

Data processing commands (extract, sort and filter)

Data processing commands (extract, sort and filter)

Remember, it's true that there are many graphics programs **that are easy to use**.

On the other hand, the following commands are generally used by **system administrators because of their** :

- **Their efficiency**
- **Their speed**
- **Their power**

Data processing commands (extract, sort and filter)

The **head** command

Data processing commands (extract, sort and filter)

The **head** command displays the first lines of a file

```
$ head [options] files
```

Example of use:

If a system administrator **wants** to read **just 10 lines** from **Double file**

Data processing commands (extract, sort and filter)

```
meskaldji@DESKTOP-GH8V564:~$ head -n10 Double  
/usr/bin/  
/usr/bin/aa-enabled  
/usr/bin/aa-exec  
/usr/bin/add-apt-repository  
/usr/bin/addpart  
/usr/bin/apport-bug  
/usr/bin/apport-cli  
/usr/bin/apport-collect  
/usr/bin/apport-unpack  
/usr/bin/appres
```

Data processing commands (extract, sort and filter)

Is there a command that displays the last lines of a file?

Data processing commands (extract, sort and filter)

The **tail** command

Data processing commands (extract, sort and filter)

The **tail** command displays the last lines of a file

```
$ tail [options] files
```

Example of use:

If a system administrator **just** wants to **read the last 10 lines** of the **list file**

Data processing commands (extract, sort and filter)

```
:~$ tail -n10 list
```

```
-rw-r--r-- 1 meskaldji meskaldji 1876096 Nov 10 21:49 system.zip
-rw-r--r-- 1 meskaldji meskaldji    55 Dec  8  2023 tecmint.txt
-rw-r--r-- 1 meskaldji meskaldji   588 Nov 16 21:28 test.txt
-rw-r--r-- 1 meskaldji meskaldji    81 Dec  2  2023 test1
-rw-r--r-- 1 meskaldji meskaldji   16 Jan 17  2024 testing
drwxr-xr-x 1 meskaldji meskaldji   512 Oct 27 11:41 testo
-rw-r--r-- 1 meskaldji meskaldji  1519 May  6  2024 testv.c
-rwxr--r-- 1 meskaldji meskaldji   29 Oct 27 12:00 tp1
drwxr-xr-x 1 meskaldji meskaldji   512 Oct 27 11:43 unix
-rwxr--r-- 1 meskaldji meskaldji   15 May  4  2024 welcome
```

Data processing commands (extract, sort and filter)

Are there any other commands that display file contents?

Data processing commands (extract, sort and filter)

The **cat**, **more**,
less, **,,,,** commands

Data processing commands (extract, sort and filter)

These commands allow viewing of file content.

- **cat:** displays the entire contents of the file.
- **more:** displays the contents of the file page by page without going back.
- **less:** displays the contents of the file page by page, so that you can go back in time.

Data processing commands (extract, sort and filter)

cat list | more

- **What happens?:** This displays the contents of list one screen at a time.
- **How to navigate:**
 - Press Space to move to the next screen.
 - Press Enter to move one line down.
 - Press q to quit.

Data processing commands (extract, sort and filter)

cat list | less

- **What happens?:** This displays the contents of list in an interactive viewer with more advanced navigation options.
- **How to navigate:**
 - Use the arrow keys to scroll up and down.
 - Press Space to move one screen down.
 - Press b to move one screen up.
 - Press / to search within the file.
 - Press q to quit.

Data processing commands (extract, sort and filter)

Is there a command to apply filters to file lines?

Data processing commands (extract, sort and filter)

The **grep** command

Data processing commands (extract, sort and filter)

The grep command is used to search for lines in files containing an expression.

