Term Work 3
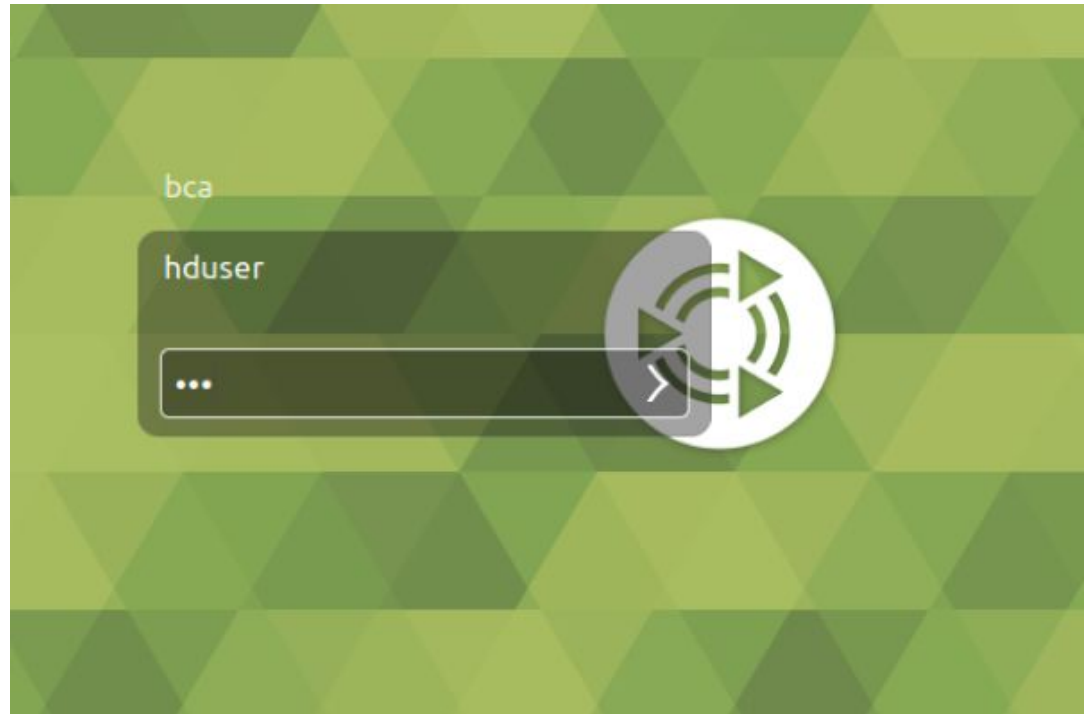
Password
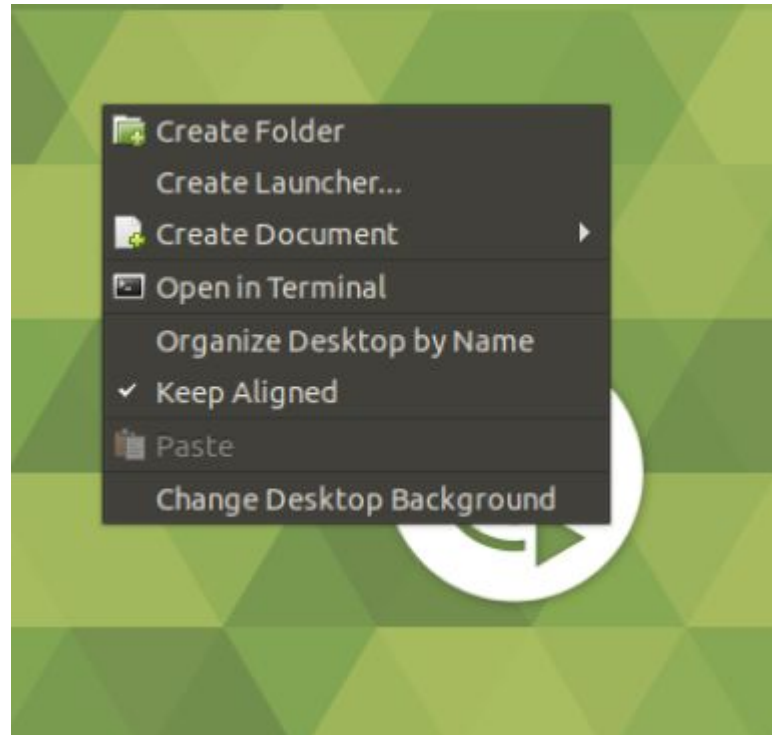dv
or
bca

