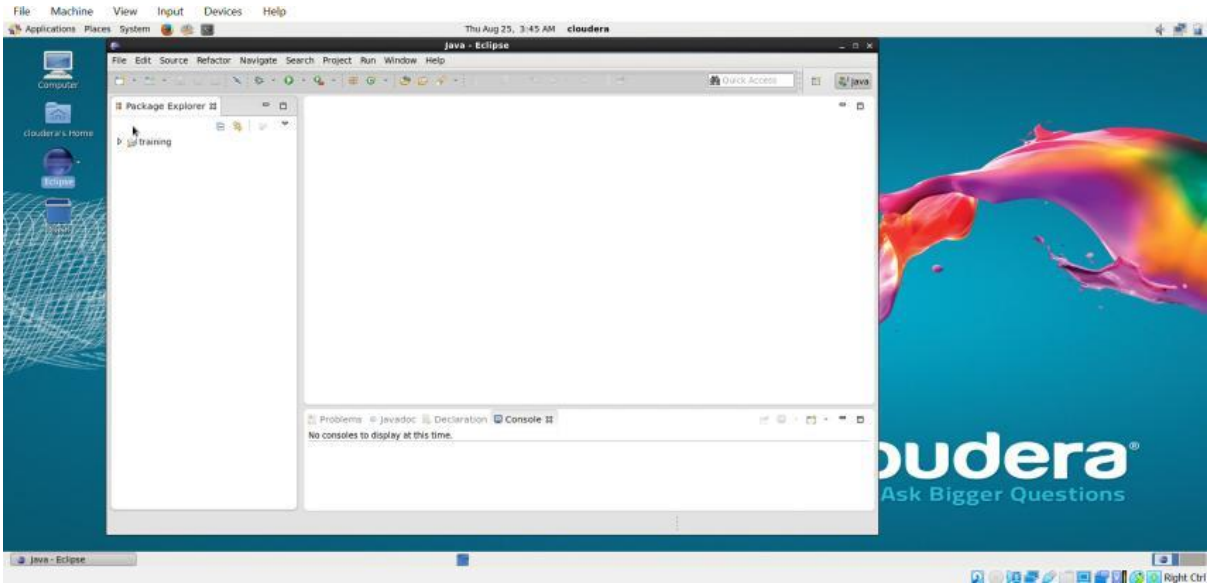
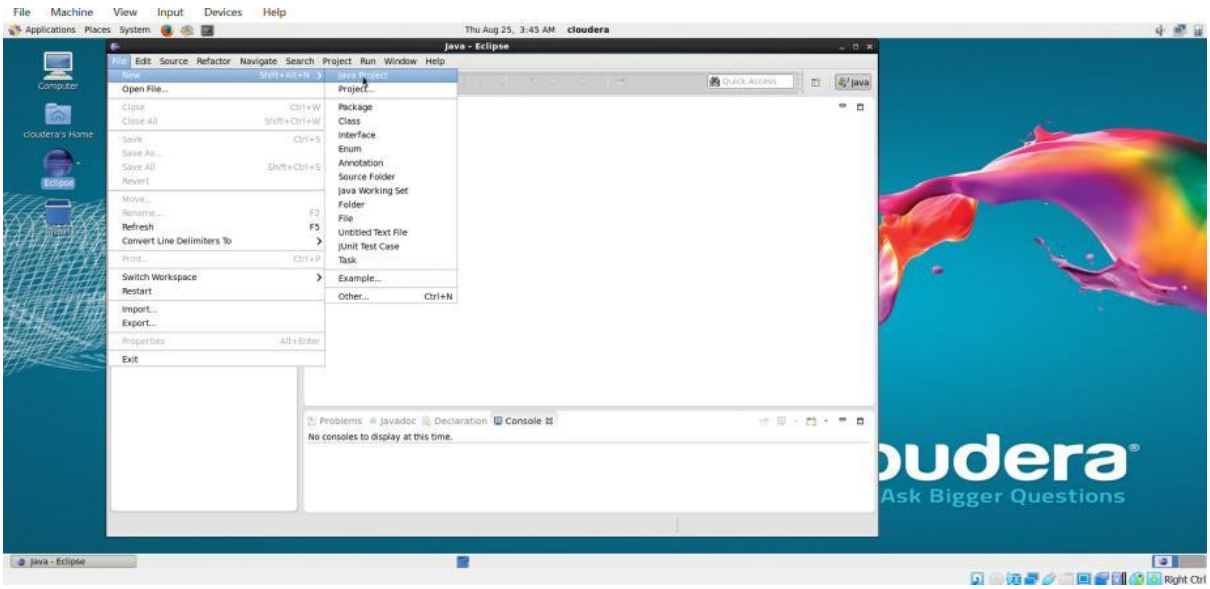
Maximum temperature Program

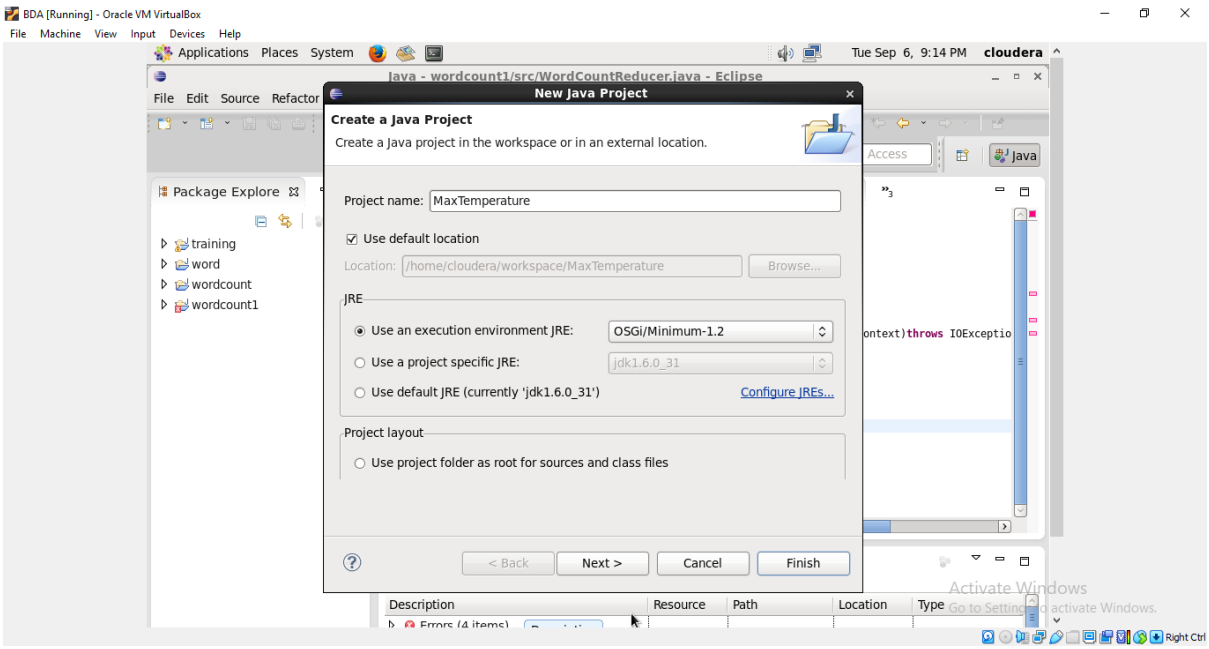
Open Oracle VM Virtual Box→clickstart →open cloudera→open eclipse



