------------------------------------ChiaLing-------------------------------------

sed -i 's/"""//g' opendata\_essays.log

sed -i '1d' opendata\_essays.log

#######replace all ", " to "##@@##" inorder to process as seperated column #####

sed -i 's/,\ /##@@##/g' opendata\_essays.log

#######replace all "^M" to "" inorder to process as seperated column #####

sed -i 's/\r//g' opendata\_essays.log

sed -i 's/\n//g' opendata\_essays.log

CREATE EXTERNAL TABLE opendata\_projects\_10 (projectid int, teacher\_acctid string, schoolid string, school\_ncesid string, school\_latitude string, school\_longitude string, school\_city string, school\_state string, school\_zip string, school\_metro string, school\_district string, school\_county string, school\_charter string, school\_magnet string, school\_year\_round string, school\_nlns string, school\_kipp string, school\_charter\_ready\_promise string, teacher\_prefix string, teacher\_teach\_for\_america string, teacher\_ny\_teaching\_fellow string, primary\_focus\_subject string, primary\_focus\_area string, secondary\_focus\_subject string, secondary\_focus\_area string, resource\_type string, poverty\_level string, grade\_level string, vendor\_shipping\_charges string, sales\_tax string, payment\_processing\_charges string, fulfillment\_labor\_materials string, total\_price\_excluding\_optional\_support string, total\_price\_including\_optional\_support string, students\_reached string, total\_donations string, num\_donors string, eligible\_double\_your\_impact\_match string, eligible\_almost\_home\_match string, funding\_status string, date\_posted string, date\_completed string, date\_thank\_you\_packet\_mailed string, date\_expiration string)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

stored as textfile;

load data local inpath "/home/cloudera/workspace/big\_data\_project/database/opendata\_projects.log" into table opendata\_projects\_10;

CREATE TABLE opendata\_essays\_10

(projectid string, teacher\_acctid string, title string, short\_description string, need\_statement string, essay string, paragraph1 string, paragraph2 string, paragraph3 string, paragraph4 string, thankyou\_note string, impact\_letter string)

ROW FORMAT DELIMITED FIELDS TERMINATED BY ','

STORED AS TEXTFILE;

load data local inpath "/home/cloudera/workspace/big\_data\_project/database/opendata\_essays.log" into table opendata\_essays\_10;

------------------------------------ChiaLing-------------------------------------