

WEEK 7 BDA LAB
USN:1BM19CS109
NAME: P PREM SAI

MapReduce Join program using DeptEmpStrength.txt and DeptName.txt files.

```
hadoop@sharat-VirtualBox:~$ hdfs dfs -copyFromLocal DeptName.txt DeptStrength.txt /
2022-06-28 01:49:34,172 WARN util.NativeCodeLoader: Unable to load native-hadoop
p library for your platform... using builtin-java classes where applicable
copyFromLocal: `DeptStrength.txt': No such file or directory
hadoop@sharat-VirtualBox:~$ hdfs dfs -copyFromLocal DeptName.txt DeptEmpStrength
.txt /
2022-06-28 01:50:03,670 WARN util.NativeCodeLoader: Unable to load native-hadoop
p library for your platform... using builtin-java classes where applicable
copyFromLocal: `/DeptName.txt': File exists
hadoop@sharat-VirtualBox:~$ hdfs dfs -copyFromLocal DeptEmpStrength.txt /
2022-06-28 01:50:14,698 WARN util.NativeCodeLoader: Unable to load native-hadoop
p library for your platform... using builtin-java classes where applicable
copyFromLocal: `/DeptEmpStrength.txt': File exists
```

```
hadoop@sharat-VirtualBox:~$ hadoop jar MapReduceJoin.jar /DeptEmpStrength.txt /D
eptName.txt /output_mapreducejoin
2022-06-28 01:54:22,260 WARN util.NativeCodeLoader: Unable to load native-hadoop
p library for your platform... using builtin-java classes where applicable
2022-06-28 01:54:22,634 INFO client.RMProxy: Connecting to ResourceManager at /
127.0.0.1:8032
2022-06-28 01:54:22,756 INFO client.RMProxy: Connecting to ResourceManager at /
127.0.0.1:8032
2022-06-28 01:54:22,936 INFO mapreduce.JobResourceUploader: Disabling Erasure C
oding for path: /tmp/hadoop-yarn/staging/hadoop/.staging/job_1656358828291_0002
2022-06-28 01:54:23,108 INFO mapred.FileInputFormat: Total input files to proce
ss : 1
2022-06-28 01:54:23,121 INFO mapred.FileInputFormat: Total input files to proce
ss : 1
2022-06-28 01:54:23,607 INFO mapreduce.JobSubmitter: number of splits:4
2022-06-28 01:54:23,771 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1656358828291_0002
2022-06-28 01:54:23,772 INFO mapreduce.JobSubmitter: Executing with tokens: []
2022-06-28 01:54:23,909 INFO conf.Configuration: resource-types.xml not found
2022-06-28 01:54:23,909 INFO resource.ResourceUtils: Unable to find 'resource-t
ypes.xml'.
2022-06-28 01:54:23,967 INFO impl.YarnClientImpl: Submitted application applica
tion_1656358828291_0002
2022-06-28 01:54:24,008 INFO mapreduce.Job: The url to track the job: http://sh
arat-VirtualBox:8088/proxy/application_1656358828291_0002/
2022-06-28 01:54:24,010 INFO mapreduce.Job: Running job: job_1656358828291_0002
2022-06-28 01:54:29,095 INFO mapreduce.Job: Job job_1656358828291_0002 running
```

```

Combine output records=0
Reduce input groups=4
Reduce shuffle bytes=157
Reduce input records=8
Reduce output records=4
Spilled Records=16
Shuffled Maps =4
Failed Shuffles=0
Merged Map outputs=4
GC time elapsed (ms)=3812
CPU time spent (ms)=4570
Physical memory (bytes) snapshot=1479266304
Virtual memory (bytes) snapshot=12898156544
Total committed heap usage (bytes)=1481113600
Peak Map Physical memory (bytes)=447643648
Peak Map Virtual memory (bytes)=2578087936
Peak Reduce Physical memory (bytes)=194940928
Peak Reduce Virtual memory (bytes)=2586525696
Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=85

```

```

Bytes Written=85
hadoop@sharat-VirtualBox:~$ hdfs dfs -ls /outputoutput_mapreducejoin
2022-06-28 01:55:29,436 WARN util.NativeCodeLoader: Unable to load native-hadoop
p library for your platform... using builtin-java classes where applicable
ls: '/outputoutput_mapreducejoin': No such file or directory
hadoop@sharat-VirtualBox:~$ hdfs dfs -ls /output_mapreducejoin
2022-06-28 01:55:36,422 WARN util.NativeCodeLoader: Unable to load native-hadoop
p library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 hadoop supergroup 0 2022-06-28 01:54 /output_mapreducejo
in/_SUCCESS
-rw-r--r-- 1 hadoop supergroup 85 2022-06-28 01:54 /output_mapreducejo
in/part-000000
hadoop@sharat-VirtualBox:~$ hdfs dfs -cat /output_mapreducejoin/part-000000
2022-06-28 01:56:01,106 WARN util.NativeCodeLoader: Unable to load native-hadoop
p library for your platform... using builtin-java classes where applicable
A11 50 Finance
B12 100 HR
C13 250 Manufacturing
Dept_ID Total_Employee Dept_Name

```