

Informatik II

Vorlesung am D-BAUG der ETH Zürich

Dr. Felix Friedrich und Dr. Hermann Lehner

Vorlesung 12, 2018

Datenbanksysteme: Das Entity Relationship (ER) Modell, das Relationale Modell und SQL.

Demo: Java & SQL: Eine GUI Applikation fuer die VBZ Daten

Literatur

Literatur: Kemper, Eickler: Datenbanksysteme: Eine Einführung. Oldenbourg Verlag, 9. Auflage, 2013.

Quellen: Basismaterial wurde von Prof. Donald Kossmann & Martin Kaufmann freundlicherweise zur Verfügung gestellt. Weitere Quelle: Folien zu *Datenbanksysteme: Eine Einführung*, Lehrstuhl III Datenbanksysteme, Prof. Kemper, TU München

2

Ziele

Nutzen von Datenbanksystemen verstehen,
Modellierungskennntnisse

- ER Modell (Modellierung der Weltsicht)
- Relationales Modell (Modellierung für die DB)

Datenbanksystem anwenden

- Java + SQL

Datenbankverwaltungssysteme

- Ein Datenbankverwaltungssystem (DBMS) ist ein Werkzeug zur Erstellung und Ausführung datenintensiver Anwendungen
- grosse Datenmengen
- grosse Datenströme



3

4

Typische Anwendungen

- **Bank**
z.B. Konten / „Geldtransfer“
- **Bibliothek**
z.B. Bücher / „Ausleihen“
- **Facebook, Twitter, ...**
z.B. Freunde, „Sende Tweet“
- **Geoinformationssysteme**
z.B. Topographische Information, "Erzeuge Karte"
- "Alles, was Sie mit einer Tabelle machen wollen, mit Excel nicht erledigen können und nicht selbst programmieren möchten".

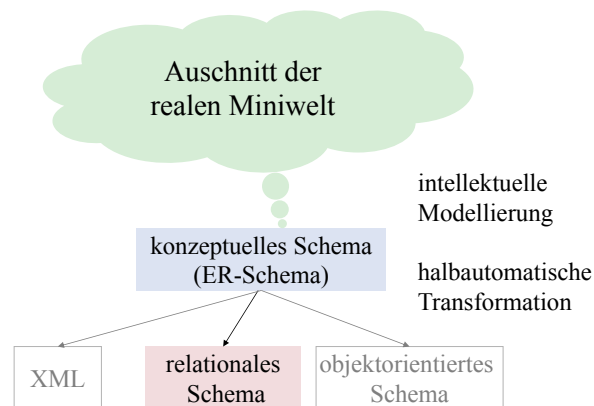
5

Wozu Datenbanksysteme?

- Vermeide Redundanz und Inkonsistenz
- Deklarativer Zugriff auf die Daten und Unabhängigkeit von der Implementation (physische Datenunabhängigkeit)
- Synchronisiere gleichzeitigen Datenzugriff
- Sicherheit, Vertraulichkeit
- **Minimiere Kosten und Aufwand**
Ähnliche Funktionalität selbst zu implementieren würde Jahre in Anspruch nehmen

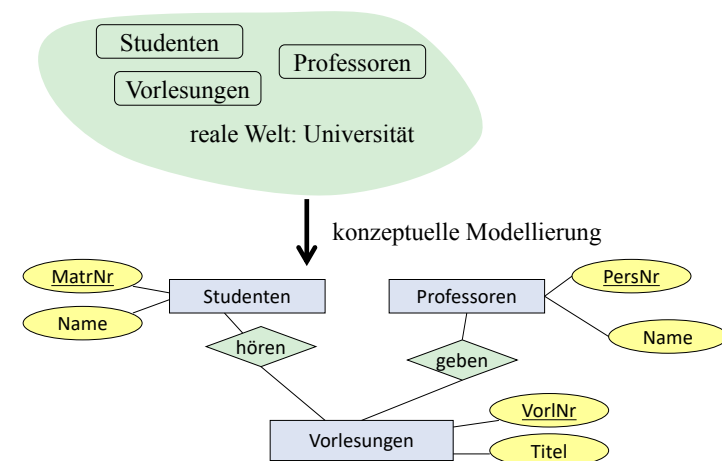
6

Datenmodellierung



7

1. Teil heutige Vorlesung: Modellierung (ER-Modell)



8

2. Teil heutige Vorlesung: Relationales Datenmodell

Studenten		hören		Vorlesungen	
Legi	Name	Legi	VorlNr	VorlNr	Titel
26120	Fichte	25403	5022	5001	Grundzüge
25403	Jonas	26120	5001	5022	Glaube und Wissen
...

```

select Name
from Studenten, hören, Vorlesungen
where Studenten.Legi= hören.Legi and
       hören.VorlNr= Vorlesungen.VorlNr and
       Vorlesungen.Titel = `Grundzüge`;
    
```

```

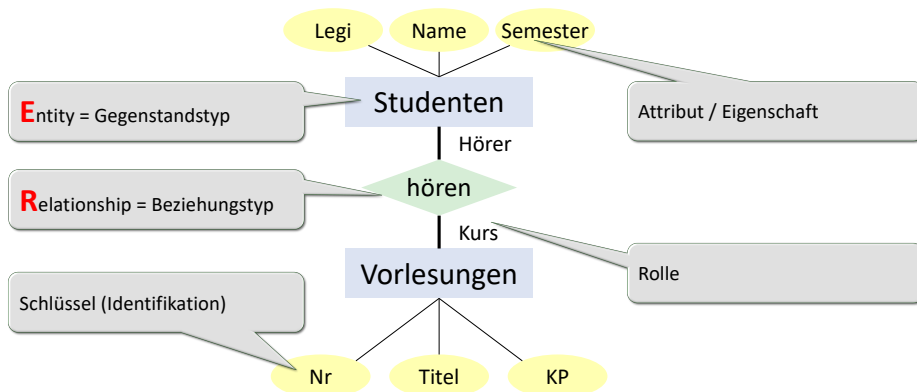
update Vorlesungen
set Title = `Grundzüge der Logik`
where VorlNr = 5001;
    
```

9

Entity Relationship Modell

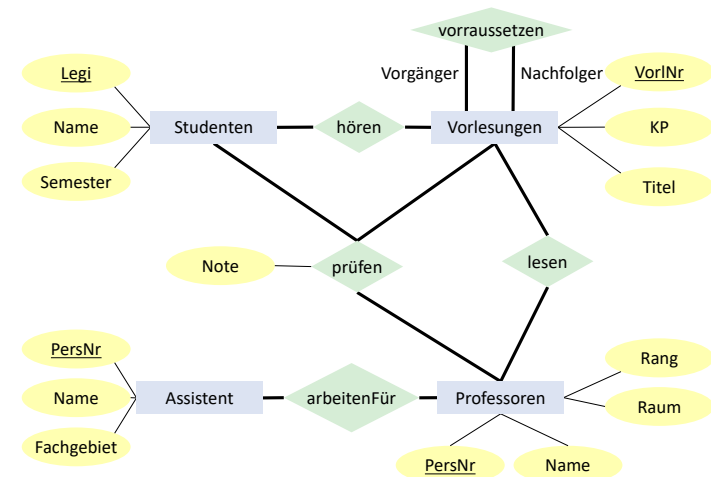
10

Entity/Relationship (ER) Modell



11

Modell einer Universität



12

... in natürlicher Sprache

- Studenten haben LegiNr, Name und Semester. Die LegiNr identifiziert einen Studenten eindeutig.
- Vorlesungen haben eine VorlNr, Kreditpunkte und einen Titel. VorlNr identifiziert eine Vorlesung eindeutig.
- Professoren haben PersNr, Name, Rang und Raum. PersNr identifiziert einen Professor eindeutig.
- Assistenten haben PersNr, Name und Fachgebiet. PersNr. identifiziert einen Assistenten eindeutig.
- Studenten hören Vorlesungen
- Vorlesungen können Voraussetzung für andere Vorlesungen sein.
- Professoren lesen Vorlesungen.
- Assistenten arbeiten für Professoren
- Studenten werden von Professoren über Vorlesungen geprüft. Studenten erhalten Noten als Teil dieser Prüfungen.
- Ist das die einzig mögliche Interpretation?
- Nein: zu einem ER-Model gehört immer noch Dokumentation und/oder gesunder Menschenverstand

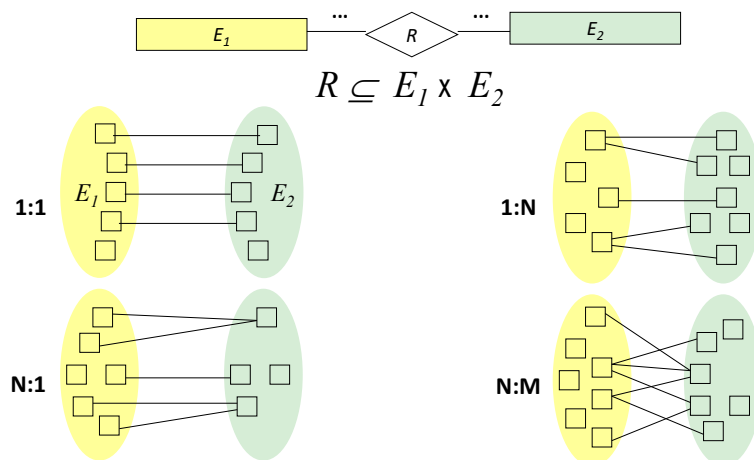
13

Warum ER?

- Vorteile
 - ER Diagramme sind einfach zu erstellen und editieren
 - ER Diagramme sind aufgrund der grafischen Darstellung einfach zu verstehen (vom Laien)
 - ER Diagramme beschreiben alle Informationsanforderungen
- Allgemeines
 - Viele Tools verfügbar
 - Kontroverse, ob ER/UML in der Praxis von Nutzen ist
 - Keine Kontroverse, dass jeder ER/UML lernen sollte

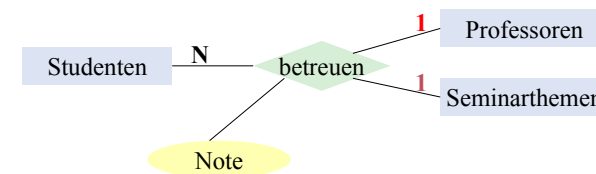
14

Funktionalitätsangaben



15

Beispiel: Seminar



Um das zu verstehen, denken Sie am besten in Abbildungen (Funktionen): eine Abbildung $f(x,y)$ hat pro Parameter x,y nur einen Wert.

betreuen: Professoren x Studenten → Seminarthemen

betreuen: Seminarthemen x Studenten → Professoren

16

Konsistenzbedingungen des Seminar

Einschränkungen

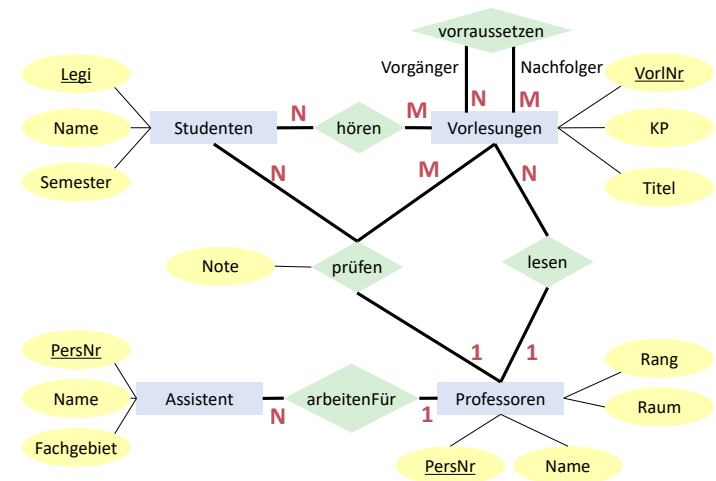
1. Studenten dürfen bei einem Professor nur ein Seminarthema bearbeiten
2. Studenten können dasselbe Seminarthema nur einmal bearbeiten

