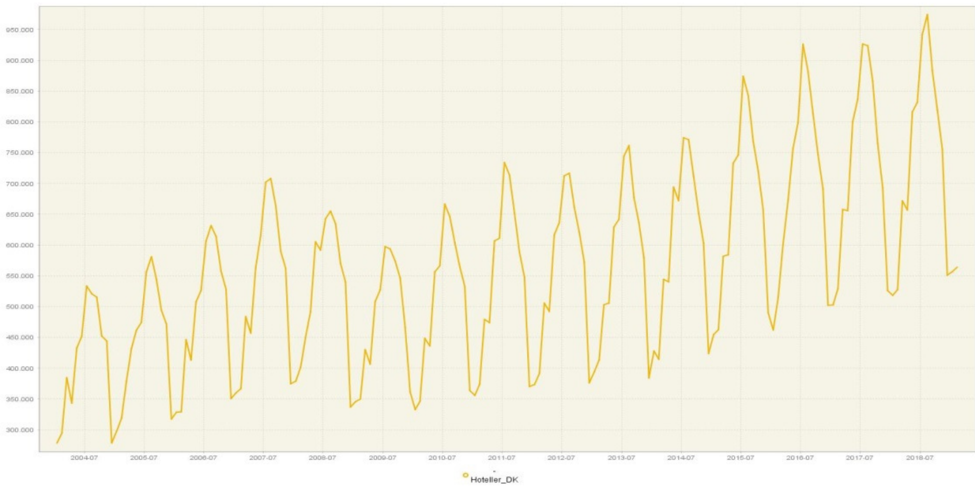
DANMARKS
STATISTIK

31.08.2020
PSL, SHE og GPM, (@dst.dk)
Methodology and Analysis

Introduction to Seasonal Adjustment

Why Seasonal Adjustment?

A time series with seasonal variation consists in part of various repeated patterns and in part of random fluctuations that correspond to short term, irregular and seasonal fluctuations (typically less than year), medium term business cycles (typically 3 - 5 years) and long cycles (more than 5 years).
Seasonal adjustment is performed in order to study the movements in a time series that are not attributable to explained by seasonal fluctuations or **seasonal effects**.
The figure below shows a time series with clear seasonal variation, i.e. peaks and troughs at the same time each year.



What is Seasonal Adjustment?

Seasonal Effects

Financial time series with monthly or quarterly frequencies are often affected by events that occur at the same time every year. This is called **seasonal effects**. One example is the rise in household shopping each year around Christmas with the result – i.e. the seasonal effect – that retail shopping increases each year from October to November and from November to December, while it decreases from December to January the following year. Seasonal adjustment aims at removing seasonal effects from a time series.

Calendar Effects

In addition, time series can be affected by the numbers of each weekday in any given month. Productivity will be higher in a month with a high number of working-days compared to a month with many Saturdays, Sundays and public holidays. These effects are called **calendar effects**. They occur in time series for which it matters which type of days (working-days or weekend days) there are four or five of in any given month. Whether Easter falls in March or April (or in the first or second quarter) is another example of a calendar effect.

Pre-adjustment

In order to compare consecutive months, a time series can be adjusted for calendar effects (e.g. standardizing the amount of working-days in each month). This is called calendar adjustment and is part of the **pre-adjustment** preparing the time series for the actual seasonal adjustment.

The Process of Seasonal Adjustment

The whole process of seasonal adjustment consists of
- Pre-adjustment (including calendar adjustment)
- Seasonal adjustment

And will result in a **seasonally adjusted time series** for which consecutive months or quarters are more comparable because calendar effects and seasonal variation has been removed.

Seasonal adjustment in Statistics Denmark is usually carried out using the X-13-ARIMA method, which is why this paper is based on this method.

Pre-adjustment

The purpose of pre-adjustment is removing the variation in the raw data (the raw or un-adjusted time series) caused by calendar effects (working day effects, trading day effects and Easter effects) as well as **outliers**.

Adjusting for Working-day Effects

The number of working days varies from month to month, and this can have large impacts on financial time series. The number of working days (Monday through Friday) in any given month varies between 18 and 23 within each year. If left un-adjusted, this can lead to spurious conclusions about changes in e.g. productivity. Adjusting for this effect is called **working-day adjustment**.

