**ELECTRICITY PRICES PREDICTION**

**Problem Statement :**

The problem is to develop a predictive model that uses historical electricity prices and relevant factors to forecast future electricity prices.The objective is to create a tool that assists both energy providers and consumers in making informed decisions regarding consumption and investment by predicting future electricity prices. This project involves data preprocessing, feature engineering, model selection, training,and evaluation.

**Design Thinking :**

The Design Thinking process for Electricity Prices Prediction involves Data Source, Data Preprocessing ,Feature Engineering, Model Selection, Model Training and Evaluation.

**Data Source:**

To develop a predictive model for forecasting electricity prices, relevant data sources play a pivotal role. These sources encompass official energy market databases, utility companies, and government agencies, which offer historical electricity price data.

**Data Preprocessing:**

In the context of forecasting electricity prices, data preprocessing involves several crucial steps to ensure data quality and model readiness. This includes handling missing values, removing duplicates, converting timestamps into datetime objects, encoding categorical variables, detecting and addressing outliers, and engineering relevant features.

**Feature Engineering:**

In the context of forecasting electricity prices, feature engineering involves crafting informative attributes from the raw data to enhance the predictive model's accuracy.

**Model Selection:**

In selecting the most suitable model for forecasting electricity prices, several options tailored to the specific dataset and problem should be considered. Time series models like ARIMA and SARIMA are adept at capturing temporal dependencies and seasonality. Machine learning models such as Random Forest and Gradient Boosting excel in capturing complex relationships and feature importance.

**Model Training:**

Model training for electricity price forecasting begins by preparing historical data and optimising hyperparameters if necessary. The model is trained on the training dataset, and its performance is assessed on a validation set using appropriate metrics. Iterative adjustments and hyperparameter tuning follow based on validation results.

**Evaluation:**

Evaluating a predictive model for electricity price forecasting entails assessing its accuracy and practicality. This involves choosing appropriate metrics like Mean Absolute Error (MAE) or Mean Squared Error (MSE) and validating the model's performance on separate test data.

**Dataset :**

We have used the [kaggle Dataset](https://www.kaggle.com/datasets/chakradharmattapalli/electricity-price-prediction) for “Electricity Prices Prediction”

