[数据文件数据加载到hive表](http://blog.chinaunix.net/uid-77311-id-4581156.html)

**本地文件加载到hive表**  
1.在hxl数据库下创建表  
hive> create table tb\_emp\_info  
    > (id int,  
    > name string,  
    > age int,  
    > tel string)  
    > ROW FORMAT DELIMITED  
    > FIELDS TERMINATED BY '|'  
    > STORED AS TEXTFILE;  
OK  
Time taken: 0.296 seconds  
hive> show tables in hxl;  
OK  
tb\_emp\_info  
Time taken: 0.073 seconds

2.准备加载数据  
[hadoop1@node1 hive]$ more tb\_emp\_info.txt  
1|name1|25|13188888888888  
2|name2|30|13888888888888  
3|name3|3|147896221  
4|name4|56|899314121  
5|name5|12|899314121  
6|name6|9|899314121  
7|name7|32|899314121  
8|name8|42|158964  
9|name9|86|899314121  
10|name10|45|789541

3.本地系统加载文件数据  
进入到tb\_emp\_info.txt文件所在的目录,然后执行hive进入到hive模式  
[hadoop1@node1 hive]$ hive  
hive> use hxl;  
OK  
Time taken: 0.103 seconds  
hive> load data local inpath 'tb\_emp\_info.txt' into table tb\_emp\_info;   
Copying data from file:/home/hadoop1/file/hive/tb\_emp\_info.txt  
Copying file: file:/home/hadoop1/file/hive/tb\_emp\_info.txt  
Loading data to table hxl.tb\_emp\_info  
OK  
Time taken: 0.694 seconds

若是分区表的话，需要指点导入的分区，如：

hive> load data local inpath 'login.txt' into table tb\_sso\_ver\_login\_day partition(statedate=20141201);

4.查看加载进去的数据  
hive> select \* from tb\_emp\_info;  
OK  
1       name1   25      13188888888888  
2       name2   30      13888888888888  
3       name3   3       147896221  
4       name4   56      899314121  
5       name5   12      899314121  
6       name6   9       899314121  
7       name7   32      899314121  
8       name8   42      158964  
9       name9   86      899314121  
10      name10  45      789541  
  
5.可以进入到hdfs目录下查看该表对应的文件  
hive> dfs -ls /user/hive/warehouse/hxl.db/tb\_emp\_info;  
Found 1 items  
-rw-r--r--   3 hadoop1 supergroup        214 2014-10-28 17:31 /user/hive/warehouse/hxl.db/tb\_emp\_info/tb\_emp\_info.txt  
  
  
**HDFS文件导入到Hive表**  
1.查看hdfs系统上的文件  
$hadoop fs -cat /user/hadoop1/myfile/tb\_class.txt  
输出部分  
0|班级0|2014-10-29 14:10:17|2014-10-29 14:10:17  
1|班级1|2014-10-29 14:10:17|2014-10-29 14:10:17  
2|班级2|2014-10-29 14:10:17|2014-10-29 14:10:17  
3|班级3|2014-10-29 14:10:17|2014-10-29 14:10:17  
4|班级4|2014-10-29 14:10:17|2014-10-29 14:10:17  
5|班级5|2014-10-29 14:10:17|2014-10-29 14:10:17  
6|班级6|2014-10-29 14:10:17|2014-10-29 14:10:17  
7|班级7|2014-10-29 14:10:17|2014-10-29 14:10:17  
8|班级8|2014-10-29 14:10:17|2014-10-29 14:10:17

2.创建表  
create table tb\_class\_info  
(id int,  
class\_name string,  
createtime timestamp ,  
modifytime timestamp)  
ROW FORMAT DELIMITED  
FIELDS TERMINATED BY '|'  
STORED AS TEXTFILE;

3.导入表  
load data inpath '/user/hadoop1/myfile/tb\_class.txt' into table tb\_class\_info;