

Ejercicios de repaso módulo 3

Introducción al tidyverse

Servicios Integrales de Recursos Biológicos, Acuáticos y Ambientales



Autores: L. Daniel Carrillo-Colín; Oscar G. Zamora-García y J. Fernando Márquez-Farías

Introducción

Esta guía de ejercicios se centra en la aplicación práctica de los conceptos fundamentales del tidyverse, con especial énfasis en el paquete *dplyr* y el operador *pipe* aprendidos durante la sesión 3 del curso de *R esencial*. A lo largo de esta guía, los asistentes del curso consolidaran su comprensión de las operaciones básicas de manipulación de datos mediante la aplicación de funciones esenciales del *dplyr* como *select()*, *mutate()*, *filter()*, *group_by()*, *summary()*.

Instrucciones

A continuación, te mostraremos una serie de resultados obtenidos a través de líneas de código de R. Es necesario que mandes al correo: [cursos.sirbaa@gmail.com](mailto: cursos.sirbaa@gmail.com), el script que produce los mismos resultados aquí mostrados.

Guarda el archivo con tu primer apellido, separado por un guion bajo y tu nombre, seguido de la palabra modulo2. Ejemplo: “**Carrillo_Daniel_modulo3.R**”

1. Información inicial del código

Comienza siempre el script con la información inicial pertinente en forma de comentario

```
#Autor: Luis Daniel Carrillo Colín  
#Fecha: 12/02/2024  
#Módulo 1: Primeros pasos en Rstudio
```

2. Carga el paquete tidyverse

Recuerda usar la función “*p_load*” del paquete “*pacman*”.

```
pacman::p_load(tidyverse)
```

3. Carga el conjunto de datos mtcars. Muestra las primeras 6 filas

##	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
## Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
## Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
## Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
## Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
## Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2
## Valiant	18.1	6	225	105	2.76	3.460	20.22	1	0	3	1

4. Ejercicios de manipulación de datos

Función select()

4.1 Selecciona las columnas “mpg” y “hp” del conjunto de datos mtcars. Recuerda usar el operador pipe. **Muestra las primeras 6 filas**

```
##           mpg    hp
## Mazda RX4      21.0 110
## Mazda RX4 Wag  21.0 110
## Datsun 710     22.8  93
## Hornet 4 Drive  21.4 110
## Hornet Sportabout 18.7 175
## Valiant        18.1 105
```

4.2 Selecciona las columnas mpg, hasta hp. **Muestra las primeras 6 filas**

```
##           mpg cyl  disp  hp
## Mazda RX4      21.0   6 160.0 110
## Mazda RX4 Wag  21.0   6 160.0 110
## Datsun 710     22.8   4 108.0  93
## Hornet 4 Drive  21.4   6 258.0 110
## Hornet Sportabout 18.7   8 360.0 175
## Valiant        18.1   6 225.0 105
```

4.3 Excluir las columnas vs y am. **Muestra las primeras 6 filas**

```
##           mpg cyl  disp  hp drat    wt  qsec gear carb
## Mazda RX4      21.0   6 160.0 110 3.90 2.620 16.46   4    4
## Mazda RX4 Wag  21.0   6 160.0 110 3.90 2.875 17.02   4    4
## Datsun 710     22.8   4 108.0  93 3.85 2.320 18.61   4    1
## Hornet 4 Drive  21.4   6 258.0 110 3.08 3.215 19.44   3    1
## Hornet Sportabout 18.7   8 360.0 175 3.15 3.440 17.02   3    2
## Valiant        18.1   6 225.0 105 2.76 3.460 20.22   3    1
```

4.4 Seleccionar solo las columnas que contienen la letra “g”. **Muestra las primeras 6 filas**

```
##           mpg gear
## Mazda RX4      21.0   4
## Mazda RX4 Wag  21.0   4
## Datsun 710     22.8   4
## Hornet 4 Drive  21.4   3
## Hornet Sportabout 18.7   3
## Valiant        18.1   3
```

4.5 Seleccionar todas las columnas excepto *gear* y *carb*. Muestra las primeras 6 filas

```
##          mpg cyl  disp  hp drat   wt  qsec vs am
## Mazda RX4      21.0   6 160.0 110 3.90 2.620 16.46  0  1
## Mazda RX4 Wag  21.0   6 160.0 110 3.90 2.875 17.02  0  1
## Datsun 710     22.8   4 108.0  93 3.85 2.320 18.61  1  1
## Hornet 4 Drive  21.4   6 258.0 110 3.08 3.215 19.44  1  0
## Hornet Sportabout 18.7   8 360.0 175 3.15 3.440 17.02  0  0
## Valiant        18.1   6 225.0 105 2.76 3.460 20.22  1  0
```

