

# Chargement des données

Denis Beauget et Hayaat Hebiret

February 2021

## 1 MongoDB et Cloud Atlas

Pour le fichier JSON représentant la 3ème classe des passagers du Titanic nous avons choisis le SGBD NoSQL MongoDB.

### Pourquoi ?

MongoDB est simple d'utilisation et est tout particulièrement adapté au format JSON (via son système de collections plutôt que les tables traditionnelles).

Il vient avec plusieurs outils pratique d'utilisation comme MongoDBCompass qui permet de visualiser rapidement les données ou de se connecter à d'autre BDD rapidement.

### Cloud Atlas

Atlas est une solution de cloud proposé par MongoDB pour héberger ses databases, gratuit jusqu'à une certaine limite il permet un accès distant sans ouvrir ses ports ou de créer un serveur dédié.

### Pourquoi ?

Même si les données sont fixés en cas de changement mineur ou d'aménagement des données il est nécessaire que chaque membre du groupe et de l'équipe dans son ensemble puisse avoir accès aux données à jours de manière simple et rapide, sans nécessairement prendre contact avec les autres membres.

Atlas permet donc d'héberger la 3ème classe et ses futurs modifications en plus de nos versions locales (c'est aussi un moyen d'avoir un "backup").

## 2 SQLite3 et SQLite JDBC

Le fichier Titanic.db représente la deuxième et troisième classe des passagers du Titanic. Nous avons choisit le SGBD SQL SQLite.

### **Pourquoi ?**

SQLite3 est compatible avec Graal et il existe une bibliothèque intéressante sous java : SQLite JDBC.

Un script java à été mis en place afin de faciliter l'insertion des fichier .csv dans une base de donnée. Cela permet notamment de ne pas avoir à installer et utiliser SQLite Studio tools ou d'autres applications tierce et d'avoir un plus grand contrôle sur les données insérées.

De plus, durant sa génération, un fichier csv contenant les lignes csv qui n'ont pas été insérées est généré afin de s'assurer que les données dans la base de données sont réglementaires et d'être signalé des soucis dans les données des fichiers.