

SMIT R PATEL

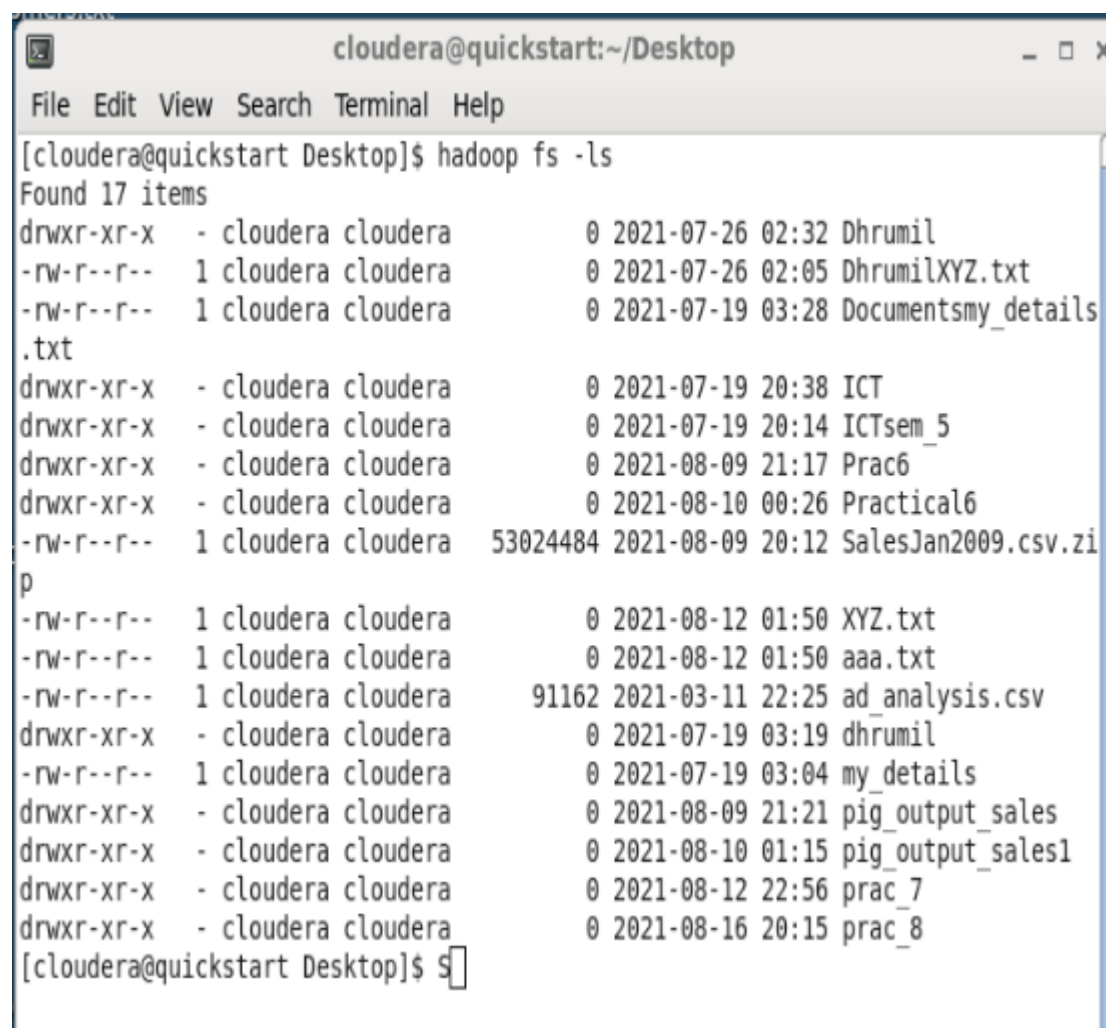
19162121031

SEM 5

BIG DATA AND ANALYTICS

PRACTICAL 1

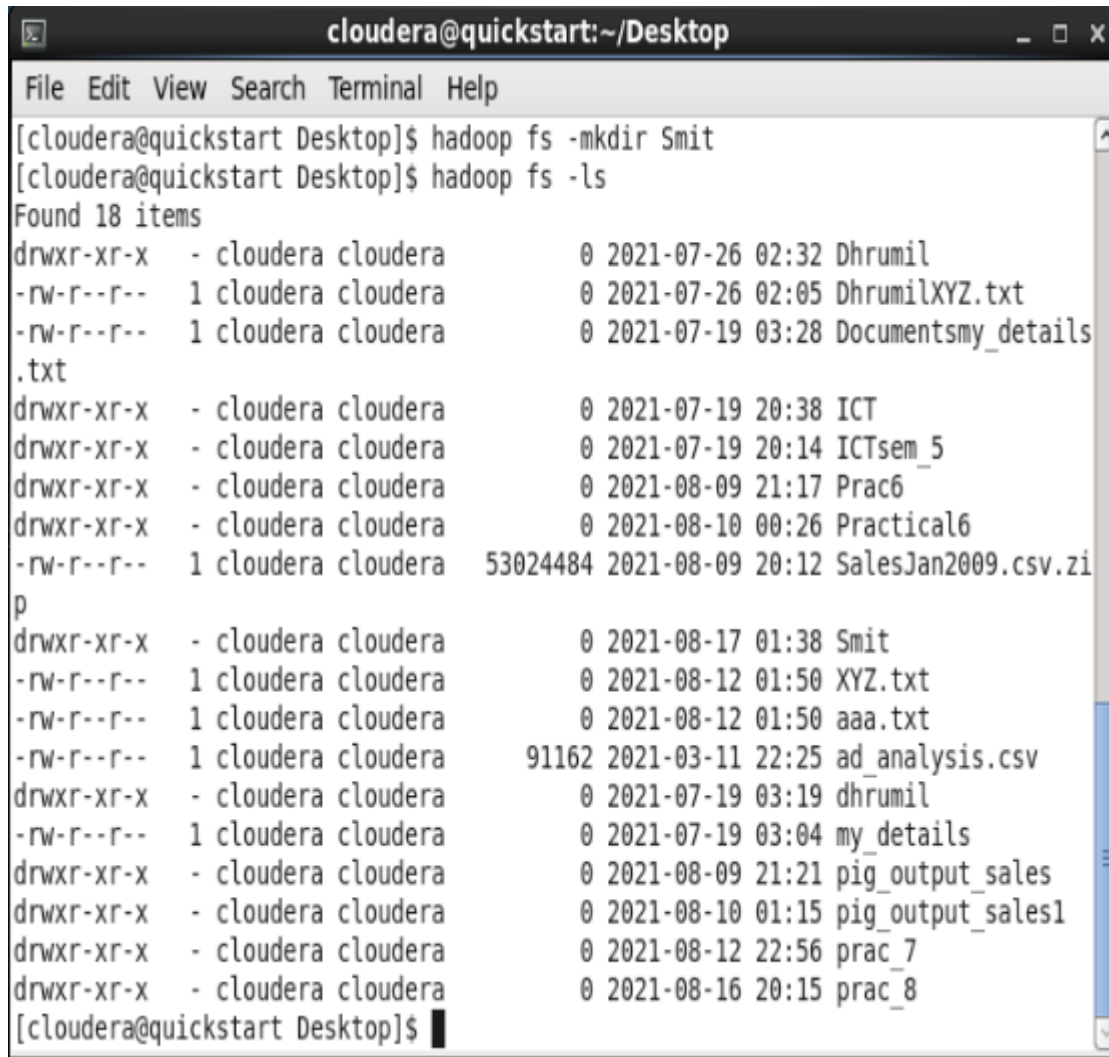
1. You are a student who is comfortable working with Linux. With the following in mind, simply try out the LS command to list all the files and directories as you would do in a Linux system. Now, try the following for listing files and directories in a Hadoop system with "hadoop fs -ls"



```
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help
[cloudera@quickstart Desktop]$ hadoop fs -ls
Found 17 items
drwxr-xr-x  - cloudera cloudera      0 2021-07-26 02:32 Dhrumil
-rw-r--r--  1 cloudera cloudera      0 2021-07-26 02:05 DhrumilXYZ.txt
-rw-r--r--  1 cloudera cloudera      0 2021-07-19 03:28 Documentsmy_details
.txt
drwxr-xr-x  - cloudera cloudera      0 2021-07-19 20:38 ICT
