

Aufgabenheft zu

Klausur: Modul 32711 – Business Intelligence

Termin: 15.03.2019, 11:30–13:30 Uhr

Prüfer: Univ.-Prof. Dr. Stefan Smolnik



Bitte lesen Sie die folgenden Hinweise sorgfältig, bevor Sie mit der Bearbeitung der Klausur beginnen!

Aufbau und Umfang der Klausur

Die Klausur umfasst vier Aufgaben mit der folgenden Punkteverteilung:

Aufgabe	1	2	3	4	Summe
Maximal erreichbare Punktzahl	20	35	35	10	100

Für die Bearbeitung dieser Klausur stehen Ihnen 120 Minuten zur Verfügung. Die Klausur besteht aus einem *Aufgabenheft* und einem *Lösungsbogen*. Das Aufgabenheft umfasst sieben Seiten. Bitte prüfen Sie Ihre Klausur auf Vollständigkeit.

Hinweise zur Verwendung des Lösungsbogens

1. Tragen Sie auf dem Deckblatt des Lösungsbogens Ihre Matrikelnummer sowie Ihren Namen und Vornamen ein. Versehen Sie zusätzlich jedes Lösungsblatt mit Ihrer Matrikelnummer.
2. Die Lösungen müssen in den vorgesehenen Raum auf dem Lösungsbogen eingetragen werden. Lösungen außerhalb des vorgesehenen Raumes werden nicht in die Bewertung einbezogen.
3. Markieren Sie bei Multiple-Choice-Aufgaben die von Ihnen gewählten Alternativen durch ein Kreuz (X) an der dafür vorgesehenen Stelle im Lösungsbogen.
4. Sorgen Sie für eindeutige Eintragungen im Lösungsbogen.
5. Notizen können auf den Rückseiten der Aufgabenblätter gemacht werden. Diese Anmerkungen werden nicht in die Bewertung einbezogen.
6. Sie dürfen das Aufgabenheft vom Lösungsbogen trennen. Das Aufgabenheft muss nicht abgegeben werden. Es müssen jedoch am Ende der Klausur *alle* Blätter des Lösungsbogens *zusammengeheftet* abgegeben werden. Trennen Sie keine einzelnen Blätter ab.

Hinweise zur Bewertung der Aufgaben

Jede vollständig richtig gelöste Aufgabe oder Teilaufgabe wird mit der an Ort und Stelle angegebenen Punktzahl bewertet.

Für die Aufgabe 1 (Einfach-Wahlaufgaben 1 aus n) gilt: Es darf nur *ein* Kreuz pro Teilaufgabe gesetzt werden. Richtig gelöste Teilaufgaben werden mit der angegebenen Punktzahl bewertet. Nicht oder falsch beantwortete Teilaufgaben werden mit null Punkten bewertet.

Für die Aufgabe 4 (Richtig – Falsch) gilt: Richtig gelöste Teilaufgaben werden mit der anteiligen Punktzahl bewertet. Nicht oder falsch beantwortete Teilaufgaben werden mit null Punkten bewertet.

Für die Aufgaben 2 und 3 gilt: Teilweise richtig gelöste Aufgaben oder Teilaufgaben können mit einer entsprechend verminderten Punktzahl bewertet werden.

Hinweise zur Verwendung eines Taschenrechners

Die Verwendung eines Taschenrechners ist dann und nur dann erlaubt, wenn dieser einer der folgenden Modellreihen angehört:

- Casio fx86 oder Casio fx87
- Texas Instruments TI 30 X II
- Sharp EL 531

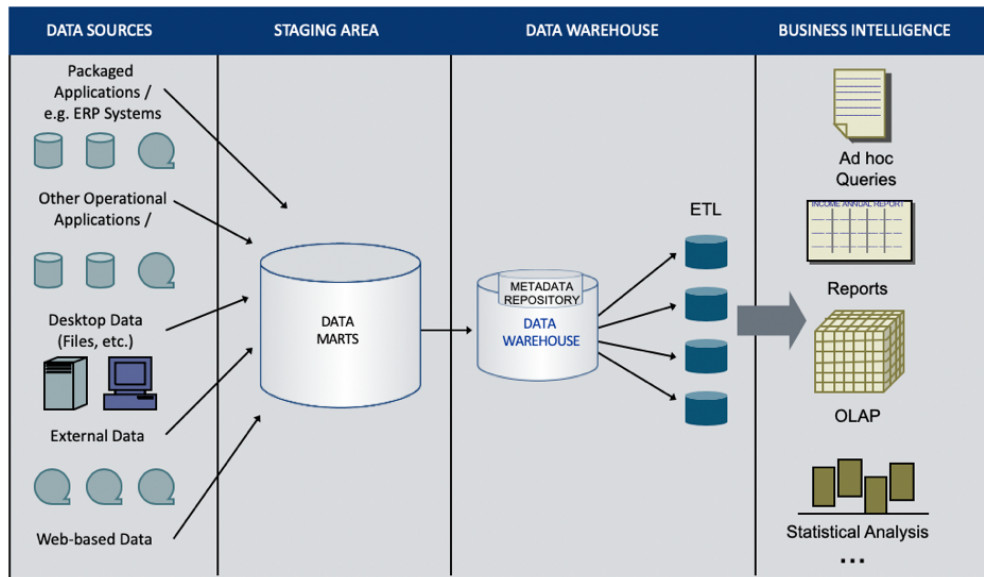
Die Verwendung anderer Taschenrechnermodelle wird als Täuschungsversuch gewertet und mit der Note „nicht ausreichend“ (5,0) sanktioniert.

