

**EXP NO: 4****CREATE UDF IN PIG****\$start-all.sh**

```

(hadoop@kali)~$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [kali]
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-11 04:59:16,429 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting resourcemanager
Starting nodemanagers

```

**\$ jps**

```

(hadoop@kali)~$ jps
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
14436 NodeManager
10772 Jps
13830 SecondaryNameNode
14211 ResourceManager
13597 DataNode
13471 NameNode

```

**\$wget <https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz>****\$ tar xvfz pig-0.16.0.tar.gz**

```

kali-linux-2023.4-vmware-amd64 - VMware Workstation 17 Player (Non-commercial use only)
Player
(hadoop@kali)~$ wget https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz
--2024-08-29 10:55:41-- https://dlcdn.apache.org/pig/pig-0.16.0/pig-0.16.0.tar.gz
Resolving dlcdn.apache.org (dlcdn.apache.org) ... 208.146.42.132, 151.101.2.132
Connecting to dlcdn.apache.org (dlcdn.apache.org)|208.146.42.132|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 17727933 (169M) [application/x-gzip]
Saving to: 'pig-0.16.0.tar.gz'

pig-0.16.0.tar.gz 100%[=====>] 169.07M 26.5MB/s in 6.55s

2024-08-29 10:55:41 (26.0 MB/s) - 'pig-0.16.0.tar.gz' saved [17727933/17727933]

(hadoop@kali)~$ tar xvfz pig-0.16.0.tar.gz
pig-0.16.0/
pig-0.16.0/bin/
pig-0.16.0/conf/
pig-0.16.0/contrib/
pig-0.16.0/contrib/piggybank/
pig-0.16.0/contrib/piggybank/java/
pig-0.16.0/contrib/piggybank/java/build/
pig-0.16.0/contrib/piggybank/java/build/classes/
pig-0.16.0/contrib/piggybank/java/build/classes/org/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/datetime/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/datetime/convert/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/datetime/diff/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/datetime/truncate/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/decode/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/match/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/stats/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/string/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/util/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/evaluation/util/apacheLogParser/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/allloader/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/apacheLog/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/avro/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/hiverc/
pig-0.16.0/contrib/piggybank/java/build/classes/org/apache/pig/piggybank/storage/partition/
pig-0.16.0/contrib/piggybank/java/build/docs/
pig-0.16.0/contrib/piggybank/java/build/docs/api/
pig-0.16.0/contrib/piggybank/java/build/test/
pig-0.16.0/contrib/piggybank/java/lib/
pig-0.16.0/contrib/piggybank/java/src/

```

**\$nano ~/.bashrc**

```

#PIG settings
export PIG_HOME=/home/hadoop/pig
export PATH=$PATH:$PIG_HOME/bin
export PIG_CLASSPATH=$PIG_HOME/conf:$HADOOP_HOME/etc/hadoop/
export PIG_CONF_DIR=$PIG_HOME/conf
#export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64
export PIG_CLASSPATH=$PIG_CONF_DIR:$PATH
#PIG setting ends

