Digital Transformation KIPA Project

Discovering Waste Price determinants

Members

Aiman Alam

Husnain Dawood

Sergey Polyarus

Suraj Karakulath

Modou Lamin Manjang

Niloufar Shokri

Overview

We want to find correlations/patterns between the price of waste (bought from suppliers) and various potential price determinants such as weather, energy and business cycle.

The wPreis dataset

The wPreis dataset

- Price of waste (from client) from around Sep 2020 until Sep 2023
- Could be negative (client got paid for taking the waste)
- Could be positive (client paid money to get the trash)
- 10 unique clusters (collection of Postleitzahl) - e.g. [‘25’, ‘24’]
- 4 unique product categories
  - 'A2 - geschreddert',
  - 'A1 & A2 - geschreddert'
  - 'A2 & A3 - geschreddert'
  - 'A3 - geschreddert'
- No null values

	week	wPreis	Plz	full	date
0	2020-37	-37.519318	['25', '24']	A2 - geschreddert	2020-09-14
1	2020-39	-40.000000	['25', '24']	A2 - geschreddert	2020-09-28
2	2020-41	-34.853543	['25', '24']	A2 - geschreddert	2020-10-12
3	2020-42	-27.560202	['25', '24']	A2 - geschreddert	2020-10-19
4	2020-43	-31.937780	['25', '24']	A2 - geschreddert	2020-10-26

- The wPreis for all 3 categories in the same cluster are somewhat correlated, with occasional deviations.
- This is observed in all clusters more or less.
- For some clusters, there is only one waste category, *A1 & A2* for [1, 4, 6, 7, 8, 9]

Exploring Potential Price Determinants

Weather - First Approach (Nilou)

Data extracted from **open meteo free API documentation.**

The parameters used to access the data (using the latitude and longitude of the cities in each cluster) :

- **temperature_2m:** Air temperature at 2 meters above ground in celsius.
- **windspeed_10m:** Wind speed at 10 meters above ground (the standard level) in kmh.
- **precipitation:** Total precipitation (rain, showers, snow) sum of the preceding hour. Data is stored with a 0.1 mm precision.
- **rain:** Only liquid precipitation of the preceding hour including local showers and rain from large scale systems in mm.
- **snowfall:** amount of the preceding hour in centimeters. For the water equivalent in millimeter, divide by 7. E.g. 7 cm snow = 10 mm precipitation water equivalent.

Weather - First Approach (Nilou)

Here is a screenshot of the dataset. Data in the main dataset was in hourly basis but converted into weekly basis by taking the average of each seven days. The time format was **string** but transformed to **daytime** object.

	time	rain	precipitation	temperature_2m	snowfall	windspeed_10m
0	2020-08-31	0.108889	0.108889	15.086667	0.0	12.055972
1	2020-09-07	0.014940	0.014940	15.696012	0.0	13.736131
2	2020-09-14	0.000000	0.000000	15.897143	0.0	9.828214
3	2020-09-21	0.097321	0.097321	14.293631	0.0	10.337440
4	2020-09-28	0.021548	0.021548	13.913393	0.0	13.442738

Weather - Second Approach (Husnain)

- Second Deutscher Wetterdienst or DWD for short, is the German Meteorological Service. They provide Weather data dating back to 1830s to as recent as yesterday. The stations currently in the Weather dataset consists of 19631 total stations. There are 438 parameters in the dataset. The list can be found here https://wetterdienst.readthedocs.io/en/latest/data/parameters.html

For now, we have accessed temperature mean and max just to play around with the dataset and queried specific servers to confirm its accessibility.

Bremen stations

index	statio...	from_date	to_date	height	latitude	longitude	name	state
610	00689	1951-01-01T00:00:0...	1998-12-31T00:00:00.000Z	3	53.1167	8.8	Bremen (Bayernstraße)	Bremen
611	00690	1941-01-01T00:00:0...	2023-08-29T00:00:00.000Z	2	53.0958	8.827	Bremen (Bürgerpark)	Bremen
612	00691	1890-01-01T00:00:0...	2023-09-11T00:00:00.000Z	4	53.0451	8.7981	Bremen	Bremen
613	00692	1951-01-01T00:00:0...	1999-12-31T00:00:00.000Z	1	53.1402	8.739	Bremen (Ritterhuder Heerstraße)	Bremen
614	00693	1951-01-01T00:00:0...	2001-03-31T00:00:00.000Z	3	53.0716	8.774	Bremen (Warturmer Heerstraße)	Bremen
615	00694	1891-09-01T00:00:0...	1992-12-31T00:00:00.000Z	8	53.1833	8.5833	Bremen-Blumenthal	Bremen
616	00695	1931-01-01T00:00:0...	2006-12-31T00:00:00.000Z	4	53.2068	8.5093	Bremen-Farge	Bremen
617	00696	1941-01-01T00:00:0...	1965-12-30T00:00:00.000Z	4	53.1	8.7333	Bremen-Lankenau	Bremen
618	00697	1971-01-01T00:00:0...	1996-12-31T00:00:00.000Z	2	53.1333	8.7167	Bremen-Mittelsbüren	Bremen
619	00698	1951-01-01T00:00:0...	1998-06-30T00:00:00.000Z	5	53.0602	8.9533	Bremen-Osterholz	Bremen
620	00699	1949-04-01T00:00:0...	1958-06-29T00:00:00.000Z	4	53.1004	8.7834	Bremen-Seefahrtschule	Bremen
621	00700	1951-01-01T00:00:0...	1989-12-31T00:00:00.000Z	2	53.0833	8.7	Bremen-Strom	Bremen
622	00701	1949-01-01T00:00:0...	2023-09-12T00:00:00.000Z	7	53.5332	8.5761	Bremerhaven	Bremen
623	00702	1931-01-01T00:00:0...	1960-06-30T00:00:00.000Z	8	53.5876	8.6015	Bremerhaven-Speckenbüttel	Bremen

Weather - Temperature and Wind (A1 & A2 Category)

Cluster: [20,21,22,23]

Cluster: [40,41,42,43,44,45,46,47]

Weather - Temperature and Wind (A2 & A3 Category)

Cluster: [24,25]

Cluster: [40,41,42,43,44,45,46,47]

Weather - Temperature and Wind (A3 Category)

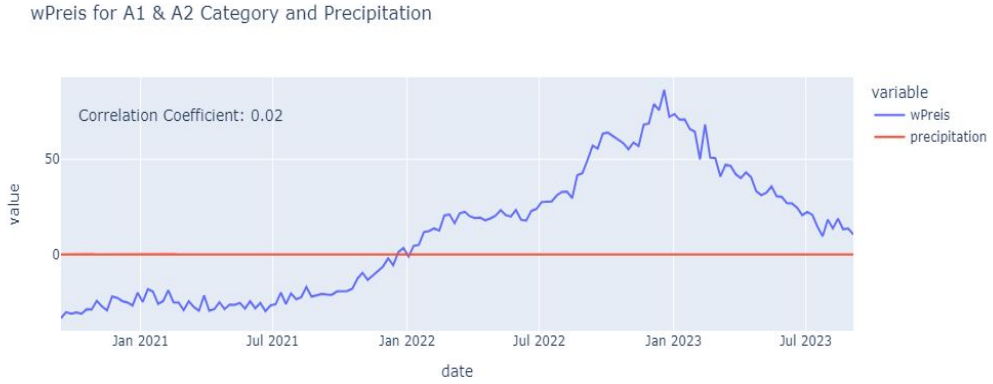
Cluster: [48,49]

