Introduction to Linux and Commands

Program

Goals

    Presentation of the Linux OS

    The basis for a good starting point with Linux

Applications

    Knowing the basic Linux command

    Files manipulations (sort, cut, wc, tr)

    Sorting and filtering data (grep / sed / awk)

    Use of bioinformatics software in command-line

Introduction to Linux

- 1970 : UNIX operating system created

- Numerous forks : Ultrix, AIX, SunOS & Linux (1991)

- **Free system, solid, stable and wide array of machines**

- **Multi-tasking/multi-user system**

    One task or process = software

   Multi-tasking : several process can be run in the same time

    Multi-user system : several users can use the system in the same time

    Tasks are protected, some can communicate

- Files within a tree representation of files and folders

Introduction to Linux

▪ The kernel manage the basic system tasks :

   System init

   Resources and processes management

   Files managements

   Inputs/Outputs managements

▪ **User communicates with the kernel through the Shell command-lines.**
Shells are also programming language
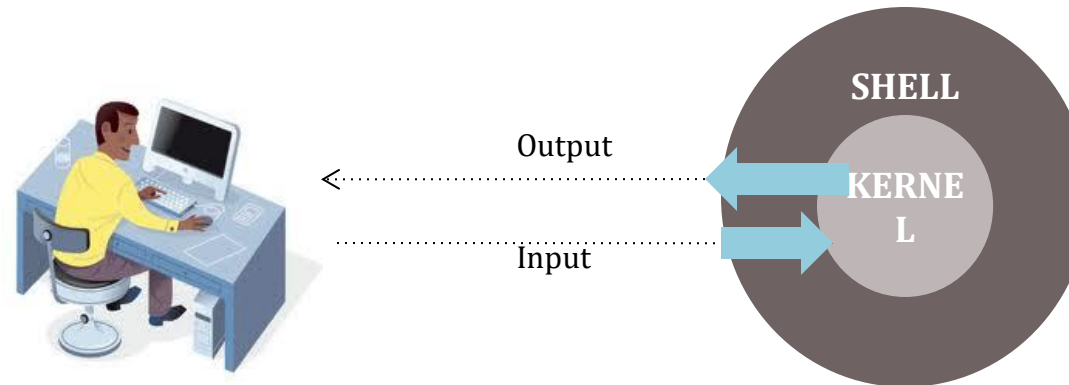
Several Linux Distribution

A distribution = Kernel with softwares
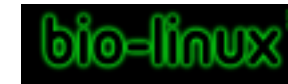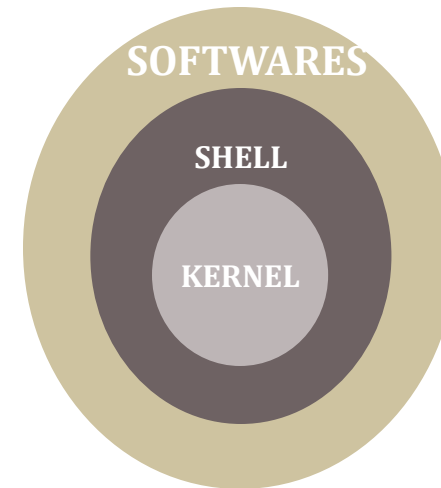
Several Linux Distribution

How to find out my linux distribution and version number?

```
cat /etc/issue        Give the distribution name
uname -a       Give the Kernel version
```

Why using Linux ?

▪ Numerous small programs/commands in the "Shell" very powerful

▪ Easy to develop workflow to link programs/commands between them

▪ A lot of free bioinformatics programs available

▪ Not necessary to waste power resources to manage graphical windows

▪ 90% of Servers are on Linux

Negative point : friendliness ? No... Graphical interfaces, high-level of user-experience.

