Lasse9 Séries temporelles

Wissal BRIKA
2A-IDF
2024-2025

Plan

** les exams kykono fhal TP

Généralités

Modeles de tendance : TP 1

Moyennes mobiles

Décomposition saisonnière : TP 2

Lissage exponentiel : TP 3

Méthode de Gardner : TP 4

Exam 2022-2023

Enregistrement du Lasse9 : click here

01
Généralités

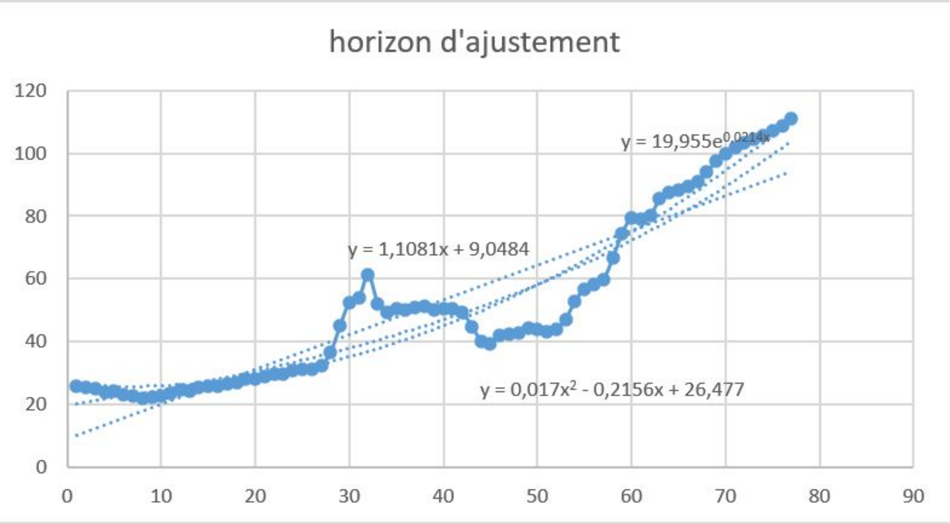
L’objectif

L’objectif principal de l’analyse d’une série temporelle est **la prévision** des ses futures réalisations . Afin de réaliser cet objectif , une première étape de modélisation est nécessaire.

Sélectionner parmi une famille de modèles , celui qui décrit le mieux la série en question

Série temporelle :

Une collection des données observés dans des intervalles de temps successifs



Une série temporelle se décompose en

1. Tendance

2. Composante saisonnière

3. Bruit

4. Cycle

** tout au long de ce cours , nous utilisons les termes MAPE , MSE(RMSE), MA : ce sont des métriques ou critères de comparaison qu'on utilise pour comparer les modèles

1. Tendance

C'est l'évolution à long terme de la série (hausse/ baisse)

⇨ Elle peut suivre l'un des modèles suivants :

- Tendance linéaire
- Tendance quadratique
- Tendance exponentielle
- Tendance exponentielle modifiée
- Tendance Gompertz
- Tendance logistique

** f TP 1 kenna étudiena had les modeles kamlin w comparaina binathom par calcul du MAPE bach n3rfo ina wahd ky approcher le plus la tendance de notre série

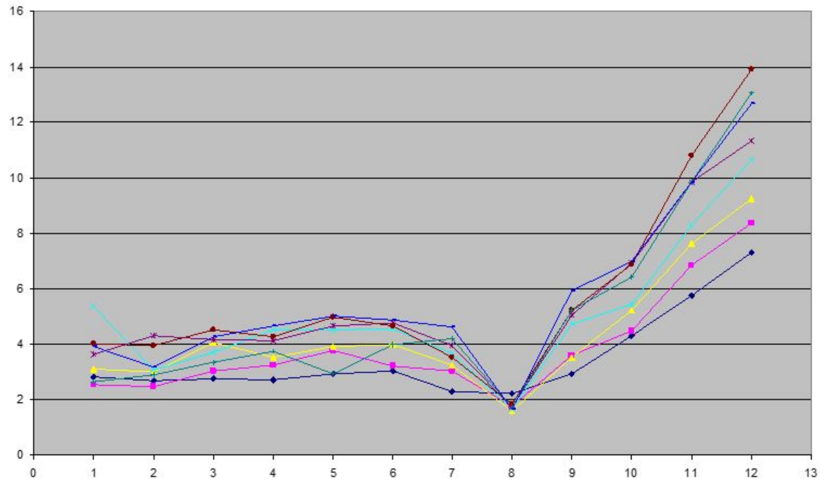
** c’est logique que l’identification de la composante saisonnière nous aide à bien prédire son comportement futur

2. Composante saisonnière

Un phénomène périodique de période identifiée

La composante saisonnière est identifiable ici par la répétition de schémas similaires sur plusieurs années .

Elle traduit l’impact régulier et cyclique de facteurs externes (saisons, comportements, événements) sur la variable étudiée



3. Le bruit

La partie aléatoire de la série

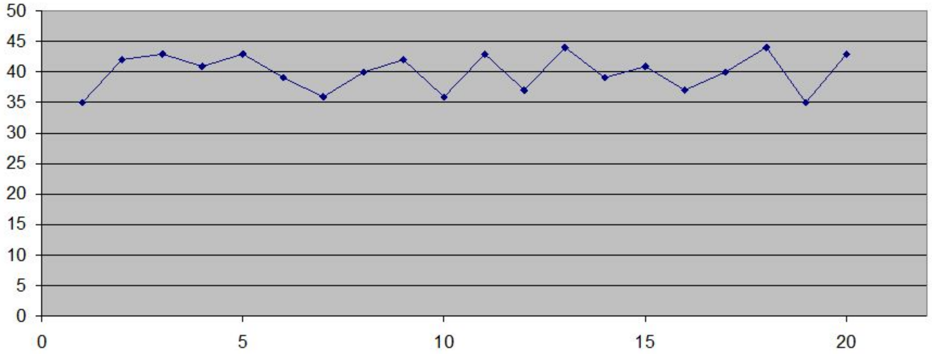
4. Le cycle

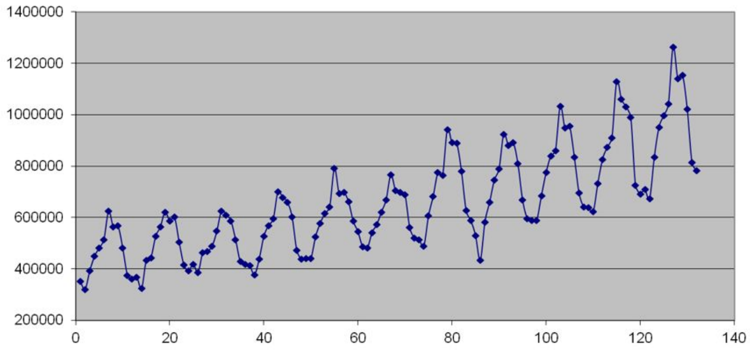
** had le cycle machi mohim bnisba lina daba hadchi li kan gal l prof w ma3mrna khdmna bih

Phénomène répétitif régulier (donc prévisible) de période inconnue ou changeante

Si le prof demande d'interpréter la courbe de la série , on peut répondre à ces questions :

- Le graphique est-il lisse ?
- Peut on identifier un phénomène périodique?
- Peut on identifier des niveaux dans le graphique ?
- La courbe tend a croitre ? a descendre ? ou fluctue autour d’une moyenne stable ?





02

Modeles de tendance

1) Tendance linéaire $y_t = a_0 + a_1 t$

2) Tendance exponentielle $y_t = \alpha\beta^t$

3) Tendance quadratique $y_t = a_0 + a_1 t + a_2 t^2$

4) Tendance exponentielle modifiée : $y_t = \alpha\beta^t + \gamma$

5) Tendance de Gompertz $y_t = e^{\alpha\beta^t + \gamma}$

6) Tendance logistique $y_t = \dfrac{1}{\alpha\beta^t + \gamma}$

** (t) c'est le numéro de l'observation (c'est pourquoi on ajoute souvent une colonne qui contient les numéros des observations)

** (Yt) c'est la valeur de la série qu'on veut estimer , variable dépendante

** Les autres variables on verra par la suite comment les estimer

Estimation des parametres

Tendance linéaire

$y_t = a_0 + a_1 t$

$$a_0 = \bar{y} - a_1\bar{t}, \text{ avec } \bar{y} = \frac{\sum_{t=1}^{T} y_t}{T} \text{ et } \bar{t} = \frac{T+1}{2}$$

$$a_1 = \frac{\sum_{t=1}^{T}(y_t - \bar{y})(t - \bar{t})}{\sum_{t=1}^{T}(t_t - \bar{t})^2} = \boxed{\frac{\operatorname{cov}(y,t)}{\operatorname{var}(t)}}$$

** on calcule la pente a1 en premier lieu pour pouvoir calculer a0 (la moyenne de y - a1 * la moyenne de t)

** sur excel , cov(y,t) = COVARIANCE.P(y,t) hia l covariance de la colonne y(valeurs de la série) avec la colonne t(observations) , et var(t) se calcule par VAR.P de la colonne t

Estimation des parametres

Tendance exponentielle

** n'oubliez pas d'appliquer EXP lors du calcul de la colonne du modèle !!

modele exponentiel = EXP(modèle linéaire)= EXP(a0' + t * a1')

Donc les coefficient se calculent de la même manière que dans le slide précédent sauf qu'ils ne se calculent pas sur la colonne y mais plutôt ln(y)

Tendance quadratique

$$y_t = a_0 + a_1 t + a_2 t^2$$

** elle est de cette forme , mais les coefficients sont tirés de la représentation graphique de excel

Pour les 3 modeles qui restent ..

Les paramètres peuvent être estimés de 2 manières différentes

1. **Méthode des 3 points**

** machi dima kankhdmo biha w y9der y3tik chi question 3liha w y golik si elle est applicable

Elle est applicable lorsque la série est monotone (croissante ou décroissante pas les deux)

- Répartir les données en trois sous ensembles I , II , III
- Calculer les médianes des Yt pour les 3 ss ensembles Y(I),Y(II),Y(III)

T multiple de 3	T-1 multiple de 3	T-2 multiple de 3
$\frac{T}{3}$	$\frac{T-1}{3}$	$\frac{T-2}{3}+1$
$\frac{T}{3}$	$\frac{T-1}{3}+1$	$\frac{T-2}{3}$
$\frac{T}{3}$	$\frac{T-1}{3}$	$\frac{T-2}{3}+1$

$$t_{III} - t_{II} = t_{II} - t_{I} = \Delta$$

** hadi l'estimation des parametres alpha , beta , gamma de la tendance exponentielle modifiée

** les formules sont à mémoriser

$$\ln(\beta) = \frac{1}{\Delta} \ln\left(\frac{y_{III} - y_{II}}{y_{II} - y_{I}} \right)$$

$$\alpha = \frac{y_{III} - y_{II}}{\beta^{t_{III}} - \beta^{t_{II}}}$$

$$\gamma = y_{I} - \alpha\beta^{t_{I}}$$

Pour ne pas se plonger dans les calculs des équations qu'on a vu dans un slide précédent , les paramètres des autres modèles sont calculés exactement de la même manière sauf que pour le **modèle logistique** on utilise **1/y** au lieu de y et dans le **modèle de Gompertz** on utilise **ln(y)**

** matnsawch encore une fois d'appliquer l'inverse et EXP lors du calcul des modèle

Tendance exponentielle modifiée

= alpha* beta^t +gamma

Tendance logistique

= 1/(alpha'* beta'^t +gamma')

Tendance de Gompertz

= EXP (alpha''* beta''^t +gamma'')

Le tableau de calcul ressemble a ca

	méthode des 3 points						
t	**Mediane**	y	1/y	ln(y)			
13,5	I	25,55	0,03914029	3,24062008		26	25,5
39	II	49,3	0,02028398	3,89792408		25	25,5
64,5	III	87,9	0,0113768	4,47618945		26	25,5
25,5	Δ						
	ln(β)	0,01904587	-0,0294114	-0,00502407			
	β	1,0192284	0,9710169	0,99498853			
	α	29,3721198	0,05315774	-5,8499094			
	γ	-12,4340067	0,00340247	8,70691616			

2. Procédures numériques

Via R , matlab ..

Nous utilisons toujours (f l cours) la procédure **optim**

Elle peut toujours être appliquée , elle permet à partir d'une solution initiale (soit aléatoire soit donnée par le prof) d'estimer les paramètres optimum pour les 3 derniers modeles de tendance (gahnsta3mloha hta f d'autres cas mais pour le même but)

**B3d lmrat t9dr tkon la solution initiale hia dok les parametres li khrjo lina b la méthode des 3 points , pour vérifier si optim peut nous donner des résultats meilleurs

On verra comment ca marche dans le code R

TP 1

gnp_annuel

03

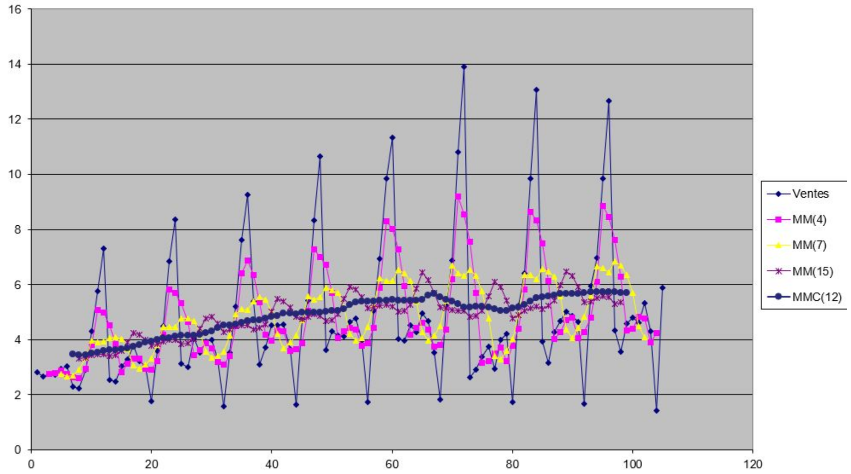
Moyenne mobile

C’est un moyen de filtrage des données qui consiste a calculer la moyenne des valeurs successives d’une série temporelle sur une fenêtre glissante de taille k

Permettent de lisser la série (éliminer les fluctuations)pour détecter la tendance

$$MMC12 = \frac{1}{12}\left[\frac{1}{2} 1 \dots 1 \frac{1}{2}\right]$$

13 termes



On calcule une moyenne mobile centrée pour les ordres paires , c'est plus pratique et pour un ordre impaire on calcule une moyenne mobile normale

La moyenne mobile d'ordre 2m+1 : $\frac{1}{2m+1}[1 \ldots 1]$ **élimine les fonctions périodiques de période 2m+1**

La moyenne mobile centrée d'ordre 2m+1: $\frac{1}{2m}\left[\frac{1}{2} 1 \ldots 1 \frac{1}{2}\right]$ **élimine les fonctions périodiques de période 2m**

Série mensuelle : MMC(12)
Série quotidienne : MM(7)
Série trimestrielle: MMC(4)

** l'ordre de la moyenne howa le nombre de termes utilisé
** en gros kankhtaro la MM selon la période de la composante saisonniere

“e” c’est l’erreur c’est à dire l’écart entre la série réelle et la série ajustée

Criteres de comparaison

$$MAPE(e) = \frac{1}{T}\sum_{t=1}^{T}\frac{|e_t| * 100}{y_t}$$

$$RMSE(e) = \sqrt{\frac{\sum_{t=1}^{T} e_t^2}{T}}$$

$$MAE(e) = \frac{\sum_{t=1}^{T} |e_t|}{T}$$

$$MSE(e) = \frac{\sum_{t=1}^{T} e_t^2}{T}$$

Convergence des criteres : tous les critères utilisés indiquent le même modèle (comme meilleur)

Divergence des criteres : au moins un des critères donne un résultat différent des autres

04

Décomposition saisonnière

La décomposition saisonnière sert à isoler et identifier les composantes d’une série temporelle : tendance , saisonnalité , cycle et bruit

Type de décomposition :

Modele additif

T + C + S + E

L’amplitude de la saisonnalité reste constante au fil du temps

Modele multiplicatif

T*C*S*E

L’amplitude de la saisonnalité varie proportionnellement au niveau de la tendance

Quel modele choisir ? $\Rightarrow$ **Test de Buys Ballot**

Test de Buys Ballot

année	moyenne	écart-type
1962	3,48	1,49
1963	3,86	1,82
1964	4,34	2,03
1965	5,02	2,28
1966	5,37	2,61
1967	5,71	3,24
1968	5,01	3,18
1969	5,64	2,87

On calcule la corrélation entre les deux colonnes

Si $R > 0,3$: modele multiplicatif
Si $0 < R < 0,2$: modele additif
Si $0,2 < R < 0,3$: les deux
(ma3rftch kifach had les deux lmohim howa hadchi li kan gal)

** 3la hsab ma chft f les exams souvent kyt3ti l modele li ghankhdmo bih

Étapes de la décomposition saisonnière

1. Calcul de MMC(12)

** hna kandiro MMC(12) hit kanhdro 3la série mensuelle w hadchi li dayr howa flcours

On dispose au début de la série , on en calcule MMC(12) pour estimer la tendance et le cycle de la série selon les schéma suivants

$$\begin{cases} \text{Modèle additif : } T + C \\ \text{Modèle multiplicatif : } T * C \end{cases}$$

** had l3yba ghir bach tfhmo layach katslah w 3lach kanhsbo dakchi li jay b dik tari9a l'essentiel howa awal haja kanbdaw biha hia moyenne mobile

2. Extraction de la composante saisonnière

$$\begin{cases} \text{Série – MMC(12) si modèle additif} \\ \text{Série/MMC(12) si modèle multiplicatif} \end{cases}$$

** hadchi ca revient l la définition li kna drna flwl 3la kifach kandécomposew une série temporelle

3. Calcul de la composante saisonnière corrigée

mois	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	médiane	comp.sais.corr
janvier		1,012303333	1,001496323	1,004119415	1,000377572	0,99956	1,00126	1,00659	1,00417	0,99759	1,0015	1,000859098
février		1,006073057	1,001534919	1,002560636	1,000102058	1,00005	1,00278	1,00567	0,99855	0,99875	1,00153	1,000897669
mars		1,002315645	0,998825757	1,002067364	0,999984333	1,00135	0,99938	1,00343	0,9977	0,99904	0,99998	0,99934807
avril		1,00071582	1,000231446	1,000707523	1,002266652	1,00091	1,00159	1,00152	0,99752	0,99925	1,00072	1,000079091
mai		1,001687654	1,000661813	1,001144383	0,999571113	1,00001	1,00263	1,00234	0,99857	0,99961	1,00066	1,000025119
juin		1,002937535	1,000517817	0,981358073	0,997479227	1,00129	1,0031	1,00247	0,99819	1,00018	1,00052	0,999881214
juillet	1,008103292	1,000597439	1,000477601	0,999269679	0,996506916	1,0012	1,00345	1,00063	0,9998	0,99969	1,00054	0,999900904
aout	0,878240298	1,001925789	0,997175973	0,993582676	0,997351424	1,00033	0,98134	1,00098	1,00196	0,99858	0,99797	0,997332795
septembre	1,013827045	1,001415793	1,000596274	1,000792468	1,00010043	0,99903	0,99384	1,00214	1,00119	0,99901	1,00069	1,000057656
octobre	1,010863847	1,001670736	1,001422385	0,999097773	1,001032723	0,99881	0,99355	1,00244	1,00123		1,00123	1,000589086
novembre	1,006039503	0,99944657	1,001080325	1,001959793	1,00164542	0,99972	0,9951	1,00156	1,00109		1,00109	1,000452121
décembre	1,004162409	0,999943924	1,002551203	1,001214224	1,001142752	1,00004	1,00664	1,00438	1,00074		1,00121	1,000577178
										somme	12,0076	12
										facteur de correctio	0,99936	

$$\begin{cases} \sum_{i=1}^{12} S_i = 0 & \text{pour le modèle additif} \\ \sum_{i=1}^{12} S_i = 12 & \text{pour le modèle multiplicatif} \end{cases}$$

$$\boxed{\text{Facteur de correction} = -\frac{1}{12}\sum_{i=1}^{12} \text{médiane}_i}$$

$$\boxed{\text{Facteur de correction} = \frac{12}{\sum_{i=1}^{12} \text{médiane}_i}}$$

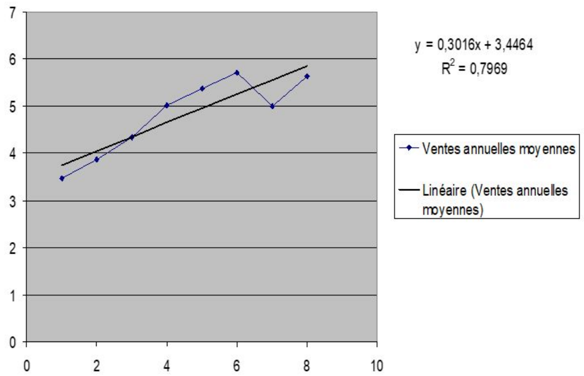
4. Calcul de la tendance

** comme ce qu'on a vu f hadok les modeles de tendance w bnisba l les coefficients kanjbdohom graphiquement

Elle peut être calculée en 2 méthodes :

1. Approximation de la tendance par les modèles de croissance
2. Dans le cas linéaire on étudie les données annuelles moyennes (tend ann)

** anchofo kifach nhsboha f slide jay



Pour calculer la tendance annuelle :

- On utilise les moyennes annuelles de notre série
- On les approche par un modèle linéaire : at+b avec a et b calculés de la même manière qu'on a fait précédemment

Données : Moyennes Annuelles		Données : Mensuelles
$y = at + b$	$\longrightarrow$	$y = \alpha t + \beta$
$t = 1 \rightarrow$ 1juillet 1962	$\longrightarrow$	15 juillet 1962
3.4781	$\longrightarrow$	3.4781 + a/24 = 3,76

Si a est l'augmentation annuelle, alors a/24 est l'augmentation de 15 jours

L'augmentation mensuelle α = a/12 = 0,3016/12

Données mensuelles : juillet 1962 $\rightarrow$ 3,76
aout 1962 $\rightarrow$ 3,76 + α

** y a l'application de ca dans le TP2 pour mieux comprendre

5. Calcul du modèle et des prévisions

$$\hat{Y}_t = \begin{cases} \hat{T}_t + S_t & \text{schéma additif} \\ \hat{T}_t S_t & \text{schéma multiplicatif} \end{cases}$$

T : tendance
S : composante saisonnière

Vient après le calcul des critères de comparaison pour comparer les résultats des différents modèles (TP2)

!! Faites attention à la période ou la saisonnalité de la série
Si elle est trimestrielle on remplace 12 par 4

Mois stables /instables/assez stables

Dans le même tableau de calcul de la composante saisonnière

Mois stable : les mois dont les valeurs de la composante saisonnière sont toutes supérieurs à 1 o toutes inférieurs à 1
Mois assez stable : les composantes sont toutes supérieurs à 1 ou toutes inférieurs à 1 sauf une valeur extrême qui diffère des autres
Mois instable : sinon

TP 2

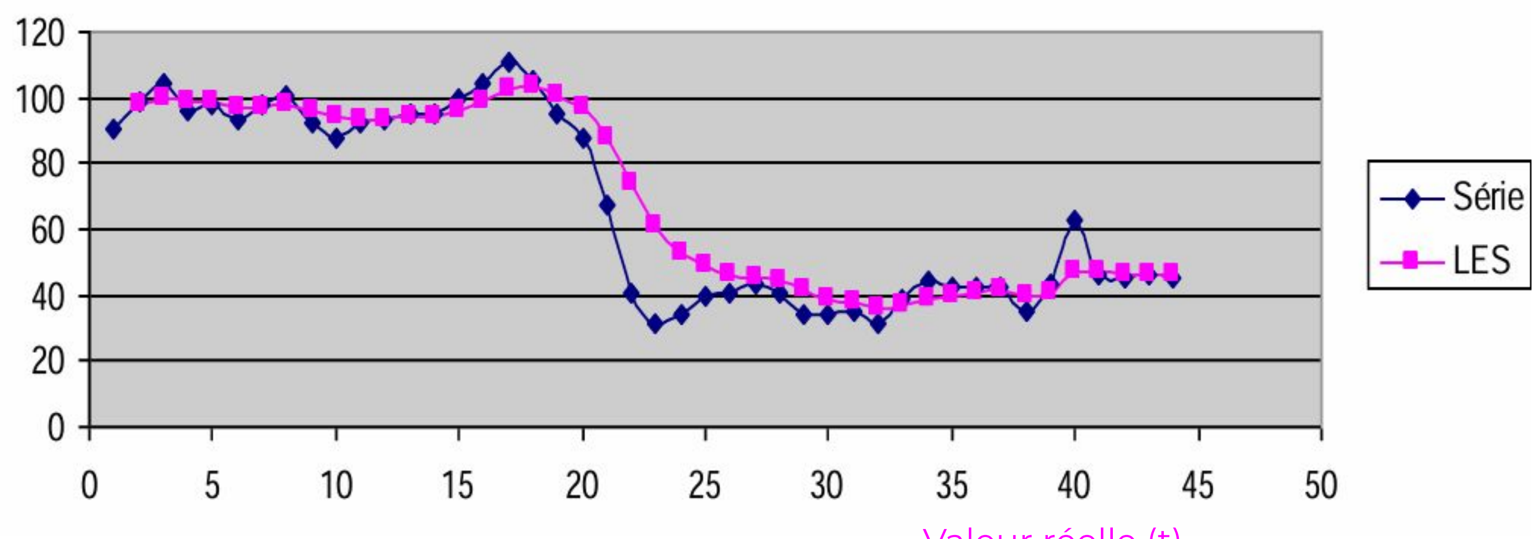
05

Lissage exponentiel

Le lissage exponentielle :

Méthode de prévision et de filtrage des séries temporelles

!! elle est plus performante que les moyennes mobiles , elle s'adapte aux différents niveaux de la série



Valeur réelle (t)

Sa formule générale : $F(t) = \alpha * Y(t) + (1-\alpha) F(t-1)$

Valeur lissée (t)

Coefficient de lissage

**A chaque fois kytbedlo fiha chi hwayj selon le type de lissage

Etapes de calcul d’un modèle de lissage exponentiel

** had les étapes homa qui définissent la structure du code R

1. Paramètres du modèle (valeurs initiales et paramètres de lissage)
2. Choisir le critère de comparaison a minimiser pour ajuster le modèle (RMSE , MAE , MAPE)
3. La fonction “critère de comparaison” dépend des paramètres du modèle
4. L’estimation des paramètres se fait en calculant le minimum de la fct critère de comparaison par application de la procédure d’optimisation optim
5. Calcul du modèle sur la période d’ajustement
6. Calcul des prévisions
7. Qualité de prévision

** ghanrj3o lihom f tp bach tchofo en quoi c’est utile

** daba b9a lina n3rfo les spécificités de chaque modèle

Lissage exponentiel simple (LES)

Rq :les observations dédiés a la prévision n'entrent dans les calculs du modele , howa flwl kygol lina par exemple kankhliw les 12 dernieres observations ghankhliwhom pour la prévision , du coup awal haja kanbdaw biha f lcode hia kann9so hadik 12 mn le nombre d'observations , ghanchofo kifach f l code R dyal tp

** makankhdmoch bih bzaf saraha

Les paramètres du modèles sont : F_2 et α

On choisit le MAPE comme critère à minimiser pour estimer le modèle, c'est une fonction dans ce cas à 2 variables

Initialisation des paramètres : $F_2 = x[1]$ et $\alpha = x[2]$

Mape = 0
Pour i = 3 à T-h

$$F_i = \alpha y_i + (1-\alpha)F_{i-1}$$

$$Mape = Mape + abs(y_i - F_{i-1}) * 100/y_i$$

Fin_pour

$$Mape = Mape/(T-h-2)$$

Optimisation de la fonction Mape par la procédure optim

Calcul du modèle sur l'horizon d'ajustement : 1,…, T-h

Calcul des prévisions pour les h derniers instants :

$$\hat{y}_t = F_{T-h},\ \text{pour } t \geq T-h+1$$

Avec :
F2 = la moyenne des 3 premières valeurs de la série
T-h= le nombre d'observations - horizon de prévision

** l'implémentation ghatl9awha fl fichier Mod_LES

Lissage exponentiel double (LED)

Série avec tendance

Avec:

La série ajustée : $F_{t-1} = S_{t-1} + T_{t-1}$

L'erreur d'ajustement : $e_t = y_t - F_{t-1}$

$$\boxed{\alpha = 0.3, S_2 = \frac{y_1 + y_2 + y_3}{3}, T_2 = \frac{y_3 - y_1}{2}}$$

Différente de hadi

** C'est l'initialisation , pour alpha machi darori 0,3 t9dro takhdo li bghito hit hado des valeurs li kandkhlohom comme valeurs initiales l optim katbda bihom w kat3tina les valeurs optimales dyalhom fhal li kna glna flwl

Les paramètres du modèles sont : $\boxed{S_2, T_2 \text{ et } \alpha}$

On choisit le MAPE comme critère à minimiser pour estimer le modèle, c'est une fonction dans ce cas à 3 variables

Initialisation des paramètres : $S_2 = x[1], T_2 = x[2]$ et $\alpha = x[3]$

$\lambda = 2\alpha - \alpha^2$ et $\mu = \frac{\alpha}{2-\alpha}$

Mape = 0

Pour i = 3 à T-h

$$S_i = \lambda y_i + (1-\lambda)(S_{i-1} + T_{i-1})$$
$$T_i = \mu(S_i - S_{i-1}) + (1-\mu)T_{i-1}$$
$$Mape = Mape + abs(y_i - S_{i-1} - T_{i-1}) * 100/y_i$$

Fin_pour

$$Mape = Mape/(T-h-2)$$

Optimisation de la fonction Mape par la procédure optim

Calcul du modèle sur l'horizon d'ajustement : 1,…, T-h

Calcul des prévisions pour les h derniers instants :

Pour i = 1 à h

$$\hat{y}_{T-h+i} = S_{T-h} + iT_{T-h}$$

Fin_pour

L'implémentation f le fichier application LED..

Lissage de Holt

Série non saisonnière

Il est inspiré du lissage double , sauf qu'ici on a deux parametres de lissage un pour S et un autre pour T

Initialisation:

$$\lambda = 0.4, \mu = 0.3, S_2 = \frac{y_1 + y_2 + y_3}{3}, T_2 = \frac{y_3 - y_1}{2}$$

** implémentation pareille à ce qui précède la même logique

** Les paramètres de lissage sont toujours entre 0 et 1

Les paramètres du modèles sont : $\boxed{S_2, T_2, \lambda \text{ et } \mu}$

On choisit le MAPE comme critère à minimiser pour estimer le modèle, c'est une fonction dans ce cas à 4 variables

Initialisation des paramètres : $S_2 = x[1], T_2 = x[2], \lambda = x[3]$ et $\mu = x[4]$

Mape = 0
Pour i = 3 à T-h

$$S_i = \lambda y_i + (1-\lambda)(S_{i-1} + T_{i-1})$$
$$T_i = \mu(S_i - S_{i-1}) + (1-\mu)T_{i-1}$$
$$Mape = Mape + abs(y_i - S_{i-1} - T_{i-1}) * 100/y_i$$

Fin_pour

$$Mape = Mape/(T-h-2)$$

Optimisation de la fonction Mape par la procédure optim

Calcul du modèle sur l'horizon d'ajustement : 1,…, T-h

Calcul des prévisions pour les h derniers instants :
Pour i = 1 à h

$$\hat{y}_{T-h+i} = S_{T-h} + iT_{T-h}$$

Fin_pour

Lissage de Holt-winters

Série saisonnière

Modele additif

Les paramètres du modèle sont :

S_t : Le niveau

T_t : La pente

I_t : La composante saisonnière

s : La période de la composante saisonnière

(s = 12 données mensuelles, s = 4 données trimestrielles)

Rq : l'ajustement ne peut etre calculé qu'a partir de la deuxieme année

α, γ, δ

** implémentation f le dossier application holt winters

**C'est le modèle le plus utilisé , et chwia m39ed 3la lokhrin

Les équations de mise à jour du **modèle additif** :

$$S_t = \alpha(y_t - I_{t-s}) + (1-\alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1}$$

$$I_t = \delta(y_t - S_t) + (1-\delta)I_{t-s}$$

La série ajustée : $F_{t-1} = S_{t-1} + T_{t-1} + I_{t-s}$

L'erreur d'ajustement : $e_t = y_t - F_{t-1}$

La prévision à la date T, pour l'horizon h est :

$$\hat{y}_T(h) = S_T + hT_T + \hat{I}_{T+h}$$

$\hat{I}_{T+h} = I_{T+h-s}$ si $0 < h \leq s$

$\hat{I}_{T+h} = I_{T+h-2s}$ si $s < h \leq 2s$, et ainsi de suite .

Lissage de Holt-winters

Modele additif

Obs	YearMont	Sweetwhit	MMC(12)	comp sais	comp corrigé	I0
1	1980-01	85				0,862681
2	1980-02	89				0,814049
3	1980-03	109				0,938607
4	1980-04	95				0,860841
5	1980-05	91				0,821775
6	1980-06	95				0,776152
7	1980-07	96	114,4583	0,83873316	1,03809671	1,038097
8	1980-08	128	118,9583	1,07600701	1,10510907	1,105109
9	1980-09	124	122,6667	1,01086957	1,041717	1,041717
10	1980-10	111	126,125	0,88007929	1,0109096	1,01091
11	1980-11	178	127,6667	1,39425588	1,24526694	1,245267
12	1980-12	140	127,625	1,09696376	1,48479696	1,484797
13	1981-01	150	128,125	1,17073171	0,86268062	
14	1981-02	132	129,7917	1,01701445	0,81404864	
15	1981-03	155	130,875	1,1843362	0,93860708	
16	1981-04	132	131,625	1,002849	0,86084082	
17	1981-05	91	131,25	0,69333333	0,82177473	
18	1981-06	94	131,0417	0,71732909	0,77615184	

Calcul des valeurs initiales :

$$S_2 = \frac{(y_3 - I_3) + (y_2 - I_2) + (y_1 - I_1)}{3}, \quad T_2 = \frac{(y_3 - I_3) - (y_1 - I_1)}{2}$$

Pour une série trimestrielle

$$T_4 = S_4 - S_3$$

$$I_4 = y_4 - S_4,\ I_3 = y_3 - S_3,\ I_2 = y_2 - (S_3 - T_4),\ I_1 = y_1 - (S_3 - 2T_4)$$

** les S3 et S4 sont obtenues par MMC(4) donc c'est la premiere valeur qu'on peut calculer

Pour (I) : elle est calculée sur excel

Calculer les moyennes mobiles centrées sur une année donnée

Déduire les composantes saisonnières en retranchant les moyennes mobiles centrées des valeurs de la série

** yla kna hsebna lcomp.sais corrigée ghaykon hsen de l'utiliser

Lissage de Holt-winters

Modele multiplicatif

$$S_t = \alpha \frac{y_t}{I_{t-s}} + (1-\alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)T_{t-1}$$

$$I_t = \delta \frac{y_t}{S_t} + (1-\delta)I_{t-s}$$

La série ajustée : $F_{t-1} = (S_{t-1} + T_{t-1})I_{t-s}$

L'erreur d'ajustement : $e_t = y_t - F_{t-1}$

Initialisation

$$S_2 = \frac{1}{3}\left(\frac{y_3}{I_3} + \frac{y_2}{I_2} + \frac{y_1}{I_1}\right), \quad T_2 = \frac{1}{2}\left(\frac{y_3}{I_3} - \frac{y_1}{I_1}\right)$$

Pour une série trimestrielle

$$T_4 = S_4 - S_3$$

$$I_4 = \frac{y_4}{S_4}, I_3 = \frac{y_3}{S_3}, I_2 = \frac{y_2}{S_3 - T_4}, \quad I_1 = \frac{y_1}{S_3 - 2T_4}$$

Prévisions

$$\hat{y}_T(h) = (S_T + hT_T)\hat{I}_{T+h}$$

$\hat{I}_{T+h} = I_{T+h-s}$ si $0 < h \leq s$

$\hat{I}_{T+h} = I_{T+h-2s}$ si $s < h \leq 2s$, et ainsi de suite …

TP 3

Vous pouvez voir également le fichier word ..
pour l’algorithme des deux derniers modeles

06

Méthode de Gardner

** hadchi se fait sur excel

Principe :

- On applique des transformations sur la série
- On calcule la variance de chaque transformation
- On choisit la transformation ayant la plus petite variance
- On choisit le modèle qui correspond à la transformation choisie a partir de ce tableau

Cas	Série ayant la plus petite variance	Modèle
A	X_t	Lissage simple
B	$(1-B)X_t$	DA-N
C	$(1-B)^2X_t$	Lissage de Holt
D	$(1-B^s)X_t$	N-M
E	$(1-B)(1-B^s)X_t$	DA-M
F	$(1-B^2)(1-B^s)X_t$	H-W multiplicatif

s : La période de la composante saisonnière

(s = 12 données mensuelles, s = 4 données trimestrielles)

$$BX_t = X_{t-1},\ B^2X_t = X_{t-2},\ B^sX_t = X_{t-s}$$

Equations des transformations :

$$(1 - B)X_t = X_t - X_{t-1}$$

$$(1 - B)^2X_t = X_t - 2X_{t-1} + X_{t-2}$$

$$(1 - B^s)X_t = X_t - X_{t-s}$$

$$(1 - B)(1 - B^s)X_t = (1 - B^s - B + B^{s+1})X_t = X_t - X_{t-s} - X_{t-1} + X_{t-s-1}$$

$$(1 - B^2)(1 - B^s)X_t = (1 - B^s - B^2 + B^{s+2})X_t = X_t - X_{t-s} - X_{t-2} + X_{t-s-2}$$

Le modèle DA-N : damped additive none

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + \varphi T_{t-1})$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\varphi T_{t-1}$$

La série ajustée : $F_{t-1} = S_{t-1} + \varphi T_{t-1}$

$$\hat{X}_t(h) = S_t + \sum_{i=1}^{h} \varphi^i T_t$$

Le modèle N-M : none trend multiplicative

$$S_t = \alpha \frac{X_t}{I_{t-s}} + (1-\alpha) S_{t-1}$$

$$I_t = \delta \frac{X_t}{S_t} + (1-\delta) I_{t-s}$$

$$\hat{X}_t(h) = S_t I_{t+h-s}$$

La série ajustée

$$F_{t-1} = S_{t-1} I_{t-s}$$

Le modèle DA-M : damped additive multiplicative

$$S_t = \alpha X_t / I_{t-s} + (1-\alpha)(S_{t-1} + \varphi T_{t-1})$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\varphi T_{t-1}$$

$$I_t = \delta(X_t/S_t) + (1-\delta) I_{t-s}$$

$$\hat{X}_t(h) = \left(S_t + \sum_{i=1}^{h} \varphi^i T_t\right) I_{t+h-s}$$

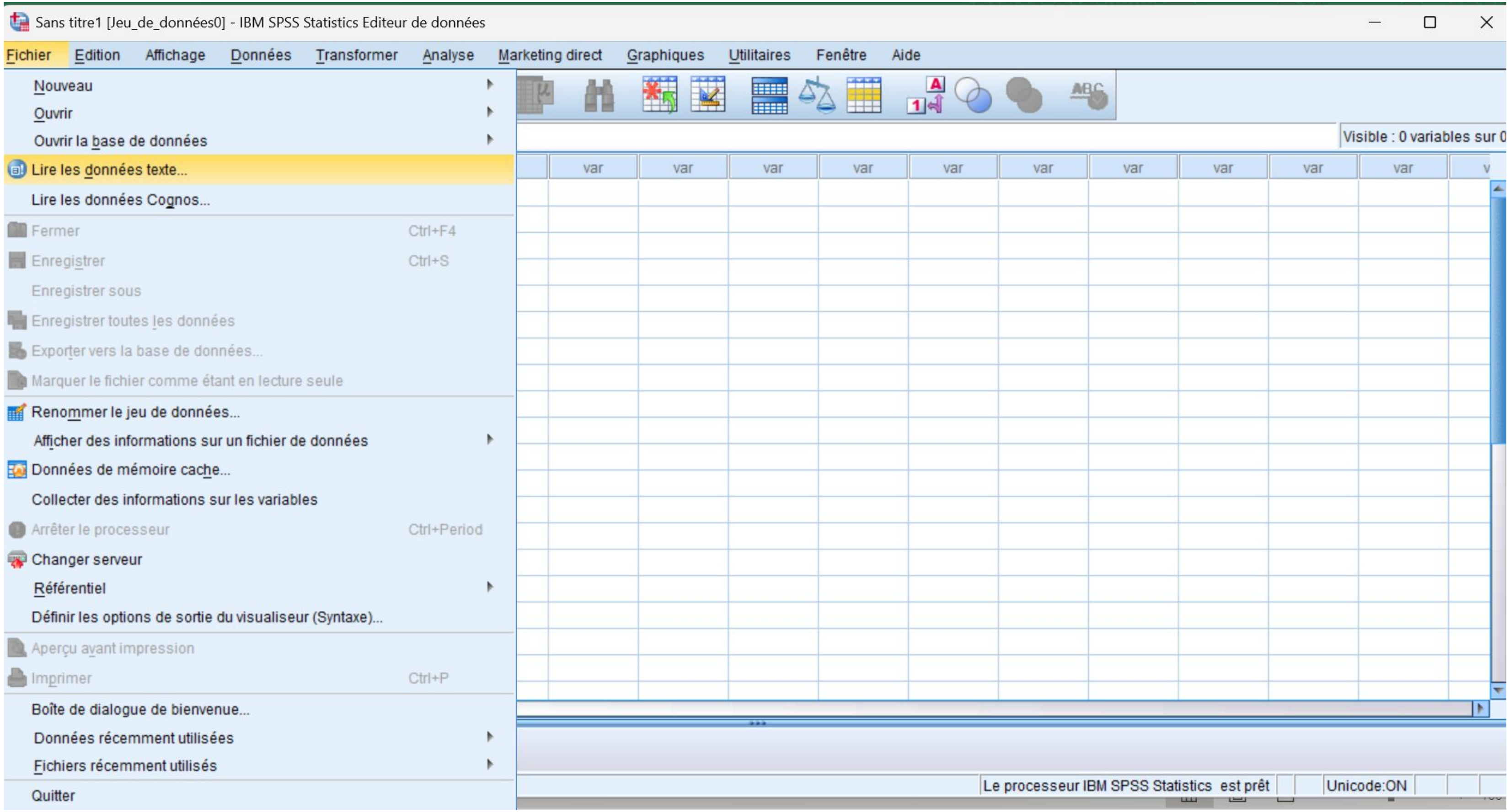
La série ajustée : $F_{t-1} = (S_{t-1} + \varphi T_{t-1}) I_{t-s}$

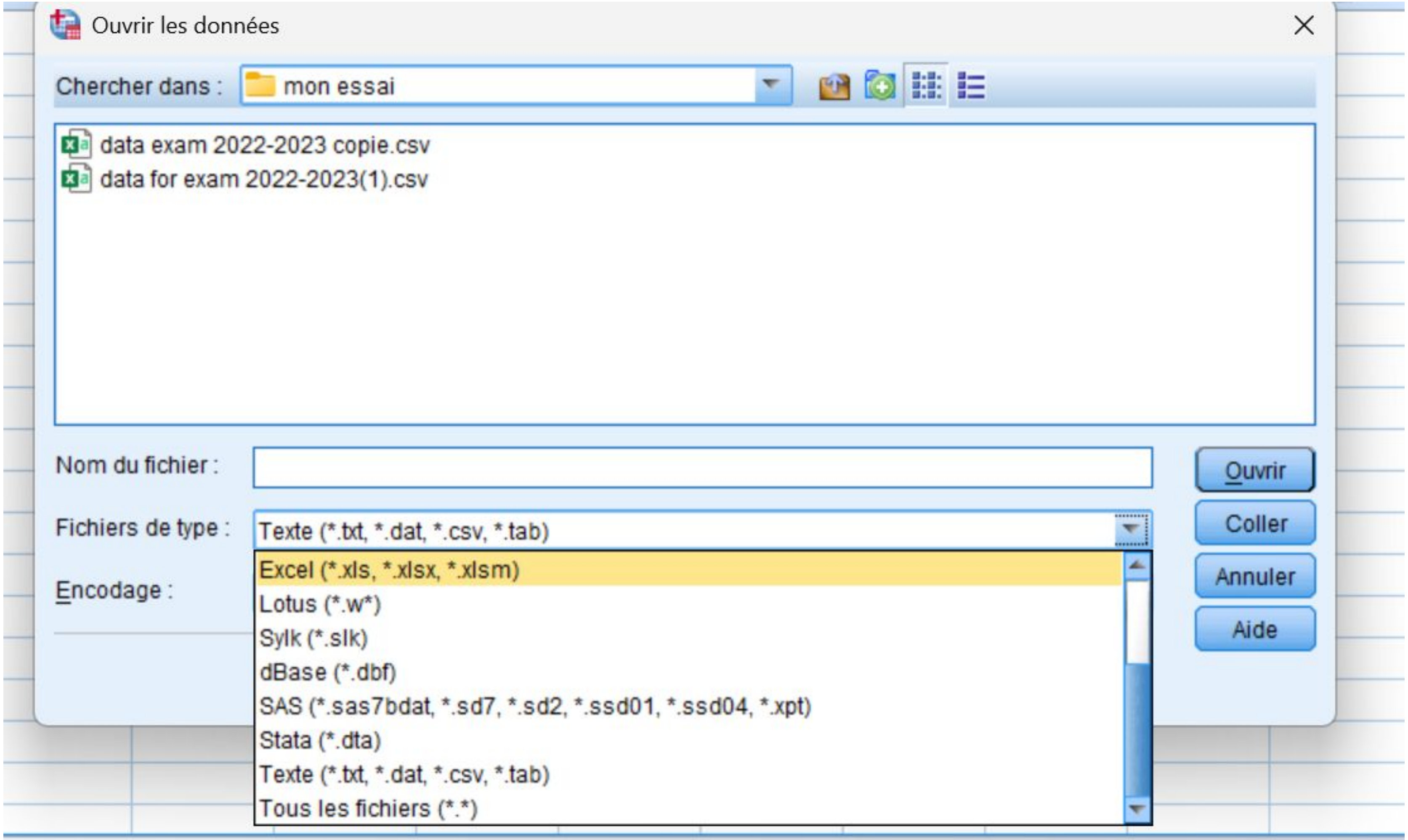
Pour tous ces modèles, l'erreur d'ajustement est définie par :
$e_t = X_t - \hat{X}_{t-1}(1) = X_t - F_{t-1}$

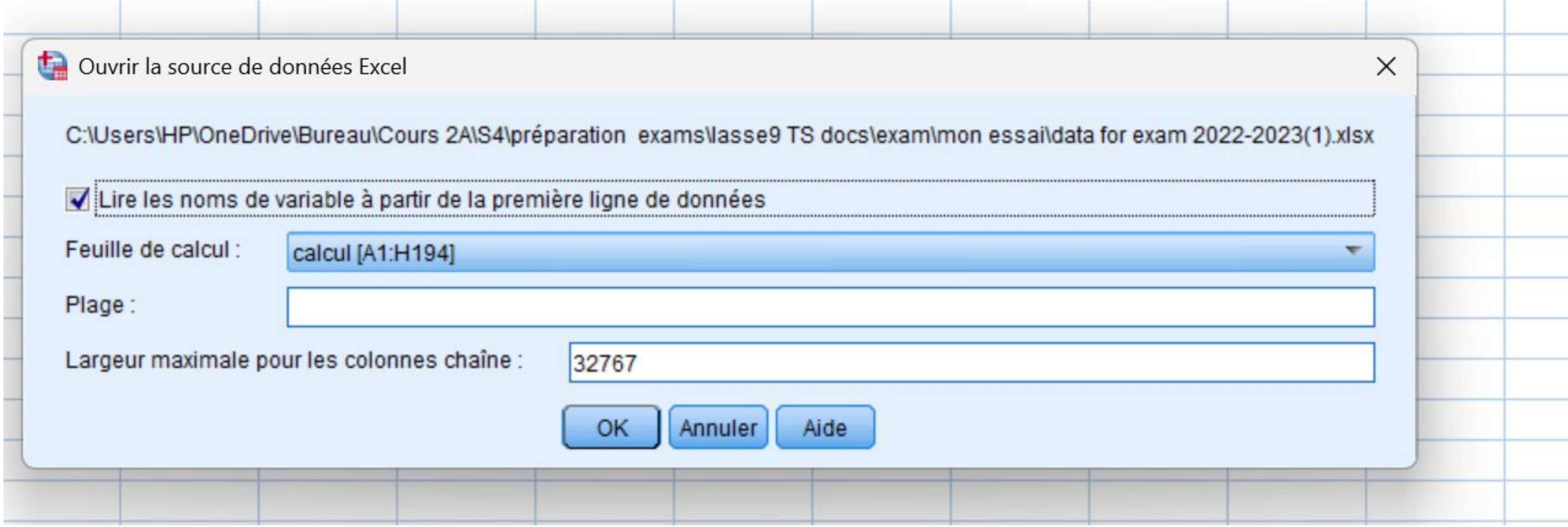
TP 4

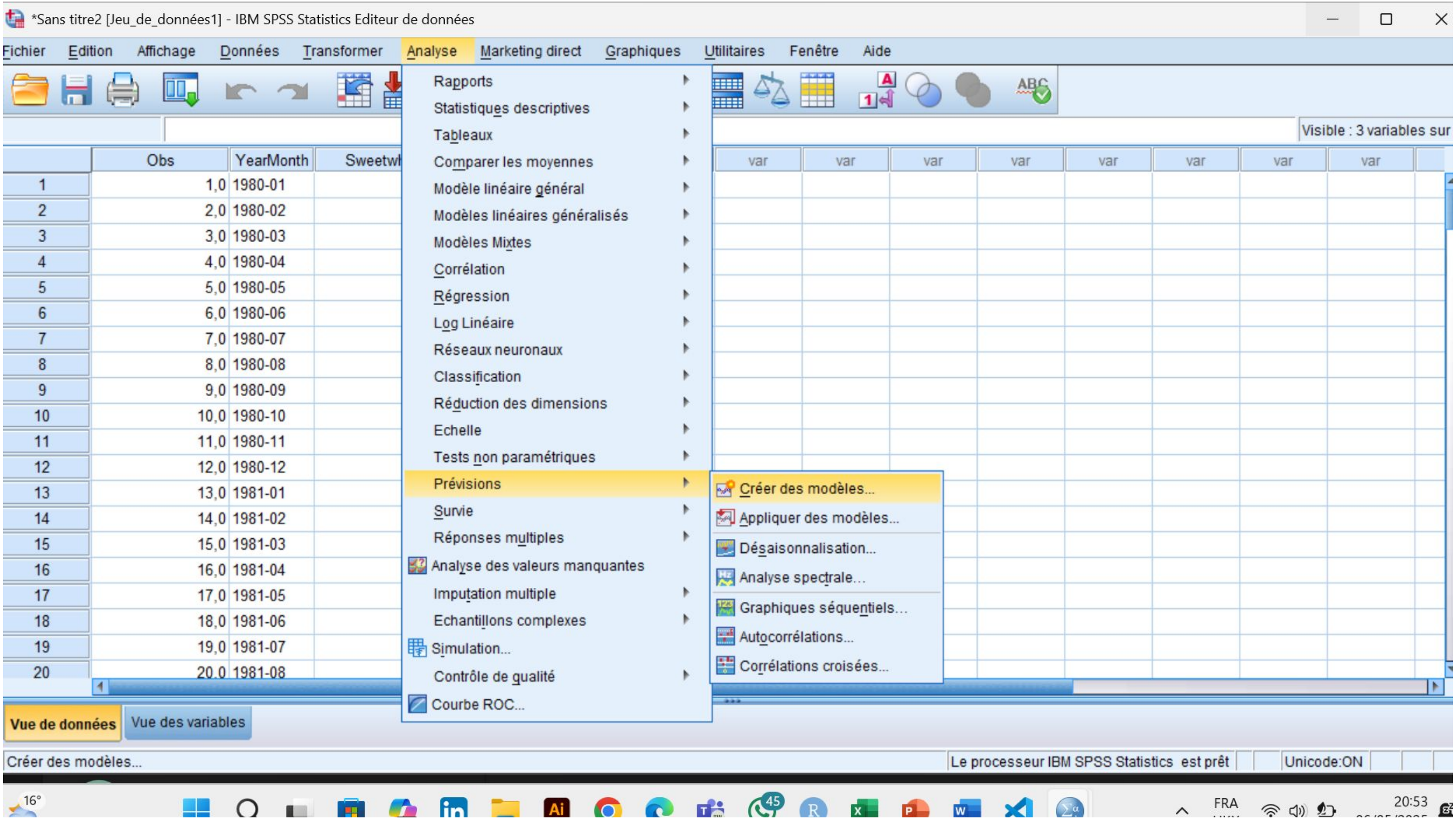
SPSS

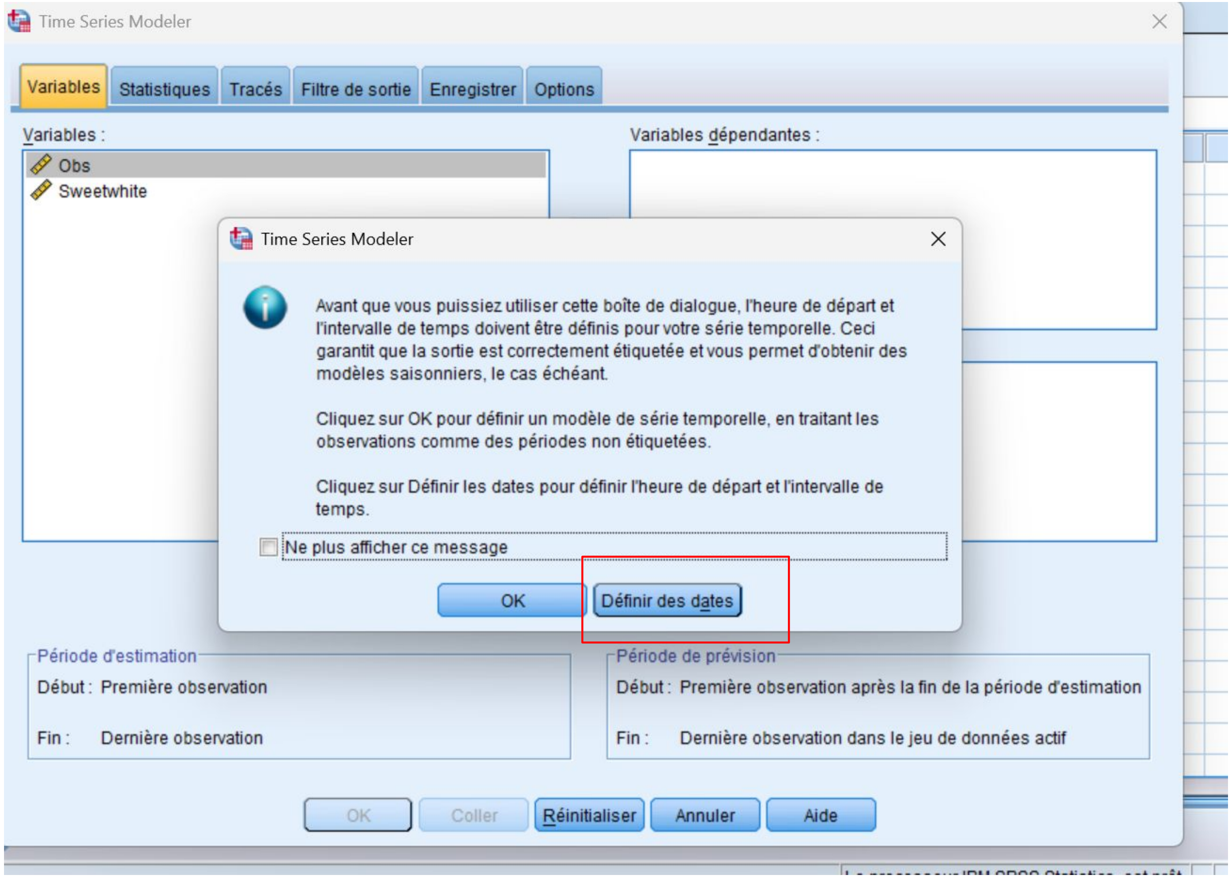
De préférence les données ykono f fichier excel



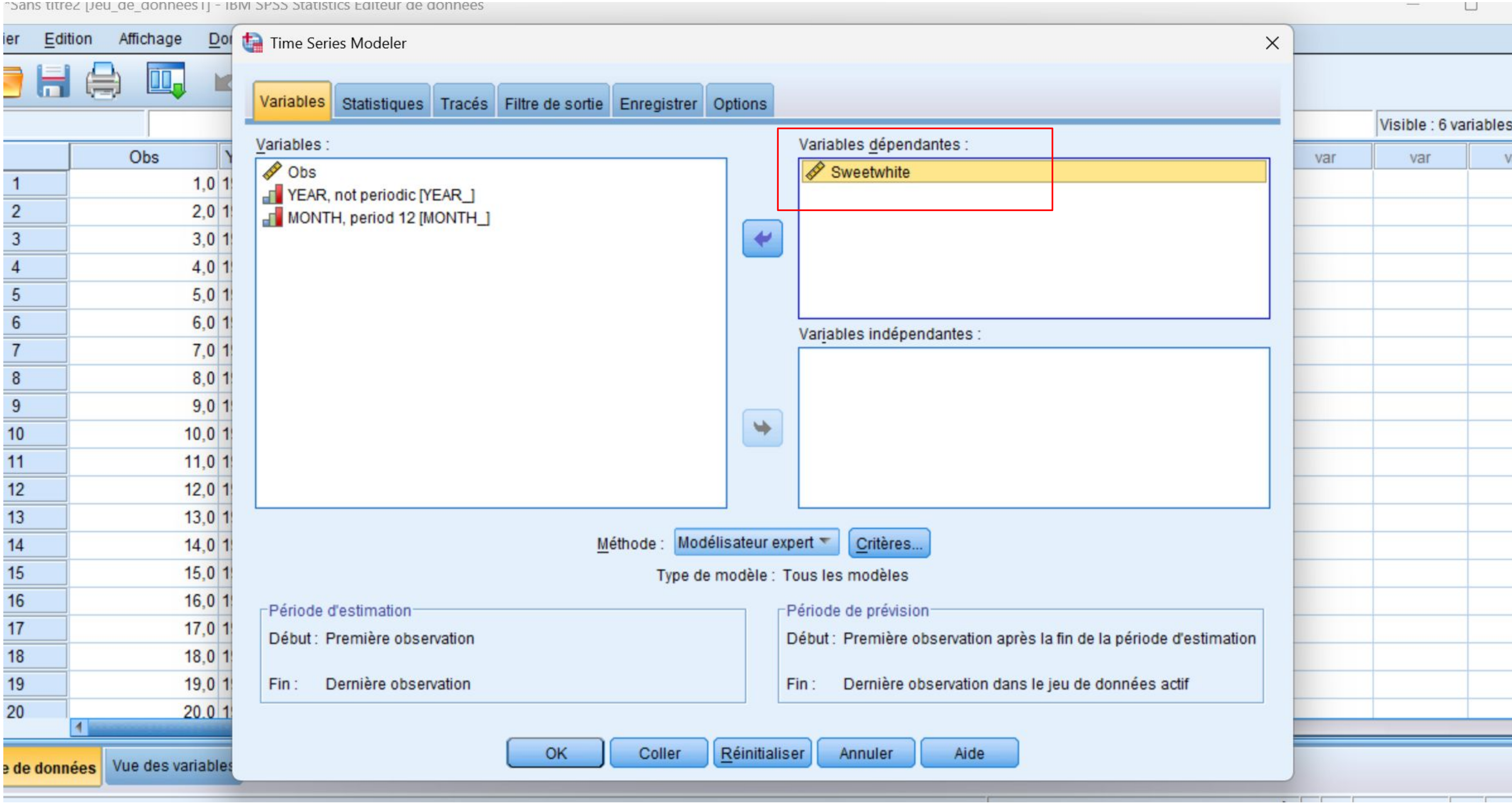

Ouvrir les données
Chercher dans :
mon essai
data exam 2022-2023 copie.csv
data for exam 2022-2023(1).csv
Nom du fichier :
Fichiers de type :
Texte (*.txt, *.dat, *.csv, *.tab)
Encodage :
Excel (*.xls, *.xlsx, *.xlsm)
Lotus (*.w*)
Sylk (*.slk)
dBase (*.dbf)
SAS (*.sas7bdat, *.sd7, *.sd2, *.ssd01, *.ssd04, *.xpt)
Stata (*.dta)
Texte (*.txt, *.dat, *.csv, *.tab)
Tous les fichiers (*.*)
Ouvrir
Coller
Annuler
Aide


Ouvrir la source de données Excel
C:\Users\HP\OneDrive\Bureau\Cours 2A\S4\préparation exams\lasse9 TS docs\exam\mon essai\data for exam 2022-2023(1).xlsx
Lire les noms de variable à partir de la première ligne de données
Feuille de calcul :
calcul [A1:H194]
Plage :
Largeur maximale pour les colonnes chaîne :
32767
OK
Annuler
Aide

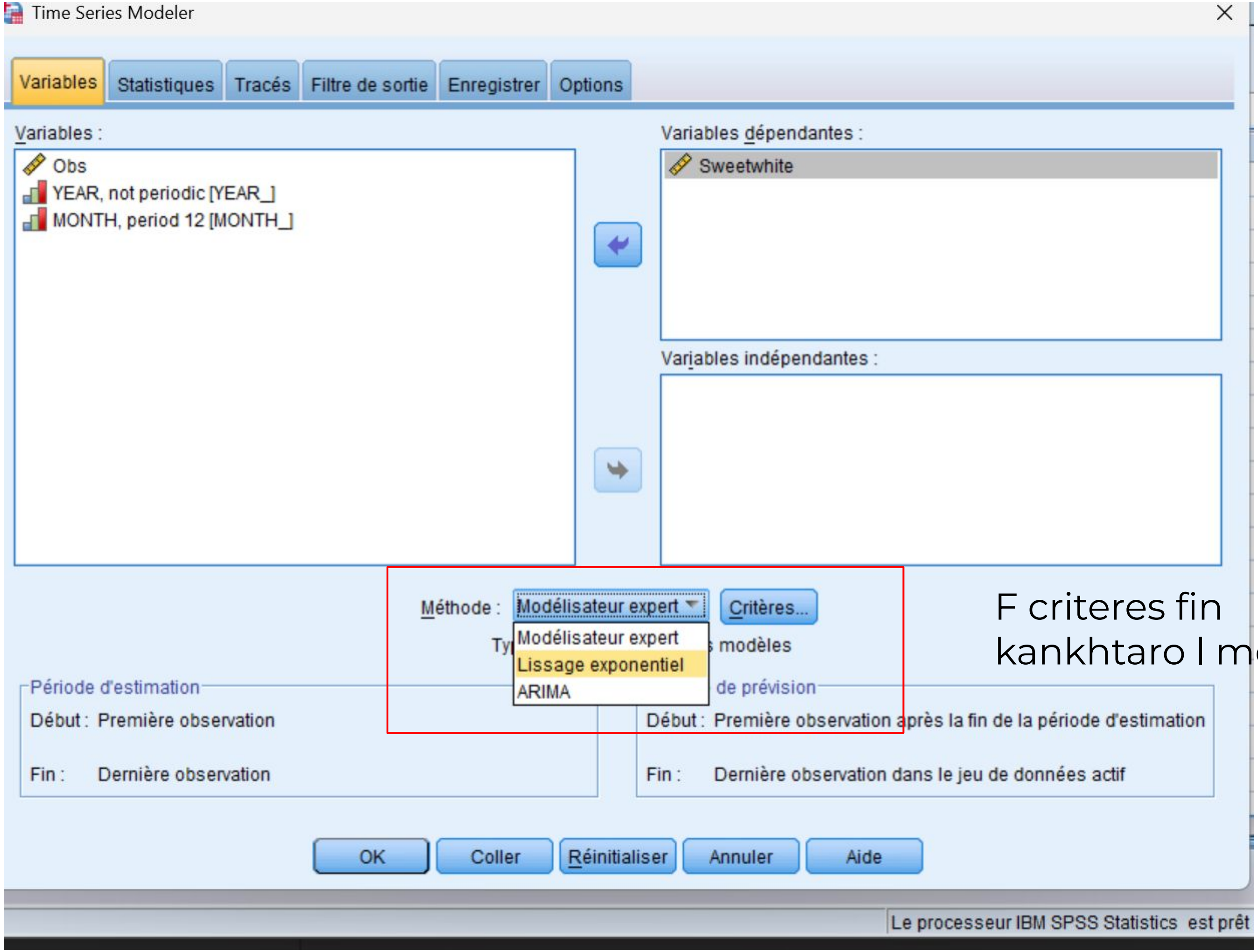

*Sans titre2 [Jeu_de_données1] - IBM SPSS Statistics Editeur de données
Fichier Edition Affichage Données Transformer Analyse Marketing direct Graphiques Utilitaires Fenêtre Aide
Visible : 3 variables sur
Obs YearMonth Sweetwh
Rapports
Statistiques descriptives
Tableaux
Comparer les moyennes
Modèle linéaire général
Modèles linéaires généralisés
Modèles Mixtes
Corrélation
Régression
Log Linéaire
Réseaux neuronaux
Classification
Réduction des dimensions
Echelle
Tests non paramétriques
Prévisions
Survie
Réponses multiples
Analyse des valeurs manquantes
Imputation multiple
Echantillons complexes
Simulation...
Contrôle de qualité
Courbe ROC...
Créer des modèles...
Appliquer des modèles...
Désaisonnalisation...
Analyse spectrale...
Graphiques séquentiels...
Autocorrélations...
Corrélations croisées...
Vue de données
Vue des variables
Créer des modèles...
Le processeur IBM SPSS Statistics est prêt
Unicode:ON



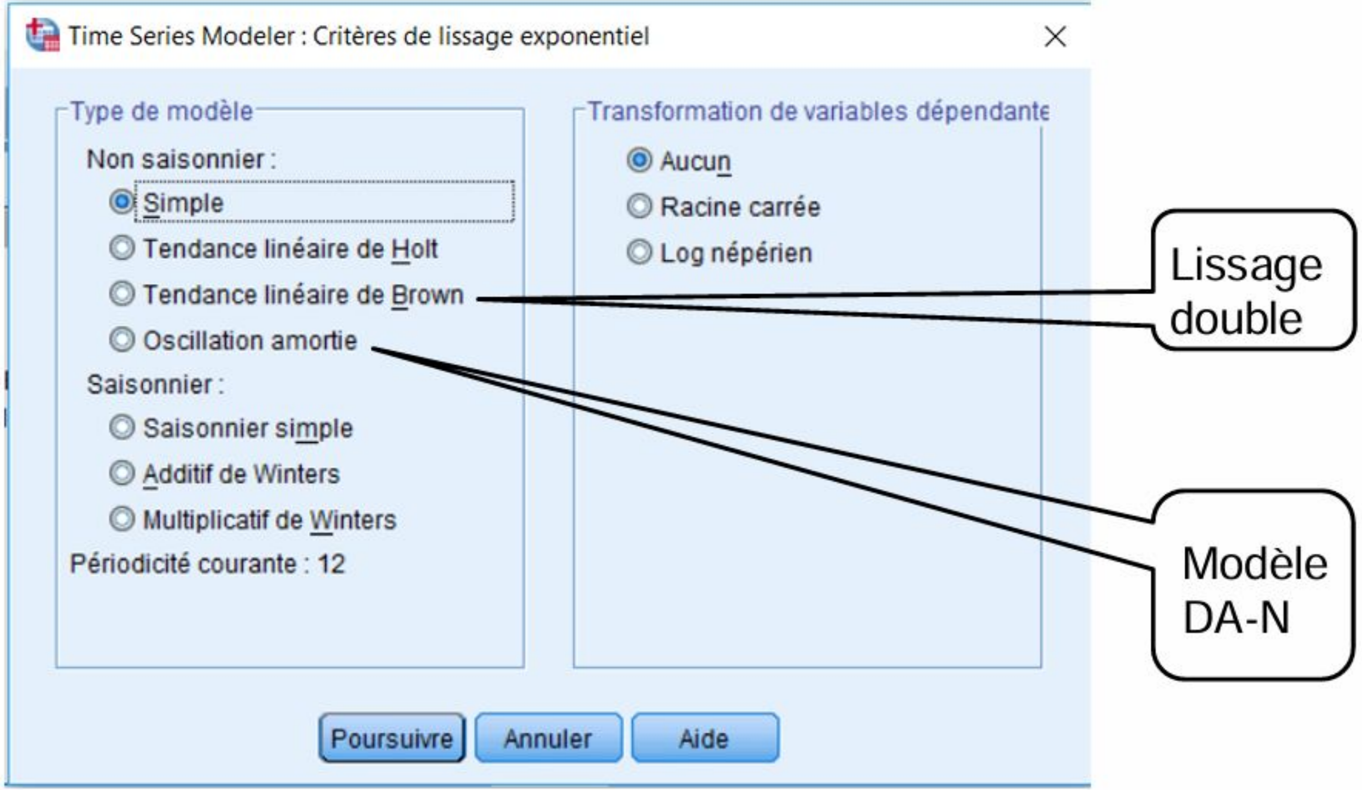
Kankhtaro Année mois yla kant mensuelle w kandéfiniw la premiere année

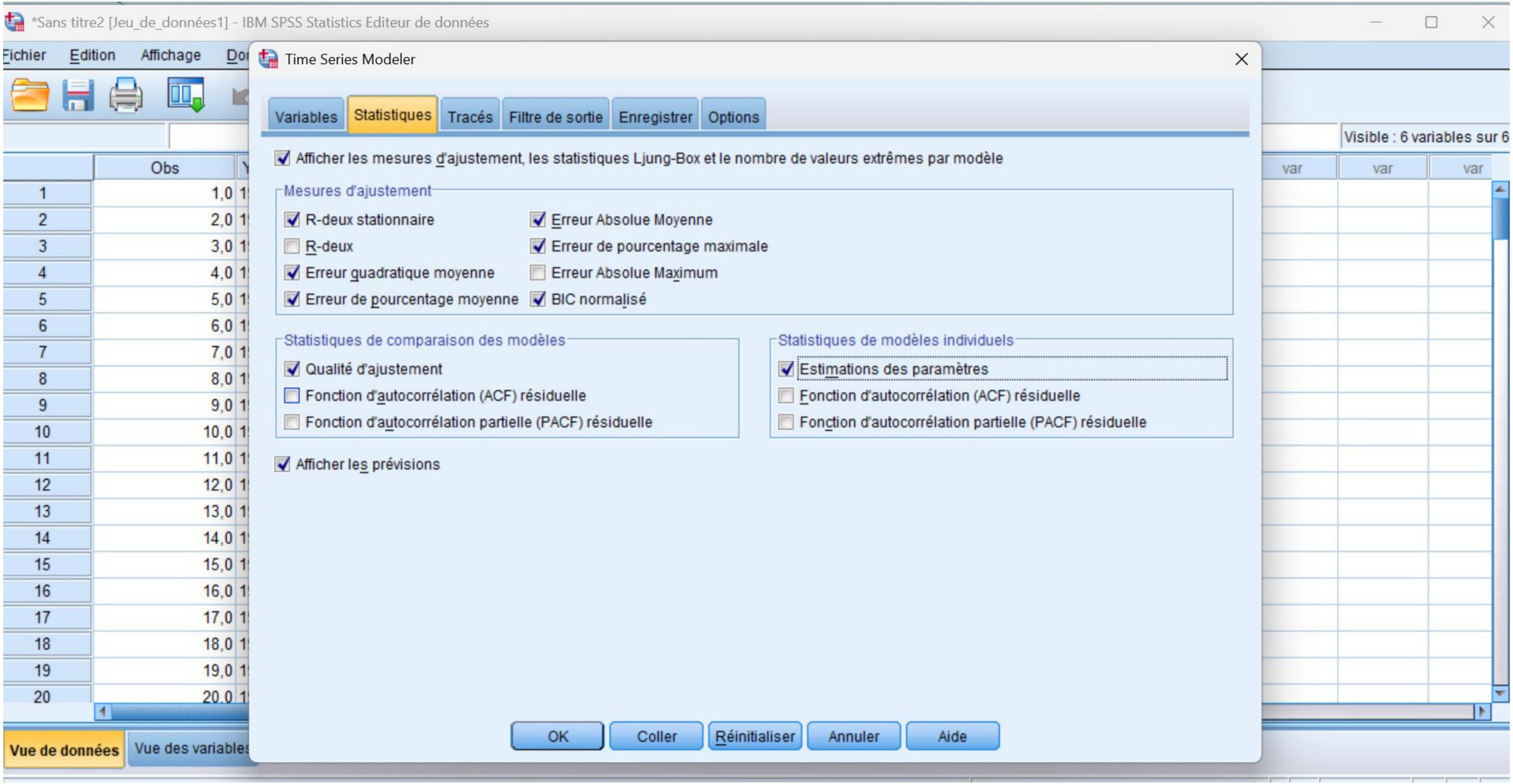


Time Series Modeler
Edition
Affichage
Variables
Statistiques
Tracés
Filtre de sortie
Enregistrer
Options
Visible : 6 variables
Variables :
Obs
YEAR, not periodic [YEAR_]
MONTH, period 12 [MONTH_]
Variables dépendantes :
Sweetwhite
Variables indépendantes :
Méthode : Modélisateur expert
Critères...
Type de modèle : Tous les modèles
Période d'estimation
Début : Première observation
Fin : Dernière observation
Période de prévision
Début : Première observation après la fin de la période d'estimation
Fin : Dernière observation dans le jeu de données actif
OK
Coller
Réinitialiser
Annuler
Aide
Vue des variables



F criteres fin kankhtaro l modele


Time Series Modeler : Critères de lissage exponentiel
Type de modèle
Non saisonnier :
Simple
Tendance linéaire de Holt
Tendance linéaire de Brown
Oscillation amortie
Saisonnier :
Saisonnier simple
Additif de Winters
Multiplicatif de Winters
Périodicité courante : 12
Transformation de variables dépendante
Aucun
Racine carrée
Log népérien
Poursuivre
Annuler
Aide
Lissage double
Modèle DA-N


*Sans titre2 [Jeu_de_données1] - IBM SPSS Statistics Editeur de données
Fichier
Edition
Affichage
Time Series Modeler
Variables
Statistiques
Tracés
Filtre de sortie
Enregistrer
Options
Visible : 6 variables sur 6
Obs
var
Afficher les mesures d'ajustement, les statistiques Ljung-Box et le nombre de valeurs extrêmes par modèle
Mesures d'ajustement
R-deux stationnaire
R-deux
Erreur quadratique moyenne
Erreur de pourcentage moyenne
Erreur Absolue Moyenne
Erreur de pourcentage maximale
Erreur Absolue Maximum
BIC normalisé
Statistiques de comparaison des modèles
Qualité d'ajustement
Fonction d'autocorrélation (ACF) résiduelle
Fonction d'autocorrélation partielle (PACF) résiduelle
Statistiques de modèles individuels
Estimations des paramètres
Fonction d'autocorrélation (ACF) résiduelle
Fonction d'autocorrélation partielle (PACF) résiduelle
Afficher les prévisions
OK
Coller
Réinitialiser
Annuler
Aide
Vue de données
Vue des variables

*Sortie1 [Document1] - IBM SPSS Statistics Viewer

Fichier Edition Affichage Données Transformer Insérer Format Analyse Marketing direct Graphiques Utilitaires Fenêtre Aide

- Sortie
  - Log
  - Time Series Modeler
    - Titre
    - Remarques
    - Jeu de données a
    - Description du mc
    - Récapitulatif des r
      - Titre
      - Qualité de l'aj
    - Statistiques du m
    - Paramètres du m
    - Prévision
    - Graphique de séri
  - Log
  - Time Series Modeler
    - Titre
    - Remarques
    - Description du mc
    - Récapitulatif des r
      - Titre
      - Qualité de l'aj
    - Statistiques du m
    - Paramètres du m
    - Prévision
    - Graphique de séri

Time Series Modeler

Description du modèle

			Type de modèle
ID de modèle	Sweetwhite	Modèle_1	Multiplicatif de Winters

Récapitulatif des modèles

Qualité de l'ajustement

Statistiques d'ajustement	Moyenne	Erreur standard	Minimum	Maximum	Percentile 5	10	25	50	75	90	95
R-deux stationnaire	,495	.	,495	,495	,495	,495	,495	,495	,495	,495	,495
R-deux	,851	.	,851	,851	,851	,851	,851	,851	,851	,851	,851
RMSE	48,373	.	48,373	48,373	48,373	48,373	48,373	48,373	48,373	48,373	48,373
MAPE	13,868	.	13,868	13,868	13,868	13,868	13,868	13,868	13,868	13,868	13,868
MaxAPE	80,114	.	80,114	80,114	80,114	80,114	80,114	80,114	80,114	80,114	80,114
MAE	33,912	.	33,912	33,912	33,912	33,912	33,912	33,912	33,912	33,912	33,912
MaxAE	151,461	.	151,461	151,461	151,461	151,461	151,461	151,461	151,461	151,461	151,461
BIC normalisé	7,846	.	7,846	7,846	7,846	7,846	7,846	7,846	7,846	7,846	7,846

Statistiques du modèle

*Sortie1 [Document1] - IBM SPSS Statistics Viewer

hier Edition Affichage Données Transformer Insérer Format Analyse Marketing direct Graphiques Utilitaires Fenêtre Aide

- Sortie
  - Log
  - Time Series Modeler
    - Titre
    - Remarques
    - Jeu de données a
    - Description du mo
    - Récapitulatif des r
      - Titre
      - Qualité de l'aj
    - Statistiques du m
    - Paramètres du m
    - Prévision
    - Graphique de séri
  - Log
  - Time Series Modeler
    - Titre
    - Remarques
    - Description du mo
    - Récapitulatif des r
      - Titre
      - Qualité de l'aj
    - Statistiques du m
    - Paramètres du m
    - Prévision
    - Graphique de séri

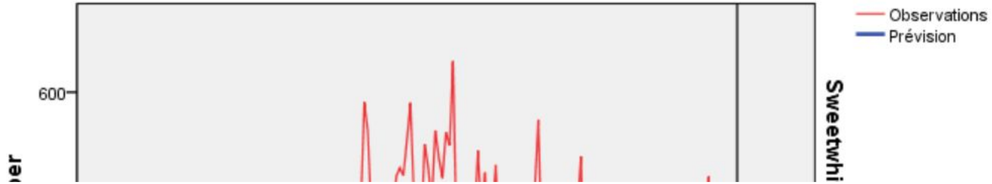
Paramètres du modèle de lissage exponentiel

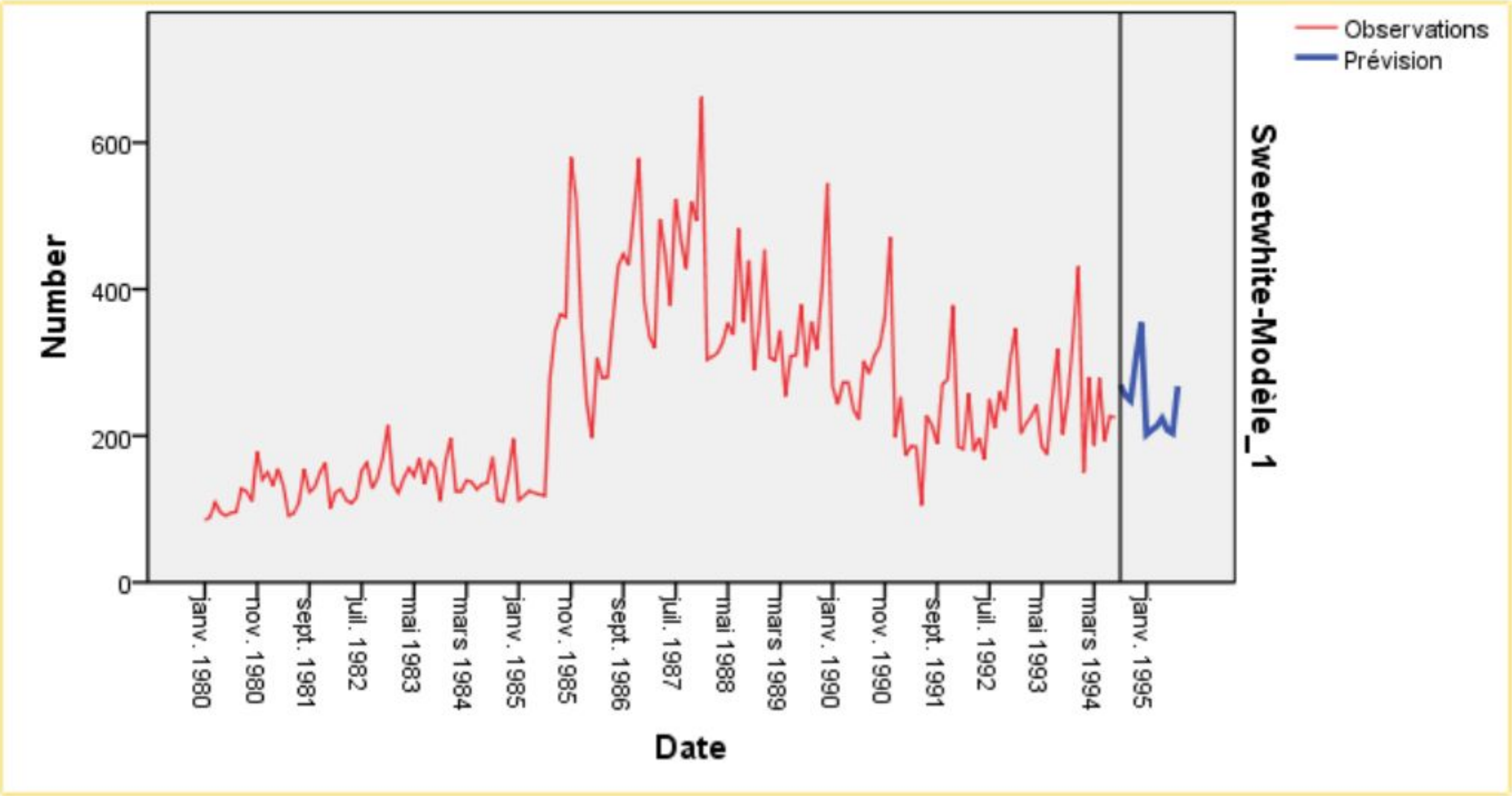
Modèle			Estimation	Erreur standard	t	Sig.
Sweetwhite-Modèle_1	Aucune transformation	Alpha (Niveau)	,481	,065	7,460	,000
		Gamma (Tendance)	,001	,019	,038	,970
		Delta (Saison)	,118	,062	1,899	,059

Prévision

Modèle		août 1994	sept. 1994	oct. 1994	nov. 1994	déc. 1994	janv. 1995	févr. 1995	mars 1995	avr. 1995	mai 1995	juin 1995
Sweetwhite-Modèle_1	Prévision	270	255	248	306	354	202	208	213	224	208	203
	UCL	365	359	361	441	509	322	339	353	375	359	358
	LCL	174	150	134	172	200	81	78	73	73	57	48

Pour chaque modèle, les prévisions commencent après la dernière valeur non manquante dans la plage de la période d'estimation requise et elles se terminent à la dernière période p
laquelle des valeurs non manquantes de tous les prédicteurs sont disponibles ou à la date de fin de la période de prévision, selon ce qui se produit en premier.





Observations
Prévision
Number
600
400
200
0
janv. 1980
nov. 1980
sept. 1981
juil. 1982
mai 1983
mars 1984
janv. 1985
nov. 1985
sept. 1986
juil. 1987
mai 1988
mars 1989
janv. 1990
nov. 1990
sept. 1991
juil. 1992
mai 1993
mars 1994
janv. 1995
Date
Sweetwhite-Modèle_1

Exam
2022-2023

HADA ma kan !!