



Distilo

RELATÓRIO DE ANÁLISE

Análise de Churn

Een uitgewerkt voorbeeld van patiënten terughalen (fictieve data).

Gerado em

Cliente

runId

16/08/2026

Voorbeeldrapport · synthetische data

demo-churn

Wie je terug moet halen: 1.274 patiënten en R\$ 825.644 aan betaalde doses die niemand heeft gezet

Ik vond **1.274 patiënten in een situatie waar je iets mee moet, 39,6% van je bestand van 3.221**, en daarbinnen **R\$ 825.644 aan doses die al betaald zijn en nooit zijn gezet**. Dat is het meest ongemakkelijke getal van het rapport, want het is geld dat al binnen is en dienst die nog niet geleverd is.

Wat het bestand laat zien

Je hebt 3.221 geregistreerde patiënten, 19.247 verkochte doses en R\$ 10.748.740 aan geregistreerde omzet tussen 26-12-2023 en 30-04-2026. De groei is echt: de aankopen gingen van 201 doses in december 2024 naar 2.446 in januari 2026.

De operatie is ook op een manier georganiseerd die de analyse makkelijk maakt. In 99,9% van de gevallen koopt de patiënt het hele schema in één keer: wie hepatitis B koopt, koopt de drie doses op dezelfde dag, wie Qdenga koopt, koopt de twee. Dat betekent dat elke aankoop een afgesloten verbintenis is, en elke niet-gezette dosis een openstaande belofte. Elk van die aankopen noem ik in de rest van de tekst een “schema”.

Peildatum voor alles hier: 30-04-2026, de laatste dag die in het bestand voorkomt. Geen enkele berekening gebruikt de datum van vandaag.

Het probleem

Van de 19.247 verkochte doses **staan er 1.467 als niet gezet: 7,6% van alles wat je hebt verkocht**. Ze vallen uiteen in twee heel verschillende problemen, en het verschil doet ertoe voor het script van de balie.

322 patiënten betaalden het complete schema en zijn nooit komen opdagen, zelfs niet voor de eerste dosis. Dat zijn 656 schema's, 1.107 doses en R\$ 622.914. De gemiddelde wachttijd van die schema's is al 200 dagen, en 203 van die 322 patiënten wachten al meer dan 180 dagen. Het meest extreme geval is een patiënt met 6 openstaande doses en R\$ 6.628 betaald.

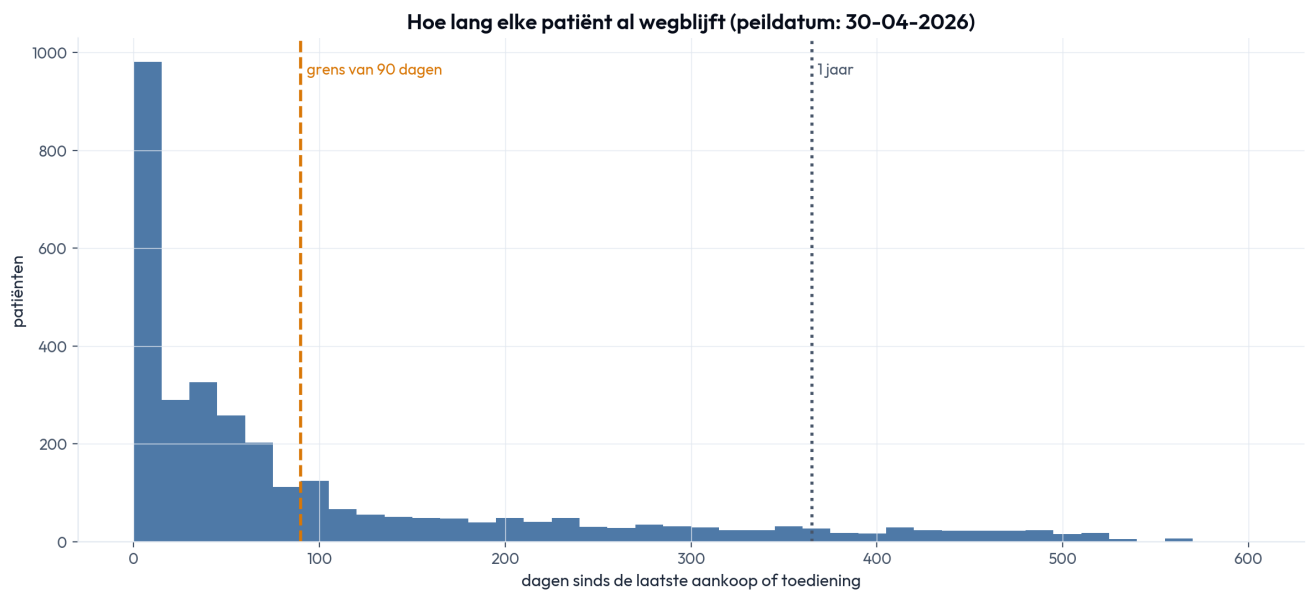
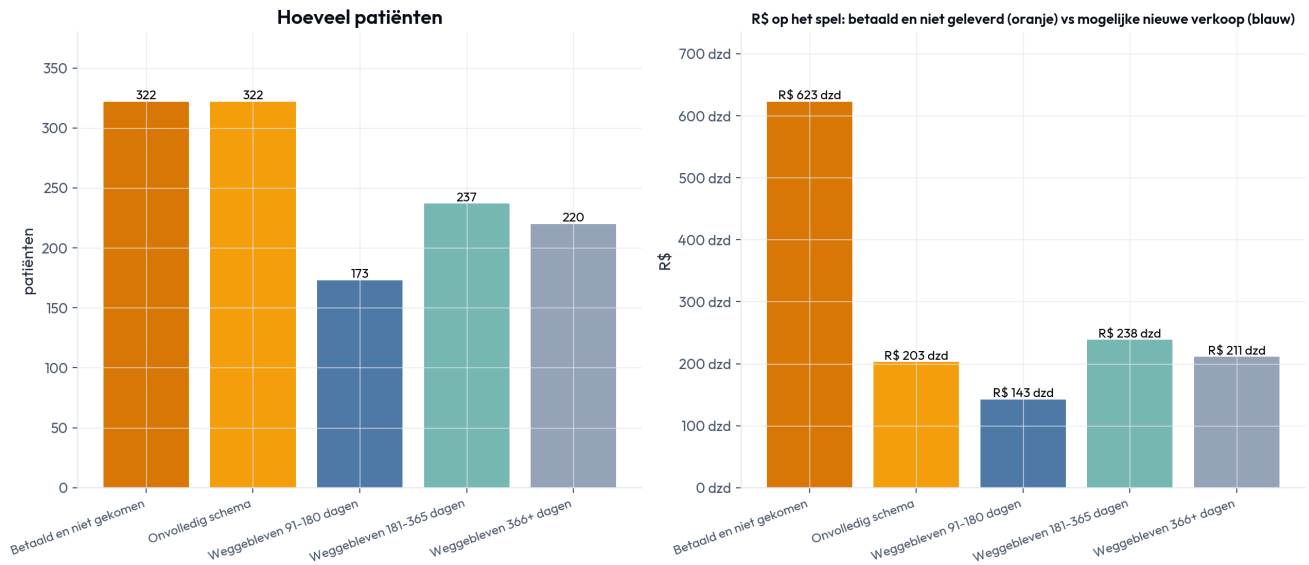
Nog eens 322 patiënten namen de eerste dosis en kwamen niet terug om het schema af te maken. Dat zijn 360 doses en R\$ 202.730. Dit zijn de makkelijkste om terug te halen: ze zijn al in de kliniek geweest, hebben al vertrouwen gegeven, hebben al een prik gehad. Het normale interval tussen de eerste en de tweede dosis is in jouw bestand 77 dagen, dus wie daar overheen is, is echt te laat.

