

Deber número 3_Bioinformática

Nombre: Angelica Flores

Grupo: G01

Orden: Resolver los ejercicios del 1.10 .

Ejercicio 1

```
aula@LAB-D-E3 MINGW64 ~/Desktop
$ cd

aula@LAB-D-E3 MINGW64 ~
$ cd Documents/

aula@LAB-D-E3 MINGW64 ~/Documents
$ cd CSB-master/

aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master
$ cd unix/

aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix
$ cd sandbox/

aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/sandbox
$ nano PrimerEjercicio.sh

aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/sandbox
$ bash PrimerEjercicio.sh
Comandos para el tamaño del archivo
PrimerEjercicio.sh: line 3: ls-lh: command not found
Comandos para realizar una copiado Marra2014_data.fasta en sandbox
Comando para determinar la clasificacion de contings del isogrupa

aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/sandbox
$ bash PrimerEjercicio.sh
Comandos para el tamaño del archivo
PrimerEjercicio.sh: line 3: ls-lh: command not found
Comandos para realizar una copiado Marra2014_data.fasta en sandbox
Comando para determinar la clasificacion de contings del isogrupa
0
Comandos para separa con una coma
>contig00001,length=527,numreads=2,gene=isogroup00001,status=it_thresh
ATCCTAGCTACTCTGGAGACTGAGGATTGAAGTTCAAAGTCAGCTCAAGCAAGAGATTG
TTTACAATTAACCCACAAAAGGCTGTACTGAAGGTGTGGCTTAAGTGTCAGAGCAACAG
Comando para sobrescribir el archivo
Comando para conocer isogrpos unicos
wc: unknown option -- 1
Try 'wc --help' for more information.
Comando que determina el contings con mayor numero de lectura
>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
Comando para ordenar segun el número
>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
```

Ejercicio2

```
aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/sandbox
$ nano SegundoEjercicio.sh
aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/sandbox
$ nano SegundoEjercicio2.sh

aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/sandbox
$ nano SegundoEjercicio.sh

aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/sandbox
$ nano SegundoEjercicio3.sh

aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/sandbox
$ nano SegundoEjercicio.sh

aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/sandbox
$ nano SegundoEjercicio.sh
```

```
GNU nano 6.4 SegundoEjercicio.sh
#Numero de niveles de individuos 3 y 27 grabados#
echo "Comando para determinar el número de veces de ives de 3 individuos y 27 grab>
head -n 3 ../data/Gesquiere2011_data.csv
cut -f 1 ../data/Gesquiere2011_data.csv | head -n 3
cut -f 1 ../data/Gesquiere2011_data.csv | grep -c -w 3

#Identificación de 3 y 27 masculina#
echo "Comando en 3"
cut -f 1 ../data/Gesquiere2011_data.csv | grep -c -w 3
echo "Comando en 27"
cut -f 1 ../data/Gesquiere2011_data.csv | grep -c -w 27

#Script con el nombre del archivo e ID#
echo "Comando para script con nombre e ID"
bash SegundoEjercicio2.sh ../data/Gesquiere2011_data.csv 27

[ Read 19 lines ]
^G Help      ^O Write Out ^W Where Is  ^K Cut       ^T Execute   ^C Location
^X Exit      ^R Read File ^\ Replace   ^U Paste     ^J Justify   ^_ Go To Line
```

```
#Comando en linea#
cut -f 1 $1 | grep -c -w $2

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

```
GNU nano 6.4 SegundoEjercicio3.sh
#Tres#
myIDS=`tail -n +2 ../data/Gesquiere2011_data.csv | cut -f 1 | sort -n | uniq`
for id in $myIDS
do
    mycounts=`bash SegundoEjercicio2.sh ../data/Gesquiere2011_data.csv $id`
    echo "ID:" $id "counts:" $mycounts
done

[ Read 7 lines ]
^G Help      ^O Write Out ^W Where Is  ^K Cut       ^T Execute   ^C Location
^X Exit      ^R Read File ^\ Replace   ^U Paste     ^J Justify   ^_ Go To Line
```

```
aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/sandbox
$ bash SegundoEjercicio.sh
Comando para determinar el número de veces de ives de 3 individuos y 27 grabados
maleID GC T
1 66.9 64.57
1 51.09 35.57
maleID
1
1
61
Comando en 3
61
Comando en 27
5
Comando para script con nombre e ID
5
Comando para determinar cuantas veces se tomó muestra de un individuo
ID: 1 counts: 10
ID: 2 counts: 2
```

