**Writing a Pig UDF**

**Problem statement:** Write a UDF to convert lowercase string to uppercase.

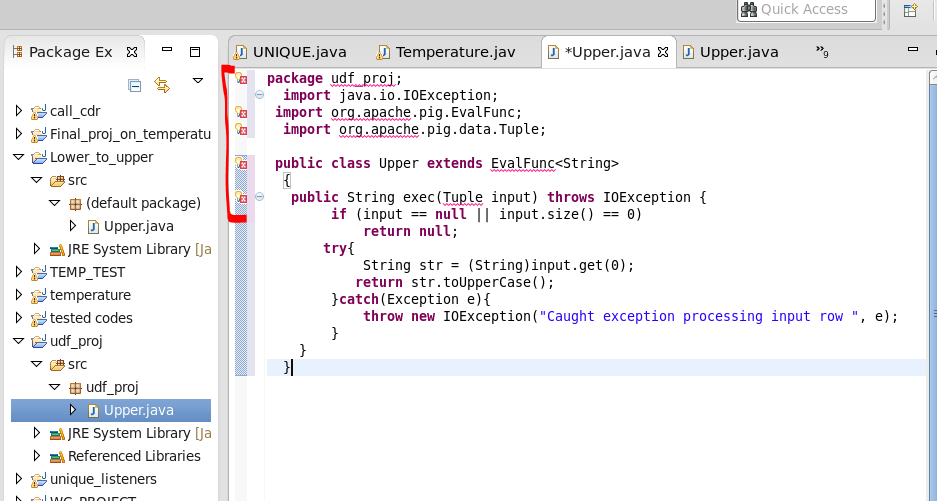
**Link for Jar file to be added in the code:** https://drive.google.com/file/d/0Bxr27gVaXO5sY1pMdjVuZUJxa2M/view?usp=sharing

**Link for Java code for converting lowercase to uppercase:**

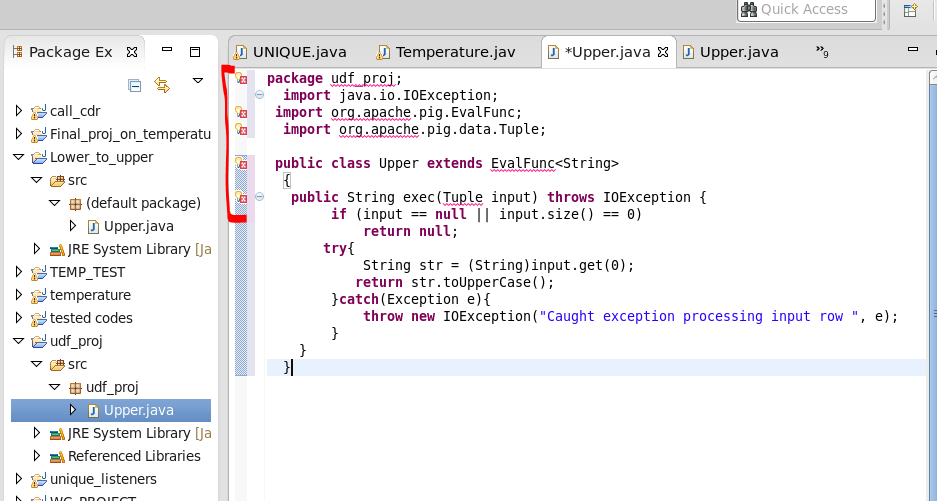
https://drive.google.com/file/d/0Bxr27gVaXO5saHZWaUd1SHJiSUE/view?usp=sharing

**Step 1:** Write the java program for converting the lowercase string to uppercase in eclipse.

Refer the below screenshot:

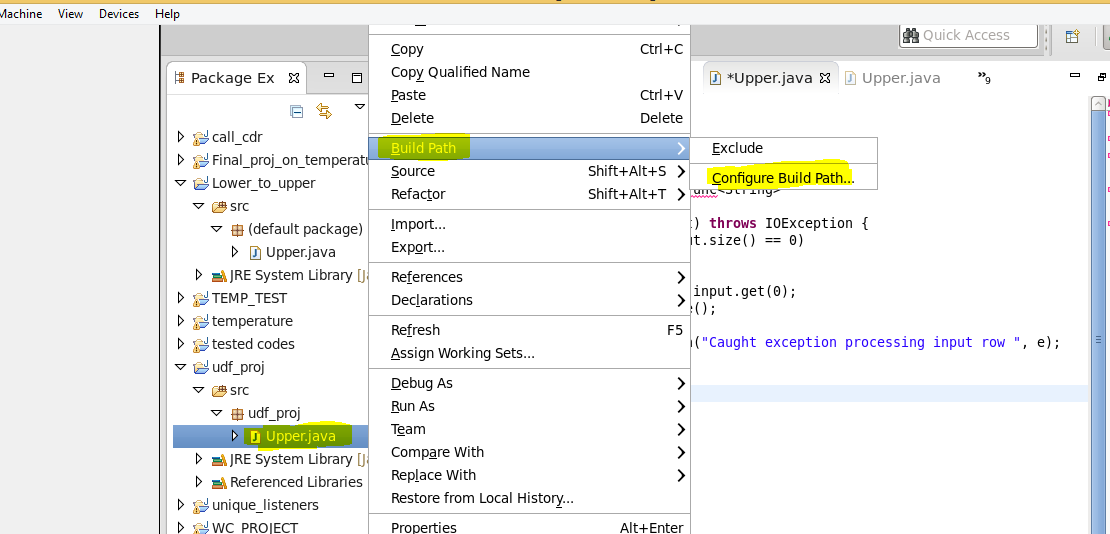


**Step 2:** To remove the error shown in the above sscreenshot add the jar, refer the below screenshot to add the Jar file.

