Feuille de travaux pratiques n° 4
Requêtes SQL (agrégation, regroupement et jointures avancées)

Exercice 4.1 — 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

5. 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

6. 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).
7. Affichez la liste des matricules et noms des employés (14 lignes).
8. 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).
9. Liste des postes des employés dont le salaire est supérieur à 1500 € (8 lignes).
10. Matricules et noms des employés dont le salaire est compris entre 1000 et 2000 € (8 lignes).
11. 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).
12. Affichez les matricules et noms des employés dont le poste est `'ingénieur'` ou bien `'commercial'` (6 lignes).
13. Affichez les matricules et noms des employés des départements 20 et 30 (11 lignes).
14. Affichez les matricules et noms des employés qui sont des commerciaux ou bien qui travaillent dans le département numéro 30 (6 lignes).
15. Affichez les employés de poste `'ingénieur'` du département numéro 20 (`'recherche'`) et ceux de poste `'secrétaire'` dans le département numéro 30 (`'vente'`) (3 lignes).
16. Affichez pour tous les employés qui sont des commerciaux, leur matricule, nom et la rémunération totale perçue (salaire + prime) (4 lignes).
17. 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).
18. Liste des employés qui touchent une prime et dont la rémunération (salaire + prime) est supérieure ou égale à 2000 € (3 lignes).
19. 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).
20. 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).
21. Matricules et noms des employés dont le nom commence par un `'B'` (3 lignes).
22. Matricules et noms des employés dont la 2^{e} lettre du nom est un `'e'` (6 lignes).
23. 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).
24. Matricule et nom des employés qui n'ont pas de supérieur (1 lignes).

Options d'affichage

25. Liste des divers postes existants, en n'affichant chaque poste qu'une seule fois. Utilisez le mot-clé `DISTINCT` (5 lignes).

26. Matricule, nom et numéro de département de tous les employés. Renommez les colonnes afin de faciliter la lecture du résultat (14 lignes).
27. 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).

Requêtes avec regroupements et agrégations

37. Affichez pour chaque département, le numéro du département et le salaire maximal des employés de ce département. Utilisez une clause `GROUP BY` pour regrouper les employés par numéro de département (3 lignes).
38. Modifiez la requête précédente pour que soient affichés également pour chaque département, le salaire minimal, le salaire moyen et le nombre d'employés de ce département (3 lignes).
39. Modifiez la requête précédente afin d'afficher également pour chaque département le nom et le lieu du département (3 lignes).
40. Affichez pour chaque employé son matricule et son nom ainsi que le nombre de projets auxquels il participe (6 lignes).
41. Affichez pour chaque couple (numéro de département, code projet) le nombre d'employés de ce département qui participent à ce projet. Spécifiez plusieurs champs de regroupement dans la clause `GROUP BY` (5 lignes).
42. Moyenne des salaires des employés qui ont le même supérieur direct que l'employé de nom `'Berger'` (1520.0).

Requêtes avec condition de sélection sur le regroupement

43. Liste des départements qui comptent strictement plus de 3 employés. On veut afficher le numéro et le nom du département suivis du nombre d'employés. Utilisez une clause `HAVING` de sélection sur le regroupement (2 lignes).
44. Liste des employés qui participent à plusieurs (strictement plus de un) projets. Afficher le matricule et le nom de l'employé ainsi que le nombre de projets auxquels il participe (2 lignes).
45. Liste des départements dont la masse salariale totale (total des salaires des employés du département) dépasse 10000 € (2 lignes).

Sous requête dans la clause `HAVING`

46. **Remarque importante :** on vous demande ici d'écrire une requête qui utilise l'opérateur `ALL`. Mais ni `ALL`, ni `ANY` ne sont définis en SQLite. Donc votre requête ne sera pas acceptée par SQLite. Vous écrirez une autre requête acceptée par SQLite à la question suivante.

    Liste des employés qui participent au plus grand nombre de projets. La sous-requête doit renvoyer le nombre de projets pour chaque employé. Vous devez définir un regroupement des projets par employés (clause `GROUP BY`). La requête principale doit sélectionner un employé si son nombre de projets est supérieur ou égal à toutes les valeurs renvoyées par la sous-requête (opérateur `ALL`). Vous devez ici aussi définir un regroupement des projets par employé dans la clause `GROUP BY` et définir le critère de sélection dans la clause `HAVING` (2 lignes).
47. Récrivez la requête précédente sans l'opérateur `ALL`. Vous pourrez utiliser l'opérateur `NOT EXISTS`. Il ne doit exister aucun autre employé qui participe à un plus grand nombre de projets (2 lignes).
48. Liste des départements dont le salaire moyen est le plus élevé. Il ne doit exister aucun autre département dont la moyenne des salaires est strictement supérieure (10, finance, paris, 3266.667).

Exercice 4.2 — Bibliothèque

On reprend la base de données des bibliothèques, le fichier `biblio.db` est disponible sur Moodle.
Ouvrez cette base de données avec *DB Browser for SQLite*.

