

Departement für Informatik
Universität Fribourg

Diplomarbeit

Data Warehouse zur Online-Analyse von Börsendaten und
Portfolioentwicklungen



1. Referent: Prof. Dr. Andreas Meier
2. Referent: Prof. Dr. Mag. Norbert Ulamec
Betreuer: Stefan Hüsemann, lic.rer.pol.

Eingereicht am 21. Mai 2002 von:

Tobias Teuffer
Rue de l'industrie 2
1700 Fribourg
tobias.teuffer@unifr.ch

Kurzfassung

Diese Diplomarbeit entstand im Zusammenhang mit dem *Börsenspiel Schweizer Universitäten* (BSU). Dieses Börsenspiel findet mehrmals jährlich statt, und bietet den Mitspielern¹ einen Einstieg in die Welt der Börse. Neu möchte die BSU ihre Dienstleistung ausbauen, und allen Teilnehmern die Möglichkeit bieten, alle während dem Spiel handelbaren Wertpapiere, die verschiedenen Portfolios und alle getätigten Transaktionen etwas genauer unter die Lupe zu nehmen.

Die Schlusskurse der Titel und alle Daten, welche während dem Spiel anfallen, sollen deshalb täglich in einem Data Warehouse gespeichert werden. Die Schlusskurse werden von *Swiss Exchange* (SWX) bereitgestellt, die Informationen bezüglich den Bewegungen innerhalb des Börsenspiels liefert die BSU selbst. Diese Datenmenge wird die Grundlage für die vorgesehene OLAP-Applikation bilden. Via Internet wird es den Mitspielern möglich sein auf diese gesammelten Daten zuzugreifen, um den Verlauf der Börse und die verschiedenen Anlagestrategien der Gruppen besser beobachten und beurteilen zu können. Zusätzlich wird es auch möglich sein, gängige Kennzahlen zu den Kursverläufen berechnen zu lassen.

Diese Dokumentation ist auf einem Theorie- und einem Praxisteil aufgebaut. Im Theorieteil werden hauptsächlich die verschiedenen Modellierungsebenen eines Data Warehouse vorgestellt. Es handelt sich dabei um die semantische, logische und physikalische Modellierungsebene, wobei insbesondere die verschiedenen Ansätze auf der logischen Ebene eingehend analysiert werden.

Das im Theorieteil erarbeitete Wissen wird im Praxisteil umgesetzt, um ein geeignetes Modell für das zukünftige Data Warehouse der BSU zu entwerfen. Schritt für Schritt wird gezeigt, wie das Modell die verschiedenen Phasen durchläuft, bis es implementiert werden kann.

Umgesetzt wird das fertige Modell schliesslich mit Hilfe von *SQL Server 2000*, auf dem die zugrundeliegende Datenbank gespeichert wird, und mit dem Programm *MicroStrategy 7*, das zur Entwicklung der OLAP-Applikation eingesetzt wird. Auch diese Phasen der Implementation werden im Praxisteil besprochen, sodass der aufmerksame Leser nach dem Studium dieser Diplomarbeit den Entstehungsprozess eines Data Warehouse, von der Modellierung bis zur Implementation, überblicken kann.

Keywords

Data Warehouse - SQL Server 2000 - MicroStrategy 7 - ETL-Prozess - OLAP - Börsenspiel - BSU - Wertpapiere - Kennzahlen - Online Analyse

¹ Das Wort *Mitspieler* steht für *Mitspieler und Mitspielerinnen*. Dieses Prinzip gilt für alle folgenden männlichen Formen.

Inhaltsverzeichnis

Kurzfassung	I
Inhaltsverzeichnis	II
Abbildungsverzeichnis	IV
1. Einleitung	1
1.1 Problemstellung	1
1.2 Zielsetzung	2
1.3 Vorgehensweise	3
2. Theorie	4
2.1 Was versteht man unter einem Data Warehouse?	4
2.2 Semantische Modellierungsebene	5
2.3 Logische Modellierungsebene	8
2.3.1 Star-Schema	8
2.3.2 Snowflake-Schema	10
2.3.3 Konsolidiertes Star-Schema	12
2.4 Physikalische Modellierungsebene	13
2.5 Aggregate	13
2.6 Metadaten	14
2.7 OLTP- versus DWH/OLAP-Systeme	15
2.8 OLAP Funktionalität	15
2.9 Das OLAP-Tool MicroStrategy 7	19
2.9.1 MicroStrategy Desktop	19
2.9.2 MicroStrategy Intelligence Server	19
2.9.3 MicroStrategy Web	20
2.10 Das Umfeld - die Börsenwelt	20
2.10.1 Die verschiedenen Arten von Wertpapieren	20
2.10.2 Typische Kennzahlen für Wertpapiere	21
2.10.3 Normalisierung der Indikatoren	25
2.11 BSU - Börsenspiel der Schweizer Universitäten	26
3. Praxis	28
3.1 Modellierung	28
3.1.1 Vorabklärungen für die Modellierung	28
3.1.2 Semantische Modellierungsebene des DWH	31
3.1.3 Logische Modellierungsebene des DWH	37

3.2 Implementation mit SQL Server 2000	41
3.2.1 Tabellendefinitionen	41
3.2.2 ETL-Prozess	45
3.3 Implementation mit MicroStrategy 7	58
3.3.1 Warehouse Katalog	59
3.3.2 Fakten	59
3.3.3 Attribute	60
3.3.4 Metriken	61
3.3.5 Filter	63
3.3.6 Transformationen	64
3.3.7 Berichte	64
3.3.8 Promts	65
3.3.9 MicroStrategy Web	66
3.4 Duplizierung des Projekts	68
4 Bedienungsanleitungen	69
4.1 Bedienungsanleitung für die Administration des DWH	69
4.1.1 Aktivierung des DWH vor einem neuen Spiel	69
4.1.2 Aktivierung des DWH während einem Spiel	71
4.1.3 Administration nach einem Spiel	72
4.1.4 Passwörter	73
4.1.5 Bekannte Probleme und ihre Behebung	73
4.2 Bedienungsanleitung für die Anwender des DWH	74
4.2.1 Beschreibung der verschiedenen Berichte	74
4.2.2 Zusätzliche Funktionen innerhalb von Berichten	76
5 Zusammenfassung und Schlussfolgerungen	77
6 Begriffserklärungen	79
7 Literaturverzeichnis	80
Anhang I: Importtabellen	81
Anhang II: DTS-Pakete	83
Anhang III: Gespeicherte Prozeduren	93
Ehrenwörtliche Erklärung	107

1. Einleitung

1.1 Problemstellung

Die 1991 gegründete Organisation *Börsenspiel Schweizer Universitäten*¹ (BSU) bietet Studierenden und Lehrlingen einmal im Jahr die Möglichkeit, virtuell an der Börse zu handeln. Neu soll den Teilnehmern ermöglicht werden, alle im Spiel handelbaren Wertpapiere, die Transaktionen und die verschiedenen Portfolios mit einem OLAP-Tool über das Internet zu analysieren. Diese Analysen umfassen auch die Berechnung verschiedener, aus dem Finanzmanagement stammender Kennzahlen.

Die Aufgabe besteht darin, ein Data Warehouse zu entwerfen und zu implementieren, das diesen Anforderungen gerecht wird. Für die Implementation stehen die beiden Programme *SQL Server 2000* und *MicroStrategy 7* zur Verfügung. Damit den Anwendern immer die neusten Börsendaten zur Verfügung stehen, muss das Data Warehouse täglich automatisch aktualisiert werden.

Die Bedienung und Administration des Data Warehouse muss einfach und mit so wenig Aufwand wie möglich verbunden sein. Der Bericht, der die Entstehung dieses Data Warehouse dokumentiert, muss sowohl das theoretische als auch praktische Wissen vermitteln, das für die Modellierung und Implementation eines Data Warehouse unabdingbar ist.

Es folgt die Auflistung der wissenschaftlichen Fragen, die später in diesem Bericht beantwortet werden:

1. Welches ist für dieses Projekt die am besten geeignete Modellierung des Data Warehouse?

Diese Frage ist von grosser Wichtigkeit, schliesslich baut das ganze Projekt auf der Modellierung auf. Um diese Frage beantworten zu können ist eine gründliche Einführung in die Theorie der Data Warehouse-Modellierung unabdingbar. In Kapitel 2 *Theorie* werden sowohl die verschiedenen Modellierungsphasen als auch die unterschiedlichen Modellierungsarten vorgestellt.

In Kapitel 3 *Praxis* wird anschliessend beschrieben und begründet, welches Modell sich als das geeignete herausgestellt hat.

2. Welche Arten von Online-Analysen machen Sinn und welche Kennzahlen (Metriken) aus dem Finanzmanagement sollen berücksichtigt werden?

¹ <http://www-bsu.unifr.ch>

In Kapitel *2.10 Das Umfeld - Die Börsenwelt* werden verschiedene Kennzahlen vorgestellt, die sich in der Börsenwelt zur Analyse von Kursverläufen verbreitet und etabliert haben. Es wird erklärt was für Arten von Indikatoren es gibt, wie sie funktionieren und was sie aussagen. In Kapitel *3 Praxis* wird versucht diese Indikatoren zu implementieren, dabei wird sich zeigen welche Kennzahlen für die Einbindung in dieses Projekt geeignet sind.

3. Wie kommen die Daten in das Data Warehouse (Extraction Transforming Loading)?

Eine kurze theoretische Einführung in den sogenannten ETL-Prozess befindet sich im Kapitel *2.1 Was versteht man unter einem Data Warehouse?* Wie und in welcher Reihenfolge die täglich neu anfallenden Daten dann tatsächlich in das Data Warehouse importiert werden, steht im Kapitel *3.2.2 ETL-Prozess*.

Wie die Daten aus der Sicht des Administrators in das Data Warehouse gelangen ist im Kapitel *4.1 Bedienungsanleitung für die Administration des DWH* nachzulesen.

4. Wie kann der Administrationsaufwand des Data Warehouse möglichst gering gehalten werden?

Da ein Spiel durchschnittlich nur ungefähr zwei Monate dauert, ändern sich die Daten im Warehouse relativ häufig. Damit der administrative Aufwand auch bei einem Spielwechsel möglichst gering bleibt, muss für diese Situation eine durchdachte Lösung gefunden werden. Das Kapitel *3.1.1 Vorabklärungen für die Modellierung* analysiert diesbezüglich verschiedene Problemfelder und versucht die optimalste Variante zu ermitteln.

Der Administrationsaufwand an sich wird im Kapitel *4.1 Bedienungsanleitung für die Administration des DWH* ersichtlich.

1.2 Zielsetzung

Das Hauptziel ist ein Data Warehouse zu entwickeln, das automatisch täglich die neusten Börsenkurse und Transaktionen der Teilnehmer des Börsenspiels importiert. Via Internet soll auf diese relativ grosse, ständig wachsende Datenmenge zugegriffen werden können, um sie nach Belieben zu analysieren. Bei der Analyse sollen Tabellen, Graphiken und verschiedene Kennzahlen zum Einsatz kommen, um die verschiedenen Daten zu visualisieren. Die Bedienung des OLAP-Tools soll möglichst einfach und intuitiv sein, damit bei einer eventuellen Einbindung in das Börsenspiel alle Teilnehmer davon profitieren können. Um obenstehendes Hauptziel zu erfüllen müssen mehrere Nebenziele erreicht werden. Von besonderer Wichtigkeit ist die Beantwortung der wissenschaftlichen Fragen, die im Kapitel *1.1 Problemstellung* gestellt wurden. Es gilt also das geeignete Modell für dieses Projekt zu finden, sinnvolle und aussagekräftige Online-Analysen zur Verfügung zu stel-

len, das Data Warehouse soweit wie möglich zu automatisieren und dabei den Administrationsaufwand so gering wie möglich zu halten.

Der Bericht soll dem interessierten Leser einen Einstieg in die Welt des *Data Warehousing* bieten, aufgeteilt in einen Theorie- und einen Praxisteil. Nach der Lektüre des Theorieteils soll der Leser unter anderem verstanden haben, wie man ein Modell für ein Data Warehouse entwirft, welches die in der Praxis verbreiteten Kennzahlen zur Analyse von Kursverläufen sind und wie diese funktionieren.