```
ID: 1 counts: 10
ID: 2 counts: 2
ID: 3 counts: 61
ID: 4 counts: 46
ID: 5 counts: 28
ID: 6 counts: 7
ID: 7 counts: 5
ID: 8 counts: 17
ID: 9 counts: 4
ID: 10 counts: 21
ID: 11 counts: 26
ID: 12 counts: 23
ID: 13 counts: 16
ID: 14 counts: 1
ID: 15 counts: 40
ID: 16 counts: 31
ID: 17 counts: 3
ID: 18 counts: 4
ID: 19 counts: 3
ID: 20 counts: 4
```

```
ID: 110 counts: 3
ID: 111 counts: 24
ID: 112 counts: 3
ID: 113 counts: 1
ID: 114 counts: 1
ID: 115 counts: 1
ID: 116 counts: 14
ID: 118 counts: 23
ID: 119 counts: 1
ID: 120 counts: 42
ID: 121 counts: 12
ID: 122 counts: 9
ID: 123 counts: 39
ID: 124 counts: 1
ID: 125 counts: 39
ID: 126 counts: 15
ID: 127 counts: 13
```

Ejercicio3

```
aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/data/Saavedra2013
$ nano TercerEjercicio.sh
```

```
aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/data/Saavedra2013
$ nano TercerEjercicio2.sh
```

```
GNU nano 6.4 TercerEjercicio.sh
#Script para filas y columnas#
echo "Filas"
wc -l n10.txt
cat n10.txt | wc -l
echo "Columnas"
head -n 1 n10.txt
head -n 1 n10.txt | tr -d ' ' | tr -d '\n'
head -n 1 n10.txt | tr -d ' ' | tr -d '\n' | wc -c
head -n 1 n10.txt
head -n 1 n10.txt | tr -d ' ' | tr -d '\n'
head -n 1 n10.txt | tr -d ' ' | tr -d '\n' | wc -c

#Imprimir filas y columnas para red#
bash TercerEjercicio2.sh

#Red con más filas y columnas#
echo "Comando red con más filas"
bash TercerEjercicio2.sh | sort -n -r -k 2 | head -n 1

echo "Comando red con más columnas"
bash TercerEjercicio2.sh | sort -n -r -k 3 | 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
```

```
GNU nano 6.4 TercerEjercicio2.sh
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

[ Read 8 lines ]

^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
```

```
aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/data/Saavedra2013
$ bash TercerEjercicio.sh
Filas
14 n10.txt
14
Columnas
0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0
0100000100000000010020
0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0
0100000100000000010020
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
```

```
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
Comando red con más filas
n58.txt 678 90
Comando red con más columnas
n56.txt 110 207
```

Ejercicio4

```
aula@LAB-D-E3 MINGW64 ~  
$ cd  
  
aula@LAB-D-E3 MINGW64 ~  
$ cd Documents/  
  
aula@LAB-D-E3 MINGW64 ~/Documents  
$ cd CSB-master/  
  
aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master  
$ cd unix/  
  
aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix  
$ cd data/  
  
aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/data  
$ nano CuartoEjercicio.sh
```

```
GNU nano 6.4 CuartoEjercicio.sh  
#Script para archivo csv y número de columnas#  
echo "Nombre de columna"  
cut -d ',' -f 7 ../data/Buzzard2015_data.csv | head -n 1  
echo "Valores distintos"  
cut -d ',' -f 7 ../data/Buzzard2015_data.csv | tail -n +2 | sort | uniq | wc -l  
echo "Valor máximo"  
cut -d ',' -f 7 ../data/Buzzard2015_data.csv | tail -n +2 | sort -n | tail -n 1  
echo "valor mínimo"  
cut -d ',' -f 7 ../data/Buzzard2015_data.csv | tail -n +2 | sort -n | head -n 1
```

```
[ Read 9 lines ]  
^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
```

```
aula@LAB-D-E3 MINGW64 ~/Documents/CSB-master/unix/data  
$ bash CuartoEjercicio.sh  
Nombre de columna  
biomass  
Valores distintos  
285  
Valor máximo  
14897.29471  
valor mínimo  
1.048466198
```