



Análise da PaySmart Solutions

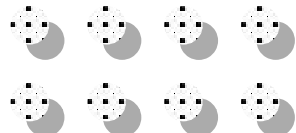
João Vítor R. S. Leite

07/09/25

Análise da PaySmart Solutions

Introdução


- 01 Overview das últimas 200 transações feitas;
- 02 Melhorias do Sistema de Análise;
- 03 Recomendações adicionais.



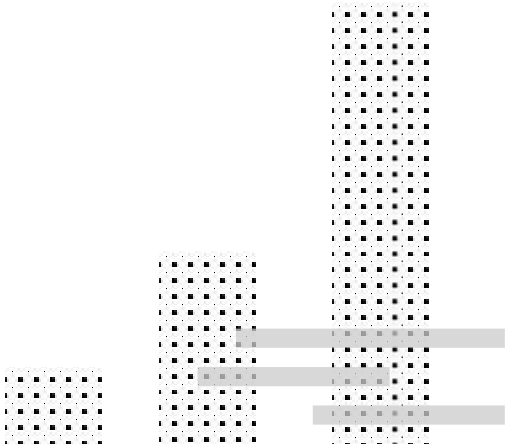
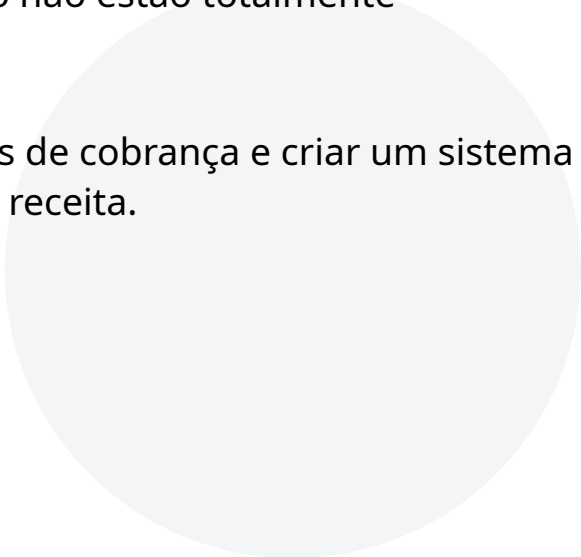


Introdução

- A taxa de aquisição é o principal produto da PaySmart Solutions.
- Falhas na cobrança geram perda direta de receita.
- As regras de isenção não estão totalmente documentadas.




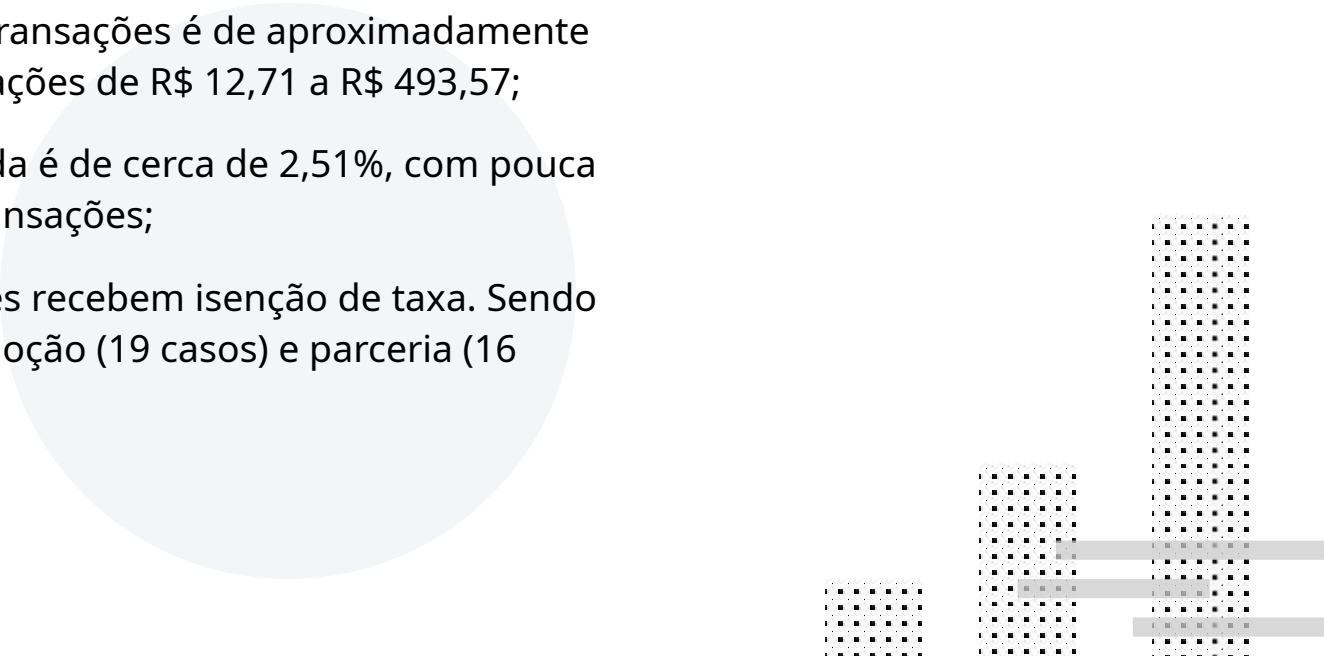
Objetivo: Identificar erros de cobrança e criar um sistema de monitoramento que garanta conformidade e proteja a receita.





