



Session:

Building the database – hands on experience

V1.0

Pier-Giorgio Zaccheddu

Federal Agency for Cartography and
Geodesy (BKG)

Richard-Strauss-Allee 11

60598 Frankfurt am Main

Tel.: +49 69 6333 305

Fax: +49 69 6333 441

Email: pier.zaccheddu@bkg.bund.de

Content

First steps before processing the names collected

- Structure for storing the data, maintaining records
- Feature categories/classification

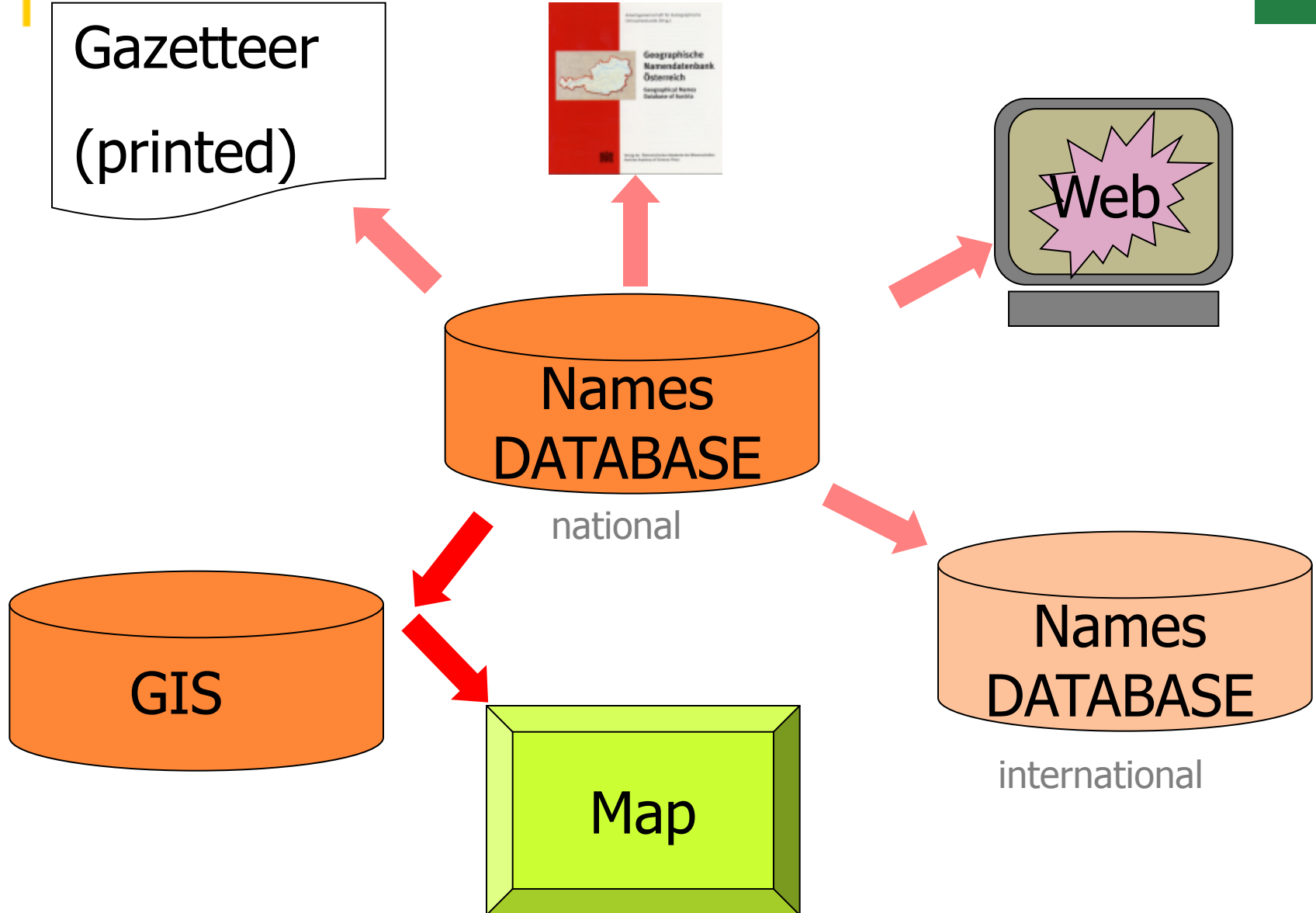
Demonstration of the processing of names collected

- Spreadsheet or Database?
- Create main tables and enter names into a database
- Database queries and reports (e.g. gazetteers as printouts)

Hands on experience



What are the purposes of a DB?



Structure for storing the data, maintaining the records

1. Gathering of names information

2. Authorizing toponyms

3. Storing the data, maintaining the records

paper:	Gazetteer
digital:	Text file, Spreadsheet, Database

Structure: Tables !!

Columns:
attribute information to the name

Rows:
One for
each name

Name	Feature Type	Coordinates
Name1			
Name2			
Name3			
.....			

Structure: Tables !!

Columns:
attribute information to the name

Rows:
One for
each name

Name	Feature Type	Coordinates
Rio de Janeiro	Populated Place	-22,92	-43,38
Rio Iguaçu	stream	-22,74	-43,25
Chácara Entrerios	mountain	-22,48	-43,05
Petrópolis	Administ. Unit	-22,48	-43,18

Structure: Tables !!

Columns:
attribute information to the name

Absolutely necessary!

Rows:
One for
each name

Field name	Data type	Description
Geographical Name	Text	the name
Feature Type	Text	e.g. River, Mountain, Populated Place. The types should be chosen from a standardized list, if exists.
Coordinates, e.g.:		
Longitude	Number	Geographical Coordinates
Latitude	Number	Geographical Coordinates Instead of Geographical Coordinates, other systems may be used, for instance UTM with X and Y coordinates.



Structure for storing the data, maintaining the records

Structure: Tables !!

Rows:
One for
each name

recommended

Field name

FeatureID
(for databases)

Variant Name
Administrative Unit

Map Sheet

Description

Source

Date

Status

Columns:

attribute information to the name

Data type

Integer

Text

Text

Text

Text

Text

Date

Text

Description

A unique identifier assigned to the name. This ID will be used to link the name with other database tables.

Other names assigned to the feature, if any

Name or Code of the Administrative Unit where the name is situated in.

Reference to a map sheet in a topographic map series.

The data type may be Integer if the sheet name contains only numbers no letters.

Comments, e.g. on the history of the name, and verbal statements on the extension of the feature.

Source of the name. e.g. captured in the field by interview

Date of the entry to the Database.
Other option: date of approval by the Board.

Comment, e.g. name is approved or not approved by the Board.



Structure for storing the data, maintaining the records

INSPIRE GN schema: elements

mandatory

- name(s) (text, spelling)
- geometry
- feature type
- unique identifier

‘voidable’

- language {three letter codes from ISO 639-3 or -5}
- nameStatus {official, standardised, historical, other}
- link to relatedSpatialObject
- script {four letters codes defined in ISO 15924}
- nativeness {endonym, exonym}
- transliterationScheme
- grammatical gender {masc., fem., neuter, common}
- grammatical number {singular, plural, dual}
- pronunciation
- sourceOfName
- typeLocal
- lifeCycleInfo (begin/end of the object in the source DB)
- ...



CODE

This column contains a specific code of maximally four letters which provides a further specification of the nature of the particular feature. It occasionally proved difficult to draw a sharp distinction between the names of polders and area names. The category 'polder' therefore includes all those names in which the word 'polder' occurs. In addition, this category also includes the names of drained lakes which are clearly visible as a landscape unit on the map, such as Schermer, Flevoland, etc.

Where no recognizable feature can be discerned from the topography, the name is coded as an area or local name.

The explanation of the 40 codes used is as follows:

ADMD	administrative division
AF	airfield, airport
AN	area name, local name
AQDT	aqueduct
BLDG	building, group of buildings
BRDG	bridge, viaduct
CEM	cemetery
CHAN	channel, fairway
CNAN	canal, waterway, ditch
DAM	dam, groyne, pier
DIKE	dike, embankment
DIST	residential district
FORT	fort, entrenchment
FRST	forest
HBR	harbour
HILL	hill, dune, mountain
HTH	(former) heathland
IND	industrial estate
INL	inlet, estuary
IS	island
LAKE	lake, pool, pond
LH	lighthouse
LOCK	sluice, lock
MISC	miscellaneous landmarks e.g. beacon, belvedere, dock, decoy, burial mound, light,

Gazetteer of The Netherlands

- Feature codes

... is it a river,
a mountain, a
populated place, etc?

NAME OF FEATURE	CODE	LONGITUDE	LATITUDE	UTM-REF.	ADM. SHEET CODE +ED.	REMARKS
Enning	BLDG	06 51 53	51 54 55	32U LC53155385	GE 410-4	
Ennipwetering	CNAN	04 52 40	52 10 11	31U FT28408170	U 310-4	
Ens	POPL	05 49 41	52 38 14	31U FU91353580	F1 21W-4	
ENSCHEDÉ	POPL	06 53 48	52 13 14	32U LC56308775	0 340-4	35
Enschotsche Akkers	AN	05 08 10	51 34 15	31U FT48001555	NB 500-4	
Ensertoct	CNAN	05 49 54	52 37 45	31U FU91653490	F1 21W-4	
Enservaart	CNAN	05 50 18	52 40 36	31U FU91854020	F1 21W-4	
Ens. Gent	BLDG	05 53 00	52 47 33	31U FU914405320	0 16W-4	

Feature types provided

EGN Feature classification

→ 8 classes and 27 sub-classes

- (1) Countries, administrative units and other areas,
- (2) Populated Places,
- (3) Non-residential structures and buildings,
- (4) Transport and telecommunication features,
- (5) Terrain features,
- (6) Hydrographic features,
- (7) Conservation areas,
- (8) Miscellaneous

→ essentially for the purpose of query-filtering

(defined by EGN Reference Group)

Code	Feature Type	Short Definition	Feature Type Examples			
1	COUNTRIES, ADMINISTRATIVE UNITS AND OTHER AREAS	Countries, territorial units of a country for administrative purposes and other man-made areas				
	1.1	Country	Country of Europe			
	1.2	Administrative units	Territorial units of every country for statistics and administrative purposes <i>Including:</i> Nomenclature of Territorial Units for Statistics in EU (NUTS 1, 2 and 3)	länder (Germany) autonomous region (Spain) province		
			Local Administrative Units (LAU 1 and 2) Other administrative units	municipality		
1.3	Other non-administrative units	Other type of man-made areas like economic, cultural, linguistic or tourist areas				
2	POPULATED PLACES	Buildings for housing of any category like cities, towns, villages, etc.				
	2.1	Administrative capitals	Populated places with capital status <i>Including:</i> Administrative capitals of NUTS 1, 2 and 3	capital of country, autonomous region (Spain), province		
			Administrative capitals of LAU 1 and 2 Other administrative capitals	capital of municipality		
			2.2	Other populated places	Populated places without administrative status <i>Including:</i> Cities, towns, villages, hamlets... Parts of them	city, town, village, hamlet hamlets... neighborhood
						11



Brazil - Feature codes?



Definitions

SOURCE: www.nga.mil

- Geographical Names INDEX
- Feedback

GIS Compatible Output Format

Field Name	Field Description	Field Type	Default Selection
RC	Region Font Code. A code that determines the character mapping used in the SHORT_FORM, Generic, and Full_Name fields (see Character Set Conversion Table [Char_Sets_20060703.pdf],: 1 = Americas/Western Europe; 2 = Eastern Europe; 3 = Africa/Middle East; 4 = Russia/ Central Asia; 5 = Asia/Pacific; 6 = Vietnam.	number	Y
UFI	Unique Feature Identifier. A number which uniquely identifies a Geoname feature.	number	Y
UNI	Unique Name Identifier. A number which uniquely identifies a name.	number	Y
LAT	Latitude of the Geoname feature in \pm decimal degrees; DD; (\pm dd.dd...): no sign (+) = North; 05°03'09"E » 5.0525 negative sign (-) = South; 05°03'09"S » -5.0525	number	Y
LONG	Longitude of the feature in \pm decimal degrees; DD; (\pm dd.dd...): no sign (+) = East; negative sign (-) = West.	number	Y
DMS_LAT	Latitude of the Geoname feature in \pm degrees, minutes, and seconds; DMS; (\pm ddmmss): no sign (+) = North; negative sign (-) = South.	number	Y

Content

First steps before processing the names collected

- Structure for storing the data, maintaining records
- Feature categories/classification

Demonstration of the processing of names collected

- Spreadsheet or Database?
- Create main tables and enter names into a database
- Database queries and reports (e.g. gazetteers as printouts)

Hands on experience



1) Spreadsheet (software: e.g. Microsoft Excel)

advantage:

extended processing capabilities

disadvantage:

digital processing limited to operations within the spreadsheet

Name	Feature Designation Name	Latitude	Longitude
Rio de Janeiro	Populated Place	-22,92	-43,38
Rio Iguaçu	Stream	-22,74	-43,25
Chácara Entrerios	Mountain	-22,48	-43,05
Petrópolis	Administrative unit	-22,48	-43,18

[demo3.xls](#)

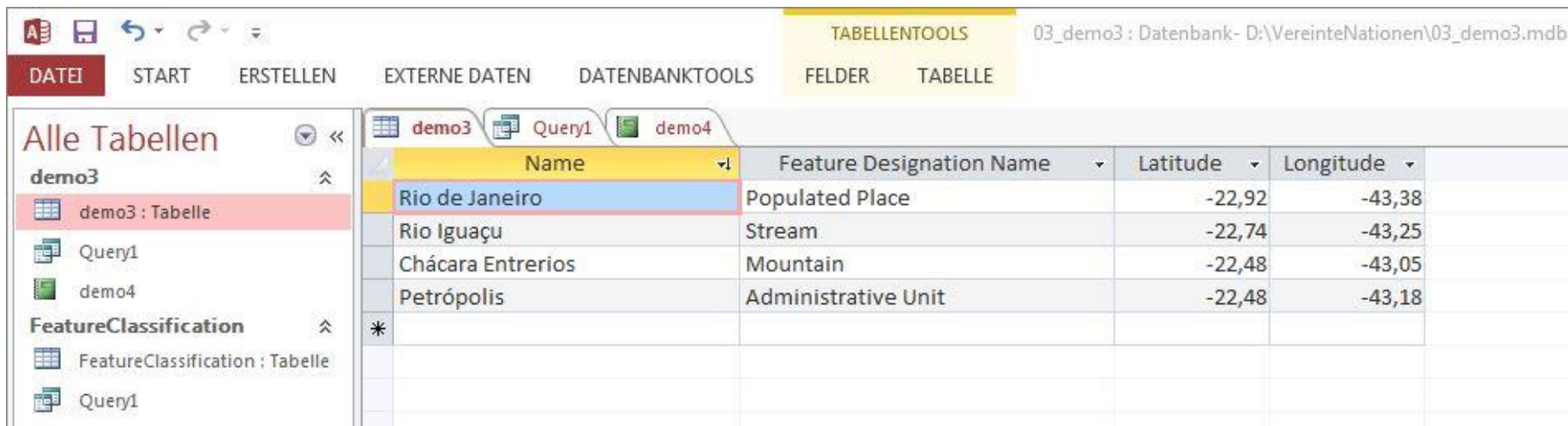
2) Database (software: e.g. Microsoft Access)

advantages:

data can be connected with other databases,
complex processing capabilities

disadvantages:

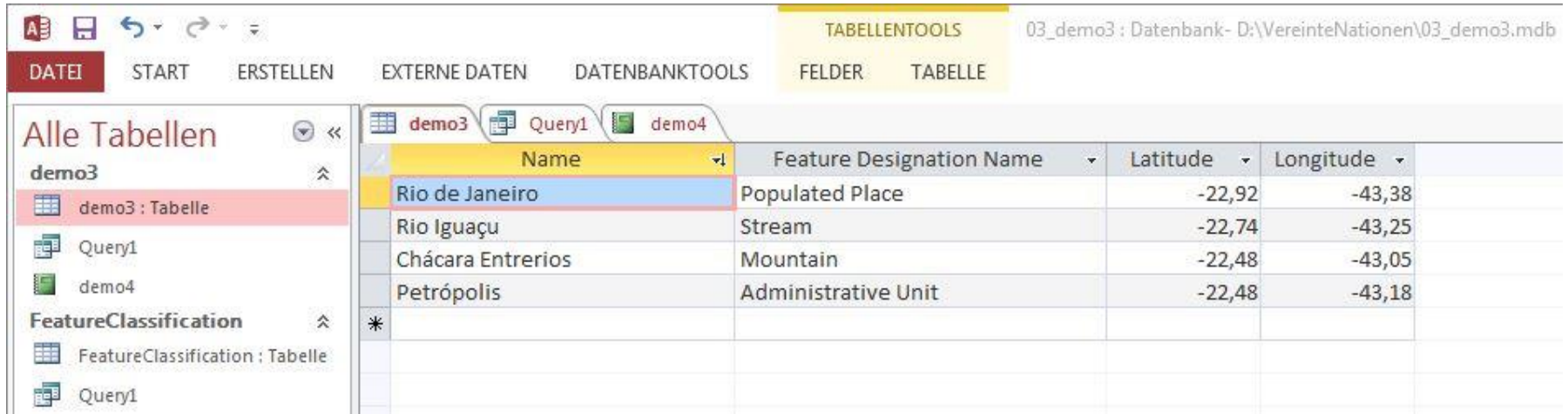
some programming and/or user skills required



The screenshot shows the Microsoft Access interface. The title bar indicates the file path: 03_demo3 : Datenbank- D:\VereinteNationen\03_demo3.mdb. The ribbon includes 'DATEI', 'START', 'ERSTELLEN', 'EXTERNE DATEN', 'DATENBANKTOOLS', 'FELDER', and 'TABELLE'. The left pane shows 'Alle Tabellen' with a tree view containing 'demo3', 'Query1', and 'demo4'. Under 'demo3', there is a table 'demo3 : Tabelle'. The main window displays a table with the following data:

Name	Feature Designation Name	Latitude	Longitude
Rio de Janeiro	Populated Place	-22,92	-43,38
Rio Iguaçu	Stream	-22,74	-43,25
Chácara Entrerios	Mountain	-22,48	-43,05
Petrópolis	Administrative Unit	-22,48	-43,18
*			

Create main tables and enter names into a database with MS Access

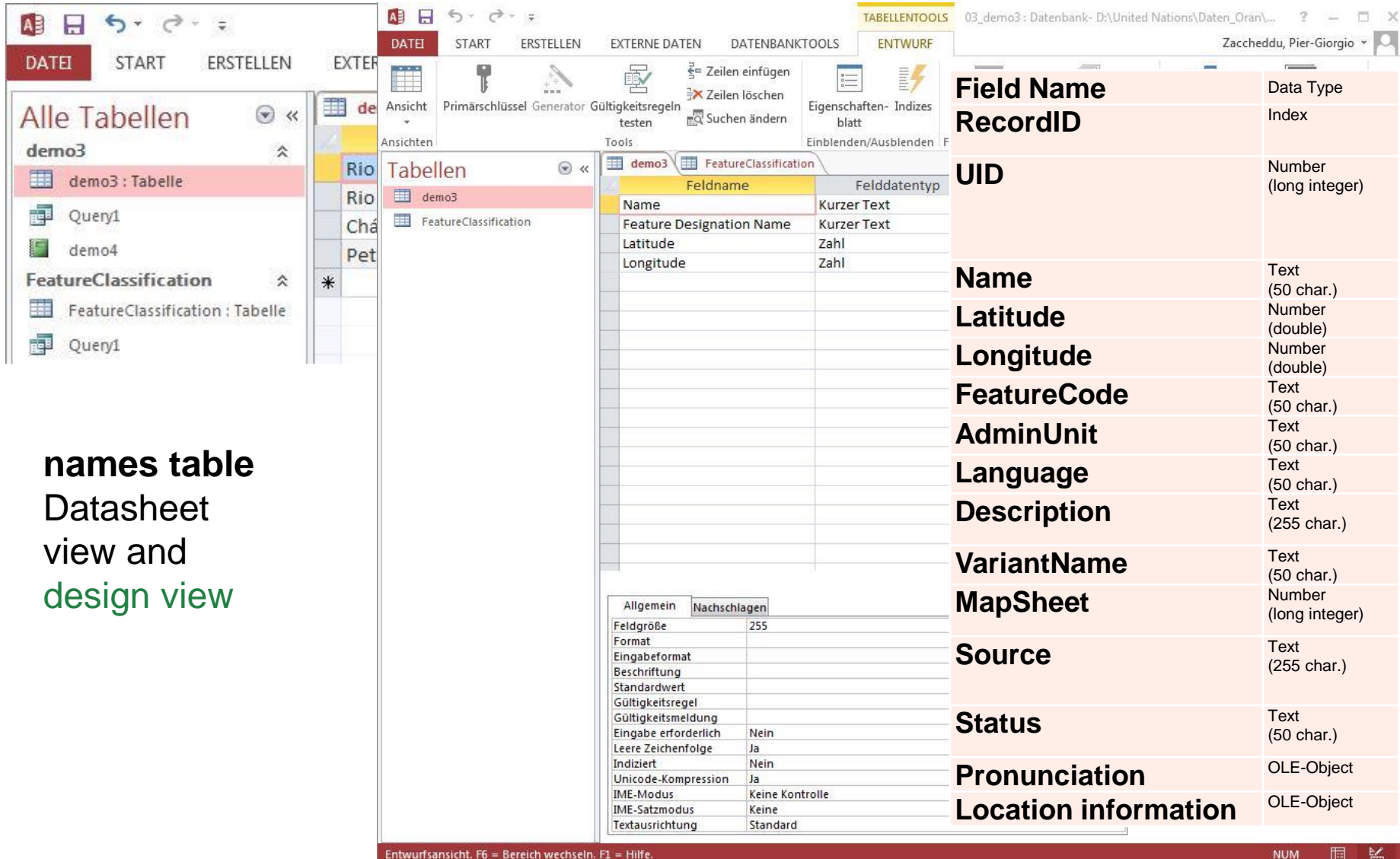


03_demo3 : Datenbank- D:\VereinteNationen\03_demo3.mdb

Name	Feature Designation Name	Latitude	Longitude
Rio de Janeiro	Populated Place	-22,92	-43,38
Rio Iguaçu	Stream	-22,74	-43,25
Chácara Entrerios	Mountain	-22,48	-43,05
Petrópolis	Administrative Unit	-22,48	-43,18
*			

names table
Datashheet
view and
design view

Create main tables and enter names into a database with MS Access



The screenshot shows the Microsoft Access interface. On the left, the 'Alle Tabellen' (All Tables) pane lists 'demo3', 'Query1', 'demo4', 'FeatureClassification', and 'FeatureClassification : Tabelle'. The main window displays the 'FeatureClassification' table in design view. The table has the following fields:

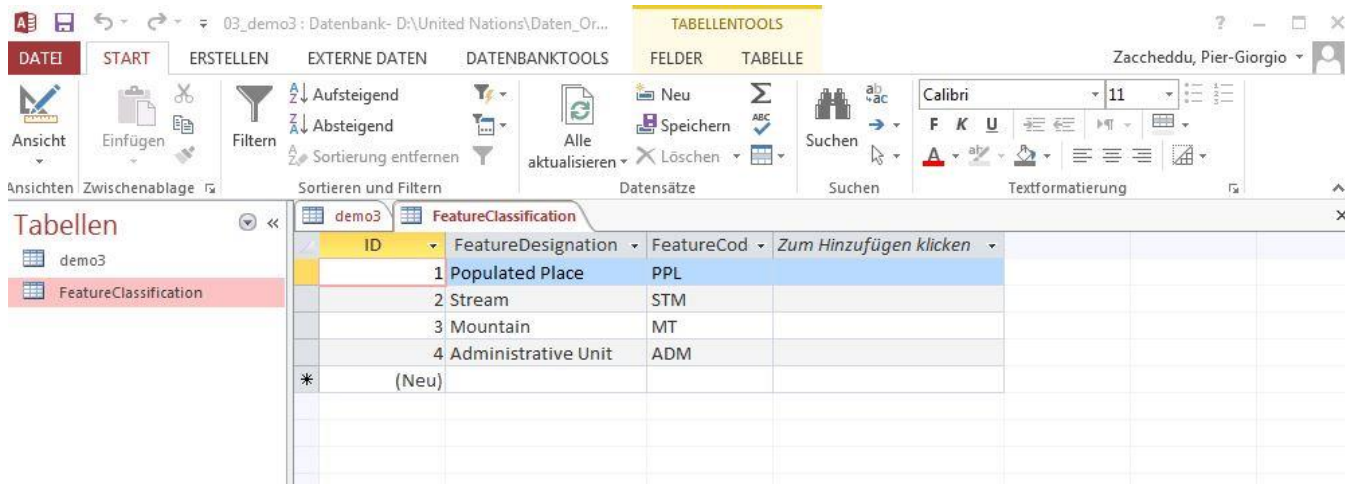
Feldname	Felddatatype
Name	Kurzer Text
Feature Designation Name	Kurzer Text
Latitude	Zahl
Longitude	Zahl

Below the table design, the 'Allgemein' (General) tab is active, showing various field properties such as 'Feldgröße' (Field Size) set to 255, 'Format', 'Eingabeformat' (Input Format), 'Beschriftung' (Caption), 'Standardwert' (Default Value), 'Gültigkeitsregel' (Validation Rule), 'Gültigkeitsmeldung' (Validation Text), 'Eingabe erforderlich' (Required), 'Leere Zeichenfolge' (Allow Zero Values), 'Indiziert' (Indexed), 'Unicode-Kompression' (Unicode Compression), 'IME-Modus' (IME Mode), 'IME-Satzmodus' (IME Sentence Mode), and 'Textausrichtung' (Text Alignment).

