Energy Data Analysis

Group Assignment

Group 5

In 2019, we had a total of 4.3% days with a price below 10 euros.

```
4.383561643835616
```

```
energy["datetime"] = pd.to_datetime(energy["datetime"])
energy["date"] = energy["datetime"].dt.date

total_period = energy[
    (energy['spot_price'] < 10)
]

percentage_of_unique_days = (len(total_period['date'].unique())
                             /(len(energy['date'].unique())))*100
```

In 2019, we had:
- an average of 415 hours per month when the price was above the monthly average;
- an average of 313 hours per month when the price was below monthly average.

```
Total hours above the monthly average per month on average:
415.0833333333333
Total hours below the monthly average per month on average:
313.0
```

```python
energy_mth_avg["monthly_avg"] = energy_mth_avg['month'].map(lambda m: month_average[m])

energy_mth_avg["above_avg"] = np.where(
    energy_mth_avg["spot_price"] >= energy_mth_avg["monthly_avg"],
    1,
    0
)
```

Gas production details in 2019, when wind was above its hourly average.

```
When wind generation was above its hourly average,
gas production was higher than its hourly average 748 times.
gas production was lower than its hourly average 2864 times.
```

```
energy_gas_wind = energy_gas_wind[energy_gas_wind["wind_above_avg"] == "above"]

gas_above = energy_gas_wind[energy_gas_wind["gas_above_avg"]=="above"]["gas_above_avg"].count()
gas_below = energy_gas_wind[energy_gas_wind["gas_above_avg"]=="below"]["gas_above_avg"].count()
```

Spot price comparison when solar production was above its hourly average.

```
When solar generation was above its hourly average,
spot price was higher than its hourly average 1590 times.
spot price was lower than its hourly average 1656 times.
```

```
energy_spot_solar = energy_spot_solar[energy_spot_solar["solar_above_avg"] == "above"]

spot_above = energy_spot_solar[energy_spot_solar["spot_above_avg"]=="above"]["spot_above_avg"].count()
spot_below = energy_spot_solar[energy_spot_solar["spot_above_avg"]=="below"]["spot_above_avg"].count()
```

Spot price comparison when power demand was above its monthly average.

When power demand was above its monthly average,
spot price was higher than its monthly average 3521 times.
spot price was lower than its monthly average 988 times.

```
energy_spot_demand = energy_spot_demand[energy_spot_demand["pd_above_avg"] == "above"]

m_spot_above = energy_spot_demand[energy_spot_demand["m_spot_above_avg"]=="above"]["m_spot_above_avg"].count()
m_spot_below = energy_spot_demand[energy_spot_demand["m_spot_above_avg"]=="below"]["m_spot_above_avg"].count()
```

Average contribution of each power source in the year 2019.

```
Average contributon of nuclear power
27.98498897271831

Average contributon of gas power
25.161466845079374

Average contributon of solar power
4.106953788920754

Average contributon of hydro power
11.839694124985558

Average contributon of coal power
5.1665458675727

Average contributon of wind power
25.740350400723198
```

```python
total_energy_gen = energy[[("nuclear"),("gas"),("solar"),("hydro"),("coal"),("wind")]].fillna(0)

total_energy_gen["total_gen"] = total_energy_gen["nuclear"] + total_energy_gen["gas"] + total_energy_gen["solar"]
                                + total_energy_gen["hydro"] + total_energy_gen["coal"] + total_energy_gen["wind"]

for x in range(6):

    total_energy_gen[f"%_{total_energy_gen.columns[x]}"]
                                = (total_energy_gen[total_energy_gen.columns[x]] / total_energy_gen["total_gen"]) * 100
```

Average contribution of each power source in the most expensive month January.

```
Average contributon of nuclear power in January:
26.189345916305744

Average contributon of gas power in January:
15.917610128321028

Average contributon of solar power in January:
2.198412708846326

Average contributon of hydro power in January:
10.614893230476957

Average contributon of coal power in January:
15.525175616219466

Average contributon of wind power in January:
29.554562399830466
```

