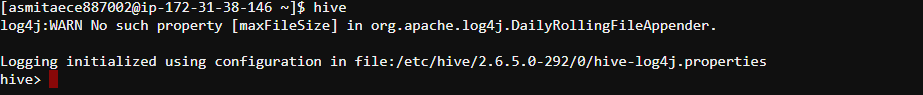
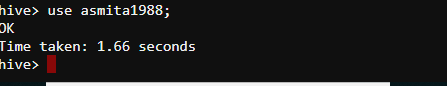
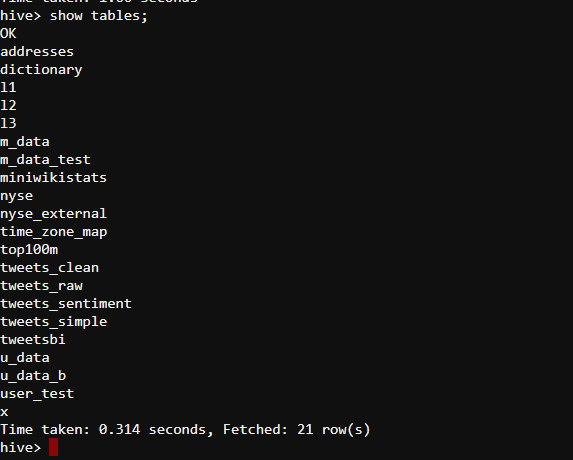
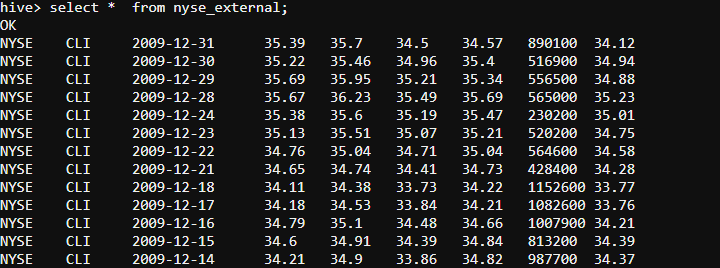
Login to Hive console :

