Hello everyone,

Recently I discovered the seasonal_decompose function of the statsmodel package and played a little bit around with it on the numerai data.

What I first did is to calculate for every feature the target correlation in every era. For a given feature, you can then use the seasonal decompose to seperate the time series data into a trend, a seasonal, and a residual noise part, to generate plots like these:

[

grafik

1268×1214 181 KB

](https://forum.numer.ai/uploads/default/original/2X/3/3a16e1392202d7d25587739825a85b6e5c82475d.png)

This is just an example for a random picked feature for a a priori period of 52 weeks (1 year).

Obviously, because the period is always set a priori, you will always get a seasonal component with said period.

So what I did next is to try to get a feeling of what periods might actually be "real" is to calculate the standard deviation of the seasonal part, and the standard deviation of the residual part, and calculate their ratio, to get a "signal/noise" ratio (ignoring the trend part for now), and do this for a range of periods from 4 weeks to 60 weeks, and also for every feature. Basically some sort of Fourier Transformation.

This is the result:

[

grafik

1167×589 25.9 KB

](https://forum.numer.ai/uploads/default/original/2X/1/10b2ce653dfb3a2f653eb5291ab183ad0f03b320.png)

The red horizontal lines within the boxes represent the median of the signal /noise ratio over all features for the specific assumed a priori period. As you can see there are peaks for the periods 13,26,39,52 weeks. Please also note the scale of the y Axis, indicating remarkably low signal to noise ratios.

It seems that most of the features have a quarterly seasonal pattern.