

# SYSTÈMES D'INFORMATION DÉCISIONNELS

DATA WAREHOUSE

Département d'informatique

Module: Systèmes d'information décisionnels

Chargé du module: Mokeddem, S

Email: [sidahmed.mokadem@univ-mosta.dz](mailto:sidahmed.mokadem@univ-mosta.dz)



Année universitaire: 2018/2019

# SYSTÈMES TRANSACTIONNELS VS SYSTÈMES DÉCISIONNELS

## Systèmes transactionnels

- Les outils traditionnels de **gestion** et **d'exploitation** des données sont du type **transactionnel** ou **OLTP** (On-Line Transaction Processing)
- L'exploitation de données tournée vers la **saisie**, le **stockage**, la **mise à jour**, la **sécurité** et **l'intégrité** des données.
  - Transactions quotidiennes
  - Exemple: inventaires de magasins, les réservations d'hôtel, etc
  - Les données sont très détaillées
- Très souvent **plusieurs** de ces **systèmes** existent indépendamment les uns des autres

# SYSTÈMES TRANSACTIONNELS VS SYSTÈMES DÉCISIONNELS

- **Opérations** dans les systèmes transactionnels
  - Ajout
  - Suppression
  - Mise à jour des enregistrements
  - Requêtes simples
  - Interrogations et modifications fréquentes des données par de nombreux utilisateurs

# SYSTÈMES TRANSACTIONNELS VS SYSTÈMES DÉCISIONNELS

## Systèmes décisionnels

- SI capable d'agréger les données internes ou externes et de les transformer en information servant à une prise de décision rapide.
- SI capable de répondre à certains types de questions:
  - Quelles sont les ventes du produit X pendant le trimestre A de l'année B dans la région C ?
  - Comment se comporte le produit X par rapport au produit Y?
  - Quel type de client peut acheter le produit X?
  - Est-ce qu'une baisse de prix de 10% par rapport à la concurrence ferait redémarrer les ventes du produit X ?

# SYSTÈMES TRANSACTIONNELS VS SYSTÈMES DÉCISIONNELS

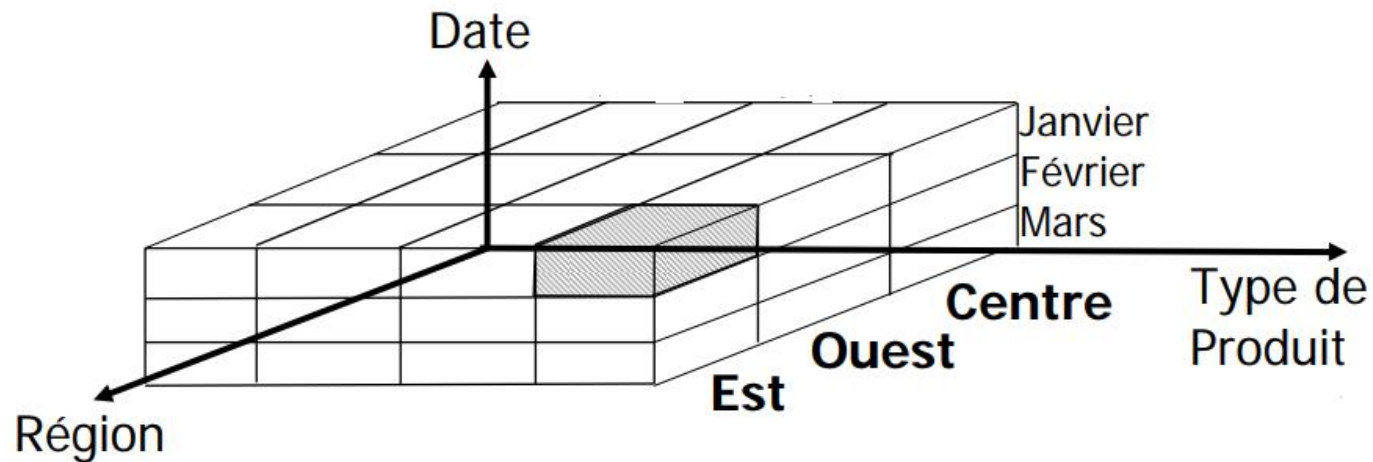
- Les questions doivent pouvoir être formulées dans le langage de l'utilisateur en fonction de son secteur d'activité:
  - Service marketing, Service économique, service relation clients...
  - La prévision des interrogations est difficile car elles sont du ressort de l'utilisateur.
  - Les questions vont varier selon les réponses obtenues:
    - Si le produit X s'est vendu moins bien que l'année précédente, il va être utile de comprendre les raisons: Détailler les ventes par région par type de magasin,...
    - Des questions ouvertes vont nécessiter la mise en place de méthodes d'extraction d'informations

# DES DONNÉES AUX DÉCISIONS

- Données :
  - Points de ventes, géographiques, démographiques, ...
- Informations :
  - I vit dans R, I est âgé de A, ...
- Connaissances :
  - Dans X%, le produit Y est vendu en même temps que le produit Z, ...
- Décisions
  - Lancer la promotion de Y & Z dans R auprès des clients plus âgés que A, ...

## OBJECTIFS (1/2)

- Transformer un SI qui avait une vocation de **production** en un SI **décisionnel**
  - Transformation des données de **production** en informations **stratégiques**
  - Gestion et visualisation des données doit être **rapide** et **intuitive**
  - visualisation **multi-dimensionnelle** des données



## OBJECTIFS (2/2)

- Analyser rapidement les données provenant de diverses sources
  - intégration de différentes BDs
- Les données:
  - Extraites
  - Groupées ensembles et organisées
  - Transformées (résumé, agrégation)



# DATA WAREHOUSE

- Ensemble de **données** destinées aux « **décideurs** », souvent une **copie** des données de **production** avec une valeur ajoutée (**agrégation**, **historique**), **intégrées** et **historisées**,
- Ensemble d'outils permettant
  - de regrouper les données
  - de nettoyer, d'intégrer les données, ...
  - de faire des requêtes, rapports, analyses
  - de faire du data mining

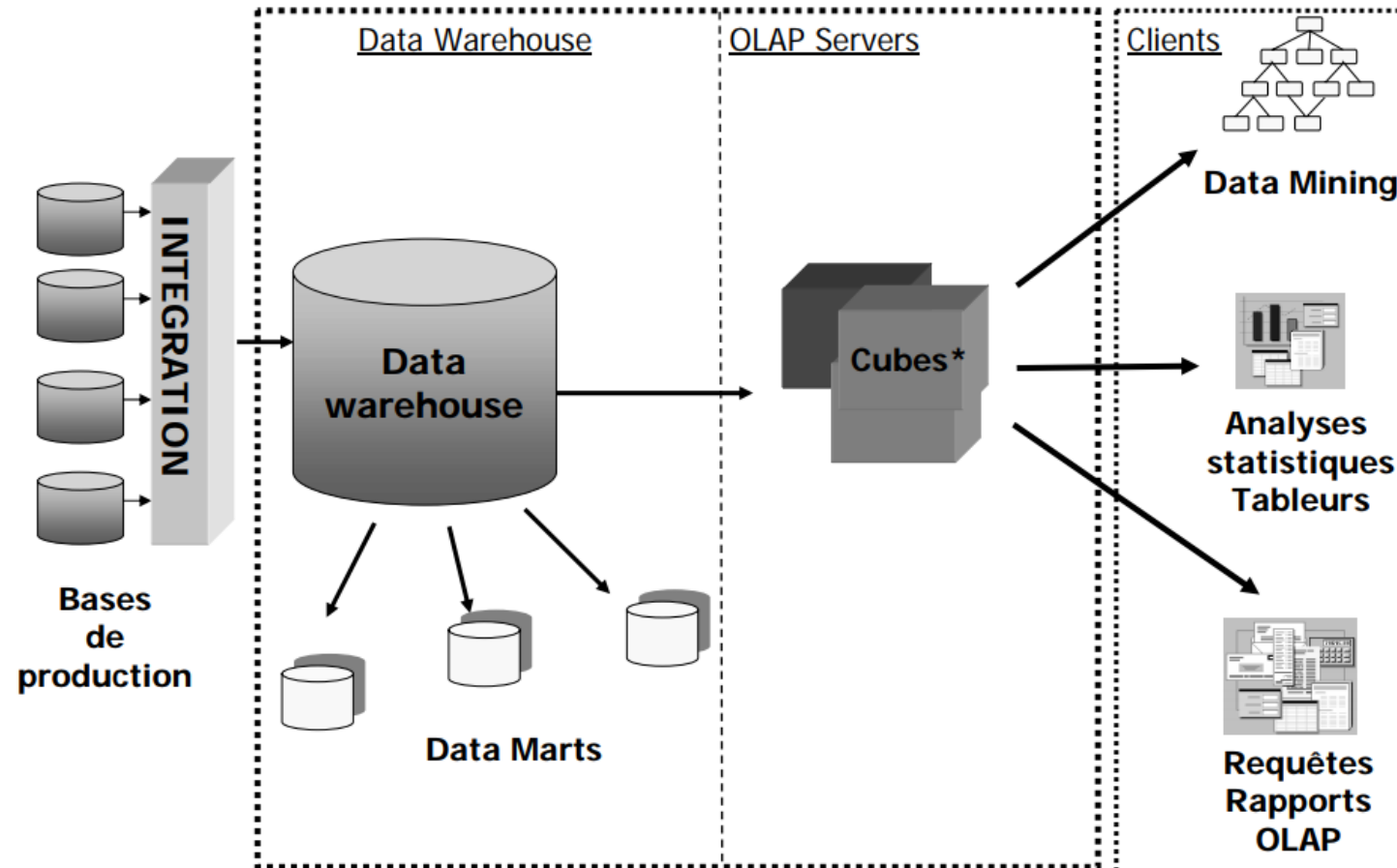
# DATA WAREHOUSE: FONCTIONS

- Récupérer des données existants dans différentes BD sources
- Stocker les données (historisées)
- Mettre à disposition les données pour :
  - Interrogation
  - Visualisation
  - Analyse

# DATA WAREHOUSE: FONCTIONS

- Récupérer des données existants dans différentes BD sources
- Stocker les données (historisées)
- Mettre à disposition les données pour :
  - Interrogation
  - Visualisation
  - Analyse

# DATA WAREHOUSE: ARCHITECTURE



# DATA WAREHOUSE: UN SGBD?

- Objectifs différents et font des traitements différents
- Stockent des données différentes
- Font l'objet de requêtes différentes
- SGBD et DW ont besoin d'une organisation différente des données
- SGBD Transactionnel et DW Decisionnel

# DATA WAREHOUSE: OBJECTIFS ET TRAITEMENTS

- Les datawarehouse sont des systèmes conçus pour l'aide à la **prise de décision**
- La plupart du temps sont utilisés en **lecture** (utilisateurs)
- Les **objectifs** principaux sont:
  - Regrouper, organiser des informations provenant de sources diverses,
  - Intégrer et stocker pour donner à l'utilisateur une vue orientée métier,
  - Retrouver et analyser l'information facilement et rapidement.

