DATAWAREHOUSE

Groep: C

KLAS: 1

Groep 1C

|  |  |
| --- | --- |
| Studentnummer | Naam + voornamen |
| 19081391 | Daniël den Heijer |
| 19131771 | Noah El Menyari |
| 19009437 | Jesse Pardoen |
| 11067578 | Wayl Hamham |
|  |  |

Inhoudsopgave

[1. Inleiding 3](#_Toc67526000)

[2. Planning 4](#_Toc67526001)

[3. Managementvragen Greatoutdoor 5](#_Toc67526002)

[4. Datawarehouse ontwerp 6](#_Toc67526003)

[4.1 Dimensie matrix 7](#_Toc67526004)

[4.2 Feitmodel 8](#_Toc67526005)

[4.3 Dimensie & dimensie hiërarchieën 9](#_Toc67526006)

[4.4 Conformed dimension 10](#_Toc67526007)

[4.5 Sterschema 11](#_Toc67526008)

[5. Datawarehouse prototype 12](#_Toc67526009)

[5.1 Create scripts 12](#_Toc67526010)

[5.2 ETL 13](#_Toc67526011)

[5.3 Test/waarborginggegevenskwaliteit 14](#_Toc67526012)

[5.4 Aanpassingen in ontwerp en datawarehouse als gevolg van beveiligingsmaatregelen 15](#_Toc67526013)

[6. Overzicht van de Dashboards 16](#_Toc67526014)

[7. Reflectie & bewijslast 17](#_Toc67526015)

[Literatuurlijst 18](#_Toc67526016)

# Inleiding

De organisatie Great Outdoor heeft het bedrijf A&C overgenomen. A&C is een bedrijf die sportieve kleding verkoopt en onderscheidt zich door merkloze kleding op de markt te brengen met grappige tekst erop. Great Outdoor is de nr.1 op het gebied van buitenrecreatie en levert alleen aan bedrijven. De CEO van Great Outdoor zou graag willen weten hoe het gaat in A&C en heeft een team ingeschakeld om een datawarehouse te realiseren voor Great Outdoor en A&C.

Dit document is bedoeld voor de ontwikkeling van de datawarehouse voor Great Outdoor en A&C. Om dat te realiseren is er een planning nodig en een takenverdeling, dit zal uiteengezet worden in hoofdstuk 2. In hoofdstuk 3 zullen de managementvragen opgesteld worden en beantwoord worden. Vervolgens zal de datawarehouse ontworpen worden in hoofdstuk 4. In hoofdstuk 5 en 6 zal er een prototype ontworpen worden voor de datawarehouse en een ontwerp uitgebeeld worden in dashboards. Als laatste zal iedere teamlid zijn werk reflecteren en documenteren voor bewijslast.

# Planning

Om de datawarehouse te realiseren zal er gewerkt worden met de projectontwikkelingsmethode scrum. Iedere week is een sprint waarin taken worden verdeeld, als er een taak niet afgemaakt is zal de taak meegenomen worden naar de volgende sprint. In totaal zullen er vijf sprints zijn om het product te realiseren. Dagelijks wordt er een daily standup georganiseerd om te praten over welke taken er gemaakt zijn en welke vandaag gemaakt zullen worden. Aan het einde van de week is er een retrospective om te kijken wat er in de week goed is gegaan en wat niet. Dit zal dan gedocumenteerd worden in dit document. Daarnaast worden de taken verdeeld en opgenomen in dit hoofdstuk.

De reden waarom er voor dit methode is gekozen omdat dit een succesvolle methode is om ermee te ontwikkelen. Het is een flexibele methode waarbij er dagelijks de kwaliteit van het product gewaarborgd wordt. Want er is immers dagelijks een vergadering en iedereen weet van elkaar wat de stand van zaken zijn.

**Sprint 1:**

In deze sprint is het belangrijk dat alle taken samen uitgevoerd worden. Het is namelijk belangrijk dat iedereen meedoet aan het opzet van het project. Zo weet iedereen immers de stand van zaken in dit project.

|  |  |  |
| --- | --- | --- |
| Datum | Tijdsduur | Opdracht |
| 29/03/2021 11:00 | 2 uur | Opzetten van SQL en database |
| 31/03/2021 13:00 | 1 uur | Managementvragen bedenken |
| 01/04/2021 14:00 | 2 uur | Opdrachten beantwoorden en goed verwoorden in het portfolio hoofdstuk 3 |

**Sprint 2:**

In deze sprint is het belangrijk dat ieder team lid naast de gezamenlijke taken een ander taak heeft. Hoofdstuk vier wordt onderverdeeld en ieder zal een taak ontvangen.

|  |  |
| --- | --- |
| Paragraaf | Naam |
| 4.1 en 4.2 | Jesse |
| 4.3 | Daniël |
| 4.4 | Wayl |
| 4.5 | Noah |

|  |  |  |  |
| --- | --- | --- | --- |
| Datum | Tijdsduur | Opdracht | Persoon |
| 07/04/2021 14:00 | 1 uur | Inlezen in de filmpjes en uitleg | Iedereen |
| 07/04/2021 15:00 | 2 uur | Individueel de opdracht bekijken en proberen eraan te beginnen. | Iedereen |
| 11/04/2021 Hele dag | 1 á 2 uur de man | Het verdelen van de taken onderling en het extra inlezen van de stof. | Iedereen |
| 11/04/2021 17:00 | 3 uur | Dimensiematrix en feitmodel maken. Samen met het introductie verhaal van het datawarehouse ontwerp. | Jesse |
| 11/04/2021 19:00 | 2 uur | Hiërarchieën proberen te begrijpen en de tabel ervoor maken | Daniël |
| 12/04/2021 10:00 | 2 uur | Hiërarchieën maken | Daniël, Jesse en Noah |
| 12/04/2021 12:00 | 2 uur | Sterschema maken | Noah |

**Sprint 3:**

In deze sprint zal de focus liggen op het afmaken van hoofdstuk vier, voornamelijk op 4.4 en 4.5. Daarnaast is het ook de bedoeling dat er een klein begin gemaakt wordt op de create scripts van volgende week.

|  |  |  |  |
| --- | --- | --- | --- |
| Datum | Tijdsduur | Opdracht | Persoon |
| 13/04/2021 | 2 uur | Inlezen in de stof en de filmpjes nader bestuderen | Iedereen |
| 14/04/2021 | 0.5 uur | Gezamelijke Planning maken | Iedereen |
| 15/04/2021 | 4 uur | Starschema afmaken | Noah |
| 16/04/2021 | 3 uur | Conformed dimensions afmaken | Wayl |
| 18/04/2021 | 2 uur | Toelichting sterdiagram uitbreiden en uitleg schrijven over methode die we gebruiken om, om te gaan met SCD's | Jesse |

**Sprint 4:**

|  |  |  |  |
| --- | --- | --- | --- |
| Datum | Tijdsduur | Opdracht | Persoon |
| 19/04/2021 | 3 uur | Inlezen in de stof en de filmpjes nader bestuderen | Iedereen |
| 20/04/2021 | 0.5 uur | Gezamelijke Planning maken | Iedereen |
| 21/04/2021 | 3 uur | Testing | Noah |
| 22/04/2021 | 3 uur | beveiliging | Danïel |
| 22/04/2021 | 3 uur | ETL Package met SSIS maken | Wayl |
| 23/04/2021 | 3 uur | Create scripts maken | Jesse |

**Sprint 5:**

|  |  |  |  |
| --- | --- | --- | --- |
| Datum | Tijdsduur | Opdracht | Persoon |
| 28/04/2021 | 3 uur | Inlezen in stof en aanmaken lokale staging database | Iedereen |
|  |  |  |
| 2 uur | Begrijpen wat de stof inhoudt en hoe ik het het beste aan kan pakken | Iedereen |
| 2 uur | Begin maken maken/oefenen met de lookup. | Iedereen |
| 29/04/2021 | 3 uur | Inlezen ETL's en ETL's maken voor de dimensietabellen in de DWH | Iedereen |
|  |  |  |
| 4 uur | Kijken wat ik met ETL’s kan en hoe ik deze toe kan voegen met de omgeving. | Iedereen |
|  |  |  |
| 30/04/2021 | 3 uur | ETL's maken voor de feittabellen in de DWH | Jesse |
| 3 uur | Begin met testing | Wayl |
| 2 uur | De ETLs werkend krijgen en de errors wegwerken | Daniël |
| 5 uur | Mapping schema problemen afmaken/bijwerken, ETL maken voor feittabellen | Noah |
| 01/05/2021 | 3 uur | Troublefixing DWH informatie overzetten vanuit staging DB | Jesse |
| 3 uur | Samen met Noah errors wegwerken in lokale ETL | Wayl |
| 1 uur | ETL errors bij lookups geprobeert te fixen | Daniël |
| 3 uur | Errors wegwerken ETL, definitieve run. | Noah |
| 02/05/2021 | 4 uur | Overzetten informatie vanuit DWH naar PowerBI en begin managementvraag 3. | Jesse |
| 2 uur | Kennismaken met PowerBI, onderzoek doen naar testen van de database | Wayl |
| 2 uur | PowerBI connecten en proberen de managementvragen te beantwoorden en inlezen in de security | Daniël |
| 2 uur | Power BI downloaden, verbinden met SQL en gegevens transformeren. Eerste grafieken van de eerste managementvraag. | Noah |
| 03/05/2021 | 5 uur | Het maken van tests | Wayl |
| 03/05/2021 | 6uur | Afmaken managementvraag 3 en begin managementvraag 4 in PowerBI, samen met uitleg over uitwerkingen. Evenals opnames ETL dataflows in het ETL hoofdstuk van het portfolio. | Jesse |
|  |  |  |
| 9 uur | Security toepassen. Alle etls opnieuw in vs2019 zetten en alles opnieuw integreren omdat het niet meer werkte. Verder heb ik management vraag 2 beantwoord en de puntjes op de I gezet. | Daniël |
|  | Managementvraag 1 beantwoorden, fouten verbeteren in PowerBI gegevens. | Noah |

# Managementvragen

In dit hoofdstuk komen de managementvragen aan bod. Dit houdt in dat de CEO van Great Outdoor een aantal vragen heeft die teruggezien zullen worden in de datawarehouse. De volgende managementvragen zullen gerealiseerd worden, namelijk:

Managementvragen:

1. Hoeveel omzet is er in een jaar behaald, per product en per bestelmethode en wat waren aantal producten die hierbij zijn verkocht?
   1. Je kan de verkoopprijs en aantal uit de database halen en vermenigvuldigen voor de omzet. Vervolgend kan je dit ook per product sorteren en de bestelmethode, deze staan in ORDER\_DETAILS en ORDER\_HEADER
2. Wat waren de opbrengsten en targets per kwartaal van elk product en welke employees hebben de targets wel en welke hebben de targets niet gehaald en welke cursussen hebben deze employees allemaal gevolgd?  
    a. Je kan de orders sorteren op datum en op employee, dit kan je met de order details tabel dan ook sorteren op product. Hierna kan je de target erbij houden uit de sales target tabel en kijken wie het wel en wie het niet gehaald hebben. Van employees die het wel gehaald hebben staan de courses in de GO Training excel data.
3. Wat waren de namen van de meest teruggebrachte producten per retailer en per jaar

en wat was het jaar waarin in totaal de meeste producten teruggebracht werden en wat waren in dit jaar de top 5 retailers met de meeste teruggebrachte producten?

Select a.Return\_date, a.order\_detail\_code, a.return\_quantity, c.product\_name, f.company\_name  
From returned\_item a join order\_details b on a.order\_detail\_code = b.order\_detail\_code  
 join product c on c.product\_number = b.product\_number

Join order\_header d on b.order\_number = d.order\_number

Join retailer\_site e on d.retailer\_site\_code = e.retailer\_site\_code

Join retailer f on e.retailer\_code = f.retailer\_code  
order\_by return\_quantity desc  
limit 5

1. Hoeveel omzet is er gemaakt per product per promotie in een maand en hoeveel producten zijn er totaal in die promotie verkocht?

Select pr\_number, start\_date, end\_date

From campaign

# Datawarehouse ontwerp

In dit hoofdstuk worden de feiten en uitleg van de relatie met managementvragen uiteengezet.

Ons eerste feit is, **het bedrijf verkoopt producten**, dit feit slaat op een groot deel van de managementvragen omdat we bij een groot deel erachter moeten komen wanneer producten verkocht worden en hoeveel dat er zijn, hierdoor bevat onze feitentabel genaamd “**Orders”, de punten product\_number, unit\_cost, quantity en unit\_price**. Bij dit feit horen de dimensies; **tijd, product, locatie en bestelmethode**. Al deze informatie wordt uit de tabel order\_details gehaald, zoals te zien in het feitmodel hieronder.

“**Klanten retourneren producten**” is ons derde feit, dit gaat vooral om het terugbrengen van de producten door de klanten, wij willen erachter komen wat wordt teruggebracht en hoeveel er wordt teruggebracht en hiervoor gebruiken wij in de **feitentabel “Retour”**, **return\_date en return\_quantity**. De dimensies hier zijn **tijd, product, retour en locatie**.

Om managementvraag 2 te beantwoorden is het handig om te weten welke werknemers cursussen hebben gelopen, die werknemers zoeken we uit het met het feit “**werknemers lopen cursussen**”, hier vallen de **sales\_staff en course categorieën van de GO training database** onder, dit is een andere database dan diegene die gebruikt wordt voor informatie van de andere feiten, dat betreft namelijk de GreatOutdoor2 database. De dimensie die hierbij hoort is de **medewerker** dimensie.

Het feit “**Medewerkers krijgen sales targets van het bedrijf**” slaat vooral op managementvraag 2, hier willen wij erachter komen of de werknemers de targets halen en of die werknemers ook cursussen gelopen hebben, zoals hierboven uitgelegd. De feitentabel voor dit feit heet **Sales\_targets** en bevat **sales\_staff\_code, product\_number en sales\_target**. De dimensies die hier gebruikt worden zijn **tijd, product, werknemer en targets**.

Het feit “**Het bedrijf maakt gebruik van promoties”** gaat over de promoties die gebruikt worden om producten goedkopere prijzen te geven en zo meer afzet te creëeren, om erachter te komen hoeveel er hierdoor verkocht wordt hebben wij de feitentabel **Sales\_promotie** gemaakt waar wij met behulp van **pr\_number, date\_start en date\_end** erachter komen wanneer welk product in de aanbieding is en zo kunnen zien of iets meer verkoopt als het in de aanbieding is en wat dan de betreffende omzet is. De dimensies hier zijn **tijd, product, promotie en bestelmethode**.

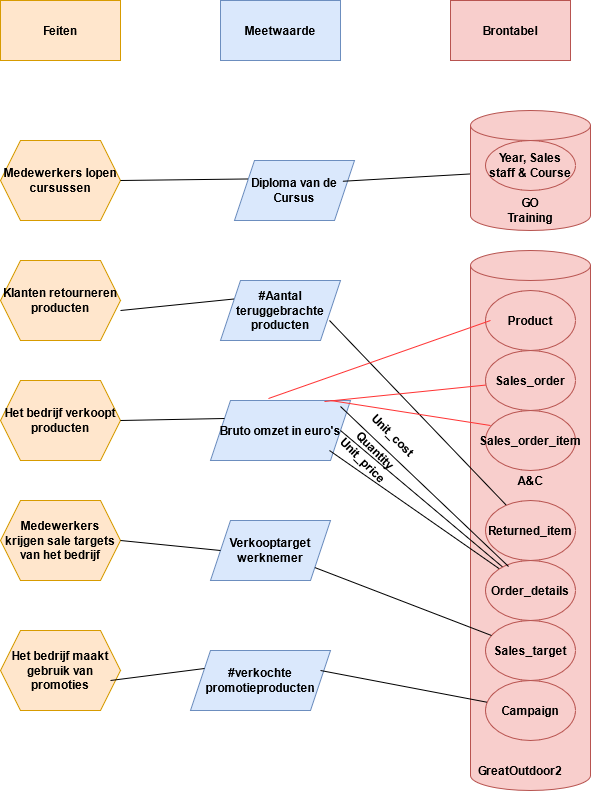
### 4.1 Dimensie matrix

Hieronder volgt de dimensie matrix,

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Dimensie ->**  **Meetwaarde v** | Tijd | Product | Werknemer | Locatie | Bestelmethode |
| Omzet in euro's | X | X |  | X | X |
| # Teruggebrachte producten | X | X |  | X |  |
| Cursus medewerker |  |  | X |  |  |
| verkooptarget werknemer | X | X | X |  |  |
| # Afzet per promotie | X | X |  |  | X |

Iedere dimensie in het tabel heeft een bepaalde meetwaarde, voor werknemer is dat alleen cursus medewerker en verkooptarget werknemer. Voor tijd en product is dat alles behalve cursus medewerker. Voor locatie is dat omzet in euro’s en teruggebrachte producten en voor bestelmethode is dat afzet per promotie en omzet in euro’s.

