**1.Create a new HDFS file structure data/dev/analysis in the HDFS home directory. Copy the covid-19 folder in to it.**

[mavricbdhoct01233@ip-10-1-1-204 ~]**hdfs dfs -mkdir -p /user/mavricbdhoct01233/data/dev/analysis**

[mavricbdhoct01233@ip-10-1-1-204 ~]$ ls

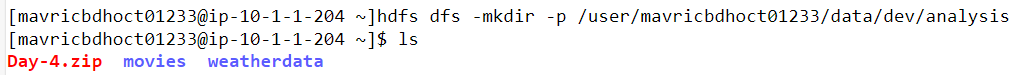
[mavricbdhoct01233@ip-10-1-1-204 ~]$ **hdfs dfs -ls**

Found 2 items

drwx------ - mavricbdhoct01233 mavricbdhoct01233 0 2023-10-26 09:21 .staging

drwxr-xr-x - mavricbdhoct01233 mavricbdhoct01233 0 2023-10-27 06:06 data

[mavricbdhoct01233@ip-10-1-1-204 ~]$



A close-up of a computer code

Description automatically generated

Copy the covid-19 folder in to it.

A screenshot of a computer code

Description automatically generated

[mavricbdhoct01233@ip-10-1-1-204 ~]$ ll

total 333216

drwxrwxr-x 3 mavricbdhoct01233 mavricbdhoct01233 4096 Oct 27 06:33 **codid-19**

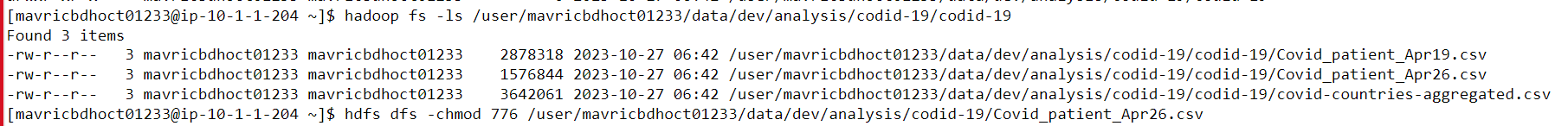
-rw-r--r-- 1 mavricbdhoct01233 mavricbdhoct01233 341195804 Oct 25 11:51 **Day-4.zip**

drwxrwxr-x 2 mavricbdhoct01233 mavricbdhoct01233 4096 Oct 25 09:54 **movies**

drwxrwxr-x 2 mavricbdhoct01233 mavricbdhoct01233 4096 Oct 25 09:52 **weatherdata**

[mavricbdhoct01233@ip-10-1-1-204 ~]$

**2.Change the permission of Covid\_patient\_Apr26.csv file to have only read/write/execute for the owner/ group and read/write permission for others.**



[mavricbdhoct01233@ip-10-1-1-204 ~]$ hadoop fs -ls /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19

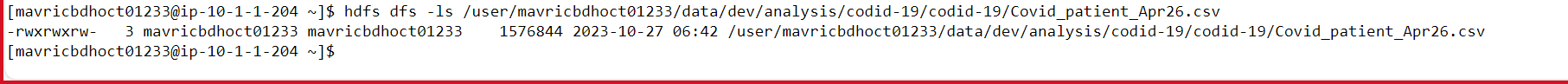
Found 3 items

-rw-r--r-- 3 mavricbdhoct01233 mavricbdhoct01233 2878318 2023-10-27 06:42 /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr19.csv

-rw-r--r-- 3 mavricbdhoct01233 mavricbdhoct01233 1576844 2023-10-27 06:42 /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr26.csv

-rw-r--r-- 3 mavricbdhoct01233 mavricbdhoct01233 3642061 2023-10-27 06:42 /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/covid-countries-aggregated.csv

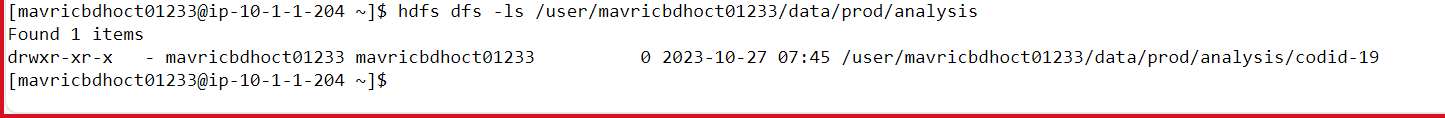
[mavricbdhoct01233@ip-10-1-1-204 ~]$ hdfs dfs -chmod 776 /user/mavricbdhoct01233/data/dev/analysis/codid-19/Covid\_patient\_Apr26.csv



[mavricbdhoct01233@ip-10-1-1-204 ~]$ **hdfs dfs -ls /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr26.csv**

-rwxrwxrw- 3 mavricbdhoct01233 mavricbdhoct01233 1576844 2023-10-27 06:42 /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr26.csv

**3. Copy covid-19 folder in to another hdfs directory data/prod/analysis**



[mavricbdhoct01233@ip-10-1-1-204 ~]$ **hdfs dfs -ls /user/mavricbdhoct01233/data/prod/analysis**

Found 1 items

drwxr-xr-x - mavricbdhoct01233 mavricbdhoct01233 0 2023-10-27 07:45 /user/mavricbdhoct01233/data/prod/analysis/codid-19

[mavricbdhoct01233@ip-10-1-1-204 ~]$

A close-up of a computer screen

Description automatically generated

[mavricbdhoct01233@ip-10-1-1-204 ~]$ **hadoop fs -cp /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19 /user/mavricbdhoct01233/data/prod/analysis**

[mavricbdhoct01233@ip-10-1-1-204 ~]$ hdfs dfs -ls /user/mavricbdhoct01233/data/data/prod/analysis

**4. Read the contents of the file Covid\_patient\_Apr19.csv and print the contents on the screen**

A text on a white background

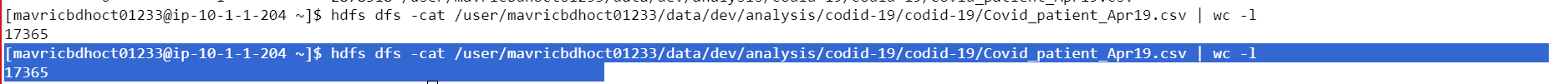
Description automatically generated

**5. Print the no of lines present in the HDFS file Covid\_patient\_Apr19.csv**

[mavricbdhoct01233@ip-10-1-1-204 **~]$ hdfs dfs -cat /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr19.csv | head -n 10**

[mavricbdhoct01233@ip-10-1-1-204 ~]$ **hdfs dfs -cat /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr19.csv | wc -l**

17365



A close-up of a line

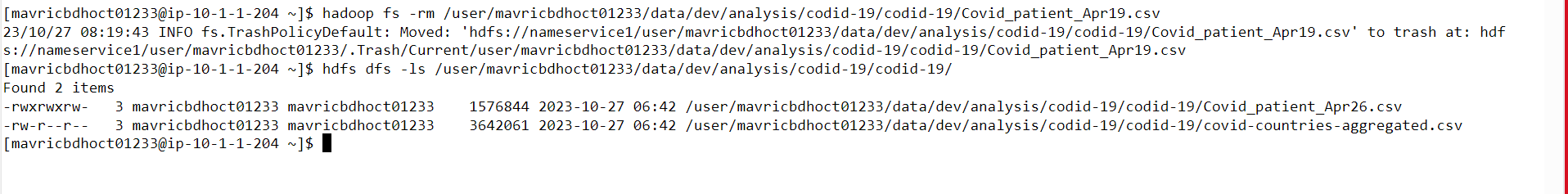
Description automatically generated

**6. Count the no of records present in the HDFS file Covid\_patient\_Apr19.csv for the state “UP”**

[mavricbdhoct01233@ip-10-1-1-204 ~]$ **hadoop fs -cat /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr19.csv | grep "UP" | wc -l**

1113

[mavricbdhoct01233@ip-10-1-1-204 ~]$



**7. Delete the file Covid\_patient\_Apr19.csv from the data/prod/analysis folder in HDFS**

[mavricbdhoct01233@ip-10-1-1-204 ~]$ **hadoop fs -rm /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr19.csv**

