Linear Time Series Assignment
ARIMA modelling of a time series

Tony LAUZE   Tien-Thinh TRAN-THUONG

2024-05-20

Part I : the data

1 - What does the chosen series represent ? (sector, potential data processing, logarithmic transformation, etc.)

We have chosen to work on the *Industrie des eaux minérales et autres eaux embouteillées et des boissons rafraîchissantes* over time. The data used is called the index of production CVS-CJO (**correction des variations saisonnières (CVS) et des jours ouvrables (CJO)**).

This data is monthly and goes from January 1990 to February 2024, it has been downloaded from the INSEE website.
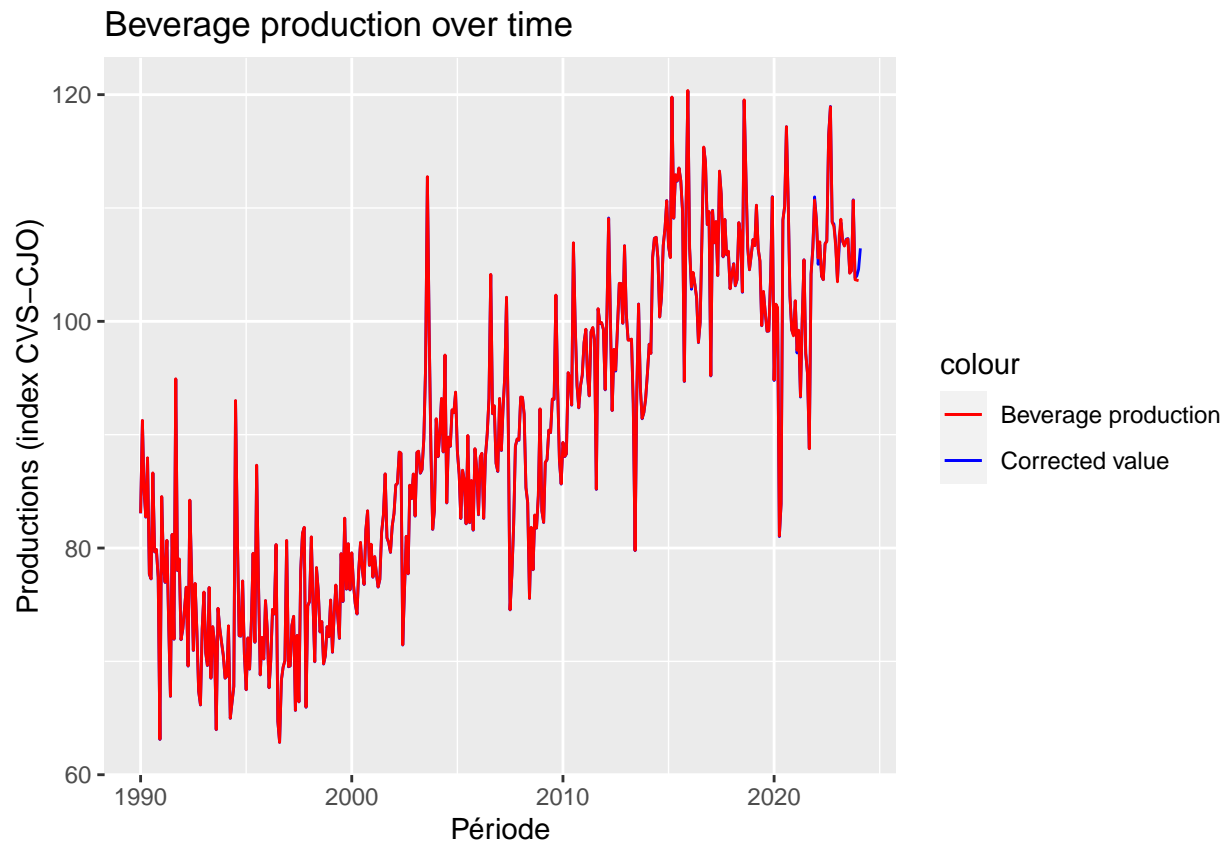

The data is not stationary, so we will have to correct it. We will use the corrected data for the rest of the analysis. It may have an up trend from 1995 to 2020.

2 - Transform the series to make it stationary if necessary (differentiate it, correct the deterministic trend, etc.). Thoroughly justify your choices.

