

Klausur BI/KMS (Teil Becker)

Fach:	Business Intelligence/KMS
Dozent:	Prof. Dr. Thomas Becker,
Datum:	15.01.2015, 15.15 – 16.45 Uhr
Anzahl der Aufgaben:	4
Hilfsmittel:	keine
Klausurpunkte:	

Hinweise zur Bearbeitung der Klausur:

Bitte prüfen Sie die Klausur auf ihre Vollständigkeit hin! Die Klausur besteht aus 4 Fragen.

Sofern Sie die **Vorleistung** erbracht haben, brauchen Sie von den 4 Aufgaben nur 3 Aufgaben zu lösen. Alle Aufgaben sind gleich gewichtet. Kennzeichnen Sie bitte die drei Aufgaben, die Sie bearbeiten.

Schreiben Sie **leserlich** und beantworten Sie die Fragen in gegebener Knappheit gemäß der Fragestellung (*eventuell stichpunktartig* oder *durch Skizzen*)!

Viel Erfolg!

Aufgabe	1	2	3	4	Summe
erreichbare Punktzahl	20	20	-	20	60
erreichte Punktzahl	10	17	-	11	38

Aufgabe 1: Grundlagen Business Intelligence

1. Was unterscheidet die transaktionale Informationsverarbeitung (OLTP) von der analytischen Informationsverarbeitung im Unternehmen (OLAP)? Nennen und erklären Sie die wesentlichen Unterschiede anhand von Kriterien.
2. Erläutern Sie die einzelnen Schritte eines BI-Prozesses (z.B. anhand der Planung und Durchführung einer Marketing-Kampagne).
3. Sie möchten ein BI-System einrichten. Welche Schritte müssen Sie nacheinander durchführen, bis ein Anwender auf Daten ihres Data Warehouses mit Hilfe von Reporting-Anwendungen (z.B. OLAP-Anwendungen) zugreifen kann (Schritte z.B. anhand unserer SAP-BW-Übungsfallstudie)?

2. Datengewinnung

Die Daten werden aus den Vorsystemen geladen. Im ETL-Prozess werden die Daten harmonisiert. Zum Beispiel die Verkaufsdaten von Produkten aus einer SQL Datenbank.

Datenspeicherung

Die gewonnen Daten werden im zentralen Data Warehouse gespeichert.

Datenverarbeitung
~~Datenanfrage~~ / Analyse

Die Daten werden ~~gewonnen~~ ^{verarbeitet} um Analysen durchzuführen und durch Berechnungen Informationen zu gewinnen. Zum Beispiel den Umsatz den man mit bestimmten Produkten erreicht.

Datenutzung

Die Daten werden genutzt um Geschäftsprozesse zu optimieren. Die Daten die ~~bei~~ der Ausführung der neuen, optimierten Prozesse entstehen, können wiederum genutzt werden. Es werden nur Produkte beim Großhändler nachbestellt, die sich häufig verkaufen, wobei die ^{zukünftigen} Daten immer im Auge behalten werden können.

1.

	OLTP	OLAP
	transaktionell	analytisch
Datenhaltung	eindimensional	multidimensional
Performance bei analytischen Auswertungen	langsam	schnell
Zweck	Datenhaltung / Speicherung	Analyse von Daten (✓) weitere Kriterien?

OLTP wird zur Datenhaltung ~~für~~ genutzt. Es dient normalerweise zur Speicherung von Daten für operative Anwendungen.

OLAP wird zur Analyse von Daten genutzt. Hierbei werden oft Daten aus transaktionellen Vorgesystemen geladen und zu Analyse Zwecken entsprechend aufbereitet.

3. fehlt!

10/20

Aufgabe 2: Multidimensionale Datenmodellierung

Der Vorstand einer Krankenkasse beauftragt Sie mit der Entwicklung eines multidimensionalen Star-Schemas. Es geht um die Erstellung eines Datenwürfels zur Analyse von Behandlungen der Versicherten.

