BEN GROENENDIJK

Aan de slag met Power BI

VAN DATA NAAR DASHBOARD

Boom

+ WEBSITE

Aan de slag met Power BI

Aan de slag met Power BI

Van data naar dashboard

Ben Groenendijk

Boom

inclusief website!

Met behulp van onderstaande unieke activeringscode krijg je toegang tot de website www.aandeslagmetpowerbi.nl voor extra materiaal. Deze code is persoonsgebonden en gekoppeld aan de eerste druk. Na activering van de code is de website twee jaar toegankelijk. De code kan tot zes maanden na het verschijnen van een volgende druk geactiveerd worden. De code is eenmalig te gebruiken.

Opmaak binnenwerk: Holland Graphics, Amsterdam
Basisontwerp omslag: Dog & Pony, Amsterdam
Omslagontwerp: Coco Bookmedia, Amersfoort
Beeld omslag: iStock

© Ben Groenendijk & Boom uitgevers Amsterdam, 2020

Behoudens de in of krachtens de Auteurswet gestelde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van reprografische verveelvoudigingen uit deze uitgave is toegestaan op grond van artikel 16h Auteurswet dient men de daarvoor wettelijk verschuldigde vergoedingen te voldoen aan de Stichting Reprorecht (Postbus 3051, 2130 KB Hoofddorp, www.reprorecht.nl). Voor het overnemen van (een) gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (art. 16 Auteurswet) kan men zich wenden tot de Stichting PRO (Stichting Publicatie- en Reproductierechten Organisatie, Postbus 3060, 2130 KB Hoofddorp, www.stichting-pro.nl).

No part of this book may be reproduced in any form, by print, photoprint, microfilm or any other means without written permission from the publisher.

ISBN 9789024428540
ISBN 9789024428557 (e-book)
NUR 123

www.aandeslagmetpowerbi.nl
www.boomhogeronderwijs.nl

Voorwoord

Informatie over Power BI is zeer gefragmenteerd te vinden op internet. Boeken over Power BI zijn erg gedetailleerd en gaan in op alle mogelijkheden die het programma kent. Voor het leren werken met Power BI zijn beide ongeschikt, er is geen logische opbouw. Je kunt Power BI alleen leren door zelf aan de slag te gaan met de app, hands-on. Dit boek behandelt alleen de essentiële onderwerpen die voor de (toekomstige) professional van belang zijn. Naast een inleidend hoofdstuk waarin de belangrijkste terminologie binnen datawarehousing wordt toegelicht, is het een hands-on boek.
De lezer gaat stap voor stap zelf aan de slag met Power BI, inclusief video-ondersteuning bij alle onderwerpen. Het is geschikt voor diegene die Power BI wil gaan toepassen voor studie of werk. Diegenen die Power BI al gebruiken leren het beter te doorgronden en professionele rapporten te bouwen.

Het is niet noodzakelijk al voorkennis te hebben over databases. In het eerste inleidende hoofdstuk worden begrippen zoals databases, relaties, datamodel, data-analytics, datawarehouses, ETL-proces, kubussen, big data, etc. toegelicht. Niet alleen het bouwen van visuals, presentatie van data, wordt toegelicht, maar ook het aanpassen en laden van data in het datawarehouse.

Alle onderwerpen worden met concrete voorbeelden stap voor stap toegelicht, zijn voorzien van relevante schermafdrukken en video-ondersteuning. De video-ondersteuning is een noodzakelijk onderdeel van het boek (blended-learning). Bij het boek worden vele bestanden geleverd om te oefenen met Power BI. Sommige met miljoenen rijen aan data, zodat de kracht van Power BI getoond kan worden. De bestanden en video's zijn alleen beschikbaar via de website www.aandeslagmetpowerbi.nl met een unieke code die in dit boek is afgedrukt. Op die website zijn ook de uitwerkingen van de voorbeelden te downloaden.

Het boek is geschikt voor de Nederlandstalige of Engelstalige versie van Power BI Desktop. In het boek is uitgegaan van de Nederlandstalige versie, maar steeds is ook het Engelstalige equivalent erbij gegeven.

Na het lezen en uitvoeren van de voorbeelden beschik je over alle kennis en tools binnen Power BI die noodzakelijk zijn om professionele rapporten en dashboards te bouwen.

