Time series
Short tutorial

ESTP training

12-14/10/2021

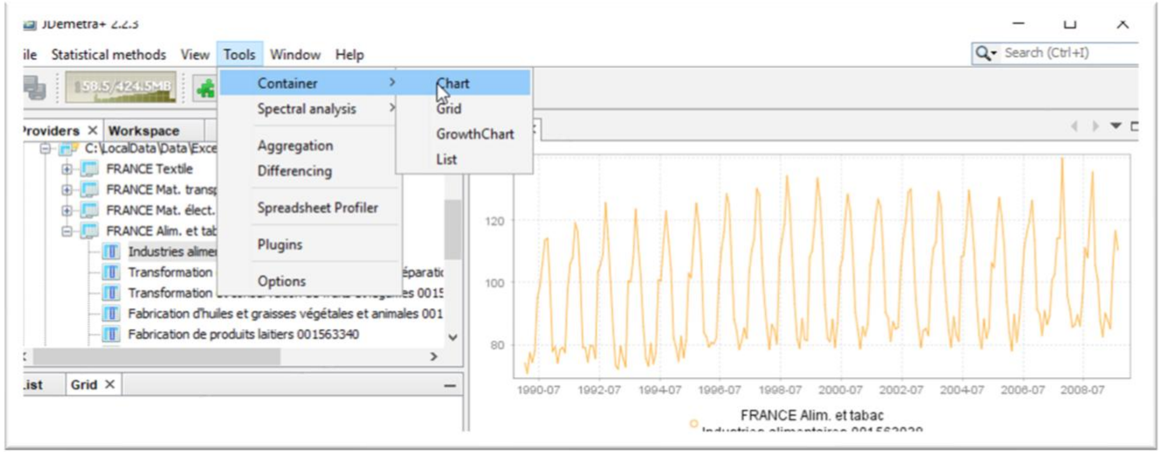
Outline

- “Regular” time series
- Auto-correlation function
- Periodogram
- Stationarity
- Arima model

Time series

- We only consider « regular » time series
  - Monthly, quarterly…[, yearly] time series
  - Annual frequency: 12, 6, 4, 3, 2, 1

- [JD+ 3.0 will deal with more general time series]



Basic tools

- Measurement of "regularities" in time series (x_t, $0 \leq t < n$)

- Auto-correlation function (ACF)

  - $acf(k) = \rho_k = \frac{\sum_{t=k}^{n-1}(x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=}^{n-1}(x_t - \bar{x})^2}$

  - If we assume that $\bar{x} = 0$, $\rho_k = \frac{\sum_{t=k}^{n-1} x_t x_{t-k}}{\sum_{t=0}^{n-1} {x_t}^2}$

Basic tools (cont.)

- Periodogram
  - Discrete Fourier's transform of the series (the series is expressed as a sum of sine and cosine functions)

  - $X_k = \sum_{t=0}^{n-1} x_t e^{-i\frac{2\pi}{n}kt} = \alpha \sum_{t=0}^{n-1} x_t (\cos\frac{2\pi}{n}kt - i\sin\frac{2\pi}{n}kt)$

  - $p_k = \alpha |X_k|^2, \qquad k = Fourier\ frequencies = \frac{2\pi}{n}j, 0 \leq j < n$

