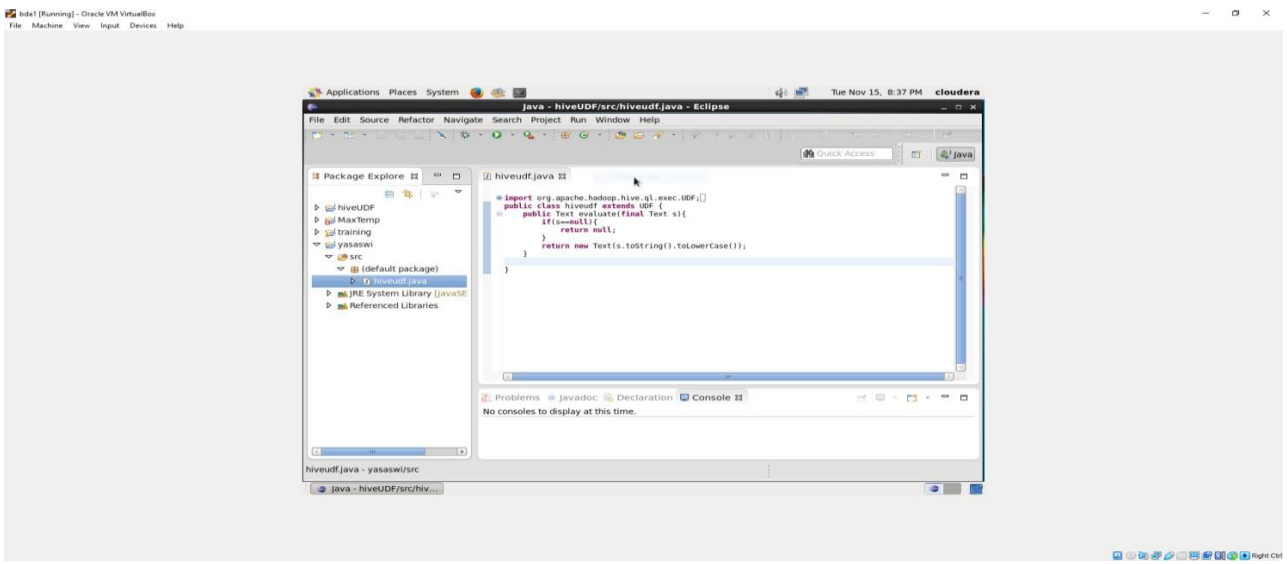
Experiment-12 *RollNo:-*

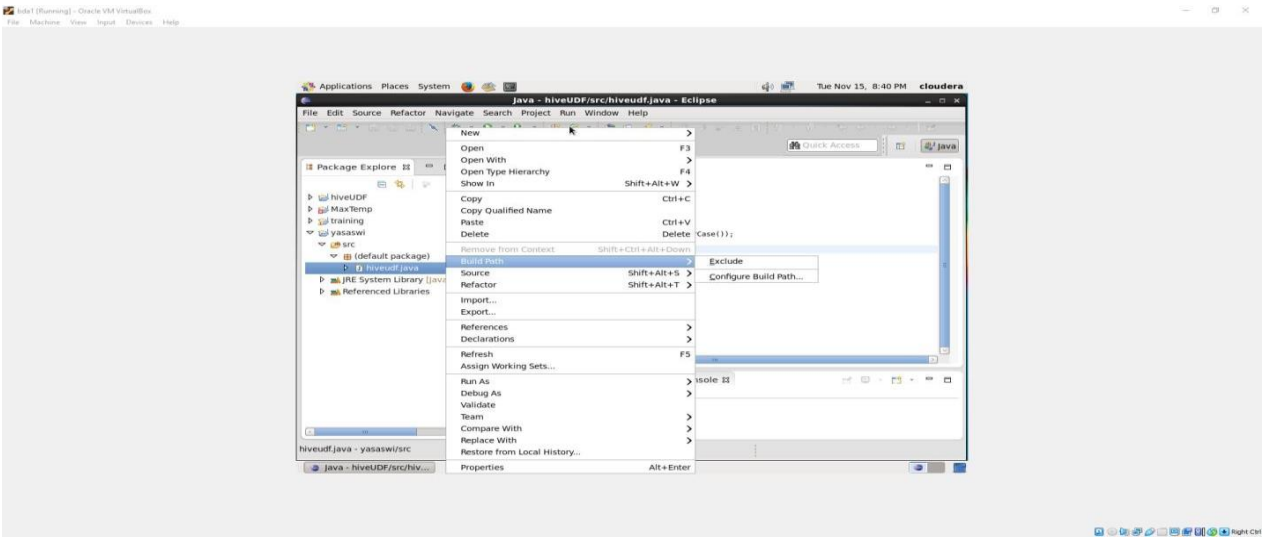
Hive udf

- Open vm virtual box and then eclipse
- Create a new java project and a new java class
- Copy the following code

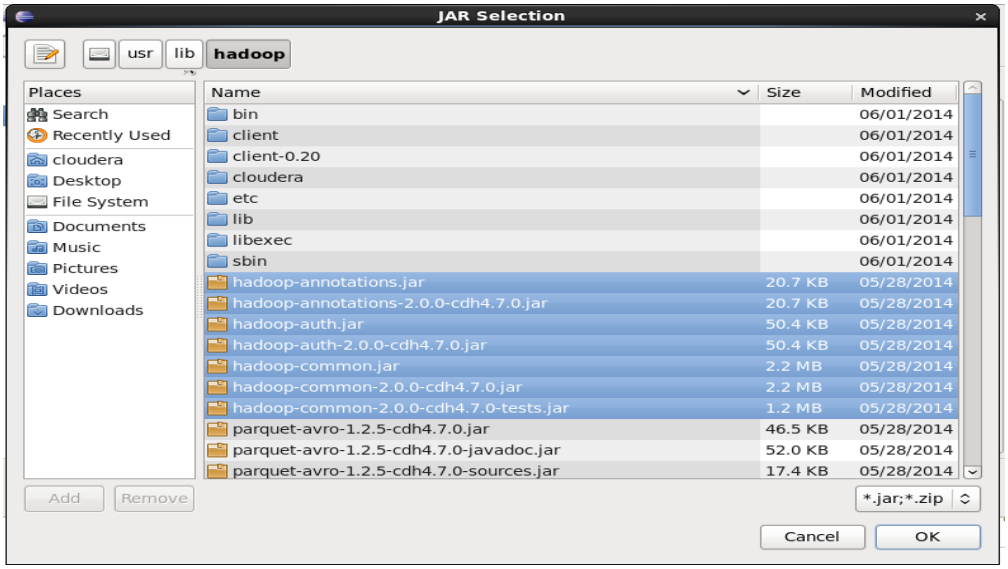
```
import org.apache.hadoop.hive.ql.exec.UDF;
import org.apache.hadoop.io.Text;
public class hiveudf extends UDF {

        public Text evaluate (final Text s ){
                if(s==null){
                        return null;

                }

        return new Text(s.toString().toLowerCase());

        }

}
```



