



Distilo

RELATÓRIO DE ANÁLISE

Análise de Churn

Exemplo de recall de pacientes (dados fictícios).

Gerado em

Cliente

runId

16/08/2026

Relatório demonstrativo · dados sintéticos

demo-churn

Quem chamar de volta: 1.274 pacientes e R\$ 825.644 de doses pagas que ninguém aplicou

Encontrei **1.274 pacientes em situação acionável, 39,6% da sua base de 3.221**, e dentro disso **R\$ 825.644 em doses que já foram pagas e nunca foram aplicadas**. Esse é o número mais desconfortável do relatório, porque é dinheiro que já entrou e serviço que ainda não saiu.

O que a base mostra

Você tem 3.221 pacientes cadastrados, 19.247 doses vendidas e R\$ 10.748.740 de receita registrada entre 26/12/2023 e 30/04/2026. O crescimento é real: as compras saíram de 201 doses em dezembro de 2024 para 2.446 em janeiro de 2026.

A operação também é organizada de um jeito que facilita a análise. Em 99,9% dos casos o paciente compra o esquema inteiro de uma vez: quem compra Hepatite B compra as três doses no mesmo dia, quem compra QDenga compra as duas. Isso quer dizer que cada compra é um compromisso fechado, e cada dose não aplicada é uma promessa em aberto. Chamo cada uma dessas compras de “esquema” no resto do texto.

Data de referência de tudo aqui: 30/04/2026, o último dia que aparece na base. Nenhum cálculo usa a data de hoje.

O problema

Das 19.247 doses vendidas, **1.467 estão marcadas como não aplicadas: 7,6% de tudo que você vendeu**. Elas se dividem em dois problemas bem diferentes, e a diferença importa para o script da recepção.

322 pacientes pagaram o esquema completo e nunca apareceram nem para a primeira dose. São 656 esquemas, 1.107 doses e R\$ 622.914. A espera média desses esquemas já é de 200 dias, e 203 desses 322 pacientes estão esperando há mais de 180 dias. O caso mais extremo é um paciente com 6 doses pendentes e R\$ 6.628 pagos.

