

Cours de Business Intelligence

Master : Traitement intelligent des systèmes

Préparé par: Mme. ELANSARI Khawla

Année Universitaire: 2021/ 2022

Informatique Décisionnelle

1. Problématiques des entreprises
2. Solutions possibles
3. Système de gestion vs. Système de décision
4. Informatique décisionnelle
5. Applications de la BI
6. Enjeux du décisionnel
7. Exploitation des données
8. L'infrastructure décisionnelle

| *Problématiques des entreprises*

1- Contexte difficile

- La mondialisation qui poursuit son bouleversement des équilibres concurrentiels et économiques;
 - Un environnement financier incertain et complexe
- => Besoin de mettre au point une stratégie SI, susceptible de mieux atteindre les objectifs des entreprises

2- Nouveaux objectifs

- Toute entreprise doit pouvoir anticiper des événements de plusieurs type : évolution du marché, ...
- => Quels outils donner au décideur pour comprendre dimensionner, piloter et gérer

| 1. *Problématiques des entreprises*

Les entreprises ne veulent plus seulement savoir :

« Combien de clients ont acheté tel produit pendant telle période ? »

Mais des nouvelles questions métiers :

- clientèle :

« Quel est leur profil ? »

« Quels autres produits les intéresseront ? »

« Quand seront-ils intéressés ? »

« Comment les conserver ou les faire revenir ? »

- marketing, actions commerciales :

« Où placer ce produit dans les rayons ? »

« Comment cibler plus précisément le mailing concernant ce produit ? »

| 1. *Problématiques des entreprises*

3- Difficultés techniques

Une **grande masse de données** (le nouveau concept Big Data) archivée: Distribuée, Hétérogène, Très Détaillée

A traiter: Synthétiser / Résumer, Analyser, Visualiser

Pour une utilisation par : Des experts et des analystes d'un domaine métier

- NON informaticiens
- NON statisticiens

4- Inconvénients des systèmes transactionnels

- Les requêtes complexes et lourdes dégradent les performances des systèmes transactionnels,
- Les données sont réparties entre données actuelles et données archivées, rendant la vue historique des données très difficile ou impossible

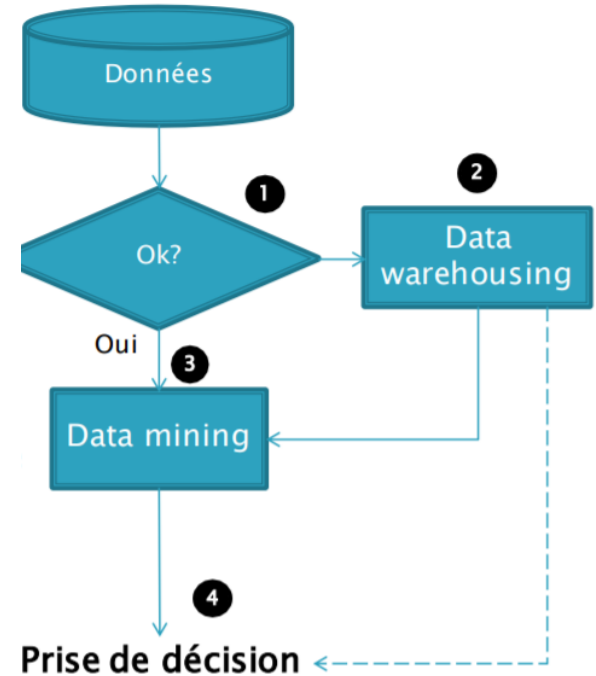
| 2. Solutions possibles

1- **Data Warehouse** : un système d'information centralisé, dédié aux applications décisionnelles.

2- **Data Mining** : un processus de prospection de données offrant un ensemble de techniques (modèles et algorithmes) pour la prédiction, la génération de liens, de patrons et de concepts.

Les données sont-elles prêtes (bien organisées, intégrées, clean, homogènes, ...) pour en extraire de nouvelles connaissances pour la prise de décision ?

- **Si oui** : on fait appel au Data Mining pour une analyse directe.
- **Sinon** : Il faut tout d'abord organiser les données puis faire appel au Data Mining pour la prise de décision => on va préparer les données (BI)



| 3. *Système de gestion vs. Système de décision*

	Système de gestion (opérationnel)	Système de décision (analyse)
Objectifs	dédié au métier et à la production ex: facturation, stock, personnel	dédié au management de l'entreprise (pilotage et prise de décision)
Pérennité	données volatiles ex: le prix d'un produit évolue dans le temps	données historisées ex: garder la trace des évolutions des prix, introduction d'une information daté
Optimisation	pour les opérations associées ex: passage en caisse (lecture de code barre)	pour l'analyse et la récapitulation ex: quels les produits achetés Ensembles
Granularité de données	Totale et atomique, on accède directement aux informations atomiques	agrégats, niveau de synthèse selon les besoins de l'analyse

| 4. *Informatique décisionnelle*

« Le système d'information décisionnel est un ensemble de données organisées de façon spécifiques, facilement accessibles et appropriées à la prise de décision [...].

La finalité d'un système décisionnel est le pilotage d'entreprise.

Les systèmes de gestion sont dédiés aux métiers de l'entreprise [...].

Les systèmes décisionnels sont dédiés au management de l'entreprise [...].»

(Goglin, 2001, pp21-22)

| 5. Applications de la BI

Marketing: Déterminer la population à cibler (âge, sexe, profession, habitation, région, ...).

Secteur bancaire: Utilisation du score de risque pour proposer le montant de crédit le plus adapté à chaque client.

Econométrie: prédiction de trafic autoroutier.

Ressources Humaines: adéquation activité / personnel.

Logistique: adéquation demande / production.

Commerce: ciblage de clientèle - aménagement des rayons

Gestion de stocks: quand commander un produit- quelle quantité demander,

| 5. *Applications de la BI*

- Analyse du comportement de consommateurs ou de citoyens, en fonction de leurs caractéristiques (sexe, age...), de critères socio-économiques (profession...), géographiques...
- Analyse de ventes en fonction de l'implantation géographique de magasins (densité, caractéristiques des régions...), de l'organisation de magasins (rayonnage, marketing, RH...)
- Analyse des structures de paniers (quel produit est vendu en même temps que quel autre à quelles conditions ?)
- Prédiction de ventes en fonctions de données conjoncturelles, gestion des stocks, des approvisionnements
- Contrôle qualité et analyse de défaut des chaînes de production en fonction des centres de production, des organisations, des fournisseurs
- Repérer les clients les plus rentables, par exemple pour une chaîne de supermarchés ou de vêtements, afin de les fidéliser.

| 6. *Enjeux du décisionnel*

La prise de décisions stratégiques dans une organisation nécessite le recours et le croisement de multiples informations qui concernent tous les départements : production, RH, achats, ventes, marketing, service après-vente, maintenance, R&D...

Or ces données sont généralement :

- **éparpillées** au sein des départements et non connectées entre elles
- **hétérogènes** dans leurs formats techniques et leurs organisations structurelles, voire leurs sémantiques
- **implémentées pour l'action** (par construction) et non pour l'analyse
- **volatiles**, au sens où leur mise à jour peut conduire à oublier des informations obsolètes

Exemple:

Un catalogue de produits sera conçu pour permettre de trouver facilement un produit en fonction de caractéristiques précises, de faire des mises à jour rapides et fiables, de gérer des stocks...

Mais un système décisionnel souhaitera : - connaître l'organisation des produits selon certaines caractéristiques et regroupements qui ne sont pas forcément premiers dans la gestion quotidienne ;



L'enjeu des systèmes décisionnels est de donner accès aux données existantes dans l'organisation, sous une forme intégrée, afin de faciliter leur interrogation croisée et massive

| 7. *Exploitation des données*

Les données agrégées dans un système décisionnel servent à trois grandes catégories d'usage :

- La production de rapport récurrents (Reporting)
- L'exploration manuelle
- L'analyse de données (descriptive ou prédictive)

1- **Reporting:**

Le principe du reporting est d'agréger et de synthétiser des données nombreuses et complexes sous forme d'indicateurs, de tableaux, de graphiques permettant d'en avoir une appréhension globale et simplifiée.

Le reporting s'appuie principalement sur les agrégats (GROUP BY en SQL par exemple) afin de faire apparaître des comptages, sommes ou moyennes en fonction de critères d'analyses.

Le reporting est généralement récurrent, le même rapport sera produit à intervalles réguliers pour reportingcontrôler les variations des indicateurs

| 7. *Exploitation des données*

2- **Exploration manuelle:**

Consiste à pouvoir explorer les données de façon peu dirigée (heuristique) afin de trouver des réponses à des questions que l'on ne s'est pas posées (sérendipité). L'idée générale est plutôt que les réponses aux premières questions que l'on se pose conduiront à se poser de nouvelles questions.