Vragen, opmerkingen of aanvragen voor trainingen naar aanleiding van dit boek zijn welkom. Stuur deze aan ben.groenendijk@outlook.com.

April 2020
Ben Groenendijk

Inhoud

Voorwoord 5

Inleiding 9
Opbouw van het boek 9
Website www.aandeslagmetpowerbi.nl 10

1 Data modeling 11
1.1 Business Intelligence 11
1.2 Data, de bron van BI 11
1.3 Data Analytics 12
1.4 Databases en tabellen 13
1.5 Relaties tussen tabellen in een database 16
1.6 Datawarehouse 18
1.7 Kubussen 20
1.8 Tabellen en relaties in een datawarehouse 23
1.9 Big Data 28

2 Power BI en visuals 33
2.1 Power BI 33
2.2 Visuals 35
2.2.1 *Visuals met één tabel* 35
2.2.2 *Visuals met meerdere tabellen* 44
2.2.3 *Visuals met berekende kolommen en metingen* 47
2.3 Publiceren 67

3 Extract, Transform en Load proces 71

4 Kardinaliteit en filterrichting 85
4.1 Een-op-een relaties 85
4.2 Een-op-veel relaties 89
4.3 Kruisfilterrichting bij relaties 94
4.4 Veel-op-veel relaties 101
4.4.1 *Een-veel-veel-een relaties* 101
4.4.2 *Veel-een-een-veel relaties* 102

5 De meest gebruikte DAX-functies 107
5.1 De dimensietabel Datum 107
5.2 Filteren met DAX 118
5.3 Berekeningen in DAX met de functie *CALCULATE()* 120
5.3.1 *CALCULATE() en annuleren filtercontext* 121
5.3.2 *CALCULATE() en wijzigen filtercontext* 124
5.3.3 *CALCULATE() en creëren filtercontext* 131
5.3.4 *CALCULATE() en rijcontext* 132
5.4 Tijdintelligentie in DAX 134
5.5 Klassenindeling in DAX 143
5.6 Rangorde in DAX 147
5.7 Variabelen, cumulatieve waarden en Pareto 152

6 Tips, trucs en nuttige visuals 161
6.1 Metingen in één tabel plaatsen 161
6.2 Snelle metingen 163
6.3 Tijdintelligentie met toekomstige datums 165
6.4 Metervisuals 168
6.5 Infographic Designer 171

7 Professionaliseren dashboard en rapport (Reporting) 175
7.1 Bladwijzers 176
7.2 Tooltips met detailinformatie 179
7.3 Drillthrough 183
7.4 Knoppen en stuurtabellen 186
7.5 Dynamische tekst en titels 193
7.6 LeMair-dashboard 197

Index 199

Inleiding

Data wordt steeds belangrijker. Organisaties krijgen steeds meer data uit verschillende bronnen: data uit databases binnen de eigen organisatie, data uit Excel, mail, Facebook, Instagram, etc. Met Power BI kunnen die data ingelezen worden in een datawarehouse, bewerkt en grafisch getoond worden in rapporten en dashboards. Met name het visualiseren van data (visuals) is belangrijk geworden. De app Power BI is wereldwijd marktleider om dat te realiseren. Microsoft biedt de app Power BI Desktop gratis aan. Via het web, telefoon of tablet kunnen de rapporten en dashboards getoond worden nadat ze zijn gepubliceerd.

Boeken over Power BI zijn zeer uitputtend en gaan uitgebreid in op alle mogelijkheden van het programma. Op internet zijn vele filmpjes te vinden over het werken met Power BI, maar dit materiaal is zeer gefragmenteerd. Voor het leren werken met Power BI zijn beide ongeschikt, er is geen logische opbouw. Je kunt Power BI alleen leren door zelf aan de slag te gaan met de app, hands-on. In dit boek is de stof met een logische opbouw beschreven, waarbij je zelf de voorbeelden stap voor stap moet uitvoeren in Power BI. Hierdoor krijg je de gewenste vaardigheid. Bij ieder voorbeeld is bovendien een video-instructie voor een toelichting.
Dit boek behandelt de essentiële onderwerpen van Power BI die voor de (toekomstige) professional van belang zijn.