\$ grep [options] expression files

Some interesting options:

- **-v** performs the reverse search
- **-c** returns only the number of lines found
- **-i** for not differentiating between upper and lower case
- **-n** to indicate line numbers
- **-E** for regular expressions

Data processing commands (extract, sort and filter)

Example of use 01 :

If a system administrator wants to search the **list file for** the expression **system** by displaying **the line number**.

```
grep -n system list
```

```
65:drwxr-xr-x 1 meskaldji meskaldji    512 Nov 10 21:48 system
```

```
66:-rw-r--r-- 1 meskaldji meskaldji 1528646 Nov 10 21:50 system.rar
```

```
67:-rw-r--r-- 1 meskaldji meskaldji 1876096 Nov 10 21:49 system.zip
```

Data processing commands (extract, sort and filter)

Example of use 02:

grep -En ^d list (searching for lines starting with d)

```
4:drwxr-xr-x 1 meskaldji meskaldji 512 Mar 17 2024 Bureau
5:drwxr-xr-x 1 meskaldji meskaldji 512 Mar  4 2024 Desk
7:drwxr-xr-x 1 meskaldji meskaldji 512 Nov  7 2023 Mydesktop
10:drwxr-xr-x 1 meskaldji meskaldji 512 Nov 10 22:01 bin
14:drwxr-xr-x 1 meskaldji meskaldji 512 Jun  2 14:04 d1
38:drwxr-xr-x 1 meskaldji meskaldji 512 Oct 26 2023 folder8
39:drwxr-xr-x 1 meskaldji meskaldji 512 Jun 26 23:18 interro
40:drwxr-xr-x 1 meskaldji meskaldji 512 Oct 27 11:44 karton
56:drwxr-xr-x 1 meskaldji meskaldji 512 Oct 28 2023 p2
57:drwxr-xr-x 1 meskaldji meskaldji 512 Oct 28 2023 p3
58:drwxr-xr-x 1 meskaldji meskaldji 512 Oct 28 2023 p4
60:drwxr-xr-x 1 meskaldji meskaldji 512 Oct 27 11:55 python
63:drw-r-xr-x 1 meskaldji meskaldji 512 Nov 23 2023 reptest
64:drwxr-xr-x 1 meskaldji meskaldji 512 Jun 26 23:05 seancell
65:drwxr-xr-x 1 meskaldji meskaldji 512 Nov 10 21:48 system
72:drwxr-xr-x 1 meskaldji meskaldji 512 Oct 27 11:41 testo
75:drwxr-xr-x 1 meskaldji meskaldji 512 Oct 27 11:43 unix
```

Data processing commands (extract, sort and filter)

Is there a command to apply filters to file columns?

Data processing commands (extract, sort and filter)

The **Cut** command

Data processing commands (extract, sort and filter)

The **cut** command is used to select **columns** and **fields** in a file.

- **\$ cut -cColumns [files]**

A column is the position of a character in the line. The first character is column 1, the second column 2, etc.

- **\$ cut -dCharacter -fChamps [files]**

Fields are delimited by tabulation by default; the -d parameter is used to specify a different delimiter.

Data processing commands (extract, sort and filter)

Example of use 01 :

If a system administrator wants to select the columns (1-4,11,13-) in the **file named «fichier»**

```
lcpig1b1@ubuntu-virtual-machine:~$ cat fichier
drwxr-xr-x 5 lcpig1b1 g1      4096 oct. 21 14:31 Bureau
-rwxr-xr-x 1 lcpig1b1 g1    6156288 oct. 7 22:21 core
drwxr-xr-x 2 lcpig1b1 g1      4096 oct. 4 12:10 Documents
-rwxr-xr-x 1 lcpig1b1 g1      8980 oct. 1 15:56 examples.desktop
-rw-rw---- 1 lcpig1b1 g1        0 oct. 24 09:07 Fich
drwxr-xr-x 2 lcpig1b1 g1      4096 oct. 2 11:51 Vidéos
lcpig1b1@ubuntu-virtual-machine:~$ cut -c1-4,11,13- fichier
drwx lcpig1b1 g1      4096 oct. 21 14:31 Bureau
-rwx lcpig1b1 g1    6156288 oct. 7 22:21 core
drwx lcpig1b1 g1      4096 oct. 4 12:10 Documents
-rwx lcpig1b1 g1      8980 oct. 1 15:56 examples.desktop
-rw- lcpig1b1 g1        0 oct. 24 09:07 Fich
drwx lcpig1b1 g1      4096 oct. 2 11:51 Vidéos
lcpig1b1@ubuntu-virtual-machine:~$
```


