

# Oficina CIF



**Visualização de dados em Farmacologia: tipos de gráficos, como plotá-los e boas práticas para publicações**

Tamires Martins

## Índice

<b>1</b>	<b>Introdução</b>	<b>1</b>
<b>2</b>	<b>Passos básicos com R</b>	<b>1</b>
2.1	Como se ajudar . . . . .	1
<b>3</b>	<b>Como rodar o código?</b>	<b>2</b>
3.1	A ordem importa (rode o código na sequência): . . . . .	2
<b>4</b>	<b>Os dados</b>	<b>3</b>
4.1	Carregando pacotes . . . . .	4
4.2	Importando os dados . . . . .	4
4.3	Visualizando os dados . . . . .	5
<b>5</b>	<b>Os gráficos</b>	<b>8</b>
5.1	Gráfico box plot simples usando ggplot2 . . . . .	8
5.2	Gráfico box plot simples usando tidyplots . . . . .	9
<b>6</b>	<b>Agora é a sua vez:</b>	<b>10</b>

# 1 Introdução

Nesta prática, vamos explorar como visualizar dados da Farmacologia em R. Para isso, também vamos aprender passos básicos com a linguagem.

## 2 Passos básicos com R

### 2.1 Como se ajudar

Abrir a documentação de uma função é essencial para entender como ela funciona e quais parâmetros você pode usar. Existem várias maneiras de acessar a documentação no R:

```
1 ?mean  
2 # ou  
3 help(mean)  
4 # ou clicar no nome da função apertar F1
```

## 3 Como rodar o código?

Coloque o cursor na linha que deseja rodar e aperte “Run” ou “Ctrl” + “Enter”. O resultado irá aparecer no console, faça o teste:

```
1 print("Olá, mundo!") # Imprimindo uma mensagem no console  
  
[1] "Olá, mundo!"  
  
1 1 + 1  
  
[1] 2
```

### 3.1 A ordem importa (rode o código na sequência):

```
1 # tire a anotação da frente e tente rodar:  
2 # numero  
3  
4 # Atribuindo um valor a um objeto  
5 numero <- 10  
6  
7 numero
```

```
[1] 10
```

Para acessar a classe de um objeto:

```
1 class(numero)
```

```
[1] "numeric"
```

Os pacotes são conjuntos de funções. Para instalar um pacote, use `install.packages("nome_do_pacote")`. Para carregar um pacote já instalado, use `library(nome_do_pacote)`. Algumas funções são da base:

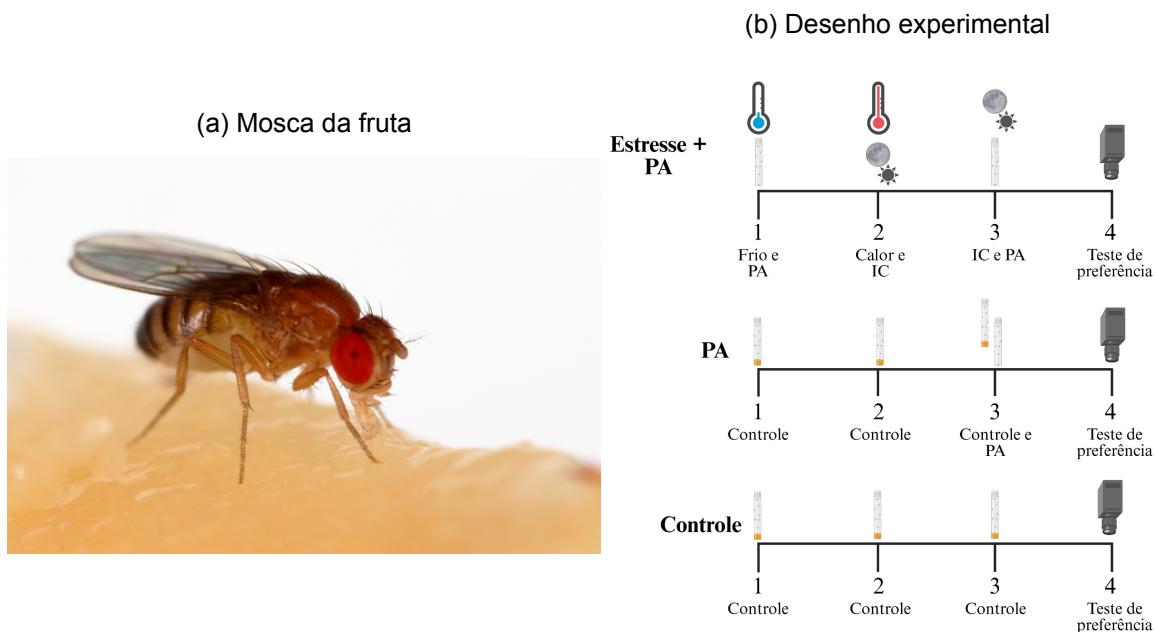
```
1 # mean() calcula a média  
2  
3 # A média de um vetor  
4 meu_vetor <- c(1, 2, 3, 4, 5)  
5 mean(meu_vetor)
```

```
[1] 3
```

