

Enquête Terrain Et Prévisions





Introduction

- **L'Enquête**

- Est une recherche d'information.

- **L'Échantillonnage**

- Est une opération consistant à sélectionner une fraction d'une population plus vaste.
 - Sondages ou Tests
 - Représentativité de l'échantillon

- **L'analyse Statistique**



Introduction

- Champ d'exercice des sondages :
 - Enquêtes sur les intentions de vote, baromètres et côtes de popularité, ...
 - Information économique et sociale (démographie, conditions de vie, emploi, consommation, santé, éducation, transports, Tendances marché, prix...).
 - Contrôle de qualité : fabrication et réception de produits industriels.



Eléments d' enquête

Avant tout travail terrain il faut :

- Le champs de l'enquête
- Les unités d'observations **(ou individus)**
- L'unité d'échantillonnage
- La base d'échantillonnage ou la base de sondage
- L'échantillonnage
- Variables d'intérêts



Etapes de l'Enquête par sondage

- **Pertinence** : s'assurer que les informations recherchées n'existent pas déjà.
- **Réflexion générale et théorique sur le sujet** : élaboration d'hypothèses qui seront confirmées ou infirmées par les observations d'enquête.
- **Faisabilité** : matériellement réalisable à un coût raisonnable. S'assurer que l'enquête produira des informations statistiques de bonne qualité.
- **Conception générale de l'enquête :**
 - Définition des objectifs de l'enquête
 - Détermination de la taille de l'échantillon selon budget et précision souhaitée.
 - Définition d'une technique de recueil de l'information
 - Définition simultanée de la méthode d'échantillonnage.
 -

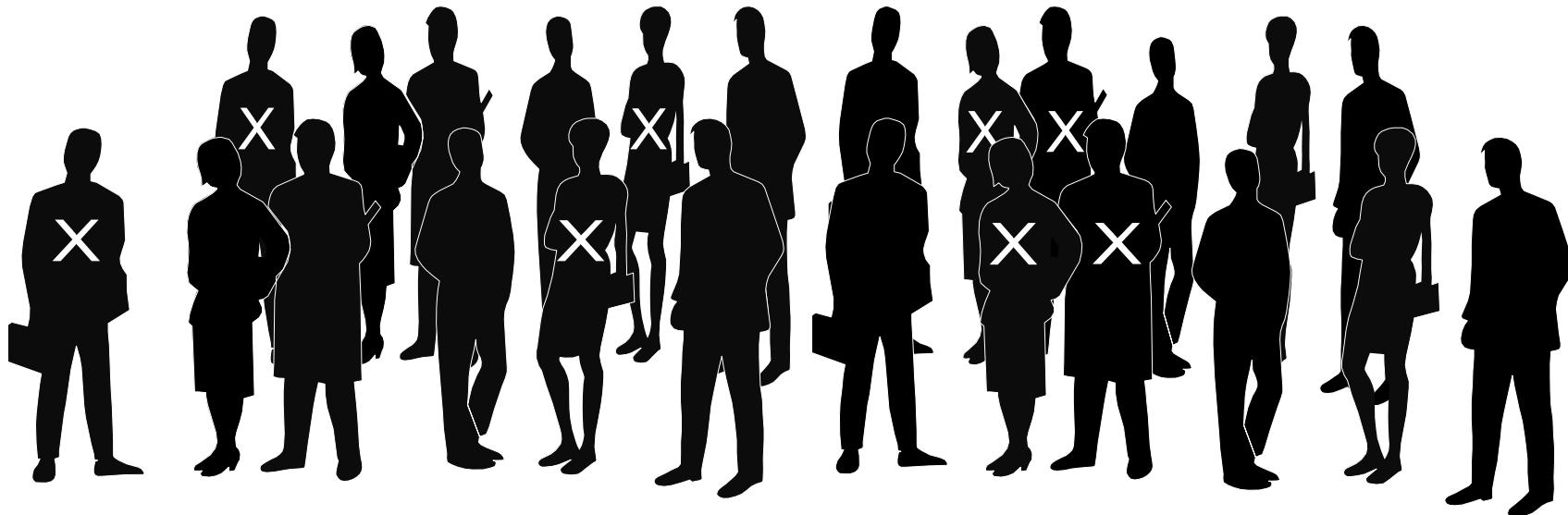
Etapes de l'Enquête par sondage

- **Rédaction du questionnaire :**
 - Première rédaction en utilisant, si possible, les résultats d'études exploratoires ou qualitatives préalables.
 - Pré-test du questionnaire
 - Rédaction définitive incluant le pré-codage
- **Administration du questionnaire :**
 - Réunion d'information des enquêteurs
 - Administration des questionnaires sur le terrain.
 - Contrôle de la qualité du travail des enquêteurs.
- **Traitement et analyse des données :**
 - Vérification de l'exhaustivité et de la vraisemblance des informations.
 - Repérage de valeurs aberrantes (apurement).
 - Codage de certaines variables (age, profession, éducation...).
 - Saisie informatique des questionnaires.
 - Traitement par logiciels spécialisés de traitement de données.

L'échantillonnage : Définition

L'échantillonnage est le processus qui permet d'étudier une petite partie d'une population pour établir un jugement sur la population entière.

L'échantillonnage implique de choisir un certain nombre d'individus au sein d'une population définie (population mère)



Un échantillon représentatif

Un échantillon représentatif a toutes les caractéristiques importantes de la population mère d'où l'extrapolation

Un échantillon est dit représentatif lorsqu'il peut traduire le plus fidèlement possible les caractéristiques importantes de la population mère

Taille de l'échantillon

Elle est souvent un compromis entre ce qui est statistiquement souhaitable et ce qui est pratiquement faisable.

Eu égard à la loi normale, la taille minimum d'un échantillon est en général de 30 individus.

Un échantillon dont la taille est inférieure à 30 ne peut pas être valorisé au plan statistique, les résultats restent valables à titre indicatif. L'extrapolation ou la généralisation deviennent déconseillées

