Time Series Decomposition and Transforms

Maria Eduarda Silva

School of Economics, University of Porto

LMU, July 2 2022

Transformations

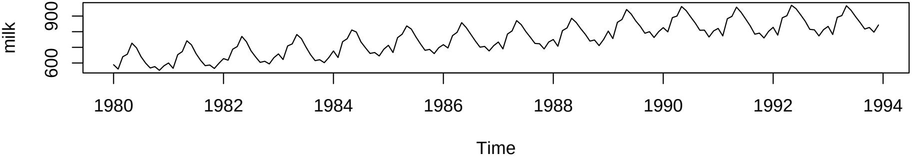
Trnasformation Are used to adjust historical data and often get a simple time series which are easier to model and lead to more accurate forecasts.

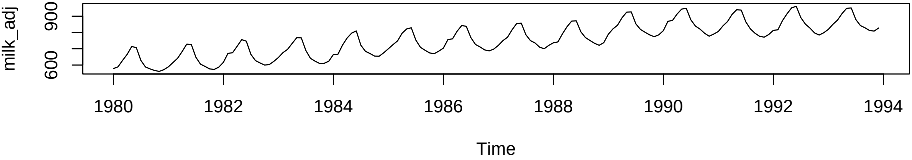
- Calendar adjustments
- Length of the month: since the different months of the year have different number of days and also because of leap year, one may adjust to the length of the month as follows:

$$W_t = \frac{X_t \times 365.25/12}{\text{no days in month } t}$$

- Number of working days
- Mathematical Transformations: to linearize the data and stabilize the variance- Box-Cox transformations (only positive data)

Adjust for length of month: Monthly Milk Prod per cow





Variance Stabilization - Box-Cox Transformations

To stabilize the variability over the series we use the Box-Cox transformations. A particular case is to log the data. Note that the multiplicative representation

$$y_t = T_t S_t R_t$$

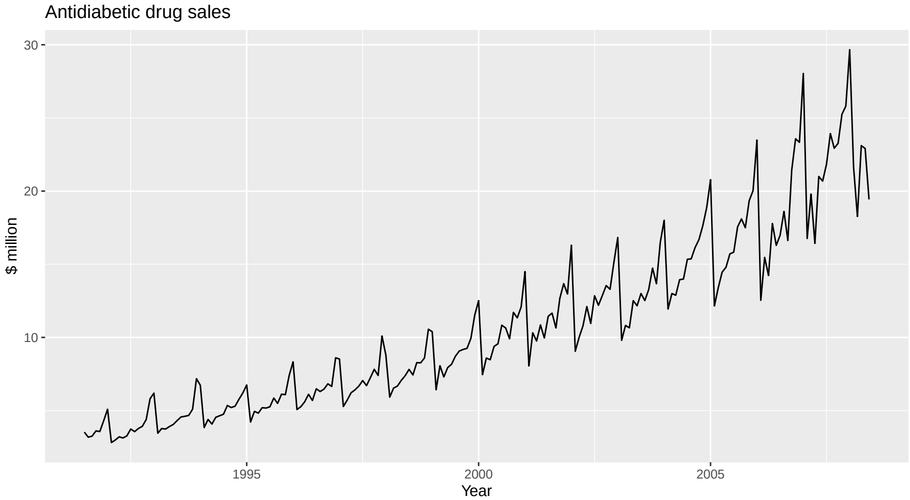
becomes

$$log(y_t) = log(T_t) + log(S_t) + log(R_t)$$

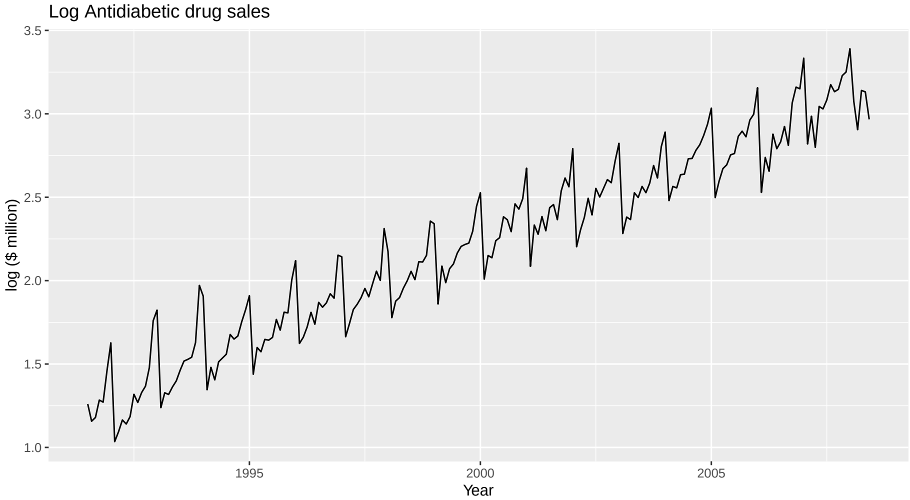
$$U_t = \begin{cases} \frac{X_t^{\lambda}-1}{\lambda} & if \quad \lambda \neq 0 \\ \log X_t & if \quad \lambda = 0 \end{cases}$$

