

Modélisation multidimensionnelle

Exercice : Achats

Nous disposons de deux sources de données relatives à l'achats d'articles

Source 1

Produit(numP, nomP, prixP, délaiLiv, stock)

Fournisseur(numF, nomF, localisation, remise) nomF est unique

Localisation de la forme (ville, province, pays)

Achat(noAchat, dateAchat, moyenPaiement, numF#)

AchatUnit(noAchat#, numP#, qte, prixUnit)

Source 2

ProdAchats(codeProd, dateAchat, descProd, catProd, nomFour, qte, dispo, prixUnit, montant)

On considère qu'il y a 1000 produits et 10 fournisseurs. Chaque produit est fourni par exactement un fournisseur et chaque achat comporte 5 produits. On considère qu'il a 20 achats par jour.

Un même produit ne peut être renseigné que dans l'une des deux sources et jamais dans les deux à la fois.

Le but est de suivre l'évolution de l'état des stocks et des montants dépensés dans le temps.

Pour chaque période de temps et chaque zone géographique mesurer les informations suivantes :

- la quantité moyenne des stocks d'un produit donné
- la somme total gagnée par chaque fournisseur
- le pourcentage du prix de chaque produit dans les achats effectués
- le taux de participation de chaque fournisseur, en terme de montant gagné, dans les achats effectués

La granularité temporelle comprend la semaine, le mois, le trimestre, et l'année. La granularité spatiale comprend la ville, la province et le pays.

Pour ce faire, il faudra :

- 1- Identifier les dimensions utiles au problème. Expliquer comment obtenir chaque dimension à partir des données sources. Indiquer quelles dimensions sont hiérarchiques et expliciter leur hiérarchie.
- 2- Définir les mesures et les fonctions d'agrégation associées
- 3- Déduisez le schéma multidimensionnel en considérant que la table de faits est normalisée et que les tables décrivant les dimensions sont dé-normalisées. Discuter le choix de la granularité choisie pour les dimensions hiérarchiques.

- 4- On considère les statistiques suivantes. Estimer la taille, en nombre de ligne, de la table de faits en considérant que les données portent sur une année.
- 5- On voudrait revoir la conception pour obtenir la quantité totale de produit vendus pour chaque catégorie. Quelle modification apporter et est-ce que peut poser un problème de cohérence et quelle solution suggérer ?
- 6- Création des vues définissant le schéma multidimensionnel en SQL

Exercice : Enseignements

Soit la BD de gestion des enseignements de l'université

Etudiant (idEtu, nom, prenom, dateNaiss, adresse, bac)

Cours (idC, nom, formation)

Enseignant(idEns, nom, prenom, specialite)

TypeExam(idType, type)

Inscription(idEtu#,idC#,anneeUniv)

Composer(idEtu#,idC#,idType#,note,date)

Enseigner(idC#,idEns#,nbH,anneeUniv)

Responsable(idEns#,idC#,decharge,anneeUniv)

Un étudiant est caractérisé par un identifiant idEtu, un nom, un prénom, une date de naissance (dateN) et une adresse. On connaît aussi son origine au travers de la spécialité du bac qu'il a obtenu. Un cours est caractérisé par un identifiant idC, un nom et est associé à une et une seule formation. Un enseignant est caractérisé par un identifiant idEns, un nom, un prénom et une spécialité. Un étudiant peut s'inscrire à un cours pour une année universitaire anneeUniv et obtenir une note pour un examen d'un cours à une date donnée date. Deux types d'examen (type) sont possibles : contrôle continu et examen terminal. On fera l'hypothèse que chaque cours sera évalué par un seul contrôle continu et un seul examen terminal. Un enseignant donne un nombre de cours nbh pour une matière et peut aussi être responsable d'un cours pour une année universitaire. Cette responsabilité lui donne droit à une décharge (décharge) d'heures d'enseignement.

Cahier des charges : On souhaite analyser les enseignements donnés à l'université selon les indicateurs suivants :

- Les notes moyennes des étudiants pour un cours à une date donnée
- Le nombre d'heures enseignées pour un enseignant et un cours
- Le nombre d'heures de décharge pour un enseignant et un cours

1- Proposez un schéma conceptuel d'une base de données multidimensionnelle permettant de répondre au cahier des charges.

2- Traduisez le schéma conceptuel en schéma logique de normalisé.

3- Écrivez les requêtes SQL permettant de construire les faits du schéma conceptuel.

Exercice : Films

Même questions que précédemment en considérant le schéma suivant :

Film(idFilm, titre, genre, annee, duree, idReal#)

Realisateur(idReal, idPersonne#)

Personne(idPersonne, nom, prenom, dateNaiss)

Role(idRole, nomRole)

Acteur(idActeur, idPersonne#)

Utilisateur(idUser, idPersonne#)

Note(idFilm#,idUser#, note, commentaire)

Et les besoins analytiques suivants

- Le nombre de films réalisés par un acteur par an, par réalisateur et par genre.
- La durée moyenne des films réalisés par un acteur par an, par réalisateur et par genre.
- Le nombre de notes déposées par un utilisateur, par film, par genre de film et par mois.
- La note moyenne déposée par un utilisateur, par film, par genre de film et par mois.