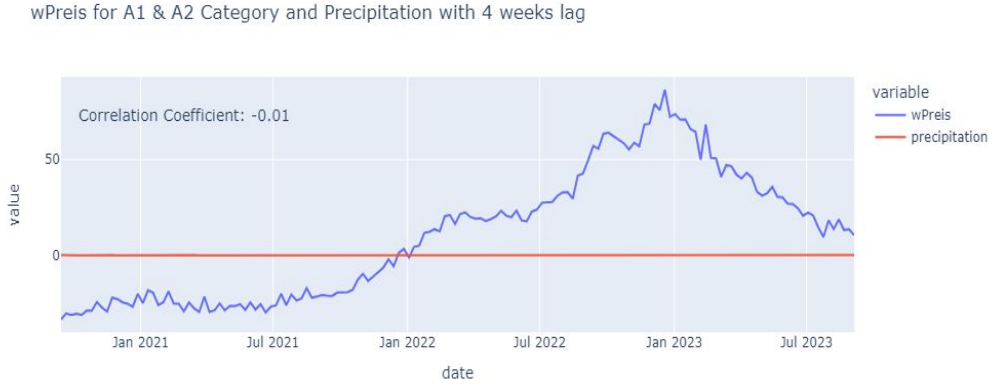
Cluster: [40,41,42,43,44,45,46,47]

Weather - Temperature and Wind (A2 Category)

Cluster: [26,27,28]

Weather - Precipitation

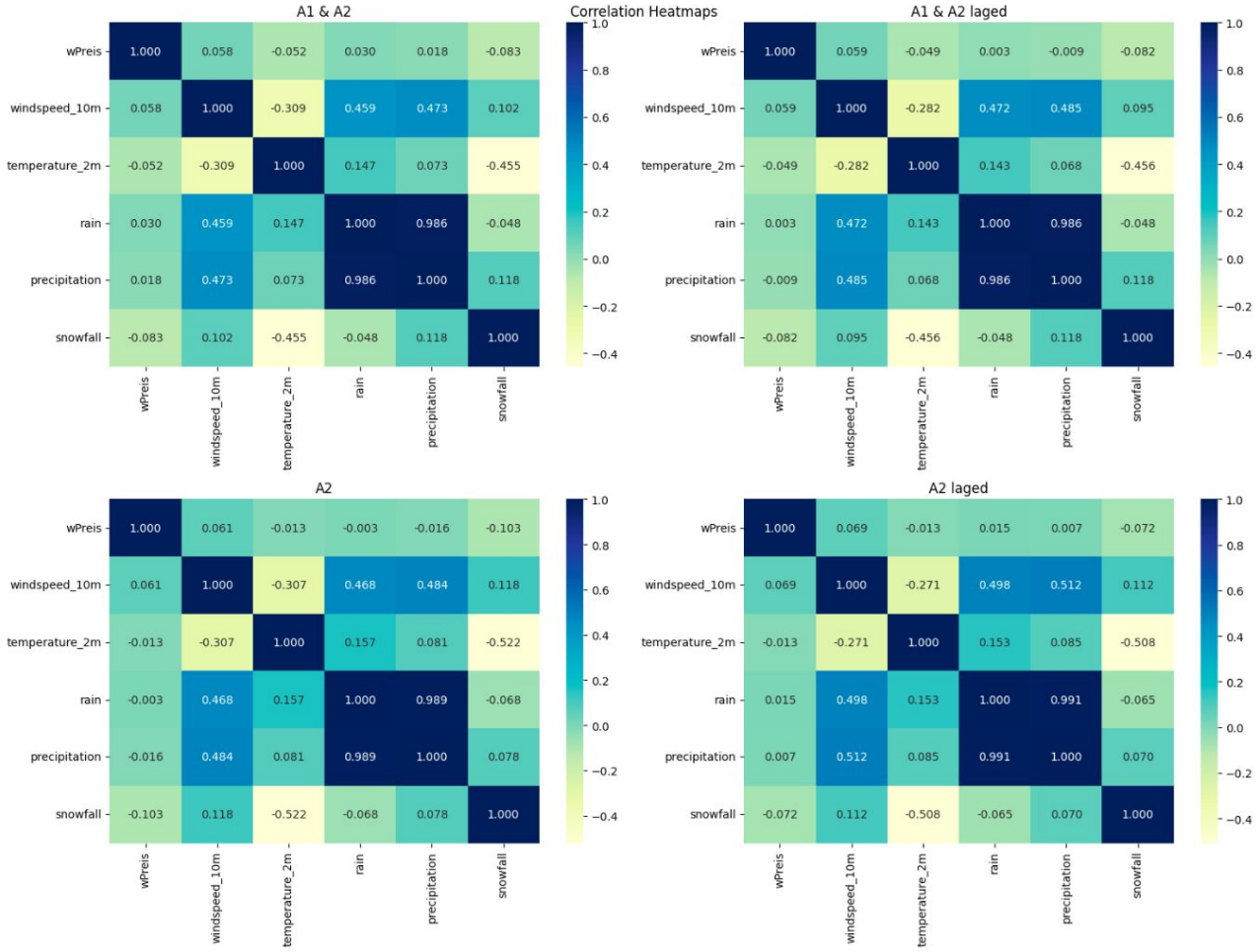




Weather - Correlation Heatmaps

The coefficients of wPreis with all other weather variables are too small.

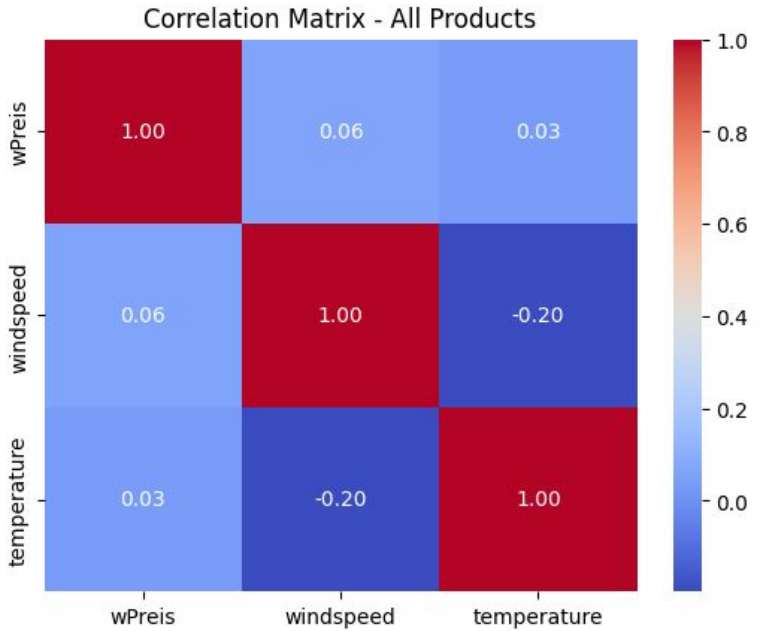
Hence this may not be a good determinant.



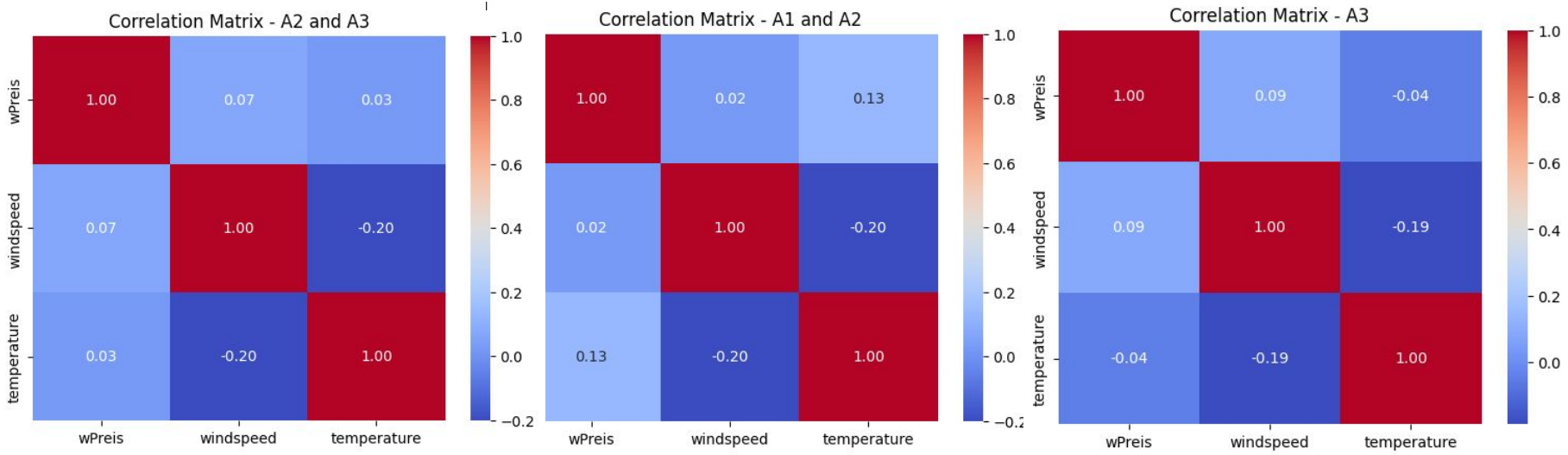
Weather - Heatmap

For the Cluster ['40', '41', '42', '44', '45', '46', '47']