Right Click & select
Open in Terminal

Type jps &


```
hduser@ubuntu:~

File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~$ jps
2472 Jps
hduser@ubuntu:~$ ▮
```

Type & press

Blue Colour - Directory

Password bca or dv

```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~$ cd Documents
hduser@ubuntu:~/Documents$ sudo mkdir termwork3
[sudo] password for hduser:
hduser@ubuntu:~/Documents$ ls
bgm20          map90.txt      map.py       -reducer      shop90.txt    word13.txt
bgm27          -mapper        -output      reducer45.py  termwork1     word23.txt
flipkart.txt   mapper45.py    red90.py     reducer.py    termwork2     words.txt
map90.py       mapper.py      red.py       shop1.txt     termwork3     word.txt
hduser@ubuntu:~/Documents$ █
```

Type cd termwork3/

& press


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents$ cd termwork3/
hduser@ubuntu:~/Documents/termwork3$ █
```

Type sudo nano age.csv

& press

File Edit View Search Terminal Help
hduser@ubuntu:~/Documents/termwork3$ sudo nano age.csv

Type the text

& press Ctrl x y & then press


```
File  Edit  View  Search  Terminal  Help
  GNU  nano  2.9.3                                age.csv

1,1,21,5.4,50
2,2,22,5.0,52
3,1,23,4.5,55
4,1,22,6.0,65
5,2,21,4.9,50
6,2,20,4.5,48
7,1,19,5.2,55
8,2,20,4.8,50
9,1,23,6.1,75
10,2,24,5.5,55
```

Type cat age.csv
& press

This shows your typed text


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ cat age.csv
1,1,21,5.4,50
2,2,22,5.0,52
3,1,23,4.5,55
4,1,22,6.0,65
5,2,21,4.9,50
6,2,20,4.5,48
7,1,19,5.2,55
8,2,20,4.8,50
9,1,23,6.1,75
10,2,24,5.5,55

hduser@ubuntu:~/Documents/termwork3$
```

Type ls
& press

This shows your file


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ ls
age.csv
hduser@ubuntu:~/Documents/termwork3$ █
```

Type sudo nano mapper.py

& press

File Edit View Search Terminal Help
hduser@ubuntu:~/Documents/termwork3$ sudo nano mapper.py

Type the code
& press Ctrl x y & then press


```
File  Edit  View  Search  Terminal  Help
  GNU  nano  2.9.3                                    mapper.py

#!/usr/bin/python
import sys
# input comes from STDIN (standard input)
for line in sys.stdin:
    line = line.strip()
    line = line.split(",")
    if len(line) >=2:
        gender = line[1]
        age = line[2]
        print '%s\t%s' % (gender, age)
```

Type cat mapper.py

& press


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ cat mapper.py
#!/usr/bin/python
import sys
# input comes from STDIN (standard input)
for line in sys.stdin:
    line = line.strip()
    line = line.split(",")
    if len(line) >=2:
        gender = line[1]
        age = line[2]
        print '%s\t%s' % (gender, age)

hduser@ubuntu:~/Documents/termwork3$
```

Type ls
& press

This shows your files

Type cat age.csv | python mapper.py

& press


```
 File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ cat age.csv | python mapper.py
1       21
2       22
1       23
1       22
2       21
2       20
1       19
2       20
1       23
2       24
hduser@ubuntu:~/Documents/termwork3$ ▮
```

Type sudo nano reducer.py

& press

File Edit View Search Terminal Help
hduser@ubuntu:~/Documents/termwork3$ sudo nano reducer.py

