

Infografiikan alkeet

AYS 18.2.2020

Julius Uusikylä
julius@krut.fi

KRUT
COLLECTIVE

“The greatest value of a picture is when it forces us to notice what we never expected to see.”

– John W. Tukey, Exploratory data analysis



Julius Uusikylä

Perustin tietomuotoilutoimisto KRUTin 2018 yhdessä AD Måns Eklundin kanssa. Teemme tietomuotoilua ja visuaalista suunnittelua:
infografiikat, dashboardit, visuaalinen ongelmanratkaisu, raportti- ja aikakauslehtidesign ja -taitto, verkkosivuja, visuaalisia identiteettejä.
Ks. krut.fi

Yhteiskuntatieteiden maisteri (Media & Global communication HY), visuaalinen journalismi Aalto, journalismin kandi (Soc&kom)

Uutisgraafikon ja visuaalisen journalismin töitä, mm. HBL ja freelancer-työt.

Termit haltuun

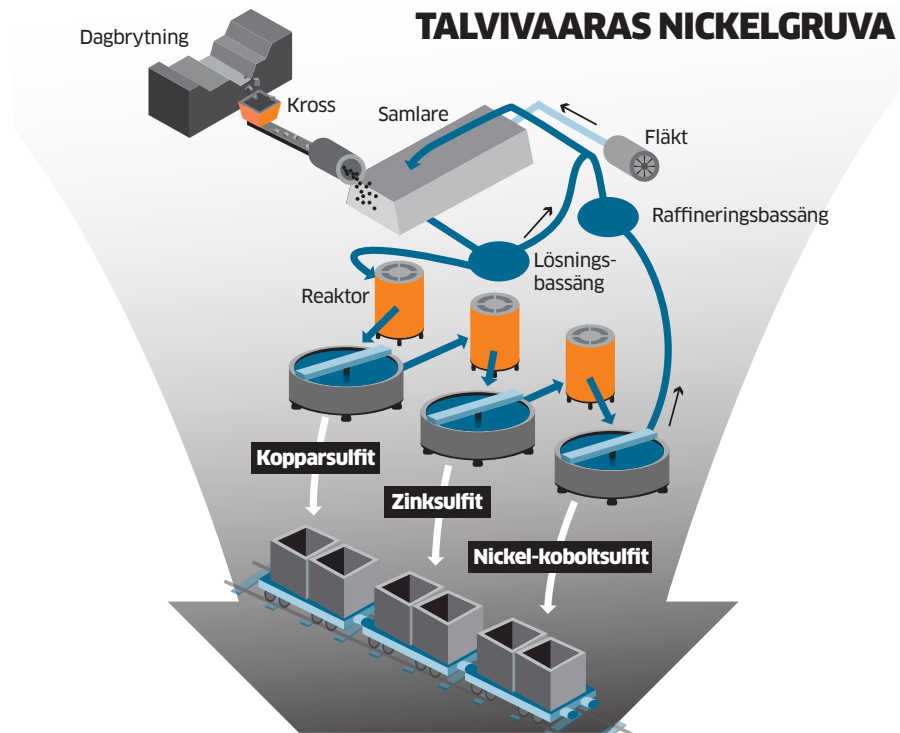
Infograafit, kaaviot,
diagrammit, ikonit ...

--> Tiedon visualisointi

Selittävät kuvat

VS

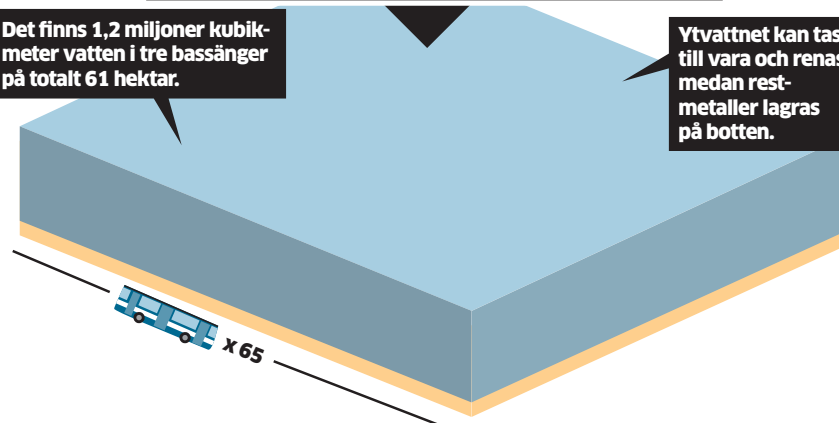
Datavisualisoinnit



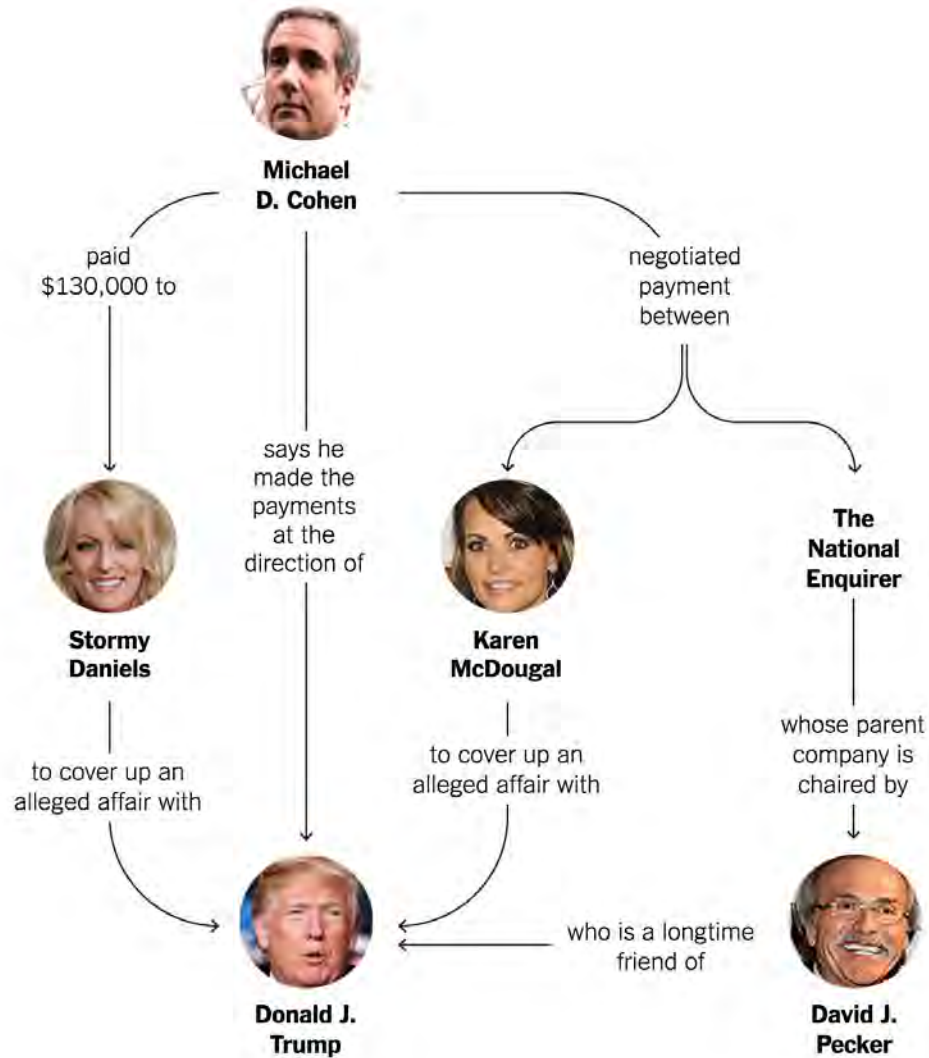
Under hela processen uppstår surt avfallsvatten som bland annat innehåller metaller. Det samlas i en bassäng för gipsfällning, som i praktiken är gruvans soptipp.

Det finns 1,2 miljoner kubikmeter vatten i tre bassänger på totalt 61 hektar.

Ytvattnet kan tas till vara och renas medan restmetaller lagras på botten.



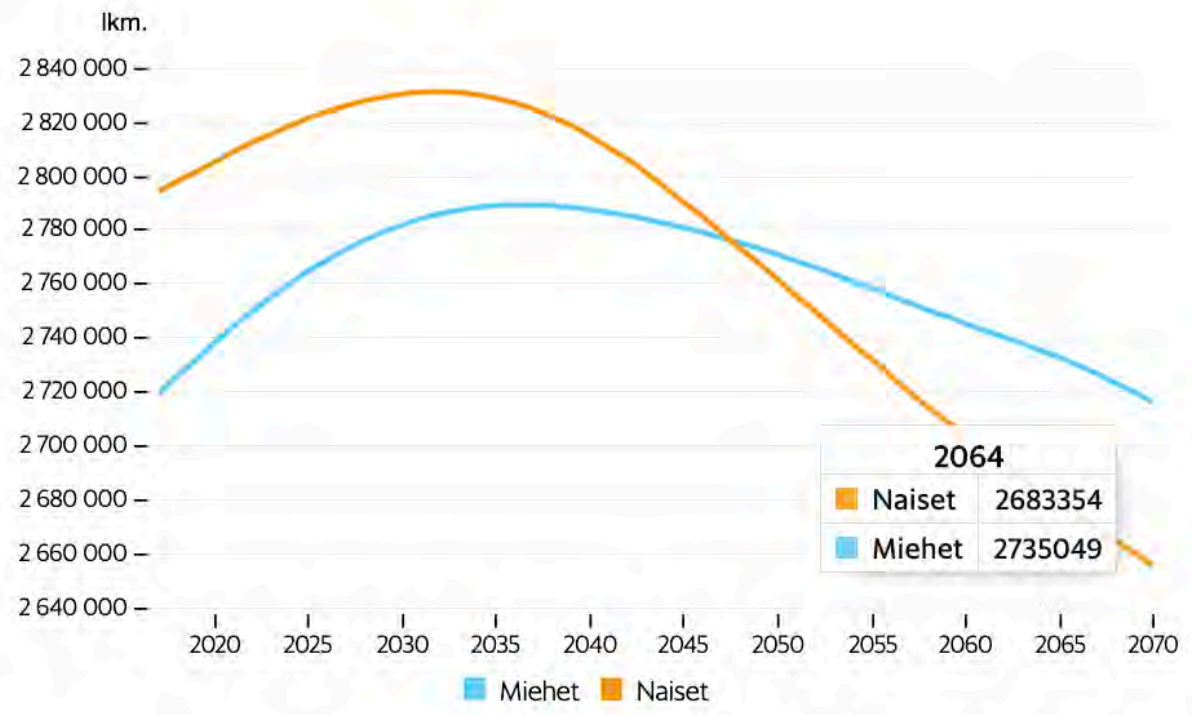
Cohen says he arranged illegal hush money payments “at the direction” of Trump and that Trump directed him to lie about them.



7

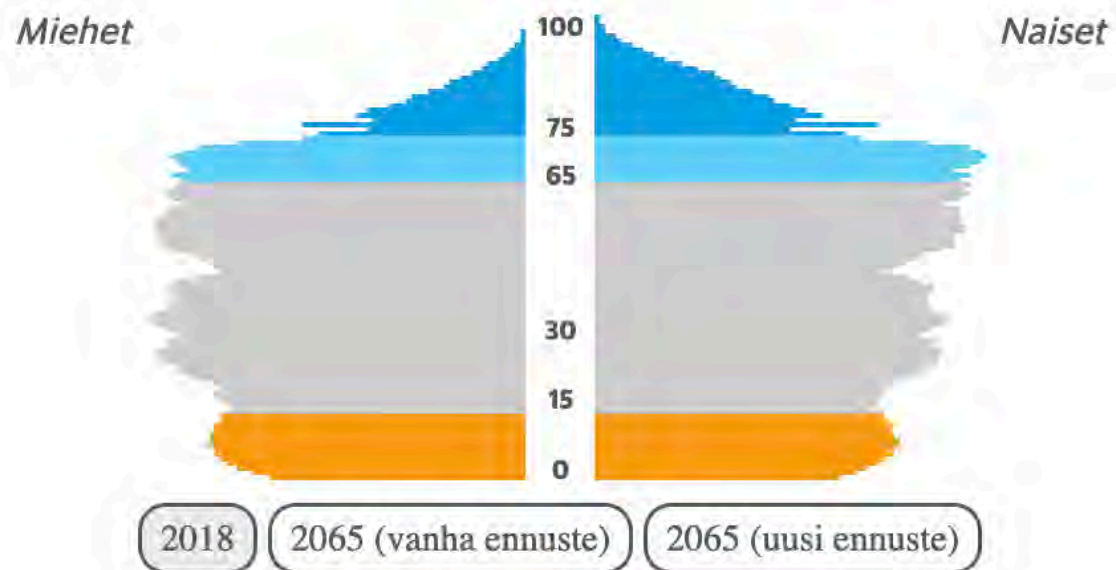
Väestöennuste 2070 asti

Sukupuolittain



Lähde: Tilastokeskus

Niin sanottu ikäpyramidi havainnollistaa, kuinka suomalaisten ikärakenne vanhenee. Kun syntyvät ikäluokat jäävät yhä pienemmiksi, pyramidin pohja kapenee.



<https://www.hs.fi/kotimaa/art-2000005901427.html>

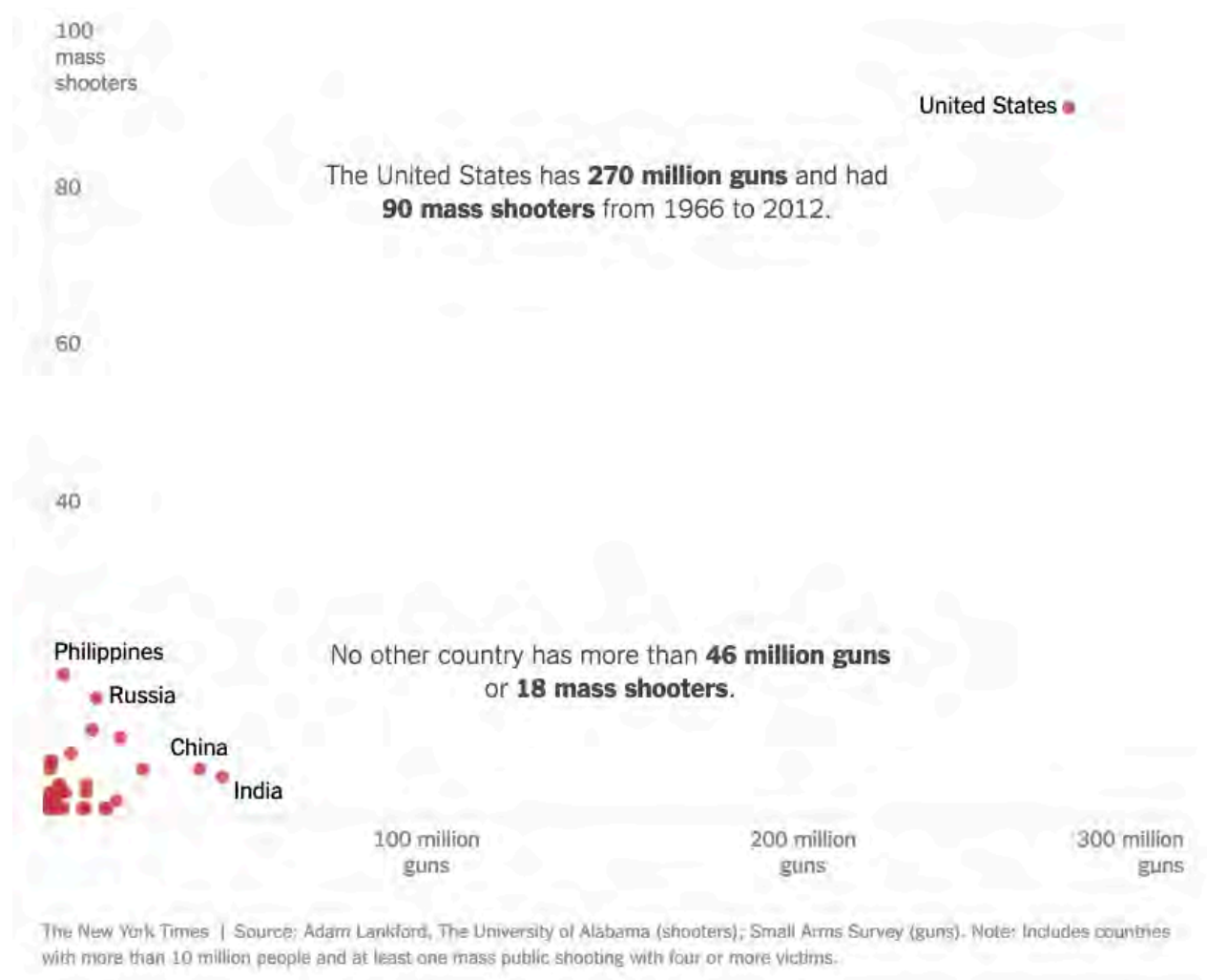
Kaaviolukutaito & data literacy

10

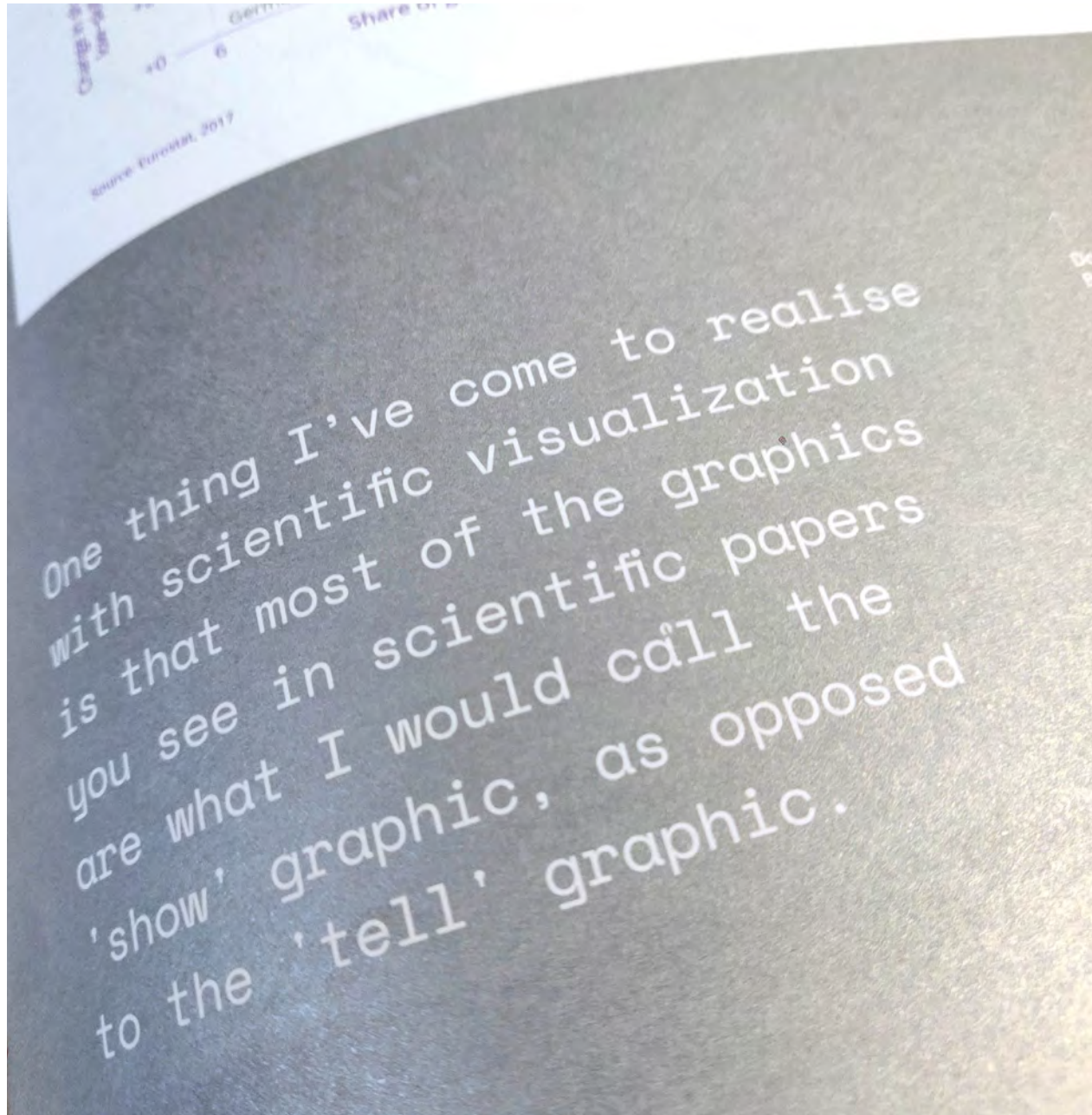
“The impact of transforming technical 70-page documents into shorter, focused, visually led pieces was remarkable, even resulting in infographics being used in Congress during budget debates.”

The Economist: <https://t.co/CgZ3fy8jSj>

“63 per cent of American adults can correctly interpret a scatter plot, a relatively common type of chart for showing the relationship between two variables.”



<https://www.nytimes.com/2017/11/07/world/americas/mass-shootings-us-international.html>



John Burn-
Murdoch,

Market Cafe
Magazine
Issue 4

Osoita yleisölle tuloksesi

“I **encourage researchers to produce more powerful graphics** that can get their often critically important findings out to more people, more quickly, and in more memorable formats. Of course, one could argue that is a job for visual journalists – people like me.

[...]

In my experience, **invaluable scientific research is often still completely missed** by mainstream media outlets, or only discovered months after the fact. This is terrible, and if one part of solving that problem is for researchers to use engaging data visualizations to get their work straight out into the world, that would **unleash a huge public benefit.**”

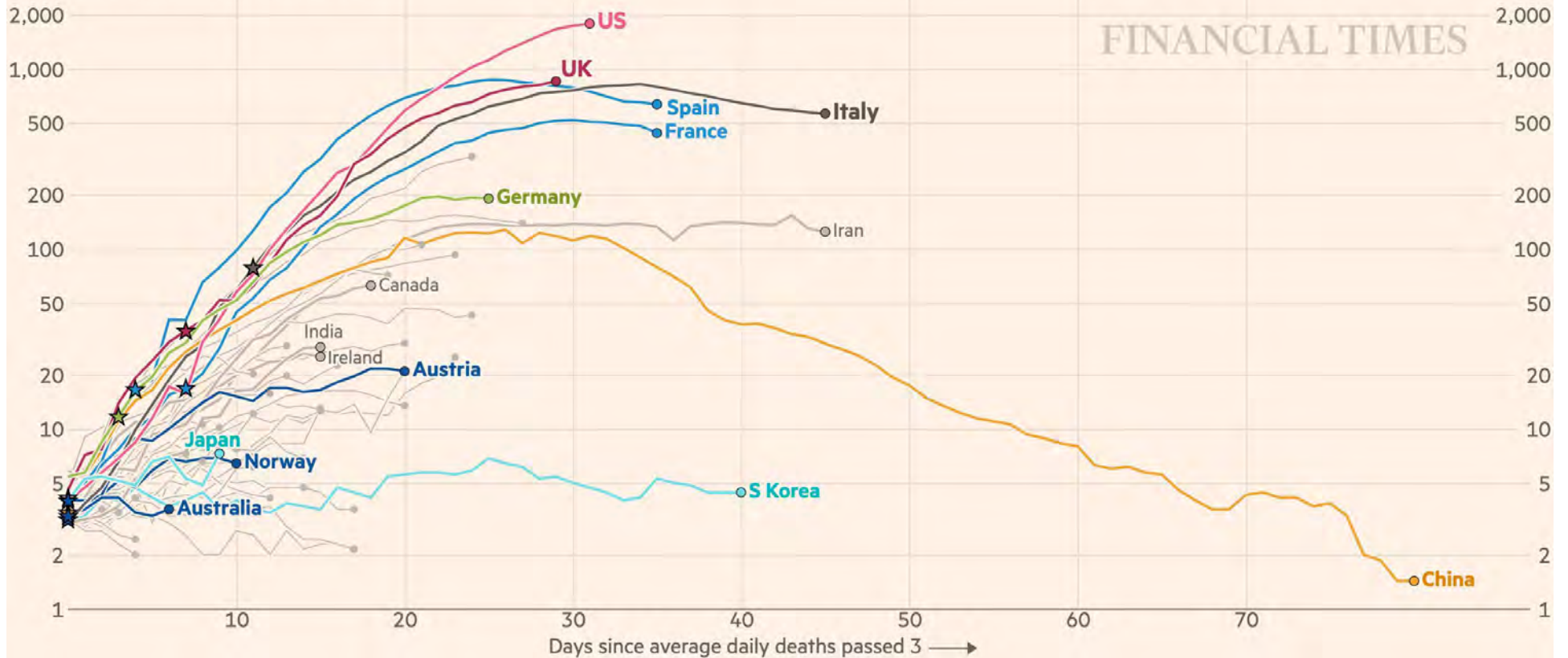
John Burn-Murdoch, Market Cafe Magazine Issue 4

Maailman ensimmäinen datavisualisointi?