Field Name	Data Type
RecordID	Index
UID	Number (long integer)
Name	Text (50 char.)
Latitude	Number (double)
Longitude	Number (double)
FeatureCode	Text (50 char.)
AdminUnit	Text (50 char.)
Language	Text (50 char.)
Description	Text (255 char.)
VariantName	Text (50 char.)
MapSheet	Number (long integer)
Source	Text (255 char.)
Status	Text (50 char.)
Pronunciation	OLE-Object
Location information	OLE-Object

names table
Datasheet
view and
design view

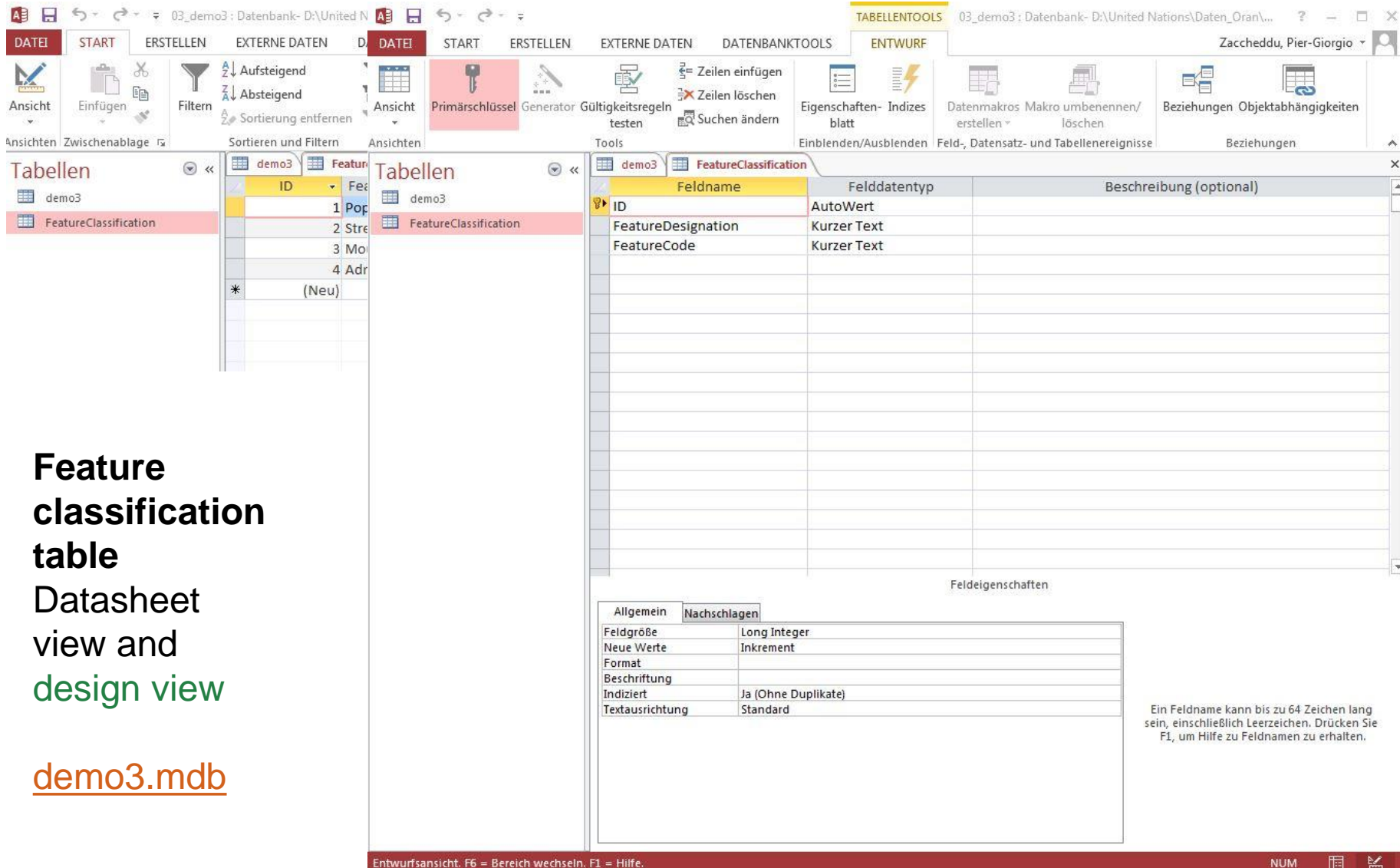
Create main tables and enter names into a database with MS Access



Feature classification table

Datasheet
view and
design view

Create main tables and enter names into a database with MS Access



The screenshot shows the Microsoft Access interface with the 'FeatureClassification' table in design view. The table structure is as follows:

Feldname	Felddatentyp	Beschreibung (optional)
ID	AutoWert	
FeatureDesignation	Kurzer Text	
FeatureCode	Kurzer Text	

The 'Allgemein' tab of the 'Nachschlagen' property sheet is visible at the bottom, showing the following properties:

Feldgröße	Long Integer
Neue Werte	Inkrement
Format	
Beschriftung	
Indiziert	Ja (Ohne Duplikate)
Textausrichtung	Standard

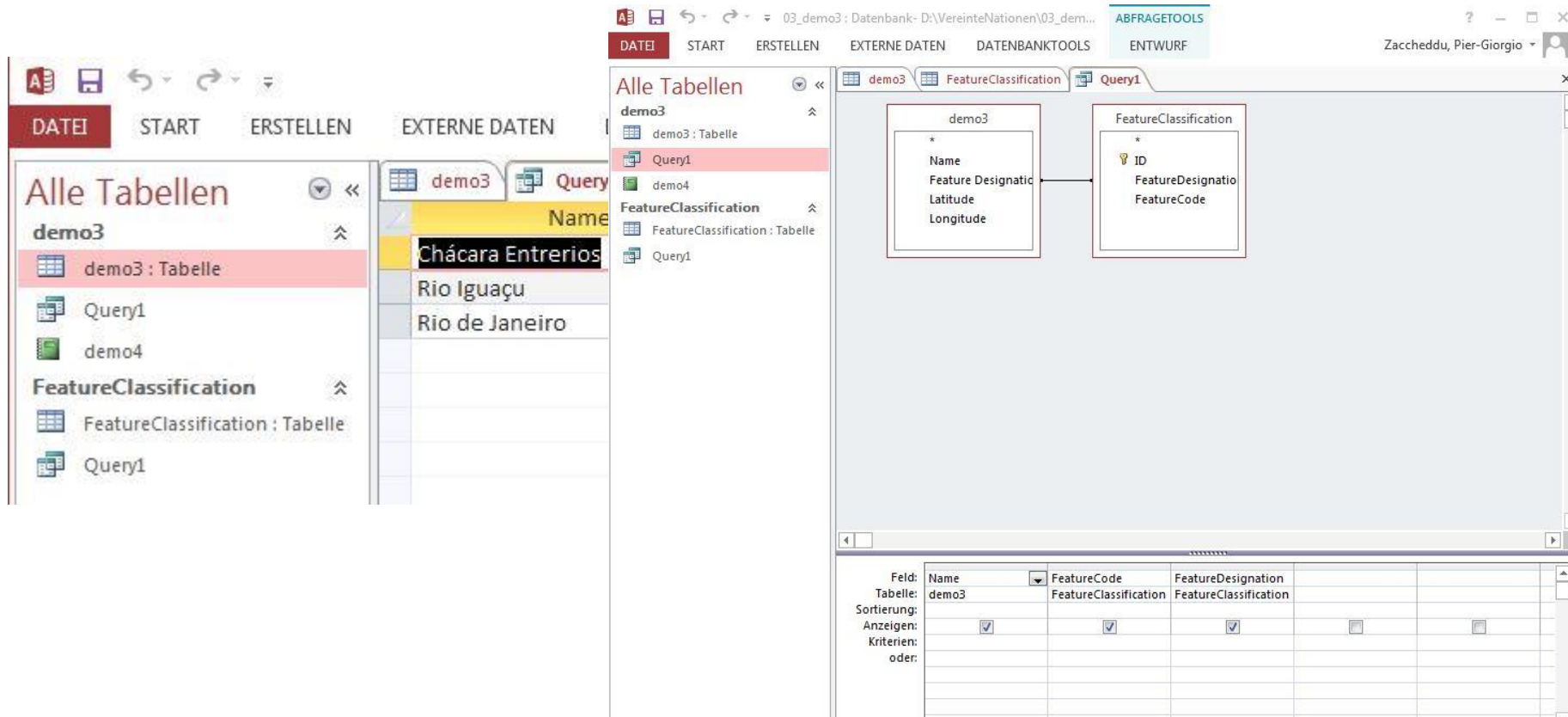
Ein Feldname kann bis zu 64 Zeichen lang sein, einschließlich Leerzeichen. Drücken Sie F1, um Hilfe zu Feldnamen zu erhalten.

**Feature
classification
table**

Datasheet
view and
design view

[demo3.mdb](#)

Database queries and reports with MS Access (e.g. gazetteers)



The screenshot displays the Microsoft Access interface. On the left, the 'Alle Tabellen' (All Tables) pane shows a list of tables: 'demo3 : Tabelle', 'Query1', 'demo4', 'FeatureClassification : Tabelle', and another 'Query1'. The main window shows a query named 'Query1' with a diagram illustrating a join between 'demo3' and 'FeatureClassification'. The 'demo3' table has fields: Name, Feature Designation, Latitude, and Longitude. The 'FeatureClassification' table has fields: ID (primary key), FeatureDesignation, and FeatureCode. A line connects the 'Feature Designation' field in 'demo3' to the 'FeatureDesignation' field in 'FeatureClassification'. Below the diagram, a data grid shows the results of the query:

Feld:	Name	FeatureCode	FeatureDesignation		
Tabellen:	demo3	FeatureClassification	FeatureClassification		
Sortierung:					
Anzeigen:	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kriterien:					
oder:					

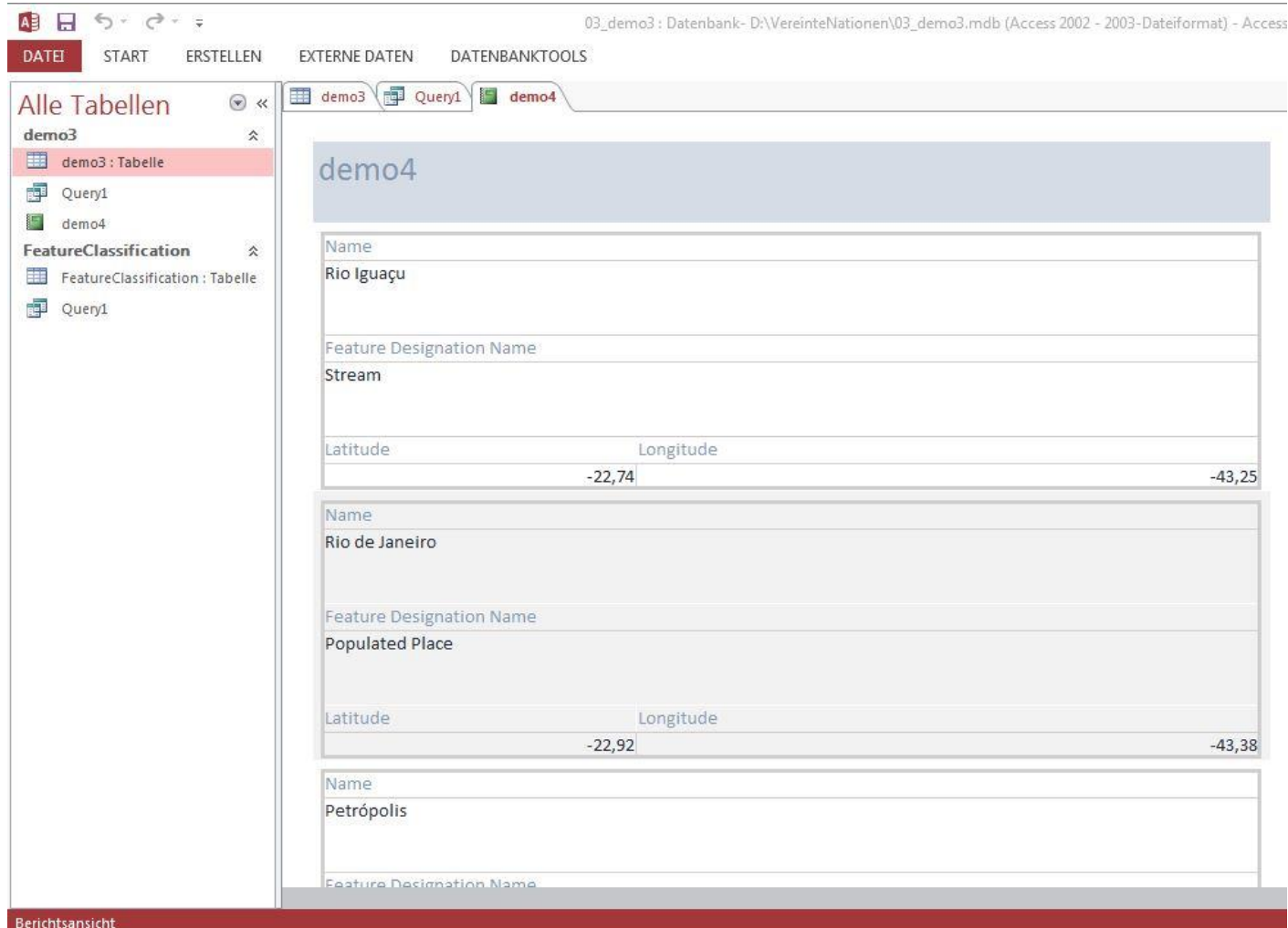
The main advantage of database **queries** is that you can join information from different tables in the database, i.e. here **joining the names table and the feature classification**

Database queries and reports with MS Access (e.g. gazetteers)

The **report tool**
can be used to
print the data in a
layout defined by
the creator.

→ **gazetteer**

It can build upon
a query or a table
and can be
created in
different designs.



