Data Collection and Preprocessing Phase

Date	07th July 2024
Team ID	SWTID1720000556
Project Title	Predicting Co2 Emission by Countries Using 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 objective of the machine learning project is to examine CO2 emissions from various nations from 1960 to 2015.Using the data set with features such as country name and its code, indicator name and its code, year and the emission value. And here we are using different types of indicators to obtain efficient data.
Data Collection Plan	Get data directly from the source websites. Most data providers offer download options in various formats.
Raw Data Sources Identified	Datasets obtained from are included in the raw data sources for this project. Kaggle is a popular platform for data science competitions and repositories. The sample data provided represents a subset of

	collected information from different countries with different indicators at different times.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Kaggle Dataset	The dataset contains the value of co2 emitted in different countries with its country codes at different times using different indicators	https://www.kaggle.com/code/ashukr/exploring-co2-emission/input?select=Indicators.csv	CSV	527 MB	Public