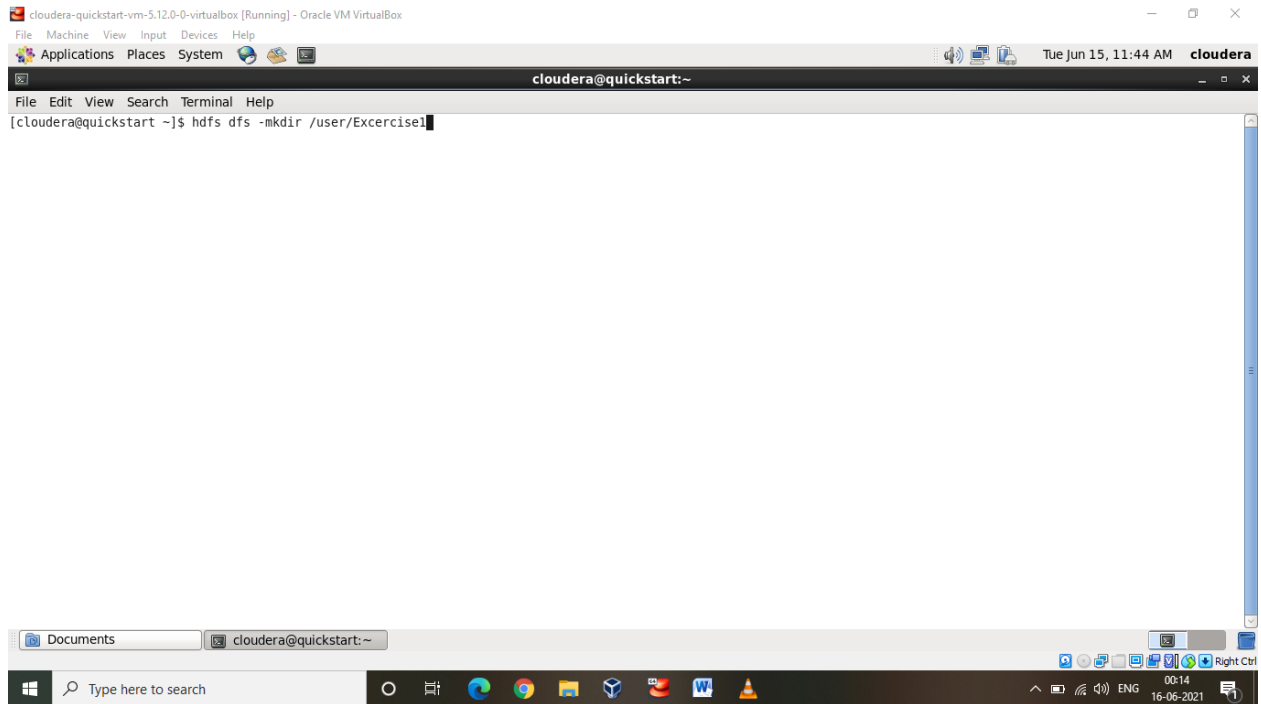


HDFS Task1

Perform

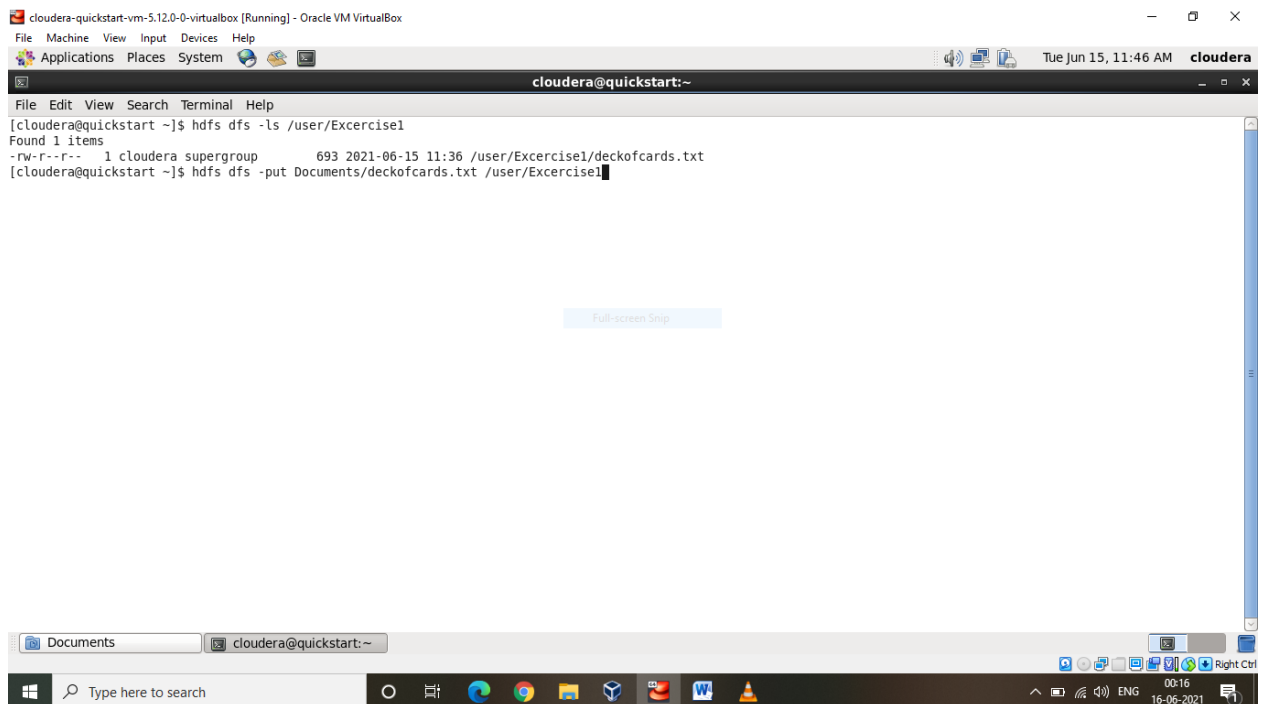
1. Start HDFS and verify that it's running
2. Create a new directory /exercise1 on HDFS



