

Datenbanken-Modul Übung 10

Erstellt von Ilhan Kapcik, Rosan Sharma und Rayan Ben Tanfous

Code und Documentation:

<https://github.com/users-pc/Datenbanken-Modul.git>

Aufgabe 1: Konzeption der Datenvisualisierung

Projekt-Idee

Unsere interaktive Visualisierung Website soll nicht nur Antworten liefern, sondern auch

Zusammenhänge verdeutlichen. Sie ermöglicht dem Benutzer, die Daten aus verschiedenen

Perspektiven zu betrachten und individuell zu analysieren. Und der Nutzer kann auch die Daten aktualisieren und verwalten.

Voraussetzungen

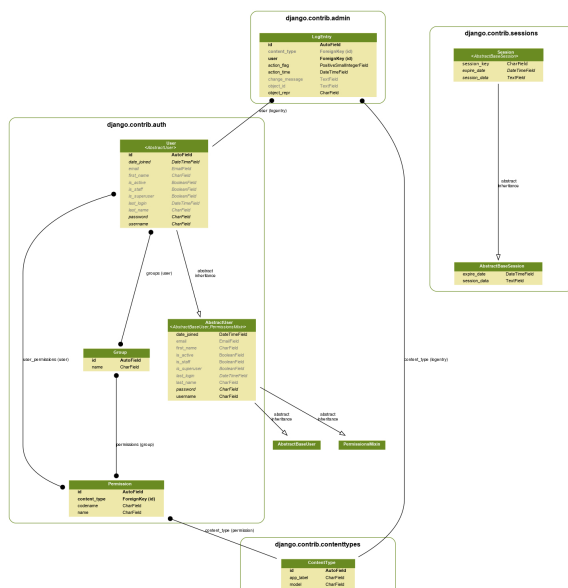
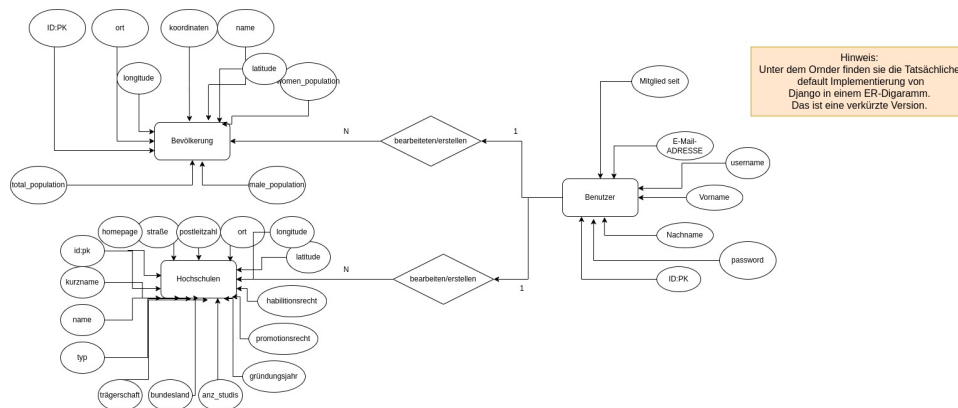
- **PostgreSQL**: Datenbank
- **Django**: Webframework für Python
- **Folium**: Kartenvisualisierung
- **Nominatim**: Koordinaten Datenbank von OpenStreetMap
- **Plotly**: Diagrammerstellung in Python
- **Bootstrap**: Vorlage von shuffle.dev

Umsetzungsschritte

1. Einrichtung von Django
2. Einrichtung von PostgreSQL
3. Erstellung der Datenbanktabellen mit Django ORM

- Import der Dateien `hochschulen.csv` und `bevölkerung.csv`
- Geocodierung der Adressen mit Nominatim
- Entwicklung der Website mit JavaScript, Bootstrap und Python

