

Respaldo Bioinformática CLASE4

Melina Vargas

1. Ejercicio 1.10.1

```
Melina@DESKTOP-C7U9BDL MINGW32 ~/Documents/GitHub/CLASE4 (main)
$ nano clase4_if.sh
```

```
MINGW32:/c/Users/user/Documents/GitHub/CLASE4
GNU nano 5.8 clase4_if.sh Modified
## Ejercicio 1.10.1 ##
## 1. Cambie de directorio a CSB/Unix/sandbox ##
echo "1. Cambie de directorio a CSB/Unix/sandbox"
cd |./../CSB-master/Unix/sandbox/
ls

## 2. Cuál es el tamaño del archivo Marra2014_data.fasta? ##
echo "2. Cuál es el tamaño del archivo Marra2014_data.fasta?"
du -h ../data/Marra2014_data.fasta

## 3. Cree una copia de Marra2014_data.sh en sandbox y asigne el nombre de my_file.fasta ##
echo "3. Cree una copia de Marra2014_data.sh en sandbox y asigne el nombre de my_f"
cp ../data/Marra2014_data.fasta my_file.fasta
ls

## 4. ¿Cuántos contigs se clasifican como isogrupo00036? ##
echo "4. ¿Cuántos contigs se clasifican como isogrupo00036?"
grep -c isogrupo00036 my_file.fasta

## 5. Reemplace el delimitador original de dos espacios con una coma ##
echo "5. Reemplace el delimitador original de dos espacios con una coma"
head -n 4 my_file.fasta
cat my_file.fasta | tr -s ' ' ',' > my_file.tmp
mv my_file.tmp my_file.fasta
echo -e "\n"
head -n 4 my_file.fasta

## 6. ¿Cuántos isogrupos únicos hay en el archivo? ##
echo "6. ¿Cuántos isogrupos únicos hay en el archivo?"
grep '>' my_file.fasta | head -n 3
grep '>' my_file.fasta | cut -d ',' -f 4 | head -n 3
echo "Total de isogrupos unicos"
grep '>' my_file.fasta | cut -d ',' -f 4 | uniq | wc -l

## 7. ¿Qué contig tiene mayor número de lecturas (numreads)? ¿Cuántas lecturas tiene? ##
echo "7. ¿Qué contig tiene mayor número de lecturas (numreads)? ¿Cuántas lecturas tiene?"
grep '>' my_file.fasta | cut -d ',' -f 1,3 | head -n 3
grep '>' my_file.fasta | cut -d ',' -f 1,3 | sort -t '=' -n -k 2 | head -n 5
echo "Conting con mayor número de lecturas"
grep '>' my_file.fasta | cut -d ',' -f 1,3 | sort -t '=' -n -k 2 -r | head -n 1

^G Help      ^O Write Out  ^W Where Is   ^K Cut        ^T Execute    ^C Location   M-U Undo
^X Exit      ^R Read File  ^\ Replace    ^U Paste      ^J Justify    ^/ Go To Line M-E Redo
```

2. Ejercicio 1.10.2

```
Melina@DESKTOP-C7U9BDL MINGW32 ~/Documents/GitHub/CLASE4 (main)
$ nano clase4_while.sh
```

```
MINGW32:/c/Users/user/Documents/GitHub/CLASE4
GNU nano 5.8 clase4_while.sh
## Ejercicio 1.10.2 Hormone Levels in Baboons ##
# 1. ¿Cuántas veces se registraron los niveles de los individuos 3 y 27? #
echo "1. ¿Cuántas veces se registraron los niveles de los individuos 3 y 27?"
cd ../../../../
grep -r "Gesquiere2011" CSB-master/unix/
cd CSB-master/unix/data/
ls
head -n 3 Gesquiere2011_data.csv
cut -f 1 Gesquiere2011_data.csv | head -n 4
echo -e "\n"
echo "Individuo 3"
cut -f 1 Gesquiere2011_data.csv | grep -c -w 3
echo "Individuo 27"
cut -f 1 Gesquiere2011_data.csv | grep -c -w 27
echo "2. Escribir un script tomando como entrada el nombre del archivo y el ID de la persona y devolviendo el número de registros para ese ID."
# 2. Escribir un script tomando como entrada el nombre del archivo y el ID de la persona y devolviendo el número de registros para ese ID. #
echo "Individuo 3"
bash ../solutions/count_baboons.sh Gesquiere2011_data.csv 3
echo "Individuo 27"
bash ../solutions/count_baboons.sh Gesquiere2011_data.csv 27
# cut -f 1 $1 | grep -c -w $2 #
echo -e "\n"
# 3. Escriba un script que devuelva el número de veces que se muestreó a cada individuo. #
echo "3. Escriba un script que devuelva el número de veces que se muestreó a cada individuo."
echo "Lista de individuos"
tail -n +2 Gesquiere2011_data.csv | cut -f 1 | sort -n | uniq
echo "Número de veces que se muestro a cada individuo"
myIDS=tail -n +2 Gesquiere2011_data.csv | cut -f 1 | sort -n | uniq
for id in $myIDS
do
    mycounts= bash ../solutions/count_baboons.sh Gesquiere2011_data.csv $id
    echo "ID:" $id "counts:" $mycounts
done
```

