

XML Projekt

für Datenbanksysteme ILV 2020

an der FH Wiener Neustadt

Gruppe 4 (Brandstätter, Forsthuber, Jernej)

Aufgabenstellung:

Each group chooses a dataset from Open Data Austria - <https://www.data.gv.at/> - and defines a use case (data based question)

- **Task 1:** Deploy the data from the file in an XML enabled database (MSSQL, Oracle, PostGre, ...) – **Jernej / Brandstätter**
- **Task 2:** Define an XML schema for the data (or a subset of interest) – **Jernej**
- **Task 3:** Produce via the SQL/XML support of your chosen RDBMS an XML version of it adhering to your schema – **Forsthuber**
- **Task 4:** Register your XML schema and afterwards store (validate) the related data in the native XML database exist - <http://exist-db.org/> - **Brandstätter**
- **Task 5:** answer your data based question using XQuery via API usage in a supported programming language – **Forsthuber / Brandstätter**

Upload your documentation (PDF between 10 and 15 pages) including

- detailed documentation of each task so somebody else could re-do it
- issues and solutions on the way

and source code all together in one zip file onto moodle!

Suche ein geeignetes Dataset über den Datenprovider Open Data Austria unter - <https://www.data.gv.at>.

Formuliere und definiere auf den Daten dementsprechende Use Cases, welches du mit Queries beantworten sollst.

Unser Team hat sich einen Datenset über das Impfverhalten der Stadt Linz ausgesucht. Dieses Datenset bietet eine Aufzeichnung von 20 Jahren von über 28 möglichen Impfungen. In Zuge der Analyse der Datenstruktur sowie des Dateninhalts haben wir folgende Use Cases definiert, die es zu beantworten gilt:

1. Wie hat sich das Impfverhalten über alle Impfmöglichkeiten generell entwickelt?
2. Wo liegt der absolute Höchstwert von Impfungen & wann begann eine mögliche Trendwende nach unten oder umgekehrt?

Github-Repo zum Projekt:

https://github.com/ThePeziBear/XML_Project_Open_Data_Austria

Task 1: Erstellung einer XML fähigen Datenbank – Auswahl Postgres

Um die von opendata heruntergeladenen Daten sinnvoll in eine Datenbank zu übertragen, wurden diese zuerst von einem wide auf ein long-Format umgewandelt.

Auszug aus dem Datenset:

Impfungen	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Diphtherie-Tetanus-Keuchhusten-Hib-Kinderlähmung-Hepatitis B	0	0	0	0	0	0	1186	1064	971	855	753	725	579	376	311	196	138	101	109	86	73	113
Diphtherie-Tetanus-Keuchhusten-Hib-Kinderlähmung	0	0	0	0	690	433	181	29	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Diphtherie-Tetanus-Keuchhusten-Kinderlähmung	0	0	0	0	0	927	241	8	0	76	0	575	801	512	775	822	1548	3362	3947	1138	1409	2399
Diphtherie-Tetanus-Keuchhusten-Hib	200	2134	2793	3787	1269	343	63	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hib (Haemophilus influenzae b)	3724	2263	1740	114	0	0	0	0	38	0	0	0	0	0	0	0	0	0	0	0	0	0
Diphtherie-Tetanus-Kinderlähmung	0	0	0	0	0	1708	5115	4498	6422	3706	3994	3410	2639	3232	1859	393	747	1471	2807	2716	1306	
Diphtherie-Tetanus-Keuchhusten	4214	2768	2155	112	0	0	0	0	0	11	2756	1693	818	1509	2835	1681	1173	1838	338	23	928	
Diphtherie-Tetanus	4876	4815	5626	5669	7826	4933	2277	1326	1150	1755	1555	1488	1175	1074	999	796	2012	960	428	370	403	291
Diphtherie	0	0	0	0	0	0	0	0	0	0	0	112	69	43	45	14	0	0	0	0	0	0
Tetanus	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Kinderlähmung (Poliomyelitis)	9002	11862	10639	10476	7421	7835	1688	2718	1636	2713	1320	1673	1386	1278	1253	557	443	573	1443	1010	577	1274
Hib-Hepatitis B	0	0	0	0	0	583	162	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Hepatitis A+B	0	0	0	0	0	0	0	0	0	0	0	1574	1758	1632	1418	1322	1023	741	565	557	732	1178
Hepatitis A	0	0	0	0	0	0	0	0	0	0	0	491	602	497	452	408	353	260	512	392	375	802
Hepatitis B	0	0	0	4054	6814	5917	5374	5283	5041	4579	3994	5716	4126	2959	3238	2781	1899	1384	1599	1937	1527	1287
Masern-Mumps-Röteln	2270	1813	1897	2667	3630	2960	2695	2481	2673	3843	3397	2524	1752	1380	558	352	565	619	486	828	1022	3435
Röteln	901	806	845	0	0	4	505	469	438	0	0	0	0	0	0	0	0	0	0	0	0	0
Masern-Mumps	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Mumps	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Tuberkulose	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Rotaviren	0	0	0	0	0	0	0	0	0	0	0	0	0	8	7	7	8	14	22	17	12	63
FSME	18682	19850	20207	19763	20650	18297	19392	18282	16727	9588	12425	16340	16195	16530	11712	11107	15769	16433	15114	13346	12144	11328
Grippe	0	969	1045	1240	2114	3660	3949	4050	5500	5750	5572	5160	3948	4423	5078	3259	3236	2620	3084	2352	2469	2336
H1N1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2378	85	0	0	0	0	714	0
Pneumokokken	0	0	0	0	0	0	0	0	0	222	231	149	276	121	158	35	44	67	65	264	64	441
Meningokokken	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1427	2031	2098	1192	2320
HPV	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2567	2677

Transformieren mittels R-Code des Datenset für das Einlesen in Postgres:

```
# File: impf_recode.R
# Date: 3.03.2020
# Copyright (C) 2020
# Description: R-Script zum Einlesen und transformieren der Impfdaten
```

```
library(reshape2)
```

```
# Einlesen
impfdat <- read.csv2("./Impf.csv", fileEncoding = "Windows-1252")
# von wide auf long
impfdat_long <- melt(impfdat)
```

```
# entfernen von X im Spaltennamen
impfdat_long$variable <- as.character(impfdat_long$variable)
impfdat_long$variable <- gsub("X", "", impfdat_long$variable)
```

```
# Schreiben ohne Index
write.csv(impfdat_long, "./Impfung_Recode.csv", row.names = FALSE)
```

Output nach Transformation des Datensets in eine neue csv-Datei Impfung_Recode.csv:

	A	B	C
1	Impfungen	variable	value
2	Diphtherie-Tetanus-Keuchhusten-Hib-Kinderlähmung-Hepatitis B	1995	0
3	Diphtherie-Tetanus-Keuchhusten-Hib-Kinderlähmung	1995	0
4	Diphtherie-Tetanus-Keuchhusten-Kinderlähmung	1995	0
5	Diphtherie-Tetanus-Keuchhusten-Hib	1995	200
6	Hib (Haemophilus influenzae b)	1995	3724
7	Diphtherie-Tetanus-Kinderlähmung	1995	0
8	Diphtherie-Tetanus-Keuchhusten	1995	4214
9	Diphtherie-Tetanus	1995	4876
10	Diphtherie	1995	0
11	Tetanus	1995	0
12	Kinderlähmung (Poliomyelitis)	1995	9002
13	Hib-Hepatitis B	1995	0
14	Hepatitis A+B	1995	0

Die Datei wurde in utf-8 kodiert. Wie man sieht gibt es Umlaute, die später noch Probleme verursachen werden. Wir haben in zwei unterschiedlichen Betriebssystemen gearbeitet, Ubuntu und Windows. Unter Windows gibt es die Enkodierung Win-1252 unter Linux ist UTF-8 häufig Standard.

