Entrepôt de Données

Jean-François Desnos

Jean-Francois.Desnos@grenet.fr

Définition (Bill Inmon 1990)

Un entrepôt de données (data warehouse) est une collection de données thématiques, intégrées, non volatiles et historisées pour la prise de décisions

Les données sont thématiques

Données sont organisées par thème (sujets majeurs, métiers),

vs systèmes de production : processus fonctionnels.

➔ analyses transversales / structures fonctionnelles et organisationnelles de l'entreprise.

Les données sont intégrées

Elles proviennent de systèmes sources hétérogènes

➔ cohérence, normalisation, maîtrise de la sémantique, prise en compte des contraintes référentielles et des règles de gestion

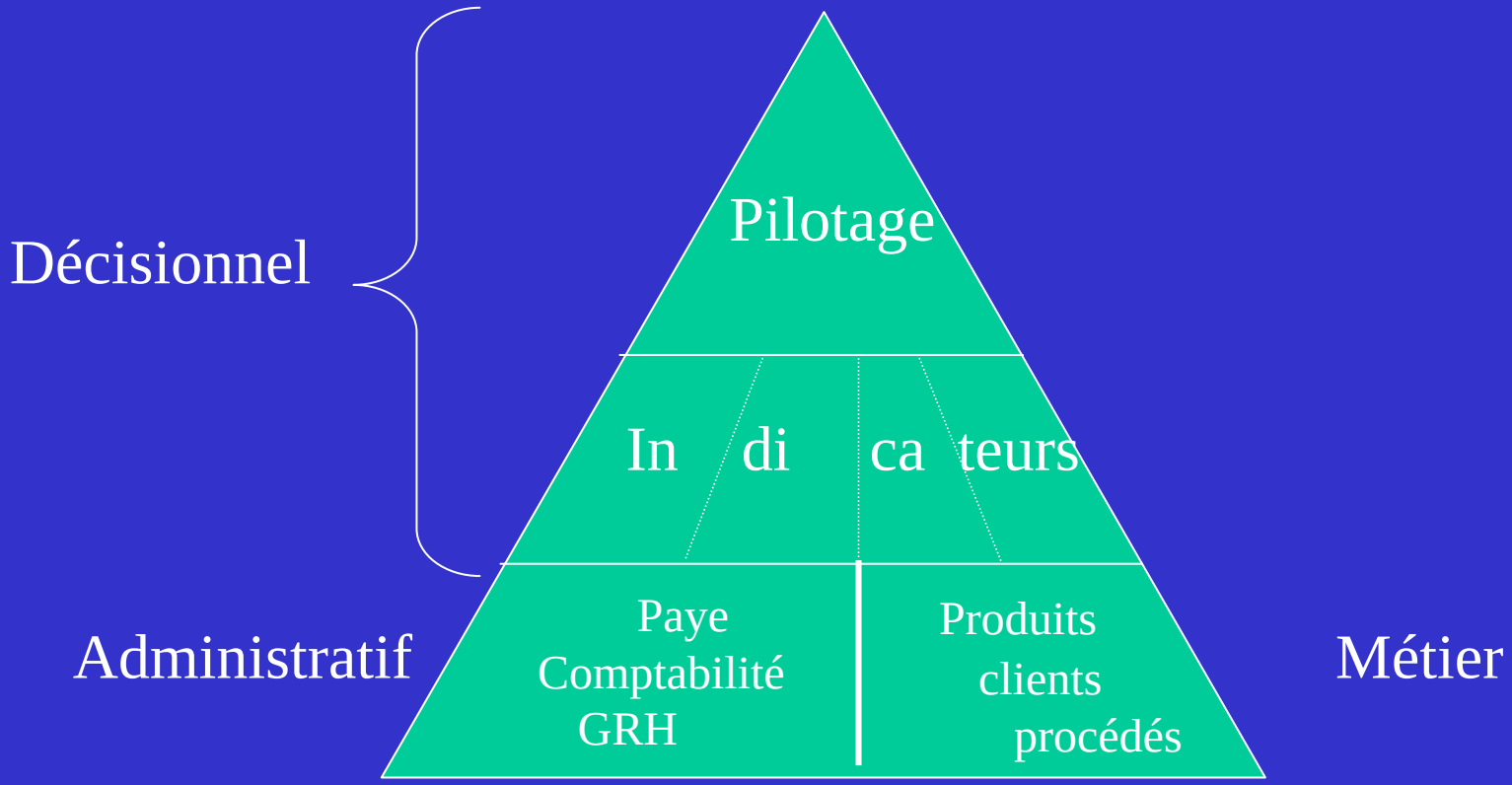
Les données sont historisées et non volatiles

- **historisation** : suivre dans le temps l'évolution des différentes valeurs des indicateurs.

➔**couches de données**

- **non volatiles :** traçabilité ➔ non suppression

La pyramide du système d'information



NB : Un système d'information réellement complet intègre des informations et des contraintes extérieures.

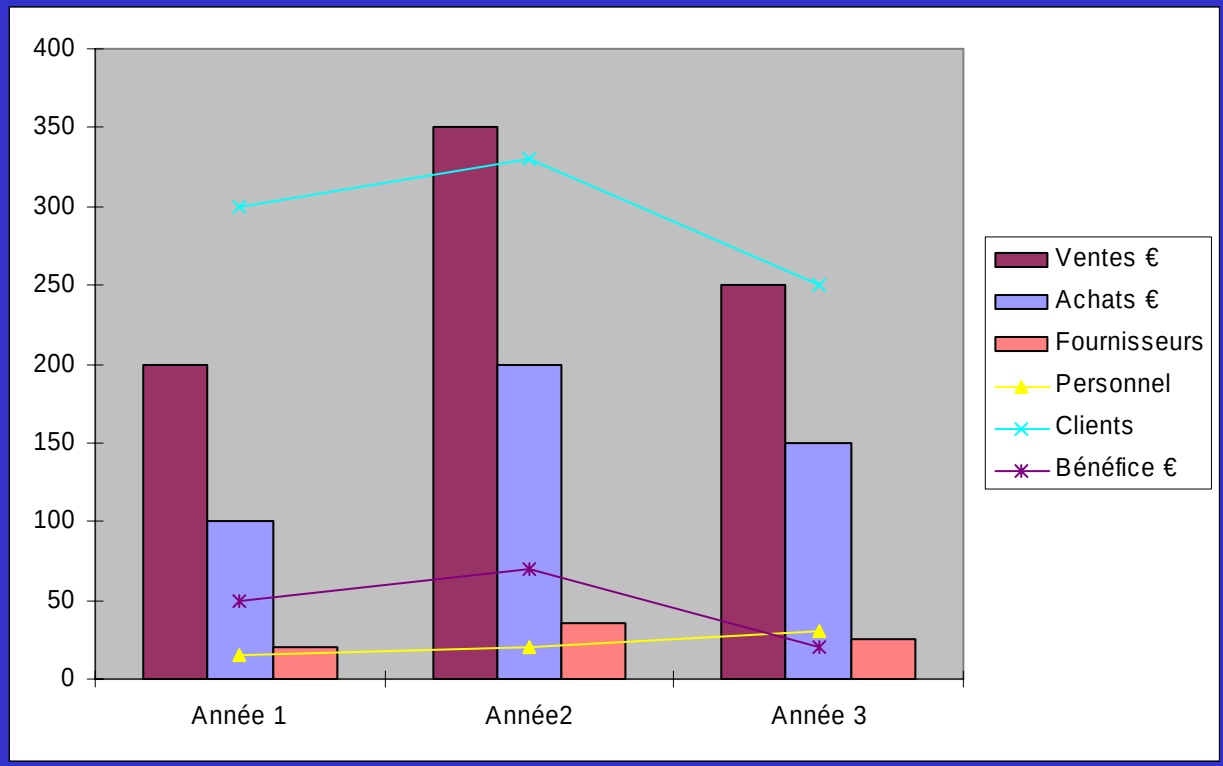
Intérêt de l’entrepôt de données

- Vision transversale de l’entreprise
- Intégration des différents bases
- Données non volatiles (pas de suppression)
- Historisation
- Organisation vers prise de décision

