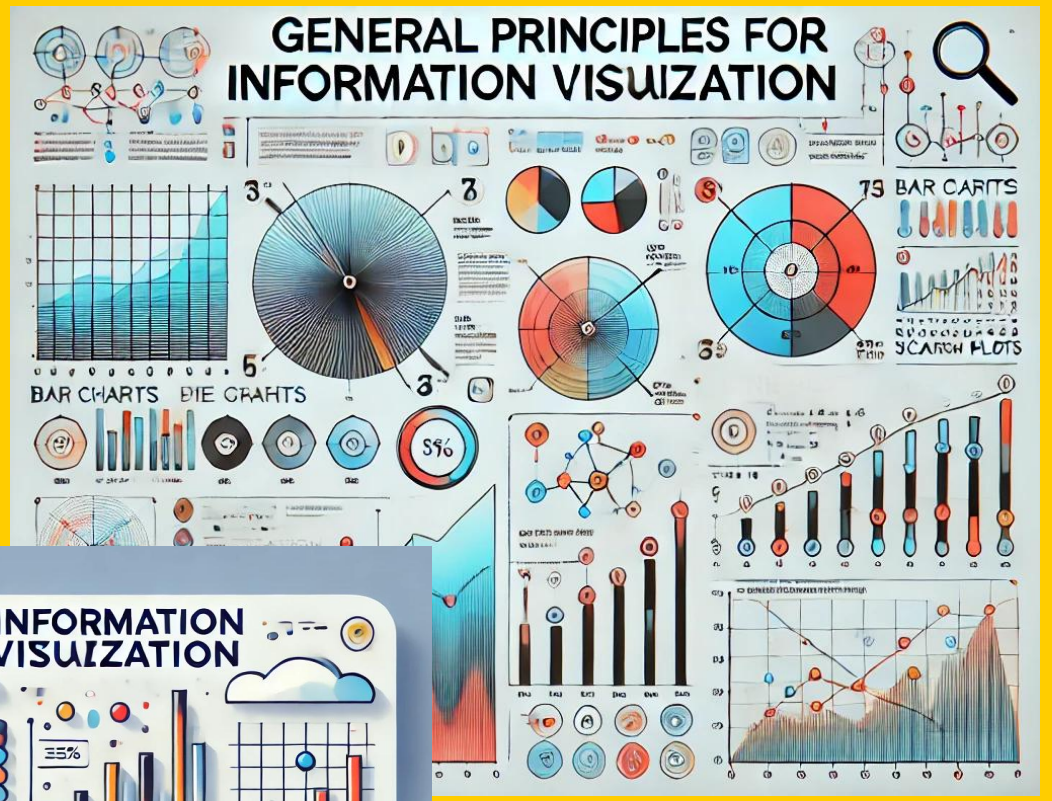
Hoorcollege 1
Algemene principes en visualisatietypes



Informatie-uitwisseling:
informatievisualisatie

Jeroen Ooge

Universiteit Utrecht

Over mij

Human-centred explainable AI met visualisaties

Mijn korte carrière

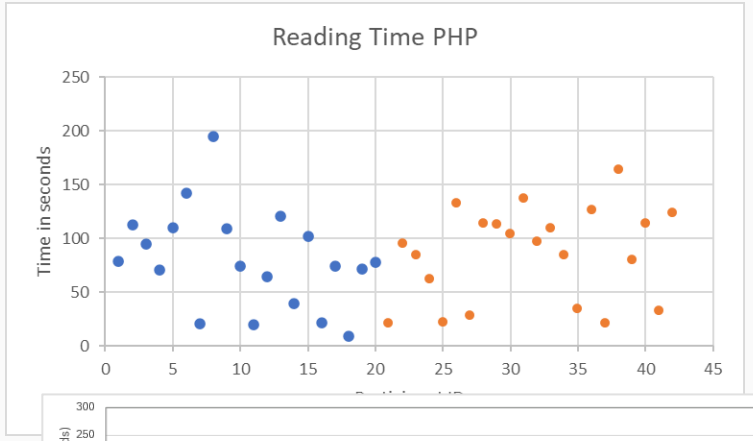
MSc fundamentele wiskunde
MSc toegepaste informatica
2012-2019

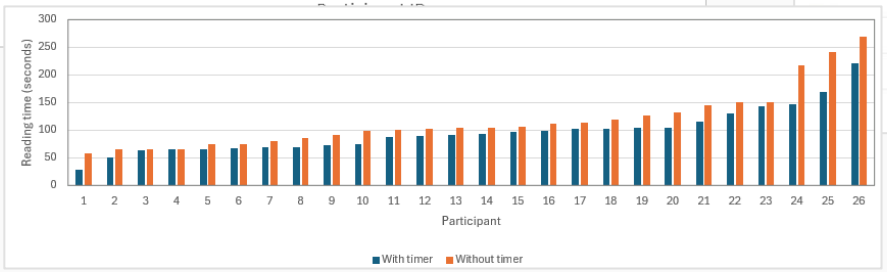
KU LEUVEN
PhD in computer science
2019-2023

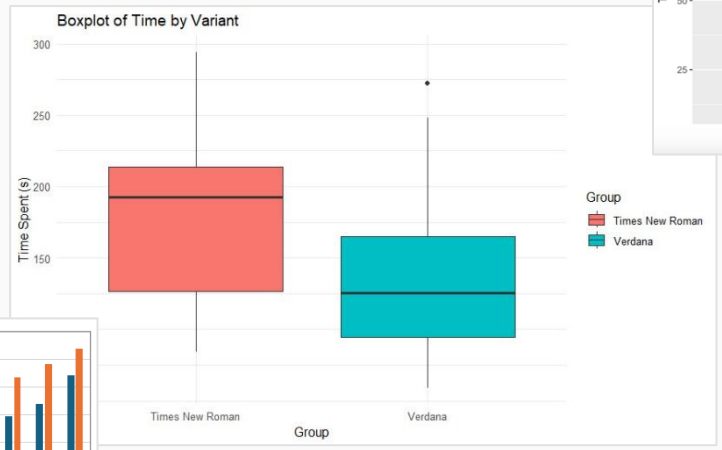
Assistent professor
in computer science
2023-nu
Universiteit Utrecht

