

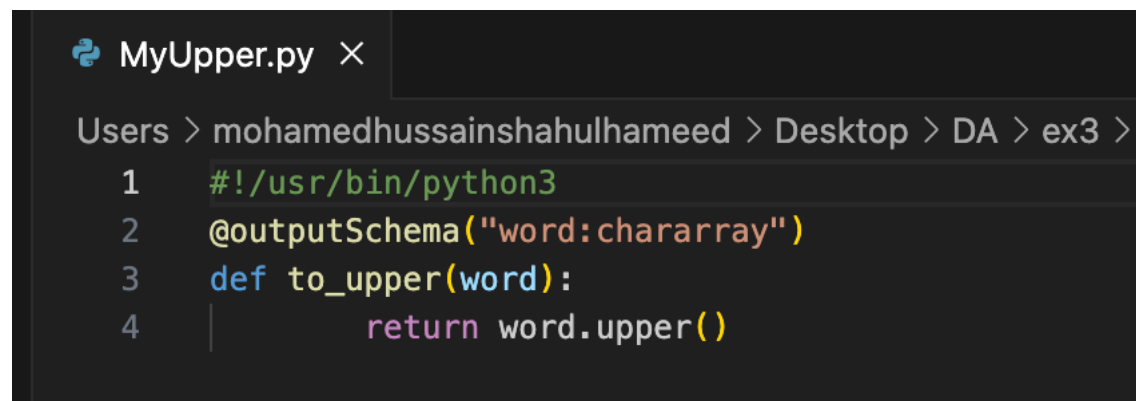
Create UDF (User Defined Functions) in Apache Pig and execute it in MapReduce/HDFS mode

AIM:

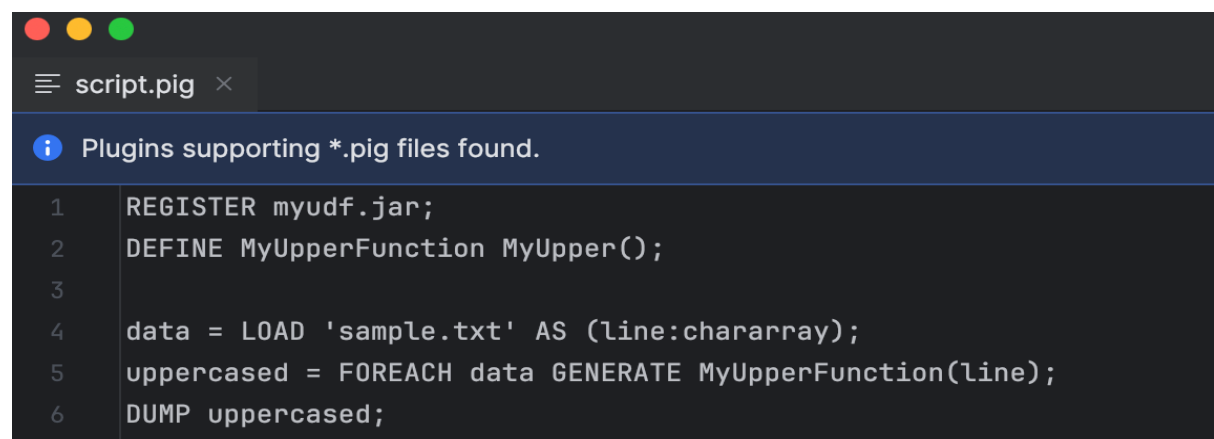
To create user defined functions in Apache Pig and execute it in MapReduce / HDFS mode.

PROCEDURE:

1. Write the Python UDF : Created a Python function that reverses strings.
2. Register the UDF in Pig : Registered the Python script in the Pig script using the REGISTER command.
3. Use the UDF in Pig Script : Applied the UDF to the data set using FOREACH ... GENERATE.
4. Execute the Pig Script : Ran the Pig script in Map Reduce mode to process the data on HDFS.