Un projet complexe

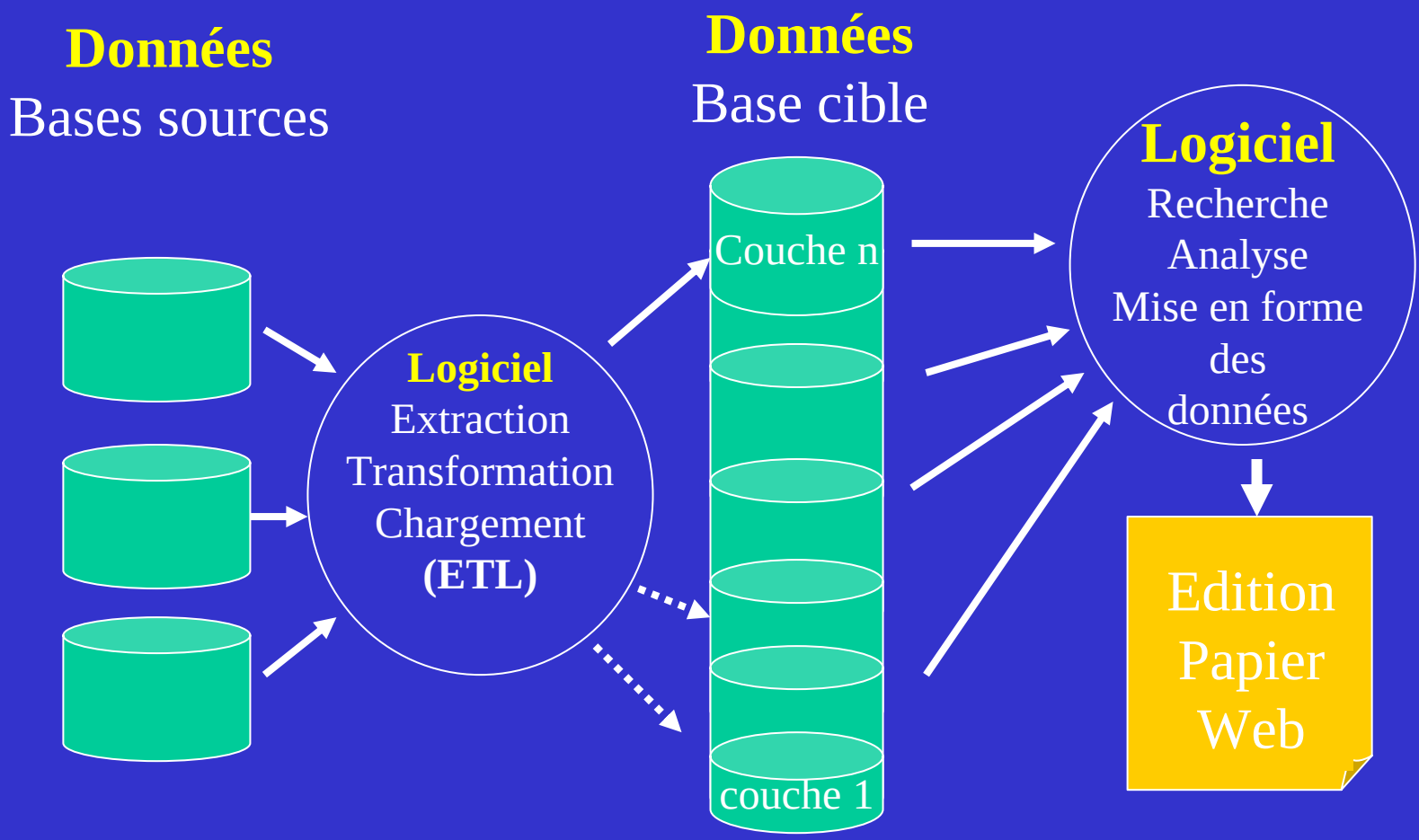
- Rassembler des données hétérogènes
- Les homogénéiser et les restructurer
- Vérifier leur fiabilité
- Les éditer (publier)

Exemple de tableau de bord



Exemples d'indicateurs d'entreprise (succursales multiples) L'année est une dimension, le magasin pourrait être une seconde dimension, le type de produit une troisième.

Schéma de principe d'un ED



Données opérationnelles vs décisionnelles

Données opérationnelles	Données décisionnelles
Orientées application, détaillées, précises au moment de l'accès	Orientées activité (thème, sujet), condensées, représentent des données historiques
Mise à jour interactive possible	Pas de mise à jour interactive
Accès par une personne à la fois	Utilisées par l'ensemble des analystes
Haute disponibilité en continu	Haute disponibilité ponctuelle
Uniques(pas de redondance en théorie)	Peuvent être redondantes
Petite quantité de données utilisées par un traitement	Grande quantité de données utilisée par les traitements
Réalisation des opérations au jour le jour	Cycle de vie différent
Forte probabilité d'accès	Faible probabilité d'accès
Utilisées de façon répétitive	Utilisées de façon aléatoire

Classes de données d’un ED

- **Métadonnées** (« données sur les données »)
- **Données détaillées** : données intégrées dans l’ED
- **Données agrégées** : sommation de données détaillées (tables d’agrégat)
- **Couches de données** : historisation

Exemple d'Entrepôt de données

L'ED doit fournir le CA des ventes d'un produit, par date, client, et vendeur, ainsi que toutes les sommations possibles de chiffre d'affaires dans une année donnée.

Une vente est caractérisée par : **produit, client, vendeur, date, prix de vente**

- **produit** **code produit, code famille (et libellés)**
- **client** **code client, type client**
- **vendeur** **code vendeur, nom, code service**
- **date** jour, **semaine, mois**

Modélisation de la base cible

le modèle relationnel est adapté aux transactions.

Dans un entrepôt de données, prévu pour l'aide à la décision, pas de modélisation en forme normale

Modèle dimensionnel :

tables de faits + tables de dimensions reliées à une table de faits par une jointure.

→ On obtient un **schéma en étoile** (dans le cas le plus simple)

Le modèle dimensionnel

- Une table contenant une *clé multiple*, la **table de faits**,
- un ensemble de tables secondaires, les **tables de dimension** (chacune possède une *clé primaire unique* correspondant à l'un des composants de la clé multiple de la table de faits).
- jointures ➔ schéma en étoile

Schéma en étoile de l'exemple

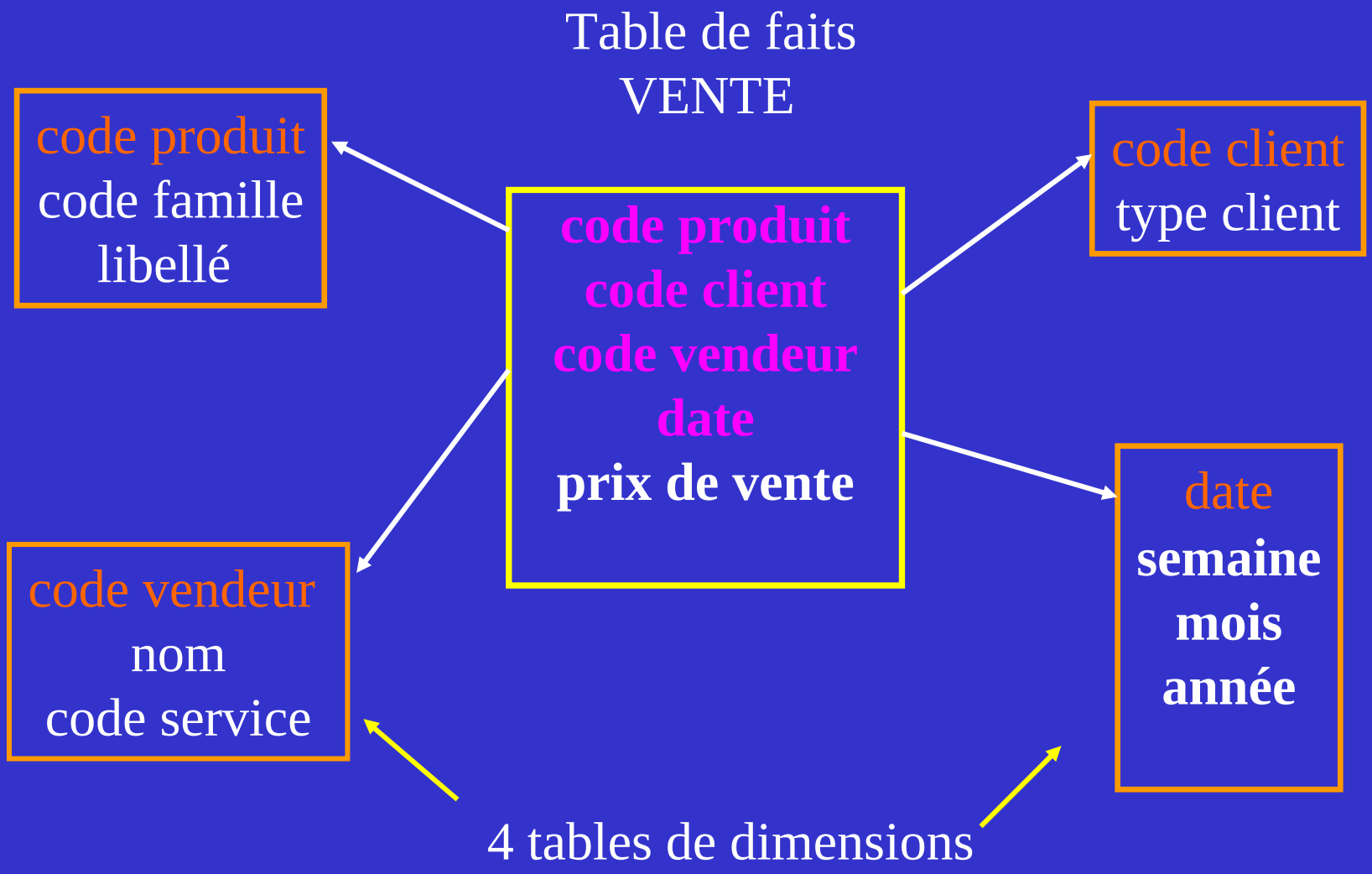


Table de faits

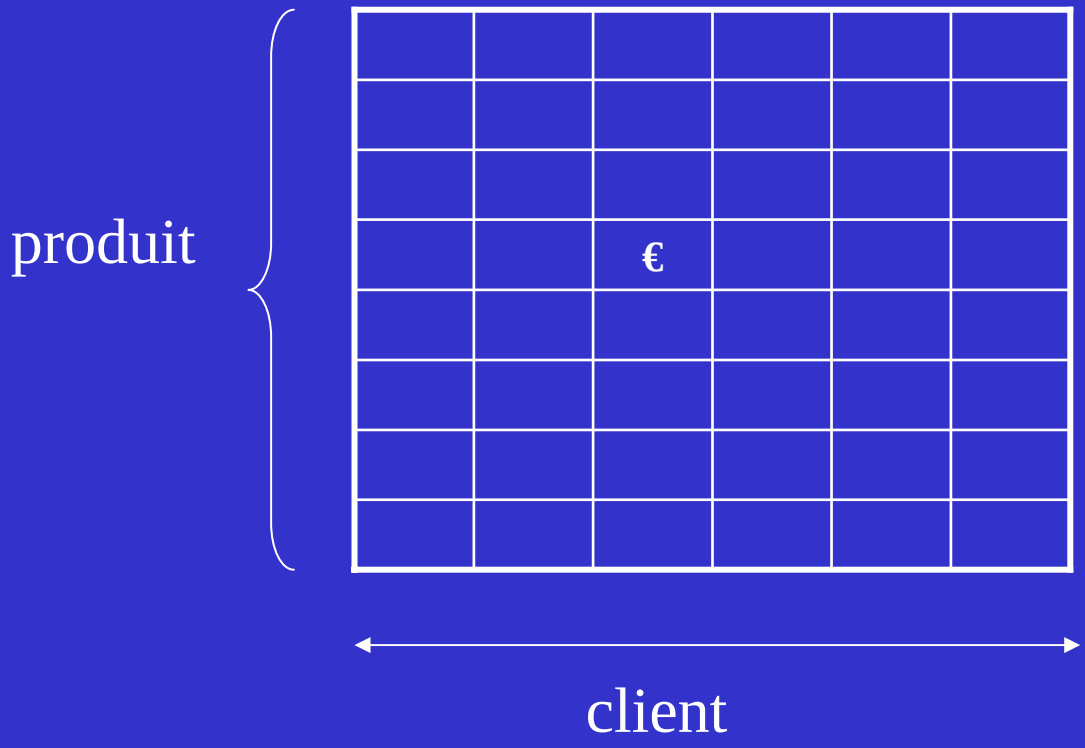
- Contient des faits numériques
- Les faits les plus utiles sont numériques et additifs.
- les agrégats, ou tables d’agrégat, sont des enregistrements récapitulatifs.

