

04-07

Le modèle ARIMA - SARIMA

**NOUS ÉCLAIRONS.
VOUS BRILLEZ.**

FORMATION CONTINUE
ET SERVICES AUX ENTREPRISES



Sommaire

1. Modèle et notations
2. La commande `auto.arima` par l'exemple
3. Digression: valeurs manquantes et aberrantes
4. Références



Modèle et notations

Le modèle ARIMA - SARIMA

- **ARIMA** (Autoregressive Integrated Moving Average) est l'une des méthodes de prévision les plus utilisées pour les séries univariées
- Cette méthode possède cependant une faiblesse majeure: elle ne supporte pas les séries présentant une composante saisonnière 😞😞😞
- **SARIMA** (S pour Seasonal) est une extension d'ARIMA prenant en compte la composante saisonnière d'une série univariée
- Le modèle **SARIMA** complet est donné par

$$\Phi_p(B)\Phi_P(B^S)\nabla^d\nabla_S^D X_t = \Theta_q(B)\Theta_Q(B^S)\epsilon_t$$

SARIMA

- Un processus satisfaisant le modèle précédent est appelé processus

$$\text{SARIMA}_s[(p, d, q), (P, D, Q)]m$$

S est la période du modèle

p ordre de la partie autorégressive

d est l'ordre de différentiation du modèle

q est l'ordre de la partie moyenne mobile du modèle

P est l'ordre de la partie autorégressive saisonnière du modèle

D est l'ordre de différentiation du modèle pour la partie saisonnière

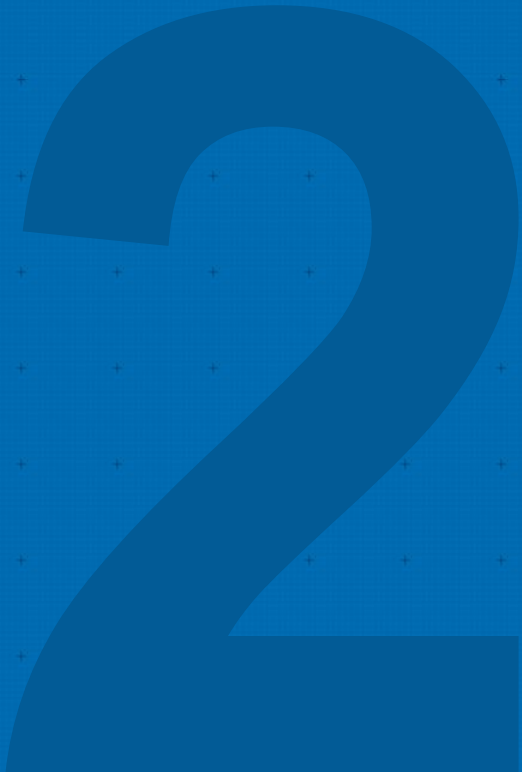
Q est l'ordre de la partie moyenne mobile saisonnière du modèle

m nombre d'observations pour une saison

Notations (optionnel)

$$\Phi_p(B)\Phi_P(B^S)\nabla^d\nabla_S^D X_t = \Theta_q(B)\Theta_Q(B^S)\epsilon_t$$

- **S** période du modèle (ou saisonnalité) Exemple: **S = 4** pour séries trimestrielles, etc...
- $\nabla = 1 - B$ et $\nabla_S = 1 - B^S$
- $\Phi_p, \Phi_P, \Theta_q, \Theta_Q$ sont des polynômes de degrés respectifs **p, P, q, Q** dont les racines sont de module strictement supérieurs à 1
- $(\epsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc

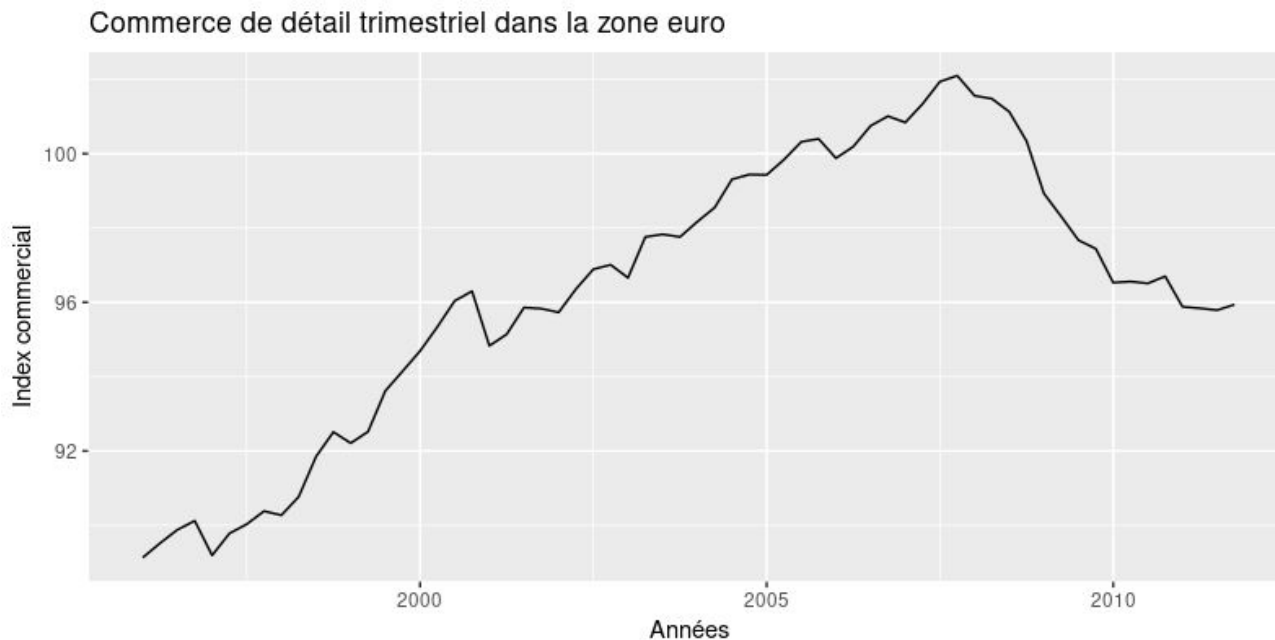


2

La commande
`auto.arima`
par l'exemple

Question 1

- Que pouvez vous dire de la série suivante ?

