



# Techniques Quantitatives et Représentations

## 1ère année Semestre 2

Tech de Co Bordeaux – 2021-2022

[Bernard.Andrucoli@u-bordeaux.fr](mailto:Bernard.Andrucoli@u-bordeaux.fr)

<http://aristeri.com/MSAG/>



Tech  
de Co

iut  
de  
BORDEAUX

RéseauTC  
France

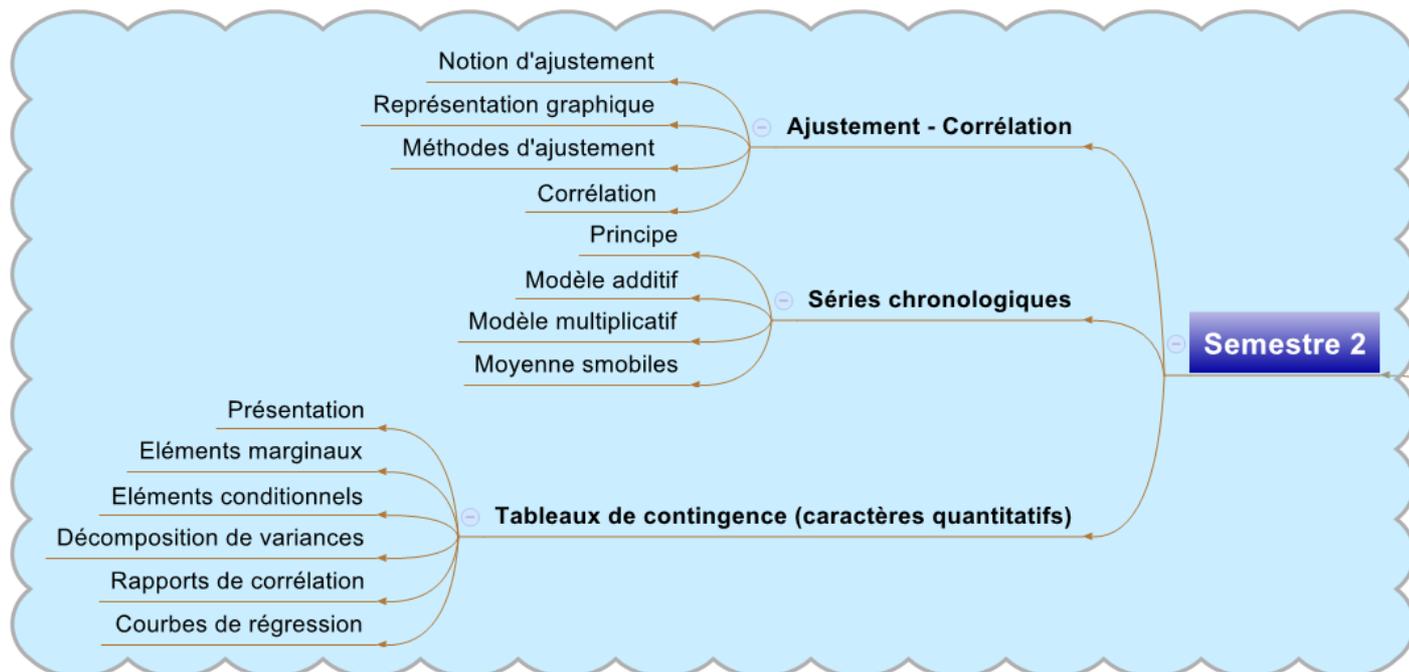
1

## o<sup>x</sup> Objectif du cours

- Mieux percevoir la difficulté que présente l'analyse des données quantitatives. (Dispersion, concentration, liens entre les caractères)
- Exploiter les données numériques avec plus de rigueur.
- Liens avec les tableurs et R
- MSAG dans une UE parfois délicate pour certains étudiants...mieux vaut éviter les handicaps.

2

# Programme indicatif



## Bibliographie indicative

The following is a list of indicative bibliographies:

- Bernard Py - Statistiques descriptives -( Economica)
- Bernard Grais - Statistiques descriptives - (Dunod)
- Bernard PY - La statistique sans formule mathématique - (Pearson Education)
- Didier Schlatchter, Comprendre les mathématiques financières, 3ème édition, (Hachette Supérieur)
- Rachid Boutti - Statistiques descriptives - Diagnostic et prise de décision (Université Cadi Ayyad)
- Henry Aubert - Manuel de statistique (Ellipses)
- Jérôme Hubler - Statistique descriptive appliquée à la gestion et à l'économie (Bréal)
- Vincent GIARD - Statistiques appliquées à la gestion (Economica)
- Gaston MIALARET - Statistiques appliquées aux sciences sociales (PUF)



# Evaluation

Test écrit 2 heures mois de mai + QCM en TD.

Grande Majorité des étudiants:

**Note TQR S2 = Note test x 70% + Note QCM x 30%**

**Possibilité** de relever certains travaux qui pourront être notés

**Malus** si participation insuffisante et/ou si comportement (Bavardages, exclusions) pas conforme à celui d'un(e) étudiant(e). De même pour les **absences non (ou mal) justifiées** et les **retards** trop fréquents. (**Malus possible!**)

*Les sonneries intempestives de téléphones portables pendant les séances entraîneront l'exclusion immédiate du cours avec les incidences évoquées ci-dessus. Calculatrice en TD sous peine d'exclusion..*



## Organisation des séances.

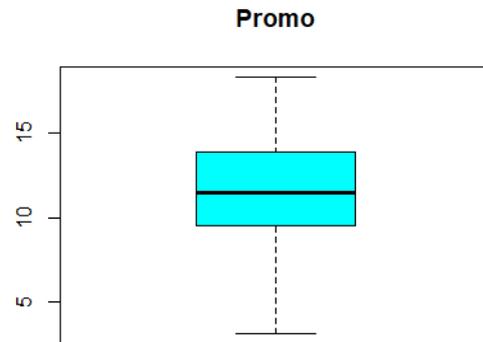
- 1 Cours (Amphi: B A)
  - C'est bien peu...
- 7 TD (BA + SB)
  - C'est très peu...
- Fin des séances bien avant les tests...

TD: exercices à préparer. Le travail personnel doit être réel, la participation active.

**Ne pas traîner.**

# Bilan des courses à mi-parcours

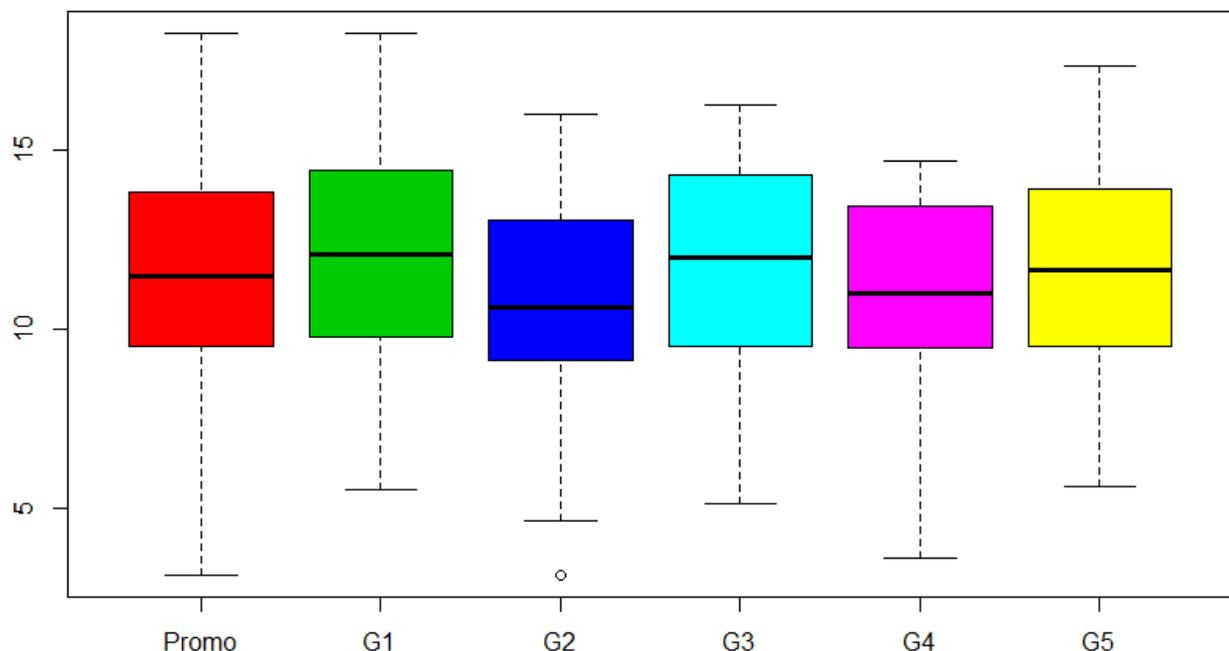
```
> summary(promo)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  3.12   9.50   11.48   11.50  13.84   18.27
> summary(g1)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  5.52  10.05   12.10   11.99  14.40   18.27
> summary(g2)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  3.12   9.14   10.60   10.91  13.06   16.01
> summary(g3)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  5.11   9.50   12.00   11.77  14.32   16.28
> summary(g4)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  3.600  9.475   11.010  10.970  13.420   14.680
> summary(g5)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  5.61   9.50   11.67   11.71  13.84   17.35
> |
```



*Résultats et graphiques issus du logiciel R.*

# Bilan des courses à mi-parcours

Répartition des notes S1 2019-2020.



## Logique des enseignements

- Un seul amphi
- Cours en ligne: <http://aristeri.com/MSAG/> (avec **étudiants** et **secouez\_vous**)
- Support **pdf** utilisé actuellement
- Une **vingtaine de vidéos** y sont présentes pour ce semestre 2
- Il y en avait autant au semestre 1...
- **A REGARDER ACTIVEMENT EN ENTIER en prenant des notes AVANT** les TD
- Fichiers de données présents DONT des corrections des TD et des minitests,
- Forum pour échanger en ligne et poser des questions.
- Aucune n'a été posée au S1...
- Un support de cours de 54 pages et 108 diapos.(66 pages et 131 diapos au S1)
- Etre actif en TD, poser des questions sur le cours en début de séance.
- Cas à traiter: illustration d'un chapitre ou d'une partie.
- Lien avec Tableurs.
- **Acquérir des connaissances pour mobiliser des compétences!**



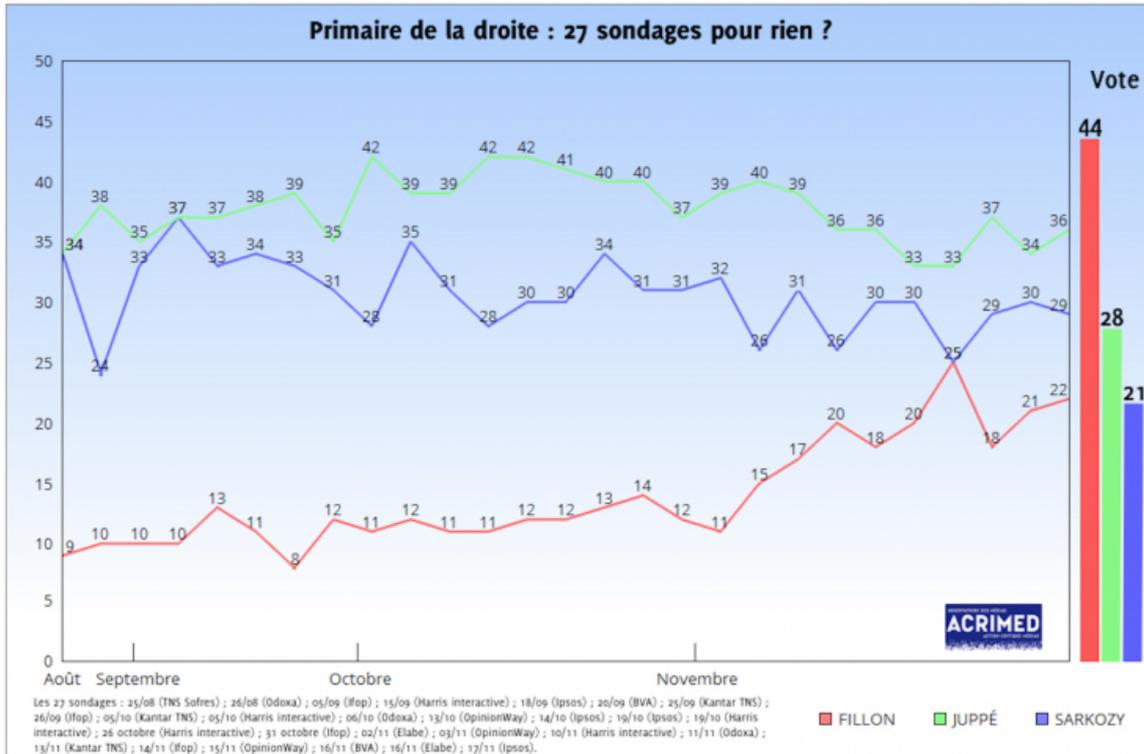
## Sondages et manipulation des chiffres: exercice périlleux!

Institut	Date	Échantillon	Manuel Valls	Benoît Hamon	Arnaud Montebourg	Vincent Peillon	Sylvia Pinel	François de Rugy	Jean-Luc Bennaïmas	Aucun / ne sait pas
Harris Interactive <a href="#">[archive]</a>	2 au 4 janvier	[6245] / 478	43 %	22 %	25 %	7 %	2 %	1 %	<0,5 %	-
Kantar Sofres - OnePoint <a href="#">[archive]</a>	3 au 6 janvier	[8011] / 488	36 %	21 %	23 %	10 %	6 %	2 %	2 %	-
OpinionWay <a href="#">[archive]</a>	9 au 11 janvier	[7758] / 453	40 %	29 %	21 %	7 %	1 %	1 %	1 %	-
Premier débat télévisé (12 janvier 2017)										
BVA <a href="#">[archive]</a>	13 au 16 janvier	[9434] / 543	34 %	27 %	26 %	7 %	3 %	2 %	1 %	-
Deuxième débat télévisé (15 janvier 2017)										
OpinionWay <a href="#">[archive]</a>	16 au 18 janvier 2017	[7002] / 536	37 %	28 %	24 %	5 %	3 %	1 %	2 %	-
Troisième débat télévisé (19 janvier 2017)										
Premier tour (22 janvier 2017)										
Résultats <a href="#">[archive]</a>	22 janvier 2017	1 630 139	31,11 %	36,35 %	17,52 %	6,85 %	1,98 %	3,88 %	1,01 %	1,3 % (blancs et nuls)

Fiche OpinionWay

Fiche BVA

# Sondages et manipulation des chiffres: exercice périlleux!



Fiche Elabe

Fiche Ifop

## Sondage et élection

		Jean-Luc Mélenchon	Benoît Hamon	Emmanuel Macron	François Fillon	Nicolas Dupont-Aignan	François Asselineau	Marine Le Pen
Résultats [archive]	Résultat	19,58%	6,36%	24,01%	20,01%	4,70%	0,92%	21,30%
Ifop-Fiducial [Sondage]	18-21 mars	11,50%	11,50%	26,00%	17,50%	5%	0,50%	26%

# Plan

## Chapitre 1: ajustement & corrélation

### I - Présentation

II - Méthodes basiques d'ajustement

III - Méthode des moindres carrés

IV - Corrélation

V - Autre exemple

VI - Ajustement, Excel et R

## Chapitre 2: Chroniques

I - Principes

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

## Chapitre 3: Tableaux de Contingence

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

VI- Courbes de régression

VII- Autre cas

13

## I - Présentation

**Séries à deux caractères:** Suppose l'étude **simultanée** de deux caractères.

Effectif ( $x_i$ )	$n_i$
5	95
15	75
<b>N</b>	<b>170</b>

Effectif moyen, modal, médian,...  
sur 170 entreprises...

...sans tenir compte du niveau du  
CA des entreprises étudiées...

CA ( $y_j$ )	$n_j$
1	45
4	55
10	70
<b>N</b>	<b>170</b>

CA moyen, modal, médian,...

...sans tenir compte du niveau de  
l'effectif des entreprises étudiées...

Ici deux caractères étudiés l'un indépendamment de l'autre...

### **Questions sans réponses:**

Nombre d'entreprises ayant un effectif de 5 et un CA de 4?

Effectif moyen des entreprises dont le CA est de 10?

14

# I - Présentation

Abordé dans un second temps

Deux approches des séries à deux caractères:

Effectif/ CA	1	4	10	Total (n <sub>i.</sub> )
5	15	35	45	95
15	30	20	25	75
Total (n <sub>.j</sub> )	45	55	70	170

Tableaux de contingence

Publicité \ CA	1	3	5	8	10	12	15	18
CA	56	72	69	72	82	76	93	80

Ajustement - Au cours de 8 périodes on a relevé les valeurs de deux indicateurs d'une entreprise: la publicité et le CA. Pour chaque période il n'y a qu'un seul montant du CA et un seul montant du budget publicitaire.

### Problème posé:

Représentation graphique, et modélisation afin de disposer d'un outil (mathématique) permettant simulation et prévision.

Abordé dans un premier temps

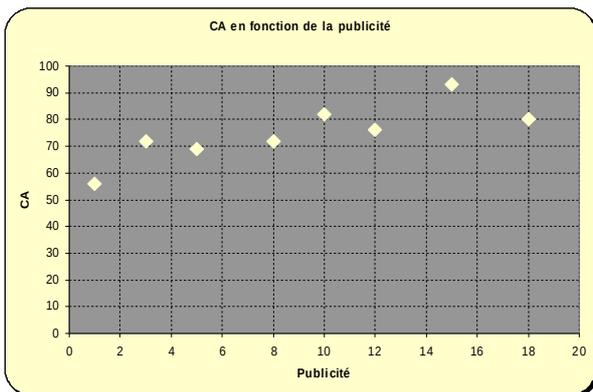
# I - Présentation-Ajustement

Représentation graphique: Attention...au sens de ce qui est représenté

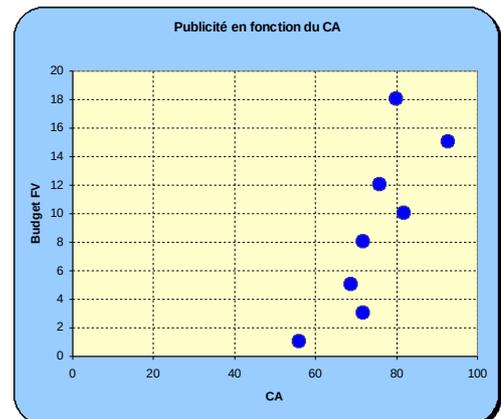
Publicité \ CA	1	3	5	8	10	12	15	18
CA	56	72	69	72	82	76	93	80

Relation dans les deux sens souvent possible. Effectuer une représentation c'est prendre partie en faveur d'une ou l'autre => En avoir conscience.

