Cloud Counselage

Cloud Counselage

GIFT A CAREER FOUNDATION

INTERNSHIP PROGRAM 2025

PROJECT REPORT

DATA ANALYTICS

Socio-economic and Academic Relationship Analysis

Created By:	Aishwarya Devendra Wankhade	**Approved By:**	<Domain Lead Name>
Created On:	15-04-2025	**Approved On:**	DD-MMM-YYYY

INDEX

S.No	Title	Pg.No
1	PROJECT DETAILS	1
2	SUMMARY	1
3	INTRODUCTION	
3.1	Background	1
3.2	Stakeholders	1
3.3	Objectives	1
4	METHODOLOGY	
4.1	Considerations & Assumptions	1-2
4.2	Approach	2
4.3	Activities	2-3
5	TARGETED VS ACHIEVED OUTPUT	3
6	CONCLUSION	3

1 PROJECT DETAILS

Project Name	Socio-economic and Academic Relationship Analysis		
Project Sponsor	Tushar Topale		
Project Manager	Harshad Topale		
Start Date	18-01-2025	**Completion Date**	18-04-2025

2 SUMMARY

This project aims to generate insights into how students' academic performance and extracurricular involvement align with their career aspirations, while examining how socio-economic and other related factors correlate with professional outcomes. The study addresses the growing challenge of high resume rejection rates by offering data-driven insights that can help students strengthen their career readiness in a competitive job market. The findings are intended to support both students and educational institutions in identifying key contributors to career success and enhancing academic planning and career guidance initiatives.

3 INTRODUCTION

3.1 Background:

Every year, millions of students apply for internships and employment opportunities, where resumes serve as a critical factor in forming first impressions. Recruiters generally spend only a few minutes reviewing a resume once it is received through email, job portals, or applicant tracking systems. Studies indicate that over 70% of resumes are eliminated during the initial screening stage. As a result, careful attention to detail and dedicating adequate effort to crafting a strong resume are vital for standing out in an increasingly competitive job market.

3.2 Stakeholders:

Harshada Topale

3.3 Objectives:

The objective of this project is to analyze student intern data to understand the relationships between academic performance, socio-economic background, extracurricular participation, and career aspirations. Through exploratory data analysis, the project aims to identify key factors influencing students' expected career outcomes and provide data-driven insights that can support improved academic planning and career guidance.

4 METHODOLOGY

4.1 Considerations & Assumptions:

a) There is a potential risk of data loss during the data cleaning and preprocessing phase, especially if the dataset contains missing, inconsistent, or inaccurate values. Such issues may impact the quality of the analysis and the reliability of the insights generated.

b) The project operates within a fixed duration of 240 hours, which may impose limitations on the depth of analysis and timely completion of deliverables. Additionally, all findings will rely entirely on the available student dataset, which may restrict the breadth and accuracy of insights.

c) It is assumed that the provided dataset is sufficiently complete and representative, and that the available open-source data analysis tools and technologies are adequate to successfully execute and deliver the project within the defined timeframe.

4.2 Approach:

a) The project began with identifying the problem statement focused on the lack of insights into the relationship between students' socioeconomic background, academic performance, skills, and expected salary.

b) The Excel dataset containing multiple attributes related to individual students was explored and analyzed to understand its structure and variables.

c) The dataset was loaded into Power BI and cleaned using the Power Query Editor to handle missing values, inconsistencies, and ensure data quality before analysis.

d) Various data visualization techniques such as scatter plots, pie charts, donut charts, tables, and cards were utilized to address both basic and moderate analytical questions.

e) As part of the project deliverables, the methodology and insights derived from answering the basic and moderate questions were documented in a Word file.

f) Additionally, a demonstration video was created to showcase the process of analysing the dataset and answering the basic and moderate questions.

4.3 Activities:

a) The project execution commenced with the preparation of a Project Charter that defined the project goals, objectives, background, and scope. This document provided a structured foundation for the subsequent development of the Software Requirements Specification (SRS).

b) The next phase of the project involved the preparation of the Software Requirements Specification (SRS) document. This document detailed the hardware, software, functional, non-functional, and external interface requirements necessary to support smooth and structured project execution.

c) The subsequent phase focused on developing a Project Schedule to support effective planning, monitoring, and control of project progress. The schedule captured details such as task names, start and end dates, estimated effort in hours, duration in days, and the status of each activity.

d) The next phase involved identifying potential risks, constraints, and dependencies that could impact project delivery. A RAID log was prepared to document Risks, Assumptions, Issues, and Dependencies, along with planned mitigation actions to support smooth project execution.

e) The project schedule was then executed as planned by adhering to the defined software requirements and aligning with the project's goals, objectives, risks, and constraints. Tools such as Microsoft Excel, Power BI, and Microsoft Word were used to analyze the data and address the basic and moderate questions outlined in the problem statement.

f) Throughout the project execution, doubts, recommendations, and voluntary improvements were consistently documented in a Lessons Learned file to capture insights and support continuous project improvement.

5 TARGETED VS ACHIEVED OUTPUT

- For questions 1Q through 10Q, the achieved outputs matched the targeted outputs as defined in the problem statement.

- The targeted output for 11Q required a single visualization illustrating how expected salary varies with GPA, family income, and experience with Python (in months). However, the achieved output consisted of three separate scatter plots: expected salary vs GPA, expected salary vs family income, and expected salary vs experience with Python.

- The targeted output for 13Q was to answer the question using a single visualization without creating any new features from the dataset. In contrast, the achieved output included a newly derived feature, “13Q CGPA” (based on CGPA), along with the use of a pie chart and a scatter plot.

- The targeted output for 17Q was to address the question using a single visualization without generating additional features from the dataset. However, the achieved output involved a derived feature, “17Q CGPA” (based on CGPA), which was used to produce the result.

6 CONCLUSIONS

This project delivers meaningful insights to stakeholders by clarifying how academic performance, extracurricular participation, and technical skills such as Python influence career outcomes, including expected salary. By identifying trends in student performance, skill development, and engagement, the analysis enables educational institutions to better design academic advising, career support services, and co-curricular initiatives. The insights can further support more targeted and data-driven career guidance, particularly for students from underrepresented backgrounds or regions with limited access to resources.
The future scope of this project may involve incorporating machine learning techniques to predict individual career success probabilities using multi-dimensional student profiles, enabling predictive analytics for career outcomes. Additionally, the scope could be extended to recommend suitable career paths or internship opportunities by analyzing academic backgrounds, skill sets, and historical placement data, leading to dynamic and personalized career recommendation systems.