

PowerPivot: doe-het-zelf oplossing voor de BI

MEER STURING GEVEN AAN INFORMATIEPRODUCTEN

Kasper de Jonge

Met de release van SQL Server 2008 R2 in mei 2010 komt Microsoft met een nieuwe versie van SQL Server met veel BI-updates en als hoogtepunt 'Managed self-service business intelligence' in de vorm van PowerPivot. PowerPivot is een gecombineerde totaaloplossing van SQL Server, SharePoint en Excel, waarmee power users en eindgebruikers zelf oplossingen op informatievraagstukken kunnen vinden, samenstellen en delen.

In veel organisaties is een corporate BI-omgeving ingericht om gefundeerd beslissingen te kunnen nemen op basis van gegevens uit de organisatie. Deze informatievoorziening wordt verzorgd met behulp van datawarehouses, datamarts, OLAP cubes, rapporten en dashboards. Ze worden onderhouden en ontwikkeld door een kleine groep specialistisch IT/BI personeel. De IT-afdeling heeft beperkte tijd en resources en kunnen lang niet alle gewenste informatievragen beantwoorden. Zo goed als alle beschikbare resources worden besteed aan het onderhouden en ontwikkelen van de meest essentiële informatie voor de ondersteuning van de primaire processen van een bedrijf.

Wat gebeurt er met de informatievraag die door tijdgebrek of kostenoverweging niet door de IT/BI-afdeling kan worden beantwoord? Eindgebruikers gaan het zelf zonder tussenkomst of inbreng van de IT-afdeling oplossen, met in de meeste gevallen Excel. Deze Excel-bestanden gaan een eigen leven leiden in de organisatie en worden verspreid via file shares, mail of usb sticks. De organisatie heeft geen sturing over deze informatieproducten, die misschien wel gebruikt worden in belangrijke beslissingstrajecten.

PowerPivot Client

Microsoft lost dit op door eindgebruikers een goed alternatief te bieden met behulp van PowerPivot Client en PowerPivot voor SharePoint. Gebruikers die bekend zijn met

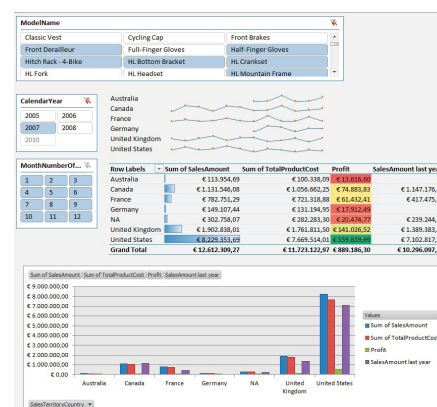
Excel kunnen met behulp van een gratis te downloaden add-in (te vinden op <http://www.powerpivot.com>) gebruik maken van de functies en visualisatiemogelijkheden van Excel uitgebreid met de krachtige functies en mogelijkheden van PowerPivot. In Excel zijn nieuwe mogelijkheden toegevoegd zoals slicers. Met slicers kun je met een druk op de knop een dwarsdoorsnede van je data maken, erg krachtig in combinatie met data uit PowerPivot. Met de komst van PowerPivot worden Excel pivot tables niet meer gebaseerd op cellen maar op tabellen zoals in een database. Gegevens uit verschillende bronnen, van bijvoorbeeld SQL Server, Azure, Oracle, Excel en Atom feeds, kunnen via een simpele interface worden ingelezen en gerelateerd. De gegevens die worden ingelezen, worden fysiek naar PowerPivot overgehaald. Wanneer er nieuwe gegevens beschikbaar zijn in de bron zal de PowerPivot gebruiker zelf zijn data moeten verversen. Door gebruik te maken van een in-memory engine (genaamd de VertiPaqengine) in Analysis service wordt het mogelijk om gegevens te analyseren en te transformeren op het werkstation van de eindgebruiker. Om miljoenen rijen aan data te kunnen gebruiken op een werkstation is gebruik gemaakt van kolom compressie die ervoor zorgt dat efficiënt gebruik kan worden gemaakt van intern geheugen. De engine is onder water een aangepaste versie van analysis services, die wordt aangeroepen vanuit de PowerPivot add-in in Excel.

