

Handboek Power BI ^{3^e} druk

Boom

PETER TER BRAAKE



Handboek Power BI

Peter ter Braake

Boom

inclusief website!

Met behulp van onderstaande unieke activeringscode krijg je toegang tot de website **www.handboekpowerbi.nl** voor extra materiaal. Deze code is persoonsgebonden en gekoppeld aan de 3e 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: Nu-nique grafische vormgeving
Basisontwerp omslag: Dog & Pony, Amsterdam
Omslagontwerp: Haagsblauw, Den Haag
Beeld omslag en hoofdstukopener: Koto_feja, Getty images

© Peter ter Braake & Boom uitgevers Amsterdam, 2021

Behoudens de in of krachtens de Auteurswet gestelde uitzonderingen mag niets uit deze uitgave worden veeelvoudigd, 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 het overnemen van (een) gedeelte(n) uit deze uitgave in bijvoorbeeld een (digitale) leeromgeving of een reader in het onderwijs (op grond van artikel 16, Auteurswet 1912) kan men zich wenden tot Stichting Uitgeversorganisatie voor Onderwijslicenties, Postbus 3060, 2130 KB Hoofddorp, www.stichting-uvo.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 9789024441419

ISBN e-book 9789024441426

NUR 173

www.handboekpowerbi.nl

www.boomhogeronderwijs.nl

Inhoudsopgave

Woord vooraf	11
1 Inleiding	15
1.1 Wat is Business Intelligence?	16
1.2 Self-Service BI	16
1.3 Voor wie is Self-Service BI?	19
1.4 Power BI	20
1.4.1 Power Query	21
1.4.2 Power Pivot	22
1.4.3 Power View	23
1.4.4 Power BI Service	24
1.5 Power BI-scenario's	25
1.6 Power BI activeren	26
1.6.1 Power BI Account	26
1.6.2 Power BI Desktop installeren	28
1.7 Een eerste snelle kennismaking met Power BI Desktop	30
1.7.1 De Northwind-database	30
1.7.2 Eerste kennismaking – Power Query	31
1.7.3 Eerste kennismaking - Visualisaties	36
1.8 Conclusie	39
2 Power Query	41
2.1 Architectuur-overview	41
2.2 Gegevens inlezen van een OData-bron	42
2.3 Gegevensprofilering	44
2.3.1 Kolomkwaliteit	44
2.3.2 Kolomdistributie	45
2.3.3 Kolomprofiel	46
2.4 Gegevens bewerken	46
2.4.1 Kolommen verwijderen	47
2.4.2 Rijen filteren	49
2.4.3 Query's samenvoegen	50
2.4.4 Waarden vervangen	55
2.4.5 Kolommen een andere naam geven	55
2.4.6 Query een naam geven	56
2.4.7 Stappen terugkijken	57
2.4.8 Gegevens laden	57

2.5	Nog meer bewerkingen	58
2.5.1	Een gegevensbron hergebruiken	59
2.5.2	Berekende kolommen maken	59
2.5.3	Gegevenstype aanpassen	62
2.5.4	Feitentabel afmaken	63
2.6	Gegevens uit CSV-bestanden inlezen	64
2.6.1	Draaitabel opheffen	66
2.6.2	Een functie maken	67
2.7	Parameters	68
2.8	Gegevens inlezen uit een map	71
2.8.1	Een functie gebruiken	73
2.8.2	Kolom splitsen	73
2.8.3	Berekende kolom vanuit voorbeelden	74
2.9	Gegevens inlezen uit de SQL Server	75
2.9.1	Gegevensverbindingsmodus	77
2.9.2	Query folding	78
2.10	Andere gegevensbronnen benaderen	80
2.10.1	Gegevens inlezen uit Excel	80
2.10.2	Gegevens inlezen uit tekstbestanden	81
2.10.3	Gegevens van het web lezen	83
2.10.4	Gegevens uit Microsoft-serverproducten	84
2.10.5	Big Data	84
2.11	Gegevensbroninstellingen	85
2.12	Andere transformaties	87
2.12.1	Tekst kolombewerkingen	87
2.12.2	Getallenbewerkingen	88
2.12.3	Datumbewerkingen	88
2.12.4	Doorvoeren (Fill up/down)	88
2.13	Excel	89
2.14	Tot slot	90
3	M, the Power Query Language	93
3.1	Formulebalk en Query Editor	93
3.2	De basis van M, het let statement	95
3.3	Gegevensbronnen	97
3.3.1	Zelf een tabel maken in M	97
3.3.2	Gegevens uit een Excel-werkboek inlezen	99
3.3.3	Gegevens uit SQL Server inlezen	101
3.3.4	Andere gegevensbronnen	103
3.4	Functies	104
3.4.1	Een functie maken	104
3.4.2	Een functie gebruiken	106
3.4.3	Azure Cognitive services gebruiken	107
3.5	Transformaties	111
3.6	Verwijzen naar query's	112

