FILTERS USING REGULAR EXPRESSION

grep:searching for a pattern

- grep scans its input for a pattern,and can display the selected pattern,the line numbers or filenames where the pattern occurs
- **Syntax:**

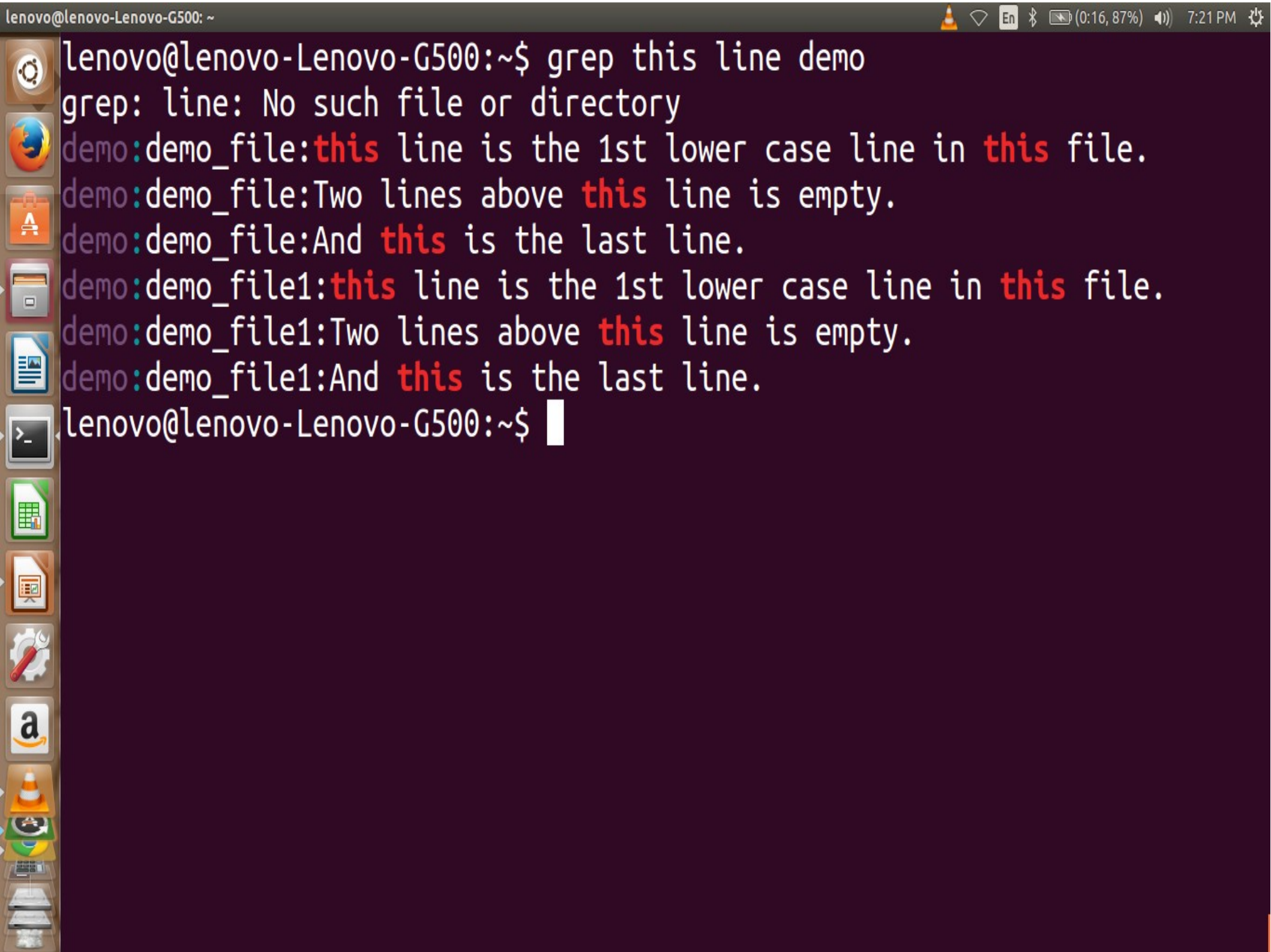
  grep *options pattern filename(s)*

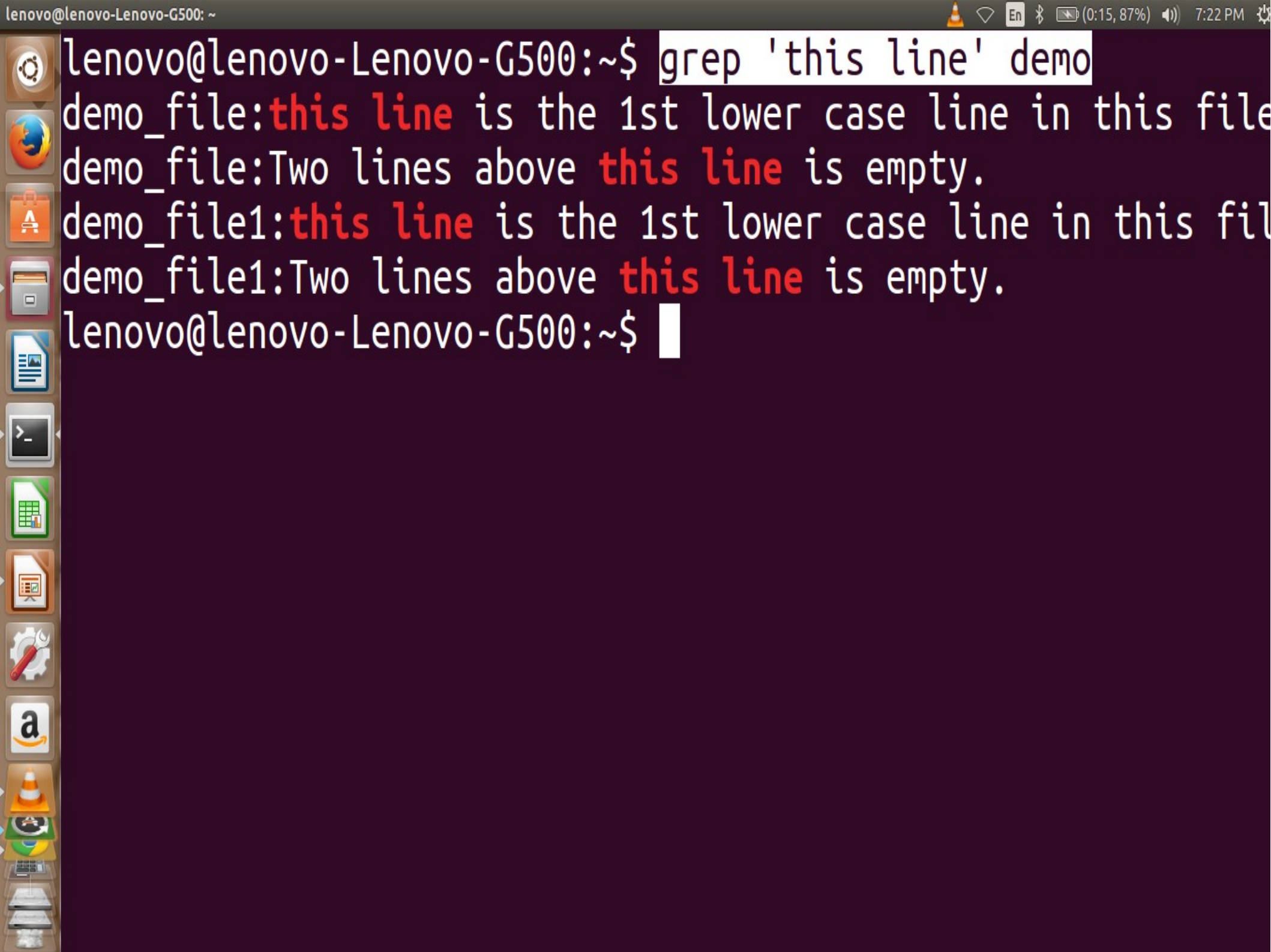
  grep “sales” emp.lst

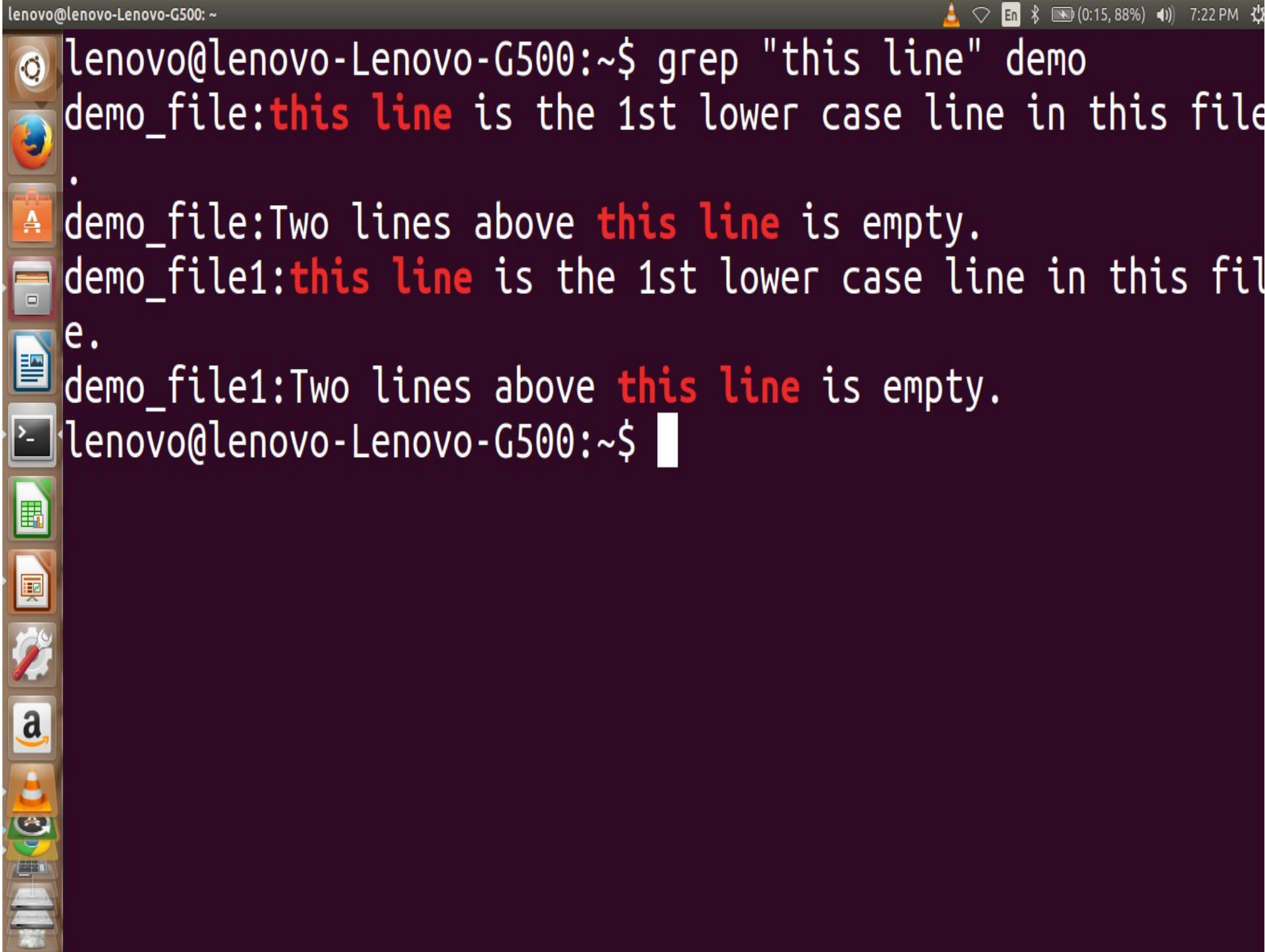
- grep is also a filter,it can search its standard input for the pattern and store the output in a file:

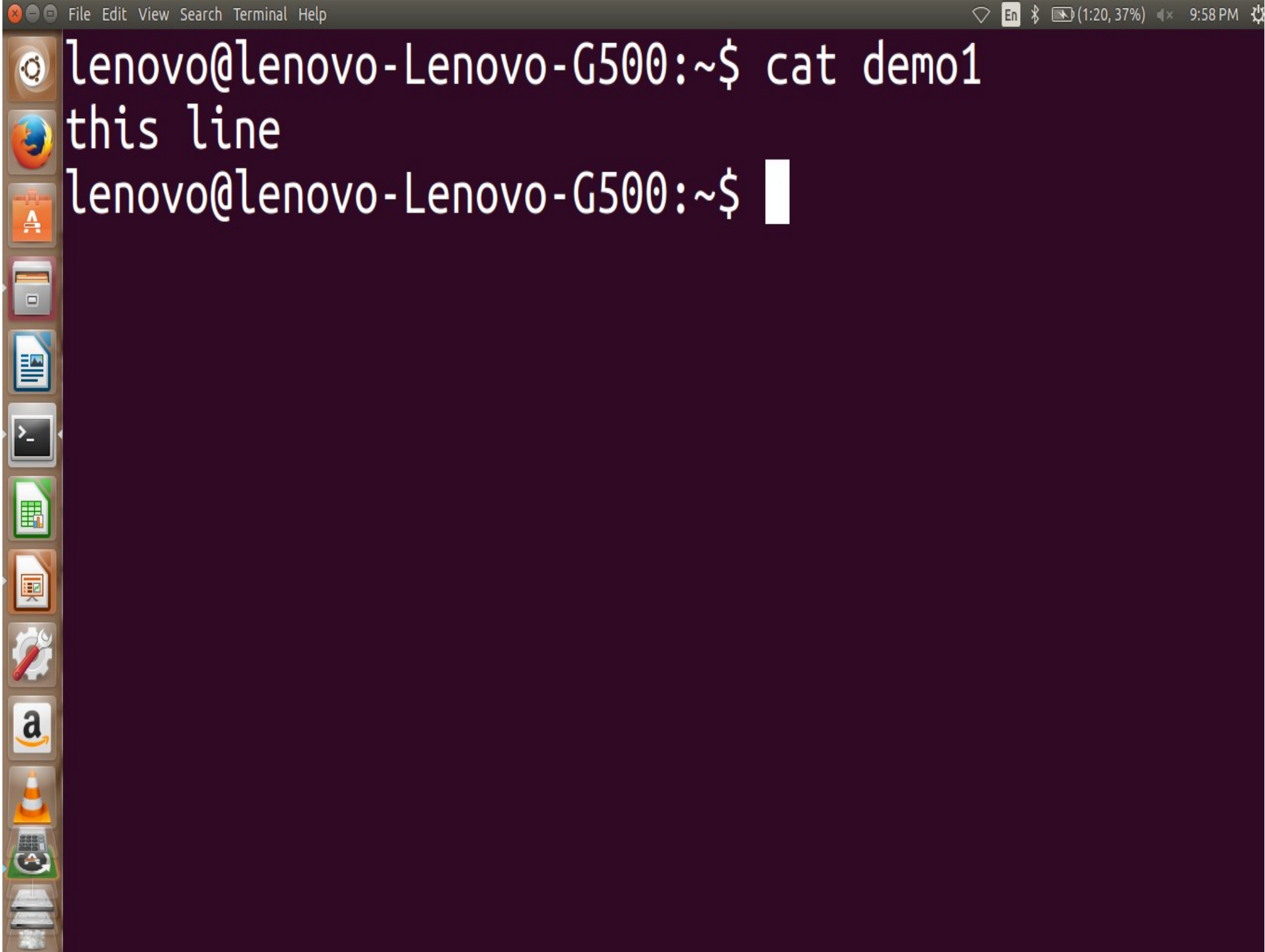
  who | grep kumar > foo

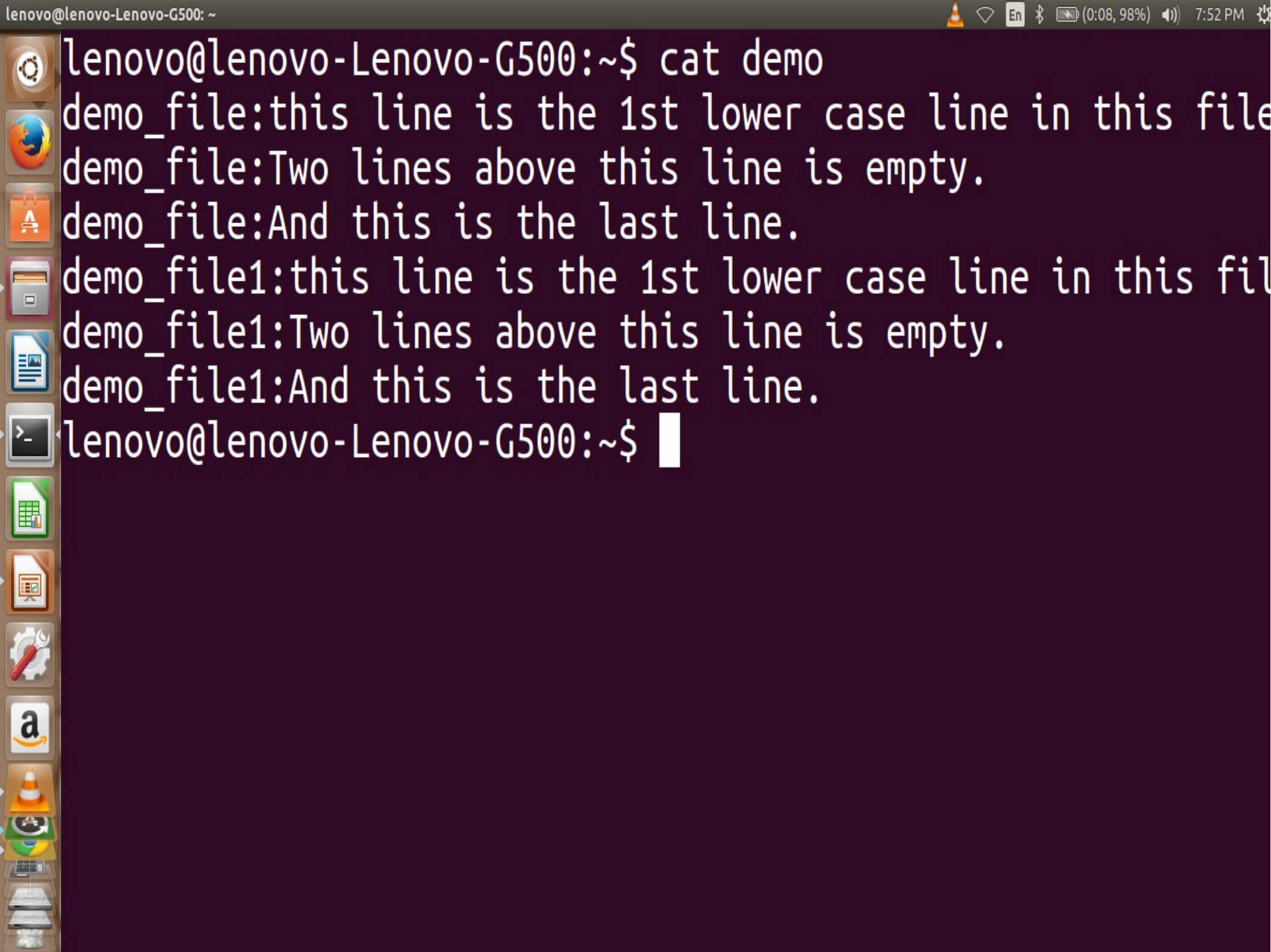
- Quoting is essential if the search string consits of more than one word or uses any of the shell's characters like *,$ etc
- grep returns the prompt in case the pattern can't be located

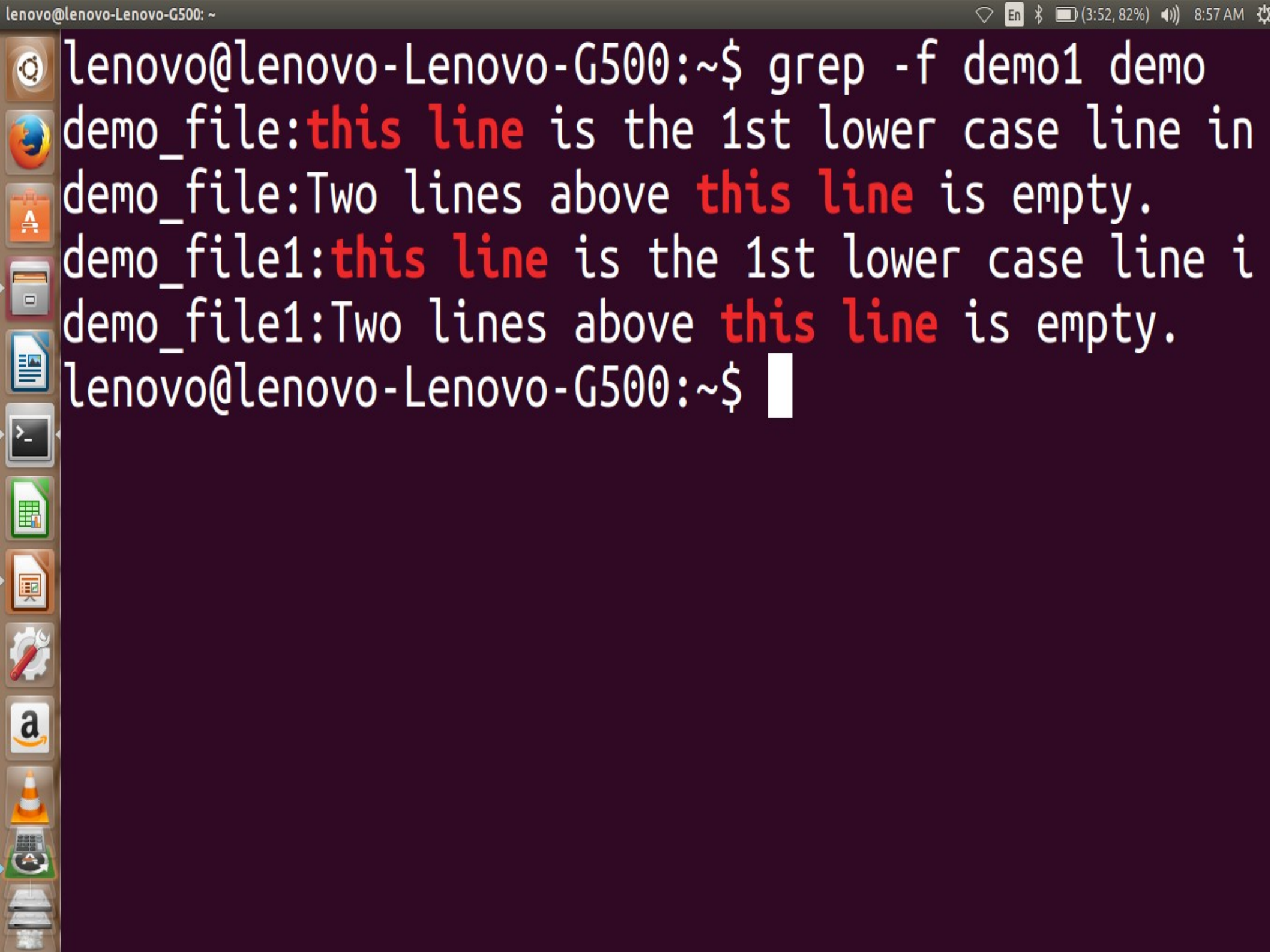

lenovo@lenovo-Lenovo-G500: ~
En
(0:16, 87%)
7:21 PM
lenovo@lenovo-Lenovo-G500:~$ grep this line demo
grep: line: No such file or directory
demo:demo_file:this line is the 1st lower case line in this file.
demo:demo_file:Two lines above this line is empty.
demo:demo_file:And this is the last line.
demo:demo_file1:this line is the 1st lower case line in this file.
demo:demo_file1:Two lines above this line is empty.
demo:demo_file1:And this is the last line.
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
En
(0:15, 87%)
7:22 PM
lenovo@lenovo-Lenovo-G500:~$ grep 'this line' demo
demo_file:this line is the 1st lower case line in this file
demo_file:Two lines above this line is empty.
demo_file1:this line is the 1st lower case line in this fil
demo_file1:Two lines above this line is empty.
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
En
(0:15, 88%)
7:22 PM
lenovo@lenovo-Lenovo-G500:~$ grep "this line" demo
demo_file:this line is the 1st lower case line in this file
.
demo_file:Two lines above this line is empty.
demo_file1:this line is the 1st lower case line in this fil
e.
demo_file1:Two lines above this line is empty.
lenovo@lenovo-Lenovo-G500:~$


File Edit View Search Terminal Help
En
(1:20, 37%)
9:58 PM
lenovo@lenovo-Lenovo-G500:~$ cat demo1
this line
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
En
(0:08, 98%)
7:52 PM
lenovo@lenovo-Lenovo-G500:~$ cat demo
demo_file:this line is the 1st lower case line in this file
demo_file:Two lines above this line is empty.
demo_file:And this is the last line.
demo_file1:this line is the 1st lower case line in this fil
demo_file1:Two lines above this line is empty.
demo_file1:And this is the last line.
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
En
(3:52, 82%)
8:57 AM
lenovo@lenovo-Lenovo-G500:~$ grep -f demo1 demo
demo_file:this line is the 1st lower case line in
demo_file:Two lines above this line is empty.
demo_file1:this line is the 1st lower case line i
demo_file1:Two lines above this line is empty.
lenovo@lenovo-Lenovo-G500:~$

Grep options

- Ignoring Case(-i)

  Ex:

  grep -i "the" demo_file

  THIS LINE IS **THE** 1ST UPPER CASE LINE IN THIS FILE.

  this line is **the** 1st lower case line in this file.

  This Line Has All Its First Character Of **The** Word With Upper Case.

  And this is **the** last line.

- Deleting Lines(-v):(-v)inverse option selects all except lines containing the pattern

  emp.lst=15 lines in that 4 lines contains director

  grep -v 'director' emp.lst > otherlist

  wc -l otherlist

  11 otherlist

Displaying line numbers(-n)

- Displays the line numbers containing the pattern along with the lines

  grep -n “marketing” emp.lst

  3:

  11:

Counting lines containing pattern(-c)

- Count (-c) option counts the number of lines containing pattern

Displaying Filenames(-l)

- Displays only the names of file containing the pattern

Matching Multiple Patterns(-e)

- $grep -e “Agarwal” -e “aggarwal” -e “agrawal” emp.lst

Sed:The stream Editor

- It is multipurpose tool which combines the work of several filters
- Syntax:

  sed options 'address action' file(s)
- Address and action are enclose with single quotes
- Addressing in sed is done in two ways:
- By one or two line numbers(like 3,7)
- By specifying a /-enclosed pattern which occurs in a line(like /From:/)

- sed processes several instructions in a sequential manner
- Each instruction operates on the output of the previous instruction