- Correlation between wPreis and windspeed: 0.06
- Correlation between wPreis and temperature: 0.03



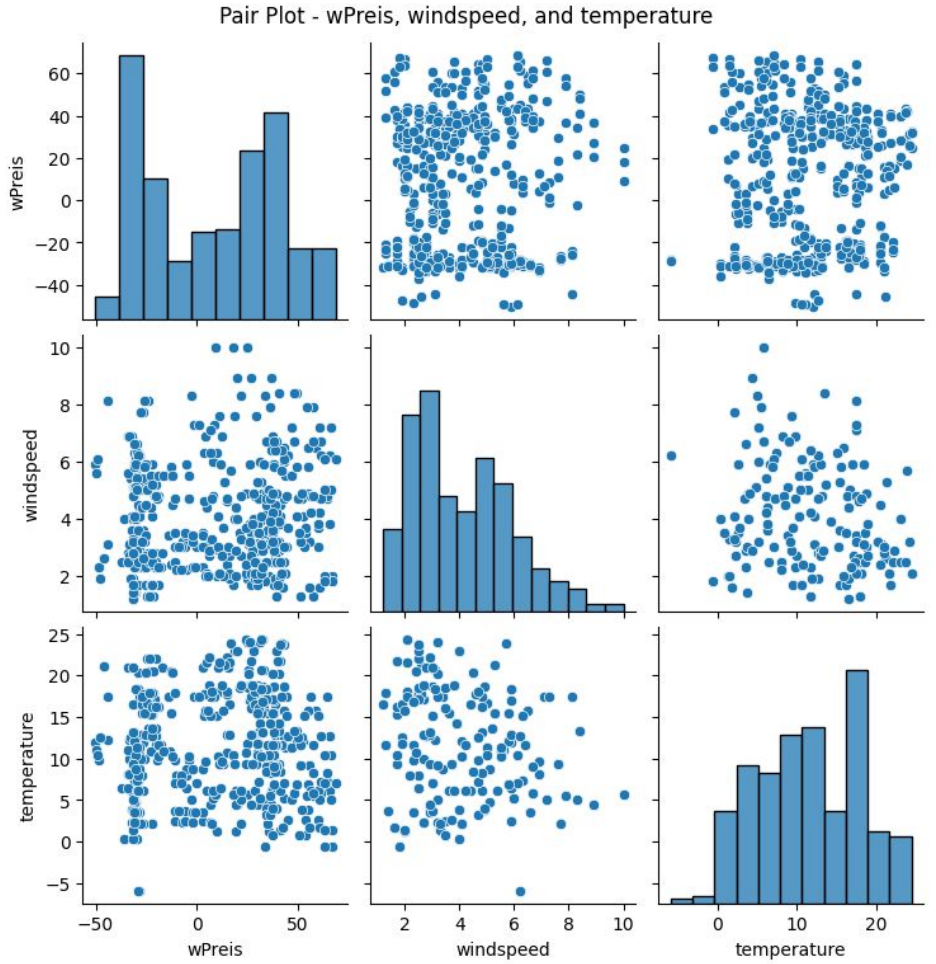
Heatmaps for individual products



Weather

For the Cluster ['40', '41', '42', '44', '45', '46', '47']

- Correlation between wPreis and windspeed: 0.06
- Correlation between wPreis and temperature: 0.03

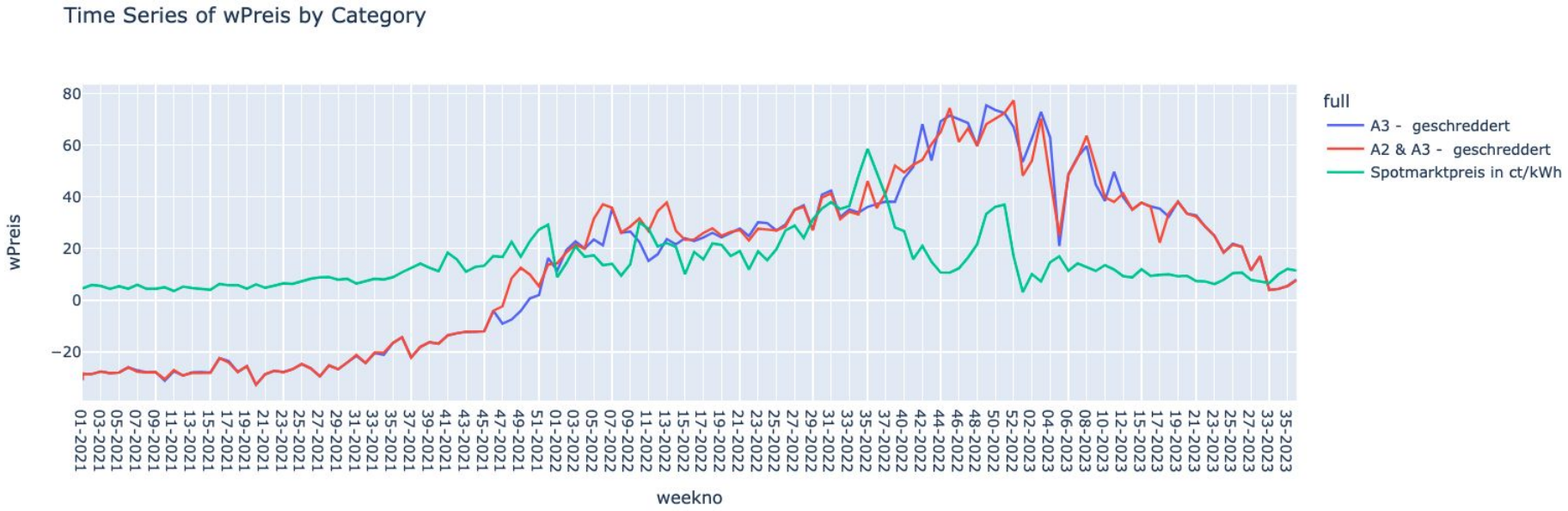


Energy - Electricity

- https://www.netztransparenz.de/EEG/Marktpraemie/Spotmarktpreis is the best available data (ct/kWh) but this is for the whole of Germany and only from Jan 2021

