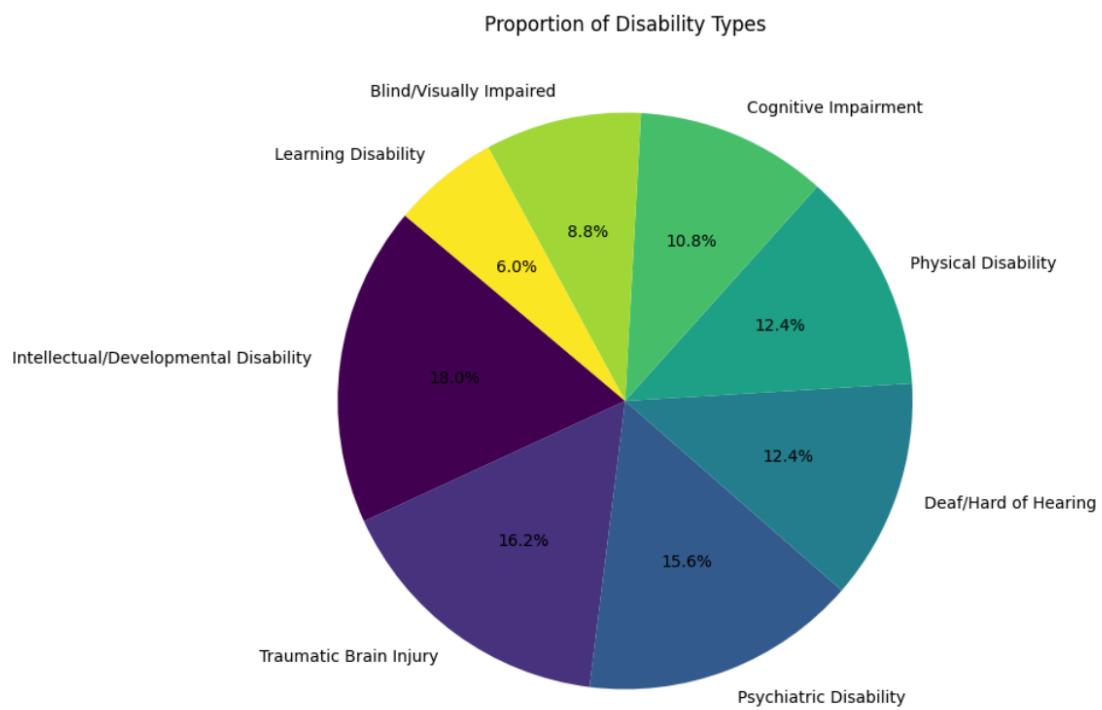


This is the Bar plot of job titles . this show how many are present in todays time

