

Capítulo 6/7 · Faltam 2 lições

## Projeto Final

# Descrição do projeto

O trabalho de um analista é analisar dados para obter percepções valiosas dos dados e tomar decisões fundamentadas neles. Esse processo consiste em várias etapas: visão geral de dados, pré-processamento de dados e testes de hipóteses.

Hipóteses são suposições feitas sobre a população com base em uma amostra de dados. Essas suposições são testadas usando métodos estatísticos para determinar se elas são verdadeiras ou falsas. Testar uma hipótese envolve fazer uma declaração sobre a população, coletar dados e usar métodos estatísticos para testar essa declaração.

Sempre que fazemos análises, nós temos que formular hipóteses que depois poderemos testar. Às vezes nós aceitamos essas hipóteses; outras vezes, nós as rejeitamos. Para tomar as decisões certas, um negócio deve ser capaz de entender se está fazendo as suposições certas ou não.

## Descrição do projeto

No contexto deste projeto, você vai testar uma hipótese relacionada às preferências musicais de duas cidades. Para isso, você vai analisar os dados de um serviço de streaming de música online para testar a hipótese apresentada abaixo e comparar o comportamento dos usuários nessas duas cidades.

Isso envolverá analisar os dados de um serviço de streaming real para comparar o comportamento dos usuários em Springfield e Shelbyville. O projeto é dividido em

três etapas, cada uma das quais tem seus objetivos específicos.

Na Etapa 1, você fornecerá uma visão geral dos dados e escreverá suas observações. Na Etapa 2, você fará o pré-processamento dos dados, limpando-os. Finalmente, na Etapa 3, você testará a hipótese dando os passos de programação necessários para testar cada declaração e comentar seus resultados nos blocos apropriados.

Depois de concluir essas etapas, você será capaz de obter percepções valiosas dos dados e tomar decisões baseadas neles.

## Hipótese

Para este projeto, reunimos os requisitos e preparamos uma hipótese que precisamos confirmar ou rejeitar.

Ao testar hipóteses, é importante perceber que elas podem ser totalmente aceitas, parcialmente aceitas, parcialmente rejeitadas ou totalmente rejeitadas.

Quando uma hipótese é totalmente aceita, isso significa que os resultados do teste confirmam a declaração feita sobre a população sem quaisquer dúvidas.

Se ela for parcialmente aceita, isso significa que os resultados confirmam a declaração até certo ponto, mas não suficientemente para aceitá-la totalmente.

Por outro lado, se uma hipótese for totalmente rejeitada, isso significa que os resultados do teste não confirmam a declaração feita sobre a população.

Finalmente, uma hipótese também pode ser parcialmente rejeitada se os dados indicarem que é falsa, mas você não puder rejeitá-la totalmente. Quando interpretamos os resultados de um teste de hipótese, é importante considerar todas essas diferentes possibilidades.

Aqui está a hipótese que precisamos aceitar ou rejeitar:

1. A atividade de desenvolvimento é diferente da atividade de teste de desenvolvimento.

## Dicionário de dados

Os dados estão armazenados no arquivo `/datasets/music_project_en.csv`. Alternativamente, você pode fazer o download [aqui](#).

### Descrição das colunas:

- `'userID'` — identifica cada usuário univocamente
- `'Track'` — título da música
- `'artist'` — nome do artista
- `'genre'` — gênero da música
- `'City'` — cidade do usuário
- `'time'` — a hora do dia em que uma música foi tocada (HH:MM:SS)
- `'Day'` — dia da semana

## Instruções para completar o projeto

Preparamos para você um modelo de notebook, em que você pode escrever seu código e descrever suas análises. Para completar o projeto, preencha cada célula de código no modelo e edite as células Markdown nos casos em que o modelo solicitar que você explique seus resultados.

É uma boa prática sempre incluir uma introdução que descreve brevemente seus objetivos e uma conclusão que resume seus resultados na forma de células Markdown.

Antes de você começar, vamos revisar as três etapas do projeto novamente:

**Etapas 1:** visão geral dos dados. O notebook tem células prontas com instruções sobre que tipo de código escrever, bem como blocos de texto onde você pode escrever suas observações.

**Etapas 2:** pré-processamento de dados. Nesta etapa, você arrumará nomes de colunas e removerá valores ausentes e duplicados. Siga o esquema fornecido no notebook e certifique-se de escrever suas observações no final desta seção.

**Etapas 3:** teste da hipótese. Esta é a parte principal do seu projeto. Dê os passos de codificação necessários para testar cada declaração e comentar seus resultados

nos blocos apropriados. Finalmente, resuma todo o projeto na seção "Conclusões".

No vídeo abaixo, abordamos alguns pontos essenciais nos quais você deve prestar atenção enquanto trabalha no seu projeto.

### Your First Project NM



## Checklist

Antes de enviar o projeto, certifique-se que:

- você concluiu todas as etapas do projeto e preencheu todas as células
- a estrutura lógica do projeto segue a estrutura do modelo.

[Vamos lá!](#)

Ou faça o projeto em seu computador e faça upload de seu trabalho quando finalizar. [Fazer upload e entregar](#)

[Avançar](#)

Quando a revisão do projeto for concluída, você vai receber um e-mail no seu endereço de e-mail cadastrado na plataforma. Você vai precisar abrir o projeto e fazer correções, se necessário.

## Seu projeto foi revisado e encontramos alguns problemas

Por favor, os corrija e reenvie o seu projeto ↓

Ok, vamos lá corrigir ele!



