SERVICES INFORMATIQUEs AUX ORGANISATIONS

BLOC 1 (BASES DE DONNEES)

SEQUENCE 1 : LES BASES DE données



Table des matières

[1. Qu’est-ce qu’une base de données 3](#_Toc113353033)

[2. Comment fonctionne une base de données 3](#_Toc113353034)

[3. Une diversité de base de données 4](#_Toc113353035)

[4. Les bases de données relationnelles 4](#_Toc113353036)

[5. Le langage SQL 6](#_Toc113353037)

[6. Principes de conception d’une base de données 6](#_Toc113353038)

# Qu’est-ce qu’une base de données

Voici la définition proposée par Oracle : Une base de données est un ensemble d'informations qui est organisé de manière à être facilement accessible, géré et mis à jour. Elle est utilisée par les organisations comme méthode de stockage, de gestion et de récupération de l’informations.

Les données sont organisées en lignes, colonnes et tableaux et sont indexées pour faciliter la recherche d'informations. Les données sont mises à jour, complétées ou encore supprimées au fur et à mesure que de nouvelles informations sont ajoutées. Elles contiennent généralement des agrégations d'enregistrements ou de fichiers de données, tels que les transactions de vente, les catalogues et inventaires de produits et les profils de clients.

Généralement, [l’administrateur de la base de données](https://www.oracle.com/fr/database/administrateur-base-de-donnees-dba.html) régule les accès des utilisateurs afin de contrôler leurs actions et d’analyser les usages. Pour garantir la cohérence des données et l'intégralité des transactions, toutes les transactions réalisées sur une base de données doivent répondre aux exigences de la conformité [ACID](https://actualiteinformatique.fr/data/definition-acid-atomicite-consistance-isolation-et-durabilite) :

**Le principe d’Atomicité** garantit la bonne exécution de la transaction. Les transactions de base de données, comme les atomes, peuvent être décomposées en plus petites parties. Si une partie d'une transaction échoue, toute la transaction sera annulée.

**La propriété de Cohérence** signifie que seules les données qui suivent des règles prédéfinies peuvent être écrites dans la base de données.

**L’isolement** fait référence à la capacité de traiter simultanément plusieurs transactions de manière indépendante.

**La durabilité** requiert de rendre les défaillances invisibles pour l'utilisateur final. Les données sont sauvegardées une fois la transaction terminée, même en cas de panne de courant ou de défaillance du système.

# Comment fonctionne une base de données

Une base de données est stockée sous la forme d'un fichier ou d'un ensemble de fichiers sur un disque ou un disque magnétique, un disque optique ou tout autre support de stockage. L'information contenue dans ces fichiers peut être divisée en enregistrements. Ces enregistrements sont constitués d'un ou de plusieurs champs. Un champ constitue une seule pièce d’information, et chaque champ contient généralement des informations se rapportant à un aspect ou attribut de l'entité décrite par la base de données.

Les enregistrements sont également organisés en tableaux qui contiennent des informations sur les relations entre les différents champs. À l'aide de mots-clés et de diverses commandes de tri, les utilisateurs peuvent rapidement rechercher, réorganiser, regrouper et sélectionner les champs dans de nombreux enregistrements pour récupérer ou créer des rapports sur des agrégats particuliers de données.

# Une diversité de base de données

Les enregistrements et les fichiers de la base de données doivent être organisés de manière à faciliter la récupération de l'information. Les requêtes sont la principale façon dont les utilisateurs récupèrent les informations de la base de données. La puissance d'un [système de gestion de base de données](https://www.oracle.com/fr/database/systeme-gestion-base-de-donnees-sgbd-definition.html) s’évalue à travers sa capacité à définir de nouvelles relations à partir celles déjà présentes dans les tableaux et à les utiliser pour obtenir des réponses aux requêtes.

Pour les entreprises possédant de grande base de données, les nombreux utilisateurs doivent être en mesure de manipuler les informations qu'elle contient rapidement et à tout moment. De plus, ces entreprises ont tendance à constituer de nombreux fichiers indépendants contenant des données liées, voire des données se superposant. Dans le cadre de leur activité, il est souvent nécessaire de coupler les données provenant de plusieurs fichiers. Plusieurs types de base de données ont été développés pour répondre à ces exigences : base de donnée orientée texte, base de donnée hiérarchique, base de donnée réseau, [base de donnée relationnel](https://actualiteinformatique.fr/cloud/definition-base-de-donnee-relationelle), [base de donnée orienté objet](https://actualiteinformatique.fr/data/definition-object-oriented-database-management-system-oodbms) et [base de donnée distribuee](https://actualiteinformatique.fr/data/definition-base-de-donnees-distribuee).

