



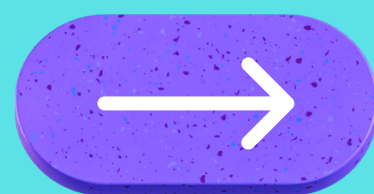
Itaú - Estágio

Case Solution

Proposta de solução para o desafio

Lucas de Araujo Souza

contact.lucasouza@gmail.com



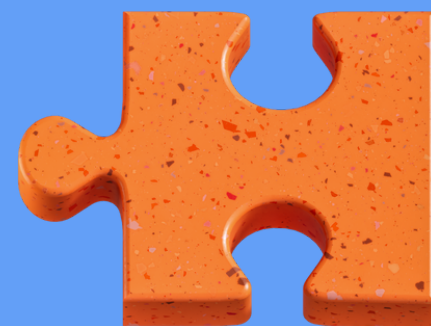


Itaú - Estágio

Etapas



Entendendo o
Problema



EDA



Propondo uma
Solução



Bônus:
Machine
Learning

[Voltar ao índice](#)



Primeiros passos



- O que é a coluna database?
- A coluna 'database' é uma identificação para monitorar os meses de trabalho com um pacote de carrinhos que combinada com a coluna 'meses de trabalho' forma uma serie temporal
- Entender e analisar a junção dessas identificações para determinar o cenário que resulta em menos danos aos carrinhos.

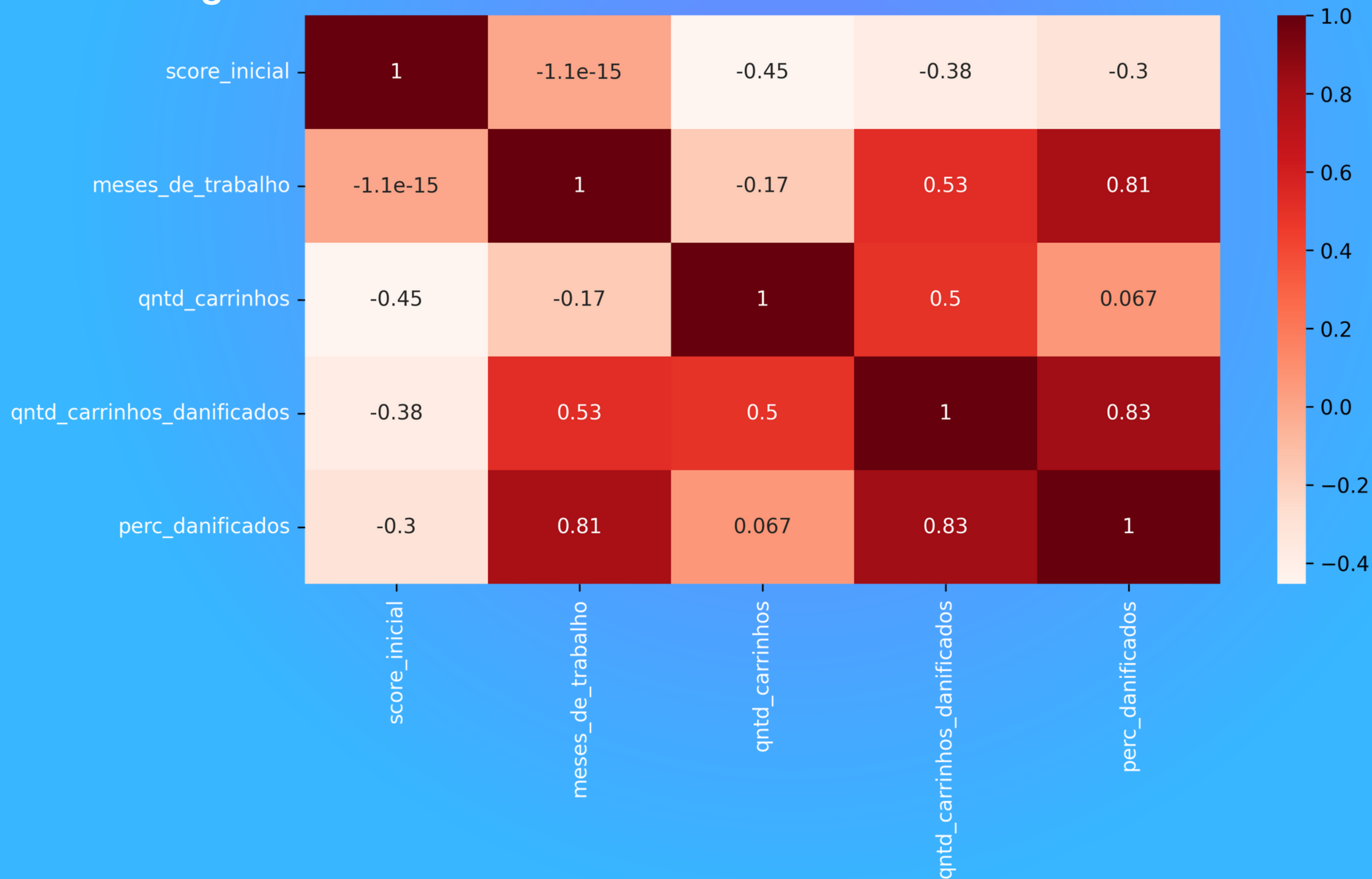
O padrão dos dados segue o seguinte:

- Identificação (20330X) Onde o X varia de 1 a 9 ->
 - Score Inicial x Final (Sempre em uma diferença $\Delta = 50$)
 - Série temporal em meses
 - Variação na quantidade total de carrinhos dispostos ao trabalho

seguir serão criadas novas colunas (engenharia de atributos) afim de facilitar a análise



Mapa de Correlação



EDA

Exploratory Data Analysis

Não teve necessidade de um tratamento de dados.

Adição de colunas (engenharia de atributos) afim de nos auxiliar na análise.

Todos os passos sempre justificados.

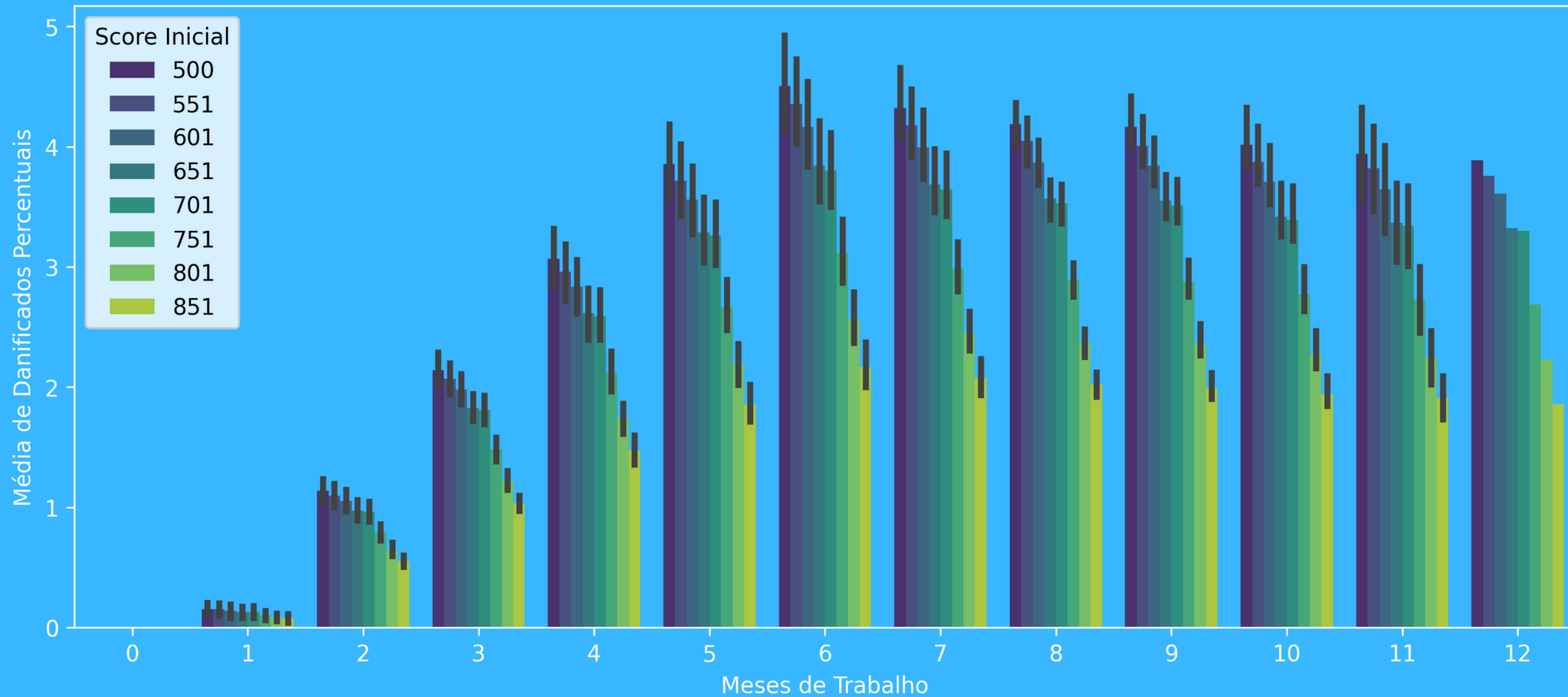
Observações:

A quantidade de carrinhos em trabalho não influencia diretamente no percentual de carrinhos danificados.

O aumento no tempo de uso de um carrinho está positivamente relacionado à porcentagem de carrinhos danificados.

O score inicial do carrinho tem uma correlação negativa com o percentual de danos, sugerindo que carrinhos com pontuações mais altas têm menor probabilidade de serem danificados.

O aumento no score inicial do carrinho está associado a uma diminuição na quantidade de carrinhos em operação, indicando possíveis desafios na obtenção de carrinhos com pontuações mais altas.



Validações

Em seguida foram criados diversos gráficos afim de validar algumas de nossas suposições



A análise revelou uma forte correlação entre o score do carrinho e sua durabilidade.



Carrinhos com pontuações mais altas demonstram uma menor probabilidade de sofrer danos ao longo do tempo.



Carrinhos com pontuações mais altas também estão em menor quantidade, possivelmente podem ter um custo maior de produção ou aquisição



Proposta de Solução

Objetivos do Rodízio:

- Minimizar danos aos carrinhos.
- Aprimorar a performance de vendas.
- Reduzir custos com manutenções.



Score	Tempo de trabalho na praia
500-600	± 3 meses
601-750	± 4 meses
751-851	± 6 meses

Desafio e Contrapartida:

- Custo adicional associado ao transporte constante





Itaú - Estágio

Bônus

Aplicando Machine Learning para prever o número de carrinhos danificados de acordo com o score inicial e o tempo de trabalho

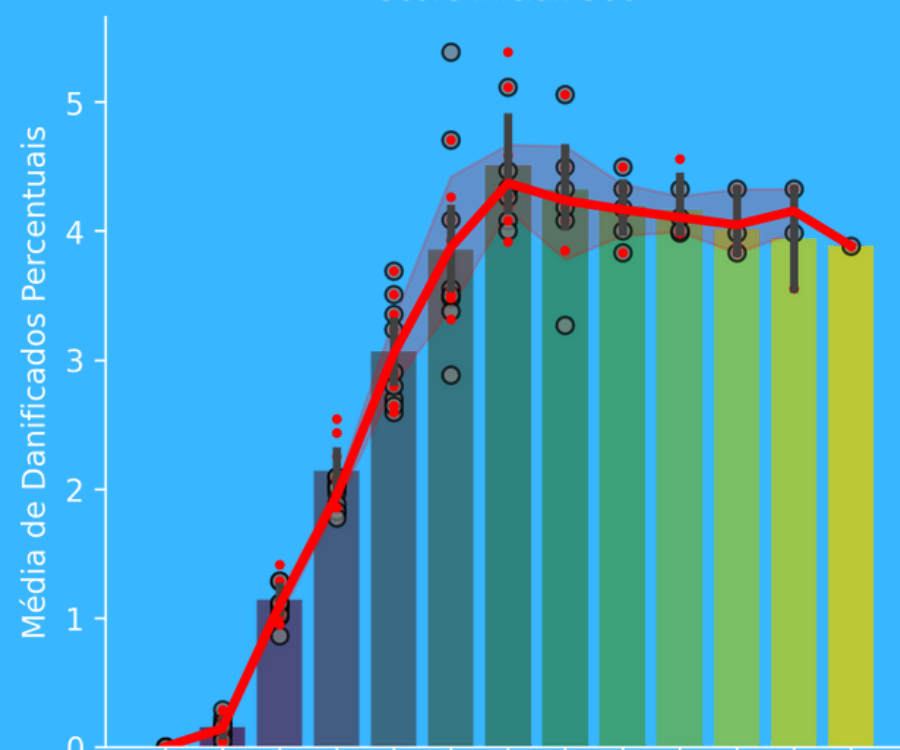
Discussão dos Resultados ML

O modelo alcançou uma acurácia de 100%. Mas não prevê corretamente os dados, os motivos podem ser:

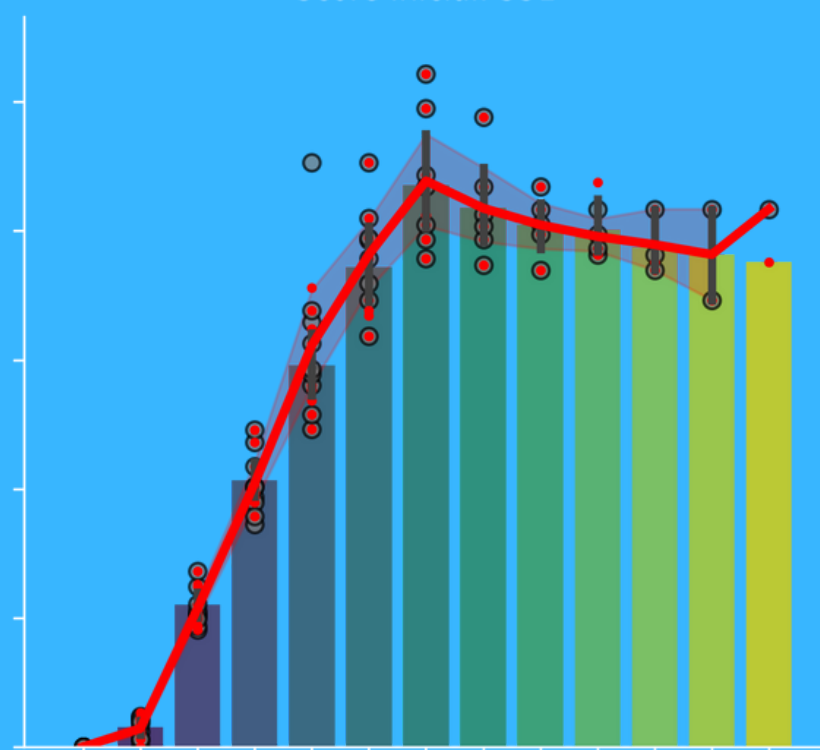
- Indica a possibilidade de overfitting
- A limitação do treinamento com poucos dados pode ter levado o modelo a memorizar padrões específicos.
- A falta de diversidade nos dados pode ter levado o modelo a memorizar em vez de aprender, comprometendo sua capacidade de generalização.



