Chapitre 3
Time series

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

Aperçu du chapitre

Chapitre sur l'analyse des séries chronologiques

1. Décomposition d'une série temporelle
2. Détermination de la tendance
3. Identification de la composante saisonnière
4. Série désaisonnalisée (série SA)
5. Prévisions

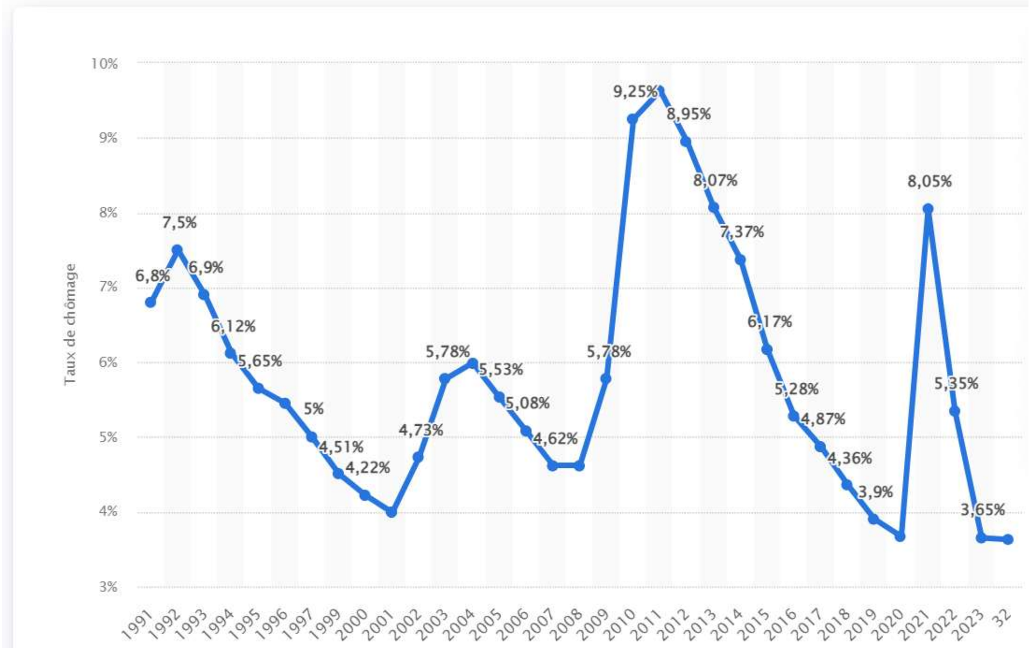
Introduction

- Evolution d'une variable dans le temps : séries temporelles. Série d'ob chaque observation est mesurée à un moment donné.
- L'ordre est important : spécificité des séries temporelles.
- Comment mesurons-nous le temps ?
  - Temps calendaire : jours, mois, trimestres, années...

Example 1

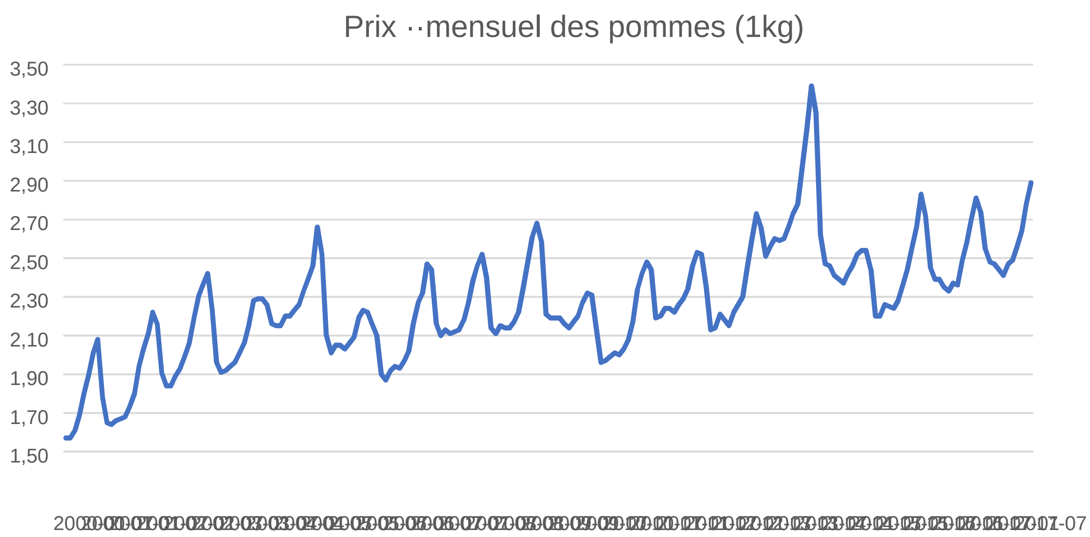
- Exemple 1 : américain taux de chômage (% de la population active), mesuré à la fin de chaque année.
- Borné (entre 0 et 100%)
- Variations dues à la economic cycle

Taux de chômage aux États-Unis entre 1991 et 2023



Example 2

- Exemple 2 : prix mensuel des pommes (1kg), en euros.
- Variations dues aux saisons : prix élevé en été, faible en hiver.
- Présence d'une tendance.



Ce que nous avons appris de ces 2 exemples

- Les séries temporelles sont fonctions du temps : séries statistiques biv
- Example 1: A cyclical variable ·uctua·ng around a stable level · stable central tendency · stationary variable.
- Example 2: The central tendency appears to be increasing · non-variable stationnaire.
- Il n'est pas toujours facile de distinguer clairement la nature et le réel évolution de la variable.

Objectifs

- Analyser une série temporelle : nous avons différents objectifs :
    - Décrire l'évolution d'une variable dans le temps
    - Interpréter cette évolution
    - Prédire son évolution future
- Décomposition de l'évolution globale