# DATA WAREHOUSE: OBJECTIFS ET TRAITEMENTS

D'après BILL Inmon: « Un DW est une collection de données orientées **sujet**, **intégrées**, non **volatiles**, **historisées**, organisées pour la prise de **décision**. »

- Orientées sujet: thèmes par activités majeures
- Intégrées: divers sources de données ;
- Non volatiles: ne pas supprimer les données du DW ;
- Historisées: trace des données, suivre l'évolution des indicateurs

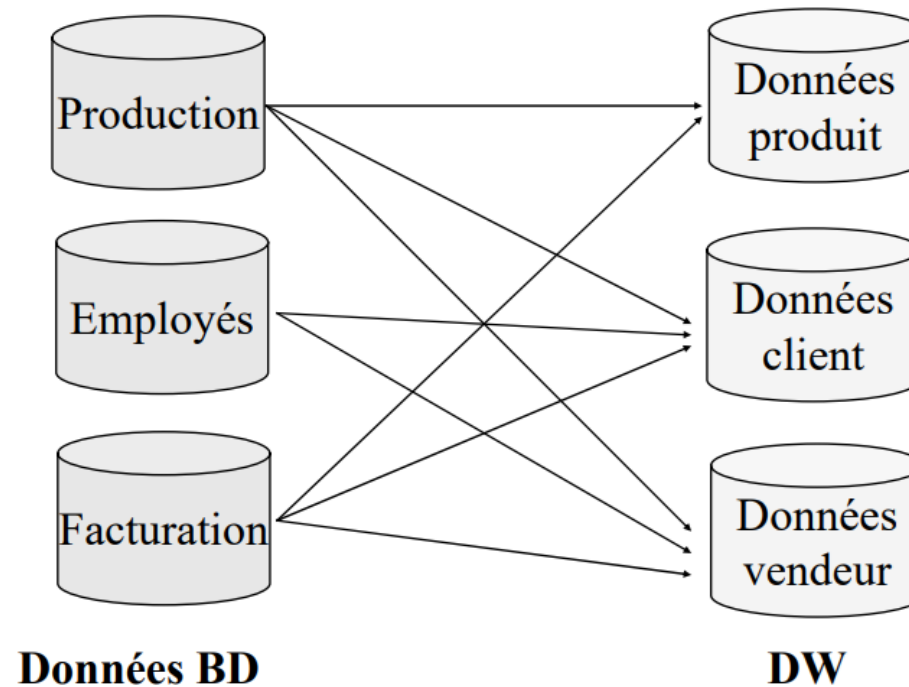
# DATA WAREHOUSE: OBJECTIFS ET TRAITEMENTS

D'après BILL Inmon: « Un DW est une collection de données orientées **sujet**, **intégrées**, non **volatiles**, **historisées**, organisées pour la prise de **décision**. »

- Orientées sujet: thèmes par activités majeures
- Intégrées: divers sources de données ;
- Non volatiles: ne pas supprimer les données du DW ;
- Historisées: trace des données, suivre l'évolution des indicateurs

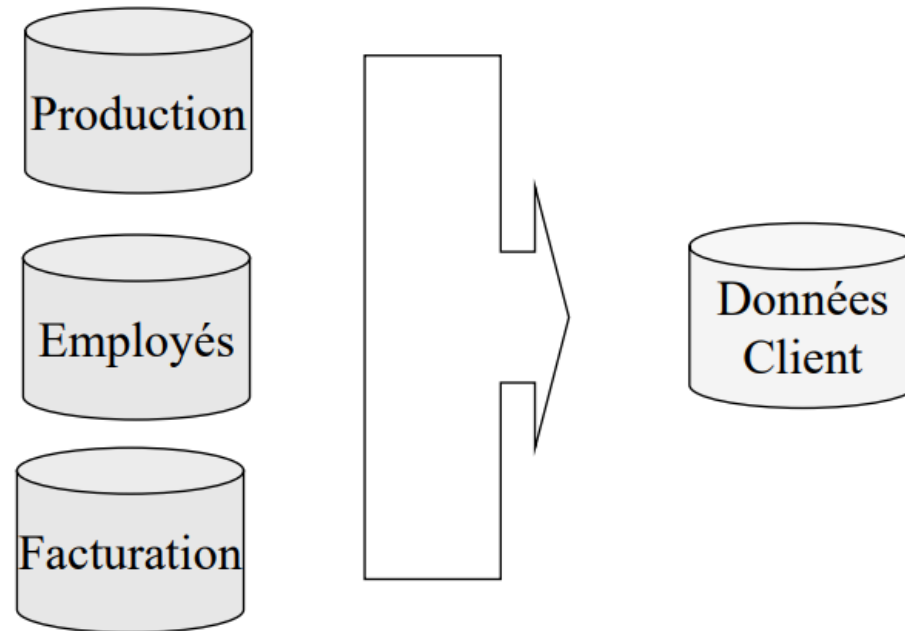


# DATA WAREHOUSE: OBJECTIFS ET TRAITEMENTS

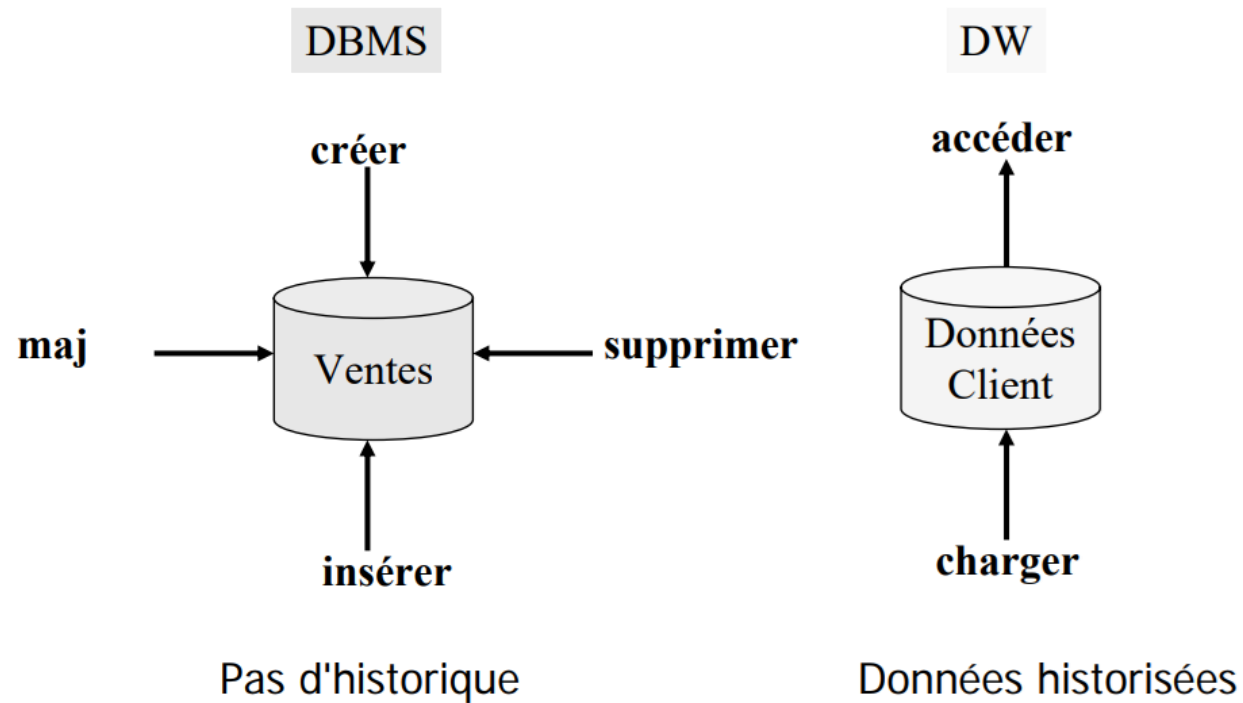


# DATA WAREHOUSE: OBJECTIFS ET TRAITEMENTS

- Techniques **d'intégration** des données
- Techniques de **nettoyage**: Cohérence entre les différentes sources des noms, unités de mesures, etc



# DATA WAREHOUSE: OBJECTIFS ET TRAITEMENTS



# DATA WAREHOUSE: OBJECTIFS ET TRAITEMENTS

produit	région	vente	date	vendeur
écrou	Est	50	01012004	X
écrou	Ouest	60	12122003	X
écrou	Centre	110	01112003	Y
vis	Est	70	01042004	Y
vis	Ouest	80	10022004	Z
vis	Centre	90	29032004	Y
boulon	Est	120	05052004	X
boulon	Ouest	10	24042004	Z
boulon	Centre	20	11022004	Y
joint	Est	50	01032004	X
joint	Ouest	40	01102003	Y
joint	Centre	70	01012003	Z

produit	prix	fournisseur
écrou	44	CC
vis	2	DD
boulon	3	VV
joint	1	BB

fournisseur	ville
...	...

- OLTP: Requêtes simples
  - les ventes de X
  - les ventes de X à quel prix de quel fournisseur
- OLAP: besoin de données agrégées, synthétisées:
  - nombre de ventes par vendeur, par région, par mois,
  - nombre de ventes par vendeur, par fournisseur, par mois,

# DATA WAREHOUSE: OBJECTIFS ET TRAITEMENTS

produit	région	vente	date	vendeur
écrou	Est	50	01012004	X
écrou	Ouest	60	12122003	X
écrou	Centre	110	01112003	Y
vis	Est	70	01042004	Y
vis	Ouest	80	10022004	Z
vis	Centre	90	29032004	Y
boulon	Est	120	05052004	X
boulon	Ouest	10	24042004	Z
boulon	Centre	20	11022004	Y
joint	Est	50	01032004	X
joint	Ouest	40	01102003	Y
joint	Centre	70	01012003	Z

produit	prix	fournisseur
écrou	44	CC
vis	2	DD
boulon	3	VV
joint	1	BB

fournisseur	ville
...	...

- OLTP: Requêtes simples
  - les ventes de X
  - les ventes de X à quel prix de quel fournisseur
- OLAP: besoin de données agrégées, synthétisées:
  - nombre de ventes par vendeur, par région, par mois,
  - nombre de ventes par vendeur, par fournisseur, par mois,

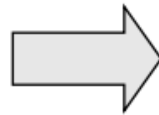
# DATA WAREHOUSE: OBJECTIFS ET TRAITEMENTS

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

# DATA WAREHOUSE: REPRÉSENTATION

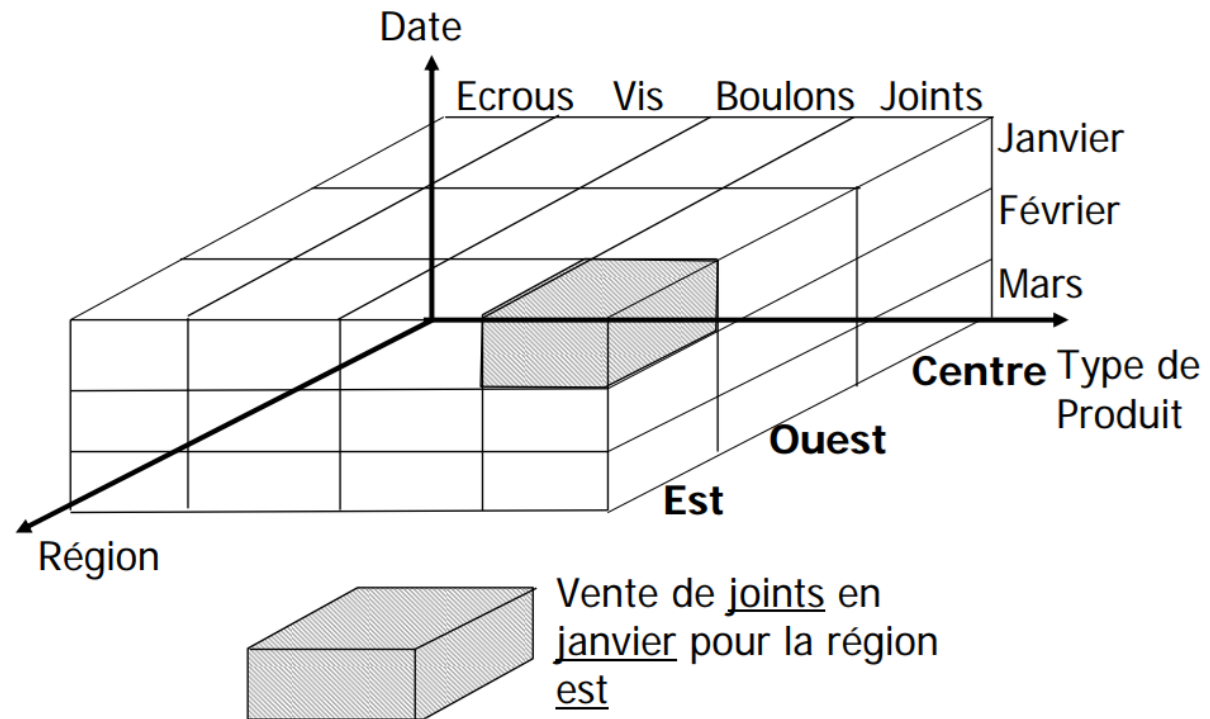
- représentation des données sous forme multidimensionnelle : ‘Cube’

produit	région	vente
écrou	Est	50
écrou	Ouest	60
écrou	Centre	110
vis	Est	70
vis	Ouest	80
vis	Centre	90
boulon	Est	120
boulon	Ouest	10
boulon	Centre	20
joint	Est	50
joint	Ouest	40
joint	Centre	70