Avantages du modèle dimensionnel

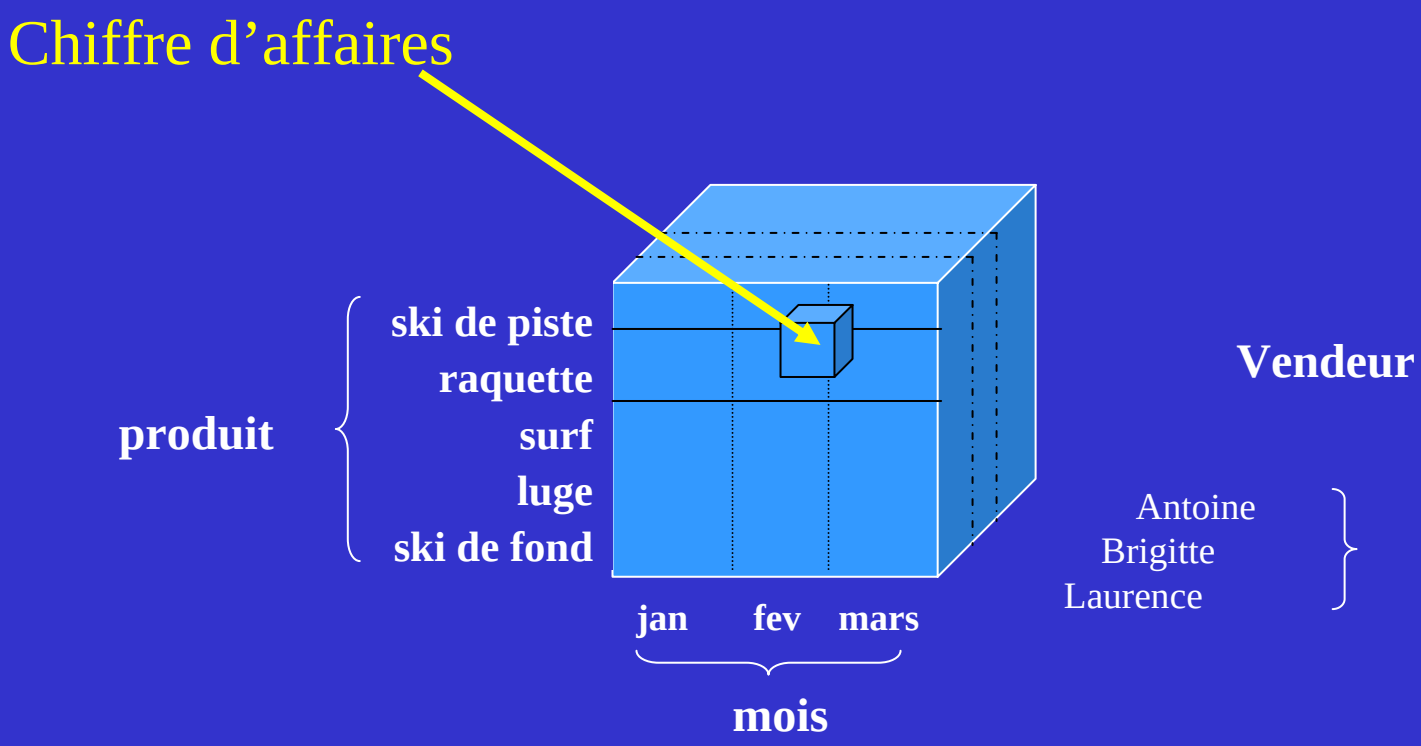
- Conçu pour un requêteur : performances;
- Peut être modifié sans peine (faits nouveaux, dimensions nouvelles ,attributs dimensionnels nouveaux, granularité variable);
- Doit être capable d'intégrer de nouvelles sources.

Résultat possible

Tableau des ventes par produit et par client :



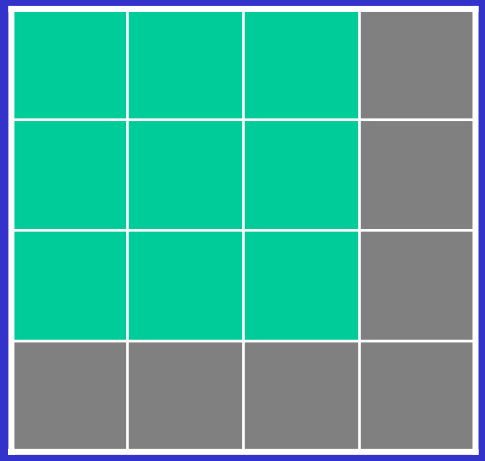
Le « cube » de données



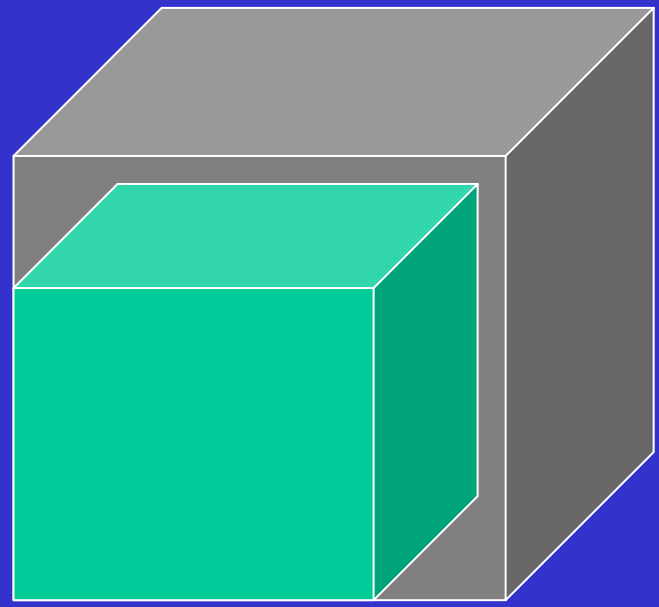
Volumétrie

L’information de synthèse peut être plus volumineuse que l’information de détail

2D



3D



Exemple de volume

Tableau 3x3 ➔ 9 éléments de détail

7 éléments de totalisation

Tableau 3x3x3 ➔ 27 éléments de détail

éléments de totalisation :

3 « tranches » 3x3 = 3x7 = 21

+ la face avant 4x4 = 16

Total = 37

Cube de données

Dans un modèle dimensionnel, on cherche à représenter les données dans un cube (ou hypercube).

- analyse ascendante : « synthétiser »
- analyse descendante : « détailler »
- rotation des dimensions et coupe : « trancher le cube »

Dimensions et indicateurs

- ***dimension*** : - produit
  - client
  - vendeur
  - date
- ***indicateur*** : - chiffre d'affaires

une dimension prend une liste de valeurs,

un indicateur est un nombre.

Additivité des indicateurs

- indicateurs : de préférence numériques et additifs.
- Certains sont semi additifs (additifs pour certaines dimensions).
- Non additifs : fonctions d'agrégations : moyenne, ratios, comptage de lignes.

Exemple d’additivité

Table de faits VENTE

code produit
code client
code vendeur
prix de vente
ristourne client (type client)
commission vendeur (CA vendeur)
retour oui/non

Cas général : Hypercube

- *Indicateur associé à n dimensions (n > 3)*
- *Le « cube » est alors appelé hypercube.*
- *On fixe n-2 dimensions.*
- *En général, plusieurs indicateurs.*
- *Ici, un deuxième indicateur pourrait être le bénéfice.*

Hiérarchie de dimensions

- Dimensions

mois

| ***semaine***

| ***jour***

Analyse multi-dimensionnelle

- *Rotation des dimensions et tranchage,*

- *Analyse descendante ou ascendante.*

Quelques définitions (1)

Entrepôt de données (ED) ou Datawarehouse :
Système d'information agrégeant des données non volatiles et historisées, dans un but d'aide à la décision.

Datamart : ED spécialisé « métier », ou ED partiel

Datamining :
Recherche et analyse d'information dans une base de données.

Quelques définitions (2)

Analyse ascendante/descendante (Drill up, drill down) :
Action d'obtention d'un niveau de détail plus global (ascendant - up) ou plus fin (descendant - down) sur un axe d'analyse multidimensionnel.

Decision Support System (DSS) : Système informatique conçu pour l'aide à la décision plutôt que pour la gestion.

Hypercube : Espace d'analyse à n dimensions (par ex. temps, lieu, produit, vendeur, prix,...).

Quelques définitions (3)

Métadonnées ou données sur les données :
Structure, contenu, localisation, règles d'agrégation
et de transformation des données.

On line analytical process (OLAP) :
Requêtage, analyse et présentation des données d'un ED.
Basé sur le « cube ».

Modèle dimensionnel : modélisation faits – dimensions
alternative à la modélisation entité / relation

Bibliographie

Manuel du designer V5, Business Objects, 1999.

Piloter l'entreprise grâce au data warehouse, J.-M. Franco et al., Eyrolles 2001.

La construction du datawarehouse, J.-F. Goglin, Hermès 1998.

Building the Data Warehouse, W. H. Inmon, Wiley 1996.

Entrepôts de données, guide pratique du concepteur, R. Kimball, Wiley 1997.