Id	Jaar	Maand	Verkoper	Aantal
1	2010	Januari	Kees	15
2	2010	Januari	Jan	20
3	2010	Februari	Vera	30
4	2010	Februari	Kees	25

MET KOLOM COMPRESSIE WORDT PER KOLOM GEKEKEN OF ER DUBBELE WAARDEN VOORKOMEN, DUBBELE VELDEN WORDEN VERVANGEN DOOR EEN VERWIJZING EN DE WAARDE WORDT EENMAAL OPGESLAGEN. IN DIT VOORBEELD KUNT U DUIDELIJK ZIEN DAT WE PER KOLOM MEER DUBBELE WAARDEN KUNNEN VINDEN DAN DAT WE PER RIJ Zouden DOEN.

Data Analysis Expressions

Een van de krachtigste mogelijkheden van PowerPivot is de nieuwe functietaal genaamd Data Analysis Expressions, afgekort DAX. Met DAX kunnen dynamische aggregaties gemaakt worden op basis van



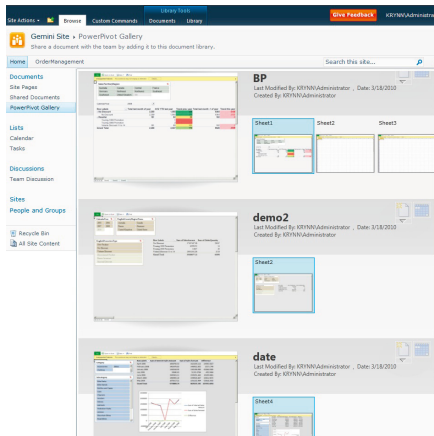
relationele database concepten en een Excel-achtige syntax. Door gebruik te maken van onder meer tijdsintelligente-, aggregatie-, tabel- en filterfuncties kunnen berekende kolommen of meetwaarden worden toegevoegd aan de bestaande data.

Enkele handige beschikbare functies in DAX zijn:

TOTALYD:	Bereken de meegegeven meetwaarden van huidig jaar tot nu
DATEADD:	Bereken de meetwaarden van een verschillende periode, zoals vorige week, vorig jaar of volgende maand
SUMX:	Bereken de som van meetwaarden geëvalueerd voor iedere regel in een tabel
RELATED:	Haalt een waarde op uit een andere tabel
FILTER:	Geeft een tabel terug gefilterd over specifieke waarden

PowerPivot Server

Gebruikers kunnen de PowerPivot bestanden gemakkelijk delen door ze te publiceren naar de SharePoint 2010 PowerPivot gallery, waar ze veranderen in een interactieve webapplicatie. De PowerPivot Gallery is gebaseerd op Silverlight en toont een thumbnail van de PowerPivot werkboeken die in de gallery zijn gepubliceerd.

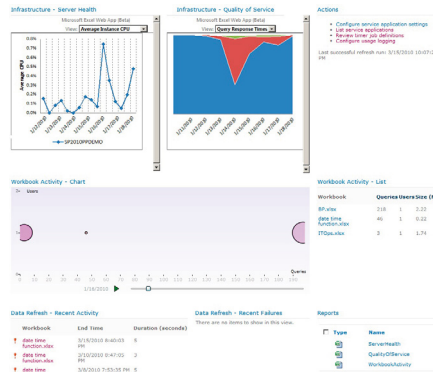


Wanneer gebruikers een werkboek selecteren wordt deze geopend in Excel Services, dit zorgt ervoor dat het PowerPivot werkboek in de browser beschikbaar wordt, gerenderd in HTML met behoud van alle interactieve functies. Er is dus geen lokale installatie van Excel of PowerPivot nodig om PowerPivot bestanden in SharePoint te kunnen bekijken.

