PA6 Cost Prediction

Schneider Electric Data Challenge

24/11/2023

Our Team

Alexis Constensoux
HEC - X-HEC DSB

Cedomir Ljubic
HEC - X-HEC DSB

Matthieu Marquis
HEC - X-HEC DSB

Lauren Segalla
McGill - X-HEC DSB

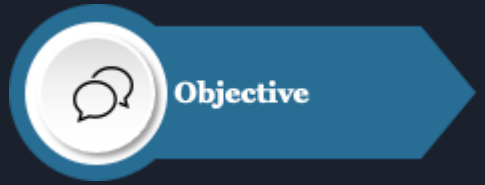
Shihan Ye
University of Melbourne- X-HEC DSB

Business Issue

Schneider Electric's reliance on plastic materials in its operations

Difficulty in accurately predicting costs due to fluctuating market factors

Inaccurate cost predictions can decrease profitability



Estimate future costs of Polyamide 6 (PA6) plastic raw material over 3, 6, and 9 months

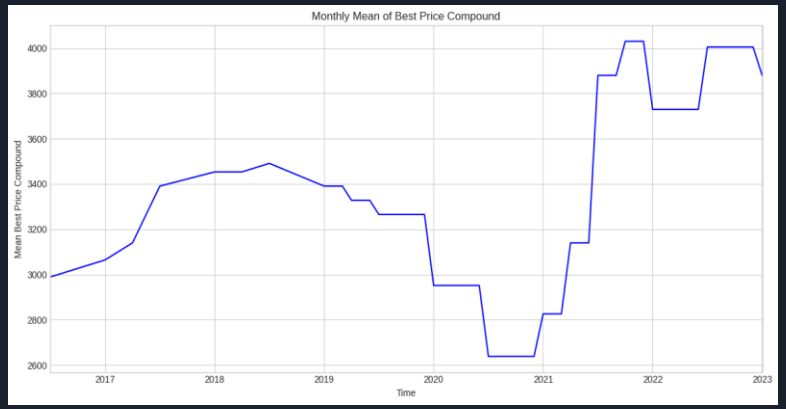
Informed decision-making, optimized budgeting and minimized financial risks

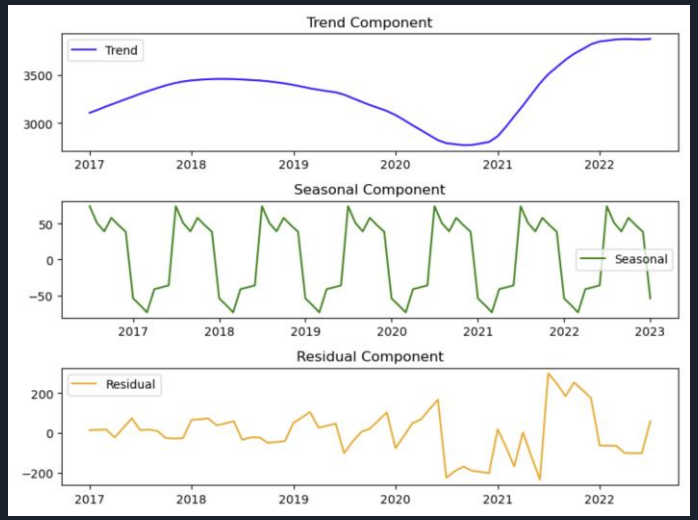
Problem Statement:

- Predict the best supplier price for 3 successive quarters, starting 01/01/2023
- We assume 3-month contracts with suppliers

EDA: Seasonal Decomposition

"Best price" varies every month initially, then every 3 or 6 months after 2019, due to supplier contracts.