Ces bases de données et les bases de données plus volumineuses sont devenues de plus en plus importantes dans la vie des entreprises, en partie parce qu'elles sont maintenant couramment conçues pour être intégrées à d'autres logiciels de bureau, y compris des tableurs.

# Les bases de données relationnelles

Pour résumer : une base de données est un ensemble de données structurées correspondant généralement à un domaine fonctionnel (facturation, ressources humaines, etc.). Physiquement, une base de données se matérialise par un ensemble de fichiers stockés sur un périphérique de stockage.

Les données d’une base de données sont gérées par un logiciel appelé système de gestion de base de données (SGBD). Ce logiciel offre plusieurs fonctionnalités : accès aux données, gestion des mises à jour, renforcement de l’intégrité, contrôle de la sécurité d’accès, etc.

Une base de données relationnelle supporte une organisation des données basée sur le modèle relationnel, développé en 1970 par Edgar Frank Codd. C’est la structure la plus répandue aujourd’hui.

Dans une base de données relationnelle, les données sont organisées en tables logiquement liées entre elles. Une table comporte un certain nombre de colonnes (ou champs) qui décrivent une ligne (ou enregistrement). La mise en relation des tables s’effectue par l’intermédiaire d’une colonne.

Exemple



Sur cet exemple, les tables livre et collection sont liées par les colonnes id\_collection de la table livre et id de la table collection.

L’interaction avec une base de données relationnelle s’effectue grâce au langage SQL. Ce langage permet la lecture et la mise à jour des données, mais aussi la définition de l’organisation des données, la gestion de la sécurité, le renforcement de l’intégrité, etc. Le langage SQL est un langage normalisé, mais les différents éditeurs de base de données ne respectent pas l’intégralité du standard.

# Le langage SQL

Le SQL est un langage informatique permettant de communiquer avec les SGBDR. La bonne nouvelle, c’est que ce langage est normé. Ainsi, le même code SQL peut être exécuté aussi bien sur un serveur SQL Server que sur un serveur Oracle. Du moins, c’est la théorie, car les SGBD ne respectent pas toujours la norme, ne l’ont pas implémentée entièrement ou ont ajouté des éléments supplémentaires. Il faut néanmoins essayer de coller le plus possible à cette norme.

SQL est un langage à base de requêtes. Elles correspondent à des ordres donnés au SGBD pour manipuler un ensemble de données : c’est un langage ensembliste et assertionnel : par exemple, modifier la structure d’une table pour changer le type d’un attribut, ou modifier les valeurs de certaines colonnes dans une table...

# Principes de conception d’une base de données

Dans une base de données relationnelle, l’objectif est de stocker dans des tables différentes les informations correspondant à des entités (objets) différentes du domaine fonctionnel. Le but est d’éviter les redondances et faire en sorte qu’une information donnée ne soit stockée qu’une fois. Sur notre exemple précédent, les informations sur l’auteur d’un livre ne sont pas stockées dans la table livre ; l’auteur d’un livre est une entité fonctionnelle à part entière qui est stockée dans une table séparée.

Ce processus de séparation des données dans plusieurs tables est appelé « normalisation ».

Une normalisation poussée à l’extrême peut nuire aux performances des requêtes d’interrogation qui doivent lire un grand nombre de tables. Pour améliorer les performances des lectures, il est alors envisageable de « dénormaliser » le modèle, en regroupant des tables, quitte à avoir des données redondantes dans les différentes lignes. Les bases de données des systèmes décisionnels, qui effectuent principalement des interrogations généralement complexes, sont très souvent dénormalisées. À l’inverse, les bases de données des systèmes transactionnels, qui effectuent principalement des petites interrogations simples et beaucoup de mises à jour, respectent bien le principe de normalisation.

Dans une base de données relationnelle, chaque table stocke les informations relatives à un objet métier concret ou abstrait qui doit être identifié.

