

```
In [1]: import pandas as pd
```

```
In [2]: df=pd.read_csv('global_energy_consumption.csv')
```

```
In [3]: df.head(20)
```

Out[3]:

	Country	Year	Total Energy Consumption (TWh)	Per Capita Energy Use (kWh)	Renewable Energy Share (%)	Fossil Fuel Dependency (%)	Industrial Energy Use (%)	Household Energy Use (%)	Er
0	Canada	2018	9525.38	42301.43	13.70	70.47	45.18	19.96	
1	Germany	2020	7922.08	36601.38	33.63	41.95	34.32	22.27	
2	Russia	2002	6630.01	41670.20	10.82	39.32	53.66	26.44	
3	Brazil	2010	8580.19	10969.58	73.24	16.71	30.55	27.60	
4	Canada	2006	848.88	32190.85	73.60	74.86	42.39	23.43	
5	UK	2016	4682.02	2491.92	40.47	69.06	20.07	23.39	
6	India	2024	9893.64	3912.56	27.07	31.70	57.59	18.56	
7	Canada	2008	7719.04	35819.59	15.03	62.73	57.93	21.51	
8	Russia	2020	9624.55	1708.80	77.36	63.39	48.42	24.01	
9	Brazil	2008	7818.38	29198.37	79.59	25.95	36.93	27.93	
10	Germany	2009	8761.33	42100.05	27.63	52.96	27.14	13.92	
11	Russia	2012	7094.64	19307.12	57.85	27.85	39.70	32.92	
12	Russia	2014	7628.97	2719.60	26.99	65.29	51.66	30.42	
13	India	2021	5279.46	35835.45	24.14	20.77	44.21	18.96	
14	Australia	2004	9249.95	46519.48	6.80	56.82	41.10	23.78	
15	Brazil	2002	8267.89	35999.98	54.06	32.82	53.98	38.41	
16	China	2007	8454.46	29133.14	17.75	12.53	32.53	10.61	
17	Russia	2019	6762.79	11886.43	6.05	18.66	54.40	34.34	
18	Australia	2009	3115.34	1261.09	13.91	55.84	37.58	25.35	
19	China	2012	452.73	34886.27	31.67	26.79	20.10	28.56	

```
In [4]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10000 entries, 0 to 9999
Data columns (total 10 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Country          10000 non-null   object  
 1   Year              10000 non-null   int64  
 2   Total Energy Consumption (TWh) 10000 non-null   float64 
 3   Per Capita Energy Use (kWh)    10000 non-null   float64 
 4   Renewable Energy Share (%)    10000 non-null   float64 
 5   Fossil Fuel Dependency (%)   10000 non-null   float64 
 6   Industrial Energy Use (%)   10000 non-null   float64 
 7   Household Energy Use (%)    10000 non-null   float64 
 8   Carbon Emissions (Million Tons) 10000 non-null   float64 
 9   Energy Price Index (USD/kWh)  10000 non-null   float64 
dtypes: float64(8), int64(1), object(1)
memory usage: 781.4+ KB
```

In [5]: `df.describe()`

Out[5]:

	Year	Total Energy Consumption (TWh)	Per Capita Energy Use (kWh)	Renewable Energy Share (%)	Fossil Fuel Dependency (%)	Industrial Energy Use (%)	Household Energy Use (%)	Carbon Emissions (Million Tons)	Energy Price Index (USD/kWh)
count	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000	10000.000000
mean	2012.151400	5142.564425	25039.950516	47.322925	44.932851	40.057585	30.217500	39.980000	12.057585
std	7.155236	2848.754040	14205.659284	24.603767	20.202800	11.538756	10.010000	20.000000	10.000000
min	2000.000000	100.480000	500.270000	5.000000	10.010000	20.000000	10.000000	20.000000	10.000000
25%	2006.000000	2713.882500	12683.220000	26.110000	27.337500	30.217500	20.000000	30.000000	10.000000
50%	2012.000000	5190.850000	25098.770000	47.150000	45.110000	39.980000	30.000000	40.000000	15.000000
75%	2018.000000	7579.977500	37113.282500	68.682500	62.430000	50.150000	40.000000	50.000000	20.000000
max	2024.000000	9999.260000	49989.570000	90.000000	80.000000	60.000000	50.000000	60.000000	20.000000

In [6]: `df_model = df.drop(['Country', 'Year'], axis=1)`

In [7]: `df_model.isnull().sum()`

```
Out[7]: Total Energy Consumption (TWh)      0
        Per Capita Energy Use (kWh)       0
        Renewable Energy Share (%)      0
        Fossil Fuel Dependency (%)     0
        Industrial Energy Use (%)     0
        Household Energy Use (%)      0
        Carbon Emissions (Million Tons)  0
        Energy Price Index (USD/kWh)    0
        dtype: int64
```

In [8]: `from sklearn.preprocessing import StandardScaler`

```
In [9]: scaler = StandardScaler()  
X_scaled = scaler.fit_transform(df_model.drop('Carbon Emissions (Million Tons)', axis=1))
```

```
In [10]: X = pd.DataFrame(X_scaled, columns=df_model.columns.drop('Carbon Emissions (Million Tons)'))  
y = df_model['Carbon Emissions (Million Tons)']
```

```
In [11]: X.head()
```

Out[11]:

	Total Energy Consumption (TWh)	Per Capita Energy Use (kWh)	Renewable Energy Share (%)	Fossil Fuel Dependency (%)	Industrial Energy Use (%)	Household Energy Use (%)	Energy Price Index (USD/kWh)
0	1.538580	1.215174	-1.366645	1.264103	0.443953	-0.591314	-1.173507
1	0.975744	0.813902	-0.556566	-0.147653	-0.497270	-0.322603	-1.479374
2	0.522165	1.170736	-1.483706	-0.277839	1.178905	0.162472	-0.102970
3	1.206772	-0.990526	1.053431	-1.397047	-0.824011	0.297410	1.502834
4	-1.507290	0.503409	1.068064	1.481411	0.202148	-0.187666	1.579301

```
In [12]: y.head()
```

Out[12]: 0 3766.11
1 2713.12
2 885.98
3 1144.11
4 842.39
Name: Carbon Emissions (Million Tons), dtype: float64

```
In [ ]:
```