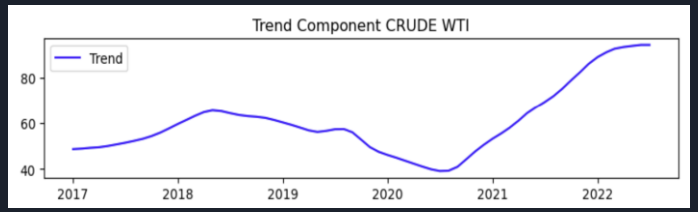




The trend seems to be macroeconomic.

Shows high growth in the past 2 years and a strong dip during the year 2020.

Seasonality is marginal but highlights a mid-year peak.

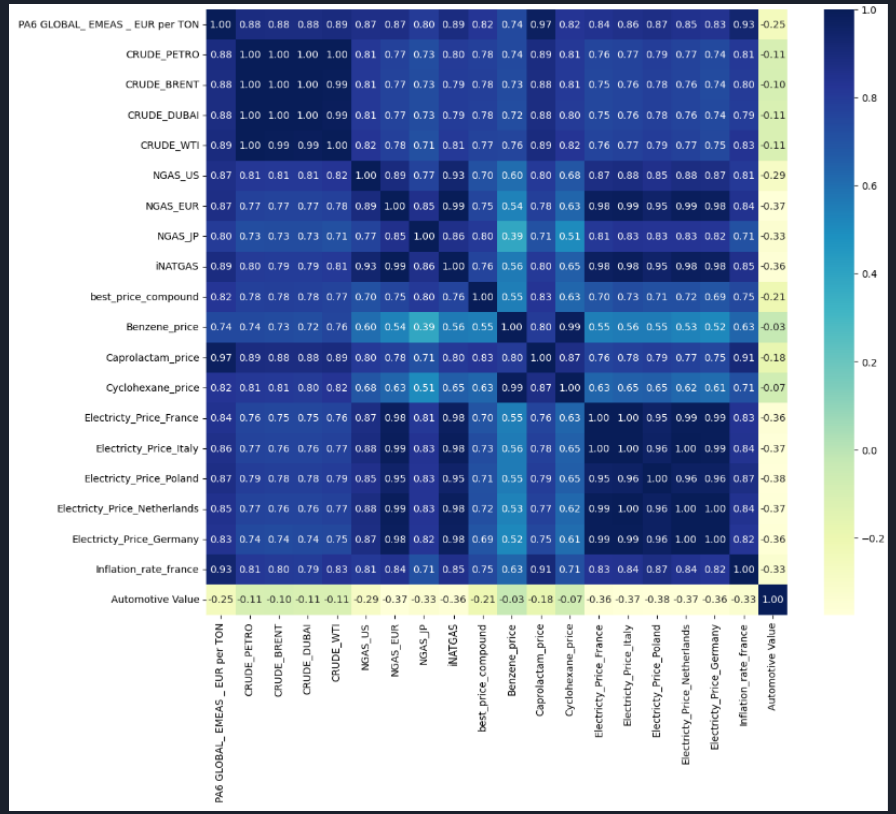


EDA: Lag Observations


```
same_month_correlation =0.83
correlation_shift_1 =0.86
correlation_shift_2 =0.91
correlation_shift_3 =0.93
```

Caprolactam prices from 3 months ago are most highly correlated with best price compound!

EDA: Correlational Analysis



The dataset shows very **high correlations** amongst almost all values.

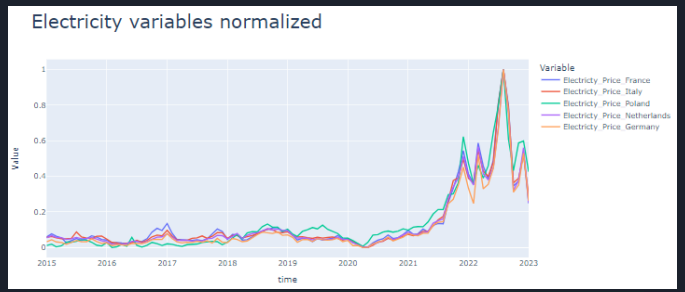
There is an almost **perfect positive correlation** between the **electricity prices** in Europe and between the **crude oil prices** from different continents.