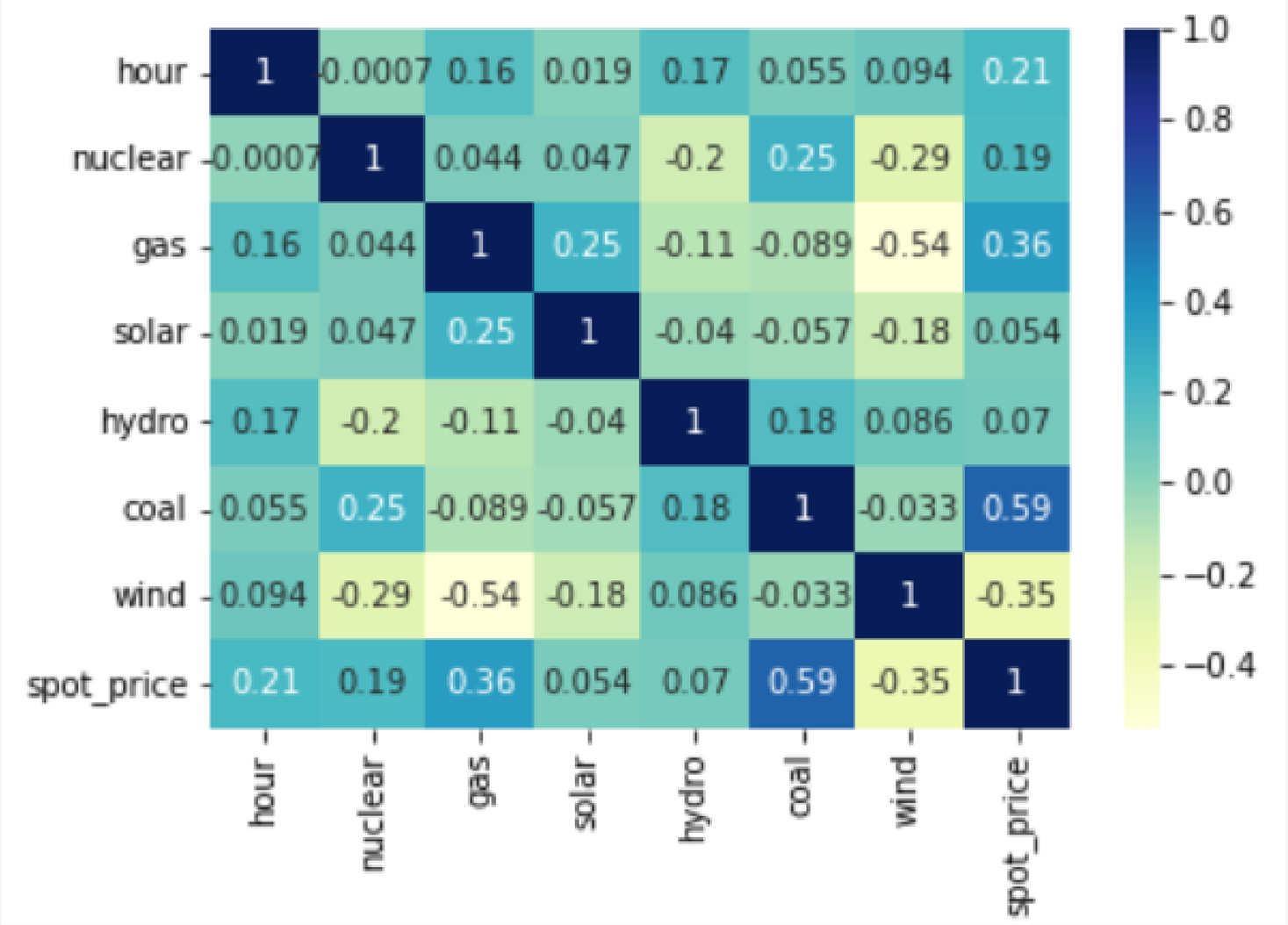
```
most_expensive = energy.groupby("month")["spot_price"].sum().idxmax()

total_energy_expensive = energy[[("nuclear"),("gas"),("solar"),("hydro"),("coal"),("wind"),("month")]].fillna(0)
                                                                                    [energy["month"] == most_expensive]
```