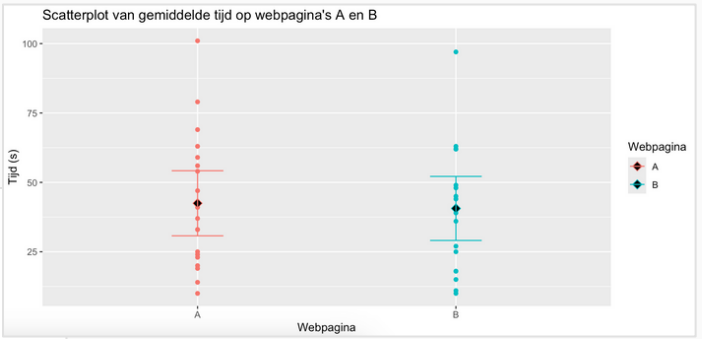
Jeroen Ooge
jeroenooge.be

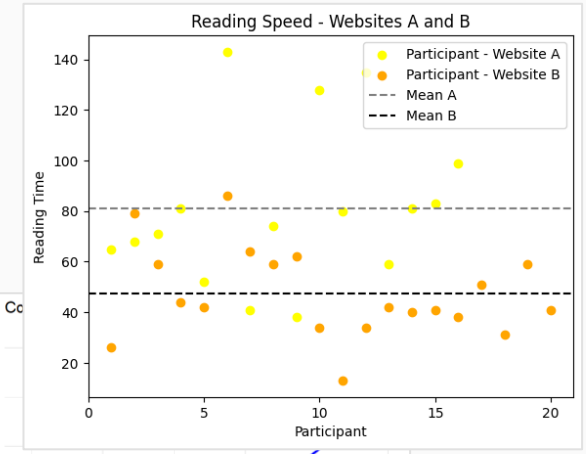
2de BA-studenten vorig jaar 🤔

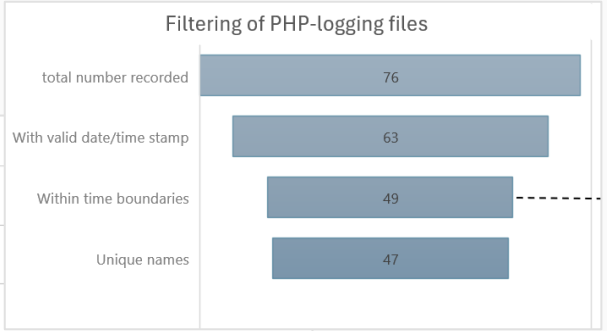


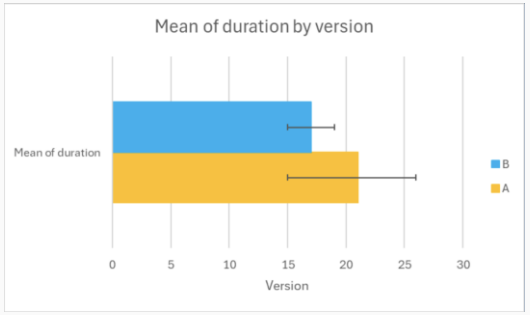


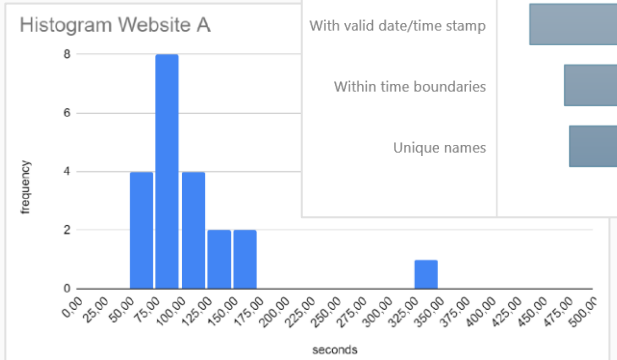


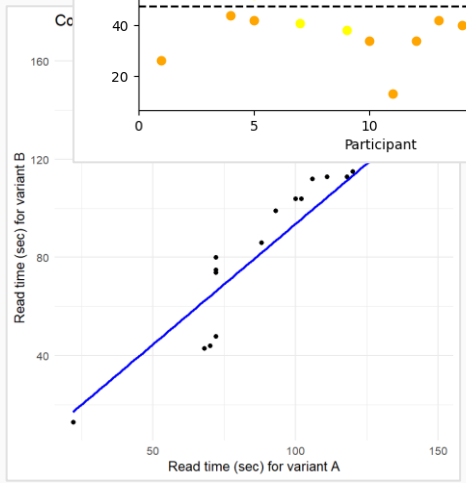












Universiteit Utrecht

Studenten vinden data voorstellen en grafieken interpreteren vaak moeilijk

Daarom deze inleiding tot informatie**vis**ualisatie

Jeroen Ooge
jeroenooge.be

Algemene principes en visualisatietypes

1. Voorbeelden en geschiedenis
2. Algemene vis-principes
3. Klassieke vis-types
4. Tools voor vis

Wat is visualisatie?

"Computer-based visualization systems provide **visual representations** of datasets designed to help people carry out tasks more effectively."

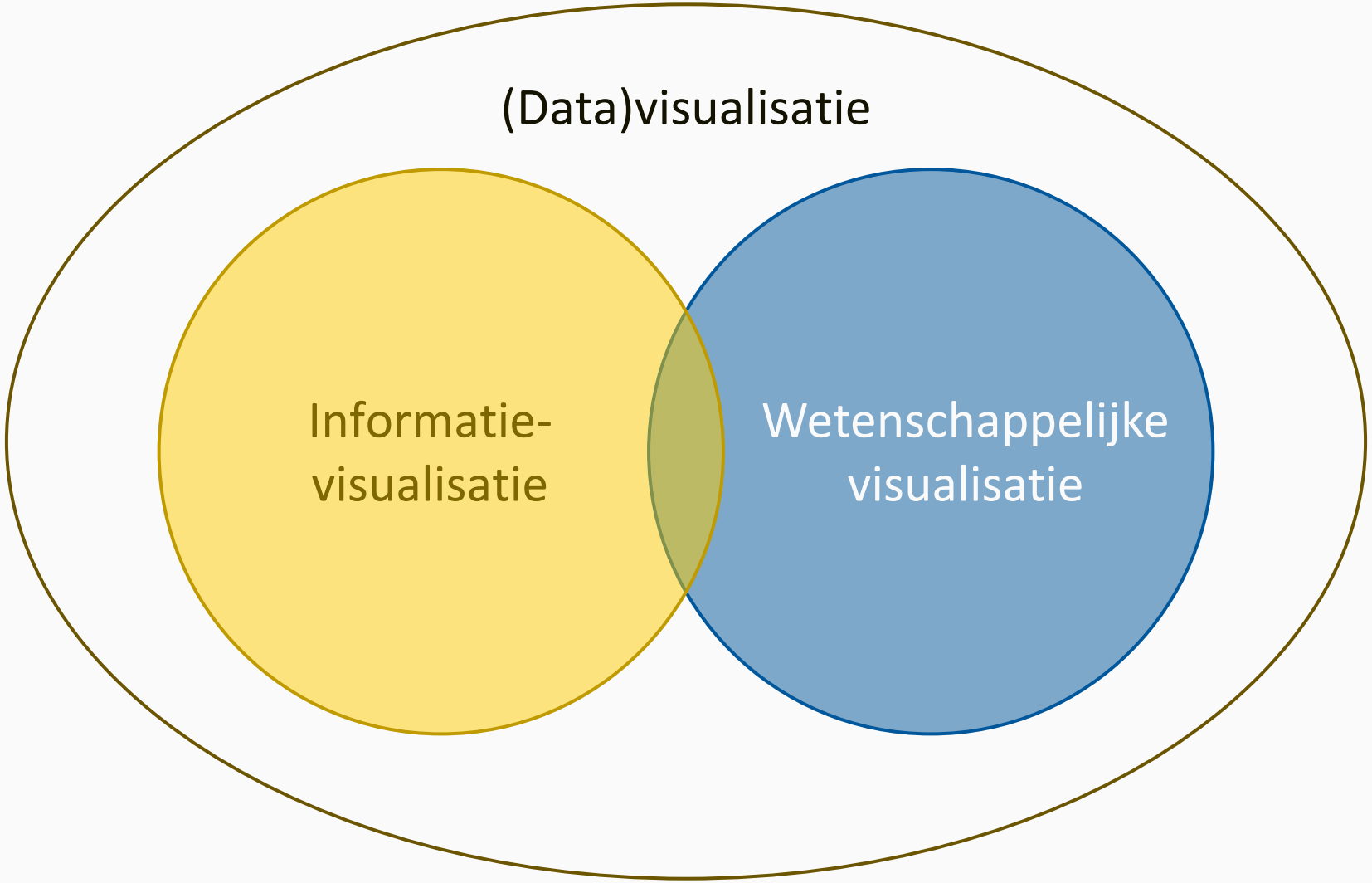
Tamara Munzner. 2014. Visualization Analysis and Design. A K Peters/CRC Press, New York. https://doi.org/10.1201/b17511

"... visually display measured quantities [with] **points, lines, a coordinate system, numbers, symbols, words, shading**, and **color**."

Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press.

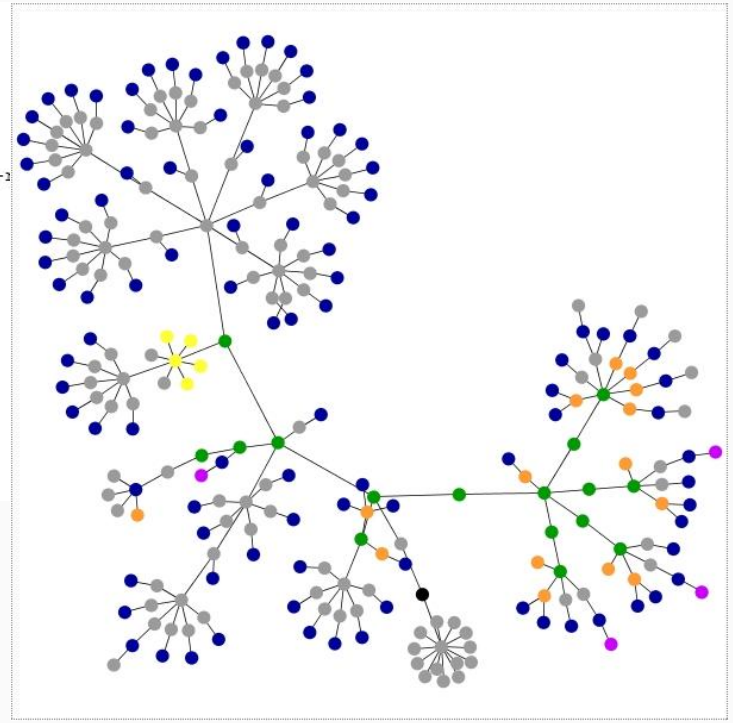
"... computer-supported interactive visual representations of **abstract data** to amplify cognition. Its purpose is ... **insight** (or rapid information assimilation or monitoring large amounts of data)."

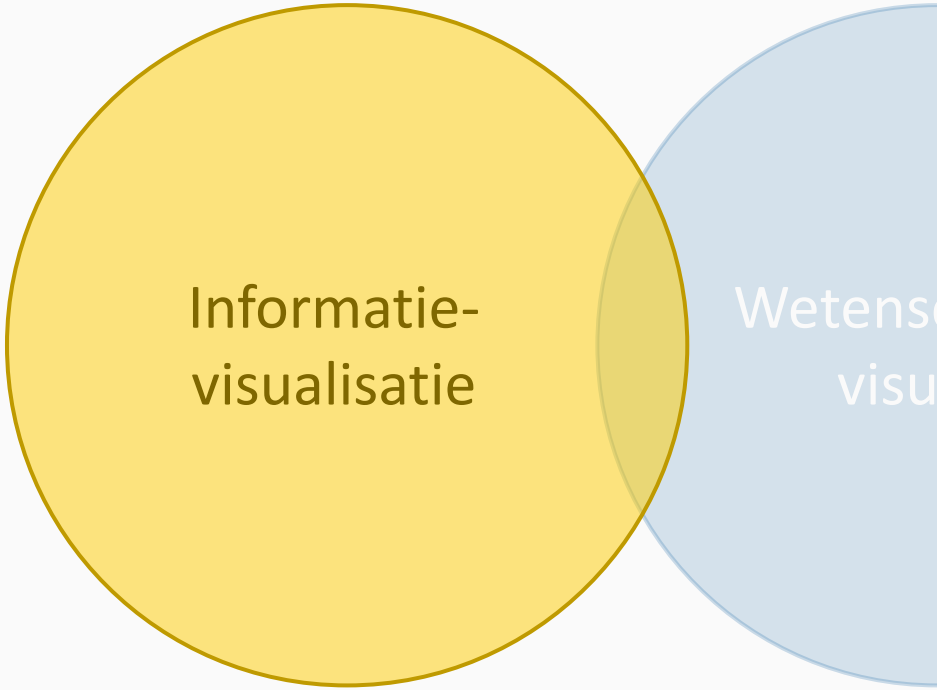
Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). Readings in information visualization: using vision to think. Morgan Kaufmann.


(Data)visualisatie
Informatie-
visualisatie
Wetenschappelijke
visualisatie

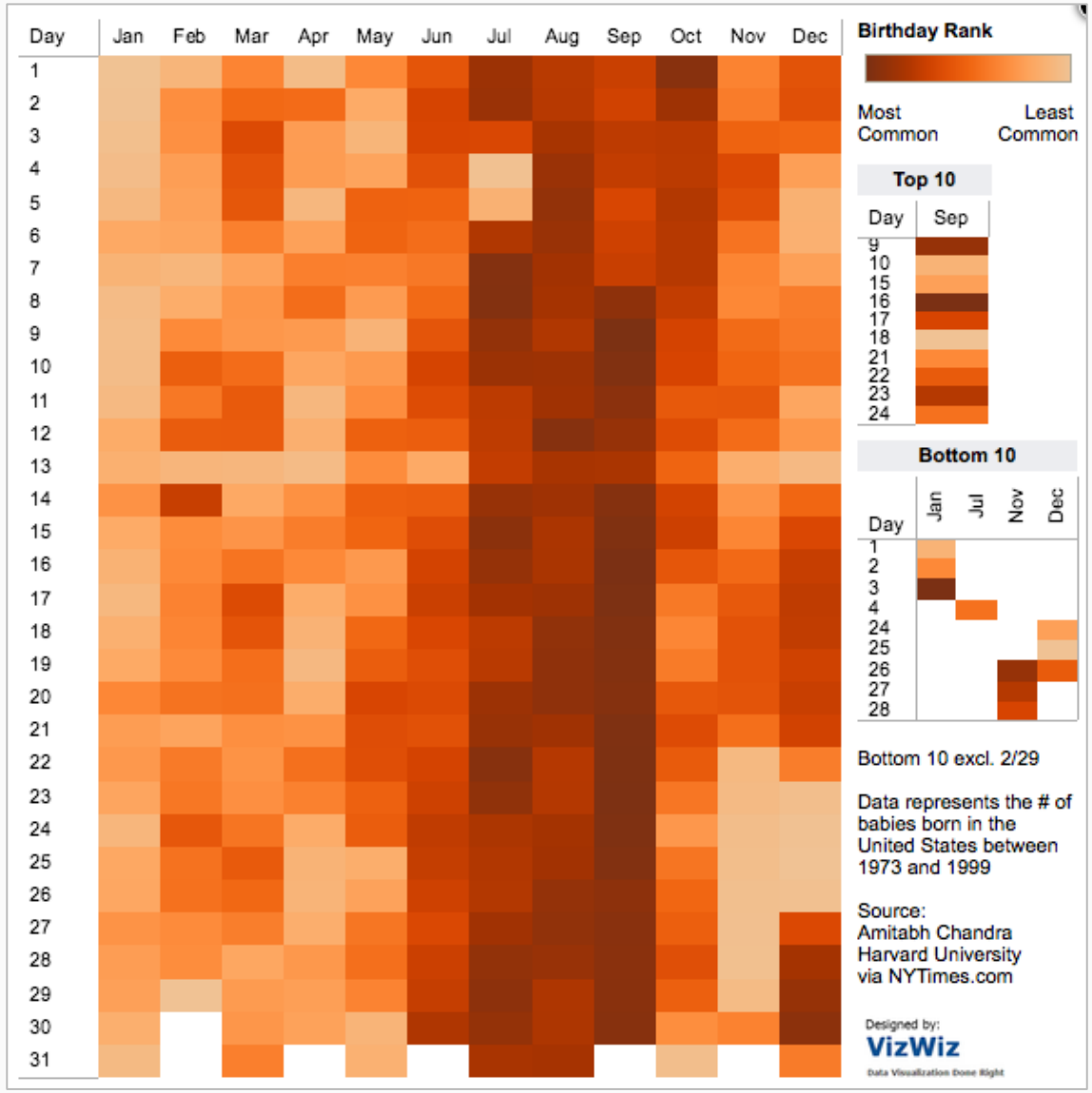
Data **zonder** welgedefinieerde representatie in 2D of 3D (= “abstracte data”)

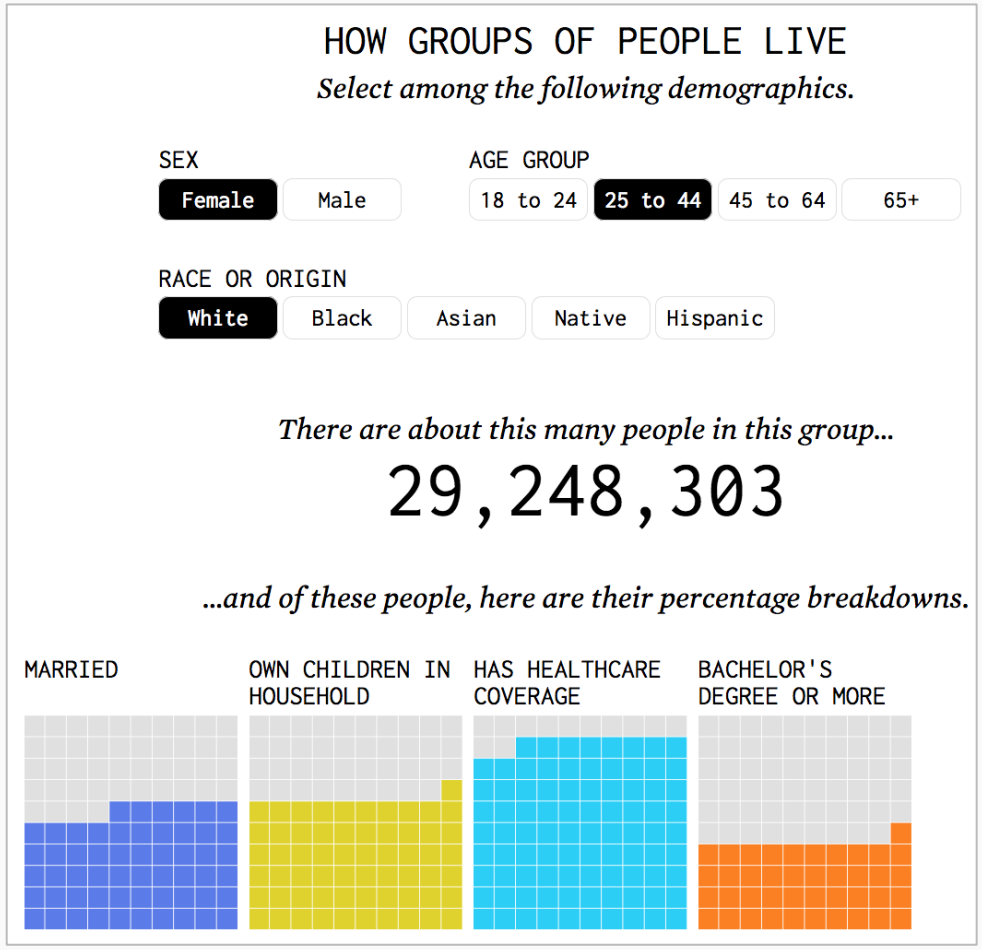
```
http://www.bu.edu/students/life/
http://www.bu.edu/students/life/housing/
http://www.bu.edu/students/life/dining-vending/
http://www.bu.edu/students/life/phone/
http://www.bu.edu/students/life/safety/
http://www.bu.edu/students/life/transportation/
http://www.bu.edu/students/life/activities/
http://www.bu.edu/students/life/bu-global-orientation-1
http://www.bu.edu/students/academics/
http://www.bu.edu/students/academics/link/
http://www.bu.edu/students/academics/admissions/
http://www.bu.edu/students/academics/registration/
http://www.bu.edu/students/academics/advising/
http://www.bu.edu/students/academics/grades/
http://www.bu.edu/students/academics/services/
http://www.bu.edu/students/academics/support/
http://www.bu.edu/students/health/
http://www.bu.edu/students/health/services/
http://www.bu.edu/students/health/counseling/
http://www.bu.edu/students/health/facilities/
http://www.bu.edu/students/health/clubsports/
http://www.bu.edu/students/health/varsitysports/
```





