

## Projet Films - Cinéma

On s'intéresse au développement d'une base de données MongoDB et d'un ensemble de fonctionnalités pour permettre de réaliser de l'analyse décisionnelle concernant les films de cinéma.

L'objectif du projet est de mettre en œuvre sur un exemple personnel les éléments vus en CM/TD et TP notamment la conception et la mise en œuvre d'une base de données en MongoDB et son requêtage, l'usage du langage python et du driver pyMongo, pour réaliser des fonctionnalités plus avancées avec un accès à la base de données MongoDB.

Les informations à gérer dans la base de données MongoDB, concernent les films qui ont été diffusés dans des salles de cinéma. Ces informations vont être utilisées à des fins d'analyse. Par exemple, pour un film, on peut vouloir connaître sa description, les salles où il a été diffusé, la durée de diffusion, le nombre d'entrées, etc. Ces informations permettront d'avoir des statistiques générales pour l'ensemble des cinémas d'une ville, mais aussi de façon plus personnalisée sur un film, ou un cinéma, etc.

En complément de ces informations, les commentaires faits sur les réseaux sociaux sur les films qui ont été diffusés, sont collectés et également stockés dans la base de données. Il peut s'agir des commentaires textuels et/ou de notes d'appréciation pour le film.

Les informations sur les films sont les informations usuelles (titre, réalisateur, durée, etc.) ainsi que sa catégorie pour permettre des recherches par catégorie. Les informations issues des réseaux sociaux pourront venir enrichir l'analyse pour un film donné, ou une catégorie de film, etc.

Le projet devra contenir les éléments suivants :

### **1<sup>ère</sup> partie : Conception et mise en œuvre de la base de données MongoDB**

L'objectif de cette partie est de choisir une modélisation cohérente des informations en adéquation avec les fonctionnalités retenues. Cette partie devra contenir un cahier des charges des fonctionnalités qui seront proposées et une explication des informations modélisées dans la base de données « en clair » et sous forme d'un diagramme UML, et des exemples de document « json » insérés dans la base.

On précisera clairement la ou les collections créées, en les justifiant et en positionnant également ces choix par rapport aux notions de schéma polymorphique, imbrication ou références vues en CM et TD.

La justification des choix de modélisation pourra s'appuyer sur les fonctionnalités attendues décrites dans le cahier des charges de l'application, qui pourront être présentées de façon concise dans cette partie et de façon détaillée dans les parties 2 et 3.

### **2<sup>ème</sup> partie : Requêtage de la base de données MongoDB**

L'objectif de cette partie est de proposer un large panel de requêtes pouvant être nécessaires pour avoir des informations sur les films et leur diffusion, et des fonctionnalités analytiques jugées pertinentes.

Il s'agit de proposer un ensemble diversifié de requêtes avec les différents opérateurs vus en CM/TD et TP, avec des instructions "find", des instructions "aggregate", et au minimum une requête avec le mécanisme Map-Reduce, en expliquant précisément votre requête.

Des requêtes de mise à jour de la base seront également données. On pourra supposer ici que des données « erronées » ont été stockées dans la base de données, et que l'on souhaite corriger ou compléter la base (y compris par exemple en ajoutant des commentaires sur un film).

### **3<sup>ème</sup> partie : Python et pyMongo**

L'objectif de cette partie est de proposer des requêtes complémentaires plus complexes que celles décrites dans la partie 2 sur les jeux mais aussi sur les commentaires et avis des utilisateurs à des fins d'analyse décisionnelle pour les gestionnaires des cinémas.

Il s'agit d'expliquer l'usage du driver pyMongo pour accéder à la base de données MongoDB et de proposer des scripts python plus avancés pour réaliser des fonctionnalités qui pourraient être utiles soit pour l'utilisateur, soit pour les gestionnaires.

**Dossier à rendre pour vendredi 3 mai sur la plateforme plubel avec tous les scripts pour la base MongoDB (insertions, mises à jour, et requêtage) et les scripts python (.py) qui devront être directement exécutables sur les machines des salles de TP. Des exemples de résultats des requêtes et des scripts pourront être intégrés dans le dossier qui comportera les 4 parties demandées dans un seul document.**

**Le projet peut être réalisé seul ou en binôme.**