CA \ Publicité	1	3	5	8	10	12	15	18
CA	56	72	69	72	82	76	93	80



Publicité => CA



CA => Publicité

# I - Présentation

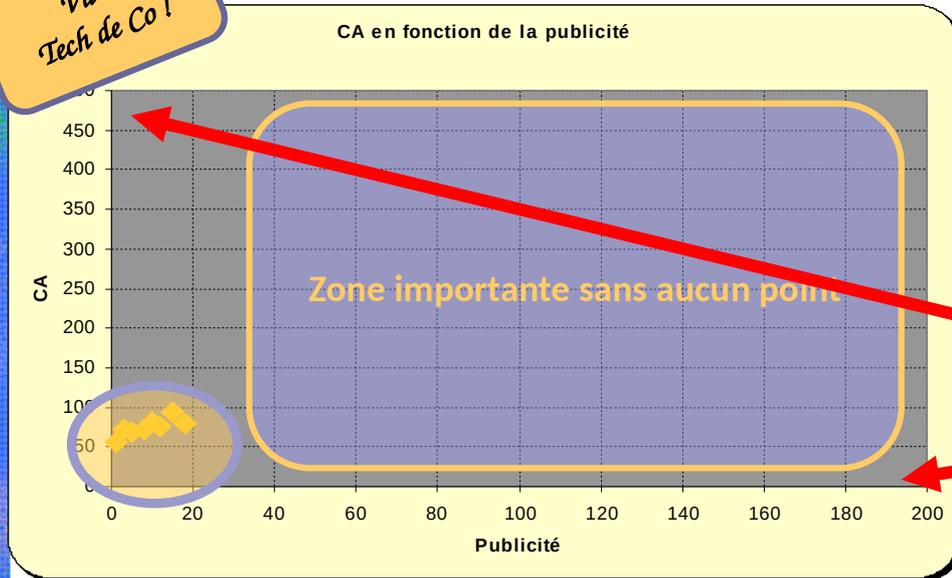
Représentation graphique: Attention...à la représentation

On s'intéresse à la relation Publicité => CA

Vu à Tech de Co !

Publicité		CA
$x_i$	$y_i$	
1		56
3		72
5		69
8		72
10		82
12		76
15		93
18		80

A l'évidence un problème d'échelle...



# I - Présentation

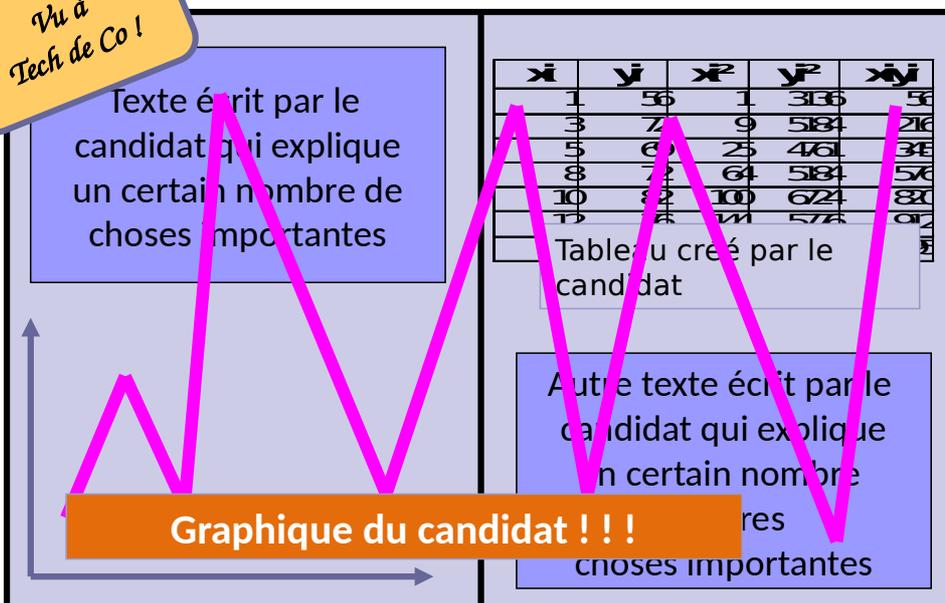
Représentation graphique: Attention...à la représentation

On s'intéresse à la relation Publicité => CA

Vu à Tech de Co !

Publicité		CA
$x_i$	$y_i$	
1		56
3		72
5		69
8		72
10		82
12		76
15		93
18		80

A l'évidence un problème d'échelle...

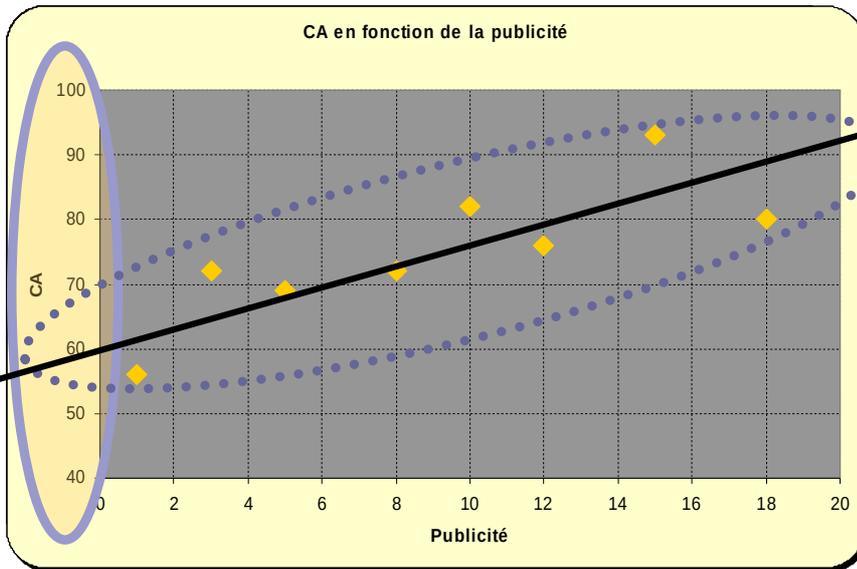


Copie double du candidat

# I - Présentation

On s'intéresse à la relation  $\text{Publicité} \Rightarrow \text{CA}$

Publicité	CA
$x_i$	$y_i$
1	56
3	72
5	69
8	72
10	82
12	76
15	93
18	80



Nuage de points

L'ajustement linéaire consiste à trouver une droite qui approche "le mieux possible" le nuage de points. Cette droite est la représentation d'une équation que l'on va rechercher.

19

## Plan

### Chapitre 1: ajustement & corrélation

#### I - Présentation

#### II - Méthodes basiques d'ajustement

III - Méthode des moindres carrés

IV - Corrélation

V - Autre exemple

VI - Ajustement, Excel et R

### Chapitre 2: Chroniques

I - Principes

II - Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

### Chapitre 3: Tableaux de Contingence

I - Présentation

II - Éléments marginaux

III - Éléments conditionnels

IV - Décomposition de la variance

V - Ajustement & Corrélation

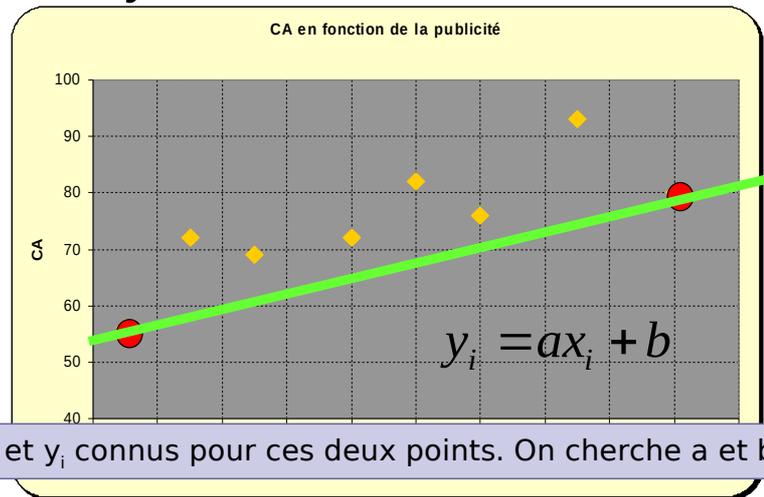
VI - Courbes de régression

VII - Autre cas

20

## II - Méthodes basiques d'ajustement

Publicité		CA
$x_i$	$y_i$	
1	56	
3	72	
5	69	
8	72	
10	82	
12	76	
15	93	
18	80	



$x_i$  et  $y_i$  connus pour ces deux points. On cherche a et b.

Donc:

$$\begin{cases} 56 = a + b \\ 80 = 18a + b \end{cases}$$

$$\begin{cases} a = 1,41 \\ b = 54,59 \end{cases}$$

Equation de cette droite:  $y_i = 1,41 x_i + 54,59$

**Rappel signification de a et b**

Contribution de la publicité dans le CA

**Outil de simulation:**

Publicité=20. CA ? =>  $y_i = 1,41 \times 20 + 54,59$  ; CA= 82,82

CA=100? Publicité? =>  $100 = 1,41x_i + 54,59$

=> Pub= $x_i = 32,17$

**Critique !!!**

21

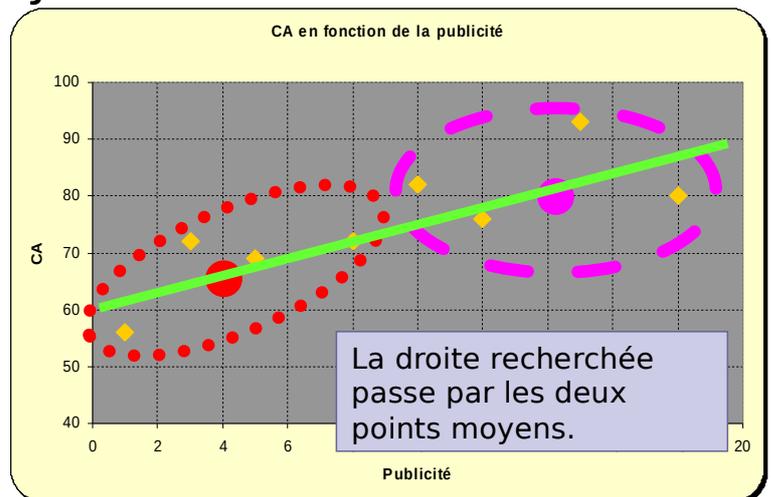
## II - Méthodes basiques d'ajustement

Meyer (Points moyens).

Publicité		CA
$x_i$	$y_i$	
1	56	
3	72	
5	69	
8	72	
10	82	
12	76	
15	93	
18	80	

Groupe de points A

Groupe de points B



**Coordonnées des points moyens:**

**Groupe A:**

$$x_A = (1+3+5+8)/4 = 4,25$$

$$y_A = (56+72+69+72)/4 = 67,25$$

**Groupe B:**

$$x_B = (10+12+15+18)/4 = 13,75$$

$$y_B = (82+76+93+80)/4 = 82,75$$

$$\text{Donc: } \begin{cases} 67,25 = 4,25a + b \\ 82,75 = 13,75a + b \end{cases} \rightarrow \begin{cases} a = 1,63 \\ b = 60,32 \end{cases}$$

Equation de cette droite:  $y_i = 1,63 x_i + 60,32$

**Critique !!!**

22

**=>Conclusion:  
n'utiliser  
aucune des deux  
méthodes  
précédentes!**

## Plan

### **Chapitre 1: ajustement & corrélation**

*I - Présentation*

*II - Méthodes basiques d'ajustement*

**III - Méthode des moindres carrés**

IV - Corrélation

V - Autre exemple

VI - Ajustement, Excel et R

### **Chapitre 2: Chroniques**

I - Principes

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

### **Chapitre 3: Tableaux de Contingence**

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

VI- Courbes de régression

VII- Autre cas

# Méthode des moindres carrés

## Bilan des méthodes précédentes.

- Droite pas déterminée par rapport à tous les points.
- Refaire ces calculs pour étudier la relation inverse.
- On n'est pas certain d'avoir obtenu la meilleure droite d'ajustement possible.
- Aucun indicateur mesurant l'intensité du lien statistique entre ces deux caractères.
- Etude de la pertinence du modèle très délicate.
- ...difficilement utilisables dans ces conditions.

25

# Méthode des moindres carrés

## Intérêt des moindres carrés.

- Meilleur ajustement possible (base de la démonstration).
- Etude de la relation dans les deux sens => 2 droites d'ajustement.
- Intégration de tous les points dans le calcul des coefficients des droites d'ajustement.
- Calcul d'un indicateur relatif à l'intensité du lien statistique possible.
- Etude de la pertinence du modèle facilitée.
- Facilement adaptable sur Tableur.
- **Seule méthode digne de confiance.**
- Deux présentations (« versions ») .

26

# Méthode des moindres carrés

## Version 1: Changement de variable.

Publicité		Version 1				
CA		$X_i$	$Y_i$	$X_i^2$	$Y_i^2$	$X_i Y_i$
1	56	-8	-19	64	361	152
3	72	-6	-3	36	9	18
5	69	-4	-6	16	36	24
8	72	-1	-3	1	9	3
10	82	1	7	1	49	7
12	76	3	1	9	1	3
15	93	6	18	36	324	108
18	80	9	5	81	25	45
		<b>Sommes</b>				
		<b>0</b>	<b>0</b>	<b>244</b>	<b>814</b>	<b>360</b>

$$\sum X_i^2 \quad \sum Y_i^2 \quad \sum X_i Y_i$$

1ère droite d'ajustement

$$y_i = ax_i + b$$

$$a = \frac{\sum X_i Y_i}{\sum X_i^2}$$

$$\bar{y} = a\bar{x} + b$$

$$\Rightarrow b = \bar{y} - a\bar{x}$$

$$a = 360 / 244 = 1,48$$

$$b = 75 - (1,48 \times 9) = 61,72$$

Equation de la 1ère droite:

$$y_i = 1,48 x_i + 61,72$$

$$X_i = x_i - \bar{x}$$

$$Y_i = y_i - \bar{y}$$

2ème droite d'ajustement

$$x_i = a' y_i + b'$$

$$a' = \frac{\sum X_i Y_i}{\sum Y_i^2}$$

$$\bar{x} = a' \bar{y} + b'$$

$$\Rightarrow b' = \bar{x} - a' \bar{y}$$

$$a' = 360 / 814 = 0,44$$

$$b' = 9 - (0,44 \times 75) = -27,17$$

Equation de la 2ème droite:

$$x_i = 0,44 y_i - 27,17$$

27

# Méthode des moindres carrés

## Version 2.

Publicité		Version 2		
CA		$x_i$	$y_i$	$x_i y_i$
1	56	1	56	56
3	72	9	216	216
5	69	25	345	345
8	72	64	576	576
10	82	100	820	820
12	76	144	912	912
15	93	225	1395	1395
18	80	324	1440	1440
		<b>Sommes</b>		
		<b>892</b>	<b>45814</b>	<b>5760</b>
<b>Moyennes</b>				
<b>9</b>	<b>75</b>			
NB Obs.		8		

$$\sum x_i^2 \quad \sum y_i^2 \quad \sum x_i y_i$$

1ère droite d'ajustement

$$y_i = ax_i + b$$

$$a = \frac{\sum x_i y_i - \bar{x} \bar{y}}{\sum x_i^2 - \bar{x}^2}$$

$$a = \frac{5760 - 9 \times 75}{892 - 9^2} = 1,48$$

$$\bar{y} = a\bar{x} + b$$

$$\Rightarrow b = \bar{y} - a\bar{x}$$

$$b = 75 - (1,48 \times 9) = 61,72$$

Equation :

