

Tarea: TC3_Forloop

1.10.1 Capturas de la Resolución del ejercicio.

```
● @EmiCherrez → /workspaces/CSB (master) $ cd unix/  
● @EmiCherrez → /workspaces/CSB/unix (master) $ cd sandbox/  
○ @EmiCherrez → /workspaces/CSB/unix/sandbox (master) $ █
```

```
GNU nano 4.8 TC3_E1.sh  
## tamaño del archivo  
echo "para el tamaño del archivo ejecutamos el siguiente código"  
ls -lh ../data/Marra2014_data.fasta  
## copia de Marra2014_data.fasta en sandbox  
echo "para crear la copia en sandbox de Marra2014_data.fasta y nombrarla se ejecuta el código"  
cp ../data/Marra2014_data.fasta my_file.fasta  
## saber cuando contigs se clasifican como isogrupos00036  
echo "para saber cuántos son isogrupos00036 se ejecuta el código"  
grep -c isogrupos00036 my_file.fasta  
## separación del delimitador original de los espacios  
echo "para separar se utiliza el siguiente código"  
cat my_file.fasta | tr -s ' ' | head -n 3  
## guardar con un sobrescrito de archivo temporal  
echo "para guardar con un sobrescrito temporal el archivo se ejecuta lo siguiente"  
cat my_file.fasta | tr -s ' ' > my_file.tmp  
mv my_file.tmp my_file.fasta  
## para saber cuántos isogrupos únicos hay en el archivo  
echo "para saber cuántos isogrupos únicos hay en el archivo se ejecuta el siguiente código"  
grep '>' my_file.fasta | cut -d ',' -f 4 | sort | uniq | wc -l  
## para saber que contig tiene el mayor número de lectura  
echo "para saber que contig tiene el mayor número de lectura se usa el siguiente código"  
grep '>' my_file.fasta | cut -d ',' -f 1,3 | head -n 3  
## para ordenar según el número de lee  
echo "para ordenar según el número de lee se utiliza el siguiente código"  
grep '>' my_file.fasta | cut -d ',' -f 1,3 | sort -t '=' -k 2 -n | head -n 5 █
```

```
● @EmiCherrez → /workspaces/CSB/unix/sandbox (master) $ nano TC3_E1.sh  
● @EmiCherrez → /workspaces/CSB/unix/sandbox (master) $ bash TC3_E1.sh  
para el tamaño del archivo ejecutamos el siguiente código  
-rw-rw-rw- 1 codespace root 553K May  9 06:45 ../data/Marra2014_data.fasta  
para crear la copia en sandbox de Marra2014_data.fasta y nombrarla se ejecuta el código  
para saber cuántos son isogrupos00036 se ejecuta el código  
0  
para separar se utiliza el siguiente código  
>contig00001,length=527,numreads=2,gene=isogroup00001,status=it_thresh  
ATCCTAGCTACTCTGGAGACTGAGGATTGAAGTTCAAAGTCAGCTCAAGCAAGAGATTG  
TTTACAATTAACCCACAAAAGGCTGTTACTGAAGGTGTGGCTTAAGTGTCAAGCAACAG  
para guardar con un sobrescrito temporal el archivo se ejecuta lo siguiente  
mv: cannot stat 'my_file.tmp': No such file or directory  
para saber cuántos isogrupos únicos hay en el archivo se ejecuta el siguiente código  
955  
para saber que contig tiene el mayor número de lectura se usa el siguiente código  
>contig00001 length=527 numreads=2 gene=isogroup00001 status=it_thresh  
>contig00002 length=551 numreads=8 gene=isogroup00001 status=it_thresh  
>contig00003 length=541 numreads=2 gene=isogroup00001 status=it_thresh  
para ordenar según el número de lee se utiliza el siguiente código  
>contig00090 length=100 numreads=3 gene=isogroup00001 status=it_thresh  
>contig00719 length=100 numreads=49 gene=isogroup00014 status=isotig  
>contig01037 length=100 numreads=38 gene=isogroup00027 status=it_thresh  
>contig01049 length=100 numreads=20 gene=isogroup00027 status=it_thresh  
>contig00309 length=101 numreads=1807 gene=isogroup00002 status=it_thresh  
○ @EmiCherrez → /workspaces/CSB/unix/sandbox (master) $ █
```