The prices of hydrocarbons in the market showcase only a moderate correlation with the prices of different types of natural gases.

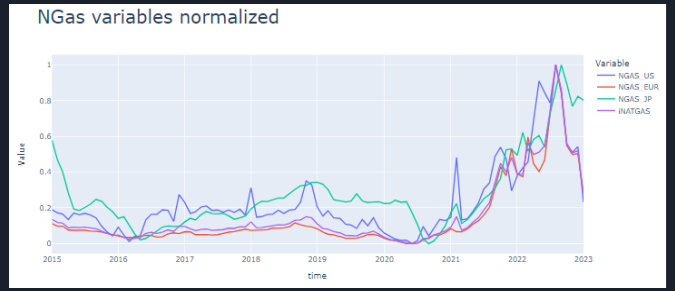
Automotive value is **weakly negatively correlated** to all other features.

Feature Engineering: Reducing Multicollinearity

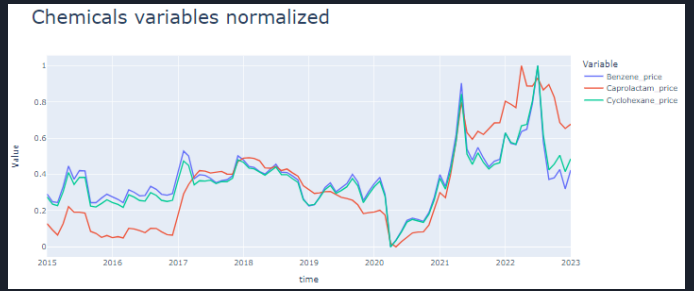
→ Keep only one of crude oil variables



→ Only keep one country at random



→ Take mean of natural gas



→ Keep either Benzene or Cyclohexane

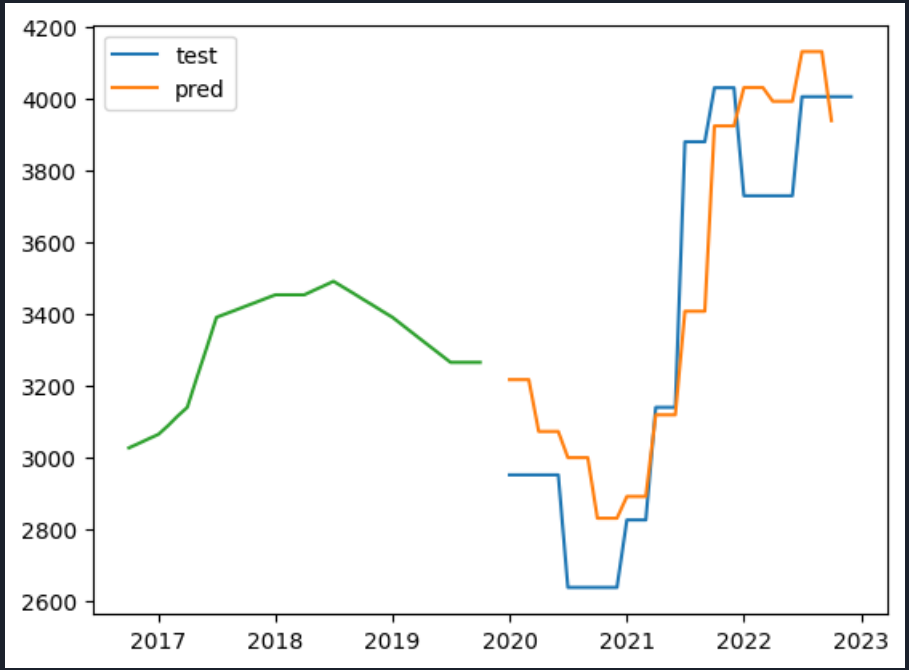
Prediction horizon 1: January-March 2023