Energy - Electricity

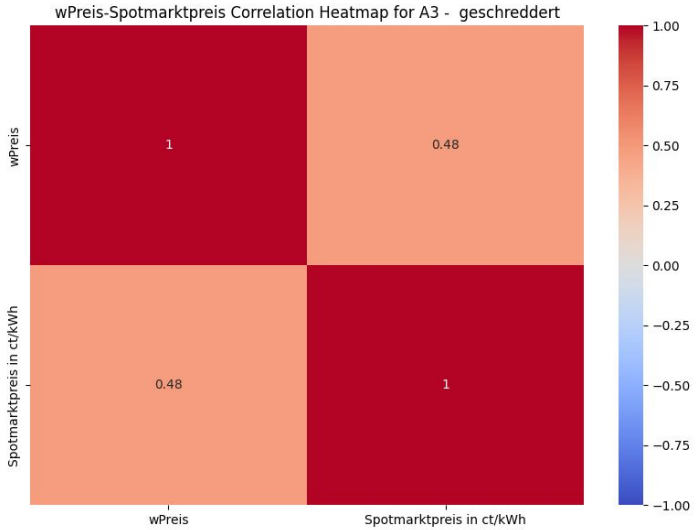
Time series for ['48', '49'] cluster

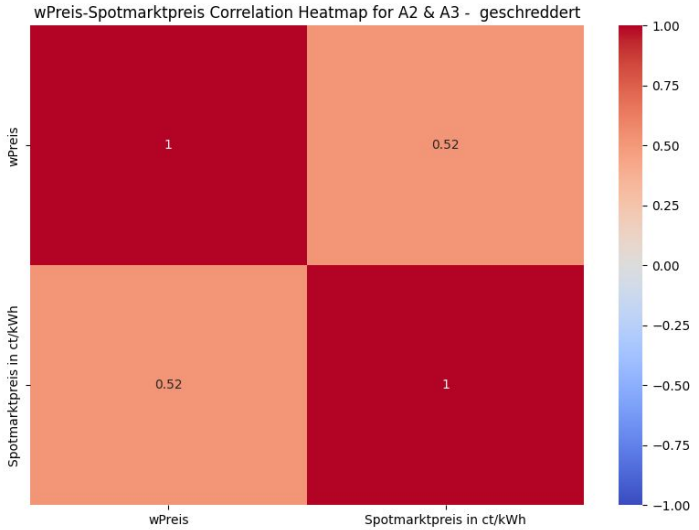


Energy - Electricity

Time series for ['48', '49'] cluster

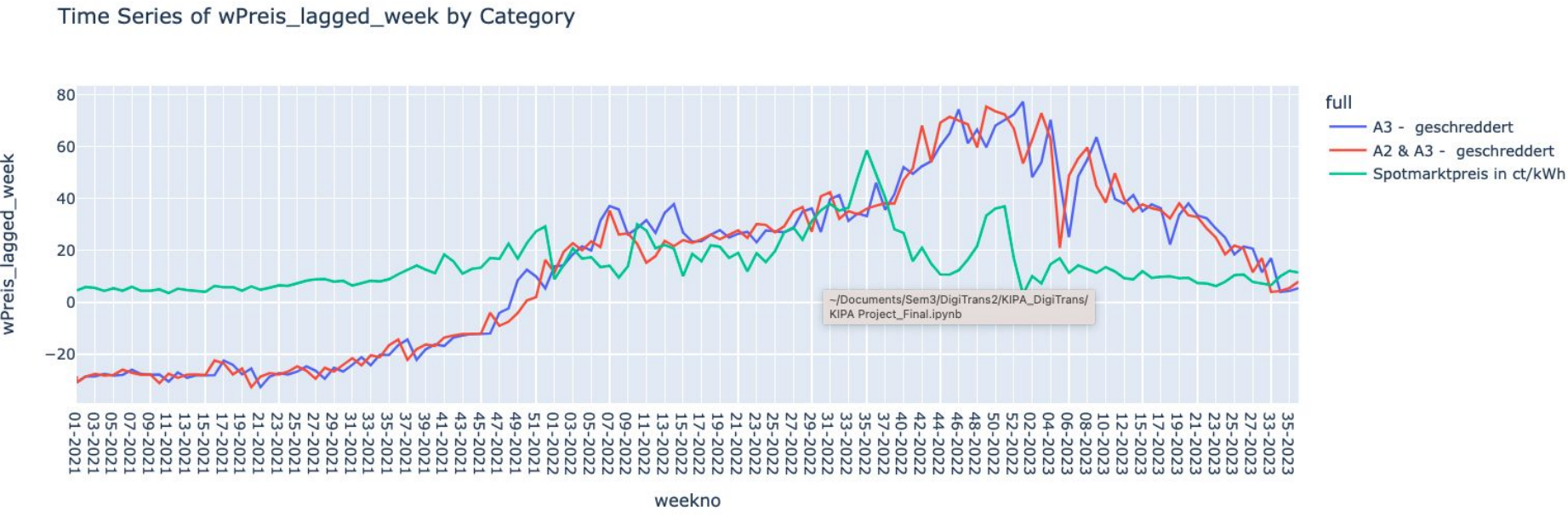
- Although visually it looks like there was some correlation in some parts, the coefficients are not signfiicant
- A2 & A3 category alone has 0.52 though





Energy - Electricity

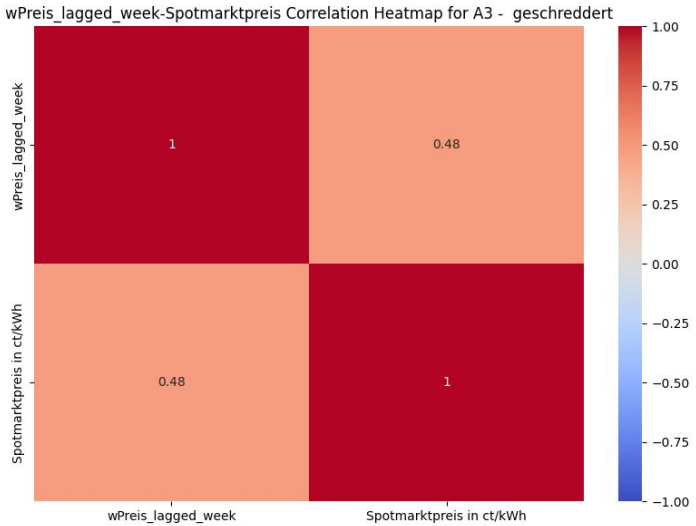
Time series for ['48', '49'] cluster with 1 week lag

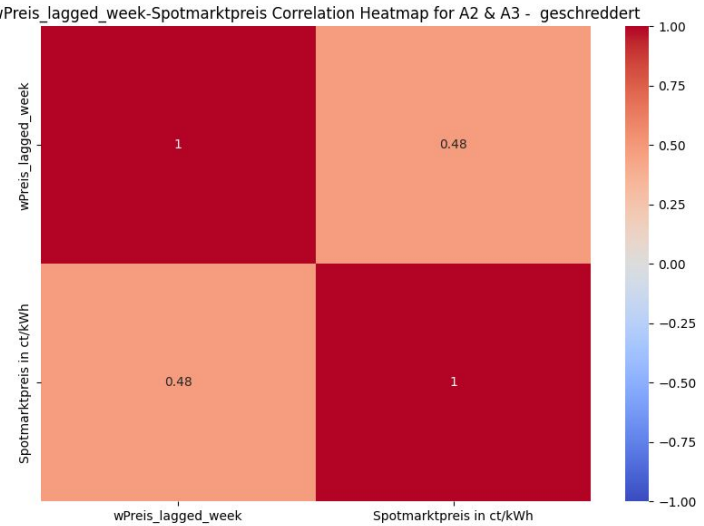


Energy - Electricity

Time series for ['48', '49'] cluster with 1-week lag

- 1-week lag also does not improve coefficients much





Energy - Electricity

- Exploring correlation with lags by up to 10 weeks also don't show much improvement