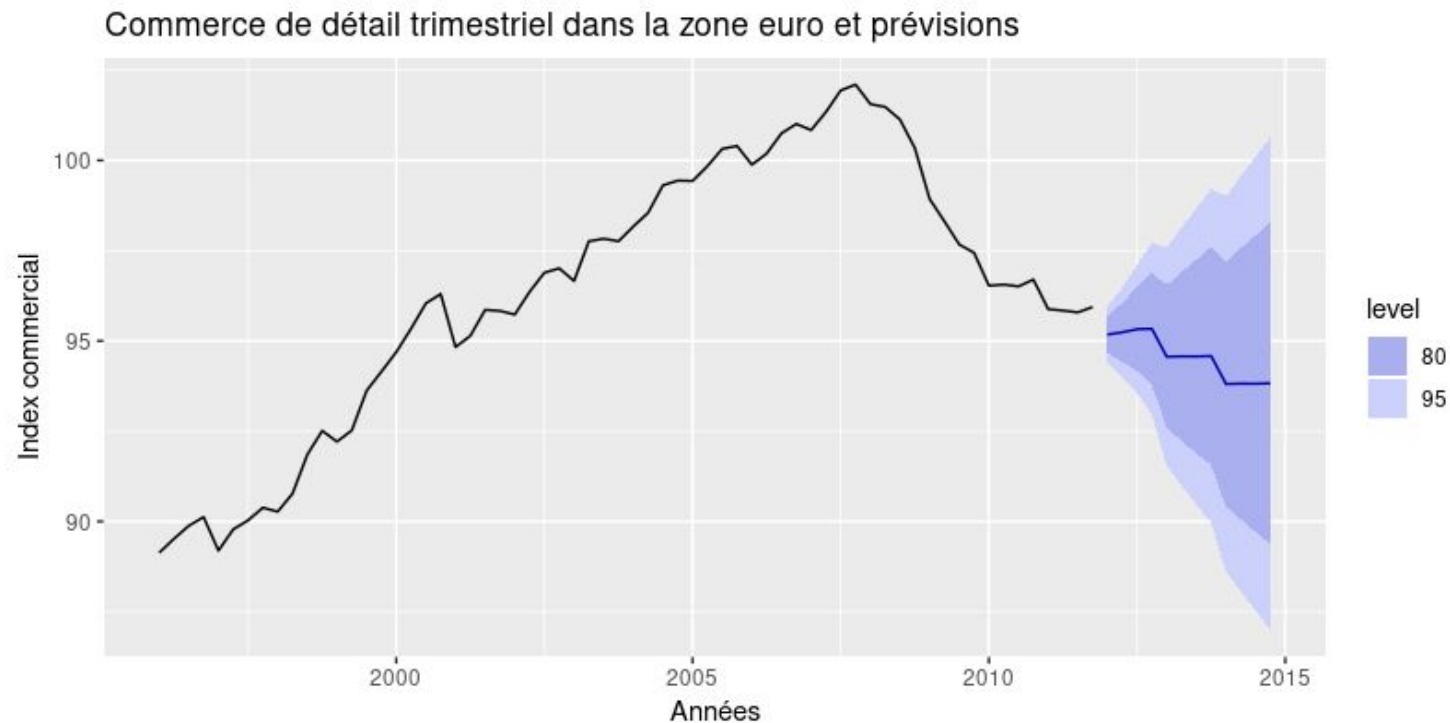


La commande `auto.arima` - Exemple #1

■ Code R

```
euretail %>%  
  auto.arima() %>%  
  forecast(h=12) %>%  
  autoplot() +  
    ylab("Index commercial") +  
    xlab("Années") +  
    ggtitle("Commerce de détail trimestriel dans la zone euro et prévisions")
```

La commande `auto.arima` - Exemple #1



La commande `auto.arima` - Exemple #1

Series: .

ARIMA(0,1,3) (0,1,1) [4]

Coefficients:

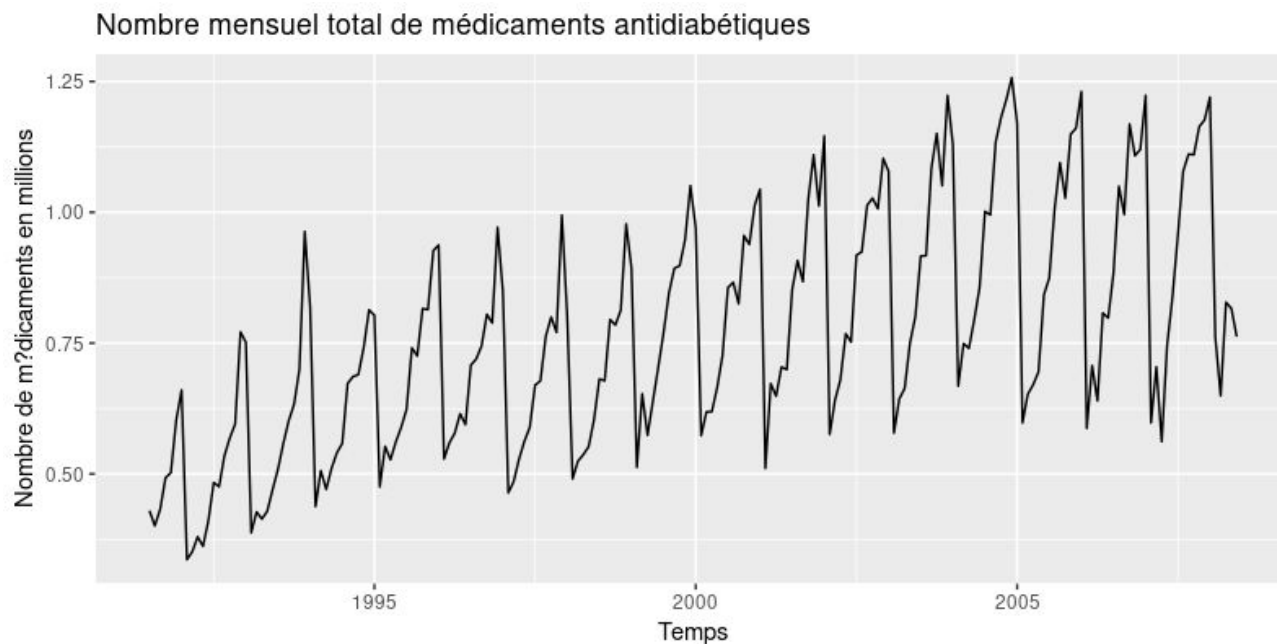
	ma1	ma2	ma3	sma1
	0.2630	0.3694	0.4200	-0.6636
s.e.	0.1237	0.1255	0.1294	0.1545

sigma^2 estimated as 0.156: log likelihood=-28.63

AIC=67.26 AICc=68.39 BIC=77.65

Question 2

- Que pouvez vous dire de la série suivante ?



La commande `auto.arima` - Exemple #2

■ Code R

```
h02 %>%
```

```
  auto.arima() %>%
```

```
  forecast(h=24) %>%
```

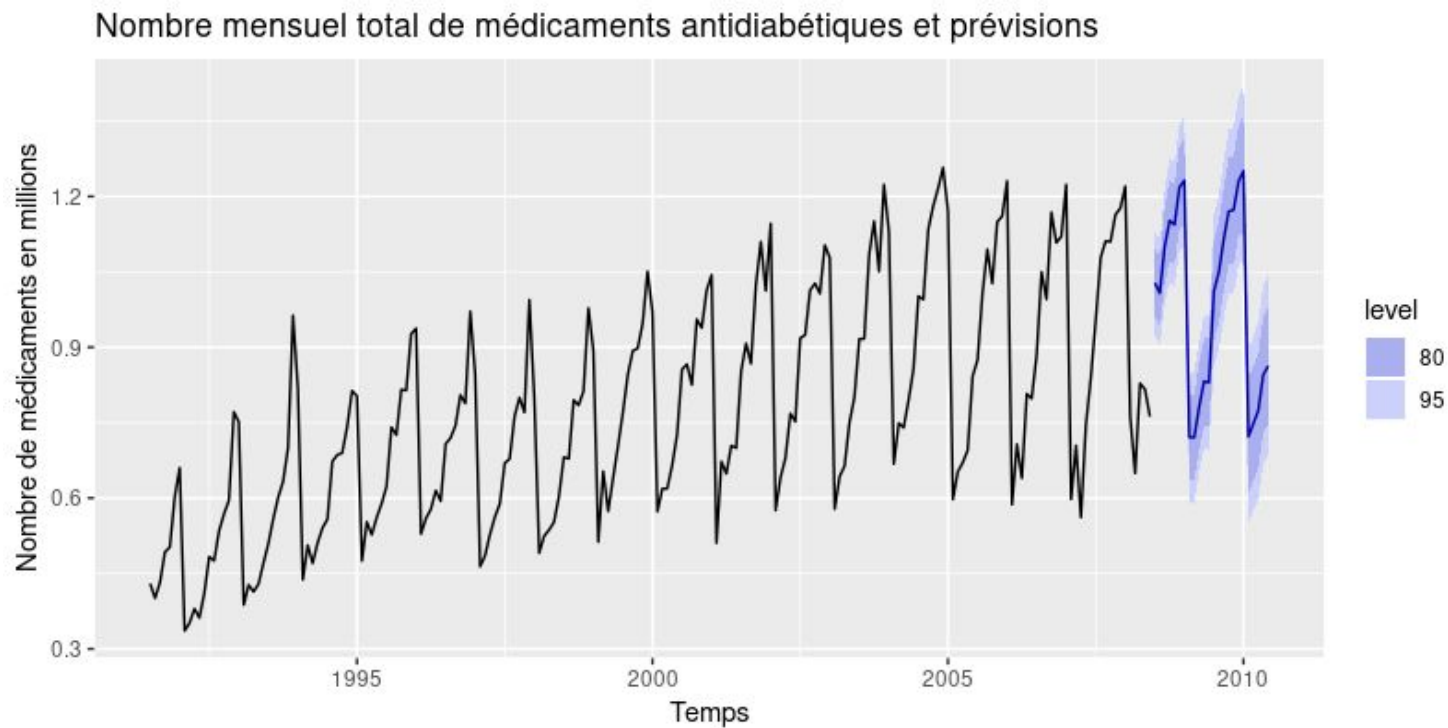
```
  autoplot() +
```

```
    ggtitle("Nombre mensuel total de médicaments antidiabétiques et prévisions") +
```

```
    xlab("Temps") +
```

```
    ylab("Nombre de médicaments en millions")
```

La commande `auto.arima` - Exemple #2



La commande `auto.arima` - Exemple #2

Series: .