- For this prediction we trained a model that predicts **next** months "best price"
- This model is most useful at end of quarter, to predict next quarter prices

Using **linear model** for simplicity and interpretability

For this horizon, the most data at our disposal

Modelling: Linear Model for 1 month prediction



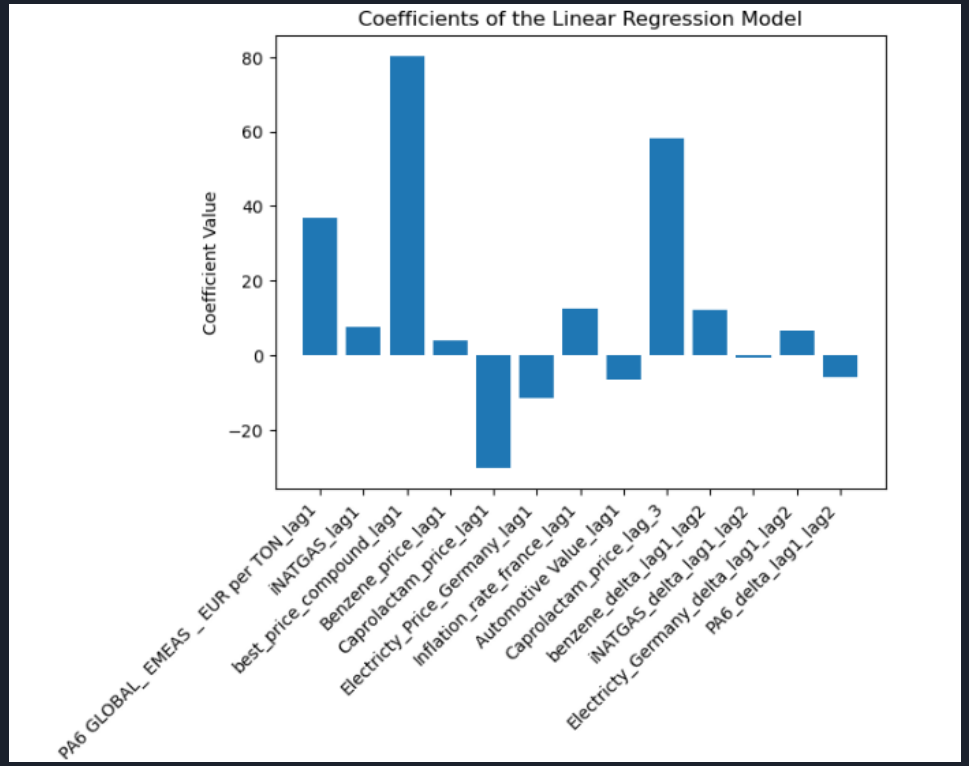
- Trained in 2016-2020
- Tested in 2020-2022