$$y_i = 1,48 x_i + 61,72$$

2ème droite d'ajustement

$$x_i = a' y_i + b'$$

$$a' = \frac{\sum x_i y_i - \bar{x} \bar{y}}{\sum y_i^2 - \bar{y}^2}$$

$$\bar{x} = a' \bar{y} + b'$$

$$\Rightarrow b' = \bar{x} - a' \bar{y}$$

$$a' = \frac{5760 - 9 \times 75}{45814 - 75^2} = 0,44$$

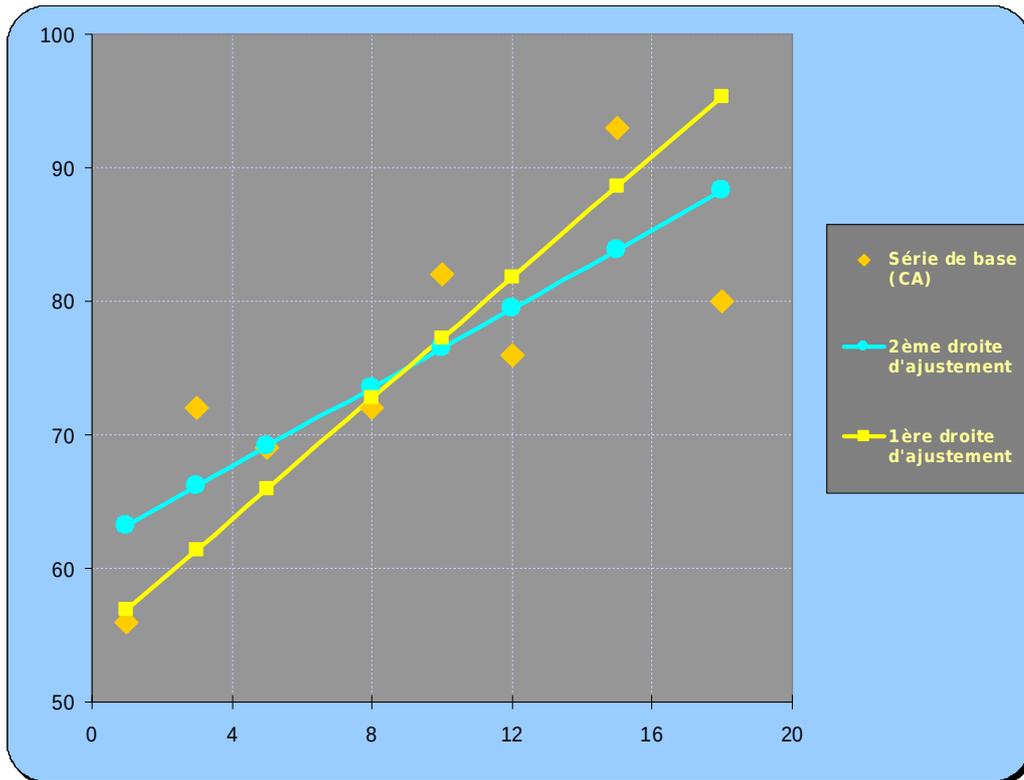
$$b' = 9 - (0,44 \times 75) = -27,17$$

Equation:

$$x_i = 0,44 y_i - 27,17$$

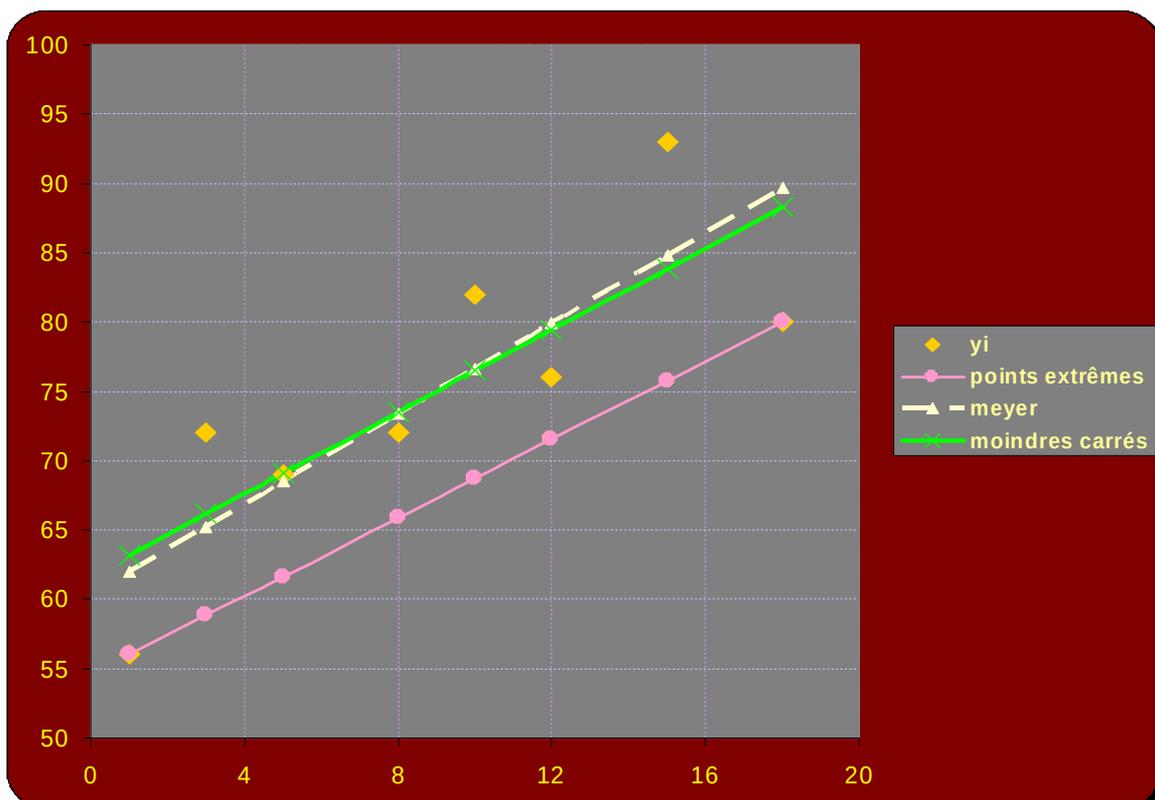
28

# Méthode des moindres carrés



29

# Comparaison des méthodes



30

# Plan

## Chapitre 1: ajustement & corrélation

I - Présentation

II - Méthodes basiques d'ajustement

III - Méthode des moindres carrés

IV - Corrélation

V - Autre exemple

VI - Ajustement, Excel et R

## Chapitre 2: Chroniques

I - Principes

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

## Chapitre 3: Tableaux de Contingence

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

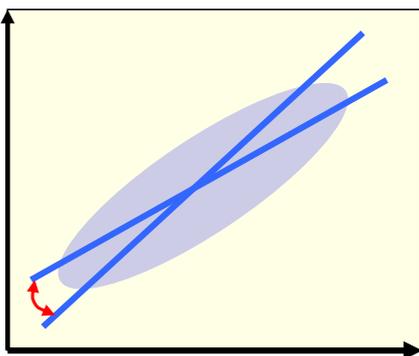
VI- Courbes de régression

VII- Autre cas

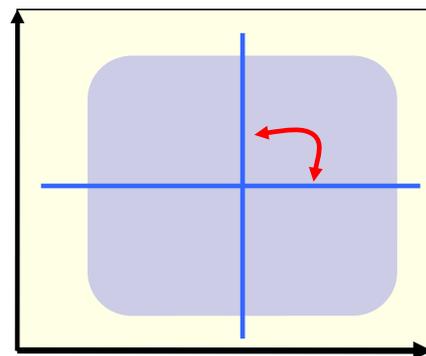
31

# Corrélation

- Corrélation: Etude de l'intensité de la liaison entre deux variables.
- Coefficient de corrélation ( $r$ ) permet de mesurer celle-ci. On a  $r \in [-1 ; 1]$
- $r$  n'a rien à voir avec  $R$ , Rapport de corrélation ( $R \in [0 ; 1]$ )
- $r$  calculé à partir des résultats intermédiaires des moindres carrés



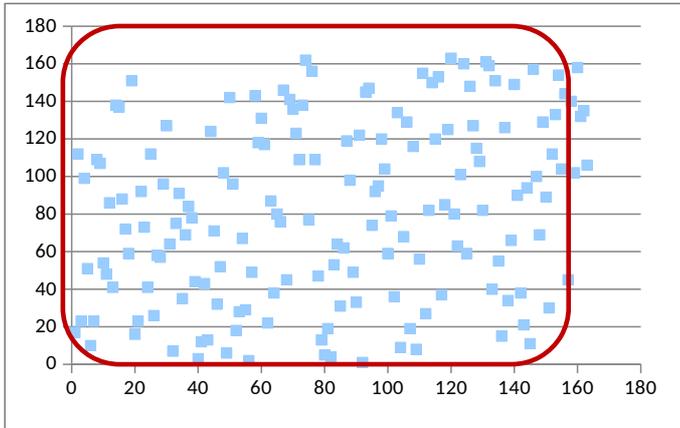
Plus l'angle est petit, plus les droites tendent à se confondre, plus les variables sont dépendantes => Liaison forte.  $r$  proche de 1 ou -1.



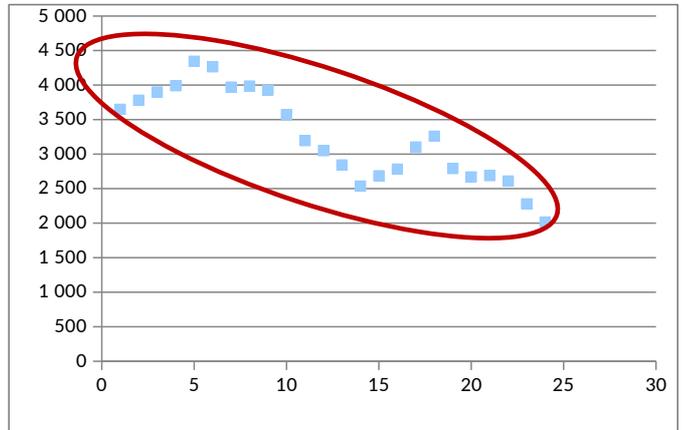
Plus l'angle est important, plus les droites sont perpendiculaires, plus les variables sont indépendantes => Liaison négligeable ou nulle.  $r$  proche de 0

32

# Corrélation



Points éparpillés partout.  
Nuage de points qui s'étale  
Droite sans intérêt



Points à peu près alignés.  
Nuage de points régulier, mince  
Droite probablement intéressante

33

# Corrélation

## Version 1

Publicité		Version 1				
$x_i$	$y_i$	$X_i$	$Y_i$	$X_i^2$	$Y_i^2$	$X_i Y_i$
1	56	-8	-19	64	361	152
3	72	-6	-3	36	9	18
5	69	-4	-6	16	36	24
8	72	-1	-3	1	9	3
10	82	1	7	1	49	7
12	76	3	1	9	1	3
15	93	6	18	36	324	108
18	80	9	5	81	25	45
<b>Moyennes</b>		<b>Sommes</b>				
<b>9</b>	<b>75</b>	<b>0</b>	<b>0</b>	<b>244</b>	<b>814</b>	<b>360</b>

$$r = \frac{\sum X_i Y_i}{\sqrt{\sum X_i^2 \sum Y_i^2}}$$

$$r = \frac{360}{\sqrt{244 \times 814}} = 0,81$$

## Version 2

Publicité	CA	Version 2		
$x_i$	$y_i$	$x_i^2$	$y_i^2$	$x_i y_i$
1	56	1	3136	56
3	72	9	5184	216
5	69	25	4761	345
8	72	64	5184	576
10	82	100	6724	820
12	76	144	5776	912
15	93	225	8649	1395
18	80	324	6400	1440
<b>Moyennes</b>		<b>Sommes</b>		
<b>9</b>	<b>75</b>	<b>892</b>	<b>45814</b>	<b>5760</b>

$$r = \frac{\sum x_i y_i - \bar{x} \bar{y}}{\sqrt{\left(\sum x_i^2 - \frac{-2}{n} \times \sum y_i^2 - \bar{y}^2\right)}}$$

$$r = \frac{5760 - 9 \times 75}{\sqrt{\frac{892}{8} - 9^2 \times \frac{45814}{8} - 75^2}} = 0,81$$

34

# Corrélation

- => a, a' et r ont toujours le même signe
- => r peut être calculé à partir de a et a' :

$$a = \frac{\sum X_i Y_i}{\sum X_i^2}$$

$$a' = \frac{\sum X_i Y_i}{\sum Y_i^2}$$

$$r = \frac{\sum X_i Y_i}{\sqrt{\sum X_i^2 \sum Y_i^2}}$$

$$r = \sqrt{a \times a'}$$

- **Attention! Formule piège!** Si a et a' < 0 on sait donc que r < 0.
- Or (a x a') est positif, ainsi que sa racine...ce qui pourrait laisser penser que r serait positif. Erreur car nous avons par exemple :

$$\sqrt{4} = 2$$

$$\sqrt{4} = -2$$

- La seconde solution n'est jamais proposée par défaut par les calculatrices ou les tableurs...

	C1			
	A	B	C	
1		4	2	

35

## Corrélation et étude de la variation.

Objectif: mesurer l'impact de x sur y, donc évaluer la pertinence du modèle.

$$SCT = \sum Y_i^2 = \sum (y_i - \bar{y})^2$$

Dispersion ou variation totale.  
Somme des carrés des écarts des observations par rapport au point moyen (SCT)

=

$$SCE = \sum (y_i \text{ Calc} - \bar{y})^2$$

Dispersion ou variation Expliquée.  
Somme des carrés des écarts des valeurs calculées par rapport au point moyen (SCE)

+

$$SCR = \sum (y_i - y_i \text{ Calc})^2$$

Dispersion ou variation non-expliquée ou Résiduelle.  
Somme des carrés des écarts des observations par rapport aux points calculés (SCR)

36

# Corrélation et étude de la variation.

Publicité	CA	Droite	SCT	SCE	SCR
$x_i$	$y_i$	$y_i$ Calc	$Y_i^2$	$(y_i \text{ calc} - y_i)^2$	$(y_i - y_i \text{ calc})^2$
1	56	63,20	361	139,32	51,79
3	72	66,15	9	78,37	34,25
5	69	69,10	36	34,83	0,01
8	72	73,52	9	2,18	2,32
10	82	76,48	49	2,18	30,52
12	76	79,43	1	19,59	11,74
15	93	83,85	324	78,37	83,68
18	80	88,28	25	176,32	68,54
<b>Moyennes</b>			<b>814</b>	<b>531,15</b>	<b>282,85</b>
9	75		<b>Sommes</b>		
NB Obs	7				

a	14
b	67
r	0,65

- $814 = 531,15 + 282,85$
- Mesure de la part de la variation expliquée par le modèle:  
 $= (SCE/SCT)$   
: Coefficient de détermination  
 $= (531,15 / 814)$   
 $= 0,65$

$69,10 = 1,48 \times 5 + 61,72$

$36 = (69 - 75)^2$

$34,83 = (69,10 - 75)^2$

$0,01 = (69 - 69,10)^2$

- Remarque importante:  $0,81^2 = 0,65$
- Donc  $r^2 =$
- Et  $= a \times a'$
- traduit donc la contribution de  $x_i$  dans les variations de  $y_i$ . Ici 65% de la variation/dispersion de  $y_i$  est expliquée par  $x_i$ .

## Corrélation

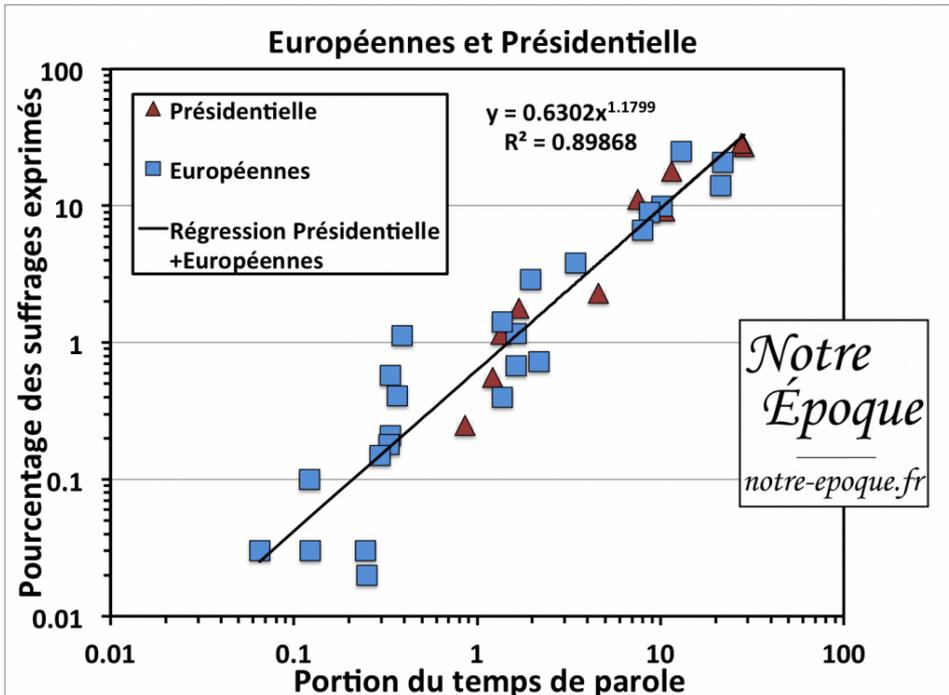
- ▶  $r$  est compris dans l'intervalle  $[-1, 1]$
- ▶  $r = 0 \Rightarrow$  Indépendance totale entre les variables, aucune corrélation
- ▶  $r = 1 \Rightarrow$  Corrélation positive totale; Droites confondues (croissantes)
- ▶  $r = -1 \Rightarrow$  Corrélation négative totale; Droites confondues (décroissantes)

- **r est petit**, compris entre -0,3 et 0,3. **Corrélation insignifiante**, indépendance (relative) des caractères
- **r est fort**,  $>0,87$  ou  $<-0,87 \Rightarrow$  Corrélation importante; Droites très proches. 75% des variations de  $y$  sont expliquées par  $x$ . (Cf B Py)
- **r important**,  $>0,75$  ou  $<-0,75$ . **Liaison relative entre les variables**. Présomption de dépendance entre les variables (à tester en fait).
- **autres valeurs de r**: Impossible de conclure à notre niveau.



# Exemple concret.

**Question:** les résultats électoraux des partis politiques sont-ils influencés par les temps de parole dont ces partis disposent dans les différents médias?



