

HDFS DATA IN AND OUT

By HadoopExam Learning Resources in Association with QuickTechie.com

Load data into and out of HDFS using the Hadoop File System (FS) commands.

1. Check All Available command

```
hdfs dfs
```

2. Get help on command

```
hdfs dfs -help get
```

3. Create Directory in HDFS

```
hdfs dfs -mkdir module2
```

4. Check Directory Created

```
hdfs dfs -ls
```

5. Now delete the directory we have just created

```
hdfs dfs -rmdir module2
```

6. Import file using put command: Copy files from the local file system into hdfs. Copying fails if the file already exists, unless the -f flag is given.

Go to local Path first using below command

```
cd /home/cloudera/Desktop/HadoopExam
```

Command to copy full directory on hdfs (It will copy entire directory content from local file System to HDFS.

```
hdfs dfs -put Module2
```

7. Check Directory has been copied successfully or not

Check Directory Created

```
hdfs dfs -ls
```

Check All files in directory copied successfully.

```
hdfs dfs -ls Module2
```

8. Remove/Delete individual file from hdfs

```
hdfs dfs -rm Module2/Purchase.txt
```

Check the file deleted or not

```
hdfs dfs -ls Module2
```

Some operation on files

9. Merge file : Now we need to merge all the files in a directory

```
hdfs dfs -getmerge Module2 All_Users.txt
```

Question: Where new file will be created?

Ans : On localdisk (Not on HDFS)

Question: In above command there is a small mistake, what is that?

Ans : You must add new line after each file , when you merge the files.

Question: When you are merging big files using getmerge command and it fails to execute, why?

Ans: As HDFS can store hugefiles and after merging it copy the files on local system, but local system is not capable of holding such a huge volume.

10. Adding new line character at the end of each file in merge process

```
hdfs dfs -getmerge -nl Module2 All_Users.txt
```

11. You see ^M character at the end of each line, you have to remove that.

```
dos2unix All_Users.txt
```

12. Exporting file from HDFS to local file system.

```
hdfs dfs -get Module2/Users.txt Module2/User_export.txt
```

13. Create a local directory

```
mkdir Module2_export
```

14. Export Entire Directory

```
hdfs dfs -get Module2 Module2_export
```

15. File Permissions (Using chmod command we can change the permission of the file in hdfs).
Current Permission of the files

```
[cloudera@quickstart Module2_export]$ hdfs dfs -ls Module2
Found 2 items
-rw-r--r-- 1 cloudera cloudera    84 2015-07-25 05:46 Module2/User_New.txt
-rw-r--r-- 1 cloudera cloudera    88 2015-07-25 05:46 Module2/Users.txt
```





Now change the permissions for a file

```
hdfs dfs -chmod 754 Module2/User_New.txt
```

Check the current permissions of the file

```
[cloudera@quickstart Module2_export]$ hdfs dfs -ls Module2
Found 2 items
-rwxr-xr-- 1 cloudera cloudera    84 2015-07-25 05:46 Module2/User_New.txt
-rw-r--r-- 1 cloudera cloudera    88 2015-07-25 05:46 Module2/Users.txt
```

Please check other Material Provided by www.HadoopExam.com

| | | | |
|---|---|--|---|
|  <p>430 Q & A Click Here</p> <p>Cloudera Hadoop Developer Certification CCD-410</p> |  <p>200 + Q & A Click Here</p> <p>Cloudera Hadoop Administrator Certification CCA-500</p> |  <p>262 Q & A Click Here</p> <p>Cloudera Hadoop HBase Certification CCB-400</p> |  <p>266 Q & A Click Here</p> <p>AWS Developer Certification Associate Level</p> |
|  <p>235 Q & A Click Here</p> <p>Cloudera Data Science Certification DS-200</p> |  <p>Click Here</p> <p>Hadoop Training With HandsOn</p> |  <p>Click Here</p> <p>Package Deal</p> |  <p>144 Q & A Click Here</p> <p>AWS Certified Solutions Architect Professional Level</p> |

| | | | |
|--|---|--|--|
|  300 + Q & A Click Here <div style="background-color: #666; color: #fff; padding: 5px; text-align: center;"> AWS Certified SysOps Administrator Associate Level </div> |  474 + Q & A Click Here <div style="background-color: #666; color: #fff; padding: 5px; text-align: center;"> AWS Certified Solutions Architect Associate Level </div> |  Click Here <div style="background-color: #666; color: #fff; padding: 5px; text-align: center;"> Click Here for AWS Package Deal </div> |  490 Q & A Click Here <div style="background-color: #666; color: #fff; padding: 5px; text-align: center;"> SAS Base Certification A00-211 </div> |
|  365 Q & A Click Here <div style="background-color: #666; color: #fff; padding: 5px; text-align: center;"> SAS Advance Certification A00-212 </div> |  86+ Q & A Click Here <div style="background-color: #666; color: #fff; padding: 5px; text-align: center;"> SAS Certified Statistical Business Analyst A00-240 </div> |  85 Q & A Click Here <div style="background-color: #666; color: #fff; padding: 5px; text-align: center;"> SAS Certified Platform Administrator 9 A00-250 </div> |  Click Here <div style="background-color: #666; color: #fff; padding: 5px; text-align: center;"> SAS Packaged Deal </div> |



234 Q & A

[Click Here](#)

**EMC Data Scientist
Associate
Certification
E20-007 (EMCDSA)**

Data Science certification really needs a good and in depth knowledge of statistics cum BigData Hadoop knowledge. It also require you to have good knowledge in like the main phases of the Data Analytics Lifecycle, analyzing and exploring data with R, statistics for model building and evaluation, the theory and methods of advanced analytics and statistical modeling, the technology and tools that can be used for advanced analytics, operationalizing an analytics project, and data visualization techniques. Successful candidates will achieve the EMC Proven Professional – Data Science Associate credential. Hence to clear the real exam it really needs very well preparation. So HadoopExam Learning Resources brings Data

Science Certification Simulator with 234 Practice Questions, which can help you to prepare for this exam in lesser time. **Practice - practice - practice!** The EMC:DS E20-007 Exam Simulator offers you the opportunity to take 4 sample Exams before heading out for the real thing. Be ready to succeed on exam day!

Upcoming Releases

1. [Apache Spark Training](#)
2. [Apache Spark Certification material](#)
3. [MongoDB Certification Material](#)
4. [Android Certification](#)
5. [Java Certification](#)

6. AWS Trainings
7. Data Science training