```

**\$mv pig-0.16.0 pig**

**\$pig**

```

kali-linux-2023.4-vmware-amd64 - VMware Workstation 17 Player (Non-commercial use only)
hadoop@kali: ~
File Actions Edit View Help
pig 0.16.0/tutorial/src/org/apache/pig/tutorial/tutorialUtil.java
pig 0.16.0/tutorial/pig
pig 0.16.0/bin/pig.cmd
pig 0.16.0/bin/pig.py

hadoop@kali: ~
$ mv pig-0.16.0 pig
hadoop@kali: ~
$ name ~/./bashrc
hadoop@kali: ~
$ source ~/./bashrc
hadoop@kali: ~
$ java
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
55024 ResourceManager
54024 SecondaryNameNode
55138 NodeManager
54393 Database
64993 JPS
54267 NameNode

hadoop@kali: ~
$ pig
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-08-29 18:58:45.221 [main] INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
2024-08-29 18:58:45.226 [main] INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE
2024-08-29 18:58:45.227 [main] INFO pig.ExecTypeProvider: Picked MAPREDUCE as the ExecType
2024-08-29 18:58:45.477 [main] INFO org.apache.pig.Main - Apache Pig version 0.16.0 (r1746530) compiled Jun 01 2016, 23:10:19
2024-08-29 18:58:45.477 [main] INFO org.apache.pig.Main - Logging error messages to: /home/hadoop/pig-1746432525450.log
2024-08-29 18:58:45.562 [main] INFO org.apache.pig.impl.util.Utils - Default hystrix file /home/hadoop/pigscript not found
