PROJECT REPORT

Data Analysis

Analysis of Resume Acceptance

Created By:	Karthik S Bhat	Approved By:	Harshada Topale
Created On:	18-10-2023	Approved On:	14-10-2023

Page left blank intentionally

INDEX

General Instructions for using the Live Project Report Template

- This template and the subsequent document created using this template is a confidential document and is the intellectual property of Cloud Counselage Pvt. Ltd. Circulating it outside of the organisation without the consent of Cloud Counselage Pvt. Ltd. is the breach of company policies and will lead to legal actions

- This template is a guideline document to communicate the implementation of design ideas and the results of the work to the stakeholders.

- The **text between inequality (< >) is to be replaced** by relevant text

- Please **remove the yellow highlight on the Text** between the inequality (< >). This is done to help you notice the text to be changed/replaced

- The text in *italics* highlighted in grey is just for reference and should be removed after adding the relevant text

1 PROJECT DETAILS

Project Name	Analysis of Resume Acceptance		
Project Sponsor	Tushar Topale		
Project Manager	Harshada Topale		
Start Date	13-10-2023	**Completion Date**	18-10-2023

2 SUMMARY

The project centred around a comprehensive data analysis initiative, utilizing a substantial dataset encompassing diverse parameters such as First Name, Email ID, Quantity, Events, Attendee Status, College Name, how attendees learned about the event, Designation, Year of Graduation, City, CGPA, Experience with Python (in months), Family Income, Expected Salary (in Lac), and Leadership Skills. The primary objective was to explore this data intricately, aiming to unravel insights regarding resume acceptance criteria.

This project was essential due to the need for a profound understanding of the factors that influence resume acceptance. By leveraging advanced data analysis techniques, the team delved into the intricate patterns within the dataset. The overarching goal was to discern the problems that lead to resume acceptance, enabling organizations and individuals to enhance their application strategies.

The long-term benefits of this endeavour are multifaceted. Firstly, it provides invaluable insights into the intricate process of resume evaluation. Understanding the pivotal elements that influence acceptance not only aids individuals in tailoring their applications but also empowers organizations to refine their hiring processes. Secondly, this project serves as a foundational framework for future data-driven analyses in the realm of recruitment, paving the way for more informed decision-making in the job market. Ultimately, the knowledge derived from this analysis holds the potential to bridge the gap between job seekers and employers, fostering a more efficient and targeted approach to talent acquisition.

3 INTRODUCTION

3.1 Background

- Millions of students apply for internships/jobs every year, resumes play an important role in playing the first impression. The recruiters spend a max of 2-3 minutes reviewing resume after it landed in their mailbox or Job board, ATS application. Surprising more than 70% of resumes get rejected in the initial screening.
- By scrutinizing a diverse dataset encompassing various parameters, the aim is to discern patterns and correlations that would shed light on what makes a resume stand out.

3.2 Stakeholders

- **Job Seekers:** Individuals who are aspiring to understand the problems that can make their resumes more appealing to potential employers.
- **Recruiters/Hiring Managers:** They are invested in optimizing the recruitment process to identify the most suitable candidates swiftly.
- **Human Resources Departments**: HRs are keen on streamlining the hiring process, ensuring it is efficient, fair, and based on relevant criteria.
- **Educational Institutions:** Colleges and universities are interested in understanding the employability factors that make their graduates more desirable to employers.
- **Employers/Organizations:** Companies looking to hire top talent are interested in innovative and data-driven approaches to identify candidates who align with their requirements.

3.3 Objectives

- We aim to conduct a comprehensive analysis of our student interns to gain insights about relationship between their academic performance, event participation, career aspiration and factors influencing their success. We have collected a dataset containing various attributes for each student.
  - **Academic Performance Analysis:** Evaluate the academic performance of students, examining factors such as CGPA. Determine if there is a correlation between academic excellence and success in internships.
  - **Event Participation Examination:** Analyse the involvement of student interns in events, including types of events attended, frequency of participation, and roles undertaken. Investigate whether active event participation correlates with enhanced internship success.
  - **Career Aspiration Assessment:** Understand the career aspirations of student interns by examining their stated goals, preferred industries, and job roles of interest. Determine if interns with clear career goals exhibit distinct patterns of success in their internships.
  - **Pattern Recognition:** Utilize statistical methods and data visualization techniques to recognize patterns within the dataset. Identify common traits among successful student interns, potentially uncovering specific combinations of academic achievements, event participation, and career aspirations that lead to positive outcomes.
  - **Recommendation Strategies:** Based on the analysis, formulate actionable recommendations for interns to enhance their prospects. This might include tailored academic guidance, suggestions for event participation, or mentorship programs catering to specific career aspirations.
  - **Insightful Reporting:** Summarize the findings in a clear and concise manner, presenting insights derived from the analysis. Ensure that the report is accessible to stakeholders, providing valuable information for decision-making processes within educational institutions and organizations hosting internships.

4 METHODOLOGY

These conventions are all about the positions of line breaks, how many characters should go on a line, and everything in between.

4.1 Considerations & Assumption

- In undertaking this project, several key considerations and assumptions guided our methodology:

Constraints:
  - **Data Quality:** Acknowledging the dataset provided, we recognized the necessity for extensive data cleaning due to potential inconsistencies, missing values, and outliers.
  - **Resource Limitations:** Constraints in terms of time and resources necessitated a focused approach, emphasizing efficiency and prioritization.