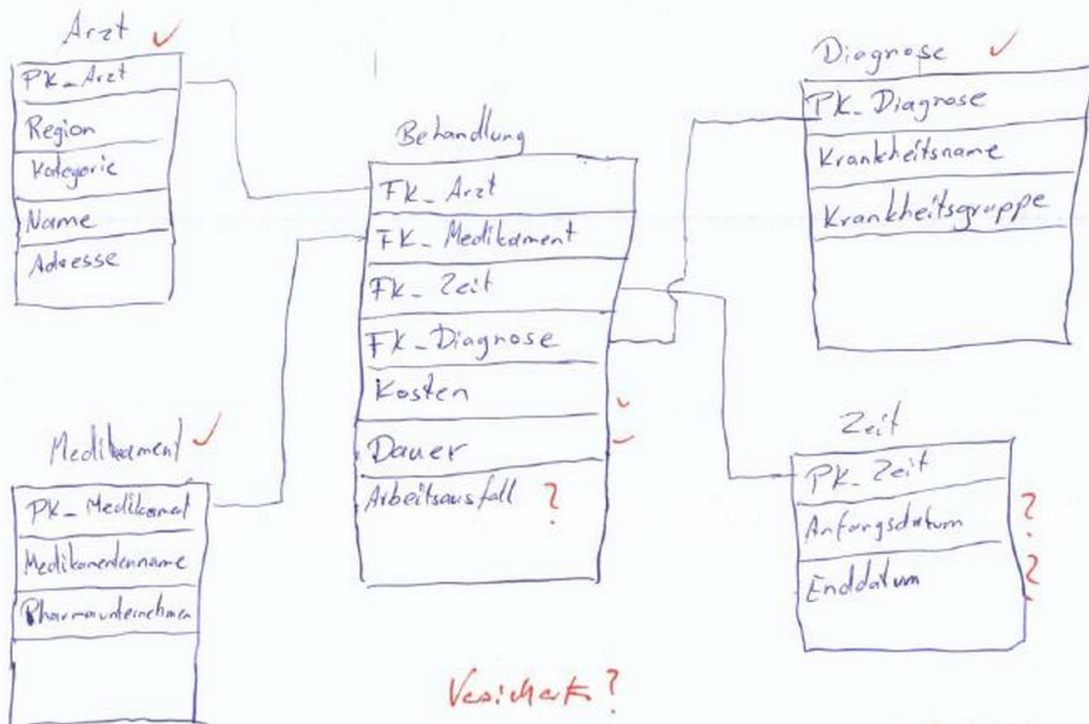
Ziele der Krankenkasse sind wirtschaftlicher Umgang mit Beiträgen und Transparenz (u.a. zur Betrugsprävention). Bestimmen Sie geeignete Kennzahlen.

Folgende Sachverhalte können wertvolle Hinweise dazu geben, welche Analyse Kriterien für die Kennzahlen relevant sein können:

- Man möchte nach Ärzten auswerten, die bestimmten Kategorien (z.B. Chirurgen, Hautärzte etc.) und Regionen zugerechnet werden können.
- Diagnosen (Krankheiten) sollen analysiert werden. Sie können Krankheitsgruppen zugerechnet werden.
- Verschiedene Medikamente und die herstellenden Pharmaunternehmen sollen auswertbar sein.

Ergänzen Sie weitere Analyse Kriterien (Merkmale), die Sie für notwendig und sinnvoll halten und stellen Sie ein Datenmodell in Form eines Star-Schemas dar.

Kennzahlen: Kosten, Dauer (der Behandlung), Arbeitsausfall (in Tagen)



Aufgabe 3: Data Warehouse Architektur /Vorgehensmodell

1. Beschreiben Sie den Aufbau und die Komponenten eines Data Warehouses i.e.S.
2. Welche besondere Bedeutung haben Metadaten in einem Data Warehouse?
3. Sie möchten ein neues BI-System in Ihrem Unternehmen als Projekt einführen. Beschreiben Sie ein dafür geeignetes Vorgehensmodell (z.B. von Kimball/Ross).

Aufgabe 4: Balanced Scorecard / Reporting

1. Skizzieren Sie den Aufbau einer Balanced Scorecard für die IT-Abteilung eines großen Telekommunikationsunternehmens. Bestimmen Sie geeignete Dimensionen. Nennen Sie zu jeder Dimension ein typisches Ziel und beschreiben dies vollständig nach der BSC-Methode.
2. Wieso unterstützt die Zielbildung mit Hilfe einer Balanced Scorecard die Konzeption und den Aufbau von Business Intelligence-Systemen?
3. Geben Sie einen kurzen Überblick über die unterschiedlichen Reporting-Anwenderwerkzeuge für BI. Wo findet sich dort eine BSC-Reporting-Anwendung?