### 4.2 Feitmodel

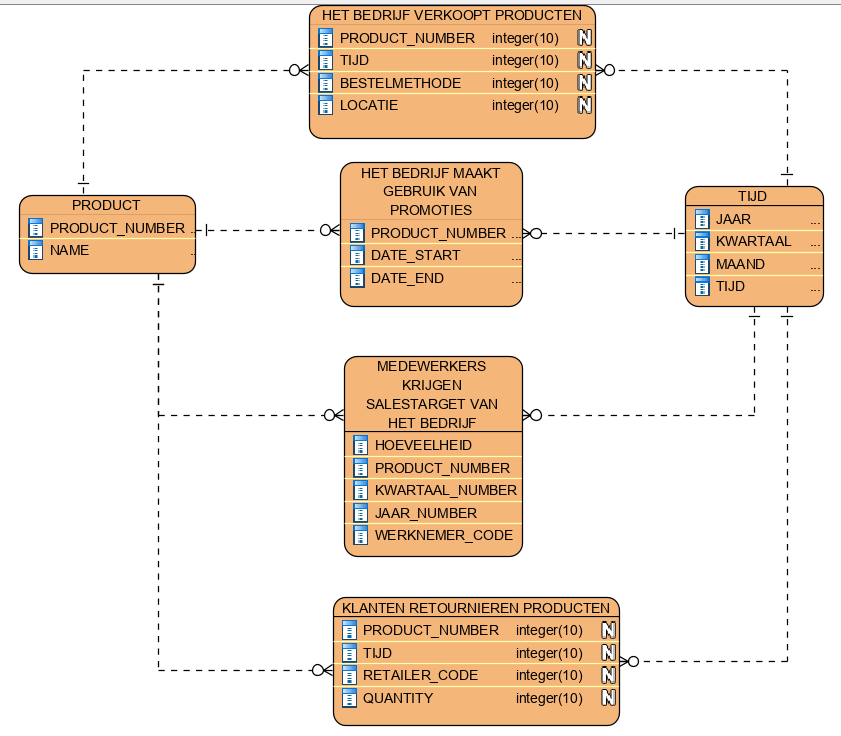


Hier zullen wij toelichting geven op het feitmodel en de keuze die gemaakt zijn bespreken ervan uitgaand dat het bovenste feit nummer 1 is.   
  
Het eerste feit is dat de medewerkers cursussen lopen, dit slaat op managementvraag 2, zoals hierboven te zien is dit feit de enige die slaat op de database van GO training, we wilden het hebben van een diploma van de cursus gebruiken als meetwaarde, maar dit kan nog aangepast worden omdat het ook mogelijk is dat er geen zinvolle wijze is om dit te kwantificeren.   
  
Daaronder is het feit genoteerd dat gaat over de retours van de klanten, de meetwaarde die wij hierbij gebruiken is het aantal teruggebrachte producten, hiermee bedoelen wij het aantal orders voor retours maar ook het aantal teruggebrachte producten binnen die orders. Al deze informatie is uit de Returned\_item tabel te halen.  
  
Om bij feit 3 achter de omzet in euro's te komen hebben wij de kosten van de producten, de prijs van de producten en de hoeveelheid bestelde producten nodig. Deze informatie is allemaal uit de order\_details tabel te halen.  
  
Bij feit 4 gaan wij ons richten op de targets die de werknemers krijgen, die targets zijn allemaal uit de sales\_target tabel te halen, samen met de betreffende werknemers.  
  
Als laatste gebruiken wij bij feit 5 de tabel campaign, deze was iets lastiger, omdat de informatie over de promoties aardig verspreidt lag, maar binnen de campaign tabel valt de benodigde informatie over de producten en over de promoties die nodig is om achter de meetwaarde te komen.

### 4.3 Dimensie & dimensie hiërarchieën

|  |  |  |  |
| --- | --- | --- | --- |
| Dimensie | Hiërarchieën | Onderbouwing | Surrogaatsleutel |
| Product | PRODUCT  Product\_number, name | Om de sales, omzet, teruggebrachte producten per product weer te kunnen geven hebben we het nummer (PK) van product nodig en de naam van het product. | Productkey |
| Locatie | LOCATIE retailer\_code, name | Locatie slaat op de retailer tabel, om per retailer teruggebrachte producten weer te geven hebben we de retailer code en naam van de retailer nodig. | Retailerkey |
| Werknemer | WERKNEMER werknemer\_code, first\_name, last\_name | De werknemer (sales\_staff) dimensie wordt gekoppeld aan de cursussen en aan de sales\_target, hiervoor hebben we de werknemer code nodig en naam van de werknemer. | Werknemerkey |
| Tijd | JAAR nummer | Tijd valt in te delen in een datum, een datum zit in een bepaalde maand, de maand bevindt zich in een kwartaal en een kwartaal in een jaar. (volgorde moet andersom van beneden naar boven) | Tijdkey |
| KWARTAAL nummer |
| MAAND naam |
| TIJD datum |
| Bestelmethode | BESTELMETHODE order\_method\_code, name | Is het methode van bestellen bijvoorbeeld creditcard, ideal of andere betaalmethodes. | bestelmethodekey |

### 4.4 Conformed dimension

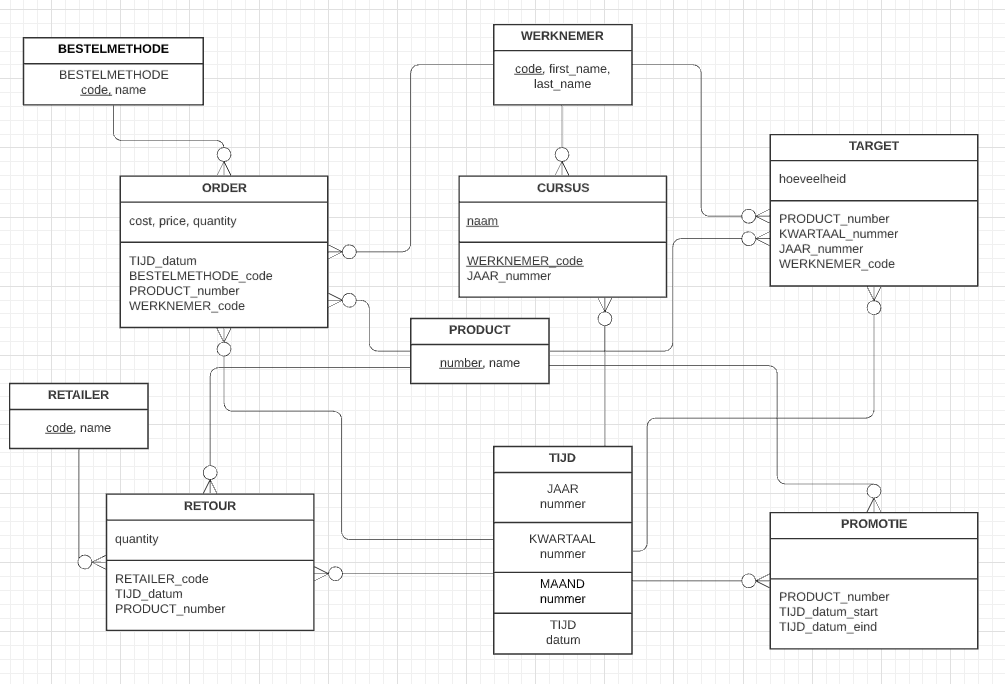


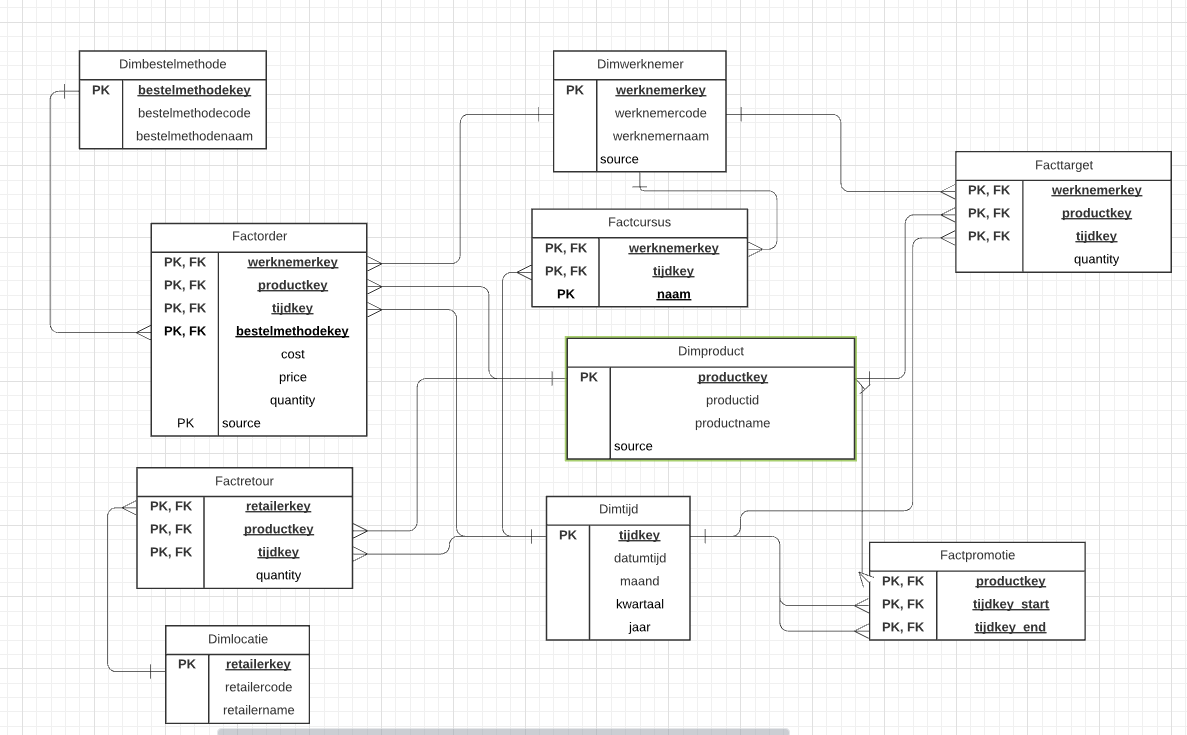
**Toelichting:**

Uit de dimensiematrix zijn er vier managementvragen die met elkaar te maken hebben. Bij bijna alle management vragen is tijd en product zeer belangrijk, die worden namelijk bij elke vraag gedeeld. Dus er kan gezegd worden dat tijd en product een zeer belangrijk rol gaat spelen bij het opstellen van de ETL's. Daarnaast is Locatie ook een belangrijk conformed dimension. Deze dimensies zullen uiteindelijk een rol spelen in het vullen van de facts.

### 4.5 Sterschema

Onderstaande foto's zijn twee uitwerkingen van ons sterschema. De bovenste was onze eerste schets, niet alle lijntjes zitten hier op de goede plek. De tweede foto is een verbeterd sterschema, waarin nu ook alle surrogaatsleutels zijn verwerkt.





**Toelichting**:  
Dit sterschema bevat vijf feiten. De feiten zijn uitgewerkt met de bijbehorende dimensies. Het eerste feit: bedrijf verkoop product wordt in het “ORDER” feittabel uitgewerkt. Het tweede feit: Klant retourneert product wordt in de “RETOUR” feittabel uitgewerkt. Het derde feit: bedrijf doet een promotie wordt in de “PROMOTIE” feittabel uitgewerkt. Het vierde feit: Medewerker haalt target wordt in de “TARGET” feittabel uitgewerkt. Het vijfde feit: Medewerker volgt een cursus wordt in de “CURSUS” feittabel uitgewerkt.

Voor onze eerste managementvraag, “Hoeveel omzet is er in een jaar behaald, per product en per bestelmethode en wat waren aantal producten die hierbij zijn verkocht?“, willen we de omzet en bijbehorende aantal producten per jaar, per bestelmethode, per product weten, hierdoor gebruiken we het feittabel “ORDER”, aangezien het gaat om het aantal bestellingen. Daarbij koppelen we het aan de TIJD (per jaar), de BESTELMETHODE (per bestelmethode) en de PRODUCT dimensie (per product).

Voor onze tweede managementvraag, “Wat waren de opbrengsten en targets per kwartaal van elk product en welke employees hebben de targets wel en welke hebben de targets niet gehaald en welke cursussen hebben deze employees allemaal gevolgd”, willen we de opbrengsten en targets per kwartaal, per product weten. We willen ook weten per medewerker of ze de targets wel hebben behaald of niet en of de betreffende medewerkers eventuele cursussen gevolgd hebben. De feittabel die wij vooral gebruiken bij deze managementvraag zijn “TARGET” en “CURSUS”. We willen namelijk weten wat de targets precies waren, en welke cursussen gevolgd worden. Hiernaast gebruiken wij de tabellen TIJD (per kwartaal), PRODUCT (per product) en WERKNEMER (per werknemer).

Onze derde managementvraag, “Wat waren de meest teruggebrachte producten per retailer per maand en wat was hier de maand waarin de meeste producten teruggebracht werden?”, slaat op de feittabel “RETOUR”, we willen hier erachter komen wat het product is dat het vaakst teruggebracht werd, per retailer, per maand. Ook willen we erachter komen in welke maand de meeste retourorders geweest zijn. Om achter al deze antwoorden te komen maken we gebruik van de TIJD (per maand), RETAILER (per retailer) en PRODUCT (per product) tabellen.

Onze laatste managementvraag luidt “Hoeveel omzet is er gemaakt per product per promotie in een maand en hoeveel producten zijn er totaal in de promotie verkocht?”. Bij deze managementvraag willen we vooral kijken naar de omzet van de promoties, om zo te zien welke promoties er het meeste binnenhalen en welke producten het meest in trek zijn tijdens deze promoties. Hiervoor gebruiken we de “PROMOTIE” feittabel. Om erachter te komen wat de meest verkochte producten waren tijdens de promoties en hoeveel er in totaal is opgebracht tijdens de promoties moeten we erachter komen wanneer de promoties plaatsvonden, die informatie halen we uit de TIJD tabel, hier pakken we de start datum en eind datum van een promotie uit en met behulp hiervan, en van de productsoort uit de PRODUCT tabel, Kunnen we erachter komen wat er het meest verkocht is in die tijd.

In het tweede diagram is het sterschema opnieuw gemaakt, dit maal met de surrogaatsleutels erin verwerkt. Dit zijn unieke keys die de oude identity vervangen. Deze oude identity komt in een nieuwe kolom die niet in de primary key zit.

