

Feuille de travaux pratiques n° 3 Requêtes SQL (suite)

Exercice 3.1 — La table des prénoms

Téléchargez l'archive `prenoms.zip` sur Moodle, puis extrayez le fichier `prenoms.db`. Lancez ensuite **DB Browser for SQLite**, puis ouvrez la base de données `prenoms.db`.

Présentation

Cette base de données contient une partie de la base de données des prénoms fournie par l'INSEE. Pour chaque prénom, il est indiqué pour chaque année de naissance (de 2000 à 2020) et chaque sexe, le nombre de personnes inscrites à l'état civil sous ce prénom. Pour la table « `prenoms` », la précision est apportée pour chaque département.

Les personnes prises en compte

Le champ couvre l'ensemble des personnes nées en France hors Mayotte et enregistrées à l'état civil sur les bulletins de naissance. Les personnes nées à l'étranger sont exclues.

Le champ des prénoms retenus

Dans les fichiers de l'état civil, en l'occurrence les bulletins de naissance, les différents prénoms sont séparés par une espace (ou blanc). Ainsi deux prénoms séparés par un tiret constituent un seul prénom composé (exemple : Anne-Laure). Le premier prénom simple ou composé figure en début de liste, et c'est celui qui sera retenu après le traitement de la protection de l'anonymat.

Conditions portant sur les prénoms retenus

1. Sur la période allant de 1900 à 1945, le prénom a été attribué au moins 20 fois à des personnes de sexe féminin et/ou au moins 20 fois à des personnes de sexe masculin
2. Sur la période allant de 1946 à 2020, le prénom a été attribué au moins 20 fois à des personnes de sexe féminin et/ou au moins 20 fois à des personnes de sexe masculin
3. Pour une année de naissance donnée, le prénom a été attribué au moins 3 fois à des personnes de sexe féminin ou de sexe masculin

Les effectifs des prénoms ne remplissant pas les conditions 1 et 2 sont regroupés (pour chaque sexe et chaque année de naissance) dans un enregistrement dont le champ prénom (`PREUSUEL`) prend la valeur spéciale `_PRENOMS_RARES`.

Documentation (fournie par l'INSEE)

Cette relation `prenoms` comporte cinq attributs : `sexe`, `preusuel`, `annais`, `dpt`, `nombre`.

Nom	Intitulé	Type	Longueur	Modalité
<code>sexe</code>	sexe	entier	1	1 pour masculin, 2 pour féminin
<code>preusuel</code>	premier prénom	chaîne de caractères	25	
<code>annais</code>	année de naissance	entier	4	2000 à 2020
<code>dpt</code>	département de naissance	entier	3	liste des départements
<code>nombre</code>	fréquence	entier	8	

Vous allez exécuter des requêtes SQL (à l'aide de l'onglet **Exécuter le SQL** pour répondre aux questions suivantes.

Conseil : sauvegardez la liste de vos requêtes SQL dans un fichier .sql. Vous pouvez inclure des commentaires dans le fichier (lignes commençant par : --).

Attention : tous les prénoms sont en capitale dans la table.

Requêtes SQL

Vous pouvez taper vos requêtes à la suite dans l'onglet « Exécuter le SQL ». Le résultat est celui de la dernière requête. À l'issue du TP, pensez à sauvegarder le fichier SQL

1. Liste des prénoms de la table. (1110019 lignes)
2. Liste des années de la table. (1110019 lignes)
3. Département et année de naissance des personnes s'appelant Laetitia. (581 lignes)
4. Département et année de naissance des garçons prénommés Camille. (1438 lignes)
5. Liste des prénoms de garçons donnés en Corse en l'an 2000. (122 lignes)
6. Liste des prénoms donnés plus de 400 fois dans un département une année donnée, on affichera prénom, département, année. (4876 lignes)
7. Liste des prénoms donnés strictement moins de 5 fois une année, avec l'année correspondante, dans les Alpes-Maritimes. (6650 lignes)
8. Liste des prénoms de filles donnés strictement plus de 10 fois entre 2000 et 2004 dans les Alpes-de-Haute-Provence. Proposez deux requêtes. (16 lignes)
9. Liste des prénoms donnés exactement 80 fois dans un département dont le numéro est strictement supérieur à 80 en 2005. (3 lignes)
10. Liste des prénoms de garçons commençant par Y donnés en 2000 dans les Alpes-Maritimes. Afficher prénom et nombre. (18 lignes)
11. Liste des prénoms commençant par « KHL ». Afficher les prénoms, départements et années. (28 lignes)
12. Liste des prénoms composés communs entre 2000 et 2003. Afficher les prénoms, départements, années et effectifs. (1932 lignes)