Ob ein Taschenrechner einer der Modellreihen angehört, können Studierende selbst überprüfen, indem sie die vom Hersteller auf dem Rechner angebrachte Modellbezeichnung mit den oben angegebenen Bezeichnungen vergleichen: Bei *vollständiger* Übereinstimmung ist das Modell erlaubt. Ist die auf dem Rechner angebrachte Modellbezeichnung umfangreicher, enthält aber eine der oben angegebenen Bezeichnungen *vollständig*, ist das Modell ebenfalls erlaubt. In allen anderen Fällen ist das Modell nicht erlaubt. *Eventuelle Vorgänger- oder Nachfolgemodelle, die nicht in der oben aufgeführten Liste enthalten sind, sind ebenfalls nicht erlaubt.*

Aufgabe 1**(20 P)**

Markieren Sie im Lösungsbogen die zutreffende Aussage. Es ist nur eine Aussage pro Teilaufgabe korrekt, d.h. Sie müssen sich für *eine* Aussage entscheiden und dürfen nur *ein* Kreuz setzen.

1.1 In der nachfolgenden Abbildung ist die Ihnen aus dem Lehrbrief bekannte typische Applikationsarchitektur der Business Intelligence (auch bekannt als Hub-and-Spoke-Architektur) dargestellt. Was fällt Ihnen an der Abbildung auf? (5 P)



- Data Marts und ETL sind vertauscht.
- OLAP und ETL sind vertauscht.
- Data Marts und Data Warehouse sind vertauscht.
- An der Abbildung ist nichts falsch.

- 1.2 Was ist *kein* Teilschritt des Ihnen bekannten Knowledge-Discovery-in-Databases-Prozesses (KDD-Prozesses)? (5 P)
- a) Datenbereinigung
 - b) Datenauswahl
 - c) Datenerstellung
 - d) Data Mining
- 1.3 Im Lehrbrief werden verschiedene Sichtweisen auf den Intelligenzbegriff vorgestellt. Worauf bezieht sich das „Intelligence“ in Business Intelligence im Entscheidungsprozess? Es bezieht sich auf ... (5 P)
- a) ... Informationen über das Entscheidungsproblem.
 - b) ... die mangelnde Fähigkeit einer Organisation, Ziele zu verfolgen.
 - c) ... die Fähigkeit des einzelnen Individuums, eigenständig und subjektiv zu handeln.
 - d) ... eine intelligente Umsetzung wie Künstliche Intelligenz oder maschinelles Lernen.
- 1.4 Das Comprehensive Decision Model (CDM) ist ein Modell für die Systematisierung der Informationsbedarfsanalyse und besteht aus drei verschiedenen Dimensionen. Was ist *keine* Aufgabe innerhalb der Dimension Organisation? (5 P)
- a) Die klare Analyse der unternehmerischen Entscheidungsprobleme und die nachvollziehbare Dokumentation der daraus resultierenden Fragestellung.
 - b) Die Berücksichtigung des Kontexts, in dem ein Akteur entscheidet.
 - c) Die Untersuchung der Anforderungen, die an die Datenerhebung und -analyse gestellt werden.
 - d) Die Berücksichtigung der Restriktionen, aber auch der Freiheitsgrade, bei der Informationsproduktion, die durch die Organisation gewählt werden.

Übertragen Sie Ihre endgültige Lösung auf den Lösungsbogen!

Aufgabe 2**(35 P)**

Ein fiktiver Finanzdienstleister bietet neben einem Gehaltskonto auch verschiedene Versicherungen an. Der Finanzdienstleister möchte nun eine gezielte Kundenkampagne schalten, bei der Kunden angesprochen werden, die neben einem Konto über eine Haftpflichtversicherung oder eine Hausratversicherung verfügen und einen regelmäßigen monatlichen Gehaltseingang von mehr als 750 Euro aufweisen. Eine Übersicht des Kundenstamms ist ausschnitthaft in Tabelle 1 festgehalten.

