

In [104]...

```
%load_ext autoreload
%autoreload 2

import sys
sys.path.append("../")

import datetime as dt
import pandas as pd
import numpy as np
import statsmodels.api as sm
from statsmodels import tsa
import statsmodels.graphics as sg
from statsmodels.stats import stattools as st
import sklearn

import matplotlib.pyplot as plt

import seaborn as sns
sns.set_theme()

%matplotlib inline
```

The autoreload extension is already loaded. To reload it, use:

```
%reload_ext autoreload
```

In [86]:

```
df = pd.read_csv('../data/ts.csv', index_col=0, parse_dates=True)
df.shape, df.head()
```

Out[86]: ((43, 11),

	rynek	inflacja_r	inflacja_q	stopa_procentowa	\
2014-03-31	6193.211456	100.6	100.2	0.0446	
2014-06-30	6227.984614	100.3	100.0	0.0447	
2014-09-30	6357.260998	99.7	99.5	0.0443	
2014-12-31	6438.874574	99.3	99.6	0.0388	
2015-03-31	6473.616559	98.5	99.5	0.0348	

	liczba_kredytow	tempo_wzrostu	ufnosc	duze_zakupy	bezrob
2014-03-31	41942.0	104.2	-20.5	-16.3	
2014-06-30	45499.0	104.1	-16.8	-15.0	
2014-09-30	43653.0	103.3	-16.9	-12.7	
2014-12-31	42993.0	103.3	-13.8	-10.7	
2015-03-31	42169.0	104.0	-14.9	-11.5	

	spr_detaliczna	pkb
2014-03-31	105.5	103.8
2014-06-30	105.1	103.9
2014-09-30	102.6	104.2
2014-12-31	103.7	103.9
2015-03-31	104.4	104.3

```
In [87]: df2 = pd.read_csv('../data/wynagrodzenie.csv', index_col=0, parse_dates=True)
df2.shape, df2.head()
```

```
Out[87]: ((58, 6),
          d_gosp \      wynag_gosp  wynag_przed  zatrud_ogolem  zatrud_przed  zatrud
2010-03-31      3241.88      3382.45      8034.0      5304.0
2730.0
2010-06-30      3270.87      3420.32      8110.0      5351.0
2759.0
2010-09-30      3258.95      3424.61      8176.0      5392.0
2784.0
2010-12-31      3371.07      3499.96      8214.0      5449.0
2765.0
2011-03-31      3392.00      3521.75      8264.0      5512.0
2752.0

          wynag
2010-03-31  3334.683495
2010-06-30  3369.477515
2010-09-30  3368.201311
2010-12-31  3456.572996
2011-03-31  3478.541868 )
```

```
In [88]: df = pd.merge(left=df, right=df2["wynag"], left_index=True, right_index=True)
df.head()
```

```
Out[88]:
```

	rynek	inflacja_r	inflacja_q	stopa_procentowa	liczba_kredytow	ten
2014-03-31	6193.211456	100.6	100.2	0.0446	41942.0	
2014-06-30	6227.984614	100.3	100.0	0.0447	45499.0	
2014-09-30	6357.260998	99.7	99.5	0.0443	43653.0	
2014-12-31	6438.874574	99.3	99.6	0.0388	42993.0	
2015-03-31	6473.616559	98.5	99.5	0.0348	42169.0	

```
In [89]: df.index
```

```
Out[89]: DatetimeIndex(['2014-03-31', '2014-06-30', '2014-09-30', '2014-12-31',
                        '2015-03-31', '2015-06-30', '2015-09-30', '2015-12-31',
                        '2016-03-31', '2016-06-30', '2016-09-30', '2016-12-31',
                        '2017-03-31', '2017-06-30', '2017-09-30', '2017-12-31',
                        '2018-03-31', '2018-06-30', '2018-09-30', '2018-12-31',
                        '2019-03-31', '2019-06-30', '2019-09-30', '2019-12-31',
                        '2020-03-31', '2020-06-30', '2020-09-30', '2020-12-31',
                        '2021-03-31', '2021-06-30', '2021-09-30', '2021-12-31',
                        '2022-03-31', '2022-06-30', '2022-09-30', '2022-12-31',
                        '2023-03-31', '2023-06-30', '2023-09-30', '2023-12-31',
                        '2024-03-31', '2024-06-30'],
                        dtype='datetime64[ns]', freq=None)
```

```
In [90]: df["inflacja_q"]
```