	Est	Ouest	Centre
<b>écrous</b>	50	60	110
<b>vis</b>	70	80	90
<b>boulons</b>	120	10	20
<b>joints</b>	50	40	70

# DATA WAREHOUSE: REPRÉSENTATION





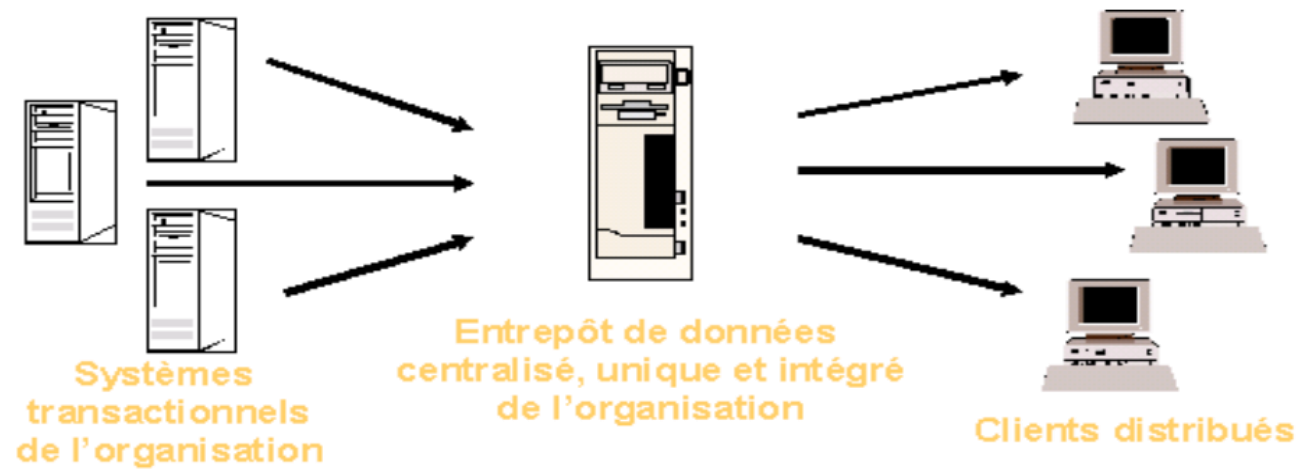
# DATA WAREHOUSE: DATA MART

- C'est un sous-ensemble de données dérivées du DW ciblé sur un sujet unique
- Caractéristiques
  - Orienté vers un sujet unique
    - Ex: comportement de la clientèle
  - Données fortement agrégées
    - Le DW joue le rôle de source et d'historique pour le Datamart
  - Organisation multidimensionnelle (cubique)
    - Dont l'une des dimensions indique souvent le temps
  - Lien dynamique avec le DW
    - Association entre valeur agrégée et valeur détaillée

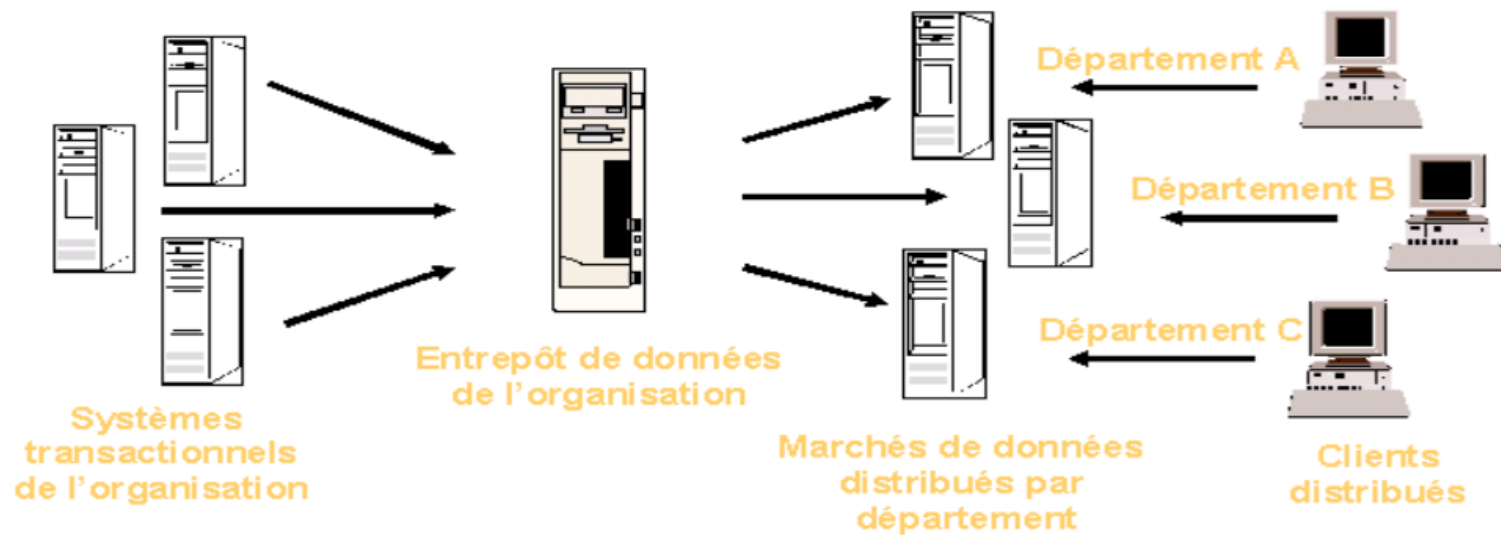
# DATA WAREHOUSE: LES DOMAINES D'APPLICATION

- La gestion de la relation client (CRM) est l'un des premiers champs d'application de la Business Intelligence.
- Le contrôle de gestion pour l'analyse des coûts, l'analyse de la rentabilité, l'élaboration budgétaire, les indicateurs de performance...
- La direction marketing pour le ciblage, le pilotage de gamme, les applications de géomarketing, de fidélisation clients...
- La direction commerciale pour le pilotage des réseaux, les prévisions des ventes, l'optimisation des territoires...
- Les ressources humaines pour la gestion des carrières,
- La direction de la production pour l'analyse qualité, la prévision des stocks, la gestion des flux, la fiabilité industrielle...
- La direction générale pour les tableaux de bord, indicateurs de pilotage, gestion d'alertes...

# DATA WAREHOUSE: ARCHITECTURE D'ACCÈS



# DATA WAREHOUSE: ARCHITECTURE D'ACCÈS



# DATA WAREHOUSE: IMPLÉMENTATION

- Relational OLAP (ROLAP)
  - Données sont stockées dans un SGBD relationnel
  - Un moteur OLAP permet de simuler le comportement d'un SGBD multi-dimensionnel
    - Exemple: Teradata, Netezza
- Multidimensional OLAP (MOLAP)
  - Structure de stockage en cube
  - Accès direct aux données dans le cube
    - Par exemple : Oracle Express, SQL Server Analysis Services
- Hybrid OLAP (HOLAP)
  - Données stockées dans SGBD relationnel (données de base)
  - Structure de stockage en cube (données agrégées)
  - Exemple: Oracle OLAP, Microsoft Analysis Services.

# DATA WAREHOUSE: ROLAP

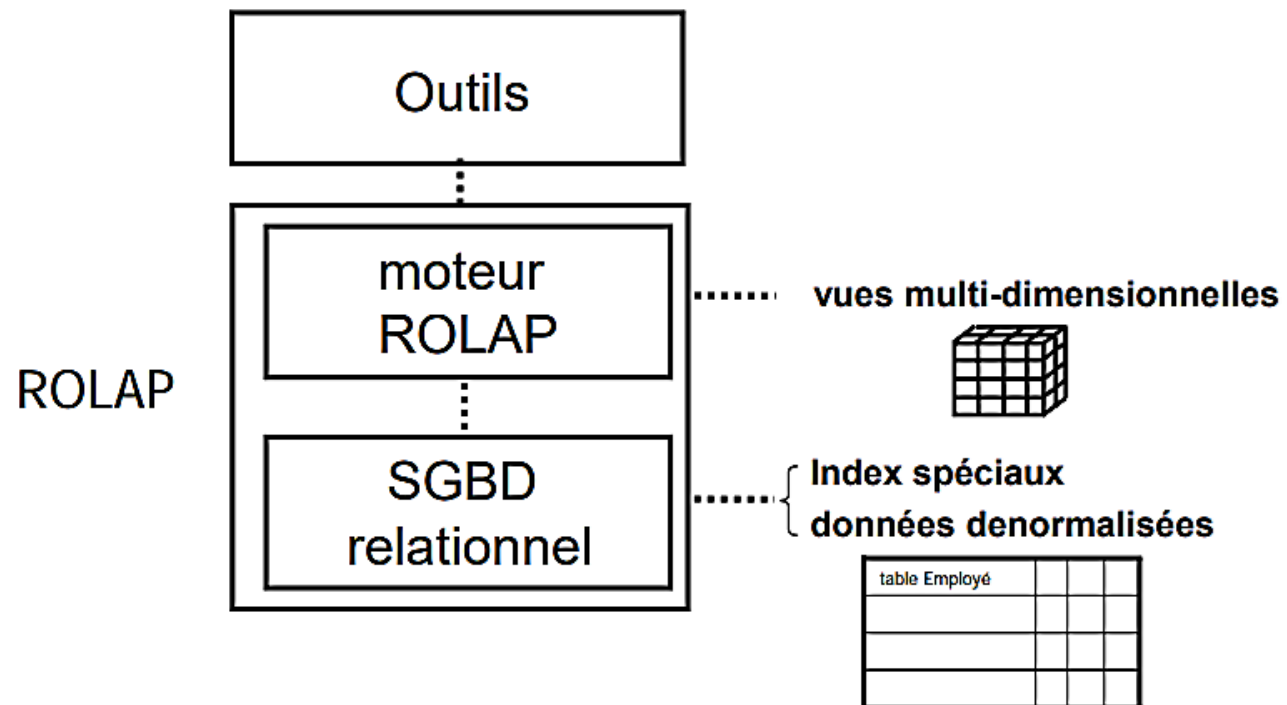
## Idée

- Données stockées en relationnel.
- La conception du schéma est particulière: schéma en étoile, schéma en flocon
- Des vues (matérialisées) sont utilisées pour la représentation multidimensionnelle
- Les requêtes OLAP (slice, rollup...) sont traduites en SQL.
- Administration (tuning) particulier de la base