Type the code

& press Ctrl x y & then press


```
File  Edit  View  Search  Terminal  Help
  GNU nano 2.9.3                                    reducer.py

#!/usr/bin/env python
#Reducer.py
import sys
gender_age = {}
#Partitoner
for line in sys.stdin:
    line = line.strip()
    gender, age = line.split('\t')
    if gender in gender_age:
        gender_age[gender].append(int(age))
    else:
        gender_age[gender] = []
        gender_age[gender].append(int(age))
#Reducer
for gender in gender_age.keys():
    ave_age = sum(gender_age[gender])*1.0 / len(gender_age[gender])
    print '%s\t%s'% (gender, ave_age)
```

Type cat reducer.py

& press


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ cat reducer.py
#!/usr/bin/env python
#Reducer.py
import sys
gender_age = {}
#Partitoner
for line in sys.stdin:
    line = line.strip()
    gender, age = line.split('\t')
    if gender in gender_age:
        gender_age[gender].append(int(age))
    else:
        gender_age[gender] = []
        gender_age[gender].append(int(age))
#Reducer
for gender in gender_age.keys():
    ave_age = sum(gender_age[gender])*1.0 / len(gender_age[gender])
    print '%s\t%s'% (gender, ave_age)
hduser@ubuntu:~/Documents/termwork3$
```

Type ls
& press

This shows your files

Type cat age.csv | python mapper.py | python reducer.py
& press


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ cat age.csv | python mapper.py | python reducer.py
1       21.6
2       21.4
hduser@ubuntu:~/Documents/termwork3$
```

Type jps
& press


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ jps
2789 Jps
hduser@ubuntu:~/Documents/termwork3$
```

Type start-all.sh
& press

Provide password 4 times (it will ask)

```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
```

Type jps
& press


```
 File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ jps
3680 ResourceManager
3488 SecondaryNameNode
3202 DataNode
3027 NameNode
3853 NodeManager
4159 Jps
hduser@ubuntu:~/Documents/termwork3$ █
```

Type hdfs dfs -mkdir /tw3
& press

If it goes into safe mode then type
hdfs dfsadmin –safemode leave
After this type
hdfs dfs –mkdir /tw3

This creates a directory in cluster


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ hdfs dfs -mkdir /tw3
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.Ke
rberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.7.1.jar) to me
thod sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.securit
y.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective acc
ess operations
WARNING: All illegal access operations will be denied in a future release
24/07/12 00:49:17 WARN util.NativeCodeLoader: Unable to load native-hadoop library for
your platform... using builtin-java classes where applicable
hduser@ubuntu:~/Documents/termwork3$ ▮
```

Type hdfs dfs -ls /

& press


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ hdfs dfs -ls /
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authenticatio
rberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.7.1.j
thod sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoo
y.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal refle
ess operations
WARNING: All illegal access operations will be denied in a future release
24/07/12 00:50:52 WARN util.NativeCodeLoader: Unable to load native-hadoop lib
your platform... using builtin-java classes where applicable
Found 20 items
drwxr-xr-x   - hduser supergroup          0 2024-07-09 04:50 /90
drwxr-xr-x   - hduser supergroup          0 2024-06-25 03:02 /TERM2
drwxr-xr-x   - hduser supergroup          0 2024-07-02 22:35 /TERM3
drwxr-xr-x   - hduser supergroup          0 2024-06-18 03:16 /TW1
drwxr-xr-x   - hduser supergroup          0 2024-06-24 03:16 /TW2
drwxr-xr-x   - hduser supergroup          0 2024-07-01 03:10 /TW3
drwxr-xr-x   - hduser supergroup          0 2024-06-06 03:38 /bcom
```

Type hdfs dfs –copyFromLocal
/home/hduser/Documents/termwork3/age.csv /tw3

& press


