

# practica\_dplyr

Sandra Rairán

2/16/2021

## Usando el conjunto de datos starwars mediante la libreria dplyr

```
library(dplyr)
```

select

```
starwars %>%  
  select(name, ends_with("color"))
```

```
## # A tibble: 87 x 4  
##   name          hair_color skin_color eye_color  
##   <chr>         <chr>      <chr>    <chr>  
## 1 Luke Skywalker blond       fair     blue  
## 2 C-3P0         <NA>       gold     yellow  
## 3 R2-D2         <NA>       white, blue red  
## 4 Darth Vader   none       white     yellow  
## 5 Leia Organa   brown     light     brown  
## 6 Owen Lars     brown, grey light     blue  
## 7 Beru Whitesun lars brown     light     blue  
## 8 R5-D4         <NA>       white, red red  
## 9 Biggs Darklighter black      light     brown  
## 10 Obi-Wan Kenobi auburn, white fair     blue-gray  
## # ... with 77 more rows
```

```
## # A tibble: 87 x 3  
##   name          height mass  
##   <chr>         <int> <dbl>  
## 1 Luke Skywalker    172    77  
## 2 C-3P0             167    75  
## 3 R2-D2             96    32  
## 4 Darth Vader      202   136  
## 5 Leia Organa      150    49  
## 6 Owen Lars        178   120  
## 7 Beru Whitesun lars 165    75  
## 8 R5-D4             97    32  
## 9 Biggs Darklighter 183    84  
## 10 Obi-Wan Kenobi   182    77  
## # ... with 77 more rows
```

filter

```
starwars %>%
  filter(species == "Droid")
```

```
## # A tibble: 6 x 14
##   name height mass hair_color skin_color eye_color birth_year sex gender
##   <chr> <int> <dbl> <chr>      <chr>      <chr>      <dbl> <chr> <chr>
## 1 C-3P0  167    75 <NA>      gold        yellow        112 none mascu~
## 2 R2-D2   96    32 <NA>      white, bl~ red          33 none mascu~
## 3 R5-D4   97    32 <NA>      white, red red          NA none mascu~
## 4 IG-88  200   140 none      metal        red          15 none mascu~
## 5 R4-P~   96    NA none      silver, r~ red, blue    NA none femin~
## 6 BB8     NA    NA none      none         black         NA none mascu~
## # ... with 5 more variables: homeworld <chr>, species <chr>, films <list>,
## #   vehicles <list>, starships <list>
```

arrange

```
starwars %>%
  arrange(desc(birth_year))
```

```
## # A tibble: 87 x 14
##   name height mass hair_color skin_color eye_color birth_year sex gender
##   <chr> <int> <dbl> <chr>      <chr>      <chr>      <dbl> <chr> <chr>
## 1 Yoda    66    17 white      green      brown        896 male mascu~
## 2 Jabb~   175  1358 <NA>      green-tan~ orange        600 herm~ mascu~
## 3 Chew~   228   112 brown      unknown    blue         200 male mascu~
## 4 C-3P0   167    75 <NA>      gold        yellow        112 none mascu~
## 5 Dooku   193    80 white      fair        brown        102 male mascu~
## 6 Qui~-   193    89 brown      fair        blue          92 male mascu~
## 7 Ki-A~   198    82 white      pale        yellow         92 male mascu~
## 8 Fini~   170    NA blond      fair        blue          91 male mascu~
## 9 Palp~   170    75 grey      pale        yellow         82 male mascu~
## 10 Clie~  183    NA brown      fair        blue          82 male mascu~
## # ... with 77 more rows, and 5 more variables: homeworld <chr>, species <chr>,
## #   films <list>, vehicles <list>, starships <list>
```

mutate (crear columnas)

```
starwars <-starwars %>% mutate(height_m = height / 100)
colnames(starwars)
```

```
## [1] "name"      "height"    "mass"      "hair_color" "skin_color"
## [6] "eye_color" "birth_year" "sex"       "gender"     "homeworld"
## [11] "species"   "films"     "vehicles"  "starships"  "height_m"
```

group\_by y summarise

```
starwars %>%
  group_by(species) %>%
  summarise(
    n = n(),
    mass = mean(mass, na.rm = TRUE)
  ) %>%
  arrange(desc(n))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)

## # A tibble: 38 x 3
##   species      n mass
##   <chr>    <int> <dbl>
## 1 Human        35  82.8
## 2 Droid         6  69.8
## 3 <NA>          4   48
## 4 Gungan        3   74
## 5 Kaminoan       2   88
## 6 Mirialan       2  53.1
## 7 Twi'lek        2   55
## 8 Wookiee        2  124
## 9 Zabrak         2   80
## 10 Aleena        1   15
## # ... with 28 more rows
```

