

TD– — Conception conceptuelle d'un Data Warehouse

Durée : 2h

Cas d'étude 1 — Consultations médicales

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 , nom, tel, adresse, sexe)
- **MEDECIN** (id , tel, adresse, spécialité)
- **CONSULTATION** (id_med , id_pers , date, prix)

Travail à faire

1. Proposer un schéma relationnel qui tienne compte de la date, du jour de la semaine, du mois, du trimestre et de l'année.
2. Quelle est la table des faits ?
3. Quels sont les faits ?
4. Combien de dimensions ont été retenues ? Quelles sont-elles ?
5. Quelles sont les hiérarchies des dimensions ? Dessinez-les.
6. Faites une représentation du cube OLAP sans tenir compte des hiérarchies.
7. À partir de ce cube, indiquez quelles opérations OLAP (*roll up*, *drill down*, *slice*, *dice*) il faut appliquer pour obtenir les informations suivantes :
 - a. le coût total des consultations par médecin en 2012 et 2013
 - b. le nombre de consultations par jour de la semaine, par spécialité et par sexe du patient
 - c. le coût des consultations par patiente pour les mois d'octobre

Cas d'étude 2 — Analyse des salaires

On dispose d'un outil OLAP pour analyser les salaires selon l'âge et le niveau d'étude des personnes et la situation géographique des entreprises.

- L'analyse selon l'âge peut se faire par **année** ou par **décade** (tranches de 10 années à partir de 14 ans et jusqu'à 73 ans).
- L'analyse du niveau d'étude peut se faire :
 - par le niveau d'enseignement atteint en fin d'études (**primaire, secondaire, supérieur**) ;
 - ou par le dernier diplôme obtenu (**certificat de fin d'étude primaire, BEPC, Bac, Licence, Master**).
- L'analyse de la situation géographique peut se faire par **ville, département, région** ou **pays**.

Travail à faire

1. Quel est le schéma relationnel en étoile le plus approprié pour cette analyse ?
2. Quelle est la table des faits ?
3. Quels sont les faits ?
4. Combien de dimensions ont été retenues ? Quelles sont-elles ?
5. Quelles sont les hiérarchies des dimensions ? Dessinez-les.
6. Faites une représentation du cube OLAP en tenant compte des hiérarchies.

S'inspirer des slides 47 et 48 du support de cours.

Cas d'étude 3 — Du relationnel au décisionnel (étoile puis flocon)

Une entreprise dispose d'une base de données relationnelle transactionnelle permettant de gérer ses commandes clients. Les données sont stockées dans les tables suivantes :

- **CLIENT** (id_client , nom, sexe)
- **PRODUIT** (id_produit , libelle, categorie)
- **COMMANDE** (id_commande , date_commande, id_client)
- **DETAIL_COMMANDE** (id_commande , id_produit , quantite)

Cette base permet de gérer les commandes, mais ne permet pas directement d'analyser les ventes (chiffre d'affaires, volumes, etc.).

On veut construire un modèle décisionnel permettant d'analyser :

- les quantités vendues ;
- le chiffre d'affaires ;
- selon le temps, les clients et les produits.

Travail à faire

1. Identifier le fait

- Quel est l'événement métier principal à analyser ?
- Existe-t-il une table correspondant déjà à une table de faits ?
- À partir de quelles tables construiriez-vous la table de faits ?
- Quelle serait la granularité du fait ?

2. Proposer une table de faits construite à partir du modèle relationnel.

3. Identifier les mesures et les dimensions nécessaires.

4. Proposer un schéma en étoile.

5. Proposer un schéma en flocon :

- Identifier une dimension pouvant être normalisée.
- Transformer cette dimension en schéma en flocon.
- Expliquer : ce que l'on gagne et ce que l'on perd par rapport au schéma en étoile.