

Grande distribution

Description de l'existant

Une entreprise de distribution dispose d'un réseau de points de vente ou magasins. La base de données, dont le schéma relationnel est donné ci-dessous, lui permet d'effectuer le suivi des ventes journalières pour les divers produits commercialisés dans ses différents points de vente. Tous les points de vente achètent **au même prix les produits quelle que soit la date d'achat**. En revanche, le prix de vente dépend du produit vendu, du magasin de vente ainsi que de la date de vente (historique des prix de vente).

PRODUIT	(<u>IdP</u> , Libelle, PrixAchat, CodeCat)
CATEGORIE	(CodeCat, Designation, CodeCatGen)
MAGASIN	(<u>IdM</u> , Nom, Type, IdL)
LIEU	(<u>IdL</u> , Nomlieu, IdLrattach)
VENTE	(<u>IdP</u> , <u>IdM</u> , <u>DateV</u> , Quantite, PrixVenteUnitaire)

Conventions :

- Les clefs primaires sont soulignées. Les clefs étrangères sont en italique gras.
- Les attributs de même nom sont définis sur le même domaine. L'attribut *IdLrattach* est compatible avec *IdL* et l'attribut *CodeCatGen* fait référence à la clef primaire *CodeCat*.

Les divers produits sont identifiés (*IdP*), caractérisés par leur libellé et leur prix unitaire d'achat et classés en catégorie. Les catégories constituent une dimension de structure hiérarchique. Par exemple, la catégorie <817, Textile> se subdivise en plusieurs sous-catégories comprenant <597, Vêtements> et <1004, Linge de maison>. Cette dernière regroupe plusieurs autres catégories dont <989, Literie> qui elle-même se décompose en <2310, Taies>, <1441, Drap>...

Pour chaque catégorie décrite dans la relation CATEGORIE, l'attribut *CodeCatGen* correspond au code de la catégorie père dans la hiérarchie.

Les magasins ont un identifiant, un nom et un type dont les valeurs appartiennent à l'ensemble : {Hypermarché, Supermarché, Superette...}.

L'attribut *IdL* dans MAGASIN correspond à l'identifiant de la commune dans laquelle se situe un magasin. Outre ces communes, la relation LIEU répertorie les départements dans lesquels elles se situent, les régions correspondant aux départements... Pour chaque lieu l'attribut *IdLrattach* indique son lieu de rattachement direct (ex : le lieu de nom Marseille est rattaché au lieu de nom Bouches du Rhône).

Dans la relation VENTE, *Quantite* est le nombre d'unités vendues pour le produit *IdP*, dans le magasin *IdM* à la date *DateV*. L'attribut *PrixVenteUnitaire* correspond au prix unitaire lors de la vente de ce produit à la date de la vente.

Dans les requêtes suivantes, le chiffre d'affaire est obtenu en faisant le produit des quantités de produits vendus par le prix unitaire des produits considérés qui dépend de la date de vente et du point de vente.

Echauffement SQL

- Q1. Créez la base de données correspondante au schéma relationnel.
Attention aux clés primaires et clés étrangères ainsi qu'aux autres contraintes.
- Q2. Donnez le chiffre d'affaire réalisé par le magasin d'identifiant 123 ?
- Q3. Donnez le chiffre d'affaire réalisé par magasin (identifiant et nom) ?
- Q4. Pour chaque magasin, donnez le chiffre d'affaire réalisé par catégorie de produit (CodeCat) ?
- Q5. Donnez la hiérarchie de lieux (IdL, NomLieu) du magasin d'identifiant 123.
- Q6. Donnez le total des quantités vendues pour le produit d'identifiant 20 dans la région de nom PACA ?
- Q7. Quels sont les magasins (IdM) qui ont vendu chaque catégorie de produits le 10/10/2013 ?
Paraphrasage : quels sont les magasins qui ont vendu, le 10/10/2013, autant de catégories de produits qu'il en existe.

Un premier pas vers l'entrepôt de données : historisation des prix de vente

- Q8. Modifier le schéma de la base de données en supprimant l'attribut PrixVenteUnitaire de la relation VENTE.
- Q9. Ajouter la relation BAREME (IdP, IdM, DateB, Prix) pour historiser les prix de vente.
- Q10. Répondez de nouveau aux questions 2 à 7 sur la nouvelle base de données.

Préliminaire à la mise en œuvre de l'entrepôt de données

A partir des questions précédentes et du contexte exposé, répondez aux questions suivantes :

- Q11. Quels sont les enjeux du projet (interne, votre entreprise ; et externe, votre client) ?
- Q12. Quels sont les objectifs de ou des analyses ?
- Q13. Quelles sont les erreurs ou anomalies pouvant apparaître dans les données sources ?
- Q14. Quels sont les axes d'analyse ? Pour chaque axe, proposer la hiérarchie associée.
- Q15. Quel est la fréquence (ou les fréquences) de mise à jour des données ?
- Q16. Déterminer les indicateurs/mesures nécessaires à l'analyse.
- Q17. Quels sont les regroupements nécessaires ?
- Q18. Proposer des requêtes SQL permettant de détecter les erreurs et de les corriger (cf. Q13).

Mise en œuvre de l'entrepôt de données

- Q19. Modéliser l'entrepôt de données correspondant.
- Q20. Utiliser les descriptions des hiérarchies de la question 14 pour écrire les requêtes SQL, procédures ou autres pour alimenter les dimensions de l'entrepôt de données à partir de la base de données. Information : un cube n'est pas toujours la solution même s'il l'englobe tout et n'oubliez pas les mises à jour (cf. question 15). Penser aux fonctions OLAP d'ORACLE comme `CUBE ()`, `GROUPING SETS (() , ())`, `ROLLUP ()` ...
- Q21. Même question que la précédente pour la table de faits.
- Q22. Créer les Datamarts associés aux objectifs et justifier vos choix de réalisation (découplé ou non découplé, matérialisé, virtuel ou partiellement matérialisé).

Requêtes d'analyse

- Q23. A partir de la nouvelle base de données comprenant la table BAREME, adaptez les requêtes déjà faites au nouveau schéma.
- Q24. Donnez les plus grandes marges réalisées pour le produit d'identifiant 13 ?
- Q25. Donnez les plus petites marges réalisées par produit ainsi que les magasins correspondants ?
- Q26. Donnez pour chaque magasin les produits pour lesquels les marges sont les plus grandes ?
- Q27. Donnez les magasins qui, tout produit confondu, réalisent les plus petites marges ?
- Q28. Quel est le produit (`IdP`, `Libelle`) le plus vendu (en nombre d'unités) ?
- Q29. Pour chaque produit (identifiant et libellé) rattaché à la catégorie Mobilier directement ou indirectement, donnez le nombre total d'unités vendues.
- Q30. Quelles sont les communes (nom) ayant tous les types de points de vente ?
- Q31. Calculez le cube correspondant aux quantités vendues par produit, par magasin et par mois.
- Q32. Quel est le chiffre d'affaire réalisé par magasin et catégorie de produit ?
- Q33. Quel est le chiffre d'affaire réalisé par magasin et catégorie de produit pour la région PACA ?
- Q34. Pour chaque produit, donnez le magasin qui en a vendu le plus grand nombre d'unités et celui qui a effectué la plus importante vente journalière.