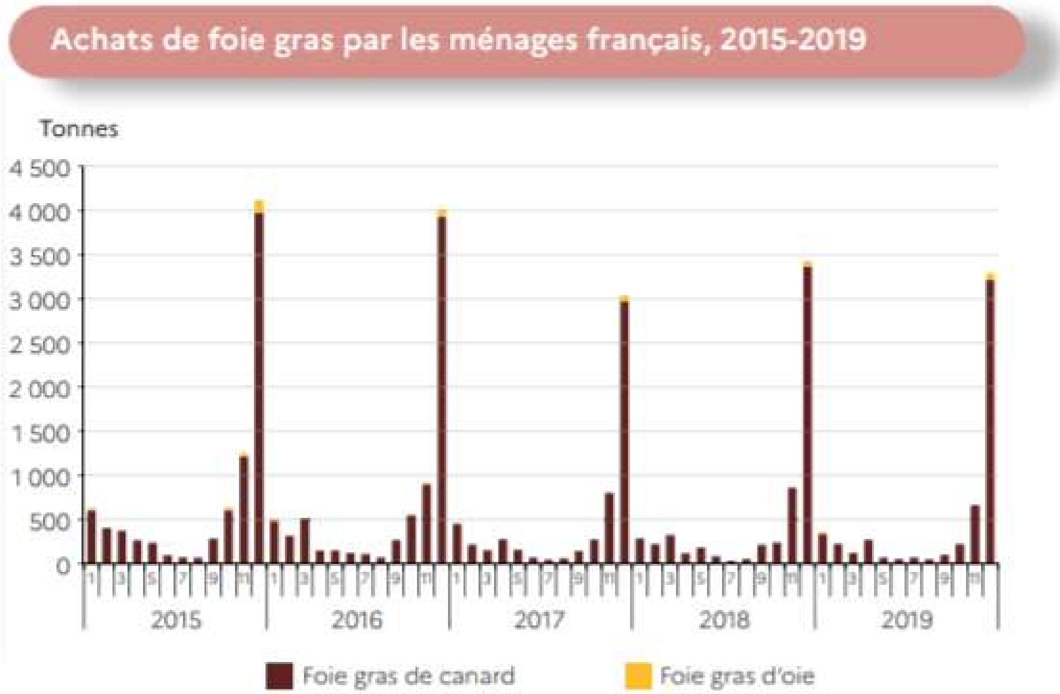
1.Décomposition d'une série temp

a) Différents éléments

b) Modèles additifs et multiplicatifs

1- a) Différents éléments

- Ex : Vente de foie gras dans un supermarché

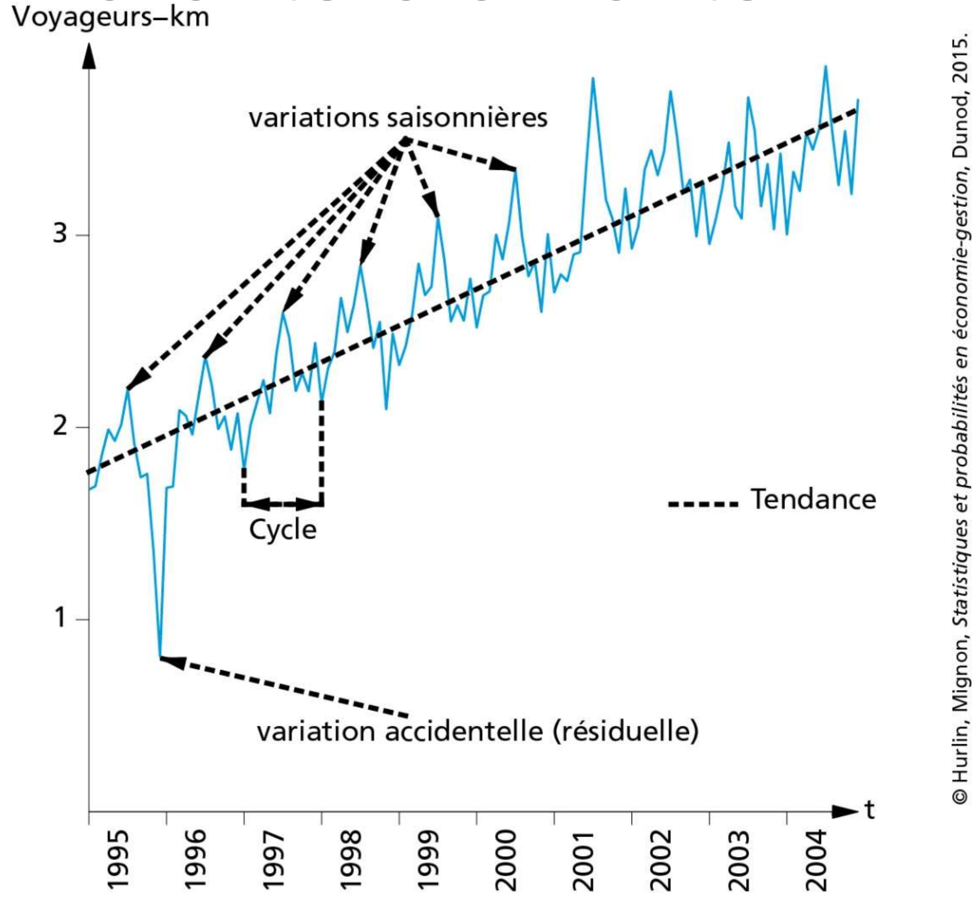


Source : FranceAgriMer d'après Kantar Worldpanel

1- a) Différents éléments

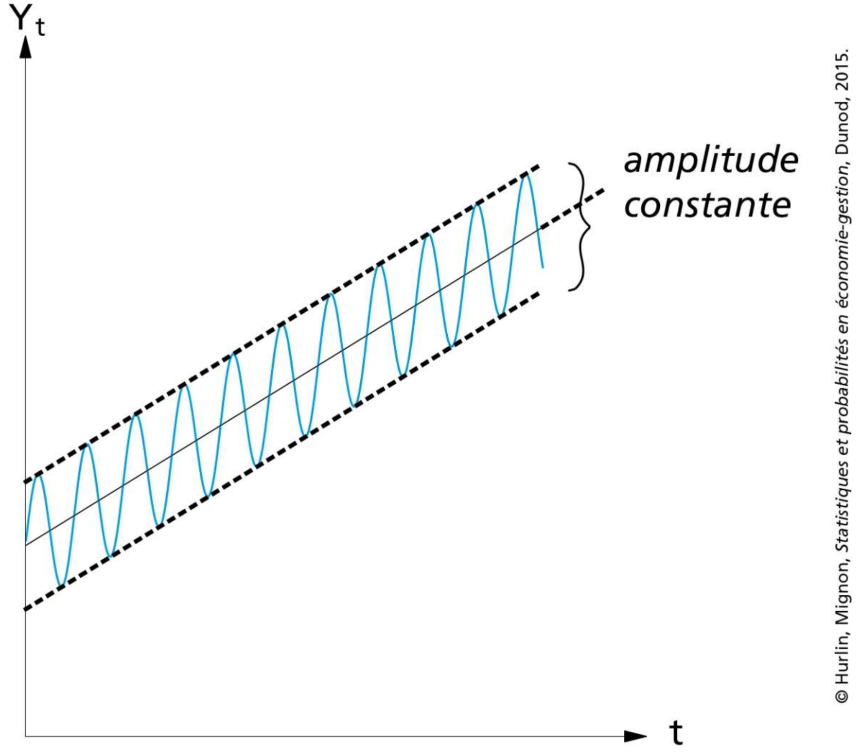
- 3 éléments :
  - La tendance : ··
  - La composante saisonnière : ··
  - L'erreur (bruit) : ··
  - Finalement, nous avons aussi une composante cyclique
- Deux modèles :
  - Modèle additif
  - Modèle multiplicatif

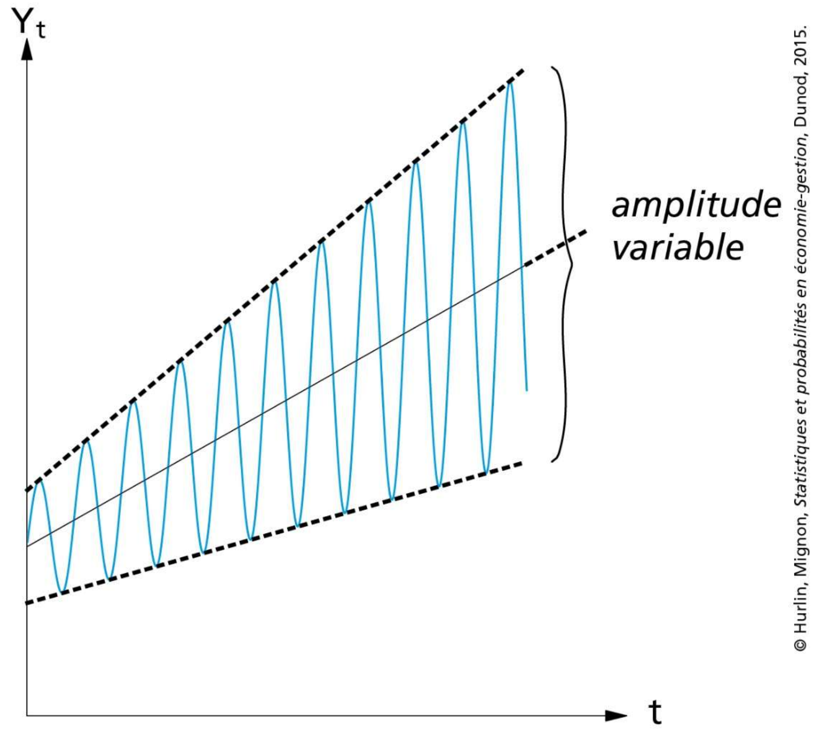
1- a) Différents éléments



Nombre de voyageurs en TGV trains entre janvier 1995 et décembre 2005, mensuel data

1- b) Modèles additifs et multiplicatifs





Modèle additif : Amplitude constante des variationsModèle multiplicatif : amplitude variable des variations

b) Modèles additifs et multiplicatifs

· Modèle additif :

··= ··+ ··+ ··

Hypothèse : les éléments sont indépendants les uns des autres.

Correspond à une suite arithmétique :

··= ····+ ·

· Modèle multiplicatif :

··= ··× ··× ··

Hypothèse : les éléments dépendent les uns des autres.

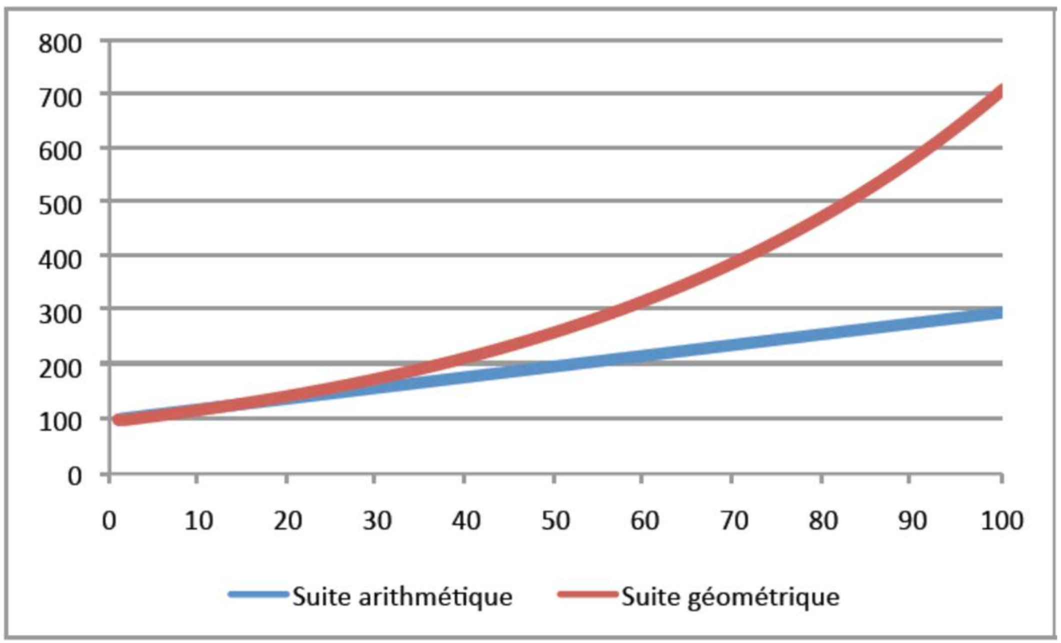
Correspond à une suite géométrique :

··= ····× ·

1- b) Modèles additifs et multiplicatifs

· Suite arithmétique : ici, ··= ····+ 2

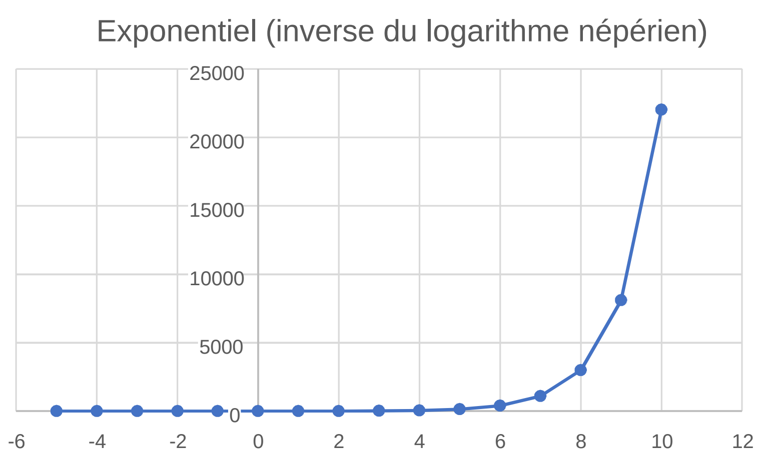
· Geometric sequence: here, ··= ····× 1,02

