

Data Processing

College 3 - 2018-11-15

Infographics should be....



Functional as a hammer

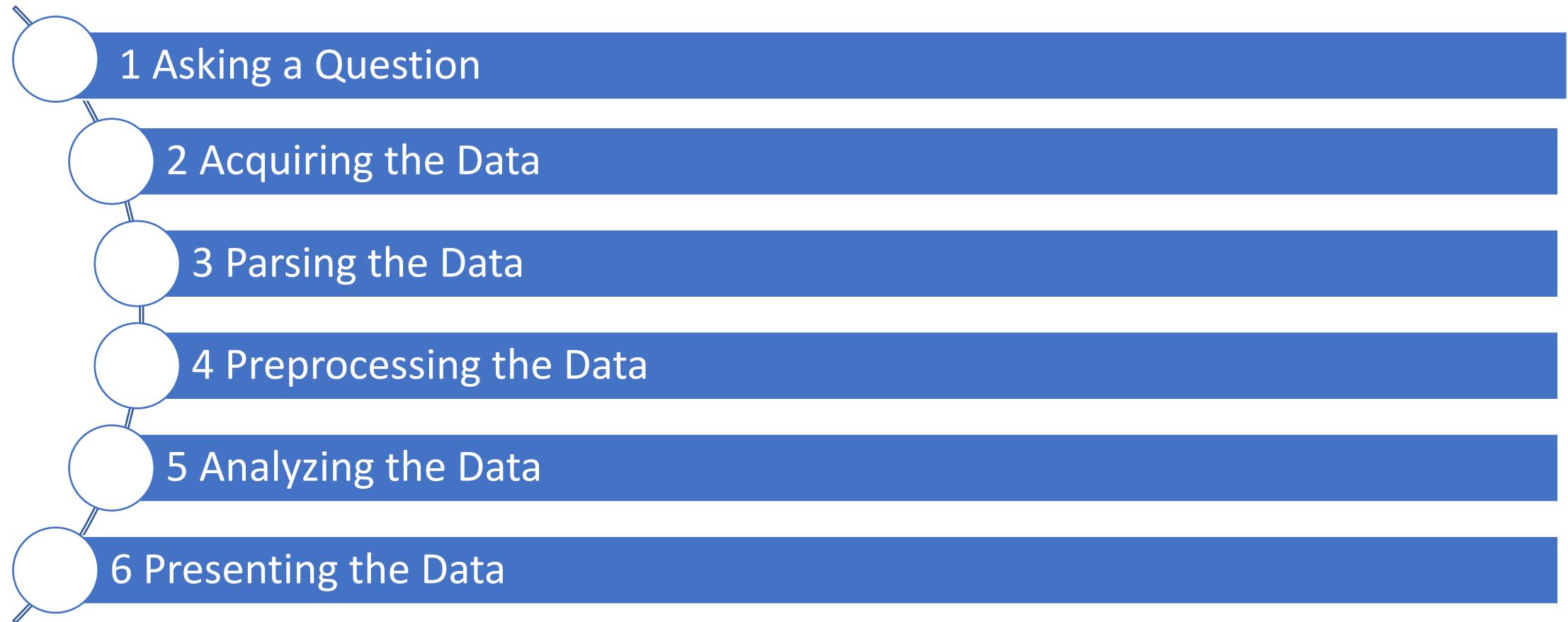


Multilayered as an onion

$$e = mc^2$$

Beautiful as equations

Data Processing Pipeline



De grote lijn

Opdrachten voor het bouwen van een Data Processing Pipeline

Design critiques:

Voorbereidingen op elkaar feedback geven op ontwerp voor eindopdracht.

Eindopdracht Data Processing:

Maak een DPP met een Linked View → verbinden van 2 interactieve visualisaties.

De grote lijn

| no | Assignment | Programming Languages | deadline |
|----|---------------------------|---|------------|
| 1 | Scraping IMDB |  | Wed 7-Nov |
| 2 | Exploratory Data Analysis |  | Wed 14-Nov |
| 3 | JavaScript line |  | Wed 21-Nov |
| 4 | D3 bar chart |   | Tue 27-Nov |
| 5 | D3 scatterplot |   | Mon 3-Dec |
| 6 | Linked views-interim |    | Mon 10-Dec |
| 7 | Linked views |    | Mon 17-Dec |

→ 17:00 (!)

De grote lijn

| no | Assignment | Programming Languages | deadline |
|----|---------------------------|---|------------|
| 1 | Scraping IMDB |  | Wed 7-Nov |
| 2 | Exploratory Data Analysis |  | Wed 14-Nov |
| 3 | JavaScript line |  | Wed 21-Nov |
| 4 | D3 bar chart |   | Tue 27-Nov |
| 5 | D3 scatterplot |   | Mon 3-Dec |
| 6 | Linked views-interim |    | Mon 10-Dec |
| 7 | Linked views |    | Mon 17-Dec |

→ 17:00 (!)

De grote lijn

| september of februari | oktober of maart | november of april | december of mei | januari of juni |
|------------------------------------|------------------------------------|---|--------------------|---------------------------------|
| Fulltime Programmeren 1 | Fulltime Programmeren 2 | Keuze: App Studio of Data Processing | | Programmeer- project |

Programmeerproject januari

Requirements voor project

Project proposal

Volgende week donderdag (22 november, 2018):

Design document

December (deadline komt nog)

Programmeer project januari Requirements

- The project should answer a **clearly defined question**, or help readers answer those questions, through a self-explanatory website. The website needs to tell a story. The story and website should be reasonably original with regards to interactive visualizations that have previously been published.
- You are invited to work on the basis of one or more **data sets that you haven't used in earlier courses**, as long as these allow a visualization of **sufficient complexity** to be created (pending approval by the course's instructor).
- The final product must implement at the **very least 3 linked interactive views** showing different aspects of the data, as well as **at least 2 other interactive components** (e.g. table, dropdown, checkbox, etc.).
- At least one view should comprise a visualization that you have not implemented in earlier courses. At least one interactive component has to be connected to data (e.g., resulting in variable or dataset change).
- The final product embodies **good visualization practices**.

Google launches new search engine to help scientists find the datasets they need

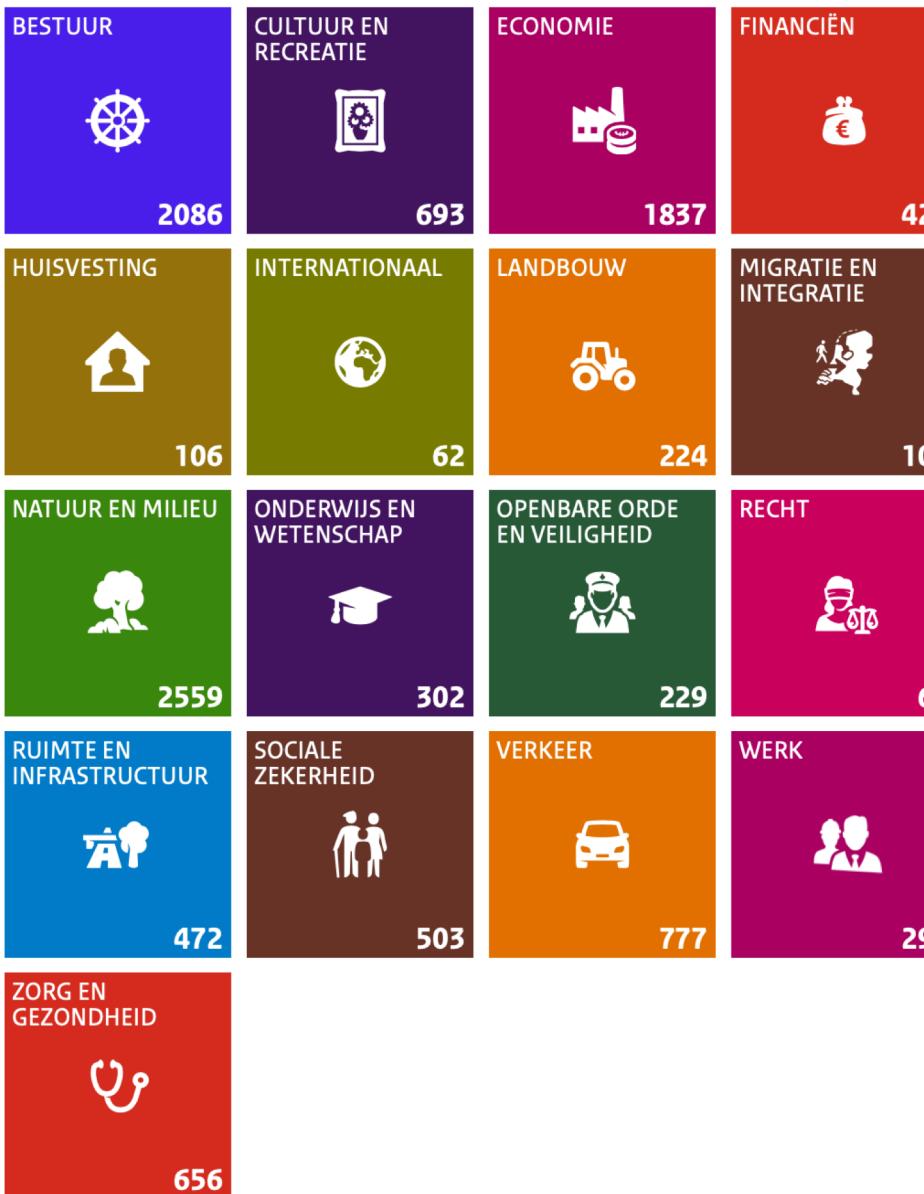
Dataset Search could be a scientist's best friend

By James Vincent | [@jjvincent](#) | Sep 5, 2018, 12:00pm EDT

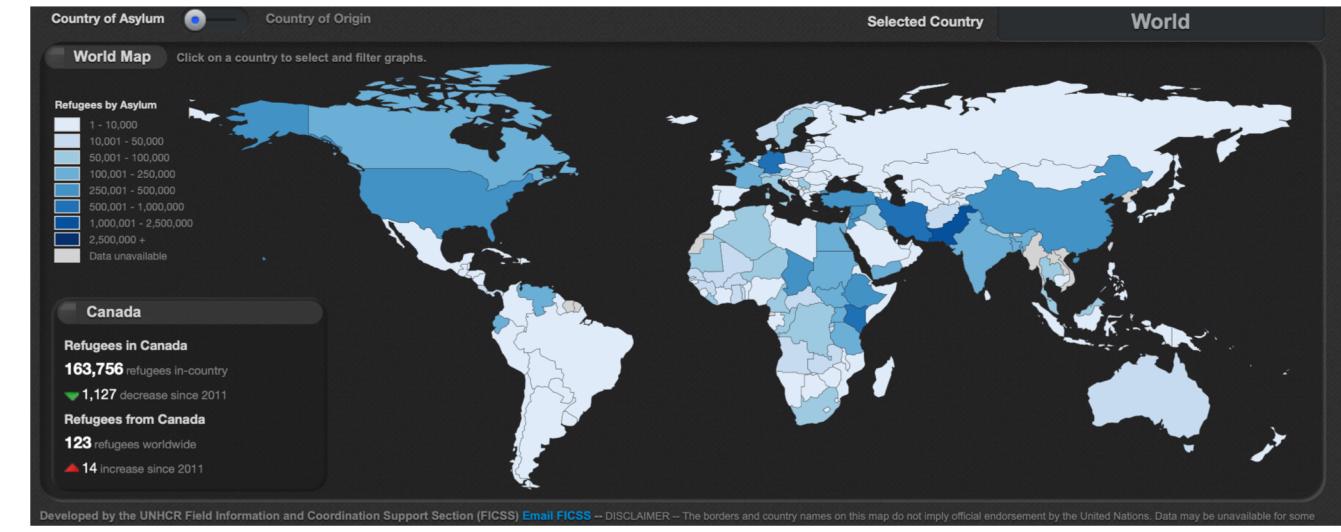
<https://www.theverge.com/2018/9/5/17822562/google-dataset-search-service-scholar-scientific-journal-open-data-access>

<https://toolbox.google.com/datasetsearch>

Blader in datasets op thema (thema's nog niet compleet)



<https://data.overheid.nl>



Refugee Data Flows

http://data.unhcr.org/wiki/index.php/API_Documentation.html

Of bijvoorbeeld...

[Centraal Bureau voor de Statistiek](#)

[KNMI DataCentrum](#)

[Happy Planet Index](#)

[The Organisation for Economic Co-operation and Development](#)

[Kaggle](#)

[World Drug Report](#)

Proposal

write in Markdown

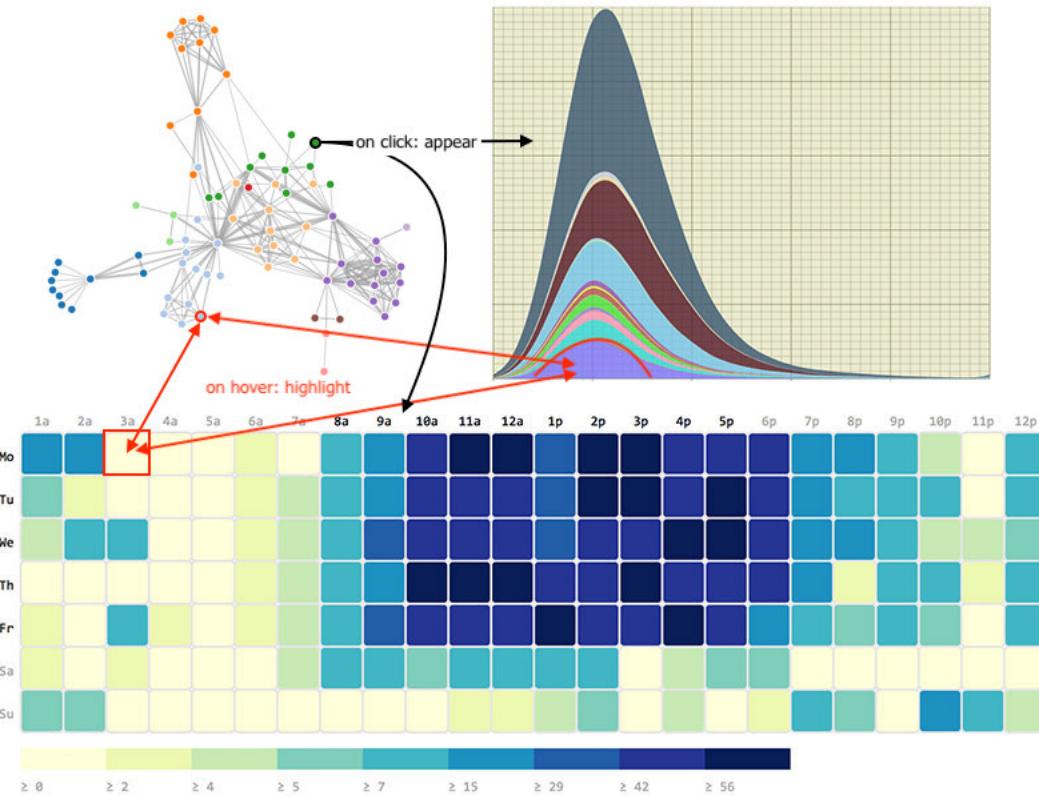
#Problem statement

- a) There is a clearly defined problem that a reasonably large group of people have, which an app or visualization can solve particularly well.
- b) There is a widespread lack of knowledge or understanding that an interactive visualization is particularly suited to remedy.

Proposal

#Solution

- Summarize your idea in a single sentence, connecting it to the “gap” that you describe.
- Include a ***visual sketch***
- List ***main features***
- Split the features into 1) the *_minimum viable product_* (MVP) and 2) parts that are optional to implement.



Proposal

#Prerequisites

- List the ***data sources***, *_links_*, and required *_preprocessing_*.
- List the ***external components*** (libraries like d3-tip or SQLite) that you need to implement certain features. Include the names, and if the component is not standard, include a link to its website.
- Include a review ***related visualizations***, in terms of features and technical aspects: what do they do? how have they implemented it? Can you do it in the same way?
- Identify the ***hardest parts*** of implementing your application: think of technical problems or limitations that could arise during development and what possibilities you have to overcome these.

Proposal

#Problem statement

#Solution

#Prerequisites

Deadline

volgende week donderdag (22 november, 2018) 23:59

Recapture vorige week

- Preattentive features
- Charts and functions
- Principles of graphic design
 - Unity
 - Variety
 - Hierarchy
- Strategy
 - Grids
 - Color
 - Type

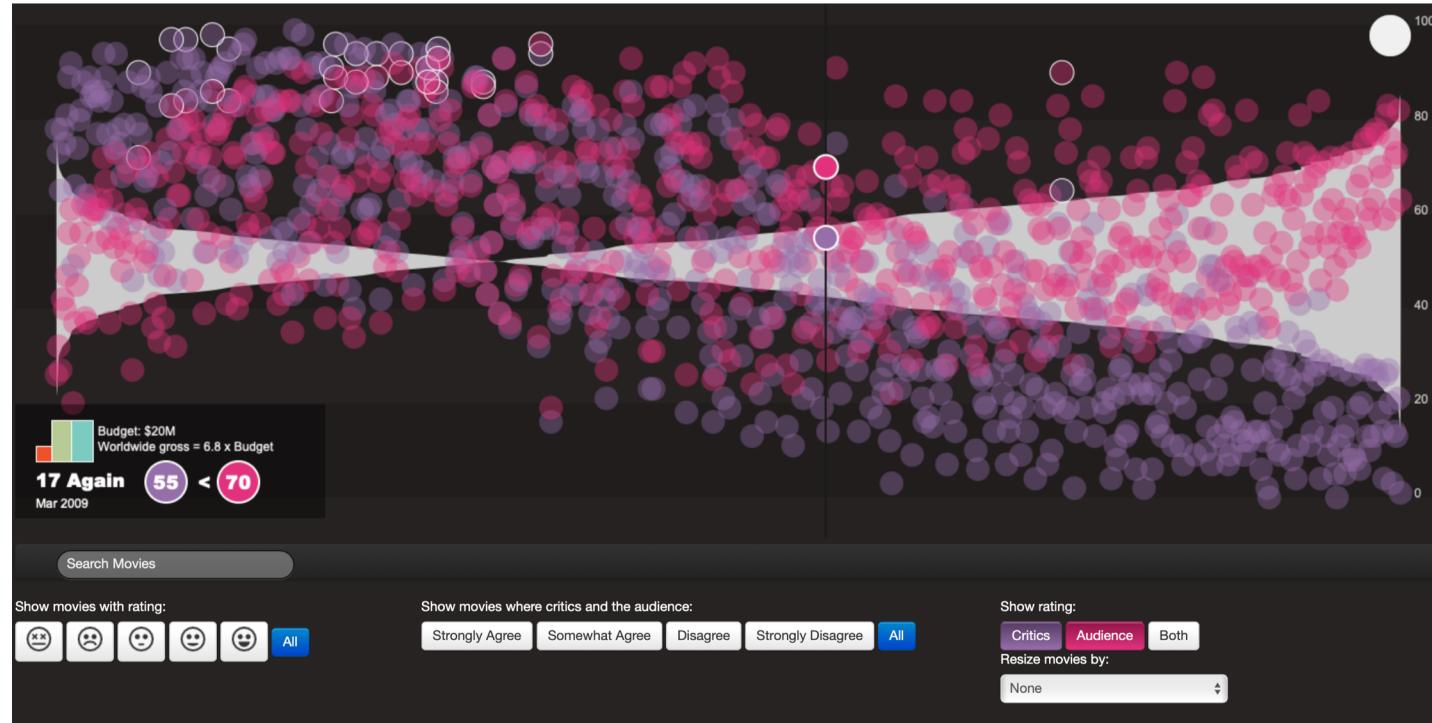
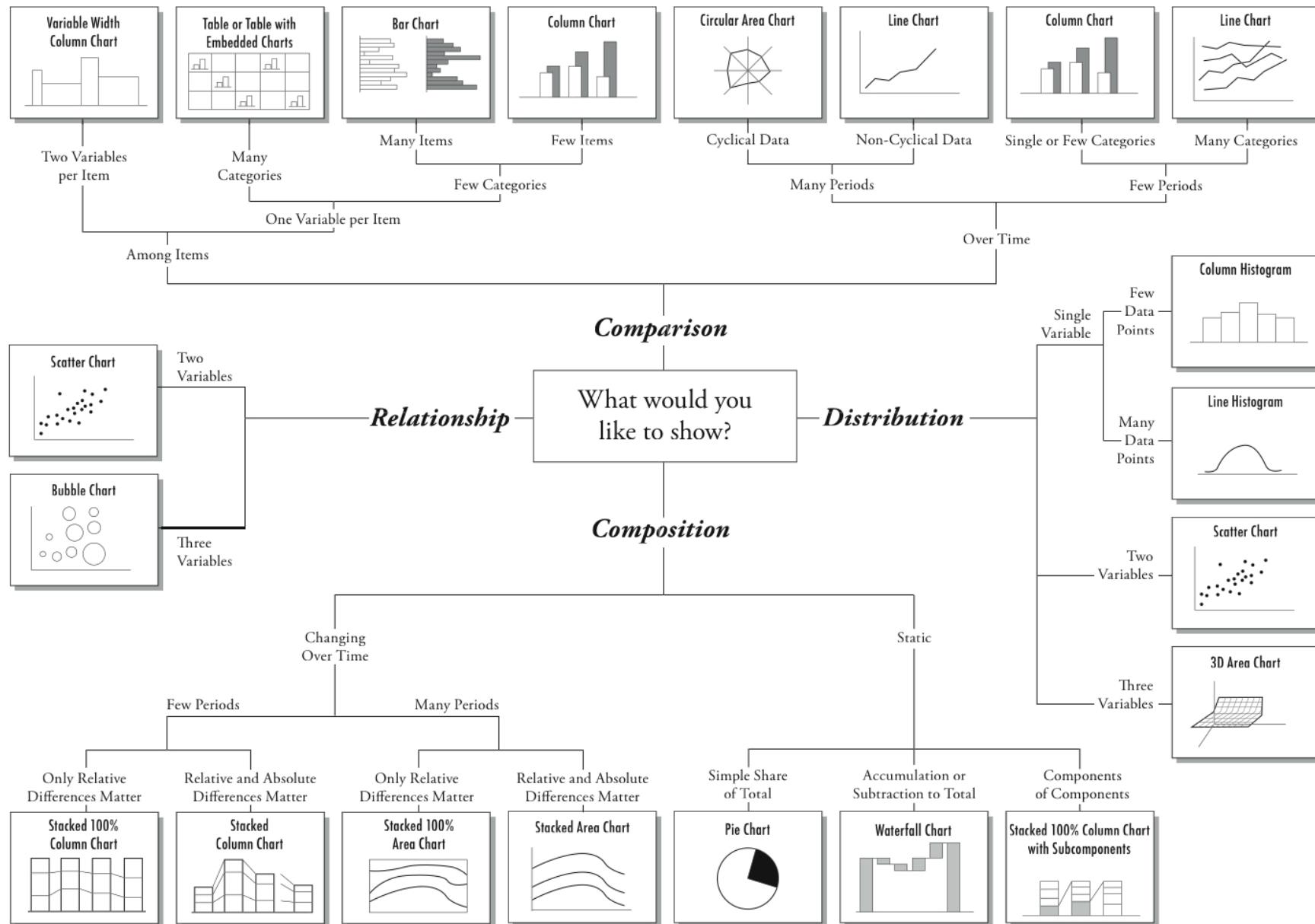


Chart Suggestions—A Thought-Starter



Design critique

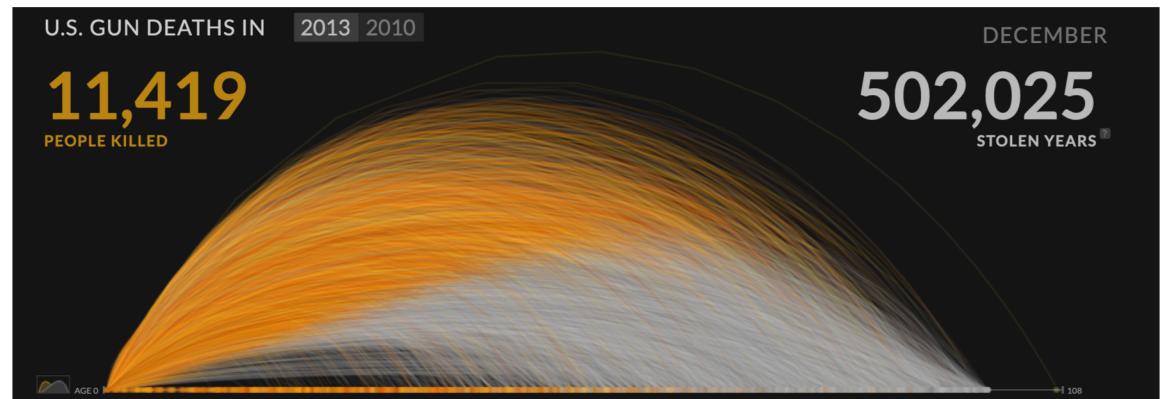
Link naar PDF gemaild om 9:55

Werk in groep van 2-3 personen

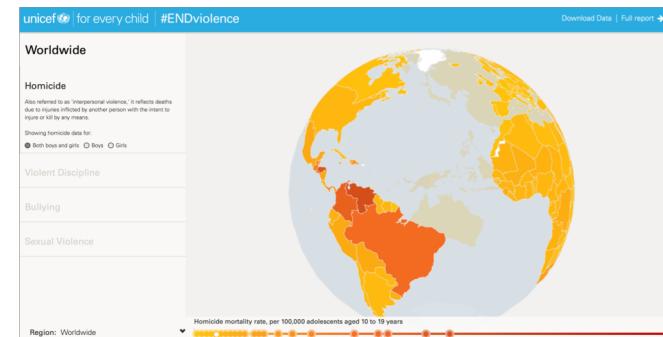
Hierna twee groepen combineren en resultaten uitwisselen

Kies een casus

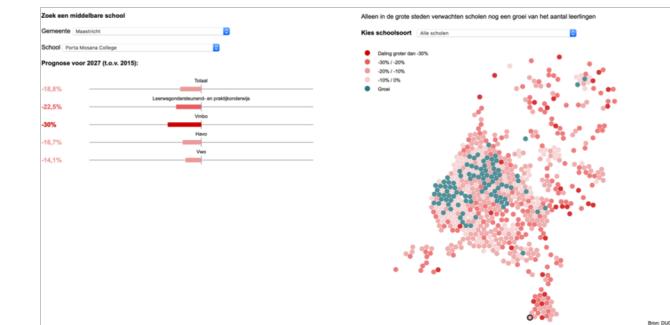
<https://guns.periscopic.com/?year=2013>



<https://works.periscopic.com/unicef-child-violence/#all>



<https://app.nos.nl/datavisualisatie/voortgezet-onderwijs/>



Vorm een beeld – wat staat er?

Analyse vragen:

- A1) Who is the intended audience?
- A2) What tasks does the visualization enable?
- A3) What data is represented in this visualization? Be specific.
- A4) How is each data type visually encoded?
- A5) How do the visual elements and user interactions support the tasks?
- A6) Why do you like / dislike this visualization?

Toepassing theoretische concepten

Toepassings vragen:

- T1) **Identificeer** van welke **preattentive features** gebruik wordt gemaakt
- T2) Voor de gekozen representatie: **bepaal** op de schaal van **Cleveland & MacGill** in hoeverre de accurate of generieke analyse mogelijk maakt.
- T3) **Reflecteer** of de classificatie voor Cleveland & MacGill in dit geval klopt
- T4) In hoeverre spelen de principes van **Unity**, **Variety** en **Hierarchy** een rol in deze visualisatie.
- T5) **Identificeer** welke **type grafisch figuur** gebruikt wordt.
- T6) Geef **sterke en zwakke punten** van dit type figuur in deze context.
- T7) Welke grafisch figuur zou je als **alternatief** kiezen. Onderbouw je keuze.

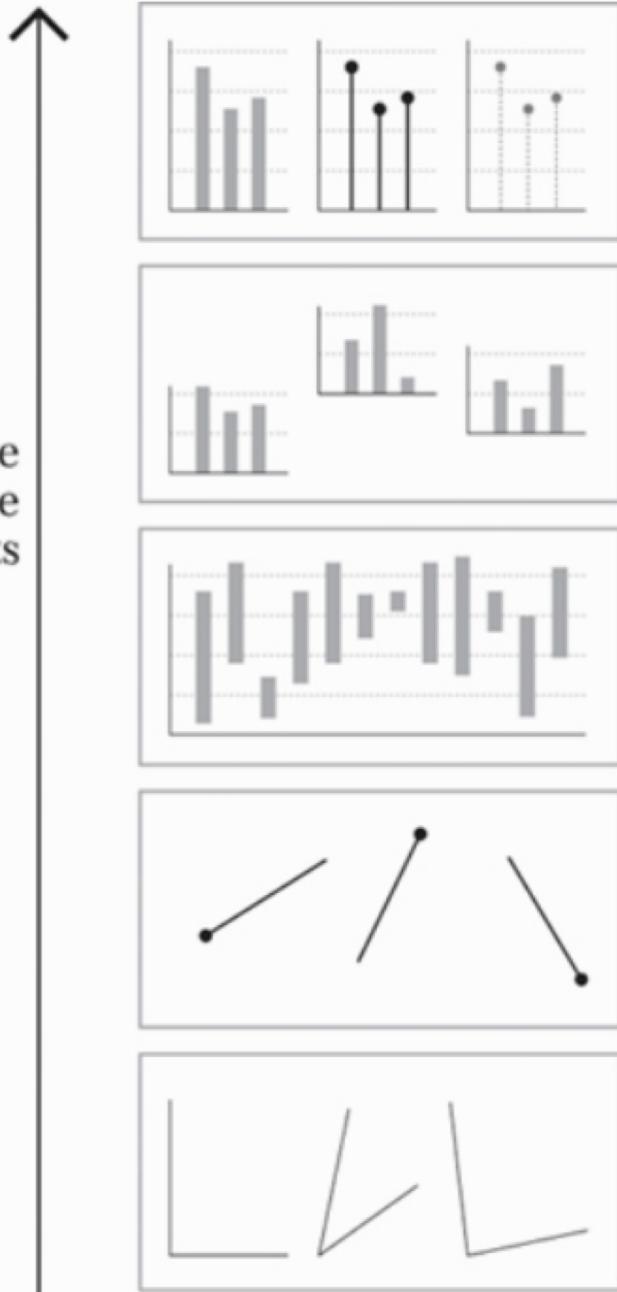
Preattentive features

- Size
- Orientation
- Color/Shade
- Proximity
- Similarity
- Connectedness
- Continuity
- Closure

De voorbeelden hiervan staan in de slides van college 2:

<https://surfdrive.surf.nl/files/index.php/s/3ayy45ieswL6S2q>

Allows more accurate judgments



Position along a common scale

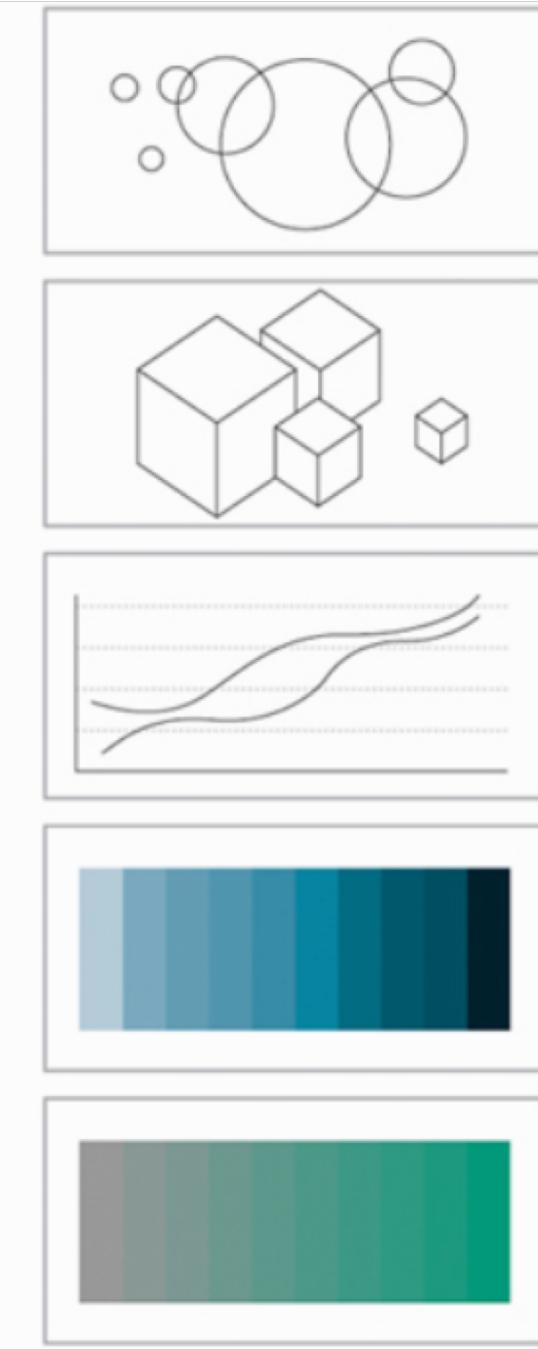
Position along nonaligned scales

Length

Direction

Angle

Allows more generic judgments



Area

Volume

Curvature

Shading

Color saturation

Main principles of graphic design

- **Unity**

Unity is the presentation of a composition as an integrated whole, not as a mere sum of its parts

- **Variety**

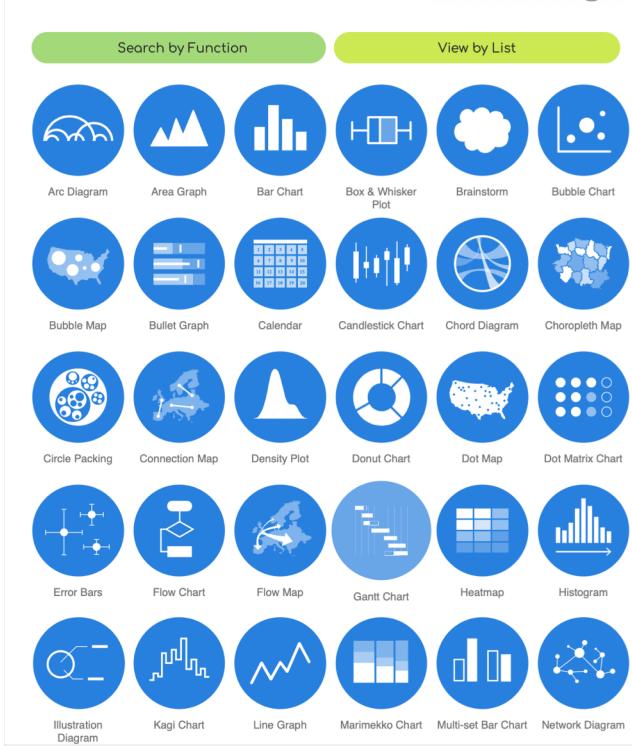
Variety is the opposite of unity, but also its complement. With too much variety, a composition will look random; with too much unity, it will look boring

- **Hierarchy**

The balance between unity and variety can lead to a good **hierarchy**

What do you want to show?

Charts and functions



<https://datavizcatalogue.com>

<https://datavizcatalogue.com/search.html>

Chart Suggestions—A Thought-Starter

