

De Kassabon

Vooraf op het einde van het jaar wordt het op de kleintjes letten. Een groot deel van je budget gaat op aan Sinterklaas- en Kerstinkopen. Niet alleen aan cadeautjes, maar ook in de supermarkt wordt het nodige besteed. Weinig omgevingen zijn vanuit informatiemanagement-oogpunt zo interessant als de supermarkt. Naast de logistieke stromen die moeten worden gecontroleerd, zijn er natuurlijk de kassasystemen, de loyalty-programma's en de koopgedrag-analyses.

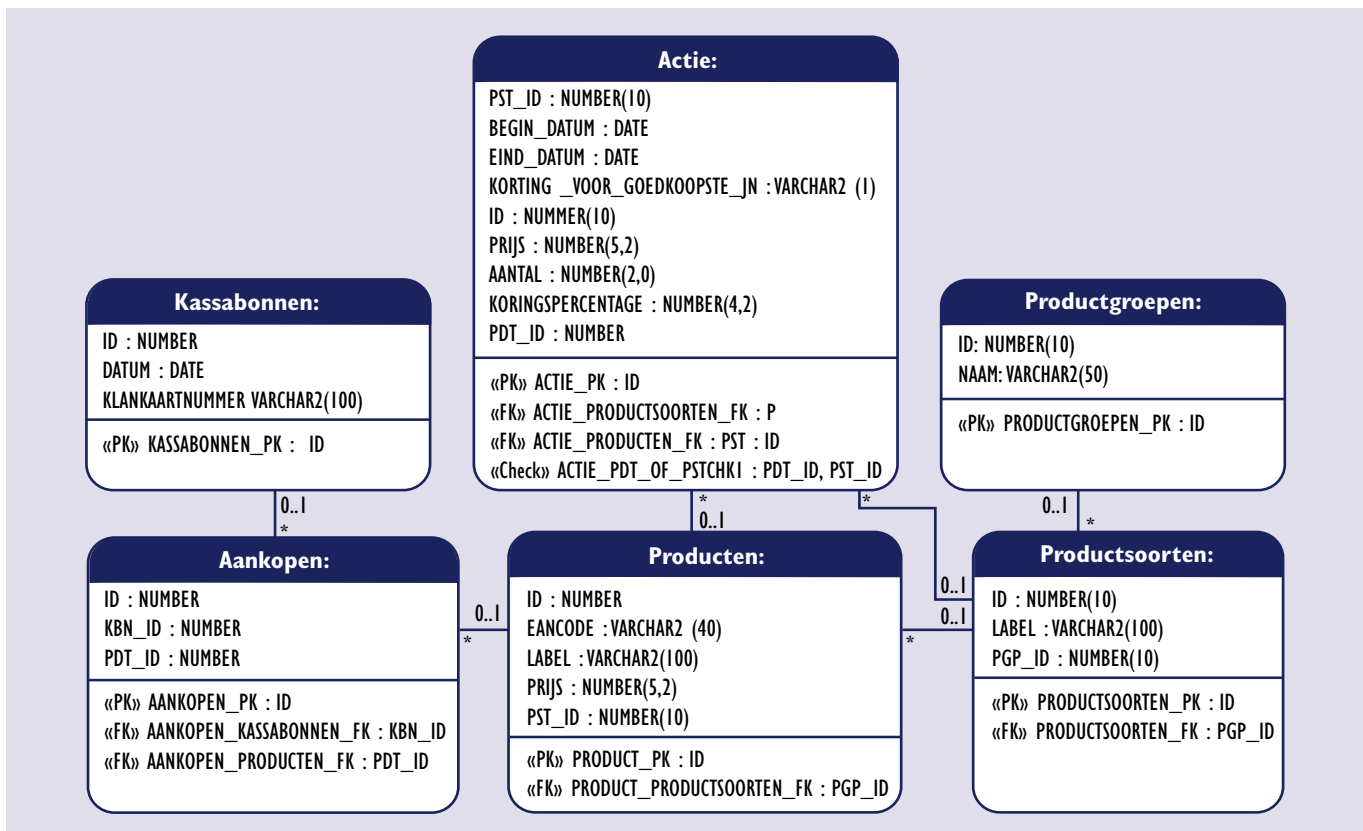
Hoe wordt optimaal gebruik gemaakt van de winkelruimte? Welke producten moet je naast elkaar zetten voor optimale verkoopmarges? Mijn vader, intussen gepensioneerd, was een kruiden-

nier die altijd zei "Dit is echt centenwerk. De marges zijn maar klein, dus hoe je winkel is ingedeeld is van wezenlijk belang".