Options d'affichage

13. Liste des prénoms avec toutes les informations en renommant les colonnes : prénom, année, département, sexe, effectif. (1105864 lignes)
14. Afficher la liste des prénoms donnés en 2001 en Corse en triant par effectif croissant. (236 lignes)
15. Afficher la liste des prénoms donnés en 2001 en Corse en triant par effectif décroissant. (236 lignes)
16. Afficher la liste des prénoms distincts de filles qui apparaissent dans la table. (20207 lignes)
17. Afficher la liste des prénoms distincts commençant par « Al » et terminant par « ce » apparaissant dans la table. (6 lignes)
18. Liste des prénoms de filles commençant par « C » donnés en 2000 en Corse. Afficher les prénoms et le nombre, trier par ordre croissant de nombre. (17 lignes)
19. Liste des prénoms (distincts) communs donnés entre 2000 et 2020 qui ont deux caractères quelconques, puis terminent par « ice ». (3 lignes)
20. Liste des prénoms (distincts) communs entre 2000 et 2020 et qui comportent exactement 6 lettres. Afficher le prénom uniquement. (2746 lignes)
21. Liste des prénoms (distincts) communs entre 2000 et 2020 et qui comportent exactement 7 lettres, dont un Z. Afficher le prénom uniquement.
22. Liste des prénoms (distincts) non composés communs entre 2000 et 2020 qui commencent par « Ab » et contiennent au moins deux fois la lettre A. Afficher le prénom, le département, l'année et l'effectif. Trié par année décroissante, par effectif croissant, puis par prénom dans l'ordre alphabétique.
23. Afficher la liste des prénoms distincts communs (pas rares) donnés entre 2000 et 2020.
24. Afficher les prénoms communs (non rares) donnés à des garçons en région Provence-Alpes-Côte-d'Azur entre 2002 et 2006. On affichera prénom et nombre, on triera par ordre décroissant de nombre.

Utilisation multiple de la table

25. Afficher le prénom, le département, le nombre de naissances en 2017, le nombre de naissances en 2018 pour les prénoms qui ont eu des naissances dans le même département en 2017 et en 2018.
26. Afficher les départements et les prénoms correspondants pour lesquels le nombre de naissances en 2018 est strictement supérieur à celui en 2017.
27. Afficher les départements pour lesquels il existe un prénom commençant par « K » tel que le nombre de naissances en 2018 dans le département est strictement supérieur à ce nombre en 2017.
28. Afficher (sans répétition) les prénoms épiciènes donnés dans les Alpes-Maritimes en 2018.
29. Afficher (sans répétition) les prénoms donnés pour une fille en 2000 dans un département et pour un garçon en 2007 dans un autre département. Préciser les deux départements.

Remarque

Ce fichier est en fait un sous-ensemble de la table fournie par l'INSEE qui recense les prénoms de 1900 à 2020. Sur le site <https://www.insee.fr/fr/statistiques/2540004>, vous pouvez trouver ces données pour la France entière ou par département, sous forme de fichiers csv. Ce fichier csv peut être importé dans **DB Browser for SQLite**.

Exercice 3.2 — La base « comparateur de territoires »

Téléchargez `commune.db` sur Moodle. Ce fichier contient une unique table : `comparateur`, qui présente une trentaine d'indicateurs décrivant la population, les logements, les revenus, l'emploi et les établissements au niveau communal.