The screenshot shows a terminal window titled 'cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox'. The terminal prompt is 'cloudera@quickstart:~'. The command 'hdfs dfs -mkdir /user/Exercise1' has been entered and executed. The window also shows a Windows taskbar at the bottom with various application icons and a search bar.

```
cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hdfs dfs -mkdir /user/Exercise1
```

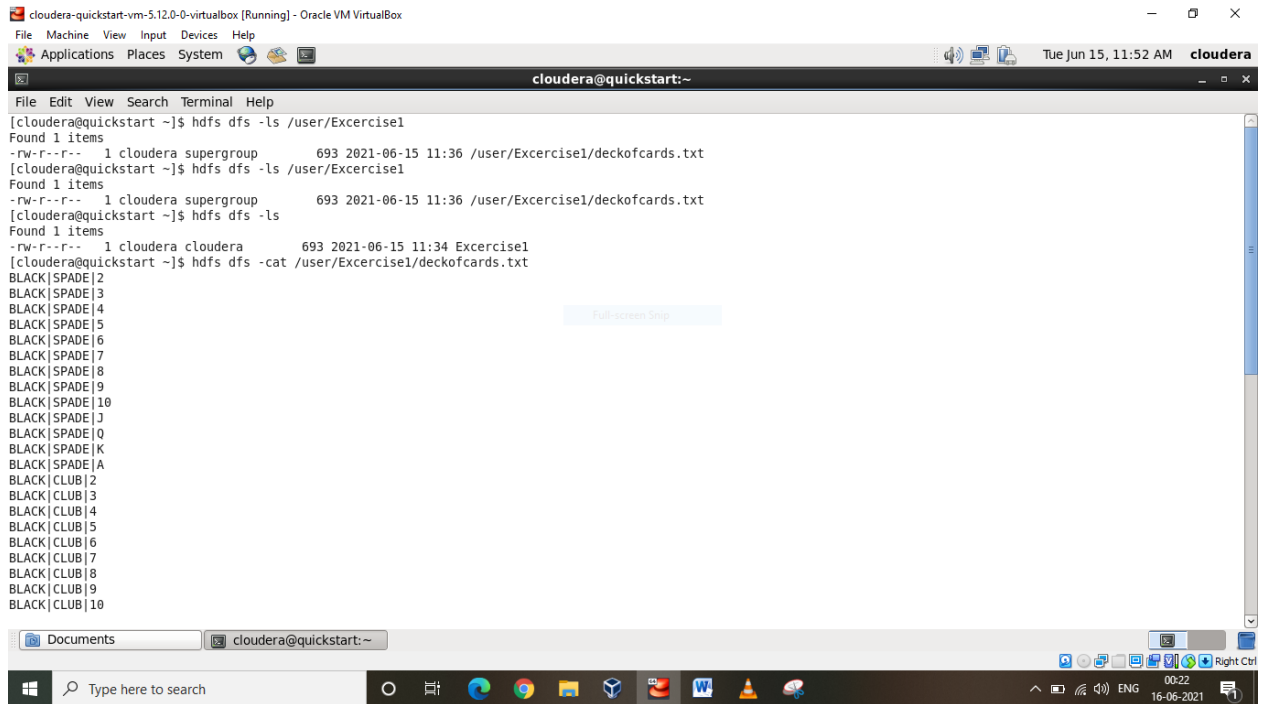
3. Upload GitHub repo sample_data/deckofcards.txt to HDFS under /exercise1 directory



The screenshot shows a terminal window titled 'cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox'. The terminal prompt is 'cloudera@quickstart:~'. The command 'hdfs dfs -ls /user/Exercise1' has been entered and executed, showing the output 'Found 1 items' and file details. The command 'hdfs dfs -put Documents/deckofcards.txt /user/Exercise1' has also been entered and executed. The window also shows a Windows taskbar at the bottom with various application icons and a search bar.