Average contribution of each power source in the cheapest month December.

```
Average contributon of nuclear power in December:
24.47696649497207

Average contributon of gas power in December:
14.997181123164474

Average contributon of solar power in December:
2.635510513873528

Average contributon of hydro power in December:
25.955194125443157

Average contributon of coal power in December:
2.2482461417390844

Average contributon of wind power in December:
29.68690160080767
```

```
cheapest_month = energy.groupby("month")["spot_price"].sum().idxmin()

total_energy_cheap = energy[[("nuclear"),("gas"),("solar"),("hydro"),("coal"),("wind"),("month")]].fillna(0)
                                                                    [energy["month"] == cheapest_month]
```

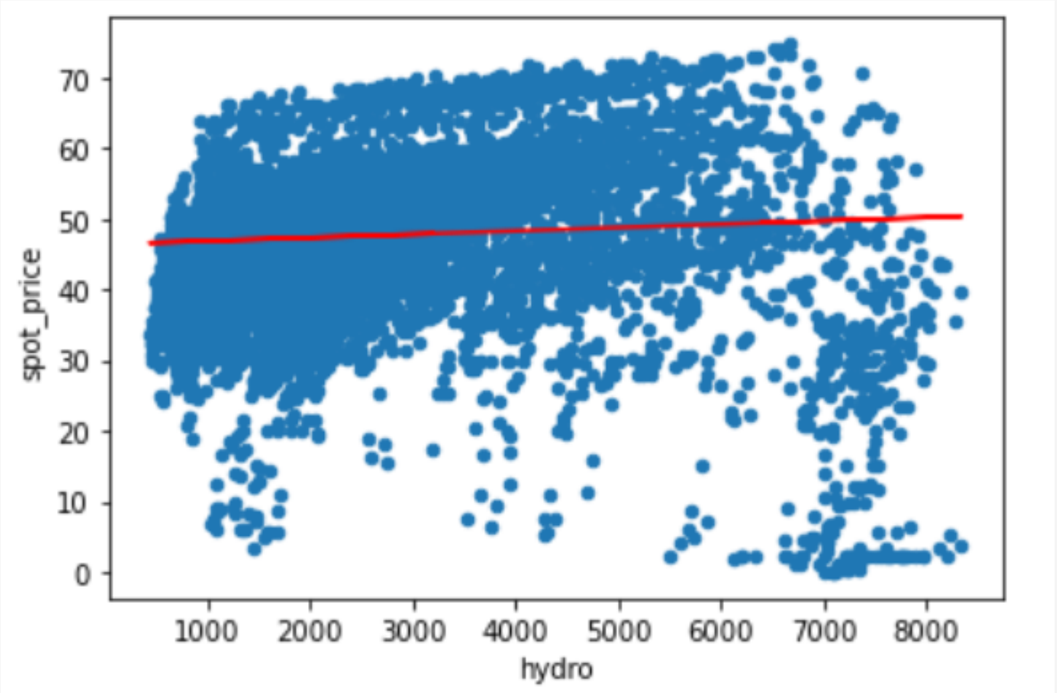
Correlation matrix of each type of energy production on spot price in the year 2019.



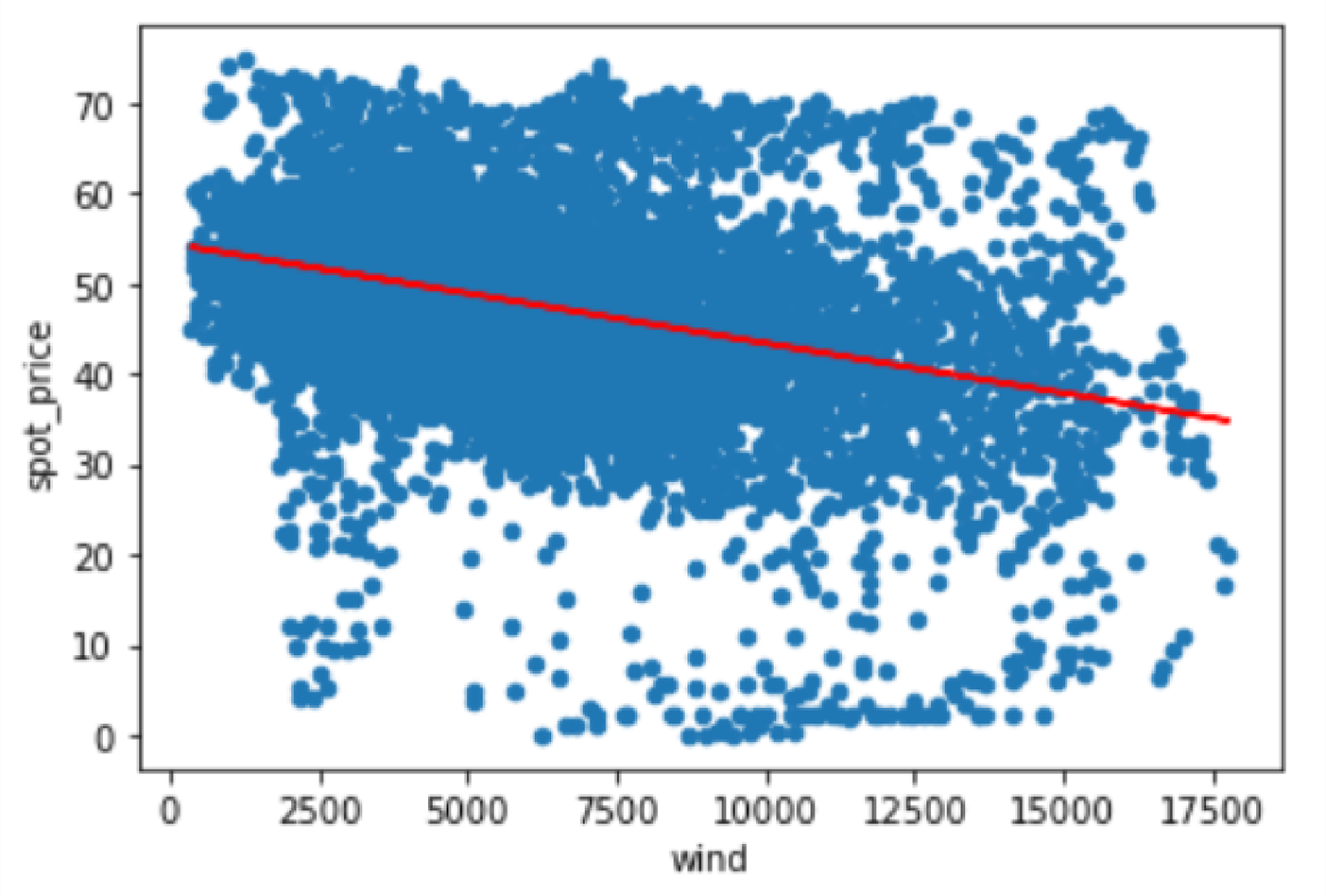
Effect of hydroelectric power generation on the price of electricity in 2019.