These transforms are also used to improve approximation to normality (Gaussian distribution).

Antidiabetic drug sales, a10



Box-Cox (log) transform of a10

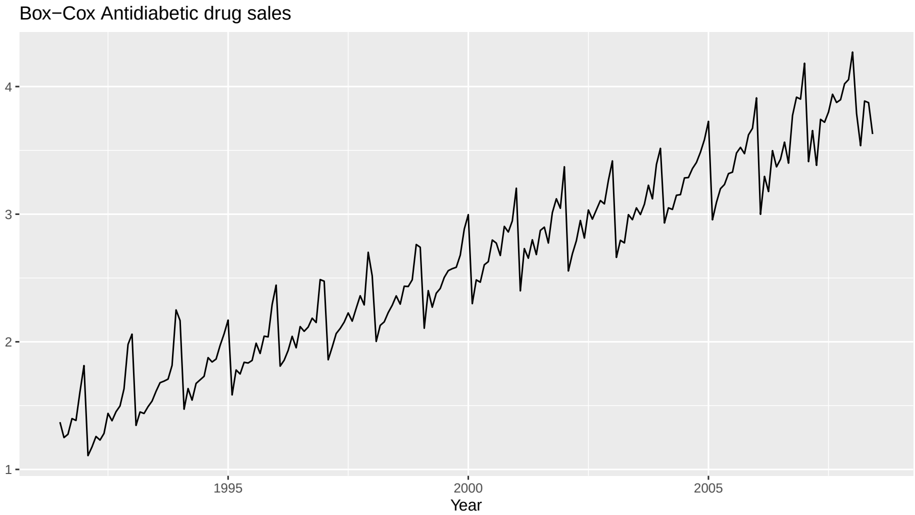


Now the variance is stable and the trend looks linear. Produce the season plots for the log data and comment.

Finding the appropriate transform

To find the best Box-Cox transform - find λ that minimizes the variance.

```
## [1] 0.1313326
```



Box-Cox Transform

- Log is the most used transformation
- When you model and forecast the transformed data then you need to back transform.
- Avoiding transforms is best

Decomposition and filtering

- **Depending on the purpose of our study, our interest may be**
  - the trend
  - the seasonal component
  - the random component
  - all together
- **Approaches**
  - Decomposition methods - estimate the trend /seasonal component
    - with deterministic functions
    - smooth functions- moving averages
    - STL
  - Filtering methods - Difference and Seasonal Difference

Time Series Decomposition

- Trend-Cycle- aperiodic changes in level over time +Seasonal - (almost) periodic changes in level due to seasonal factors
- **Additive decomposition

$$y_t = T_t + S_t + R_t$$

  - y_t data at time t
  - T_t trend-cycle component at time t
  - S_t seasonal component at time t
  - R_t remainder component at time t

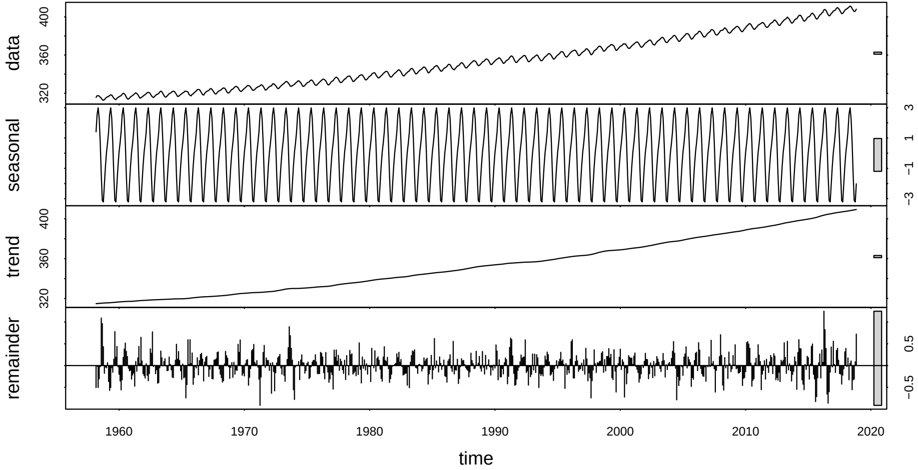
Time Series Decomposition by Loess - STL

