

# Informatique Décisionnelle



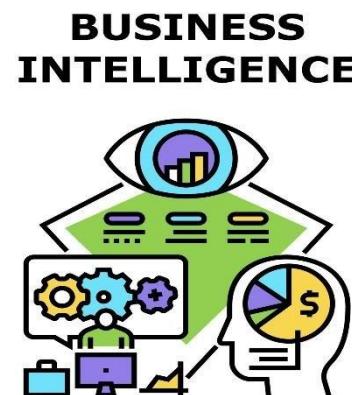
4 ERP-BI

Equipe BI

Dernière modification :2024-  
2025

# Partie 1:

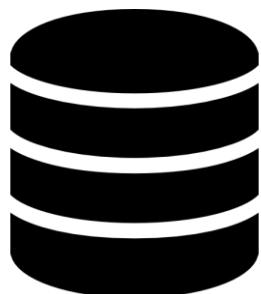
## Introduction Générale



## Mise en situation (1/4)

**Une entreprise de fabrication et distribution de parapluie a des informations clients sont stockées comme suit:**

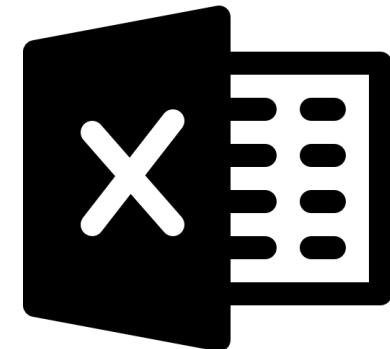
Les commandes sont stockées dans une base de données relationnelle



Les factures sont gérées dans un système ERP.



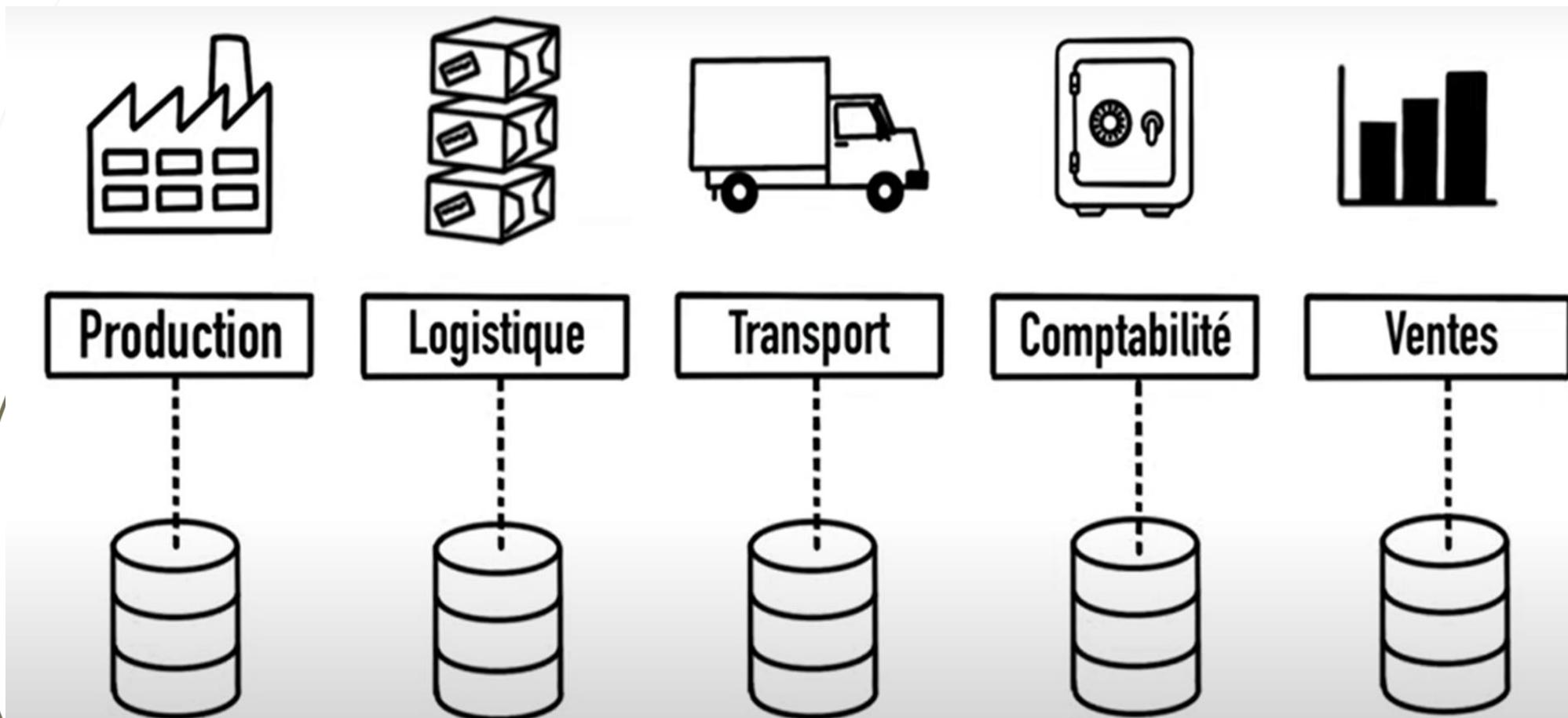
Les interactions avec le service client sont enregistrées dans un fichier Excel



1

## Mise en situation (2/4)

Une entreprise de fabrication et distribution de parapluie a des informations stockées comme suit:



1

## Mise en situation (2/4)



1

## Mise en situation (3/4)

**Une société immobilière utilise les formats suivants pour gérer ses biens :**



Les descriptions des biens sont conservées dans des documents Word.



Les feedbacks des clients sont collectés en temps réel à partir de leur page facebook.



Les informations financières (loyers, charges) sont dans des classeurs.

1

# Mise en situation (4/4)

**Vous êtes intéressé par l'achat d'actions. Quelle entreprise choisiriez-vous avec une somme X ?**

**IBOURSA**

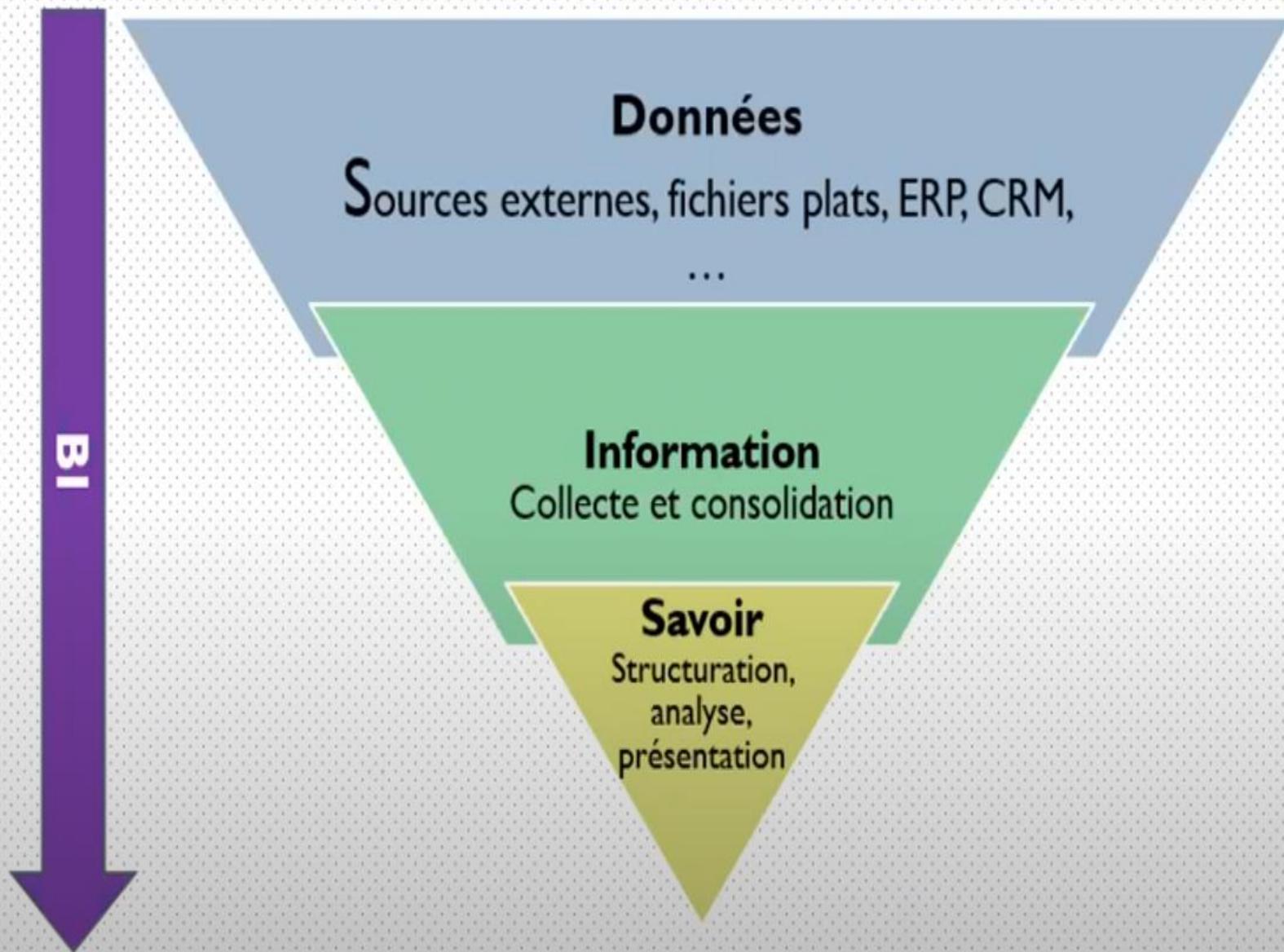


ilBoursa.com	ECONOMIE	BOURSE	OPCVM	START UP	LIFESTYLE	AFRIQUE	FORUM	FCP ILBOURSA CEA	
								Tunis	Choisir une vale
BH BANK		14.00	14.02	14.00	40	561	14.02	0.14%	
BH LEASING		4.16	4.16	4.15	395	1639	4.15	-0.24%	
BIAT		99.99	100.40	99.95	2 227	222700	100.00	0.01%	
BTE (ADP)		4.80	4.80	4.80	210	1008	4.80	0.00%	
CARTHAGE CEMENT		2.03	2.03	1.99	20 071	40142	2.00	-1.48%	
CELLCOM		2.03	2.08	2.03	7 333	15253	2.08	6.67%	
CERÉALIS		13.30	13.30	13.30	42	559	13.30	0.00%	
CIL		21.02	21.03	21.02	140	2944	21.03	0.05%	
CIMENTS DE BIZERTE		0.61	0.61	0.60	2 911	1747	0.60	-1.64%	
CITY CARS		12.70	12.70	12.60	19 297	244686	12.68	-0.16%	
DELICE HOLDING		14.18	14.25	14.15	6 884	97546	14.17	-0.07%	
ELBENE		2.28	2.28	2.28	20	46	2.28	10.14%	
ELECTROSTAR		0.31	0.31	0.30	555	167	0.30	-3.23%	
ENNAKL AUTOMOBILES		10.65	10.65	10.60	625	6656	10.65	0.00%	
ESSOUKNA		1.68	1.75	1.68	20	35	1.75	4.17%	
EURO-CYCLES		13.49	13.88	13.30	4 852	65939	13.59	-1.45%	
GIF FILTER		0.42	0.42	0.42	101	42	0.42	-2.33%	
HANNIBAL LEASE		7.50	7.50	7.50	1 080	8100	7.50	0.00%	
HEXABYTE		7.55	7.55	7.55	1	8	7.55	0.00%	
ICF		76.30	77.00	75.10	1 290	99330	77.00	0.92%	
LAND'OR		9.70	9.72	9.70	4 258	41303	9.70	0.00%	
MAGASIN GENERAL		7.40	7.40	7.40	200	1480	7.40	0.00%	
MAGHREBIA VIE		6.03	6.03	6.00	500	3000	6.00	-0.50%	
MEUBLES INTERIEURS		6.10	6.16	6.00	378	2287	6.05	-0.82%	
MIP		0.08	0.08	0.08	50	4	0.08	0.00%	
MONOPRIX		4.18	4.18	4.18	15 196	63519	4.18	4.50%	

# Problématique

- La prise de décisions stratégiques dans une organisation nécessite le recours et le croisement de multiples informations qui concernent tous les départements : production, RH, achats, ventes, marketing, service après-vente, maintenance, R&D...
- Or ces données sont généralement :
  - **éparpillées** au sein des départements et non connectées entre elles
  - **hétérogènes** dans leurs formats techniques et leurs organisations structurelles, voire leurs sémantiques
  - **implémentées pour l'action** et non pour l'analyse
  - **volatiles**, au sens où leur mise à jour peut conduire à oublier des informations obsolètes

# Le processus de l'informatique décisionnelle



# Nouveaux besoins!

Pour prendre de «**bonnes décisions** », on doit pouvoir:

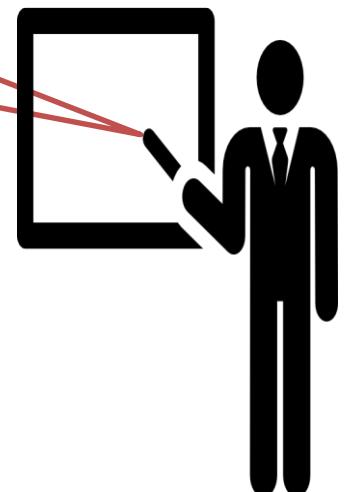
- Accéder en temps réel aux données de l'entreprise,
- Traiter ces données, Extraire l'information pertinente de ces données, par exemple pour savoir :
  1. **Quels** sont les résultats des ventes par gamme de produit et par région de l'année dernière ? !
  2. **Quelle** est l'évolution des chiffres d'affaires par type de magasin et par période ? !
  3. **Comment** qualifier les acheteurs de mon produit X ?



# C'est quoi l'informatique décisionnelle?

DÉFINITION 1 : (SOURCE : GARTNER)

Terme générique qui englobe les applications, l'infrastructure, les outils et les meilleures pratiques permettant l'accès et l'analyse de l'information afin d'améliorer et d'optimiser les décisions et les performances.

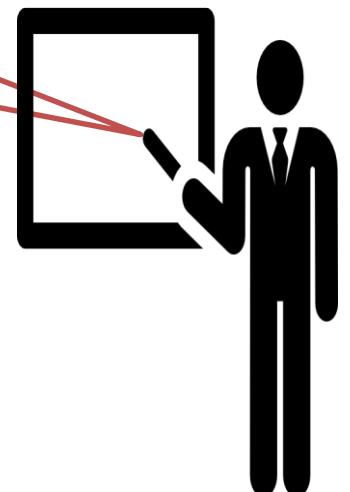


1

# C'est quoi l'informatique décisionnelle?

DÉFINITION 2 : (SOURCE : SMILE (OPEN SOURCE SOLUTIONS))

Un ensemble de solutions informatiques permettant l'analyse des données de l'entreprise, afin d'en dégager les informations qualitatives nouvelles qui vont fonder des décisions, qu'elles soient tactiques ou stratégiques.



