

DESS DCISS
Examen « Entrepôt de Données »
1^{er} avril 2004 16h15–18h15
Tous documents autorisés

I. Questions de cours (1/4 des points)

- a. Définir les termes suivants : modèle dimensionnel, datamart, agrégat.
- b. Dans le cas d'un schéma en étoile comportant plusieurs tables de fait, comment sont-elles reliées ? Quel est l'intérêt qu'elles soient reliées ?

II. Exercice. (3/4 des points)

Énoncé

Une banque distribue une carte de paiement (« carte de crédit ») à ses clients. Elle décide de réaliser un entrepôt de données afin de faire le suivi des paiements suivants effectués avec la carte :

- a) voyages en avion,
- b) locations de voiture,
- c) hôtellerie.

Elle veut faire un suivi indépendant de chacun des paiements a), b) ou c), mais aussi avoir la possibilité d'un suivi global.

A chaque déplacement en avion, la compagnie aérienne lui envoie un fichier contenant les éléments suivants : identification de la carte de paiement, coordonnées du client et de la compagnie aérienne; ville de départ, ville d'arrivée, n° du vol, date du vol, n° du billet, classe du siège, distance parcourue, date d'achat et prix payé.

Pour simplifier, on supposera qu'un voyage en avion se réduit à un vol (voyage aller effectué dans la même journée, pas de correspondances), et que le client voyage seul.

Les loueurs de véhicule transmettent après chaque location : identification de la carte de paiement, coordonnées du client et de la société de location de véhicules; catégorie du véhicule, date de début de location, date de fin de location, nombre de jours, distance parcourue, date de réservation et prix payé.

L'hôtel transmet à chaque séjour : identification de la carte de paiement, coordonnées du client et de l'hôtel; catégorie de chambre, date de début de séjour, date de fin de séjour, nombre de nuitées, date de réservation, prix de l'hébergement et prix de la restauration.

Questions

1. Un premier entrepôt de données (ED) ne concerne que les déplacements en avion. Établir le modèle dimensionnel. Faire clairement apparaître les dimensions et les indicateurs. Cet ED doit permettre de répondre aux questions suivantes : 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 CA : tous clients confondus, et/ou toutes dates, et/ou toutes compagnies, et/ou toutes destinations.

2. De même, établir deux autres modèles dimensionnels, l'un pour les locations de voiture, l'autre pour l'hôtellerie.
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 sommations de la même manière que pour les déplacements.
Dans le cas de l'hôtellerie, on veut des tableaux de bord par client, hôtel, ville, date de début de séjour, catégorie de véhicule, faisant apparaître le nombre de nuitées, le prix total payé et le ratio « prix de l'hébergement / prix de la restauration ».
3. Quels sont les nombres de dimensions et d'indicateurs respectifs de chacun des trois ED ? Que peut on dire des représentations en cubes ?
4. On veut maintenant regrouper ces trois ED 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 désire ici pouvoir éditer les détails de CA par période de temps et par client, ville de destination, ville de location (si différente), ville d'hébergement (si différente), 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 hôtellerie, si elles existent, à partir d'un déplacement en avion. On notera qu'un voyage en avion n'implique pas forcément location de voiture et/ou hôtellerie, et inversement.
5. Faire le schéma d'une table d'agrégat permettant de stocker le nombre de kilomètres parcourus par catégorie de véhicule et par date (jour de début de location, mois, trimestre et année), tous clients, loueurs et villes de location confondus.

Quelles sont les tables de dimension associées à cette table d'agrégat ?