- The Seasonal Decomposition of Time Series by Loess is implemented in R in the stl() function
- Decomposes a time series into seasonal, trend and irregular components using loess
- Loess is a robust weighted regression smoothing method close to nearest neighbour regression
- The seasonal component is found by loess smoothing the seasonal sub-series (the series of all January values, . . .)
- If s.window = "periodic" smoothing is effectively replaced by taking the mean
- The seasonal values are removed, and the remainder smoothed to find the trend
- The overall level is removed from the seasonal component and added to the trend component. This process is iterated a few times
- The remainder component is the residuals from the seasonal plus trend fit.

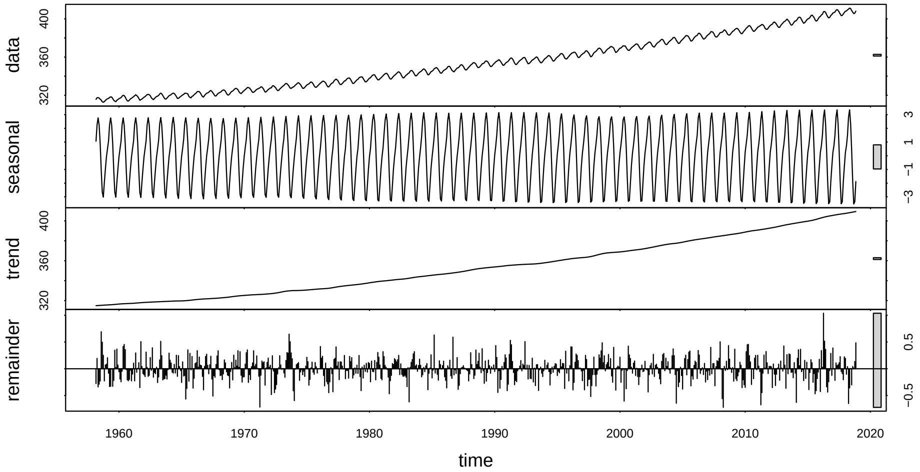
STL decomposition

- Versatile and Robust
- Seasonal component allowed to change overtime and rate of change controllled by the user
- Smoothness of trend-cycle controlled by the user
- Optionally robust to outliers
- No calendar adjustments
- Only additive
- You need to log the data if multiplicative decomposition is needed

STL decomposition of Cardox - fixed periodicity



STL decomposition of Cardox - not fixed periodicity

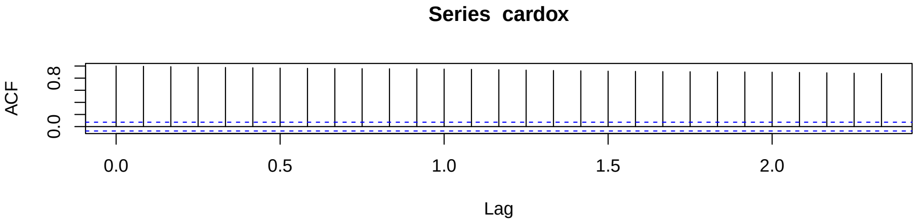


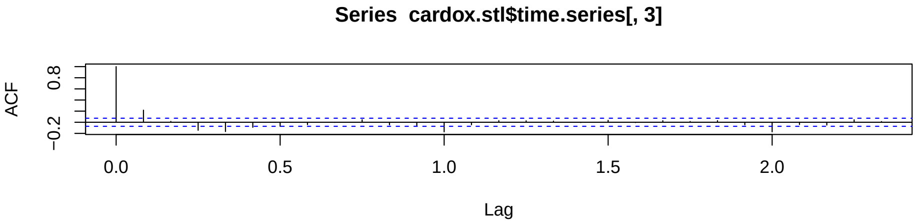
STL components

```
head(cardox.stlper$time.series)
```

```
##            seasonal    trend   remainder
## Mar 1958  1.4039091 314.8115 -0.50536160
## Apr 1958  2.5512431 314.9140 -0.01528735
## May 1958  2.9864460 315.0166 -0.50308202
## Jun 1958  2.2977101 315.1059 -0.30364134
## Jul 1958  0.6929085 315.1952 -0.02813505
## Aug 1958 -1.4385111 315.2765  1.09200191
```

Investigate the correlation behaviour of the remainder.