OLS Regression Results

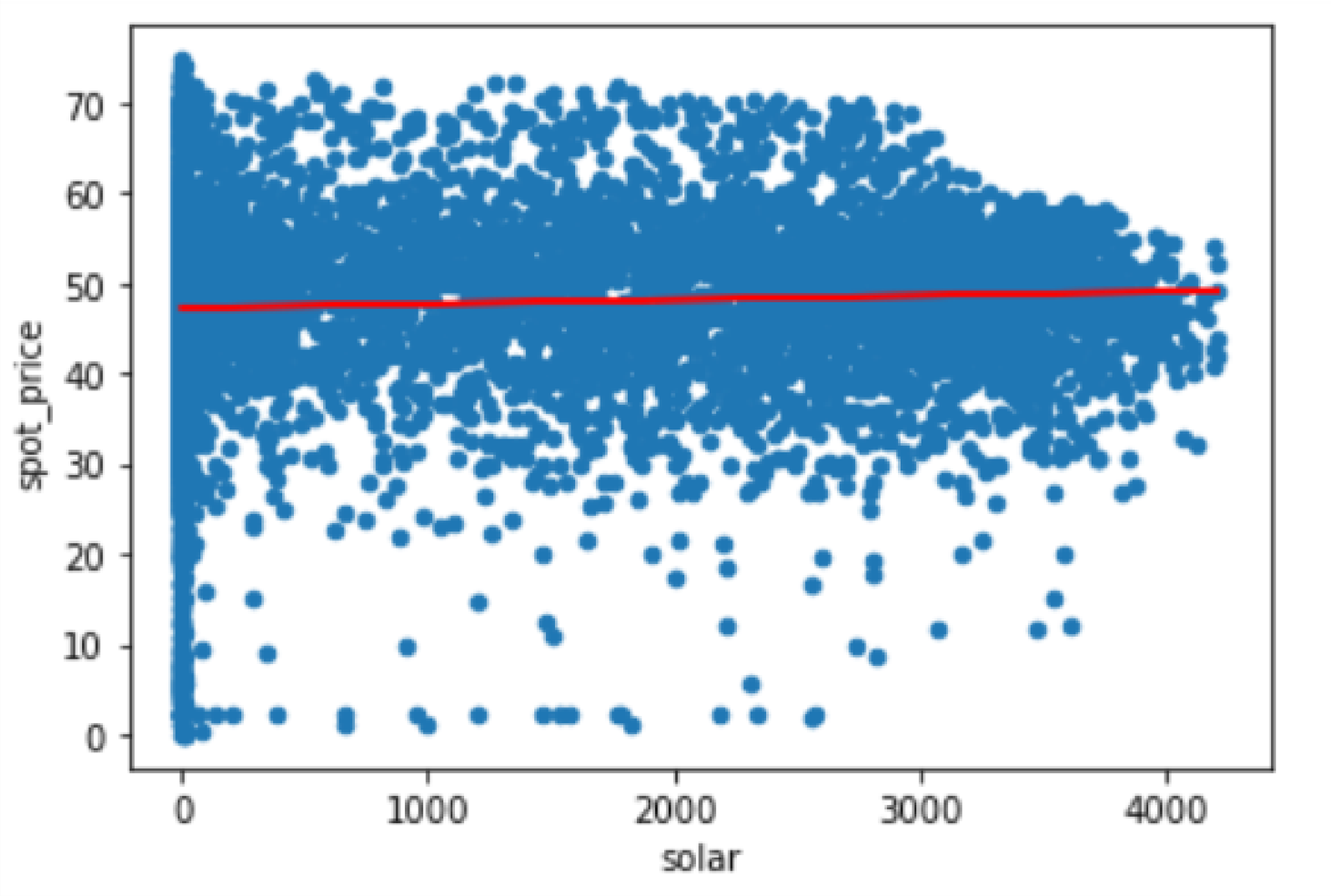
Dep. Variable:	hydro	R-squared:	0.005
Model:	OLS	Adj. R-squared:	0.005
Method:	Least Squares	F-statistic:	43.03
Date:	Tue, 06 Dec 2022	Prob (F-statistic):	5.68e-11
Time:	12:27:59	Log-Likelihood:	-76886.
No. Observations:	8737	AIC:	1.538e+05
Df Residuals:	8735	BIC:	1.538e+05
Df Model:	1		
Covariance Type:	nonrobust		



