

## TD : 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 sommes 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.
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).
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?
4. Supposons un cube D représentant une coupe de l'hypercube à 4 dimensions, selon une valeur de la variable **date**. Combien de tableaux à deux dimensions on peut obtenir à partir de D. En tout combien y a-t-il de types différents de tableaux à deux dimensions ?

## 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, ainsi que toutes les sommes de la même manière que pour les déplacements.

### 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.

### 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)

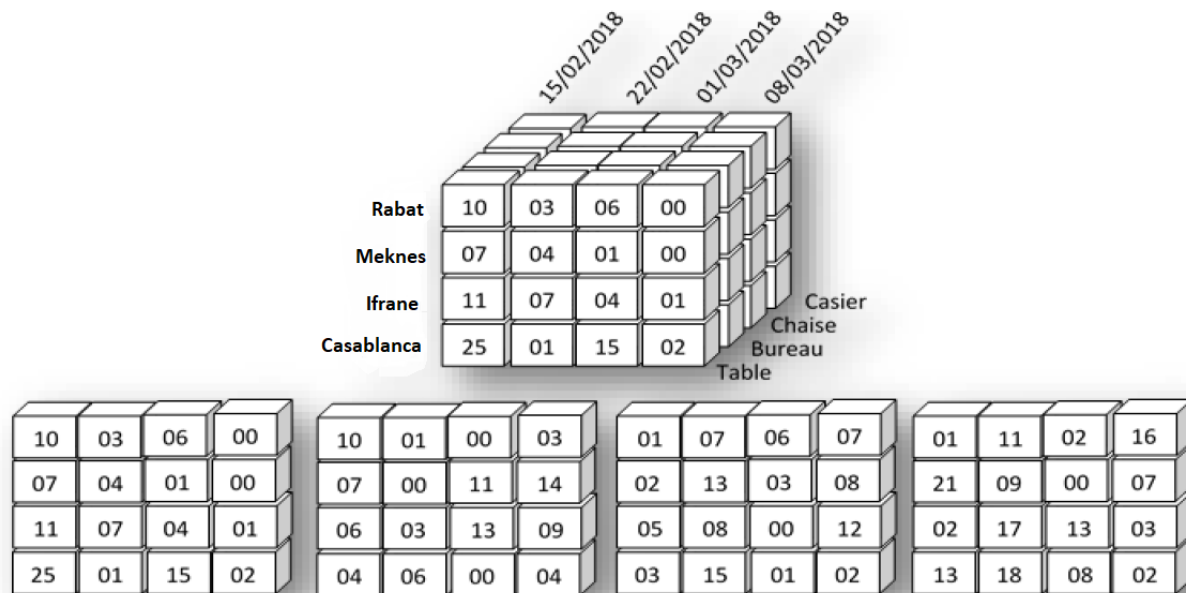
**MEDECIN** (id\_medecin, tel, adresse, spécialité)

**CONSULTATION** (id\_med, id\_pers, date, prix)

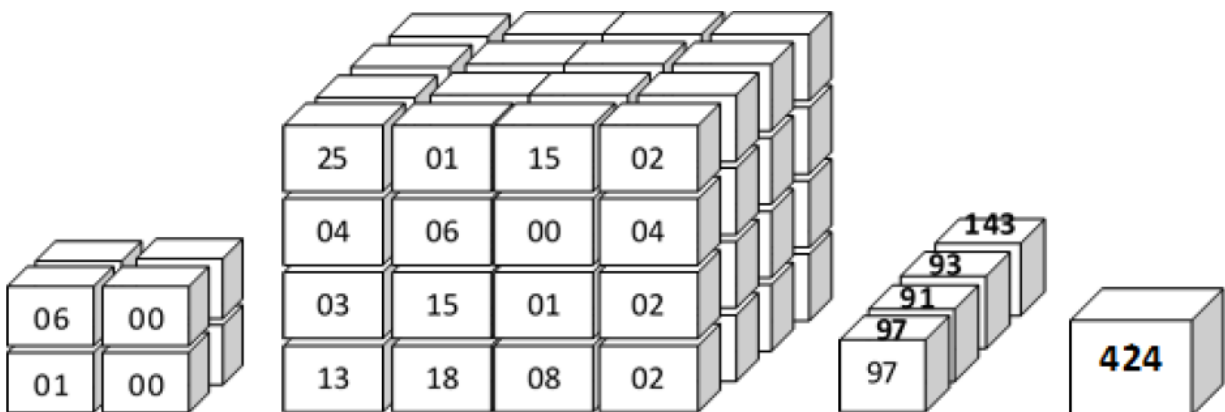
1. Proposer un schéma relationnel qui tienne compte de la date, du jour de la semaine, du mois, du trimestre et de l'année.
  - Quelle est la table des faits ?
  - Quels sont les faits ?
  - Combien de dimensions ont été retenues ? Quelles sont-elles ?
  - Quelles sont les hiérarchies des dimensions ? Dessinez-les.
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.
  - Le nombre de consultations par jour de la semaine, par spécialité et par sexe du patient.
  - Le coût des consultations par patiente pour les mois d'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