

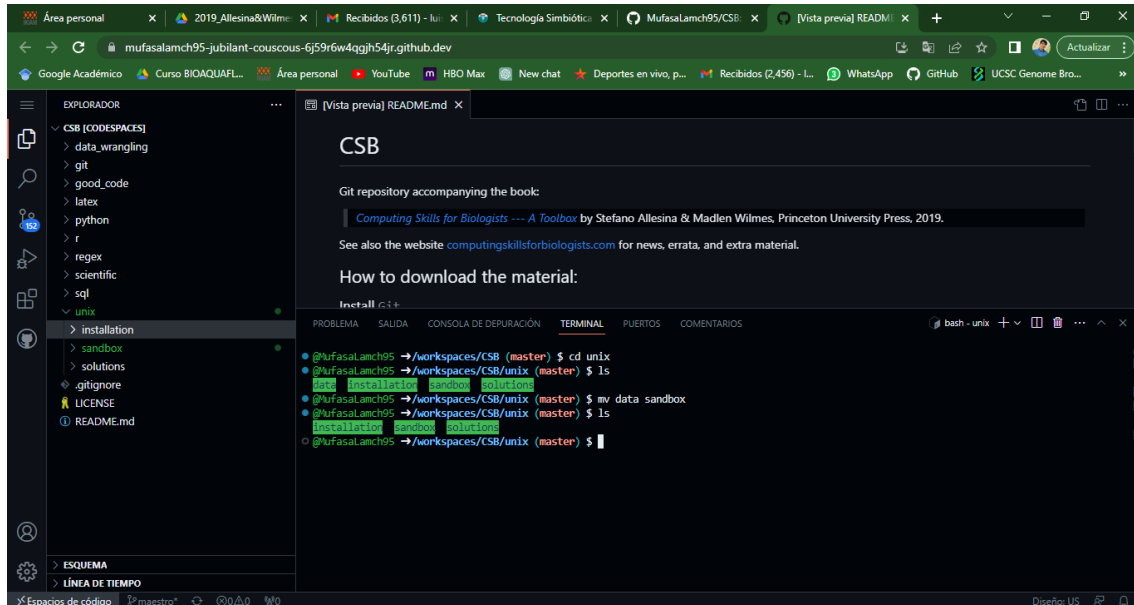
Nombre: Luis Antonio Maldonado

Bioinformática

TC3_ForLoop

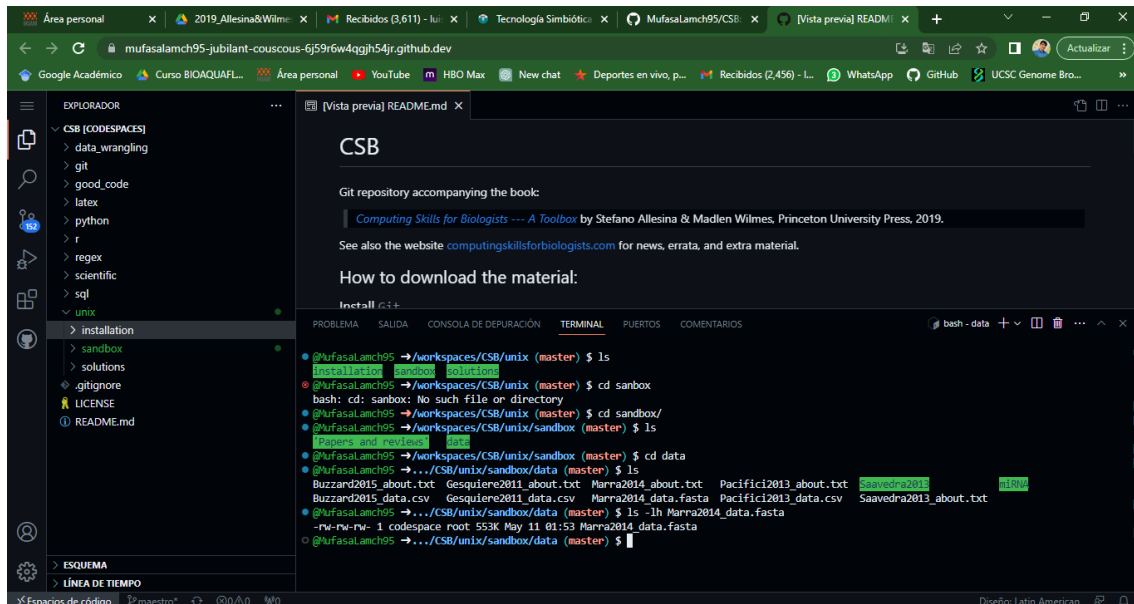
Eje 1.10.1

1. Cambiar directorio a sandbox.



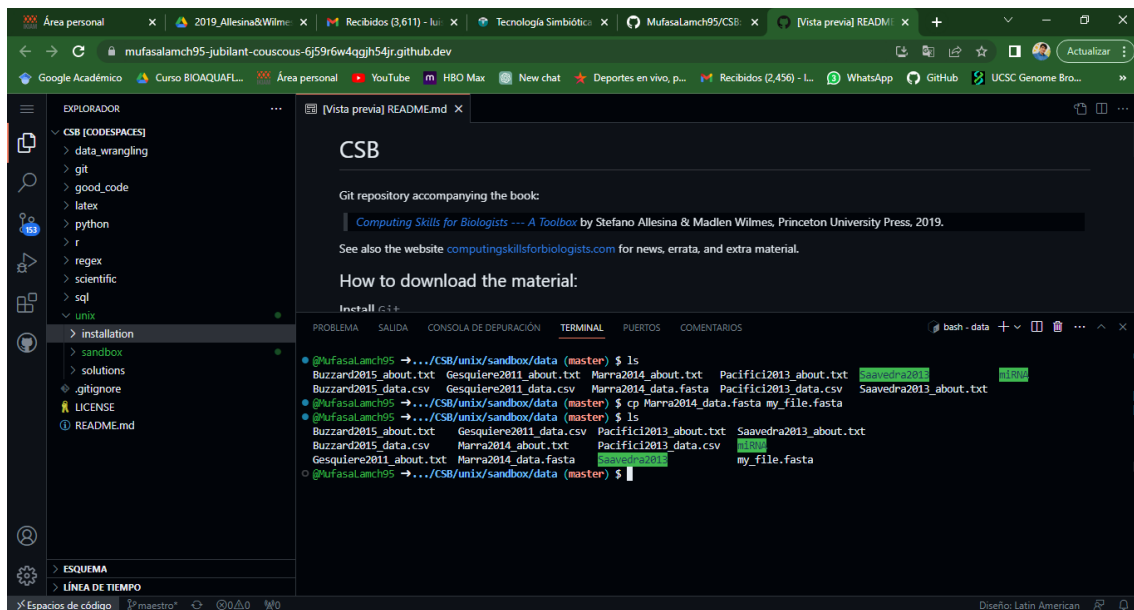
```
@%ufasalanch95 →/workspaces/CSB (master) $ cd unix
@%ufasalanch95 →/workspaces/CSB/unix (master) $ ls
data  installation  sandbox
@%ufasalanch95 →/workspaces/CSB/unix (master) $ mv data sandbox
@%ufasalanch95 →/workspaces/CSB/unix (master) $ ls
installation  papers  sandbox
@%ufasalanch95 →/workspaces/CSB/unix (master) $
```

