

Prüfung 4: Datenbanken

Aufgabe 1: (6 Punkte)

Die Tabelle unten enthält eine Statistik mit Bestellungen von Produkten eines Telekom-Unternehmens. Ähnliche Produkte gibt es bei verschiedenen Telekoms; die Preise sind frei erfunden. Die einzelnen Spalten bedeuten:

- Datum: An welchem Tag die Bestellung erfasst wurde.
- Festnetz: Ob ein analoger Anschluss oder ein digitaler ISDN-Anschluss bestellt wurde.
- DSL: Das bestellte Internet-Abonnement.
- TV: Das bestellte Digital-TV-Abonnement.
- Kombiangebot: Welches Kombiangebot (mit Rabatt) die Produkte darstellen.
- Preis: Wie viel die bestellten Produkte kosten.
- Verarbeitung: Ob die Bestellung vollständig automatisiert verarbeitet werden kann oder ob sie von Mitarbeitern manuell bearbeitet werden muss.
- Bestellkanal: Wie die Bestellung erfasst wurde, in einem Shop der Telekom, durch Direkt Marketing-Mitarbeiter, online durch den Kunden selbst oder durch eine Partner-Firma.

Datum	Festnetz	DSL	TV	Kombiangebot	Preis	Verarbeitung	Bestellkanal
2010-08-22	analog	max	plus	triobundle	100	automatisch	Shop
2010-08-28	isdn	mini	plus		120	manuell	Kunde Online
2010-08-29	analog	max	plus	triobundle	100	manuell	Kunde Online
2010-08-26	analog	standard	plus		70	automatisch	Kunde Online
2010-08-22	analog	max	plus	triobundle	100	automatisch	Kunde Online
2010-08-26	analog	mini			30	automatisch	Kunde Online
2010-08-22	analog	max	plus		90	automatisch	Partner
2010-08-23	isdn	max	plus	triobundle	110	automatisch	Kunde Online
2010-08-23	analog	start	basic		70	manuell	Kunde Online
2010-08-25	analog	max	plus	triobundle	100	automatisch	Partner
2010-08-26	analog	max	basic		80	manuell	Kunde Online
2010-08-25	analog	max	plus	triobundle	90	automatisch	Direkt Marketing
2010-08-23	analog	standard			70	automatisch	Direkt Marketing
2010-08-25	analog	standard			70	automatisch	Direkt Marketing
2010-08-23	analog				20	manuell	Partner

Aufgabe 1.1 (2 Punkte)

Für eine Auswertung sollen nur diejenigen Bestellungen angezeigt werden, bei denen der Kunde einen **analogen Festnetzanschluss** mit **DSL** Abonnement **standard** oder **max** bestellt, und die **automatisch** verarbeitet werden können. Zudem sollen auch alle Bestellungen angezeigt werden, bei denen ein Kunde im Bestellkanal „Kunde Online“ einen ISDN-Anschluss bestellt hat, sofern der Preis mindestens CHF 120.- betrug und die Verarbeitung manuell geschah.

Tragen Sie in der folgenden Tabelle den dafür notwendigen Spezialfilter ein (nicht benötigte Zeilen können Sie einfach leer lassen):

Datum	Festnetz	DSL	TV	Kombiangebot	Preis	Verarbeitung	Bestellkanal

Aufgabe 1.2 (2 Punkte)

Beschreiben Sie in wenigen Sätzen anschaulich, welche Filtermöglichkeiten ein Autofilter in Excel bietet:

Erklären Sie nun möglichst prägnant, warum für die Auswertung von Aufgabe 1.1 ein Auto-Filter nicht ausreichte, warum ein Spezialfilter verwendet werden musste:

Aufgabe 1.3 (2 Punkte)

Sie finden in dem Excel mit den Bestelldaten der Aufgabe 1.1 in einem weiteren Sheet folgende Tabelle:

Zeilenbeschriftungen	Summe von Preis	Mittelwert von Preis	Anzahl von Preis
analog	990.00	76.15	13
automatisch	720.00	80.00	9
Direkt Marketing	230.00	76.67	3
Kunde Online	200.00	66.67	3
Partner	190.00	95.00	2
Shop	100.00	100.00	1
manuell	270.00	67.50	4
Kunde Online	250.00	83.33	3
Partner	20.00	20.00	1
isdn	230.00	115.00	2
automatisch	110.00	110.00	1
Kunde Online	110.00	110.00	1
manuell	120.00	120.00	1
Kunde Online	120.00	120.00	1
Gesamtergebnis	1220.00	81.33	15