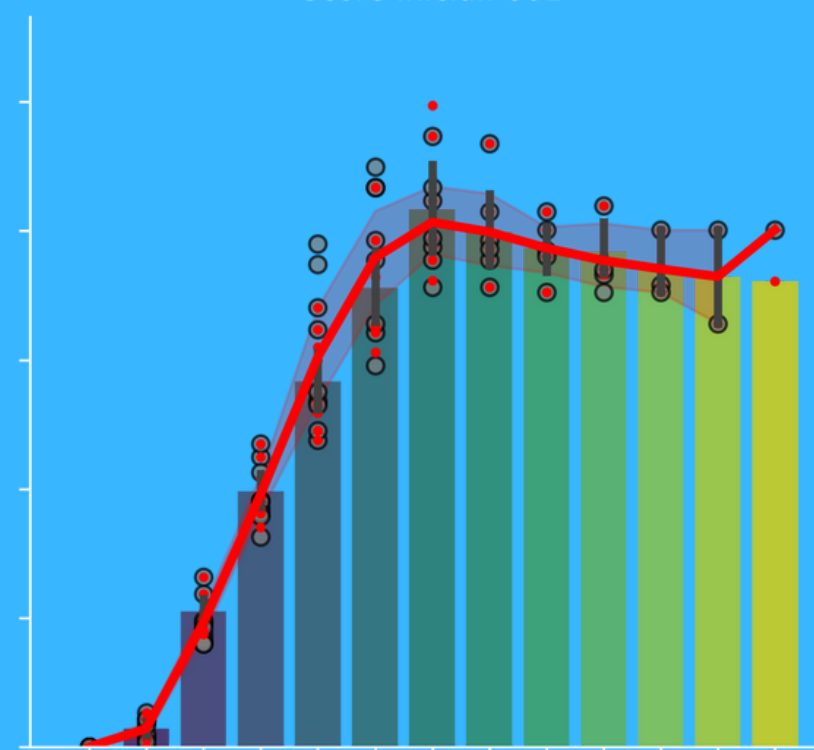
Score Inicial: 500



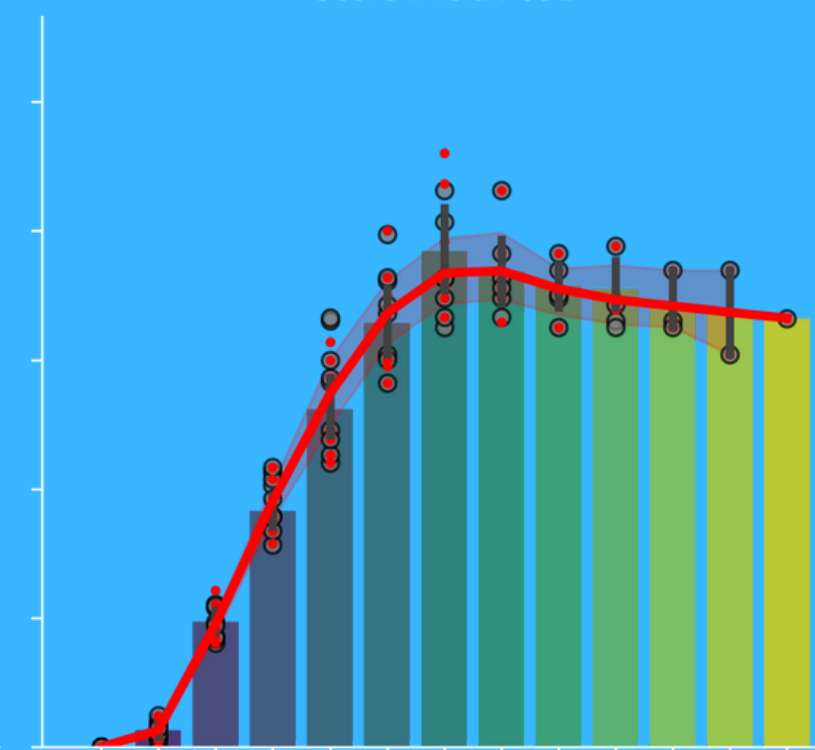
Score Inicial: 551



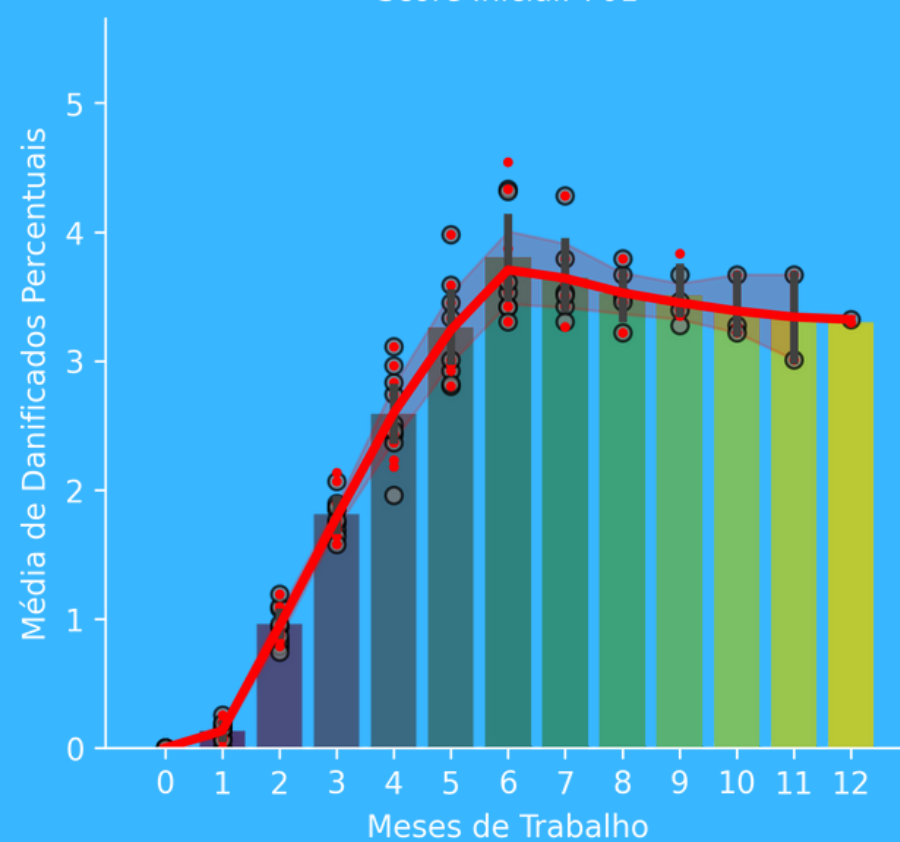
Score Inicial: 601



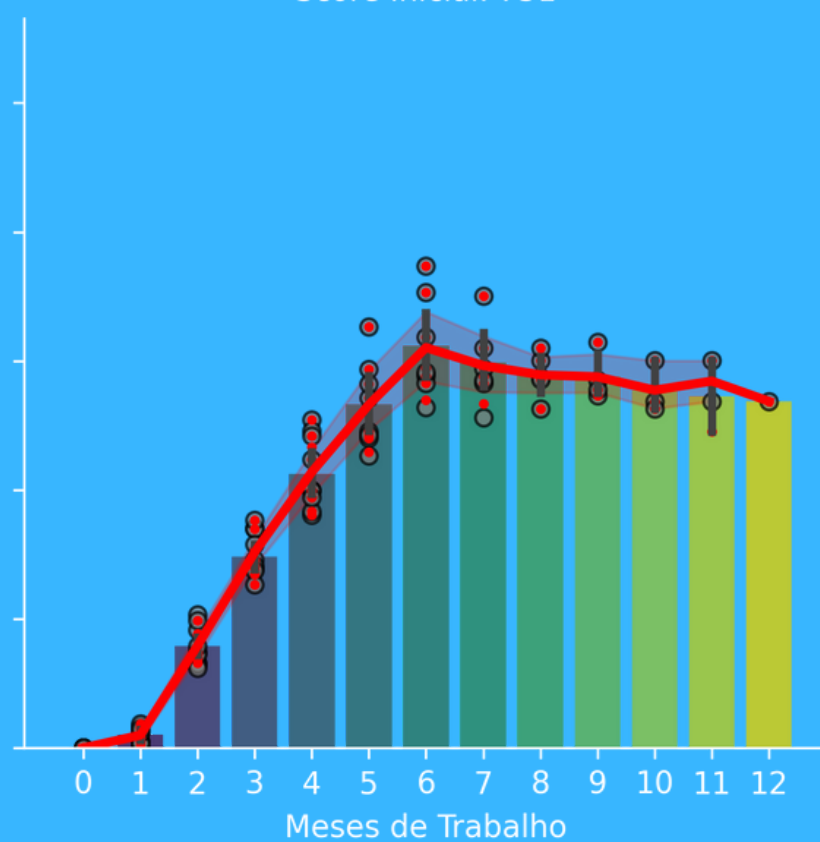