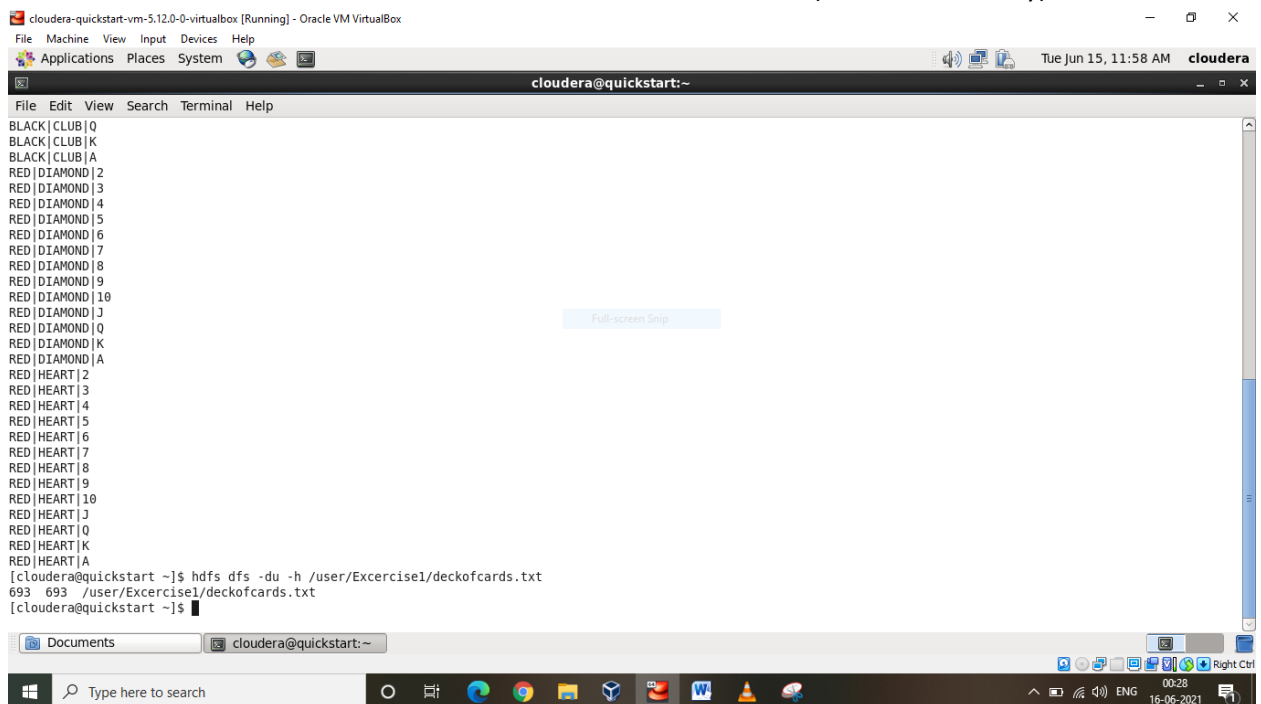
```
cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hdfs dfs -ls /user/Exercise1
Found 1 items
-rw-r--r-- 1 cloudera supergroup      693 2021-06-15 11:36 /user/Exercise1/deckofcards.txt
[cloudera@quickstart ~]$ hdfs dfs -put Documents/deckofcards.txt /user/Exercise1
```

