

Henk Vlootman
Michiel Rozema

Handboek

Power Pivot

DATA-ANALYSE

IN EXCEL

EN POWER BI

[vakmedianet](http://vakmedianet.nl)

HENK VLOOTMAN

MICHIEL ROZEMA

Handboek Power Pivot

DATA-ANALYSE

IN EXCEL EN POWER BI

Alle rechten in deze uitgave zijn voorbehouden aan Vakmedianet. Niets uit deze uitgave mag worden veeelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van kopieën uit deze uitgave is toegestaan op grond van art. 16h t/m 16m Auteurswet jo. Besluit van 27 november 2002, Stb. 575, dient men de daarvoor wettelijk verschuldigde vergoeding te voldoen aan de Stichting Reprorecht te Hoofddorp (Postbus 3060, 2130 KB).

Hoewel aan de totstandkoming van deze uitgave de uiterste zorg is besteed, aanvaarden de auteur(s), redacteur(en) en uitgever(s) geen aansprakelijkheid voor eventuele fouten en onvolkomenheden, noch voor gevolgen hiervan.

Ontwerp omslag en binnenwerk: Justus Bottenheft

ISBN 978 94 6276 128 5

© 2016 Vakmedianet, Deventer

Voorwoord 11

DEEL I **Business intelligence, Excel en Power BI** 13

1 Business intelligence en business analytics 15

- 1.1 De rol van informatie 15
- 1.2 Organisaties sturen met prestatie-indicatoren 15
- 1.3 Het vijflagenmodel voor business intelligence 16
- 1.4 Standaard- en ad-hocrapportages 19
- 1.5 Enterprise en end-user business intelligence 20
- 1.6 Business analytics 22
- 1.7 Data science en 'big data' 23

2 Power BI 25

- 2.1 Onderdelen van Power BI 25
- 2.2 Invoer en voorbereiden: Power Query 25
- 2.3 Analyseren: Power Pivot 27
- 2.4 Uitvoer: Power View 28
- 2.5 Uitvoer: Power Map 31
- 2.6 Het Power BI-cloudplatform 32
- 2.7 Power BI apps 36

3 Excel als analyse-instrument 37

DEEL II **Modelleren in Power Pivot** 49

4 De principes van het relationele datamodel 51

- 4.1 De platte database 51
- 4.2 Nadelen van een platte database 53
- 4.3 Het concept van een relationele database 54
- 4.4 Waarin onderscheidt Power Pivot zich ten opzichte van andere databaseproducten? 56

5 De Power Pivot-werkomgeving 57

- 5.1 Hoe werkt Power Pivot binnen Excel? 57
- 5.2 Power Pivot activeren 57

- 5.3 De tab Power Pivot in het Excel-lint 62
- 5.4 De tab Start in het Power Pivot-lint 64
- 5.5 De tab Ontwerpen in het Power Pivot-lint 77
- 5.6 De tab Geavanceerd in het Power Pivot-lint 83
- 5.7 De tab Gekoppelde tabel in het Power Pivot-lint 87

6 Gegevens importeren 89

- 6.1 Introductie 89
- 6.2 Selectie van gegevens bij import 90
- 6.3 De importmogelijkheden in het lint 92
- 6.4 Importeren uit database 93
- 6.5 Importeren uit gegevensservice 95
- 6.6 Uit een andere bron 97
- 6.7 Uit Power Query 106
- 6.8 Uit een gekoppelde tabel 108
- 6.9 Bestaande verbindingen 110
- 6.10 Vernieuwen 113

7 Tabellen en relaties in Power Pivot 115

- 7.1 Gegevenstypes en notatie 115
- 7.2 Beschikbare gegevenstypes in Power Pivot 118
- 7.3 Hoe gebruik je gegevenstypes in een tabel? 122
- 7.4 Wat is een relatie? 123
- 7.5 Types tabellen binnen een relatie 124
- 7.6 Een speciale filtertabel: de datumtabel 127
- 7.7 Hoe leg je relaties tussen tabellen? 128
- 7.8 Gegevens filteren met relaties 134

8 Uitvoer 137

- 8.1 Draaitabel 137
- 8.2 Draaigrafiek 167
- 8.3 Power View (Excel 2013) 176
- 8.4 Power Map (Excel 2013)/3D maps (Excel 2016) 184

DEEL III Berekeningen met DAX 187

9 Berekeningen met DAX 189

- 9.1 Berekeningen in Power Pivot 189
- 9.2 Context 192
- 9.3 Berekende kolommen 195
- 9.4 Metingen 196

10	Basis DAX-functies	201
10.1	Operatoren, bewerkingen en constante functies	201
10.2	Basisaggregaties	207
10.3	Logische functies	208
10.4	Foutafhandeling	210
10.5	Opzoeken en verwijzen	211
11	Filters	215
11.1	Filtercontext	215
11.2	De functie CALCULATE	216
11.3	ALL-functies	222
11.4	Actieve en niet-actieve relaties	230
12	Tabelfuncties	233
12.1	Tabelargumenten en tabelresultaten	233
12.2	Basisfuncties	235
12.3	Functies voor berekende tabellen	236
12.4	Filteren met tabelfuncties	242
12.5	Nogmaals: CALCULATE - Rijcontext en filtercontext	244
12.6	De ins en outs van filteren met tabellen	249
13	Intelligente tijdfuncties	255
13.1	De datumtabel	255
13.2	Het berekenen van year-to-date totalen	255
13.3	Intelligente tijdfuncties	259
13.4	Filteren met tijdfuncties	261
14	Geavanceerde DAX-functies	271
14.1	Informatiefuncties	271
14.2	Geavanceerde tabelfuncties	274
14.3	Het gebruik van KEEPFILTERS	281
14.4	Variabelen in DAX	283
DEEL IV	Geavanceerde scenario's in DAX	287
15	Stuurtabellen	289
15.1	Het concept van stuurtabellen	289
15.2	Stuurtabellen en tabelfuncties	293
15.3	Hoe maak je een stuurtablel?	294

16	Voorraad- en omzetberekeningen	297
16.1	Overzicht van het model	297
16.2	Basisberekeningen	298
16.3	Vergelijkingen	301
16.4	Tijdgerelateerde berekeningen	305
16.5	Actuele omzet en targets	308
16.6	Uitstaande betalingen	310
16.7	Voorraadstatus	316
17	Forecasting en toekomstige waarde	319
18	Werken met alternatieve kalenders	329
18.1	Overzicht van het model	329
18.2	Weeknummers volgens de ISO-standaard	330
18.3	Year-to-date en Period-to-date	332
18.4	Groecijfers	336
18.5	Voortschrijdend gemiddelde	340
19	Pareto-analyse en watervalgrafieken	345
19.1	Pareto-analyse	345
19.2	Datamodel voor Pareto-analyse	347
19.3	DAX-functies voor de Pareto-grafiek	348
19.4	Watervalgrafieken	354
19.5	Datamodel voor een watervalgrafiek	355
19.6	DAX-functies voor een watervalgrafiek	356

DEEL V **Professionele Power Pivot-modellen** 361

20	Metadata	363
20.1	Hiërarchieën	363
20.2	Key Performance Indicators of KPI's	374
20.3	Perspectieven	378
21	Power Pivot onder de motorkap	381
21.1	DAX Query	381
21.2	Data management views	385
22	Snelle modellen	393
22.1	Omvang van het model	393
22.2	Berekende kolommen	394
22.3	Relaties	395
22.4	Slicers en tijdlijnen	395

22.5	Draaitabellen	397
22.6	DAX-berekeningen	398
22.7	Alternatieven voor IF	400

	Trefwoorden register	407
--	----------------------	-----

	Over de auteurs	409
--	-----------------	-----

Voor je ligt het Handboek Power Pivot. Gefeliciteerd met deze aankoop! Power Pivot is zonder twijfel de meest krachtige en waardevolle toevoeging aan Excel van de laatste tien jaar. Wij zijn al zeven jaar heel enthousiast over Power Pivot en we kunnen niet wachten om je er kennis mee te laten maken.

Power Pivot helpt je om krachtige modellen te bouwen, die veel verder gaan dan wat je in Excel ooit zult bereiken. Modellen waarmee je grote aantallen gegevens kunt combineren en analyseren, zodat je waardevolle inzichten verkrijgt. Ben je Excel-gebruiker, dan zul je merken dat je met Power Pivot veel sneller veel betere resultaten oplevert. Bovendien is kennis van Power Pivot ook de basis voor werken met Power BI, het populaire nieuwe platform voor Business Intelligence van Microsoft.

Dit is ons tweede boek over Power Pivot. Het eerste hebben we geschreven omdat we een behoefte zagen aan goede, Nederlandstalige informatie over het onderwerp. In de tussentijd is Power Pivot alleen maar populairder geworden en hebben wij zelf heel veel modellen ontwikkeld. Onze ervaringen hebben we in dit boek verwerkt; daarom gaat het boek niet alleen over hoe Power Pivot werkt, maar geven we ook veel tips hoe je goede modellen maakt.

■ Onze tips en ‘best practices’ herken je aan een blauwe streep in de kantlijn.