Nach erfolgreichem Setup der Installationsdatei von Postgres (pgAdmin4) wurde über die Konsole eine entsprechende Database und ein Table erstellt:

```
#Einloggen mit Benutzer 'postgres'
psql -U postgres
(Für Zugriff mit der Console muss die Systemumgebungsvariable Path für PSQL erweitert werden.
(C:/Programfiles/Postgres/bin)

#Passwort für Benutzer
Test

#Datenbank anlegen
CREATE DATABASE XML WITH ENCODING 'UTF8' LC_COLLATE='German_Germany'
LC_CTYPE='German_Germany';
(Bei Fehlermeldung LC_COLLATE='German-Austria & LC_CTYPE='German_Austria)

#scope auf erstellte db legen
\c xml






#Tabelle impf anlegen
CREATE TABLE impf (id serial PRIMARY KEY, impfung character varying(100), jahr smallint, anzahl
Integer); (Für Zugriff von Postgres auf Datei muss der Zugriff unter Dateieigenschaften/Sicherheit ein
neuer Benutzer "jeder" angelegt werden mit Schreib & Leserechte)

#csv file in die tabelle impf laden
COPY impf(impfung, jahr, anzahl) FROM 'C:/Users/Dell/Github/XML_Projekt/Impfung_Recode.csv'
DELIMITER ',' CSV Header;
```

#alle daten aus impf anzeigen

```
SELECT * FROM impf
```

Darstellung der Daten in pgAdmin4 unter Query Editor:

					
	id	impfung	jahr	anzahl	
	[PK] integer	character varying (100)	smallint	integer	
1	1	Diphtherie-Tetanus-Keuchhust...	1995	0	
2	2	Diphtherie-Tetanus-Keuchhust...	1995	0	
3	3	Diphtherie-Tetanus-Keuchhust...	1995	0	
4	4	Diphtherie-Tetanus-Keuchhust...	1995	200	
5	5	Hib (Haemophilus influenzae b)	1995	3724	
6	6	Diphtherie-Tetanus-KinderlÃ¤r...	1995	0	
7	7	Diphtherie-Tetanus-Keuchhust...	1995	4214	
8	8	Diphtherie-Tetanus	1995	4876	
9	9	Diphtherie	1995	0	
10	10	Tetanus	1995	0	
11	11	KinderlÃ¤hmung (Poliomyeliti...	1995	9002	
12	12	Hib-Hepatitis B	1995	0	
13	13	Hepatitis A+B	1995	0	
14	14	Hepatitis A	1995	0	
15	15	Hepatitis B	1995	0	
16	16	Masern-Mumps-RÃ¶teln	1995	2270	
17	17	RÃ¶teln	1995	901	
18	18	Masern-Mumps	1995	0	
19	19	Mumps	1995	0	
20	20	Tuberkulose	1995	10	
21	21	Rotaviren	1995	0	
22	22	FSME	1995	10000	

Abgesehen vom erfolgreichen Import fallen die seltsamen Sonderzeichen auf →
Enkodierungsproblem.

Task 2: Definition eines XML-Schema für die Daten

```
<!ELEMENT root (impf+)>
<!ELEMENT impf (impfung, jahr, anzahl)>
<!ELEMENT impfung (#PCDATA)>
<!ELEMENT jahr (#PCDATA)>
<!ELEMENT anzahl (#PCDATA)>
```

Das Schema wurde als Validierung_Datenschema.dtd gespeichert. Nach einigen missglückten Versuchen der Validierung, wurde es etwas vereinfacht (Header entfernt).

Task 3: Erstelle über den SQL/XML-Support der RDBMS eine XML-Version der Daten die dem XML-Schema entspricht.

Query über Query Editor pgAdmin4:

```
select xmlelement(name impf, xmlelement(name impfung, impfung), xmlelement(name jahr, jahr),
xmlelement(name anzahl, anzahl)) FROM impf;
```

Ausgabe der XML-Datei:

```
\copy (select table_to_xml('impfung', true, false, '')) TO
'/home/christian/wd/fhwn/db_systeme/XML_Project_Open_Data_Austria/result_neu3.xml';
```

Output der XML-Datei (mit etwas manueller Nachbesserung):

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE root SYSTEM "Validierung_Datenschema.dtd">

<root>

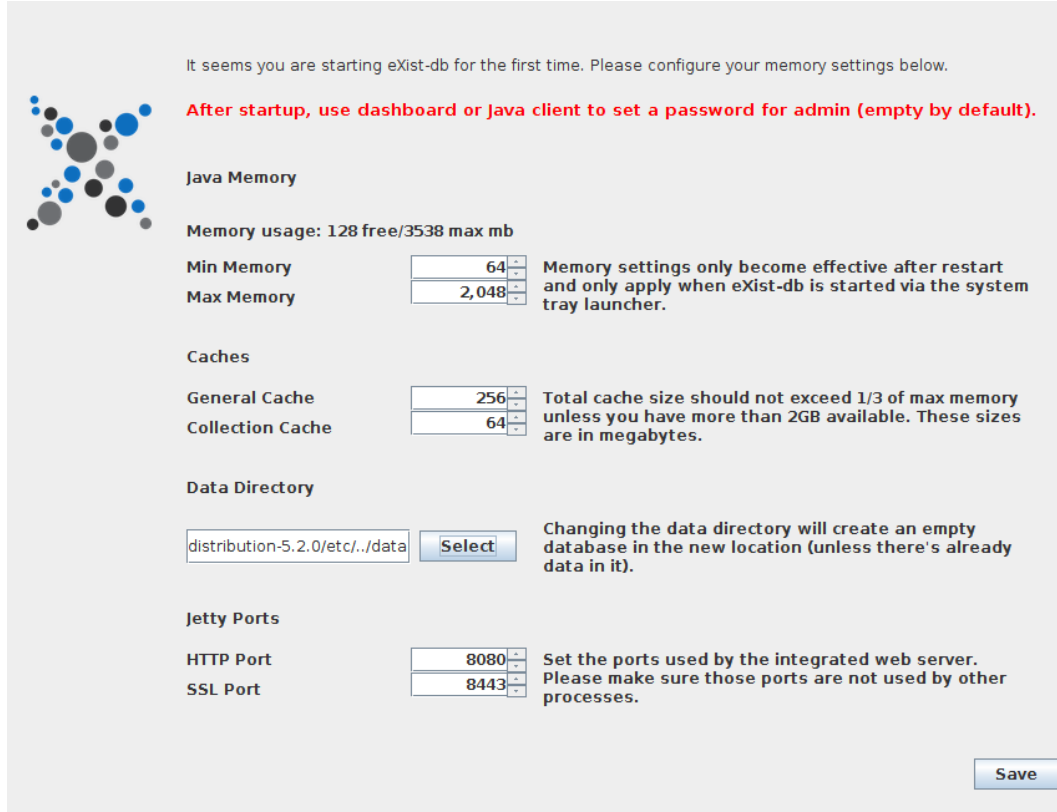
<impf><impfung>Diphtherie-Tetanus-Keuchhusten-Hib-Kinderlaehmung-Hepatitis B </impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Diphtherie-Tetanus-Keuchhusten-Hib-Kinderlaehmung</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Diphtherie-Tetanus-Keuchhusten-Kinderlaehmung</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Diphtherie-Tetanus-Keuchhusten-Hib</impfung><jahr>1995</jahr><anzahl>200</anzahl></impf>
<impf><impfung>Hib (Haemophilus influenzae b)</impfung><jahr>1995</jahr><anzahl>3724</anzahl></impf>
<impf><impfung>Diphtherie-Tetanus-Kinderlaehmung </impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Diphtherie-Tetanus-Keuchhusten</impfung><jahr>1995</jahr><anzahl>4214</anzahl></impf>
<impf><impfung>Diphtherie-Tetanus</impfung><jahr>1995</jahr><anzahl>4876</anzahl></impf>
<impf><impfung>Diphtherie</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Tetanus</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Kinderlaehmung (Poliomyelitis)</impfung><jahr>1995</jahr><anzahl>9002</anzahl></impf>
<impf><impfung>Hib-Hepatitis B </impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Hepatitis A+B</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Hepatitis A</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Hepatitis B</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Masern-Mumps Roeteln</impfung><jahr>1995</jahr><anzahl>2270</anzahl></impf>
<impf><impfung>Roeteln</impfung><jahr>1995</jahr><anzahl>901</anzahl></impf>