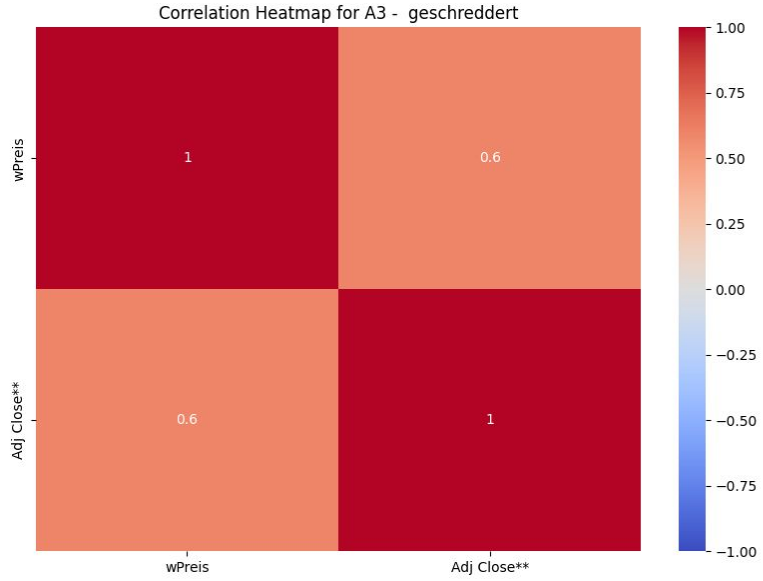
	Category	Lag	Correlation
0	A3 - geschreddert	2	0.432391
1	A3 - geschreddert	3	0.416149
2	A3 - geschreddert	4	0.401785
3	A3 - geschreddert	5	0.385725
4	A3 - geschreddert	6	0.359828
5	A3 - geschreddert	7	0.333825
6	A3 - geschreddert	8	0.301300
7	A3 - geschreddert	9	0.270280
8	A3 - geschreddert	10	0.241397
9	A2 & A3 - geschreddert	2	0.450800
10	A2 & A3 - geschreddert	3	0.426034
11	A2 & A3 - geschreddert	4	0.417254
12	A2 & A3 - geschreddert	5	0.395097
13	A2 & A3 - geschreddert	6	0.372961
14	A2 & A3 - geschreddert	7	0.343700
15	A2 & A3 - geschreddert	8	0.308703
16	A2 & A3 - geschreddert	9	0.273671
17	A2 & A3 - geschreddert	10	0.248491

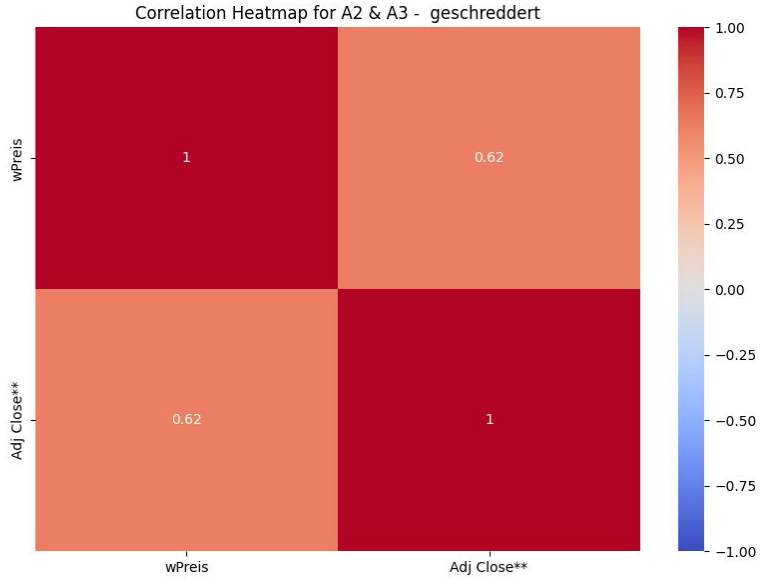
Energy - Oil

- No data on oil prices per region in Germany (may not make sense either)
- So we look at Global Oil and Gas Market Prices as a proxy from Yahoo finance
- And assume that relative changes in global prices have a proportional impact on the same prices in Germany
- **Adjusted Close Price (Adj Close)** - accounts for events such as stock splits and dividend payments.

- Close correlation from Sep 2020 to around Jan 2022 for both categories
- For A3 - geschreddert, corr. roughly continues until Jun 2022, then inverse corr. until Dec 2022 and then low corr. until Sep 2023
- For A2 & A3 - geschreddert, it continues from Jan 2022 to around Mar 2022, then inverse corr. until Dec 2022, and similar to the other category, loses the correlation after that.
- Something remained consistent until Jan 2022. After that things began going in the opposite direction. And from Dec 2022 onwards they become less correlated.

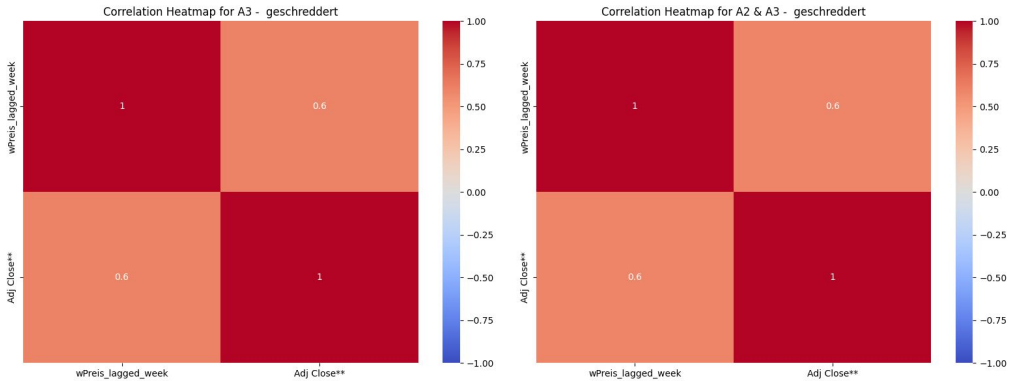
- Correlation coefficients also give us a measure of how correlated they are.
- For this cluster, they are 0.6 for A3 - geschreddert category and 0.62 for A2 & A3 - geschreddert which are above 0.5 and hence significant.





Correlation with lag

- We explore correlation with a lag of 1 week to see if that is a better predictor.
- But 1-week lag coefficients don't improve much
- Neither does coefficients for lags of 2 to 10 weeks

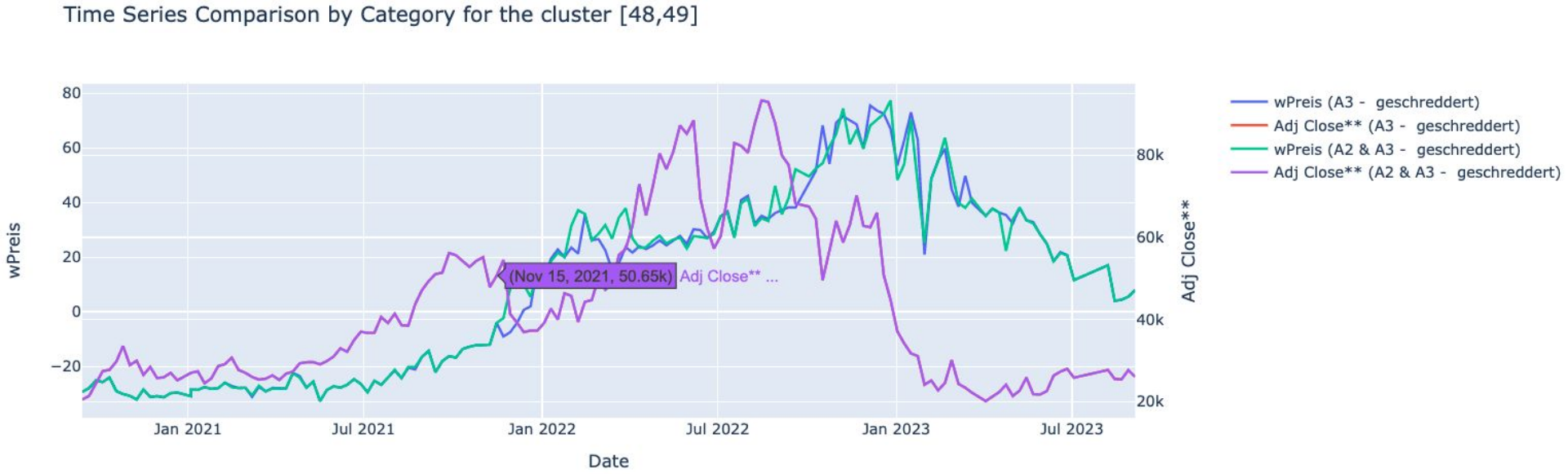



```
Correlation for A3 -  geschreddert: 0.60
Correlation for A2 & A3 -  geschreddert: 0.60
```

