

TP n° 15 : Bases de données

Ouvrir **SQLiteMan** puis y ouvrir la base de données **presidentielles.db** qui se trouve quelque part dans un répertoire intitulé **CPGE** (ou à télécharger depuis le site de M. Schleck).

On voit que cette base de données est constituée de dix tables :

- **communes** avec 12 attributs regroupant différentes informations à propos des communes de France métropolitaine (numéro de département, superficie, population, etc.) ;
- **salaires** avec 4 attributs dont **salaire** qui représente le salaire moyen dans la commune ;
- 8 tables dont le nom commence par **pres**, une pour chacun des deux tours des quatre dernières élections présidentielles en France, qui contiennent des informations sur les résultats des élections (nombres d'inscrits, de votants, etc.).

Le but de ce TP est de formuler des requêtes SQL permettant d'extraire des informations de ces différentes tables.

A. Avec une seule table

Pour commencer, on s'intéresse seulement à la table **communes**.

1. Écrire une requête qui détermine les numéros de département (**code_dep**) des villes s'appelant Troyes (*ne pas oublier les guillemets pour les chaînes de caractères et respecter la casse, c'est-à-dire les capitales et les minuscules*).
2. Même question avec les villes s'appelant Allonnes. Est-ce que l'attribut **nom_commune** est une clé primaire ?
3. Écrire une requête qui donne la population et la superficie de votre ville préférée. *On fera attention au fait que le nom de la commune ne suffit pas à l'identifier de façon unique. En revanche, il n'y a pas deux villes ayant le même nom dans un même département.*
4. Écrire une requête qui détermine la population française.
5. Écrire une requête donnant le numéro et le nombre d'habitants de chaque département.
6. Écrire une requête qui détermine le nombre de villes moyennes en France, c'est-à-dire comptant entre 20 000 et 100 000 habitants.
7. Écrire une requête permettant d'obtenir le nom et le nombre d'habitants des dix villes les moins peuplées de France.
Aparté historique : oui il y a des villes sans aucun habitant dans la Meuse, ce sont des villages détruits lors de la première guerre mondiale qui n'ont pas été reconstruits.

B. Toutes les tables

On va maintenant s'intéresser à toutes les tables ou, plus précisément, soit l'une des dix tables soit plusieurs tables en même temps grâce à des jointures.

B.1. La dernière présidentielle

8. Écrire une requête donnant le numéro du département, le nom de la commune et le nombre de votants des villes où personne ne s'est abstenu au second tour de la dernière présidentielle.
9. Écrire une requête permettant de connaître le taux de participation (c'est-à-dire le pourcentage de votants par rapport au nombre d'inscrits) au second tour de la dernière présidentielle.

B.2. Quelques jointures

Comme vu au début du TP, le nom de la commune n'est pas une clé primaire. On utilisera donc le couple (code_dep, code_commune) pour effectuer les jointures.

10. Reprendre la question précédente mais en considérant seulement les villages de moins de 1000 habitants.
11. Écrire une requête déterminant le nombre de communes dans lesquelles il y a eu strictement plus de votants au second tour qu'au premier tour de la dernière présidentielle.
12. Écrire une requête donnant le nom de chaque département accompagné du nombre de communes de celui-ci dans lesquelles le nombre de bulletins blancs ou nuls au second tour de la présidentielle de 2017 est au moins deux fois supérieur à celui du second tour de la présidentielle de 2012.
13. Écrire une requête qui donne le pourcentage de votants pour E. Macron au second tour de la dernière présidentielle parmi les villes de plus de 10 000 habitants dont le salaire horaire net moyen est d'au moins 20 euros par personne. (*Dans la table salaires, l'attribut salaire correspond au salaire horaire net moyen, en euros.*)

B.3. HAVING

Pour filtrer selon une condition qui dépend d'une fonction d'agrégation, il faut utiliser la condition **HAVING** au lieu de **WHERE**. Très souvent, cette utilisation est couplée à l'utilisation de **GROUP BY**.

14. Écrire une requête qui détermine le numéros des départements et leur population à condition que celle-ci soit supérieure ou égale à un million.
15. On a vu au début du TP que plusieurs communes peuvent avoir le même nom. Écrire une requête donnant les noms qui sont commun à au moins dix communes.

C. Pour aller plus loin

16. Que fait la requête ci-dessous ?

```
SELECT nom_commune
FROM communes
WHERE z_moyen = (SELECT z_moyen
                 FROM communes
                 WHERE nom_commune = "TROYES")
```

Remarque : on parle de « sous-requête » ou de requêtes « imbriquées ». L'utilisation des parenthèses est indispensable, pas celle de l'indentation.

17. Écrire une requête (constituée de deux requêtes imbriquées) qui calcule le nombre moyen d'habitants dans un département français.
18. Écrire une requête qui permet de déterminer le département le moins peuplé de France. On pourra donner deux versions : avec ou sans sous-requête.
19. Même question en ajoutant le nom du département (qui est présent par exemple dans la table `presidentielles2017_t2`).