Het zou erg goed kunnen dat er sprake is van SCD's in onze database, denk hierbij aan productnamen die veranderen of een werknemer die zijn voornaam verandert. Er zijn meerdere manieren om, om te gaan met SCD's, maar ons lijkt het, het beste om gebruik te maken van de methode die het meest toekomstbestendig is en waarbij zo min mogelijk onduidelijkheid kan plaatsvinden. Wij willen nieuwe records aanmaken wanneer nodig, hierbij geven wij dan tevens ook de begin- en einddatum van de records aan waardoor er zo min mogelijk onduidelijkheid ontstaat.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Target Table** | **Target Column** | **Data-type** | **Source Table** | **Source Column** | **Data-type** | **Expression** | **Default Value if Null** | **Data Issues/Quality/Comments** |
| Name of target table | Column in target table | Data type for this target column | Name of the Source table | Column in source table where you would get this data from | Data type for this source column | Decodes, aggregates, conversions, if statements, lookup functions…… | Value to use in target field when source field is null | Used to document Not null, value if looked up, upper case, pk, fk, etc…comments, issues |
| DimProduct | ProductKey | Int |  |  |  | SK - identity |  | PK |
|  | ProductId | Int | GO: product | Product\_number | Int |  |  |  |
| AC: product | id | int |
|  | ProductNaam | Varchar(255) | Go: product | Product\_name | Varchar(255) |  |  |  |
| AC: product | Description | Varchar(255) |
|  | | | | | | | | |
| DimLocatie | RetailerKey | Int |  |  |  | SK - identity |  | PK |
|  | RetailerCode | Int | GO: Retailer | Retailer\_code |  |  |  |  |
|  | RetailerNaam | Varchar(255) | GO: Retailer | Company\_name | Varchar(50) |  |  |  |
|  | | | | | | | | |
| DimWerknemer | WerknemerKey | Int |  |  |  | SK - identity |  | PK |
|  | WerknemerCode | Int | GO: sales\_staff | Sales\_staff\_code | Int |  |  |  |
| AC: employee | Emp\_id | Int |
|  | WerknemerNaam | Varchar(255) | GO: sales\_staff | First\_name + Last\_name | Varchar(255) |  |  |  |
| AC: employee | Emp\_fname + Emp\_lname | Varchar(255) |
| DimBestelmethode | Bestelmethodekey | Int |  |  |  | SK - identity |  | PK |
|  | Bestelmethodecode | Int | GO: order\_method | Order\_method\_code | Int |  |  |  |
|  | Bestelmethodenaam | Varchar(255) | GO: order\_method | Order\_method\_en | Varchar(255) |  |  |  |
| FactOrder | Bestelmethodekey | Int | GO: order\_header |  |  | Lookup naar DimBestelmethode met order\_method\_code (GO) |  | FK, PK |
|  | Werknemerkey | Int | GO: order\_header, AenC: sales\_order |  |  | Lookup naar DimWerknemer met sales\_staff\_code(GO) / sales\_rep (AenC) |  | FK, PK |
|  | Productkey | Int | GO: product, AenC: sales\_order\_item |  |  | Look up naar DimProduct met product\_number (GO) / prod\_id (AenC) |  | FK, PK |
|  | Tijdkey | Int | GO: order\_header, AenC: sales\_order |  |  | Lookup naar DimTijd met order\_date |  | FK, PK |
|  | Price | Float | GO: order\_details | Unit\_price | Float |  |  |  |
| AenC: product | Unit\_price | Double(15,5) |
|  | Quantity | Int | GO: order\_details | Quantity | Smallint |  |  |  |
| AenC: product en sales\_order\_item | Product(Quantity) \* sales\_order\_item(quantity) | Int | Quantity uit product vermenigvuldigen met quantity uit sales\_order\_item |  |
| FactPromotie | Productkey | Int | GO: campaign |  |  | Lookup naar DimProduct met product\_number (GO) |  | FK, PK |
|  | Tijdkey\_start | Int | GO: promotion |  |  | Lookup naar DimTijd met date\_start (GO) |  | FK, PK |
|  | Tijdkey\_eind | Int | GO: promotion |  |  | Lookup naar DimTijd met date\_end (GO) |  | FK, PK |
|  | Discount | Float | GO: campaign | Discount | Float |  |  |  |
| FactRetour | Retailerkey | Int |  |  |  | Lookup |  | FK, PK |
|  | Productkey | Int |  |  |  | Lookup |  | FK, PK |
|  | Tijdkey | Int |  |  |  | Lookup |  | FK, PK |
|  | Quantity | Smallint | GO: returned\_item | Return\_quantity | Smallint |  |  |  |
| FactTarget | Werknemerkey | Int |  |  |  | Lookup |  | FK, PK |
|  | Productkey | Int |  |  |  | Lookup |  | FK, PK |
|  | Tijdkey | Int |  |  |  | Lookup |  | FK, PK |
|  | Quantity | Float | GO: Sales\_target | Sales\_target |  |  |  |  |
| FactCursus | Werknemerkey | Int |  |  |  | Lookup |  | FK, PK |
|  | Tijdkey | Int |  |  |  | Lookup |  | FK, PK |
|  | Naam | Varchar(255) | GO Training Excel sheet | Course | Plain text |  |  |  |

# Datawarehouse

### 5.1 Create scripts

**Staging database**

****

**OLAP database**



### 5.2 ETL



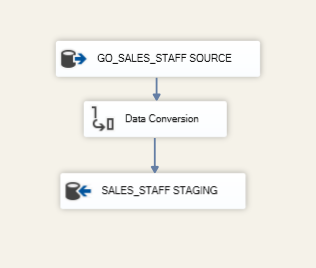
In 5.1 zijn de create scripts aangemaakt dat dient als bestemming voor de data die uit A&C gehaald gaan worden. Ook is er een script aangemaakt voor de staging area. De tabellen die gebruikt zullen worden worden vanuit de Great Outdoor en A&C database gehaald en hierin als het ware gekopiëerd. In onderstaande afbeeldingen zijn de verschillende data flow taken te zien die hiervoor worden uitgevoerd. Allereerst is er een execute SQL statement, die ervoor zorgt dat de staging area wordt geleegd voordat deze weer wordt gevuld. Zo vermijden we dubbele data. Elke data flow task bevat een source object, en een destination object. In een aantal gevallen vinden er ook data conversies plaats, namelijk van Unicode string (WSTR) naar een reguliere string (STR). Dit is een botsing die vaker voorkomt en tot nu toe ook de enige. De mapping spreekt voor zich, de attributen hebben nog steeds dezelfde naam in de staging area kopie.  
De staging database heeft geen foreign key constraints, zo voorkomen we foutmeldingen bij het leegmaken en vullen van de database. De referentiele integriteit wordt hersteld bij het omzetten van de data van de staging naar de uiteindelijke bestemming database.

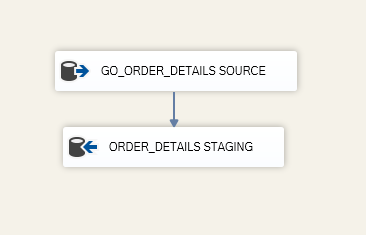


Hierboven ons overzicht van de dataflows die de staging DB opmaken.

Hieronder is het overzicht te zien van de dataflows die de DWH maken. 

Voorbeeld van een data flow task met en zonder data conversie:





**ETL**

Om vanuit de staging database de dimensies en feiten in te vullen gebruiken we meerdere data flow tasks. Echter voor één tabel, DimTijd, worden de gegevens niet opgehaald uit de staging database. DimTijd moet alle data bevatten waarin de data van promoties, targets, orders, etc. Kunnen vallen. Daarom wordt DimTijd handmatig met een script gevuld. Hiervoor maken we een stored procedure. De code die we voor de stored procedure gebruiken volgt hieronder.





Hierboven staat een embedded object met alle dataflows in ETL dataflow ontwerp, hieronder zullen wij nog enkele speciale gevallen toelichten.

**FactOrder**

‘not available’ als bestelmethode zetten voor Orders uit AC: INSERT INTO DimBestelmethode (bestelmethodecode, bestelmethodenaam) VALUES (-999, 'not available').   
Dit moet ingevuld worden, omdat de orders die gemaakt werden bij AenC, geen bestelmethode aangaven. We hebben eerder besloten om bestelmethode onderdeel van de primary key van FactOrder te maken, dus moet er wel een waarde zijn voor elke order. Deze query voegt een bestelmethode toe, genaamd ‘not available’ voor alle A&C orders.

**FactCursus**

Doordat alle informatie over de cursussen in een excel document stond moesten we een excel source gebruiken voor deze tabel. Door de afwezigheid van speciale tekens als de é en andere letters met eventuele puntjes of streepjes erop in het excel document moesten wij bij het overzetten van de gegevens bij de lookup van de werknemerkey de optie “Redirect rows to no match output” aanvinken, in plaats van de optie “Fail component” die wij hiervoor altijd gebruikten. Hierdoor kregen we voor de matches die wel succesvol zijn deze wel in de tabel. Om ook de namen met speciale tekens te matchen zouden deze handmatig moeten worden aangepast in de Excel gegevensbron.

### 5.3 Test/waarborginggegevenskwaliteit

Het is zeer belangrijk om eerst het ontwerp van de ETL te testen. Want het verkeerd maken van het ETL-proces kan leiden tot een corrupte datawarehouse. Door eerst ETL-schema's te maken wordt er voor gezorgd dat het ETL-proces succes uitgevoerd zal worden. Hiervoor gebruikt men ETL-schema's, een ETL-schema is een tabel waarin de relatie tussen de gebruikte brontabellen en een doeltabel uiteengezet worden.

Omdat ETL's moeilijk te testen zijn, zijn er softwarepakketten beschikbaar om tests uit te voeren. Hieronder de belangrijkste ETL-tests:

**Metadata testing**

* Data type check; hier worden de tabellen en column data types definities getest of zij gelijk zijn aan de data model design.
* Data lenght check; hier wordt het lengte van de database column getest of zij gelijk zijn aan het data model design.

**Data completeness**

* Record count validation; hier worden de source tabel en de eindbestemming tabel vergeleken of zij hetzelfde record count hebben.
* Compare entire source and target data; hier worden de source en de eindbestemming volledig vergeleken of zij de data bevatten die overgezet moeten worden. **Data quality testing**
* Duplicate data check; hier wordt getest op dubbele rijen met dezelfde primairy key.

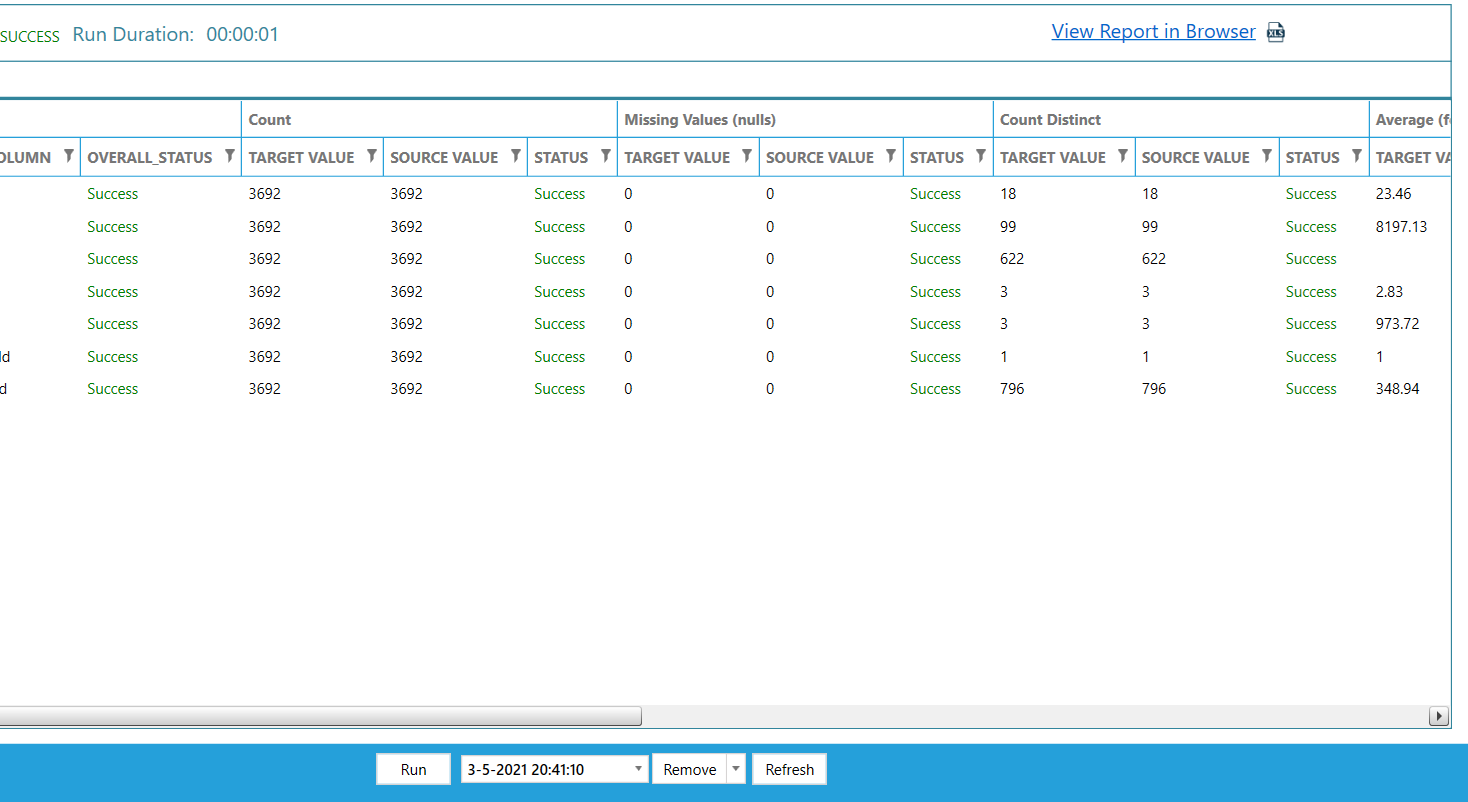
**ETL intergration testing**

* End to end data testing; is het handmatig reviewen van het source en target databases.

Dit zijn enkele technieken van het ETL datawarehouse testing. Voor het testen van onze data en ETL zal er gebruikt gemaakt worden van de ETL validator van DATAGAPS. DATAGAPS is een organisatie die een softwarepakket heeft ontwikkeld die het testen van databases en ETL-flow's makkelijker heeft gemaakt. Op de website is het mogelijk om aangemeld te worden en een 30 dagen trial aan te vragen om het software pakket uit te proberen. Er zullen een aantal van de hierboven genoemde tests hieronder uiteengezet worden.

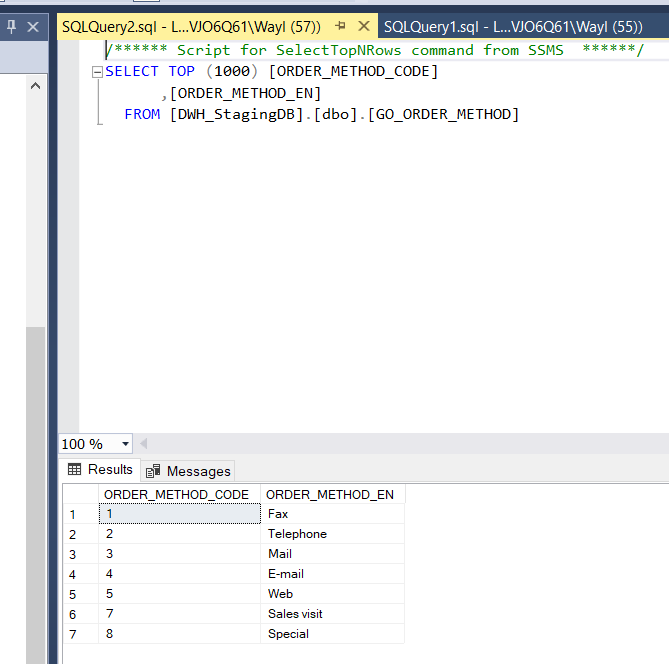
Voor dit project zal er een record count validation test gedaan worden en een end to end data test. Dit om ervoor te zorgen dat alle data die in de staging en in de datawarehouse goed erin zitten. Want dit programma is niet makkelijk in gebruik waardoor er maar één test is opgesteld. Het installeren van ETL validator vergt niet alleen een installatie wizard volgen, er zullen een aantal in totaal drie pakketten ernaast via mscmd geïnstalleerd worden. Allereerst moet de postgresql uitgepakt worden en uitgevoerd worden in cmd, daarna de pgbouncer en als laatst moet de tomcat server uitgepakt worden in de server map. Deze taken allemaal waren trial and error en hebben veel tijd in beslag genomen.

