

AUSZUG AUS...
IT MITTELSTAND
 AUSGABE 11|2018

DATENANALYSEN

AUF DIE VORBEREITUNG KOMMT ES AN

Ohne die entsprechende Datengrundlage lassen sich aussagekräftige Analysen nur schwerlich realisieren. Die Methoden der „*Data Preparation*“ können die Vorbereitungen unterstützen.

LAUT DEM „FIRST HALF 2018 MARKET SURVEY“ von Nucleus Research gehen die Ausgaben für Business-Intelligence-Tools (BI) im Jahr 2018 um 80 Prozent zurück. Das Problem ist nicht, dass es ein geringeres Interesse gibt, Informationen zu analysieren. Stattdessen wollen 83 Prozent der Anwender Analytics-Werkzeuge unmittelbar in den Applikationen, die sie ohnehin nutzen, anstatt kostbare Zeit damit zu verbringen, zu BI-Stand-alone-Lösungen zu wechseln.

Ferner bleiben Reports, Dashboards und Visualisierungen ein wichtiger Bestandteil des Analyseprozesses. Mehr und mehr Nutzer sind aber nicht mehr nur mit diesen „Frontend“-Ergebnissen der Analyse befasst, sondern in wachsendem Maße mit den für die Analyse grundlegenden „Backend“-Aktivitäten. Hier kommt die sogenannte „Data Preparation“ (siehe Infokasten) ins Spiel. Das Würzburger Forschungs- und Beratungsinstitut Barc hat 2017 in einer weltweiten Studie das Thema „Data Preparation im Fachbereich“ untersucht (695 Teilnehmer in

KEINE KOSTBARE ZEIT VER-SCHWENDEN

50 Ländern). Danach nutzen bereits heute sieben von zehn Unternehmen die Methode, eines von zehn Unternehmen befindet sich einer Projekt- oder Planungsphase dazu. Die wichtigsten Treiber für Data Preparation sind:

- › **49 PROZENT:** höhere Erwartungen an die Performance, Agilität und Flexibilität der Unternehmensbereiche;
- › **47 PROZENT:** höhere Erwartungen an greifbare Geschäftsauswirkungen und bessere Wettbewerbsfähigkeit durch Analytics;
- › **46 PROZENT:** wachsende Zahl von Datenquellen mit steigenden Datenmengen, Vielfalt der Datentypen und -quellen und zunehmender

Geschwindigkeit, mit der Daten generiert, ausgewertet und weiterverarbeitet werden können (Big Data).

Diese Ergebnisse klingen laut Zetvions nicht nach einem neuen Hype, sondern nach handfesten geschäftlichen Anforderungen. Dabei ist die Datenvorverarbeitung durch IT- und Fachabteilungen keine neue Erfindung. Was Data Preparation von seinen technischen Vorfahren unterscheidet, ist seine Relevanz und Nutzbarkeit für ein breiteres Spektrum nicht-technischer Entscheidungsträger. Das hat die Bostoner Aberdeen Group in zwei Untersuchungen (März 2016, März 2017) deutlich gemacht. Das klassische Datenmanagement



Was versteht man unter Data Preparation?

↳ „Data Preparation“ umfasst alle Aktivitäten, die dazu dienen, die Qualität, Nutzbarkeit, Zugänglichkeit und Übertragbarkeit von Daten zu verbessern. Zu diesen Aktivitäten gehören (sind aber nicht darauf begrenzt):

- › **Datenintegration** (das Zusammenführen von Informationen aus verschiedenen Datenbeständen),
- › **Data Profiling** (der weitgehend automatisierte Prozess zur Qualitätsanalyse vorhandener Datenbestände),
- › **Datenbereinigung** (Verfahren zum Entfernen und Korrigieren von Datenfehlern in Datenbanken oder anderen Informationssystemen) und
- › **Data Governance** (unternehmensinterne Richtlinien für den Umgang mit Daten).
- › **Das Ziel von Data Preparation** ist, Nutzer und analytische Systeme mit sauberen und nutzbaren Daten zu versorgen – als Grundlage für die Ermittlung relevanter Kundenbedürfnisse. ↩

Quelle: Zetvvisions AG

Druck, in immer kürzerer Zeit die dafür erforderlichen „richtigen“ Informationen bereitzustellen.