Möglichkeiten

1. Professoren können das Seminarthema für andere Studenten wiederverwenden
2. Dasselbe Thema kann von verschiedenen Professoren verwendet werden

17

Universität mit Funktionalitäten



18

Daumenregeln

Wann Attribut, wann Entität?

- Entität, wenn das Konzept mehr als eine Beziehung hat
- Attribut, wenn das Konzept nur eine 1:1 Beziehung hat

Partitionierung von ER-Modellen

- Realistische Modelle sind grösser als eine Seite
- Nach Bereichen / Organisationseinheiten partitionieren
- Kein gutes automatisches Graphenpartitionierungstool bekannt

Tipps

- Keine Redundanz modellieren
- Je weniger Entitäten desto besser

19

ER Modellierung: Zusammenfassung

ER beschreibt eine Miniwelt

- Das "was" und die Regeln
- ER ist statisch. Es beschreibt keine Übergänge

Nützlich zum Erstellen von Software zur Beantwortung von (An)fragen über die Miniwelt

- es folgt nun: ER-Modell → relationales Modell

Ähnliche Modellierungsmöglichkeiten bietet [UML](#) (mehr auf OOP zugeschnitten)

Auch andere graphische Darstellungen des ER Modells gebräuchlich, z.B. "Krähfußnotation" optisch näher bei UML

20

Relationales Modell

Die Welt in Tabellen

Relationales Modell, Formalismus

- **Relation R**

- $R \subseteq D_1 \times \dots \times D_n$
- D_1, D_2, \dots, D_n sind Domänen

Beispiel: *Telefonbuch* \subseteq *string* \times *string* \times *integer*

- **Tupel:** $t \in R$

Beispiel: $t = (\text{„Mickey Mouse“}, \text{„Main Street“}, 4711)$

- **Relationenschemata** werden wie folgt beschrieben

Telefonbuch: $\{\{ \text{Name: string, Strasse: string, Telefon#: integer } \}$



21

22

Relationales Modell

Telefonbuch		
Name	Strasse	Telefon#
Mickey Mouse	Main Street	4711
Minnie Mouse	Broadway	94725
Donald Duck	Broadway	95672
...

Ausprägung: Zustand der Datenbank

Schlüssel: minimale Menge von Attributen, welche ein Tupel eindeutig identifizieren
z.B. {Telefon#} oder {Name, Geburtstag}

Primärschlüssel (durch Unterstreichung hervorgehoben): Ausgewählter Schlüssel, welcher üblicherweise zur Identifikation eines Tupels in einer Relation verwendet wird.

23

Regel #1: Darstellung von Entities

Studenten:

$\{\{ \text{Legi: integer, Name: string, Semester: integer} \}$

Vorlesungen:

$\{\{ \text{VorlNr: integer, Titel: string, KP: integer} \}$

Professoren:

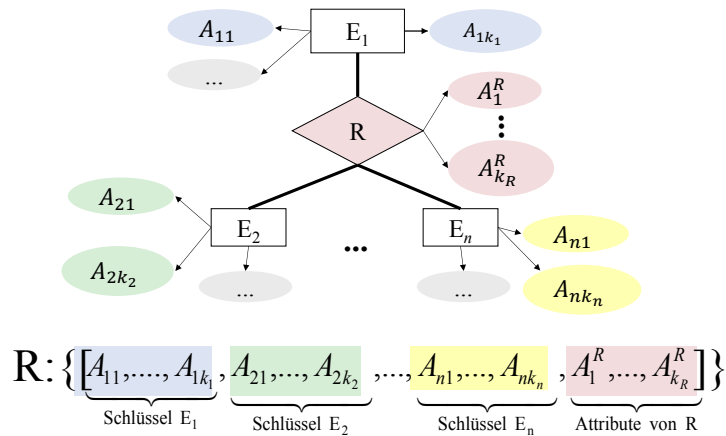
$\{\{ \text{PersNr: integer, Name: string, Rang: string, Raum: integer} \}$