2.



```
@%ufasalanch95 →/workspaces/CSB/unix (master) $ ls
@%ufasalanch95 →/workspaces/CSB/unix (master) $ cd sandbox
bash: cd: sandbox: No such file or directory
@%ufasalanch95 →/workspaces/CSB/unix (master) $ cd sandbox/
@%ufasalanch95 →/workspaces/CSB/unix/sandbox (master) $ ls
papers  reviews  txt
@%ufasalanch95 →/workspaces/CSB/unix/sandbox (master) $ cd data
@%ufasalanch95 →.../CSB/unix/sandbox/data (master) $ ls
Buzzard2015_about.txt  Gesquiere2011_about.txt  Marra2014_about.txt  Pacific12013_about.txt  Saavedra2013_about.txt
Buzzard2015_data.csv  Gesquiere2011_data.csv  Marra2014_data.fasta  Pacific12013_data.csv
@%ufasalanch95 →.../CSB/unix/sandbox/data (master) $ ls -lh Marra2014_data.fasta
-rw-rw-rw- 1 codespace root 553K May 11 01:53 Marra2014_data.fasta
@%ufasalanch95 →.../CSB/unix/sandbox/data (master) $
```

3.



CSB

Git repository accompanying the book:

[Computing Skills for Biologists --- A Toolbox](#) by Stefano Allesina & Madlen Wilmes, Princeton University Press, 2019.

See also the website [computingskillsforbiologists.com](#) for news, errata, and extra material.

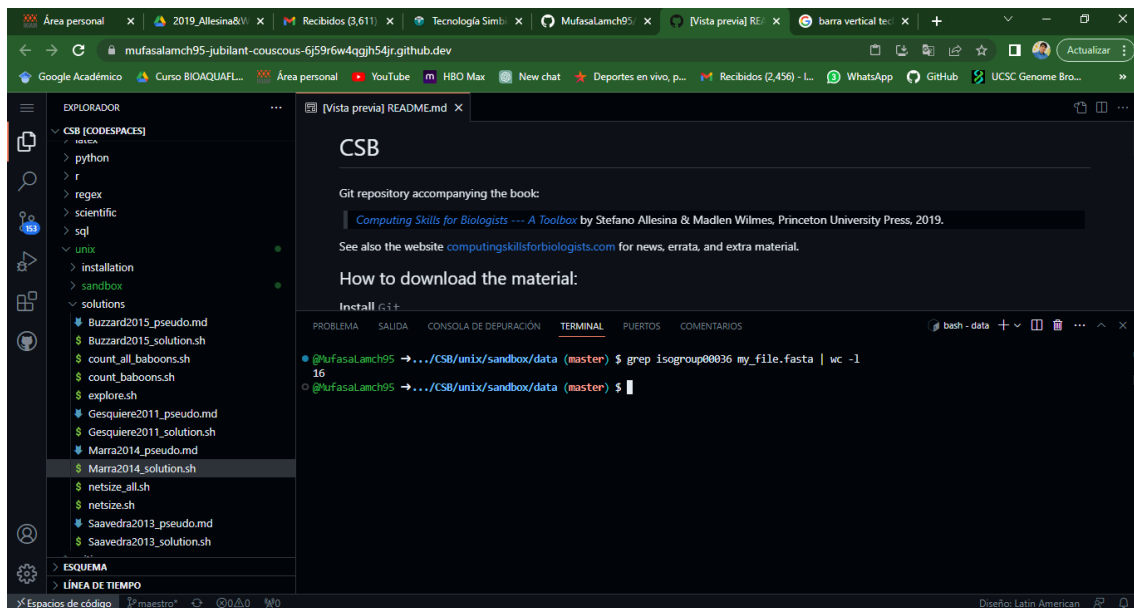
How to download the material:

Install `git`

PROBLEMA SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS COMENTARIOS

```
@%ufasalanch95 → .../CSB/unix/sandbox/data (master) $ ls
Buzzard2015_about.txt  Gesquiere2011_about.txt  Marra2014_about.txt  Pacifici2013_about.txt  Saavedra2013_about.txt
Buzzard2015_data.csv  Gesquiere2011_data.csv  Marra2014_data.fasta  Pacifici2013_data.csv  Saavedra2013_data.csv
@%ufasalanch95 → .../CSB/unix/sandbox/data (master) $ cp Marra2014_data.fasta my_file.fasta
@%ufasalanch95 → .../CSB/unix/sandbox/data (master) $ ls
Buzzard2015_about.txt  Gesquiere2011_data.csv  Pacifici2013_about.txt  Saavedra2013_about.txt
Buzzard2015_data.csv  Marra2014_about.txt  Pacifici2013_data.csv  my_file.fasta
Gesquiere2011_about.txt  Marra2014_data.fasta  my_file.fasta
@%ufasalanch95 → .../CSB/unix/sandbox/data (master) $
```

4.



CSB

Git repository accompanying the book:

[Computing Skills for Biologists --- A Toolbox](#) by Stefano Allesina & Madlen Wilmes, Princeton University Press, 2019.

See also the website [computingskillsforbiologists.com](#) for news, errata, and extra material.

How to download the material:

Install `git`

PROBLEMA SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS COMENTARIOS

```
@%ufasalanch95 → .../CSB/unix/sandbox/data (master) $ grep isogroup00036 my_file.fasta | wc -l
16
@%ufasalanch95 → .../CSB/unix/sandbox/data (master) $
```

Área personal x 2019_Allesina&V x Recibidos (3,611) x Tecnología Simb x MufasaLamch95 x [Vista previa] RE x barra vertical tec x +

← → ↺ mufasalamch95-jubilant-couscous-6j59r6w4qjih54jr.github.io

Google Académico Curso BIOQAFL... Área personal YouTube HBO Max New chat Deportes en vivo, p... Recibidos (2,456) - L... WhatsApp GitHub UCSC Genome Bro... »

EXPLORADOR

- CSB [CODESPACES]
 - data_wrangling
 - git
 - good_code
 - latex
 - python
 - r
 - regex
 - scientific
 - sql
 - unix
 - installation
 - sandbox
 - solutions
 - .gitignore
 - LICENSE
 - README.md

[Vista previa] README.md x

CSB

Git repository accompanying the book:

[Computing Skills for Biologists --- A Toolbox](#) by Stefano Allesina & Madlen Wilmes, Princeton University Press, 2019.

See also the website [computingskillsforbiologists.com](#) for news, errata, and extra material.

How to download the material:

Install [Git](#).

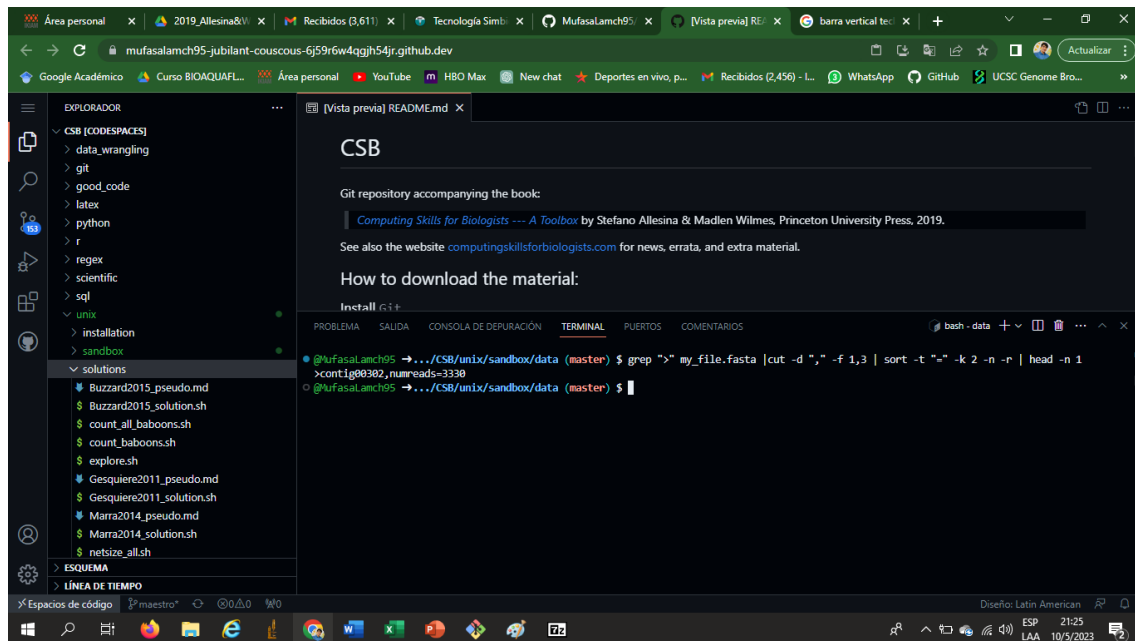
PROBLEMA SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS COMENTARIOS

```
bash - data + ▢ ▢ ... ^ x
```

```
@MufasaLamch95 →.../CSB/unix/sandbox/data (master) $ grep ">" my_file.fasta | cut -d "," -f 4 | sort | uniq | wc -l
43
@MufasaLamch95 →.../CSB/unix/sandbox/data (master) $
```

