```
import pandas as pd


df21 = pd.read_csv('ASI - 2021.csv')
df21 = df21.drop(columns = [
    " Energy BOARD 1 3MW",
    " Energy BOARD 1 5MW",
    " Energy BOARD 10 3MW",
    " Energy BOARD 10 5MW",
    " Energy BOARD 11 3MW",
    " Energy BOARD 11 5MW",
    " Energy BOARD 12 5MW",
    " Energy BOARD 13 5MW",
    " Energy BOARD 14 5MW",
    " Energy BOARD 15 5MW",
    " Energy BOARD 16 5MW",
    " Energy BOARD 17 5MW",
    " Energy BOARD 18 5MW",
    " Energy BOARD 19 5MW",
    " Energy BOARD 2 3MW",
    " Energy BOARD 2 5MW",
    " Energy BOARD 3 3MW",
    " Energy BOARD 3 5MW",
    " Energy BOARD 4 3MW",
    " Energy BOARD 4 5MW",
    " Energy BOARD 5 3MW",
    " Energy BOARD 5 5MW",
    " Energy BOARD 6 3MW",
    " Energy BOARD 6 5MW",
    " Energy BOARD 7 3MW",
    " Energy BOARD 7 5MW",
    " Energy BOARD 8 3MW",
    " Energy BOARD 8 5MW",
    " Energy BOARD 9 3MW",
    " Energy BOARD 9 5MW",
    " Energy MSB  5MW 3200A",
    " Energy MSB 3MW",
    " Energy MSB 5MW 6300A"
]
)

df21["Time"] = pd.to_datetime(df21["Time"], format='mixed')
df21.set_index("Time", inplace = True)
df21.describe()
```

	Power BOARD 1 3MW	Power BOARD 1 5MW	Power BOARD 10 3MW	Power BOARD 10 5MW	Power BOARD 11 3MW	Power BOARD 11 5MW	Power 1
count	105120.000000	105120.000000	105120.000000	105120.00000	105120.000000	105120.000000	105120.0
mean	59.258257	16.343084	50.578415	41.58027	28.084237	42.159446	47.7
std	87.781739	23.733659	74.322410	59.72337	41.413960	62.427369	70.4
min	0.000000	0.000000	0.000000	0.00000	0.000000	0.000000	0.0
25%	0.000000	0.000000	0.000000	0.00000	0.000000	0.000000	0.0
50%	0.000000	0.000000	0.000000	0.00000	0.000000	0.000000	0.0
75%	99.000000	30.000000	87.000000	76.00000	49.000000	72.000000	82.0
max	312.000000	119.000000	266.000000	306.00000	160.000000	232.000000	270.0

8 rows × 33 columns

```
env21 = pd.read_csv("Weather Sensor - 2021.csv", encoding='unicode_escape')
env21 = env21[['Time', 'W/m^2', '°C']]

env21["Time"] = pd.to_datetime(env21["Time"], format='mixed')
env21.set_index("Time", inplace=True)
env21.describe()
```

```
<ipython-input-33-5637374e3e47>:1: DtypeWarning: Columns (3,5) have mixed types. Specify dtype option
  env21 = pd.read_csv("Weather Sensor - 2021.csv", encoding='unicode_escape')
```

	W/m^2	°C
count	105120.000000	105120.000000
mean	212.040099	31.129503
std	316.694294	10.176531
min	0.000000	10.600000
25%	0.000000	23.700000
50%	1.470000	25.350000
75%	358.972500	38.200000
max	1515.000000	70.190000

```
df22 = pd.read_csv('ASI - 2022.csv')
df22 = df22.drop(columns = [
    " Energy BOARD 1 3MW",
    " Energy BOARD 1 5MW",
    " Energy BOARD 10 3MW",
    " Energy BOARD 10 5MW",
    " Energy BOARD 11 3MW",
    " Energy BOARD 11 5MW",
    " Energy BOARD 12 5MW",
    " Energy BOARD 13 5MW",
    " Energy BOARD 14 5MW",
    " Energy BOARD 15 5MW",
    " Energy BOARD 16 5MW",
    " Energy BOARD 17 5MW",
    " Energy BOARD 18 5MW",
    " Energy BOARD 19 5MW",
    " Energy BOARD 2 3MW",
    " Energy BOARD 2 5MW",
    " Energy BOARD 3 3MW",
    " Energy BOARD 3 5MW",
    " Energy BOARD 4 3MW",
    " Energy BOARD 4 5MW",
    " Energy BOARD 5 3MW",
    " Energy BOARD 5 5MW",
    " Energy BOARD 6 3MW",
    " Energy BOARD 6 5MW",
    " Energy BOARD 7 3MW",
    " Energy BOARD 7 5MW",
    " Energy BOARD 8 3MW",
    " Energy BOARD 8 5MW",
    " Energy BOARD 9 3MW",
    " Energy BOARD 9 5MW",
    " Energy MSB  5MW 3200A",
    " Energy MSB 3MW",
    " Energy MSB 5MW 6300A"
]
)
```

```
df22["Time"] = pd.to_datetime(df22["Time"], format='mixed')
df22.set_index("Time", inplace = True)
df22.describe()
```