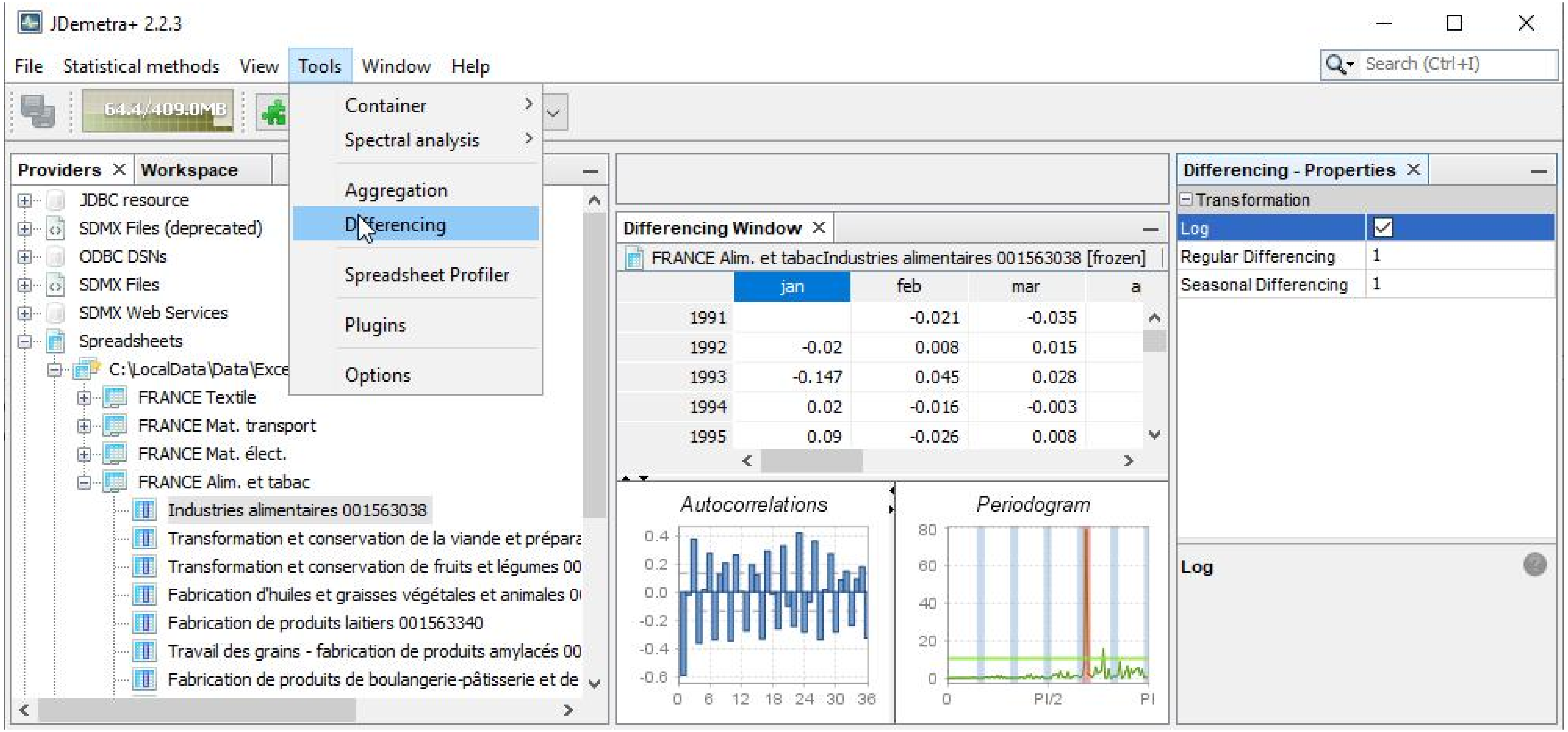
  - The periodogram gives the importance of each frequency in the series

Stationarity

- A series is said stationary at the second order, if its mean and variance do not depend on time and if the covariance between lagged series only depends on the difference between lags.
- Many statistical tools/algorithms only apply on stationary time series

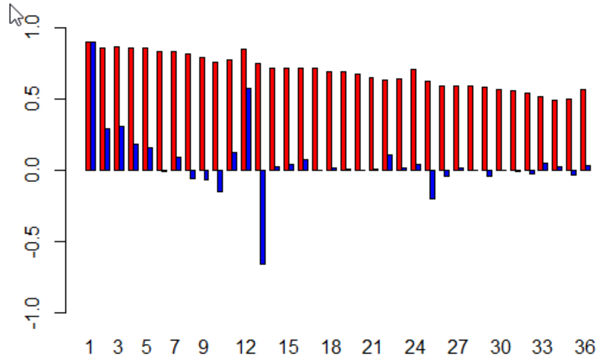
- Most economic time series a non-stationary
- (Simple) solutions
  - Log-transformation
  - Differencing
    - $y_t = x_t - x_{t-1}$
    - $y_t = x_t - x_{t-s}$