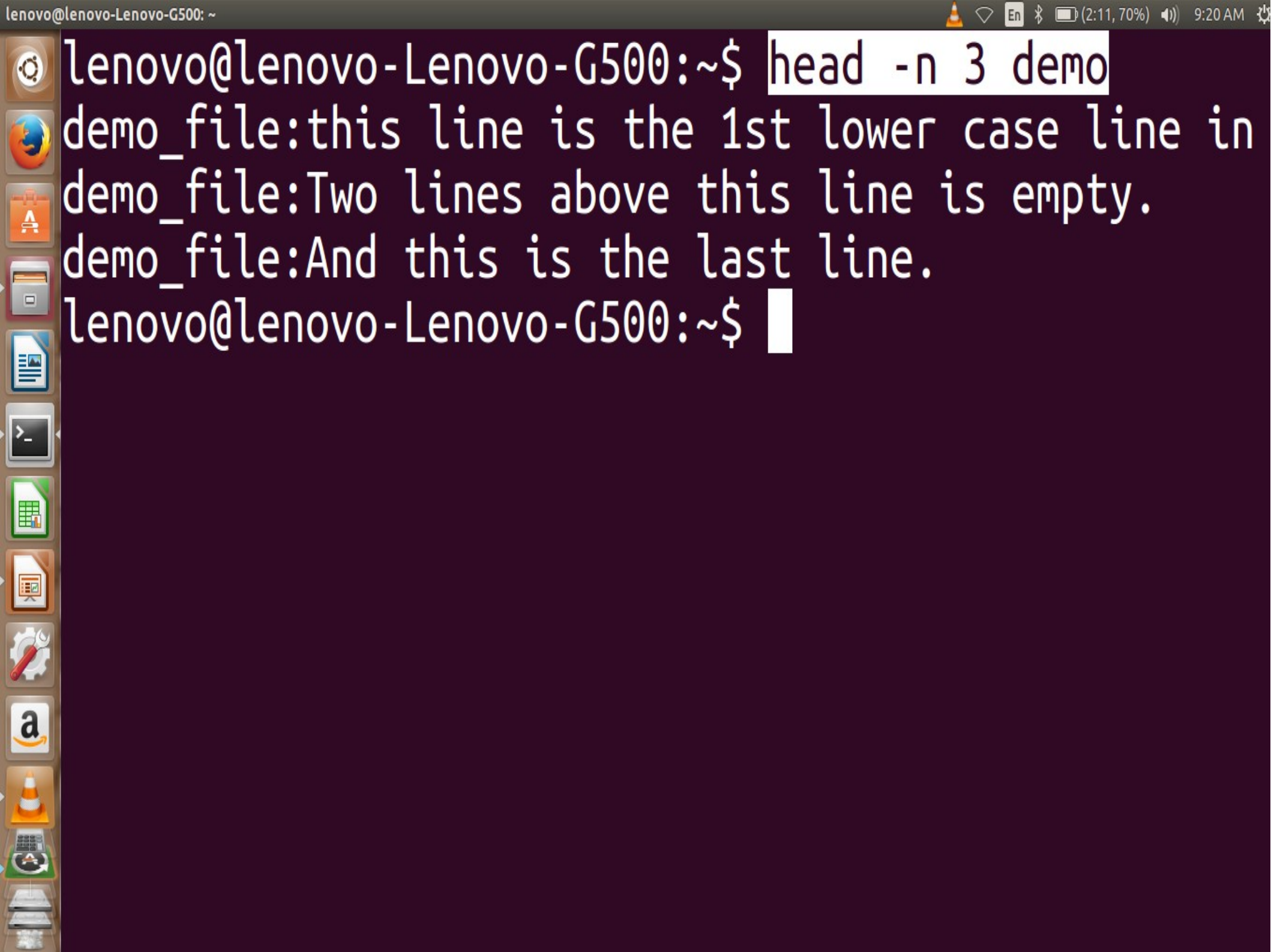
Line addressing

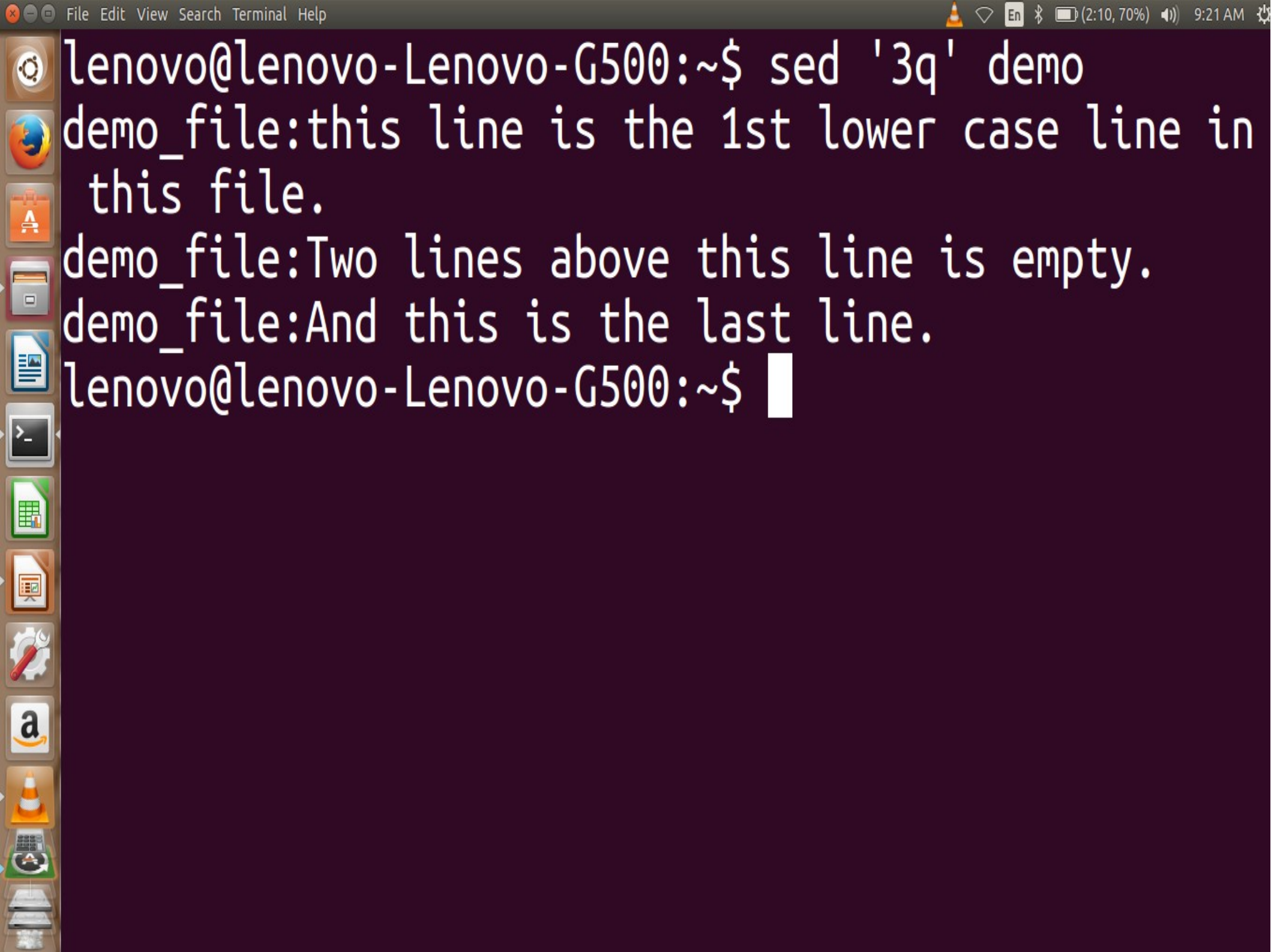
- $ sed '3q' emp.lst

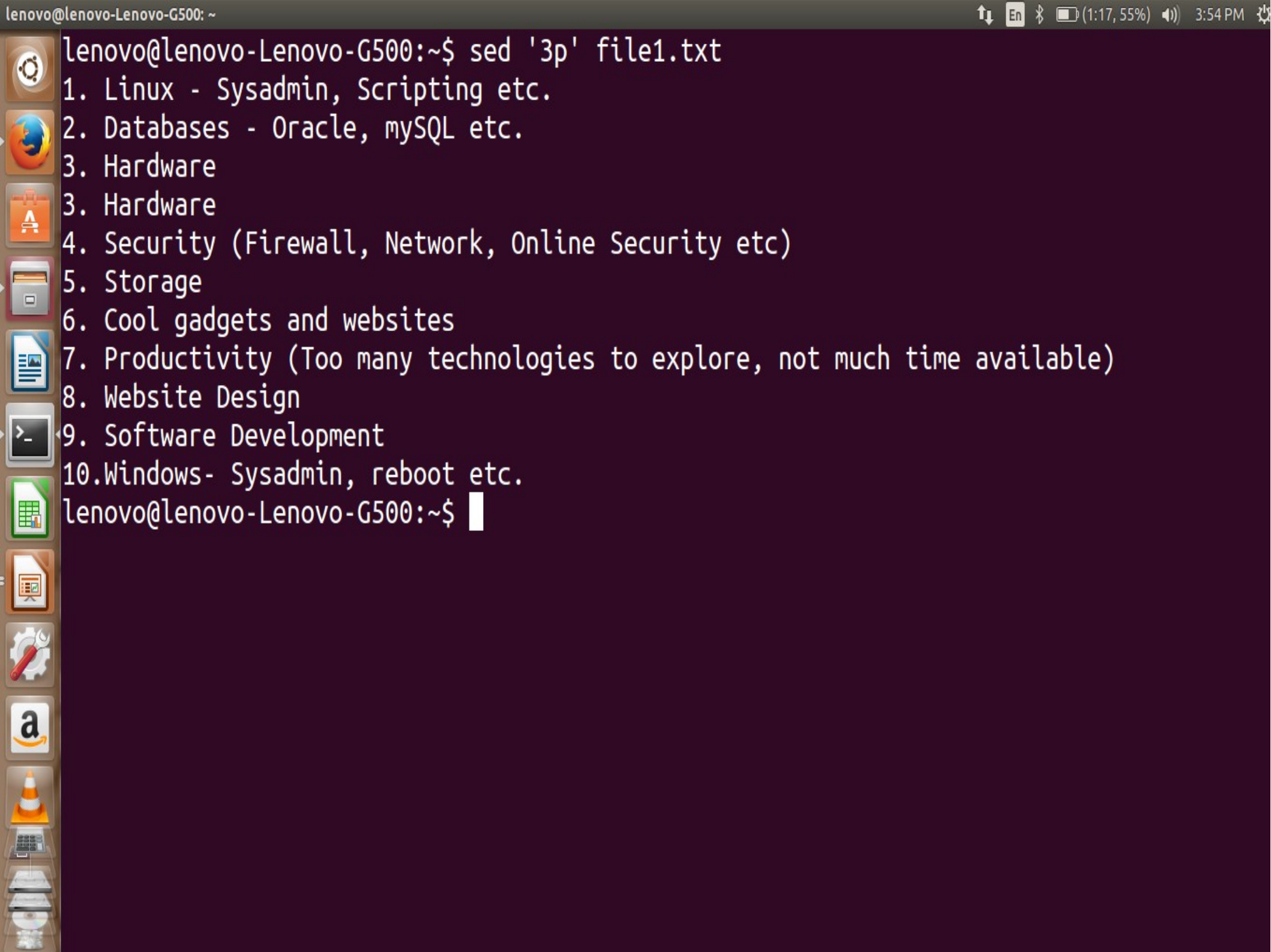
  where address=3

  action =q(quit)

  quits after line number 3


lenovo@lenovo-Lenovo-G500: ~
En
(2:11, 70%)
9:20 AM
lenovo@lenovo-Lenovo-G500:~$ head -n 3 demo
demo_file:this line is the 1st lower case line in
demo_file:Two lines above this line is empty.
demo_file:And this is the last line.
lenovo@lenovo-Lenovo-G500:~$


File Edit View Search Terminal Help
(2:10, 70%)
9:21 AM
lenovo@lenovo-Lenovo-G500:~$ sed '3q' demo
demo_file:this line is the 1st lower case line in
this file.
demo_file:Two lines above this line is empty.
demo_file:And this is the last line.
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
En
(1:17, 55%)
3:54 PM
lenovo@lenovo-Lenovo-G500:~$ sed '3p' file1.txt
1. Linux - Sysadmin, Scripting etc.
2. Databases - Oracle, mySQL etc.
3. Hardware
3. Hardware
4. Security (Firewall, Network, Online Security etc)
5. Storage
6. Cool gadgets and websites
7. Productivity (Too many technologies to explore, not much time available)
8. Website Design
9. Software Development
10.Windows- Sysadmin, reboot etc.
lenovo@lenovo-Lenovo-G500:~$

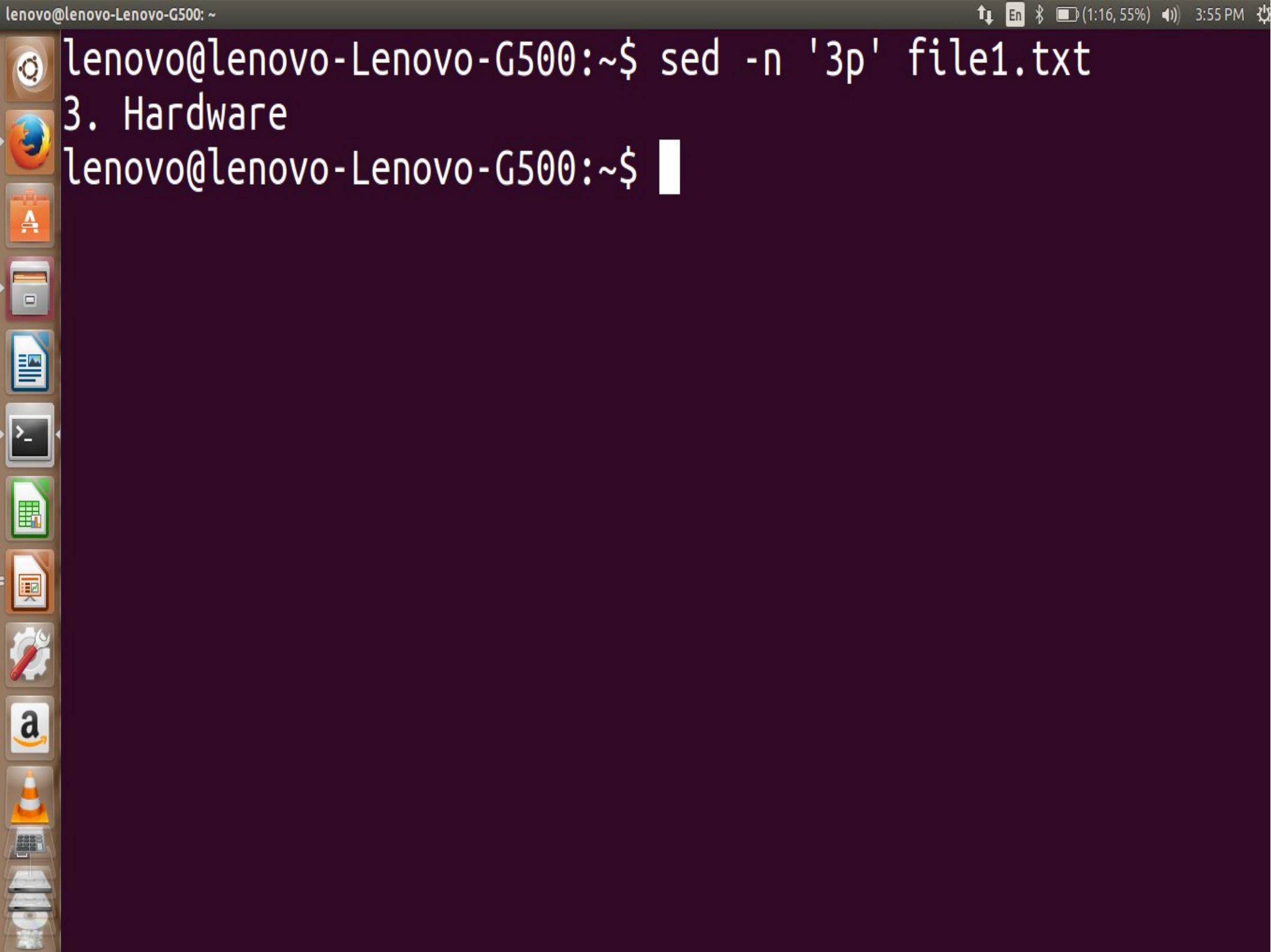
- To suppress automatic printing of pattern space use -n command with sed. sed -n option will not print anything, unless an explicit request to print is found.
- $ sed -n '1,2p' emp.lst

