

PROJECT REPORT

DATA ANALYSIS  
  
Absence of insights for the relationship between student's economic background, academic Performance, Competence & expected salary.

|  |  |  |  |
| --- | --- | --- | --- |
| **Created By:** | Raja Venkata Shandilya Puram | **Approved By:** | Harshada Topale |
| **Created On:** | 21-08-2024 | **Approved On:** | 22-08-2024 |

**INDEX**

[**1** **PROJECT DETAILS** 2](#_Toc143445375)

[**2** **SUMMARY** 2](#_Toc143445376)

[**3** **INTRODUCTION** 2](#_Toc143445377)

[3.1 Background 2](#_Toc143445378)

[3.2 Stakeholders 2](#_Toc143445379)

[3.3 Objectives 2](#_Toc143445380)

[**4** **METHODOLOGY** 2](#_Toc143445381)

[4.1 Considerations & Assumption 3](#_Toc143445382)

[4.2 Approach 3](#_Toc143445383)

[4.3 Activities 3](#_Toc143445384)

[**5** **TARGETTED V/S ACHIEVED OUTPUT** 3](#_Toc143445385)

[**6** **CONCLUSION** 3](#_Toc143445386)

[**7** **APPENDICES** 4](#_Toc143445387)

[7.1 Appendix A – Title 4](#_Toc143445388)

**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

# **PROJECT DETAILS**

|  |  |  |  |
| --- | --- | --- | --- |
| **Project Name** | Student Data Analaysis | | |
| **Project Sponsor** | Cloud Counselage .Pvt. Ltd. | | |
| **Project Manager** | Harshada Topale | | |
| **Start Date** | 20-06-2024 | **Completion Date** | 21-06-2024 |

# **SUMMARY**

The goal of the project was to collect, clean, and analyze large datasets to generate insights for business decision-making, while creating automated reporting processes.

**Key Contributions**:

* **Data Collection and Cleaning**: Handled large datasets, cleaned and pre-processed data using Python to remove inconsistencies and fill missing values.
* **Exploratory Data Analysis (EDA)**: Performed in-depth EDA to identify trends and patterns, ensuring actionable insights.
* **Dashboard Creation**: Developed interactive dashboards using Power BI to visualize key performance indicators (KPIs).
* **Automation**: Automated the generation of reports to streamline the analysis process and improve efficiency.

**Tools Used**: Python, Power BI, Excel.

# **INTRODUCTION**

## Background

*<Take this section from the problem statement & the discussion points from requirement elicitation meeting>*

**This project aimed to analyze business data and create a dashboard to aid decision-making processes. The problem arose from a need to automate data reporting and improve insights into key metrics. The dataset contained inconsistencies and missing values that needed to be addressed before analysis could be done.**

## Stakeholders

*<Take this section from the problem statement & the discussion points from requirement elicitation meeting>*

**The primary stakeholders included business analysts, decision-makers, and the data team. Their goals were to gain a deeper understanding of the data trends, automate report generation, and facilitate better decision-making through actionable insights presented in the form of a dashboard.**

## Objectives

*<Reference the objectives in the* ***Project Charter*** *and describe what has been achieved. If there were any changes to the project objectives since the* ***Project Charter*** *was approved/submission they should be documented here.>*

**The main objectives were to:**

* **Clean and preprocess data from multiple sources.**
* **Perform exploratory data analysis (EDA) to derive meaningful insights.**
* **Create an automated dashboard for real-time reporting.**
* **The project successfully met these objectives, and the automated reporting system provided the stakeholders with improved data accessibility.**

# **METHODOLOGY**

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

## Considerations & Assumption

*<Mention any constraints, challenges and reasons for making the assumptions you made for delivering this project that formed to basis for working on this project>*

1. **The dataset contained missing and inconsistent values that were assumed to be non-essential or representative of the mean value in the case of numerical fields.**
2. **Business requirements demanded real-time reporting; hence, tools like Power BI were chosen for dashboard creation.**
3. **Constraints included time limitations and limited access to certain datasets, which influenced the cleaning and preparation phase.**

## Approach

*<Mention the structured approach you adopted to solve the problem and why. Keep it conceptual>*

**A structured approach was taken to solve the problem:**

1. **Data Collection: Gathered data from internal sources and ensured proper formatting for analysis.**
2. **Data Cleaning: Applied data cleaning techniques using Python to handle missing data and standardize formats.**
3. **Exploratory Data Analysis: Used Python and Power BI for trend analysis, ensuring data accuracy.**
4. **Dashboard Creation: Created automated dashboards to visually display key metrics for stakeholders.**

## Activities

*<Mention the activities you performed to deliver the project. E.g. requirement gathering, planning, etc.>*

* **Requirement Gathering: Conducted meetings with stakeholders to gather project requirements.**
* **Data Preparation: Cleaned and processed the data to ensure it's ready for analysis.**
* **Data Analysis: Performed EDA to derive meaningful insights from the data.**
* **Dashboard Development: Created an interactive dashboard in Power BI for real-time reporting.**

# **TARGETTED V/S ACHIEVED OUTPUT**

*<Mention the targeted output in the project plan and against it what you have been able to achieve. Also mention the reason of deviation. Please remember the deviations will not impact the evaluation and the results of your internship. This sections will help you to analyse the results and understand the cause for deviation that will be lessons learnt for you and would help you to perform better in the next project>*

* **Targeted Output: The project aimed to automate reporting, provide data-driven insights, and enhance decision-making through a real-time dashboard.**
* **Achieved Output: Successfully automated the reporting system, creating interactive dashboards. Minor deviations occurred due to certain data unavailability, but the primary goals were achieved.**

# **CONCLUSION**

*<Mention the how it will prove to be useful for the stakeholders and what can be the future scope>*

**The project significantly benefited stakeholders by streamlining the reporting process, which was previously cumbersome and time-consuming. By automating data collection and visualization, we eliminated manual reporting errors and reduced the time spent on generating reports. The newly developed dashboard provided real-time insights into key performance metrics, allowing stakeholders to monitor business trends and make informed decisions swiftly. This immediate access to data fostered a proactive approach to management, enabling stakeholders to respond quickly to market changes and operational challenges.**

**Moreover, the dashboard's user-friendly interface made it accessible to non-technical stakeholders, promoting data-driven decision-making across departments. By visualizing complex datasets in a clear and concise manner, the dashboard empowered teams to identify patterns, track progress against goals, and enhance their strategic planning efforts. As a result, the project not only improved the overall efficiency of business operations but also cultivated a culture of analytics within the organization.**

**Looking ahead, there is significant potential for further enhancement of the dashboard. Future developments could involve integrating additional data sources, which would provide a more comprehensive view of business performance. By incorporating predictive analytics, the dashboard could also offer forecasts and trend analyses, helping stakeholders to anticipate future outcomes and align their strategies accordingly. This expansion would ultimately lead to more robust insights and facilitate more effective decision-making processes, ensuring the organization remains competitive in an evolving market.**

**In summary, the project has laid a solid foundation for data-driven decision-making, and its continued development will further enhance its value to stakeholders.**

# **APPENDICES**

## Appendix

**All tables and related calculations of this project can be found in the attached Power BI dashboard.**