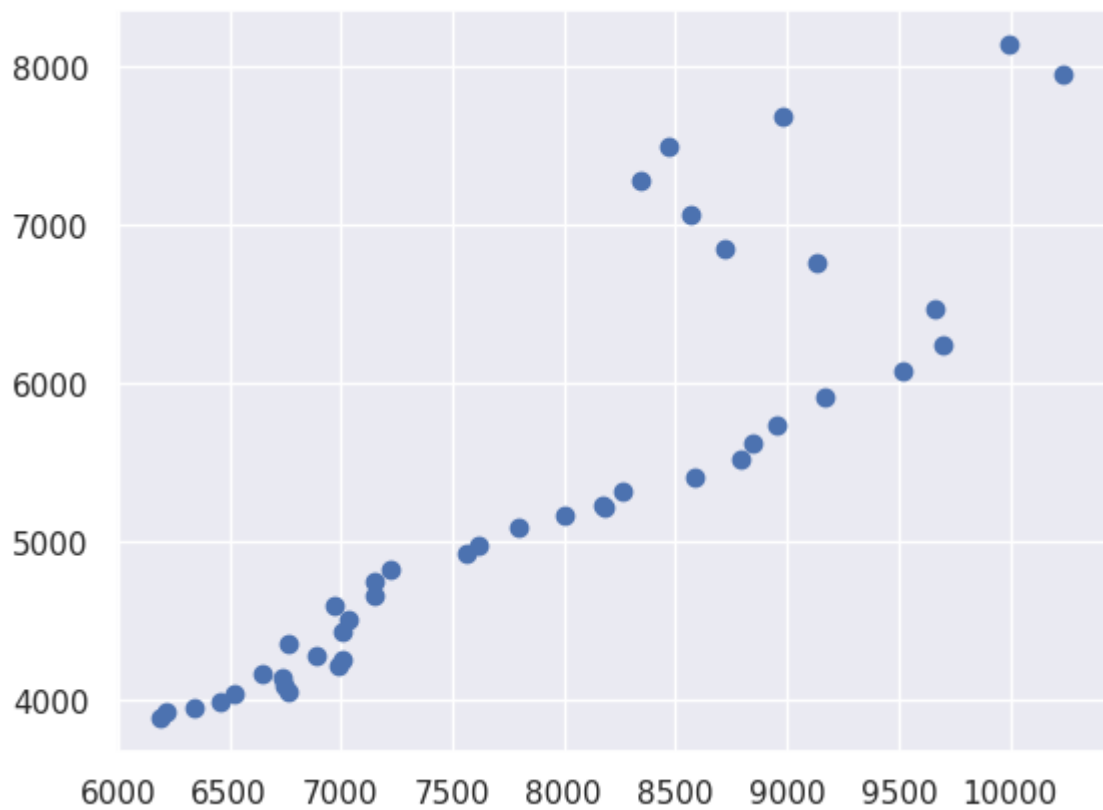
```
Out[90]: 2014-03-31    100.2
          2014-06-30    100.0
          2014-09-30     99.5
          2014-12-31     99.6
          2015-03-31     99.5
          2015-06-30    100.5
          2015-09-30     99.6
          2015-12-31     99.7
          2016-03-31     99.3
          2016-06-30    100.5
          2016-09-30     99.7
          2016-12-31    100.7
          2017-03-31    101.1
          2017-06-30    100.3
          2017-09-30     99.8
          2017-12-31    101.1
          2018-03-31    100.4
          2018-06-30    100.5
          2018-09-30    100.0
          2018-12-31    100.5
          2019-03-31    100.2
          2019-06-30    101.7
          2019-09-30    100.3
          2019-12-31    100.6
          2020-03-31    102.0
          2020-06-30    100.3
          2020-09-30    100.1
          2020-12-31    100.4
          2021-03-31    102.1
          2021-06-30    101.9
          2021-09-30    101.0
          2021-12-31    102.6
          2022-03-31    103.8
          2022-06-30    105.8
          2022-09-30    103.1
          2022-12-31    103.6
          2023-03-31    104.3
          2023-06-30    101.9
          2023-09-30     99.7
          2023-12-31    100.5
          2024-03-31    100.9
          2024-06-30    101.4
          Name: inflacja_q, dtype: float64
```

```
In [91]: infl_arr = df["inflacja_q"].to_numpy()
          inflacja_q_acc_arr = np.ones_like(infl_arr)
          num = len(infl_arr)
          for i in range(1, num):
              inflacja_q_acc_arr[i] = inflacja_q_acc_arr[i-1] / 100 * infl_arr[i-1]

          df["inflacja_q_kum"] = inflacja_q_acc_arr
          df.tail()
```