2024-08-29 18:58:45.667 [main] WARN org.apache.hadoop.util.NativeCodeLoader - Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2024-08-29 18:58:46.721 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2024-08-29 18:58:46.723 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-29 18:58:46.724 [main] INFO org.apache.pig.backend.hadoop.executionengine.HExecutionEngine - Converting to Hadoop file system at: hdfs://localhost:9000
2024-08-29 18:58:48.975 [main] INFO org.apache.pig.PigServer - Pig Script ID for the session: PIG-default-5c28e2d3-7330-47ba-a30b-c0b6274bfc2
2024-08-29 18:58:48.975 [main] WARN org.apache.pig.PigServer - ATS is disabled since yarn.timeline-service.enabled set to false
grunts> quit
2024-08-29 18:59:00.426 [main] INFO org.apache.pig.Main - Pig script completed in 15 seconds and 392 milliseconds (15392 ms)

hadoop@kali: ~

```

**\$cd DA-Lab**

**\$mkdir exp4**

**\$cd exp4**

**\$nano sample.txt**

```

hadoop@kali: ~/DA-Lab/exp4
File Actions Edit View Help
GNU nano 7.2 sample.txt
1, John
2, Jane
3, Joe
4, Emma

```

**\$nano demo\_pig.pig**

```

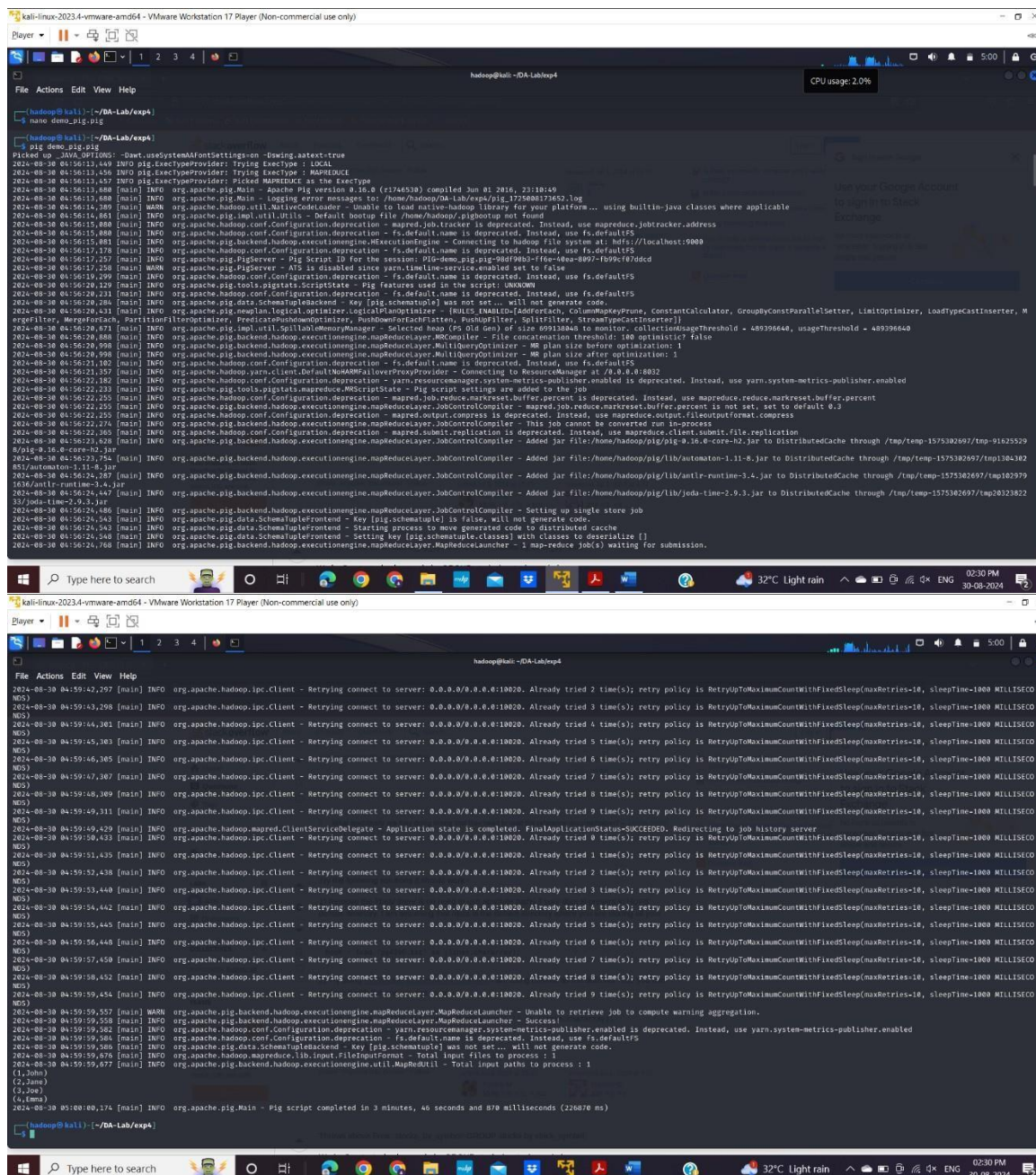
hadoop@kali: ~/DA-Lab/exp4
File Actions Edit View Help
GNU nano 7.2 demo_pig.pig
-- Load the data from HDFS
data = LOAD '/exp4/sample.txt' USING PigStorage(',') AS (id:int, name:chararray);
-- Dump the data to check if it was loaded correctly
DUMP data;

