

3) Tables de données

Nous avons découvert lors du cours 1 quelques formats pour stocker des données : JSON, xml ou CSV. Le JSON est actuellement le plus utilisé des trois dans les bases de données. Cependant, une majorité des bases de données stockent celles-ci en tables.

a) Qu'est-ce qu'une table de données ?

Une table de données est tout simplement une organisation des données dans un tableau.

Par exemple, pour enregistrer les données relatives à des fichiers mp3 de morceaux de musique, on les classe de cette façon dans une table :

ID	DURÉE	TITRE	STYLE	DATE	INTERPRÈTE
MORCEAU#1	04:34	ROUX COOL	REGGAE	1993	BOB
MORCEAU#2	04:34	LOVE BIRD	JAZZ	1956	ELSA
MORCEAU#3	03:25	RE VOLE LOVE	RAP	2017	MC
MORCEAU#4	08:45	BLUE PIGEON	ROCK	1987	PIXEES

Regardez la vidéo suivante : <https://www.youtube.com/watch?v=IJJgcZ2DEs0>

Dans une table, chaque ligne correspond à un **élément**. Les colonnes représentent les **attributs (ou descripteurs)** d'un élément.

Pour notre exemple, (morceau#1 , 04:34 , ROUX COOL , REGGAE , 1993 , BOB) est un élément. " STYLE " est un attribut.

Pour un attribut donné, les valeurs sont du même type.

Pour notre exemple, toutes les dates sont des entiers (int), tous les styles sont des chaînes de caractères (string).

Ce passage des données brutes à des données classées s'appelle une indexation de table : l'index permet un accès rapide à un élément de la table. Un peu comme l'index à la fin d'un livre permet de retrouver facilement une information dans ce livre.

Pour s'assurer d'avoir bien compris, faites le quiz suivant (non ramassé) :

<https://parcours.algorea.org/contents/4707-4702-1352246428241737349-1070539248445908636-35415640138424232-1655405614123744253/>

b) Intérêt d'une table de données

Une table de données est intéressante car elle permet de :

- sélectionner des données selon un ou des critères (exemple : toutes les chansons avant 1980)
- trier des données (exemple : classer les chansons par ordre croissant de durée)
- retrouver ou pas une donnée (exemple : y a-t-il un titre de DAFT PUNK dans la table?)
- calculer (exemple : durée moyenne des morceaux, nombre de morceaux de reggae)

Nous verrons aussi qu'en classant des données dans plusieurs tables en leur donnant un attribut commun, cela permet de limiter les modifications à faire dans une table si une donnée change. De plus, en croisant plusieurs tables, on parvient à isoler de nouvelles informations : essayez de répondre à cette énigme : <https://parcours.algorea.org/contents/4707-4702-1352246428241737349-1070539248445908636-1194788383492460195-624462471231889418/>

Incroyable ! On retrouve le nom des malades en croisant ces 2 tables.