

Pourquoi analyser ? Exemple de stockage de données dans une seule table

La table ci-dessous contient de données concernant les bureaux d'une entreprise internationale et des employés qu'y travaillent :

(un bureau a un certain volume de ventes)

bureau	ville	adresse	ventes	idNational	nom	prenom	telephone
3	Gent	55 Hoofkerkstraadt	330.294,00 €	1201 AG	Oyler	Jeanne	128654413
3	Gent	55 Hoofkerkstraadt	330.294,00 €	8383 AP	Plantz	Ed	205927177
3	Gent	55 Hoofkerkstraadt	330.294,00 €	7296 TQ	Plantz	Sofia	362620947
5	Bruxelles	80 Rue de la Tulipe	804.024,00 €	7055 QS	Naley	Linnea	022224780
6	Lisboa	99 Rue Peixe	444.369,00 €	7524 LA	Troher	Maria	169694594
7	Vienne	90 LeopoldStraat	676.802,00 €	6045 WX	Bogdanovich	Izzy	872362038
7	Vienne	90 LeopoldStraat	676.802,00 €	5321 EN	Willis	Sigrid	
8	Milan	100 Via del templo di Giove	541.366,00 €	0792 HS	Baltec	Andrea	762612545
8	Milan	100 Via del templo di Giove	541.366,00 €	1108 ST	Dumont	Marie	318380090
8	Milan	100 Via del templo di Giove	541.366,00 €	1677 RL	Rubio	Ricky	147397717
9	Frankfurt	13 Berliner Strasse	119.018,00 €	5308 FU	De Sousa	Lena	733401808
10	Milan	12 Via Meravigli	900.366,00 €	5969 EF	Dean	Will	149741538
10	Milan	12 Via Meravigli	900.366,00 €	2200 YD	Stannard	Albert	837986543
11	Londres	12 Market Street	828.859,00 €	8667 TT	Archer	Malgorzata	059974511
11	Londres	12 Market Street	828.859,00 €	0741 NC	Gildersleeve	Babet	641205438
11	Londres	12 Market Street	828.859,00 €	9223 CD	Meterson	Cathy	460870468

A première vue notre méthode de stockage est correcte (une seule table contenant toutes les informations)... mais considérez maintenant les problèmes suivantes:

1. Redondance

- Pour chaque nouvel employé de la société on doit répéter toutes les informations de la société. S'il s'agit d'une grande base de données, ça impliquera un énorme gaspillage d'espace de stockage (disque).

2. Modification

- Si on modifie l'adresse d'un certain bureau on devra la modifier dans tous les enregistrements qui correspondent à ce bureau. Si on oublie une ligne ou on fait une erreur de saisie, nos données deviendront incohérentes.

3. Insertion

- Pour introduire les données d'un employé on est obligé d'introduire tout de suite les données du bureau où il travaille. Et s'il n'a pas encore un bureau et on veut juste stocker les données de l'employé?
Laisser vides les champs concernant son bureaux ce n'est pas une option : ça nous obligerait à remplir tous ces champs manuellement plus tard ! (bureau, ville, adresse et ventes)
- Même chose pour les bureaux... et si on vient de créer un bureau mais il n'y a pas encore d'employés qu'y travaillent ?

4. Elimination

- Si on efface tous les employés d'un bureau de la table... est-ce que le bureau existe encore ? On pourrait remplir les champs concernant les bureaux et laisser vides les champs concernant l'employé, mais ce n'est pas cohérent vu que chaque enregistrement de la table lie un bureau et un employé.

Dans cet exemple, pour éviter ces quatre types de problèmes il suffit de séparer les données concernant les **employés** et les **bureaux** dans deux tables comme ci-dessous :

Table **Bureaux** :

bureau ▾	ville ▾	adresse ▾	ventes ▾
1	Francistown	32844 West Sierra Leone Way	376.904,00 €
2	Oslo	35325 West Central African Republic Ln.	362.936,00 €
3	Gent	55 Hoofkerkstraadt	330.294,00 €
5	Bruxelles	80 Rue de la Tulipe	804.024,00 €
6	Lisboa	99 Rue Peixe	444.369,00 €
7	Vienne	90 LeopoldStraat	676.802,00 €
8	Milan	100 Via del templo di Giove	541.366,00 €
9	Frankfurt	13 Berliner Strasse	119.018,00 €
10	Milan	12 Via Meravigli	900.366,00 €
11	Londres	12 Market Street	828.859,00 €

Table **Employés** :

idNational ▾	nom ▾	prenom ▾	telephone ▾	bureau ▾
1201 AG	Oyler	Jeanne	128654413	3
7296 TQ	Plantz	Sofia	362620947	3
8383 AP	Plantz	Ed	205927177	3
7055 QS	Naley	Linnea	022224780	5
7524 LA	Troher	Maria	169694594	6
5321 EN	Willis	Sigrid		7
6045 WX	Bogdanovich	Izzy	872362038	7
0792 HS	Baltec	Andrea	762612545	8
1108 ST	Dumont	Marie	318380090	8
1677 RL	Rubio	Ricky	147397717	8
5308 FU	De Sousa	Lena	733401808	9
2200 YD	Stannard	Albert	837986543	10
5969 EF	Dean	Will	149741538	10
0741 NC	Gildersleeve	Babet	641205438	11
8667 TT	Archer	Malgorzata	059974511	11
9223 CD	Meterson	Cathy	460870468	11

champ de liaison

- Observez qu'on a rajouté un champ **bureau** ou **idBureau** dans la table **Employés** pour indiquer le bureau où chaque employé travaille

Grace à cette séparation en plusieurs tables, on va résoudre les quatre problèmes précédant :

1. Redondance

- Résolu : chaque bureau est présent une seule fois dans la base de données, ses données ne se répètent plus.

2. Modification

- Résolu : si on modifie l'adresse d'un certain bureau il faut la modifier une seule fois, aucun risque d'incohérence.

3. Insertion

- Résolu : on peut introduire les données d'un employé sans le lier tout de suite à un bureau en laissant juste un champ vide. L'employé se trouvera juste dans la table employés.

Note : on établira une règle dans le SGBD pour empêcher à l'utilisateur d'indiquer un code de bureau inexistant pour un employé

4. Elimination

- Résolu : On peut effacer les employés de la table Employés et il n'y a aucune conséquence dans la table Bureaux.

Besoin d'analyser

Dans l'exemple précédant on a uniquement deux tableaux. C'est très facile de séparer les données en deux domaines (Employés et Bureaux), et trouver la liaison n'est pas une tâche compliquée. Dans une BD réelle il peut y avoir de dizaines de domaines et on ne peut pas créer la base de données « manuellement » (ex : la BD pour gérer un aéroport aura des informations concernant les avions, les pilotes, les vols, les aéroports, les compagnies, les terminales... comment séparer tout ça en tableaux et réaliser les liaisons ?). Heureusement on compte sur les outils d'analyse pour pouvoir mener à bonne fin de BD réelles.