Opbouw van het boek

In dit boek wordt in hoofdstuk 1 een theoretische toelichting gegeven over begrippen binnen datawarehousing. Databases, relaties, datamodel, data-analytics, datawarehouses, ETL-proces, kubussen, big data, etc. worden toegelicht. Zonder deze basiskennis kan geen professionele datawarehouse ingericht worden.
In het hoofdstuk 2 ga je diverse visuals in Power BI te maken. De benodigde bestanden en video-ondersteuning worden bij het boek geleverd via de website www.aandeslagmetpowerbi.nl.
Even belangrijk als het maken van visuals is consistente data in het warehouse van Power BI laden. Zonder juiste data, geen visuals of visuals met onjuiste informatie. Het bewerken van data en vervolgens inlezen in het datawarehouse van Power BI zijn belangrijke handelingen. Dit wordt het ETL-proces (Extract-Transform-Load) genoemd. Hoofdstuk 3 gaat over het ETL-proces. Een datawarehouse beschikt over data uit verschillende bronnen. Naast de database

van de organisatie, data uit Excel, e-mail, Facebook, Instagram, etc. Vaak moet die data bewerkt worden (transform, data cleansing) voordat die geschikt zijn voor het datawarehouse van Power BI.
Binnen een datawarehouse komen alle soorten relaties voor, één-op-één, één-op-veel en veel-op-veel. Bovendien heeft Power BI een filterrichting. Iedere relatievorm heeft zijn bijzondere eigenschappen. Ook de filterrichting kan voor verrassingen zorgen. Kennis hierover is belangrijk. In hoofdstuk 4 worden deze begrippen met concrete voorbeelden toegelicht.
In het hoofdstuk 5 wordt dieper ingegaan op de DAX-taal van Power BI en de datumtabel met tijdintelligentie. Een zeer belangrijk hoofdstuk.
In hoofdstuk 6 worden enkele handige tips uitgewerkt en nog enkele nuttige visuals.
Als laatste draait hoofdstuk 7 om het bouwen van professionele rapporten en dashboards met behulp van bladwijzers, fraaie tooltips, drillthrough opties, knoppen & stuurtabellen en dynamische titels.

Website www.aandeslagmetpowerbi.nl

Bij het boek hoort de website www.aandeslagmetpowerbi.nl. Gebruik de code op de colofonpagina van je boek om in te loggen. Hierop zijn alle bestanden die in het boek gebruikt worden te downloaden. Er zijn bestanden met miljoenen rijen aan data zodat de kracht van Power BI goed getoond kan worden. Naast deze bestanden zijn op het portaal ook de ondersteunde video's beschikbaar bij de verschillende onderwerpen. Tot slot vind je er de uitwerkingen van de opgaven.

Data modeling 1

Met Power BI kunnen gegevens worden gevisualiseerd. Die gegevens komen uit gemodelleerde tabellen die onderlinge relaties hebben. In dit hoofdstuk wordt de theorie uitgelegd van hoe die gegevenstabellen werken in een datamodel. Hierbij worden veelvoorkomende begrippen en termen die gebruikt worden bij databases en datawarehousing met voorbeelden toegelicht. Het valt buiten het bestek van dit boek om hier diep op in te gaan, maar om goed met Power BI te kunnen werken is kennis van de werking van een datawarehouse en het modeleren van data noodzakelijk. Ook worden in dit hoofdstuk Big Data, Business Intelligence en Data Analytics toegelicht.

Bekijk de video *Introductie Power BI* op www.aandeslagmetpowerbi.nl.

1.1 Business Intelligence

Voor het begrip Business Intelligence (BI) zijn vele definities. Als we deze definities samenvatten dan verstaan we onder BI: uit gegevens (data) managementinformatie creëren die leidt tot meer kennis en inzicht om betere beslissingen te kunnen nemen. BI wordt gezien als de opvolger van Management Informatie Systemen (MIS). Bij die systemen worden periodiek managementrapportages gemaakt. BI is het proces van het verzamelen van data uit diverse bronnen en hiermee snel de gewenste informatie op een visuele manier tonen. Met visueel worden dashboards en grafische weergaven van informatie bedoeld.