▪ Interpreter for command-lines and programming language

▪ Interface between user and kernel/system on behalf of command-lines

▪ Various shells : *sh* (*Bourne shell), bash (Bourne again shell), ...*
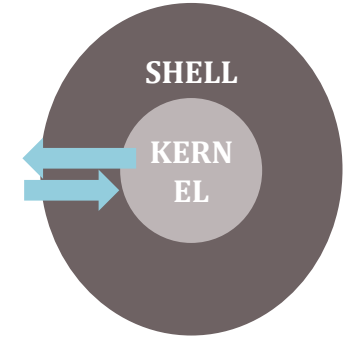
```
echo $SHELL          Give the Default Shell
```

Prompt `[tranchant@node6 data]$ █`

user name

Server name

current directory

```
command [ -options ] [ arguments or target]
```

Ex: *pwd* : print name of current directory

  ls : list all files in a directory

A few commands

Help

man ls	To get help (Manual)
ls --help	

Basics

|	
pwd	Display absolute path
ls	List all files/directories [only show names]
ls –l	Long listing: show other information too
who	Connected users list
whoami	Display the full name of current user
uname	Version and Name of the system
exit	Exit the shell session

Practical 1

Open a terminal :

      - Check with pwd command the name of your working directory.
      - What is the linux distibution on server?
      - What is the shell?
   - Display the ls command help

File tree

Main Directories

```
                              /
        ┌────┬─────┬──────┬──────┬──────┐
       bin   etc   lib   sbin   usr   home
```
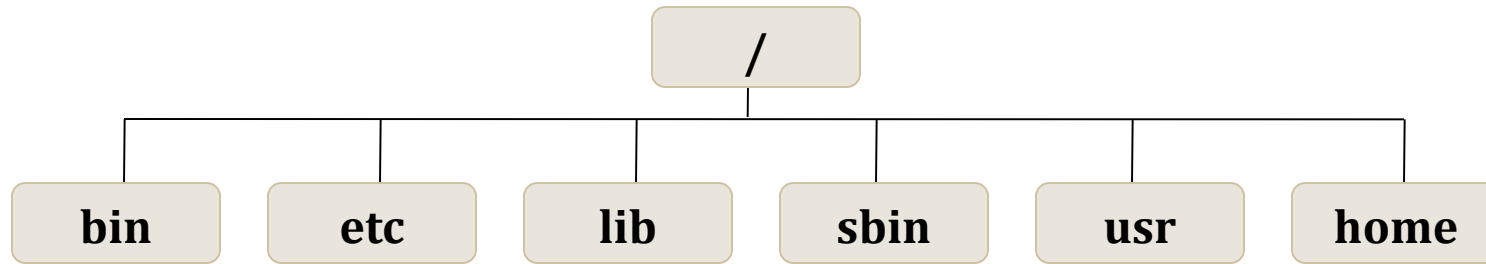
/	Root directory (slash)
/bin	**Main commands, shell, programs**
/etc	Configuration files for the system
/lib	Programming Libraries
/mnt	Mount point
/usr, /opt	**Applications and user libraries**
/usr/bin	**Other commands**
/var	Logfiles
/tmp	Temporary files
/home	**User directory (one per user, name = login)**