Challenges:
  - **Assumptions:** We assumed that the provided dataset, despite potential issues, contained valuable insights. Additionally, we assumed that the cleaned data would adequately represent the underlying trends and patterns.

4.2 Approach

- Our approach was structured to address the complexities of the project and focused on a systematic exploration of the data:

  Structured Approach:

  - **Exploratory Data Analysis (EDA):** Initiated with EDA to understand the dataset's characteristics, identifying outliers and missing values. This step was crucial in gaining insights into the data's nature.
  - **Data Cleaning:** Employed rigorous data cleaning techniques to rectify inconsistencies, handle missing values, and remove outliers. This step ensured the dataset's reliability and accuracy.
  - **Descriptive Statistics**: Utilized descriptive statistics to summarize the dataset, providing a clear overview of its key features, distributions, and central tendencies.
  - **Correlation Analysis:** Investigated relationships between variables through correlation analysis, identifying potential patterns and dependencies crucial for subsequent analysis.
  - **Visualization:** Leveraged data visualization tools to represent trends, outliers, and correlations visually, making complex patterns more accessible to stakeholders.

  Why this Approach?

o **Comprehensive Understanding:** This structured methodology allowed for a thorough comprehension of the dataset's intricacies, ensuring that subsequent analyses were built on a solid foundation.

o **Visual Representation:** Visualization played a pivotal role, translating raw data into understandable visuals, aiding in effective communication of findings.

o **Iterative Process:** Recognizing the iterative nature of data analysis, our approach facilitated revisiting cleaning and analysis steps as new insights emerged, ensuring a dynamic and adaptable process.

4.3 Activities

The project involved a series of activities, each integral to its successful delivery:

- **Requirement Gathering:** Attended meeting to understand the project's objectives, defining the scope and expectations.
- **Data Cleaning and Preprocessing:** Dedicated significant effort to clean and preprocess the dataset, ensuring data integrity and reliability.
- **Exploratory Data Analysis:** Conducted in-depth exploratory analysis to uncover trends, outliers, and correlations.
- **Statistical Analysis:** Utilized statistical techniques to validate hypotheses and draw meaningful conclusions from the data.
- **Visualization:** Developed a variety of visualizations to represent key insights, aiding in the interpretation of complex data patterns.
- **Documentation:** Thoroughly documented the methodology, findings, and insights to facilitate clear communication of results to stakeholders.

5 TARGETTED V/S ACHIEVED OUTPUT

Targeted Output:

- The project plan aimed to achieve a comprehensive analysis of the provided dataset, encompassing thorough data cleaning, exploratory data analysis (EDA), descriptive statistics, correlation analysis, and insightful visualizations. The goal was to extract meaningful patterns and relationships from the data, providing valuable insights into the factors influencing resume acceptance and intern success.

Achieved Output:

- The project successfully met its targeted objectives. Extensive data cleaning and preprocessing were performed to ensure data integrity. Exploratory data analysis provided a deep understanding of the dataset, including key trends and correlations. Descriptive statistics and visualizations were employed effectively to represent the data's nuances, enabling clear communication of findings. While no

machine learning or modelling techniques were used, the analysis provided valuable insights into the dataset's characteristics and highlighted significant factors influencing resume acceptance and intern success.

Reasons for Deviation:

- There were no significant deviations from the targeted output. The project adhered closely to the planned methodology. However, minor deviations occurred during the data cleaning phase due to unexpected complexities in handling certain data inconsistencies. These challenges necessitated additional time and effort for thorough cleaning, ensuring the reliability of the subsequent analyses. Despite these deviations, the overall objectives were achieved within the stipulated timeframe.

6 CONCLUSION

Usefulness for Stakeholders:

- The insights derived from this project hold significant value for stakeholders involved in recruitment processes and educational institutions. For recruiters and hiring managers, the analysis provides a nuanced understanding of resume acceptance factors, enabling them to make more informed decisions during candidate evaluations. Educational institutions can use these findings to enhance their curriculum, emphasizing skills and experiences that align with industry expectations, thereby improving students' employability.

Future Scope:

- The project's results pave the way for future endeavours in similar domains. The methods and techniques employed, especially in data cleaning and exploratory analysis, serve as a foundation for more complex analyses in recruitment and educational research. Future projects could delve deeper into specific subgroups within the dataset, exploring tailored factors influencing different industries or job roles. Moreover, ongoing data collection and analysis could provide longitudinal insights, tracking trends and changes in recruitment patterns over time.

In conclusion, this project not only provides valuable insights into resume acceptance and intern success factors but also lays the groundwork for continued research in this area. The knowledge gained can inform future decisions, contributing to more effective recruitment processes and better alignment between educational institutions and industry requirements.

7   APPENDICES

7.1   Appendix A – Data Cleaning Summary

Component	Description
Data Validation	Checked for missing or inconsistent entries.
Handling Missing Data	Filled missing values using appropriate methods such as mean or mode.
Handling Inconsistencies	Corrected inconsistencies in formatting and structure.
Outlier Detection	Identified and removed outliers using statistical methods.
Data Transformation	Converted categorical variables into numerical formats for analysis.
Data Integrity Check	Ensured dataset consistency and accuracy post-cleaning.