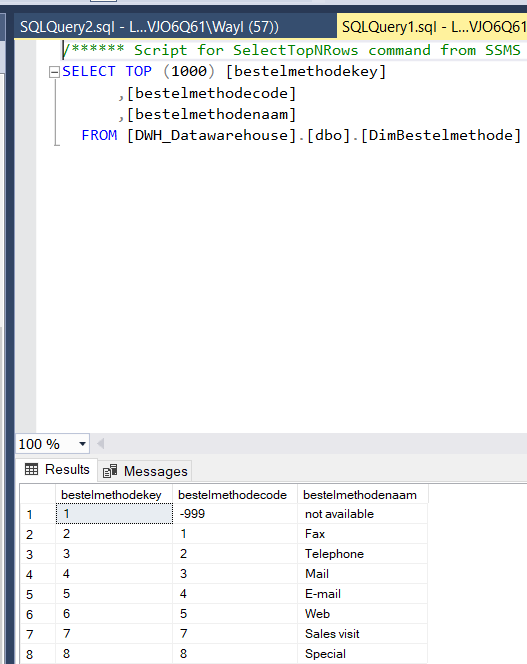
**Record count validation test**



Hierboven is er een figuur die aangeeft wat de target value is. De targetvalue van de source en staging database is 3692 en als resultaat is er te zien dat er geen gemiste values in de database zijn, dat is te zien aan de waarde 0. Dus de stagingdatabase is één op één overgenomen van de source database. Dus dit geeft aan dat de staging database goed is gevuld met de database van greatoutdoor.

**End to end data test**





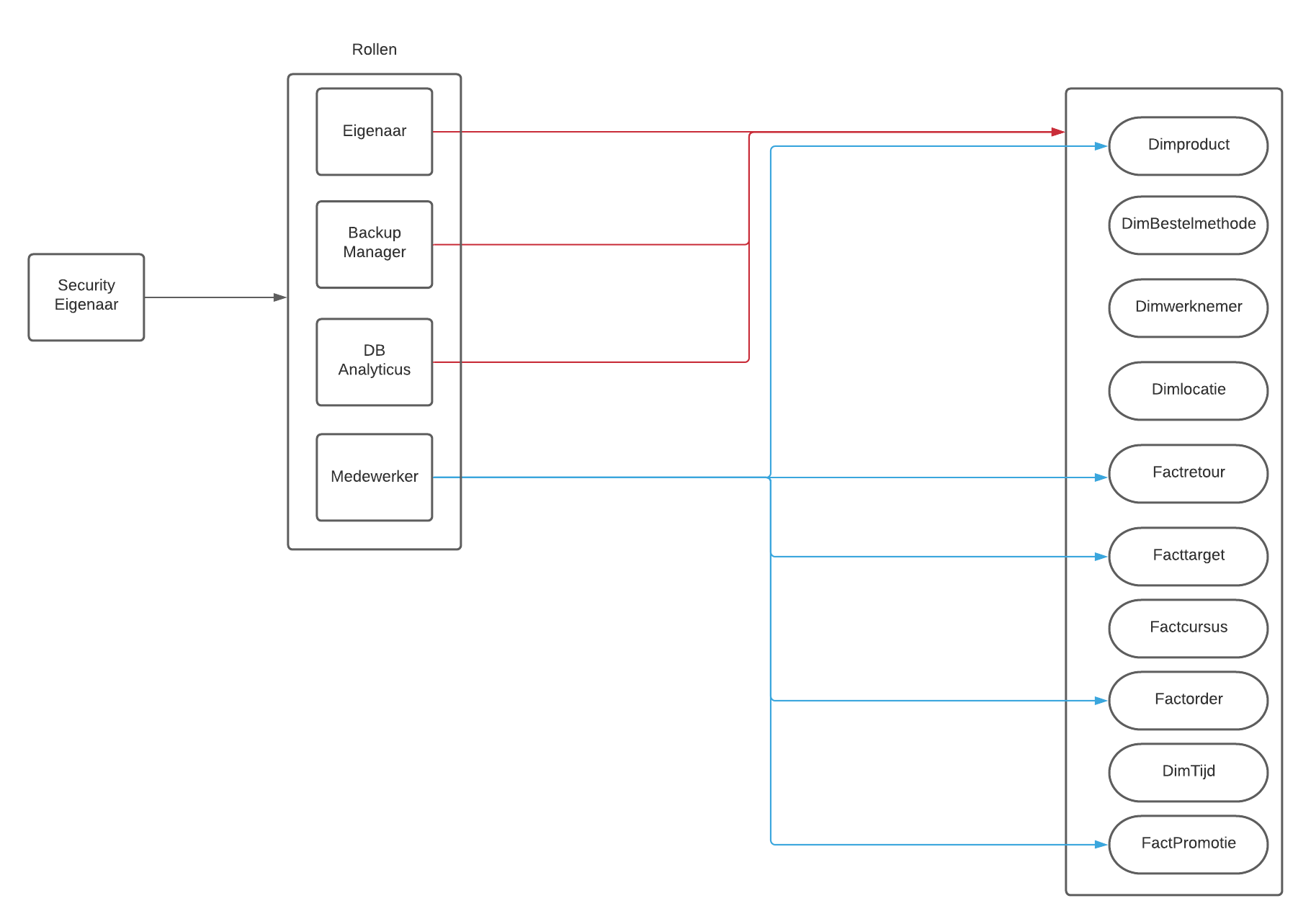
Hierboven zijn er twee figuren één figuur over de tabel order methode code en de ander over dimbestelmethode. In de ETL’s is dimbestelmethode gevuld met data uit order\_methode\_code. Door zelf de top 1000 queries van beide tabellen uit te voeren is het mogelijk om de records te vergelijken en zo te checken of de data transformatie goed is gelukt. Dit is natuurlijk geen best practice voor een grote datawarehouse, maar is prima voor een éénmanszaak.

### 5.4 Beveiliging

Denk eraan dat de securitymaatregelen ook leiden tot aanpassingen in het ontwerp en de datawarehouse. Neem dus hieronder ook de aangepaste ontwerpen en scripts op voor de volledigheid.

Beveiliging heeft invloed op de hele applicatieontwikkeling en ook op het ontwerp van belangrijke componenten van het DWH. Denk bijvoorbeeld aan de load manager, warehouse manager en query manager. Omdat het programma langs allemaal scripts moet voordat je bij een gedeelte kan komen. Dit kost over het algemeen veel tijd. Programmeurs moeten hier een efficiënte manier voor zoeken zodat het bijna geen effect heeft op het development van de app.

We moeten eerst gaan classificeren welke gebruikers er zijn in deze database. Daarna kunnen we gaan beginnen met waar elke gebruiker toegang zou moeten hebben. Dan kunnen we rollen aan gaan maken en aanwijzen waar welke rol toegang tot zou moeten hebben. De beveiliging wordt geregeld in Power BI.

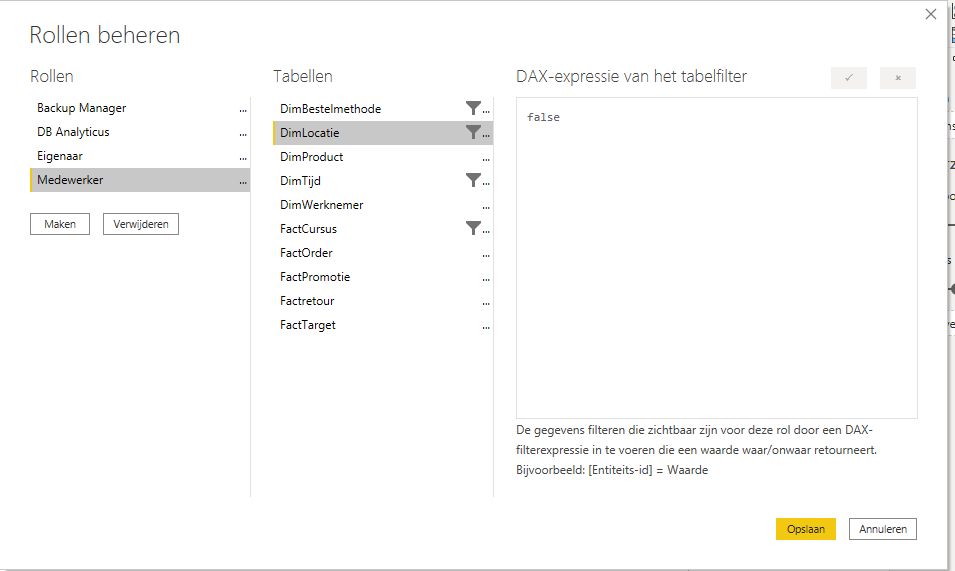


Gemaakt aan de hand van de Microsoft docs: (Database-Level Roles - SQL Server, 2020).

Bij beveiliging moet je denken aan login security(verbinden met de server), database security(toegang krijgen tot de database) en Database objecten(toegang tot individuele objecten en data) (Harkins, 2004). Wij gaan ons vooral bezig houden met security omtrent database objecten. Dit gaan we doen d.m.v. Power BI zoals is getoond in het hoorcollege.

We wilde proberen de rollen aan te geven via de Power BI workspace maar kwamen er al gauw achter dat wanneer je hiermee wilt werken, je de pro versie nodig hebt en een Azure abonnement nodig hebt. Wanneer we in een professionelere omgeving zitten waarin dit ook geregeld is via het bedrijf zouden we hier zeker gebruik van maken. Maar met ons budget en onze tools lijkt het ons beter om dit achter ons te laten en gewoon intern de rollen aan te geven.

In de onderstaande afbeelding vier soorten gebruikers die goed van pas komen voor een bedrijf. We hebben in Power BI false ingevoerd waar de gebruiker geen toegang tot mag hebben. De rest is bij default true.



Verder zou ik mijn teamgenoten toe kunnen voegen, om ze daarna een rol te geven zodat ze overal toegang hebben.   
Zo krijgen medewerkers die inloggen alleen gegevens te zien waar ze toegang tot hebben, zo niet worden de gegevens verborgen. Op deze manier kunnen we de security waarborgen.

We zouden nog rollen aan kunnen maken op de database maar aangezien onze database lokaal runt denk ik niet dat het nodig is. Er wordt standaard een DBO(Database owner) aangemaakt en een user.

# Overzicht van de Dashboards

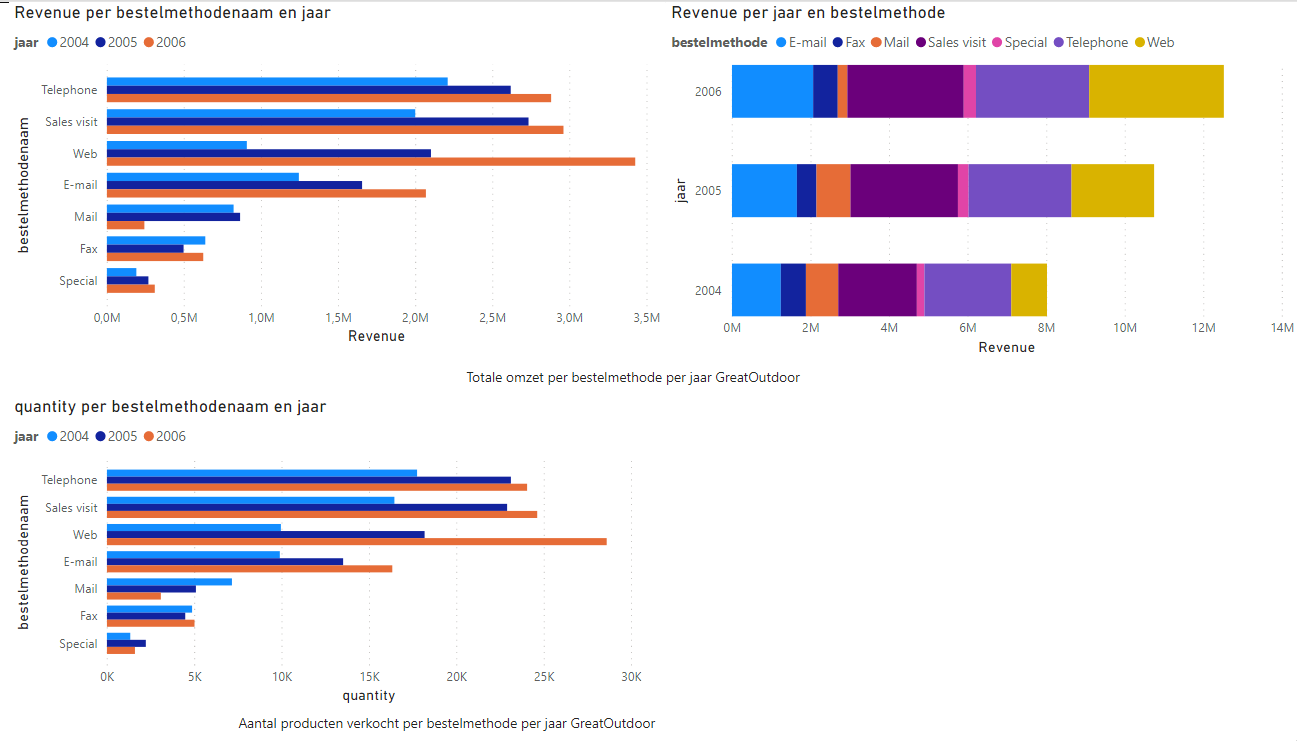
## Managementvraag 1

Managementvraag 1 luidt: “Hoeveel omzet is er in een jaar behaald, per product en per bestelmethode en wat waren aantal producten die hierbij zijn verkocht?” Om deze vraag te beantwoorden kunnen alleen gegevens van GreatOutdoor worden gebruikt. De data ontvangen van A&C bevat namelijk geen data over de bestelmethode. Om toch ook een beeld te krijgen van de omzet van de producten van A&C wordt er wel de omzet per jaar per product uitgezet.  
Om deze vraag te beantwoorden met zinvolle statistieken hebben we de managementvraag opgedeeld in een aantal onderdelen.

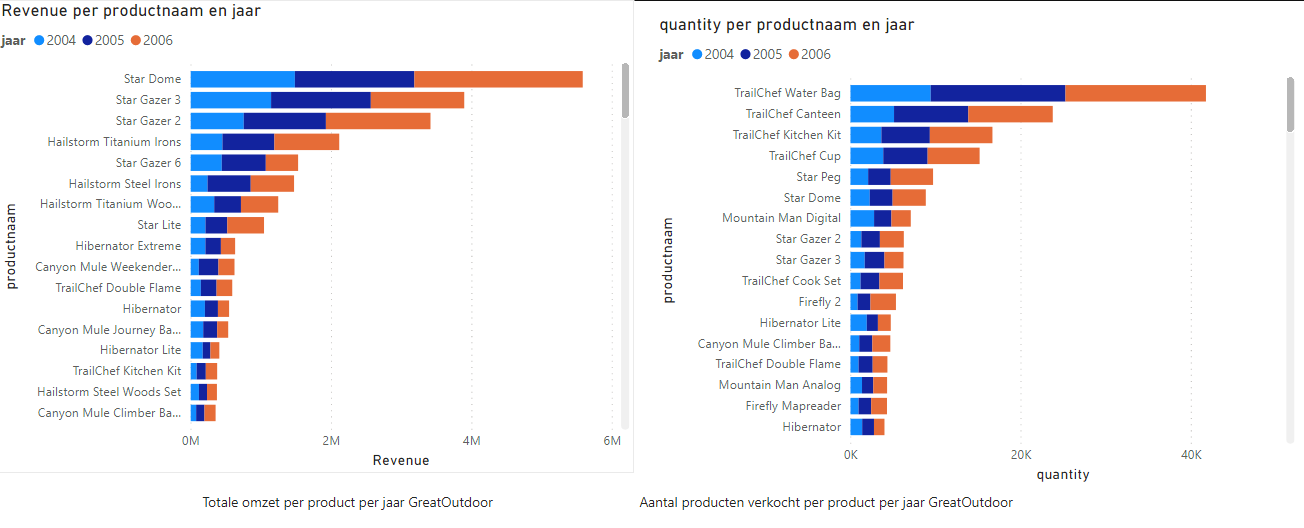
* Hoeveel omzet is er in een jaar behaald per product en per bestelmethode?
* Hoeveel producten zijn er verkocht in een jaar per product en per bestelmethode?