3.7	Een datumdimensie maken	114
3.8	Plaatjes toevoegen	118
3.9	Tot slot	119
4	Power BI-visualisaties	121
4.1	Introductie	121
4.1.1	Doel en doelgroep	121
4.1.2	Te gebruiken visualisaties	123
4.1.3	Kleuren	123
4.1.4	Aantal visualisaties per rapport	124
4.1.5	Standaardiseren	124
4.2	De Power BI-rapportweergave	124
4.3	Eerste Power BI-rapport	126
4.3.1	Een kolomdiagram maken	127
4.3.2	Opmaak – het verfrullertje	129
4.3.3	Slicer	131
4.3.4	Tabel	133
4.3.5	Donut	135
4.3.6	Drill-down	135
4.3.7	Paginaopmaak	136
4.4	Meer interactie	137
4.4.1	Tooltips	137
4.4.2	Drill-through	139
4.4.3	Relaties	143
4.4.4	Knoppen	145
4.4.5	Bladwijzers	146
4.4.6	Filters	148
4.4.7	Sync slicers	149
4.5	Huisstijl	151
4.6	Tot slot	151
5	DAX	155
5.1	Theorie vooraf	155
5.1.1	Berekende kolom	155
5.1.2	Berekende meetwaarde	156
5.1.3	Impliciete meetwaarde	157
5.1.4	Context	159
5.2	Berekende kolommen	160
5.2.1	Related-functie	161
5.2.2	Relatedtable-functie	162
5.2.3	PATH, PATHITEM en LOOKUPVALUE	164
5.3	Berekende meetwaarden (calculated measures)	167
5.3.1	Simpele expliciete meetwaarden	167
5.3.2	Calculate-functie	169
5.3.3	If- en Info-functies	172
5.3.4	Variabelen	174

5.4	Time Intelligence	175
5.4.1	TotalYTD-functie	175
5.4.2	Sameperiodlastyear-functie	177
5.4.3	Averagex- en Summarize-functie	179
5.5	Snelle meting (Quick measures)	181
5.6	Een datumdimensie maken	182
5.7	Tot slot	184
6	Gegevensmodel	187
6.1	Inleiding	187
6.2	Relaties	189
6.2.1	Actieve en inactieve relaties	189
6.2.2	Userrelationship	191
6.2.3	Soorten relaties	192
6.2.4	Kruisfilterrichting	194
6.3	Tabeleigenschappen	196
6.4	Veldeigenschappen	200
6.5	Aggregatietabellen	205
6.6	Datumdimensie	209
6.7	Row Level Security	210
6.8	Tot slot	211
7	Power BI Service	213
7.1	Beginnen met Power BI Service	213
7.1.1	Aanmelden	213
7.1.2	Gratis versus betaald	215
7.1.3	Cloud of lokaal	215
7.2	Power BI Desktop publiceren naar Power BI	216
7.2.1	Werkruimte (Workspace)	216
7.2.2	Publiceren	217
7.2.3	Row level security	218
7.2.4	Gegevensbronreferenties	219
7.3	Dashboards	220
7.3.1	Een dashboard maken	220
7.3.2	Extra elementen aan je dashboard toevoegen	223
7.3.3	Een dashboard opmaken	224
7.3.4	Commentaar	225
7.3.5	Telefoonweergave	226
7.3.6	Aanbevolen en favorieten rapporten en dashboards	227
7.3.7	Volledig scherm gebruiken	228
7.3.8	Tegels aanpassen	228
7.3.9	Gegevenssets, rapporten en dashboards	229
7.4	Gegevens ophalen	229

7.5	Gegevens vernieuwen	231
7.5.1	Import versus DirectQuery	232
7.5.2	Gegevens verversen	233
7.6	Rapporten en dashboards delen	237
7.6.1	Delen van rapporten en dashboards	237
7.6.2	Werkruimte delen	239
7.6.3	Power BI App	241
7.6.4	Publiceren op internet	243
7.6.5	Insluiten in SharePoint online	244
7.6.6	Power BI Embedded	244
7.7	Werkmappen en gegevensstromen	244
7.7.1	Excel	244
7.7.2	Gegevensstromen (Dataflows)	246
7.8	Tot slot	246
	Register	249

Woord vooraf

Dit boek gaat over Power BI. Het gaat over Power BI Desktop, de Power BI-onderdelen van Excel en de Power BI Service. Voor het lezen en begrijpen van dit boek is voorkennis niet strikt noodzakelijk. Affiniteit met het werken met gegevens en enige ervaring daarmee is wel aan te raden. Dat kan Excel-ervaring zijn met tabellen en draaitabellen. Maar die ervaring kan ook zijn opgedaan met andere programma's zoals Access of SQL Server.

Microsoft is al enkele jaren hard bezig om Power BI op de markt te zetten. De releases volgen elkaar snel op. Maandelijkse releases brengen nieuwe functionaliteiten. Tegen de tijd dat je dit leest, heeft Power BI alweer meer mogelijkheden dan op het moment dat dit geschreven wordt. Ook zullen enkele screenshots die je in dit boek aantreft ondertussen niet meer overeenkomen met wat je in de praktijk ziet. Maar de basis is ondertussen stabiel. De kennis die je opdoet in dit boek zal je verder brengen, ook met nieuwere versies.

Dit boek zal je vertrouwd maken met de basisfunctionaliteit van Power BI en het zal je inzicht geven in wanneer je wat moet doen. Die doelstelling is niet afhankelijk van de meest recente versie. Na het lezen van dit boek heb je genoeg kennis en kunde om met Power BI oplossingen te bouwen en daarbij ook gebruik te maken van nieuwere functionaliteiten, zelfs als die in dit boek nog niet zijn beschreven. Na het lezen van dit boek kan je met Power BI Desktop data-analyses uitvoeren en rapporten en dashboards bouwen.

Er is op het internet heel veel te vinden over Power BI. Veel demo's, blogs, video's en de nodige tutorials. In dit boek is alles bij elkaar gebracht. Van A tot Z doorloop je met dit boek alle relevante zaken van Power BI. Je zult echter nooit uitgeleerd raken. Als je eenmaal de basis onder de knie hebt, kan je verder leren met alle informatie die op het internet beschikbaar is.

