

Bases de Données Relationnelle 2, TD
L'optimisation et les index

Objectif :

- Utilisation d'EXPLAIN pour savoir comment la base de données évalue le fonctionnement d'une requête.
- Création des index qui couvrent le filtrage et le tri que vous souhaitez effectuer sur vos requêtes.
- Évaluation du meilleur ordre pour les colonnes lors de la création d'un index.

Partie 1 : Familiarisation avec EXPLAIN

Lisez la documentation sur EXPLAIN :

<https://dev.mysql.com/doc/refman/8.0/en/using-explain.html>

Prenez votre temps pour lire la documentation sur le format de sortie d'EXPLAIN, particulièrement, les explications des colonnes « i d », « select_type », « table », « possible_keys », « key », « ref », « rows » et « filtered ».

<https://dev.mysql.com/doc/refman/8.0/en/explain-output.html>

Pour ce TD, vous pouvez également utiliser le EXPLAIN visuel de MySQLWorkbench (voir slides du cours).

Partie 2 : Instanciation et familiarisation avec la BD

Téléchargez et exécutez le script « employees.sql » disponible sur coursenligne. Regardez la structure de la table « employees ».

Q1. Formulez une requête SQL pour compter le nombre de lignes dans la table « employees ».

Q2. Quels sont les index de la table « employees ». Vous pouvez utiliser le « Table Inspector » de MySQLWorkbench.

Partie 3 : Utilisation d'EXPLAIN

Q3. Formulez une requête SQL pour trouver tous les employés avec le prénom « Joe ».

Q4. Utilisez EXPLAIN avec la requête précédente.

- Combien de lignes doit parcourir la requête ?
- Qu'est-ce que vous pensez de ce nombre de lignes par rapport au résultat de la requête ?
- Quel index elle utilise ?

Q5. Avec « CREATE INDEX », créez un index pour la recherche d'un employé par prénom. Donnez-lui un nom selon la convention : nomtable_nomcolonne_idx.

- Vérifiez les index de la table employee.

Q6. Exécutez à nouveau EXPLAIN pour trouver tous les employés avec le prénom « Joe ».

- Quelle est la différence la plus importante de ce résultat avec celui obtenu en Q4 ?

Partie 3 : Des index sur plusieurs colonnes

Q7. Nous voulons maintenant rechercher un employé par nom et prénom. Créez un index sur le nom des employés.

Q8. Écrivez une requête pour trouver l'employé du nom « Halloran » et prénom « Aleksander » et utilisez EXPLAIN.

- Qu'est-ce que vous remarquez concernant l'utilisation des index ?

Q9. Créez un index qui regroupe les deux colonnes nom et prénom. Attention : indiquez d'abord la colonne nom, puis prénom lors de la création de l'index.

Q10. Exécutez à nouveau EXPLAIN pour trouver l'employé du nom « Halloran » et prénom « Aleksander ».

- Que constatez-vous par rapport à Q8 ?

Partie 3 : Choix de colonnes pour les index

Le choix des index dépend de la manière dont vous interrogez la table et les colonnes qui font partie des requêtes. Lorsque nous recherchons des employés, nous voulons les trouver par nom et prénom, nous avons donc un index sur les deux. Lorsque l'on crée un index pour plusieurs colonnes, l'ordre des colonnes est important.

Si nous avons un employé appelé Georgi Facello, notre index le référencera sous l'entrée l'entrée *Facello-Georgi* (l'index est nom, prénom). Cela ne fonctionne donc que pour les requêtes dans lesquelles vous recherchez nom et prénom ou uniquement le nom de famille. Une requête qui ne recherche que le prénom ne peut pas utiliser cet index, car l'index est de gauche à droite. Pour cela, vous avez besoin d'un index qui commence par prénom.

Maintenant, si vous interrogez toutes les colonnes de l'index ou la plupart d'entre elles, quel ordre doivent-elles avoir ? Les colonnes avec la **plus grande variété de valeurs** devraient venir en premier pour un filtrage optimal. Pour savoir quelles sont ces colonnes, vous devrez faire quelques requêtes.

Q11. Écrivez une requête pour calculer la moyenne des identifiants différents des employés. Puisque les identifiants sont des valeurs uniques, vous devriez obtenir « 1 » comme réponse à votre requête.

Q12. Écrivez une requête pour calculer la moyenne des prénoms différents des employés.

Q13. Écrivez une requête pour calculer la moyenne des noms différents des employés.

Q14. Quelle colonne est plus adaptée pour la gauche de l'index (puisque'elle offre un filtrage meilleur, une plus grande variété de valeurs) ?

Q15. Avec ALTER TABLE, supprimez tous les index que vous avez créés.