

Normalisation des bases de données

Pourquoi normaliser

- ✓ Éliminer la redondance
- ✓ Éviter les incohérences
- ✓ Éviter la perte d'informations

- ✓ Données atomiques
- ✓ Non décomposables
(nom = Pierre Legrand)
- ✓ Pas de listes
(hobbies = lecture, sport, voyages)
- ✓ Données stables dans le temps
- ✓ Pas de colonnes multiples avec la même sémantique
(hobby1, hobby2, hobby3)
- ✓ Toutes les valeurs d'une colonne sont de même type

Exemple

ID_Client	Nom	Téléphones
1	Alice	123456789, 987654321
2	Bob	456789123

données non atomiques

Solution

ID_Client	Nom	Téléphone
1	Alice	123456789
1	Alice	987654321
2	Bob	456789123

Doublons sur id

Table clients

clé primaire

ID_Client	Nom
1	Alice
2	Bob

Table telephones

clé étrangère

ID_Client	Téléphone
1	123456789
1	987654321
2	456789123

Autre exemple

ID_Etudiant	Nom	Cours1	Cours2	Cours3
1	Alice	Math	Physique	Chimie
2	Bob	Biologie	-	-
3	Claire	Histoire	Littérature	Philosophie

Table "etudiants"

ID_Etudiant	Nom
1	Alice
2	Bob
3	Claire

Table "affectations"

ID_Etudiant	Cours
1	Math
1	Physique
1	Chimie
2	Biologie
3	Histoire
3	Littérature
3	Philosophie

Normaliser cette table

parent1	parent2	parent1_age	parent2_age	enfant1	enfant1_age	Enfant2	enfant2_age
Pierre Konrad	Aoife Konrad	28	29	Siobhan	3	Seamus	1

La dépendance fonctionnelle

Définition : il existe une Dépendance Fonctionnelle (DF) entre deux attributs A et B d'une relation R si à chaque valeur a_i de A est toujours associée la même valeur b_j de B.

Pour simplifier :
si on connaît A, on connaît B

Notation : $A \rightarrow B$ (A détermine B)

Exemple de DF

ID_Employé	Nom	Département	Salaire
101	Dupont	Comptabilité	3000
102	Martin	Marketing	3500
103	Dubois	Comptabilité	3000
104	Leroy	Marketing	3500

ID_employé détermine toutes les autres colonnes

Département pourrait déterminer Salaire

(à condition que celui-ci soit identique pour tous les employés d'un même département)

Dépendance transitive

Département	Responsable
Comptabilité	Mme Dupuis
Marketing	M. Bernard

ID_Employé	Nom	Département	Salaire
101	Dupont	Comptabilité	3000
102	Martin	Marketing	3500
103	Dubois	Comptabilité	3000
104	Leroy	Marketing	3500

ID_employé → Responsable

Dépendance partielle

ID_Etudiant	ID_Cours	Nom_Etudiant	Nom_Cours	Date_Inscription
1	101	Alice	Mathématiques	2024-09-01
2	102	Bob	Physique	2024-09-02
1	102	Alice	Physique	2024-09-01

Ici la clef primaire est composée des colonnes
`ID_Etudiant et ID_Cours`

`Nom_Etudiant` ne dépend que d'une partie de la clef

- ✓ Est conforme à 1NF
- ✓ Toutes les colonnes non clés dépendent entièrement de la clé primaire (pas de dépendances fonctionnelles partielles).

Exemple

ID_Commande	ID_Produit	Nom_Produit	Quantité	Prix_Unitaire	Date_Commande
1	A123	Chaise	4	25	2024-12-01
1	B456	Table	1	100	2024-12-01
2	A123	Chaise	2	25	2024-12-05

La clef primaire est composée de ID_Commande et ID_Produit
Il y a des dépendances partielles sur les autres colonnes

ID_Produit → Nom_Produit et Prix_Unitaire

ID_Commande → Date_Commande

Totalité de la clef → Quantité

1. Table Commandes :

ID_Commande	ID_Produit	Quantité	Date_Commande
1	A123	4	2024-12-01
1	B456	1	2024-12-01
2	A123	2	2024-12-05

2. Table Produits :

ID_Produit	Nom_Produit	Prix_Unitaire
A123	Chaise	25
B456	Table	100

Exercice

Vendeur_ID	Client_ID	Nom_Client	Ville_Client	Montant_Vente
101	1	Dupont	Paris	500
102	2	Martin	Lyon	300
101	3	Leroy	Marseille	700

- ✓ Identifier la clé primaire.
- ✓ Identifier les dépendances fonctionnelles partielles.
- ✓ Diviser la table pour la mettre en 2NF.

- ✓ Est conforme à 1NF et 2NF
- ✓ Aucune colonne non clef ne dépend transivement de la clef primaire.

Une dépendance transitive existe si $A \rightarrow B$ et $B \rightarrow C$, donc $A \rightarrow C$, cela introduit une redondance inutile

Exemple

ID_Employé	Nom	ID_Departement	Nom_Departement	Responsable_Departement
1	Dupont	D01	Comptabilité	Mme Durand
2	Martin	D02	Marketing	M. Bernard
3	Leroy	D01	Comptabilité	Mme Durand

La clef primaire est ID_Employé

ID_Produit -> {Nom, ID_Departement, Nom_Departement,
Responsable_Departement}

ID_Departement -> {Nom_Departement, Responsable_Departement}

Il y a redondance sur les colonnes Nom_Departement et Reponsable_Departement

Solution

1. Table Employés :

ID_Employé	Nom	ID_Departement
1	Dupont	D01
2	Martin	D02
3	Leroy	D01

2. Table Départements :

ID_Departement	Nom_Departement	Responsable_Departement
D01	Comptabilité	Mme Durand
D02	Marketing	M. Bernard

Exercice

ID_Vente	ID_Client	Nom_Client	Ville_Client	Montant_Vente
1	101	Dupont	Paris	500
2	102	Martin	Lyon	300
3	101	Dupont	Paris	700

- ✓ Identifier la clé primaire.
- ✓ Identifier les dépendances transitives et autres.
- ✓ Diviser la table pour la mettre en 3NF.

- ✓ Est conforme à 1NF, 2NF et 3NF
- ✓ La clef primaire est une super clef, c'est à dire qu'elle détermine de façon unique toutes les colonnes de la table

Exemple

ID_Employé	ID_Projet	Nom_Projet
1	P1	Alpha
2	P1	Alpha
3	P2	Beta

$ID_{Projet} \rightarrow Nom_{Projet}$

Hors ici la clef primaire est :

soit $\{ID_{Employé}, ID_{Projet}\}$

soit juste $ID_{Employé}$

ID_{Projet} n'est donc pas une super clef

1. Table Employés_Projets :

ID_Employé	ID_Projet
1	P1
2	P1
3	P2

2. Table Projets :

ID_Projet	Nom_Projet
P1	Alpha
P2	Beta

Table : Livres_Auteurs

ISBN	Auteur_ID	Nom_Auteur
123	1	Hugo
123	2	Dumas
456	1	Hugo

- ✓ Identifier la clé primaire.
- ✓ Identifier les dépendances.
- ✓ Diviser la table pour la rendre conforme à BCNF.

- ✓ Est conforme à 3NF et inférieur
- ✓ Ne contient aucune dépendance multivaluée non triviale.

Une dépendance multivaluée existe lorsqu'une colonne dans une table peut être associée à plusieurs valeurs d'une autre colonne, indépendamment des autres colonnes.

Exemple

Table : Compétences_Projets

Employé_ID	Compétence	Projet
1	Programmation	Application Web
1	Gestion de projet	Application Web
1	Programmation	Maintenance
1	Gestion de projet	Maintenance

La clef primaire est : {Employé_ID, Compétence, Projet}

$\text{Employé_ID} \rightarrow\!\!> \text{Compétence}$

Un employé peut avoir plusieurs compétences

$\text{Employé_ID} \rightarrow\!\!> \text{Projet}$

Un employé peut travailler sur plusieurs projets

Problème

La présence des deux dépendances multivaluées indépendantes entraîne une redondance. Chaque compétence est associée à chaque projet, même si ces deux informations ne sont pas liées.