## Avantages/inconvénients

- Souplesse, évolution facile, permet de stocker de gros volumes
- Mais peu efficace pour les calculs complexes

# DATA WAREHOUSE: ROLAP



# DATA WAREHOUSE: MOLAP

## Idée

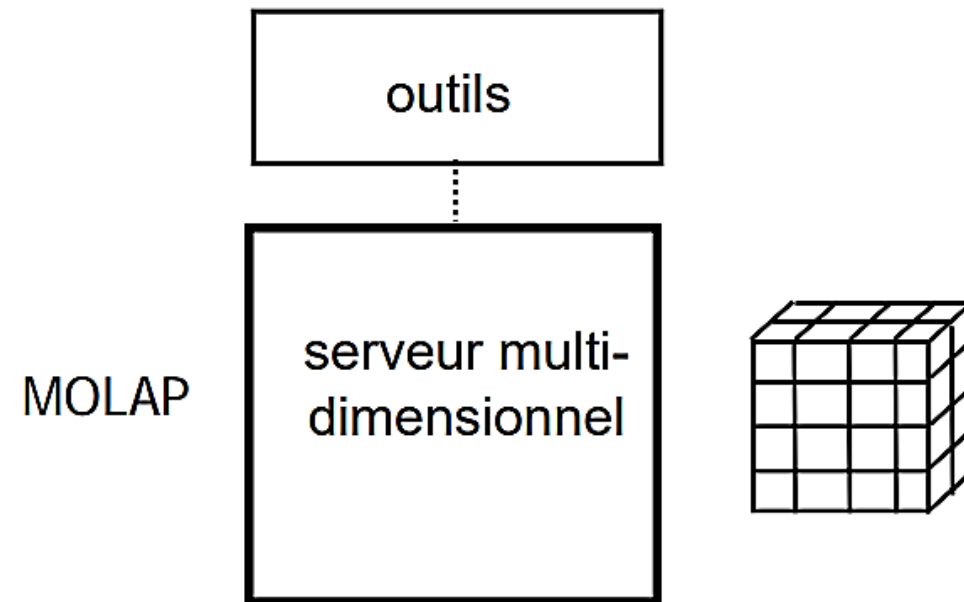
- Modélisation directe du cube
- Ces cubes sont implémentés comme des matrices à plusieurs dimensions
- Le cube est indexé sur ses dimensions

## Avantages/inconvénients

- Rapide
- formats propriétaires
- ne supporte pas de très gros volumes de données



# DATA WAREHOUSE: MOLAP

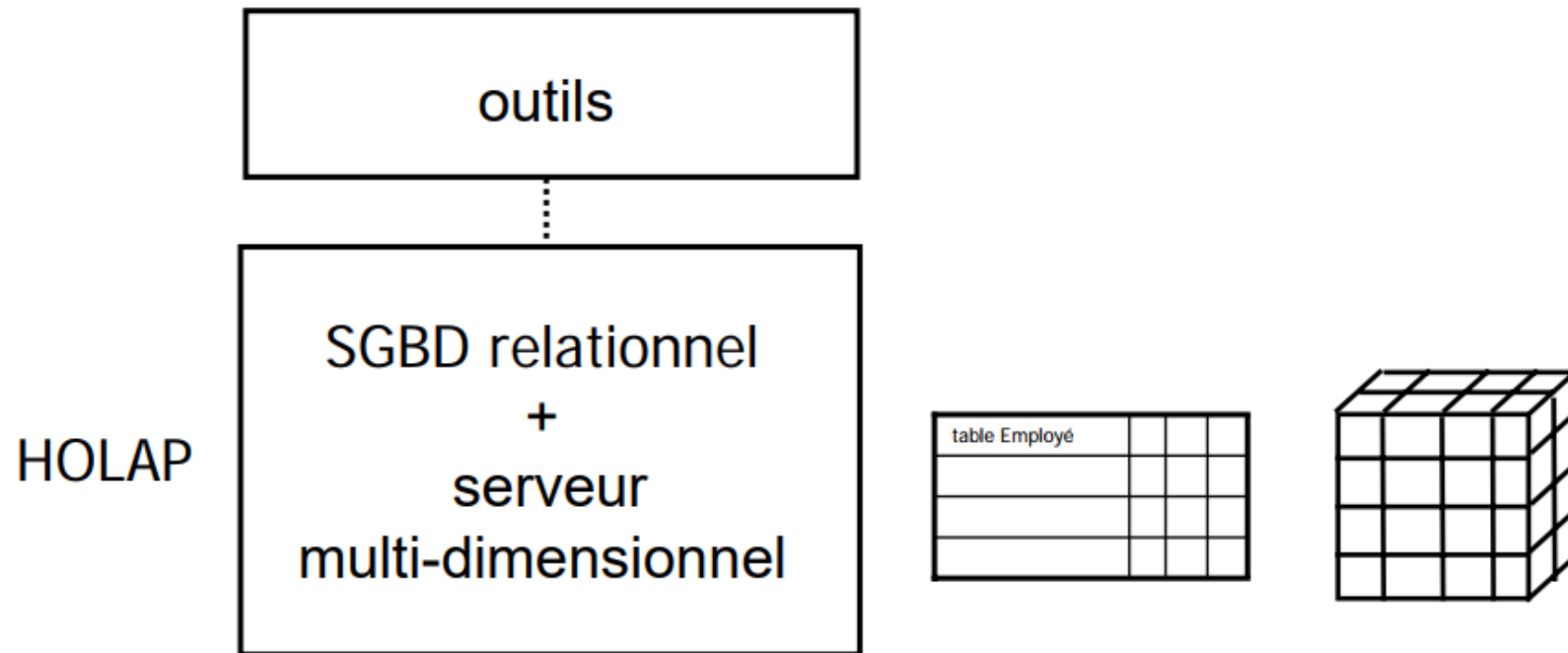


# DATA WAREHOUSE: HOLAP

## Idée

- MOLAP + ROLAP
- Données stockées dans des tables relationnelles
- Données agrégées stockées dans des cubes
- Les requêtes vont chercher les données dans les tables et les cubes

# DATA WAREHOUSE: HOLAP



# DATA WAREHOUSE: MODÉLISATION

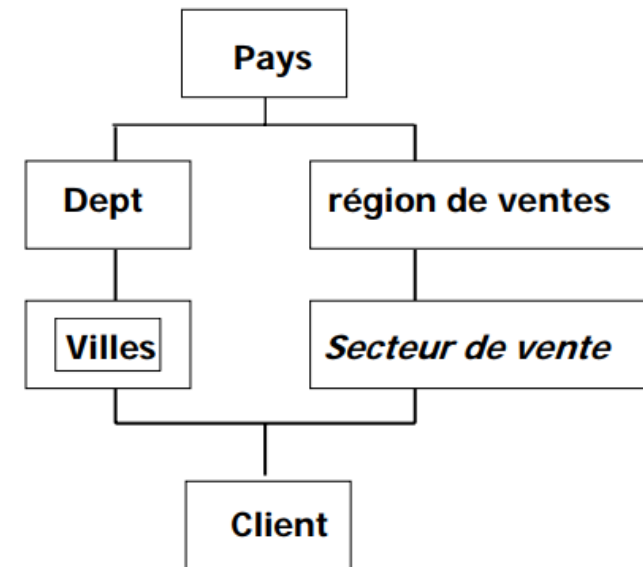
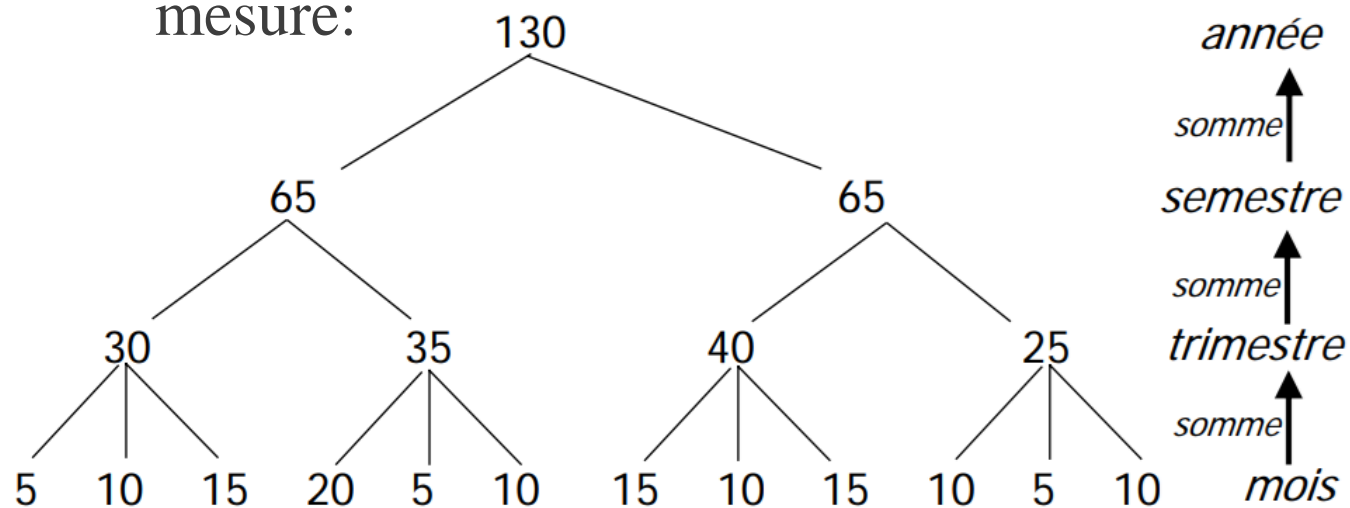
- Modélisation conceptuelle BD : entité et relation
- Modélisation de DW : dimension et mesure
- Les mesures sont les valeurs numériques que l'on compare
  - (ex : montant\_ventes, qte\_vendue)
  - Ces valeurs sont le résultat d'une opération d'agrégation des données
- Les dimensions sont les points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

# DATA WAREHOUSE: DIMENSION

- Dimension = liste d'éléments
- Dimension contient des membres organisés en hiérarchie
  - Chacun des membres appartient à un niveau hiérarchique (ou niveau de granularité) particulier
- Granularité d'une dimension : nombre de niveaux hiérarchiques
  - Temps : année – semestre – trimestre - mois

# DATA WAREHOUSE: DIMENSION

- Les axes de dimension doivent fournir des règles de calcul d'agrégat pour chaque mesure:



# DATA WAREHOUSE: MANIPULATION DE CUBE

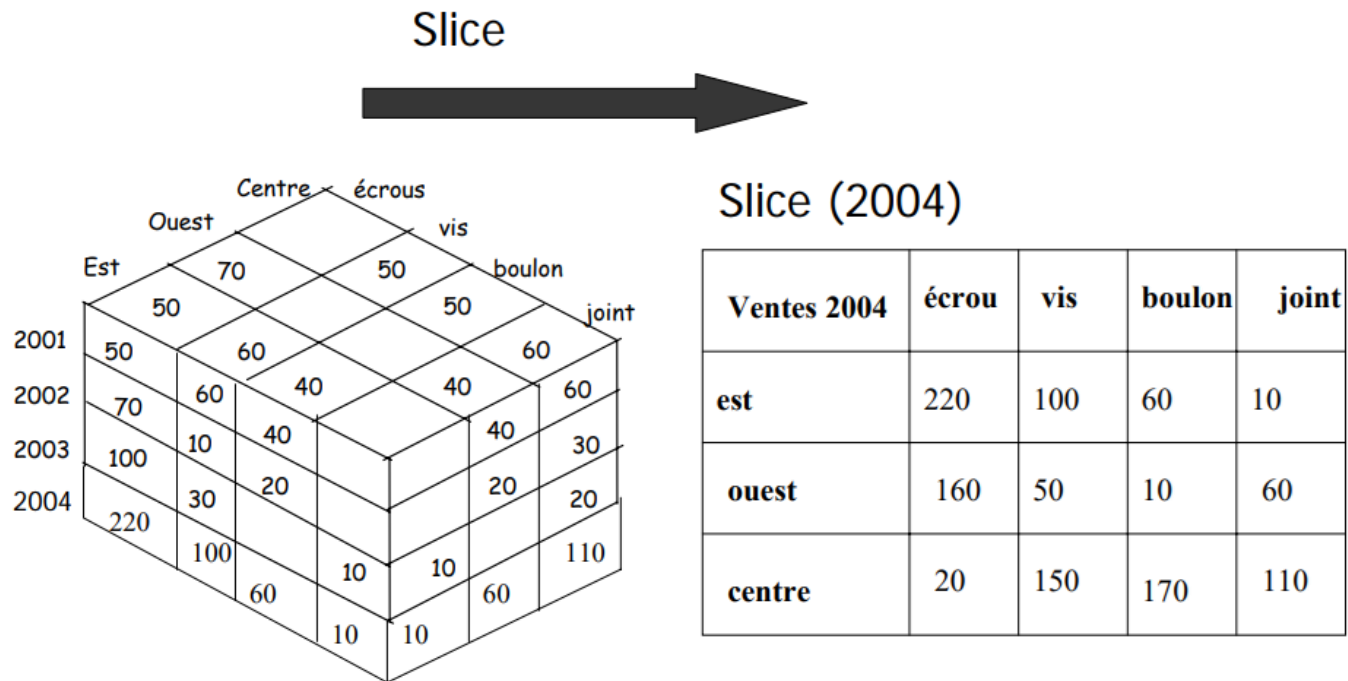
- Opérateurs appliqués sur le cube sont **algébriques** (le résultat est un autre cube) et peuvent être combinés
  - **Slicing & Dicing** (extraction)
  - Changement de la granularité d'une dimension
    - **Roll up** (agrégation d'une dimension => résumé)
    - **Drill down** (plus détaillées)

# DATA WAREHOUSE: SLICING & DICING

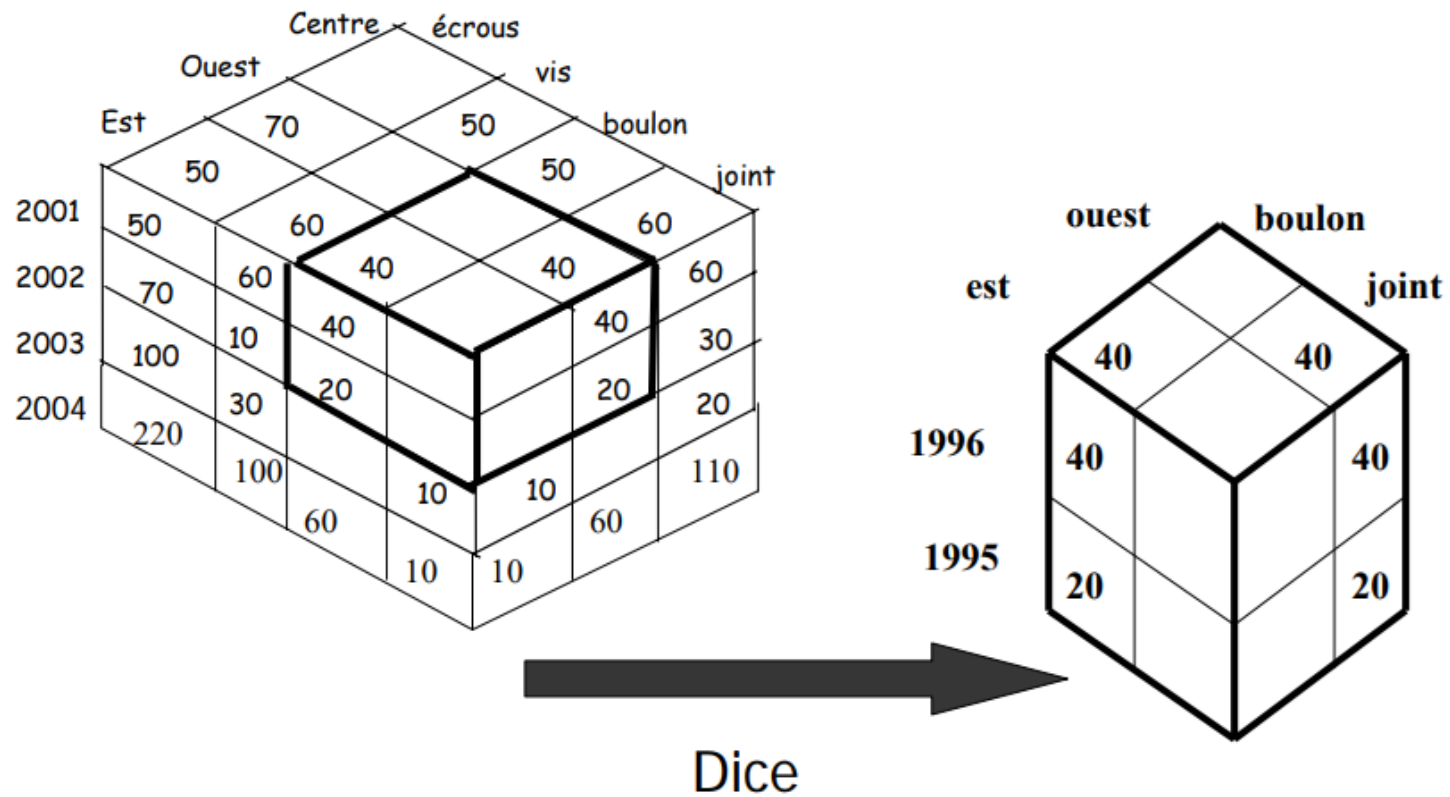
- Slicing: Sélection de tranches du cube par des prédicats selon une dimension
  - filtrer une dimension selon une valeur
    - Exemple: Slice (2004) : on ne retient que la partie du cube qui correspond à cette date
- Dicing: extraction d'un sous-cube



# DATA WAREHOUSE: SLICING



# DATA WAREHOUSE: DICING



# DATA WAREHOUSE: OPÉRATIONS LIÉES À LA GRANULARITÉ

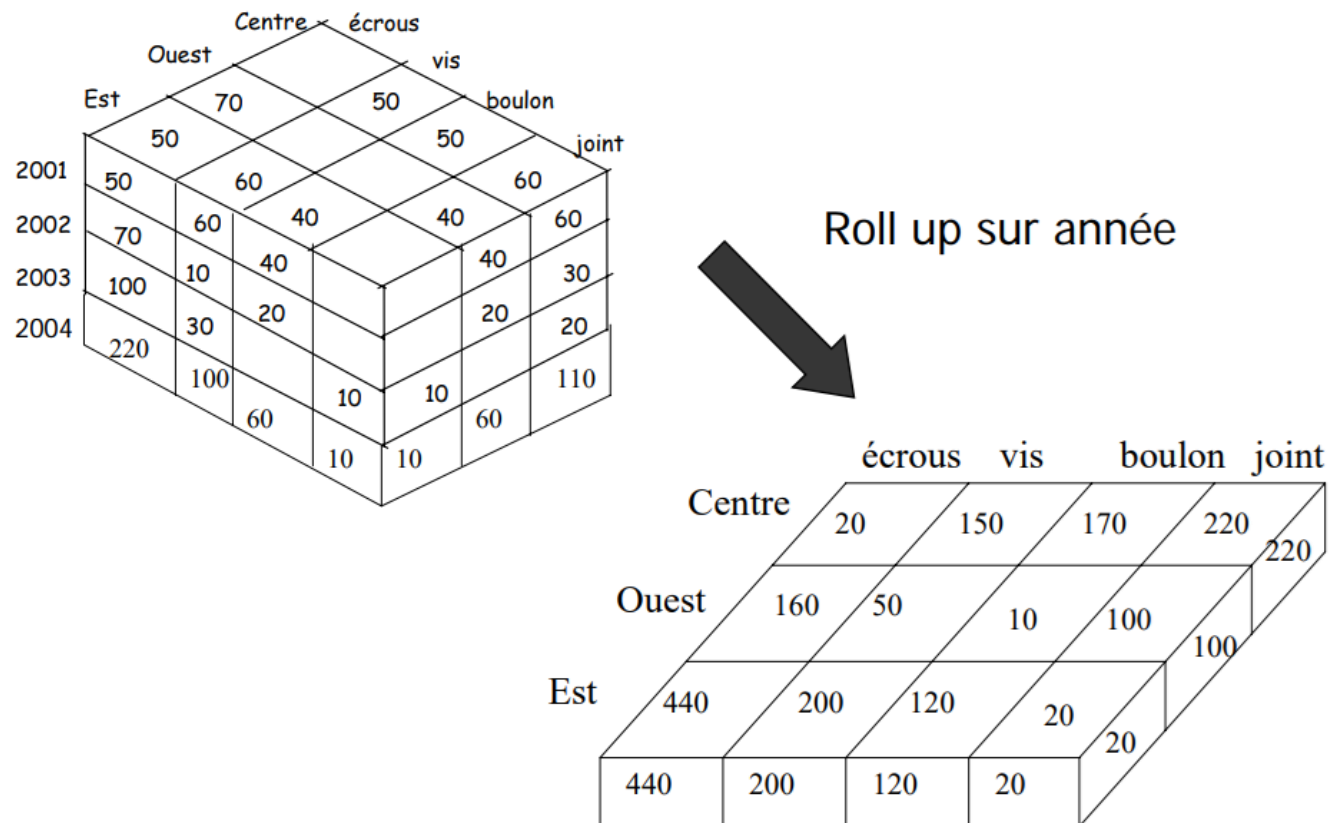
## ■ **Roll-up**

- Représenter les données du cube à un niveau de granularité supérieur conformément à la hiérarchie définie sur la dimension.
  - Utilisation de la fonction d'agrégation (somme, moyenne, etc) spécifiée pour la mesure et la dimension

## ■ **Drill-down:**

- Représenter les données du cube à un niveau de granularité de niveau inférieur, donc sous une forme plus détaillée.

