

Jürgen Noe

Schnelleinstieg in SAP® Business Warehouse (BW)

- ▶ Grundlagen von Business Intelligence
- ▶ Die Data Warehouse Workbench
- ▶ Die Business Explorer Suite
- ▶ Extraktion, Transformation und Lade
- ▶ Problembehebung
- ▶ Durchgängiges Fallbeispiel

Inhaltsverzeichnis

Einleitung	5
1 Business Intelligence	9
1.1 Die Bedeutung von Business Intelligence	9
1.2 Data Warehouse	11
1.3 Die Corporate Information Factory	13
1.4 SAP Business Warehouse	15
1.5 Abgrenzung zu SAP BO Business Intelligence	19
1.6 Neuere Ansätze	23
2 Fallbeispiel	31
2.1 Aufgabenstellung	31
2.2 Kennzahlen- und Kontenmodell	32
2.3 Die wichtigsten Modellierungselemente	36
2.4 Datenflussdiagramme	45
3 Datenflussdiagramm	49
3.1 Namenskonvention	49
3.2 Anlegen der InfoAreas	51
3.3 Anlegen der Anwendungskomponenten	55
3.4 Das Datenflussdiagramm	57
4 Datenmodellierung	69
4.1 Anlegen der InfoObjectCataloge	69
4.2 Anlegen der InfoObjects	76
4.3 Navigationsattribute	113
4.4 Anlegen der DataSources	116
4.5 Anlegen der DataStoreObjects	129

4.6	Anlegen der InfoCubes	143
4.7	Anlegen der MultiProvider	153
5	Extraktion, Transformation und Laden (ETL)	163
5.1	Begriffserklärung ETL-Prozess	163
5.2	Extraktion	165
5.3	Transformation	177
5.4	Laden	205
6	Reporting	239
6.1	BEx Query Designer	239
6.2	BEx Analyzer	253
7	Trouble-Shooting	263
7.1	InfoObject im Business Content aktivieren	263
7.2	Request aus PSA löschen	266
7.3	Daten im DSO löschen	267
7.4	Daten im Cube löschen	269
7.5	Steuereintrag der Stammdaten löschen	270
8	Fazit	273
A	Über den Autor	279
B	Index	281
C	Literaturhinweise	287
D	Disclaimer	289
	Weitere Bücher von Espresso Tutorials	290

2 Fallbeispiel

Im Folgenden beschreibe ich Ihnen ein Fallbeispiel aus der Pharmabranche und stelle grundlegende Überlegungen zur Umsetzung der darin gestellten Aufgaben an. In diesem Zusammenhang werde ich auf das Kennzahlen- sowie das Kontenmodell eingehen – die wichtigsten Modelle für die prinzipielle Organisation Ihrer Daten in einem InfoProvider.

Es folgt eine kurze Darstellung der zentralen Modellierungselemente InfoObject, DataStoreObjekt, InfoCube und MultiProvider. Schließlich entwerfen wir einen kompletten Datenfluss für eine vereinfachte Mehrschichtenarchitektur, wie sie im letzten Kapitel vorgestellt wurde, inklusive der Datenflussdiagramme.

2.1 Aufgabenstellung

Stellen Sie sich folgende Situation vor: Sie arbeiten in der Business Intelligence-Abteilung eines weltweit operierenden Pharmaunternehmens. Bei einem Meeting mit Vertretern aus dem zentralen Vertrieb werden Anforderungen für einen neuen Bericht formuliert: Der zentrale Vertrieb benötigt schnellstmöglich eine Übersicht der unternehmenseigenen umsatzstärksten Produkte (Top 3) über die letzten 36 Monate im Vergleich zu denen der Wettbewerber. Es soll eine flexible, monatsgenaue Darstellung erfolgen. Im Unternehmen steht den Mitarbeitern des Vertriebs SAP Business Warehouse als zentrales Tool für die monatliche Ausführung dieser und weiterer Auswertungen zur Verfügung. Zusätzlich erhalten Sie die Information, dass IMS Health-Daten gekauft und im Projektlaufwerk abgelegt wurden. Anhand des National Drug Codes wurde aus den IMS Health-Daten eine aktuelle Liste erstellt, die Ihnen anzeigt, bei welchen Produkten es sich um die eigenen bzw. die der Wettbewerber handelt.

Zurück am Arbeitsplatz öffnen Sie die IMS Health-Daten im Projektlaufwerk. Sehen Sie in Abbildung 2.1 verkürzt einen Auszug der Daten in MS Excel.