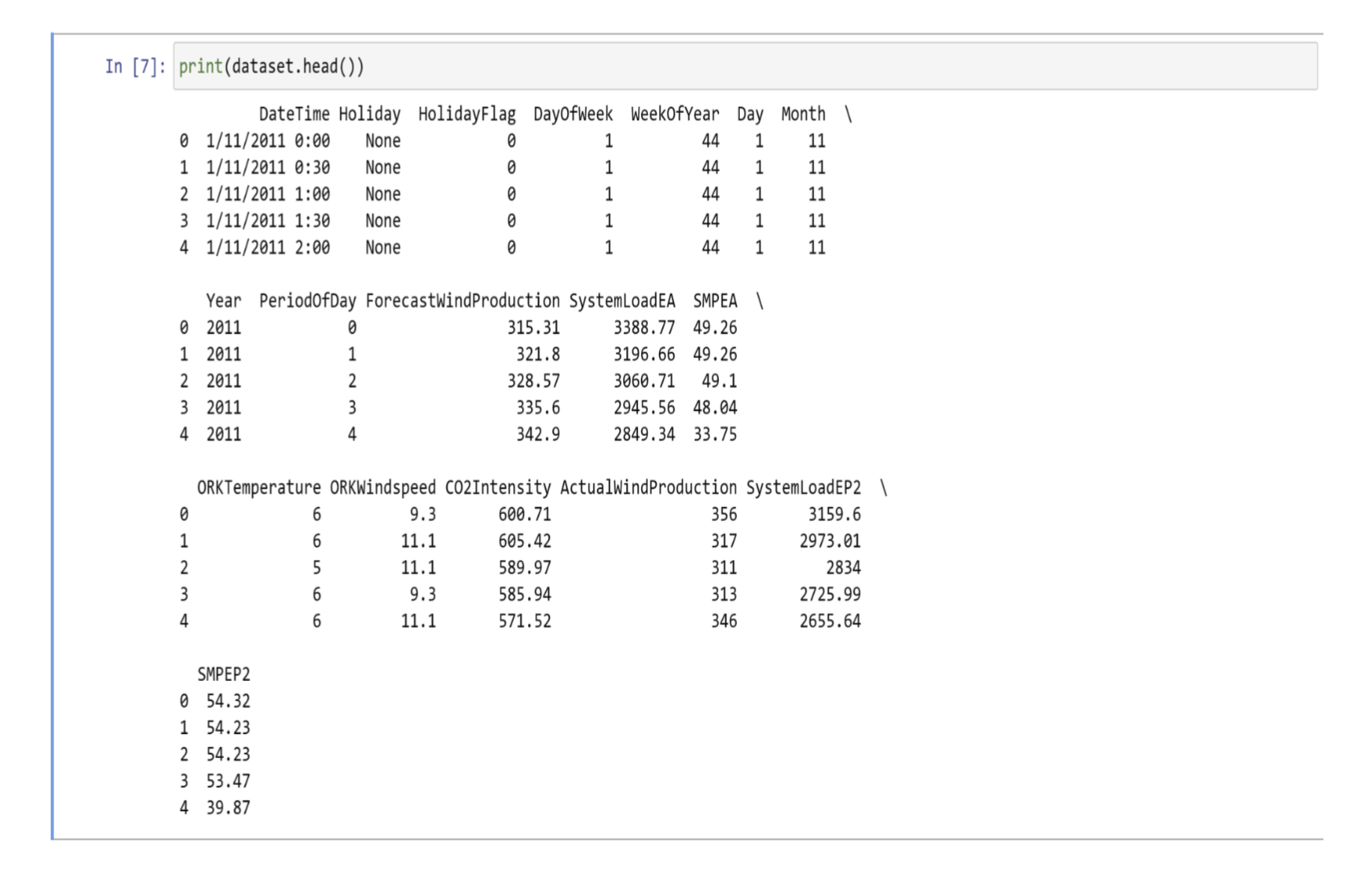
**Import Required Libraries:**

Start by importing the necessary Python libraries.



**Load the Dataset:**

Load the electricity price dataset into a Pandas DataFrame.



**Data Exploration:**

It's important to understand our data. Check for missing values,explore

the structure of the dataset, and look at some sample records.

print(dataset.head()) # Display the first few rows

print(dataset.info()) # Check for missing values

**Data Preprocessing:**

Depending on the dataset, you might need to perform various

preprocessing steps, including:

- Handling missing values: We can remove or impute missing values.

- Feature selection: Choose relevant features for prediction.

- Encoding categorical variables: Convert categorical data into

numerical form.

- Feature scaling: Normalise or standardise numerical features.

**Split the Dataset:**

Split our dataset into training and testing subsets to evaluate our model's

performance. A common split is 80% for training and 20% for testing.

**Model Building:**

Select a machine learning or deep learning model for electricity price

prediction, train it on the training data, and evaluate it on the testing

data.We might choose regression models, time series models, or neural

networks based on the nature of your data.