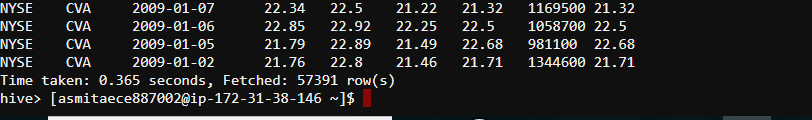


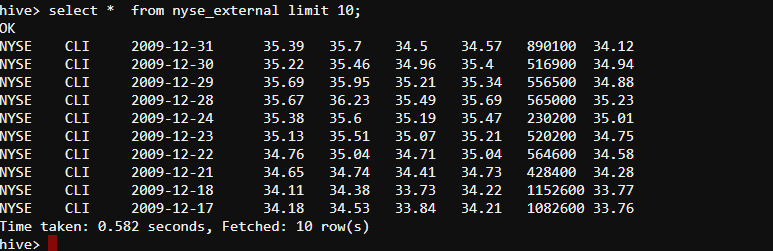
Use my database



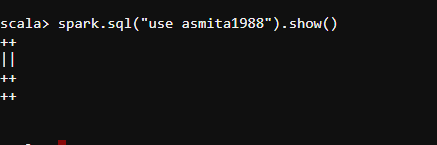


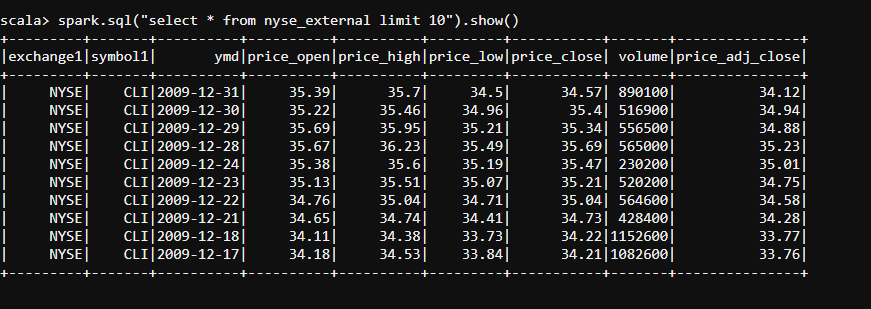






Now in another CLI , we open the scalaa prompt

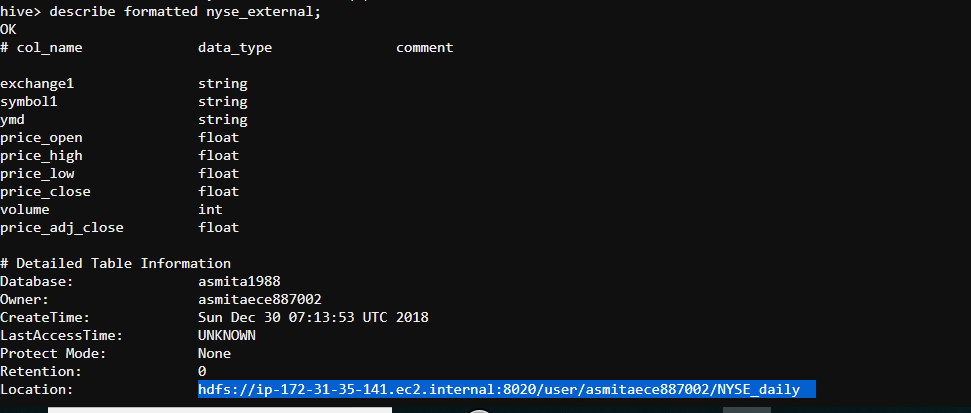




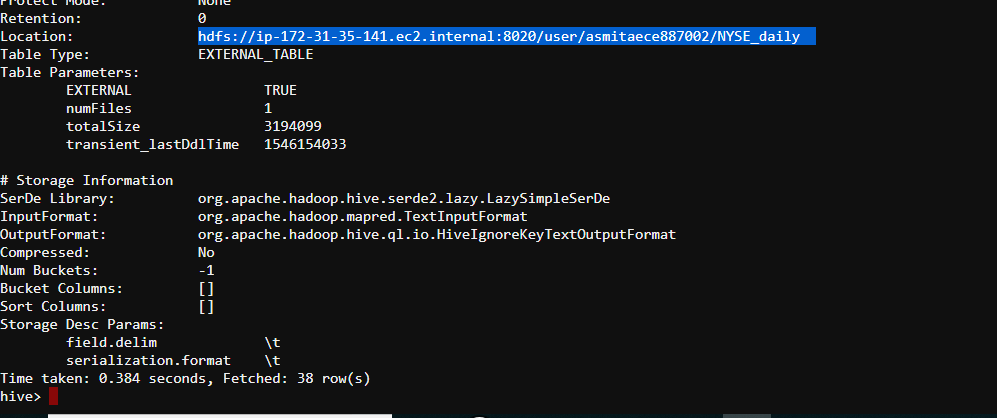
Hence the spark sql is not talking to hadoop at all , it directly finds out the query result from metastore

IF we check from HIVE -

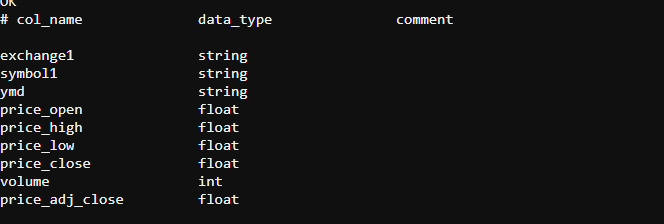
Describe formatted nyse\_external;



The location of the table is as follows :

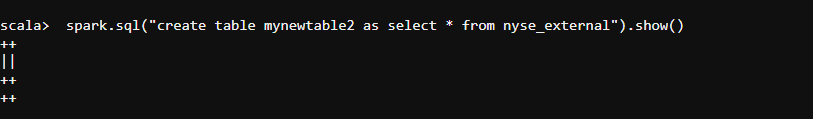


Below are the various columns formats :