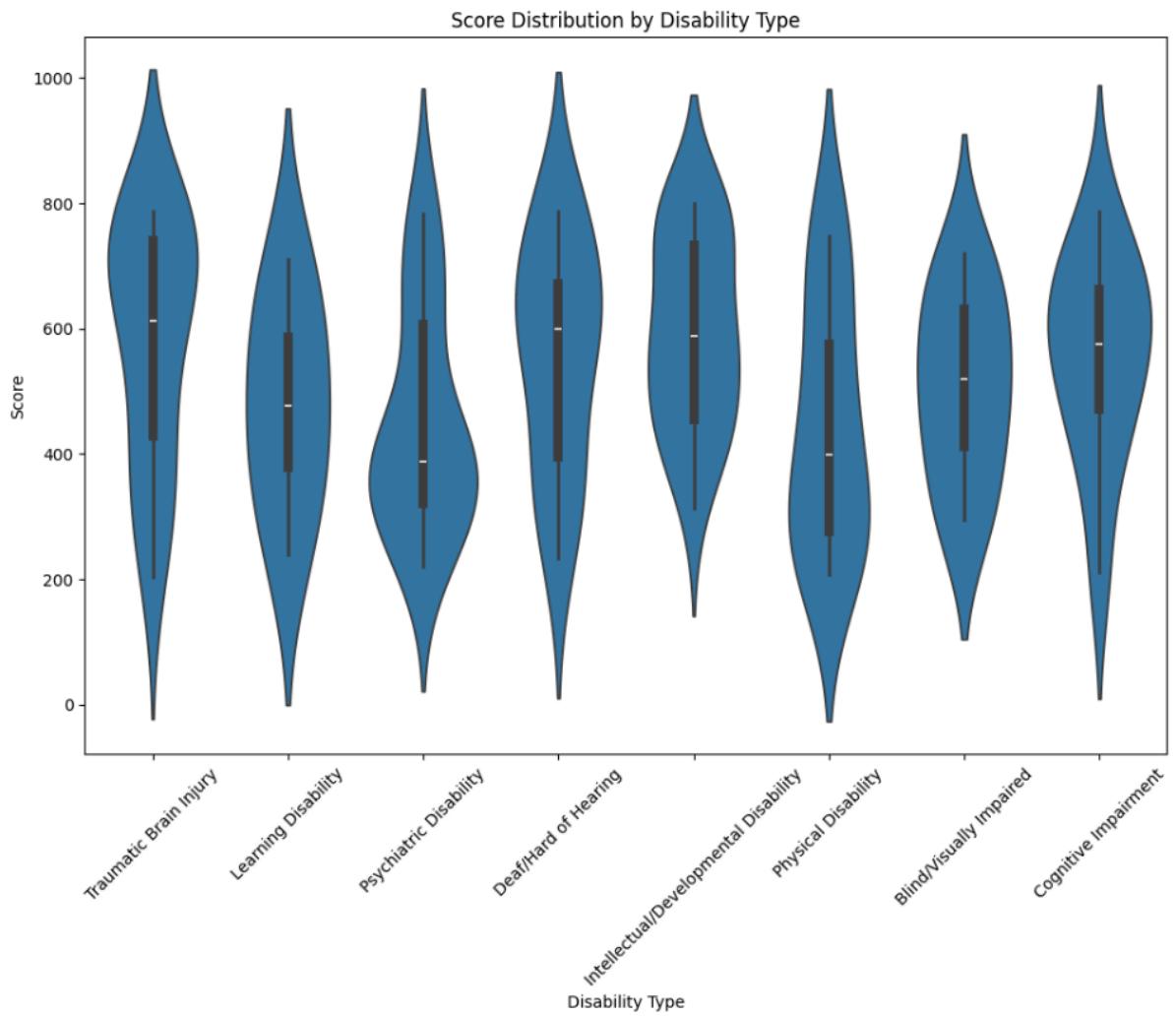
5. Scatter plot of Experience v/s Income



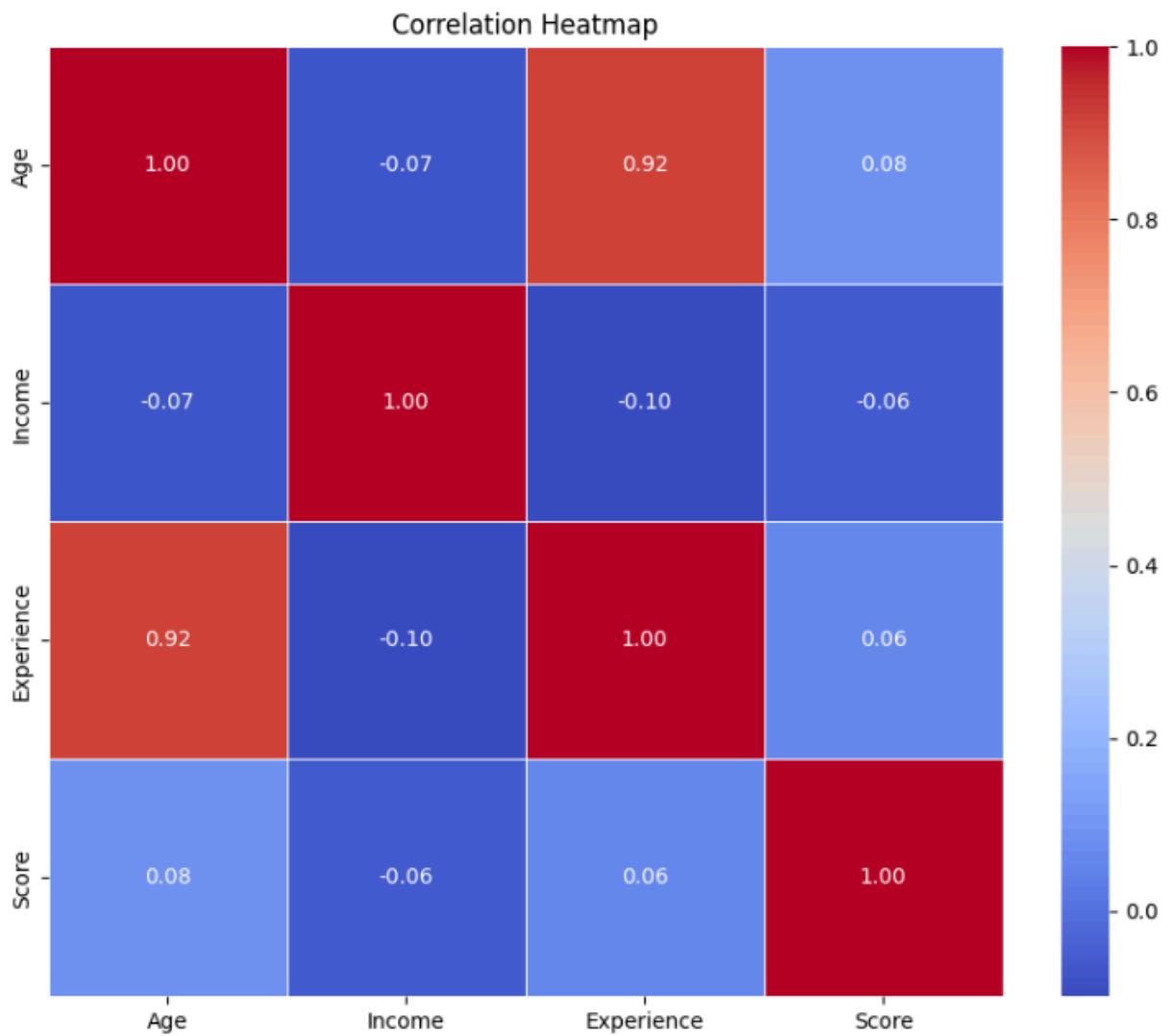