The image shows a screenshot of an Excel spreadsheet containing a large table of data. The table has many columns, with the first few containing text and the rest containing numerical values. The data appears to be organized in a grid format, typical of a data export from a database or system. The text is small and difficult to read, but the structure is clear.

Abbildung 2.1: Auszug der IMS Health-Daten

Auch wenn ich hier keine Details sichtbar machen kann, erkennen Sie, dass die Datei aus einer Vielzahl von Spalten besteht: In der ersten Spalte finden Sie den Handelsnamen, in der zweiten den eindeutigen National Drug Code, in der nächsten die Kundengruppe, gefolgt von der Dosierung der Inhaltsstoffe, der Verpackungsgröße und dem Einsatzgebiet. Die nächsten 36 Spalten beschreiben den Gesamtumsatz, gefolgt von den Gesamtmengen der letzten 36 Monate in Eaches. Eaches, abgekürzt EA, ist im SAP-Customizing eine Einheit für Mengen, die frei vom Kunden definiert werden kann. Es handelt sich hierbei um eine alternative Einheit, die in der Praxis recht häufig verwendet wird. Um diese Daten in SAP BW abzubilden und auszuwerten, kommen prinzipiell das Kennzahlen- und das Kontenmodell infrage.

2.2 Kennzahlen- und Kontenmodell

SAP BW bietet zwei prinzipielle Modellierungsstrategien, nach denen Sie Ihre Daten organisieren können. Das *Kennzahlenmodell* modelliert die Kennzahlen möglichst direkt. Die eben vorgestellte IMS Health-Datei ist ein Beispiel dafür. Jeder monatliche Umsatz, jede Menge inklusive Mengeneinheit ist als eigene Kennzahl definiert. Im Beispiel der IMS Health-Daten werden die Kennzahlen im Wesentlichen durch den National Drug Code sowie die Dosierung und Darreichungsform eindeutig beschrieben. Dies führt zu einem sehr breiten Datensatz. Die Anzahl der Zeilen wiederum ist im Kennzahlenmodell im Vergleich zum Kontenmodell sehr klein gehalten. Sie erhal-

ten pro National Drug Code mit Dosierung und Darreichungsform eine Zeile und 180 Kennzahlen-Spalten.

Im Gegensatz hierzu sind beim *Kontenmodell* die Kennzahlen einzelnen Konten zugeordnet, daher der Name. Das Kontenmodell hat den Vorteil, dass es mit nur wenigen Kennzahlen auskommt. Über eine Hilfsgröße kann eine Kennzahl eindeutig einem Konto zugeordnet werden. Damit minimiert sich zwar die Datensatzgröße, die Anzahl der benötigten Zeilen wird jedoch wesentlich erhöht.

Im Beispiel der IMS Health-Daten dient der Monat als Konto. Die Kennzahlen werden auf jeden einzelnen Monat gebucht. Es ergeben sich folgende Kennzahlen:

- ▶ Umsatz in der jeweiligen Währung,
- ▶ Gesamtmengen in Eaches.

Bei der Darstellung im Kontenmodell würden Sie für das Konto »Monat« pro National Drug Code inklusive Dosierung und Darreichungsform 36 Zeilen und fünf Kennzahlen erhalten. Der Umfang Ihrer Datei wird im Kontenmodell also um den Faktor 36 aufgebläht – kaum geeignet für den E-Mail-Versand oder Download. Aus diesem Grund liegen die IMS Health-Daten im Kennzahlenmodell vor.

Wenn man sich die IMS Health-Daten aus Abbildung 2.1 genauer anschaut, kann man feststellen, dass viele Kennzahlenspalten den Wert 0 haben. Dies ist einer der Nachteile im Kennzahlenmodell: Selbst wenn keine Werte für eine Kennzahl vorliegen, z. B., weil es sich um ein neues Produkt ohne Umsatzzahlen älter als ein Jahr handelt, müssen alle Kennzahlenspalten gefüllt werden. Im Kontenmodell können Sie diese Zeilen einfach auslassen. Für ein neues Produkt werden Werte in der entsprechenden Zeile erst ab dem Monat eingetragen, ab dem die ersten Umsätze vorliegen. Vergangenheitswerte sind nicht gefüllt.