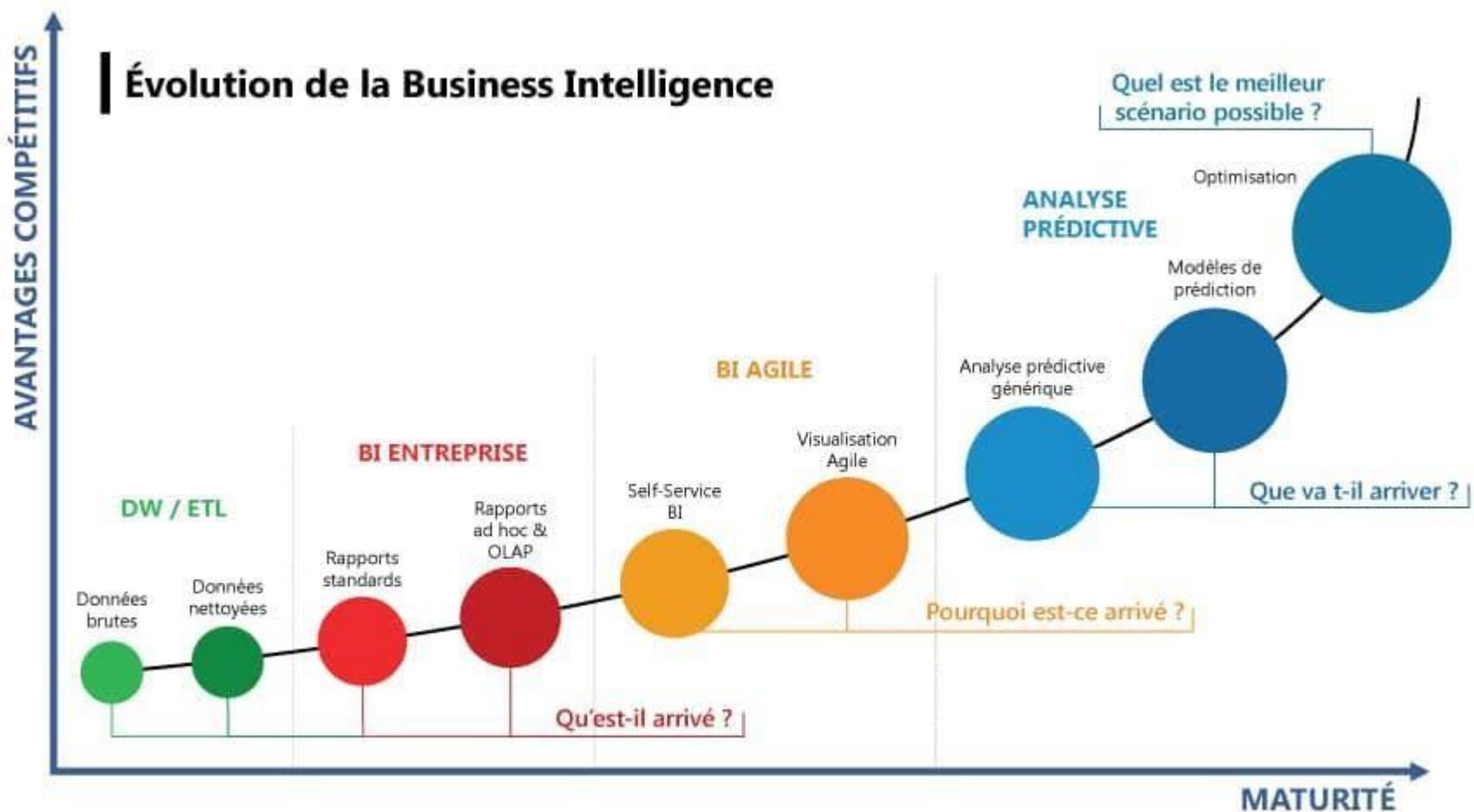
## C'est quoi l'informatique décisionnelle?

- L'informatique décisionnelle est l'ensemble des méthodes, moyens et outils informatiques utilisés pour piloter une entreprise et aider à la prise de décision.
- L'aide à la prise de décision est la responsabilité de quelques personnes dans l'entreprise appelés décideurs.
- L'informatique décisionnelle (BI) est une solution informatique permet de **collecter, de traiter** et **d'analyser** toutes les données de l'entreprise selon des critères spécifiques ou externe . Les résultats de cette analyse permettent aux **managers** d'obtenir **une vue globale** sur leur activité



3

# Évolution de l'informatique décisionnelle



# L'Intérêt de l'informatique décisionnelle?

Les décideurs d'une entreprise doivent pouvoir répondre à un certain nombre de question pour diriger leur entreprise :

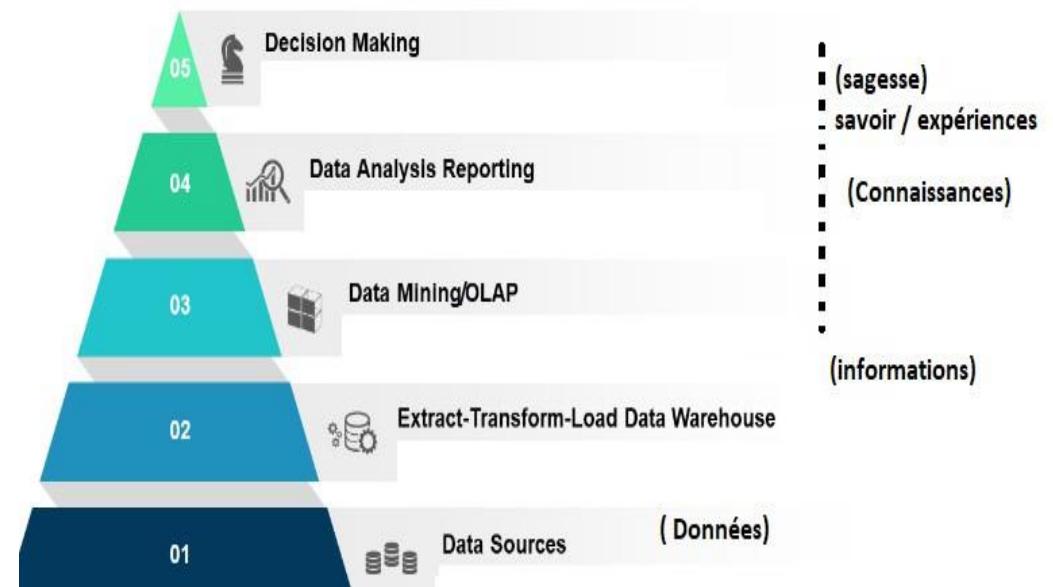
- Qui sont mes clients ?
- Pourquoi sont ils mes clients ?
- Comment cibler ma clientèle ?
- Quel est l'évolution de tel produit ?
- Qui sont mes employés ? ...



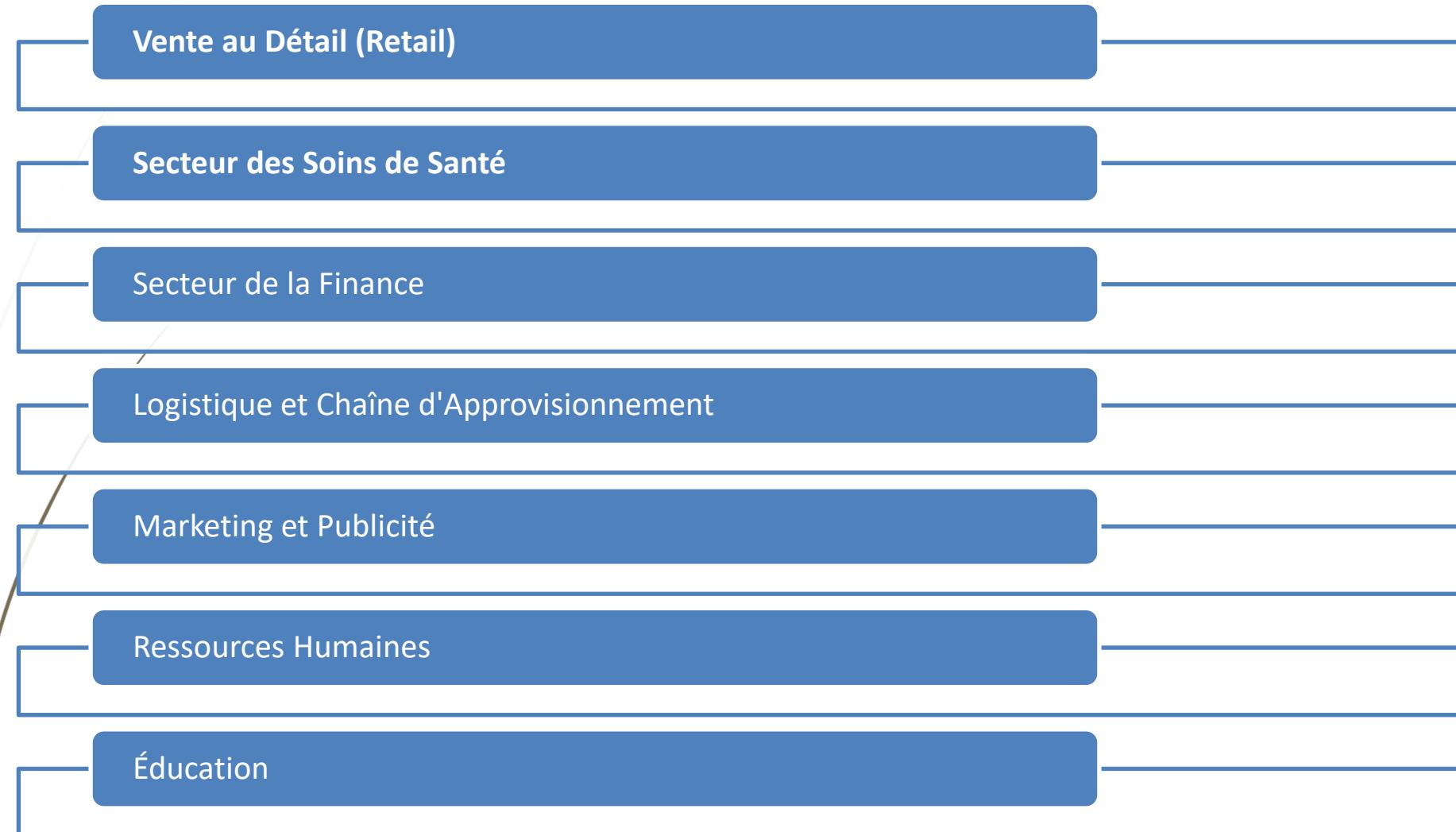
L'objectif est donc d'apporter aux décideurs d'une entreprise les moyens de répondre à ces questions, de connaître les tendances des clients pour ainsi bien anticiper les réactions de ses derniers et donc de fidéliser les clients.

# Objectifs d'un système décisionnel

- Transformer un système d'information qui avait une vocation de production en un système d'information décisionnel.
- Transformation des données de production en informations stratégiques
- Les données doivent être:
  - ◆ Extraites
  - ◆ Groupées et organisées
  - ◆ Corrélées
  - ◆ Transformées (résumé, agrégation)



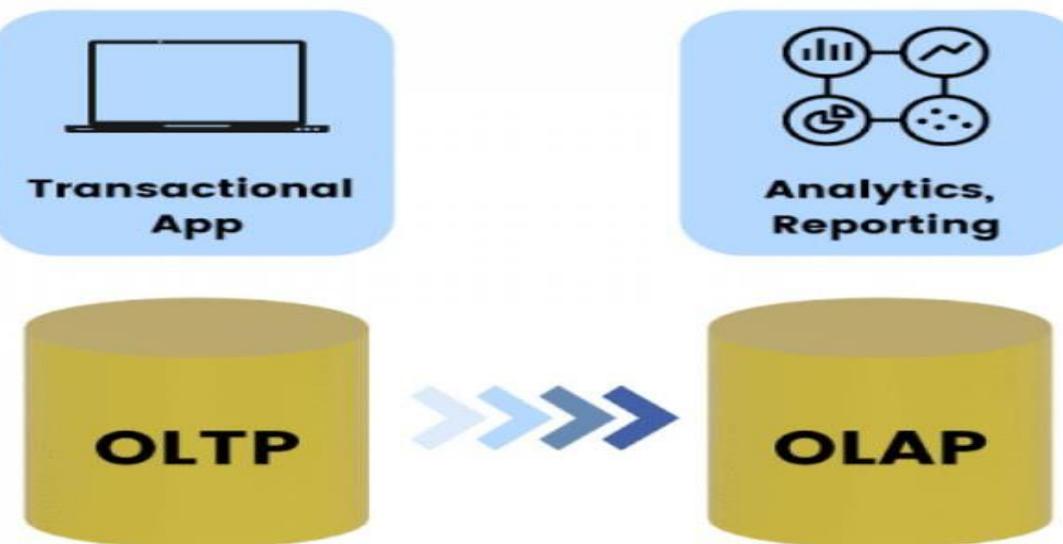
# Domaines d'application de la BI



## OLTP et OLAP (1/3)

Les systèmes informatiques peuvent se subdiviser en deux.

- Le **système transactionnel OLTP** (On-Line Transaction Processing)
- Le **système analytique OLAP**. (On-Line Analytical Processing)



# Pourquoi ne pas réutiliser les SGBD?

- Les SGBD sont des systèmes dont le mode de travail est transactionnel (**OLTP** On-Line Transaction Processing).
- Permet d'insérer, modifier, interroger des informations rapidement, efficacement, en sécurité.
- Deux objectifs principaux :
  - ◆ Sélectionner, ajouter, mettre à jour et supprimer des tuples
  - ◆ Ces opérations doivent pouvoir être effectuées très rapidement, et par de nombreux utilisateurs simultanément.

**Pour conclure, il convient de noter que les systèmes OLTP ne sont pas bien adaptés à l'analyse de données**

- Les systèmes **OLTP** servent, en général, de source de données pour les systèmes **OLAP** qui sont quant à eux, source d'analyse des données qui vont permettre d'aboutir à la décision.
- Vous comprendrez alors qu'**OLTP** et **OLAP** ont des objectifs opposés et ont un stockage de données différent faisant l'objet de requêtes différentes.