4. View the content of the /exercise1 directory



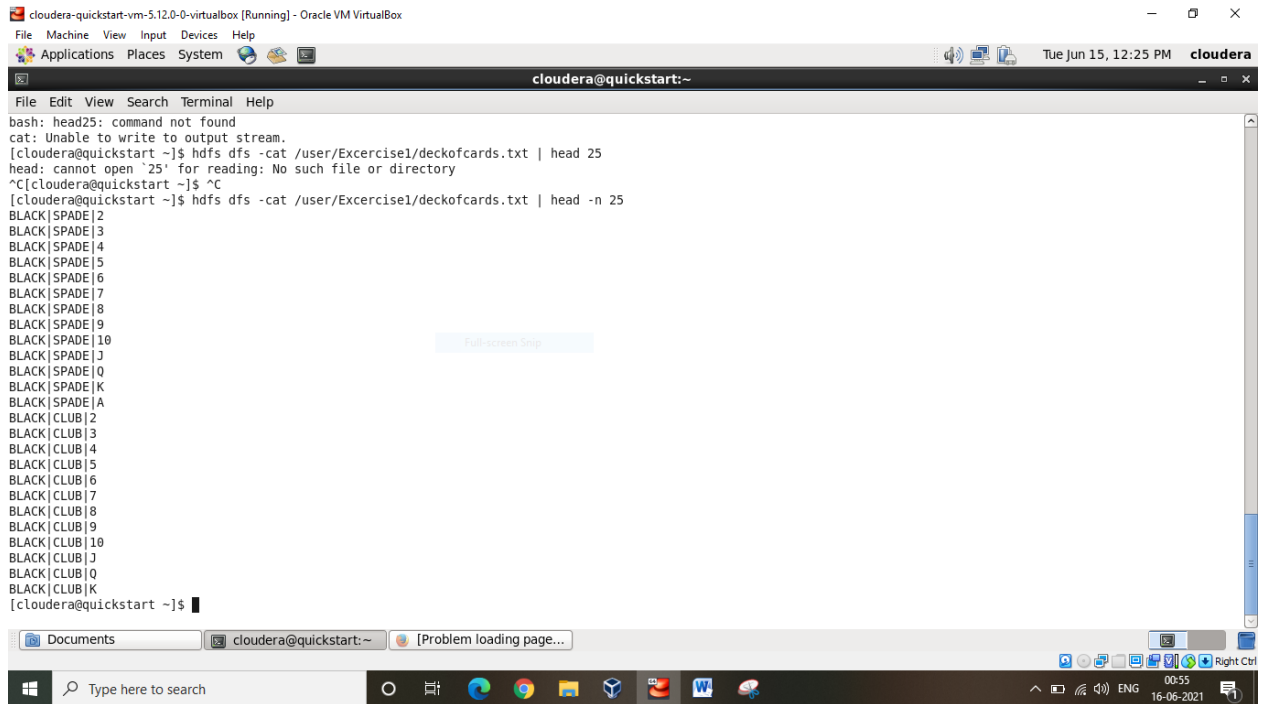
```
cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hdfs dfs -ls /user/Exercise1
Found 1 items
-rw-r--r-- 1 cloudera supergroup      693 2021-06-15 11:36 /user/Exercise1/deckofcards.txt
[cloudera@quickstart ~]$ hdfs dfs -ls /user/Exercise1
Found 1 items
-rw-r--r-- 1 cloudera supergroup      693 2021-06-15 11:36 /user/Exercise1/deckofcards.txt
[cloudera@quickstart ~]$ hdfs dfs -ls
Found 1 items
-rw-r--r-- 1 cloudera cloudera        693 2021-06-15 11:34 Exercise1
[cloudera@quickstart ~]$ hdfs dfs -cat /user/Exercise1/deckofcards.txt
BLACK|SPADE|2
BLACK|SPADE|3
BLACK|SPADE|4
BLACK|SPADE|5
BLACK|SPADE|6
BLACK|SPADE|7
BLACK|SPADE|8
BLACK|SPADE|9
BLACK|SPADE|10
BLACK|SPADE|J
BLACK|SPADE|K
BLACK|SPADE|A
BLACK|CLUB|2
BLACK|CLUB|3
BLACK|CLUB|4
BLACK|CLUB|5
BLACK|CLUB|6
BLACK|CLUB|7
BLACK|CLUB|8
BLACK|CLUB|9
BLACK|CLUB|10
```

5. Determine the size of the hamlet.txt file in KB that resides on HDFS (not local directory)



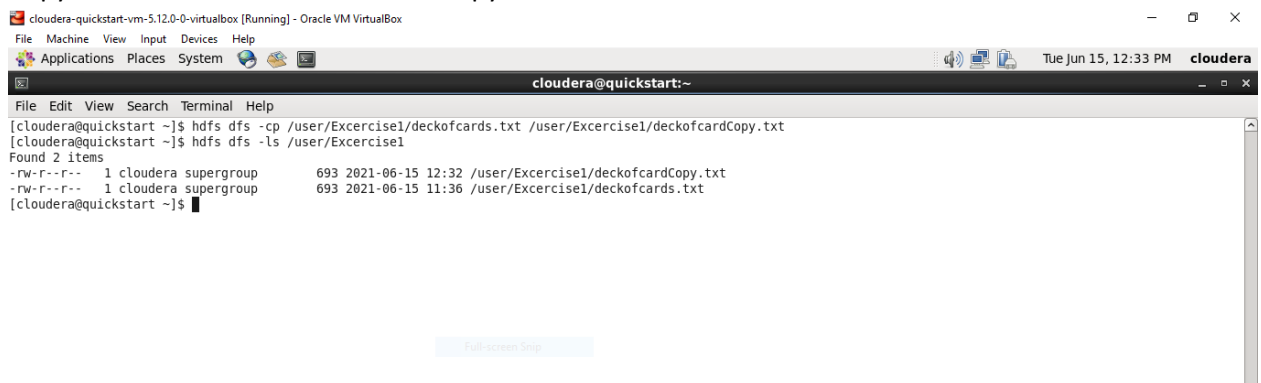
```
cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
BLACK|CLUB|Q
BLACK|CLUB|K
BLACK|CLUB|A
RED|DIAMOND|2
RED|DIAMOND|3
RED|DIAMOND|4
RED|DIAMOND|5
RED|DIAMOND|6
RED|DIAMOND|7
RED|DIAMOND|8
RED|DIAMOND|9
RED|DIAMOND|10
RED|DIAMOND|J
RED|DIAMOND|Q
RED|DIAMOND|K
RED|DIAMOND|A
RED|HEART|2
RED|HEART|3
RED|HEART|4
RED|HEART|5
RED|HEART|6
RED|HEART|7
RED|HEART|8
RED|HEART|9
RED|HEART|10
RED|HEART|J
RED|HEART|Q
RED|HEART|K
RED|HEART|A
[cloudera@quickstart ~]$ hdfs dfs -du -h /user/Exercise1/deckofcards.txt
693 693 /user/Exercise1/deckofcards.txt
[cloudera@quickstart ~]$
```

