DRG afzetvoorspelmodel

# Business doelstellingen

Wekelijks worden pre-ROI berekeningen uitgevoerd bij DRG om de verwachte ROI per promotie te bepalen. Cruciale input hiervoor is de voorspelde promotie afzet. Deze wordt bepaald door middel van het afzetvoorspelmodel.

Dit model is verwerkt in de pre-ROI tool (welke vervangen zal worden door de pre-ROI applicatie), waarin de verschillende category managers wekelijks de promoties (ca 9 weken voordat deze plaatsvindt) doorrekenen. Op basis van de uitkomsten uit de tool, kunnen nog aanpassingen gedaan worden in deze promoties (bv andere plek in de folder, ander kortingsbedrag, multibuy, …)

Eind 2016 is het afzetvoorspelmodel voor Dirk opgeleverd. Echter, in Q1 en Q2 2017 is de modelvoorspelling dusdanig verslechterd en onstabiel dat besloten werd het hele model grondig te redesignen. Tegelijkertijd werden deze verbeterslagen toegepast op het afzetvoorspelmodel voor Dekamarkt dat nog ontwikkeld moest worden.

# Onderliggend datamodel



# Modelkeuzes en parameters

## Wat te voorspellen?

Het uiteindelijke doel is om de verwachte afzet per artikel te berekenen. Dit kan op verschillende manieren:

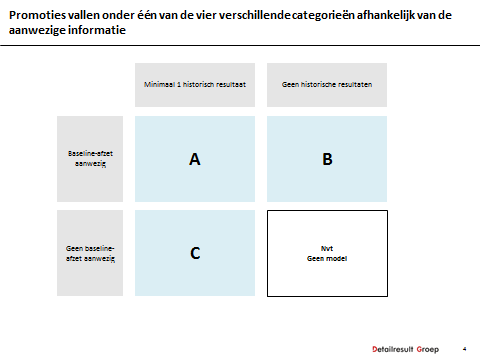
1. Absolute afzet
2. Baseline afzet \* uplift factor
3. Baseline afzet + absolute uplift

Alle drie de bovenstaande manieren zijn getest, en gemodelleerd met een lineair model. Hieruit kwam duidelijk naar voren dat methodiek 1, voorspelling van de absolute afzet, de meest nauwkeurige en robuuste voorspelling afgaf.

*NB. Dit kan komen doordat het baseline model nog niet stabiel/accuraat genoeg is. In de toekomst zou dus onderzocht kunnen worden wat de impact is van een verbeterde baseline voorspelling. Verwachting is wel dat methodiek 3 dan beter zal werken dan methodiek 2. Dit omdat bij methodiek 2 en kleine afwijking in de uplift factor een grote voorspelfout op kan leveren.*

## Welke modelcategorieën?

Voor een goede voorspelkracht is voldoende historische informatie cruciaal. Des te vaker een artikel in de promotie is geweest, des te meer informatie beschikbaar is. Om dit zo optimaal mogelijk te vatten zijn alle promoties ingedeeld in vier groepen:



**Hierbij geldt dat het volgende voorspeld wordt**

1. Delta absolute afzet t.o.v. best matches
2. Absolute afzet
3. Delta absolute afzet t.o.v. best matches

Per model zijn verschillende verklarende variabelen meegenomen:

**Model A**



**Model B**



**Model C**



## Matches:

Voor zowel model A als C is de bepaling van best matches cruciaal.

Een match is een historische promotie van hetzelfde artikel. Op basis van verschil in promotie-eigenschappen t.o.v. deze historische actie wordt dan een voorspelling gedaan over de verwachte delta.

Om een stabiel model te hebben is er voor gekozen om niet één, maar meerdere best matches meegenomen, als deze van voldoende kwaliteit zijn.

Voor de matchbepaling geldt het volgende stapppenplan:

1. Koppel alle historische acties van hetzelfde artikelnummer
2. Bepaalf of het een zuivere match betreft: als acties te verschillend zijn, dan willen we de historische actie niet gebruiken om een voorspelling mee te maken. Deze matches worden ‘onzuiver’ genoemd en uitgesloten:
   1. Voorpagina vs vloeraanbieding
   2. I/O-actie vs reguliere actie
   3. Weekend-actie vs reguliere actie
   4. A-actie (overal bestelbaar) vs B-actie (niet overal bestelbaar)
   5. Lay-in vs reguliere actie
   6. Schapactie vs regulier actie
   7. Gratis artikel vs betaald artikel
   8. Delta van meer dan 20 filialen
3. Bereken het matchpercentage o.b.v. verschillende gewichten (o.a. op basis van actieprijs, kortingspercentage, recentheid en plaats in de folder)
4. Bepaal welke matches meegenomen moeten worden:
   1. Beste match
   2. De tweede en derde match indien het matchpercentage hoger is dan 80%

