Business Intelligence (BI)

Ziane.S

Année universitaire 2024-2025

Business Intelligence

Semestre: 6 Parcours ISIL

Unité d'enseignement Méthodologique Matière : Business intelligence (BI)

Crédits : 2 Coefficient : 1

- Objectifs de l'enseignement : l'objectif de cette matière est d'apporter des connaissances à l'étudiant pour : Comprendre le business intelligence
- Déterminer les différentes étapes dans la mise en place d'un projet B.I
- Examiner le marché de la BI et les outils leaders

Connaissances préalables recommandées : maitrise de base de données relationnelles

Contenu de la matière :

- I. BI vs ERP
- II. Architecture classique d'une solution BI
- III. ETL et Data
- IV. Modèles multi-dimensionnels et OLAP
- V. Data Warehouse Modeling
- VI. Data Warehouse architecture
- VII. BI et Data Mining
- VIII. KPI, Dashboard, Scorecards et Cockpit
- IX. Real-time BI

Données (Data): Ce sont des éléments bruts, non organisés, qui n'ont pas encore de signification.

• Exemple: Une liste de chiffres « 23, 45, 78, 90 » ou un fichier contenant des noms et des dates sans contexte.

Information: Les données sont transformées en information lorsqu'elles sont organisées et mises dans un contexte. Elles deviennent compréhensibles et utiles.

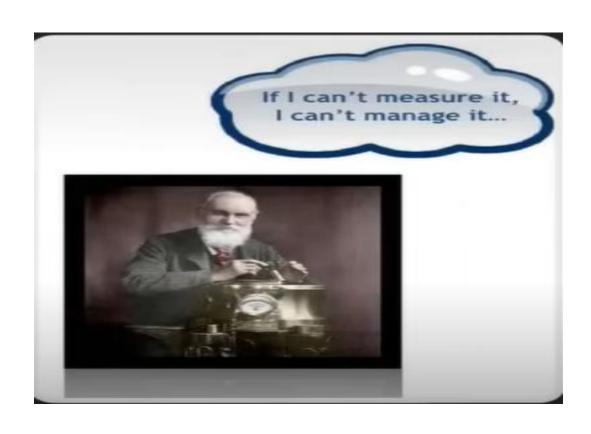
- Exemple: "23°C, 45°C, 78°C, 90°C" -> Ce sont des températures.
- Explication : On a ajouté une signification aux données en précisant qu'il s'agit de températures.

Connaissance (Knowledge): L'information devient connaissance lorsqu'on peut l'utiliser pour comprendre une situation ou prendre des décisions.

- Exemple: "Si la température dépasse 50°C, la machine risque de surchauffer."
- Explication : On interprète l'information et on déduit une règle ou une relation utile.

Sagesse (Wisdom) : C'est la capacité à utiliser la connaissance de manière judicieuse pour prendre les meilleures décisions.

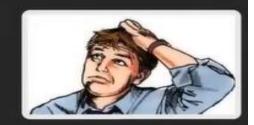
- Exemple : "Pour éviter la surchauffe, nous allons programmer un système d'alerte lorsque la température dépasse 45°C."
- Explication : La sagesse repose sur l'expérience et le jugement pour appliquer la bonne solution.