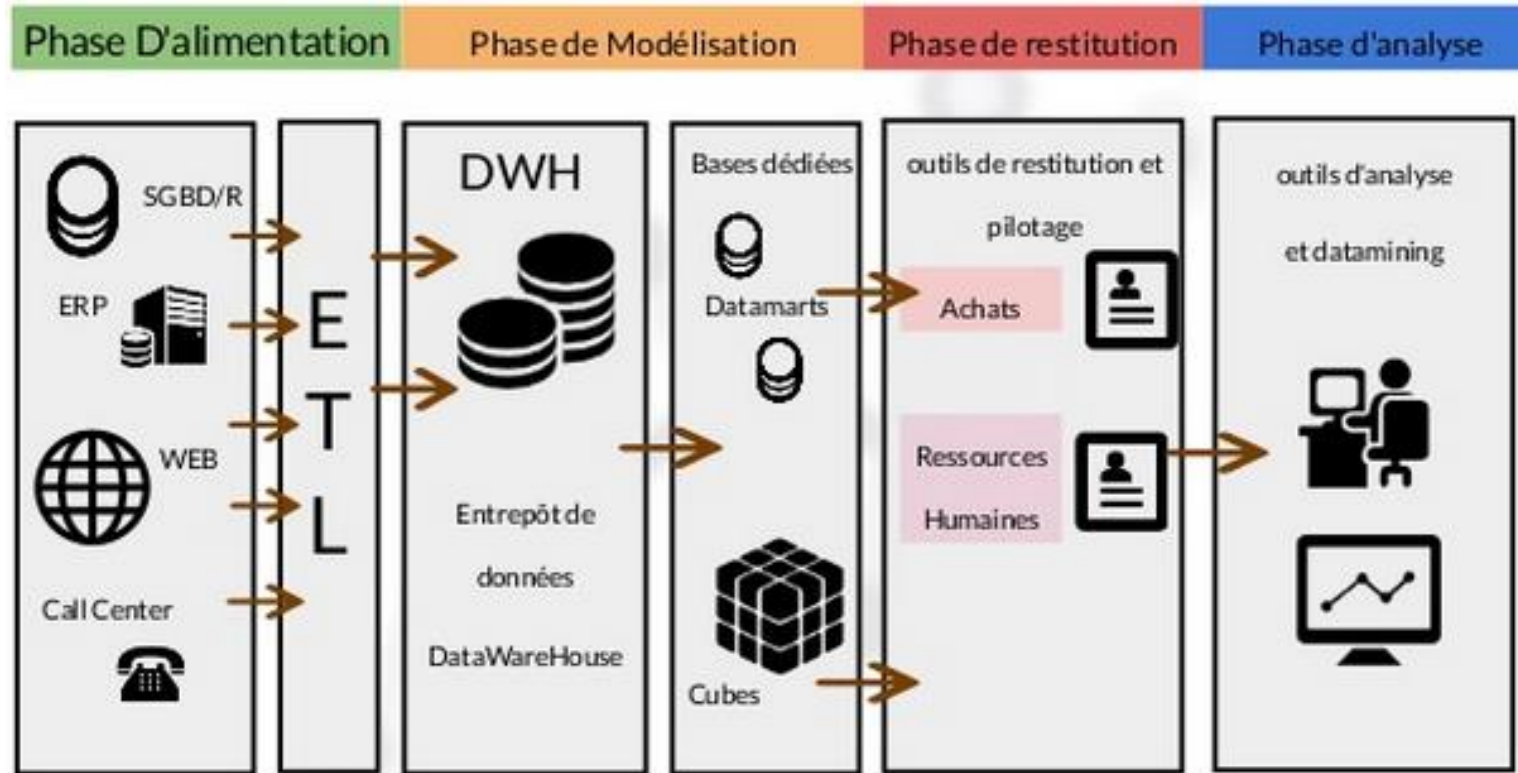
L'exploration de données s'appuie sur des outils permettant de manipulation (IHM) et de visualiser (infovis) les données selon des requêtes dynamiquement produites par des utilisateurs experts du domaine

3- **Analyse des données:**

L'analyse de données est une branche de la statistique qui permet de mettre en évidence des tendances des données ou corrélations entre les données non évidentes a priori.

- Dans le cas de l'analyse descriptive, Il s'agit de rechercher une information statistique "cachée" que l'on ne connaît pas a priori.
- L'approche prédictive consiste à réaliser un modèle statistique des corrélations entre les données à partir d'échantillons d'apprentissage, puis à appliquer le modèle à des données nouvelles pour prédire leur comportement.

| 8. L'infrastructure décisionnelle



Data Warehouse

1. Data Warehouse
2. Data Warehouse : Caractéristiques
3. SGBD et Datawarehouse
4. Architecture Datawarehouse
5. Intégration des données
6. Modèle dimensionnel
7. ETL vs. ELT
8. Granularité
9. Data Mart

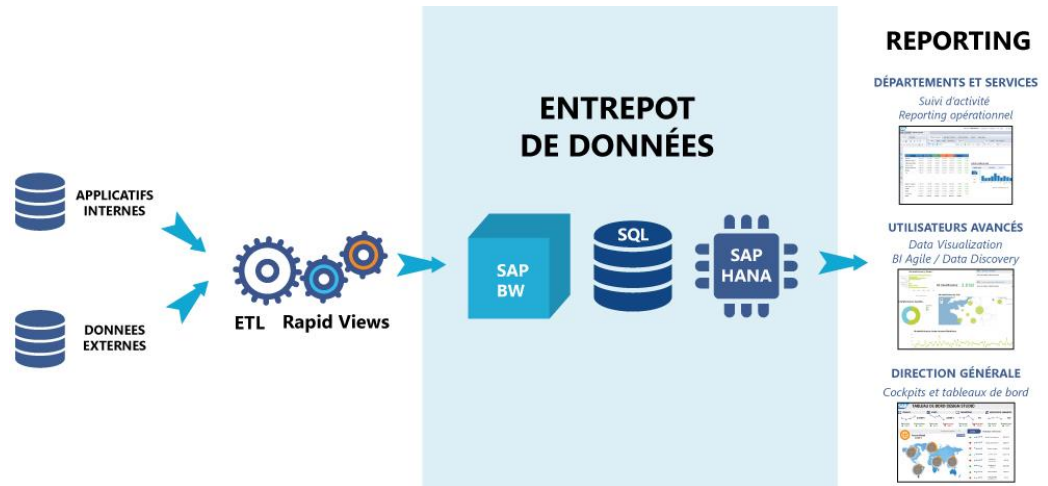
| 1. Data Warehouse

Data Warehouse est une base de données centralisée d'une entreprise.

Spécialement conçu pour recueillir, stocker et intégrer des données provenant de multiples sources de données. Afin de les rendre disponibles pour l'analyse le reporting et la prise de décision.

Définition de Bill Inmon (1996):

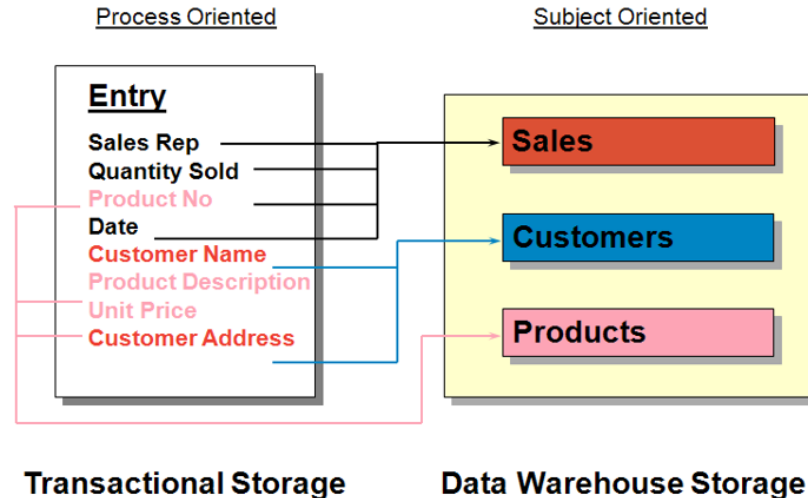
« Le Data Warehouse est une collection de données **orientées sujet, intégrées, non volatiles** et **historisées**, organisées pour le support d'un processus d'aide à la décision. »



| 2. Data Warehouse : Caractéristiques

1- **Orienté sujet:** Un entrepôt de données utilise un thème et fournit des informations sur un sujet particulier et plus défini au lieu des opérations actuelles d'une entreprise. En d'autres termes, le processus d'entreposage de données est mieux équipé pour traiter un thème spécifique.

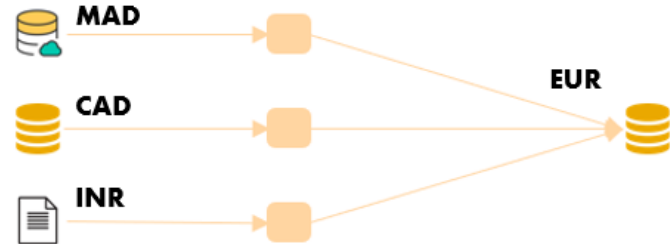
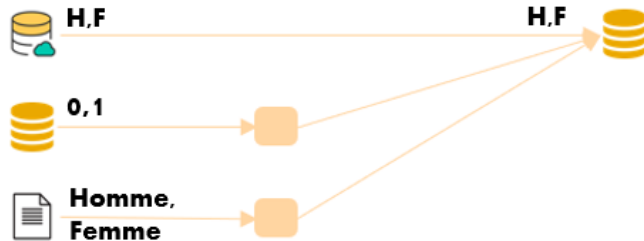
Des exemples de thèmes ou de sujets incluent les ventes, les distributions, le marketing, etc.



| 2. *Data Warehouse : Caractéristiques*

2- Intégré: L'intégration est définie comme l'établissement d'une connexion entre des données provenant de plusieurs bases de données ou sources. Cependant, il est également essentiel que les données soient stockées dans l'entrepôt de données de manière unifiée.

Le processus d'entreposage de données intègre des données provenant de sources multiples, telles qu'un mainframe, des bases de données relationnelles, des fichiers plats, etc. En outre, il permet de maintenir des codes cohérents, des mesures d'attributs, des conventions de dénomination et des formats.



| 2. Data Warehouse : Caractéristiques

3- **Variante temporelle:** La variation temporelle dans un entrepôt de données est plus étendue par rapport aux autres systèmes. Les données stockées dans un entrepôt de données sont rappelées avec une période de temps spécifique et fournissent des informations d'un point de vue historique.