Assistenten:

$\{\{ \text{PersNr: integer, Name: string, Fachgebiet: string} \}$

24

Regel #2: Darstellung von Beziehungen



25

Darstellung von Beziehungen

hören:

$\{ [\text{Legi: integer, VorlNr: integer}] \}$

lesen:

$\{ [\text{PersNr: integer, VorlNr: integer}] \}$

Fremdschlüssel, identifizieren
Tupel aus anderen Entitäten

arbeitenFür:

$\{ [\text{AssiPersNr: integer, ProfPersNr: integer}] \}$

voraussetzen:

$\{ [\text{Vorgänger: integer, Nachfolger: integer}] \}$

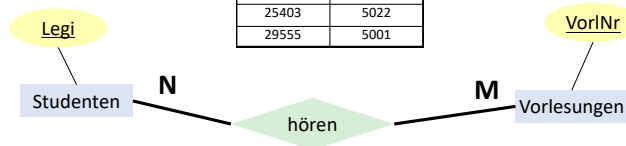
prüfen:

$\{ [\text{Legi: integer, VorlNr: integer, PersNr: integer, Note: decimal}] \}$

26

Ausprägung von hören

Studenten		hören		Vorlesungen	
Legi	...	Legi	VorlNr	VorlNr	...
26120	...	26120	5001	5001	...
27550	...	27550	5001	4052	...
...	...	28106	5041
		28106	5052
		28106	5216
		28106	5259
		29120	5001
		29120	5041
		29120	5049
		29555	5022
		25403	5022
		29555	5001



27

Zur Regel #2: Benennung der Attribute?

Spezifiziert das ER-Modell **Rollen**, dann

- nimm den Namen der jeweiligen Rolle

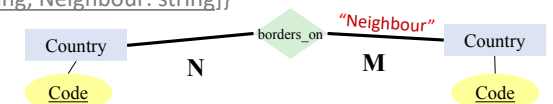
$\text{voraussetzen:} \{ [\text{Vorgänger: string, Nachfolger: string}] \}$



andernfalls

- benutze die Namen der Schlüsselattribute der Entitäten
- bei Mehrdeutigkeit erfinde aussagekräftigen Namen

$\text{borders_on:} \{ [\text{Code: string, Neighbour: string}] \}$



28

Regel #3: Zusammenfassung von Relationen

Vorlesungen : {[VorlNr, Title, CP]}
Professoren : {[PersNr, Name, Level, Room]}
lesen: {[VorlNr, PersNr]}

Fasse (nur) Relationen mit **gleichem Schlüssel** zusammen
 (also auch nur (N:1), (1:N) oder (1:1) Beziehungen)

Zusammenfassen:

Vorlesungen : {[VorlNr, Title, CP, **gelesenVon**]}
Professoren : {[PersNr, Name, Level, Room]}



Ausprägung von Professoren und Vorlesungen

Professoren				Vorlesungen			
PersNr	Name	Rang	Raum	VorlNr	Titel	KP	gelesenVon
2125	Sokrates	FP	226	5001	Grundzüge	4	2137
2126	Russel	FP	232	5041	Ethik	4	2125
2127	Kopernikus	AP	310	5043	Erkenntnistheorie	3	2126
2133	Popper	AP	52	5049	Mäeutik	2	2125
2134	Augustinus	AP	309	4052	Logik	4	2125
2136	Curie	FP	36	5052	Wissenschaftstheorie	3	2126
2137	Kant	FP	7	5216	Bioethik	2	2126
				5259	Der Wiener Kreis	2	2133
				5022	Glaube und Wissen	2	2134
				4630	Die 3 Kritiken	4	2137