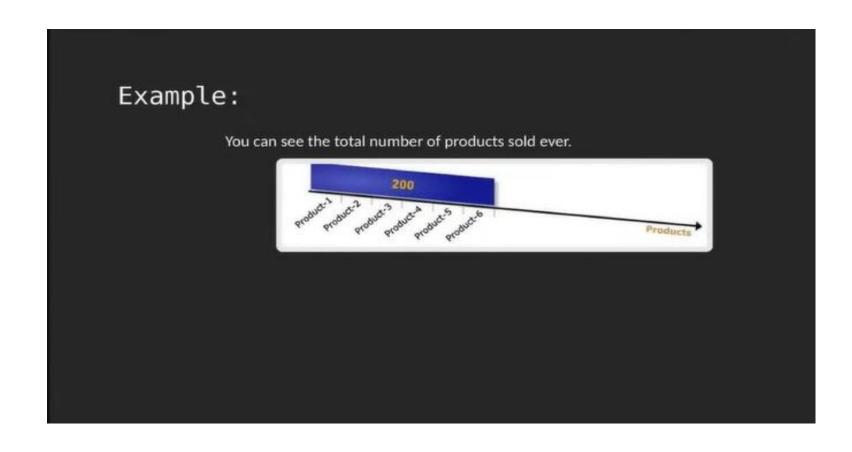


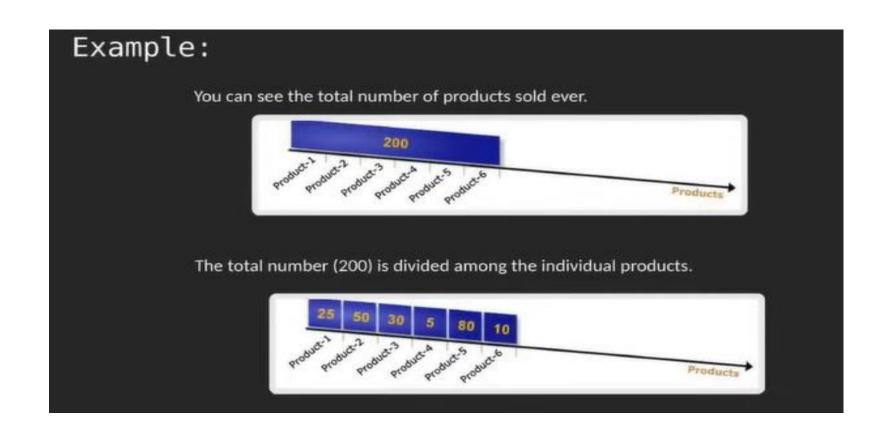


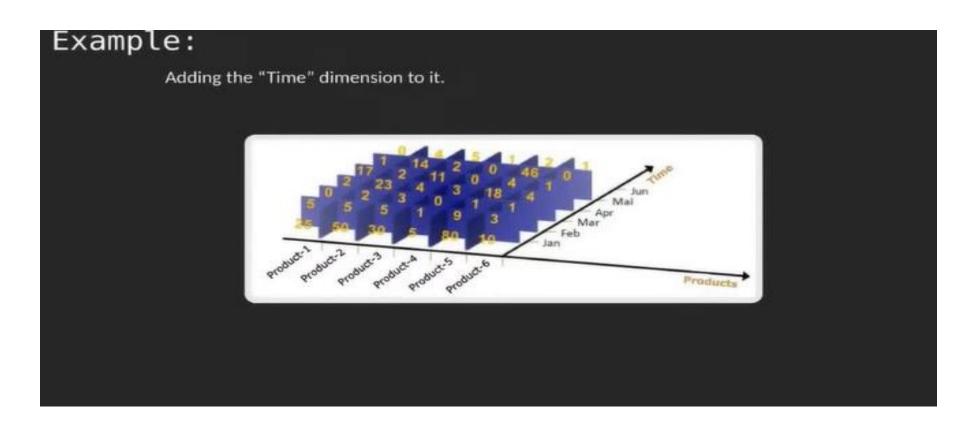
Common Pain..

- Business data is distributed across the enterprise.
- Data everywhere, information nowhere.
- Takes too long wasted resources/efforts.
- So many excuses Why I just can't get it to you when you want it.
- Different users have different needs:
 - · Excel versus PDF
 - Pull versus push
 - On demand on schedule
 - · Your format My format

