Inflation data

As inflation data, we use monthly data for the Czech republic in the time period from January 2004 up to February 2023. In the choice of which attitude to adopt, we decided to use month-on-month inflation data, as they are much rather able to capture the monht-on-month trend for which it makes most sense to look for in the Google Trends data. Next, the data are gotten rid of seasonality.

GT

Google Trends data serve as our source of external regressors. We begin by specifying our inflation-related terms. We decided to use for now to use just several; 3 specific inflation related terms, word „“cena“ and 8 most searched combinations of word “cena“ with something else. Next, Principal Component Analysis (PCA) is introduced. Out of 13 PCA components, only 5 turn out to be valuable with standard deviations being higher than 1. Thus, these are added to our dataset.

Něco o granger causalite

Then, many Arima models are run. According to Autocorroleation and Partial Autocorrelation function, the best attitude would be to try different Arimas with different settings for its parameters. The autoregressive and moving average settings of the model were both ordered to domain of [1,2,3] and the degree of differencing to [0,1]. As input, we use inflation data and external regressor and three ways: aligned, lagged or lagged but to opposite direction. We discard the models with external regressor having p-value higher than 0.1. For every model that comes out with external regressor being statistically significant we build benchmark model without the external regressor in order to observe and compare information criteria. We save all these models.

We repeat this process for three different time intervals due to recent years of instability that might not provide good information background in contrast of recently calmer decades before locally in the Czech Republic. These time intervals start at January 2014 and end in December 2019, February 2022 and February 2023 respectively.