```
File Edit View Search Terminal Help
hduser@ubuntu:~/Documents/termwork3$ hdfs dfs -copyFromLocal /home/hduser/Documents/ter
mwork3/age.csv /tw3
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.Ke
rberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.7.1.jar) to me
thod sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.securit
y.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective acc
ess operations
WARNING: All illegal access operations will be denied in a future release
24/07/12 00:55:07 WARN util.NativeCodeLoader: Unable to load native-hadoop library for
your platform... using builtin-java classes where applicable
hduser@ubuntu:~/Documents/termwork3$
```

Type hdfs dfs -ls /tw3

& press


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ hdfs dfs -ls /tw3
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.Ke
rberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.7.1.jar) to me
thod sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.securit
y.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective acc
ess operations
WARNING: All illegal access operations will be denied in a future release
24/07/12 00:56:16 WARN util.NativeCodeLoader: Unable to load native-hadoop library for
your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r--   1 hduser supergroup        142 2024-07-12 00:55 /tw3/age.csv
hduser@ubuntu:~/Documents/termwork3$ ▯
```

Type sudo chmod 777 mapper.py reducer.py
& press

Provide password dv or bca

Type ls -l
& press


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ ls -l
total 12
-rw-r--r-- 1 root root 142 Jul 12 00:28 age.csv
-rwxrwxrwx 1 root root 254 Jul 12 00:32 mapper.py
-rwxrwxrwx 1 root root 462 Jul 12 00:41 reducer.py
hduser@ubuntu:~/Documents/termwork3$
```

Type


```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ hadoop jar /home/hduser/Downloads/hadoop-streaming
-2.7.3.jar \
> -input /tw3/age.csv \
> -output /tw3/output \
> -mapper /home/hduser/Documents/termwork3/mapper.py \
> -reducer /home/hduser/Documents/termwork3/reducer.py
```

& press

Everything working fine, it will show the output directory & press


```
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=213
        File Output Format Counters
                Bytes Written=14
24/07/12 01:05:52 INFO streaming.StreamJob: Output directory: /tw3/output
hduser@ubuntu:~/Documents/termwork3$
```

Type hdfs dfs –cat /tw3/output/part-00000
& press

This shows the final output

