

05

Selecionando as estatísticas descritivas

A criação de grupamentos com o método `groupby()` facilita bastante na sumarização das informações de um `DataFrame`. O método `describe()` aplicado a um grupamento gera um conjunto de estatísticas descritivas bastante útil no processo de análise de dados, conforme o exemplo abaixo:

```
precos = pd.DataFrame([['Feira', 'Cebola', 2.5],
                      ['Mercado', 'Cebola', 1.99],
                      ['Supermercado', 'Cebola', 1.69],
                      ['Feira', 'Tomate', 4],
                      ['Mercado', 'Tomate', 3.29],
                      ['Supermercado', 'Tomate', 2.99],
                      ['Feira', 'Batata', 4.2],
                      ['Mercado', 'Batata', 3.99],
                      ['Supermercado', 'Batata', 3.69]],
                     columns = ['Local', 'Produto', 'Preço'])

precos
```

	Local	Produto	Preço
0	Feira	Cebola	2.5
1	Mercado	Cebola	1.99
2	Supermercado	Cebola	1.69
3	Feira	Tomate	4
4	Mercado	Tomate	3.29
5	Supermercado	Tomate	2.99
6	Feira	Batata	4.2
7	Mercado	Batata	3.99
8	Supermercado	Batata	3.69

```
produtos = precos.groupby('Produto')
produtos.describe().round(2)
```

Produto	Preço	mean	std	min	25%	50%	75%	max
Cebola	3.0	3.96	0.26	3.69	3.84	3.99	4.10	4.2
Tomate	3.0	2.06	0.41	1.69	1.84	1.99	2.24	2.5
Batata	3.0	3.43	0.52	2.99	3.14	3.29	3.64	4.0

Utilizando o `DataFrame` (`precos`) acima, que código seria necessário para gerar a visualização abaixo:

Produto	Média	Desvio Padrão	Mínimo	Máximo
Cebola	3.96	0.26	3.69	4.2
Tomate	2.06	0.41	1.69	2.5
Batata	3.43	0.52	2.99	4.0

Selecione uma alternativa

A

```
estatisticas = ['mean', 'std', 'min', 'max']
nomes = {'mean': 'Média', 'std': 'Desvio Padrão',
         'min': 'Mínimo', 'max': 'Máximo'}
produto['Preço'].aggregate(estatisticas)
    .rename(columns = nomes)
```

B

```
estatisticas = ['mean', 'std', 'min', 'max']
nomes = {'mean': 'Média', 'std': 'Desvio Padrão',
         'min': 'Mínimo', 'max': 'Máximo'}
produto['Preço'].aggregate(estatisticas)
    .rename(columns = nomes).round(2)
```

C

```
estatisticas = ['mean', 'std', 'min', 'max']
nomes = ['Média', 'Desvio Padrão', 'Mínimo', 'Máximo']
produto['Preço'].aggregate(estatisticas)
    .rename(columns = nomes).round(2)
```

D

```
estatisticas = ['mean', 'devp', 'min', 'max']
nomes = {'mean': 'Média', 'devp': 'Desvio Padrão',
         'min': 'Mínimo', 'max': 'Máximo'}
produto['Preço'].aggregate(estatisticas)
    .rename(columns = nomes).round(2)
```