Deze twee onderdelen worden geanalyseerd om vervolgens van de populairste/opbrengende producten per bestelmethode en per jaar te kijken naar de omzet, en de managementvraag zo te beantwoorden.

GreatOutdoor:

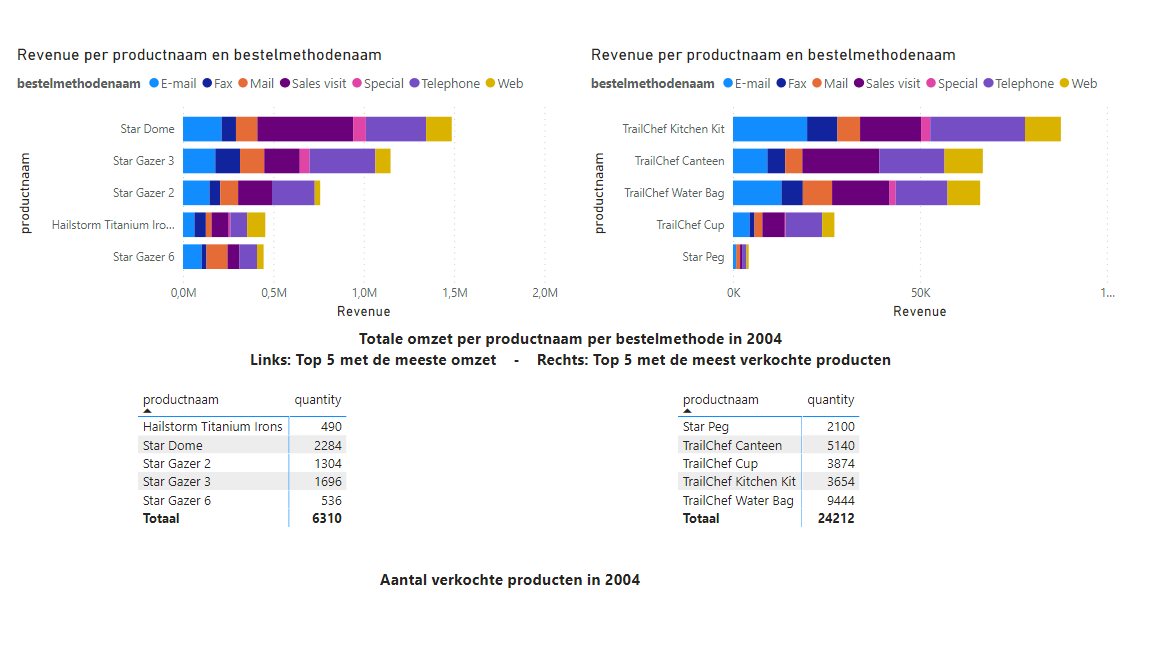


In de bovenstaande tabellen zijn twee weergaven gemaakt voor de totale omzet per bestelmethode. Zo is er te zien dat bijvoorbeeld de bestelmethode ‘Web’ met het jaar populairder wordt. Bestelmethode ‘Mail’ wordt steeds minder gebruikt.   
De onderste tabel geeft het aantal verkochte producten per bestelmethode per jaar aan. Hier wordt dus geen rekening gehouden met de prijzen.

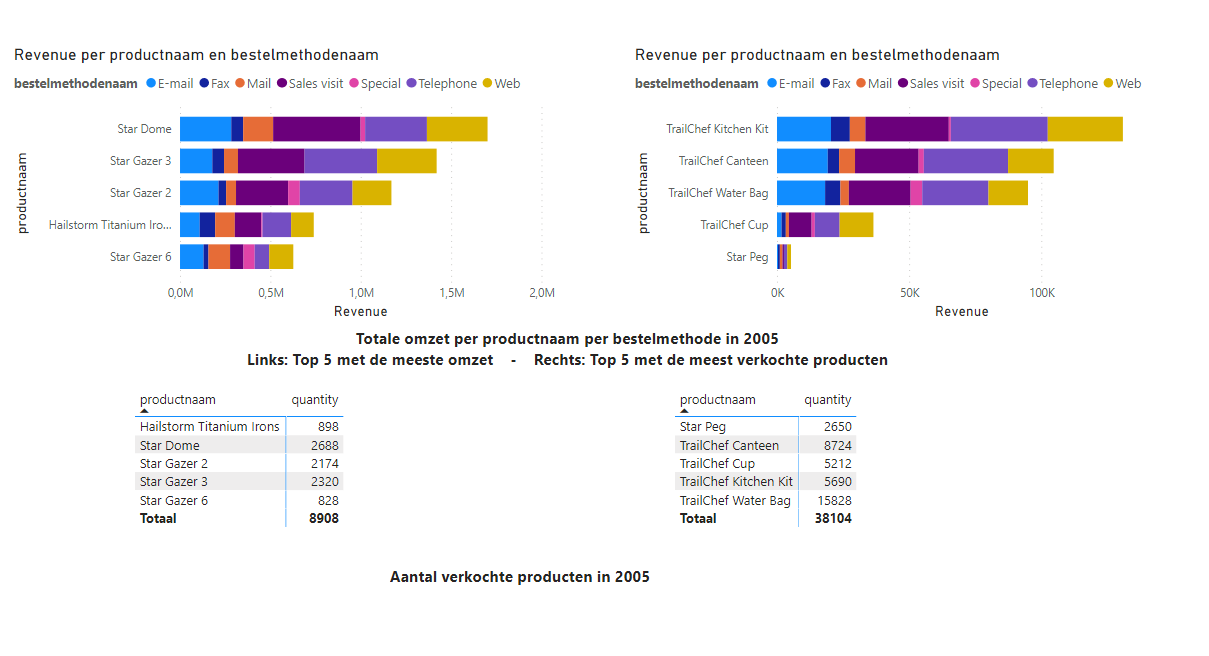


In deze twee grafieken staan links en rechts respectievelijk de totale omzet en het totale aantal producten per product per jaar. Beiden grafieken laten eigenlijk een heel verschillende lijst zien. We kiezen er daarvoor om beiden te combineren en de top 5 van zowel producten met de meeste omzet als meest verkochte producten mee te nemen naar de laatste grafieken. Deze tien grafieken worden geanalyseerd op bestelmethode en product.

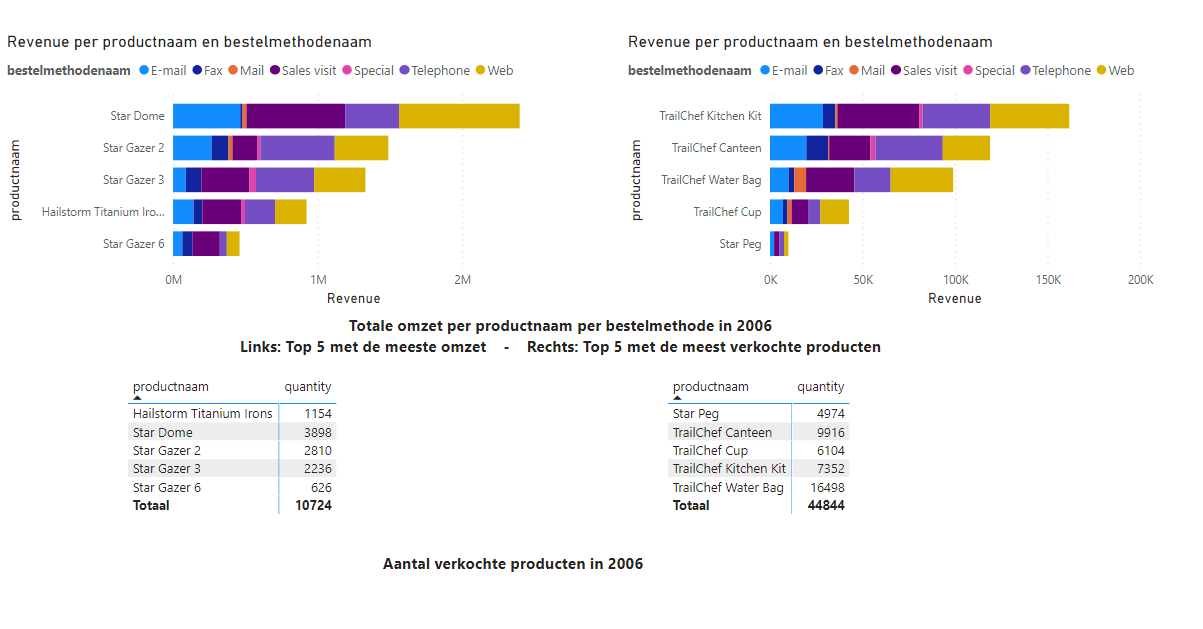
2004:



2005:



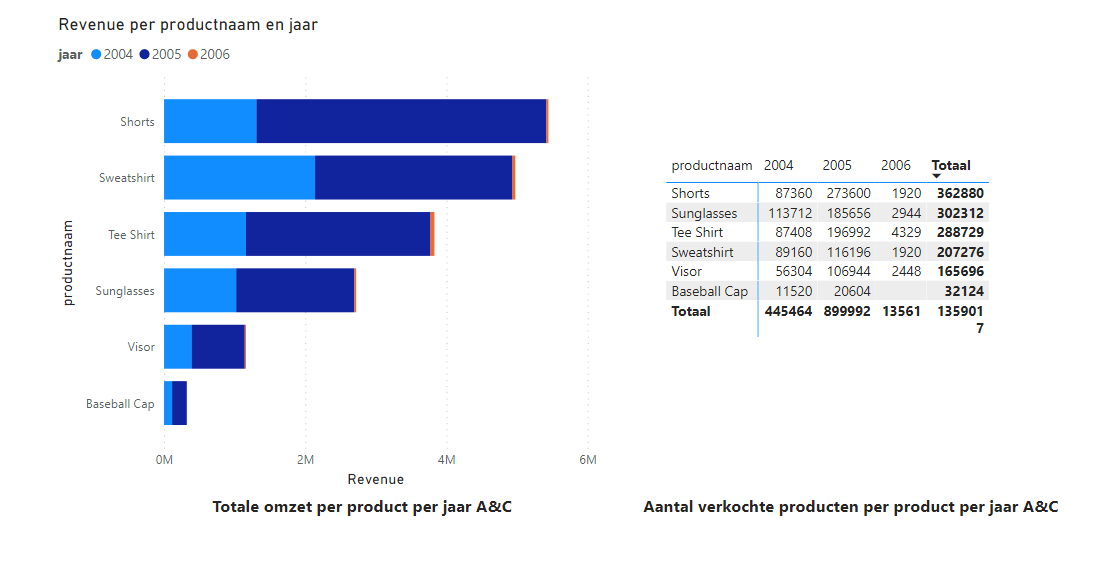
2006:



In de bovenstaande drie afbeeldingen wordt per jaar de totale omzet per product per bestelmethode uitgezet. Onder elke grafiek is een matrix tabel gemaakt waarin van de desbetreffende producten het aantal dat dat jaar verkocht is wordt weergegeven. In deze overzichten is de totale omzet per productnaam en per bestelmethode in een oogopslag te zien, i.c.m. het aantal verkochten producten wat daarbij hoort.

A&C:

Voor A&C is één overzicht gemaakt, aangezien er geen gegevens over de bestelmethode aanwezig zijn. Het aantal producten is ook klein genoeg, waardoor alle producten in een overzichtje passen.



In bovenstaand overzicht is links de totale omzet per product, per jaar te zien. Rechts is een matrix tabel gemaakt waarin per product het aantal verkochte producten per jaar te zien is. Zo kunnen deze twee vergeleken worden en wordt voor A7C beantwoordt welke producten de meeste omzet per jaar hebben gemaakt en wat de meest verkochte producten zijn.

Hieronder is het Power BI bestand als embedded object opgenomen.



## Managementvraag 2

Managementvraag 2 kan je opdelen in 3 vragen zodat je 3 tabellen krijgt. “Wat waren de opbrengsten en targets per kwartaal van elk product en welke employees hebben de targets wel en welke hebben de targets niet gehaald en welke cursussen hebben deze employees allemaal gevolgd?” Bestaat uit:

* Wat waren de opbrengsten en targets per kwartaal van elk product
* welke employees hebben de targets gehaald en welke niet
* en welke cursussen hebben deze employees allemaal gevolgd?

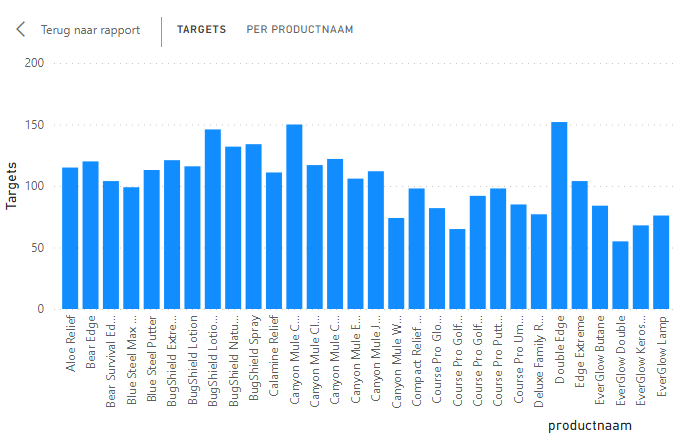
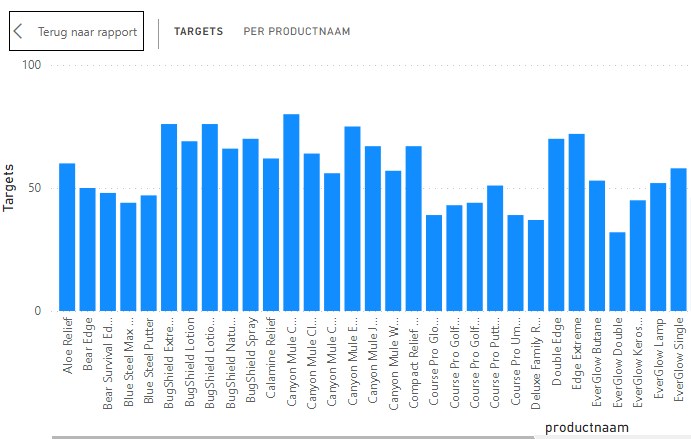
Als je al deze vragen in 1 tabel stopt krijg je een hele drukke grafiek/tabel waar je eigenlijk niks aan hebt.

Daarom hebben wij het verdeelt in 3 vragen. Als eerste vraag: “Wat waren de opbrengsten en targets per kwartaal van elk product”.

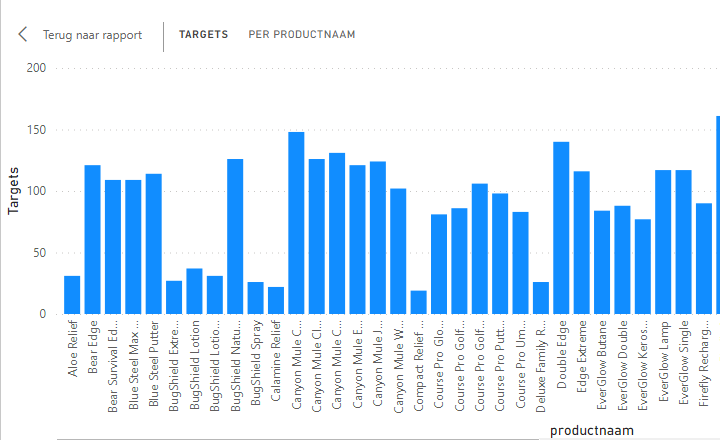
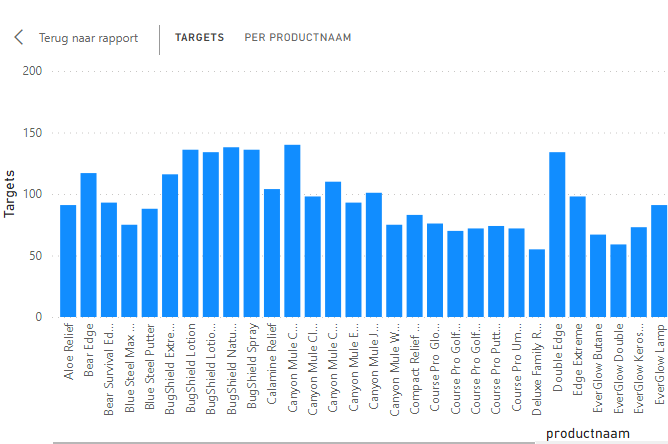


Als je deze tabel eenmaal hebt kan je heel makkelijk verkleinen naar kwartaal 1 of kwartaal 2 etc.