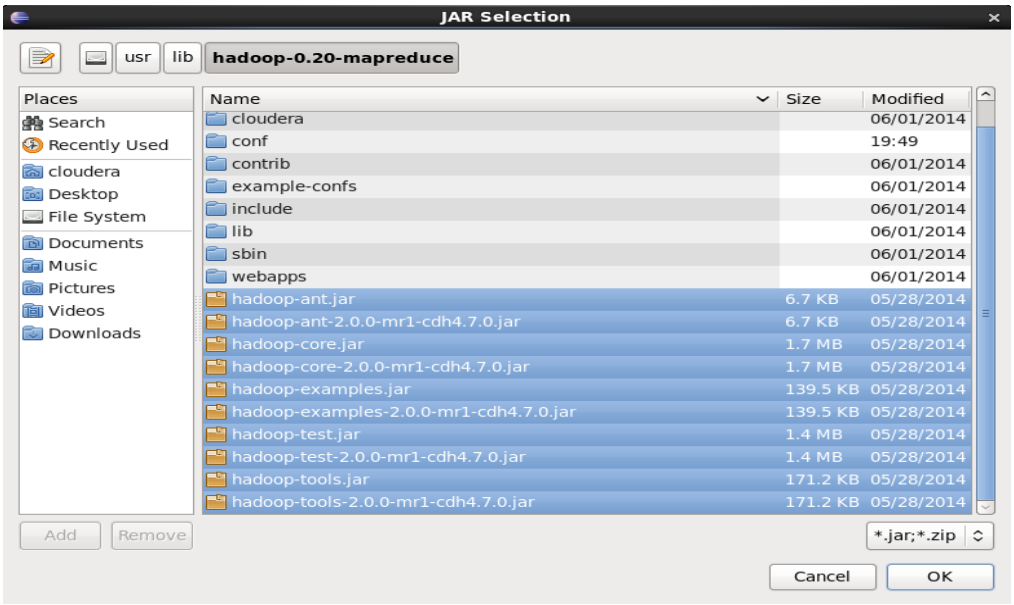
- Now we need to add external jar files.
- Rightclick->Build Path ->Configure Biuld Path….



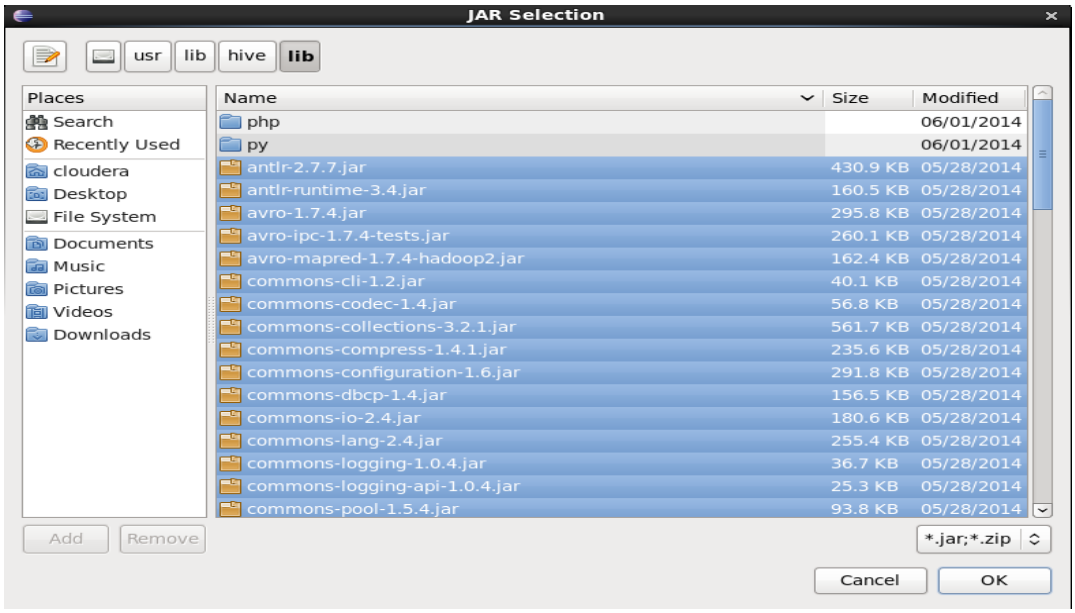
- **Now click on libraries-> add external jars -> Hadoop**



