

LISTA DE EXERCÍCIOS CÁLCULO COM R

Para cada pergunta você deve apresentar, por meio de prints de tela, o código completo em R e o resultado obtido, quer na saída da console ou em forma gráfica.

Upar o arquivo ao final

Se precisar divida a lista para upar até três arquivos.

Questões em vermelho são desafios

Fazer em dupla, mas entrega é individual

- 1) A linguagem R tem uma especificidade de aplicações:
Justifique por que é aplicada para
 - a) Aplicações estatísticas. Dê um exemplo prático
 - b) Criação de visualizações gráficas. Dê um exemplo prático
 - c) Implementação de modelos matemáticos. Dê um exemplo de modelagem matemática na prática.

- 2) Em um contexto de cursos superior de bacharelado, um sistema está sendo criado para se tornar inteligente e apontar o desempenho dos alunos. Uma prova de conceito (POC) é realizada à partir do trecho de código e dos seguintes requisitos a seguir. Apresente os **resultados dos alunos aprovados em forma gráfica** obedecendo ao código e requisitos. Faça insights.
 - a) Trecho de código

```
set.seed(seed)
notas <- data.frame(Matricula = sample(x = 011475:012990, size = 30, replace = FALSE),
ac1 = sample(x = 1:10, size = 30, replace = TRUE),
ac2 = sample(x = 1:10, size = 30, replace = TRUE),
prova_final= sample(x = 1:10, size = 30, replace = TRUE))
```
 - b) **Requisitos:**
seed=20
data.frame é uma função no R que cria uma tabela ou array de 2 dimensões.
sample (é uma função que retira uma amostra do conjunto de dados, quando a função replace = TRUE a amostra tem substituição)
Pesos das avaliações:
ac1 = 40%
ac2 = 40%
Prova_final = 20%
A aprovação é quando a média de ac1, ac2 e prova_final >= 7

- 3) Dados dois quadros, análise e utilize o quadro 2 para criar colunas novas no quadro 1. E plote uma representação gráfica de barras do faturamento de cada vendedor. Faça insights

Quadro 1 - produtos x vendedor			
Vendedor	Dia	Produto	Quantidade
Ana	1	saia	1
Flavia	1	casaco	2
Pedro	1	vestido	1
Mariana	1	meia	5
Ana	2	calçaFlaire	3
Flavia	2	calçaSkinny	0
Pedro	2	calçaSkinny	3
Mariana	2	manta	4

Quadro2 - Preços unitários		
Produto	Preço Unitário	Custo Unitário
calça	R\$ 240,00	R\$ 210,00
vestido	R\$ 320,00	R\$ 160,00
blusa	R\$ 98,00	R\$ 60,00
manta	R\$ 69,00	R\$ 39,00
saia	R\$ 145,00	R\$ 89,00
casaco	R\$ 289,00	R\$ 144,00
meia	R\$ 39,00	R\$ 15,00

- 4) Crie o seguinte código em R:

Dados dois valores, retorna o maior entre eles. Obs.: Insira mensagens de aviso para o usuário caso o valor que ele insira não seja numérico.

- 5) Utilizando seed =10, crie um vetor de valores numéricos com o seguinte código:

```
set.seed(seed)
vetorNumerico <- sample(x = 1:25, size = 800, replace = TRUE)
```

Calcule a frequência em que cada número aparece.

Lembre-se sample é uma função em R que gera uma amostra de dados. Consulte documentação (<https://www.rdocumentation.org/packages/base/versions/3.6.2/topics/sample>).

- 6) Para estudar o comportamento das vendas, um administrador coleta informações sobre o número de itens vendidos nos últimos 30 dias. Simule a coleta destes dados, onde as vendas variem entre 28 e 52 unidades. Use set.seed(50), de forma que possamos reproduzir os resultados dos chamados de Geradores de Números Aleatórios. Estes são, na verdade, **Pseudo-Aleatórios** porque são totalmente algorítmicos: dada a mesma semente, você obtém a mesma sequência.

```
vendas <- data.frame(Qtd = sample(x = 28:52, size = 30, replace = TRUE),
DiaMes = seq(from=as.Date("2021-09-01"), to=as.Date("2021-09-30"), by="day"),
DiaSemana = weekdays(seq(from=as.Date("2021-09-01"), to=as.Date("2021-09-30"),
by="day")))
```

- a) Plote um gráfico de dispersão da quantidade de vendas por dia.

```
plot(vendas$DiaMes, vendas$Qtd, main = "Frequência da Qtd de vendas", xlab = "Dia", ylab =
"Qtd vendas")
```

- b) Qual foi a variação das vendas no mês de análise?

```
varmes <- max(vendas$Qtd) - min(vendas$Qtd)
```

- c) Qual é a média de quantidade de vendas?

```
medven<- round(mean(vendas$Qtd))
```

- d) Construa o histograma da frequência de vendas. Coloque o título e os rótulos nos eixos e determine uma cor para o histograma.

```
hist(vendas$Qtd, main = "Frequência da Qtd de vendas", xlab = "Qtd vendas", ylab =
"Frequência", col = "green")
```

Você deverá analisar e justificar os itens a,b,c e d.