Das Mantra der Datenprofis

Die richtigen Informationen, zum richtigen Zeitpunkt, an der richtigen Stelle – das ist das Mantra der Datenprofis. Denn: Mit Blick auf die aus den Daten gewonnenen Einsichten und die auf ihrer Basis getroffenen Entscheidungen gilt auch schon immer: „Garbage in – garbage out“. Die Qualität und Nutzbarkeit der in der Analyse verwendeten Daten hat einen direkten Einfluss auf die Richtigkeit der Erkenntnisse, die auf der anderen Seite herauskommen, wie Aberdeen zutreffend bemerkt. Zusätzlich kämpfen Unternehmen heute mit der Dauer, die sie für die Bereitstellung der Daten für die Analyse benötigen. Schlechte Datenqualität, mangelhafte Zugänglichkeit zu den richtigen Datenquellen und unzureichende Geschwindigkeit bringen den Analyseprozess ins Stocken. Es sind genau diese Probleme, die Unternehmen mit dem Einsatz von Data Preparation beheben oder gar nicht entstehen lassen wollen. Dabei, so Aberdeen, sei eine Spitzenleistung auf diesem Gebiet mitunter dadurch definiert, den Zeitraum zu verkürzen, um die Daten für die Analyse bereitzustellen.

Allerdings lassen zwei Ergebnisse der Barc-Studie aufhorchen. Die Untersuchungsteilnehmer sind sich zwar der Relevanz von Data Governance für Data Preparation bewusst, aber nur 32 Prozent sind diesbezüglich mit dem aktuellen Zustand zufrieden oder sehr zufrieden. Ferner erstaunt, welches Werkzeug für Data Preparation mit Abstand an erster Stelle steht: Excel. Über die begrenzten Funktionalitäten von Excel für alles, was mit dem Management von Daten zu tun hat, braucht hier nicht diskutiert zu werden. Dem Anspruch von Data Preparation, die Qualität, Nutzbarkeit, Zugänglichkeit und Übertragbarkeit von Daten zu verbessern, kann auf Dauer nur gerecht werden, wer eine Data Governance einführt und zudem bereit ist, in professionelle Software-Lösungen – nicht zuletzt auch für das Stammdatenmanagement – zu investieren. ↩

MONIKA PÜRSING

- › war eine Domäne der IT-Fachleute und Datenbankadministratoren. Data Preparation ist nicht nur etwas für Datenprofis, sondern gibt einem breiteren Kreis von Verantwortlichen die Kontrolle über die Qualität und Nutzbarkeit ihrer Daten. Dazu gehören auch Stammdaten zu Produkten, Lieferanten, Kunden, Mitarbeitern und Finanzen. Marketing-Manager, Vertriebsdirektoren oder CFOs – sie alle übernehmen eine aktive Rolle bei der Verbesserung ihres Datenbestands.

Doch es gilt noch einige Stolpersteine aus dem Weg zu räumen. Stichwort „Datensilos“: 46 Prozent der von Aberdeen 2016 befragten Unternehmen

sagen, die Schwierigkeit, Daten bereichsübergreifend zugänglich zu machen, sei ihre größte Herausforderung. Neben den daraus erwachsenden Datenqualitätsproblemen spiele der Faktor Zeit eine immer größere Rolle. Da Dynamik und Tempo in der Geschäftswelt immer mehr zunehmen, bleibt weniger Zeit, um die hoffentlich „richtigen“ Entscheidungen zu treffen. Daher wächst der