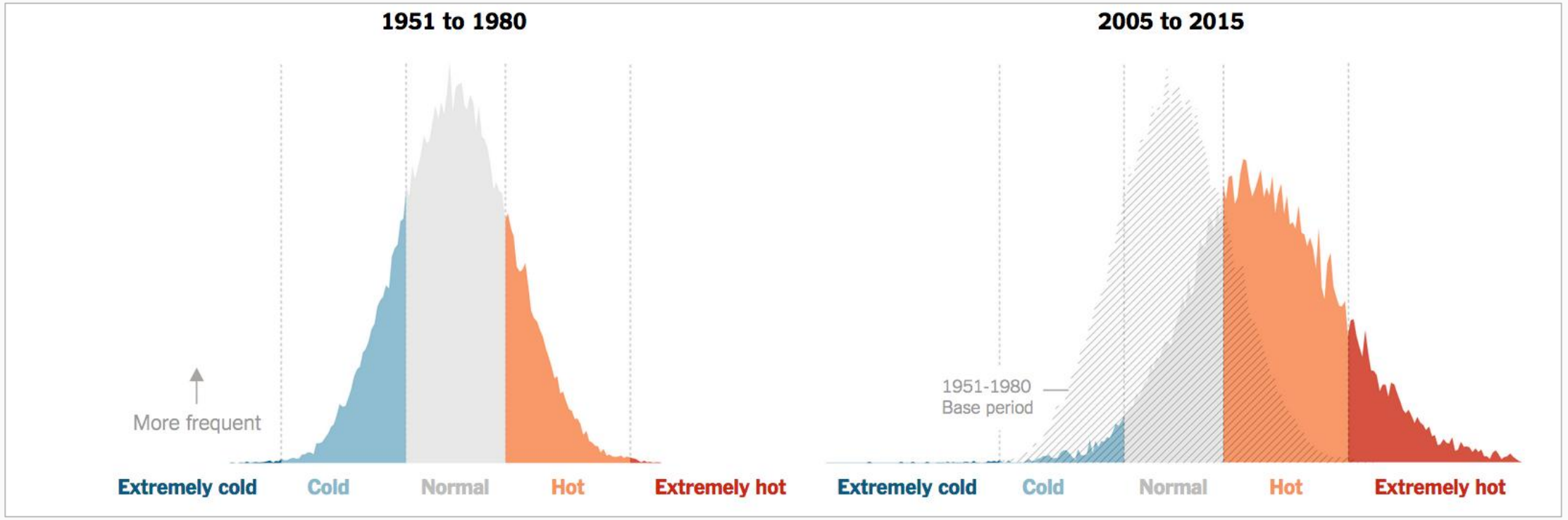
Statische vis





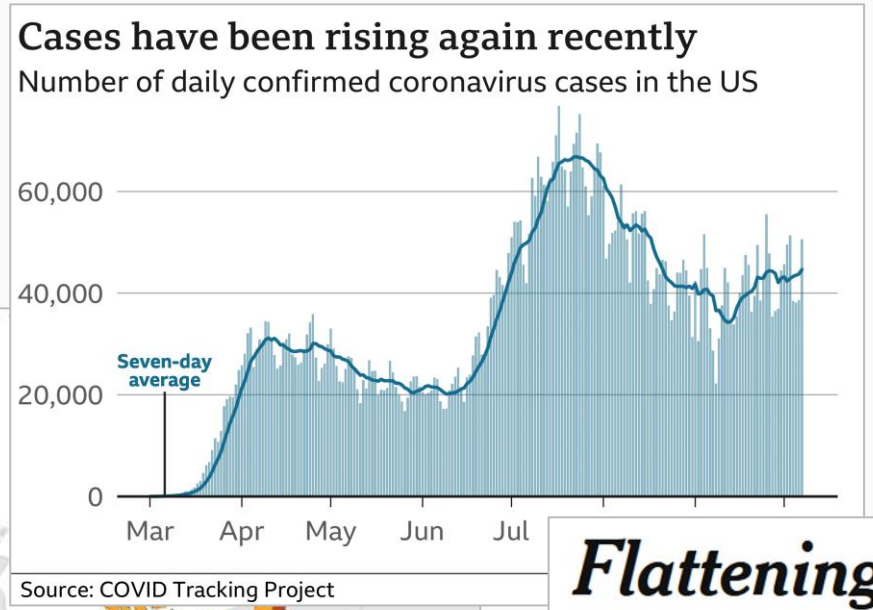
http://vizwiz.blogspot.in/2012/05/how-common-is-your-birthday-find-out.html
https://flowingdata.com/2018/01/23/the-demographics-of-others/

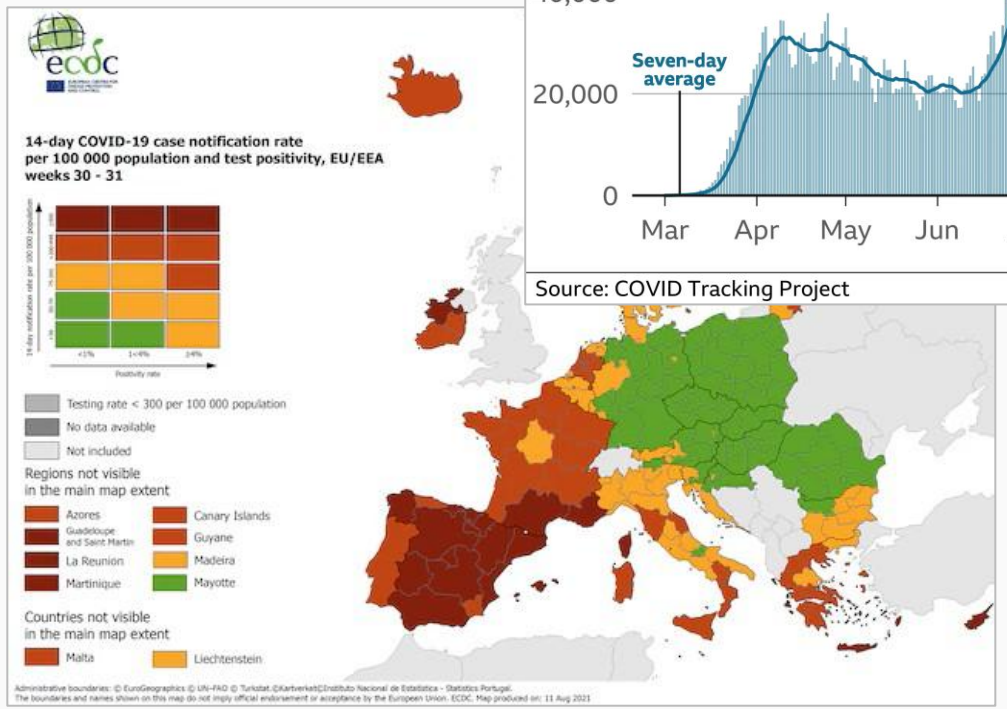
Statische vis

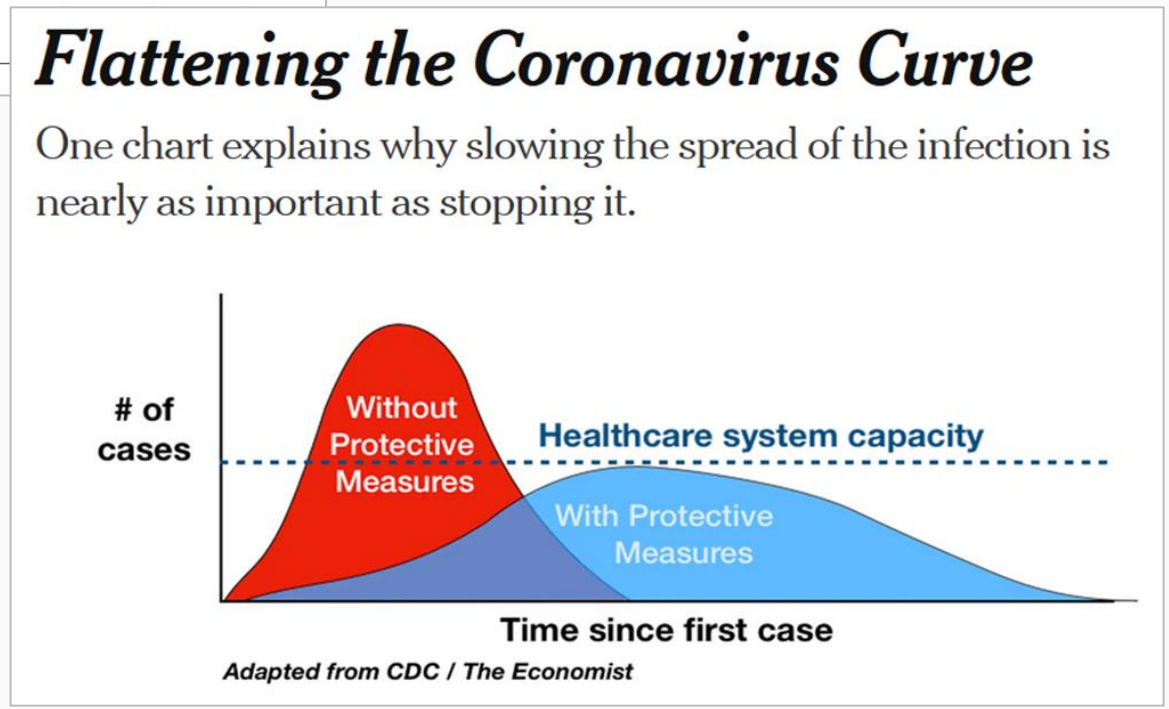


https://www.nytimes.com/interactive/2017/07/28/climate/more-frequent-extreme-summer-heat.html