=> Interpréter les données du passé, relier les informations au présent, avoir un aperçu sur les comportements futurs

Base de Production

Données Mai 2021

Nom	Ville
Sébastien	Québec
Francis	Ottawa

Données Septembre 2021

Nom	Ville
Sébastien	Québec
Francis	Québec

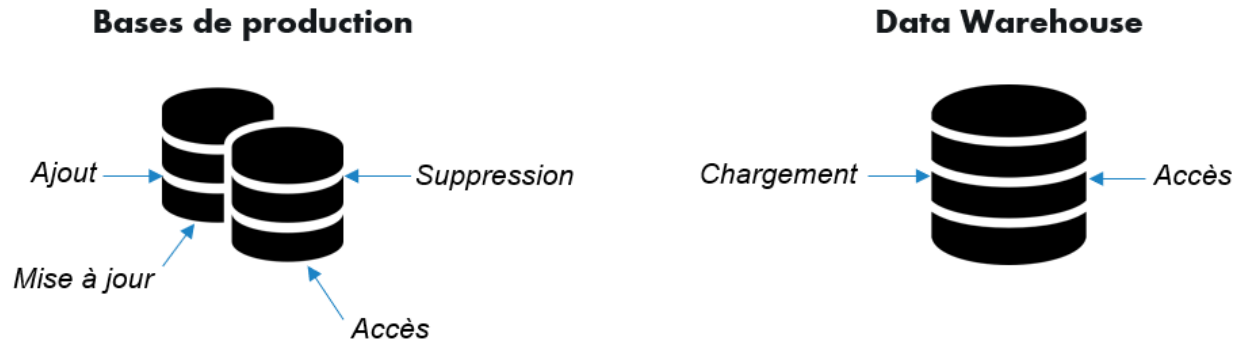
Data Warehouse

Code Cal	Année	Mois
1	2021	Mai
2	2021	Septembre

Code Date	Nom	Ville
1	Sébastien	Québec
1	Francis	Ottawa
2	Francis	Québec

| 2. *Data Warehouse : Caractéristiques*

4- **Non volatile:** Dans l'entrepôt de données non volatile, les données sont permanentes, c'est-à-dire que lorsque de nouvelles données sont insérées, **les données précédentes ne sont pas remplacées**, omises ou supprimées. Dans cet entrepôt de données, les données sont en lecture seule et ne sont actualisées qu'à certains intervalles. Les deux opérations de données effectuées dans l'entrepôt de données sont **l'accès aux données** et le **chargement des données, pas de mise à jour ni de suppression**.



| 3. *SGBD et Datawarehouse*

En réalité SGBD et datawarehouse ont des objectifs différents. Ils stockent les données de manière différentes et font l'objet de requêtes différentes. Ils sont ainsi basés sur deux systèmes différents : OLTP et OLAP

OLTP: OLTP (*On Line Transaction Processing*) est le modèle utilisé par les SGBD. Le mode de travail est transactionnel. L'objectif est de pouvoir insérer, modifier et interroger rapidement et en sécurité la base. Ces actions doivent pouvoir être effectuées très rapidement par de nombreux utilisateurs simultanément.

OLAP: Les datawarehouses eux reposent sur le système OLAP (*On Line Analytical Processing*). Ce système travail en lecture seulement. Les objectifs principaux sont regrouper, organiser des informations provenant de sources diverses, les intégrer et les stocker pour donner à l'utilisateur une vue orientée métier, retrouver et analyser l'information facilement et rapidement.

| 3. *SGBD et Datawarehouse*

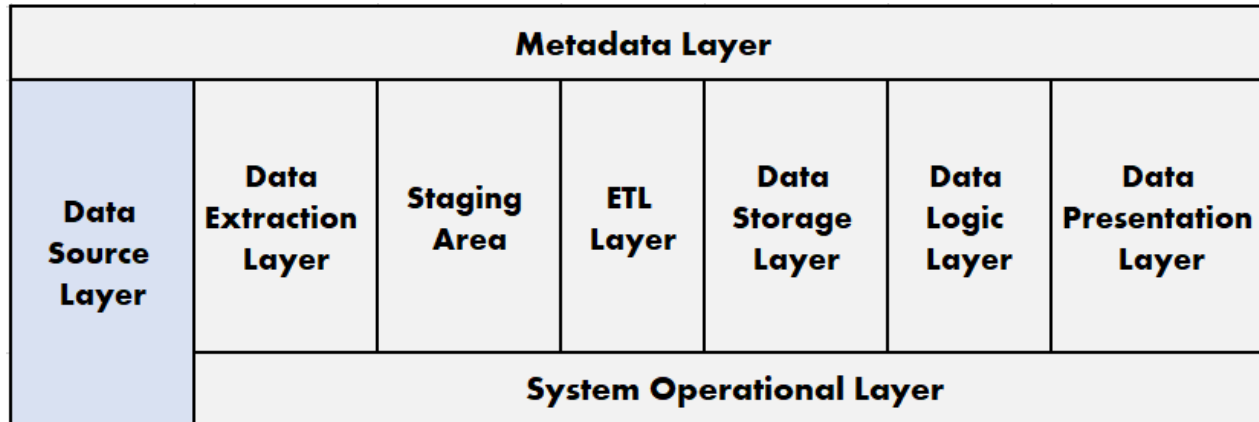
Caractéristiques	OLTP	OLAP
Utilisation	SGBD (base de production)	Datawarehouse
Opération typique	Mise à jour	Analyse
Type d'accès	Lecture écriture	Lecture
Niveau d'analyse	Elémentaire	Global
Quantité d'information échangées	Faible	Importante
Orientation	Ligne	Multidimension
Taille BD	Faible (max qq GB)	Importante (pouvant aller à plusieurs TB).
Ancienneté des données	Récentes	Historiques

Exemples:

- **OLTP** : les ventes du produit X.

- **OLAP** : les ventes du produit X par vendeur, région et par mois.

| 4. *Architecture Datawarehouse*

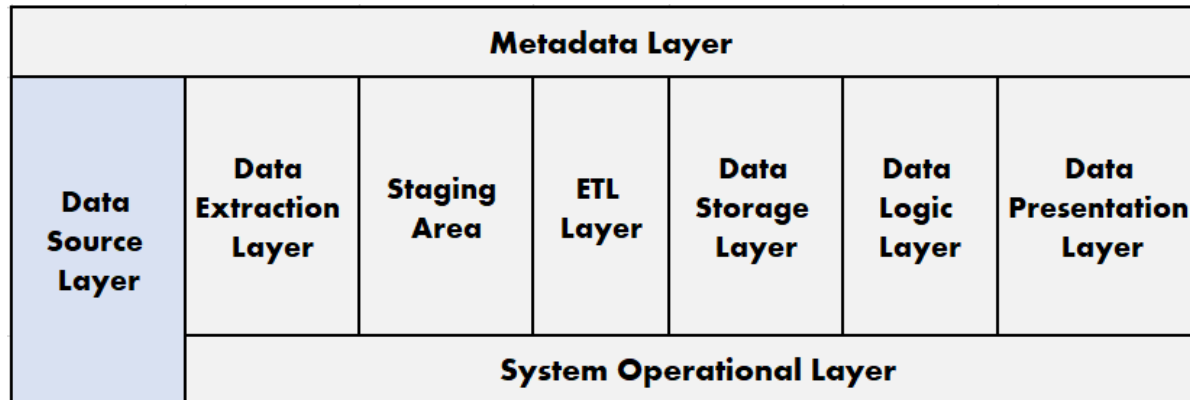


| 4. *Architecture Datawarehouse*

Data Source Layer

Cette couche contient les différentes sources de données qui alimentent les données dans l'entrepôt de données.

La source de données peut être n'importe quel format de fichier texte brut. Base de données relationnelle ou autre type de bases de données, fichier Excel, base de données SQL, MySQL, PostgreSQL.

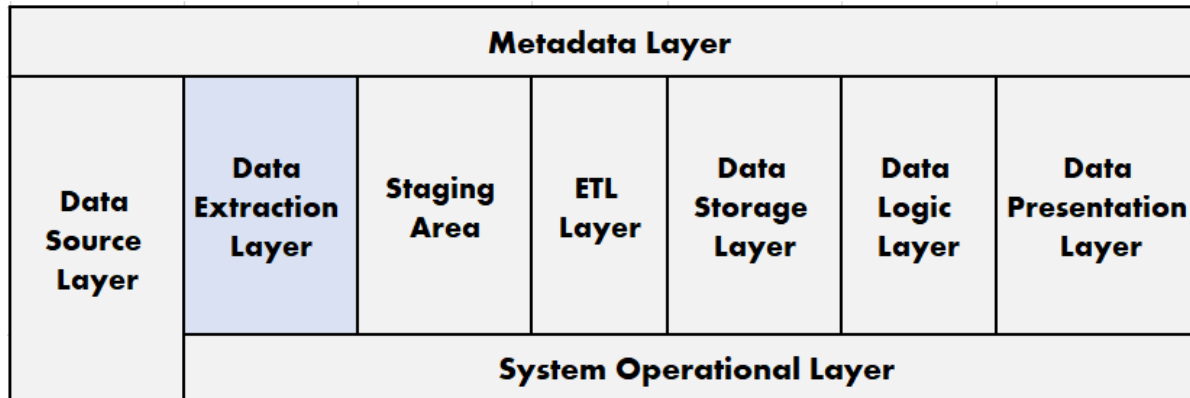


| 4. *Architecture Datawarehouse*

Data Extraction Layer

Les données sont extraites de la source de données vers l'entrepôt de données.

Un nettoyage minimal des données peut être fait à ce niveau.