Het is helaas niet mogelijk om voor elke promotie een even nauwkeurige voorspelling af te geven. Dit is sterk afhankelijk van fit van de matches. Om de gebruikers van de pre-ROI tool een indicatie te geven van de betrouwbaarheid van de voorspelling, wordt een stoplicht meegegeven. Hierbij geldt:

1. Groen: minimaal 1 historisch resultaat met matchpercentage <= 90%  
   🡪 Neem voorspelling 1-op-1 over
2. Oranje: minimaal 1 historisch resultaat met matchpercentage tussen 60% en 90%  
   🡪 Check historische matches en scherp voorspelling aan indien benodigd
3. Rood: Geen zuiver historisch resultaat, of matchpercentage < 60%  
   🡪 Geen betrouwbare voorspelling mogelijk; vul eigen prognose in

# Modeluitkomsten

## Validatie

* We berekenen wekelijks het verschil tussen de verwachte afzet uit de pre-ROI en de gerealiseerde afzet
* Dit gebeurt voor alle producten met afzet > 2.500 (Dirk) of afzet > 1.500 (Deka) waarvoor een pre-ROI berekening gemaakt is en die ook daadwerkelijk in promotie zijn geweest met de eigenschappen (in-en verkooprijs, plek in folder etc.) zoals die zijn ingegeven in de actielijst
* We bereken de op gerealiseerde afzet gewogen gemiddelde absolute afwijking van voorspelde afzet met gerealiseerde afzet
* Naast de vergelijking van de realisatie t.o.v. het gestelde target, wordt de voorspelfout ook vergeleken met de delta die momenteel gerealiseerd wordt vanuit DRG (prognose CGC)
* De uitdaging hierbij is dat de prognose vanuit CGC gebruikt wordt voor de insturing van artikelen naar de filialen. Hierdoor is te zien dat de prognose vanuit CGC bijna nooit significant lager ligt dan de realisatie.
* Om hiervoor te corrigeren wordt in het model ook rekening gehouden met de verdeelde aantallen, plus een marge voor eventuele voorraad in de filialen zelf.

De resultaten hiervan, ten tijde van de oplevering van het model, zijn hieronder te zien:





*NB. In de weekvalidatie is de afzetondergrens verwijderd, en worden alle producten meegenomen waarvoor de pre-ROI tool gebruikt wordt.*

## Modeloutput

To do: link naar database met volledige versie-historie

* + - Check met Tjeerd en Floris

# Model management

Om te waarborgen dat de modellen nauwkeurig en up to date blijven, is model management ingericht. Dit bestaat uit drie componenten:

1. **Model update** (wekelijks)
   1. Uitvoeren van de update (meest recente week)
   2. Weekvalidatie (totaal en outliers)
   3. Uitzetten van niet-representatieve matches (o.b.v. gecontroleerde outliers)
   4. Update tool met nieuwe promoties
2. **Model herijking** (maandelijks)
   1. Herschatting model parameters o.b.v. toegevoegde weken
   2. Kleine modelfixes en wensen
   3. Vergelijking oude met nieuwe model
   4. Update tool

*NB. Wordt hoofdzakelijk uitgevoerd door DRG*

1. **Model herijking** (half jaarlijks)
   1. Modeldoorontwikkeling o.b.v. wishlist
   2. Vergelijking oude met nieuwe model
   3. Update tool

Zie voor meer informatie: \\Mic-eas-fs01\data\Projecten\DetailResult\afzetvoorspelmodel\4. Model management

To do:

1. Stappenplan update
2. Handleiding tool / applicatie

# Referenties

Alle scripts die gebruikt zijn in de ontwikkeling en voor het updaten van het model staan in

[\\Mic-eas-fs01\data\Projecten\DetailResult\afzetvoorspelmodel](file:///\\Mic-eas-fs01\data\Projecten\DetailResult\afzetvoorspelmodel)

Betrokken personen

* Willemijn van Waarden (DRG)
* Dorthe van Waarden
* Esmee de Jong
* Jesper de Groot
* Tijmen Veltman
* Tjeerd van Campen