Een ander deel van de meerwaarde van dit boek is dat het een Nederlandstalig boek is. Voor sommige mensen is het immers gemakkelijker om een boek in het Nederlands te lezen dan in het Engels. Los daarvan kan je een Nederlands- of een Engelstalige versie van Power BI Desktop (of Excel) hebben. Om beide talen te ondersteunen noemen we zo veel mogelijk Engelse namen tussen haakjes achter de Nederlandse.

In dit boek werken we veel praktische voorbeelden uit. Je kan deze voorbeelden stapsgewijs meemaken. De stappen die je moet volgen worden met bullets aangegeven. Veel van deze voorbeelden maken gebruik van de Northwind-voorbeelddatabase die als ODATA-feed beschikbaar is op het internet. Screenshots zijn gebaseerd op de Power BI Desktop-versie van december 2020. De voorbeelddatabases zijn Engelstalig. Ondanks dat het boek in principe Nederlandstalig is, is ervoor gekozen deze te gebruiken omdat ze gemakkelijk voorhanden zijn en mooie gegevens bevatten om mee te spelen. Nadeel is dat we als gevolg van deze keuze hier en daar Engelse termen gebruiken.

Technische vereisten

Om de meeste voorbeelden in dit boek zelf te kunnen maken, heb je een aantal dingen nodig. Deze voorwaarden zijn:

- Een computer met Windows.
- Power BI desktop of de mogelijkheid dit te downloaden en installeren.
- Internettoegang.

Verder zijn er enkele extra's. Deze omvatten:

- In het boek maak je met een tijdelijk e-mailadres een gratis Power BI-account aan. Wil je rapporten kunnen delen met anderen (bijvoorbeeld medestudenten) zal je een betaalde versie van Power BI moeten nemen (er is een trial beschikbaar met beperkte geldigheidsduur) en met je medestudenten in eenzelfde Office365 tenant zitten. In dat geval moet je een school- of werk-e-mail gebruiken om een account aan te maken.
- Een enkele opdracht maakt gebruik van een SQL Server-database. Wil je die voorbeelden zelf uitproberen moet je de beschikking hebben over een SQL Server. SQL Server Express is gratis te downloaden evenals de gebruikte voorbeelddatabase.



Bij dit boek hoort de website www.handboekpowerbi.nl. Hierop vind je kennisvragen om mee te oefenen, datasets als downloads en links naar externe bronnen. Ook is hier het boek online te raadplegen. Docenten kunnen op de website terecht voor PowerPoints.

Veel plezier met leren en het ontdekken van de Power BI-mogelijkheden.



Welkom in het *Handboek Power BI*. Als je dit leest, is de kans groot dat je (regelmatig) met gegevens werkt en informatie wilt presenteren. Excel is al jaren het meest gebruikte programma om met gegevens te werken. Dat komt enerzijds doordat het gemakkelijk is om mee te werken en anderzijds door de vele krachtige opties, zoals draaitabellen, grafieken en voorwaardelijke opmaak. Power BI is begonnen als nieuwe extra functionaliteit in Excel. Hoewel veel van die nieuwe functionaliteit nog altijd in Excel zit, raadt Microsoft tegenwoordig aan om Power BI Desktop te gebruiken.

Met Power BI wil Microsoft het makkelijk maken om met gegevens te werken.

- Je moet met grote gegevensverzamelingen (datasets) kunnen werken.
- Je moet gemakkelijk willekeurige bronnen kunnen benaderen. Daar horen ‘eigen’ databases bij, maar bijvoorbeeld ook gegevens van Facebook of andere websites.
- Je moet gemakkelijk de gegevens kunnen aanpassen aan je behoeftes.
- Je moet al die gegevens gemakkelijk kunnen omtoveren tot zinvolle informatie met intuïtieve rapporten.
- En tot slot moet je je werk gemakkelijk kunnen delen met anderen.

In dit boek lees je wat de mogelijkheden van Power BI zijn. Je leert ook met alle mogelijkheden te werken. En tot slot zie je de mogelijkheden van de Power BI Service. Dat is een web portal waar je rapporten naar kan uploaden, waar je dashboards kan maken en waar je rapporten en dashboards kan delen met anderen.

Omdat BI in Power BI staat voor Business Intelligence krijg je in het eerste hoofdstuk achtergrond over wat Business Intelligence is en de beoogde rol van Power BI in het grotere geheel. De volgende hoofdstukken gaan in detail in op de verschillende onderdelen.

1.1 Wat is Business Intelligence?

De letters BI achter Power BI staan voor Business Intelligence. Hieronder volgt een definitie van Business Intelligence (vrij vertaald naar de definitie van Gartner op www.gartner.com/it-glossary/business-intelligence-bi):



Business Intelligence is een overkoepelende term waarmee applicaties, infrastructuur en hulpmiddelen en aangeraden werkwijzen worden bedoeld, die als doel hebben om gegevens beschikbaar te stellen om de juiste beslissingen te kunnen nemen op basis van correcte, betrouwbare informatie.

Bovenstaande definitie is samen te vatten als: van gegevens informatie maken. Bedrijven hebben databases vol met gegevens. Ook zijn er veel gegevens beschikbaar in Excel-bestanden of in bijvoorbeeld CSV (Comma Separated Values)-bestanden. Daarnaast is er tegenwoordig een oneindige hoeveelheid gegevens beschikbaar via het internet. Maar wat kan je met al die gegevens? Ben je in staat die gegevens te gebruiken? Soms hoor je mensen zeggen: data (gegevens) zijn het nieuwe goud. Kan je informatie maken van al die gegevens? Informatie die jou, of anderen binnen het bedrijf, in staat stelt om betere beslissingen te nemen?