Das Kennzahlenmodell ist zudem wenig flexibel, wenn eine weitere Kennzahl mit aufgenommen werden soll. Wenn Sie etwa bei den IMS Health-Daten nicht die Historie über die letzten 36 Monate, son-

dern über einen längeren Zeitraum von 60 Monaten betrachten möchten, müssen Sie 72 zusätzliche Kennzahlenspalten anlegen. Dies bedeutet aber auch: Sie müssen alle Datenstrukturen und InfoProvider, die Sie zur Abbildung im SAP BW benötigen, um 72 zusätzliche Kennzahlen erweitern. Dies stellt einen nicht zu vernachlässigenden Aufwand dar.

Im Kontenmodell dagegen müssten Sie weder am Modell noch an der Datenstruktur etwas ändern. Für eine Langzeitbetrachtung über fünf Jahre hätten Sie pro Kennzahl 60 Zeilen, sofern Daten für diesen Zeitraum vorhanden sind.

Wenden wir uns wieder unseren Aufgaben im Fallbeispiel zu: Welches Modell eignet sich eher für das Erstellen beider Berichte mit allen geforderten Details? Die erste Anforderung fragt nach den eigenen umsatzstärksten Top 3-Produkten. Betrachten wir hierzu zunächst, wie das gewünschte Ergebnis der Top 3-Produkte in Excel später in SAP Business Explorer Excel aussehen soll:

Für unseren Bericht ist die gewünschte Struktur nach Umsatz, National Drug Code, Hersteller sowie zusätzlich nach Monat gegliedert. Diese Berichtsanforderung lässt sich über ein Kennzahlenmodell lösen. Die für die Ausgabe gewünschte Struktur im SAP Business Explorer Analyzer wird auch *Startaufriss* genannt.

Startaufriss



Der Startaufriss einer Abfrage beschreibt, auf welchen Merkmalskombinationen die Auswertung im Bericht erfolgt. Die Merkmalskombination in den Zeilen einer Spalte beschreibt den Startaufriss. Die Merkmale in den sogenannten freien Merkmalen einer Abfrage beschreiben die Navigationsmöglichkeiten.

Um in unserem Beispiel einen Startaufriss hinzubekommen, müssen wir den Hersteller zum jeweiligen National Drug Code nachlesen. Dazu benötigen wir aus dem Vertrieb die Liste mit der Zuordnung

des NDC zu den Herstellern. Wie Sie im SAP BW den Hersteller zu einem National Drug Code nachlesen, zeige ich Ihnen im Abschnitt 5.3. Abschließend muss nach Eigenprodukten gefiltert und nach Umsatz sortiert werden. Als Ergebnis erhält man die eigenen umsatzstärksten Produkte. Die monatsgenaue Darstellung lässt sich über eine entsprechende Kennzahl darstellen. Dies funktioniert allerdings nur für die in der IMS Health-Datei geladenen Monate. Weiter zurückliegende Zeiten, wie z. B. 36 Monate von heute in die Vergangenheit gerechnet, können nicht mehr dargestellt werden. Die geforderte flexible Darstellung nach einzelnen Monaten oder Perioden ließe sich im Kennzahlenmodell nur schwer bis gar nicht erreichen. Dies ist ein erstes Indiz dafür, dass für unsere Aufgabe ein Kontenmodell benötigt wird.

Gibt es also eine allgemein gültige Aussage, wann man das Kennzahlenmodell und wann das Kontenmodell verwenden sollte? Diese Frage lässt sich in der Tat nach folgenden Kriterien entscheiden:

Verwendung Kennzahlenmodell vs. Kontenmodell

Kennzahlenmodell

- ▶ beschränkte, konstante Anzahl an Kennzahlen, die sich nicht oder nur selten ändern,
- ▶ in allen Datensätzen ist die Mehrheit der Kennzahlen mit Werten ungleich Null ausgeprägt,
- ▶ wenige Datensätze, große Datensatzbreite.



Kontenmodell

- ▶ hohe und vor allem flexible Anzahl von Kennzahlen in Berichten,
- ▶ im Kennzahlenmodell sind viele Kennzahlen mit Null ausgeprägt,
- ▶ Berichte selektieren nur wenige Datensätze,
- ▶ viele Datensätze, geringe Datensatzbreite.

