

## electricity prices:

1. DateTime: Date and time of the record
2. Holiday: contains the name of the holiday if the day is a national holiday
3. HolidayFlag: contains 1 if it's a bank holiday otherwise 0
4. DayOfWeek: contains values between 0-6 where 0 is Monday
5. WeekOfYear: week of the year
6. Day: Day of the date
7. Month: Month of the date
8. Year: Year of the date
9. PeriodOfDay: half-hour period of the day
10. ForecastWindProduction: forecasted wind production
11. SystemLoadEA forecasted national load
12. SMPEA: forecasted price
13. ORKTemperature: actual temperature measured

- 14. ORKWindspeed: actual windspeed measured
- 15. CO2Intensity: actual CO2 intensity for the electricity produced
- 16. ActualWindProduction: actual wind energy production
- 17. SystemLoadEP2: actual national system load
- 18. SMPEP2: the actual price of the electricity consumed (labels or values to be predicted)

So your task here is to use this data to train a machine learning model to predict the price of electricity consumed by the machines. In the section below, I will take you through the task of electricity price prediction with machine learning using Python.

---

# Electricity Price

## Prediction using Python

I will start the task of electricity price prediction by importing the necessary Python libraries and the dataset that we need for this task:

```
1 import pandas as pd
2 import numpy as np
3 data = pd.read_csv("https://",
4 print(data.head()))
```

	DateTime	Holiday	...
	SystemLoadEP2	SMPEP2	
0	01/11/2011 00:00	None	...
	3159.60	54.32	
1	01/11/2011 00:30	None	...
	2973.01	54.23	
2	01/11/2011 01:00	None	...
	2834.00	54.23	
3	01/11/2011 01:30	None	...
	2725.99	53.47	
4	01/11/2011 02:00	None	...
	2655.64	39.87	

[5 rows x 18 columns]

## electricity price prediction model:

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

---

```
RandomForestRegressor(bootstrap=True,  
                        ccp_alpha=0.0, criterion='mse',
```

```
                        max_depth=None,  
                        max_features='auto',  
                        max_leaf_nodes=None,
```

```
                        max_samples=None,  
                        min_impurity_decrease=0.0,
```

```
                        min_impurity_split=None,  
                        min_samples_leaf=1,
```

```
                        min_samples_split=2,  
                        min_weight_fraction_leaf=0.0,
```

```
                        n_estimators=100, n_jobs=None,  
                        oob_score=False,
```

```
                        random_state=None, verbose=0,  
                        warm_start=False)
```

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

```
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
```

So there are some columns with null values, I will drop all these rows containing null values from the dataset: