

Kernmodule 3

Datavisualisatie

1



GitHub

`https://github.com/edovino/KM3IxD`

eigen project/opdrachten ook weer daar - mail me de link!

BBC FOUR



Welke variabelen
gebruikt Hans
Rosling?



Gestructureerde data

Dit soort visualisaties maken gebruik van *gestructureerde data*

- presentatie
bv kleuren afhankelijk van waarden in de data
- animatie
delen van de visualisaties veranderen als de data verandert
- selectie
delen van de data verbergen voor overzicht
inzoomen

2704 Reclamation Bldg 5 St. Louis, MO 63127
 824-0389
 DR. VINCE BISHOP, MD, Overstays Hospital, grad. med.
 Overstays Hospital, Overstays Hospital, grad. med.
 Overstays Hospital, Overstays Hospital, grad. med.

211 West 104 St. NYC 10023
 Tel 9-6278
 Days 10-6000 1700
 NYS 10017

Electrical Engineering Institute, Physics, Optics, Light
 Tunnel in Business address
 Tunnel, Photography, Radar, Electronics, Common, Optical

Gestructureerde data gebruiken

Makkelijke manier nodig om (delen van) die data op te kunnen halen

- kan heel veel data zijn: selecteren

Mechanisme om data op te slaan, te verwerken & terug te halen: database

'Aansturen' van zo'n database doe je (o.a.) met de programmeertaal **SQL**

Relationele database

- data met een onderlinge samenhang
- data met een vaste structuur

In een database zitten de gegevens opgeslagen in *tabellen*. Een tabel heeft kolommen. Iedere kolom heeft een *naam*, en een *type* (*TEXT*, *NUMBER*, *DATE*)

#	year	country	population	life_expectancy	wealth
1	1980	NL	192931	64	10.000
2	1990	AU	5384958	54	15.000

SQL - Structured Query Language

- aparte programmeertaal, ergens mid jaren 70
- ontwikkeld voor *relationele* databases

relationeel: op de een of andere manier zijn verbanden tussen gegevens vastgelegd (en op te vragen)

Relationele databases die SQL gebruiken (o.a.)

- Oracle
- IBM DB2
- mySQL, MariaDB
- Postgresql

Werken met data uit de database: Queries

Query is een commando om data uit database te halen of te bewerken

Ophalen data: **SELECT**

Invoegen: **INSERT**

Bewerken: **UPDATE**

Verwijderen: **DELETE**

Daarnaast nog allerlei functies: **SUM()** , **COUNT()** , **AVG()** , **MIN()** , **MAX()** : je kunt ook (beperkt) rekenen binnen queries

SQLite Browser

Download versie voor juiste systeem vanaf

<http://sqlitebrowser.org/>

Anatomie van een simpel **SELECT** statement

```
SELECT kolommen FROM tabel(len) WHERE voorwaarde(n)
```

kolommen welke kolommen wil je in 't resultaat zien? Gebruik '*' voor alles.

tabellen uit welke tabellen komen de kolommen?

where aan welke voorwaard(en) moeten een rij voldoen om in 't resultaat te komen?

```
SELECT name, email FROM gebruikers WHERE leeftijd < 25
```

geeft naam & email-adres van alle gebruikers onder de 25

SQLite

Simpele, kleine, lichte 'embedded' database

- Oracle, DB2, MariaDB, MySQL zijn aparte, zelfstandige programma's ('servers')
- vaak op aparte machines: vanuit je programma maak je verbinding met de database-server

Bij SQLite hoeft dat niet: wordt ingebouwd in je programma

SQLite data zit in een los file in de `./bin/data` directory van je OF programma

Opdracht volgende keer

Zoek een *interactieve* datavisualisatie / infographic die je aanspreekt of fascineert.

Analyseer wat voor data er gebruikt wordt

- wat voor variabelen worden er getoond?
- hoe zijn die variabelen te beïnvloeden?
- wat voor keuzes zijn er gemaakt in de weergave van die variabelen?
- waar zou de data vandaan gehaald kunnen zijn?
- ontwerp een databasetabel/databasetabellen waar je die data in op zou kunnen slaan