Concevoir et déployer un data warehouse, R. Kimball et al., Eyrolles 2000.

http://www.tdwi.org

Entrepôts de données

Modélisation dimensionnelle

- La modélisation dimensionnelle souvent appelée modélisation OLAP (Codd 1993) se présente comme une alternative au modèle relationnel. Il correspond mieux aux besoins du décideur tout en intégrant la modélisation par sujet.
- C’est une méthode de conception logique qui vise à présenter les données sous une forme standardisée intuitive et qui permet des accès hautement performants. Elle aboutit à présenter les données non plus sous forme de tables mais de *cube* centré sur une activité.

Modélisation entité-relation

- Éliminer la redondance des données
- Adaptée aux transactions (ex : mise à jour d'une adresse client), mais pas aux interrogations
- Modèle complexe : des milliers de tables
- Pas de compréhension pour l'utilisateur
- Nécessité de performances

Modélisation dimensionnelle

- Une table de faits (clé multiple)
- Tables de dimensions : clé primaire unique qui correspond à l'un des composants de la clé multiple de la table de faits

➔ jointure en étoile

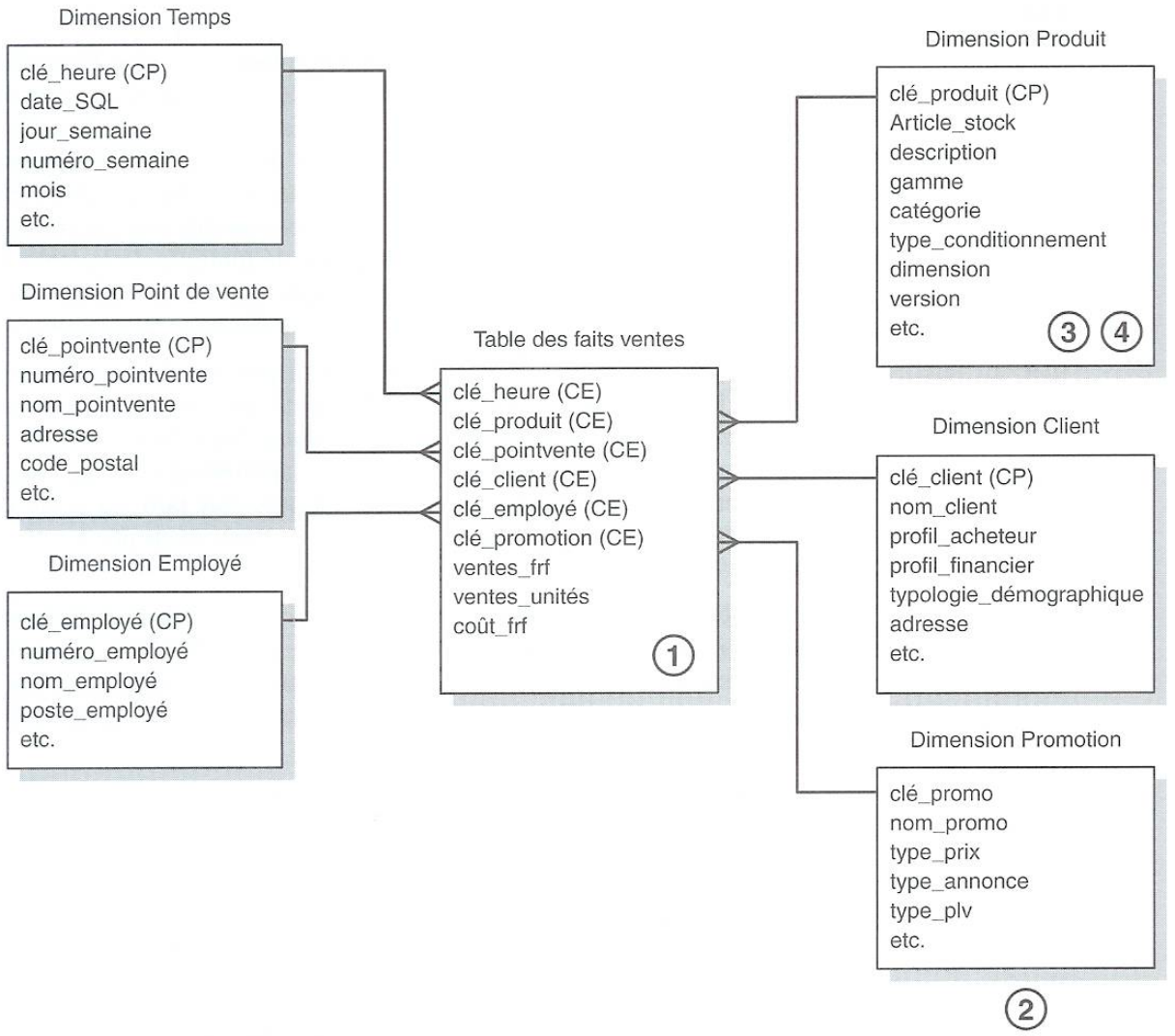
Table de faits

- Contient un ou plusieurs faits numériques qui se produisent pour la combinaison de clés dé finissant chaque enregistrement

- Ex.

  Vente_euros, Vente_unités, Coût_euros

Ex. : vente en grande distribution



Un schéma entité-relation = plusieurs tables de faits

- Détecter les « processus métier » et les modéliser l'un après l'autre
- Sélectionner les relations n,n contenant des faits numériques et additifs. En faire autant de tables de faits
- Les tables dimensionnelles reliées à plusieurs schémas sont dites ***conformes*** (partagées)

Avantages de la modélisation dimensionnelle

- Structure prévisible
- Structure standardisée

➔Requêteur simple et optimisé

- Toutes les dimensions sont équivalentes

➔ Symétrie

Evolutions du modèle

- Ajout de faits nouveaux possible (si compatible avec grain)
- Ajout d'une dimension nouvelle (si une seule valeur par enregistrement de la table de faits)
- Ajout d'attributs dimensionnels
- Outils d'agrégat (enreg. récapitulatifs)

Planification

- Construction de l'ED datamart par datamart, pour éviter une trop grande complexité
- Eviter les « tuyaux de poêle »

- ➔ élaborer un « bus décisionnel » grâce aux dimensions conformes

Dimension conforme

- Une table de dimension en relation avec plusieurs tables de fait est dite *conforme*
- Cohérence des interfaces utilisateurs et des contenus
- Cohérence de l'interprétation des attributs

➔ Grande importance dans la conception

Fait conforme

- Fait ayant la même définition dans tous les datamarts, même unité de mesure, même contexte dimensionnel.
- Ex pour le fait « *recettes* » : périodes, régions de ventes cohérentes
- Ex : conditionnement en unités et boîtes

- Fait non conforme : noms distincts

Exemple de l'agence de voyage
1 - voyages aériens

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.

Schéma dimensionnel
1 - voyages aériens

Exemple de l’agence de voyage
2 – location de voiture

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.

Schéma dimensionnel
2 – location de voiture

Exemple de l’agence de voyage
3 - hôtel

Dans le cas de l’hôtellerie, on veut des tableaux de bord par client, hôtel, ville, date de début de séjour, faisant apparaître le nombre de nuitées et le prix total payé .

Schéma dimensionnel
3 - hôtel

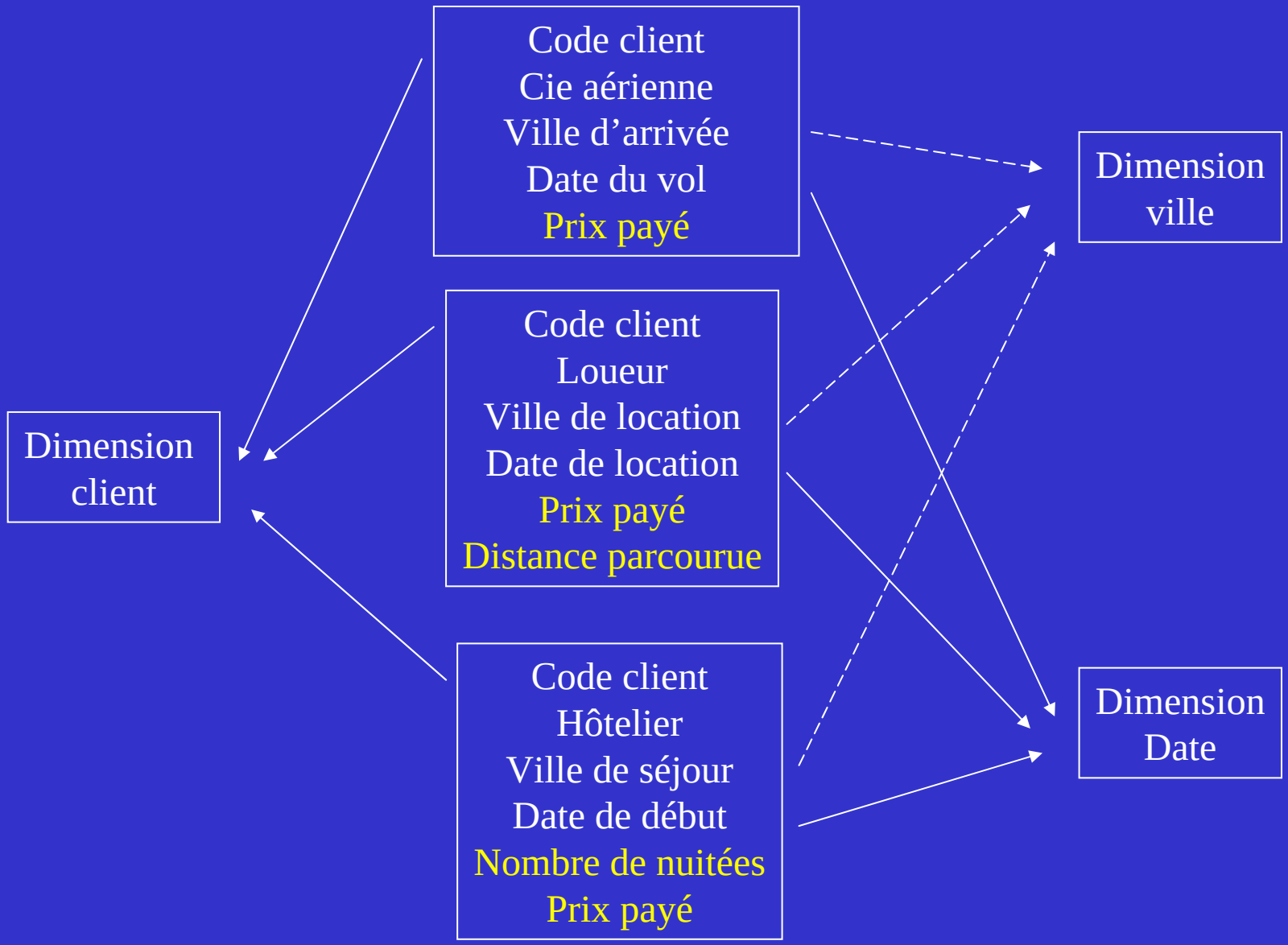
Exemple de l’agence de voyage regroupement

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 veut pouvoir éditer les détails de CA par période de temps et par client, ville, 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/ou hôtellerie, si elles existent, à partir d’un déplacement en avion. Un voyage en avion n’implique pas forcément location de voiture et/ou hôtellerie, et inversement.


