

### Passo 3/3

Muito bem! Seu projeto foi aprovado. 🙌

O revisor pode ter deixado algumas sugestões para que você possa tornar o projeto ainda melhor, então não se esqueça de verificá-las.

[Visualizar](#)

Parabéns! Você concluiu o sprint de pré-processamento de dados. É hora de aplicar o conhecimento e as habilidades que você adquiriu em um projeto: um estudo de caso analítico da vida real que você concluirá por conta própria.

Após terminar o projeto, envie seu trabalho para o revisor do projeto para avaliação por meio da plataforma. Ele te dará feedback dentro de 24 horas. Use o feedback para fazer alterações e, em seguida, envie a nova versão de volta ao revisor do projeto.

Você pode obter mais comentários sobre a nova versão. Isso é completamente normal. Não é incomum passar por vários ciclos de feedback e revisão.

Seu projeto será considerado concluído assim que o revisor do projeto o aprovar.

## Descrição do Projeto

Seu projeto é preparar um relatório para a divisão de empréstimos de um banco. Você precisará descobrir se o estado civil de um cliente e o número de filhos têm impacto sobre a inadimplência de um empréstimo. O banco já tem alguns dados sobre a capacidade de crédito dos clientes.

Seu relatório será considerado ao criar uma **pontuação de crédito** de um cliente em potencial. Uma **pontuação de crédito** é usada para avaliar a capacidade de um devedor em potencial de pagar seu empréstimo.

## Instruções para concluir o projeto

**Passo 1.** Abra o arquivo de dados `/datasets/credit_scoring_eng.csv` e dê uma olhada nas informações gerais.

## Passo 2. Pré-processe os dados:

- Identifique e preencha os valores ausentes.
- Substitua o tipo de dados de número real pelo tipo inteiro.
- Excluir dados duplicados.
- Categorize os dados.

Certifique-se de explicar:

- quais valores ausentes você identificou;
- possíveis razões que esses valores ausentes estavam presentes;
- qual método você usou para preencher os valores ausentes;
- qual método você usou para localizar e excluir dados duplicados e por quê;
- possíveis razões para a presença de dados duplicados;
- qual método você usou para alterar o tipo de dados e por quê;
- quais dicionários você selecionou para este conjunto de dados e por quê.

Os dados podem conter artefatos ou valores que não correspondem à realidade – por exemplo, um número negativo de dias empregados. Esse tipo de coisa acontece quando você está trabalhando com dados reais. Você precisa descrever os possíveis motivos pelos quais esses dados podem ter aparecido e processá-los.

### Passo 3. Responda a estas perguntas:

- Existe alguma relação entre ter filhos e pagar um empréstimo em dia?
- Existe alguma relação entre o estado civil e o pagamento de um empréstimo no prazo estipulado?
- Existe uma relação entre o nível de renda e o pagamento de um empréstimo no prazo?
- Como as diferentes finalidades do empréstimo afetam o pagamento pontual do empréstimo?

Interprete suas respostas. Explique o que significam os resultados obtidos.

#### Passo 4. Escreva uma conclusão geral.

### Formato:

Conclua o projeto no notebook Jupyter (ele é aberto quando você clica em Next - Avançar). Escreva o código nas células de **código** e notas com explicações e interpretações nas células de **marcação**.

Use formatação e títulos.

## Descrição de dados

- `children` : o número de crianças na família
- `days_employed` : quanto tempo o cliente trabalhou
- `dob_years` : a idade do cliente
- `education` : o nível de educação do cliente
- `education_id` : identificador da educação do cliente
- `family_status` : estado civil do cliente
- `family_status_id` : identificador do estado civil do cliente
- `gender` : o sexo do cliente
- `income_type` : o tipo de renda do cliente
- `debt` : se o cliente já deixou de pagar um empréstimo
- `total_income` : renda mensal
- `purpose` : motivo para fazer um empréstimo

## Como meu projeto será avaliado?

Seu mentor usará critérios de avaliação para avaliar seu projeto. Antes de começar, leia estes critérios com atenção. Eles irão ajudá-lo a garantir que você está fazendo um bom trabalho.

Veja o que os mentores analisam ao avaliar seu projeto:

- Como você descreve os problemas identificados nos dados?
- Que métodos você usa para substituir tipos de dados e para processar valores ausentes e dados duplicados?
- Você categorizou os dados? Por que você fez do jeito que fez?
- Você exportou os dados finais para tabelas dinâmicas?
- Você usou as cláusulas *try-except* em seu código para lidar com erros inesperados?
- Você seguiu a estrutura correta do projeto e escreveu um código limpo?
- Você tirou conclusões precisas e úteis?
- Você deixou comentários sobre cada passo que deu?

Ao clicar em "Vamos!" você será levado ao JupyterHub da TripleTen. Você pode concluir seu projeto lá.

Há outra opção: você pode instalar o Python em seu computador e fazer seu projeto lá. Existem muitos guias para instalar o Python e suas bibliotecas em seu computador.

Uma maneira fácil de instalar tudo o que você precisa de uma só vez é baixar o kit de distribuição do Anaconda (<https://www.anaconda.com/distribution/>) (os materiais estão em inglês). Durante a instalação, verifique as seguintes opções nas configurações avançadas:

- Adicione o Anaconda à variável de ambiente PATH do sistema
- Registre o Anaconda como o sistema Python 3.n (3.n será substituído pelo número da versão do software Python incluído no download do Anaconda).

Existem duas maneiras de abrir o Jupyter Notebook:

- Encontre o Jupyter Notebook no Anaconda Navigator e inicie-o clicando no ícone
- Escreva *jupyter notebook* no Prompt de Comando (Windows) ou Terminal (macOS)

O conhecimento e as habilidades que você adquiriu nas lições anteriores o ajudarão a concluir seu projeto com sucesso.

Mantenha suas folhas de conclusões e resumos à mão.

Boa sorte!

Avancar

O que você achou do processo de revisão de código? Por favor nos informe. Isso nos ajuda a melhorar a experiência educacional.

## Share your thoughts