1. Finanzperspektive

Ziel	Kennzahl	Vorgabe	Maßnahme
Umsatz steigern	Umsatz pro Mitarbeiter	+10%	Trainings für Mitarbeiter ✓

Lernperspektive

mehr Know-How im eigenen Unternehmen	Anzahl von Mitarbeitern die sich mit Virtualisierung auskennen	+8%	Aufsetzen der virtuellen Server nicht mehr auslagern an andere Unternehmen ✓
--------------------------------------	--	-----	--

Produktperspektive

Verbesserung der Leistung	kW ?	+5%	mehr Mitarbeiter für Forschung und Entwicklung einstellen
---------------------------	------	-----	---

Kundenperspektive

Qualität steigern	Anzahl von Anschüssen der bei Produktion entsteht Anteil von defekt ausgelieferten Produkten reduzieren	-2%	Sichtprüfung vor Versand (7) ✓
-------------------	---	----------------	---------------------------------------

2. Es lässt sich erkennen welche Kennzahlen ~~von~~ von Bedeutung sind und welche Auswertungen somit später wichtig sein können. Die hilft bei der Konzeptionierung und dem Aufbau eines entsprechenden BI-Systems

(3)

(✓)

Rückseite
→

3.

- Dashboards

- Monitore

- ~~• Reports~~

- Reports

Systematik?

①

Die BSC-Reporting-Anwendung ist hauptsächlich für die Planung der Geschäftsprozesse und BI-Anwendungen benötigt. Daher ist sie eher beim Management anzusehen. ?

Einordn.
in Systematik?

Business Intelligence Teil II

Wirtschaftsinformatik (Bachelor) - awis
15.01.2015

Prof. Dr. Gunther Piller

Hinweise:

- Kontrollieren Sie, ob die Klausur vollständig ist: Teil II besteht aus 4 Aufgaben
- Geben Sie oben Ihren Namen und Ihre Matrikelnummer an.
- Geben Sie die Lösung direkt unter der Aufgabe an.
- Hilfsmittel: Taschenrechner
- Viel Erfolg!

Auszug aus der Prüfungsordnung, § 12, Abs. (1), (2), (5):

(1) Eine Prüfungsleistung gilt als mit „nicht ausreichend“ bewertet, wenn Studierende bei Vorliegen selbst zu vertretender Gründe zu einem Prüfungstermin nicht erscheinen oder nach Beginn der Prüfung von der Prüfung zurücktreten. Dasselbe gilt, wenn eine Prüfungsleistung nicht innerhalb der vorgegebenen Bearbeitungszeit erbracht wird.

(2) Die für das Versäumnis eines Prüfungstermins oder für den Rücktritt nach Beginn einer Prüfung geltend gemachten Gründe müssen dem Prüfungsamt unverzüglich schriftlich angezeigt und glaubhaft gemacht werden. Bei Krankheit hat ein ärztliches Attest unverzüglich, d. h. ohne schuldhaftes Zögern, spätestens bis zum dritten Tag nach dem Prüfungstermin beim Prüfungsamt vorzuliegen. Das Attest muss die Prüfungsunfähigkeit der oder des Studierenden erkennen lassen. Die Vorlage eines amtsärztlichen Attests kann verlangt werden.

(5) Versuchen Studierende, das Ergebnis der Prüfungsleistung durch Täuschung oder Benutzung nicht zugelassener Hilfsmittel zu beeinflussen, gilt die betreffende Prüfungsleistung für diese Studierenden als mit „nicht ausreichend“ bewertet. Studierende, die den ordnungsgemäßen Ablauf der Prüfung stören, können von den jeweils Prüfenden oder Aufsichtsführenden von der Fortsetzung der Prüfung ausgeschlossen werden. In diesem Fall gilt die betreffende Prüfungsleistung als mit „nicht ausreichend“ bewertet. In besonderen Fällen kann der Prüfungsausschuss gravierendere Maßnahmen bestimmen.

bitte nicht ausfüllen

21	8	8	18	Σ	
5	8	6	17	36	
					Note 2

Aufgabe 1: Designprinzipien von SAP HANA als In-Memory DB 5121

A) Nennen und beschreiben Sie vier Designprinzipien von SAP HANA, als typisches Beispiel für eine In-Memory Datenbank.

2112

Designprinzipien
alle Daten liegen im Hauptspeicher wann, was Beschreibung ✓
es können sowohl transaktionale als auch analytische Operationen ausgeführt werden auch im ERP mit RDBMS? Beschreibung ✓
Daten werden spaltenorientiert abgelegt Daten werden spaltenorientiert abgelegt ✓
Daten werden spaltenorientiert abgelegt Beschreibung ✓

B) Welche typischen Nutzenpotentiale kann In-Memory-Datenmanagement aus Anwendersicht im Bereich Business Intelligence möglicherweise erschließen? Nennen Sie drei Nutzenpotentiale und erklären Sie diese anhand beispielhafter betriebswirtschaftlicher Anwendungsfälle.