Out[91]:

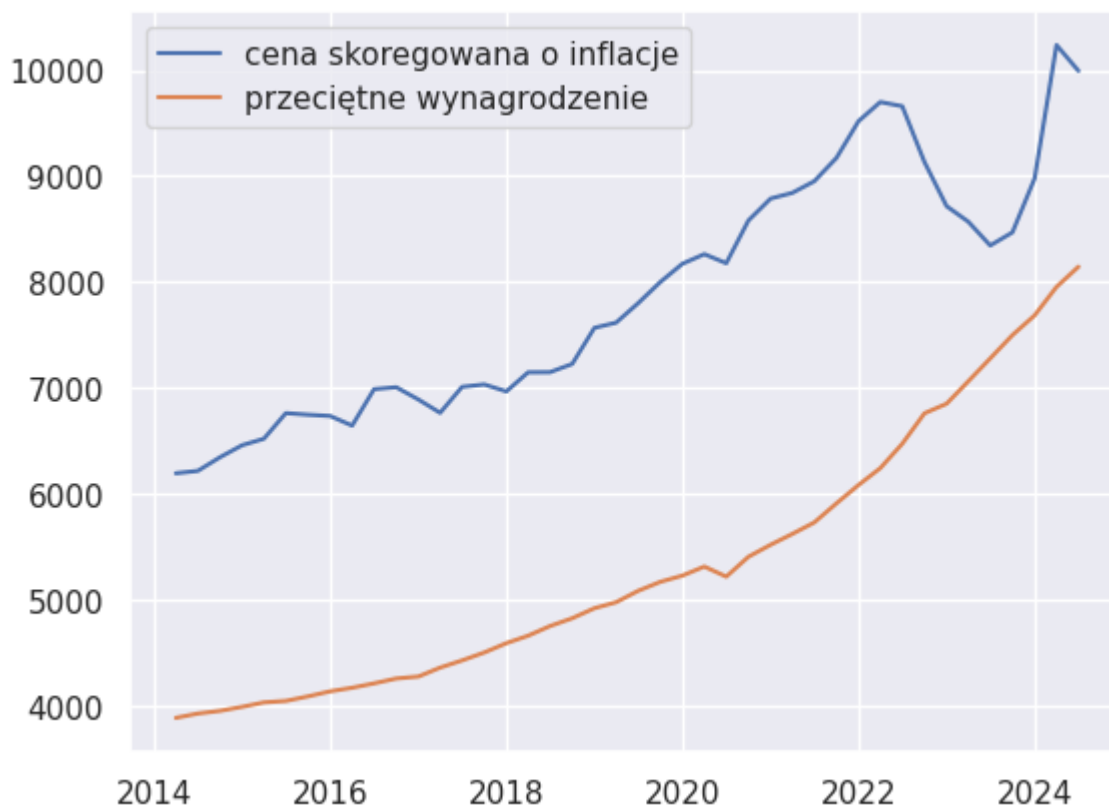
	rynek	inflacja_r	inflacja_q	stopa_procentowa	liczba_kredytow	te
2023-06-30	11928.471147	113.1	101.9	0.0877	30798.0	
2023-09-30	12335.591393	109.7	99.7	0.0779	40749.0	
2023-12-31	13037.926475	106.4	100.5	0.0808	68860.0	
2024-03-31	14942.844839	102.8	100.9	0.0796	64504.0	
2024-06-30	14718.380953	102.5	101.4	0.0819	45434.0	