6. Print the first 25 lines to the screen from deckofcards.txt on HDFS



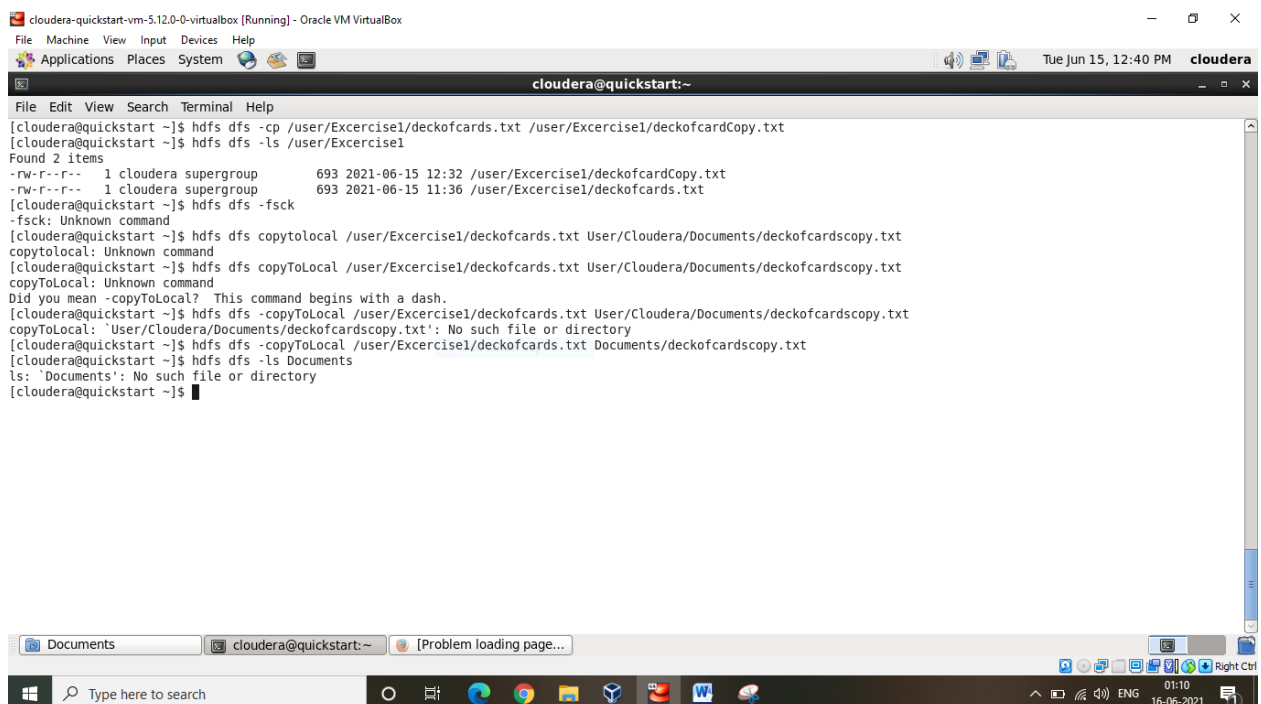
```
cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
bash: head25: command not found
cat: Unable to write to output stream.
[cloudera@quickstart ~]$ hdfs dfs -cat /user/Exercisel/deckofcards.txt | head 25
head: cannot open '25' for reading: No such file or directory
^C[cloudera@quickstart ~]$ ^C
[cloudera@quickstart ~]$ hdfs dfs -cat /user/Exercisel/deckofcards.txt | head -n 25
BLACK|SPADE|2
BLACK|SPADE|3
BLACK|SPADE|4
BLACK|SPADE|5
BLACK|SPADE|6
BLACK|SPADE|7
BLACK|SPADE|8
BLACK|SPADE|9
BLACK|SPADE|10
BLACK|SPADE|J
BLACK|SPADE|Q
BLACK|SPADE|K
BLACK|SPADE|A
BLACK|CLUB|2
BLACK|CLUB|3
BLACK|CLUB|4
BLACK|CLUB|5
BLACK|CLUB|6
BLACK|CLUB|7
BLACK|CLUB|8
BLACK|CLUB|9
BLACK|CLUB|10
BLACK|CLUB|J
BLACK|CLUB|Q
BLACK|CLUB|K
[cloudera@quickstart ~]$
```

