

## Atelier 3 : Hive

### Exercice 1 :

1. Créer une table `livres` dont le schéma est le suivant : (`id INT`, `title STRING`, `publishDate STRING`)
2. importer les données à partir du fichier `/atelier3/books/books.txt` (stocker localement) dans la table `livres`;
3. Afficher les enregistrements dont l'id est 2.
4. Créer une nouvelle table `ventes` dans l'emplacement HDFS suivant `‘/user/cloudera/atelier3’`;  
  
La table `ventes` contient 3 colonnes : (`id INT`, `buyer STRING`, `purchaseDate STRING`)
5. Afficher 5 lignes de la table `ventes`.
6. Créer la table `vente_livres` avec le schéma suivant (`id INT`, `title STRING`, `buyer STRING`, `purchaseDate STRING`)
7. Remplir la table `vente_livres` en utilisant une jointure selon la colonne `id` entre les tables : `livres` et `ventes`.
8. Afficher 10 enregistrements de la table `vente_livres`
9. Supprimer les tables `livres`, `ventes` et `vente_livres`

### Exercice 2 :

Dans cet exercice nous allons créer une table partitionnée `media`.

Le schéma de la table `media` est le suivant : (`id int`, `title string`, `releaseDate string`, `type string`)

La colonne `type` serait la colonne de partition. Charger les données à partir de 3 fichiers locaux dans trois partitions :

Partition `book` => `/atelier3/media/books.txt`

Partition `cd` => `/atelier3/media/cds.txt`

Partition `dvd` => `/atelier3/media/dvds.txt`

1. Afficher les enregistrements de type `cd`
2. Afficher les enregistrements de type `dvd`
3. Afficher les enregistrements de type `book`
4. Afficher toutes les partitions de la table `media`
5. Ouvrir un nouveau terminal et afficher la structure des répertoires de partition sous HDFS
6. Supprimer la table `media`.