~ Espacios de código maestro* ↺ ↻ ↺ ↻ ↺ ↻ W/O

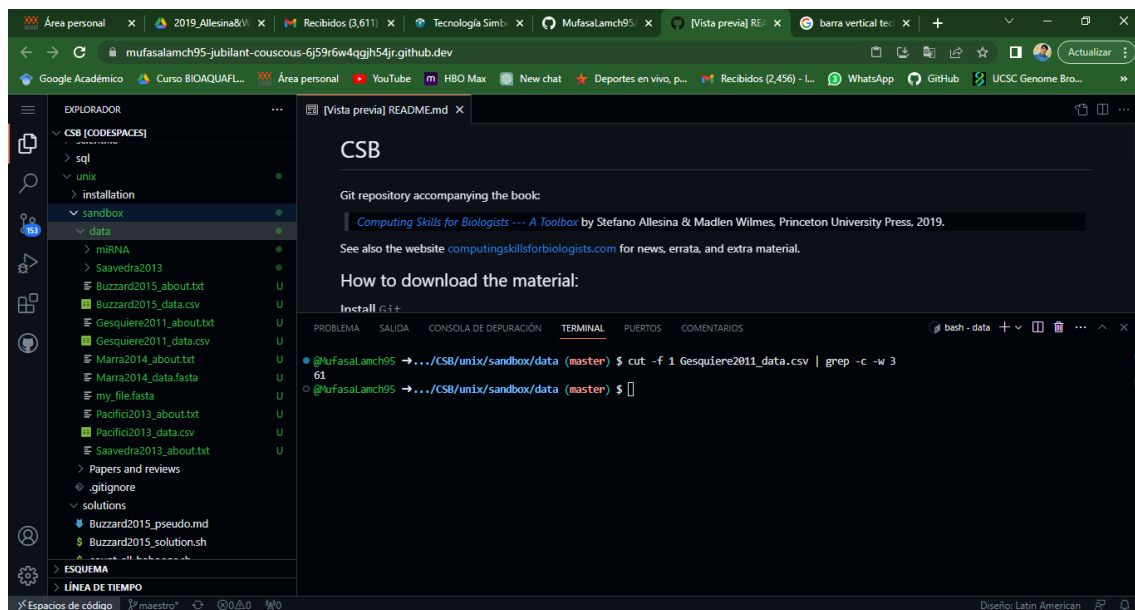
Diseño: Latin American