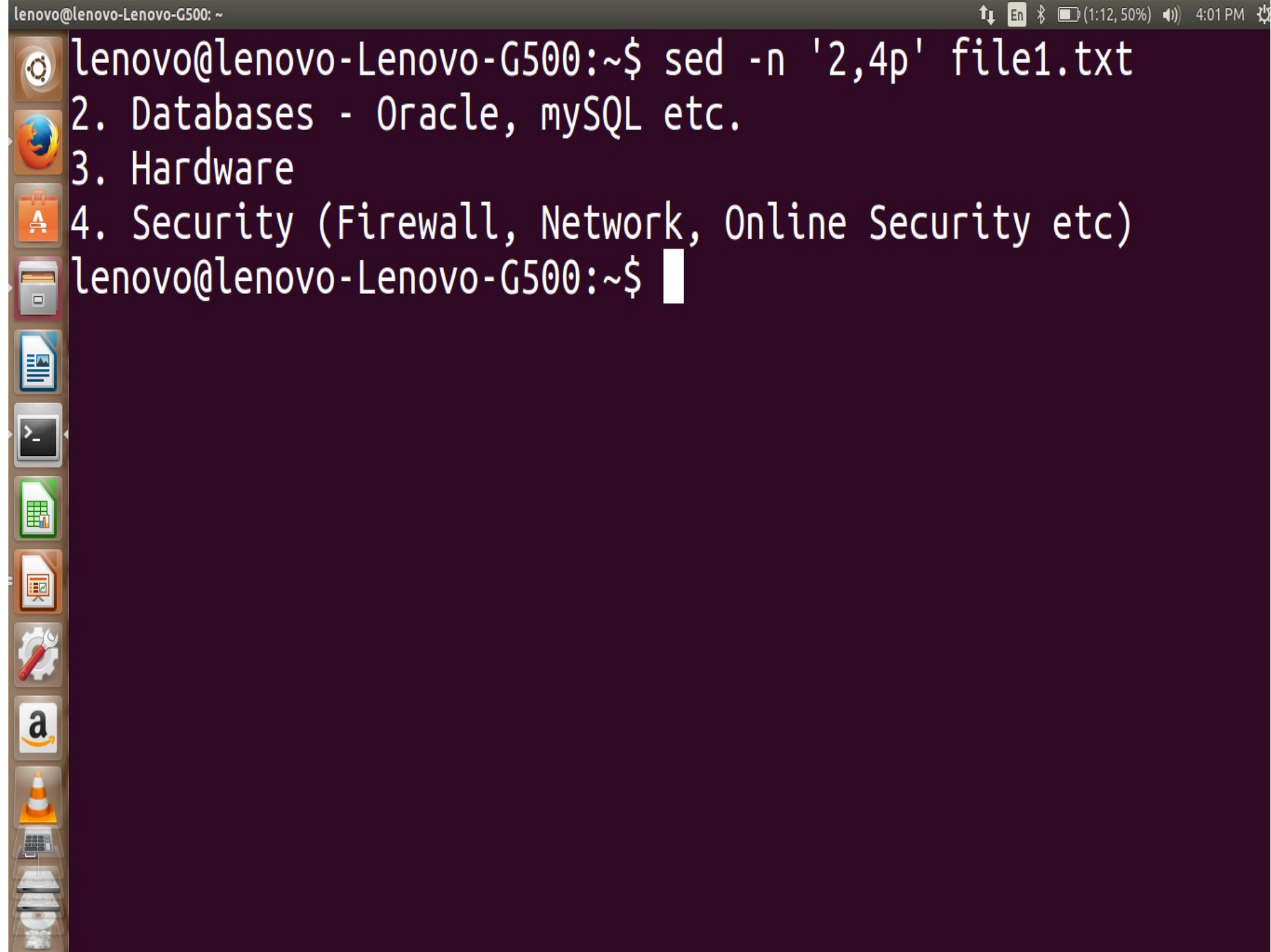
  where p(print)
- To select the last line of file, use the $

  $sed -n '$p' emp.lst

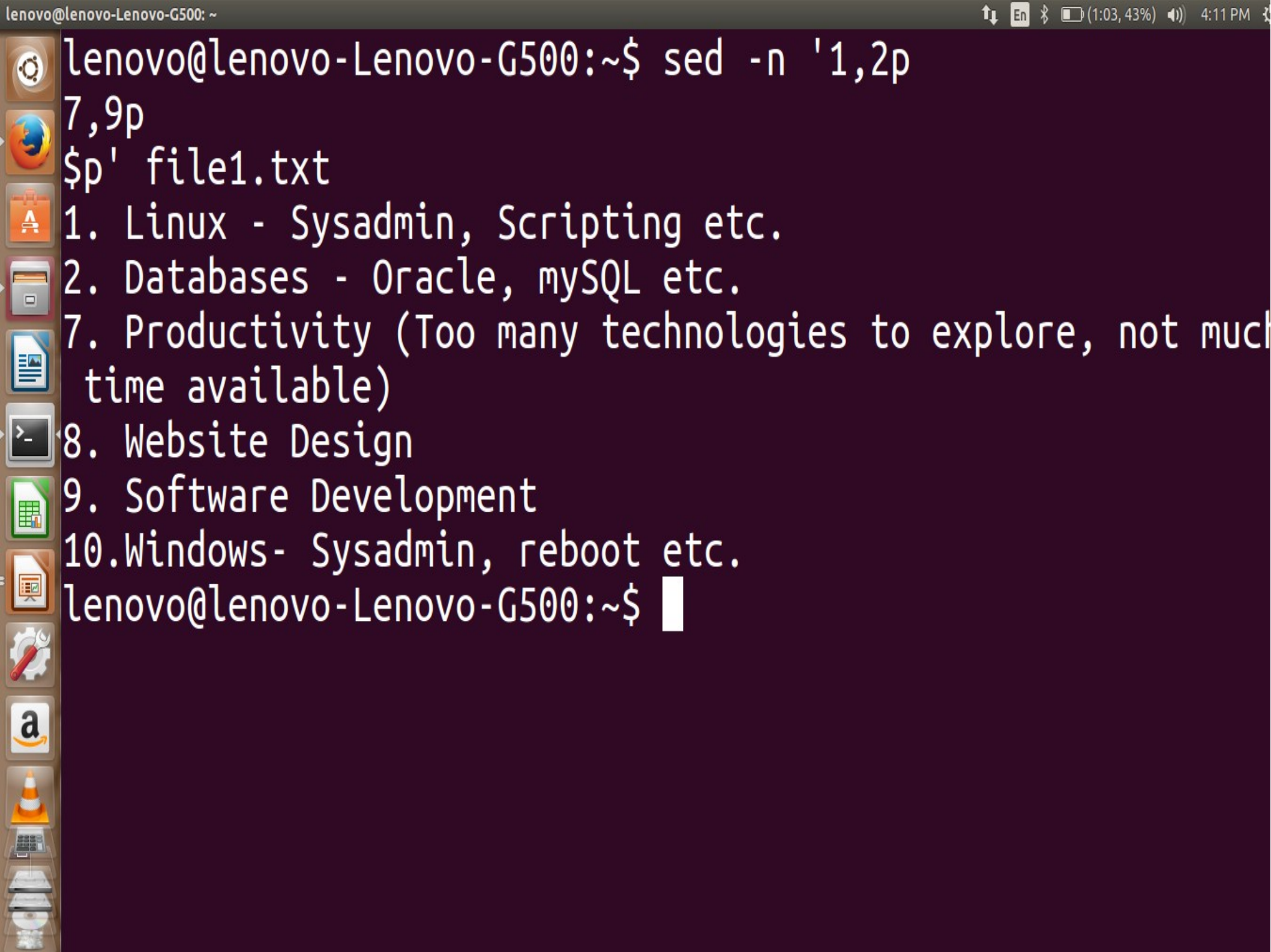
Selecting Line from anywhere

- $sed -n '3,7p' emp.lst


lenovo@lenovo-Lenovo-G500: ~
En
(1:16, 55%)
3:55 PM
lenovo@lenovo-Lenovo-G500:~$ sed -n '3p' file1.txt
3. Hardware
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
En
(1:12, 50%)
4:01 PM
lenovo@lenovo-Lenovo-G500:~$ sed -n '2,4p' file1.txt
2. Databases - Oracle, mySQL etc.
3. Hardware
4. Security (Firewall, Network, Online Security etc)
lenovo@lenovo-Lenovo-G500:~$

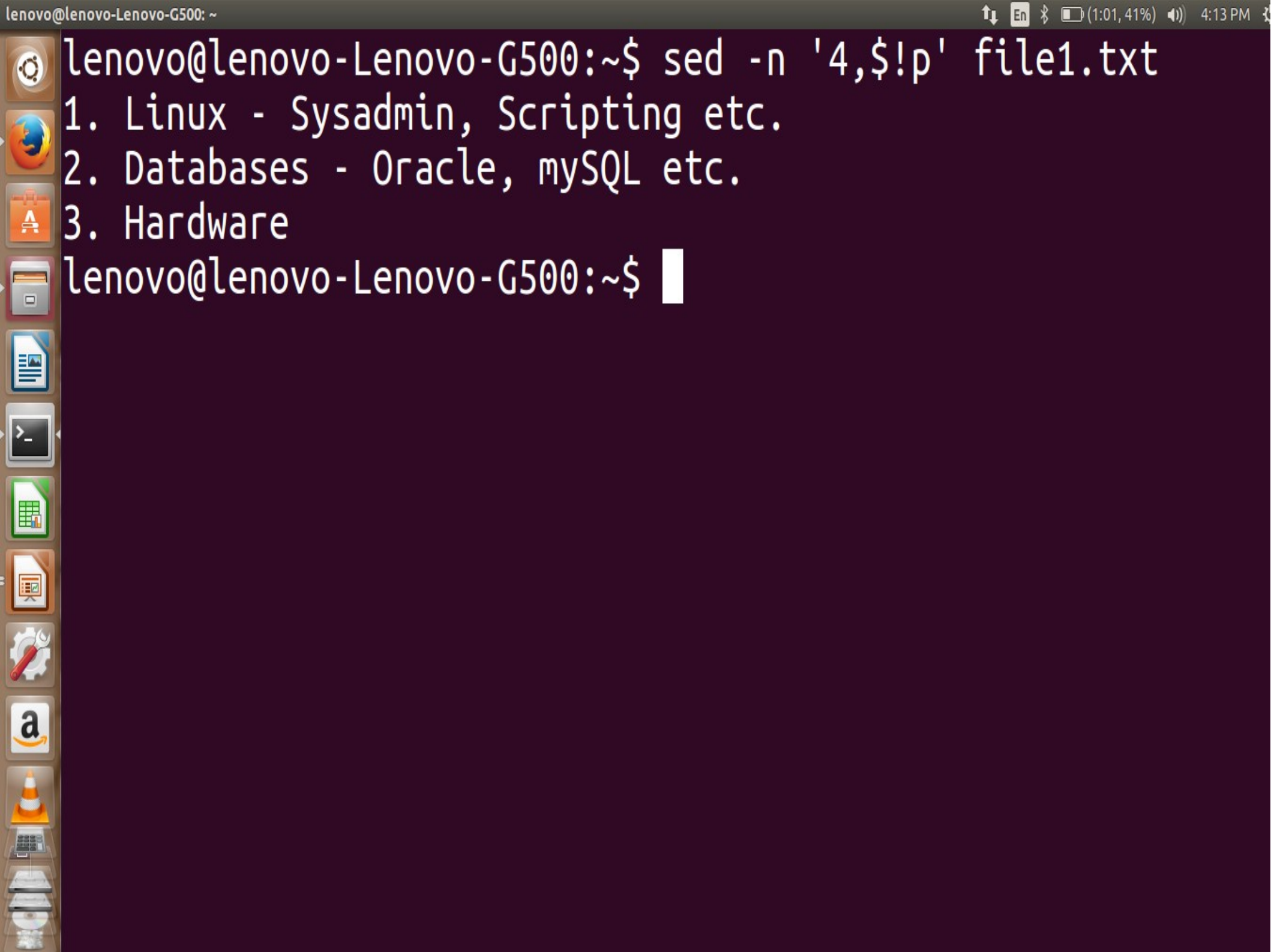
Selecting Multiple groups of lines


lenovo@lenovo-Lenovo-G500: ~
(1:03, 43%)
4:11 PM
lenovo@lenovo-Lenovo-G500:~$ sed -n '1,2p
7,9p
$p' file1.txt
1. Linux - Sysadmin, Scripting etc.
2. Databases - Oracle, mySQL etc.
7. Productivity (Too many technologies to explore, not much
time available)
8. Website Design
9. Software Development
10.Windows- Sysadmin, reboot etc.
lenovo@lenovo-Lenovo-G500:~$

Negating the action(!)

