

# BIG DATA LAB

Name: VANSH SHARMA

USN: 1NT19IS181

Batch: C2

Date:26/05/2022

Tasks To complete

To start with hadoop

>hadoop version

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop version
Hadoop 3.2.1
Source code repository https://gitbox.apache.org/repos/asf/hadoop.git -r b3cbbb467e22ea829b3808f4b7b01d07e0bf3842
Compiled by rohithsharmaks on 2019-09-10T15:56Z
Compiled with protoc 2.5.0
From source with checksum 776eaf9eee9c0ffc370bcbc1888737
This command was run using /home/hadoop/hadoop-3.2.1/share/hadoop/common/hadoop-common-3.2.1.jar
```

>start-dfs.sh

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [admin1-HP-280-G4-MT-Business-PC]
```

1. Create a directory named by your name

>hadoop fs -mkdir /vansh

>hadoop fs -ls /

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -mkdir /vansh
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -ls /
Found 3 items
drwxrwxr-x - hadoop supergroup 0 2021-07-28 12:58 /tmp
drwxr-xr-x - hadoop supergroup 0 2021-07-28 12:56 /user
drwxr-xr-x - hadoop supergroup 0 2022-06-02 15:06 /vansh
```

2. Add two empty files into it

>hadoop fs -touchz /vansh/one.txt

>hadoop fs -touchz /vansh/two.txt

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -touchz /vansh/one.txt
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -touchz /vansh/two.txt
```

3. List all the files in the directory that you have created

>hadoop fs -ls /vansh

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -ls /vansh
Found 2 items
drwxr-xr-x   - hadoop supergroup          0 2022-06-02 15:15 /vansh/1NT19IS181
-rw-r--r--   1 hadoop supergroup          0 2022-06-02 15:07 /vansh/one.txt
```

4. Create one more directory by your usn and put that inside the directory of your name

>hadoop fs -mkdir /vansh/1NT19IS181

>hadoop fs -ls /vansh

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -mkdir /vansh/1NT19IS181
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -ls /vansh
Found 3 items
drwxr-xr-x   - hadoop supergroup          0 2022-06-02 15:09 /vansh/1NT19IS181
-rw-r--r--   1 hadoop supergroup          0 2022-06-02 15:07 /vansh/one.txt
-rw-r--r--   1 hadoop supergroup          0 2022-06-02 15:07 /vansh/two.txt
```

5. Create a file in your local system by the name\_usn.txt and add the contents as "Syntax to list ls command is hadoop fs -ls "

>cat > vansh\_1NT19IS181.txt

>"Syntax to list ls command is hadoop fs -ls"

>hadoop fs -touchz /vansh/1NT19IS181/1NT19IS181\_vansh.txt

>hadoop fs -ls /vansh/1NT19IS181

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -touchz /vansh/1NT19IS181/1NT19IS181_vansh.txt
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -ls /vansh/1NT19IS181
Found 1 items
-rw-r--r--   1 hadoop supergroup          0 2022-06-02 15:12 /vansh/1NT19IS181/1NT19IS181_vansh.txt
```

➤Put empty file of hdfs into the local file system

>hadoop fs -copyToLocal /vansh/1NT19IS181/1NT19IS181\_vansh.txt

>ls

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -copyToLocal /vansh/1NT19IS181/1NT19IS181_vansh.txt
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ ls
1NT19IS181_vansh.txt  Desktop  Downloads  hadoop-3.2.1.tar.gz  Public  tmpdata
apache-hive-3.1.2-bin  dfsdata  eclipse-workspace  Music  snap  vansh_1NT19IS181.txt
apache-hive-3.1.2-bin.tar.gz  Documents  hadoop-3.2.1  Pictures  Templates  Videos
```

► Use mv command to move files from local file system and hdfs file system

>hadoop fs -mv /vansh/two.txt /vansh/1NT19IS181

>hadoop fs -ls /vansh/1NT19IS181

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -mv /vansh/two.txt /vansh/1NT19IS181
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -ls /vansh/1NT19IS181
Found 2 items
-rw-r--r-- 1 hadoop supergroup 0 2022-06-02 15:12 /vansh/1NT19IS181/1NT19IS181_vansh.txt
-rw-r--r-- 1 hadoop supergroup 0 2022-06-02 15:07 /vansh/1NT19IS181/two.txt
```

► Experiment of cp command

>hadoop fs -cp /vansh/one.txt /vansh/1NT19IS181

>hadoop fs -ls /vansh/1NT19IS181

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -cp /vansh/one.txt /vansh/1NT19IS181
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -ls /vansh/1NT19IS181
Found 3 items
-rw-r--r-- 1 hadoop supergroup 0 2022-06-02 15:12 /vansh/1NT19IS181/1NT19IS181_vansh.txt
-rw-r--r-- 1 hadoop supergroup 0 2022-06-02 15:15 /vansh/1NT19IS181/one.txt
-rw-r--r-- 1 hadoop supergroup 0 2022-06-02 15:07 /vansh/1NT19IS181/two.txt
```

► Display the total size of all the files in name directory

>hadoop fs -du /vansh

>hadoop fs -du -s /vansh

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -du /vansh
0 0 /vansh/1NT19IS181
0 0 /vansh/one.txt
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -du -s /vansh
0 0 /vansh
```

► Display each file size in usn directory Use stat command on name and usn directories

>hadoop fs -du /vansh/1NT19IS181

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -du /vansh/1NT19IS181
0 0 /vansh/1NT19IS181/1NT19IS181_vansh.txt
0 0 /vansh/1NT19IS181/one.txt
0 0 /vansh/1NT19IS181/two.txt
```

➤ Illustrate the use of append command

```
cat > one.txt
```

```
Hello
```

```
^Z
```

```
[4]+ Stopped
```

```
cat > one.txt
```

```
>hadoop fs -appendToFile one.txt /vansh/1NT19IS181/two.txt
```

```
>hadoop fs -cat /vansh/1NT19IS181/two.txt
```

```
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -appendToFile one.txt /vansh/1NT19IS181/two.txt
2022-06-02 15:18:25,052 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
hadoop@admin1-HP-280-G4-MT-Business-PC:~$ hadoop fs -cat /vansh/1NT19IS181/two.txt
2022-06-02 15:18:43,371 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
HELLO
hadoop@admin1-HP-280-G4-MT-Business-PC:~$
```