Beschreiben Sie prägnant und präzise, was die Tabelle darstellt und wie eine solche Tabelle in Excel erstellt werden kann:

Aufgabe 2: Analyse von Geschwindigkeitsmessungen mit SQL [6 Punkte]

Die Tabelle unten enthält die Aufzeichnung einer Radarkontrolle: Geschwindigkeitsmessung in km/h, Zeitpunkt und Datum der Messung, in welche Richtung das Fahrzeug fuhr, wie lang das Fahrzeug in mm war, daraus abgeleitet der Fahrzeugtyp, um welche Stunde die Messung stattfand, und ob eine Busse erteilt wird oder nicht.

Tempo	Zeit	Datum	Fahrt- richtung	Fahrzeuglänge	Typ	Stunde	Busse
29	15:32:36	05-27-03	1	5'100	Lieferwagen	15	nein
63	15:33:58	05-27-03	1	1'100	Töff	15	ja
30	15:34:02	05-27-03	1	4'800	PKW	15	nein
32	15:34:59	05-27-03	1	4'900	PKW	15	nein
37	15:35:49	05-27-03	1	5'000	Lieferwagen	15	nein
34	15:36:30	05-27-03	1	4'700	PKW	15	nein
65	15:36:39	05-27-03	1	2'600	Töff	15	ja
45	15:37:27	05-27-03	1	3'500	PKW	15	nein
46	15:37:29	05-27-03	1	3'700	PKW	15	nein
56	16:14:05	05-27-03	2	5'500	Lieferwagen	16	ja
59	16:14:19	05-27-03	2	4'800	PKW	16	ja
45	16:14:32	05-27-03	2	5'500	Lieferwagen	16	nein
47	16:14:38	05-27-03	2	5'500	Lieferwagen	16	nein
62	16:15:35	05-27-03	2	6'500	Lieferwagen	16	ja
50	16:15:42	05-27-03	2	4'600	PKW	16	nein
51	16:15:45	05-27-03	2	5'200	Lieferwagen	16	nein
75	16:16:19	05-27-03	2	5'800	Lieferwagen	16	ja
50	16:16:40	05-27-03	2	5'500	Lieferwagen	16	nein
45	16:16:47	05-27-03	2	4'500	PKW	16	nein
48	16:16:48	05-27-03	2	5'900	Lieferwagen	16	nein
53	16:17:59	05-27-03	1	4'100	PKW	16	nein
54	16:19:43	05-27-03	1	5'100	Lieferwagen	16	ja
82	16:32:42	05-27-03	1	4'800	PKW	16	ja
47	17:12:10	05-27-03	2	4'400	PKW	17	nein
48	17:14:34	05-27-03	2	5'200	Lieferwagen	17	nein
60	17:16:46	05-27-03	2	4'200	PKW	17	ja

Aufgabe 2.1 (2 Punkte)

Schreiben Sie eine SQL-Anweisung, um die folgende Frage zu beantworten: Wie schnell fuhr der schnellste PKW zwischen 16:00 und 16:59? Ihre Anweisung soll das Tempo und die Fahrzeuglänge der gesuchten Messung ausgeben (Hinweis: geschickte Sortierung und Limitierung helfen Ihnen weiter):

Aufgabe 2.2 (2 Punkte)

Schreiben Sie eine SQL-Anweisung, um exakt die folgende Tabelle zu erzeugen:

TYP	ANZAHL	LAENGE
Lieferwagen	12	5483
Töff	2	1850
PKW	12	4416

Es soll für jeden Fahrzeugtyp die Anzahl Messungen zu diesem Typ ausgegeben werden sowie die durchschnittliche Länge der Fahrzeuge, und die Spaltenbeschriftungen sollen der Tabelle oben entsprechen (Gross-/Kleinschreibung ist irrelevant):