We hebben ervoor gekozen om vooral in te gaan op Power Pivot in Excel 2016. Maar ook als je met Excel 2010 of Excel 2013 werkt, kun je de technieken in dit boek gebruiken. Dit geldt ook als je gebruikmaakt van het programma Power BI Desktop. En zelfs BI-professionals die met SQL Server werken, hebben baat bij dit boek en met name alle hoofdstukken over DAX.

Tekst in dit kader geeft informatie over verschillen tussen de versies van Power Pivot.

Om je in staat te stellen om zelf de besproken technieken uit te proberen, is een groot deel van de modellen die we in dit boek gebruiken beschikbaar via de website van Vakmedianet. Ga hiervoor naar handboekpowerpivot.nl. We wensen je veel leesplezier en vooral veel succes met Power Pivot!

HENK VLOOTMAN
MICHIEL ROZEMA

Business intelligence, Excel en Power BI

Excel is een van de meest gebruikte softwareproducten ter wereld. Er zijn maar weinig kantoormedewerkers die nog nooit een Exceldocument in handen hebben gehad. Excel wordt op veel manieren gebruikt: van eenvoudig lijstjes bijhouden tot complete applicaties voor bijvoorbeeld plannings en administratie.

Maar waar Excel misschien wel het meest voor wordt gebruikt, is het verkrijgen van inzicht in informatie. Daarmee is Excel een business intelligence instrument: je gebruikt het om informatie te analyseren en om de resultaten daarvan te gebruiken om je organisatie verder te helpen.

Excel heeft echter een aantal nadelen in het gebruik als business intelligence (BI) oplossing en Power BI is Microsofts antwoord op die nadelen.

In het eerste deel van dit boek schetsen we een aantal grote lijnen die je duidelijk maken hoe Power Pivot past in een BI-oplossing en waar je Power Pivot wel en niet voor gebruikt. Hoofdstuk 1 introduceert en verklaart een aantal begrippen en concepten in BI en legt uit waarom BI een steeds belangrijker onderwerp wordt.

In hoofdstuk 2 geven we een overzicht van het Power BI-platform en welke plaats Power Pivot daarin inneemt.

Hoofdstuk 3 is gewijd aan het gebruik van Excel. We zetten uiteen hoe een professioneel model in Excel eruitziet en hoe de verhouding tussen analyses en applicaties in Excel is.

Business intelligence en business analytics

1.1 De rol van informatie

Informatie is vandaag de dag ontegenzeggelijk een van de belangrijkste bezittingen van een organisatie. Als consument kennen we het verschijnsel allemaal: bedrijven staan in de rij om, tegen een kleine tegenprestatie, onze persoonlijke gegevens te mogen ontvangen. De reden is dat het beschikken over grote hoeveelheden klantgegevens voor bedrijven van onschatbare waarde is om effectief te adverteren, de juiste producten en diensten op de markt te brengen en het gewenste imago te creëren in de markt.

Maar niet alleen voor commerciële bedrijven is informatie van levensbelang. Ook binnen de overheid en het onderwijs wordt informatie gebruikt om de primaire processen efficiënter in te richten. Of het nu gaat om het besparen van energie door verlichting uit te zetten wanneer dat kan, het aansturen van verkeerslichten om een optimale doorstroming van het verkeer te waarborgen, of het vroegtijdig signaleren van leerlingen die dreigen uit te vallen, overal is informatie de basis voor vooruitgang en innovatie.

Maar informatie ontstaat niet vanzelf. Het begint met bijhouden van *gegevens*, of *data*. Maar data alleen levert niets op: echte waarde ontstaat door het combineren van gegevens, het plaatsen van gegevens in context en het vinden van structuur en verbanden in de data. Dat is waar business intelligence en business analytics over gaan.

1.2 Organisaties sturen met prestatie-indicatoren

Om als organisatie effectief te kunnen functioneren, moet ze duidelijk kunnen vaststellen of de juiste dingen gedaan worden. Daarvoor moet je wel weten wat de juiste dingen dan zijn. Business intelligence begint

daarom vaak met nadenken over de strategie van de organisatie: wat zijn de doelstellingen van de organisatie, hoe denkt de organisatie die doelen te bereiken en welke processen dienen daarvoor te worden uitgevoerd?

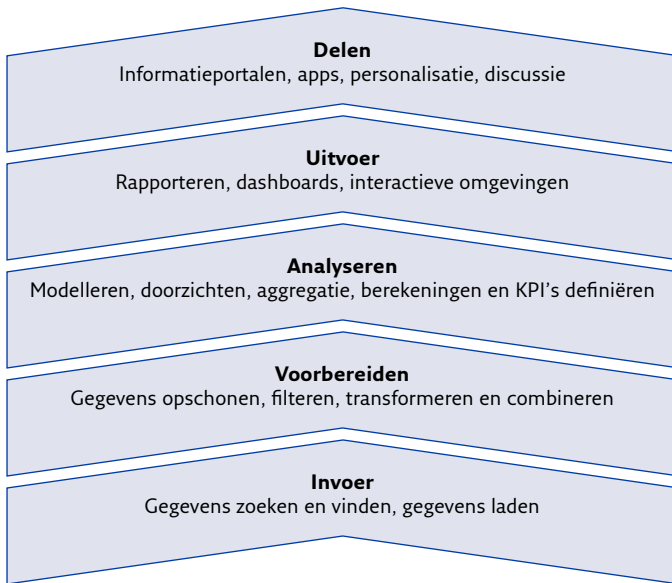
Als duidelijk is welke processen voor de organisatie belangrijk zijn, kun je de processen zo inrichten dat je op verschillende punten in een proces kunt meten wat de voortgang is in het proces. Een onderwijsinstelling bijvoorbeeld, die als doel heeft studenten op te leiden voor een diploma, heeft in het algemeen als belangrijkste ‘bedrijfsproces’ het geven van onderwijs. Op verschillende punten in dat proces kan de voortgang gemeten worden, bijvoorbeeld door vakken af te sluiten met toetsen waarmee studenten laten zien dat ze de gewenste kennis hebben opgedaan. Voor een individuele student is dat al een duidelijke indicator van zijn of haar voortgang richting een diploma, maar voor de instelling is de waarde nog groter: het geeft niet alleen de voortgang van de studenten weer, maar levert ook inzicht in de punten waarop het proces verbetering behoeft. Als veel studenten in het tweede jaar van de opleiding vertraging oplopen, is het zinvol om na te gaan of het onderwijsproces op dat punt aangepast moet worden.

Het meten van de voortgang van bedrijfsprocessen werkt met prestatie-indicatoren die inzicht geven in hoe goed het proces verloopt. Ook hier geldt weer dat ‘goed’ gedefinieerd moet zijn, ofwel: alleen meten is niet genoeg, je hebt ook een doelwaarde nodig waar je een meting tegenaan kunt houden. In ons onderwijsvoorbeeld: de doelwaarde voor een eindcijfer is minimaal 6 en op een hoger niveau zou de indicator ‘percentage studenten dat een vak haalt binnen een half jaar’ als doelwaarde 90 procent kunnen zijn.

Er zijn vele prestatie-indicatoren denkbaar en in de praktijk is het ondoenlijk om ze allemaal continu in de gaten te houden. Daarom beperk je je in het algemeen tot de belangrijkste, of in het Engels: de key performance indicators, ofwel KPI's. De KPI's binnen een organisatie kunnen op elk niveau anders zijn, hoewel ze op verschillende niveaus natuurlijk wel met elkaar in lijn moeten zijn. Zo zal de directie de organisatie sturen op geaggregeerde en strategische KPI's en zullen op lagere niveaus meer gedetailleerde en operationele KPI's gebruikt worden.

1.3 Het vijflagenmodel voor business intelligence

Om op een gestructureerde manier te kunnen praten over een oplossing voor business intelligence hebben we een model ontwikkeld met vijf lagen, die op hoofdlijnen beschrijven hoe een BI-oplossing is opgebouwd. Het vijflagenmodel is hierna weergegeven.

AFBEELDING 1.1 *Het vijflagenmodel voor business intelligence*

De onderste laag, **Invoer**, is natuurlijk de basis voor een BI-oplossing: als je gegevens wilt analyseren, moeten die gegevens in het BI-systeem ingevoerd worden. Voor een eenvoudig Excel-model voer je vaak met de hand gegevens in, maar meestal zijn de gegevens al in een systeem opgeslagen. Dit kan een database zijn, maar ook Excel-bestanden, tekst- of andere bestanden. En ook kan het zijn dat de informatie die je nodig hebt, op een website te vinden is. Er is een enorme diversiteit in hoe gegevens beschikbaar zijn en elke bron van data heeft zijn eigen manier om de gegevens op te vragen.

Als je gegevens gaat analyseren, kun je er niet vanuit gaan dat alle benodigde data ook in de goede vorm beschikbaar zijn om analyses te kunnen uitvoeren. Dat geldt des te meer wanneer je meerdere databronnen gebruikt. Het is daarom nodig om de data voor te bereiden voor analyse. Dat is wat in het model in de laag **Voorbereiden** gebeurt. Hier wordt 'ruwe data' getransformeerd tot een verzameling gegevens die geanalyseerd kan worden.

In de praktijk zit er veel werk in deze laag. Met name bij het gebruik van verschillende databronnen is niet altijd duidelijk welke gegevens bij elkaar horen. Veel bedrijfsapplicaties gebruiken nummers om verschillende entiteiten te identificeren, bijvoorbeeld een klantnummer in een verkoopsysteem. Maar het komt veel voor dat verschillende systemen andere nummers gebruiken voor dezelfde klant. Of er wordt wel hetzelfde nummer gebruikt,

maar andere informatie over de klant is niet hetzelfde, bijvoorbeeld het telefoonnummer. Het is een hele klus om deze gegevens dan bij elkaar te brengen. In feite is dit specialistenwerk en is er speciale software voor het combineren en opschonen van data.

Een andere functie in de laag Voorbereiden is het filteren van data. Zo zijn in de bronsystemen misschien gegevens van vele jaren beschikbaar, maar willen we in een analyse alleen naar de laatste twee of drie jaar kijken. Of we willen uit een financieel systeem wel alle winst- en verliesposten uit het grootboek, maar niet de balansposten.

Ten slotte kan het bij het voorbereiden van data gaan om een verandering van de structuur van de data. Wat vooral in Excel-bestanden veel voorkomt, is dat gelijksoortige gegevens in verschillende kolommen naast elkaar staan: bijvoorbeeld een kolom voor elke maand. Veel analysetools, en ook de tools waar dit boek over gaat, werken beter met een tabelstructuur waar het aantal kolommen vast is. In dit geval wil je dan een transformatie op de gegevens uitvoeren, waarin een tabel wordt gemaakt met een kolom die aangeeft over welke maand het gaat en een kolom met de waarde zelf:

AFBEELDING 1.2 *Van kolommen naar rijen*

Maanden	Waarde
Januari	457
Februari	913
Maart	859
April	507
Mei	196
Juni	549
Juli	418
Augustus	844
September	913
Oktober	377
November	433
December	746

Als je gegevens eenmaal in een werkbare vorm staan, kun je met het eigenlijke **Analyseren** beginnen. Dit is de kern van elke BI-oplossing. Je bepaalt hier welke doorsnedes je wilt maken van je data, bijvoorbeeld per maand en jaar, per klantsegment, of welke doorsnede voor jouw organisatie maar relevant is. Een andere factor is welke aggregaties over de gegevens nodig zijn: veel rapporten worden gevormd door de gegevens op een bepaalde manier samen te vatten. Door allerlei berekeningen toe te voegen wordt het mogelijk nog meer inzichten uit de gegevens te halen.

Al deze vormen van analyse worden uitgevoerd door het samenstellen van een analysemodel, waarin doorsnedes, aggregaties, berekeningen en nog

meer worden opgenomen. Een goed analysemodel maakt het heel eenvoudig om duidelijke en inzichtelijke rapporten en dashboards te maken.

De laag **Uitvoer** heeft betrekking op deze rapporten en dashboards. Op basis van een analysemodel worden de resultaten van het model zichtbaar gemaakt. Deze uitvoer kan allerlei vormen aannemen, van een eenvoudig lijstje cijfers in Excel, via interactieve grafische overzichten, tot professioneel opgemaakte rapporten die volautomatisch worden gecreëerd. De laatste jaren zijn ook infographics populair: hierin worden resultaten op een heel visuele manier weergegeven, waarbij de manier van visualisatie nauw aansluit bij de gegevens zelf. Een voorbeeld is wanneer de zetelverdeling van partijen in de Tweede Kamer wordt weergegeven door op een plattegrond van de Tweede Kamer de zetels de kleur van een partij te geven.

De laatste laag draait om het **Delen** van de uitvoer. In de laag delen denk je na over hoe de resultaten van de analyse bij de juiste personen terechtkomen. Een voor de hand liggende manier is om die personen een mail te sturen, wat bij gebruik van Excel veel wordt gedaan. Dit is echter een slechte keuze: elke mail produceert een kopie van het Excel-bestand, waardoor in no-time heel veel versies van hetzelfde bestand ontstaan. Een betere optie is gebruik te maken van een informatieportaal, waarop de rapporten en dashboards beschikbaar zijn voor iedereen die daar toegang toe moet hebben. Daarnaast zijn mobiele apps sterk in opkomst.