1.2 Data, de bron van BI

Organisaties verzamelen en raadplegen steeds meer data. Data die een organisatie nodig heeft om te functioneren, zoals het registreren van patiëntgegevens, urenregistratie, inkoopgegevens, verkoopgegevens, financiële gegevens, studentgegevens of gegevens over verkeersstromen.
Die data kunnen worden omgezet in informatie om inzicht te krijgen in de organisatie. Welk verkoopkanaal genereert de meeste omzet? Wat zijn de kosten per afdeling in een zorginstelling? Wat is per tijdsinterval de gemiddelde snelheid op een snelweg?

Met informatie kan een organisatie ook inzicht krijgen in hoe ze presteert. Op basis van de juiste informatie kan bijvoorbeeld gemonitord worden of je op de goede weg bent op basis van Key Performance Indicators (KPI's). Zijn de doelen bereikt? Dit kunnen kortetermijndoelen zijn, zoals maandcijfers, aantal bezoekers per dag op de Facebookpagina, gemiddeld aantal clicks op een webpagina, of langetermijndoelen, zoals het vergroten van het marktaandeel.
Met informatie kan ook onderzoek gedaan worden om een doelgroep beter te begrijpen. Welk marketingkanaal heeft het meest bijgedragen aan de omzet? Via welk mediakanaal komen de meeste klachten binnen? Hoeveel procent van de orders komt binnen via mobiele telefoon, tablet of laptop?

De data kunnen opgeslagen liggen in databases op lokale computersystemen (on-premise), maar data worden steeds meer opgeslagen in de Cloud. Naast de (basis)data uit informatiesystemen zoals Customer Relationship Management (CRM) of Enterprise Resource Planning (ERP) zijn de gebruikte data afkomstig uit mediabronnen, zoals e-mail, website, Facebook, Twitter, Instagram. Wellicht wordt ook nog gebruikgemaakt van Big Data. Naast de eigen bronnen kunnen via internet beschikbare externe gegevensbronnen nodig zijn, zoals juridische databanken, data via portals van andere organisaties, of landingstijden van de vliegtuigen op Schiphol. Als je al die gegevens uit de verschillende bronnen samenbrengt, ontstaat een datawarehouse.

1.3 Data Analytics

Om aan de gewenste informatie te komen, kan de data uit de aanwezige bronnen geanalyseerd worden (Data Analytics) en getoond worden in dashboards of grafische rapporten. De persoon die zich daarmee bezighoudt, wordt de data-analist genoemd. Power BI is een tool om data analytics mee uit te voeren. Je kunt hiermee uit verschillende databronnen de gewenste managementinformatie visueel tonen via dashboards of grafische rapporten, zie figuur 1.1. Er zijn meerdere tools voor Data Analytics, zoals Tableau, maar volgens onderzoeksbureau Gartner is Power BI de onbetwiste marktleider.

Figuur 1.1

1.4 Databases en tabellen

De meeste gegevens komen uit relationele databases die lokaal of in de Cloud zijn opgeslagen. Dit zijn databases waarin de gegevens van de dagelijkse processen worden opgeslagen. Voor het werken met Power BI is het daarom van belang dat je kennis hebt over hoe gegevens in databases zijn opgeslagen en de werking daarvan. Zonder die kennis begeef je je op glad ijs in Power BI. Het foutief koppelen van gegevens uit databases, datawarehouses en andere bronnen, zoals gegevens uit een Excelbestand, gegevens van een website of gegevens uit e-mail, kan leiden tot onjuiste informatie.

Aan de hand van een voorbeeld zal duidelijk gemaakt worden waarom gegevens in een database worden opgeslagen. Stel, we gaan een lang weekend weg en reserveren een hotel bij Booking.com. Op de site van Booking.com worden incheckdatum, uitcheckdatum, land, stad en een geschikt hotel geselecteerd. Vervolgens een kamertype, aantal kamers, aantal personen per kamer, de persoonsgegevens, e-mailadres en uiteindelijk de betaalgegevens. Na een geslaagde boeking krijg je direct een e-mail met de boekingsgegevens in PDF-formaat. Je zou misschien denken dat Booking.com iedere boeking als PDF-bestand in zijn boekingssysteem opslaat. De bestandsnaam zou dan het unieke boekingsnummer kunnen zijn, zodat het ook snel teruggevonden kan worden. Het grootste

nadeel van die methode is dat er geen managementinformatie of marketinginformatie uit de PDF-bestanden te halen is. Bijvoorbeeld: wat is de omzet over de afgelopen maand? Wat is per land het aantal boekingen over het afgelopen jaar? Welke hotels hebben het meeste opgebracht in het laatste kwartaal? Op basis van het boekingsgedrag kan Booking.com de klant verleiden met een gerichte e-mail. Bij tien boekingen per maand zouden de benodigde gegevens nog achterhaald kunnen worden door de PDF-bestanden een voor een na te pluizen, maar bij duizenden boekingen per dag is dat onuitvoerbaar! Met een database kan dat wel.