5.Función filter()

5.1 Filtrar las filas donde el rendimiento (mpg) es mayor a 20.

```
##          mpg cyl  disp  hp drat   wt  qsec vs am gear carb
## Mazda RX4      21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag  21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710     22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive  21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
## Merc 240D      24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
## Merc 230       22.8   4 140.8  95 3.92 3.150 22.90  1  0    4    2
## Fiat 128       32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
## Honda Civic    30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
## Toyota Corolla 33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
## Toyota Corona  21.5   4 120.1  97 3.70 2.465 20.01  1  0    3    1
## Fiat X1-9      27.3   4  79.0  66 4.08 1.935 18.90  1  1    4    1
## Porsche 914-2  26.0   4 120.3  91 4.43 2.140 16.70  0  1    5    2
## Lotus Europa   30.4   4  95.1 113 3.77 1.513 16.90  1  1    5    2
## Volvo 142E     21.4   4 121.0 109 4.11 2.780 18.60  1  1    4    2
```

5.2 Filtrar las filas donde el número de cilindros (cyl) es igual a 6.

```
##          mpg cyl  disp  hp drat   wt  qsec vs am gear carb
## Mazda RX4      21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag  21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## Hornet 4 Drive  21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
## Valiant        18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
## Merc 280       19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
## Merc 280C      17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
## Ferrari Dino   19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
```

5.3 Filtrar las filas donde la potencia (hp) es menor a 100 y el peso (wt) es mayor a 2.5

```
##          mpg cyl  disp  hp drat   wt  qsec vs am gear carb
## Merc 240D      24.4   4 146.7  62 3.69 3.19 20.0  1  0    4    2
## Merc 230       22.8   4 140.8  95 3.92 3.15 22.9  1  0    4    2
```

5.4 Filtrar las filas donde el modelo (rownames(mtcars)) contiene la palabra “Merc”

Pista: Usa la función “grepl”

```
##      mpg cyl  disp  hp drat   wt  qsec vs am gear carb
## Merc 240D  24.4  4 146.7  62 3.69 3.19 20.0  1  0    4    2
## Merc 230   22.8  4 140.8  95 3.92 3.15 22.9  1  0    4    2
## Merc 280   19.2  6 167.6 123 3.92 3.44 18.3  1  0    4    4
## Merc 280C  17.8  6 167.6 123 3.92 3.44 18.9  1  0    4    4
## Merc 450SE  16.4  8 275.8 180 3.07 4.07 17.4  0  0    3    3
## Merc 450SL  17.3  8 275.8 180 3.07 3.73 17.6  0  0    3    3
## Merc 450SLC 15.2  8 275.8 180 3.07 3.78 18.0  0  0    3    3
```

5.5. Filtrar las filas donde la transmisión (am) es automática (0). **Muestra las primeras 6 filas**

```
##      mpg cyl  disp  hp drat   wt  qsec vs am gear carb
## Hornet 4 Drive  21.4  6 258.0 110 3.08 3.215 19.44  1  0    3
## Hornet Sportabout 18.7  8 360.0 175 3.15 3.440 17.02  0  0    3
## Valiant 18.1  6 225.0 105 2.76 3.460 20.22  1  0    3
## Duster 360 14.3  8 360.0 245 3.21 3.570 15.84  0  0    3
## Merc 240D 24.4  4 146.7  62 3.69 3.190 20.00  1  0    4
## Merc 230 22.8  4 140.8  95 3.92 3.150 22.90  1  0    4
```

6. Función mutate()

6.1. Crea una nueva columna (disp_per_cyl) que sea el cociente entre la cilindrada (disp) y el número de cilindros (cyl). **Muestra las primeras 6 filas**

```
##      cyl disp  hp drat vs am gear carb disp_per_cyl
## Mazda RX4      6  160 110 3.90  0  1    4    4      26.66667
## Mazda RX4 Wag   6  160 110 3.90  0  1    4    4      26.66667
## Datsun 710      4  108  93 3.85  1  1    4    1      27.00000
## Hornet 4 Drive   6  258 110 3.08  1  0    3    1      43.00000
## Hornet Sportabout 8  360 175 3.15  0  0    3    2      45.00000
## Valiant         6  225 105 2.76  1  0    3    1      37.50000
```

6.2. Crear una nueva columna “power_to_weight” que sea el cociente entre la potencia (hp) y el peso (wt). **Muestra las primeras 6 filas**