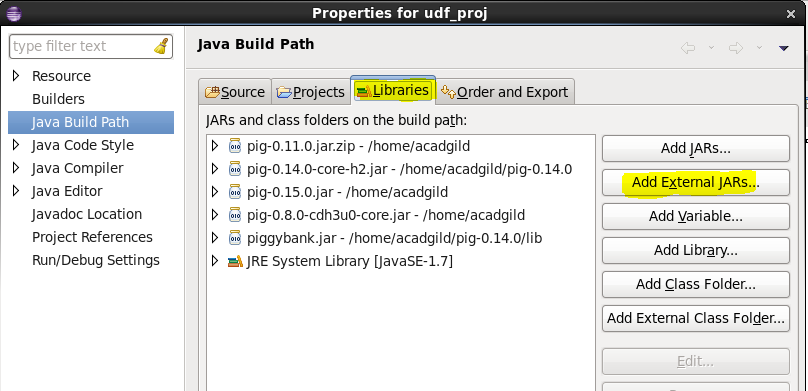


**3.** Steps to add the Jar file:

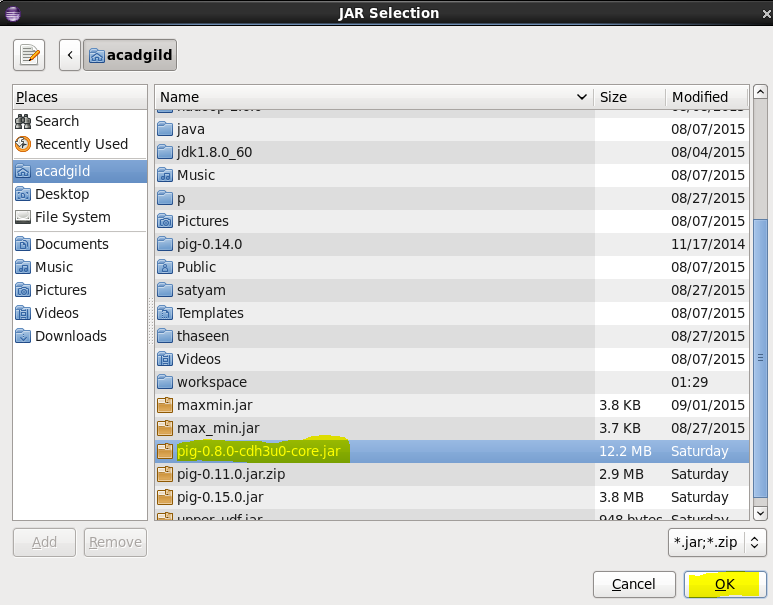
**3.1** Right click on the program -----> Build Path ----🡪 Configure Build path



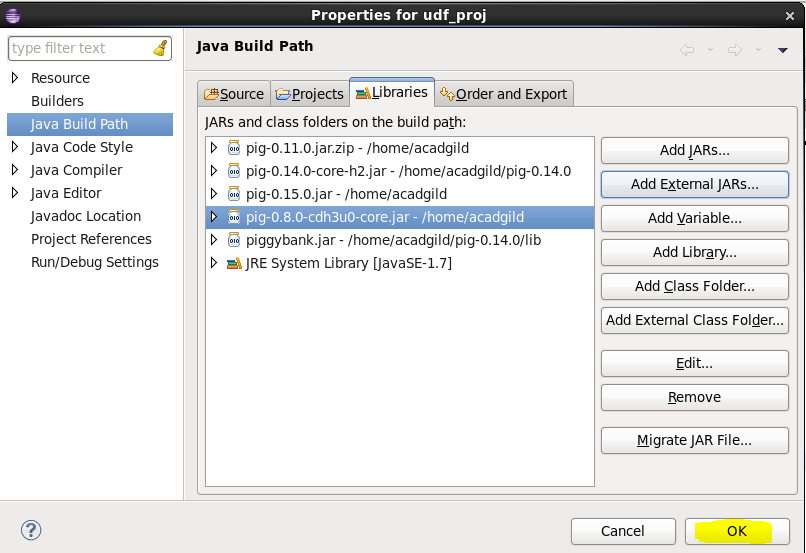
**3.2** Click on libraries ---🡪 Add external Jars