drwxr-xr-x  - cloudera cloudera      0 2021-07-19 20:14 ICTsem_5
drwxr-xr-x  - cloudera cloudera      0 2021-08-09 21:17 Prac6
drwxr-xr-x  - cloudera cloudera      0 2021-08-10 00:26 Practical6
-rw-r--r--  1 cloudera cloudera 53024484 2021-08-09 20:12 SalesJan2009.csv.zip
-rw-r--r--  1 cloudera cloudera      0 2021-08-12 01:50 XYZ.txt
-rw-r--r--  1 cloudera cloudera      0 2021-08-12 01:50 aaa.txt
-rw-r--r--  1 cloudera cloudera 91162 2021-03-11 22:25 ad_analysis.csv
drwxr-xr-x  - cloudera cloudera      0 2021-07-19 03:19 dhrumil
-rw-r--r--  1 cloudera cloudera      0 2021-07-19 03:04 my_details
drwxr-xr-x  - cloudera cloudera      0 2021-08-09 21:21 pig_output_sales
drwxr-xr-x  - cloudera cloudera      0 2021-08-10 01:15 pig_output_sales1
drwxr-xr-x  - cloudera cloudera      0 2021-08-12 22:56 prac_7
drwxr-xr-x  - cloudera cloudera      0 2021-08-16 20:15 prac_8
[cloudera@quickstart Desktop]$ S
```

```
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help
drwxr-xr-x - cloudera cloudera 0 2021-08-12 22:56 prac_7
drwxr-xr-x - cloudera cloudera 0 2021-08-16 20:15 prac_8
[cloudera@quickstart Desktop]$ hdfs dfs -ls
Found 17 items
drwxr-xr-x - cloudera cloudera 0 2021-07-26 02:32 Dhrumil
-rw-r--r-- 1 cloudera cloudera 0 2021-07-26 02:05 DhrumilXYZ.txt
-rw-r--r-- 1 cloudera cloudera 0 2021-07-19 03:28 Documentsmy_details
.txt
drwxr-xr-x - cloudera cloudera 0 2021-07-19 20:38 ICT
drwxr-xr-x - cloudera cloudera 0 2021-07-19 20:14 ICTsem_5
drwxr-xr-x - cloudera cloudera 0 2021-08-09 21:17 Prac6
drwxr-xr-x - cloudera cloudera 0 2021-08-10 00:26 Practical6
