

Manipulação e Apresentação de Dados - Exercícios

Guilherme Bovi Ambrosano

Exercícios do dplyr

Exercício 1



Usando o conjunto de dados `agridat::walsh.cottonprice`:

- ▶ Obter um conjunto de dados com as colunas `year` (ano), `cotton` (cents por libra de algodão), `cottonseed` (dólares por toneladas de sementes) e `combined` ($\text{cotton} + 1,857 \times \text{cottonseed}$, em cents por libra)
- ▶ Sabendo que 1 ton = 2000 pounds e que 1 dólar = 100 cents, converter `cottonseed` para a mesma unidade de `cotton`
- ▶ Criar a coluna `combined2`, com a mesma fórmula da coluna `combined`
- ▶ Criar a coluna `década` a partir da coluna `year`,

##		year	cotton	cottonseed	combined	combined2	década
## 1	1910	13.52	1.2075	15.76	15.762327	1910	
## 2	1911	13.96	1.2995	16.37	16.373172	1910	
## 3	1912	9.65	0.8575	11.24	11.242377	1910	
## 4	1913	11.50	0.9165	13.20	13.201940	1910	
## 5	1914	12.47	1.0950	14.50	14.503415	1910	
## 6	1915	7.35	0.7730	8.79	8.785461	1910	
## 7	1916	11.22	1.5065	14.02	14.017571	1910	

Exercício 1 (continuação)



Usando o conjunto de dados `agridat::walsh.cottonprice`:

- Obter a média e a mediana da coluna `combined2` para cada década

```
## # A tibble: 4 x 3
##   década média mediana
##   <chr>   <dbl>   <dbl>
## 1 1910    18.3    15.1
## 2 1920    24.5    23.0
## 3 1930    12.2    11.4
## 4 1940    16.9    16.7
```

Exercício 2



Usando o conjunto de dados `dplyr::starwars`:

- Obter um vetor com o número de filmes em que cada personagem aparece

##	Luke Skywalker	C-3PO	R
##	5	6	
##	Darth Vader	Leia Organa	Owen
##	4	5	
##	Beru Whitesun Lars	R5-D4	Biggs Darklighter
##	3	1	
##	Obi-Wan Kenobi	Anakin Skywalker	Wilhuff Tarkin
##	6	3	
##	Chewbacca	Han Solo	Gr
##	5	4	
##	Jabba Desilijic Tiure	Wedge Antilles	Jek Tono Por
##	3	3	
##	Yoda	Palpatine	Boba
##	5	5	
##	IG-88	Bossk	Lando Calrissian

Exercício 2 (continuação)



Usando o conjunto de dados `dplyr::starwars`:

- Criar uma `tibble` nova, em que se repitam nas linhas os nomes dos personagens tantas vezes quanto o número de filmes em que aparecem

```
## # A tibble: 173 x 1
##   name
##   <chr>
## 1 Luke Skywalker
## 2 Luke Skywalker
## 3 Luke Skywalker
## 4 Luke Skywalker
## 5 Luke Skywalker
## 6 C-3P0
## 7 C-3P0
## 8 C-3P0
## 9 C-3P0
## 10 C-3P0
## # i 163 more rows
```

Exercício 2 (continuação)



Usando o conjunto de dados `dplyr::starwars`:

- Adicionar na tibble criada uma coluna com os nomes dos filmes e um id por filme de cada personagem

```
## # A tibble: 173 x 3
##   name                id films
##   <chr>              <int> <chr>
## 1 Luke Skywalker     1 A New Hope
## 2 Luke Skywalker     2 The Empire Strikes Back
## 3 Luke Skywalker     3 Return of the Jedi
## 4 Luke Skywalker     4 Revenge of the Sith
## 5 Luke Skywalker     5 The Force Awakens
## 6 C-3P0              1 A New Hope
## 7 C-3P0              2 The Empire Strikes Back
## 8 C-3P0              3 Return of the Jedi
## 9 C-3P0              4 The Phantom Menace
## 10 C-3P0             5 Attack of the Clones
## # i 163 more rows
```

Exercícios

Exercício 1



Transformar os objetos criados abaixo na estrutura a seguir.

```
pValores <- tibble(  
  contraste = c("Trat 1 - Trat 2", "Trat 1 - Trat 3",  
                "Trat 2 - Trat 3"),  
  `p-valor` = c(0.9999, 0.050, 0.0001)  
)
```

```
vetor.ex1
```

```
## Trat 1-Trat 2 Trat 1-Trat 3 Trat 2-Trat 3  
##          0.9999          0.0500          0.0001
```

```
str(vetor.ex1)
```

```
##   Named num [1:3] 1e+00 5e-02 1e-04  
##   - attr(*, "names")= chr [1:3] "Trat 1-Trat 2" "Trat 1-Trat 3"
```

Exercício 2



Usando as funções do tidyverse e partindo do data-frame dados presente no arquivo Dados.RData, crie a tabela abaixo:

```
## # A tibble: 4 x 3
##   Trat  média desvio
##   <chr> <dbl>  <dbl>
## 1 1      15.5   0.521
## 2 2      17.0   0.534
## 3 3      18.2   0.395
## 4 4      21.0   0.906
```

