Bing

Bonjour, je suis heureux de vous aider avec votre examen sur la normalisation des données. Voici un QCM de 5 questions que j'ai généré à partir des notions clef du cours et des ressources en ligne annexes. Pour chaque question, je vous donne la réponse ainsi qu'un feedback sur les pièges à éviter en conception de base de données.

Q1: Quel est le but principal de la normalisation des données ?
A) Réduire la redondance et les anomalies dans les bases de données relationnelles
B) Ramener les valeurs numériques sur une échelle commune pour faciliter l'analyse
C) Organiser les données selon un format standard pour augmenter leur cohésion
D) Transformer les données en informations exploitables pour la prise de décision

Réponse: A) Réduire la redondance et les anomalies dans les bases de données relationnelles
Feedback: La normalisation des données est une approche de conception de base de données utilisée dans les bases de données relationnelles pour éviter la redondance, c'est-à-dire la duplication inutile des données, et les anomalies, c'est-à-dire les erreurs ou les incohérences qui peuvent survenir lors des opérations d'insertion, de mise à jour ou de suppression des données. La normalisation des données n'a pas pour but de changer l'échelle des valeurs numériques, ce qui relève plutôt du prétraitement des données pour le machine learning. La normalisation des données n'a pas non plus pour but d'organiser les données selon un format standard, ce qui relève plutôt de la normalisation des formats. Enfin, la normalisation des données n'a pas pour but de transformer les données en informations exploitables, ce qui relève plutôt de l'analyse ou de la visualisation des données.

Q2: Quelle est la condition nécessaire pour qu'une relation soit en première forme normale (1FN ou 1NF) ?
A) Tous les attributs doivent être atomiques, c'est-à-dire indivisibles
B) Tous les attributs doivent dépendre entièrement de la clé primaire
C) Tous les attributs doivent dépendre uniquement de la clé primaire
D) Tous les attributs doivent être uniques, c'est-à-dire sans doublons

Réponse: A) Tous les attributs doivent être atomiques, c'est-à-dire indivisibles
Feedback: La première forme normale (1FN ou 1NF) est le niveau le plus élémentaire de normalisation. Elle impose que tous les attributs d'une relation soient atomiques, c'est-à-dire qu'ils ne contiennent pas plusieurs valeurs séparées par des virgules, des tirets ou d'autres symboles. Par exemple, si un attribut adresse contient le numéro, la rue, le code postal et la ville, il n'est pas atomique et doit être décomposé en plusieurs attributs distincts. La condition B) correspond à la deuxième forme normale (2FN ou 2NF), la condition C) correspond à la troisième forme normale (3FN ou 3NF) et la condition D) correspond à la forme normale de Boyce-Codd (FNBC ou BCNF).

Q3: Quelle est la différence entre la troisième forme normale (3FN ou 3NF) et la forme normale de Boyce-Codd (FNBC ou BCNF) ?
A) La 3FN élimine les dépendances transitives, tandis que la FNBC élimine les dépendances partielles
B) La 3FN élimine les dépendances partielles, tandis que la FNBC élimine les dépendances transitives
C) La 3FN élimine les dépendances transitives, tandis que la FNBC élimine toutes les dépendances non triviales
D) La 3FN élimine toutes les dépendances non triviales, tandis que la FNBC élimine les dépendances multivaluées

Réponse: C) La 3FN élimine les dépendances transitives, tandis que la FNBC élimine toutes les dépendances non triviales
Feedback: Une dépendance fonctionnelle est une contrainte qui exprime qu'une valeur d'un attribut détermine une valeur d'un autre attribut. Par exemple, si A -> B signifie que la valeur de A détermine la valeur de B, alors on dit que B dépend fonctionnellement de A. Une dépendance fonctionnelle est dite triviale si elle est toujours vraie, par exemple A -> A ou A,B -> A. Une dépendance fonctionnelle est dite partielle si elle concerne une partie d'une clé candidate, c'est-à-dire un ensemble minimal d'attributs qui identifient de manière unique chaque tuple d'une relation. Par exemple, si A,B est une clé candidate et que A -> C, alors on dit que C dépend partiellement de A. Une dépendance fonctionnelle est dite transitive si elle concerne un attribut qui n'est pas une clé candidate et qui dépend d'un autre attribut qui n'est pas une clé candidate. Par exemple, si A est une clé candidate et que A -> B et B -> C, alors on dit que C dépend transitivement de B. La troisième forme normale (3FN ou 3NF) élimine les dépendances transitives, c'est-à-dire qu'elle impose que tous les attributs qui ne sont pas des clés candidates dépendent uniquement de la clé primaire. La forme normale de Boyce-Codd (FNBC ou BCNF) élimine toutes les dépendances non triviales, c'est-à-dire qu'elle impose que tous les attributs qui déterminent d'autres attributs soient des clés candidates. La condition A) est fausse car elle inverse les rôles de la 3FN et de la FNBC. La condition B) est fausse car elle confond la 2FN et la 3FN. La condition D) est fausse car elle confond la FNBC et la quatrième forme normale (4FN – NF4).

Q4: Quel est le principal avantage de la quatrième forme normale (4FN – NF4) ?
A) Elle permet d'éviter les anomalies d'insertion, de mise à jour et de suppression
B) Elle permet d'éviter les redondances dues aux relations many-to-many
C) Elle permet d'éviter les redondances dues aux dépendances multivaluées
D) Elle permet d'éviter les redondances dues aux dépendances cycliques

Réponse: C) Elle permet d'éviter les redondances dues aux dépendances multivaluées
Feedback: Une dépendance multivaluée est une contrainte qui exprime qu'une valeur d'un attribut implique plusieurs valeurs possibles pour un autre attribut, indépendamment des autres attributs. Par exemple, si un étudiant peut avoir plusieurs cours et plusieurs hobbies, alors on dit qu'il existe une dépendance multivaluée entre l'attribut étudiant et les attributs cours et hobbies. La quatrième forme normale (4FN – NF4) élimine les redondances dues aux dépendances multivaluées, c'est-à-dire qu'elle impose que pour chaque dépendance multivaluée X ->> Y, X soit une clé candidate. La condition A) est fausse car elle s'applique à toutes les formes normales. La condition B) est fausse car elle confond les dépendances multivaluées avec les associations many-to-many, qui sont des concepts différents. La condition D) est fausse car elle introduit une notion de dépendance cyclique qui n'existe pas dans le modèle relationnel.

Q5: Quel est le principal inconvénient de la normalisation des données ?
A) Elle augmente la complexité des requêtes SQL
B) Elle diminue la performance des algorithmes de machine learning
C) Elle modifie la signification des données originales
D) Elle nécessite un stockage supplémentaire pour les données

Réponse: A) Elle augmente la complexité des requêtes SQL
Feedback: La normalisation des données a pour principal inconvénient d'augmenter la complexité des requêtes SQL, car elle implique de créer plus de tables et de faire plus de jointures pour accéder aux données souhaitées. La normalisation des données n'a pas pour inconvénient de diminuer la performance des algorithmes de machine learning, au contraire elle peut améliorer la qualité des données en évitant les redondances et les anomalies. La normalisation des données n'a pas pour inconvénient de modifier la signification des données originales, au contraire elle respecte le principe de conservation des informations, c'est-à