Aufgabe 2: Datenvisualisierung



Population Statistics	Hochschulen Deutschland																																																				
<table> <tr> <th colspan="2">Bevölkerung</th></tr> <tr> <td>id</td><td>string pk</td></tr> <tr> <td>ort</td><td>string</td></tr> <tr> <td>longitude</td><td>float</td></tr> <tr> <td>latitude</td><td>float</td></tr> <tr> <td>name</td><td>string</td></tr> <tr> <td>total_population</td><td>int</td></tr> <tr> <td>male_population</td><td>int</td></tr> <tr> <td>female_population</td><td>int</td></tr> </table>	Bevölkerung		id	string pk	ort	string	longitude	float	latitude	float	name	string	total_population	int	male_population	int	female_population	int	<table> <tr> <th colspan="2">Hochschulen</th></tr> <tr> <td>id</td><td>pk</td></tr> <tr> <td>kurzname</td><td>string</td></tr> <tr> <td>name</td><td>string</td></tr> <tr> <td>typ</td><td>string</td></tr> <tr> <td>trägerschaft</td><td>string</td></tr> <tr> <td>bundesland</td><td>string</td></tr> <tr> <td>anz_studis</td><td>int</td></tr> <tr> <td>gründungsjahr</td><td>int</td></tr> <tr> <td>promotionsrecht</td><td>boolean</td></tr> <tr> <td>habilitationsrecht</td><td>boolean</td></tr> <tr> <td>straße</td><td>string</td></tr> <tr> <td>postleitzahl</td><td>string</td></tr> <tr> <td>ort</td><td>string</td></tr> <tr> <td>longitude</td><td>float</td></tr> <tr> <td>latitude</td><td>float</td></tr> <tr> <td>homepage</td><td>string</td></tr> </table>	Hochschulen		id	pk	kurzname	string	name	string	typ	string	trägerschaft	string	bundesland	string	anz_studis	int	gründungsjahr	int	promotionsrecht	boolean	habilitationsrecht	boolean	straße	string	postleitzahl	string	ort	string	longitude	float	latitude	float	homepage	string
Bevölkerung																																																					
id	string pk																																																				
ort	string																																																				
longitude	float																																																				
latitude	float																																																				
name	string																																																				
total_population	int																																																				
male_population	int																																																				
female_population	int																																																				
Hochschulen																																																					
id	pk																																																				
kurzname	string																																																				
name	string																																																				
typ	string																																																				
trägerschaft	string																																																				
bundesland	string																																																				
anz_studis	int																																																				
gründungsjahr	int																																																				
promotionsrecht	boolean																																																				
habilitationsrecht	boolean																																																				
straße	string																																																				
postleitzahl	string																																																				
ort	string																																																				
longitude	float																																																				
latitude	float																																																				
homepage	string																																																				

Kartenvisualisierung

- **hochschule.csv:**
- Adresse
- Name
- Bundesland
- Anzahl Studierende
- Koordinaten
- Homepage

Flächenkartogramm

- **bevölkerung.csv:**
- Anzahl
- Insgesamt
- Männlich
- Weiblich
- Koordinaten der Orte
- Koordinaten der Bundesländer (Über Geofabrik nicht als Punkt)

Diagramme

- **Balkendiagramm:** Anzahl der Hochschulen im Vergleich zur Bevölkerungsdichte nach Bundesland.
- **Kreisdiagramme:** Kategorien der Hochschulen, Anzahl der Studierenden per Privat- und öffentlichen Hochschulen
- **Tabellen:** Wir wollen unseren Nutzern alle Quellen transparent darstellen über die Website, deswegen nehmen wir alle Attribute auf.

Aufgabe 3: Datenzugriff

Die Datenbereinigung ist unter "Übung10/main/views.py" unter Zeile 18 bis 72 und 92-131

zu finden.

Der PostgreSQL-Server ist remote. Sie können, wenn sie das Projekt aufsetzen, alle Daten über Django-Admin einsehen oder über die Weboberfläche.

Aufgabe 4: Weitere Schritte

1. Abfrage:

SQL: "

SELECT

trägerschaft,

SUM(anz_studis) AS Anzahl_Studierende

FROM Hochschulen

GROUP BY trägerschaft;

" Diese Abfrage summiert die Anzahl der Studierenden für jede

Trägerschaft

(öffentlich oder privat).

2. Abfrage:

SQL: "SELECT DISTINCT typ FROM Hochschulen;" → Gibt alle Hochschul

Kategorien aus

3. Abfrage:

"SELECT bundesland, SUM(anz_studis) AS Anzahl_Studierende

FROM Hochschulen

GROUP BY bundesland;" → Gruppiert nach Bundesländern und zählt die

Anzahl der

Studierenden

4. Abfrage:

SQL : "SELECT

h.bundesland,

SUM(h.anz_studis) AS Anzahl_Studierende,

b.total_population AS Bevölkerung

FROM Hochschulen h

JOIN Bevölkerung b ON h.bundesland = b.ort

GROUP BY h.bundesland, b.total_population;" → Verknüpfung der

Bundesland

Bevölkerungsdichte und Hochschulen mit der Anzahl der Studierenden in diesem

Bundesland.