Der Praxisteil soll zeigen, wie sich das theoretische Wissen in die Praxis übertragen lässt. Es soll ersichtlich werden, wie man mit der Kombination von *SQL Server 2000* und *MicroStrategy 7* ein Data Warehouse-Modell implementieren kann, und welche Problemfelder dabei auftauchen können.

1.3 Vorgehensweise

Die Vorgehensweise kann in vier verschiedene Phasen unterteilt werden:

- Die erste Phase besteht darin Informationen zu sammeln, Bücher zu lesen und das ganze Projekt zu überblicken. Literatur aus dem Bereich der Wirtschaft (für das Verständnis der Börse und der verbreiteten Kennzahlen) und Literatur aus dem Bereich des *Data Warehousing* kommt dabei zum Einsatz.²
- Die zweite Phase betrifft die Modellierung. Mit Hilfe des angeeigneten Wissens muss das geeignete Modell für dieses Projekt gefunden und erarbeitet werden, welches den Anforderungen so gerecht wie möglich wird.
- In der dritten Phase wird das entworfene Modell implementiert. Zuvor müssen natürlich der Funktionsumfang und die jeweiligen Möglichkeiten der beiden Programme *SQL Server 2000* und *MicroStrategy 7* kennengelernt und verstanden werden. Danach gilt es zuerst die modellierten Tabellen auf dem *SQL Server 2000* zu implementieren, um danach die OLAP-Anwendung mit *MicroStrategy 7* darauf aufzusetzen.
- Die vierte Phase betrifft das Verfassen dieses Dokuments, welches die einzelnen Phasen beschreibt und alle gewonnenen Erkenntnisse sammelt.

² In Kapitel 7 *Literaturverzeichnis* sind alle konsultierten Bücher aufgeführt.

2 Theorie

Dieses Kapitel widmet sich der Data Warehouse Theorie. Hier wird erklärt was ein Data Warehouse ist, wie es aufgebaut ist und wie es modelliert wird. Das hier gesammelte theoretische Wissen wird später im Kapitel 4 *Praxis* angewendet, um das Data Warehouse für die BSU zu entwerfen und zu implementieren.

2.1 Was versteht man unter einem Data Warehouse?

Der Begriff Data Warehouse wurde erstmals von W. H. Inmon [INMO92] geprägt:

"A Data Warehouse is a subject-oriented, integrated, time-variant, and nonvolatile collection of Data in support of managements Decision support process."

Die Definition von A. Kurz [KURZ98] lautet wie folgt:

"Ein Data Warehouse repräsentiert ein, von den operativen Datenbanken getrennte Decision Support-Datenbank (Analyse-Datenbank), die primär zur Unterstützung des Entscheidungsprozesses im Unternehmen genutzt wird. Ein Data Warehouse wird immer multidimensional modelliert und dient zur langfristigen Speicherung von historischen, bereinigten, validierten, synthetisierten, operativen, internen und externen Datenbeständen."

Ein Data Warehouse ist also dazu da, um möglichst viele aussagekräftige, bereinigte, historische und aktuelle Daten zu sammeln, um sie optimal analysieren zu können. Oft wird diese Technologie zur Bereitstellung von Informationen für Kontroll- und Entscheidungsprozesse in Unternehmen eingesetzt. Dabei werden die vielen, täglich anfallenden Daten des Unternehmens zentral gespeichert. Dank der Analyse dieser Daten können Mängel in Bereichen wie Kundenbetreuung, Kundenbindung, Marketing, Qualitätssicherung, usw. frühzeitig entdeckt und behoben werden. Natürlich kann ein Data Warehouse aber grundsätzlich überall eingesetzt werden, wo viele Daten anfallen, so wie das beim *Börsenspiel Schweizer Universitäten* der Fall ist. Das Ziel bleibt dabei immer dasselbe, die scheinbar unüberblickbare Fülle an Daten trotzdem zu überblicken und Tendenzen frühzeitig zu erkennen, um daraus einen Wissensvorsprung gegenüber der Konkurrenz zu erzielen.

Die unterste und somit erste Schicht der allgemeinen technischen Reverenzarchitektur eines Data Warehouse (Abb. 2-1) bilden die Quelldatensysteme. Sie versorgen das Data Warehouse mit Informationen. Es ist wichtig, dass die im Warehouse gesammelten Daten physikalisch von den Quelldatensystemen getrennt werden. Somit entsteht zwar Redundanz, da dieselben Informationen an verschiedenen Orten gespeichert sind, dafür werden aber bei den rechenintensiven Analysen die Quelldatensysteme nicht belastet.

Die zweite Schicht besteht aus den Tabellen, in denen die Daten und Metadaten gespeichert sind. Auch der *Extraction Transformation Loading* Prozess, kurz ETL-Prozess gehört dieser Ebene an. Der ETL-Prozess ist verantwortlich für Selektion, Transformation, Bereinigung und Laden der relevanten Daten aus den operativen Quelldatensystemen. Der ETL-Prozess ist ereignisgesteuert oder findet periodisch statt. Aus Erfahrung weiss man, dass die Implementation dieser Schicht sehr arbeitsintensiv ist und einen erheblichen Teil des Gesamtaufwandes eines Data Warehouse Projektes in Anspruch nimmt.

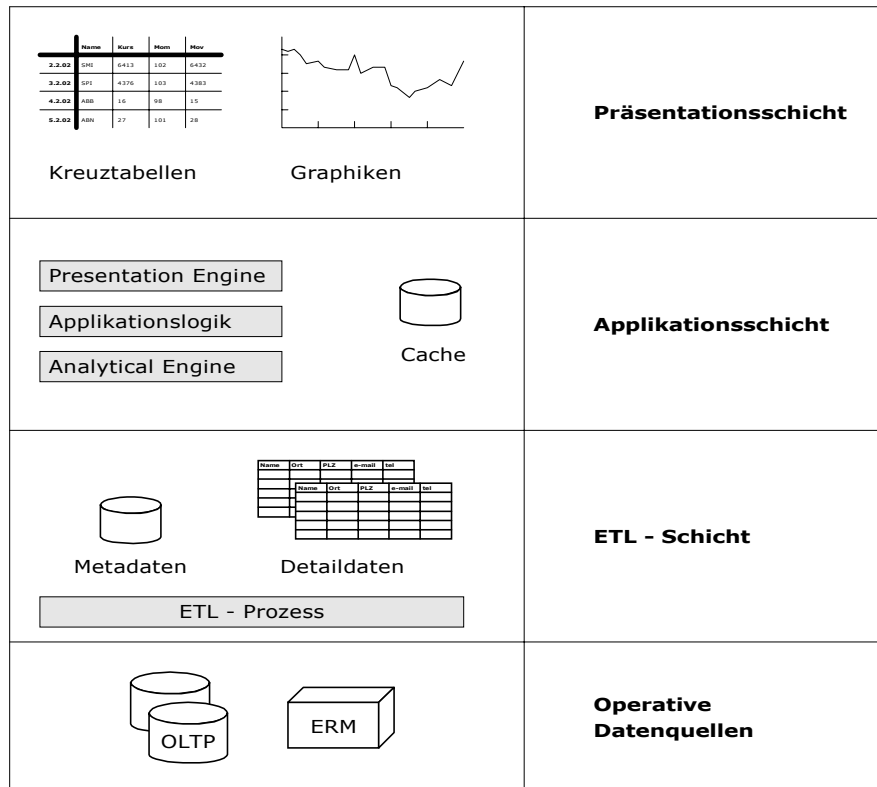


Abb. 2-1, Technische Referenzarchitektur, [vgl. KURZ99]

Die Applikationsschicht übersetzt die multidimensionalen Abfragen der Anwender in eine Abfragesprache (z.B. ANSI-SQL/92) und liefert die Ergebnisdatensätze zurück. Dieser Übersetzungsprozess wird durch die Metadaten gesteuert, zu jedem Data Warehouse Objekt existiert nämlich auch eine entsprechende Metadaten-Objektbeschreibung. Meistens verfügt diese Schicht auch über einen Zwischenspeicher (Cache). Dieser verbessert das Antwortzeitverhalten des Systems nachhaltig.

Die Präsentationsschicht ist die oberste dieser vier Schichten, sie ist verantwortlich für die Informationsvisualisierung. Sie muss möglichst gut auf den Anwender und seine Bedürfnisse zugeschnitten sein. Um die Ergebnisse aussagekräftig darzustellen bieten stehen Kreuztabellen und Graphiken zur Verfügung.

2.2 Semantische Modellierungsebene

Die semantische Modellierungsebene ist die erste von insgesamt drei Modellierungsebenen. Sie dient als eine Art Schnittstelle zwischen der Sicht der Anwender und der Sicht des Entwicklers auf das Data Warehouse Projekt. Wie der Name bereits sagt, werden im semantischen Datenmodell die semantischen (inhaltlichen) Zusammenhänge veranschaulicht.

Aus der Perspektive des Anwenders muss das Modell intuitiv verständlich und leicht zu erfassen sein. Der Entwickler hingegen strebt nach einem möglichst genauen Modell mit hohem Detaillierungsgrad. Das semantische Modell bietet die Diskussionsgrundlage für diese beiden unterschiedlichen Sichtweisen.

Die hauptsächliche Aufgabe besteht darin, die Fakten und Dimensionen für das Projekt zu definieren. Als Fakten gelten allgemein betriebswirtschaftliche Kennzahlen, wie Erlöse, Gewinne oder Verluste eines Unternehmens. Auch der *Kurswert* einer Aktie ist ein Faktum, oder etwa die *Performance* zum Vortag. Zwischen diesen beiden Fakten besteht jedoch ein kleiner, interessanter Unterschied. Der *Kurswert* ist in einer Tabelle physikalisch festgehalten, die *Performance* jedoch ist nirgends gespeichert, weil sie berechnet werden kann. Nur die Art und Weise, wie die *Performance* berechnet wird, ist in den Metadaten abgespeichert. Bereits in der semantischen Modellierungsebene werden die gespeicherten und die berechneten Fakten (auch Kennzahlen genannt) unterschieden.

Die Dimensionen (Auswertedimensionen) definieren die möglichen Sichten des Anwenders auf die Fakten. Eine typische Dimension ist die Zeit. Für die oben kurz angesprochene Kennzahl *Performance* heisst das, dass man sie entlang der Zeitachse auswerten kann. Somit kann sie für jeden beliebigen Tag angezeigt werden. Aber auch die monatliche oder jährliche Performance kann berechnet werden. Eine Dimension besteht nämlich aus einem sogenannten Aggregationspfad, der die verschiedenen Ebenen innerhalb der Dimension definiert (Abb. 2-2).

Im folgenden wird die einfache und intuitiv verständliche Darstellungsform von A. Kurz [KURZ98] verwendet. Diese graphische Darstellungsform besteht aus einer übergeordneten Überblicksgraphik mit den Fakten im Zentrum und mehreren einzelnen detaillierten Aggregationsgraphen zu den Dimensionen. Die Überblicksgraphiken beschreiben immer eine Kennzahlengruppe. Abbildung 2-2 zeigt die Kennzahlengruppe *Kurse* und zwei dazugehörige Dimensionen *Datum* und *Titel*.

Der Abbildung kann entnommen werden, dass die Kennzahlengruppe *Kurse* sowohl aus gespeicherten als auch berechneten Fakten besteht. Die beiden Pfeile weisen darauf hin, dass diese Kennzahlengruppe entlang der beiden Dimensionen *Titel* und *Datum* ausgewertet werden können. Den kleinen Kreisen innerhalb der Dimensionen ist immer ein Name zugeordnet. Sie stehen für die verschiedenen Ebenen oder Blickwinkel, von denen aus die Kennzahlen betrachtet werden können, und werden Hierarchieobjekte genannt.

Die Verbindungen unter den Hierarchieobjekten stellen den Aggregationspfad dar. Entlang dieses Pfades kann man sich durch eine Dimension hindurchbewegen. Es ist also möglich, die *Relative Stärke* - oder ein beliebiges anderes Faktum - für eine ganze Kategorie berechnen zu lassen, oder für eine Branche innerhalb dieser Kategorie und selbstverständlich auch für einzelne Titel.