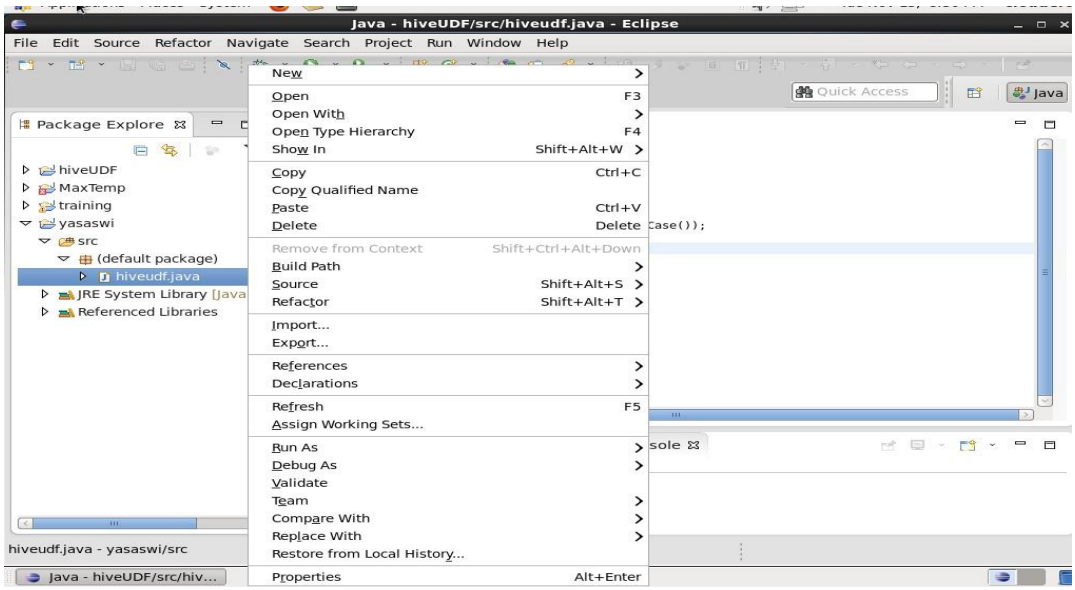
- Now click on **-> hadoop0.20 mapreduce folder**
  add hadoop jar files



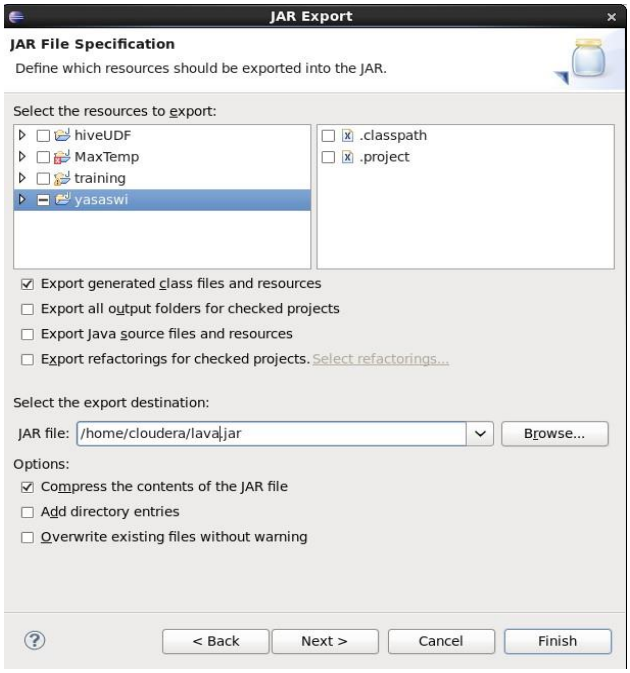
- Click on **lib->hive->lib……..** now select all the jar files in the lib folder.



- Export the jar file by clicking **Export…**

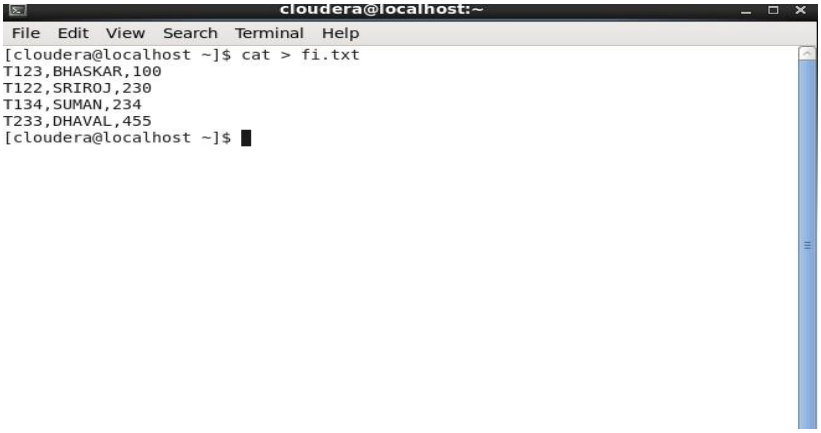


- Give name to the jar file in path directory **"/home/cloudera/example.jar"**
- And click on finish
- The jar file is exported



- Open the terminal and create a text file and copy the data.(otherwise you can add the data of your own in the following format only).

