

Business rules by family (SpP)

Here is the full technical validation deck for the latest review of the models at SpP:

Please also note that we conduct outliers exclusion (detailed [here](#)).

You will find below the included specificity of each family.

Ketaspire

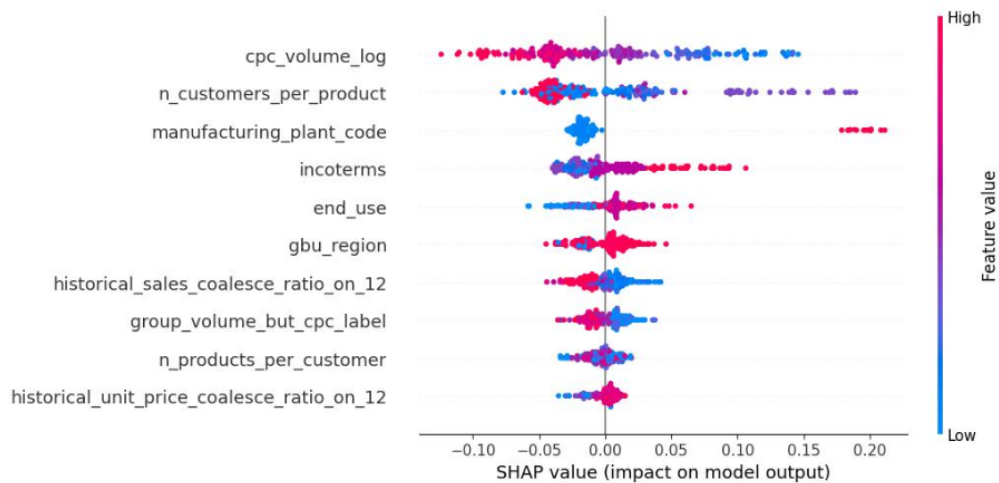
Exclusions

PK KT-857 NT (Z58-41636) => This product is driving the R2 score of the Ketaspire family down for no obvious reason (its price is not an outlier, nor are any of its features used in this family). We decided to excluded it because it only represents 0.5% of the sales of the family.

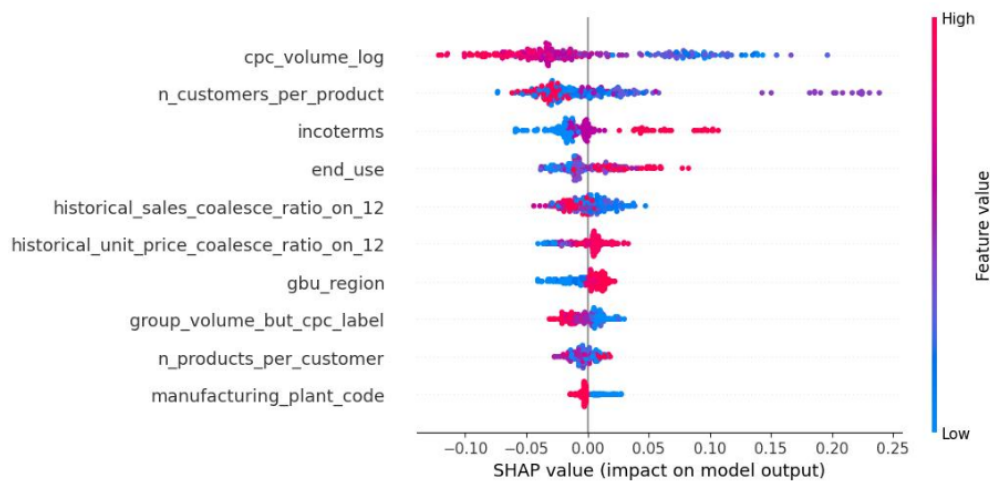
PVDC

Exclusions

EMULDUR 381 A (Z68-40339) => This product is the only one tagged with a (3RD PARTY) manufacturing plant and has a high price, therefore it brings by its own the SHAP value of the manufacturing_plant_code feature way higher than it should be. To avoid this happening, we decided to exclude this product for now.