-rw-r--r-- 1 cloudera cloudera 53024484 2021-08-09 20:12 SalesJan2009.csv.zip
-rw-r--r-- 1 cloudera cloudera 0 2021-08-12 01:50 XYZ.txt
-rw-r--r-- 1 cloudera cloudera 0 2021-08-12 01:50 aaa.txt
-rw-r--r-- 1 cloudera cloudera 91162 2021-03-11 22:25 ad_analysis.csv
drwxr-xr-x - cloudera cloudera 0 2021-07-19 03:19 dhrumil
-rw-r--r-- 1 cloudera cloudera 0 2021-07-19 03:04 my_details
drwxr-xr-x - cloudera cloudera 0 2021-08-09 21:21 pig_output_sales
drwxr-xr-x - cloudera cloudera 0 2021-08-10 01:15 pig_output_sales1
drwxr-xr-x - cloudera cloudera 0 2021-08-12 22:56 prac_7
drwxr-xr-x - cloudera cloudera 0 2021-08-16 20:15 prac_8
[cloudera@quickstart Desktop]$
```

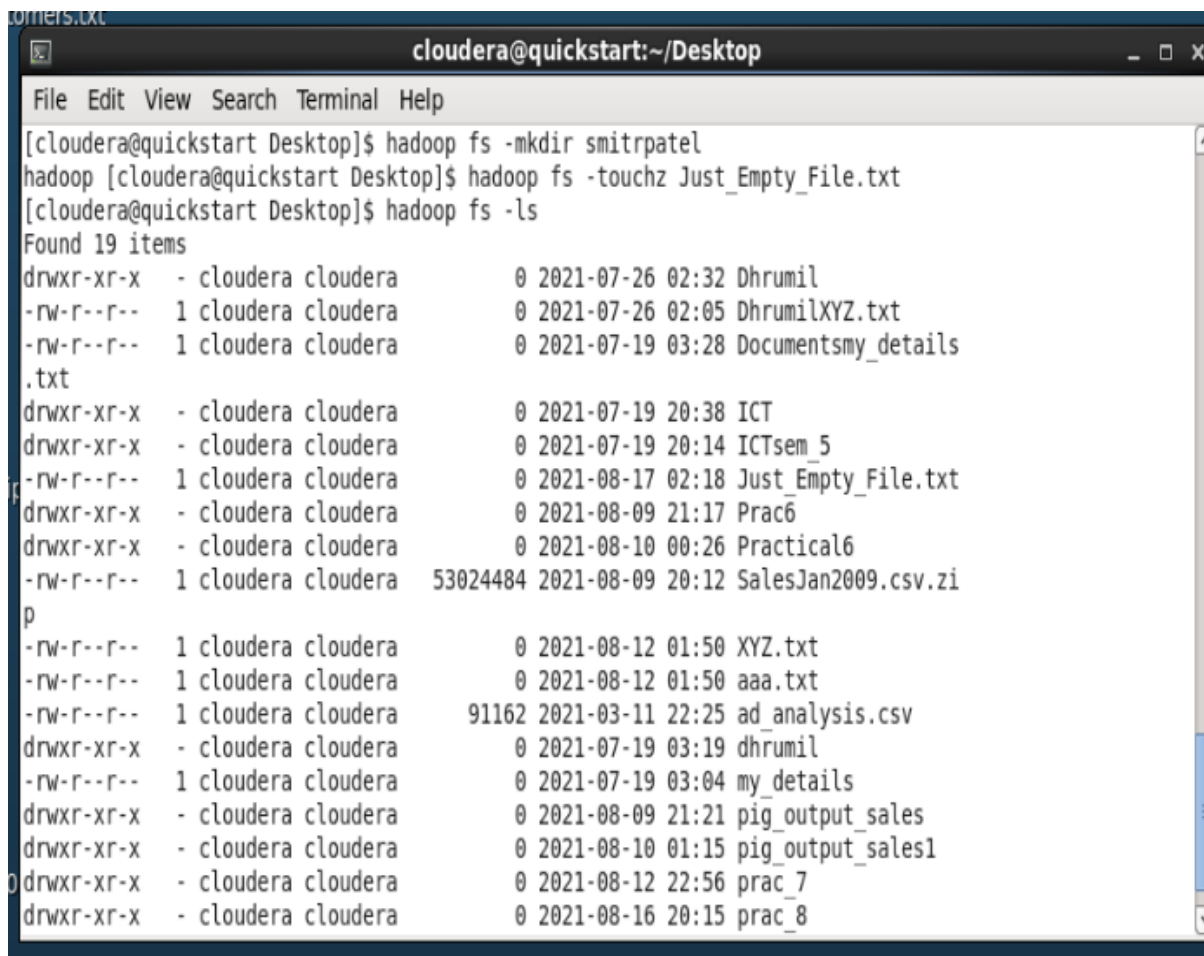
2 You are now aware you cannot change directories in Hadoop . Hence, create a folder on your Hadoop home directory with your name.

A terminal window titled 'cloudera@quickstart:~/Desktop' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal shows the execution of 'hadoop fs -mkdir Smit' followed by 'hadoop fs -ls'. The output lists 18 items in the Hadoop file system, including directories like 'DhruMil', 'ICT', 'Prac6', 'Practical6', 'SalesJan2009.csv.zip', 'Smit', 'XYZ.txt', 'aaa.txt', 'ad\_analysis.csv', 'my\_details', 'pig\_output\_sales', 'pig\_output\_sales1', 'prac\_7', and 'prac\_8'. Each entry shows permissions, owner, group, size, and timestamp.

```
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help
[cloudera@quickstart Desktop]$ hadoop fs -mkdir Smit
[cloudera@quickstart Desktop]$ hadoop fs -ls
Found 18 items
drwxr-xr-x  - cloudera cloudera      0 2021-07-26 02:32 DhruMil
-rw-r--r--  1 cloudera cloudera      0 2021-07-26 02:05 DhruMilXYZ.txt
-rw-r--r--  1 cloudera cloudera      0 2021-07-19 03:28 Documentsmy_details
.txt
drwxr-xr-x  - cloudera cloudera      0 2021-07-19 20:38 ICT
drwxr-xr-x  - cloudera cloudera      0 2021-07-19 20:14 ICTsem_5
drwxr-xr-x  - cloudera cloudera      0 2021-08-09 21:17 Prac6
drwxr-xr-x  - cloudera cloudera      0 2021-08-10 00:26 Practical6
-rw-r--r--  1 cloudera cloudera 53024484 2021-08-09 20:12 SalesJan2009.csv.zi
p
drwxr-xr-x  - cloudera cloudera      0 2021-08-17 01:38 Smit
-rw-r--r--  1 cloudera cloudera      0 2021-08-12 01:50 XYZ.txt
-rw-r--r--  1 cloudera cloudera      0 2021-08-12 01:50 aaa.txt
-rw-r--r--  1 cloudera cloudera 91162 2021-03-11 22:25 ad_analysis.csv
drwxr-xr-x  - cloudera cloudera      0 2021-07-19 03:19 dhruMil
-rw-r--r--  1 cloudera cloudera      0 2021-07-19 03:04 my_details
drwxr-xr-x  - cloudera cloudera      0 2021-08-09 21:21 pig_output_sales
drwxr-xr-x  - cloudera cloudera      0 2021-08-10 01:15 pig_output_sales1
drwxr-xr-x  - cloudera cloudera      0 2021-08-12 22:56 prac_7
drwxr-xr-x  - cloudera cloudera      0 2021-08-16 20:15 prac_8
[cloudera@quickstart Desktop]$
```

```
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help
drwxr-xr-x - cloudera cloudera 0 2021-08-16 20:15 prac_8
[cloudera@quickstart Desktop]$ hadoop fs -rmdir Smit
[cloudera@quickstart Desktop]$ hadoop fs -ls
Found 17 items
drwxr-xr-x - cloudera cloudera 0 2021-07-26 02:32 Dhrumil
-rw-r--r-- 1 cloudera cloudera 0 2021-07-26 02:05 DhrumilXYZ.txt
-rw-r--r-- 1 cloudera cloudera 0 2021-07-19 03:28 Documentsmy_details
.txt
drwxr-xr-x - cloudera cloudera 0 2021-07-19 20:38 ICT
drwxr-xr-x - cloudera cloudera 0 2021-07-19 20:14 ICTsem_5
drwxr-xr-x - cloudera cloudera 0 2021-08-09 21:17 Prac6
drwxr-xr-x - cloudera cloudera 0 2021-08-10 00:26 Practical6
-rw-r--r-- 1 cloudera cloudera 53024484 2021-08-09 20:12 SalesJan2009.csv.zip
-rw-r--r-- 1 cloudera cloudera 0 2021-08-12 01:50 XYZ.txt
-rw-r--r-- 1 cloudera cloudera 0 2021-08-12 01:50 aaa.txt
-rw-r--r-- 1 cloudera cloudera 91162 2021-03-11 22:25 ad_analysis.csv
drwxr-xr-x - cloudera cloudera 0 2021-07-19 03:19 dhrumil
-rw-r--r-- 1 cloudera cloudera 0 2021-07-19 03:04 my_details
drwxr-xr-x - cloudera cloudera 0 2021-08-09 21:21 pig_output_sales
drwxr-xr-x - cloudera cloudera 0 2021-08-10 01:15 pig_output_sales1
drwxr-xr-x - cloudera cloudera 0 2021-08-12 22:56 prac_7
drwxr-xr-x - cloudera cloudera 0 2021-08-16 20:15 prac_8
[cloudera@quickstart Desktop]$
```

