

```

PS C:\Uni\HL\hl_2022\lab_4> hdfs dfs -ls
ls: .': No such file or directory
PS C:\Uni\HL\hl_2022\lab_4> hdfs dfs -ls /
Found 1 items
drwxr-xr-x - Denys_Kudriev supergroup 0 2023-01-14 20:41 /input_dir
PS C:\Uni\HL\hl_2022\lab_4> hdfs dfs -copyFromLocal /sample.txt /input_dir
copyFromLocal: /sample.txt': No such file or directory
PS C:\Uni\HL\hl_2022\lab_4> hdfs dfs -copyFromLocal sample.txt /input_dir
PS C:\Uni\HL\hl_2022\lab_4> hdfs dfs -ls /input_dir
Found 1 items
-rw-r--r-- 1 Denys_Kudriev supergroup 222 2023-01-14 20:51 /input_dir/sample.txt
PS C:\Uni\HL\hl_2022\lab_4>

```

```

2023-01-14 20:51:12,893 INFO namenode.FSEditLog: Number of transactions: 15 Total time for transactions(ms): 15 Number of transactions batched in Syncs: 1 Number of syncs: 14 SyncTimes(ms): 12
2023-01-14 20:51:12,928 INFO hdfs.StateChange: BLOCK* allocate blk_1073741827_1003, replicas=127.0.0.1:9866 for /input_dir/sample.txt._COPYING_
2023-01-14 20:51:13,011 INFO hdfs.StateChange: DIR* completeFile: /input_dir/sample.txt._COPYING_ is closed by DFSClient_NONMAPREDUCE_1685842791_1

```

```

PS C:\Uni\HL\hl_2022\lab_4> hadoop jar units.jar hadoop.ProcessUnits /input_dir output_dir
2023-01-15 17:22:02,373 INFO client.DefaultNoHARMAFailoverProxyProvider: Connecting to ResourceManager at /127.0.0.1:8032
2023-01-15 17:22:02,514 INFO client.DefaultNoHARMAFailoverProxyProvider: Connecting to ResourceManager at /127.0.0.1:8032
2023-01-15 17:22:02,678 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2023-01-15 17:22:02,703 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/Denys_Kudriev/.staging/job_1673795225699_0005
2023-01-15 17:22:02,910 INFO mapred.FileInputFormat: Total input files to process : 1
2023-01-15 17:22:02,966 INFO mapreduce.JobSubmitter: number of splits:2
2023-01-15 17:22:03,114 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1673795225699_0005
2023-01-15 17:22:03,115 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-01-15 17:22:03,281 INFO conf.Configuration: resource-types.xml not found
2023-01-15 17:22:03,282 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2023-01-15 17:22:03,691 INFO impl.YarnClientImpl: Submitted application application_1673795225699_0005
2023-01-15 17:22:03,729 INFO mapreduce.Job: The url to track the job: http://EPUAKYIW0F4B:8088/proxy/application_1673795225699_0005/
2023-01-15 17:22:03,730 INFO mapreduce.Job: Running job: job_1673795225699_0005
2023-01-15 17:22:12,858 INFO mapreduce.Job: Job job_1673795225699_0005 running in uber mode : false
2023-01-15 17:22:12,859 INFO mapreduce.Job: map 0% reduce 0%
2023-01-15 17:22:19,022 INFO mapreduce.Job: map 100% reduce 0%
2023-01-15 17:22:25,105 INFO mapreduce.Job: map 100% reduce 100%
2023-01-15 17:22:25,112 INFO mapreduce.Job: Job job_1673795225699_0005 completed successfully
2023-01-15 17:22:25,200 INFO mapreduce.Job: Counters: 54

```

```
2023-01-15 17:22:12,859 INFO mapreduce.Job: map 0% reduce 0%
2023-01-15 17:22:19,022 INFO mapreduce.Job: map 100% reduce 0%
2023-01-15 17:22:25,105 INFO mapreduce.Job: map 100% reduce 100%
2023-01-15 17:22:25,112 INFO mapreduce.Job: Job job_1673795225699_0005 completed successfully
2023-01-15 17:22:25,200 INFO mapreduce.Job: Counters: 54
```

File System Counters

```
FILE: Number of bytes read=61
FILE: Number of bytes written=796678
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=521
HDFS: Number of bytes written=75
HDFS: Number of read operations=11
HDFS: Number of large read operations=0
HDFS: Number of write operations=2
HDFS: Number of bytes read erasure-coded=0
```

Job Counters

```
Launched map tasks=2
Launched reduce tasks=1
Data-local map tasks=2
Total time spent by all maps in occupied slots (ms)=7934
Total time spent by all reduces in occupied slots (ms)=3065
Total time spent by all map tasks (ms)=7934
Total time spent by all reduce tasks (ms)=3065
Total vcore-milliseconds taken by all map tasks=7934
Total vcore-milliseconds taken by all reduce tasks=3065
Total megabyte-milliseconds taken by all map tasks=8124416
Total megabyte-milliseconds taken by all reduce tasks=3138560
```

Map-Reduce Framework

```
Map input records=5
Map output records=65
Map output bytes=585
Map output materialized bytes=67
Input split bytes=188
Combine input records=65
Combine output records=5
Reduce input groups=5
Reduce shuffle bytes=67
Reduce input records=5
Reduce output records=5
Spilled Records=10
Shuffled Maps =2
Failed Shuffles=0
Merged Map outputs=2
GC time elapsed (ms)=10
CPU time spent (ms)=1056
Physical memory (bytes) snapshot=1079238656
Virtual memory (bytes) snapshot=2223587328
Total committed heap usage (bytes)=1591738368
Peak Map Physical memory (bytes)=389890048
Peak Map Virtual memory (bytes)=736477184
Peak Reduce Physical memory (bytes)=301715456
Peak Reduce Virtual memory (bytes)=750940160
```

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=333
```

File Output Format Counters

```
Bytes Written=75
```

```
PS C:\Uni\HL\hl_2022\lab_4>
```

```

PS C:\Uni\HL\hl_2022\lab_4> hdfs dfs -ls /
Found 4 items
drwxr-xr-x - Denys_Kudriev supergroup      0 2023-01-15 17:09 /input_dir
drwxr-xr-x - Denys_Kudriev supergroup      0 2023-01-15 17:26 /output_dir
drwx----- - Denys_Kudriev supergroup      0 2023-01-15 17:07 /tmp
drwxr-xr-x - Denys_Kudriev supergroup      0 2023-01-15 17:22 /user
PS C:\Uni\HL\hl_2022\lab_4> hdfs dfs -ls /output_dir
Found 2 items
-rw-r--r--  1 Denys_Kudriev supergroup      0 2023-01-15 17:26 /output_dir/_SUCCESS
-rw-r--r--  1 Denys_Kudriev supergroup    75 2023-01-15 17:26 /output_dir/part-00000
PS C:\Uni\HL\hl_2022\lab_4> hdfs dfs -cat output_dir/part-00000
1979      24.615385
1980      29.153847
1981      33.615383
1984      39.615383
1985      36.923077
PS C:\Uni\HL\hl_2022\lab_4> hdfs dfs -copyToLocal output_dir/part-00000 output.txt
PS C:\Uni\HL\hl_2022\lab_4>

```

output.txt - Notepad

File	Edit	Format	View	Help
------	------	--------	------	------

1979	24.615385
1980	29.153847
1981	33.615383
1984	39.615383
1985	36.923077