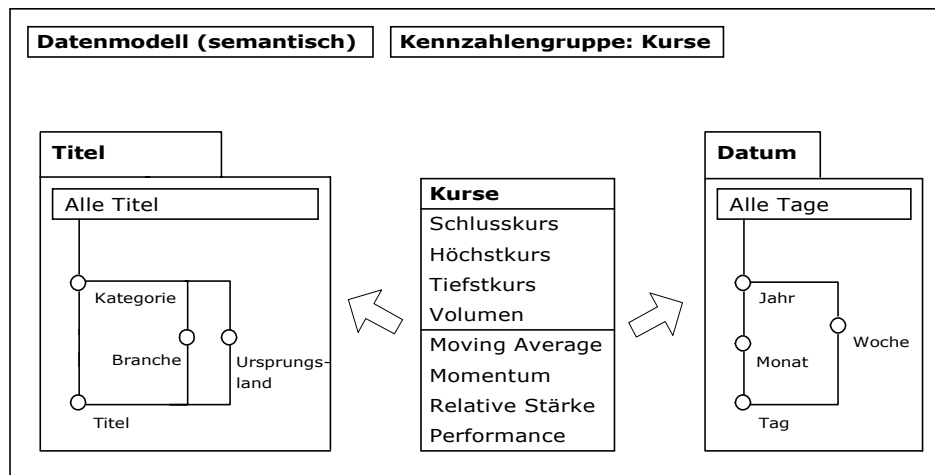


Abb. 2-2, Überblicksgraphik, Semantische Modellierungsebene

Es ist möglich, dass die verschiedenen Kennzahlen innerhalb einer Kennzahlengruppe unterschiedliche Hierarchieobjekte einer assoziierten Dimension ansprechen. An dieser Stelle ist das semantische Modell nicht exakt. Die Kennzahl *Schlusskurs* zum Beispiel wird immer auf dem Hierarchieobjekt *Tag* der Dimension *Datum* basieren. Der *Höchstkurs* hingegen kann mit allen Hierarchieobjekten innerhalb der Dimension *Datum* assoziiert werden.

In einem Projekt wird für jede Kennzahlengruppe eine Überblicksgraphik erstellt. Somit kann der Anwender die Möglichkeiten der zukünftigen Analyse überblicken, auf fehlende Kennzahlen aufmerksam machen oder eigene Ideen miteinbringen.

Weil die Dimensionen in den Überblicksgraphiken im Detaillierungsgrad beschränkt sind, wird jede Dimension in einer separaten Graphik noch einmal genauer beschrieben. Wie das aussieht zeigt Abbildung 2-3 für die Dimension *Datum*.

Diese Darstellung liefert mehr Informationen zu den einzelnen Hierarchieobjekten. Es wird bereits vermerkt, was genau pro Hierarchieobjekt gespeichert wird. Die Informationen über das Jahr beschränken sich in diesem Beispiel auf die Jahreszahl. Beim Monat wird nebst der Nummer des Monats im aktuellen Jahr (1-12) auch die Bezeichnung mitgespeichert. Wie man Abbildung 2-3 entnehmen kann, werden diese Informationen zu den einzelnen Hierarchieobjekten in Rechtecken dargestellt und mit dem betreffenden Knoten verbunden.

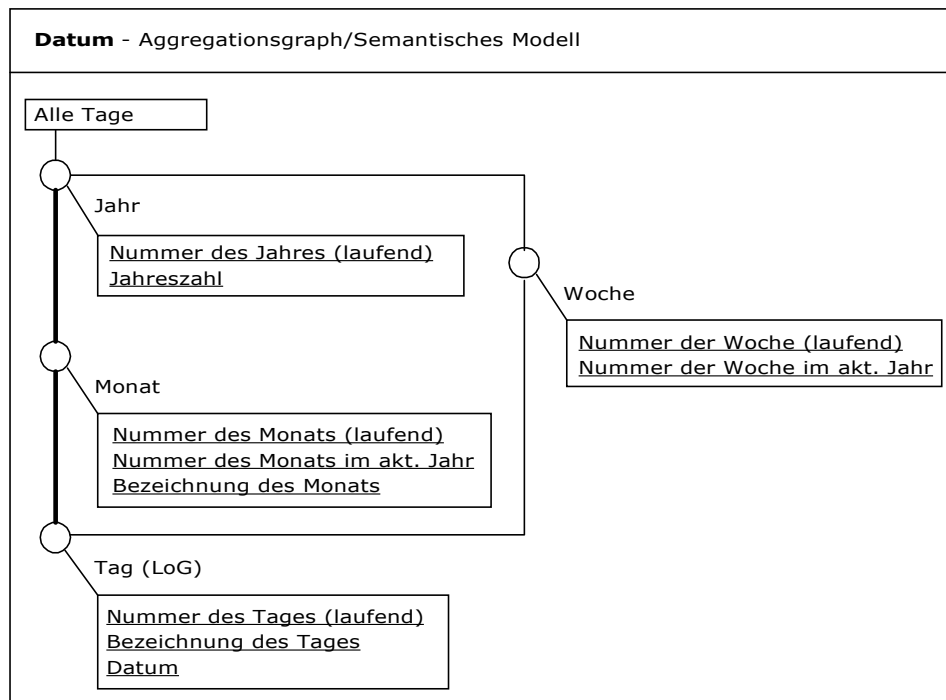


Abb. 2-3, Aggregationsgraph, Semantische Modellierungsebene

Jede Dimension besitzt einen Standard-Aggregationspfad, der etwas dicker dargestellt wird. Der Standard-Aggregationspfad der Dimension *Datum* verläuft entlang den Hierarchieobjekten *Jahr*, *Monat* und *Tag*. Der einzelne Tag repräsentiert den höchsten Detaillierungsgrad der Dimension *Datum* und wird deshalb LoG-Knoten (Level of Granularity) genannt. In der Dimension *Titel* wäre der einzelne Titel der LoG-Knoten.

2.3 Logische Modellierungsebene

Die logische Modellierungsebene wird vom semantischen Datenmodell abgeleitet. Im Unterschied zur semantischen Modellierungsebene wird hier die Sicht des Entwicklers auf das Projekt repräsentiert. In den folgenden Abschnitten werden die verbreiteten logischen Schemata vorgestellt, mit denen diese Ebene modelliert wird.

2.3.1 Star-Schema

Das *Star-Schema* kann als Standardform der logischen Modellierungsebene für ein relationales Data Warehouse angesehen werden. Pro Kennzahlengruppe steht eine hochnormalisierte Faktentabelle im Zentrum, die von mehreren denormalisierten (redundanten) Dimensionstabellen umgeben ist. Jede Dimension wird durch genau eine relationale Tabelle logisch abgebildet. Die zentrale Faktentabelle ist über eindeutige Schlüssel mit den umgebenden Dimensionstabellen verbunden. Der Primärschlüssel der Faktentabelle setzt sich aus allen Schlüssel der assoziierten Dimensionen zusammen.

Erst dieser zusammengesetzte Primärschlüssel identifiziert einen Eintrag eindeutig. Die Primärschlüssel der Dimensionstabellen werden in der Regel künstlich generiert, mittels fortlaufender Nummerierung.

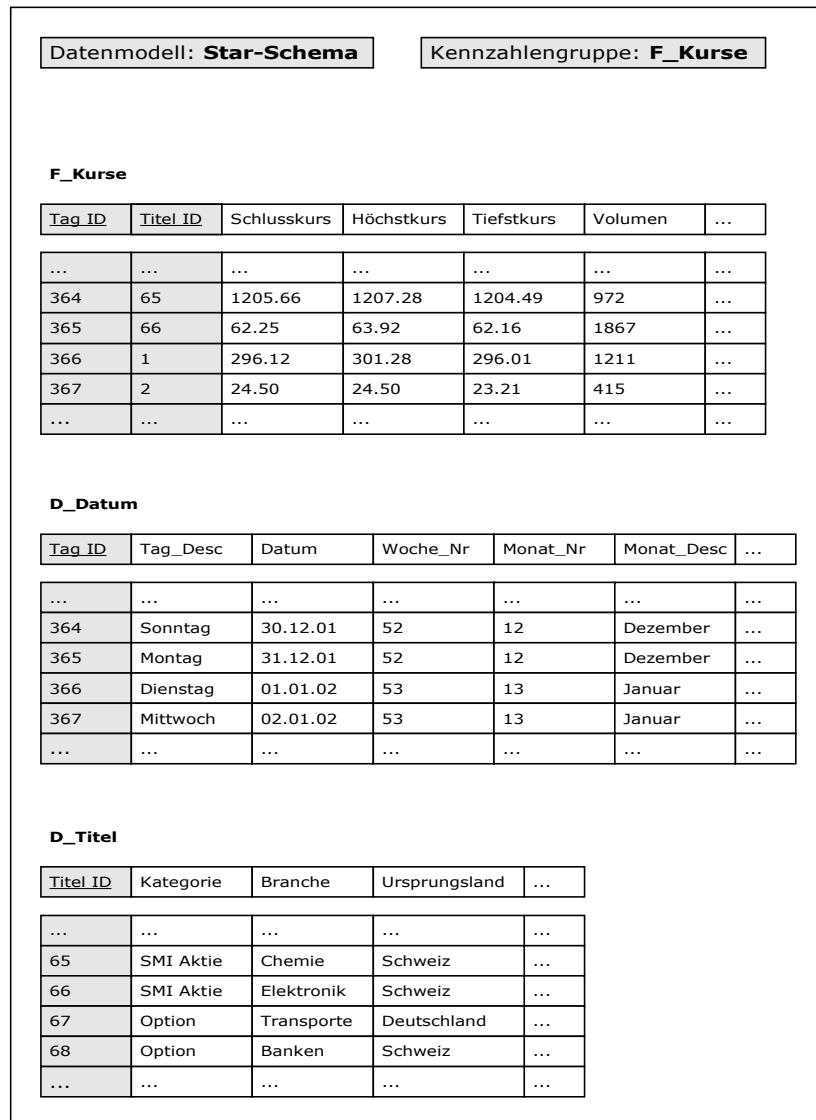


Abbildung 2-4, Star-Schema, Logische Modellierungsebene

Abbildung 2-4 zeigt, wie das entsprechende *Star-Schema* der in Abbildung 2-2 dargestellten semantischen Überblicksgraphik aussieht. Um Übersicht zu gewinnen sind die Schlüsselspalten in der Graphik grau dargestellt und den Dimensions- und Faktentabellen wurde der Präfix 'D_' beziehungsweise 'F_' vorangestellt. In der Graphik kann man erkennen, dass in den Dimensionen redundante Einträge vorhanden sind, die Faktentabelle jedoch hochnormalisiert ist. Die Redundanz in den Dimensionstabellen ist beabsichtigt, auf diese Weise werden nicht unnötig viele Tabellen gebraucht um eine Dimension darzustellen. Da die Grösse der Dimensionstabellen begrenzt ist, muss man sich auch keine Gedanken über den verschwendeten Speicherplatz machen, den die Redundanz mit sich

bringt. Nur die Faktentabellen werden mit der Zeit anwachsen, diese sind jedoch wie bereits erwähnt normalisiert.

Zwischen den Dimensionen und der Faktentabelle bestehen immer 1:N Beziehungen. In der Tabelle *Datum* gibt es beispielsweise pro Tag genau einen Eintrag, während in der Tabelle *F_Kurse* derselbe Tag mehrmals vorkommen kann.

Vorteile des Star-Schemas:

- Das Datenmodell ist einfach und intuitiv, weil für die Fakten und die verschiedenen Dimensionen je eine Tabelle genügt.
- Bei Abfragen werden nur wenige Join-Operationen benötigt, weil jede Dimension mit einer Tabelle abgebildet wird.
- Da nur wenige physikalische Tabellen existieren ist auch der Wartungsaufwand entsprechend gering.
- Weil das Datenmodell einfach ist, ist auch die Definition der Metadaten einfach.

Nachteile des Star-Schemas:

- Die Bildung von Aggregaten (siehe Kapitel 2.5 *Aggregate*) ist heikel, weil es zu falschen Aggregatbildungen durch mehrfaches Zählen von Einträgen kommen kann.
- Bei sehr grossen Dimensionstabellen kann es bei häufiger Abfrage von Elementen der Hierarchieobjekten zu Einbussen des Antwortzeitverhaltens kommen (weil die Dimensionstabellen denormalisiert sind).