# OLTP vs OLAP (3/3)

Caractéristiques	OLTP	OLAP
Applications	production	aide à la décision
Utilisateurs	un département professionnel IT	transversal (entreprise) décideur non IT
Données	normalisées, non agrégées	dénormalisées, agrégées
Requêtes	simples, nombreuses, régulières, prévisibles, répétitives	complexes, peu nombreuses, irrégulières, non prévisibles
Nb tuples invoqués par requête (moyenne)	dizaines	millions
Taille données	100 MB à 1 GB	1 GB à 1 TB
Ancienneté des données	récente, mises à jour	historique
Rapidité de traitement	typiquement très rapide	dépend de la quantité de données impliquées



# **Partie 2:**

# **Methodologies de Gestion de projet BI**



# Méthodes appliquées pour un projet BI

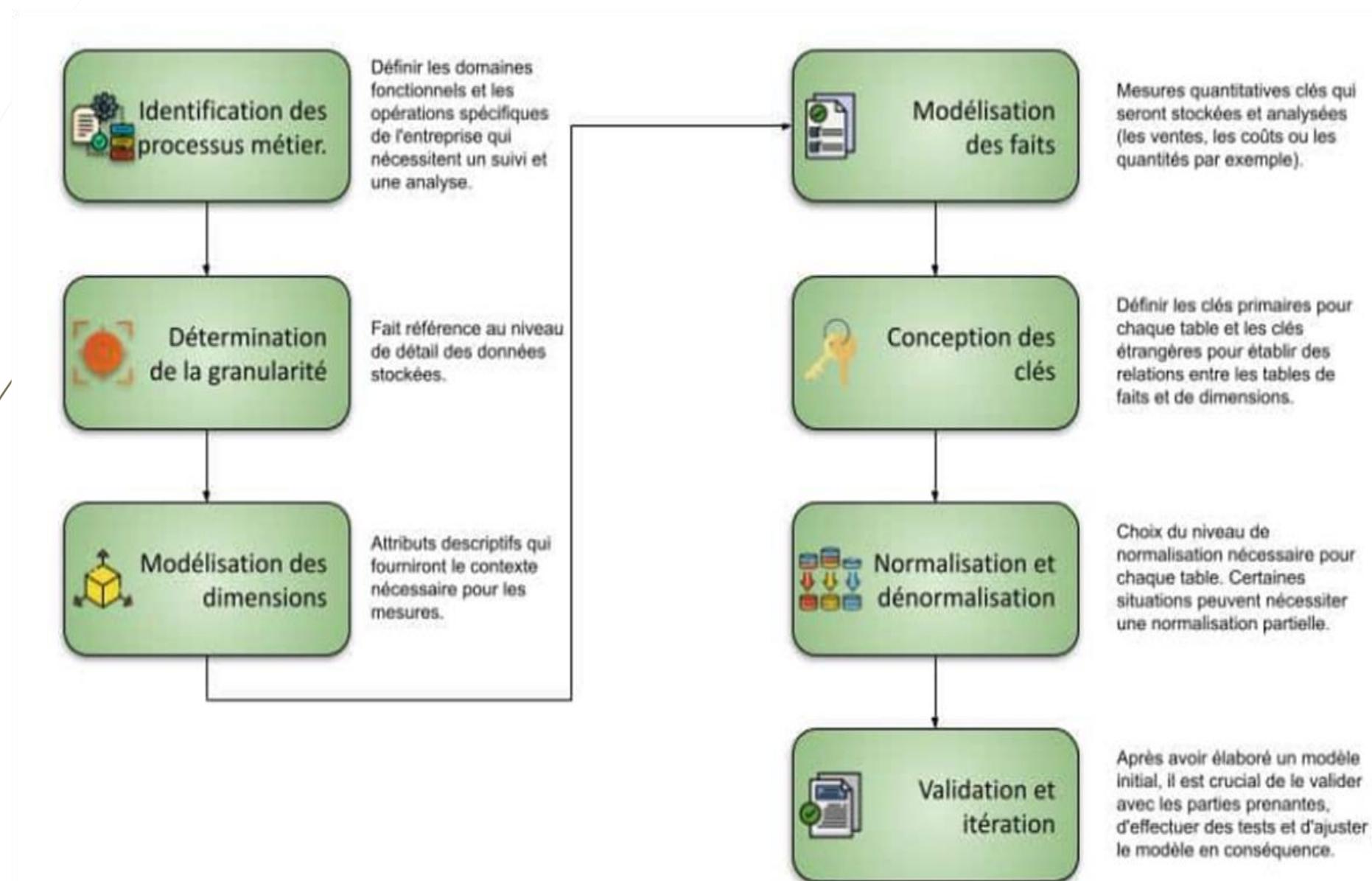
Une **méthodologie de gestion de projet** est un ensemble structuré de pratiques, principes, et processus utilisés pour planifier, exécuter, et finaliser un projet avec succès. Elle fournit un cadre pour organiser les tâches, gérer les ressources, et contrôler les risques tout en respectant les délais et le budget.

Pourquoi l'utiliser ?

- **Clarté et organisation** : Elle structure le projet en étapes définies, facilitant la gestion.
- **Gestion des risques** : Elle aide à anticiper et à résoudre les problèmes potentiels.
- **Efficacité** : Elle optimise l'utilisation des ressources et améliore la communication entre les équipes.
- **Suivi des objectifs** : Elle permet de mesurer les progrès et d'assurer que les objectifs sont atteints.

# Méthodes appliquées pour un projet BI

## Méthode KIMBALL

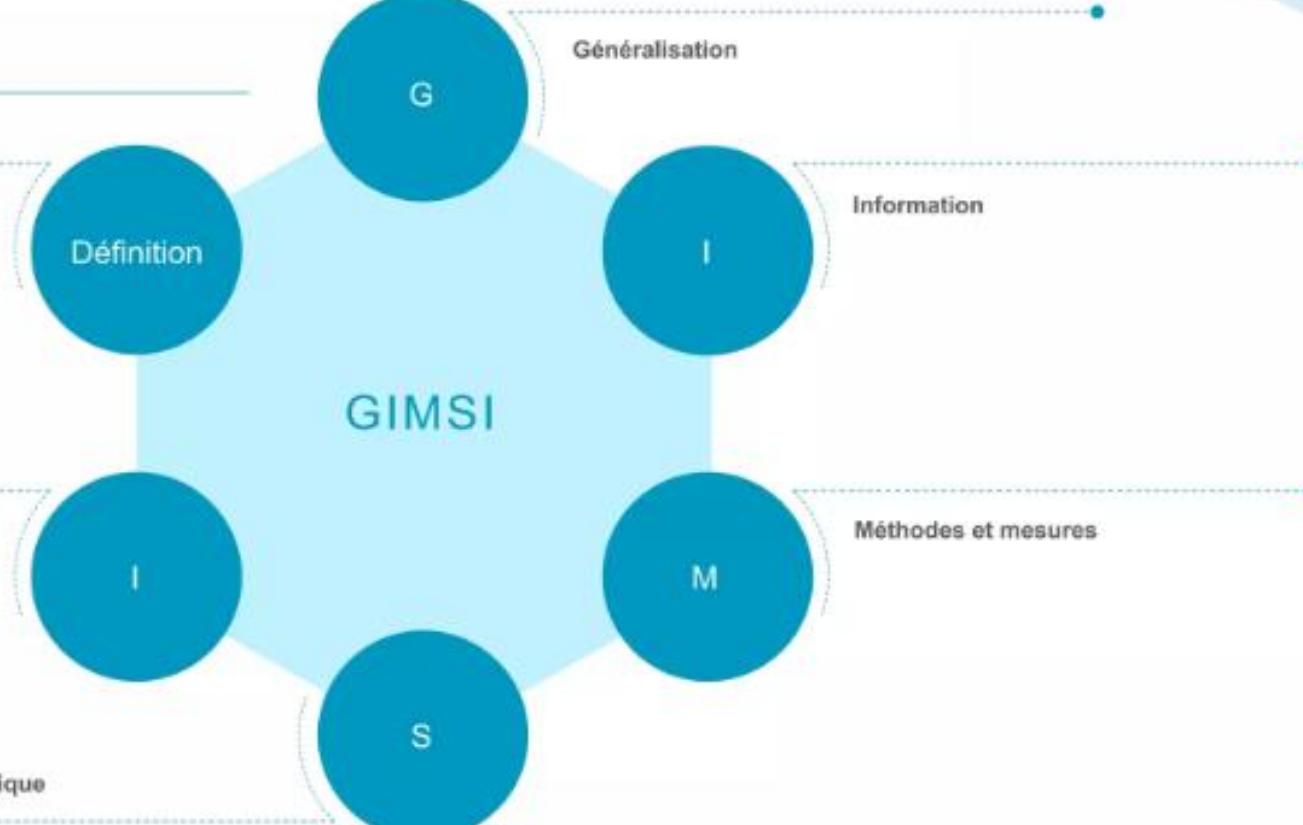


# Méthodes appliquées pour un projet BI

## Méthode GIMSI

### GIMSI

- Définit un cadre méthodologique afin de mieux formaliser les conditions de réussites du projet BI centré sur la problématique « du tableau de bord»



# Méthodes appliquées pour un projet BI

## Méthode GIMSI

Project management method : GIMSI



### Identification

- Business Environment
- Company Identification

### Design

- Objective Definition
- Dashboard Design
- Choosing Key Performance Indicators (KPIs)
- Data Collection
- Dashboard system

### Implementation

- Software Selection
- Integration and Deployment

### Improvement

- System Audit

# Partie 2: Entrepôt de données (Datawarehouse)



# Data Warehouse ou entrepôt de données

## Définition et objectifs (1/4)

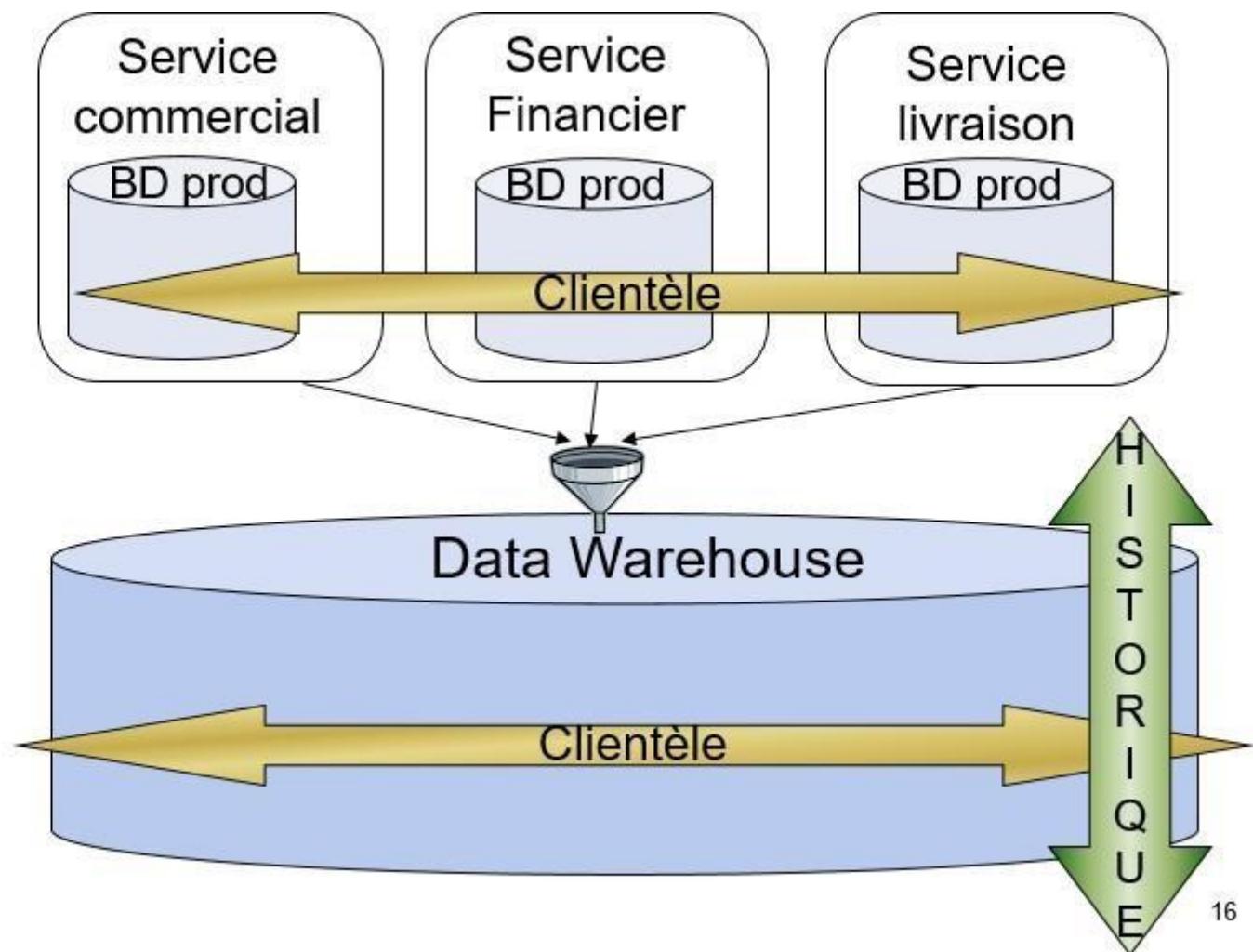
- Les datawarehouse sont des systèmes conçus pour l'aide à la prise de décision. (Mode de travail:**OLAP** On-Line Analytical Processing)
- D'après BILL Inmon (1996) :  

**« Un DW est une collection de données orientées sujet, intégrées, non volatiles, historisées, organisées pour la prise de décision. »**
- Les objectifs principaux sont
  - ◆ regrouper, organiser des informations provenant de sources diverses,
  - ◆ les intégrer et les stocker pour donner à l'utilisateur une vue orientée métier (**sujet**),
  - ◆ retrouver et analyser l'information facilement et rapidement.

# Data Warehouse ou entrepôt de données

## Définition et objectifs (2/4)

OLTP: On-Line  
Transactional  
Processing



OLAP: On-Line  
Analytical  
Processing

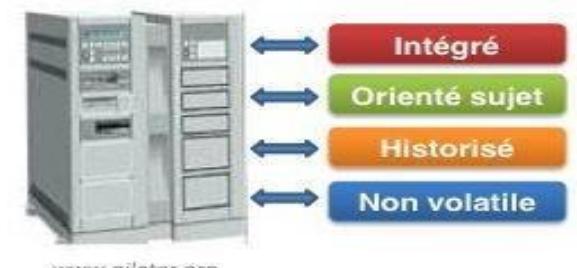
# Data Warehouse ou entrepôt de données

## Définition et objectifs (3/4)

Dans un entrepôt de données les données sont:

**Orientées sujet:** thèmes par activités majeures ; Le datawarehouse est lui organisé autour des sujets majeurs de l'entreprise.

**Intégrées:** Les données proviennent de plusieurs sources différentes. Avant d'être intégrées au sein du datawarehouse elles doivent être mise en forme et unifiées afin d'en assurer la cohérence. Cela nécessite une forte normalisation.



# Datawarehouse ou entrepôt de données

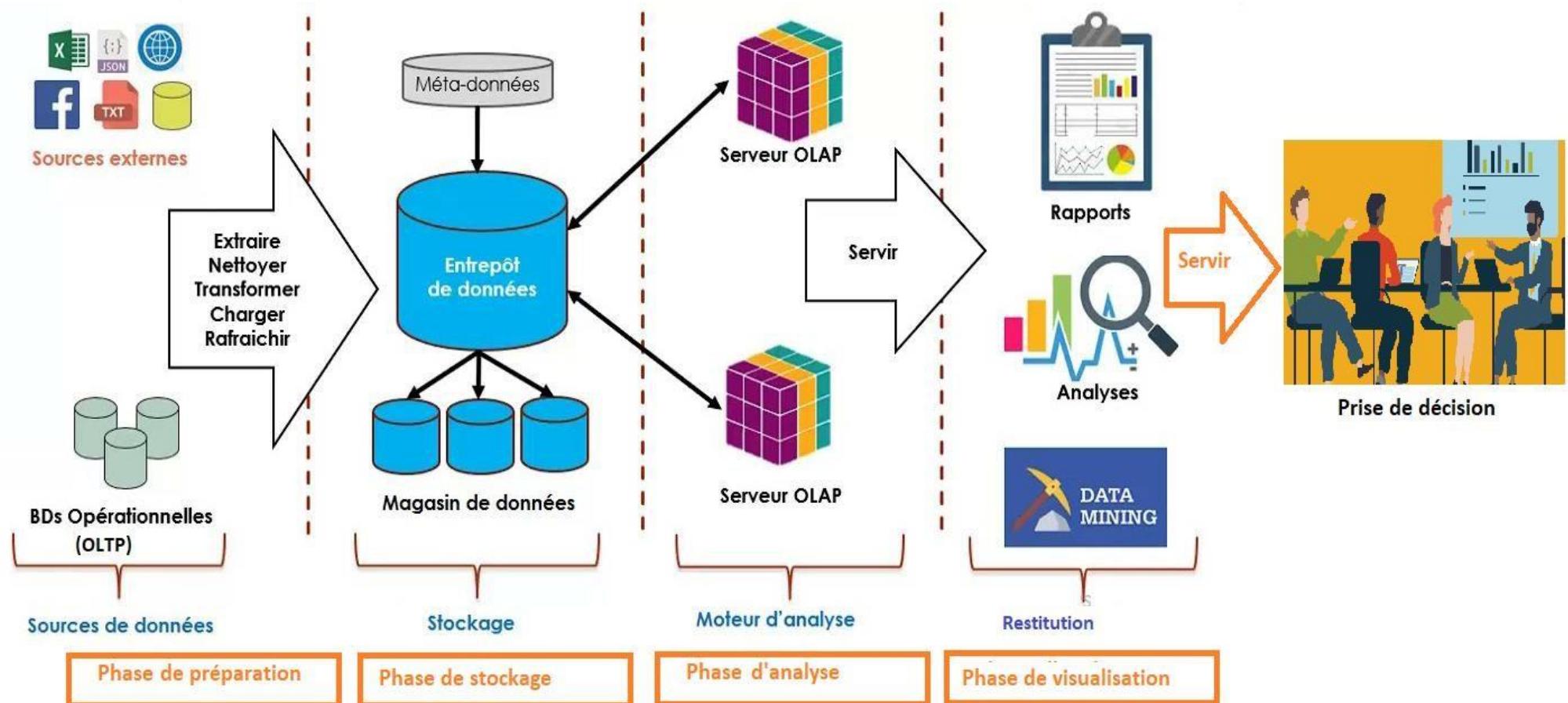
## Définition et objectifs (4/4)

**Non volatiles:** Un datawarehouse veut conserver la traçabilité des informations et des décisions prises.

Les données ne sont ni modifiées ni supprimées. Une requête émise sur les mêmes données à plusieurs mois d'intervalles doit donner le même résultat.

