

Bases de Données avancées Projet BDA SQL3-Oracle et NoSQL (MongoDB)

Indications

Le projet se fera **en binôme (ou monôme)**. Un rapport devra être rédigé comportant les points essentiels avec la structure suivante :

- Page de garde
- Introduction
- Réponses aux questions posées dans le sujet (le modèle de classes, le schéma relationnel, la définition des types, les associations, ...)
- Requêtes SQL correspondant aux questions posées
- Explication des parties non triviales des requêtes complexes,
- Captures (résultats de l'exécution des requêtes) et commentaires.
- Modélisation orientée-documents et illustration par des exemples
- Explication des requêtes complexes,
- Analyse par rapport à la conception
- Conclusion

-
- Le délai pour envoyer le rapport ainsi que les différents Scripts (pour les deux parties) est fixé au **samedi 10 mai 2025**
 - Le travail devra être envoyé à l'adresse MasterWorks213@gmail.com
Objet du mail: **Projet BDA+ Nom et prénoms des Etudiants**
 - Le projet comporte deux parties (La partie I porte sur le relationnel-objet et la partie II sur le NoSQL orienté documents avec MongoDB)
 - Les affichages à l'écran et les messages doivent être clairs et explicites
 - Il est demandé d'utiliser un menu pour le choix des requêtes à exécuter.
-

Dans ce projet, on s'intéresse à la gestion d'un réseau de transport urbain dans une ville intelligente (smart city). Dans la ville, sont offerts plusieurs moyens de transport (métro, tramway, bus, train), par ces moyens de transport plusieurs lignes sont desservies.

Un *moyen de transport* est décrit par une abréviation (MET, TRA, BUS, TRN), possède une heure d'ouverture, une heure de fermeture et un nombre moyen de voyageurs dans la journée.

Une *ligne* de transport comporte deux stations spécifiques : une station de départ et une station d'arrivée, elle est composée d'un ensemble de tronçons. Une ligne est identifiée par un code unique et est rattachée à un *seul* moyen de transport.

Une *station* est identifiée par un code unique, et possède un nom ainsi que des coordonnées spatiales (longitude, latitude) sur la carte du réseau. Une station peut être principale ou secondaire ; Une station principale est une station de correspondance où on peut avoir plusieurs moyens de transport (i.e les voyageurs peuvent descendre dans une station principale pour changer le moyen de transport et la destination éventuellement)

Un *tronçon* est un morceau (une partie) de la ligne qui relie deux stations consécutives (intermédiaires), il possède un numéro unique, une station de début, une station de fin et une longueur (kilométrage). Un tronçon peut appartenir à deux lignes différentes.

Pour les voyages, on prévoit pour chaque ligne, un ensemble de navettes selon le moyen de transport (bus, tram, métro, train) qui desservent la ligne. La *navette* possède un numéro unique, une marque et une année de mise en circulation.

Une navette effectue plusieurs voyages dans le temps. Un *voyage* est décrit par un numéro, une durée, une date, une heure de début, un sens (aller/retour), un nombre de voyageurs (selon le nombre de tickets vendus) et une observation (RAS, Panne, Retard, ...etc).

Par ailleurs, une application de géolocalisation permet à tout moment d'estimer la durée pour aller d'une station A à une station B (avec le moyen de transport adéquat). On simulera une méthode CalculerDurée qui donnera à tout moment la durée de déplacement sur un tronçon selon le moyen de transport utilisé.

Travail demandé

Partie I : Relationnel-Objet

A- Modélisation orientée objet

1. Etablir un diagramme de classes UML relatif à cette étude de cas (préciser les attributs et méthodes)
2. Transformez ce diagramme en un schéma relationnel (préciser les clés primaires et les clés étrangères)

B- Création des TableSpaces et utilisateur

3. Créer deux TableSpaces *SQL3_TBS* et *SQL3_TempTBS*
4. Créer un utilisateur SQL3 en lui attribuant les deux tableSpaces créées précédemment
5. Donner tous les privilèges à cet utilisateur.