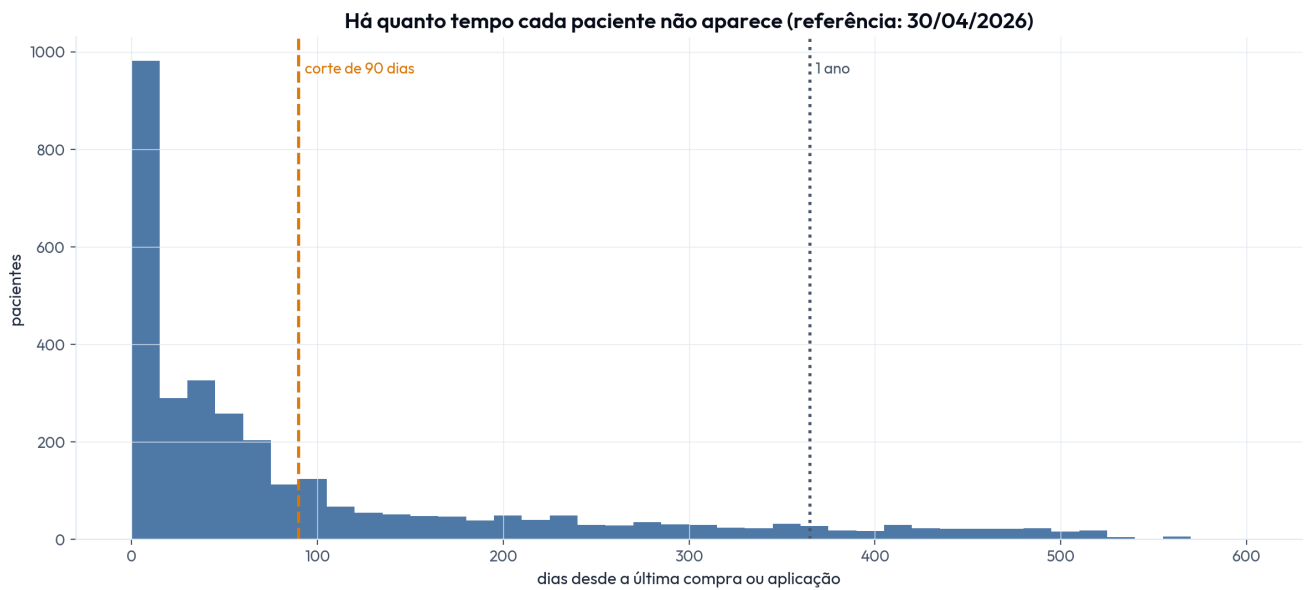
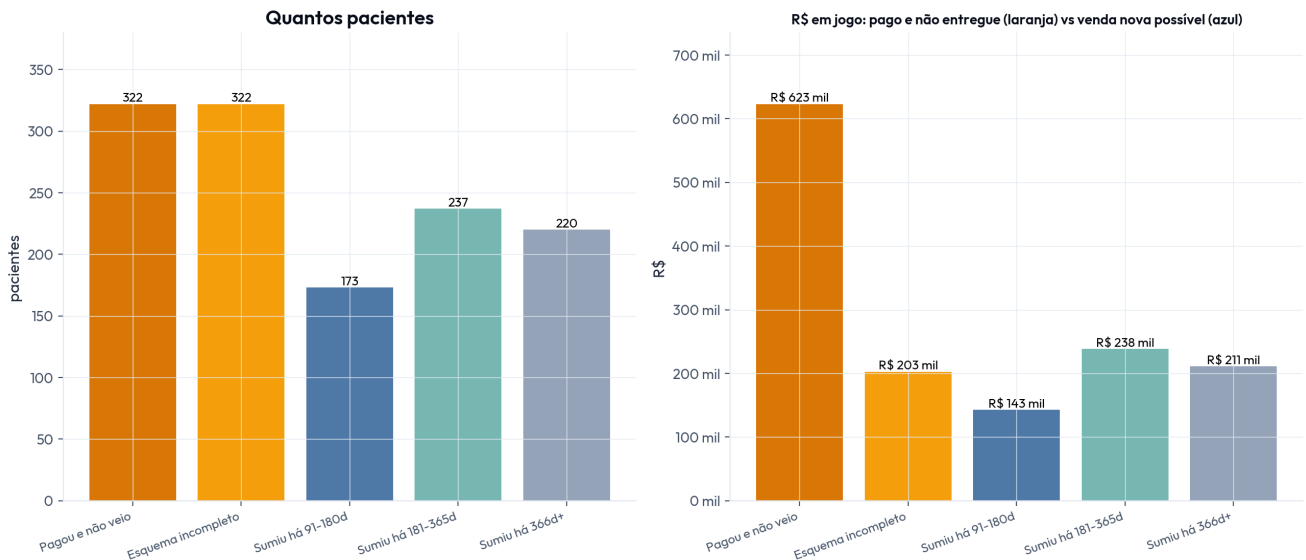
Outros 322 pacientes tomaram a primeira dose e não voltaram para completar. São 360 doses e R\$ 202.730. Esses são os mais fáceis de trazer de volta: já entraram na clínica, já confiaram, já tomaram uma agulhada. O intervalo normal entre a primeira e a segunda dose na sua base é de 77 dias, então quem passou disso está atrasado de verdade.

Os dois grupos não se sobrepõem: são 644 pacientes distintos.

Mais 630 pacientes não têm nenhuma pendência, mas sumiram. Compraram, tomaram tudo e não voltaram. Separei em três faixas porque a conversa muda em cada uma: 173 sumiram há 91 a 180 dias, 237

há 181 a 365 dias e 220 há mais de um ano. Juntos deixaram R\$ 1.800.096 no caixa ao longo do tempo, e 494 deles compraram duas ou mais vacinas diferentes, ou seja, não foram cliente de uma vez só.

Os cinco grupos para chamar de volta



A primeira barra do gráfico acima está inflada, e o próximo bloco explica por quê.

Antes de discar: duas coisas para conferir no sistema

Essas duas verificações são portões. Enquanto elas não forem feitas, uma parte das listas pode estar apontando para o paciente errado, e nada irrita mais um cliente do que ser cobrado por algo que ele já fez. Estimo 2 a 3 horas de trabalho para as duas.

Portão 1: as 851 aplicações carimbadas no último dia da base

No dia 30/04/2026 a base registra **851 aplicações**. Em um dia normal dos dois meses anteriores são 40. É vinte e uma vezes o movimento normal, o que já pediria explicação, mas há dois detalhes que fecham o caso.

Primeiro: **nenhuma dessas 851 é primeira dose**. São todas segundas e terceiras doses. Em qualquer outro dia da base, 62% das aplicações são primeira dose. Um mutirão de verdade traria gente nova junto.

Segundo: o intervalo entre a compra e a aplicação despenca. Na base inteira, a segunda dose nunca é aplicada antes de 60 dias da compra e a terceira nunca antes de 120 dias, o que faz sentido clínico. Nesse dia específico aparecem segundas doses com 30 dias e terceiras com 36 dias.

