Prof. Dr. Stefan Brass PD Dr. Alexander Hinneburg Institut für Informatik MLU Halle-Wittenberg

# Einführung in Datenbanken — Übungsblatt 13 (Outer Join, ER-Diagramm) —

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

#### Einsendeschluss ist Montag, der 27.01.2025, 18<sup>00</sup>.

Hausaufgaben können einzeln oder in Zweier-Gruppen bearbeitet werden. Sie können die Gruppe für jede Aufgabe neu wählen.

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

- "VORRECHNEN: O": 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 SQL-Aufgaben dieses Übungsblattes beziehen sich wieder auf die "EMP-DEPT-Datenbank" mit Informationen über Angestellte und Abteilungen einer Firma (eine bekannte Beispiel-Datenbank von Oracle). Das Schema "empdept\_public" besteht aus folgenden Tabellen:

- Abteilungen: dept(<u>deptno</u>, dname, loc)
- Angestellte: emp(<u>empno</u>, ename, job, mgr<sup>°</sup>→emp, hiredate, sal, comm<sup>°</sup>, deptno<sup>°</sup>→dept)
- Gehaltsstufen: salgrade(grade, losal, hisal)

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?pgsql=db& username=student\_gast&db=postgres&ns=empdept\_public]

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

### Aufgabe 1 (4 Punkte)

Schreiben Sie folgende Anfrage in SQL: Geben Sie für jede Abteilung die Anzahl Angestellte und die Summe der Gehälter aus (und natürlich Nummer und Name der Abteilung). Dabei sollen auch leere Abteilungen (ohne Angestellte) mit ausgegeben werden (jeweils mit dem Wert 0 in beiden Spalten). <u>Sortieren</u> Sie das Ergebnis nach der Abteilungsnummer und nennen Sie die beiden letzten Spalten "Angestellte" und "Gehaltssumme".

Das erwartete Ergebnis ist:

deptno	dname	Angestellte	Gehaltssumme
10	ACCOUNTING	3	8750
20	RESEARCH	4	10075
30	SALES	6	9400
40	OPERATIONS	0	0

## Aufgabe 2 (4 Punkte)

Geben Sie für jeden Manager (Angestellten mit Job "MANAGER") die Anszahl der direkt untergeordneten Sachbearbeiter (Angestelle mit Job "CLERK") aus.

Die Anfrage soll so gestaltet sein, dass sie auch dann funktionieren würde, wenn ein eventuell neu eingestellter Manager noch gar keine Untergebenen hätte (das kommt in den Beispieldaten nicht vor).

Das erwartete Ergebnis ist:

empno	ename	Num_Clerks		
7566	JONES	0		
7698	BLAKE	1		
7782	CLARK	1		

## Aufgabe 3 (4 Punkte)

Geben Sie für jede Abteilung die Anzahl verschiedener Gehaltsstufen sowie die minimale und die maximale Gehaltsstufe aus. Dabei sollen auch leere Abteilungen in der Ausgabe erscheinen (mit der Zahl 0 für die Anzahl und einem Nullwert für Minimum und Maximum). <u>Sortieren</u> Sie das Ergebnis nach der Abteilungsnummer.

Das erwartete Ergebnis ist:

deptno	dname	count	min	max
10	ACCOUNTING	3	2	5
20	RESEARCH	2	1	4
30	SALES	4	1	4
40	OPERATIONS	0		

Hinweis: Beim Adminer sind die hier leer dargestellten Tabelleneinträge mit dem Wort "*NULL*" gefüllt. Beide Repräsentationen des Nullwerts sind möglich.

## Aufgabe 4 (8 Punkte)

Zeichnen Sie ein ER-Diagramm in der Barker-Notation (wie in der Vorlesung) für eine Datenbank mit (Youtube-)Videos von Feuerwerksartikeln:

- Zu jedem Video ist die URI (Webadresse) zu speichern, außerdem ein Titel, die Länge in Sekunden und eine Bewertung der Videoqualität. Die URI identifiziert die Videos eindeutig. Die Bewertung fehlt manchmal.
- Zu jedem Feuerwerksartikel sind Name und Hersteller zu speichern, sowie die Brenndauer in Sekunden (der offizielle, vom Hersteller angegebene Wert). Name und Hersteller zusammen identifizieren den Artikel eindeutig. Der Name allein reicht dagegen nicht unbedingt aus (es kann z.B. einen "Goldvulkan" von verschiedenen Herstellern geben).
- Ein Video zeigt immer mindestens einen Artikel, eventuell aber auch mehrere Artikel (z.B. bei einem Vergleich mehrerer ähnlicher Artikel).
- Zu Artikeln muss es nicht unbedingt ein Video geben, das den Artikel zeigt. Es kann aber natürlich auch mehrere Videos geben.
- Jedes Video stammt aus genau einem Kanal ("Youtuber", Video-Filmer bzw. Video-Anbieter). Zu jedem Kanal ist der Name des Kanals zu speichern, der den Kanal auch eindeutig identifiziert. Außerdem soll es möglich sein, eine optionale Bemerkung (beliebiger Text) zu speichern.
- Ein Kanal kann beliebig viele Videos enthalten, aber mindestens eins (der Kanal wird in der Datenbank erst angelegt, wenn auch ein Video des Kanals gespeichert wird).