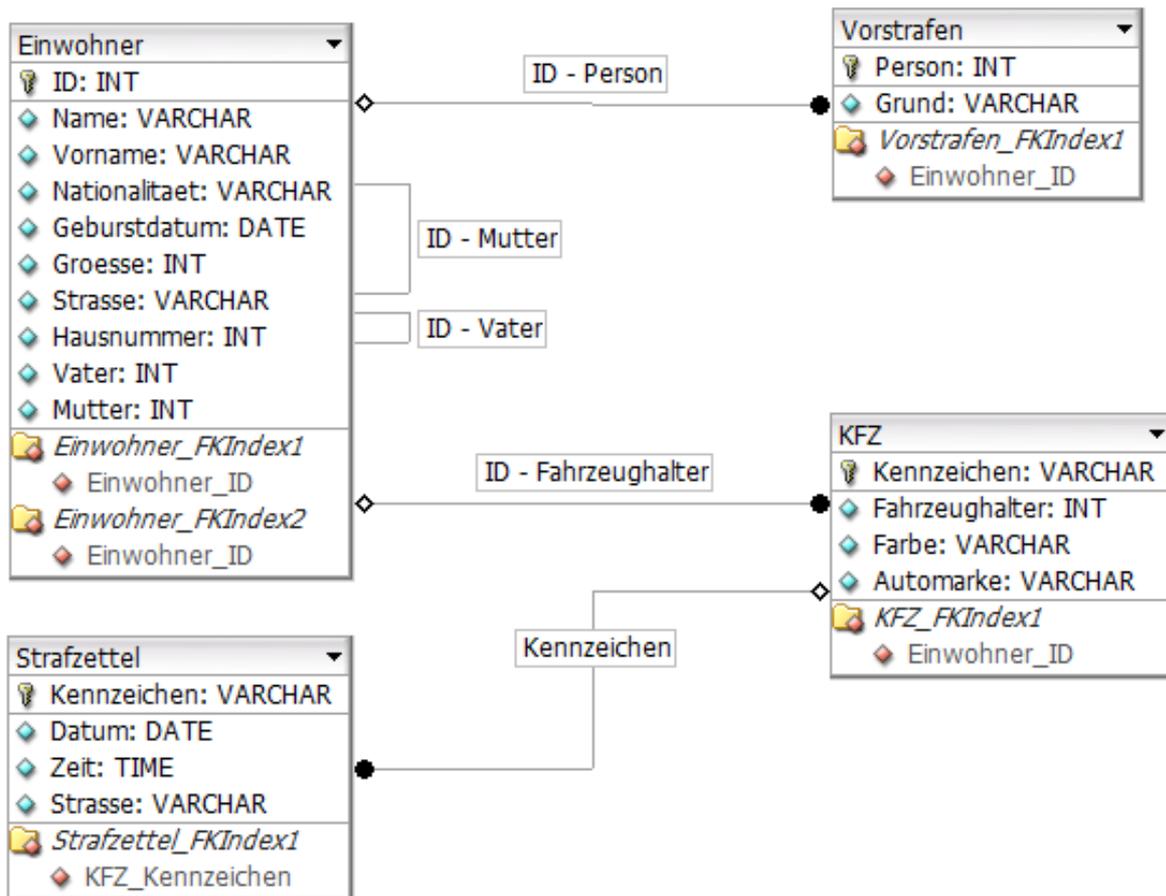
Aufgabe 2.3 (2 Punkte)

Beschreiben Sie in Worten präzise, welche Auswertung die folgende SQL-Anweisung ausgibt:

```
SELECT tempo, COUNT(tempo)
FROM messungen
WHERE (typ='PKW')
GROUP BY tempo
HAVING COUNT(tempo) > 1
ORDER BY tempo ASC
```

Aufgabe 3: Abfragen über mehrere Tabellen mit SQL (6 Punkte)

Wir betrachten bei den folgenden beiden Aufgaben eine kleine Datenbank mit Angaben zu Einwohnern, ihren Vorstrafen, ihren Autos und den Strafzetteln, die auf diese Autos ausgestellt wurden:



Die Tabellen beinhalten folgende Daten:

- Einwohner enthält die Angaben zu den Einwohnern einer fiktiven Ortschaft
- Vorstrafen enthält die Vorstrafen (ein Einwohner kann mehrere Vorstrafen haben)
- KFZ beschreibt die Fahrzeuge der Einwohner (ein Einwohner kann mehrere KFZ haben)
- Strafzettel beschreibt die auf ein KFZ ausgestellten Strafzettel (ein KFZ kann mehrere Strafzettel haben)

Es gelten folgende Beziehungen:

- Vorstrafen.person ist ein Fremdschlüssel zu Einwohner.ID
- Einwohner.Mutter ist ein Fremdschlüssel zu Einwohner.ID
- Einwohner.Vater ist ein Fremdschlüssel zu Einwohner.ID
- KFZ.Fahrzeughalter ist ein Fremdschlüssel zu Einwohner.ID
- Strafzettel.Kennzeichen ist ein Fremdschlüssel zu KFZ.Kennzeichen