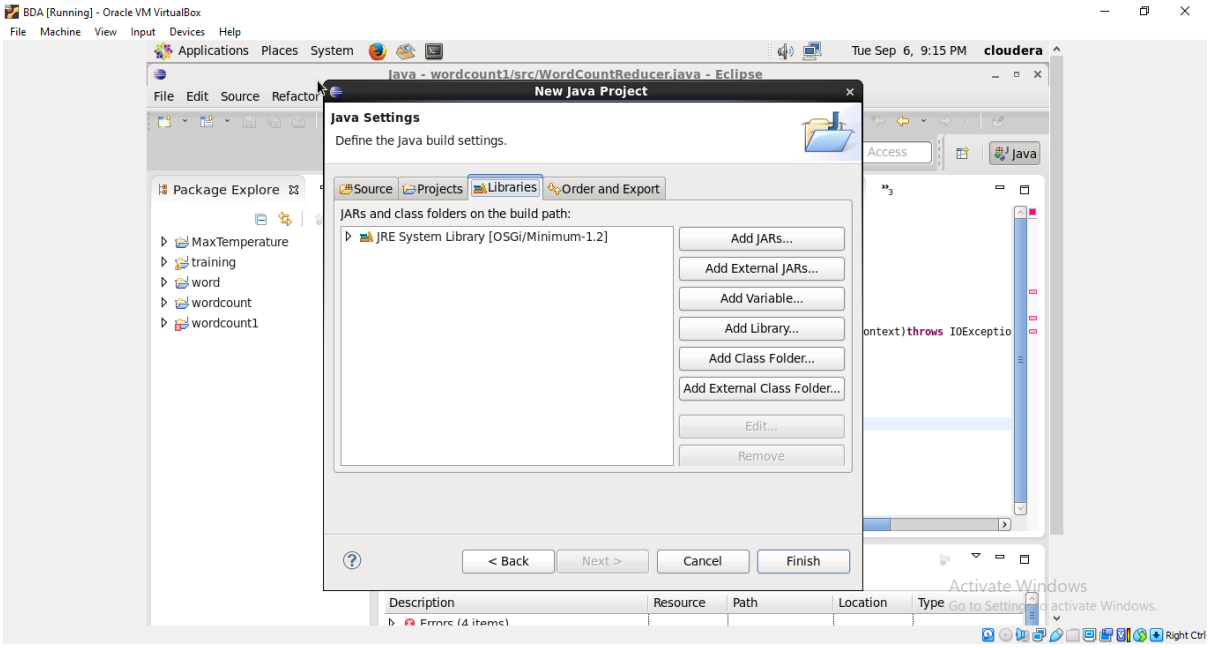
Click on file->New->Java Project



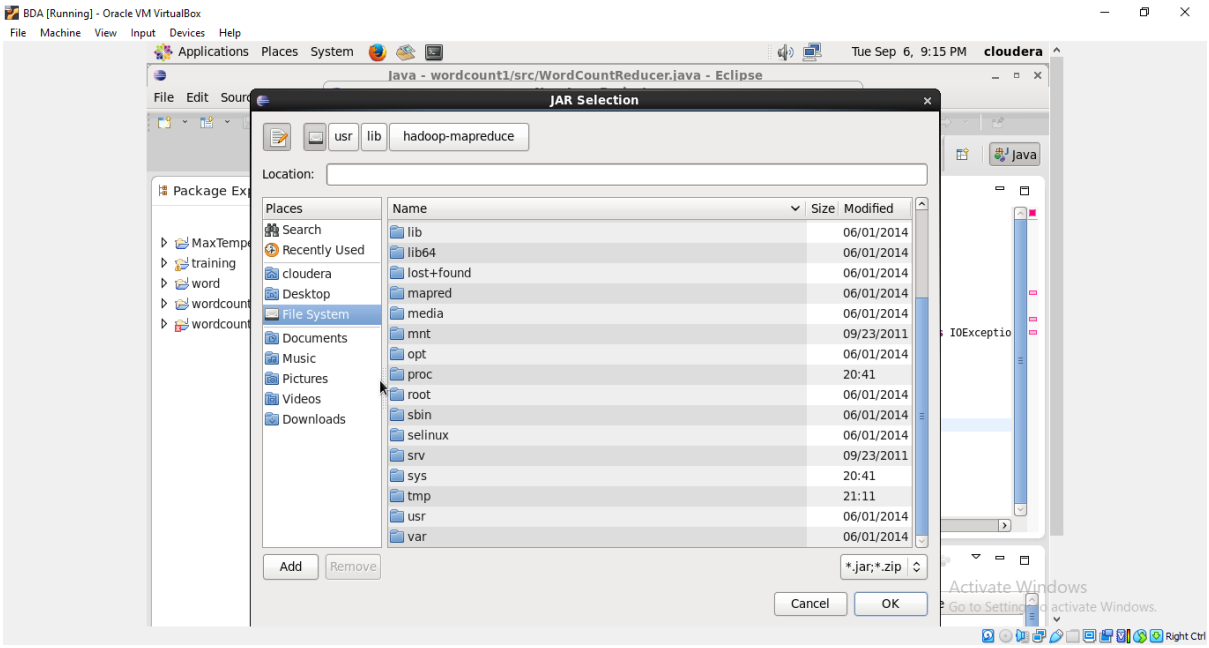
Give your project name as “MaxTemperature” and click on next.



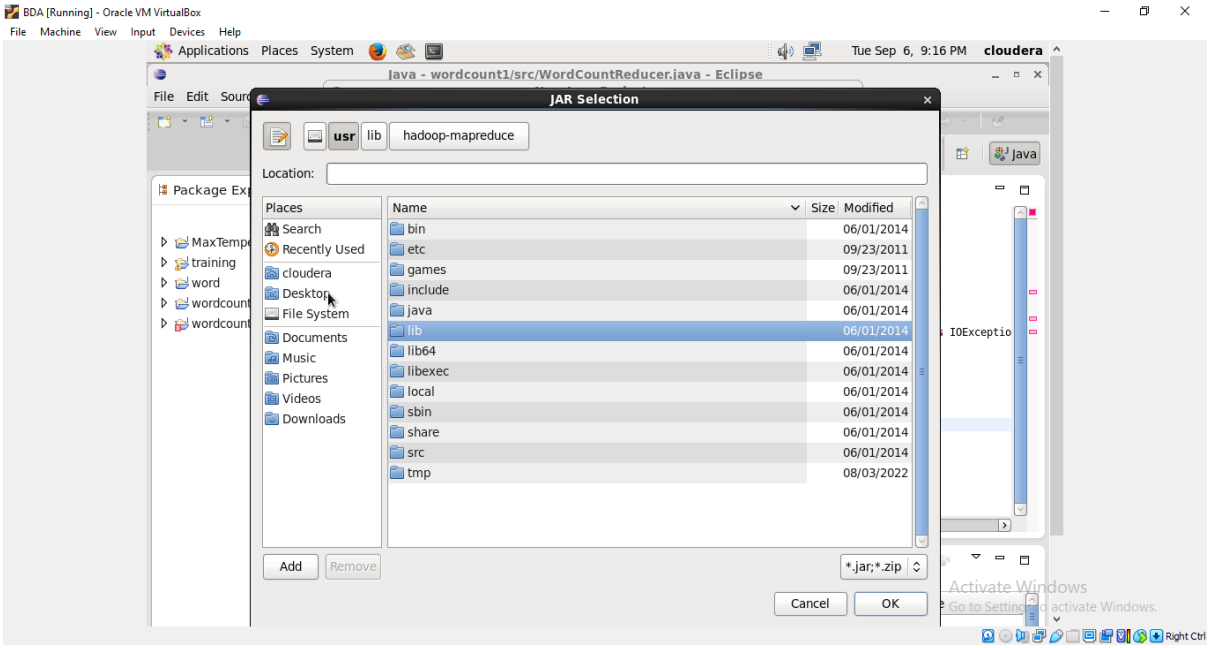
After that, click on libraries on the top and click on add external jars on the right side



