

## TD 2 : Conception d'un entrepôt de données

### Exercice 1

**Rappel Définition 1 :** Un Entrepôt de données est une collection de données orientées sujet , intégrées, non volatile et historisées, organisées pour le support d'un processus d'aide à la décision .

**Rappel Définition 2** Un modèle en étoile est constitué de :

- Une **table de faits** : **identifiants des tables de dimension ; une ou plusieurs mesures** .
- Plusieurs **tables de dimension** : **descripteurs des dimensions**.
- Une **granularité définie par les identifiants dans la table des faits**.

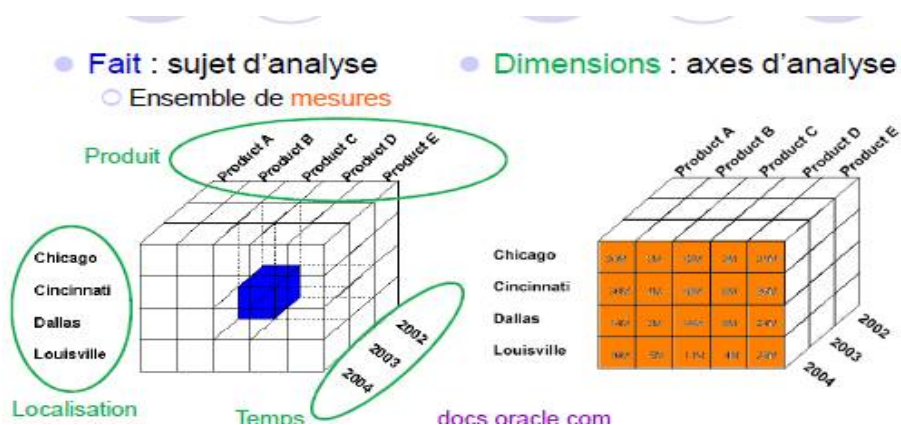
**Exemple :** Une entreprise de vente souhaite mettre en place un système d'information décisionnel sous la forme d'un data mart (un mini entrepôt de données) pour observer son activité de ventes au niveaux des différents lieux de distributions de ses articles et cela dans plusieurs villes. Ces lieux de distributions sont renseignés par leur enseigne, leur type (en fonction de leur surface), leur adresse (code postal et ville), leur département, leur région. Les ventes sont renseignées selon une période qui se décline en mois, en trimestre et année. Les ventes sont observées par le nombre d'articles selon le type, et le chiffre d'affaire

Les chiffres d'affaires réalisés par une entreprise sont regroupés par années comme suit :

Années N

Produit	VILLE 1	VILLE 2	VILLE 3
Produit A	4	8	5
Produit B	10	7	9
Produit C	8	6	7

Le CUB OLAP peut être définie comme suit



### Question

- 1- Quels sont les axes d'analyse, et les mesures
- 2- Construire le modèle en étoile de ce Cube en utilisant la table VENTE

### Exercice 2

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 d 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 ch Les objets de l'ED sont les suivants :

- 1) produit, caractérisé par : code\_produit, code\_famille
- 2) client, caractérisé par : code\_client, nom, csp
- 3) vente, caractérisée par : date, code\_produit, code montant\_de\_vente
- 4) vendeur, caractérisé par : code\_vendeur, nom, code\_service
- 5) date, caractérisée par : semaine, mois, année

### Questions

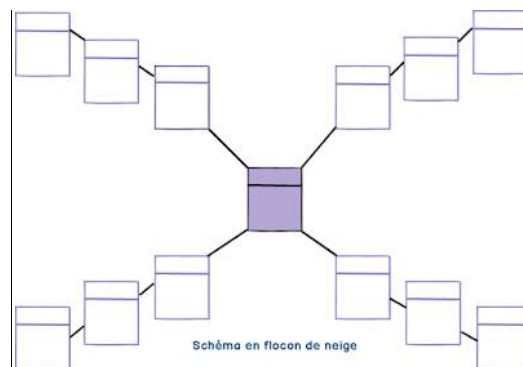
1. Tracer le schéma en étoile dimensionnel de l'ED, en précisant pour chaque table sa nature dimensionnelle (table de faits ou table de dimension) que la nature des champs.
2. Cube de données : Dans l'exemple traité, et représenté par le schéma en étoile, le cube de données est un hyper cube à 4 dimensions : Produit, client, vendeur, date. Graphiquement, on peut dessiner en perspective 4 types de cube à 3 dimensions. Définir les 4 types et à quoi correspond chaque type?

### Exercice 3

1. Concevoir un modèle en étoile qui permet d'analyser les ventes d'une entreprise de restauration rapide. Le principe est de mesurer les ventes grâce aux quantités vendues et aux bénéfices, en fonction des ventes réalisées par jour, dans un restaurant donné, pour un aliment donné. L'objectif est de pouvoir analyser les ventes par jour, par semaine, par mois et par année.

Les restaurants peuvent être regroupés en fonction de leur ville et de leur pays

**Définition 3 : Le schéma en flocon de neige** est un type de **schéma en étoile** qui inclut la forme hiérarchique des tables dimensionnelles. Dans ce schéma, il existe une table de faits composée de différentes tables de dimension et de sous-dimension reliées par des clés primaires et étrangères à la table de faits. Il porte le nom de **flocon de neige** car sa structure ressemble à un flocon de neige.



2. Modifier le modèle proposé en un modèle en flocon de neige pour modéliser explicitement les hiérarchies des dimensions représentant le temps et la localisation géographique des magasins.
3. On souhaite à présent mesurer le nombre de commandes qui est donné par jour et par restaurant. Etendre le modèle précédent afin de prendre en compte cet aspect.