

Distribuições de frequência com variáveis quantitativas

Quando construímos distribuições de frequências com variáveis quantitativas, precisamos, na maior parte dos casos, definir as classes em que desejamos agrupar os nossos dados. O **pandas** nos oferece a função `cut()` para realizar esta tarefa.

Analise o seguinte código:

```
classes = [dados.Altura.min(), 1.65, 1.75, dados.Altura.max()]
labels = ['1 - Baixa', '2 - Média', '3 - Alta']

frequencia = pd.value_counts(
    pd.cut(
        x = dados.Altura,
        bins = classes,
        labels = labels,
        include_lowest = True
    )
)

percentual = pd.value_counts(
    pd.cut(
        x = dados.Altura,
        bins = classes,
        labels = labels,
        include_lowest = True
    ), normalize = True
) * 100

dist_freq_altura = pd.DataFrame(
    {'Frequência': frequencia, 'Porcentagem (%)': percentual}
)

dist_freq_altura.rename_axis('Estaturas', axis= 'columns', inplace = True)

dist_freq_altura.sort_index(ascending = True, inplace = True)

dist_freq_altura
```

Utilizando a variável **Altura**, do nosso *dataset* de aula, indique a alternativa que contém a distribuição de frequências gerada pelo código acima.

Seleciona uma alternativa

A

Estaturas Frequência Porcentagem (%)

1 - Baixa	20529	26.716554
2 - Média	36162	47.061426
3 - Alta	20149	26.222020

B**Estaturas Freqüência Porcentagem (%)**

1 - Baixa	20529	0.267166
2 - Média	36162	0.470614
3 - Alta	20149	0.262220

C**Estaturas Freqüência Porcentagem (%)**

1 - Baixa	20149	26.222020
2 - Média	36162	47.061426
3 - Alta	20529	26.716554

D**Estaturas Freqüência Porcentagem (%)**

1 - Baixa	20528	26.715600
2 - Média	36162	47.062039
3 - Alta	20149	26.222361