

```
[root@sandbox-hdp ~]# hdfs dfs -du -s /test_data/hello.dat
0 /test_data/hello.dat
```

- To determine the size of a file; shows size in kb

```
[root@sandbox-hdp ~]# hdfs dfs -cat /test_data/hello.dat
how are you
Hope you are doing well
hope you are fine and doing well
```

- Shows the content of the file.

```
[root@sandbox-hdp ~]# cat>>hello1.dat
hi
hope you are well
[root@sandbox-hdp ~]# ls hello1.dat
hello1.dat
[root@sandbox-hdp ~]# ls
anaconda-ks.cfg hello1.dat
[root@sandbox-hdp ~]# hdfs dfs -copyFromLocal hello1.dat /test_data/
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data
Found 3 items
-rw-r--r-- 1 root hdfs 0 2020-05-23 05:44 /test_data/hello
-rw-r--r-- 1 root hdfs 69 2020-05-23 06:20 /test_data/hello.dat
-rw-r--r-- 1 root hdfs 21 2020-05-23 06:35 /test_data/hello1.dat
```

- Cat>> is used to create files in the local system
- Ls and file name can be used to find out if file present in system
- -copyFromLocal used to move files from local to HDFS (similar to get)

```
[root@sandbox-hdp ~]# cat>>how2.dat
how are u doing these days?
[root@sandbox-hdp ~]# ls
anaconda-ks.cfg hello1.dat hello2.dat hello3.dat how2.dat how.dat
[root@sandbox-hdp ~]# hdfs dfs -put how2.dat /test_data/
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data
ls: '/test_data': No such file or directory
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data
Found 7 items
-rw-r--r-- 1 root hdfs 0 2020-05-23 05:44 /test_data/hello
-rw-r--r-- 1 root hdfs 69 2020-05-23 06:20 /test_data/hello.dat
-rw-r--r-- 1 root hdfs 21 2020-05-23 06:35 /test_data/hello1.dat
-rw-r--r-- 1 root hdfs 18 2020-05-23 06:38 /test_data/hello2.dat
-rw-r--r-- 1 root hdfs 26 2020-05-23 07:04 /test_data/hello3.dat
-rw-r--r-- 1 root hdfs 13 2020-05-23 08:39 /test_data/how.dat
-rw-r--r-- 1 root hdfs 29 2020-05-23 08:47 /test_data/how2.dat
```

- -put is similar to -copy from local.

```
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data
```

```
Found 8 items
```

```
-rw-r--r-- 1 root hdfs 0 2020-05-23 05:44 /test_data/hello
-rw-r--r-- 1 root hdfs 69 2020-05-23 06:20 /test_data/hello.dat
-rw-r--r-- 1 root hdfs 21 2020-05-23 06:35 /test_data/hello1.dat
-rw-r--r-- 1 root hdfs 18 2020-05-23 06:38 /test_data/hello2.dat
-rw-r--r-- 1 root hdfs 26 2020-05-23 07:04 /test_data/hello3.dat
-rw-r--r-- 1 root hdfs 13 2020-05-23 08:39 /test_data/how.dat
-rw-r--r-- 1 root hdfs 29 2020-05-23 08:47 /test_data/how2.dat
-rw-r--r-- 1 root hdfs 0 2020-05-23 08:54 /test_data/test.dat
```

```
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data1
```

```
[root@sandbox-hdp ~]# hdfs dfs -cp /test_data/hello.dat /test_data1/
```

```
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data1
```

```
Found 1 items
```

```
-rw-r--r-- 1 root hdfs 69 2020-05-23 09:15 /test_data1/hello.dat
```

- -cp used to copy files from one HDFS directory to another.

```
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data
```

```
Found 7 items
```

```
-rw-r--r-- 1 root hdfs 0 2020-05-23 05:44 /test_data/hello
-rw-r--r-- 1 root hdfs 69 2020-05-23 06:20 /test_data/hello.dat
-rw-r--r-- 1 root hdfs 21 2020-05-23 06:35 /test_data/hello1.dat
-rw-r--r-- 1 root hdfs 26 2020-05-23 07:04 /test_data/hello3.dat
-rw-r--r-- 1 root hdfs 13 2020-05-23 08:39 /test_data/how.dat
-rw-r--r-- 1 root hdfs 29 2020-05-23 08:47 /test_data/how2.dat
-rw-r--r-- 1 root hdfs 0 2020-05-23 08:54 /test_data/test.dat
```

```
[root@sandbox-hdp ~]# hdfs dfs -rm /test_data/hello.dat
```

```
20/05/23 09:41:00 INFO fs.TrashPolicyDefault: Moved: 'hdfs://sandbox-hdp.hort
.com:8020/user/root/.Trash/Current/test_data/hello.dat'
```

```
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data
```

```
Found 6 items
```

```
-rw-r--r-- 1 root hdfs 0 2020-05-23 05:44 /test_data/hello
-rw-r--r-- 1 root hdfs 21 2020-05-23 06:35 /test_data/hello1.dat
-rw-r--r-- 1 root hdfs 26 2020-05-23 07:04 /test_data/hello3.dat
-rw-r--r-- 1 root hdfs 13 2020-05-23 08:39 /test_data/how.dat
-rw-r--r-- 1 root hdfs 29 2020-05-23 08:47 /test_data/how2.dat
-rw-r--r-- 1 root hdfs 0 2020-05-23 08:54 /test_data/test.dat
```

