Mustapha NEZZARI

Dans ce compte rendu, vous trouverez les réponses au TP2.

Compte rendu du TP2

SQL OLAP

Sommaire

[Modélisation dimensionnelle 3](#_Toc467357500)

[Schéma en étoile 3](#_Toc467357501)

[L’entrepôt de données 3](#_Toc467357502)

[La dimension Produit 3](#_Toc467357503)

[La dimension Client 3](#_Toc467357504)

[La dimension Lieu 3](#_Toc467357505)

[La dimension Temps 3](#_Toc467357506)

[La table des faits Vente 3](#_Toc467357507)

[Les mesures 3](#_Toc467357508)

[Les vues matérialisées 3](#_Toc467357509)

[Requêtes 3](#_Toc467357510)

[Mode de rafraichissement 3](#_Toc467357511)

[Les clés 3](#_Toc467357512)

[Tests des vues 4](#_Toc467357513)

[Optimisations 4](#_Toc467357514)

[Les index 4](#_Toc467357515)

[Les hiérarchies 4](#_Toc467357516)

# Modélisation dimensionnelle

Dans le cas présenté, l’entreprise cherche à analyser ses ventes selon des mesures différentes. Elle s’intéresse aux **quantités** vendues, au **prix** de ventes (avec éventuellement des **remises**) et aux **chiffres d’affaire**.

Elle souhaite aussi pouvoir avoir des informations sur les produits vendus comme leur **désignation**, leur **catégorie**, leur **sous-catégorie**.

Elle s’intéresse aussi aux caractéristiques de ses clients. Les questions qu’elle se pose montrent un intérêt pour leur **âge**, leur **tranche d’âge**, leur **sexe**. Afin de cibler ses campagnes promotionnelles, elle s’intéresse à des critères géographiques locaux et internationaux comme le **pays**, la **ville** ou encore le **code postal**.

Enfin, elle veut archiver le tout afin de pouvoir effectuer des analyses temporelles sur certaines **années**, **mois**, **semaines** et **jours**.

En tenant compte de toutes ces informations, j’ai réalisé le schéma en étoile en Figure 1.

INSERER SHEMA EN ETOILE

On peut y voir la table de fait Vente et les différentes dimensions qui représentent les axes possibles d’analyse à savoir Client, Lieu, Produit et Temps.

# L’entrepôt de données

Afin de créer cet entrepôt de données, j’ai utilisé les vues matérialisées. Il s’agit ici d’expliquer certains choix ainsi que certaines méthodes utilisées.

## La dimension Produit

Cette dimension représente un produit à vendre.

### Script SQL

Explications

Dans le SIO on dispose de la table Produit qui possède trois attributs :

* Sa référence
* Sa désignation qui est composé du nom du produit, de sa catégorie et, s’il y en a une, de sa sous-catégorie.
* Son stock

Il a donc fallu découper la désignation pour séparer le nom, la catégorie et la sous-catégorie grâce au regex.

## La dimension Client

## La dimension Lieu

## La dimension Temps

## La table des faits Vente

### Les mesures

La mesure « age »

## Les vues matérialisées

### Requêtes

### Mode de rafraichissement

Selon moi, il n’est pas nécessaire de rafraichir les vues à chaque fois qu’il y a une entrée de fait dans le SIO car les données du SID seront utilisées ponctuellement et non de façon continue. J’ai donc décidé d’utiliser la méthode REFRESH FORCE ON DEMAND sur toutes mes vues, qui me permet de manuellement rafraîchir mes vues de façon incrémentale (si c’est possible, sinon toute la vue est rafraîchie).

Il faut donc utiliser les méthodes suivantes pour rafraîchir les vues à chaque fois que l’on veut s’en servir pour prendre des décisions :



### Tests

Afin de tester si mes vues fonctionnent, j’ai créé une procédure qui permet d’insérer 4 ventes dans les tables du SIO. Le script de cette procédure est disponible dans le fichier *procedure\_insert.sql*.

# Les clés

# Tests des vues

# Optimisations

## Les index

## Les hiérarchies