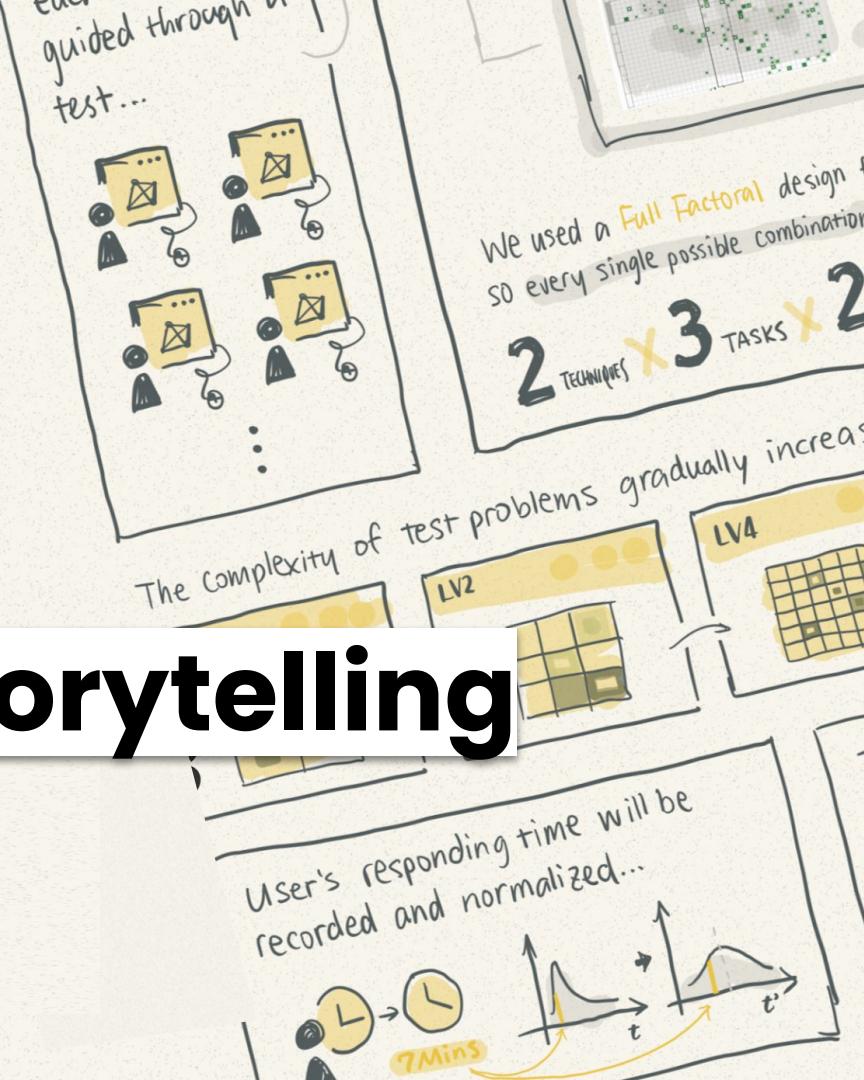


Data Comics for Data-Driven Storytelling

Benjamin Bach

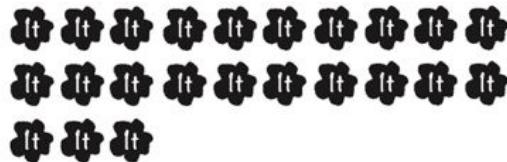
University of Edinburgh



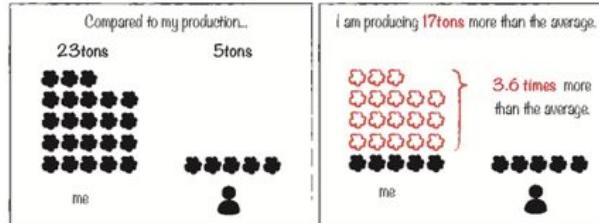
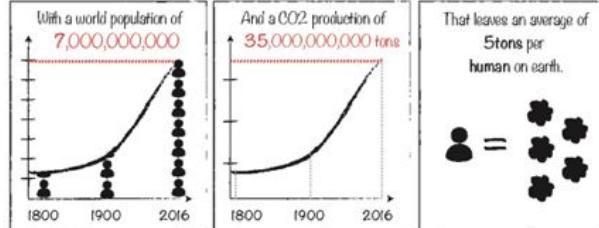
CO Footprint



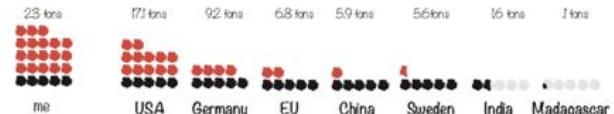
Which produced...



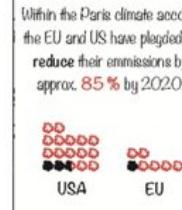
23 tons of CO₂.



Thus, my travels in 2016 alone produced more CO₂ than the average person in the most countries:



Within the Paris climate accord, the EU and US have pledged to reduce their emissions by approx. 85% by 2020.



As European, this leaves me 1 trans-atlantic flight a year.



Just that:

no other transports
no computer, no mobile,
no heating,
no clothes,
no food.

MY LAST WEEK'S SLEEP RECORD
07.May-13.May



Context, Motivation & Problem Study

This is a
THREE CARDS

To test which one performs better, we designed three tasks.
Tasks like these are frequently used in domains like brain connectivity analysis:

11

TASKS

CONNECTIVITY

REGION

Identify the region with most changes

- 1 Examine each region and observe the graph change

- 2 Estimate the changes in each region

- 3 Click on the region with the highest edge weight variation

Assess the connectivity of common neighbors of two nodes

Find the common neighbors, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the

nodes, meaning the nodes that are

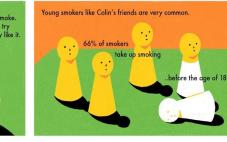
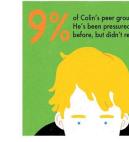
neighbors of both of the

nodes, meaning the nodes that are

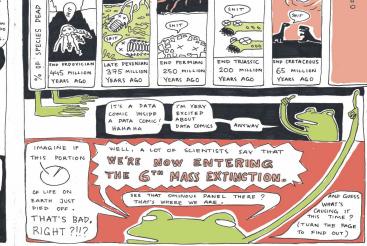
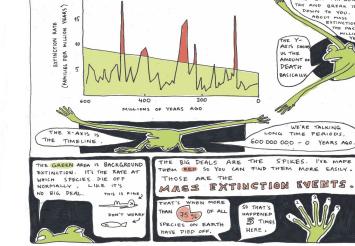
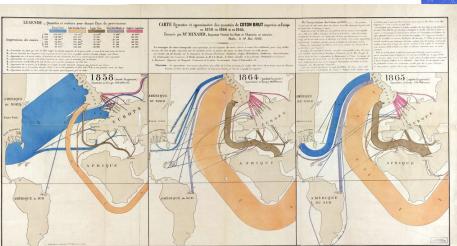
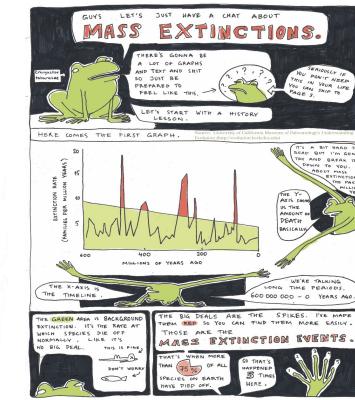
neighbors of both of the

nodes, meaning the nodes that are

neighbors of both of the



TONNES





Visual+
Interactive
Data



THE UNIVERSITY
of EDINBURGH



University
of Glasgow

Microsoft
Research

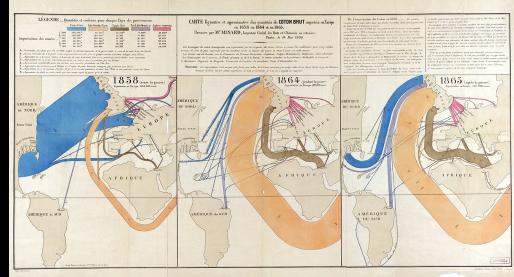
UNIVERSITY OF
TORONTO



HARVARD
UNIVERSITY



UNIVERSITY OF
CALGARY



What are data comics?

physical, social, political, and economic structures of a region can place residents at varying risks for vulnerability. Areas susceptible to violence or natural disaster pose clear threats to individuals. An individual's environment also affects his or her development and behavioral choices. Resources available in the physical and social environments create the contexts within which decisions are made about health, education, and employment. Political and social environments also dictate whether resources are accessible to all adolescents. An examination of the residential distribution of adolescents provides a baseline for comparing geographical patterns of vulnerability. Within Uganda, by type of residence, the majority of adolescents (87 percent) live in rural versus urban areas. Figure 6 shows the distribution of adolescents aged 10 to 19 living in Uganda. Regional distributions show Karamoja contains only four percent of the adolescent population. Kampala with a much denser population contains 4.6 percent of the population. The Eastern and Western regions contain the largest proportions of the adolescent population.

Household factors influencing vulnerability

Household-level factors have direct impacts on the well-being of adolescents. Households are the primary setting where adolescents live and engage in activities. For this reason, the household environment and the people who live there have significant impacts on the lives of adolescents. Physical conditions of the home influence the health of residents. Family structures and demographic characteristics of household members affect the knowledge, decisions, behaviors and interactions in the environment of the adolescent.

Access to improved water sources and sanitation

Unsafe water, inadequate sanitation, and poor hygiene are among the five leading risk factors responsible for one quarter of all deaths in the world (WHO 2009). Unsafe water supplies and inadequate sanitation in homes increase exposure to water-borne diseases and can cause diarrhea. Ensuring access to clean water sources and sanitation is key to maintaining hygiene and health. Improved water sources are those that either naturally protect water from contamination or are constructed to do so. These include piped water, public taps, standpipes, boreholes, tube wells, protected wells and springs, and rainwater collection. Improved sanitation includes constructs and systems that prevent fecal contamination. These include flush or pour toilets, ventilated pit latrines, pit latrines with slabs, and composting toilets (UNICEF 2013b).

Housing conditions across East and Southern Africa are largely in need of improvement, and lack of improved sanitation varies by country. In nearly all of East and Southern Africa, over half of adolescents either do not have improved sanitation or share facilities with other households. Conditions are worst in Madagascar and Mozambique where fewer than four percent of adolescents live in households with improved sanitation that is not shared (Figure 7). Rwanda has the lowest proportion of adolescents affected—35 percent—which is still unacceptably high. Lack of access to improved water sources affects lower proportions but is still a problem in the region. In five countries, fewer than half of adolescents have access to improved water sources (Figure 8). Water conditions are best in Namibia, where only 15 percent of adolescents have no access to improved water.

In Uganda, overall access to improved water and sanitation increased by a small but significant percentage between 2006 and 2011 (Figure 9). In 2006, 33 percent of adolescents had no access to improved water; in 2011, it is 30 percent. The proportion of adolescents without access to improved

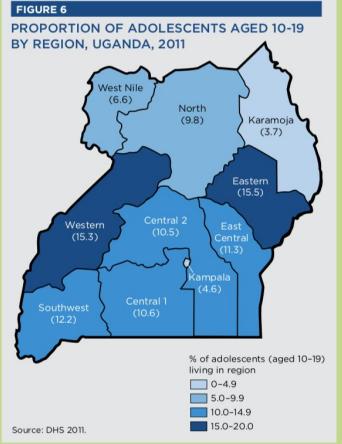


FIGURE 7
PERCENT OF ADOLESCENTS AGED 10-19
LIVING IN HOUSEHOLDS WITH NO
IMPROVED OR WITH SHARED SANITATION,
EAST AND SOUTHERN AFRICA



FIGURE 8
PERCENT OF ADOLESCENTS AGED
10-19 LIVING IN HOUSEHOLDS WITH
NO IMPROVED WATER SOURCE, EAST
AND SOUTHERN AFRICA

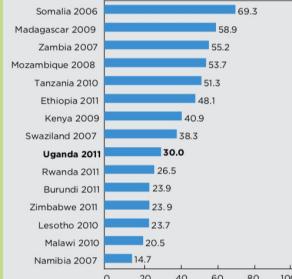
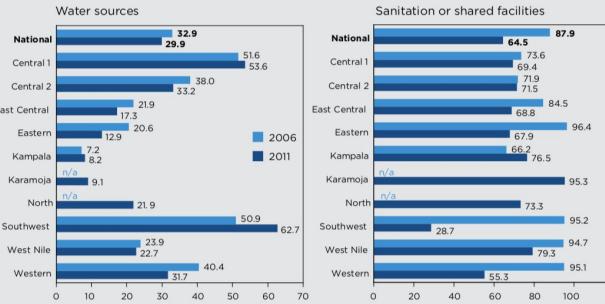


FIGURE 9
PERCENT OF ADOLESCENTS AGED 10-19 LIVING IN HOUSEHOLDS WITHOUT ACCESS TO
IMPROVED WATER AND WITHOUT ACCESS TO IMPROVED OR WITH SHARED SANITATION, IN
UGANDA, BY REGION, 2006 AND 2011



Data Comics

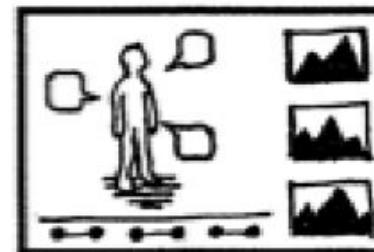
Seven Genres



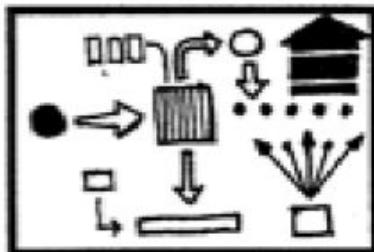
Magazine Style



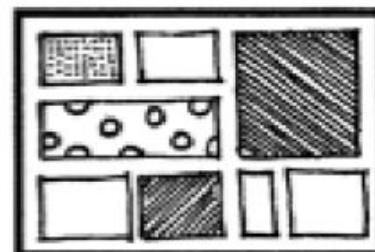
Annotated Chart



Partitioned Poster



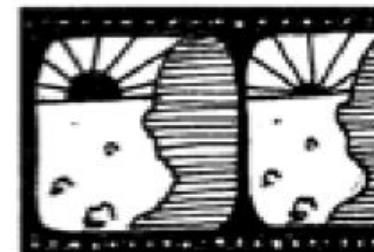
Flow Chart



Comic Strip



Slide Show



Film/Video/Animation

Data Comics

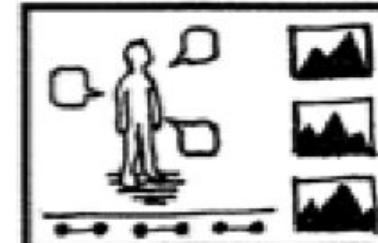
Seven Genres



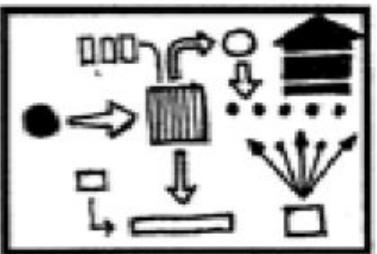
Magazine Style



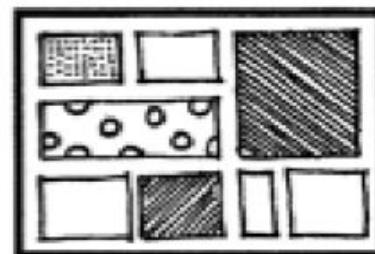
Annotated Chart



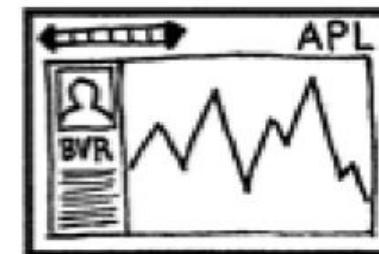
Partitioned Poster



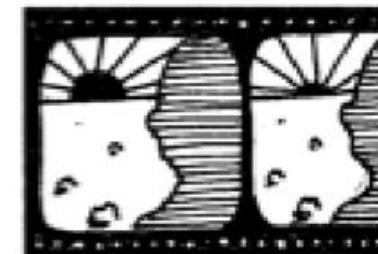
Flow Chart



Comic Strip

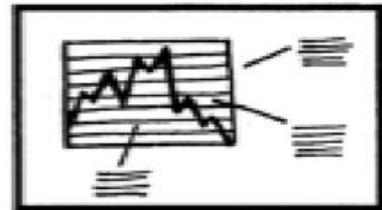
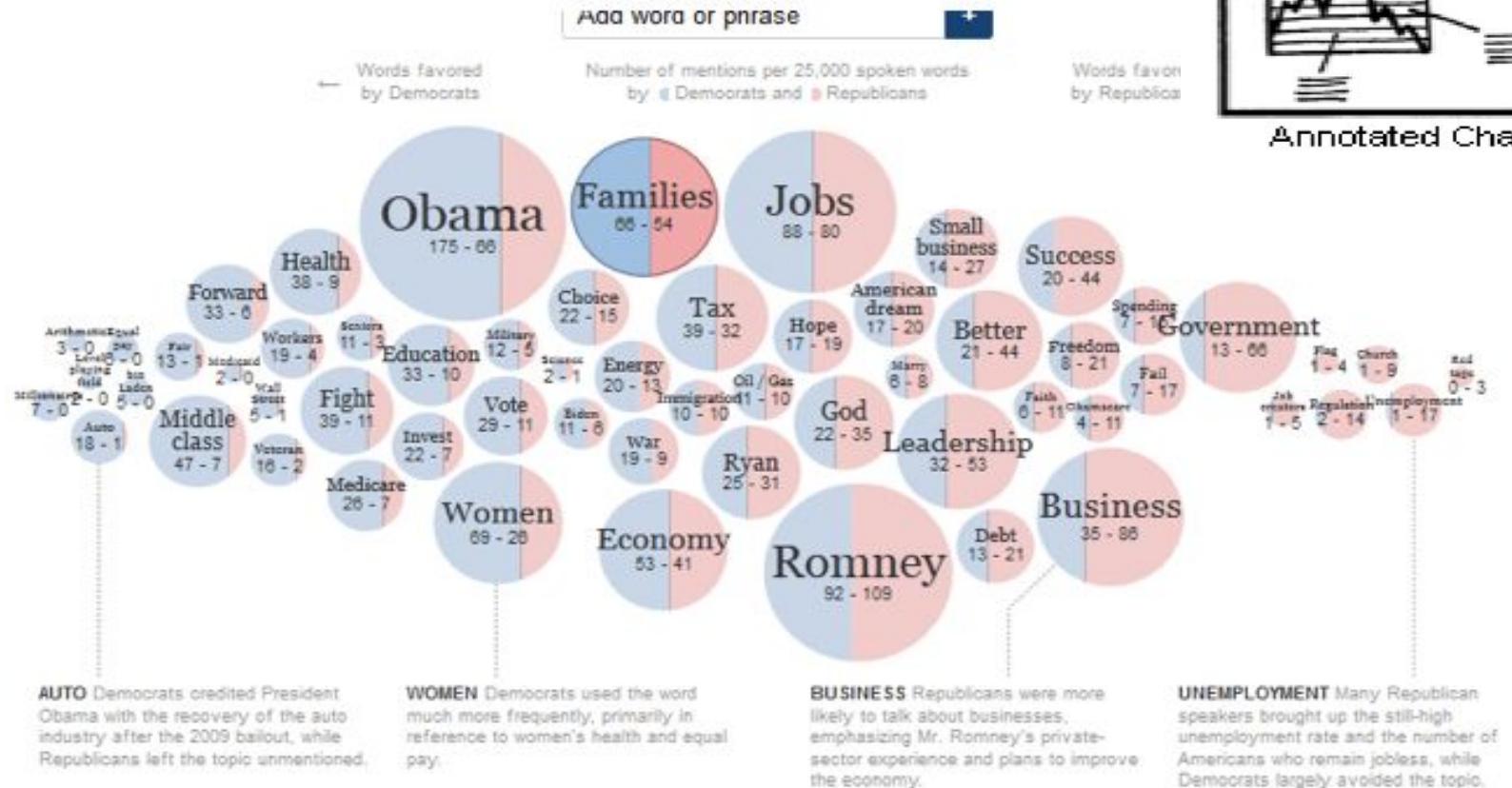


Slide Show



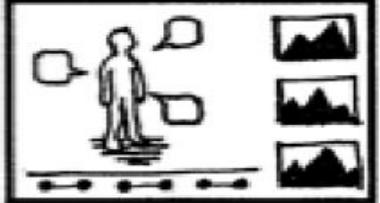
Film/Video/Animation

Annotated Chart



Annotated Chart

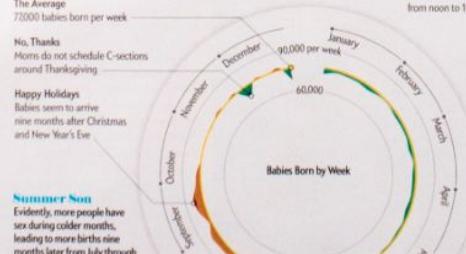
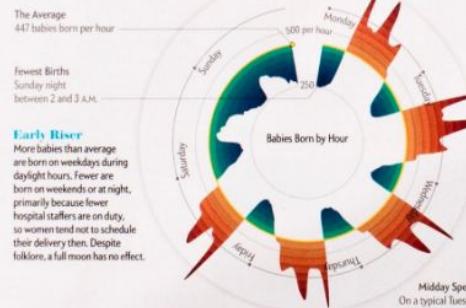
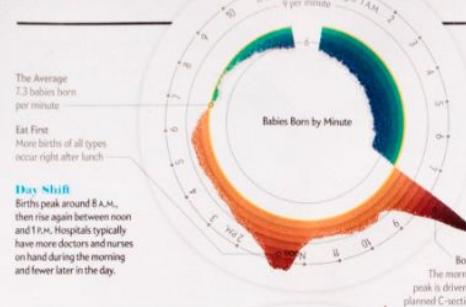
Partitioned Poster (Infographic)



Partitioned Poster

<https://www.visualcinnamon.com/portfolio/baby-spike>

GRAPHIC SCIENCE



The Baby Spike

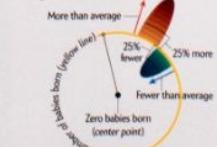
Births peak on weekdays during daytime work hours

Two generations ago babies were born pretty much spontaneously, around the clock. But today in the U.S., about half of all births are cesarean sections prescheduled by Mom or deliveries induced by doctors concerned about the mother's or baby's health. These medical procedures have skewed the days of the week, and hours of the day, during which those little bundles of joy arrive.

The procedures dominate because more than 98 percent of infants are born in a hospital, despite what seems to be the rising popularity of home births. Far more babies now arrive on weekdays than on weekends, most between 8 A.M. and 6 P.M. "We can't schedule spontaneous labor, obviously," says Neel Shah, a physician and professor at Harvard Medical School. "But we can schedule delivery."

—Mark Fischetti and Zan Armstrong

Each graph shows U.S. data averaged across 2014



SOURCE: NATIONAL CENTER FOR HEALTH STATISTICS AND U.S. SOCIAL SECURITY ADMINISTRATION

Data Comics

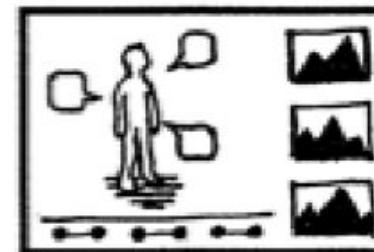
Seven Genres



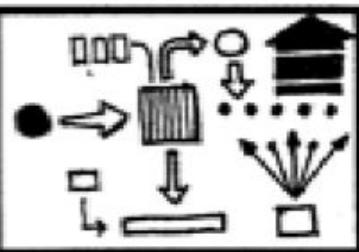
Magazine Style



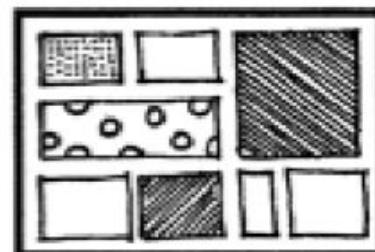
Annotated Chart



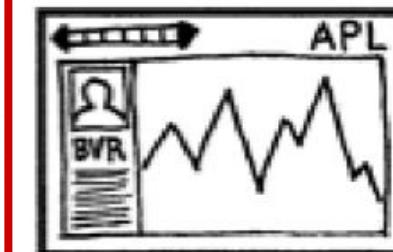
Partitioned Poster



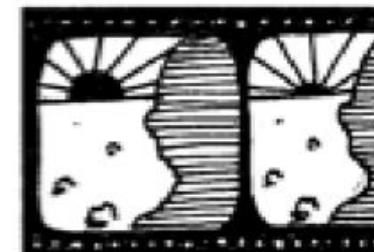
Flow Chart



Comic Strip

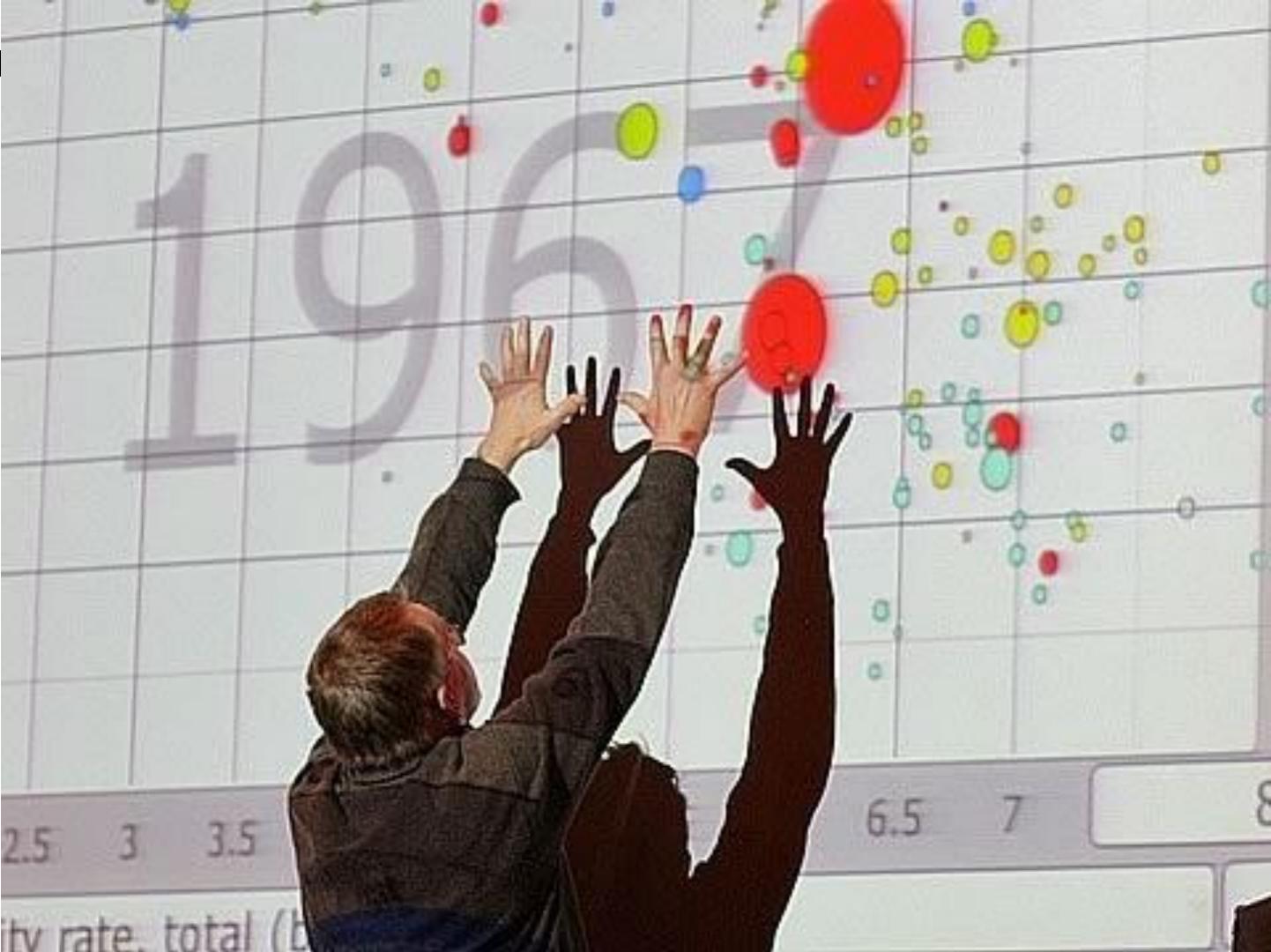


Slide Show



Film/Video/Animation

Life Presentation



[https://www.ted.com/talks/
hans_rosling_shows_the_b
est_stats_you_ve_ever_se
en?language=en](https://www.ted.com/talks/hans_rosling_shows_the_best_stats_you_ve_ever_seen?language=en)

Data Videos



<https://www.youtube.com/watch?v=QPKKQnijnsM>

<https://www.youtube.com/watch?v=QPKKQnijnsM>

Space vs. Time oriented formats

Space oriented

Use space to structure information

Large and detailed images

Space == importance

Encourage exploration

Adapt to readers' pace

> Reader driven

Time oriented

Use time to structure information

Sequences of images

Time+order == importance

Encourage explanation

Follow author's narrative

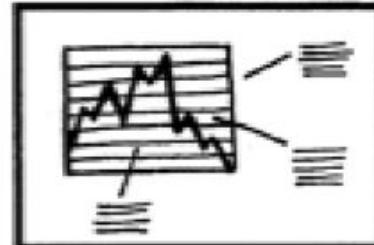
> Author driven

Data Comics

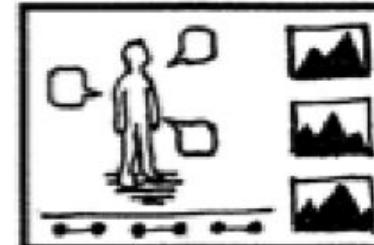
Seven Genres



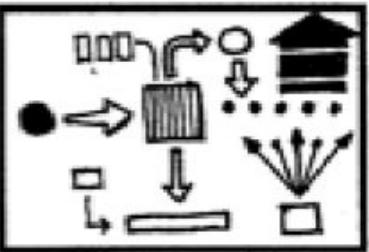
Magazine Style



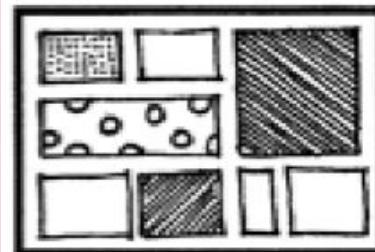
Annotated Chart



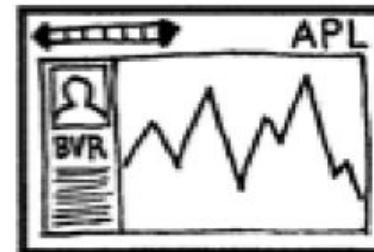
Partitioned Poster



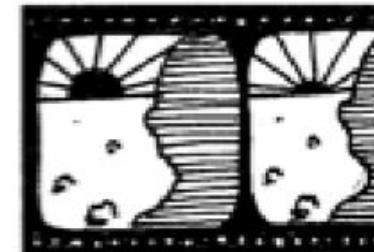
Flow Chart



Comic Strip

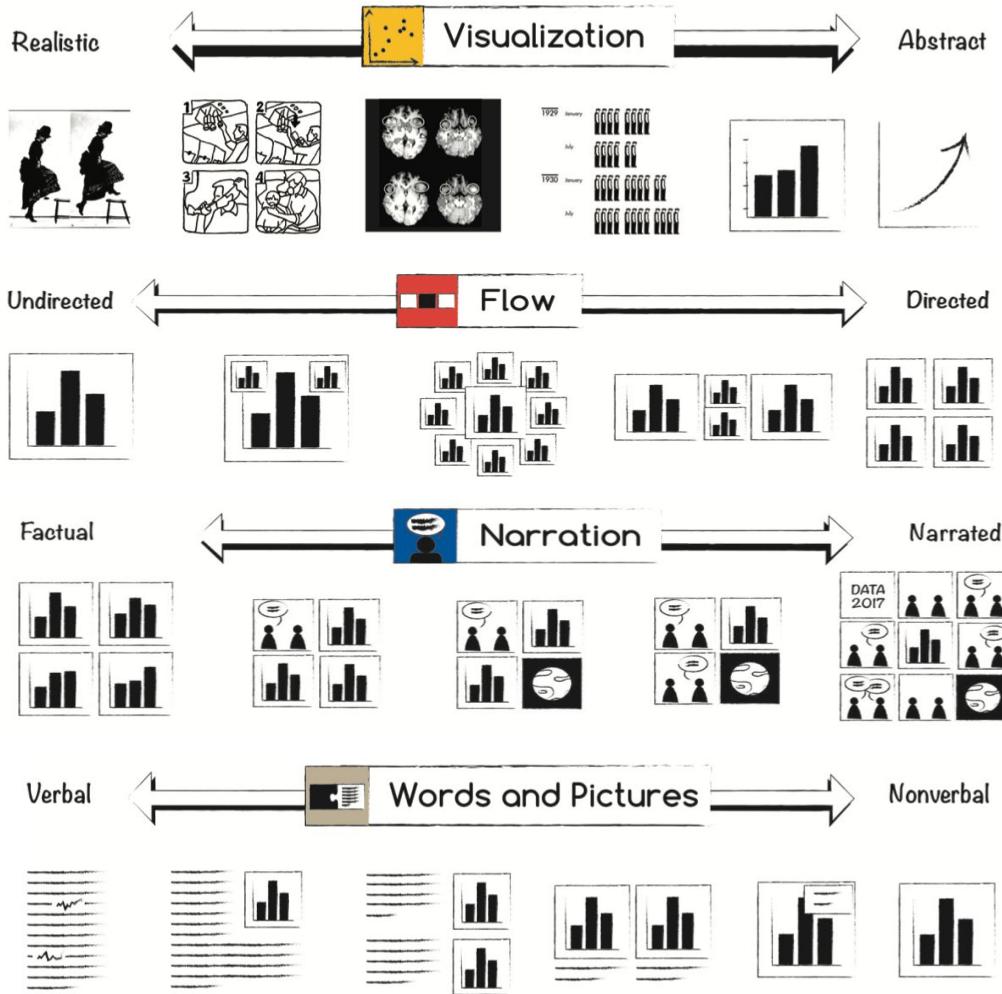


Slide Show



Film/Video/Animation

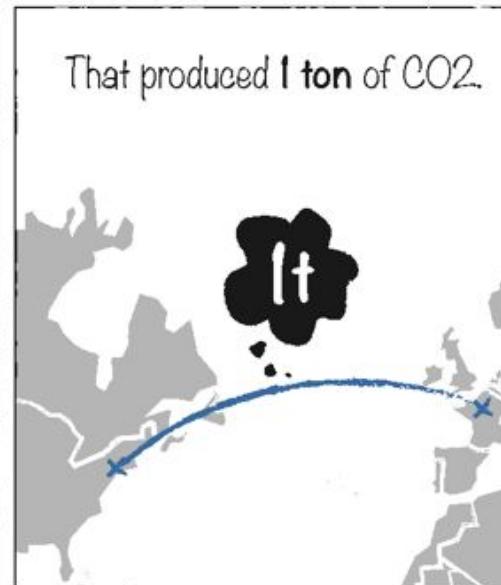
Dimensions of data comics



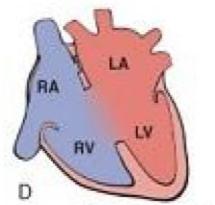
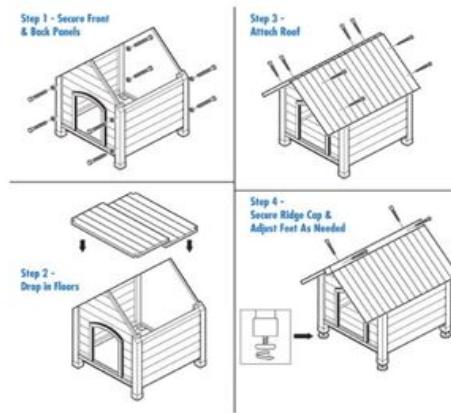
Panel (sequences)



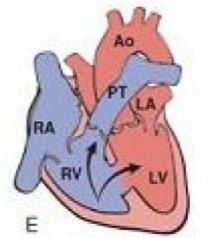
Panel (sequences)



Sequential Explanations

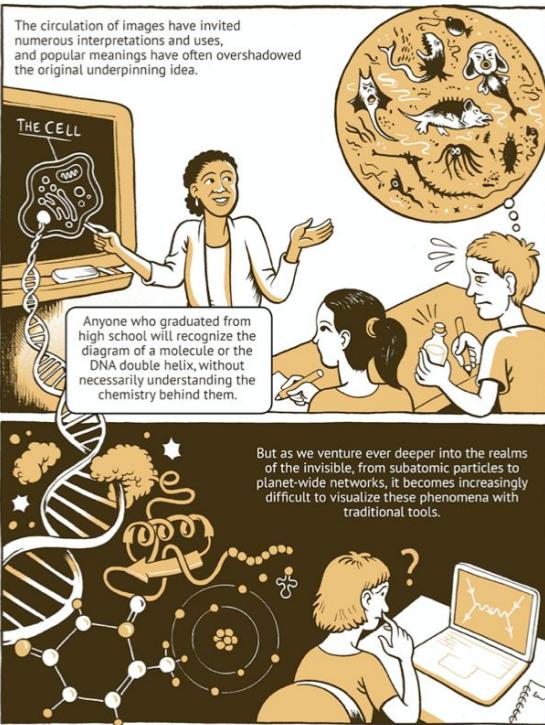


D
Complete Atrioventricular
Canal Defect



E
Large VSD with
Irreversible Pulmonary
Hypertension

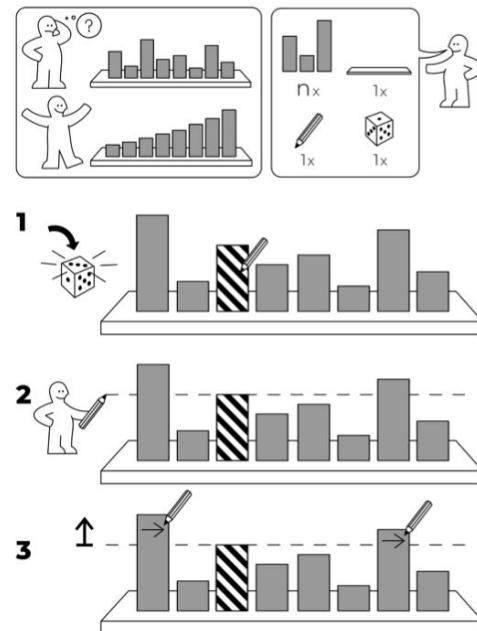
Scientific comics



Farinella, Matteo. "Of Microscopes and Metaphors: Visual Analogy as a Scientific Tool." *The Comics Grid: Journal of Comics Scholarship* 8 (2018).

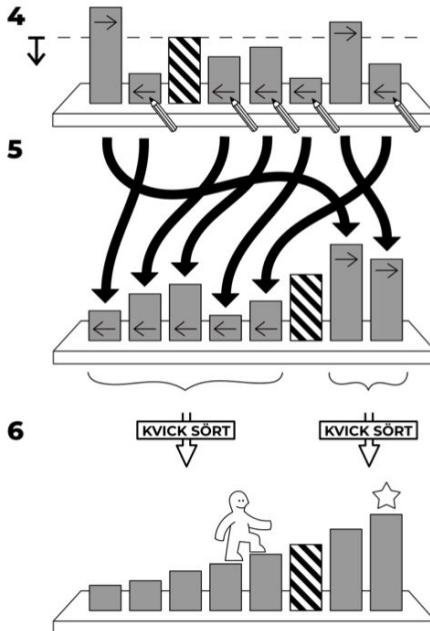
Visual instructions

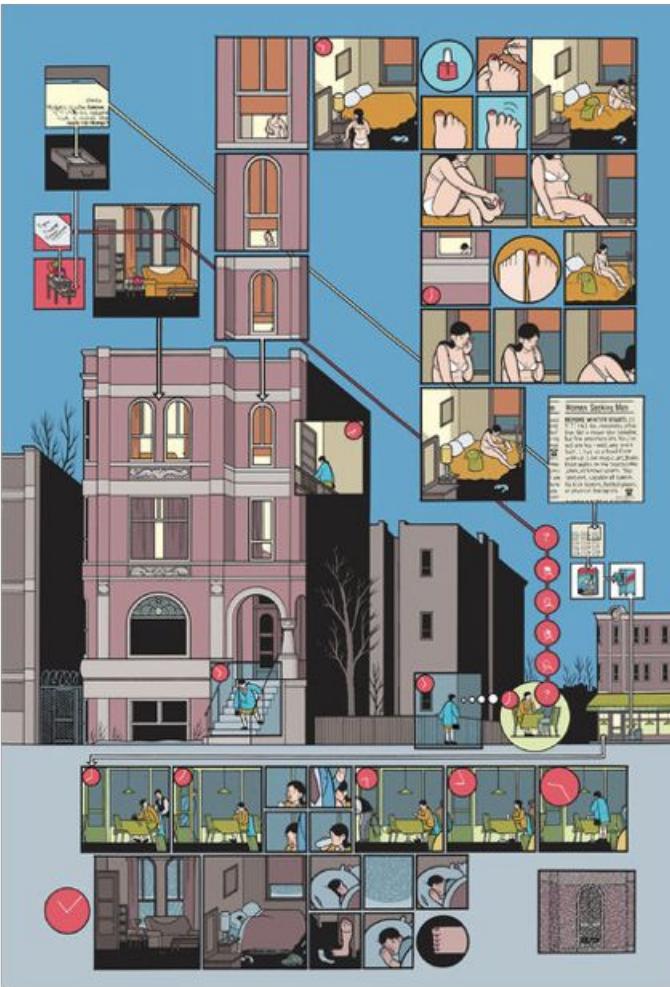
KVICK SÖRT

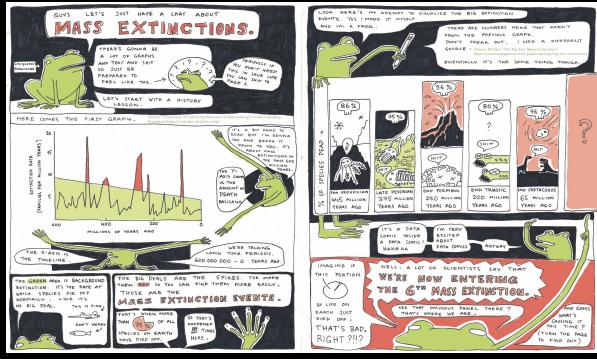


idea-instructions.com/quick-sort/

idea-instructions.com/quick-sort/
v1.1, CC by-nc-sa 4.0 **IDEA**

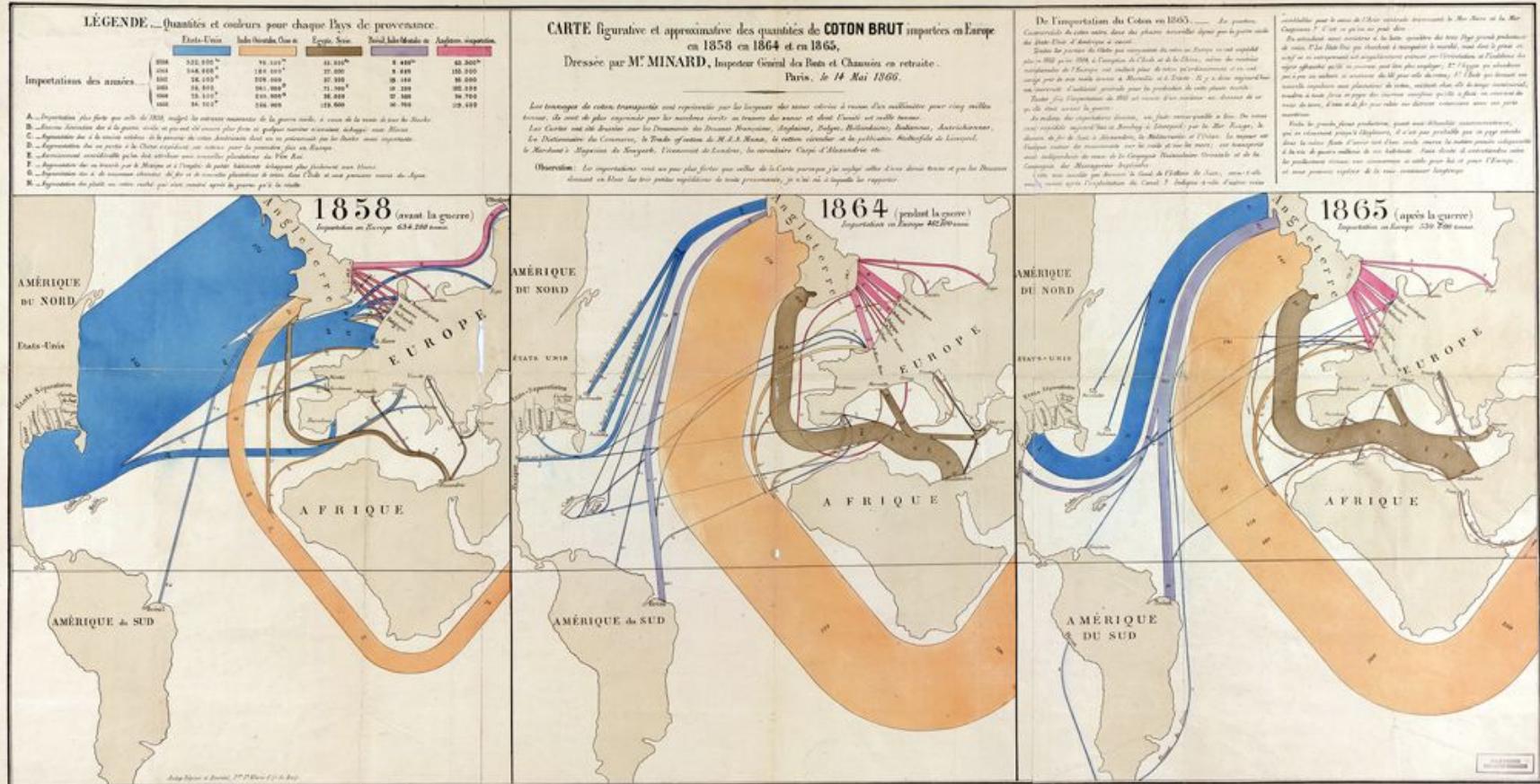






Examples of Data Comics

Joseph Minard, ~1880s



Isotype, Arnold Gantz, ~1920



Maps can show the way birds go when they fly north or south for the season.

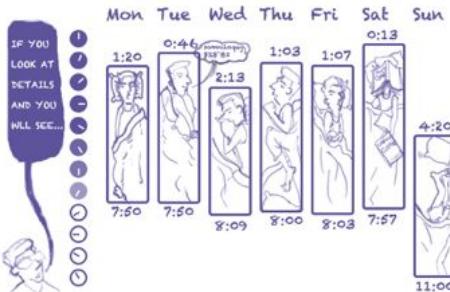


MY LAST WEEK'S SLEEP RECORD

07.May-13.May

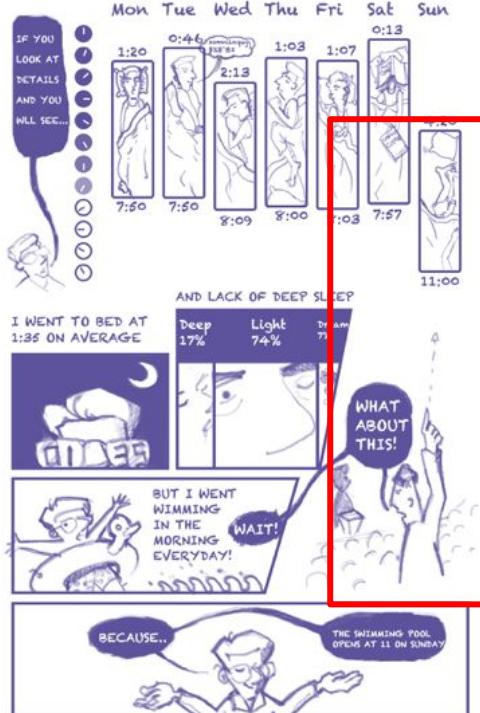
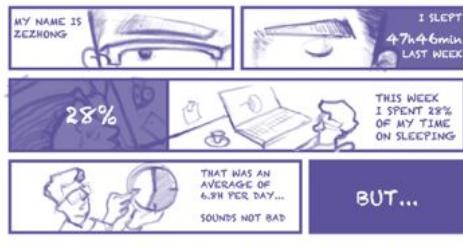


BUT...



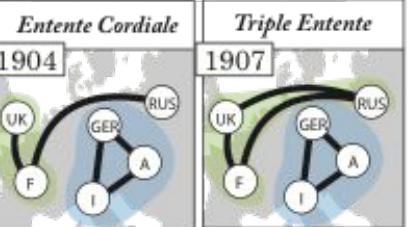
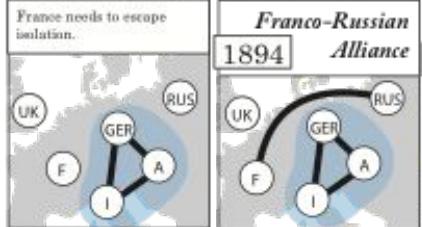
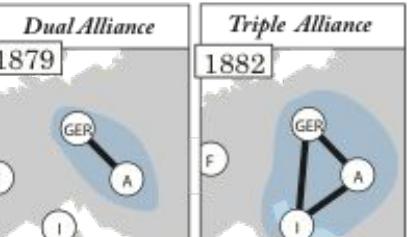
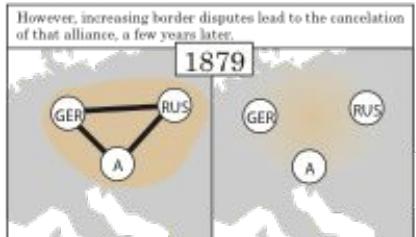
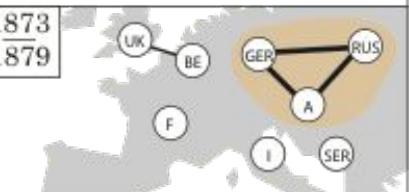
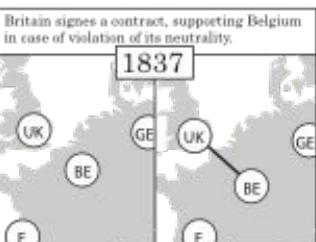
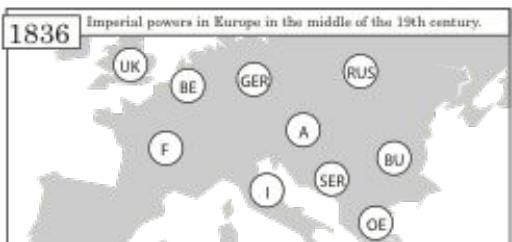
MY LAST WEEK'S SLEEP RECORD

07.May-13.May

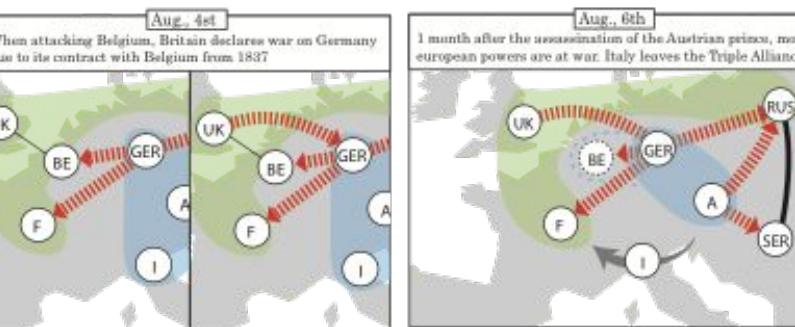
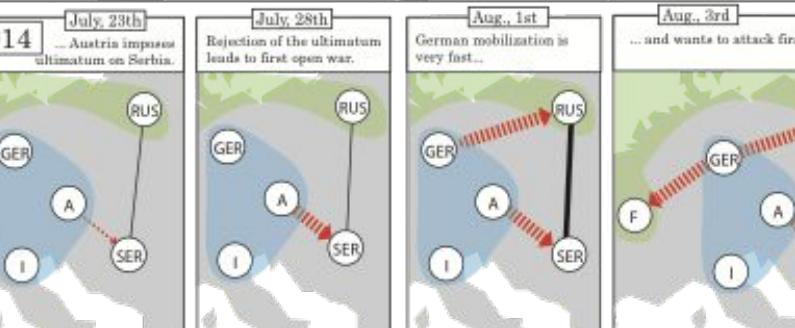
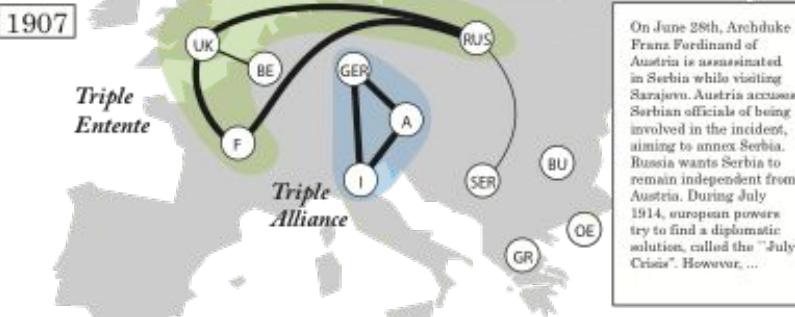


(c) Zezhong Wang, in Bach, B., Wang, Z., Farinella, M., Murray-Rust, D. and Henry Riche, N., 2018, April. Design patterns for data comics. In *Proceedings of the 2018 chi conference on human factors in computing systems* (pp. 1-12).

European Alliances before World War I (1836-1914)



In 1907 the most important European powers are divided into major alliances. On the Balkan, Serbia, Austria Greece and the Ottoman empire gamble for influence.



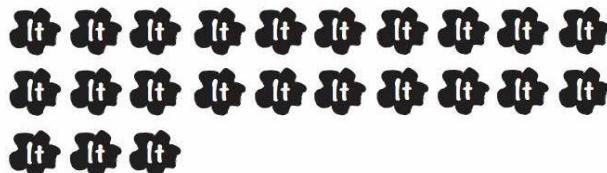
On June 28th, Archduke Franz Ferdinand of Austria is assassinated in Serbia while visiting Sarajevo. Austria accuses Serbian officials of being involved in the incident, aiming to annex Serbia. Russia wants Serbia to remain independent from Austria. During July 1914, european powers try to find a diplomatic solution, called the "July Crisis". However, ...

(c) Benjamin Bach, in Bach, B., Kerracher, N., Hall, K. W., Carpendale, S., Kennedy, J., & Henry Riche, N. (2016, May). Telling stories about dynamic networks with graph comics. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (pp. 3670-3682). ACM.

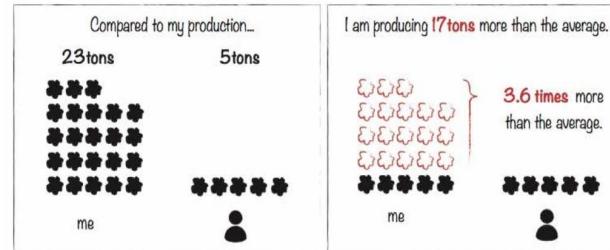
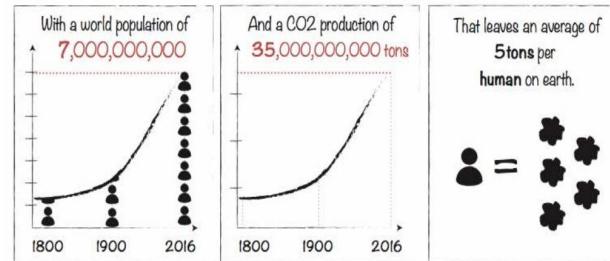
CO Footprint



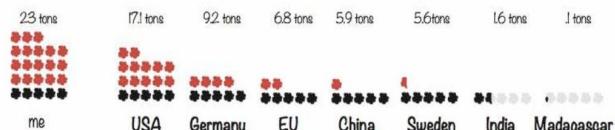
Which produced...



23 tons of CO2.



Thus, my travels in 2016 alone produced more CO2 than the average person in the most countries:



Within the Paris climate accord, the EU and US have pledged to reduce their emissions by approx. 85 % by 2020.



As European, this leaves me 1 trans-atlantic flight a year.



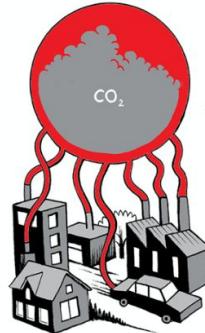
Just that.

no other transports
no computer, no mobile,
no heating,
no clothes,
no food.

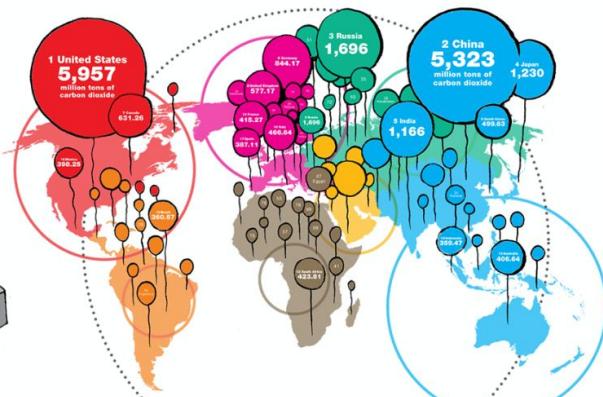
(c) Benjamin Bach

Hot spots – the carbon atlas

Imagine if we could capture all the annual CO₂ emissions of a country in a giant balloon...

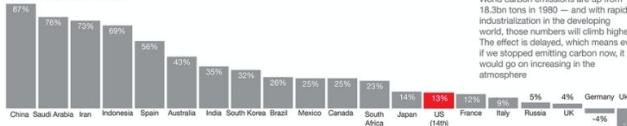


...this is what the world would look like:



Moreover, some counties are taking active steps to curb their CO₂ emissions, while others are raising rapidly:

CO₂ emission growth of the highest 20 emitters, 1995 to 2005



World total

28.19bn tons of CO₂
28% growth in carbon emissions;
1995-2005

World carbon emissions are up from 18.3bn tons in 1990 — and with rapid industrialization in the developing world, they're likely to rise even higher. The effect is delayed, which means even if we stopped emitting carbon now, it would go on increasing in the atmosphere



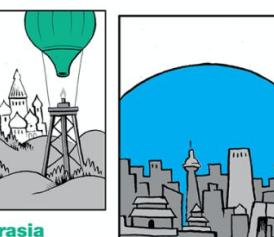
North America
6.99bn tons of CO₂
14% growth in carbon emissions;
1995-2005

The US as a major producer of greenhouse gases has been reluctant to accept responsibility for climate change. It existed — and refused to accept the Kyoto protocol. But fresh weather events and an avalanche of scientific evidence have forced it to rethink its position.



Europe
4.67bn tons of CO₂
9% growth in carbon emissions;
1995-2005

For the first time, there is hard scientific evidence of climate change affecting Europe. The Kyoto protocol was agreed in 1992 to 405 million metric tons in 1990, due to then-deteriorating scientific evidence. Now, the energy giant may make clean up trading carbon credits



Eurasia
2.58bn tons of CO₂
4% growth in carbon emissions;
1995-2005

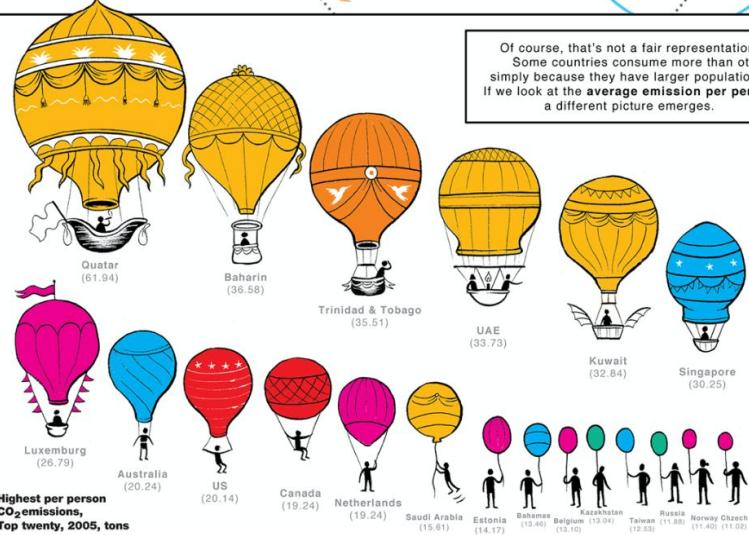
The US' carbon emissions dropped from 1992 to 405 million metric tons in 1990, due to then-deteriorating scientific evidence. Now, the energy giant may make clean up trading carbon credits



Middle East

1.45bn tons of CO₂
62% growth in carbon emission
1995-2005

The region is a major contributor to global greenhouse gas emissions through an oil and gas industry which produces over 30 percent of world oil supply and over 10 percent of its gas



Highest per person CO₂ emissions,
Top twenty, 2005, tons

(c) Matteo Farinella, in Wang, Z., Wang, S., Farinella, M., Murray-Rust, D., Henry Riche, N. and Bach, B., 2019, May. Comparing effectiveness and engagement of data comics and infographics. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).



(c) Terri Po

GUYS LET'S JUST HAVE A CHAT ABOUT

MASS EXTINCTIONS.



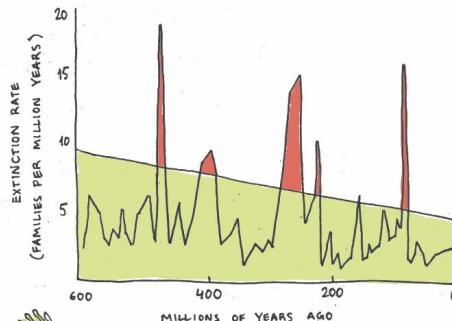
THERE'S GONNA BE
A LOT OF GRAPHS
AND TEXT AND SHIT
SO JUST BE
PREPARED TO
FEEL LIKE THIS.



LET'S START WITH A HISTORY LESSON.

HERE COMES THE FIRST GRAPH.

Source: University of California Museum of Paleontology's Understanding Evolution (<http://evolution.berkeley.edu>)



THE X-AXIS IS
THE TIMELINE.

WE'RE TALKING
LONG TIME PERIODS.
600 000 000 - 0 YEARS AGO.

THE GREEN AREA IS BACKGROUND EXTINCTION. IT'S THE RATE AT WHICH SPECIES DIE OFF NORMALLY, LIKE IT'S NO BIG DEAL.

THIS IS FINE.



THE BIG DEALS ARE THE SPIKES. I'VE MADE THEM RED SO YOU CAN FIND THEM MORE EASILY. THOSE ARE THE MASS EXTINCTION EVENTS.

THAT'S WHEN MORE THAN 75% OF ALL SPECIES ON EARTH HAVE DIED OFF.

SO THAT'S HAPPENED 5 TIMES HERE.



LOOK HERE'S MY ATTEMPT TO VISUALISE THE BIG EXTINCTION EVENTS. YES I MADE IT MYSELF.

AND I'M A FROG.

THERE ARE NUMBERS HERE THAT AREN'T FROM THE PREVIOUS GRAPH.

DON'T FREAK OUT. I USED A DIFFERENT

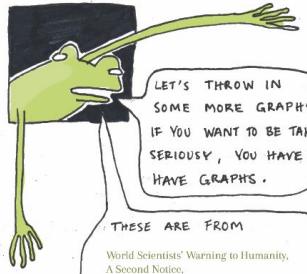
SOURCE : Viviane Richter: "The Big Five Mass Extinctions"
<https://cosmosmagazine.com/paleontology/big-five-extinctions>

ESSENTIALLY IT'S THE SAME THING THOUGH.



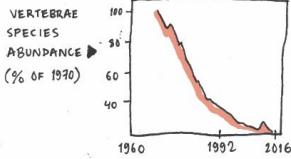
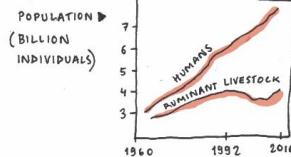
(c) Heidi Tammi

THAT'S RIGHT. THE HUMANS.

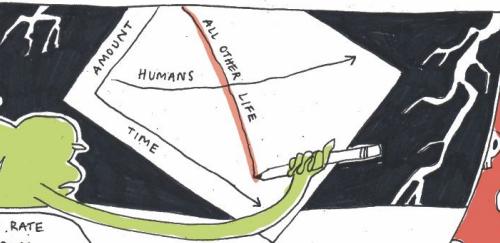


THESE ARE FROM
World Scientists' Warning to Humanity,
A Second Notice,
<https://doi.org/10.1093/biosci/bix125>

IT WAS PUBLISHED IN 2017.
SO IT'S PRETTY RECENT DATA.

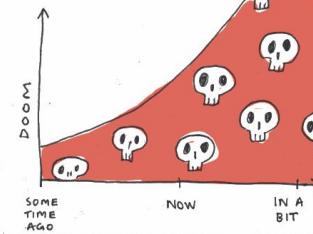


YOU SEE WHERE THIS IS GOING?



THE CURRENT EXTINCTION RATE IS ESTIMATED TO BE FROM 10 TO 1000 TIMES HIGHER THAN IT SHOULD BE (THAT REFERS TO THE BACKGROUND EXTINCTION RATE). MAKING ACCURATE ESTIMATES IS TRICKY...

... BUT THIS IS THE GENERAL CONSENSUS.



IUCN (INTERNATIONAL UNION FOR CONSERVATION OF NATURE) HAS DATA ON SPECIES THAT ARE **THREATENED** WITH EXTINCTION RIGHT NOW.

THAT IS,

NE	DD	LC	NT	VU	EN	CR	EW	EX
----	----	----	----	----	----	----	----	----

SPECIES CLASSIFIED AS VULNERABLE, ENDANGERED OR CRITICALLY ENDANGERED.



HELP TAMMINEN 2016
© Heidi Tamminen

The IUCN Red List of Threatened Species, <https://www.iucnredlist.org/>

University of California Museum of Palaeontology's Understanding Evolution, <http://evolution.berkeley.edu>

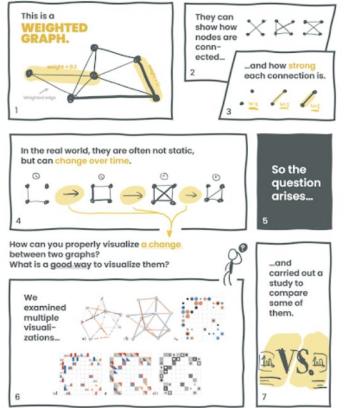
Viviane Richer: The Big Five Mass Extinctions, <https://cosmosmagazine.com/paleontology/big-five-extinctions>

William J. Ripple, Christopher Wolf, Thomas M. Newsome, Mauro Galetti, Mohammed Alamgir, Eileen Crist, Mahmoud I. Mahmoud, William F. Laurance, 15,364 scientist signatures from 184 countries; World Scientists' Warning to Humanity: A Second Notice, BioScience, Volume 67, Issue 12, 1 December 2017, Pages 1026–1028, <https://doi.org/10.1093/biosci/bix125>

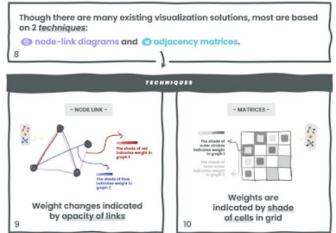
(c) Heidi Tammi

Explaining data analysis <https://statscomics.github.io>

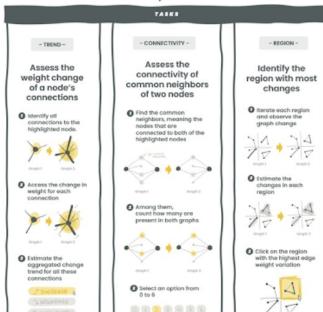
Context, Motivation & Problem Study



Tasks & Conditions



To test which one performs better, we designed three tasks. Tasks like these are frequently used in domains like brain connectivity analysis:



Additionally, we arranged different sizes & densities for the datasets we created.

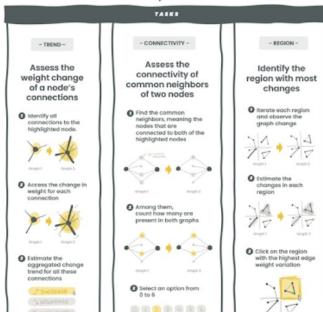
Hypotheses

For each task, we measured performance as number of correct trials. We sought to verify the following hypotheses:

H1: For the trend task, Matrices outperform node-link diagrams for dense networks.

H2: For the connectivity task, Matrices DO NOT outperform node-link diagrams.

To test which one performs better, we designed three tasks. Tasks like these are frequently used in domains like brain connectivity analysis:



For the trend task, Matrices outperform node-link diagrams for dense networks.

For the connectivity task, Matrices DO NOT outperform node-link diagrams.

For the region task, Matrices always outperform node-link diagrams.

Overall, We expect node-link diagrams to decrease in performance for dense datasets.

IV Study Setup, Data Collection & Data Transformation

We recruited 11 participants to complete the study.

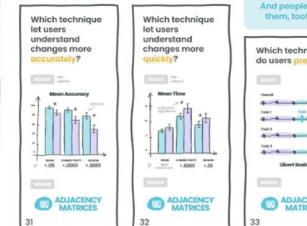
Where we estimate their performance on understanding the weighted graph changes.

The complexity of test problems gradually increases in the test...

A user's responding time will be recorded and log-transformed... At the end, they will also be asked to give answer on which technique they prefer.

V Results

Result shows that Adjacency matrices perform better for visualizing weighted graph changes under most conditions...



VI Hypotheses Evaluation

We expected node-link diagrams to decrease in performance for dense datasets. This is verified true, but we also find it will decrease for large datasets as well.

We thought matrices would outperform node-link diagrams for trend (H1) and Region (H3) tasks on dense datasets... Take note this applies to all tasks across all datasets!

We thought for the Connectivity task, matrices should NOT outperform node-link diagrams... But well, it did.



H4 Supported



H1, H3 Supported



H2 Unsupported

Wang, Z., Ritchie, J., Zhou, J., Chevalier, F. and Bach, B., 2020. Data Comics for Reporting Controlled User Studies in Human-Computer Interaction. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), pp.967-977.

Explaining data analysis

Data Collection & Data Transformation

We recruited 11 participants to complete the study... 

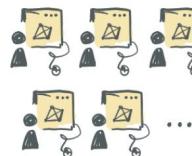
Where we estimate their performance on understanding the weighted graph changes.



We used a full factorial design for the test problems, so every single possible combination was present.

2 TECHNIQUES \times 3 TASKS
 \times 2 SIZES \times 2 DENSITIES

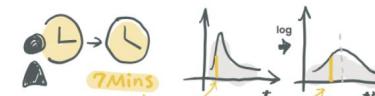
Also, every participant was doing all conditions.
(within-subject study)



The complexity of test problems gradually increases in the test...



A user's responding time will be recorded and log-transformed...



At the end, they will also be asked to give answer on which technique they prefer.

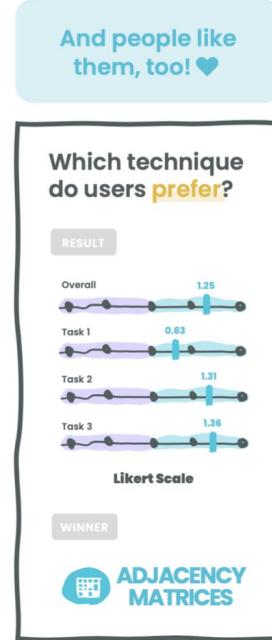
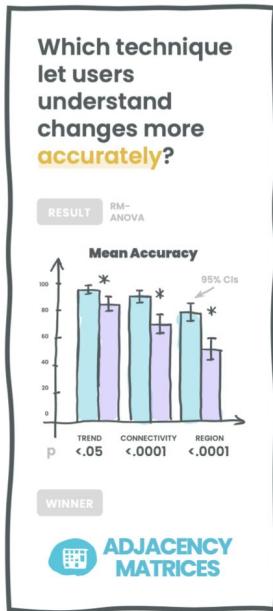


Explaining data analysis

V

Results

Result shows that  **Adjacency matrices** perform better for visualizing weighted graph changes under most conditions...

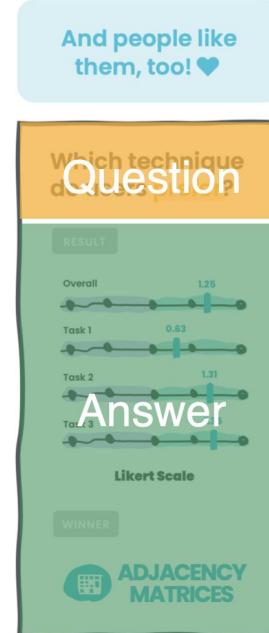
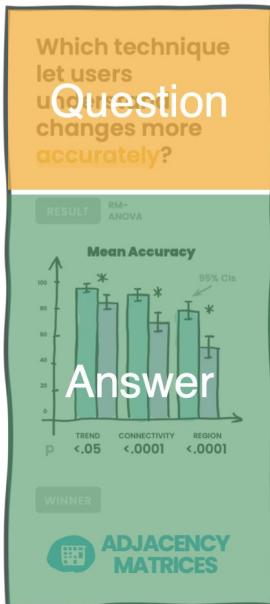


Explaining data analysis

V

Results

Result shows that  **Adjacency matrices** perform better for visualizing weighted graph changes under most conditions...

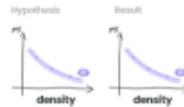


Explaining data analysis

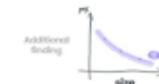
Hypotheses Evaluation

H4

We expected node-link diagrams to decrease in performance for dense datasets.



This is verified true, but we didn't expect it will decrease for large datasets as well.



34



H4 Supported

With additional findings

H1 & H3

We thought matrices would outperform node-link diagrams for Trend (H1) and Region (H3) tasks on dense datasets...



Turns out this applies to all tasks across all datasets!

35

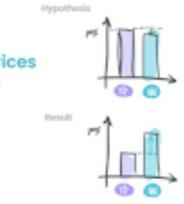


H1, H3 Supported

With additional findings

H2

We thought for the Connectivity task, matrices should NOT outperform node-link diagrams...



But well, it did.

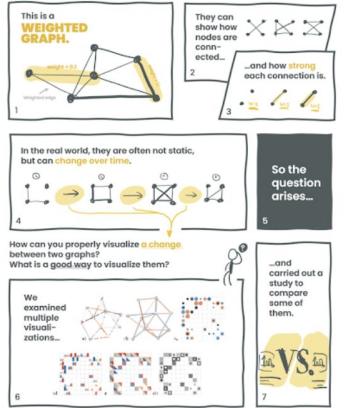
36



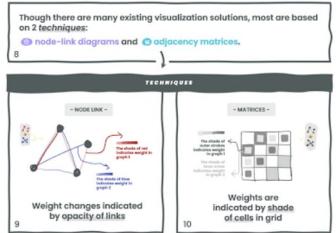
H2 Unsupported

Explaining data analysis <https://statscomics.github.io>

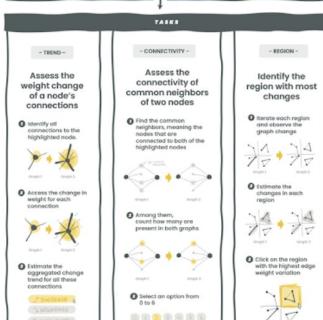
Context, Motivation & Problem Study



Tasks & Conditions



To test which one performs better, we designed three tasks. Tasks like these are frequently used in domains like brain connectivity analysis:



Additionally, we arranged different sizes & densities for the datasets we created.

Hypotheses

For each task, we measured performance as number of correct trials. We sought to verify the following hypotheses:

H1: For the trend task, Matrices outperform node-link diagrams for dense networks.

H2: For the connectivity task, Matrices DO NOT outperform node-link diagrams.

Weight changes indicated by opacity of links

Weights are indicated by shade of cells in grid

For the region task, Matrices always outperform node-link diagrams.

Overall, We expect node-link diagrams to decrease in performance for dense datasets.

IV Study Setup, Data Collection & Data Transformation

We recruited 11 participants to complete the study.

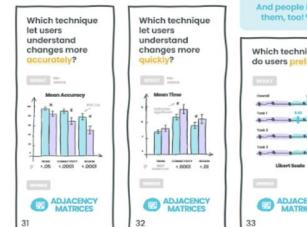
Where we estimate their performance on understanding the weighted graph changes.

The complexity of test problems gradually increases in the test...

A user's responding time will be recorded and log-transformed... At the end, they will also be asked to give answer on which technique they prefer.

V Results

Result shows that Adjacency matrices perform better for visualizing weighted graph changes under most conditions...



VI Hypotheses Evaluation

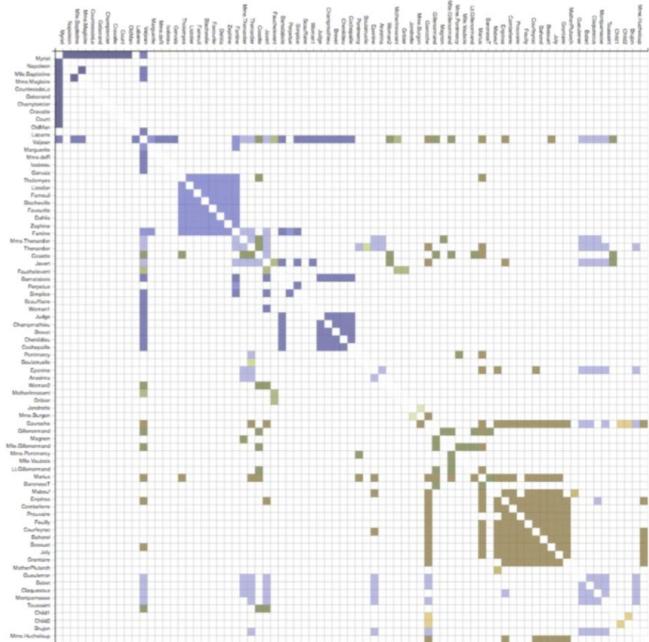
We expected node-link diagrams to decrease in performance for dense datasets. This is verified true, but we also find it will decrease for large datasets as well.

We thought matrices would outperform node-link diagrams for Trend (H1) and Region (H2) tasks on dense datasets... Take note this applies to all tasks across all datasets!

We thought for the Connectivity task, matrices should NOT outperform node-link diagrams... But well, it did.

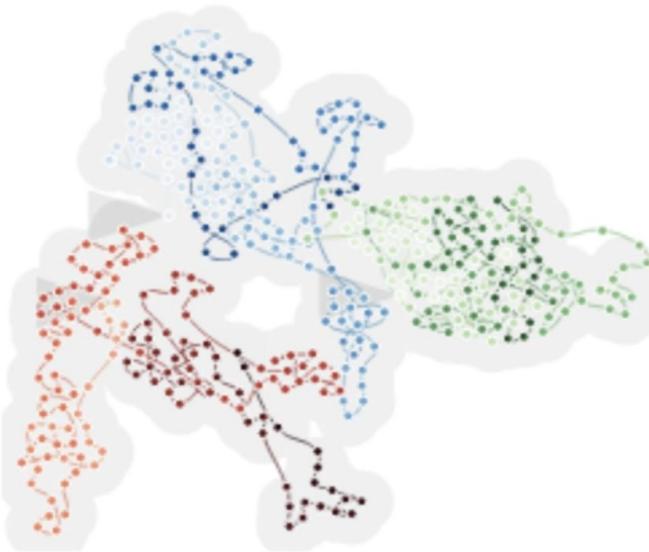


Wang, Z., Ritchie, J., Zhou, J., Chevalier, F. and Bach, B., 2020. Data Comics for Reporting Controlled User Studies in Human-Computer Interaction. *IEEE Transactions on Visualization and Computer Graphics*, 27(2), pp.967-977.



Les Misérables co-occurrence network

<https://github.com/micahstubbs/d3-adjacency-matrix-layout>



Bach, B., Shi, C., Heilot, N., Madhyastha, T., Grabowski, T., & Dragicevic, P. (2015). Time curves: Folding time to visualize patterns of temporal evolution in data. *IEEE transactions on visualization and computer graphics*, 22(1), 559-568.

Find support....

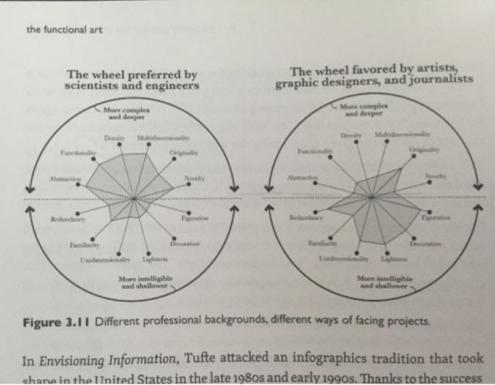
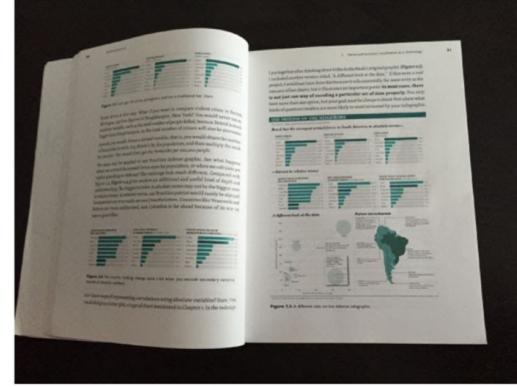


Figure 3.11 Different professional backgrounds, different ways of facing projects.

In *Envisioning Information*, Tufte attacked an infographics tradition that took shape in the United States in the late 1980s and early 1990s. Thanks to the success

The Functional Art by Alberto Cairo



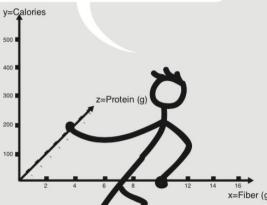
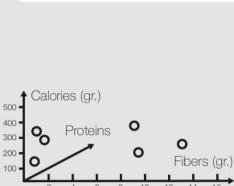
Visual Language for Designers by Connie Malamed

Visualization Cheat Sheets

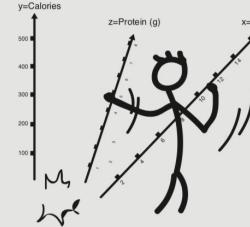


Parallel Coordinates

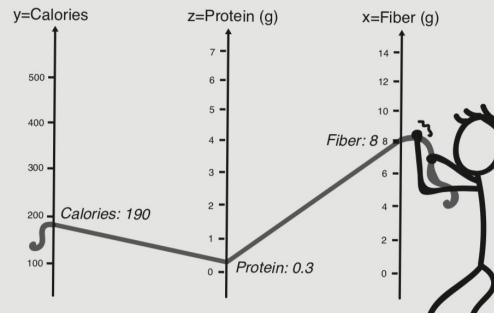
Let's change the arrangement of the axes...



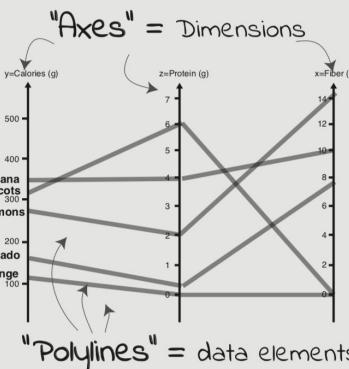
...like so :



We now connect the values for each fruit with a line.



And obtain our Parallel Coordinates Plot (abbreviated: PCP).



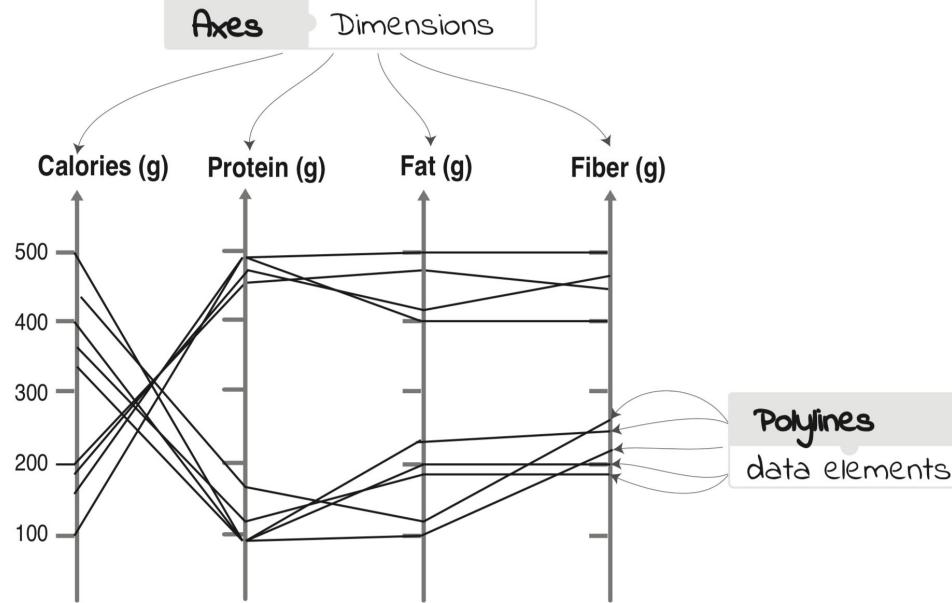
Wang, Z., Sundin, L., Murray-Rust, D. and Bach, B., 2020, April. Cheat Sheets for Data Visualization Techniques. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

Visualization Cheat Sheets



Parallel Coordinates

Anatomy



<https://visualizationcheatsheets.github.io>

Wang, Z., Sundin, L., Murray-Rust, D. and Bach, B., 2020, April. Cheat Sheets for Data Visualization Techniques. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

Visualization Cheat Sheets



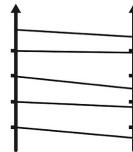
Parallel Coordinates

Parallel lines

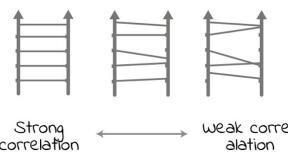
Positive Correlation

Correlations indicate that high values in one data dimension co-occur with high values in another dimension.

Correlations are not causations!



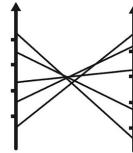
A correlation is visible through rather parallel polylines between two axes.



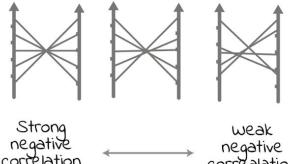
Crossing lines

Negative Correlation

Inverse correlations indicate that high values in one data dimension co-occur with low values in another data dimension.



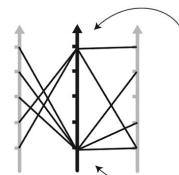
An inverse correlation is visible through lots of crossing polylines between two axes.



Converging lines

Groups

Groups indicate many elements with the same value or similar values.



Groups are visible by many lines intersecting an axis at the same position.

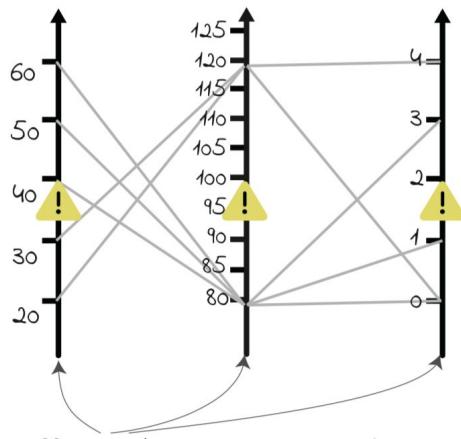
Wang, Z., Sundin, L., Murray-Rust, D. and Bach, B., 2020, April. Cheat Sheets for Data Visualization Techniques. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

Visualization Cheat Sheets



Parallel Coordinates

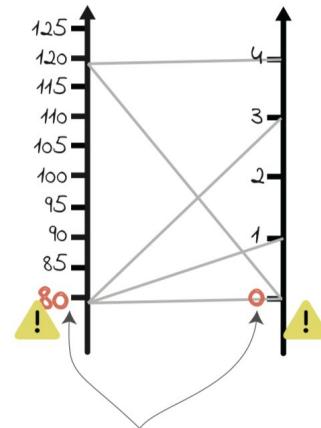
Axis scales



Different dimensions usually have different scales and units.

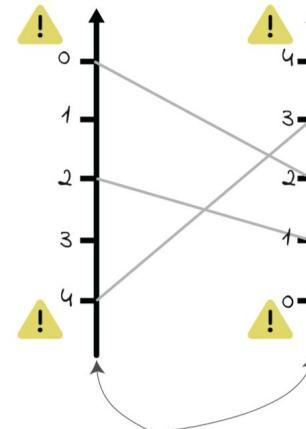
Pitfalls

Truncated axes



Values on axes can start from values other than '0'.

Axes order

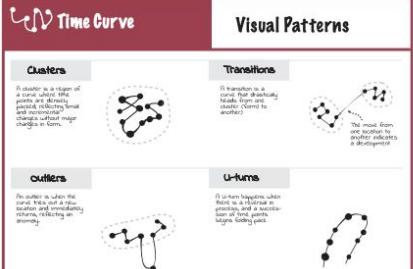
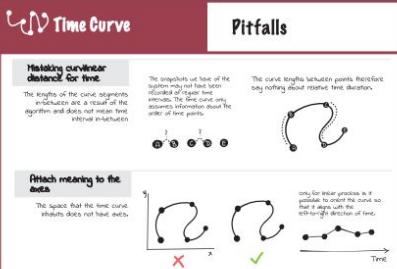
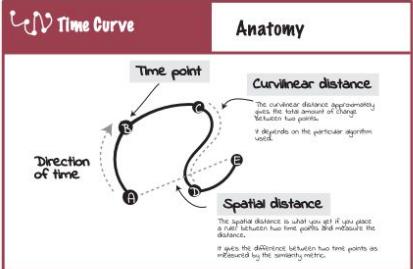


values on axes can be either descending or ascending.



Visualization Cheat Sheets

<https://visualizationcheatsheets.github.io>



Transitions

Outliers

U-turns

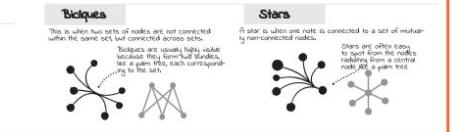
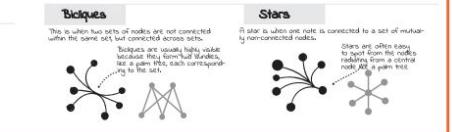
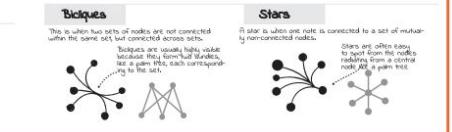
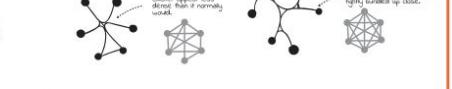
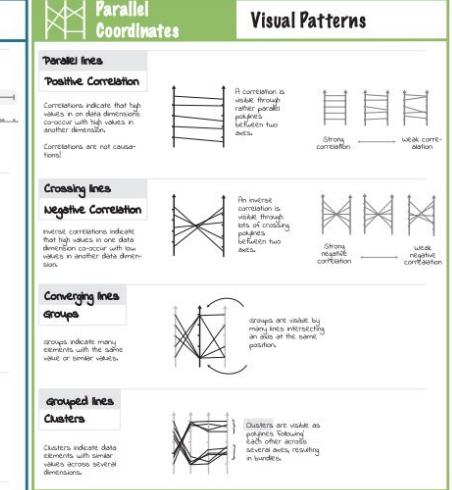
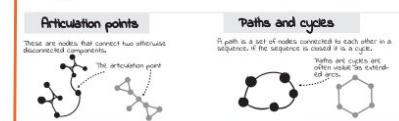
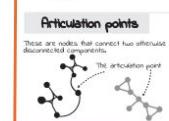
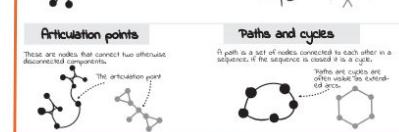
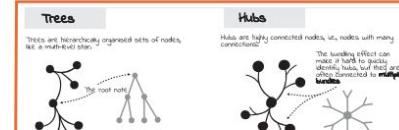
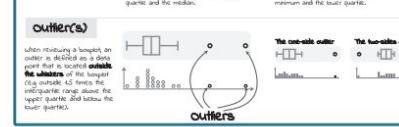
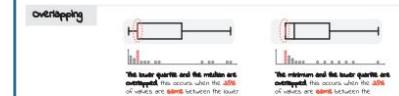
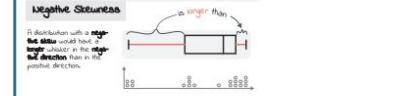
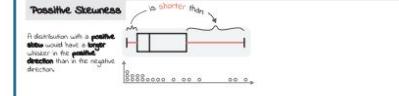
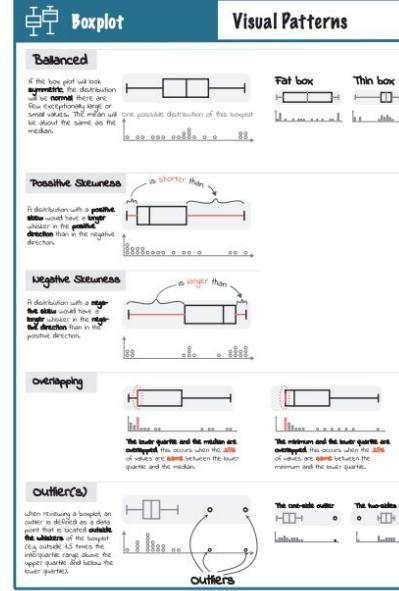
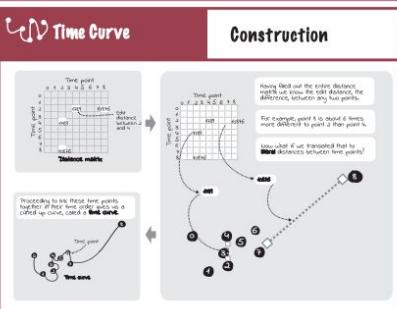
Cycles

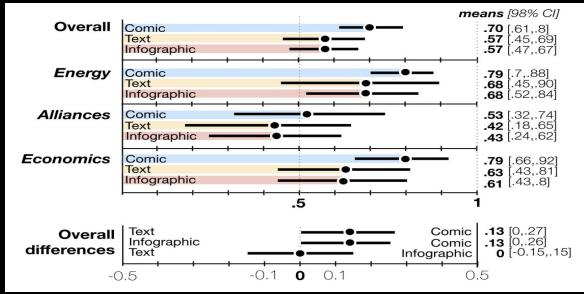
Oscillations

Alternations

Effective development

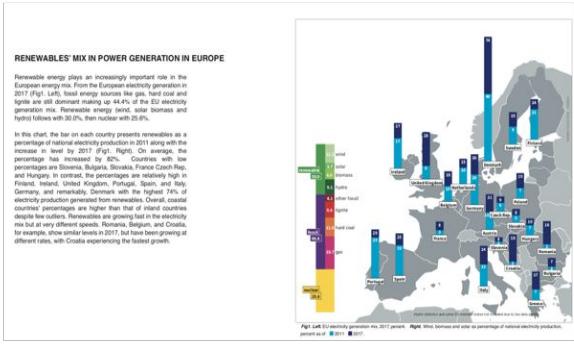
Ineffective development



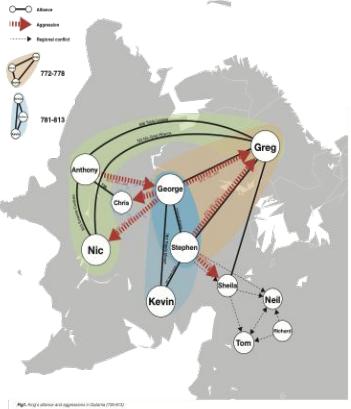


Are data comics effective?