Click on file system →usr

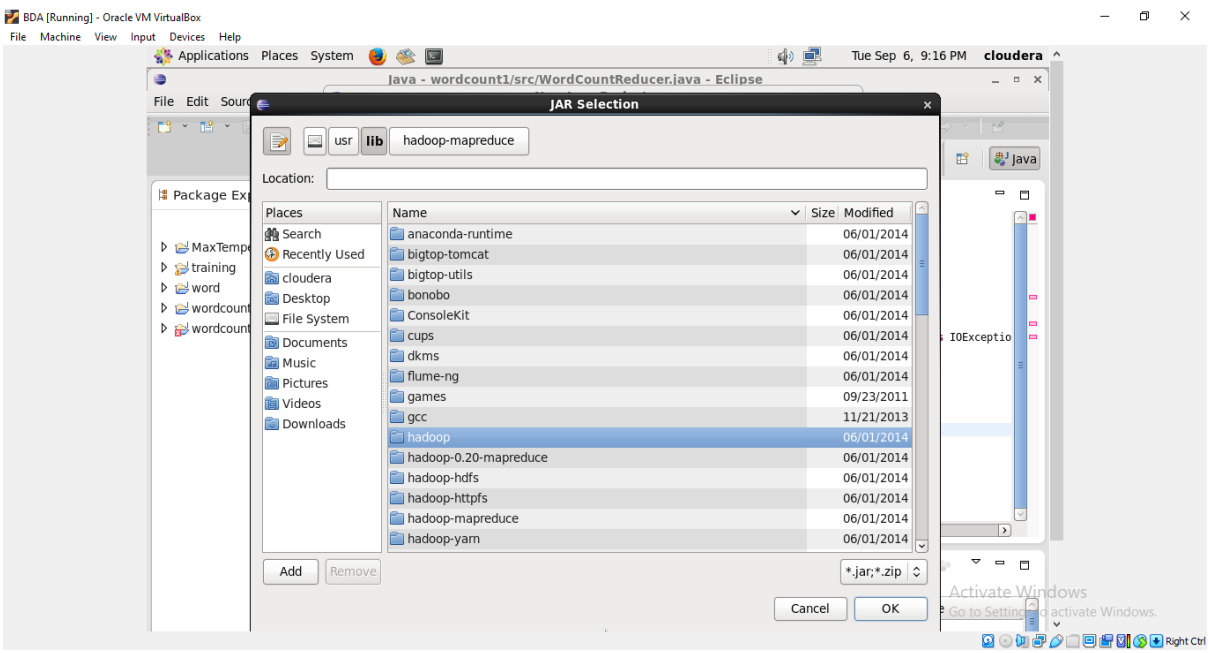


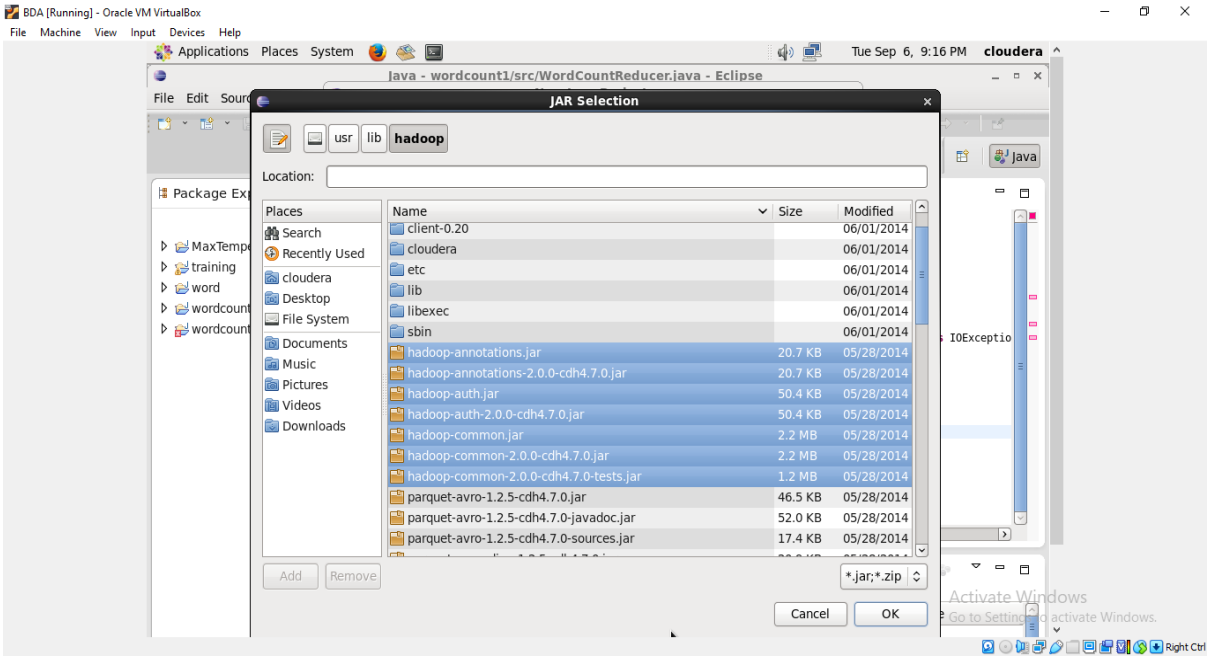
Click on lib



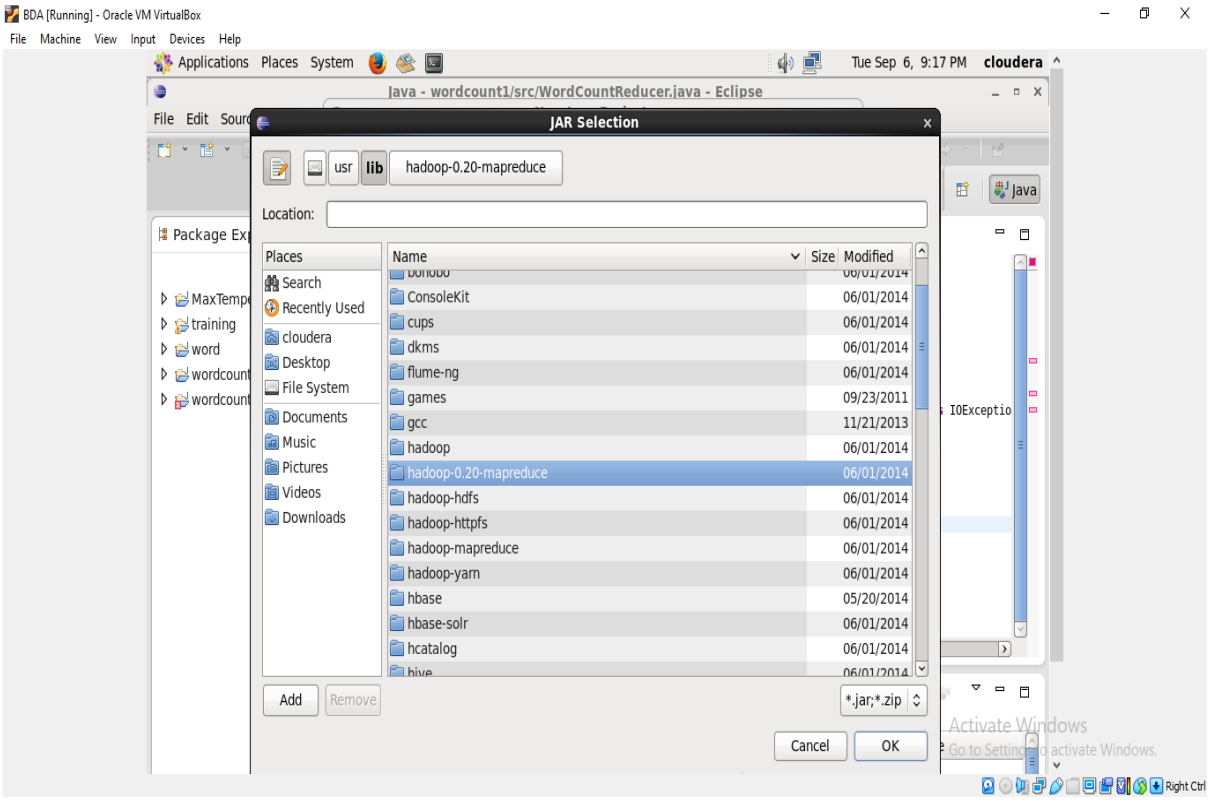
Click on hadoop

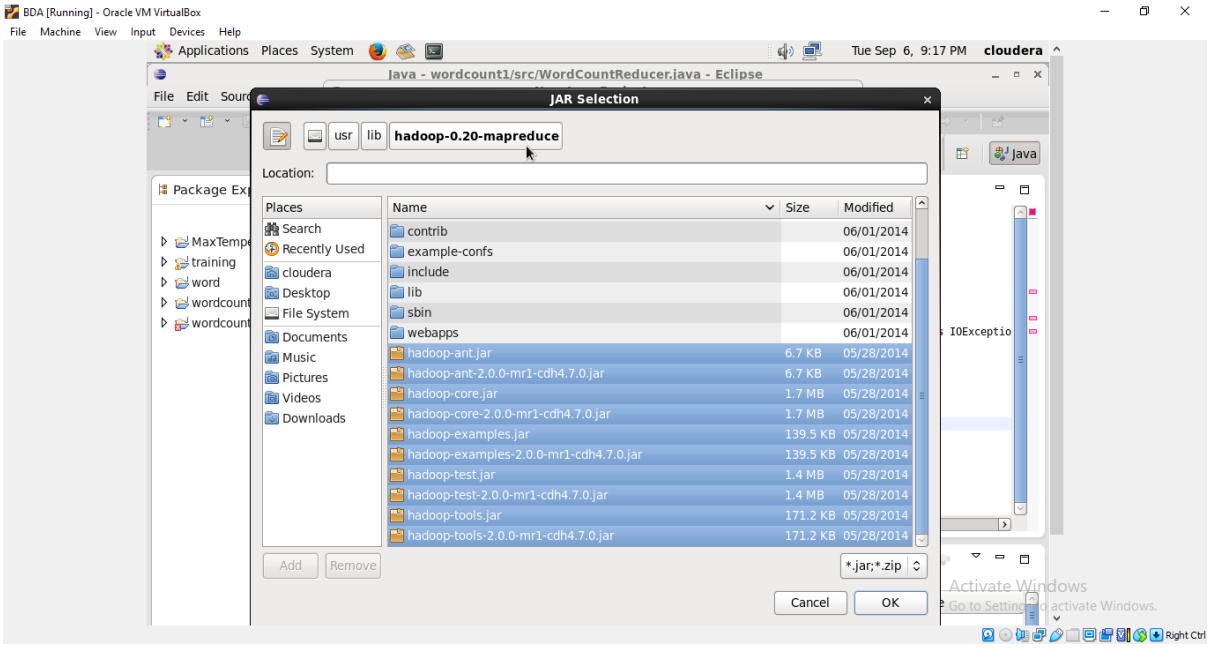
Select all the files named with hadoop and click ok



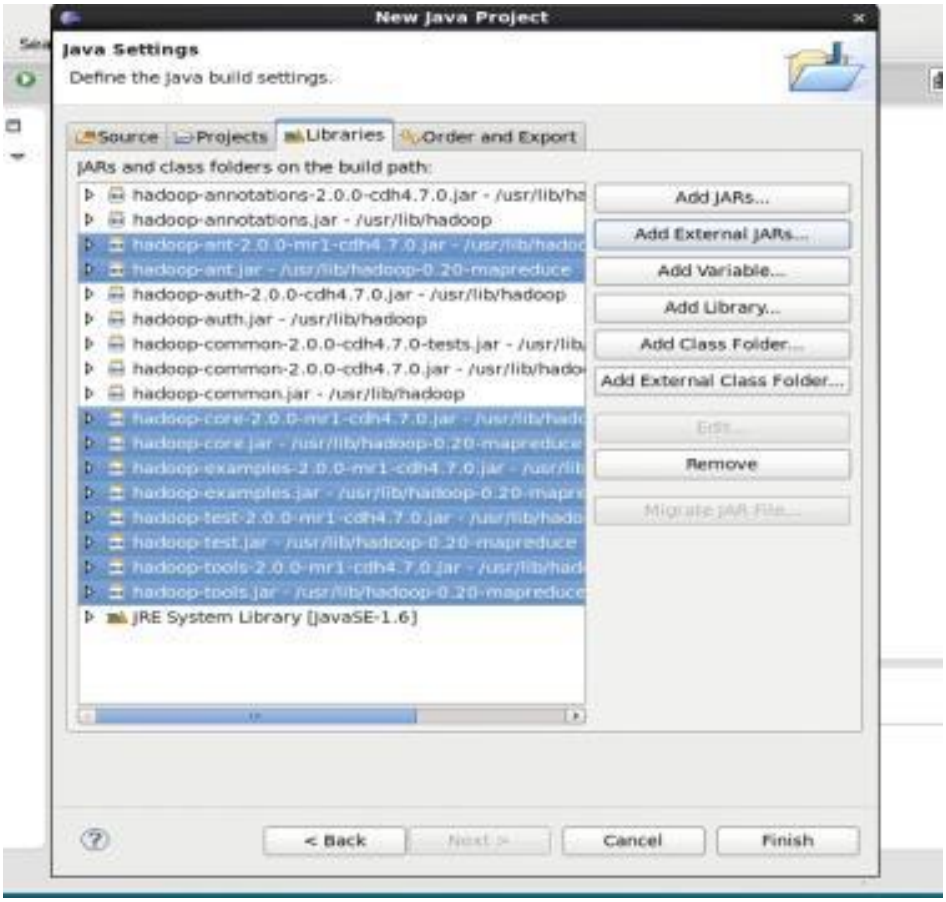


Go back to lib and click hadoop-0.20-mapreduce and select all the files named with hadoop and click ok.

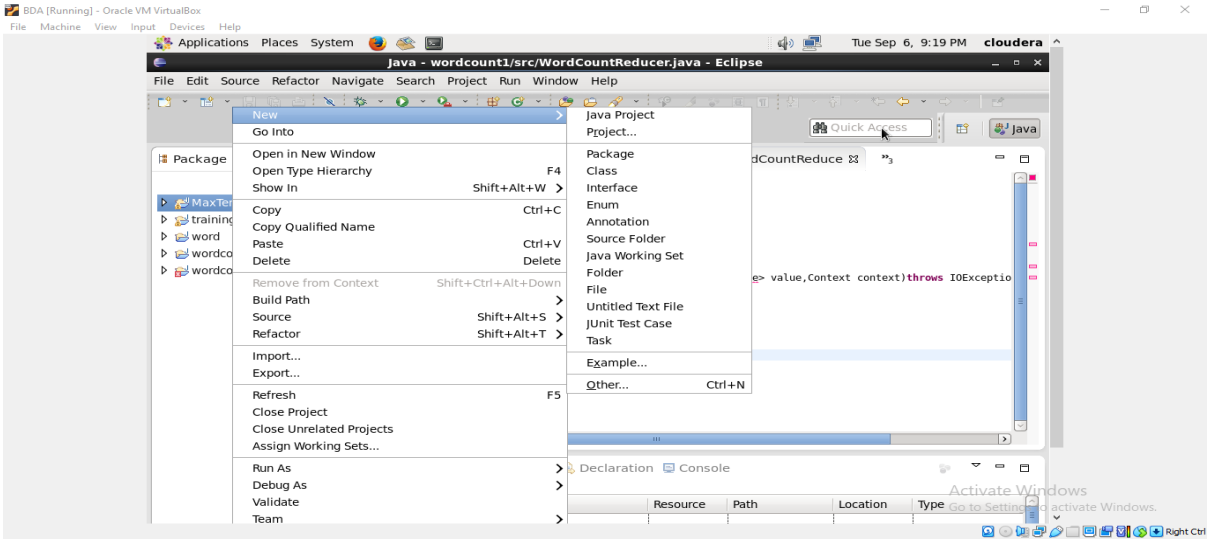


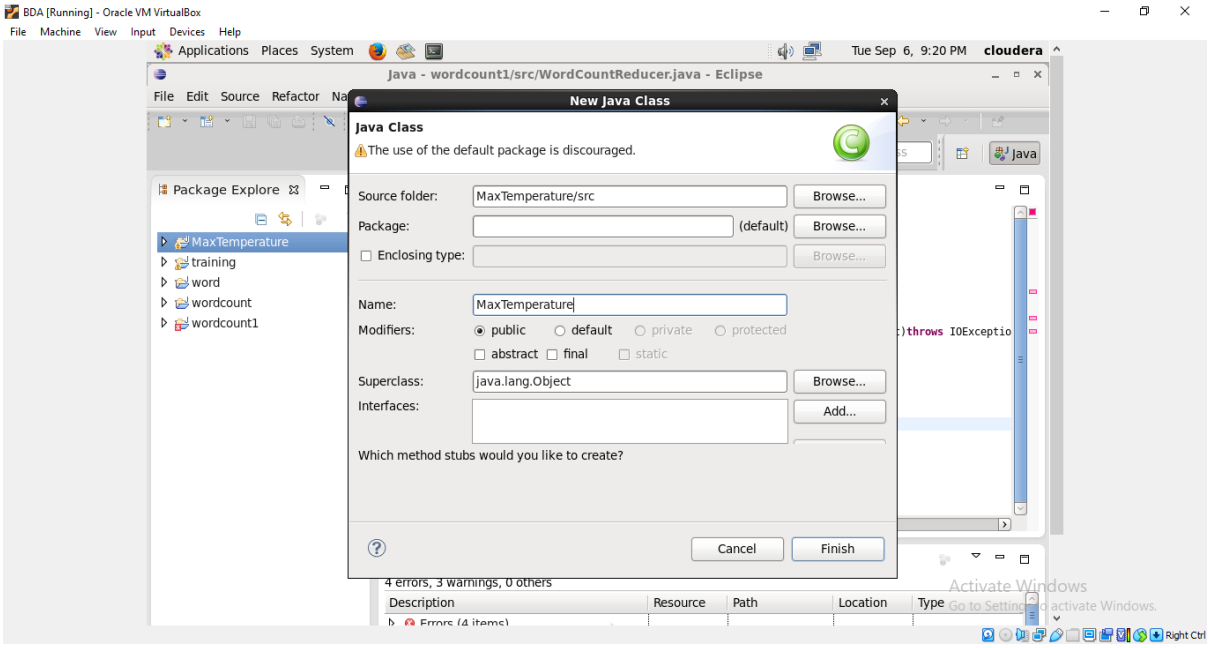


All the external files are added and now click on finish. The project is created successfully.