3. Ejercicio 1.10.3

```
Melina@DESKTOP-C7U9BDL MINGW32 ~/Documents/GitHub/CLASE4 (main)
$ nano clase4_until.sh
MINGW32:/c/Users/user/Documents/GitHub/2022II_gbi6/CLASE4
GNU nano 5.8 clase4_until.sh
## Ejercicio 1.10.3 Plant-Pollinator Networks ##
# 1. Escriba un script que tome uno de estos archivos y determine el número de filas (polinizadores) y el número de columnas (plantas). #
echo "1. Escriba un script que tome uno de estos archivos y determine el número de filas (polinizadores) y el número de columnas (plantas). "
cd ../../../../CSB-master/unix/data/Saavedra2013
ls -v

bash ../scriptconcatenado.sh n5.txt
echo -e "\n"
# 2. Escriba un script que imprima los números de filas y columnas para cada red. #
echo "2. Escriba un script que imprima los números de filas y columnas para cada red."
Files=.txt

for f in $Files
do
    mifila=cat $f | wc -l
    micolumna= head -n 1 $f |tr -d ' ' | tr -d '\n' | wc -c
    echo "Mis archivos:" $f "Filas:" $mifila "Columnas:" $micolumna
done
echo -e "\n"
# 3. ¿Qué archivo tiene el mayor número de filas? ¿Cuál tiene el mayor número de columnas?
echo "3. ¿Qué archivo tiene el mayor número de filas? ¿Cuál tiene el mayor número de columnas?"
echo "El mayor número de columnas"
bash ../scriptconcatenado2.sh | sort -n -r -k 3 | head -n 1
echo -e "\n"
echo "El mayor número de filas"
bash ../scriptconcatenado2.sh | sort -n -r -k 2 | head -n 1

[ Read 26 lines (Converted from DOS Format) ]
AG Help      AO Write Out  AW Where Is   AK Cut        AT Execute    AC Location  M-U Undo     M-A Set Mark  M-J To B
AX Exit      AR Read File  A\ Replace    AU Paste      AJ Justify    A/ Go To Line M-E Redo     M-6 Copy     AQ Where
```

4. Archivos csv

```
Melina@DESKTOP-C7U9BDL MINGW32 ~/Documents/GitHub/CLASE4 (main)
$ bash clase4_if.sh > clase4_if.csv

Melina@DESKTOP-C7U9BDL MINGW32 ~/Documents/GitHub/CLASE4 (main)
$ |
```

[illegible]

clase4_if - Excel

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista ¿Qué desea hacer?