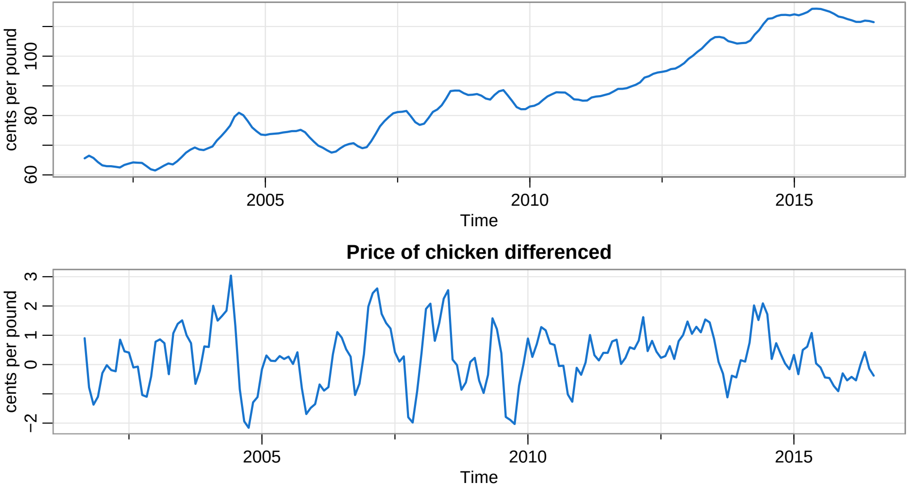




Some notation

- x_t denotes the observation at time t
- B is called the **backshift or lag** operator: $Bx_t = x_{t-1}$
- $B^{12}x_t = x_{t-12}$
- $\nabla = 1 - B$ the **difference operator**:
  $\nabla x_t = (1 - B)x_t = x_t - Bx_t = x_t - x_{t-1}$
- $$y_t = \nabla x_t = x_t - x_{t-1}$$
  represents the increments or change of the variable x on consecutive time points
- Applying the difference operator is said **Differencing** and can help stabilise the mean of a time series by removing changes in the level of a time series and therefore eliminating (or reducing) trend

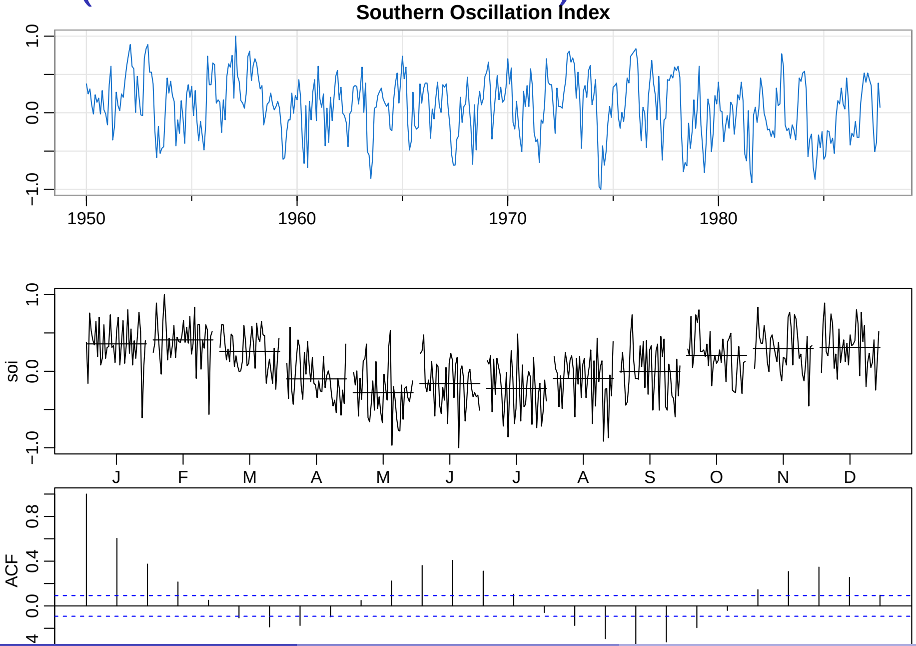
The difference operator applied to the Monthly Price of Chicken



