

## HDFS Commands

Open a terminal window to the current working directory.

```
# /home/training
```

# 1. Print the Hadoop version

```
hadoop version
```

# 2. List the contents of the root directory in HDFS

```
#
```

```
hadoop fs -ls /
```

# 3. Report the amount of space used and

# available on currently mounted filesystem

```
#
```

```
hadoop fs -df hdfs:/
```

# 4. Count the number of directories,files and bytes under

# the paths that match the specified file pattern

```
#
```

```
hadoop fs -count hdfs:/
```

# 5. Run a DFS filesystem checking utility

```
#
```

```
hadoop fsck - /
```

# 6. Run a cluster balancing utility

```
#
```

```
hadoop balancer
```

# 7. Create a new directory named "hadoop" below the

# /user/training directory in HDFS. Since you're

# currently logged in with the "training" user ID,

# /user/training is your home directory in HDFS.

```
#
```

```
hadoop fs -mkdir /user/training/hadoop
```

# 8. Add a sample text file from the local directory

# named "data" to the new directory you created in HDFS

# during the previous step.

```
#
```

```
hadoop fs -put data/sample.txt /user/training/hadoop
```

# 9. List the contents of this new directory in HDFS.

#

```
hadoop fs -ls /user/training/hadoop
```

# 10. Add the entire local directory called "retail" to the

# /user/training directory in HDFS.

#

```
hadoop fs -put data/retail /user/training/hadoop
```

# 11. Since /user/training is your home directory in HDFS,

# any command that does not have an absolute path is

# interpreted as relative to that directory. The next

# command will therefore list your home directory, and

# should show the items you've just added there.

#

```
hadoop fs -ls
```

# 12. See how much space this directory occupies in HDFS.

#

```
hadoop fs -du -s -h hadoop/retail
```

# 13. Delete a file 'customers' from the "retail" directory.

#

```
hadoop fs -rm hadoop/retail/customers
```

# 14. Ensure this file is no longer in HDFS.

#

```
hadoop fs -ls hadoop/retail/customers
```

# 15. Delete all files from the "retail" directory using a wildcard.

#

```
hadoop fs -rm hadoop/retail/*
```

# 16. To empty the trash

#

```
hadoop fs -expunge
```

# 17. Finally, remove the entire retail directory and all

# of its contents in HDFS.

#

```
hadoop fs -rm -r hadoop/retail
```

# 18. List the hadoop directory again

```
#
hadoop fs -ls hadoop

# 19. Add the purchases.txt file from the local directory
# named "/home/training/" to the hadoop directory you created in HDFS
#
hadoop fs -copyFromLocal /home/training/purchases.txt hadoop/

# 20. To view the contents of your text file purchases.txt
# which is present in your hadoop directory.
#
hadoop fs -cat hadoop/purchases.txt

# 21. Add the purchases.txt file from "hadoop" directory which is present in HDFS directory
# to the directory "data" which is present in your local directory
#
hadoop fs -copyToLocal hadoop/purchases.txt /home/training/data

# 22. cp is used to copy files between directories present in HDFS
#
hadoop fs -cp /user/training/*.txt /user/training/hadoop

# 23. '-get' command can be used alternatively to '-copyToLocal' command
#
hadoop fs -get hadoop/sample.txt /home/training/

# 24. Display last kilobyte of the file "purchases.txt" to stdout.
#
hadoop fs -tail hadoop/purchases.txt

# 25. Default file permissions are 666 in HDFS
# Use '-chmod' command to change permissions of a file
#
hadoop fs -ls hadoop/purchases.txt
sudo -u hdfs hadoop fs -chmod 600 hadoop/purchases.txt

# 26. Default names of owner and group are training,training
# Use '-chown' to change owner name and group name simultaneously
#
hadoop fs -ls hadoop/purchases.txt
sudo -u hdfs hadoop fs -chown root:root hadoop/purchases.txt

# 27. Default name of group is training
# Use '-chgrp' command to change group name
```

```
#  
hadoop fs -ls hadoop/purchases.txt  
sudo -u hdfs hadoop fs -chgrp training hadoop/purchases.txt
```

# 28. Move a directory from one location to other

```
#  
hadoop fs -mv hadoop apache_hadoop
```

# 29. Default replication factor to a file is 3.

# Use '-setrep' command to change replication factor of a file

```
#  
hadoop fs -setrep -w 2 apache_hadoop/sample.txt
```

# 30. Copy a directory from one node in the cluster to another

# Use '-distcp' command to copy,

# -overwrite option to overwrite in an existing files

# -update command to synchronize both directories

```
#  
hadoop fs -distcp hdfs://namenodeA/apache_hadoop hdfs://namenodeB/hadoop
```

# 31. Command to make the name node leave safe mode

```
#  
hadoop fs -expunge  
sudo -u hdfs hdfs dfsadmin -safemode leave
```

# 32. List all the hadoop file system shell commands

```
#  
hadoop fs
```

# 33. Last but not least, always ask for help!

```
#  
hadoop fs -help
```