

Cours Entrepôts des données

Enseignante: Salma DAMMAK

2 ème année Licence GSI

Année Universitaire: 2020/2021

Partie 2:Modélisation Multidimensionnelle

01

Les concepts de base

02

Les schémas dimensionnelles

03

Les Modèles logiques

Les concepts de base

- Un élément fondamental dans la démarche de spécification d'un système d'information.
- Les données doivent être organisées d'une manière qui soit à la fois conforme au point de vue de l'utilisateur et compatible avec les contraintes techniques de mise en œuvre.

Modélisation entité-relation??

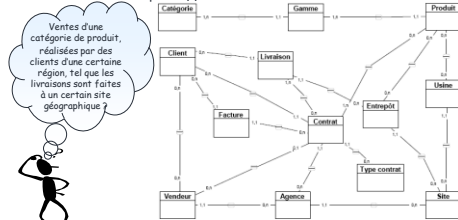
Cours entrepôt de données

3

Les concepts de base

Modélisation E/R

Exemple : Application d'administration de ventes



4

Les concepts de base

■ Avantages du modélisation E/R:

- ☺ Normalisation : Eliminer les redondances, Préserver la cohérence des données
- ☺ Optimisation des transactions
- ☺ Réduction de l'espace de stockage

Exemple: Une simple recherche dans une table principale d'adresses clients.

→ Cette recherche est contrôlée par une clé d'adresse client, qui définit l'unicité de l'enregistrement et permet une recherche indexée extrêmement rapide.

→ Le respect des formes normales fait que cette mise à jour soit faite en une itération, sans risque d'oublier des enregistrements.

5

Les concepts de base

■ Limites du modèle ER:

- ☹ Schéma trop complet: plusieurs tables.
- ☹ Pas d'interface graphique rendre utilisable le E/R.
- ☹ Données historiques difficilement représentées.
- ☹ Difficulté de compréhension par les utilisateurs non informaticien.
- ☹ Inadapté pour l'analyse.

Cours entrepôt de données

6

Les concepts de base

« Les modèles entité-relation ne peuvent pas servir de base pour construire des entrepôts de données de l'entreprise » (Kimball, 1997)



Besoin d'une **modélisation spécifique** aux Entrepôts des données

Cours entrepôt de données

7

Les concepts de base

■ Nouvelle méthode conceptuelle autour des concepts métiers

- Ne pas normaliser au maximum

■ Introduction de nouveaux types de tables

- Table de faits
- Table de Dimensions

■ Définition de nouveaux modèles

- Schéma en étoile
- Schéma en flocon

Cours entrepôt de données

8

Les concepts de base

© Compréhensibilité:

- Les données sont regroupées selon des catégories d'affaires

© Performance:

■ Un modèle dimensionnel est le résultat :

- d'une analyse des besoins : **ce que je souhaite étudier.**
- d'une analyse des données disponibles : **ce que je peux étudier.**

Cours entrepôt de données

9

Les concepts de base

■ La modélisation multidimensionnelle

- appelée modélisation OLAP (Codd 1993)

- correspond mieux aux besoins du décideur en intégrant la modélisation par sujet.

- permet des accès hautement performants

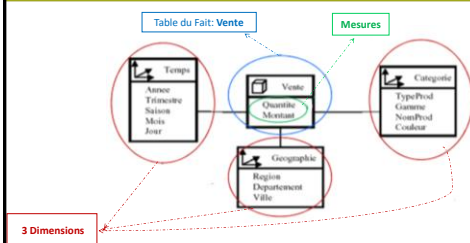
- considère les données comme des points dans un espace à plusieurs dimensions

- Ces points représentent les centres d'intérêts décisionnels (sujets) analysés en fonction des différents axes d'analyse.

Cours entrepôt de données

10

Les concepts de base



Cours entrepôt de données

11

Les concepts de base

■ Table de Fait

- un centre d'intérêt décisionnel.