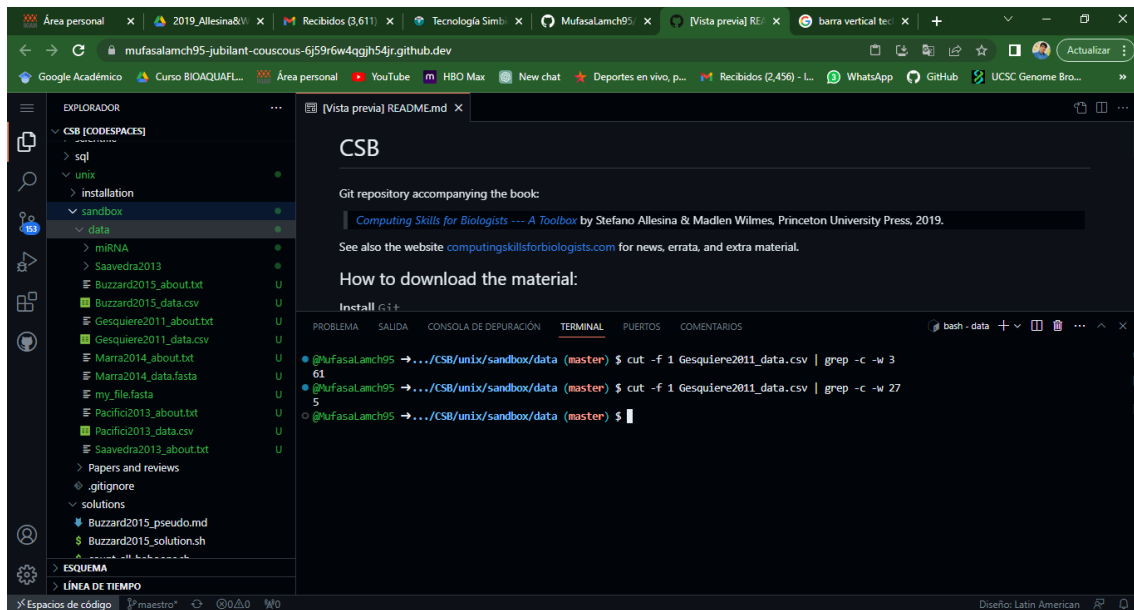


Ejercicio 7

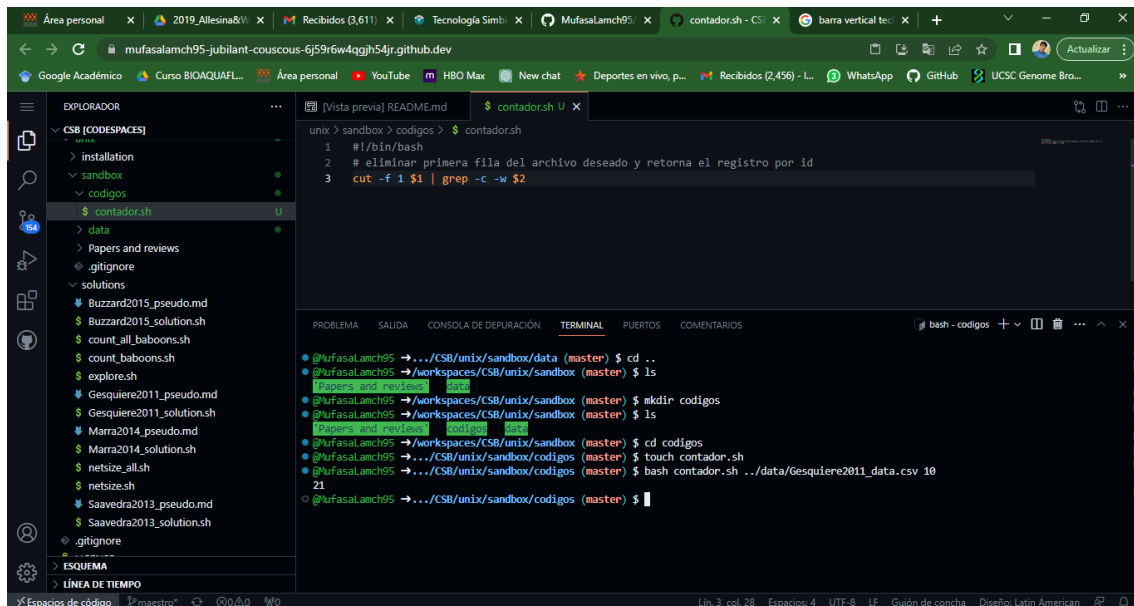
Literat 1.10.2

1.