Score Inicial: 651



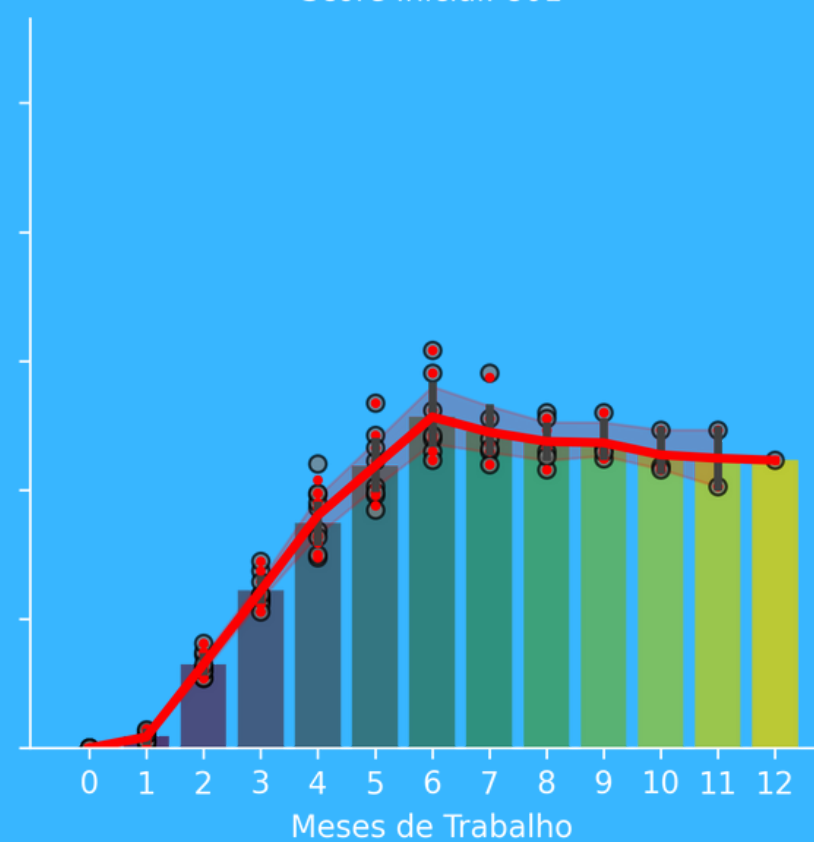
Score Inicial: 701



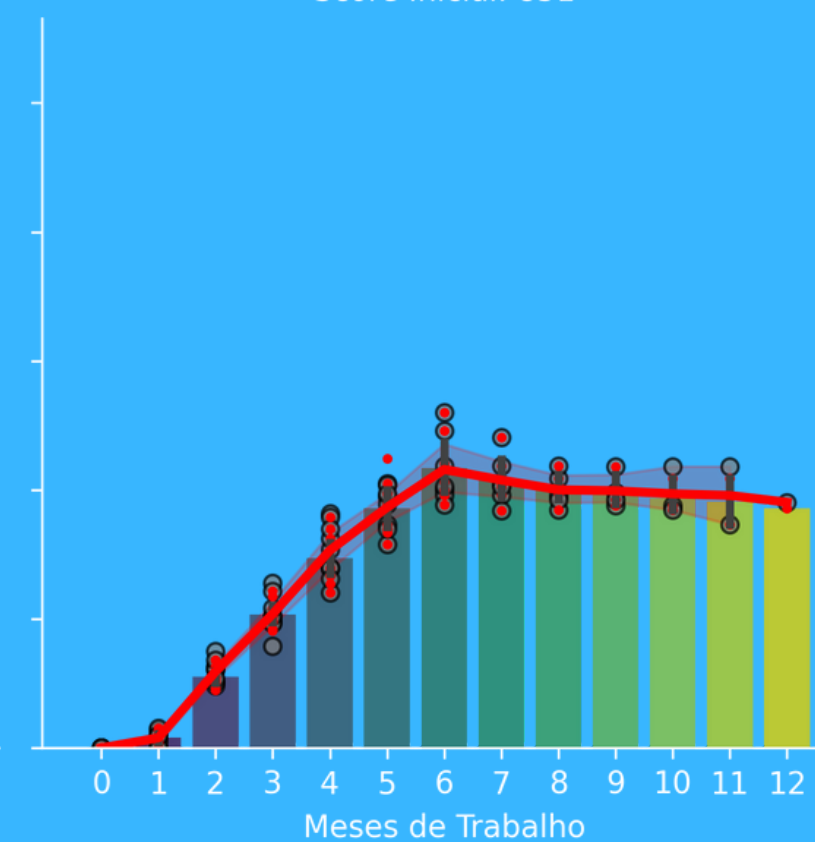
Score Inicial: 751



Score Inicial: 801



Score Inicial: 851



○ Previsão
● Esperado



Itaú - Estágio



Obrigado por acompanhar até o fim

As queries e transformações também foram disponibilizadas em linguagem SQL no GitHub. Para mais detalhes ambos os códigos estão devidamente comentados.

Todas as ações tomadas estão justificadas.

[Voltar ao índice](#)

