

```

1 ---
2 title: "PEC1_ADO"
3 output:
4   pdf_document: default
5   word_document: default
6 date: "2024-11-04"
7 ---
8
9 ###PEC1_ADO
10
11 ##Ejercicio 1
12
13 #Instalamos bioconductor y SummarizedExperiment
14
15 ```{r install-packages}
16 if (!requireNamespace("BiocManager", quietly = TRUE))
17   install.packages("BiocManager")
18 BiocManager::install("SummarizedExperiment")
19 ```
20
21
22 ```{r}
23 BiocManager::install("SummarizedExperiment")
24 ```
25
26 #Librerias
27 ```{r cars}
28 library(readr)
29 library(SummarizedExperiment)
30 ```
31
32 #Introducimos los datos
33
34
35 #Asignamos nombre a las dos bases con las que vamos a trabajar
36 ```{r}
37 data_info <- DataInfo_S013
38 data_values <- DataValues_S013
39 ```
40
41 #Fijamos semilla (hacemos una prueba)
42 ```{r}
43 set.seed(123456)
44 seleccion <- sample(1:nrow(data_values), 1)
45 Indice <- data_values[seleccion, 1]
46 ColumnaSujeto <- data_values[seleccion, 2]
47 print(seleccion)
48 print(ColumnaSujeto)
49 ```
50
51
52 ##Ejercicio 2
53
54 #Hacemos dataframe de columnas. Como nos ha dado problemas con la longitud de las
columnas de datos (al intentar crear después el SummarizedExperiment), escogemos una
columna que sabemos que tiene el número correcto, y además escogemos la función
length.out para indicar el número de elementos que necesitamos que será =counts.
55 ```{r}
56 colData <- DataFrame(
57   Col = rep(data_info$...1, length.out = ncol(counts)),
58   row.names = colnames(counts)
59 )
60 print(colData)
61 ```

```

```

[1] 7
[1] 7

```

```
DataFrame with 696 rows and 1 column
  Col
<character>
1    SUBJECTS
2    SURGERY
3    AGE
4    GENDER
5    Group
...
692 SM.C18.1_T5
693 SM.C20.2_T5
694 SM.C24.0_T5
695 SM.C24.1_T5
696 SUBJECTS
```

```
62
63 #Hacemos dataframe de líneas, en este caso escogemos dos filas aleatorias.
64 ```{r}
65 rowData <- DataFrame(
66   AGE = data_values$AGE,
67   Surgery = as.factor(data_values$SURGERY)
68 )
69 print(rowData)
70
71 ^
```

```
DataFrame with 39 rows and 2 columns
  AGE Surgery
<numeric> <factor>
1      27 by pass
2      19 by pass
3      42 by pass
4      37 by pass
5      42 tubular
...
35     39 tubular
36     35 tubular
37     46 by pass
38     41 tubular
39     26 by pass
```

```
72 #Ejecutamos la función Summarized Experiment para juntar las dos dataframe construidas
    anteriormente.
73 ```{r}
74 se <- SummarizedExperiment(
75   assays = list(counts = counts),
76   colData = colData,
77   rowData = rowData
78 )
79 ^
80
81 #Recogemos la información de nuestra SummarizedExperiment.
82 ```{r}
83 print(se)
84 ^
```

```
class: SummarizedExperiment
dim: 39 696
metadata(0):
assays(1): counts
rownames: NULL
rowData names(2): AGE Surgery
colnames: NULL
colData names(1): Col
```

```
85 #Mostramos los datos del SummarizedExperiment. Ha sido posible crear la matriz pero
los nombres de las filas y columnas no aparecen, aparecen como 1,2,3,4 y 5. A pesar de
que en rowData y colData aparecían bien. He intentado solucionarlo pero no he podido.
```

```
86
87 ```{r}
88 assays(se)[[1]][1:5, 1:5]
89 ```
```

```
      [,1] [,2] [,3]      [,4] [,5]
[1,] " 1" " 1" "by pass" "27" "F"
[2,] " 2" " 2" "by pass" "19" "F"
[3,] " 3" " 3" "by pass" "42" "F"
[4,] " 4" " 4" "by pass" "37" "F"
[5,] " 5" " 5" "tubular" "42" "F"
```

```
90 |
91
92
93 ```{r}
94 summary(se)
95 ```
```

```
[1] "SummarizedExperiment object of length 39 with 2 metadata columns"
```

```
96
97
98
99 NOTAS: Siguiendo el tutorial proporcionado, no había manera de contruir el
SummarizedExpression. Tenía muchos problemas al montar el número de filar y columnas
así que he tenido que recurrir a otros tutoriales y vídeos que me permitieran más
información. Dejo la bibliografía abajo.
100
101 BB
102 https://www.youtube.com/watch?v=D8IVRmbMjyc
103 https://www.bioconductor.org/help/course-materials/2019/BSS2019/04\_Practical\_CoreApproachesInBioconductor.html
104
```