De twee groepen overlappen niet: het zijn 644 verschillende patiënten.

Nog eens 630 patiënten hebben niets openstaan, maar zijn weggebleven. Ze kochten, namen alles en kwamen niet terug. Ik heb ze in drie banden verdeeld omdat het gesprek per band anders is: 173 zijn 91 tot 180 dagen weg, 237 zijn 181 tot 365 dagen weg en 220 meer dan een jaar. Samen lieten ze in de loop van de

tijd R\$ 1.800.096 in de kas achter, en 494 van hen kochten twee of meer verschillende vaccins, ze waren dus geen eenmalige klant.

De vijf groepen om terug te bellen



De eerste balk van de grafiek hierboven is opgeblazen, en het volgende blok legt uit waarom.

Voordat je gaat bellen: twee dingen om in het systeem na te kijken

Deze twee controles zijn poorten. Zolang ze niet gedaan zijn, kan een deel van de lijsten naar de verkeerde patiënt wijzen, en niets irriteert een klant meer dan aangesproken worden op iets wat hij al gedaan heeft. Ik schat 2 tot 3 uur werk voor de twee samen.

Poort 1: de 851 toedieningen met de stempel van de laatste dag van het bestand

Op 30-04-2026 registreert het bestand **851 toedieningen**. Op een normale dag in de twee maanden daarvoor zijn dat er 40. Dat is eenentwintig keer de normale drukte, wat op zich al om uitleg vraagt, maar twee details maken de zaak rond.

Ten eerste: **geen van die 851 is een eerste dosis**. Het zijn allemaal tweede en derde doses. Op elke andere dag in het bestand is 62% van de toedieningen een eerste dosis. Een echte mutirão (een eendaagse massale vaccinatieactie) zou ook nieuwe mensen meebrengen.

Ten tweede: het interval tussen aankoop en toediening stort in. In het hele bestand wordt de tweede dosis nooit eerder dan 60 dagen na aankoop gezet en de derde nooit eerder dan 120 dagen, wat klinisch logisch is. Op die specifieke dag verschijnen tweede doses na 30 dagen en derde doses na 36 dagen.

De twee mogelijke lezingen zijn: **(a) een artefact van de export**, waarbij ingeplande of nog openstaande doses de datum hebben gekregen waarop het bestand is aangemaakt, of **(b) een echte mutirão** van opgespaarde tweede en derde doses. Met alleen dit bestand valt dat niet te beslissen, en ik ga het niet voor je beslissen.

De test die de twee scheidt: **open de agenda van 30-04-2026 en tel hoeveel mensen echt door de prikkamer zijn gegaan**. Zijn het er zo'n 30 of 40, dan is het een artefact. Zijn het er 851, dan was het een mutirão. De volledige lijst staat in `aplicacoes_ultimo_dia_conferir.csv`, gesorteerd op het interval aankoop-toediening, zodat de meest onwaarschijnlijke gevallen bovenaan staan.