- $sed -n '3,$!p' emp.lst

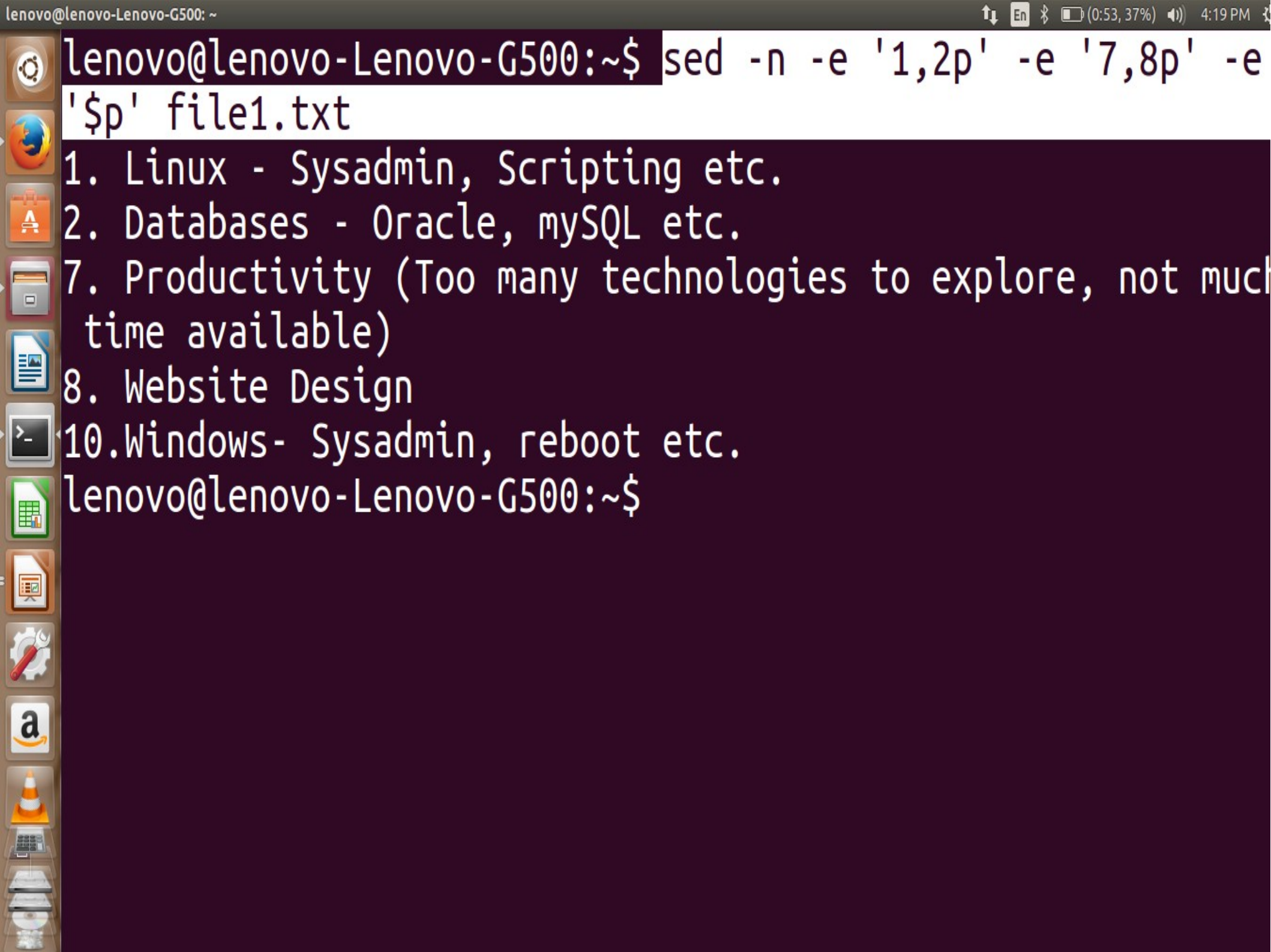
  or sed -n '1,2p' emp.lst

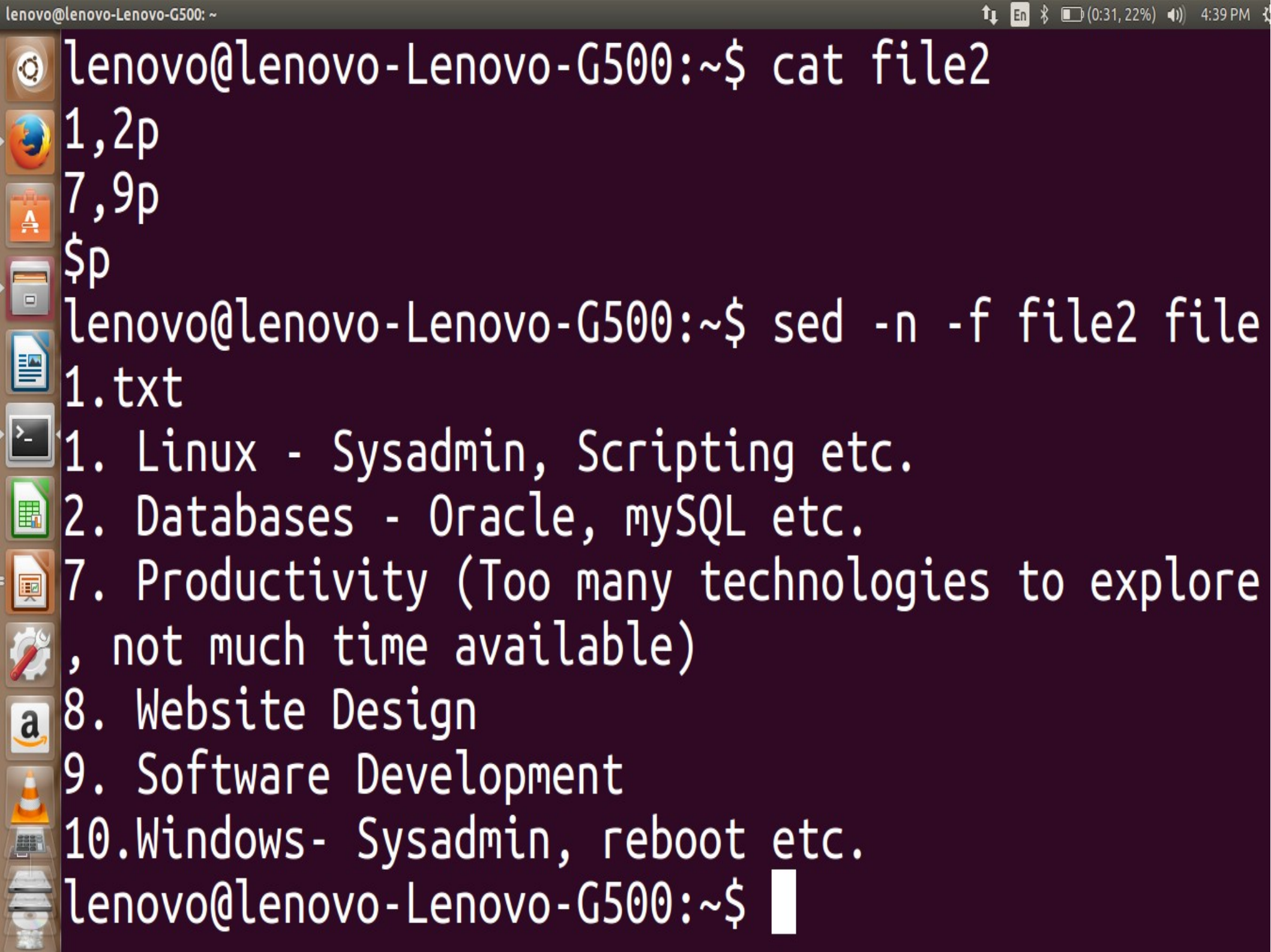

lenovo@lenovo-Lenovo-G500: ~
(1:01, 41%)
4:13 PM
lenovo@lenovo-Lenovo-G500:~$ sed -n '4,$!p' file1.txt
1. Linux - Sysadmin, Scripting etc.
2. Databases - Oracle, mySQL etc.
3. Hardware
lenovo@lenovo-Lenovo-G500:~$

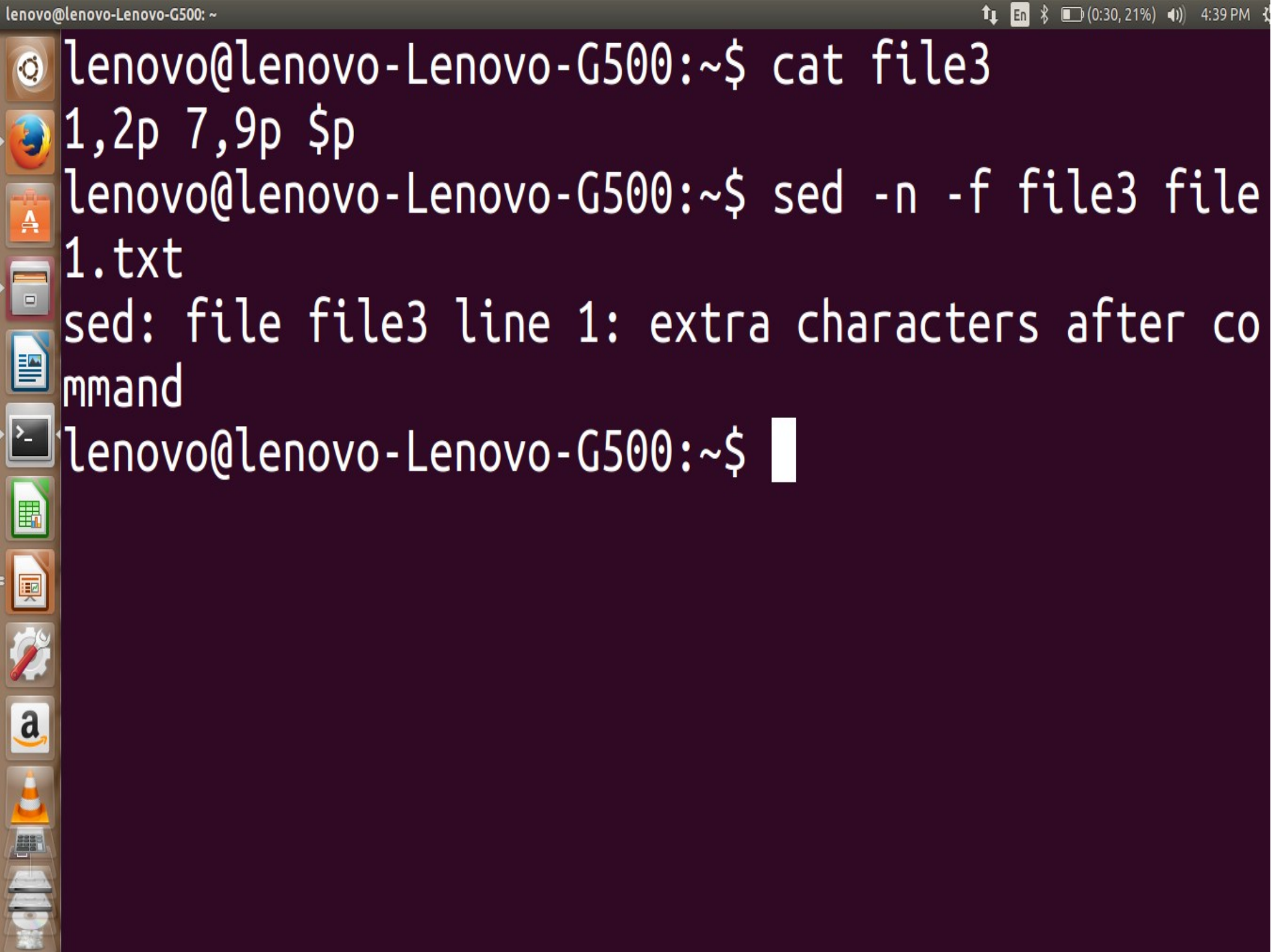
Using Multiple instructions(-e and -f)

