

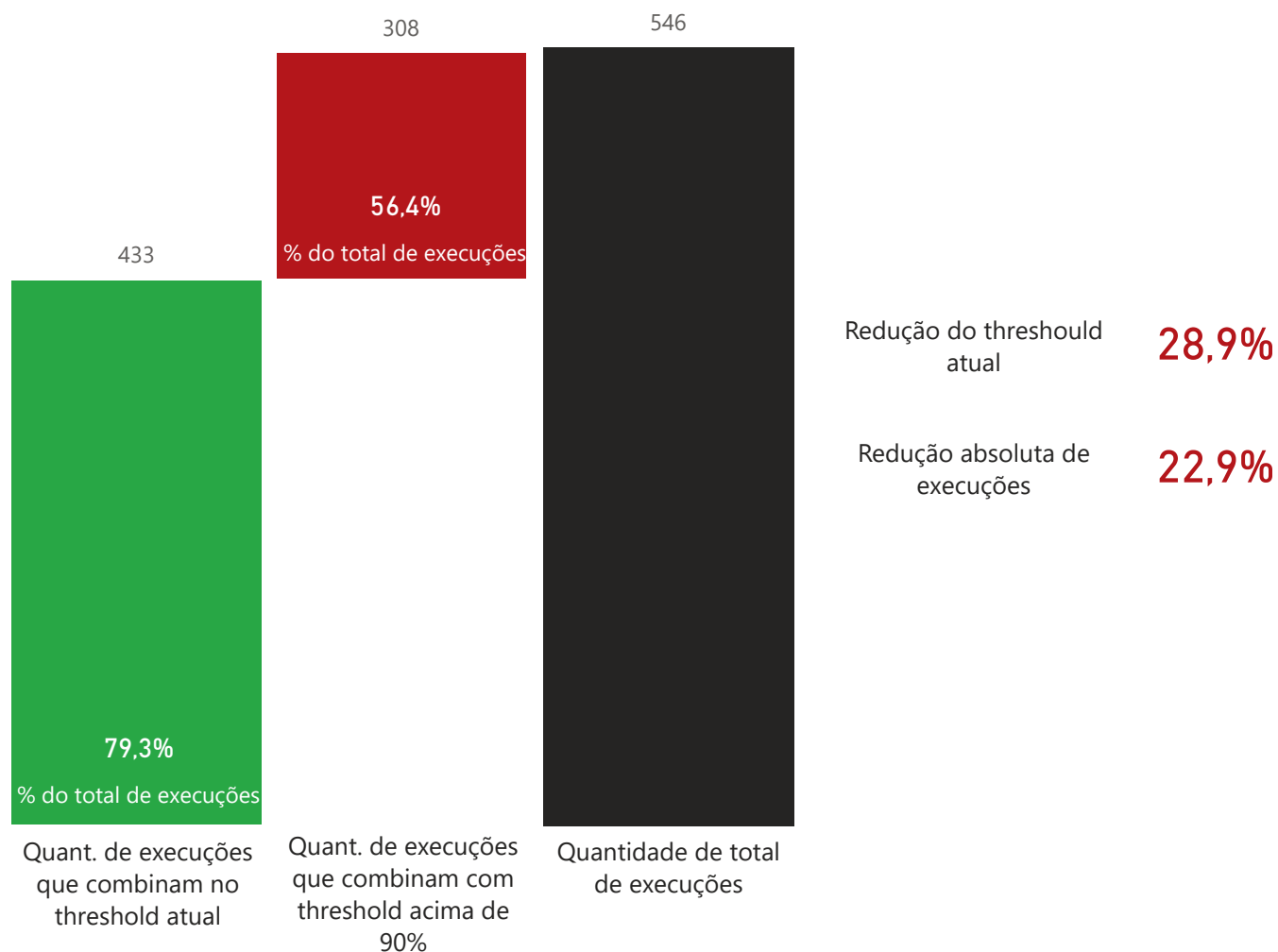
# Análise de Efetividade e Falhas na Biometria Facial



# Análise de Efetividade e Falhas na Biometria Facial



Como ficaria o índice geral de match com *threshold* maior que 90?