5. Abfrage:

SQL: "SELECT

'Male' AS Geschlecht,

SUM(male_population) AS Anzahl_Männlich,

'Female' AS Geschlecht,

SUM(female_population) AS Anzahl>Weiblich,

'Total' AS Geschlecht,

SUM(total_population) AS Gesamtbevölkerung

FROM Bevölkerung;"

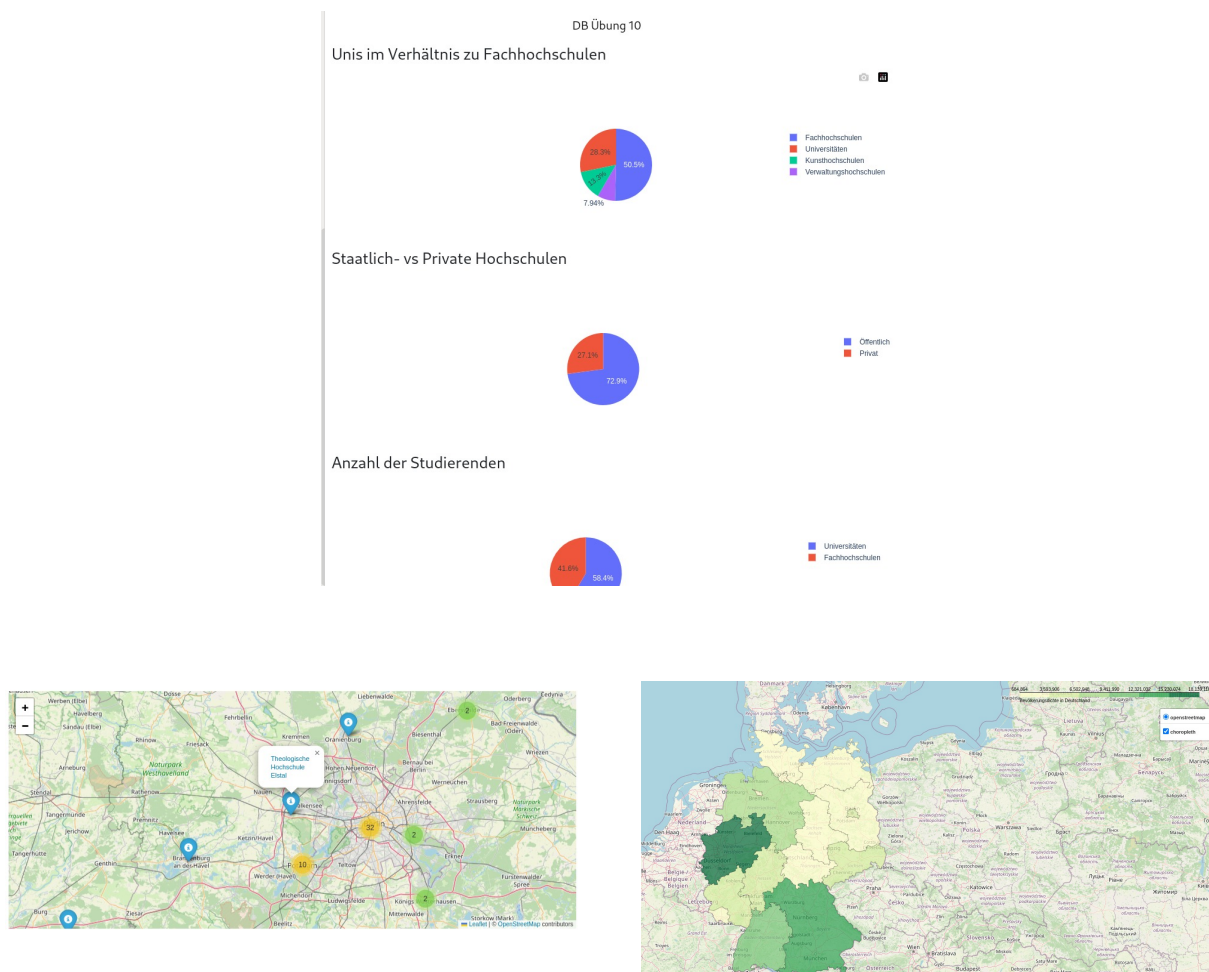
→ Gibt die Gesamtbevölkerung von Deutschland aus nach Geschlecht und der