Italy and Spain's daily death tolls are falling; in the UK and US daily deaths still trend upward

Daily deaths with coronavirus (7-day rolling average), by number of days since 3 daily deaths first recorded
Stars represent national lockdowns ★



FT graphic: John Burn-Murdoch / @jburnmurdoch
Source: FT analysis of European Centre for Disease Prevention and Control; FT research. Data updated April 13, 19:00 GMT
© FT

<https://medium.com/nightingale/how-john-burn-murdochs-influential-dataviz-helped-the-world-understand-coronavirus-6cb4a09795ae>

Laadukas visualisointi...

Alberto
Cairo,
The truthful
art.

Käännökset
omiani.

Truthful

... on totuudenmukainen

Functional

... kuvaa datan tarkasti ja auttaa yleisöä tekemään mielekkäitä johtopäätöksiä

Beautiful

... on huokutteleva ja esteettisesti puhutteleva yleisölleen

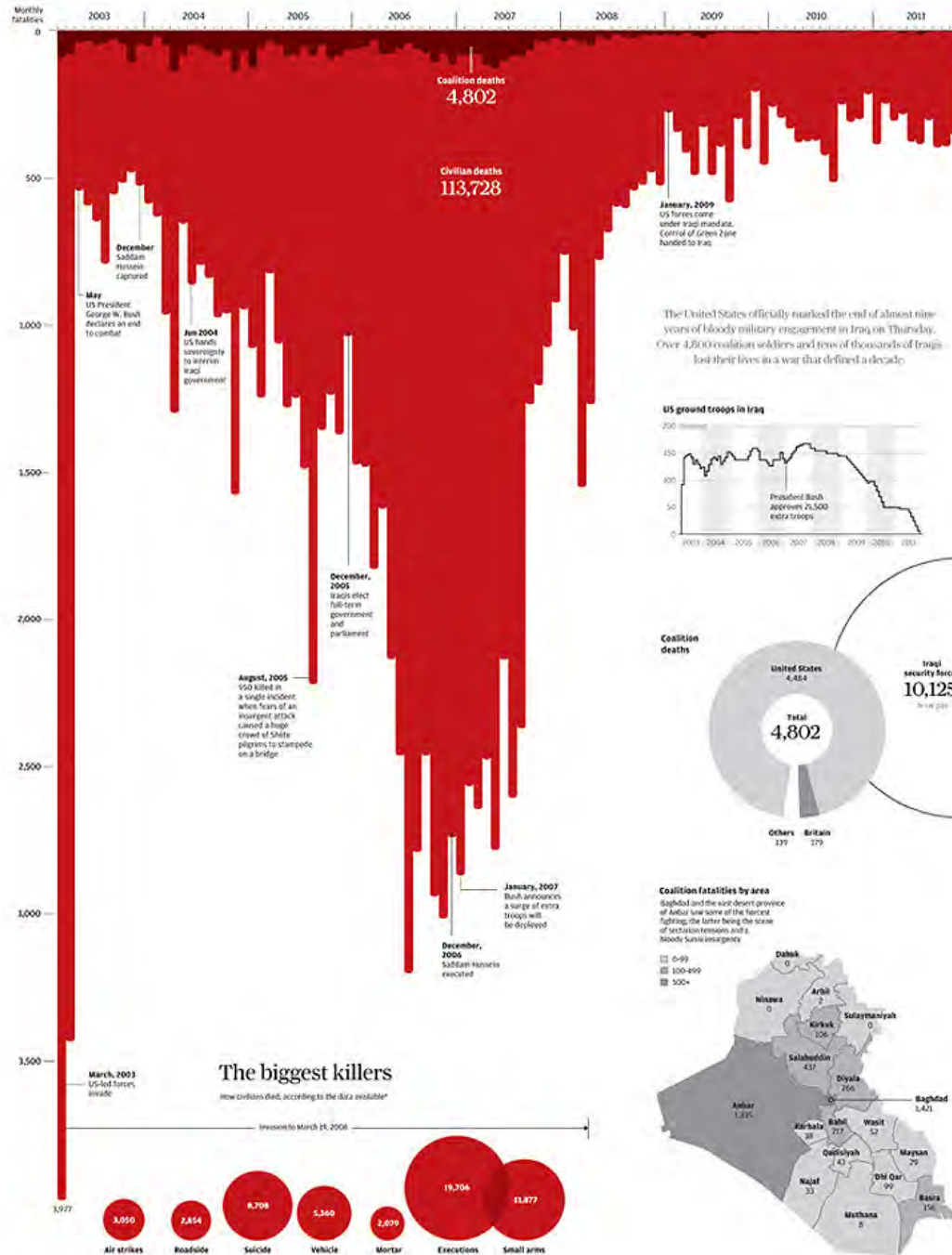
Insightful

... auttaa lukijaa uusien oivallusten äärelle (Tukey!)

Enlightening

... valistaa. “Jos ymmärrämme ja hyväksymme ne todisteet joita visualisointi esittelee, tämä muuttaa ajatteluamme parempaan päin.”

Iraq's bloody toll



South China
Morning
Post.

*Data from the Iraq Body Count project, which is a non-profit organization that tracks civilian deaths in Iraq. The data is based on reports from the public and is not a complete record of all deaths. The data is also subject to change as more information becomes available.

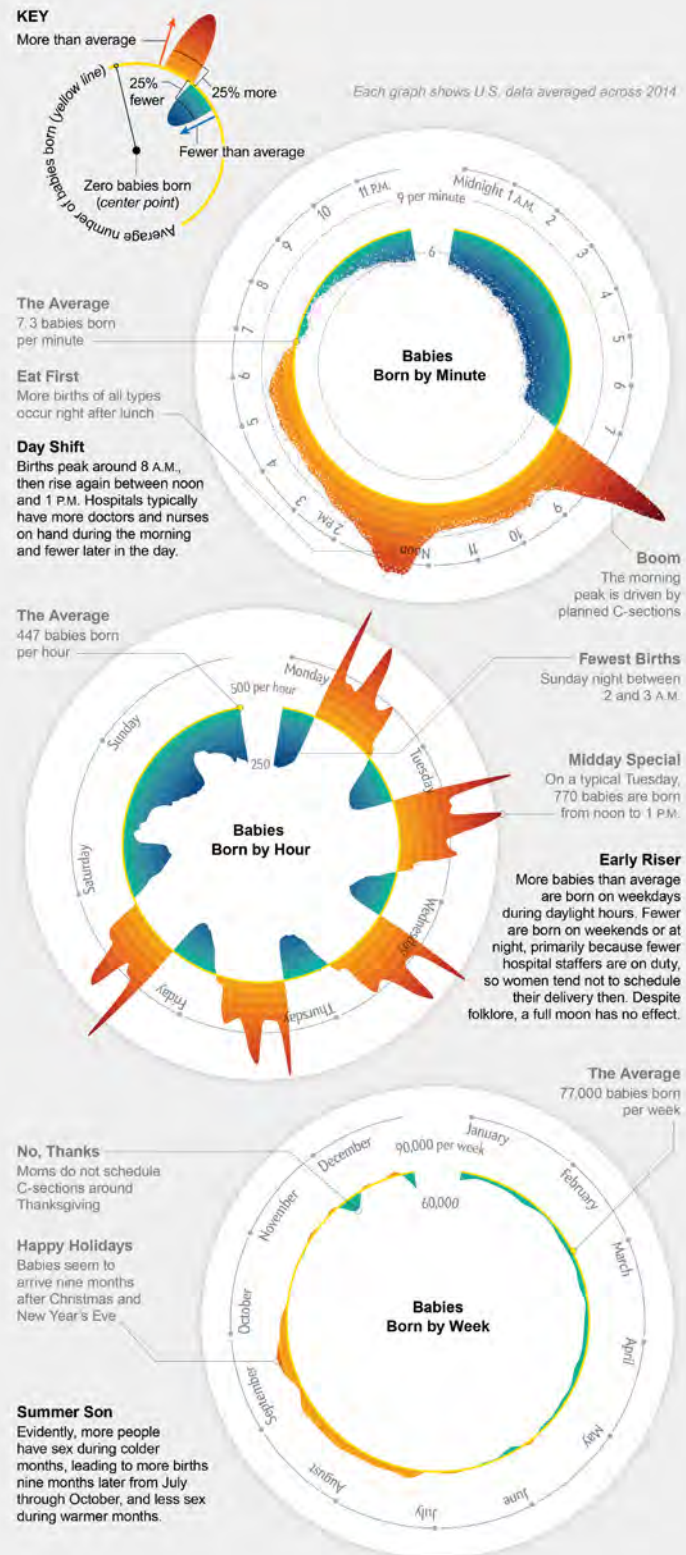


Nature magazine

<https://www.nature.com/news/trillions-of-trees-1.18333>

Scientific American
Information is Beautiful
Awards 2017

informationisbeautifulawards.com/showcase/2100-why-are-so-many-babies-born-around-8-00-a-m

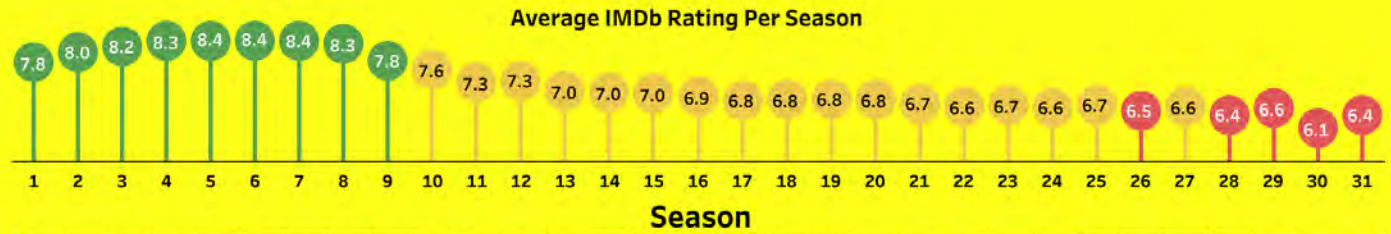


the DECLINE OF THE SIMPSONS

The Simpsons is one of the most iconic TV shows ever created. It first aired in 1989 and continues to do so to this day, spanning over 31 seasons. It soon became one of the most comedically successful shows in television history. However, over time, the comedic material and overall appeal of the show began to decline.

I created a heat chart to show the audience ratings for every episode of every season. I also showed the average rating of each season. The data for this visualization was scraped from IMDb using Python.

Visualized by Roshaan Khan.

