FILE FORMATES:

1, TEXT FORMATES

2, AVRO FORMATES

3, PARQUET FORMATES

4, RC AND ORC FORMATES

------------------------------------TEXT FILE ----------------------------

TEXT FILE IS ITWILL TAKE LATE OF SPACE FOR STORAGE AND WHILE RPOCESSING IT WILL TAKE LATE OF TIME . THAT'S WHY IT WILL INTRADUCE LATE OF FILE FORMATES LIKE .

1 SEQUENCE, AVRO,PARQUET,RC AND ORC

HIVE IS INTRADUCE TWO FILE FORMATES RC AND ORC THIS ARE FILE FORMETS

TEXT FILE IS HUMANREADABLE AND STORAGE DATA IS BULKY AND NOT EFFICIENT TOQUERY

DONOT SUPPORT BLOCK COMPESSIONS

---------------------------------SEQUENCE FILE---------------------------------------------------

SEQUENCE FILE IS PROVIED PERSISTENT DATA STUCTURE FOR BINARY KEY VALUE PAIRES

ROW BASED

COMMONLY USED TO TRANSFRE DATA BERWEEN MAP REDUCE JOBS

SUPPORT SPLITTING EVEN WHEN THE DATA IS COMPRESSED

------------------------------AVRO FILE FORMAT--------------------------

AVRO IS SERILIZATION FLATFORM IT MEANS IF YOU WANT ANY FILE TRANSFOR OR DATA ,

WHILE TRANSFORING DEFFIRENT NODES CAN TRASFORE THE DIRECTLY THAT'S THERE IS UESING SERILAIZATION

---------------------------OPTIMIZED ROW COLUMNAR-----------------------------------

ORC IS BETTER PERFOMANCE AND BETTER COMPESSION ,ORC IS COLUMN ORIENTED BINARYFILE FORMAT , IF YOU WANT STORING TB OF DATA REDUCE CAPACITY IT HAVE ORC , ORC IT CAN HAVE INDEXSING

---------------------------PARTITION AND BUCKETING------------------------------------