In een goede BI-oplossing zijn de vijf lagen duidelijk onderscheiden en worden benodigde bewerkingen op data in de juiste laag uitgevoerd. Dit heeft veel voordelen, zoals het voorkomen dat dezelfde bewerking op data, of dezelfde berekening meerdere keren wordt geïmplementeerd. Dit zorgt ervoor dat het gemakkelijker is om veranderingen in bijvoorbeeld bronsystemen te verwerken in de BI-oplossing. In een slecht opgebouwde BI-oplossing kan een verandering in een bronsysteem op allerlei plekken wijzigingen noodzakelijk maken, tot in de uitvoer aan toe.

1.4 Standaard- en ad-hocrapportages

Wanneer een organisatie gestuurd wordt op basis van KPI's, kunnen deze in een rapport of dashboard worden opgenomen, dat op elk moment geraadpleegd kan worden. Deze vorm van rapportage is in hoge mate gestandaardiseerd: er is goed nagedacht over de strategie, de bedrijfsprocessen, het meten van de prestatie-indicatoren en de wijze waarop deze gerapporteerd worden. Een geautomatiseerde rapportage kan in dit geval voor elke betrokkene in de organisatie dezelfde informatie weergeven, en deze rapportage is relatief stabiel en zal niet vaak veranderen. Meer en meer dringt

business intelligence echter ook door tot het dagelijkse werk. Niet alleen strategische beslissingen rond bijvoorbeeld het veranderen van bedrijfsprocessen, of zelfs het veranderen van bedrijfsstrategie, worden op basis van informatie genomen. Steeds meer dringt het besef door dat in feite elke beslissing, of deze nu strategische impact heeft of heel operationeel is, gebaseerd zou moeten zijn op concrete informatie. Dit betekent ook dat er behoefte ontstaat aan een dynamischer vorm van analyse, waarmee ad-hocvraagstukken snel beantwoord kunnen worden.

Tegen deze achtergrond is er sinds enkele jaren veel aandacht voor zogenaamde 'Self-Service BI'. De aanleiding hiervoor is dat traditionele BI-platforms slechts een klein deel van de medewerkers van de organisatie weten te bereiken. Hierdoor moeten de meeste gebruikers zelf in hun informatiebehoefte voorzien, waarvoor ze vaak Microsoft Excel gebruiken. Deze situatie komt nog steeds in vrijwel elke organisatie voor.

De oorspronkelijke belofte van Self-Service BI was dat medewerkers van een organisatie zelf hun BI-oplossingen zouden kunnen maken, zonder dat ze daar de IT-afdeling voor nodig hebben. Uit het vijflagenmodel wordt al duidelijk dat dit slechts zeer beperkt realiseerbaar is: wie wil verantwoordelijk zijn voor zowel het ontsluiten van brondata, het voorbereiden daarvan, het maken van een analysemodel, het vormgeven van de uitvoer en ook nog het op een goede manier delen van de resultaten? Vooral in de onderste twee lagen is het erg prettig als er binnen de organisatie gewoon 'nette' data beschikbaar zijn.

Beter is het daarom om onderscheid te maken tussen twee soorten BI die elkaar aanvullen en versterken. We hanteren hiervoor de termen Enterprise BI en End-user BI.

1.5 Enterprise en end-user business intelligence

Enterprise BI

Rapporten met de key performance indicators van een organisatie worden met veel zorg opgebouwd. En terecht natuurlijk: de prestaties van de organisatie en zelfs het voortbestaan hangen ervan af. Deze rapporten worden binnen professioneel opgezette projecten gerealiseerd. Dit is het domein van zogenaamde Enterprise business intelligence. Deze vorm van BI maakt gebruik van zeer krachtige – en dure – server platforms, grote databases en

rapportageplatforms, die grote hoeveelheden gegevens kunnen verwerken en duizenden gebruikers kunnen bedienen.

Enterprise BI is het werkterrein van de IT-afdeling, die oplossingen ontwikkelt in professioneel opgezette projecten. Hierin is veel aandacht voor kwaliteit van de data, voor beschikbaarheid van het platform en voor een gestructureerd en beheersbaar proces voor zowel ontwikkeling als beheer.

Enterprise BI is een vakgebied met een eigen jargon, eigen methodieken en concepten en natuurlijk ook eigen tools. In gesprek met een Enterprise BI-specialist kunnen termen vallen als ETL-processen, datawarehouses, OLAP, cubes, sterschema's en nog meer. Het is ons opgevallen dat in veel publicaties over Power BI deze begrippen ook gebruikt worden; dit heeft ongetwijfeld te maken met de Enterprise BI-achtergrond van de betreffende auteurs. Wij proberen in dit boek deze begrippen zoveel mogelijk te vermijden, omdat dit boek niet primair gericht is op BI-professionals én omdat Power Pivot vooral een End-user BI-product is.