**3.3** Browse to the location where pig jar is present and then click on OK.

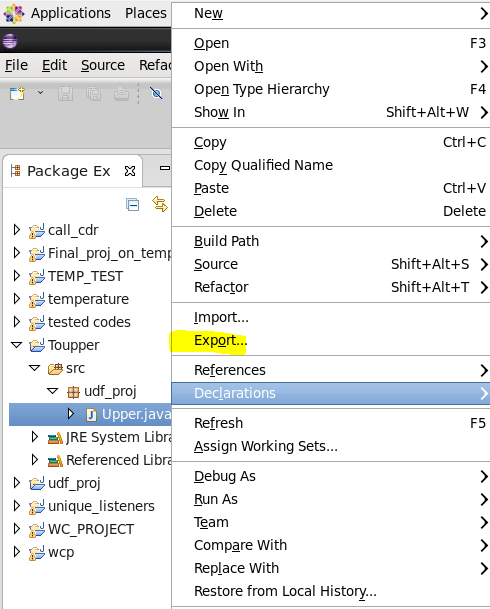


**3.4** Click on ok shown in the below screenshot to finish the importing of the Jar.

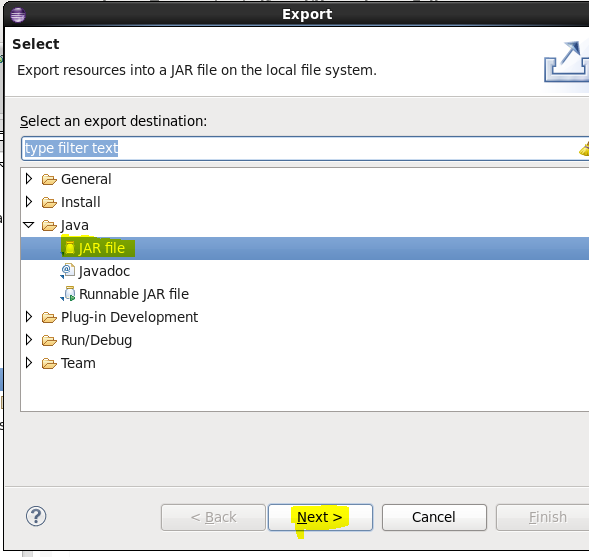


**Step 4**. Exporting of JAR file.

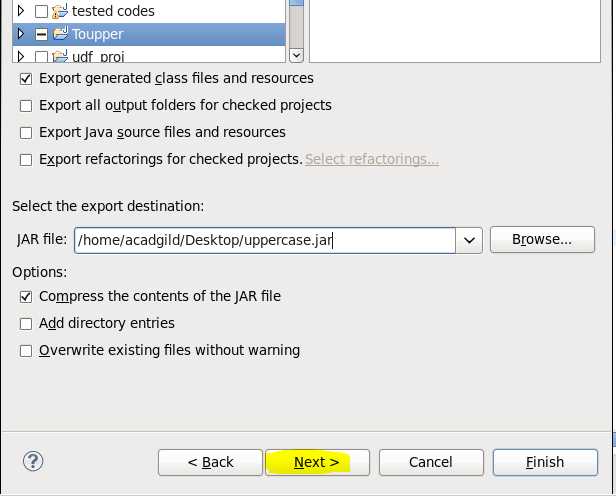
**4.1** Right click on the program file ---🡪 click on export



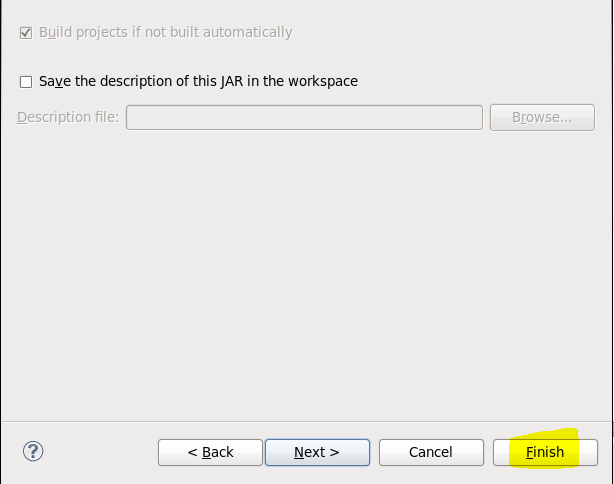
**4.2** Select the Jar file and then click on Next.



Browse to the location where you want to store the Jar file and then click on Next.



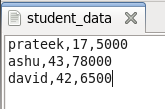
**4.3** Click on Finish to complete the exporting of Jar file.



**Step 5.** Open the grunt shell and then write the below Pig Scripts.

**Note:** Take a dummy comma separated dataset with three columns respectively as name, age and gpa.

Refer the below screenshot for the dataset.



Copy the dataset into hdfs.

**Command:** hadoop dfs -put student\_data /

**Scripts for Pig:**

**REGISTER /home/acadgild/uppercase1.jar //Register the Jar file which was exported from eclipse.**

**A = LOAD '/student\_data' using PigStorage(‘,’) AS (name: chararray, age: int, gpa: float); //Dataset is loaded into the relation A**

**B = FOREACH A GENERATE udf\_proj.UPPER(name); //udf\_proj is the fully qualified name of the package and UPPER is the name of the class.**

**DUMP B;**