- -rm used to delete files

```
[root@sandbox-hdp ~]# hdfs dfs -rmdir /test_data1
rmdir: `/test_data1': Directory is not empty
```

```
[root@sandbox-hdp ~]# hdfs dfs -mkdir /hello1
[root@sandbox-hdp ~]# hdfs dfs -ls /
Found 19 items
drwxrwxrwx - yarn      hadoop      0 2018-06-18 15:18 /app-logs
drwxr-xr-x - hdfs      hdfs        0 2018-06-18 16:13 /apps
drwxr-xr-x - yarn      hadoop      0 2018-06-18 14:52 /ats
drwxr-xr-x - root      hdfs        0 2020-05-23 03:10 /bigdatatest1
drwxr-xr-x - root      hdfs        0 2020-05-23 02:31 /hadoop2
-rwxr-xr-x 1 root      hdfs        0 2020-05-23 02:10 /hadoop3
-rwxr-xr-x 1 root      hdfs        0 2020-05-23 02:12 /hadoop4
drwxr-xr-x - hdfs      hdfs        0 2018-06-18 14:52 /hdp
drwxr-xr-x - root      hdfs        0 2020-05-23 05:35 /hello
drwxr-xr-x - root      hdfs        0 2020-05-23 09:49 /hello1
drwx----- - livy      hdfs        0 2018-06-18 15:11 /livy2-recovery
drwxr-xr-x - mapred    hdfs        0 2018-06-18 14:52 /mapred
drwxrwxrwx - mapred    hadoop      0 2018-06-18 14:52 /mr-history
drwxr-xr-x - hdfs      hdfs        0 2018-06-18 15:59 /ranger
drwxrwxrwx - spark     hadoop      0 2020-05-23 00:27 /spark2-history
drwxr-xr-x - root      hdfs        0 2020-05-23 09:41 /test_data
drwxr-xr-x - root      hdfs        0 2020-05-23 09:24 /test_data1
drwxrwxrwx - hdfs      hdfs        0 2018-06-18 16:06 /tmp
drwxr-xr-x - hdfs      hdfs        0 2018-06-18 16:08 /user
[root@sandbox-hdp ~]# hdfs dfs -rmdir /hello
[root@sandbox-hdp ~]# hdfs dfs -rmdir /hello1
[root@sandbox-hdp ~]# hdfs dfs -ls /
Found 17 items
drwxrwxrwx - yarn      hadoop      0 2018-06-18 15:18 /app-logs
drwxr-xr-x - hdfs      hdfs        0 2018-06-18 16:13 /apps
drwxr-xr-x - yarn      hadoop      0 2018-06-18 14:52 /ats
drwxr-xr-x - root      hdfs        0 2020-05-23 03:10 /bigdatatest1
drwxr-xr-x - root      hdfs        0 2020-05-23 02:31 /hadoop2
-rwxr-xr-x 1 root      hdfs        0 2020-05-23 02:10 /hadoop3
-rwxr-xr-x 1 root      hdfs        0 2020-05-23 02:12 /hadoop4
drwxr-xr-x - hdfs      hdfs        0 2018-06-18 14:52 /hdp
drwx----- - livy      hdfs        0 2018-06-18 15:11 /livy2-recovery
drwxr-xr-x - mapred    hdfs        0 2018-06-18 14:52 /mapred
drwxrwxrwx - mapred    hadoop      0 2018-06-18 14:52 /mr-history
drwxr-xr-x - hdfs      hdfs        0 2018-06-18 15:59 /ranger
drwxrwxrwx - spark     hadoop      0 2020-05-23 00:27 /spark2-history
drwxr-xr-x - root      hdfs        0 2020-05-23 09:41 /test_data
drwxr-xr-x - root      hdfs        0 2020-05-23 09:24 /test_data1
drwxrwxrwx - hdfs      hdfs        0 2018-06-18 16:06 /tmp
drwxr-xr-x - hdfs      hdfs        0 2018-06-18 16:08 /user
```

