Data Collection and Preprocessing Phase

Date	17 July 2024
Project Title	Machine Learning Approach For Predicting The Price Of Natural Gas
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 natural gas prediction project aims to forecast natural gas consumption and production based on various environmental, operational, and economic features. Using datasets that include factors such as temperature, pressure, historical consumption, and production metrics, the goal is to build a model that accurately predicts natural gas usage and supply, enhancing operational decisions and efficiency.
Data Collection Plan	• Search for datasets related to natural gas consumption, production, and environmental factors. • Prioritize datasets with comprehensive feature sets, including variables such as temperature, pressure, historical usage, and production data. • Ensure datasets cover a diverse range of geographic locations and operational conditions for robustness.

Raw Data Sources Identified	The raw data sources for this project include datasets obtained from Github, well-known platforms for data science competitions and repositories. The provided sample data represents a subset of the collected information, encompassing variables such as temperature, pressure, historical consumption, and production metrics.

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Dataset	Contains natural gas consumption and production data including features such as temperature, pressure, and historical metrics.	https://drive.google.com/file/d/1jb4n2QgMR5GpL101Cv9AasZPrjf8l1Oo/view?usp=sharing	CSV	364 KB	Public