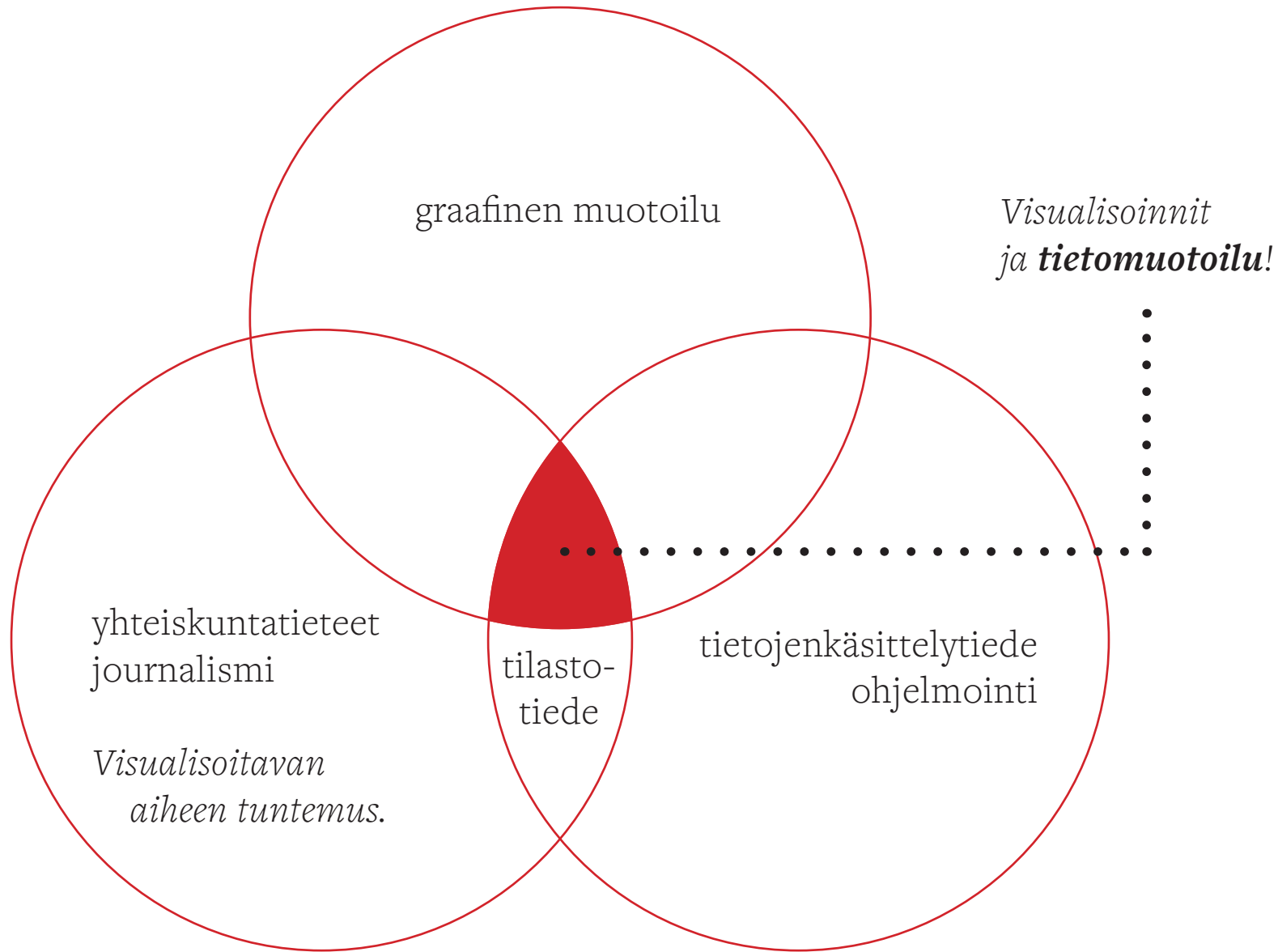


	Season																															
Episode	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
1	8.2	8.2	8.6	8.5	8.5	8.7	9.0	8.5	9.1	7.6	7.3	7.6	7.5	7.6	7.5	7.4	6.1	7.5	6.6	7.2	7.0	6.0	7.0	6.3	6.8	5.6	5.5	6.0	6.7	6.2	5.7	
2	7.7	8.2	7.8	8.2	9.2	8.1	8.3	9.2	6.9	8.1	7.5	7.4	7.0	7.3	7.3	6.8	6.8	6.6	6.5	6.8	6.6	6.7	7.0	7.2	6.9	6.8	6.6	6.5	6.5	6.0	5.9	
3	7.4	8.3	8.2	9.0	8.6	5.9	8.4	8.1	8.1	7.5	7.5	7.0	7.1	7.1	6.8	7.1	6.9	6.6	7.1	6.8	6.1	6.7	6.4	6.8	6.3	6.4	6.9	7.0	6.1	6.1	6.5	
4	7.7	8.1	8.6	7.8	8.9	8.6	8.8	7.7	8.2	8.0	7.7	7.2	7.0	7.0	7.0	6.4	7.3	6.8	6.7	7.1	7.3	7.0	6.4	6.6	6.4	7.4	7.6	6.8	7.1	5.8	6.9	
5	8.0	7.4	8.5	8.5	8.8	8.3	8.5	8.3	8.2	7.2	7.8	7.0	7.6	6.7	7.0	7.0	6.9	6.6	6.9	6.5	6.6	7.1	7.0	6.2	6.8	6.1	6.5	6.3	6.3	7.0	6.2	
6	7.6	8.0	7.7	8.2	8.1	9.2	8.6	8.1	7.7	7.6	7.3	7.7	7.1	7.2	6.4	7.1	6.8	6.6	6.5	7.0	6.5	6.3	7.9	6.2	6.3	7.9	6.6	6.7	6.3	6.5	5.4	
7	7.8	7.8	8.3	7.8	7.7	8.1	9.0	7.8	7.7	8.1	7.0	7.2	6.8	7.0	7.0	6.8	6.6	6.9	6.9	7.0	6.7	6.2	6.6	6.9	6.7	6.7	5.8	6.1	6.3	5.6	6.3	
8	7.7	8.4	7.9	8.2	8.7	8.6	8.6	8.8	8.1	7.3	7.1	7.9	6.9	7.2	6.6	6.1	7.0	7.1	7.1	6.5	7.1	6.7	6.0	7.1	6.3	6.3	6.8	6.6	6.8	6.4	7.3	
9	7.5	8.1	7.9	8.9	8.5	9.0	8.0	8.6	7.6	8.2	7.3	8.1	7.2	6.7	7.3	6.5	6.8	6.1	8.2	5.9	6.6	6.9	8.2	6.6	7.6	6.7	8.4	6.3	7.2	6.7	6.5	
10	7.4	7.8	8.8	8.7	8.6	8.1	7.5	9.1	7.6	7.8	7.3	7.1	7.3	6.7	7.1	6.8	6.6	6.5	6.8	6.9	6.5	6.1	5.9	6.9	6.9	6.6	6.5	5.8	6.5	6.7	6.6	
11	7.8	8.8	8.3	8.7	8.3	7.8	8.4	7.8	5.0	7.8	6.8	7.4	7.0	6.6	7.0	6.6	6.9	6.9	6.1	6.9	7.0	6.6	6.9	6.4	7.1	7.1	6.5	6.8	6.2	5.5	6.0	
12	8.4	8.3	8.3	9.1	8.2	9.0	8.3	8.6	7.7	7.0	7.3	6.8	6.4	7.2	7.2	6.9	6.4	6.7	6.5	6.7	6.7	6.9	5.6	6.1	6.2	5.6	6.5	6.1	6.6	6.3	6.6	
13	7.8	8.0	8.5	8.0	8.3	8.8	8.6	7.7	8.3	7.8	6.4	7.1	6.4	6.9	7.1	6.7	7.6	7.2	7.6	7.1	6.4	6.7	6.2	6.9	6.6	6.1	6.3	6.3	6.6	6.9	5.9	
14		7.5	8.0	8.2	8.2	8.7	7.7	8.0	8.2	7.3	7.7	7.2	7.1	6.7	6.7	6.7	6.9	6.8	7.3	6.3	6.9	7.1	6.9	6.1	6.6	6.5	6.5	6.4	6.0	6.1	7.3	
15		8.3	8.1	8.4	8.9	8.6	7.7	8.8	7.5	7.3	7.3	7.5	7.0	7.1	6.6	7.2	6.4	6.1	6.6	6.6	6.9	6.6	6.8	6.8	6.8	6.4	6.3	6.5	6.3	6.4	6.5	
16		7.5	8.3	8.4	8.5	8.5	8.2	8.2	7.6	7.5	7.3	6.6	7.7	7.1	7.0	7.4	6.3	7.0	5.9	7.1	5.6	5.9	7.2	7.0	6.7	6.6	7.1	6.5	6.3	5.7	6.5	
17		7.6	8.7	9.1	7.9	7.9	8.9	8.0	8.0	7.7	7.1	6.8	5.4	7.1	6.5	7.1	7.1	7.2	6.7	7.1	6.5	6.1	6.9	6.3	6.3	6.5	6.3	6.8	6.7	6.3	7.0	
18		8.0	8.2	7.1	8.3	8.6	8.1	8.9	7.9	7.4	7.3	8.6	7.7	6.9	7.0	6.6	6.7	5.8	7.0	6.9	6.8	6.8	7.0	6.3	6.8	6.5	6.2	6.1	7.0	4.5	6.3	
19		8.6	7.8	8.1	8.4	8.3	8.2	8.3	8.2	7.6	6.5	7.2	6.7	6.6	7.3	7.2	7.1	6.9	7.0	7.1	6.8	6.4	7.5	6.7	5.7	6.7	7.1	6.8	6.6	6.2	6.0	
20		7.9	8.1	8.1	8.3	8.2	8.6	7.6	7.8	7.3	7.1	7.2	7.1	6.7	7.2	7.0	6.6	6.7	6.3	6.3	7.1	7.0	6.7	6.1	7.9	5.9	6.7	6.1	6.7	5.3	5.8	
21		8.3	8.3	7.8	7.5	8.1	8.9	7.9	7.9	7.1	7.1	7.0	7.0	6.6	6.2	7.2	7.0	8.1		7.2	6.9	7.1	6.5	7.0	6.5	6.5	6.3	6.4	7.6	5.3	6.7	
22		7.9	7.7	8.2	8.1	8.2	8.4	8.1	8.4	7.3	7.9		7.8	7.7	7.2		6.3	7.1			7.2	6.8	3.9	6.8	6.5	6.7	6.7			5.9	7.4	
23			7.8			8.2	8.1	9.3	8.2	8.0												6.4									5.9	
24			8.3			8.7	8.0	7.2	7.8																							
25					9.2	8.4	7.9	8.0																								

https://www.linkedin.com/posts/roshaankhan_datascience-analytics-python-activity-6740165181831208960-pT9Q

Mistä aloitan?

Visualisoinnin perusteita



Kuinka tieto järjestellään?

Lähteenä *Tieto näkyväksi* (Koponen et. al. 2016)

Järjestä *itse data* – 4 asteikkoa

Muuttujat ja asteikot:

Kaikella, jota visualisoidaan (esim kaupungit) on eri ominaisuuksia (esim. asukasluku) = *muuttujat*

Muuttujien *arvot* voidaan asetella **luokiteltaville** tai **määrällisille** *asteikoille*:



Laatu- / nominaaliasteikko = ei suuruus- tai järjestysskaalaa (esim. asukkaiden ammatit kaupungissa). Voidaan luokitella eri tyyppien mukaan.

Järjestä *itse data* – 4 asteikkoa

Muuttujat ja asteikot:

Kaikella, jota visualisoidaan (esim kaupungit) on eri ominaisuuksia (esim. asukasluku) = *muuttujat*

Muuttujien *arvot* voidaan asetella **luokiteltaville** tai **määrällisille** *asteikoille*:



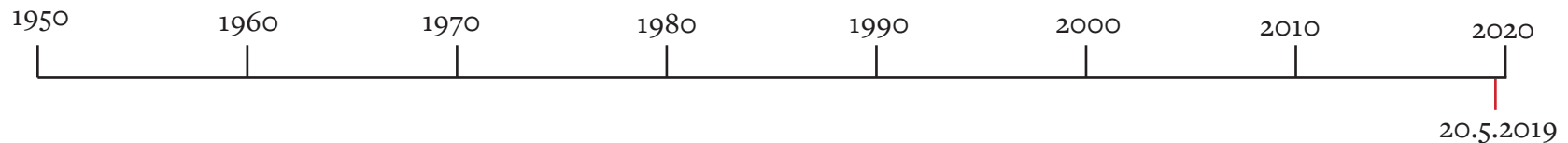
Järjestys- / ordinaaliasteikko = ei suuruusskaalaa, mutta luokat voi järjestää tiettyyn järjestykseen (Likert-asteikko)
pituus, bkt, työttömyysprosentti, lämpötila Kelvineinä)

Järjestä *itse data* – 4 asteikkoa

Muuttujat ja asteikot:

Kaikella, jota visualisoidaan (esim kaupungit) on eri ominaisuuksia (esim. asukasluku) = *muuttujat*

Muuttujien *arvot* voidaan asetella **luokiteltaville** tai **määrällisille** *asteikoille*:



Intervalliasteikko = ei luokkia, vaan yksittäisiä kohteita jotka voi asetella asteikolle tietyin etäisyyksin (vuosiluvut, lämpötilat). Eroja ja etäisyyksien välillä voi laskea, EI MUITA LASKUTOIMITUKSIA (vuosilukujen prosentuaalinen ero <-- dadaa)

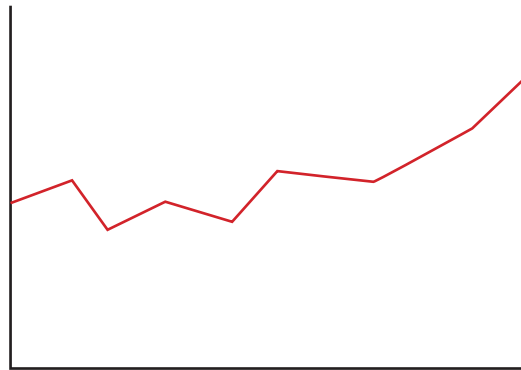
Järjestä *itse data* – 4 asteikkoa

Muuttujat ja asteikot:

Kaikella, jota visualisoidaan (esim kaupungit) on eri ominaisuuksia (esim. asukasluku) = *muuttujat*

Muuttujien *arvot* voidaan asettaa **luokiteltaville** tai **määrällisille** *asteikoille*:

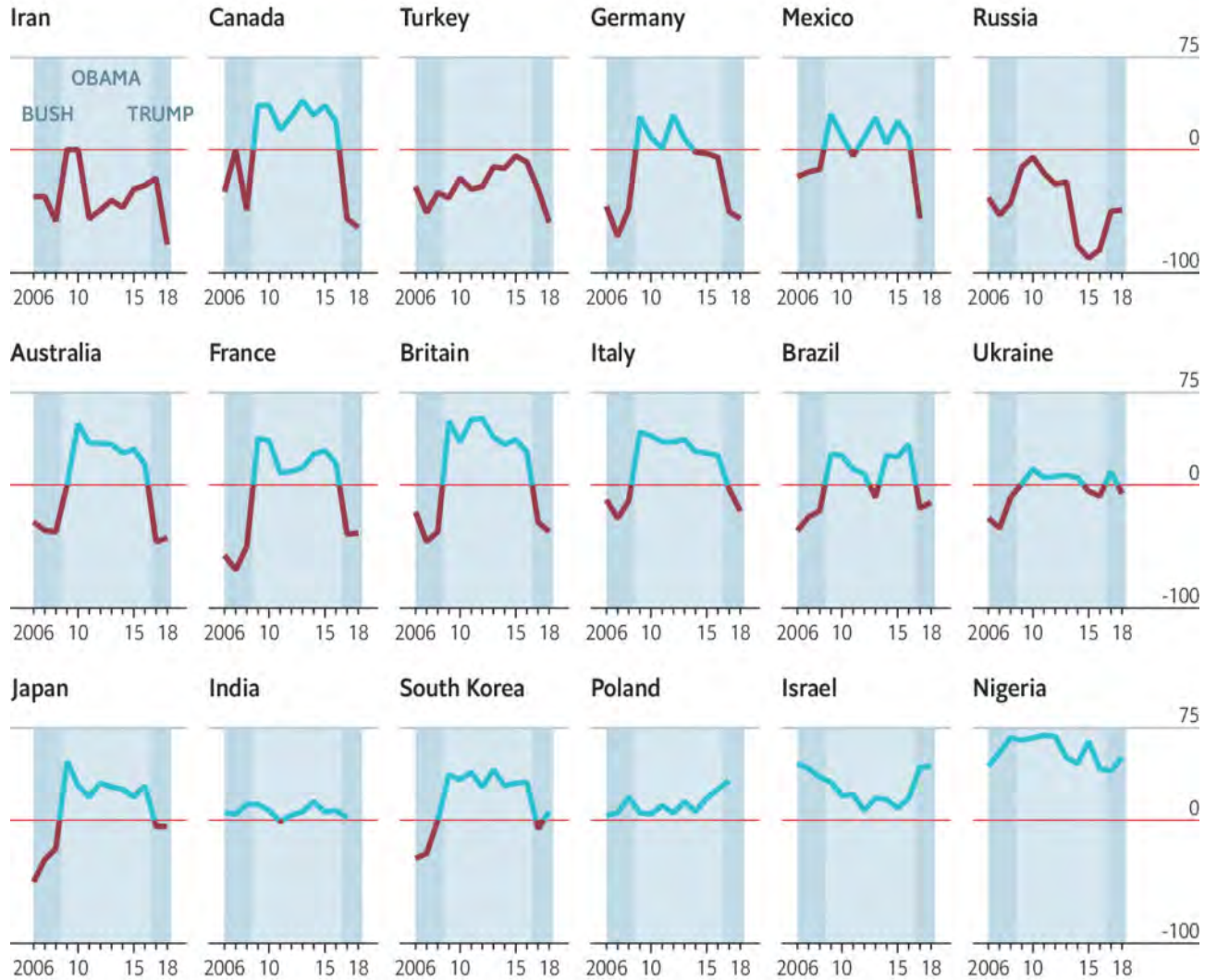
29



Suhdeasteikko = yksittäisten kohteiden asteikko jossa absoluuttinen nollepiste. Kaikki laskutoimitukset mahdollisia (fyysiset suureet kuten paino ja pituus, bkt, työttömyysprosentti, lämpötila Kelvineinä)

Hail to the chief

Net approval of American leadership, %



Source: Gallup

The Economist

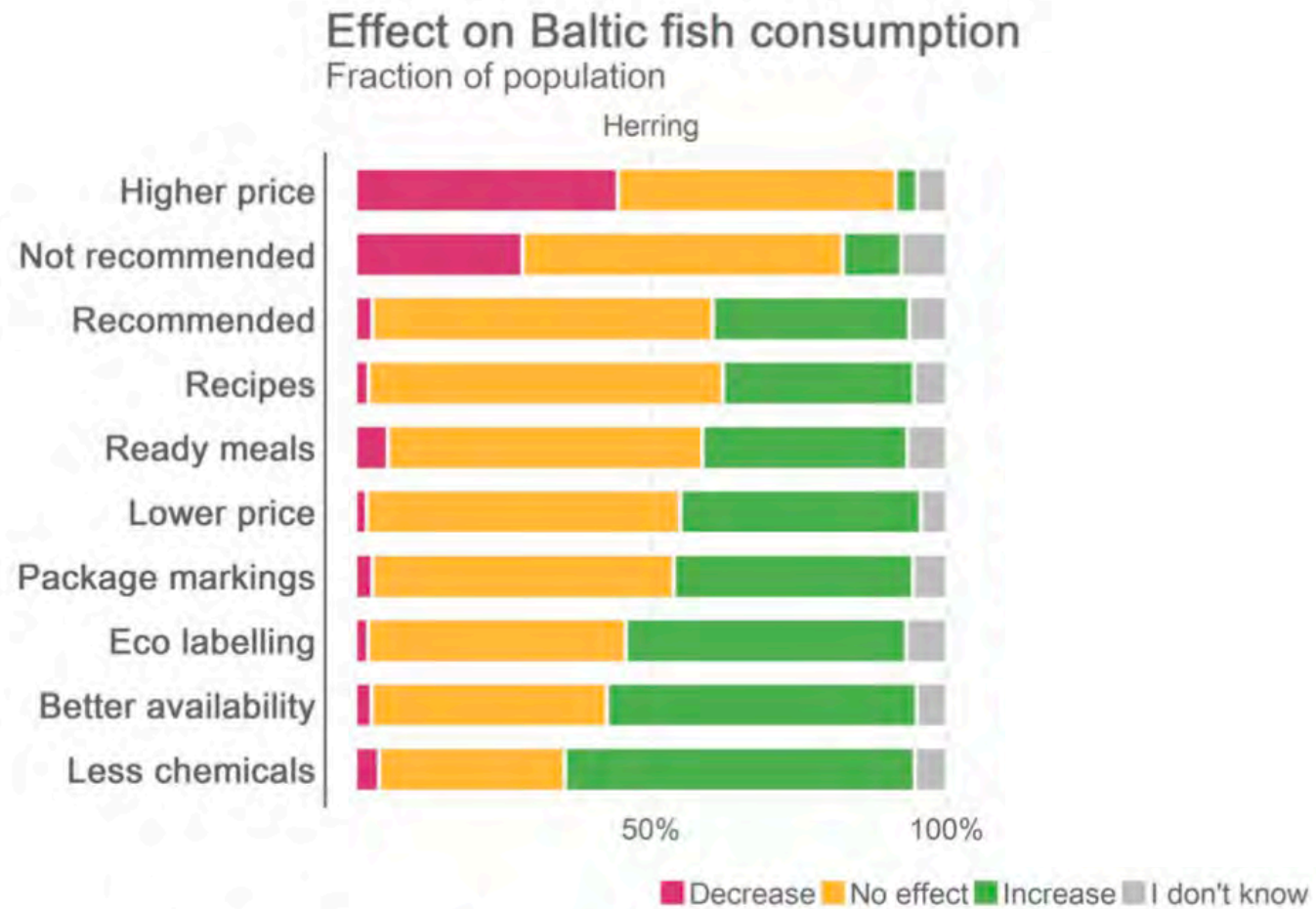
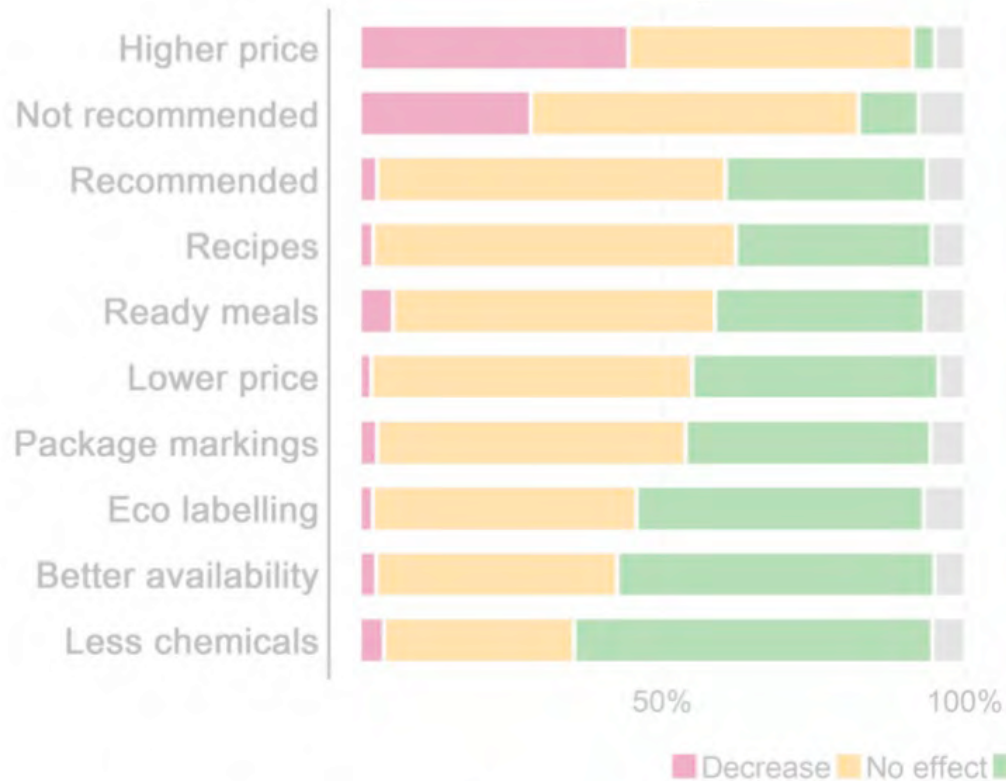


Figure 6. The influence of different determinants on potential changes in Baltic herring and salmon consumption.

Effect on Baltic fish consumption

Fraction of population

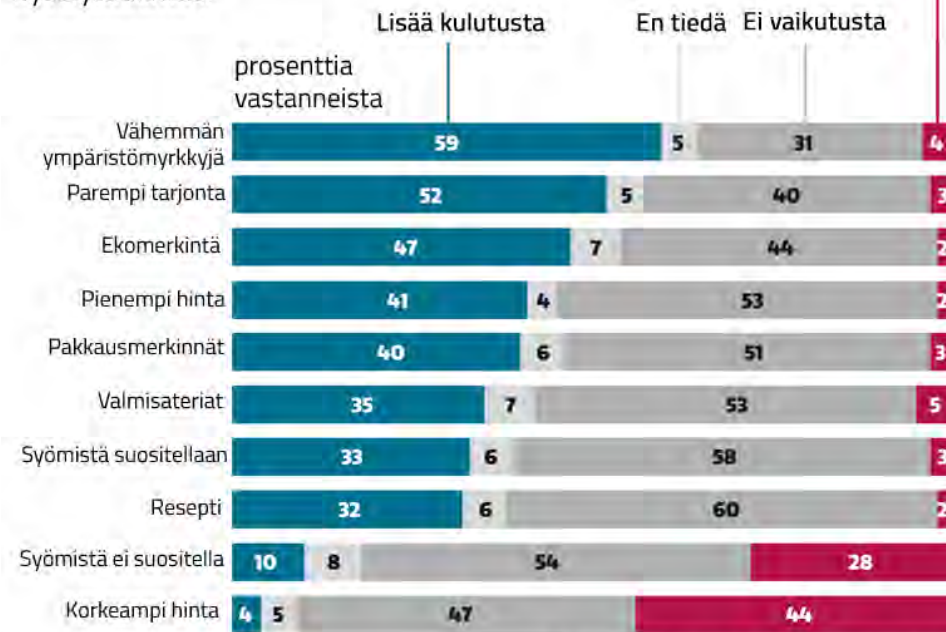
Herring



MIKÄ SAISI IHMISET KÄYTTÄMÄÄN ENEMMÄN SILAKKAA?

Myrkkysten väheneminen lisää mielenkiintoa silakansyönttiin eniten, kohonneet hinnat vähentävät sitä eniten, kertoo kyselytutkimus.