As duas leituras possíveis são: **(a) artefato de extração**, em que doses agendadas ou ainda pendentes foram carimbadas com a data em que o arquivo foi gerado, ou **(b) mutirão real** de segundas e terceiras doses represadas. Não dá para decidir só com esse arquivo, e não vou decidir por você.

O teste que separa as duas: **abra a agenda de atendimento do dia 30/04/2026 e conte quantas pessoas realmente passaram pela sala de aplicação**. Se forem umas 30 ou 40, é artefato. Se forem 851, foi mutirão. A lista completa está em `aplicacoes_ultimo_dia_conferir.csv`, ordenada pelo intervalo compra-aplicação, então os casos mais implausíveis aparecem primeiro.

O que está em jogo: se for artefato, essas 851 doses (R\$ 446.514, 736 pacientes) estão hoje contadas como entregues e na verdade são pendências. O grupo “pagou e não veio” e o grupo “esquema incompleto” cresceriam, e o número de R\$ 825.644 subiria. Note que a maioria desses 736 pacientes comprou há pouco tempo, então provavelmente não são casos atrasados, e sim doses ainda dentro do prazo. Por isso não os incluí nas listas de chamada: incluir agora seria cobrar gente que não deve nada.

Portão 2: os 1.324 esquemas repetidos

Encontrei **1.324 esquemas em que o mesmo paciente comprou de novo uma vacina que já tinha comprado antes**, envolvendo 1.073 pacientes e R\$ 1.530.840, ou 14,2% da receita registrada. O intervalo mediano entre a compra anterior e a repetida é de **92 dias**.

Isso é clinicamente estranho. Excluí a Gripe Tetra da conta justamente por ser a única de reforço periódico da sua tabela. O que sobra é Hepatite B repetida, HPV-9 repetido, Herpes Zoster repetido, QDenga repetida: esquemas que se toma uma vez na vida, comprados de novo três meses depois. As repetições aparecem em todas as vacinas mais ou menos na mesma proporção (334 em QDenga, 180 em Herpes Zoster, 153 em Hepatite B, 150 em HPV-9), o que enfraquece a hipótese de ser um comportamento de um produto específico.

As duas leituras: **(a) registro duplicado**, em que o mesmo esquema foi lançado duas vezes no sistema, ou **(b) venda real**, com pessoas diferentes usando o mesmo cadastro (mãe agendando para filho no próprio nome, por exemplo) ou revacinação por perda de comprovante.

O teste que separa: **pegue 10 casos do topo de `cursos_repetidos_conferir.csv` e confira no prontuário se as duas compras têm nota fiscal separada e lote de vacina diferente**. Duas notas com lotes distintos é venda real; uma nota só é duplicidade.

O que está em jogo para a chamada de volta: **52 das pendências que listei estão em esquemas repetidos** (42 no grupo “pagou e não veio” e 10 no “esquema incompleto”), somando R\$ 45.906. Se for duplicidade, ligar para esses pacientes é convidá-los a tomar uma dose que já tomaram. Todas as listas de chamada trazem a coluna `repeticao_a_conferir` marcada como verdadeira nesses casos: **filtre por ela e deixe esses pacientes para o fim da fila até o portão 2 estar respondido.**

O plano de retomada de contato

Com os dois portões respondidos, a fila é esta. A recepção faz 40 contatos por dia, cerca de 5 minutos cada, o que dá 3h20 de trabalho diário e custa R\$ 100 por dia a R\$ 30 a hora. A operação inteira, com os quatro toques descritos abaixo aplicados aos 1.274 pacientes, sai por cerca de R\$ 12.740 e 425 horas de recepção. O custo em dinheiro é irrelevante perto do que está em jogo; o limitante é a agenda da recepção, e é por isso que a ordem abaixo importa.

Quanto cada grupo devolve por real gasto, a 30% de conversão e com os quatro toques inclusos no custo:

Grupo	Contatos	Custo	Retorno a 30%	Relação
Pagou e não veio	322	R\$ 3.220	R\$ 186.874	58x
Esquema incompleto	322	R\$ 3.220	R\$ 60.819	19x
Sumidos 91-180 dias	173	R\$ 1.730	R\$ 42.785	25x
Sumidos 181-365 dias	237	R\$ 2.370	R\$ 71.548	30x
Sumidos 366 dias ou mais	220	R\$ 2.200	R\$ 63.390	29x
Gripe vencida	146	R\$ 1.460	R\$ 8.322	6x

