



universität  
wien

## Business Intelligence Seminar, WS 2007/08

Prof. Dr. Knut Hinkelmann

Fachhochschule Nordwestschweiz  
knut.hinkelmann@fhnw.ch



## Business Intelligence

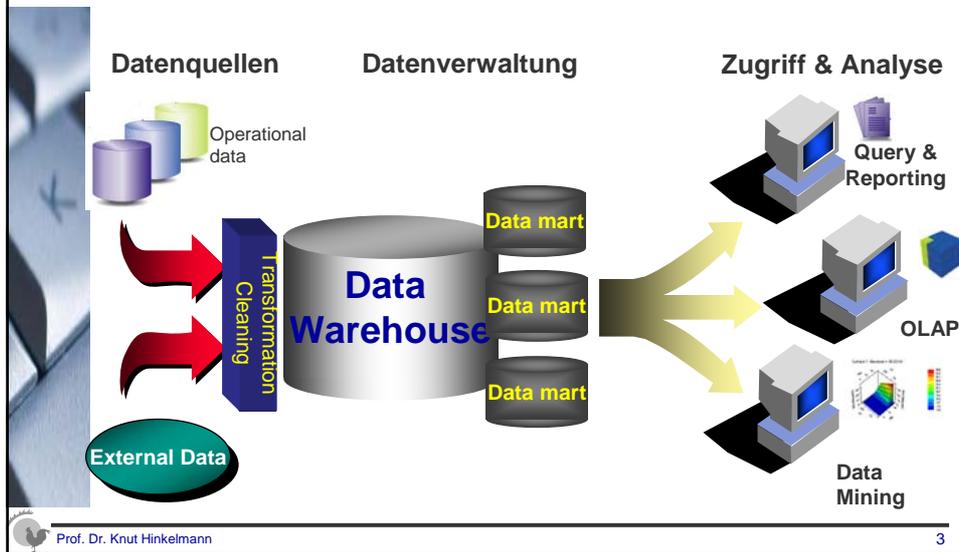
Entscheidungsorientierte Sammlung,  
Aufbereitung und Darstellung  
geschäftswichtiger Informationen.



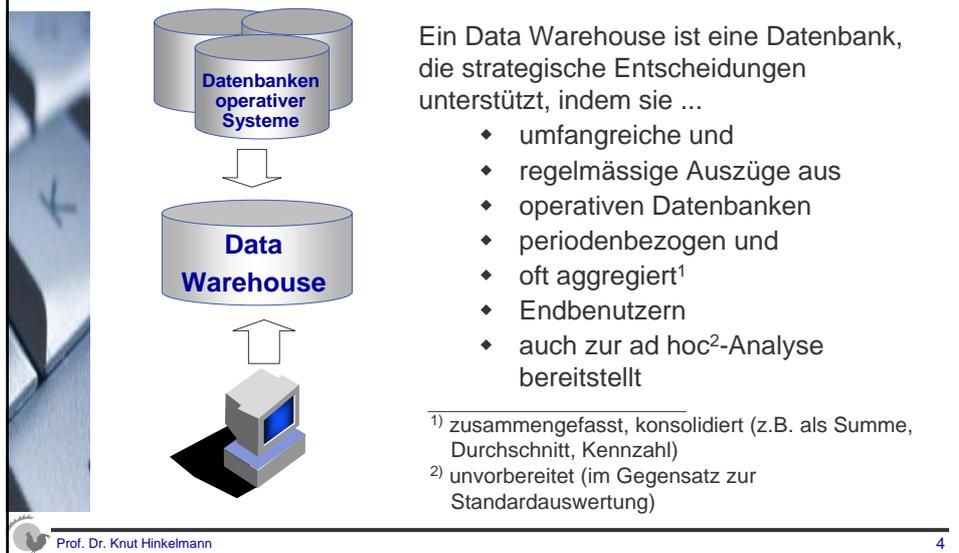
Prof. Dr. Knut Hinkelmann

2

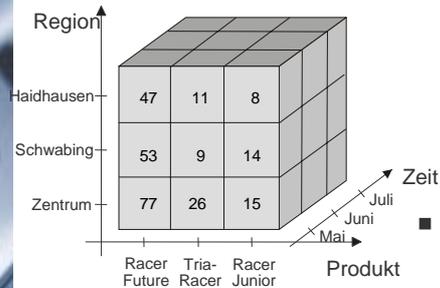
## Business Intelligence



## Data Warehouse - ein mögliche Definition



## Inhalt eines Data Warehouse

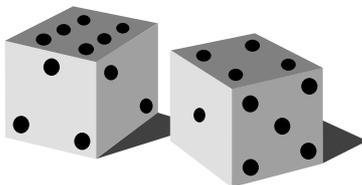


- Typisch für ein Data Warehouse ist die Unterscheidung von
  - ◆ Fakten (Indikatoren)
    - Betriebswirtschaftlicher Erfolgsfaktor, z.B. Umsatz
  - ◆ Dimensionen
    - Kriterien der Beurteilung, z.B. Region, Produkt, Zeit
- Der **Würfel**<sup>1</sup> veranschaulicht die Dimensionen eines Indikators:
  - ◆ an den Schnittpunkten der drei Dimensionen steht der Indikator
- Aufgaben: Einen Indikator nach seinen Dimensionen messen, z.B. "Wie hoch ist der Umsatz nach Periode, Produkt und Region?"