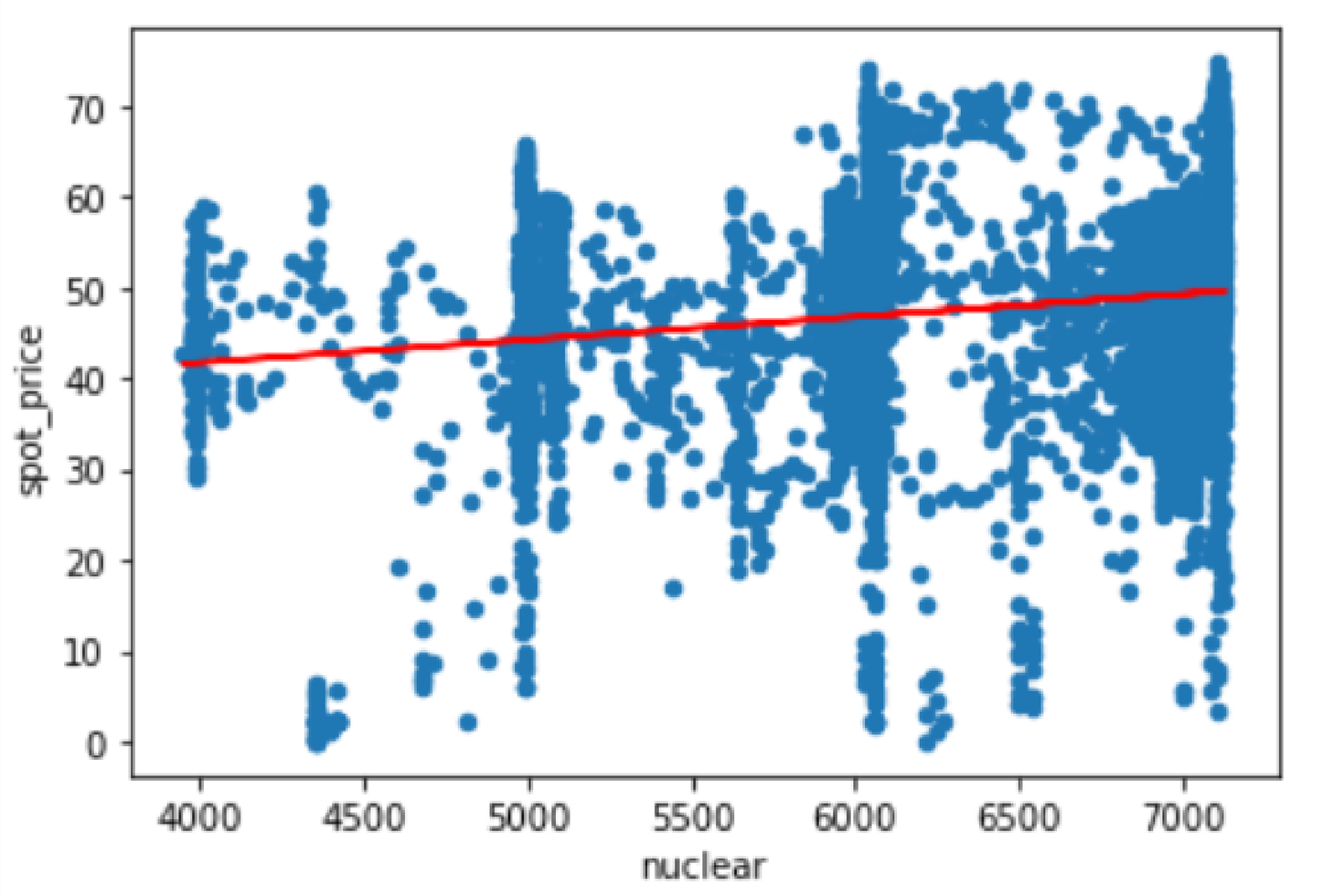
Effect of wind power generation on the price of electricity in 2019.