- e option allows you to enter as many instructions

- -f option to direct sed to take its instructions from file

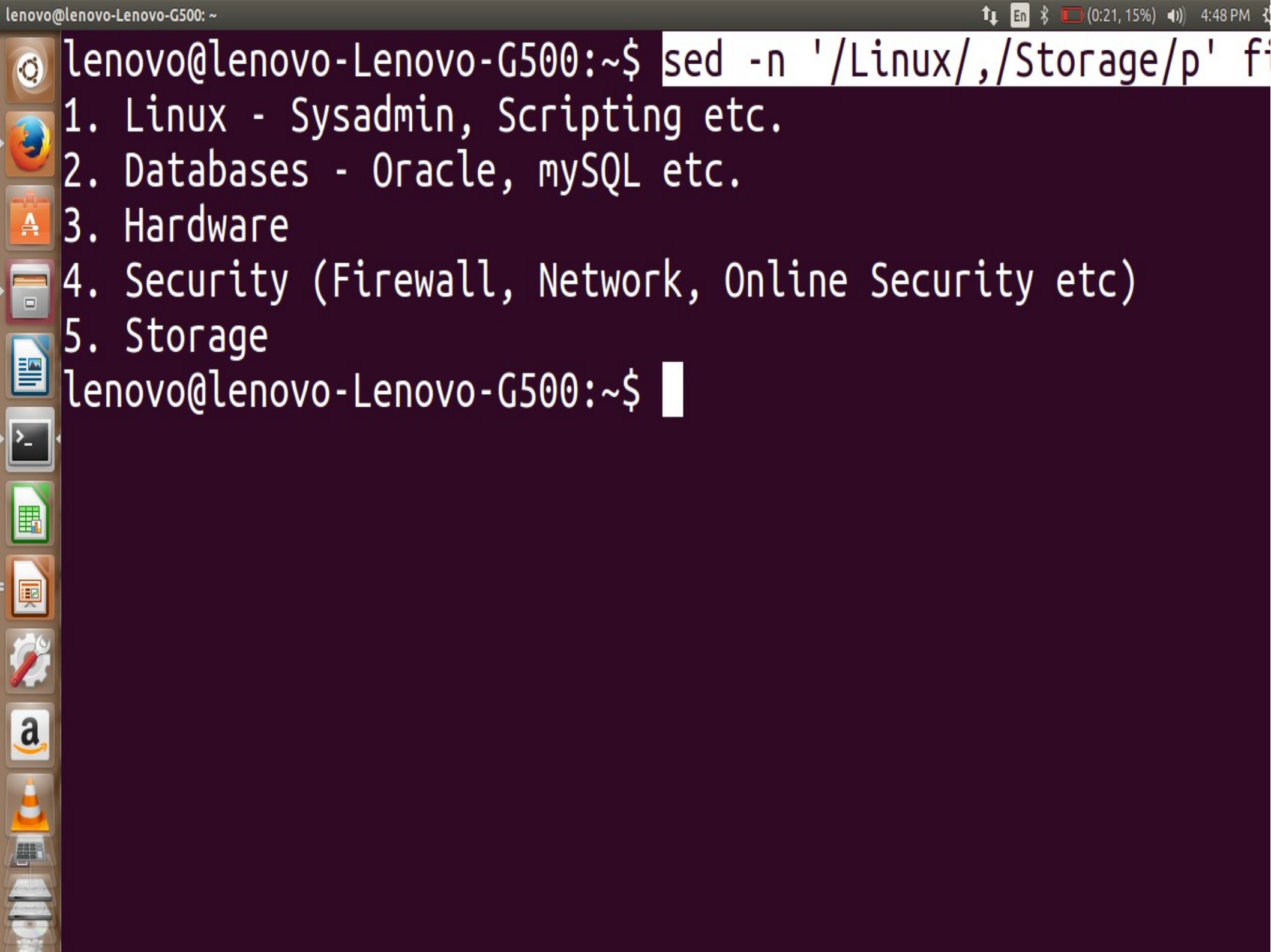

lenovo@lenovo-Lenovo-G500: ~
(0:53, 37%)
4:19 PM
lenovo@lenovo-Lenovo-G500:~$ sed -n -e '1,2p' -e '7,8p' -e '$p' file1.txt
1. Linux - Sysadmin, Scripting etc.
2. Databases - Oracle, mySQL etc.
7. Productivity (Too many technologies to explore, not much time available)
8. Website Design
10.Windows- Sysadmin, reboot etc.
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
(0:31, 22%)
4:39 PM
lenovo@lenovo-Lenovo-G500:~$ cat file2
1,2p
7,9p
$p
lenovo@lenovo-Lenovo-G500:~$ sed -n -f file2 file
1.txt
1. Linux - Sysadmin, Scripting etc.
2. Databases - Oracle, mySQL etc.
7. Productivity (Too many technologies to explore
, not much time available)
8. Website Design
9. Software Development
10.Windows- Sysadmin, reboot etc.
lenovo@lenovo-Lenovo-G500:~$

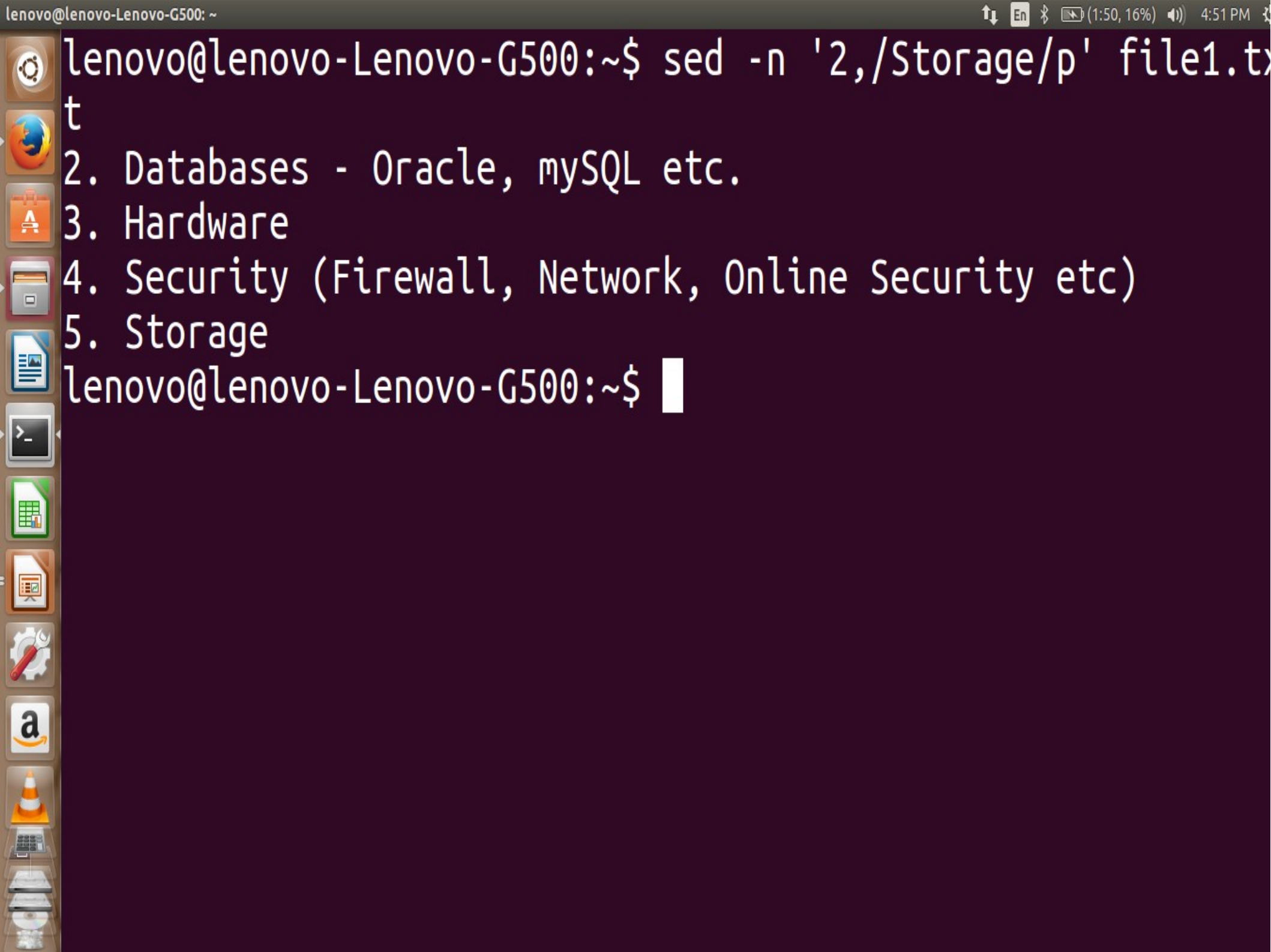

lenovo@lenovo-Lenovo-G500: ~
(0:30, 21%)
4:39 PM
lenovo@lenovo-Lenovo-G500:~$ cat file3
1,2p 7,9p $p
lenovo@lenovo-Lenovo-G500:~$ sed -n -f file3 file
1.txt
sed: file file3 line 1: extra characters after co
mmand
lenovo@lenovo-Lenovo-G500:~$

Context Addressing

- One or two patterns to locate lines.The pattern must be bounded by a / on either side


lenovo@lenovo-Lenovo-G500: ~
(0:21, 15%)
4:48 PM
lenovo@lenovo-Lenovo-G500:~$ sed -n '/Linux/,/Storage/p' fi
1. Linux - Sysadmin, Scripting etc.
2. Databases - Oracle, mySQL etc.
3. Hardware
4. Security (Firewall, Network, Online Security etc)
5. Storage
lenovo@lenovo-Lenovo-G500:~$

- Comma separated pair of context addresses to select a group of lines


lenovo@lenovo-Lenovo-G500: ~
En
(1:50, 16%)
4:51 PM
lenovo@lenovo-Lenovo-G500:~$ sed -n '2,/Storage/p' file1.tx
t
2. Databases - Oracle, mySQL etc.
3. Hardware
4. Security (Firewall, Network, Online Security etc)
5. Storage
lenovo@lenovo-Lenovo-G500:~$

Using Regular expression

$sed -n '/[aA]gg*[ar][ar]wal/p' emp.lst

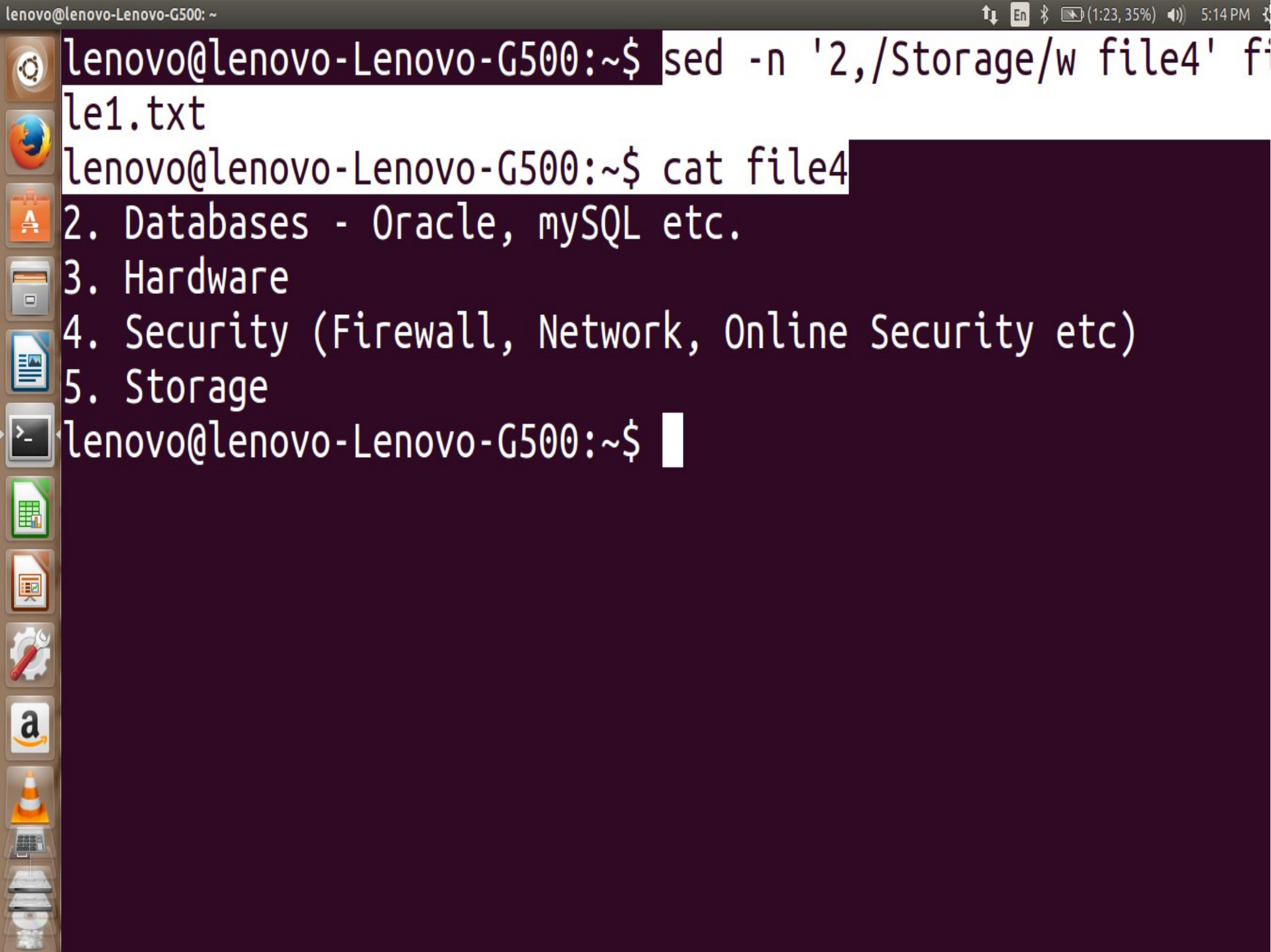
$sed -n '/50.....$/p' emp.lst

50| 7000

to locate all people born in 1950

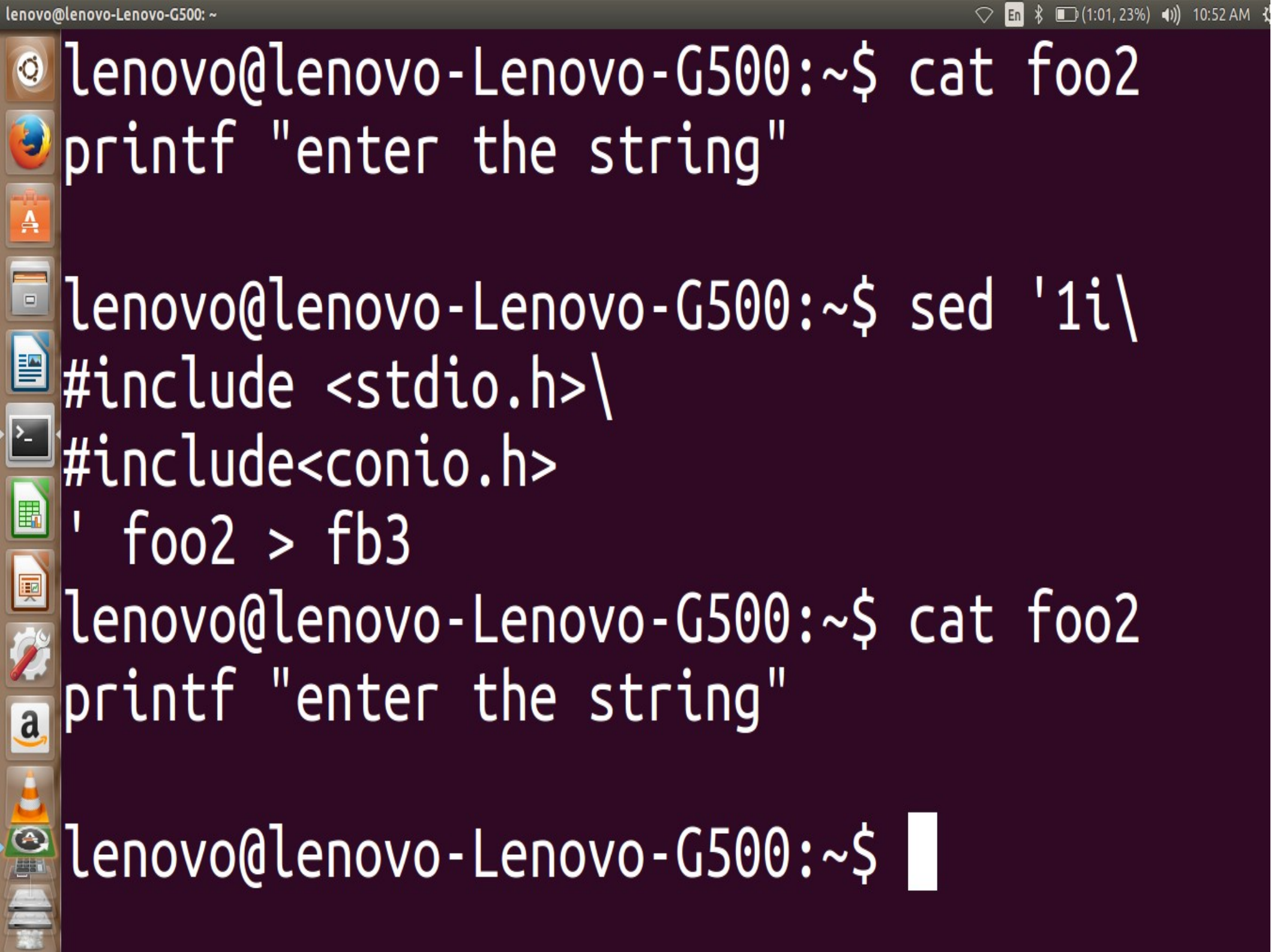
Writing Selected lines to a file(w)