Wat er op het spel staat: is het een artefact, dan tellen die 851 doses (R\$ 446.514, 736 patiënten) vandaag als geleverd terwijl ze eigenlijk openstaan. De groep "betaald en niet gekomen" en de groep "onvolledig schema" zouden groeien, en het getal van R\$ 825.644 zou stijgen. Let wel: de meeste van die 736 patiënten kochten kort geleden, dus het zijn waarschijnlijk geen achterstallige gevallen, maar doses die nog binnen de termijn vallen. Daarom heb ik ze niet in de bellijsten opgenomen: ze nu opnemen zou mensen aanspreken die niets schuldig zijn.

Poort 2: de 1.324 herhaalde schema's

Ik vond **1.324 schema's waarin dezelfde patiënt opnieuw een vaccin kocht dat hij al eerder had gekocht**, verdeeld over 1.073 patiënten en R\$ 1.530.840, oftewel 14,2% van de geregistreeerde omzet. Het mediane interval tussen de eerdere aankoop en de herhaalde is **92 dagen**.

Dat is klinisch vreemd. Ik heb griep (Gripe Tetra) juist buiten de telling gelaten omdat het de enige periodieke herhaalprik in je assortiment is. Wat overblijft is herhaalde hepatitis B, herhaalde HPV-9, herhaalde gordelroos, herhaalde Qdenga: schema's die je één keer in je leven neemt, drie maanden later opnieuw gekocht. De herhalingen komen bij alle vaccins in ongeveer dezelfde verhouding voor (334 bij Qdenga, 180 bij gordelroos, 153 bij hepatitis B, 150 bij HPV-9), wat de hypothese verzwakt dat het gedrag bij één specifiek product hoort.

De twee lezingen: **(a) dubbele registratie**, waarbij hetzelfde schema twee keer in het systeem is ingevoerd, of **(b) echte verkoop**, met verschillende mensen onder hetzelfde dossier (een moeder die op eigen naam voor haar kind boekt, bijvoorbeeld) of hervaccinatie omdat het bewijs kwijt was.

De test die scheidt: **neem 10 gevallen van bovenaan `cursos_repetidos_conferir.csv` en controleer in het prontuário (het patiëntendossier) of de twee aankopen een aparte factuur en een ander vaccinlot hebben.** Twee facturen met verschillende lots is echte verkoop; één factuur is een dubbeling.

Wat er voor het terughalen op het spel staat: **52 van de openstaande doses die ik heb opgesomd zitten in herhaalde schema's** (42 in de groep “betaald en niet gekomen” en 10 in “onvolledig schema”), samen R\$ 45.906. Is het een dubbeling, dan is die patiënten bellen hen uitnodigen voor een dosis die ze al hebben gehad. Alle bellijsten bevatten de kolom `repeticao_a_conferir`, die in die gevallen op waar staat: **filter erop en laat die patiënten achteraan in de rij tot poort 2 beantwoord is.**

Het plan om het contact te hervatten

Met de twee poorten beantwoord is dit de rij. De balie doet 40 contacten per dag, ongeveer 5 minuten elk, wat 3 uur en 20 minuten werk per dag is en R\$ 100 per dag kost bij R\$ 30 per uur. De hele operatie, met de vier contactmomenten hieronder toegepast op de 1.274 patiënten, komt uit op ongeveer R\$ 12.740 en 425 uur balie. De kosten in geld zijn verwaarloosbaar naast wat er op het spel staat; de beperking is de agenda van de balie, en daarom doet de volgorde hieronder ertoe.

Hoeveel elke groep teruggeeft per uitgegeven real, bij 30% conversie en met de vier contactmomenten in de kosten inbegrepen:

Groep	Contacten	Kosten	Opbrengst bij 30%	Verhouding
Betaald en niet gekomen	322	R\$ 3.220	R\$ 186.874	58x
Onvolledig schema	322	R\$ 3.220	R\$ 60.819	19x
Weggebleven 91-180 dagen	173	R\$ 1.730	R\$ 42.785	25x
Weggebleven 181-365 dagen	237	R\$ 2.370	R\$ 71.548	30x
Weggebleven 366 dagen of meer	220	R\$ 2.200	R\$ 63.390	29x
Griep verlopen	146	R\$ 1.460	R\$ 8.322	6x

Voordat je de tabel leest, een waarschuwing over optellen: de griepregel brengt geen nieuwe mensen, de 146 zitten al in de drie regels van weggebleven patiënten, en daarom telt de kolom contacten op tot 1.420 voor 1.274 verschillende mensen. Tel de kolommen op als plafond, nooit als bodem. Twee lezingen van deze tabel. Ten eerste: **de opbrengst van de eerste twee groepen is geen nieuw geld:** die R\$ 186.874 en R\$ 60.819 zitten al in je kas en wat het terughalen doet, is een dienstdenkt omzetten in geleverde dienst, en dat bedrag weghalen uit het risico dat het een terugbetalingsverzoek wordt. Alleen de regels van weggebleven patiënten en griep zijn nieuwe verkoop. Ten tweede: **geen enkele groep zakt onder 2x**, dus geen enkele hoeft om de kosten afgeschreven te worden. Wat de volgorde bepaalt, is balie-uren, niet geld. Alle opbrengstregels dragen dezelfde zwakte: de 30% is een marktreferentie, niet jouw getal (laag vertrouwen). En bij de regels met nieuwe verkoop (weggebleven en griep) is de opbrengst de volle prijs: wat de dosis de

kliniek kost om in te kopen staat niet in het bestand, dus de werkelijke verhouding in die vier regels is lager dan de tabel laat zien, al is er genoeg marge dat geen van hen onder 2x zakt.

