

Corrigé TD1 : Business Intelligence

Exercice 01

Il s'agit de modéliser l'entrepôt de données (ED) des ventes d'une entreprise commerciale. Cette entreprise vend des produits regroupés par familles de produits. Une vente correspond à un produit et un seul ; la vente est effectuée par l'un des vendeurs du service de vente spécialisé dans le produit. La semaine de vente est le numéro de semaine dans l'année. L'ED doit pouvoir fournir le chiffre d'affaires des ventes d'un produit, par date, client, et vendeur, ainsi que toutes les sommations possibles de chiffre d'affaires.

Les objets de l'ED sont les suivants :

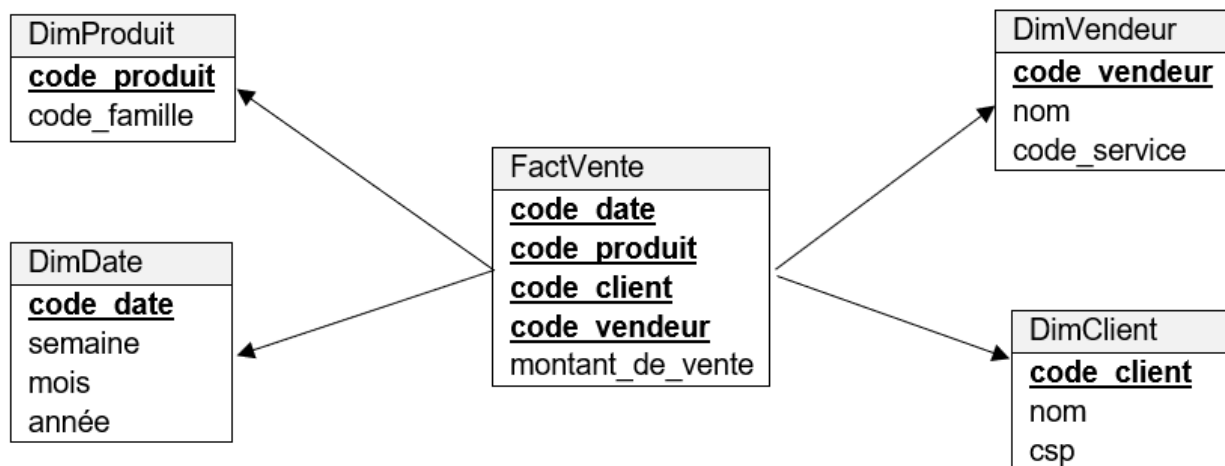
- **produit**, caractérisé par : **code_produit**, **code_famille**
- **client**, caractérisé par : **code_client**, **nom**, **csp** (categorie socio-professionnelle)
- **vente**, caractérisée par : **date**, **code_produit**, **code_client**, **code_vendeur**, **montant_de_vente**
- **vendeur**, caractérisé par : **code_vendeur**, **nom**, **code_service**
- **date**, caractérisée par : **semaine**, **mois**, **année** (la date s'écrit par exemple 20020402)

1. **Donner les définitions** des quatre termes suivants : table de faits, table de dimension, indicateur, hiérarchie.

Table de faits : Table principale du modèle dimensionnel . Elle contient les données observables (les faits) sur le sujet étudié selon divers axes d'analyse (les dimensions) dimensions)

2. **Schéma en étoile** : tracer le schéma en étoile dimensionnel de l'ED, en précisant pour chaque table sa nature dimensionnelle (table de faits ou table de dimension).

Sol :



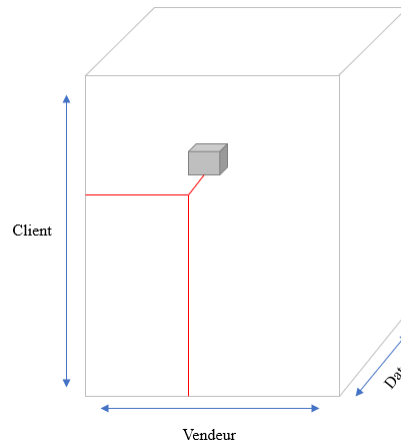
3. **Cube de données** : Dans l'exemple traité, et représenté par le schéma en étoile, le cube de données est un hyper cube à 4 dimensions :

Graphiquement, on peut dessiner en perspective 4 types de cubes à 3 dimensions. Définir les 4 types.
à quoi correspond chaque type?

Sol :

Graphiquement, on peut dessiner en perspective 4 types de cubes à 3 dimensions :

A. client, vendeur, date (pour chaque valeur de produit)



B. produit, vendeur, date (pour chaque valeur de client)

C. produit, client, date (pour chaque valeur de vendeur)

D. produit, client, vendeur (pour chaque valeur de date)

Dans chaque cube, l'élément de base est l'indicateur « montant de la vente ».

Exercice 02

Part 1

Donner le schéma en étoile associé à l'activité « **Voyages aériens** » d'une agence de voyage pour répondre au besoin suivant :

- Quel est le **chiffre d'affaires** (CA) par **client**, par **date** de voyage (et par mois, trimestre et année), par **compagnie aérienne**, par **ville de destination** ? Les tableaux de bord doivent pouvoir présenter les totaux et sous totaux de chiffre d'affaires (CA) : tous clients confondus, et/ou toutes dates, et/ou toutes compagnies, et/ou toutes destinations.

Part 2

Donner le schéma en étoile associé à l'activité « **Location de voiture** » d'une agence de voyage pour répondre au besoin suivant :

- Dans le cas de la location de voiture, on souhaite éditer le **CA**, le **nombre de jours de location**, et le **kilométrage** pour chaque **client**, date de **réservation**, **ville**, **loueur**, et **catégorie** de véhicule.

Part 3

Donner le schéma en étoile associé à l'activité « **Hôtel** » d'une agence de voyage pour répondre au besoin suivant :

- Dans le cas de l'hôtellerie, on veut des tableaux de bord par **client**, **hôtel**, **ville**, **date** de début de séjour, faisant apparaître le **nombre de nuitées** et le **prix total** payé.

- chiffre d'affaires** (CA) / **client**, **date**, **compagnie aérienne**, **ville**
- CA**, **nombre de jours de location**, **kilométrage** / **client**, **date**, **ville**, **loueur**, et **catégorie du véhicule**
- nombre de nuitées**, **prix total** / **client**, **hôtel**, **ville**, **date**

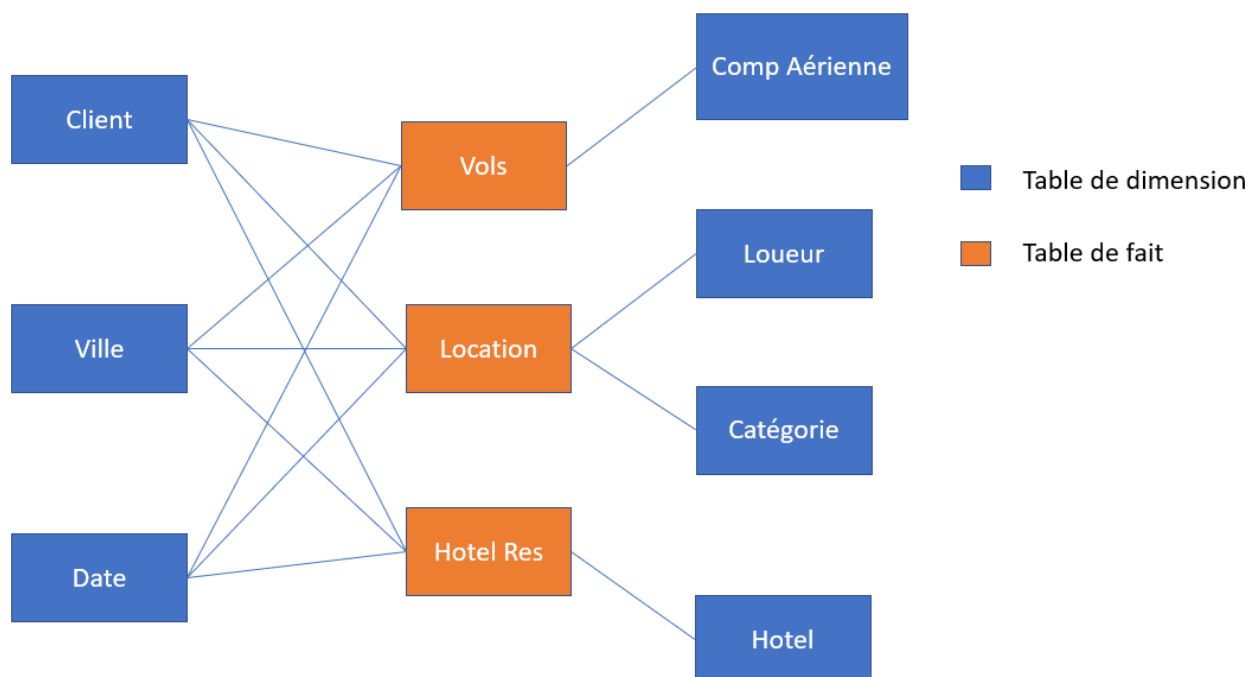
Part 4

Regrouper les trois ED (exercice 3, 4 et 5) en un seul, afin de répondre aux questions supplémentaires suivantes :

- Quel est le CA total induit par un déplacement en avion ?
- Quelle est la durée du séjour ?
- Quel est le CA en location de voiture ? En hôtellerie ?

On veut pouvoir éditer les détails de CA par période de temps et par client, ville, compagnie aérienne, loueur et hôtelier, et faire tous les regroupements utiles.

- Figurer le modèle dimensionnel d'un tel ED, en montrant en particulier comment l'on peut retrouver location de voiture et/ou hôtellerie, si elles existent, à partir d'un déplacement en avion. Un voyage en avion n'implique pas forcément location de voiture et/ou hôtellerie, et inversement.



Exercice 03

On veut construire un entrepôt de données afin de stocker les informations sur les consultations d'un pays. On veut notamment connaître le nombre de consultations, par rapport à différents critères (personnes, médecins, spécialités, etc).

Ces informations sont stockées dans les relations suivantes :

PERSONNE (id_personne, nom, tel, adresse, sexe)

MEDECIN (id_medecin, tel, adresse, spécialité)

CONSULTATION (id_med, id_pers, date, prix)

1. Proposer un schéma relationnel qui tienne compte de la date, du jour de la semaine, du mois, du trimestre et de l'année.

→ Quelle est la table des faits ?

Consultation

→ Quels sont les faits ?

Le prix et le nombre de consultations

→ Combien de dimensions ont été retenues ? Quelles sont-elles ?

3 dimensions :

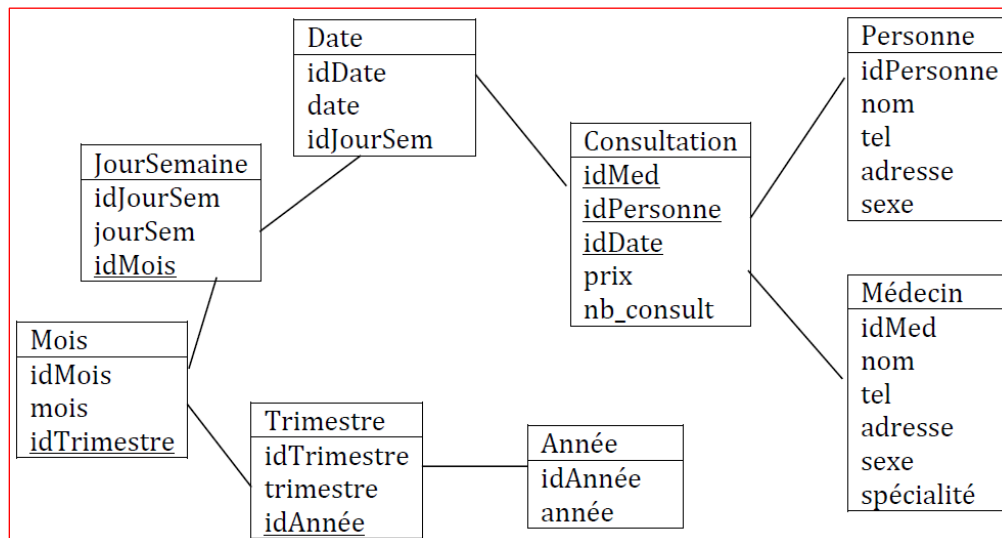
Médecin,

Personne,

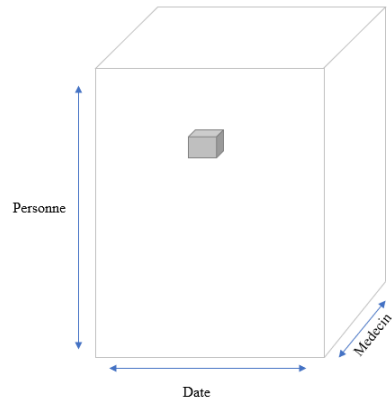
Temps

→ Quelles sont les hiérarchies des dimensions ? Dessinez-les.

Date: Année -> Trimestre -> Mois -> JourSemaine -> Date



2. Faites une représentation du cube OLAP sans tenir compte des hiérarchies.

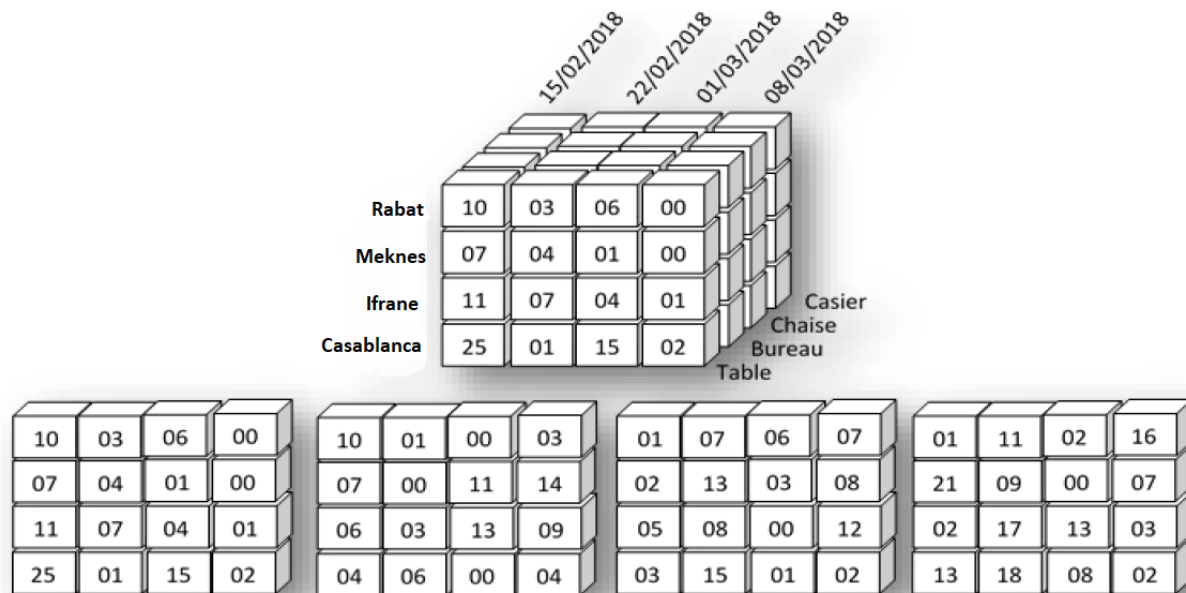


3. A partir de ce cube, indiquez quelles opérations OLAP (roll up, drill down, slice, dice) il faut appliquer pour obtenir les informations suivantes :

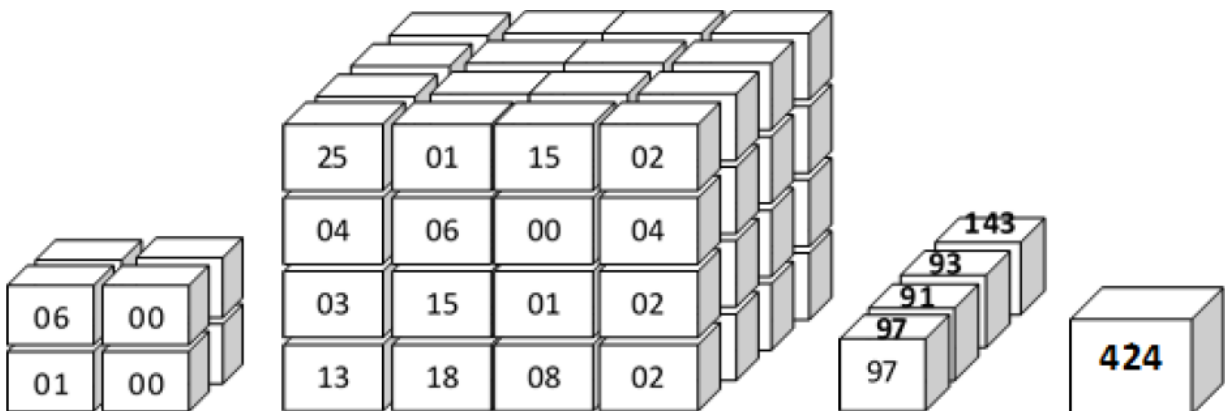
- Le coût total des consultations par médecin en 2012 et 2013.
Roll up on Date to Année
Slice Année = 2012
Roll up on Date to Année
Slice Année = 2013
- Le nombre de consultations par jour de la semaine, par spécialité et par sexe du patient.
Roll up on Date to jourSemaine
Roll up on Médecin to spécialité
Roll up on Personne to sexe
- Le coût des consultations par patiente pour les mois d'octobre.
Roll up on Date to mois
Slice mois = octobre

Exercice 04

Soit le cube de données suivant :



1. Déterminer les dimensions et les hiérarchies des dimensions de chacun des cuboïdes suivants.
2. Donner le détail des opérations qui ont permis d'avoir ces cuboïdes.

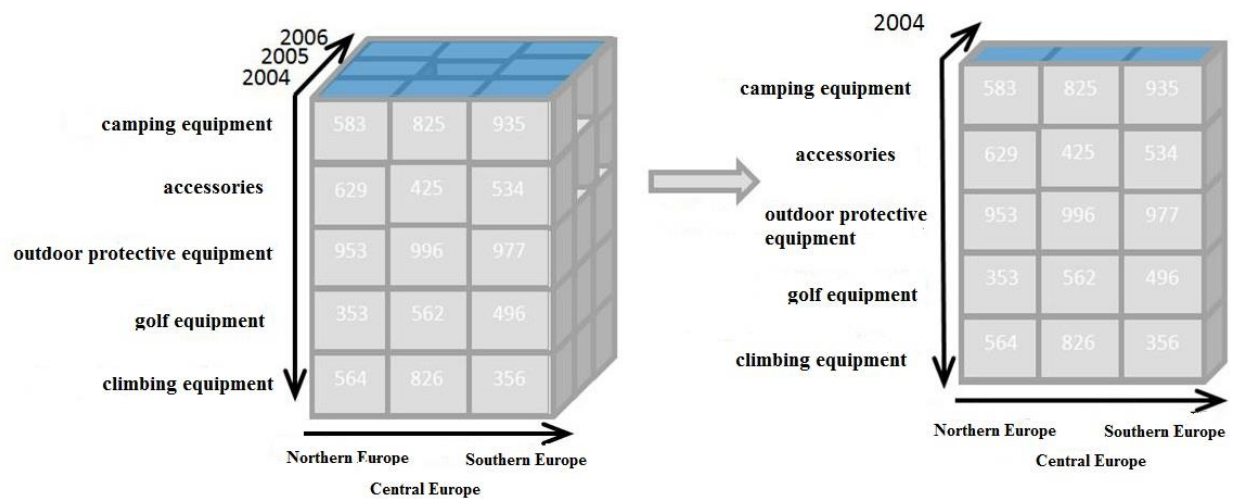


25	01	15	02
04	06	00	04
03	15	01	02
13	18	08	02

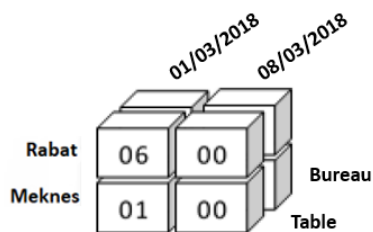
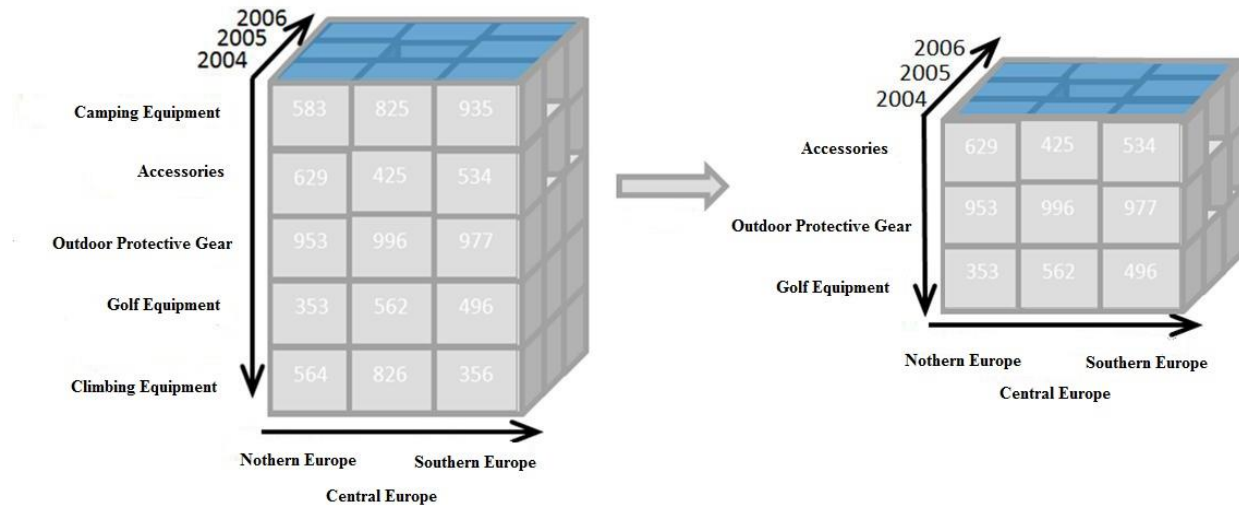
11	07
07	04
04	01
01	00