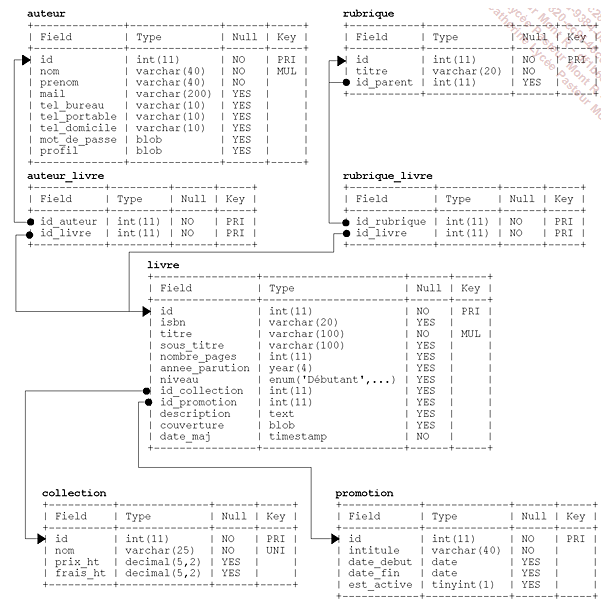
Dans la table, chaque colonne stocke une information unitaire (attribut, propriété) qui caractérise une ligne de la table. Chaque colonne possède un type de données (entier, chaîne de caractères, date, etc.) et peut être obligatoire ou non.

Une colonne ou combinaison de colonnes qui identifie de manière unique une ligne d’une table est appelée clé candidate. La valeur d’une clé candidate est différente pour toutes les lignes de la table (pas de doublon autorisé). Une clé candidate peut être constituée par une colonne arbitraire utilisée spécifiquement pour cela.

La clé primaire d’une table est une des clés candidates de la table, choisie plus ou moins arbitrairement, si ce n’est que les colonnes de la clé primaire doivent aussi être obligatoires ; il y a une seule clé primaire par table. Les autres clés candidates de la table sont alors appelées clés uniques.

Une colonne ou combinaison de colonnes d’une table qui référence une clé candidate d’une autre table (en général la clé primaire) est appelée clé étrangère. Une table peut avoir plusieurs clés étrangères.

Exemple



Le schéma ci-dessus présente le modèle de la base de données utilisée dans cet ouvrage. Ce modèle est un modèle simplifié de gestion des livres d’un éditeur.

Ce modèle comporte les tables suivantes :

|  |  |
| --- | --- |
| auteur | Auteurs des livres. |
| rubrique | Rubriques permettant le classement des livres dans différentes catégories (base de données, langage de développement, etc.). Les rubriques sont organisées sur deux niveaux : rubrique principale et sous-rubrique. Une sous-rubrique est rattachée à une rubrique parent par l’intermédiaire de la colonne id\_parent. Pour une rubrique parent, la colonne id\_parent est vide. |
| collection | Collections de l’éditeur dans lesquelles les livres sont publiés. |
| promotion | Promotions sur les livres. |
| livre | Livres publiés par l’éditeur. |
| auteur\_livre | Relation entre les auteurs et les livres : un auteur peut écrire plusieurs livres et un livre peut avoir plusieurs auteurs. |
| rubrique\_livre | Relation entre les rubriques et les livres : une rubrique peut contenir plusieurs livres et un livre peut appartenir à plusieurs rubriques. |

Dans toutes les tables, à l’exception de auteur\_livre et rubrique\_livre, la clé primaire est la colonne id. Pour les tables auteur\_livre et rubrique\_livre, la clé primaire est la combinaison des deux colonnes, respectivement (id\_auteur,id\_livre) et (id\_rubrique,id\_livre).

Dans l’état actuel du modèle, il y a une clé unique sur la colonne nom de la table collection. Dans le chapitre Construire une base de données dans MySQL, nous ajouterons des clés uniques sur d’autres tables.

La table livre comporte deux clés étrangères : id\_collection (vers la table collection) et id\_promotion (vers la table promotion).

La table auteur\_livre comporte deux clés étrangères : id\_auteur (vers la table auteur) et id\_livre (vers la table livre).

La table rubrique\_livre comporte deux clés étrangères : id\_rubrique (vers la table rubrique) et id\_livre (vers la table livre).

La table rubrique comporte une clé étrangère : id\_parent (vers la table rubrique).