De reeks van vier contactmomenten

Geldt voor alle groepen, alleen de tekst en het kanaal waarmee je begint verschillen.

Op **dag 0** gaat een licht bericht, zonder aanbod en zonder aanmaning, alleen een herinnering dat er een dosis op naam van de persoon gereserveerd staat. Op **dag 3** krijgt wie niet reageerde een herinnering met een concrete termijn, in de trant van “ik kan dit tijdstip tot vrijdag voor je vasthouden”. Op **dag 7** komt een echte prikkel: een voorrangstijdstip buiten de piek, of toediening aan huis als de kliniek dat aanbiedt, of een attentie. Op **dag 14** is het een menselijk telefoontje, ook als de drie eerdere contactmomenten via WhatsApp gingen, want dat is de laatste troef. Wie op het vierde contactmoment niet reageert, verlaat de actieve rij en komt pas terug in de massacampagne van het kwartaal erop. Dring niet aan: verder gaan dan dit beschadigt het merk en converteert niet.

Vandaag: de 50 grootste van de groep “betaald en niet gekomen”

Open `pagou_e_nao_veio.csv`, dat al gesorteerd is van het hoogste naar het laagste openstaande bedrag, en neem de eerste 50 regels. **Dat is R\$ 206.166 betaald en niet geleverd, geconcentreerd bij 50 mensen.** Deze partij is niet voor de balie: jij of het management belt, want het gaat om veel geld en het gesprek kan een terugbetalingsverzoek worden.

Goed om te weten wat die selectie op geld met de rij doet: 50 namen gekozen op bedrag zijn geen afspiegeling van de groep. Bela Vista komt erin met 22,6% van zijn patiënten uit de groep en Riacho Doce met 5,3%; wie ouder is dan 60 komt erin met 21,3% en kinderen tot 12 jaar met 8,7%. Dat is niet jouw keuze en niet de mijne, het is de vaccinprijs die het bepaalt, want dure schema's concentreren zich bij wie gordelroos en HPV-9 neemt. Omdat de andere 272 uit dezelfde groep binnen acht werkdagen hetzelfde script krijgen, duurt de scheefheid een week en wordt ze geen beleid. Loopt week 1 uit, stel de partij dan opnieuw samen door de 5 grootste per wijk te nemen in plaats van de 50 grootste in het algemeen.

Het script is dienstverlening, geen verkoop: de persoon heeft betaald, de kliniek staat in het krijt, jij belt om in te plannen. Het doel is **30 van de 50 met een afspraak vóór vrijdag**. Het getal dat je vrijdag nakijkt, is hoeveel van die 50 namen een tijdstip in het systeem hebben. Blijft het antwoord onder de 20, dan ligt het niet aan de lijst maar aan het script, en is het de moeite de opening te veranderen voordat je opschaalt naar de andere 272.

Doe je dit niet: dan blijven die 203 patiënten die al meer dan 180 dagen wachten met hun geld in jouw hand zitten, en elke maand die verstrijkt vergroot de kans dat het een verzoek om teruggave wordt in plaats van een afspraak.

Week 1: de andere 272 uit dezelfde groep, met een scripttest

De resterende 272 gaan per bericht, 40 per dag, acht werkdagen. Hier loont het **de lijst doormidden te delen, 136 en 136, en twee verschillende scripts te gebruiken**: één gericht op het geld (“je hebt R\$ X aan betaalde doses klaarstaan”) en één gericht op de bescherming (“je schema hepatitis B is onvolledig en de bescherming is pas rond bij de derde dosis”). Meet hoeveel afspraken elke helft opleverde en gebruik de

winnaar in alle volgende golven. Weekdoel: **80 afspraken, ongeveer 30% conversie** (laag vertrouwen, het is een marktreferentie en geen getal uit jouw bestand).

Wie op het vierde contactmoment niet reageert, komt op een wachtlijst en krijgt een laatste poging samen met de seizoenscampagne.

Week 2: de 322 met een onvolledig schema

`esquema_incompleto.csv`, 40 per dag, ook per bericht. Deze groep is goedkoper om te converteren omdat de persoon al een keer is geweest en al een dosis in de arm heeft. Het argument is klinisch en kort: een half schema beschermt niet. Doel: **100 afspraken in de week**, en het vrijdaggetal is hoeveel tweede en derde doses er echt zijn gezet, niet hoeveel berichten er uitgingen.

Het grootste stuk van de groep is hepatitis B, met 84 van de 322, gevolgd door Qdenga met 69, gordelroos met 52, HPV-9 en meningokokken B met elk 40 en meningokokken ACWY met 37. Omdat hepatitis B het enige vaccin met drie doses is, verdient het een eigen script: wie bij de tweede stopte heeft een ander klinisch argument dan wie bij de eerste stopte.