1.10.2

```
GNU nano 4.8 TC3_E2.sh
## cuantas veces fueron los niveles de los individuos 3 y 27 grabados?
echo "para saber cuantas veces fueron los niveles de los individuos 3 y 27 se utiliza el codigo"
echo "se puede ver el archivo de la siguiente manera"
head -n 3 ../data/Gesquiere2011_data.csv
echo "para este caso solo se usa la columna 1"
cut -f 1 ../data/Gesquiere2011_data.csv | head -n 3
echo "para contar el numero de ocurrencia"
## identificacion masculina 3
echo "para la identificacion masculina 3 se usa el siguiente codigo"
cut -f 1 ../data/Gesquiere2011_data.csv | grep -c -w 3
## identificacion masculina 27
echo "para la identificacion masculina 27 se usa el siguiente codigo"
cut -f 1 ../data/Gesquiere2011_data.csv | grep -c -w 27
## ESCRIBA UN SCRIPT tomando como entrada el nombre del archivo y el DNI del individuo, devolviendo el numero
echo "Script tomando como entrada el nombre del archivo y su DNI"
bash TC3_E2.2.sh ../data/Gesquiere2011_data.csv 27
```

```
GNU nano 4.8 TC3_E2.2.sh
## para poder leer la entrada del comando de linea

cut -f 1 $1 | grep -c -w $2
```

```
GNU nano 4.8 TC3_E2.3.sh Modified
myIDS=tail -n +2 ../data/Gesquiere2011_data.csv | cut -f 1 | sort -n | uniq`
for id in $myIDS
do
    mycounts=`bash TC3_E2.2.sh
    ../data/Gesquiere2011_data.csv $id`
    echo "ID:" $id "counts:" $mycounts
done
```

- @EmiCherrez → /workspaces/CSB/unix/sandbox (master) \$ bash TC3_E2.sh
para saber cuantas veces fueron los niveles de los individuos 3 y 27 se utiliza el codigo
se puede ver el archivo de la siguiente manera
maleID GC T
1 66.9 64.57
1 51.09 35.57
para este caso solo se usa la columna 1
maleID
1
1
para contar el numero de ocurrencia
para la identificacion masculina 3 se usa el siguiente codigo
61
para la identificacion masculina 27 se usa el siguiente codigo
5
Script tomando como entrada el nombre del archivo y su DNI
5
- @EmiCherrez → /workspaces/CSB/unix/sandbox (master) \$ █

1.10.3

```

• @EmiCherrez → /workspaces/CSB/unix/sandbox (master) $ cd ../
• @EmiCherrez → /workspaces/CSB/unix (master) $ cd data/
• @EmiCherrez → /workspaces/CSB/unix/data (master) $ cd Saavedra2013
○ @EmiCherrez → ../CSB/unix/data/Saavedra2013 (master) $ █

```

```

GNU nano 4.8 TC3_E3.sh Modified
## Escriba un script que tome uno de estos archivos y determine el numero de filas y columnas
echo "conteo de las filas"
wc -l n10.txt
echo "regresion"
cat n10.txt | wc -l
echo "conteo del numero de columnas"
echo "se utiliza la primera linea"
head -n 1 n10.txt
echo "eliminacion de los espacios y del terminador de linea"
head -n 1 n10.txt | tr -d ' ' | tr -d '\n'
echo "conteo de caracteres"
head -n 1 n10.txt | tr -d ' ' | tr -d '\n' | wc -c
## Escriba u script que imprima el numero de filas y columnas para cada red
echo "impresion del numero de filas y columnas para cada red"
bash TC3_E3.2.sh
## Cual es la red con el mayot numero de filas? Cual es el que tiene mayor numero de columna?
echo "red con mas filas"
bash TC3_E3.2.sh | sort -n -r -k 2 | head -n 1
echo "red con mas columnas"
bash TC3_E3.2.sh | sort -n -r -k 3 | head -n 1