Dictionnaire de variables fourni par l'INSEE

- CODGEO : Code du département suivi du numéro de commune ou du numéro d'arrondissement municipal
- LIBGEO : Libellé de la commune ou de l'arrondissement municipal pour Paris, Lyon et Marseille
- REG : Région
- DEP : Département
- P16_POP : Population en 2016
- P11_POP : Population en 2011
- SUPERF : Superficie (en km²)
- NAIS1116 : Nombre de naissances entre le 01/01/2011 et le 01/01/2016
- DECE1116 : Nombre de décès entre le 01/01/2011 et le 01/01/2016
- P16_MEN : Nombre de ménages en 2016
- NAISD18 : Nombre de naissances domiciliées en 2018
- DECESD18 : Nombre de décès domiciliés en 2018
- P16_LOG : Nombre de logements en 2016
- P16_RP : Nombre de résidences principales en 2016
- P16_RSECOCC : Nombre de résidences secondaires et logements occasionnels en 2016
- P16_LOGVAC : Nombre de logements vacants en 2016
- P16_RP_PROP : Nombre de résidences principales occupées par propriétaires en 2016
- NBMENFISC16 : Nombre de ménages fiscaux en 2016
- PIMP16 : Part des ménages fiscaux imposés en 2016
- MED16 : Médiane du niveau de vie en 2016
- TP6016 : Taux de pauvreté en 2016
- P16_EMPLT : Nombre d'emplois au lieu de travail en 2016
- P16_EMPLT_SAL : Nombre d'emplois salariés au lieu de travail en 2016
- P11_EMPLT : Nombre d'emplois au lieu de travail en 2011
- P16_POP1564 : Nombre de personnes de 15 à 64 ans en 2016
- P16_CHOM1564 : Nombre de chômeurs de 15 à 64 ans en 2016

- P16_ACT1564 : Nombre de personnes actives de 15 à 64 ans en 2016
- ETTOT15 : Total des établissements actifs au 31 décembre 2015
- ETAZ15 : Établissements actifs de l'agriculture, sylviculture et pêche au 31/12/2015
- ETBE15 : Établissements actifs de l'industrie au 31/12/2015
- ETFZ15 : Établissements actifs de la construction au 31/12/2015
- ETGU15 : Établissements actifs du commerce, transports et services divers au 31/12/2015
- ETGZ15 : dont Établissements actifs du commerce et réparation automobile au 31/12/2015
- ETOQ15 : Établissements actifs de l'administration publique, enseignement, santé et action sociale au 31/12/2015
- ETTEF115 : Établissements actifs de 1 à 9 salariés au 31 décembre 2015
- ETTEFP1015 : Établissements actifs de 10 salariés ou plus au 31 décembre 2015

Requêtes SQL

1. Code géographique de Nice.
2. Communes de Côte-d'Or. Afficher code commune, nom, population de 2016. (700 lignes)
3. Même question mais en portant la condition sur CODGEO. (700 lignes)
4. Communes ne comptant aucun habitant en 2011, ni en 2016. Afficher seulement le nom et le département. (6 lignes)
5. Communes comptant au moins deux fois plus d'habitants en 2016 qu'en 2011 (et habitée en 2011). Afficher seulement le nom. (5 lignes)
6. Communes dont le nombre d'habitants a été divisé par 2 ou plus de 2011 à 2016. Afficher le nom, la population en 2011 et 2016. Trier par population en 2016 décroissante. (4 lignes)
7. Communes des Alpes-Maritimes de strictement plus de 3000 habitants en 2011 dont la population a baissé de 2011 à 2016. Afficher le nom et les nombres d'habitants en 2011 et 2016. (19 lignes)
8. Communes de Gironde pour lesquelles le nombre de logements est strictement supérieur à la population en 2016. (9 lignes)
9. Communes du Finistère dont le nom contient « Ker ». Afficher le nom et la population en 2016. (7 lignes)
10. Communes du Cantal de 1000 habitants au plus en 2016 avec au moins 5 établissements actifs du commerce et de la réparation automobile. Afficher le nom, la population en 2016 et le nombre de tels établissements. Trier par ce dernier nombre dans l'ordre décroissant. (45 lignes)
11. Départements comprenant au moins une commune dont le nom commence par « Saint-Maurice ». Afficher les numéros de département sans répétition. (35 lignes)
12. Communes des Alpes-Maritimes dont la part de population de 15 à 64 ans en 2016 est supérieure ou égale à 70%. Afficher le nom de la commune, le pourcentage de la population de 15 à 64 ans. Nommer cette dernière colonne « Pourcentage ». (5 lignes)
13. Communes de Corse dont la population est supérieure à 20000 habitants en 2011 et en 2016. Afficher le code géographique de la commune, le nom, la population en 2011 et en 2016. (2 lignes)
14. Évolution de la population des communes du Val-de-Marne (94) de 2011 à 2016. Afficher le nom de la commune, la différence de population. (47 lignes)
15. Taux de croissance de la population de 2011 à 2016 des communes du Bas-Rhin (67) dont le nom commence par S. Afficher, nom de la commune, taux de variation de la population (en %). Afficher les résultats triés par taux décroissant. (65 lignes)
16. Communes du Lot-et-Garonne (47) dont la population a augmenté au moins de 1% de 2011 à 2016. Afficher le nom, la population en 2016 et le taux de variation de 2011 à 2016 (en %). (161 lignes)
17. Densité de population des communes du Puy-de-Dôme (63) en 2011 et 2016. Afficher le nom de la commune, le nombre d'habitants par km² en 2011 et 2016. Trier par densité décroissante en 2016. (464 lignes)
18. Nombre de logements occupés dans les communes du Rhône (69) en 2016. Afficher le nom de la commune, le département, le nombre de logements occupés. (267 lignes)
19. Nombre moyen d'habitants par logement dans les communes de Haute-Garonne dont le nom contient la lettre « T ». Afficher le nom de la commune et le nombre recherché. Trier par nom dans l'ordre inverse de l'ordre alphabétique. (30 lignes)