Sol

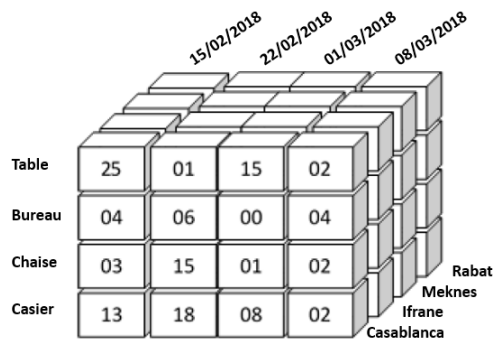
Slicing : extraction d'une tranche d'information,



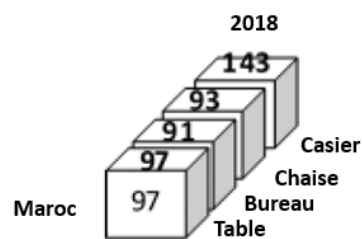
Scoping (ou Dicing) : extraction d'un bloc de données (opération plus générale que le slicing),



Dice on Date, Ville, Equipment

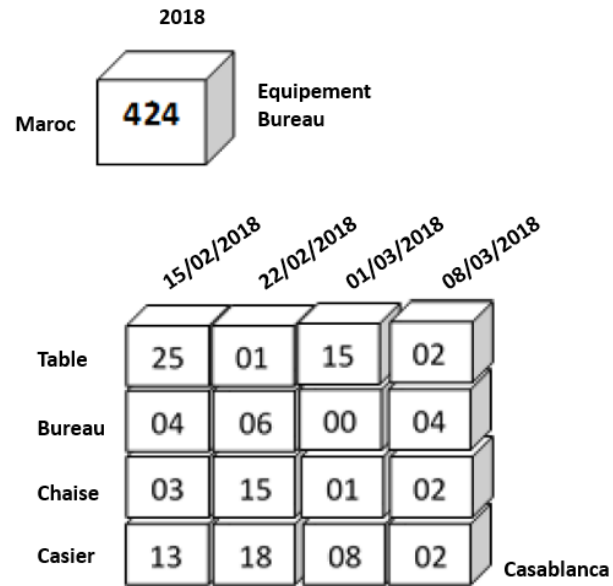


Pivot Ville et Equipement



Roll up on Date (Année)

Roll up on Ville (Pays)



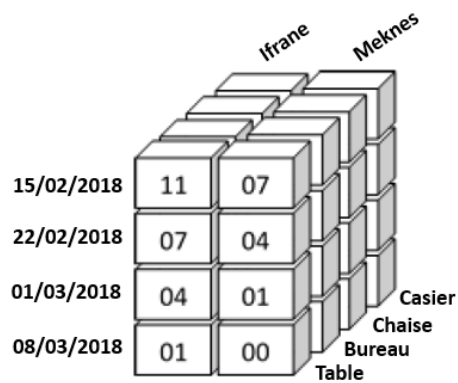
Roll up on Date (Année)

Roll up on Ville (Pays)

Roll up on Equipement (Catégorie)

Pivot Ville et Equipement

Slice on Ville = Casablanca



Pivot Ville et Date

Slice on Ville = Ifrane

Slice on Ville = Meknes