

Chapitre n° 14 : commentaires fin

1 Tri d'une table

1.1 Présentation

1.2 Trier des données en fonction d'une clef

1.3 Ordre lexicographique et stabilité

1.4 Exemple d'application

2 Fusion de tables

2.1 Présentation

Le but est de regrouper des données de plusieurs tables différentes dans une seule même table, pour ensuite pouvoir sélectionner, trier, etc.

Le regroupement fonctionnera différemment selon que les tables à regrouper possèdent les mêmes attributs (*C.F. 2.2 Réunion de tables*) ou des attributs différents (mais avec au moins un attribut en commun, *C.F. 2.3 Opération de jointure*).

2.2 Réunion de tables

Le point essentiel à retenir est que *la réunion de tables n'est possible que si les tables présentent les mêmes attributs (même nom et même type de données pour chaque attribut)*.

Si cette condition est vérifiée, la réunion se fait grâce par une concaténation des tableaux, grâce à l'opérateur « + ».

Le reste du paragraphe montre un exemple pratique de réunion de tables avec des données réelles. les fichiers CSV et les programmes sont mis à votre disposition, vous pouvez tester par vous même (ne pas hésiter à ouvrir les fichiers CSV dans un tableur, et tester par vous même les programmes Python, que vous devez être capable de comprendre).

Il montre également un contre-exemple déclenchant une erreur, car un des tableaux ne possède pas l'attribut demandé au moment de la sélection de lignes.

2.3 Opération de jointure

Le point essentiel à retenir est que *la jointure de tables n'est possible que si les tables présentent au moins un attribut en commun. La jointure permettra alors de compléter les lignes pour qu'elles possèdent les attributs des deux tables*.

Il faut comprendre l'algorithme avec la double boucle. Il faut également comprendre la façon de le coder avec une double compréhension.

En plus de l'exemple présenté dans ce paragraphe, voici une autre utilisation possible de l'opération de jointure. On pourrait imaginer une table de donnée avec une liste de films et leurs années de sortie, et une autre table de donnée avec une liste de films avec les acteurs. La jointure de ces tables permettraient de réunir toutes ces informations sur les films en une seule table.

2.4 Utilisation d'un identifiant unique

L'exemple présenté illustre les problèmes qui se posent lorsqu'il y a ambiguïté sur l'identification des données (ce qui n'est pas rare). La solution est assez simple : il s'agit d'ajouter un attribut (donc une clef avec la valeur correspondante à chaque dictionnaire, c'est à dire chaque ligne de la table de données). La numérotation donne ainsi un identifiant unique à chaque ligne de la table.