EX.NO. 4 **210701521**

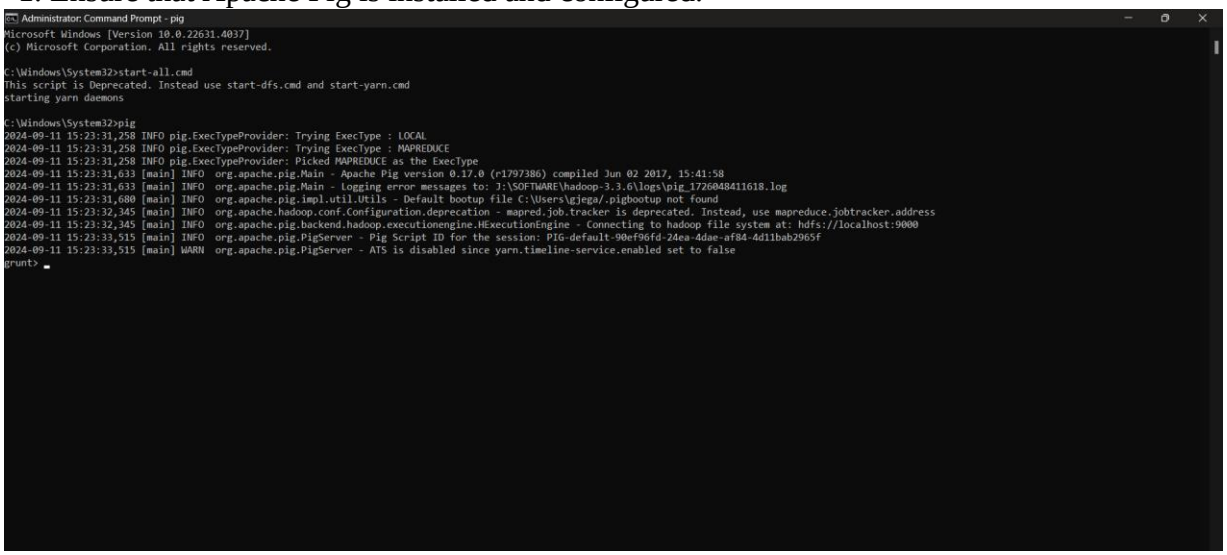
Create UDF (User Defined Functions) in Apache Pig and execute it in MapReduce/HDFS mode

AIM:

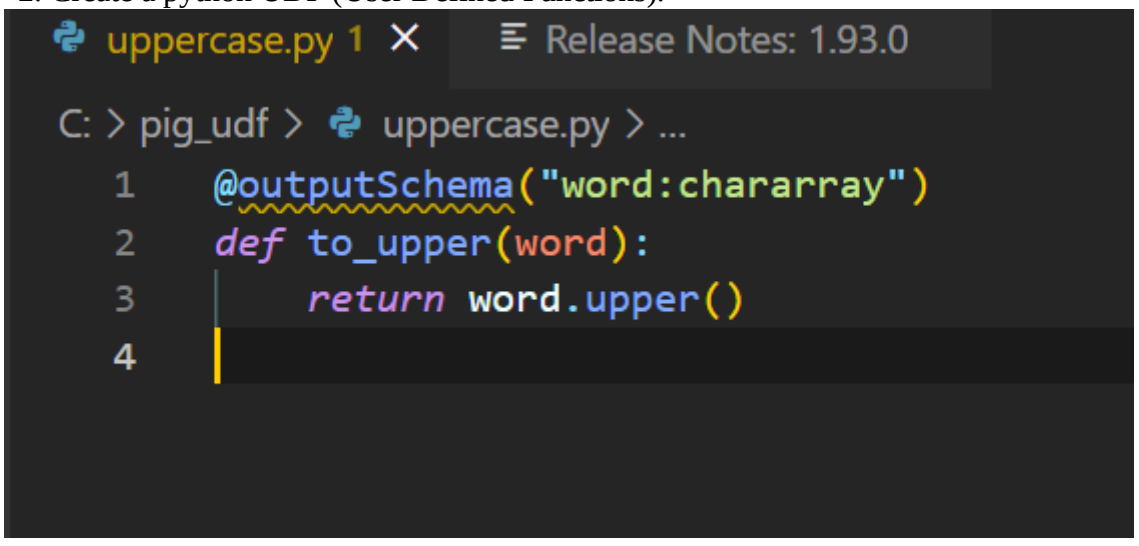
To create UDF (User Defined Functions) in Apache Pig and execute it in MapReduce/HDFS mode.

