In [12]:

```
import pandas as pd
import numpy as np
data = pd.read_csv(r"C:\Users\lenovo\Downloads\electricity.csv")
print(data.head())
```

```
           DateTime Holiday  HolidayFlag  DayOfWeek  WeekOfYear  Day  M
onth  \
0  01/11/2011 00:00    None            0          1          44    1
11
1  01/11/2011 00:30    None            0          1          44    1
11
2  01/11/2011 01:00    None            0          1          44    1
11
3  01/11/2011 01:30    None            0          1          44    1
11
4  01/11/2011 02:00    None            0          1          44    1
11

   Year  PeriodOfDay ForecastWindProduction SystemLoadEA  SMPEA  \
0  2011            0                315.31      3388.77  49.26
1  2011            1                321.80      3196.66  49.26
2  2011            2                328.57      3060.71  49.10
3  2011            3                335.60      2945.56  48.04
4  2011            4                342.90      2849.34  33.75

  ORKTemperature ORKWindspeed CO2Intensity ActualWindProduction SystemL
oadEP2  \
0           6.00         9.30       600.71               356.00       3
159.60
1           6.00        11.10       605.42               317.00       2
973.01
2           5.00        11.10       589.97               311.00       2
834.00
3           6.00         9.30       585.94               313.00       2
725.99
4           6.00        11.10       571.52               346.00       2
655.64

  SMPEP2
0  54.32
1  54.23
2  54.23
3  53.47
4  39.87

C:\Users\lenovo\anaconda3\lib\site-packages\IPython\core\interactiveshe
ll.py:3165: DtypeWarning: Columns (9,10,11,14,15,16,17) have mixed type
s.Specify dtype option on import or set low_memory=False.
  has_raised = await self.run_ast_nodes(code_ast.body, cell_name,
```

In [13]:

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 38014 entries, 0 to 38013
Data columns (total 18 columns):
 #   Column                  Non-Null Count  Dtype 
---  ------                  --------------  ----- 
 0   DateTime                38014 non-null  object
 1   Holiday                 38014 non-null  object
 2   HolidayFlag             38014 non-null  int64 
 3   DayOfWeek               38014 non-null  int64 
 4   WeekOfYear              38014 non-null  int64 
 5   Day                     38014 non-null  int64 
 6   Month                   38014 non-null  int64 
 7   Year                    38014 non-null  int64 
 8   PeriodOfDay             38014 non-null  int64 
 9   ForecastWindProduction  38014 non-null  object
 10  SystemLoadEA            38014 non-null  object
 11  SMPEA                   38014 non-null  object
 12  ORKTemperature          38014 non-null  object
 13  ORKWindspeed            38014 non-null  object
 14  CO2Intensity            38014 non-null  object
 15  ActualWindProduction    38014 non-null  object
 16  SystemLoadEP2           38014 non-null  object
 17  SMPEP2                  38014 non-null  object
dtypes: int64(7), object(11)
memory usage: 5.2+ MB
```

In [3]:

```
data["ForecastWindProduction"] = pd.to_numeric(data["ForecastWindProduction"], errors=
data["SystemLoadEA"] = pd.to_numeric(data["SystemLoadEA"], errors= 'coerce')
data["SMPEA"] = pd.to_numeric(data["SMPEA"], errors= 'coerce')
data["ORKTemperature"] = pd.to_numeric(data["ORKTemperature"], errors= 'coerce')
data["ORKWindspeed"] = pd.to_numeric(data["ORKWindspeed"], errors= 'coerce')
data["CO2Intensity"] = pd.to_numeric(data["CO2Intensity"], errors= 'coerce')
data["ActualWindProduction"] = pd.to_numeric(data["ActualWindProduction"], errors= 'c
data["SystemLoadEP2"] = pd.to_numeric(data["SystemLoadEP2"], errors= 'coerce')
data["SMPEP2"] = pd.to_numeric(data["SMPEP2"], errors= 'coerce')
```

In [4]:

```
data.isnull().sum()
```

Out[4]:

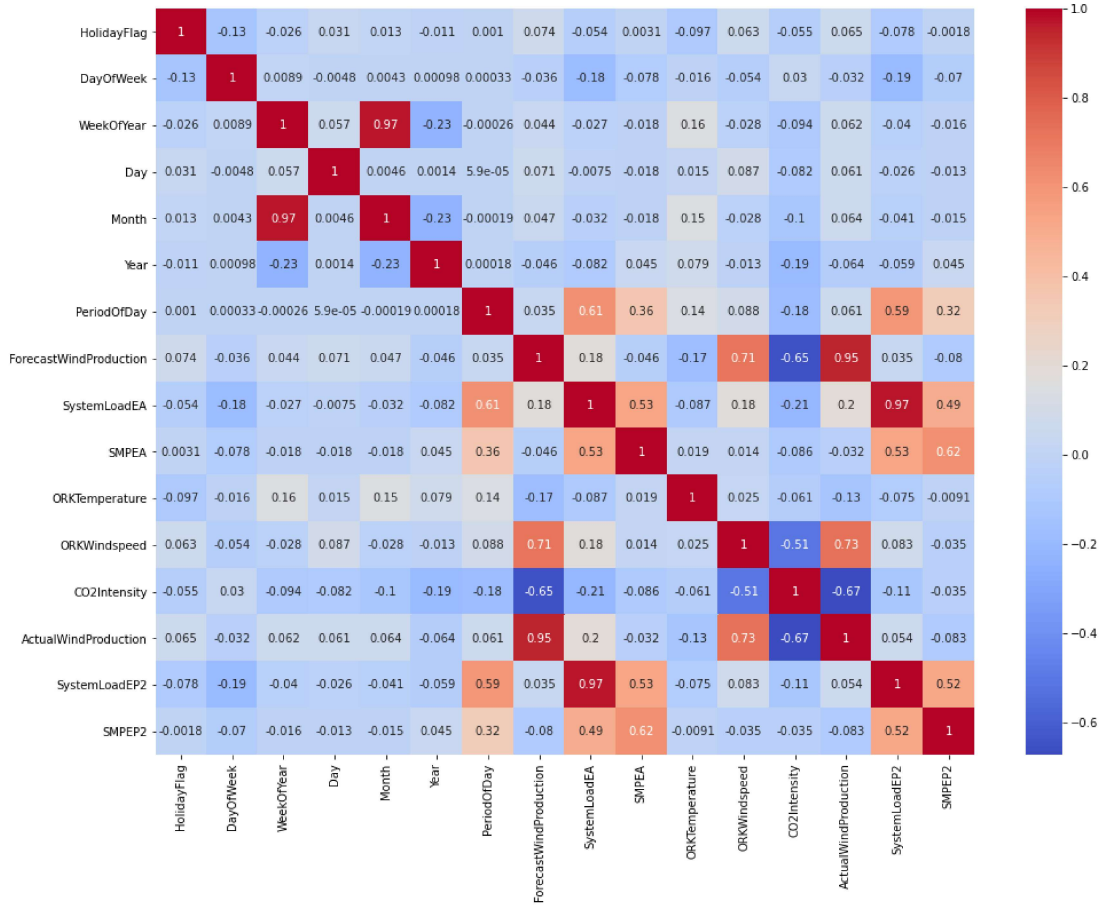
```
DateTime                    0
Holiday                     0
HolidayFlag                 0
DayOfWeek                   0
WeekOfYear                  0
Day                         0
Month                       0
Year                        0
PeriodOfDay                 0
ForecastWindProduction      5
SystemLoadEA                2
SMPEA                       2
ORKTemperature            295
ORKWindspeed              299
CO2Intensity                7
ActualWindProduction        5
SystemLoadEP2               2
SMPEP2                      2
dtype: int64
```

In [5]:

```
data = data.dropna()
```

In [6]:

```
import seaborn as sns
import matplotlib.pyplot as plt
correlations = data.corr(method='pearson')
plt.figure(figsize=(16, 12))
sns.heatmap(correlations, cmap="coolwarm", annot=True)
plt.show()
```



In [7]:

```
x = data[["Day", "Month", "ForecastWindProduction", "SystemLoadEA",
          "SMPEA", "ORKTemperature", "ORKWindspeed", "CO2Intensity",
          "ActualWindProduction", "SystemLoadEP2"]]
y = data["SMPEP2"]
from sklearn.model_selection import train_test_split
xtrain, xtest, ytrain, ytest = train_test_split(x, y,
                                                test_size=0.2,
                                                random_state=42)
```

In [8]:

```
from sklearn.ensemble import RandomForestRegressor
model = RandomForestRegressor()
model.fit(xtrain, ytrain)
```

Out[8]:

```
RandomForestRegressor()
```

In [9]:

```
luction", "SystemLoadEA", "SMPEA", "ORKTemperature", "ORKWindspeed", "CO2Intensity", "A
9.56, 9.0, 14.8, 491.32, 54.0, 4426.84]])
```

Out[9]:

```
array([68.9074])
```

In [ ]: