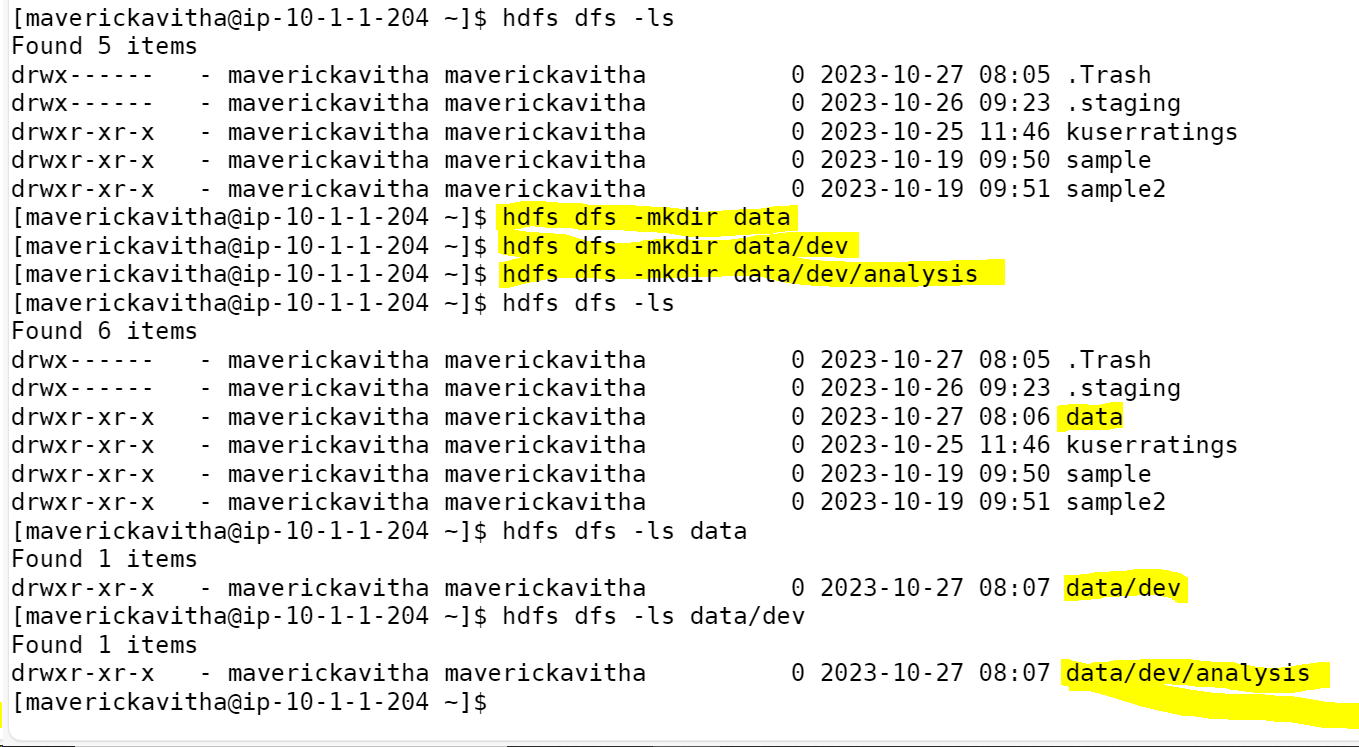
**HDFS Assignment -1**

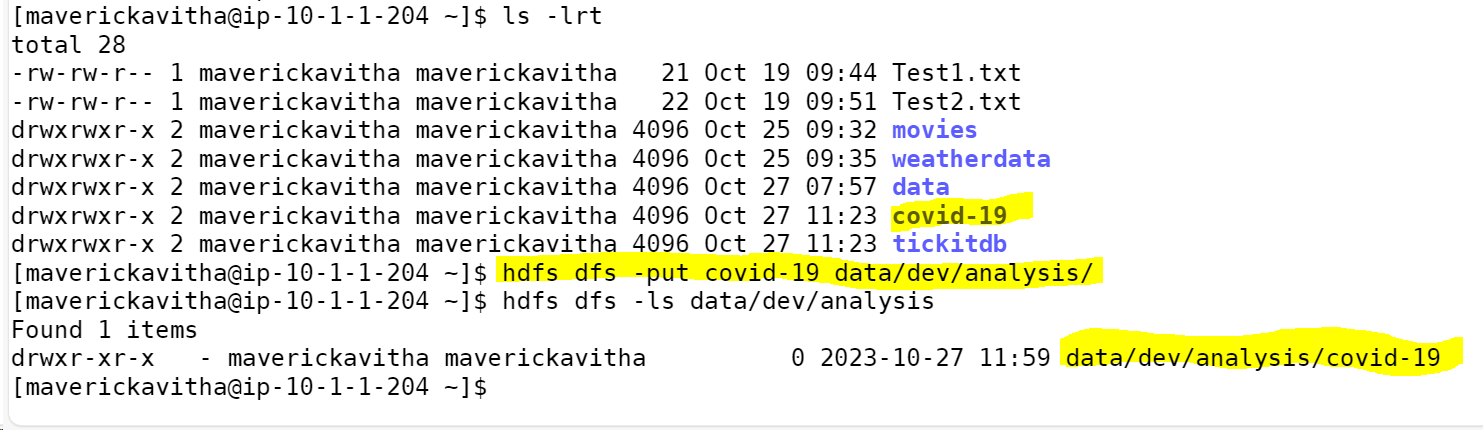
1. **Create a new HDFS file structure data/dev/analysis in the HDFS home directory. Copy the covid-19 folder into it.**

**Solution:**

**Step 1**: Creating the file structure **data/dev/analysis** in HDFS home directory.

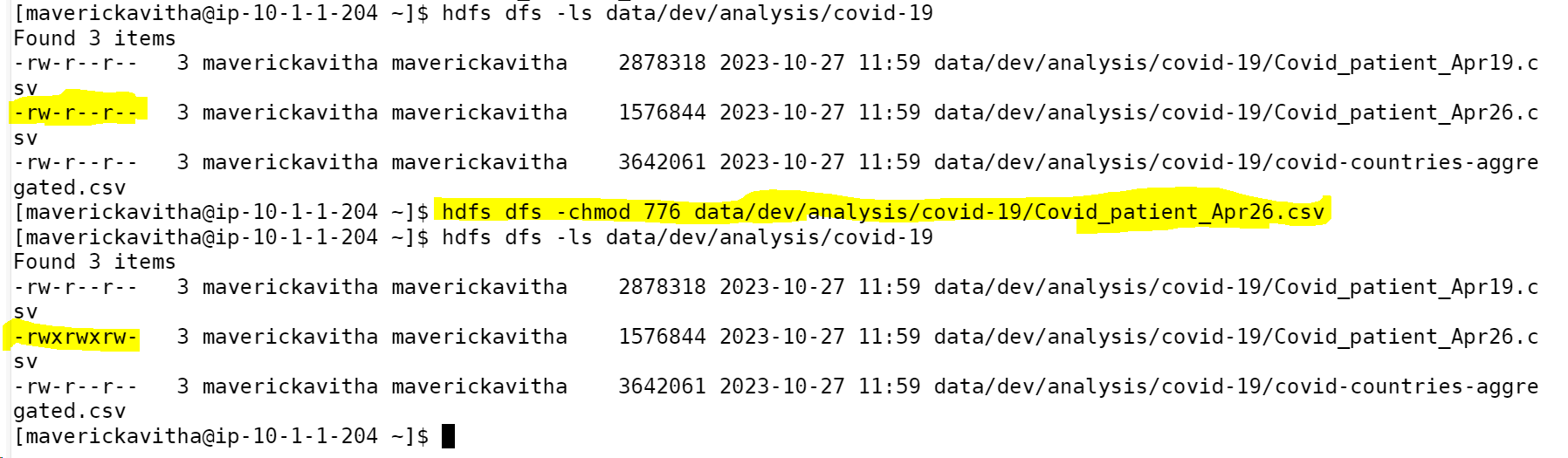


**Step 2:** Copying Covid-19 folder into **data/dev/analysis**



1. **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.**

**Solution:**



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

**Solution:**

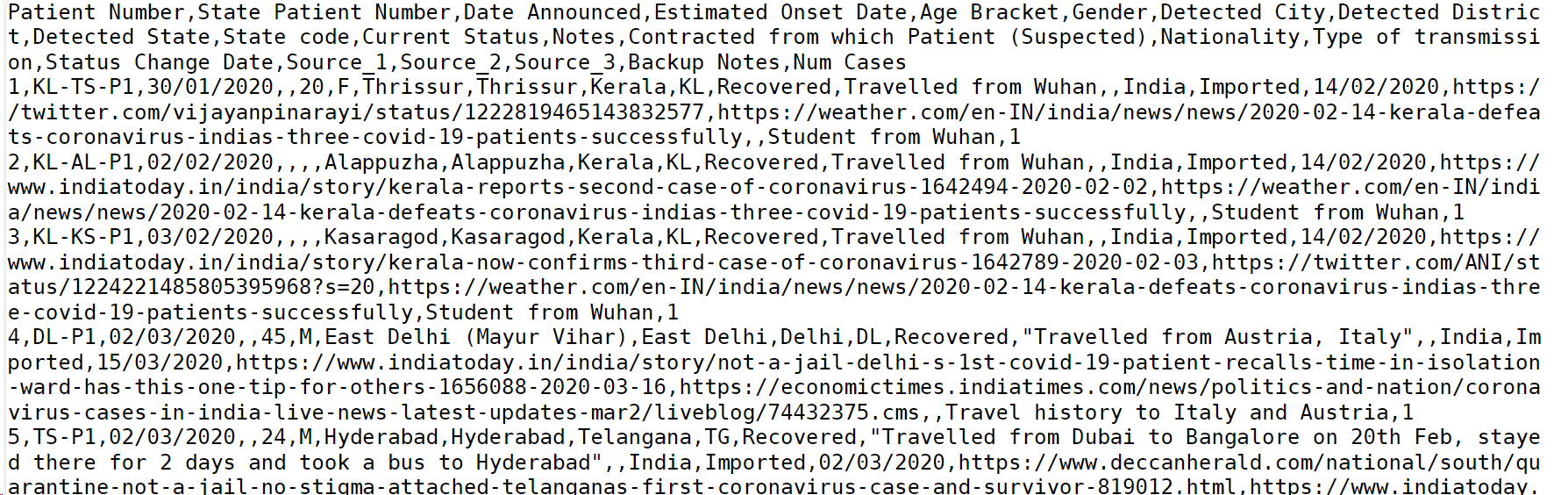
A yellow line with black text

Description automatically generated

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

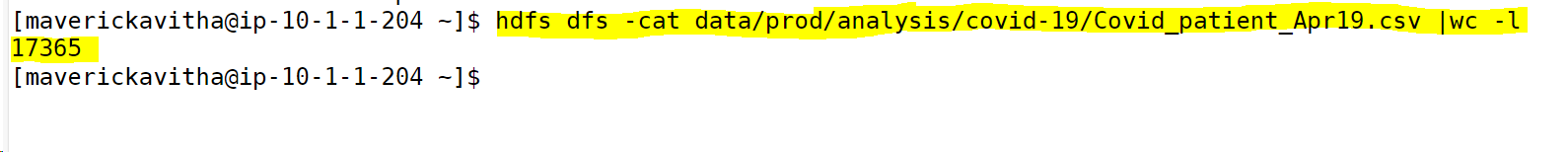
**Solution:**





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

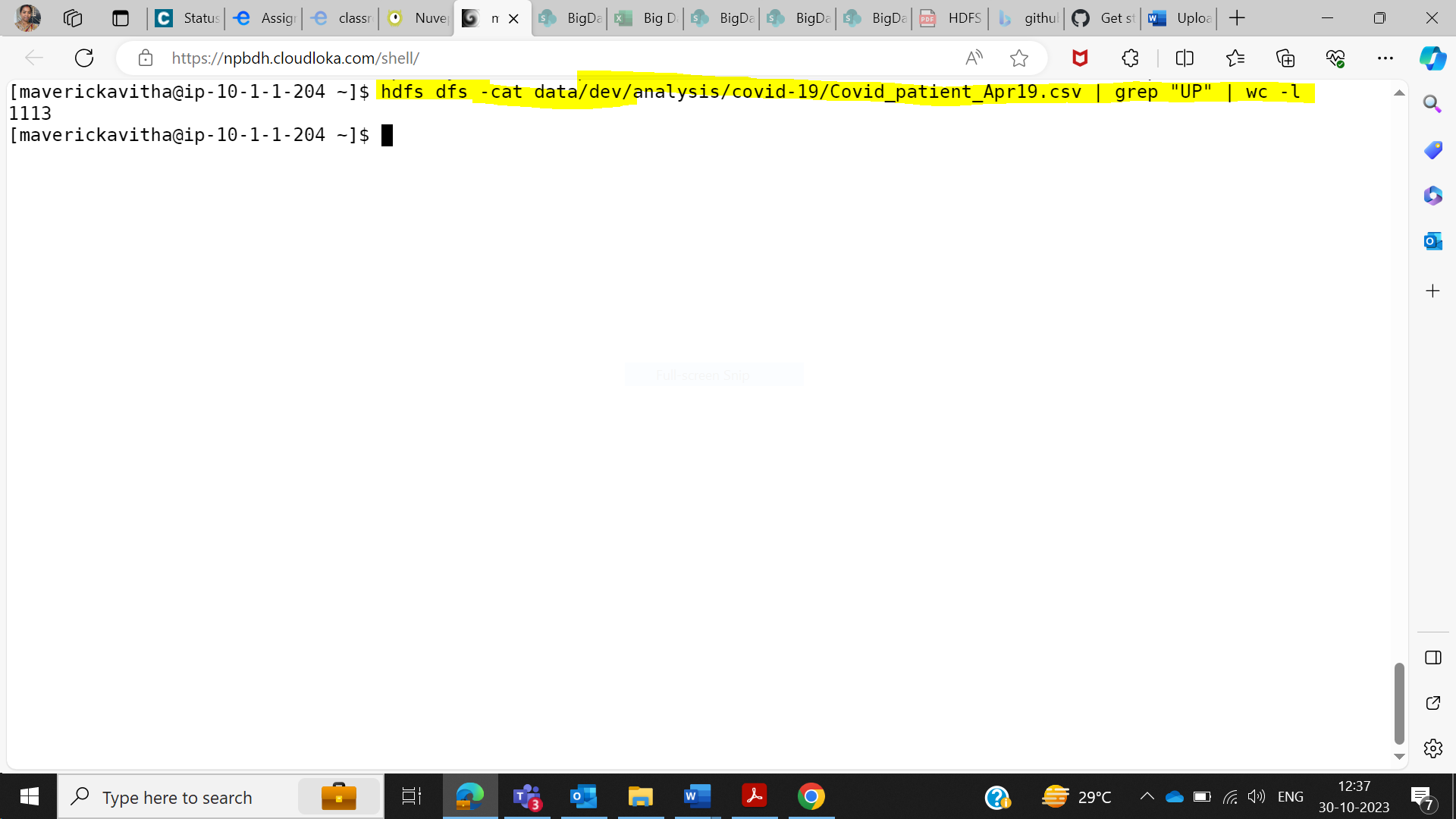
**Solution:**



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

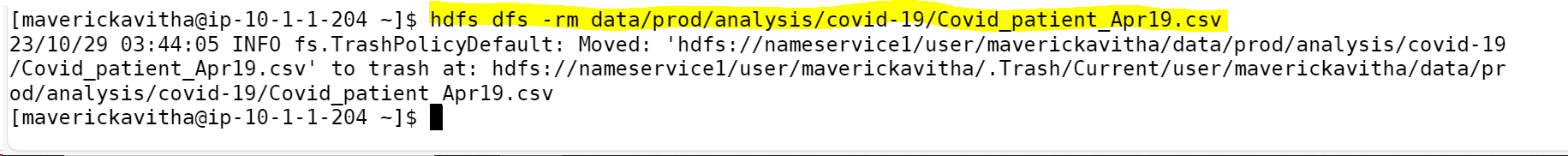
**state “UP”**

**Solution:**



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

**Solution:**



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

**Solution:**

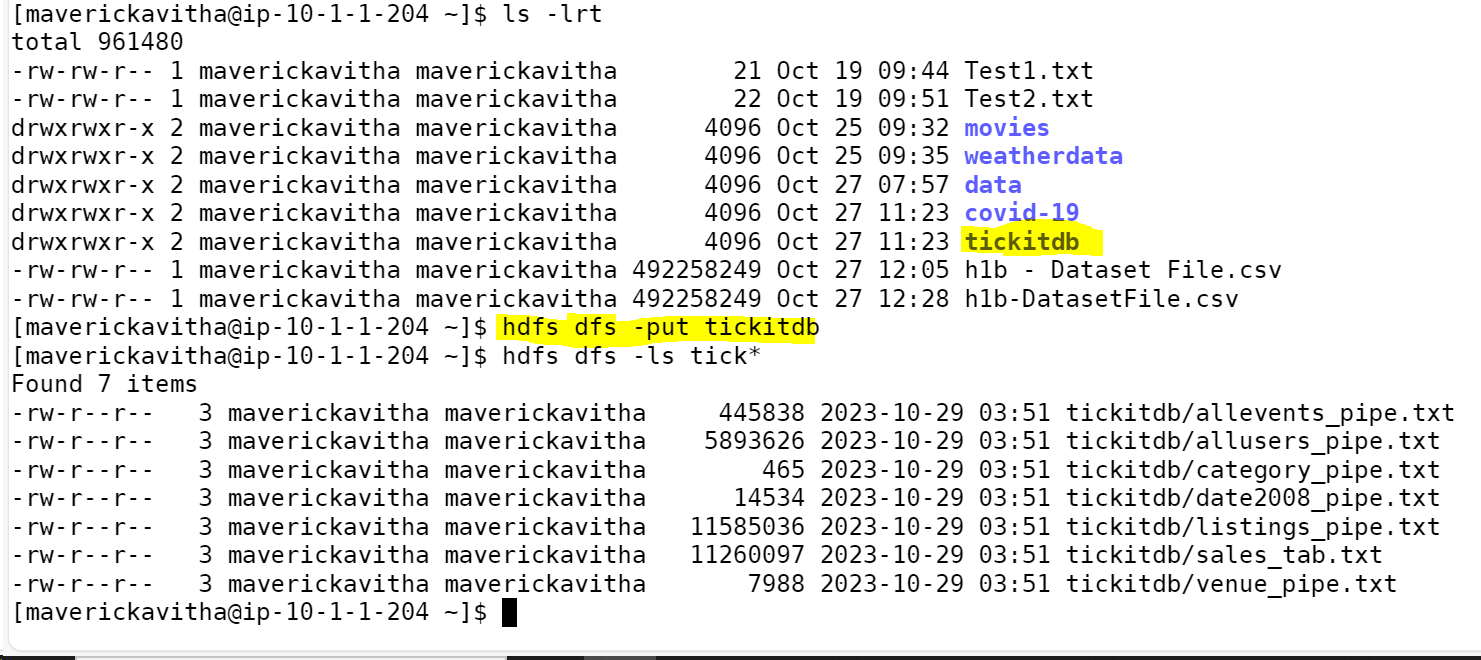
Approach followed is , inorder to delete the folder we need to empty the folder , so I have first deleted the files inside the covid-19 folder and then deleted the folder.

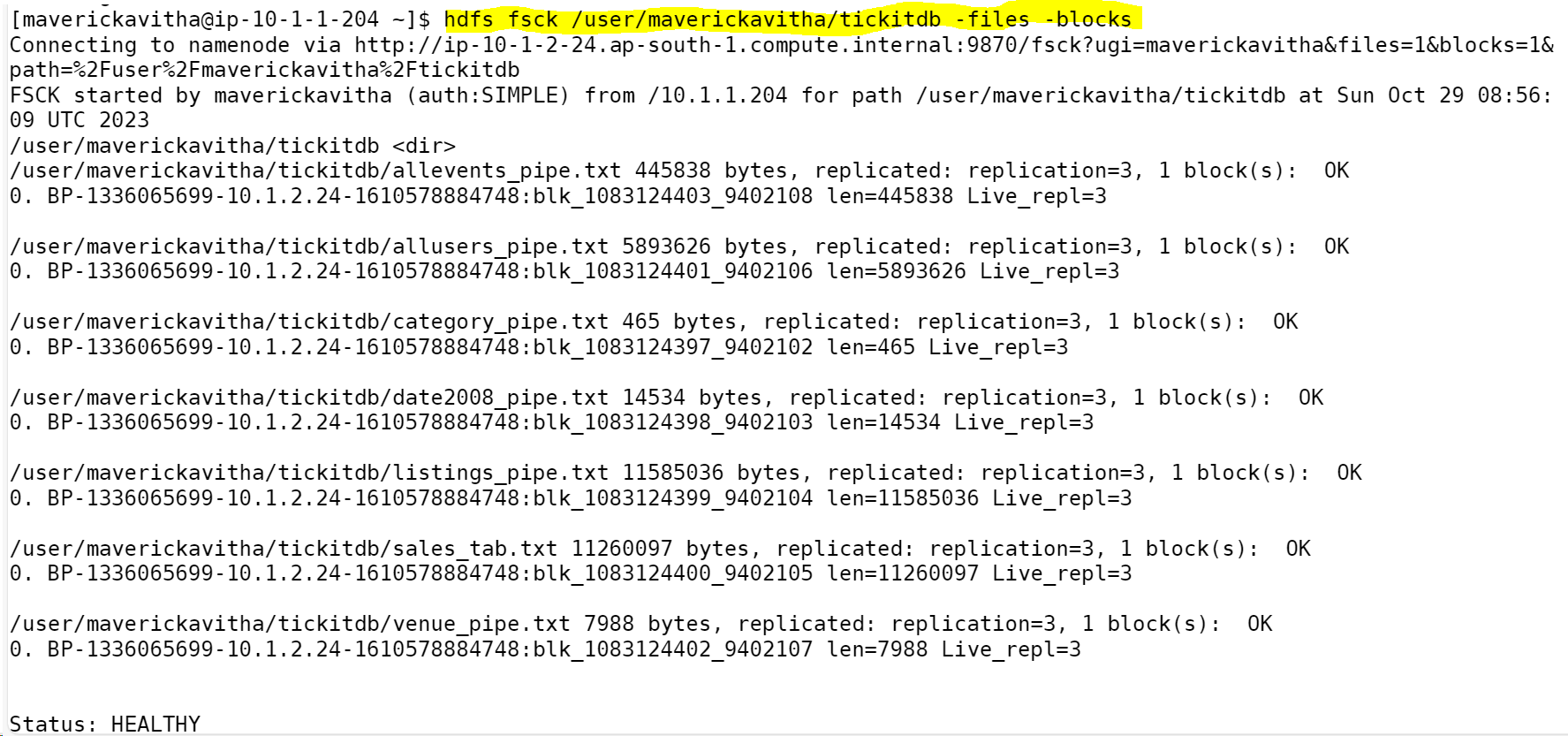
A yellow text with black text

Description automatically generated

**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?**

**Solution:**





A screenshot of a computer program

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.**

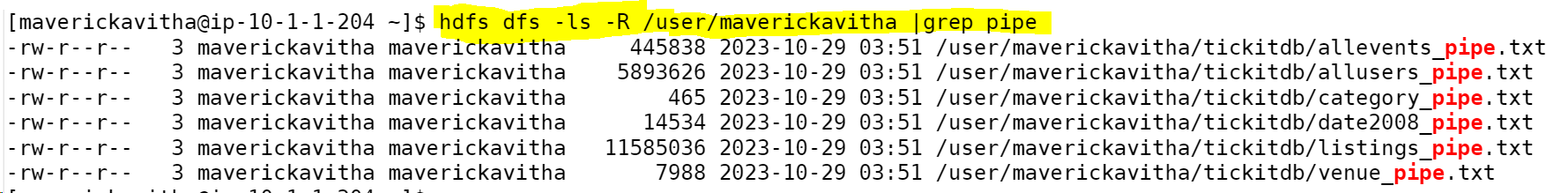
**Solution:**

**Ans 1: File names with pipe in tickitdb**

A black and yellow text on a white background

Description automatically generated

**Ans 2: File names with pipe in hdfs home directory**



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

**Solution:**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

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?**

**Solution:**

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

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

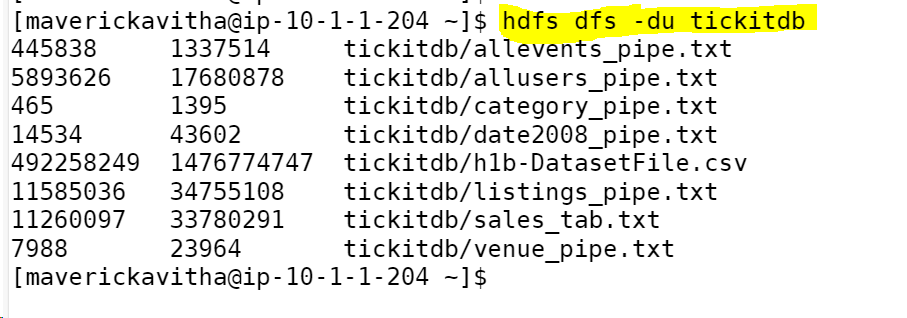
**Solution:**

**Storage occupied by total hdfs fike directory and its content**

A close up of a number

Description automatically generated

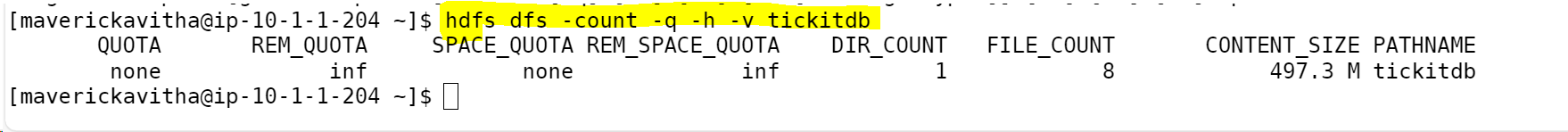
**Storage occupied by tickitdb**



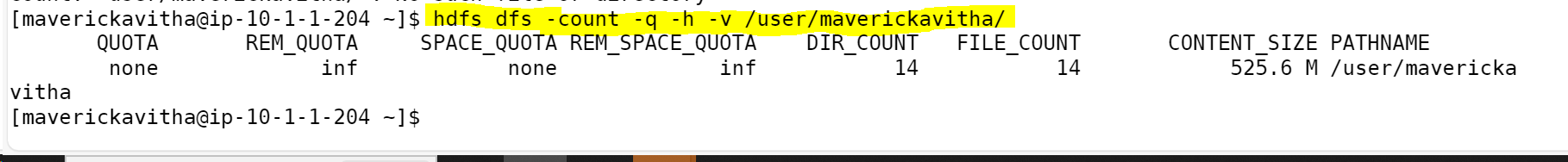
**14. Check the storage quota for the file directory.**

**Solution:**

**Storage quota for tickitdb folder is calculated**



**Storage quota calculated for the total file directory of user**

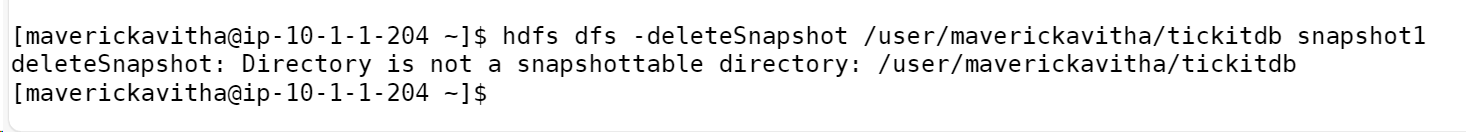


**15. Delete a specific snapshot 'snapshot1' from the file directory.**

**Solution:**

**As there is no access to create snapshot , no access to delete as well.**

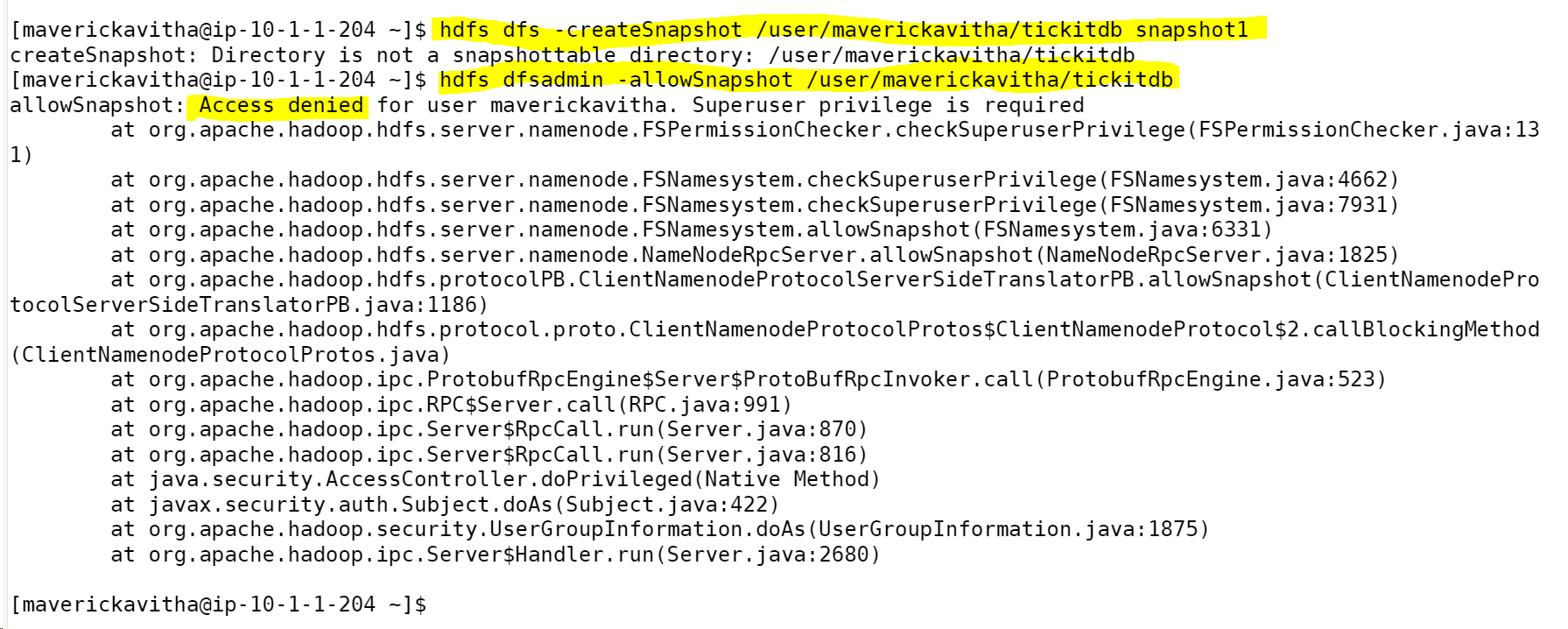
**HDFS command : >> hdfs dfs -deleteSnapshot /user/maverickavitha/tickitdb snapshot1**



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

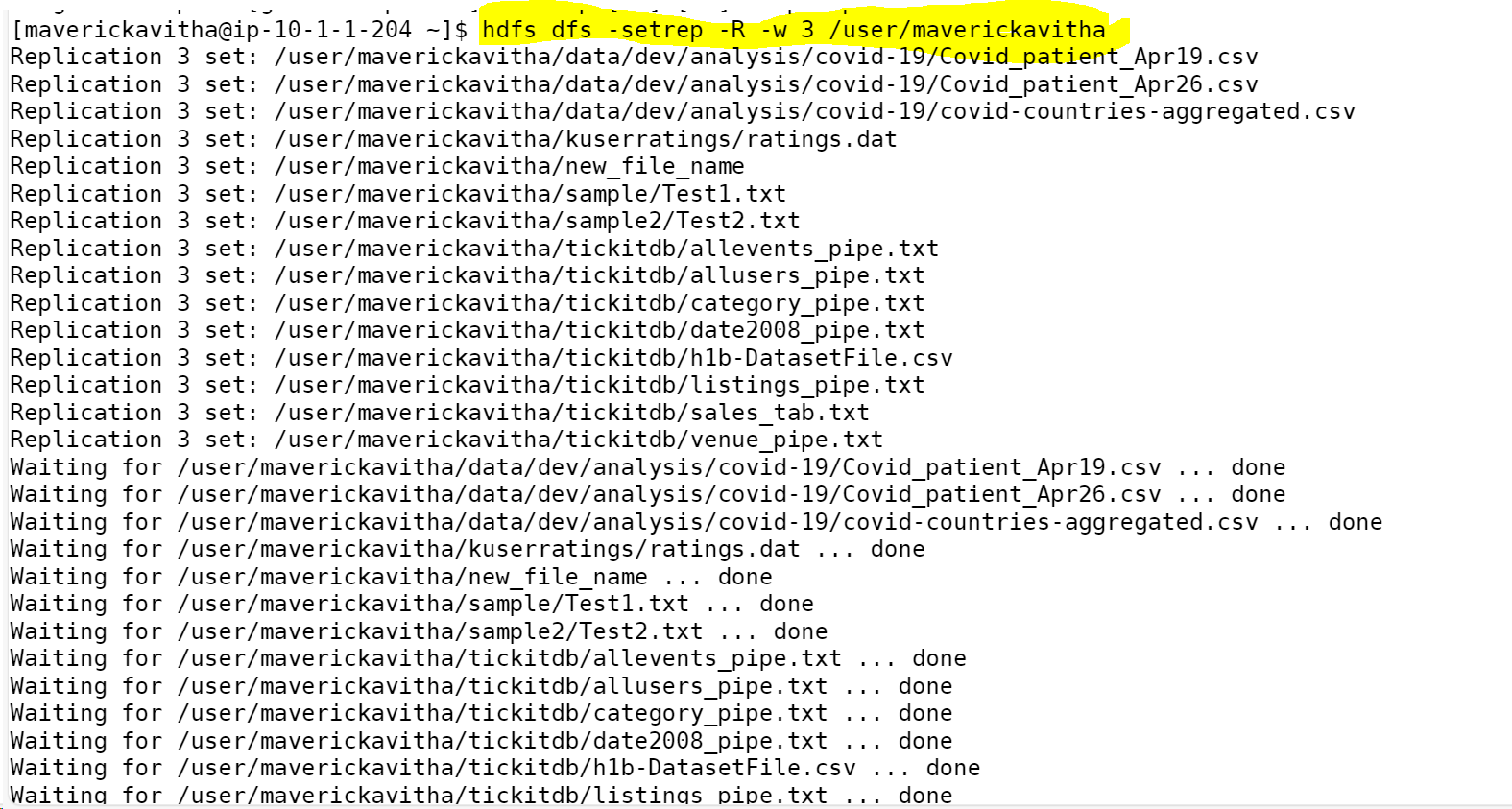
**Solution:**

**No access to create snapshot**



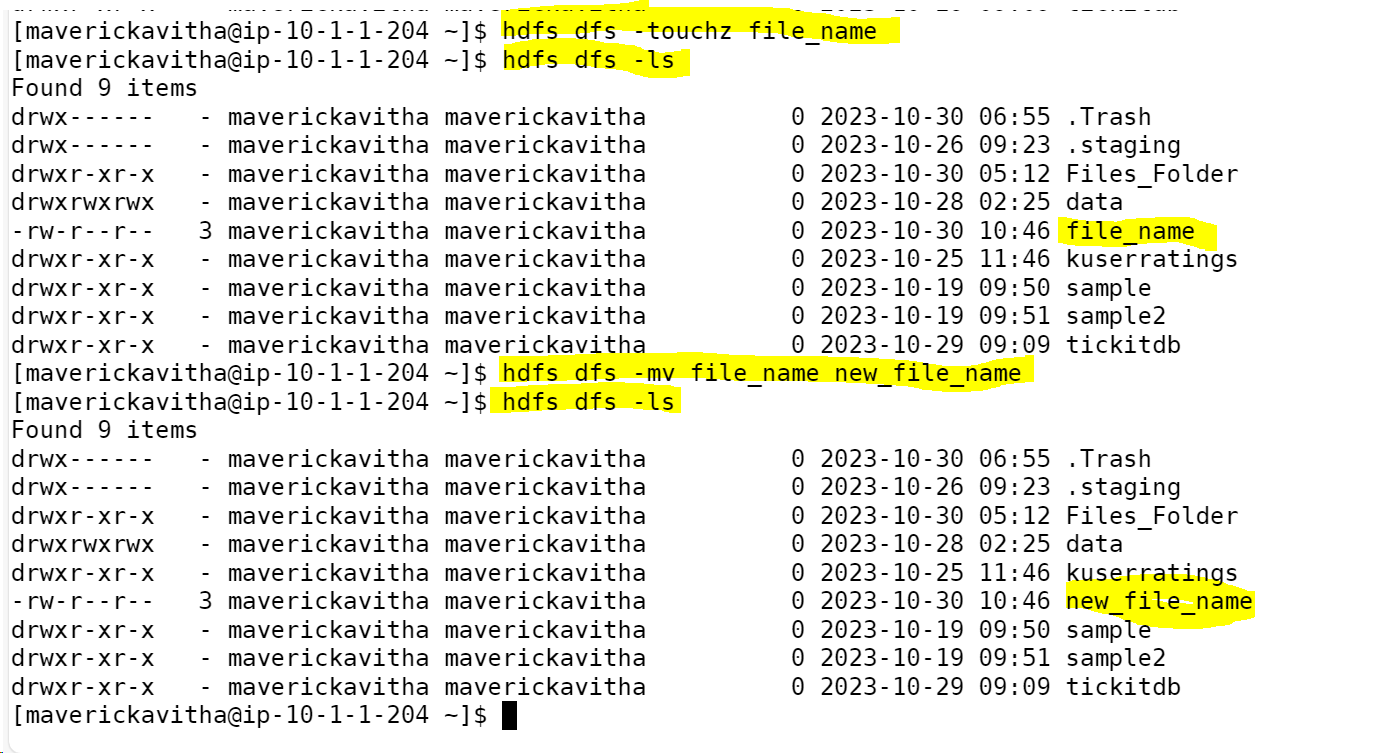
**17. Set a replication factor of 3 for all files within the file directory.**

**Solution:**



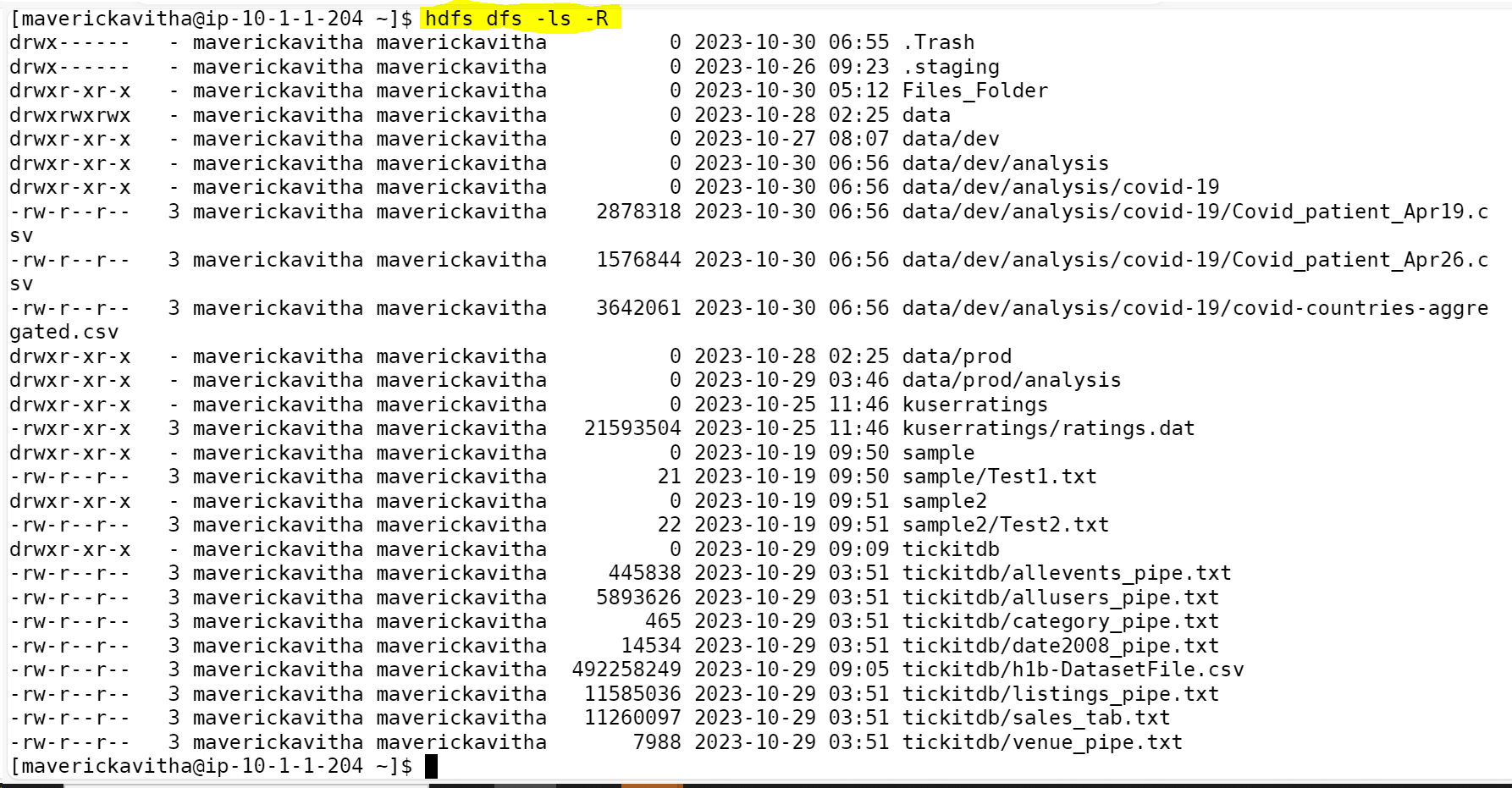
**18. Rename the 'file name' file to 'new file name' within the file directory in HDFS.**

**Solution:**



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

**Solution:**



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

**Solution:**



A screenshot of a computer

Description automatically generated