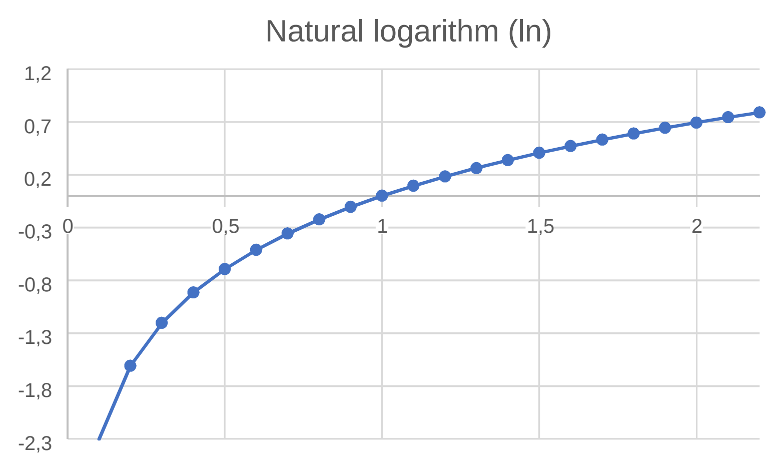


Modèles additifs et multiplicatifs

- Le modèle multiplicatif peut être converti en modèle additif en appliquer une transformation logarithmique naturelle :

··= ··× ··× ··becomes ln ··= ln ··+ ln ··+ ln ··





Modèles additifs et multiplicatifs

- Comment choisir entre le modèle additif et le modèle multiplicatif :
- Graphiquement : on regarde si la variable a une amplitude constante d variations (additif) ou croissant/décroissant (multiplicatif)
- Algébriquement : modèle Achats-Ballot. Nous examinons les différenc moyennes (par an par exemple) et les écarts types (chaque année également), qui représentent les écarts avec la moyenne. Ensuite nou calculer les coefficients de variations (écart type / moyennes).
  - Si les coefficients de variation sont stables dans le temps, le modèle additif est préféré, autre modèle multiplicatif.
- Nous étudierons principalement le modèle additif.

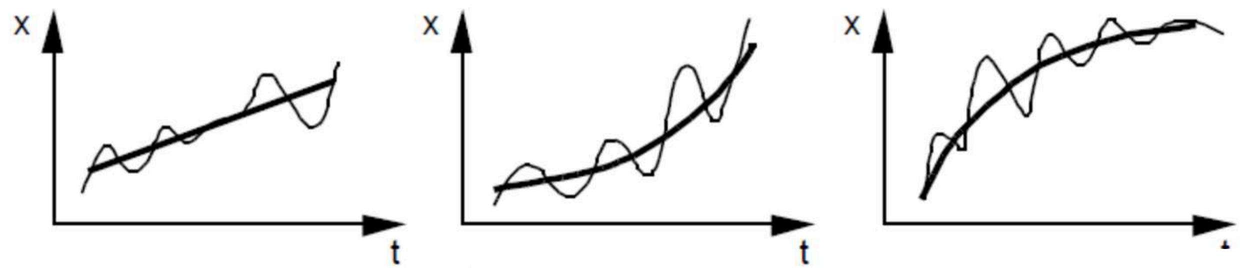
2- Détermination de la tendance

a) Moindres carrés ordinaires

b) Moyennes mobiles

2- a) Moindres carrés ordinaires (OLS)

- Ajustement paramétrique de la tendance : basé sur une inspection vis
  série, nous choisissons une forme de courbe (par exemple linéaire, qu
  puis ajustez cette courbe aux données.



