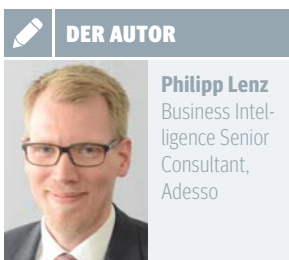


Datenmodellierung und Präsentation mit Microsoft Power BI

Die Power-BI-Plattform bietet die Möglichkeit, Daten in einem Datenmodell zu orchestrieren und in beeindruckenden Dashboards zu visualisieren. Die Herkunft der Daten ist nicht nur auf Datenquellen von Microsoft beschränkt, sie erstreckt sich auch auf andere wie Oracle, OData, Flatfiles etc.



DER AUTOR

Philipp Lenz
Business Intelligence Senior Consultant,
Adesso

Der Aufbau eines Datenmodells ist in der Praxis einfach umgesetzt. Die Komplexität des Modells steigt jedoch durch stetig wachsende Anforderungen und die Datenmenge. Die Praxis strapaziert somit das Datenmodell und dessen Wartbarkeit. Im Folgenden werden Best-Practice-Lösungsansätze für optimale Datenmodelle gezeigt und Möglichkeiten individueller Visualisierungsformen erläutert, um optimale Darstellungen zu zaubern. Gezeigt wird zudem, wie Informationen in Organisationen verteilt sind und wie sich bestehende SQL-Server-Berichte in Power BI einbinden lassen.

Datenmodelle von Anfang bis Ende richtig planen und sie auf die Quellen abstimmen

Verschiedenste Daten werden in ein Datenmodell durch die Vielzahl an Verbindungsmanagern in ein Power-BI-Modell importiert. Diese Datenquellen haben unterschiedlichste Aktualisierungsrhythmen und Granularitäten. Das ist eine der Herausforderungen für die ETL-Prozesse und den Datenimport, um die Daten im Datenmodell auf ein gemeinsames fachliches Level zu heben. Es empfiehlt sich, eine «High Level Map» als Dokumentation mit den Aktualisierungsintervallen zu erstellen und bei Bedarf nur einzelne Datenquellen zu aktualisieren, um die Daten nach Aktualität zu organisieren.

Lässt sich dauerhaft kein gemeinsamer Nenner finden (weil sich manche Datenquellen schneller und andere langsamer oder unzuverlässiger aktualisieren), ist es sinnvoll, eine ETL-Strecke in ein Data-Warehouse zu implementieren, um eine Analyseschicht für Auswertungen bereitzustellen.

Datenimporte richtig umsetzen

Ist diese Hürde genommen, lässt sich mit dem Importieren und Erstellen eine Microsoft-Power-BI-Mappe starten. Für den Import in Power BI lassen sich die Daten direkt in das Datenmodell importieren oder via Power Query beziehungsweise mit der Programmiersprache «M» transformiert in das Datenmodell einfügen, und zwar indem man Daten filtert, aggregiert oder durch Berechnungen anreichert. Durch Filtern der Daten lässt sich die Datenspar-

samkeit in den Vordergrund stellen und vorab bestimmen, welche Zeilen und Spalten importiert werden. Power BI zeichnet sich durch eine Engine aus, die unter anderem die Daten komprimiert speichert. Trotzdem empfiehlt es sich, nicht alle Daten zu importieren, um dauerhaft eine optimale Performance und Übersichtlichkeit zu haben. Der Vorteil: Das Aktualisieren der Datenquellen wird verkürzt.

Zu beachten ist, dass es sich um eine spaltenorientierte Datenbank-Engine handelt, die die jeweiligen Spalten in je einem Speicherbereich hält, das heisst, weniger Daten, kleinere Speicherbereiche, bessere Indizierung, weniger Arbeitsspeicher-Verbrauch und bessere Performance. Nutzer, die Datenquellen aus Datenbanken verwenden, die bereits in der Abfrage das Filtern von Daten ermöglichen, sollte diese bevorzugt verwenden.

Microsoft Power BI bietet viele Möglichkeiten, um Dashboards und Berichte zu verteilen.

So lassen sich über Sichten beide Filterarten direkt auf dem Server anwenden. Spätere Datenaggregationen lassen sich einerseits auf dem Server durchführen, andererseits beim Verbinden weiterer Datenquellen, die nicht aus der Datenbank entnommen werden, innerhalb von Power Query.

Die Modellierung bestimmt die Funktionalität

Was für performante OLAP Cubes gilt, gilt auch für Power BI. Je nach Möglichkeit empfiehlt es sich, Star-beziehungsweise Snowflake-Schemata einzusetzen, die dimensional modelliert sind. Dies ermöglicht übersichtliche Datenmodelle und das Einstellen der Filter-Richtung.