Statische vis



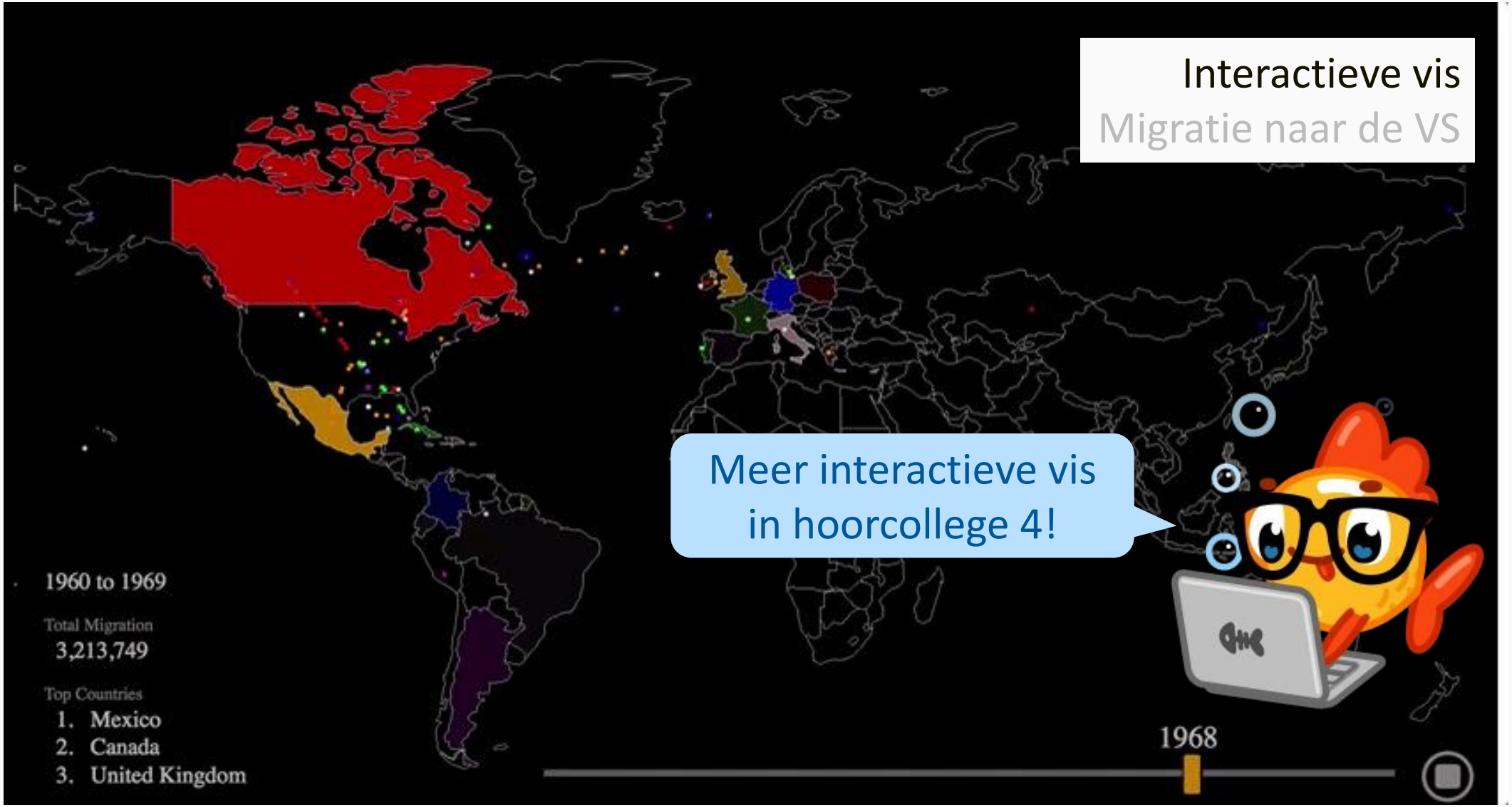




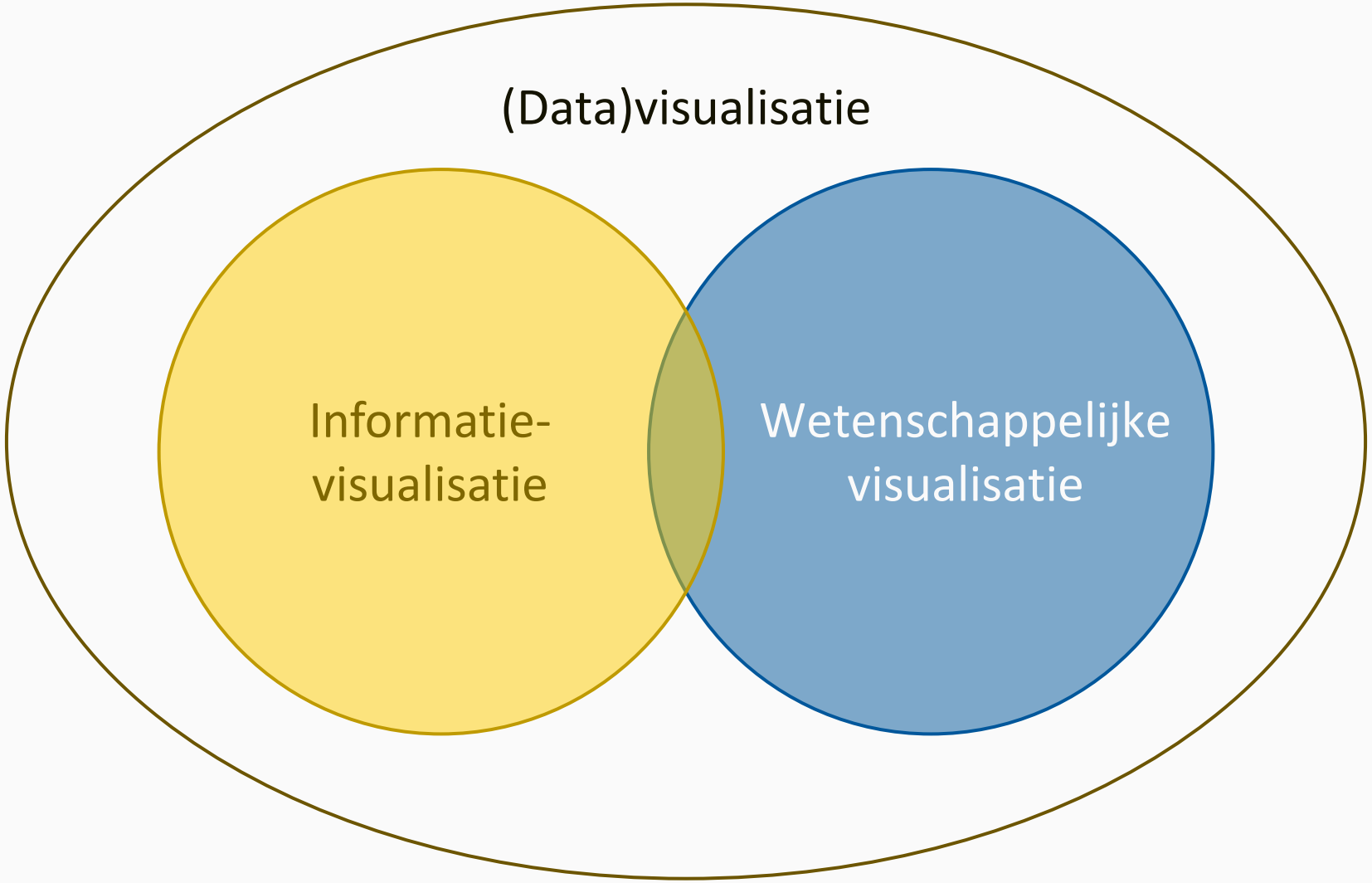
Visuele dashboards

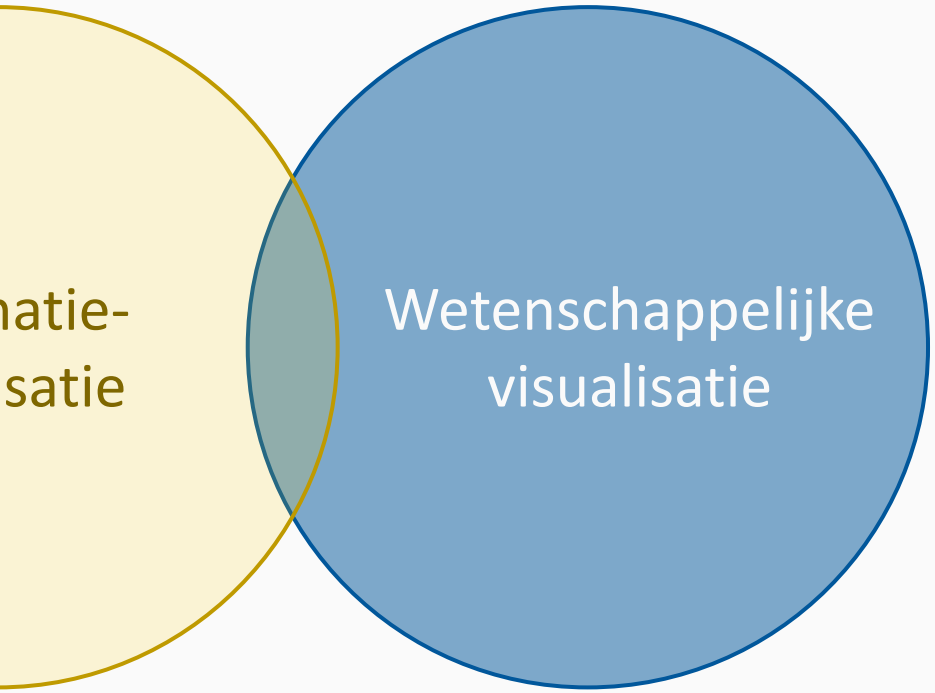
https://www.informationisbeautifulawards.com/showcase/4646-hr-attrition-dashboard