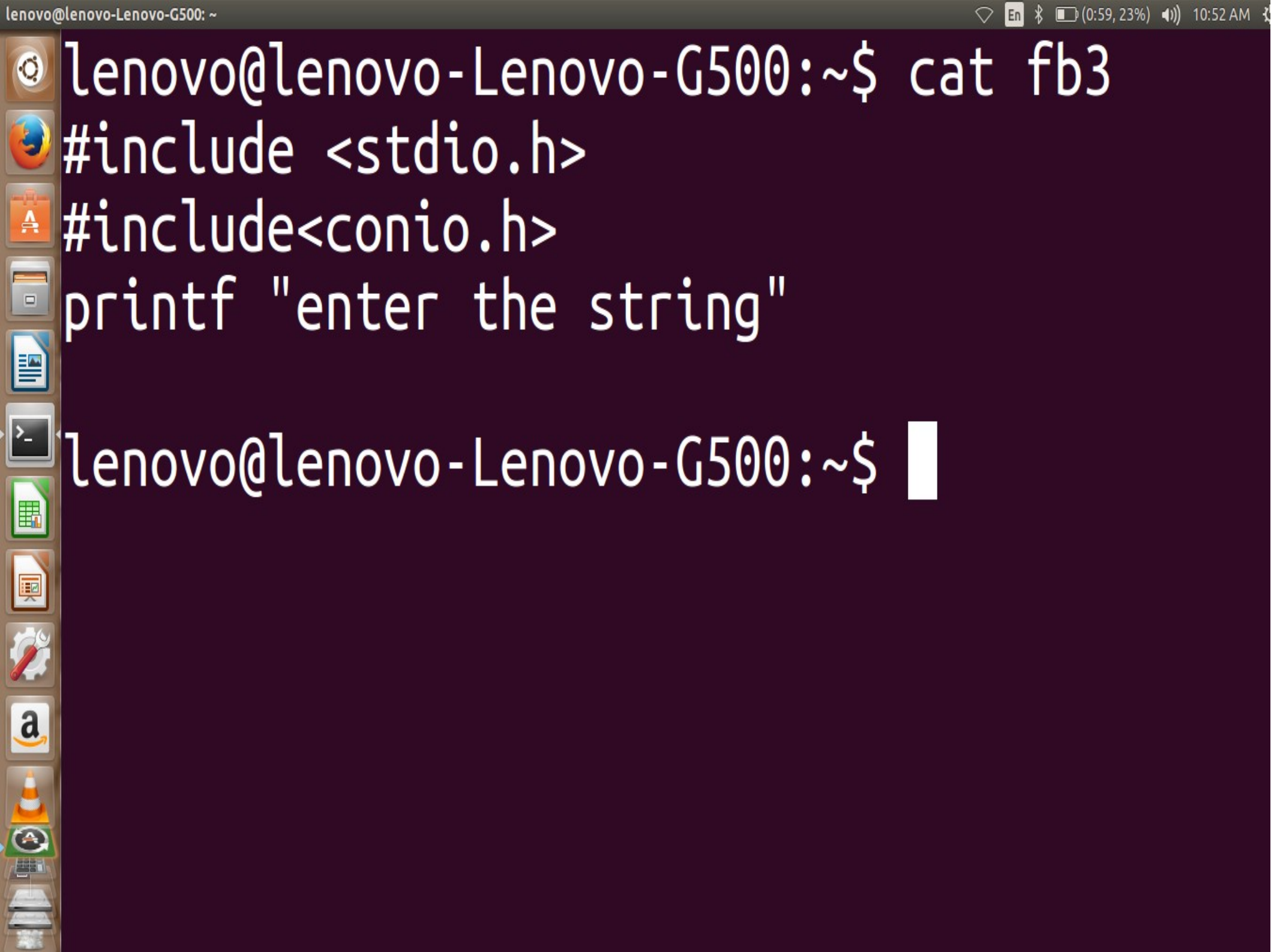
- w command to write the selected lines to a separate file

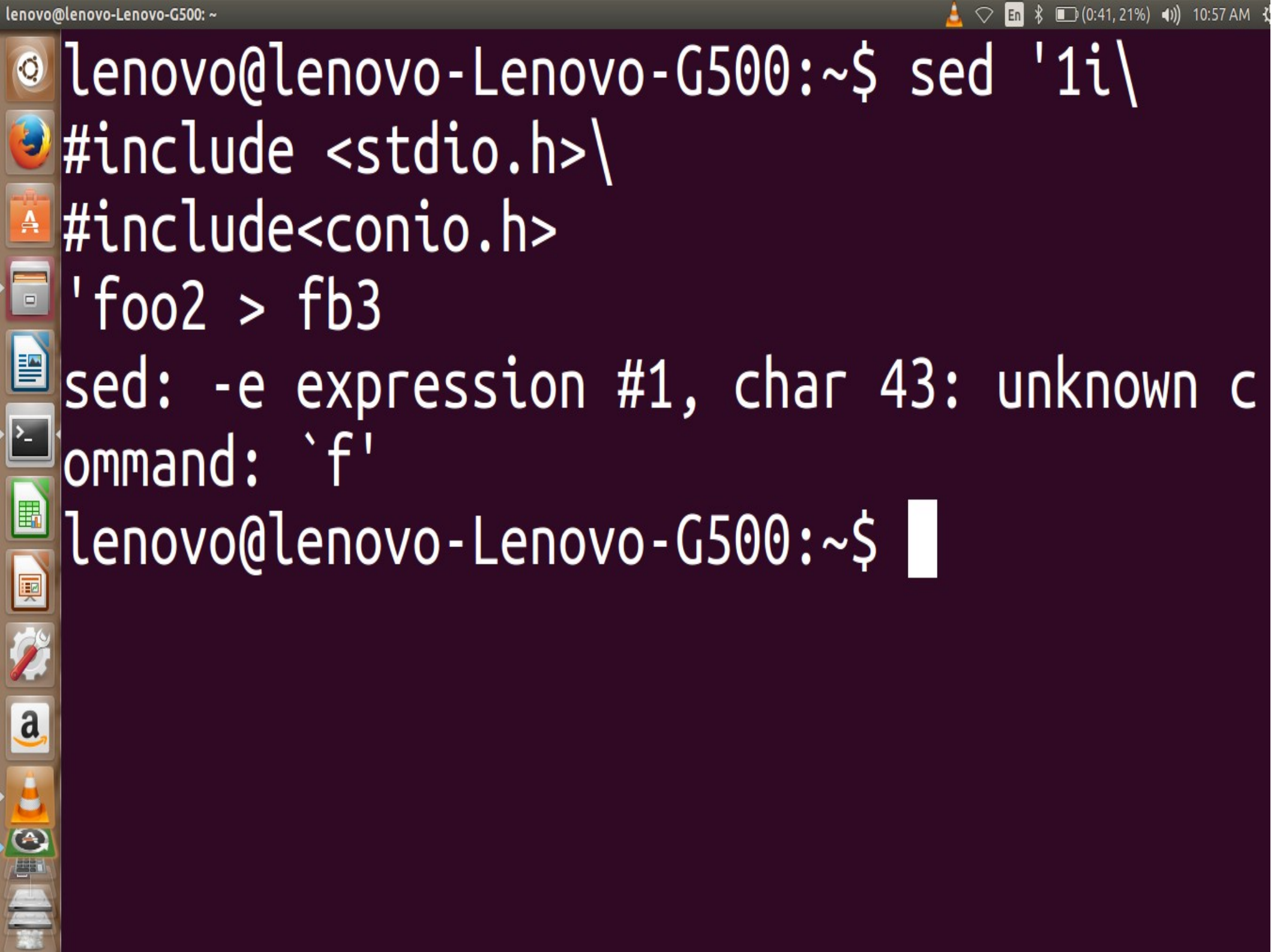

lenovo@lenovo-Lenovo-G500: ~
En
(1:23, 35%)
5:14 PM
lenovo@lenovo-Lenovo-G500:~$ sed -n '2,/Storage/w file4' fi
le1.txt
lenovo@lenovo-Lenovo-G500:~$ cat file4
2. Databases - Oracle, mySQL etc.
3. Hardware
4. Security (Firewall, Network, Online Security etc)
5. Storage
lenovo@lenovo-Lenovo-G500:~$