Kwartaal 1: Kwartaal 2:

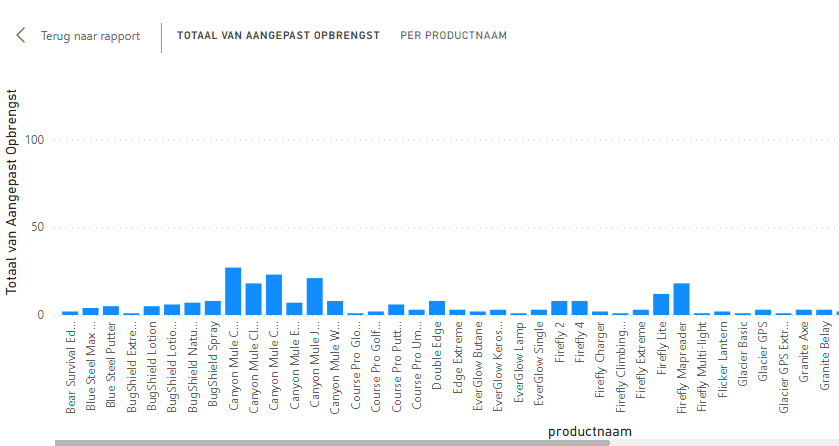
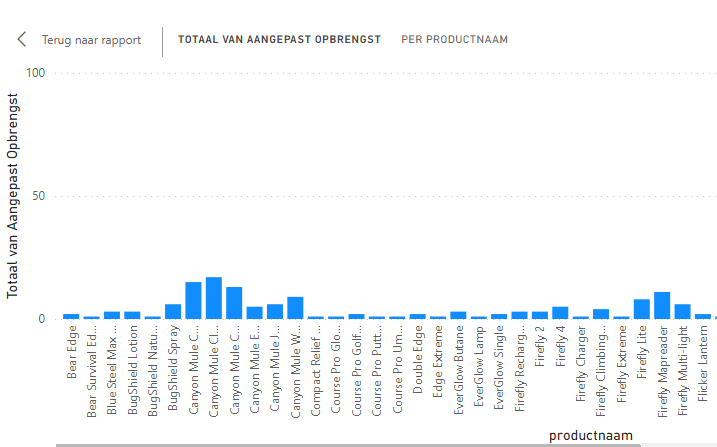


Kwartaal 3: Kwartaal 4:

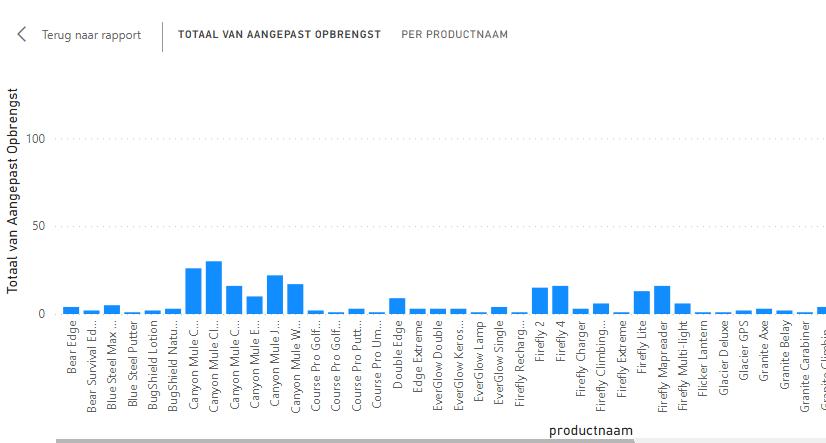
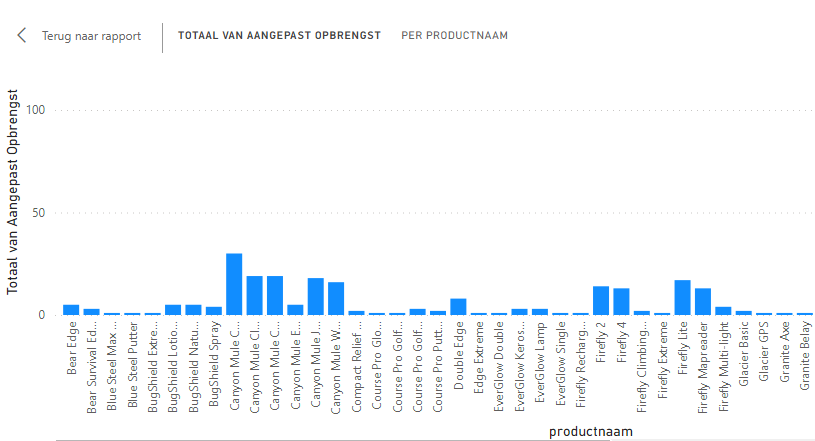


Daarna kan je deze tabel vergelijken met de opbrengsten

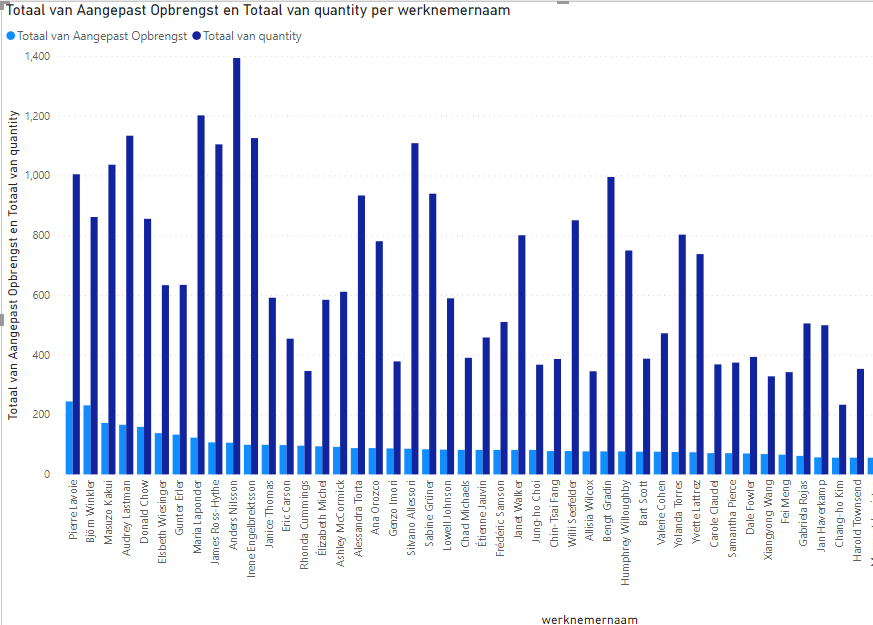
Kwartaal 1: Kwartaal 2



Kwartaal 3: Kwartaal 4:



Als tweede moeten we gaan kijken welke employees het wel en niet hebben gehaald. We hebben quantity \* price gedaan om de totale opbrengsten te krijgen en deze te vergelijken met wat ze hadden moeten behalen om het target te behalen



Volgens onze berekeningen heeft niemand van de werknemers de targets gehaald.

Aangezien deze targets ver boven de behaalde statistieken liggen heb ik de conclusie getrokken dat er iets fout is gegaan bij de berekening van power BI en de tijd erbij. Of het bedrijf moet hele slechte medewerkers hebben.

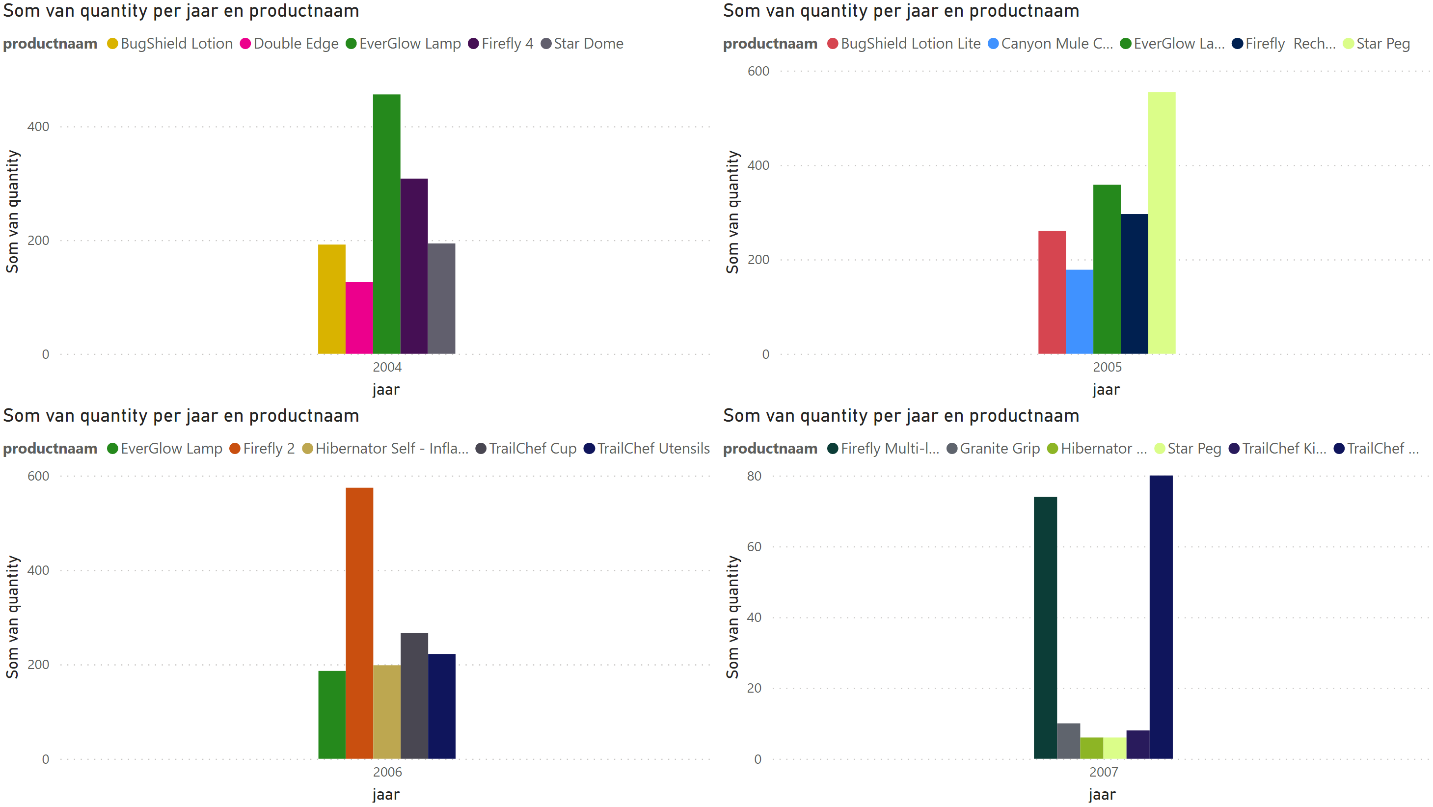


## Managementvraag 3

“Wat waren de namen van de 5 meest teruggebrachte producten per jaar, wat was het jaar waarin in totaal de meeste producten teruggebracht werden en wat waren in dit jaar de top 5 retailers met de meeste teruggebrachte producten. Wat waren bij deze 5 retailers de namen van de 5 meest teruggebrachte producten voor alle jaren bij elkaar.”

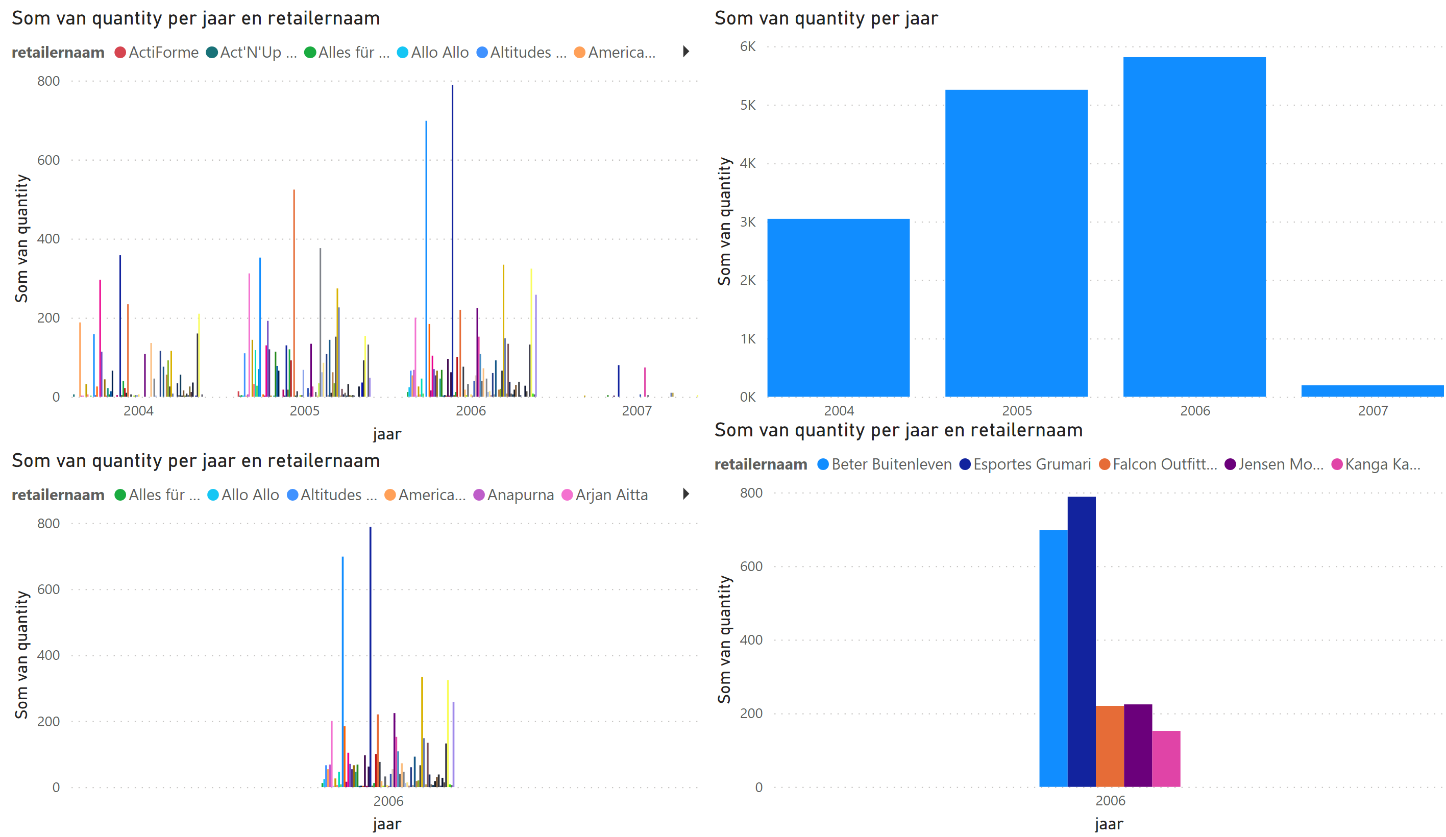
Om deze managementvraag te beantwoorden zullen we 3 grafieken nodig hebben. Grafiek 1 laat zien wat de meest teruggebrachte producten zijn per jaar, los van eventuele retailers, dit kan je gebruiken om te zien wat de producten zijn die het minst in de smaak vallen of op z'n minst het vaak teruggebracht worden. Grafiek 2 wordt gebruikt om te kijken bij welke retailers er het vaakst producten teruggebracht worden, bij deze grafiek nemen we de 5 retailers met de grootste waardes en deze retailers nemen wij mee naar grafiek 3, waar wij gaan kijken wat, per retailer, de meest teruggebrachte producten zijn om zo te zien welke producten het minst in de smaak vallen.

