

PRACTICAL 3 : To Implement the following file management tasks in Hadoop:

- Adding files and directories.
- Retrieving files from HDFS to local file system.
- Deleting Files from HDFS.

```
maria_dev@sandbox-hdp:~$
--archives <comma separated list of archives>    specify comma separated archives
to be unarchived on the compute machines.

The general command line syntax is
bin/hadoop command [genericOptions] [commandOptions]

Usage: hadoop fs [generic options] -mkdir [-p] <path> ...
[maria_dev@sandbox-hdp ~]$ hadoop fs -mkdir deepanshi
[maria_dev@sandbox-hdp ~]$ hadoop fs -ls
Found 1 items
drwxr-xr-x   - maria_dev hdfs                0 2025-09-25 05:00 deepanshi
[maria_dev@sandbox-hdp ~]$ ls
[maria_dev@sandbox-hdp ~]$ pwd
/home/maria_dev
[maria_dev@sandbox-hdp ~]$ ls
[maria_dev@sandbox-hdp ~]$ wget http://media.sundog-soft.com/hadoop/ml-100k/u.data
--2025-09-25 05:01:45-- http://media.sundog-soft.com/hadoop/ml-100k/u.data
Resolving media.sundog-soft.com (media.sundog-soft.com)... 3.5.17.132, 52.216.21
1.169, 52.217.199.73, ...
Connecting to media.sundog-soft.com (media.sundog-soft.com)[3.5.17.132]:80... co
nected.
HTTP request sent, awaiting response... 200 OK
Length: 2079229 (2.0M) [application/octet-stream]
Saving to: 'u.data'

100%[=====>] 2,079,229   7.85KB/s   in 4m 0s

2025-09-25 05:05:57 (8.48 KB/s) = 'u.data' saved [2079229/2079229]

[maria_dev@sandbox-hdp ~]$ ls
u.data
[maria_dev@sandbox-hdp ~]$ ls -la
total 2060
drwx----- 1 maria_dev maria_dev   4096 Aug 25 05:01 .
drwxr-xr-x 1 root      root        4096 Jun 18 2018 ..
-rw----- 1 maria_dev maria_dev  1193 Aug 25 04:57 .bash_history
-rw-r--r-- 1 maria_dev maria_dev    18 Sep  6 2017 .bash_logout
-rw-r--r-- 1 maria_dev maria_dev   193 Sep  6 2017 .bash_profile
-rw-r--r-- 1 maria_dev maria_dev   619 Jun 18 2018 .bashrc
-rw-rw-r-- 1 maria_dev maria_dev 2079229 Nov 11 2016 u.data
[maria_dev@sandbox-hdp ~]$ hadoop fs -copyFromLocal u.data ml-100k/u.data
copyFromLocal: 'ml-100k/u.data': No such file or directory: 'hdfs://sandbox-hdp.
hortonworks.com:8020/user/maria_dev/ml-100k/u.data'
[maria_dev@sandbox-hdp ~]$ ls -l u.data
-rw-rw-r-- 1 maria_dev maria_dev 2079229 Nov 11 2016 u.data
[maria_dev@sandbox-hdp ~]$ ^C
[maria_dev@sandbox-hdp ~]$ hadoop fs -mkdir ml-100k
[maria_dev@sandbox-hdp ~]$
[maria_dev@sandbox-hdp ~]$ hadoop fs -copyFromLocal u.data ml-100k/
[maria_dev@sandbox-hdp ~]$
[maria_dev@sandbox-hdp ~]$ hadoop fs -ls ml-100k
```

```

[maria_dev@sandbox-hdp ~]$ hadoop fs -copyFromLocal u.data ml-100k/

[maria_dev@sandbox-hdp ~]$
[maria_dev@sandbox-hdp ~]$ hadoop fs -ls ml-100k

Found 1 items
-rw-r--r-- 1 maria_dev hdfs 2079229 2025-08-25 05:11 ml-100k/u.data
[maria_dev@sandbox-hdp ~]$
[maria_dev@sandbox-hdp ~]$ hadoop fs -ls
Found 2 items
drwxr-xr-x - maria_dev hdfs 0 2025-08-25 05:00 deepanshi
drwxr-xr-x - maria_dev hdfs 0 2025-08-25 05:11 ml-100k
[maria_dev@sandbox-hdp ~]$ hadoop fs -rm ml-100k/u.data
25/08/25 05:13:51 INFO fs.TrashPolicyDefault: Moved: 'hdfs://sandbox-hdp.hortonworks.com:8020/user/maria_dev/ml-100k/u.data' to trash at: hdfs://sandbox-hdp.hortonworks.com:8020/user/maria_dev/.Trash/Current/user/maria_dev/ml-100k/u.data
[maria_dev@sandbox-hdp ~]$ ^C
[maria_dev@sandbox-hdp ~]$ hadoop fs -rmdir ml-100k
[maria_dev@sandbox-hdp ~]$ hadoop fs -ls
Found 2 items
drwx----- - maria_dev hdfs 0 2025-08-25 05:13 .Trash
drwxr-xr-x - maria_dev hdfs 0 2025-08-25 05:00 deepanshi
[maria_dev@sandbox-hdp ~]$ ^C
[maria_dev@sandbox-hdp ~]$ Hadoop fs
-bash: Hadoop: command not found
[maria_dev@sandbox-hdp ~]$ hadoop fs
Usage: hadoop fs [generic options]
    [-appendToFile <localsrc> ... <dst>]
    [-cat [-ignoreCrc] <src> ...]
    [-checksum <src> ...]
    [-chgrp [-R] GROUP PATH...]
    [-chmod [-R] <MODE[,MODE]... | OCTALMODE> PATH...]
    [-chown [-R] [OWNER][:[GROUP]] PATH...]
    [-copyFromLocal [-f] [-p] [-l] <localsrc> ... <dst>]
    [-copyToLocal [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]
    [-count [-q] [-h] [-v] [-t [<storage type>]] [-u] <path> ...]
    [-cp [-f] [-p | -p[topax]] <src> ... <dst>]
    [-createSnapshot <snapshotDir> [<snapshotName>]]
    [-deleteSnapshot <snapshotDir> <snapshotName>]
    [-df [-h] [<path> ...]]
    [-du [-s] [-h] <path> ...]
    [-expunge]
    [-find <path> ... <expression> ...]
    [-get [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]
    [-getfacl [-R] <path>]
    [-getfattr [-R] {-n name | -d} [-e en] <path>]
    [-getmerge [-nl] <src> <localdst>]
    [-help [cmd ...]]
    [-ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [<path> ...]]
    [-mkdir [-p] <path> ...]
    [-moveFromLocal <localsrc> ... <dst>]
    [-moveToLocal <src> <localdst>]