23/10/27 08:19:43 INFO fs.TrashPolicyDefault: Moved: 'hdfs://nameservice1/user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr19.csv' to trash at: hdf

s://nameservice1/user/mavricbdhoct01233/.Trash/Current/user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr19.csv

[mavricbdhoct01233@ip-10-1-1-204 ~]$ hdfs dfs -ls /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/

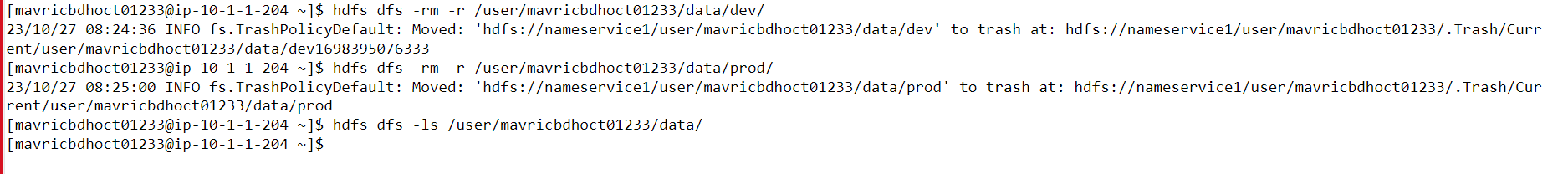
Found 2 items

-rwxrwxrw- 3 mavricbdhoct01233 mavricbdhoct01233 1576844 2023-10-27 06:42 /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/Covid\_patient\_Apr26.csv

-rw-r--r-- 3 mavricbdhoct01233 mavricbdhoct01233 3642061 2023-10-27 06:42 /user/mavricbdhoct01233/data/dev/analysis/codid-19/codid-19/covid-countries-aggregated.csv

[mavricbdhoct01233@ip-10-1-1-204 ~]$

**8. Delete the covid-19 folder from both the dev and prod folders created above.**



[mavricbdhoct01233@ip-10-1-1-204 ~]$ **hdfs dfs -rm -r /user/mavricbdhoct01233/data/dev/**

23/10/27 08:24:36 INFO fs.TrashPolicyDefault: Moved: 'hdfs://nameservice1/user/mavricbdhoct01233/data/dev' to trash at: hdfs://nameservice1/user/mavricbdhoct01233/.Trash/Curr

ent/user/mavricbdhoct01233/data/dev1698395076333

[mavricbdhoct01233@ip-10-1-1-204 ~]$ hdfs dfs -rm -r /user/mavricbdhoct01233/data/prod/

23/10/27 08:25:00 INFO fs.TrashPolicyDefault: Moved: 'hdfs://nameservice1/user/mavricbdhoct01233/data/prod' to trash at: hdfs://nameservice1/user/mavricbdhoct01233/.Trash/Cur

rent/user/mavricbdhoct01233/data/prod

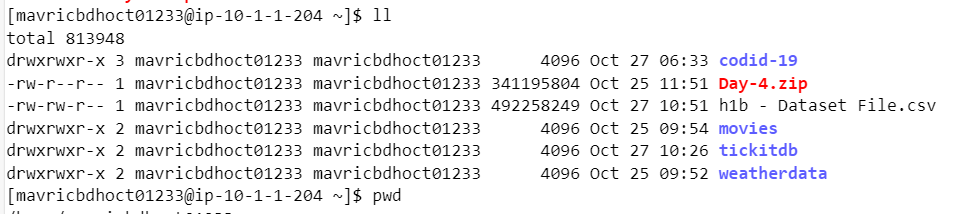
[mavricbdhoct01233@ip-10-1-1-204 ~]$ hdfs dfs -ls /user/mavricbdhoct01233/data/

[mavricbdhoct01233@ip-10-1-1-204 ~]$

**9.Copy the tickitdb folder(in the local file system) in to HDFS home directory. How many blocks will be created for the files in this directory?**