Maar ook: hoe moet je, met al die kortingsacties, de totaalprijs op de kassabon berekenen? Dat laatste gaan we doen in deze puzzel.

Het onderstaande model doet, op primitieve wijze, drie dingen:

- Registratie van het assortiment
- Vastleggen van voordeelacties
- Registratie van alle aankopen en berekeningen van het totaal op de kassabon



De acties zijn van uiteenlopende aard – deze vormen het meest complexe onderdeel van het datamodel. Een actie kan van toepassing zijn op een speciaal product (Calvé Pindakaas, 400g) of op een productsoort (alle potten pindakaas).

Een actie kan een product simpelweg afprijzen: “Van prijs X voor Prijs Y”, of een bepaalde korting geven: “Nu product A met X% korting”. Vaak is een kortingspas van toepassing bij aankoop van meerdere exemplaren: “koop nu X exemplaren van product A voor prijs Y” of: “Drie halen, twee betalen (hetgeen hetzelfde is als bij aankoop van 3 exemplaren krijg je 33% korting).

Een actie kan ook betrekking hebben op een productsoort: “Drie potten pindakaas naar keuze voor 5 euro”. Of: “25% korting bij aankoop van 6 flessen rode wijn”. Vaak zit er een addertje onder het gras bij kortingen op productsoort, en slaat de korting alleen op het goedkoopste product: “Bij aankoop van 6 flessen rode wijn naar keuze, de goedkoopste met 50% korting”.

1.1 Uitdagingen

1) Wat is het (wat beperkte) assortiment van deze supermarkt? Toon de artikelen per (en op volgorde van) productgroep en productsoort.

Om deze vraag te kunnen beantwoorden hebben we in ieder geval drie tabellen nodig, te weten de Productgroepen, de productsoorten en de producten. Een query tegen deze drie tabellen is dan ook zo geschreven.

```
select pgp.naam
      ,pst.label
      ,pdt.label
      ,pdt.prijs
  from productgroepen pgp
      ,productsoorten pst
      ,producten      pdt
 where pdt.pst_id = pst.id
       and pgp.id = pst.pgp_id
 order by pgp.naam
        ,pst.label
        ,pdt.label;
```

Zoals je kunt zien, en wat ook te verwachten is, is dat de

NAAM	LABEL	LABEL	PRIJS
Alcoholische Dranken	Port	Finest Ruby Port	7.99
Alcoholische Dranken	Rode Wijn	Bordeaux Mas de la Maume, 2005	2.99
Alcoholische Dranken	Rode Wijn	Huiswijn, Bodega de las Torres	1.99
Alcoholische Dranken	Rode Wijn	Tempranillo Berberana	4.26
Alcoholische Dranken	Rode Wijn	Vino Tinto de Penedes, 2004	6.78
Broodbeleg	Hagelslag	De Ruyter Pure Hagelslag, 400g	2.89
Broodbeleg	Jam	Eigen Merk Jam, Abrikozen, 600g	1.69
Broodbeleg	Jam	Hero, Extra Confiture, Granaatappeljam, 375g	2.56
Broodbeleg	Pindakaas	Calve Pindakaas,	3.19
...			

Productgroep namen en de Productsoort labels meerdere keren in de resultaat set te zien zijn. Als je frequent gebruiker bent van SQL*Plus dan weet je dat misschien dat er een BREAK commando tot je beschikking staat. Met behulp van dit BREAK commando kun je repeterende namen laten onderdrukken.

```
SQL> break on naam
SQL> select pgp.naam
      2      ,pst.label
      3      from productgroepen pgp
      4      ,productsoorten pst
      5      ,producten      pdt
      6      where pdt.pst_id = pst.id
      7      and pgp.id = pst.pgp_id
      8      order by pgp.naam
      9      ,pst.label
     10      ,pdt.label
     11 /
```

NAAM	LABEL
Alcoholische Dranken	Port
	Rode Wijn
	Rode Wijn
	Rode Wijn
	Rode Wijn
Broodbeleg	Hagelslag
	Jam
	Jam
	Pindakaas
	Pindakaas
	Pindakaas

11 rows selected.

Hierboven zie je het resultaat van het BREAK commando. Nu heeft SQL*Plus deze mogelijkheid wel, en waarschijnlijke vele reporting tools ook, maar dit effect is ook mogelijk door handig gebruik te maken van Analytische functies.