Présidentielle 2012: 94%  
Présidentielle 2007: 97%  
Présidentielle 2002: 82%  
Européennes 2014: 86%

<http://notre-epoque.fr/>

## Plan

### Chapitre 1: ajustement & corrélation

I - Présentation

II - Méthodes basiques d'ajustement

III - Méthode des moindres carrés

IV - Corrélation

V - Autre exemple

VI - Ajustement, Excel et R

### Chapitre 2: Chroniques

I - Principes

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

### Chapitre 3: Tableaux de Contingence

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

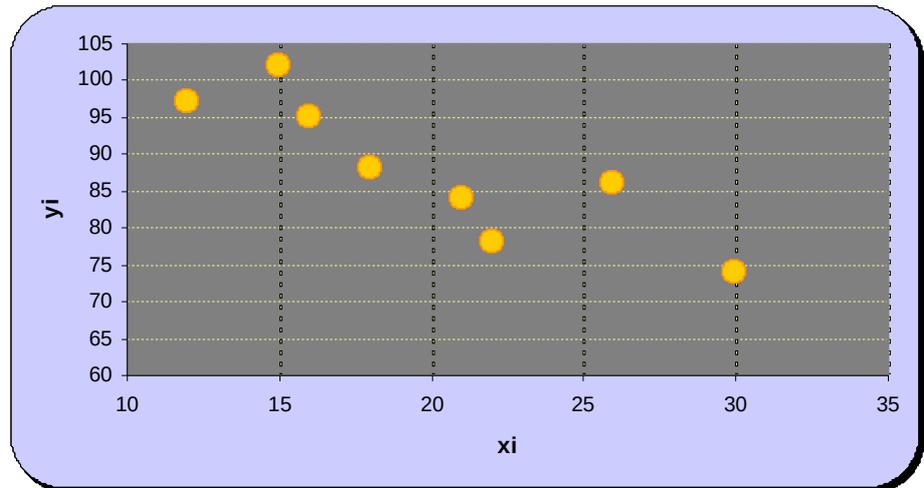
V- Ajustement & Corrélation

VI- Courbes de régression

VII- Autre cas

# Autre exemple

$x_i$	$y_i$
12	97
15	102
16	95
18	88
21	84
22	78
26	86
30	74
<b>Moyennes</b>	
<b>20</b>	<b>88</b>
NB Observations	8



41

# Autre exemple: version 1

		Version 1				
$x_i$	$y_i$	$X_i$	$Y_i$	$X_i^2$	$Y_i^2$	$X_i Y_i$
12	97	-8	9	64	81	-72
15	102	-5	14	25	196	-70
16	95	-4	7	16	49	-28
18	88	-2	0	4	0	0
21	84	1	-4	1	16	-4
22	78	2	-10	4	100	-20
26	86	6	-2	36	4	-12
30	74	10	-14	100	196	-140
<b>Moyennes</b>		<b>Sommes</b>				
<b>20</b>	<b>88</b>	<b>0</b>	<b>0</b>	<b>250</b>	<b>642</b>	<b>-346</b>
NB Observations	8					

- $a = -846 / 250 = -1,38$
- $b = 88 - 20 \times (-1,38) = 115,68$
- $a' = -846 / 642 = -0,54$
- $b' = 20 - 88 \times (-0,54) = 67,43$
- $r = -0,86$
- $\rho = 0,75$

42

# Autre exemple: version 2

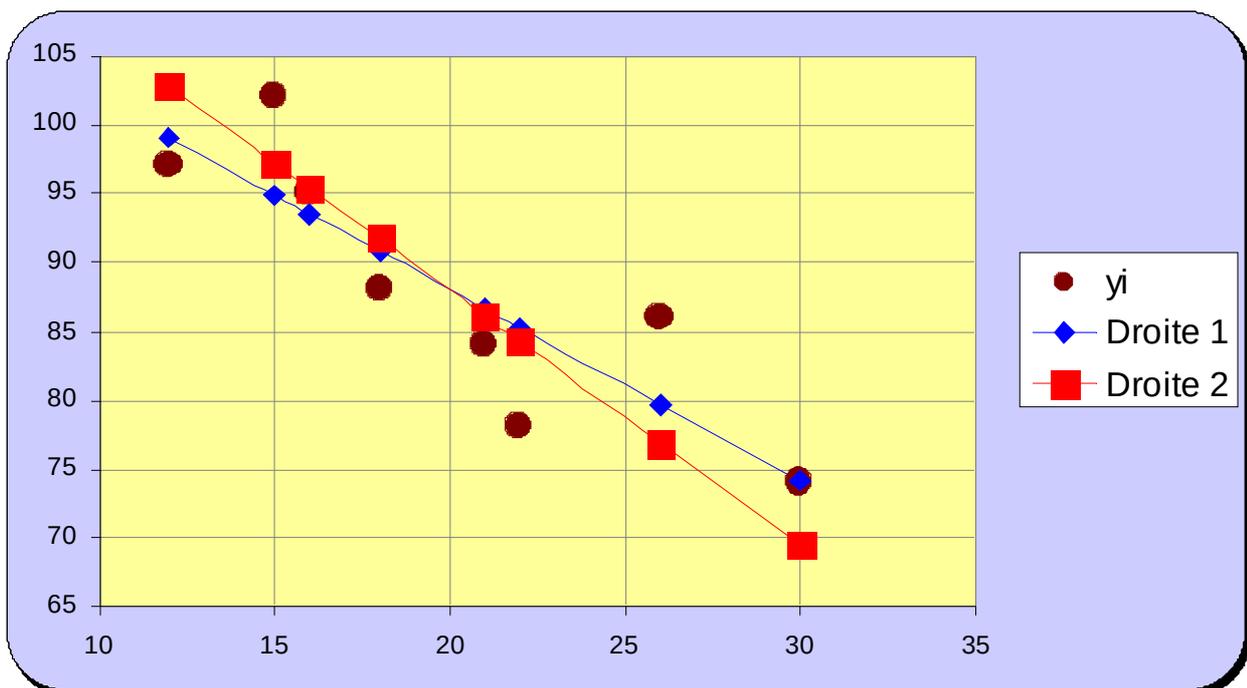
		Version 2		
xi	yi	xi <sup>2</sup>	yi <sup>2</sup>	xiyi
12	97	144	9409	1164
15	102	225	10404	1530
16	95	256	9025	1520
18	88	324	7744	1584
21	84	441	7056	1764
22	78	484	6084	1716
26	86	676	7396	2236
30	74	900	5476	2220
Moyennes		Sommes		
20	88	3450	62594	13734
NB Observations	8			

$$a = (1734 / 8 - 20 \times 88) / (3450 / 8 - 20^2) = -1,38$$

$$a' = (1734 / 8 - 20 \times 88) / (62594 / 8 - 88^2) = -0,54$$

43

# Autre exemple



44

# Plan

## Chapitre 1: ajustement & corrélation

*I - Présentation*

*II - Méthodes basiques d'ajustement*

*III - Méthode des moindres carrés*

*IV - Corrélation*

*V - Autre exemple*

**VI - Ajustement, Excel et R**

## Chapitre 2: Chroniques

I - Principes

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

## Chapitre 3: Tableaux de Contingence

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

VI- Courbes de régression

VII- Autre cas

45

# Ajustement, Corrélation et Tableur

Publicité		CA		
$x_i$	$y_i$		<u>Droite 1</u>	<u>Droite 2</u>
1	56		63,20	0,60
3	72		66,15	7,67
5	69		69,10	6,35
8	72		73,52	7,67
10	82		76,48	12,10
12	76		79,43	9,44
15	93		83,85	16,96
18	80		88,28	11,21
Moyennes			a	a'
9	75		1,48	0,44
Nb Obs	8		b	b'
			61,72	-24,17
r	0,80778426			
$\rho$	0,65251541			
	0,65251541			

46

# OpenOffice.org & Libreoffice



	A	B	C	D	E
1	Publicité	CA			
2	<b>x<sub>i</sub></b>	<b>y<sub>i</sub></b>	<b>Droite 1</b>	<b>Droite 2</b>	
3	1	56	63,2	0,6	
4	3	72	66,15	7,67	
5	5	69	69,1	6,35	
6	8	72	73,52	7,67	
7	10	82	76,48	12,1	
8	12	76	79,43	9,44	
9	15	93	83,85	16,96	
10	18	80	88,28	11,21	
11	<b>Moyennes</b>				
12	9	75		a	b
13	<b>Nb Obs.</b>	8		1,48	61,72
14				a'	b'
15	r	0,81		0,44	-24,17
16		0,65			
17	p	0,65			

Seules valeurs saisies

Aucun calcul intermédiaire.

47

# OpenOffice.org & LibreOffice

	A	B	C	D
1	Publicité	CA		
2	<b>x<sub>i</sub></b>	<b>y<sub>i</sub></b>	<b>Droite 1</b>	<b>Droite 2</b>
3	1	56	{=TENDANCE(B3:B10;A3:A10;A3:A10)}	{=TENDANCE(A3:A10;B3:B10;B3:B10)}
4	3	72	{=TENDANCE(B3:B10;A3:A10;A3:A10)}	{=TENDANCE(A3:A10;B3:B10;B3:B10)}
5	5	69	{=TENDANCE(B3:B10;A3:A10;A3:A10)}	{=TENDANCE(A3:A10;B3:B10;B3:B10)}
6	8	72	{=TENDANCE(B3:B10;A3:A10;A3:A10)}	{=TENDANCE(A3:A10;B3:B10;B3:B10)}
7	10	82	{=TENDANCE(B3:B10;A3:A10;A3:A10)}	{=TENDANCE(A3:A10;B3:B10;B3:B10)}
8	12	76	{=TENDANCE(B3:B10;A3:A10;A3:A10)}	{=TENDANCE(A3:A10;B3:B10;B3:B10)}
9	15	93	{=TENDANCE(B3:B10;A3:A10;A3:A10)}	{=TENDANCE(A3:A10;B3:B10;B3:B10)}
10	18	80	{=TENDANCE(B3:B10;A3:A10;A3:A10)}	{=TENDANCE(A3:A10;B3:B10;B3:B10)}
11	<b>Moyennes</b>			
12	=MOYENNE(A3:A10)	=MOYENNE(B3:B10)	a	b
13	<b>Nb Obs.</b>	=NB(B3:B10)	{=DROITEREG(B3:B10;A3:A10)}	{=DROITEREG(B3:B10;A3:A10)}
14			a'	b'
15	r	=COEFFICIENT.CORRELATION(A3:A10;B3:B10)	{=DROITEREG(A3:A10;B3:B10)}	{=DROITEREG(A3:A10;B3:B10)}
16		=COEFFICIENT.DETERMINATION(A3:A10;B3:B10)		
17	p	=B15^2		

Attention: fonctions matricielles ici sous OpenOffice.

48

# Excel

	A	B	C	D
1	Publicité	CA		
2	$x_i$	$y_i$	Droite 1	Droite 2
3	1	56	63,20	0,60
4	3	72	66,15	7,67
5	5	69	69,10	6,35
6	8	72	73,52	7,67
7	10	82	76,48	12,10
8	12	76	79,43	9,44
9	15	93	83,85	16,96
10	18	80	88,28	11,21
11	<b>Moyennes</b>			
12	9	75	a	b
13	Nb Obs.	8	1,48	61,72
14			a'	b'
15	r	0,81	0,44	-24,17
16		0,65		
17	$\square$	0,65		
18				

Aucun calcul intermédiaire.

49

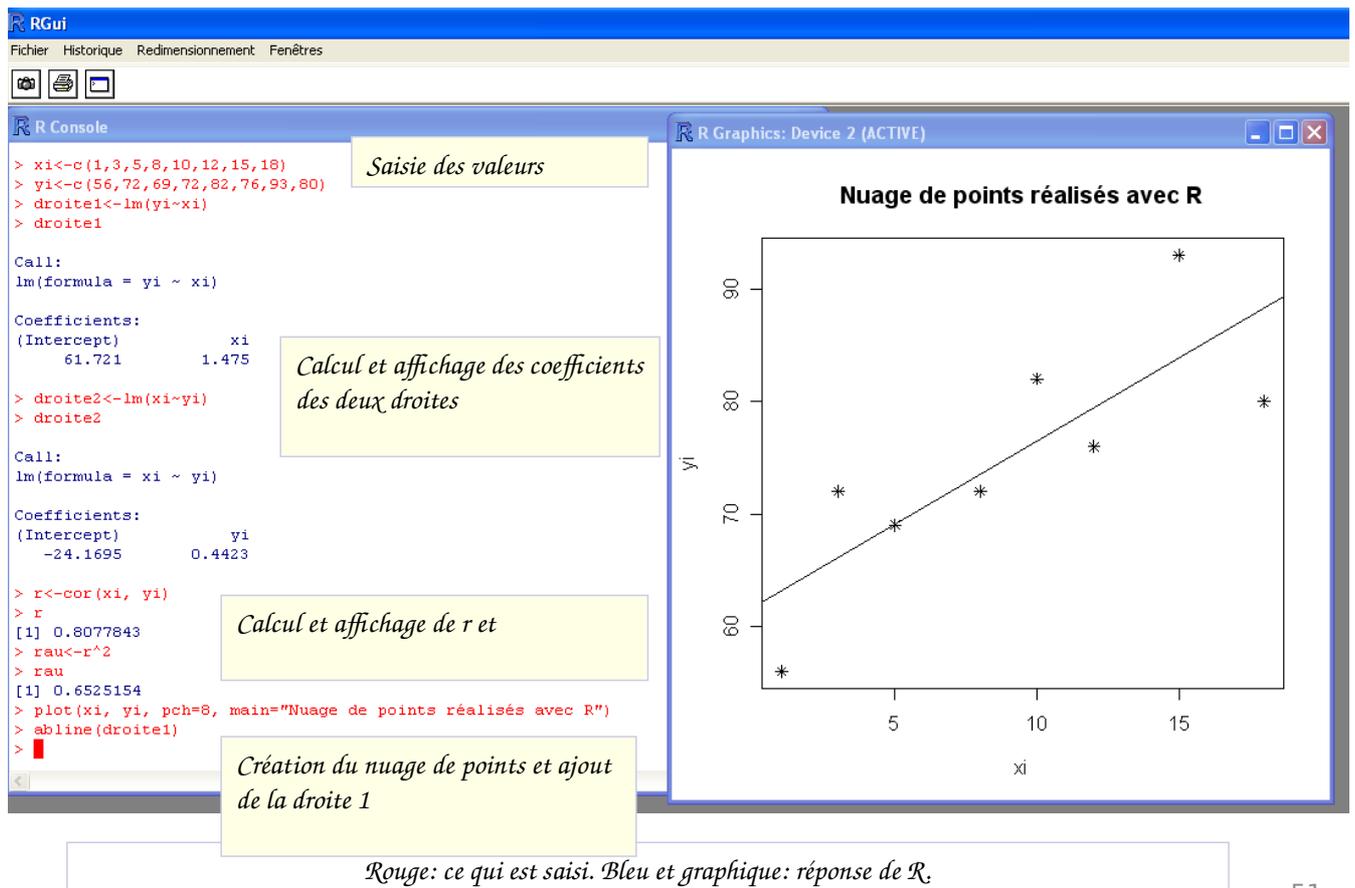
# Excel

	A	B	C	D
1	Publicité	CA		
2	$x_i$	$y_i$	Droite 1	Droite 2
3	1	56	=TENDANCE(\$B\$3:\$B\$10,\$A\$3:\$A\$10,A3)	=TENDANCE(A3:A10,B3:B10,B3:B10)
4	3	72	=TENDANCE(\$B\$3:\$B\$10,\$A\$3:\$A\$10,A4)	=TENDANCE(A3:A10,B3:B10,B3:B10)
5	5	69	=TENDANCE(\$B\$3:\$B\$10,\$A\$3:\$A\$10,A5)	=TENDANCE(A3:A10,B3:B10,B3:B10)
6	8	72	=TENDANCE(\$B\$3:\$B\$10,\$A\$3:\$A\$10,A6)	=TENDANCE(A3:A10,B3:B10,B3:B10)
7	10	82	=TENDANCE(\$B\$3:\$B\$10,\$A\$3:\$A\$10,A7)	=TENDANCE(A3:A10,B3:B10,B3:B10)
8	12	76	=TENDANCE(\$B\$3:\$B\$10,\$A\$3:\$A\$10,A8)	=TENDANCE(A3:A10,B3:B10,B3:B10)
9	15	93	=TENDANCE(\$B\$3:\$B\$10,\$A\$3:\$A\$10,A9)	=TENDANCE(A3:A10,B3:B10,B3:B10)
10	18	80	=TENDANCE(\$B\$3:\$B\$10,\$A\$3:\$A\$10,A10)	=TENDANCE(A3:A10,B3:B10,B3:B10)
11	<b>Moyennes</b>			
12	=MOYENNE(A3:A10)	=MOYENNE(B3:B10)	a	b
13	Nb Obs.	=NB(B3:B10)	=DROITEREG(B3:B10;A3:A10)	=ORDONNEE.ORIGINE(B3:B10;A3:A10)
14			a'	b'
15	r	=COEFFICIENT.CORRELATION(A3:A10;B3:B10)	=DROITEREG(A3:A10;B3:B10)	=ORDONNEE.ORIGINE(A3:A10;B3:B10)
16		=COEFFICIENT.DETERMINATION(A3:A10;B3:B10)		
17	P	=B15^2		
18				

Fonctions matricielles pas obligatoires sous Excel.

50

# Ajustement, Corrélacion et R



51

## Ajustement, Corrélacion

- Application à quelques données
  - Régions: PIB, Population, Superficie, Nombre d'entreprises
  - 200 premiers groupes mondiaux: CA, Résultat Net, Effectif

52

# Plan

## *Chapitre 1: ajustement & corrélation*

*I - Présentation*

*II - Méthodes basiques d'ajustement*

*III - Méthode des moindres carrés*

*IV - Corrélation*

*V - Autre exemple*

*VI - Ajustement, Excel et R*

## **Chapitre 2: Chroniques**

**I - Principes**

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

## **Chapitre 3: Tableaux de Contingence**

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

VI- Courbes de régression

VII- Autre cas

53

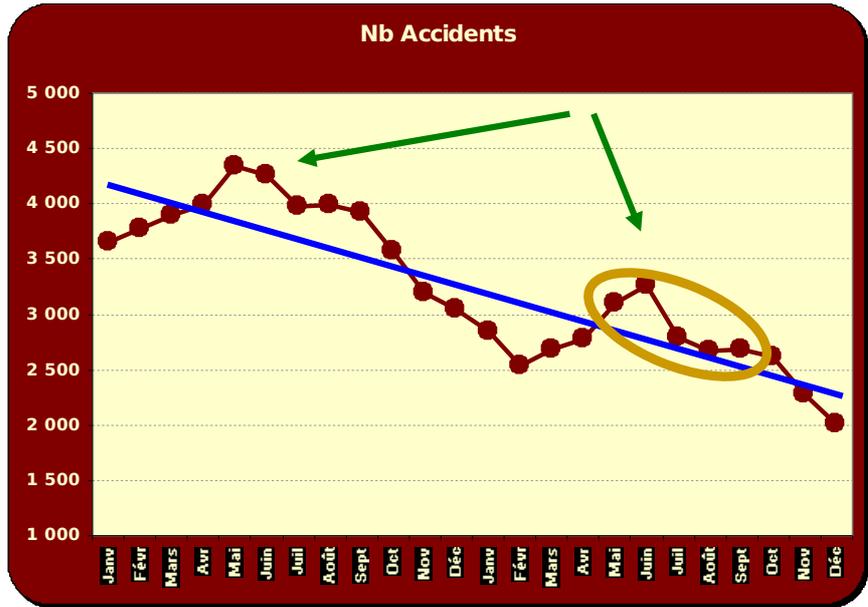
## Chroniques - Principes

- *série chronologique*: une série statistique ordonnée en fonction du temps.
- Etude intéressante: peut permettre de prévoir l'évolution du phénomène observé dans le temps.
- Cas particulier des séries à un caractère.
- Fluctuations: manifestations du temps.