```

**\$hdfs dfs -mkdir /exp4**

**\$hdfs dfs -copyFromLocal ~/DA-Lab/exp4/sample.txt /exp4**

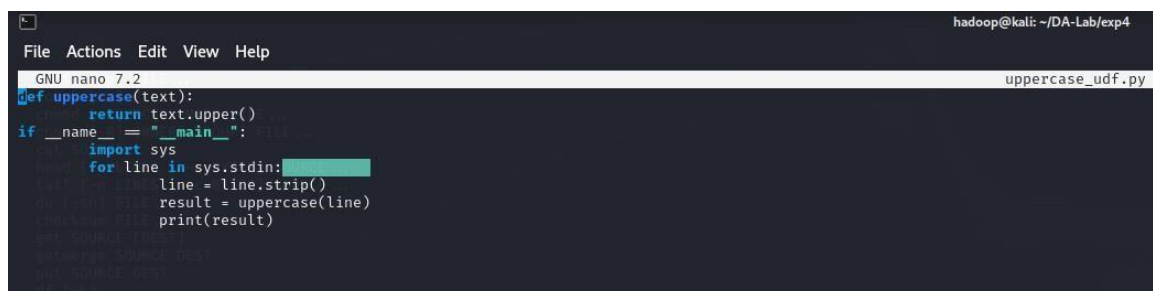
## \$pig demo\_pig.pig



```

kali@kali:~/DA-Lab/exp4$ pig demo_pig.pig
Picked up JAVA_OPTIONS: -Dant.useSystemAfontSettings -Dswing.aatext=true
2024-08-30 04:56:12.649 INFO org.apache.pig.ExecTypeProvider: Trying ExecType: LOCAL
2024-08-30 04:56:13.456 INFO org.apache.pig.Main: Apache Pig version 0.16.0 (r1744538) compiled Jun 01 2016, 23:10:49
2024-08-30 04:56:13.680 INFO org.apache.pig.Main: Logging error messages to: /home/hadoop/DA-Lab/exp4/pig-175080617652.log
2024-08-30 04:56:15.209 INFO org.apache.hadoop.util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2024-08-30 04:56:16.861 INFO org.apache.pig.impl.util.Utils: Default bootstrap file /home/hadoop/pigbootstrap not found
2024-08-30 04:56:15.880 INFO org.apache.hadoop.conf.Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2024-08-30 04:56:15.880 INFO org.apache.pig.backend.hadoop.executionengine.MRConfiguration: fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 04:56:15.881 INFO org.apache.pig.backend.hadoop.executionengine.MRConfiguration: Connecting to Hadoop file system at: hdfs://localhost:9000
2024-08-30 04:56:17.178 INFO org.apache.pig.PigServer: Pig Script ID for the session: PIG-demo_pig.pig-98d9803-f6e-4897-f696-f6dcd
2024-08-30 04:56:17.257 INFO org.apache.pig.PigServer: AJS is disabled since yarn.timeline-service.enabled set to false
2024-08-30 04:56:18.221 INFO org.apache.hadoop.conf.Configuration.deprecation: fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 04:56:18.129 INFO org.apache.pig.tools.pigstats.ScriptState: Pig features used in the script: UNKNOWN
2024-08-30 04:56:18.221 INFO org.apache.pig.data.SchemaTableBackend: fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 04:56:18.284 INFO org.apache.pig.data.SchemaTableBackend: Key [pig.schema.table] was not set... will not generate code.
2024-08-30 04:56:18.431 INFO org.apache.pig.mapplan.logical.optimizer.LogicalPlanOptimizer: [RULES [EnableDag, Adaptor, ColumnMapKeyPrune, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCaster, M
2024-08-30 04:56:18.671 INFO org.apache.pig.impl.util.SpillableMemoryManager: Selected heap (PS Old Gen) of size 89913840 to monitor. collectionUsageThreshold = 489396648, usageThreshold = 489396640
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: File concatenation threshold: 100 optimistic: false
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: MR plan size before optimization: 1
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: MR plan size after optimization: 1
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: yarn.resourcemanager.system-metrics-publisher.enabled is deprecated. Instead, use yarn.system-metrics-publisher.enabled
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: Connecting to ResourceManager at /8.0.0.0:8032
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: Pig script settings are added to the job
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: mapred.job.reduce.markreset.buffer.percent is deprecated. Instead, use mapreduce.reduce.markreset.buffer.percent
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: mapred.output.compress is deprecated. Instead, use mapreduce.output.fileoutputformat.compress
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: mapred.submit.replication is deprecated. Instead, use mapreduce.client.submit.file.replication
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: Added jar file:/home/hadoop/pig/lib/automaton-1.11-8.jar to DistributedCache through /tmp/tempt-1575382697/tmp-91625529
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: Added jar file:/home/hadoop/pig/lib/antlr-runtime-3.4.jar to DistributedCache through /tmp/tempt-1575382697/tmp1304382
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: Added jar file:/home/hadoop/pig/lib/antlr-runtime-3.4.jar to DistributedCache through /tmp/tempt-1575382697/tmp102979
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: Added jar file:/home/hadoop/pig/lib/joda-time-2.9.3.jar to DistributedCache through /tmp/tempt-1575382697/tmp20223822
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: Setting up single store job
2024-08-30 04:56:18.680 INFO org.apache.pig.data.SchemaTableFrontend: Key [pig.schema.table] is false, will not generate code.
2024-08-30 04:56:18.680 INFO org.apache.pig.data.SchemaTableFrontend: Starting process to move generated code to distributed cache
2024-08-30 04:56:18.680 INFO org.apache.pig.data.SchemaTableFrontend: Setting key [pig.schema.table.classes] with classes to serialize: []
2024-08-30 04:56:18.680 INFO org.apache.pig.backend.hadoop.executionengine.mapreduce.lamr.Compiler: 1 map-reduce job(s) waiting for submission.
2024-08-30 04:56:18.680 INFO org.apache.pig.Main: Pig script completed in 3 minutes, 46 seconds and 870 milliseconds (220870 ms)
  
```

## \$nano uppercase\_udf.py



```

File Actions Edit View Help
GNU nano 7.2 uppercase_udf.py
def uppercase(text):
    return text.upper()
if __name__ == "__main__":
    import sys
    for line in sys.stdin:
        line = line.strip()
        result = uppercase(line)
        print(result)
  