<sup>1</sup>) Mehrdimensionale Würfel heissen "Hypercubes"



## OLAP (OnLine Analytical Processing)



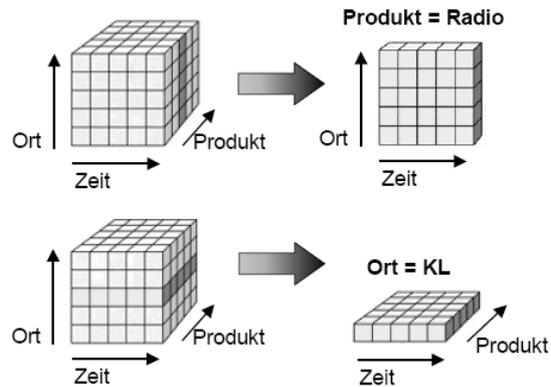
- OLAP<sup>1</sup>) ist eine Abfragemethode, die Endbenutzern einen
  - ◆ **mehrdimensionalen**, schnellen Zugriff und eine
  - ◆ benutzerfreundliche **interaktive Analyse** grosser Datenmengen ermöglicht
- Funktionen
  - ◆ Roll-Up
  - ◆ Drill-Down
  - ◆ Dicing (würfeln)
  - ◆ Slicing (schneiden)

<sup>1</sup>) Eingeführt von E.F.Codd, 1993



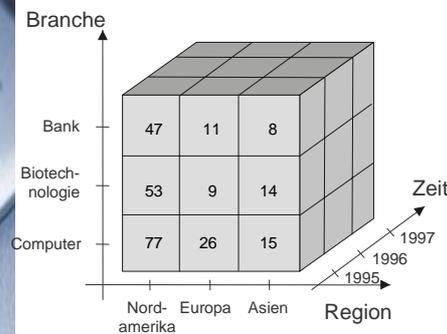
## Slicing (in Scheiben schneiden)

Auswahl eines kleineren Würfels oder Reduktion der Dimensionen in einem mehrdimensionalen Würfel



## Slicing

Auswahl eines kleineren Würfels oder Reduktion der Dimensionen in einem mehrdimensionalen Würfel



Zweidimensionale Darstellung als Tabelle:

1995

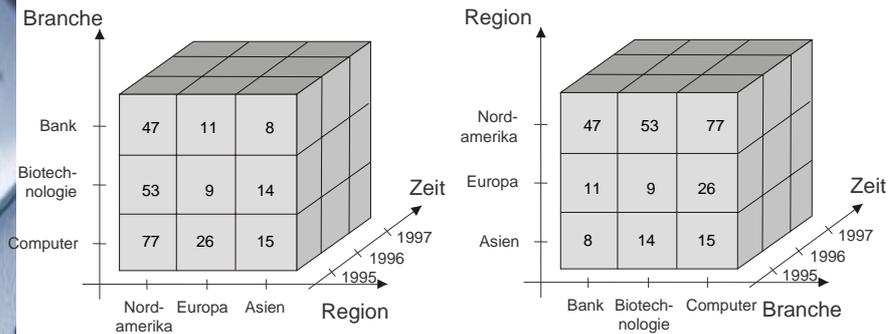
	Nord-amerika	Europa	Asien
Bank	47	11	8
Biotechn.	53	9	14
Computer	77	26	15

1996

	Nord-amerika	Europa	Asien
Bank	43	15	8
Biotechn.	54	13	17
Computer	71	45	28

## Dicing (würfeln)

- Rotation der Resultate um andere Sichten zu zeigen, z.B.
  - ◆ Anzeigen anderer Dimensionen
  - ◆ Austausch von Zeilen und Spalten (auch Pivoting)



## Rolling Up and Drilling Down

**Rolling Up** (zusammenfassend)

**Drilling Down** (detaillierend)

Beispiel: Umsatzzahlen für ...

