



Chapitre 3:

Analyse multidimensionnelle et OLAP

OLAP

- ▶ Catégorie de logiciel (moteur) dont le rôle est l'exploration et l'analyse rapide des données dans le DW suivant une approche multidimensionnelle utilisant plusieurs niveaux d'agrégation
- ▶ **Avantage:**
 - ▶ l'utilisateur n'a pas à maîtriser des langages d'interrogation et des interfaces complexes
 - ▶ L'utilisateur interagit avec les données



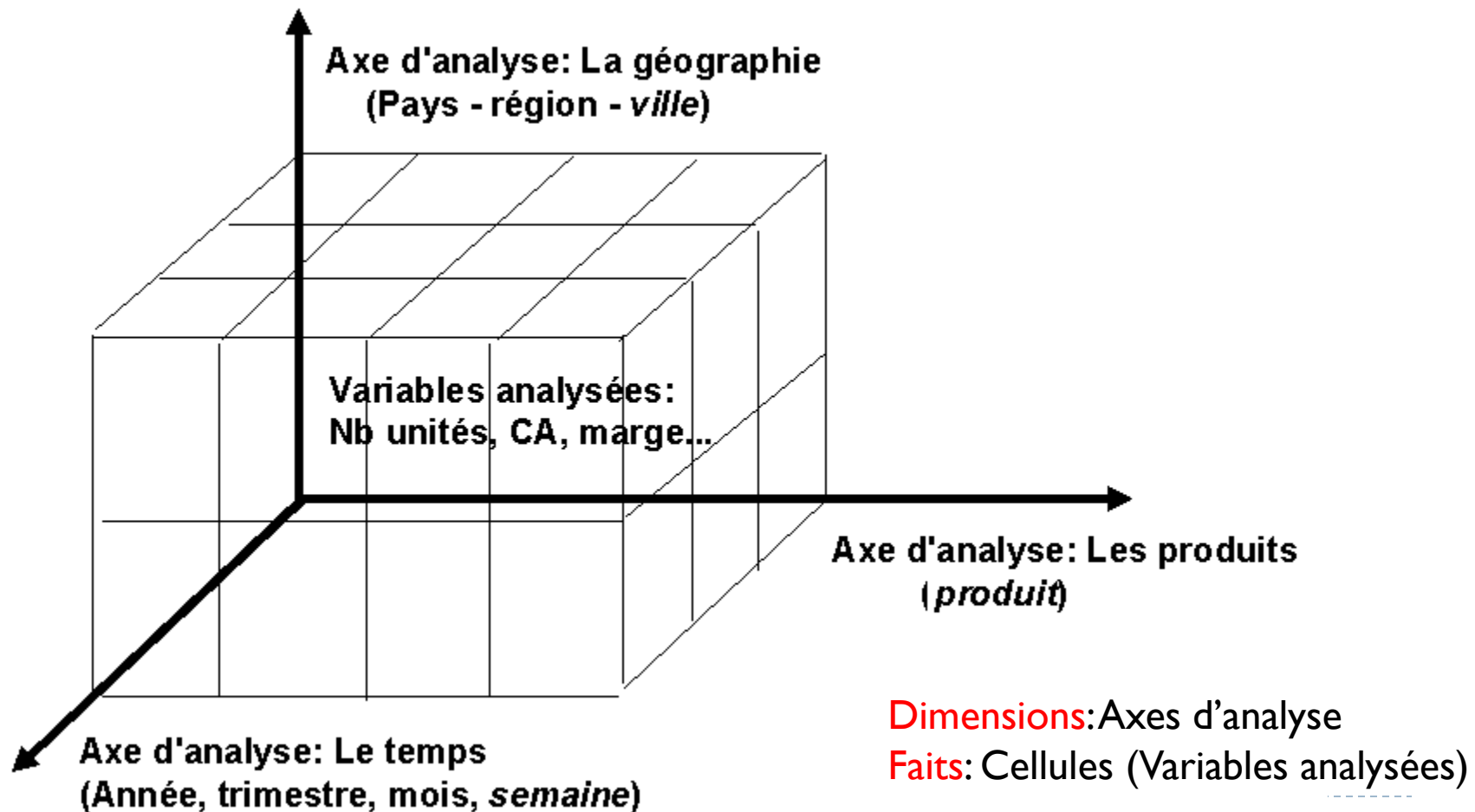
Approche multidimensionnelle

- ▶ Les informations dans un système décisionnel sont représentées par une structure à plusieurs dimensions, où les DIMENSIONS sont un ensemble d'attribut de la base de données:
 - ▶ Client
 - ▶ Produit
 - ▶ Vendeurs
 - ▶ Géographie
 - ▶ Temps
- ▶ Les cellules continent des données agrégées, qu'on appelle FAITS
 - ▶ Quantité vendue
 - ▶ Chiffre d'Affaire
 - ▶ Coût
- ▶ Le résultat est représenté sous forme:
 - ▶ Relations
 - ▶ Cube de données
 - ▶ Hyper cube de données



Représentation multidimensionnelle

- Le résultat de l'analyse serait de la forme suivante:



Représentation multidimensionnelle

▶ Agrégation des données

- ▶ Les données seront groupées à différents niveaux de granularité (niveau de détail des données)
- ▶ Les groupement sont pré-calculés / granularité
 - ▶ total des ventes / mois
 - ▶ total des ventes / années
 - ▶ total des ventes / mois 12
- ▶ Sur les granularités:
 - ▶ Temps: Jours , Mois, Trimestre, Semestre, Années, ...
 - ▶ Produit: Numéros, Types, Gammes, Marques, ...
 - ▶ Zone géographique: Quartiers, Villes, Régions, Pays, ...
 - ▶ ...



Manipulation d'un cube

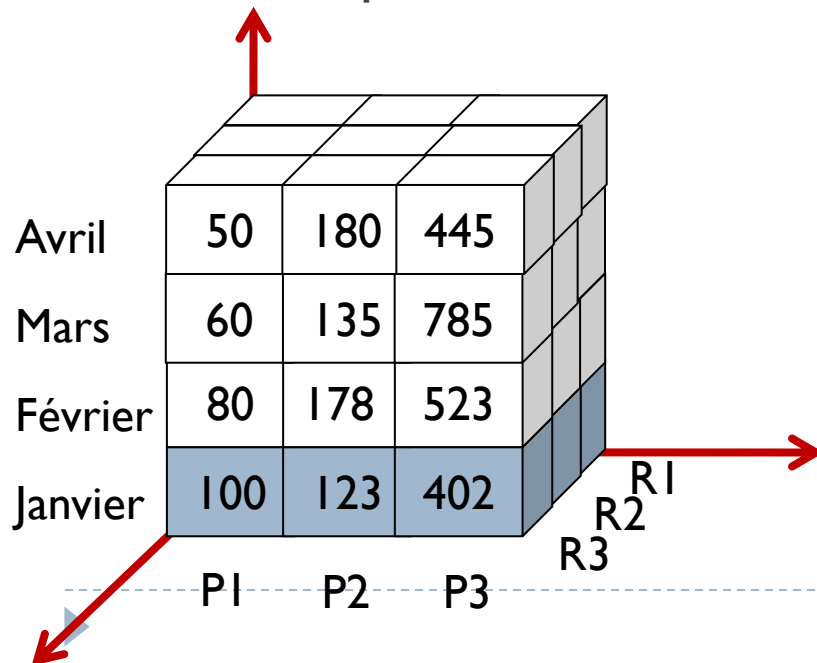
- ▶ Les Opérateurs appliqués sur le cube sont algébriques (le résultat est un autre cube) et peuvent être combinés
- ▶ Les Opérateurs sont:
 - ▶ Slicing: Extraction d'une tranche d'information
 - ▶ *Scoping ou Dicing*: Extraction d'un bloc de données
 - ▶ Roll-up (agrégation d'une dimension => résumé): passage au grain inférieur
 - ▶ Drill-down (plus détaillées): passage au grain supérieur



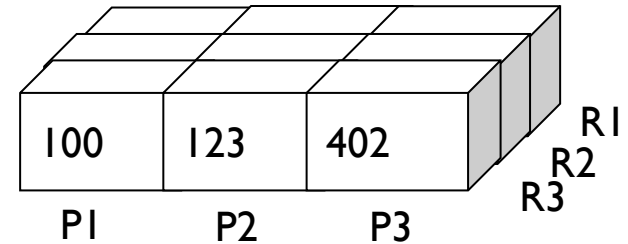
Manipulation d'un cube

► 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 2004



Slicing

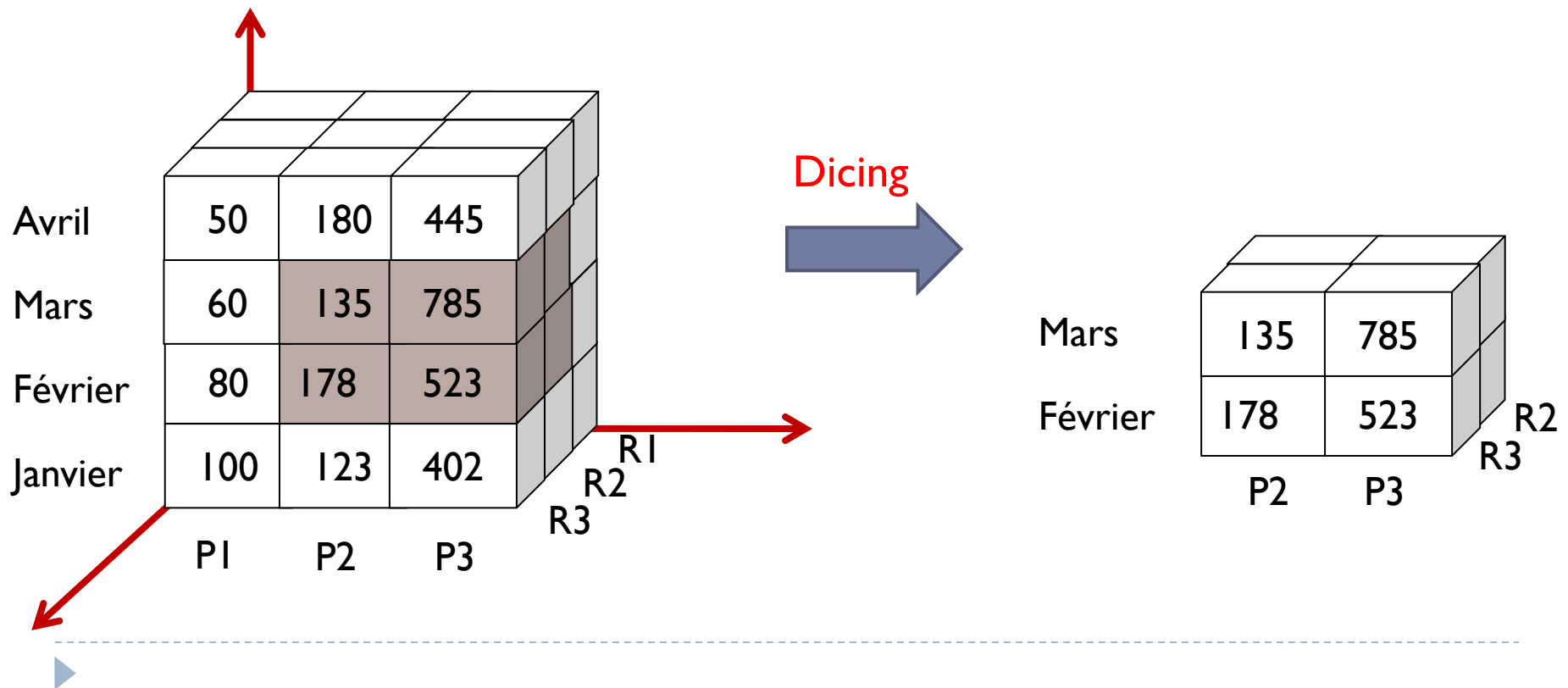


Ventes Janvier	R1	R2	R3
P1	77	79	100
P2	289	157	123
P3	54	97	402

Manipulation d'un cube

► Scoping ou Dicing

- Extraction d'un sous-cube (bloc de données). C'est une opération plus générale que le slicing



