

Einführung in Datenbanken

— Übungsblatt 7 (Selbstverbund, Unteranfragen) —

Organisatorische Hinweise:

- **Einsendeschluss ist Montag, der 04.01.2021, 18⁰⁰.**
- 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 7 (15 Punkte)

Wählen Sie die schon bekannte Datenbanken der US-Präsidenten (bis George Busch senior). Im Adminer ist es das Schema `president_public`. Die Datenbank besteht aus den folgenden Tabellen:

- `state(state_name, admin_entered, year_entered)`
- `president(pres_name, birth_year, years_serv, death_age, party, state_born→state)`
- `pres_hobby(pres_name→president, hobby)`
- `administration(admin_nr, pres_name→president, year_inaugurated)`
- `admin_pr_vp((admin_nr, pres_name)→administration, vice_pres_name)`
- `pres_marriage(pres_name→president, spouse_name, pr_age, sp_age, nr_children, mar_year)`
- `election(election_year, candidate, votes, winner_loser_indic)`

Sie bekommen jeweils 5 Punkte für jede der folgenden drei Anfragen, davon einen für den Kommentar zu `DISTINCT`.

- a) Wer hat eine Wahl verloren (Wert 'L' für `winner_loser_indic` in `election`) und später noch eine Wahl gewonnen ('W')?

Falls Sie `DISTINCT` verwenden, schreiben Sie als Kommentar eine kurze Begründung dazu, warum die Anfrage sonst Duplikate liefern könnte. Falls Sie kein `DISTINCT` verwenden, schreiben Sie als Kommentar eine Erklärung, warum die Anfrage niemals Duplikate liefern kann.

Die erwartete Antwort ist:

candidate

Adams J

1 Datensatz

(Die Daten zu Wahlen in der Datenbank sind leider unvollständig, wundern Sie sich also nicht über die geringe Anzahl von Ergebnissen.)

- b) Welche Präsidenten haben die Hobbies „Riding“ und „Golf“? Gesucht sind nur Präsidenten die beide Hobbies haben, nicht etwa Präsidenten, die das eine oder das andere Hobby haben. Geben Sie das Geburtsjahr des Präsidenten mit aus und sortieren Sie die Ausgabe nach diesem Geburtsjahr.

Sie dürfen bei dieser Anfrage keine Unteranfragen verwenden.

Erklären Sie wieder, warum bzw. warum nicht Sie DISTINCT verwenden.

Die erwartete Antwort ist:

pres_name	birth_year
Wilson W	1856
Taft W H	1857
Harding W G	1865

3 Datensätze

- c) Welche Präsidenten (bzw. Präsidentschaftskandidaten) haben keine Wahl verloren, bevor Sie eine Wahl gewonnen haben? Geben Sie das Jahr der gewonnenen Wahl mit aus und sortieren Sie die Ausgabe danach. (Hinweis: Clinton steht in der Datenbank noch als Kandidat, der die Wahl 1992 gewonnen hat, aber nicht als Präsident. Er soll mit berücksichtigt werden.)

Begründen Sie bitte kurz die Verwendung bzw. Nicht-Verwendung von DISTINCT. Die erwartete Antwort ist:

candidate	election_year
Washington G	1789
Washington G	1792
Johnson L B	1964
Nixon R M	1968
Nixon R M	1972
Carter J E	1976
Reagan R	1980
Reagan R	1984
Bush G	1988
Clinton W J	1992

10 Datensätze

(Wieder fällt auf, dass die Datenbank bei den Wahlen eine große Lücke hat.)