ARIMA(4,1,1) (0,1,2) [12]

Coefficients:

	ar1	ar2	ar3	ar4	ma1	sma1	sma2
	0.0888	0.3386	0.2302	-0.2233	-0.9068	-0.4798	-0.1624
s.e.	0.1063	0.0976	0.0894	0.0850	0.0853	0.0913	0.0930

sigma^2 estimated as 0.00276: log likelihood=291.7

AIC=-567.4 AICc=-566.6 BIC=-541.38

La commande `auto.arima` - Exemple #2

■ On analyse les résidus

```
fit.arima <- auto.arima(h02)
```

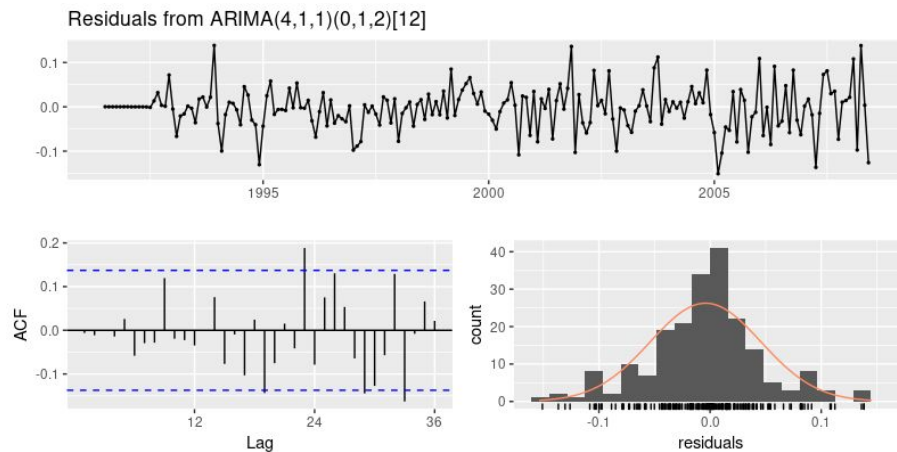
```
checkresiduals(fit.arima)
```

Ljung-Box test

```
data: Residuals from ARIMA(4,1,1)(0,1,2)[12]
```

```
Q* = 26.081, df = 17, p-value = 0.07299
```

```
Model df: 7. Total lags used: 24
```



La commande `auto.arima` - Exemple #2

- On peut faire un test d'évaluation de la pertinence des prévisions. On va retrancher ici 2 ans

```
train <- window(h02, end=c(2005, 12))
```


```
fit.arima <- auto.arima(train)
```

```
diab1 <- fit.arima %>%
```

```
  forecast(h=24) %>%
```

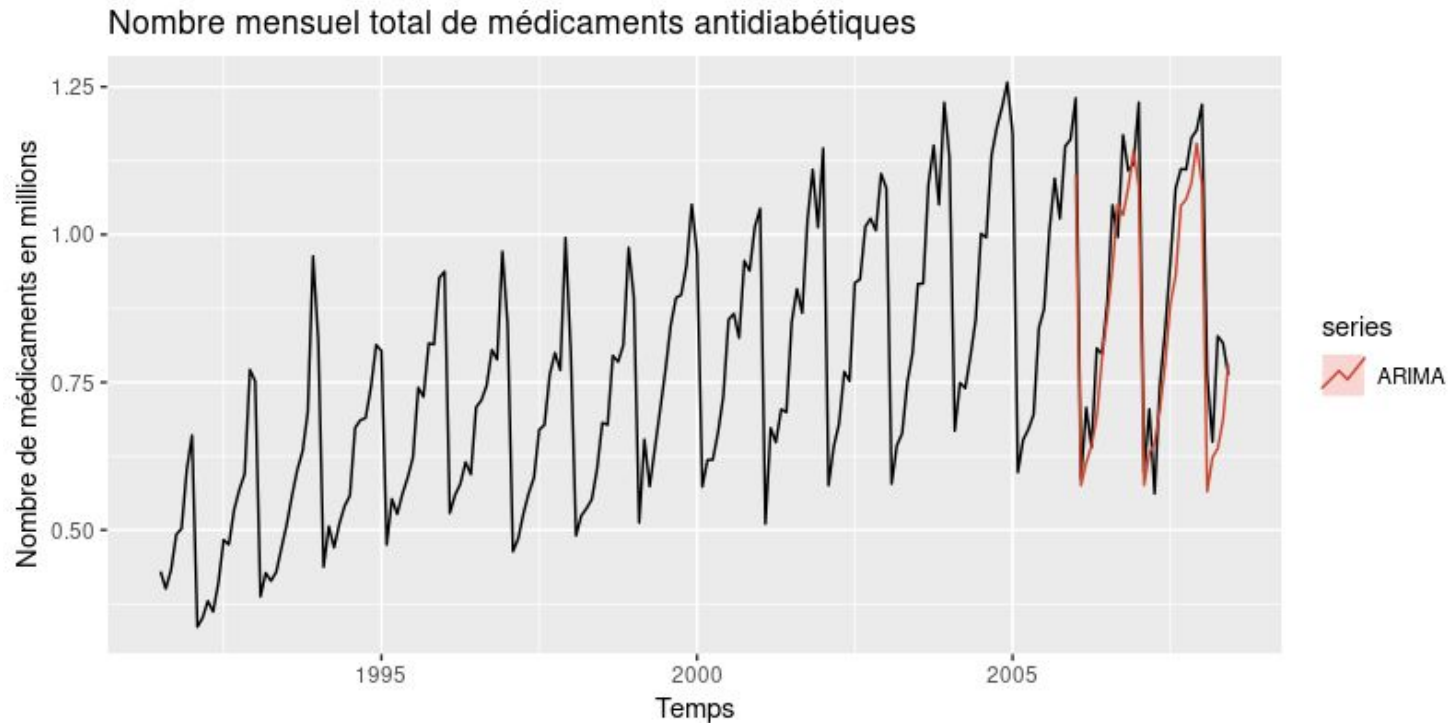
```
  accuracy(h02)
```

```
diab1[,c("RMSE", "MAPE", "MASE")]
```