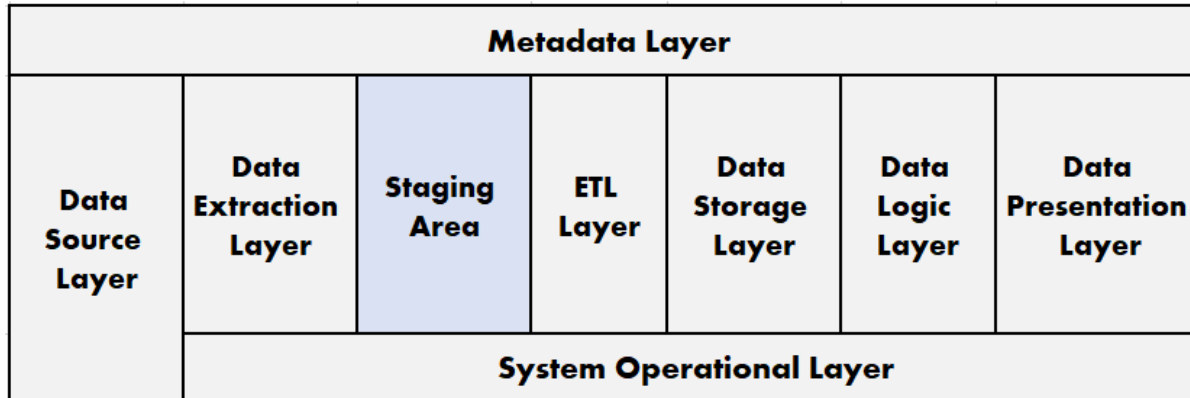


| 4. *Architecture Datawarehouse*

Staging Area

C'est là que les données seront chargées avant d'être nettoyées et transformées en Data Warehouse/Data Mart.

Avoir toutes les données chargées dans une zone de données commune facilite l'intégration des données.

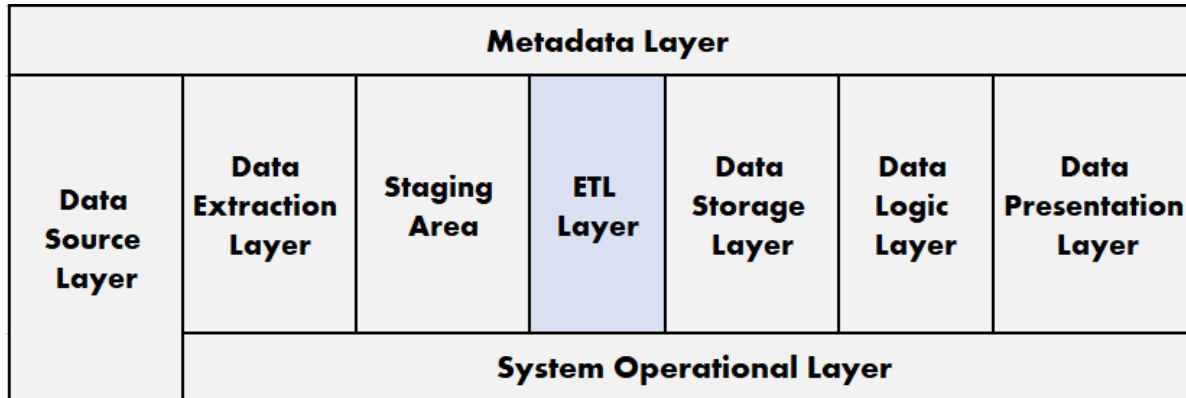


4. Architecture Datawarehouse

ETL Layer



Les termes « Extract, Transform, Load (ETL) » désignent une séquence d'opérations portant sur les données : collecte à partir d'un nombre illimité de sources, structuration et centralisation dans un référentiel unique.



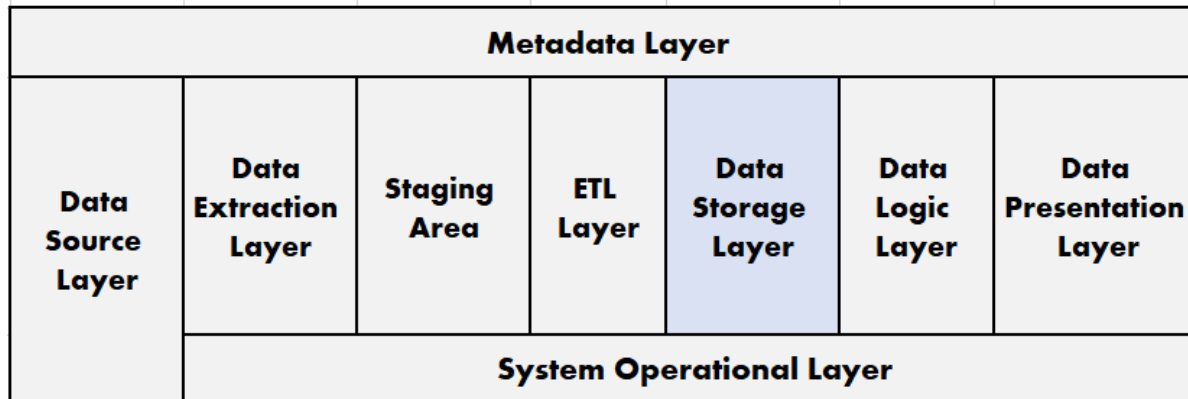
| 4. *Architecture Datawarehouse*

Data Storage Layer

C'est là que sont chargées les données transformées et nettoyées.

En fonction de l'étendue des fonctionnalités, 3 types ou entités peuvent être trouvés ici. (1) Entrepôt de données (2) Data Mart (3) Magasin de données opérationnelles (ODS).

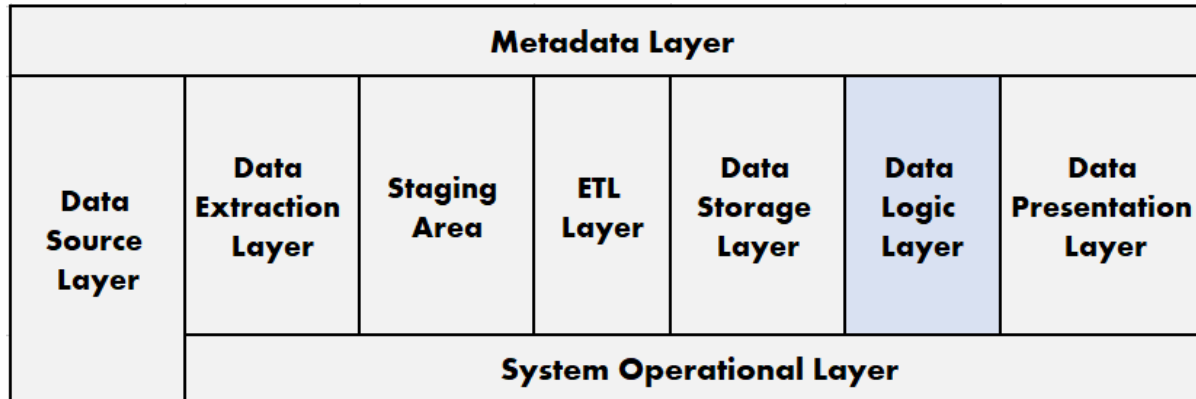
Note: Les 3 types d'unités peuvent coexister dans un même système.



| 4. *Architecture Datawarehouse*

Data Logic Layer

C'est là que les règles métier sont stockées. Les règles métier stockées ici n'affectent pas les règles de transformation des données sous-jacentes, mais affectent l'apparence du rapport.



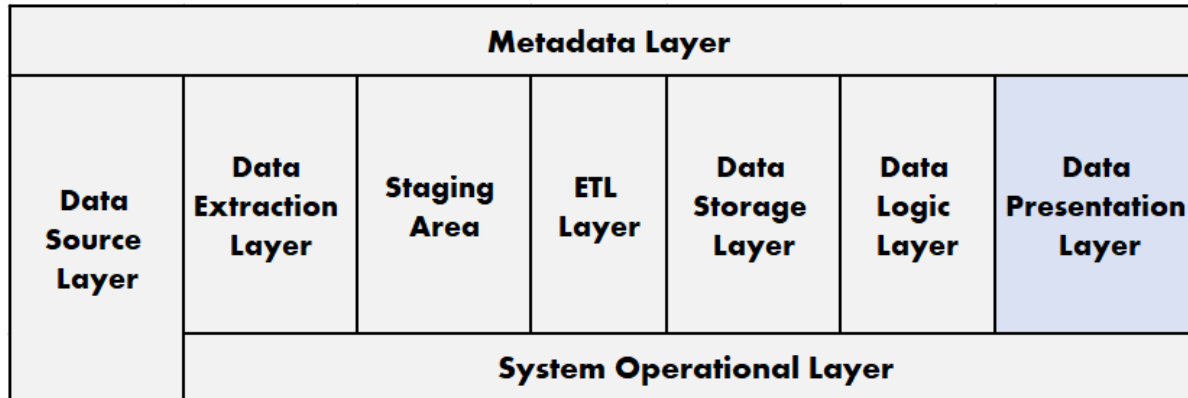
| 4. *Architecture Datawarehouse*

Data Presentation Layer

Il s'agit des informations qui parviennent aux utilisateurs.

Cela peut prendre la forme d'un rapport tabulaire/graphique dans un navigateur, d'un rapport envoyé par e-mail qui est automatiquement généré et envoyé tous les jours, ou d'une alerte qui avertit les utilisateurs des exceptions, entre autres.

Généralement, un outil OLAP et/ou un outil de reporting est utilisé dans cette couche.