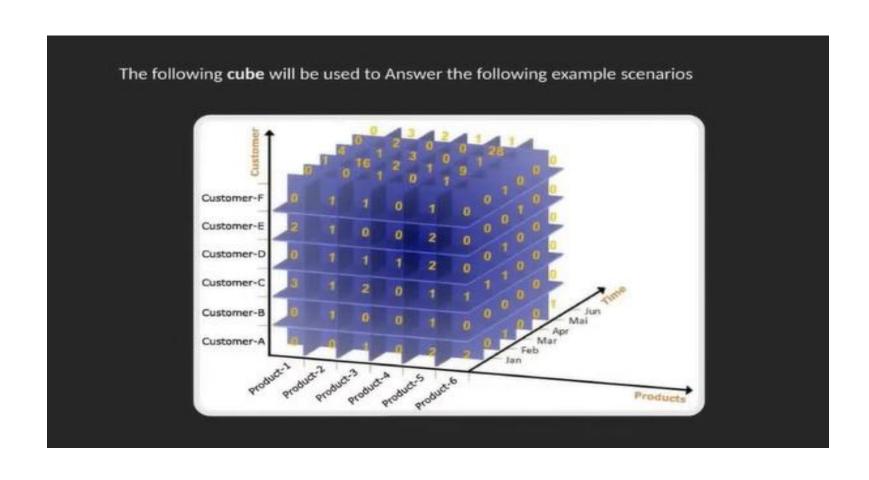


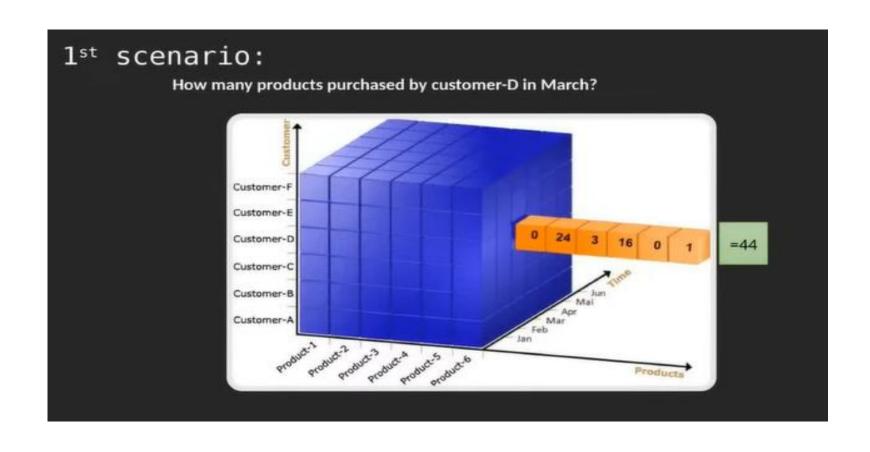


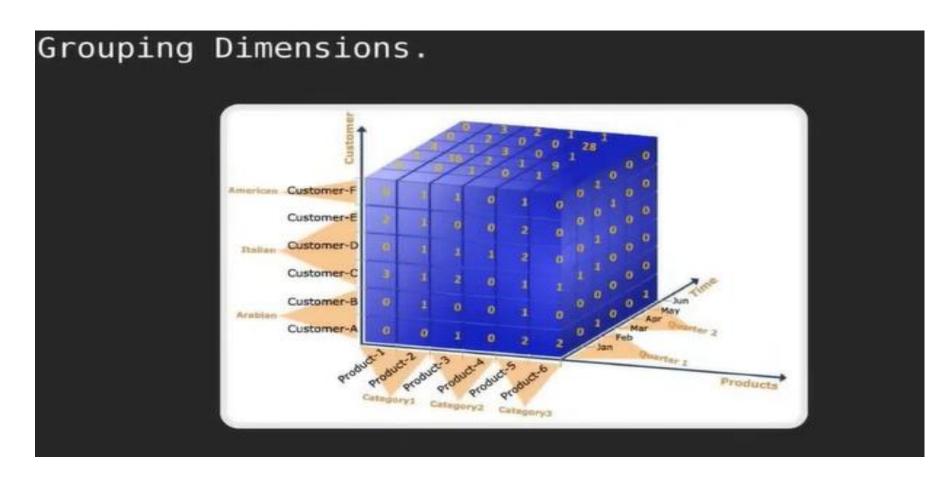




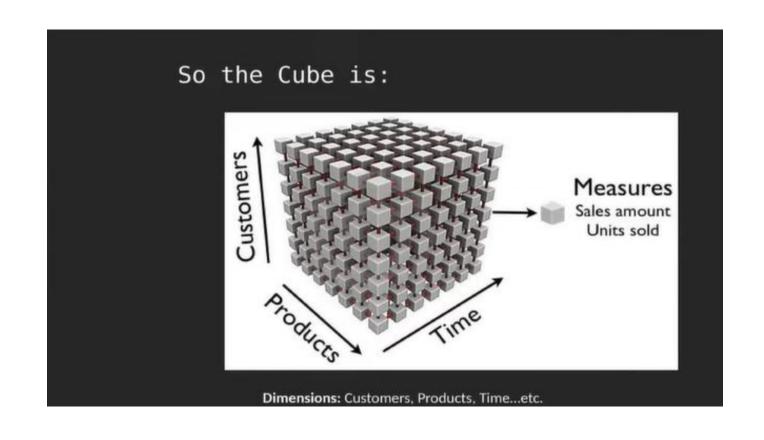
when i drill down







the more information you get as a manager the more you are capable to take actions





What is Business Intelligence?

- Looking for information?
- BI transforms raw <u>Data</u> into valuable <u>Information</u>.

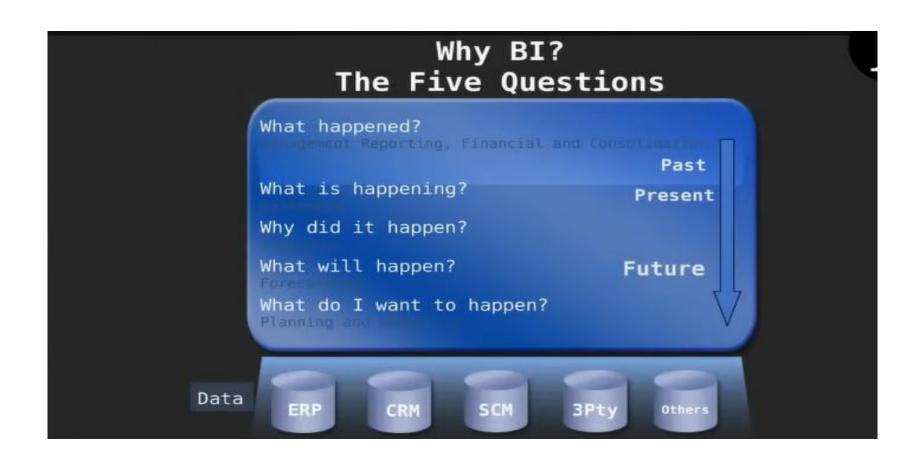
• <u>Information</u> can be transformed into <u>Knowledge</u> through discovery.

Business Intelligence (BI) is about getting the <u>right information</u>, to the right decision makers, at the right time.

What is Business Intelligence?

- BI is an enterprise-wide platform that supports reporting, analysis and decision making.
- •BI Leads to:
 - Fact-based decision making not assumption-based decision making.
 - "Single version of the truth"





Benefits..

- Give user the means to make better decisions.
- Leverage your investment in your legacy systems (e.g. ERP) or data warehouse, it is the best way to get value form existing applications.
- Challenge assumption with factual information.
- Eliminate report backlog and delays.
- Find root causes and take action.
- Identify waste resources and reduce cost.
- Improve operational efficiency.
- Creates organizational language.
- Negotiate better contracts with suppliers and customers.
- Improve strategies with better marketing analysis.

I. BI vs ERP

1. Définition de l'ERP (Enterprise Resource Planning - Progiciel de Gestion Intégrée - PGI)

Un **ERP** est un logiciel de gestion qui permet à une entreprise de gérer ses activités quotidiennes de manière intégrée. Il centralise et automatise différents processus métier, tels que :

- Gestion comptable et financière (facturation, paiements)
- Gestion des stocks (suivi des produits disponibles)
- Gestion des ressources humaines (paie, recrutement)
- Gestion des fournisseurs (commandes, contrats)
- Gestion des ventes et de la distribution (suivi des commandes, logistique)
- Gestion de l'e-commerce (ventes en ligne, marketing)

Exemple:

Une entreprise de fabrication utilise un ERP pour gérer ses matières premières, suivre la production, gérer les salaires des employés et envoyer les factures aux clients, le tout dans un même système.

