

Einführung in Datenbanken

— Übungsblatt 11 (SQL Test) —

Ihre Lösungen laden Sie bitte in die Übungsplattform in StudIP hoch ([StudIP-Eintrag der Vorlesung], Reiter „Übungsplattform“).

Einsendeschluss ist Montag, der 13.01.2025, 18⁰⁰.

Hausaufgaben können einzeln oder in Zweier-Gruppen bearbeitet werden. (Da dieses Blatt eine Art Probeklausur ist, ist Einzelabgabe empfohlen, aber nicht verpflichtend.)

Denken Sie daran, dass Sie bei jeder Aufgabe angeben müssen, ob Sie bereit sind, vorzurechnen:

- „VORRECHNEN:0“: Ich werde nicht zur Übung kommen.
- „VORRECHNEN:1“: Ich möchte diese Aufgabe nicht vorrechnen.
- „VORRECHNEN:2“: Ich möchte diese Aufgabe nur ungern vorrechnen.
- „VORRECHNEN:3“: Ich kann vorrechnen, lasse aber gern anderen den Vortritt.
- „VORRECHNEN:4“: Ich kann problemlos vorrechnen.
- „VORRECHNEN:5“: Ich möchte gerne, dass meine Abgabe besprochen wird.

Falls Sie als Gruppe abgeben, muss jedes Gruppenmitglied einzeln die Bereitschaft zum Vorrechnen erklären (VORRECHNEN1:N ist der Wert für den Studierenden, der die Aufgabe in die Übungsplattform hochgeladen hat, und VORRECHNEN2:M der Wert für den anderen Studierenden).

Vergessen Sie nicht, eventuell verwendete Quellen wie ChatGPT (oder auch die Lösung eines anderen Studierenden) anzugeben (mit der in Kapitel 0, Folien 22 bis 27 beschriebenen Codierung). Ohne Quellenangabe werden „zu ähnliche Lösungen“ als Plagiat behandelt. Mit Quellenangabe werden sie normal korrigiert. In der Klausur müssen Sie aber ähnliche Aufgaben ohne Hilfe lösen! Sie sind erwachsene Menschen und können sich ja denken, dass Sie nur durch eigene Beschäftigung mit den Aufgaben lernen.

Wir würden uns freuen, wenn Sie die für die jeweilige Aufgabe verwendete Zeit in Minuten in der Form „ZEIT:N“ angeben würden (bei Gruppenarbeit „ZEIT1:N“ und „ZEIT2:M“). Diese Angabe ist freiwillig.

Hinweis: SQL-Anfragen mit Syntaxfehlern werden automatisch mit 0 Punkten bewertet! Testen Sie also Ihre Anfragen (z.B. im Adminer).

Datenbank

Die Aufgaben dieses Übungsblattes stammen zum großen Teil aus der Zwischenklausur zu „Datenbanken I“ im WS 2003/2004. Die ersten drei Aufgaben sind daher eher unterhalb des Niveaus, das bei unserer Klausur vorausgesetzt wird (bei uns gibt es nur eine Klausur am Ende). Die Original-Klausur finden Sie auf der Webseite der Vorlesung:

[<https://users.informatik.uni-halle.de/~brass/db24/exams/exam7.pdf>]

Es gibt in dieser Klausur noch eine Reihe interessanter Ankreuzaufgaben.

Unten auf der Webseite

[<https://users.informatik.uni-halle.de/~brass/db24/pruefung.html>]

gibt es auch noch weitere alte Klausuren.

Bedenken Sie, dass Sie sich in der Klausur immer in ein neues DB-Schema einarbeiten müssen. Es empfiehlt sich, dieses Übungsblatt am Stück zu bearbeiten, und auch die Gesamtzeit zu messen. In der Klausur sind ca. 10 min für jede SQL-Anfrage eingeplant (mit etwas Luft am Schluss, aber nicht sehr viel).

In der Klausur gilt die Regel „0 Punkte bei Syntaxfehlern“ nicht, aber in dieser Hausaufgabe. Sie könnten also überlegen, wann es günstiger wäre, zur nächsten Aufgabe überzugehen, und das beim Probedurchlauf auch tun. Anschließend sollten Sie aber die Syntaxfehler noch entfernen, bevor Sie die Anfragen abgeben. Die meisten Klausur-Abgaben mit Syntaxfehlern enthalten nicht nur einen einzigen Syntaxfehler, sondern mehrere Fehler, so dass es am Ende oft doch sehr wenig Punkte gibt. Aber das Zeitmanagement ist auch eine Kompetenz, die bei der Klausur getestet wird. Sie sollten sich nicht in einer einzelnen Aufgabe verrennen, sondern spätestens nach 15 Minuten zur nächsten Aufgabe übergehen. ILIAS zeigt die verbleibenden Minuten an, und Sie können sich z.B. auf einem Schmierzettel notieren, wann Sie mit einer Aufgabe begonnen haben. Smartwatches sind bei der Klausur verboten (alles, was mehr kann, als Datum und Uhrzeit anzuzeigen).