7. Copy deckofcards.txt to deckofcardsCopy.txt



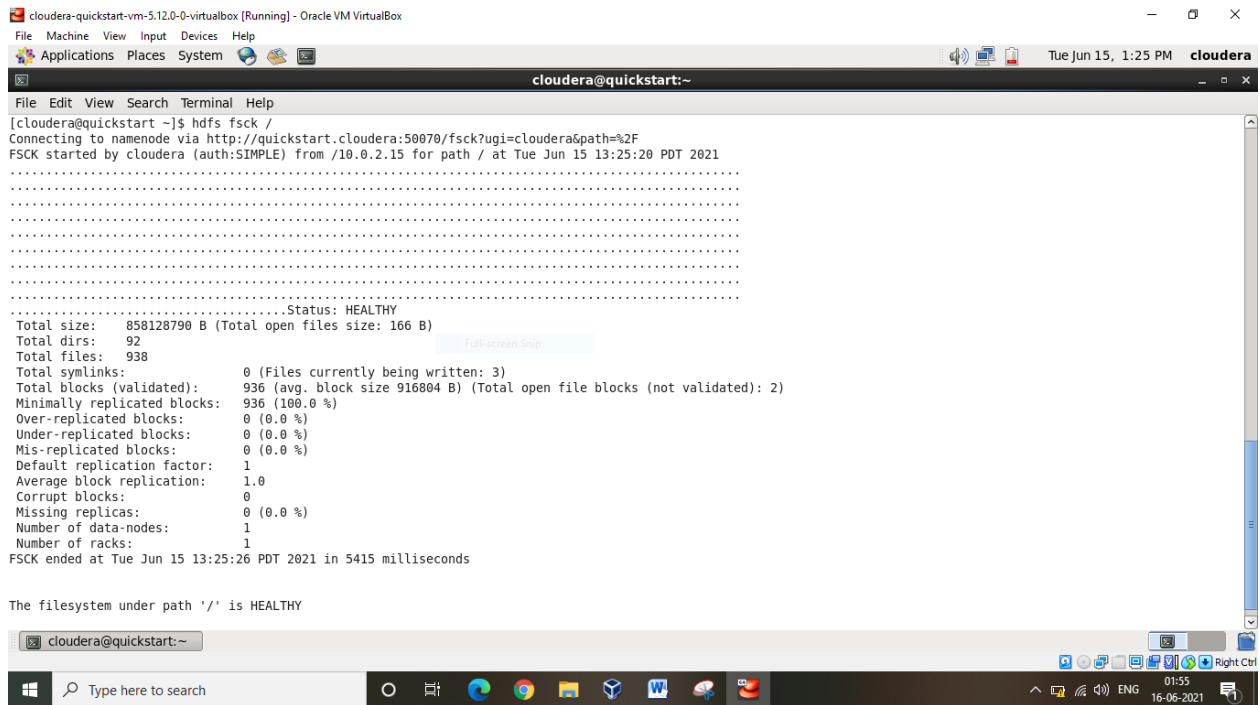
```
cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hdfs dfs -cp /user/Exercisel/deckofcards.txt /user/Exercisel/deckofcardCopy.txt
[cloudera@quickstart ~]$ hdfs dfs -ls /user/Exercisel
Found 2 items
-rw-r--r-- 1 cloudera supergroup 693 2021-06-15 12:32 /user/Exercisel/deckofcardCopy.txt
-rw-r--r-- 1 cloudera supergroup 693 2021-06-15 11:36 /user/Exercisel/deckofcards.txt
[cloudera@quickstart ~]$
```