2.3.2 Snowflake-Schema

Der Unterschied zwischen dem *Snowflake-Schema* und dem *Star-Schema* liegt primär darin, dass die Dimensionstabellen nicht mehr nur mit einer Tabelle umgesetzt werden, sondern auch normalisiert werden.

In Abbildung 2-5 ist die Kennzahlengruppe Kurse und die normalisierte Dimension *Datum* dargestellt, welche im *Snowflake-Schema* nicht aus einer, sondern aus vier Tabellen besteht. Die Primärschlüssel sind wie beim *Star-Schema* grau hinterlegt. Der Präfix *LU_* steht für Lookup-Tabelle. Aus Platzgründen wurde auf die Darstellung der Dimensionstabelle *Titel* verzichtet.

Das *Snowflake-Schema* hat auf die Faktentabelle *Kurse* keinen Einfluss, sie ist immer noch über dieselben Fremdschlüssel mit den Dimensionen verbunden. In der Tabelle *LU_Tag* sind jetzt nicht mehr alle Informationen der Dimension *Datum* gespeichert, möchte man zum Beispiel auf die Beschreibung des Jahres zugreifen, gelangt man nur noch via dem Fremdschlüssel *Jahr_ID* der Tabelle *LU_Woche* oder *LU_Monat* zu dieser Information. Es sind also im Vergleich zum *Star-Schema* deutlich mehr Join-Operationen innerhalb der Dimensionstabellen nötig.

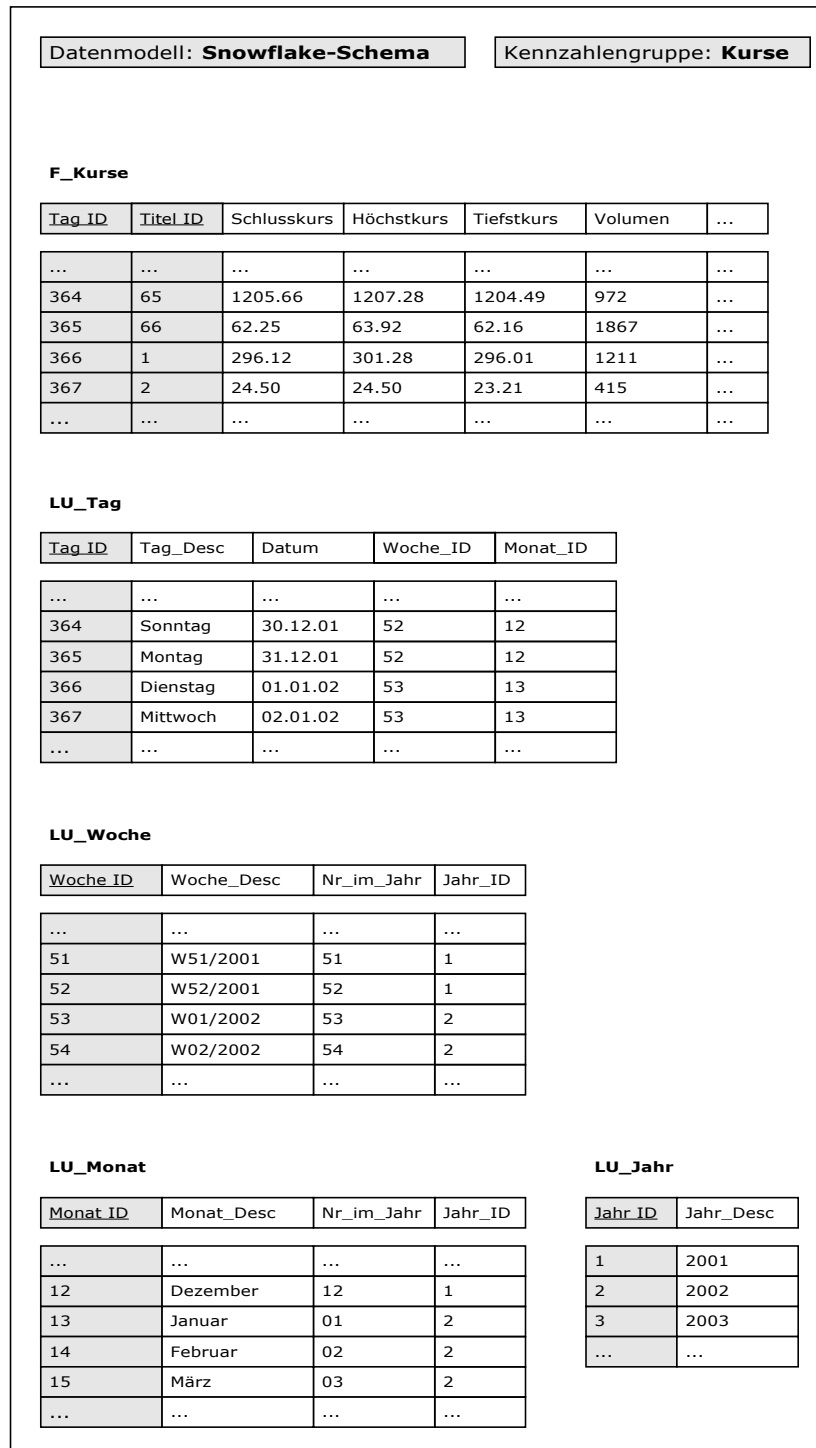


Abb. 2-5, Snowflake-Schema, Logische Modellierungsebene

Natürlich kann die Zahl der Join-Operationen in den Lookup-Tabellen einfach reduziert werden. Man speichert auf der Ebene des LoG-Knotens einfach alle Schlüssel der Lookup-Tabellen dieser Dimension, damit man direkt auf diese zugegriffen kann. Für die Tabelle *LU_Tag* würde das bedeuten, dass zusätzlich noch eine Fremdschlüsselbeziehung zu der Tabelle *LU_Jahr* gespeichert werden müsste.

Vorteile des Snowflake-Schemas:

- Die Bildung von Aggregaten ist einfach, weil pro Hierarchie eine Tabelle gespeichert wird.
- Die Dimensionstabellen enthalten keine redundanten Einträge, sie werden bis zur 3. Normalform normalisiert.

Nachteile des Snowflake-Schemas:

- Die Anzahl der physikalischen Data Warehouse Tabellen ist grösser, was sich negativ auf den Wartungsaufwand auswirkt.
- Es werden mehr Join-Operationen benötigt, weil es mehr Tabellen gibt.
- Die Definition der Metadaten gestaltet sich schwieriger, weil die Struktur des ganzen Datenmodells komplexer ist.

Laut A.Kurz [KURZ99] ist das normalisierte (3. Normalform) *Snowflake-Schema* trotzdem die beste Wahl, weil es die Bildung von Aggregaten gut unterstützt und keine redundanten Einträge in den einzelnen Lookup-Tabellen der Hierarchieelemente enthält. Die oben besprochene Optimierung der Join-Operationen innerhalb einer Dimension wird selbstverständlich auch vorgenommen.

2.3.3 Konsolidiertes Star-Schema

Die Basisidee dieses Schemas besteht darin, Aggregate bereits in der zentralen Faktentabelle zu speichern. Dies erfordert den Einsatz von sogenannten Level-Attributen in den assoziierten Dimensionstabellen. Da alle Aggregatfunktionen bereits im Voraus berechnet wurden, zeichnet sich das *Konsolidierte Star-Schema* durch ein sehr gutes Antwortzeitverhalten aus.

Der Tabelle *D_Datum* wurde das zusätzliche Feld *H-Level* hinzugefügt, dieses gibt Aufschluss über den Grad der Konsolidierung. Der Eintrag mit der ID = 366 in Abbildung 2-7 ist kein einzelner Tag, sondern steht für eine ganze Woche, diese ID fasst den ganzen Monat Dezember des Jahres 2001 zusammen.

In der Tabelle *F_Kurse* werden die Ergebnisse dieser Aggregationen physikalisch gespeichert. Deshalb kann zum Beispiel das Volumen eines Titels für jede Aggregationsebene (Tag, Woche, Monat, Jahr) der Dimension *D_Titel* ohne Berechnung ausgelesen werden.

Vorteile des Konsolidierten Star-Schemas:

- Sehr gutes Antwortzeitverhalten, weil die Aggregationen bereits im Voraus berechnet wurden.
- Bei Abfragen werden nur wenige Join-Operationen benötigt, weil jede Dimension mit einer Tabelle abgebildet wird.

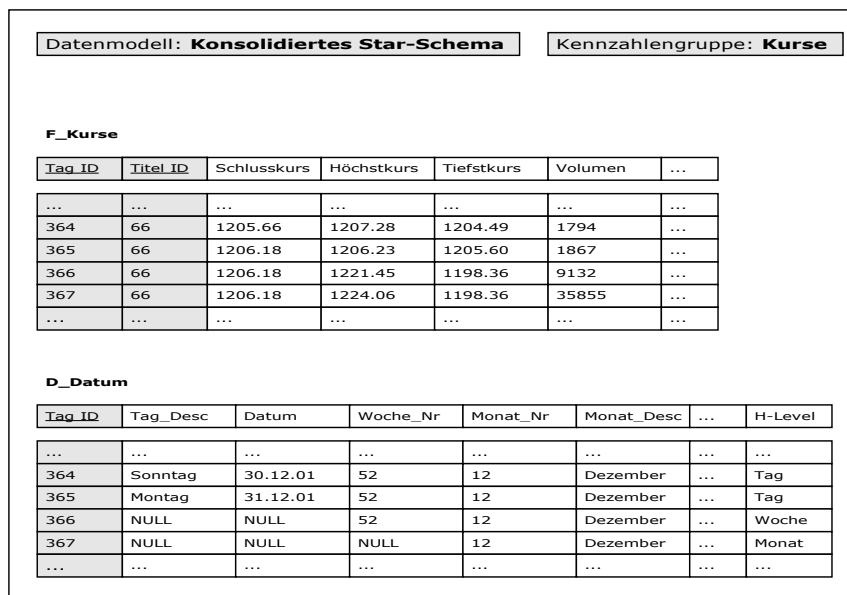


Abbildung 2-6, Konsolidiertes Star-Schema, Logische Modellierungsebene

Nachteile des Konsolidierten Star-Schemas:

- Sehr grosser Aufwand bei der Wartung des Data Warehouse, da der ETL-Prozess sehr komplex ist.
- Alle möglichen Aggregate müssen vorab berechnet werden, was bei einer grösseren Anzahl von Dimensionen sehr zeitintensiv wird.
- Die Grösse der Faktentabelle wird durch die Aggregate stark vergrössert.

2.4 Physikalische Modellierungsebene

Die letzte der drei Modellierungsebenen ist die physikalische Ebene, sie wird vom logischen Datenmodell abgeleitet und individuell auf das spezifische Datenbank Management System (DBMS) angepasst. Die Grundstruktur der Tabellen ist bereits definiert, zu definierten bleiben noch die Datentypen der verschiedenen Felder der Tabellen.

Der automatisierte Import der Daten aus den Quelldatensystemen (ETL-Prozess) sowie die Entwicklung der OLAP-Anwendung gehören auch zur physikalischen Modellierungsebene, welche beim Aufbau eines Data Warehouse am meisten Zeit in Anspruch nimmt.

2.5 Aggregate

Die Hierarchieobjekte innerhalb einer Dimension können bekanntlich entlang einem definierten Aggregationspfad durchlaufen werden, wobei der LoG-Knoten den höchsten Detaillierungsgrad aufweist. Alle anderen Hierarchieobjekte erlauben eine aggregierte Sicht auf die Kennzahlengruppe. Ob diese Aggregate erst berechnet werden, wenn man sie

benötigt, oder ob sie bereits vorab berechnet und gespeichert wurden, hat einen Einfluss auf das Antwortzeitverhalten des Data Warehouse.

Aggregate werden pro Hierarchieobjekt und zentraler Faktentabelle gebildet und meistens als separate Tabellen gespeichert. Vergleicht man die unterschiedliche Repräsentationen einer Dimension im Bezug auf die Wahl des logischen Modells (vgl. Abb. 2-4 und 2-5), wird schnell ersichtlich, dass das *Snowflake-Schema* besser für Aggregationen geeignet ist als das *Star-Schema*. Im *Snowflake-Schema* können die verschiedenen Hierarchie-Objekte (z.B. Woche, Monat) mit einer eindeutigen ID bezeichnet werden, was beim *Star-Schema* nicht möglich ist. Diese ID ist von grossem Vorteil, da sie in einer Aggregationstabelle als Fremdschlüssel dienen kann.

Unabhängig davon ob Aggregate in speziellen Tabellen gespeichert werden, oder erst bei Gebrauch berechnet werden, können nicht alle Fakten auf dieselbe Art und Weise verdichtet werden. Die Aggregation der Kennzahlen *Höchstkurs* und *Volumen* (aus der Faktentabelle *Kurse*) entlang der Zeit-Dimension unterscheidet sich darin, dass für die Berechnung des Höchstkurses der höchste Wert der Kennzahl *Höchstkurs* in das Aggregat aufgenommen werden muss, wobei die Aggregation des Volumens aus der Summe der einzelnen Volumina besteht.

Physikalisch gespeicherte Aggregate verbessern zwar das Antwortzeitverhalten drastisch, verursachen jedoch mehr Wartungsaufwand und einen erhöhten Speicherplatzbedarf.

2.6 Metadaten

Metadaten sind Daten über Daten, also Informationen über die Datenbestände eines Data Warehouse, ähnlich einem Inhaltsverzeichnis. Die Struktur und der Inhalt des Data Warehouse befinden sich in den Metadaten. Die operativen Datenquellen und die Transformationsregeln werden mittels der zentralen Metadaten beschrieben, aber auch das Navigieren in den verschiedenen Dimensionen und Hierarchieobjekten wird von ihnen gesteuert. Schliesslich ist auch noch das Sicherheitskonzept eines Data Warehouse, das die Rechte der Anwender definiert, in den Metadaten abgelegt.

Die Metadaten übernehmen demnach viele Aufgaben, bilden den Fugenkitt, der die unterschiedlichen Komponenten der Data Warehouse Architektur zu einer Einheit verbindet, und gleichzeitig dienen sie dem Anwender als eine Art Strassenkarte durch die Menge der historischen und aktuellen Detaildaten.

Es existiert eine Vielzahl von unterschiedlichen Metadaten-Beschreibungsformaten. Zwar gab es einige Bestrebungen einen Standard zu definieren, jedoch scheiterten diese zu meist an den unterschiedlichen Herstellerinteressen. Es könnte jedoch sein, dass sich die *eXtensible Markup Language* (XML) als neuer Metadaten-Standard durchsetzen wird und somit unterschiedliche Komponenten (auch von verschiedenen Herstellern) eines Data Warehousing-Systems besser miteinander kommunizieren können.

2.7 OLTP- versus DWH/OLAP-Systeme

OnLine Transactional Processing (OLTP) -Systeme dienen dazu, operative Daten effizient und zuverlässig zu verwalten. Beim Datenbankentwurf steht hauptsächlich die effiziente Speicherung der Daten im Vordergrund. Mit Hilfe eines Transaktionskonzeptes (Commit und Rollback-Mechanismen) wird sichergestellt, dass einzelne Transaktionen entweder vollständig oder gar nicht durchgeführt werden. Dies hat zur Folge, dass die Integrität der Daten gewährleistet ist. OLTP-Systeme bestehen normalerweise aus mehreren Tabellen, die jeweils in dritter Normalform sind. Typisch für OLTP-Systeme ist, dass nur begrenzte Datenmengen verarbeitet werden, und dass mit dem jüngsten, aktuell gültigen Datenbestand gearbeitet wird.

Demgegenüber können *OnLine Analytical Processing* (OLAP) -Anwendungen grosse Datenmengen (im Bereich von mehreren Terabytes) verarbeiten und insbesondere auf historische Daten zugreifen, um daraus Rückschlüsse, zum Beispiel auf die betriebswirtschaftliche Entwicklung eines Unternehmens, zu gewinnen. Eine OLAP-Anwendung ermöglicht es dem Entscheidungsträger mit den historischen Daten zu 'spielen'.

Es besteht mittlerweile weitgehender Konsens, dass man OLTP- und OLAP-Anwendungen nicht auf demselben Datenbestand (d.h. auf derselben physischen Datenbank) ausführen sollte. OLTP-Datenbanken sind hinsichtlich logischem und physischem Entwurf auf Änderungstransaktionen mit Zugriff auf sehr begrenzte Datenmengen hin optimiert. Dahingegen sind OLAP-Anfragen von historischen Daten meistens sehr komplex. Würden beide Anwendungen auf dieselbe Datenbank zugreifen, wäre die Leistungsfähigkeit beider Anwendungen deutlich geringer. Deshalb wird heute der Aufbau eines autonomen Data Warehouse propagiert, das für die OLAP-Anwendung alle Daten zur Verfügung stellt.

2.8 OLAP Funktionalität

In diesem Kapitel werden die sechs standard OLAP-Operationen vorgestellt, die dabei helfen, durch die Datenmenge in einem Data Warehouse zu navigieren. Diese Operationen werden an einem symbolischen, dreidimensionalen Datenwürfel graphisch veranschaulicht.

Drill-Down: Das Hierarchieobjekt innerhalb des Aggregationsgraphen wird gewechselt, indem in Richtung des LoG-Knotens gewandert wird. In der Abbildung 2-7 befindet man sich in der Dimension *Gruppe* auf der Ebene der Universitäten. Ein *Drill-Down* würde nun zu den einzelnen Gruppen innerhalb einer Universität (z.B. Fribourg) führen, diese Situation ist in der Abbildung 2-8 dargestellt.

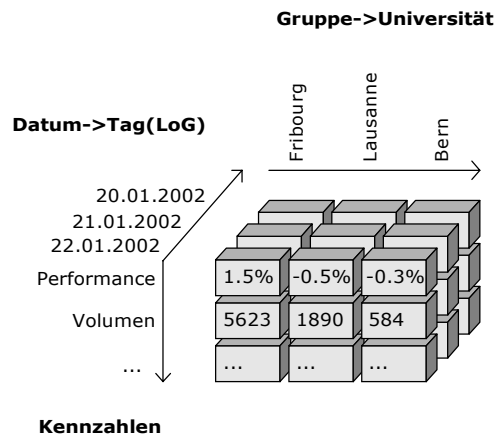


Abb. 2-7, OLAP-Datenwürfel

Roll-Up: Diese Operation wird auch *Drill-Up* genannt und ist die inverse Operation zu *Drill-Down*, man bewegt sich demnach innerhalb des Aggregationsgraphen in die Richtung des All-Knotens. Der Detaillierungsgrad der Daten wird dabei verringert. Der Übergang von Abbildung 2-8 zu Abbildung 2-7 stellt ein *Roll-Up* dar.

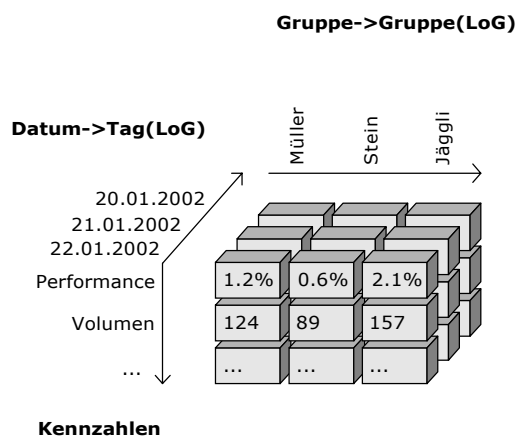


Abb. 2-8, nach Drill-Down

Pivot: Die Perspektive, von der aus die Berichtsdaten betrachtet werden, wird gewechselt. Das entspricht einem Vertauschen der aktuell dargestellten Dimensionen, jedoch ohne die Dimensionalität der ursprünglichen OLAP-Abfrage zu verändern. Das Resultat einer *Pivot-Operation* bezüglich den Dimensionen *Datum* und *Gruppe* ist im Übergang von Abbildung 2-8 zu Abbildung 2-9 dargestellt.

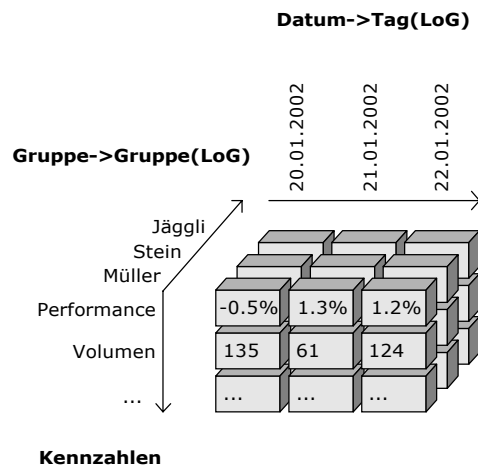


Abb. 2-9, nach Pivot-Operation

Slice: Entspricht dem Abschneiden einer Scheibe des Datenwürfels, zeigt also nur die Daten an, die ein bestimmtes Kriterium erfüllen. Die ursprüngliche OLAP-Abfrage wird um eine Dimension verringert. In Abbildung 2-10 ist das Resultat einer *Slice-Operation* für das Kriterium 'Datum = 22.01.2002' dargestellt.

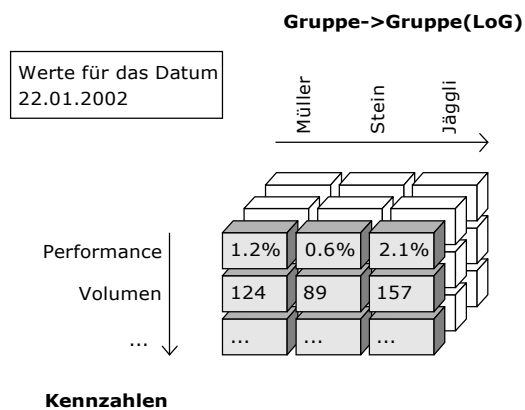


Abb. 2-10, nach Slice-Operation

Dice: Diese Operation setzt sich aus mehreren *Slice-Operationen* zusammen. Das entspricht dem Ausschneiden eines bestimmten Teiles aus dem bestehenden Datenwürfel. Die ursprüngliche OLAP-Anfrage bleibt erhalten, jedoch werden die dargestellten Elemente der Hierarchieobjekte verändert, indem ungewünschte Elemente herausgefiltert werden. Das unerwünschte Element in der Abbildung 2-11 ist die Gruppe *Stein* und wurde deshalb herausgefiltert.

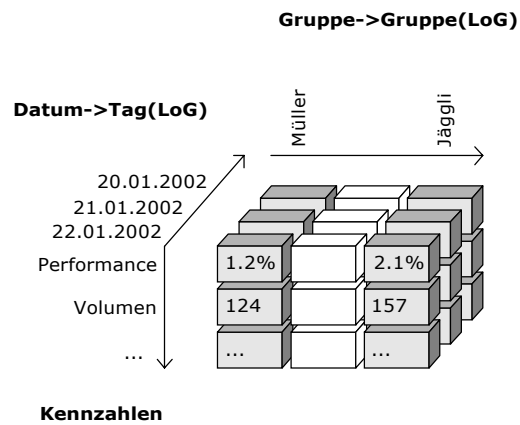


Abb. 2-11, nach Dice-Operation

Drill-Across: Daten aus Faktentabellen, welche über gleiche Auswertedimensionen verfügen, können in einem OLAP-Bericht dargestellt werden. Ein *Drill-Across* ist deshalb erst möglich, wenn mindestens zwei Kennzahlengruppen existieren, die mindestens eine gemeinsame Dimension haben.

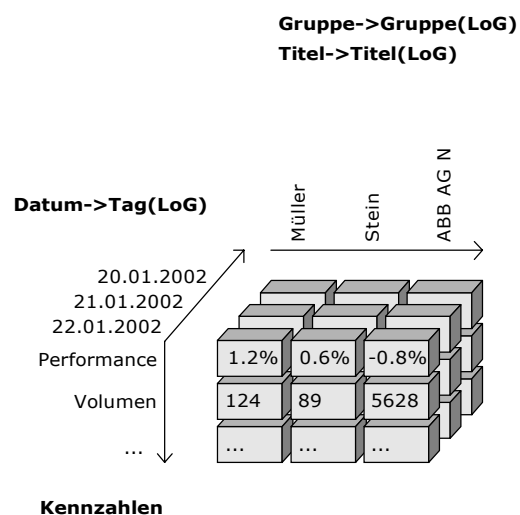


Abb. 2-12, nach Drill-Across

In der Abbildung 2-12 ist neben den beiden Gruppen *Müller* und *Stein* auch noch eine Aktie im Resultat vorhanden. Das ist möglich, weil die Kennzahlengruppen *Portfolios* und *Titel* dieselbe Dimension *Datum* und dieselben Kennzahlen *Performance* und *Volumen* besitzen.