A white paper with black text

Description automatically generated





A screenshot of a computer

Description automatically generated

**10. Find all the file names which has the word “pipe” in their names.**

A black text on a white background

Description automatically generated

hdfs dfs -find /user/mavricbdhoct01233/tickitdb -name "\*pipe\*" -print

**11. Copy the h1b.csv file to hdfs home directory. How many blocks will be created for this file?**

Number of Blocks = File Size / HDFs block size

Number of Blocks = 469.75 MB / 128 MB ≈ 3.66

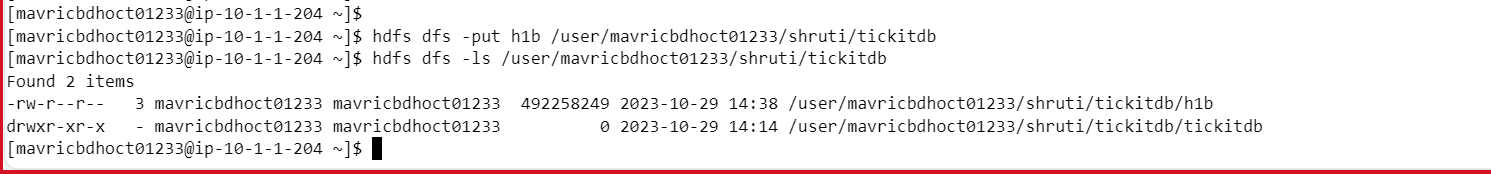
A screenshot of a computer

Description automatically generated

A white screen with black text

Description automatically generated

**12. Move the h1b.csv file (in the HDFS) to the tickitdb folder in hdfs. How many blocks will be there for the tickitdb folder?**



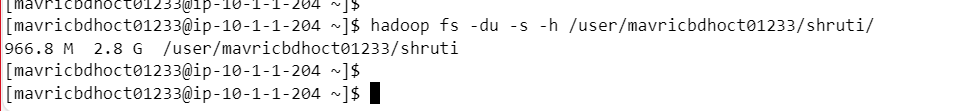
A screen shot of a computer

Description automatically generated

hdfs dfs -ls -mv /user/mavricbdhoct01233/h1b.csv /user/mavricbdhoct01233/tickitdb

**13. Calculate the total storage space occupied by the files directory and its contents.**

hadoop fs -du -s -h /user/mavricbdhoct01233/tickitdb



**14. Check the storage quota for the file directory.**

A white sheet with black text

Description automatically generated

**15. Delete a specific snapshot 'snapshot1' from the file directory.**

hadoop fs -deleteSnapshot /user/mavricbdhoct01233/shruti/ snapshot1

**16. Create a snapshot of the file directory in HDFS.**

hdfs dfs -createSnapshot /user/mavricbdhoct01233/shruti/ snapshot1

**17. Set a replication factor of 3 for all files within the file directory.**

A screenshot of a computer

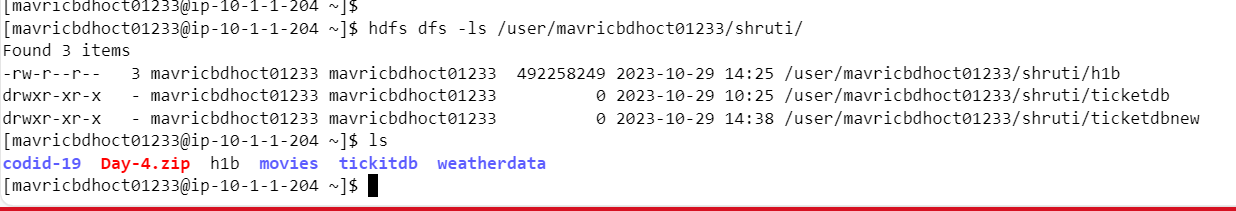
Description automatically generated

**18. Rename the 'file name' file to 'new file name' within the file directory in HDFS.**

A screenshot of a computer

Description automatically generated

**19. List all the files and directories in the files directory in HDFS.**



**20. Create an HDFS directory file names in your home directory and upload a local file 'test.txt' to it.**

A screenshot of a computer

Description automatically generated

A black text on a white background

Description automatically generated