Building a project to load and preprocess the dataset to measure the energy consumption

1. Import Necessary Libraries:

Start by importing the necessary Python libraries. You'll need libraries like Pandas for data manipulation, Matplotlib for visualization, and any specific libraries for working with energy consumption data.

2.Load the Dataset:

Start by importing the necessary Python libraries. You'll need libraries like Pandas for data manipulation, Matplotlib for visualization, and any specific libraries for working with energy consumption data.

3. Data Preprocessing:

Data preprocessing is essential to clean and prepare the dataset for analysis. Depending on the dataset, you may need to handle missing values, remove duplicates, and format columns.This include:

- Handling missing data.
- Converting data types.
- Removing outliers.
- Normalizing or scaling data.

Program to measure the energy consumption (Only Loading and Preprocessing):

#Step 1: Importing the necessary libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error, mean_squared_error


# Step 2: Load the Dataset
data = pd.read_csv("PJMW_hourly.csv")


# Step 3:Data Preprocessing
# Drop missing values, if any
data.dropna(inplace=True)
# Convert the date column to a datetime object
data['Date'] = pd.to_datetime(data['Date'])
# Calculate daily energy consumption
daily_energy_consumption = data.groupby('Date')['Consumption'].sum().reset_index()
# Calculate monthly energy consumption
monthly_energy_consumption =
data.groupby(data['Date'].dt.to_period('M'))['Consumption'].sum().reset_index()
# Save the preprocessed data to new CSV files, if needed
daily_energy_consumption.to_csv('daily_energy_consumption.csv', index=False)
monthly_energy_consumption.to_csv('monthly_energy_consumption.csv', index=False)
```