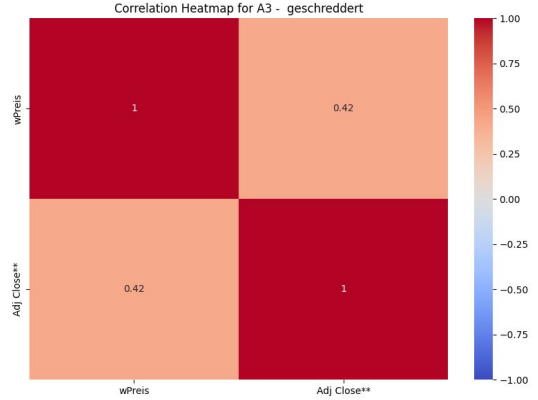
	Category	Lag	Correlation
0	A3 - geschreddert	2	0.570127
1	A3 - geschreddert	3	0.551046
2	A3 - geschreddert	4	0.529585
3	A3 - geschreddert	5	0.514884
4	A3 - geschreddert	6	0.503870
5	A3 - geschreddert	7	0.480222
6	A3 - geschreddert	8	0.454619
7	A3 - geschreddert	9	0.439824
8	A3 - geschreddert	10	0.420328
9	A2 & A3 - geschreddert	2	0.590616
10	A2 & A3 - geschreddert	3	0.577857
11	A2 & A3 - geschreddert	4	0.564648
12	A2 & A3 - geschreddert	5	0.552081
13	A2 & A3 - geschreddert	6	0.544712
14	A2 & A3 - geschreddert	7	0.524520
15	A2 & A3 - geschreddert	8	0.501037
16	A2 & A3 - geschreddert	9	0.487971
17	A2 & A3 - geschreddert	10	0.467319

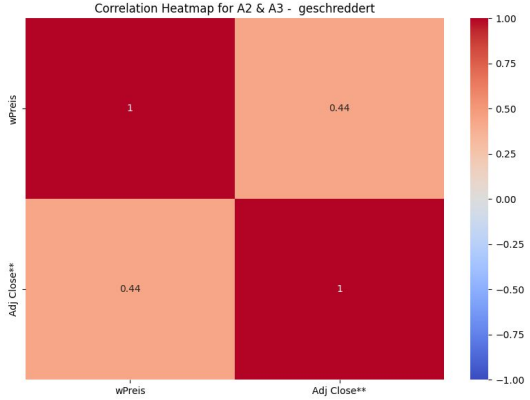
Energy - Gas

- Similar to oil, gas prices been collected from Yahoo, for the whole of Germany.
- And we explore correlation **Adjusted Close Price (Adj Close)** with wPreis



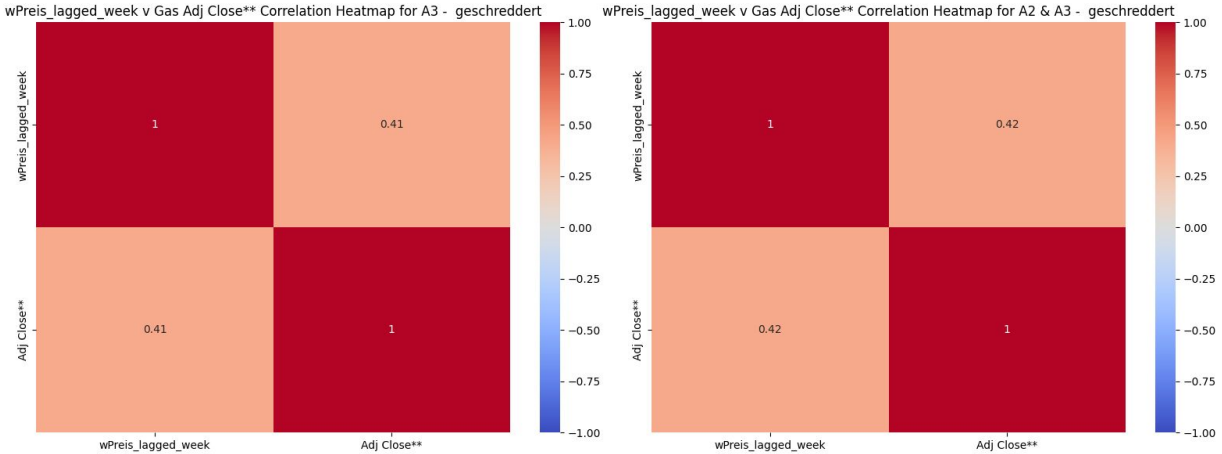
- Unlike oil, for gas, the correlation is not that much.
- Best we can say is that the general trend is rising for both from Jan 2021 until around Dec 2022 and after that the wPreis continues to rise while gas prices rise sharply and fall sharply.
- This is validated by lower coefficients (around 0.4)





Correlation with lag

- Exploring lags by 1 to 10 weeks also don't improve coefficients



wPreis_lagged_week v Gas Adj Close** Correlation for A3 - geschreddert: 0.41

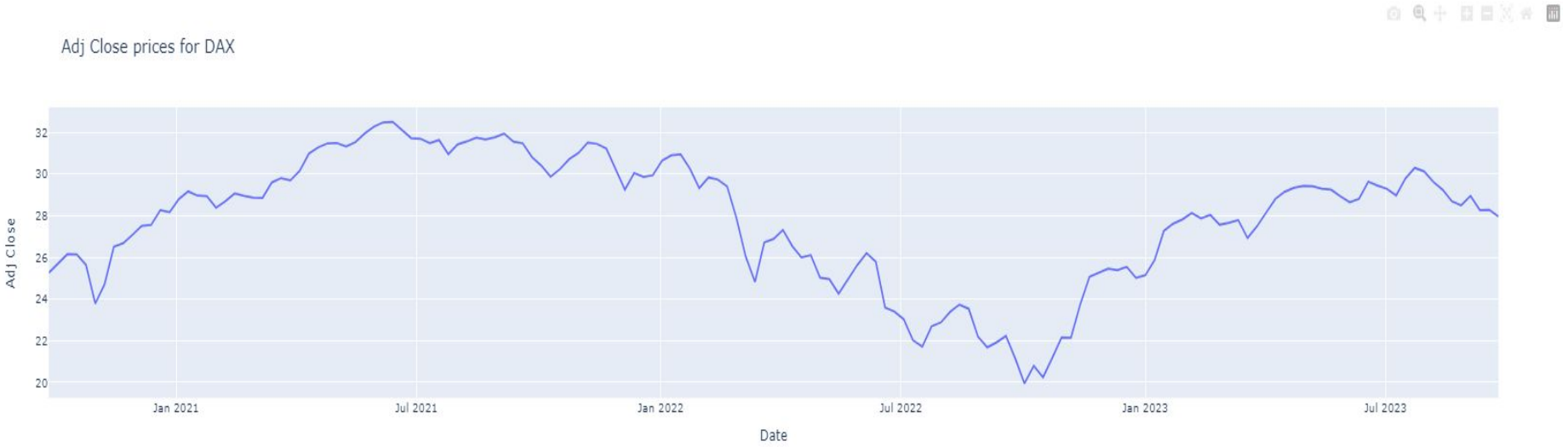
wPreis_lagged_week v Gas Adj Close** Correlation for A2 & A3 - geschreddert: 0.42

	Category	Lag	Correlation
0	A3 - geschreddert	2	0.570127
1	A3 - geschreddert	3	0.551046
2	A3 - geschreddert	4	0.529585
3	A3 - geschreddert	5	0.514884
4	A3 - geschreddert	6	0.503870
5	A3 - geschreddert	7	0.480222
6	A3 - geschreddert	8	0.454619
7	A3 - geschreddert	9	0.439824
8	A3 - geschreddert	10	0.420328
9	A2 & A3 - geschreddert	2	0.590616
10	A2 & A3 - geschreddert	3	0.577857
11	A2 & A3 - geschreddert	4	0.564648
12	A2 & A3 - geschreddert	5	0.552081
13	A2 & A3 - geschreddert	6	0.544712
14	A2 & A3 - geschreddert	7	0.524520
15	A2 & A3 - geschreddert	8	0.501037
16	A2 & A3 - geschreddert	9	0.487971
17	A2 & A3 - geschreddert	10	0.467319

Business Cycle - (DAX)

- We use the **DAX** - the **Deutscher Aktien Index** or the **GER40** : a stock index that represent 40 of the largest and most liquid German companies that trade on the Frankfurt Exchange. Source Yahoo finance.
- From the dataset, we extract the adjusted close price of the weekly average of the GER40 and then merge it with different clusters-category to visualize and calculate the correlation between the DAX and the wPries.

