

# Einführung in Datenbanken

## — Übungsblatt 10 (Aggregationen, Rel. Algebra) —

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

**Einsendeschluss ist Montag, der 24.01.2022, 18<sup>00</sup>.**

### Hausaufgabe 10 (20 Punkte)

Die Aufgaben dieses Übungsblattes beziehen sich auf die Datenbank mit Informationen über Nahrungsergänzungsmittel (insbesondere Multivitamin-Tabletten):

- Stoff\_Kategorie(Kat, Bezeichnung, Sort\_Nr)
- Stoff(Vit, Einheit, Tagesdosis<sup>°</sup>, Kat→Stoff\_Kategorie)
- Praeparat(Pid, Name, Hersteller, PZN<sup>°</sup>, Anz<sup>°</sup>, Einheit<sup>°</sup>, Tagesdosis<sup>°</sup>, Gewicht<sup>°</sup>, Preis<sup>°</sup>, glutenfrei<sup>°</sup>, lactosefrei<sup>°</sup>)
- Inhalt(Pid → Praeparat, Vit → Stoff, Menge, Prozent<sup>°</sup>, Anmerkung<sup>°</sup>)
- Zutat(Pid → Praeparat, Seq, Name, Anmerkung<sup>°</sup>)

Im Adminer ist es das Schema „vit\_public“:

[[https://dbs.informatik.uni-halle.de/edb?pgsql=db&username=student\\_gast&db=postgres&ns=vit\\_public](https://dbs.informatik.uni-halle.de/edb?pgsql=db&username=student_gast&db=postgres&ns=vit_public)]

Es gibt auch ein SQL-Skript, um die Tabellen in einer eigenen Datenbank anzulegen:

[<https://users.informatik.uni-halle.de/~brass/db21/sql/vit.sql>]

Probieren Sie die SQL-Anfragen wie immer mit PostgreSQL aus. Anfragen, die Syntaxfehler enthalten, werden höchstwahrscheinlich mit 0 Punkten bewertet.

- a) **(5 Punkte)** Welche Nahrungsergänzungsmittel enthalten maximal 10 Inhaltsstoffe (sind also vermutlich kein Multivitamin-Präparat)? Geben Sie Hersteller, Name und die Anzahl der Inhaltsstoffe aus.

„Inhaltsstoffe“ bezieht sich auf die Wirkstoffe (Vitamine, Mineralstoffe, u.s.w.), nicht auf die Zutaten.

Das Ergebnis sollte so aussehen:

Hersteller	Name	Inhaltsstoffe
Verla-Pharm Arzneimittel	Magnesium Verla 300	1

Die Ausgabespalten sollen so heißen wie hier gezeigt (das schließt immer auch die Groß-/Kleinschreibung ein).

Leider ist die Datensammlung noch sehr klein. Vielleicht werden bei zukünftigen Updates Präparate z.B. mit zwei oder drei Inhaltsstoffen hinzugefügt. Ihre Anfrage müsste dann natürlich auch funktionieren.

Sie dürfen die Ausgabe sortieren (z.B. nach Hersteller und Name), müssen das aber nicht. Es gibt nicht mehr oder weniger Punkte, wenn Sie eine (vernünftige) `ORDER BY`-Klausel schreiben. Der Vergleich mit dem erwarteten Ergebnis ist ja offensichtlich auch ohne Sortierung möglich.

Auch in der Klausur gilt: Falls die Aufgabenstellung keine Sortierung vorschreibt, gibt es keinen Punktabzug, wenn Sie trotzdem eine sinnvolle Sortierung durchführen (das könnte z.B. den Vergleich mit der erwarteten Antwort erleichtern). Falls die `ORDER BY`-Klausel aber eindeutig mangelhaft ist, ist ein Punktabzug nicht ausgeschlossen. Ein Punktabzug wäre auch möglich, wenn die Anfrage beweisbar in allen Datenbank-Zuständen nur ein einziges Ausgabebetupel liefert. Dann wäre die Sortierung ja nicht sinnvoll. Man kann sich auf den Standpunkt stellen, dass alle Anfragen ein sinnvoll sortiertes Ergebnis liefern sollten. Das würde auch plötzliche Änderungen der Tupel-Reihenfolge bei einem Update des DBMS ausschließen (vielleicht hat der alte Optimierer `GROUP BY` mit Sortierung implementiert, der neue mit einer Hashtabelle zum Finden der Gruppen-Werte). In dieser Vorlesung wird eine nicht explizit verlangte Sortierung zukünftig also als Frage des persönlichen Stils angesehen. (In der Übung wurde gefragt, ob es Punktabzug für nicht verlangte Sortierungen gibt. Das ist hiermit beantwortet.)