2. Définition de la Business Intelligence (BI)

La BI (Business Intelligence), ou informatique décisionnelle, est un ensemble de technologies et de processus permettant d'analyser des données pour aider à la prise de décision dans une entreprise. Contrairement à l'ERP qui gère les opérations quotidiennes, la BI se concentre sur l'analyse et l'interprétation des données pour détecter des tendances et améliorer la stratégie.

Exemple:

• Un responsable des ventes utilise un outil de BI (comme Power BI ou Tableau) pour analyser les performances des ventes des six derniers mois et identifier les produits les plus rentables.

3. Différence entre BI et ERP

Critère	ERP (Enterprise Resource Planning)	BI (Business Intelligence)
Objectif	Gérer et automatiser les opérations de l'entreprise	Analyser les données pour la prise de décision
Données traitées	Données opérationnelles (transactions, commandes)	Données analytiques (rapports, KPI, tendances)
Exemples d'outils	SAP, Oracle ERP, Odoo, Microsoft Dynamics	Power BI, Tableau, Qlik Sense, Google Data Studio
Question à laquelle il répond	"Que se passe-t-il en ce moment dans mon entreprise ?"	"Pourquoi cela s'est-il passé et que peut-on améliorer ?"
Utilisatio n	Saisie et gestion des données en temps réel (factures, stocks, employés)	Analyse des données historiques et actuelles pour détecter des tendances

3. Différence entre BI et ERP

Cas d'une entreprise de vente de vêtements en ligne:

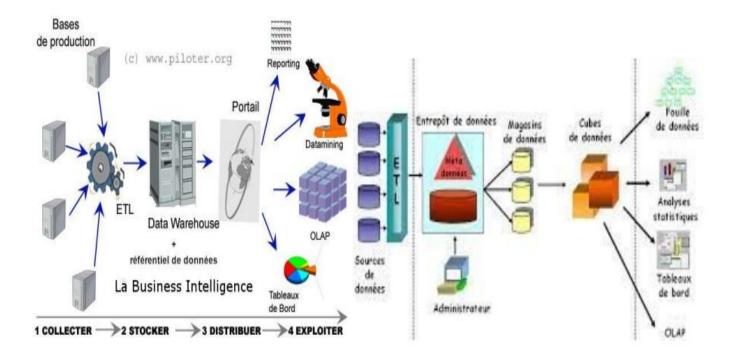
- L'ERP enregistre chaque vente, met à jour le stock et génère des factures pour les clients.
- La BI analyse les données de ventes des 12 derniers mois pour voir quels vêtements se vendent le mieux et à quelle période, permettant ainsi d'anticiper la demande et d'optimiser les achats.

II. Architecture classique d'une solution BI

Une solution de Business Intelligence (BI) repose sur une chaîne de traitement de l'information, depuis la collecte des données jusqu'à leur exploitation pour la prise de décision.

On peut simplifier cette architecture en quatre grandes étapes logiques:

- 1. Collecter les données
- 2. Stocker les données
- 3. Distribuer les données
- 4. Exploiter les données



1. Collecter: Les outils d'ETL (Extract, Transform, Load)

La première étape d'un projet BI est la collecte des données. Ces données proviennent de différentes sources (bases de données, ERP, CRM, fichiers Excel, API, etc.). Cependant, elles doivent être nettoyées et transformées avant d'être utilisées pour la prise de décision (ETL).

Sources de données :

Bases de données opérationnelles : MySQL, Oracle, PostgreSQL (ex : base clients, base commandes)

Fichiers: Excel, CSV, JSON

Web: APIs, réseaux sociaux, données ouvertes (Open Data)

Capteurs (IoT) : données temps réel

1. Collecter: Les outils d'ETL (Extract, Transform, Load)

ETL signifie Extract – Transform – Load :

Extract (extraire) : se connecter aux sources de données

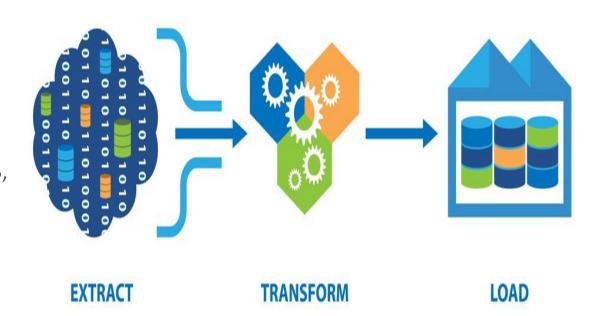
Transform (transformer): nettoyer, corriger, formater

Load (charger) : insérer les données dans un entrepôt

Exemples d'outils ETL : Talend, Apache NiFi, Microsoft SSIS, Pentaho, Azure Data Factory.

Exemple : Une entreprise de e-commerce collecte :

- Les commandes depuis sa base MySQL
- Les retours clients depuis un fichier Excel
- Les visites du site depuis Google Analytics (via API)



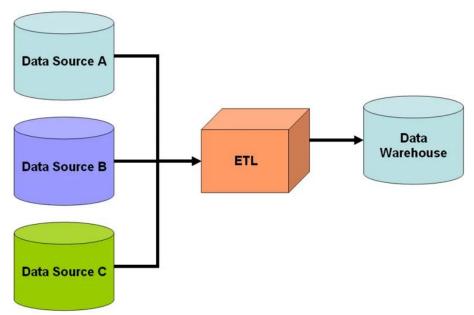
2. Stocker: Le Data Warehouse et le Data Mart

Une fois les données nettoyées et transformées, elles doivent être stockées dans un environnement optimisé pour l'analyse.