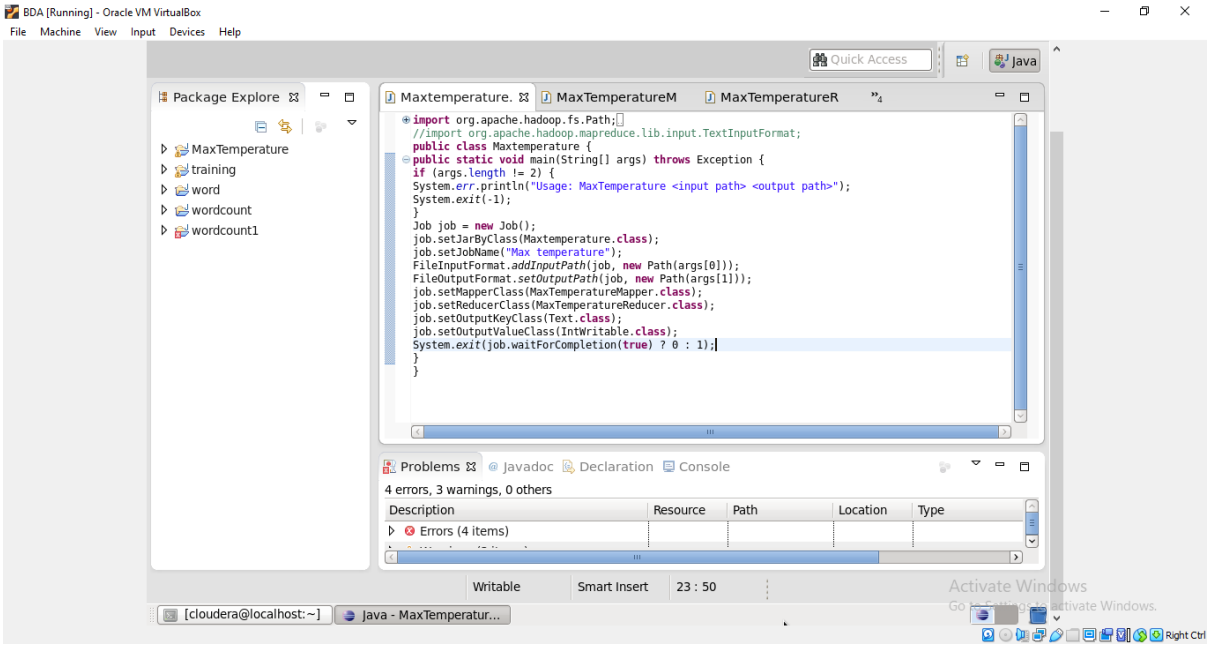


Create a class with name “MaxTemperature”

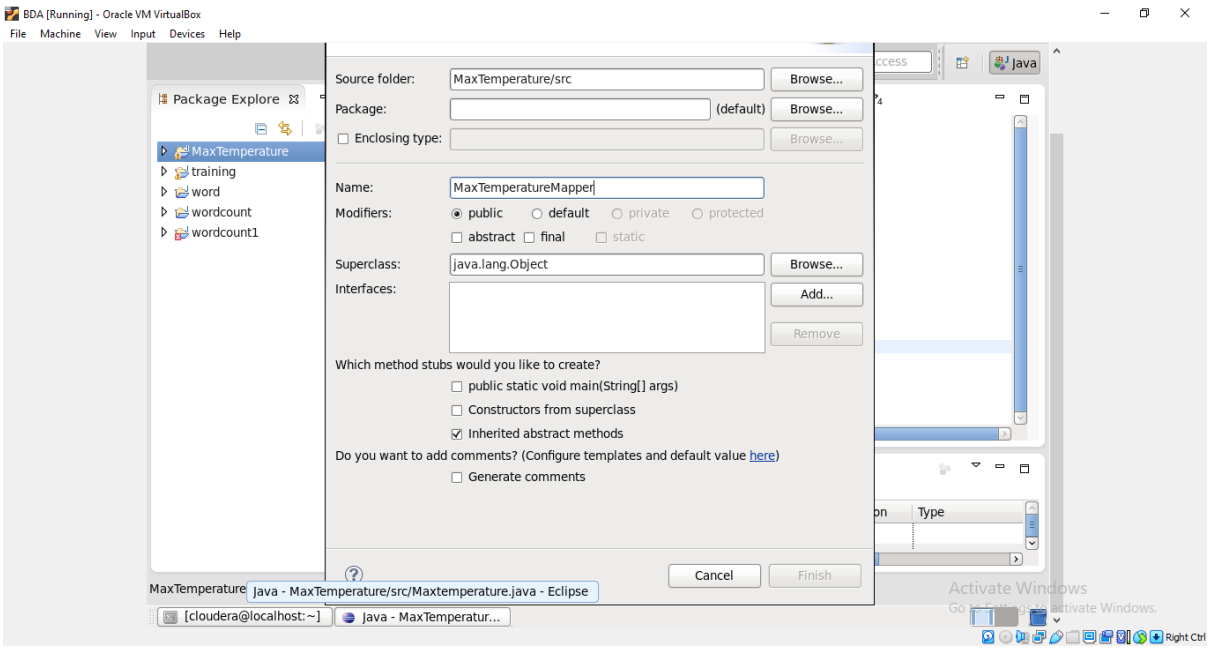




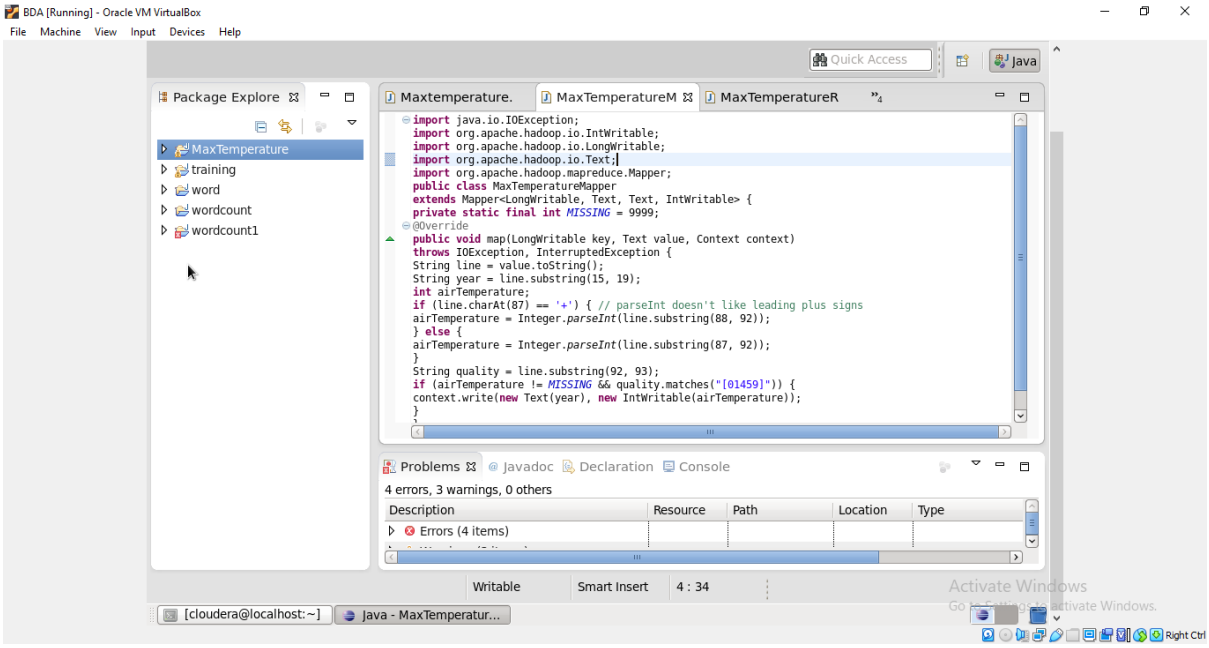
Copy and paste the program



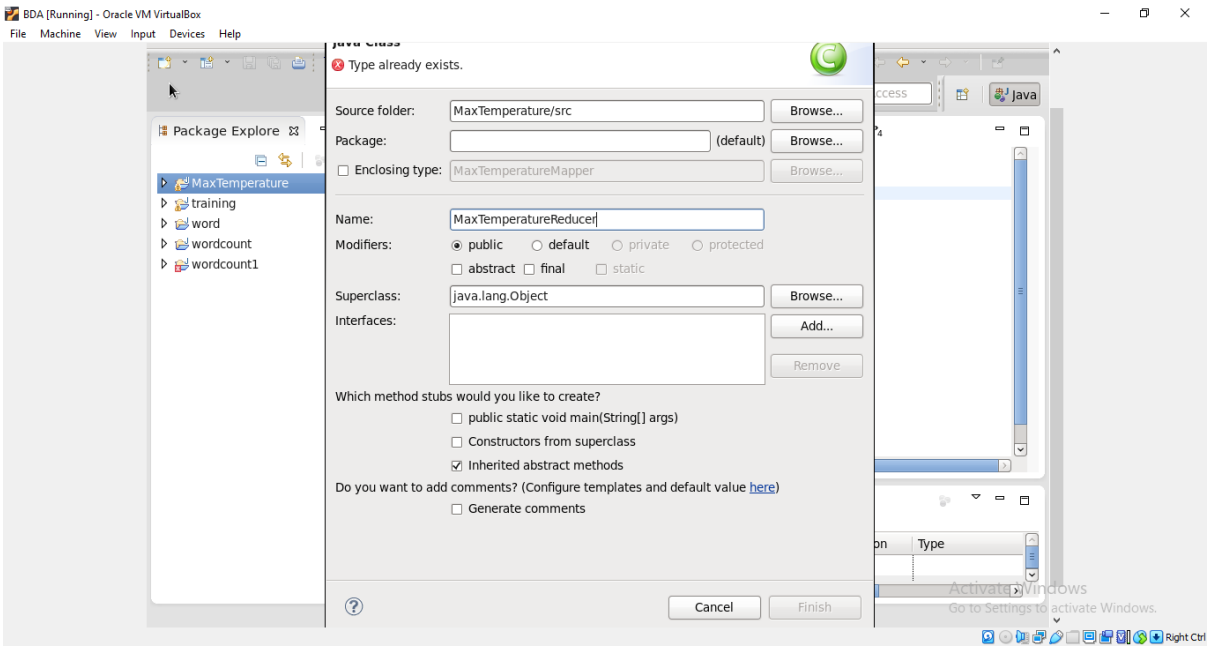
Create another class “MaxTEmperatureMapper”