```



## \$hdfs dfs -copyFromLocal ~/DA-Lab/exp4/uppercase\_udf.py /exp4

```
(hadoop@kali)-[~/hadoop/bin]
$ ./hdfs dfs -ls /exp4
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-21 00:26:01,736 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
Found 3 items
drwxr-xr-x   - hadoop supergroup          0 2024-08-30 05:07 /exp4/output
-rw-r--r--   1 hadoop supergroup        27 2024-08-30 04:43 /exp4/sample.txt
-rw-r--r--   1 hadoop supergroup       172 2024-08-30 05:02 /exp4/uppercase_udf.py
```

## \$nano udf\_example.pig

```
hadoop@kali: ~/DA-Lab/exp4
File Actions Edit View Help
GNU nano 7.2 udf_example.pig
-- Register the Python UDF script
REGISTER 'hdfs:///exp4/uppercase_udf.py' USING jython AS udf;
-- Load some data
data = LOAD 'hdfs:///exp4/sample.txt' AS (text:chararray);
-- Use the Python UDF
uppercased_data = FOREACH data GENERATE udf.uppercase(text) AS uppercase_text;
-- Store the result
STORE uppercased_data INTO 'hdfs:///exp4/output';
```

## \$pig -f udf\_example.pig

```
kali-linux-2023.4-vmware-jam64 - VMware Workstation 17 Player (Non-commercial use only)
Player
hadoop@kali: ~/DA-Lab/exp4
$ nano uppercase_udf.py
$ cd ../hadoop/bin
$ ./hdfs dfs -put ~/DA-Lab/exp4/uppercase_udf.py /exp4
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-08-30 05:02:54,541 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
(hadoop@kali)-[~/hadoop/bin]
$ cd ../hadoop/bin
$ ./hdfs dfs -put ~/DA-Lab/exp4/uppercase_udf.py /exp4
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-08-30 05:02:54,541 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
(hadoop@kali)-[~/hadoop/bin]
$ cd ../hadoop/bin
$ nano udf_example.pig
$ pig -f udf_example.pig
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-08-30 05:06:18,501 INFO pig.ExecTypeProvider: Trying ExecType : LOCAL
2024-08-30 05:06:18,601 INFO pig.ExecTypeProvider: Trying ExecType : MAPREDUCE
2024-08-30 05:06:18,636 [main] INFO org.apache.pig.Main - Apache Pig version 0.16.0 (r1740538) compiled Jun 01 2016, 23:10:49
2024-08-30 05:06:18,665 [main] WARN org.apache.hadoop.util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2024-08-30 05:06:20,127 [main] INFO org.apache.pig.impl.util.Utils - Default bootstrap file /home/hadoop/.pigbootstrap not found
2024-08-30 05:06:20,382 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
2024-08-30 05:06:20,383 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:06:20,383 [main] INFO org.apache.pig.backend.hadoop.executionengine.MExecutionEngine - Connecting to hadoop file system at: hdfs://localhost:9000
2024-08-30 05:06:22,040 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:06:22,513 [main] INFO org.apache.pig.PigServer - Pig Script ID for the session: PIG-udf_example.pig-1dec08cc-3e11-4763-a9e8-501b5b6b38a1
2024-08-30 05:06:22,513 [main] WARN org.apache.pig.PigServer - ATS is disabled since yarn.timeline-service.enabled set to false
2024-08-30 05:06:22,725 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:06:24,290 [main] INFO org.apache.pig.scripting.jython.JythonScriptEngine - Created tmp python.cachedir/tmp/pig_jython_37818a2655894475167
2024-08-30 05:06:26,094 [main] WARN org.apache.pig.scripting.jython.JythonScriptEngine - pig.cmd.args.remaining is empty. This is not expected unless on testing.
2024-08-30 05:06:30,044 [main] INFO org.apache.pig.scripting.jython.JythonScriptEngine - Register scripting UDF: udf_uppercase
2024-08-30 05:06:31,784 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:06:31,867 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:06:38,487 [main] INFO org.apache.pig.scripting.jython.JythonFunction - No schema defined for function 'uppercase' in /tmp/pig2113342629392933085tmp/uppercase_udf.py
2024-08-30 05:06:38,503 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:06:38,743 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - mapred.textoutputformat.separator is deprecated. Instead, use mapreduce.output.textoutputformat.separator
2024-08-30 05:06:38,830 [main] INFO org.apache.pig.tools.pigstats.ScriptState - Pig features used in the script: UNKNOWN
2024-08-30 05:06:38,908 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
2024-08-30 05:06:38,982 [main] INFO org.apache.pig.data.SchemaTupleBackend - Key [pig_schematuple] was not set... will not generate code.