Data processing commands (extract, sort and filter)

Example of use 02 :

If a system administrator wants to select fields (1,3,4) from **the file named «fichier»**

```
lcpig1b1@ubuntu-virtual-machine:~$ cat fichier
drwxr-xr-x 5 lcpig1b1 g1      4096 oct. 21 14:31 Bureau
-rwxr-xr-x 1 lcpig1b1 g1    6156288 oct. 7 22:21 core
drwxr-xr-x 2 lcpig1b1 g1      4096 oct. 4 12:10 Documents
-rwxr-xr-x 1 lcpig1b1 g1      8980 oct. 1 15:56 examples.desktop
-rw-rw---- 1 lcpig1b1 g1         0 oct. 24 09:07 Fich
drwxr-xr-x 2 lcpig1b1 g1      4096 oct. 2 11:51 Vidéos
lcpig1b1@ubuntu-virtual-machine:~$ cut -d" " -f1,3,4 fichier
drwxr-xr-x lcpig1b1 g1
-rwxr-xr-x lcpig1b1 g1
drwxr-xr-x lcpig1b1 g1
-rwxr-xr-x lcpig1b1 g1
-rw-rw---- lcpig1b1 g1
drwxr-xr-x lcpig1b1 g1
lcpig1b1@ubuntu-virtual-machine:~$
```


Data processing commands (extract, sort and filter)

Is there a command for sorting?

Data processing commands (extract, sort and filter)

The sort command

Data processing commands (extract, sort and filter)

The **sort** command sorts the lines of a file in ascending order by default.

\$ sort [options] [files]

Some interesting options:

- **-k** to specify the column number used for sorting
- **-o** to write the result to a file
- **-r** sort in reverse order (sorting in descending order)
- **-n** numeric sorting, ideal for columns of numbers

Data processing commands (extract, sort and filter)

Example of use:

```
user1@PC:~$ cat fichier
-rw-rw-r-- 1 user1 user1      76473 11:33 présentation.pptx
-rw-rw-r-- 1 user1 user1 75075316 21:29 linux64-8.0.3_en-US.tar.xz
-rw-rw-r-- 1 user1 user1 105631957 14:24 linux-en.tar.gz
-rw-rw-r-- 1 user1 user1 130956158 11:17 112.mp4
user1@PC:~$ sort -n -r -k 5 fichier
-rw-rw-r-- 1 user1 user1 130956158 11:17 112.mp4
-rw-rw-r-- 1 user1 user1 105631957 14:24 linux-en.tar.gz
-rw-rw-r-- 1 user1 user1 75075316 21:29 linux64-8.0.3_en-US.tar.xz
-rw-rw-r-- 1 user1 user1 76473 11:33 présentation.pptx
user1@PC:~$
```

Column number 05

Data processing commands (extract, sort and filter)

Is there a command to delete duplicates?

Data processing commands (extract, sort and filter)

The **uniq** command

Data processing commands (extract, sort and filter)

The **uniq** command removes duplicates from **sorted files**, i.e. it only finds successive lines that are identical.

● **\$ uniq [options] [source [destination]]**

Example:

```
user1@PC:~$ cat fichier
mohammed
ali
ali
ali
bob
bob
user1@PC:~$ uniq fichier
mohammed
ali
bob
user1@PC:~$ □
```

Data processing commands (extract, sort and filter)

Is there a command to replace one expression with another?

Data processing commands (extract, sort and filter)

The **sed** command

Data processing commands (extract, sort and filter)

The **sed** command is used to replace a **expression** by **another**.

1. Basic Text Substitution:

Replace the first occurrence of "oldword" with "newword" on each line in a file:

```
sed 's/oldword/newword/' filename
```

Example of use:

```
sed 's/cat/dog/' animals.txt
```

Data processing commands (extract, sort and filter)

2. Global Substitution

Replace **all occurrences** of "oldword" with "newword" on each line:

`sed 's/oldword/newword/g' filename`

Example:

`sed 's/cat/dog/g' animals.txt`

- Replaces **all** occurrences of "cat" with "dog".