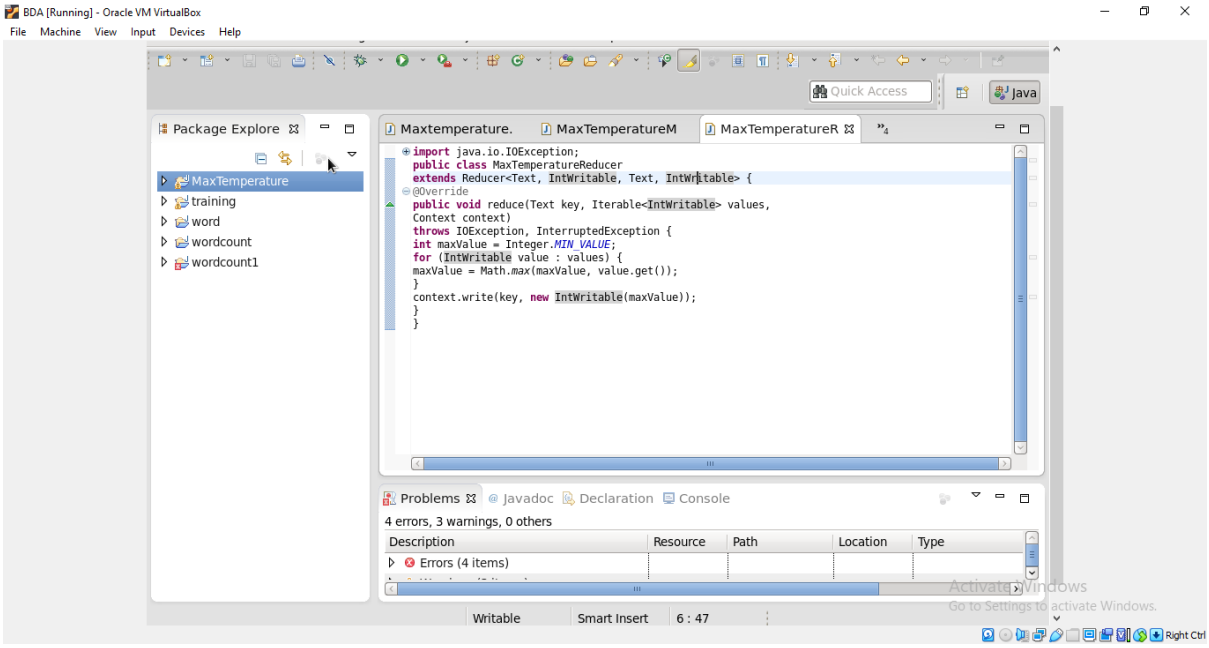
Copy and paste the program maxTemperaturemapper



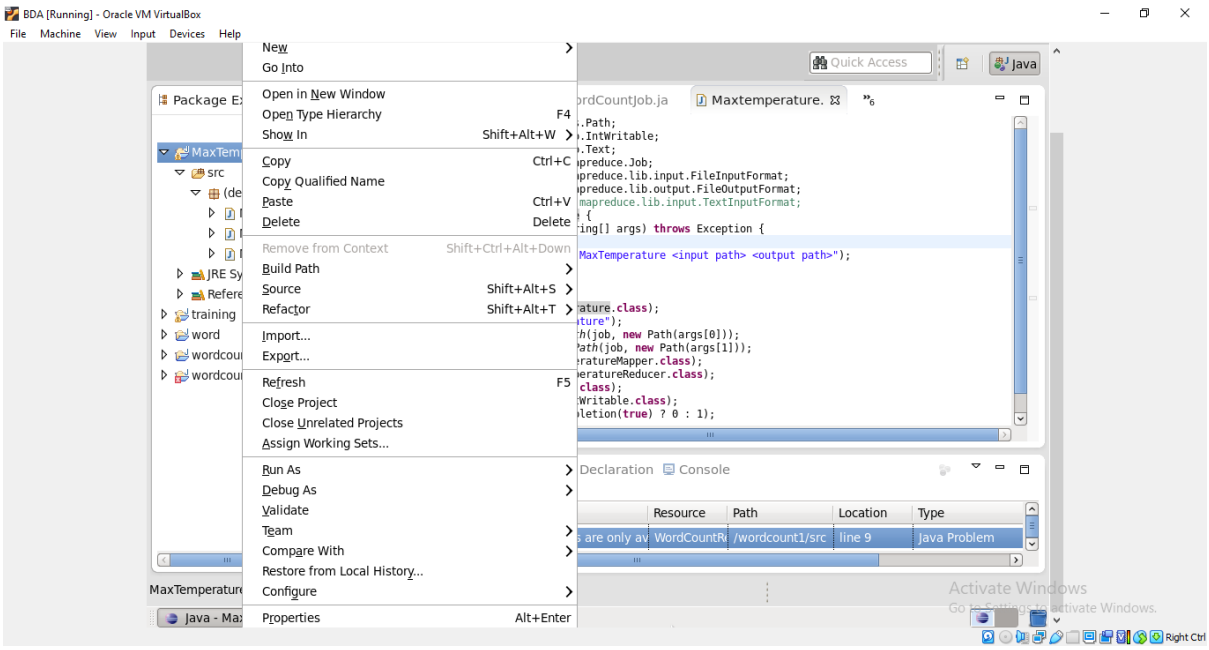
Create another class “MaxTEmperatureReducer”



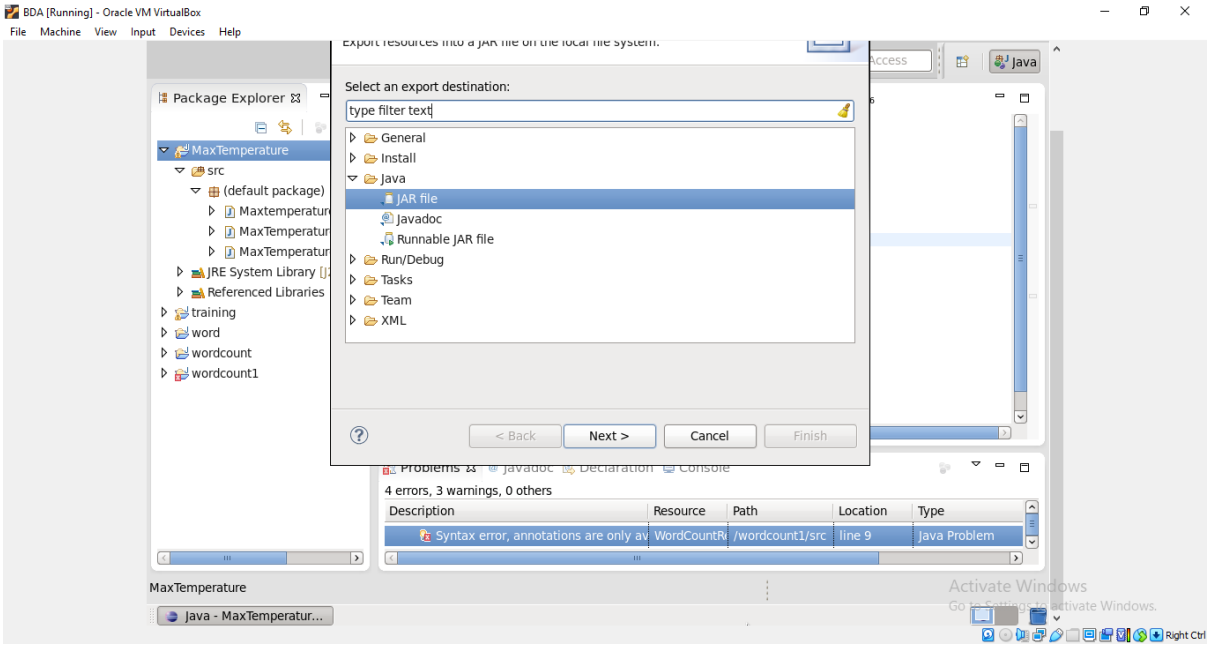
Copy and paste the program MaxTemperatureReducer



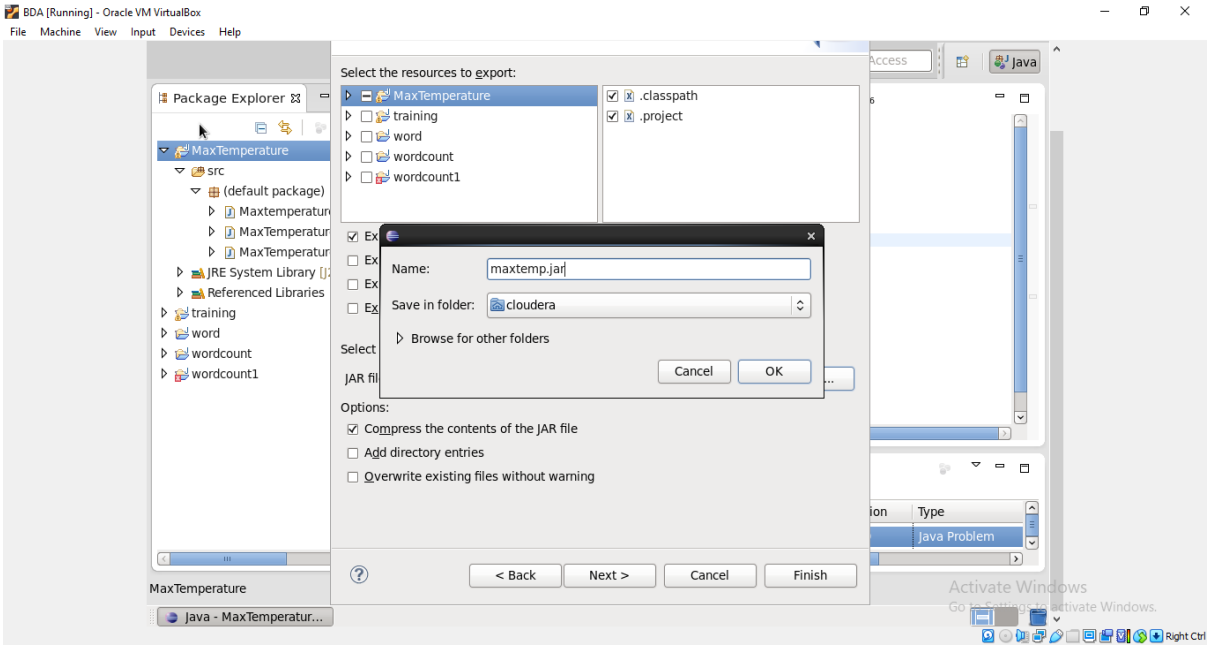
After finishing all the programs, right click on your project and select export



Select java-> JAR file->next



Click on browse and a dialog box will appear, give any name for the jar file and click ok.