Differentiated beverage production over time
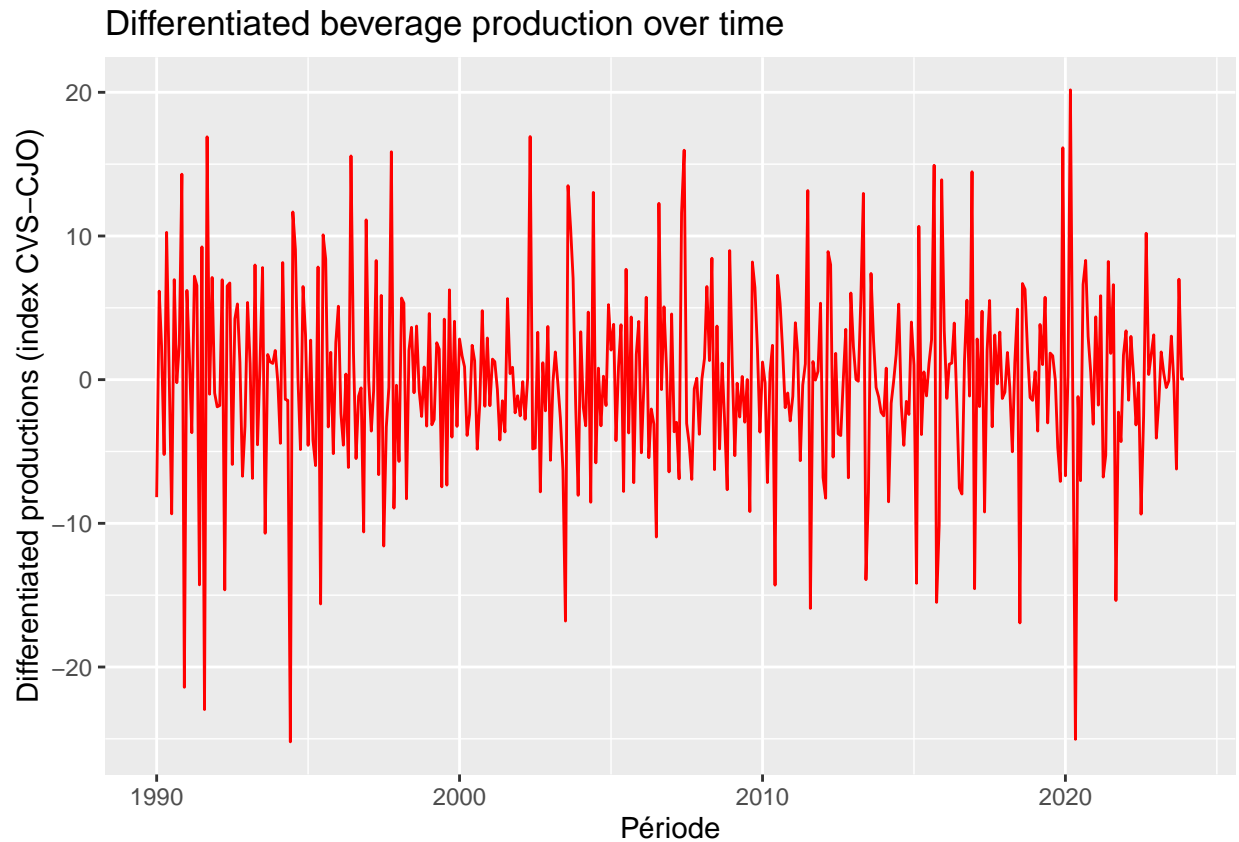
We have done a differentiation, there is no longer any trend and the data seems to be around 0. We will plot ACF and PACF to check if the data is stationary.