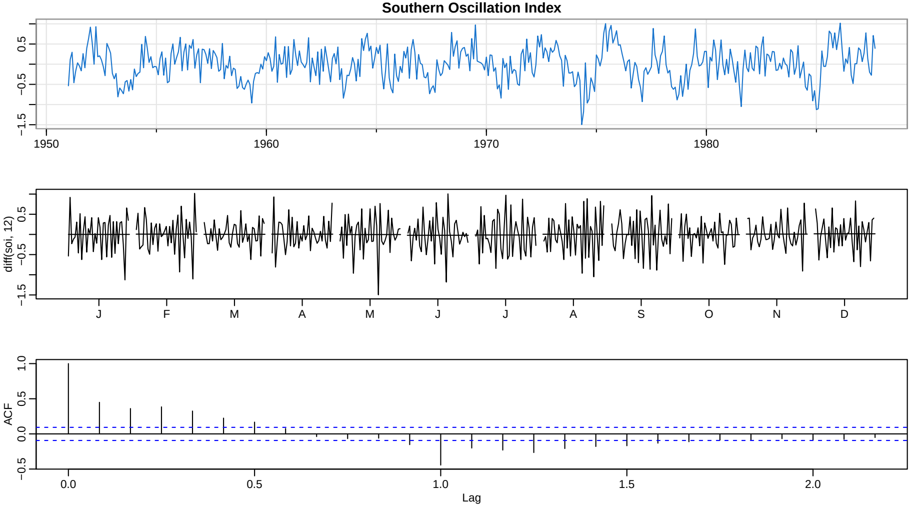
Seasonal difference operator

- Seasonal difference operator $\nabla_S = 1 - B^S$, where S is the seasonality
- $y_t = \nabla_S x_t = x_t - x_{t-S}$ represents the increments or change of the variable x over consecutive seasonal periods.
- If $S = 12$ then y_t represents the increments from one year to the next
- Applying the seasonal difference operator can help stabilise the mean of a time series by removing changes in the level due to seasonality

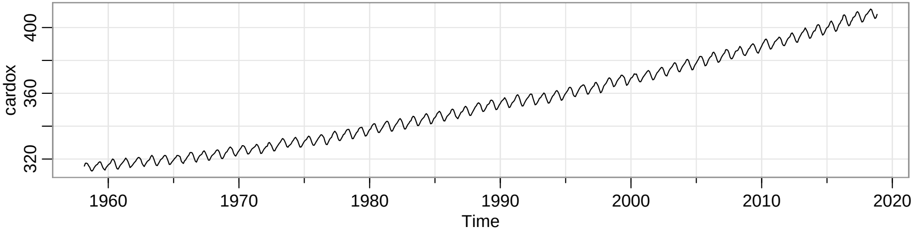
SOI (Southern Oscillation Index)

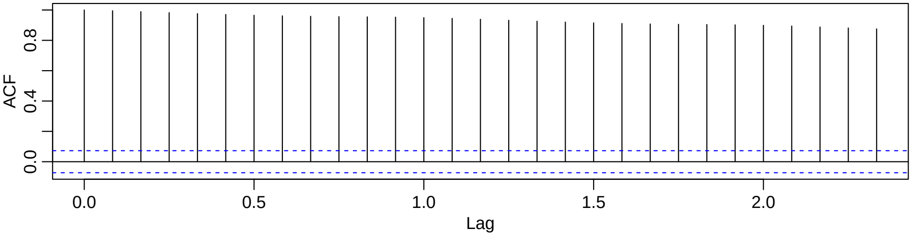


Seasonaly differenced SOI (Southern Oscillation Index)