Tabelle 1: Stammdaten

Kdnr.	Hausratversicherung [0=nein; 1=ja]	Haftpflichtversicherung	Kontonummer	regelmäßiger mtl. Gehaltseingang in Euro
577298	0	ja	4711	2500,00
242536	0	nein	2412	1000
337835	1	ja	-	0
483024	1	ja	3010	5000
18_820	1	nein	1234	2000
234432	1	ja	0815	1.500
989875	0	ja	0104	2750
280934	0	ja	6666	666
301085	1	nein	...	12 €
241218	1	Ja	8888	1500
32168	1	ja	5555	Keinen
234892	1	Ja	1996	500

Des Weiteren liegen folgende Informationen vor:

- Es dürfen nur 6-stellige numerische Werte verwendet werden und jeder Kunde hat mindestens eine Versicherung oder ein Konto.
- Das Vorhandensein einer Hausrat- oder Haftpflichtversicherung ist mit „ja“ oder „nein“ anzugeben.
- Die Kontonummern sind immer 4-stellig und numerisch. Wenn kein Konto vorhanden ist, wird eine „0“ hinterlegt.
- Der regelmäßige monatliche Gehaltseingang wird numerisch und ohne Sonderzeichen angegeben und beruht auf Informationen aus den Kontoumsätzen.
- Die Tabelle bildet nur einen Ausschnitt der Gesamtmenge ab.

- 2.1 Benennen Sie die vier grundsätzlichen Schritte des Entscheidungsbaumverfahrens und beschreiben Sie jeden der vier Schritte kurz mit 1-3 Sätzen. (8 P)
- 2.2 Welche fünf Datenqualitätsprobleme finden sich in Tabelle 1? Benennen und erläutern Sie diese. (5 P)

Tabelle 2: Umfrageergebnisse

ID	Geschlecht	Alter	Bereits vorhandene Versicherung	Produktinteresse für Rechtsschutzversicherung
1	weiblich	46	Haftpflichtversicherung	Ja
2	weiblich	34	Lebensversicherung	Nein
3	männlich	44	Hausratversicherung	Nein
4	männlich	45	Haftpflichtversicherung	Ja
5	männlich	27	Lebensversicherung	Nein
6	männlich	55	Hausratversicherung	Nein
7	männlich	32	Haftpflichtversicherung	Ja
8	weiblich	33	Hausratversicherung	Ja
9	weiblich	49	Hausratversicherung	Ja
10	männlich	24	Lebensversicherung	Nein
11	weiblich	65	Lebensversicherung	Nein
12	weiblich	26	Haftpflichtversicherung	Ja

- 2.3 In einem weiteren Schritt möchte der Finanzdienstleister gezielt Werbung für eine Rechtsschutzversicherung machen. Im Vorfeld führt er dazu eine Umfrage bei seinen Kunden durch. Das Ergebnis ist in Tabelle 2 dargestellt. Aufgrund dieser Daten möchte der Finanzdienstleister einen Entscheidungsbaum aufstellen. Ziel ist es, mittels des Entscheidungsbaums entscheiden zu können, ob einem gegebenen Kunden ein Werbeprospekt für eine Rechtsschutzversicherung gesendet werden soll. Dazu macht der Finanzdienstleister folgende Vorgaben:

- Zur Berechnung der Impurity soll das vorliegende Maß verwendet werden:

$$\text{Impurity} = \min\{\text{Anteil mit Interesse} = \text{ja}, \text{Anteil mit Interesse} = \text{nein}\}$$

- Qualitative Attribute sollen in ihre möglichen Werte aufgespalten werden, quantitative Attribute am arithmetischen Mittel der Werte dieses Attributs innerhalb der Trainingsinstanzen des betreffenden Knotens.
- Zur Vereinfachung nehmen Sie an, dass das Alter nur am Durchschnittsalter aufgespalten werden kann.
- Early Stopping ist nicht vorgesehen.
- Auf Pruning soll vollständig verzichtet werden.

Erstellen Sie einen Entscheidungsbaum nach den Vorgaben des Finanzdienstleisters. Dokumentieren Sie die von Ihnen vorgenommenen Zwischenschritte, um Teilpunkte zu erlangen. Sie können für die Berechnung und Zeichnung Abkürzungen (Männlich = M, Weiblich = W, Lebensversicherung = LV, Haftpflichtversicherung = HPV, Hausratversicherung = HRV) verwenden. (18 P)

- 2.4 Skizzieren Sie je zwei Vor- und Nachteile von Entscheidungsbäumen im praktischen Einsatz in je 1-2 Sätzen. (4 P)

Übertragen Sie Ihre endgültige Lösung auf den Lösungsbogen!