319

Nutzenpotential	Erklärung anhand eines Beispiels
höhere Aktualisierungsrate ✓	Berechnungen die sonst nur über Nacht/am Wochenende ausgeführt werden konnten, können jetzt mehrmals täglich ausgeführt werden BIW-Anwendungsfall ✓
höhere Aktualität ✓	Zugriff auf Daten in Echtzeit BIW-Anwendungsfall ✓
höhere Bandbreite ✓	Zugriff auf mehr Daten ohne diese vorher verknüpfen zu müssen (in OLAP) ✓

Aufgabe 2: Spaltenorientierte Datenhaltung

818

Eine Tabelle besteht aus Einträgen für die Attribute A, B und C:

	Col1	Col2	Col3
Row1	A	B	C
Row2	A	B	C
Row3	A	B	C
Row4	A	B	C

Die Tabelle wird spaltenorientiert in den Hauptspeicher einer In-Memory-Anwendung geladen. Skizzieren Sie, in welcher Reihenfolge die Datenelemente abgelegt werden:

A A A A | B B B B | C C C C

212

616

Für welche Art von Datenoperationen erwarten Sie aufgrund der spaltenorientierten Datenspeicherung einen Vorteil, für welche einen Nachteil? Begründen Sie Ihre Antwort.

Analytische Auswertungen (wie z.B. der Umsatz für alle Mitarbeiter) sind sehr schnell, da die entsprechenden Daten direkt "nebeneinander" liegen.

Warum ist das ein Vorteil?

Transaktionale Auswertungen (wie z.B. alle Daten zu einem bestimmten Mitarbeiter) sind langsamer, da die Daten "weit auseinander liegen". Entsprechend ist auch das Einfügen von Datensätzen langsam.

Warum ist das ein Nachteil?

Aufgabe 3: Datenübertragung in SAP HANA

Welche zwei Möglichkeiten bestehen, SAP HANA mit Daten aus einem oder mehreren ERP-Systemen, die traditionelle Datenbanksysteme verwenden, zu versorgen? Beschreiben Sie jede kurz. Erklären Sie insbesondere wesentliche Unterschiede bezüglich der Verfügbarkeit von ERP-Daten in HANA.

Möglichkeit	Beschreibung und Erklärung
ETL	Im ETL-Prozess werden die Daten aus Vorgesystemen/aus Dateien harmonisiert und geladen.
Trigger	Trigger reagieren wenn schon Daten aus einem Vorgesystem geladen wurden und Änderungen am Datenbestand in diesem System auftreten oder zu einem bestimmten Zeitpunkt

618

(am Wochenende). Sie gleichen die Daten in HANA mit denen aus dem entsprechenden System ab.

Aufgabe 4: Dictionary & Prefix Encoding

17118

Nehmen Sie an, Sie hätten eine Tabelle mit 80 Millionen Einwohnern von Deutschland und deren Zuordnung zu ihren Wohnorten vorliegen. Rechnen Sie mit ca. 12 000 Orten. (Rechnen Sie näherungsweise mit MB=1000 KB, 1 KB=1000 B, B steht für Byte.)

A) Wieviel Speicherplatz benötigt der zur Spalte „Wohnort“ gehörende Attributvektor unter Verwendung von Dictionary Encoding? **Erläutern Sie die Schritte Ihrer Rechnung** und geben Sie das Ergebnis an.

wenn die Spalte Wohnort kodiert wird, benötigt sie:

$$2^n \Rightarrow 2^{12.000} \Rightarrow \log_2(12.000) \approx 14 \text{ Bit (pro Wohnort)}$$

$$14 \text{ bit} \cdot 80.000.000 = 140 \text{ MB Speicherplatz.}$$

616

B) Verwenden Sie Prefix Encoding mit Berlin als Prefix-Wert. Berlin hat ca. 3,5 Millionen Einwohner. Welchen Speicherplatz benötigt der komprimierte Attributvektor? **Erläutern Sie die Schritte Ihrer Rechnung** und geben Sie das Ergebnis an. Falls Sie hierbei den Speicherplatzbedarf bestimmter Größen vernachlässigen, ist dies kurz zu begründen.

Beim Prefix Encoding wird der am häufigsten auftretende Wert komprimiert, indem nur der Wert und die Anzahl ~~der~~ wie oft der Wert auftritt gespeichert wird. 11112

$$3.500.000 \cdot 14 \text{ bit} \approx 6 \text{ MB} \leftarrow \text{was ist das?}$$

Der Speicherplatz des Wertes und der Anzahl der Vorkommisse wird vernachlässigt, da die Größe vernachlässigbar ist.

$$140 \text{ MB} - 6 \text{ MB} = 134 \text{ MB}$$

Der komprimierte Attributvektor benötigt 134 MB Speicherplatz.