

Data Collection and Preprocessing Phase

Date	14 July 2024
Team ID	739755
Project Title	Blood Donation Prediction
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 Blood Donation Prediction project aims to develop a predictive model to determine whether an individual is likely to donate blood in the future. By analyzing historical donation data and other relevant factors, the model can help blood banks and donation centers manage their resources efficiently, schedule donation drives, and maintain a steady supply of blood.
Data Collection Plan	search for datasets related to blood donation, Identify Data Sources, blood group and applicant details. Collection of Historical Donation data, Demographic Data, Health Data, Behavioural Data.

Raw Data Sources Identified	The raw data sources for this project are included datasets are obtained from Kaggle and GitHub repositories.
-----------------------------	---

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	The dataset include the features Recency, Frequency, Monetary, Time	https://www.kaggle.com/code/alperkaraca1/blood-donor-prediction/input	CSV	8.54 MB	Public
GitHub repository	Collected the flask part code from repository.	https://github.com/smartinternz02/SI-GuidedProject-579989-1694076075	CSV	4 kB	Public