A. Data Warehouse (Entrepôt de données): Un Data Warehouse est une base de données centralisée contenant des données historiques et consolidées pour l'analyse décisionnelle.

Caractéristiques :

- Stocke des données structurées et optimisées pour la lecture.
- Facilite les analyses complexes et multidimensionnelles.
- Permet une meilleure performance que les bases de données transactionnelles.
- Historisation (suivi dans le temps)
- Accès rapide aux données via requêtes complexes



Exemple de Data Warehouse.

Une grande entreprise de vente au détail possède plusieurs magasins physiques et une plateforme de vente en ligne. Elle collecte des données provenant de différentes sources : ventes en magasin, ventes en ligne, stocks, clients, fournisseurs, etc.

Data Warehouse:

L'entreprise décide de créer un Data Warehouse pour centraliser toutes ces données dans un seul endroit. Ce Data Warehouse stocke des informations historiques et actuelles provenant de toutes les sources.

Par exemple, il contient des données sur les ventes des 5 dernières années, les stocks actuels, les informations sur les clients (historique d'achats, préférences), et les données des fournisseurs.

Les équipes de l'entreprise utilisent ce Data Warehouse pour analyser les tendances de ventes, prévoir les demandes futures, et prendre des décisions stratégiques comme l'ouverture de nouveaux magasins ou l'ajustement des stocks.

Avantage : Le Data Warehouse permet une vue globale et unifiée de toutes les données de l'entreprise, ce qui facilite l'analyse et la prise de décision. Aussi le Data Warehouse sert de source centrale, tandis que les Datamarts permettent des analyses plus ciblées et efficaces pour des équipes ou des projets spécifiques.

2. Stocker: Le Data Warehouse et le Data Mart

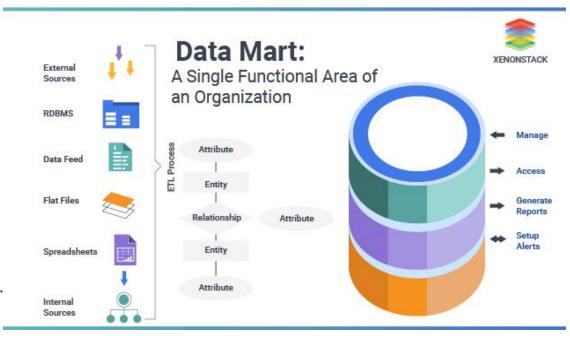
B. Data Mart: Un Data Mart est une version réduite d'un Data Warehouse, dédiée à un service ou une fonction spécifique (exemple : un Data Mart Finance, un Data Mart RH).

Avantages:

- Plus rapide à mettre en place
- Spécifique aux besoins d'un département métier

Exemple: Un Data Warehouse stocke toutes les données d'une entreprise (ventes, RH, logistique). Un Data Mart Finance extrait uniquement les données financières pour le service comptabilité.

Exemples de technologies : Snowflake, Amazon Redshift, Google BigQuery, Microsoft SQL Server



Data Mart can be called as a mini <u>Data</u>

<u>Warehouse design</u> that shows the contents to be needed only to the client-side, i.e. it holds the overview of the data. The primary purpose of using it is to get the only information required that can be not very easy to get from the Warehouse.

Exemple de Datamart :

Dans la même entreprise de vente au détail, le service marketing souhaite analyser spécifiquement les données clients pour améliorer les campagnes publicitaires.

Datamart:

Un Datamart est créé à partir du Data Warehouse, mais il est spécialisé et contient uniquement les données liées aux clients : historique des achats, préférences, données démographiques, etc.

Ce Datamart est plus petit et plus facile à utiliser pour l'équipe marketing, car il ne contient que les informations pertinentes pour leurs analyses.

Par exemple, le service marketing peut utiliser ce Datamart pour identifier les clients les plus fidèles, segmenter les clients par catégorie d'âge, ou analyser l'efficacité des campagnes publicitaires passées.

Avantage : Le Datamart est plus ciblé et plus facile à manipuler pour des analyses spécifiques, ce qui permet aux équipes de se concentrer sur leurs besoins .

3. Distribuer : Rendre les données accessibles

Une fois les données stockées et structurées dans l'entrepôt de données, il faut les rendre accessibles aux utilisateurs finaux de manière sécurisée, efficace et adaptée à leurs besoins.

Cette étape est souvent négligée, alors qu'elle est fondamentale pour la démocratisation de la donnée en entreprise.

Pourquoi faut-il distribuer les données ?

Les données brutes, même bien stockées, sont inutiles si elles ne sont pas accessibles aux bonnes personnes, au bon moment, dans un format compréhensible.

Les enjeux :

Accessibilité : comment permettre aux utilisateurs d'accéder aux bonnes données ?

Sécurité : comment garantir que seules les personnes autorisées y accèdent ?

Performance : comment permettre des requêtes rapides même sur de gros volumes ?

Pertinence : comment adapter les données au profil et aux besoins de chaque utilisateur ?