Interactieve vis
Migratie naar de VS



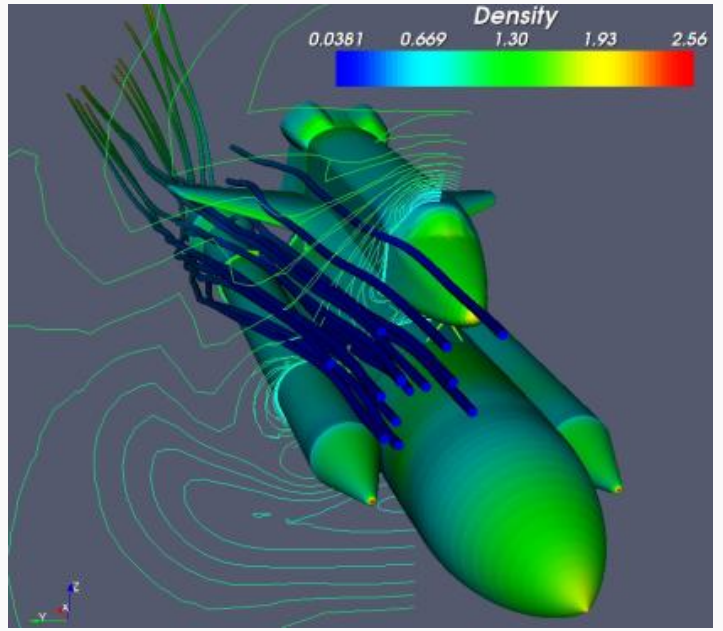
http://metrocosm.com/us-immigration-history-map.html

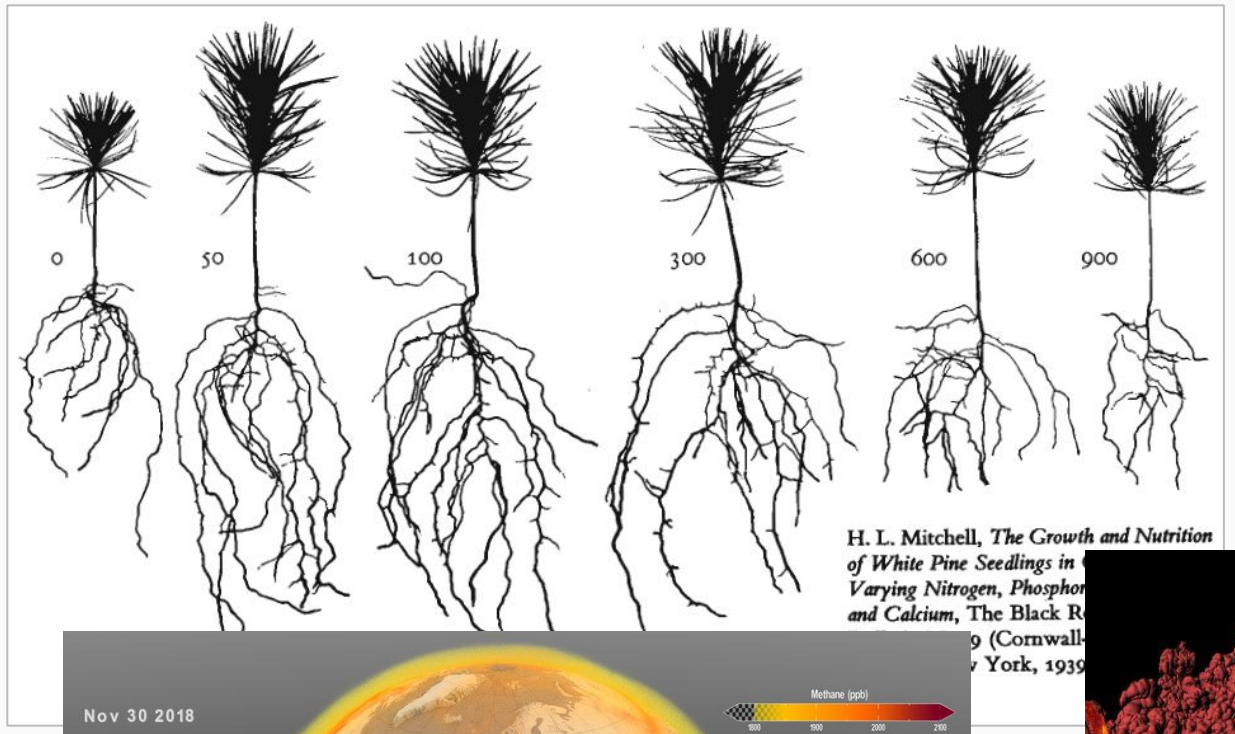

(Data)visualisatie
Informatie-
visualisatie
Wetenschappelijke
visualisatie

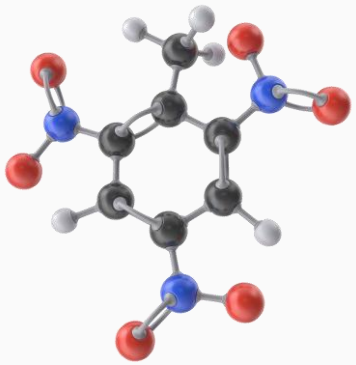


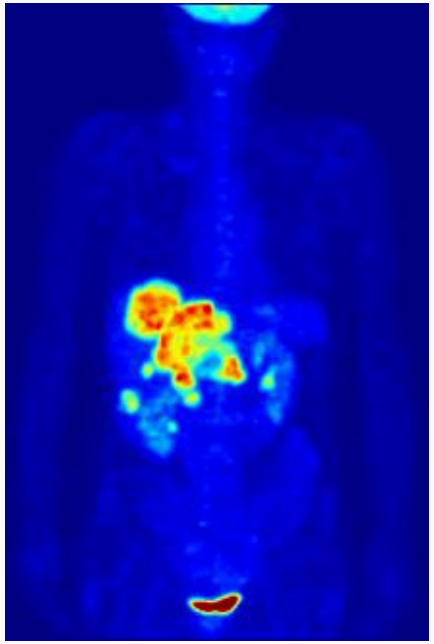
Data **met** welgedefinieerde representatie in 2D of 3D (vb. scan, simulaties ...)

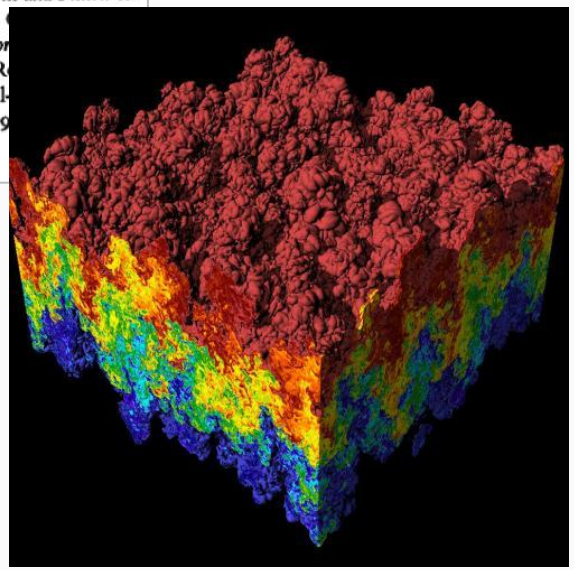
```
0265640 132304 133732 032051 037334 024721 015013 052226 001662
0265660 025537 064663 054606 043244 074076 124153 135216 126614
0265700 144210 056426 044700 042650 165230 137037 003655 006254
0265720 134453 124327 176005 027034 107614 170774 073702 067274
0265740 072451 007735 147620 061064 157435 113057 155356 114603
0265760 107204 102315 171451 046040 120223 001774 030477 046573
0266000 171317 116055 155117 134444 167210 041405 147127 050505
0266020 004137 046472 124015 134360 173550 053517 044635 021135
0266040 070176 047705 113754 175477 105532 076515 177366 056333
0266060 041023 074017 127113 003214 037026 037640 066171 123424
0266100 067701 037406 140000 165341 072410 100032 125455 056646
0266120 006716 071402 055672 132571 105645 170073 050376 072117
0266140 024451 007424 114200 077733 024434 012546 172404 102345
0266160 040223 050170 055164 164634 047154 126525 112514 032315
0266200 016041 176055 042766 025015 176314 017234 110060 014515
0266220 117156 030746 154234 125001 151144 163706 136237 164376
0266240 137055 062276 161755 115466 005322 132567 073216 002655
0266260 171466 126161 117155 065763 016177 014460 112765 055527
0266300 003767 175367 104754 036436 172172 150750 043643 145410
0266320 072074 000007 040627 070652 173011 002151 125132 140214
0266340 060115 014356 015164 067027 120206 070242 033065 131334
0266360 170601 170106 040437 127277 124446 136631 041462 116321
0266400 020243 005602 004146 121574 124651 006634 071331 102070
0266420 157504 160307 166330 074251 024520 114433 167273 030635
0266440 133614 106171 144160 010652 007365 026416 160716 100413
0266460 026630 007210 000630 121224 076033 140764 000737 003276
0266500 114060 042647 104475 110537 066716 104754 075447 112254
0266520 030374 144251 077734 015157 002513 173526 035531 150003
0266540 146207 015135 024446 130101 072457 040764 165513 156412
0266560 166410 067251 156160 106406 136770 030516 064740 022032
0266600 142166 123707 175121 071170 076357 037233 031136 015232
0266620 075074 016744 044055 102230 110063 033350 052765 172463
```