Antes de ler a tabela, um aviso de soma: a linha da gripe não traz gente nova, os 146 já estão contados dentro das três linhas de sumidos, e por isso a coluna de contatos soma 1.420 para 1.274 pessoas distintas. Some as colunas como teto, nunca como piso. Duas leituras dessa tabela. Primeiro, **o retorno dos dois primeiros grupos não é dinheiro novo**: aqueles R\$ 186.874 e R\$ 60.819 já estão no seu caixa e o que a chamada faz é transformar dívida de serviço em serviço prestado, tirando esse valor do risco de virar pedido de reembolso. Só as linhas de sumidos e de gripe são venda nova. Segundo, **nenhum grupo fica abaixo de 2x**, então nenhum precisa ser descartado por custo. O que decide a ordem é hora de recepção, não dinheiro. Todas as linhas de retorno carregam a mesma fragilidade: os 30% são referência de mercado, não número seu (confiança baixa). E nas linhas de venda nova (sumidos e gripe) o retorno é preço cheio: o que a dose custa para a clínica comprar não está na base, então a relação real nessas quatro linhas é menor do que a tabela mostra, ainda que sobre folga para nenhuma delas ficar abaixo de 2x.

A sequência de quatro toques

Vale para todos os grupos, mudando só o texto e o canal de entrada.

No **dia 0** vai uma mensagem leve, sem oferta e sem cobrança, só lembrando que existe uma dose reservada em nome da pessoa. No **dia 3**, quem não respondeu recebe um lembrete com prazo concreto, do tipo “consigo segurar esse horário até sexta”. No **dia 7**, entra um incentivo real: horário prioritário fora do pico, ou aplicação em casa se a clínica oferece, ou um brinde. No **dia 14** é ligação humana, mesmo que os três toques anteriores tenham sido por WhatsApp, porque essa é a última cartada. Quem não responde ao quarto toque sai da fila ativa e volta só na campanha de massa do trimestre seguinte. Não insista: passar disso queima a marca e não converte.

Hoje: os 50 maiores do grupo “pagou e não veio”

Abra `pagou_e_nao_veio.csv`, que já vem ordenado do maior para o menor valor pendente, e pegue as 50 primeiras linhas. **São R\$ 206.166 pagos e não entregues concentrados em 50 pessoas.** Esse lote não é da recepção: quem liga é você ou a gerência, porque o assunto é dinheiro alto e a conversa pode virar pedido de reembolso.

Vale saber o que esse recorte por dinheiro faz com a fila: 50 nomes escolhidos por valor não são um retrato do grupo. Bela Vista entra com 22,6% dos seus pacientes do grupo e Riacho Doce com 5,3%; quem tem mais de 60 anos entra a 21,3% e as crianças até 12 anos a 8,7%. Não é escolha sua nem minha, é o preço da vacina mandando, porque esquema caro se concentra em quem toma Herpes Zoster e HPV-9. Como os outros 272 do mesmo grupo recebem o mesmo roteiro dentro de oito dias úteis, o desequilíbrio dura uma semana e não vira política. Se a semana 1 escorregar, refaça o lote pegando os 5 maiores de cada bairro em vez dos 50 maiores no geral.

O roteiro é de serviço, não de venda: a pessoa pagou, a clínica está devendo, você ligou para marcar. A meta é **30 dos 50 com data marcada até sexta-feira**. O número que você confere na sexta é quantos nomes desses 50 têm horário no sistema. Se a resposta ficar abaixo de 20, o problema não é a lista, é o roteiro, e vale trocar a abertura antes de escalar para os outros 272.

Se você não fizer isso: esses 203 pacientes que já esperam mais de 180 dias continuam com dinheiro parado na sua mão, e cada mês que passa aumenta a chance de virarem pedido de devolução em vez de agendamento.

Semana 1: os outros 272 do mesmo grupo, com teste de roteiro

Os 272 restantes vão por mensagem, 40 por dia, oito dias úteis. Aqui vale **dividir a lista ao meio, 136 e 136, e usar dois roteiros diferentes**: um focado no dinheiro (“você tem R\$ X em doses pagas aguardando”) e outro focado na proteção (“seu esquema de Hepatite B está incompleto e a proteção só fecha na terceira dose”). Meça quantos agendamentos cada metade gerou e use o vencedor em todas as ondas seguintes. Meta da semana: **80 agendamentos, cerca de 30% de conversão** (confiança baixa, é referência de mercado e não número da sua base).

Quem não responder ao quarto toque entra numa lista de espera e recebe uma última tentativa junto com a campanha sazonal.

Semana 2: os 322 do esquema incompleto

`esquema_incompleto.csv`, 40 por dia, também por mensagem. Esse grupo é mais barato de converter porque a pessoa já veio uma vez e já tem uma dose no braço. O argumento é clínico e curto: o esquema

pela metade não protege. Meta: **100 agendamentos na semana**, e o número da sexta é quantas segundas e terceiras doses foram aplicadas de fato, não quantas mensagens saíram.