# DATA WAREHOUSE: OPÉRATIONS LIÉES À LA GRANULARITÉ



# DATA WAREHOUSE: OPÉRATIONS LIÉES À LA GRANULARITÉ

				Oran	Mosta	Alger	Annaba	Bechar	Setif	Tlemcen	Chlef	Relizane	Constantine
	joint	30			30				30	30		30	
	boulon	20				10		20	30	20	50	10	10
	vis	10	30		30					40		70	
	écrous	30	20		30	40		20			10	10	
2001		30			30							10	10
2002		30	20		20		30	30	20	50	10		10
2003		10	10	20		60	20	40	20	30	10	20	
2004		10	20	10	10	70		40	10		10	10	

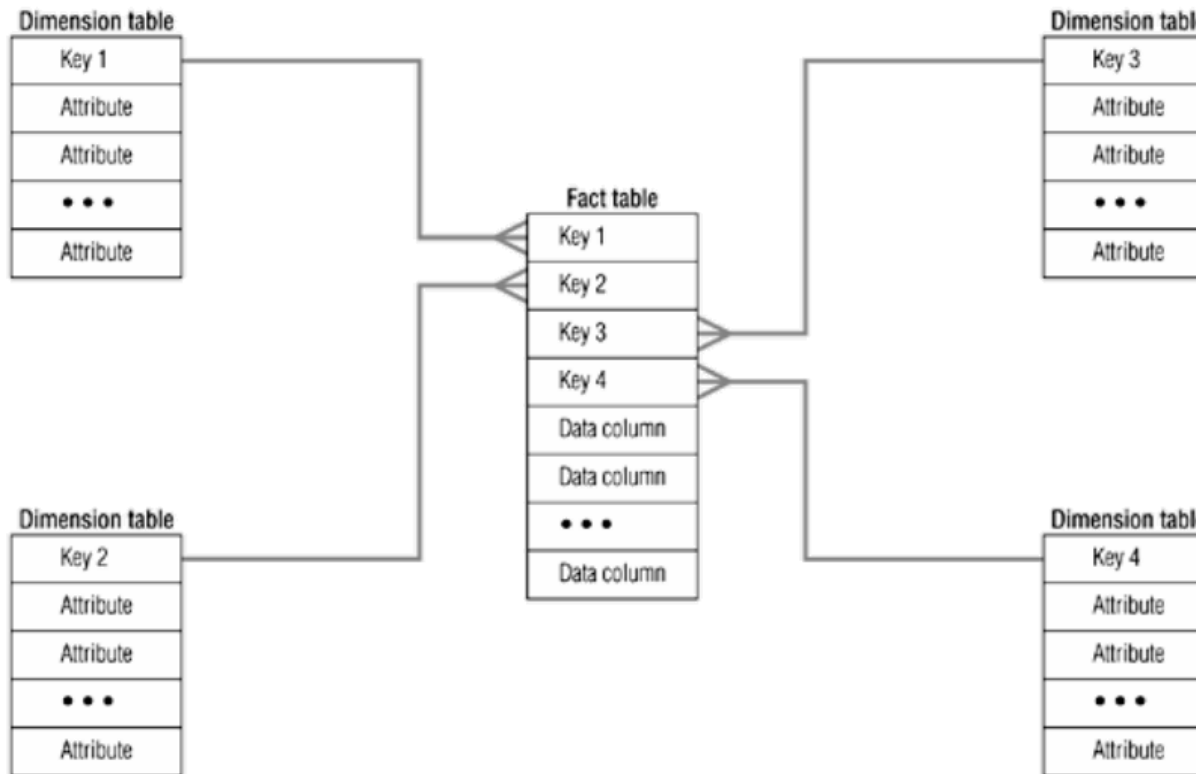
# DATA WAREHOUSE: FAITS

- Un fait représente la valeur d'une mesure, mesurée ou calculée, selon un membre de chacune des dimensions
  - Exemple : «250 000 DA » est un fait qui exprime la valeur de la mesure « coût des travaux » pour le membre «2017» du niveau année de la dimension « temps » et le membre « Mostaganem » 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

# DATA WAREHOUSE: LE MODÈLE EN ÉTOILE

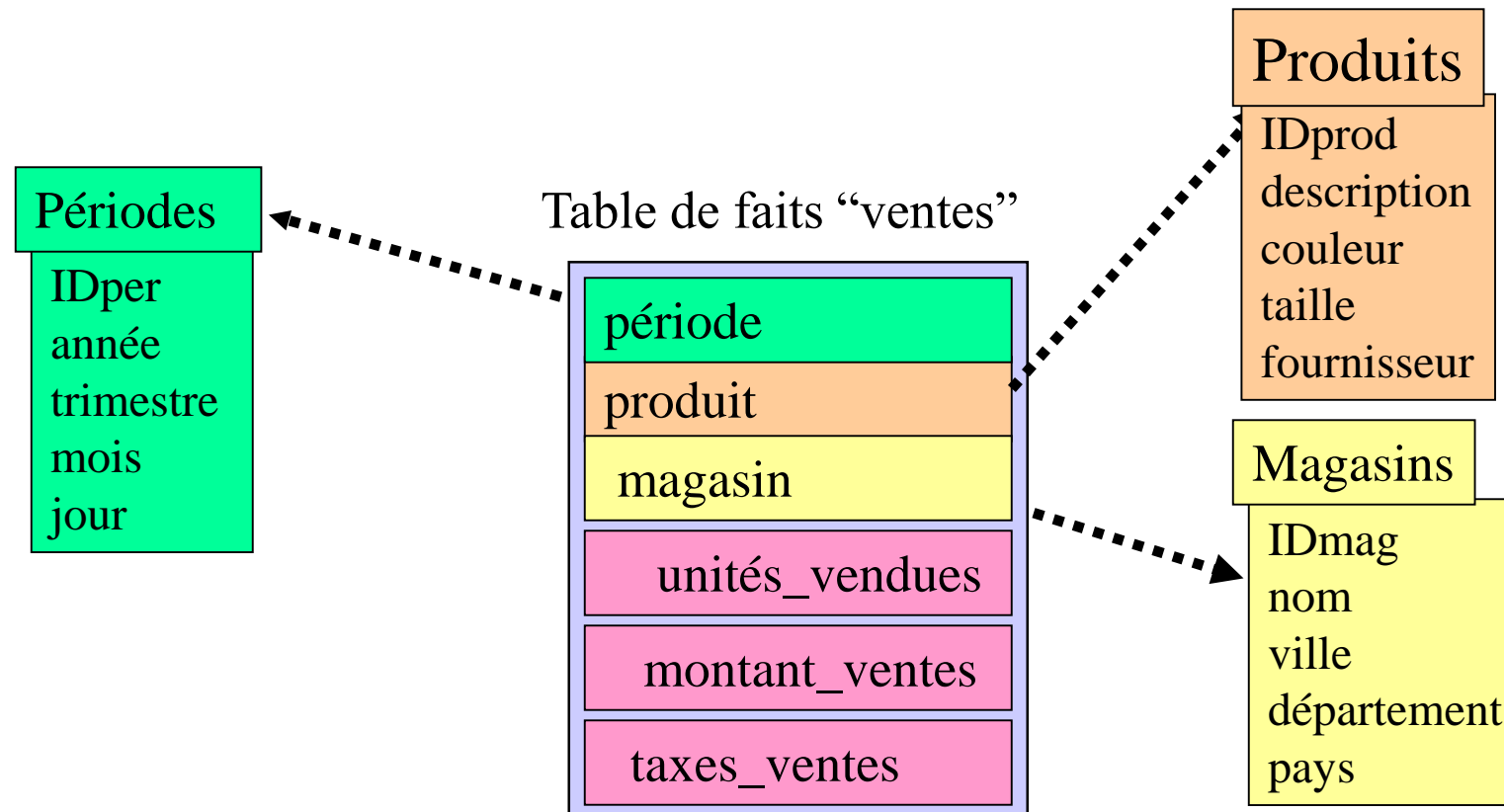
- Une (ou plusieurs) table(s) de faits comprenant une ou plusieurs mesures.
- Plusieurs tables de dimension dénormalisées: descripteurs des dimensions.
- Les tables de dimension n'ont pas de lien entre elles