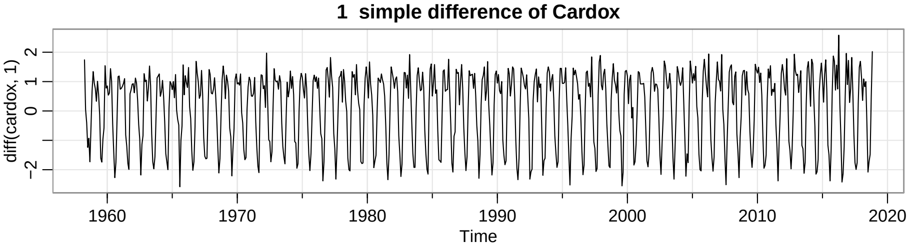


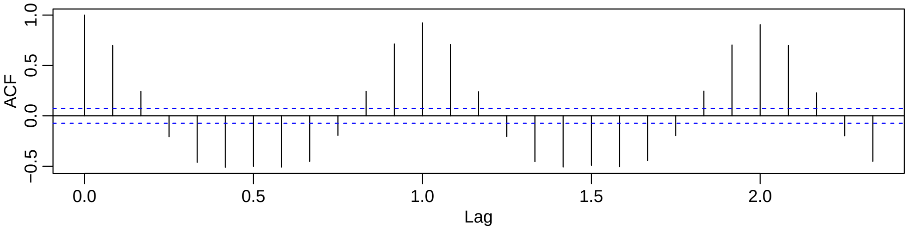
Differencing time series with trend and seasonality





Differencing time series with trend and seasonality - CARDOX

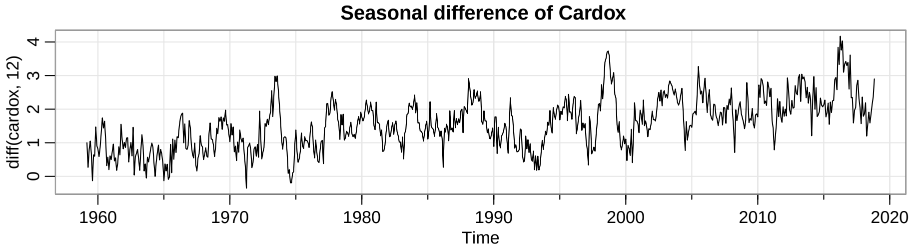


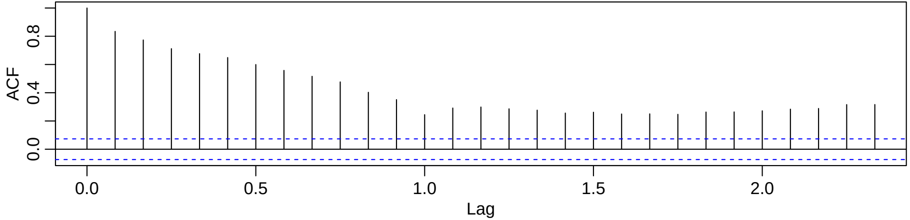


Differencing time series with trend and seasonality - CARDOX

Simple differences, changes in carbon dioxide from month to month, show a seasonal cycle

Differencing time series with trend and seasonality - CARDOX

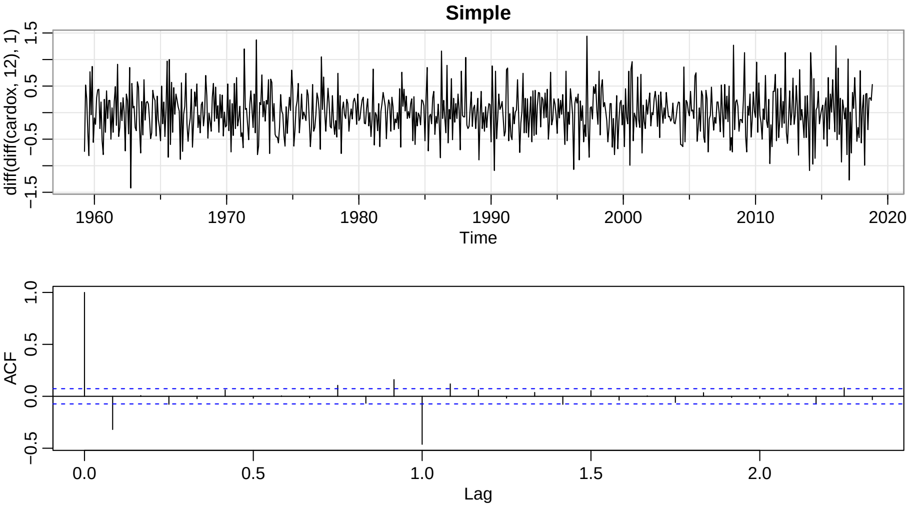




Differencing time series with trend and seasonality

Seasonal differences, changes in carbon dioxide from one month to the same month next year , show some trend

Differencing time series with trend and seasonality



Differencing time series with trend and seasonality

- When we apply both operators, simple and seasonal the filtered data data no longer presents trend or seasonality.
- Why is this important?