

LE PROJET

Découverte et exploitation de données OpenData de la ville de Lyon (Velo'V)

Module de référence : Bases de Données – MySQL Administration / NoSQL

Durée estimée de la mission : 15 heures

1. Mission

La ville de Lyon collecte régulièrement des données concernant l'état des stations Velo'V dont une partie est issue de ses partenaires industriels. Ces données sont ensuite mises à disposition sur la plateforme OpenData de la ville (<https://data.grandlyon.com>).

Afin de plébisciter les modes de transports doux, la métropole vous mandate pour mettre en place une plateforme permettant de consulter/rechercher/visualiser les données en effectuant des croisements entre plusieurs sources (disponibilité temps-réel des stations Velo'V, quartiers de la ville de Lyon, points d'intérêt de la ville de Lyon).

2. Enjeux Objectifs

L'objectif de cette mission est de mettre en œuvre sur un cas concret les compétences acquises en Bases de Données Relationnelles et NoSQL.

Vous partirez des sources de données brutes que vous structurerez et intégrerez à une Base de Données MySQL (qui servira de référentiel de données) et une Base de Données MongoDB (qui servira aux recherches sur votre plateforme).

Enfin votre plateforme Web (la technologie étant à votre discrétion, elle n'est d'ailleurs pas obligatoirement Web) viendra se connecter aux Bases de Données pour répondre aux recherches des utilisateurs, visualiser l'état des stations aux alentours, suggérer la station la plus adaptée, ...

(Note : votre plateforme peut éventuellement se limiter à la connexion à la Base de Données MongoDB. La Base de Données MySQL étant un référentiel, on peut considérer qu'elle est

utilisée par d'autres applications que votre plateforme. Vous devez néanmoins modéliser les données dans la Base de Données MySQL et l'alimenter en parallèle de MongoDB)

3. Règles du jeu

Les grandes lignes du projet sont les suivantes :

- Collecte des données de la plateforme OpenData de la métropole de Lyon
 - Format geo-json
 - Système de projection par défaut RGF93(EPSG:4171)
 - Modifier le lien de téléchargement pour télécharger toutes les données (maxfeatures=30 modifié en maxfeatures=10000)
- Nettoyage et transformation des données téléchargées
 - Pour MongoDB : enlever le début avant le "[" (inclus) et la fin après le "]" (inclus), ainsi que les virgules à la fin de chaque ligne.
 - Pour MySQL : déterminer les schémas de données à mettre en œuvre pour stocker les documents JSON téléchargés dans une Base de Données Relationnelle correctement structurée.
- Automatiser l'alimentation en données par la mise en œuvre de scripts (shell, python ou autre) qui téléchargent, nettoient et insèrent les données dans les 2 Bases de Données MySQL et MongoDB en parallèle. Vous devez permettre l'alimentation indépendante de MySQL et de MongoDB.
- Sécuriser la Base de Données MongoDB.
- Mettre en place la plateforme Web (<https://docs.mongodb.com/ecosystem/drivers/>) :
 - Interface Web permettant d'interagir avec la Base de Données
 - Recherche de station (les plus proches d'un point d'intérêt, ...) avec critères et tris divers
 - Visualisation de l'état de stations (liste / carte / ...)
 - Statistiques sur les stations (valeurs numériques / graphiques / ...)
 - ...

Vous devez obligatoirement travailler avec Git durant l'intégralité du projet.

Cette mission est réalisée par groupe de 3 à 4 personnes.

4. Livrables

A rendre directement via le projet Git ainsi qu'une archive téléchargeable :

- L'application fonctionnelle (le code source), un document d'installation (readme.md) ainsi que vos scripts.
- Une documentation technique expliquant les différents choix technologiques et notamment les modèles de données est également attendue.
- Les éléments supports (powerpoint/pdf) de votre présentation et démonstration orale.

5. Évaluation

L'évaluation de la mission est composée d'un rendu écrit et d'une présentation orale du projet.

Les documents rendus seront évalués selon les critères suivants :

Structure et forme du document : <ul style="list-style-type: none">Pertinence des justifications des choix technologiques (schémas de données, ...)Scripts d'automatisation de l'alimentation en donnéesRecherches mises en oeuvres	/20
Bonus : <ul style="list-style-type: none">Originalité, inventivitéMise en place d'une Base de Données MongoDB hautement disponible en plus de la simple sécurisation	/4
Malus : <ul style="list-style-type: none">Qualité du rendu rédigé (documentation d'installation, technique)	/3
NOTE	/20

L'oral sera évalué selon les critères suivants :

Support de présentation : <ul style="list-style-type: none">Qualité du support (pdf/pptx), structure, plan, illustration, gestion du temps	/5
Aisance orale : <ul style="list-style-type: none">Expression orale, présentation et attitude, mise en scène, originalité, répartition des temps de parole	/5
Transmission du message : <ul style="list-style-type: none">Organisation du projet, capacité à vulgariser la technique, pertinence des technos, démonstration, schémas	/5
Questions-réponses : <ul style="list-style-type: none">Pertinence, apport des réponses, gestion du stress	/5
NOTE	/20

6. Ressources en lignes pour aller plus loin !

Documentation et exemples :

- <https://docs.mongodb.com/ecosystem/drivers/>
- <https://docs.mongodb.com/manual/geospatial-queries/>

Données de la métropole de Lyon :

- Disponibilité en temps-réel des stations Velo'V de la métropole de Lyon :
<https://data.grandlyon.com/equipements/stations-vflov-de-la-mftropole-de-lyon-disponibilitf-temps-rfel/>

- Quartiers des communes de la métropole de Lyon :
<https://data.grandlyon.com/citoyennete/quartiers-des-communes-de-la-mftropole-de-lyon/>
- Points d'intérêt touristiques de la métropole de Lyon :
<https://data.grandlyon.com/culture/points-dintfrft-touristiques-de-la-mftropole-de-lyon/>

Vous pouvez compléter votre plateforme avec toute autre source de données que vous jugeriez intéressante pour le projet.

7. Ouverture d'esprit

Source d'inspiration

Vous pouvez vous inspirer de :

- <https://docs.mongodb.com/manual/tutorial/geospatial-tutorial/>
- <https://velov.grandlyon.com/mapping>
- <https://www.tripadvisor.fr>

Challenge personnel

La Base de Données MongoDB devra être montée en cluster d'instances MongoDB afin de fournir haute disponibilité et permettre la parallélisation des requêtes d'accès aux données. Vous prendrez soin de sécuriser cette Base de Donnée.

Trucs et astuces :

Un cluster est composé d'au moins 3 instances.

Documentation de MongoDB : <https://docs.mongodb.com/manual/>