3. Distribuer : Rendre les données accessibles

Structuration des données pour l'accès :

Cubes OLAP (Online Analytical Processing): Les cubes OLAP sont des structures multidimensionnelles créées à partir du Data Warehouse. Ils permettent aux utilisateurs d'analyser les données selon plusieurs axes (appelés "dimensions").

Avantages:

Requêtes très rapides

Navigation intuitive (drill-down, roll-up)

Résumé des données pré-calculé (agrégation)

Exemples de dimensions : Temps (année, trimestre, mois), Géographie (pays, région, ville), Produit (catégorie, référence)

Outils: Microsoft SSAS, Oracle OLAP, SAP BW

3. Distribuer : Rendre les données accessibles

Gestion des accès et de la sécurité

La sécurité des données est un pilier essentiel de la distribution.

Principes de sécurité :

Authentification : vérifier l'identité de l'utilisateur (ex : login / mot de passe, SSO, LDAP)

Autorisation : gérer ce qu'il a le droit de voir ou de faire (ex : lecture seule, accès aux données d'un seul service)

Traçabilité : journaliser les connexions et les actions effectuées sur les données.

Exemple : Un responsable RH ne peut accéder qu'aux informations du personnel, mais pas aux données financières.

4. Exploiter : Assister l'utilisateur dans l'analyse des données

L'objectif final d'une solution BI est d'aider les utilisateurs à prendre des décisions basées sur les données.

Comment exploiter les données efficacement ?

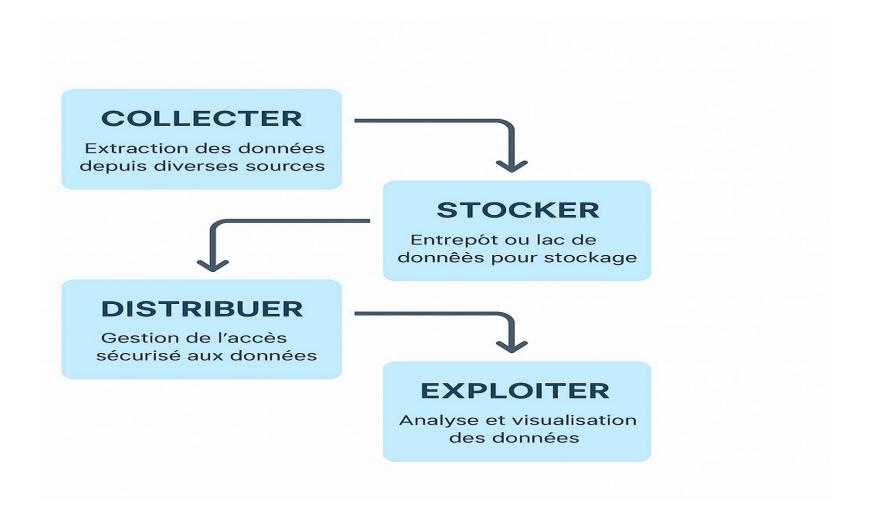
- Utilisation d'outils de visualisation de données (graphiques, cartes interactives, tableaux de bord)
- Fonctionnalités avancées d'exploration des données (filtrage, drill-down, comparaison temporelle)
- Automatisation des rapports et alertes (exemple : envoyer un email si les ventes baissent de 10%)

Outils BI pour l'exploitation des données : Power BI (Microsoft) Tableau Software, Qlik Sense, Google Data Studio

Exemple concret : Un analyste financier utilise Tableau pour identifier les produits les plus rentables, et décide d'augmenter la publicité sur ces produits pour maximiser les profits.



II. Architecture classique d'une solution BI



III. ETL et Data : Le Processus Extract-Transform-Load (ETL)

Le processus ETL (Extract, Transform, Load) est une étape clé dans la gestion des données en Business Intelligence (BI). Il permet de transférer des données depuis plusieurs sources, les nettoyer et les charger dans un entrepôt de données (Data Warehouse) pour analyse.

Pourquoi utiliser ETL:

- > Les données proviennent de sources variées (bases de données, fichiers Excel, API, ERP...).
- Les données doivent être nettoyées et homogénéisées avant d'être utilisées pour la prise de décision.
- Un Data Warehouse a besoin de données bien structurées pour offrir des analyses fiables.

1. Extraction : Récupérer les données

L'extraction consiste à collecter des données brutes depuis différentes source:

- > Identifier les sources de données pertinentes (bases de données, fichiers plats, API, services web, etc.).
- Récupérer des données brutes qui peuvent être incomplètes ou non structurées.

Méthodes d'extraction :

Extraction complète : Récupération de toutes les données d'une source.

Extraction incrémentale : Seules les nouvelles données ou mises à jour sont extraites (exemple : récupérer uniquement les ventes du dernier jour).

Extraction en temps réel : Mise à jour continue des données (utilisé dans les systèmes nécessitant des décisions rapides).

2. Transformation : Nettoyer et structurer les données

La transformation est une étape essentielle pour garantir la qualité et l'homogénéité des données avant de les stocker.

- > Convertir les données dans un format exploitable par le Data Warehouse.
- > Nettoyer les erreurs et supprimer les incohérences.
- > Standardiser les formats (dates, devises, unités de mesure, etc.).
- > Filtrer et supprimer les doublons pour éviter les erreurs d'analyse.