- All predictions made EOQ for next month

- **MAPE: 5.77%**
- **MAE: 196.8**

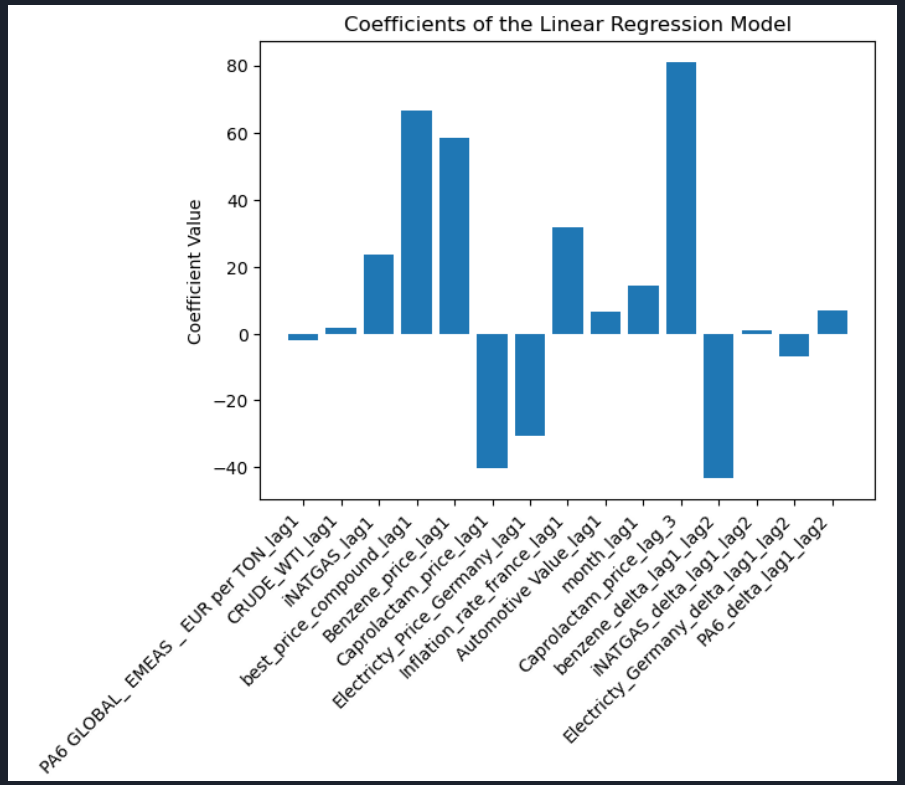
- Target prediction for 01/01/2023: 3665
- True value: 3879

Modelling: Feature Importance up to 2020



- **High autocorrelation** (dependance on best price lag 1)
- High importance of **Caprolactam lag 3** as anticipated
- High importance of **PA6 Global lag 1**

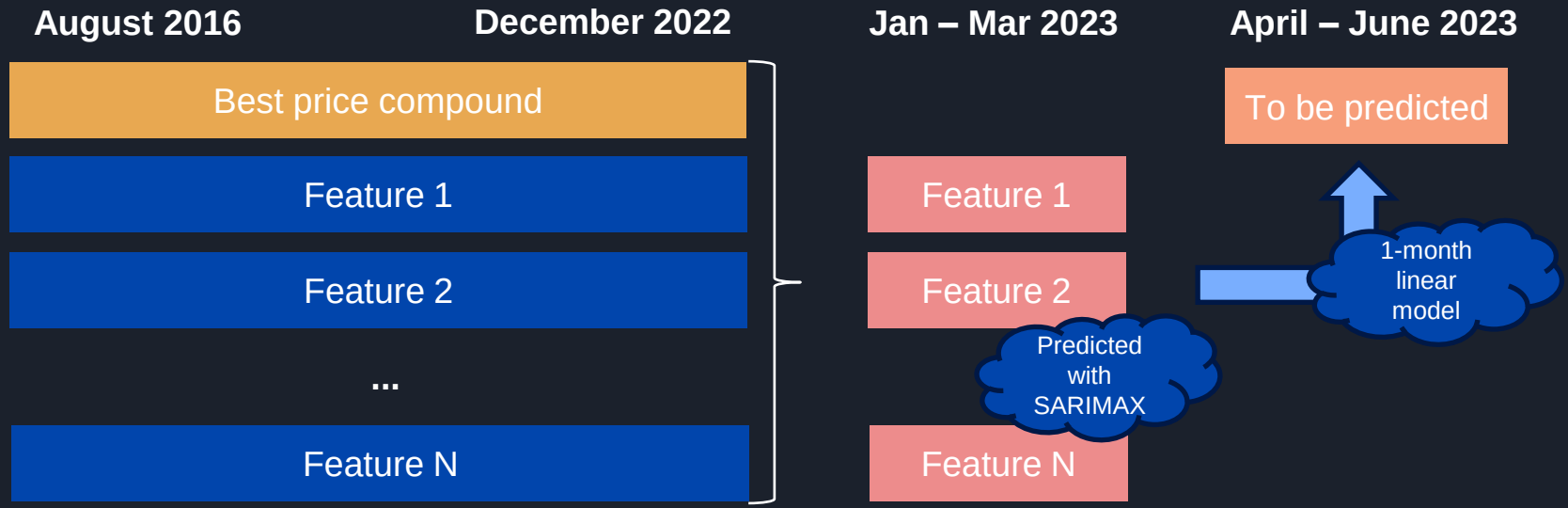
Modelling: Feature Importance trained up to 2022



