Objective:

To analyze the correlation between AEP (Annual/Average Energy Production) and various input features in the dataset. This helps in identifying the most significant features that influence AEP, enabling effective feature selection for modeling.

Introduction:

Correlation analysis is a fundamental statistical tool used to understand the strength and direction of linear relationships between variables. In this lab, we aim to explore how various factors such as temperature, wind speed, pressure, and other environmental or operational variables correlate with AEP. This is especially useful in renewable energy systems such as wind farms or solar plants, where accurate modeling of AEP is essential.

```
import pandas as pd
import numpy as np
#from dataprep.datasets import load_dataset   # Used to import / load
dataset
#from dataprep.eda import create_report   # Used to load Report
formate

df=pd.read_csv(r'C:\Users\PMLS\ML\LAB5\5_features_extracted.csv',
index_col=['Datetime'], parse_dates=['Datetime'])
df.head()
```

Datetime	aep	year_day	holiday	weekend	winter	spring
2004-10-01 01:00:00	12379.0	275	0	0	0	0
2004-10-01 02:00:00	11935.0	275	0	0	0	0
2004-10-01 03:00:00	11692.0	275	0	0	0	0
2004-10-01 04:00:00	11597.0	275	0	0	0	0
2004-10-01 05:00:00	11681.0	275	0	0	0	0

Datetime	summer	fall	hour	month	day_of_week
2004-10-01 01:00:00	0	1	1	10	4
2004-10-01 02:00:00	0	1	2	10	4
2004-10-01 03:00:00	0	1	3	10	4
2004-10-01 04:00:00	0	1	4	10	4
2004-10-01 05:00:00	0	1	5	10	4

```
print('pearson\n\n',df.corrwith(df["aep"],method="pearson"))
```

```
pearson

 aep             1.000000
year_day        -0.124617
holiday         -0.053212
weekend         -0.267287
winter           0.328332
spring          -0.246582
summer           0.139793
fall            -0.220913
hour             0.421008
month           -0.125896
day_of_week     -0.220016
dtype: float64
```

```
print('kendall\n\n',df.corrwith(df["aep"],method="kendall"))
```

```
kendall

 aep             1.000000
year_day        -0.082592
holiday         -0.041101
weekend         -0.218066
winter           0.278561
spring          -0.197749
summer           0.098140
fall            -0.178467
hour             0.292545
month           -0.085584
day_of_week     -0.158129
dtype: float64
```

```
print('spearman\n\n',df.corrwith(df["aep"],method="spearman"))
```

```
spearman

 aep             1.000000
year_day        -0.123249
holiday         -0.050335
weekend         -0.267060
winter           0.341147
spring          -0.242178
summer           0.120189
fall            -0.218564
hour             0.431821
month           -0.123443
day_of_week     -0.219619
dtype: float64
```

```
df.isnull().sum()
```

```
aep             0
year_day        0
holiday         0
weekend         0
winter          0
spring          0
summer          0
fall            0
hour            0
month           0
day_of_week     0
dtype: int64
```