In een database worden gegevens (data) in logische groepen opgeslagen. Een vereenvoudigd databasemodel (datamodel) van ons voorbeeld is weergegeven in figuur 1.2.

Figuur 1.2

Er is een bestand *Klant* met alleen (geregistreerde) klantgegevens. Verder een bestand *Boeking* met alleen boekingsgegevens en een bestand *Hotel* met gegevens van de te boeken hotels. In het bestand *Klant* worden de basisgegevens van alle klanten gestructureerd opgeslagen in de vorm van een tabel, zie figuur 1.3. In databaseterminologie wordt gesproken van tabel *Klant*.

KLANT

Klantnr	Naam	Adres	Postcode	Plaats	E-mail
100001	T. de Vries	Steenstraat 41	1380 VB	Weesp	vries34@gmail.com
100002	D. Aanraad	Leiweg 12	1621 AP	Hoorn	d.aanraad@outlook.com
100003	M. den Hoed	Poststraat 111	3011 VD	Rotterdam	m.den.hoed@bijenkorf.nl
……..					

Figuur 1.3

Een tabel heeft rijen en kolommen. In de kolomkoppen van de tabel staan alle gewenste gegevens van een klant. In databases worden die kolomkoppen velden genoemd. Zo is er het veld *Klantnr*, het veld *Naam*, enzovoort, tot en met het veld *E-mail*. In een rij staan de klantgegevens van één klant. In een kolom staan alle waarden van het betreffende veld. Zo bevat de kolom *Naam* alle namen van onze klanten. In een relationele database moeten alle rijen uniek zijn. Twee of meer rijen met exact dezelfde informatie opnemen is niet toegestaan. Het veld *Klantnr* zorgt voor unieke rijen in tabel *Klant*. Hierdoor ontstaat, indien ook de

We leven in een datatijdperk. We hebben steeds meer behoefte aan analyse van grote hoeveelheden data – *data analytics* – en visualisatie in grafische rapporten en dashboards. Dit wordt Business Intelligence genoemd. De BI-tools zijn in de loop der jaren sterk verbeterd. De eerste BI-tools konden alleen door specialisten gebruikt worden, maar Power BI is een tool voor de gebruikers: Self Service BI.

In dit boek worden eerst de onmisbare basisbegrippen uitgelegd, zoals gegevensbronnen, databases, relaties, datawarehouses, kubussen en big data, met veel voorbeelden en illustraties. Hierna start je met Power BI. De meest gebruikte visualisaties (visuals) en de DAX-taal worden toegelicht. Na het lezen van dit boek heb je een gedegen kennis van Power BI, die je direct in de praktijk kunt toepassen.

Het leren werken met een applicatie (app), zoals Power BI, kan alleen als je er hands-on mee aan de slag gaat. Om je hierbij te ondersteunen, maken we gebruik van *blended learning.* Ieder voorbeeld wordt vooraf-gegaan door een video-instructie op de website die bij het boek hoort. Na deze instructie ga je naar het boek. Daar wordt het voorbeeld stap-voor-stap toegelicht en voorzien van relevante schermafdrukken. Op de website vind je veel meer informatie, zoals data en uitwerkingen. Er staat ook een introductievideo op waarop je kunt zien dat miljoenen rijen aan data in fracties van seconden in Power BI tot een grafische weerwave worden verwerkt.

Ben Groenendijk is docent informatica aan de Hogeschool Rotterdam. Hij is auteur van diverse boeken over o.a. databases en Excel.

ISBN 978-90-244-2854-0

9 789024 428540 >

www.aandeslagmetpowerbi.nl
www.boomhogeronderwijs.nl