## Análise:

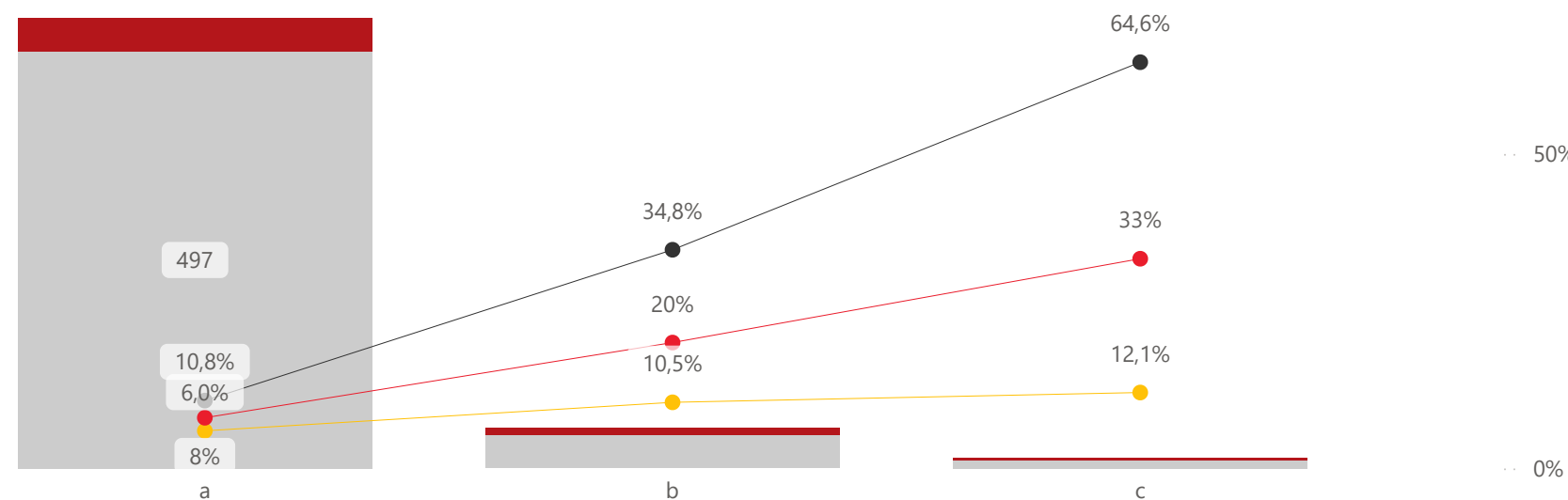
- Ao aumentar o threshold de similaridade da biometria facial para acima de 90%, a efetividade do sistema diminui consideravelmente;
- Ao elevar o threshold para 90%, essa taxa de sucesso cai para 56,4%. Isso representa uma redução relativa de 28,9% em relação ao cenário atual.
- Considerando o total das execuções (546), a mudança do threshold significaria uma queda absoluta na efetividade de 22,9%.
- A análise indica que um threshold acima de 90% tornaria o sistema significativamente mais rigoroso, aumentando os casos rejeitados e, possivelmente, gerando mais falsos negativos.

# Análise de Efetividade e Falhas na Biometria Facial



## Serviços com mais falhas

● Quantidade de execuções ● Quantidade de execuções falhadas ● Intervalo de confiança inferior ● Taxa de falhas ● Intervalo de confiança superior



Provedor	Quantidade de execuções	Quantidade de execuções falhadas	Intervalo de confiança inferior	Taxa de falhas	Intervalo de confiança superior
a	497	40	6,0%	8%	10,8%
b	40	8	10,5%	20%	34,8%
c	9	3	12,1%	33%	64,6%
Total	546	51	7,2%	9%	12,1%

### Análise:

- O Provedor A apresentar a menor taxa de falhas e um intervalo de confiança (IC) bastante estreito, indicando alta estabilidade;
- O Provedor B exibe uma taxa de falhas significativamente maior e um IC mais amplo, o que aponta tanto para um desempenho inferior quanto para a necessidade de investigação adicional das causas de erro.
- O Provedor C, com apenas nove tentativas, registra uma taxa de falhas maior, mas seu intervalo de confiança é tão extenso que não permite conclusões estatisticamente seguras, sugerindo que a amostra é insuficiente.
- No momento, o provedor C é o provedor com mais falhas, entretanto, recomenda-se, ampliar o volume de testes do Provedor C para estreitar seu IC e validar seu comportamento, investigar eventuais gargalos no Provedor B para reduzir falhas e manter o Provedor A como referência de confiabilidade até que novas evidências indiquem mudança.

