

BDA Experiment 1

Name: Sahil Dinesh Patil	Class: BE3
Roll no: 38	Batch: B

- \$ hadoop fs -ls
It will list available files

```
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hadoop fs -ls /
Found 10 items
drwxr-xr-x - cloudera supergroup      0 2024-07-18 03:31 /607
drwxr-xr-x - cloudera supergroup      0 2024-07-18 22:22 /bda
drwxrwxrwx - hdfs supergroup          0 2017-10-23 09:15 /benchmarks
drwxr-xr-x - cloudera supergroup      0 2024-07-18 22:22 /cloudera
drwxr-xr-x - hdfs supergroup          0 2024-07-14 23:29 /dezyre
drwxr-xr-x - hbase supergroup         0 2024-07-28 22:34 /hbase
drwxr-xr-x - solr solr                0 2017-10-23 09:18 /solr
drwxrwxrwt - hdfs supergroup          0 2024-07-14 22:17 /tmp
drwxr-xr-x - hdfs supergroup          0 2017-10-23 09:17 /user
drwxr-xr-x - hdfs supergroup          0 2017-10-23 09:17 /var
[cloudera@quickstart ~]$
```

```
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hadoop fs -ls
Found 7 items
drwxr-xr-x - cloudera cloudera      0 2024-07-18 03:31 607
drwxr-xr-x - cloudera cloudera      0 2024-07-14 22:45 dezyre
drwxr-xr-x - cloudera cloudera      0 2024-07-14 23:35 dezyrel
drwxr-xr-x - cloudera cloudera      0 2024-07-15 22:24 dezyrel
drwxr-xr-x - cloudera cloudera      0 2024-07-18 03:27 khushboo
drwxr-xr-x - cloudera cloudera      0 2024-07-18 03:26 lord
drwxr-xr-x - cloudera cloudera      0 2024-07-18 03:19 satshil
[cloudera@quickstart ~]$
```

- `$ hadoop fs -usage ls`

It gives options used with particular hdfs command

```

File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hadoop fs -usage ls
Usage: hadoop fs [generic options] -ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [<path> ...]
[cloudera@quickstart ~]$

```

- `$ hadoop fs -help ls`

Lists usage information along with options

```

File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hadoop fs -help ls
-ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [<path> ...] :
  List the contents that match the specified file pattern. If path is not
  specified, the contents of /user/<currentUser> will be listed. For a directory a
  list of its direct children is returned (unless -d option is specified).

  Directory entries are of the form:
      permissions - userId groupId sizeOfDirectory(in bytes)
  modificationDate(yyyy-MM-dd HH:mm) directoryName

  and file entries are of the form:
      permissions numberOfReplicas userId groupId sizeOfFile(in bytes)
  modificationDate(yyyy-MM-dd HH:mm) fileName

  -C  Display the paths of files and directories only.
  -d  Directories are listed as plain files.
  -h  Formats the sizes of files in a human-readable fashion
      rather than a number of bytes.
  -q  Print ? instead of non-printable characters.
  -R  Recursively list the contents of directories.
  -t  Sort files by modification time (most recent first).
  -S  Sort files by size.
  -r  Reverse the order of the sort.
  -u  Use time of last access instead of modification for
      display and sorting.
[cloudera@quickstart ~]$

```

- \$ `hadoop fs -help`
To figure out all the available hadoop commands

```

[cloudera@quickstart ~]$ hadoop fs -help
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] [-x] <path> ...]
    [-cp [-f] [-p | -p[topax]] <src> ... <dst>]
    [-createSnapshot <snapshotDir> [<snapshotName>]]
    [-deleteSnapshot <snapshotDir> <snapshotName>]
    [-df [-h] [<path> ...]]
    [-du [-s] [-h] [-x] <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] <src> ...]
    [-rmdir [--ignore-fail-on-non-empty] <dir> ...]
    [-setfacl [-R] [{-b|-k} {-m|-x <acl_spec>} <path>]|[--set <acl_spec> <path>]]
    [-setfattr {-n name [-v value] | -x name} <path>]
    [-setrep [-R] [-w] <rep> <path> ...]
    [-stat [format] <path> ...]
    [-tail [-f] <file>]
    [-test [-defsz] <path>]
    [-text [-ignoreCrc] <src> ...]
    [-touchz <path> ...]
    [-usage [cmd ...]]

-appendToFile <localsrc> ... <dst> :
    Appends the contents of all the given local files to the given dst file. The dst
    file will be created if it does not exist. If <localSrc> is -, then the input is
    read from stdin.

-cat [-ignoreCrc] <src> ... :
    Fetch all files that match the file pattern <src> and display their content on
    stdout.

-checksum <src> ... :

