

Numérique et sciences informatiques
Classe de terminale

III Bases de données
1. Modèle relationnel

I La notion de base de données

I.1 Principes d'une base de données

- Les services attendus : stockage des données, structuration des données, intégrité et cohérence de l'information, possibilité de mise à jour, possibilité d'interrogation, sécurisation de l'accès aux données, etc.
- Le principe d'une décomposition de l'information en un ensemble de données :
 - * décomposition en données « ponctuelles » qui peuvent se représenter par un type informatique élémentaire : nombre entier, nombre flottant, chaîne de caractère, booléen, date, heure, etc.
 - * il ne s'agit pas de tout dire sur une personne, un objet, un état, etc. : le choix des informations pertinentes dépend du contexte dans lequel on souhaite les utiliser

I.2 Le modèle relationnel

- Le modèle relationnel a été créé en 1970 par Edgar F. Codd, ingénieur chez IBM.
- Le principe est de représenter un ensemble d'objets, de personnes, d'états, etc. par un ensemble de tuples (ou de n-uplets).
 - * Chacun des tuples représente une des entités de cet ensemble
 - * Tous les tuples sont structurés de la même façon comme une succession de valeurs simples, chacune correspondant à une information particulière sur cette entité.
- On peut se représenter un tel ensemble d'entités comme un tableau à deux dimensions :
 - * Chaque ligne représente une des entités de cet ensemble
 - * Chaque colonne représente une information particulière sur les entités de cet ensemble.
- Un exemple : on stocke des informations sur les livres d'une bibliothèque dans un tableau suivant :

Id_ex	Isbn	titre	auteur	année
1	...	Le rouge et le noir	Stendhal	1830
2	...	Le père Goriot	Honoré de Balzac	1835
3	...	Thérèse Desqueyroux	François Mauriac	1927

- Le vocabulaire :
 - * un tel ensemble structuré est appelé une **relation** (ou une **table**).
 - * la structure d'une relation, c'est-à-dire la décomposition de chaque ligne en une succession d'informations bien spécifiées est appelé le **schéma de la relation**.
 - * Une ligne correspond à une **entité** ou un **enregistrement**.
 - * une colonne correspond à un **attribut**.
- Un attribut dans une relation est caractérisée par :
 - * son **nom**

* son **domaine de valeur**, ce qui correspond approximativement au type informatique de ses valeurs : Int (entier) String, Boolean, Date, etc.

- On utilise la notation suivante pour exprimer le schéma d'une relation :

nom_relation (attr1 Dom1, nom_attr2 Dom2, ...)

Remarque : nous ajouterons par la suite quelques éléments supplémentaires à cette notation quand nous aurons introduits les notions correspondantes.

- Une **base de données relationnelles** est un ensemble de relations. On peut donc se la représenter comme une succession de tableaux à deux dimensions, chacun correspondant un type d'entité différent : des personnes, des objets, des états, des liens entre entités, etc.

- Le **schéma d'une base de données** relationnelles est formé des schémas des relations qui la composent.

II Contraintes d'intégrité d'une base de données relationnelle

II.1 Contrainte de domaine

- Pour toutes entités d'une relation, la valeur associée à un attribut doit appartenir au domaine de valeur de cet attribut : être un entier ou un « flottant » ou une chaîne de caractères, etc.

II.2 Contraintes d'entité : la notion de clé primaire

- Dans une relation, chaque tuple représente un « objet » distinct des autres. Une entité (une ligne du tableau) doit donc être distincte des autres entités, c'est-à-dire que deux tuples distincts ne peuvent pas avoir exactement les mêmes valeurs à tous les attributs. Autrement dit, il ne doit pas y avoir de doublons dans le tableau.

- Pour assurer cette condition, on définit un attribut (ou un ensemble d'attributs) comme **clé primaire** de la relation.

* Si un attribut de la relation constitue la clé primaire alors chaque enregistrement de la relation possède une valeur distincte pour cet attribut. Autrement dit, deux enregistrements distincts ne peuvent pas avoir la même valeur pour cet attribut

* Si un ensemble d'attribut forment la clé primaire d'une relation alors deux enregistrements ne peuvent pas avoir les mêmes valeurs pour tous les attributs qui forment la clé primaire. Il doit donc y avoir au moins une différence.

Remarque : en pratique, on prend le plus souvent un seul attribut pour former la clé primaire.

- Attention il ne peut y avoir **qu'une seule** clé primaire pour une relation donnée.
- Très souvent on crée un attribut identifiant dont le domaine de valeur est Int et dont la fonction principale est de former la clé primaire de la relation.