2024-08-30 05:06:39,132 [main] INFO org.apache.pig.mapgen.LogicalOptimizer.LogicalPlanOptimizer - [RULES_ENABLED: [InferPiggy, ColumnMajorPiggy, ConstantCalculator, GroupByConstParallelSetter, LimitOptimizer, LoadTypeCastInserter, M
ergeFilter, MergeForEach, PartitionFilterOptimizer, PredicatePushdownOptimizer, PushdownForEachFlatten, PushdownFilter, SplitFilter, StreamTypeCastInserter]]
2024-08-30 05:06:39,403 [main] INFO org.apache.pig.impl.util.SpillableMemoryManager - Selected heap (PS Old Gen) of size 699138048 to monitor: collectionUsageThreshold = 489396048, usageThreshold = 489396048
2024-08-30 05:06:39,604 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreduceLayer.MRCCompiler - File concatenation threshold: 184 optimizer: false
2024-08-30 05:06:39,858 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreduceLayer.MultiQueryOptimizer - MR plan size before optimization: 1
2024-08-30 05:06:39,899 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapreduceLayer.MultiQueryOptimizer - MR plan size after optimization: 1
2024-08-30 05:06:39,959 [main] INFO org.apache.hadoop.conf.Configuration.deprecation - fs.default.name is deprecated. Instead, use fs.defaultFS
```

```

kali-linux-2023.4-vmware-amd64 - VMware Workstation 17 Player (Non-commercial use only)
Player
File Actions Edit View Help
2024-08-30 05:10:12,729 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 4 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:13,731 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 5 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:14,733 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 6 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:15,735 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 7 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:16,738 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 8 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:17,740 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 9 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:17,859 [main] INFO org.apache.hadoop.mapred.ClientServiceDelegate - Application state is completed. FinalApplicationStatus=SUCCEEDED. Redirecting to job history server
2024-08-30 05:10:18,862 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 10 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:19,864 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 1 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:20,866 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 2 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:21,868 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 3 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:22,871 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 4 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:23,874 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 5 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:24,876 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 6 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:25,879 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 7 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:26,881 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 8 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:27,883 [main] INFO org.apache.hadoop.ipc.Client - Retrying connect to server: 0.0.0.0/0.0.0.0:10020. Already tried 9 time(s); retry policy is RetryUpToMaximumCountWithFixedSleep(maxRetries=10, sleepTime=1000 MILLISECS)
2024-08-30 05:10:27,987 [main] WARN org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Unable to retrieve job to compute warning aggregation.
2024-08-30 05:10:27,988 [main] INFO org.apache.pig.backend.hadoop.executionengine.mapReduceLayer.MapReduceLauncher - Success!
2024-08-30 05:10:28,108 [main] INFO org.apache.pig.Main - Pig script completed in 4 minutes, 9 seconds and 767 milliseconds (249747 ms)

(hadoop@kali) ~/DB-Lab/exp4
$ cd ../~/hadoop/bin

(hadoop@kali) ~/hadoop/bin
$ ./hdfs dfs -cat /exp4/output/part-m-00000
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-08-30 05:12:25,908 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
1,JOHN
2,JANE
3,JOE
4,EMMA

(hadoop@kali) ~/hadoop/bin

```

**\$hdfs dfs -cat /exp4/output/\***

```

(hadoop@kali) ~/hadoop/bin
$ ./hdfs dfs -cat /exp4/output/*
Picked up _JAVA_OPTIONS: -Dawt.useSystemAAFontSettings=on -Dswing.aatext=true
2024-09-21 00:33:32,731 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform...
1,JOHN
2,JANE
3,JOE
4,EMMA

```