IT Management Dashboard

De IT-afdeling kan met behulp van de Po-

werPivot Management Dashboard meer inzicht krijgen in het gebruik van PowerPivot applicaties in de organisatie. Het dashboard geeft inzicht in hoe en wanneer de gebruikers gebruik maken van PowerPivot. Om tussen alle PowerPivot applicaties er die ene PowerPivot applicatie uit te halen, die bedrijfskritisch aan het worden is, kan gebruik worden gemaakt van de Workbookactivity-chart. De meeste PowerPivot applicaties



zullen voor een korte periode (ad hoc) gebruikt worden maar op het moment dat een PowerPivot applicatie structureel gebruikt wordt door veel mensen als onderdeel van het bedrijfsproces, zal dit opvallen. In dit geval kan deze applicatie worden opgepakt door de IT/BI-afdeling.

PowerPivot applicatie maken

In dit artikel gaan we een informatievraag oplossen met behulp van de PowerPivot client. Het scenario hier is als volgt: Contoso is een grootspeler in de wereldmarkt met winkels over de hele wereld. Om hun positie in de markt te versterken willen ze een betere klanttevredenheid bereiken om zodoende meer te kunnen verkopen aan bestaande klanten en nieuwe klanten te bereiken door mond op mond reclame. Om inzicht te krijgen in de huidige stand van zaken is door een extern bureau een onderzoek gedaan naar de klanttevredenheid per winkel. Er is gemeten wat de tevredenheid was van de producten die de klanten kochten, hulpvaardigheid van de medewerkers, de snelheid van de levering

en de totale klanttevredenheid. Het externe bureau levert de resultaten aan in een Excel bestand.

Het Contoso BI-team is op dit moment bezig met het oplossen van structurele data-problemen en het toevoegen van een nieuw HRM-systeem aan het centrale datawarehouse en heeft geen tijd om een ad hoc informatieverzoek op korte termijn te kunnen oplossen. Zodoende heeft het management aan jou, als Informatie medewerker, gevraagd om deze gegevens inzichtelijk te maken.

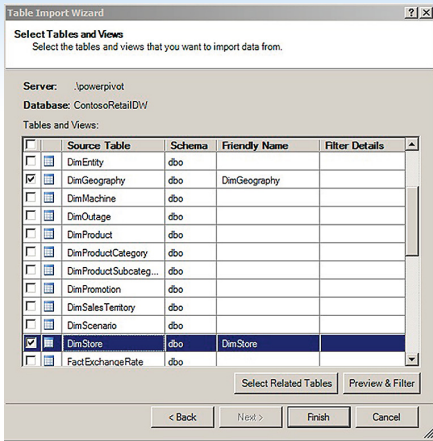
Om deze PowerPivot rapportage te kunnen maken hebben we een aantal gegevens nodig, om te beginnen het klanttevredenheidsonderzoek in Excel. Deze gegevens willen we kunnen verrijken met winkelgegevens, die beschikbaar zijn in het centrale datawarehouse.

De eerste stap in het maken van een PowerPivot applicatie is het laden van de gegevensbronnen in PowerPivot. Aangezien de gegevens van het klanttevredenheidsonderzoek al aanwezig zijn in een Excel bestand is het handig om daarmee te beginnen.

StoreName	Employee helpfulness	Ontime delivery	Product satisfaction	Overall Satisfaction
Contoso Tehran No.2 Store	8,1	10	9,1	
Contoso Saint Petersburg Store	8,1	8,2	9,7	
Contoso Damascus No.2 Store	9,8	8,4	10	
Contoso Beijing Store	8,1	8,2	8,1	
Contoso Moscow No.2 Store	8,7	8	8,2	
Contoso Sydney No.2 Store	9,8	9,1	8,3	
Contoso Timphu No.2 Store	9,5	8,9	8,2	
Contoso Hong Kong No.2 Store	9,2	9,3	8,9	
Contoso Toronto No.3 Store	9,5	8,1	9,6	