Adjusting for Trading Day Effects

The intensity of shopping varies with the days of the week. More people tend to do their shopping on Saturdays than on Mondays. This means that for some time series for which trading days matter, more retail shopping will occur in months with five Saturdays than in months with five Mondays. Adjusting for this is called **trading day adjustment**.

Adjusting for Easter Effects

Apart from working-days and trading days, moving holidays – especially Easter – can affect economic activity. Easter may fall in March, in April or in part in both months, or in the first, second or partly in both quarters for quarterly time series. Adjusting for this effect is called **Easter adjustment** or **moving holiday adjustment**.

Outliers

During pre-adjustment the raw time series is adjusted for extreme values or outliers. The most important types of outliers are:

- **Additive outlier** is an extreme value that is only observed once (a single month or quarter), and which is not observed for adjacent time-points (before or after). An example is a short strike or lock-out.
- **Level shifts** are caused by events that affect the raw time series permanently so that all later observations are at a higher or lower level. An example could be car sales affected by changes in the registration tax.
- **Transitory changes** are events that affect the raw time series at one point in time, while the effect gradually disappears afterwards. An example is an increase in coffee prices, which would cause a sharp decrease in the amount of coffee sold, although the sale of coffee would most likely stabilize at the previous level after some time.

Calculations of endpoint values are based on the latest observations as well as forecasts, which means that outliers (unpredictable events) will lead to increased uncertainty in calculating the last or latest values. This affects calculations even more when the last observations differ from the expected outcomes, meaning that the forecasted values may not accurately reflect our expectations for the future developments of the time series. When the (inaccurately) forecasted values are replaced by actual, observed values, there's a risk of large revisions to published data.

Pre-adjustment Model

Pre-adjustment is performed using a regARIMA model, which is a regression model in which the residual (the pre-adjusted time series) is modelled by an ARIMA model. An ARIMA model is a time series model which catches the temporal dependence between adjacent periods. Mathematically it is written as:

$$X_t = \sum \beta_i x_{it} + O_t$$

where X_t denotes the raw (un-adjusted) time series at time t
$\sum \beta_i x_{it}$ denotes the pre-adjustment at time t
O_t denotes the pre-adjusted time series at time t

Model Specification

The variable x_{it} denotes the effect that is adjusted for, e.g. the number of trading days in month t, while the parameter β_i denotes the relative size of its impact on the time series. The pre-adjustment might comprise of several effects (indexed by i), and thus $\sum \beta_i x_{it}$ denotes the entire adjustment and O_t denotes the pre-adjusted time series at time t.

Temporary and Permanent Adjustment

Technically, the adjustments for calendar effects and outliers are performed in the same way, but there is an important difference: Adjustment for outliers is only **temporary**, and the extreme value is added back to the time series after the seasonal adjustment is carried out, since the outliers identify actual events (see next paragraph), whereas adjustment for calendar effects is **permanent**, and these effects remain removed from the final, seasonally adjusted time series.

Which Effects are Adjusted for?

Which adjustments are relevant for a particular statistics is determined when processing it, primarily based on a number of statistical tests. Calendar effects ought to have a financial interpretation as well as statistical significance, however.

If no calendrer effects or outliers are identified, no pre-adjustment is performed and the pre-adjusted time series remains identical to the raw (unadjusted) time series ($X_t = O_t$).

