Data Collection and Preprocessing Phase

Date	06 July 2024
Team ID	739973
Project Title	SmartLender – Envisioning Success: Predicting University Scores With Machine Learning
Maximum Marks	2 Marks

Data Collection Plan & Raw Data Sources Identification Template

Elevate your data strategy with the Data Collection plan and the Raw Data Sources report, ensuring meticulous data curation and integrity for informed decision-making in every analysis and decision-making endeavor.

Data Collection Plan Template

Section	Description
Project Overview	The goal of this project is to develop a machine learning model that can predict university scores based on various student attributes. The model will help identify the most important factors that influence student performance, enabling educators and policymakers to make data-driven decisions to improve student outcomes.

Data Collection Plan	Search for datasets related to Envisioning Success: Predicting University Scores With Machine Learning, and College Score predicting.

Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Kaggle, the popular platform for data science competitions and repositories. The provided sample data represents a subset of the collected information, encompassing variables such rainfall, temperature.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	The dataset comprises country, national_rank, quality_of_educatio -n, quality_of_faculty, and score	https://www.kaggle.com/datasets/mylesoneill/world-universityrankings	CSV	84KB	Public