

Exercícios 01

PJ/WZ

1 Aritmética

1.1 Calcular as seguintes expressões no R

- $\frac{32}{2} + 4$
- $3^2 \cdot 5 + 2$
- $2 + 5 \cdot 3^2$
- $\frac{39-4^2}{5}$
- $\frac{39-4^2}{5 \cdot 6}$
- $\frac{39-4^2}{5} \cdot 6$
- $1 - \frac{20}{20} \frac{19}{20} \frac{18}{20} \frac{17}{20} \frac{16}{20}$
- $\frac{e^{-2} 2^3}{3!}$
- $5 + \log_2 8$
- $3 \cdot \log_{10} 1000$
- $-2 \cdot \log_e(4)$
- $\log_5(80)$

2 Objetos: vetores

2.1 Criar vetores com os seguintes elementos

- a. 12 14 16 11 12 18 15 12 15
- b. 3 7 11 15 19 23
- c. 1 2 4 7 11 16 22 29 37 46 56 67

2.2 Fornecer comandos que forneçam a partir dos vetores criados no exemplo anterior:

- a. Quantos valores únicos há em a).
- b. Os valores em a) maiores que 13.
- c. O resultado em somar o vetor em b) com em c)?
- d. Quantos são valores de a) são divisíveis por 3? Quais são eles?
- e. Quantas vezes ocorrem cada valor de a).
- f. Qual o valor mais frequente em a).

Dica `unique()` e `which.max()`

2.3 Mais sobre criar vetores

Fornecer comandos para criar vetores como se segue.

- a. 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5
- b. 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3
- c. 1 2 3 4 5 6 7 8 9 10 1 2 3 4 5 6 7 8 9 10
- d. 0.5 1.5 2.5 3.5 4.5 5.5 6.5 7.5
- e. -3 -2 -1 0 1 2 3
- f. 10 10 15 15 15 15 15 13 13 13
- g. 1 2 2 3 3 3 4 4 4 4 5 5 5 5 5
- h. 10 10 10 10 10 9 9 9 9 8 8 8 7 7

Dica: `rep()` e `seq()`

2.4 Manipulando vetores

A tabela a seguir mostra o *PIB per capita* de alguns estados brasileiros.

Forneça os comandos necessários para os itens solicitados.

Estado	RS	SC	PR	SP	MG	RJ	ES
PIB per Capta	42.406	45.118	40.788	51.140	30.794	45.174	34.177

2.4.1 Usando dois vetores

- a. Entre com os dados no R definindo dois vetores, um vetor para as siglas dos estados e outro para os valores. Para entrar com os valores considere usar a função `scan()`.
- b. Encontre os valores médios, medianos e ainda o desvio padrão e o coeficiente de variação dos *PIB per capita*.
- c. Qual o menor e maior PIB? A quais estados correspondem?
- d. Quantos estados possuem o *PIB per capita* acima de 40.000 ? Quais são eles? Qual o *PIB per capita* mediano destes estados?
- e. Qual a proporção de estados com *PIB per capita* abaixo de 42.000? Quais são eles? Qual o *PIB per capita* médio destes estados?
- f. Quais os estados com *PIB per capita* superior à média?
- g. Ordene os vetores dos estados e dos valores por ordem alfabética dos estados. Dica: `order()`.
- h. Ordene os vetores dos estados e dos valores por ordem decrescente de valores.

2.4.2 Usando um vetor nomeado (*named vector*)

Uma representação alternativa é considerar um *vetor nomeado*.

Neste caso os valores seriam os de *PIB per capita* e os nomes seriam as siglas dos estados. Definir o vetor nomeado e refazer os itens anteriores

Hide

```
names(PIBpC) <- estados ## ou ...
(PIBpC <- c(RS = 42406, SC = 45118, PR = 40788, SP = 51140, MG = 30794, RJ = 45174, ES = 34177))
```

2.4.3 Usando um *data-frame*

Os dados podem ser organizados em uma estrutura ainda mais geral de *dataframe* que neste caso é melhor do que a de dois vetores separados.

- a. Ordene nesta estrutura e refaça os itens anteriores utilizando dados neste novo formato.

Hide

```
PIB <- data.frame(Estado = c("RS", "SC", "PR", "SP", "MG", "RJ", "ES"),
                  PIBpC = c(42406, 45118, 40788, 51140, 30794, 45174, 34177))
```

- Acrescente uma coluna no *data-frame* com a região (sul ou sudeste) na qual cada estado está inserido.
 - Obtenha o *PIB per capita* médio e máximo em cada região.
 - Reorganize o dataframe de forma que as siglas dos estados passem a ser os nomes das linhas (*rownames*).
- Qual(ais) a(s) diferença(s) com o formato anterior?