1. Table Employé_Compétences :

Employé_ID	Compétence
1	Programmation
1	Gestion de projet

2. Table Employé_Projets :

Employé_ID	Projet
1	Application Web
1	Maintenance

Exercice

ID_Etudiant	Cours	Club
1	Math	Tennis
1	Physique	Tennis
1	Math	Théâtre
1	Physique	Théâtre

- ✓ Identifier la clé primaire.
- ✓ Identifier les dépendances.
- ✓ Diviser la table pour la rendre conforme à 4NF.

- ✓ Est conforme à 4NF et inférieur
- ✓ Toutes les dépendances de jointure dans la table sont triviales.
Cela signifie que si une table est divisée en plusieurs parties, il est possible de la reconstruire uniquement en effectuant des jointures sans perdre d'informations ni introduire d'incohérences.

Une dépendance multivaluée existe lorsqu'une colonne dans une table peut être associée à plusieurs valeurs d'une autre colonne, indépendamment des autres colonnes.

Exemple

Table : Enseignants_Cours_Étudiants

Enseignant	Cours	Étudiant
Dupont	Math	Alice
Dupont	Math	Bob
Martin	Physique	Alice
Martin	Physique	Bob

Les relations :

- ✓ Un enseignant peut enseigner plusieurs cours.
- ✓ Un cours peut être suivi par plusieurs étudiants.
- ✓ Chaque enseignant enseigne son cours à tous les étudiants inscrits.

Cette table est redondante, car la relation entre Enseignant et Étudiant peut être dérivée des relations entre Enseignant-Cours et Cours-Étudiant.

Solution

1. Table Enseignants_Cours :

Enseignant	Cours
Dupont	Math
Martin	Physique

2. Table Cours_Étudiants :

Cours	Étudiant
Math	Alice
Math	Bob
Physique	Alice
Physique	Bob

Exercice

Partenaire	Projet	Service
Société A	Projet X	Service 1
Société A	Projet X	Service 2
Société B	Projet Y	Service 1
Société B	Projet Y	Service 3

- ✓ Identifier la clé primaire.
- ✓ Identifier les dépendances.
- ✓ Diviser la table pour la rendre conforme à 5NF.

- ✓ Est conforme à 5NF et inférieur
- ✓ La table ne contient aucune dépendance jointe non triviale, même dans des situations très spécifiques.

Une dépendance jointe signifie que les données d'une table peuvent être décomposées en plusieurs tables plus petites, et ces tables peuvent être recomposées (reconstruites) par une jointure sans perte d'information.

La 6NF introduit le concept de représentation irréductible : chaque table ne contient qu'une seule relation atomique entre colonnes. Cela est particulièrement utile dans les bases temporelles pour gérer des relations qui évoluent dans le temps.

Table : Contrats_Employés

Employé_ID	Contrat_ID	Salaire	Date_Début	Date_Fin
1	C01	3000	2024-01-01	2024-06-30
1	C01	3200	2024-07-01	2024-12-31

Salaire → Employé_ID et Date_Début.

Contrat_ID → Employé_ID.

Ces dépendances sont mélangées, ce qui peut entraîner des redondances si les contrats ou les salaires changent fréquemment.

Solution

1. Table Employé_Contrat :

Employé_ID	Contrat_ID
1	C01

2. Table Salaire_Temps :

Employé_ID	Salaire	Date_Début	Date_Fin
1	3000	2024-01-01	2024-06-30
1	3200	2024-07-01	2024-12-31

Projet_ID	Ressource	Date_Dispo	Date_Fin
P01	Serveur A	2024-01-01	2024-06-30
P01	Serveur B	2024-01-01	2024-06-30

- ✓ Identifier la clé primaire.
- ✓ Identifier les dépendances.
- ✓ Diviser la table pour la rendre conforme à 6NF.