C- Langage de définition de données

6. En se basant sur le diagramme de classes établi, définir tous les **types abstraits** nécessaires. Définir toutes les **associations** qui existent.
7. Définir les **méthodes** permettant de :
 - Calculer pour chaque navette, le nombre total de voyages effectués.
 - Retourner pour chaque ligne, la liste des navettes qui la desservent.
 - Calculer pour une ligne (de numéro donné), le nombre de voyages effectués durant une période (Exemple : du 01-01-2025 au 15-02-2025).
 - Changer le nom de la station « BEZ » par « Univ » dans toutes les lignes/tronçons comportant cette station.
 - Calculer pour un moyen de transport donné (Exemple Métro), le nombre de voyages effectués à une date donnée (Exemple le 28-02-2025) et le nombre de voyageurs total.
8. Définir les tables nécessaires à la base de données.

D- Création des instances dans les tables

9. Remplir (via des scripts) toutes les tables par des instances (4 moyens de transport, plusieurs lignes pour chaque moyen de transport, plusieurs stations sur les lignes et plusieurs voyages par jour, sur une période de **deux mois** au minimum du 01-01-2025 au 01-03-2025), en respectant les **contraintes** d'intégrité.

Les valeurs des attributs doivent être sensées et significatives

Le numéro de la ligne devra avoir la forme suivante B001, B002, ... M001, M002, ... TR001, TN001, ...

Les initiales B, M, TR et TN représentent respectivement le type de transport : Bus, Métro, Tramway et Train. Une ligne possède une station de départ, une station d'arrivée et un ensemble de stations intermédiaires. Les stations sont numérotées par S001, S002, ... Les navettes sont numérotées par N001, N002, N003, ... Les numéros de voyages sont V0001, V0002, ...

E- Langage d'interrogation de données

10. Lister tous les voyages (num, date, moyen de transport, navette) ayant enregistré un quelconque problème (panne, retard, accident, ...)
11. Lister toutes les lignes (numéro, début et fin) comportant une station principale
12. Quelles sont les navettes (numéro, type de transport, année de mise en service) ayant effectué le maximum de voyages durant le mois de janvier 2025 ? Préciser le nombre de voyages.
13. Quelles sont les stations offrant au moins 2 moyens de transport ? (préciser la station et les moyens de transport offerts)

Partie II : NoSQL – Modèle orienté « documents »

A- Modélisation orientée document

On suppose que la plupart des requêtes sur la base vont porter sur les voyages (voir exemples de requêtes plus bas).

- Proposer une modélisation orientée document de la base de données décrite dans la partie I, dans ce cas.
- Illustrez votre modélisation sur un exemple (ou plus) de la BD que vous avez générée
- Justifiez votre choix de conception
- Quelles sont les inconvénients de votre conception

B- **Remplir la base de données** (via un script, ajouter d'autres données afin d'augmenter le volume de la base, augmenter le nombre de voyages, la période, ... etc.)

C- Répondre aux requêtes suivantes :

- Afficher tous les voyages effectués en date du 01-01-2025 (préciser les détails de chaque voyage)
- Dans une collection **BON-Voyage**, récupérer tous les voyages (numéro, numLigne, date, heure, sens) n'ayant enregistré aucun problème, préciser le moyen de transport, le numéro de la navette associés au voyage.
- Récupérer dans une nouvelle collection **Ligne-Voyages**, les numéros de lignes et le nombre total de voyages effectués (par ligne). La collection devra être ordonnée par ordre décroissant du nombre de voyages. Afficher le contenu de la collection.
- Augmenter de 100, le nombre de voyageurs sur tous les voyages effectués par métro avant la date du 15 janvier 2025.
- Reprendre la 3^{ème} requête à l'aide du paradigme **Map-Reduce**
- Avec votre conception, peut-on répondre aux requêtes suivantes (justifiez vos réponses) :
 - a- *Afficher les navettes ayant effectué un nombre maximum de voyages, en précisant le moyen de transport associé.*
 - b- *Afficher les moyens de transport dont le nombre de voyageurs dépasse toujours un seuil S donné (par exemple 10000) par jour.*

D- Analyse

Donnez votre analyse de ces requêtes par rapport à la conception que vous avez proposée.
Peut-on avoir une meilleure conception ?