Week 3: de recent en middellang weggebleven patiënten

`sumidos_sem_pendencia.csv`, gefilterd op de banden “recente” (recent: 173 mensen, 91 tot 180 dagen weg) en “medio” (middellang: 237 mensen, 181 tot 365 dagen). Hier is geen enkele schuld: het is nieuwe verkoop. Het schema dat in jouw kliniek wordt verkocht loopt van R\$ 190 tot R\$ 4.468, dus je kunt de twee banden niet met dezelfde prijs projecteren: ik heb het mandje gebruikt dat elke band werkelijk kocht, **R\$ 824 per schema in de recente band en R\$ 1.006 in de middellange** (gemiddelde van de historische schema’s van de groep zelf, niet van de catalogus). Bij 30% conversie is de recente band **R\$ 42.785** waard en de middellange **R\$ 71.548** (laag vertrouwen bij beide, het conversiepercentage is een marktreferentie; de prijs heeft hoog vertrouwen, die komt uit het echte mandje dat elke band kocht).

De aanvalsvolgorde is precies zo: wie 4 maanden weg is converteert beter dan wie 10 maanden weg is, en je wilt van de makkelijke band leren voordat je balie-uren aan de moeilijke besteedt.

Week 4 en verder: de koude groep en de griepcampagne

De 220 die meer dan een jaar weg zijn, verdienen geen vier individuele contactmomenten. **Doe één massale verzending** met een concreet aanbod en laat wie reageert de normale rij instromen.

Daar gaat `gripe_anual_vencida.csv` mee, met **146 patiënten van wie de laatste aankoop van griep meer dan 365 dagen geleden is**. Een belangrijke waarschuwing: **alle 146 zitten al in de lijst van weggebleven patiënten**, het zijn geen nieuwe mensen. En de aanname dat de dosis jaarlijks is, komt uit de bijsluiters, niet uit jouw data: in jouw bestand is het mediane interval tussen twee aankopen van griep **81 dagen**, wat te kort is om een jaarlijkse herhaalprik te zijn en waarschijnlijk eerder een symptoom is van het probleem van poort 2. Behandel deze lijst als een alternatief script voor de weggebleven patiënten, niet als een eigen segment. Bij R\$ 190 per dosis en 30% conversie is het R\$ 8.322 (laag vertrouwen, aanname niet bevestigd in de data).

Nog steeds over griep: er is een segment dat in geen enkele lijst hierboven voorkomt: **1.896 van je 3.221 patiënten (58,9%) hebben nooit griep gekocht**, en **1.027 van hen zijn nog actief**, zonder openstaande

dosis en zonder weggebleven te zijn. De andere 869 staan al in de bellijsten, dus tel de twee groepen niet op. Die 1.027 staan in `sem_gripe_ativos.csv`, gaven gemiddeld R\$ 4.050 uit in de kliniek en namen nooit het enige vaccin uit je assortiment dat terugkeert. De balie stuurt hen één bericht bij de start van de volgende seizoenscampagne, 100 namen per dag in twee weken, met een aanbod voor een griepstijp. Bij R\$ 190 per dosis en 15% conversie is dat **R\$ 29.270** (laag vertrouwen: het percentage is een referentie voor seizoenscampagnes en komt niet uit jouw bestand, en de aannahme van een jaarlijkse herhaalprik hangt nog af van poort 2). Het getal dat je vrijdag nakijkt, is hoeveel griepdoses er zijn gezet, niet hoeveel berichten er uitgingen; blijft de eerste week onder de 50 gezette doses, stop dan de verzending en test een ander aanbod op 100 namen voordat je de rest van de lijst opmaakt. Doe je het niet, dan blijft dit stuk van het bestand één keer in zijn leven kopen en komt het nooit in een jaarcyclus.

De lijst met hoge waarde, en waarom er 31 namen bij kwamen

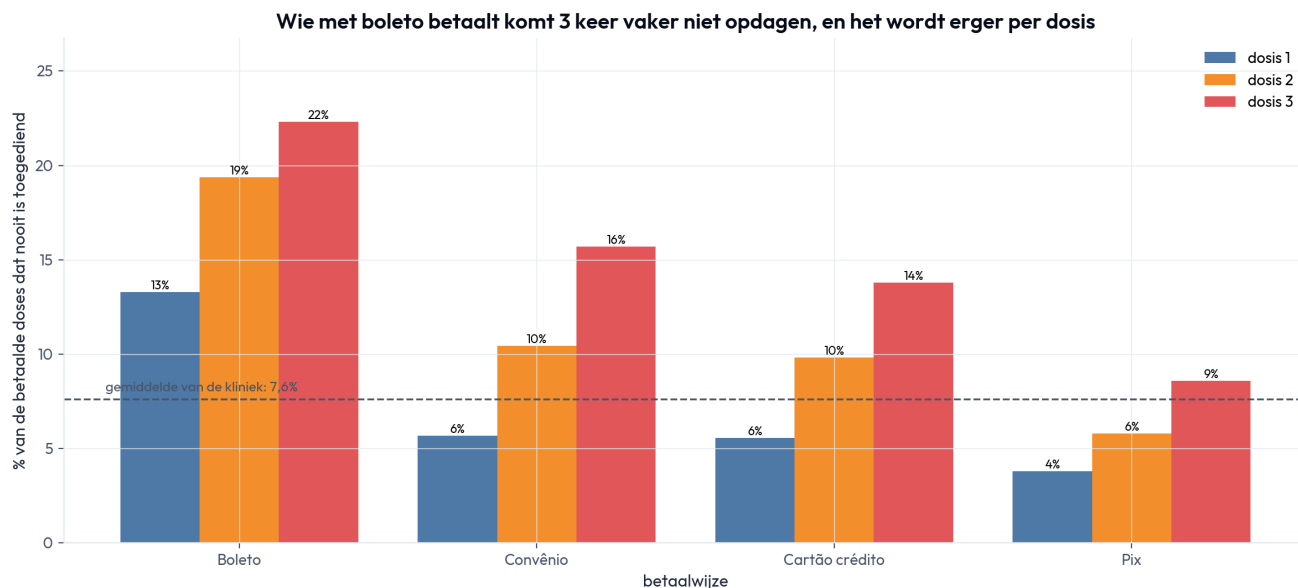
`alto_valor_em_risco.csv` bevat **159 patiënten**: de 128 uit het bovenste deciel van historische uitgaven (de 10% die het meest uitgaven) onder alle te benaderen patiënten (afkapgrens R\$ 4.702,80, gemiddelde waarde R\$ 5.788 per patiënt) plus 31 die ik er bewust aan heb toegevoegd. Het is geen aparte groep: het is een selectie uit de drie voorgaande groepen, dus **tel haar R\$ niet op bij de andere om niet dubbel te tellen**.