2. Transformation : Nettoyer et structurer les données

Transformation	Explication	Exemple
Formatage	Uniformiser les formats de données	Convertir "01/03/2025" en "2025-03-01"
Nettoyage	Supprimer les incohérences et erreurs	Supprimer les valeurs nulles ou erronées
Fusionner les sources	Regrouper les données de plusieurs bases	Fusionner les données clients de SAP et d'Excel
Suppression des doublons	Éliminer les données en double	Un même client enregistré plusieurs fois
Filtrage	Garder uniquement les données utiles	Exclure les clients inactifs depuis 5 ans
Traduction	Convertir des données d'une langue à une autre	Transformer "USD" en "Dollars américains"

Exemple technique:

Un ETL reçoit des données clients avec des formats de date différents. La transformation va convertir toutes les dates en YYYY-MM-DD pour garantir l'uniformité dans l'analyse.

3. Chargement : Insérer les données transformées dans un Data Warehouse

Une fois les données extraites et transformées, elles sont chargées dans un Data Warehouse ou un Data Mart.

- Insérer les données préparées dans le système cible.
- > Optimiser la performance des requêtes BI.
- > Sécuriser l'accès aux données en fonction des utilisateurs.

Méthodes de chargement :

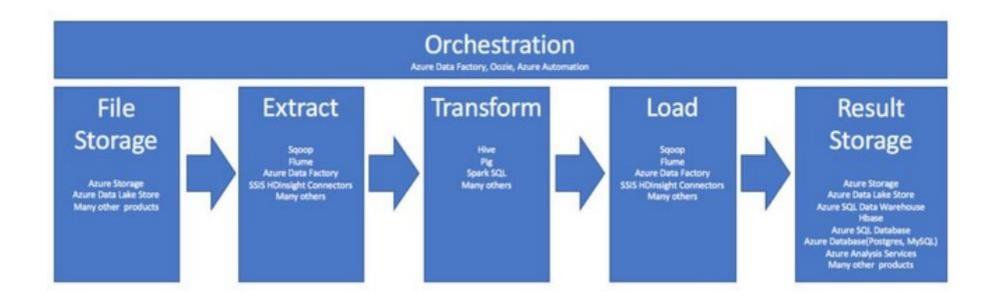
Chargement complet : Remplacement total des anciennes données par les nouvelles (utile pour les petits ensembles de données).

Chargement incrémental : Ajout uniquement des nouvelles données pour éviter d'écraser les anciennes (très utilisé en BI).

Chargement en temps réel : Mise à jour continue des données pour des systèmes nécessitant une actualisation immédiate.

Exemple technique : Un ETL met à jour quotidiennement un Data Warehouse avec les nouvelles commandes clients, permettant aux équipes BI d'analyser les tendances de vente en temps réel.

3. Chargement : Insérer les données transformées dans un Data Warehouse



Exemple concret d'un ETL en entreprise

Imaginons une entreprise de e-commerce qui veut analyser ses ventes.

Extraction : Récupération des données de ventes depuis un site web, un ERP et une base SQL.

Transformation: Conversion des prix en euros, suppression des doublons, correction des erreurs de date.

Chargement : Stockage des données dans un Data Warehouse pour créer des tableaux de bord avec Power BI.

1. Définition du Data Warehouse

Un Data Warehouse (entrepôt de données) est une base de données spécialisée dédiée au stockage et à l'analyse des données. Contrairement aux bases de données transactionnelles (OLTP), qui gèrent les opérations quotidiennes d'une entreprise (ventes, facturation, gestion des stocks), un Data Warehouse est conçu pour faciliter l'analyse et les prises de décision.

Exemple:

- Une entreprise de e-commerce utilise un ERP pour gérer les commandes en temps réel.
- Mais pour analyser l'évolution des ventes sur plusieurs années, elle stocke ces données dans un Data Warehouse.
- Les responsables marketing peuvent ensuite utiliser cette base pour identifier les tendances de vente par région, par produit ou par période.

Un Data Warehouse n'est pas une simple copie des bases de données opérationnelles. Il est structuré selon plusieurs principes fondamentaux qui garantissent la qualité des données pour l'analyse décisionnelle.

1. Orienté sujet

- Les données sont organisées par thèmes (ventes, clients, produits, finances, etc.)
- Contrairement aux bases transactionnelles OLTP qui gèrent des opérations, le Data Warehouse regroupe les données par sujets stratégiques.
- Exemple : Toutes les données liées aux ventes (commandes, produits, clients, régions) sont centralisées dans un même ensemble.

2. Intégré

- Les données proviennent de sources multiples et sont uniformisées.
- Un Data Warehouse reçoit des données hétérogènes de plusieurs systèmes (ERP, bases SQL, fichiers Excel...).
- Il harmonise ces données pour les rendre cohérentes et exploitables.

• Exemple :

- Une entreprise collecte les ventes de trois pays différents, où les devises sont en euros, dollars et livres sterling.
- Le Data Warehouse convertit toutes les valeurs en euros pour garantir une analyse cohérente.

3. Non volatile

- Les données ne sont pas modifiées ni supprimées.
- Contrairement aux bases transactionnelles, où les informations changent fréquemment, les données d'un Data Warehouse restent en lecture seule.
- Cela garantit l'intégrité des informations pour l'analyse.
- Exemple :
 - Lorsqu'un client passe une commande, un ERP peut mettre à jour le stock en temps réel.
 - Mais dans le Data Warehouse, les données sont conservées sans modification pour que les analystes puissent suivre l'évolution des stocks dans le temps.