## 4 Os dados

Os dados que iremos trabalhar são simulações baseadas em dados reais (Eckert et al. (não publicado)). O experimento consiste em avaliar comportamentos das moscas-das-frutas (*Drosophila melanogaster*) machos e fêmeas, em diferentes condições de tratamento.

Figura 1: Experimento com *Drosophila melanogaster*



#### 4.1 Carregando pacotes

```
1 # seguido de texto são anotações, não código.  
2  
3 # Instale se ainda não tiver  
4 # install.packages("ggplot2")  
5 # install.packages("tidyplots")  
6 # install.packages("patchwork")  
7  
8 library(ggplot2)  
9 library(tidyplots)  
10 library(patchwork)
```

#### 4.2 Importando os dados

Os dados estão disponíveis no arquivo `dados_moscas.xlsx`, que deve estar na mesma pasta do arquivo `.qmd`.

```

1 # Importar os dados do arquivo Excel
2 dados <- readxl::read_excel("dados_moscas.xlsx")
3
4 dados

# A tibble: 259 x 31
#>   video animal_no_video sexo  tratamento animal_number tempo test_duration_s
#>   <chr> <chr>       <chr> <chr>      <chr>      <dbl>
#> 1 G1V3  2             M     controle    1           Total    2400
#> 2 G1V2  8             M     controle    2           Total    2400
#> 3 G1V1  1             M     controle    3           Total    2400
#> 4 G1V1  3             M     controle    4           Total    2400
#> 5 G1V3  12            M     controle    5           Total    2400
#> 6 G2V1  2             M     controle    6           Total    2400
#> 7 G2V3  13            M     controle    7           Total    2400
#> 8 G2V4  1             M     controle    8           Total    2400
#> 9 G2V4  3             M     controle    9           Total    2400
#> 10 G3V1 7             M    controle    10          Total    2400
#> i 249 more rows
#> i 24 more variables: total_distance_travelled_m <dbl>,
#>   total_time_mobile_s <dbl>, total_time_immobile_s <dbl>,
#>   total_mobile_episodes <dbl>, total_immobile_episodes <dbl>,
#>   number_of_entries_to_the_agar_zone <dbl>, time_in_the_agar_zone_s <dbl>,
#>   distance_travelled_in_the_agar_zone_m <dbl>,
#>   time_mobile_in_the_agar_zone_s <dbl>, ...

```

### 4.3 Visualizando os dados

```

1 dplyr::glimpse(dados)

Rows: 259
Columns: 31
$ video
$ animal_no_video
$ sexo
$ tratamento
$ animal_number
$ tempo
$ test_duration_s
$ total_distance_travelled_m
<chr> "G1V3", "G1V2", "G1V1", "G1V~  

<chr> "2", "8", "1", "3", "12", "2~  

<chr> "M", "M", "M", "M", "M", "M"~  

<chr> "controle", "controle", "con~  

<chr> "1", "2", "3", "4", "5", "6"~  

<chr> "Total", "Total", "Total", "~  

<dbl> 2400, 2400, 2400, 2400, 2400~  

<dbl> 17.768994, 9.695978, 7.65609~