Het doel van Business Intelligence is om deze gegevens te gebruiken en er informatie van te maken zodat jij, of anderen binnen je bedrijf, hun werk beter kunnen doen.



Business Intelligence heeft als doel competitief voordeel te creëren en organisaties slimmer te laten werken.

1.2 Self-Service BI

In het verleden was Business Intelligence vaak een zaak van de IT-afdeling. Technenuten bouwden datawarehouses, maakten analytische/semantische modellen en stelden rapportages samen voor de organisatie. De mensen binnen een bedrijf die beslissingen moeten nemen (dat zijn niet alleen maar managers) waren afhankelijk van wat de technenuten maakten. Deze situatie kan een aantal problemen opleveren:

1. IT levert de gewenste informatie te laat op.
2. IT levert de verkeerde informatie op.
3. Business Intelligence-professionals zijn duur.

In het eerste geval heeft het Business Intelligence-team de vraag om informatie goed begrepen. Door technische complexiteit of simpelweg door drukte van de betrokken mensen, worden de gevraagde rapporten te laat geleverd. In

sommige gevallen is de vraag om informatie al niet meer relevant, omdat de beslissingnemer zijn of haar besluit al heeft genomen.

In het tweede geval hebben de makers van de Business Intelligence-oplossing de vraag niet goed begrepen en verkeerde rapporten gemaakt. De vraag om informatie vanuit de organisatie wordt dus helemaal niet beantwoord. De IT'ers zijn minder goed op de hoogte van de bedrijfsprocessen dan de mensen op de werkvloer. Daardoor zijn ze soms minder goed in staat om de juiste informatie te leveren.

Een derde probleem is dat Business Intelligence-professionals duur zijn in termen van uurloon. Veel vragen om informatie zullen achteraf bezien eenmalige vragen zijn. Het loont dan niet om dure professionals in te zetten om de vragen te beantwoorden. De mensen die met de vraag zitten, kunnen beter zelf in staat gesteld worden om met antwoorden te komen.

Microsoft noemt mensen die met gegevens werken kenniswerkers. Deze kenniswerkers worden vaak in staat geacht om zelf sneller, beter en goedkoper bruikbare informatie te maken van de beschikbare gegevens. Er wordt niet van hen verwacht om Basel II Compliant Credit Risk Datawarehouses te bouwen. Wel kunnen ze de gegevens uit zo'n database halen en tot nuttige informatie verwerken.

Bovenstaande korte probleemschets is de verantwoording van Self-Service BI.

Self-Service Business Intelligence draait om het in staat stellen van eindgebruikers om zelf in hun informatiebehoefte te voorzien zonder anderen te vragen oplossingen voor ze te bouwen om de juiste informatie op het juiste moment ter beschikking te hebben.



Bij Self-Service BI is het de bedoeling dat beslissingnemers/kenniswerkers zelf in hun eigen informatiebehoefte voorzien. Een van de achterliggende gedachten daarbij is dat mensen steeds handiger worden met computers. Veel (maar niet alle) mensen kunnen goed genoeg overweg met een computer om zelf hun eigen informatie te vergaren. Bovendien weten de kenniswerkers zelf vaak goed waar welke informatie voorhanden is. Belangrijk is wel dat mensen in staat gesteld worden om deze taak op zich te nemen. Daar zijn twee belangrijke aspecten aan:

1. De gegevens moeten wel voorhanden zijn.
2. De mensen moeten de juiste programma's krijgen om mee te werken.

De gegevens moeten voorhanden zijn

Het klinkt leuk: mensen zelf hun rapporten laten maken. Maar waar komen de gegevens vandaan?

In wat grotere organisaties zal er nog steeds sprake zijn van een (corporate) datawarehouse. Dit is een database waar veel van de relevante gegevens zijn verzameld met informatievoorziening als doel.

Steeds vaker wordt het datawarehouse aangevuld met geprepareerde datasets (gegevensverzamelingen) die voor gebruikers beschikbaar worden gemaakt. Deze datasets kunnen op veel verschillende manieren aangeboden worden.

Het kunnen databases zijn die speciaal zijn ingericht voor een analyse-doel (datamarts) of bijvoorbeeld exports in de vorm van CSV-bestanden. Een andere mooie optie zou Power BI DataFlow zijn, waar de gegevens speciaal voor gebruik in Power BI zijn geprepareerd en klaargezet in de cloud.

Uiteraard kunnen datasets ook van buiten de organisatie komen. Op het internet zijn bijvoorbeeld gegevens over het weer of demografische gegevens beschikbaar. Beide kunnen interessant zijn om te gebruiken in bijvoorbeeld omzetanalyses.

Een veel gehoord voorbeeld, vaak genoemd in combinatie met big data, is Twitter. Het zou interessant kunnen zijn om alle tweets waarin een bedrijf wordt genoemd te verzamelen en te analyseren op sentiment. Wordt er positief of negatief getweet over het bedrijf, waar gaan die tweets over en wat kunnen we eraan doen?

Een belangrijk punt om hier te noemen is security. Databasebeheerders zijn jarenlang bezig geweest om databases af te schermen. Als alle mogelijke bronnen van informatie zijn afgeschermd, wordt Self-Service BI onmogelijk. Deze opmerking is geenszins een pleidooi om security af te schaffen. Welke gegevens waar, wanneer en hoe beschikbaar te maken is een belangrijk onderdeel van een Business Intelligence-implementatie binnen een organisatie. Een organisatie zal bijvoorbeeld wetgeving zoals de AVG (Algemene Verordening Gegevensbescherming) moeten naleven.

