

Klausur Datenbanken
Wintersemester 2016/2017
Prof. Dr. Wolfgang May
8. Februar 2017, 14-16 Uhr
Bearbeitungszeit: 90 Minuten

Vorname:

Nachname:

Matrikelnummer:

Bei der Klausur sind **keine Hilfsmittel** (Skripten, Taschenrechner, etc.) erlaubt. Handies müssen ausgeschaltet sein. Papier wird gestellt. Benutzen Sie nur die **ausgeteilten**, zusammengehefteten **Blätter** für Ihre Antworten. Schreiben Sie mit blauem/schwarzem Kugelschreiber, Füller, etc.; Bleistift ist nicht erlaubt.

Zum **Bestehen** der Klausur sind **45** Punkte hinreichend.

- meine Note soll mit Matrikelnummer so bald wie möglich auf der Vorlesungs-Webseite veröffentlicht werden.
- meine Note soll nicht veröffentlicht werden; ich erfahre sie dann aus FlexNever oder beim zuständigen Prüfungsamt.

	Max. Punkte	Erreichte Punkte
Aufgabe 1 (ER-Modell)	16	
Aufgabe 2 (Transformation in das Relationale Modell)	18	
Aufgabe 3 (SQL und Relationale Algebra)	42	
Aufgabe 4 (Verschiedenes)	14	
Summe	90	

Note:

Themenstellung: Formel-1

Alle Klausuraufgaben basieren auf einem gemeinsamen “Auftrag”: In der Klausur soll eine Datenbank über die Formel-1 entworfen werden, die Daten seit deren Einführung im Jahr 1950 enthält:

1. Formel-1-*Teams* haben einen Namen, sind in einem Land zuhause (hier wird angenommen, dass sie nie in ein anderes Land wechseln), und außerdem ist gespeichert, wann sie gegründet wurden (was auch vor der Einführung der Formel-1 gewesen sein kann; einige davon sind nicht mehr aktiv, aber das spielt an dieser Stelle keine Rolle): *Ferrari* (seit 1929) und *Alfa Romeo* (seit 1913) sind Teams aus *Italien*. *Mercedes* (seit 1894) ist ein Team aus *Deutschland*. *Red Bull*³ ist ein 2005 gegründetes Team aus *Österreich*. *McLaren* (gegründet 1965) ist ein Team aus *Großbritannien*.
2. Zu jedem Land ist außer dem Namen auch gespeichert, auf welchem (eindeutigen) Kontinent es liegt (Russland und die Türkei werden zu Europa gerechnet).
3. Zu jedem *Fahrer* ist der Name (es wird angenommen, dass es keine zwei Fahrer mit demselben Namen gibt), sein Geburtsdatum, und sein *Heimatland* gespeichert (auch wenn er schon lange in Monaco wohnt). *Sebastian Vettel* wurde am 3.7.1987 geboren, und ist aus *Deutschland*. *Nico Rosberg* wurde am 27.6.1985 geboren, und ist ebenfalls aus *Deutschland*. *Giuseppe Farina* wurde am 30.10.1906 geboren und war *Italiener*.
4. Zu jedem Fahrer ist gespeichert, bei welchen Teams er wann unter Vertrag stand bzw. steht. *Sebastian Vettel* stand vom 1.1.2007 bis 31.7.2007 bei *Sauber* unter Vertrag, dann vom 1.8.2007-31.12.2008 bei *Toro Rosso*⁴, vom 1.1.2009-31.12.2014 bei *Red Bull*, und seit 1.1.2015 bei *Ferrari*. *Nico Rosberg* stand vom 1.1.2006 bis 31.12.2009 bei *Williams* unter Vertrag, und vom 1.1.2010 bis 31.12.2016 bei *Mercedes*. *Giuseppe Farina* stand vom 1.1.1950 bis 31.12.1951 bei *Alfa Romeo* unter Vertrag, und vom 1.1.1952 bis 31.12.1955 bei *Ferrari*.
Es kann vorkommen, dass ein Fahrer erst bei einem Team T_1 unter Vertrag steht, dann bei einem oder mehreren anderen, und später nochmal bei T_1 .
5. Eine Saison entspricht jeweils einem Kalenderjahr. In jeder Saison finden mehrere (5 bis 25) Rennen auf unterschiedlichen Rennstrecken statt; auf jeder Strecke maximal ein Rennen pro Saison:
6. Jede Strecke hat einen Namen (meistens der Name einer Stadt), ausserdem ist gespeichert, in welchem Land sie liegt. Der *Hockenheimring* und der *Nürburgring* liegen in *Deutschland*, *Monza* liegt in *Italien*, *Suzuka* liegt in *Japan*.
7. In der Saison 2016 fanden u.a. Rennen auf dem *Hockenheimring* am 31.7.2016, in *Monza* am 4.9.2016 und in *Suzuka* am 9.10.2016 statt. In der Saison 2015 fanden u.a. Rennen in *Monza* am 6.9.2015 und in *Suzuka* am 27.9.2015 (und kein Rennen auf dem *Hockenheimring*) statt.
In der ersten Saison fand u.a. am 3.9.1950 ein Rennen in *Monza* statt.
8. Für jede Teilnahme eines Fahrers an einem Rennen wird gespeichert, welche Platzierung er erreicht hat, oder ob er ausgefallen ist.
In der Saison 2016 hat *Nico Rosberg* das Rennen in *Monza* auf dem 1. Platz beendet; *Lewis Hamilton* wurde Zweiter, *Sebastian Vettel* wurde Dritter. In der Saison 2012 hat *Sebastian Vettel* das Rennen in *Suzuka* auf dem 1. Platz beendet, *Nico Rosberg* und *Fernando Alonso* sind in diesem Rennen ausgefallen. In der Saison 2008 hat *Sebastian Vettel* das Rennen in *Monza* auf dem 1. Platz beendet. Das Rennen in *Monza* in der Saison 1950 hat *Giuseppe Farina* gewonnen.