3 Matrizes e arrays

3.1 Dados em tabela

Em uma entrevista com um conjunto de estudantes foram coletados dados e parte deles gerou a tabela a seguir.

Forneça comandos em R para obter as quantidades pedidas.

	Exatas	Humanas	Biológicas
< 20 anos	58	42	38
≥ 20 anos	45	32	64

- Use o formato de uma matriz para representar os dados.
- Qual o número total de entrevistados?
- Qual o número de entrevistados de cada área?
- Qual o número de entrevistados de cada faixa etária?
- Isole o valor do número de estudantes de Humanas abaixo de 20 anos.
- Isole os valores dos entrevistados de idade igual acima de 20 anos para cada área.
- Qual a probabilidade de selecionar ao acaso um estudante de biológicas abaixo de 20 anos?
- Qual a probabilidade de selecionar ao acaso um estudante 20 ou mais anos da área de Exatas?
- Qual a probabilidade de selecionar ao acaso um estudante de humanas sabendo que é abaixo de 20 anos?
- Qual a probabilidade de selecionar ao acaso um estudante abaixo de 20 sabendo que é de Exatas?

Para entrar com os dados escolhemos o formato de *matriz* que, por *default* é preenchida por coluna. Nos comandos a seguir iniciamos mostrando como manter a matriz e definir seus *atributos*.

Hide

```
##a.
(MAT <- matrix(c(58, 45, 42, 32, 38, 64), ncol=3))
str(MAT)
attributes(MAT)
rownames(MAT)
rownames(MAT) <- c("m20", "M20")
colnames(MAT) <- c("Exatas", "Humanas", "Biologicas")
MAT
str(MAT)
attributes(MAT)
## OBS: pode-se
dimnames(MAT) <- list(c("m20", "M20"),
                      c("Exatas", "Humanas", "Biologicas"))
```

Os itens podem ser respondidos como a seguir. Em alguns casos há mais de uma maneira de se obter o resultado.

Code

Algumas funções pré implementadas podem facilitar o manuseio de matrizes.

Dica: `addmargins()` e `prop.table()`.

Code

OBS: funções de soma ou seleção de valores podem ser descritas por operações de vetores e matrizes. Examine e descreva o que cada operação a seguir representa.

Code

3.2 Uma curiosidade sobre operações com matrizes

Examine os comandos a seguir e interprete os resultados por eles gerados.

Os comandos em comentário (depois de "#") são alternativos aos mostrados, experimente estes também.

Hide

```
y <- c(22, 44, 52, 34)
U4 <- rep(1,4)
t(U4) %*% U4                ## crossprod(U4)
t(U4) %*% y                  ## crossprod(U4, y)
(t(U4) %*% y)/(t(U4) %*% U4) ## solve(crossprod(U4), crossprod(U4, y))
```

3.3 Dados em array

Em uma entrevista com um conjunto de estudantes foram coletados dados e com parte deles gerou-se a tabela a seguir.

Forneça comandos em R para obter as quantidades pedidas.

Universidade A			Universidade B		
Exatas	Humanas	Biológicas	Exatas	Humanas	Biológicas

	Universidade A			Universidade B		
< 20 anos	58	42	38	36	72	20
≥ 20 anos	45	32	64	24	28	30

- Entre com os dados na forma de um *array*.
- Obtenha o número de estudantes de Exatas da Universidade B com menos que 20 anos.
- Obtenha o número de estudantes de Humanas da Universidade B.
- Qual o total em cada instituição? E a proporção?
- Qual o total geral de estudantes de exatas? E de todas áreas?
- Quais as proporções de estudantes, por área, em cada Universidade?
- Qual o número total e a proporção de estudantes por faixa etária?

Dica: `apply()`

[Hide](#)

```
##a
##a.
(ARR <- array(c(58, 45, 42, 32, 38, 64,
               36, 24, 72, 28, 20, 30), dim=c(2, 3, 2)))
str(ARR)
dimnames(ARR) <- list(c("M20", "M20"),
                      c("Exatas", "Humanas", "Biologicas"),
                      paste0("Universidade", LETTERS[1:2]))
ARR

ARR[, , 1]
ARR[, , "UniversidadeA"]
```

3.4 Dados Titanic

3.4.1 Titanic, como *array*

Carregue os dados *Titanic* e verifique a documentação do objeto.
Forneça comandos para obter as quantidades pedidas.