O maior pedaço do grupo é Hepatite B, com 84 dos 322, seguido de QDenga com 69, Herpes Zoster com 52, HPV-9 e Meningo B com 40 cada e Meningo ACWY com 37. Como a Hepatite B é a única de três doses, vale um roteiro próprio para ela: quem parou na segunda tem um argumento clínico diferente de quem parou na primeira.

Semana 3: os sumidos recentes e médios

`sumidos_sem_pendencia.csv`, filtrando as faixas “recente” (173 pessoas, sumidas há 91 a 180 dias) e “medio” (237 pessoas, 181 a 365 dias). Aqui não há dívida nenhuma: é venda nova. O esquema vendido na sua clínica vai de R\$ 190 a R\$ 4.468, então não dá para projetar as duas faixas com o mesmo preço: usei a cesta que cada uma comprou de fato, **R\$ 824 por esquema na faixa recente e R\$ 1.006 na média** (média dos esquemas históricos do próprio grupo, não do catálogo). A 30% de conversão, a faixa recente vale **R\$ 42.785** e a média vale **R\$ 71.548** (confiança baixa nos dois, a taxa de conversão é referência de mercado; o preço tem confiança alta, sai da cesta real que cada faixa comprou).

A ordem de ataque é essa mesmo: gente que sumiu há 4 meses converte melhor que gente que sumiu há 10, e você quer aprender com a faixa fácil antes de gastar hora de recepção na difícil.

Semana 4 em diante: os frios e a campanha de gripe

Os 220 que sumiram há mais de um ano não merecem quatro toques individuais. **Faça um disparo único em massa** com oferta concreta e deixe quem responder entrar na fila normal.

Junto vai `gripe_anual_vencida.csv`, com **146 pacientes cuja última compra de Gripe Tetra passou de 365 dias**. Um aviso importante: **todos os 146 já estão dentro da lista de sumidos**, não são gente nova. E a premissa de que a dose é anual vem da bula, não do seu dado: na sua base, o intervalo mediano entre duas compras de Gripe Tetra é de **81 dias**, o que é curto demais para ser reforço anual e provavelmente é mais um sintoma do problema do portão 2. Trate essa lista como um roteiro alternativo para os sumidos, não como um segmento próprio. A R\$ 190 a dose e 30% de conversão, são R\$ 8.322 (confiança baixa, premissa não confirmada no dado).

Ainda na gripe, há um segmento que não aparece em nenhuma lista acima: **1.896 dos seus 3.221 pacientes (58,9%) nunca compraram Gripe Tetra, e 1.027 deles seguem ativos**, sem pendência e sem terem sumido. Os outros 869 já estão nas listas de chamada, então não some os dois grupos. Esses 1.027 estão em `sem_gripe_ativos.csv`, gastaram em média R\$ 4.050 na clínica e nunca levaram a única vacina do seu catálogo que se repete. A recepção dispara mensagem única para eles na abertura da próxima campanha sazonal, 100 nomes por dia em duas semanas, oferecendo horário de gripe. A R\$ 190 a dose e 15% de conversão, são **R\$ 29.270** (confiança baixa: a taxa é referência de campanha sazonal, não sai da sua base, e a premissa de reforço anual ainda depende do portão 2). O número que você confere na sexta é quantas doses de gripe foram aplicadas, não quantas mensagens saíram; se a primeira semana ficar abaixo de 50 doses aplicadas, pare o disparo e teste outra oferta em 100 nomes antes de gastar o resto da lista. Se você não fizer, esse pedaço da base continua comprando uma vez na vida e nunca entra num ciclo anual.