Sie müssen das ER-Diagramm im PDF-Format abgeben. Mit handgezeichneten (und eingescannten) Diagrammen können Sie maximal 7 Punkte erreichen (von 8). Es können Punkte auch für schlechtes Aussehen abgezogen werden.

Das Relationship muss wie in der Vorlesung gezeigt beschriftet werden, die Position der beiden Namen kann aber anders sein. Wenn es gar nicht anders geht, reicht auch ein Name in der Mitte.

#### Hinweise zum Oracle SQL Developer Data Modeler

Sie können, müssen aber nicht, den "Oracle SQL Developer Data Modeler" zum Zeichnen verwenden:

[https://www.oracle.com/database/sqldeveloper/technologies/sql-data-modeler/]

[https://www.oracle.com/de/database/sqldeveloper/technologies/sql-data-modeler/]

Es ist ein professionelles Entwurfs-Werkzeug, braucht aber etwas Zeit zur Einarbeitung. Das Programm ist in Java geschrieben und sollte unter allen gängigen Betriebssystemen laufen (auch unter Linux). Das Programm ist kostenlos.

Das ER-Diagramm ist das "Logical Model" in diesem Werkzeug. Wählen Sie den entsprechenden Tab.

Oben in der Palette wählen Sie das Symbol "Kasten mit Zahnrad" ("New Entity"). Dann können Sie auf der Zeichenfläche den Kasten für einen Entity-Typ groß ziehen (in der oberen linken Ecke drücken Sie die Maustaste, und halten sie gedrückt, während sie den Mauszeiger an die untere rechte Ecke bewegen). Es erscheint die "Entity Properties" Dialogbox. Das Entity heißt zunächst "Entity\_1", aber Sie können den Namen ändern. Im Menu der Dialogbox links können Sie unter "Attributes" Attribute anlegen (mit dem grünen Pluszeichen). Auch die heißen zunächst einfach "Attribute\_1" u.s.w. und müssen umbenannt werden. Sie brauchen für diese Aufgabe keinen Datentyp zu wählen (er erscheint ja nicht im Diagramm), aber wenn Sie das tun wollen, sollten Sie "Logical" anklicken, dann haben Sie eine Auswahlbox mit SQL-Datentypen unter "Source Type". Für Schlüsselattribute klicken Sie "Primary UID" an, für Attribute, die nicht Null sind, "Mandatory".

Für Relationships gibt es in der Zeichenpalette oben u.a. den Doppelpfeil mit Sternen auf beiden Seiten für Viele-zu-viele-Beziehungen (n:m) und den einfachen Pfeil mit 1 und Sternchen für Eins-zu-viele-Beziehungen (1:n). Klicken Sie auf das entsprechende Symbol und anschließend auf die beiden beteiligten Entity-Typen. Es öffnet sich die "Relation Properties" Dialogbox. Darin können Sie dem Relationship als Ganzes einen Namen geben (brauchen Sie hier nicht), und auch Namen in den beiden Richtungen ("Name on Source" und "Name on Target"). Außerdem können Sie mit "Source Optional" und "Target Optional" auswählen, ob die entsprechende Hälfte der Line gestrichelt sein soll. Damit die beiden Namen des Relationships auch im Diagramm angezeigt werden, müssen Sie mit der rechten Maustaste auf den Diagramm-Hintergrund klicken und in dem Menu unter "Show" den Unterpunkt "Labels" auswählen.

Wenn Sie die Kästen nachträglich verschieben wollen, müssen Sie oben in der Palette den Pfeil auswählen ("Select"). Sonst erzeugen Sie immer weiter Entity-Typen oder Relationships, wenn Sie im Diagramm klicken.

Sie speichern das Diagramm unter "File"/"Save AS … ". Durch das Speichern bekommt Ihr Diagramm einen Namen. Unter "File"/"Print Diagram" können Sie verschiedene Ausgabeformate wählen, u.a. PDF.