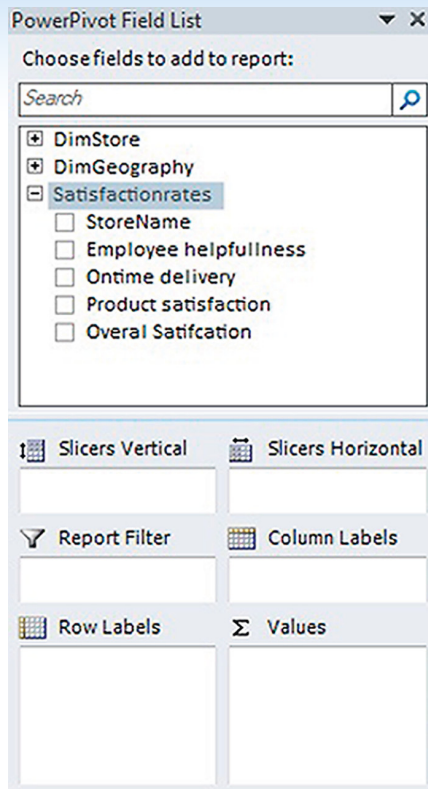
Een handige manier om gegevens vanuit Excel in te lezen als bron is gebruik te maken van de 'Createlinkedtable' functie in PowerPivot. Deze functie is beschikbaar in de PowerPivot Ribbon van Excel 2010: Naast deze functie bevat de Ribbon nog twee belangrijke knoppen, namelijk 'PowerPivot window' en de 'Pivottable'. De PowerPivot window knop opent de daadwerkelijke PowerPivot interface waar het PowerPivot model kan worden gemaakt en aangepast. Wanneer het model is gemaakt kan met behulp van de Pivottable knop PivotTable of PivotCharts worden toegevoegd aan het werkboek.

	A	B	C	D	E
	StoreName	Employee helpfulness	Ontime delivery	Product satisfaction	Overall Satisfaction
2	Contoso Albany Store	4,3	4	5,7	5,7
3	Contoso Alexandria Store	5,7	4,3	4,4	5
4	Contoso Amsterdam Store	4,8	6	5,2	4,4
5	Contoso Anchorage Store	4,7	4,4	4	5,7
6	Contoso Annapolis Store	5,9	5,3	5,9	5,4
7	Contoso Appleton Store	5,5	4,7	6	4,7
8	Contoso Arlington Store	4,9	5,8	5,6	4,5
9	Contoso Ashgabat No.2 Store	8,3	9,2	8	8,7
10	Contoso Ashgabat No.1 Store	6,5	7,1	6,4	6,9
11	Contoso Asia Online Store	5,1	5,4	5,5	4,5
12	Contoso Asia Reseller	5,4	5,1	4,7	4,2
13	Contoso Athens Store	4,2	4,4	5,1	4,8



Om de gegevens uit het aangeleverde Excel sheet naar PowerPivot te halen kunnen we het werkboek selecteren en dan op de knop 'Createlinkedtable' klikken. De gegevens worden nu ingelezen in de PowerPivot omgeving, de PowerPivot interface word automatisch geopend.

De volgende stap is het verrijken van deze gegevens met gegevens uit ons centrale datawarehouse. We willen dus gegevens uit onze database inlezen in PowerPivot. In het PowerPivot scherm kunnen we de knop 'From Database' vinden, wanneer we daar op klikken kunnen we kiezen voor 'From SQL Server'. In dit scherm kan verbinding gemaakt worden met de SQL Server waar ons datawarehouse op staat. Zo-



en score in de Pivottable slepen. We zien dan meteen iets vreemds: de getallen zijn niet netjes uitgesplitst per winkel. Hij toont per winkel de totalen. Dit komt omdat er geen relaties zijn gelegd tussen onze handmatige Excel-tabel en de gegevens uit het datawarehouse. Gelukkig heeft Power-

Nu geeft PowerPivot ons de resultaten per winkel, we kunnen nu de Top 10 winkels met de hoogste 'overall klanttevredenheid' score tonen.

Als informatiemedewerker van Contoso valt meteen iets op, alle winkels die goed scoren, hebben allemaal veel medewerkers in dienst. Zou er een verband zijn tussen het krijgen van een hoge klanttevredenheid score en veel medewerkers in dienst hebben?

Om dit te kunnen uitzoeken moeten we wat meer analyses uitvoeren, om te beginnen willen we kijken wat de gemiddelde score is per land. Omdat de 'winkels'-tabel een relatie heeft met de 'geografie'-tabel in het datawarehouse heeft PowerPivot deze automatisch bij het inlezen ook in PowerPivot meegenomen.

We kunnen de velden land en totale klanttevredenheid en aantal medewerkers in een nieuwe PowerPivot tabel slepen. PowerPivot maakt nu de som van de totale klanttevredenheid en aantal medewerkers per land. Om hier een gemiddelde van te maken kunnen we rechtermuisknop op de value in de PowerPivot field list klikken en dan kiezen voor 'Summarize by Average'. Op dat moment berekent PowerPivot het gemiddelde per land. Om meer inzicht te krijgen in de aantallen kunnen we gebruik maken van de Excel visualisatie mogelijkheden zoals databars. Nadat we de sortering hebben aangepast en de databars hebben toegevoegd zien we dat een hoge score inderdaad een relatie heeft met een groot aantal medewerkers.

Via het PowerPivot scherm kan verbinding worden gemaakt met de SQL Server waar ons datawarehouse op staat

dra de verbinding gemaakt is kan gekozen worden uit de tabellen die we willen inlezen. In ons geval willen we de gegevens uit de Winkel (DimStore) en Geografie (DimGeography) tabel.

Na deze geselecteerd te hebben worden de tabellen toegevoegd aan ons PowerPivot model, waarna we de gegevens kunnen gaan gebruiken in onze Pivottable in Excel. Met behulp van de Pivottable knop kunnen we een Pivottable toevoegen aan onze Excel-sheet. In de PowerPivot Field list kunnen we de gegevens selecteren die we in onze Pivottable willen tonen.

Om een pivot te maken met de tien winkels met de hoogste 'algemene klanttevredenheid' score kunnen we de winkelnaam

Pivot dit ook voor ons opgemerkt en zien we een melding 'Relationships may be needed' en een knop 'create'. Wanneer we hier op klikken zal PowerPivot proberen relaties te ontdekken tussen de verschillende tabellen. In ons geval heeft PowerPivot zelf een relatie ontdekt en aangemaakt.

Om meer inzicht te krijgen in de verschillen willen we een vergelijking maken van de score per land ten opzichte van de totale score. Dit kunnen we niet doen met de standaard Excel pivottable functies. Hier zullen we gebruik moeten maken van een berekende meetwaarde in DAX. We maken een nieuwe meetwaarde aan door op

Row Labels	Overall Satication	Employee helpfulness	Product satisfaction
Contoso Saint Petersburg Store	10	8,1	9,7
Contoso Beijing Store	10	8,1	8,1
Contoso Tehran No.2 Store	10	8,1	9,1
Contoso Damascow No.2 Store	10	9,8	10
Contoso Moscow No.2 Store	9,9	8,7	8,2
Contoso Sydney No.2 Store	9,8	9,8	8,3
Contoso Thimphu No.2 Store	9,7	9,5	8,2
Contoso Hong Kong No.2 Store	9,7	9,2	8,9
Contoso Shanghai No.1 Store	9,6	9,1	8
Contoso Montreal No.1 Store	9,6	8,1	9,2
Grand Total	98,3	88,5	87,7

Row Labels	Average of Overall Satisfaction	Average of EmployeeCount
Taiwan	9,50	95
Portugal	8,70	95
Canada	8,78	88
Russia	8,85	88
Pakistan	7,30	83
Turkmenistan	7,80	82
Thailand	7,25	79
China	7,39	68
Japan	6,75	67
Iran	7,60	65
Syria	7,55	65
Bhutan	7,50	65
Ireland	7,10	55
Switzerland	7,40	53
South Korea	5,65	50
Kyrgyzstan	5,50	50
Australia	6,33	48
Malta	5,70	42
United Kingdom	5,61	37
Slovenia	4,20	36
Spain	5,20	35
Sweden	4,40	33
India	5,03	33
Italy	5,18	29
United States	4,99	29
Singapore	5,00	27
the Netherlands	4,40	26
Denmark	5,90	25
Romania	4,20	25
Germany	4,76	21
France	4,86	16
Poland	4,50	15
Greece	4,80	15
Armenia	5,20	15
Grand Total	5,47	36