A lista de alto valor, e por que ela ganhou 31 nomes

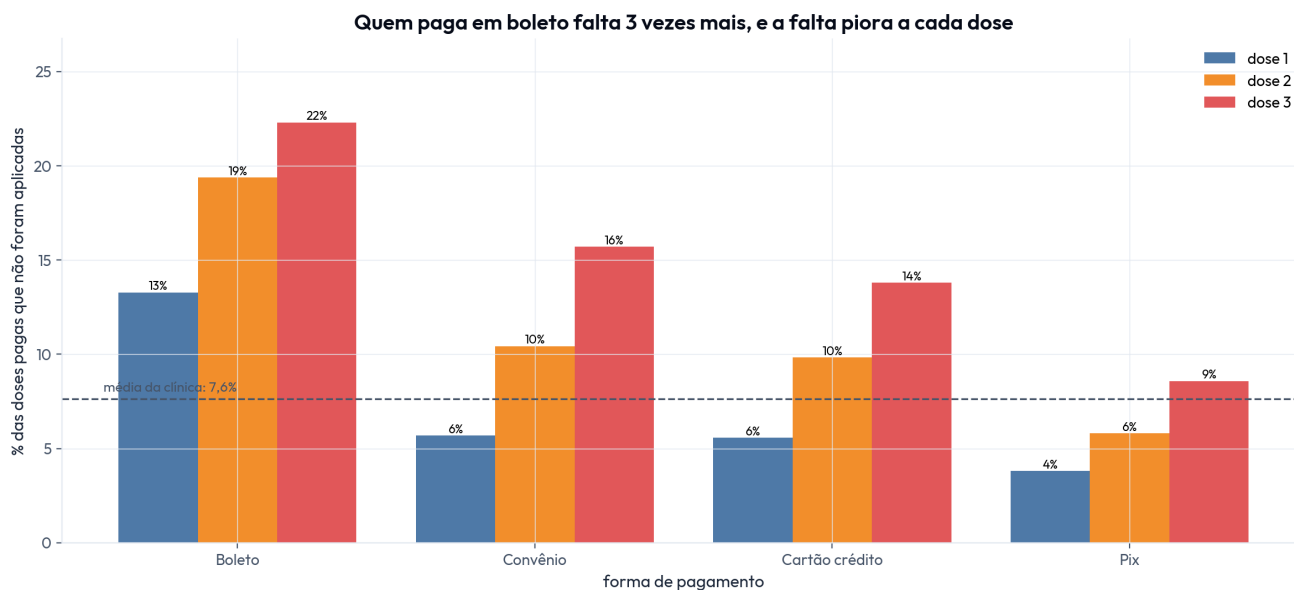
`alto_valor_em_risco.csv` traz **159 pacientes**: os 128 do decil superior de gasto histórico (os 10% que mais gastaram) entre todos os acionáveis (corte em R\$ 4.702,80, valor médio de R\$ 5.788 por paciente) mais 31 que eu acrescentei de propósito. Ela não é um grupo à parte: é um recorte dos três grupos anteriores, então **não some o R\$ dela com os outros para não contar duas vezes**.

O motivo dos 31: quando se ordena uma fila por dinheiro gasto, o dinheiro carrega junto quem ele favorece. Medi a taxa de seleção dessa fila por forma de pagamento, canal de aquisição, bairro, sexo e faixa etária. O maior desvio é por **faixa etária**: adolescentes de 13 a 18 anos entram na lista a 14,0%, enquanto adultos de 46 a 60 entram a 8,7%. A razão entre a menor e a maior taxa é 0,62, abaixo do patamar de 0,80 que costuma ser usado como piso aceitável. Faz sentido: HPV-9 custa R\$ 1.117 e é vacina de adolescente, então a idade entrou na fila de carona no preço do produto. Só que arrumar a idade sozinha não arruma a fila: ao puxar as faixas de adulto para dentro, o desvio migra para o bairro, que sai de 0,67 para 0,61. Por isso a cota roda em rodadas, sempre no corte mais desequilibrado do momento, até nenhum dos cinco ficar abaixo de 0,80.

O arquivo que você vai usar já sai equilibrado: **acrescentei os 31 pacientes de maior gasto de cada grupo sub-representado**, 20 pela faixa etária (7 da faixa 46-60, 4 da 31-45, 3 de 0-12, 3 de 19-30 e 3 de 60+) e 11 pelo bairro (4 do Alto da Serra, 3 da Boa Vista, 2 da Vila Nova, 1 de Morada do Sol e 1 de São Cristóvão), até cada grupo alcançar 80% da taxa do grupo mais favorecido. Custa **R\$ 77,50 de hora de recepção** e esses 31 têm valor médio de R\$ 4.570, então não é caridade, é gente cara que a ordenação por decil ia deixar de fora. No arquivo entregue nenhum dos cinco cortes fica abaixo de 0,80: o mais apertado continua sendo a faixa etária, agora em 0,80, e o bairro sobe para 0,81. A coluna `entrada` do arquivo diz se cada linha veio pelo decil de valor ou pela cota de equilíbrio.

O que faz o paciente faltar: a forma de pagamento

Sua taxa geral de falta é de 7,6% das doses vendidas. Ela não é distribuída por igual, e o corte que mais separa é a forma de pagamento.



Quem paga em boleto falta 16,0% das vezes. Quem paga em Pix falta 4,8%. É 3,3 vezes mais. O boleto responde por 14,2% das suas doses vendidas e por **28,6% de todo o valor pendente, R\$ 236.418**. O efeito não é ilusão de mistura de público: ele aparece dentro de cada canal de aquisição separadamente, de 22,8% no boleto vindo do Google até 2,5% no Pix vindo de indicação.