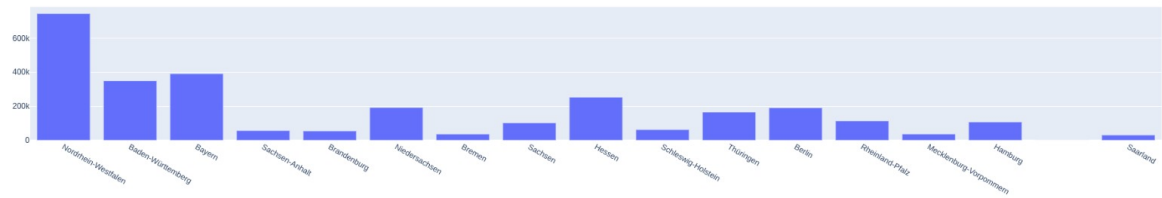
Gesamtbevölkerung

Dokumentation/Herangehensweise

Aufgabe 5:

Folgendes sind unsere Visualisierungen. Alles über die Projektwebsite sichtbar





Hochschulen

Vector Search

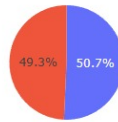
#id	Kurzname	Name	Typ	Trägerschaft	Bundesland	Anz. Studis	Gründungsjahr	Promotionsrecht	Habilitationsrecht	Straße	Postleitzahl	Ort	Homepage
1	Aachen FH	Fachhochschule Aachen	Fachhochschulen / HAW	öffentlich-rechtlich	Nordrhein-Westfalen	14086	1971	False	False	Bayernallee 11	52066	Aachen	http://www.fh-aachen.de
2	Aachen TH	Rheinisch-Westfälische Technische Hochschule Aachen	Universitäten	öffentlich-rechtlich	Nordrhein-Westfalen	45527	1870	True	True	Templergraben 55	52062	Aachen	https://www.rwth-aachen.de
3	Aalen H	Hochschule Aalen - Technik und Wirtschaft	Fachhochschulen / HAW	öffentlich-rechtlich	Baden-Württemberg	4559	1962	False	False	Beethovenstraße 1	73430	Aalen	http://www.hs-aalen.de
4	Albstadt-Sigmaringen H	Hochschule Albstadt-Sigmaringen	Fachhochschulen / HAW	öffentlich-rechtlich	Baden-Württemberg	2666	1971	False	False	Anton-Günther-Straße 51	72488	Sigmaringen	www.hs-albsig.de
5	Alfter HFuG	Alanus Hochschule für Kunst und Gesellschaft	Künstlerische Hochschulen	privat, staatlich anerkannt	Nordrhein-Westfalen	1564	1973	True	False	Villestraße 3	53347	Alfter	https://www.alanus.edu.de/home
6	Amberg-Weiden OTH	Ostbayerische Technische Hochschule Amberg-Weiden	Fachhochschulen / HAW	öffentlich-rechtlich	Bayern	4077	1994	True	False	Kaiser-Wilhelm-Ring 23	92224	Amberg	http://www.oth-ab.de/
7	Anhalt H	Hochschule Anhalt - Anhalt University of Applied Sciences	Fachhochschulen / HAW	öffentlich-rechtlich	Sachsen-Anhalt	7067	1991	True	False	Bernburger Straße 55	6366	Köthen	http://www.hs-anhalt.de
8	Ansbach H	Hochschule für angewandte Wissenschaften Ansbach	Fachhochschulen / HAW	öffentlich-rechtlich	Bayern	3898	1996	True	False	Residenzstraße 8	91522	Ansbach	http://www.hs-ansbach.de
9	Aschaffenburg TH	Technische Hochschule Aschaffenburg	Fachhochschulen / HAW	öffentlich-rechtlich	Bayern	3559	1995	False	False	Würzburger Straße 55	63743	Aschaffenburg	http://www.th-ab.de

