

Correction Atelier 3: HIVE

Exercice 1 :

1. CREATE TABLE livres (id INT, title STRING, publishDate STRING)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE;

LOAD DATA LOCAL INPATH '/home/cloudera/Atelier3/books/books.txt'
OVERWRITE INTO TABLE livres;
2. select * from livres where id=2;
3. CREATE EXTERNAL TABLE ventes (id INT, buyer STRING, purchaseDate STRING)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY '\t'
STORED AS TEXTFILE
LOCATION '/user/biadmin/atelier3';

LOAD DATA LOCAL INPATH
'/home/cloudera/Atelier3/books/book-purchases.txt'
OVERWRITE INTO TABLE ventes;
4. select * from ventes limit 5;
5. CREATE TABLE livres_ventes (id INT, title STRING, buyer STRING, purchaseDate STRING)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE;
6. INSERT OVERWRITE TABLE livres_ventes
SELECT b.id, b.title, p.buyer, p.purchaseDate
FROM livres b JOIN ventes p ON (b.id=p.id);
7. select * from livres_ventes limit 10;
8. drop table livres_ventes;
drop table ventes;
drop table livres;

Exercice 2:

```
CREATE TABLE media (id INT, title STRING, releaseDate STRING)
PARTITIONED BY (type STRING)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','
STORED AS TEXTFILE;
```

```
LOAD DATA LOCAL INPATH '/home/cloudera/Atelier3/media/books.txt'
OVERWRITE INTO TABLE media PARTITION (type='book');
```

```
LOAD DATA LOCAL INPATH '/home/cloudera/Atelier3/media/cds.txt'
OVERWRITE INTO TABLE media PARTITION (type='cd');
```

```
LOAD DATA LOCAL INPATH '/home/cloudera/Atelier3/media/dvds.txt'
OVERWRITE INTO TABLE media PARTITION (type='dvd');
```

1. select * from media where type = "cd";
2. select * from media where type = "dvd";
3. select * from media where type = "book";
4. show partitions media;
5. \$ hadoop fs -ls /user/hive/warehouse/media
6. drop table media;