

Nome completo legível

Número RA

Resolução e pontuação

(1) Resolva cada item, apresentando apenas as respostas finais. O primeiro item está resolvido como exemplo. Use este código e substitua M pelo **último** dígito de seu RA:

```
import pandas as pd
import numpy as np
df = pd.DataFrame({'Idade': [3,4,5,6], 'Tempo': [5,4,M,5]})
```

Ex.: Identifique o comando que ordena a planilha por Tempo.

```
df.sort_values(by="Tempo")
```

(a) Identifique o comando que filtra os valores de Idade no mínimo 4. (1pto)

```
df.query("Idade >= 4")
```

(b) Identifique o comando que calcula a média dos valores de Tempo. (1pto)

```
df["Tempo"].mean()
```

(c) Escreva a equação de regressão linear de Tempo em função de Idade. Calcule os coeficientes usando este código e arredonde-os com uma casa após a vírgula: (1pto)

```
(a, b) = np.polyfit(df["Idade"], df["Tempo"], deg = 1)
print(a, b)
```

```
Tempo = (a) * Idade + (b)
```

(Ou $y = a \cdot x + b$, mas o enunciado pede a equação, não o par de números.)

$$M=0 \Rightarrow \text{Tempo} = -0.4 \times \text{Idade} + 5.3$$

$$M=1 \Rightarrow \text{Tempo} = -0.3 \times \text{Idade} + 5.1$$

$$M=2 \Rightarrow \text{Tempo} = -0.2 \times \text{Idade} + 4.9$$

$$M=3 \Rightarrow \text{Tempo} = -0.1 \times \text{Idade} + 4.7$$

$$M=4 \Rightarrow \text{Tempo} = 0.0 \times \text{Idade} + 4.5$$

$$M=5 \Rightarrow \text{Tempo} = 0.1 \times \text{Idade} + 4.3$$

$$M=6 \Rightarrow \text{Tempo} = 0.2 \times \text{Idade} + 4.1$$

$$M=7 \Rightarrow \text{Tempo} = 0.3 \times \text{Idade} + 3.9$$

$$M=8 \Rightarrow \text{Tempo} = 0.4 \times \text{Idade} + 3.7$$

$$M=9 \Rightarrow \text{Tempo} = 0.5 \times \text{Idade} + 3.5$$

(2) O que é a mediana de um conjunto de números? Por que a calcularíamos, além da média? (2pts)

É o valor maior que metade dos elementos e menor que a outra metade. (1pt); não precisa ser a definição formal.)

A média pode ser distorcida por poucos valores que sejam extremos ou excepcionais. (1pt)

(3) Escreva um programa em Python que lê um inteiro n e verifica se n é par ou divide 3960, respondendo Lilás se sim e Verde se não. (Obs.: $a \% b == 0$ é verdadeiro exatamente quando b divide a , sendo positivos. Você pode assumir que n é positivo.) (3pts)

```
n = int(input("Digite um número:"))
if n % 2 == 0 or 3960 % n == 0:
    print("Lilás")
else:
    print("Verde")
```

1pt: estrutura lógica e indentação / newlines

1pt: conectivo or

1pt: testes com $a \% b == 0$

(4) Na planilha da Questão (1), os valores de Idade e Tempo estão correlacionados positiva ou negativamente? Justifique. O que isso significa? (2pts)

O coeficiente de correlação é dado por $df["Idade"].corr(df["Tempo"])$, mas o sinal já é o mesmo de a no item 1c. (1pt) Se positivo/negativo, os valores de Tempo tendem a crescer/decrescer conforme os valores de Idade crescem. (1pt)