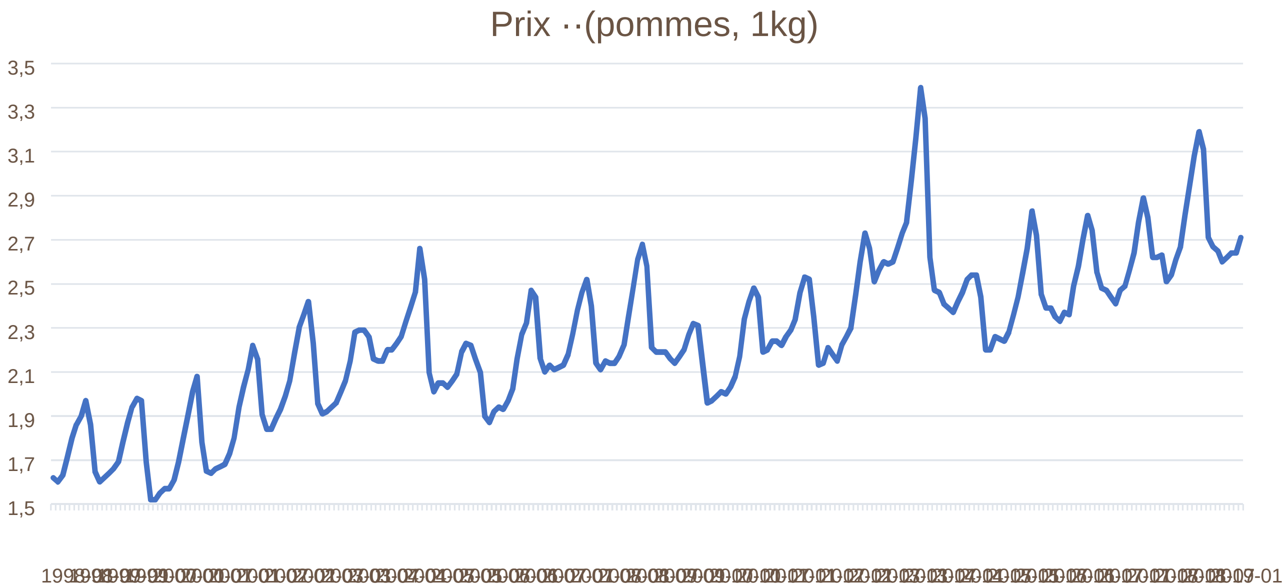
Tendance linéaire $g_t = a \cdot t + b$

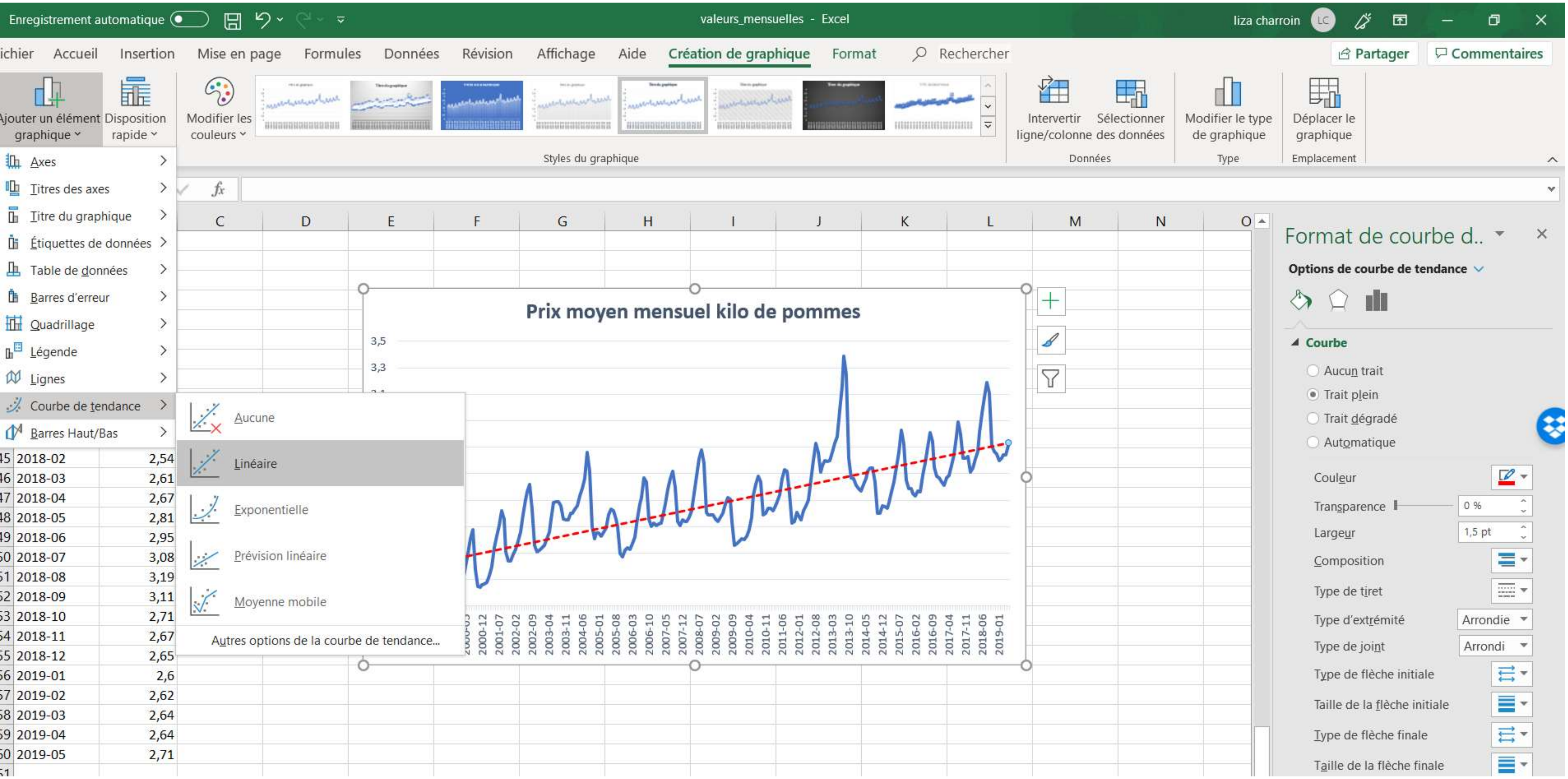
Exponential trend $g_t = b \cdot a^t$

Tendance logarithmique $g_t = \ln(a \cdot t + b)$

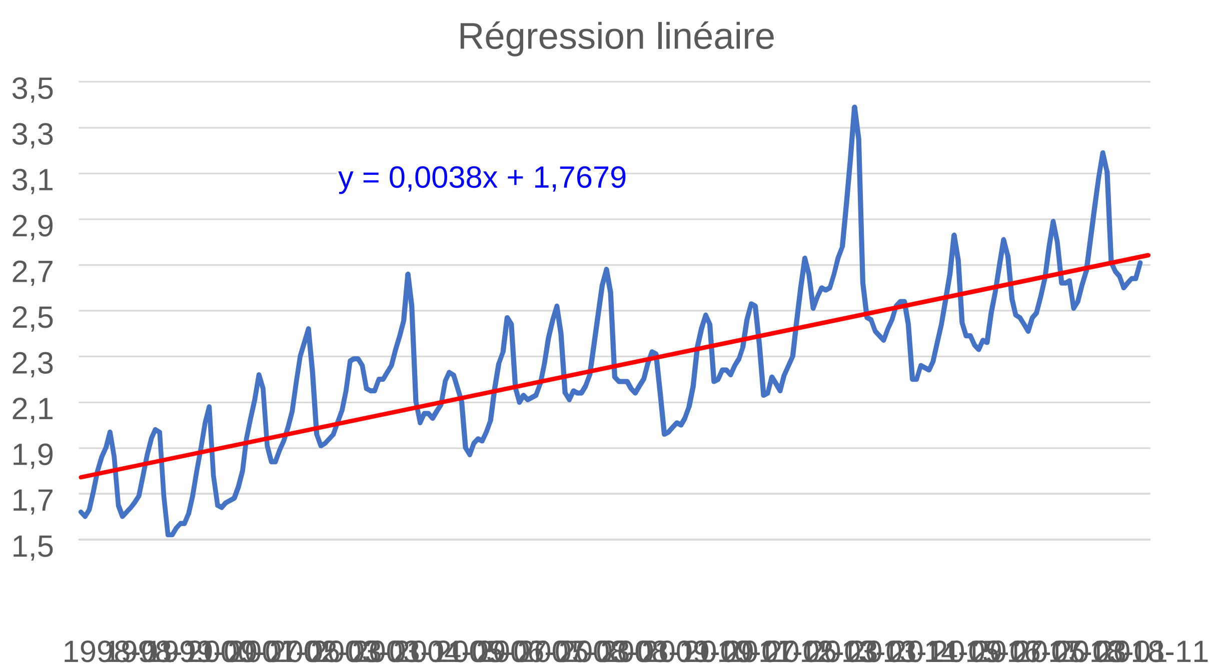
2- a) Moindres carrés ordinaires (OLS)

- https://www.insee.fr/fr/statistiques/serie/000641367#Telechargement




Enregistrement automatique
valeurs_mensuelles - Excel
liza charroin
Fichier Accueil Insertion Mise en page Formules Données Révision Affichage Aide Création de graphique Format
Rechercher
Partager
Commentaires
Ajouter un élément graphique
Disposition rapide
Modifier les couleurs
Styles du graphique
Intervertir ligne/colonne
Sélectionner des données
Données
Modifier le type de graphique
Type
Déplacer le graphique
Emplacement
Axes
Titres des axes
Titre du graphique
Étiquettes de données
Table de données
Barres d'erreur
Quadrillage
Légende
Lignes
Courbe de tendance
Barres Haut/Bas
Aucune
Linéaire
Exponentielle
Prévision linéaire
Moyenne mobile
Autres options de la courbe de tendance...
Prix moyen mensuel kilo de pommes
3,5
3,3
2018-02 2,54
2018-03 2,61
2018-04 2,67
2018-05 2,81
2018-06 2,95
2018-07 3,08
2018-08 3,19
2018-09 3,11
2018-10 2,71
2018-11 2,67
2018-12 2,65
2019-01 2,6
2019-02 2,62
2019-03 2,64
2019-04 2,64
2019-05 2,71
Format de courbe d..
Options de courbe de tendance
Courbe
Aucun trait
Trait plein
Trait dégradé
Automatique
Couleur
Transparence
0 %
Largeur
1,5 pt
Composition
Type de tiret
Type d'extrémité
Arrondie
Type de joint
Arrondi
Type de flèche initiale
Taille de la flèche initiale
Type de flèche finale
Taille de la flèche finale

2- a) Moindres carrés ordinaires (OLS)



Interpretation?

2- a) Moindres carrés ordinaires (OLS)

La tendance linéaire ··est donnée par équation de la régression linéaire

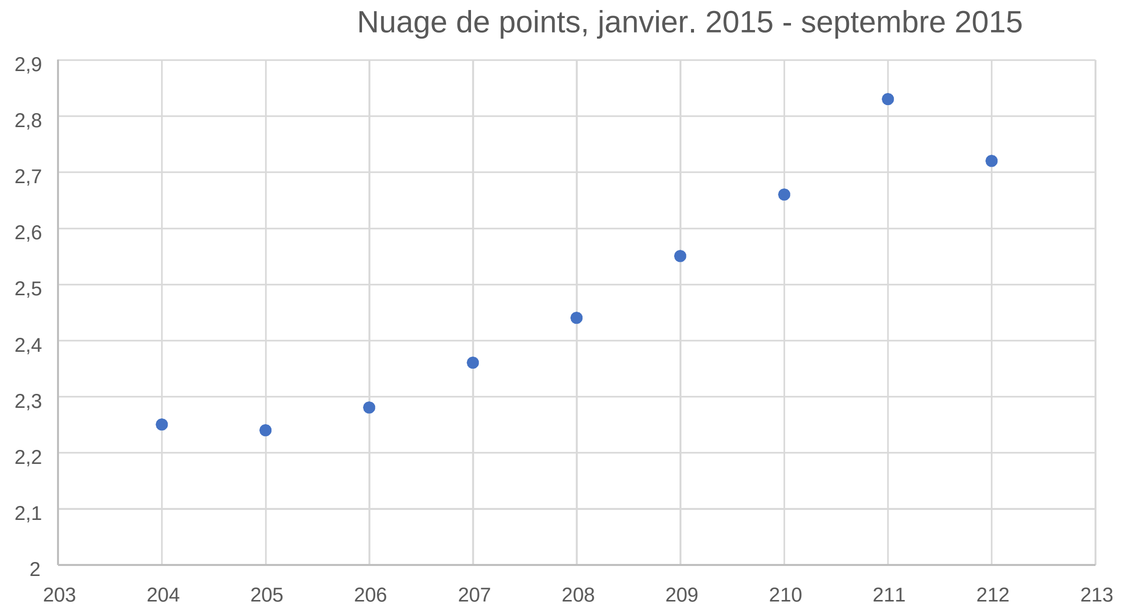
SOMME =0.0038*E4+1.7679

	A	B	C	D	E	F	G
1						Average monthly prices in France -	
2							
3	Date	Year	Month	Month#	t	Price	Linear trend (g_t)
4	1998-01	1998	january	1	0	1.62	=0.0038*E4+1.7679
5	1998-02	1998	february	2	1	1.6	1.7717
6	1998-03	1998	march	3	2	1.63	1.7755
7	1998-04	1998	april	4	3	1.71	1.7793
8	1998-05	1998	may	5	4	1.8	1.7831

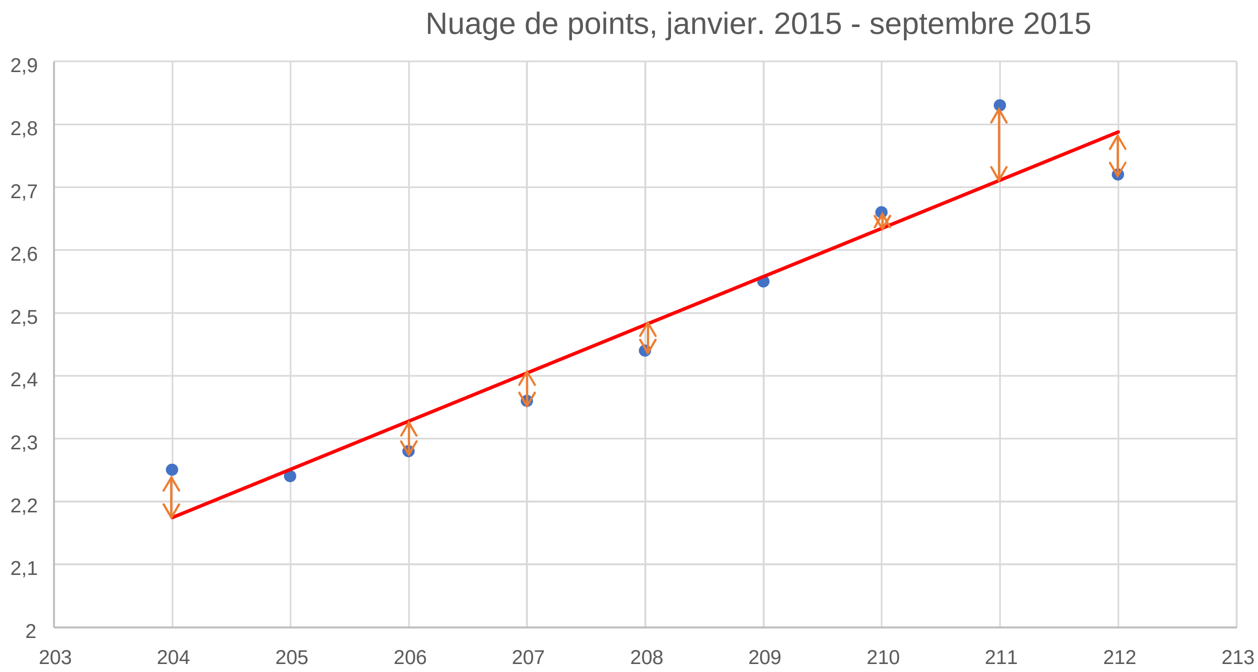
2- a) Moindres carrés ordinaires (OLS)

- Comment détermine-t-on la fonction OLS ?
  - Here, we focus on linear regressions: ·= ··+ ·(here ·represents time, so nous l'écrivons habituellement ·= ··+ ·)
  - ·représente la pente de la courbe (si >0, relation positive, si <0 négatif)
  - ·représente l'ordonnée à l'origine
- L'objectif est de tracer la ligne qui minimise les écarts entre chaque point et cette ligne.