	Power BOARD 1 3MW	Power BOARD 1 5MW	Power BOARD 10 3MW	Power BOARD 10 5MW	Power BOARD 11 3MW	Power BOARD 11 5MW	Power 1
count	105120.000000	105120.000000	105120.00000	105120.000000	105120.000000	105120.000000	105120.0
mean	57.415525	20.282677	49.89493	51.872451	27.865059	40.771766	46.8
std	85.148061	30.216308	73.40230	75.668523	40.991699	60.019159	69.7
min	0.000000	0.000000	0.00000	0.000000	0.000000	0.000000	0.0
25%	0.000000	0.000000	0.00000	0.000000	0.000000	0.000000	0.0
50%	0.000000	0.000000	0.00000	0.000000	0.000000	0.000000	0.0
75%	98.000000	35.000000	87.00000	92.000000	49.000000	70.000000	81.0
max	312.000000	119.000000	266.00000	306.000000	160.000000	232.000000	270.0

8 rows × 34 columns

```
env22 = pd.read_csv("Weather Sensor - 2022.csv", encoding='unicode_escape')
env22 = env22[['Time', 'W/m^2', '°C']]

env22["Time"] = pd.to_datetime(env22["Time"], format='mixed')
env22.set_index("Time", inplace=True)
env22.describe()
```

	W/m^2	°C
count	103103.000000	103103.000000
mean	219.076441	30.612099
std	330.253935	9.374339
min	0.000000	19.800000
25%	0.000000	23.700000
50%	1.000000	25.300000
75%	369.000000	37.400000
max	1621.000000	64.100000

```
df23 = pd.read_csv('Generation - 2023 (January - August).csv')
df23 = df23.drop(columns = [
    " Energy BOARD 1 3MW",
    " Energy BOARD 1 5MW",
    " Energy BOARD 10 3MW",
    " Energy BOARD 10 5MW",
    " Energy BOARD 11 3MW",
    " Energy BOARD 11 5MW",
    " Energy BOARD 12 5MW",
    " Energy BOARD 13 5MW",
    " Energy BOARD 14 5MW",
    " Energy BOARD 15 5MW",
    " Energy BOARD 16 5MW",
    " Energy BOARD 17 5MW",
    " Energy BOARD 18 5MW",
    " Energy BOARD 19 5MW",
```

```
    " Energy BOARD 2 3MW",
    " Energy BOARD 2 5MW",
    " Energy BOARD 3 3MW",
    " Energy BOARD 3 5MW",
    " Energy BOARD 4 3MW",
    " Energy BOARD 4 5MW",
    " Energy BOARD 5 3MW",
    " Energy BOARD 5 5MW",
    " Energy BOARD 6 3MW",
    " Energy BOARD 6 5MW",
    " Energy BOARD 7 3MW",
    " Energy BOARD 7 5MW",
    " Energy BOARD 8 3MW",
    " Energy BOARD 8 5MW",
    " Energy BOARD 9 3MW",
    " Energy BOARD 9 5MW",
    " Energy MSB  5MW 3200A",
    " Energy MSB 3MW",
    " Energy MSB 5MW 6300A"
]
)

df23["Time"] = pd.to_datetime(df23["Time"], format='mixed')
df23.set_index("Time", inplace = True)
df23.describe()
```

	Power BOARD 1 3MW	Power BOARD 1 5MW	Power BOARD 10 3MW	Power BOARD 10 5MW	Power BOARD 11 3MW	Power BOARD 11 5MW	Power BOARD 12 5MW
count	69984.000000	69984.000000	69984.000000	69984.000000	69984.000000	69984.000000	69984.000000
mean	57.411637	20.226966	49.915852	51.893890	28.083833	42.150434	47.288637
std	84.443151	29.934526	73.037588	74.891627	41.071480	62.022557	69.640838
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
50%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
75%	98.000000	34.000000	86.000000	92.000000	49.000000	73.000000	81.000000
max	312.000000	118.000000	266.000000	306.000000	160.000000	231.000000	269.000000

8 rows × 33 columns

```
env23 = pd.read_csv("Weather sensor - 2023.csv", encoding='unicode_escape')
env23 = env23[['Time', 'W/m^2', '°C']]

env23["Time"] = pd.to_datetime(env23["Time"], format='mixed')
env23.set_index("Time", inplace=True)
env23.describe()
```

	W/m^2	°C
count	65664.000000	65664.000000
mean	226.456110	30.950434
std	337.325469	16.617770
min	0.000000	19.800000
25%	0.000000	23.900000
50%	1.000000	25.500000
75%	383.000000	37.800000
max	1521.000000	3536.000000

```
rows_with_nan = env22[env22.isna().any(axis=1)]
print(rows_with_nan.to_string())
```

```
2022-05-06 14:25:00    NaN NaN
2022-05-06 14:30:00    NaN NaN
2022-05-06 14:35:00    NaN NaN
2022-05-06 14:40:00    NaN NaN
2022-05-06 14:45:00    NaN NaN
2022-05-06 14:50:00    NaN NaN
2022-05-06 14:55:00    NaN NaN
2022-05-06 15:00:00    NaN NaN
2022-05-06 15:05:00    NaN NaN
```

```
rows_with_nan = env23[env23.isna().any(axis=1)]
print(rows_with_nan.to_string())
```