7) Gere as seguintes sequências e atribua a uma variável:

- a) $x1 = 1, 2, 3, 4$
- b) $x2 = 2, 5, 8, 11$
- c) $x3 = A, B, C, D$
- d) $x4 = a, b, c, d$
- e) $x5 = A, C, E, G$
- f) $x6 = 10, 9, 8, 7$
- g) $x7 = 10, 8, 6, 4$
- h) $x8 = 2, 2, 2, 2, 5, 5, 5, 5, 8, 8, 8, 8$
- i) $x9 =$ recebe todas as sequências anteriores ($x1, x2, x3, x4, x5, x6, x7, x8$)

8) Atribua a data de hoje para a variável x e a data do seu nascimento para a variável y e verifique a sua idade em dias. Verifique se sua idade está correta em anos.

9) Dados dois quadros, análise e utilize o quadro 2 para criar colunas novas no quadro 1. E plote uma representação gráfica de barras do faturamento de cada vendedor. Faça insights

Quadro 1 - produtos x vendedor			
Vendedor	Dia	Produto	Quantidade
Ana	1	saia	1
Flavia	1	casaco	2
Pedro	1	vestido	1
Mariana	1	meia	5
Ana	2	calçaFlaire	3
Flavia	2	calçaSkinny	0
Pedro	2	calçaSkinny	3
Mariana	2	manta	4

Quadro2 - Preços unitários		
Produto	Preço Unitário	Custo Unitário
calça	R\$ 240,00	R\$ 210,00
vestido	R\$ 320,00	R\$ 160,00
blusa	R\$ 98,00	R\$ 60,00
manta	R\$ 69,00	R\$ 39,00
saia	R\$ 145,00	R\$ 89,00
casaco	R\$ 289,00	R\$ 144,00
meia	R\$ 39,00	R\$ 15,00

- a) Quantas calças (flaire + skinny) foram vendidas? Qual o Faturamento total em calças?
 - b) Qual vendedor obteve maior desempenho de vendas (maior faturamento)? E qual vendedor obteve menor desempenho?
 - c) Qual faturamento gerado pelo vendedor Pedro?
- 10) Para estudar o desempenho dos vendedores, o gerente de uma loja de vestuário coleta informações sobre o número de peças vendidas por cada vendedor nos últimos 30 dias. Simule a coleta destes dados, conforme as instruções:
- a) Use `set.seed(20)`;
 - b) os vendedores são: Ana, Flavia, Pedro e Mariana. Cada vendedor deve apresentar dados referentes às suas vendas para os 30 dias; (Dica: use a função `rep()` e a cada 4 elementos repita a lista de vendedores; Faça o vetor dias ter tamanho $30*4$);
 - c) os produtos comercializados são: calçaFlaire, calçaSkinny, vestido, blusa, manta, saia, casaco, meia; as vendas devem variar entre 0 e 5 unidades diárias;

Quadro 1 - produtos x vendedor			
Vendedor	Dia	Produto	Quantidade
Ana	1	saia	1
Flavia	1	casaco	2
Pedro	1	vestido	1
Mariana	1	meia	5
Ana	2	calçaFlaire	3
Flavia	2	calçaSkinny	0
Pedro	2	calçaSkinny	3
Mariana	2	manta	4

11) Crie as seguintes funções:

Dado um vetor, retorna um novo vetor de somas acumuladas a cada posição.

Exemplo: Entrada: `x <- c(1,3,8,2,1,3)` ; Saída: 1, 4, 12, 14, 15, 18.

12) **Projetinho:** Dada a base de dados de temperatura das cidades do mundo, com o arquivo em csv, aplique técnicas de ETL e sumarizações para criar uma nova tabela que apresente a média do último ano de coleta das temperaturas de cada cidade. Utilize as ferramentas que você domina. Feito isso, apresente um gráfico das cidades que tem as 5 maiores temperatura mundiais.

Avalie se vc precisa tratar os dados, limpar, excluir, para que surjam novas tabelas, mais reduzidas, de modo a facilitar os insights. Mostre com prints de tela todas fases do projetinho.