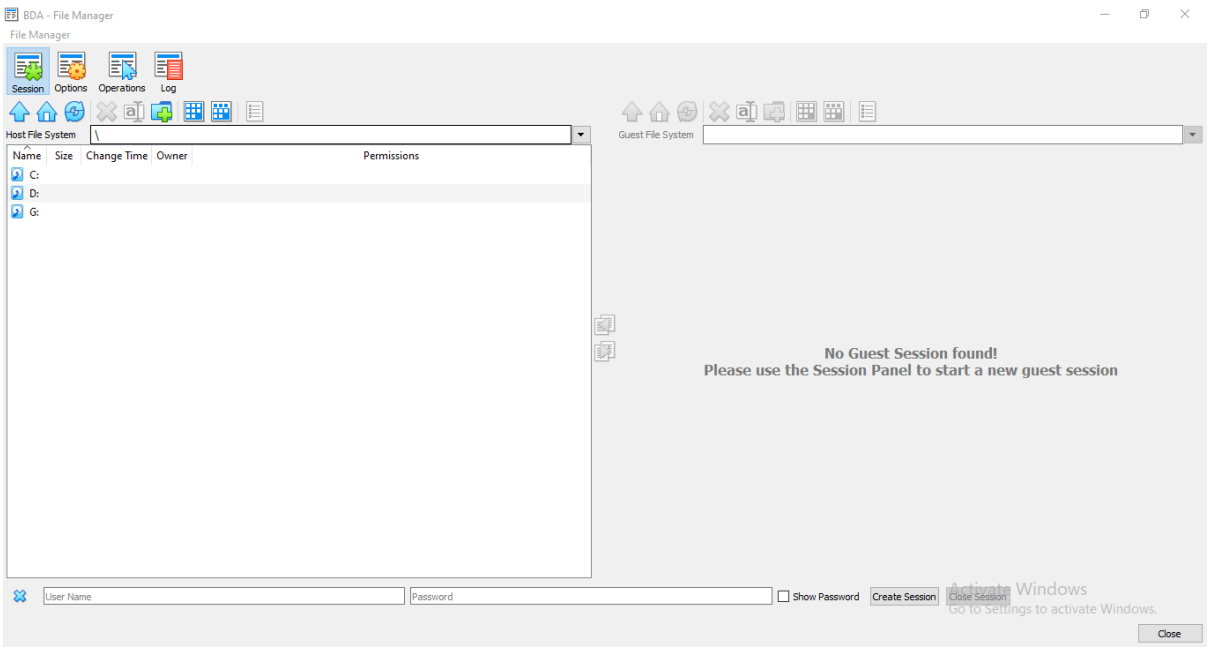


Click on finish.

Now for the input data file ,we need to copy the file from windows operating system(host) to linux operating system<guest>

Click on machine→file manager on the cloudera

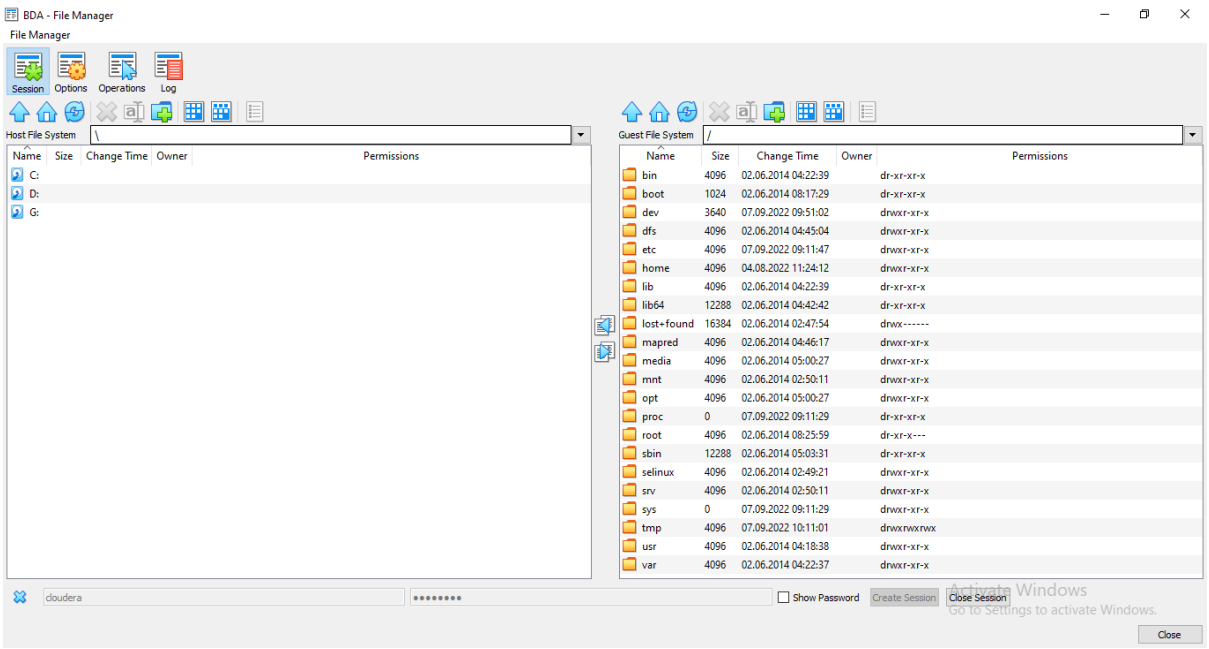
A window appears like



Now create a session

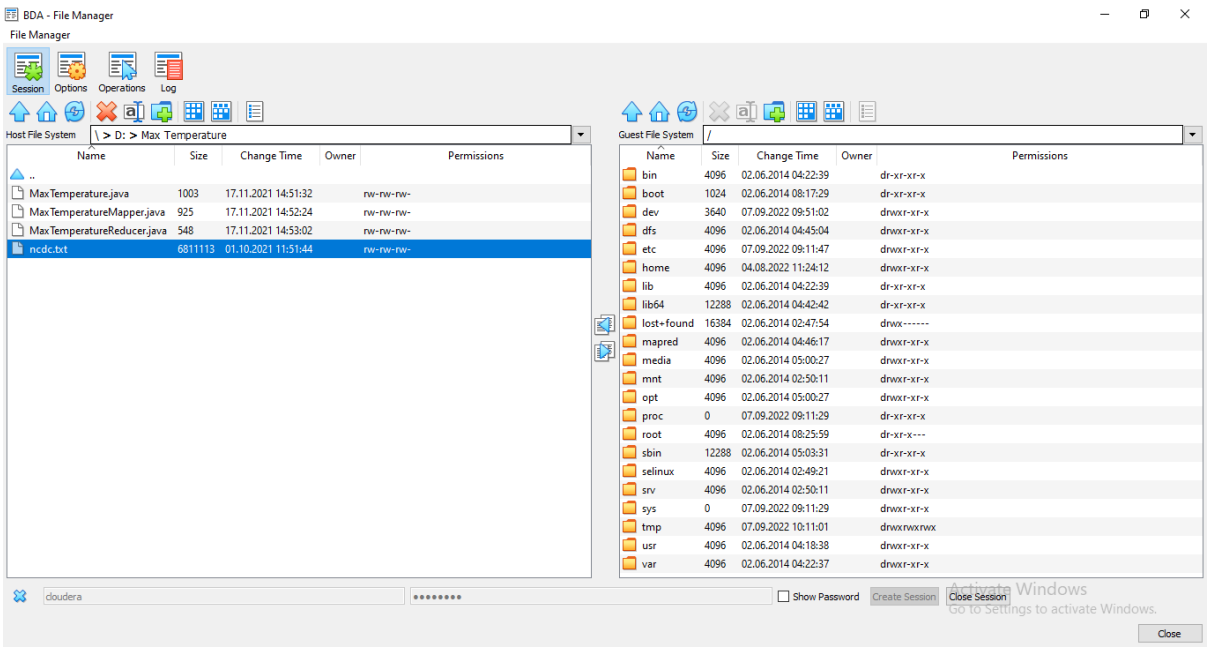
Username:cloudera

Password:cloudera



On the left side we have host(windows) and on the right side guest(linux)