Als we aan bovenstaande query de volgende Analytische Functie toevoegen dan weten we op welk niveau we een break moeten uitvoeren, we willen namelijk alleen de eerste tonen en de overige namen willen we onderdrukken.

```
row_number() over (partition by pgp.naam
                  order by pgp.naam
                  , pst.label
                  , pdt.label
                  ) pst_rn

SQL> select pgp.naam
      2      , pst.label
      3      , row_number() over (partition by pgp.naam
      4      order by pgp.naam
      5      , pst.label
      6      , pdt.label
      7      ) pst_rn
      8      from productgroepen pgp
      9      ,productsoorten pst
     10      ,producten      pdt
     11      where pdt.pst_id = pst.id
     12      and pgp.id = pst.pgp_id
     13      order by pgp.naam
     14      ,pst.label
```

```

15      ,pdt.label
16 /

```

NAAM	LABEL	PST_RN
Alcoholische Dranken	Port	1
Alcoholische Dranken	Rode Wijn	2
Alcoholische Dranken	Rode Wijn	3
Alcoholische Dranken	Rode Wijn	4
Alcoholische Dranken	Rode Wijn	5
Broodbeleg	Hagelslag	1
Broodbeleg	Jam	2
Broodbeleg	Jam	3
...		

eenvoudig. Met behulp van onderstaande query word een overzicht getoond van artikelen en de aantallen zoals deze vermeld staan op de kassabon.

Als we dan alleen de eerste uit de groep laten tonen, bijvoorbeeld via een CASE statement, dan hebben we hetzelfde effect bereikt als de SQL*Plus BREAK.

```

SQL> select case
2      when pst_rn = 1
3      then naam
4      end naam
5      , label
6      from (
7      select pgp.naam
8      , pst.label
9      , row_number() over (partition by pgp.naam
10     order by pgp.naam
11     , pst.label
12     , pdt.label
13     ) pst_rn
14     from productgroepen pgp
15     ,productsoorten pst
16     ,producten pdt
17     where pdt.pst_id = pst.id
18     and pgp.id = pst.pgp_id
19     order by pgp.naam
20     ,pst.label
21     ,pdt.label
22 )
23 /

```

NAAM	LABEL
Alcoholische Dranken	Port
	Rode Wijn
	Rode Wijn
	Rode Wijn
	Rode Wijn
Broodbeleg	Hagelslag
	Jam
	Jam
	Pindakaas
...	

Uiteraard kunnen we dezelfde truc voor de Productsoorten Label uithalen.

2) Welke artikelen zijn aangeschaft bij de transactie die kassabon I heeft opgeleverd? Geef de aantallen per product.

Aangezien alle transacties worden geregistreerd in de KASSABONNEN-tabel wordt het terugzoeken van transacties heel

```

with items as
( select pdt.label
  , pdt.prijs
  , kbn.datum
  from productsoorten pst
  , producten pdt
  , aankopen akp
  , kassabonnen kbn
  where akp.pdt_id = pdt.id
  and pdt.pst_id = pst.id
  and akp.kbn_id = kbn.id
  and kbn.id = 1
  order by pst.label
  , pdt.label
)
select label
  , count(*) aantal
  from items
  group by label
/

```

LABEL	AANTAL
Euroshopper, Supervoordeel, 1 kg	2
Tempranillo Berberana	1
Huiswijn, Bodega de las Torres	2
Eigen Merk, Met nootjes, 450 g	1
Bordeaux Mas de la Maume,2005	1
Vino Tinto de Penedes, 2004	1
Calve Pindakaas, grootverpakking, 700 gr	2

De query begint met de WITH Clause, in de Oracle documentatie wordt dit Subquery Factoring genoemd. Deze techniek is het best te vergelijken met een inline view. Vooraf definieer je een naam voor de inline view en deze kun je gebruiken in de rest van het SQL statement. Uiteraard is deze query ook met een inline view te schrijven.

De reden dat er meer tabellen worden gebruikt dan strikt noodzakelijk is vooruitlopend op de derde uitdaging.

Tot nu toe hebben we nog geen gebruik gemaakt van “speciale aanbiedingen”. Hiervoor verwijst ik u naar de online-versie van dit artikel, waar u ook de scripts kunt downloaden om zelf met de verschillende uitdagingen aan de slag te gaan.

De rest van de opgaven en de online-versie kunt u vinden op <http://technology.amis.nl/blog/?p=2625>

Anders nog iets?

Anton Scheffer en Alex Nuijten zijn werkzaam bij AMIS.

Advertentie