3. Create an empty file in your Hadoop system using the touchz command.



```
cloudera@quickstart:~/Desktop
File Edit View Search Terminal Help
[cloudera@quickstart Desktop]$ hadoop fs -mkdir smitrpatel
hadoop [cloudera@quickstart Desktop]$ hadoop fs -touchz Just_Empty_File.txt
[cloudera@quickstart Desktop]$ hadoop fs -ls
Found 19 items
drwxr-xr-x - cloudera cloudera      0 2021-07-26 02:32 Dhrumil
-rw-r--r-- 1 cloudera cloudera      0 2021-07-26 02:05 DhrumilXYZ.txt
-rw-r--r-- 1 cloudera cloudera      0 2021-07-19 03:28 Documentsmy_details
.txt
drwxr-xr-x - cloudera cloudera      0 2021-07-19 20:38 ICT
drwxr-xr-x - cloudera cloudera      0 2021-07-19 20:14 ICTsem_5
-rw-r--r-- 1 cloudera cloudera      0 2021-08-17 02:18 Just_Empty_File.txt
drwxr-xr-x - cloudera cloudera      0 2021-08-09 21:17 Prac6
drwxr-xr-x - cloudera cloudera      0 2021-08-10 00:26 Practical6
-rw-r--r-- 1 cloudera cloudera 53024484 2021-08-09 20:12 SalesJan2009.csv.zi
p
-rw-r--r-- 1 cloudera cloudera      0 2021-08-12 01:50 XYZ.txt
-rw-r--r-- 1 cloudera cloudera      0 2021-08-12 01:50 aaa.txt
-rw-r--r-- 1 cloudera cloudera 91162 2021-03-11 22:25 ad_analysis.csv
drwxr-xr-x - cloudera cloudera      0 2021-07-19 03:19 dhrumil
-rw-r--r-- 1 cloudera cloudera      0 2021-07-19 03:04 my_details
drwxr-xr-x - cloudera cloudera      0 2021-08-09 21:21 pig_output_sales
drwxr-xr-x - cloudera cloudera      0 2021-08-10 01:15 pig_output_sales1
drwxr-xr-x - cloudera cloudera      0 2021-08-12 22:56 prac_7
drwxr-xr-x - cloudera cloudera      0 2021-08-16 20:15 prac_8
```

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
[cloudera@quickstart Desktop]$ cd ..  
