

Exercice 01 : Gestion des Transactions

1. Pourquoi les systèmes vérifient-ils les contraintes d'intégrité uniquement à la fin de la transaction ? Indiquez également la propriété ACID qui doit être assurée pour cela et expliquez pourquoi ?
2. Justifiez l'énoncé suivant : L'exécution simultanée d'une transaction est plus importante lorsque les données doivent être extraites du disque (lent) ou lorsque les transactions sont longues et moins importante lorsque les données sont en mémoire et que les transactions sont très courtes.
3. Expliquez pourquoi la reprise de transaction est nécessaire?
4. Dans un contexte transactionnel, les opérations UNDO et REDO sont nécessaires pour laquelle des quatre propriétés ACID?
5. Citer les deux stratégies d'optimisation de plan d'exécution des requêtes SQL ?
6. Comparez NoSQL et la base de données relationnelle ?
7. Qu'est ce qu'il démontre le théorème CAP?
8. Que signifie MapReduce et Hadoop ? Décrivez le principe de fonctionnement de MapReduce avec un exemple approprié ?

Exercice 02: Datawarehouse

Le système d'information de l'UIK « gestion de la scolarité » accumule depuis des années un grand nombre de données concernant les résultats de délibération. Ces données peuvent être utilisées pour tirer des informations précieuses de pilotage.

1. Proposer un schéma en étoile (ou en *flocon neige*) pour répondre à la requête suivante : **RQ** : Quel est le taux de réussite des étudiant de master en 2019 par série de bac et par sexe du étudiant et par département ?
2. Donnez la requête OLAP qui retourne ce résultat ?

Exercice 03: Applications de Data Mining

Sélectionnez l'un des secteurs suivants et répondez à la question ci-dessous :

Commerce, Secteur Bancaire, Assurance, Médicale, Gouvernement, Éducation

1. Décrivez la nature des sources de données dans ce secteur d'activité
2. Énumérez et discutez les principaux problèmes qui doivent être abordés
3. Décrivez une application de datamining possible dans ce secteur d'activité
4. Décrire un exemple de secteur pour laquelle l'analyse des données volumineuses (Big Data Analytics) est un problème essentiel et des enjeux impliqués.
5. Comment pouvons-nous apporter de la valeur ajoutée ?