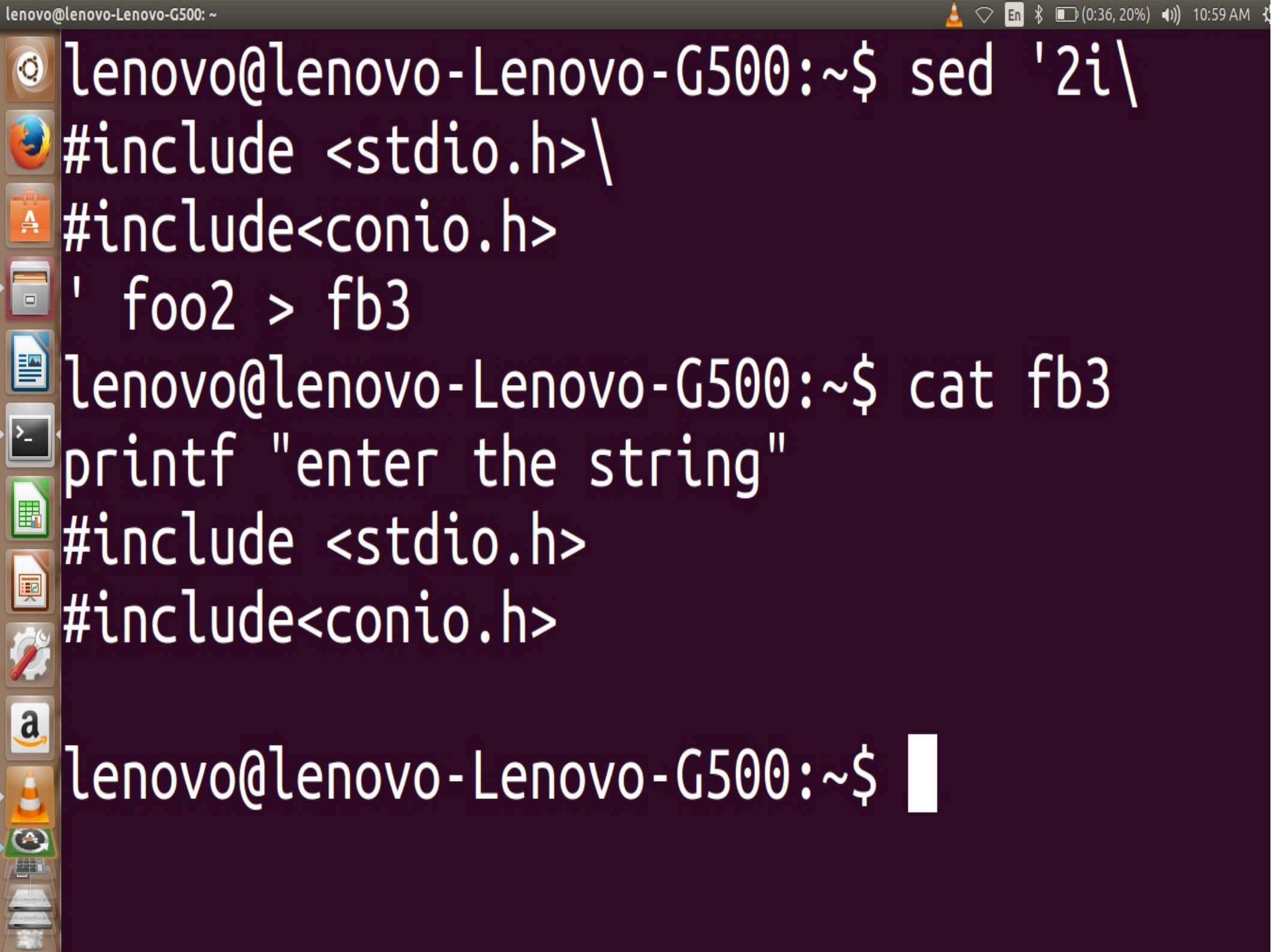
Text Editing

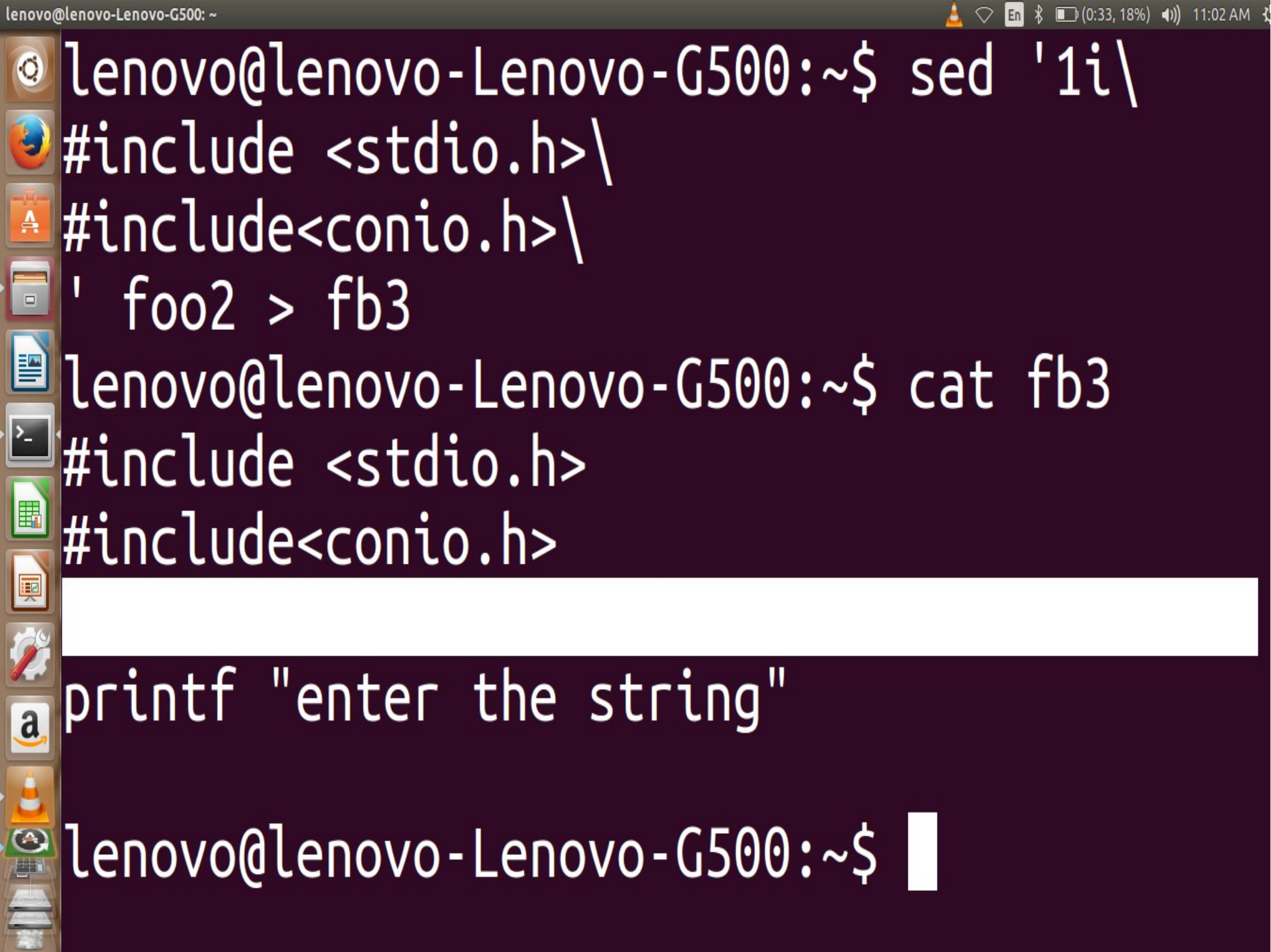
- Sed can insert text and change existing text in a file(i,a,c)

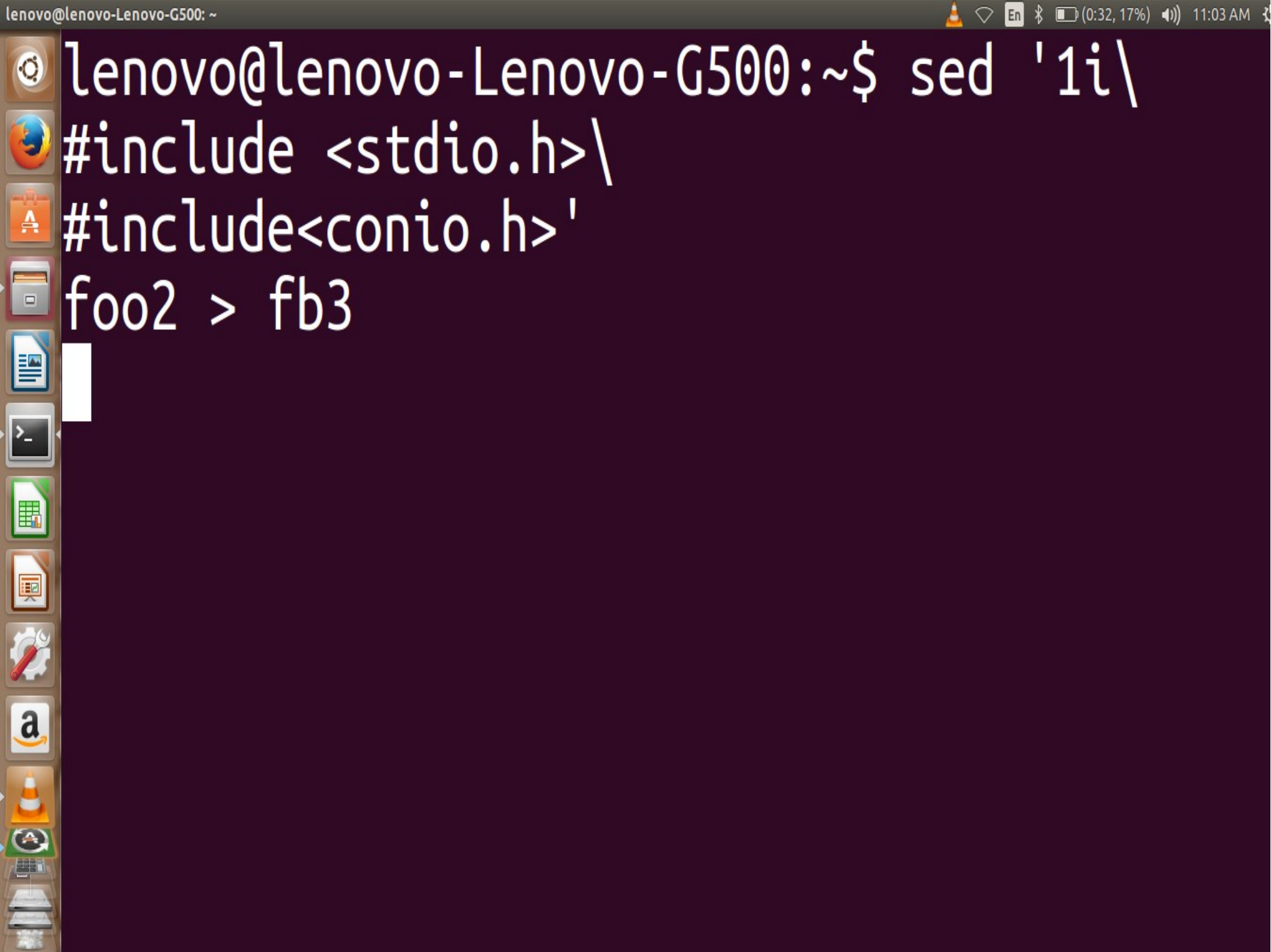

lenovo@lenovo-Lenovo-G500: ~
En
(1:01, 23%)
10:52 AM
lenovo@lenovo-Lenovo-G500:~$ cat foo2
printf "enter the string"
lenovo@lenovo-Lenovo-G500:~$ sed '1i\
#include <stdio.h>\
#include<conio.h>
' foo2 > fb3
lenovo@lenovo-Lenovo-G500:~$ cat foo2
printf "enter the string"
lenovo@lenovo-Lenovo-G500:~$

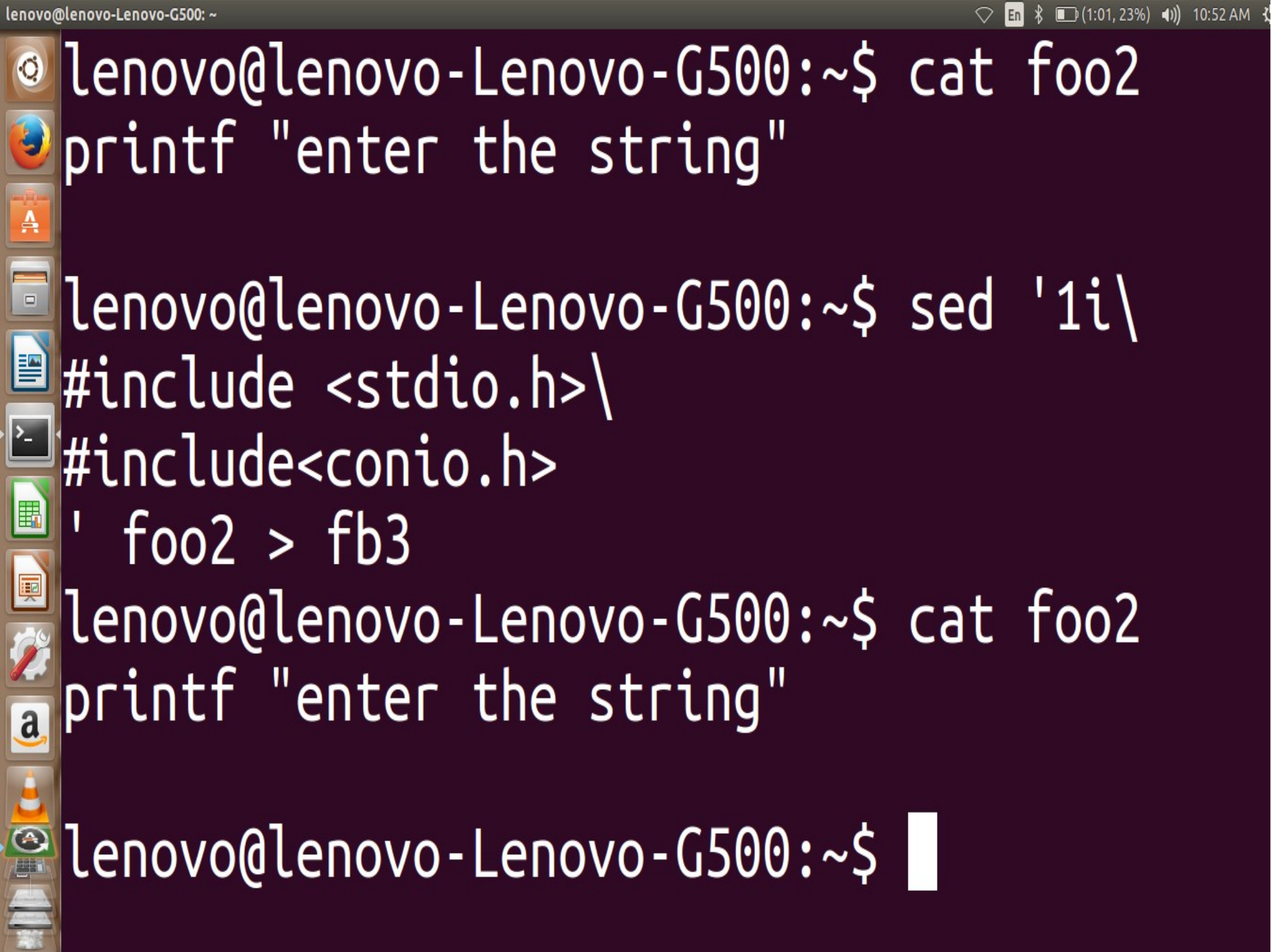

lenovo@lenovo-Lenovo-G500: ~
En
(0:59, 23%)
10:52 AM
lenovo@lenovo-Lenovo-G500:~$ cat fb3
#include <stdio.h>
#include<conio.h>
printf "enter the string"
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
En
(0:41, 21%)
10:57 AM
lenovo@lenovo-Lenovo-G500:~$ sed '1i\
#include <stdio.h>\
#include<conio.h>
'foo2 > fb3
sed: -e expression #1, char 43: unknown c
ommand: `f'
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
(0:36, 20%)
10:59 AM
lenovo@lenovo-Lenovo-G500:~$ sed '2i\
#include <stdio.h>\
#include<conio.h>
' foo2 > fb3
lenovo@lenovo-Lenovo-G500:~$ cat fb3
printf "enter the string"
#include <stdio.h>
#include<conio.h>
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
(0:33, 18%)
11:02 AM
lenovo@lenovo-Lenovo-G500:~$ sed '1i\
#include <stdio.h>\
#include<conio.h>\
' foo2 > fb3
lenovo@lenovo-Lenovo-G500:~$ cat fb3
#include <stdio.h>
#include<conio.h>
printf "enter the string"
lenovo@lenovo-Lenovo-G500:~$


lenovo@lenovo-Lenovo-G500: ~
(0:32, 17%)
11:03 AM
lenovo@lenovo-Lenovo-G500:~$ sed '1i\
#include <stdio.h>\
#include<conio.h>'
foo2 > fb3


lenovo@lenovo-Lenovo-G500: ~
En
(1:01, 23%)
10:52 AM
lenovo@lenovo-Lenovo-G500:~$ cat foo2
printf "enter the string"
lenovo@lenovo-Lenovo-G500:~$ sed '1i\
#include <stdio.h>\
#include<conio.h>
' foo2 > fb3
lenovo@lenovo-Lenovo-G500:~$ cat foo2
printf "enter the string"
lenovo@lenovo-Lenovo-G500:~$

Deleting Lines(d)

- $sed '/director/d' emp.lst > olist

  -n option not to be used with d
- $sed -n '/director/!p' emp.lst > olist

  selects all lines except those containing director, and saves them in olist