PROCEDURE:

1. Ensure that Apache Pig is installed and configured.



2. Create a python UDF (User Defined Functions).



3. Jython should be installed as Pig will use it to interpret the Python UDFs.

4. Create a Pig script that registers and uses the Python UDF.

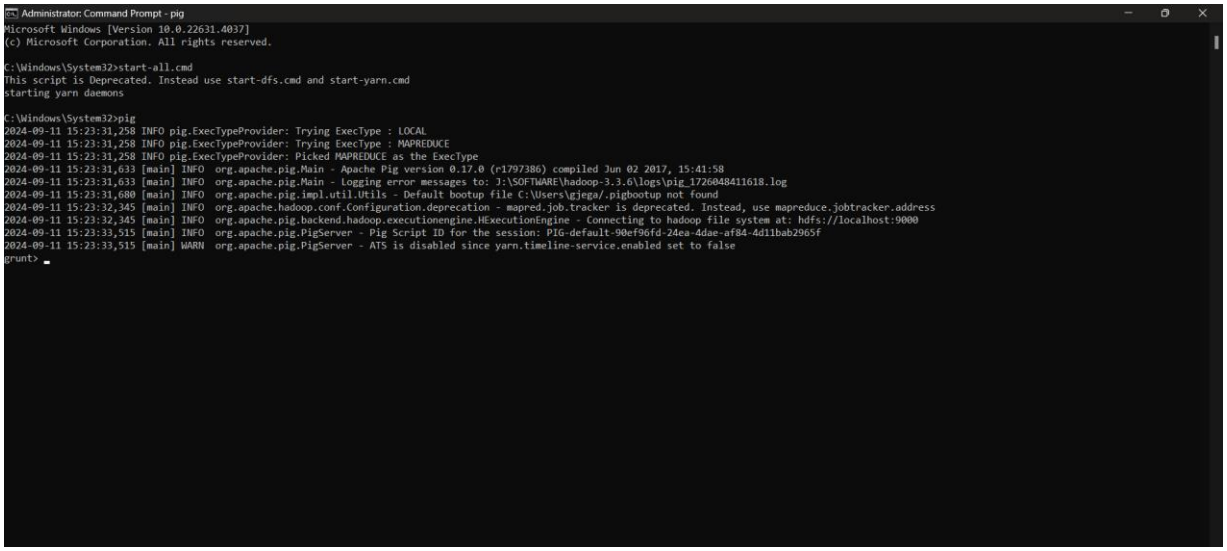
C: > pig_udf > script.pig

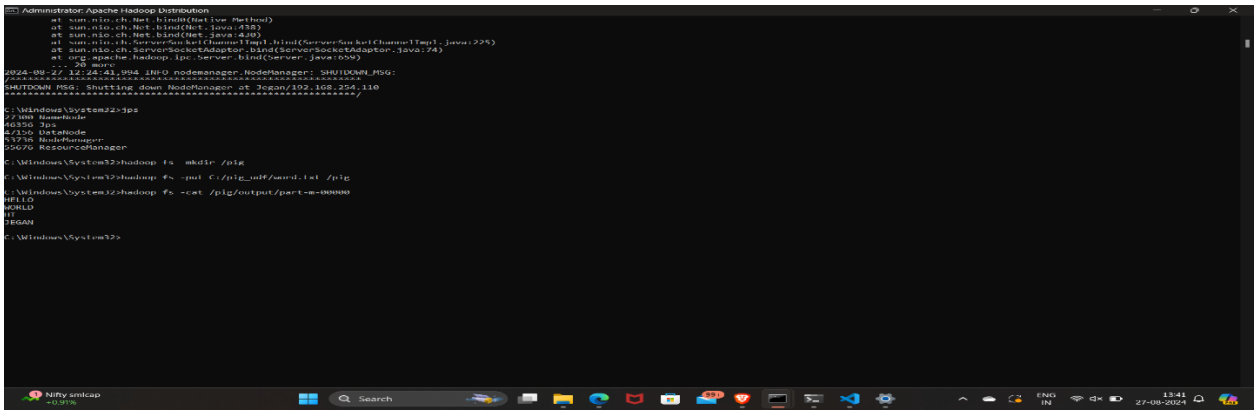
```
-- Register the Python UDF script
REGISTER 'uppercase.py' USING jython AS myudf;

-- Load the input file from HDFS
data = LOAD 'hdfs:///pig/word.txt' USING PigStorage(',') AS (line: chararray);

-- Apply the UDF to convert each line to uppercase
uppercased_data = FOREACH data GENERATE myudf.to_upper(line);

-- Store the result in HDFS
STORE uppercased_data INTO 'hdfs:///pig/output' USING PigStorage(',');
```

