⚙️ HADOOP HDFS COMPLETE COMMAND GUIDE (Cloudera VM)

🔧1️⃣Required Pseudo Services to Start

```
sudo service hadoop-hdfs-namenode start
sudo service hadoop-hdfs-datanode start
sudo service hadoop-yarn-resourcemanager start
sudo service hadoop-yarn-nodemanager start
sudo service hadoop-mapreduce-historyserver start
```

✅ **Meaning:** Starts Hadoop's main file system (NameNode manages metadata, DataNode stores data).

🧠 *Without these, no file/directory operation works.*

📁2️⃣Create Directory in HDFS

```
hdfs dfs -mkdir /user/cloudera/demo
```

✅ **Meaning:** Creates a directory named demo inside your HDFS user folder.

💡 *Use -p if parent directories don't exist.*

📄3️⃣Create File in HDFS

```
hdfs dfs -touchz /user/cloudera/demo/sample.txt
```

✅ **Meaning:** Creates an empty file (0 bytes) named sample.txt in /user/cloudera/demo.

💻4️⃣Create File in Local System

```
echo "Hello Hadoop" > sample1.txt
```

✅ **Meaning:** Creates a file named sample1.txt in your local Desktop (current directory) with text inside.

💡 *Verify using ls.*

⬆️5️⃣ Move / Copy File from Local ⟶ HDFS

🅐 Using put

`hdfs dfs -put sample1.txt /user/cloudera/demo/`
✅ Uploads file from local ⟶ HDFS (keeps local copy).

🅑 Using copyFromLocal

`hdfs dfs -copyFromLocal sample1.txt /user/cloudera/demo/`
✅ Same as put, another syntax.

🅒 Using moveFromLocal

`hdfs dfs -moveFromLocal sample1.txt /user/cloudera/demo/`
✅ Moves file from local ⟶ HDFS (deletes local copy after upload).

🗂️6️⃣ View Directory and Files

`hdfs dfs -ls /user/cloudera/demo/`
✅ **Meaning:** Lists all files inside the /user/cloudera/demo directory.

📜7️⃣ View File Content (HDFS)

`hdfs dfs -cat /user/cloudera/demo/sample1.txt`
✅ **Meaning:** Displays the content of the file directly on the terminal.

⬇️8️⃣ Move / Copy File from HDFS ⟶ Local

🅐 Using get

`hdfs dfs -get /user/cloudera/demo/sample1.txt /home/cloudera/Desktop/`
✅ Copies file from HDFS to local (keeps it in HDFS).

🅑 Using copyToLocal

`hdfs dfs -copyToLocal /user/cloudera/demo/sample1.txt /home/cloudera/Desktop/`
✅ Same as get.

🅒 Using moveToLocal

`hdfs dfs -moveToLocal /user/cloudera/demo/sample1.txt /home/cloudera/Desktop/`
✅ Moves file from HDFS ⟶ local (deletes it from HDFS).

🧾9️⃣Check Disk Usage (du) and Filesystem Info (df)

Disk Usage (per file)

`hdfs dfs -du /user/cloudera/demo/`
✅ Shows file size and directory usage.

Filesystem Info

`hdfs dfs -df`
✅ Shows total, used, and available HDFS space.

🗑️ 🔟 Delete File or Directory

Delete Single File

`hdfs dfs -rm /user/cloudera/demo/sample1.txt`

Delete Entire Folder

`hdfs dfs -rm -r /user/cloudera/demo/`

✅ **Meaning:** Removes files or directories from HDFS.

♻️ 1️⃣1️⃣Empty Trash (Permanent Delete)

`hdfs dfs -expunge`

✅ **Meaning:** Empties the HDFS Trash — permanently deletes removed files.

📂 1️⃣2️⃣Copy / Move Files Inside HDFS

Copy File

`hdfs dfs -cp /user/cloudera/demo/sample1.txt /user/cloudera/demo/sample_copy.txt`

Move File

`hdfs dfs -mv /user/cloudera/demo/sample_copy.txt /user/cloudera/exp1/`
✅ **Meaning:** Copies or moves files within HDFS, not local.

🧠 1️⃣3️⃣ Common Fixes for Issues You Faced

Problem	Cause	Fix
No such file or directory	Wrong file name or folder doesn't exist	Check using ls before upload
Caught exception / InterruptedException	Hadoop services not running	Run sudo service hadoop-hdfs-namenode start & sudo service hadoop-hdfs-datanode start
mkdir: File exists	Directory already exists	Use another name or delete with -rm -r
Permission denied	Need admin rights	Add sudo -u hdfs before command
Local path confusion	Used /user/cloudera/Desktop (HDFS) instead of /home/cloudera/Desktop (local)	Always use /home/cloudera/Desktop for local paths
-pwd: Unknown command	pwd only works in local Linux, not HDFS	Use pwd locally, not hdfs dfs -pwd

---

⚡ 1️⃣4️⃣Quick Memory Table (For Exam/Viva)

Task	Command	Meaning
Create Directory	hdfs dfs -mkdir /path/	Make folder in HDFS
Create File in HDFS	hdfs dfs -touchz /path/file.txt	Create empty file
Create File in Local	echo "Text" > file.txt	Make local text file
Upload Local ⟶ HDFS	hdfs dfs -put file.txt /path/	Copy to HDFS
Download HDFS ⟶ Local	hdfs dfs -get /path/file.txt /home/cloudera/Desktop/	Copy to local
Move in HDFS	hdfs dfs -mv /src /dest	Move within HDFS
Copy in HDFS	hdfs dfs -cp /src /dest	Copy within HDFS
Check Size	hdfs dfs -du /path	Shows file sizes
Check Filesystem	hdfs dfs -df	Shows HDFS usage
Remove	hdfs dfs -rm /path	Delete file
Empty Trash	hdfs dfs -expunge	Permanent delete

🏁 1️⃣5️⃣Your Final Verification Flow (Full Practice Round)

```
# 1. Start Services
sudo service hadoop-hdfs-namenode start
sudo service hadoop-hdfs-datanode start

# 2. Make HDFS Directory
hdfs dfs -mkdir /user/cloudera/demo
```

```
# 3. Make Local File
echo "Hello Hadoop" > sample.txt

# 4. Upload to HDFS
hdfs dfs -put sample.txt /user/cloudera/demo

# 5. Verify Upload
hdfs dfs -ls /user/cloudera/demo

# 6. Check Content
hdfs dfs -cat /user/cloudera/demo/sample.txt

# 7. Move from HDFS → Local
hdfs dfs -moveToLocal /user/cloudera/demo/sample.txt /home/cloudera/Desktop/

# 8. Verify in Local
ls /home/cloudera/Desktop/
```

✅ **Done** — you've now completed the core HDFS cycle (Local → HDFS → Local) perfectly.