2.9 Das OLAP-Tool MicroStrategy 7

Nachdem das Modell eines Data Warehouse entworfen wurde, die Tabellen implementiert sind und der ETL-Prozess funktioniert, kann auf diesem Fundament eine OLAP-Applikation erstellt werden. Für das Data Warehouse der BSU kommt die Software *MicroStrategy 7* zum Einsatz, welche aus mehreren verschiedenen Komponenten besteht. In den folgenden Unterkapiteln wird eine Auswahl von drei dieser Komponenten vorgestellt.

2.9.1 MicroStrategy Desktop

Um die OLAP-Anwendung zu erstellen, wird mit der Komponente *MicroStrategy Desktop* gearbeitet. Eine *Project Source* (ODBC-Datenquelle) dient zum Zugriff auf die Datenbank, in der sich die Tabellen des Data Warehouse befinden. Basierend auf einer *Project Source* kann man ein neues Projekt anlegen und definieren, welche Tabellen der Datenbank in den *Warehouse Katalog* aufgenommen werden. Dies betrifft immer die Dimensions- und Faktentabellen.

Anschliessend können Fakten, Attribute, Filter, Promts, Metriken, Transformationen und vieles mehr für das Projekt definiert werden. Auf die Bedeutung dieser Begriffe wird im Kapitel *3.3 Implementation mit MicroStrategy 7* näher eingegangen. Diese verschiedenen Objekte können in Berichte eingebunden werden, über welche die Anwender auf die Informationen im Data Warehouse zugreifen und diese analysieren können.

2.9.2 MicroStrategy Intelligence Server

Der *Intelligence Server* ist das Herz der *MicroStrategy 7* Architektur. Er ermöglicht die Interaktion mit dem Data Warehouse und kontrolliert alle anfallenden Anfragen. Der *Intelligence Server* besteht aus einer *Analytical Engine* und einer *SQL Engine*.

Die *Analytical Engine* stellt über 150 vordefinierte statistische, finanztechnische und mathematische Funktionen zur Verfügung. Zudem ist sie erweiterbar, es können benutzerdefinierte Funktionen festgelegt und ins System eingebunden werden.

Die *SQL Engine* dient als Übergangsmechanismus zwischen der Applikation und dem Data Warehouse. User-Anfragen werden übernommen und in SQL übersetzt, damit die Anfrage an die Datenbank weitergeleitet werden kann. Damit ein effizienter Datentransfer zwischen dem Data Warehouse und der Applikation gewährleistet werden kann, generiert die *SQL Engine* nur optimierten SQL Code.

Normalerweise wird der *MicroStrategy Intelligence Server* auf einem separaten Server installiert. Die Anwendung (*MicroStrategy Desktop* oder *Web*) schickt die Anfragen an den *Intelligence Server*, der sie in optimierten SQL Code übersetzt und anschliessend an das Data Warehouse weiterleitet (dessen Tabellen wiederum auf einem separaten Rech-

ner platziert sein sollten). Für Testzwecke oder kleine Datenmengen ist es allerdings auch möglich die Anwendung, den *Intelligence Server* und das Data Warehouse auf ein und demselben Server zu platzieren.

2.9.3 MicroStrategy Web

MicroStrategy Web ist die Komponente, welche den Zugriff auf das Projekt via Internet ermöglicht. Die Konfiguration dieser Komponente erfolgt via Internet, dabei kann definiert werden, welche Projekte aus dem Internet zugänglich sind. Man kann verschiedene Benutzer mit verschiedenen Zugriffsrechten definieren, es ist sogar möglich via Internet neue Berichte zu erstellen.

Im *MicroStrategy Desktop* erstellte Berichte können quasi 1:1 ins Internet übernommen werden. Dank der Komponente *MicroStrategy Web* ist es mit wenig Aufwand möglich, ein Data Warehouse internetfähig zu machen. Damit diese Komponente fehlerfrei funktioniert, muss man unbedingt beachten, welche Treiber- und Browsersoftware von MicroStrategy 7 zertifiziert ist.

2.10 Das Umfeld - die Börsenwelt

Wie bereits erwähnt, enthält das im Rahmen dieser Diplomarbeit realisierte Data Warehouse hauptsächlich Daten über den Verlauf von Börsenkursen, Transaktionen und Portfolios der Teilnehmer. In den folgenden Kapiteln werden die verschiedenen Arten der handelbaren Wertpapiere und verbreitete Kennzahlen für die Analyse von Kursverläufen vorgestellt.

2.10.1 Die verschiedenen Arten von Wertpapieren

Zum jetzigen Zeitpunkt werden im Börsenspiel der BSU vier verschiedene Kategorien von Wertpapieren unterschieden: Aktien, Optionen, Obligationen und Devisen.

SMI-, SPI-, New Market- und Ausland-Aktien: Eine Aktie ist eine Urkunde über den Besitzanteil am Grundkapital einer Aktiengesellschaft. Sie hat einen Nennwert und einen Kurswert. Der Nennwert gibt an mit welchem Anteil der Aktionär am Grundkapital und damit am gesamten Vermögen der Aktiengesellschaft beteiligt ist. Der Kurswert ist die aktuelle Bewertung des Wertpapiers durch die Börse und ist täglichen Schwankungen ausgesetzt. Für die Visualisierung der Aktienkurse sind nur die Kurswerte von Bedeutung.
EUREX Optionen: Die EUREX ist eine vollständig elektronische Börse, an der nur Derivate gehandelt werden. Eine Option berechtigt den Inhaber ein bestimmtes Wertpapier während einer bestimmten Periode oder zum Ende der Laufzeit zu einem vorab bestimmten

Kurs zu erwerben (Call Option) oder zu verkaufen (Put Option). Dieses Recht kann sich der Anleger gegen Zahlung einer Optionsprämie sichern.

Obligationen: Eine Schuldverschreibung mit festgesetzter (fester oder variabler) Verzinsung und fester Laufzeit, meist zwischen 5 und 30 Jahren. Dem Anleger bieten Obligationen die Möglichkeit, Kapital langfristig und risikoarm zu investieren.

Devisen: Da auch die Währungen ständigen Kursschwankungen ausgesetzt sind, kann mit dem Kauf oder Verkauf von ausländischen Zahlungsmitteln gehandelt werden.

2.10.2 Typische Kennzahlen für Wertpapiere

Kennzahlen sind dazu da, den zukünftigen Kursverlauf von Wertpapieren besser einschätzen zu können. Es wird davon ausgegangen, dass genügend Informationen, die für eine Prognose erforderlich sind, in den Kursen und deren bisheriger Entwicklung implizit enthalten sind. Dennoch ist gegenüber allen Indikatoren Vorsicht angebracht, keiner von ihnen ist ein sicheres Signal für die weitere Entwicklung eines Kurses, sondern nur ein Indiz für die bisherige Trendrichtung. In trendlosen Märkten kommen normalerweise sogenannte *Oszillatoren* zum Einsatz, um einen aufkommenden Trend frühzeitig festzustellen. Ist ein Trend vorhanden, kann dieser mit einem sogenannten *Trendfolger* genauer analysiert werden.

High Low

Der *High Low* Indikator ist keine eigentliche Kennzahl, weil für seine Visualisierung keine Berechnungen vorgenommen werden müssen. In einer Graphik werden schlicht und einfach der tägliche Höchst-, Tiefst- und Schlusskurs eines Titels dargestellt. Liegen die Kurven des Höchst- und Tiefstkurses weit auseinander, weist das auf eine hohe Volatilität hin.

Umsatzvolumen

Veränderungen im Umsatzvolumen, also der Menge der An- und Verkäufe einer Aktie, sind ein wichtiges Indiz für die Stärke der Kursbewegung nach oben oder nach unten. Steigende Aktienkurse bei gleichzeitig steigendem Umsatzvolumen gelten als eindeutiges Trendsignal für eine Kursbewegung nach oben - und umgekehrt. Wichtiger als das absolute Umsatzvolumen ist dabei die Stärke der Veränderung im Verhältnis zum Kurstrend [MOEL00].

Gleitender Durchschnitt (Moving Average) - Trendfolger

Der Gleitende Durchschnitt gehört zu den ältesten und mit Abstand am meisten benutzten Indikatoren. Für einen Aktienwert (oder eine Anzahl von Aktienwerten) wird für eine bestimmte Anzahl von Tagen, meistens 38 oder 200, ein Mittelwert aus den jeweiligen

Kursen gebildet. Jeden Tag fällt der älteste Tageskurs weg und der neue Kurs fließt in die Berechnung ein. Der Zweck des gleitenden Durchschnitts ist es, die Extreme in den Kursschwankungen zu glätten, damit der Haupttrend in der Börsenentwicklung schärfer wird. Der 38-Tage-Durchschnitt gibt Aufschluss über den kurzfristigen Trend, der 200-Tage-Durchschnitt über den langfristigen Trend.

Durchstößt der 38-Tage-Durchschnitt den 200-Tage-Durchschnitt nach oben oder nach unten (cross over), ist dies ein Indiz für eine endgültige Trendwende in der bisherigen fallenden oder steigenden Kursentwicklung.

$$\text{Moving Average} = \frac{\sum_{j=1}^n \text{Schlusskurs}_j}{n}$$

Der gleitende Durchschnitt errechnet sich aus der Summe einer Anzahl Kurswerte dividiert durch die Anzahl der Kurswerte [MOEL00].

Relative Stärke nach Levy (RSL) - Oszillator

Levy geht in seiner Betrachtung davon aus, dass sich Kurse, die sich in der Vergangenheit positiv entwickelt haben, auch in der Zukunft positiv entwickeln werden. Die Relative Stärke bezieht sich also nicht auf den Vergleich mit einem anderen Titel oder Index, sondern vielmehr auf den Titel selbst.

$$\text{Relative Stärke} = \frac{\text{Schlusskurs}}{\text{Moving Average}}$$

Ein Wert größer als 1 besagt, dass die aktuelle Performance besser, ein Wert kleiner als 1 besagt, dass die aktuelle Performance schlechter als in der Vergangenheit ist [MARK02].

Volumen Oszillator

Der Volumen Oszillator basiert auf der Berechnung der Differenz zweier Gleitenden Durchschnitts. Durch das Bilden der Differenz schwanken die Werte des Indikators um die Nulllinie.

$$\text{Volumen Oszillator} = \text{Moving Average}_n (\text{Volumen}) - \text{Moving Average}_m (\text{Volumen})$$

Die Variable n wird immer kleiner als die Variable m gewählt, beispielsweise kann man für n den Wert 30, und für m den Wert 50 wählen.

Ein Wert oberhalb der Nulllinie zeigt, dass das kürzere Moving Average über dem längeren liegt und damit das Umsatzvolumen steigend ist. Kommt der Indikator unter die Nulllinie zu liegen, fällt das Umsatzvolumen. Interessant ist zu beobachten, ob ein steigendes Umsatzvolumen auch mit einem steigenden Kurs übereinstimmt. Ist das der Fall, scheint der Trend in Ordnung zu sein. Bilden sich jedoch Divergenzen, ist mit einer Trendwende zu rechnen [MARK02].

Das Momentum - Trendfolger

Mit dem Momentum wird der aktuelle Kurs eines Wertpapiers, im Verhältnis einer beliebigen Anzahl Tage zuvor, bezeichnet. Normalerweise liegt diese Anzahl der Tage zwischen 20 und 30 Tagen.

$$\text{Momentum} = \frac{\text{Schlusskurs}}{\text{Schlusskurs (vor n Tagen)}}$$

Das Momentum verdeutlicht die Stärke der zyklischen Kursschwankungen. Ein Momentum unter 1 bedeutet eine negative Änderung des Aktienkurses, umgekehrt bedeutet ein Momentum über 1 eine positive Kursänderung. Der Trend eines Aktienkurses bleibt erhalten, solange der Kurs und das Momentum die gleiche Richtung aufweisen. Ein Auseinanderlaufen der Richtung von Kurs und Momentum bedeutet eine Trendwende.