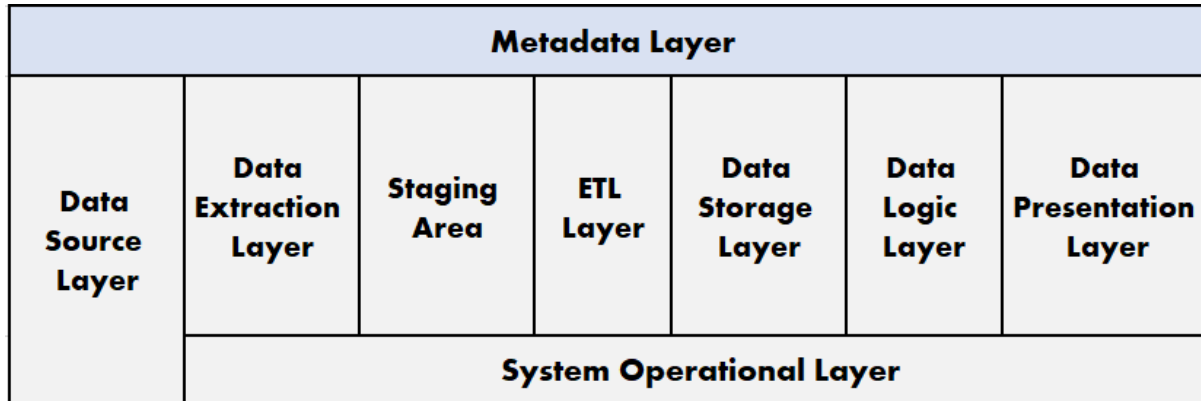


| 4. *Architecture Datawarehouse*

Metadata Layer

C'est là que sont stockées les informations sur les données stockées dans le système d'entrepôt de données.

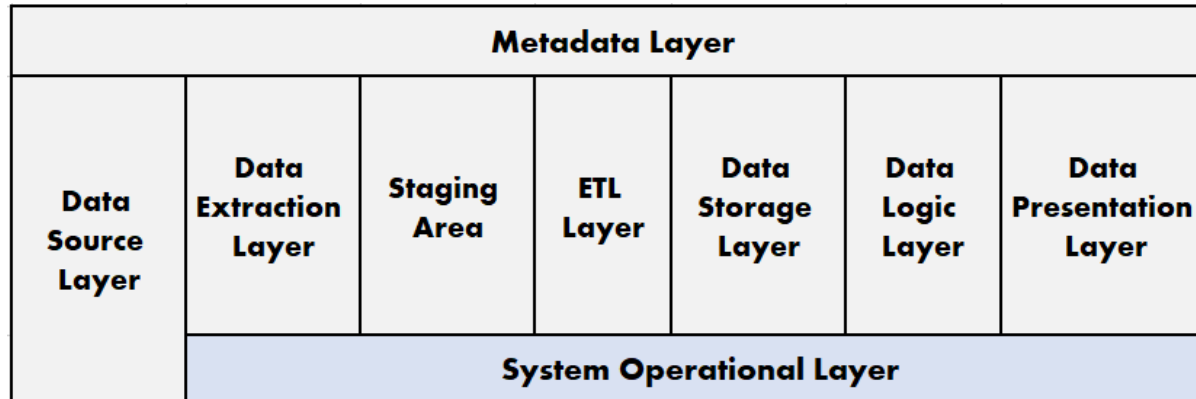
Un modèle de données logique serait un exemple de quelque chose qui se trouve dans la couche de métadonnées.



| 4. *Architecture Datawarehouse*

System operational Layer

Cette couche comprend des informations sur le fonctionnement du système d'entrepôt de données, telles que l'état du travail ETL, les performances du système et l'historique d'accès des utilisateurs.



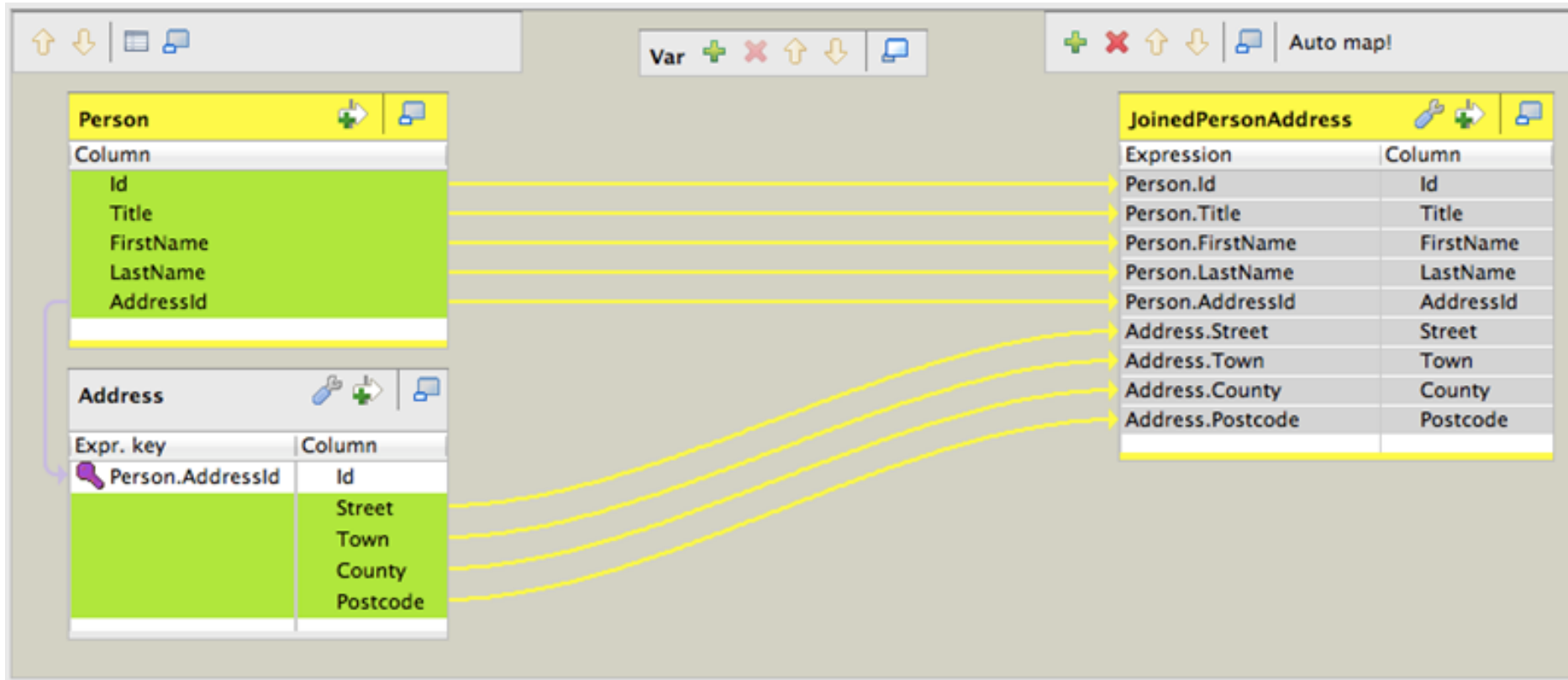
| 5. *Intégration des données (Alimentation du DW)*

L'intégration des données vise à assurer la cohérence entre les données et à ne stocker dans l'entrepôt de données que des informations préalablement mises en relation les unes avec les autres (mapping).

Outils ETL/ ELT:

- Pentaho PDI
- Talend
- Oracle Data Integrator (ODI)
- Business Objects Data Integrator
- Microsoft SSIS - SQL Server Integration Services

| 5. Intégration des données (Alimentation du DW)



| 5. *Intégration des données*

ETL = Extract + Transform + Load

Extraction

- 1^{ère} étape du processus ETL : Récupération des données depuis les sources de données.

Transformation

- filtrer
- trier
- homogénéiser
- nettoyer
- ...

Chargement

- Chargement des données dans le DW

| 5. *Intégration des données*

Extraction

Avant de pouvoir commencer à organiser vos données, la première étape du processus de données ETL consiste à extraire ou à extraire les données de toutes les sources pertinentes et à les compiler.

L'extraction doit Être périodique et répétée.

| 5. *Intégration des données*

Extraction

Extraction complète:

- Capture l'ensemble des données à un certain instant (snapshot de l'état opérationnel);
- Normalement employée dans deux situations:
 1. Chargement initial des données;
 2. Rafraîchissement complet des données (ex: modification d'une source).
- Peut être très coûteuse en temps (ex: plusieurs heures/jours).

Extraction incrémentale:

- Capture uniquement les données qui ont changées ou ont été ajoutées depuis la dernière extraction;
- Peut être faite de deux façons:
 1. Extraction temps-réel;
 2. Extraction différée (en lot).

| 5. *Intégration des données*

Extraction

1. Extraction temps-réel:

S'effectue au moment où les transactions surviennent dans les systèmes sources.

2. Extraction différée:

Extrait tous les changements survenus durant une période donnée (ex: heure, jour, semaine, mois).

| 5. *Intégration des données*

Transformation

L'opération la plus importante de l'étape de transformation consiste à appliquer aux données brutes les règles internes de l'entreprise de manière à répondre aux exigences en matière de reporting : les données brutes sont nettoyées et converties aux formats de rapport qui conviennent (si les données ne sont pas nettoyées, il devient plus difficile d'appliquer les règles internes de reporting).

Exemples de transformation:

1. Révision de format:

- Ex: Changer le type ou la longueur de champs individuels.

2. Décodage de champs:

- Consolider les données de sources multiples
 - Ex: ['homme', 'femme'] vs ['M', 'F'] vs [1,2].
- Traduire les valeurs cryptiques
 - Ex: 'AC', 'IN', 'SU' pour les statuts actif, inactif et suspendu.

| 5. *Intégration des données*

Transformation

3. Pré-calcul des valeurs dérivées:

- Ex: profit calculé à partir de ventes et coûts.

4. Découpage de champs complexes:

- Ex: extraire les valeurs prénom, secondPrénom et nomFamille à partir d'une seule chaîne de caractères nomComple.

5. Fusion de plusieurs champs:

- Ex: information d'un produit
 - Source 1: code et description;
 - Source 2: types de forfaits;
 - Source 3: coût.