In het Microsoft-platform wordt Enterprise BI voornamelijk ingevuld met onderdelen van het SQL Server-dataplatform.

End-user BI

Voor ad-hocanalyses worden in het algemeen geen grote projecten ingericht, omdat snelheid van groot belang is. Deze vorm van BI wordt niet door IT-ers gerealiseerd, maar door personen die dicht op de bedrijfsprocessen zitten, zoals een businessanalist, een controller of een marketingspecialist. Dit is het domein van End-user BI.

De bouwer van End-user BI-oplossingen maakt gebruik van andere instrumenten dan de IT-er, en vaak is dat Excel. Vaak worden de oplossingen echter zo complex dat de bouwer veel tijd kwijt is aan het up-to-date houden van de rapporten en aan het beheer en de verdere ontwikkeling van de oplossing. Bovendien raken steeds meer Excel-gebruikers aan de grenzen van wat Excel nog aankan.

Microsoft pakt deze problemen aan door de mogelijkheden van Excel uit te breiden, zodat het een echt End-user BI-instrument wordt. Excel wordt hiermee onderdeel van een compleet End-user BI-platform: Power BI.

Beide soorten BI versterken elkaar doordat op elke laag gekozen kan worden voor het gebruik van de laag daaronder in óf de End-user BI-, óf de

Enterprise BI-variant. Ook is het mogelijk om rapporten en analyses die in de End-user BI-omgeving zijn gemaakt, over te brengen naar de Enterprise BI-omgeving als daar behoefte aan is.

1.6 Business analytics

Business intelligence is gericht op het verkrijgen van inzicht op basis van wat er gebeurd is. Wanneer je business intelligence goed toepast, weet je wat er gebeurd is en ben je ook in staat om je een beeld te vormen waarom dat zo is. De reden voor de wens om dit inzicht te hebben, is dat je zo kunt vaststellen wat je vooral moet blijven doen en wat je voortaan anders moet gaan doen. Een sportwinkel kan bijvoorbeeld meten wat de omzet is van verschillende sportartikelen en daarbij ook bijhouden op welke plek in de winkel deze artikelen staan. Hierbij kan blijken dat tennisballen die bij de ingang van de winkel liggen, minder goed verkocht worden dan tennisballen die dicht bij de kassa liggen. Een mogelijke conclusie kan zijn dat alle tennisballen in het assortiment bij de kassa moeten liggen.

Op dit eenvoudige voorbeeld is natuurlijk van alles af te dingen: is er prijsverschil tussen de verschillende soorten tennisballen? Is het merk verschillend en zijn de goed verkochte tennisballen van het merk waar de nummer één van de wereld mee speelt? Of is de kleur misschien anders? Er kunnen vele kenmerken van invloed zijn op de omzet van een artikel en het is moeilijk te voorspellen wat het effect zal zijn als je een artikel op een andere plek in de winkel aanbiedt. Met een business intelligence-oplossing kun je in feite alleen maar uitproberen wat er gebeurt als je iets verandert en dat goed meten, zodat je er een analyse op kan doen.

Het vakgebied *business analytics* is erop gericht om niet het verleden goed weer te geven, maar juist voorspellingen te doen over wat er in de toekomst zal gebeuren. In het voorbeeld van de sportzaak zou een business analytics-oplossing kunnen voorspellen hoeveel artikelen van een bepaalde soort en merk de komende twee weken zullen worden verkocht. Deze voorspelling kan gebruikt worden om de juiste voorraad in te kopen, wat een zeer positief effect heeft op de financiële resultaten.

Om een ander voorbeeld te gebruiken: een routeplanner die een business analytics-oplossing gebruikt, zou de reistijd kunnen voorspellen op een traject, waarbij gebruikgemaakt wordt van de actuele verkeerssituatie, maar ook van historische gegevens over het spitsverloop op dit traject. Vooraf weten hoe lang je over je reis gaat doen, is erg prettig. Beter nog is natuurlijk als de routeplanner je vertelt hoe je moet rijden om zo snel

mogelijk op je bestemming aan te komen. In feite zijn dit twee vormen van business analytics: de eerste vertelt je wat er gaat gebeuren (de reistijd) en wordt in het Engels wel 'predictive analytics' genoemd. De tweede vertelt je wat je moet doen om je doel zo goed mogelijk te bereiken (de snelste route) en wordt wel 'prescriptive analytics' genoemd.