Manipulation d'un cube

- ▶ **Drill-up (Roll-up) / Drill-down**

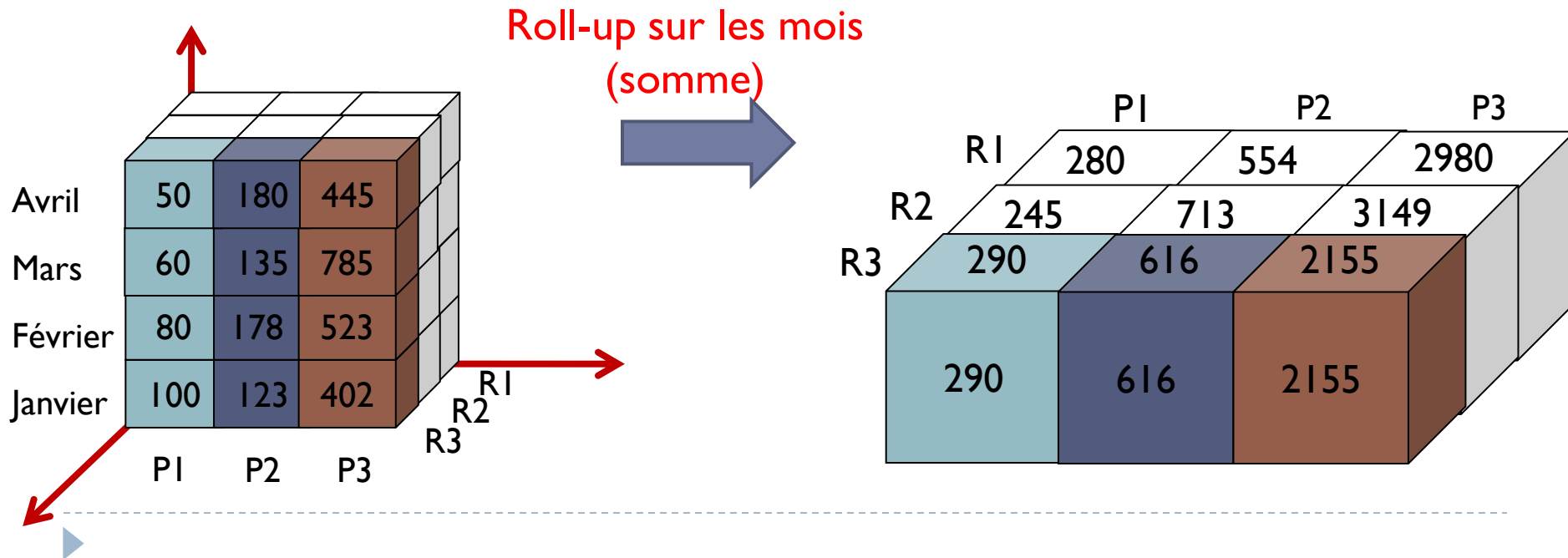
- ▶ Sont des opérations liées à la granularité.
- ▶ Ils permettent à l'utilisateur de naviguer entre les différents niveaux de données (granularité) allant de la plus résumé (granularité supérieure) au plus détaillé (granularité inférieure)



Manipulation d'un cube

► Drill-up (Roll-up)

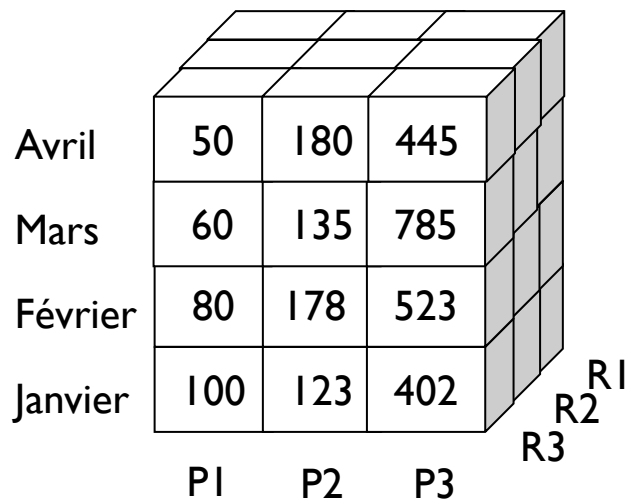
- Consiste à représenter les données du cube à un niveau de granularité supérieur sur une dimension
 - Utilisation des fonctions d'agrégation (somme, moyenne, etc) spécifiées pour la mesure et la dimension