File tree
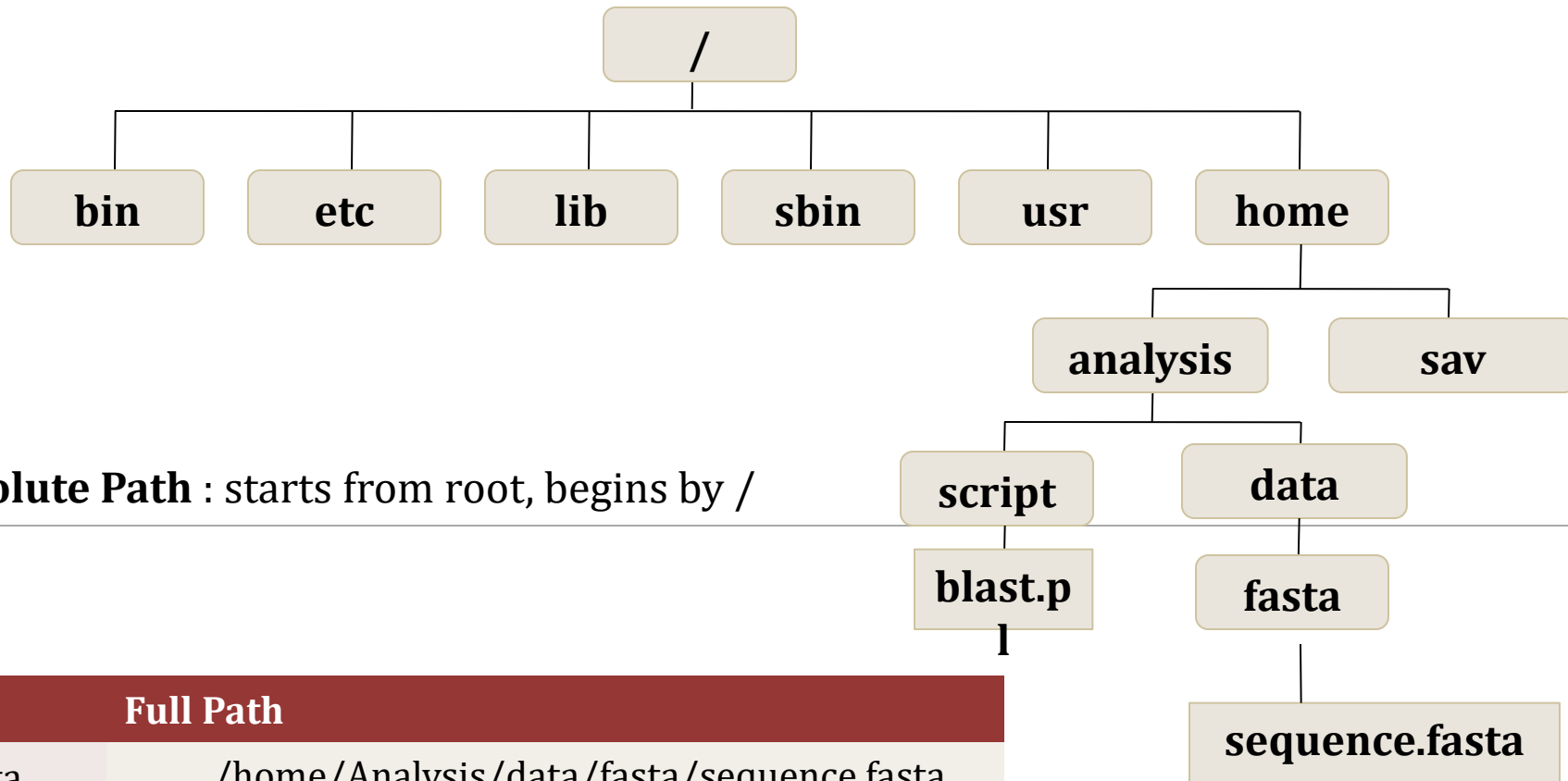
Relative path : give the position of a file/folder based on the current directory

Example :

Current directory	Relative path
data	fasta/sequence.fasta

File tree

Relative path : give the position of a file/folder based on the current directory