54

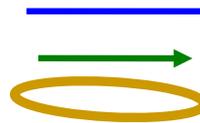
# Chroniques - Principes

Mois	Nb Accidents
Janv	3 650
Févr	3 780
Mars	3 900
Avr	3 991
Mai	4 345
Juin	4 265
Juil	3 970
Août	3 986
Sept	3 928
Oct	3 572
Nov	3 195
Déc	3 050
Janv	2 843
Févr	2 537
Mars	2 686
Avr	2 782
Mai	3 103
Juin	3 260
Juil	2 791
Août	2 669
Sept	2 689
Oct	2 610
Nov	2 280
Déc	2 016



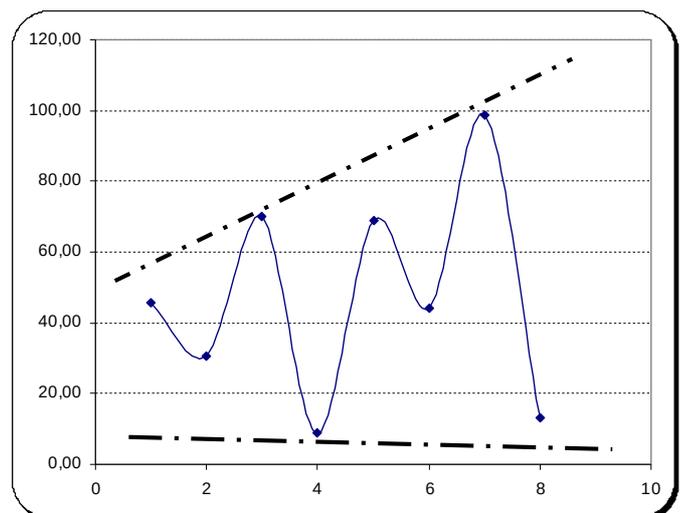
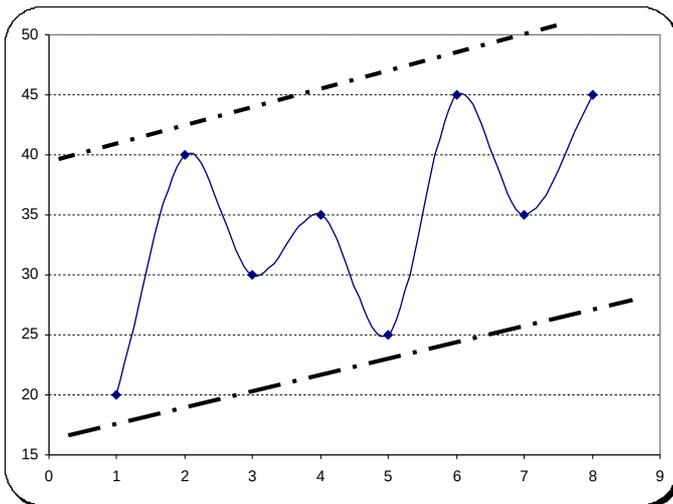
Phénomène caractérisé par:

- Tendance ou trend
- Variations saisonnières
- Variations accidentelles



55

# Chroniques - Principes



Modèle additif

Modèle Multiplicatif

Phénomènes caractérisés par:

- Tendance ou trend
- Variations saisonnières
- Variations accidentelles

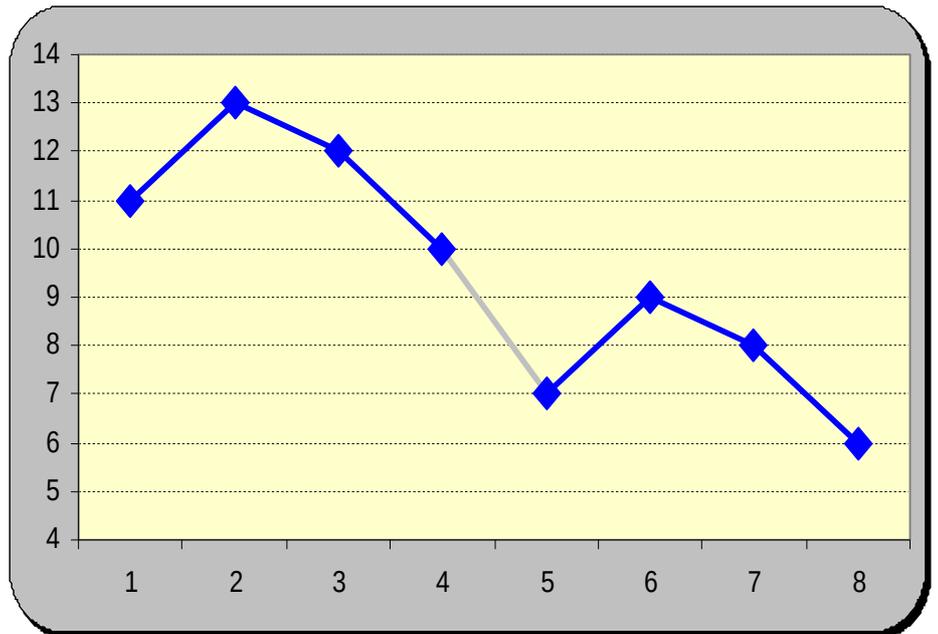
**Difficulté** (relative): savoir identifier le modèle le plus approprié.

56

# Chroniques - Principes

**Exemples suivants:**  
Périodicité: Trimestres  
Durée: 2 ou 3 ans

Trimestres	Accidents
$x_i$	$y_i$
1	11
2	13
3	12
4	10
5	7
6	9
7	8
8	6
<b>Moyennes</b>	
4,5	9,5



## Objectif:

Disposer d'un modèle permettant simulation et prévision  
=>Intégrant **tendance, saisonnalité, et variations accidentelles.**

57

# Chroniques - Principes

## Notations:

- Temps. (Variable déterminante) :  $t$
- Variable ou (série déterminée) :  $y_t$
- Série ajustée:  $y^*_t$
- Trend (tendance):  $f(t)$
- Coefficients partiels:  $S_t$
- Coefficients saisonniers :  $S_j$  (et j: nombre de saisons)
- Correction des variations saisonnières :  $CVS_t$
- Variations accidentelles :  $\epsilon_t$

58

# Plan

## Chapitre 1: ajustement & corrélation

I - Présentation

II - Méthodes basiques d'ajustement

III - Méthode des moindres carrés

IV - Corrélation

V - Autre exemple

VI - Ajustement, Excel et R

## Chapitre 2: Chroniques

I - Principes

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

## Chapitre 3: Tableaux de Contingence

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

VI- Courbes de régression

VII- Autre cas

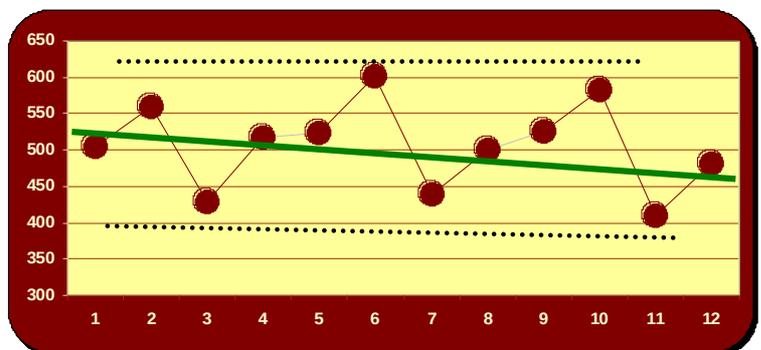
59

# Modèle additif

**Données: Immatriculations** de Voitures particulières au cours des 3 dernières années.

**Source:** <http://www.statistiques.equipement.gouv.fr/>

Années	Trimestres	Immatriculations	Milliers
2004	1	506 884	504
	2	560 923	560
	3	428 531	428
	4	517 371	517
2005	5	526 233	524
	6	601 510	601
	7	439 006	439
	8	501 040	501
2006	9	526 502	525
	10	582 727	582
	11	409 138	409
	12	482 182	482



Total pour chaque année: 2009, 2065, et 1998. => Stabilité, voire légère baisse.

Périodicité des immatriculations chaque année.

Amplitude des fluctuations comparables pour chaque année...

...Mais quand même pas à l'identique!

Modèle devant intégrer:

Tendance

Mouvements saisonniers

✓ Mouvements « accidentels » ou résiduels.

✓ Equation du modèle:  $y^*_t = f(t) + S_j + \epsilon_t$

60

# Modèle additif

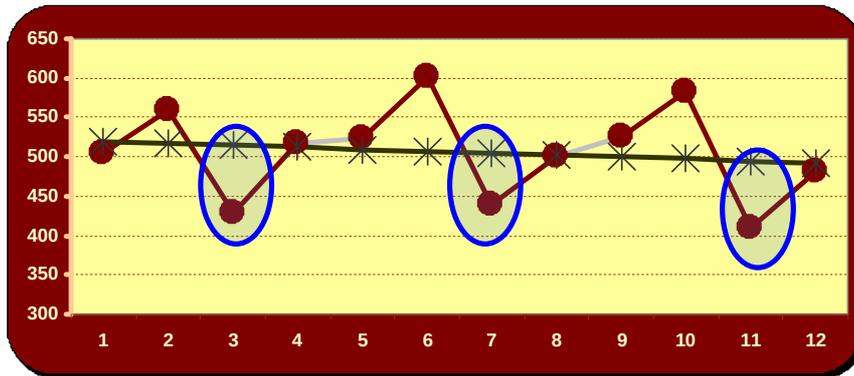
## Détermination du trend: Moindres Carrés.

t	yt	Tt	Yt	Tt <sup>2</sup>	Yt <sup>2</sup>	TtYt
1	504	-5,5	-2	30,25	4	11
2	560	-4,5	54	20,25	2916	-243
3	428	-3,5	-78	12,25	6084	273
4	517	-2,5	11	6,25	121	-27,5
5	524	-1,5	18	2,25	324	-27
6	601	-0,5	95	0,25	9025	-47,5
7	439	0,5	-67	0,25	4489	-33,5
8	501	1,5	-5	2,25	25	-7,5
9	525	2,5	19	6,25	361	47,5
10	582	3,5	76	12,25	5776	266
11	409	4,5	-97	20,25	9409	-436,5
12	482	5,5	-24	30,25	576	-132
<b>Moyennes</b>		<b>Sommes</b>				
6,5	506	0	0	143	39110	-357

droite (Trend)	
a	-2,50
b	522,23
r	-0,15

a	-2,50
b	522,23
r	-0,15

$$f(t) = -2,50t + 522,23$$



**Droite:**  
Position identique par rapport aux données pour des trimestres comparables  
=> Base de la détermination des coefficients saisonniers.

61

# Modèle additif

## Détermination des coefficients saisonniers: Différences par rapport au trend.

t	yt	f(t) (Trend)	St=yt-f(t)	Sj	yt estimée	CVS <sub>t</sub>	CVS <sub>t</sub>	ε <sub>t</sub>	ε <sub>t</sub>
1	504	519,73	-15,73	7,92	527,65	496,08	496,08	23,65	23,65
2	560	517,23	42,77	73,75	590,99	486,25	486,25	30,99	30,99
3	428	514,74	-86,74	-79,42	435,32	507,42	507,42	7,32	7,32
4	517	512,24	4,76	-2,26	509,99	519,26	519,26	-7,01	-7,01
5	524	509,74	14,26	7,92	517,67	516,08	516,08	-6,33	-6,33
6	601	507,25	93,75	73,75	581,00	527,25	527,25	-20,00	-20,00
7	439	504,75	-65,75	-79,42	425,33	518,42	518,42	-13,67	-13,67
8	501	502,26	-1,26	-2,26	500,00	503,26	503,26	-1,00	-1,00
9	525	499,76	25,24	7,92	507,68	517,08	517,08	-17,32	-17,32
10	582	497,26	84,74	73,75	571,01	508,25	508,25	-10,99	-10,99
11	409	494,77	-85,77	-79,42	415,35	488,42	488,42	6,35	6,35
12	482	492,27	-10,27	-2,26	490,01	484,26	484,26	8,01	8,01

$$y_t \text{ estimée: } y_t^* = f(t) + S_j$$

$$s_t = y_t - f(t)$$

S<sub>j</sub>: moyenne des S<sub>t</sub> des trimestres correspondants

$$CVS_t = y_t - S_j \quad \text{ou} \quad CVS_t = f(t) + S_t - S_j$$

$$\epsilon_t = f(t) - CVS_t \quad \text{ou} \quad \epsilon_t = y_t - y_t^*$$

62

# Modèle additif

Equation du modèle:

$$y_t^* = -2,50t + 522,23 + \begin{bmatrix} 7,92 \\ 73,75 \\ -79,42 \\ -2,26 \end{bmatrix}$$

Prévision année 4: t = 13 à 16

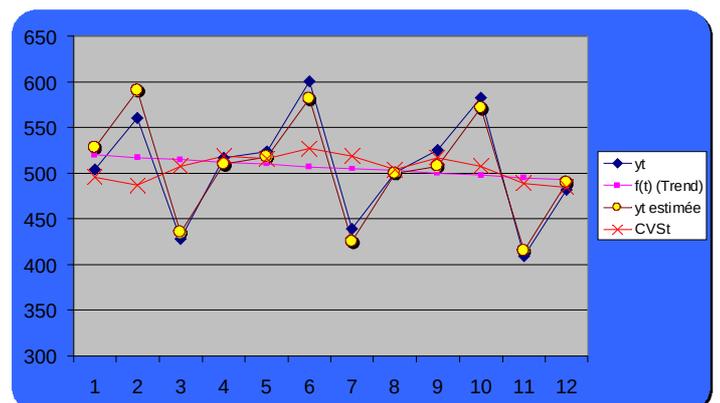
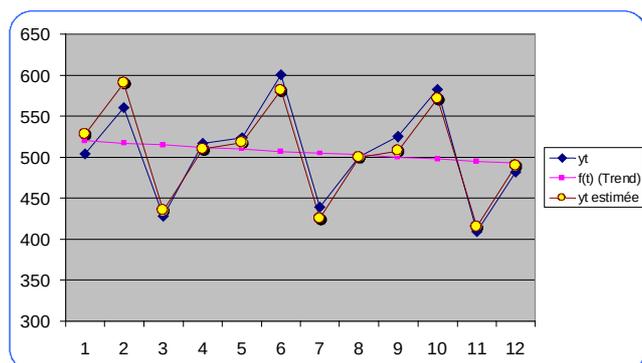
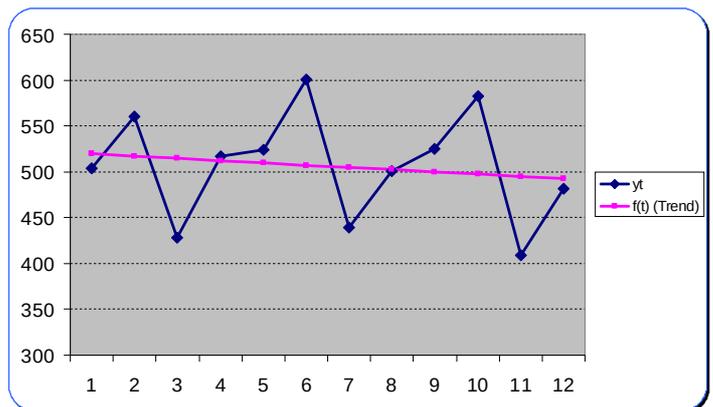
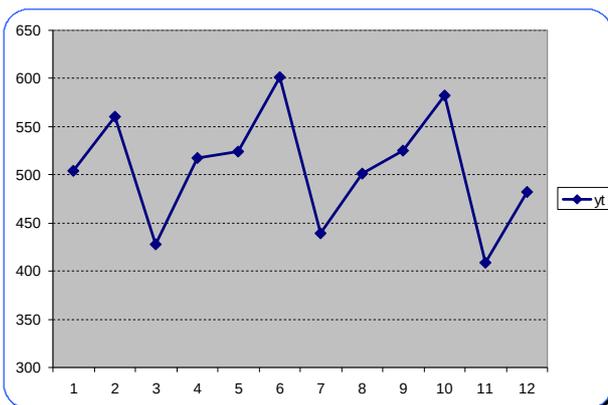
t	Prévisions
13	497,69
14	561,03
15	405,36
16	480,03