- -rmdir only works for empty directories

```
[root@sandbox-hdp ~]# hdfs dfs -appendToFile - /test_data/hello.dat
how are you
[root@sandbox-hdp ~]# hdfs dfs -du -s /test_data/hello.dat
12 /test_data/hello.dat
[root@sandbox-hdp ~]# hdfs dfs -appendToFile - /test_data/hello.dat
Hope you are doing well
[root@sandbox-hdp ~]# hdfs dfs -du -s /test_data/hello.dat
36 /test_data/hello.dat
```

- After entering the required change to the file, press enter and then ctrl+d to go back to root

```
[root@sandbox-hdp ~]# hdfs dfs -touchz /test_data/test.dat
h[root@sandbox-hdp ~]#hdfs dfs -ls /test_data
Found 8 items
-rw-r--r-- 1 root hdfs 0 2020-05-23 05:44 /test_data/hello
-rw-r--r-- 1 root hdfs 69 2020-05-23 06:20 /test_data/hello.dat
-rw-r--r-- 1 root hdfs 21 2020-05-23 06:35 /test_data/hello1.dat
-rw-r--r-- 1 root hdfs 18 2020-05-23 06:38 /test_data/hello2.dat
-rw-r--r-- 1 root hdfs 26 2020-05-23 07:04 /test_data/hello3.dat
-rw-r--r-- 1 root hdfs 13 2020-05-23 08:39 /test_data/how.dat
-rw-r--r-- 1 root hdfs 29 2020-05-23 08:47 /test_data/how2.dat
-rw-r--r-- 1 root hdfs 0 2020-05-23 08:54 /test_data/test.dat
```

- -touchz is used to create files within a directory

```
[root@sandbox-hdp ~]# pwd
/root
[root@sandbox-hdp ~]# ls
anaconda-ks.cfg hello1.dat hello2.dat hello3.dat how2.dat how.dat
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data
Found 8 items
-rw-r--r-- 1 root hdfs 0 2020-05-23 05:44 /test_data/hello
-rw-r--r-- 1 root hdfs 69 2020-05-23 06:20 /test_data/hello.dat
-rw-r--r-- 1 root hdfs 21 2020-05-23 06:35 /test_data/hello1.dat
-rw-r--r-- 1 root hdfs 18 2020-05-23 06:38 /test_data/hello2.dat
-rw-r--r-- 1 root hdfs 26 2020-05-23 07:04 /test_data/hello3.dat
-rw-r--r-- 1 root hdfs 13 2020-05-23 08:39 /test_data/how.dat
-rw-r--r-- 1 root hdfs 29 2020-05-23 08:47 /test_data/how2.dat
-rw-r--r-- 1 root hdfs 0 2020-05-23 08:54 /test_data/test.dat
[root@sandbox-hdp ~]# hdfs dfs -copyToLocal /test_data/test.dat /root
[root@sandbox-hdp ~]# ls
anaconda-ks.cfg hello1.dat hello2.dat hello3.dat how2.dat how.dat test.dat
[root@sandbox-hdp ~]# hdfs dfs -get /test_data/hello.dat /root
[root@sandbox-hdp ~]# ls
anaconda-ks.cfg hello1.dat hello2.dat hello3.dat hello.dat how2.dat how.dat test.dat
```

- Pwd gives the current path
- -copyToLocal and -get copies files from HDFS to local storage


```
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data
Found 8 items
-rw-r--r-- 1 root hdfs 0 2020-05-23 05:44 /test_data/hello
-rw-r--r-- 1 root hdfs 69 2020-05-23 06:20 /test_data/hello.dat
-rw-r--r-- 1 root hdfs 21 2020-05-23 06:35 /test_data/hello1.dat
-rw-r--r-- 1 root hdfs 18 2020-05-23 06:38 /test_data/hello2.dat
-rw-r--r-- 1 root hdfs 26 2020-05-23 07:04 /test_data/hello3.dat
-rw-r--r-- 1 root hdfs 13 2020-05-23 08:39 /test_data/how.dat
-rw-r--r-- 1 root hdfs 29 2020-05-23 08:47 /test_data/how2.dat
-rw-r--r-- 1 root hdfs 0 2020-05-23 08:54 /test_data/test.dat
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data1
Found 2 items
-rw-r--r-- 1 root hdfs 69 2020-05-23 09:15 /test_data1/hello.dat
-rw-r--r-- 1 root hdfs 21 2020-05-23 09:17 /test_data1/hello1.dat
[root@sandbox-hdp ~]# hdfs dfs -mv /test_data/hello2.dat /test_data1/
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data
Found 7 items
-rw-r--r-- 1 root hdfs 0 2020-05-23 05:44 /test_data/hello
-rw-r--r-- 1 root hdfs 69 2020-05-23 06:20 /test_data/hello.dat
-rw-r--r-- 1 root hdfs 21 2020-05-23 06:35 /test_data/hello1.dat
-rw-r--r-- 1 root hdfs 26 2020-05-23 07:04 /test_data/hello3.dat
-rw-r--r-- 1 root hdfs 13 2020-05-23 08:39 /test_data/how.dat
-rw-r--r-- 1 root hdfs 29 2020-05-23 08:47 /test_data/how2.dat
-rw-r--r-- 1 root hdfs 0 2020-05-23 08:54 /test_data/test.dat
[root@sandbox-hdp ~]# hdfs dfs -ls /test_data1
Found 3 items
-rw-r--r-- 1 root hdfs 69 2020-05-23 09:15 /test_data1/hello.dat
-rw-r--r-- 1 root hdfs 21 2020-05-23 09:17 /test_data1/hello1.dat
-rw-r--r-- 1 root hdfs 18 2020-05-23 06:38 /test_data1/hello2.dat
```