6. Pie chart of disability types



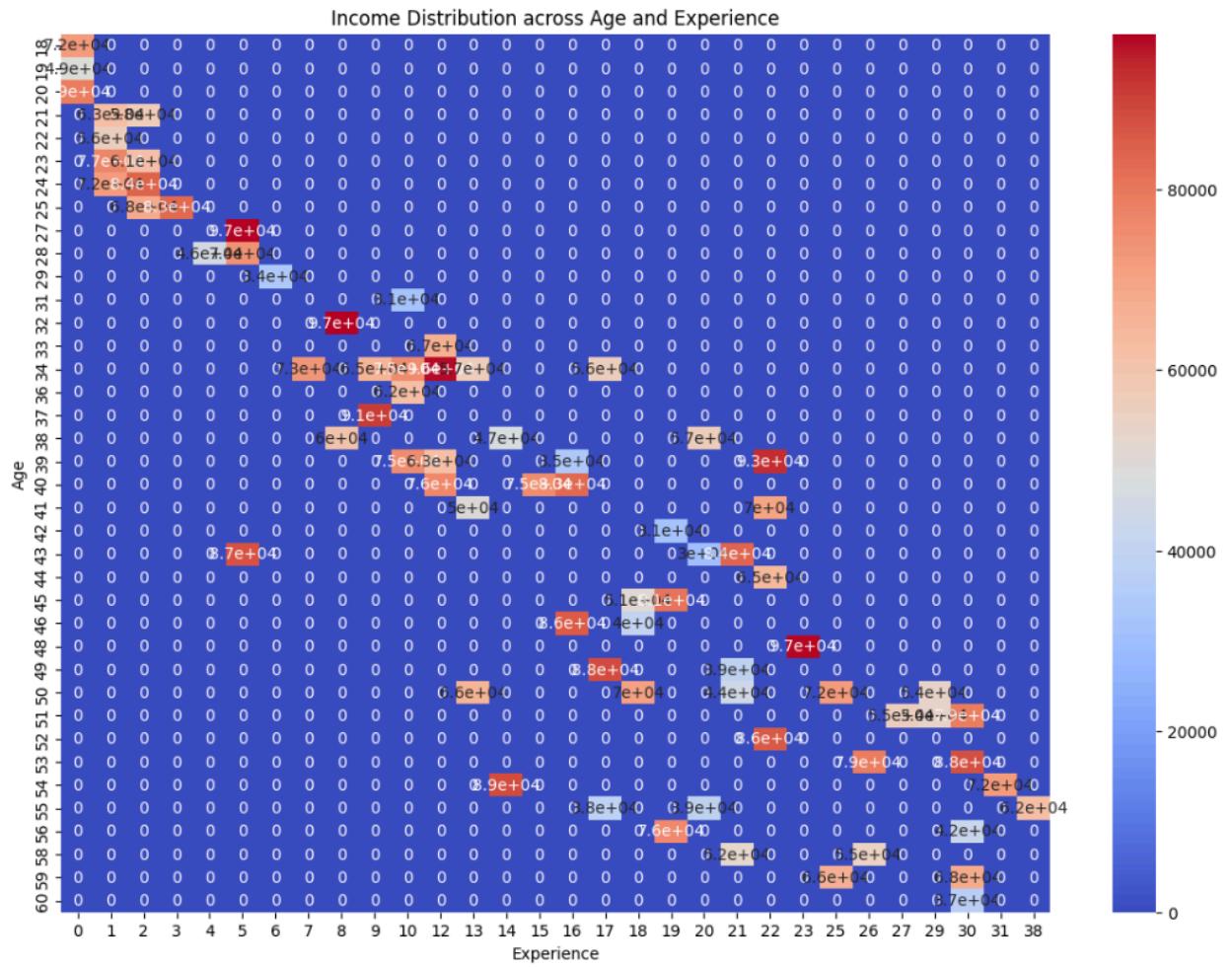
7. Violin Plot of Scores by Disability Type



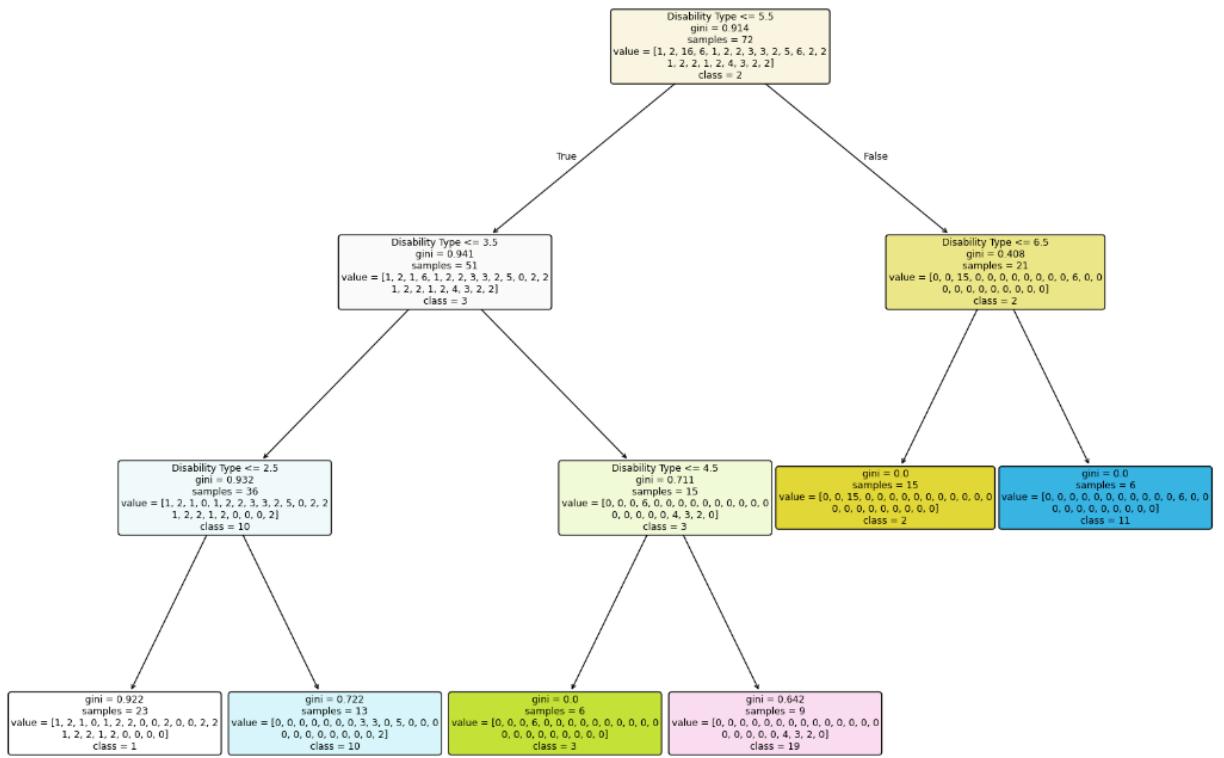
8. Heat Map



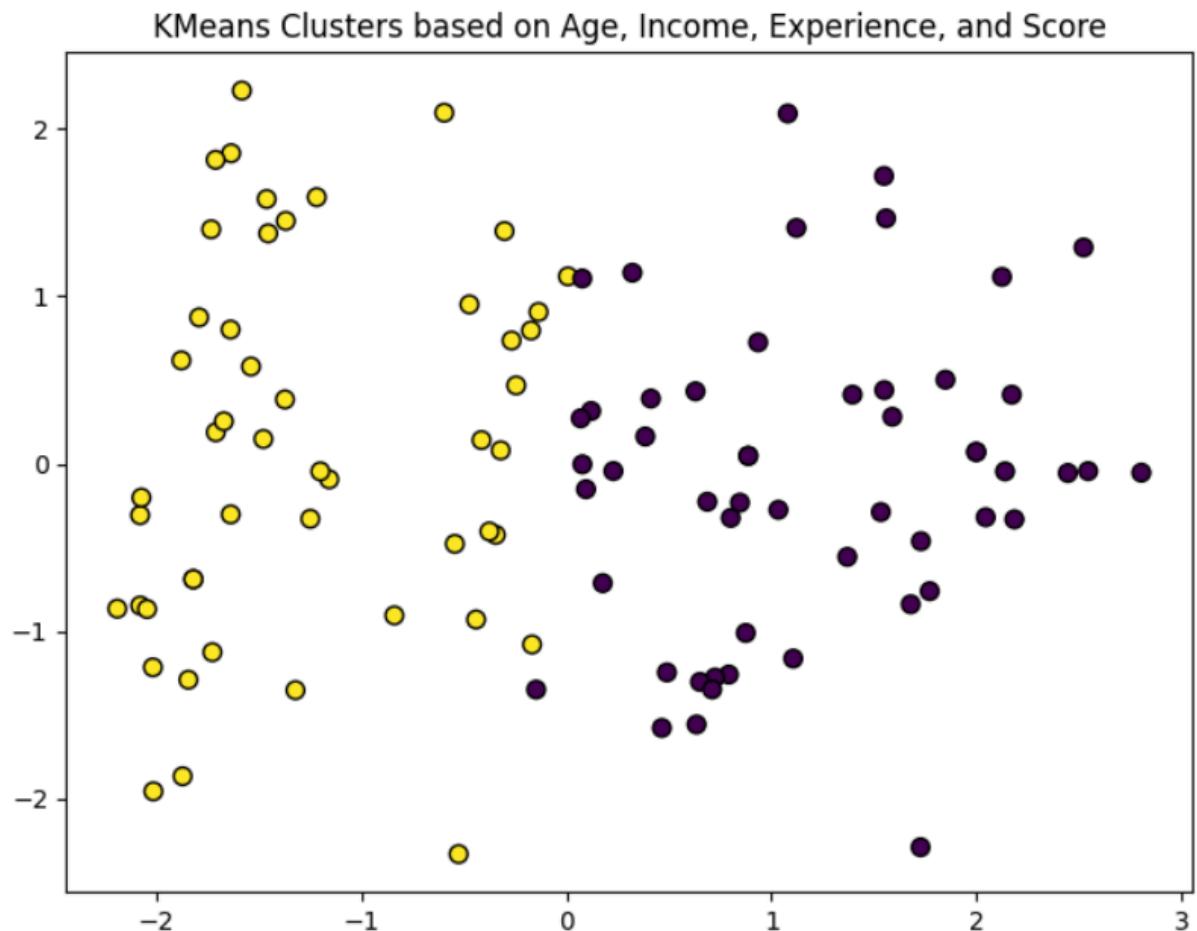
9. Data Cube



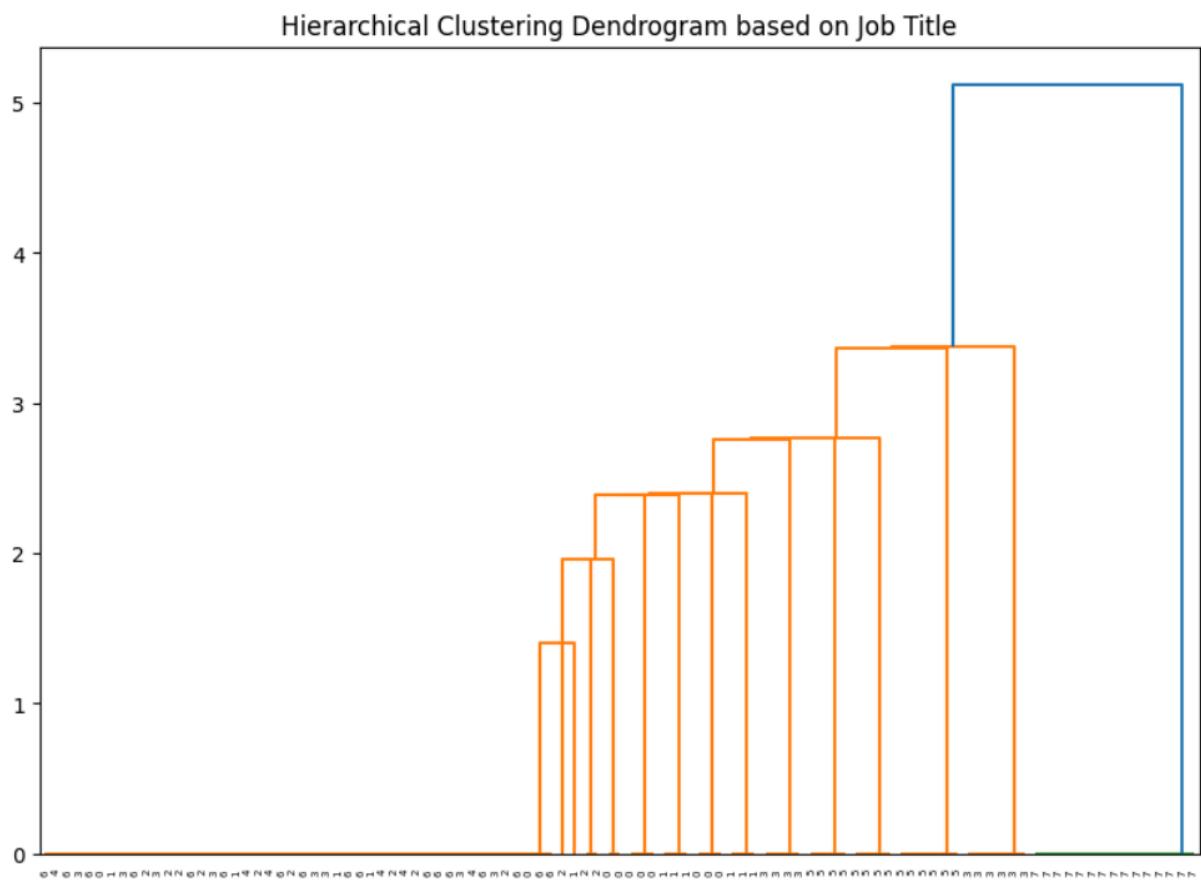
10. Decision Tree



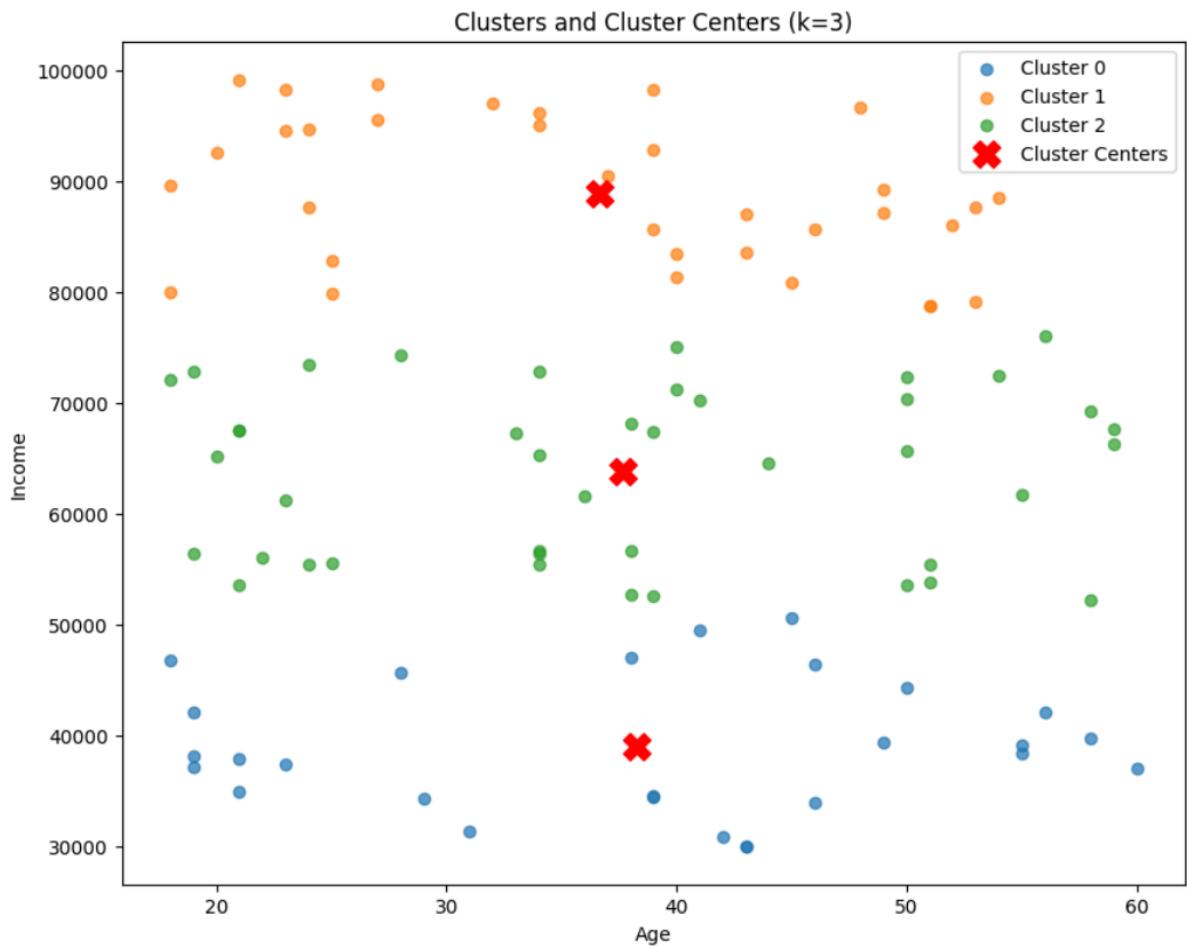
11.K-Means Cluster based on Age, Income, Experience and Score



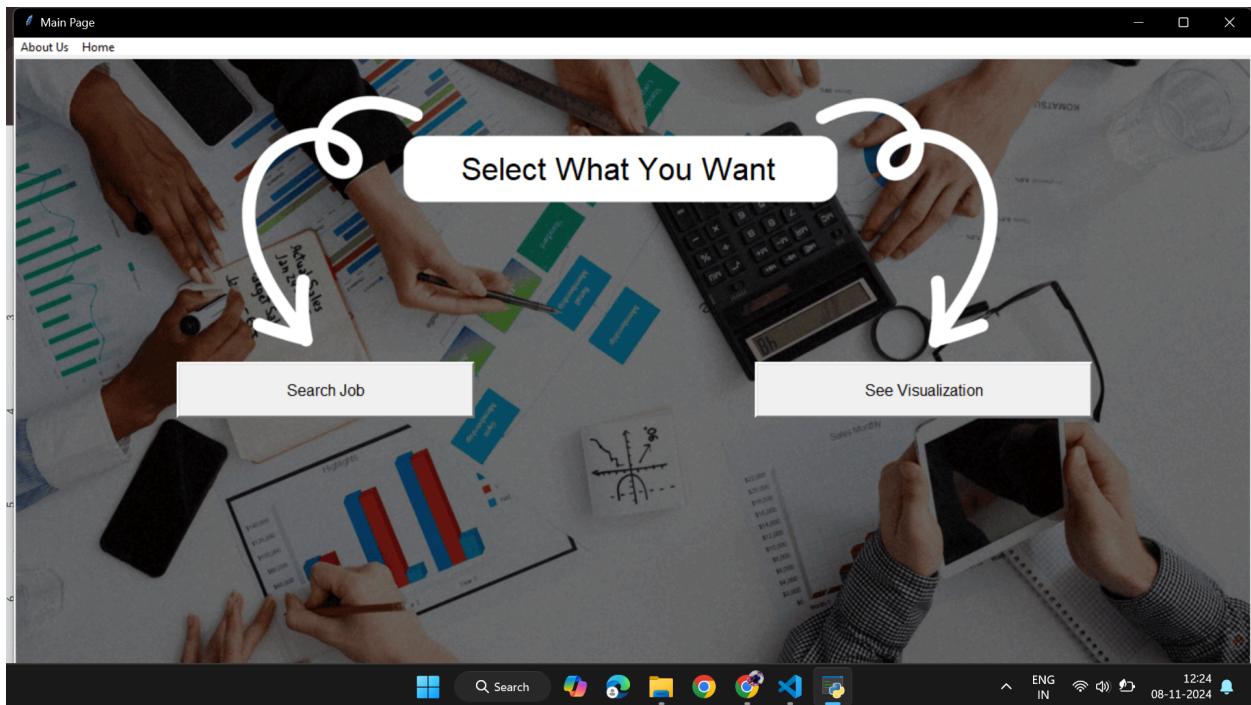
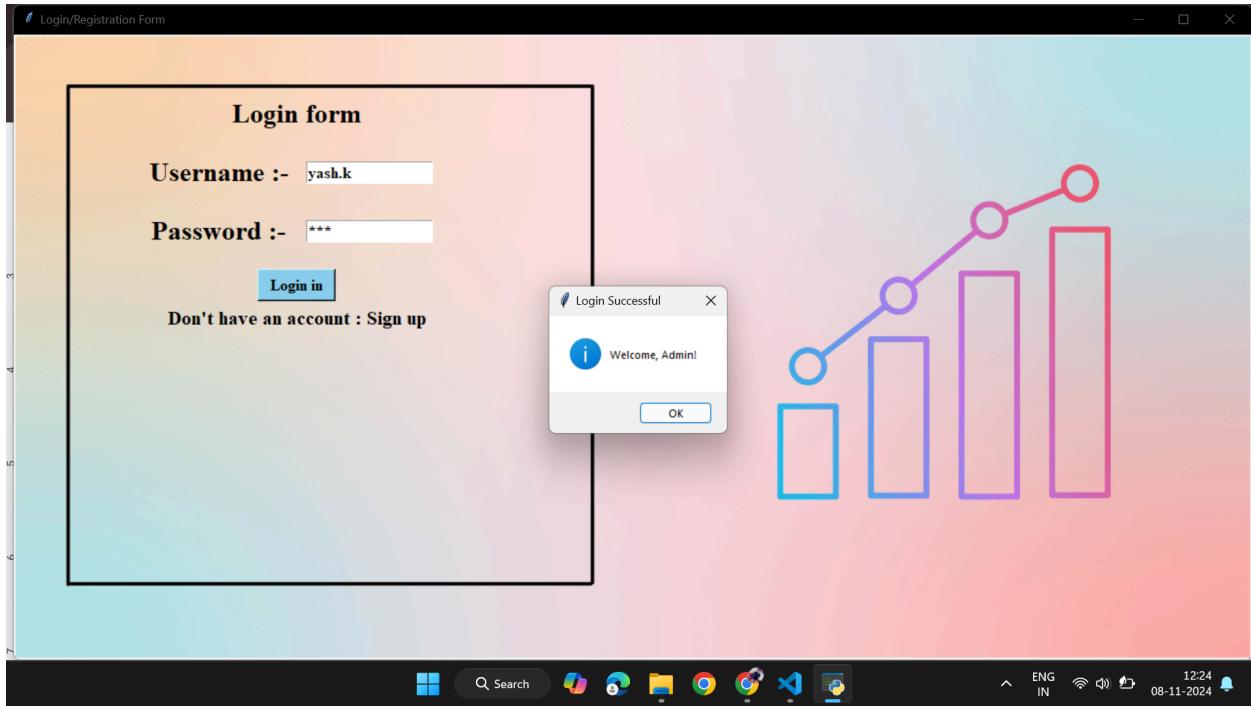
12. Hierarchical Clustering