Das Momentum ist ein nützlicher Indikator für die frühzeitige Anzeige von Trendwenden in der Kursentwicklung. Umso kleiner die Variable n gewählt wird, desto sensibler wird das Momentum auf Kursveränderungen. Ein steigendes Momentum gilt als Kaufsignal, ein stagnierendes als ein Haltesignal und ein fallendes als ein Verkaufssignal [MOEL00].

Stochastik Indikator - Oszillator

Der Stochastik Indikator zählt zu den gebräuchlichsten Indikatoren und wurde von George C. Lane entwickelt. Diese Kennzahl misst das Verhältnis des aktuellen Kurses zu den Höchst- und Tiefstkursen der letzten Tage. Üblicherweise wird der Stochastik Indikator zweimal berechnet, für die Dauer von 3 beziehungsweise 5 Tage.

$$\text{Stochastik Indikator} = \left(\frac{\text{Schlusskurs} - \text{Tiefstkurs (letzte n Tage)}}{\text{Höchstkurs (letzte n Tage)} - \text{Tiefstkurs (letzte n Tage)}} \right) * 100$$

Der Klammerausdruck (*letzte n Tage*) bezeichnet eine Menge von Tagen, der Ausdruck *Höchst-* bzw. *Tiefstkurs (letzte n Tage)* hingegen bezeichnet den Höchst- bzw. Tiefstkurs dieser Menge. Die Werte des Indikators liegen immer im Bereich von 0 bis 100. Liegen

beide Werte (für $n=3$ und $n=5$) unter 30 bedeutet das, dass sich der Markt in einem überverkauften Stadium befindet, ein Hinweis also, dieses Wertpapier zu kaufen. Falls beide Werte über 70 liegen, sollte man verkaufen [MARK02].

Money Flow Index - Trendfolger

Diese Kennzahl bezieht neben den eigentlichen Kursen auch noch die Umsätze in die Berechnung ein. Er stellt eine Art Umsatz Momentum dar. Für die Berechnung benötigt man zuerst den durchschnittlichen Wert des Titels pro Tag.

$$\text{Durchschnittlicher Kurs} = \frac{\text{Höchstkurs} + \text{Tiefstkurs} + \text{Schlusskurs}}{3}$$

Als nächstes berechnet man den Money Flow:

$$\text{Money Flow} = \text{Durchschnittlicher Kurs} * \text{Volumen}$$

Man unterscheidet zwei verschiedene Money Flows, den Positiv Money Flow (PMF) und den Negativ Money Flow (NMF). Ist der durchschnittliche Preis im Vergleich zu gestern gestiegen, zählt man den Money Flow als PMF, ist er gesunken zählt man den Money Flow als NMF. Pro Tag gibt es also entweder einen PMF oder einen NMF aber nie beides. Der nächste Schritt in der Berechnung besteht darin, die Money Ratio zu berechnen. Normalerweise zieht man die PMF's und NMF's der letzten 20 Tage mit in die Berechnung ein.

$$\text{Money Ratio} = \frac{\sum_{j=1}^n \text{Positive Money Flow}_j}{\sum_{j=1}^n \text{Negative Money Flow}_j}$$

Aus der Money Ratio lässt sich schliesslich der gewünschte Money Flow Index berechnen:

$$\text{Money Flow Index} = 100 - \frac{100}{1 + \text{Money Ratio}}$$

Dank der Normalisierung wird sich der Money Flow Index immer zwischen 0 und 100 bewegen. Als Standardsignal wird das Kreuzen mit der 50er Linie gedeutet. Ein Durchbruch von unten nach oben liefert ein Kaufsignal, ein Durchbruch von oben nach unten ein Verkaufsignal [MARK02].

Bollinger Bands - Trendfolger

Dieser Indikator besteht aus drei Bändern, die versuchen den Kursverlauf mit Hilfe einer oberen und unteren Begrenzung einzufangen. Das mittlere Band ist ein normales Moving Average, während mit Hilfe der Standardabweichung das obere und untere Band definiert werden. Die Formel für das Moving Average (Mittleres Band) ist bereits bekannt, deshalb wird an dieser Stelle nur noch auf die Berechnung der beiden Bänder eingegangen.

$$\text{Oberes Band} = \text{Mittleres Band} + D * \sqrt{\frac{\sum_{j=1}^n (\text{Schlusskurs}_j - \text{Mittleres Band})^2}{n}}$$

Die Berechnung des unteren Bandes ist analog, die Standardabweichung wird einfach vom mittleren Band subtrahiert anstatt addiert.

D ist eine Variable mit der man den Einfluss der Standardabweichung bestimmen kann. Wählt man $D=1$ so bewegen sich rund 70% der Kurse innerhalb der Begrenzungen, für $D=2$ sind es rund 95%. John Bollinger, der Erfinder dieses Indikators, empfiehlt den Wert D auf zwei festzulegen. Die Anzahl Tage n, welche in die Berechnung einfließen, können grundsätzlich frei gewählt werden, Bollinger empfiehlt mit 20 Tagen zu rechnen. Berührt oder durchstösst ein Kurs die Bollinger Bänder, kann man davon ausgehen, dass er sich in die entgegengesetzte Richtung weiterbewegen wird [TECH01].

2.10.3 Normalisierung der Indikatoren

Es sollte möglich sein, den Kurs eines Wertpapiers zusammen mit Indikatoren darzustellen. Bewegt sich ein Kurs auf einer Höhe von ungefähr 200 Punkten, hat eine Kennzahl, die um den Wert 1 schwankt, keine Aussagekraft, weil ihr Verlauf in der Graphik nicht mehr zu erkennen ist. Dieses Problem kann gelöst werden, indem man alle Indikatoren normalisiert, damit ihr Ergebnisbereich immer in der Nähe des Kursverlaufs liegt.

Am Beispiel des Stochastik Indikator (Wertebereich zwischen 0 und 100) wird gezeigt, wie man den Ergebnisbereich dieser Kennzahl auf jeden beliebigen Kursverlauf normalisieren kann. In diesem Fall symbolisiert der Kursverlauf die 50er Linie, die 0er und 100er Linie befinden sich 50% unter beziehungsweise über dem eigentlichen Kursverlauf. Analoges gilt für die beiden ausschlaggebenden Linien die neu bei 20% unter beziehungsweise 20% über dem Kursverlauf zu liegen kommen.

$$\text{Zwischenresultat} = \left(\frac{\text{Schlusskurs} - \text{Tiefstkurs (letzte n Tage)}}{\text{Höchstkurs (letzte n Tage)} - \text{Tiefstkurs (letzte n Tage)}} \right)$$

Obenstehende Formel liefert Werte zwischen 0 und 1. Eine Multiplikation mit dem Schlusskurs ergibt einen Wertebereich zwischen 0 und dem Schlusskurs. Addiert man dazu noch 50% des Schlusskurses, so liegen die Ergebnisse des Stochastik Indikators genau im oben definierten Wertebereich. Die Formel für diese Berechnung lautet dann folgendermassen:

$$\text{Stochastik Indikator} = (\text{Zwischenresultat} * \text{Schlusskurs}) + \frac{\text{Schlusskurs}}{2}$$

In derselben Graphik müssen auch die beiden ausschlaggebenden Linien dargestellt werden, diese liegen wie bereits erwähnt 20% über beziehungsweise unter dem Kursverlauf. Diese Methode kann in derselben Art und Weise auf die anderen Indikatoren angewandt werden.

2.11 BSU - Börsenspiel der Schweizer Universitäten

Der NPO³-Verein BSU wurde am 20. Oktober 1991 von den Wirtschaftsstudenten Andreas Hüchting und Harald K. Berg der Universität Freiburg (Schweiz) unter dem Namen *SEFU* (Stock Exchange of Fribourg University) gegründet. Ziel des Vereins ist es: "Einen Einblick in die Wertpapiermärkte und die marktbeeinflussenden Kräfte zu ermöglichen; gesamtwirtschaftliches Denken zu fördern; eine Brücke zwischen Theorie und Praxis zu schlagen; gruppenspezifische Vorgänge in Gang zu setzen und dadurch die Teamarbeit der Teilnehmer zu fördern." (Art. 5 der Vereinsstatuten).

Der Verein setzt sich aus einer Administration (Wirtschaftsstudenten der Universität Freiburg) und einem Vorstand (Gründungsmitglieder und andere ehemalige Administrationsmitglieder) und Vereinsmitgliedern (Teilnehmer an den Aktivitäten) zusammen. Zusätzlich können die Organisatoren auf die beratende Unterstützung von Professoren der Universität Freiburg zählen.

1992 wurde an der Universität Freiburg das erste jährlich stattfindende Börsenspiel durchgeführt. Im Verlaufe der Zeit wurde auch den Studenten anderer Schweizer Universitäten und Hochschulen die Möglichkeit gegeben, am Börsenspiel teilzunehmen, womit sich der Verein vor allem unter den Wirtschaftsstudenten der Deutschschweiz einen Namen machen konnte. Seit 1997 wurde das *Börsenspiel* durch die *Portfolio Management Simulation* ersetzt, welche alle Studenten der Schweiz (somit auch der Westschweiz und des Tessins) anspricht.

BSU-Frifin organisiert ebenfalls Börsensimulationen im Auftrag Dritter. So wurde für den ehemaligen Schweizer Bankverein der *Magic Fonds Cup* aus der Taufe gehoben. Seit

³ NPO steht für "Non Profit Organisation"

1998 wird zusätzlich für die Lehrlinge der Regionalbanken (RBA) ein Börsenspiel zu Ausbildungszwecken organisiert.

Weitere Informationen sind unter der Adresse <http://www.unifr.ch/bsu> zu finden.

CONFIDENTIAL

MicroStrategy Web verliert nach einem Neustart des iiufpc06 die Verbindung zum Data Warehouse BSU

Laut Konfiguration der *MicroStrategy Web* Komponente müssten die Verbindungen nach einem Neustart des Servers automatisch wiederhergestellt werden. Jedoch ist das nicht der Fall, weshalb die Verbindungen manuell aktiviert werden müssen. Hierzu muss man sich unter der Adresse <http://iiufpc06.unifr.ch/microstrategy7/admin/admin.asp> einloggen.

Bleibt die Verbindung inaktiv, können die Mitspieler nicht mehr auf das Data Warehouse zugreifen.

Die neusten Daten wurden nicht in das Data Warehouse importiert

Folgende beiden Ursachen kommen in Frage:

- Der *SQL Server-Agent* ist inaktiv

In diesem Fall muss er manuell gestartet werden. Wie man fehlende Informationen in das Data Warehouse importiert, ist unter Punkt 2 im Abschnitt *SQL Server 2000* des Kapitels *4.1.2 Aktivierung des DWH während einem Spiel* erklärt.

- Wurden die neusten Daten nicht importiert obwohl der *SQL Server-Agent* aktiv ist, fehlen die Daten in der BSU-Datenbank oder der ETL-Prozess wurde falsch angepasst.

4.2 Bedienungsanleitung für die Anwender des DWH

Alle Daten, die während einem Börsenspiel anfallen, werden in einem Data Warehouse gespeichert. Sie haben jetzt die Möglichkeit diese relativ grosse Datenmenge zu analysieren. Dazu stehen Ihnen verschiedene Berichte zur Verfügung, welche Ihnen an dieser Stelle kurz vorgestellt werden. Den Link und die Zugangsdaten zu dieser OLAP-Anwendung finden Sie auf der Homepage der BSU.

Alle Analysen sind ohne Gewähr, nur die Rangliste auf der Homepage der BSU ist spielentscheidend.

4.2.1 Beschreibung der verschiedenen Berichte

Einige Berichte stellen Ihnen eine Auswahl zur Verfügung. Beachten Sie, dass Einträge innerhalb einer Auswahl gesucht werden können, wobei es genügt, wenn Sie nur einen Teil des gesuchten Ausdrucks angeben. Benutzen Sie die kleinen Pfeile unterhalb der Auswahl um durch die Einträge zu navigieren.