Das funktioniert NICHT

Professoren					Vorlesungen		
PersNr	Name	Rang	Raum	liest	VorlNr	Titel	KP
2125	Sokrates	FP	226	5041	5001	Grundzüge	4
2125	Sokrates	FP	226	5049	5041	Ethik	4
2125	Sokrates	FP	226	4052	5043	Erkenntnistheorie	3
...	5049	Mäeutik	2
2134	Augustinus	AP	309	5022	4052	Logik	4
2136	Curie	FP	36				3
							2
							2
							2
							4

Problem: Redundanz und Anomalien
 PersNr ist kein gültiger Schlüssel für Professoren mehr



Relationales Modell der Uni-DB

Professoren				Studenten			Vorlesungen				hören	
PersNr	Name	Rang	Raum	Legi	Name	Semester	VorlNr	Titel	KP	gelesenVon	Legi	VorlNr
2125	Sokrates	FP	226	24002	Xenokrates	18	5001	Grundzüge	4	2137	26120	5001
2126	Russel	FP	232	25403	Jonas	12	5041	Ethik	4	2125	27550	5001
2127	Kopernikus	AP	310	26120	Fichte	10	5043	Erkenntnistheorie	3	2126	27550	4052
2133	Popper	AP	52	26830	Aristoxenos	8	5049	Mäeutik	2	2125	28106	5041
2134	Augustinus	AP	309	27550	Schopenhauer	6	4052	Logik	4	2125	28106	5052
2136	Curie	FP	36	28106	Carnap	3	5052	Wissenschaftstheorie	3	2126	28106	5216
2137	Kant	FP	7	29120	Theophrastos	2	5216	Bioethik	2	2126	28106	5259
				29555	Feuerbach	2	5259	Der Wiener Kreis	2	2133	29120	5001
							5022	Glaube und Wissen	2	2134	29120	5041
								Die 3 Kritiken	4	2137	29120	5049
											29555	5022

Assistenten				voraussetzen		prüfen			
PersNr	Name	Fachgebiet	Boss	Vorgänger	Nachfolger	Legi	Nr	PersNr	Note
3002	Platon	Ideenlehre	2125	5001	5041	28106	5001	2126	1
3003	Aristoteles	Sylogistik	2125	5001	5043	25403	5041	2125	2
3004	Wittgenstein	Sprachtheorie	2126	5001	5049	27550	4630	2137	2
3005	Rhetikus	Planetenbewegung	2127	5041	5216				
3006	Newton	Keplersche Gesetze	2127	5043	5052				
3007	Spinoza	Gott und Natur	2126	5041	5052				
				5052	5259				



```

public static void loadDatabaseFile(File file, String query) {
    if (file == null) return;
    records.clear();
    model.setRowCount(0);
    Connection connection = null;
    try {
        String url = "jdbc:mysql://" + file.getAbsolutePath();
        connection = DriverManager.getConnection(url);
        ResultSet rs = connection.createStatement().executeQuery(query);
        while (rs.next()) {
            Record r = new Record();
            r.setString("name", rs.getString("name"));
            r.getDouble("delay", rs.getDouble("delay"));
            r.getDouble("latitude", rs.getDouble("latitude"));
            r.getDouble("longitude", rs.getDouble("longitude"));
            r.setDouble("maxdelay", rs.getDouble("maxdelay"));
            records.add(r);
        }
    } catch (SQLException e) {
        Alerts.showExceptionInfo("Ein Problem ist aufgetreten bei Abfragen der Datenbank", e);
    } finally {
        if (connection != null) {
            connection.close();
        }
    } catch (SQLException e) {
        Alerts.showExceptionInfo("Ein Problem ist aufgetreten beim Schliessen der Datenbankverbindung", e);
    }
}