##		cyl	hp	wt	qsec	vs	am	gear	carb	power_to_weight
##	Mazda RX4	6	110	2.620	16.46	0	1	4	4	41.98473
##	Mazda RX4 Wag	6	110	2.875	17.02	0	1	4	4	38.26087
##	Datsun 710	4	93	2.320	18.61	1	1	4	1	40.08621
##	Hornet 4 Drive	6	110	3.215	19.44	1	0	3	1	34.21462
##	Hornet Sportabout	8	175	3.440	17.02	0	0	3	2	50.87209
##	Valiant	6	105	3.460	20.22	1	0	3	1	30.34682

6.3. Crear una columna “km_l” que represente el rendimiento en km por litro usando la columna “mpg”. Es decir, multiplícalo por 0.425144. **Muestra las primeras 6 filas**

##		mpg	cyl	qsec	vs	am	gear	carb	km_l
##	Mazda RX4	21.0	6	16.46	0	1	4	4	8.928024
##	Mazda RX4 Wag	21.0	6	17.02	0	1	4	4	8.928024
##	Datsun 710	22.8	4	18.61	1	1	4	1	9.693283
##	Hornet 4 Drive	21.4	6	19.44	1	0	3	1	9.098082
##	Hornet Sportabout	18.7	8	17.02	0	0	3	2	7.950193
##	Valiant	18.1	6	20.22	1	0	3	1	7.695106

6.4. Crear una nueva columna “wt_mpg” que sea el producto de “mpg” y “wt”. **Muestra las primeras 6 filas**

##		mpg	cyl	wt	qsec	vs	am	gear	carb	wt_mpg
##	Mazda RX4	21.0	6	2.620	16.46	0	1	4	4	55.020
##	Mazda RX4 Wag	21.0	6	2.875	17.02	0	1	4	4	60.375
##	Datsun 710	22.8	4	2.320	18.61	1	1	4	1	52.896
##	Hornet 4 Drive	21.4	6	3.215	19.44	1	0	3	1	68.801
##	Hornet Sportabout	18.7	8	3.440	17.02	0	0	3	2	64.328
##	Valiant	18.1	6	3.460	20.22	1	0	3	1	62.626

6.5. Crea una nueva columna “high_power” que sea TRUE si la potencia (hp) es mayor a 150 y FALSE si es lo contrario. **Muestra las primeras 6 filas**

##		hp	drat	wt	qsec	vs	am	gear	carb	high_power
##	Mazda RX4	110	3.90	2.620	16.46	0	1	4	4	FALSE
##	Mazda RX4 Wag	110	3.90	2.875	17.02	0	1	4	4	FALSE
##	Datsun 710	93	3.85	2.320	18.61	1	1	4	1	FALSE
##	Hornet 4 Drive	110	3.08	3.215	19.44	1	0	3	1	FALSE
##	Hornet Sportabout	175	3.15	3.440	17.02	0	0	3	2	TRUE
##	Valiant	105	2.76	3.460	20.22	1	0	3	1	FALSE

7.Funciones “group_by()” y “summarise()”

7.1 Agrupa por el número de cilindros (cyl) y calcular el promedio del rendimiento (mpg) para cada grupo

```
mtcars_group_summarise1 <- mtcars %>%  
  group_by(cyl) %>%  
  summarise(avg_mpg = mean(mpg))
```

```
mtcars_group_summarise1
```

```
## # A tibble: 3 × 2  
##   cyl avg_mpg  
##   <dbl>   <dbl>  
## 1     4    26.7  
## 2     6    19.7  
## 3     8    15.1
```

7.2. Agrupa por la transmisión (am) y contar el número de modelos en cada categoría

```
## # A tibble: 2 × 2  
##   am count  
##   <dbl> <int>  
## 1     0    19  
## 2     1    13
```

7.3 Agrupar por el número de velocidades (gear) y encontrar la potencia (hp) máxima en cada grupo

```
## # A tibble: 3 × 2  
##   gear max_hp  
##   <dbl>   <dbl>  
## 1     3    245  
## 2     4    123  
## 3     5    335
```

7.4. Agrupar por la cantidad de cilindros (cyl) y calcular la mediana del peso (wt) para cada grupo

```
## # A tibble: 3 × 2  
##   cyl median_wt  
##   <dbl>   <dbl>  
## 1     4     2.2  
## 2     6     3.22  
## 3     8     3.76
```

7.5. Agrupar por la presencia de marchas manuales (`am==1`) y calcular el promedio del rendimiento y la desviación estándar del peso (`wt`) para cada grupo.

```
## # A tibble: 2 × 3
##   manual_transmission avg_mpg sd_wt
##   <chr>              <dbl> <dbl>
## 1 Automatic          17.1 0.777
## 2 Manual             24.4 0.617
```