Taille optimale d'un échantillon représentatif d'une population étudiée

$$n = \frac{(1.96)^2 \times N}{(1.96)^2 + l^2 \times (N-1)}$$

n = taille de l'échantillon à interroger

N = taille de l'univers investigué

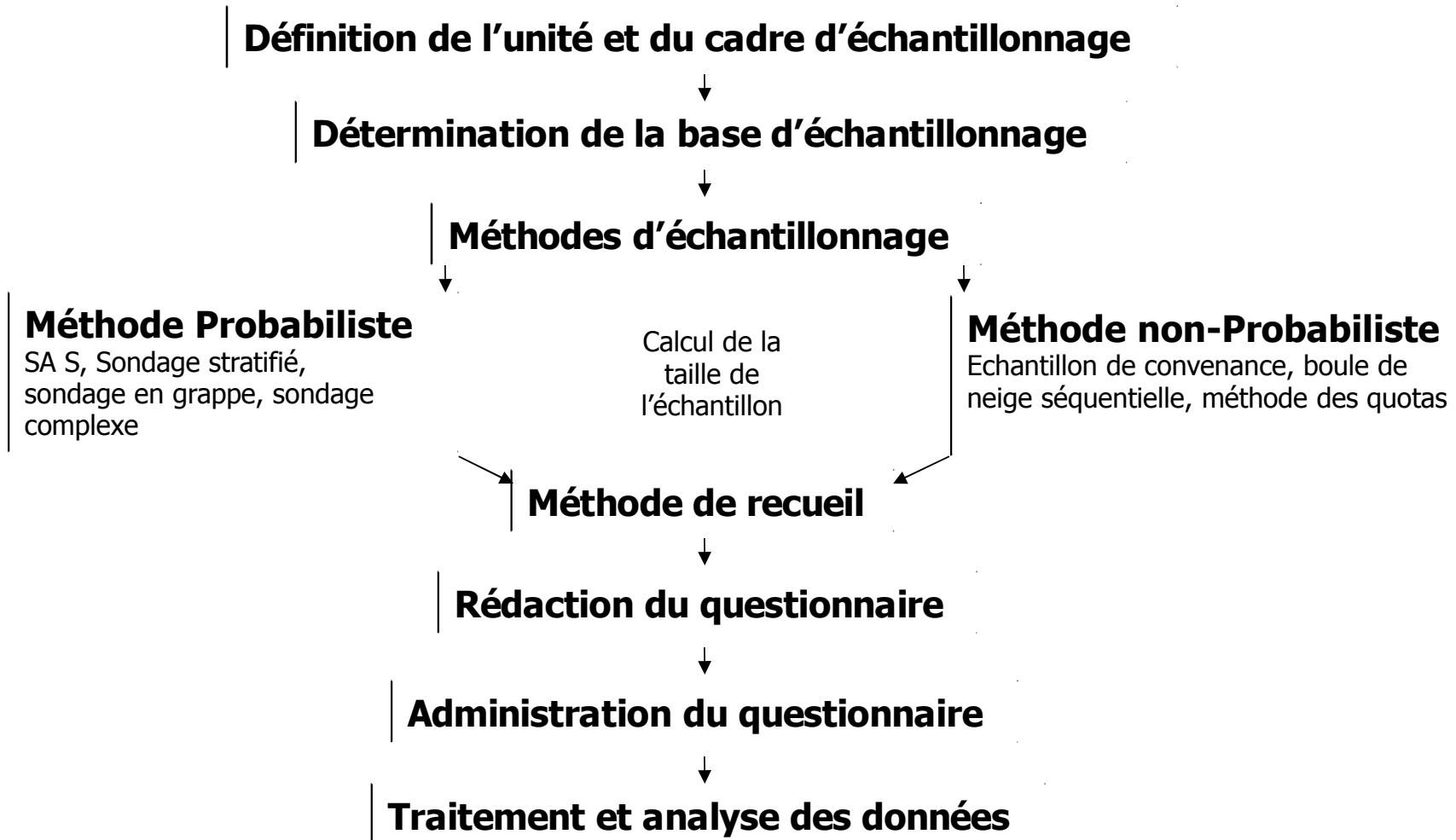
l = largeur de la fourchette exprimant la marge d'erreur

Exemples :

Pour une population mère (univers) de 5 000 individus et en cas d'erreur de 3% au-dessus et en dessous de la valeur (soit une fourchette de 6%), nous devons étudier un échantillon de 878 personnes.

Univers	Largeur de la fourchette				
	4%	<u>6%</u>	8%	10%	12%
10.000.000	2.400	1.067	600	384	267
100.000	2.345	1.056	597	383	266
50.000	2.291	1.045	593	381	265
<u>5.000</u>	1.622	<u>878</u>	536	357	253
500	414	341	273	217	174