8. Copy deckofcards.txt back to local file system and name it deckofcards.copy.txt

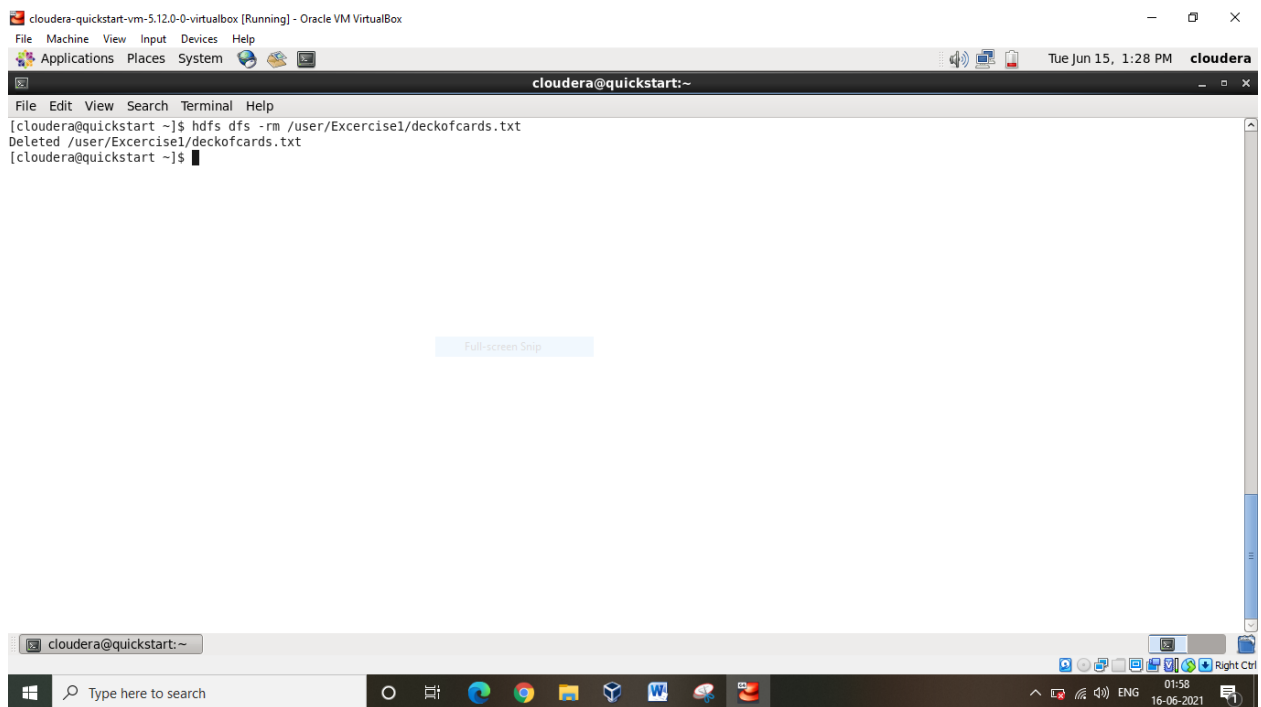


```
cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hdfs dfs -cp /user/Exercisel/deckofcards.txt /user/Exercisel/deckofcardCopy.txt
[cloudera@quickstart ~]$ hdfs dfs -ls /user/Exercisel
Found 2 items
-rw-r--r-- 1 cloudera supergroup 693 2021-06-15 12:32 /user/Exercisel/deckofcardCopy.txt
-rw-r--r-- 1 cloudera supergroup 693 2021-06-15 11:36 /user/Exercisel/deckofcards.txt
[cloudera@quickstart ~]$ hdfs dfs -fsck
-fsck: Unknown command
[cloudera@quickstart ~]$ hdfs dfs copyToLocal /user/Exercisel/deckofcards.txt User/Cloudera/Documents/deckofcardscopy.txt
copyToLocal: Unknown command
[cloudera@quickstart ~]$ hdfs dfs copyToLocal /user/Exercisel/deckofcards.txt User/Cloudera/Documents/deckofcardscopy.txt
copyToLocal: Unknown command
Did you mean -copyToLocal? This command begins with a dash.
[cloudera@quickstart ~]$ hdfs dfs -copyToLocal /user/Exercisel/deckofcards.txt User/Cloudera/Documents/deckofcardscopy.txt
copyToLocal: 'User/Cloudera/Documents/deckofcardscopy.txt': No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -copyToLocal /user/Exercisel/deckofcards.txt Documents/deckofcardscopy.txt
[cloudera@quickstart ~]$ hdfs dfs -ls Documents
ls: 'Documents': No such file or directory
[cloudera@quickstart ~]$
```

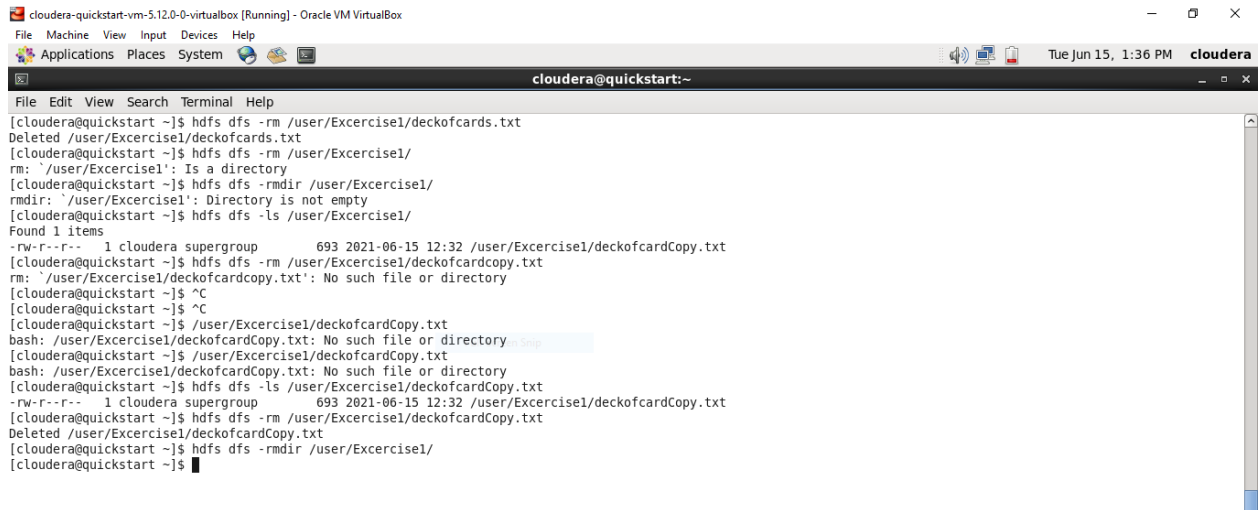
9. Check the entire filesystem for inconsistencies/problems