**Model Evaluation:**

Assess the model's performance using appropriate metrics (e.g., Mean

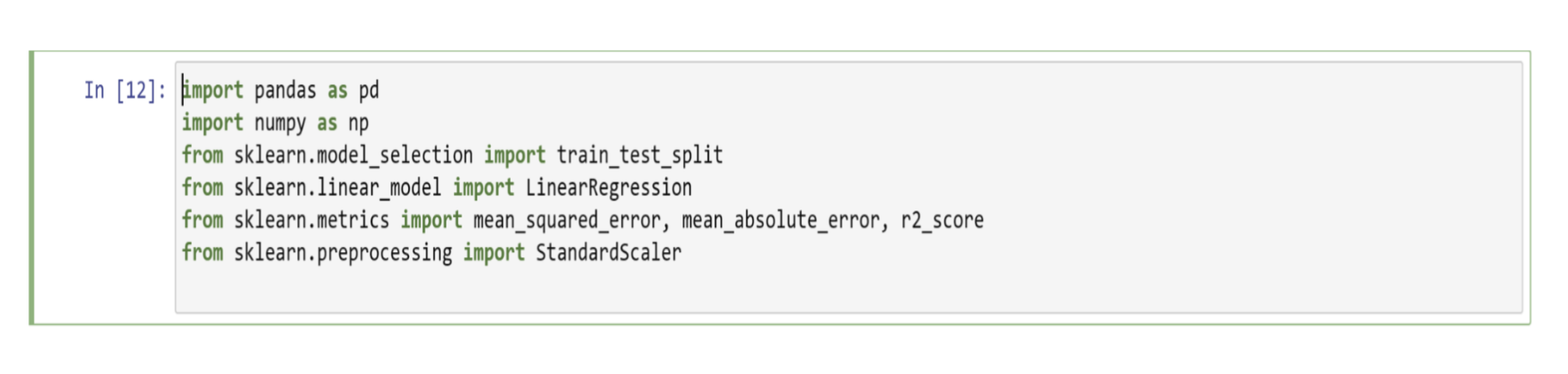
Absolute Error, Root Mean Squared Error) and make any necessary

improvements.

**Feature Engineering:**

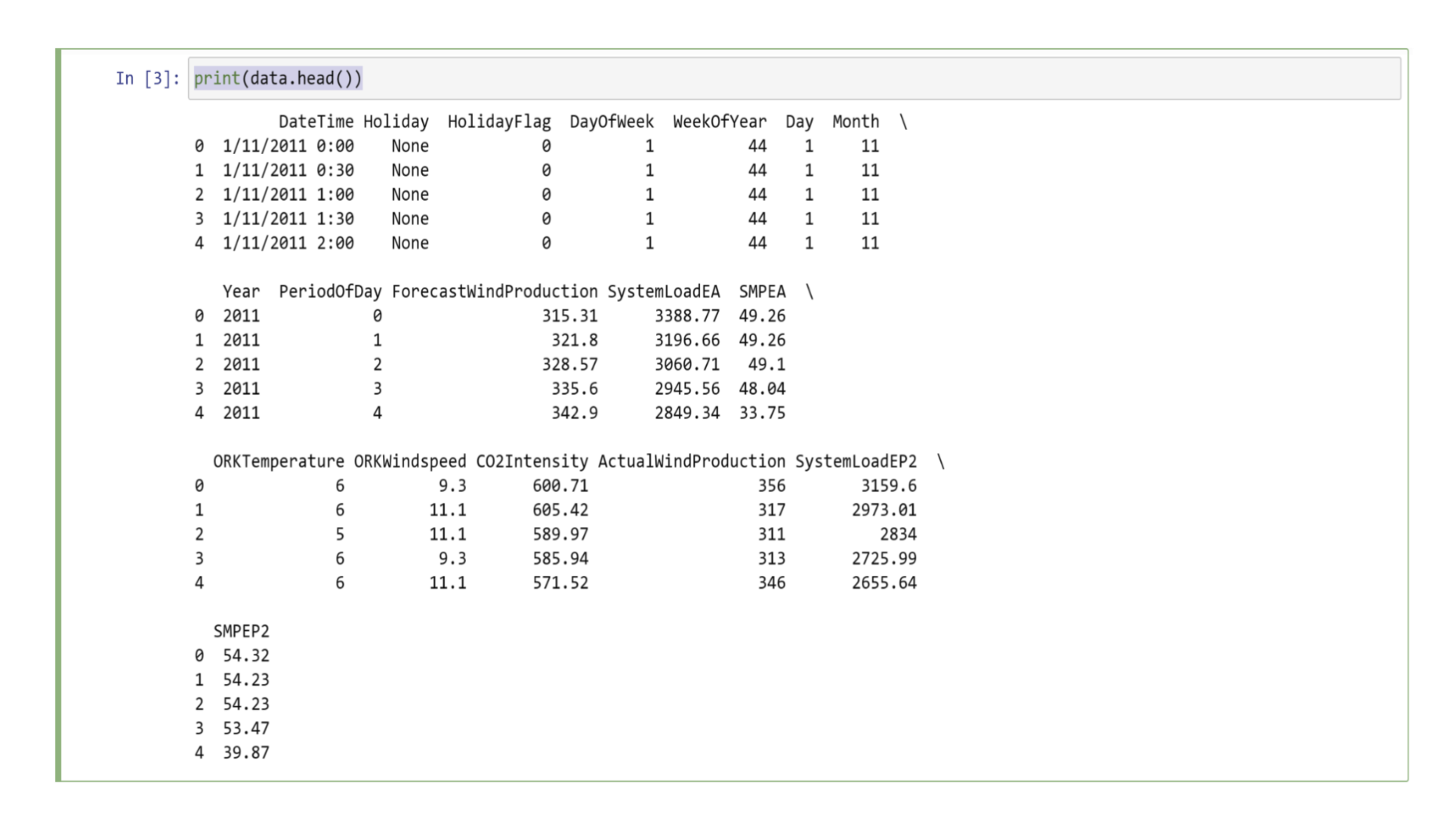
Choose the most relevant features for our prediction task