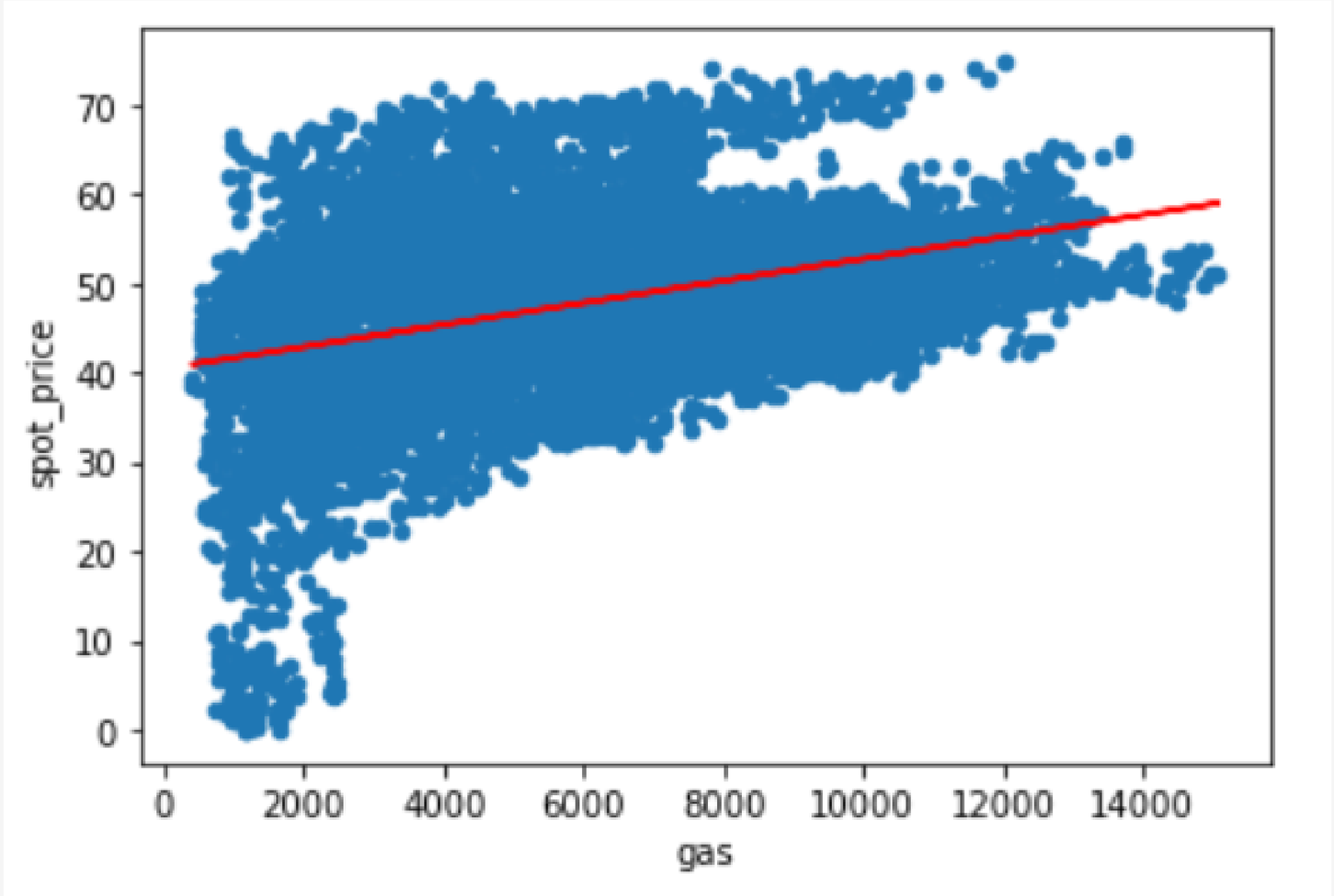
Effect of solar power generation on the price of electricity in 2019.