De reden voor de 31: wanneer je een rij ordent op uitgegeven geld, neemt het geld mee wie het bevoordeelt. Ik heb het selectiepercentage van die rij gemeten per betaalwijze, acquisitiekanaal, wijk, geslacht en leeftijdsgroep. De grootste afwijking is per **leeftijdsgroep**: tieners van 13 tot 18 jaar komen op de lijst met 14,0%, terwijl volwassenen van 46 tot 60 erop komen met 8,7%. De verhouding tussen het laagste en het hoogste percentage is 0,62, onder de drempel van 0,80 die doorgaans als aanvaardbare ondergrens wordt gebruikt. Het is logisch: HPV-9 kost R\$ 1.117 en is een tienervaccin, dus de leeftijd lifte de rij in op de prijs van het product. Alleen lost alleen de leeftijd rechtzetten de rij niet op: trek je de volwassen groepen naar binnen, dan verschuift de afwijking naar de wijk, die van 0,67 naar 0,61 gaat. Daarom draait het quotum in rondes, steeds op de meest scheve dimensie van dat moment, tot geen van de vijf onder 0,80 zit.

Het bestand dat je gaat gebruiken komt al in balans: **ik heb de 31 patiënten met de hoogste uitgaven uit elke ondervetegenwoordigde groep toegevoegd**, 20 op leeftijdsgroep (7 uit de groep 46-60, 4 uit 31-45, 3 uit 0-12, 3 uit 19-30 en 3 uit 60+) en 11 op wijk (4 uit Alto da Serra, 3 uit Boa Vista, 2 uit Vila Nova, 1 uit Morada do Sol en 1 uit São Cristóvão), tot elke groep 80% van het percentage van de meest bevoordeelde groep haalt. Het kost **R\$ 77,50 aan balie-uren** en die 31 hebben een gemiddelde waarde van R\$ 4.570, dus het is geen liefdadigheid, het zijn dure mensen die de ordening op deciel buiten de boot zou laten vallen. In het geleverde bestand zit geen van de vijf dimensies onder 0,80: de krapste blijft de leeftijdsgroep, nu op 0,80, en de wijk stijgt naar 0,81. De kolom `entrada` in het bestand zegt of elke regel via het waardedeciel of via het evenwichtsquotum is binnengekomen.

Wat maakt dat de patiënt niet komt: de betaalwijze

Je algemene no-showpercentage is 7,6% van de verkochte doses. Het is niet gelijk verdeeld, en de dimensie die het sterkst scheidt, is de betaalwijze.



Wie met boleto betaalt (een offline betaald bankformulier, vergelijkbaar met een acceptgiro) komt 16,0% van de keren niet opdagen. Wie met Pix betaalt (directe bankoverschrijving) komt 4,8% niet opdagen. Dat is 3,3 keer zoveel. Boleto is goed voor 14,2% van je verkochte doses en voor **28,6% van al het openstaande bedrag, R\$ 236.418**. Het effect is geen illusie door een mengeling van publiek: het verschijnt binnen elk acquisitiekanaal apart, van 22,8% bij boleto via Google tot 2,5% bij Pix via doorverwijzing.

Het kanaal scheidt ook goed: **wie via Google kwam komt 13,0% niet opdagen, wie via doorverwijzing kwam 4,3%**, en Google alleen draagt R\$ 339.371 aan openstaande doses. Een patiënt die via zoeken binnenkomt heeft minder binding en blijft makkelijker weg.

En de no-show stijgt met elke dosis van het schema: 5,9% bij de eerste, 9,5% bij de tweede, 12,7% bij de derde. Een kanttekening bij het lezen: **de derde dosis bestaat in jouw assortiment alleen voor hepatitis B**, dus die laatste trede mengt “het is de derde dosis” met “het is hepatitis B”. De stijging van de eerste naar de tweede, die wel, verschijnt bij zes verschillende vaccins en is solide.