Code client
Cie aérienne
Ville d'arrivée
Date du vol
Prix payé
Dimension ville
Code client
Loueur
Ville de location
Date de location
Prix payé
Distance parcourue
Dimension client
Code client
Hôtelier
Ville de séjour
Date de début
Nombre de nuitées
Prix payé
Dimension Date

Niveau de détail

- On privilégie le niveau le plus fin

- Evolutivité
- Puissance
- Efficacité du Data mining

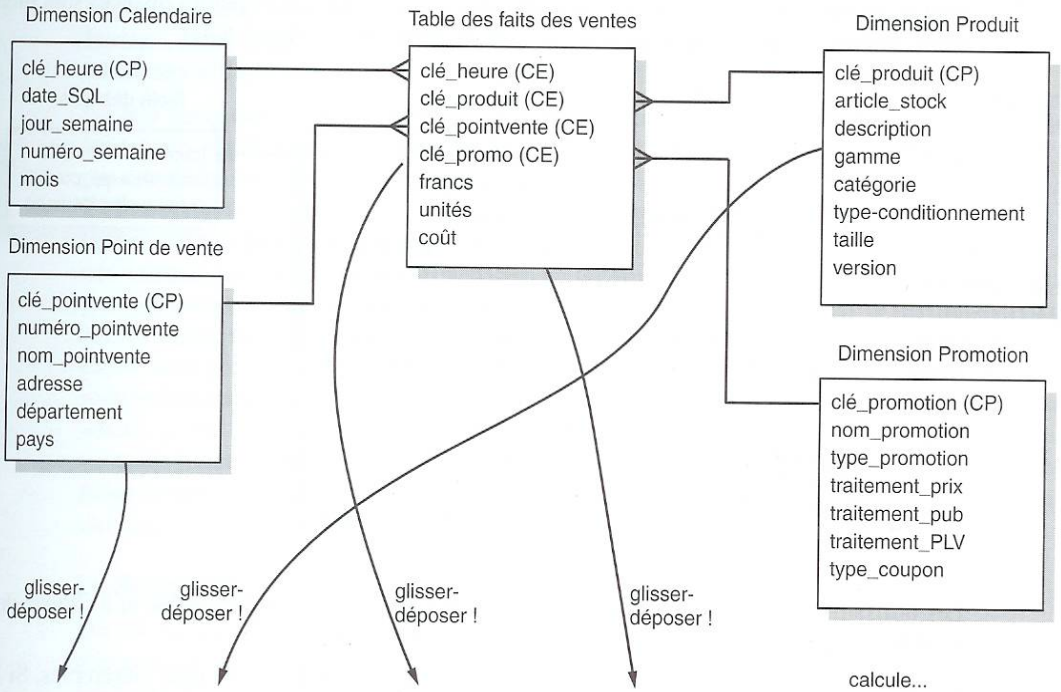
Multisources

- Commencer par un datamart monosource
- Réparer les tuyaux de poêle : chercher à créer des dimensions conformes
- Bus décisionnel : dimensions et faits conformes

Dimensions, faits et attributs

- Plausible : 10 dimensions pour un datamart
  - Si 2 ou 3 dimensions : le concepteur en a-t-il oubliées ?
  - Si 20 dimensions : devient trop complexe
- Un fait est une « observation du marché », la plupart du temps un champ numérique de la source
- L'attribut est un champ textuel (apparaît dans les dimensions)

Parcourir les dimensions



Département	Marque	Total Francs	Coût total	Marge brute
Loire-Atlantique	Lave Vit	F 5 422,00	F 4 821,00	F 601,00
Loire-Atlantique	Lave Plus	F 12 323,00	F 11 422,00	F 901,00
Loire-Atlantique	Extra	F 4 110,00	F 3 648,00	F 462,00
Creuse	Lave Vit	F 4 646,00	F 3 861,00	F 785,00
Creuse	Lave Plus	F 10 956,00	F 9 633,00	F 1 323,00
Creuse	Extra	F 3 520,00	F 2 990,00	F 530,00

Figure 5.3

Relations entre un schéma de modélisation en étoile et un état.

Forage

- Forage vers le bas = Drill down = donner des détails

- Forage vers le haut = Drill up = sommer

Un véritable **forage** mélange les attributs hiérarchisés et non hiérarchisés de toutes les dimensions disponibles

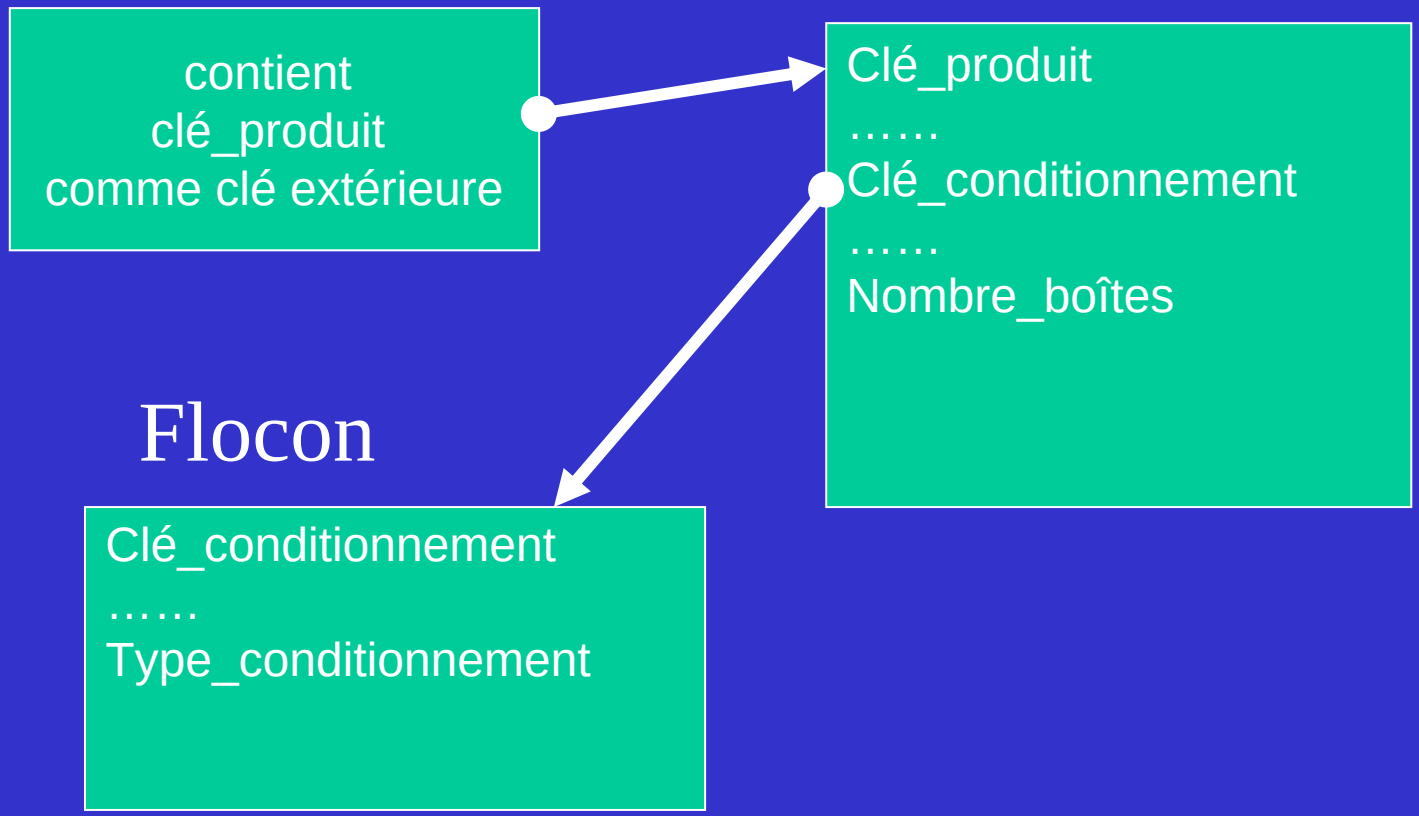
Floconnage

- Définition : dimension dont les champs à faible cardinalité sont dans des tables séparées, reliées à la table d'origine au moyen de clés artificielles.
- Non recommandé : performances, complexité
- Gain en espace disque non déterminant

Exemple de floconnnage

Table de faits

Table de dimension



Attributs des tables de dimensions

- Pas de codes …
- Littéraux (mots complets)
- Descriptifs
- Soignés (orthographe, valeurs)
- Indexés
- Documentés (métadonnées)

Métadonnées

- Ensemble d'informations nécessaires à l'accès, à la compréhension et à l'exploitation des données du data warehouse.

- Le référentiel de l'entrepôt de données = métadonnées + outils d'administration

  Il collecte l'ensemble des modèles de données nécessaires à la construction et à l'exploitation du data warehouse.