Schéma relationnel

```
Abonné(NumAbo, NomAbo, PrénomAbo, AdrAbo, DateAbo)
Livre(ISBN, Titre, Éditeur, Année)
Emprunt(NumEmp, DateEmp, DateRet, NumAbo, ISBN)
Auteur(NumAut, NomAut, PrénomAut)
Écrit(NumAut, ISBN)
```

Requêtes SQL

Écrivez des requêtes dans l'onglet *Exécuter le SQL* pour répondre aux questions suivantes.

Jointures

1. Afficher la table complète des abonnés et de leurs emprunts. Les abonnés qui n'ont effectué aucun emprunt doivent aussi apparaître.

    Vous devez effectuer une jointure entre les tables Emprunt et Abonné. Utilisez l'opérateur `LEFT JOIN`.
2. Afficher la table complète des livres avec leurs emprunts. Les livres qui n'ont jamais été empruntés doivent aussi apparaître.

    Vous devez effectuer une jointure entre les tables `Livre` et `Emprunt`.
3. Afficher la table complète des abonnés ainsi que le titre des livres qu'ils ont empruntés. Les abonnés n'ayant jamais emprunté un livre doivent aussi apparaître.

    Vous devez effectuer deux jointures entre les tables `Abonné`, `Livre` et `Emprunt`.

Manipulation des dates avec SQLite

4. Affichez toutes les informations ainsi que la durée d'emprunt de chaque emprunt. Qu'observez-vous ?
5. En fait, par défaut, SQLite calcule uniquement la différence des années. Mais il existe la fonction SQLite `julianday` pour convertir une date en nombre de jours et faire les calculs qu'on veut.

    Remplacez `dateret - dateemp` par `julianday(dateret) - julianday(dateemp)` et observez le résultat.

    Remarque : cette syntaxe n'est pas le SQL standard, mais propre à SQLite.

    Avec d'autres SGBD, pour avoir ce qu'on s'attend à obtenir avec `dateret - dateemp`, on peut écrire `DATEDIFF(dateret, dateemp)`. Il faut lire la documentation selon le SGBD avec lequel vous travaillez.

    Par exemple, pour SQLite, vous pouvez consulter

— https://sqlite.org/lang_datefunc.html pour les dates,
— https://www.sqlite.org/lang_corefunc.html pour d'autres fonctions de base.

6. Affichez la durée moyenne des emprunts.
7. Affichez les informations des emprunts dont la durée est strictement plus grande que la moyenne.

Requêtes avec regroupements et agrégations

8. Affichez pour chaque numéro d'abonné le nombre d'emprunts qu'il a réalisé. Utilisez une clause `GROUP BY` pour regrouper les emprunts par numéro d'abonné.
9. Modifiez la requête précédente pour que soient affichés également le nom et le prénom de l'abonné ainsi que les dates de son premier et de son dernier emprunts.
10. Durée moyenne des emprunts de chaque abonné. Les abonnés doivent être affichés par ordre décroissant des moyennes des durées des emprunts.
11. Affichez pour chaque abonné et pour chaque livre qu'il a emprunté le nombre d'emprunts de ce livre par cet abonné. Affichez, le numéro et le nom de l'abonné ainsi que le numéro et le titre du livre. Utilisez plusieurs champs dans la clause `GROUP BY`.
12. Affichez pour chaque éditeur le nombre d'emprunts de livres édités par cet éditeur.

Requêtes avec condition de sélection sur le regroupement

13. Liste des abonnés qui ont réalisé strictement plus de 2 emprunts. Affichez le numéro et le nom de l'abonné suivis du nombre d'emprunts. Utilisez une clause `HAVING` de sélection sur le regroupement.
14. Liste des auteurs qui ont participé à l'écriture de plusieurs (strictement plus de un) livres. Afficher les numéro et nom de l'auteur ainsi que le nombre de livres.
15. Liste des livres qui ont été écrits par plusieurs (strictement plus de un) auteurs.
16. Liste des abonnés dont la durée maximale des emprunts est supérieure à 30 jours.

Sous requête dans la clause `HAVING`

17. Liste des abonnés qui ont réalisé le plus grand nombre d'emprunts.
    — La sous-requête doit renvoyer le nombre d'emprunts pour chaque abonné à partir d'un copie `E2` d'`Emprunt`. Vous devez définir un regroupement des emprunts par abonné (clause `GROUP BY`).
    — La requête principale doit sélectionner un abonné s'il n'existe pas d'abonné ayant un nombre d'emprunts strictement supérieur. Pour ce faire, vous devez ici aussi définir un regroupement des emprunts par abonné dans la clause `GROUP BY` et définir le critère de sélection dans la clause `HAVING`.
    — Dans la sous-requête, vous devez rajouter une clause `HAVING` stipulant que le nombre d'emprunts pour `E2` est supérieur au nombre d'emprunts de l'abonné de la requête principale.

    Remarque importante : une autre solution en SQL standard serait d'utiliser l'opérateur `ALL`. Mais ni `ALL`, ni `ANY` ne sont définis en SQLite.
18. Liste des abonnés qui n'ont pas réalisé le plus grand nombre d'emprunts. Il doit exister au moins un abonné ayant réalisé plus d'emprunts.
19. Liste des livres dont la moyenne des durées d'emprunt est la plus élevée. Il ne doit exister aucun autre livre dont la durée moyenne des emprunts est supérieure.