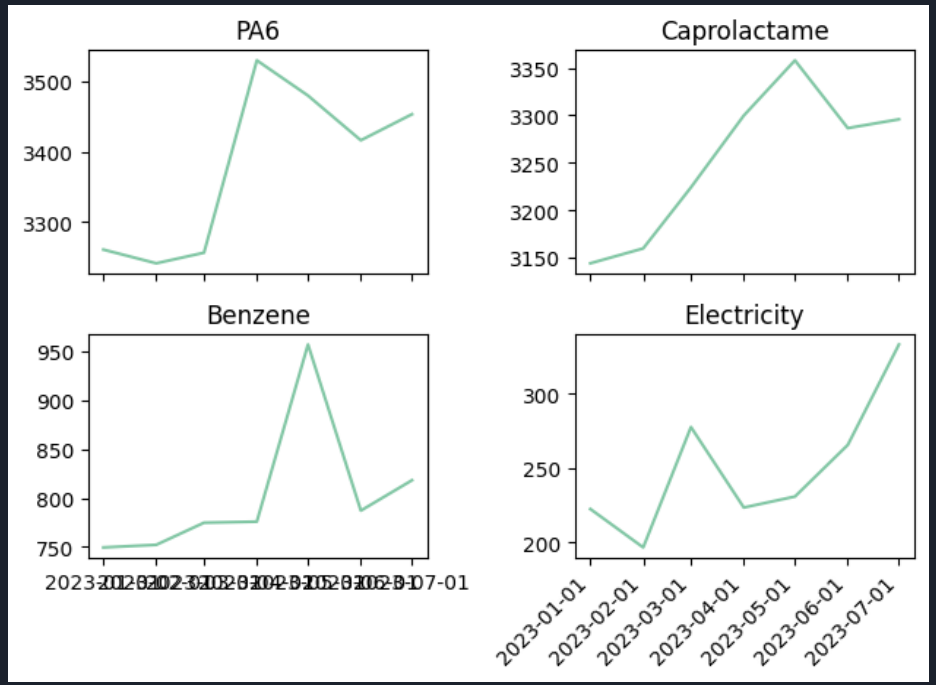
- Still high autocorrelation
- Even higher importance of **Caprolactam lag 3** as anticipated
- Big increase Benzene importance

Prediction horizon 2: April-June 2023

- Model predicts best price **4 months ahead**
- First predicts X features 3 months ahead using SARIMAX
- Then uses 1-month linear model on top of SARIMAX predictions