Data processing commands (extract, sort and filter)

3. In-Place Editing (Modify the File Directly)

Replace text directly in the file (overwrites the file):

```
sed -i 's/oldword/newword/g' filename
```

Example:

```
sed -i 's/cat/dog/g' animals.txt
```

- Replaces all occurrences of "cat" with "dog" in animals.txt.

Data processing commands (extract, sort and filter)

Is there a command to count words?

Data processing commands (extract, sort and filter)

The WC command

Data processing commands (extract, sort and filter)

The **wc (word count)** command is used to count **lines, words, bytes** and **characters**.

```
$ wc [Options] [files]
```

Some interesting options:

- **-l** counts the number of lines
- **-c** counts the number of bytes
- **-w** counts the number of words
- **-m** counts the number of characters

Data processing commands (extract, sort and filter)

Example of use:

System administrator wants to compare two files
/etc/passwd and **/etc/shadow**

```
root@PC:~# wc /etc/passwd /etc/shadow
```

45	73	2597	/etc/passwd
45	45	1680	/etc/shadow
90	118	4277	total

root@PC:~#

lines words characters

Data processing commands (extract, sort and filter)

**Is there command for
file processing?**

Data processing commands (extract, sort and filter)

The **paste, split, ...**
commands

Data processing commands (extract, sort and filter)

- Line-to-line join

\$ paste [options] [files]

it concatenates the lines of files with each other, line1 of file1 with line1 of file2, etc.

the -d option is used to specify the **delimiter**

- Cutting a file into pieces

\$ split [options] [file [prefix]]

It cuts a large file into several pieces.

Data processing commands (extract, sort and filter)

Example of use 01 :

System administrator wants to make a **line-by-line** join of files **file1**, **file2** and **file3**

```
user1@PC:~$ cat fichier1
```

```
-rwxr-sr-x  
-rw-rw-r--  
-rwxr-sr-x
```

```
user1@PC:~$ cat fichier2
```

```
user1 user1 0  
user1 user1 158  
user1 user1 8560
```

```
user1@PC:~$ cat fichier3
```

```
19:55 fichier2  
19:58 fichier3  
10:50 fichier.txt
```

```
user1@PC:~$ paste -d" " fichier1 fichier2 fichier3
```

```
-rwxr-sr-x user1 user1 0 19:55 fichier2  
-rw-rw-r-- user1 user1 158 19:58 fichier3  
-rwxr-sr-x user1 user1 8560 10:50 fichier.txt
```

```
user1@PC:~$
```

Data processing commands (extract, sort and filter)

Example of use 02 :

System administrator wants to cut a large file **file** in **10** pieces each in one file

```
user1@PC:~$ split -n 10 fichier fich
user1@PC:~$ ls -l fich*
```

-rw-rw-r--	1	user1	user1	364275212	16:44	fichaa
-rw-rw-r--	1	user1	user1	364275212	16:44	fichab
-rw-rw-r--	1	user1	user1	364275212	16:44	fichac
-rw-rw-r--	1	user1	user1	364275212	16:45	fichad
-rw-rw-r--	1	user1	user1	364275212	16:45	fichae
-rw-rw-r--	1	user1	user1	364275212	16:45	fichaf
-rw-rw-r--	1	user1	user1	364275212	16:45	fichag
-rw-rw-r--	1	user1	user1	364275212	16:45	fichah
-rw-rw-r--	1	user1	user1	364275212	16:46	fichai
-rw-rw-r--	1	user1	user1	364275218	16:46	fichaj
-rw-rw-r--	1	user1	user1	3642752126	16:38	fichier

```
user1@PC:~$
```

Data processing commands (extract, sort and filter)

Other commands
join, diff, cmp

Data processing commands (extract, sort and filter)

- **Joint on common fields**

\$ join [options] file1 file2

It performs a join of two files based on a common field.

- **File comparison**

\$ diff [options] file1 file2

\$ cmp [options] file1 file2

They can be used to compare the contents of two files

etc ●●●