- b) **(5 Punkte)** Welches Präparat bzw. welche Präparate haben die maximale Anzahl Inhaltsstoffe?

Wie bei a) sind mit „Inhaltsstoffen“ die Wirkstoffe gemeint.

Das Ergebnis sollte so aussehen:

Hersteller	Name	Inhaltsstoffe
gesundleben	A-Z Vital	26

Die Ausgabespalten sollen so heißen wie hier gezeigt.

Wie bei a) ist die Sortierung optional.

- c) **(5 Punkte)** Geben Sie für jedes Präparat die Summe der Mengen der Inhaltsstoffe aus (also wie viel Wirkstoff eine Tablette etc. enthält). Als kleine Komplikation müssen Sie die Einheiten **mg** und **ug** (für Mikrogramm  $\mu\text{g}$ ) beachten, die in der Tabelle **Stoff** eingetragen sind. Sie können sich darauf verlassen, dass es nur diese beiden Einheiten gibt. Das Ergebnis soll in **mg** mit einer Nachkommastelle sein.

Geben Sie die **pid** und den Namen des Präparates aus sowie die Gesamtmenge an Wirkstoffen. Außerdem drucken Sie bitte das Gewicht einer Einheit (Tablette, Beutel mit Getränkepulver, etc.) in Gramm (auch mit einer Nachkommastelle). Sie müssen dazu den Wert in der Spalte **Gewicht** durch den Wert in der Spalte **Anz** teilen.

Das Ergebnis sollte so aussehen:

pid	name	gesamtmenge	gewicht
5	Kardiodrink	1430.3	6.0
6	Mineraldrink	1066.6	6.0
1	A-Z Vital	635.1	1.6
2	Centrum	548.5	1.2
4	A-Z Complete	511.6	1.5
3	A-Z Komplett	326.3	1.2
7	Magnesium Verla 300	300.0	NULL

Die Ausgabespalte soll so heißen wie hier gezeigt.

Sortieren Sie das Ergebnis nach der Gesamtmenge (größte Menge zuerst).

- d) **(5 Punkte)** Die folgende Anfrage ist in relationaler Algebra zu schreiben. Sie bezieht sich auf die Vitamin-Datenbank wie oben (aktuell noch ohne die Tabelle **Zutat**). Unter folgender Webadresse können Sie Relax mit dieser Datenbank öffnen:

[<http://dbis-uibk.github.io/relax/calc/gist/aeb39915904003f10caa6e53bbc5f19a>]

Falls Sie die Daten im für RelaX nötigen Format sehen wollen (das ist aber nur interessant, wenn Sie eigene Datenbanken als „Gist“ anlegen wollen):

[<https://gist.github.com/aeb39915904003f10caa6e53bbc5f19a>]

Die Daten sind in einer Art vereinfachtem CSV-Format: Wenn ein Spaltenwert nicht nur ein einzelnes Wort oder eine einzelne Zahl ist, muss er in "... " geschrieben werden.

Schreiben Sie folgende Anfrage in relationaler Algebra: Welche Präparate enthalten mindestens 10 mg Eisen? Drucken Sie den Namen des Präparats und die im Präparat enthaltene Menge an Eisen.

Sie brauchen nicht zu kontrollieren, dass die Einheit für Eisen „mg“ ist.

Das Ergebnis sollte so aussehen:

Praeparat.Name	Inhalt.Menge
'A-Z Komplett'	14

Der Präfix der Spaltennamen mit dem Relationenname darf auch anders sein oder fehlen (RelaX druckt einen solchen Präfix).

Tipps:

- Es ist hier empfohlen, einen natürlichen Verbund zu nutzen, der zwei Tabellen über den gleich benannten Spalten verknüpft. Die Syntax (hier bezogen auf die Beispiel-Datenbank der Vorlesung) ist:

```
STUDENTEN join BEWERTUNGEN.
```

In der Vorlesung würde der natürliche Verbund so geschrieben:

```
STUDENTEN ⋈ BEWERTUNGEN
```

- Ein Beispiel für die Syntax der Selektion ist:

```
sigma SID = 101 (STUDENTEN).
```

Leider ist der Index in RelaX nicht speziell markiert. In der Vorlesung würde es  $\sigma_{SID=101}(STUDENTEN)$  geschrieben.

- Ein Beispiel einer Projektion ist:

```
pi SID, EMAIL (STUDENTEN).
```

Dies entspricht:  $\pi_{SID,EMAIL}(STUDENTEN)$ .

- Beachten Sie, dass die Selektion stärker bindet als der Join. Sie müssen ggf. Klammern setzen.
- Denken Sie daran, dass die Groß-/Kleinschreibung der Spalten- und Tabellennamen bei RelaX wichtig ist.