$$405,36 = -2,50 \times 15 + (-79,42)$$



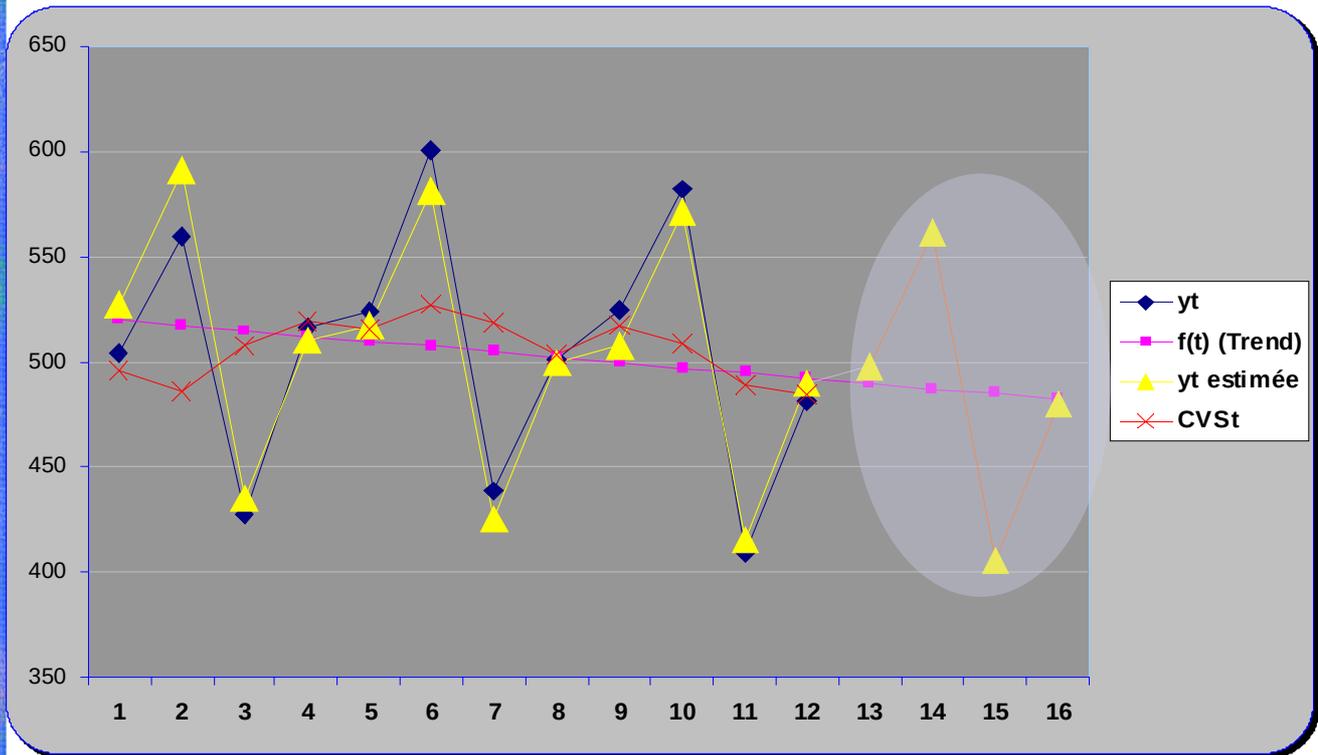
63

## Modèle additif: Graphique



64

# Modèle additif: Graphique



65

## Modèle additif

Evaluer le modèle

t	$y_t$	$y_t^*$	Ecart (%)
1	504	527,65	4,69%
2	560	590,99	5,53%
3	428	435,32	1,71%
4	517	509,99	-1,36%
5	524	517,67	-1,21%
6	601	581,00	-3,33%
7	439	425,33	-3,11%
8	501	500,00	-0,20%
9	525	507,68	-3,30%
10	582	571,01	-1,89%
11	409	415,35	1,55%
12	482	490,01	1,66%

$$1,71\% = (517 - 509,99)/517 \times 100$$

66

# Plan

## *Chapitre 1: ajustement & corrélation*

*I - Présentation*

*II - Méthodes basiques d'ajustement*

*III - Méthode des moindres carrés*

*IV - Corrélation*

*V - Autre exemple*

*VI - Ajustement, Excel et R*

## **Chapitre 2: Chroniques**

*I - Principes*

*II- Modèle additif*

**III - Modèle multiplicatif**

IV - Moyennes mobiles

## **Chapitre 3: Tableaux de Contingence**

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

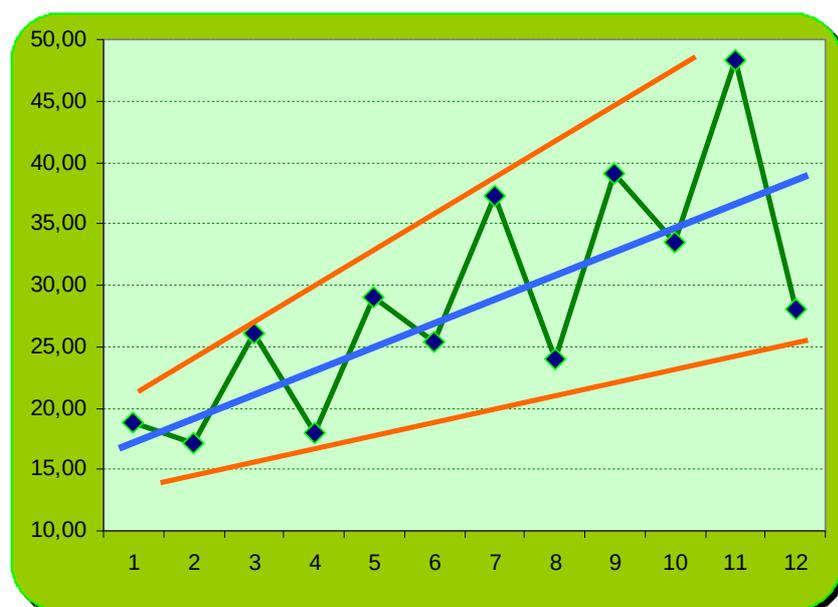
VI- Courbes de régression

VII- Autre cas

67

# Modèle Multiplicatif

t	y <sub>t</sub>
1	18,84
2	17,20
3	26,14
4	18,00
5	28,97
6	25,34
7	37,21
8	24,05
9	39,09
10	33,49
11	48,28
12	28,00



Tendance, Fluctuations saisonnières, amplitude divergente avec le temps => A reproduire dans le modèle.

68

# Modèle Multiplicatif

## Détermination du trend: Moindres Carrés.

t	y <sub>t</sub>	t <sub>t</sub>	Y <sub>t</sub>	t <sub>t</sub> <sup>2</sup>	Y <sub>t</sub> <sup>2</sup>	t <sub>t</sub> Y <sub>t</sub>	f(t)
1	18,84	-9,88	-5,50	30,25	97,57	54,33	18,44
2	17,20	-11,52	-4,50	20,25	132,65	51,83	20,31
3	26,14	-2,58	-3,50	12,25	6,64	9,02	22,18
4	18,00	-10,72	-2,50	6,25	114,86	26,79	24,05
5	28,97	0,25	-1,50	2,25	0,06	-0,38	25,91
6	25,34	-3,38	-0,50	0,25	11,41	1,69	27,78
7	37,21	8,49	0,50	0,25	72,12	4,25	29,65
8	24,05	-4,67	1,50	2,25	21,79	-7,00	31,52
9	39,09	10,37	2,50	6,25	107,59	25,93	33,39
10	33,49	4,77	3,50	12,25	22,78	16,70	35,26
11	48,28	19,56	4,50	20,25	382,69	88,03	37,13
12	28,00	-0,72	5,50	30,25	0,51	-3,95	39,00
<b>Moyennes</b>				<b>Sommes</b>			
6,5	28,72	0,00	0,00	143,00	970,68	267,25	

a	1,87
b	16,57
r	0,72



### Droite:

Position identique par rapport aux données pour des trimestres comparables  
 => **Base** de la détermination des coefficients saisonniers.

69

# Modèle Multiplicatif

## Détermination des coefficients saisonniers:

Rapports au trend.

t	y <sub>t</sub>	Trend (f <sub>t</sub> )	S <sub>t</sub> =y <sub>t</sub> /f <sub>t</sub>	S <sub>j</sub>	y <sub>t</sub> estimée	CVS <sub>t</sub>	CVS <sub>t</sub>	ε <sub>t</sub>	ε <sub>t</sub>
1	18,84	18,44	1,02	1,10	20,35	17,07	17,07	0,93	0,93
2	17,2	20,31	0,85	0,90	18,34	19,05	19,05	0,94	0,94
3	26,14	22,18	1,18	1,24	27,60	21,00	21,00	0,95	0,95
4	18	24,05	0,75	0,74	17,87	24,22	24,22	1,01	1,01
5	28,97	25,91	1,12	1,10	28,60	26,25	26,25	1,01	1,01
6	25,34	27,78	0,91	0,90	25,09	28,06	28,06	1,01	1,01
7	37,21	29,65	1,25	1,24	36,91	29,90	29,90	1,01	1,01
8	24,05	31,52	0,76	0,74	23,43	32,36	32,36	1,03	1,03
9	39,09	33,39	1,17	1,10	36,84	35,42	35,42	1,06	1,06
10	33,49	35,26	0,95	0,90	31,84	37,09	37,09	1,05	1,05
11	48,28	37,13	1,30	1,24	46,21	38,79	38,79	1,04	1,04
12	28	39,00	0,72	0,74	28,98	37,68	37,68	0,97	0,97

$$y_t \text{ estimée: } y_t^* = f(t) \times S_j$$

$$s_t = y_t / f(t)$$

S<sub>j</sub>: moyenne des S<sub>t</sub> des trimestres correspondants

$$CVS_t = y_t / S_j \quad \text{ou} \quad CVS_t = f(t) \times S_t / S_j$$

$$\epsilon_t = CVS_t / f(t) \quad \text{ou} \quad \epsilon_t = y_t / y_t^*$$

70

# Modèle multiplicatif

Equation du modèle:

$$y_t^* = (1,87t + 16,57) \times \begin{bmatrix} 1,10 \\ 0,9 \\ 1,24 \\ 0,74 \end{bmatrix}$$

Prévision année 4: t = 13 à 16

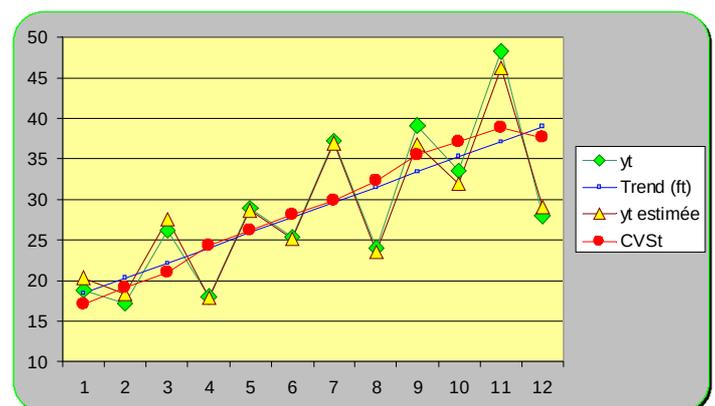
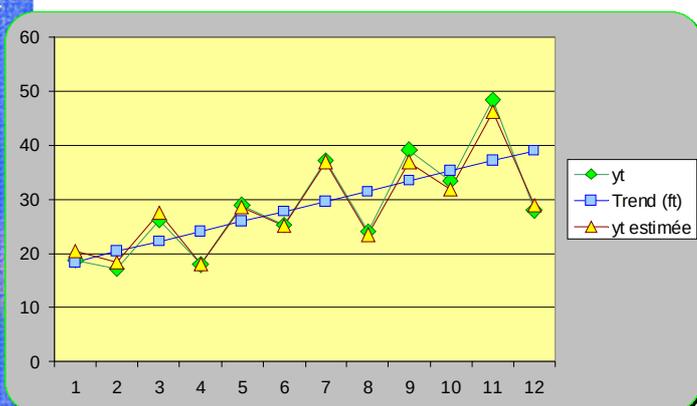
t	Prévisions
13	45,09
14	38,59
15	55,52
16	34,54

$$55,52 = (1,87 \times 15 + 16,57) \times 1,24$$



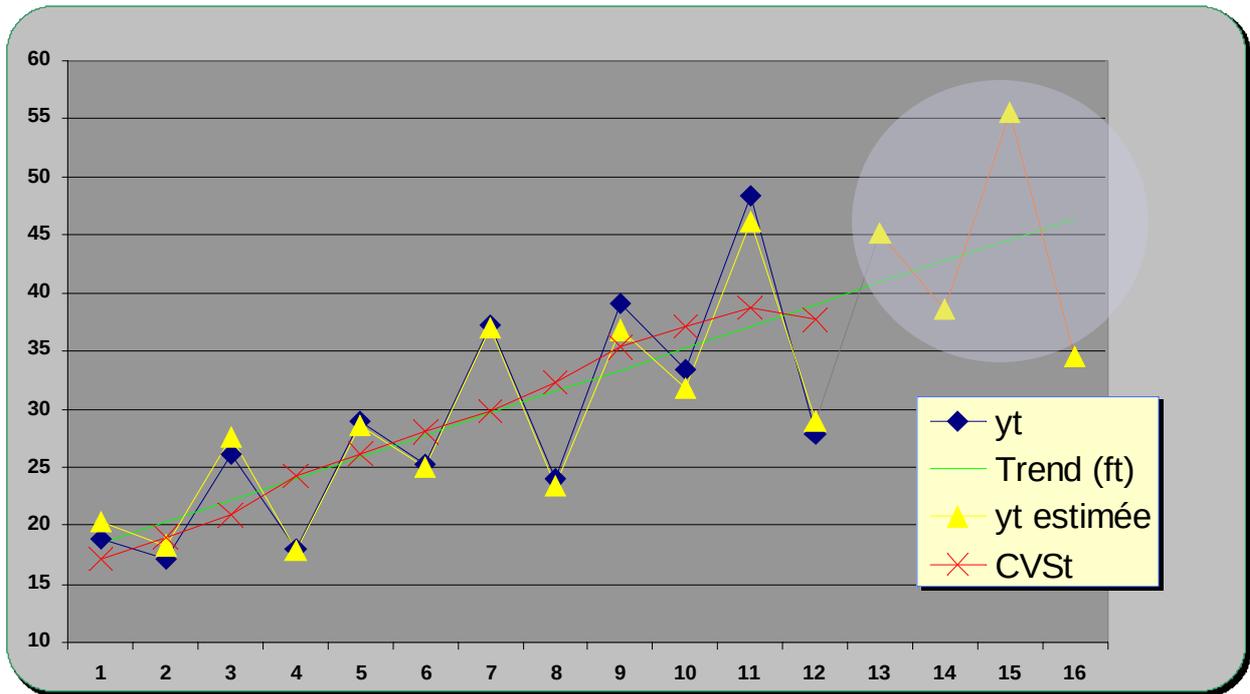
71

## Modèle multiplicatif: Graphique



72

# Modèle multiplicatif



73

# Modèle multiplicatif

Evaluer le modèle

t	yt	yt estimée	Ecart
1	18,84	20,35	-8,00%
2	17,2	18,34	-6,61%
3	26,14	27,60	-5,59%
4	18	17,87	0,72%
5	28,97	28,60	1,29%
6	25,34	25,09	1,00%
7	37,21	36,91	0,81%
8	24,05	23,43	2,59%
9	39,09	36,84	5,74%
10	33,49	31,84	4,94%
11	48,28	46,21	4,28%
12	28	28,98	-3,51%

$$0,72\% = (18-17,87)/18 \times 100$$

74

# Plan

## *Chapitre 1: ajustement & corrélation*

*I - Présentation*

*II - Méthodes basiques d'ajustement*

*III - Méthode des moindres carrés*

*IV - Corrélation*

*V - Autre exemple*

*VI - Ajustement, Excel et R*

## **Chapitre 2: Chroniques**

*I - Principes*

*II- Modèle additif*

*III - Modèle multiplicatif*

**IV - Moyennes mobiles**

## **Chapitre 3: Tableaux de Contingence**

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

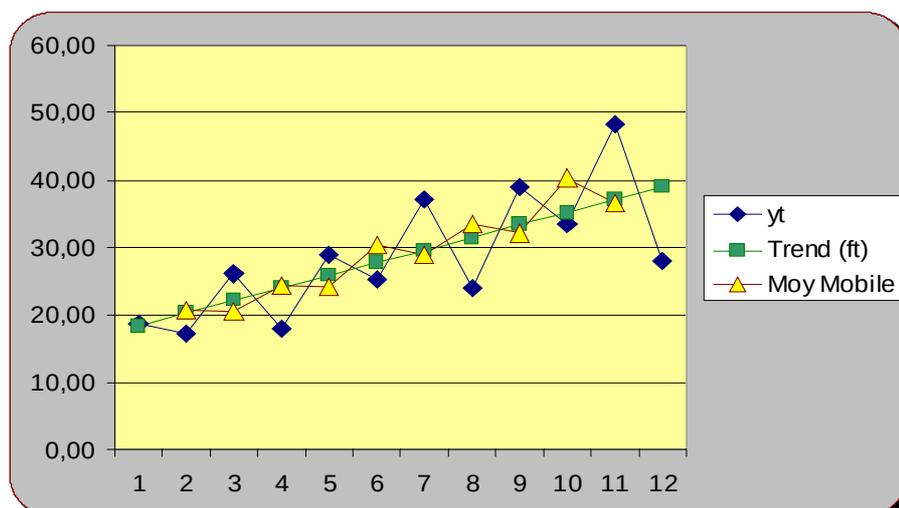
VI- Courbes de régression

VII- Autre cas

75

# Moyennes mobiles

- Critique de la méthode précédente: Tendence est linéaire
- Aucune raison que ce soit le cas systématiquement
- Tendence peut avoir de légères fluctuations