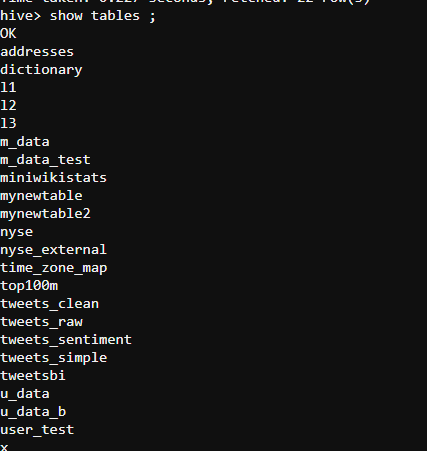


Now based on the column type , the dataframe was created in spark shell .

Now create a new table in spark sql --

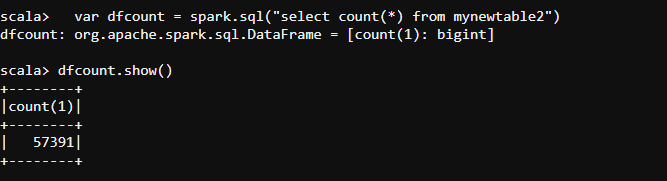


Now check in hive command

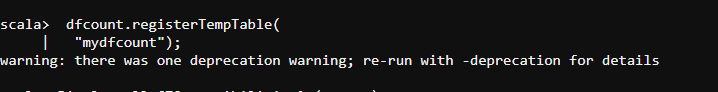


The table mynewtables2 is created

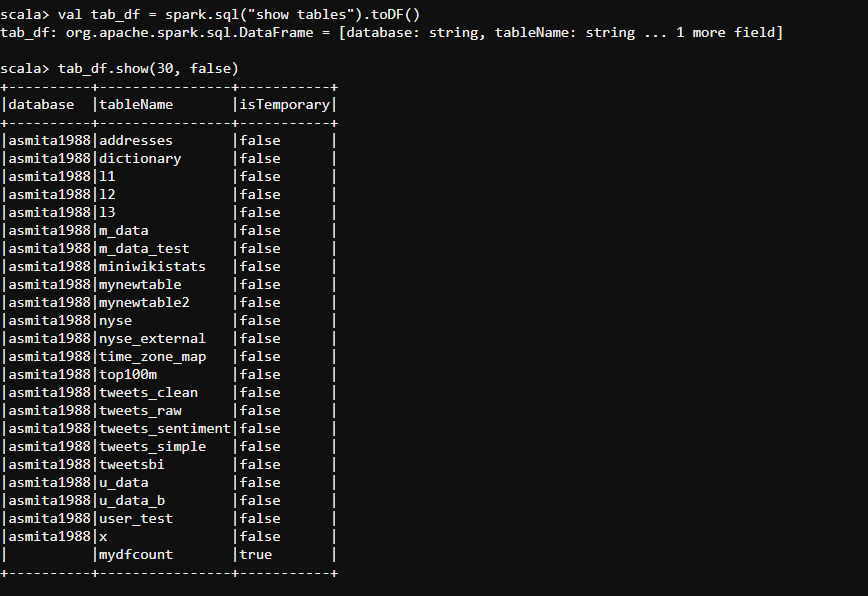
From spark – shell , check the count