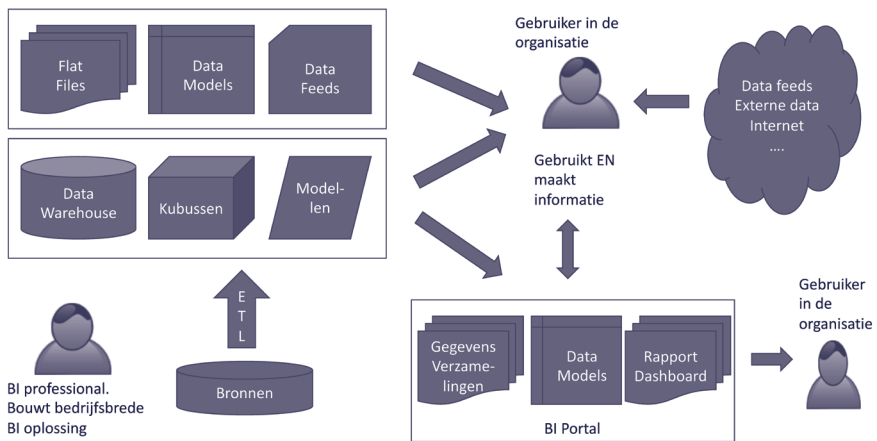
De mensen moeten de juiste programma's krijgen om mee te werken

Een ander belangrijk onderdeel van Self-Service BI is de mensen van de juiste programma's (tools) te voorzien. Binnen de producten die Microsoft maakt is Excel al sinds jaar en dag het favoriete programma van veel mensen. Microsoft heeft er in eerste instantie voor gekozen de meest gebruikte tool nog beter geschikt te maken voor Business Intelligence-doeleinden. Een voordeel ten opzichte van concurrerende programma's is dat je geen nieuw programma hoeft te leren. Uiteraard moet je wel de nieuwe functionaliteiten die Power BI biedt onder de knie krijgen. Toen de Power BI-capaciteiten zoals die binnen Excel beschikbaar zijn eenmaal populair waren geworden, veranderde de strategie van Microsoft. Nu wordt aangeraden om Power BI Desktop als tool te gebruiken. Je moet dus toch een nieuw programma leren gebruiken. Vandaar dit boek!

Het voordeel van Power BI Desktop is dat het specifiek gemaakt is voor Power BI-oplossingen (rapporten), terwijl Excel een veel generieker programma is. Bovendien is Power BI Desktop gratis.

Self-Service BI is geen vervanging van de klassieke corporate Business Intelligence. Het is er eerder een aanvulling op. Business Intelligence-initiatieven kunnen nu door gebruikers zelf op de werkvloer opgepakt worden. De succesvolle initiatieven kan je in een later stadium opnemen in de bedrijfsbrede Business Intelligence-oplossing.

De totaaloplossing kan er dan komen uit te zien zoals in figuur 1.1.



Figuur 1.1 Self-Service Business Intelligence

1.3 Voor wie is Self-Service BI?

In zijn algemeenheid kunnen we binnen een organisatie verschillende soorten gebruikers van rapportages onderkennen. Bill Inmon definieert vier soorten gebruikers:

1. farmers;
2. tourists;
3. explorers;
4. miners.

Farmers

Voor farmers zijn rapporten in de vorm van vooraf gedefinieerde pdf-bestanden misschien wel genoeg. Ze hebben vaste, terugkerende informatiebehoeften. Een manager op het gebied van personeelszaken wil bijvoorbeeld wekelijks een overzicht van het aantal fte's, uitgesplitst naar externen en mensen in loondienst en het ziekteverzuim in beide groepen.

Tourists

Tourists hebben naast de vaste informatiebehoeften van farmers behoefte aan ad-hoc-informatie. Voor werkzaamheden die maar af en toe uitgevoerd worden, ontstaat opeens een andere informatiebehoefte dan de dagelijkse vraag. Rapporten moeten wat interactiever worden. Misschien dat deze mensen zelf rapporten moeten kunnen aanpassen.

Explorers

Explorers zijn mensen die meer gaan analyseren. Rapporten zijn vrij statisch. Rapporten geven een antwoord op een vooraf gestelde vastomlijnde vraag. Dat sluit precies aan bij farmers en is grotendeels afdoende voor tourists. Explorers

willen op zoek naar verbanden en willen niet alleen het ‘wat’ weten (de droge cijfers), maar ook het hoe en waarom (de cijfers achter de cijfers). Een tool als Excel of Power BI sluit beter aan bij deze behoefte dan een vast rapport. De tool stelt deze mensen in staat om zelf te zoeken naar nieuwe informatie.

Miners

Miners gaan nog een stap verder. Dit zijn de mensen die aan de slag gaan met grote hoeveelheden gegevens om te proberen patronen in de gegevens te herkennen. Hoe kan een verzekering een frauduleuze claim onderscheiden van een terechte claim? Waarin onderscheiden beide zich en kan ik met die wetenschap van een nieuwe claim voorspellen of het om een frauduleuze claim gaat? Dit soort mensen hebben statistische programma's en dataminingtools nodig. Volgens bovenstaande indeling is Power BI Desktop (of Excel) niet geschikt voor de eerste groep mensen.

