*Import sqoop hive(default)*

\*database in mysql.

mysql> use mysql\_hive;

mysql> create table mysqll(id int,f\_name varchar(20),l\_name varchar(20));

mysql> insert into mysqll values(101,'nikhil','nik');

mysql> select \* from mysqll;

\*sqoop import command, for mysql to hive

$ sqoop import --connect 'jdbc:mysql://localhost:3306/mysql\_hive' --username root --password cloudera --table mysqll --hive-import -m 1

\*hadoop fs -cat /user/hive/warehouse/mysqll/part-m-00000

\*mysql data goes to default in hive bcoz we didn't give any proper path

hive> show databases;

hive> use default;

hive> set hive.cli.print.current.db=true;

hive (default)> show tables;

hive> select \* from mysqll;

{same mysql table imported into hive, as it is, into warehouse by default}

*Import sqoop hive (into – hive - database)*

\*in hive already database&tables have to exists..

hive> create database hive\_mysql;

hive> use hive\_mysql;

hive> set hive.cli.print.current.db=true;

hive (hive\_mysql)> create table hivee(id int,f\_name varchar(20),l\_name varchar(20));

\*sqoop command mysql database to hive(table) database

$ sqoop import --connect 'jdbc:mysql://localhost:3306/mysql\_hive' --username root --password cloudera --table mysqll --hive-import -m 1 --hive-table hive\_mysql.hivee

\*$ hadoop fs -cat /user/hive/warehouse/hive\_mysql.db/hivee/part-m-00000

\*to check the table data

hive (hive\_mysql)> show tables;

hive (hive\_mysql)> select \* from hivee;

--\* now same sqoop command runs many time

--\* check table data, that data table appended next to next

$ sqoop import --connect 'jdbc:mysql://localhost:3306/mysql\_hive' --username root --password cloudera --table mysqll --hive-import -m 1 --hive-table hive\_mysql.hivee

--\* mapper makes the copy and data atteched next to next

$ hadoop fs -cat /user/hive/warehouse/hive\_mysql.db/hivee/part-m-00000\_copy\_2

\*to check the table data

hive (hive\_mysql)> show tables;

hive (hive\_mysql)> select \* from hivee;

\*

- for describe and all details of the table

hive (hive\_mysql)> show create table hivee;

'hdfs://quickstart.cloudera:8020/user/hive/warehouse/hive\_mysql.db/hivee'

*Export sqoop*

\*export - (hive to mysql)

\*in mysql already database & tables have to exists.

mysql> create database mysql\_htab;

mysql> create table mysqlhive(id int,f\_name varchar(20),l\_name varchar(20));

\* in hive

- for describe and all details of the table

hive (hive\_mysql)> show create table hivee;

'hdfs://quickstart.cloudera:8020/user/hive/warehouse/hive\_mysql.db/hivee'

\*

$ sqoop export --connect 'jdbc:mysql://localhost:3306/mysql\_htab' --username root --password cloudera --table mysqlhive -m 1 --hcatalog-database hive\_mysql --hcatalog-table hivee

mysql> select \* from mysqlhive;

*export hive table*

hive (tabl)> create table emp(id int,name string,sal int,gen string,dno int)

> row format delimited

> fields terminated by ','

> ;

hive (tabl)> load data local inpath '/home/cloudera/xyz.txt' into table emp;

hive (tabl)> select \* from emp;

OK

101 manish 15000 m 11

102 pranjal 17000 f 12

103 dheeraj 18000 m 13

104 ak 19000 m 11

105 prachi 15000 f 12

106 nikhil 16000 m 11

hive (tabl)> create table summary(dno int,gen string,tot int,avg int,cnt int,max int,min int);

hive (tabl)> insert overwrite table summary

> select dno, gen, sum(sal), avg(sal), count(\*), max(sal), min(sal)

> from emp group by dno,gen;

hive (tabl)> select \* from summary;

OK

11 m 50000 16666 3 19000 15000

12 f 32000 16000 2 17000 15000

13 m 18000 18000 1 18000 18000

\*mysql> create table abc(dno int,gen varchar(2),tot int,avg int,cnt int,max int,min int);

\*$ sqoop export --connect 'jdbc:mysql://localhost:3306/sql\_data' --username root --password cloudera --table abc --export-dir '/user/hive/warehouse/tabl.db/summary/000000\_0' --input-fields-terminated-by '\001'