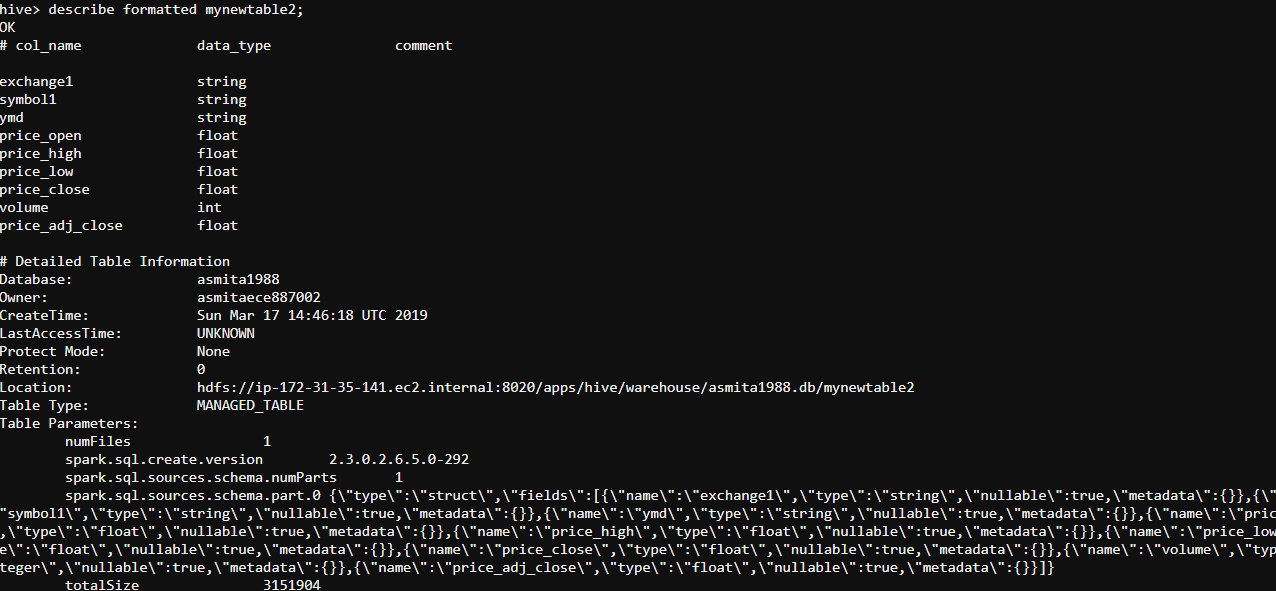
The dataframe is registered as a temporary view

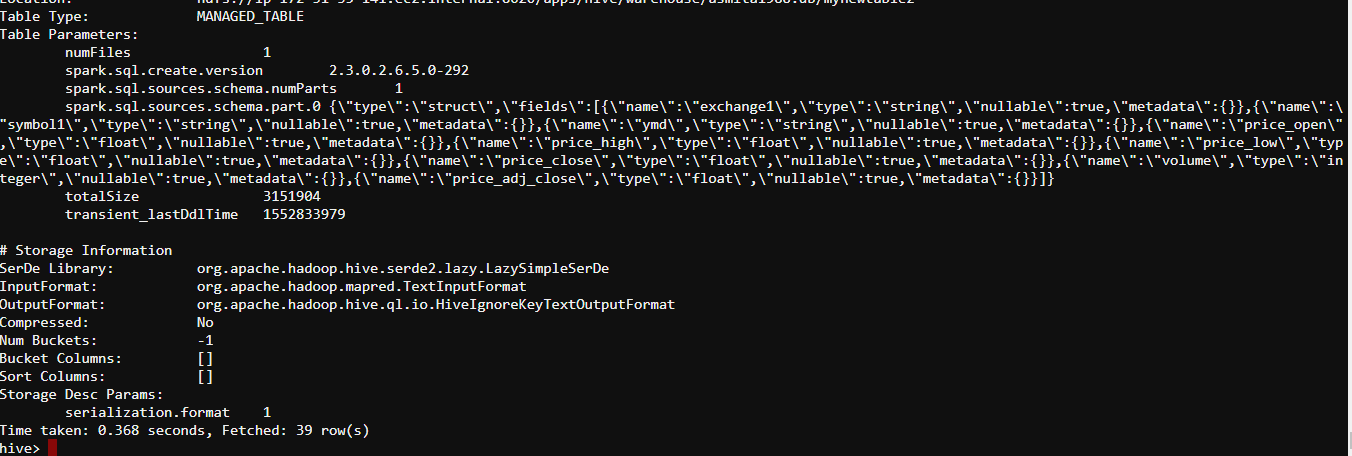


Now checking the mydfcount , its uts a cached temporary data in the memory



Now checking the new table metadata from hive





Location of the file :

hdfs://ip-172-31-35-141.ec2.internal:8020/apps/hive/warehouse/asmita1988.db/mynewtable2