In [92]: `plt.scatter(df["rynek"]/inflacja_q_acc_arr, df["wynag"])`Out[92]: `<matplotlib.collections.PathCollection at 0x724fc889b7d0>`In [93]: `df.wynag / inflacja_q_acc_arr`

```
Out[93]: 2014-03-31    3882.041728
          2014-06-30    3914.975186
          2014-09-30    3940.430945
          2014-12-31    3998.108998
          2015-03-31    4056.473908
          2015-06-30    4090.768809
          2015-09-30    4113.869572
          2015-12-31    4178.640019
          2016-03-31    4225.437307
          2016-06-30    4297.411260
          2016-09-30    4322.843980
          2016-12-31    4354.716840
          2017-03-31    4409.671599
          2017-06-30    4429.780547
          2017-09-30    4491.084020
          2017-12-31    4588.612698
          2018-03-31    4608.197746
          2018-06-30    4679.868900
          2018-09-30    4728.949529
          2018-12-31    4821.672008
          2019-03-31    4852.798668
          2019-06-30    4947.681161
          2019-09-30    4946.487823
          2019-12-31    4988.298310
          2020-03-31    5039.541700
          2020-06-30    4852.340365
          2020-09-30    5011.815138
          2020-12-31    5111.930765
          2021-03-31    5185.957566
          2021-06-30    5177.280618
          2021-09-30    5240.658997
          2021-12-31    5341.937071
          2022-03-31    5342.535822
          2022-06-30    5336.626699
          2022-09-30    5268.284068
          2022-12-31    5178.588088
          2023-03-31    5153.387216
          2023-06-30    5092.113515
          2023-09-30    5148.150930
          2023-12-31    5292.040294
          2024-03-31    5450.365993
          2024-06-30    5531.079724
          Name: wynag, dtype: float64
```