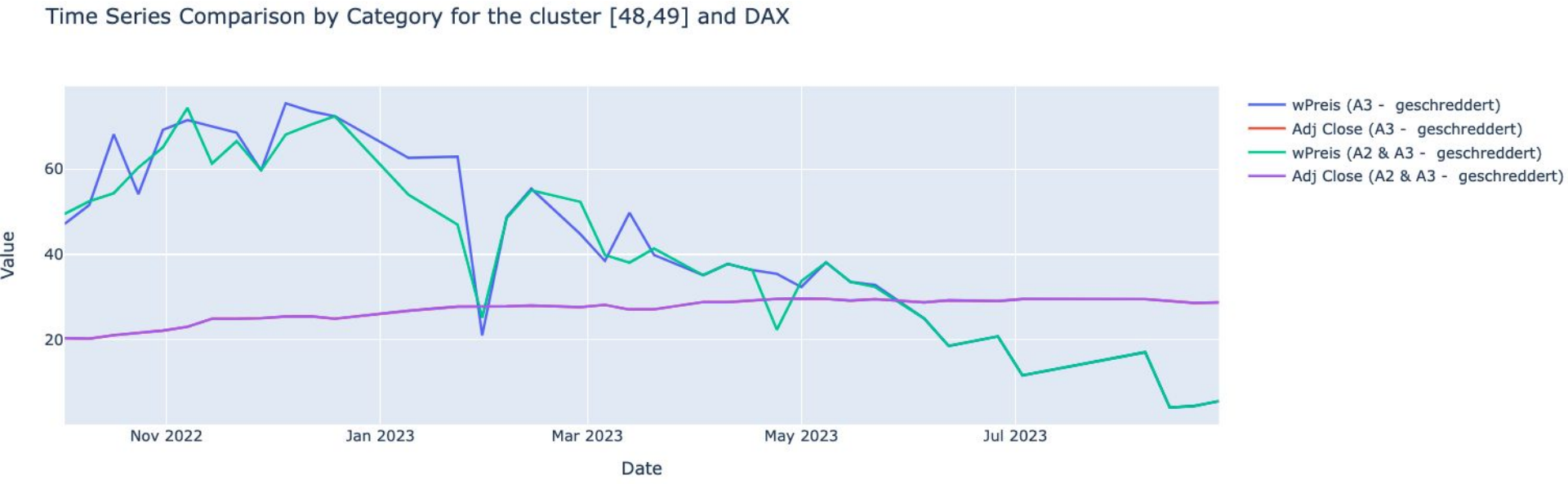
DAX Visual


Adj Close prices for DAX
Adj Close
32
30
28
26
24
22
20
Jan 2021
Jul 2021
Jan 2022
Jul 2022
Jan 2023
Jul 2023
Date

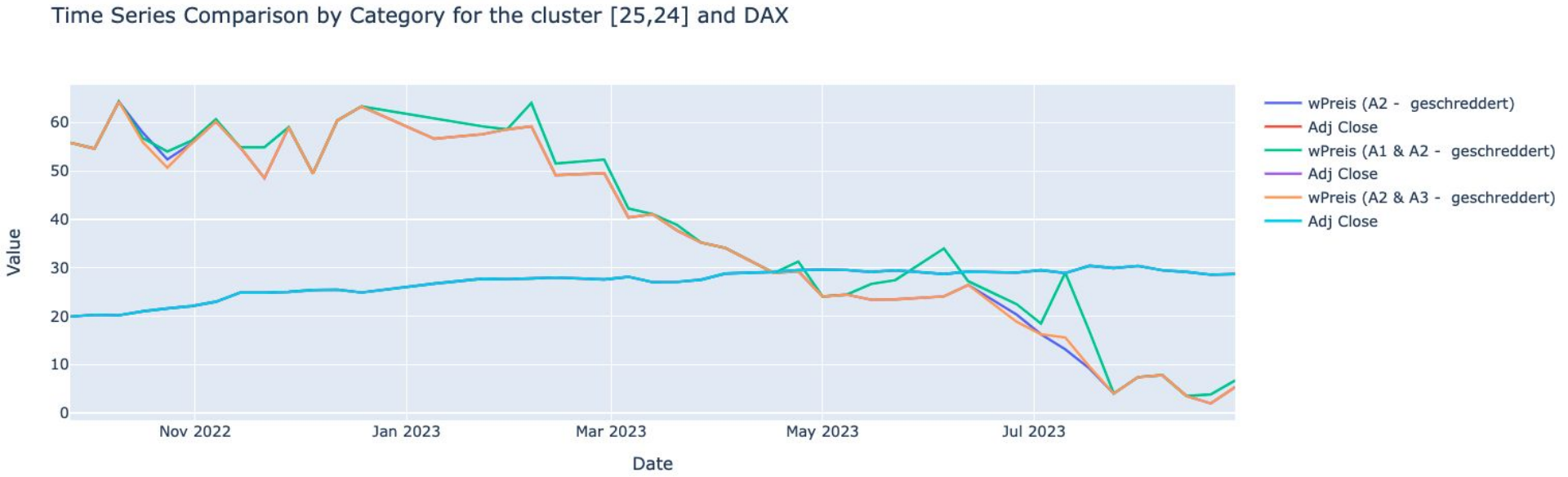
Business Cycle - (DAX)

Time Series Comparison by Category for the cluster [1, 4, 6, 7, 8, 9] andthe daily DAX

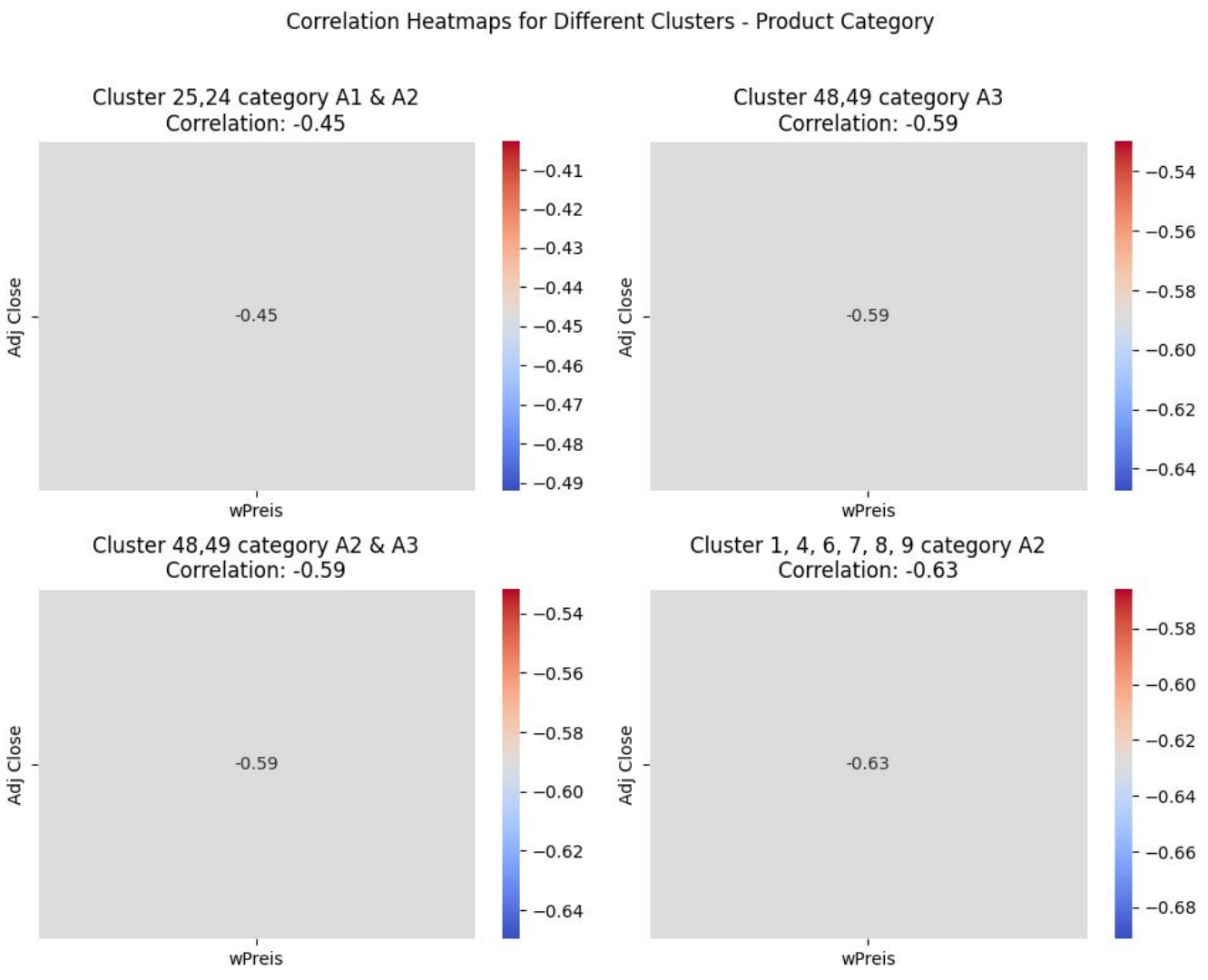
Business Cycle - (DAX)