```

```

GNU nano 4.8 TC3_E3.2.sh Modified
FILES=*.txt
for f in $FILES
do
    myrow=`cat $f | wc -l`
    mycol=`head -n 1 $f | tr -d ' ' | tr -d '\n' | wc -c`
    echo $f $myrow $mycol
done

```

```

• @EmiCherrez → ../CSB/unix/data/Saavedra2013 (master) $ bash TC3_E3.sh
conteo de las filas
14 n10.txt
regresion
14
conteo del numero de columnas
se utiliza la primera linea
0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0
eliminacion de los espacios y del terminador de linea
01000001000000000100conteo de caracteres
20
impresion del numero de filas y columnas para cada red
n1.txt 97 80
n10.txt 14 20
n11.txt 270 91
n12.txt 7 72
n13.txt 61 17
n14.txt 35 15
n15.txt 38 11
n16.txt 118 24
n17.txt 76 31
n18.txt 13 14
n19.txt 10 16
n2.txt 62 41
n20.txt 18 7
n21.txt 19 45
n22.txt 19 36
n23.txt 179 26
n24.txt 80 28
n25.txt 17 16
n26.txt 82 40
n27.txt 27 5
n28.txt 90 19
n29.txt 61 25
n3.txt 25 36
n30.txt 8 19
n31.txt 28 25
n32.txt 45 21
n33.txt 70 20
n34.txt 79 25
n35.txt 14 8
n36.txt 40 169

```

```
n37.txt 44 13
n38.txt 51 99
n39.txt 33 25
n4.txt 101 11
n40.txt 28 18
n41.txt 12 10
n42.txt 42 8
n43.txt 55 29
n44.txt 56 9
n45.txt 36 61
n46.txt 58 17
n47.txt 139 41
n48.txt 118 49
n49.txt 47 23
n5.txt 21 7
n50.txt 45 46
n51.txt 8 15
n52.txt 33 7
n53.txt 34 13
n54.txt 126 25
n55.txt 14 50
n56.txt 110 207
n57.txt 14 11
n58.txt 678 90
n59.txt 663 130
n6.txt 9 31
n7.txt 16 25
n8.txt 19 33
n9.txt 12 22
red con mas filas
n58.txt 678 90
red con mas columnas
n56.txt 110 207
@EmiCherrez → ../CSB/unix/data/Saavedra2013 (master) $
```

1.10.4

```
● @EmiCherrez → ../CSB/unix/data/Saavedra2013 (master) $ cd ../
○ @EmiCherrez → /workspaces/CSB/unix/data (master) $
```

```
GNU nano 4.8 TC3_E4.sh Modified
## -nombre de la columna
## -numero de valores distintos
## -valor minimo
## -valor maximo
echo "extraccion del nombre de la columna"
cut -d ',' -f 7 ../data/Buzzard2015_data.csv | head -n 1
echo "numero de valores distintos"
cut -d ',' -f 7 ../data/Buzzard2015_data.csv | tail -n +2 | sort | uniq | wc -l
echo "valor maximo"
cut -d ',' -f 7 ../data/Buzzard2015_data.csv | tail -n +2 | sort -n | tail -n 1
echo "valor minimo"
cut -d ',' -f 7 ../data/Buzzard2015_data.csv | tail -n +2 | sort -n | tail -n 1
```

```
@EmiCherrez → /workspaces/CSB/unix (master) $ cd data
@EmiCherrez → /workspaces/CSB/unix/data (master) $ nano TC3_E4.sh
● @EmiCherrez → /workspaces/CSB/unix/data (master) $ nano TC3_E4.sh
● @EmiCherrez → /workspaces/CSB/unix/data (master) $ bash TC3_E4.sh
extraccion del nombre de la columna
biomass
numero de valores distintos
285
valor maximo
14897.29471
valor minimo
14897.29471
○ @EmiCherrez → /workspaces/CSB/unix/data (master) $
```