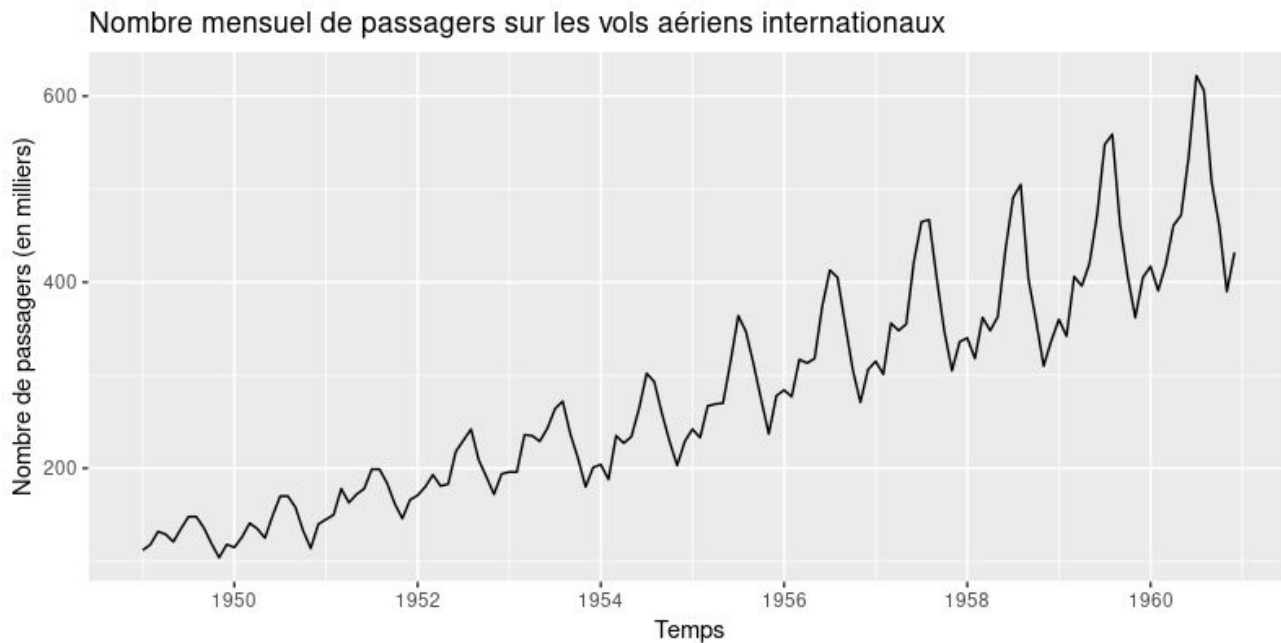
	RMSE	MAPE	MASE
Training set	0.04588769	4.568735	0.5600418
Test set	0.07927912	7.013943	1.0942005

La commande `auto.arima` - Exemple #2



Question 3

- Que pouvez vous dire de la série suivante ?



La commande `auto.arima` - Exemple #3

■ Code R

```
airline %>%
```

```
  auto.arima(lambda="auto") %>%
```

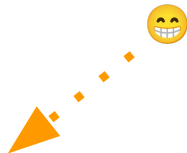
```
  forecast(h=24) %>%
```

```
  autoplot() +
```

```
    ggtitle("Nombre mensuel de passagers sur les vols aériens internationaux") +
```

```
    xlab("Temps") +
```

```
    ylab("Nombre de passagers (en milliers)")
```



La commande `auto.arima` - Exemple #3

Series: airline

ARIMA(0,1,1) (0,1,1) [12]