[cloudera@quickstart ~]$ hadoop fs -touchz Just_Empty_File.txt  
[cloudera@quickstart ~]$ hadoop fs -ls  
Found 19 items  
drwxr-xr-x - cloudera cloudera 0 2021-07-26 02:32 Dhrumil  
-rw-r--r-- 1 cloudera cloudera 0 2021-07-26 02:05 DhrumilXYZ.txt  
-rw-r--r-- 1 cloudera cloudera 0 2021-07-19 03:28 Documentsmy_details.txt  
drwxr-xr-x - cloudera cloudera 0 2021-07-19 20:38 ICT  
drwxr-xr-x - cloudera cloudera 0 2021-07-19 20:14 ICTsem_5  
-rw-r--r-- 1 cloudera cloudera 0 2021-08-17 02:20 Just_Empty_File.txt  
drwxr-xr-x - cloudera cloudera 0 2021-08-09 21:17 Prac6  
drwxr-xr-x - cloudera cloudera 0 2021-08-10 00:26 Practical6  
-rw-r--r-- 1 cloudera cloudera 53024484 2021-08-09 20:12 SalesJan2009.csv.zip  
-rw-r--r-- 1 cloudera cloudera 0 2021-08-12 01:50 XYZ.txt  
-rw-r--r-- 1 cloudera cloudera 0 2021-08-12 01:50 aaa.txt  
-rw-r--r-- 1 cloudera cloudera 91162 2021-03-11 22:25 ad_analysis.csv  
drwxr-xr-x - cloudera cloudera 0 2021-07-19 03:19 dhrumil  
-rw-r--r-- 1 cloudera cloudera 0 2021-07-19 03:04 my_details  
drwxr-xr-x - cloudera cloudera 0 2021-08-09 21:21 pig_output_sales  
drwxr-xr-x - cloudera cloudera 0 2021-08-10 01:15 pig_output_sales1  
drwxr-xr-x - cloudera cloudera 0 2021-08-12 22:56 prac_7  
drwxr-xr-x - cloudera cloudera 0 2021-08-16 20:15 prac_8  
drwxr-xr-x - cloudera cloudera 0 2021-08-17 02:17 smitpatel  
[cloudera@quickstart ~]$
```



4. Create a sample file in your LOCAL SYSTEM and name it- ABC, and execute the “copyFromLocal” command in your Hadoop system to copy this file to the folder you created in your name.

```
cloudera@quickstart:~  
File Edit View Search Terminal Help  
drwxr-xr-x - cloudera cloudera 0 2021-08-09 21:17 Prac6  
drwxr-xr-x - cloudera cloudera 0 2021-08-10 00:26 Practical6  
-rw-r--r-- 1 cloudera cloudera 53024484 2021-08-09 20:12 SalesJan2009.csv.zi  
p  
-rw-r--r-- 1 cloudera cloudera 0 2021-08-12 01:50 XYZ.txt  
-rw-r--r-- 1 cloudera cloudera 0 2021-08-12 01:50 aaa.txt  
-rw-r--r-- 1 cloudera cloudera 91162 2021-03-11 22:25 ad_analysis.csv  
drwxr-xr-x - cloudera cloudera 0 2021-07-19 03:19 dhrumil  
-rw-r--r-- 1 cloudera cloudera 0 2021-07-19 03:04 my_details  
drwxr-xr-x - cloudera cloudera 0 2021-08-09 21:21 pig_output_sales  
drwxr-xr-x - cloudera cloudera 0 2021-08-10 01:15 pig_output_sales1  
drwxr-xr-x - cloudera cloudera 0 2021-08-12 22:56 prac_7  
drwxr-xr-x - cloudera cloudera 0 2021-08-16 20:15 prac_8  
drwxr-xr-x - cloudera cloudera 0 2021-08-17 02:21 smitrpatel  
[cloudera@quickstart ~]$ hadoop fs -copyFromLocal /home/cloudera/Desktop/ABC.txt  
smitrpatel  
[cloudera@quickstart ~]$ hadoop fs -ls smitrpatel  
Found 2 items  
-rw-r--r-- 1 cloudera cloudera 6 2021-08-19 00:43 smitrpatel/ABC.txt  
-rw-r--r-- 1 cloudera cloudera 0 2021-08-17 02:21 smitrpatel/Just_Emp  
ty File.txt  
[cloudera@quickstart ~]$ hadoop fs -copyFromLocal /home/cloudera/Desktop/ABC.txt smitrpatel/  
copyFromLocal: `smitrpatel/ABC.txt': File exists  
[cloudera@quickstart ~]$
```