Pegar Fuente Alineación Número

Calibri 11 A A

N K S Fuente A

Ajustar texto General

Combinar y centrar \$ % 000 0,00 0,00

1. Cambio de directorio a CSB/Unix/sandbox

	A	B	C	D	E	F	G	H	I
18	16								
19	5. Reemplace el delimitador original de dos espacios con una coma								
20	>contig00001 length=527 numreads=2 gene=isogroup00001 status=it_thresh								
21	ATCCTAGCTACTCTGGAGACTGAGGATTGAAGTTCAAAGTCAGCTCAAGCAAGAGATTG								
22	TTTACAATTAACCCACAAAAGGCTGTTACTGAAGGTGTGGCTTAAGTGTGAGAGCAACAG								
23	CTATGAGTGGAGGAATTTCTATTACAATATAATTCATCTCTGGTAAATTGACCAATTA								
24									
25									
26	>contig00001,length=527,numreads=2,gene=isogroup00001,status=it_thresh								
27	ATCCTAGCTACTCTGGAGACTGAGGATTGAAGTTCAAAGTCAGCTCAAGCAAGAGATTG								
28	TTTACAATTAACCCACAAAAGGCTGTTACTGAAGGTGTGGCTTAAGTGTGAGAGCAACAG								
29	CTATGAGTGGAGGAATTTCTATTACAATATAATTCATCTCTGGTAAATTGACCAATTA								
30	6. ¿Cuántos isogrupos únicos hay en el archivo?								
31	>contig00001,length=527,numreads=2,gene=isogroup00001,status=it_thresh								
32	>contig00002,length=551,numreads=8,gene=isogroup00001,status=it_thresh								
33	>contig00003,length=541,numreads=2,gene=isogroup00001,status=it_thresh								
34	gene=isogroup00001								
35	gene=isogroup00001								
36	gene=isogroup00001								
37	Total de isogrupos unicos								
38	43								
39	7. ¿Qué contig tiene mayor número de lecturas (numreads)? ¿Cuántas lecturas tiene?								
40	>contig00001,numreads=2								

clase4_if

Listo

clase4_while - Excel

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista ¿Qué desea hacer?

Portapapeles Fuente Alineación Número Formato condicional Dar formato como Estilos

Calibri 11 Ajustar texto General \$ % 000 ,00 0,00

A1 1.¿Cuántas veces se registraron los niveles de los individuos 3 y 27?

	A	B	C	D	E	F	G	H	I	J
1	1.¿Cuántas veces se registraron los niveles de los individuos 3 y 27?									
2	CSB-master/unix/solutions/count_all_baboons.sh:myIDS='tail -n +2 ../data/Gesquiere2011_data.csv cut -f 1 sort -n uniq`									
3	CSB-master/unix/solutions/count_all_baboons.sh: mycounts='bash count_baboons.sh ../data/Gesquiere2011_data.csv \$id`									
4	CSB-master/unix/solutions/Gesquiere2011_solution.sh:head -n 3 ../data/Gesquiere2011_data.csv									
5	CSB-master/unix/solutions/Gesquiere2011_solution.sh:cut -f 1 ../data/Gesquiere2011_data.csv head -n 3									
6	CSB-master/unix/solutions/Gesquiere2011_solution.sh:cut -f 1 ../data/Gesquiere2011_data.csv grep -c -w 3									
7	CSB-master/unix/solutions/Gesquiere2011_solution.sh:cut -f 1 ../data/Gesquiere2011_data.csv grep -c -w 27									
8	CSB-master/unix/solutions/Gesquiere2011_solution.sh:bash count_baboons.sh ../data/Gesquiere2011_data.csv 27									
9	CSB-master/unix/solutions/Gesquiere2011_solution.sh:tail -n +2 ../data/Gesquiere2011_data.csv cut -f 1 sort -n uniq									
10	CSB-master/unix/solutions/Gesquiere2011_solution.sh:myIDS='tail -n +2 ../data/Gesquiere2011_data.csv cut -f 1 sort -n uniq`									
11	CSB-master/unix/solutions/Gesquiere2011_solution.sh: mycounts='bash count_baboons.sh ../data/Gesquiere2011_data.csv \$id`									
12	Buzzard2015_about.txt									
13	Buzzard2015_data.csv									
14	Gesquiere2011_about.txt									
15	Gesquiere2011_data.csv									
16	Marra2014_about.txt									
17	Marra2014_data.fasta									
18	Pacifici2013_about.txt									
19	Pacifici2013_data.csv									
20	Saavedra2013									
21	Saavedra2013_about.txt									
22	miRNA									
23	scriptconcatenado.sh									

clase4_while

Listo

clase4_while - Excel

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista ¿Qué desea hacer?

Pegar Fuente Alineación Número Estilos

Formato condicional Dar formato como tabla Estilos de celda

A1 1. ¿Cuántas veces se registraron los niveles de los individuos 3 y 27?