Vähentää kulutusta



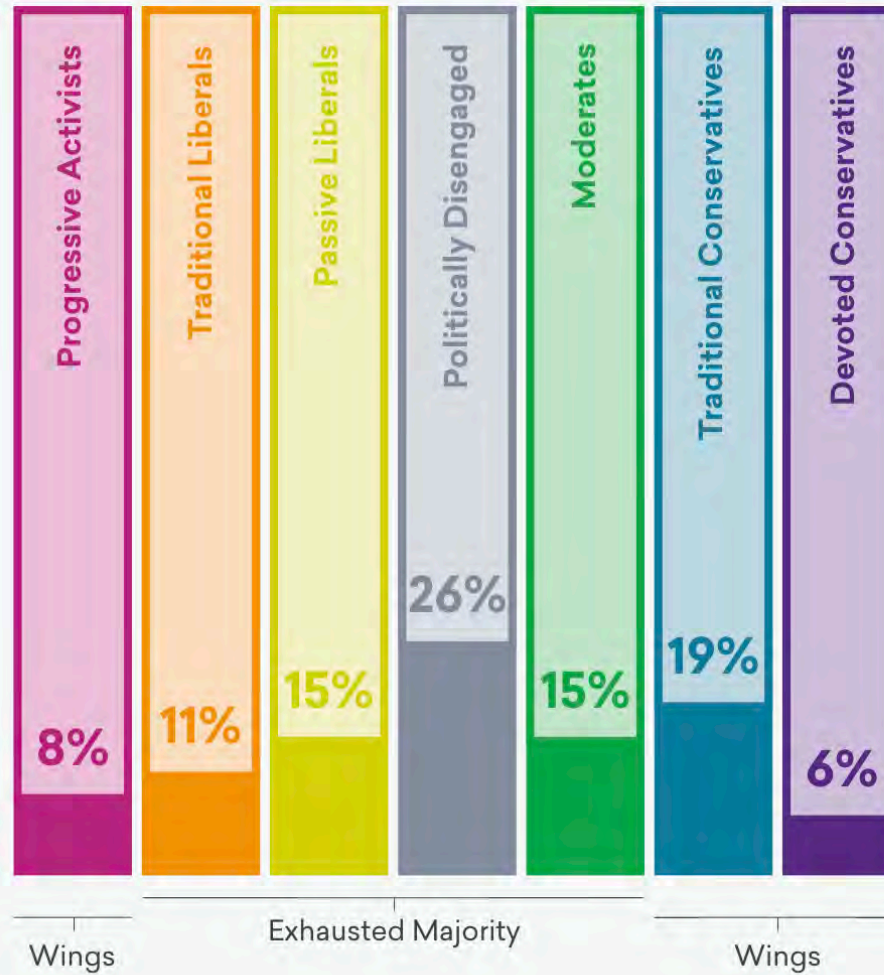
GRAFIikka KRUT COLLECTIVE / KRUT.FI

Figure 6. The influence of different determinants on potential changes in Baltic herring and salmon consumption.

Less is more

Keep it simple?

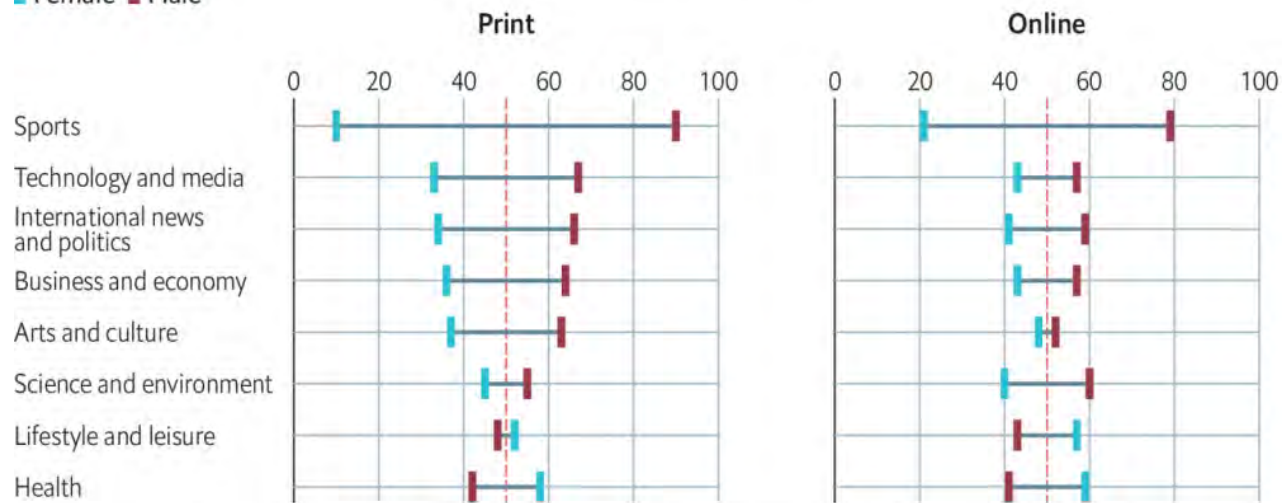
The Hidden Tribes of America



A gender setting

United States, share of male and female bylines by subject*, selected

■ Female ■ Male

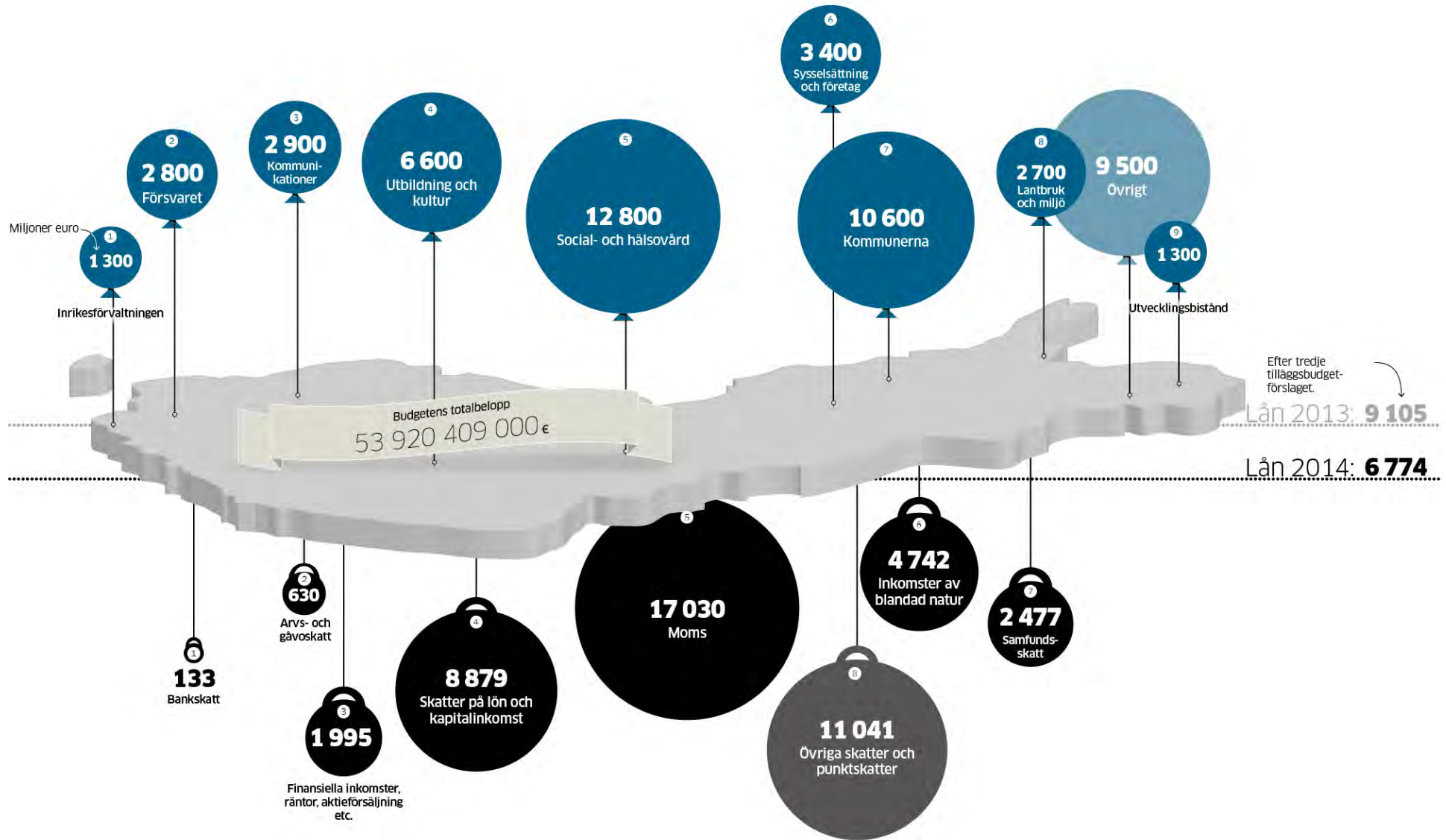


Source: Women's Media Centre

*Analysis of 52,584 pieces of content from September 1st to November 30th, 2017

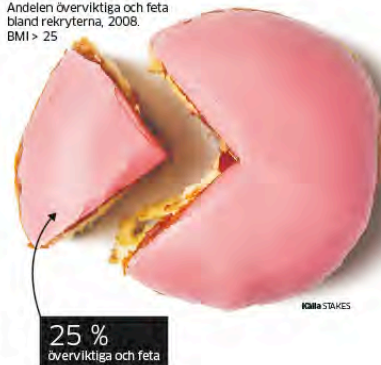
The Economist

Suomen valtionbudjetti 2014

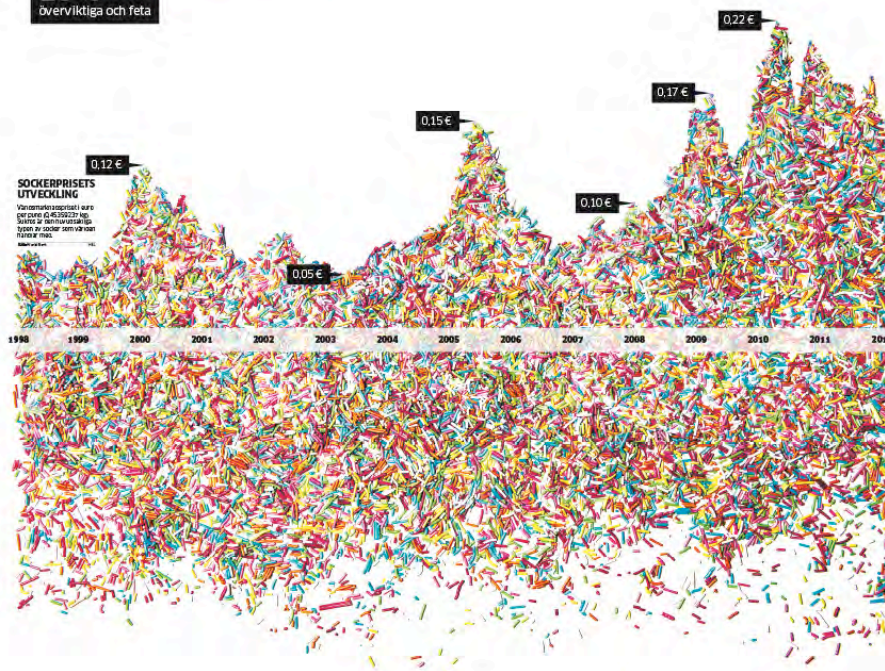


Ylipainoisten alokkaiden osuus, 2008

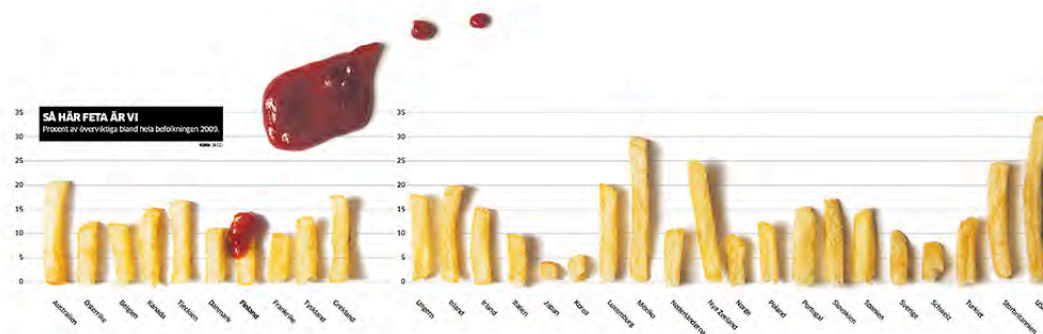
FÖRSVARET "VÄXER"
Andelen överviktiga och feta
bland rekryterna, 2008.
BMI > 25