2- a) Moindres carrés ordinaires (OLS)



2- a) Moindres carrés ordinaires (OLS)



Nous voulons dessine
tendance (ligne) tandi
minimizing the
deviations between
nos données et les
doubler. Nous voulons
minimiser la taille de
les flèches.

2- a) Moindres carrés ordinaires (OLS)

- · est la pente (coefficient) et · est l'origine.
- On choisit · et · tels que :

$$\min_{\cdot,\cdot\cdot} \{\cdot\cdot\cdot\cdot \times \cdot + \cdot\}^2$$

Nous voulons minimiser la distance entre chaque point et la tendance line ··= ·× ·+ ·.

Nous irons plus loin dans le chapitre 5.

2- b) Moyennes mobiles

- Autre méthode : moyennes mobiles. Objectif : obtenir la tendance en lissage. Méthode non paramétrique.
- Aucune hypothèse sur le modèle sous-jacent (linéaire, exponentiel, etc
- Principe du calcul : on définit une série de moyennes qui sont
  - Calculé sur un certain nombre d'observations consécutives, appelé le order.
  - Décalé d'une unité de temps à chaque nouveau calcul (« déplacement »)

Moyennes mobiles : ordre 3

Formule générale :

Moving averages

· Moyennes mobiles d'ordre 5

·· estimation de la tendance à ·= 3

·· estimation de la tendance à ·= 4

etc.

Formule générale : ··

Les moyennes mobiles avec un ordre impair sont faciles à calculer.

Moyennes mobiles d'ordre pair (par exemple,

- Problème : quelle date (t) associer à une moyenne de 4 valeurs ?
- Solutions : on fait la moyenne de deux moyennes consécutives d'ordre

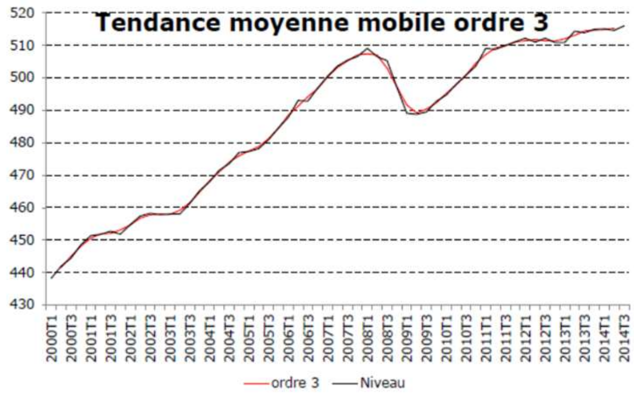
$$\frac{1}{2}\left[\frac{\cdot\cdot+\cdot\cdot+\cdot\cdot+\cdot s}{4}+\frac{\cdot\cdot+\cdot\cdot+\cdot\cdot+\cdot\cdot}{4}\right]$$

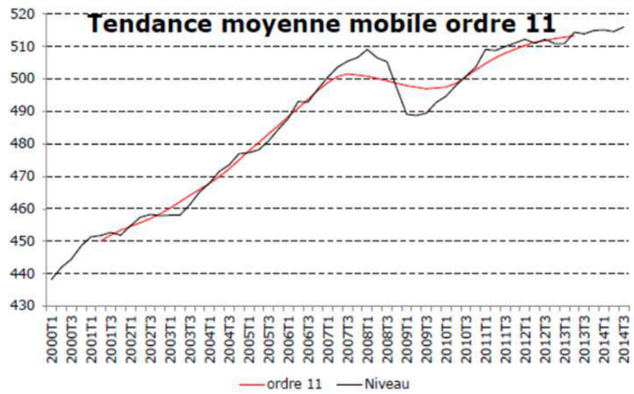
$$=\frac{\frac{1}{2}\qquad\qquad 2\,\frac{\cdot\cdot}{}}{4}$$

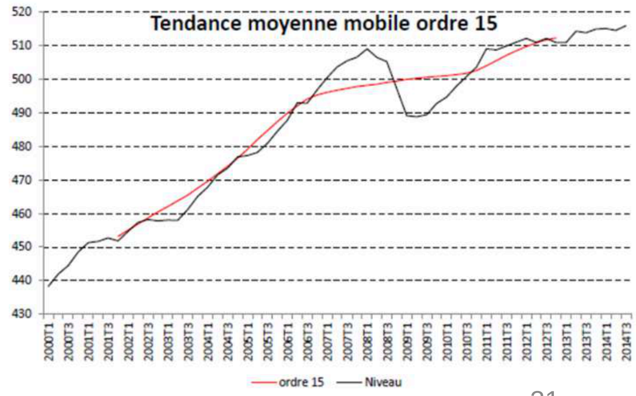
Formule générale : $\cdot\cdot\cdot\cdot\cdot = \dfrac{\cdots}{\cdot}$

Moving averages

- Quel ordre choisir ?
- La commande doit être adaptée à la périodicité des variations saisonn
- Cependant, plus la commande est importante,
  - Plus le lissage est fort
  - plus les valeurs sont « perdues »







Example

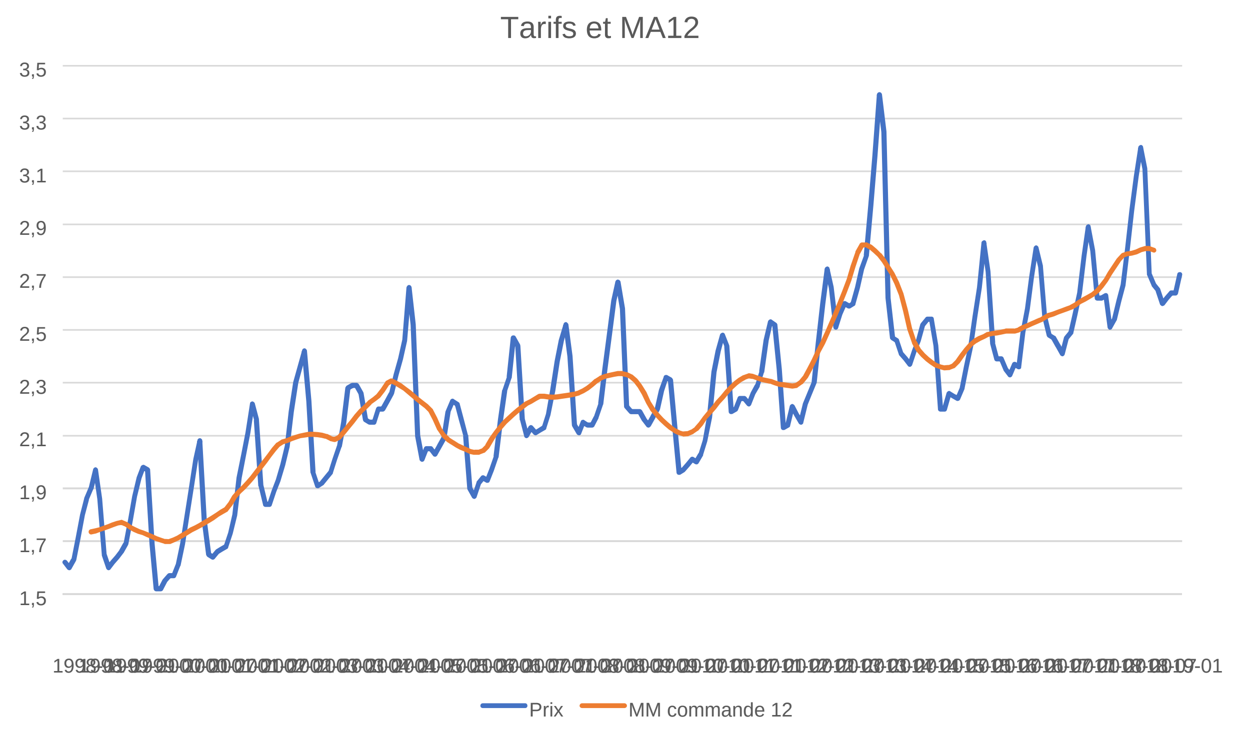
2- b) Moyennes mobiles

- Moyennes mobiles d'ordre 12

	Average monthly prices in France - Apples (1kg)			
t	Price	Linear trend (g_t)	Moving averages (order 12)	Deviations g_t)
0	1.62	1.7679		
1	1.6	1.7717		
2	1.63	1.7755		
3	1.71	1.7793		
4	1.8	1.7831		
5	1.86	1.7869		
6	1.9	1.7907	=(1/2*F4+SOMME(F5:F15)+1/2*F16)/12	
7	1.97	1.7945	SOMME(**nombre1**; [nombre2]; ...)	
8	1.86	1.7983	1.744166667	
9	1.65	1.8021	1.749583333	
10	1.6	1.8059	1.755416667	
11	1.62	1.8097	1.761666667	
12	1.64	1.8135	1.768333333	