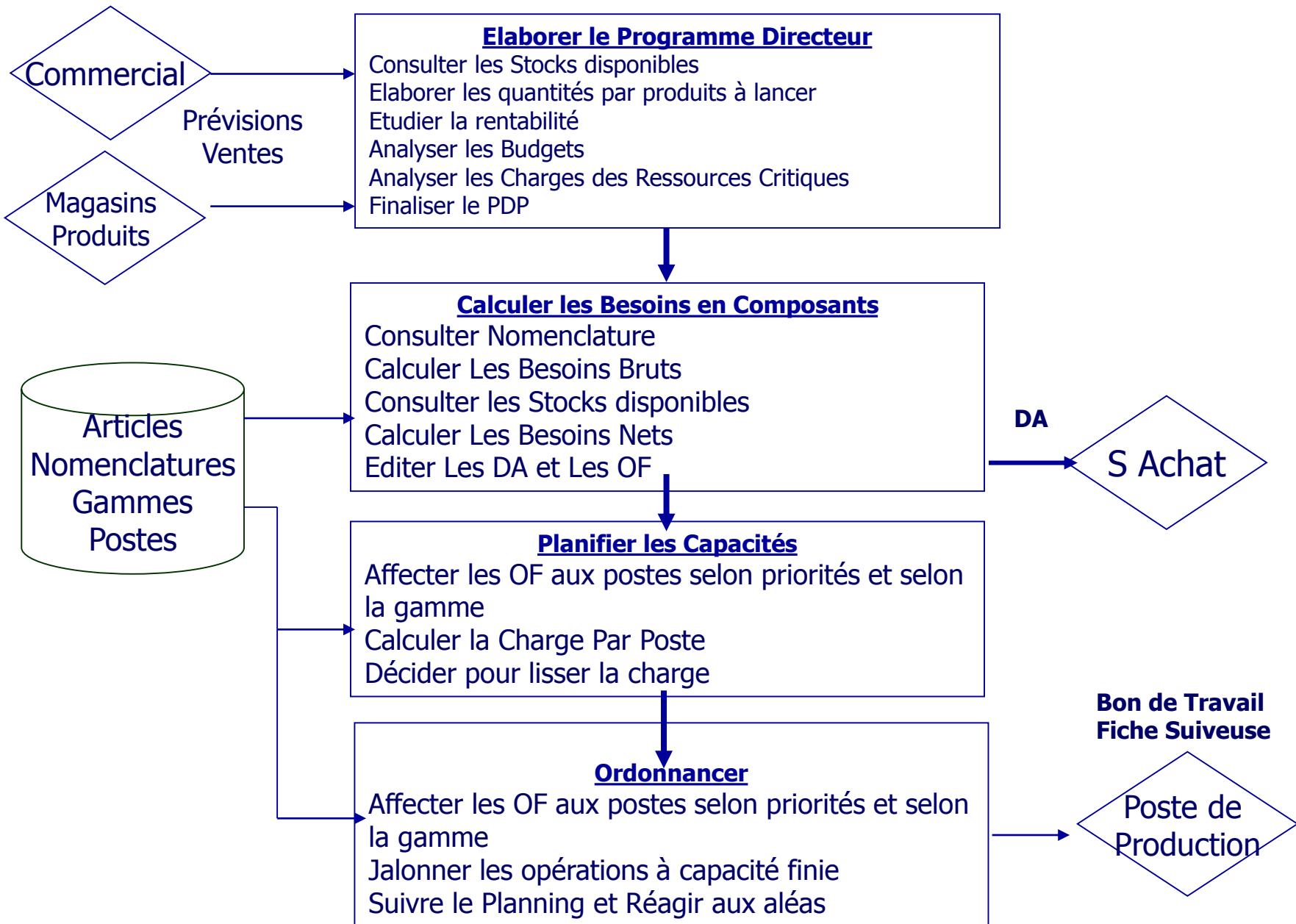
L'Échantillonnage : succession d'étapes



Les prévisions de la demande

PREVISIONS

LA PROCEDURE DE LA GESTION DE PRODUCTION

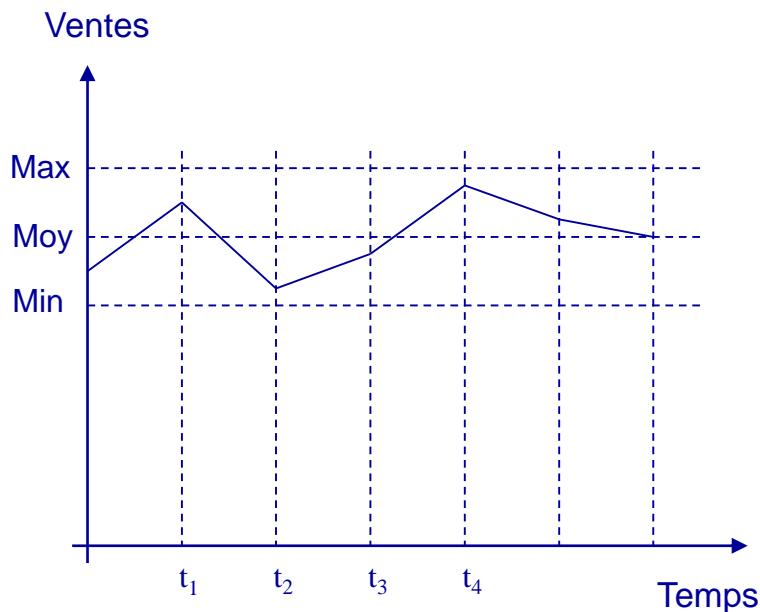


Prévisions

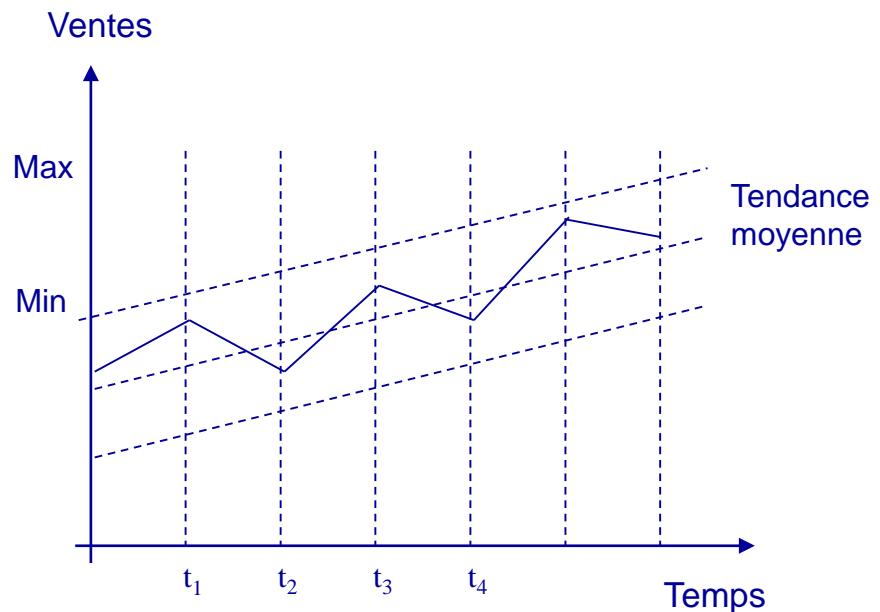
- Éléments essentiel pour le bon fonctionnement de l'entreprise :
 - Conditionne la prise de décision à long, moyen et cours terme.
 - Long terme: prévisions pouvant engager des décisions stratégiques sur les investissements et la diversification des produits.
 - Moyen terme: permettent d'ajuster les capacités globales de production et d'approvisionnement.
 - Court terme: planification opérationnelle de la production, ordonnancement.
 - .
- Deux types de méthode de prévision: qualitatives et quantitatives.
 - Qualitatives: utilisées pour les prévisions moyen et long terme. Décisions commerciales basées sur des études de marché, sur des intentions d'achat en cas de lancement d'un nouveau produit.
 - Quantitatives: fondées sur des modèles mathématiques provenant de données historiques sur le produit.

Les lois du marché

- Consommation constante

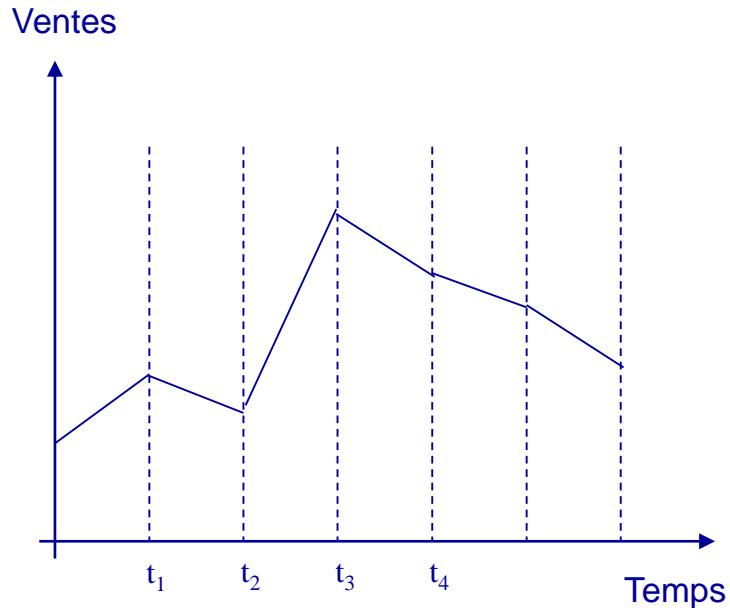


- Consommation à tendance

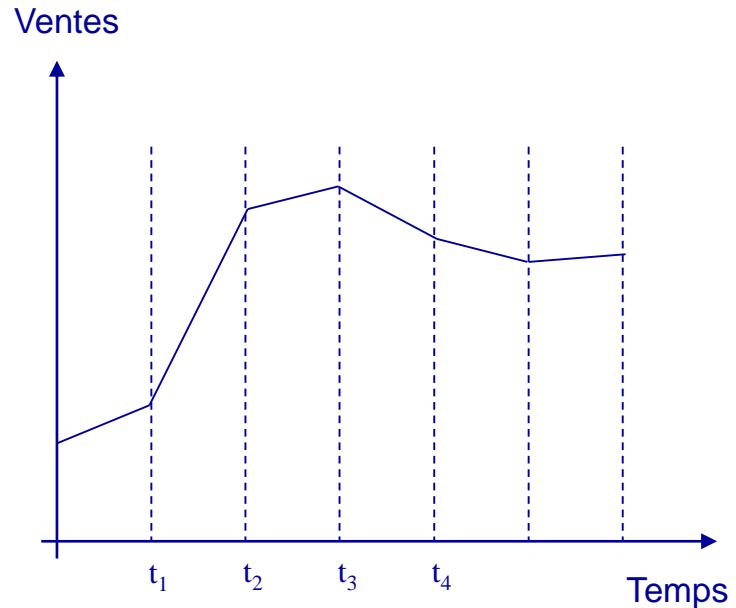


Les lois du marché

Consommation saisonnière



Consommation saisonnière à tendance



Connaissance des consommations passées

- ▶ Il faut analyser la structure de la demande
- ▶ Exemple:

Mois	1	2	3	4	5	6	7	8	9	10	11	12
Ventes	80	90	60	90	100	140	100	100	110	100	110	120

Moyenne arithmétique

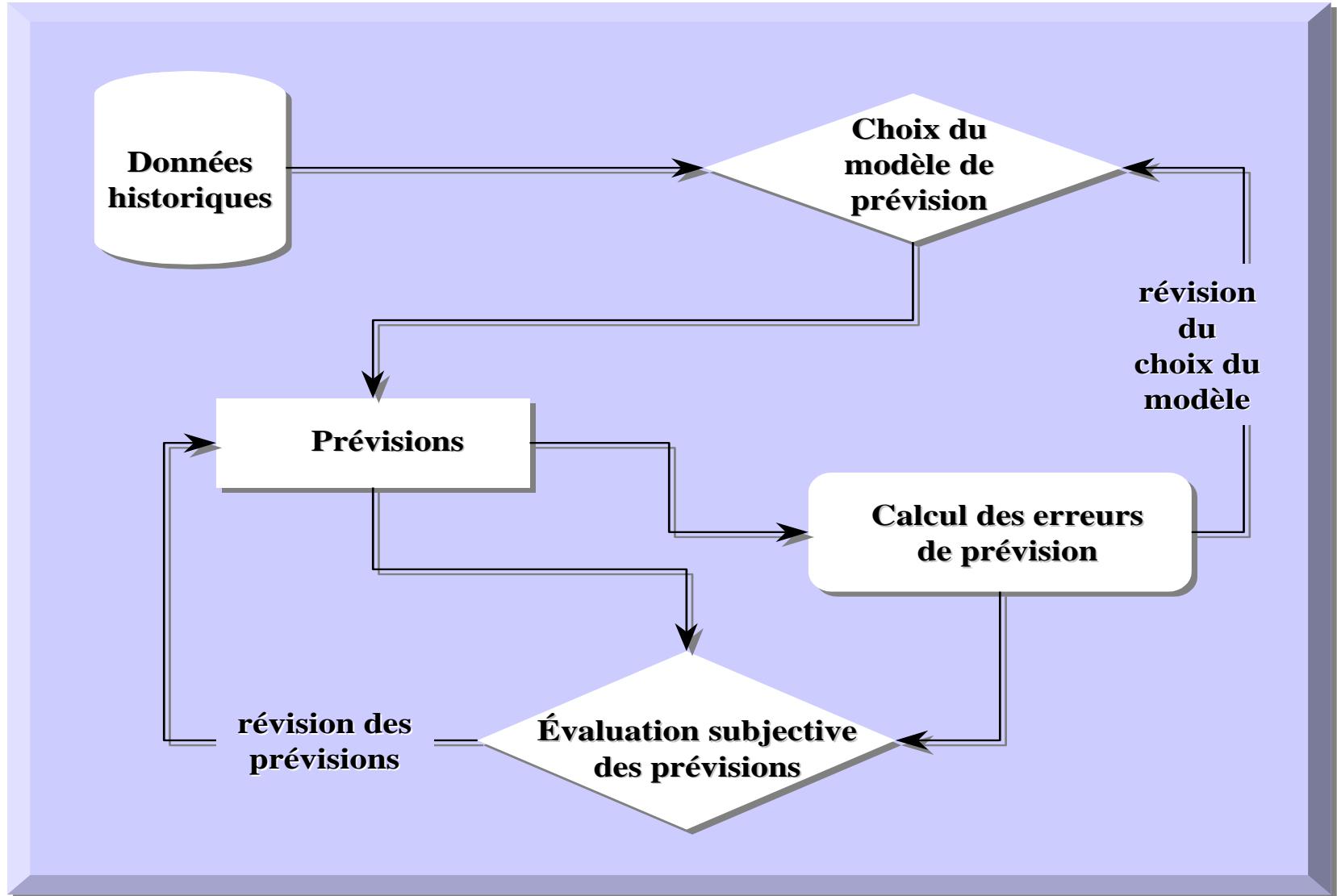
$\bar{V} = 100$ Utile si on connaît la dispersion des valeurs

Écart type:

$$\sigma = \sqrt{\text{variance}} = \sqrt{\frac{\sum_{i=1}^n (V_i - \bar{V})^2}{n - 1}}$$

$$\sigma = 20$$

Le processus de prévision ...