ganz Europa

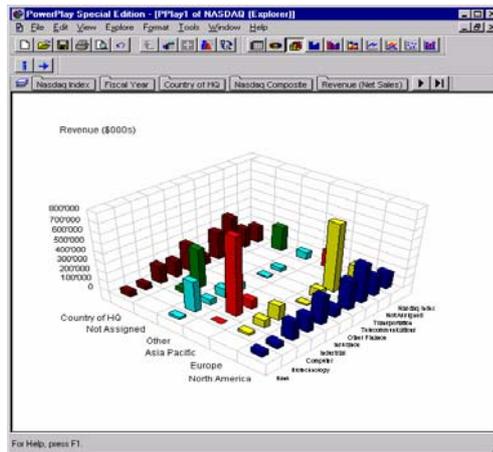


Mitteleuropa

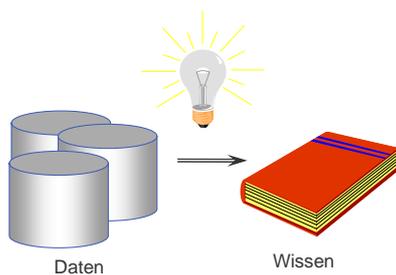


Deutschland

## Visualisierung



## Data Mining: Entdecken von Wissen in Datenbanken



- Data Mining: Identifikation von wettbewerbsrelevantem Wissen aus grossen Datenbanken
- Automatische Extraktion von Informationen zur **Vorhersage** von Trends und Verhalten
- Automatische Erkennung vorher unbekannter **Muster**

## Beispiel aus der Automobilindustrie

- Daten über 7 - 10 Jahre Historie für ca. 7.000.000 Fahrzeuge
  - ◆ Fahrzeugdaten (Produktionsdaten; Daten über Motor, Getriebe, ...)
  - ◆ Beanstandungen (Schadensteil, Schadensart, ...)
  - ◆ Werkstattaufenthalte
- Frage: Wie kann man das Auto zuverlässiger machen?
- Mustererkennung: Suche in Datenbank nach möglichen Gründen für Ausfälle
- mögliche Umsetzung des Wissens:
  - ◆ Änderung in Konstruktion
  - ◆ Wechsel des Zulieferers
  - ◆ Kundendienst: vorbeugende Wartung
  - ◆ usw.

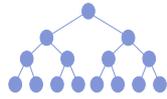


## Arten von Data Mining Problemen

- Klassifikation
  - ◆ Lernen einer Menge von Regeln, die Objekte aufgrund ihrer Attribute vorgegebenen Klassen zuordnen
  - ◆ Beispiel: Klassifikation von Bankkunden ob Kreditwürdig oder nicht
- Clustering
  - ◆ Aufteilung einer Datenmenge in disjunkte Gruppen ähnlicher Objekte
  - ◆ Beispiel: Kundenprofile
- Regression (Vorhersage)
  - ◆ Lernen der Beziehung einer Zielvariable von Ausgangsvariablen
  - ◆ Beispiel: Potentielle Käufergruppe für ein Produkt (in Abhängigkeit z.B. von Alter, Bildung, Einkommen)
- Assoziationen (Warenkorbanalyse)
  - ◆ Lernen von Abhängigkeiten zwischen Daten,
  - ◆ Beispiel: Freitags kaufen Kunden, die Windeln kaufen, oft auch Bier.
- Abweichungs- und Änderungsentdeckung
  - ◆ Entdecken von Unterschieden zu früher aufgezeichneten Daten oder vorgegebenen normativen Werten
  - ◆ Beispiel: Erkennung von Kreditkartenbetrug, Identifikation besonderer Kundengruppen.
- Zeitreihenanalyse
  - ◆ Lernen von Regeln für typische Abläufe zur Prognose zeitlicher Veränderungen
  - ◆ Beispiele: Kontenentwicklungen, Aktienkurse, Vorhersage von Systemausfällen auf Basis von Messdaten