20. Départements qui comportent une commune dont le taux de logements vacants est supérieur ou égal à 50% en 2016. Ne pas considérer les communes sans habitant. Afficher le numéro de département, on ne veut pas de répétition. (2 lignes)
21. Nom des communes des Vosges (88) avec le nombre d'établissements actifs par habitant (sans distinguer le nombre de salariés) en 2016. (507 lignes)
22. Départements qui comportent une commune dont le nom commence par Y et dont le nombre d'établissements de plus de 10 salariés est supérieur ou égal à celui de 1 à 9 salariés. Éviter les répétitions éventuelles. (1 ligne)

Exercice 3.3 — Projets

Analyse de l'existant

Un employé est identifié par son matricule et on mémorise son nom, son poste, son salaire, la prime qu'il reçoit et sa date d'embauche.

Chaque employé travaille dans un département (commercial, direction, production, etc.) et a ou non un autre employé pour supérieur.

On stocke pour chaque département son numéro identifiant, son nom et son lieu. Chaque projet est réalisé par un ou plusieurs employés et chaque employé participe à un ou plusieurs projets.

Pour chaque projet, on stocke son code identifiant et son nom.

Dictionnaire de données

Attribut	Type	Description
matr	entier(11)	Matricule de l'employé
nome	char(20)	Nom de l'employé
poste	char(10)	Poste de l'employé
dateemb	date	Date d'embauche de l'employé
salaire	réel(10,2)	Salaire en euros
prime	réel(10,2)	Prime en euros
sup	entier(11)	Matricule du supérieur
numd	entier(6)	Numéro identifiant un département
nomd	char(20)	Nom du département
lieu	char(10)	Lieu du département
codep	entier(11)	Code identifiant un projet
nomp	char(20)	Nom du projet
fonction	char(20)	Fonction de l'employé dans le projet

Schéma relationnel

```
emp(matr, nome, poste, dateemb, salaire, prime, sup, numd)
dept(numd, nomd, lieu)
projet(codep, nomp)
participe(matr, codep, fonction)
```

1. Récupérez le fichier `projets.sql` sur Moodle. Ouvrez ce fichier avec un éditeur de texte (**gedit** par exemple).

Création des tables et insertion des données

2. Lancez **DB Browser for SQLite**. Créez une base de données `projets.db`. Allez sur l'onglet **Exécuter le SQL**.
3. Copiez les requêtes de création des tables depuis le fichier `sql` et exécutez-les.
4. Pour l'insertion des données, il manque la requête correspondant à `dept`. Écrivez la requête SQL pour insérer les données suivantes :

numd	nomd	lieu
10	finance	Paris
20	recherche	Nice
30	vente	Lyon
40	fabrication	Marseille

- Insérez les autres données (en les copiant depuis le fichier SQL). Attention à l'ordre d'insertion des données.

