Seasonality

Introduction

This report investigates whether it is feasible to make accurate predictions of the energy prices. The goal of this report is not to actually produce such predictions, but merely to demonstrate, using statistical techniques whether it is possible. Specifically, hourly prices on electricity in the period 2022-10-26 to 2025-04-09 are used to test whether there is *seasonality* at the hourly, monthly and day of week level. If these vary systematically, then there is room to build predictions based on this (using regressions analysis introduced later in this course).

Descriptive statistics

Table 1: Summary statistics

|  |  |
| --- | --- |
| **Standard summary stats** | |
| Mean | 0.086967 |
| Standard dev. | 0.069661 |
| Observations | 21504 |

Notes: This table contains standard summary statistics

Table 2: Summary statistics by year

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 2025 | 2024 | 2023 | 2022 |
| Mean | 0.098766 | 0.070863 | 0.081255 | 0.191502 |
| Standard dev. | 0.055489 | 0.052209 | 0.050119 | 0.139442 |
| Obsevations | 2351 | 8784 | 8760 | 1585 |

Notes: This table contains standard summary statistics by year

Figure 1: CI on energy prices

Notes: This figure shows the 95% confidence interval using the normal distribution around the mean for each year

Interpretation (add more explanation):

* The average energy price is 0.087 EUR per kWh with a standard deviation of 0.07 (Table 1)
* The standard deviation is relatively stable over time except for in 2022, which is also a year, with higher energy prices than the other years (Table 2)
* A 95% confidence interval around the average prices shows no overlap between any of the years, although the year 2022 also displays visibly higher prices (Figure 1)
* A concern with the data for 2022 and 2025 is that we have fewer observations. If there are systematic differences across the month of the year, then this skew our results

Statistical testing

We set up three null hypotheses to test one by one, which probes for different types of seasonality:

1. Day of week

1. Month of the year

1. Hour of the day

We test each of these using ANOVA

Table 3: ANOVA by day of week

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 2,607956 | 6 | 0,434659 | 91,84231 | 2,6E-114 | 2,099017 |
| Within Groups | 101,7382 | 21497 | 0,004733 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 104,3461 | 21503 |  |  |  |  |

Table 4: ANOVA by month of the year

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 9,618658 | 11 | 0,874423 | 198,3913 | 0 | 1,789093 |
| Within Groups | 94,72749 | 21492 | 0,004408 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 104,3461 | 21503 |  |  |  |  |

Table 5: ANOVA by hour of the day

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *Source of Variation* | *SS* | *df* | *MS* | *F* | *P-value* | *F crit* |
| Between Groups | 7,127623 | 23 | 0,309897 | 68,47028 | 0 | 1,529742 |
| Within Groups | 97,21853 | 21480 | 0,004526 |  |  |  |
|  |  |  |  |  |  |  |
| Total | 104,3461 | 21503 |  |  |  |  |