**Historisées:** Contrairement au système de production les données ne sont jamais mises à jour. Chaque nouvelle données est insérées. Un référentiel de temps doit être mis en place afin de pouvoir identifier chaque donnée dans le temps.

# Architecture Globale d'un système décisionnel



# Architecture Globale d'un système décisionnel



Cette phase fait intervenir des processus ETL qui se chargeront de récupérer les données issues de différentes sources de stockage, de les formater, nettoyer et consolider.



Cette phase permet de stocker les données sous une forme adaptée(l'entrepôt de données ou datawarehouse).

Le datawarehouse se charge de stocker et de centraliser les données en vue de la constitution du système d'information décisionnel.



Cette phase permet de structurer les données sous formes de cubes à l'aide des serveurs OLAP afin de faciliter l'analyse multidimensionnelle.

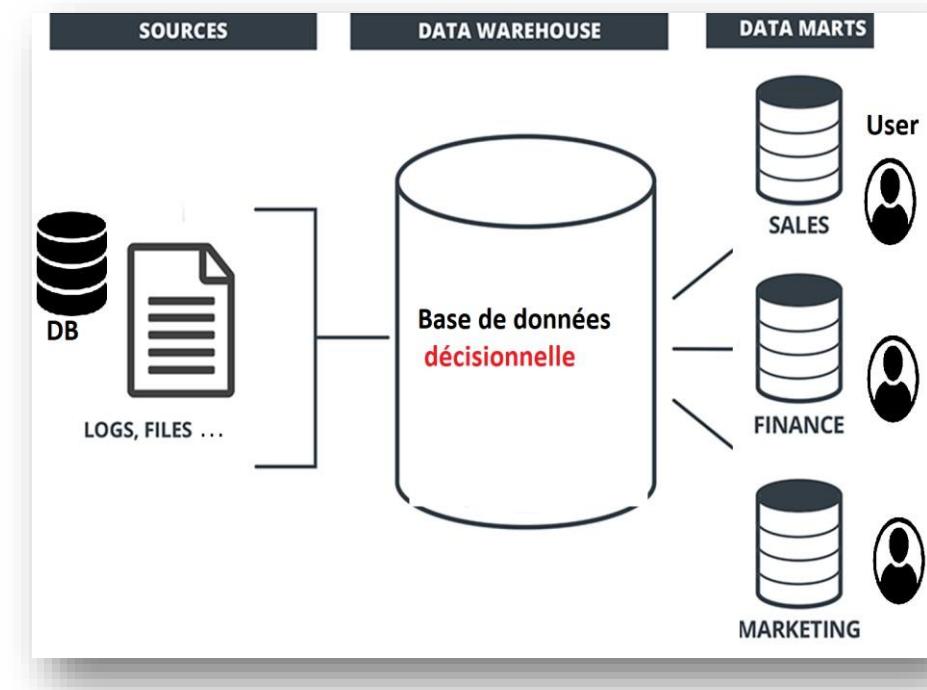


Les utilisateurs finaux interviennent dans cette dernière étape pour exploiter et analyser les données qui leur sont fournies.

# Data Warehouse et Datamart

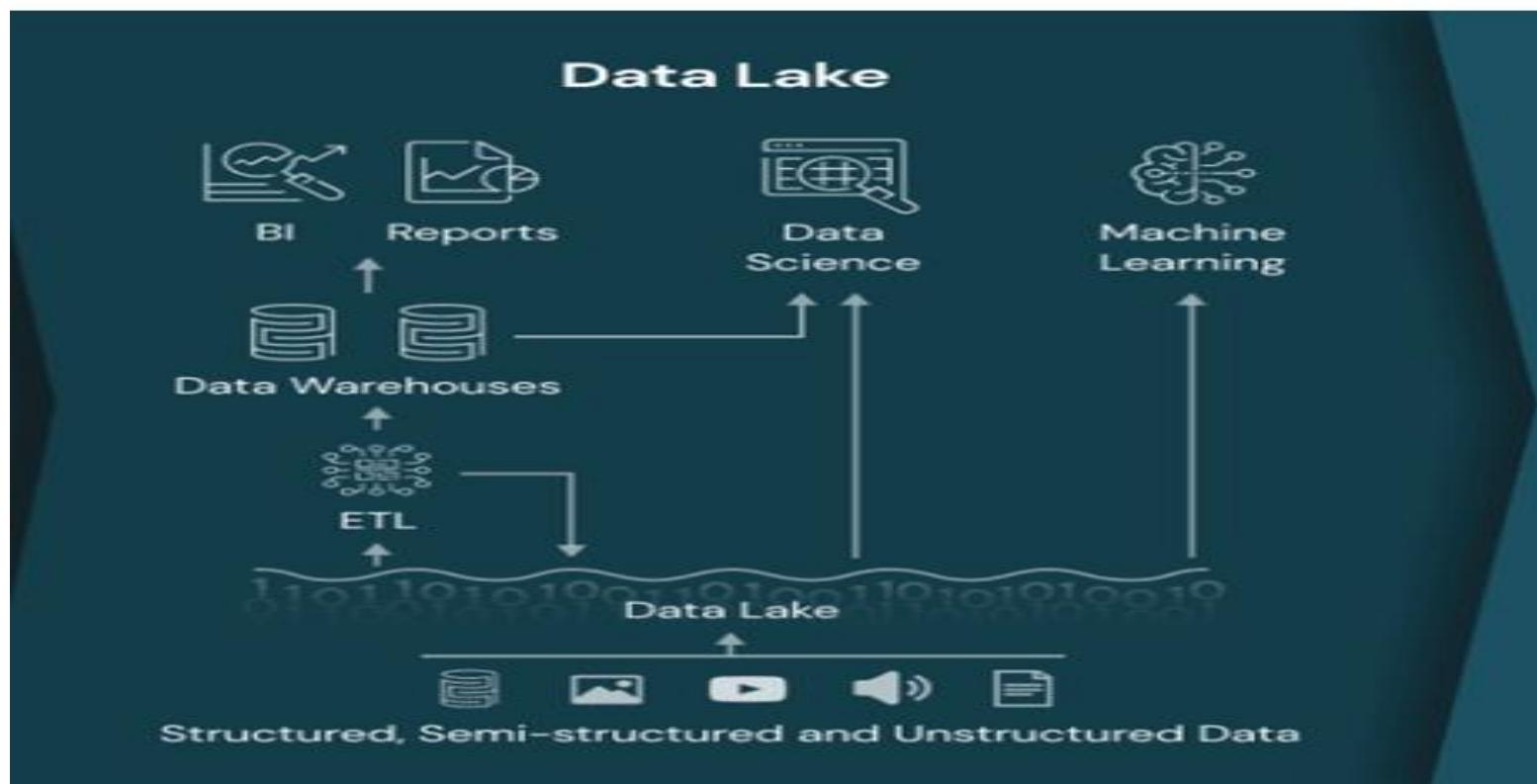
- Un Data Warehouse et un Data Mart se distinguent par le spectre qu'il recouvre :
  - Le Data Warehouse recouvre l'ensemble des données et problématiques d'analyse visées par l'entreprise.
  - Le Datamart recouvre une partie des données et problématiques liées à un métier ou un sujet d'analyse en particulier.
  - Un Datamart est un sous-

ensemble du Data Warehouse de l'entreprise, obtenu par extraction et agrégation des données de celui-ci.



# Data lake

- Un data lake est un emplacement de stockage centralisé.
- Il peut stocker des données structurées, semi-structurées ou non structurées, ce qui signifie que les données peuvent être conservées sous des formats plus souples pour une utilisation ultérieure.
- Lorsqu'il importe les données, le data lake les associe à des identificateurs et des balises de métadonnées pour une récupération plus rapide.



# Data Warehouse vs Data lake

<p>Le schéma d'un datawarehouse est défini et structuré avant le stockage.</p>	<p>Un data lake n'applique pas de schéma prédéfini, ce qui lui permet de stocker les données dans leur format natif.</p>
<p>Un data warehouse est facilement accessible aux utilisateurs techniques et non techniques grâce à son schéma clairement défini et documenté.</p>	<p>Accessibilité complexe les données ne sont pas organisées sous forme simplifiée avant leur stockage, un data lake a souvent besoin d'un expert ayant une compréhension approfondie des différents types de données existants.</p>
<p>Avec un data warehouse, vous devez prévoir non seulement du temps pour définir le schéma initial, mais aussi des ressources considérables pour modifier ce schéma à l'avenir chaque fois que les besoins de l'entreprise évoluent.</p>	<p>Les data lakes s'adaptent très facilement aux changements. Enfin, lorsque les besoins en capacité de stockage augmentent, il est plus facile de faire évoluer les serveurs d'un cluster de data lake.</p>

# Datawarehouse ou entrepôt de données

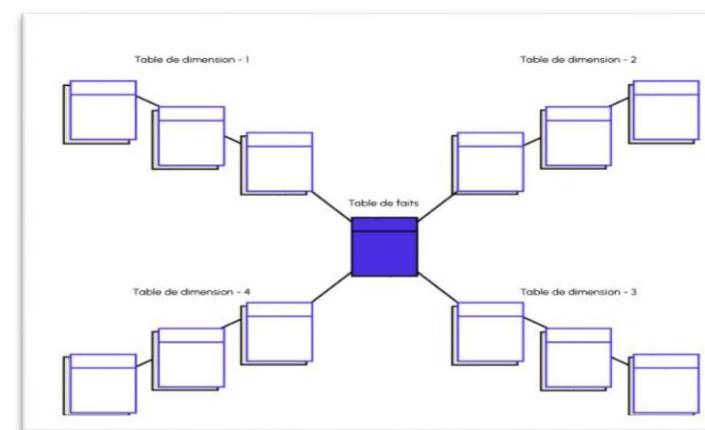
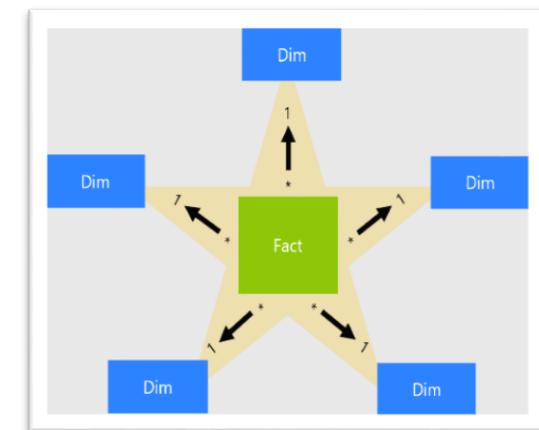
## Modélisation (1/2)

- Il existe trois schémas possibles pour la modélisation de l'entrepôt :

- Le schéma en étoile
- le schéma en flocon.
- Le schéma en constellation

- Quel que soit le modèle considéré, on distinguera :

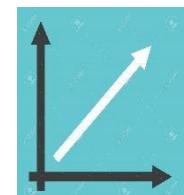
- **la table des faits** qui contient l'information à analyser (par exemple les ventes)
- **des tables de dimensions** qui contiennent les informations sur les dimensions d'analyse (par exemple le lieu, le temps, la description du produit).



# Datawarehouse ou entrepôt de données

## Modélisation (2/2)

- Pour la Modélisation conceptuelle BD : entité et relation
- Pour la Modélisation de **DW** : dimension et mesure
  - Les mesures sont les valeurs numériques que l'on compare (ex : montant\_ventes, qte\_vendue)
    - ◆ Ces valeurs sont le résultat d'une opération d'agrégation des données
  - Les dimensions sont les points de vues depuis lesquels les mesures peuvent être observées :
    - ◆ Ex : date, localisation, produit, etc.
    - ◆ Elles sont stockées dans les tables de dimensions



# Modélisation du DW

## Dimension

- Une **dimension** peut être définie comme :
  - ◆ un thème, ou un axe (attributs), selon lequel les données seront analysées
  - ◆ Ex : Temps, Région, Produits...
- Une dimension contient des membres organisés en hiérarchie:
  - ◆ Chacun des membres appartient à un niveau hiérarchique (ou niveau de granularité) particulier
  - ◆ Ex : pour la dimension Temps:
    - ◆ année – semestre – mois – jour

# Modélisation du DW

## Mesure – fait – table de faits

- Une mesure est un élément de donnée sur lequel portent les analyses, en fonction des différentes dimensions
  - ◆ **Exemple:** coût des travaux, nombre d'accidents, ventes
- Un fait représente la valeur d'une mesure, mesurée ou calculée, selon un membre de chacune des dimensions
  - ◆ **Exemple :** «250 000 euros » est un fait qui exprime la valeur de la mesure « coût des travaux » pour le membre « 2002 » du niveau année de la dimension « temps » et le membre « Versailles » du niveau « ville » de la dimension « découpage administratif »
- Les mesures sont stockées dans les tables de faits
  - ◆ Table de fait contient les valeurs des mesures et les clés vers les tables de dimensions

# Le Modèle en étoile

## Avantages et inconvénients

- Une table de faits comprenant une (ou plusieurs) mesures.