is called as Feature Engineering

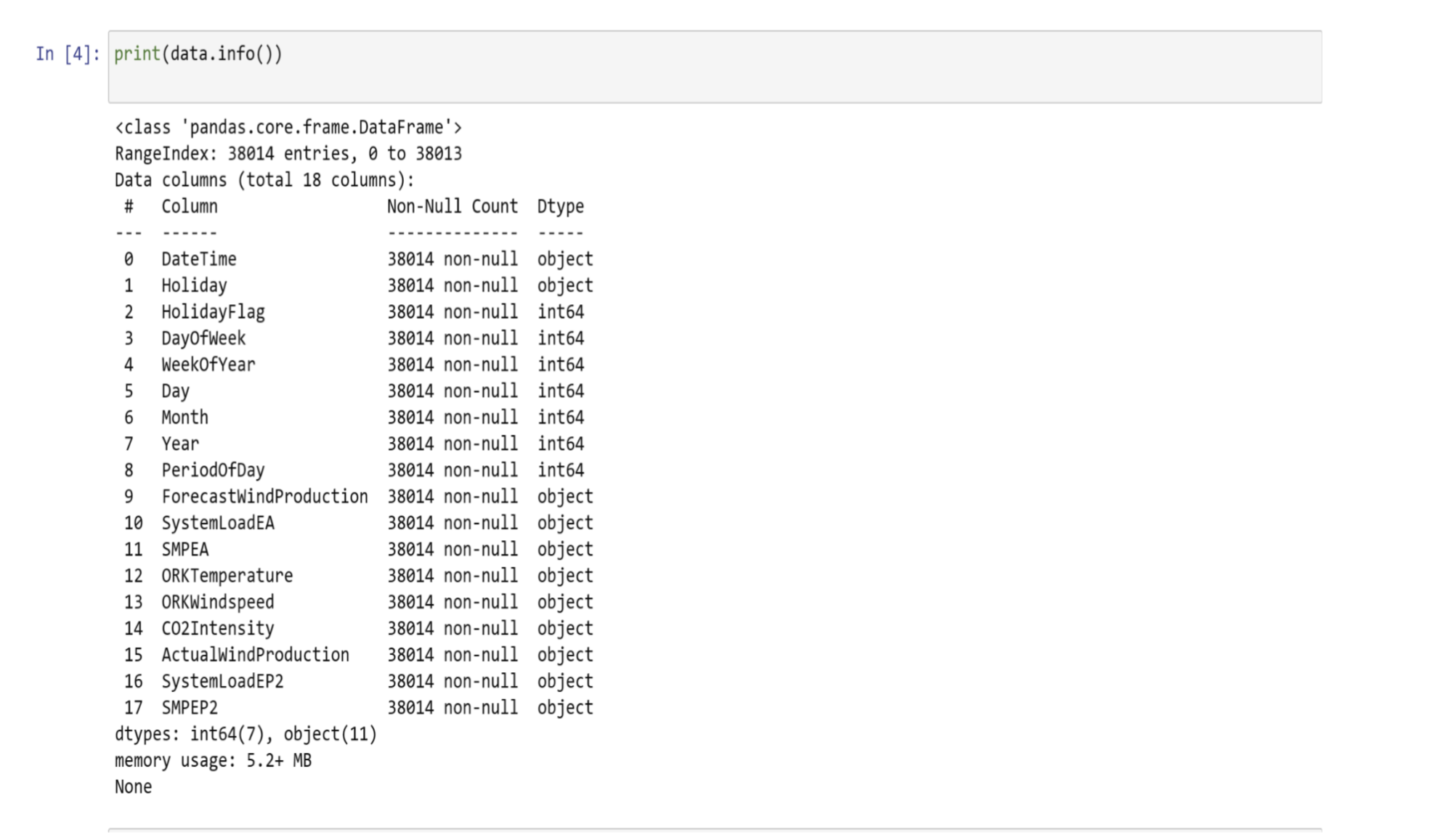
Before Feature Engineering,We have to load the dataset