```

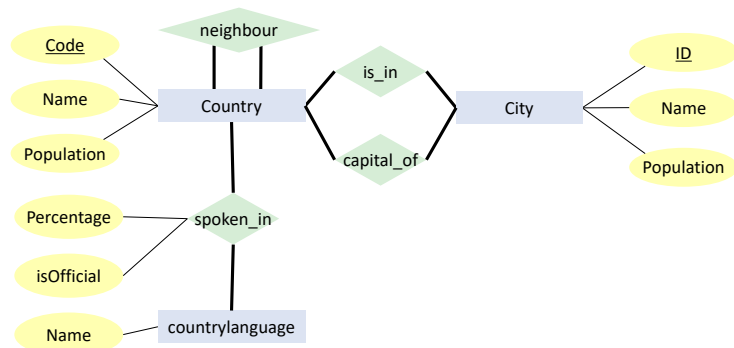
HANDS-ON: ZUGRIFF AUF SQL DATENBANKEN AUS EINER JAVA APPLIKATION

33

ANHANG: WELT-DATENBANK

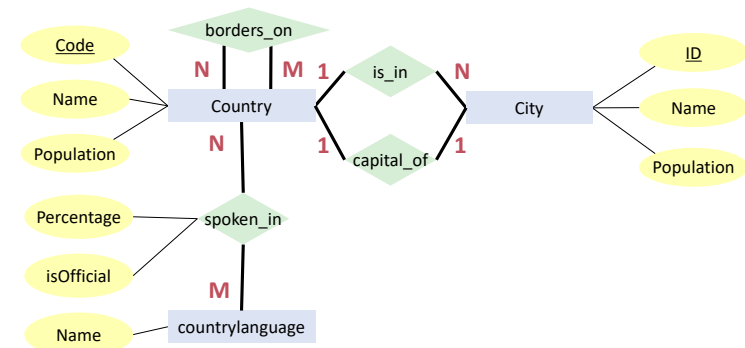
34

[Ein Modell der Welt]



35

[Modell der Welt mit Funktionalitäten]



36

[Die Welt in Tabellen]

ID	Name	CountryCode	District	Population	Code	Name	Continent	Region	SurfaceArea	IndepYear	Population	LifeExpectancy	GNP	GNPOID	LocalName
3248	Bern	CHE	Bern	122700	ABW	Aruba	North America	Caribbean	193.00	NULL	103000	78.4	828.00	793.00	Aruba
3247	Basel	CHE	Basel-Stadt	166700											
3246	Geneve	CHE	Geneve	173500	AFG	Afghanistan	Asia	Southern and Central Asia	652080.00	1919	22720000	45.0	5976.00	NULL	Afghanistan/Afg
3245	Zürich	CHE	Zürich	336800											
3249	Lausanne	CHE	Vaud	114500											
3117	Oldenburg	DEU	Niedersachsen	154125	AGO	Angola	Africa	Central Africa	1249700.00	1975	12878000	38.3	6648.00	7984.00	Angola
3130	Remscheid	DEU	Nordrhein-Westfalen	120125											
3131	Heilbronn	DEU	Baden-Württemberg	119526	AIA	Anguilla	North America	Caribbean	96.00	NULL	8000	76.1	63.20	NULL	Anguilla
3132	Pforzheim	DEU	Baden-Württemberg	117227	ALB	Albania	Europe	Southern Europe	28748.00	1912	3401200	71.6	3205.00	2500.00	Shqipëria
3133	Offenbach am Main	DEU	Hessen	116627											
3134	Ilm	DEU	Baden-Württemberg	116103	AND	Andorra	Europe	Southern Europe	468.00	1278	78000	83.5	1630.00	NULL	Andorra
					ANT	Netherlands Antilles	North America	Caribbean	800.00	NULL	217000	74.7	1941.00	NULL	Nederlandse A

CountryCode	Language	IsOfficial	Percentage
ABW	Dutch	T	5.3
ABW	English	F	9.5
ABW	Papiamentu	F	76.7
ABW	Spanish	F	7.4
AFG	Balochi	F	0.9
AFG	Pashtu	T	55.4

- (borders_on nicht modelliert)