<impf><impfung>Masern-Mumps</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Mumps</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>Tuberkulose</impfung><jahr>1995</jahr><anzahl>10</anzahl></impf>
<impf><impfung>Rotaviren</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>
<impf><impfung>FSME</impfung><jahr>1995</jahr><anzahl>18682</anzahl></impf>
<impf><impfung>Grippe</impfung><jahr>1995</jahr><anzahl>0</anzahl></impf>

```

Da es immer noch Encoding-Probleme gab, wurden kurzerhand die Umlaute in ASCII-taugliche Formate ersetzt. Da die Validierung in der existdb auf „auto“ gesetzt wurde, musste das Validierungs-Schema mitgegeben werden. Außerdem wurde ein für den später notwendigen ExistDB-Import ein root-Tag hinzugefügt.

Task 4: Register your XML schema and afterwards store (validate) the related data in the native XML database exist - <http://exist-db.org/>

Installation Datenbank (Unix/Windows)



It seems you are starting eXist-db for the first time. Please configure your memory settings below.

After startup, use dashboard or java client to set a password for admin (empty by default).

Java Memory

Memory usage: 128 free/3538 max mb

Min Memory	<input type="text" value="64"/>	Memory settings only become effective after restart and only apply when eXist-db is started via the system tray launcher.
Max Memory	<input type="text" value="2,048"/>	

Caches

General Cache	<input type="text" value="256"/>	Total cache size should not exceed 1/3 of max memory unless you have more than 2GB available. These sizes are in megabytes.
Collection Cache	<input type="text" value="64"/>	

Data Directory

Changing the data directory will create an empty database in the new location (unless there's already data in it).

Jetty Ports

HTTP Port	<input type="text" value="8080"/>	Set the ports used by the integrated web server. Please make sure those ports are not used by other processes.
SSL Port	<input type="text" value="8443"/>	

Alle Einstellungen wurden auf Standard gesetzt.

Um das bearbeitete XML-File hochladen und (implizit) validieren zu können, war es notwendig:

1. Ein dtd-Schema zu erstellen (Siehe Task 2)
2. Die Datei conf.xml im Verzeichnis /etc zu ändern (validation mode = „auto“)