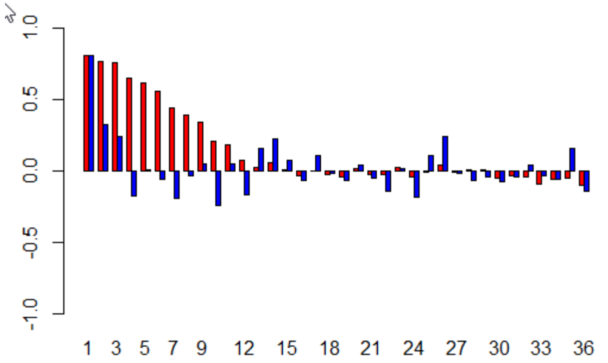
Stationary series

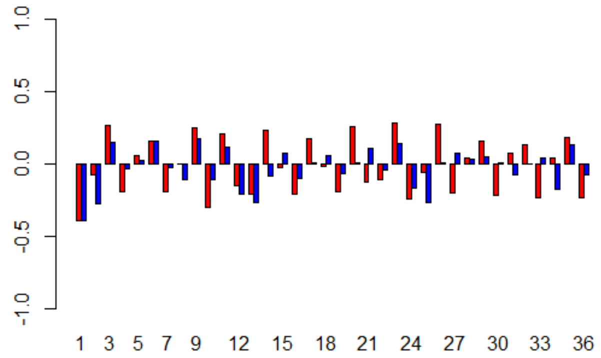


Stationary series in R

```
suppressPackageStartupMessages(library(rjd3modelling))

s<-log(retail$RetailSalesTotal)

plot(s)

ac<-rjd3toolkit::autocorrelations(s, T, n=36)
pac<-rjd3toolkit::autocorrelations.partial (s, T, n=36)

all<-cbind(ac, pac)
barplot(t(all), beside = T, col = c("red", "blue"), names.arg = c(1:36), ylim=c(-1,1))

ds=differences(s,12,F)
#plot(ds, type='l')

ac<-rjd3toolkit::autocorrelations(ds, T, n=36)
pac<-rjd3toolkit::autocorrelations.partial(ds, T, n=36)

all<-cbind(ac, pac)
barplot(t(all), beside = T, col = c("red", "blue"), names.arg = c(1:36), ylim=c(-1,1))

ds=differences(ds,1,F)
#plot(ds, type='l')

ac<-rjd3toolkit::autocorrelations(ds, T, n=36)
pac<-rjd3toolkit::autocorrelations.partial(ds, T, n=36)

all<-cbind(ac, pac)
barplot(t(all), beside = T, col = c("red", "blue"), names.arg = c(1:36), ylim=c(-1,1))

spec<-spec.pgram(ds)
```







Arima models

- Auto-projective Models:
  - $x_t = f(x_{t-1}, x_{t-2}, \cdots, \varepsilon_t)$
- ARIMA models (Box-Jenkins)
  - Good approximation of f
  - Flexible, parsimonious

Arima models (definitions)

- Backward operator
  - $B^k x_t = x_{t-k}$
  - $(1 - B^k)x_t = x_t - x_{t-k}$
- Auto-regressive model
  - $(1 + \varphi_1 B + \cdots + \varphi_p B^p)x_t = \Phi(B)x_t = \varepsilon_t$
  - $x_t = \varepsilon_t - \varphi_1 x_{t-1} - \cdots - \varphi_p x_{t-p}$ (≠R, X13)
- Moving average model
  - $x_t = \Theta(B)x_t = (1 + \theta_1 B + \cdots + \theta_q B^q)\varepsilon_t$
  - $x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$

Arima models (definitions)

- ARMA (p, q)
  - $\Phi(B)x_t = \Theta(B)\varepsilon_t$
- SARIMA$(p, d, q)(bp, bd, bq)_s$
  - $(1-B)^d(1-B^s)^{bd}\Phi(B)\Phi_s(B^s)x_t = \Theta(B)\Theta_s(B^s)\varepsilon_t$

- Airline model
  - $(1-B)(1-B^s)x_t = (1+\theta B)(1+\theta_s B^s)\ \varepsilon_t$

  - Interpretation of $x_t = x_{t-12} + \theta_s\varepsilon_{t-12} + \varepsilon_t$: *same as previous year, partially corrected for the "committed error"*

Arima model properties

- Autocovariance function(ACF): only defined for stationary models
  - $x_t = \sum_{k=0}^{\infty} \psi_{t-k} \varepsilon_{t-k}$ $\quad$ ψ *generated by* $\frac{\Theta(B)}{\Phi(B)}$
  - $acgf = \frac{\Theta(B)}{\Phi(B)} \frac{\Theta(F)}{\Phi(F)}$
- (Pseudo-)spectrum = Fourier's transform of the ACF (with extension to non-stationary models)
  - Counterpart of the periodogram
  - Contribution of each frequency to the variance of the series

Final remarks

- Importance of the representation in the frequency domain for series decomposition
  - Trend ≈ low frequencies
  - Seasonal ≈ seasonal frequencies (related to the periodicity)
  - Irregular ≈ high-frequencies

- Especially true for SEATS