Plusieurs tables de dimension dénormalisées

- Les tables de dimension n'ont pas de lien entre elles

### Avantages :

Facilité de navigation

- Alimentation facile.
- Performances : nombre de jointures limité ; gestion des données creuses.
- Gestion des agrégats

### Inconvénients :

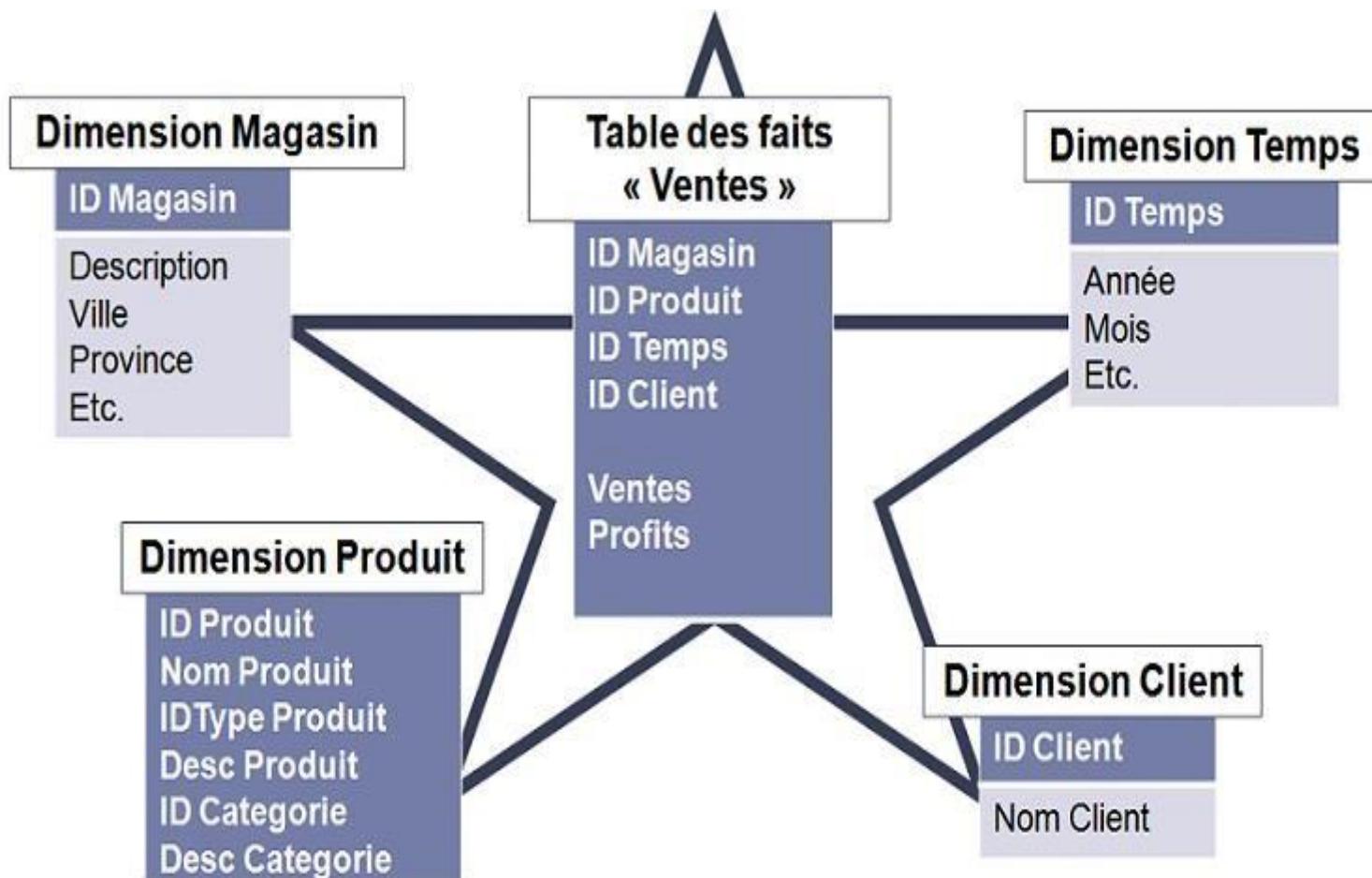
Toutes les dimensions ne concernent pas les mesures

- Redondances dans les dimensions

-

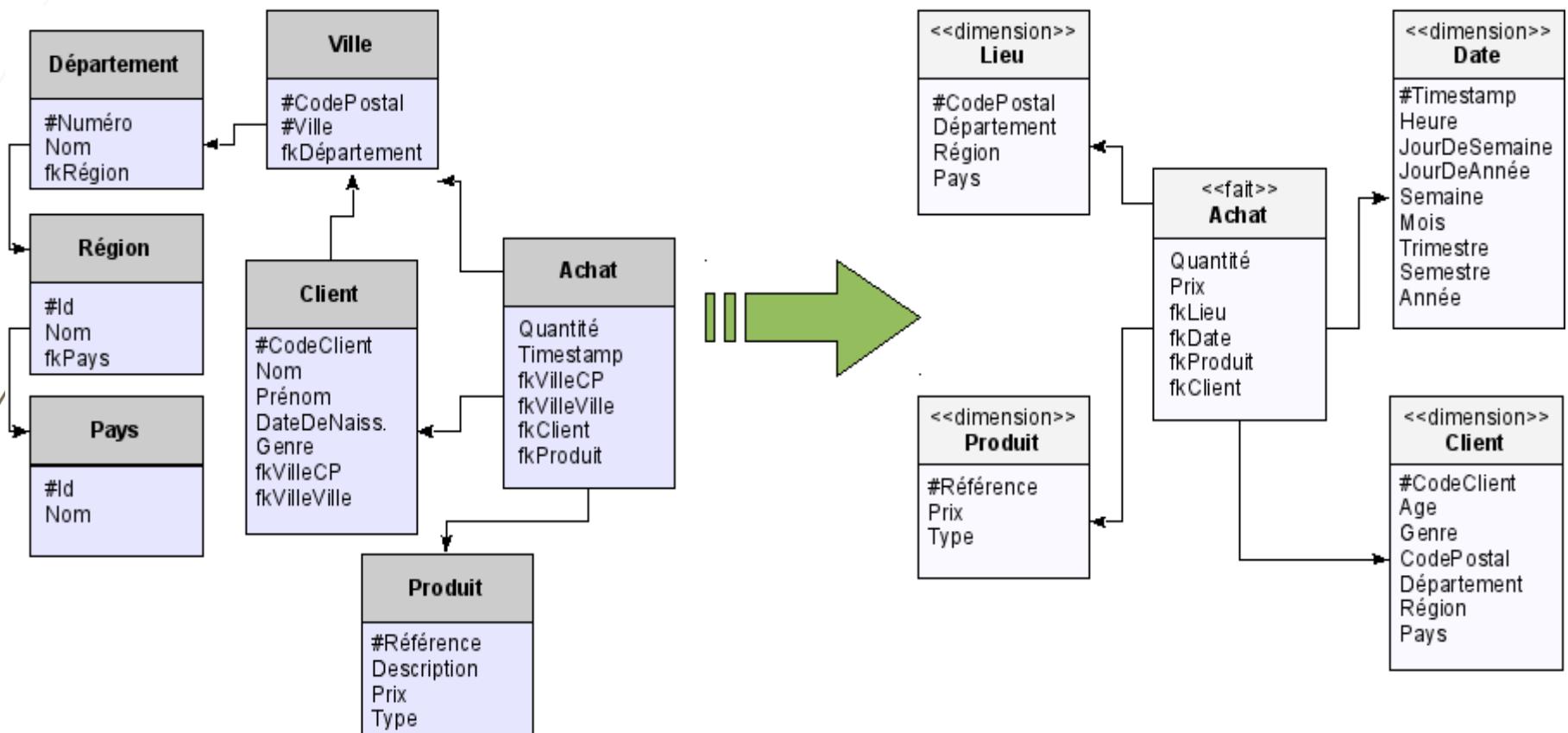
# Le Modèle en étoile

## Exemple1



# Le Modèle en étoile

## Exemple2



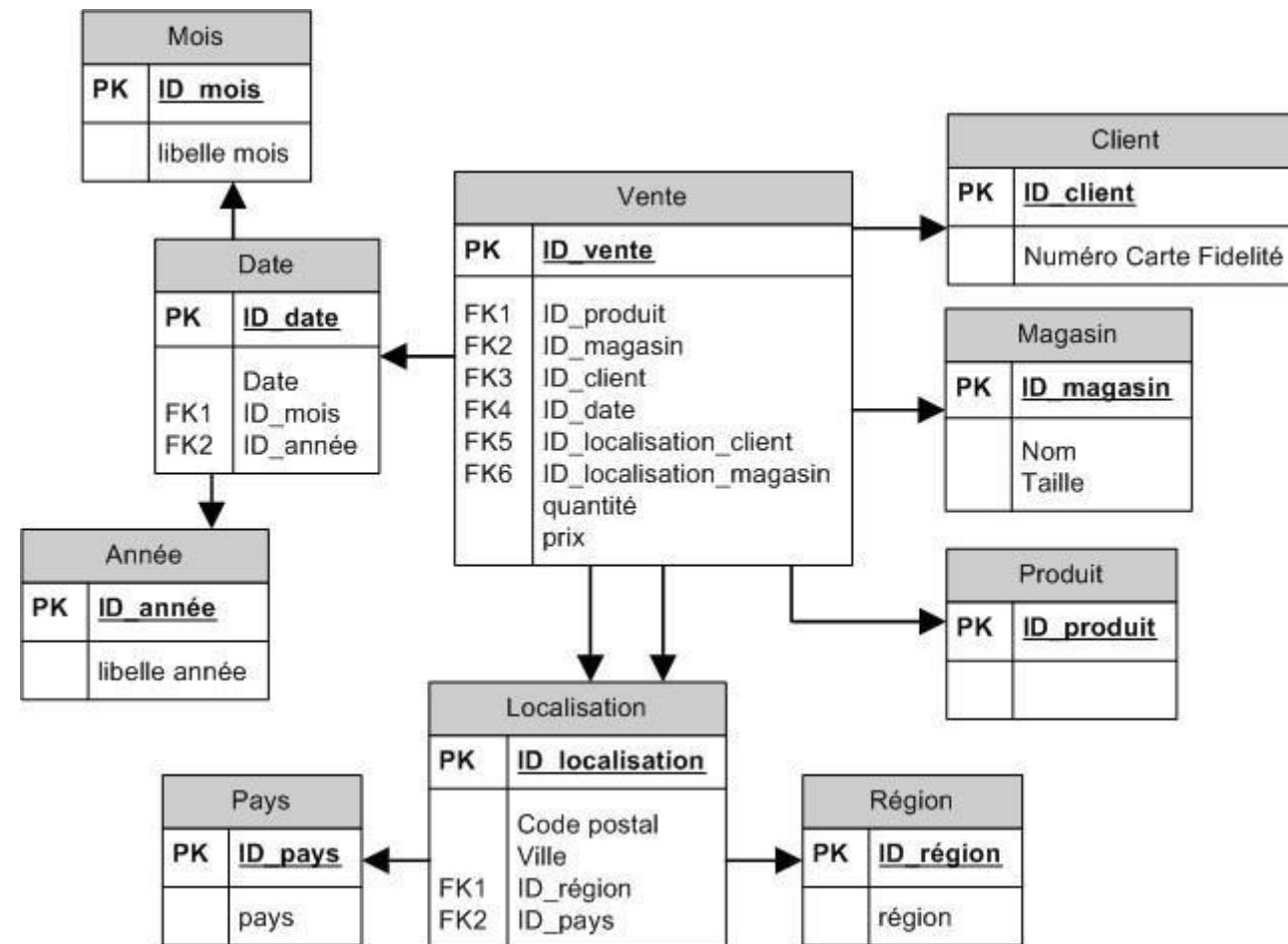
# Le Modèle en flocons

## Avantages et inconvénients

- Le schéma en flocon est dérivé du schéma en étoile où les tables de dimensions sont normalisées (la table des faits reste inchangée).
- Avec ce schéma, chacune des dimensions est décomposée selon sa (ou ses) hiérarchie(s).
  - Exemple : Commune, Département, Région, Pays, Continent
- Utilisé lorsque les tables sont trop volumineuses
  - ◆ Avantages :
    - réduction du volume,
  - ◆ Inconvénients :
    - navigation difficile,
    - nombreuses jointures.
    - Alimentation complexe.

# Le Modèle en flocons

## Exemple

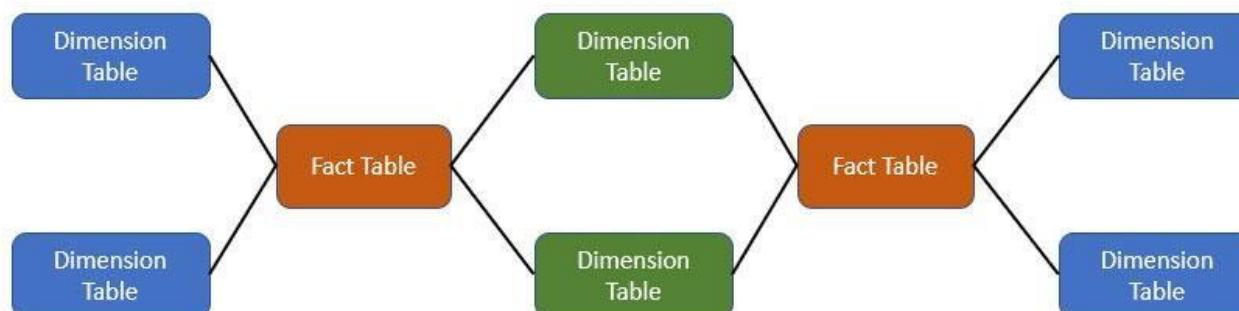


# Modèle en constellation

## Présentation

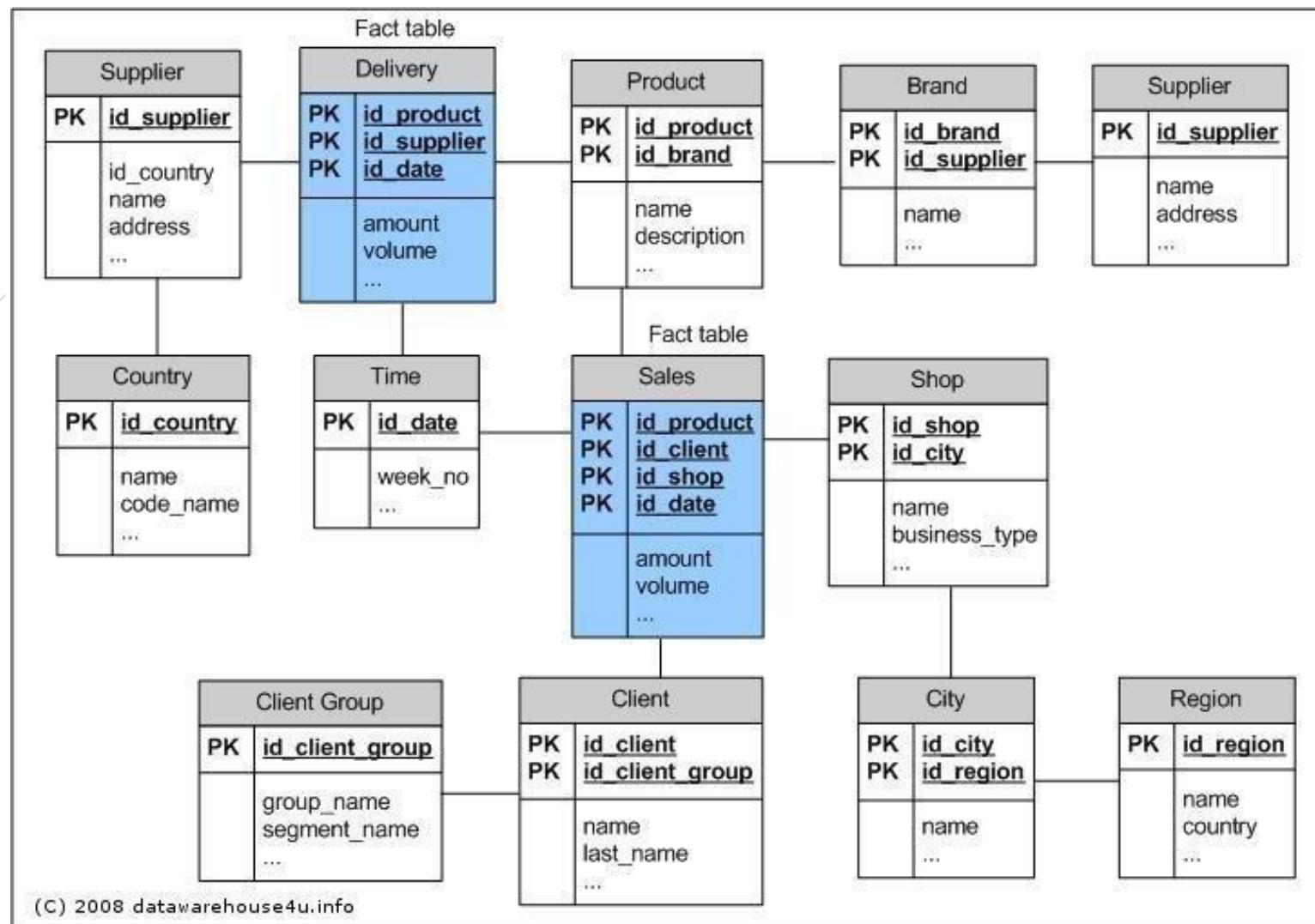
La modélisation en constellation consiste à fusionner plusieurs modèles en étoile qui utilisent des dimensions communes.