1. De 5 meest teruggebrachte producten per jaar



In het dashboard hierboven zijn 4 grafieken te zien, elke grafiek staat voor een jaartal en voor elk jaar zijn de 5 producten eruit gepikt die het meest teruggebracht zijn, los van retailers. Het antwoord op dit deel van de managementvraag luidt dan ook (op volgorde van meest teruggebrachte product naar minst teruggebrachte product):  
2004: Everglow Lamp, Firefly 4, Star Dome, Bugshield Lotion, Double Edge  
2005: Star Peg, EverGlow Lamp, Firefly rechargeable battery, bugshield lotion lite, canyon mule  
2006: firefly 2, trailchef cup, trailchef utensils, hibernator self, everglow lamp  
2007: trailchef utensils, firefly multi-light, granite grip, trailchef kitchen kit en een gedeelde 5e plek tussen star peg en Hibernator Lite

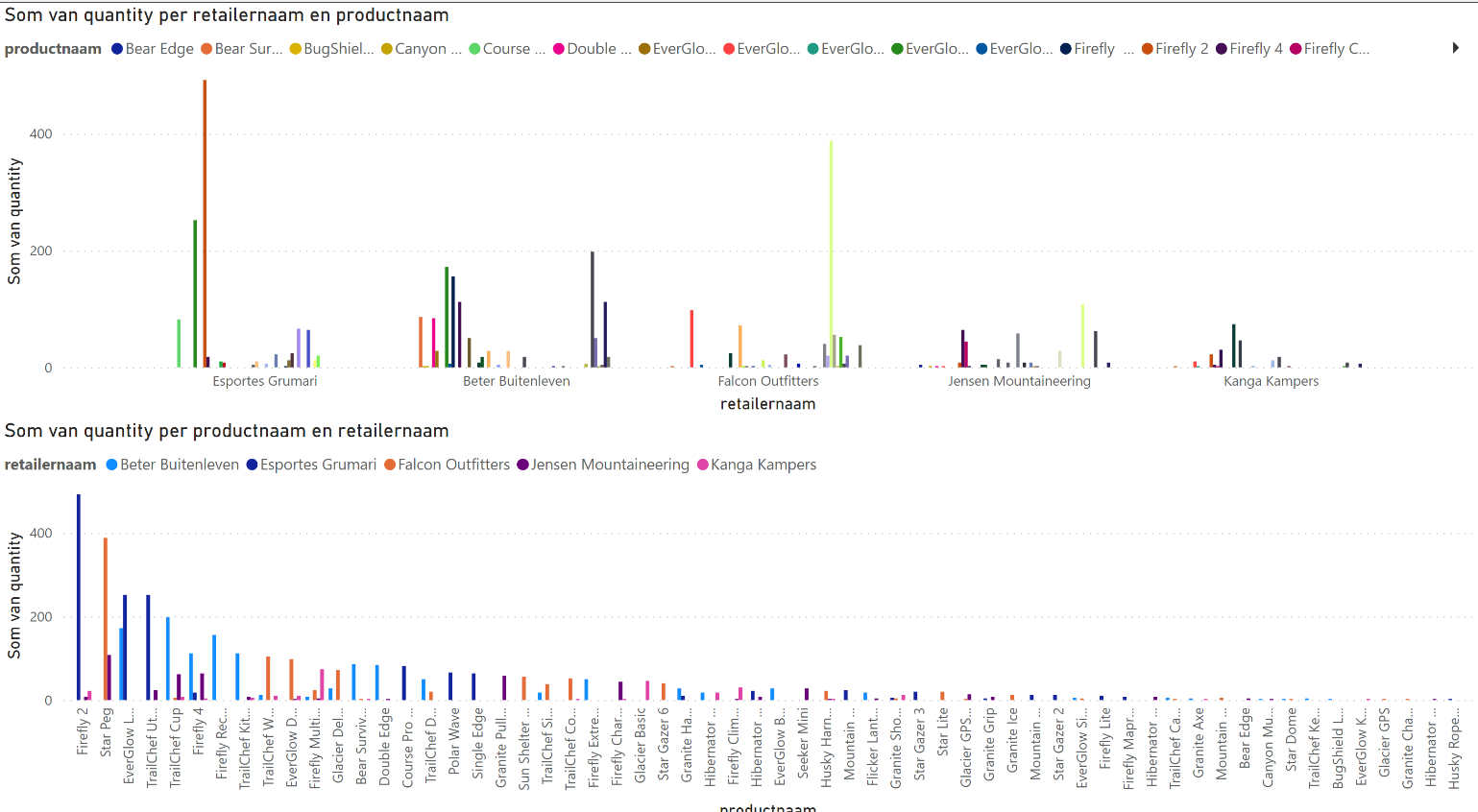
1. Top 5 retailers, gebaseerd op het aantal teruggebrachte producten

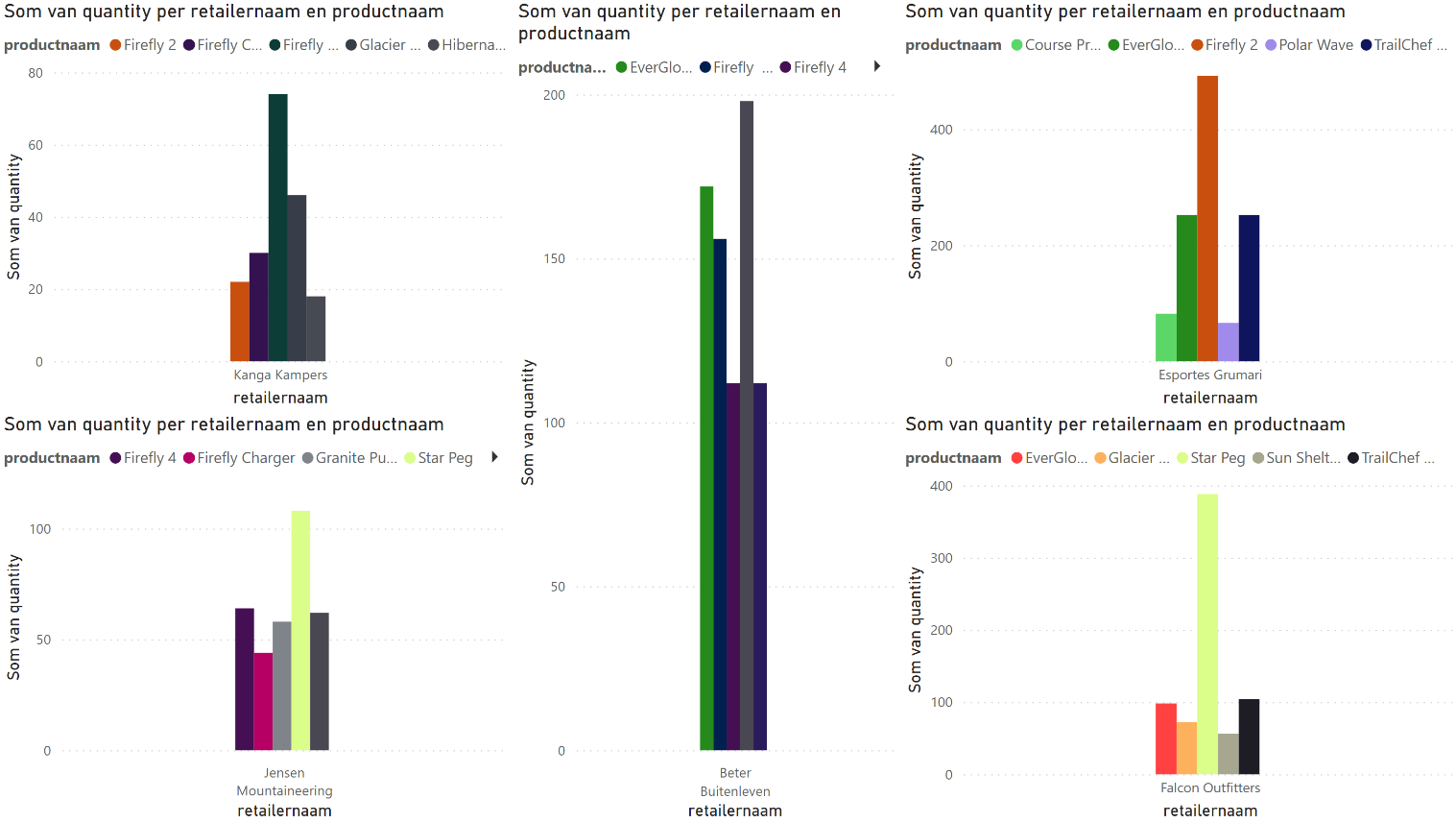


Ook dit dashboard heeft 4 grafieken, om erachter te komen wat de top 5 retailers zijn gebaseerd op het aantal teruggebrachte producten moeten we er eerst achter komen in welk jaar er de meeste producten teruggebracht werden.   
Dit zoeken we uit in de grafiek rechtsboven, hieruit is op te merken dat in 2006 er de meeste producten teruggebracht werden.   
In de grafiek linksonder is het jaar 2006 alleen genomen zodat het duidelijker is wat de uitschieters zijn qua het anatal teruggebrachte producten.   
Als laatste is er bij de grafiek rechtsonder een filter geplaatst op de retailers, waardoor je alleen de top 5 retailers ziet gebaseerd op het aantal teruggebrachte producten.

Het antwoord op dit deel van de managementvraag luidt:  
1. Esportes Grumari  
2. Beter Buitenleven  
3. Jensen Mountaineering  
4. Falcon Outfitters  
5. Kanga Kampers

1. De namen van de 5 meest teruggebrachte producten voor alle jaren bij de 5 retailers die hierboven zijn gevonden





Voor het derde deel van deze managementvraag hebben we 2 dashboards opgenomen, dit doordat er enige onduidelijkheid kan ontstaan door de vraagstelling. Het is mogelijk dat de vraag alleen gaat over het totaal van teruggebrachte producten van de 5 retailers, maar het is ook mogelijk dat de vraag slaat op het aantal teruggebrachte producten per retailer, hierom hebben we beide opties meegenomen in het dashboard.

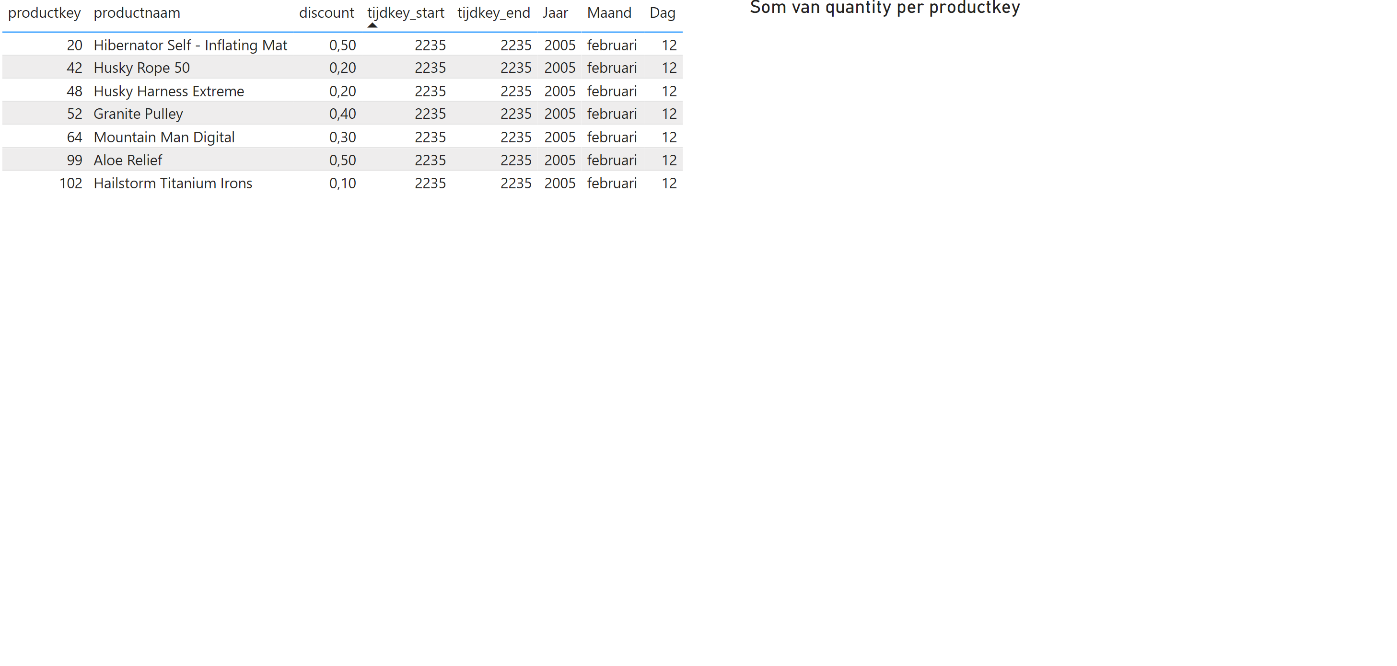
Het antwoord op dit deel van de managementvraag luidt (als je het totaal neemt van alle retailers):  
1: Firefly 2  
2: Star Peg  
3: Everglow Lamp  
4: Trailchef Utensils  
5: Trailchef Cup

## Managementvraag 4

“Hoeveel omzet is er gemaakt per promotie, wat zijn de top 5 producten tijdens iedere promotie en hoeveel producten zijn er totaal in die promotie verkocht?”

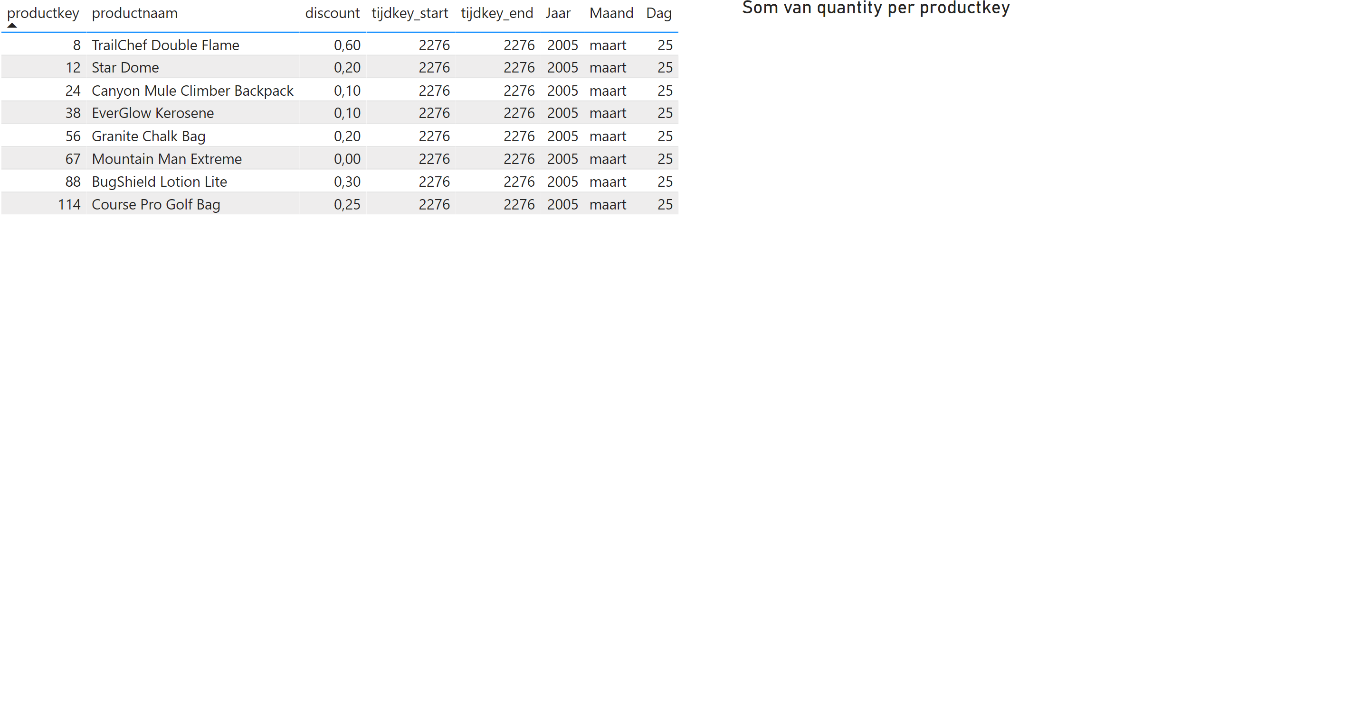
Great Outdoor heeft in het jaar 2005 6 promoties gehad, deze verschilden in tijd, van 1 dag naar meerdere maanden. Wij zullen hieronder aangeven per promotie wat de top 5 verkochte producten waren en hoeveel omzet er in die promoties gemaakt is. Ook zullen wij een totaal van verkochte producten geven (dit geldt alleen voor producten die onder de aanbieding vallen).

