Frequently used HDFS shell commands

```
# Open a terminal window to the current working directory.
# /home/hadoop

# 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 “cse5” below the
# /user/hadoop directory in HDFS. Since you’re
# currently logged in with the “hadoop” user ID,
# /user/hadoop is your home directory in HDFS.
#
hadoop fs -mkdir /user/hadoop/cse5

# 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/hadoop/cse5

# 9. List the contents of this new directory in HDFS.
#
hadoop fs -ls /user/hadoop/cse5
```

```
# 10. Add the entire local directory called “retail” to the
# cse5 directory in HDFS.
#
hadoop fs -put data/retail /user/training/cse5

# 11. Since /user/hadoop  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 cse5/retail

# 13. Delete a file ‘customers’ from the “retail” directory.
#
hadoop fs -rm cse5/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 cse5/retail

# 18. List the cse5 directory again
#
hadoop fs -ls cse5

# 19. Add the purchases.txt file from the local directory
# named “/home/hadoop/” to the cse5 directory you created in HDFS
#
hadoop fs -copyFromLocal /home/hadoop/purchases.txt cse5/

# 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 “cse5” directory which is present in HDFS directory
# to the directory “data” which is present in your local directory
#
hadoop fs -copyToLocal cse5/purchases.txt /home/hadoop/data

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

# 23. ‘-get’ command can be used alternaively to ‘-copyToLocal’ command
#
hadoop fs -get hadoop/sample.txt /home/hadoop/

# 24. Display last kilobyte of the file “purchases.txt” to stdout.
#
hadoop fs -tail cse5/purchases.txt

# 25. Move a directory from one location to other
#
hadoop fs -mv hadoop apache_hadoop

# 26. 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

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