1-Méthodes de prévision

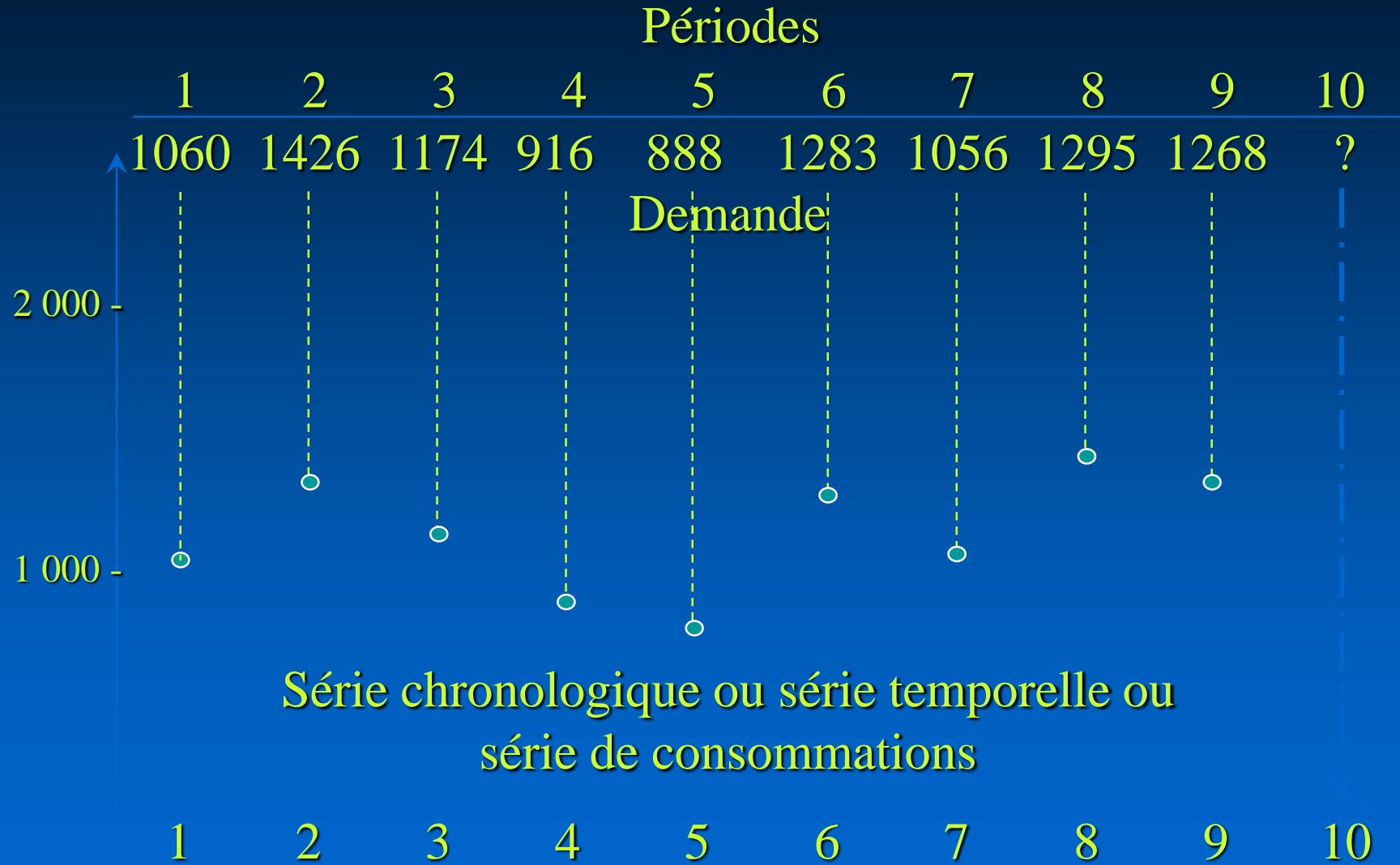
- Méthodes explicatives (L'approche causale en prévision)
 - Comparaison d'une série d'observations de la variable X à laquelle on s'intéresse à d'autres séries d'observations Y, Z, W
 - Calcul des valeurs futures de X à partir des valeurs estimées de Y, Z, W
- Méthodes par extrapolation (Prévisions dans les séries chronologiques)
 - Etude de l'historique d'une série d'observations de la variable à laquelle on s'intéresse
 - Extrapolation du passé sur le futur
- Méthode par expertise
 - Demander l'avis d'experts sur la variable à laquelle on s'intéresse
 - Synthétiser les avis

2- Les séries chronologiques

Les méthodes statistiques de prévision se basent sur l'analyse de données historiques appelées les séries chronologiques.

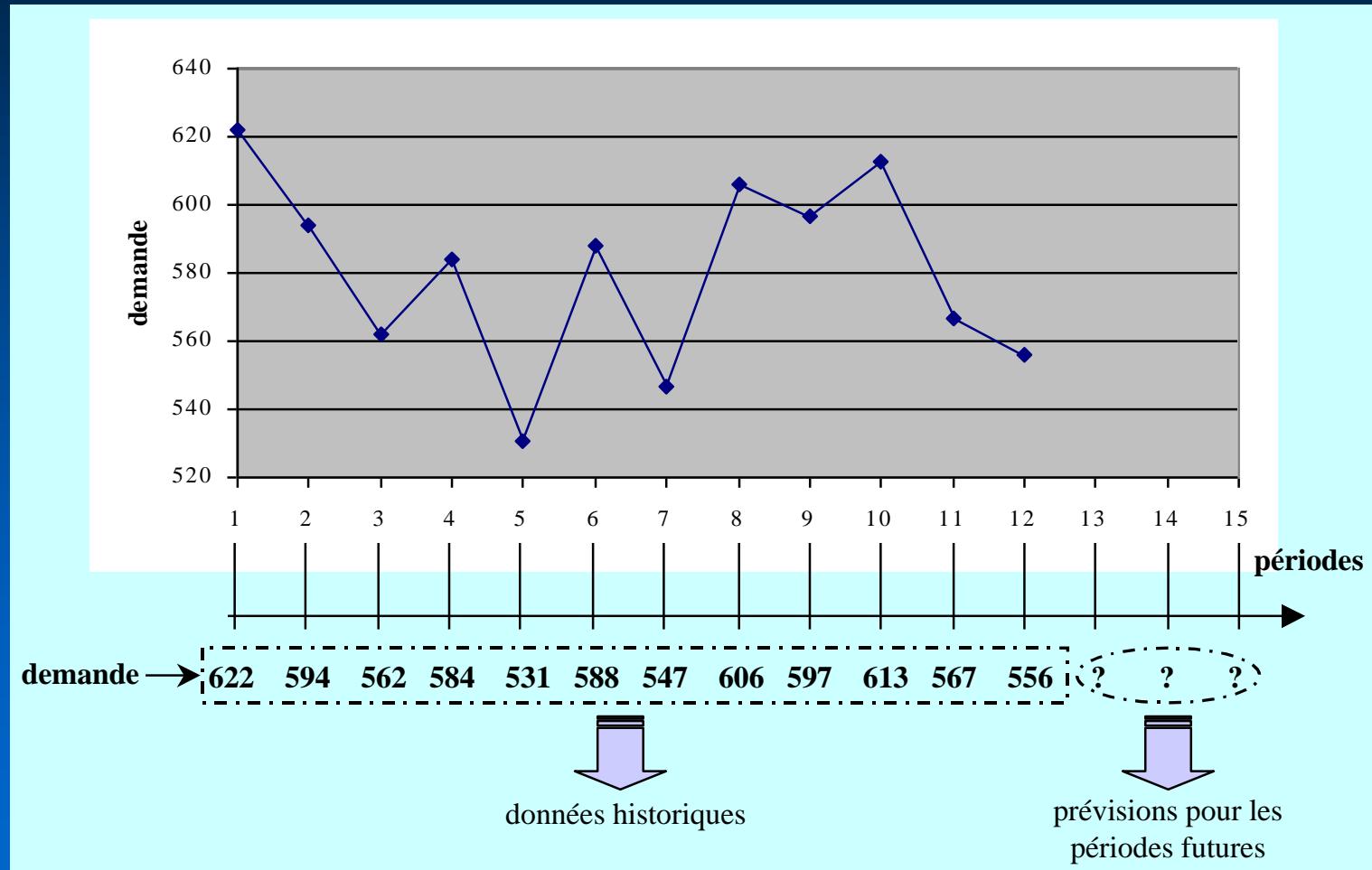
Une série chronologique est un ensemble d'observations faites à différentes périodes successives dans le temps.

Qu'est-ce qu'une série chronologique ...

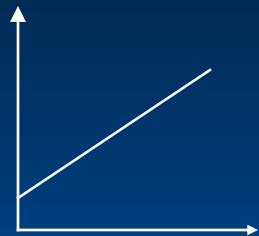


Série chronologique ...