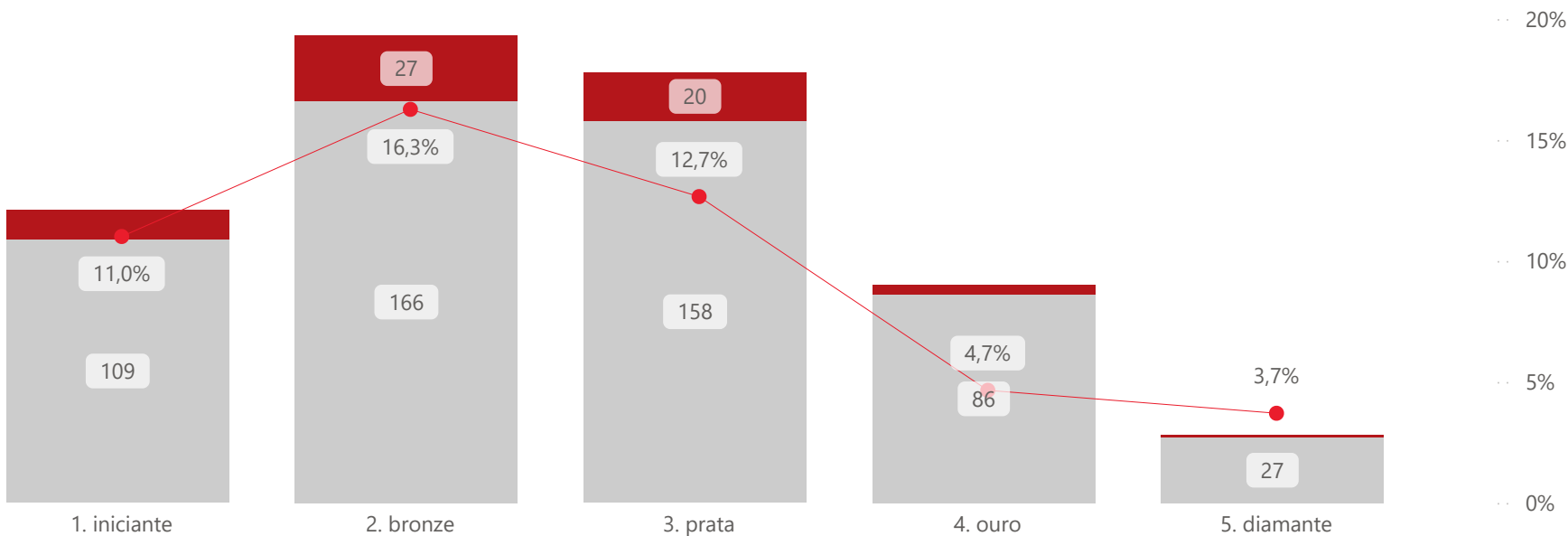


# Análise de Efetividade e Falhas na Biometria Facial



## Categoria com maior índice de sem combinação

● Quantidade de execuções ● Quantidade de execuções sem combinações ● Taxa de execuções sem combinações