[Hide](#)

```
data(Titanic)
Titanic
help(Titanic)
dim(Titanic)
str(Titanic)
```

- Qual o número e a proporção de sobreviventes?
- Quantos eram da tripulação (*crew*) e destes, qual a proporção de sobreviventes?
- Qual a proporção de sobreviventes entre mulheres?
- E entre as que viajavam na 1a classe?
- E entre mulheres e crianças? A taxa da sobrevivência foi maior nestes grupos comparado com demais?
- Qual a proporção de homens e mulheres no total? E a mesma proporção entre os que viajavam de

segunda classe?

3.4.2 Titanic, como *data.frame*

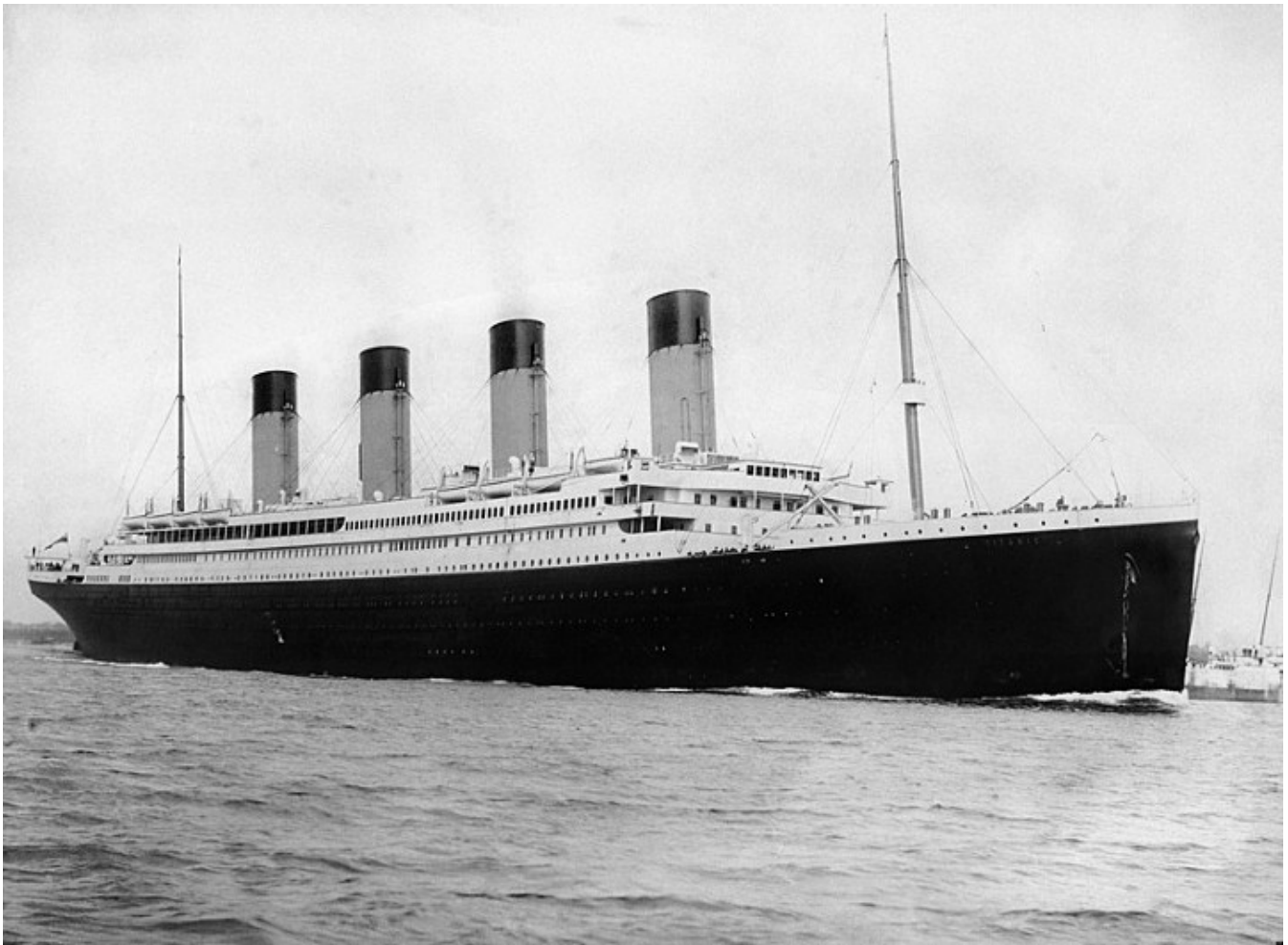
Objetos podem ser convertidos de um para outro formato.

Hide

```
Titanic.df <- as.data.frame(Titanic)
object.size(Titanic)
object.size(Titanic.df)
```

Recalcule itens anteriores utilizando agora os dados no formato *data.frame*.

Dica: `prop.table()` e `aggregate()`.



RMS Titanic, By F.G.O. Stuart (1843–1923)