

# Time series

## Short tutorial

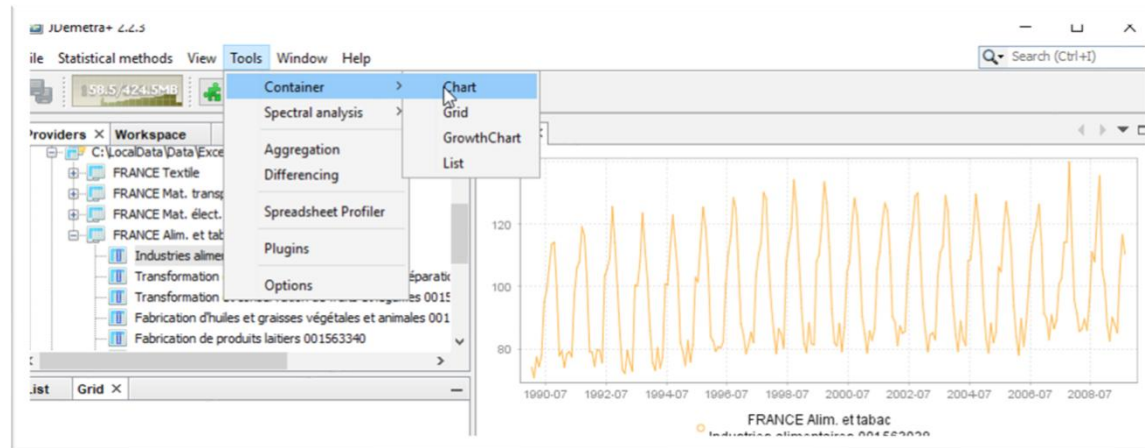
**ESTP training**

# 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)

- $acf(k) = \rho_k = \frac{\sum_{t=k}^{n-1} (x_t - \bar{x})(x_{t-k} - \bar{x})}{\sum_{t=0}^{n-1} (x_t - \bar{x})^2}$

- 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)

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

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

- 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

The screenshot displays the JDemetra+ 2.2.3 software interface. The 'Tools' menu is open, showing options like Container, Spectral analysis, Aggregation, Differencing (highlighted), Spreadsheet Profiler, Plugins, and Options. The 'Differencing Window' is open, showing a table of data for 'FRANCE Alim. et tabacIndustries alimentaires 001563038 [frozen]'. The table has columns for 'jan', 'feb', 'mar', and 'a' (annual), and rows for years 1991 to 1995. Below the table are two plots: 'Autocorrelations' and 'Periodogram'. The 'Differencing - Properties' window is also open, showing the 'Transformation' tab with 'Log' checked, 'Regular Differencing' set to 1, and 'Seasonal Differencing' set to 1.

**Differencing Window**

	jan	feb	mar	a
1991		-0.021	-0.035	
1992	-0.02	0.008	0.015	
1993	-0.147	0.045	0.028	
1994	0.02	-0.016	-0.003	
1995	0.09	-0.026	0.008	

**Autocorrelations**

**Periodogram**

**Differencing - Properties**

Transformation

Log ☒

Regular Differencing 1

Seasonal Differencing 1

# 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)

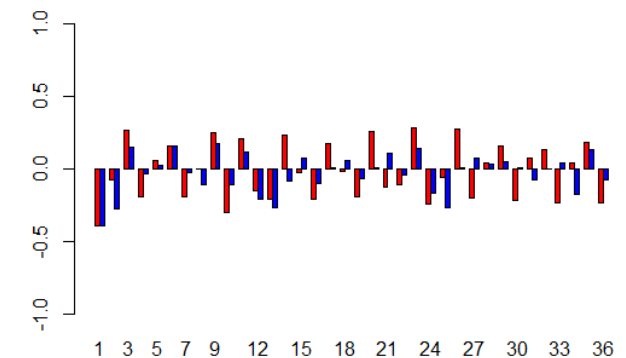
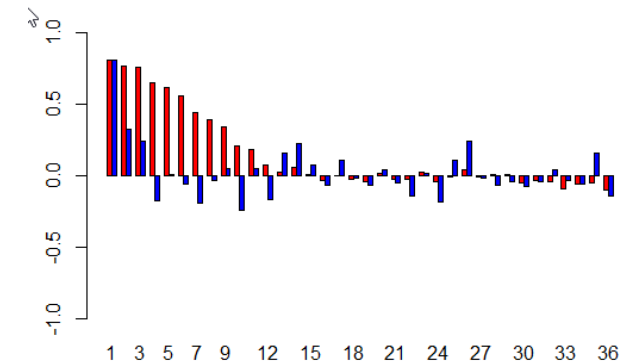
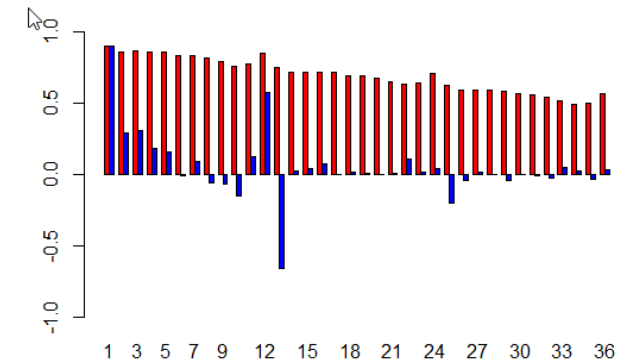
all<-cbind(ac, pac)
barplot(t(all), beside = T, col = c("red", "blue"), names.arg = c(1:36), ylim=c(-1,1))

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.ngram(ds)
```





# Arima models

- Auto-projective Models:
  - $x_t = f(x_{t-1}, x_{t-2}, \dots, \varepsilon_t)$
- 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

- $(1 + \varphi_1 B + \dots + \varphi_p B^p)x_t = \Phi(B)x_t = \varepsilon_t$
- $x_t = \varepsilon_t - \varphi_1 x_{t-1} - \dots - \varphi_p x_{t-p}$

(≠R, X13)

- Moving average model

- $x_t = \Theta(B)x_t = (1 + \theta_1 B + \dots + \theta_q B^q)\varepsilon_t$
- $x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}$

# 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
  - $x_t = \sum_{k=0}^{\infty} \psi_{t-k} \varepsilon_{t-k}$        $\psi$  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  $\approx$  low frequencies
  - Seasonal  $\approx$  seasonal frequencies (related to the periodicity)
  - Irregular  $\approx$  high-frequencies
- Especially true for SEATS