Effect of nuclear power generation on the price of electricity in 2019.



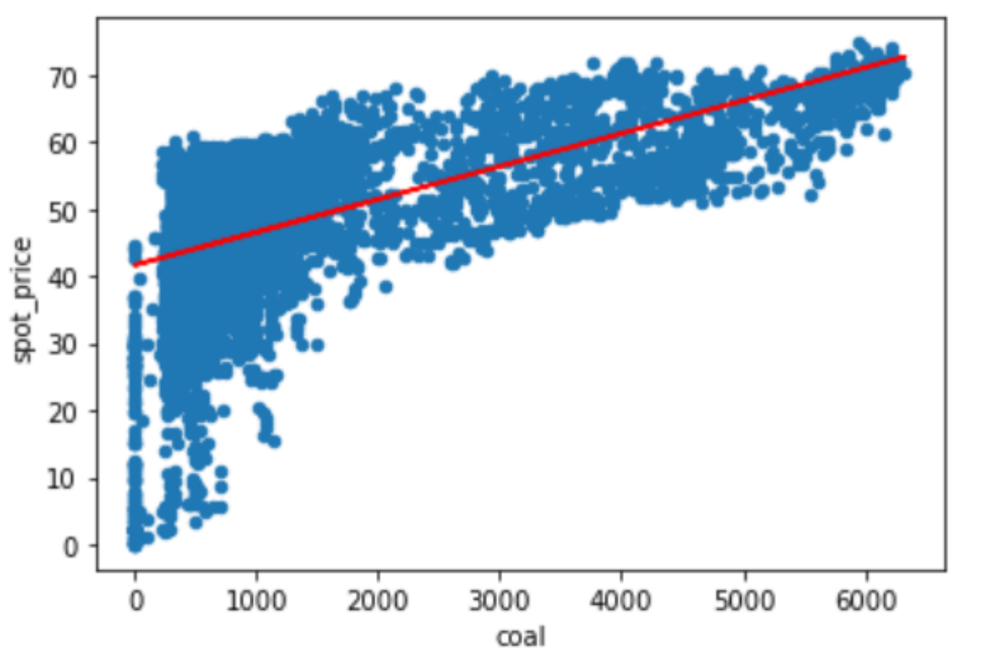
Effect of gas power generation on the price of electricity in 2019.



Effect of coal power generation on the price of electricity in 2019.

OLS Regression Results

Dep. Variable:	coal	R-squared:	0.350
Model:	OLS	Adj. R-squared:	0.350
Method:	Least Squares	F-statistic:	4699.
Date:	Tue, 06 Dec 2022	Prob (F-statistic):	0.00
Time:	12:28:20	Log-Likelihood:	-73235.
No. Observations:	8737	AIC:	1.465e+05
Df Residuals:	8735	BIC:	1.465e+05
Df Model:	1		
Covariance Type:	nonrobust		



Expense difference in average between weekdays and weekends.

weekend	spot_price
False	49.171389
True	44.041915

```
Difference between weekdays and weekends:
spot_price    5.129474
```

```python
weekend_weekday["weekend"] = np.where(
    weekend_weekday["weekday"].isin([0, 1, 2, 3, 4]),
    "False",
    "True"
)
```

Thanks for your attention

GROUP 5

LIA DOLLISON
LUDOVICO GANDOLFI
LAKSHMI PRIYANKA JAKKA
HANG CHI KU
IVAN LOPEZ CANSADO
IGNACIO PIRE RUBIO
ROMAN ZOTKIN