

DBS - Übungszettel 10

Gen Ogura, Jurek Foltys, Onur Pikel

Konzeption

- Deutschlandkarte
- Verknüpfung der Tabellen durch Bundesländer
- Bsp.-Abfrage: Anteil Studierende an Gesamtbevölkerung

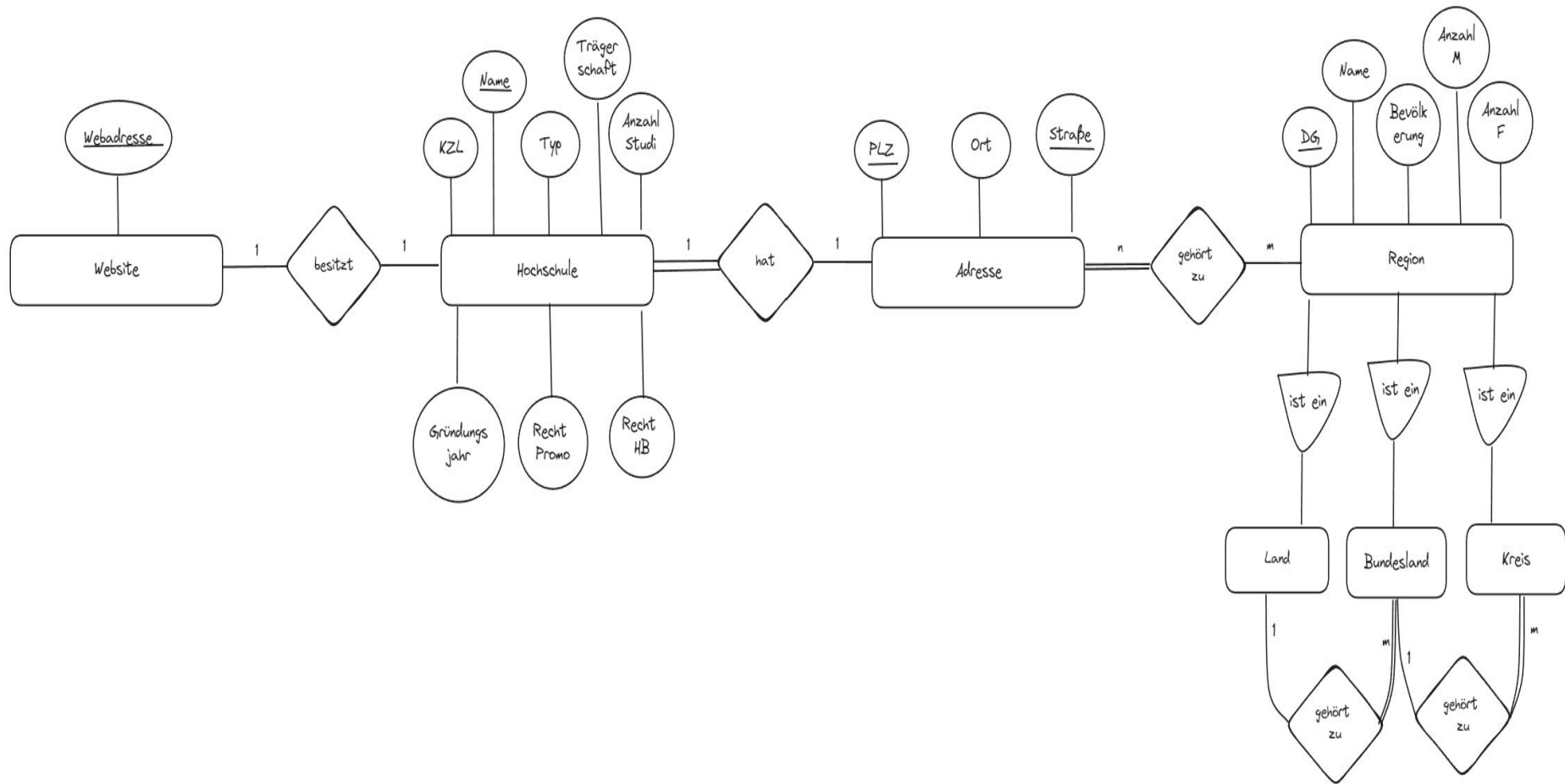
Relevante Datenfelder

Für verknüpfung der Tabellen:

- bundesland (hochschule)
- name (bevölkerung)

Für die Datenerhebung:

- Hochschulname (hochschule)
- Anzahl studierende (hochschule)
- Insgesamt (bevölkerung)
- etc.