76

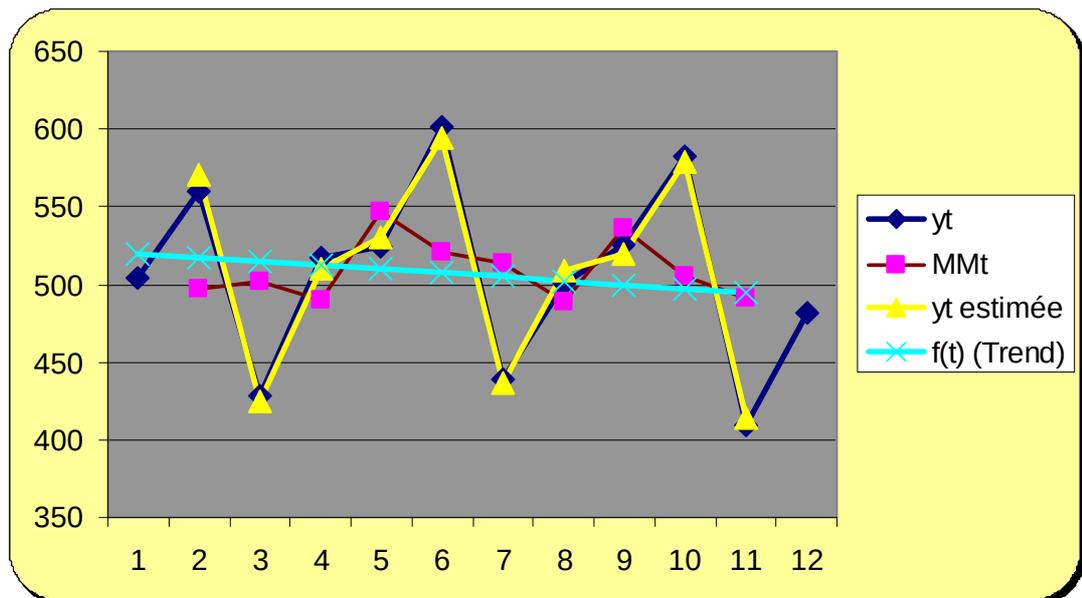
# Moyennes mobiles - Principe

t	yt	MMt	St=yt-MMt	Sj	Sj rectifiés	yt estimée
1	504			-17,17	-16,93	
2	560	497,33	62,67	73,00	73,24	570,57
3	428	501,67	-73,67	-76,78	-76,54	425,13
4	517	489,67	27,33	20,00	20,24	509,90
5	524	547,33	-23,33	-17,17	-16,93	530,40
6	601	521,33	79,67	73,00	73,24	594,57
7	439	513,67	-74,67	-76,78	-76,54	437,13
8	501	488,33	12,67	20,00	20,24	508,57
9	525	536,00	-11,00	-17,17	-16,93	519,07
10	582	505,33	76,67	73,00	73,24	578,57
11	409	491,00	-82,00	-76,78	-76,54	414,46
12	482			20,00	20,24	

-0,94	0,00
-0,24	

77

# Moyennes mobiles - Principe



78

# Conclusion

- Autre exemple
- Modèles simples, et efficaces
- Facilement portables sur informatique
- Résultats intéressants
- Calculs des Coef. Sais. adaptables selon les situations et dans le temps
- Possibilité d'insérer des outils statistiques élaborés
- => Très utilisé

79

Tableaux de  
contingence

Caractères quantitatifs

80

# Plan

## Chapitre 1: ajustement & corrélation

I - Présentation

II - Méthodes basiques d'ajustement

III - Méthode des moindres carrés

IV - Corrélation

V - Autre exemple

VI - Ajustement, Excel et R

## Chapitre 2: Chroniques

I - Principes

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

## Chapitre 3: Tableaux de Contingence

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

VI- Courbes de régression

VII- Autre cas

81

**Présentation:** Étude de deux caractères sur un échantillon de 170 entreprises.

### Série à 1 caractère ( $x_i$ ): Effectif

	Effectif ( $x_i$ )	$n_i$	$n_i x_i$	$n_i x_i^2$
[0-10[	5	95	475	2375
[10-20[	15	75	1125	16875
	<b>N</b>	<b>170</b>	<b>1600</b>	<b>19250</b>



Moyenne x	9.41
Variance x	24.65
Ecart-type x	4.97

### Série à 1 caractère ( $y_j$ ): CA

	CA ( $y_j$ )	$n_j$	$n_j y_j$	$n_j y_j^2$
[0-2[	1	45	45	45
[2-6[	4	55	220	880
[6-14[	10	70	700	7000
	<b>N</b>	<b>170</b>	<b>965</b>	<b>7925</b>



Moyenne y	5.68
Variance y	14.40
Ecart-type y	3.79

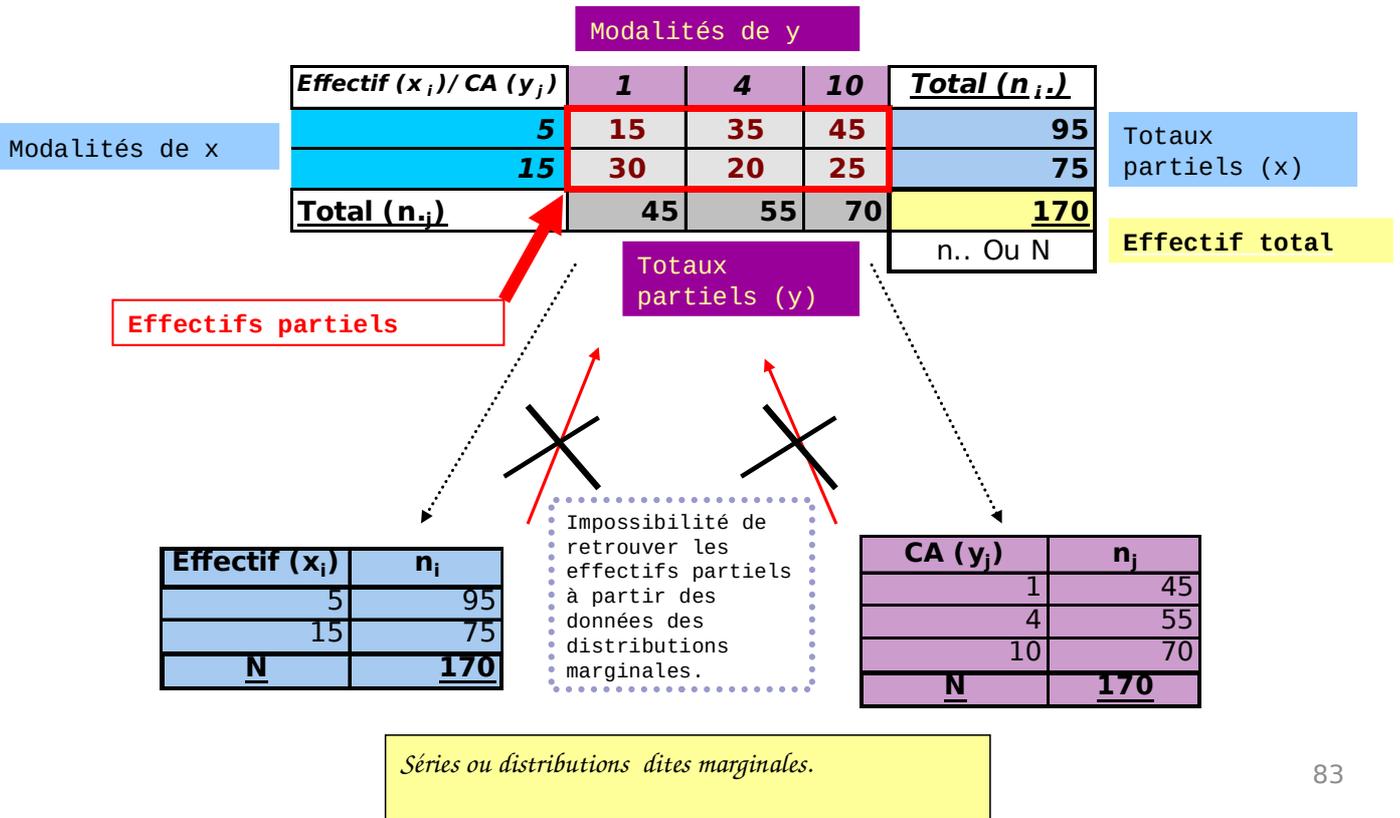
**Bilan:** Séparation des deux caractères étudiés l'un indépendamment de l'autre. Aucune interaction entre ceux-ci.

Questions sans réponses: exemple avec le nombre d'entreprises dont le CA est de 4 et l'effectif de 5

82

# Présentation

Tableau de contingence=tableau croisé=tri croisé=tableau à double entrée. Attention: 3 dimensions.



83

# Notations

Modalités de y

$x_i   y_j$	$y_1$	$y_2$	$y_3$	...	$y_i$	...	$y_q$	$n_{i.}$
$x_1$	$n_{11}$	$n_{12}$			$n_{1j}$			$n_{1.}$
$x_2$	$n_{21}$	$n_{22}$						$n_{2.}$
$x_3$		$n_{31}$			$n_{3j}$		$n_{3q}$	$n_{3.}$
...								
$x_i$	$n_{i1}$	$n_{i3}$			$n_{ij}$		$n_{iq}$	$n_{i.}$
...								
$x_p$	$n_{p1}$				$n_{pj}$		$n_{pq}$	$n_{p.}$
<b><math>n_{.j}</math></b>	<b><math>n_{.1}</math></b>	<b><math>n_{.2}</math></b>	<b><math>n_{.3}</math></b>		<b><math>n_{.j}</math></b>		<b><math>n_{.q}</math></b>	<b><math>n_{..}</math></b>

Modalités de x

Totaux partiels => Effectifs marginaux (x)

Effectifs partiels

Totaux partiels => Effectifs marginaux (y)

Effectif total

i: indice de ligne  
 j: indice de colonne  
 $x_i$ : valeurs du caractère x  
 $y_j$ : valeurs du caractère y  
 p: nombre de lignes  
 q: nombre de colonnes

$n_{.j}$ : Totaux partiels du caractère  $y_j$   
 $n_{i.}$ : totaux partiels du caractère  $x_i$   
 $n_{..}$ : Effectif total (Somme de tous les effectifs partiels)

$$\sum_{i=1}^p n_{ij} = n_{.j}$$

$$\sum_{j=1}^q n_{ij} = n_{i.}$$

$$\sum_{i=1}^p \sum_{j=1}^q n_{ij} = n_{..}$$

84

# Fréquences

15 entreprises  
ont un CA de 1  
et 5 salariés..

Tableau de contingence

Effectif/ CA	1	4	10	Total ( $n_{i.}$ )
5	15	35	45	95
15	30	20	25	75
Total ( $n_{.j}$ )	45	55	70	170

...soit 8,82% de  
(toutes) ces  
entreprises.

Fréquences absolues

Effectif/ CA	1	4	10	Total ( $n_{i.}$ )
5	8.82%	20.59%	26.47%	55.88%
15	17.65%	11.76%	14.71%	44.12%
Total ( $n_{.j}$ )	26.47%	32.35%	41.18%	100.00%

$\frac{n_{ij}}{n_{.}}$

15,71% des  
entreprises de  
5 salariés ont  
un CA de 1.

Fréquences conditionnelles

Effectif/ CA	1	4	10	Total ( $n_{i.}$ )
5	15.79%	36.84%	47.37%	100.00%
15	40.00%	26.67%	33.33%	100.00%
Total ( $n_{.j}$ )	26.47%	32.35%	41.18%	100.00%

$\frac{n_{ij}}{n_{i.}}$

33,33% des  
entreprises  
ayant un CA  
de 1 ont 5  
salariés.

Effectif/ CA	1	4	10	Total ( $n_{i.}$ )
5	33.33%	63.64%	64.29%	55.88%
15	66.67%	36.36%	35.71%	44.12%
Total ( $n_{.j}$ )	100.00%	100.00%	100.00%	100.00%

$\frac{n_{ij}}{n_{.j}}$



85

## Plan

### Chapitre 1: ajustement & corrélation

I - Présentation

II - Méthodes basiques d'ajustement

III - Méthode des moindres carrés

IV - Corrélation

V - Autre exemple

VI - Ajustement, Excel et R

### Chapitre 2: Chroniques

I - Principes

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

### Chapitre 3: Tableaux de Contingence

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

VI- Courbes de régression

VII- Autre cas

86

# Éléments marginaux

Caractère  $x_i$

Effectif/CA	1	4	10	Total (n <sub>i</sub> )
5	15	35	45	95
15	30	20	25	75
<b>Total (n<sub>j</sub>)</b>	45	55	70	170
n <sub>j</sub> y <sub>j</sub>	45	220	700	<u>965</u>
n <sub>j</sub> y <sub>j</sub> <sup>2</sup>	45	880	7000	<u>7925</u>

n <sub>i</sub> x <sub>i</sub>	n <sub>i</sub> x <sub>i</sub> <sup>2</sup>
475	2375
1125	16875
<u>1600</u>	<u>19250</u>

$$\bar{x} = \frac{\sum n_i x_i}{n_{..}}$$

$$\text{var}(x_i) = \frac{\sum n_i x_i^2}{n_{..}} - \bar{x}^2$$

$$\sigma_{x_i} = \sqrt{\text{var}(x_i)}$$

Caractère  $y_j$

$$\bar{y} = \frac{\sum n_j y_j}{n_{..}}$$

$$\text{var}(y_j) = \frac{\sum n_j y_j^2}{n_{..}} - \bar{y}^2$$

$$\sigma_{y_j} = \sqrt{\text{var}(y_j)}$$

Ce qui donne ici:

	Moyenne	Variance	Ecart_type
x <sub>i</sub>	1600/170	9.41	24.65
y <sub>j</sub>	965/170	5.68	3.79

Page 7 du poly.  
87

## Plan

### Chapitre 1: ajustement & corrélation

#### I - Présentation

#### II - Méthodes basiques d'ajustement

#### III - Méthode des moindres carrés

#### IV - Corrélation

#### V - Autre exemple

#### VI - Ajustement, Excel et R

### Chapitre 2: Chroniques

#### I - Principes

#### II- Modèle additif

#### III - Modèle multiplicatif

#### IV - Moyennes mobiles

### Chapitre 3: Tableaux de Contingence

#### I- Présentation

#### II- Éléments marginaux

#### III- Éléments conditionnels

#### IV- Décomposition de la variance

#### V- Ajustement & Corrélation

#### VI- Courbes de régression

#### VII- Autre cas

# Éléments conditionnels: caractère $x_i$

Effectif/CA	1	4	10	Total ( $n_i$ )	$n_i \cdot x_i$	$n_i \cdot x_i^2$
5	15	35	45	95	475	2375
15	30	20	25	75	1125	16875
<b>Total (<math>n_j</math>)</b>	<b>45</b>	<b>55</b>	<b>70</b>	<b>170</b>	<b>1600</b>	<b>19250</b>
$n_j \cdot y_j$	45	220	700	965		
$n_j \cdot y_j^2$	45	880	7000	7925		
$\sum n_{ij} x_i$	525	475	600	1600		
$\sum n_{ij} x_i^2$	7125	5375	6750	19250		

Objectifs: Déterminer l'effectif moyen des entreprises en fonction du niveau de CA. Idem pour les variances correspondantes.

Combien peut-on calculer de moyennes conditionnelles pour ce caractère?

$$\bar{x}_j = \frac{\sum n_{ij} x_i}{n_j}$$

$$\text{Var}(x_j) = \frac{\sum n_{ij} x_i^2}{n_j} - \bar{x}_j^2$$

Moyennes conditionnelles		
11,67	8,64	8,57
Variances conditionnelles		
22,22	21,14	22,96

89

# Éléments conditionnels: caractère $y_j$

Effectif/CA	1	4	10	Total ( $n_i$ )	$n_i \cdot x_i$	$n_i \cdot x_i^2$	$\sum n_{ij} y_j$	$\sum n_{ij} y_j^2$
5	15	35	45	95	475	2375	605	5075
15	30	20	25	75	1125	16875	360	2850
<b>Total (<math>n_j</math>)</b>	<b>45</b>	<b>55</b>	<b>70</b>	<b>170</b>	<b>1600</b>	<b>19250</b>	<b>965</b>	<b>7925</b>
$n_j \cdot y_j$	45	220	700	965				
$n_j \cdot y_j^2$	45	880	7000	7925				
$\sum n_{ij} x_i$	525	475	600	1600				
$\sum n_{ij} x_i^2$	7125	5375	6750	19250				

$$\bar{y}_i = \frac{\sum n_{ij} y_j}{n_i}$$

$$\text{var}(y_i) = \frac{\sum n_{ij} y_j^2}{n_i} - \bar{y}_i^2$$

Objectifs: Déterminer le CA moyen des entreprises en fonction de leur effectif. Idem pour les variances correspondantes. Combien peut-on calculer de moyennes conditionnelles pour ce caractère?

Moyennes conditionnelles	
6,37	4,8
Variances conditionnelles	
12,86	14,96

90

# Plan

## Chapitre 1: ajustement & corrélation

I - Présentation

II - Méthodes basiques d'ajustement

III - Méthode des moindres carrés

IV - Corrélation

V - Autre exemple

VI - Ajustement, Excel et R

## Chapitre 2: Chroniques

I - Principes

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

## Chapitre 3: Tableaux de Contingence

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

VI- Courbes de régression

VII- Autre cas

91

## Décomposition de la variance: Caractère $x_i$

Effectif/CA	1	4	10	Total ( $n_i$ )
5	15	35	45	95
15	30	20	25	75
<b>Total (<math>n_j</math>)</b>	<b>45</b>	<b>55</b>	<b>70</b>	<b>170</b>

Moyennes conditionnelles		
$x_1$	$x_2$	$x_3$
<b>11.67</b>	<b>8.64</b>	<b>8.57</b>
Variances conditionnelles		
$x_1$	$x_2$	$x_3$
<b>22.22</b>	<b>23.14</b>	<b>22.96</b>

	Moyenne marginale	Variance marginale
$x_i$	1600/170	9.41
		1925/170 - 9,41 <sup>2</sup>
		<b>24.65</b>

Var **INTER** = variance des moyennes conditionnelles

$$=(45 \times 11,67^2 + 55 \times 8,64^2 + 70 \times 8,57^2) / 170 - 9,41^2 = \mathbf{1,83}$$

Variance Totale  
(24,65)

$$Var\_Inter(x_i) = \frac{\sum_j n_j \bar{x}_j^2}{n_{..}} - \bar{x}^2$$

Var **INTRA** = moyenne des variances conditionnelles

$$=(45 \times 22,22 + 55 \times 23,14 + 70 \times 22,96) / 170 = \mathbf{22,82}$$

$$Var\_Intra(x_i) = \frac{\sum_j n_j \text{var}(x_j)}{n_{..}}$$

92

## Décomposition de la variance: Caractère $x_i$

□ **Vérification:** Var Totale (**24,65**) = Var Inter (**1,83**) + Var Intra (**22,82**)

↳ **Rapport de Corrélation (R):**  $R = (\text{Var Inter} / \text{Var Totale}) \times 100$

$$R = \frac{1,83}{24,65} \times 100 = 7,4\%$$

□ **Interprétation**

Moyennes conditionnelles		
$x_1$	$x_2$	$x_3$
11.67	8.64	8.57
Variances conditionnelles		
$x_1$	$x_2$	$x_3$
22.22	23.14	22.96

→ Moyennes et variances (de l'effectif) des entreprises classées en 3 groupes de CA.