Tabelle Einwohner

ID	NAME	VORNAME	NATIONALITAET	GEBURTSDATUM	GROESSE	VATER	MUTTER
1	Anderson	George	USA	02.11.1932	172	(null)	(null)
10	Armstrong	Lars	Kanada	01.09.1977	182	(null)	(null)
6	Anderson	Julia	USA	23.04.1956	162	3	4
4	Smith	Sara	Australia	16.07.1937	167	(null)	(null)
3	Smith	Tom	Australia	07.12.1938	181	(null)	(null)
8	Jones	Sahra	USA	16.09.1956	154	(null)	8
12	Faith	Tom	USA	22.07.1951	176	(null)	(null)
13	Faith	Salma	USA	31.08.1954	161	(null)	(null)
14	Faith	Larry	USA	02.11.1976	180	12	13
15	Thommas	Max	USA	07.12.1978	182	15	(null)
18	Jang	Lie	China	15.08.1982	182	(null)	19
19	Jang	Chia	China	12.06.1971	165	18	18

Tabelle Vorstrafen

PERSON	GRUND
2	Ladendiebstahl
8	Erpressung
11	Raub und schwerer Einbruch
12	Raub
1	Raub und Erpressung
2	Raub und schwerer Einbruch
2	Körperverletzung
10	Raub und Ladendiebstahl
11	Ladendiebstahl
18	Körperverletzung
11	Steuerhinterziehung
1	schwerer Einbruch

Tabelle KFZ

KENNZEICHEN	FAHRZEUGHALTER	FARBE	AUTOMARKE
E405AD3	1	blau	GM
E4D4557	2	weiss	Opel
593FDNS	3	lila	Opel
4576557	8	weiss	Toyota
X345557	10	blau	BMW
4334DDD	11	gelb	Ford
Y65478D	12	schwarz	Peugot
L49210F	14	silber	Mercedes
L344DF3	16	rot	BMW
X3332SD	18	rot	VW
X4440AQ	20	braun	VW

Tabelle Strafzettel

KENNZEICHEN	DATUM	ZEIT	STRASSE
E405AD3	06.07.2009	10:41	St Jones Road
E405AD3	07.09.2009	06:43	St Jones Road
E405AD3	15.08.2009	12:52	St Jones Road
L49210F	14.09.2009	14:23	Placa Avante
X3332SD	10.09.2009	12:13	August Street
X4440AQ	26.08.2009	16:07	St Jones Road
X4440AQ	15.08.2009	18:15	St Jones Road
4334DDD	04.07.2009	09:20	Elmar Street
593FDNS	08.08.2009	12:20	Clester Street

Aufgabe 3.1 (2 Punkte)

Welche Farben haben die Autos von Einwohnern, die einen Mercedes oder einen BMW fahren, und welche Nationalität haben diese Einwohner? Die Antwort lautet (die Reihenfolge der Resultate spielt keine Rolle):

NAME	VORNAME	NATIONALITAET	AUTOMARKE	FARBE
Armstrong	Lars	Kanada	BMW	blau
Faith	Larry	USA	Mercedes	silber

Schreiben Sie eine SQL-Anweisung, welche die obige Frage beantwortet. Die Ausgabe Ihrer Anweisung soll exakt wie oben aussehen:

Achtung: Die Datenqualität ist nicht über alle Zweifel erhaben! Es hat auch Fremdschlüssel, zu denen es keinen Primärschlüssel gibt... da haben die Programmierer nicht aufgepasst... lassen Sie sich dadurch nicht verwirren.

Aufgabe 3.2 (2 Punkte)

Sie finden in einem Programm folgende SQL-Anweisung:

```
SELECT name, count(*) AS anzahl  
FROM einwohner  
JOIN kfz ON id = fahrzeughalter  
JOIN strafzettel USING (kennzeichen)  
WHERE farbe = 'blau'  
GROUP BY name  
HAVING count(*) > 1
```

Tragen Sie in folgender Tabelle das exakte Resultat dieser Abfrage ein. Nicht-benötigte Spalten und Zeilen können Sie einfach leer lassen:

Formulieren Sie präzise, 1. welche Frage die SQL-Anweisung beantwortet und 2. beschreiben Sie, welche Daten die Anweisung wie ausgibt:

Aufgabe 3.3 (2 Punkte)

Erklären Sie anhand von der obigen SQL-Abfrage prägnant, was der Unterschied zwischen WHERE und HAVING ist:

Aufgabe 4: Abfragen über mehrere Tabellen mit SQL (6 Punkte)

Wir betrachten für die folgenden Aufgaben eine kleine Filmdatenbank mit Angaben zu Personen (Schauspielern und Regisseuren), Filmen und wer welche Rolle in den Filmen spielt.

personen			
id	vorname	nachname	geschlecht
1	Sean	Connery	m
2	Alec	Baldwin	m
3	John	McTiernan	m
4	Robert	Redford	m
5	Sydney	Pollack	m
6	Faye	Dunaway	w
7	Craig	Sheffer	m

personen			
id	vorname	nachname	geschlecht
8	Brad	Pitt	m
9	Kathryn	Bigelow	w
10	Jeremy	Renner	m
11	Anthony	Mackie	m
12	Clint	Eastwood	m
13	Meryl	Streep	w
14	Hilary	Swank	w

filme					
id	titel	jahr	bewertung	anzahl_bewertungen	regisseur
1	The Hunt for Red October	1990	7.6	56882	3
2	ThreeDays of the Condor	1975	7.5	14777	5
3	A River RunsThrough It	1992	7.1	20429	4
4	The Hurt Locker	2008	7.8	92696	9
5	The Bridges of Madison County	1995	7.8	22833	12
6	Million Dollar Baby	2004	8.2	151594	12

Rollen				
film_id	schauspieler_id	rollen_vorname	rollen_nachname	reihenfolge
1	1	Marko	Ramius	1
1	2	Jack	Ryan	2
2	4	Joseph	Turner	1
2	6	Kathy	Hale	2
3	7	Norman	Maclean	1
3	8	Paul	Maclean	2
4	10	William	James	1
4	11	JT	Sanborn	2
5	12	Robert	Kincaid	1
5	13	Francesca	Johnson	2
6	12	Frankie	Dunn	1
6	14	Maggie	Fitzgerald	2

Es gelten folgende Beziehungen:

- filme.regisseur ist ein Fremdschlüssel in die Tabelle personen
- rollen.film_id ist ein Fremdschlüssel in die Tabelle filme
- rollen.schauspieler_id ist ein Fremdschlüssel in die Tabelle personen

Aufgabe 4.1 (2 Punkte)

Quizfrage: Wer ist der Regisseur des Filmes mit den meisten Bewertungen, wie heisst dieser Film, wann erschien er, und wie viele Bewertungen erhielt dieser Film? Dabei sollen nur Filme berücksichtigt werden, die vor 2000 erschienen sind.

nachname	vorname	titel	jahr	anzahl_bewertungen

Schreiben Sie eine SQL-Anfrage, um die obige Quizfrage zu beantworten. Ihre Anfrage soll exakt die Daten ausgeben, die für die obige Tabelle benötigt werden:

Aufgabe 4.2 (2 Punkte)

Gesucht sind die Namen aller Schauspielerinnen sowie ihrer jeweiligen Rollennamen, inkl. der Reihenfolge, in der die Rolle aufgelistet ist:

NACHNAME	VORNAME	ROLLEN_NACHNAME	ROLLEN_VORNAME	REIHENFOLGE
Dunaway	Faye	Hale	Kathy	2
Streep	Meryl	Johnson	Francesca	2
Swank	Hilary	Fitzgerald	Maggie	2

Die Tabelle ist nach Nachnamen aufsteigend sortiert. Schreiben Sie eine SQL-Anfrage, die exakt obige Daten ausgibt:

Aufgabe 4.3 (1 Punkt)

In einem Programm finden Sie folgende SQL-Anweisung:

```
SELECT schauspieler.nachname, schauspieler.vorname, COUNT(*) AS anzahl  
FROM personen schauspieler  
JOIN rollen rolle ON rolle.schauspieler_id = schauspieler.id  
GROUP BY schauspieler.nachname, schauspieler.vorname  
HAVING COUNT(*) > 1
```

Welche Resultate liefert diese Anfrage? Füllen Sie die folgende Tabelle aus, inklusive Spaltenbeschriftung. Überflüssige Zeilen und Spalten einfach leer lassen:

Aufgabe 4.4 (1 Punkt)

Beschreiben Sie in wenigen Worten präzise, welche Frage die obige SQL-Anfrage beantwortet:

--