Ook de tweede groep, de tourists, zijn waarschijnlijk meer gebaat bij door anderen gemaakte formele rapporten. Binnen Power BI kunnen BI-professionals voor deze doelgroep zogenaamde ‘Paginated reports’ maken. Daar gaat dit boek niet verder op in. In de Power BI Service (de web portal waar we rapporten en dashboards gebruiken en delen) kunnen deze mensen wel goed gebruikmaken van de rapporten en dashboards die anderen hebben gemaakt en gedeeld. Excel en Power BI Desktop zijn uitstekend geschikt voor de explorers en de miners. Met beide programma's kan je zelf grasduinen in de beschikbare informatie en de analyses doen die je wilt.

1.4 Power BI

Microsoft koos er in eerste instantie voor om Microsoft Office Excel hét programma te laten zijn om Self-Service Business Intelligence te doen. Microsoft heeft veel nieuwe functionaliteit in Excel gestopt om dat te bereiken. Die nieuwe onderdelen werden gezamenlijk bekend onder de naam Power BI. Tegenwoordig wordt met Power BI meestal de Power BI Service bedoeld, de cloud-omgeving van Power BI waarover je later in dit boek meer leert. Ook is Excel als primaire tool vervangen door Power BI Desktop. De functionaliteit is echter voor het overgrote deel vergelijkbaar en in te delen in drie onderdelen:

1. Power Query;
2. Power Pivot;
3. Power View.

De individuele namen van de onderdelen worden steeds minder gebruikt. Qua functionaliteit zijn ze nog duidelijk herkenbaar en te onderscheiden. In deze paragraaf maak je kort kennis met de verschillende onderdelen. De diepgang komt in de volgende hoofdstukken.

1.4.1 Power Query

Met Excel heb je via de Gegevens (Data)-tab altijd al veel databronnen kunnen ontsluiten, zoals:

- MS Access;
- SQL Server databases en SQL Server Analysis services-kubussen;
- tekstbestanden;
- XML en OData;
- ODBC en OLEDB-bronnen.

Hoewel dit een heleboel verschillende bronnen zijn, zeker als je bedenkt dat je met ODBC en/of OLEDB bijna alle soorten databases kan aanspreken, wordt deze lijst tegenwoordig toch beperkt. Big Data en de bijbehorende databases vormen een steeds belangrijker onderdeel van de gegevens die we willen gebruiken bij onze analyses. Veel verschillende websites zoals Facebook en andere sociale media-sites worden ook steeds belangrijker om te gebruiken.

Daarnaast zijn sommige databases van bijvoorbeeld CRM (Customer Relationship Management)-applicaties niet rechtstreeks benaderbaar. Toch zit in dat soort databases een schat aan informatie. Dit soort informatie moet (gemakkelijk) bruikbaar zijn.

Het onderdeel Power Query van Power BI heeft twee doelen:

1. meer verschillende gegevensbronnen ontsluiten;
2. gemakkelijk gegevens kunnen aanpassen aan de behoeftes.

Meer verschillende gegevensbronnen ontsluiten

Power Query geeft je vele mogelijkheden om gegevens te benaderen, van klassieke databases via SAAS-services tot Big Data. Microsoft is van plan de lijst met mogelijke bronnen steeds verder uit te breiden. In Excel 365 en Excel 2016 en later is Power Query al volledig geïntegreerd in het lint Gegevens (Data). Je ziet de naam alleen nog terug in de menuoptie Power Query editor starten.

Gemakkelijk gegevens kunnen aanpassen aan de behoeftes

Power Query is de ETL-tool van Power BI. ETL staat voor Extract, Transform, Load en is een heel belangrijk onderdeel van Business Intelligence-oplossingen. Extract betekent in deze context het lezen van gegevens uit bronsystemen. Daar hebben we het net over gehad.

Transform betekent het aanpassen van de gegevens aan hoe de gegevens gebruikt moeten gaan worden. De gegevens zoals je ze aantreft in de bron zullen niet altijd voldoen aan de eisen die je eraan stelt. Regelmatig tref je gegevens aan in de vorm van een draaitabel, terwijl je de gegevens nodig hebt in de vorm van een gewone tabel. Misschien komen de datatypes niet overeen met wat je wilt doen. Datums kunnen bijvoorbeeld zijn opgeslagen als tekst in plaats van als 'echte' datums. Of er moeten nog berekeningen gedaan worden voordat je met de gegevens kan werken. Power Query stelt je in staat om deze aanpassingen (en

vele andere aanpassingen) aan gegevens gemakkelijk te doen voordat je verder werkt met de gegevens op rapporten.

De L van ETL staat voor Load, ofwel Laden. De gegevens die je inleest moeten ergens naartoe. 'Ergens' is in de context van Power BI naar de Power BI-gegevensset (dataset) (Power Pivot-datamodel genoemd in Excel). De Power BI-gegevensset is een in-memory database waarin de gegevens worden opgeslagen.

Bovenstaande is strikt genomen niet helemaal juist. Power BI kent verschillende soorten verbindingen. Je leert hier later meer over. Binnen Excel kunnen gegevens ook naar een Excel-werkblad geschreven worden.

In hoofdstuk 2 leer je werken met Power Query.

1.4.2 Power Pivot

Veel mensen, waaronder bijvoorbeeld data-analisten, willen met gegevens kunnen 'spelen'. Je wilt snel en gemakkelijk, met simpel slepen van gegevens, mooie, sprekende visualisaties kunnen maken. Power Pivot is een datamodel dat tot doel heeft deze wens te verwezenlijken.

Power Pivot heeft twee doelen:

1. het gemakkelijk maken om met gegevens te werken;
2. goede performance bieden voor Analytics.

Het gemakkelijk maken om met gegevens te werken

Power Pivot is wat ook wel een semantisch model wordt genoemd. De definitie van semantisch model is als volgt:



Een **semantisch model** vertaalt een technische database-implementatie in een voor eindgebruikers leesbaar en begrijpelijk model dat bruikbaar is zonder over (veel) technische kennis te beschikken.