Box Cox transformation: lambda= -0.2947156

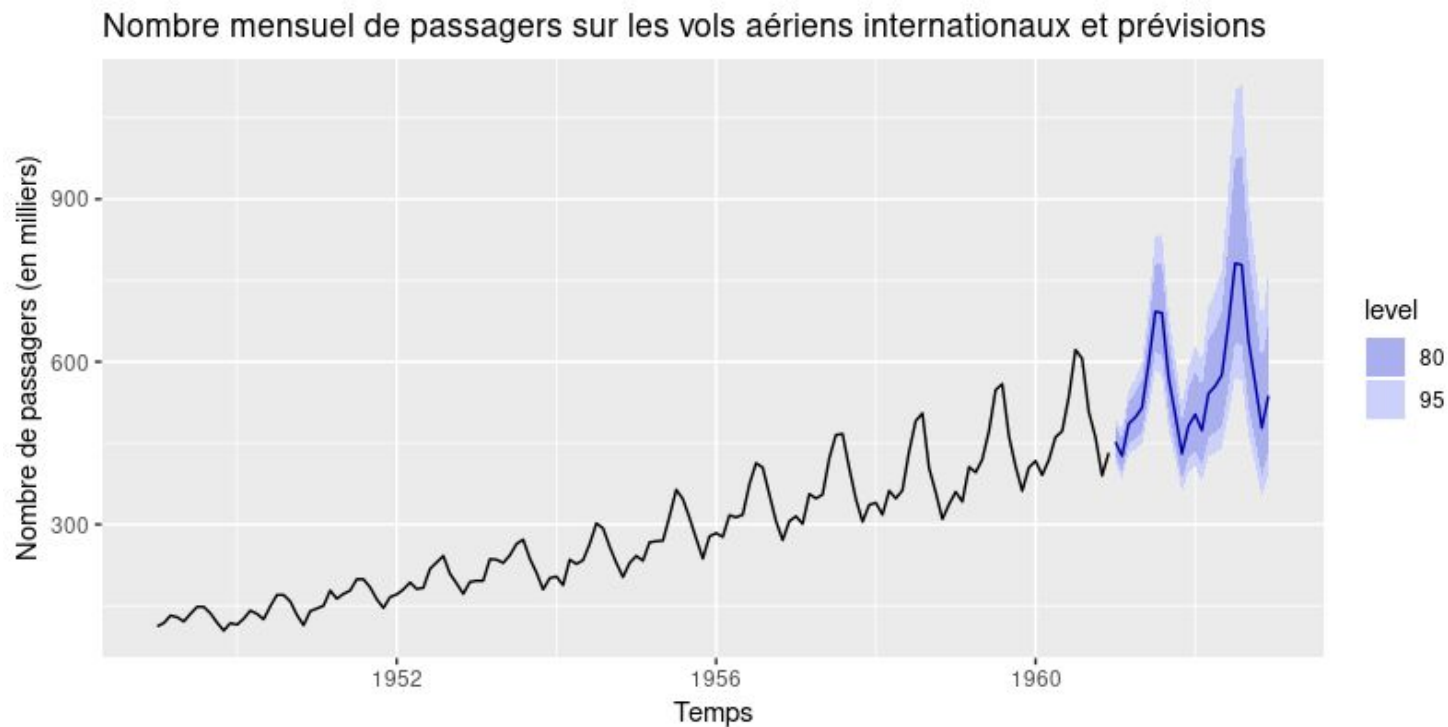
Coefficients:

	ma1	sma1
	-0.4355	-0.5847
s.e.	0.0908	0.0725

sigma^2 estimated as 5.855e-05: log likelihood=451.6

AIC=-897.19 AICc=-897.01 BIC=-888.57

La commande `auto.arima` - Exemple #3



La commande `auto.arima` - Exemple #3

■ On analyse les résidus

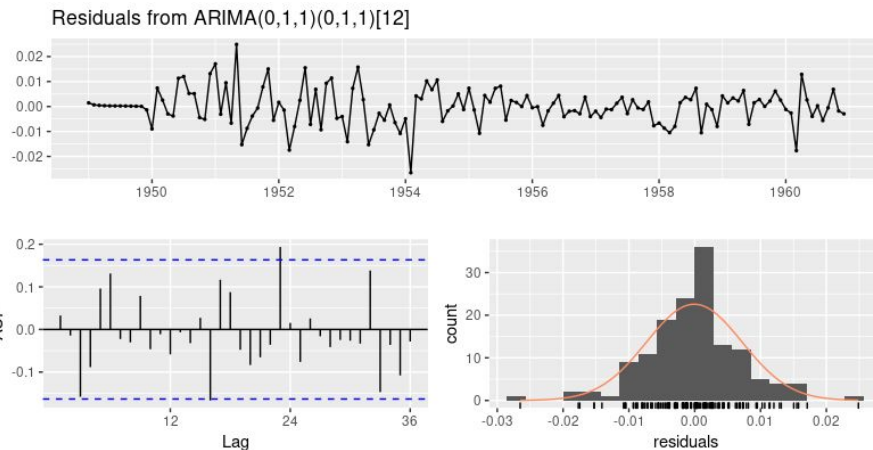
```
Fit1.arima <- auto.arima(airline, lambda="auto")  
summary(fit1.arima)  
checkresiduals(fit1.arima)
```

Ljung-Box test

data: Residuals from ARIMA(0,1,1)(0,1,1)[12]