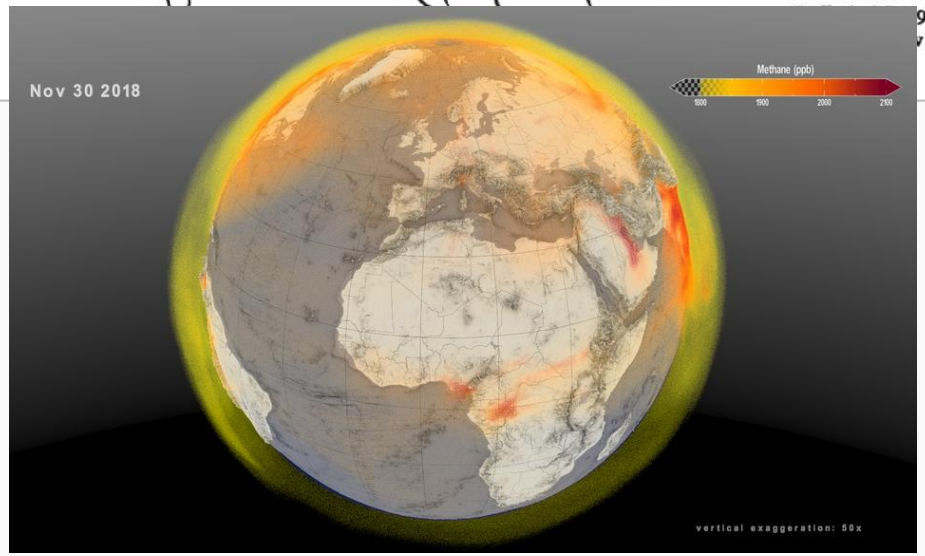


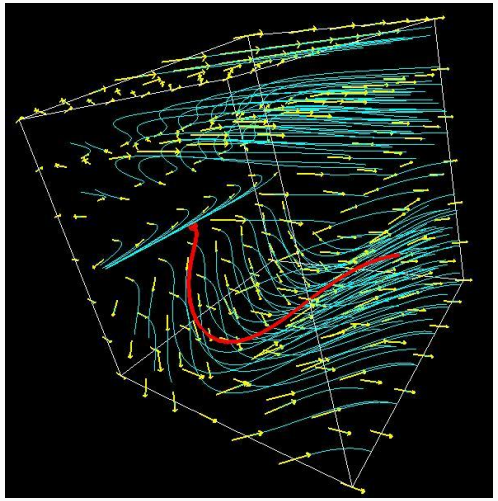




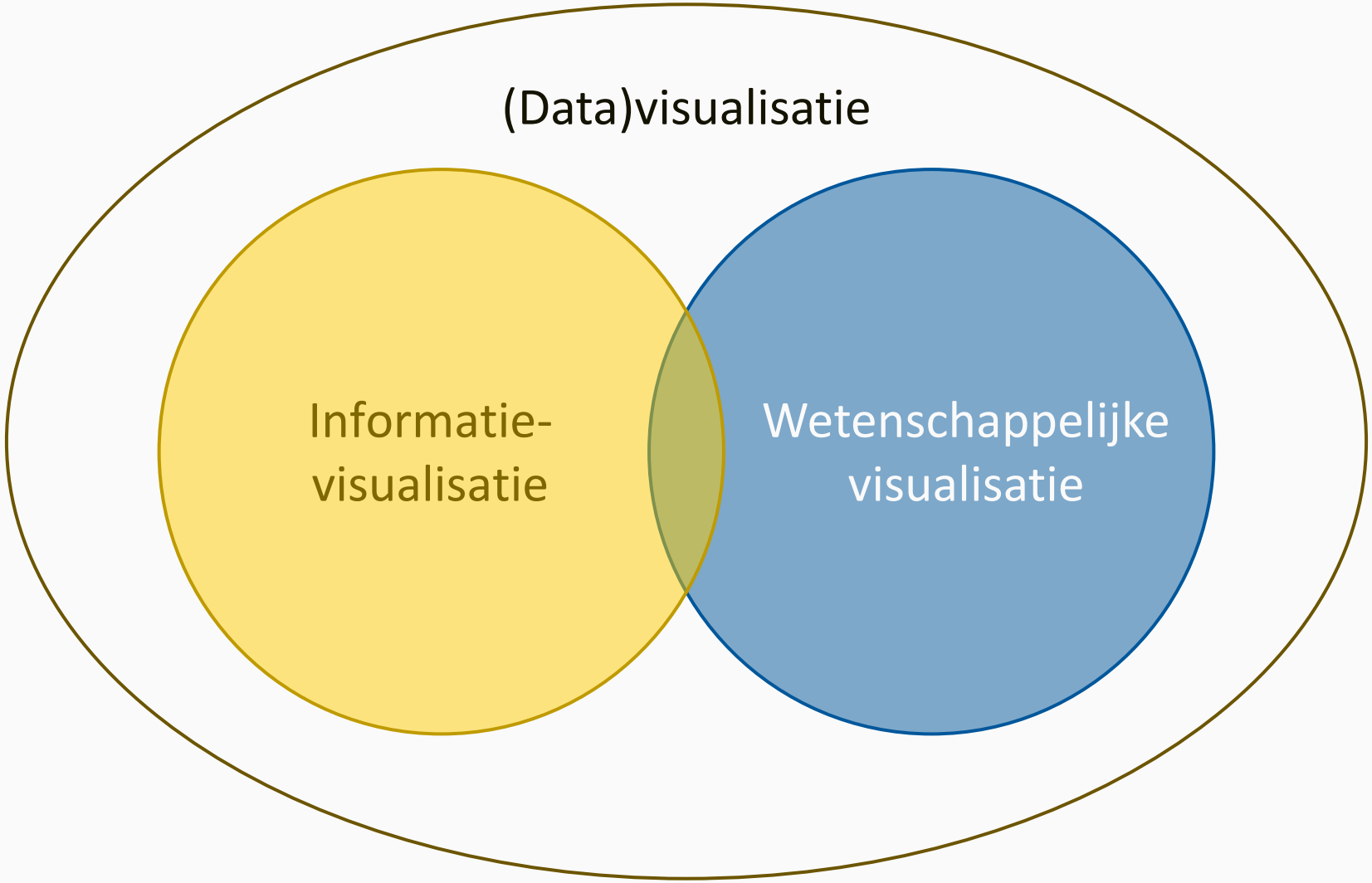






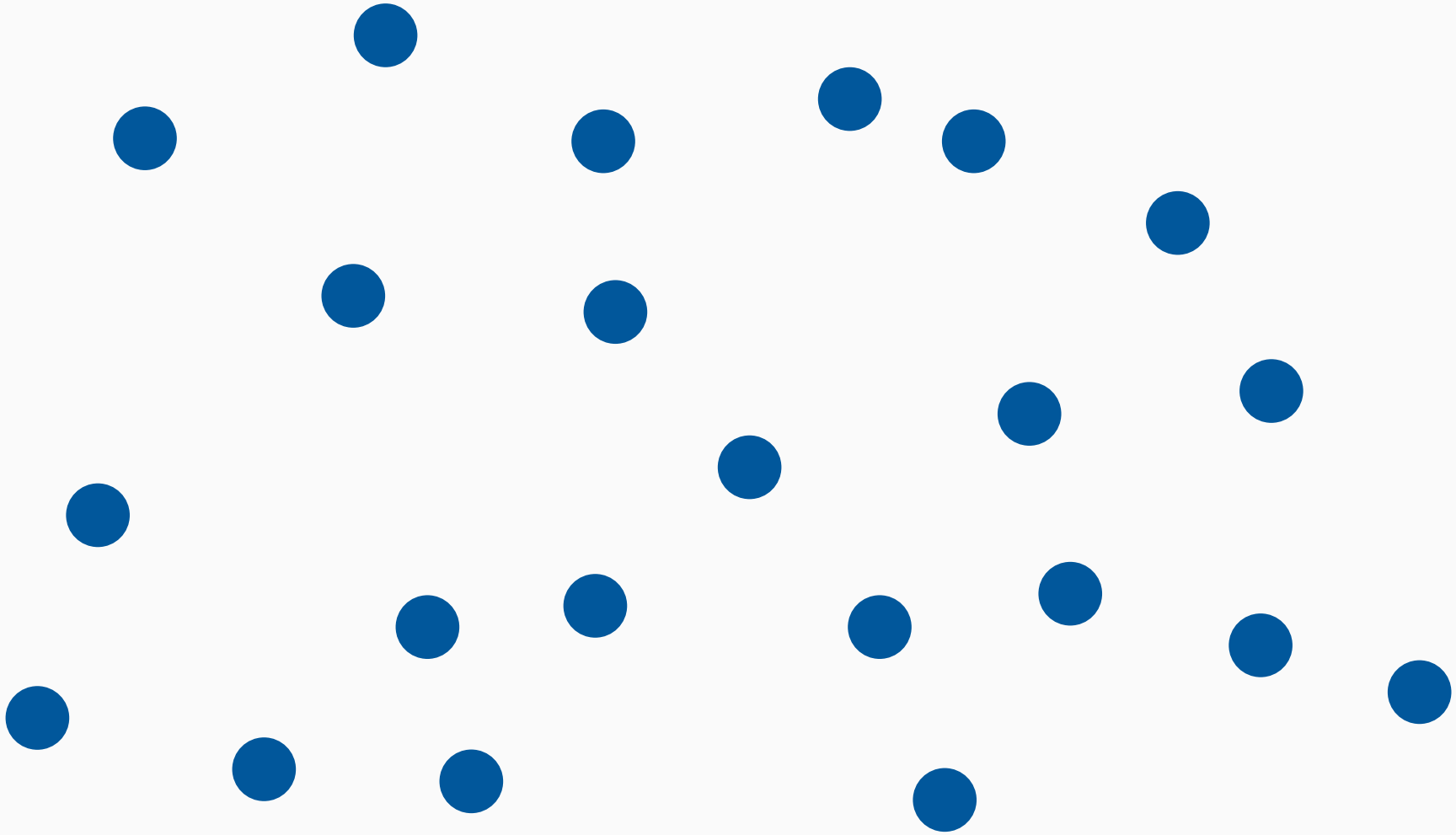


Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press. p. 50


(Data)visualisatie
Informatie-
visualisatie
Wetenschappelijke
visualisatie