Un modèle en constellation comprend donc plusieurs tables de faits et des tables de dimensions communes ou non à ces tables de faits.



# Modèle en constellation

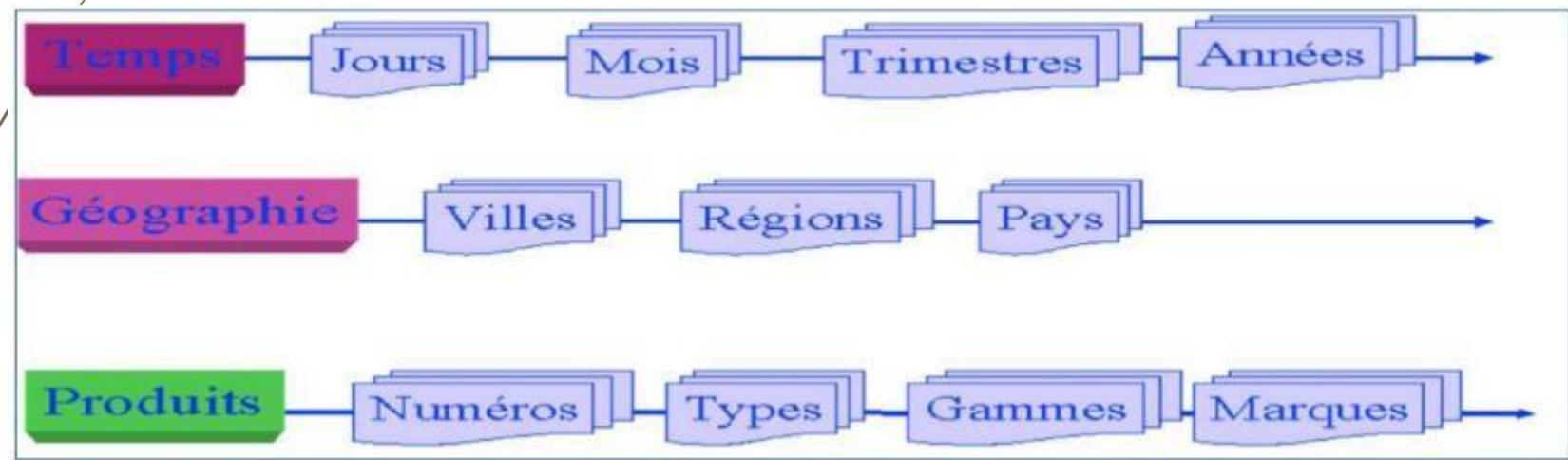
## Exemple



## 30 Hiérarchie des paramètres d'une dimension

- Une mesure est généralement analysée selon les dimensions qui le caractérisent.
- Nécessaire de définir pour chaque dimension ses différents niveaux de détail définissant ainsi une (ou plusieurs) hiérarchie(s) de paramètres

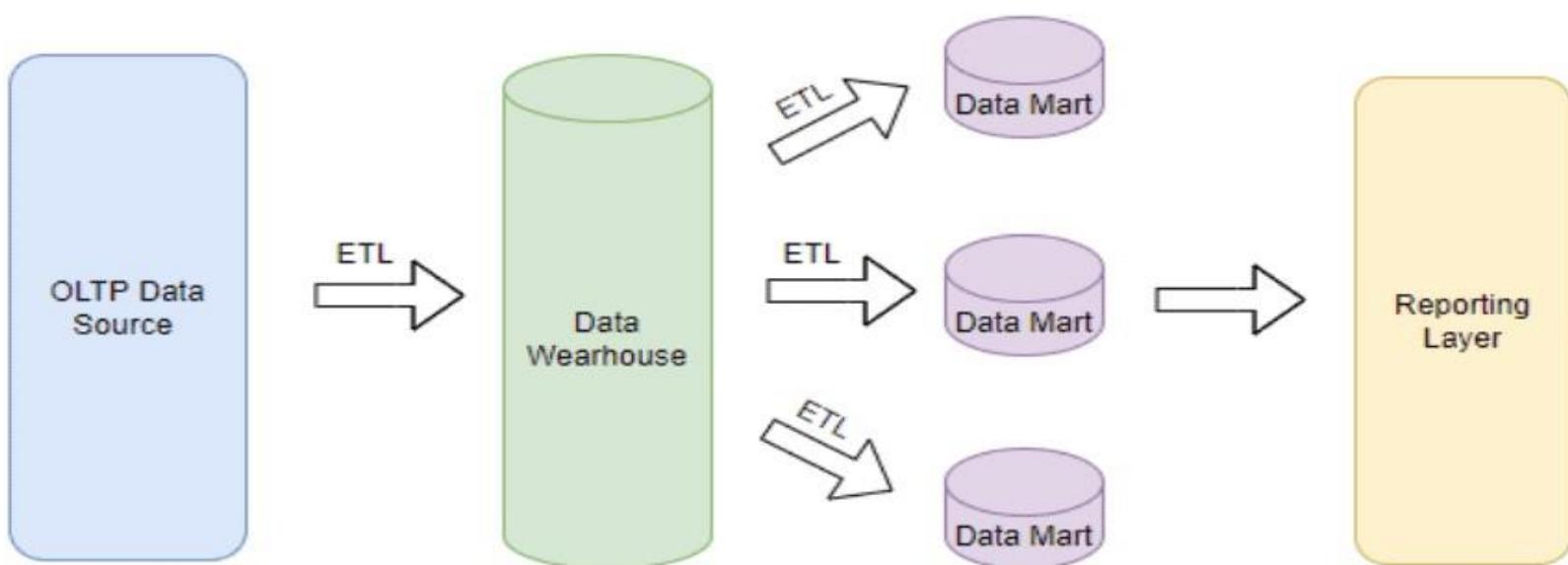
Exemple :



# Approches Modélisation DataWarehouse

- **L'approche d'Inmon (ou approche Top-Down):**

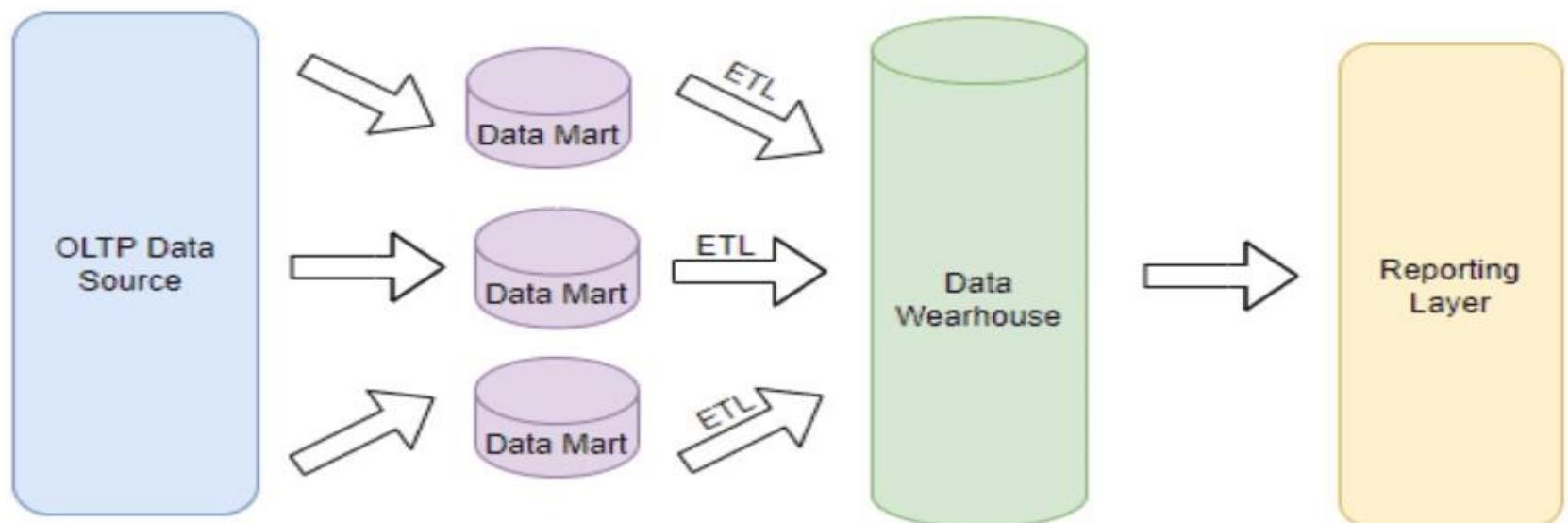
le Data Warehouse est un référentiel centralisé d'entreprise stockant l'information au niveau le plus détaillé. Des Datamarts modélisés sous forme de schémas en étoile sont ensuite créés à partir de ce Data Warehouse.



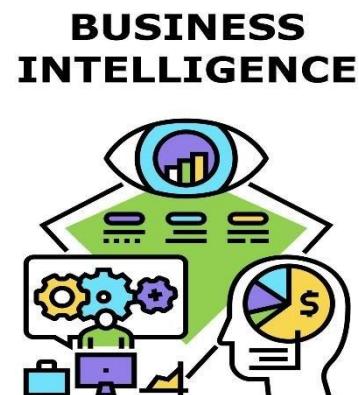
# Approches Modélisation DataWarehouse

- **L'approche de Kimball (ou approche Bottom-Up)**

Le Data Warehouse peut être vu comme l'union des datamarts cohérents entre eux grâce aux dimensions conformes (data warehouse bus). La modélisation dimensionnelle permet un requetage facile et performant pour les utilisateurs.

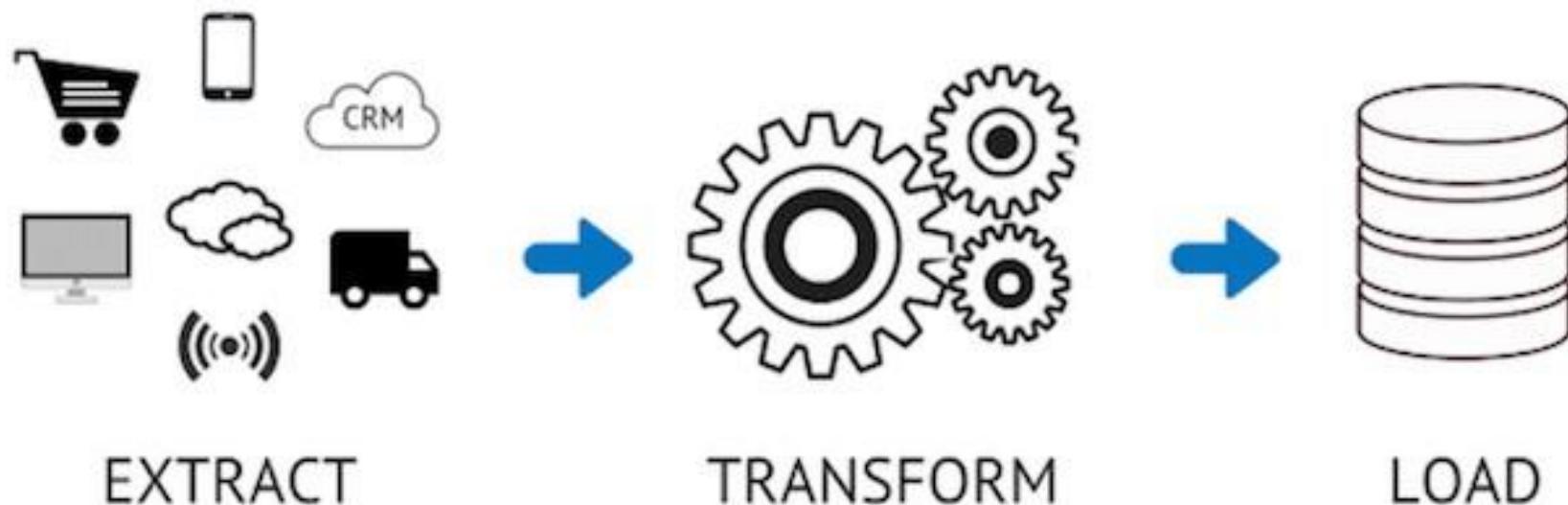


## Partie 3: Notions de la BI



# ETL: Extract Transform Load

Processus permettant de collecter des données en provenance de sources multiples pour ensuite les convertir dans un format adapté à une Data Warehouse et les y transférer



# ETL: Extract Transform Load



L'extraction est la première étape du processus d'apport de données à DW.

Extraire cela veut dire lire et copier les données sources dans la zone de préparation en vue de manipulations ultérieures

La transformation est la seconde phase du processus. Elle sert à la:

- Consolidation des données.
- Correction des données et élimination de toute ambiguïté.
- Elimination des données redondantes.
- Compléter et renseigner les valeurs manquantes.

C'est la dernière phase de l'alimentation d'un entrepôt de données, le chargement est une étape indispensable. Elle reste toutefois très délicate et exige une certaine connaissance des structures du système de gestion de base de données.

# ODS vs STAGING AREA

## I Points Communs:

Permettent de stocker les données extraites des SI sources

Faire des opérations sur ces données

## I Différences:

Dans le cas du staging Area, les données sont détruites directement après avoir été chargées dans le Datawarehouse mais pas pour l'ODS où les données auront quand même une durée de vie plus longue.

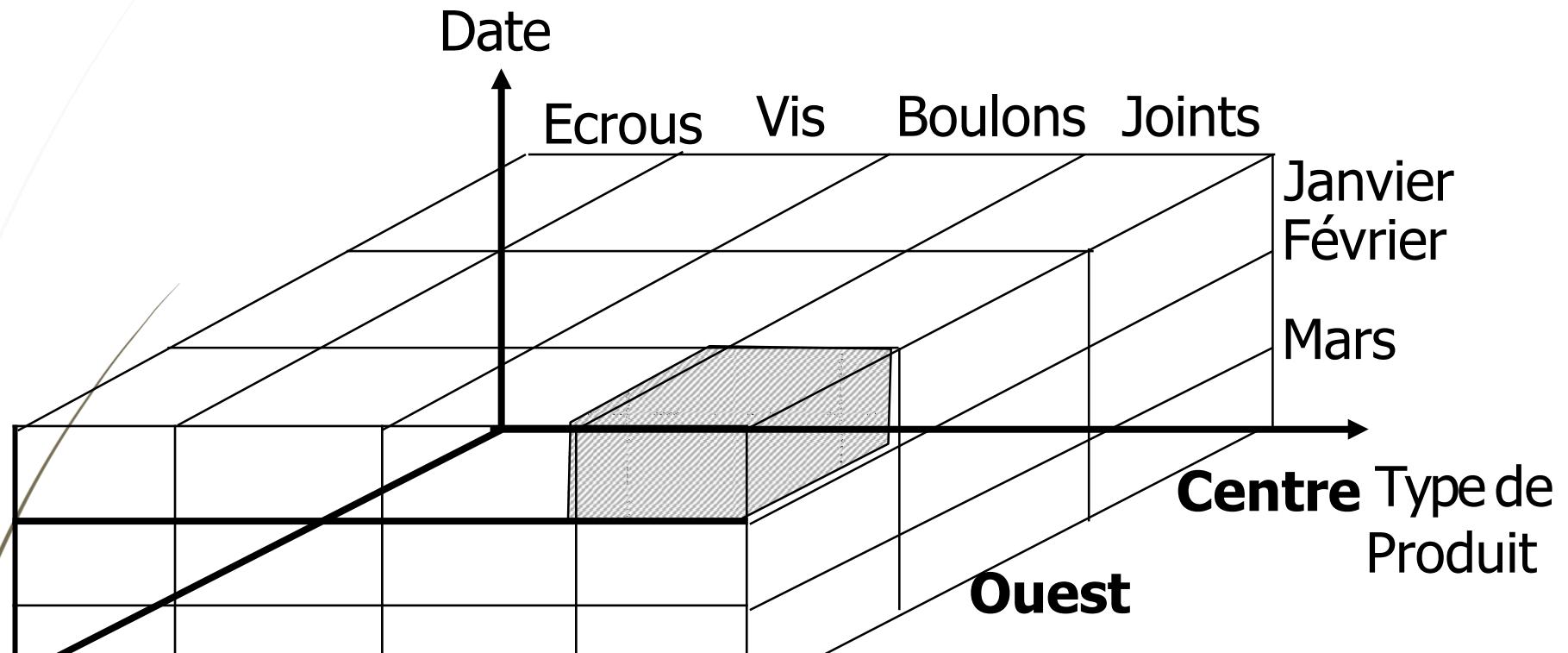
l'ODS répond plus à une problématique de reporting immédiat dans le sens où l'ODS sera mis à jour plus souvent que le datawarehouse : on pourrait dire que l'ODS pourrait être alimenté toutes les semaines et le Datawarehouse une fois par mois.