Seasonal Adjustment

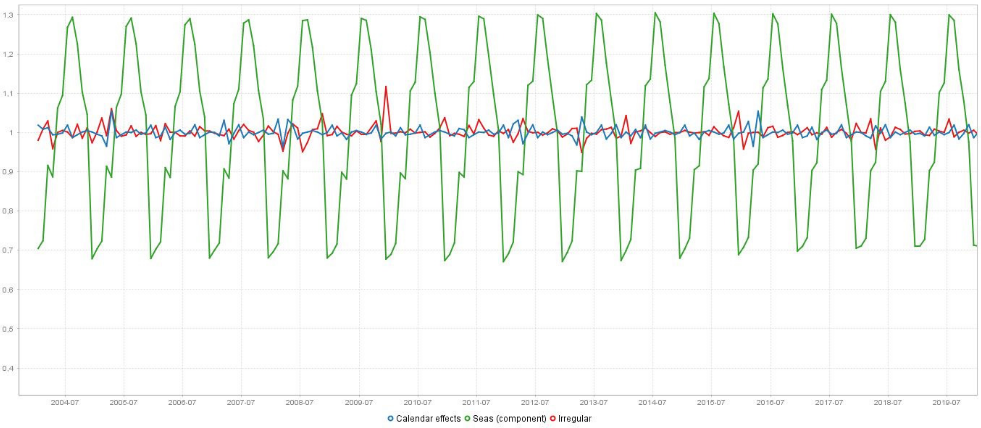
Decomposition

After pre-adjusting the time series, the seasonal adjustment is performed. It is performed on the pre-adjusted time series, which is decomposed into three components:

- **The trend** (T_t) denotes the long term movement of the time series. Any possible level shifts are added back to the trend in the last stages of the seasonal adjustment (see the paragraph on outliers above).
- **The seasonal component** (S_t) denotes the part of the time series variation that occurs within a one year period. These annual patterns are quite stable over time with regards to direction and magnitude. I.e. more or less stable monthly or quarterly variations.

- **The irregular component** (I_t) denotes the part of the raw time series that is neither contained by the trend nor the seasonal component (or has been removed during pre-adjustment). It is often referred to as random or unexplained variation. In the last stages of the seasonal adjustment process, outliers (except for level shifts) are added back to this component.

The figure below shows the seasonal component (green), the irregular component (red) and the calendar effects (blue) of a financial time series. There is no clear or repeated pattern in the irregular component, whereas the seasonal component shows a very clear, repeated pattern.



Choice of Decomposition Model

Normally, it is assumed that a time series can be written either as sum or as a product of its three components:

$$O_t = T_t + S_t + I_t \qquad \text{additive model}$$
$$O_t = T_t \cdot S_t \cdot I_t \qquad \text{multiplicative model}$$

where O_t denotes the pre-adjusted time series at time t (or just the raw time series in case no pre-adjustment was necessary) . In case the magnitude of the seasonal component is proportional to the level of the time series, a multiplicative model is preferable. This is tested for automatically by the seasonal adjustment software.

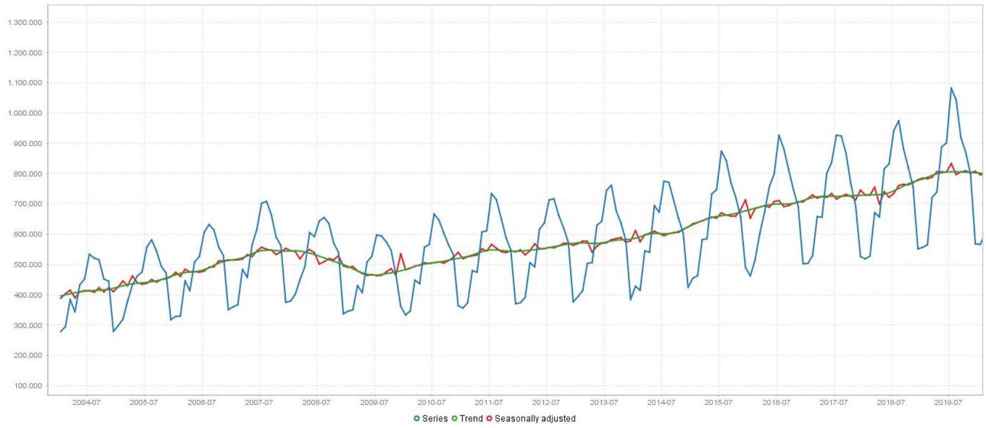
Technically, decomposition is carried out as follows:

- The pre-adjusted time series is modelled by an ARIMA model as part of the pre-adjustment process
- The pre-adjusted time series is forecasted by the ARIMA model, i.e. an indirect prognosis
- The three components are determined in an iterative process using various types of moving average filters applied to the forecasted, pre-adjusted time series

When the model has been determined and the three components have been identified, the **seasonally adjusted time series** (A_t) is identified by removing the seasonal component from the pre-adjusted time series:

$$A_t = O_t - S_t = T_t + I_t \qquad \text{if the model is additive}$$
$$A_t = O_t / S_t = T_t \cdot I_t \qquad \text{if the model is multiplicative}$$

The figure below shows the actual time series (blue), the seasonally adjusted time series (red) and the trend (green) of the financial time series shown earlier. As the seasonal effects are very clear in this case, the seasonal adjustment has smoothed out the time series quite a lot.