Example :
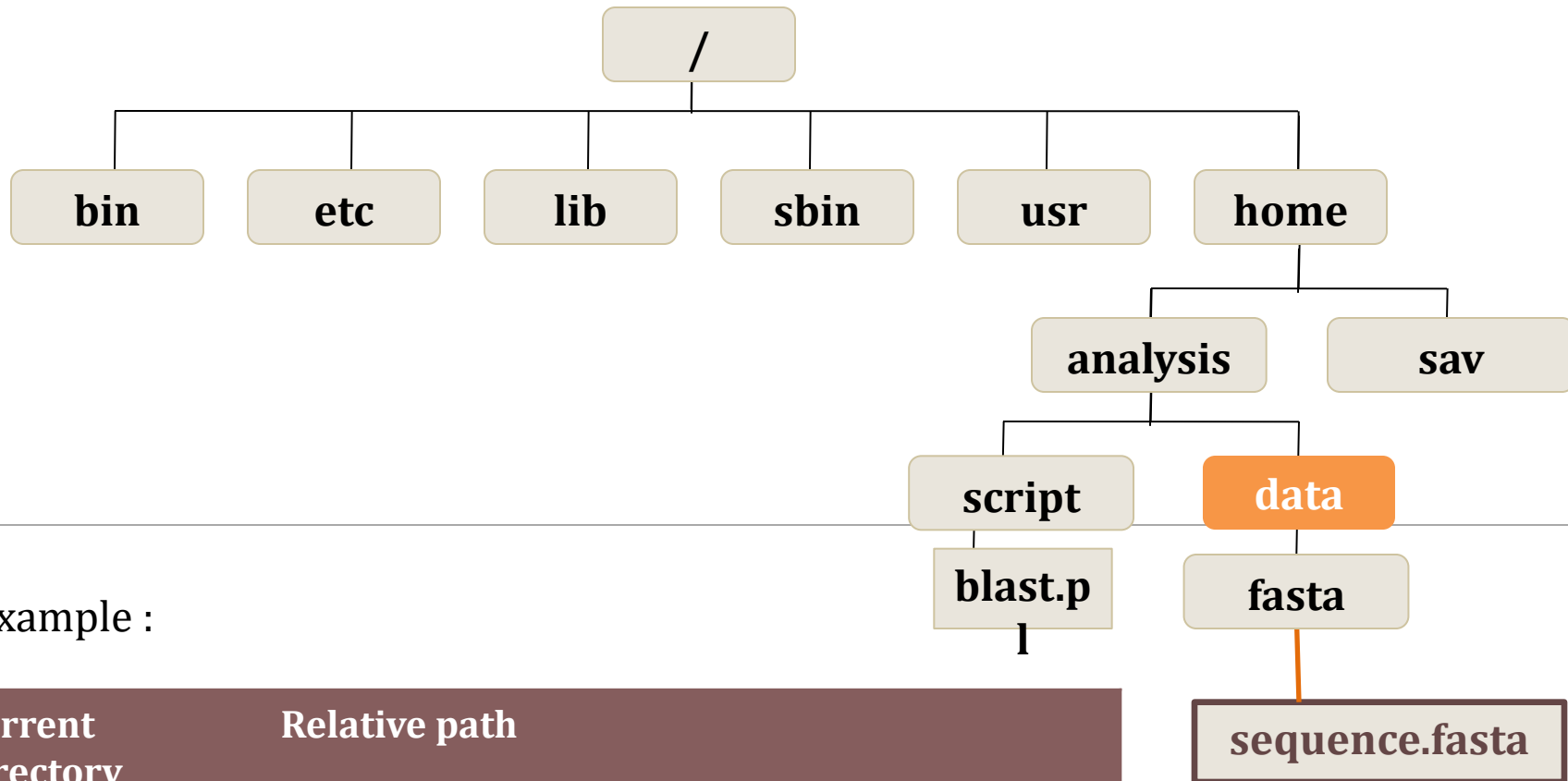
Current directory	Relative path
script	