Series df$value_corr

Series df$value_corr
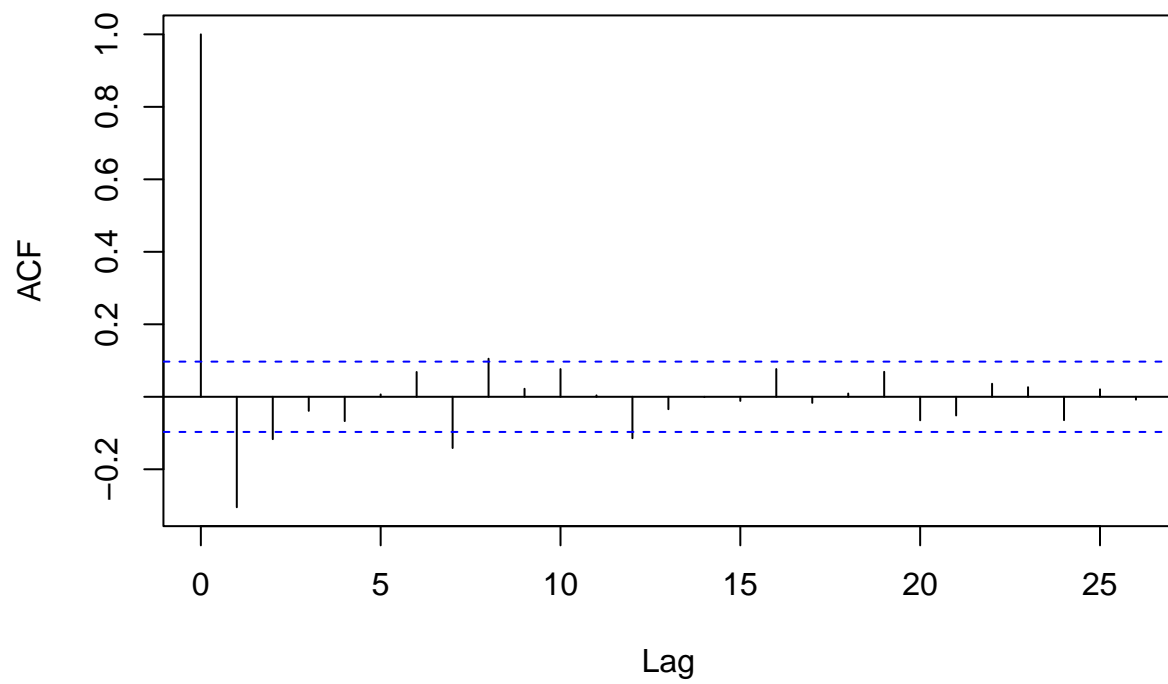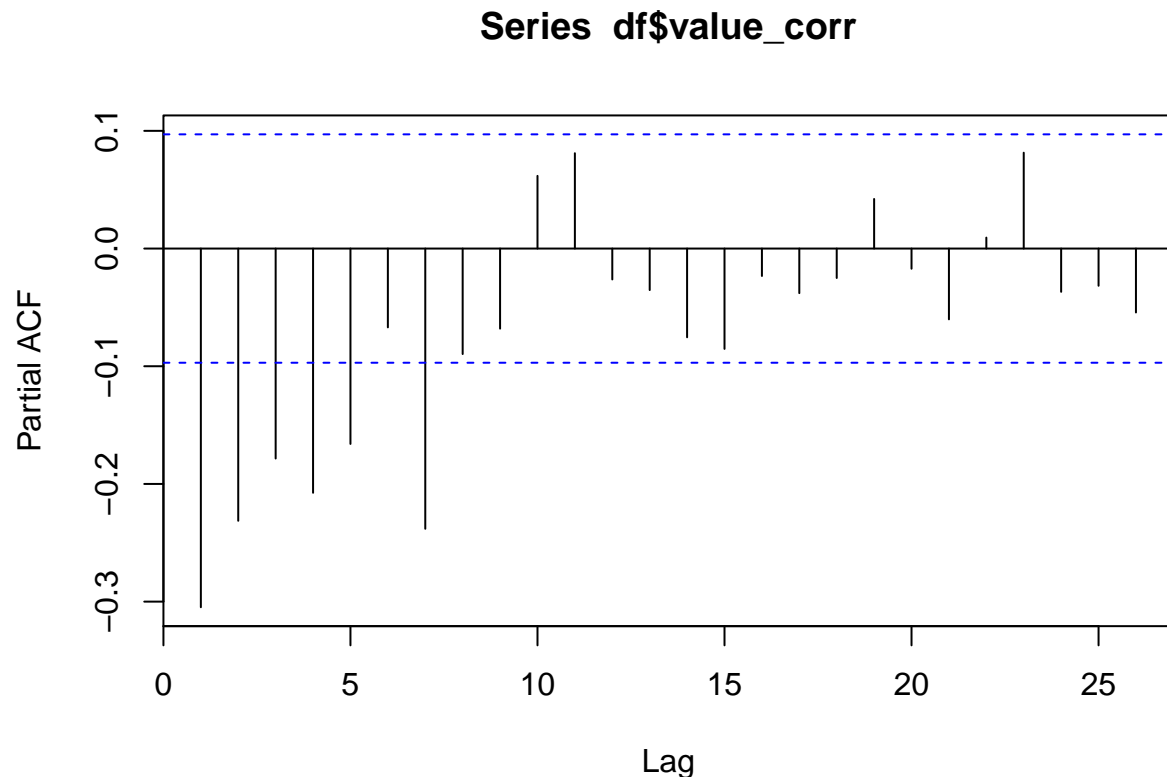
3 - Graphically represent the chosen series before and after transforming it.

Part II : ARMA models

4 - Pick (and justify your choice) an ARMA(p,q) model for your corrected time series Xt. Estimate the model parameters and check its validity

5 - Write the ARIMA(p,d,q) model for the chosen series.

Part III : Prediction

6 - Write the equation verified by the confidence region of level α on the future values (XT+1,XT+2).

7 - Give the hypotheses used to get this region.

8 - Graphically represent this region for $\alpha = 95\%$. Comment on it.

9 - Open question : let Yt a stationary time series available from t = 1 to T. We assume that YT+1 is available faster than XT+1. Under which condition(s) does this information allow you to improve the prediction of XT+1 ? How would you test it/them ?