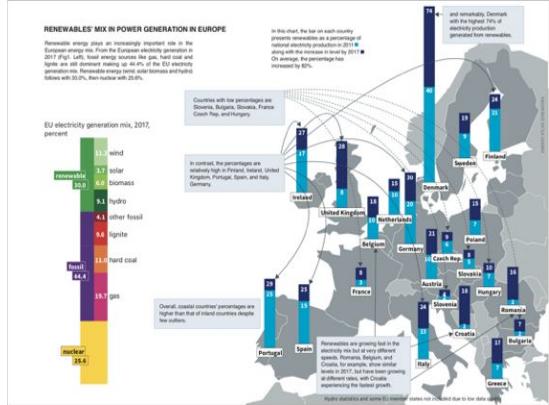
Illustrated Text



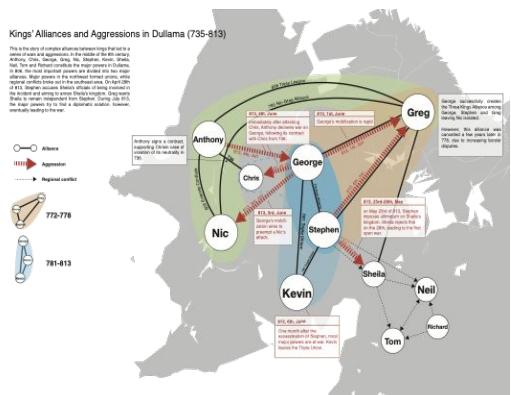
Kings' Alliances and Aggressions in Dullama (735-813)



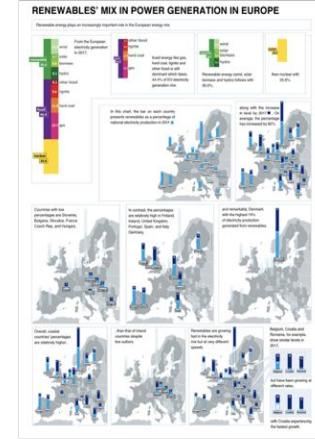
Infographics



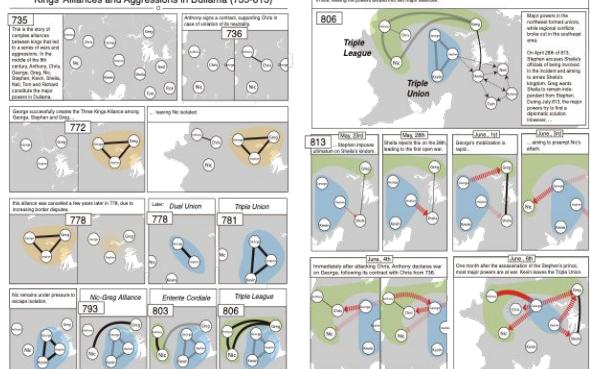
Kings' Alliances and Aggressions in Dullama (735-813)

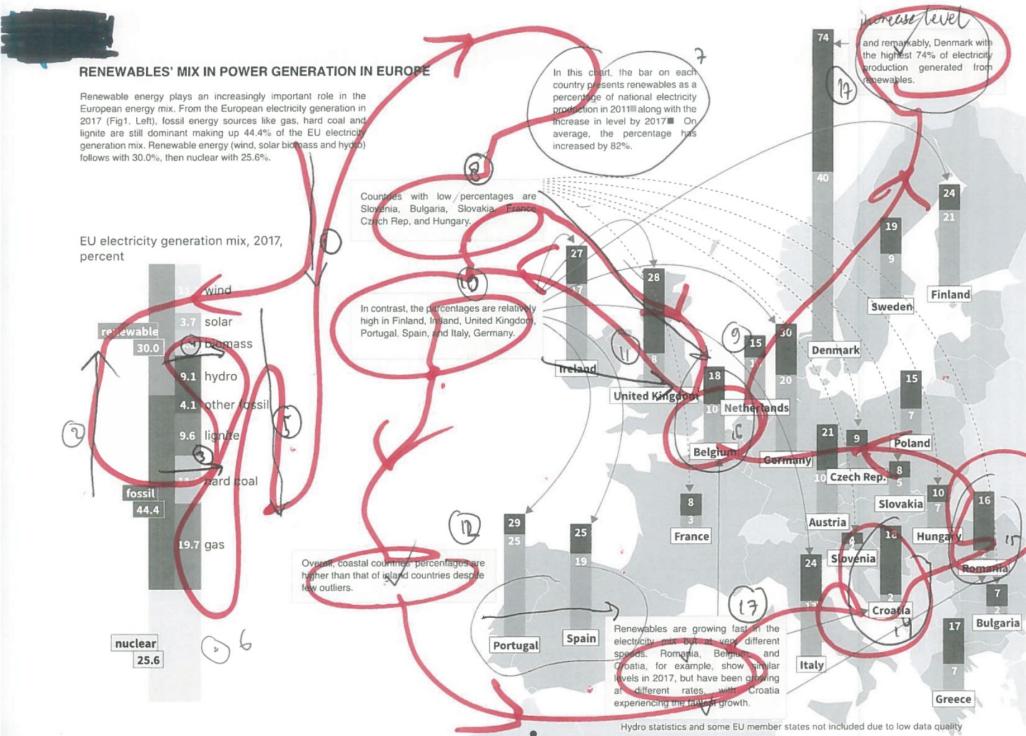


Data comics



Kings' Alliances and Aggressions in Dullama (735-813)





Effectiveness of data comics

Pros:

Visual overview

Guide readers

Adapt to reader pace

Space == narration

Include illustrations besides vis

Memorable layout == narration

Can be shared easily

Effectiveness of data comics

Cons:

Careful with detailed images -> size

Avoid repetition -> highlight changes

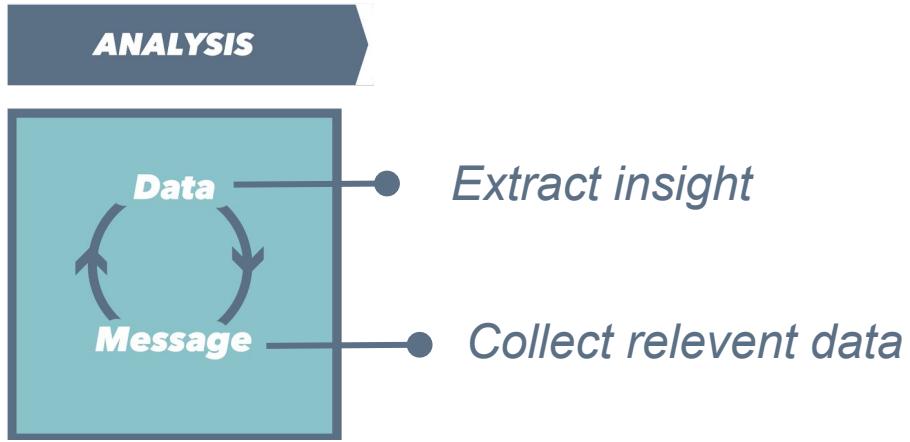
Require space

Non-interactive

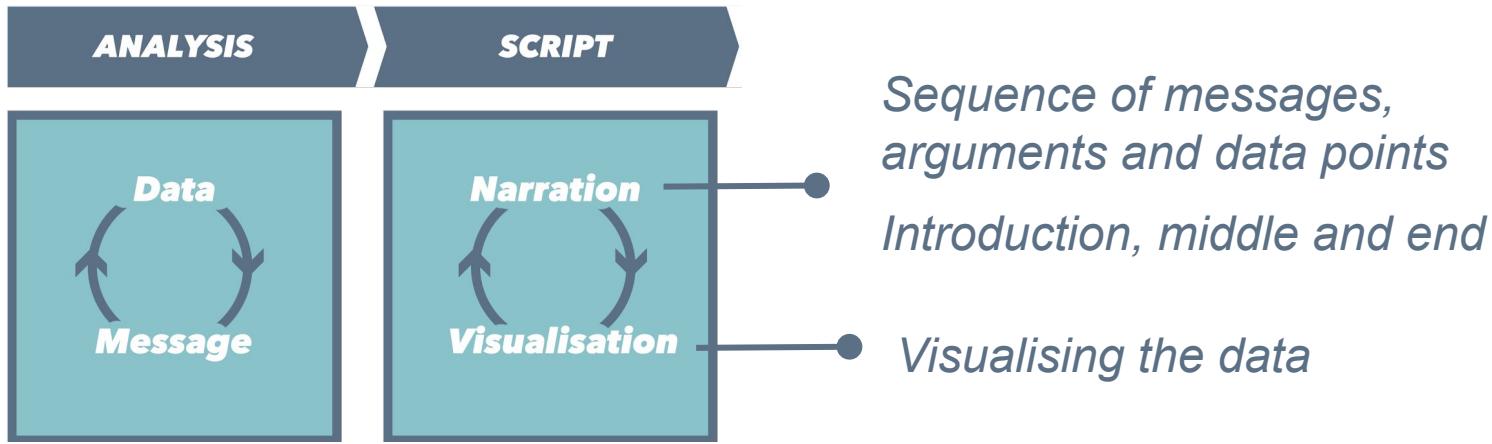


Creating data comics

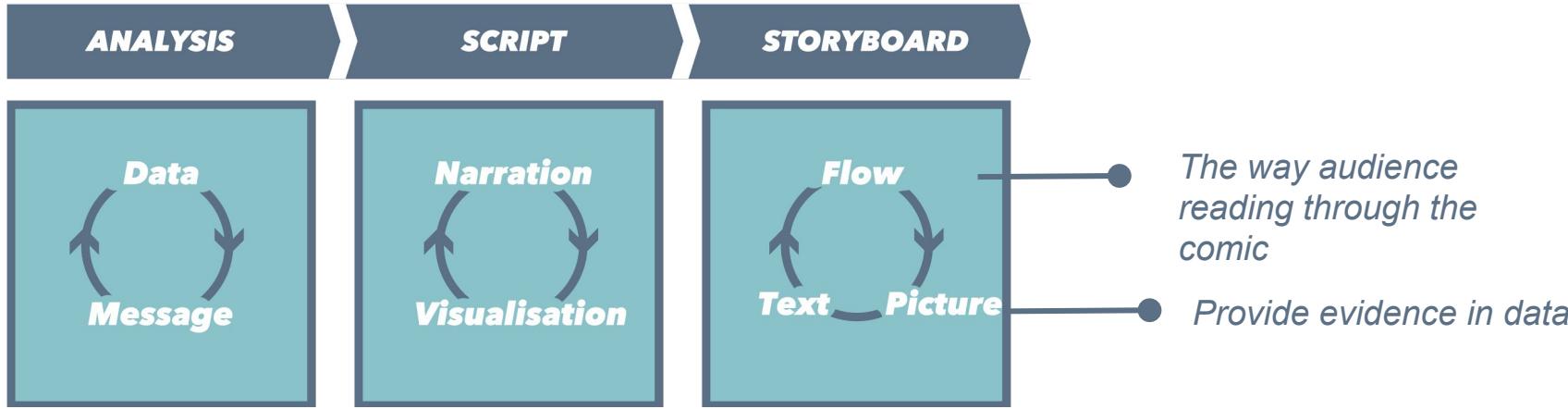
Creation process



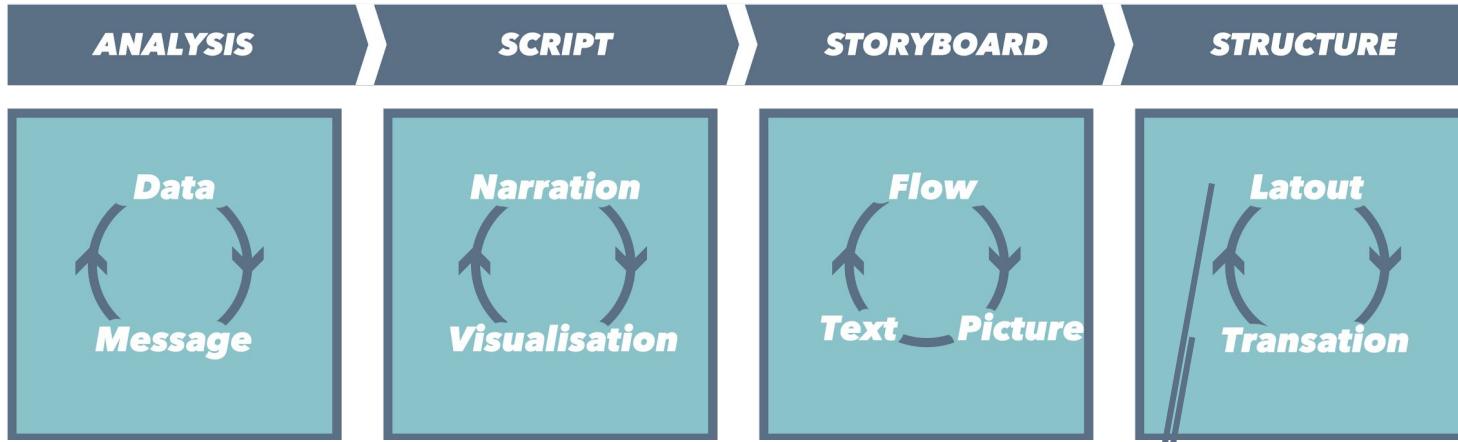
Creation process



Creation process



Creation process



placement of panels / contents in each panel

Connections between information

graphic design

storytelling

Drawing & illustration

storytelling

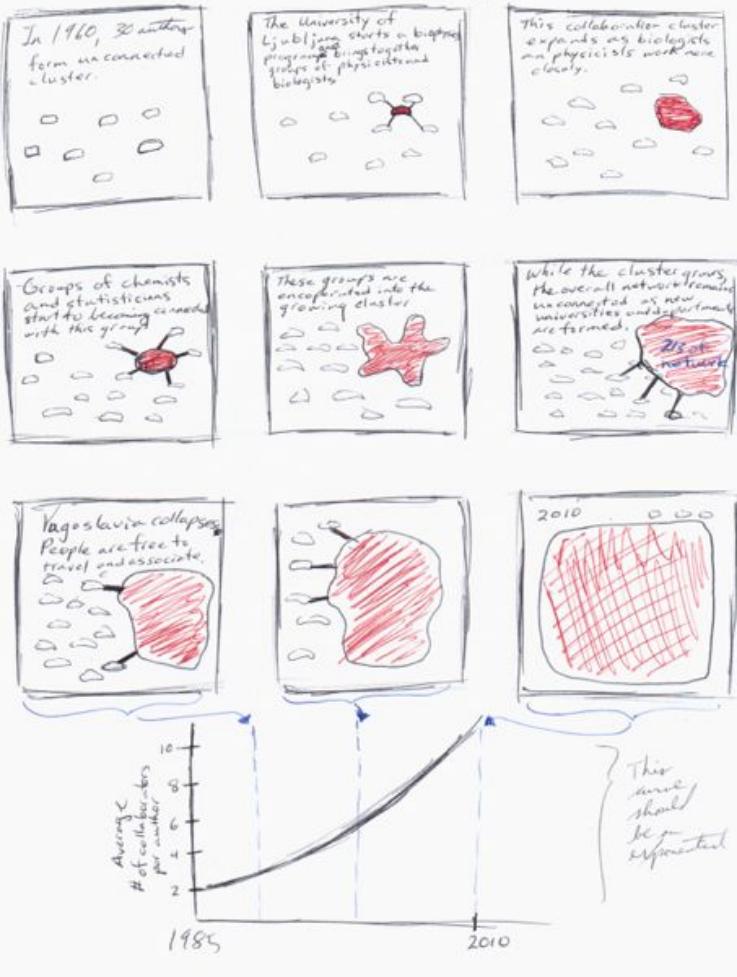
Skills

visualization
design

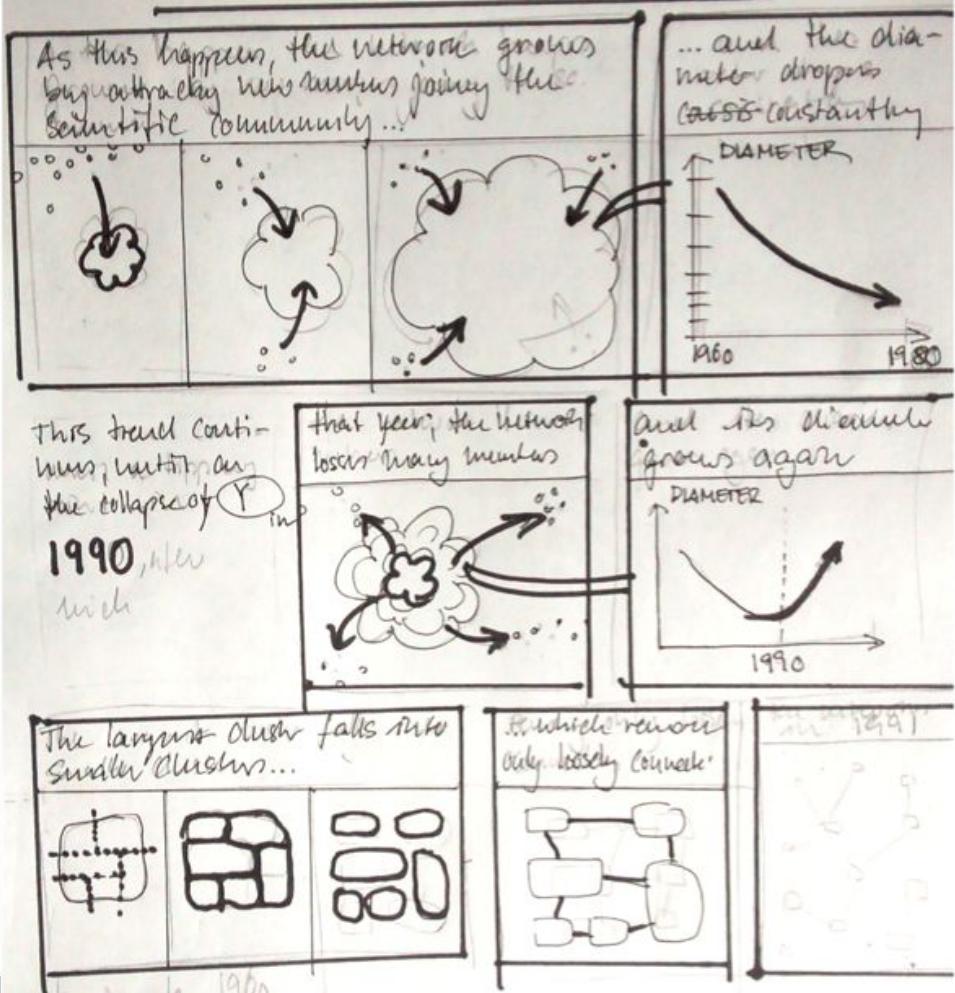
data analysis

creating
visualizations

Slovene Co-authorship (Version 3)



SLOVENE COLL. II



Data Comics Design Patterns

Narrative



Exposé



Multiple-explanations

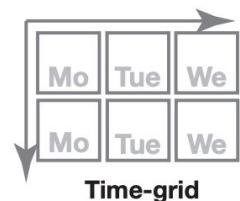


Question & answer

Temporal



Temporal sequence



Time-grid



Time-nesting

Faceting

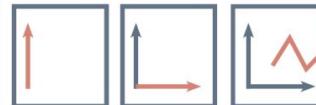


Multiple facets



Contrast

Visual Encoding



Build-up



Legend

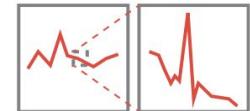


Annotated transition

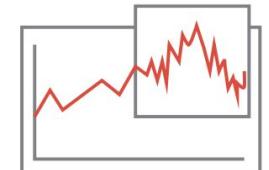
Granular



Zoom



Cut-out



Lens

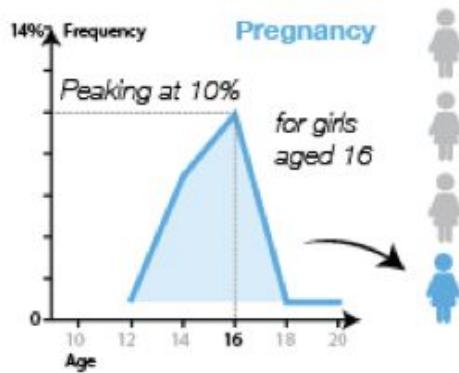
Zoom



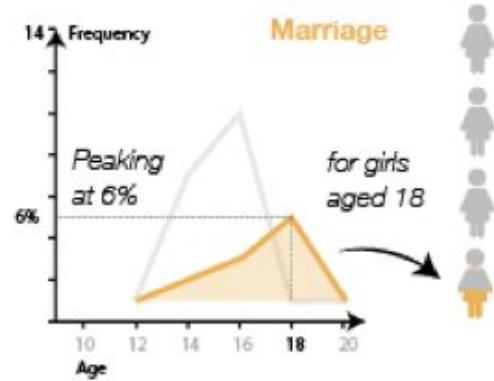
down from 127 in 2010

Multiple Explanation

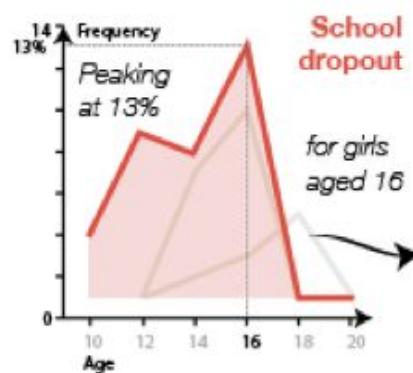
She has a one-in-four risk of becoming pregnant during adolescence,



is at high risk of being engaged in early marriage,



and will likely drop out of school before reaching secondary level.



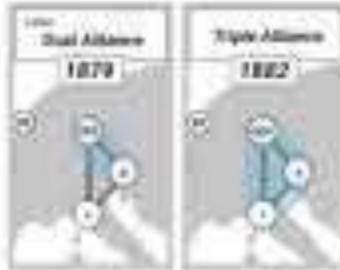
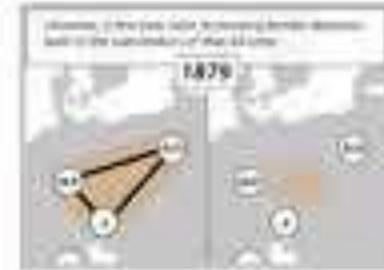
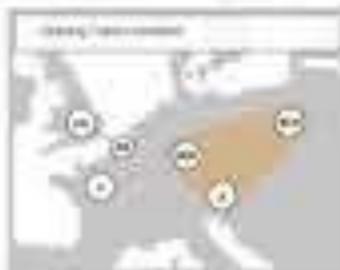
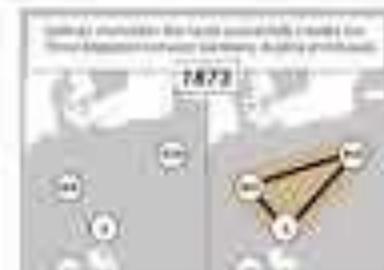
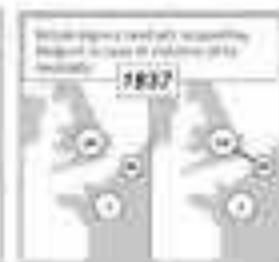
Temporal Change



European Alliances before World War I



Replace



This operation will replace a panel with new panels after doing "click" or "mouseover" on the "Element".

Documentation



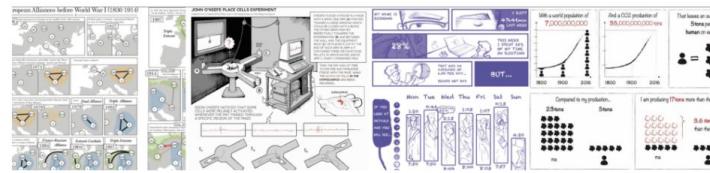
Interactive Data Comics

Bringing interactivity to
data-driven storytelling
with data comics

Home
Publication
Examples
Tutorial
Get Started
Documentation
Tips
Online editor
About



Interactive Data Comics



The screenshot shows a timeline of European alliances from 1834 to 1907-1914. It includes maps of Europe with alliance groups highlighted in orange, blue, and red. A large play button indicates a video component. On the right, there are buttons for 'Replace', 'Watch later', and 'Share'. Below the main content is a note about replacing a panel with new content.

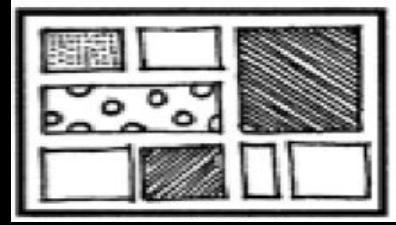
Data comics are a way of effectively communicating with data through data visualizations. They are inspired by the visual language of comics. This project adds interactivity to data comics to **support exploration, explanation, and engagement**.

Interactions are specified using a **JSON** specification and which can be rendered using our [online editor](#).

Interactions include:

- Hiahlaht

<https://interactivedatacomics.github.io/>



Wrapping up

Data Comics

- Combining time and space oriented storytelling
- Familiar to many people
- Easily accessible through many media
- Widely applicable
- Effective for breaking down complexity
- Huge design space: expression, style, ...
- Design patterns to help creation
- Data-driven storytelling *is* complicated

Data Comics

- Combining time and space oriented storytelling
- Familiar to many people
- Easily accessible through many media
- Widely applicable
- Effective for breaking down complexity
- Huge design space: expression, style, ...
- Design patterns to help creation
- Data-driven storytelling *is* complicated

Data Comics

- Combining time and space oriented storytelling
- Familiar to many people
- Easily accessible through many media
- Widely applicable
- Effective for breaking down complexity
- Huge design space: expression, style, ...
- Design patterns to help creation
- Data-driven storytelling *is* complicated

Data Comics

- Combining time and space oriented storytelling
- Familiar to many people
- Easily accessible through many media
- Widely applicable
- Effective for breaking down complexity
- Huge design space: expression, style, ...
- Design patterns to help creation
- Data-driven storytelling *is* complicated

Data Comics

- Combining time and space oriented storytelling
- Familiar to many people
- Easily accessible through many media
- Widely applicable
- Effective for breaking down complexity
- Huge design space: expression, style, ...
- Design patterns to help creation
- Data-driven storytelling *is* complicated

Data Comics

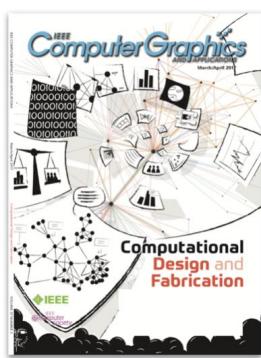
- Combining time and space oriented storytelling
- Familiar to many people
- Easily accessible through many media
- Widely applicable
- Effective for breaking down complexity
- Huge design space: expression, style, ...
- Design patterns to help creation
- Data-driven storytelling *is* complicated

Data Comics

- Combining time and space oriented storytelling
- Familiar to many people
- Easily accessible through many media
- Widely applicable
- Effective for breaking down complexity
- Huge design space: expression, style, ...
- Design patterns to help creation
- Data-driven storytelling *is* complicated

Data Comics

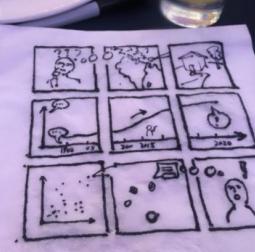
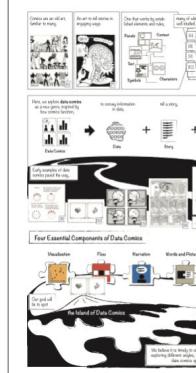
- Combining time and space oriented storytelling
- Familiar to many people
- Easily accessible through many media
- Widely applicable
- Effective for breaking down complexity
- Huge design space: expression, style, ...
- Design patterns to help creation
- Data-driven storytelling *is* complicated



The Emerging Genre of Data Comics

Benjamin Bach, Harvard University
Nathalie Henry-Riche, Microsoft Research
Sheelaagh Carpendale, University of Harvard University
[Download PDF](#)

This article explores the emerging genre of data comics, which combine data visualization with narrative and humor. It discusses the four essential components of data comics: flow, narration, words and pictures, and verbal.



Napkin Comic

Zhezhong Wang



United States of Plastic

The Guardian

The rich drive up the mean far above what the typical income is. Therefore, statisticians instead let the median represent the average. While the mean is the mid-point that splits the total amount of income into equal halves..

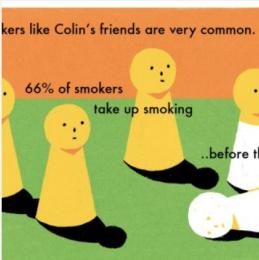
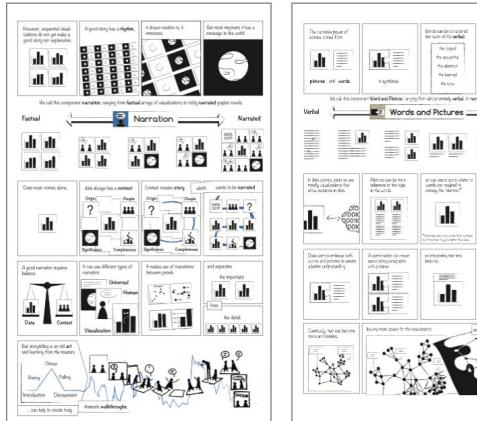
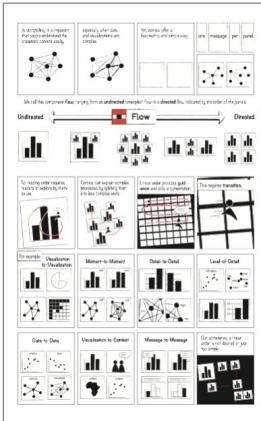


the median splits the population into equal halves, even though above-median people have a much higher income in total.



Child Poverty

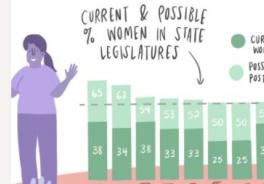
Lovisa Sundin



Smoking

Terri Po

CORD NUMBERS OF WOMEN ARE WINNING FOR OFFICE... AND NO THEY'RE WINNING PRIMARIES.



The Political Future is Female

Olivia Walch - The NIB

China's rise is so significant because of its sheer enormity. If we talk to a random person on Earth, you likely speak to a native Chinese speaker.



The Future Sounds Chinese

Josh Kramer - The NIB

Further Reading

McCloud, Scott, and A. D. Manning. "Understanding comics: The invisible art." *IEEE Transactions on Professional Communications* 41.1 (1998): 66-69.

Farinella, M., 2018. The potential of comics in science communication. *Journal of science communication*, 17(01), pp.Y01-01.

Zhao, Z., Marr, R. and Elmquist, N., 2015. Data comics: Sequential

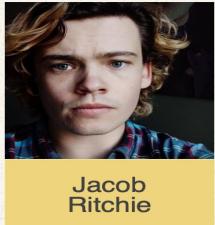
Bach, B., Riche, N.H., Carpendale, S. and Pfister, H., 2017. The emerging genre of data comics. *IEEE computer graphics and applications*, 37(3), pp.6-13.

Bach, B., Wang, Z., Farinella, M., Murray-Rust, D. and Henry Riche, N., 2018, April. Design patterns for data comics. In Proceedings of the 2018 chi conference on human factors in computing systems (pp. 1-12).

Wang, Z., Wang, S., Farinella, M., Murray-Rust, D., Henry Riche, N. and Bach, B., 2019, May. Comparing Effectiveness and Engagement of Data Comics and Infographics. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).

Wang, Z., Sundin, L., Murray-Rust, D. and Bach, B., 2020, April. Cheat Sheets for Data Visualization Techniques. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-13).

Wang, Z., Ritchie, J., Zhou, J., Chevalier, F. and Bach, B., 2020. Data Comics for Reporting Controlled User Studies in Human-Computer Interaction.



<http://visualinteractive.data.github.io>

Data Comics for Data-Driven Storytelling

Benjamin Bach

University of Edinburgh