6. Conversion de dates:

- Ex: '24 FEB 2011' vs '24/02/2011' vs '02/24/2011'.

| 5. *Intégration des données*

Chargement

Types de chargement:

- **Chargement initial:**

- Fait une seule fois lors de l'activation de l'entrepôt de données;
- Les indexes et contraintes d'intégrité référentielle (clé étrangères) sont normalement désactivés temporairement;
- Peut prendre plusieurs heures

- **Chargement incrémental:**

- Fait une fois le chargement initial complété;
- Tient compte de la nature des changements (ex: SCD Type 1, 2 ou 3);
- Peut être fait en temps-réel ou en lot.

- **Rafraîchissement complet:**

- Employé lorsque le nombre de changements rend le chargement incrémental trop complexe;
- Ex: lorsque plus de 20% des enregistrements ont changé depuis le dernier chargement.

| 6. *Modèle Dimensionnel*

Dans un Data Warehouse (et au niveau de chaque Data mart), les données et leurs relations sont organisées suivant un modèle de données spécifique. Le choix du modèle de données structure et définit le design du Data Warehouse.

Il existe trois modélisations possibles :

- La modélisation en étoile,
- La modélisation en flocons,
- La modélisation en constellation.

Les dimensions

Les tables de dimensions sont utilisées pour décrire les données que l'on souhaite stocker dans le Data Warehouse.

Exemple: Vous êtes un retailer et vous voulez enregistrer la **date**, le **magasin** et l'**employé** associé à chaque achat. A chaque catégorie de données correspond une table de dimension : la table de **dimension date**, la table de **dimension magasin** et la table de **dimension employé**.

| 6. *Modèle Dimensionnel*

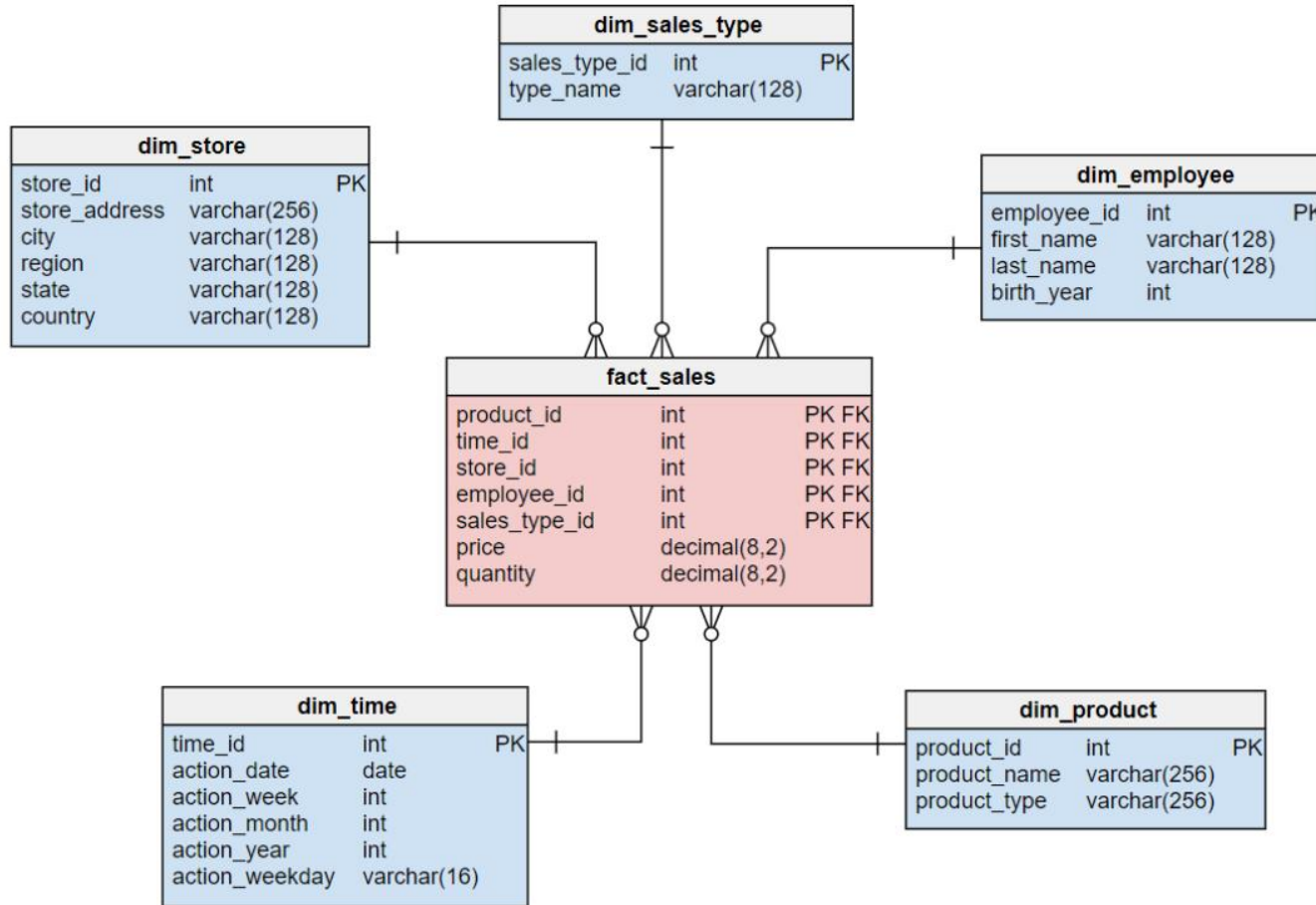
Les faits

Les tables de faits contiennent les données que l'on souhaite voir apparaître dans les rapports d'analyse, sous forme de métriques. Les données des tables de faits sont agrégées à partir des tables de dimensions qui leur sont associées.

Une table de faits se présente sous la forme d'un ensemble de colonnes stockant les valeurs et les clés étrangères (identifiants) associées aux tables de dimensions.

La combinaison des différentes clés étrangères est ce qui forme l'identité de la table de faits, sa clé primaire. La table de faits n'a donc pas besoin d'avoir une clé primaire propre : elle est définie et identifiée par une combinaison unique de clés étrangères.

