

1) External And Internal Tables

External Table :

- #Hive is owner of metadata only.
- # If we drop only metadata will be lost, actual data inside hdfs wont be dropped

Create External Table ...

Internal Table :

- # Hive owns hdfs file (actual data inside hdfs) + metadata
- # if we drop both will be lost
- # Default in Hive

Create Table ...

2) Sample Create Table Statement

Local File -> /home/ruby/files/employee_data.txt

499,Poole:GBR,England,141000

501,Blackburn:GBR,England,140000

500,Bolton:GBR,England,139020

502,Newport:GBR,Wales,139000

503,PrestON:GBR,England,135000

504,Stockport:GBR,England,132813

create table if not exists employeetable1 (col1 int,col2 array<string>,col3 string,col4 int)

row format delimited fields terminated by','

collection items terminated by':'

lines terminated by'\n'

stored as textfile;

create external table if not exists employeetable1(col1 int,col2 array<string>,col3 string,col4 int)

row format delimited fields terminated by','

collection items terminated by':'

lines terminated by'\n'

stored as textfile

location'/user/ruby/employee';

- By default hive will store metadata in below path variable

*set hive.metastore.warehouse.dir=/user/hive/warehouse

*location attribute can be used to override above

- Loading data into Tables using load command - INTO/OVERWRITE

```
load data local inpath '/home/ruby/files/employee_data.txt' into table employeetable1;
```

*local -> input file in local path not in hdfs

*hdfs -> input file in hdfs path, omit local keyword

*into/overwrite -> into will append

overwrite clear and load new data

- Verification

```
hive > select * from employeetable1
```

- hive > drop table employeetable1
- To see table headers

```
*set hive.cli.print.header=true
```

```
hive > select * from employeetable1
```