O canal também separa bem: **quem chegou pelo Google falta 13,0%, quem veio por indicação falta 4,3%**, e o Google sozinho carrega R\$ 339.371 de doses pendentes. Paciente que chega por busca tem menos vínculo e some mais fácil.

E a falta cresce a cada dose do esquema: 5,9% na primeira, 9,5% na segunda, 12,7% na terceira. Uma ressalva de leitura: **a terceira dose só existe para Hepatite B na sua tabela**, então esse último degrau mistura “é a terceira dose” com “é Hepatite B”. A subida da primeira para a segunda, essa sim, aparece em seis vacinas diferentes e é sólida.

Isso vira duas mudanças de processo, e são elas que impedem a lista de reaparecer daqui a seis meses:

A primeira é **agendar todas as doses do esquema no ato da compra**, com data marcada no sistema, especialmente para quem paga em boleto e para quem chegou pelo Google. Hoje a pessoa paga o pacote inteiro e vai embora sem data para a segunda dose; 77 dias depois ninguém lembra. Responsável: recepção, a partir da próxima segunda, valendo para 100% das vendas novas.

A segunda é **um alerta automático no painel**: esquema pago sem aplicação em 30 dias entra sozinho na fila de contato, e dose de esquema começado que passa 100 dias da anterior gera aviso. Assim a recepção trabalha com 5 a 10 nomes por dia em vez de descobrir 1.467 doses paradas daqui a dois anos. Responsável: quem cuida do sistema, montado até o fim do mês.

Como medir se funcionou

Na sexta-feira de cada semana, calcule por grupo: quantos foram contatados, quantos responderam, quantos marcaram e quantos apareceram. O número que manda é o último. Divida os agendamentos pelas horas de recepção gastas naquele grupo e compare os grupos entre si.

Na segunda seguinte, ataque primeiro o grupo com melhor resultado por hora e desça na fila o que veio fraco. Se um grupo entregar menos de 10% de conversão em duas semanas seguidas, tire ele da fila individual e jogue para disparo em massa. Minha aposta é que “esquema incompleto” ganha de “pagou e não veio” por hora gasta, mesmo tendo menos dinheiro em jogo, porque a conversa é mais curta e o paciente já conhece a clínica. Se acontecer o contrário, inverta a ordem das semanas 1 e 2 nas próximas rodadas.

Limitações que você precisa conhecer

A base termina em 30/04/2026 e hoje é 16/08/2026. São 108 dias que não estão no arquivo. Tudo que eu chamo de “sumido há 100 dias” pode ter voltado no meio tempo. Não embuti esses 108 dias em nenhum número, então cruze as listas com o sistema atual antes de discar.

As compras param antes das aplicações. A última compra registrada é de 10/04/2026, mas o fluxo já afina em 03/04: entre 3 e 10 de abril há 42 compras no total, contra uma média de 32,3 por dia em março. Na prática o registro de vendas foi cortado em 02/04/2026. Isso não afeta as listas de pendência, mas significa que qualquer coisa que eu diga sobre “vendas do último mês” está incompleta, e por isso não digo.

O grupo dos sumidos depende de ninguém ter comprado por fora do sistema. Se a clínica registra aplicação de vacina trazida pelo paciente ou atendimento sem venda, esses eventos não estão aqui e algumas pessoas da lista podem estar ativas.

A coluna `data_cadastro` foi descartada. Ela vai de 05/01/2025 a 21/03/2026, mas as compras começam em 26/12/2023, e **1.922 dos 3.221 pacientes têm compra anterior à própria data de cadastro.** A coluna parece registrar uma migração de sistema, não a entrada do paciente. Nenhum cálculo aqui a usa.

A coluna de sexo não é confiável. Ela concorda com a terminação do primeiro nome em só 52,3% dos casos, o mesmo que sorteio. Não usei sexo para priorizar ninguém, e a cota de equilíbrio da fila de alto valor foi decidida por faixa etária e bairro, que são os cortes que de fato desequilibravam. O detalhe está no apêndice.

Não há telefone nem e-mail nos arquivos. As listas saem com identificador e nome; a recepção precisa cruzar com a agenda para pegar o contato.

Os percentuais de conversão de 30% a 40% não vêm da sua base, e sim de referência de mercado para chamada de volta. São o número mais frágil do relatório. Depois da primeira semana você terá o seu próprio número, e ele vale mais que o meu.