B Index

A

ABAP 6, 66, 69, 108, 191,
194, 279
Funktionsbausteine 103
Abbruch Ladeprozess 266
Anwendungskomponenten 55
anlegen 55–57
Anzeige­komponenten anlegen
55–57
Applikations­komponenten 46,
49, 51

B

Berichtswesen 11, 13, 15, 19,
27
Bewegungsdaten
laden 236–38
Business Content 36, 46
Business Explorer
BEx Analyzer 253–61
BEx Query Designer 239–
53
Business Explorer Suite 19,
90
Business Information
Accelerator 21
Business Intelligence 21, 9–
30
Definition 9

SAP BO Business
Intelligence Suite 273
SAP BO Business
Intelligence Suite 4.1 21
Business Layer 46, 54, 76,
91, 111, 141, 143, 145, 148,
164, 199, 216
Business Objects
SAP BO Business
Intelligence 5
SAP BO Web Intelligence
20
Business Transformation 13,
24
Business Warehouses
Business Content 37, 38,
40, 78, 86, 263, 264
Business Warehouse
Accelerator 26, 42, 44
Dimensionen 17
BW Query 239

C

Corporate Identity 19
Corporate Information Factory
13–15
Cross-Selling 9
Crystal Reports 19, 20

D

Dashboard Designer 19, 21
Data Acquisition Layer 24
Data Marts 13, 24
 Departmental Data Marts
 15
Data Warehouses 10, 14, 28,
 164
 Architektur 13
 Business Warehouse 2.0
 28
 Data Warehousing
 Workbench 49, 51, 55,
 97, 114, 116, 165, 205,
 256, 263
 Definition 11
 Ebenen 12
 Enterprise Data
 Warehouses 13
Data-Mining 9, 19
DataSources
 anlegen 116–29
 Bewegungsdaten- 116
 DataSource-Pflege 128
 Stammdaten 124–28
DataStore Objekte (DSO) 39–
 41, 133
 anlegen 131
DataStores 134
Datenbanken 146
 In-Memory- 26
 spaltenbasierte 26
Datenfluss 57, 131, 138, 143
Datenflussdiagramme 47, 45–
 47, 57, 59, 129, 162
 anlegen 57–68

Datenmodellierung 69–162
Datensilos 13
Datentransferprozesse (DTP)
 16, 164, 184–91, 199, 198–
 204
Datentypen 79, 86, 90, 98,
 108, 121, 126
 -Vorschlag 120
Datenvolumen 43
Decision Support Systeme 15
Dimensionen
 anlegen und umbenennen
 146–48
 Merkmale tauschen 148–49
Dimensionstabellen 17, 42,
 146
DSO 27
 Requestnummer 135
 schreiboptimiertes 131, 133
 schreib-optimiertes 40
 semantischer Schlüssel
 135, 141
 Vorlagen 134

E

Ebene, virtuelle 46
Einplanung 232
Empfehlungsdienst 9
Endroutine 194, 198
Engpässe 148
Expertenroutine 194, 198
Extraktion, Transformation und
 Laden (ETL) 13
 Definition 163–65
 Extraktion 167

- IMS-Daten ins BW laden
165–73
- Laden 205–38
- NDC-Stammdaten laden
173–77
- Transformation *Siehe*
Transformation
- F**
- Fallbeispiel 30–47
- Flatfile-Extraktoren 164
- G**
- Gewinn- und Verlustrechnung
10
- H**
- Hierarchien 36, 57, 78, 87,
90, 92, 207, 257
- Historienkubes 42, 44
- I**
- IMS Health 31, 33, 60, 116,
120, 123, 131, 140, 165,
192
- Prozessketten 209–26
- Index 223
- InfoAreas 46, 49, 51, 57, 69,
144
- anlegen 51–55
- InfoCubes 41–44, 113, 143,
150, 152
- anlegen 143–53
- degenerierte 146
- InfoObjectCataloge 46, 84,
108, 134–38, 141
- anlegen 69–75
- InfoObjects 16, 36–39, 46,
134, 149, 265
- Abnehmer 90
- aktivieren 109
- anlegen (Business Layer)
91–92
- anlegen (Kennzahlen
Umsatz/Menge) 96–102
- anlegen (Massenanlage)
103–12
- anlegen (Propagation
Layer) 76–91
- anlegen (Reporting Layer)
92–96
- Dosierung 89
- Einsatzgebiet 91
- im Business Content
aktivieren 263–66
- Kennzahlen *Siehe*
Kennzahlen
- Klammerung 38
- Konzentration 89
- löschen *Siehe* Löschen
fehlerhafter Angaben
- Massenanlage 103
- Merkmale *Siehe* Merkmale
- Navigationsattribute 113–
15, 152–53, 158
- Pflege 78
- Referenz- 78
- StandardInfoObjects 38
- Template- 78, 91
- Verpackungsgröße 90