4. Historisé

- Les données sont horodatées pour suivre leur évolution dans le temps.
- Un Data Warehouse stocke des données historiques pour analyser des tendances.
- On peut ainsi comparer les performances actuelles avec celles des années précédentes.

• Exemple :

- Une entreprise suit ses ventes mois par mois pour identifier les saisons où les produits se vendent le mieux.
- Un analyste peut alors observer les tendances sur 5 ans et adapter la stratégie commerciale en conséquence.

Les Tableaux de Bord en Business Intelligence (BI)

Les tableaux de bord sont une étape essentielle du processus de Business Intelligence (BI). Ils permettent de visualiser et interpréter les données transformées et stockées dans le Data Warehouse.

Tableaux de bords (MS Powerpivot)



Les Tableaux de Bord en Business Intelligence (BI)



Les Tableaux de Bord en Business Intelligence (BI)

À quel moment intervient le tableau de bord dans la BI?

Dans le processus de BI, un tableau de bord est utilisé à la dernière étape : l'exploitation des données.

Processus BI:

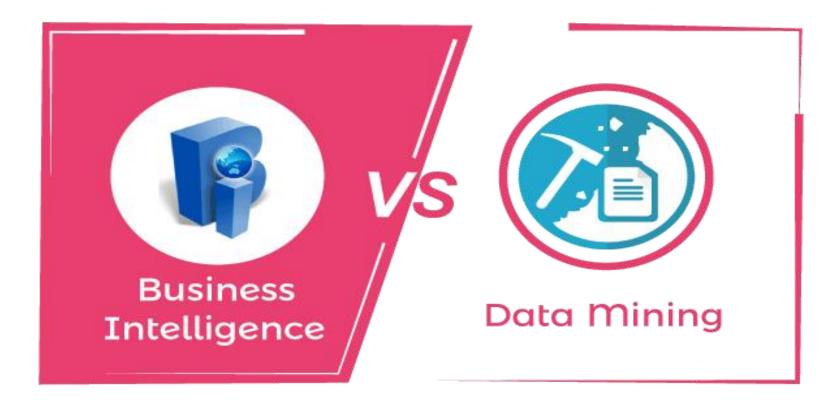
- Collecte des données (ETL) → On extrait, transforme et charge les données.
- Stockage (Data Warehouse / Datamarts) → Les données sont centralisées et organisées.
- Accès et analyse (OLAP, reporting, tableaux de bord) → On analyse et visualise les informations pour aider à la prise de décision.
- Les tableaux de bord font partie de cette troisième étape, où l'utilisateur final consulte et interprète les données pour prendre des décisions stratégiques

Pourquoi utiliser un tableau de bord ?

- Facilite la compréhension des données grâce aux graphiques et indicateurs visuels.
- Permet de prendre des décisions rapides grâce aux données mises à jour en temps réel.
- Aide à identifier les tendances et anomalies pour anticiper les actions à mener.

Bl vs Data Mining: Deux Approches Complémentaires

La Business Intelligence (BI) et le Data Mining sont deux disciplines liées à l'analyse des données, mais elles ont des objectifs différents et fonctionnent de manière complémentaire.



Bl vs Data Mining: Deux Approches Complémentaires

Business Intelligence (BI)	Data Mining
But : Présenter des données organisées pour faciliter la prise de décision.	But : Découvrir des motifs cachés et tendances dans de grands ensembles de données.
Outils : Tableaux de bord, reporting, OLAP, Data Warehouse.	Outils : Algorithmes de machine learning, statistiques, modélisation prédictive.
Utilisateurs : Managers, décideurs, analystes métiers.	Utilisateurs : Data scientists, analystes de données.
Approche : Analyse descriptive et visuelle des performances passées.	Analyse prédictive pour anticiper des comportements futurs.

Exemple : La BI analyse les ventes passées pour voir quel produit s'est le mieux vendu. Le Data Mining identifie des facteurs cachés influençant ces ventes (ex : les clients qui achètent un smartphone achètent souvent un avec).

2. Comment BI et Data Mining fonctionnent ensemble ?

Le Data Mining alimente la BI:

En analysant de grandes quantités de données, le Data Mining extrait des tendances cachées.

Ces découvertes sont ensuite intégrées dans les rapports et tableaux de bord BI.

La BI facilite le Data Mining :

La BI collecte et organise les données dans des Data Warehouses.

Ces bases structurées sont ensuite exploitées par le Data Mining pour construire des modèles prédictifs.

Exemple d'un site e-commerce :

BI: Affiche les produits les plus vendus sous forme de tableau de bord.

Data Mining: Analyse les comportements d'achat pour proposer des recommandations personnalisées.

BI: Intègre ces recommandations dans une interface utilisateur (ex: "Produits suggérés pour vous").

2. Comment BI et Data Mining fonctionnent ensemble ?

3. Exemples Concrets de Data Mining en Bl

Marketing et Vente:

BI: Analyse les campagnes publicitaires performantes.

Data Mining : Prédit quel type de clients est le plus susceptible de cliquer sur une publicité.

Secteur Bancaire:

BI: Suivi des transactions et détection des fraudes.

Data Mining: Repère des motifs anormaux pour identifier des fraudes potentielles.

Santé:

BI: Visualisation des dossiers patients et diagnostics courants.

Data Mining : Détecte des facteurs de risque pour certaines maladies.

Thank you