

Case de Financiamento de Veículos

Cenário

Você foi contratado como estagiário de dados em uma empresa que oferece financiamento para compra de veículos. A empresa está interessada em entender melhor os padrões de concessão de crédito e os fatores que influenciam a inadimplência dos clientes. Um dos objetivos principais é otimizar o processo de análise de crédito, minimizando riscos e maximizando a rentabilidade.

A empresa também sabe que clientes de regiões com rendas menores tendem a apresentar maior risco de inadimplência. Portanto, é importante identificar padrões nos dados que possam ajudar a prever comportamentos de pagamento e inadimplência.

Instruções

Você receberá um dataset fictício com as seguintes informações:

- **Clientes:** ID do cliente, idade, gênero, localização (região/cidade), faixa de renda.
- **Financiamentos:** ID do financiamento, ID do cliente, valor financiado, data da concessão, prazo de financiamento, taxa de juros aplicada, valor da parcela.
- **Pagamentos:** ID do financiamento, data do pagamento, status do pagamento (em dia, atrasado, inadimplente).
- **Características do Veículo:** ID do veículo, modelo, ano, valor de compra.
- **Informações Regionais:** Renda média da região, taxa de inadimplência histórica da região.

O seu objetivo é realizar uma análise detalhada dos dados, responder às perguntas de negócio e construir um modelo preditivo simples para identificar clientes com maior risco de inadimplência.

Parte 1: Análise Exploratória de Dados (EDA)

1. Entendimento do Dataset:

- Carregue o dataset em Python e forneça uma visão geral dos dados (estrutura, colunas, tipos de dados, valores nulos).
- Realize a limpeza de dados se necessário (remoção de valores nulos, tratamento de outliers, etc.).

2. Análise de Clientes:

- Qual é a distribuição dos clientes por gênero, faixa etária e faixa de renda?
- Quais são as características dos clientes que tomam empréstimos maiores?

3. Análise de Financiamentos e Veículos:

- Quais são as regiões com maior volume de financiamentos?
- Há diferenças significativas nas taxas de inadimplência entre regiões com diferentes rendas médias?
- Qual o perfil dos veículos financiados (valor médio dos veículos, modelo e ano mais financiados)?

4. Análise de Inadimplência:

- Qual é a taxa de inadimplência geral e como ela varia de acordo com a renda da região e as características dos clientes?
- Quais variáveis (idade, renda, tipo de veículo, valor financiado) parecem estar mais associadas à inadimplência?

Parte 2: Visualização de Dados

Para responder aos itens 2 a 4, crie visualizações para apresentar os insights obtidos na análise exploratória de dados. Utilize bibliotecas de seu conhecimento e defina a melhor maneira de apresentar as respostas e resuma os principais insights obtidos na análise exploratória e nas visualizações.

Item Extra: Modelagem Preditiva

Construa, caso seja de seu conhecimento, um modelo preditivo simples para identificar os clientes com maior probabilidade de inadimplência. Esse modelo ajudará a empresa a ajustar suas políticas de crédito e reduzir riscos.

Entrega

- **Código:** Submeta o código em Python
- **Relatório:** Entregue uma apresentação de no máximo 3 slides apresentando as respostas
- **Prazo:** até o dia da entrevista presencial