LG data science home test

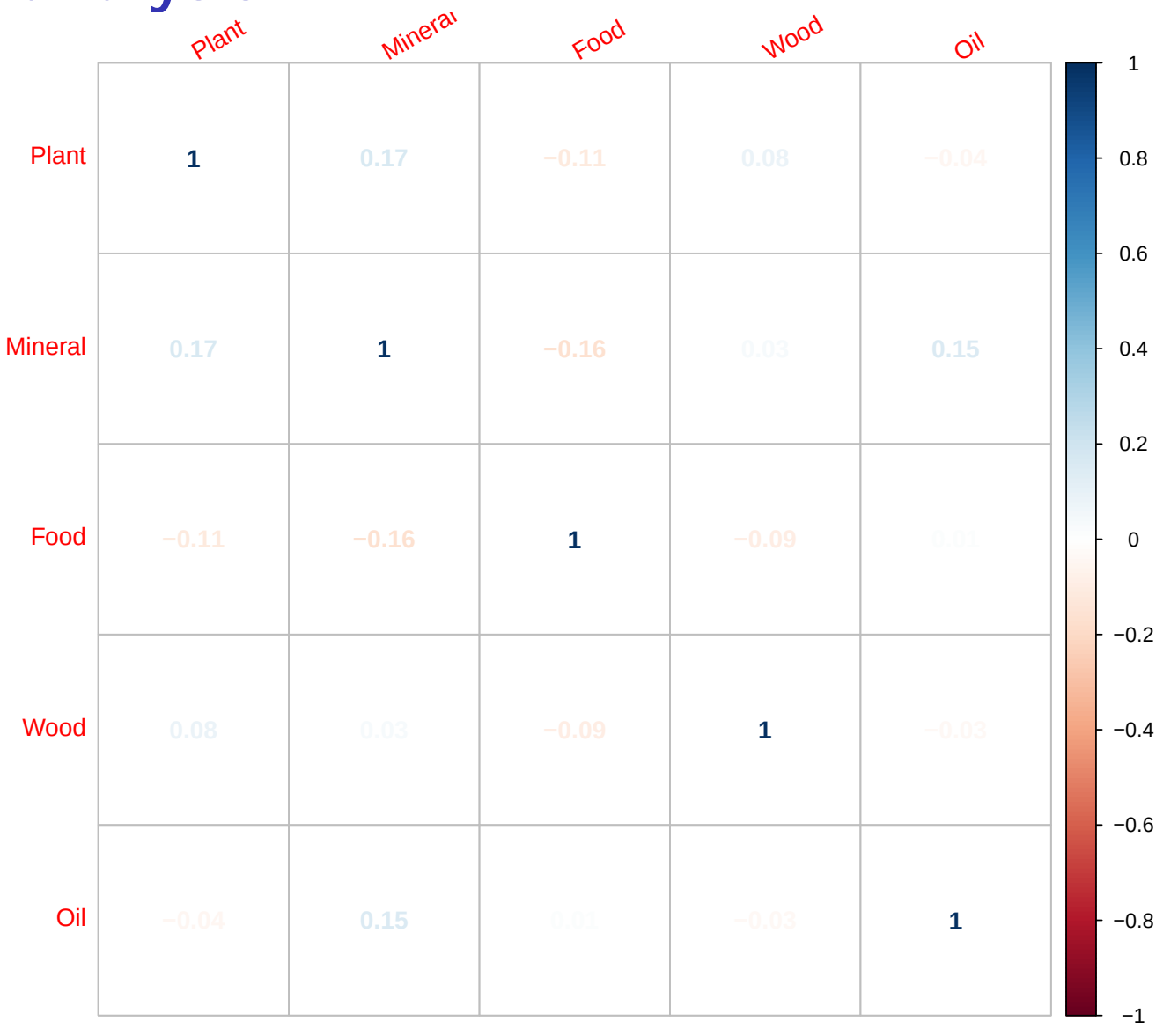
Edgaras Juodsnukis

2020-05-11

Content

- Exploratory analysis: correlation analysis.
- Modelling results: ME (mean error), RMSE (root mean square error), MAE (mean absolute error) and percentage errors.
- Additional: included in code and further analysis.

Exploratory analysis



Modelling results (univariate model, quarterly)

In case of the one-dimensional ARIMA models, the largest percentage error was found in the modeling of food industry products (~38%). In other cases, the percentage error ranged between 6% and 17%.

	Plant	Mineral	Food	Wood	Oil
ME	138092.5	-52056.8	-171667.2	8518.9	-113731.5
RMSE	167392.8	171175.4	203443.5	41584.7	212467.7
MAE	145983.7	144847.0	174149.8	30955.4	205472.7
MPE	15.8	-4.9	-38.1	1.3	-3.4
MAPE	16.8	10.7	38.4	7.9	5.8

Modelling results (multivariate model, quarterly)

In case of VAR model, the largest percentage error was found in the modeling of food industry products. The smallest error was found in forecasting oil (~7%).

	Plant	Mineral	Food	Wood	Oil
ME	80816.4	243572.8	-431135.9	36285.4	258482.1
RMSE	266237.2	282548.2	489704.9	46287.8	366903.0
MAE	244451.3	243572.8	431135.9	36285.4	258482.1
MPE	10.4	16.7	-92.4	9.7	7.0
MAPE	28.8	16.7	92.4	9.7	7.0

Additional

- Dickey-Fuller Unit Root test showed that all of the variables are I(1). This means that first differences of the all variables are stationary.
- Impulse-response analysis showed that products of plant origin and solid mineral fuel could be related in short-term relationships. Furthermore, products of plant origin reacts to the oil and petroleum products shocks.
- Forecast error variance decomposition showed that proportions of changes in the variable due to their own shocks suppress to 70% in 10th period, while this proportion seeks 95% in the 1st period.