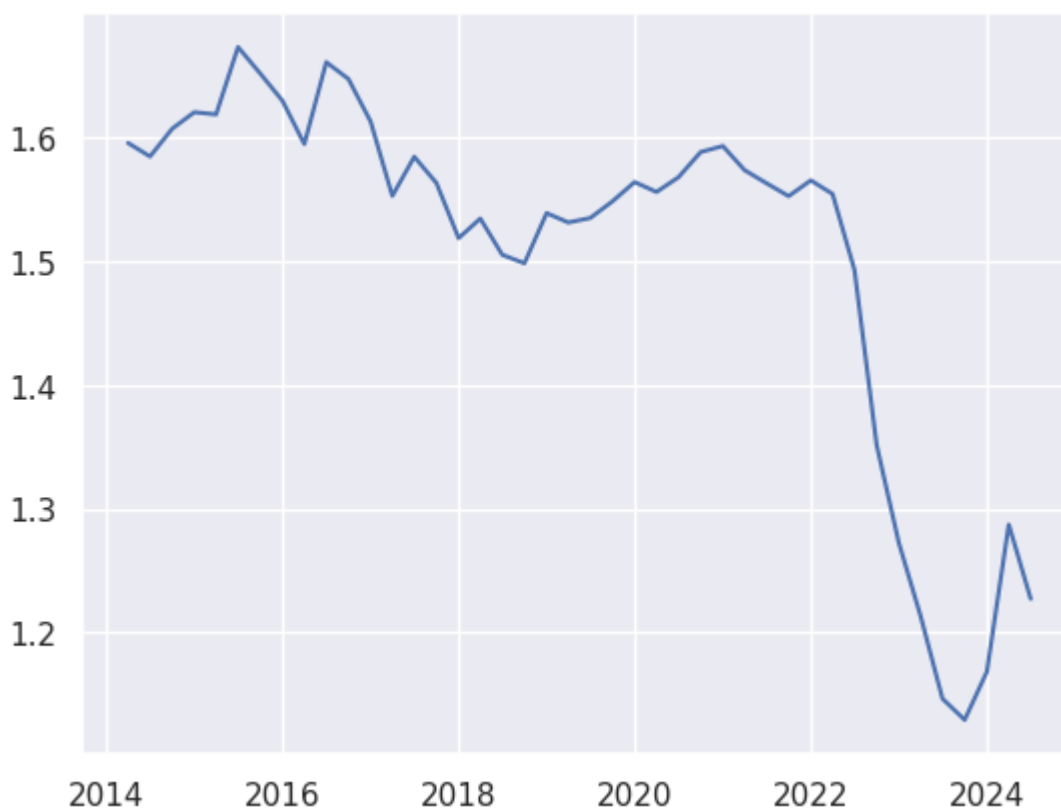
```
In [111]: fig = plt.Figure()
          ax = fig.gca()
          ax.plot(df.index, df.rynek / df.inflacja_q_kum, label="cena skoregowana o
          ax.plot(df.index, df.wynag, label="przeciętne wynagrodzenie")
          ax.legend()

          fig.savefig('../images/ceva_vs_wynag.png')
          display(fig)
```



```
In [95]: plt.plot(df.index, df.rynek / (df.inflacja_q_kum * df.wynag) )
```

```
Out[95]: [<matplotlib.lines.Line2D at 0x724fc8789790>]
```



```
In [96]: ddf = pd.read_csv('../data/gus_budowa.csv', header=0, parse_dates=["data"]
ddf.set_index(keys="data", drop=True, inplace=True)
ddf.dtypes, ddf.head()
```

```

Out[96]: (ogolem                float64
          indywidualne         float64
          sprzedaz_lub_wynajem float64
          spoldzielcze          float64
          pozostale             float64
          dtype: object,
          ogolem  indywidualne  sprzedaz_lub_wynajem  spoldzielcze
\
data
2014-01-01  12511.0          7155.0          4449.0          415.0
2014-02-01  12462.0          6714.0          5068.0          358.0
2014-03-01  10589.0          6530.0          3687.0          239.0
2014-04-01  11646.0          6102.0          5095.0          193.0
2014-05-01   8893.0          5516.0          3185.0           46.0

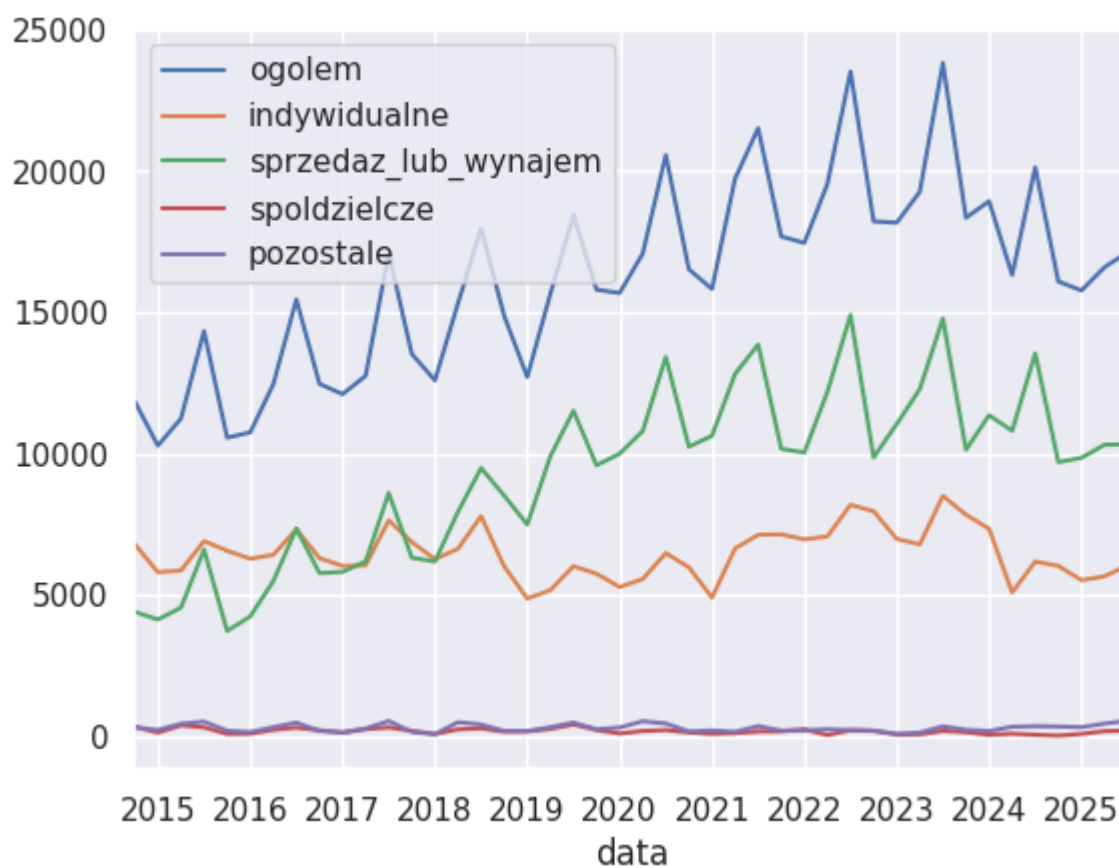
          pozostale
data
2014-01-01    492.0
2014-02-01    322.0
2014-03-01    133.0
2014-04-01    256.0
2014-05-01    146.0 )