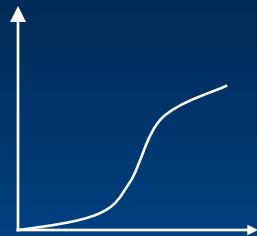
$X_t, t = 1, \dots, T$



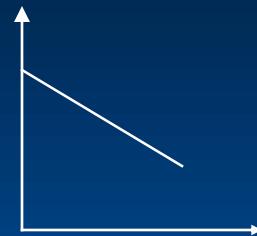
Voici quelques comportements de séries chronologiques



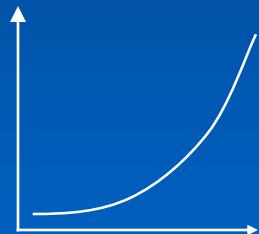
linéaire



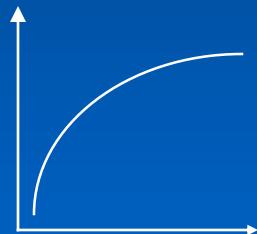
courbe en S



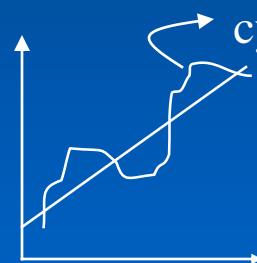
linéaire décroissante



exponentielle



asymptotique



cycle avec tendance

Étapes de la méthode des séries chronologiques

1. Collecte des données

- ➡ cueillette d'observations sur les valeurs de la variable de prévision sur plusieurs périodes.
- ➡ mettre de côté les données non représentatives (ex.: lors de grève.)

Étapes de la méthode des séries chronologiques

2. Analyse des données

→ définir le modèle sous-jacent représentant le mieux l'évolution de la demande passée par:

i) technique de l'observation visuelle

Étapes de la méthode des séries chronologiques

3. Choix de la meilleure méthode de prévision

- ⇒ test systématique des méthodes se rapportant au modèle sous-jacent choisi.
- ⇒ étude des résultats obtenus à l'aide d'une technique qu'on verra plus tard (mesures d'erreurs).

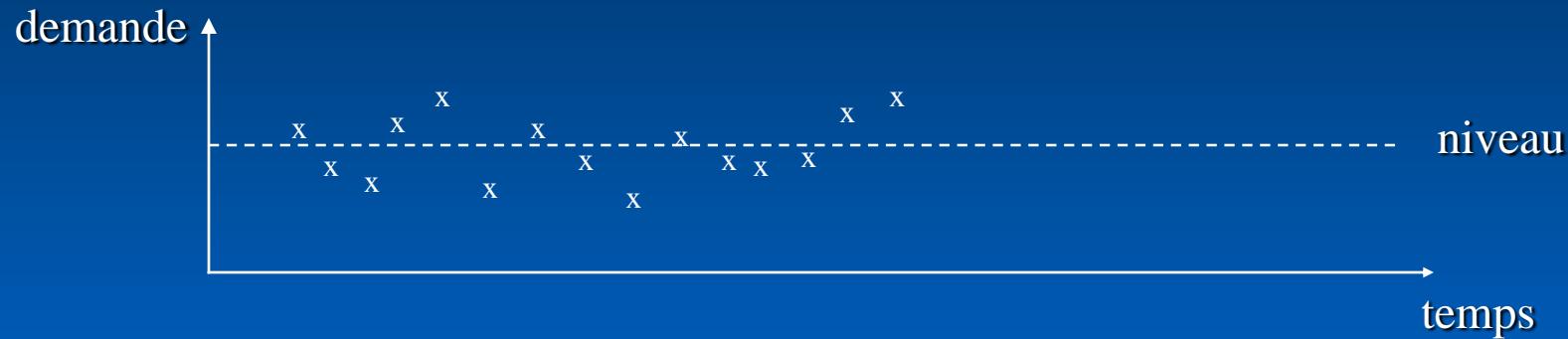
Étapes de la méthode des séries chronologiques

4. Obtention des prévisions

- ⇒ utiliser les équations déterminées précédemment pour déterminer les prévisions.

Modèle avec niveau constant

Une étude est faite de l'évolution aléatoire des données observées autour d'une valeur centrale stable dite niveau.



Le niveau correspond en fait à une moyenne

Différentes façons d'établir ce niveau

- Moyenne statistique
- Moyenne mobile
- Moyenne pondérée
- Lissage exponentiel

Moyenne statistique

La valeur du niveau est égale à la moyenne des observations retenues.

$$\bar{X}_t = \frac{\sum_{i=1}^n X_i}{n}$$

où \bar{X}_t = moyenne statistique au temps t

X_i = valeur observée au temps i

n = #observations

Moyenne statistique (suite)

La prévision P de la demande pour les périodes futures j est donnée par

$$P_{t+j} = \bar{X}_t$$

La prévision est donc la même pour toutes les périodes à venir.

Moyenne mobile

Pour cette moyenne, seules les observations les plus récentes sont utilisées pour calculer la prévision.

Cette méthode nécessite de conserver un grand nombre de données en mémoire.

Moyenne mobile (suite)

Les prévisions se calculent de la façon suivante

$$P_{t+1} = \frac{\sum_{i=t-m+1}^t X_i}{m}$$

où m = #observations considérées (ordre de la moyenne mobile)
t = la dernière période pour laquelle nous considérons
une observation

Exemple 2

Nous avons les données suivantes concernant les ventes en 1000 gallons d'essence par semaine.

Semaine	Ventes (1000 gallons)	Semaine	Ventes (1000 gallons)
1	17	7	20
2	21	8	18
3	19	9	22
4	23	10	20
5	18	11	15
6	16	12	22

Considérons une moyenne mobile basée sur 3 observations.
Quelles sont les prévisions des semaines 4 et 5?
Quelle serait la prévision pour la semaine 7?

Lissage exponentiel simple

Cette méthode permet de calculer une moyenne pondérée qui tient compte du poids attaché aux observations, le poids s'estompant quand on avance dans le temps.

Cette méthode est une des plus utilisées.

Lissage exponentiel simple

Soient

P_t = prévision au temps t.

X_t = observation au temps t.

α = facteur de pondération compris entre 0 et 1
(appelé aussi constante de lissage)

Lissage exponentiel simple

La prévision au temps P_t se calcule ainsi:

$$P_t = \alpha X_{t-1} + \alpha(1-\alpha) X_{t-2} + \alpha(1-\alpha)^2 X_{t-3} + \dots + \alpha(1-\alpha)^{n-1} X_{t-n}$$

Cette formule se réécrit sous la forme

$$P_t = \alpha X_{t-1} + (1-\alpha) P_{t-1} = P_{t-1} + \alpha (X_{t-1} - P_{t-1})$$

Lissage exponentiel simple

Trois types de données sont nécessaires pour appliquer la méthode:

- 1) La prévision pour la période précédente.
- 2) La demande réelle pour cette même période.
- 3) Facteur de pondération α

Exemple 3

Une firme utilise un lissage exponentiel simple avec un coefficient α de 0,1 pour prévoir une demande. La prévision pour la première semaine de février était de 500 unités alors que la demande réelle était de 450.

Prévoyez la demande pour la semaine du 8 février.