```
File  Edit  View  Search  Terminal  Help
hduser@ubuntu:~/Documents/termwork3$ hdfs dfs  -cat /tw3/output/part-00000
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util
rberosUtil (file:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-2.7.1.jar) to
thod sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.secu
y.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective
ess operations
WARNING: All illegal access operations will be denied in a future release
24/07/12 01:07:17 WARN util.NativeCodeLoader: Unable to load native-hadoop library f
your platform... using builtin-java classes where applicable
1       21.6
2       21.4
hduser@ubuntu:~/Documents/termwork3$
```

Now open firefox and type localhost:50070/

& press
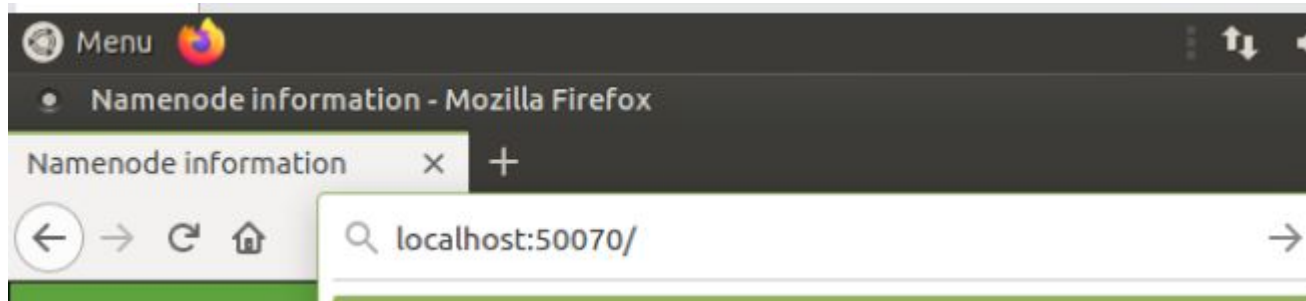
Click utilities & click Browse the file system
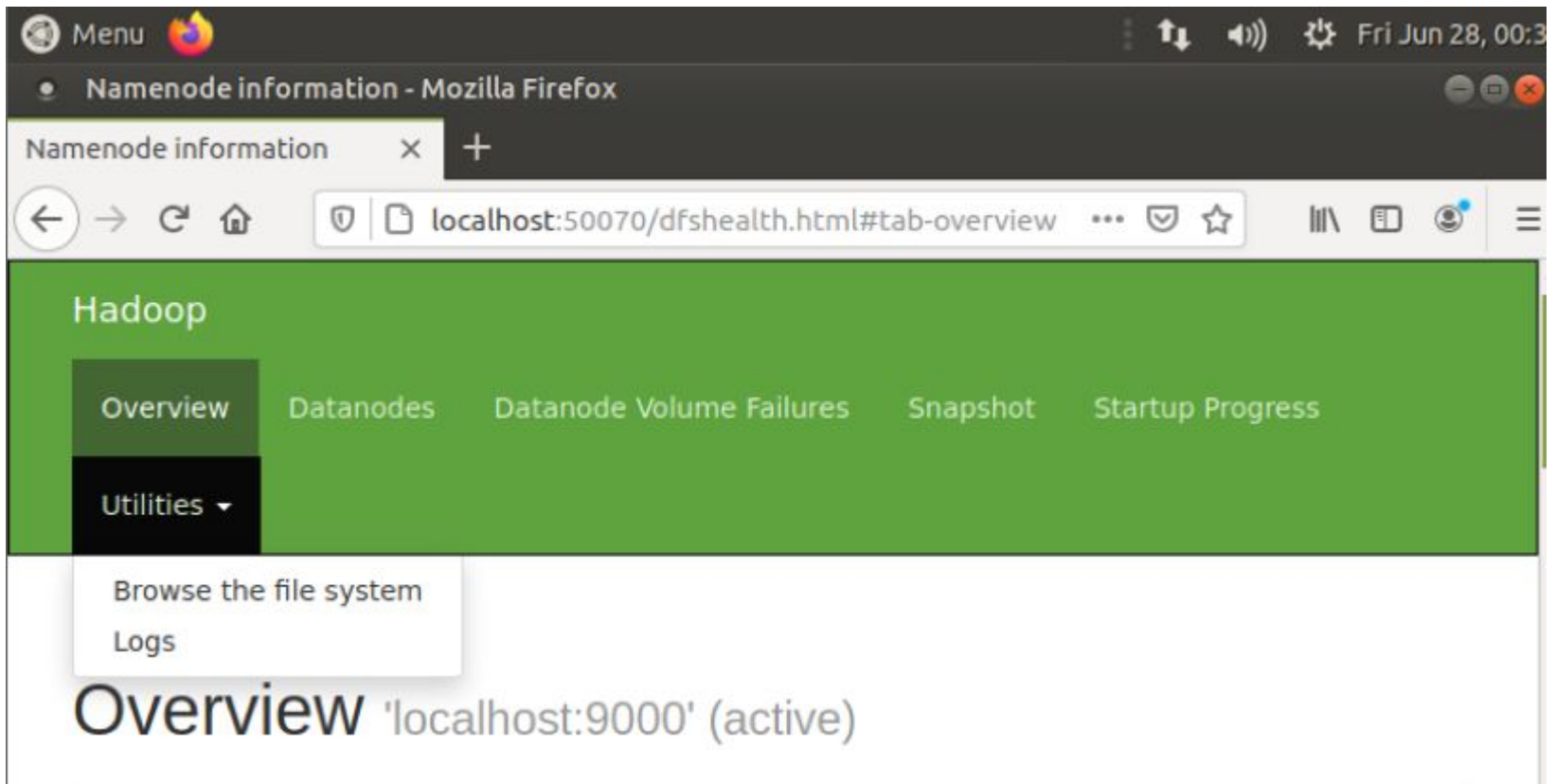
