

Data Collection and Preprocessing Phase

Date	11 July 2024
Team ID	SWTID1720359900
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 project aims to develop a machine learning model for predicting the price of natural gas. Natural gas prices are influenced by various factors, such as supply and demand dynamics, geopolitical events, weather conditions, and economic indicators. By harnessing historical price data along with relevant contextual information, the machine learning model can learn patterns and relationships to forecast future price movements accurately.
Data Collection Plan	We plan to obtain the data containing natural gas prices provided to us by SmartInternz and use it to predict future price movements.

Raw Data Sources Identified	Raw data obtained from Google SmartInternz contains natural gas prices from year 1997 – 2020 for each date apart from a few null values
-----------------------------	---

Raw Data Sources Template

Source Name	Description	Location/URL	Format	Size	Access Permissions
Gas prices by SmartInternz (daily_csv.csv)	This data gives us price of Natural gas from the year 1997 – 2002.	https://drive.google.com/file/d/1jb4n2QgMR5GpL101Cv9AasZPrjf8l1Oo/view	CSV	365 KB	Public