OUTPUT:

```
MyUpper.py ×
Users > mohamedhussainshahulhameed > Desktop > DA > ex3 >
1  #!/usr/bin/python3
2  @outputSchema("word:chararray")
3  def to_upper(word):
4      return word.upper()
```



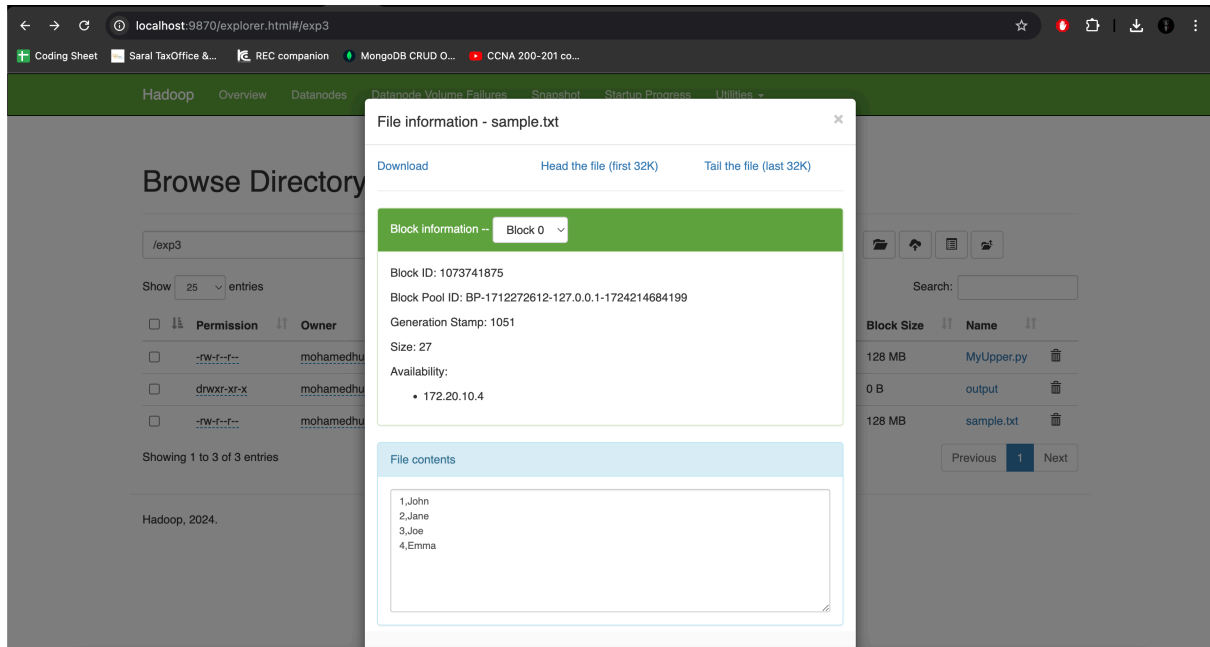
```
script.pig ×
Plugins supporting *.pig files found.
1  REGISTER myudf.jar;
2  DEFINE MyUpperFunction MyUpper();
3
4  data = LOAD 'sample.txt' AS (line:chararray);
5  uppercased = FOREACH data GENERATE MyUpperFunction(line);
6  DUMP uppercased;
```

```

mohamedhussainshahulhameed@Mohameds-Laptop ~ % pig -x local udf_example.pig

SLF4J: Failed to load class "org.slf4j.impl.StaticLoggerBinder".
SLF4J: Defaulting to no-operation (NOP) logger implementation
SLF4J: See http://www.slf4j.org/codes.html#StaticLoggerBinder for further details.
2024-09-03 14:41:52,247 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
2024-09-03 14:41:52,248 INFO pig.ExecTypeProvider: Picked LOCAL as the ExecType
2024-09-03 14:41:52,307 [main] INFO org.apache.pig.Main - Apache Pig version 0.17.0 (r1797386) compiled Jun 02 2017, 15:41:58
2024-09-03 14:41:52,307 [main] INFO org.apache.pig.Main - Logging error messages to: /Users/mohamedhussainshahulhameed/pig_1725354712304.log
2024-09-03 14:41:52,472 [main] ERROR org.apache.pig.Main - ERROR 2997: Encountered IOException. File udf_example.pig does not exist
Details at logfile: /Users/mohamedhussainshahulhameed/pig_1725354712304.log
2024-09-03 14:41:52,484 [main] INFO org.apache.pig.Main - Pig script completed in 395 milliseconds (395 ms)

```



```

mohamedhussainshahulhameed@Mohameds-Laptop ~ % hadoop fs -cat /exp3/output/part-m-00000
SLF4J: Failed to load class "org.slf4j.impl.StaticLoggerBinder".
SLF4J: Defaulting to no-operation (NOP) logger implementation
SLF4J: See http://www.slf4j.org/codes.html#StaticLoggerBinder for further details.
1,JOHN
2,JANE
3,JOE
4,EMMA
mohamedhussainshahulhameed@Mohameds-Laptop ~ %

```

RESULT:

Thus Create UDF (User Defined Functions) in Apache Pig and execute it in MapReduce/HDFS mode has been executed successfully.