  T123,BHASKAR,100
  T122,SRIROJ,230
  T134,SUMAN,234
  T233,DHAVAL,455



- Create a hive file in the **vi** editor(hive file extension is **'.q'**). Now copy and paste the code in the hive file.(Note: click **'I'** to paste the code in the vi editor . To close the vi editor click **esc**
  ':wq' and click on enter)

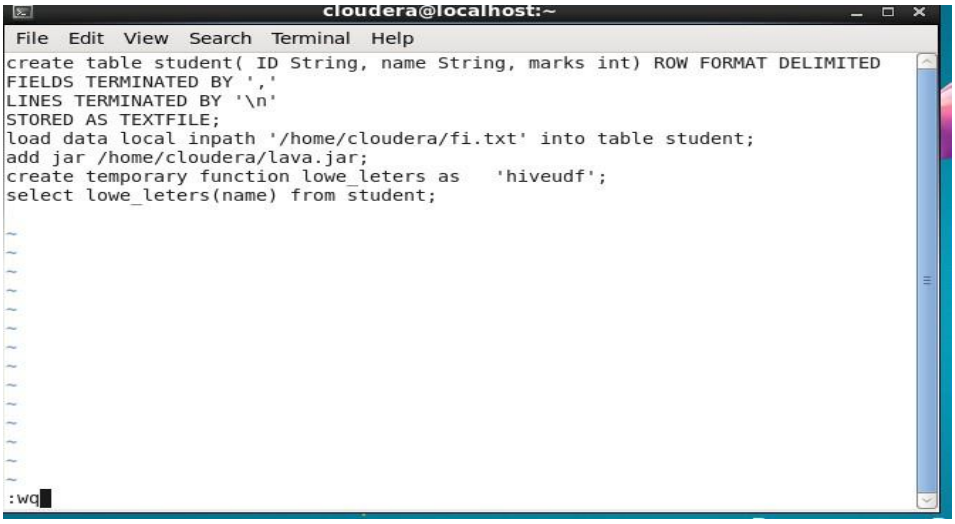
create table student(ID String, name String, marks int) ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','

LINES TERMINATED BY '\n'
STORED AS TEXTFILE;
load data local inpath '/home/cloudera /fi.txt' into table student;

add jar /home/cloudera /hivefunction.jar;

create temporary function lowe_leters as 'hiveudf';
select lowe_leters(name) from student;

- Change the highlighted words.



Now we need to execute the hive file using the command ' hive –f filename.q '.

output is :

2022-11-08 02:44:21,243 Stage-1 map = 0%, reduce = 0%
2022-11-08 02:44:32,352 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.08 sec
2022-11-08 02:44:33,386 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.08 sec
2022-11-08 02:44:34,409 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.08 sec
2022-11-08 02:44:35,428 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.08 sec
2022-11-08 02:44:36,447 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.08 sec
2022-11-08 02:44:37,468 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.08 sec
2022-11-08 02:44:38,490 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.08 sec
2022-11-08 02:44:39,514 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.08 sec
2022-11-08 02:44:40,533 Stage-1 map = 100%, reduce = 0%, Cumulative CPU 2.08 sec
2022-11-08 02:44:41,555 Stage-1 map = 100%, reduce = 100%, Cumulative CPU 2.08 sec MapReduce Total cumulative CPU time: 2 seconds 80 msec
Ended Job = job_202211080215_0002 MapReduce Jobs Launched:

Job 0: Map: 1 Cumulative CPU: 2.08 sec HDFS Read: 290 HDFS Write: 28 SUCCESS Total MapReduce CPU Time Spent: 2 seconds 80 msec
OK
bhaskar sriroj suman dhaval
Time taken: 38.087 seconds.