03_demo3 : Datenbank- D:\VereinteNationen\03_demo3.mdb (Access 2002 - 2003-Dateiformat) - Access

DATEI START ERSTELLEN EXTERNE DATEN DATENBANKTOOLS

Alle Tabellen

demo3

- demo3 : Tabelle
- Query1
- demo4

FeatureClassification

- FeatureClassification : Tabelle
- Query1

demo4

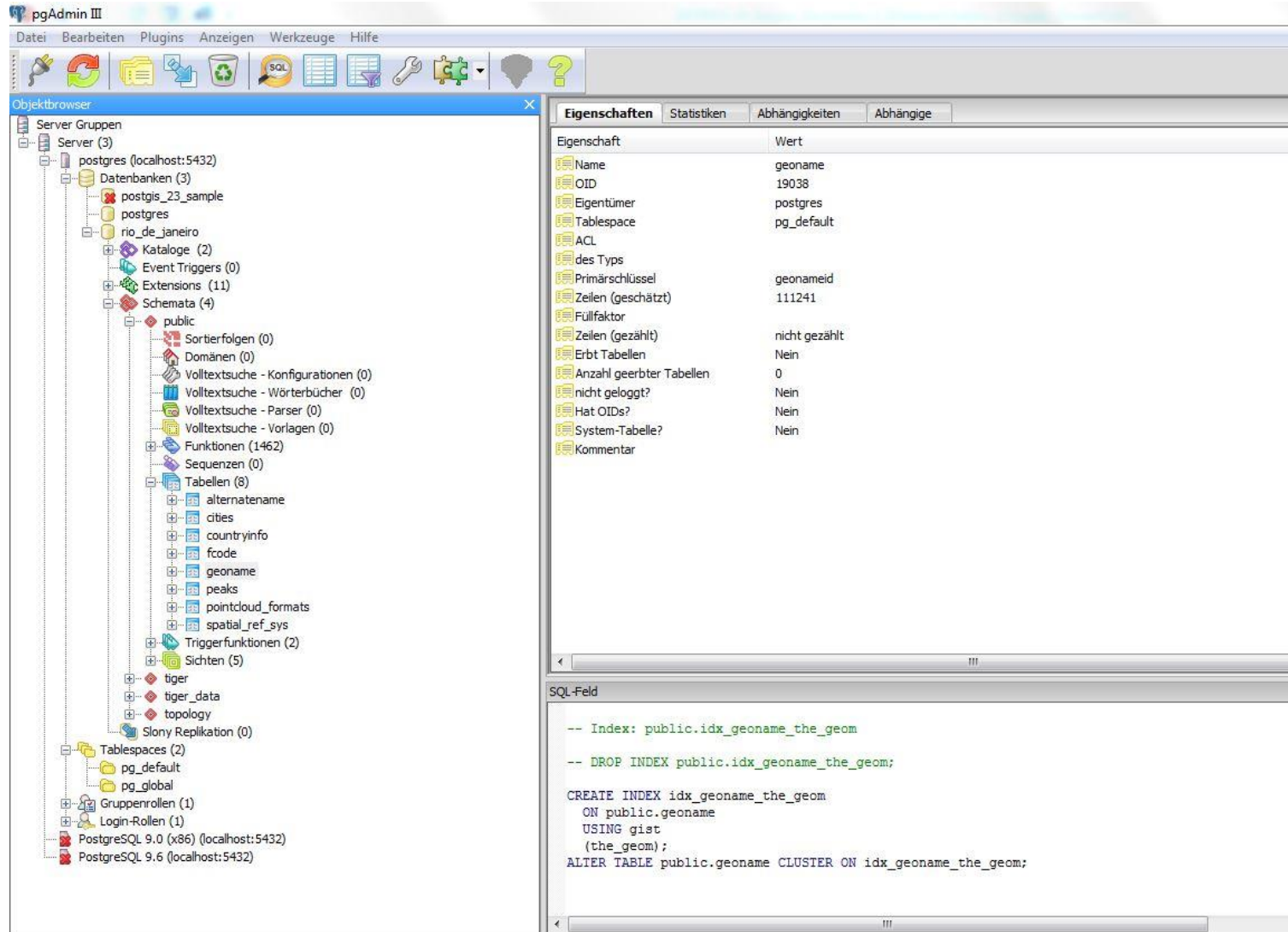
Name	Feature Designation Name	Latitude	Longitude
Rio Iguazu	Stream	-22,74	-43,25
Rio de Janeiro	Populated Place	-22,92	-43,38
Petrópolis			

Berichtsansicht

Create main tables and enter names into a database with PostgreSQL/PostGIS

Definition and creation of a sql-database with the tool pgAdmin

soft- and hardware skills are needed!



The screenshot shows the pgAdmin III interface. On the left, the 'Objektbrowser' (Object Browser) displays a tree view of the database structure. The 'public' schema is expanded, showing various objects including tables. The 'geoname' table is visible under the 'Tabellen' (Tables) category.

On the right, the 'Eigenschaften' (Properties) tab is active, showing the properties of the selected 'geoname' table. The properties are as follows:

Eigenschaft	Wert
Name	geoname
OID	19038
Eigentümer	postgres
Tablespace	pg_default
ACL	
des Typs	
Primärschlüssel	geonameid
Zeilen (geschätzt)	111241
Fullfaktor	
Zeilen (gezählt)	nicht gezählt
Erbt Tabellen	Nein
Anzahl geerbter Tabellen	0
nicht geloggt?	Nein
Hat OIDs?	Nein
System-Tabelle?	Nein
Kommentar	

