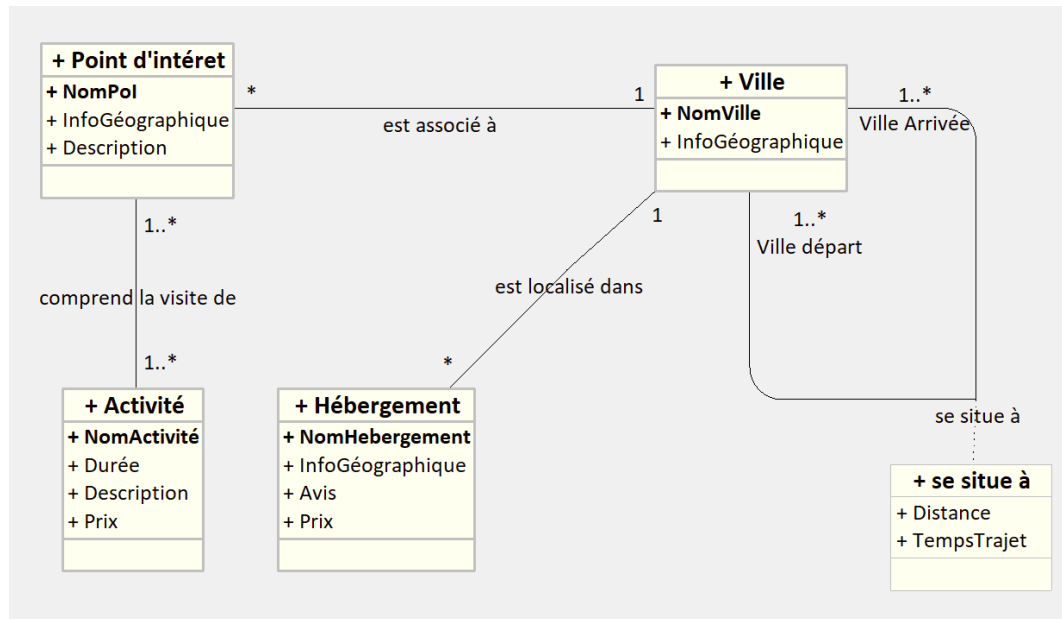


Projet de NoSQL M2 MIAGE-IF 2024-2025

Enoncé

On souhaite modéliser une base de données polyglotte permettant de gérer des informations sur des points d'intérêts afin d'aider à planifier un voyage¹.

La figure ci-dessous présente un diagramme de classes.



Un voyage comprend la visite de plusieurs points d'intérêt (ex. le biodôme à Montréal), pouvant faire l'objet d'activités (ex. la visite du biodôme), chaque point d'intérêt étant localisé dans une ville disposant d'hébergements. A chaque ville, hébergement et point d'intérêt sont associées des informations géographiques. Les informations géographiques des villes permettent de calculer des distances (en km) entre elles. Toutes les distances entre toutes les villes ne sont pas forcément calculées. On supposera que les distances entre un point d'intérêt ou un hébergement avec sa ville de rattachement sont négligeables. La description des points d'intérêts et la description des activités ne sont pas uniformes et peuvent varier en fonction du type de point d'intérêts (ex. musée ou espace naturel) ou des activités (ex. certaines activités peuvent avoir un lieu de départ ou des informations sur les saisons et d'autres non).

La base de données polyglotte doit permettre de répondre aux requêtes suivantes :

1. Quels sont les villes situées à moins de 10km d'une ville donnée ?
2. Quelles sont les activités associées à un point d'intérêt donné ?
3. Quels sont les hébergements d'une ville ?
4. Quelle est le temps de trajet et la distance entre 2 villes données ?
5. Quelles activités se font entre avril et juin ?
6. Etant données une ville de départ et une ville d'arrivée, quels sont les différentes villes possibles à visiter entre les 2 ?

Il est possible d'ajouter d'autres classes ou attributs au modèle ci-dessus.

¹ Un exemple de site est disponible sur : <https://www.authentikcanada.com/fr-fr>

Travail à réaliser

Pour ce projet, vous devez, pour la base de données décrite précédemment, **définir les instructions de création, suppression et mises à jour des données et les requêtes d'interrogation associées. Votre base de données doit être polyglotte et utiliser au minimum trois modèles différents** parmi les modèles vus en cours, c'est-à-dire :

1. Un modèle clé-valeur, sous *Redis* par exemple.
2. Un modèle colonne, sous *Cassandra* par exemple.
3. Un modèle relationnel, sous *PostgreSQL* avec éventuellement certains attributs pouvant être des couples clé-valeur ou des données en *JSON* par exemple.
4. Un modèle document, sous *MongoDB* par exemple.
5. Un modèle graphe, sous *Neo4j* par exemple.

Le choix du modèle doit être fait afin de répondre au mieux aux requêtes et à la description des données. Vous pouvez utiliser des données réelles² ou générer des données fictives via un programme externe, ou faire générer des données (respectant votre modèle) via un outil d'IA (type *ChatGPT*). Vous pouvez utiliser les moteurs vus en cours ou tout autre moteur de votre choix.

Rendu du TP

Pour ce projet, vous devrez notamment :

1. Rédiger un **rapport³ (10-15 pages maximum)** contenant :
 - Une description des différents modèles de données avec les commandes de création/suppression et mises à jour des données.
 - La description de vos choix de modélisation en indiquant les avantages et les limites de vos choix.
 - L'expression de vos requêtes d'interrogation et le résultat obtenu sur des exemples.
 - Une conclusion indiquant les difficultés rencontrées, la répartition du travail entre les membres du binôme et le temps passé sur ce projet.
2. **Réaliser une présentation orale de 5-7 mn avec des transparents et éventuellement une démonstration**, rappelant certains points évoqués dans le rapport. Chaque membre du binôme devra intervenir de manière équitable lors de cette présentation.

Modalités

Le projet est **seul.e ou en binôme⁴**. Vos fichiers (**dont le nom indiquera les noms de familles des membres du binôme**) devront être **déposé sur moodle avant le 16/06/25 12h**. Ces fichiers devront contenir :

- La version électronique du rapport (**1 fichier pdf**)
- La version électronique de vos transparents (**1 fichier pdf**).
- Un **fichier zip contenant** : Les scripts de création de vos modèles et les requêtes d'interrogation (en vue de les rejouer sur les moteurs utilisés).

La **présentation orale aura lieu pendant le cours du 20/06/2025** matin en séance de cours (selon un planning, qui vous sera communiqué).

² Exemples : <https://www.data.gouv.fr/fr/datasets/points-dinteret-touristique/>

³ Rappel : tout rapport doit comporter une introduction et une conclusion.

⁴ Aucun groupe de 3 n'est autorisé.