Dimension temps

- Clé_date (clé principale)
- Date complète
- Jour de la semaine
- Numéro du jour dans la semaine
- Numéro du jour dans l’année
- ……
- Indicateur jour ouvrable
- Indicateur dernier jour du mois

Sous dimension temps (flocon)

- Clé_date (clé principale)
- Pays (clé principale)
- Indicateur de jour férié
- Indicateur de fête religieuse
- Indicateur de fête civile
- …….

Dimensions changeantes

- Clé_produit ou clé_client ne changent pas, mais les attributs évoluent.
- On peut :

1. Réécrire sur l'enregistrement (historique perdu)
2. Ajouter un enreg. Avec nouvelle valeur de clé
3. Créer un nouveau champ « ancien » dans l'enreg et y stocker l'ancienne valeur d'attribut

Dimensions dégénérées

- N° bon de commande, n° de facture ?
- Souvent à conserver dans la base de faits, mais pas d'attributs associés

➔ pas de table de dimension associée

Clés

- Toutes les clés : clés « de substitution » dépourvues de signification, et sans rapport avec les bases sources.

Faits

- Additifs de préférence
- Faits textuels (exemple : conditions météos lors d’un relevé de sinistre d ’assurance) = peu d’intérêt de comptage et regroupement si texte libre
- Préférer les « témoins »

Conception d'une table de faits

1. Commencer par un data mart à source unique
2. Granularité : un enreg. de la table de faits sera créé pour chaque :
   a. Vente
   b. Déclaration de sinistre
   c. Inscription d'étudiant
   d. Ligne de facture

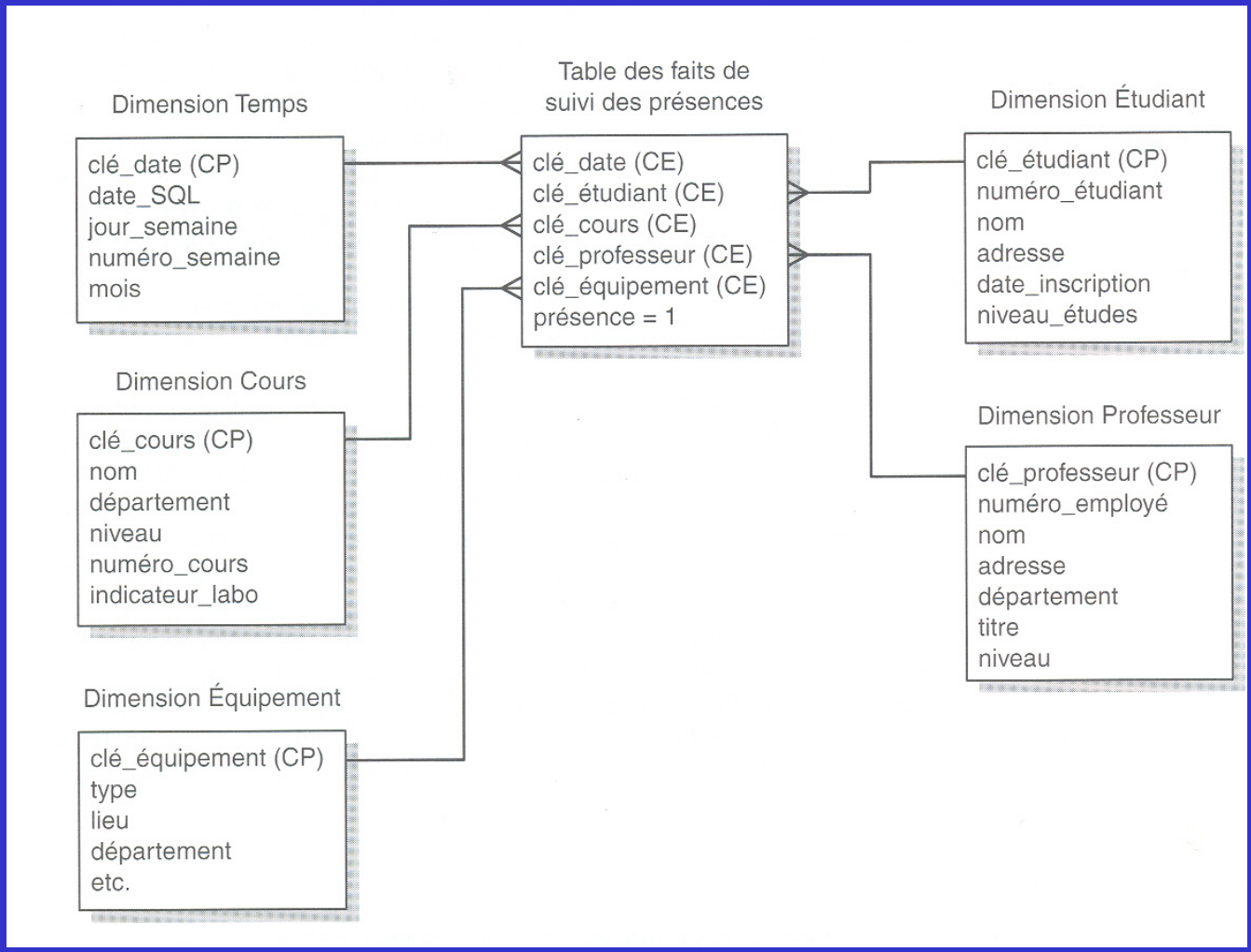
Conception d'une table de faits (suite)

3. Choisir les dimensions : descripteurs à valeur unique (de préférence)
4. Choisir les faits : les faits (indicateurs) doivent correpondre à la granularité de la table de faits

Agrégat

- Table récapitulative (sommations) destinée à améliorer les performances du requêtage.
- Il s’agit d’une table de faits qui possède des tables dimensionnelles.

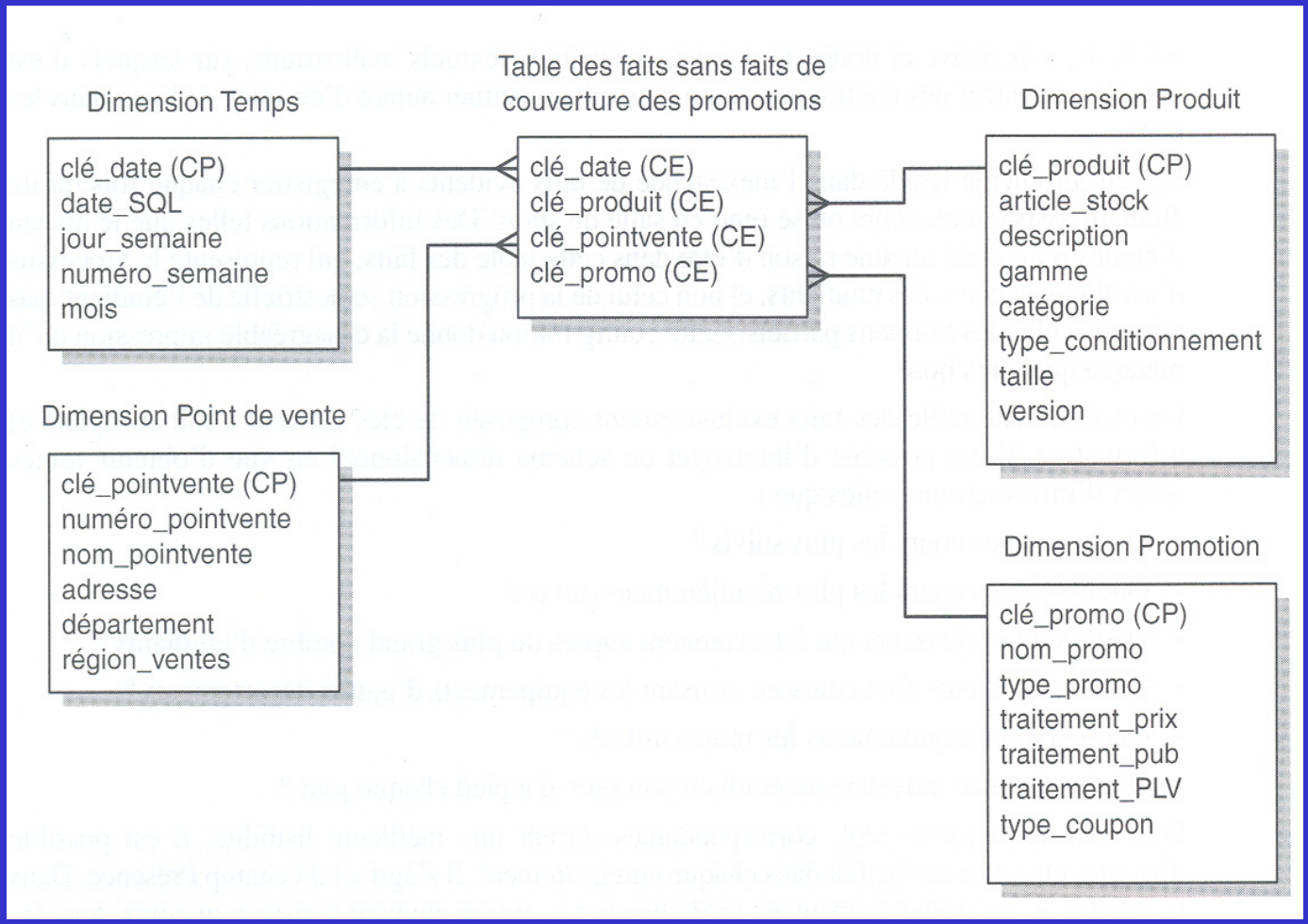
Table de faits sans fait exemple 1



Interrogations sur l’exemple 1

1. Cours les plus suivis ?
2. Cours les plus régulièrement suivis ?
3. Profs qui ont le plus d’étudiants ?
4. Profs qui utilisent des équipements d’autres départements ?
5. Equipements les moins utilisés ?

