



67%

03

Teste de hipótese

ATIVIDADES
3 de 6DISCORD
ALURAFÓRUM DO
CURSOVOLTAR
PARA
DASHBOARDMODO
NOTURNO

14.5k xp

a

Transcrição

Anteriormente, vimos que a média do valor da gorjeta no domingo é maior do que a do sábado. Porém, também observamos que a quantidade de pessoas no restaurante é maior no sábado do que no domingo.

A questão é se os valores das contas nos dois dias são **diferentes** também. Para verificarmos isso, escreveremos mais uma célula textual para organizar nosso projeto, com o escrito "## Teste de Hipótese".

Nossa **hipótese nula** busca descobrir se realmente a distribuição do valor da conta é **igual** no sábado e no domingo. Já a **hipótese alternativa** é justamente o **contrário**, e escreveremos suas descrições em células textuais da mesma forma que fizemos anteriormente.

Agora, montaremos a `query()` que trará o valor da conta de cada um dos dias em questão, como realizamos em nossos testes relacionados a gorjeta. Começaremos com `'Domingo'` e guardaremos os valores em uma variável `valor_conta_domingo`.

```
valor_conta_domingo = gorjetas.query("dia_da_semana == 'Domingo']").valor_da_conta
```

COPIAR CÓDIGO



67%

ATIVIDADES
3 de 6

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODOS
NOTURNO



14.5k xp

a

Em seguida, copiaremos a sentença e colaremos em uma nova célula de edição, apenas substituindo por 'Sábado' para a nova variável `valor_conta_sabado`.

```
valor_conta_sabado = gorjetas.query("dia_da_semana == 'Sábado']").valor_da_conta
```

COPIAR CÓDIGO

Assim, obteremos os valores das contas nos dois dias. Agora, rodaremos o `ranksums()` para ver se os dois grupos são iguais ou diferentes.

Para isso, escreveremos `ranksums()` recebendo as duas variáveis.

```
ranksums(valor_conta_domingo, valor_conta_sabado)
```

COPIAR CÓDIGO

O resultado de `pvalue` será `0.35`, ou seja, apenas aceitaremos a hipótese alternativa se este número for menor do que `0.05`; como é maior, apenas nossa **hipótese nula** é válida. Isso significa que a distribuição do valor da conta **é igual** no sábado e no domingo.

Portanto, mesmo que o sábado seja mais frequentado e o domingo possua a maior média de gorjetas, não são diferentes no que diz respeito ao valor da conta.

Para finalizarmos e melhorarmos a visualização, atribuiremos este valor para uma variável `r2` e aplicaremos `print()` dizendo que 'O valor do p-value é {}' seguido de `format()` com



67%

r2.pvalue .

```
r2 = ranksums(valor_conta_domingo, valor_conta_sabado)
print('O valor do p-value é {}'.format(r2.pvalue))
```

COPIAR CÓDIGO

ATIVIDADES
3 de 6

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODOS
NOTURNO



14.5k xp

a

Desta forma, teremos nosso relatório mais bem organizado.