- Ce que l'on souhaite mesurer

- Quantités vendues, montant des ventes

- Contient les clés étrangères des axes d'analyses (dimensions)

- Date, produit

Cours entrepôt de données

12

Les concepts de base

- **Mesure**
 - Un indicateur d'analyse de type numérique et cumulable.
 - Accompagnée d'un ensemble de fonctions d'agrégation qui permettent de l'agréger en fonction des axes d'analyse.
- Mesure **Additive**: additionnable selon toutes les dimensions
 - Quantité vendue, chiffre d'affaire
- Mesure **semi additive**: additionnable suivant certaines dimensions
 - Solde d'un compte bancaire
- Mesure **non additive**: non additionnable quelque soit la dimension
 - Prix unitaire

Cours entrepôt de données

13

Les concepts de base

- Exemple:
 - Fait: Montant des ventes, chaque jour pour chaque produit dans chaque magasin
 - en général plusieurs lignes et peu de colonnes



Cours entrepôt de données

14

Les concepts de base

- **Dimension**
 - Axe d'analyse selon lequel vont être étudiées les données.
 - Dimension = axe d'analyse Client, produit, période de temps...
- Contient une clé primaire unique qui correspond à l'un des composants de la clé multiple de la table des faits.

Cours entrepôt de données

15

Les concepts de base

Produit (Clé produit (CP), Code produit, Description du produit, Famille du produits, Marque Emballage, Poids)

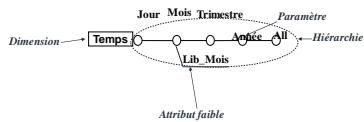
Produit
Clé Produit
Descripteur produit
Descripteur marque
Descripteur catégorie
Descripteur type emballage
Table emballage
Poids
Unité de mesure du poids
Type de stockage
Type de durée rayon
Longueur sur étagère
Hauteur sur étagère
Profondeur sur étagère

Cours entrepôt de données

16

Les concepts de base

La dimension Temps Commune à l'ensemble du DW Reliée à toute table de faits



Cours entrepôt de données

17

Les concepts de base

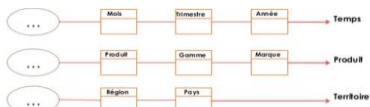
- Paramètre
 - un attribut appartenant à une dimension.
- représente un niveau de détail selon lequel sont visualisées les mesures d'activité d'un sujet d'analyse.
- Les paramètres d'une dimension peuvent être accompagnés de descripteurs appelés **attributs faibles** qui n'est pas utilisé dans les calculs de regroupement
 - Exemple, l'identifiant d'une agence Code_Ag peut être accompagné par le nom de celle-ci.

Cours entrepôt de données

18

Les concepts de base

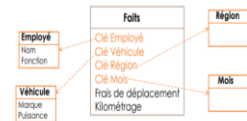
- Une **hiérarchie**
 - une perspective d'analyse définie dans une dimension.
 - regroupe un ensemble de paramètres organisés de la granularité la plus fine vers la granularité la plus générale.
 - Exemple: année – semestre – trimestre - mois



19

Les concepts de base

- Exemple 1:
 - Requête: Quels sont les frais de **déplacement** et le **kilométrage** des **commerciaux** de la région **nord** ayant des véhicules de **10 à 14 CV** en **avril 2004**?
 - Frais de déplacement
 - Kilométrage
 - Par Employé (fonction)
 - Par Région
 - Par Véhicule (puissance)
 - Par Mois



Cours entrepôt de données

20

Les concepts de base

Exemple 2:

Requête: Quelles ont été les marges sur les ventes du produit 'P023' pour le client Ben Salah Ahmed à Hammamet durant le mois de Janvier?

- Marge
- Par Produit
- Par Client
- Par Ville
- Par Mois

Cours entrepôt de données

21

Les concepts de base

Exemple 3:

Requête: Quels ont été les revenus sur les ventes de la marque 'Teams' en Tunisie durant l'année 2011?

- Revenu
- Marque
- Pays
- Année

Cours entrepôt de données

22

Les concepts de base

Exemple 4:

Requête: Quels ont été les quantités vendues de la gamme 'G006' durant le Trimestre 2 pour la région du nord ?