Durch die Filter-Richtungen lässt sich festlegen, ob alle Möglichkeiten dargestellt sind, die die Dimensionen anbieten oder nur die, die im aktuellen Filter-Kontext noch zur Verfügung stehen. Werden alle Daten in einer breiten Tabelle bereitgestellt, ist eine solche Einstellung nicht möglich. Nach Möglichkeit sollten Nutzer numerische Schlüsselattribute einsetzen. Das Datenmodell kann sie schneller verarbeiten und sie ermöglichen eine bessere Kompression und Indizierung.

Weiterhin ist es empfehlenswert, auch eine Datumsdimensionstabelle zu verwenden, mit der sich zum Beispiel



2-Fakten-Tabellen miteinander verbinden und filtern lassen. Ein Beispiel ist die «CALENDAR()»-Funktion in DAX, die eine entsprechende Tabelle erzeugt und die sich über die Funktionen Day(), Month() etc. ausbauen lässt.

Definierte Kennzahlen anstatt Zufallsprodukte

Um das Bedienen zu erleichtern und eine Governance von Kennzahlen zu gewährleisten, sollten alle Spalten aus dem Datenmodell ausgeblendet werden, mit denen Kennzahlen gebildet werden können. Dafür sind in DAX entsprechende Measures (berechnete Kennzahlen) zu erstellen, die in einer eigenen Measure-Tabelle abgelegt werden. Dafür wird eine leere Tabelle über den Menüpunkt «Daten eingeben» erstellt. In dieser Tabelle werden dann die Measures erstellt.

Dieses Vorgehen erlaubt es, komplexe sowie einfache Berechnungen im Datenmodell zentral zu platzieren. Dies wird für alle sinnvollen Zahlen, die in den Fakten enthalten sind, durchgeführt, um weiterhin den «Self-Service-BI»-Aspekt zu erhalten. Explizite Measures ermöglichen ein schnelles Wiederverwenden sowie das Sicherstellen der korrekten Berechnung. Damit das alles übersichtlich ist, sollte das Datenmodell noch «feingetunt» werden, indem alle technischen Spalten, die für Verbindungen zwischen Tabellen notwendig sind, ausgeblendet werden. Jedoch ist hier eine entsprechende Governance eindringlich empfohlen, um auf bestehende Datawarehouse-Lösungen zurückgreifen, beziehungsweise auf den dortigen Cubes die vorhandenen berechneten Measures und KPIs wiederverwenden zu können.

Verteilen von Dashboards und Berichten

Microsoft Power BI bietet viele Möglichkeiten, um Dash-

boards und Berichte zu verteilen. Berichte lassen sich in sogenannten Arbeitsbereichen organisieren und innerhalb vorhandener Office-365-Gruppen aus dem Azure Active Directory teilen. Dashboards lassen sich zudem innerhalb und ausserhalb der Organisation teilen. Letzteres lässt sich einstellen, damit Daten nicht in falsche Hände geraten. Dashboards an Externe zu verteilen, ist beispielsweise dann eine Option, wenn Unternehmen Dienstleister oder Zulieferer in ihre Prozesse einbinden und über den aktuellen Stand informieren wollen. Eine weitere Möglichkeit ist das Veröffentlichen von Berichten im Internet. Der Bericht steht als öffentlicher Link zur Verfügung und lässt sich etwa als IFrame in eine Webseite einbinden. Zusätzlich bietet Power BI Schnittstellen zum Einbinden der Berichte via REST-Schnittstelle in Webseiten und Apps an.

Integration von Reporting-Services in Power BI

Mit Microsoft SQL Server 2016 kamen die Mobile Reports. Bei den klassischen Paginated Reports besteht nun die Möglichkeit, den Reporting-Services-Dienst über die Verwaltungskonsolle mit einem Power BI Tenant zu verbinden. Von dort können die Benutzer Elemente aus den Berichten an vorhandene Power-BI-Dashboards «anheften». Voraussetzung hierfür ist jedoch, dass die Datenquellen gespeicherte Anmeldeinformationen beinhalten und sich der jeweilige Benutzer mit seinen Power-BI-Anmeldeinformationen einloggt. Ab dem Einloggen aktualisiert der SQL-Server-Agent die Reporting-Services-Kachel des Berichts automatisch. Sobald der Nutzer auf eine Reporting-Services-Kachel innerhalb eines Power-BI-Berichts klickt, wird er auf den Berichtsserver navigiert. Hierzu muss der Benutzer eine Verbindung zum Berichtsserver haben.

Power BI zeichnet sich durch eine Engine aus, die unter anderem die Daten komprimiert speichert.

Grafiken: freepik.com

Der Aufbau eines Datenmodells ist in der Praxis einfach umgesetzt. Die Komplexität des Modells steigt jedoch durch stetig wachsende Anforderungen und die Datenmenge.