2.



3.

```
1 #salta la primera línea del archivo csv, ordena y busca los identificadores unicos
2 Identificadores="tail -n +2 ../data/Gesquiere2011_data.csv | cut -f 1 | sort -n | uniq"
3
4 #repetición de ids
5 for id in $Identificadores
6 do
7     contadores="bash contador.sh ../data/Gesquiere2011_data.csv $id"
8     echo "Identificador:" $id "conteos:" $contadores
9 done
```

```
@ufasalanch95 → .../CSB/unix/sandbox/codigos (master) $ bash contador_veces.sh
Identificador: 1 conteos: 10
Identificador: 2 conteos: 2
Identificador: 3 conteos: 61
Identificador: 4 conteos: 46
Identificador: 5 conteos: 28
Identificador: 6 conteos: 7
Identificador: 7 conteos: 5
Identificador: 8 conteos: 17
Identificador: 9 conteos: 4
Identificador: 10 conteos: 21
Identificador: 11 conteos: 26
Identificador: 12 conteos: 23
Identificador: 13 conteos: 16
Identificador: 14 conteos: 1
Identificador: 15 conteos: 40
Identificador: 16 conteos: 31
```

Ejercicios 1.10.3

1.

```
1 #!/bin/bash
2
3 echo "nombre del archivo:"
4 echo $1
5 echo "Numero de filas:"
6 #imprime la cantidad total de filas
7 cat $1 | wc -l
8 echo "Numero de columnas:"
9 #saltar la fila 1 que es la cabecera e imprime las columnas
10 head -n 1 $1 | tr -d ' ' | tr -d '\n' | wc -c
11
```

```
@ufasalanch95 → .../CSB/unix/sandbox/codigos (master) $ touch plantas_filas_columnas.sh
@ufasalanch95 → .../CSB/unix/sandbox/codigos (master) $ bash plantas_filas_columnas.sh ../data/Saavedra2013/n1.txt
nombre del archivo:
../data/Saavedra2013/n1.txt
Numero de filas:
97
Numero de columnas:
80
@ufasalanch95 → .../CSB/unix/sandbox/codigos (master) $
```

2.

The screenshot shows a VS Code editor with a file explorer on the left and a terminal at the bottom. The file explorer shows a project structure with folders like 'CSB [CODESPACES]', 'sandbox', 'codigos', 'data', 'papers and reviews', 'gitignore', and 'solutions'. The 'codigos' folder is selected, showing files like 'contador_veces.sh', 'contador.sh', 'plantas_filas_columnas.sh', and 'todos_plantas_filas_columnas.sh'. The 'todos_plantas_filas_columnas.sh' file is open in the editor, showing a shell script that processes text files in the 'data' directory. The terminal shows the command 'bash todos_plantas_filas_columnas.sh' being executed, resulting in a list of file names and their dimensions (lines and columns).

```
1 #!/bin/bash
2
3 archivos=../data/Saavedra2013/*.txt
4 for f in $archivos
5 do
6     filas=$(cat $f | wc -l)
7     columnas=$(head -n 1 $f | tr -d ' ' | tr -d '\n' | wc -c)
8     echo $f $filas $columnas
9 done
```

PROBLEMA SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS COMENTARIOS

```
@%fufasaLamch95 →.../CSB/unix/sandbox/codigos (master) $ bash todos_plantas_filas_columnas.sh
../data/Saavedra2013/n1.txt 97 80
../data/Saavedra2013/n10.txt 14 20
../data/Saavedra2013/n11.txt 270 91
../data/Saavedra2013/n12.txt 7 72
../data/Saavedra2013/n13.txt 61 17
../data/Saavedra2013/n14.txt 35 15
../data/Saavedra2013/n15.txt 38 11
../data/Saavedra2013/n16.txt 118 24
../data/Saavedra2013/n17.txt 76 31
../data/Saavedra2013/n18.txt 13 14
../data/Saavedra2013/n19.txt 10 16
../data/Saavedra2013/n2.txt 62 41
../data/Saavedra2013/n21.txt 19 45
../data/Saavedra2013/n22.txt 19 36
../data/Saavedra2013/n23.txt 179 26
```

3.

