FUNÇÃO ARRANGE E SLICE NO R

Leonardo Machado

2022-06-27

###Funções###

As funções, arrange() e slice(), permitem, respectivamente, ordenar e selecionar dados de um data frame.

## Arrange()

starwars %>%   
 arrange(height)

## # A tibble: 87 × 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 Ratts T… 79 15 none grey, blue unknown NA male mascu…  
## 3 Wicket … 88 20 brown brown brown 8 male mascu…  
## 4 Dud Bolt 94 45 none blue, grey yellow NA male mascu…  
## 5 R2-D2 96 32 <NA> white, bl… red 33 none mascu…  
## 6 R4-P17 96 NA none silver, r… red, blue NA none femin…  
## 7 R5-D4 97 32 <NA> white, red red NA none mascu…  
## 8 Sebulba 112 40 none grey, red orange NA male mascu…  
## 9 Gasgano 122 NA none white, bl… black NA male mascu…  
## 10 Watto 137 NA black blue, grey yellow NA male mascu…  
## # … with 77 more rows, and 5 more variables: homeworld <chr>, species <chr>,  
## # films <list>, vehicles <list>, starships <list>

O comando acima resulta em uma saída em ordem crescente, como deaful. Para saída em ordem decrescente , temos duas formas:

starwars %>%   
 arrange(- height)

## # A tibble: 87 × 14  
## name height mass hair\_color skin\_color eye\_color birth\_year sex gender  
## <chr> <int> <dbl> <chr> <chr> <chr> <dbl> <chr> <chr>   
## 1 Yarael … 264 NA none white yellow NA male mascu…  
## 2 Tarfful 234 136 brown brown blue NA male mascu…  
## 3 Lama Su 229 88 none grey black NA male mascu…  
## 4 Chewbac… 228 112 brown unknown blue 200 male mascu…  
## 5 Roos Ta… 224 82 none grey orange NA male mascu…  
## 6 Grievous 216 159 none brown, wh… green, y… NA male mascu…  
## 7 Taun We 213 NA none grey black NA fema… femin…  
## 8 Rugor N… 206 NA none green orange NA male mascu…  
## 9 Tion Me… 206 80 none grey black NA male mascu…  
## 10 Darth V… 202 136 none white yellow 41.9 male mascu…  
## # … with 77 more rows, and 5 more variables: homeworld <chr>, species <chr>,  
## # films <list>, vehicles <list>, starships <list>

Ou

starwars %>%   
 arrange(desc(height))

## # A tibble: 87 × 14  
## name height mass hair\_color skin\_color eye\_color birth\_year sex gender  
## <chr> <int> <dbl> <chr> <chr> <chr> <dbl> <chr> <chr>   
## 1 Yarael … 264 NA none white yellow NA male mascu…  
## 2 Tarfful 234 136 brown brown blue NA male mascu…  
## 3 Lama Su 229 88 none grey black NA male mascu…  
## 4 Chewbac… 228 112 brown unknown blue 200 male mascu…  
## 5 Roos Ta… 224 82 none grey orange NA male mascu…  
## 6 Grievous 216 159 none brown, wh… green, y… NA male mascu…  
## 7 Taun We 213 NA none grey black NA fema… femin…  
## 8 Rugor N… 206 NA none green orange NA male mascu…  
## 9 Tion Me… 206 80 none grey black NA male mascu…  
## 10 Darth V… 202 136 none white yellow 41.9 male mascu…  
## # … with 77 more rows, and 5 more variables: homeworld <chr>, species <chr>,  
## # films <list>, vehicles <list>, starships <list>

Ordenando múltiplas colunas

starwars %>%   
 arrange(hair\_color,height)

## # A tibble: 87 × 14  
## name height mass hair\_color skin\_color eye\_color birth\_year sex gender  
## <chr> <int> <dbl> <chr> <chr> <chr> <dbl> <chr> <chr>   
## 1 Mon Mot… 150 NA auburn fair blue 48 fema… femin…  
## 2 Wilhuff… 180 NA auburn, g… fair blue 64 male mascu…  
## 3 Obi-Wan… 182 77 auburn, w… fair blue-gray 57 male mascu…  
## 4 Watto 137 NA black blue, grey yellow NA male mascu…  
## 5 Shmi Sk… 163 NA black fair brown 72 fema… femin…  
## 6 Barriss… 166 50 black yellow blue 40 fema… femin…  
## 7 Luminar… 170 56.2 black yellow blue 58 fema… femin…  
## 8 Eeth Ko… 171 NA black brown brown NA male mascu…  
## 9 Lando C… 177 79 black dark brown 31 male mascu…  
## 10 Biggs D… 183 84 black light brown 24 male mascu…  
## # … with 77 more rows, and 5 more variables: homeworld <chr>, species <chr>,  
## # films <list>, vehicles <list>, starships <list>