³engl. für “rotes männliches unkastriertes Rindvieh”

⁴ital. für “rotes männliches unkastriertes Rindvieh”

Aufgabe 1 (ER-Modell [16 Punkte])

Entwickeln Sie ein ER-Modell für das Szenario. Geben Sie darin die Schlüsselattribute sowie die Beziehungskardinalitäten an.

Aufgabe 2 (Transformation in das Relationale Modell [18 Punkte])

- a) **Lösen Sie diesen Aufgabenteil auf dem *letzten* Blatt und trennen dieses ab (und geben es am Ende mit ab!).** Dann haben Sie dieses Blatt separat zugreifbar um später damit die Aufgaben 2b, 3 und 4 (SQL, Relationale Algebra+SQL, Diverses) zu lösen.

Geben Sie an, welche Tabellen (mit Attributen, Schlüsseln etc.) Ihre Datenbank enthält (keine SQL CREATE TABLE-Statements, sondern einfach grafisch). (12 P)

Markieren Sie dabei auch Schlüssel (durch unterstreichen) und Fremdschlüssel (durch überstreichen).

Geben Sie die Tabellen mit jeweils mindestens zwei Beispieldupeln (z.B. denen, die sich aus dem Aufgabentext ergeben, und weiteren erfundenen) an.

- b) Geben Sie die CREATE TABLE-Statements für diejenige Tabelle, in der die Rennergebnisse gespeichert sind, so vollständig wie möglich an (6 P).

Aufgabe 3 (SQL und Relationale Algebra [42 Punkte])

Verwenden Sie für diese Aufgabe die von Ihnen entworfene relationale Datenbasis. Keine der Antworten soll Duplikate enthalten.

- a) Geben Sie eine SQL-Anfrage **und** einen Algebra-Ausdruck oder -Baum an, die alle Paare (Fahrer, Rennstrecke) ausgeben, bei denen ein Fahrer bei einem Rennen in Asien einen der ersten 3 Plätze belegt hat. (2+2 P)
- b) Geben Sie eine SQL-Anfrage an, die für jeden Fahrer, der *mindestens in zwei Saisons* ein Rennen gewonnen hat, dessen Namen und die Anzahl der in seiner gesamten Karriere gewonnenen Rennen ausgibt. (3 P)
- c) Geben Sie eine SQL-Anfrage **und** einen Algebra-Ausdruck oder -Baum an, die die Namen aller Fahrer ausgeben, die nie einen Vertrag bei einem deutschen Team hatten. (3+4 P)
- d) Geben Sie eine SQL-Anfrage **und** einen Algebra-Ausdruck oder -Baum an, die die Namen aller Fahrer ausgeben, die auf allen in Asien gelegenen Rennstrecken schon mindestens einmal gefahren sind und dabei auch im Ziel angekommen sind. (5+5 P)
- e) Geben Sie eine SQL-Anfrage **und** einen Algebra-Ausdruck oder -Baum an, die die Namen aller Fahrer ausgeben, die für jedes Team, bei dem sie einmal unter Vertrag standen, mindestens ein Rennen gewonnen haben. (6+6 P)
- f) Ein bisschen Theorie (6 P).

Seien $R(A, B)$, $S_1(B)$ und $S_2(B)$ Relationen.

Beweisen oder widerlegen Sie, ob die folgende Gleichung allgemeingültig ist:

$$(R \div S_1) \cap (R \div S_2) = R \div (S_1 \cup S_2)$$

Name:

MatNr.:

Aufgabe 4 (Verschiedenes [14 Punkte])

- a) Geben Sie eine SQL-Anfrage **und** einen Algebra-Ausdruck oder -Baum an, die alle Tupel (Fahrername, Teamname, Strecke) ausgeben, so dass ein Fahrer für ein Team seines Heimatlandes ein Rennen auf einer Strecke in seinem Heimatland gewonnen hat. (3+3P)
- b) Begründen Sie, warum die obige Anfrage als ein “zyklisches Join” bezeichnet wird. (2P)
- c) Skizzieren (und begründen) Sie, wie die interne Auswertung obiger Anfrage am effizientesten durchgeführt werden könnte. (6 P)
(Die Datenbank enthält ca. 760 Fahrer aus 100 Ländern, 70 Strecken in 25 Ländern, 300 Teams aus 20 Ländern; 70 Saisons mit jeweils 5-25 Rennen; insgesamt ca. 1000 Rennen mit je ca. 20 Teilnehmern, und damit 20000 Ergebnis-Einträgen, davon 1000 Einträge für einen 1. Platz.)