Infopackage 16, 164, 166,
169

InfoProvider 16, 113

Data Store Objects 16

InfoCubes *Siehe* InfoCubes

InfoSource 134

J

Jahresbilanz 10

K

Kennzahlen 37, 74, 96, 111,
137, 140

aktivieren 100

Betrag 140

Kennzahl-Pflege 98

Marktanteile 111

Menge 111, 141

selektieren 159

Umsatz 111, 140, 141

Kennzahlenmodell 32

Kontenmodell 33, 140, 238

Kostenstelle 17, 38

L

Laden *Siehe* Extraktion,
Transformation und Laden
(ETL)

Layered Scalable Architecture
(LSA) 23, 27

Line Items 43

Löschen

fehlerhafter Angaben (DSO)
267–69

fehlerhafter Angaben
(InfoCube) 269–70

fehlerhafter Angaben (PSA)
266–67

Steuereintrag der
Stammdaten 270–72

LSA *Siehe* skalierbare
Architektur

M

Mehrschichtenarchitektur 69

Ebenen 51

Mengenfelder 122

Merkmale 74, 95, 113

Attribute 17

Hierarchien 17

referenzierte 78

Zeitmerkmal 140

MultiProvider 44–45, 46, 60,
243

aktivieren 161

anlegen 153–62

prüfen 161

N

Name, technischer 140

Namenskonvention 49–51

National Drug Code 31, 32,
92, 94, 114, 252

Navigationsattribute *Siehe*
InfoObjects

O

Objekte

aktivieren 72

- bestehende übernehmen
 - 129
 - globale 49
 - lokale 49
 - Operational Data Store (ODS)
 - 25
 - Optimierung 42, 65, 146
- P**
- Persistent Staging Area 16, 40
 - Probe 123
 - Propagation Layer 24, 45, 46, 54, 70, 73, 74, 76, 86, 91, 96, 111, 164
 - Provider, semantisch
 - partitioniert 43
 - Prozesskette 165
 - Prüfung 128
- Q**
- Qualitätssicherung 28, 29
 - Quality & Harmonisation Layer
 - 24
 - Quellsystem 15, 38, 45, 49, 55, 116, 163, 164, 264
 - Queries 27, 239–61
 - Queryausführung 148
- R**
- Recommendation Engine 9
 - Reporting 19, 113, 152, 238–61
 - Reporting Layer 27, 46, 54, 76, 92, 164
 - Request löschen 266
- S**
- SAP
- Business Information Warehouse 15, 21
 - SAP Business Explorer Analyzer 34
 - SAP Business Explorer Excel 34
 - SAP Business Warehouse 15–19, 38, 39, 46
 - SAP BW 49
 - SAP Customizing 32
 - SAP ERP 49, 163
 - SAP HANA 26, 42, 43, 44
 - SAP R/3 15
- Schattencubes 42
 - semantischer Schlüssel *Siehe* DSO
 - skalierbare Architektur 23, 27
 - Snowflake-Schema *Siehe* Sternschema
 - Social Media Networks 28
 - Sofortstart 227
 - Staging 15
 - Staging Area 13, 24
 - Stammdaten 28, 86
 - laden 231–35
 - Prozessketten 226–30
 - Startauftritt 34
 - Startroutine 194, 198
 - Sternschema 17, 41
 - Dimensionstabellen 17
 - erweitertes 17
 - Faktentabelle 17

T

- Tag-Cloud 20
- Templates 47, 60, 63, 68, 76
- Textarten 92
- Transaktionen
 - Transaktion RSA1 51, 57, 201
 - Transaktion RSPC 205
 - Transaktion SE38 108
- Transformation 177–205
 - von DataSource in schreiboptimiertes DSO 178–91
 - von Kennzahlen- zu Kontenmodell 191–99
 - von NDC-Stammdaten 202–5

U

- Umsatzfelder 121

- Unternehmensplanung 15, 19, 40
- Update 168
- Upselling 9
- Usability 19

V

- Versionsmanagement 65, 66, 68
- Vertrieb 9, 10, 28, 31, 34, 279
- Virtual Layer 54, 57, 242
- Vorschau 123
- Vorschlag 119, 126

W

- Werke 78
 - Vertriebswerk 78