```

```
In [97]: ddf = ddf.resample(pd.offsets.QuarterEnd()).mean()
```

```
In [98]: ddf.plot()
```

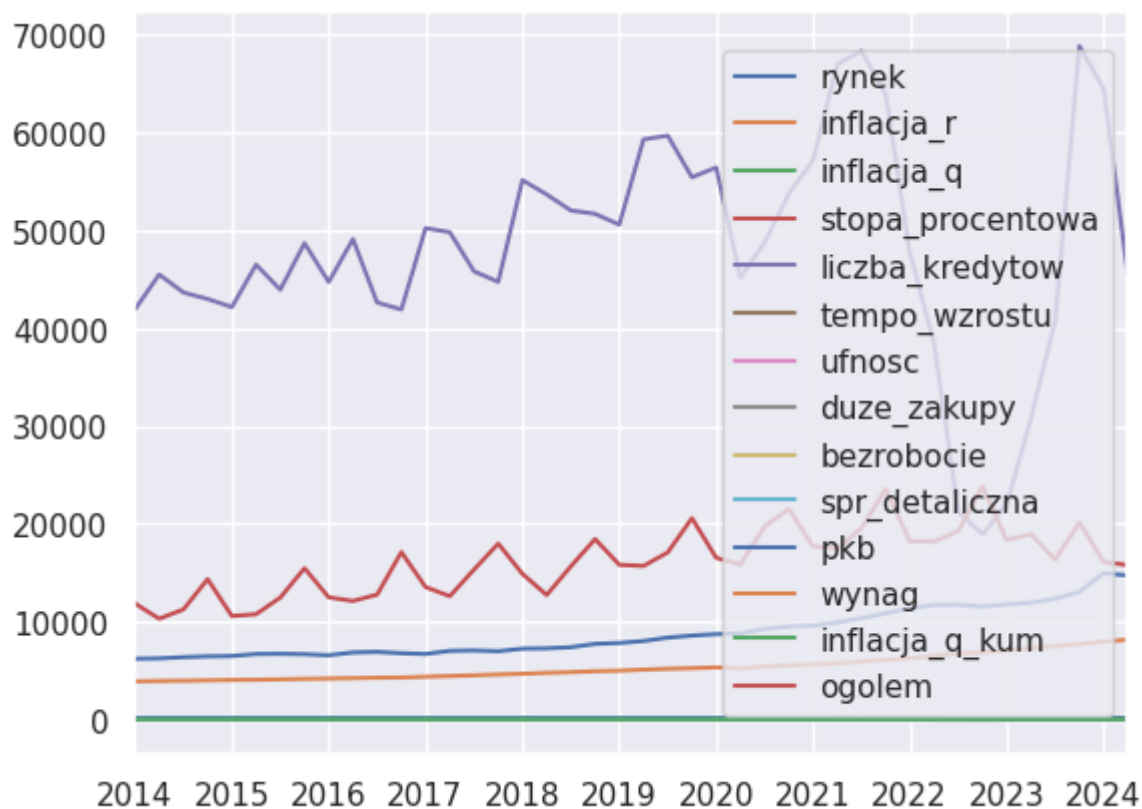
```
Out[98]: <Axes: xlabel='data'>
```



```
In [99]: df = pd.merge(left=df, right=ddf["ogolem"], left_index=True, right_index=True)
df.head()
```