Lesen Sie jetzt nur weiter, wenn Sie gut 60 min Zeit haben. Mehr als 90 min hätten Sie in der Klausur auf keinen Fall für die folgenden 6 Aufgaben (in der Klausur gibt es nur 5 SQL-Anfragen, dafür aber noch weitere Arten von Aufgaben).

Der Dozent möchte sein Skript in ein Online-Lernsystem umwandeln. Dies soll u.a. erlauben, dass Studierende selbst die Teile des Skripts auswählen, die sie lernen wollen.

- Die Tabelle „ABSCHNITTE“ enthält einen Eintrag für jeden Abschnitt des Skripts (Lerneinheit). Jeder Abschnitt hat eine eindeutige Nummer „AB“ (Primärschlüssel). Die Überschrift „TITEL“ des Abschnitts ist Sekundärschlüssel.

ABSCHNITTE			
AB	TITEL	URL	SEITEN
101	Einleitung	c1_intro/s1.pdf	5
201	Relationales Modell: Basis	c2_relmo/s1.pdf	3
202	Integritätsbedingungen	c2_relmo/s2.pdf	7
203	Relationale Algebra	c2_relmo/s3.pdf	12
301	Einfache Anfragen in SQL	c3_sql_i/s1.pdf	6

- Die zweite Tabelle enthält Datenbank-Begriffe, die im Skript eingeführt oder verwendet werden. Begriffe werden über eine eindeutige Nummer „BEG“ identifiziert, „NAME“ ist Sekundärschlüssel. Begriffe sind in eine Hierarchie (wie ein Verzeichnisbaum) eingeordnet. Das Attribut OBERBEG ist ein Fremdschlüssel, der auf die Tabelle selbst verweist. Es ist das einzige Attribut der Datenbank, das null sein kann.

BEGRIFFE		
BEG	NAME	OBERBEG ^o
1000	Datenmodell	(null)
1001	Relationales Modell	1000
1002	Schlüssel	1001
1003	Fremdschlüssel	1001
1004	Datei	(null)
1005	Anfragesprachen	(null)
1006	Relationale Algebra	1005
1007	Projektion	1006

- Die dritte Tabelle setzt Abschnitte und Begriffe in Beziehung (wie ein Index in einem Buch). „AB“ ist ein Fremdschlüssel, der auf „ABSCHNITTE“ verweist, und „BEG“ verweist auf „BEGRIFFE“. In der Spalte „ART“ steht der Wert „DEF“ für ein definierendes Vorkommen des Begriffs, der Wert „BSP“ für ein Beispiel, und der Wert „VORAUS“ für ein Vorkommen des Begriffes, bei dem seine Kenntnis bereits vorausgesetzt wird.

VORKOMMEN		
AB	BEG	ART
101	1004	VORAUS
101	1000	DEF
101	1005	DEF
201	1000	VORAUS
202	1002	DEF
202	1002	BSP
202	1003	DEF
202	1007	VORAUS
203	1007	BSP
203	1007	DEF

Probieren Sie die SQL-Anfragen wie immer mit PostgreSQL aus, z.B. über die Adminer-Webschnittstelle (diese Testmöglichkeit besteht auch in der Klausur):

[https://dbs.informatik.uni-halle.de/edb-probeklausur-ws1920?pgsql=db&username=student_gast&db=postgres&ns=exam7]

Das Passwort steht Sie in StudIP-Eintrag der Vorlesung, Reiter „Informationen“.

Schema in Kurznotation:

- ABSCHNITTE(AB, TITEL, URL, SEITEN)
- BEGRIFFE(BEG, NAME, OBERBEG^o → BEGRIFFE)
- VORKOMMEN(AB → ABSCHNITTE, BEG → BEGRIFFE, ART)

Aufgabe 1 (4 Punkte)

Geben Sie alle Begriffe aus (jeweils Nummer und Name), die irgendwo im Skript vorausgesetzt werden (Art des Vorkommens „VORAUS“), aber nirgendwo definiert werden (Art „DEF“).

Das erwartete Ergebnis ist:

BEG	NAME
1004	Datei

Die Groß-/Kleinschreibung der Ausgabespalten ist nicht vorgeschrieben. Die Klausur stammt noch aus einer Zeit, in der Oracle in der Vorlesung benutzt wurde. Dort ist Großschreibung normal.

Aufgabe 2 (4 Punkte)

In welchen Titeln von Abschnitten kommen sowohl die Teilzeichenkette „SQL“ als auch die Teilzeichenkette „Anfragen“ vor? Es ist nicht bekannt, ob zuerst „SQL“ und dann „Anfragen“ kommt, oder umgekehrt. Geben Sie jeweils die Abschnittsnummer, den Titel und die URL aus.