```

- \$ `hadoop fs -mkdir /user/cloudera/DF`
To create new directory

```
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hadoop fs -mkdir /user/cloudera/DF
[cloudera@quickstart ~]$
```

- \$ `hadoop fs -ls -R /user/cloudera`
Returns all available files and subdirectories

```
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hadoop fs -ls -R /user/cloudera
drwxr-xr-x - cloudera cloudera      0 2024-07-18 03:31 /user/cloudera/607
drwxr-xr-x - cloudera cloudera      0 2024-07-14 22:45 /user/cloudera/dezyre
drwxr-xr-x - cloudera cloudera      0 2024-07-14 23:35 /user/cloudera/dezyrel
-rw-r--r-- 1 cloudera cloudera 50515 2024-07-14 23:35 /user/cloudera/dezyrel/express-deployment.json
drwxr-xr-x - cloudera cloudera      0 2024-07-15 22:24 /user/cloudera/dezyrel
drwxr-xr-x - cloudera cloudera      0 2024-07-18 03:27 /user/cloudera/khushboo
drwxr-xr-x - cloudera cloudera      0 2024-07-18 03:26 /user/cloudera/lord
drwxr-xr-x - cloudera cloudera      0 2024-07-18 03:19 /user/cloudera/satshil
[cloudera@quickstart ~]$
```

- \$ `sudo -u hdfs hadoop fs -mkdir /DF`
This will create a new directory named 'DF'

```
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ sudo -u hdfs hadoop fs -mkdir /DF
[cloudera@quickstart ~]$
```

- \$ `hadoop fs -copyFromLocal Sample1.txt /user/cloudera/DF`
Copy/Upload Sample1.txt

```
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hadoop fs -copyFromLocal Sample1.txt /user/cloudera/DF
[cloudera@quickstart ~]$
```

- \$ ls
List all files present in the directory

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
[cloudera@quickstart ~]$ ls  
cloudera-manager cm_api.py Desktop Documents Downloads eclipse enterprise-deployment.json express-deployment.json file kerberos lib Music parcels Pictures Public Templates Videos workspace  
[cloudera@quickstart ~]$
```

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
[cloudera@quickstart ~]$ ls  
cloudera-manager Desktop Downloads enterprise-deployment.json file lib parcels Public Sample2.txt Templates workspace  
cm_api.py Documents eclipse express-deployment.json kerberos Music Pictures Sample1.txt Sample3.txt Videos  
[cloudera@quickstart ~]$
```

- \$ hadoop fs -copyFromLocal Sample2.txt /user/cloudera/DF
Copy/Upload Sample2.txt

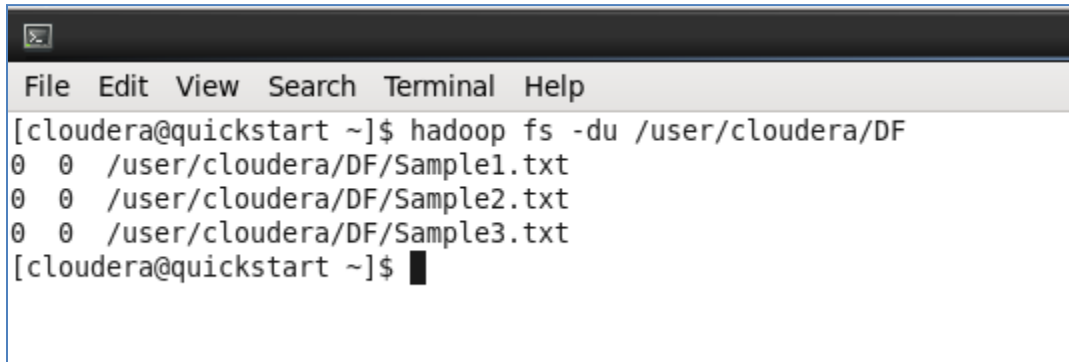
```
File Edit View Search Terminal Help  
[cloudera@quickstart ~]$ hadoop fs -put Sample2.txt /user/cloudera/DF  
[cloudera@quickstart ~]$
```