Exercício 3



Usando o conjunto de dados presente no site
<http://www.leb.esalq.usp.br/leb/exceldados/DCE2023.TXT>, obter os
dias mais quentes do ano:

##	2023-11-14	2023-11-17	2023-09-25	2023-11-12	2023-11-13
##	30.4	30.2	30.1	30.0	29.9

Exercício 4



Usando o conjunto de dados presente no site <http://www.leb.esalq.usp.br/leb/exceldados/DCE2023.TXT>, obter os dias com precipitação entre 30mm e 40mm:

##	2023-01-29	2023-01-31	2023-03-20	2023-06-14	2023-11-23
##	38.6	37.3	33.5	32.3	31.2

Exercício 5



Transformar os objetos criados abaixo na estrutura a seguir, usando os pacotes do tidyverse.

```
tempo1 <- c(15, 14, 17, 14, 17, 13)
tempo2 <- c(14, 13, 16, 14, 12, 16)
tempo3 <- c(16, 14, 15, 12, 17, 15)
```

```
## # A tibble: 3 x 2
##   Tempo      Média
##   <fct>      <dbl>
## 1 1 semana    15
## 2 2 semanas  14.2
## 3 3 semanas  14.8
```

Exercício 6



Obter a tabela abaixo partindo do conjunto de dados `mtcars`, usando os pacotes do `tidyverse`.

Carros cujos nomes começam com M:

##	carro	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	ge
## 1	Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	
## 2	Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	
## 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	
## 5	Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	
## 6	Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	
## 7	Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	
## 8	Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	
## 9	Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	
## 10	Maserati Bora	15.0	8	301.0	335	3.54	3.570	14.60	0	1	

Exercício 7



Obter a tabela abaixo partindo do conjunto de dados `mtcars`, usando os pacotes do `tidyverse`.

Carros cujos nomes terminam em números:

##		carro	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	ca
## 1	Datsun	710	22.8	4	108.0	93	3.85	2.32	18.61	1	1	4	
## 2	Duster	360	14.3	8	360.0	245	3.21	3.57	15.84	0	0	3	
## 3	Merc	230	22.8	4	140.8	95	3.92	3.15	22.90	1	0	4	
## 4	Merc	280	19.2	6	167.6	123	3.92	3.44	18.30	1	0	4	
## 5	Fiat	128	32.4	4	78.7	66	4.08	2.20	19.47	1	1	4	

Exercício 8



Usando o conjunto de dados `archbold.apple` do pacote `agridat`, obter uma tabela como a Tabela 4 de Archbold, Brown, Cornelius (1987).

<https://journals.ashs.org/jashs/view/journals/jashs/112/2/article-p219.xml>

```
library(agridat)
```

```
## # A tibble: 7 x 4
##   spacing stock      Golden Redspur
##   <fct>    <fct>    <dbl>    <dbl>
## 1 1.8 m    <NA>        121.     120.
## 2 3.0m     <NA>        151.     160.
## 3 4.3 m    <NA>        165.     164.
## 4 <NA>     M0007        153.     131.
## 5 <NA>     MM106        153.     188.
## 6 <NA>     MM111        114.     146.
## 7 <NA>     Seedling      159.     119.
```


Exercício 9



Obter a tabela abaixo partindo do conjunto de dados iris, usando os pacotes do tidyverse.

```
## # A tibble: 8 x 3
##   Species      name      value
##   <chr>      <chr>      <dbl>
## 1 Iris setosa Petal.Length_média 1.46
## 2 Iris setosa Petal.Width_média 0.246
## 3 Iris setosa Petal.Length_desvio 0.174
## 4 Iris setosa Petal.Width_desvio 0.105
## 5 Iris versicolor Petal.Length_média 4.26
## 6 Iris versicolor Petal.Width_média 1.33
## 7 Iris versicolor Petal.Length_desvio 0.470
## 8 Iris versicolor Petal.Width_desvio 0.198
```

Exercício 10



Obter a estrutura abaixo partindo do conjunto de dados iris, usando os pacotes do tidyverse.

```
lista.ex10
```

```
## $setosa
## Mediana (Petal.Length) Mediana (Petal.Width) Mediana (Sepal.Length)
##           1.5           0.2           5.0
## Mediana (Sepal.Width)
##           3.4
##
## $versicolor
## Mediana (Petal.Length) Mediana (Petal.Width) Mediana (Sepal.Length)
##           4.35           1.30           5.90
## Mediana (Sepal.Width)
##           2.80
```

```
str(lista.ex10)
```

```
## List of 2
## $ setosa : Named num [1:4] 1.5 0.2 5 3.4
## ..- attr(*, "names")= chr [1:4] "Mediana (Petal.Length)" "Mediana (Petal.W
## $ versicolor: Named num [1:4] 4.35 1.3 5.9 2.8
## ..- attr(*, "names")= chr [1:4] "Mediana (Petal.Length)" "Mediana (Petal.W
```