../data/fasta/sequence.fasta
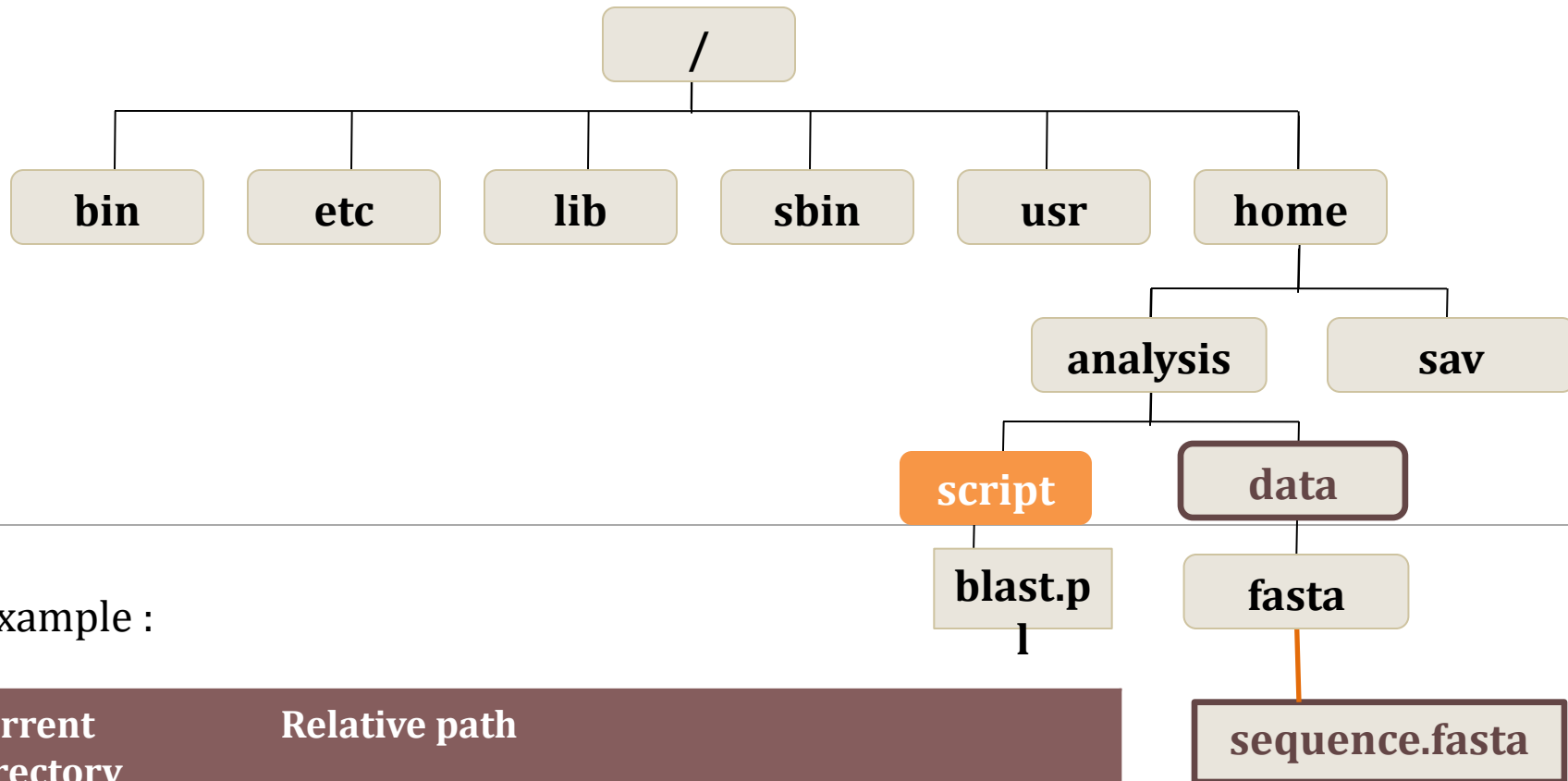
Directory management : some commands

