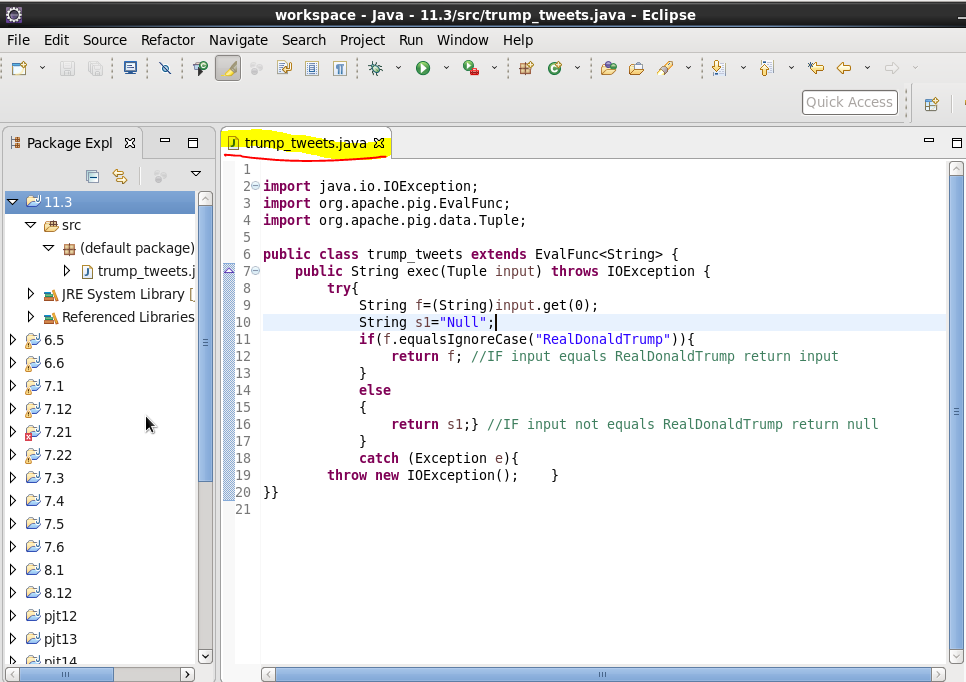
**Assignment – 11.3**

****

****

Udf code :



In the user define function for checking To find the tweets posted by @realDonaldTrump from the given dataset. We will extend the evaluation function that will take string as argument.We will check if that argument is equal to RealDonaldTrump or not,if it is equal than it will return the input that is RealDonaldTrump otherwise it will return null.and when we will run the pig script we will remove the null part and count the remaining number of RealDonaldTrump.

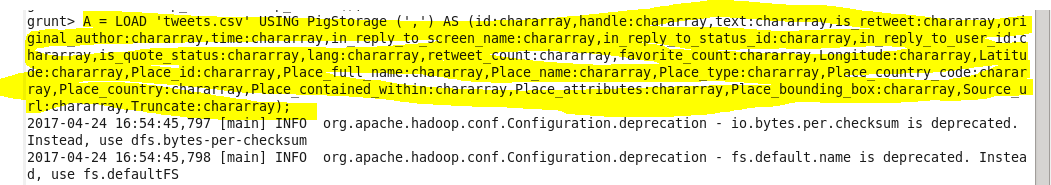
**Pig code :**

First we will register our UDF jar file and after that we will define our class that is trump\_tweets

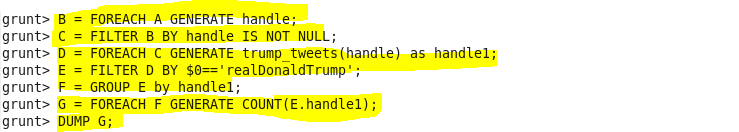


Now we can write our pig script.

* First we will load our tweet.csv file using LOAD command.

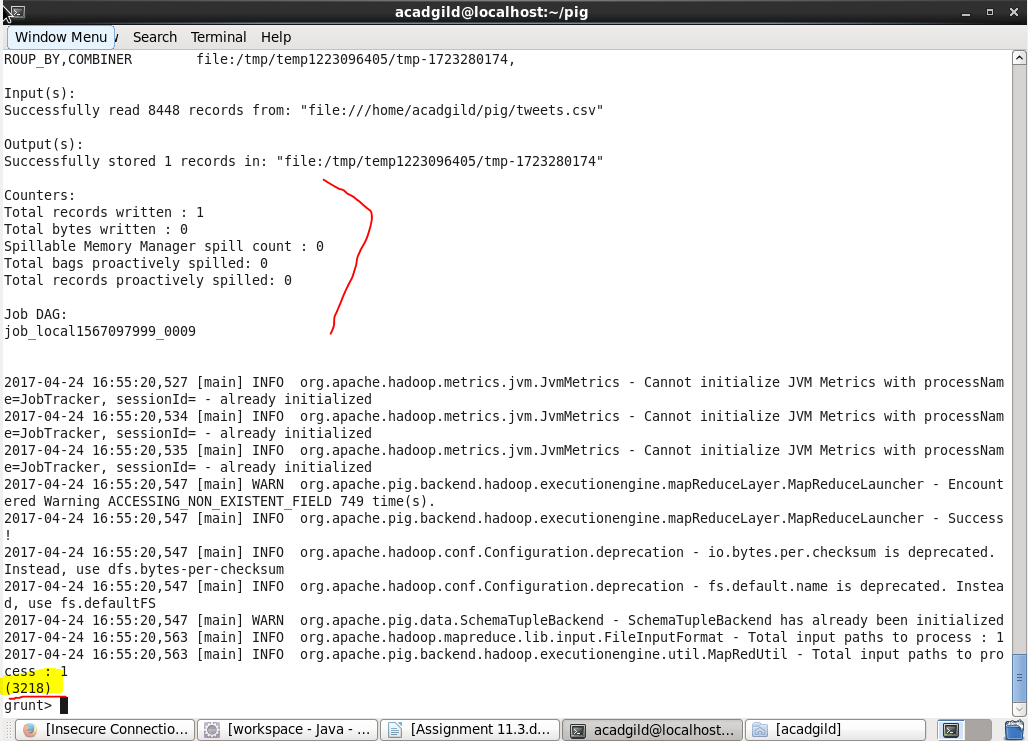


Now for our loaded data we will generate handle and then we will filter out our data where handle is null. Then we will use our user define function that is trump\_tweets and will pass the handle as argument ,after that we will check if it is RealDonaldTrump or not for that we will filter our data by the handle equal to realdonaldtrump and after that we will generate the count of handle that is total number of realdonaldtrump using COUNT function.



OUTPUT :

So after running the above pig script we got the following as result.



So total 3218 tweets from RealDonaldTrump.