Consequences of Forecasting

Forecasting the time series is necessary as symmetric moving averages are used in the decomposition process. Forecasting correctly determines the quality of the seasonal adjustment of the last observations. During volatile periods in which large changes in economy or other societal factors occur, such as is presently the case with the COVID-19 pandemic, the forecasting is more uncertain. This uncertainty carries over to the seasonal adjustment and should be taken into account when interpreting the seasonally adjusted data (cf. the paragraph on outliers).

Other aspects of seasonal adjustment

Seasonal adjustment of time series which are sums of other time series

If a time series is a sum of other time series (sub-series), it can be seasonally adjusted in two ways: **Direct seasonal adjustment** refers to performing seasonal adjustment on the sum series directly (separately from the sub-series, if seasonal adjustment of those is needed). In this case, the sum of seasonally adjusted sub-series is usually not equal to the directly seasonally adjusted sum series. **Indirect seasonal adjustment refers** to seasonally adjusting the sub-series separately and then adding the seasonally adjusted sub series to achieve the seasonally adjusted sum series. In this case, the sum of seasonally adjusted sub series is equal to the seasonally adjusted sum series.

Benchmarking

Seasonal adjustment with X-13-ARIMA does not guarantee that the sum of several observations of the raw time series corresponds exactly to sums of the same observations of the seasonally adjusted time series. However, the difference between those sums over a one year period will normally be negligible. Benchmarking refers to adjusting the seasonally adjusted time series so that the sums of the seasonally adjusted and raw time series observations correspond exactly over a calendar year. The calendar adjusted time series can be benchmarked to the seasonally adjusted time series as an alternative to benchmarking it to the raw time series. However, very few time series are benchmarked. Benchmarking is generally not recommended.

Revisions

Seasonally adjusted data is usually changed each time new, raw data is added to the model. When new observations are added to the raw data of a monthly or quarterly time series, seasonal adjustment is technically carried out on a new data set, which can necessitate changing previously published data. In some cases, this can lead to changes in growth rates and tendencies observed within the last few months. In Statistics Denmark the recommendation is to publish revisions to data from the current calendar year and three years back in time.

Quality of seasonal adjustment

The first question to answer in connection with seasonal adjustment is whether seasonal patterns occur in a time series. This is determined by a series of tests as well as visual inspection. A large number of time series do not display clear, seasonal variation, and in some cases determining whether there is seasonal variation at all can be difficult.

If a time series is found not to contain seasonal variation, Statistics Denmark's policy is **not** to seasonally adjust the time series, but just to publish the raw and possible the pre-adjusted time series. This corresponds to the recommendations of the "guidelines" of EuroStat, containing general directions for seasonal adjustment such as pre-adjustment, direct/indirect seasonal adjustment, benchmarking, revisions etc. This document can be downloaded from EuroStat's website (https://ec.europa.eu/eurostat/cros/content/ess-guidelines-sa-2015-edition_en).

If seasonal variation occurs in the time series, it can be seasonally adjusted, and the X-13-ARIMA algorithm comes with a number of diagnostic tools displaying the quality of the seasonal adjustment. These are used to determine whether to accept the seasonal adjustment or to look for other seasonal adjustment models. If no seasonal adjustment model of good quality can be identified, seasonal adjustment of the time series in question should not be carried out,

Software

In Statistics Denmark, seasonal adjustment is almost exclusively performed using the seasonal adjustment software JDemetra+ (presently version 2.2.2). This agrees with EuroStat's official recommendation for seasonal adjustment software. JDemetra+ can be downloaded for free from EuroStat's website (https://ec.europa.eu/eurostat/cros/content/software-jdemetra_en).