pwd	Name of current Directory
ls rep_name	Display the list of files in the folder
cd rep_name	Change working directory
mkdir rep_name	Create the directory
rmdir rep_name	Remove the directory
rm –r rep_name	Remove the directory and all the files
cp source target	Copy source to target
mv old_name new_name	Change the file name

Use with caution

File management : some commands (1)

touch file_name   Create a file

nano file_name.   Create a file (open in nanoeditor)

cp source target   Copy source to target

mv old_name new_name   Change the file name

cat filename   Print the entire contents of a file to screen

less filename   Display file content one screen length at a time. (for big file)

- Certain commands are then used to go back and forward through the file:

- space bar: display the next page

- b: display the previous page

- enter/return: display the next line

- k: display the previous line

- q: quit the viewer

Practical 2 :

Navigating and creating

1. Create a directory in your home folder where you will conduct your practice exercises « practice_directory »

2. Navigate into that directory using the cd command.

3. Inside this directory, create several files « file1.txt » « file2.txt » and directories « dir1 » « dir2 » «dir1/subdir » using commands like touch for files and mkdir for directories.

4. Rename some of the files and directories using the mv command. file1.txt to renamed_file.txt and dir1 renamed_dir1

5. Copy file2.txt in dir2

6. Copy dir2 to dir2.new.

7. Delete dir2.new directory (including its contents) with the rm -r command.

8. Return to your home directory with the cd command.

9. Move file2.txt (in dir2) to dir1/subdir.

10. Move file2.txt (in dir1/subdir) in dir1 rename it simultaneously file3.txt .

Nano

1. Create a file called file5.txt in the unix folder by typing « nano file5.txt »

2. Write in here 'this is the contents of file5.txt'.

3. Save the file by pressing ctrl-o (press control and then press o).

4. confirm the file name and you can press return/enter to confirm this.

5. Exit nano by pressing ctrl-x (and hitting return/enter).

File management : some commands (2)

more filename.   Display file content one screen length at a time. (for big file)

      Certain commands are then used to go back and forward through the file:

- enter/return: display the next screen

- q: quit the viewer

head -n x filename.   Display the x first line of a file

tail -n x filename.   Display the x last lines of a file

Download management

Curl link   transfer data from or to a server

      Options:

- -O : save the Downloaded with the same name (curl -O http://example.com/file.txt)

- -o : Save the Downloaded File with a New Name (curl -o http://example.com/file.txt)

wget link   transfer data from or to a server

Compress/decompress file(s)

zip / gzip filename   Compress/ decompress file(s) and directorie(s (.zip format)

gzip / gunzip filename. Compress/ decompress file(s) (.zip format)

wc grep

wc   transfer data from or to a server

Grep   transfer data from or to a server

redirect and pipe

Command > file   Redirect output in a newly created file (will erase existing file with this name)

Command >> file Redirect output to a new file (creation) or at the end of an already existing file with this name (append to file)

Command < file   Redirect the Input from a file

Command1 | Command2   Redirect the Standard Output from one command to the Standard Input of another without using a file

Generic characters

Some Special Characters

?   **Every single character**
*   **Whatever is the character chain**
[ensemble]   **All characters in ensemble**
[!ensemble]   **All not in ensemble**

Exemple :

programme.c programme.log   programme.o
programmes.pl   fichier.contig

```
   ls programme.c      #programme.c
ls programme.?     #programme.c   programme.o
ls *.c*            #programme.c   fichier.contig
ls programme.[co] #programme.c programme.o
```

Let's practice with some real data

1.Create a folder « « practice_directory2 »

2.Navigate into that directory

3.Download data: Use the following command:

curl -L -o Data.zip https://tinyurl.com/KrispLinuxWorkshopData2020

4. Unzip « data.zip »

5. Delete zip file

6. Create a new directory name 'Data'

7. Move all « .fasta » files to that new directory

8. View the contents of some files (sequence1.fasta - sequence10.fasta)

9. View the first few lines of some files (sequence1.fasta - sequence10.fasta)

10. Merge sequence files

11. Download from thises two files using wget

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR256/023/SRR25627323/SRR25627323_1.fastq.gz

ftp://ftp.sra.ebi.ac.uk/vol1/fastq/SRR256/023/SRR25627323/SRR25627323_2.fastq.gz

12.Uncompress SRR25627323_1.fastq.gz

13 . Count the number of read in SRR25627323_1.fastq.gz using « grep » or « grep + wc »