Die Berichte befinden sich alle im Ordner *Shared Reports* und sind in vier Kategorien unterteilt.

Kurs-Analysen

- *Aktien-Analyse, Devisen-Analyse, Obligationen-Analyse und Optionen-Analyse*, diese Berichte liefern die Kurswerte der im Spiel handelbaren Titel. Wählen Sie hierzu einfach die gewünschten Titel aus und folgen Sie den Anweisungen. Die Aktien-Analyse greift auch auf historische Kurse zurück, wobei sich die Analyse der anderen Titel auf die Dauer des Börsenspiels beschränkt. Die historischen Kurse der Aktien können zusätzlich mit einigen gängigen Indikatoren analysiert werden.
- *Overall Performance*, diese Analyse ist erst nach dem Spielende aussagekräftig. Alle im Spiel handelbaren Titel werden aufgelistet und nach ihrer Performance bewertet.
- *Rank Kategorien Performance (pro Tag)*, wählen Sie einen Tag (gestern oder früher) und der Bericht liefert Ihnen die durchschnittliche Performance dieser Kategorie. Klicken Sie anschliessend auf die Kategorie um die Performance der einzelnen Titel zu sehen.
- *Rank Titel Performance (pro Tag)*, wählen Sie einen Tag (gestern oder früher) und der Bericht liefert Ihnen die Rangliste bezüglich der Performance aller im Spiel handelbarer Titel.

Portfolio-Analysen

- *Portfolio-Zusammensetzung*, wählen Sie eine Gruppe aus und der Bericht liefert Ihnen die Zusammensetzung des Portfolios der betreffenden Gruppe, das heisst Sie sehen wie das investierte Geld angelegt ist.

Transaktions-Analysen

- *Alle Transaktionen einer Gruppe*, wählen Sie eine Gruppe aus und der Bericht listet das bis jetzt angefallene Kaufs- und Verkaufsvolumen dieser Gruppe auf.
- *Gesamtes Spielvolumen*, dieser Bericht ist erst nach dem Spielende aussagekräftig und bietet eine Übersicht aller Transaktionen.
- *Rangliste Courtage*, dieser Bericht zeigt wer bis zum aktuellen Zeitpunkt wieviel Courtage gezahlt hat.
- *Volumen einer Aktie, Volumen einer Devise, Volumen einer Obligation und Volumen einer Option*, wählen Sie einen Titel und diese Berichte stellen dessen Kaufs- und Verkaufsvolumen dar.
- *Tägliches Gesamtvolumen, Tägliches Kaufsvolumen und Tägliches Verkaufsvolumen*, wählen Sie einen Tag (gestern oder früher) und der Bericht rangiert die verschiedenen Titel nach dem jeweiligen Kriterium.

Vermögens-Analysen

- *Overall Performance*, dieser Bericht ist erst nach dem Spielende aussagekräftig und rangiert die Performance aller Gruppen.

- *Performance pro Tag*, wählen Sie ein Datum aus (gestern oder früher) und der Bericht zeigt Ihnen die Performance aller Gruppen.
- *Rangliste -> Gruppe, Rangliste -> Sprache (Average) und Rangliste -> Uni (Average)*, wählen Sie eine Tag aus (gestern oder früher) und der Bericht rangiert die Gruppen nach ihrem Vermögen.
- *Vermögen pro Gruppe*, wählen Sie eine Gruppe aus und der Bericht liefert Ihnen die Entwicklung des Vermögens der betreffenden Gruppe.

4.2.2 Zusätzliche Funktionen innerhalb von Berichten

Tabellarische Berichte bieten eine Ausgangslage für weitere Analysen. So können die Ergebnisse beispielsweise nach jeder Spalte auf- und absteigend sortiert werden. Klicken Sie hierzu auf das kleine Dreieck links vom Spaltennamen.

Es stehen auch sogenannte *Drill-Funktionen* zur Verfügung, um die Ergebnisse unter einem anderen Blickwinkel darzustellen. Klicken Sie beispielsweise in der Kurs-Analyse *Rank Kategorien Performance (pro Tag)* auf eine Kategorie, um die dieser Kategorie angehörigen Titel und ihre Performance zu sehen.

Zum Schluss noch ein Beispiel einer etwas anspruchsvolleren *Drill-Funktion* innerhalb der Transaktions-Analyse *Alle Transaktionen (einer Gruppe)*. Dieser Bericht listet Ihnen für jeden Tag das Volumen aller Transaktionen einer Gruppe auf. Möchten Sie wissen welche einzelnen Transaktionen hinter diesen Volumina stecken, gehen Sie wie folgt vor:

Klicken Sie oben auf 'Drill-Funktion' und wählen Sie anschliessend 'weitere Optionen...'. Scrollen Sie jetzt in der Systemhierarchie bis Sie den Eintrag 'Titel' gefunden haben und wählen Sie ihn aus. Aktivieren Sie rechts davon die Checkbox 'Bei Drill-Funktion das Elternobjekt beibehalten'. Markieren Sie jetzt unten in der Tabelle die Tage, über deren Transaktionen Sie mehr erfahren möchten. Klicken Sie auf 'GO' um die Analyse zu starten.

5 Zusammenfassung und Schlussfolgerungen

Die in Kapitel 1.2 *Zielsetzung* definierten Ziele wurden erfreulicherweise erreicht. Ein passendes Modell für das Data Warehouse konnte gefunden und implementiert werden. Der ETL-Prozess läuft während der ganzen Dauer eines Börsenspiels automatisch ab, nur bei einem Spielwechsel fällt ein geringer Administrationsaufwand an. Die Analysen wurden in die vier Bereiche Kurs-, Portfolio-, Transaktions- und Vermögensanalysen aufgeteilt und sind via Internet ausführbar. Für die Analyse der historischen Aktienkurse konnten zusätzlich diverse Indikatoren zur Verfügung gestellt werden.

Die Arbeit mit dem *SQL Server 2000* wurde als sehr angenehm empfunden. Mit gespeicherten Prozeduren, DTS-Paketen und dazugehörigen Aufträgen konnte der ETL-Prozess flexibel definiert werden. Die graphische Oberfläche des DTS-Editors ermöglichte es, auch bei komplexen Imports die Übersicht zu bewahren. Einziger negativer Kritikpunkt beim *SQL Server 2000* betrifft den *SQL Server-Agent*, welcher nach einem Neustart des Servers nicht automatisch startet, obwohl er so konfiguriert wurde.

Auch *MicroStrategy 7* konnte insgesamt überzeugen. Die vorgesehenen Analysen konnten alle bis auf einen Indikator implementiert werden und bringen Ergebnisse zum Vorschein, die zuvor in der Informationsflut untergegangen sind. Dank der Komponente *MicroStrategy Web* konnte das Data Warehouse fast ohne zusätzlichen Aufwand internetfähig gemacht werden, was besonders positiv aufgefallen ist. Negativ aufgefallen ist, dass die Web-Komponente bei einem Neustart des Servers (auf dem sie installiert ist) die Verbindungen zu den Data Warehouses verliert und nicht automatisch wiederherstellt, obwohl sie so konfiguriert wurde.

Die BSU-Datenbank, von welcher die Daten für das Data Warehouse stammen, ist nicht ganz über alle Zweifel erhaben. Im varchar-Format gespeicherte Datumswerte und doppelte Einträge in der Kurstabelle *bsu.Kurs_Detail* sorgten für zusätzlichen Aufwand während dem ETL-Prozess. Auch existiert zum jetzigen Zeitpunkt keine Tabelle mit zusammengesetztem Primärschlüssel, was erstaunlich ist. Auf der Ebene des Data Warehouse sorgen zusammengesetzte Primärschlüssel dafür, dass keine Redundanz in den Faktentabellen entstehen kann.

Die Praxis hat bestätigt, was bereits in Kapitel 2 *Theorie* vorausgesagt wurde. Die Modellierung ist eine verhältnismässig kurze, aber sehr entscheidende Etappe im Aufbau eines Data Warehouse. Spätere Änderungen am Datenmodell bringen viel zusätzlichen Aufwand mit sich, weshalb es sich lohnt, das Modell bereits zu Beginn gut zu überdenken. Im Gegensatz zur Modellierung dauerte die Implementation des ETL-Prozesses um ein Vielfaches länger. Die Struktur der Quelldaten und die notwendigen Transformationen vor dem Import beeinflussen den Zeitaufwand nachhaltig.

Sofern man sich gut mit der betreffenden Software auskennt, ist der Aufbau der OLAP-Anwendung relativ schnell beendet. Muss man die Software jedoch von Grund auf im Selbststudium erlernen, nimmt auch diese letzte Etappe einige Zeit in Anspruch.

Abschliessend kann festgehalten werden, dass die beiden Produkte *SQL Server 2000* und *MicroStrategy 7* sich gut ergänzen, und sich zumindest für den Aufbau eines kleinen bis mittleren Data Warehouse gut eignen.

6 Begriffserklärungen

BSU	Börsenspiel Schweizer Universitäten, www-bsu.unifr.ch
DBMS	Datenbank Management System
ERM	Enterprice Resource Management
ETL-Prozess	<i>Extraction Transforming Loading</i> -Prozess, zuständig für den Import der Daten in das Data Warehouse, siehe Kapitel 3.2.2 <i>ETL-Prozess</i> .
EUREX	Eurex wurde im Dezember 1996 gemeinsam von der Deutsche Börse AG und der Schweizer Börse ins Leben gerufen und durch die Fusion der DTB Deutsche Terminbörse und SOFFEX (Swiss Options and Financial Futures Exchange) im Jahr 1998 gegründet. Die Zielsetzung dieser Fusion war die Entwicklung und Umsetzung einer gemeinsamen Plattform der beiden Terminmärkte sowie die Harmonisierung der Produktpalette.
OLAP	OnLine Analytical Processing
OLTP	OnLine Transactional Processing
SMI	Swiss Market Index, der SMI ist ein kapitalgewichteter, nicht dividendenkorrigierter Index. Er umfasst maximal die 30 liquidesten Titel hochkapitalisierter Unternehmungen mit Sitz in der Schweiz oder Liechtenstein (sog. Blue Chips).
SPI	Swiss Performance Index, der SPI ist ein dividendenkorrigierter Index, der alle an der SWX Swiss Exchange kotierten Aktien inländischer Gesellschaften (inkl. Fürstentum Liechtenstein) enthält. Nicht berücksichtigt werden die an der SWX kotierten Investmentgesellschaften (Vermeidung von Doppelzählungen).
SWX	Swiss Exchange, die SWX Swiss Exchange gehört zu den technologisch führenden Börsen der Welt. Auf der Basis des starken Finanzplatzes Schweiz verfolgt sie eine konsequent auf Internationalität ausgerichtete Strategie.
XML	eXtensible Markup Language

7 Literaturverzeichnis

- [DEWS01] Robin Dewson. **Beginning SQL Server 2000 Programming**. Wrox Press, Birmingham, 2001
- [INMO92] W.H. Inmon. **Building the Data Warehouse**. Publisher John Wiley & Sons, Inc., 1992.
- [KURZ98] A. Kurz. **Data Warehousing im Internet/Intranet: Prototypische Implementierung eines Web- basierten Executive Information System für Entscheidungsträger**. Ph.D. Thesis, Institute of Software Technology, Technical University of Vienna, October 1998, Vienna, Austria, pp. 313.
- [KURZ99] A. Kurz. **Data Warehousing Enabling Technology**. MITP-Verlag, 1999.
- [MARK02] **MARKET MAKER Software AG**, Börsensoftware für professionelle Anwender, 2002.
<http://www.market-maker.de/08.Boerse/Lexikon/>
- [MICR00-1] Microstrategy: **Certified Decision Support Engineer Programme**. Kursunterlagen, Microstrategy Inc., Köln, 2000.
- [MICR00-2] Microstrategy: **Project Designer Guide**. Microstrategy Inc., Vienna (VA), USA, 2000.
- [MOEL00] H-W. Moeller. **Das Börsenseminar, Ein Wegweiser durch den Anlage-Dschungel**. Heyne Verlag, München, 2000.
- [TECH01] Steven B. Achelis. **Technical Analysis from A to Z**, 2001.
<http://www.equis.com/free/taaz/>

CONFIDENTIAL