

## SQL

Toutes les questions (à part la première) nécessitent d'écrire une requête SQL, ou une expression d'algèbre relationnelle.

Faites attention à la présentation (indentation en python...).

## 1 Base de données

Une grande chaîne de magasins souhaite pouvoir analyser les achats faits par ses différents clients.

On définit une base de données avec le schéma relationnel suivant :

- clients (`idc: integer`; `nom: text`; `prenom: text`; `annee_naissance: integer`)
- produits (`idp: integer`; `nom: text`; `marque: text`; `prix: real`) (le prix est en euros)
- achats (`client: integer`; `produit: integer`; `magasin: integer`; `date: string`)
- magasin (`idm: integer`; `ville: string`; `departement: integer`)

Remarques :

- Un type de produit donné apparaîtra une seule fois dans la table produits mais pourra correspondre à plusieurs achats.
- Si un client achète plusieurs produits, cela correspondra à plusieurs lignes dans la table achats

- QUESTION 1 — Dans la table `achats`, les attributs `'client'` et `'produit'` sont des clés étrangères : expliquer.  
— Que représentent les attributs soulignés ?

## 2 SQL

- QUESTION 2 SQL : Quels sont les noms/prénoms des clients nés après l'an 2000 ?  
Écrire l'expression d'algèbre relationnelle associée.

- QUESTION 3 SQL : Combien de noms de familles différents apparaissent dans la table `'clients'` ?

- QUESTION 4 SQL : On veut la liste des prénoms et noms de tous les clients, triés par prénom croissant et, en cas d'égalité, par nom décroissant.

## 3 Jointures

- QUESTION 5 SQL : Pour chaque achat, on veut le nom de famille de chaque client, et la ville où il a fait ses courses. Sans doublons.  
Écrire l'expression d'algèbre relationnelle associée.

- QUESTION 6 On veut la liste des achats des clients dont le nom commence par un "A" et qui ont payé moins de 5€.

## 4 Agrégats

- QUESTION 7 Pour chaque département, quel est le nombre de magasins ? (avec le numéro du département)  
Écrire l'expression d'algèbre relationnelle associée.

- QUESTION 8 Quel est le prix total payé par chaque client, ainsi que son nom et son prénom ?

- QUESTION 9 Donner la liste des (identifiants des) clients ayant acheté plus de 7 produits à un prix unitaire supérieur à 1000€.  
Écrire l'expression d'algèbre relationnelle associée.

- QUESTION 10 Donner le prix moyen d'achat, pour chaque client (seulement le numéro du client).

- QUESTION 11 Donner le prix moyen d'achat, pour chaque département (oui, il faut faire une jointure).

## 5 Requêtes imbriquées - Maxima

- QUESTION 12 Quel est le dernier achat effectué ? (toutes les informations)

- QUESTION 13 Donner l'identifiant du deuxième produit le plus cher, de deux manières différentes (dans un des cas, on pourra utiliser `'LIMIT a, b'` qui permet de donner 'a' lignes après avoir "oublié" les 'b' premières).

## 6 Plus dur

- QUESTION 14 On veut la liste des prénoms des personnes les plus jeunes (celles qui ont l'année de naissance la plus...)

- QUESTION 15 Rappel : un booléen en SQL est une valeur 0 ou 1.  
Calculer, en une seule requête SQL, le prix moyen des achats de strictement moins de 100€ (<) et de ceux de plus de 100€ (>=).

- QUESTION 16 On rajoute une table  
`echanges (client_donneur: integer, client_releveur: integer, produit: integer)`  
qui permet d'organiser des échanges d'objets entre clients.  
Écrire une requête SQL qui permet de donner le nom des deux clients et du produit.

- QUESTION 17 Donner les numéros de départements dans lesquels on a payé au total plus que la moyenne nationale (moyenne du total département par département).

## 7 Python

Dans cette section, on travaille sur les mêmes données, mais on oublie SQL. On suppose que les informations sont disponibles directement dans Python.

- QUESTION 18 Écrire une fonction `moyenne(t)` python qui calcule la moyenne des éléments d'une liste. Puis une autre pour la variance, `variance(t)`, qu'on peut calculer avec  $V = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2$  où  $m$  est la moyenne des  $x_i$ .

- QUESTION 19 On suppose qu'on a les informations de la table `achats` sous forme de liste de quadruplets en python. Écrire une fonction `plus_chers(achats)` qui renvoie la liste des identifiants des produits qui sont plus chers que la médiane des prix.

Remarque : On suppose pour l'instant que la fonction `médiane` existe. Attention à la 'forme' des objets manipulés.

- QUESTION 20 Écrire une fonction `mediane(t)` qui prend en entrée une liste de flottants et renvoie "la" médiane<sup>1</sup>.  
Remarque : si vous utilisez une fonction de python utile (comme `sorted`), expliquez pourquoi.

- QUESTION 21 On suppose qu'on a la table `produits` sous forme d'une liste de triplets, triés par date. Expliquez comment chercher efficacement l'ensemble des produits qui sont étiquetés avec une date donnée. Quel algorithme classique peut-on utiliser et quelle est sa complexité ?

- QUESTION 22 On travaille encore avec les produits, mais pas forcément triés, et on souhaite obtenir l'identifiant du deuxième achat le plus cher. Écrire une fonction python (on cherchera à écrire une version en temps linéaire).

1. Wikipedia : la médiane d'un ensemble de valeurs est une valeur  $x$  qui permet de couper l'ensemble des valeurs en deux parties égales : mettant d'un côté une moitié des valeurs, qui sont toutes inférieures ou égales à  $x$  et de l'autre côté l'autre moitié des valeurs, qui sont toutes supérieures ou égales à  $x$  (s'il y a un nombre impair de valeurs, la valeur centrale sera mise des deux côtés).