- -mv used to move (cut and paste not copy) from one HDFS directory to another
- Hello2.dat moved from test_data to test_data1

```
[root@sandbox-hdp ~]# hdfs dfs -ls /
Found 19 items
drwxrwxrwx - yarn 0 2018-06-18 15:18 /app-logs
drwxr-xr-x - hdfs 0 2018-06-18 16:13 /apps
drwxr-xr-x - yarn 0 2018-06-18 14:52 /ats
drwxr-xr-x - root 0 2020-05-23 03:15 /bigdatatest
drwxr-xr-x - root 0 2020-05-23 03:10 /bigdatatest1
drwxr-xr-x - root 0 2020-05-23 02:31 /hadoop2
-rwxr-xr-x 1 root 0 2020-05-23 02:10 /hadoop3
drwxr-xr-x - hdfs 0 2020-05-23 02:12 /hadoop4
drwxr-xr-x - hdfs 0 2018-06-18 14:52 /hdp
drwxr-xr-x - root 0 2020-05-23 05:35 /hello
drwx----- - livy 0 2018-06-18 15:11 /livy2-recovery
drwxr-xr-x - mapred 0 2018-06-18 14:52 /mapred
drwxrwxrwx - mapred 0 2018-06-18 14:52 /mr-history
drwxr-xr-x - hdfs 0 2018-06-18 15:59 /ranger
drwxrwxrwx - spark 0 2020-05-23 00:27 /spark2-history
drwxr-xr-x - root 0 2020-05-23 09:41 /test_data
drwxr-xr-x - root 0 2020-05-23 09:24 /test_data1
drwxrwxrwx - hdfs 0 2018-06-18 16:06 /tmp
drwxr-xr-x - hdfs 0 2018-06-18 16:08 /user
[root@sandbox-hdp ~]# hdfs dfs -rm -r /bigdatatest
20/05/23 09:46:25 INFO fs.TrashPolicyDefault: Moved: 'hdfs://sandbox-hdp.h
0/user/root/.Trash/Current/bigdatatest'
[root@sandbox-hdp ~]# hdfs dfs -ls /
Found 18 items
drwxrwxrwx - yarn 0 2018-06-18 15:18 /app-logs
drwxr-xr-x - hdfs 0 2018-06-18 16:13 /apps
drwxr-xr-x - yarn 0 2018-06-18 14:52 /ats
drwxr-xr-x - root 0 2020-05-23 03:10 /bigdatatest1
drwxr-xr-x - root 0 2020-05-23 02:31 /hadoop2
-rwxr-xr-x 1 root 0 2020-05-23 02:10 /hadoop3
drwxr-xr-x - hdfs 0 2020-05-23 02:12 /hadoop4
drwxr-xr-x - hdfs 0 2018-06-18 14:52 /hdp
drwxr-xr-x - root 0 2020-05-23 05:35 /hello
drwx----- - livy 0 2018-06-18 15:11 /livy2-recovery
drwxr-xr-x - mapred 0 2018-06-18 14:52 /mapred
drwxrwxrwx - mapred 0 2018-06-18 14:52 /mr-history
drwxr-xr-x - hdfs 0 2018-06-18 15:59 /ranger
drwxrwxrwx - spark 0 2020-05-23 00:27 /spark2-history
drwxr-xr-x - root 0 2020-05-23 09:41 /test_data
drwxr-xr-x - root 0 2020-05-23 09:24 /test_data1
drwxrwxrwx - hdfs 0 2018-06-18 16:06 /tmp
drwxr-xr-x - hdfs 0 2018-06-18 16:08 /user
```

- -rm -r can be used to remove delete directories even if they are not empty