### Promotie 1: Commercial TV



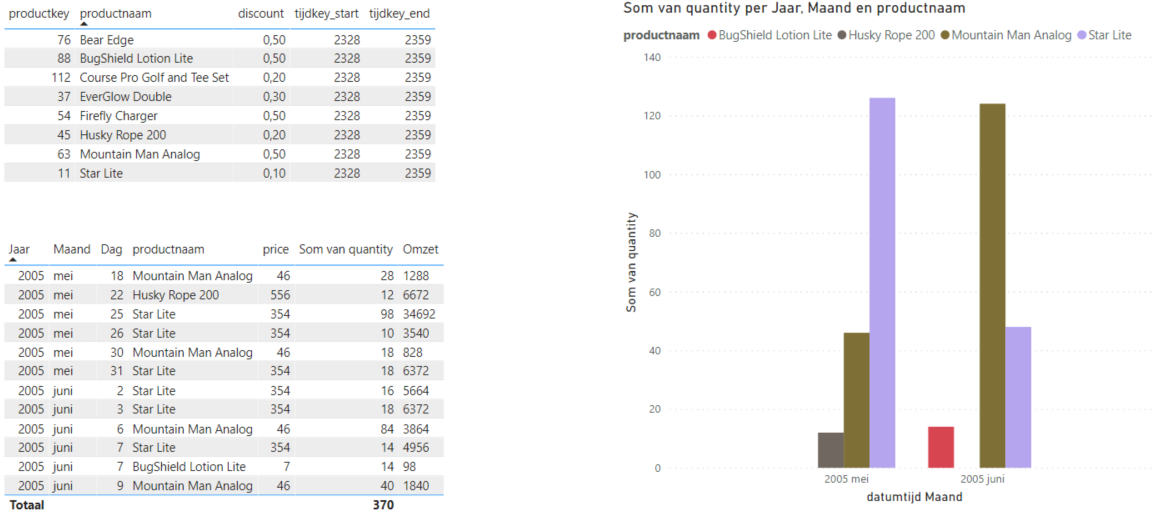
De eerste promotie van het jaar, genaamd Commercial TV, duurde maarliefst 1 dag. Tijdens deze commercial werden 6 producten aangeprezen met verschillende kortingsaantallen. Helaas zijn er deze datum geen orders geweest die de betreffende producten bevatten, dus hier kunnen wij niet uitrekenen wat de omzet geweest is en er is dus ook geen top 5 te maken van de hoeveelheid producten.

### Promotie 2: One Day Profit



Bij de tweede promotie, genaamd One Day Profit is hetzelfde gebeurd als bij de eerste promotie, ook hier was de promotietijd maar één dag en zijn er tijdens die dag geen producten verkocht die binnen de promotie vielen. Hierdoor zijn we bij deze promotie ook niet in staat om de omzet te berekenen en de top 5 op te stellen.

### Promotie 3: Summer Sale



De derde promotie is de eerste waar tijdens de promotie ook daadwerkelijk producten worden verkocht. Dit betekent dat we hier eindelijk de overzichten kunnen gaan geven die benodigd zijn om de managementvraag te beantwoorden. De informatie is in de grafieken te zien maar we zullen de informatie nog even compact hieronder beschrijven om onduidelijkheden te voorkomen.

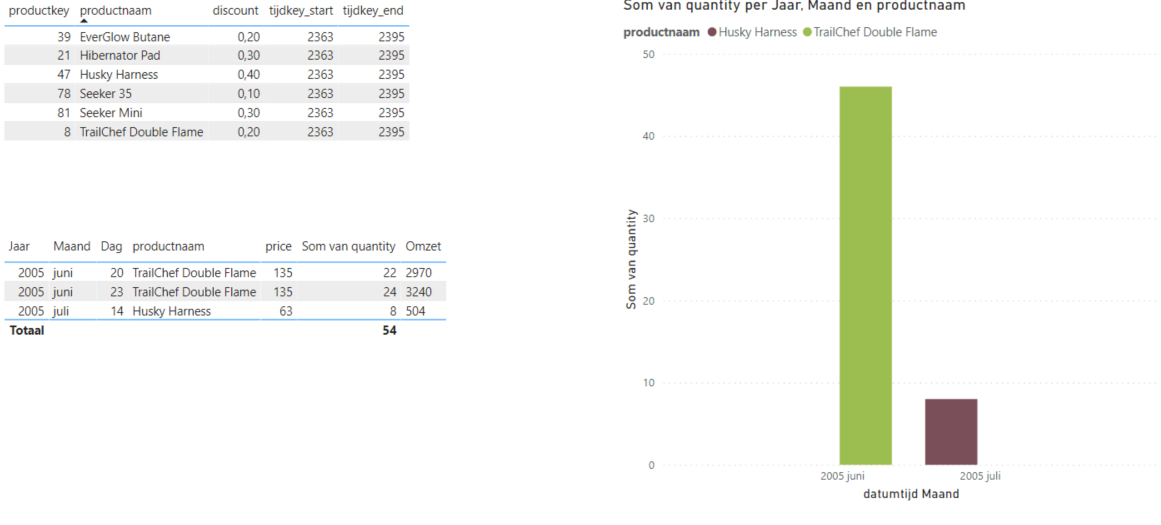
Vraag 1, Wat is de omzet van deze promotie: Deze vraag is makkelijk te beantwoorden door de prijs en de kwantiteit van de verschillende producten te multipliceren en daarbij dan ook de discount te multipliceren, dit gaat als volgt:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Naam** | **Prijs** | **Hoeveelheid** | **Korting** | **Omzet** |
| **BugShield Lotion Lite** | 7 | 14 | 50% | €49 |
| **Husky Rope 200** | 556 | 12 | 20% | €1.334,40 |
| **Mountain Man Analog** | 46 | 170 | 50% | €3.910 |
| **Star Lite** | 354 | 174 | 10% | €55.436,40 |
| **Totaal** |  | **370** |  | **€60.729,80** |

De top 5 producten, gebaseerd op omzet, zijn dan ook:  
1.Star lite  
2.Mountain man analog   
3.Husky rope 200   
4.Bugshield lotion lite

Het totaal aantal verkochte producten binnen deze promotie is 370.

### Promotie 4: Summer Promo



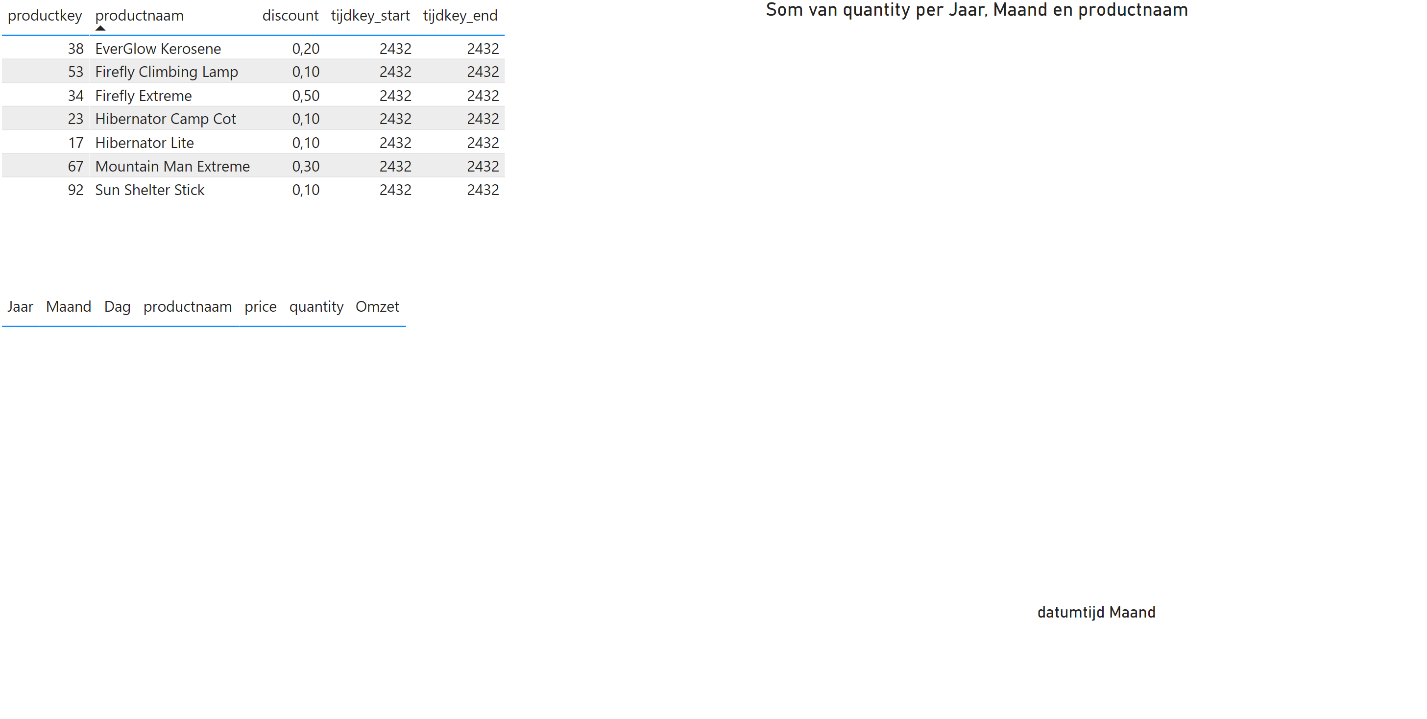
Promotie 4 was minder succesvol dan promotie 3 en dat is duidelijk te zien aan de verkoopcijfers. Zoals je op de grafieken kan zien, zijn er maar 2 soorten producten die deze periode verkocht zijn, namelijk de Husky Harness en de TrailChef Double Flame. Uit deze grafieken is ook direct een top 5 (nu een top 2) te lezen, namelijk dat TrailChef Double Flame veel beter verkocht dan de Husky Harness.

Om de behaalde omzet uit te rekenen gebruiken we dezelfde methode als bij de vorige promotie en die gaat als volgt:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Naam** | **Prijs** | **Hoeveelheid** | **Korting** | **Omzet** |
| **Husky Harness** | 63 | 8 | 40% | €302,40 |
| **TrailChef Double Flame** | 135 | 46 | 20% | €4.968 |
| **Totaal** |  | **54** |  | **€5.270,40** |

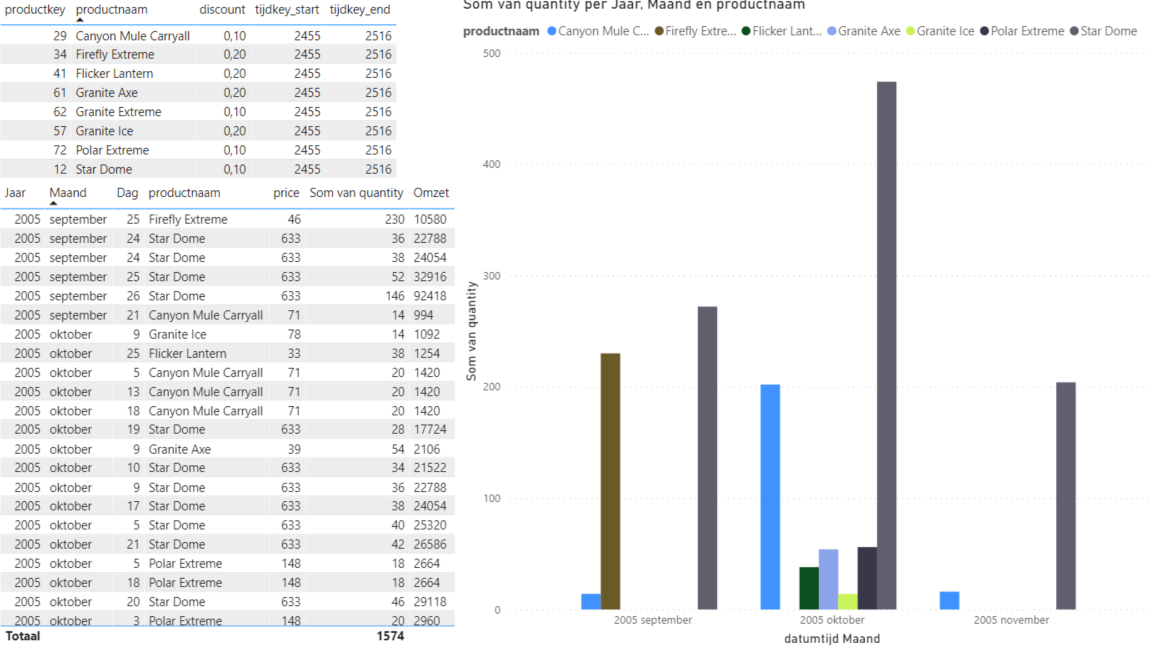
Het totale aantal verkochte producten binnen deze promotie is 54.

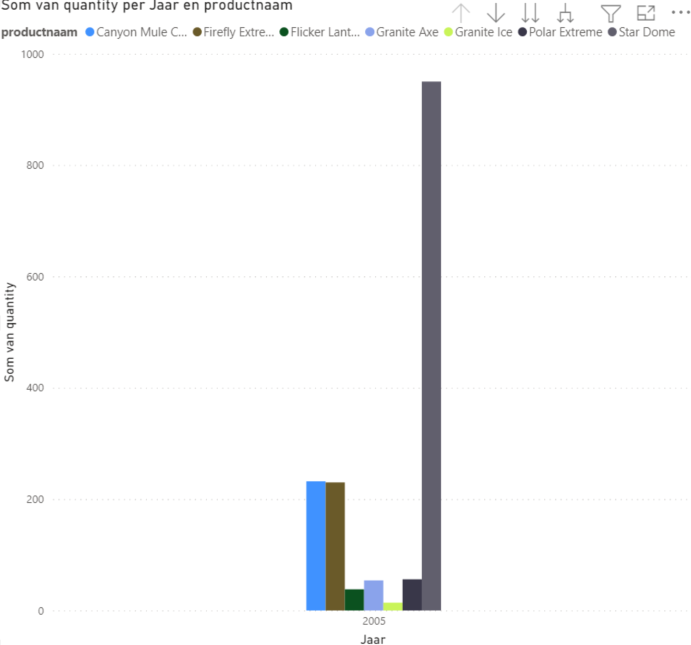
### Promotie 5: One Day Profit



Deze promotie had wederom geen verkochte producten tijdens de promotietijd, hierdoor kunnen wij weer geen duidelijk overzicht maken van de meest verkochte producten.

### Promotie 6: Indian Summer Sale





De indian summer sale was met afstand de meest succesvolle promotie. Tijdens deze 2 maanden durende promotie zijn er maarliefst **1574 producten verkocht**. De totale omzet gaan wij hieronder uitrekenen op dezelfde manier die we gebruikten bij de vorige promoties.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Naam** | **Prijs** | **Hoeveelheid** | **Korting** | **Omzet** |
| **Canyon Mule Carryon** | 71 | 232 | 10% | €14.824,80 |
| **Firefly Extreme** | 46 | 230 | 20% | €8.464 |
| **Flicker Lantern** | 33 | 38 | 20% | €1.003,20 |
| **Granite Axe** | 39 | 54 | 20% | €1.684,70 |
| **Granite Ice** | 78 | 14 | 20% | €873,60 |
| **Polar Extreme** | 148 | 56 | 10% | €7.459,20 |
| **Star Dome** | 633 | 950 | 10% | €541.215 |
| **Totaal** |  | **1574** |  | **€575.524,50** |



Hierboven een embedded object met de grafieken voor managementvraag 3 en 4