Dit wordt twee procesveranderingen, en zij zijn het die voorkomen dat de lijst over zes maanden weer opduikt:

De eerste is **alle doses van het schema inplannen op het moment van aankoop**, met een datum in het systeem, vooral voor wie met boleto betaalt en wie via Google kwam. Vandaag betaalt de persoon het hele pakket en vertrekt zonder datum voor de tweede dosis; 77 dagen later weet niemand het meer. Verantwoordelijk: de balie, vanaf aanstaande maandag, voor 100% van de nieuwe verkopen.

De tweede is **een automatische melding op het dashboard**: een betaald schema zonder toediening na 30 dagen komt vanzelf in de contactrij, en een dosis van een begonnen schema die meer dan 100 dagen na de vorige ligt, geeft een waarschuwing. Zo werkt de balie met 5 tot 10 namen per dag in plaats van over twee

jaar 1.467 stilliggende doses te ontdekken. Verantwoordelijk: wie het systeem beheert, ingericht vóór het einde van de maand.

Hoe je meet of het werkte

Bereken op de vrijdag van elke week per groep: hoeveel zijn benaderd, hoeveel reageerden, hoeveel maakten een afspraak en hoeveel kwamen opdagen. Het getal dat telt, is het laatste. Deel de afspraken door de balie-uren die aan die groep zijn besteed en vergelijk de groepen onderling.

Pak de maandag erop eerst de groep met het beste resultaat per uur aan en zet wat zwak was lager in de rij. Levert een groep twee weken achter elkaar minder dan 10% conversie, haal hem dan uit de individuele rij en zet hem op massale verzending. Mijn inzet is dat “onvolledig schema” wint van “betaald en niet gekomen” per besteed uur, ook al staat er minder geld op het spel, omdat het gesprek korter is en de patiënt de kliniek al kent. Gebeurt het omgekeerde, draai dan de volgorde van week 1 en 2 om in de volgende rondes.

Beperkingen die je moet kennen

Het bestand eindigt op 30-04-2026 en vandaag is 16-08-2026. Dat zijn 108 dagen die niet in het bestand zitten. Alles wat ik “100 dagen weggebleven” noem, kan in de tussentijd teruggekomen zijn. Ik heb die 108 dagen in geen enkel getal verwerkt, dus leg de lijsten naast het huidige systeem voordat je belt.

De aankopen stoppen eerder dan de toedieningen. De laatste geregistreerde aankoop is van 10-04-2026, maar de stroom dunt al uit op 03-04: tussen 3 en 10 april zijn er in totaal 42 aankopen, tegen een gemiddelde van 32,3 per dag in maart. In de praktijk is de verkoopregistratie afgekapt op 02-04-2026. Dat raakt de lijsten met openstaande doses niet, maar het betekent dat alles wat ik over “verkopen van de laatste maand” zou zeggen onvolledig is, en daarom zeg ik het niet.

De groep weggebleven patiënten hangt ervan af dat niemand buiten het systeem om heeft gekocht. Registreert de kliniek de toediening van een door de patiënt meegebracht vaccin of een consult zonder verkoop, dan staan die gebeurtenissen hier niet in en kunnen sommige mensen op de lijst actief zijn.

De kolom `data_cadastro` is buiten beschouwing gelaten. Hij loopt van 05-01-2025 tot 21-03-2026, maar de aankopen beginnen op 26-12-2023, en **1.922 van de 3.221 patiënten hebben een aankoop van vóór hun eigen registratiedatum.** De kolom lijkt een systeem migratie vast te leggen, niet de instroom van de patiënt. Geen enkele berekening hier gebruikt hem.

De kolom geslacht is niet betrouwbaar. Hij komt in maar 52,3% van de gevallen overeen met de uitgang van de voornaam, hetzelfde als loting. Ik heb geslacht voor niemand als prioriteit gebruikt, en het evenwichtsquotum van de rij met hoge waarde is bepaald op leeftijdsgroep en wijk, de dimensies die werkelijk scheef trokken. De details staan in de bijlage.

Er staan geen telefoonnummers en geen e-mailadressen in de bestanden. De lijsten komen met identicator en naam; de balie moet ze naast de agenda leggen om de contactgegevens te vinden.

De conversiepercentages van 30% tot 40% komen niet uit jouw bestand, maar uit een marktreferentie voor terughaalacties. Ze zijn het zwakste getal van het rapport. Na de eerste week heb je je eigen getal, en dat is meer waard dan het mijne.

Bijlage

Geproduceerde bestanden

- `pagou_e_nao_veio.csv` (322 regels): gesorteerd op openstaand bedrag aflopend. Bevoordeelt wie een duur vaccin kocht: wie nog een BMR van R\$ 210 tegoed heeft, valt achteraan, ook al wacht hij al langer.
- `esquema_incompleto.csv` (322 regels): dezelfde sortering en dezelfde vertekening.
- `sumidos_sem_pendencia.csv` (630 regels): gesorteerd op tijdsband en daarbinnen op historische uitgaven aflopend. Bedient wie weinig uitgaf en altijd terugkwam minder goed.
- `gripe_anual_vencida.csv` (146 regels): gesorteerd op dagen sinds de laatste griepdosis, van de oudste naar de meest recente. Selectie uit de lijst van weggebleven patiënten, geen nieuwe namen.
- `alto_valor_em_risco.csv` (159 regels): gesorteerd op historische uitgaven aflopend, met het evenwichtsquotum op leeftijdsgroep en wijk aan het eind. Selectie uit de drie hoofdgroepen; niet bij hen optellen.
- `sem_gripe_ativos.csv` (1.027 regels): gesorteerd op historische uitgaven aflopend. Actieve patiënten die nooit griep kochten; bevoordeelt wie al veel uitgaf, wie alleen een goedkoop vaccin kocht valt achteraan.
- `cursos_repetidos_conferir.csv` (1.324 regels): gesorteerd op schemawaarde aflopend. Controlelijst voor poort 2.
- `aplicacoes_ultimo_dia_conferir.csv` (851 regels): gesorteerd op het interval tussen aankoop en toediening, oplopend, zodat de meest onwaarschijnlijke gevallen bovenaan staan. Controlelijst voor poort 1.