- Quantité
- Gamme
- Trimestre
- Région

Cours entrepôt de données

23

Les concepts de base

😊 Diminution du nombre de tables et de jointures

😊 Modèle évolutif qui peut être modifié sans peine

Cours entrepôt de données

24

Les schémas dimensionnelles

- Schéma en étoile
- Schéma en flocon de neige
- Schéma en constellation

Cours entrepôt de données

25

Les schémas dimensionnelles

- **Schéma en étoile**
 - **1 sujet d'analyse** (Table de Fait) comportant un ou plusieurs indicateurs (mesures)
 - **n axes d'analyse** (Dimensions), comportant les descripteurs des dimensions (paramètres)
 - **m perspectives d'analyse** (Hiérarchies) organisant les paramètres en différentes granularités

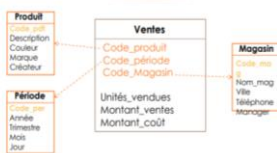
Cours entrepôt de données

26

Les schémas dimensionnelles

▪ Schéma en étoile

Exemple : « Analyse des ventes en fonction du temps, de produits et de magasins »

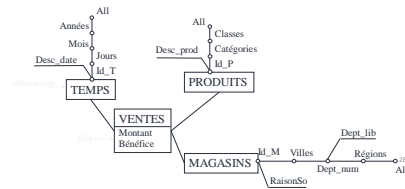


Cours entrepôt de données

27

Les schémas dimensionnelles

▪ Schéma en étoile



Cours entrepôt de données

28

Les schémas dimensionnelles

■ Schéma en étoile

- ⊗ Facilité de navigation
- ⊗ Performances : nombre de jointures limité ; gestion des données creuses.
- ⊗ Gestion des agrégats

- ⊗ Redondances dans les dimensions.
- ⊗ Alimentation complexe..

Cours entrepôt de données

29

Les schémas dimensionnelles

■ Schéma en flocon de neige

- = Modèle en étoile + normalisation des dimensions

- Utilisé lorsque les tables sont très volumineuses
- Dérivé du schéma en étoile où les tables de dimensions sont normalisées
- La table des faits reste inchangée
- Chacune des dimensions est décomposée selon sa (ou ses) hiérarchie(s)
- Exemple : Commune, Département, Région, Pays, Continent

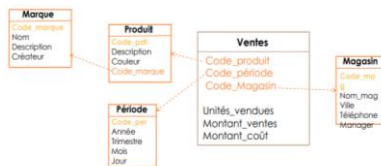
Cours entrepôt de données

30

Les schémas dimensionnelles

■ Schéma en flocon de neige

- Exemple: Normalisation de la dimension Produits

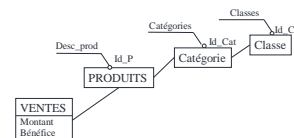


31

Les schémas dimensionnelles

■ Schéma en flocon de neige

- Exemple: Normalisation de la dimension Produits



Cours entrepôt de données

32

Les schémas dimensionnelles

■ Schéma en flocon de neige

- ☹ Réduction du volume
- ☹ Permettre des analyses par pallier (drill down) sur la dimension hiérarchisée
- ☹ Navigation difficile
- ☹ Nombreuses jointures

Cours entrepôt de données

33

Les schémas dimensionnelles

■ Schéma en constellation

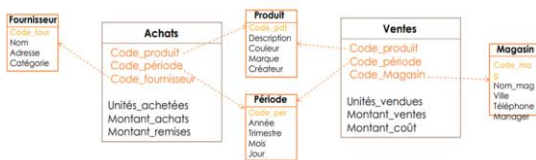
- Fusionner plusieurs modèles en étoile qui utilisent des dimensions communes
- Un modèle en constellation comprend donc :
 - Plusieurs tables de faits
 - Des tables de dimensions communes ou non à ces tables de faits.

Cours entrepôt de données

34

Les schémas dimensionnelles

■ Schéma en constellation



Cours entrepôt de données

35

Les schémas dimensionnelles

■ Schéma en constellation

- ☺ Facilite les corrélation entre les différents sujets d'analyse.
- ☺ Simplifie la modélisation avec la possibilité de partager les dimensions.

Cours entrepôt de données

36