Business Cycle - (DAX)

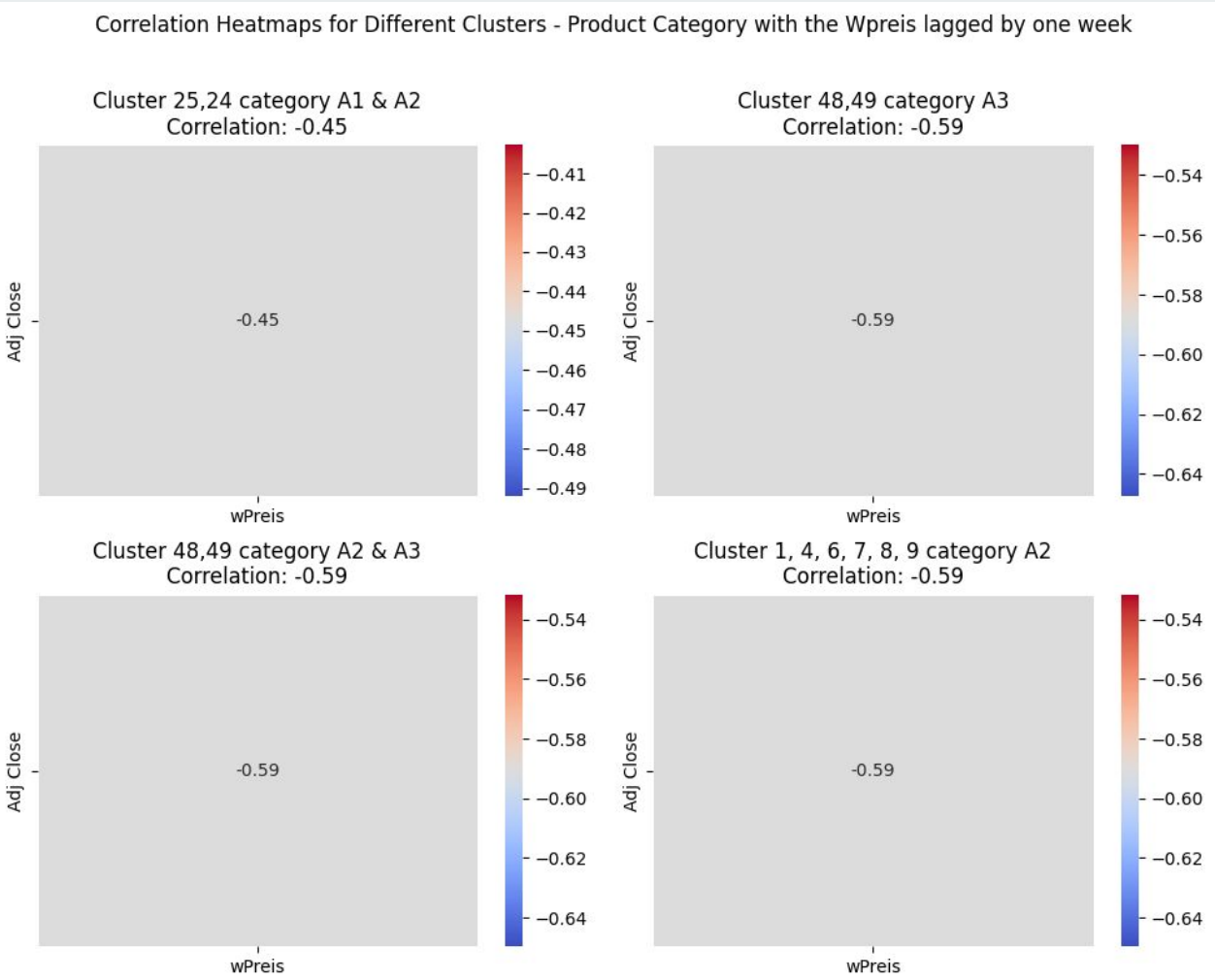


Correlation between DAX and specific Cluster-categories



Correlation of wPreis lagged by a week

There isn't any significant difference between the normal wPreis and the lagged wPries, they both range somewhat between **-0.45 to -0.63**



Construction

- Data on construction permits number (per land per month) - taken from Statistik der Baugenehmigungen (code 31111) placed on https://www-genesis.destatis.de/. This variable was taken because it was the only one available with monthly frequency and per land.
- **NB!** The PLZs are converted into lands despite not always matching the land exactly (some include two or more lands, some encompass not entire land). The accuracy may have been better if there waste prices were sorted by lands.

Construction

Waste Prices in ['Niedersachsen', 'Bremen']

Corr. coeff. between the A2 & A3 - geschreddert prices and the number of construction permits issues is -0.75
Corr. coeff. between the A3 - geschreddert prices and the number of construction permits issues is -0.73

Construction

Corr. coeff. between the A1 & A2 - geschreddert prices and the number of construction permits issues is -0.79
Corr. coeff. between the A2 & A3 - geschreddert prices and the number of construction permits issues is -0.78
Corr. coeff. between the A2 - geschreddert prices and the number of construction permits issues is -0.78

Construction

Corr. coeff. between the A1 & A2 - geschreddert prices and the number of construction permits issues is -0.82

Construction

- Overall insight is that the pattern of correlations is the same
- The correlation between actual waste prices and number of permits is large and negative.
- We assume it has something to do with the waste offer: usually the timespan between getting a construction permit and starting the construction itself is forced to be as short as possible => construction works start as soon as the permit is obtained => the waste is produced immediately => there is more waste offered on the market => the price decreases.

Conclusions and Recommendations

Conclusions

- Weather determinants show very low correlation with wPreis (which makes sense somewhat)
- Energy determinants show different variations
  - Oil shows a somewhat significant correlation of around 0.6 (that too overall - including ‘black swan’ events (like the Russia-Ukraine situation?)
  - Gas shows lesser correlation than oil
  - Electricity shows some high correlations but for lags upto 17 weeks etc.
- Business Cycle (DAX) shows some significant inverse correlation
- Construction permits show high inverse correlation for some clusters

Recommendations

- For the client's forecasting, we suggest the following
  - Avoid weather as a determinant as there is very little correlation
  - Avoid gas for a similar reason
  - Consider oil as a determinant but keep in mind major global events that might upend the trend completely
  - Consider electricity but further exploration may be needed to see if it was a coincidence. And for best results, consider lags of around 17 weeks
  - Consider DAX and Construction permits but in an inverse correlation

Thank you.