01 Overview das últimas 200 transações feitas

Esta análise teve como objetivo identificar lacunas operacionais que afetam a saúde financeira da empresa e indicar soluções. Os parâmetros são:

- O ticket médio das transações é de aproximadamente R\$ 247,16, com variações de R\$ 12,71 a R\$ 493,57;
 - A taxa média aplicada é de cerca de 2,51%, com pouca variação entre as transações;
 - 17,5% das transações recebem isenção de taxa. Sendo motivadas por promoção (19 casos) e parceria (16 casos).
- 
- 

01 Overview das últimas 200 transações feitas

Outras análises foram feitas para complementar o diagnóstico e estas, por sua vez, indicaram que:

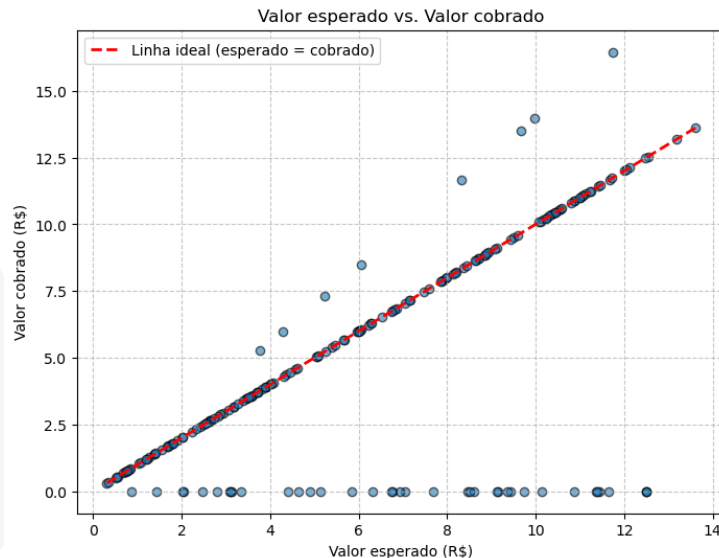
- Na maioria das transações (77%) o valor cobrado é exatamente o que era esperado. Entretanto, há casos em que o valor cobrado foi menor, ou maior, do que o esperado, indicando que os erros não são sistêmicos;
- De cada 5 transações, uma tem a taxa de cobrança incorreta (21,5%). Esse é o principal problema a ser investigado;
- A soma do que foi cobrado a mais (R\$ 23,60) é muito menor do que o total não cobrado (R\$ -267,45). O impacto financeiro líquido e negativo traduz a falha técnica nas cobranças de taxa.

01 Overview das últimas 200 transações feitas

Por fim, a análise dos **outliers** nos mostra que:

O gráfico de dispersão mostra vários pontos afastados da linha ideal (esperado = cobrado). Visualmente temos bem mais de 7 outliers, indicados pelo método com o desvio-padrão. O método do desvio-padrão marca como anomalia apenas valores além de 3 desvios da média.