10. Delete deckofcards.txt from HDFS

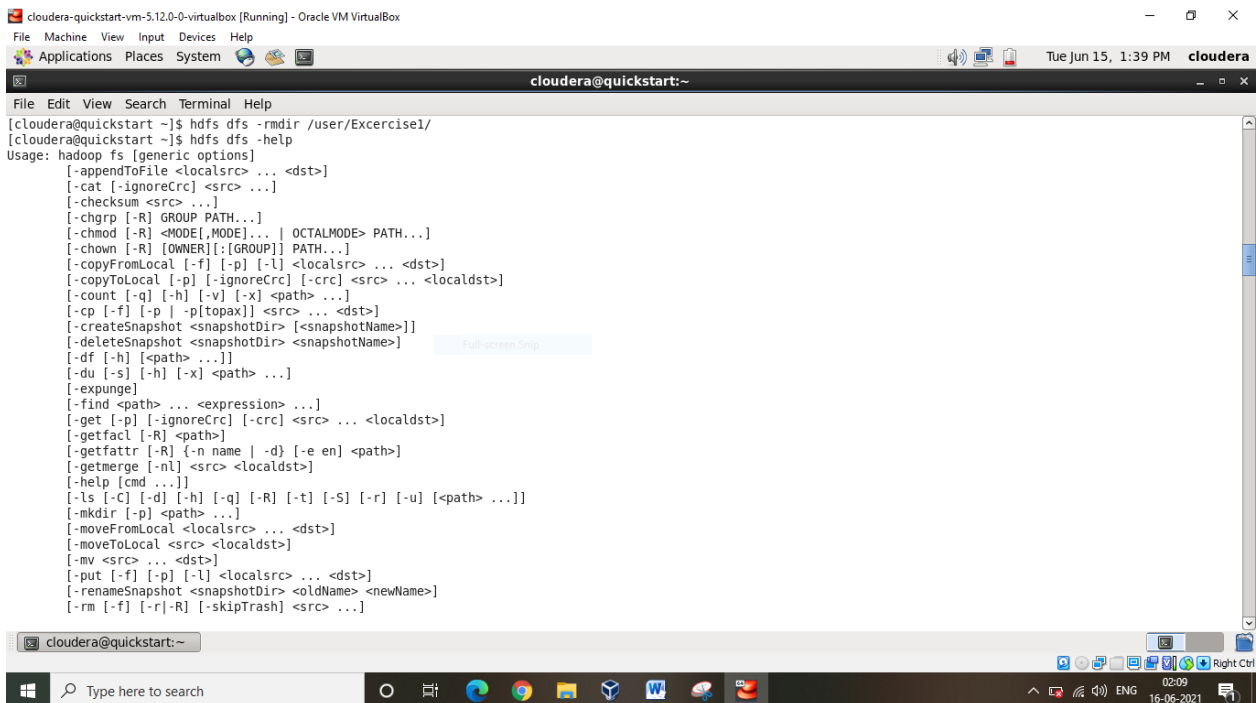


11. Delete the /exercise1 directory from HDFS



```
cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hdfs dfs -rm /user/Excercise1/deckofcards.txt
Deleted /user/Excercise1/deckofcards.txt
[cloudera@quickstart ~]$ hdfs dfs -rm /user/Excercise1/
rm: /user/Excercise1: Is a directory
[cloudera@quickstart ~]$ hdfs dfs -rmdir /user/Excercise1/
rmdir: /user/Excercise1: Directory is not empty
[cloudera@quickstart ~]$ hdfs dfs -ls /user/Excercise1/
Found 1 items
-rw-r--r-- 1 cloudera supergroup 693 2021-06-15 12:32 /user/Excercise1/deckofcardCopy.txt
[cloudera@quickstart ~]$ hdfs dfs -rm /user/Excercise1/deckofcardCopy.txt
rm: /user/Excercise1/deckofcardCopy.txt: No such file or directory
[cloudera@quickstart ~]$ ^C
[cloudera@quickstart ~]$ ^C
[cloudera@quickstart ~]$ /user/Excercise1/deckofcardCopy.txt
bash: /user/Excercise1/deckofcardCopy.txt: No such file or directory
[cloudera@quickstart ~]$ /user/Excercise1/deckofcardCopy.txt
bash: /user/Excercise1/deckofcardCopy.txt: No such file or directory
[cloudera@quickstart ~]$ hdfs dfs -ls /user/Excercise1/deckofcardCopy.txt
-rw-r--r-- 1 cloudera supergroup 693 2021-06-15 12:32 /user/Excercise1/deckofcardCopy.txt
[cloudera@quickstart ~]$ hdfs dfs -rm /user/Excercise1/deckofcardCopy.txt
Deleted /user/Excercise1/deckofcardCopy.txt
[cloudera@quickstart ~]$ hdfs dfs -rmdir /user/Excercise1/
[cloudera@quickstart ~]$
```

12. Take a second to look at other available shell options



```
cloudera-quickstart-vm-5.12.0-0-virtualbox [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places System
cloudera@quickstart:~
File Edit View Search Terminal Help
[cloudera@quickstart ~]$ hdfs dfs -rmdir /user/Excercise1/
[cloudera@quickstart ~]$ hdfs dfs -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> ...]
```