

# ME115 - Linguagem R

## Atividade Prática 05 - Gabarito

1º semestre de 2023

Exercícios inspirados em <https://www.r-exercises.com/>

### Introdução

Nessa atividade, exploraremos:

1. Criação de documento em R Markdown;
2. Boas práticas na formatação de códigos.

### Atividade

Crie um novo arquivo R Markdown (.Rmd) no RStudio e salve-o com um nome apropriado. Insira um cabeçalho YAML com título, autor e data de sua escolha no topo do seu arquivo .Rmd. Neste cabeçalho também inclua a informação de que o arquivo a ser gerado pelo Knit é do formato html.

Usando o conjunto de dados `cars` disponível no R, execute as tarefas seguintes.

### Atividade

Usando o conjunto de dados `cars` disponível no R, execute as tarefas seguintes.

1. Escreva uma sentença falando sobre o conjunto de dados: quantas observações, quantas colunas e nomes das colunas. Essas informações devem ser inseridas diretamente no texto.

#### Solução:

```
O conjunto de dados cars possui r ncol(cars) variáveis (colunas) e r nrow(cars) observações. Os nomes das variáveis são: r names(cars).
```

O conjunto de dados `cars` possui 2 variáveis (colunas) e 50 observações. Os nomes das variáveis são: `speed`, `dist`.

2. Imprima as primeiras 6 linhas do conjunto de dados. Dica: função `head()`.

#### Solução:

```
head(cars)
```

```
##   speed dist
## 1     4    2
## 2     4   10
## 3     7    4
## 4     7   22
## 5     8   16
## 6     9   10
```

3. Carregue o pacote `knitr`. Se você não quiser que o código e seu resultado apareçam no relatório, use a opção `include=FALSE`.

**Solução:**

```
```{r, include=FALSE}
library(knitr)
```
```

4. Uma solução melhor para formatar uma tabela/matriz/data frame pode ser usar a função `kable()` do pacote `knitr`, que você acabou de carregar. Imprima as primeiras 6 linhas dos dados `cars` usando esse recurso e compare com a saída do item 2.

**Solução:**

```
kable(head(cars), caption = "Primeiras linhas do conjunto de dados `cars`")
```

Table 1: Primeiras linhas do conjunto de dados `cars`

| speed | dist |
|-------|------|
| 4     | 2    |
| 4     | 10   |
| 7     | 4    |
| 7     | 22   |
| 8     | 16   |
| 9     | 10   |

5. Exiba um resumo do conjunto de dados usando a função `summary()`.

**Solução:**

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   :  2
##  1st Qu.:12.0    1st Qu.: 26
##  Median :15.0    Median : 36
##  Mean   :15.4    Mean   : 43
##  3rd Qu.:19.0    3rd Qu.: 56
##  Max.   :25.0    Max.   :120
```

```
kable(summary(cars), caption = "Resumo do conjunto de dados `cars`")
```

Table 2: Resumo do conjunto de dados `cars`

| speed        | dist        |
|--------------|-------------|
| Min. : 4.0   | Min. : 2    |
| 1st Qu.:12.0 | 1st Qu.: 26 |
| Median :15.0 | Median : 36 |
| Mean :15.4   | Mean : 43   |
| 3rd Qu.:19.0 | 3rd Qu.: 56 |
| Max. :25.0   | Max. :120   |

6. Faça um gráfico de dispersão usando a função `plot()`, no qual a distância (`dist`) deve estar no eixo x e velocidade (`speed`) no eixo y. Controle a saída do gráfico usando as opções `fig.height` e `fig.width`. Insira também uma descrição da figura usando `fig.cap`.

**Solução:**

```
```{r, fig.height=3.5, fig.width=5, fig.cap="Gráfico de dispersão da velocidade
por distância", fig.align='center'}
```

```
par(mar = c(5, 5, 2, 1))
plot(x = cars$dist, y = cars$speed,
     las = 1, pch = 19, col = 4, cex.axis = 0.8, cex.main = 0.8,
     xlab = "Distância (ft)", ylab = "Velocidade (mph)",
     main = " ")
```