Row Labels	Average of Overall Satisfaction	Average of EmployeeCount	Compared to overall
Taiwan	9,50	95	174%
Portugal	8,70	95	159%
Canada	8,78	88	161%
Russia	8,85	88	162%
Pakistan	7,30	83	133%
Turkmenistan	7,80	82	143%
Thailand	7,25	79	133%
China	7,39	68	135%
Japan	6,75	67	123%
Iran	7,60	65	139%
Syria	7,55	65	138%
Bhutan	7,50	65	137%
Ireland	7,10	55	130%
Switzerland	7,40	53	135%
South Korea	5,65	50	103%
Kyrgyzstan	5,50	50	101%
Australia	6,33	48	116%
Malta	5,70	42	104%
United Kingdom	5,61	37	103%
Slovenia	4,20	36	77%
Spain	5,20	35	95%
Sweden	4,40	33	80%
India	5,03	33	92%
Italy	5,18	29	95%
United States	4,99	29	91%
Singapore	5,00	27	91%
the Netherlands	4,40	26	80%
Denmark	5,90	25	108%
Romania	4,20	25	77%
Germany	4,76	21	87%
France	4,86	16	89%
Poland	4,50	15	82%
Greece	4,80	15	88%
Armenia	5,20	15	95%
Grand Total	5,47	36	100%

de knop 'New measure' in de PowerPivot Ribbon te klikken. Om het gemiddelde van alle winkels berekenen kunnen we de functie CALCULATE gebruiken. Met behulp van deze functie kunnen we een berekening laten uitvoeren over een andere context. In ons geval is de context van de berekening niet alle winkels in een land zoals in de pivottable is geselecteerd, maar echt alle winkels. Onderstaande functie zou ons het gemiddelde 'Overall Satisfaction' geven voor alle winkels in een land in de pivottable

```
=CALCULATE(AVERAGE(Satisfactionrates[OverallSatisfaction]))
```

Maar aan de calculate tabel kan een extra parameter worden meegegeven waarin we de filter context kunnen overschrijven. In ons geval moet onze context alle winkels zijn. We kunnen een lijst van alle winkels krijgen door gebruik te maken van ALL(DimStore). De volledige berekening is als volgt:

```
=CALCULATE(AVERAGE(Satisfactionrates[OverallSatisfaction]), ALL(DimStore))
```

Nu kunnen we deze berekening gebruiken om te delen door het gemiddelde per land

```
=AVERAGE(Satisfactionrates[OverallSatisfaction]) / CALCULATE(AVERAGE(Satisfactionrates[OverallSatisfaction]), ALL(DimStore))
```

Dit geeft uiteindelijk het percentage van de rating van het huidige land ten op zichte van de totale rating. Ook hier kunnen we weer gebruik maken van de standaard Excel "conditionalformatting" functionaliteit om tot betere inzichten te komen.

Conclusie

We hebben gezien hoe een doorsnee informatiemedewerker gebruik kan maken van PowerPivot in zijn of haar dagelijks werk. Dit is maar een facet van de vele mogelijkheden van PowerPivot. Het kan ook worden gebruikt door BI Professionals om een 'proof of concept' van een bepaalde informatievraag snel op te lossen of voor het snel inzichtelijk krijgen van de verstopte data in een database.

Referenties

Voor meer informatie over PowerPivot kunt u op internet volgende interessante sites vinden: <http://powerpivotFAQ.com>, waar verschillende PowerPivot enthousiastelingen, Microsoft medewerkers en SQL MVP's vragen beantwoorden over PowerPivot.

<http://www.powerpivotPro.com> is een blog gerund door een voormalig medewerker van het PowerPivot team en is voornamelijk gericht op het maken van PowerPivot rapporten in de client.

<http://www.powerpivotwins.com> is de PowerPivot en SharePoint site voor de IT Pro, deze wordt gerund door twee Microsoft medewerkers met veel kennis over SharePoint en PowerPivot.

Kasper de Jonge, is Microsoft BI consultant voor ADA ICT. Op zijn blog <http://www.powerpivotblog.nl> kunt U meer informatie over PowerPivot, SSRS 2008 R2, SharePoint 2010 PowerPivot en PerformancePoint vinden.