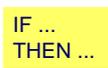
## Wichtige Methoden des Data Mining



Entscheidungsbäume



Neuronale Netze



Regelinduktion



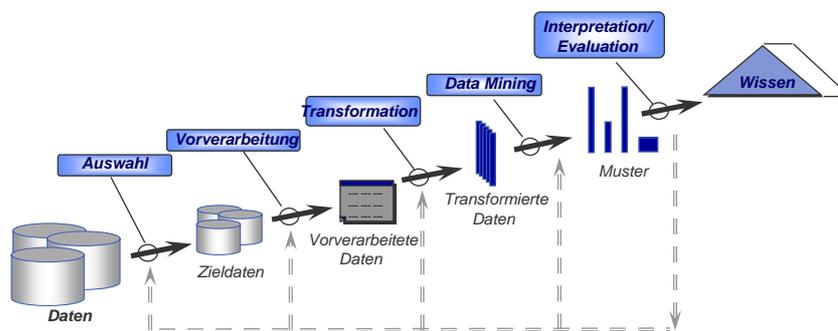
Nearest Neighbor

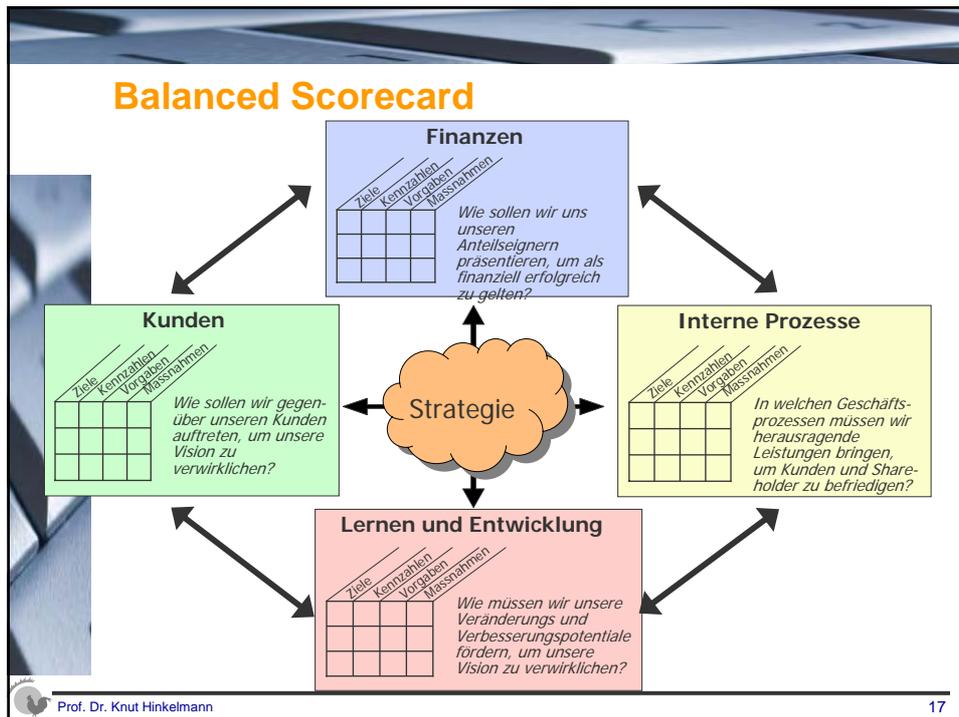


Genetische Algorithmen



## Data Mining ist eine Phase im Prozess der Wissensentdeckung aus Datenbanken





## Business Intelligence: Zusammenfassung

**Ziel:** Aus Unternehmensdaten entscheidungsrelevantes Wissen herleiten

**Data Warehousing/OLAP**

- ◆ Integration von Daten (mehrerer) operativer Systeme
- ◆ Benutzerfreundliche Analyse historischer Daten
- ◆ mehrdimensionale Auswertungen mit OLAP-Werkzeugen
- ◆ Reporting/OLAP sagt, **was** passiert ist

**Data Mining**

- ◆ Ableitung von Modellen für Vorhersage künftiger Ereignisse
- ◆ Data Mining sagt, **warum** Dinge passiert sind
- ◆ Liefert handlungsrelevante Informationen

**Balanced Scorecard**

- ◆ Strategisches Management
- ◆ mehrdimensionales Kennzahlensystem

Prof. Dr. Knut Hinkelmann

18

## Zielsetzung

- Aufgabenstellung
  - ◆ mündliche Präsentation: ca. 35 Minuten
  - ◆ schriftliche Seminararbeit: ca. 20 Seiten
- Hinweise:
  - ◆ Die Seminararbeit ist eine Einzelarbeit.
  - ◆ Die Sprache für die Seminararbeit und die Präsentation ist englisch
- Wissenschaftliche Ausarbeitung:
  - ◆ Das Seminar zielt auf das eigenständige und wissenschaftliche Arbeiten zu einer bestimmten Fragestellung aus dem Bereich Business Intelligence.
  - ◆ Die Arbeiten sollen sowohl wissenschaftliche Erkenntnisse als auch praktische Implikationen an Fallbeispielen oder Anwendungsszenarien berücksichtigen.
  - ◆ Die Arbeit sollte in den grösseren Kontext des Business Intelligence eingeordnet werden
  - ◆ Einstiegsliteratur zu einzelnen Themen wird im Laufe der ersten Woche elektronisch zur Verfügung gestellt. Diese ist lediglich als Einstieg zu betrachten. Es wird jedoch eine darüber hinaus gehende Beschäftigung mit dem Thema erwartet.



## Termine und Kommunikation

- Termine:

Themenvergabe:	2.10.07
Exposee per Email:	15.11.07
Zwischenbesprechung:	20.11.07
Seminar:	24./25.1.08
	Präsentationen
	Abgabe Seminararbeiten
- Kommunikation
  - ◆ Kommunikation im Wesentlichen via Email (bitte Email-Adressen angeben)
  - ◆ Die Arbeiten sind sowohl in Papierversion als auch elektronisch abzugeben.