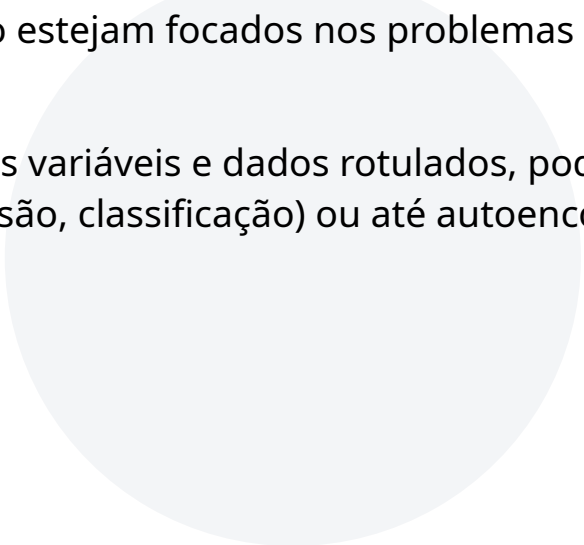
Observamos uma concentração de pontos na linha do Valor Cobrado igual a R\$ 0,00. Esses casos são os outliers mais relevantes, pois representam transações em que a cobrança falhou totalmente. Como existem muitos desvios pequenos e alguns grandes, a média e o desvio-padrão acabam ficando elevados, o que faz com que o método estatístico identifique menos outliers do que realmente existem.





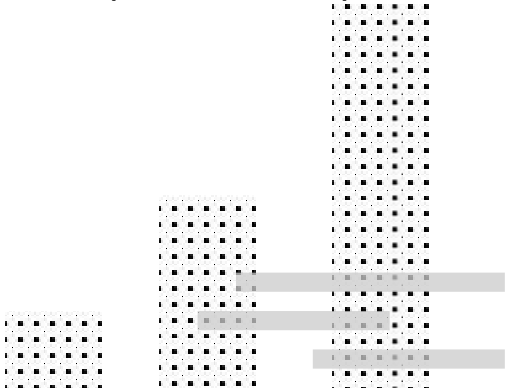
02 Melhorias do Sistema de Análise

Com os dados atuais, a melhor técnica de Machine Learning encontrada para apoiar a análise é o Isolation Forest, pois ele não depende de rótulos e consegue detectar padrões de anomalias automaticamente.



Essa abordagem nos permitiu identificar as falhas de forma automatizada, garantindo que nossos esforços para a correção estejam focados nos problemas que causam o maior impacto para o negócio.

Em um cenário com mais variáveis e dados rotulados, poderíamos avançar para modelos supervisionados (regressão, classificação) ou até autoencoders para detectar padrões complexos.



02 Melhorias do Sistema de Análise

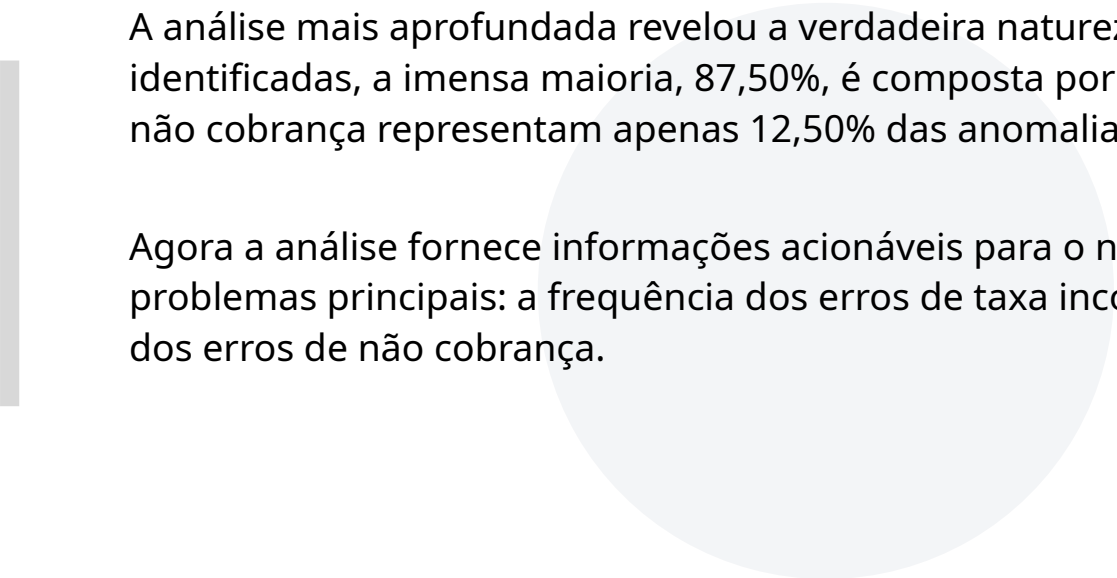
O script desenvolvido permite que ele faça as seguintes ações:

- **Aprende o que é normal:** Em vez de usarmos regras manuais ("se o valor for X, é um erro"), o algoritmo aprende, de forma automática, o que é um comportamento de transação padrão no nosso negócio.
- **Identifica o que é diferente:** Ele marca como "anomalia" qualquer transação que se desvie significativamente desse padrão. Isso nos permite encontrar problemas que poderiam passar despercebidos.
- **Foca nos erros reais:** O script é aprimorado para que ele ignore transações isentas que foram processadas corretamente. Com isso, o algoritmo se concentra apenas nos erros que realmente importam, resultando em uma análise muito mais precisa.



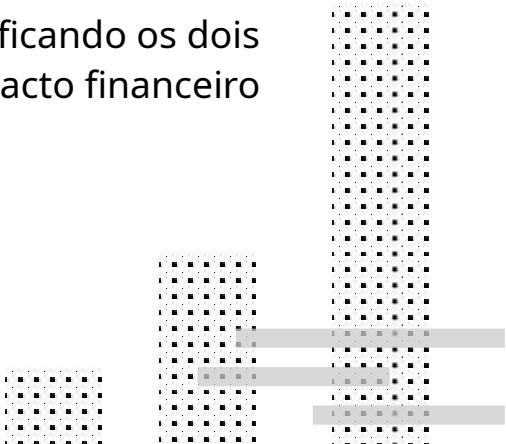
02 Melhorias do Sistema de Análise

O novo modelo de análise se mostrou significativamente mais preciso ao se concentrar exclusivamente nos erros reais de negócio. Removendo as transações isentas a porcentagem de falhas detectadas caiu de 8% para apenas 5%. Essa otimização valida que o modelo agora está mais alinhado com o que o negócio considera um problema, permitindo uma análise mais focada e relevante.



A análise mais aprofundada revelou a verdadeira natureza dos erros. Entre as anomalias identificadas, a imensa maioria, 87,50%, é composta por falhas de taxa incorreta. E que erros de não cobrança representam apenas 12,50% das anomalias.

Agora a análise fornece informações acionáveis para o negócio, identificando os dois problemas principais: a frequência dos erros de taxa incorreta e o impacto financeiro dos erros de não cobrança.



Alternativas



- A partir desse ponto, o foco não é mais encontrar falhas, mas sim investigar suas causas-raiz e priorizar as soluções, seja para reduzir a alta frequência de erros de taxa ou mitigar as grandes perdas causadas por erros de não cobrança;
- Para investigar a causa-raiz dos problemas, o ideal é avançar para o aprendizado supervisionado, utilizando algoritmos de classificação. Modelos como Árvores de Decisão ou Random Forests são particularmente úteis, pois, além de prever se uma transação é um erro, eles revelam as regras que levam a essa classificação;
- Alternativamente, você pode usar algoritmos de agrupamento (clustering), como o K-Means, para categorizar as anomalias detectadas e identificar se existem padrões ocultos entre elas, como um cluster de erros que ocorre apenas em um determinado dia da semana ou para um produto específico.

Obrigado

João Vítor R S Leite

E-mail: joaovictorrs194@gmail.com

Linkedin: <https://www.linkedin.com/in/joao-rodrigues-leite/>

Github: <https://github.com/joaorsleite>