Contar el numero de valores perdidos

```
sum(is.na(starwars$mass))
```

```
## [1] 28
```

slice\_max (seleccione filas con los valores más altos)

```
starwars %>% slice_max(height, n = 5)
```

```
## # A tibble: 5 x 15
##   name height mass hair_color skin_color eye_color birth_year sex gender
##   <chr>   <int> <dbl> <chr>      <chr>      <chr>      <dbl> <chr> <chr>
## 1 Yara~   264    NA none       white       yellow        NA male mascul~
## 2 Tarf~   234   136 brown      brown       blue        NA male mascul~
## 3 Lama~   229    88 none       grey        black        NA male mascul~
## 4 Chew~   228   112 brown      unknown     blue       200 male mascul~
## 5 Roos~   224    82 none       grey        orange        NA male mascul~
## # ... with 6 more variables: homeworld <chr>, species <chr>, films <list>,
## #   vehicles <list>, starships <list>, height_m <dbl>
```

across (agrupa por genero y calcula el promedio de todas las columnas numericas)

```
starwars %>%
  group_by(gender) %>%
  filter(n() > 1) %>%
  summarise(across(where(is.numeric), ~ mean(.x, na.rm = TRUE)))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)

## # A tibble: 3 x 5
##   gender    height mass birth_year height_m
##   <chr>    <dbl> <dbl>    <dbl>    <dbl>
## 1 feminine  165.  54.7    47.2    1.65
## 2 masculine 177. 106.    97.8    1.77
## 3 <NA>     181.  48     62     1.81
```

distinct y across

```
distinct(starwars, across(contains("color")))
```

```
## # A tibble: 67 x 3
##   hair_color    skin_color eye_color
##   <chr>         <chr>      <chr>
## 1 blond        fair        blue
## 2 <NA>         gold        yellow
## 3 <NA>         white, blue red
## 4 none         white        yellow
## 5 brown        light        brown
## 6 brown, grey  light        blue
## 7 brown        light        blue
## 8 <NA>         white, red  red
## 9 black        light        brown
## 10 auburn, white fair        blue-gray
## # ... with 57 more rows
```

pull extraer variables

```
pull(starwars, name, birth_year)
```

```
##           19           112           33
##   "Luke Skywalker"   "C-3PO"   "R2-D2"
##           41.9           19           52
##   "Darth Vader"     "Leia Organa" "Owen Lars"
##           47           <NA>           24
##   "Beru Whitesun lars" "R5-D4" "Biggs Darklighter"
##           57           41.9           64
##   "Obi-Wan Kenobi"   "Anakin Skywalker" "Wilhuff Tarkin"
##           200           29           44
##   "Chewbacca"       "Han Solo"   "Greedo"
##           600           21           <NA>
##   "Jabba Desilijic Tiure" "Wedge Antilles" "Jek Tono Porkins"
##           896           82           31.5
##   "Yoda"             "Palpatine"   "Boba Fett"
##           15           53           31
##   "IG-88"            "Bossk"     "Lando Calrissian"
##           37           41           48
##   "Lobot"            "Ackbar"     "Mon Mothma"
##           <NA>           8           <NA>
##   "Arvel Crynyd" "Wicket Systri Warrick" "Nien Nunb"
##           92           <NA>           91
##   "Qui-Gon Jinn"   "Nute Gunray" "Finis Valorum"
##           52           <NA>           <NA>
##   "Jar Jar Binks"  "Roos Tarpals" "Rugor Nass"
##           <NA>           <NA>           <NA>
##   "Ric Olié"       "Watto"     "Sebulba"
##           62           72           54
##   "Quarsh Panaka"  "Shmi Skywalker" "Darth Maul"
##           <NA>           48           <NA>
##   "Bib Fortuna"    "Ayla Secura"   "Dud Bolt"
##           <NA>           <NA>           72
```

##	"Gasgano"	"Ben Quadinaros"	"Mace Windu"
##	92	<NA>	<NA>
##	"Ki-Adi-Mundi"	"Kit Fisto"	"Eeth Koth"
##	<NA>	<NA>	<NA>
##	"Adi Gallia"	"Saesee Tiin"	"Yarael Poof"
##	22	<NA>	<NA>
##	"Plo Koon"	"Mas Amedda"	"Gregar Typho"
##	<NA>	82	<NA>
##	"Cordé"	"Cliegg Lars"	"Poggle the Lesser"
##	58	40	<NA>
##	"Luminara Unduli"	"Barriss Offee"	"Dormé"
##	102	67	66
##	"Dooku"	"Bail Prestor Organa"	"Jango Fett"
##	<NA>	<NA>	<NA>
##	"Zam Wesell"	"Dexter Jettster"	"Lama Su"
##	<NA>	<NA>	<NA>
##	"Taun We"	"Jocasta Nu"	"Ratts Tyerell"
##	<NA>	<NA>	<NA>
##	"R4-P17"	"Wat Tambor"	"San Hill"
##	<NA>	<NA>	<NA>
##	"Shaak Ti"	"Grievous"	"Tarfful"
##	<NA>	<NA>	<NA>
##	"Raymus Antilles"	"Sly Moore"	"Tion Medon"
##	<NA>	<NA>	<NA>
##	"Finn"	"Rey"	"Poe Dameron"
##	<NA>	<NA>	46
##	"BB8"	"Captain Phasma"	"Padmé Amidala"