# Partie 4: Cube OLAP

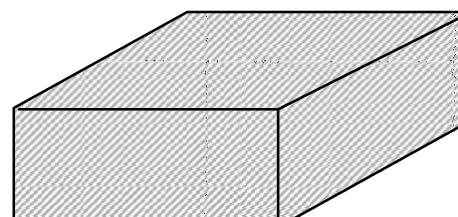


# Cube OLAP

I représentation des données sous forme multidimensionnelle

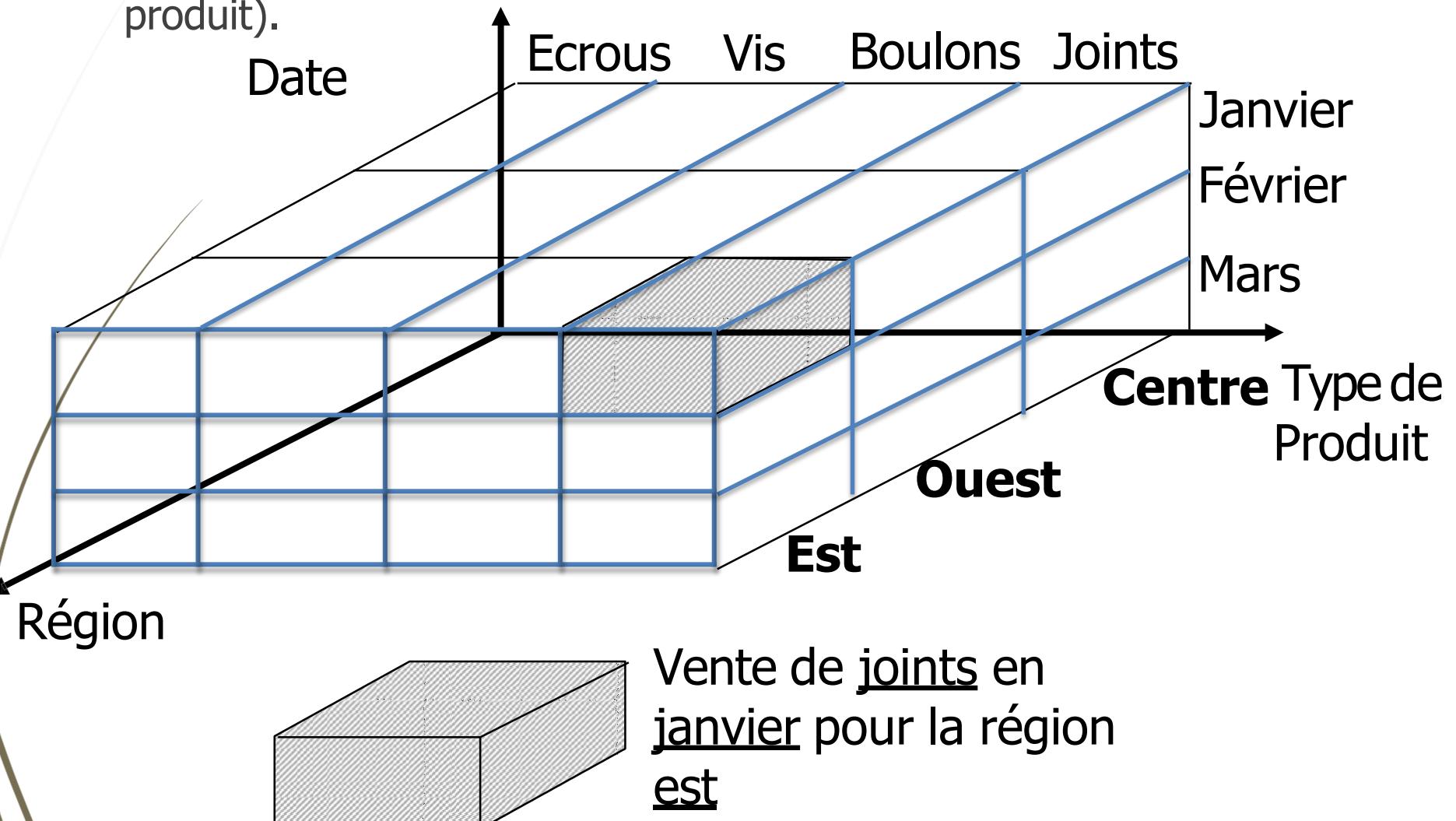


Vente de joints en  
janvier pour la région  
est



# Cube OLAP

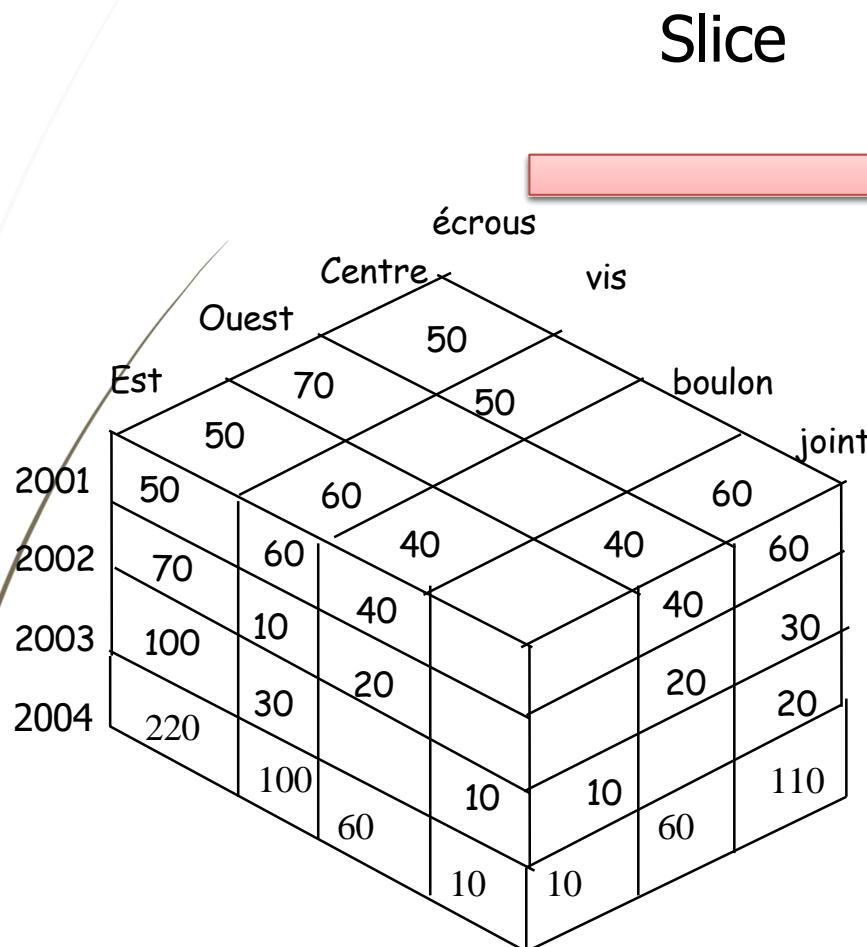
- ◆ Axes: dimensions (date, type de produits, région),
- ◆ Chaque cellule de l'hypercube contient une mesure calculée (vente de produit).



# Manipulation d'un cube OLAP

39

**SLICING:** Sélection de tranches du cube par des prédictats selon une dimension



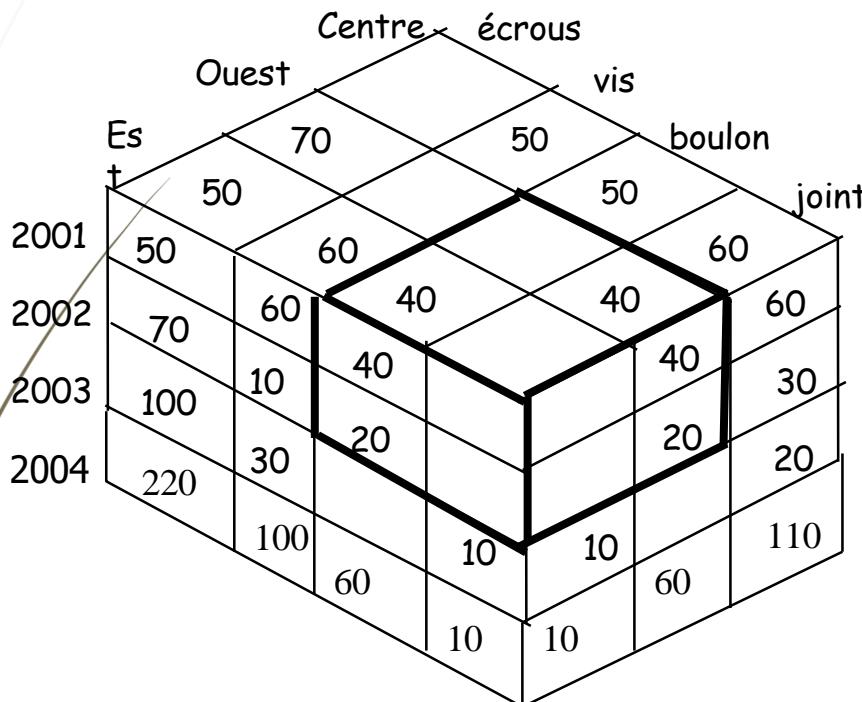
Slice (2004)

Ventes 2004	écrou	vis	boulon	joint
est	220	100	60	10
ouest	160	50	10	60
centre	20	150	170	110

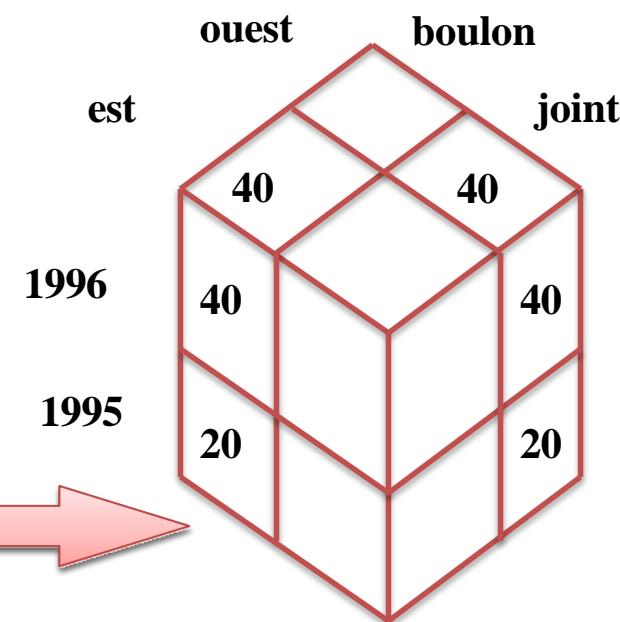
# Manipulation d'un cube OLAP

40

**DICING:** extraction d'un sous-cube

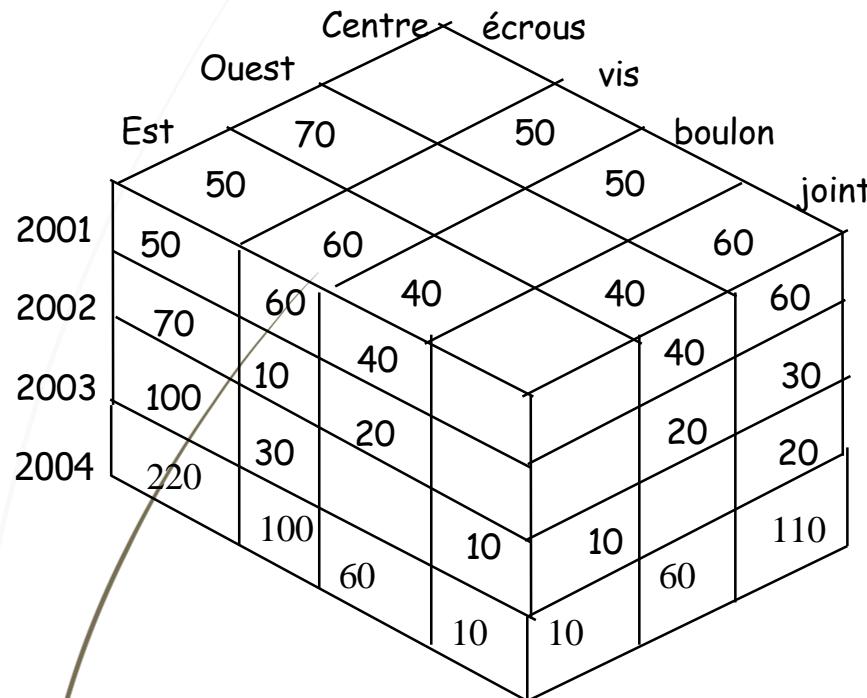


Dice

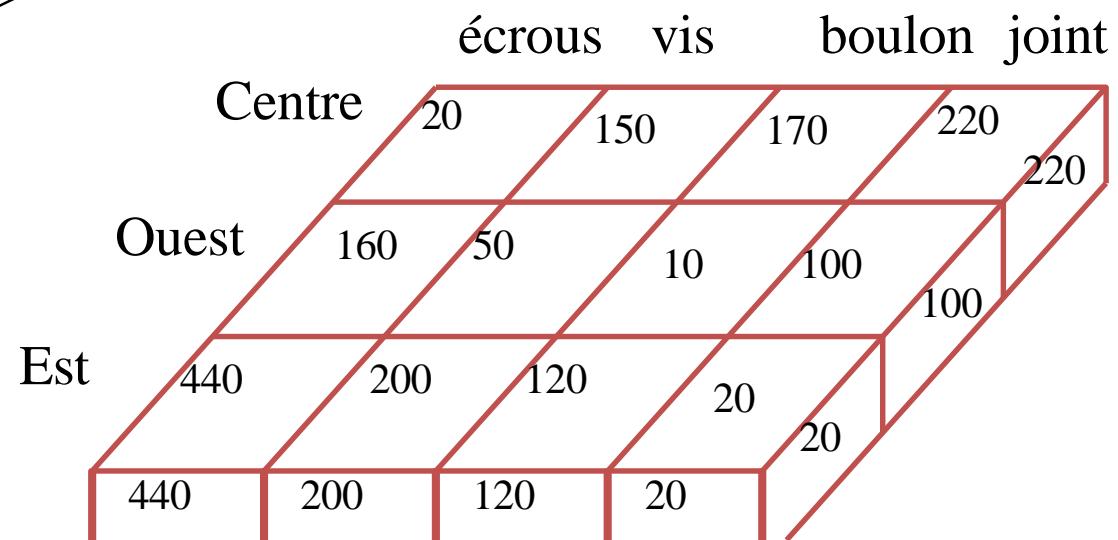


# Manipulation d'un cube OLAP

- **Roll-Up:** représenter les données du cube à un niveau de granularité supérieur conformément à la hiérarchie définie sur la dimension.