Bevölkerung

Vector Search

#id	DG	Ort	Ingesamt	Männliche Bevölkerung	Weibliche Bevölkerung
1	1	Schleswig-Holstein	2953270	1443269	1510001
2	1001	Flensburg, kreisfreie Stadt	92550	45874	46676
3	1002	Kiel, kreisfreie Stadt	247717	120149	127568
4	1003	Lübeck, kreisfreie Stadt, Hansestadt	218095	104723	113372
5	1004	Neumünster, kreisfreie Stadt	79502	39256	40246
6	1051	Dithmarschen, Kreis	135252	66561	68691
7	1053	Herzogtum Lauenburg, Kreis	203712	99689	104023
8	1054	Nordfriesland, Kreis	169043	82669	86374
9	1055	Ostholstein, Kreis	203606	97822	105784
10	1056	Pinneberg, Kreis	322130	157663	164467
11	1057	Plön, Kreis	131266	63357	67909
12	1058	Rendsburg-Eckernförde, Kreis	278979	137154	141825
13	1059	Schleswig-Flensburg, Kreis	206038	101614	104424
14	1060	Segeberg, Kreis	284988	140654	144334
15	1061	Steinburg, Kreis	132419	65268	67151
16	1062	Stormarn, Kreis	247973	120816	127157
17	2	Hamburg	1892122	925616	966506
18	3	Niedersachsen	8140242	4009822	4130420
19	31	Braunschweig, Statistische Region	1610865	796058	814807
20	3101	Braunschweig, kreisfreie Stadt	251804	124770	127034

Previous 1 2 3 Last Page Next



■ Weiblich
■ Männlich

[DB Übung 10](#)[Documentation](#) [Github](#)

von Ilhan Kapcik, Rosan Sharma & Rayan Ben Tanfous

Aufgabe 7: Dokumentation

1. Einrichtung von PostgreSQL und Django

Problem: Die Installation und Konfiguration von PostgreSQL kann am Anfang

verwirrend sein. Wir haben einen Remote-Server genommen und ein Postgresql

erstellt, damit wir als Team zusammenarbeiten können. Wie ?

Lösung: Wir haben uns in der Dokumentation umgesehen, Tutorials angeschaut und

uns gegenseitig unterstützt.

2. Datenbanktabellen erstellen

Problem: Wie erstellen wir überhaupt Tabellen? Welche Felder sollten wir verwenden? Wie definieren wir Beziehungen zwischen den Tabellen?

Lösung: Wir haben uns mit dem Django ORM vertraut gemacht und unsere Modelle

definiert. Dabei lernen wir, wie wir Primärschlüssel, Fremdschlüssel und Embedded

SQL verwenden. Dazu gehört auch der Vorgang der Datenbereinigung.

3. Daten importieren und geografische Koordinaten

hinzufügen

Problem: Wir haben CSV-Dateien mit Hochschuldaten und Bevölkerungszahlen. Wie

importieren wir diese in unsere Datenbank? Und wie fügen wir den Hochschulen

geografische Koordinaten hinzu?

Lösung: Wir haben Python-Skripte geschrieben, um die Daten zu

importieren und
Nominatim zu verwenden, um Adressen in Koordinaten umzuwandeln
(Breiten- und
Längengrade).

4. Entwicklung der Website

Problem: Wie gestalten wir eine benutzerfreundliche Website? Welche
Technologien

verwenden wir? Wie verknüpfen wir Frontend und Backend?

Lösung: Wir setzen auf Bootstrap für das Design, JavaScript für die
Interaktivität und

Django für die Logik. Dabei bietet Django eine gute Umgebung.

5. Datenvisualisierung

Problem: Wie präsentieren wir die Daten auf unserer Website? Welche
Diagramme

und Karten sind am besten geeignet?

Lösung: Wir haben uns mit Folium für Karten, Plotly für Diagramme und
anderen

Tools vertraut gemacht. Dabei haben wir gelernt, wie man Daten sinnvoll
und gut

Darstellen kann.

6. Weitere Schritte

Problem: Wie können wir fehlende Daten hinzufügen und bessere
Visualisierungen
machen?

Lösung: Wir haben das Internet durchforstet und viele Opendata Quellen
gefunden,

wie Nominatim und Geofabrik zur Aufwertung unserer Daten.