Requêtes SQL sur une seule table

- Sélectionnez tous les champs (colonnes) et tous les n-uplets (lignes) de la table emp. Votre requête doit comporter des clauses SELECT et FROM seulement (14 lignes).
- Affichez la liste des matricules et noms des employés (14 lignes).
- Liste des matricules et noms des employés du département 20. Une clause WHERE doit permettre de sélectionner seulement les employés du département numéro 20 (5 lignes).
- Liste des postes des employés dont le salaire est supérieur à 1500 € (8 lignes).
- Matricules et noms des employés dont le salaire est compris entre 1000 et 2000 € (8 lignes).
- Matricules et noms des employés embauchés entre le 1er janvier 2017 exclus et le 1er Juillet 2017 inclus. Les dates sont au format 'AAAA-MM-JJ' (7 lignes).
- Affichez les matricules et noms des employés dont le poste est 'ingénieur' ou bien 'commercial' (6 lignes).
- Affichez les matricules et noms des employés des départements 20 et 30 (11 lignes).
- Affichez les matricules et noms des employés qui sont des commerciaux ou bien qui travaillent dans le département numéro 30 (6 lignes).
- Affichez les employés de poste 'ingénieur' du département numéro 20 ('recherche') et ceux de poste 'secrétaire' dans le département numéro 30 ('vente') (3 lignes).
- Affichez pour tous les employés qui sont des commerciaux, leur matricule, nom et la rémunération totale perçue (salaire + prime) (4 lignes).
- Matricules et noms des employés qui ne touchent pas de prime. Un employé ne touche aucune prime si le champ prime prend la valeur NULL (10 lignes).
- Liste des employés qui touchent une prime et dont la rémunération (salaire + prime) est supérieure ou égale à 2000 € (3 lignes).
- Matricules et noms des employés dont le salaire est inférieur à la prime. Vérifiez bien le résultat en faisant attention aux employés qui ne touchent pas de prime (2 lignes).
- Matricules et noms des employés dont le salaire est supérieur à la prime. Vérifiez bien le résultat pour les employés qui ne touchent pas de prime (12 lignes).
- Matricules et noms des employés dont le nom commence par un 'B' (3 lignes).
- Matricules et noms des employés dont la 2^e lettre du nom est un 'e' (6 lignes).
- Matricule et nom des employés qui ont un supérieur. Un employé qui n'a pas de supérieur a la valeur NULL pour le champ sup (matricule de l'employé qui est le supérieur) (13 lignes).
- Matricule et nom des employés qui n'ont pas de supérieur (1 lignes).

Options d'affichage

- Liste des divers postes existants, en n'affichant chaque poste qu'une seule fois. Utilisez le mot-clé DISTINCT (5 lignes).
- Matricule, nom et numéro de département de tous les employés. Renommez les colonnes afin de faciliter la lecture du résultat (14 lignes).
- Liste de tous les employés par ordres décroissants des salaires et des primes : la prime sera utilisée en cas d'égalité sur le salaire. Utilisez une clause ORDER BY (14 lignes commençant par (1090, Leroy, president, NULL, 2017-01-01, 5000, NULL, 10)).

28. Numéros des départements dont au moins un employé touche une prime. Chaque numéro de département doit n'apparaître qu'une seule fois (30).

Requêtes SQL avec jointures

29. Affichez toutes les informations sur chaque employé et le département où il travaille. Utilisez le champ * dans la clause SELECT afin d'afficher tous les champs des 2 tables (14 lignes).
30. Modifiez le requête précédente afin que le numéro du département n'apparaisse qu'une seule fois (14 lignes).
31. Matricules, noms, rémunérations (i.e. salaire + prime) et nom du département pour tous les employés du département 'recherche'. Faites afficher la somme directement dans la clause SELECT (5 lignes).
32. Matricules, noms et salaire des employés qui travaillent dans le département 'recherche' et dont le salaire est supérieur à 2000 € (3 lignes).
33. Noms des départements dans lesquels au moins un employé touche une prime. Chaque nom de département doit n'apparaître qu'une seule fois (vente).
34. Affichez pour chaque participation d'un employé à un projet, le matricule et le nom de l'employé, le code et le nom du projet, ainsi que la fonction de l'employé dans ce projet (8 lignes).
35. Modifiez la requête précédente pour qu'elle n'affiche que les participations des employés du département 20 (6 lignes).
36. Affichez la liste des départements des employés qui participent à un projet dont le nom contient le terme « tramway » (3 lignes).