Business analytics maakt gebruik van beschikbare gegevens uit het verleden en extrapoleert deze via complexe wiskundige modellen naar de toekomst. Het probleem van de toekomst voorspellen is echter dat het erg moeilijk is. Dat komt doordat er heel veel factoren van invloed kunnen zijn op die toekomst. En je weet niet eens welke factoren relevant zijn. Bijvoorbeeld: is het weer van invloed op de reistijd? Dat zou best kunnen: als het regent of het vriest, is te verwachten dat je reistijd langer zal zijn. Maar heeft de voetbalkalender ook invloed? Dat zou best kunnen. En de beurskoersen dan?

De enige manier om uit te maken welke gegevens nodig zijn om goede voorspellingen te doen, is alle gegevens waar je over kunt beschikken, mee te nemen in de analyse en het wiskundige model te laten vertellen welke gegevens ertoe doen en welke niet. De consequentie hiervan is dat er behoefte ontstaat om heel veel gegevens te bewaren en te analyseren.

1.7 Data science en 'big data'

Het opslaan, beheren en analyseren van enorme hoeveelheden informatie wordt aangeduid met de term 'big data'. Big data is eigenlijk een modewoord en wordt door zo'n beetje elke technologieleverancier gebruikt om nieuwe producten aan de man te brengen.

Data science is het (wetenschappelijke) vakgebied dat draait om het analyseren van enorme hoeveelheden gegevens. De term data science wordt de laatste paar jaar steeds meer gebruikt, maar de wetenschappelijke technieken zijn al veel ouder. Er is een aantal oorzaken voor de toenemende populariteit van data science.

Allereerst dalen de kosten voor het opslaan van data steeds meer. We kunnen dat zelf vaststellen door te kijken naar de prijs en opslagcapaciteit van een USB-stick of een SD-kaart: anno 2016 koop je voor een paar tientjes een kaart met 128 GB opslagcapaciteit. Zo'n twintig jaar geleden kostte een schijfje met 120 MB capaciteit rond 75 gulden. Ofwel: de prijs per eenheid is in twintig jaar tijd met een factor 500 gedaald. Dit betekent dat het steeds meer haalbaar wordt om veel gegevens te bewaren.

Een andere reden is dat de technologie om ook echt iets met al die informatie te doen, ook steeds toegankelijker wordt. Microsoft speelt hierin een belangrijke rol en maakt zelfs van Excel een 'big data tool'. Niet alleen kan Excel tegenwoordig miljoenen rijen gegevens opslaan en analyseren (met Power Pivot!), maar in Excel 2016 zijn ook al eenvoudige functies voor predictive analytics ingebouwd. Daarmee wordt data science minder het domein van zeldzame specialisten in bedrijven en instellingen die over de middelen beschikken om enorme investeringen te doen in tijd en geld voor data science.

De belangrijkste reden voor de aandacht voor data science is wel dat dit het vakgebied is dat de wiskundige modellen levert om voorspellingen te doen op basis van data. Hiermee helpt data science ons om effectiever en efficiënter gebruik te maken van beschikbare middelen, of dat nu geld is, tijd, of andere middelen.

Het is nog nauwelijks te overzien welke impact data science gaat hebben op onze maatschappij. De ontwikkelingen zijn al overal te zien: handschriftherkenning, in real-time vertalen van gesprekken tussen twee talen, detecteren van frauduleuze creditcardtransacties, zelfrijdende auto's: vrijwel elke innovatie bestaat momenteel uit het slim gebruiken van data, of is ontwikkeld met behulp van data science. Er zouden hele boeken over dit onderwerp geschreven kunnen worden (en dat gebeurt ook volop) maar het is niet het onderwerp van dit boek. Maar met Excel en Power Pivot heb je wel tools in handen die volop in deze ontwikkelingen meegaan.

Power Pivot is een uitbreiding op Microsoft Excel, een van de meest gebruikte computerprogramma's. Sinds Excel 2010 betekent Power Pivot een fundamentele verandering en verrijking van de mogelijkheden van Excel voor het analyseren van gegevens. Power Pivot is hét gereedschap voor Business Intelligence voor Excel-gebruikers, maar vormt ook het hart van Microsoft Power BI.

Handboek Power Pivot is het tweede boek van de auteurs over dit onderwerp. Zij putten uit hun ervaring in het werken met Power Pivot om nog dieper in te gaan op het analyseren van grote aantallen gegevens. Dit boek is geschikt voor zowel de Excel-gebruiker die wil leren hoe Power Pivot kan worden gebruikt, als de gevorderde Power Pivot-gebruiker die meer uit het product wil halen. In begrijpelijke taal en met praktische voorbeelden helpt *Handboek Power Pivot* je om het uiterste uit je gegevens te halen.

