

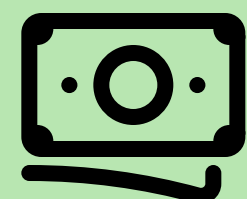
MARIAH LISBOA



JULHO/2025

# ANÁLISE DE CRÉDITO BANCÁRIO

SCORE DE RISCO COM CLASSIFICAÇÃO INTELIGENTE DE CLIENTES INADIMPLENTES



CONTEXTO E OBJETIVOS

CONTEXTO

queda da taxa de juros % ↓  
aumento na demanda de pedidos por crédito ↗

DIFICULDADE

processo manual  
pressão por decisões mais ágeis e assertivas

OBJETIVOS

como identificar, de forma simples e eficiente, os clientes com maior risco de inadimplência?

# METODOLOGIAS E PROCESSOS APLICADOS

Tratamento e limpeza1	Análise exploratória2	Técnicas aplicadas3
<ul style="list-style-type: none"><li>• <b>Identificação e tratamento</b> de variáveis-chave com <b>dados faltantes</b>.</li><li>• <b>Identificação de inconsistências</b> no cálculo do índice de dívida sobre patrimônio, dificultado pela ausência de variáveis essenciais como salário.</li><li>• <b>Tratamento de variáveis cruciais</b>, como salário e número de dependentes, <b>preservando sua tipologia</b> e a <b>consistência da análise</b>.</li></ul>	<ul style="list-style-type: none"><li>• Mapeamento dos <b>perfis com maior taxa de inadimplência</b> por faixa etária, renda e dívidas.</li><li>• <b>Segmentação das variáveis em grupos de risco</b> com base na distribuição dos dados.</li><li>• <b>Criação do grupo extremo (Q5)</b>, com risco significativamente “acima da média”.</li></ul>	<ul style="list-style-type: none"><li>• Cálculo do <b>risco relativo</b> para identificação de <b>grupos com maior probabilidade de inadimplência</b>.</li><li>• Construção de um <b>score simplificado</b> com variáveis binárias (alta ou baixa exposição ao risco).</li><li>• Validação do modelo com <b>matriz de confusão</b> e análise da força preditiva do score.</li></ul>

COMPOSIÇÃO DOS DADOS

População

35.575  
clientes

Apenas **622 clientes** são  
classificados como  
**inadimplentes**.

Volume de empréstimos

305.335  
empréstimos

Cada cliente possui, em média,  
**9 empréstimos registrados**.

Taxa de inadimplência

1,75%  
inadimplentes

**Cenário altamente  
desbalanceado**, o que exige  
atenção nos modelos preditivos

# PANORAMA GERAL DA BASE

TAXA DE DADOS FALTANTES

19,77%

Clientes não informaram dados essenciais como **salário e número de dependentes**, o que pode impactar diretamente a análise de perfil e renda.



FAIXA ETÁRIA PREDOMINANTE

21 - 42 Anos

A base é composta majoritariamente por **adultos jovens**, com maior propensão ao consumo de crédito.



MEDIANA SALARIAL

R\$4.416,00

A **média é distorcida por outliers**, como valores extremos de até R\$ 1.560.000,00, que não refletem o perfil geral da base.



TIPO DE EMPRÉSTIMO MAIS SOLICITADO

99,4%

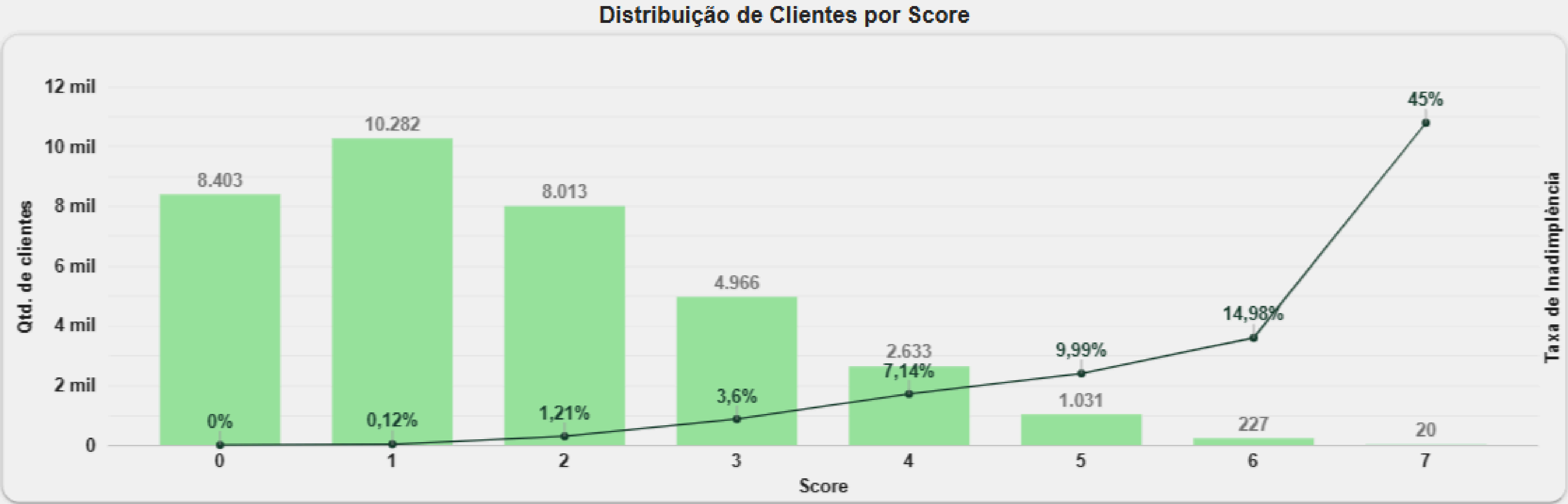
A maioria dos registros está classificada como **"Other"**, limitando a interpretação sobre a finalidade do crédito solicitado.



---

# Score Tradicional: Metodologia e Resultados

SCORE INICIAL: ESTRUTURA E PREMISSAS



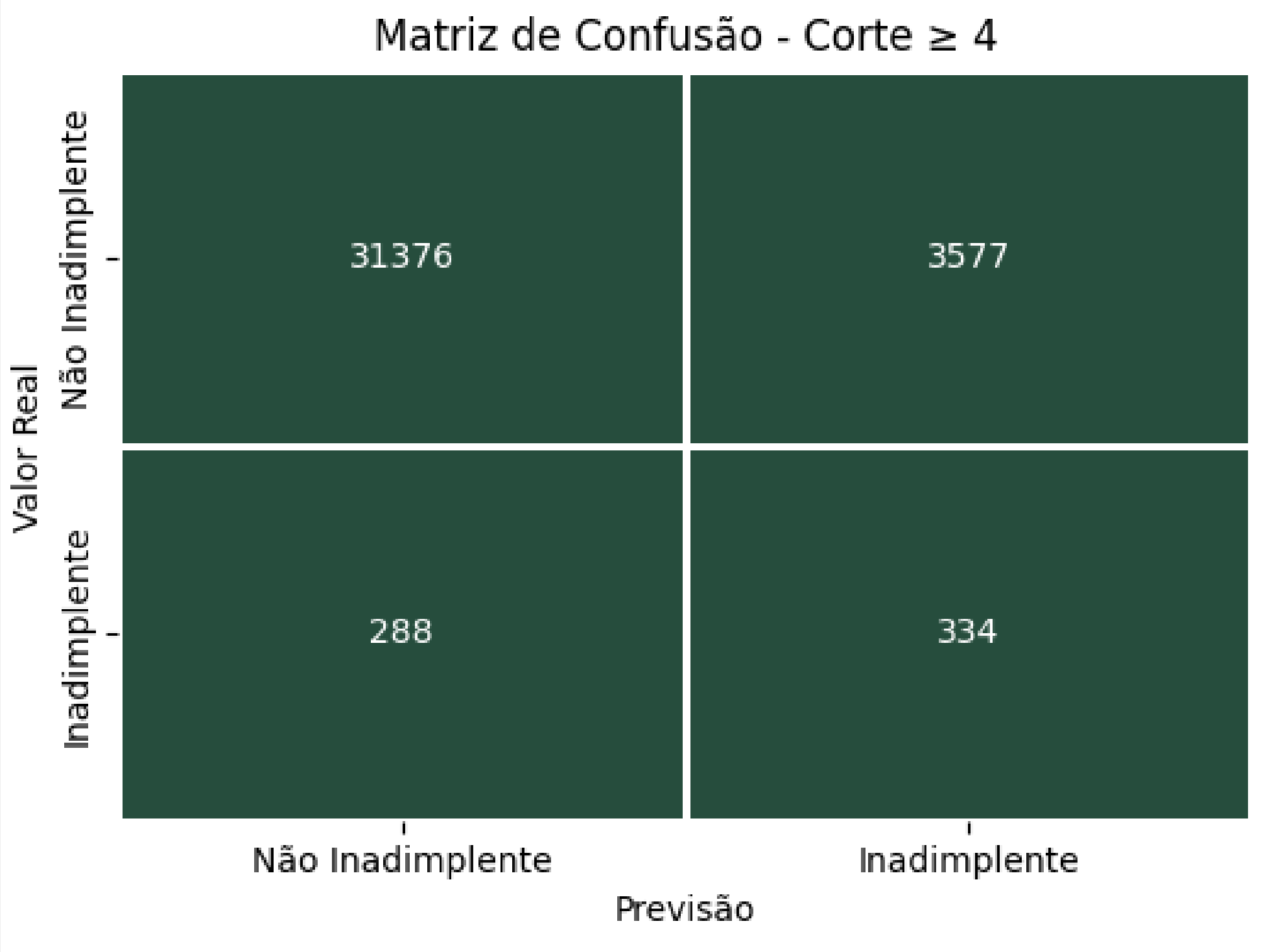
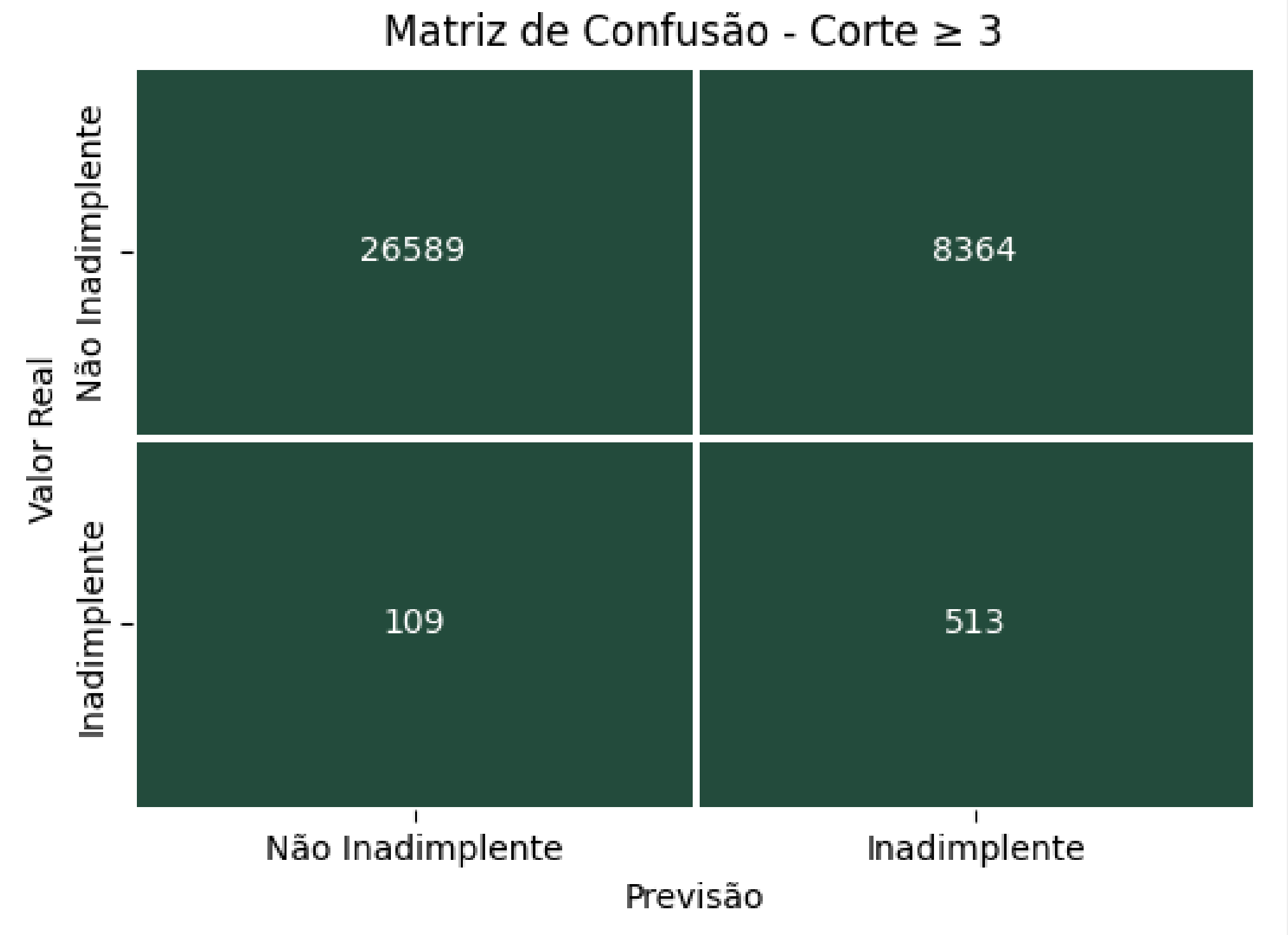
- 1

Identificar os quartis com maior **taxa de inadimplência** na base, através do cálculo de **risco relativo**.
- 2

Uma **pontuação binária** por variável foi elaborada, **marcando os clientes presentes nesses grupos** de maior risco.
- 3

Cada cliente **recebeu uma soma de pontos**, proporcional à quantidade de **características de risco**.

# Matriz de Confusão - Score Tradicional



Acurácia	76,2%
Recall	82,4%
Precisão	5,7%

Acurácia	89,1%
Recall	53,6%
Precisão	8,5%



# Matriz de Confusão - Score Tradicional



Acurácia	76,2%
Recall	82,4%
Precisão	5,7%

Acurácia	89,1%
Recall	53,6%
Precisão	8,5%

SCORE AGRAVANTE: EXPONDO RISCOS CAMUFLADOS

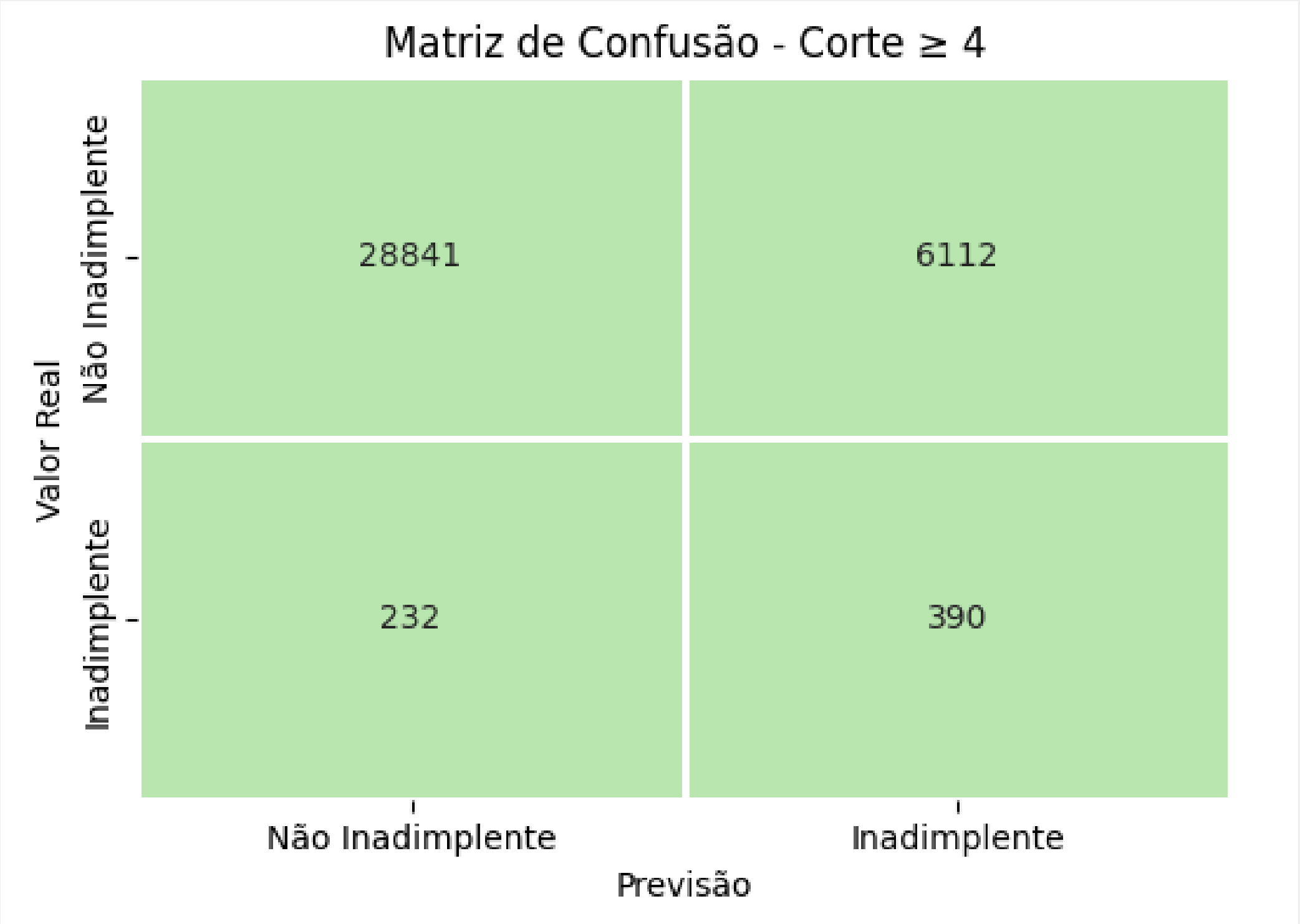
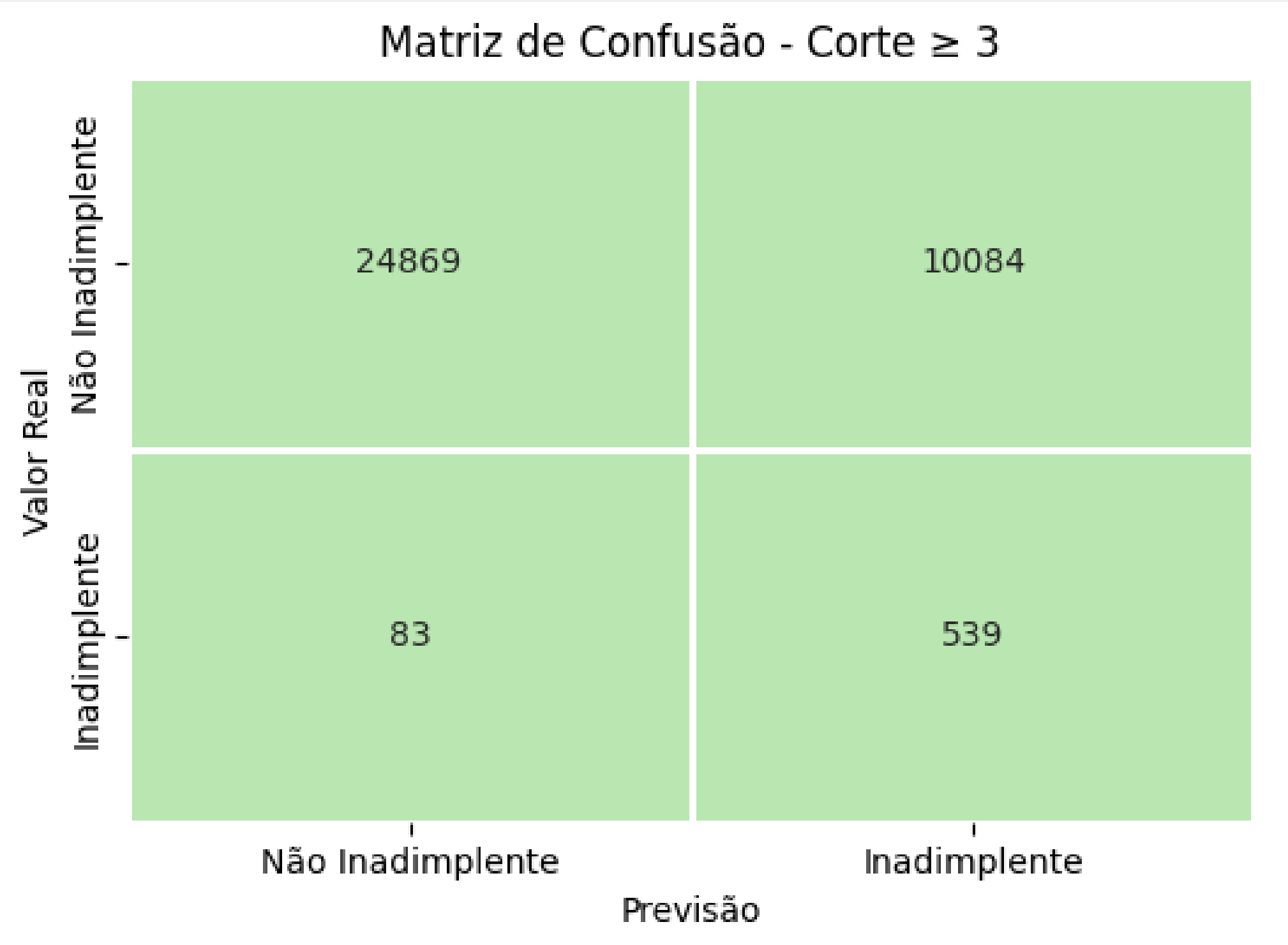
- Perfis de risco camuflados:**  
Alguns clientes não atingem a pontuação mínima do score tradicional, mas apresentam sinais claros de risco. Como o uso excessivo de crédito, múltiplos empréstimos em aberto, atrasos recorrentes e ausência de informações cruciais.
- Desequilíbrio entre recall e precisão:**  
Cortes mais baixos identificavam mais inadimplentes (alta sensibilidade), mas com muitos falsos positivos. Cortes mais altos eram mais precisos, mas deixavam inadimplentes passarem.
- Trade-off estratégico na classificação:**  
A escolha do ponto de corte envolve uma decisão de negócio: priorizar a redução de inadimplência ou evitar reprovações injustas. O Q5 entra como uma camada complementar para equilibrar esse jogo.

Preenchidos

Faltantes



# Matriz de Confusão - Score com Risco Agravado



Acurácia	71,4%
Recall	86,6%
Precisão	5,07%

Acurácia	82,1%
Recall	62,7%
Precisão	5,1%

# Matriz de Confusão - Score com Risco Agravado



## TRADE-OFF

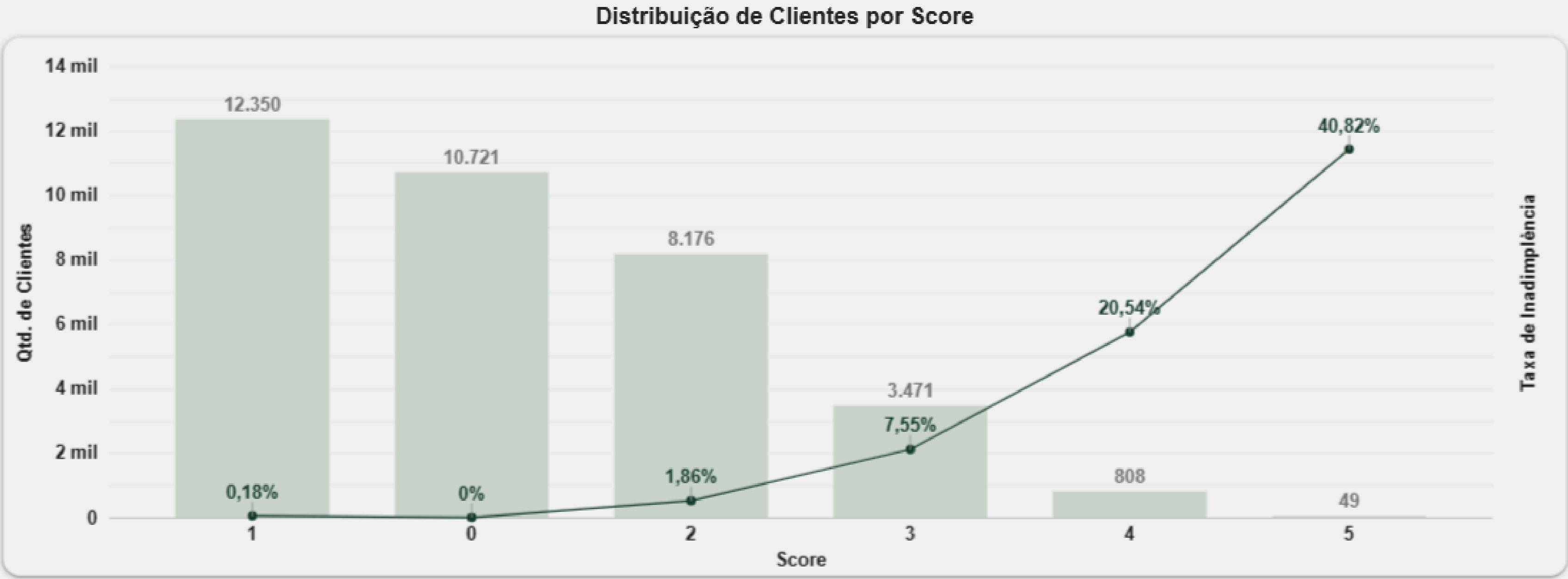
OUTRO TRADE-OFF? SERÁ QUE ALGUMAS VARIÁVEIS  
ESTAVAM INFLUENCIANDO NEGATIVAMENTE O  
MODELO, **COMPROMETENDO SUA CAPACIDADE  
PREDITIVA?**

# SCORE REFINADO: PERFORMANCE COM MENOS VARIÁVEIS

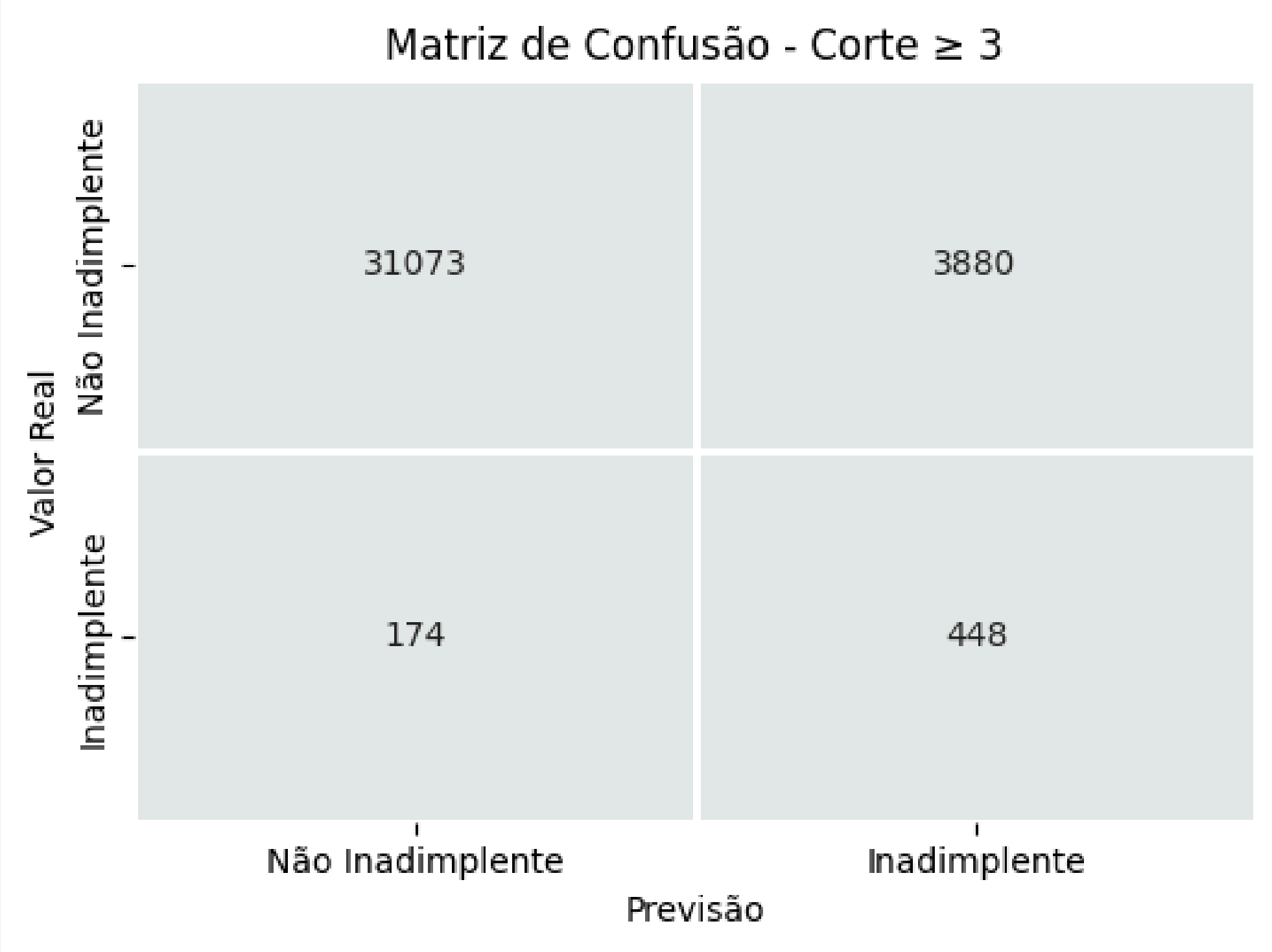
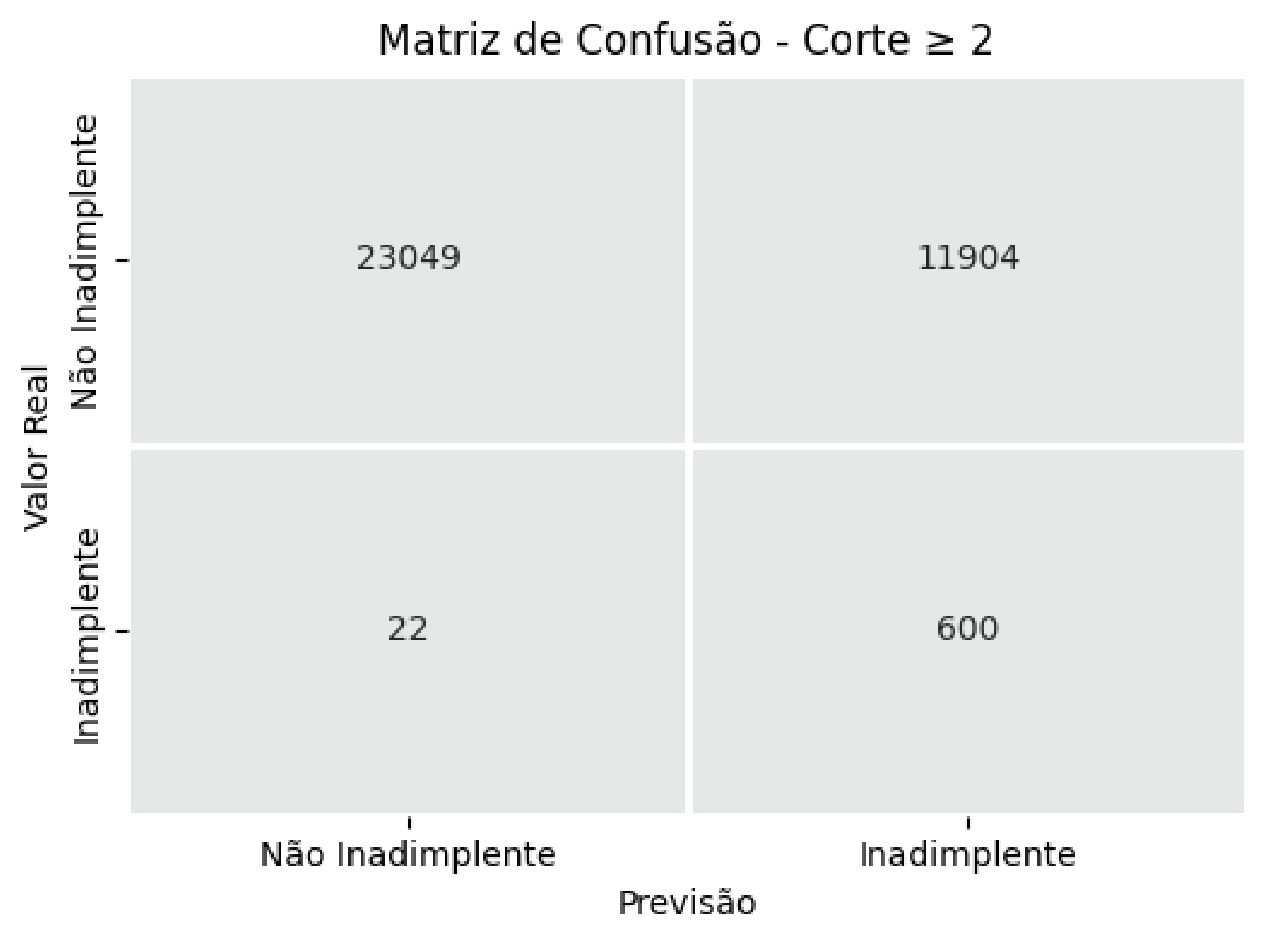
- **POR QUE CRIAR UM SCORE REFINADO?**