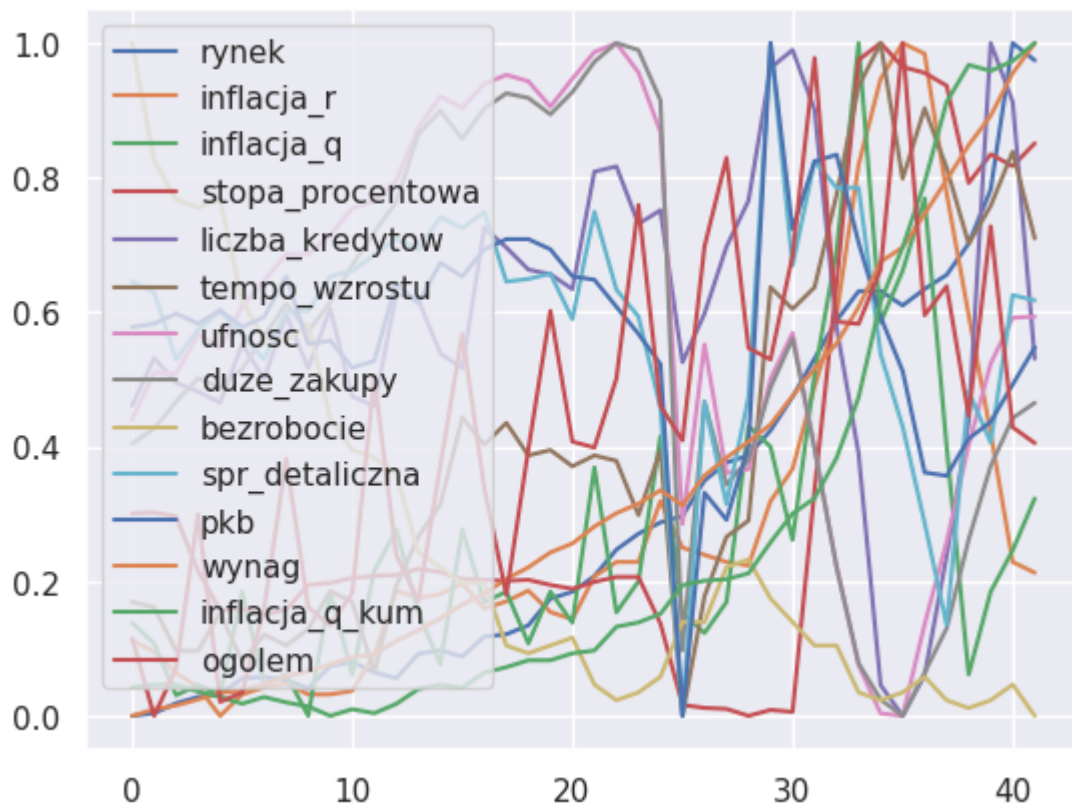
Out[99]:

	rynek	inflacja_r	inflacja_q	stopa_procentowa	liczba_kredytow	ten
2014-03-31	6193.211456	100.6	100.2	0.0446	41942.0	
2014-06-30	6227.984614	100.3	100.0	0.0447	45499.0	
2014-09-30	6357.260998	99.7	99.5	0.0443	43653.0	
2014-12-31	6438.874574	99.3	99.6	0.0388	42993.0	
2015-03-31	6473.616559	98.5	99.5	0.0348	42169.0	

In [100... `df.plot()`Out[100... `<Axes: >`

```
In [101... min_max_scaler = sklearn.preprocessing.MinMaxScaler()
ndf_arr = min_max_scaler.fit_transform(df)
ndf = pd.DataFrame(data=ndf_arr, columns=df.columns)
ndf.plot()
```

Out[101... `<Axes: >`



```
In [102... corr = df.corr()
corr.style.background_gradient(cmap='coolwarm').format(precision=3)
```

```
Out[102...
           rynek  inflacja_r  inflacja_q  stopa_procentowa  liczba_kredytow  te
rynek      1.000      0.697      0.573              0.735             -0.020
inflacja_r  0.697      1.000      0.813              0.772             -0.477
inflacja_q  0.573      0.813      1.000              0.588             -0.286
stopa_procentowa  0.735      0.772      0.588              1.000             -0.513
liczba_kredytow -0.020     -0.477     -0.286             -0.513              1.000
tempo_wzrostu  0.850      0.844      0.737              0.806             -0.152
ufnosc       -0.544     -0.708     -0.591             -0.590              0.521
duze_zakupy  -0.556     -0.648     -0.497             -0.602              0.514
bezrobocie   -0.680     -0.585     -0.501             -0.375             -0.133
spr_detaliczna -0.216     -0.259      0.049             -0.195              0.377
pkb         -0.177     -0.099      0.137             -0.105              0.302
wynag        0.988      0.730      0.557              0.757             -0.055
inflacja_q_kum  0.942      0.702      0.456              0.841             -0.188
ogolem       0.616      0.651      0.543              0.317              0.037
```

```
In [105... df["rpk"] = 0.0
df.loc[dt.datetime(year=2023, month=7, day=1): dt.datetime(year=2024, mon
df.tail(12)
```

Out[105...

	rynek	inflacja_r	inflacja_q	stopa_procentowa	liczba_kredytow	te
2021-09-30	10341.384137	105.4	101.0	0.0246	68353.0	
2021-12-31	10841.440443	107.7	102.6	0.0470	63918.0	
2022-03-31	11333.612554	109.7	103.8	0.0678	47767.0	
2022-06-30	11716.052702	113.9	105.8	0.0903	38398.0	
2022-09-30	11719.917172	116.3	103.1	0.0920	21218.0	
2022-12-31	11530.415342	117.3	103.6	0.0896	18932.0	
2023-03-31	11744.826140	117.0	104.3	0.0890	21968.0	
2023-06-30	11928.471147	113.1	101.9	0.0877	30798.0	
2023-09-30	12335.591393	109.7	99.7	0.0779	40749.0	
2023-12-31	13037.926475	106.4	100.5	0.0808	68860.0	
2024-03-31	14942.844839	102.8	100.9	0.0796	64504.0	
2024-06-30	14718.380953	102.5	101.4	0.0819	45434.0	

In [106...

```
df.to_csv("../data/nts.csv")
```