Table de faits sans fait exemple 2

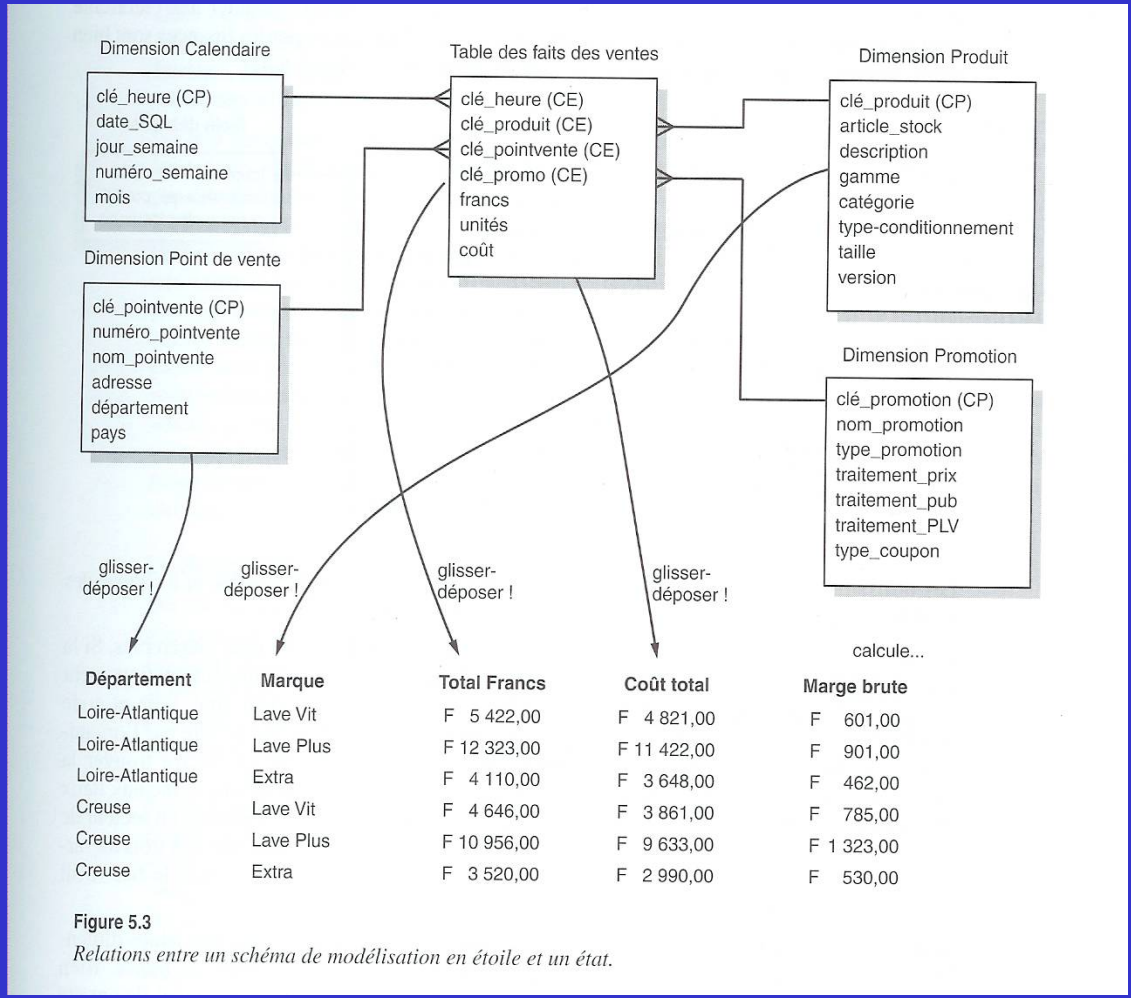


Interrogations sur l’exemple 2

Quels sont les produits qui étaient en promotion et ne se sont pas vendus ?

1. Consultation de la table « de couverture » : liste des produits en promotion tel jour dans tel point de vente
2. Consultation de la table des ventes pour recenser les produits qui se sont vendus

Table des ventes de l'exemple 2



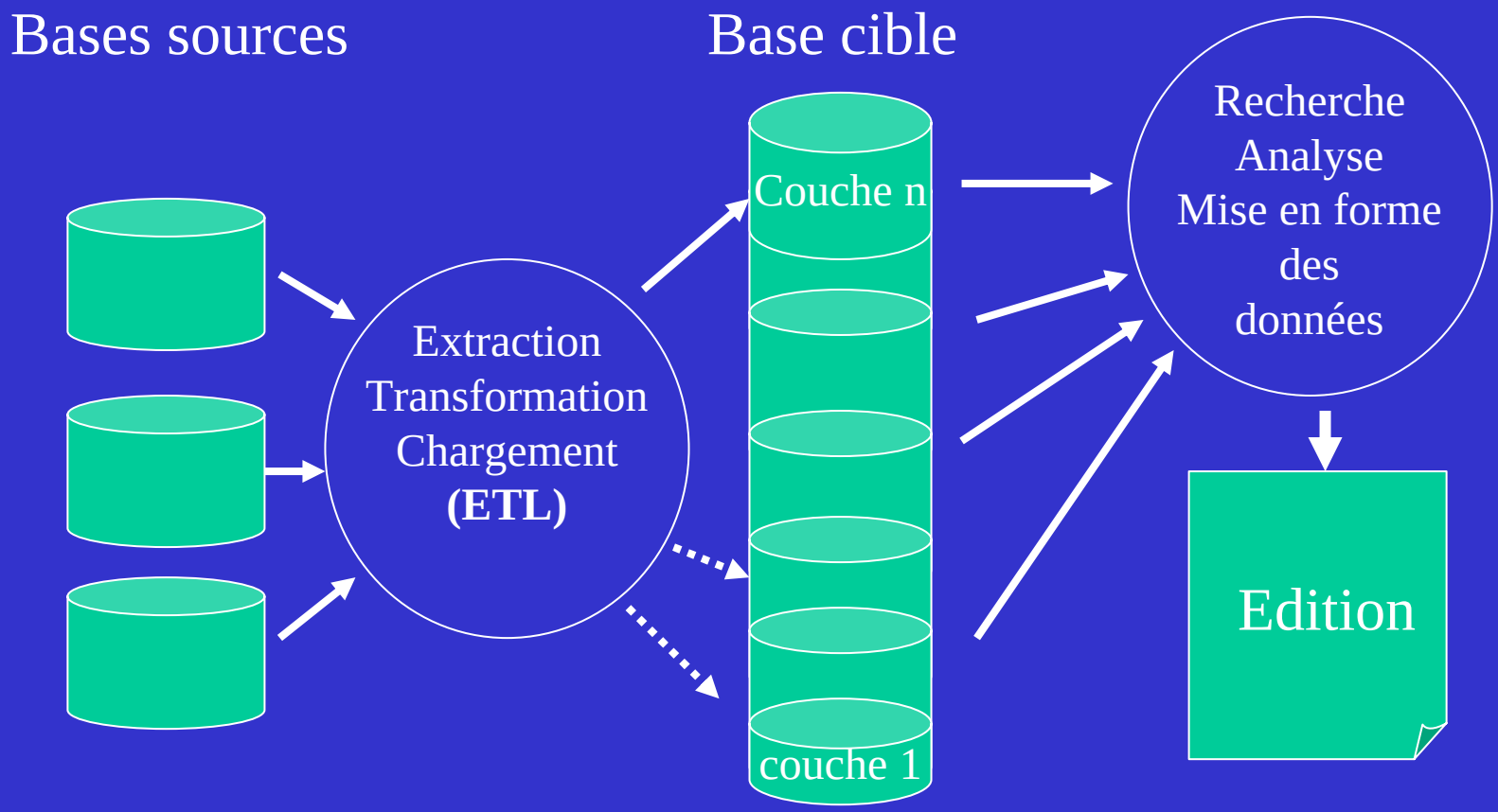
Département	Marque	Total Francs	Coût total	Marge brute
Loire-Atlantique	Lave Vit	F 5 422,00	F 4 821,00	F 601,00
Loire-Atlantique	Lave Plus	F 12 323,00	F 11 422,00	F 901,00
Loire-Atlantique	Extra	F 4 110,00	F 3 648,00	F 462,00
Creuse	Lave Vit	F 4 646,00	F 3 861,00	F 785,00
Creuse	Lave Plus	F 10 956,00	F 9 633,00	F 1 323,00
Creuse	Extra	F 3 520,00	F 2 990,00	F 530,00

Figure 5.3

Relations entre un schéma de modélisation en étoile et un état.

Analyse des données de l’ED : Exemple de B.O.

Schéma de principe d'un ED



L’outil BO

Business Objects : extraire, mettre en forme et analyser les données de la base cible.

→ intranet de l’établissement

- univers = architecture des informations extraites.
- utilisateurs : créent de nouvelles requêtes, paramétrent et en exécutent des requêtes existantes.

Univers BO

Le designer d’univers :

Informaticien

Compétences

» SQL

» BD de l’entreprise

Développement d’un univers

- Planification
  - découper le SI en domaines
- Analyse
  - Analyser les besoins des utilisateurs
- Conception
  - Faire le schéma conceptuel; spécifier l’univers
- Application
  - Créer l’univers
- Maintenance

Les objets dans BO

- Indicateur : nombre
- Dimension : entité (une variable)

Exemple

UFR / Ens.		Chimie	
Prof Univ		5	

Classes et univers

- Les objets peuvent être regroupés en classes et sous-classes
- L’ « univers » est un ensemble d’objets créés sur une BD

Hiérarchies

- Une dimension appartient souvent à une hiérarchie.
- La hiérarchie ordonne la dimension et permet de changer le niveau d’analyse.

Exemple de hiérarchie

- Exemple

Hiérarchie dans l'organisation de la scolarité.

Analyse dimensionnelle

- Elle est possible si les dimensions de l'univers sont hiérarchisées

- Exemple : nombre d'étudiants par composante,....., étape
- ➔ changement du niveau d'analyse

Création d’un univers

- La structure d’un univers doit s’adapter à la logique de l’utilisateur, et non à la structure des tables de la BD.

- Autant que possible, les dimensions doivent être définies dans des hiérarchies.