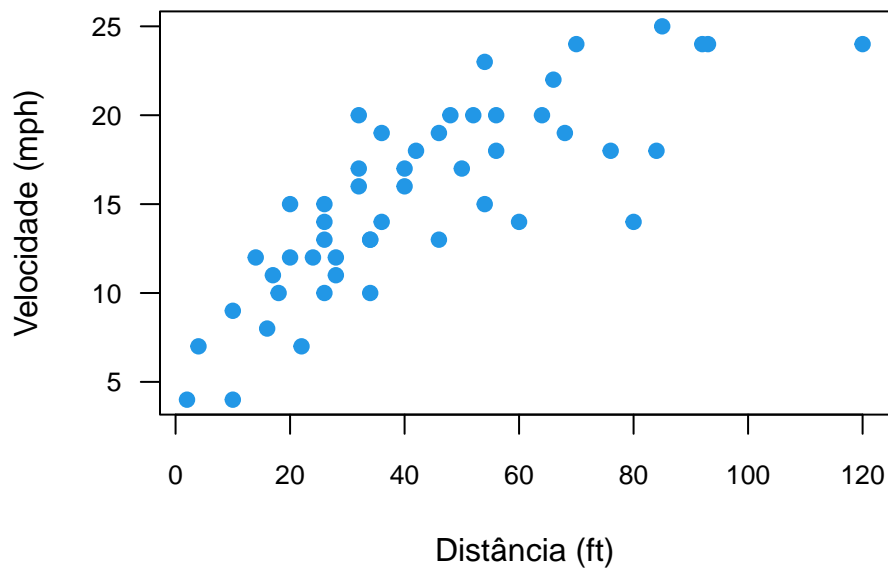


Figure 1: Gráfico de dispersão da velocidade por distância

7. Crie um objeto da classe data.frame e nomeie-o apropriadamente. O objeto deve conter o conjunto de dados cars e as variáveis a seguir:

- $speed2 = speed^2$
- $log.dist = \log(dist)$
- $sqrt.dist = \sqrt{dist}$

**Solução:**

```
mycars <- cars
mycars$speed2 <- cars$speed^2
mycars$log.dist <- log(cars$dist)
mycars$sqrt.dist <- sqrt(cars$dist)
str(mycars)
```

```
## 'data.frame': 50 obs. of 5 variables:
## $ speed : num 4 4 7 7 8 9 10 10 11 ...
## $ dist : num 2 10 4 22 16 10 18 26 34 17 ...
## $ speed2 : num 16 16 49 49 64 81 100 100 121 ...
## $ log.dist : num 0.693 2.303 1.386 3.091 2.773 ...
## $ sqrt.dist: num 1.41 3.16 2 4.69 4 ...
```

```
kable(head(mycars))
```

speed	dist	speed2	log.dist	sqrt.dist
4	2	16	0.693	1.41
4	10	16	2.303	3.16
7	4	49	1.386	2.00
7	22	49	3.091	4.69
8	16	64	2.773	4.00
9	10	81	2.303	3.16

8. Crie uma aplicação para uma função família `apply`, mas que seja diferente dos exemplos da aula prática 4. Apresente os resultados no seu documento.

**Solução:**

```
apply(cars, 2, function(x) c("Média" = mean(x), "Desvio Padrão" = sd(x)))
```

```
##           speed dist
## Média      15.40 43.0
## Desvio Padrão 5.29 25.8
```

9. Reproduza o resultado abaixo:

- Média e desvio padrão da velocidade (`speed`):

```
media.speed = mean(cars$speed)
sd.speed = sd(cars$speed)
```

A média e desvio padrão da variável `speed` são 15.4 e 5.3, respectivamente.

**Solução:**

A média e desvio padrão da variável `speed` são `media.speed` e `round(sd.speed, 1)`, respectivamente.

A média e desvio padrão da variável `speed` são 15.4 e 5.3, respectivamente.

- Faça o mesmo para calcular a média e desvio padrão da distância (`dist`).

**Solução:**

A média e o desvio padrão da da variável distância (`dist`) são, respectivamente, `mean(cars$dist)` e `sd(cars$dist)`.

A média e o desvio padrão da da variável distância (`dist`) são, respectivamente, 42.98 e 25.769.

10. Oculte o código do seu relatório usando a opção `echo`.

**Solução:**

```
```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = FALSE)
```
```

## Agradecimento

O material foi produzido pela Profa. Tatiana Benaglia para o curso de ME115.