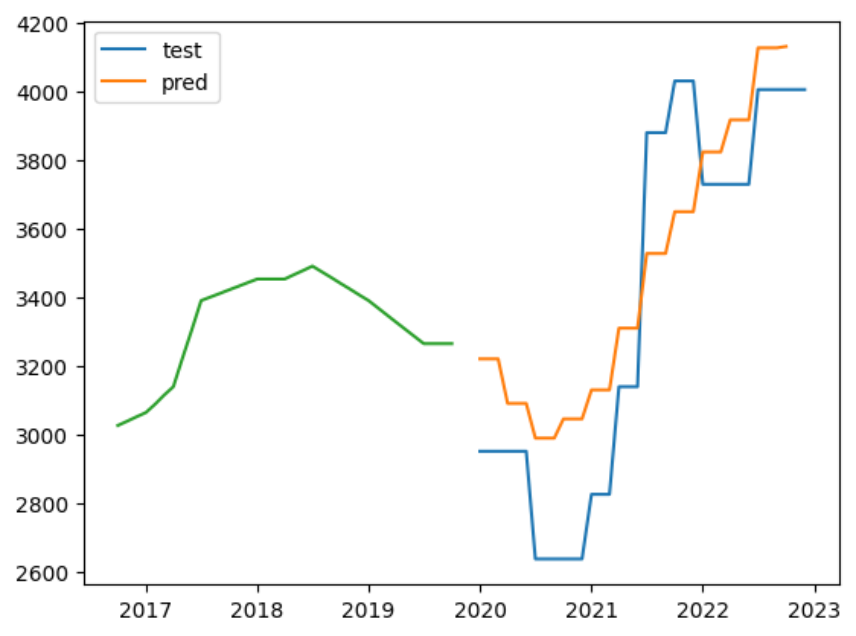


Predicted features with SARIMAX



- Prediction of the 4 most important features using SARIMAX model

Modelling: linear model for long term prediction



- Only trained on the important selected features

- **Trained in 2016-2020**
- **Tested in 2020-2022**

- MAPE: 7.50%
- MAE: 242

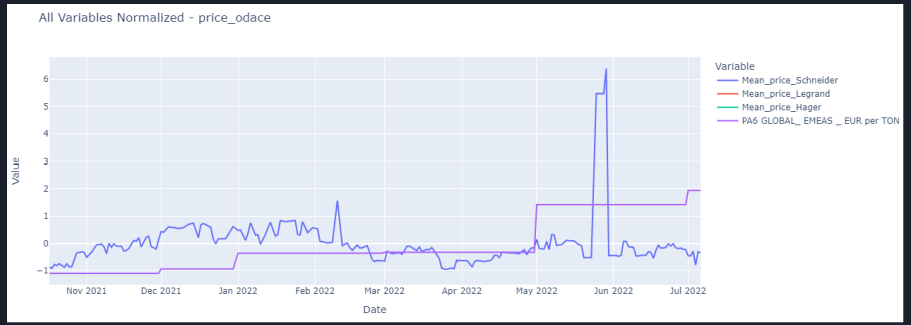
- Less precise than 1 month model

Link between Raw Material and Business Trends

Expectation 1 : Prices for products with highest PA6 % are most correlated with PA6 global compound prices

Expectation 2 : Most likely lag between PA6 global price and finished product prices, due to production time

Odace is made of 20% of PA6



PA6 price correlation

Correlation Coefficients for price_odace_PA6	
Mean_price_Schneider	0.148754
Mean_price_Legrand	NaN
Mean_price_Hager	NaN

Best price correlation

Correlation Coefficients for price_odace_	
Mean_price_Schneider	-0.055981
Mean_price_Legrand	NaN
Mean_price_Hager	NaN

Price analysis – other products

IC60 product (33,2% PA6)

Correlation Coefficients for price_IC60_PA6:	
Mean_price_SE	0.759527
Mean_price_Hager	-0.120425

GV2 product (2,2% of PA6)

Correlation Coefficients for price_GV2_PA6:	
Mean_price_SE	0.841660
Mean_price_ABB	0.418522
Mean_price_Siemens	0.060293

Executive Summary

- Objective: using time series data to predict next 3/6/9 months PA6 cost and provide pricing strategies
- Methods:
  1. EDA highlighted importance of feature engineering to remove multicollinearity and add lagged features, such as Caprolactam_lag_3
  2. Utilization of 3 simple models, with decreasing precision, one for each horizon
  3. Using predicted features from SARIMAX model together with a linear model to predict the 6/9-months price

- Prediction results:

3-month	6-month	9-month
3665	3539	3726

- Recommendations:
  1. Whenever the 1-month prediction is significantly lower than 4-month prediction, negotiate a 6-month supplier contract instead of 3-months
  2. Keep an eye on the highlighted compounds from our analysis, and macroeconomic trends