Apêndice

Arquivos gerados

- `pagou_e_nao_veio.csv` (322 linhas) — ordenado por valor pendente decrescente. Enviesa a favor de quem comprou vacina cara: quem deve uma Triviral de R\$ 210 cai para o fim mesmo esperando há mais tempo.
- `esquema_incompleto.csv` (322 linhas) — mesma ordenação e mesmo viés.
- `sumidos_sem_pendencia.csv` (630 linhas) — ordenado por faixa de tempo e, dentro dela, por gasto histórico decrescente. Sub-atende quem gastava pouco e voltava sempre.
- `gripe_anual_vencida.csv` (146 linhas) — ordenado por dias desde a última dose de gripe, do mais antigo para o mais recente. Recorte da lista de sumidos, sem nomes novos.
- `alto_valor_em_risco.csv` (159 linhas) — ordenado por gasto histórico decrescente, com a cota de equilíbrio por faixa etária e por bairro no fim. Recorte dos três grupos principais; não somar com eles.
- `sem_gripe_ativos.csv` (1.027 linhas) — ordenado por gasto histórico decrescente. Pacientes ativos que nunca compraram Gripe Tetra; enviesa a favor de quem já gastou muito, quem comprou uma vacina barata só cai para o fim.
- `cursos_repetidos_conferir.csv` (1.324 linhas) — ordenado por valor do esquema decrescente. Lista de auditoria do portão 2.
- `aplicacoes_ultimo_dia_conferir.csv` (851 linhas) — ordenado pelo intervalo entre compra e aplicação, crescente, então os casos mais implausíveis vêm primeiro. Lista de auditoria do portão 1.

Imagens: `buckets_acionaveis.png` , `dias_desde_ultima_atividade.png` , `falta_por_pagamento.png` .

Como cheguei aos grupos

Termo	Definição usada
Data de referência	30/04/2026, a última data de qualquer coluna de data da base
Esquema	Conjunto de linhas do mesmo paciente, mesma vacina e mesma data de compra
Pagou e não veio	Esquema com zero doses aplicadas
Esquema incompleto	Esquema com pelo menos uma dose aplicada e pelo menos uma pendente
Sumido	Paciente sem nenhuma dose pendente e sem compra ou aplicação há mais de 90 dias
Faixas de sumido	Recente 91-180 dias, médio 181-365 dias, frio 366 dias ou mais
Repetição suspeita	Segunda compra ou posterior da mesma vacina pelo mesmo paciente, excluindo Gripe Tetra
Valor médio de esquema	R\$ 969 na base inteira (11.088 esquemas vendidos); nas projeções de venda nova, a média da própria faixa: R\$ 824 recente, R\$ 1.006 médio, R\$ 960 frio
Custo de contato	R\$ 2,50, de R\$ 30 por hora e 12 contatos por hora

A regra de “esquema incompleto” usa o protocolo declarado na própria base, na coluna `protocolo_doses`, que é consistente por vacina (Hepatite B sempre 3, HPV-9 sempre 2, Gripe Tetra sempre 1, e assim por diante). Como o protocolo vem do dado e não de heurística, a confiança nesse grupo é alta.

Taxa de falta por corte

Percentual das doses vendidas que não foram aplicadas, e o valor parado em cada recorte.

Corte	Maior taxa	Menor taxa
Forma de pagamento	Boleto 16,0% (R\$ 236.418)	Pix 4,8% (R\$ 236.084)
Canal de aquisição	Google 13,0% (R\$ 339.371)	Indicação 4,3% (R\$ 193.379)
Número da dose	Terceira 12,7%	Primeira 5,9%
Vacina	Hepatite B 9,0%	Pneumo 20 6,1%
Faixa etária	13-18 anos 9,3%	46-60 anos 7,0%
Bairro	Jardim das Acácias 8,7%	Morada do Sol 6,4%
Sexo	Masculino 7,9%	Feminino 7,4%

Duas leituras negativas que também valem: **bairro não concentra nada**. A proporção de pacientes acionáveis por bairro fica entre 36% e 43% nos dez bairros, e os três maiores juntos são 32,7% da fila, praticamente o peso deles na base. Campanha de bairro não tem vantagem sobre campanha geral, e por isso não gerei arquivo de bairro. **A coluna de sexo não deve ser usada para nada**. Além de quase não separar (7,9% contra 7,4%), ela não bate com os nomes: entre os 2.514 pacientes cujo primeiro nome termina em “a” ou “o”, a coluna concorda com a terminação em apenas 52,3% dos casos, o mesmo que jogar uma moeda. Ou o campo foi preenchido errado na origem, ou está trocado em parte dos registros. Deixei a coluna nos arquivos apenas para você conferir, mas não priorize nem escreva mensagem baseado nela.

Colunas do arquivo que não entraram em nenhum cálculo

`data_cadastro` (inconsistente, explicada nas limitações) e `nome` das planilhas de vacinação, usado apenas para exibição nas listas. Todas as demais colunas de `pacientes.csv` e `vacinacoes.csv` entraram em pelo menos um corte acima.