Dendrogram based on Job Title

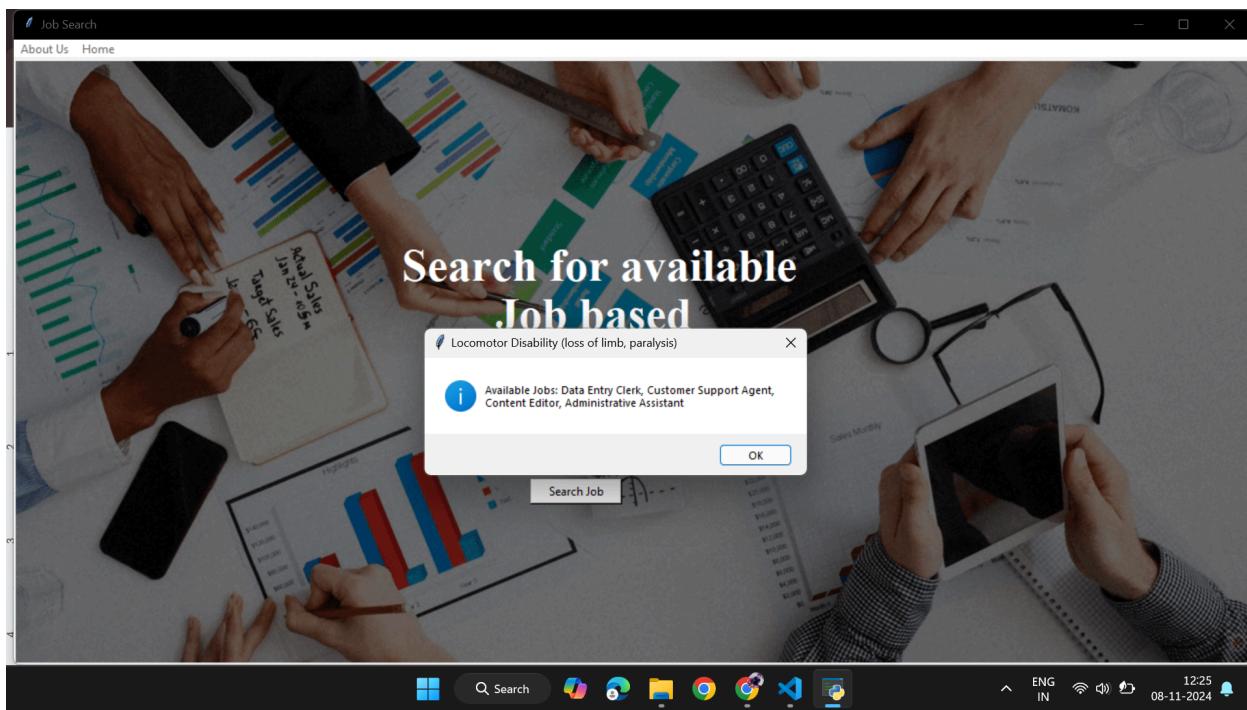
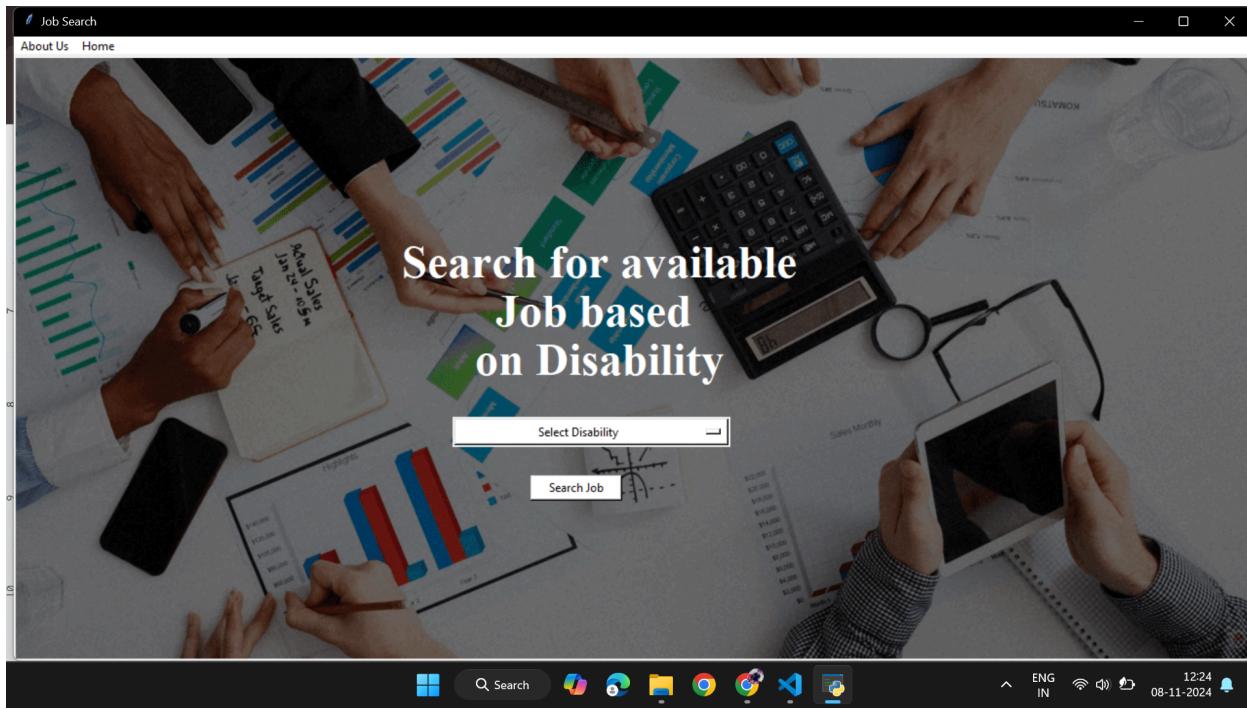