SHAP values before exclusion



SHAP values after exclusion

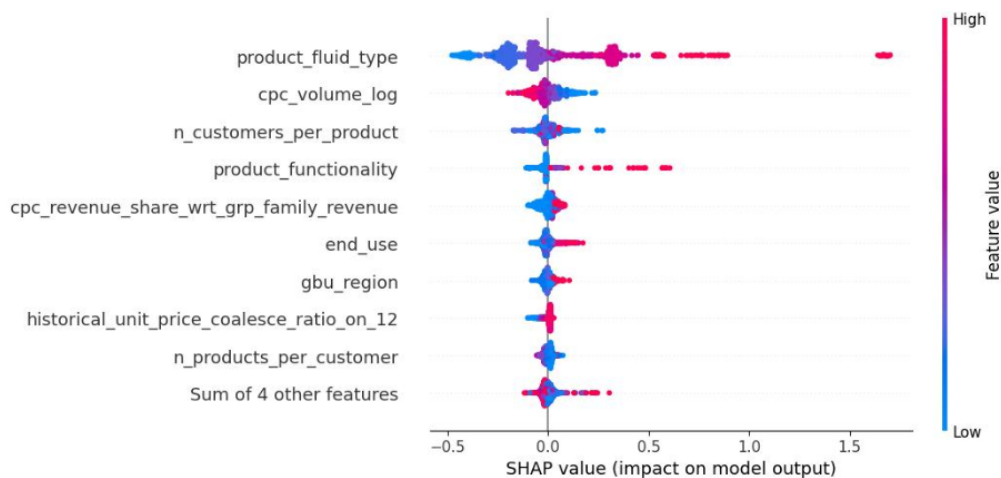
Fluids

A big part of the weights (almost 50%) is focused on *product_fluid_type* only. This feature alone will define most of the comparable set. This should be considered when analyzing results of this family.

Exclusions

The following products in the fluids family (Fomblin, Fluorolink) are sold at very high prices and low volumes. Having a few of them as comparable for a CPC could bring the median way higher than expected. If we include these products, their associated CPCs also end up with a price decrease recommendation of about 95% of the price.

Note that since they often have dedicated fluid type, we might be able to use "product_fluid_type" or a derived new column of the taxonomy to identify these products to either filter them out or add a hard-boundary to compare them altogether and not with the rest of the data.



The red dots on the far right of the "product_fluid_type" SHAP values correspond to these products

List of products (name // code)

FOMB SD10 // Z61-39202

FOMB ZMF 402 // Z61-34500

FOMB T4 (T4) // Z61-25358

FOMB ZMF 520(ZMF520) // Z61-40370

FOMBLIN ZMF 23 // Z61-39572

FOMBLIN R&D // Z61-26326

FLUOROLINK D 4000 // Z61-25252

Galden

Galden was initially part of the fluids family but was split due to having too different behavior than the other h4. Analysis for this separation is done [here](#).

There is a specific behavior of the *n_customers_per_product* feature to be noted for Galden, details are in [this ticket](#).

Exclusions

Products with the "Out of scope" value in the taxonomy for *product_boiling_point* are excluded from the model.

Tecnoflon FFKM

Some products for Tecnoflon FFKM have very high prices (details [here](#)). They have therefore been filtered out of the scope.

List of products (name // code)

TECNOFLON SHP LT // Z61-41519

TECNOFLON PFR LT // Z61-29317

TECNOFLON PFR95HT // Z61-24239

TECNOFLON PFR 94 // Z61-24236

Udel & Radel

Hard boundary

- the **Market segment** feature is used to create a new HB feature to separate "HC" from the rest of market segments (details [here](#)).

Ryton

Hard boundary

- the **Product is CN** feature from [SpP product taxonomy file](#) is used to create a new HB feature to separate "CN" from the rest of products (details [here](#)).

Amodel

Hard boundary

- the **Product additive** feature from [SpP product taxonomy file](#) is used to separate "HFFR" from the rest of products (details [here](#)).
- the **Manual region** feature from the [SpP manual data sources file](#) is used to separate "Greater China" and "Americas" from the rest of regions (details [here](#)).