```
-->
<validation mode="auto">
  <!--
    Specify the location of one or more catalog files. Catalogs are
    used to resolve external entities in XML documents.

    "${WEBAPP_HOME}" and "${EXIST_HOME}" can be used as magic string.
  -->
  <entity-resolver>
    <catalog uri="${WEBAPP_HOME}/WEB-INF/catalog.xml"/>
  </entity-resolver>
</validation>
```


3. Die hochzuladende xml-Datei noch einmal zu modifizieren:

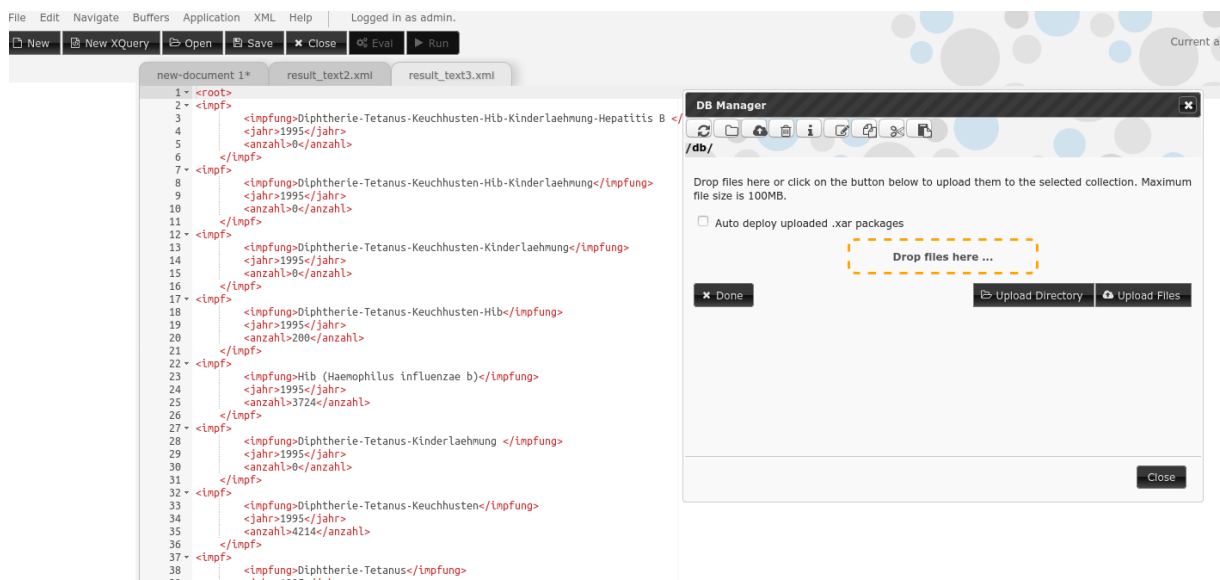
```
<!DOCTYPE root SYSTEM "Validierung_Datenschema.dtd">
```

Dabei wurde das Validierungs-Schema mitgegeben.

4. Die Datei catalog.xml im Verzeichnis `/etc/webapp/WEB-INF` zu editieren, damit das Schema gefunden wird:

```
<!-- NEU VON UNS -->
<system systemId="Validierung_Datenschema.dtd" uri="entities/Validierung_Datenschema.dtd"/>
```

5. Den eigentlichen Upload auszuführen.



Das ganze Prozedere war etwas Trial und Error, die Log-Datei exist.loc im Verzeichnis `./logs` war dabei sehr hilfreich. Das Encoding-Problem erschwerte die Sache etwas.

Wenn man die Validierung auf „Auto“ stellt muss man explizit das Schema und dessen Pfad mitgeben mittels `<system systemID="" uri="">`. Alternativ könnte man in der config-Datei „yes“ auswählen, aber wir hatten mit „auto“ erste Erfolge beim Import erzielt. Der Grund war, dass ohne Angabe des Schemas keine Validierung stattgefunden hat.

Task 5: answer your data based question using XQuery via API usage in a supported programming language

Als Programmiersprache der Wahl wurde Python gewählt. Als Entwicklungsumgebung wurde Anaconda3 mit Jupyter Notebook verwendet.

Die folgend aufgeführten Packages wurden importiert.

pyexistdb wird für die Schnittstelle zur exist Datenbank verwendet.

pandas wird zur Verwendung von DataFrames importiert.

lxml wird verwendet um die Daten aus der Datenbank verarbeiten zu können.

matplotlib wird für das Auswertungsdiagramm gebraucht.

```
from pyexistdb import db
import pandas as pd
from lxml import etree, objectify
import matplotlib.pyplot as plt
```

Mit folgendem Befehl wird die Verbindung zur Datenbank hergestellt. So kann über die Variable „exdb“ auf die Datenbank zugegriffen werden. Es wird die Adresse der Datenbank, der Benutzername und das entsprechende Passwort übergeben.

```
exdb = db.ExistDB("http://127.0.0.1:8080/exist", username = "admin", password = "")
```

Die Methode „QueryToTree“ schickt einen xquery Ausdruck an die Datenbank und gibt den erhaltenen xml-Baum aus und zurück. Die Abfrage wird in die <result> Tags gepackt, um zu vermeiden, dass die Datenbank nur 10 Ergebnisse liefert. Der Parameter „pretty_print“ wird True gesetzt, damit der xml-Baum durch den „tostring“ Befehl in eine lesbare Form umgewandelt wird.

```
def QueryToTree(query):
    result = exdb.query("<result>{" + query + "}</result>")
    print(etree.tostring(result.results[0], pretty_print=True).decode())
    return result.results[0]
```

Folgende Abfrage würde den Inhalt des Datenbankeintrags „result_text3.xml“ ausgeben.

```
QueryToTree("for $x in doc(\"result_text3.xml\") return $x")
```

Folgender Befehl beantwortet die oben in der Aufgabenstellung formulierten Fragen die wie folgt lauten:

Wie verhält sich die Gesamtzahl der Impfungen über die Jahre?

Wo liegt der absolute Höchstwert von Impfungen & wann begann eine mögliche Trendwende nach unten oder umgekehrt?

```
root = QueryToTree("for $x in doc(\"result_text3.xml\")/root/impf let $jahr := $x/jahr group by $jahr order by sum($x/jahr) return <impf><impfung>{data($jahr)}</impfung><impfgesamtzahl>{sum($x/anzahl)}</impfgesamtzahl></impf>")
```

Anschließend wird der Antwort-Baum zur besseren Übersicht in ein DataFrame umgewandelt und die Daten in einem Diagramm dargestellt.

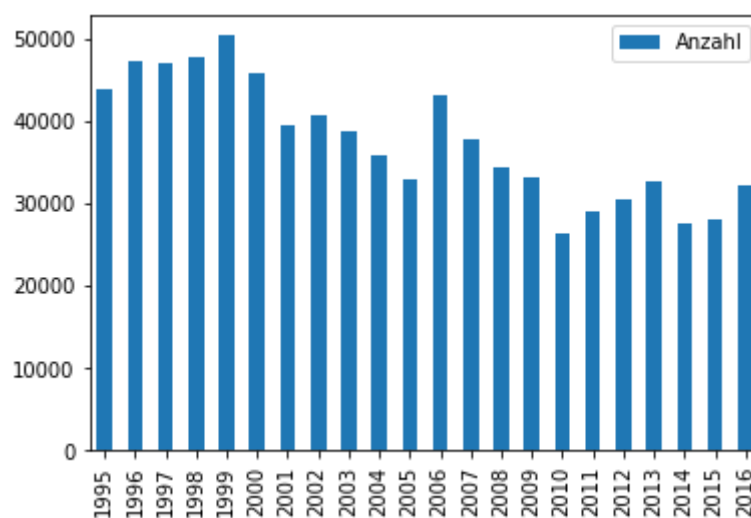
```
impfungen = [child.getchildren() for child in root.getchildren()]
jahre = [child[0].text for child in impfungen]
anzahl = [int(child[1].text) for child in impfungen]
df = pd.DataFrame(anzahl,jahre, columns=['Anzahl'])
df.plot(kind='bar')
plt.show
```

```
<result xmlns:exist="http://exist.sourceforge.net/NS/exist">
  <impf>
    <impfung>1995</impfung>
    <impfgesamtzahl>43879</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>1996</impfung>
    <impfgesamtzahl>47280</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>1997</impfung>
    <impfgesamtzahl>46947</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>1998</impfung>
    <impfgesamtzahl>47882</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>1999</impfung>
    <impfgesamtzahl>50414</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2000</impfung>
    <impfgesamtzahl>45892</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2001</impfung>
    <impfgesamtzahl>39421</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2002</impfung>
    <impfgesamtzahl>40833</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2003</impfung>
    <impfgesamtzahl>38672</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2004</impfung>
    <impfgesamtzahl>35803</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2005</impfung>
```

```

    <impfgesamtzahl>32964</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2006</impfung>
    <impfgesamtzahl>43277</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2007</impfung>
    <impfgesamtzahl>37770</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2008</impfung>
    <impfgesamtzahl>34290</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2009</impfung>
    <impfgesamtzahl>33123</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2010</impfung>
    <impfgesamtzahl>26435</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2011</impfung>
    <impfgesamtzahl>29112</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2012</impfung>
    <impfgesamtzahl>30481</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2013</impfung>
    <impfgesamtzahl>32714</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2014</impfung>
    <impfgesamtzahl>27540</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2015</impfung>
    <impfgesamtzahl>28019</impfgesamtzahl>
  </impf>
  <impf>
    <impfung>2016</impfung>
    <impfgesamtzahl>32178</impfgesamtzahl>
  </impf>
</result>

```



Dabei zeigt sich eine Abnahme der absoluten Anzahl der Impfungen in Linz. Es gab 1999 den Höchstwert. So eine Abnahme kann viele verschiedene Ursachen haben: es könnten etwa mehrere Impfungen zu einer zusammengefasst worden sein. Tatsächlich zeigt sich etwa, dass

zum Beispiel ab etwa 2000 eine Kombinationsimpfung „Diphtherie-Tetanus-Keuchhusten-Hib-Kinderlähmung-Hepatitis B“ vorhanden war, die es vorher nur ohne Hepatitis B gab, und vor 1999 auch noch ohne Kinderlähmung.

Eine weitere Ursache für die Abnahme könnte die demographische Veränderung der Bevölkerungsstruktur sein, d.h. es wurden weniger Kinder geboren. Das müsste man mit demographischen Daten abgleichen.

Es könnte auch daran liegen, dass Familien mit Kindern vermehrt aufs Land ziehen, statt in der Stadt zu bleiben. Um das zu verifizieren, könnte man die Geburtenanzahl gemeinsam mit den Schulgängern/Kindergartenkindern (idealerweise korrigiert um das Alter) in OÖ betrachten. Eine mögliche Frage wäre: sind die Schulgänger/Kindergartenkinder in Linz zurückgegangen und die Geburten gleichgeblieben? Das könnte unter Umständen diese Hypothese belegen. Man müsste natürlich auch bedenken, dass möglicherweise die Hausgeburten zurückgegangen sind, und auch die Kinder vom Land vermehrt in der Stadt geboren wurden. Daher sollte man die Schulbesucher/Kindergartenkinder in ganz OÖ betrachten.

Eine weitere Hypothese könnte die Zunahme der Impfgegnerschaft untersuchen: sollten sich die demographischen und impftechnischen Faktoren als vernachlässigbar erweisen (oder auch nicht), müsste man die einzelnen Impfstoffe einzeln betrachten. Vielleicht zeigen sich Abnahmen bei Impfungen, die Impfskeptikern besonders gefährlich erscheinen.

Hier noch eine abschließende Betrachtung nach Impfungen (mit R anhand der Rohdaten erstellt):

Etwa bei Meningokokken hat es 2010 einen ziemlichen Rückgang gegeben, was 2016 wieder deutlich anstieg. Auch in 2016 gab es einen starken Anstieg bei der Masern-Impfung. FSME, die deutlich häufigste Anzahl, wirkte einigermaßen konstant. Die Kinderlähmungsimpfung war bis etwa 2000 separiert und dürfte dann in die Diphtherie-Tetanus-Kinderlähmung-Impfung aufgenommen worden sein.

So auf den ersten oberflächlichen Blick ergeben sich die größten Unterschiede aufgrund der Impftechnik (Zusammenfassung zu einer Impfung), kleinere Unterschiede aufgrund von Saisonalen Trends (etwa 2016 die Masernepidemie). Insgesamt bleibt jedoch ein schwach abnehmender Trend.

Impfungen in Linz
von 1995 bis 2016