	A	B	C	D	E	F	G	H	I	J	K
23	scriptconcatenado.sh										
24	maleIDGCT										
25	166.964.57										
26	151.0935.57										
27	maleID										
28	1										
29	1										
30	1										
31											
32											
33	Individuo 3										
34	61										
35	Individuo 27										
36	5										
37	2. Escribir un script tomando como entrada el nombre del archivo y el ID de la persona y devolviendo el número de registros para ese ID.										
38	Individuo 3										
39	61										
40	Individuo 27										
41	5										
42											
43											
44	3. Escriba un script que devuelva el número de veces que se muestra a cada individuo.										
45	Lista de individuos										

clase4_while

Listo

Escribe aquí para buscar

clase4_while - Excel

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista ¿Qué desea hacer?

Pegar Fuente Alineación Número

A1 1. ¿Cuántas veces se registraron los niveles de los individuos 3 y 27?

	A	B	C	D	E	F	G	H	I
42									
43									
44	3. Escriba un script que devuelva el número de veces que se muestra a cada individuo.								
45	Lista de individuos								
46	1								
47	2								
48	3								
49	4								
50	5								
51	6								
52	7								
53	8								
54	9								
55	10								
56	11								
57	12								
58	13								
59	14								
60	15								
61	16								
62	17								
63	18								
64	19								

clase4_while

Listo

Excel interface showing a spreadsheet titled "clase4_while". The formula bar displays the formula: `=1.¿Cuántas veces se registraron los niveles de los individuos 3 y 27?`

The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I
169	126								
170	127								
171	Número de veces que se muestro a cada individuo								
172	ID: 1 counts: 10								
173	ID: 2 counts: 2								
174	ID: 3 counts: 61								
175	ID: 4 counts: 46								
176	ID: 5 counts: 28								
177	ID: 6 counts: 7								
178	ID: 7 counts: 5								
179	ID: 8 counts: 17								
180	ID: 9 counts: 4								
181	ID: 10 counts: 21								
182	ID: 11 counts: 26								
183	ID: 12 counts: 23								
184	ID: 13 counts: 16								
185	ID: 14 counts: 1								
186	ID: 15 counts: 40								
187	ID: 16 counts: 31								
188	ID: 17 counts: 3								
189	ID: 18 counts: 4								
190	ID: 19 counts: 3								
191	ID: 20 counts: 4								

The taskbar shows a terminal window with the following commands:

```
Melina@DESKTOP-C7U9BDL MINGW32 ~/Documents/GitHub/CLASE4 (main)
$ bash clase4_util.sh > clase4_util.csv
Melina@DESKTOP-C7U9BDL MINGW32 ~/Documents/GitHub/CLASE4 (main)
$
```

Excel interface showing the ribbon (Archivo, Inicio, Insertar, Diseño de página, Fórmulas, Datos, Revisar, Vista) and the task pane (¿Qué desea hacer?). The active sheet is 'clase4_until'.

The task pane contains the following instructions:

1. Escriba un script que tome uno de estos archivos y determine el número de filas (polinizadores) y el número de columnas (plantas).

The spreadsheet data is as follows:

	A	B	C	D	E	F	G	H	I	J	K
1	1. Escriba un script que tome uno de estos archivos y determine el número de filas (polinizadores) y el número de columnas (plantas).										
2	n1.txt										
3	n2.txt										
4	n3.txt										
5	n4.txt										
6	n5.txt										
7	n6.txt										
8	n7.txt										
9	n8.txt										
10	n9.txt										
11	n10.txt										
12	n11.txt										
13	n12.txt										
14	n13.txt										
15	n14.txt										
16	n15.txt										
17	n16.txt										
18	n17.txt										
19	n18.txt										
20	n19.txt										
21	n20.txt										
22	n21.txt										
23	n22.txt										

The task pane also shows the status bar: Listo.

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista ¿Qué?