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
[cloudera@quickstart ~]$ ls  
cloudera-manager cm_api.py Desktop Documents Downloads eclipse enterprise-deployment.json express-deployment.json file kerberos lib Music parcels Pictures Public Sample1.txt Sample2.txt Templates Videos workspace  
[cloudera@quickstart ~]$
```

- \$ hadoop fs -moveFromLocal Sample3.txt /user/cloudera/DF
Move Sample3.txt

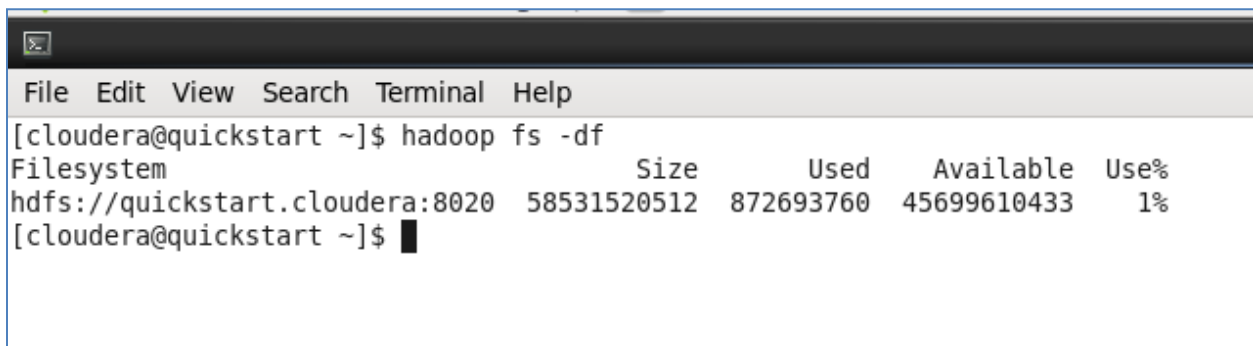
```
File Edit View Search Terminal Help  
[cloudera@quickstart ~]$ hadoop fs -moveFromLocal Sample3.txt /user/cloudera/DF  
[cloudera@quickstart ~]$
```

- \$ `hadoop fs -du /user/cloudera/DF`
Display the disk usage for all files



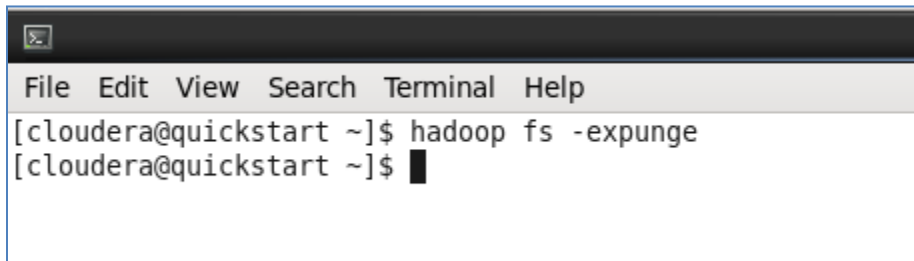
```
[cloudera@quickstart ~]$ hadoop fs -du /user/cloudera/DF
0 0 /user/cloudera/DF/Sample1.txt
0 0 /user/cloudera/DF/Sample2.txt
0 0 /user/cloudera/DF/Sample3.txt
[cloudera@quickstart ~]$
```

- \$ `hadoop fs -df`
Display disk usage of current hadoop system



```
[cloudera@quickstart ~]$ hadoop fs -df
Filesystem                                Size      Used    Available  Use%
hdfs://quickstart.cloudera:8020 58531520512 872693760 45699610433    1%
[cloudera@quickstart ~]$
```

- \$ `hadoop fs -expunge`
It empties the trash by deleting all the files



```
[cloudera@quickstart ~]$ hadoop fs -expunge
[cloudera@quickstart ~]$
```