Exemple 4

Considérons les valeurs observées suivantes pour les 12 prochaines périodes.

t	X _t
1	90
2	105
3	95
4	110
5	95
6	95
7	105
8	120
9	120
10	115
11	125
12	115

Soit $\alpha=0,1$ et $0,3$.

Quelles sont les prévisions pour les périodes 1 à 13?

Exemple 4 (solution)

	Xt	Pt ($\alpha = 0,1$)	Pt ($\alpha = 0,3$)
1	90	90.00	90.00
2	105	90.00	90.00
3	95	91.50	94.50
4	110	91.85	94.65
5	95	93.67	99.26
6	95	93.80	97.98
7	105	93.92	97.08
8	120	95.03	99.46
9	120	97.52	105.62
10	115	99.77	109.94
11	125	101.29	111.45
12	115	103.67	115.52
		104.80	115.36

Exemple

Les données trimestrielles des pannes de certains moteur d'avion dans une base militaire local durant les deux dernières années sont les suivantes : 200 ; 250 ; 175 ; 186 ; 225 ; 285 ; 305 ; 190.

Utiliser la méthode des moyennes mobiles sur 3 et sur 6 mois pour prévoir les pannes des moteurs.

Utiliser la méthode du lissage exponentiel pour $\alpha=0.1$ et $\alpha=0.4$ pour prévoir les

Lissage exponentiel simple

Le **facteur de pondération**, α , détermine le niveau de lissage et la vitesse de réaction à la différence entre la prévision et la demande réelle.

Le choix de α dépend de l'allure de la demande.

Initialisation

Raisons pour expliquer le succès des méthodes de lissage exponentiel

- 1) Le modèles sont assez précis.
- 2) La formulation des modèles se fait aisément.
- 3) L'utilisateur peut comprendre comment le modèle fonctionne.
- 4) Le modèle requiert peu de calculs.

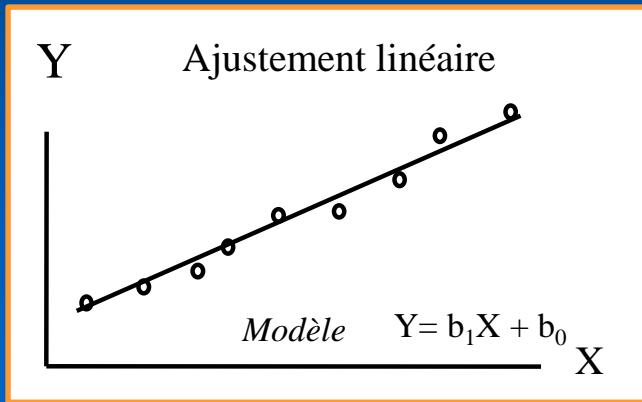
Modèle avec tendance

Pour ce type de modèle, on ne considère plus une moyenne stable mais plutôt la tendance de la demande en fonction du temps.

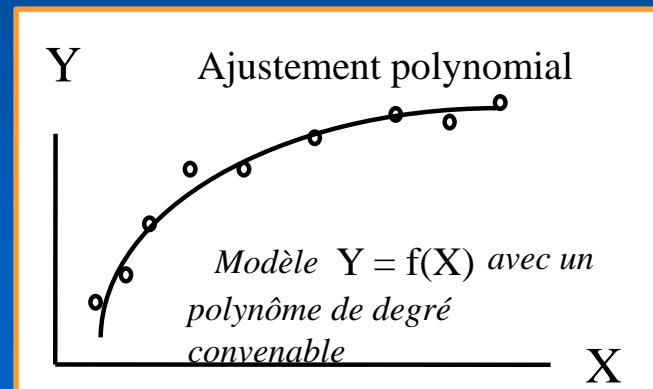
En fait, la moyenne ne peut nous être utile dans ce cas.

1.1 Régression simple

Une régression est dite simple si elle permet de prédire les valeurs d'une variable dite dépendante (expliquée (Y)) à partir des valeurs prises par une autre variable dite indépendante (explicative (X)).
Le modèle mathématique postulé peut être :



Une droite si Y varie linéairement avec X.



Sinon un polynôme de degré convenable.

Régression simple

$$a = \bar{Y} - b\bar{X}$$

$$b = \frac{\left(\sum XY \right) - n\bar{X}\bar{Y}}{\left(\sum X^2 \right) - n\bar{X}^2}$$

où \bar{X}, \bar{Y} sont les moyennes respectives des n observations des X et des Y

Régression simple

- a et b sont obtenus par les équations normales de la méthode des moindres carrés.
- Cette méthode tente de trouver la droite représentant le mieux les données en minimisant la somme des carrés de la distance verticale entre chaque point et son point correspondant sur la droite.

Régression simple exercice

Une usine de fabrication de composant électronique a établi son Bilan de vente annuel tel donné sur le tableau
Elle veux faire ses prévisions pour le premier mois de l'année qui suit
Pouvez vous l'aider à le faire?



X	Y
1	600
2	1550
3	1500
4	1500
5	2400
6	3100
7	2600
8	2900
9	3800
10	4500
11	4000
12	4900

Exemple 1

Le tableau suivant présente les données et les calculs nécessaires pour faire des prévisions.

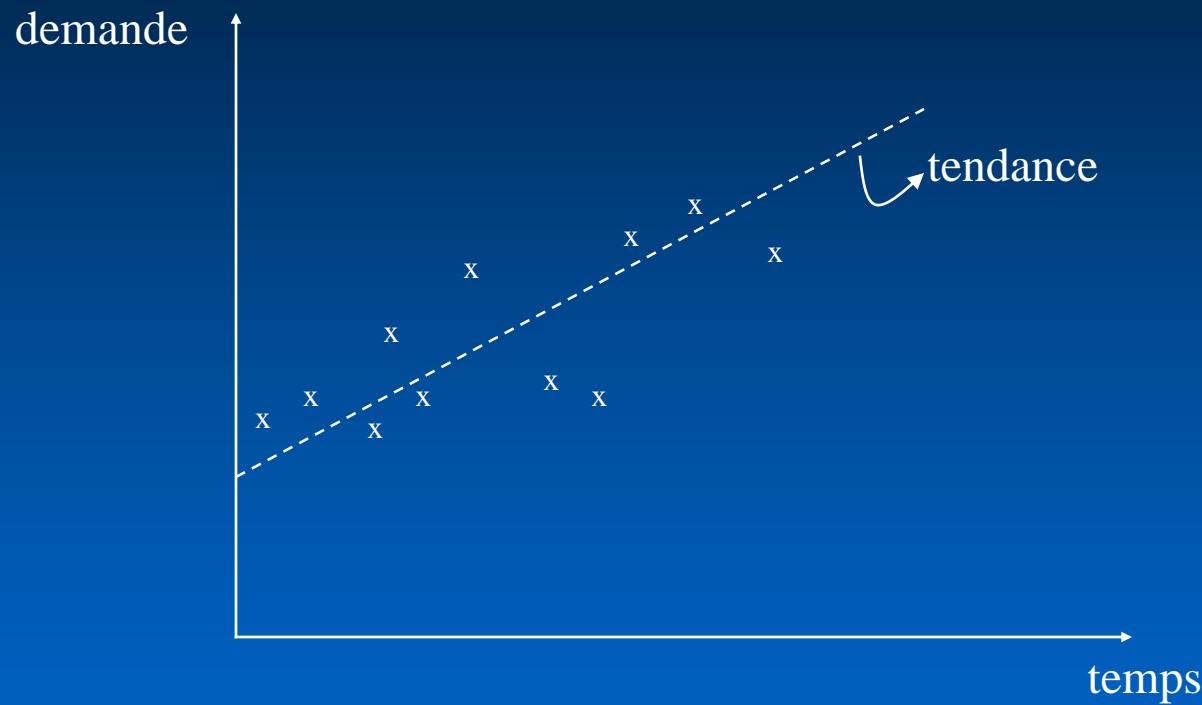
X	Y	XY	X ²	Y ²	Prévision
1	600	600	1	360000	803.1
2	1550	3100	4	2402500	1160.9
3	1500	4500	9	2250000	1520.5
4	1500	6000	16	2250000	1880.1
5	2400	12000	25	5760000	2239.7
6	3100	18600	36	9610000	2599.4
7	2600	18200	49	6760000	2959.0
8	2900	23200	64	8410000	3318.6
9	3800	34200	81	14440000	3678.2
10	4500	45000	100	20250000	4037.8
11	4000	44000	121	16000000	4397.4
12	4900	58800	144	24010000	4757.1
78	33350	268200	650	112502500	