Website (Webadresse)

Hochschule (KZL, Name, Typ, Trägerschaft, AnzahlStudi, Gründungsjahr, RechtPromo,
RechtHB)

Adresse (PLZ, Ort, Straße)

Region (DG, Name, Bevölkerung, AnzahlM, AnzahlF)

Land (DG)

Bundesland (DG)
....

Kreis (DG)
....

besitzt (Webadresse, Name)
.....

hat (Name, PLZ, Straße)
.....

gehörtzuAR(PLZ, Straße, DG)
.....

gehörtzuLB(DGL, DGB)
.....

gehörtzuBK(DGB, DGK)
.....

Datenbank erstellen

- Postgresql herunterladen
- Datenbank mit CREATE DATABASE erstellen
- Tabellen mit geeignetem schema erstellen
- Primary Key und Foreign Key angeben

Datensätze bereinigen

- üs durch u ersetzen
- Leerzeichen entfernen
- “-” entfernen
- Überflüssige Daten entfernen
- Falsche/falsch gespeicherte Daten entfernen

Daten importieren

```
COPY tabellenname(spalte1, spalte2,...)  
FROM dateipfad  
DELIMITER ':'  
CSV HEADER;
```


SQL-Abfragen

Zahl aller Hochschulen in Berlin:

```
SELECT COUNT(hochschulname)  
FROM hochschule  
WHERE bundesland='Berlin';
```

SQL-Abfragen

Bundesländer nach Studierendenanzahl geordnet:

```
SELECT bundesland, SUM(studis) AS studierendenanzahl  
FROM hochschule  
GROUP BY bundesland  
ORDER BY SUM(studis) DESC;
```

SQL-Abfragen

Hochschulen mit Gesamtbevölkerung des Bundeslands:

```
SELECT hochschulname, insgesamt AS  
Gesamtbevölkerung  
FROM hochschule, bevölkerung  
WHERE name = bundesland;
```

SQL-Abfragen

Hochschulen mit Bundesland geordnet nach Einwohner des Bundeslandes:

```
SELECT hochschulname AS Name, name AS Bundesland  
FROM bevölkerung  
RIGHT JOIN hochschule ON name = bundesland  
Order By insgesamt DESC;
```

ACHTUNG!: Hochschulen ohne zugeordnete Bevölkerungszahl fallen raus.

SQL-Abfragen

Bundesland mit Frauenquote, absoluter Studentenzahl und relativer Studentenzahl:

```
SELECT bundesland AS Bundesland, SUM(studis) AS Studierendenzahl,  
(MAX(weiblich)*100)/MAX(insgesamt) AS Frauenquote,  
(MAX(studis)*10000)/MAX(insgesamt) AS Studierendenquote_pro_10000_Einwohner  
FROM bevölkerung  
JOIN hochschule ON name = bundesland  
GROUP BY bundesland  
ORDER BY SUM(studis) DESC;
```

bundesland	studierendenanzahl	frauenquote	studierendenquote_pro_10000_einwohner
Nordrhein-Westfalen	743388	50.98879129501129	34.911293361815424
Bayern	390639	50.482396620400046	40.82085102891358
Baden-Württemberg	348718	50.3963074600162	28.94171648748783
Hessen	252834	50.6965966554849	64.31651479497322
Niedersachsen	192482	50.740751933419176	33.85894424268959
Berlin	191058	50.91683618485156	112.0830538358155
Thuringen	165464	50.561441684071156	545.5355018652032
Rheinland-Pfalz	113320	50.60880227931188	71.20926150776
Hamburg	107313	51.08053286204589	220.9265575898383
Sachsen	102371	50.79632377845954	72.39574054024422
Schleswig-Holstein	62425	51.12979849455011	87.83484070200151
Sachsen-Anhalt	56820	50.896831352900314	87.92930533242051
Brandenburg	54289	50.8323115576913	79.28460807536332
Mecklenburg-Vorpommern	36608	50.890088173630446	77.85661560153723
Bremen	36330	50.61311442855808	267.7173862255864
Saarland	30761	50.93002077234437	160.6683416174222

(16 Zeilen)

Code