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

Hoe ik tot de groepen kwam

Term	Gebruikte definitie
Peildatum	30-04-2026, de laatste datum in welke datumkolom van het bestand dan ook
Schema	Set regels van dezelfde patiënt, hetzelfde vaccin en dezelfde aankoopdatum
Betaald en niet gekomen	Schema met nul gezette doses
Onvolledig schema	Schema met minstens één gezette dosis en minstens één openstaande
Weggebleven	Patiënt zonder enige openstaande dosis en zonder aankoop of toediening in meer dan 90 dagen
Banden van weggebleven	Recent 91-180 dagen, middellang 181-365 dagen, koud 366 dagen of meer
Verdachte herhaling	Tweede of latere aankoop van hetzelfde vaccin door dezelfde patiënt, griep uitgezonderd
Gemiddelde schemawaarde	R\$ 969 in het hele bestand (11.088 verkochte schema's); in de projecties van nieuwe verkoop het gemiddelde van de band zelf: R\$ 824 recent, R\$ 1.006 middellang, R\$ 960 koud
Contactkosten	R\$ 2,50, uit R\$ 30 per uur en 12 contacten per uur

De regel voor “onvolledig schema” gebruikt het protocol dat in het bestand zelf staat, in de kolom `protocolo_doses`, dat per vaccin consistent is (hepatitis B altijd 3, HPV-9 altijd 2, griep altijd 1, enzovoort). Omdat het protocol uit de data komt en niet uit een vuistregel, is het vertrouwen in deze groep hoog.

No-showpercentage per dimensie

Percentage van de verkochte doses dat niet is gezet, en het stilliggende bedrag in elke selectie.

Dimensie	Hoogste percentage	Laagste percentage
Betaalwijze	Boleto 16,0% (R\$ 236.418)	Pix 4,8% (R\$ 236.084)
Acquisitiekanaal	Google 13,0% (R\$ 339.371)	Doorverwijzing 4,3% (R\$ 193.379)
Dosisnummer	Derde 12,7%	Eerste 5,9%
Vaccin	Hepatitis B 9,0%	Pneumokokken 20 6,1%
Leeftijdsgroep	13-18 jaar 9,3%	46-60 jaar 7,0%
Wijk	Jardim das Acácias 8,7%	Morada do Sol 6,4%
Geslacht	Man 7,9%	Vrouw 7,4%

Twee negatieve lezingen die ook gelden: **wijk concentreert niets**. Het aandeel te benaderen patiënten per wijk ligt tussen 36% en 43% in de tien wijken, en de drie grootste samen zijn 32,7% van de rij, praktisch hun gewicht in het bestand. Een wijkcampagne heeft geen voordeel boven een algemene campagne, en daarom heb ik geen wijkbestand gemaakt. **De kolom geslacht moet nergens voor gebruikt worden**. Behalve dat ze bijna niets scheidt (7,9% tegen 7,4%), klopt ze niet met de namen: onder de 2.514 patiënten van wie de voornaam eindigt op “a” of “o”, komt de kolom in maar 52,3% van de gevallen overeen met de uitgang, hetzelfde als een munt opgooien. Of het veld is aan de bron verkeerd ingevuld, of het is in een deel van de registraties verwisseld. Ik heb de kolom in de bestanden gelaten, alleen zodat jij kunt nakijken, maar prioriteer er niet op en schrijf er geen bericht op.

Kolommen uit het bestand die in geen enkele berekening zijn meegenomen

`data_cadastro` (inconsistent, toegelicht bij de beperkingen) en `nome` uit de vaccinatietabellen, alleen gebruikt voor weergave in de lijsten. Alle andere kolommen van `pacientes.csv` en `vacinacoes.csv` zijn in minstens één dimensie hierboven meegenomen.

* **Over dit voorbeeld**. Dit is een echt rapport uit de Distilo-pipeline, gedraaid op synthetische data voor een fictief bedrijf. Dat bedrijf is Braziliaans: een vaccinatieliniek, dezelfde als in het voorbeeld over vaccins die samen gehaald worden; dit is het rapport over patiënten terughalen. Daarom staan de bedragen in Braziliaanse real, zijn de wijk- en patiëntnamen Braziliaans, en weerspiegelt een deel van het vaccinatieschema een Braziliaans programma. Het is gepubliceerd zoals het geproduceerd is en niet aangepast, omdat de analyse is wat we je laten zien en de analyse niet verandert met het land. Je eigen rapport is geschreven in jouw taal, over jouw bedrijf, in jouw valuta.