Formule Excel : SOMME en français et SUM en Anglais

2- b) Moyennes mobiles

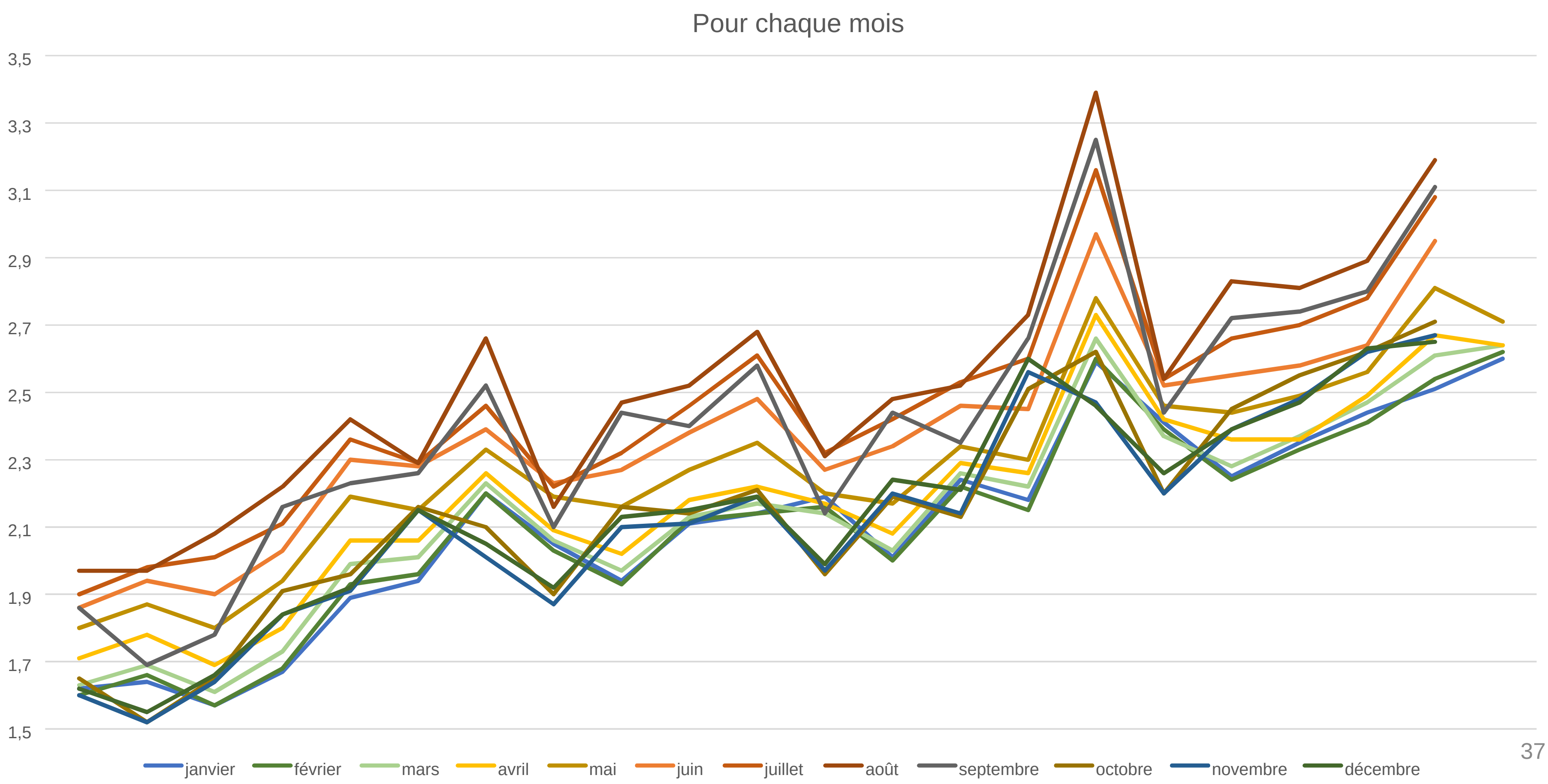


Comparaison entre OLS et déménagement averages

- OLS:
  - Simple à calculer sur Excel
  - Peut être calculé pour toute la période
  - Facile à interpréter
  - Permet d'extrapoler la tendance
  - Mais : pas flexible ET quelle tendance choisir (linéaire, exponentielle, etc.) ?
- Moyennes mobiles :
  - Flexible (peut s'adapter à de grandes variations dans le temps)
  - Mais : difficile de choisir l'ordre, ne peut pas être calculé sur toute la période et ne permet pas d'extrapoler la tendance (donc pas de prévision possible)

3. Identification des saisonniers component

Saisons


Pour chaque mois
3,5
3,3
3,1
2,9
2,7
2,5
2,3
2,1
1,9
1,7
1,5
janvier
février
mars
avril
mai
juin
juillet
août
septembre
octobre
novembre
décembre

3. Identification de la composante saisonnière

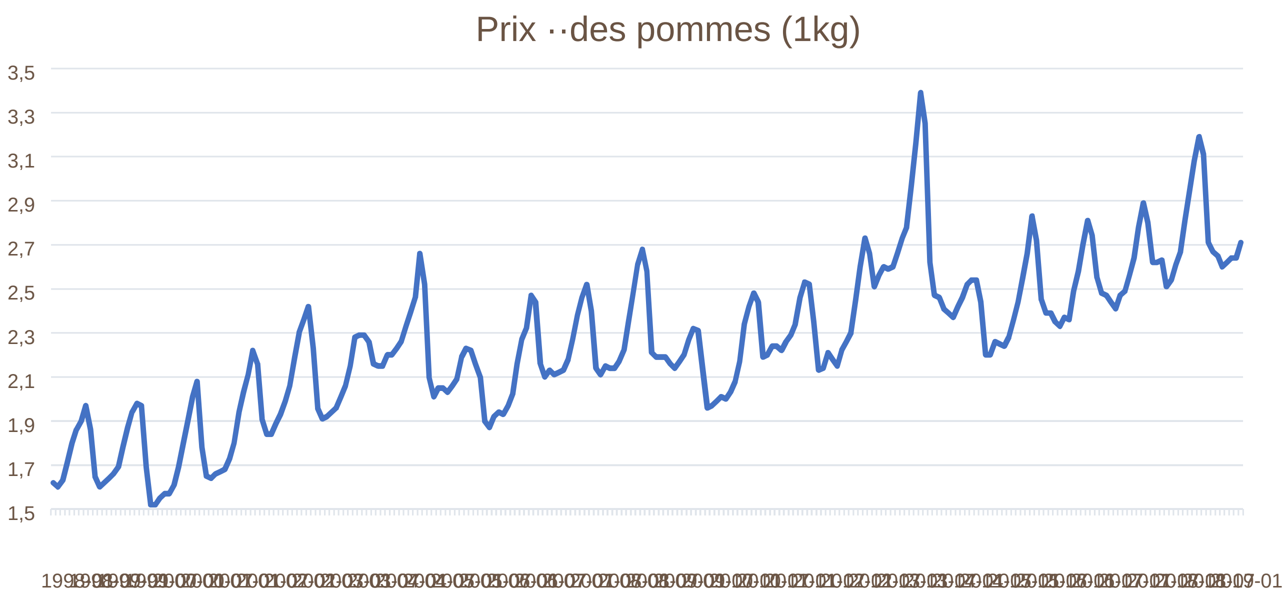
- Pour chaque date ·, nous avons
  - Une observation ··
  - Une estimation de la tendance linéaire ··
- La composante saisonnière est déterminée à partir de ces deux éléme
  Nous calculons les écarts par rapport à la tendance. Pour le :
  - Additive model: ··= ·····
  - Multiplicative model: ··= $\frac{\cdot\cdot}{\cdot\cdot}$

3. Identification de la composante saisonnière

- Ensuite, nous calculons la moyenne de · ·pour chaque saison de l'anne mois, chaque trimestre· selon les données)
- Cela correspond au coefficient saisonnier de chaque mois (ou quarter·):
  - 12 coefficients saisonniers (SC) avec données mensuelles
  - 4 SC avec données trimestrielles

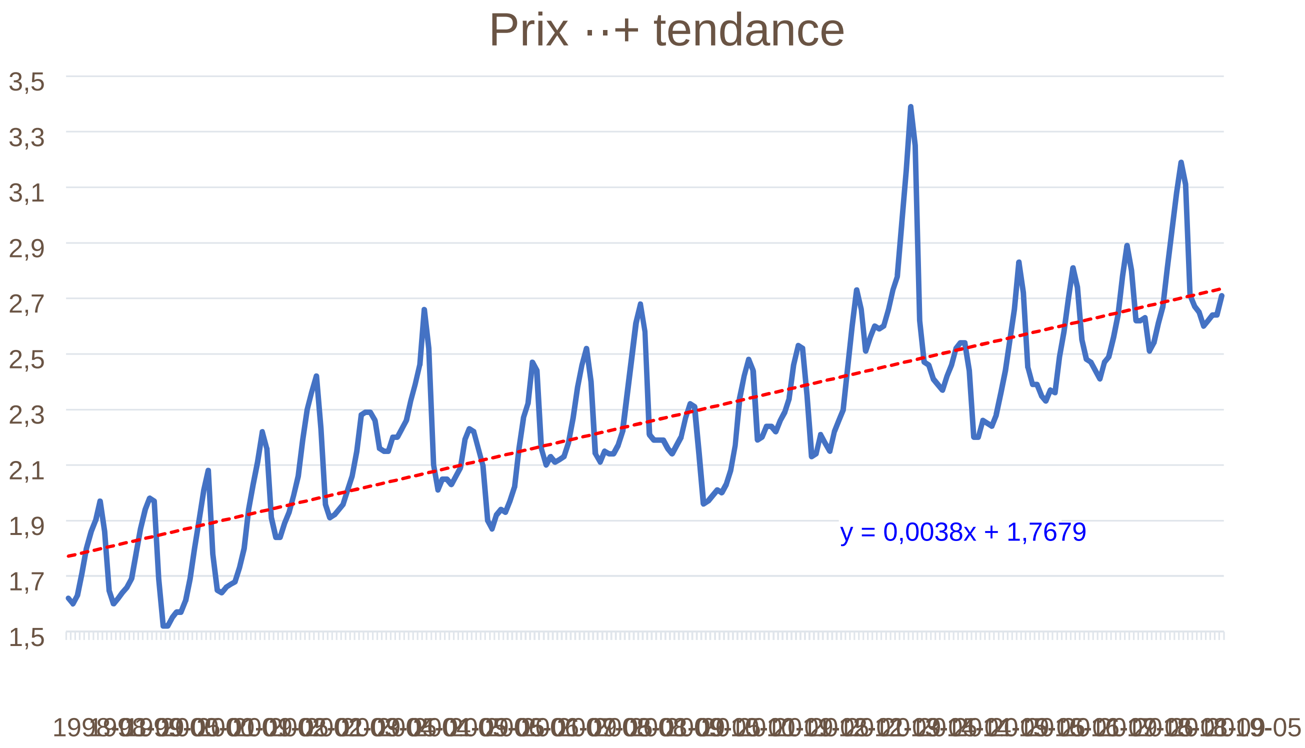
Example

- Quel modèle choisissons-nous ?