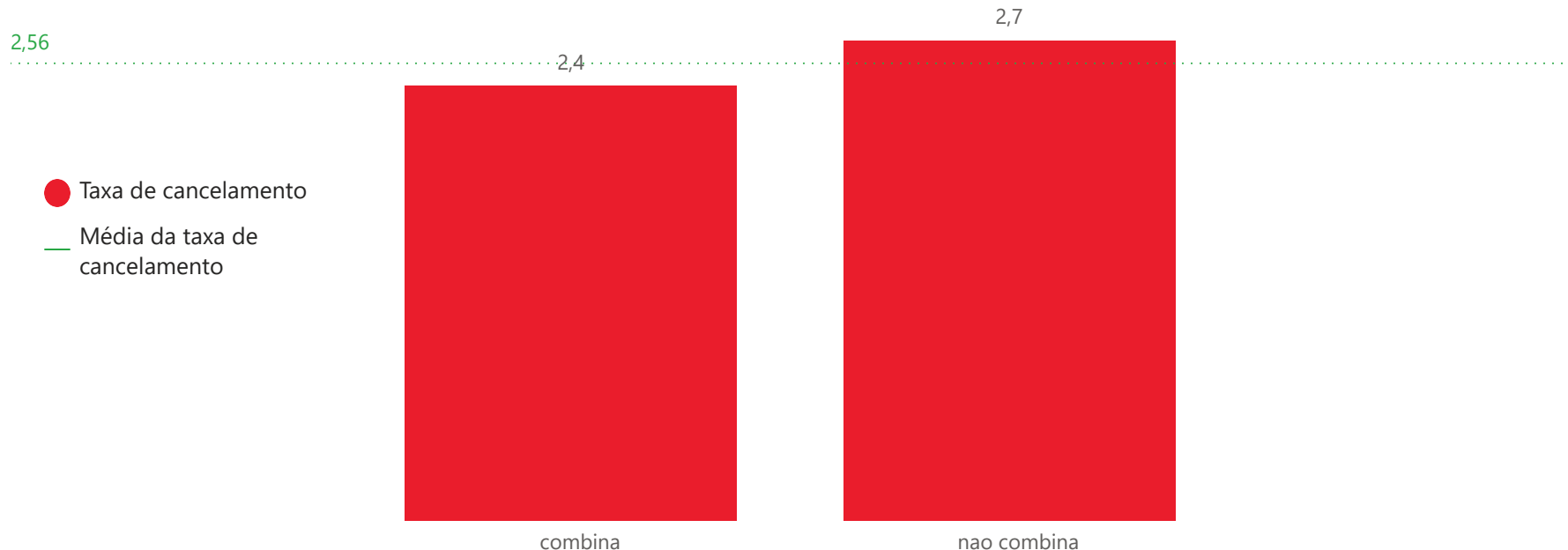


### Análise:

- **Categorias de menor "nível" (bronze e prata)** apresentam as maiores taxas de não reconhecimento, sugerindo maior índice de falha ou menor qualidade das capturas nesses perfis;
- **Entregadores de alto nível (ouro e diamante)** têm desempenho muito superior, com taxas de falha abaixo de 5%.
- A diferença marcada nas taxas indica que pode haver reforço de captura de imagem ou **ajuste de thresholds** especificamente para os perfis com maior índice de falhas.

Nível/Categoria	Quantidade execuções sem combinações	Quantidade execuções	Taxa de execuções sem combinações
1. iniciante	12	109	11,0%
2. bronze	27	166	16,3%
3. prata	20	158	12,7%
4. ouro	4	86	4,7%
5. diamante	1	27	3,7%
Total	64	546	11,7%

## Relação pedido cancelado com biometria

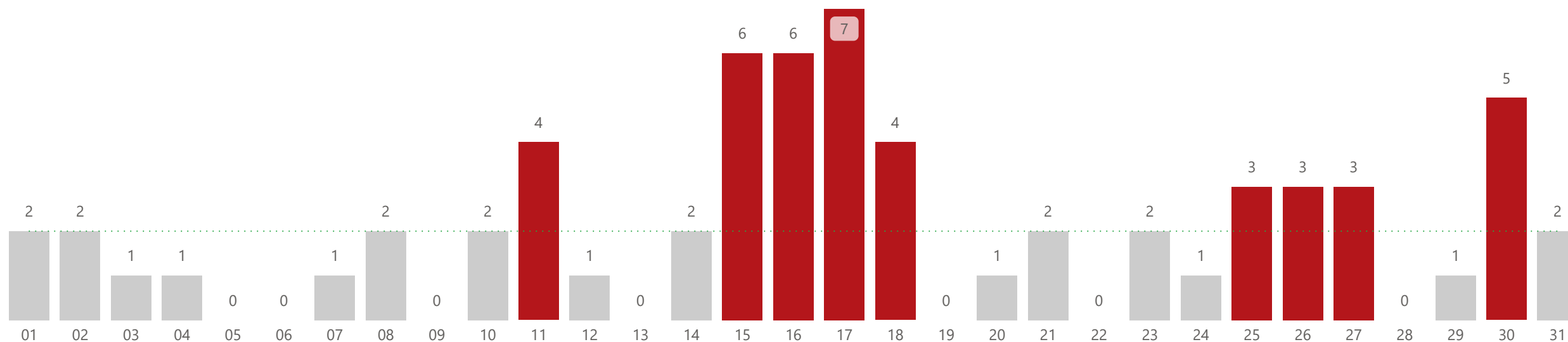


### Análise:

- A diferença é de **0,30** cancelamentos por sessão (2,70 vs. 2,40), um aumento relativo de ~12,5%;
- **Apesar de haver um leve incremento**, essa variação é pequena em termos absolutos e pode ser influenciada por outliers (alguns entregadores com muitos cancelamentos);
- **Conclusão:** não há evidência forte de correlação entre volume de cancelamentos e falhas na validação biométrica.
- **Próximo passo:** rodar um teste estatístico (ex.: correlação de pearson) para checar se essa diferença é significativa e analisar a distribuição via boxplot para entender possíveis outliers;

## Dias de aumento do comportamento indevido “empréstimo de conta”

● Quantidade de execuções sem combinações ● Média de falhas



### Análise:

- Observa-se um aumento significativo de execuções sem combinações, especialmente nos dias 15 a 17, e 30, indicando picos claros de tentativas possivelmente fraudulentas ou indevidas nesses períodos;
- O dia 17 apresenta o maior pico;
- Os dias subsequentes e anteriores aos picos (dias 14-18) também mostram valores elevados comparados aos demais, sugerindo períodos de atividade concentrada ou campanhas específicas para verificar fraudes ou comportamento suspeito;
- Em contraste, na primeira semana (dias 01 a 07), há menos casos identificados, indicando um padrão mais estável ou baixa incidência desse comportamento indevido nesse período;
- Também foi buscado execuções do mesmo motorista em devices diferentes, e execuções em devices iguais repetidas e não foram localizados pontos de atenção;