In een semantisch model worden alle gegevens die nodig zijn om zinvolle data-analyses te doen, samengebracht in hun onderlinge verband en in de taal van de gebruiker. Een goed semantisch model is intuïtief te gebruiken. De gegevens staan logisch gegroepeerd in de taal van de gebruiker in het model. De gebruiker sleept wat hij of zij nodig heeft naar een rapport. Doordat het model de onderlinge verbanden tussen de gegevens kent, wordt automatisch de juiste query gegenereerd om de gegevens op te halen.

Power Pivot is binnen Power BI dit semantische model dat je straks in staat stelt intuïtief met de gegevens te werken.

Goede performance bieden voor Analytics

Laat ik beginnen met een definitie te geven van Analytics.

Analytics is het ontdekken en communiceren van betekenisvolle patronen in gegevensverzamelingen.



Wat ik eerder luchtig ‘spelen’ met gegevens heb genoemd, valt onder Analytics. Het gaat hier niet om vooraf gedefinieerde vragen en daarmee om vooraf geschreven query’s. Het gaat om het ontdekken, om het vinden van nieuwe verbanden. Dat vergt een dynamisch proces.

Vastomlijnde vragen kunnen goed voor performance getuned worden. Willekeurige ad-hoc-analyses zijn veel moeilijker te tunen. Power Pivot is een techniek waarbij alle gegevens die in het model opgenomen zijn in het geheugen van de computer zitten. Daar worden ze, sterk gecomprimeerd, bewaard zolang je met jouw model werkt. Power View, het rapport- of visualisatiegedeelte, kan er dan heel gemakkelijk en snel bij, wat een goede performance oplevert, hoe de gegevens ook gebruikt worden.

Als ik rapportagetools zou moeten classificeren zou ik zeggen dat Power BI een datamodelleertool is. Het datamodel is het belangrijkste onderdeel van een Power BI-oplossing. Het bepaalt de functionaliteit, de mogelijkheden en de performance van de rapporten, het einddoel van de oplossing.

In hoofdstuk 4 leer je de basis van het maken van een semantisch (Power Pivot-) model.

1.4.3 Power View

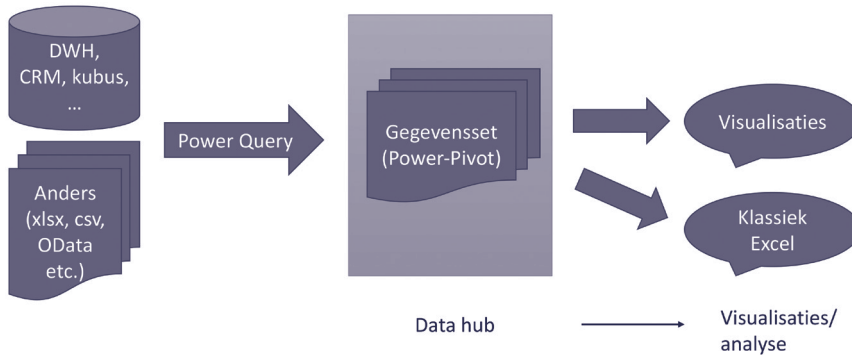
In de definitie van Analytics in de vorige paragraaf gebruiken we het woord ‘communiceren’. De ontdekte patronen moeten zichtbaar gemaakt kunnen worden. Anders valt er niets te communiceren. Bovendien zegt een plaatje meer dan duizend woorden, wat hier betekent dat sommige verbanden verborgen blijven totdat je de gegevens visualiseert met de juiste grafiekvorm.

Power View is een drag-and-drop (slepen en neerzetten) rapportageomgeving die gebruikmaakt van een Power Pivot-model. Met Power View kan je gemakkelijk visualisaties maken die patronen in je gegevens duidelijk inzichtelijk maken. Ook eenvoudiger zaken als omzet per maand-rapportages kan je gemakkelijk en snel maken.

Hoewel ik van Power Pivot heb gezegd dat dat het belangrijkste onderdeel is, zou ik dat ook kunnen zeggen van Power View. Het gaat om het inzichtelijk maken, het beschikbaar stellen en het delen van informatie in de vorm van heldere, begrijpelijke rapporten. Power View is daarmee het eindresultaat. Power Pivot moet dat mogelijk maken.

In Power BI Desktop worden de drie zojuist genoemde onderdelen niet meer apart met eigen naam benoemd. Power Query zie je nog veel, omdat je veel zal

werken met de Power Query editor. De namen Power Pivot en Power View worden alleen nog in Excel gebruikt. Power View is zelfs ‘deprecated’ verklaard en zal dus geheel uit Excel gaan verdwijnen. Power BI bestaat echter wel degelijk uit drie componenten: een datapreparatie-tool, een datamodel en de visualisaties. In hoofdstuk 6 leer je werken met visualisaties en leer je om mooie visualisaties en rapporten te bouwen. Als je alle onderdelen op een rij zet, kan je onderstaand architectuurplaatje tekenen (zie figuur 1.2).



Figuur 1.2 Power BI-architectuur

1.4.4 Power BI Service

De Power BI Service is een website waar je Excel-werkboeken en Power BI Desktop-bestanden naar kan uploaden. Power BI Desktop is de tool waarin je een gegevensset maakt en (een deel van) een rapport. De Power BI Service is waar mensen het gemaakte rapport gebruiken en misschien zelfs aanpassen of uitbreiden.