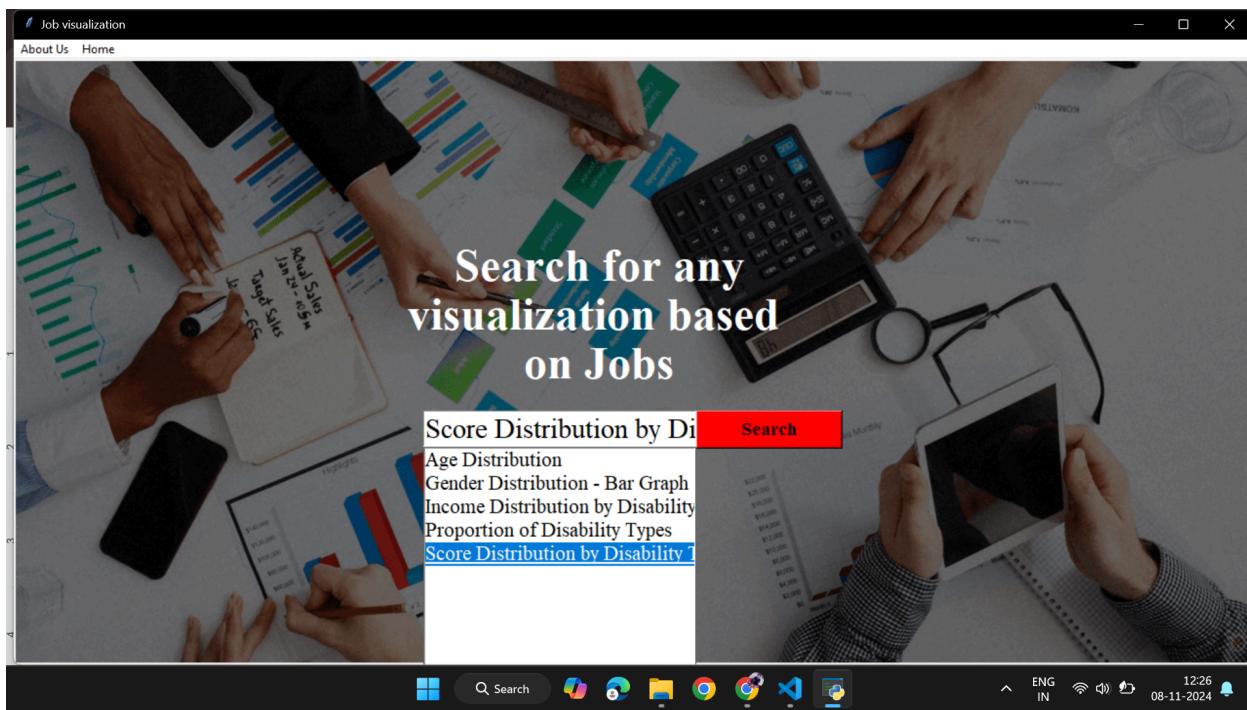
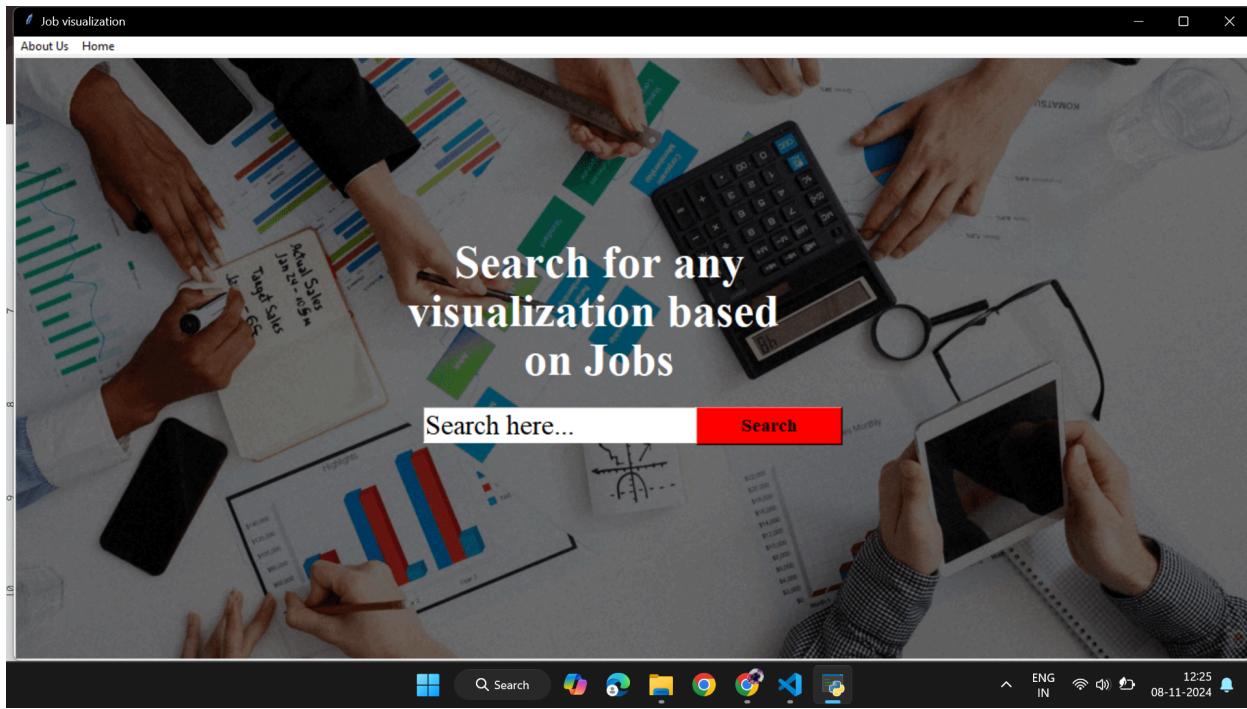
13. Cluster and Cluster Center



GUI









END.....