

Feuille de travaux pratiques n° 2

Premières requêtes SQL

Exercice 2.1 — Joueuses de tennis

1. Téléchargez le fichier `wta2021.db` qui contient une base de données des résultats des matches du circuit féminin WTA en 2021, puis ouvrez-le dans **DBBrowser for SQLite**.
2. Utilisez DBML pour décrire les données correspondant à cette base de données.
3. À l'aide de l'onglet **Parcourir les données** de **DBBrowser for SQLite**, répondez aux questions suivantes.
 - (a) Afficher tous les noms de tournois.
 - (b) Trier les tournois par date, du dernier au premier.
 - (c) Afficher tous les tournois de mars 2021.
 - (d) Afficher toutes les joueuses dont le nom est Ostapenko.
 - (e) Trier les joueuses français figurant dans la table, de la plus jeune à la plus ancienne.
 - (f) Afficher toutes les joueuses nées un 1^{er} avril.
4. Reprendre les questions précédentes et écrire les requêtes SQL correspondantes.

Exercice 2.2 — 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 « prénoms », 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 prénoms comporte cinq attributs : `sexe`, `preusuel`, `annais`, `dpt`, `nombre`.

Nom	Intitulé	Type	Longueur	Modalité
sexe	sexe	entier	1	1 pour masculin, 2 pour féminin
preusuel	premier prénom	chaîne de caractères	25	
annais	année de naissance	entier	4	2000 à 2020
dpt	département de naissance	entier	3	liste des départements
nombre	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. (14 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. (2887 lignes)
21. Liste des prénoms (distincts) communs entre 2000 et 2020 et qui comportent exactement 7 lettres, dont un Z. Afficher le prénom uniquement. (115 lignes)
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. (2855 lignes)
23. Afficher la liste des prénoms distincts communs (pas rares) donnés entre 2000 et 2020. (10727 lignes)
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. (5283 lignes)

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 2.3 — 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)