

Projet
« SpaceCompanies »
(Travail personnel)

L'objectif du projet est d'étudier et d'analyser une collection de documents concernant des entreprises en lien avec le domaine spatial.

Il s'agit dans un premier temps de comprendre les documents décrits dans cette collection nommée « SpaceCompanies », ensuite de fournir un ensemble de requêtes permettant son analyse, et à terme la création d'un tableau de bord pour présenter les résultats (non demandé dans ce projet).

Le projet devra contenir les éléments suivants :

1^{ère} partie : Compréhension et rétro-conception de la base de données

Cette partie devra contenir une explication en clair (en français), des informations modélisées dans la base de données, un diagramme UML, et des exemples de documents json pour illustrer les explications. La modélisation retenue des documents pourra être discutée.

2^{ème} 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 la compréhension des informations contenues dans la collection.

Il s'agit ici de présenter un panel 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.

(de l'ordre de 10 requêtes sans aggregate, 10 avec aggregate, et 1 map-reduce ou plus)

3^{ème} 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 la collection.

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 (de l'ordre de 4 scripts).

Dossier nommé « SGD_Nom_Prenom.zip » à rendre pour le vendredi 17 mai directement par mail avec un lien de téléchargement à nadine.cullot@u-bourgogne.fr ou richard.genestier@u-bourgogne.fr selon votre groupe de TP, ou en téléchargement par teams si vous préférez, avec tous les scripts pour la base MongoDB (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.