```

```

$ total_time_mobile_s <dbl> 2266.728, 2164.461, 1879.804~
$ total_time_immobile_s <dbl> 456.56888, 45.38675, 373.661~
$ total_mobile_episodes <dbl> 92, 189, 158, 61, 61, 144, 1~
$ total_immobile_episodes <dbl> 63, 112, 176, 189, 145, 54, ~
$ number_of_entries_to_the_agar_zone <dbl> 63, 21, 81, 80, 21, 0, 60, 3~
$ time_in_the_agar_zone_s <dbl> 392.3685023, 392.9342318, 12~
$ distance_travelled_in_the_agar_zone_m <dbl> 0.92832895, 0.46652902, 0.89~
$ time_mobile_in_the_agar_zone_s <dbl> 140.29460, 297.49660, 77.520~
$ time_immobile_in_the_agar_zone_s <dbl> 53.039574, 5.436460, 32.6709~
$ immobile_episodes_in_the_agar_zone <dbl> 0, 7, 20, 9, 11, 1, 3, 5, 6, ~
$ number_of_entries_to_the_raia_zone <dbl> 204, 174, 217, 247, 153, 300~
$ time_in_the_raia_zone_s <dbl> 2343.337, 1472.244, 1206.188~
$ distance_travelled_in_the_raia_zone_m <dbl> 19.313651, 13.688998, 6.3153~
$ time_mobile_in_the_raia_zone_s <dbl> 2359.2774, 1849.0043, 1327.5~
$ time_immobile_in_the_raia_zone_s <dbl> 38.64488, 566.04737, 83.4141~
$ immobile_episodes_in_the_raia_zone <dbl> 94, 163, 34, 159, 193, 120, ~
$ number_of_entries_to_the_sacarose_zone <dbl> 91, 92, 142, 62, 89, 11, 184~
$ time_in_the_sacarose_zone_s <dbl> 192.105235, 318.022605, 35.3~
$ distance_travelled_in_the_sacarose_zone_m <dbl> 0.46174388, 0.76953020, 0.55~
$ time_mobile_in_the_sacarose_zone_s <dbl> 210.65218, 202.56256, 249.58~
$ time_immobile_in_the_sacarose_zone_s <dbl> 69.913487, 31.602566, 6.4040~
$ immobile_episodes_in_the_sacarose_zone <dbl> 10, 8, 5, 4, 10, 9, 4, 6, 3, ~
$ pindex <dbl> 0.99111436, 0.14099125, 0.78~

```

```
1 skimr::skim(dados)
```

Tabela 1: Data summary

Name	dados
Number of rows	259
Number of columns	31
Column type frequency:	
character	6
numeric	25
Group variables	None

#### Variable type: character

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
video	0	1	4	4	0	22	0
animal_no_video	0	1	1	2	0	14	0
sexo	0	1	1	1	0	2	0
tratamento	0	1	4	8	0	5	0
animal_number	0	1	1	3	0	259	0
tempo	0	1	5	5	0	1	0

#### Variable type: numeric

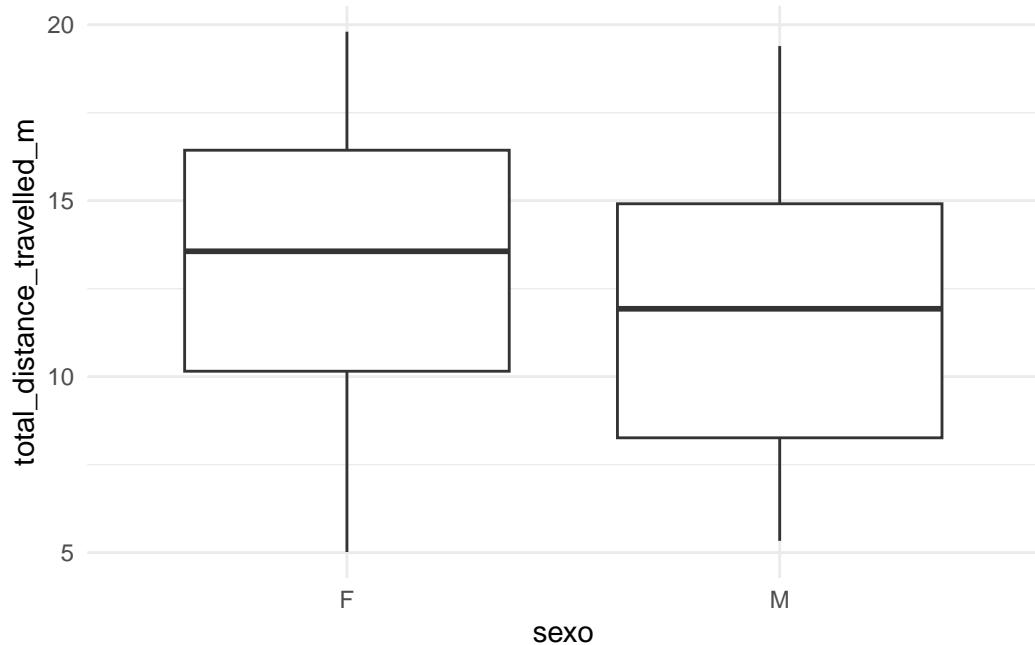
skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
test_duration_s	0	1	2400.00	0.00	2400.00	2400.00	2400.00	2400.00	2400.00	□□□□□
total_distance_travelled_m	0	1	12.49	4.11	5.02	9.20	12.36	15.78	19.80	□□□□□

	skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
∞	total_time_mobile_s	0	1	2115.62	174.80	1808.06	1962.93	2131.77	2275.08	2397.82	□□□□□
	total_time_immobile_s	0	1	319.91	174.32	2.43	175.97	317.93	481.22	594.83	□□□□□
	total_mobile_episodes	0	1	125.07	44.12	51.00	86.00	122.00	162.00	200.00	□□□□□
	total_immobile_episodes	0	1	123.66	44.63	50.00	84.00	125.00	158.50	200.00	□□□□□
	number_of_entries_to_the_agar_zone	0	1	46.40	29.17	0.00	20.50	45.00	72.00	99.00	□□□□□
	time_in_the_agar_zone_s	0	1	199.09	119.91	0.79	86.13	199.68	306.81	398.72	□□□□□
	distance_travelled_in_the_agar_zone_m	0	1	0.50	0.28	0.02	0.27	0.49	0.75	1.00	□□□□□
	time_mobile_in_the_agar_zone_s	0	1	190.64	109.52	1.59	96.00	187.74	279.97	396.42	□□□□□
	time_immobile_in_the_agar_zone_s	0	1	52.08	29.09	0.35	27.00	53.04	77.70	99.49	□□□□□
	immobile_episodes_in_the_agar_zone	0	1	10.47	6.10	0.00	5.00	11.00	16.00	20.00	□□□□□
	number_of_entries_to_the_raia_zone	0	1	193.84	58.55	100.00	147.00	183.00	250.00	300.00	□□□□□
	time_in_the_raia_zone_s	0	1	1713.17	421.12	1008.61	1329.61	1742.91	2078.82	2385.72	□□□□□
	distance_travelled_in_the_raia_zone_m	0	1	12.31	4.24	5.01	8.57	12.11	15.82	19.84	□□□□□
	time_mobile_in_the_raia_zone_s	0	1	1613.42	481.35	801.06	1203.63	1616.75	2050.79	2399.29	□□□□□
	time_immobile_in_the_raia_zone_s	0	1	297.35	178.04	2.75	137.88	295.15	451.37	598.57	□□□□□
	immobile_episodes_in_the_raia_zone	0	1	98.51	56.71	0.00	46.50	99.00	145.50	199.00	□□□□□
	number_of_entries_to_the_sacarose_zone	0	1	104.86	55.90	2.00	61.50	109.00	153.00	200.00	□□□□□
	time_in_the_sacarose_zone_s	0	1	210.76	115.86	1.13	115.42	212.04	311.36	399.89	□□□□□
	distance_travelled_in_the_sacarose_zone_m	0	1	0.48	0.28	0.00	0.23	0.46	0.73	1.00	□□□□□
	time_mobile_in_the_sacarose_zone_s	0	1	201.64	115.06	1.22	92.13	218.22	292.37	399.75	□□□□□
	time_immobile_in_the_sacarose_zone_s	0	1	49.25	28.08	0.25	23.78	50.44	71.73	99.74	□□□□□
	immobile_episodes_in_the_sacarose_zone	0	1	5.26	3.16	0.00	2.50	5.00	8.00	10.00	□□□□□
	pindex	0	1	0.52	0.30	0.00	0.25	0.51	0.79	1.00	□□□□□