# DATA WAREHOUSE: LE MODÈLE EN ÉTOILE

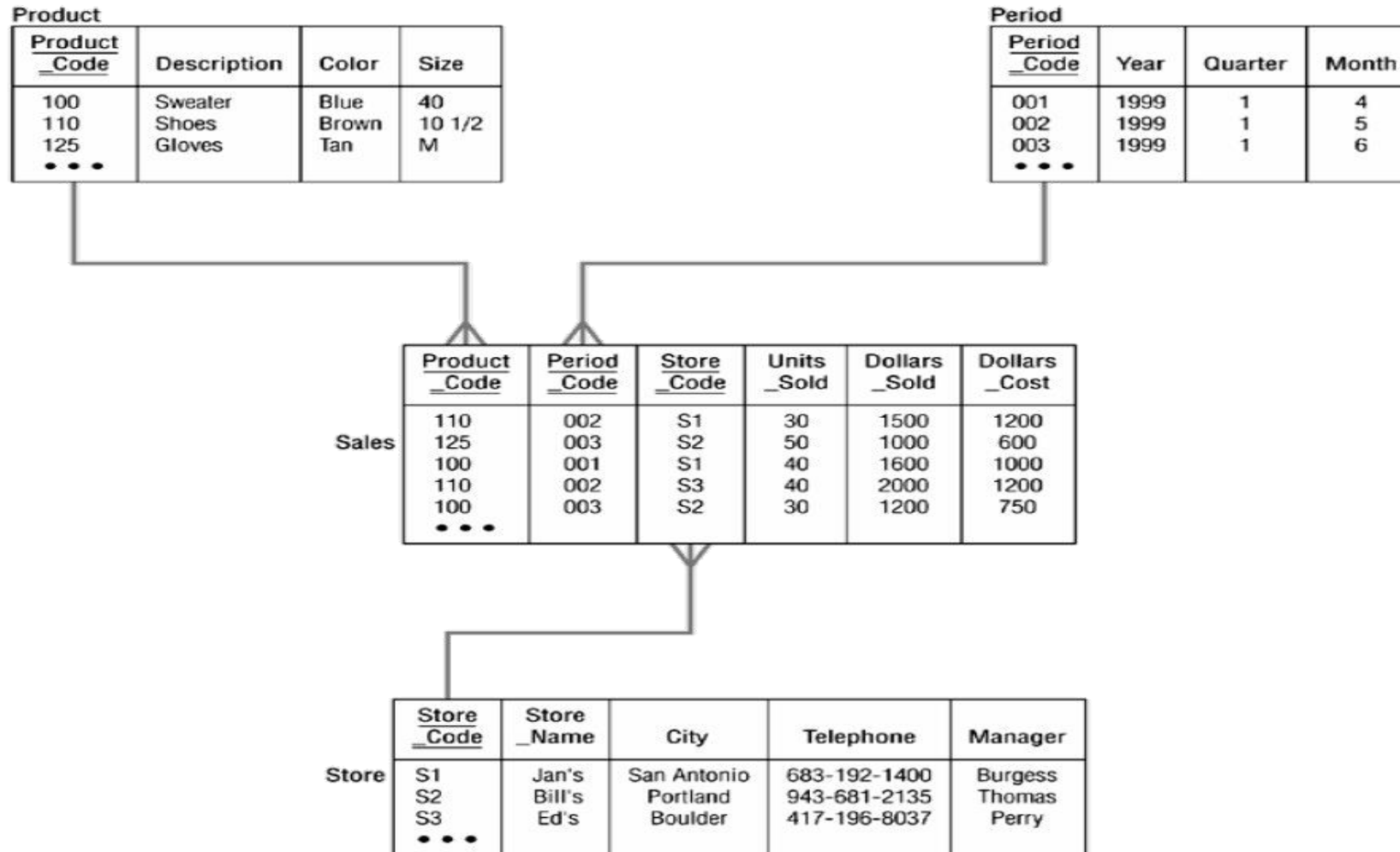




# DATA WAREHOUSE: LE MODÈLE EN ÉTOILE



# DATA WAREHOUSE: LE MODÈLE EN ÉTOILE



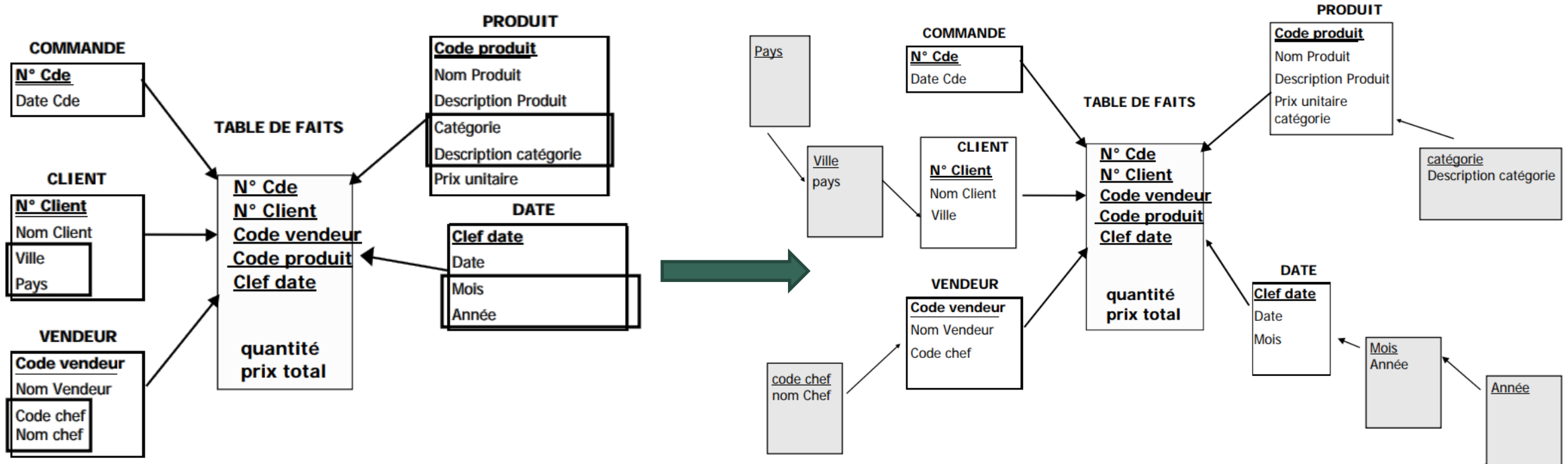
# DATA WAREHOUSE: LE MODÈLE EN ÉTOILE

- Avantages
  - Facilité de navigation
  - 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
  - Alimentation complexe.

# DATA WAREHOUSE: LE MODÈLE EN FLOCON

- 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

# DATA WAREHOUSE: LE MODÈLE EN FLOCON



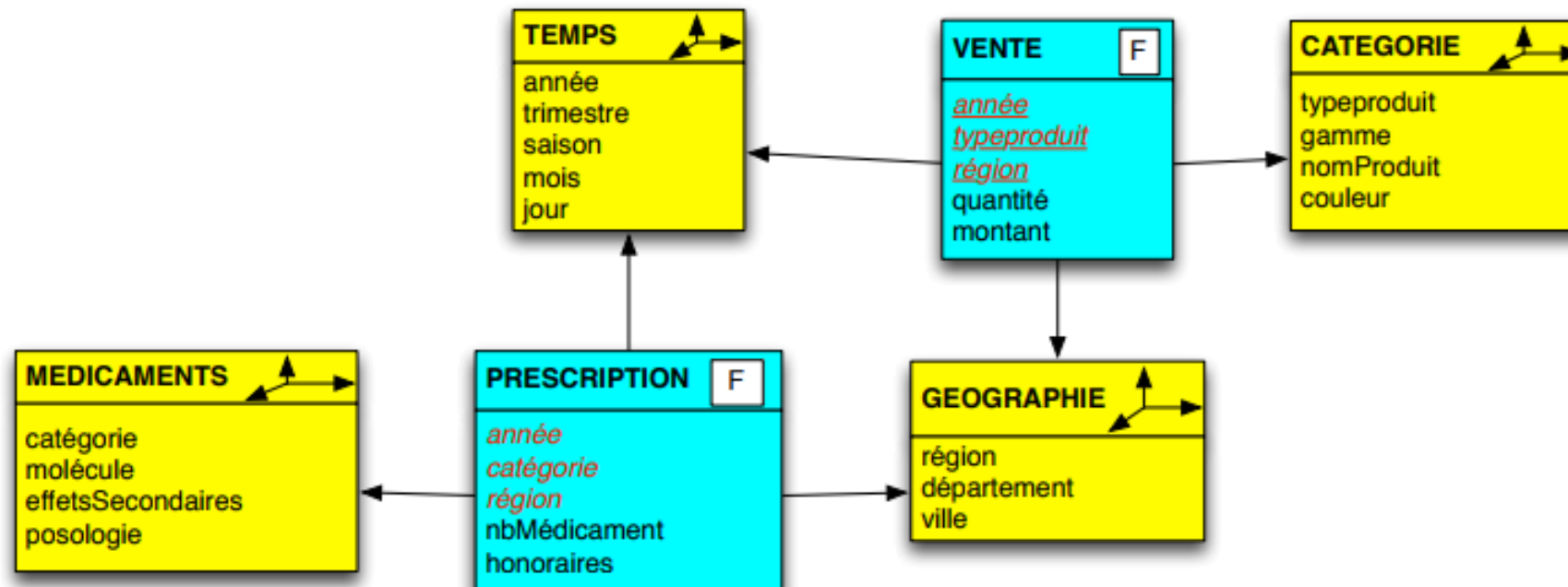
# DATA WAREHOUSE: LE MODÈLE EN FLOCON

- Modèle en étoile + normalisation des dimensions
- Lorsque les tables sont trop volumineuses
  - Avantages
    - réduction du volume,
  - Inconvénients
    - navigation difficile
    - nombreuses jointures

# DATA WAREHOUSE: LE MODÈLE EN CONSTELLATION

- La modélisation en constellation consiste à fusionner plusieurs modèles en étoile qui utilisent des dimensions communes.
  - comprend en conséquence **plusieurs faits** et des **dimensions communes** ou non
- Un modèle en constellation comprend donc plusieurs tables de faits et des tables de dimensions communes ou non à ces tables de faits.

# DATA WAREHOUSE: LE MODÈLE EN CONSTELLATION



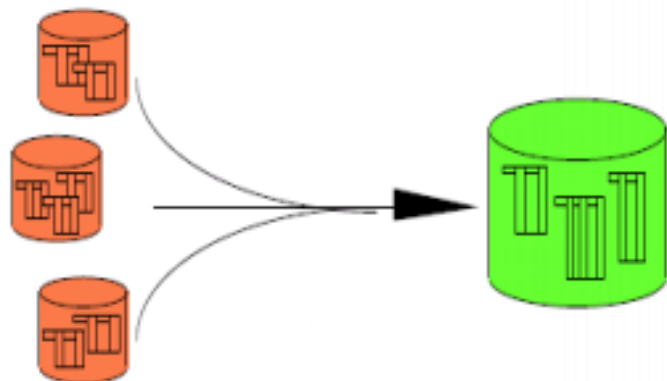


# DATA WAREHOUSE: CONCEPTION

- Conception
  - Définir la finalité du DW : Piloter quelle activité de l'entreprise ?
  - Définition du modèle de données (modèle en étoile/flocon ou cubes)
- Acquisition des données
  - Déterminer et recenser les données à entreposer: recherche des données dans les sources de l'entreprise
  - Nettoyage des données
  - Démarches d'alimentation
- Définir les aspects techniques de la réalisation
- Définir les modes de restitution, ...
- Stratégies d'administration, évolution, maintenance

# PROCESSUS D'ALIMENTATION D'UN DW : ACQUISITION DES DONNÉES (ETL)

- Extraction (Extract)
- Transformation (Transform)
- Chargement (Loading)



TACHES	SUPPORT
Extraction	accès aux différentes sources
Nettoyage	recherche et résolution des inconsistances dans les sources
Transformation	entre différents formats, langages, etc.
Chargement	des données dans l'entrepôt
Réplication	des sources dans l'entrepôt
Analyse	Ex : détection de valeurs non valides ou inattendues
Transfert de données haut débit	pour les très grands entrepôts
Test de qualité	Ex : pour correction et complétude
Analyse des méta données	aide à la conception

# DATA WAREHOUSE: ACQUISITION DES DONNÉES (ETL)

## Extraction (Extract)

- Depuis les bases sources
- Périodique et Répétée
- Dater ou marquer les données envoyées
- Utilisation d'interfaces comme ODB, OCI, JDBC.
- Difficulté:
  - Ne pas perturber les applications OLTP
- Différentes techniques d'extraction:
  - Méthode **Push**: Le système opérationnel qui au fil des transactions alimente le DW
  - Méthode **Pull** : Le système décisionnel cherche périodiquement les données dans les bases de production

# DATA WAREHOUSE: ACQUISITION DES DONNÉES (ETL)

## Transformation

- C'est une suite d'opérations qui a pour but de rendre les données cibles homogènes et puissent être traitées de façon cohérente.
- Convertir / uniformiser les noms des attributs
  - fautes de frappe
  - différents formats dans une même colonne
- Uniformiser les valeurs d 'attributs
- Nettoyer ( Valeurs manquantes, aberrantes...)

# DATA WAREHOUSE: ACQUISITION DES DONNÉES (ETL)

## Transformation

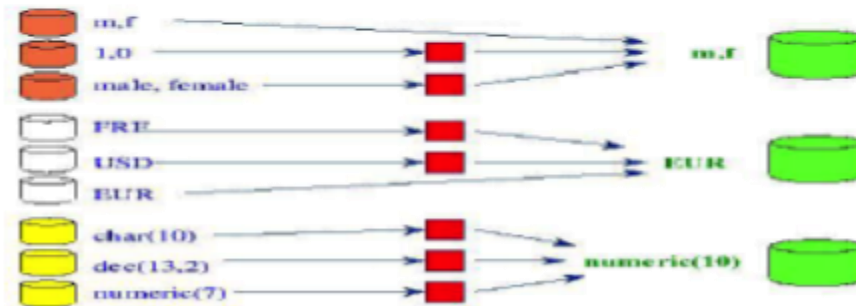
- persistance de données obsolètes
- confrontation de données sémantiquement équivalentes mais syntaxiquement différentes
- fonctions de normalisation
- fonctions de conversion
- usage de dictionnaires de synonymes ou d'abréviations

# DATA WAREHOUSE: ACQUISITION DES DONNÉES (ETL)

*Définition de table de règles :*

valeur source	remplacé par	Valeur cible
Mr		M
monsieur		M
Masculin		M
M		M
Msieur		M

*Exemple de conversions :*



**nettoyage = jointure + projection**

# DATA WAREHOUSE: ACQUISITION DES DONNÉES (ETL)

- Objectifs : Suppression des incohérences sémantiques entre les sources pouvant survenir lors de l'intégration :
  - **des schémas :**
    - problèmes de terminologie : un objet est désigné par 2 noms différents, un même nom désigne 2 objets différents
    - incompatibilités de contraintes : 2 concepts équivalents ont des contraintes incompatibles
    - conflit sémantique : choix de différents niveaux d'abstraction pour un même concept
    - conflits de structures : choix de différentes propriétés pour un même concept
  - **des données :**
    - Equivalence de champs
    - Equivalence d'enregistrements : fusion d'enregistrements

# DATA WAREHOUSE: ACQUISITION DES DONNÉES (ETL)

## Chargement

- C'est l'opération qui consiste à charger les données nettoyées et préparées dans le DW.
- C'est une opération qui peut être longue :
  - Mettre en place des stratégies pour assurer de bonnes conditions à sa réalisation
  - Définir la politique de rafraîchissement.
- C'est une phase plutôt mécanique et la moins complexe



# DATA WAREHOUSE: ACQUISITION DES DONNÉES (ETL)

## Outils ETL

- BusinessObjects, Data Integrator,
- Oracle Corporation, Warehouse builder,
- IBM, Websphere Datastage

# DATA WAREHOUSE: STOCKAGE

- Choix d'implémentation
  - MOLAP
  - ROLAP
  - HOLAP
- Implémentation du modèle en étoile et/ou des cubes et/ou des vues matérialisées
- Définition des index
- Stockage les données

# DATA WAREHOUSE: RESTITUTION

- C'est le but du processus d'entreposage des données.
- Elle conditionne souvent le choix de l'architecture du DW et de sa construction.
- Elle doit permettre toutes les analyses nécessaires pour la construction des indicateurs recherchés.