Atenção que a ordem das colunas fazem diferença. No caso acima a saída será ordenado primeiramente a primeira coluna e dentro de cada classe dessa ordem a oredm dos pesos

## slice

No primeiro caso utilizaremos o comando slice.head para encontrarmos os três personagens mais baixos. Veja , primeiramente organizo os dados em ordem crescente, assim os três primeiros serão o desejado.

starwars %>%   
 arrange(height) %>%  
 slice\_head(n=3)

## # A tibble: 3 × 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 Ratts Ty… 79 15 none grey, blue unknown NA male mascu…  
## 3 Wicket S… 88 20 brown brown brown 8 male mascu…  
## # … with 5 more variables: homeworld <chr>, species <chr>, films <list>,  
## # vehicles <list>, starships <list>

O comando slice.head necessita definir o parâmetro n, o número de observação desejada.

Agora, os três mais altos

Duas maneiras possível: ordenando de forma decrescente e utilizano o slice.head, ou mantendo o ordenamento crescente e utilizando o slice.tail(n = ?)

starwars %>%   
 arrange(-height) %>%   
 slice\_head(n=3)

## # A tibble: 3 × 14  
## name height mass hair\_color skin\_color eye\_color birth\_year sex gender  
## <chr> <int> <dbl> <chr> <chr> <chr> <dbl> <chr> <chr>   
## 1 Yarael P… 264 NA none white yellow NA male mascu…  
## 2 Tarfful 234 136 brown brown blue NA male mascu…  
## 3 Lama Su 229 88 none grey black NA male mascu…  
## # … with 5 more variables: homeworld <chr>, species <chr>, films <list>,  
## # vehicles <list>, starships <list>

De outra forma:

starwars %>%   
 arrange(height) %>%   
 slice\_tail(n=3)

## # A tibble: 3 × 14  
## name height mass hair\_color skin\_color eye\_color birth\_year sex gender  
## <chr> <int> <dbl> <chr> <chr> <chr> <dbl> <chr> <chr>   
## 1 Poe Dame… NA NA brown light brown NA male mascu…  
## 2 BB8 NA NA none none black NA none mascu…  
## 3 Captain … NA NA unknown unknown unknown NA <NA> <NA>   
## # … with 5 more variables: homeworld <chr>, species <chr>, films <list>,  
## # vehicles <list>, starships <list>

## slice.sample(n=?)

O comando acima permite selecionar uma amostra aleatória do data frame, operação muito utilizada para retirar-se amostra controle X amostra teste.

starwars %>%   
 slice\_sample(n=10)

## # A tibble: 10 × 14  
## name height mass hair\_color skin\_color eye\_color birth\_year sex gender  
## <chr> <int> <dbl> <chr> <chr> <chr> <dbl> <chr> <chr>   
## 1 Qui-Gon… 193 89 brown fair blue 92 male mascu…  
## 2 Luke Sk… 172 77 blond fair blue 19 male mascu…  
## 3 Leia Or… 150 49 brown light brown 19 fema… femin…  
## 4 Raymus … 188 79 brown light brown NA male mascu…  
## 5 Eeth Ko… 171 NA black brown brown NA male mascu…  
## 6 Lando C… 177 79 black dark brown 31 male mascu…  
## 7 R5-D4 97 32 <NA> white, red red NA none mascu…  
## 8 Finis V… 170 NA blond fair blue 91 male mascu…  
## 9 Quarsh … 183 NA black dark brown 62 <NA> <NA>   
## 10 Anakin … 188 84 blond fair blue 41.9 male mascu…  
## # … with 5 more variables: homeworld <chr>, species <chr>, films <list>,  
## # vehicles <list>, starships <list>