$Q^* = 28.635$, $df = 22$, $p\text{-value} = 0.1556$

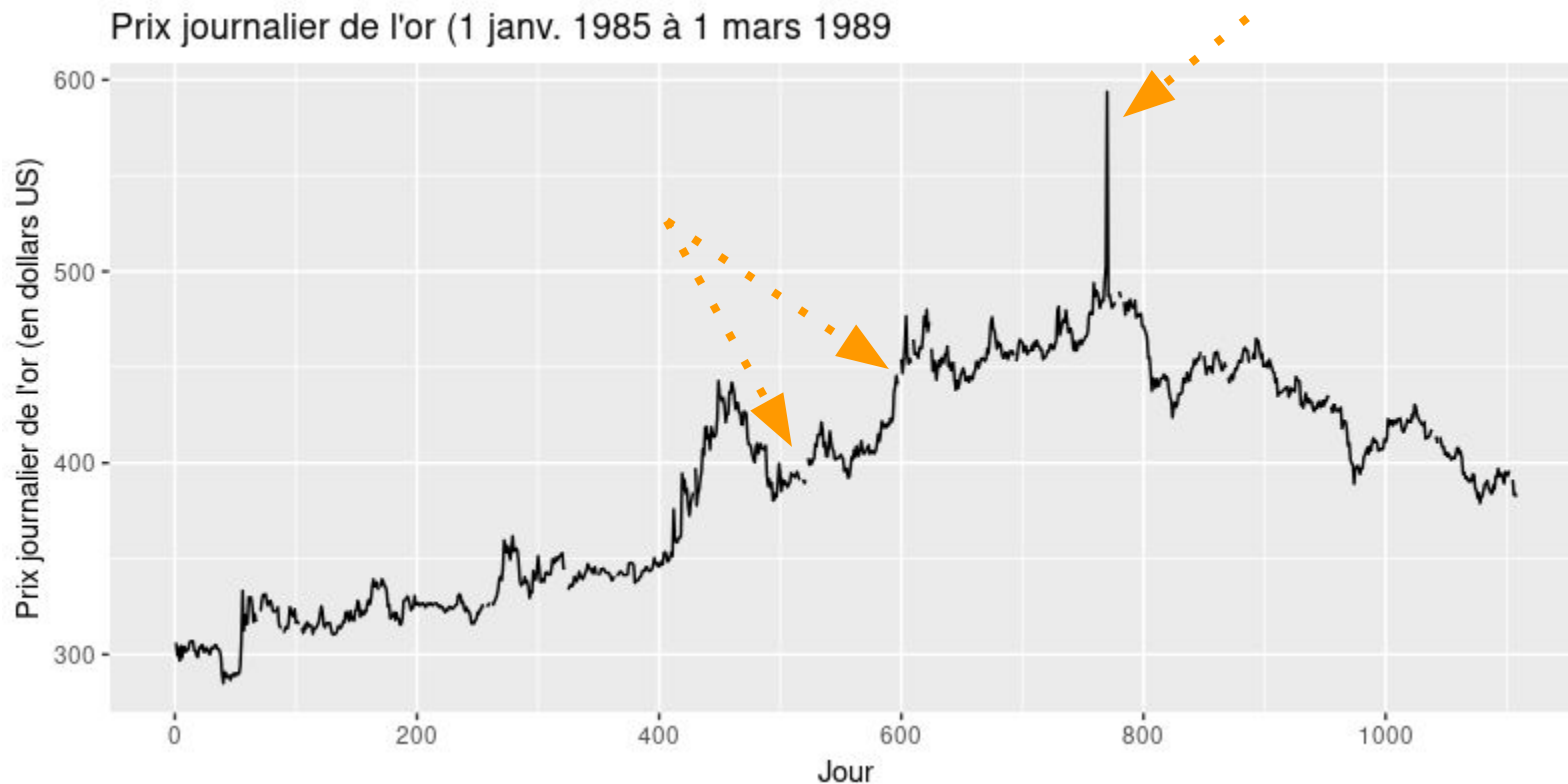
Model df: 2. Total lags used: 24





Digression: valeurs manquantes et aberrantes

Valeurs manquantes et aberrantes

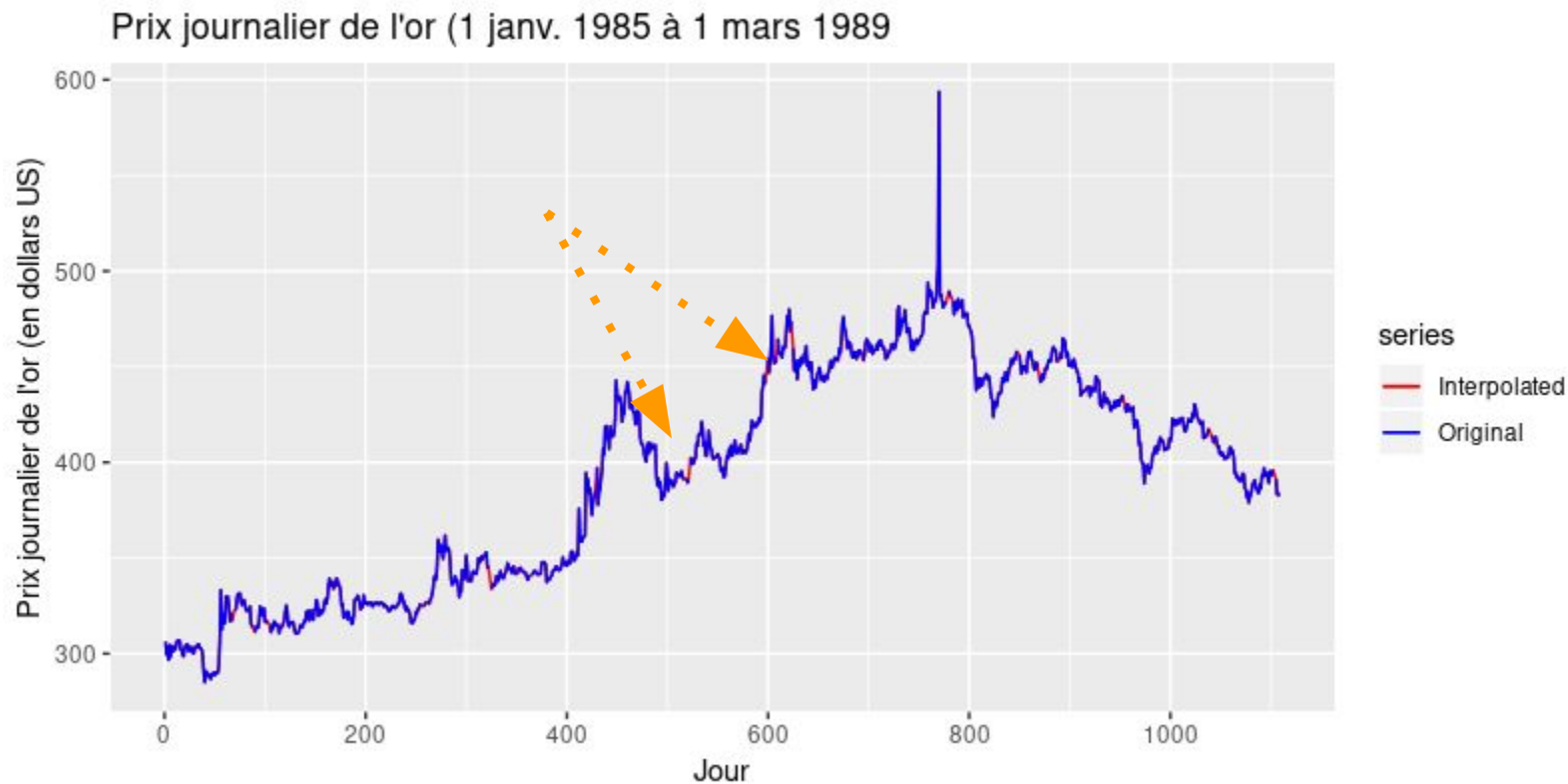


La commande `na.interp`

- Interpolation automatique des valeur manquantes (voir cours de préparation des données)
- Code R

```
gold1 <- na.interp(gold)
```

La commande `na.interp`

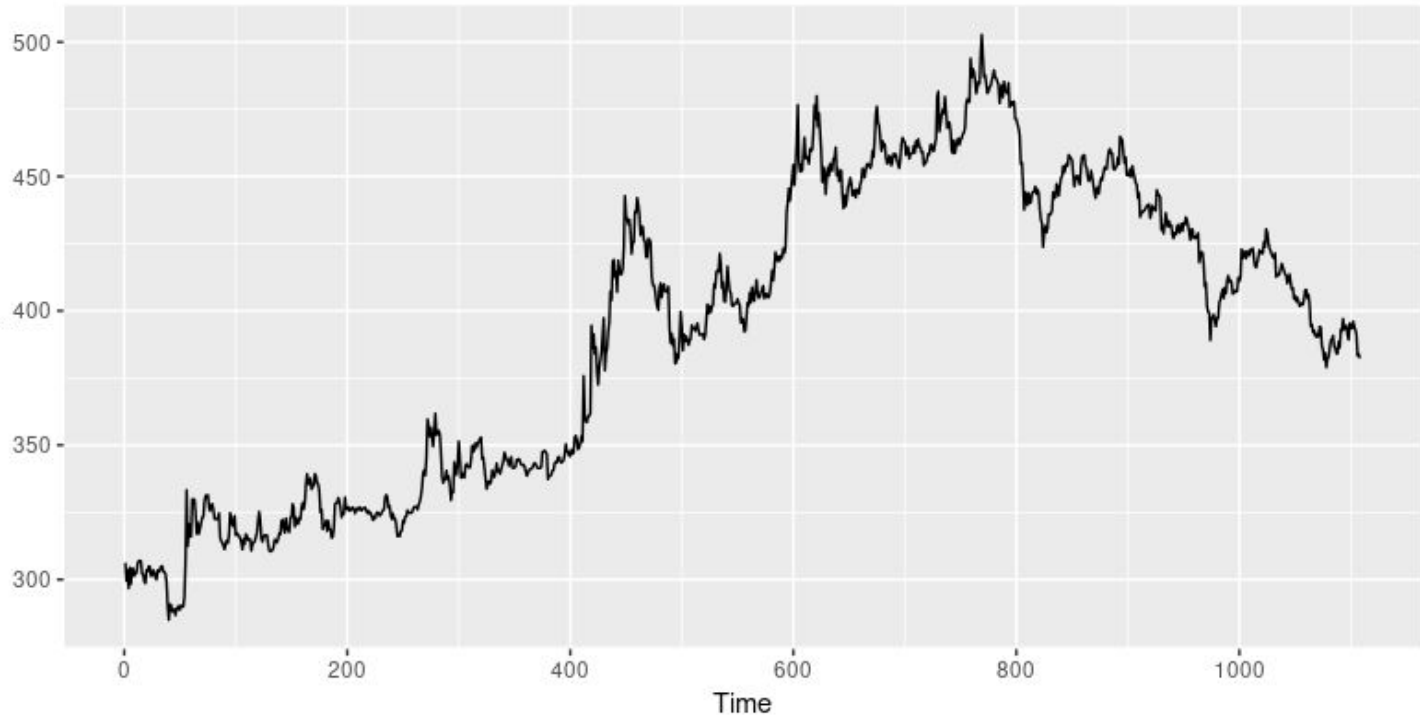


La commande `tsclean`

- Supprime automatiquement les valeurs aberrantes (voir cours de préparation des données)
- Code R

```
gold %>% tsclean()
```

La commande tsclean





Pull de <https://github.com/mswawola-cegep/420-a58-sf.git>

04-07-TP



Références

Références

- [1] Cours “R et la prévision de séries temporelles” de Michel Carbon - Université Laval
- [2] [A Gentle Introduction to SARIMA for Time Series Forecasting in Python](#)
- [3] [MODÈLE DE TYPE ARIMA](#)
- [4] [Analysez et modélisez des séries temporelles - Entraînez des modèles SARIMA](#)