Mesure de la tendance

La mesure de la tendance =

Amplitude de la variation moyenne
observée d'une période à l'autre.

Mesure de la tendance



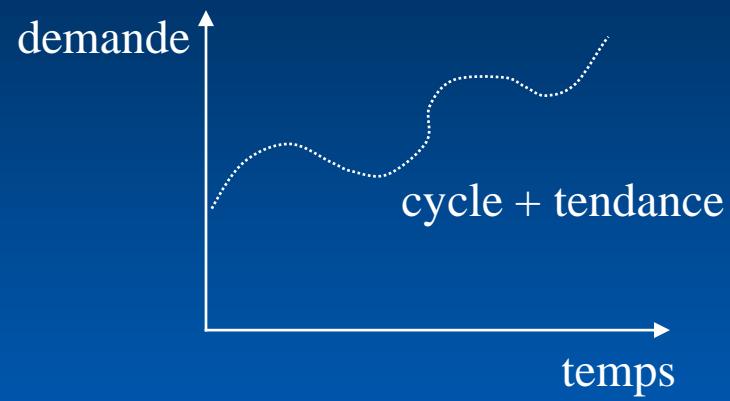
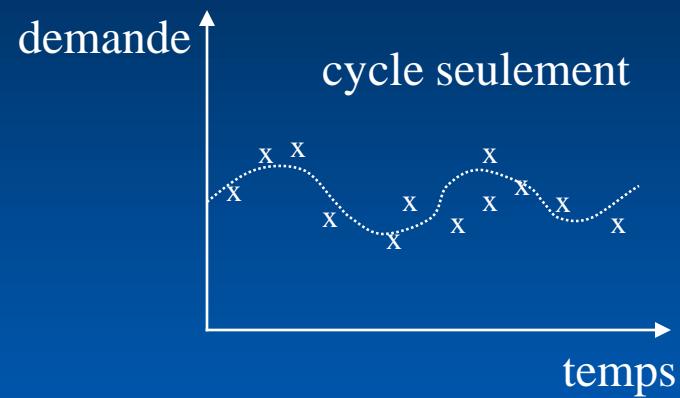
Modèle avec cycle

Ici, la demande ne varie pas de façon constante.
On ne peut donc plus parler de tendance linéaire
comme le modèle précédent.

La demande varie de façon **cyclique** ou saisonnière.

Le cycle est décelable par visionnement des données ou par l'analyse de l'autocorrélation.

Modèle avec cycle



3.3 Exemple: Le tableau suivant représente les ventes trimestrielles d'un distributeur de gaz.

Trim. Année	1980	1981	1982	1983	1984	1985	1986
1	91707	96794	116250	106878	122915	150682	129762
2	63048	74949	71998	71800	92079	96967	82597
3	57041	56791	59620	65880	80241	85492	74167
4	78667	89127	98985	94254	118075	126312	103340

La première étude à faire est descriptive.

Des graphiques permettent de visualiser une **tendance** et des **variations saisonnières**

Choix de la meilleure technique de prévision

Principales mesures d'erreurs

La meilleure méthode de prévision doit donner les prévisions les plus précises possibles.

Pour évaluer une méthode, on se base sur les erreurs de prévision passées.

Principales mesures d'erreurs

L'écart entre une donnée passée et la prévision faite par le modèle pour la période correspondante est mesuré pour s'assurer de la justesse du modèle.

Les différentes mesures que nous allons voir peuvent être évaluées pour différentes méthodes. La méthode donnant les meilleurs résultats serait la plus appropriée.

Différentes mesures d'écart

- Écart quadratique moyen
- Écart absolu moyen
- Écart absolu moyen en %
- Biais

Écart quadratique moyen

L'écart quadratique moyen se calcule ainsi

$$EQM_n = \frac{\sum_{i=1}^n (P_i - X_i)^2}{n}$$

où P_i = valeur prévue et X_i = valeur réelle.

Écart absolu moyen

Ce type de mesure d'erreur tient compte des écarts sans égard au signe des valeurs.

L 'écart absolu moyen se calcule ainsi

$$EAM_t = \frac{\sum_{i=1}^n |P_i - X_i|}{n}$$

Erreur absolue moyenne

Par ailleurs, l'**erreur absolue moyenne en %** se calcule un peu différemment que l'écart absolu moyen.

$$MAPE = \frac{100 \sum_{i=1}^n \frac{|P_i - X_i|}{X_i}}{n}$$

Biais ou erreur moyenne

Pour le calcul du biais, les écarts tiennent compte du signe des valeurs i.e. négatifs ou positifs.

Le biais devrait être près de 0.

Biais ou erreur moyenne

Si le biais est $> 0 \Rightarrow$ les prévisions ont tendance à dépasser les valeurs réelles.

Si le biais est $< 0 \Rightarrow$ les prévisions ont tendance à être sous les valeurs réelles.

$$\text{Si } e_t = P_t - X_t$$

Biais ou erreur moyenne

Le biais se calcule ainsi

$$Biais_n = \frac{\sum_{i=1}^n (P_i - X_i)}{n}$$

Exemple 5

Considérons les données sur les ventes du modèle d'automobile Z pour les années 1974 à 1980. Les prévisions obtenues à l'aide de la méthode du lissage simple et à l'aide d'une régression linéaire sont également données.

	X	LE	Reg
1	180	180	192
2	205	180	194
3	185	200	195
4	200	188	197
5	220	198	199
6	210	216	200
7	180	211	202

Exemple 5 (suite)

Calculez les différentes mesures d'erreur.

Quelle méthode de prévision est la meilleure?