Click tw3

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	hduser	supergroup	0 B	7/9/2024, 4:50:45 AM	0	0 B	90
drwxr-xr-x	hduser	supergroup	0 B	6/25/2024, 3:02:27 AM	0	0 B	TERM2
drwxr-xr-x	hduser	supergroup	0 B	7/2/2024, 10:35:30 PM	0	0 B	TERM3
drwxr-xr-x	hduser	supergroup	0 B	6/18/2024, 3:16:15 AM	0	0 B	TW1
drwxr-xr-x	hduser	supergroup	0 B	6/24/2024, 3:16:43 AM	0	0 B	TW2
drwxr-xr-x	hduser	supergroup	0 B	7/1/2024, 3:10:25 AM	0	0 B	TW3

Click output

Browse Directory

/TW3							Go!

Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name
drwxr-xr-x	hduser	supergroup	0 B	7/1/2024, 3:11:35 AM	0	0 B	output
-rw-r--r--	hduser	supergroup	141 B	7/1/2024, 3:03:14 AM	1	128 MB	textdata3.csv

Click part-00000

Click Download
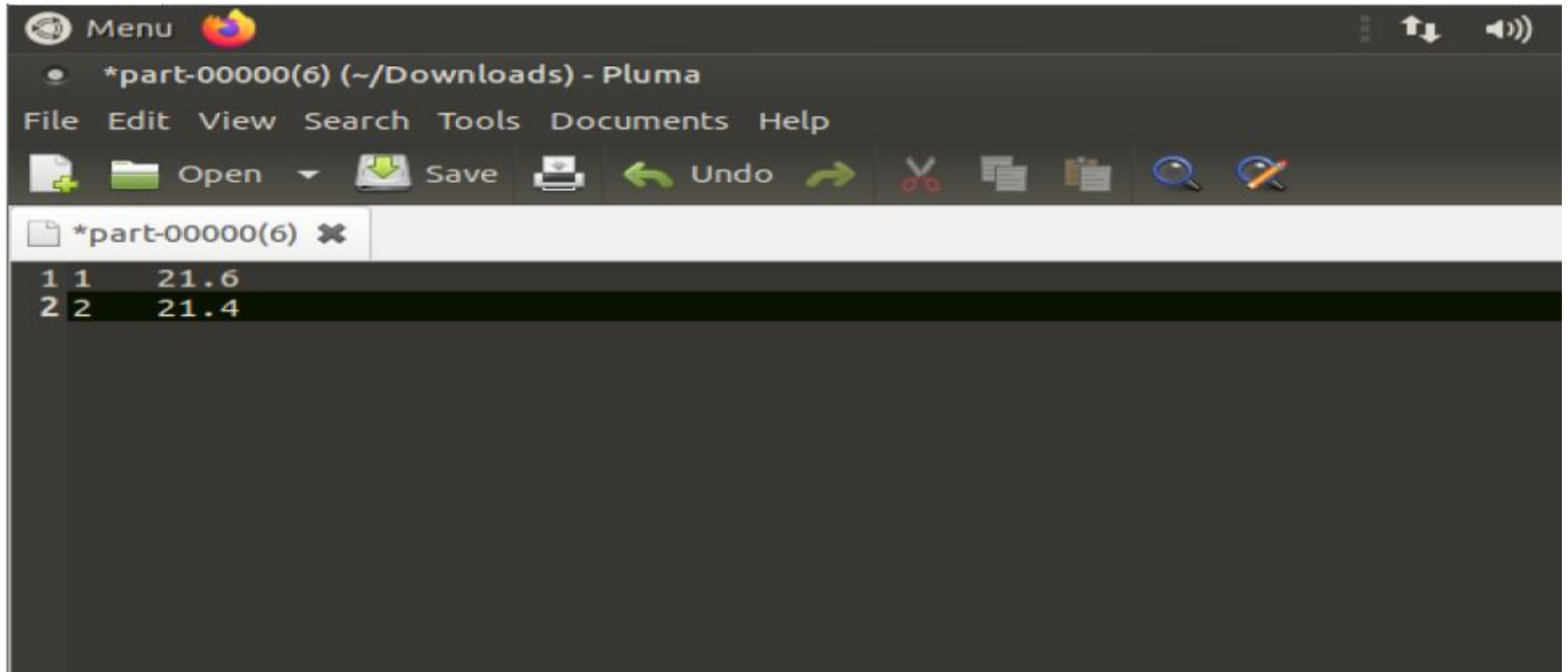
Click savefile

Go to download and open the file part-00000 to view the output