Das erwartete Ergebnis ist:

AB	TITEL	URL
301	Einfache Anfragen in SQL	c3_sql_i/s1.pdf

Aufgabe 3 (4 Punkte)

Geben Sie Nummer und Titel aller Abschnitte aus, die mindestens einen Begriff definieren, der in der Begriffshierarchie ganz oben steht, für den also „OBERBEG“ einen Nullwert enthält. Im Beispiel würde Abschnitt 101 die Begriffe 1000 (Datenmodell) und 1005 (Anfragesprachen) definieren. Denken Sie wie bei allen Anfragen darüber nach, ob eine explizite Duplikateliminierung notwendig ist.

Das erwartete Ergebnis ist:

AB	TITEL
101	Einleitung

Aufgabe 4 (4 Punkte)

Es ist schlecht, wenn Begriffe definiert werden, nachdem Sie bereits vorausgesetzt wurden. Schreiben Sie eine Anfrage, die Abschnitte A und B findet, so dass B eine größere Nummer (AB) als A hat, aber in B ein Begriff definiert wird, der in A vorausgesetzt wird. Geben Sie den Namen des Begriffs und die Nummern der beiden Abschnitte A und B aus. Die Ausgabe-Spalte mit der Abschnittsnummer von A soll „VORAUSGESETZT_IN“ heißen, und die mit der Abschnittsnummer von B soll „DEFINIERT_IN“ heißen.

Das erwartete Ergebnis ist:

NAME	VORAUSGESETZT_IN	DEFINIERT_IN
Projektion	202	203

Bei dieser Aufgabe sind die Spaltennamen vorgeschrieben, inklusive der Großschreibung. In der echten Klausur ist üblicherweise der Default, dass die Spaltennamen genau so heißen müssen, wie im Beispiel gezeigt (inklusive Groß-/Kleinschreibung). Da neuere Klausuren mit PostgreSQL entwickelt sind, werden die meisten Spaltennamen dort aber klein geschrieben sein, so dass oft keine Umbenennung nötig ist.

Aufgabe 5 (4 Punkte)

Die Abschnitte in der Tabelle AB sind durch Aufteilung von größeren Kapiteln entstanden. Die Kapitelnummer ist die Abschnittsnummer AB ohne die letzten zwei Dezimalstellen. Man kann Sie berechnen, indem man die Abschnitts-Nummer AB durch 100 teilt (mit Integer-Division, d.h. Abrundung — glücklicherweise ist „AB“ als INTEGER deklariert).

Geben Sie für jedes Kapitel Folgendes aus:

- die Kapitel-Nummer,
- die Anzahl der Abschnitte des Kapitels und
- die Gesamtanzahl der Seiten (d.h. Summe der einzelnen Seitenzahlen)

Sortieren Sie die Ausgabe nach der Kapitelnummer.

Bitte wählen Sie die Spaltenüberschriften wie in der Beispiel-Ausgabe gezeigt (auch mit der vorgegebenen Groß-/Kleinschreibung).

Wenn Sie wollen, können Sie noch freiwillig das Verzeichnis des Kapitels aus der URL ausgeben. In der Klausur würde das wohl einen Bonuspunkt geben, bei der Hausaufgabe zur administrativen Vereinfachung leider nicht. Sie können voraussetzen, dass jede URL einen „/“ enthält, und dass der Präfix der URL bis zum Schrägstrich für alle Abschnitte eines Kapitels gleich ist. Zur Erinnerung:

- `POSITION(s_1 IN s_2)`: Position der Zeichenkette s_1 als Teilzeichenkette von s_2 .
- `SUBSTRING(s FROM n FOR m)`: Teilzeichenkette von s ab Position n mit einer Länge von m Zeichen.

Die erwartete Antwort (mit der optionalen Spalte „Verzeichnis“) ist:

Kapitel	Abschnitte	Seiten	Verzeichnis
1	1	5	c1_intro
2	3	22	c2_relmo
3	1	6	c3_sql_i

Aufgabe 6 (4 Punkte)

Geben Sie zu jedem vorkommenden Oberbegriff die Anzahl seiner Unterbegriffe an, allerdings nur, wenn er mindestens zwei Unterbegriffe hat. Drucken Sie die Nummer und den Namen des Oberbegriffs, die Anzahl seiner Unterbegriffe und den alphabetisch ersten Unterbegriff (als ein Beispiel eines Unterbegriffs). Nennen Sie die Ausgabespalten wie im Beispiel gezeigt.

Die erwartete Antwort ist:

Nr	Name	Unterbegriffe	Beispiel
1001	Relationales Modell	2	Fremdschlüssel