The screenshot shows a VS Code editor with a file explorer on the left and a terminal at the bottom. The file explorer shows a project structure with folders like 'CSB [CODESPACES]', 'sandbox', 'codigos', 'data', 'papers and reviews', 'gitignore', and 'solutions'. The 'codigos' folder is selected, showing files like 'contador_veces.sh', 'contador.sh', 'plantas_filas_columnas.sh', and 'todos_plantas_filas_columnas.sh'. The 'todos_plantas_filas_columnas.sh' file is open in the editor, showing a shell script that processes text files in the 'data' directory. The terminal shows the command 'bash todos_plantas_filas_columnas.sh | sort -n -r -k 3 | head -n 1' being executed, resulting in the output 'n56.txt 110 207'.

```
1 #!/bin/bash
2
3 archivos=../data/Saavedra2013/*.txt
4 for f in $archivos
5 do
6     filas=$(cat $f | wc -l)
7     columnas=$(head -n 1 $f | tr -d ' ' | tr -d '\n' | wc -c)
8     echo $f $filas $columnas
9 done
```

PROBLEMA SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS COMENTARIOS

```
@%fufasaLamch95 →.../CSB/unix/sandbox/codigos (master) $ bash todos_plantas_filas_columnas.sh | sort -n -r -k 3 | head -n 1
../data/Saavedra2013/n56.txt 110 207
@%fufasaLamch95 →.../CSB/unix/sandbox/codigos (master) $
```

Ejercicio 1.10.4

1.

The screenshot shows a VS Code editor with a terminal window open. The terminal is running a shell script named `cervezalanch.sh` in a directory `unix/sandbox/codigos`. The script processes a CSV file `../data/Buzzard2015_data.csv` and outputs the following results:

```
unix > sandbox > codigos > $ cervezalanch.sh
4 #Busque el archivo seleccionado y la posición de la columna que se requiere
5 cut -d ',' -f $2 $1 | head -n 1
6 echo "valores distintos :"
7 #Busca los valores unicos y distintos de la columna seleccionada
8 cut -d ',' -f $2 $1 | tail -n +2 | sort | uniq | wc -l
9 echo "valor mínimo:"
10 #Busca el valor mínimo de la columna seleccionada
11 cut -d ',' -f $2 $1 | tail -n +2 | sort -n | head -n 1
12 echo "valor máximo:"
13 #Busca el valor máximo
14 cut -d ',' -f $2 $1 | tail -n +2 | sort -n | tail -n 1

PROBLEMA SALIDA CONSOLA DE DEPURACIÓN TERMINAL PUERTOS COMENTARIOS
@%ufasalanch95 →.../CSB/unix/sandbox/codigos (master) $ bash cervezalanch.sh ../data/Buzzard2015_data.csv 7
nombre columna :
biomass
valores distintos :
285
valor mínimo:
1.048466198
@%ufasalanch95 →.../CSB/unix/sandbox/codigos (master) $ bash cervezalanch.sh ../data/Buzzard2015_data.csv 7
nombre columna :
biomass
valores distintos :
285
valor mínimo:
1.048466198
valor máximo:
14897.29471
@%ufasalanch95 →.../CSB/unix/sandbox/codigos (master) $
```

The terminal output shows the results of running the script for two different column indices (7 and 7). The first run shows the column name 'biomass', the number of distinct values (285), the minimum value (1.048466198), and the maximum value (14897.29471).