Now select on left side d drive→Maxtemperature→ncdc.txt

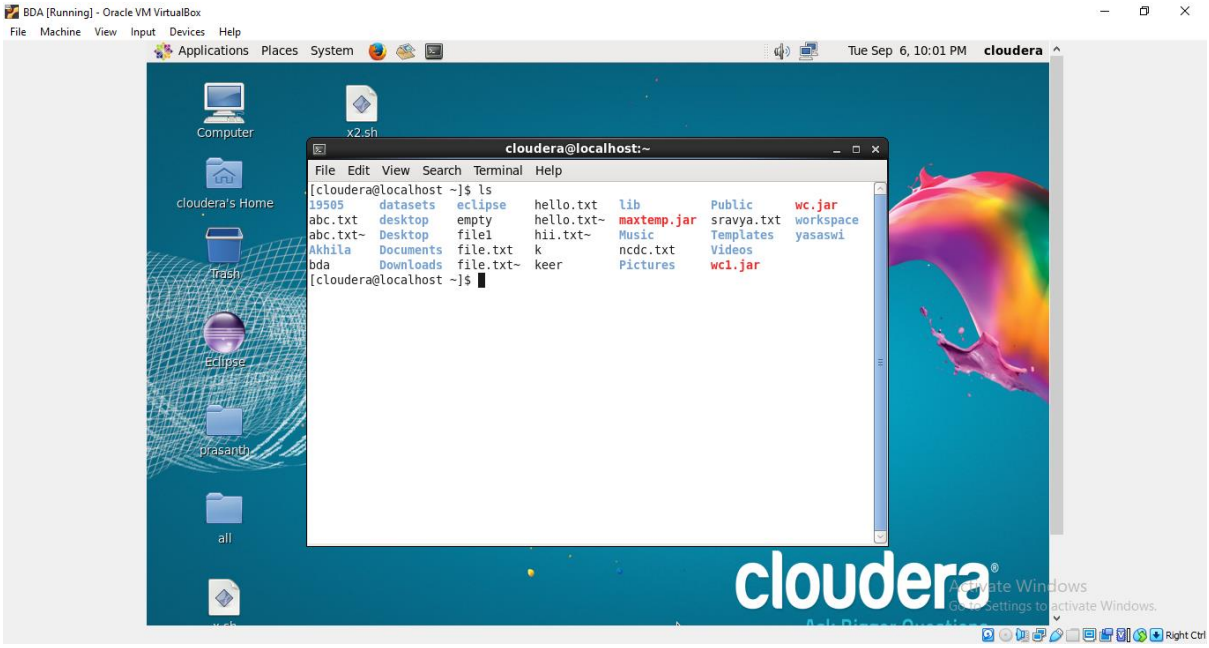


On right side click home→cloudera

Copy the ncdc.txt and close the session

Open terminal in the cloudera

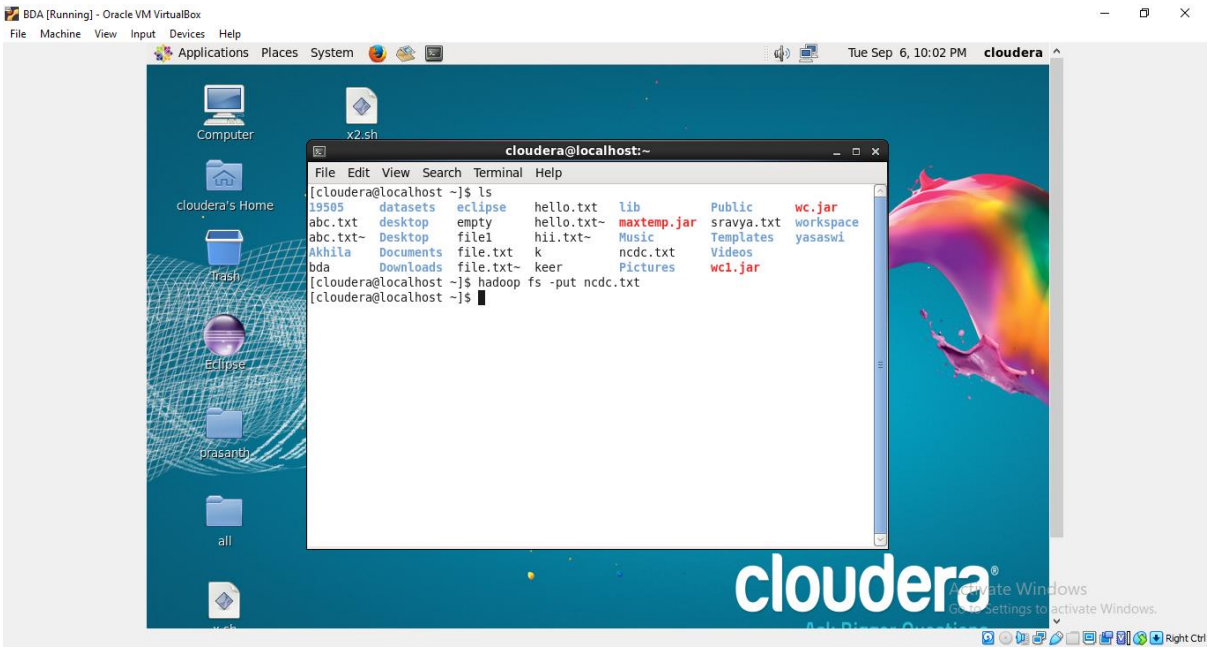
Run the command ls to check the jar file and and the input file ncdc.txt



We need to copy the file from linux os to hdfs

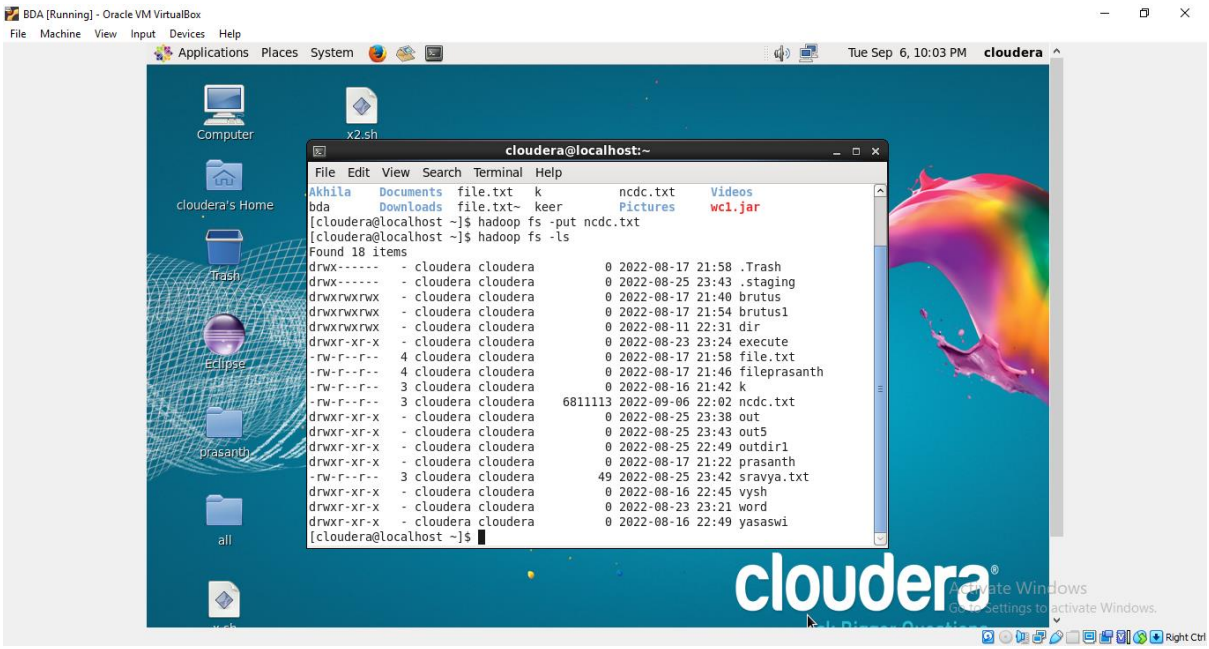
Run the command

hadoop fs –put ncdc.txt



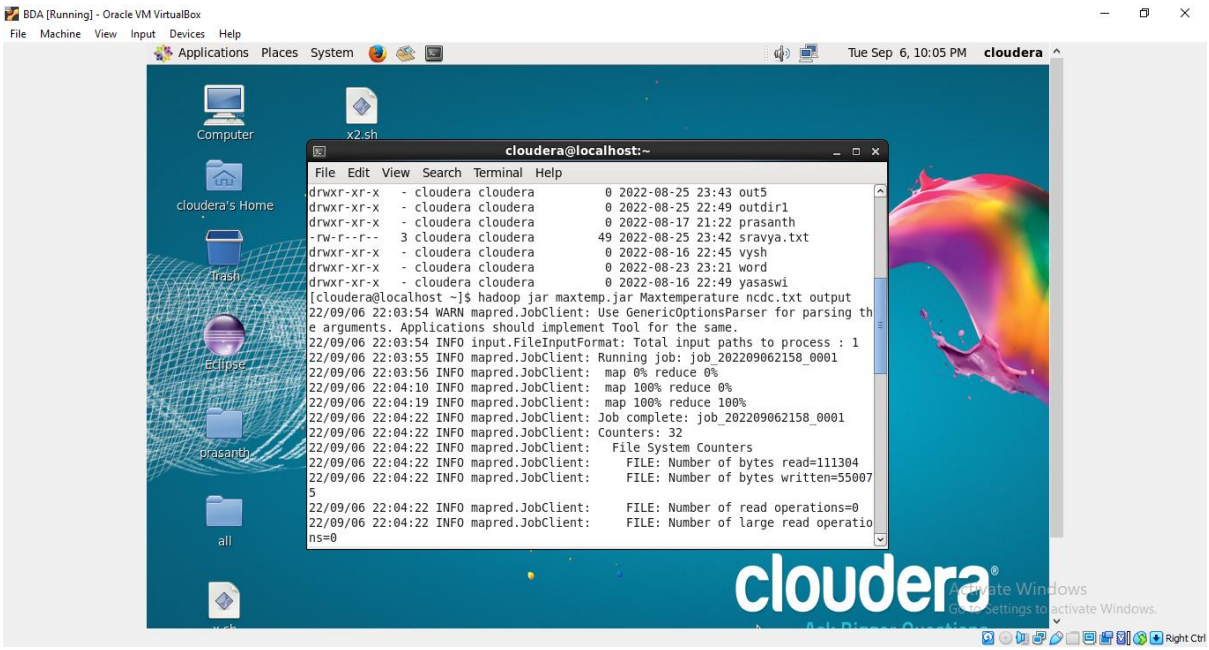
Check the hdfs for the ncdc.txt

Hadoop fs –ls



Now execute the command

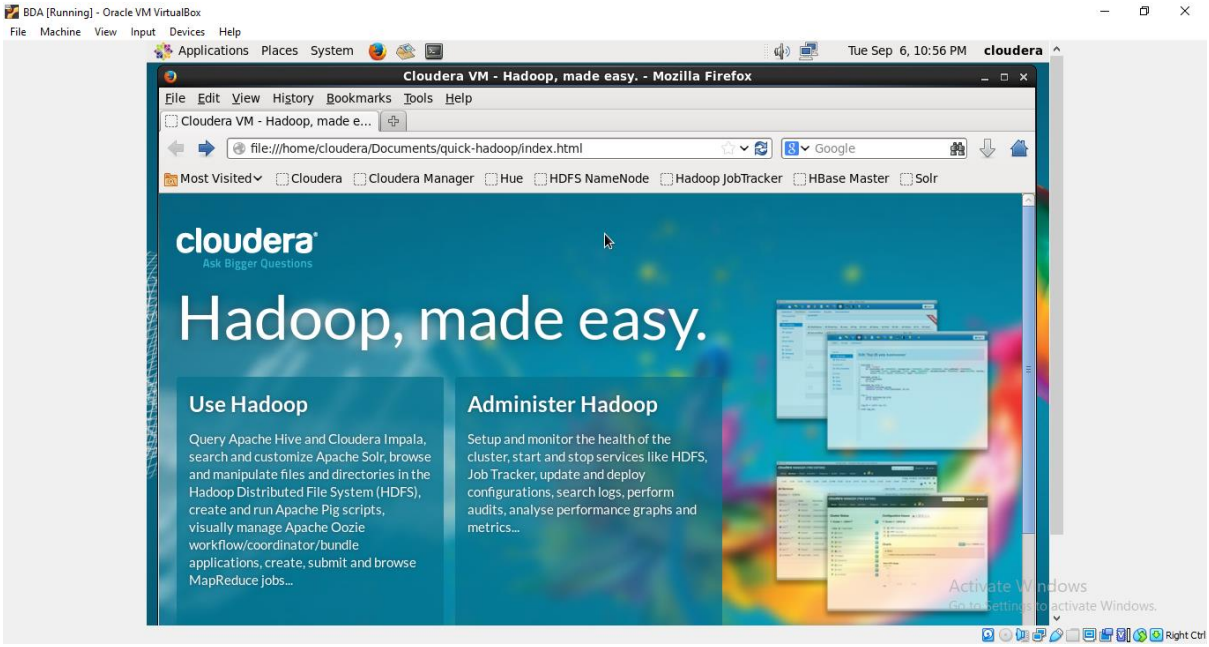
Hadoop jar <jar file name> <mainclassname> inputfile outputdirectoryname



