

# Travaux pratiques : Cassandra

Jonathan Lejeune



## Objectifs

Ce sujet de travaux pratiques vous permet de vous initier au requêtage de données via CQL et à l'utilisation de Cassandra à travers le driver Spark.

## Prérequis

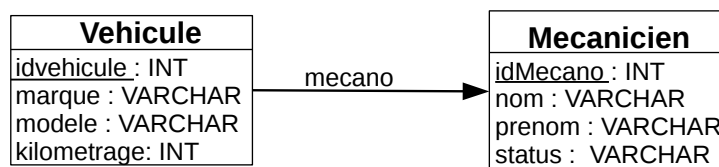
Vous devez avoir fait la première partie du TP préparatoire de Cassandra

## Consignes générales

Démarrez un serveur Cassandra sur votre machine locale et ouvrez un client `cqlsh`. La commande `help` vous permettra d'afficher la documentation et la grammaire des instructions CQL sur votre navigateur.

## Exercice 1 – Réparation automobile

Nous considérons un garage automobile. Le garage reçoit des voitures à réparer et les affecte à un mécanicien. Nous suivrons le schéma relationnel suivant :



### Question 1

Créez un keyspace `garage` avec la stratégie de réplication `SimpleStrategy` et un facteur de réplication de 1 (faire `help CREATE_KEYSPACE` si besoin). Prenez ce keyspace comme keyspace courant avec la commande `USE` ;

### Question 2

Créez les tables `Vehicule` et `Mecanicien` (faire `help CREATE_TABLE` si besoin).

### Question 3

Copiez-collez dans le shell CQL, le fichier *insertion\_cassandra.txt* fourni dans les ressources pour remplir la base.

### Question 4

Listez tous les véhicules (faire `help SELECT` si besoin)

### Question 5

Listez toutes les marques de véhicules.

### Question 6

Donnez le nom du mécanicien numéro 3.

### Question 7

Donnez les marques et les modèles de véhicule du mécanicien numéro 2. Pourquoi ceci affiche une erreur ? Modifiez votre requête afin d'afficher le résultat attendu.

### Question 8

Créez un index sur la colonne `mecano` de la table `Vehicule` (faire `help CREATE_INDEX;`) et refaite la requête de la question précédente (sans la clause `ALLOW FILTERING`). Que remarquez-vous ? Pourquoi ?

### Question 9

Donnez les véhicules avec un identifiant inférieur à 5. Que se passe-t-il si au lieu d'utiliser la clause `ALLOW FILTERING`, on comparait les id non pas par leur valeur mais par leur token ( voir fonction `TOKEN`) ?

## Exercice 2 – Spark et Cassandra

### Question 1

Écrire un programme spark qui copie la table `mecanicien` dans une table `mecanicien_cpy`.

### Question 2

Écrire un programme spark qui permet de fusionner les deux tables `mecanicien` et `vehicule` en une table `reparation`. On considérera qu'une réparation est identifiée par la clé primaire du véhicule concerné.

### Question 3

Écrire un programme Spark qui modifie la table `mecanicien` en ajoutant une colonne `vehicules` qui liste l'ensemble des véhicules affectés au mécanicien.