Sokerin hintakehitys

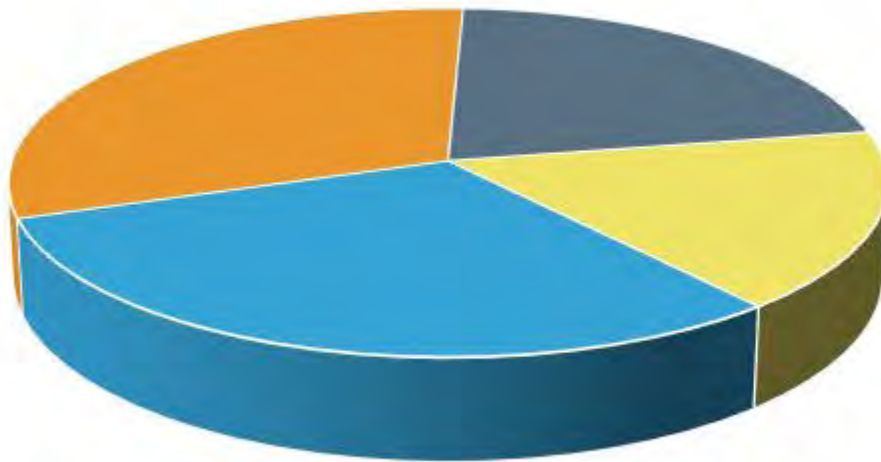


Ylipainoisten osuus väestöstä 2009, maittain



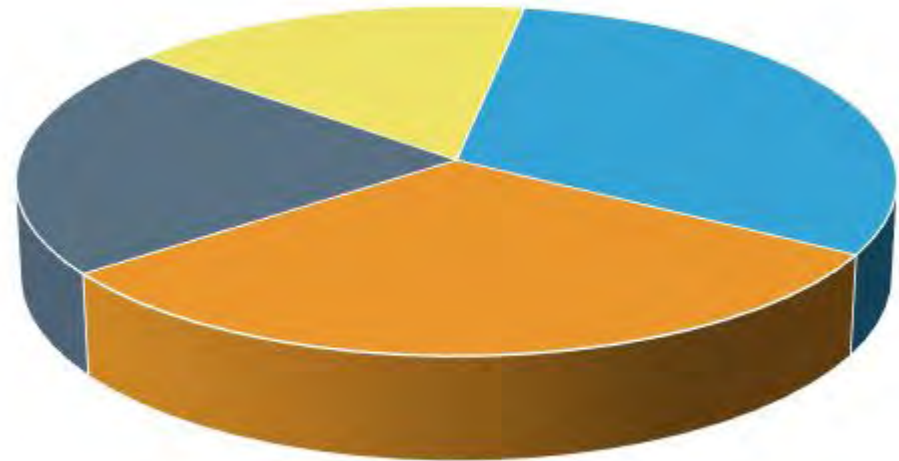
~~3D koristeču~~

Sales Q1



■ Paul ■ Bryan ■ Susan ■ Jane

Sales Q2



■ Paul ■ Bryan ■ Susan ■ Jane

Person	Sales
Paul	50
Bryan	50
Susan	35
Jane	27

<https://www.getnerdyhr.com/3d-pie-charts-are-evil/>

Interaktiivisuus

Vähemmän tutkimustietoa kuin printtigraafeista, mutta interaktiivinen “palastelu” auttaa hyvin käytettynä yleisöä havainnollistamaan useampia ulottuvuuksia (esim. Hans Roslingin animaatiot)

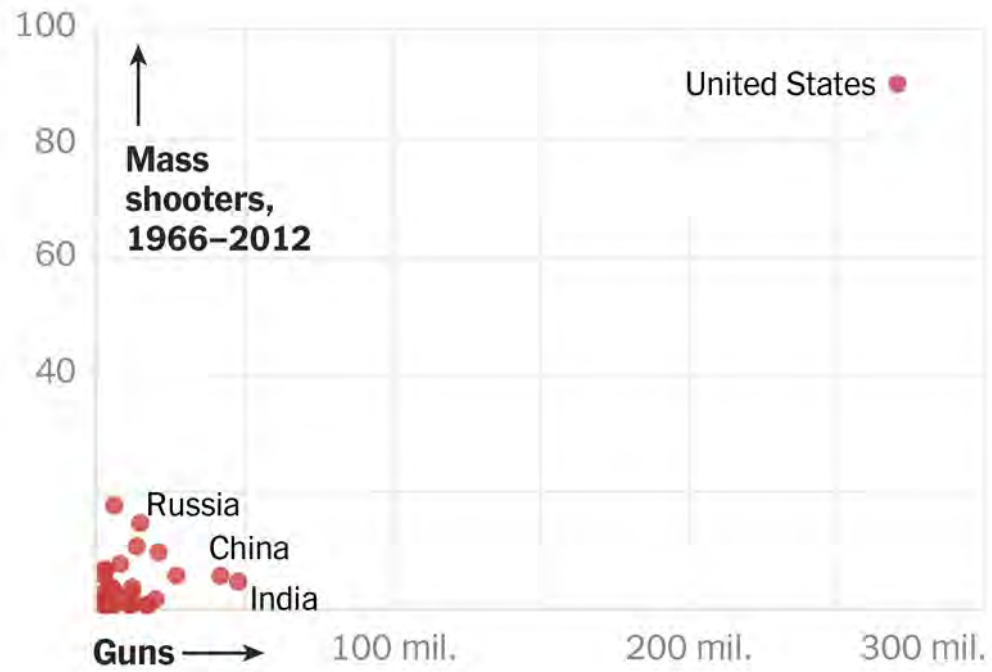
SVT Datajournalistik:

Väljarströmmarna

<https://www.svt.se/special/valu2018-valjarstrommar/>

Vad har hänt med CO2-utsläpp på 10 år

<https://www.svt.se/special/co2-utslapp-tio-ar/>



Hans Rosling

45

<https://www.gapminder.org/answers/how-does-income-relate-to-life-expectancy/>

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

GDP per person v self-reported happiness

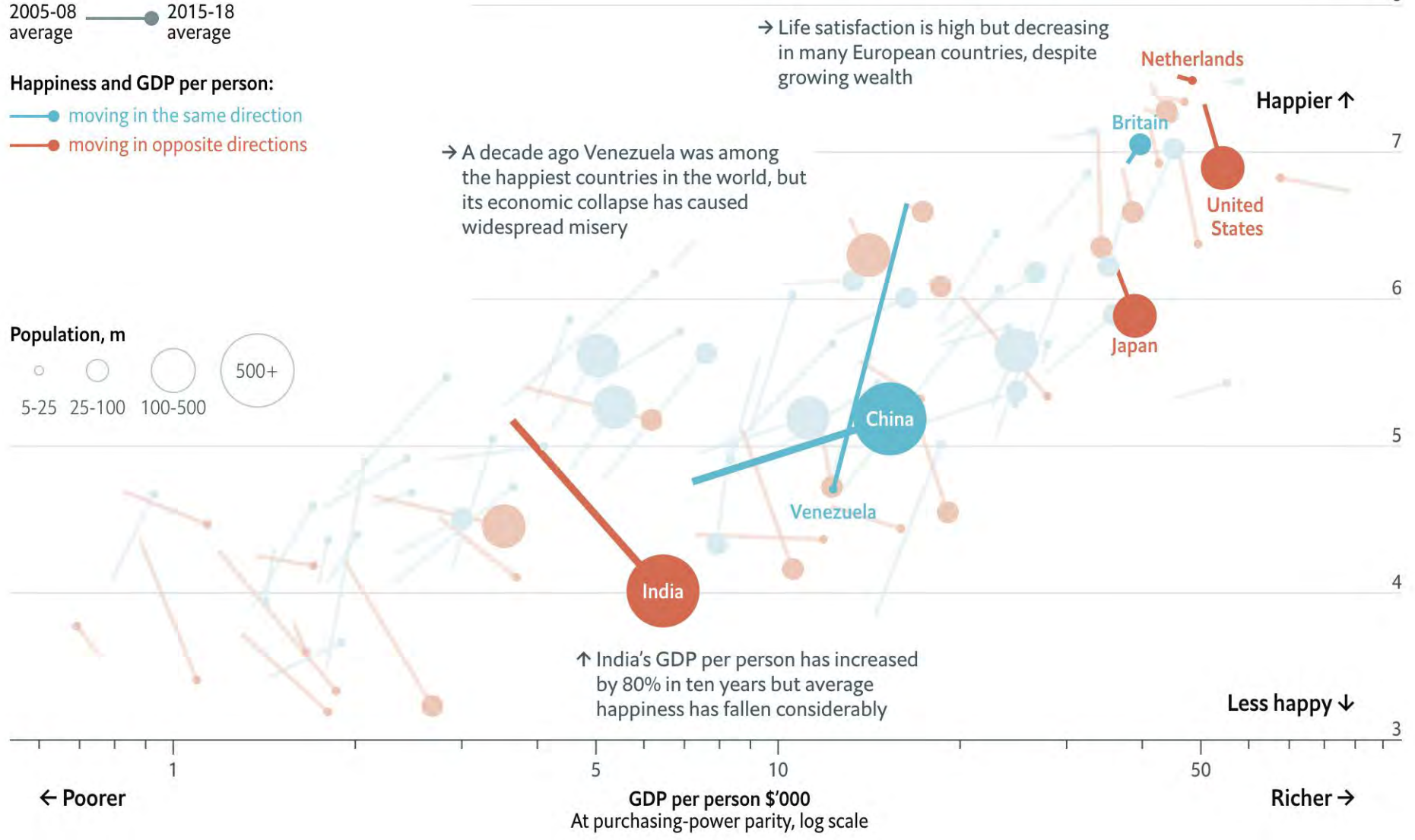
85 countries with adult population over 5m

Happiness
0-10 scale

2005-08 average —● 2015-18 average

Happiness and GDP per person:
● moving in the same direction
● moving in opposite directions

Population, m
○ 5-25 ○ 25-100 ○ 100-500 ○ 500+



Punainen lanka:

Trillions of Trees

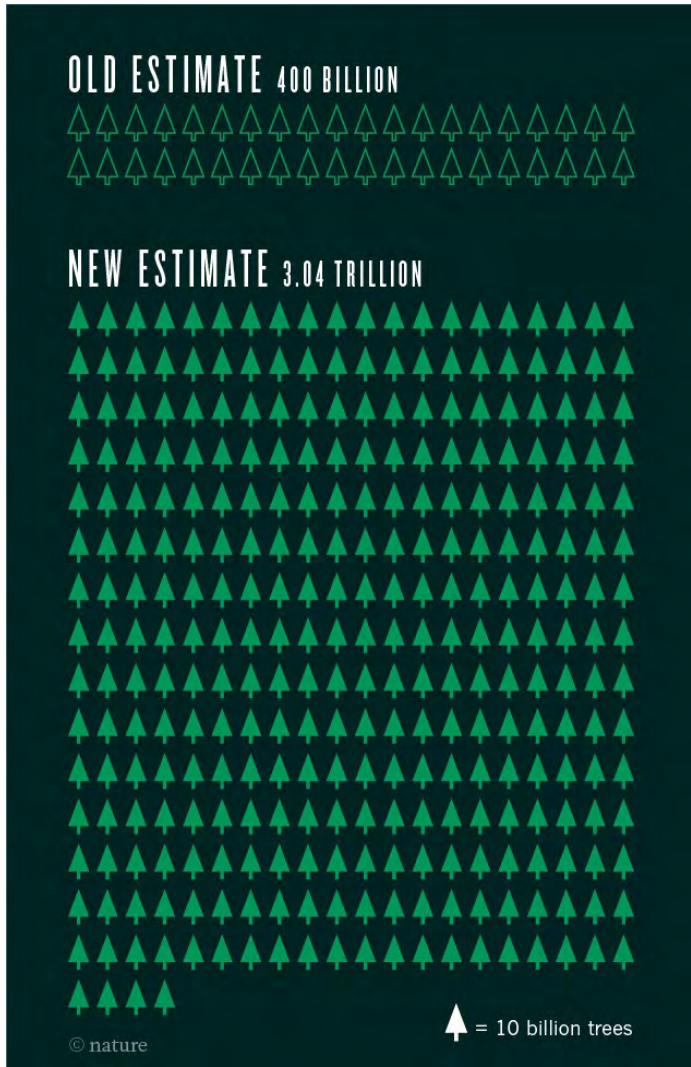
<https://www.youtube.com/watch?v=jqdOkXQngw8>