Raw data

Example

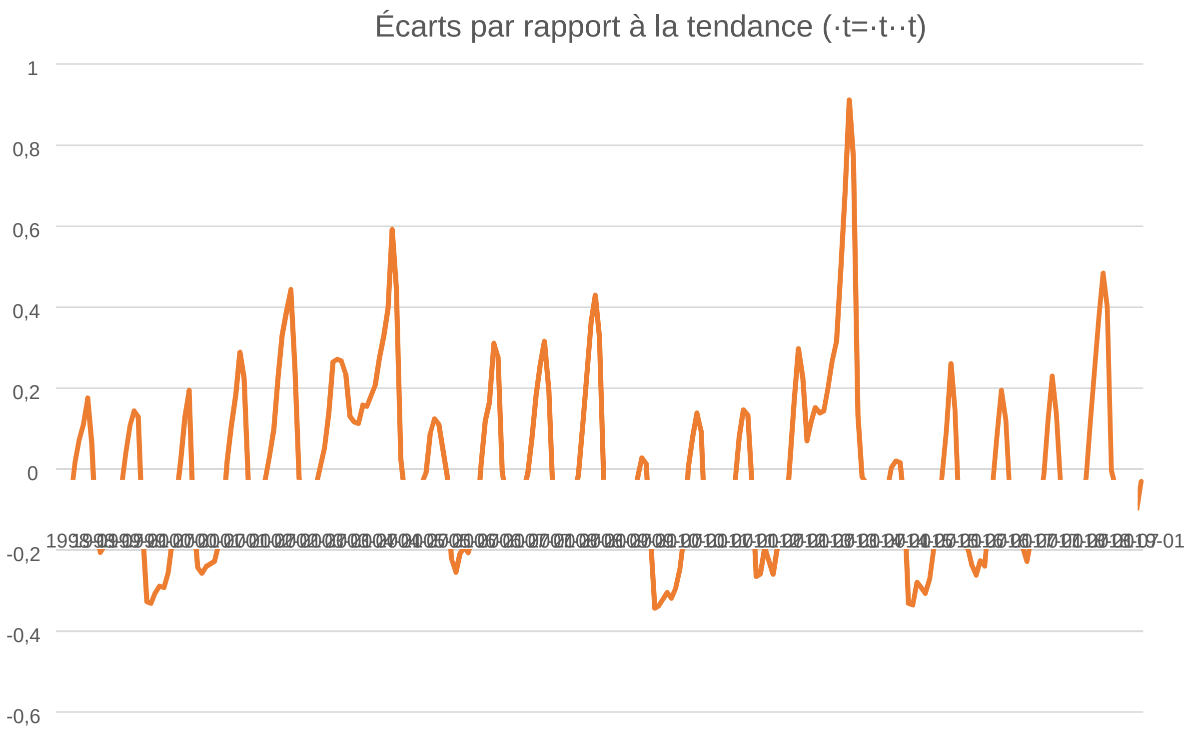


Tendance linéaire

Calcul des écarts à la tendance

t	Price	Tendance linéaire (g·)	Moyennes mobiles (ordre 12)	Écarts par rapport à la tendance (s· = x· · g·)
0	1.62	1,7679		-0,1479
1	1.6	1,7717		-0,1717
2	1.63	1,7755		-0,1455
3	1.71	1,7793		-0,0693
4	1.8	1,7831		0,0169
5	1.86	1,7869		0,0731
6	1.9	1.7907	1.735833333	0,1093
7	1.97	1,7945	1.739166667	0,1755
8	1.86	1,7983	1.744166667	0,0617
9	1.65	1.8021	1.749583333	-0,1521
10	1.6	1.8059	1.755416667	-0,2059

Écarts par rapport à la tendance



3. Identification de la composante saisonnière

- Données mensuelles de janvier. 1998 à mai 2019
  - Nous avons 22 observations chaque mois pour janvier, février· jusqu'en mai
  - 21 observations chaque mois de juin, juillet, · jusqu'en décembre
- SC en janvier : moyenne de 22 écarts par rapport à la tendance obser Janvier (idem en février, etc. jusqu'en mai).
- SC en juin : moyenne de 21 écarts à la tendance (idem en juillet, · december).

3- a) Coefficients saisonniers

3- b) Correction éventuelle de la saisonnalité coefficients

- Hypothèse de neutralité des variations saisonnières sur une période
- Moyenne des coefficients saisonniers (SC) :
  - Zéro dans le modèle additif
  - Égal à 1 dans le modèle multiplicatif
- Si cette hypothèse n'est pas satisfaite, il est nécessaire d'ajuster le coefficients saisonniers :
  - Additive model · subtract the mean of the SCs from each seasonal coefficient
  - Multiplicative model · divide each seasonal coe·cient by the mean of the SCs

3- b) Correction éventuelle de la saisonnalité coefficients

- Si nous disposons de données trimestrielles :
  - Moyens de SC : ⋯= $\frac{\cdots\cdots\cdots\cdots\cdots\cdots}{\cdot}$
  - Modèle additif :
    - ⋯est significativement différent de 0 ?
    - Si oui, SC ajusté : ⋯·= ·······, ····   ·= ········, etc.
  - Modèle multiplicatif :
    - ⋯est significativement différent de 1 ?
    - Si oui, SC ajusté : ⋯·= ···/···, ····   ·= ····/···, etc.

Example

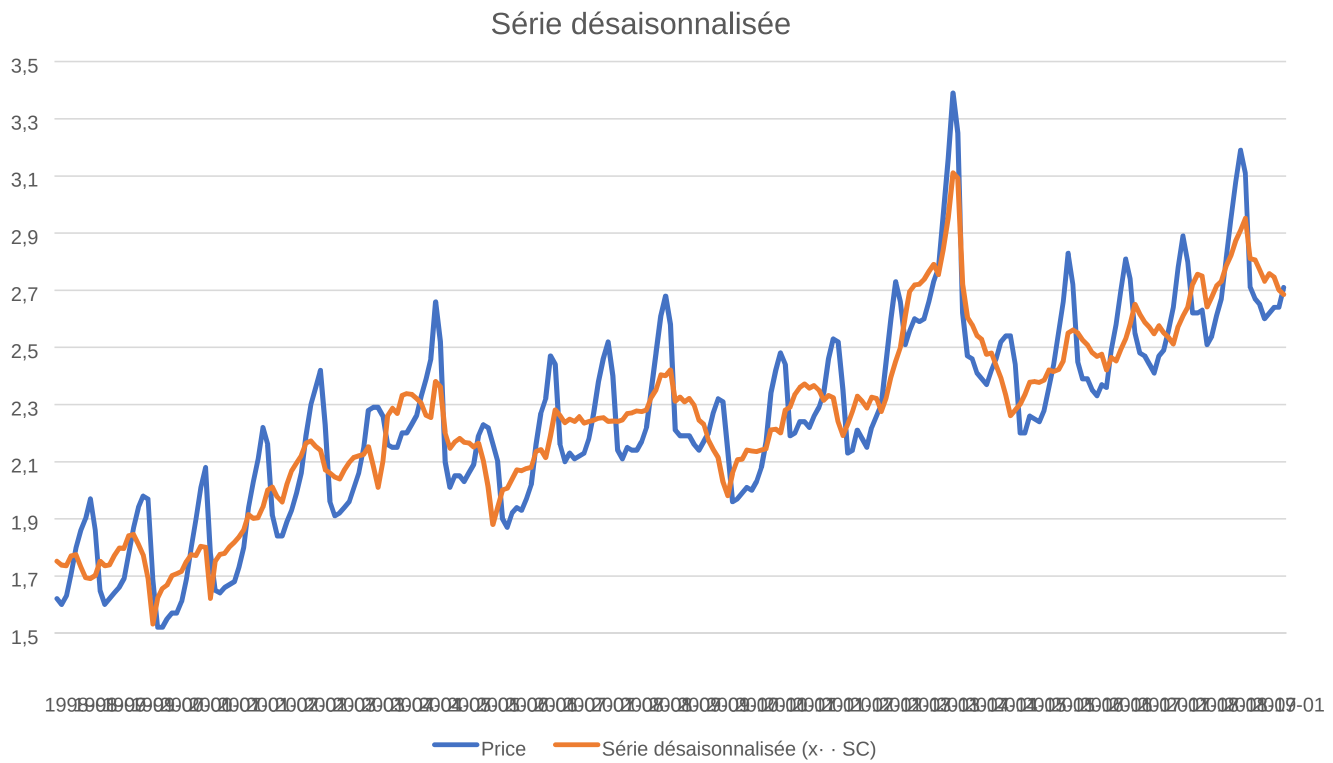
Coefficients saisonniers
-0,131245455
-0,137772727
-0,105663636
-0,060827273
0,024918182
0,128052381
0,206157143
0,2795
0,158557143
-0,100957143
-0,136185714
-0,119033333
0,000458297