5. Execute the Pig Script in MapReduce Mode using the command:

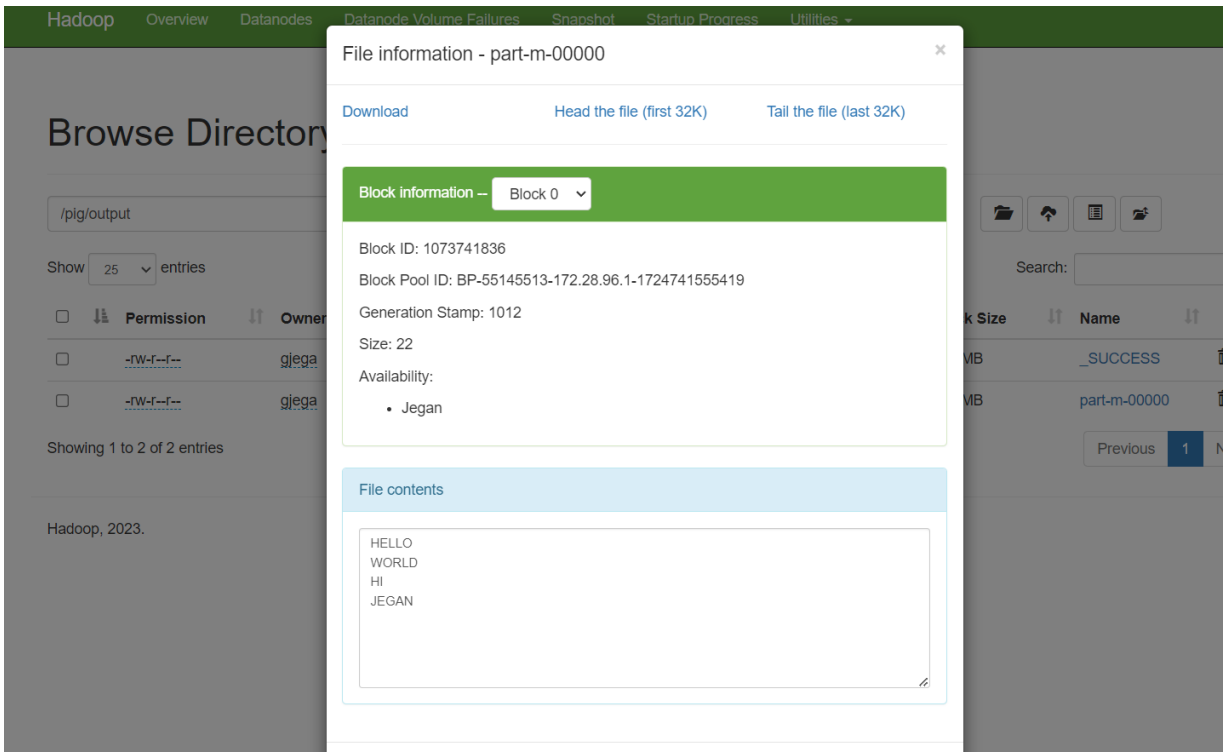
   pig -x mapreduce script.pig

OUTPUT:





EX.NO. 4 **210701521**



RESULT:

Thus, to create a UDF in Apache Pig and execute in MapReduce mdoe has been executed successfully