# DATA WAREHOUSE: RESTITUTION

- C'est le but du processus d'entreposage des données.
- Elle conditionne souvent le choix de l'architecture du DW et de sa construction.
- Elle doit permettre toutes les analyses nécessaires pour la construction des indicateurs recherchés.

# DATA WAREHOUSE: RESTITUTION

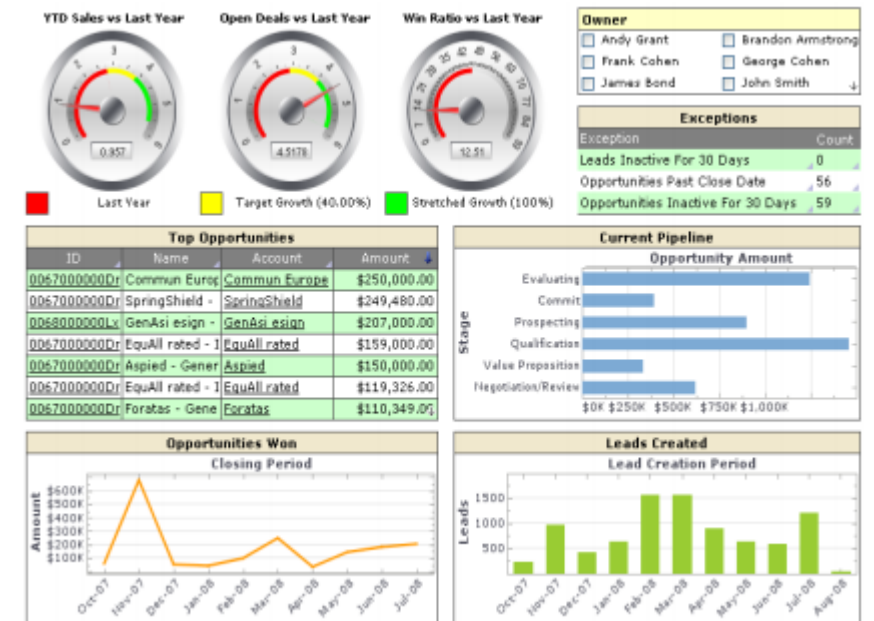
- Réalisation de rapports divers (Reporting)
- Réalisation de tableaux de bords (Dashboards)
- Analyse en ligne diverses (OLAP)
- Fouille de données (Data Mining)
- Visualisations autour d'un DW (visualizations)

# DATA WAREHOUSE: RESTITUTION

- Rapports (Reporting) :
  - Pour les utilisateurs qui ont besoin d'un accès régulier à des informations d'une manière presque statique
  - Un rapport est défini par une requête (plusieurs requêtes) et une mise en page
  - Les rapports peuvent être exécutés automatiquement ou manuellement

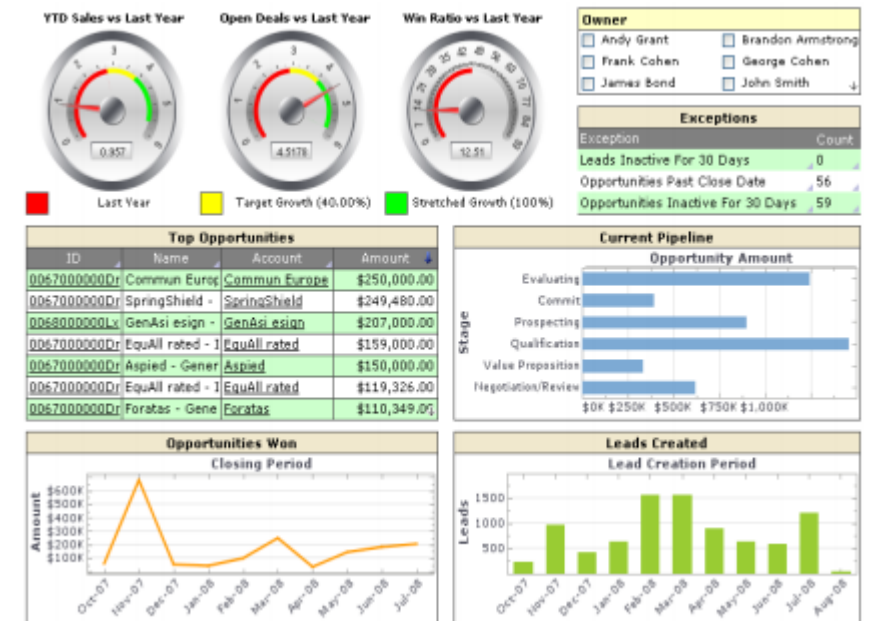
# DATA WAREHOUSE: RESTITUTION

- Tableaux de bords (Dashboards) :
  - Affichent une quantité limitée d'informations dans un format graphique facile à lire
  - Fréquemment utilisé par les cadres supérieurs qui ont besoin d'un rapide aperçu des changements les plus importants
  - Pas vraiment utile pour une analyse complexe et détaillée



# DATA WAREHOUSE: RESTITUTION

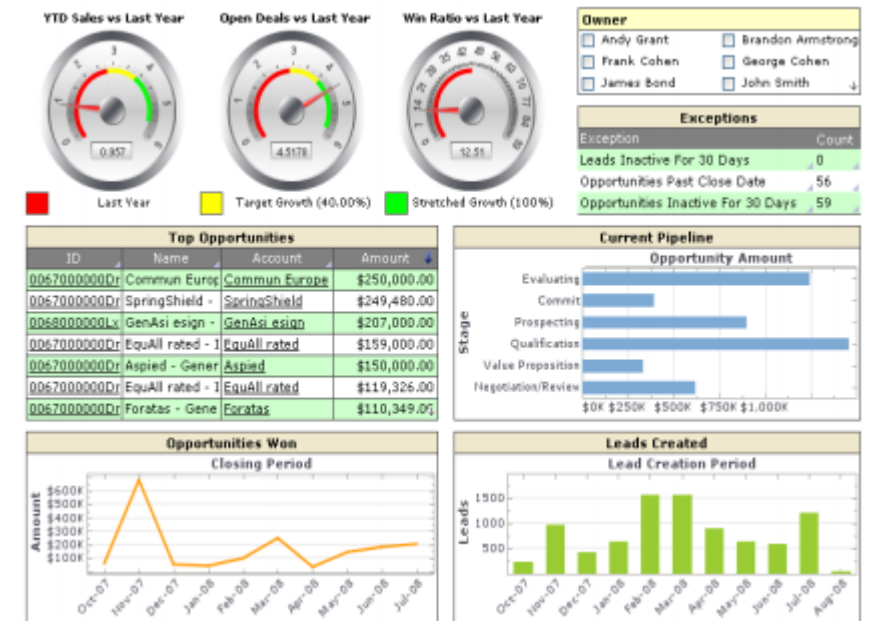
- Tableaux de bords (Dashboards) :
  - Affichent une quantité limitée d'informations dans un format graphique facile à lire
  - Fréquemment utilisé par les cadres supérieurs qui ont besoin d'un rapide aperçu des changements les plus importants
  - Pas vraiment utile pour une analyse complexe et détaillée





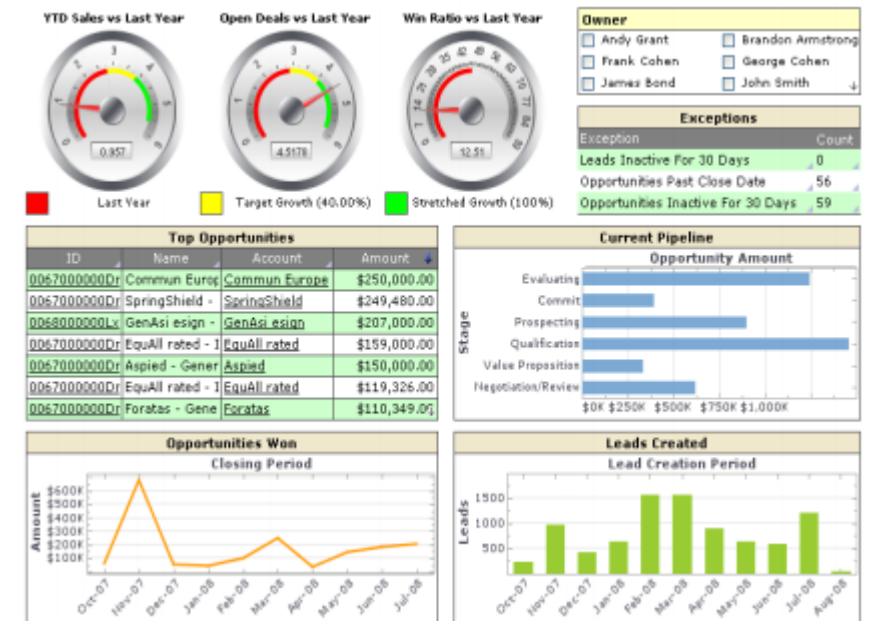
# DATA WAREHOUSE: RESTITUTION

- Tableaux de bords (Dashboards) :
  - Affichent une quantité limitée d'informations dans un format graphique facile à lire
  - Fréquemment utilisé par les cadres supérieurs qui ont besoin d'un rapide aperçu des changements les plus importants
  - Pas vraiment utile pour une analyse complexe et détaillée



# DATA WAREHOUSE: RESTITUTION

- Tableaux de bords (Dashboards) :
  - Affichent une quantité limitée d'informations dans un format graphique facile à lire
  - Fréquemment utilisé par les cadres supérieurs qui ont besoin d'un rapide aperçu des changements les plus importants
  - Pas vraiment utile pour une analyse complexe et détaillée



# DATA WAREHOUSE: RESTITUTION

- Analyse OLAP (On-Line Analytical processing) :
- Techniques OLAP apparues en recherche dans les années 70 mais ont été développées dans les années 90 dans l'industrie
- Permettent de réaliser des synthèses, des analyses et de la consolidation dynamique de données multidimensionnelles
- Constitue la façon la plus naturelle d'exploiter un DW du fait de son organisation multidimensionnelle

# DATA WAREHOUSE: RESTITUTION

- Analyse OLAP (On-Line Analytical processing) :
  - Techniques OLAP apparues en recherche dans les années 70 mais ont été développées dans les années 90 dans l'industrie
  - Permettent de réaliser des synthèses, des analyses et de la consolidation dynamique de données multidimensionnelles
  - Constitue la façon la plus naturelle d'exploiter un DW du fait de son organisation multidimensionnelle

# DATA WAREHOUSE: RESTITUTION

- Fouille de données (Data Mining) :
  - Recherche de connaissance, sous forme de modèle de comportement, cachés dans les données
  - Domaine jeune à l'intersection de l'intelligence Artificielle, les Statistiques, les BD
  - Nombreuses techniques de fouille : régression linéaire, induction d'arbres de décision, algorithmes génériques, réseaux de neurones, ...
  - Les techniques de fouille sont en pleine évolution et sont de plus en plus intégrées dans les ED

# DATA WAREHOUSE: ARCHITECTURE