**Var Inter:** hétérogénéité entre ces groupes

**Var Intra:** homogénéité entre ces groupes

→ **R:** Indicateur de l'hétérogénéité

Effectifs moyens en fonction d'un niveau de CA

→ **Hétérogénéité**

⊗  $\mathcal{R}$ : impact éventuel du CA ( $y_j$ ) sur l'effectif ( $x_i$ )

93

## Décomposition de la variance: Caractère $y_j$

Effectif/CA	1	4	10	Total ( $n_i$ )
5	15	35	45	95
15	30	20	25	75
Total ( $n_j$ )	45	55	70	170

	Moyenne marginale	Variance marginale
$y_j$	$965/170$ <b>5.68</b>	$7925/170 - 5,68^2$ <b>14.40</b>

Moyennes conditionnelles	
$y_1$	$y_2$
6.37	4.80
Variances conditionnelles	
$y_1$	$y_2$
12.86	14.96

Var **INTER** = variance des moyennes conditionnelles

$$= (95 \times 6,37^2 + 75 \times 4,80^2) / 170 - 5,68^2 = 0,61$$

Variance Totale (14,40)

$$\text{Var}_{\text{Inter}}(y_j) = \frac{\sum_i n_i \cdot y_i^2}{n_{..}} - \bar{y}^2$$

Var **INTRA** = moyenne des variances conditionnelles

$$= (95 \times 12,86 + 75 \times 14,96) / 170 = 13,79$$

$$\text{Var}_{\text{Intra}}(y_j) = \frac{\sum_i n_i \cdot \text{var}(y_j)}{n_{..}}$$

94

# Décomposition de la variance: Caractère $y_j$

□ **Vérification:** Var Totale (**14,40**) = Var Inter (**0,61**) + Var Intra (**13,79**)

↳ **Rapport de Corrélation (R):**  $R = (\text{Var Inter} / \text{Var Totale}) \times 100$

$$R = \frac{0,61}{13,79} \times 100 = 4,2\%$$

□ **Interprétation**

Moyennes conditionnelles	
$y_1$	$y_2$
6.37	4.80
Variances conditionnelles	
$y_1$	$y_2$
12.86	14.96



Moyennes et variances (du CA) des entreprises classées en 2 groupes d'effectif.

**Var Inter:** hétérogénéité entre ces groupes

**Var Intra:** homogénéité entre ces groupes



**R:** Indicateur de l'hétérogénéité

CA moyens en fonction d'un niveau de l'effectif



**Hétérogénéité**

⊗  $\mathcal{R}$ : impact éventuel de l'effectif ( $x_i$ ) sur le CA ( $y_j$ )

95

## Plan

### Chapitre 1: ajustement & corrélation

I - Présentation

II - Méthodes basiques d'ajustement

III - Méthode des moindres carrés

IV - Corrélation

V - Autre exemple

VI - Ajustement, Excel et R

### Chapitre 2: Chroniques

I - Principes

II- Modèle additif

III - Modèle multiplicatif

IV - Moyennes mobiles

### Chapitre 3: Tableaux de Contingence

I- Présentation

II- Éléments marginaux

III- Éléments conditionnels

IV- Décomposition de la variance

V- Ajustement & Corrélation

VI- Courbes de régression

VII- Autre cas

96

# Ajustement & Corrélation

## Calculs intermédiaires

Effectif/CA	1	4	10	Total (n <sub>i</sub> )	n <sub>i</sub> ·x <sub>i</sub>	n <sub>i</sub> ·x <sub>i</sub> <sup>2</sup>	Σn <sub>ij</sub> y <sub>j</sub>	Σn <sub>ij</sub> y <sub>j</sub> <sup>2</sup>	Σn <sub>ij</sub> x <sub>i</sub> y <sub>j</sub>
5	15	35	45	95	475	2375	605	5075	3025
15	30	20	25	75	1125	16875	360	2850	5400
<b>Total (n<sub>j</sub>)</b>	<b>45</b>	<b>55</b>	<b>70</b>	<b>170</b>	<b>1600</b>	<b>19250</b>	<b>965</b>	<b>7925</b>	<b>8425</b>
n <sub>j</sub> y <sub>j</sub>	45	220	700	965					
n <sub>j</sub> y <sub>j</sub> <sup>2</sup>	45	880	7000	7925					
Σn <sub>ij</sub> x <sub>i</sub>	525	475	600	1600					
Σn <sub>ij</sub> x <sub>i</sub> <sup>2</sup>	7125	5375	6750	19250					

$$(35 \times 5 \times 4) + (20 \times 15 \times 4)$$

Ou 475 x 4

$$\sum \sum n_{ij} x_i y_j$$

$$(15 \times 5 \times 1) + (35 \times 5 \times 4) + (45 \times 5 \times 10)$$

ou  
605 x 5

## Rappel: Eléments marginaux

	Moyenne	Variance	Ecart_type		
x <sub>i</sub>	1600/170	9.41	1925/170 - 9,41 <sup>2</sup>	24.65	4.97
y <sub>j</sub>	965/170	5.68	7925/170 - 5,68 <sup>2</sup>	14.40	3.79

97

# Ajustement & Corrélation

## 1ère droite d'ajustement

$$y_j = -0,16x_i + 7,15$$

$$a = \frac{\sum \sum n_{ij} x_i y_j - \bar{x} \bar{y} n_{..}}{\sum n_{i.} x_i^2 - \bar{x}^2 n_{..}} \rightarrow a = \frac{\sum \sum n_{ij} x_i y_j - \bar{x} \bar{y} n_{..}}{\text{var}(x)} \rightarrow a = \frac{8425 - 9,41 \times 5,68 \times 170}{24,65} = -0,16$$

$$b = \bar{y} - a \bar{x} \rightarrow b = 5,68 - (-0,17 \times 9,41) = 7,15$$

## 2ème droite d'ajustement

$$x_i = -0,27y_j + 10,94$$

$$a' = \frac{\sum \sum n_{ij} x_i y_j - \bar{x} \bar{y} n_{..}}{\sum n_{.j} y_j^2 - \bar{y}^2 n_{..}} \rightarrow a' = \frac{\sum \sum n_{ij} x_i y_j - \bar{x} \bar{y} n_{..}}{\text{var}(y)} \rightarrow a' = \frac{8425 - 9,41 \times 5,68 \times 170}{14,40} = -0,27$$

$$b' = \bar{x} - a' \bar{y} \rightarrow b' = 9,41 - (-0,27 \times 5,68) = 10,94$$

## Coefficient de corrélation

$$r = \sqrt{a \times a'} = \sqrt{(-0,16) \times (-0,27)} = -0,21$$

$$r = \frac{\sum \sum n_{ij} x_i y_j - \bar{x} \bar{y} n_{..}}{\sqrt{(\sum n_{i.} x_i^2 - \bar{x}^2 n_{..})(\sum n_{.j} y_j^2 - \bar{y}^2 n_{..})}} \rightarrow r = \frac{\sum \sum n_{ij} x_i y_j - \bar{x} \bar{y} n_{..}}{\sqrt{\text{var}(x) \times \text{var}(y)}} \rightarrow r = \frac{8425 - 9,41 \times 5,68 \times 170}{\sqrt{24,65 \times 14,40}} = -0,21$$

98

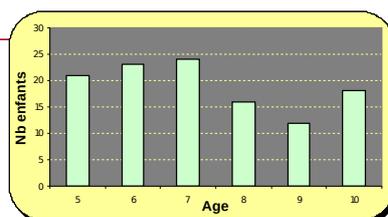
**Chapitre 1: ajustement & corrélation****I - Présentation****II - Méthodes basiques d'ajustement****III - Méthode des moindres carrés****IV - Corrélation****V - Autre exemple****VI - Ajustement, Excel et R****Chapitre 2: Chroniques****I - Principes****II- Modèle additif****III - Modèle multiplicatif****IV - Moyennes mobiles****Chapitre 3: Tableaux de Contingence****I- Présentation****II- Éléments marginaux****III- Éléments conditionnels****IV- Décomposition de la variance****V- Ajustement & Corrélation****VI- Courbes de régression****VII- Autre cas**

99

## Courbes de régression

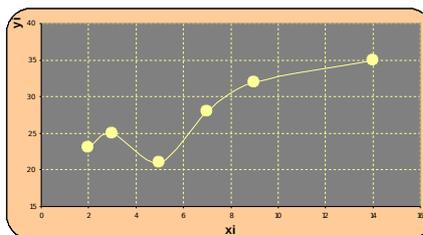
Représentations graphiques: séries à un caractère.

Age	$n_i$
5	21
6	23
7	24
8	16
9	12
10	18



Représentations graphiques: Ajustement.

$x_i$	$y_i$
2	23
3	25
5	21
7	28
9	32
14	35

**Conclusion**

- ❖ Deux dimensions à représenter dans les deux cas
- ❖ => Pas de difficulté majeure
- ❖ Tableaux de contingence: 3 dimensions ( $x_i$ ,  $y_j$ ,  $n_{ij}$ )
- ❖  $x_i$  et  $y_j$  ont (souvent) un nombre de modalités différents (2 et 3 ici)
- ❖ => Représentation plus délicate

# Courbes de régression

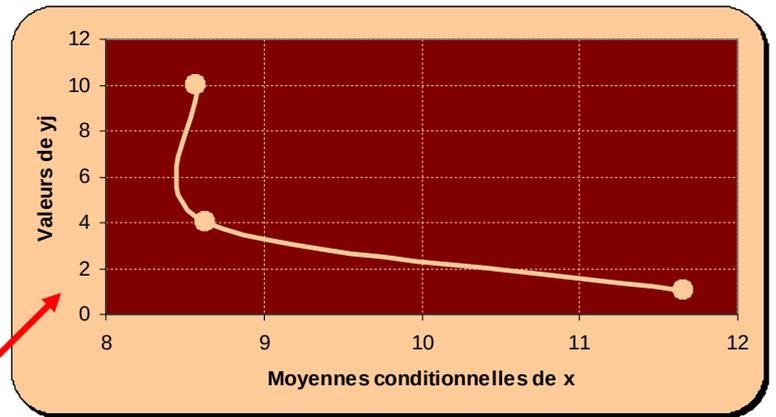
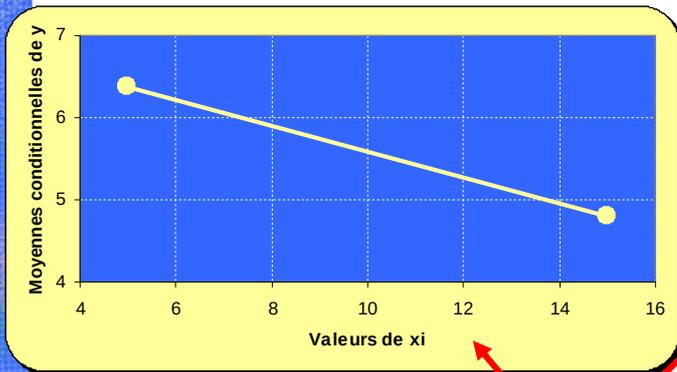
- **Principe:** Représenter sur un plan (2 dimensions donc) une situation qui en compte 3...
- « Projeter » sur un axe deux dimensions (effectifs et un caractère)
- On peut donc associer chaque caractère avec les effectifs...
- Deux courbes de régression donc...

1<sup>ère</sup> courbe: associer  $x_i$  avec  $\bar{y}_i$

$x_i$	$\bar{y}_i$
5	6.37
15	4.8

2<sup>ème</sup> courbe: associer  $y_j$  avec  $\bar{x}_j$

$\bar{x}_j$	$y_j$
11.67	1
8.64	4
8.57	10



Axes « dynamiques »

101

## Rappel

contingence

Entrez le nombre de lignes:  et de colonnes:  Envoyer

Saisir les valeurs des caractères et des effectifs:


Effectifs Observés (Valeurs saisies)

En gras les totaux en ligne et en colonne. Gras et souligné le total général. En bleu les valeurs du caractère  $x_i$ , et en rouge celles du caractère  $y_j$ .

$x_i \backslash y_j$	1	4	10	Total	$n_{ijx_i}$	$n_{ijx_i}^2$	$\text{Somn}_{ijy_j}$	$\text{Somn}_{ijy_j}^2$	$\text{Somn}_{ijy_j x_i}$
5	15	35	45	95	475	2375	605	5075	3025
15	30	20	25	75	1125	16875	360	2850	5400
Total	45	55	70	170	1600	19250	965	7925	8425
$n_{ijy_j}$	45	220	700	965					
$n_{ijy_j}^2$	45	880	7000	7925					
$\text{Somn}_{ijx_i}$	525	475	600	1600					
$\text{Somn}_{ijx_i}^2$	7125	5375	6750	19250					
$\text{Somn}_{ijx_i y_j}$	525	1900	6000	8425					

Eléments relatifs à  $x_i$

102

# Plan

## **Chapitre 1: ajustement & corrélation**

**I - Présentation**

**II - Méthodes basiques d'ajustement**

**III - Méthode des moindres carrés**

**IV - Corrélation**

**V - Autre exemple**

**VI - Ajustement, Excel et R**

## **Chapitre 2: Chroniques**

**I - Principes**

**II- Modèle additif**

**III - Modèle multiplicatif**

**IV - Moyennes mobiles**

## **Chapitre 3: Tableaux de Contingence**

**I- Présentation**

**II- Éléments marginaux**

**III- Éléments conditionnels**

**IV- Décomposition de la variance**

**V- Ajustement & Corrélation**

**VI- Courbes de régression**

**VII- Autre cas**

103

Autre cas: CA ( $x_i$ ) et Investissement ( $Y_j$ )

$x_i \backslash y_j$	1	3	5	10	ni.
5	78	48	2	0	128
10	110	85	3	12	210
15	1	3	12	56	72
20	2	0	11	125	138
50	0	2	1	79	82
n.j	191	138	29	272	630

- ✓ *Liens entre le CA et l'investissement ?*
- ✓ *Impact du CA sur le financement de l'investissement ?*
- ✓ *Impact de l'investissement sur le CA ?*

104

## Autre cas: CA (xi) et Investissement (Yj)

Calculs intermédiaires

$x_i y_j$	1	3	5	10	$n_i$	$n_i \cdot x_i$	$n_i \cdot x_i^2$	$\sum n_{ij} y_j$	$\sum n_{ij} y_j^2$	$\sum n_{ij} x_i y_j$
5	78	48	2	0	128	640	3200	232	560	1160
10	110	85	3	12	210	2100	21000	500	2150	5000
15	1	3	12	56	72	1080	16200	630	5928	9450
20	2	0	11	125	138	2760	55200	1307	12777	26140
50	0	2	1	79	82	4100	205000	801	7943	40050
<b>n.j</b>	<b>191</b>	<b>138</b>	<b>29</b>	<b>272</b>	<b>630</b>	<b>10680</b>	<b>300600</b>	<b>3470</b>	<b>29358</b>	<b>81800</b>
<b>n.jy</b>	191	414	145	2720	<b>3470</b>					
<b>n.jy<sup>2</sup></b>	191	1242	725	27200	<b>29358</b>					
<b><math>\sum n_{ij} x_i</math></b>	1545	1235	490	7410	<b>10680</b>					
<b><math>\sum n_{ij} x_i^2</math></b>	13975	15375	9950	261300	<b>300600</b>					
<b><math>\sum n_{ij} x_i y_j</math></b>	1545	3705	2450	74100	<b>81800</b>					

105

## Autre cas: CA (xi) et Investissement (Yj)

Éléments marginaux

	x	y
<b>Moyennes</b>	16.95	5.51
<b>Variances</b>	189.76	16.26

$\rightarrow = 3470/630$   
 $\rightarrow = 29358/630 - 5,51^2$

Éléments conditionnels

Caractère x	x1	x2	x3	x4
<b>Moyennes</b>	8.09	8.95	16.90	27.24
<b>Variances</b>	7.74	31.32	57.61	218.50

$\rightarrow = 7410/272$   
 $\rightarrow = 261300/272 - 27,34^2$

Caractère y	y1	y2	y3	y4	y5
<b>Moyennes</b>	1.81	2.38	8.75	9.47	9.77
<b>Variances</b>	1.09	4.57	5.77	2.89	1.45

$\rightarrow = 801/82$   
 $\rightarrow = 7943/82 - 9,77^2$

106

## Autre cas: CA (xi) et Investissement (Yj)

### Décomposition des variances

#### Décomposition variance de x

Var INTER (x)	83.56	R	44.04%
Var INTRA (x)	106.19		

**Var totale: 189,76**

#### Décomposition variance de y

Var INTER (y)	13.04	R	80.17%
Var INTRA (y)	3.22		

**Var totale: 19,26**

107

## Autre cas: CA (xi) et Investissement (Yj)

### Corrélation & ajustement

#### Corrélation

r	0.66	0.66
$\rho$	0.43	0.43

#### Ajustement

1ère droite	
a	b
0.19	2.25

2ème droite	
a'	b'
2.24	4.60

108

# Fin cours ...

- ...MSAG en S3
- ...Septembre prochain...
- ...Soyez patients...