```

```

[-chmod [-R] <MODE[,MODE]... | OCTALMODE> PATH...]
[-chown [-R] {OWNER}[:{GROUP}] PATH...]
[-copyFromLocal [-f] [-p] [-l] <localsrc> ... <dst>]
[-copyToLocal [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]
[-count [-q] [-h] [-v] [-t <storage type>]] [-u] <path> ...]
[-cp [-f] [-p | -p[topex]] <src> ... <dst>]
[-createSnapshot <snapshotDir> [<snapshotName>]]
[-deleteSnapshot <snapshotDir> <snapshotName>]
[-df [-h] [<path> ...]]
[-du [-s] [-h] <path> ...]
[-expunge]
[-find <path> ... <expression> ...]
[-get [-p] [-ignoreCrc] [-crc] <src> ... <localdst>]
[-getfacl [-R] <path>]
[-getfattr [-R] [-n name | -d] [-e en] <path>]
[-getmerge [-nl] <src> <localdst>]
[-help [cmd ...]]
[-ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [<path> ...]]
[-mkdir [-p] <path> ...]
[-moveFromLocal <localsrc> ... <dst>]
[-moveToLocal <src> <localdst>]
[-mv <src> ... <dst>]
[-put [-f] [-p] [-l] <localsrc> ... <dst>]
[-renameSnapshot <snapshotDir> <oldName> <newName>]
[-rm [-f] [-r|-R] [-skipTrash] [-safely] <src> ...]
[-rmdir [--ignore-fail-on-non-empty] <dir> ...]
[-setfacl [-R] [[-b|-k] (-m|-x <acl_spec>| <path>)] [--set <acl_spec> <pa
th>]]

[-setfattr [-n name [-v value] | -x name] <path>]
[-setrep [-R] [-w] <rep> <path> ...]
[-stat [format] <path> ...]
[-tail [-f] <file>]
[-test -[defaz] <path>]
[-text [-ignoreCrc] <src> ...]
[-touchz <path> ...]
[-truncate [-w] <length> <path> ...]
[-usage [cmd ...]]

```

Generic options supported are

```

-conf <configuration file>      specify an application configuration file
-D <property=value>             use value for given property
-fs <local|namenode:port>       specify a namenode
-jt <local|resourcemanager:port> specify a ResourceManager
-files <comma separated list of files> specify comma separated files to be co
pied to the map reduce cluster
-libjars <comma separated list of jars> specify comma separated jar files to
include in the classpath.
-archives <comma separated list of archives> specify comma separated archives
to be unarchived on the compute machines.

```

The general command line syntax is

```
bin/hadoop command [genericOptions] [commandOptions]
```

```
[maria_dev@sandbox-hdp ~]$ ^C
```

```
[maria_dev@sandbox-hdp ~]$
```