Moyenne assez proche de 0 ici

4- Série désaisonnalisée (SA series)

Seasonal adjustment

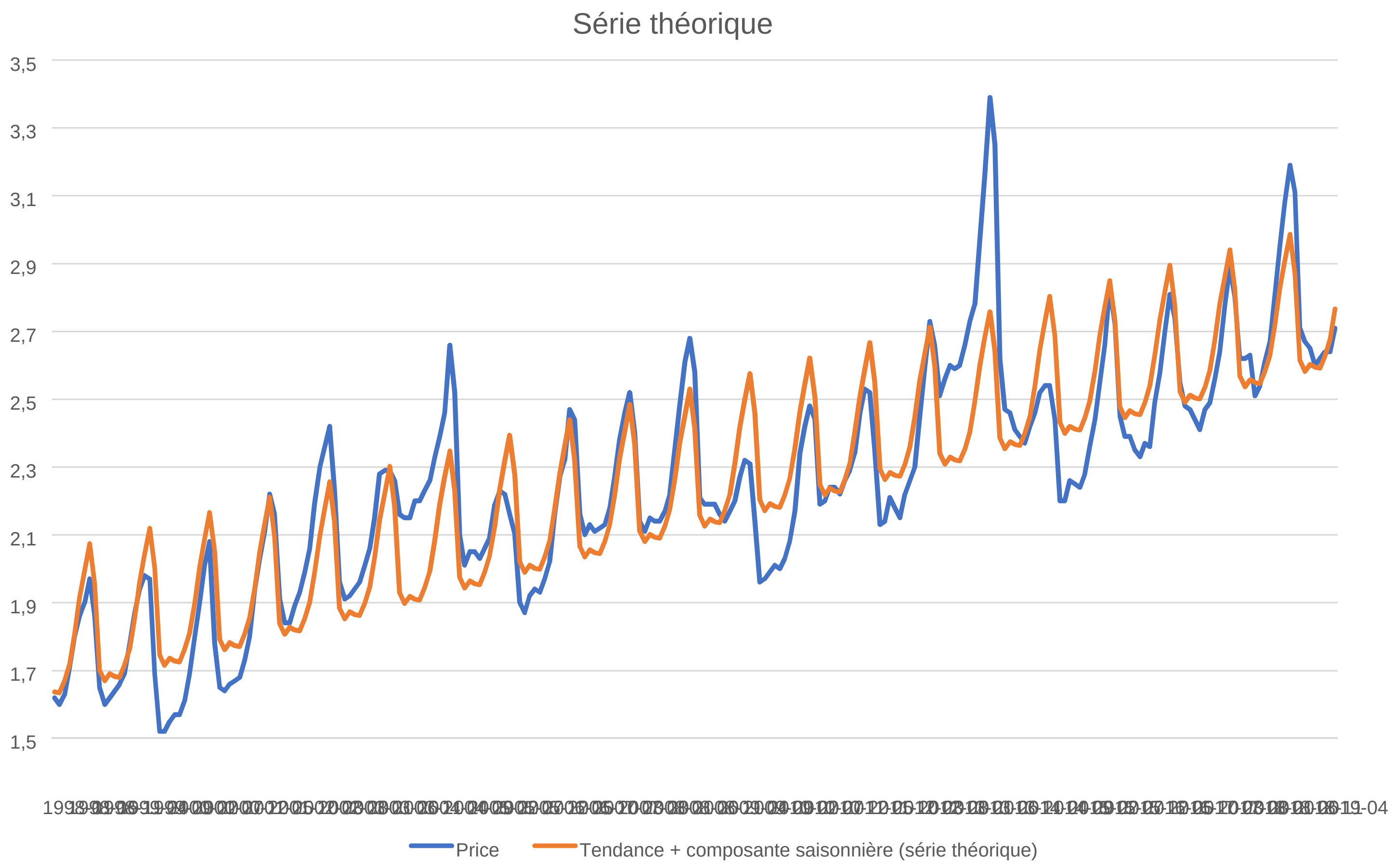
- Pourquoi procéder à une désaisonnalisation ? Élimine l·effet des fluctu pour mieux observer les tendances de fond.
- Comment procéder à la désaisonnalisation ?
  - Modèle additif : soustraire de la valeur de la variable chaque mois (ou trimestre) le coefficient saisonnier associé à ce mois (ou trimestre).
    - La série désaisonnalisée (série SA) comprend ·· · ···· où · représente le mois et t la date.
  - Modèle multiplicatif : divisez la valeur de la variable dans chaque mois (ou trimestre) par le coefficient saisonnier associé à ce mois (ou trimestre).
    - La série SA comprend ··/····


Série désaisonnalisée
3,5
3,3
3,1
2,9
2,7
2,5
2,3
2,1
1,9
1,7
1,5
Price
Série désaisonnalisée (x· · SC)

5- Prévisions

Prévision

- Nous reconstituons une série théorique à partir de la tendance et de la coefficients
- Avec:
  - Le modèle additif : ·· ·= ··+ ····
  - Le modèle multiplicatif : ·· ·= ··× ····
- Pour faire des prédictions, nous étendons les séries théoriques
- Il n'est possible d'étendre la tendance ··qu'avec la méthode OLS



Série théorique
3,5
3,3
3,1
2,9
2,7
2,5
2,3
2,1
1,9
1,7
1,5
Price
Tendance + composante saisonnière (série théorique)

Prévision

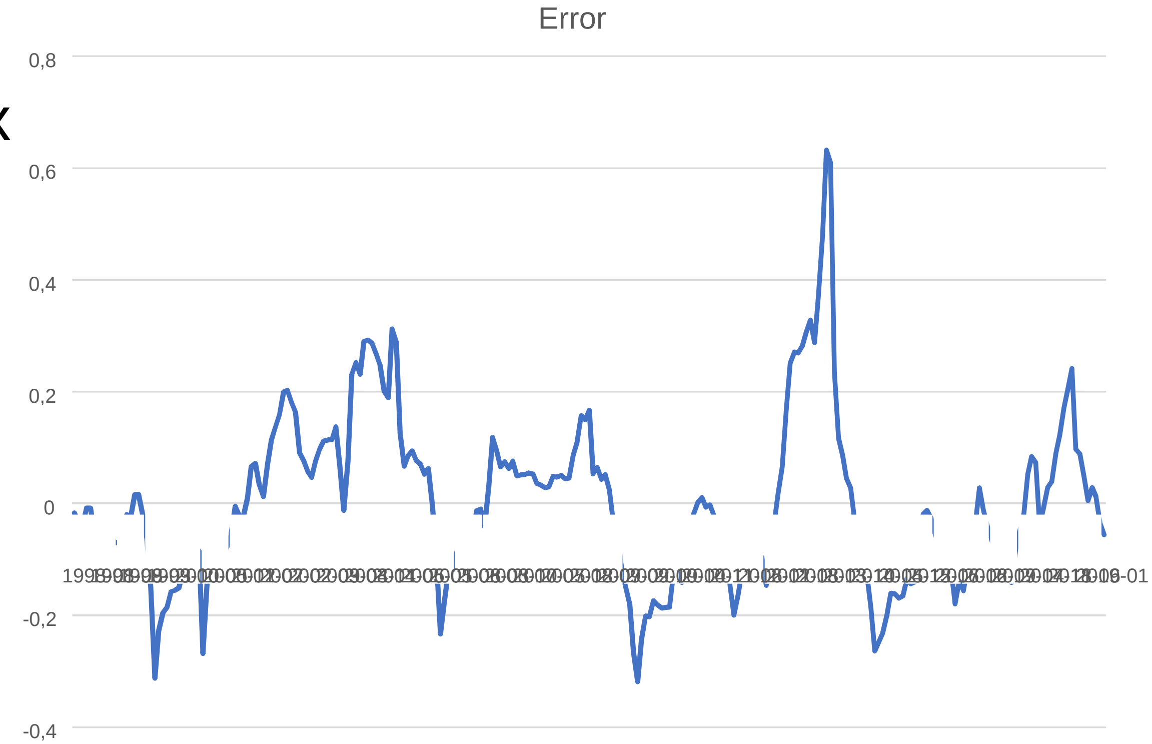
- Quel devrait être le prix d'1 kg de pommes en juin 2019 ?
  - Tendance (OLS) : g·= 0,0038·+ 1,7679
  - Juin 2019 correspond à · = 257
  - g···= 0.0038 × 257 + 1.7679 = 2.7445 ·linear trend
  - On ajoute le coefficient saisonnier correspondant : ······= 0,1280
- Prévisions de prix en juin 2019 : = ·2.7445 + 0.1280 = 2.8725.
- Prix observé en juin 2019 ? 2.79. Nous sommes assez proches. Comr expliquer l'écart entre les deux valeurs ?

Pour améliorer nos prévisions

- Pour être plus précis, on peut prendre en compte la composante cyclic
  (par exemple, tous les 5 ans la récolte des pommes peut être meilleure
- Choisissez une autre forme fonctionnelle pour estimer la tendance (no
  exponential, etc.).
- Focus sur les déterminants du prix des pommes (climat, etc.).

Erreur et/ou cycle

- Erreur (et/ou cycle) = prix observé - théorique series



Résumé pour analyser une série temporelle

1. On choisit le modèle (additif ou multiplicatif) grâce à la variations (constantes ou variables)
2. Nous déterminons la tendance (OLS ou Moyennes mobiles) pour voi évolution des séries temporelles sur le long terme
3. On calcule les coefficients saisonniers (grâce aux écarts à la tendance)
4. Nous pouvons calculer la série désaisonnalisée en supprimant le SC à partir des observations.
5. Nous déterminons la série théorique (tendance + SC) et pouvons pré valeurs futures (uniquement avec tendance OLS).

Résumé

Série théorique

=

+ ou x

+ ou x