Aufgabe 3**(35 P)**

Um aus Daten Informationen und letztlich Wissen generieren zu können, sind mehrere Schritte notwendig. Ein Data Warehouse (DWH) bildet dabei das elementare Bindeglied zwischen der Datensammlung und der Datenanalyse.

- 3.1 Beschreiben Sie kurz die drei Ziele eines DWH. (6 P)
- 3.2 Ein DWH besteht aus verschiedenen Architekturkomponenten. Nennen und beschreiben Sie *drei* dieser Komponenten in 1-2 Sätzen. (6 P)
- 3.3 Manche Unternehmen nutzen Data Marts zusätzlich zum Einsatz eines zentralen DWH. Vergleichen Sie anhand von *sechs* ausgewählten Kriterien die Vor- und Nachteile von Data Marts gegenüber DWHs. (12 P)
- 3.4 Die Einsatzmöglichkeiten von Business-Intelligence-Lösungen sind vielfältig. Diskutieren Sie anhand von jeweils *einer* Möglichkeit beispielhaft, welche Chancen sich durch den Einsatz von Business-Intelligence-Lösungen in der *Konsumgüterherstellung* und welche sich in der *Energiebranche* ergeben. (11 P)

Übertragen Sie Ihre endgültige Lösung auf den Lösungsbogen!

Aufgabe 4**(10 P)**

Überprüfen Sie die folgenden Aussagen auf ihre Richtigkeit. Kennzeichnen Sie im Lösungsbogen *uneingeschränkt* zutreffende Aussagen mit einem Kreuz bei „Richtig“ und bei allen anderen Aussagen mit einem Kreuz bei „Falsch“ .

- 4.1 Business Intelligence (BI) ist die an einem Entscheidungsproblem orientierte Analyse und adressatengerechte, technikgestützte Bereitstellung von wettbewerbsrelevanten Informationen, mit dem Ziel, den Entscheidungsprozess zu unterstützen.
- 4.2 Die Abkürzung ETL steht in der BI für Extraktion – Tauchen – Laden.
- 4.3 Business Intelligence 2.0 umfasst unter anderem OpenSource BI.
- 4.4 Right-Time BI bedeutet, dass Informationen zur richtigen Zeit am richtigen Ort für den richtigen Adressaten zur Verfügung gestellt werden.
- 4.5 Die Regressionsanalyse dient der statistischen Modellierung von Approximationsproblemen.
- 4.6 Traditionelle BI beschäftigt sich vor allem mit der Analyse unstrukturierter Daten
- 4.7 Bei OpenSource BI wird der Betrieb der BI-Lösung an einen externen Dienstleister ausgelagert.
- 4.8 Metadaten zeichnen sich durch zwei wesentliche Eigenschaften aus: ihr Alleinstellungsmerkmal und ihr Detaillierungsmerkmal.
- 4.9 Anders als bei DWHs können Anwender in den Fachabteilungen auf Data Marts nicht direkt zugreifen.
- 4.10 Roll-Down und Drill-Up gehören zu den wesentlichen Operationen in einem OLAP-Würfel.

Übertragen Sie Ihre endgültige Lösung auf den Lösungsbogen!

Fakultät für Wirtschaftswissenschaft

Matrikelnr.

--	--	--	--	--	--	--

Name

Vorname

Lösungsbogen

Klausur: Modul 32711 - Business Intelligence

Termin: 15.03.2019 11:30 – 13:30 Uhr

Prüfer: Univ.-Prof. Dr. Stefan Smolnik

Aufbau und Bewertung der Klausur

Aufgabe	1	2	3	4	Summe
Maximal erreichbare Punktzahl	20	35	35	10	100
Erreichte Punktzahl					

Datum:

Note:

--	--	--	--	--	--	--

Aufgabe 1 (ankreuzen)

(20 P)

	a)	b)	c)	d)
1.1				
1.2				
1.3				
1.4				

Aufgabe 2

(35 P)

--	--	--	--	--	--	--

--	--	--	--	--	--	--

--	--	--	--	--	--	--

--	--	--	--	--	--	--

--	--	--	--	--	--	--

--	--	--	--	--	--	--

Aufgabe 3

(35 P)

--	--	--	--	--	--	--

--	--	--	--	--	--	--

--	--	--	--	--	--	--

--	--	--	--	--	--	--

--	--	--	--	--	--	--

--	--	--	--	--	--	--

Aufgabe 4 (Ein Kreuz bei Richtig oder Falsch eintragen)

(10 P)

	Richtig	Falsch
4.1		
4.2		
4.3		
4.4		
4.5		
4.6		
4.7		
4.8		
4.9		
4.10		