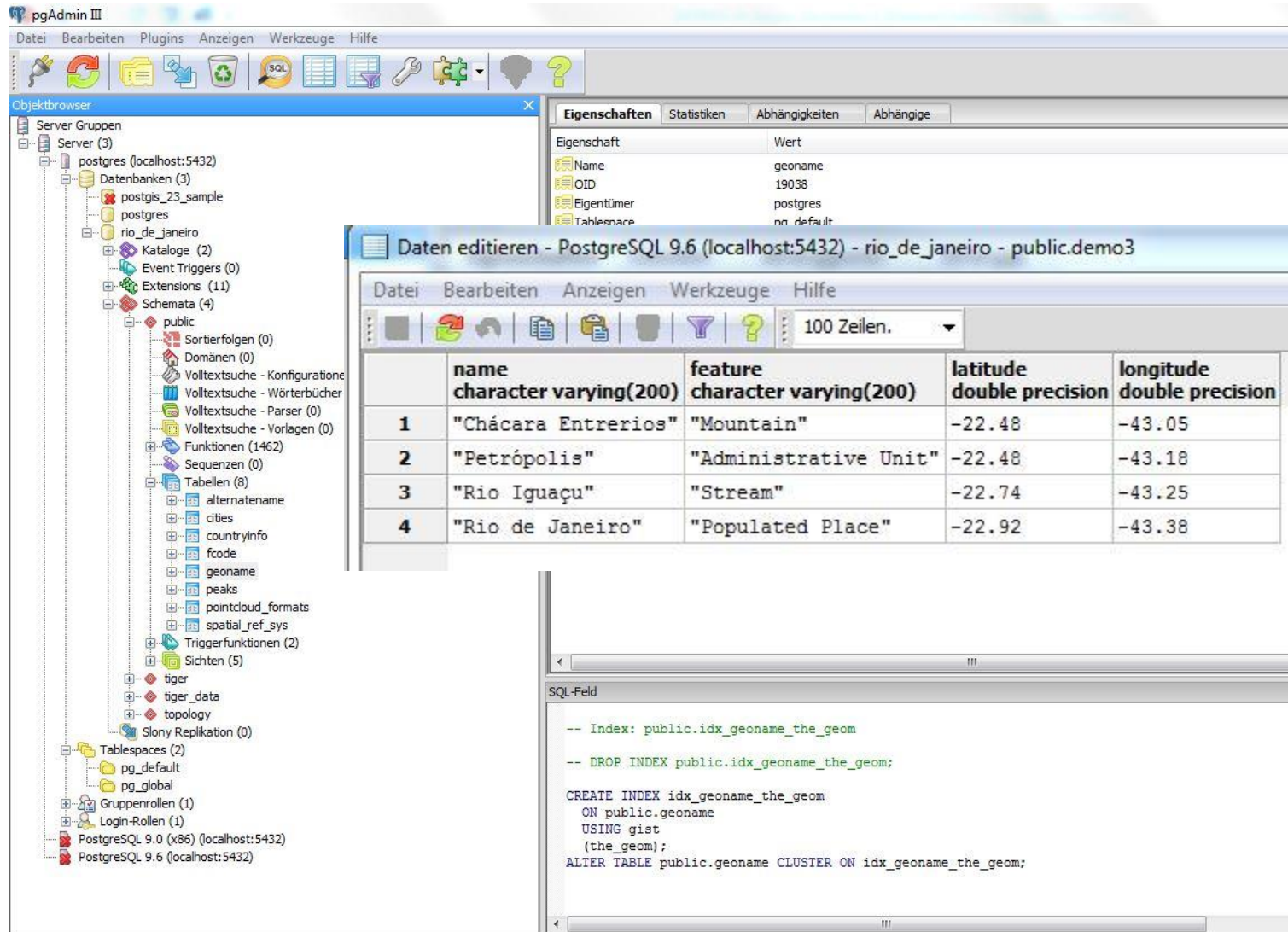
At the bottom of the interface, the 'SQL-Feld' (SQL Field) window contains the following SQL commands:

```
-- Index: public.idx_geoname_the_geom
-- DROP INDEX public.idx_geoname_the_geom;
CREATE INDEX idx_geoname_the_geom
ON public.geoname
USING gist
(the_geom);
ALTER TABLE public.geoname CLUSTER ON idx_geoname_the_geom;
```

Create main tables and enter names into a database with PostgreSQL/PostGIS

Definition and creation of a sql-database with the tool pgAdmin

soft- and hardware skills are needed!



The screenshot shows the pgAdmin III interface. On the left, the 'Objektbrowser' (Object Browser) displays a tree view of the database structure. The 'public' schema is expanded, showing various objects including tables like 'geoname'. On the right, the 'Eigenschaften' (Properties) window shows details for the 'geoname' table, including its name, OID (19038), owner (postgres), and tablespace (pg_default).

In the foreground, the 'Daten editieren' (Edit Data) window for the 'public.demo3' table is open. It displays a table with 4 rows and 5 columns. The columns are: 'name' (character varying(200)), 'feature' (character varying(200)), 'latitude' (double precision), and 'longitude' (double precision). The rows contain data for 'Chácara Entrerios', 'Petrópolis', 'Rio Iguaçu', and 'Rio de Janeiro'.

Below the table editor, the 'SQL-Feld' (SQL Field) window shows the following SQL commands:

```
-- Index: public.idx_geoname_the_geom
-- DROP INDEX public.idx_geoname_the_geom;
CREATE INDEX idx_geoname_the_geom
ON public.geoname
USING gist
(the_geom);
ALTER TABLE public.geoname CLUSTER ON idx_geoname_the_geom;
```



Databases with MS Access and PostgreSQL/PostGIS

PostgreSQL/PostGIS

The same result with both database tools!!!

MS Access

Daten editieren - PostgreSQL 9.6 (localhost:5432) - rio_de_janeiro - public.demo3

Datei Bearbeiten Anzeigen Werkzeuge Hilfe

100 Zeilen.

	name character varying(200)	feature character varying(200)	latitude double precision	longitude double precision
1	"Chácara Entrerios"	"Mountain"	-22.48	-43.05
2	"Petrópolis"	"Administrative Unit"	-22.48	-43.18
3	"Rio Iguaçu"	"Stream"	-22.74	-43.25
4	"Rio de Janeiro"	"Populated Place"	-22.92	-43.38

TABELLENTOOLS 03_demo3 : Datenbank - D:\VereinteNationen\03_demo3.mdb

DATEI START ERSTELLEN EXTERNE DATEN DATENBANKTOOLS FELDER TABELLE

Alle Tabellen

demo3

demo3 : Tabelle

Query1

demo4

FeatureClassification

FeatureClassification : Tabelle

Query1

Name	Feature Designation Name	Latitude	Longitude
Rio de Janeiro	Populated Place	-22,92	-43,38
Rio Iguaçu	Stream	-22,74	-43,25
Chácara Entrerios	Mountain	-22,48	-43,05
Petrópolis	Administrative Unit	-22,48	-43,18
*			

Content

First steps before processing the names collected

- Structure for storing the data, maintaining records
- Feature categories/classification

Demonstration of the processing of names collected

- Spreadsheet or Database?
- Create main tables and enter names into a database
- Database queries and reports (e.g. gazetteers as printouts)

Hands on experience



Training Course on Toponymy.

Creating a Geographical Names Database and producing Gazetteers and Maps

– Exercise –

Pier-Giorgio Zaccheddu, Andreas Illert

Knowledge:

Steps in designing a database

- Determine the purpose of your database.
- Determine the tables you need in the database.
- Determine the fields you need in the tables.
- Identify fields with unique values.
- Determine the relationships between tables.
- Refine your design.
- Add data and create other database objects.

Scope of the Exercise:

1. (a) create main tables for a names database or
(b) use of a predefined names database
2. insert the data from the field collection into a database
3. print a gazetteer from the database
4. create a map with the geographical names in a Geographical Information System (GIS)
5. Publish the names data using Google Maps/Earth

Thank you for your attention!