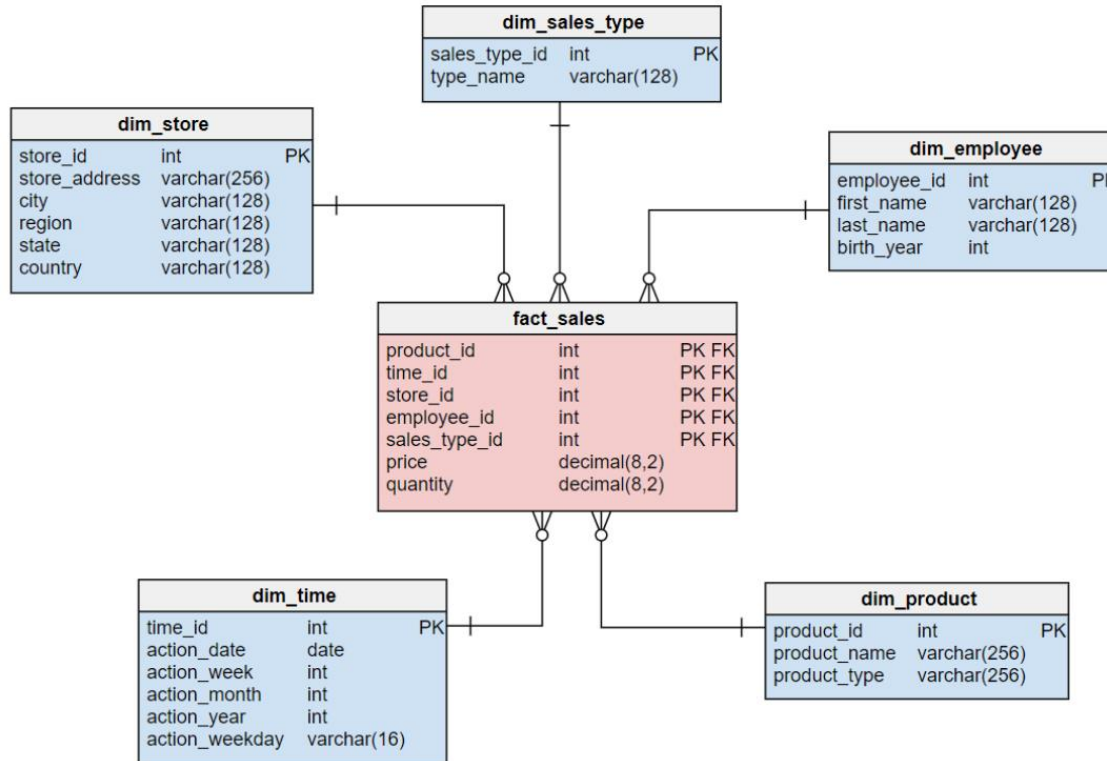
6. Modèle Dimensionnel



6. Modèle Dimensionnel

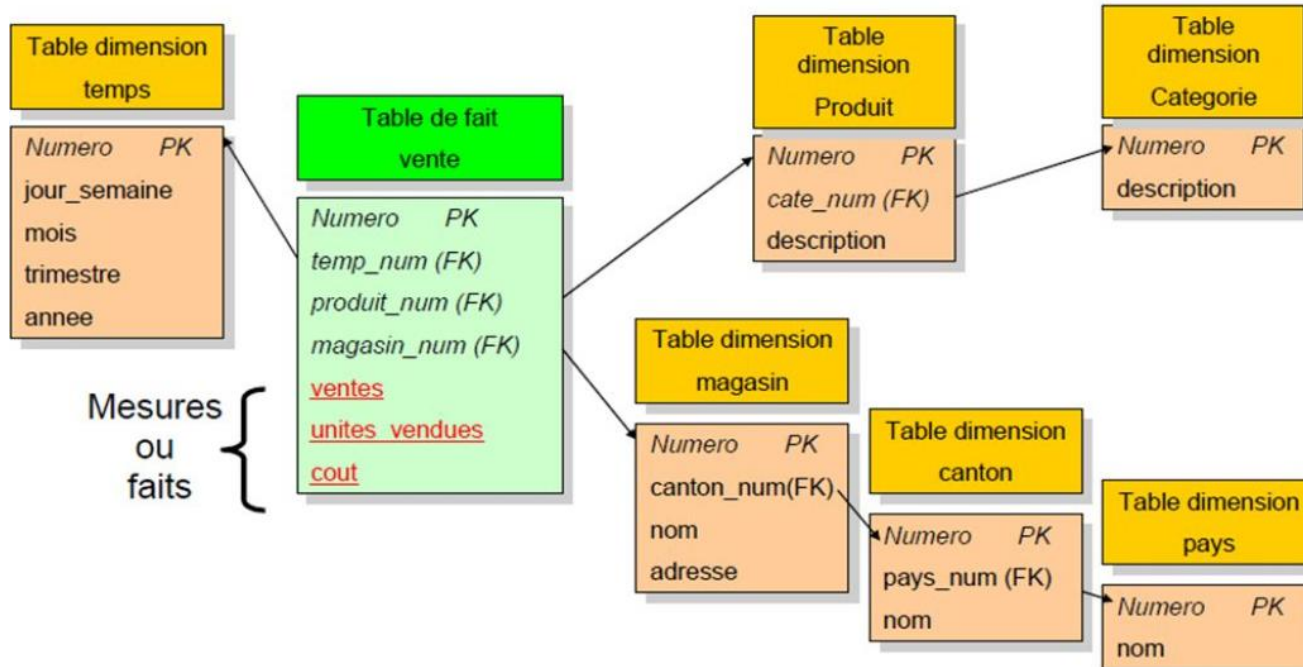
Modèle en étoile

Un modèle en étoile est constitué d'une table centrale, dite table des faits, et de nombreuses tables dimensionnelles autour



6. Modèle Dimensionnel: Modèle en Flocon

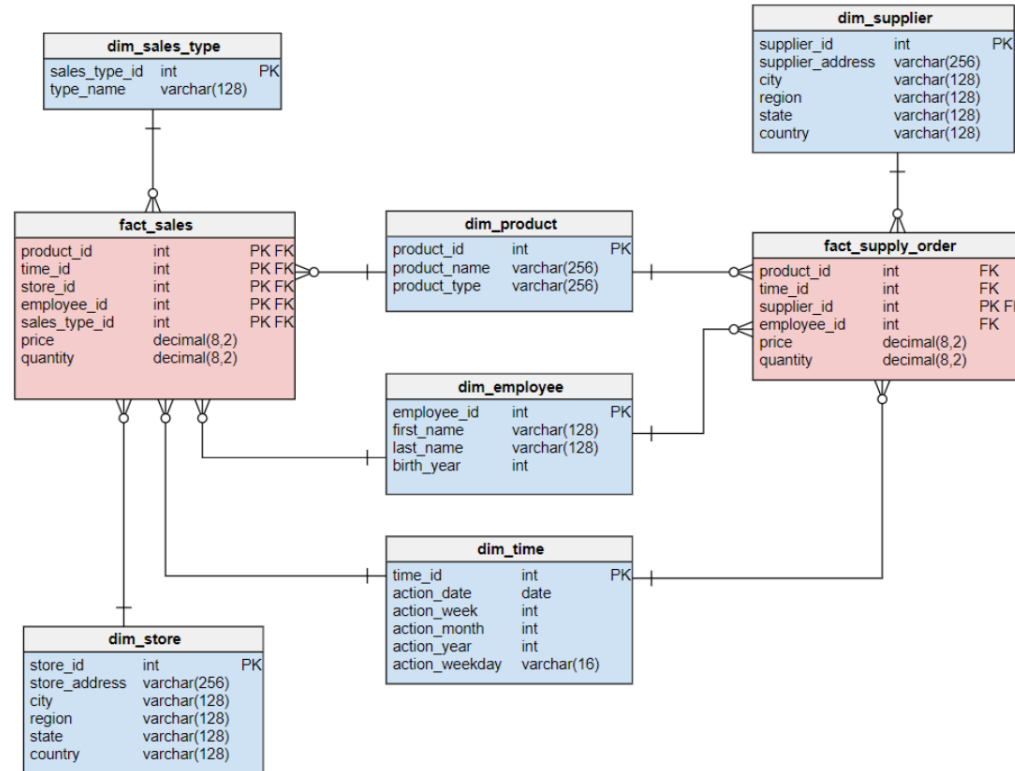
Le modèle de données dit « en flocons » est une variante du modèle en étoile : chaque table de dimension est re-normalisée pour faire apparaître la hiérarchie sous-jacente (nomenclature, etc.). La normalisation n'est pas indispensable car ni mises à jour ni suppressions ne sont effectuées directement sur l'entrepôt de données. L'intérêt principal du modèle en flocons réside dans le gain en espace de stockage qui est de l'ordre de 5 à 10 %.



6. Modèle Dimensionnel

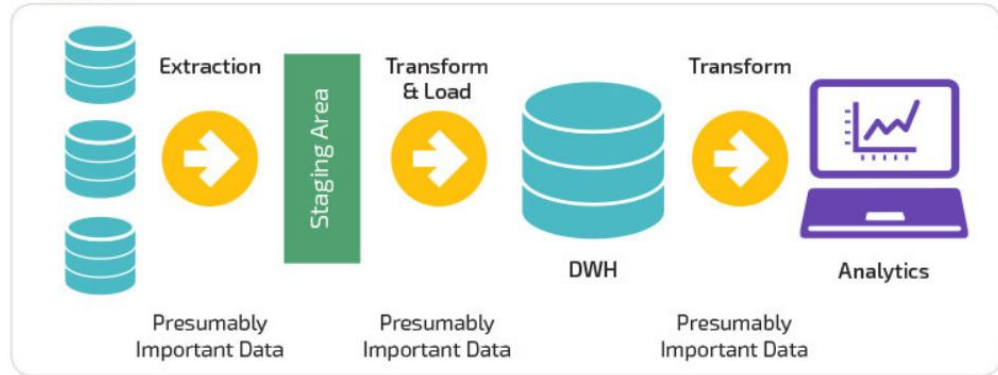
Modèle en constellation

Le modèle en constellation se compose donc de plusieurs tables des faits avec leurs tables de dimensions respectives. Les tables de dimensions communes aux différentes tables des faits ne font pas l'objet de redondances : c'est l'un des principaux avantages cette modélisation.

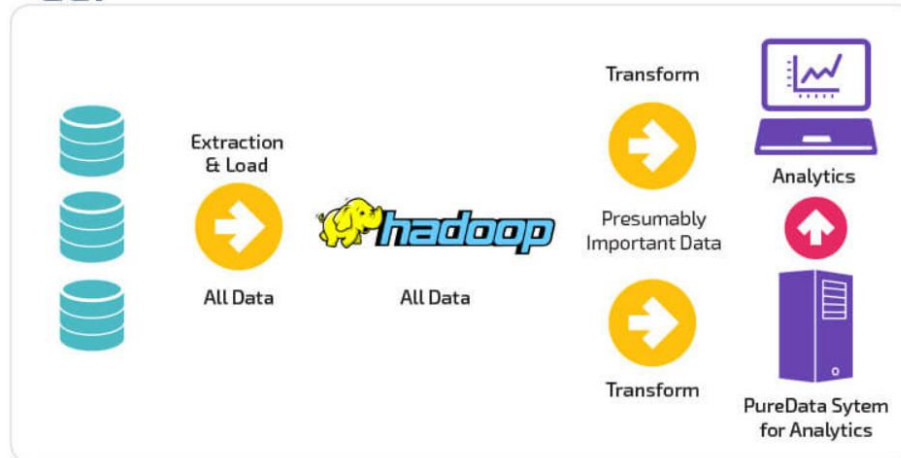


7. ETL vs. ELT

ETL

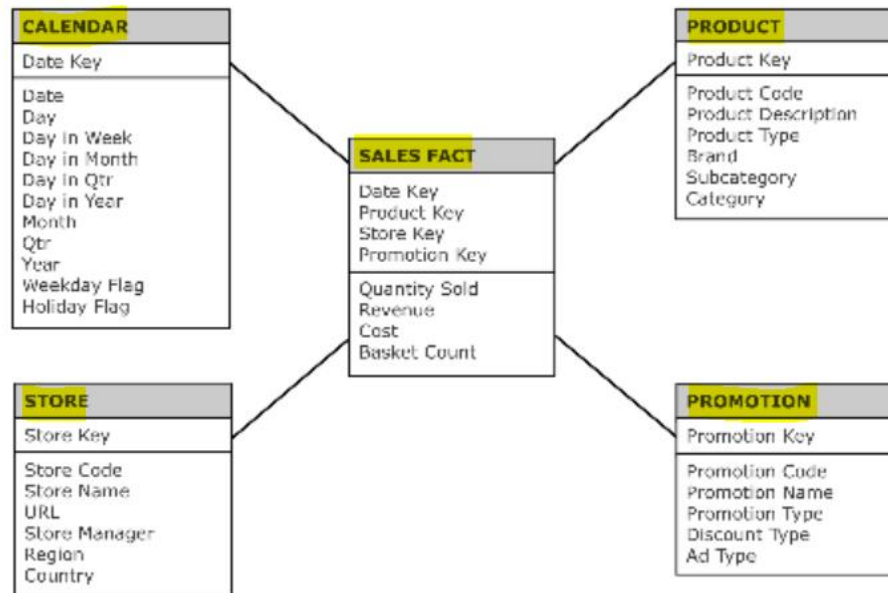
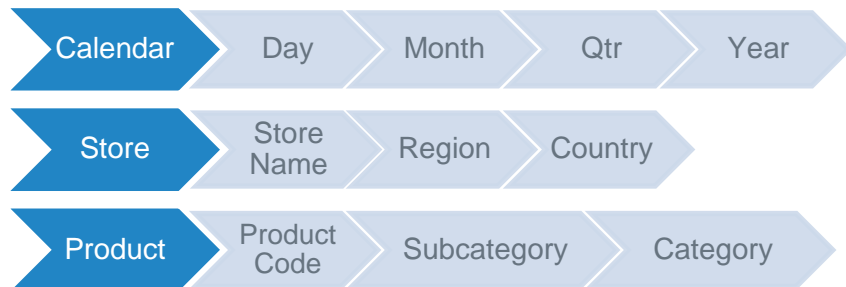


ELT



| 8. Granularité

La granularité est ce qui permet de définir le niveau de détail des informations présentes dans une ligne d'une table de faits. Il est défini par un ensemble minimal de dimensions.