Manipulation d'un cube

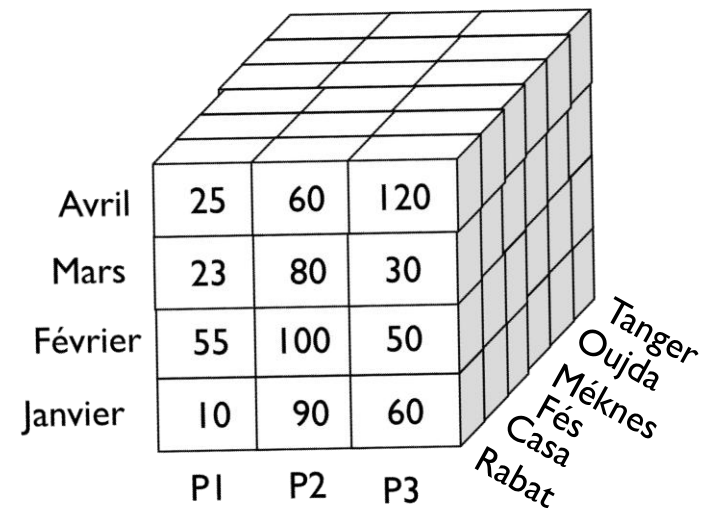
► Drill-down

- Consiste à représenter les données du cube à un niveau de granularité de niveau inférieur, donc sous une forme plus détaillée.
- On fait un « zoom » sur une dimension



Avril	50	180	445
Mars	60	135	785
Février	80	178	523
Janvier	100	123	402
	P1	P2	P3

Drill-down sur les
Régions



Avril	25	60	120
Mars	23	80	30
Février	55	100	50
Janvier	10	90	60
	P1	P2	P3

Types de serveurs OLAP

- ▶ On peut citer trois types de serveurs OLAP
 - ▶ Relational OLAP(ROLAP)
 - ▶ Multidimensional OLAP (MOLAP)
 - ▶ Hybrid OLAP (HOLAP)



ROLAP

- ▶ Dans un serveur de type ROLAP, les données sont stockées dans un SGBD relationnel et le OLAP permet de simuler le comportement d'un SGBD multi-dimensionnel
 - ▶ Le schéma de conception utilisée est soit schéma en étoile ou schéma en flocon
 - ▶ Pour la représentation multidimensionnelle, on utilise des vues (matérialisées)
 - ▶ Les requêtes OLAP (slice, rollup...) sont traduites en SQL
 - ▶ Utilisation d'index spéciaux de type bitmap
 - ▶ Administration (tuning) particulier de la base
- ▶ **Avantage**
- ▶ **Avantages**
 - ▶ Souple,
 - ▶ permet une évolution facile,
 - ▶ permet de stocker de gros volumes.
- ▶ **Inconvénient**
 - ▶ Peu efficace pour les calculs complexes

MOLAP

- ▶ Dans un serveur MOLAP, utilise des moteurs de stockage multidimensionnels pour des vues multidimensionnelles des données.
 - ▶ la structure de stockage est en cube et avec un accès direct aux données dans le cube
 - ▶ Ces cubes sont implémentés sous forme des matrices à plusieurs dimensions
 - ▶ Chaque cube est indexé sur ses dimensions
- ▶ **Avantage:**
 - ▶ Rapidité
- ▶ **Inconvénients:**
 - ▶ Formats propriétaires
 - ▶ Ne supporte pas de très gros volumes de données → utiliser plusieurs serveurs MOLAP



HOLAP

- ▶ **C'est une combinaison de ROLAP et MOLAP**
 - ▶ Données stockées dans SGBD relationnel (données de base)
 - ▶ Les données agrégées sont stockées dans serveurs MOLAP (des cubes)
- ▶ **Avantages:**
 - ▶ Il possède à la fois la grande évolutivité de ROLAP
 - ▶ La possibilité de stocker de gros volumes de ROLAP
 - ▶ L'efficacité des calculs complexes de MOLAP
 - ▶ La rapidité de MOLAP