Printing First Few rows,



Let’s have a look at all the columns of this dataset:



We can see that so many features with numerical values are

string values in the dataset and not integers or float values. So

before moving further, we have to convert these string values to

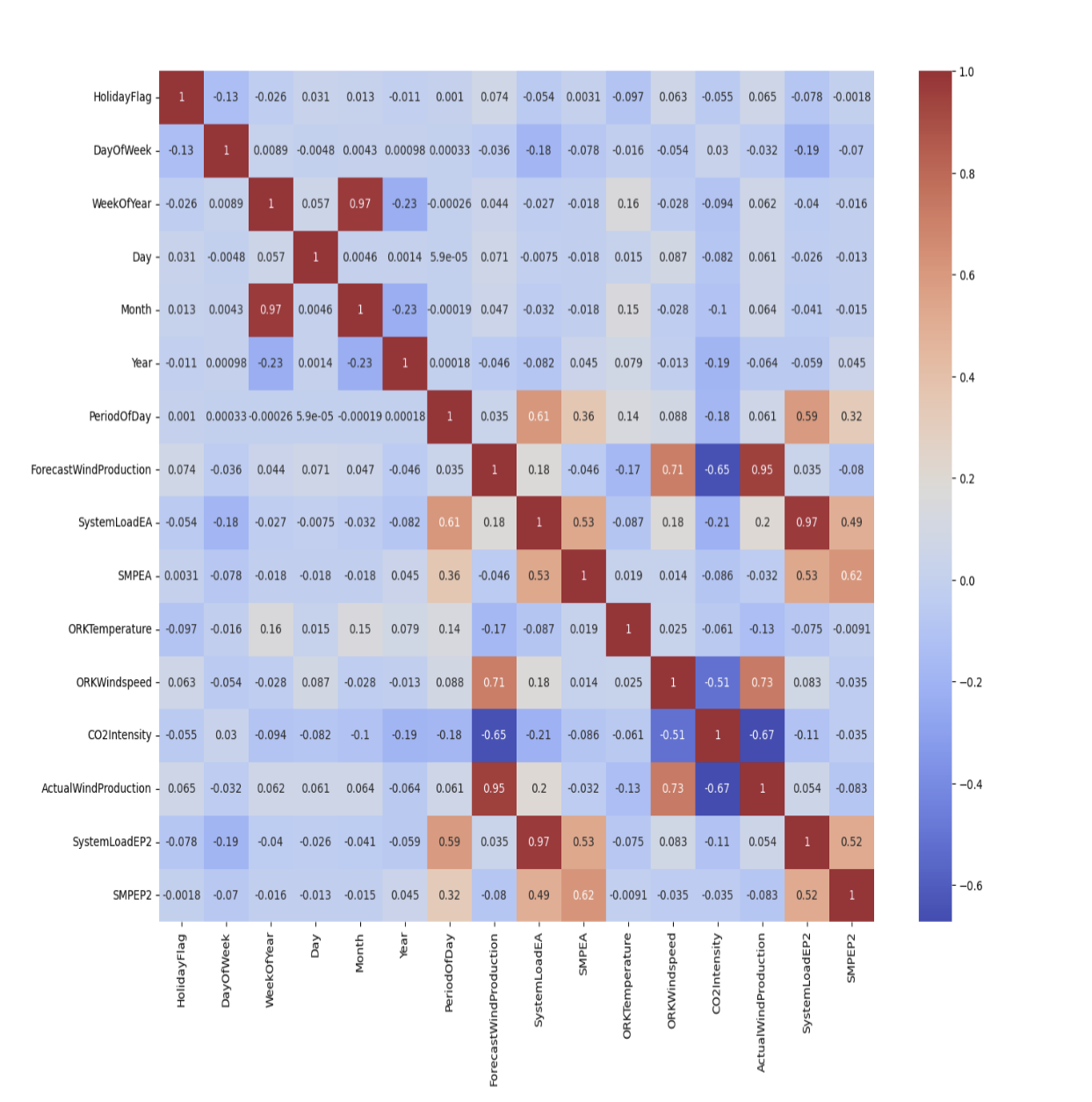
float values:



Now let’s have a look at whether this dataset contains any null

values or not:





**Model Training:**

Now let’s move to the task of training an electricity price

prediction model. Here we will first add all the important

features to x and the target column to y, and then we will split

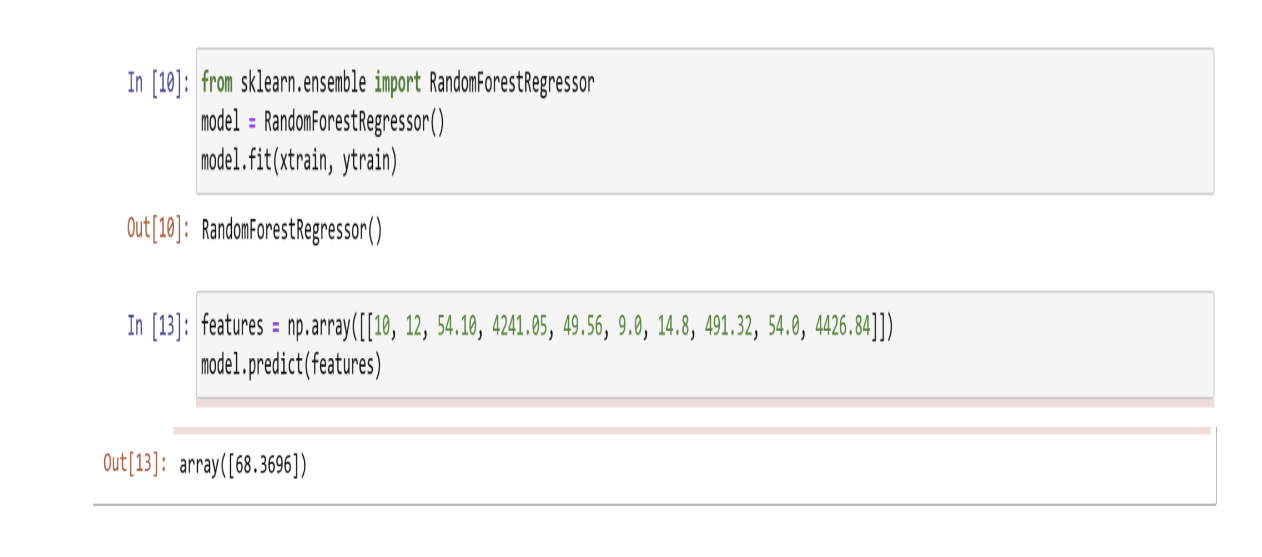
the data into training and test sets:



As this is the problem of regression, so here I will choose the

Random Forest regression algorithm to train the electricity price

prediction model:



**Model Evaluation:**

Evaluate the model's performance on the testing dataset

using appropriate metrics for regression tasks. Common

evaluation metrics include Mean Absolute Error (MAE), Root

Mean Squared Error (RMSE), and R-squared (R^2) for

regression models