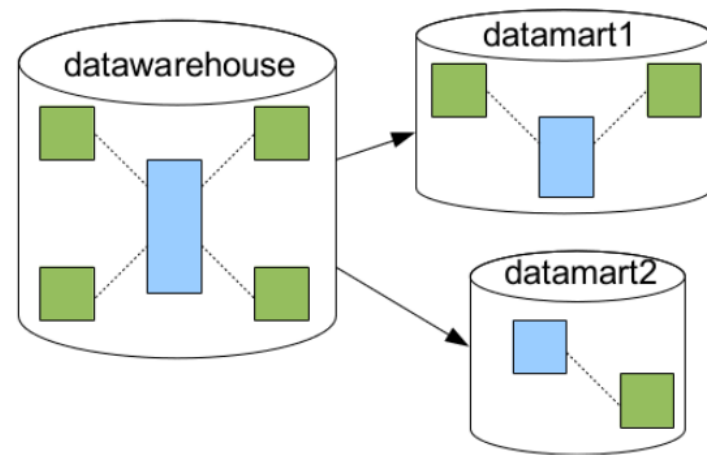


| 9. Data Mart

Les **Data Marts** sont des compartiments du Data Warehouse conçus pour répondre à un besoin métier spécifique. Un Data Mart regroupe et organise toutes les données nécessaires au traitement d'une problématique précise. Par exemple : toutes les données relatives aux ventes, ou bien toutes celles relatives à la logistique, au marketing, à la production, etc.

Pour prendre une image, le Data Warehouse peut être comparé à un immeuble de bureaux et les Data Marts aux différents bureaux qui le constituent.

Les **Data Marts** ont donc une utilité du point de vue des utilisateurs mais aussi en terme de sécurité. Au lieu que tous les utilisateurs aient accès à toutes les données stockées, l'organisation en Data Marts permet de contrôler et de restreindre les accès aux données en fonction des besoins des utilisateurs.



TD₁ : Elaboration de modèles d'entrepôt de données

| **Exercice 1**

Une entreprise de fabrication de vaisselle jetable souhaite mettre en place un système d'information décisionnel sous la forme d'un Data Mart (un mini entrepôt de données) pour observer son activité de ventes au niveaux des différents lieux de distributions de ses articles et cela dans plusieurs villes.

Ces lieux de distributions sont renseignés par leur enseigne, leur type (en fonction de leur surface), leur adresse (code postal et ville), leur département, leur région. Les ventes sont renseignées selon une période qui se décline en mois, en trimestre et année. Les ventes sont observées par le nombre d'articles selon le type, et le chiffre d'affaire.

- Quel est le fait à observer ?
- Quels sont les axes d'analyse, et les mesures ?
- Construire le modèle en étoile de ce Data Mart.

| **Exercice 1**

Une entreprise de fabrication de vaisselle jetable souhaite mettre en place un système d'information décisionnel sous la forme d'un Data Mart (un mini entrepôt de données) pour observer son activité de **ventes** au niveaux des différents lieux de distributions de ses articles et cela dans plusieurs villes.

Ces lieux de distributions sont renseignés par leur enseigne, leur type (en fonction de leur surface), leur adresse (code postal et ville), leur département, leur région. Les ventes sont renseignées selon une période qui se décline en mois, en trimestre et année. Les ventes sont observées par le nombre d'articles selon le type, et le chiffre d'affaire.

- Quel est le fait à observer ?
- Quels sont les axes d'analyse, et les mesures ?
- Construire le modèle en étoile de ce Data Mart.

| **Exercice 1**

Une entreprise de fabrication de vaisselle jetable souhaite mettre en place un système d'information décisionnel sous la forme d'un Data Mart (un mini entrepôt de données) pour observer son activité de **ventes** au niveaux des différents lieux de distributions de ses articles et cela dans plusieurs villes.

Ces lieux de distributions sont renseignés par leur enseigne, leur type (en fonction de leur surface), leur adresse (code postal et ville), leur département, leur région. Les ventes sont renseignées selon une période qui se décline en mois, en trimestre et année. Les ventes sont observées par le **nombre d'articles** selon le type, et le **chiffre d'affaire**.

- Quel est le fait à observer ?
- Quels sont les axes d'analyse, et les mesures ?
- Construire le modèle en étoile de ce Data Mart.

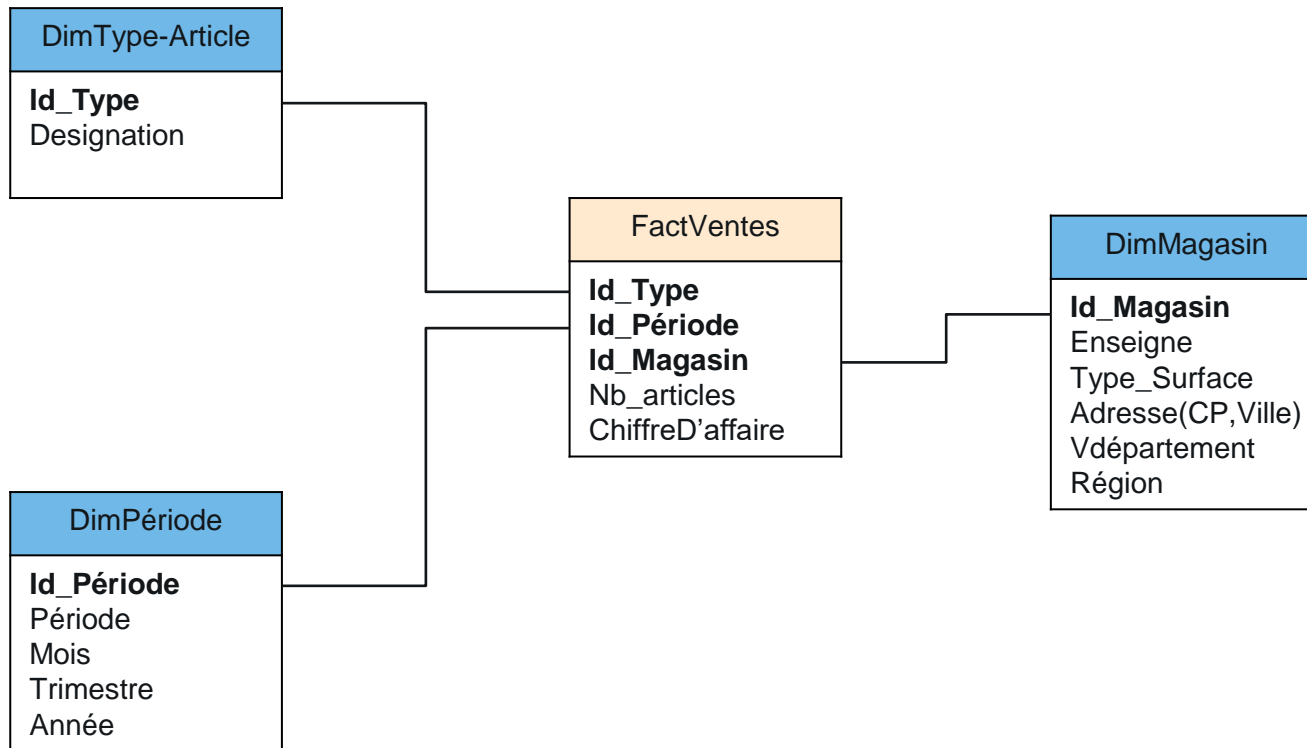
| **Exercice 1**

Une entreprise de fabrication de vaisselle jetable souhaite mettre en place un système d'information décisionnel sous la forme d'un Data Mart (un mini entrepôt de données) pour observer son activité de **ventes** au niveaux des différents **lieux de distributions** de ses articles et cela dans plusieurs villes.

Ces lieux de distributions sont renseignés par leur enseigne, leur type (en fonction de leur surface), leur adresse (code postal et ville), leur département, leur région. Les ventes sont renseignées selon une **période** qui se décline en mois, en trimestre et année. Les ventes sont observées par le **nombre d'articles** selon le **type**, et le **chiffre d'affaire**.

- Quel est le fait à observer ?
- Quels sont les axes d'analyse, et les mesures ?
- Construire le modèle en étoile de ce Data Mart.

| *Exercice 1*



| **Exercice 2**

1- Concevoir un modèle en étoile qui permet d'analyser les ventes d'une entreprise de restauration rapide. Le principe est de mesurer les ventes grâce aux quantités vendues et aux bénéfices, en fonction des ventes réalisées par jour, dans un restaurant donné, pour un aliment donné.

L'objectif est de pouvoir analyser les ventes par jour, par semaine, par mois et par année. Les restaurants peuvent être regroupés en fonction de leur ville et de leur pays.

2- On souhaite à présent mesurer le nombre de commandes qui est donné par jour et par restaurant.

Etendre le modèle précédent afin de prendre en compte cet aspect.



Questions ?