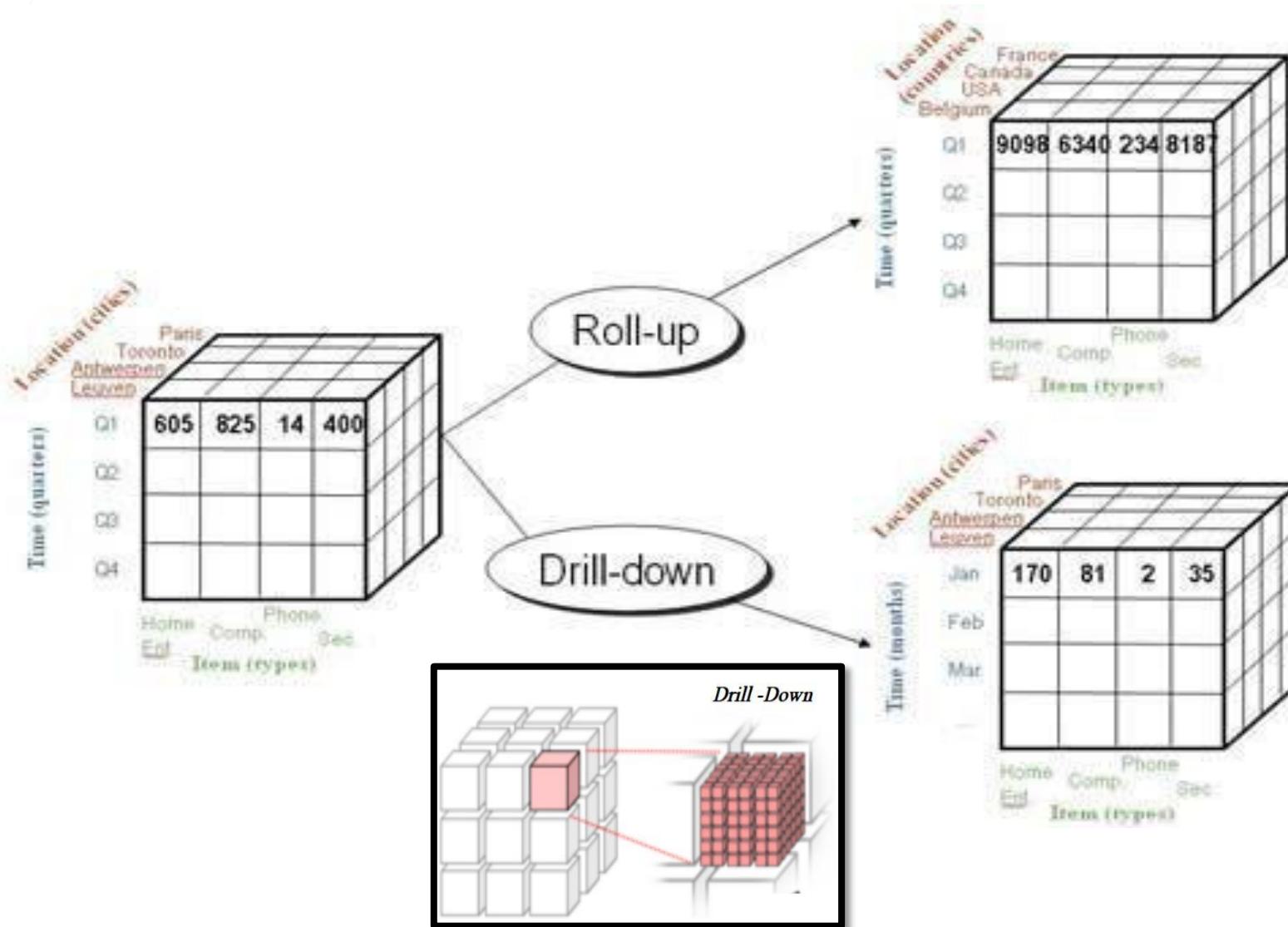
Roll up sur année



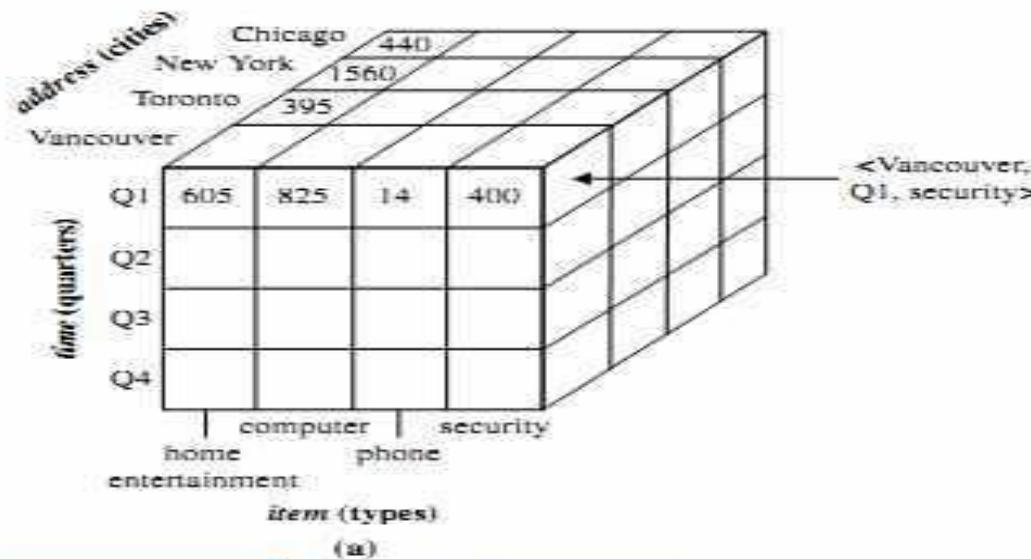
# Manipulation d'un cube OLAP

42

- **Drill-Down:** à représenter les données du cube à un niveau de granularité de niveau inférieur, donc sous une forme plus détaillée.



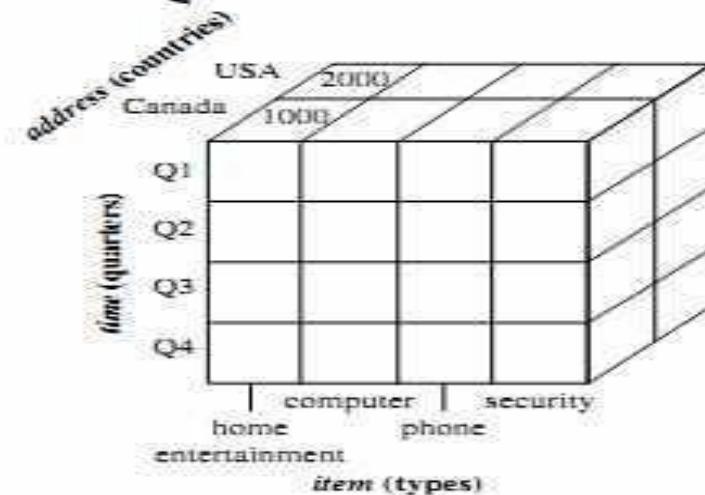
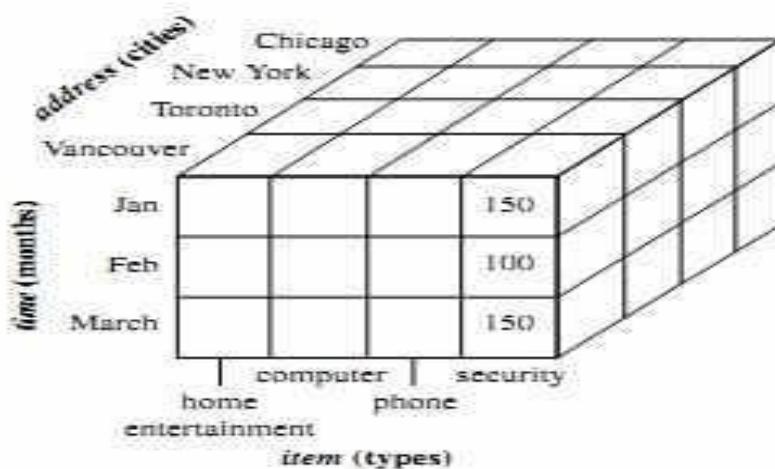
# Manipulation d'un cube OLAP



(a)

Drill-down  
on time data for Q1

Roll-up  
on address

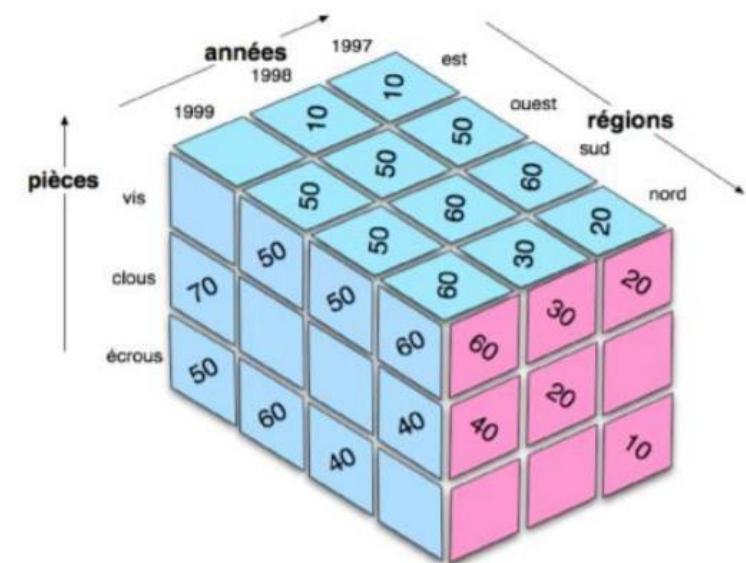
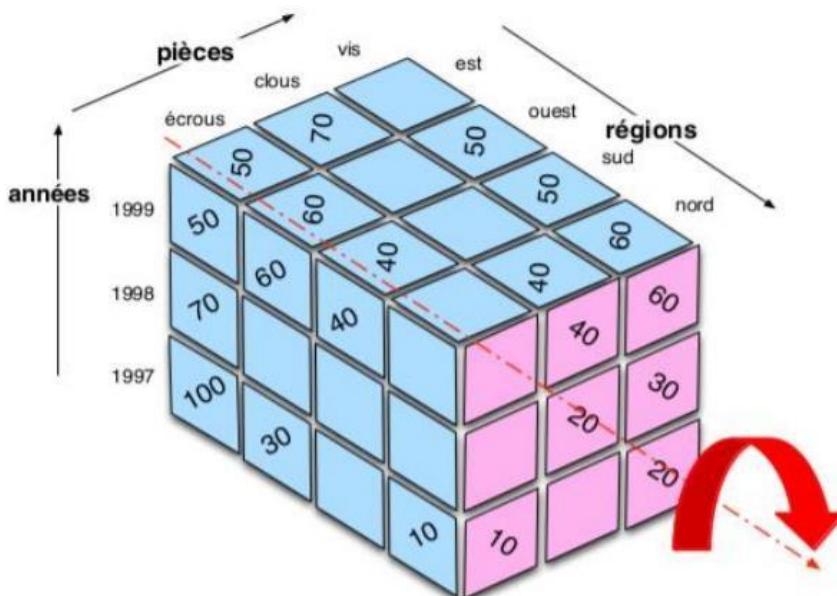


# Manipulation d'un cube OLAP

- Concerne la représentation, permet un changement de points de vue selon différentes dimension.

Exemple 1:

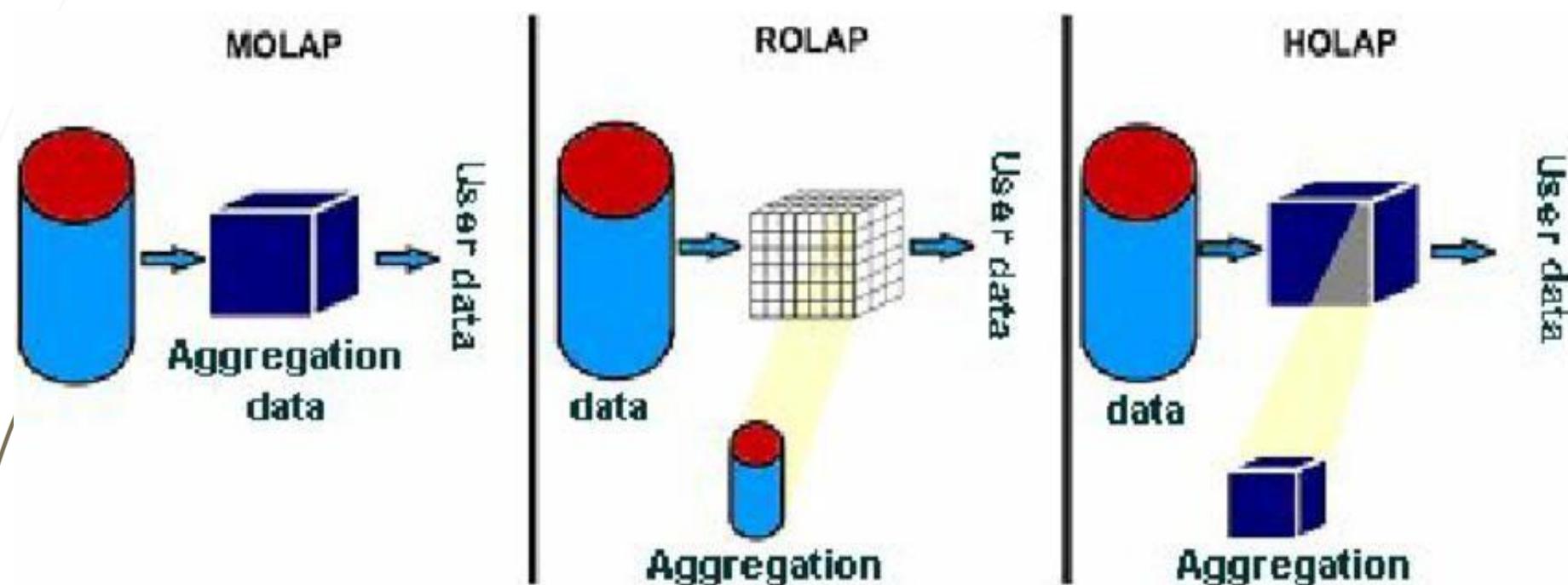
- Rotate (ou pivot):** consiste à faire une rotation à un cube autour d'un de ses axes, de manière à présenter un ensemble de faces différent.



# Types de Cubes

- **Relational OLAP (ROLAP)**
  - ◆ Données sont stockées dans un SGBD relationnel
  - ◆ Un moteur OLAP permet de simuler le comportement d'un SGBD multi-dimensionnel
- **Multidimensional OLAP (MOLAP)**
  - ◆ Structure de stockage en cube
  - ◆ Accès direct aux données dans le cube
- **Hybrid OLAP (HOLAP)**
  - ◆ Données stockées dans SGBD relationnel (données de base) + structure de stockage en cube (données agrégées)

## MOLAP vs ROLAP vs HOLAP



# Références

[Modes of data storing based on \[15\]. | Download Scientific Diagram \(researchgate.net\)](#)

[Kimball vs Inman - Data Analytics Team \(d-a-team.com\)](#)

[Concept de Business Intelligence à l'aide de conception pyramidale. Étapes du flux de traitement : sources de données, ETL, OLAP datawarehouse -- l'exploitation de données, analyse des données - 2002 Image Vectorielle Stock - Alamy \(alamyimages.fr\)](#)