Työkalu /

käyttäjä tutkii itse:

World Water Atlas:

<http://www.worldwateratlas.org/>



Visualisointien luonti – first steps

Valitse muoto

Financial Times

Visual Vocabulary

Kaaviotyyppejä:

[https://github.com/ft-interactive/chart-doctor/tree/master/
visual-vocabulary](https://github.com/ft-interactive/chart-doctor/tree/master/visual-vocabulary)

Chart Continuum

<https://policyviz.com/2014/09/09/graphic-continuum/>

<h3>Deviation</h3> <p>Emphasise variation (V) from a fixed reference point. Typically, the reference point is zero but it can also be a target or a stop-start value. Use this to point to a change or a trend.</p> <p>Example FT use Track surpluses/shortfalls change</p> <p>Diverging bar A simple standard bar chart that can handle both positive and negative values.</p> <p>Diverging stacked bar Perfect for presenting survey results when positive and negative responses are being compared.</p> <p>Split Splitting a single value into two contrasting components (eg. pros/cons).</p> <p>Barstack/Barfill plot The widest area of most charts allows a balance to be shown either above or below a baseline or between two bars.</p>	<h3>Correlation</h3> <p>Show the relationship between two or more variables. Be mindful that, unless you tell your audience, many readers will assume the relationship you show them to be causal (it often isn't).</p> <p>Example FT use Inflation and unemployment, income and life expectancy</p> <p>Scatterplot The standard way to show the relationship between two continuous variables, eg. IQ (y-axis) vs. test score (x-axis).</p> <p>Column + line timeline A good way of showing the relationship between an amount (columns) and a rate (line).</p> <p>Connected scatterplot Usually used to allow the relationship between two variables to be tracked over time.</p> <p>Bubble Like a scatterplot, but each individual data point is sized according to a third variable.</p> <p>X-Y heatmap A good way of showing the relationship between two categories of data, less effective at showing big differences in volume.</p>	<h3>Ranking</h3> <p>Use when an item's position in an ordered list is more important than its absolute or relative value. Don't forget to highlight the items of interest in the data.</p> <p>Example FT use Market penetration, league tables, construction sector results</p> <p>Ordered bar Standard bar charts display the value of each item, but most want to highlight the shape of the data.</p> <p>Ordered column See above.</p> <p>Ordered proportional symbol Use when there are too many variations between values and/or ranking differences are important.</p> <p>Dot strip plot Data points in order to a strip and a space-efficient method of laying out multiple individual data points.</p> <p>Slope Perfect for showing how wide have changed over temporary different components.</p> <p>Lollipop Lollipop shows more attention to the data value than ordered bar charts and top-down flow charts.</p> <p>Barco Effective for showing changing rankings across multiple dates. For large datasets, consider grouping lines using colour.</p>	<h3>Distribution</h3> <p>Show values in a dataset and how often they occur. The values for each distribution can be a combination of highlighting the best or worst in the data.</p> <p>Example FT use Income distribution, population density, distribution, revealing inequality</p> <p>Histogram The standard way to show a distribution of data. The bars represent the frequency of data points.</p> <p>Dot plot A simple way of showing the range of data across multiple categories.</p> <p>Dot strip plot Good for showing individual values in a distribution, can be a problem when the number of data points is high.</p> <p>Barcode plot Like dot strip plots, but more effective with complex distributions and points that come in a sequence.</p> <p>Stripplot Stacking multiple distributions by showing the median, quartiles, and range of the data.</p> <p>Violin plot Similar to a box plot but more effective with complex distributions and points that come in a sequence.</p> <p>Population pyramid A standard way of showing the age and sex distribution of a population, effectively showing back-to-back histograms.</p> <p>Cumulative curve A good way of showing the distribution of a data set, where cumulative frequency is used to show the total.</p> <p>Frequency polygons For displaying multiple distributions of data, line graphs are often used to show the relationship between two variables.</p> <p>Seesaw Use to emphasise individual points in a distribution. Points can be sized by an additional variable, but with medium-sized datasets.</p>	<h3>Change over Time</h3> <p>Use especially for changing trends. These can be shown using line graphs or area charts. Use these to show trends or seasonal variations.</p> <p>Example FT use Share price movements, economic trends, sectoral changes in a market</p> <p>Line The standard way to show a changing trend. Use this to show trends or seasonal variations.</p> <p>Column Columns work well for showing change over time, but usually used with only one series of data at a time.</p> <p>Column + line timeline A good way of showing the relationship between an amount (columns) and a rate (line).</p> <p>Slope Good for showing changing trends, as the data can be presented as a line graph.</p> <p>Area chart Use with care - these are good at showing changes over time, but showing change in components can be very difficult.</p> <p>Candlestick Usually based on day-to-day activity, candlestick charts are used to show price movements and volatility.</p> <p>Fan chart (projections) Use to show the changing size of a projection - usually the growth rate, but can be used for other projections.</p> <p>Connected scatterplot A good way of showing changing data for a series of variables over a period of time.</p> <p>Calendar heatmap A great way of showing important patterns (days, weeks, months) - at the expense of showing proportional quantities.</p> <p>Priority timeline Good when data sets go through an evolution over time.</p> <p>Circle timeline Good for showing the values of a series of data across multiple categories and different time points.</p> <p>Vertical timeline Resists that on the Y-axis is used for showing the relationship between two variables, such as the relationship between a variable and a time period.</p> <p>Salesogram Another alternative to the circle timeline for showing trends, where there are a large number of variables in the data.</p> <p>Streamgraph A type of area chart, but when using change in proportions, you can see the most important individual values.</p>	<h3>Magnitude</h3> <p>Use for comparisons. These can be relative (not using size) or absolute (using size). Use these to show trends or seasonal variations.</p> <p>Example FT use Company performance, market capitalisation, volume in general</p> <p>Column The standard way to show a comparison between two or more items. Use this to show trends or seasonal variations.</p> <p>Bar See above. Good when the data is not too large and you have long category names.</p> <p>Paired column A standard way of showing two items for comparison. Can be used to show trends over time.</p> <p>Paired bar See above.</p> <p>Horizontal A good way of showing the size of a comparison between two or more items. Use this to show trends or seasonal variations.</p> <p>Proportional symbol Use when there are big differences between values and/or ranking differences are important.</p> <p>Barpyramid (histogram) Excellent addition to some bar charts - use only with whole numbers (to show the relationship between two variables).</p> <p>Lollipop Lollipop shows more attention to the data value than ordered bar charts and top-down flow charts.</p> <p>Radar A space-efficient way of showing the values of a series of data across multiple categories and different time points.</p> <p>Parallel coordinates An alternative to parallel charts - right up to an important difference is important. Usually presented as a line graph, but can be used to show trends over time.</p> <p>Barbell Good for showing a relationship between two or more items, but with a performance target.</p> <p>Ground symbol An alternative to bar charts, where the data is shown as a series of horizontal lines, with individual elements in a row.</p>	<h3>Part-to-whole</h3> <p>Show how a single entity can be broken down into its component elements. The reader's interest is likely to be in the size of the components, rather than the magnitude of the part instead.</p> <p>Example FT use Fiscal budgets, company structure, national election results</p> <p>Stacked column/bar A simple way of showing part-to-whole relationships, but can be difficult to read with more than a few components.</p> <p>Horizontal A good way of showing part-to-whole relationships, but can be difficult to read with more than a few components.</p> <p>Pie A common way of showing part-to-whole relationships, but can be difficult to read with more than a few components.</p> <p>Donut Similar to a pie chart, but the center is cut out. Good for showing part-to-whole relationships, but can be difficult to read with more than a few components.</p> <p>Treemap Use for hierarchical part-to-whole relationships, but can be difficult to read with more than a few components.</p> <p>Venn A way of showing the relationship between two or more items, but can be difficult to read with more than a few components.</p> <p>Arc A friendly, often used for showing part-to-whole relationships, but can be difficult to read with more than a few components.</p> <p>Sankey Good for showing the flow of data between two or more items, but can be difficult to read with more than a few components.</p> <p>Venn Generally preferred for showing part-to-whole relationships, but can be difficult to read with more than a few components.</p> <p>Waterfall Can be used for showing part-to-whole relationships, but can be difficult to read with more than a few components.</p>	<h3>Spatial</h3> <p>Apply from location maps, or use when the spatial location of data is more important than the magnitude of the data.</p> <p>Example FT use Population density, natural resources, agricultural levels, weather in a region</p> <p>Basic choropleth (categorical) The standard way of showing spatial data, but can be difficult to read with more than a few categories.</p> <p>Proportional symbol (counting method) Use for non-numeric data, but can be difficult to read with more than a few categories.</p> <p>Flow map A good way of showing the relationship between two or more items, but can be difficult to read with more than a few components.</p> <p>Contour map A good way of showing the relationship between two or more items, but can be difficult to read with more than a few components.</p> <p>Equalized cartogram Concerning area, use a map to show the relationship between two or more items, but can be difficult to read with more than a few components.</p> <p>Scaled cartogram (value) Concerning area, use a map to show the relationship between two or more items, but can be difficult to read with more than a few components.</p> <p>Dot density Used to show the density of individual events/locations, but can be difficult to read with more than a few components.</p> <p>Heat map Good for showing the relationship between two or more items, but can be difficult to read with more than a few components.</p>	<h3>Flow</h3> <p>Show the major elements of a process or system. These might be stages or components of a process, or the flow of data between two or more items.</p> <p>Example FT use Manufacturing process, supply chain, network</p> <p>Sankey Shows things that flow from one position to another, but can be difficult to read with more than a few components.</p> <p>Waterfall Used to show the relationship between two or more items, but can be difficult to read with more than a few components.</p> <p>Chart A common way of showing the relationship between two or more items, but can be difficult to read with more than a few components.</p> <p>Network Used for showing the relationship between two or more items, but can be difficult to read with more than a few components.</p>
--	---	--	--	--	---	---	---	---

Visual vocabulary

Designing with data

There are so many ways to visualise data - how do we know which one to pick? Use the categories across the top to decide which data relationship is most important in your story, then look at the different types of chart within the category to form some initial ideas about what might work best. This list is not meant to be exhaustive, nor a wizard, but is a useful starting point for making informative and meaningful data visualisations.

FT Visual Vocabulary Chart created by Alex de Waard, based on work by the Information Systems Research Group at the University of Cambridge. Used under the Creative Commons Attribution 4.0 International License.



Datan käsittely

Google sheets, Excel, Openoffice yms

Openrefine

geocod.io

Airtable

Pandas

Python

CartoDB

R

Visualisointi

Tableau

Data Studio

Adobe Illustrator, Dataly

HTML/CSS , Javascript, d3.js

RAW Graphs

Flourish

Datawrapper

Infogram

Highcharts

R

Python

QGIS

Carto

Leaflet

Vain muutama esimerkki!

Flourish.studio

53

Aloita täältä:

<https://flourish.studio/2018/09/28/choosing-the-right-visualisation/>

Lukemista & inspiraatiota

Muutama artikkeli

7 tips for better data visualizations
Subtleties of Color (NASA)
Data literacy for all
7 principles of Icon Design
data-to-viz.com - valitse oikea muoto!

Kirjoja

Cairo, Alberto. How charts Lie
Bergstrom, Carl T. Calling Bullshit
Rosling, Hans & Ola, Rönnlund, Anna.
Factfulness (suom. Faktojen maailma)
Visualizing data: a handbook for data driven
design

Esimerkkejä ja tekijöitä

Alberto Cairo (twitter)
Data is beautiful (Reddit)
Economist - graphic detail
Marketcagemag.com

#dataviz




Information is Beautiful Awards
Gapminder tool
Accurat.it
Periscopic
Data Visualization Society - newsletter
Data Stories podcast

Lähteet

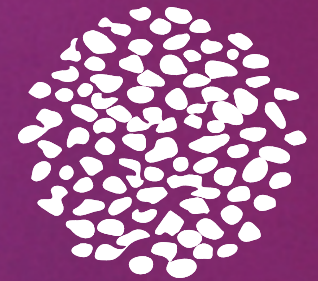
Cairo, Alberto (2016). *The Truthful Art. Data, Charts and Maps for Communication*. New Riders, USA.
Koponen, Juuso; Hildén, Jonatan;
Vapaasalo, Tapio (2016). *Tieto näkyväksi*.
Informaatiomuotoilun perusteet. Aalto ARTS
Books, Finland.
Market Cafe Magazine, Issue 4

“Best pie chart ever”



-  Sky
-  Sunny side of pyramid
-  Shady side of pyramid

KITOS!



KRUT
COLLECTIVE