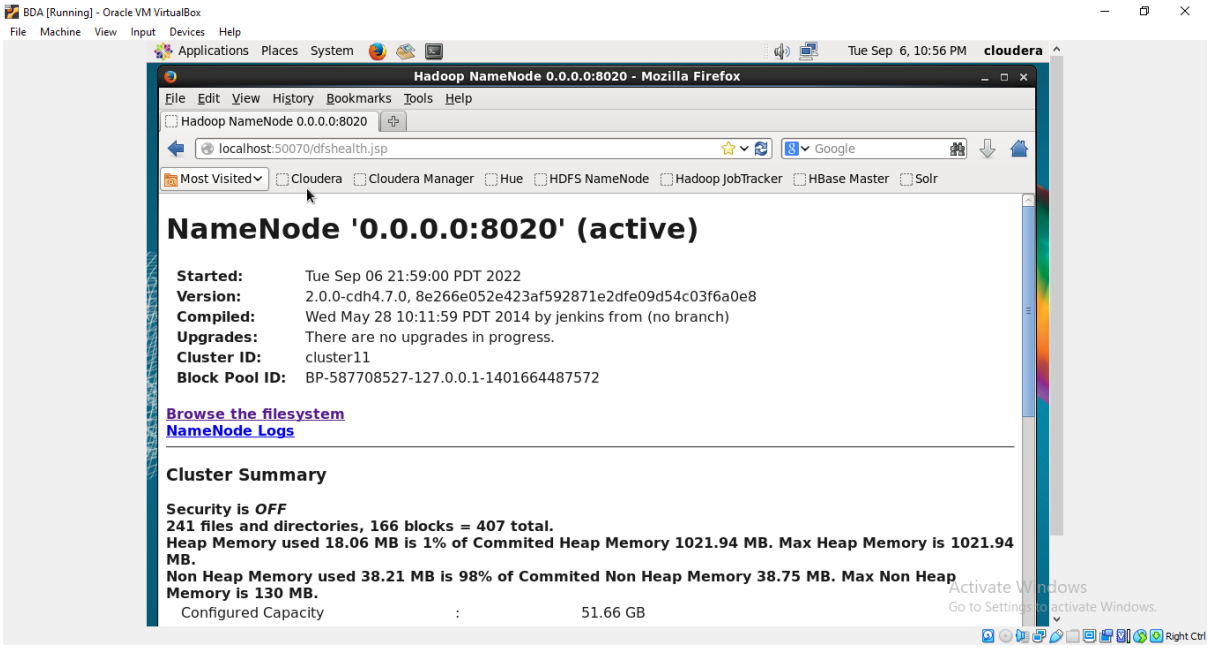
OUTPUT:

CHECKING USING BROWSER

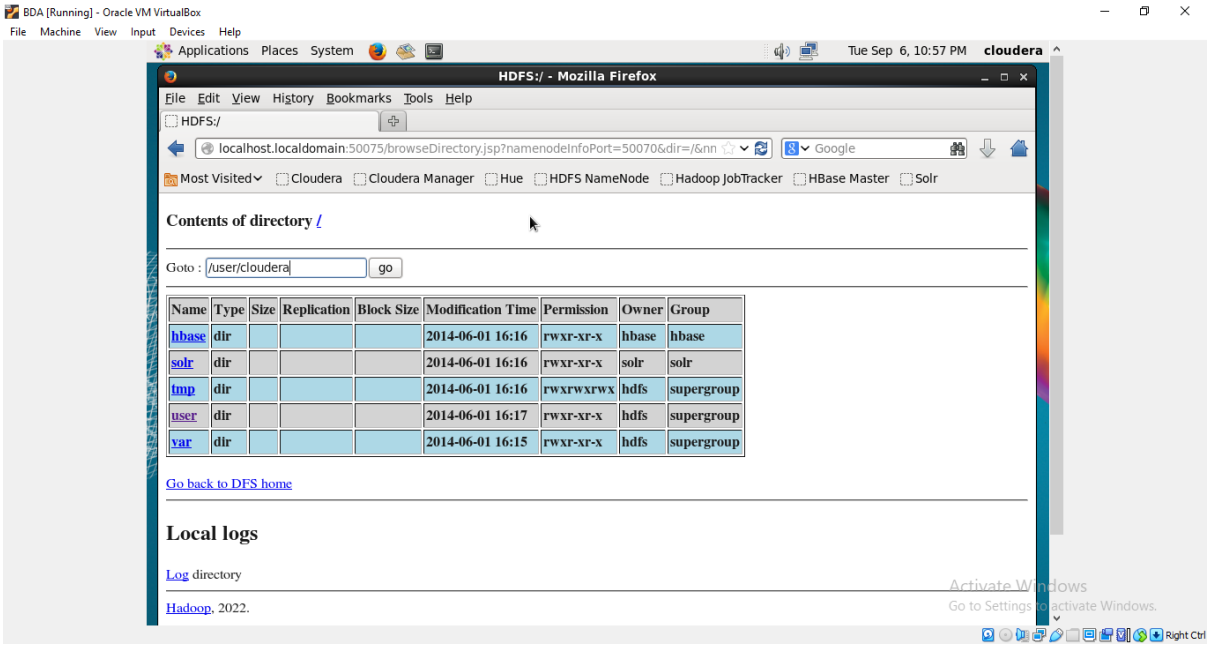
Open mozila browser→click HDFS namenode



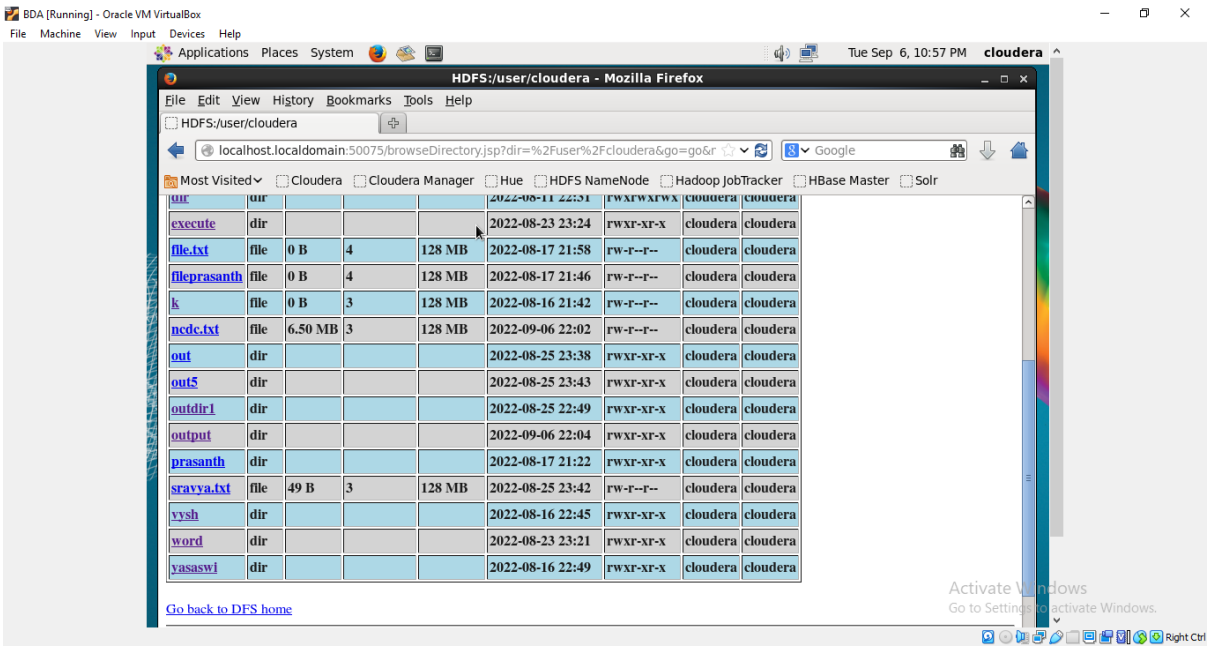
Click on browse file system



Goto user→cloudera

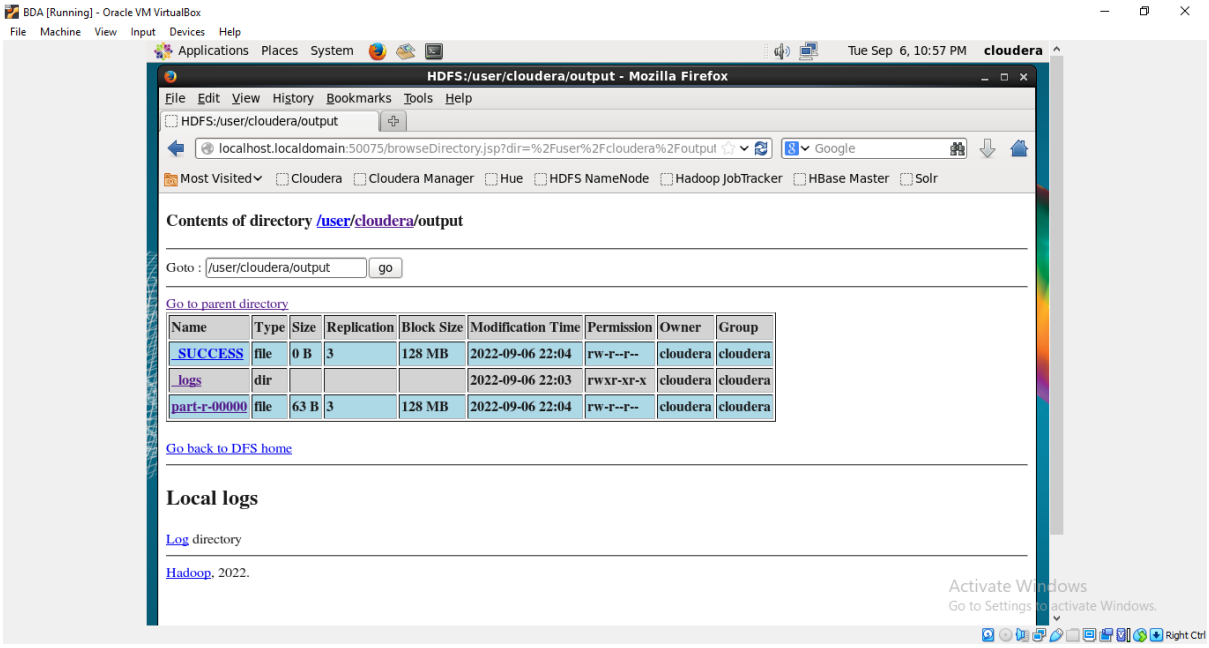


Open the output directory that you created



It contains the following files

Click part-r-00000



The output is displayed