Com a análise, foi possível identificar que algumas variáveis estavam camuflando o perfil real de risco, gerando muitos falsos positivos e prejudicando o equilíbrio entre as métricas.

- **Variáveis removidas:**
  - **Dependentes** (200 inadimplentes)
  - **Faixa de salário** (201 inadimplentes)
- **Motivo:**
  - Alta taxa de falsos positivos
  - Baixa contribuição preditiva
  - Pior desempenho em testes com matriz de confusão



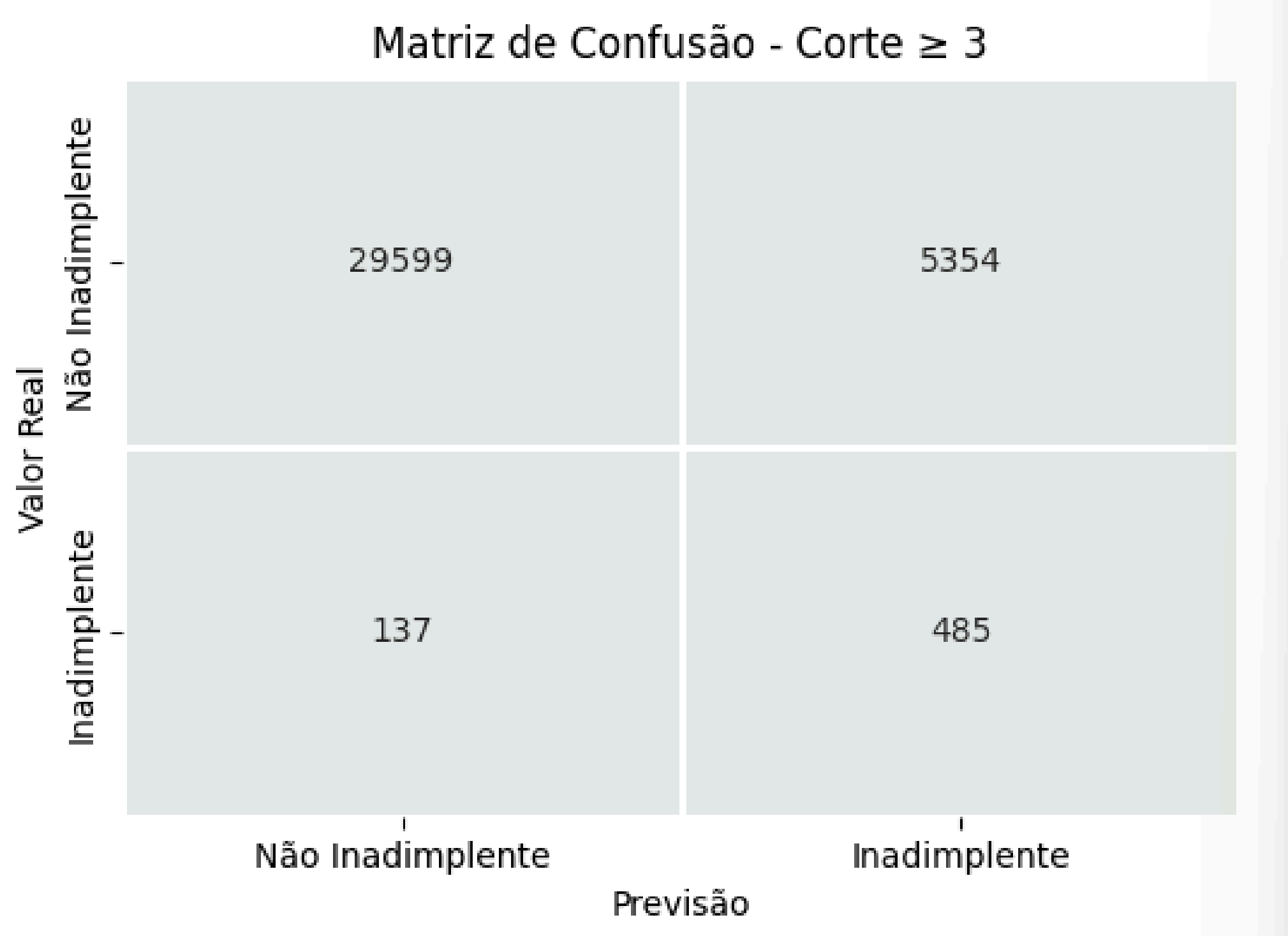
# Matriz de Confusão - Score Refinado



Acurácia	66,2%
Recall	96,4%
Precisão	4,8%

Acurácia	88,3%
Recall	72,0%
Precisão	10,3%

# Matriz de Confusão - Score Refinado com Risco Agravado



Acurácia

85%

Recall

77,9%

Precisão

8,3%

# CONCLUSÕES

- Dados faltantes aumentaram falsos positivos e dificultaram a classificação de bons pagadores.
- Variáveis com baixa força preditiva, quando removidas, podem melhorar o equilíbrio entre precisão e recall.
- A criação da flag Q5 trouxe uma nova camada de análise, revelando perfis camuflados com alto risco apesar do score tradicional baixo.
- O modelo evoluiu para uma versão mais interpretável e eficaz, focando no risco real e não apenas na ausência de dados.





# Recomendações Estratégicas

<p><b>01 - Modelo de Score</b></p> <p>Adotar o <b>modelo refinado com corte <math>\geq 3 + Q5</math></b>, que oferece melhor equilíbrio entre <b>acurácia (85%)</b>, <b>recall (77,9%)</b> e <b>precisão (8,3%)</b>.</p>	<p><b>02 - Idade (21–42 anos)</b></p> <p><b>Jovens têm maior risco de inadimplência</b>, possivelmente por menor estabilidade financeira.</p> <p><b>Estratégia:</b> crédito progressivo, limites menores, educação financeira e acompanhamento mais próximo.</p>	<p><b>03 - Renda (até R\$ 3.400)</b></p> <p><b>Baixa renda = risco 4x maior.</b> Capacidade de pagamento é fator-chave.</p> <p><b>Estratégia:</b> comprovação rigorosa de renda, política de crédito compatível, incentivo à sustentabilidade financeira.</p>	<p><b>04 - Dependentes (0 ou +2)</b></p> <p><b>Risco elevado nos extremos.</b> Pode indicar sobrecarga ou isolamento financeiro.</p> <p><b>Estratégia:</b> avaliar composição familiar + renda, atenção a perfis com vulnerabilidades invisíveis.</p>
<p><b>05 - Poucos empréstimos (1 a 5)</b></p> <p><b>Menor histórico = maior risco.</b> Pode indicar novos clientes ou uso limitado do crédito.</p> <p><b>Estratégia:</b> análise reforçada para iniciantes, foco em educação financeira e renegociação antes de novas concessões.</p>	<p><b>06 - Endividamento ( &gt; 46%)</b></p> <p><b>Quanto mais endividado, maior o risco.</b></p> <p><b>Estratégia:</b> monitoramento contínuo, renegociação preventiva e gestão ativa do risco antes de requalificar.</p>	<p><b>07 - Uso do limite de crédito (&gt; 52%)</b></p> <p><b>Uso extremo = grupo mais arriscado.</b></p> <p><b>Estratégia:</b> alertas automáticos, bloqueios preventivos, renegociação e análise manual antes de renovação.</p>	<p><b>08 - Histórico de atraso (&gt; 0)</b></p> <p>Atrasos prévios = risco 559x maior. <b>Mesmo valores baixos indicam alta propensão à inadimplência.</b></p> <p><b>Estratégia:</b> monitoramento em tempo real, reclassificação no score e campanhas de negociação proativa.</p>



**OBRIGADA!**