## 5 Os gráficos

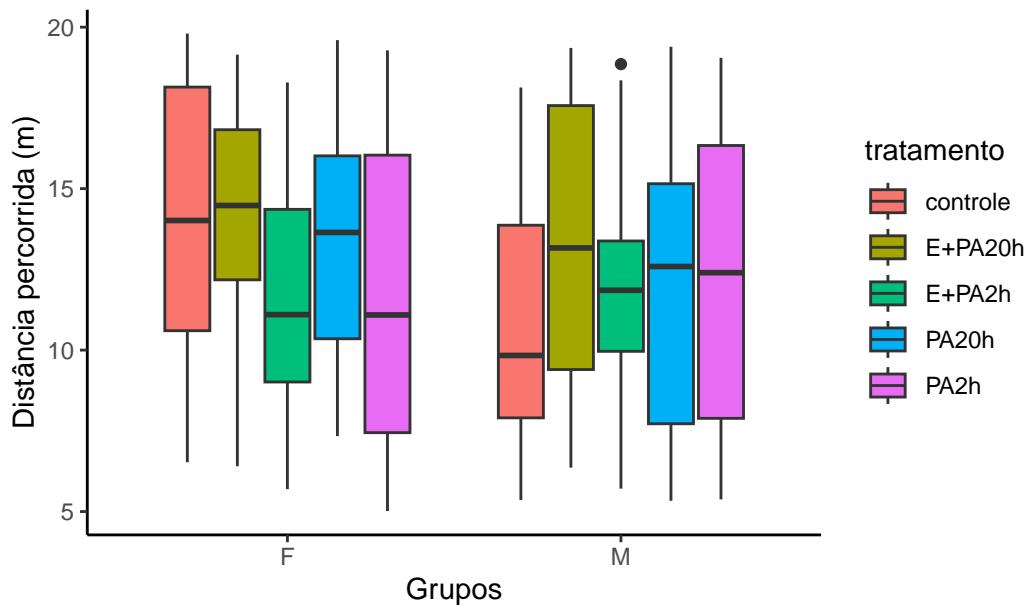
### 5.1 Gráfico box plot simples usando ggplot2

```
1 ggplot(dados, aes(x = sexo, y = total_distance_travelled_m)) +  
2   geom_boxplot() +  
3   theme_minimal()
```



```
1 # dando nome ao objeto, adicionando mais uma variável categórica, e mudando o tema:  
2 grafico <- ggplot(  
3   dados,  
4   aes(x = sexo, y = total_distance_travelled_m, fill = tratamento)) +  
5   geom_boxplot() +  
6   labs(  
7     title= "Quem caminha mais? machos ou fêmeas? com ggplot2",  
8     x = "Grupos",  
9     y = "Distância percorrida (m)") +  
10  theme_classic()  
11  
12 grafico
```

## Quem caminha mais? machos ou fêmeas? com ggplot2



Depois de satisfeitos, podemos salvar:

```
1 ggsave(  
2   plot = grafico, # o default é o ultimo gráfico gerado  
3   filename = "ggplot2.png",  
4   path = "figuras",  
5   dpi = 600 # o default é 300  
6 )
```

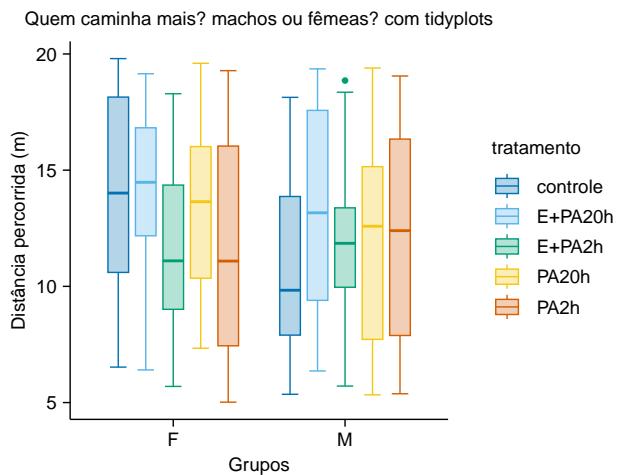
## 5.2 Gráfico box plot simples usando tidyplots

```
1 grafico <- tidyplot(  
2   dados,  
3   x = sexo,  
4   y = total_distance_travelled_m,  
5   fill = tratamento  
6 ) |>  
7   add_boxplot() |>  
8   adjust_title("Quem caminha mais? machos ou fêmeas? com tidyplots") |>  
9   adjust_x_axis_title("Grupos") |>
```

```

10  adjust_y_axis_title("Distância percorrida (m)") |>
11  theme_tidyplot() |> # tema padrão do tidyplots
12  save_plot("figuras/tidyplot.png")

```



```

1 # ou salvando com ggplot2
2 ggsave(
3   plot = grafico,
4   filename = "tidyplots.png",
5   path = "figuras",
6   dpi = 600
7 )

```

## 6 Agora é a sua vez:

1. Faça uma pergunta sobre os dados.
2. Crie um gráfico que responda a essa pergunta.
3. Personalize o gráfico com títulos, rótulos e cores.
4. Salve o gráfico.

1 # [ADICIONE O CÓDIGO AQUI]