

# Einführung in Datenbanken

## — Übungsblatt 9 (Aggregations-Anfragen) —

### Organisatorische Hinweise:

- **Einsendeschluss ist Montag, der 18.01.2021, 18<sup>00</sup>.**
- Ihre Lösungen laden Sie bitte in die Übungsplattform in StudIP hoch ([StudIP-Eintrag der Vorlesung], Reiter „Übungsplattform“, dann auf „Weiter zu ... Übungsplattform“).
- Bitte geben Sie die Hausaufgaben als einfache Textdatei mit der Endung `.sql` ab. Alle Zusatzangaben müssen mit `--` als Kommentar gekennzeichnet werden.
- Hausaufgaben müssen einzeln bearbeitet werden. „Zu ähnliche Lösungen“ führen automatisch zu 0 Punkten für alle Beteiligten. Das gilt auch dann, wenn Sie nicht direkt abgeschrieben haben, sondern nur zufällig die gleiche Quelle benutzt haben.
- Probieren Sie Ihre Anfragen mit dem Adminer aus:

```
[https://dbs.informatik.uni-halle.de/edb?pgsql=db&
    username=student.gast&db=postgres&ns=]
```

Die Zugangsdaten unserer Installation stehen in StudIP, Reiter „Adminer“.

- Anfragen, die Syntaxfehler enthalten, werden höchstwahrscheinlich mit 0 Punkten bewertet. Nutzen Sie notfalls rechtzeitig das Forum in StudIP, um Hilfe zu bekommen.
- Verwenden Sie in Ihren Anfragen nur Informationen, die in der Aufgabenstellung erwähnt sind. Ihre Anfrage muss mit beliebigen DB-Zuständen funktionieren, nicht nur dem Beispiel-Zustand.
- Schreiben Sie keinen Schema-Namen in die Anfrage. Bei der automatischen Vorkorrektur heißt das Schema eventuell anders.
- Es können auch Punkte für schlechten Stil abgezogen werden. Z.B. sollten Sie sinnvolle Namen für die Tupelvariablen wählen. Bemühen Sie sich auch um gute Formatierung. Vermeiden Sie unnötige Verbunde oder andere unnötigen Komplikationen.
- Ihre Anfragen sollen keine Duplikate liefern (in keinem Zustand, der die Integritätsbedingungen erfüllt). Für ein überflüssiges `DISTINCT` wird allerdings auch ein Punkt abgezogen.

## Hausaufgabe 9 (15 Punkte)

Wählen Sie im Adminer das Schema „komponist\_public“. Diese Sammlung von Daten klassischer Musik-CDs besteht aus den folgenden Tabellen:

- KOMPONIST(KNR, NAME, VORNAME, GEBOREN, GESTORBEN<sup>°</sup>)
- STUECK(SNR, KNR<sup>°</sup>→KOMPONIST, TITEL, TONART<sup>°</sup>, OPUS<sup>°</sup>)
- CD(CDNR, NAME, HERSTELLER<sup>°</sup>, ANZ\_CDS, GESAMTSPIELZEIT<sup>°</sup>)
- AUFNAHME(CDNR→CD, SNR→STUECK, ORCHESTER<sup>°</sup>, LEITUNG<sup>°</sup>)
- SOLIST((CDNR, SNR)→AUFNAHME, NAME, INSTRUMENT<sup>°</sup>)

Sie bekommen jeweils 5 Punkte für jede der folgenden drei Anfragen.

- a) Geben Sie Name und Vorname von allen Komponisten aus, von denen es mindestens 5 Stücke in der Datenbank gibt. Geben Sie die Anzahl der Stücke mit aus, nennen Sie die Spalte `anz_stuecke` und sortieren Sie die Ausgabe danach absteigend (größte Anzahl zuerst).

Stücke mit unbekanntem Komponisten (d.h. mit einem Nullwert in der Spalte `knr`) können Sie natürlich ignorieren. Sie können davon ausgehen, dass Name und Vorname zusammen einen Alternativschlüssel der Komponisten-Tabelle bilden.

Die erwartete Antwort ist:

name	vorname	anz_stuecke
Händel	Georg Friedrich	25
Telemann	Georg Philipp	15
Prokofiev	Serge	13
Bach	Johann Sebastian	10
Monteverdi	Claudio	10
Mozart	Wolfgang Amadeus	9
Biber	Heinrich Ignaz Franz	9
Vivaldi	Antonio	8
Mozart	Leopold	6
Beethoven	Ludwig van	5

10 Datensätze

- b) Was ist die minimale, maximale und durchschnittliche Anzahl Stücke pro CD?

Sie müssen also die Anzahl Stücke pro `cdnr` bestimmen, und jeweils durch `anz_cds` teilen (z.B. haben Doppel-CDs nur eine `cdnr`), und anschließend den Durchschnitt über diesen Anzahlen bilden.

Runden Sie die Zahlen auf eine Nachkommestelle (geht mit der Funktion `ROUND` mit der Anzahl Nachkommastellen als zweitem Argument). Nennen Sie die Spalten „Min“, „Max“ und „Durchschnitt“ — auch mit dieser Groß-/Kleinschreibung.

Die erwartete Antwort ist:

Min	Max	Durchschnitt
0.5	23.0	3.8

1 Datensatz

- c) Geben Sie für jeden Komponisten, von dem mindestens drei Stücke mit bekannter Tonart (nicht Null) in der Datenbank stehen, aus, wie viele Stücke es sind, und was der Prozentsatz in Dur-Tonarten und was der Prozentsatz in Moll-Tonarten ist. Den Komponisten identifizieren Sie durch Name und Vorname und ordnen die Ausgabe nach dem Namen, bei gleichem Namen nach dem Vornamen. Dur-Tonarten sind solche, die in „-dur“ enden, und Moll-Tonarten solche, die in „-moll“ enden.

Bei dieser Aufgabe sollen nur Stücke mit definierter Tonart betrachtet werden (auch in den Prozent-Berechnungen). Geben Sie die Prozente gerundet ohne Nachkommastellen aus (nutzen Sie `ROUND`, und zwar die Variante mit nur einem Parameter).

Achten Sie darauf, dass `COUNT` bei PostgreSQL einen Integer-Typ liefert (`bigint`), und PostgreSQL die Integer-Division verwendet, wenn beide Argumente einen Integer-Typ haben (es wird dann also abgerundet). Sie könnten z.B. die Anzahl Dur/Moll-Stücke mit `100.0` multiplizieren, bevor Sie durch die Gesamtanzahl dividieren.

Die erwartete Antwort ist:

name	vorname	anz	prozent_dur	prozent_moll
Bach	Johann Sebastian	10	70	30
Händel	Georg Friedrich	21	62	38
Mozart	Leopold	6	100	0
Mozart	Wolfgang Amadeus	5	80	20
Prokofiev	Serge	4	75	25
Schubert	Franz	3	67	33
Telemann	Georg Philipp	15	87	13
Vivaldi	Antonio	8	50	50
Wolf-Ferrari	Ermanno	4	100	0

9 Datensätze