Objets de l’univers

- Normaliser les noms des objets

  ex. : Nationalité-code

  Nationalité-libellé
- Utiliser la langue du métier

- Chaque dimension doit avoir une **liste de valeurs** qui associe code et libellé.

Structure de la base ED

- Objectif : produire des indicateurs au carrefour des dimensions
- Une table de faits comporte des dimensions et des indicateurs
- A chaque dimension d'une table de faits, on associe une table de dimension (sauf dimension dégénérée)

La dimension historique

- Elle est sous-jacente dans tous les faits : date, trimestre, année selon les faits mesurés : nb d'inscriptions, budget, appartenance à un groupe.
- La date d'extraction est également importante.

➔ 2 dimensions historiques :

les faits, l'extraction

Création des tables de faits

- Si les dimensions sont dans des hiérarchies différentes, on peut avoir intérêt à créer plusieurs tables de faits.

Structure des tables de dimension

- Il peut y avoir une hiérarchie liée à la dimension. Ex. : Nationalité

code-nat ➔ Lib-nat

code-groupe de pays ➔ Lib-groupe

Schéma en ***flocon***

Séparation indicateur et dimension

- mieux vaut créer une variable avec liste de valeurs, plutôt que des indicateurs.
- l'ajout d'une nouvelle valeur est alors simple.
- Il est plus simple de modifier des

valeurs (lignes)

plutôt que des

structures (colonnes)

Exemple indicateur et dimension

Si l'on créé les indicateurs :

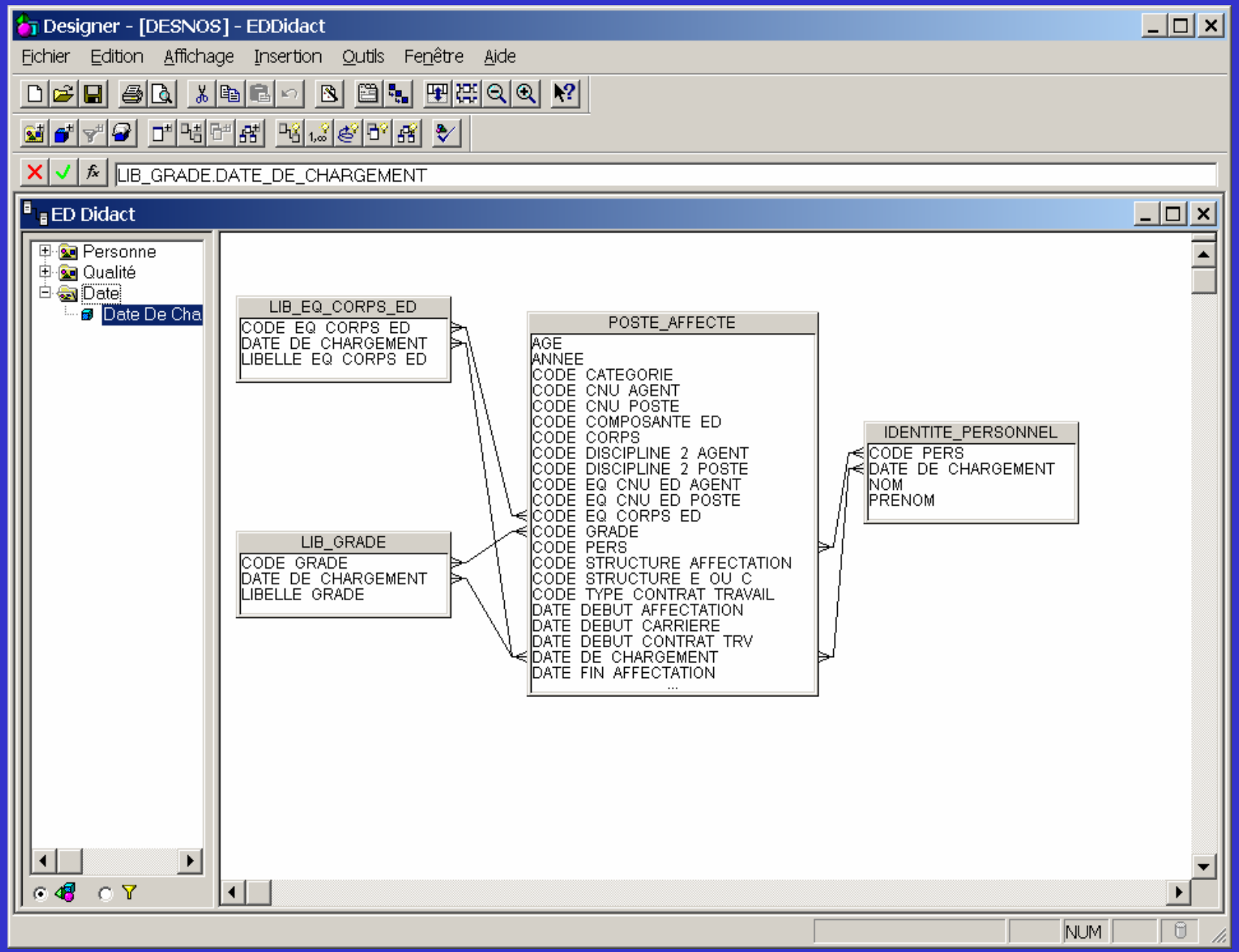
- professeur
- assistant
- On veut créer « agrégé », il faut modifier la table de faits

Si l'indicateur est « enseignant », on modifie seulement la liste de valeurs.

Agrégats

- Deux solutions :

1. Tables d’agrégats = calculs pré-établis optimisant les performances
2. Fonction *aggregate aware* de BO


Designer - [DESNOS] - EDDidact
Fichier Edition Affichage Insertion Outils Fenêtre Aide
LIB_GRADE.DATE_DE_CHARGEMENT
ED Didact
Personne
Qualité
Date
Date De Cha
LIB_EQ_CORPS_ED
CODE EQ CORPS ED
DATE DE CHARGEMENT
LIBELLE EQ CORPS ED
LIB_GRADE
CODE GRADE
DATE DE CHARGEMENT
LIBELLE GRADE
POSTE_AFFECTE
AGE
ANNEE
CODE CATEGORIE
CODE CNU AGENT
CODE CNU POSTE
CODE COMPOSANTE ED
CODE CORPS
CODE DISCIPLINE 2 AGENT
CODE DISCIPLINE 2 POSTE
CODE EQ CNU ED AGENT
CODE EQ CNU ED POSTE
CODE EQ CORPS ED
CODE GRADE
CODE PERS
CODE STRUCTURE AFFECTATION
CODE STRUCTURE E OU C
CODE TYPE CONTRAT TRAVAIL
DATE DEBUT AFFECTATION
DATE DEBUT CARRIERE
DATE DEBUT CONTRAT TRV
DATE DE CHARGEMENT
DATE FIN AFFECTATION
IDENTITE_PERSONNEL
CODE PERS
DATE DE CHARGEMENT
NOM
PRENOM
NUM

ED cours 4

Eléments d’architecture

Administration

Stratégie

Construction

Data warehouse et dataweb

Intranet : réseau de données privatif de l'institution

Architecture 3 (n)-tiers :

- Client léger (navigateur)
- Serveur(s) d'applications
- Serveur(s) de données

Intégration au portail

A partir du client universel, accès à toutes les applications de l'entreprise :

Applis décisionnelles, bureau virtuel, workflow, docflow, BD,…

- Point d'accès unique,
- Architecture simplifiée,
- Réduction des coûts.

Administrer

Qualité de service (disponibilité, fiabilité, sécurité), gestion des anomalies

Configuration, paramétrage, maintenance du système d'information :

- gestion financière, administrative, technique; maintenance et support.

Administration fonctionnelle

- Traduire le besoin de chaque utilisateur
- Assurer la cohérence globale
- Faire évoluer l’ensemble
- Etablir des profils
- Editer : coordonner la diffusion
- Maîtriser la qualité : le métadictionnaire et son évolution

Administration technique

- Volumétrie
- Puissance des traitements
- Evolutions
- ➔ outils d'administration des systèmes, sécurité, sauvegardes, métrologie
- Ces outils s'appuient sur la politique des SI de l'entreprise

Choix stratégiques

- Référentiel intégré
- Solution centralisée ou non
- Espaces privés
- Maîtriser les coûts
- Commencer par un datamart ?

Référentiel intégré

Multiplicité des dictionnaires : ETL, SGBD, OLAP

Pour garantir la cohérence, 1 seul dictionnaire si possible

Solution centralisée ou non

- Si l’organisation comporte un siège et des agences : 2 niveaux d’ED ?
- La solution centralisée est coûteuse (réseau) et moins évolutive
- La solution décentralisée est plus souple, mais plus complexe

Du datamart à l’ED

- Développement RAD
- L’architecture retenue doit permettre le passage à l’échelle

La démarche RAD

Phase 1
- Analyse des besoins
- Mise en œuvre 1ère version
- Retour d'expériences

.....

Phase N
- Besoins complémentaires
- Mise en œuvre Nième version
- Retour d'expériences

Les itérations du RAD

A chaque étape (itération), il faut vérifier les cohérences :

- du dictionnaire
- de l’architecture et des flux (non régression)
- des plate-formes et environnements

La charge de mise en œuvre de l'ED

- La « restitution » de l'information, visible par l'utilisateur, ne représente que 15 %
- La partie immergée :
  - Alimentation de l'ED : 50 %
  - OLAP : 20 %
  - Administration : 15 %

Le chef de projet

- doit être orienté « utilisateurs »
- soutenu par la direction et les utilisateurs
- ne pas promettre la lune
- soigner la conception (modèle, données à intégrer)
- ne pas faire le jouet du président
- Ne pas s’arrêter à la première phase

Les difficultés liées à l'axe « temps »

- Le calendrier
  - Définitions de l'année, du mois; irrégularités
  - Additivité des variables
  - Variations de périodicité
- Consolidation, agrégation (mouvements internes)

Agrégats

Il faut sélectionner les agrégats à retenir :

Compromis

volume

temps de réponse

Référentiel