



64%

02

Análise descritiva

ATIVIDADES
2 de 6DISCORD
ALURAFÓRUM DO
CURSOVOLTAR
PARA
DASHBOARDMODO
NOTURNO

14.5k xp

a

Transcrição

Agora que já realizamos várias análises visuais gerando alguns gráficos, faremos **análises descritivas** em relação ao valor da conta, porcentagem, média de gorjetas e etc.

Começaremos exibindo a média geral das gorjetas escrevendo `gorjetas.gorjeta.mean()` . Para organizar melhor o resultado, atribuiremos este valor à variável `media_geral_gorjetas` .

```
media_geral_gorjetas = gorjetas.gorjeta.mean()
```

COPIAR CÓDIGO

Para deixarmos um descrição melhor, criaremos um `print` recebendo 'A média geral das gorjetas é de {}' seguido de `format()` , passando `media_geral_gorjetas` .

```
print('A média geral das gorjetas é de {}'.format(media_geral_gorjetas))
```

COPIAR CÓDIGO

Agora, faremos a média geral dos valores para quinta, sexta, sábado e domingo. Pegaremos primeiro a média de todos os dias com `gorjetas.groupby()` , recebendo a especificação de



64%

'dia_da_semana' entre colchetes. Em seguida, aplicamos a média `mean()`.

```
gorjetas.groupby(['dia_da_semana']).mean()
```

[COPIAR CÓDIGO](#)

| | valor_da_conta | gorjeta | total_de_pessoas | porcentagem |
|---------------|----------------|----------|------------------|-------------|
| dia_da_semana | | | | |
| Domingo | 21.410000 | 3.255132 | 2.842105 | 0.166974 |
| Quinta | 17.682742 | 2.771452 | 2.451613 | 0.161129 |
| Sexta | 17.151579 | 2.734737 | 2.105263 | 0.169474 |
| Sábado | 20.441379 | 2.993103 | 2.517241 | 0.153678 |

Pelas informações da tabela, poderemos ver que o dia com a maior média de gorjeta é domingo. A média do `total_de_pessoas` deve ser feita por números inteiros, já que não é possível ter números decimais neste caso. Portanto, removeremos este campo de nossa análise adicionando dois colchetes após `mean()` para inserir os valores que queremos exibir.

```
gorjetas.groupby(['dia_da_semana']).mean()[['valor_da_conta', 'gorjeta', 'porcer
```

[COPIAR CÓDIGO](#)

Com isso, não teremos mais os dados de `total_de_pessoas`, pois não são relevantes para esta nossa análise. A média de `porcentagem` possui o menor valor no sábado, em comparação com os demais

ATIVIDADES
2 de 6

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODOS
NOTURNO



14.5k xp

a



64%

dias.

Estamos olhando somente a média dos valores, sem levar em consideração o número de clientes em cada dia. Para saber a frequência desses dias, inserimos `gorjetas.dia_da_semana` com `value_counts()` .

```
gorjetas.dia_da_semana.value_counts()
```

[COPIAR CÓDIGO](#)

Assim poderemos visualizar a quantidade de pessoas que foram ao restaurante nos quatro dias. O sábado é o dia mais frequentado e a sexta-feira é o dia menos frequentado.

Para deixar nosso projeto mais bem organizado, adicionamos outro `print()` com a frase 'Frequência dos dias' .

```
print('Frequência dos dias')
gorjetas.dia_da_semana.value_counts()
```

[COPIAR CÓDIGO](#)

```
➜ Frequência dos dias
Sábado      87
Domingo     76
Quinta      62
Sexta       19
Name: dia_da_semana, dtype: int64
```

ATIVIDADES
2 de 6

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODO
NOTURNO



14.5k xp

a



64%

Compararemos com a tabela de médias e veremos que a gorjeta é maior para o domingo, o segundo dia mais frequentado. Enquanto o sábado que é o dia mais cheio não possui a maior média de gorjeta.

A seguir, veremos se os dois grupos são de fato diferentes.

ATIVIDADES
2 de 6

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODOS
NOTURNO



14.5k xp

a