Pegar Fuente Alineación

Calibri 11 A A

N K S Fuente Alineación

Ajustar texto Combinar y centrar

1. Escriba un script que tome uno de estos archivos y dete

	A	B	C	D	E	F	G
57	n56.txt						
58	n57.txt						
59	n58.txt						
60	n59.txt						
61	Nombre de archivo: n5.txt						
62	Número de filas:21						
63	Número de columnas:7						
64							
65							
66	2. Escriba un script que imprima los números de filas y columnas para cada red.						
67	Mis archivos: n1.txt Filas: 97 Columnas: 80						
68	Mis archivos: n10.txt Filas: 14 Columnas: 20						
69	Mis archivos: n11.txt Filas: 270 Columnas: 91						
70	Mis archivos: n12.txt Filas: 7 Columnas: 72						
71	Mis archivos: n13.txt Filas: 61 Columnas: 17						
72	Mis archivos: n14.txt Filas: 35 Columnas: 15						
73	Mis archivos: n15.txt Filas: 38 Columnas: 11						
74	Mis archivos: n16.txt Filas: 118 Columnas: 24						
75	Mis archivos: n17.txt Filas: 76 Columnas: 31						
76	Mis archivos: n18.txt Filas: 13 Columnas: 14						
77	Mis archivos: n19.txt Filas: 10 Columnas: 16						
78	Mis archivos: n2.txt Filas: 62 Columnas: 41						
79	Mis archivos: n20.txt Filas: 18 Columnas: 7						

clase4_until

Listo

clase4_until - Excel

Archivo Inicio Insertar Diseño de página Fórmulas Datos Revisar Vista ¿Qué desea hacer?

Pegar Fuente Alineación Número Formato condicional Dar formato como tabla Estilos

A1 1. Escriba un script que tome uno de estos archivos y determine el número de filas (polinizadore

	A	B	C	D	E	F	G	H	I	J	K
112	Mis archivos: n50.txt	Filas: 45	Columnas: 46								
113	Mis archivos: n51.txt	Filas: 8	Columnas: 15								
114	Mis archivos: n52.txt	Filas: 33	Columnas: 7								
115	Mis archivos: n53.txt	Filas: 34	Columnas: 13								
116	Mis archivos: n54.txt	Filas: 126	Columnas: 25								
117	Mis archivos: n55.txt	Filas: 14	Columnas: 50								
118	Mis archivos: n56.txt	Filas: 110	Columnas: 207								
119	Mis archivos: n57.txt	Filas: 14	Columnas: 11								
120	Mis archivos: n58.txt	Filas: 678	Columnas: 90								
121	Mis archivos: n59.txt	Filas: 663	Columnas: 130								
122	Mis archivos: n6.txt	Filas: 9	Columnas: 31								
123	Mis archivos: n7.txt	Filas: 16	Columnas: 25								
124	Mis archivos: n8.txt	Filas: 19	Columnas: 33								
125	Mis archivos: n9.txt	Filas: 12	Columnas: 22								
126											
127											
128	3. ¿Qué archivo tiene el mayor número de filas? ¿Cuál tiene el mayor número de columnas?										
129	El mayor número de columnas										
130	n56.txt 110 207										
131											
132											
133	El mayor número de filas										
134	n58.txt 678 90										

clase4_until

Listo

Escribe aquí para buscar

CLASE4

Archivo Inicio Compartir Vista

Anclar al acceso rápido Copiar Pegar Mover a Eliminar Copiar a Cambiar nombre Nueva carpeta Propiedades Abrir Seleccionar todo No seleccionar ninguno Invertir selección

Portapapeles Organizar Nuevo Abrir Seleccionar

← → ↕ ↑ << GitHub > CLASE4 🔍 Buscar en CLASE4

	Nombre	Fecha de modificación	Tipo	Tamaño
★ Acceso rápido				
Escritorio	.gitattributes	14/11/2022 21:08	Archivo GITATTRIB...	1 KB
Descargas	clase4_if	14/11/2022 21:45	Archivo de valores...	2 KB
Documentos	clase4_if.sh	14/11/2022 21:44	Archivo SH	2 KB
Imágenes	clase4_until	14/11/2022 22:05	Archivo de valores...	4 KB
Clase4	clase4_until.sh	14/11/2022 22:03	Archivo SH	2 KB
CLASE4	clase4_while	14/11/2022 21:55	Archivo de valores...	5 KB
sandbox	clase4_while.sh	14/11/2022 21:53	Archivo SH	2 KB
tc2_VargasMelin	README.md	14/11/2022 21:24	Archivo MD	6 KB
OneDrive - Person				
Este equipo				
Descargas				
Documentos				
Escritorio				
Imágenes				

8 elementos | 1 elemento seleccionado 3,50 KB