In de service krijg je de extra mogelijkheid om dashboards te maken. Dashboards zijn een soort startpunt waarop de meest relevante informatie samengebracht kan worden in de vorm van tegels. Door op tegels te klikken, kom je bij het rapport waar die informatie ontsloten wordt.

De Power BI Service is volledig interactief. Gebruikers kunnen nieuwe rapporten maken op bestaande gegevenssets. Gebruikers kunnen ook bestaande rapporten aanpassen. Het meest voorkomende is dat gebruikers de rapporten consumeren (lezen). Daarbij kunnen zij filters aanpassen of toevoegen, grafieken selecteren die daarmee andere grafieken filteren en ga zo maar door. Alles wat een explorer-gebruiker nodig heeft om met de gegevens te ‘spelen’.

Vanuit de Power BI Service kan je ook het verversen van de gegevens regelen. Kijk je tegen oude gegevens aan of is de informatie helemaal up-to-date?

Ten slotte kan je vanuit de service dashboards en rapporten delen met anderen. In hoofdstuk 7 lees je meer over de Power BI Service.

1.5 Power BI-scenario's

Tot nu toe heeft in dit hoofdstuk de verantwoording voor Power BI gelegen bij de Self-Service BI. Er zijn echter vier duidelijke scenario's te definiëren waar Power BI een grote rol kan spelen:

1. Self-Service BI in grotere organisaties;
2. midden- en kleinbedrijf;
3. prototyping;
4. corporate BI.

Self-Service BI in grotere organisaties

De redenen om Self-Service BI te willen in grotere organisaties zijn kort besproken in paragraaf 1.2. Belangrijk om te beseffen is dat Self-Service BI bij grotere organisaties geen vervanging is van klassieke of corporate Business Intelligence. Self-Service BI met Power BI is een aanvulling op de corporate BI-oplossing, die een organisatie meer dan ooit in staat stelt om Business Intelligence niet als een project te zien maar als een proces, een werkwijze, met de intelligente organisatie als ultiem doel.

Midden- en kleinbedrijf

Business Intelligence is voor elk denkbaar bedrijf interessant. Het doel is immers competitief voordeel te verwerven. Welk bedrijf wil dat niet? Het is misschien wel de sleutel tot succes.

Business Intelligence kan echter een heel dure tak van sport zijn. Dure professionals gaan met dure tools datawarehouses bouwen, kubussen bouwen, ETL-procesen programmeren en front-ends maken met rapporten en dashboards. Deze oplossingen vragen vaak nieuwe hardware of extra cloud-abonnementen en als ze eenmaal gebouwd zijn, moeten ze onderhouden worden.

Dat is allemaal leuk en aardig, maar veel bedrijven uit het mkb hebben de benodigde kennis en budgetten niet. Maar bijna alle mkb-bedrijven hebben wel Office. En vanaf Office 2010 is Power BI beschikbaar. Als alternatief kan een bedrijf voor niet al te hoge maandelijkse bedragen Office 365-abonnementen nemen en zo Power BI verkrijgen. Power BI Desktop is zelfs helemaal gratis en ook de Power BI Service heeft een gratis instapabonnement.

Met andere woorden, als 'gewone' Business Intelligence qua kennis of qua budget buiten het bereik van een organisatie ligt, kan je met Power BI de meerwaarde van Business Intelligence toch in huis halen.

Prototyping

Een derde scenario waar Power BI meerwaarde kan hebben is als prototyping-gereedschap voor Business Intelligence-consultants. Een belangrijk onderdeel van Business Intelligence is de informatieanalyse. Informatieanalyse is uitzoeken wat mensen binnen een bedrijf aan informatie nodig hebben om hun werk beter te doen. Dit klinkt niet zo moeilijk, maar kan in de praktijk toch erg lastig zijn.

De hoeveelheid en verscheidenheid van beschikbare gegevens neemt dagelijks toe. De wil en de noodzaak deze gegevens te gebruiken wordt daarmee steeds groter.

Microsoft Power BI is een belangrijk en veelgebruikt hulpmiddel om toepassingsgerichte inzichten te maken en zo meerwaarde te creëren met deze gegevens. Dat kan vanuit Excel of met de gratis Power BI Desktop-applicatie. Je kunt gegevens uit alle denkbare bronnen inlezen en bewerken, onderlinge verbanden van verschillende gegevens definiëren en visualisaties maken. Met dit boek leer je met Power BI te werken en de onderlinge samenhang van de verschillende Power BI-functies te begrijpen. Je leert zo goede selfservice BI-oplossingen te bouwen: met de Power BI-desktopapplicatie en in de cloud met de Power BI-service.

De derde druk sluit aan bij de ontwikkelingen die binnen Power BI hebben plaatsgevonden. Daarbij gaat deze druk niet meer uit van Excel als tool om BI-oplossingen mee te bouwen, maar van Power BI-desktop.

Bij dit boek hoort de website www.handboekpowerbi.nl. Daar vinden studenten en docenten extra materiaal, zoals meerkeuzevragen met feedback en de benodigde bestanden om de opdrachten te maken. Deze elementen zijn voor de derde druk geactualiseerd.

Handboek Power BI is geschreven voor studenten in het hoger onderwijs die een bedrijfskundige studie volgen, of verbonden zijn aan een opleiding binnen de informatica. Het boek is ook geschikt voor professionals die zelf met Power BI aan de slag willen gaan.

Peter ter Braake is zelfstandig SQL Server en Business Intelligence docent/consultant. Hij is auteur van de boeken *Leerboek Business Intelligence* en *Database modelleren*.



9 789024 441419

www.handboekpowerbi.nl
www.boomhogeronderwijs.nl