Waarom is vis belangrijk?

Hoeveel cirkels zijn er?



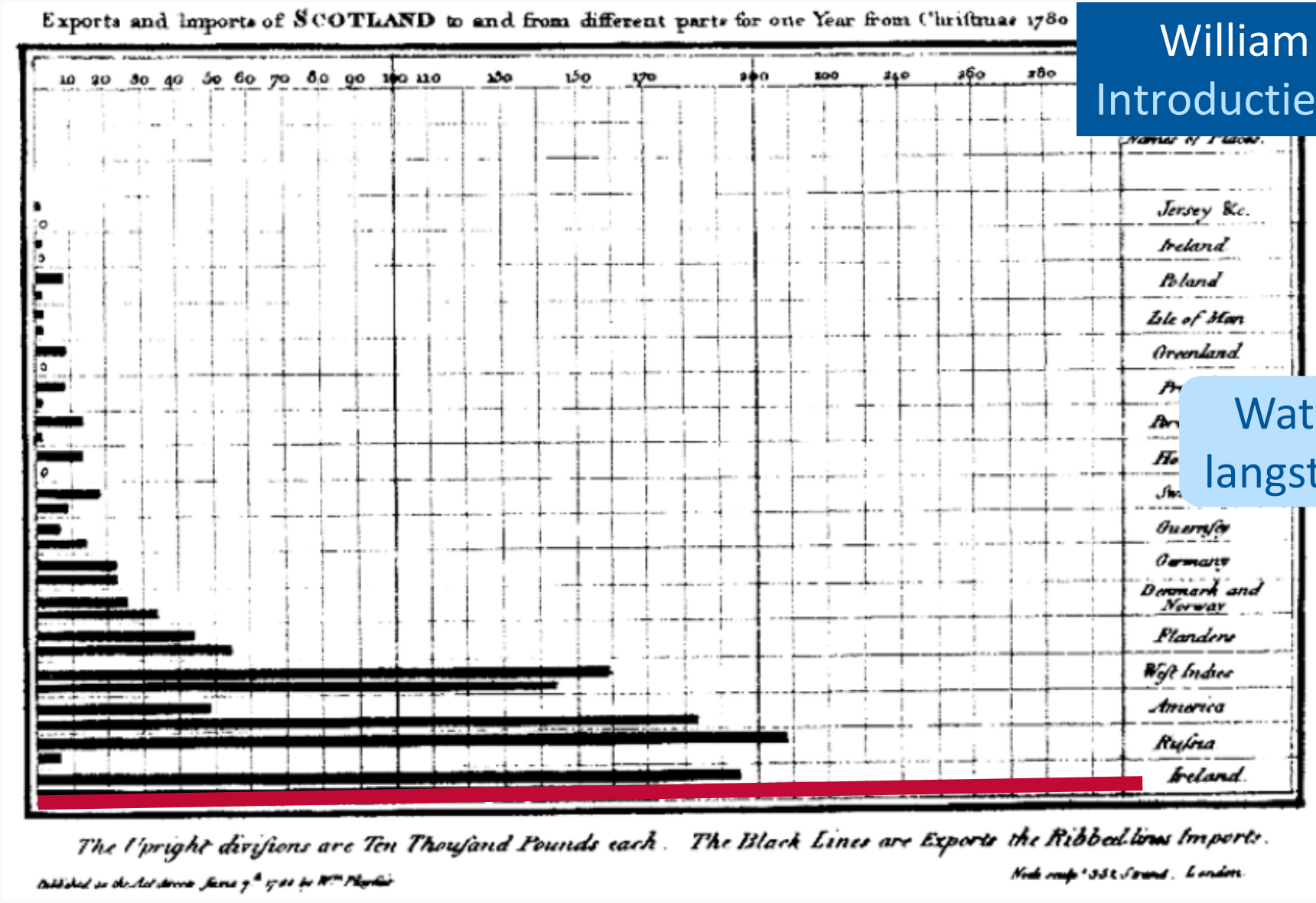
Hoeveel cirkels zijn er?

Mensen hebben sterke perceptuele vaardigheden
en kunnen erg goed patronen herkennen

10 10 2

Slide source: Joris Klerkx

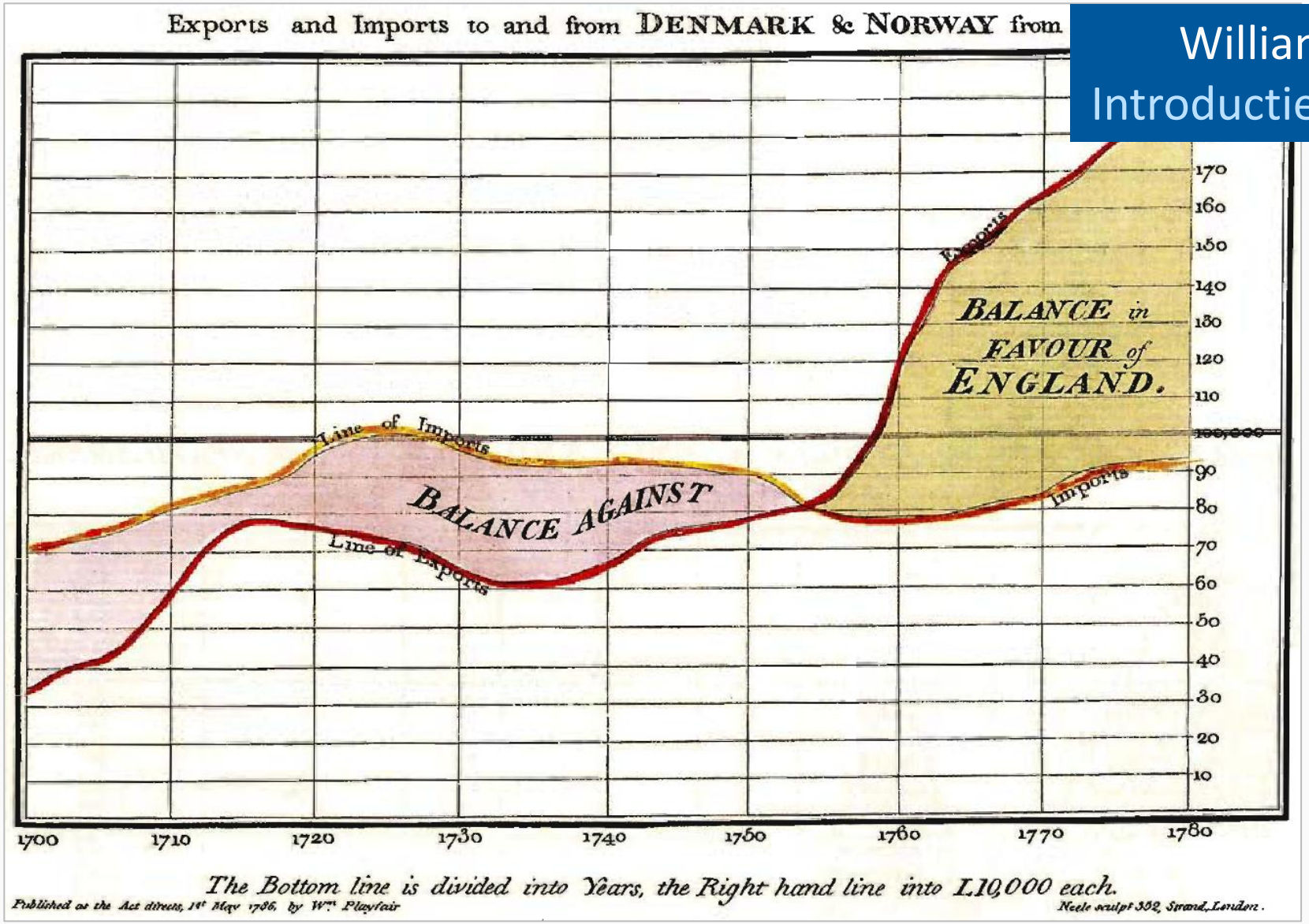
William Playfair (1786)
Introductie van bar charts



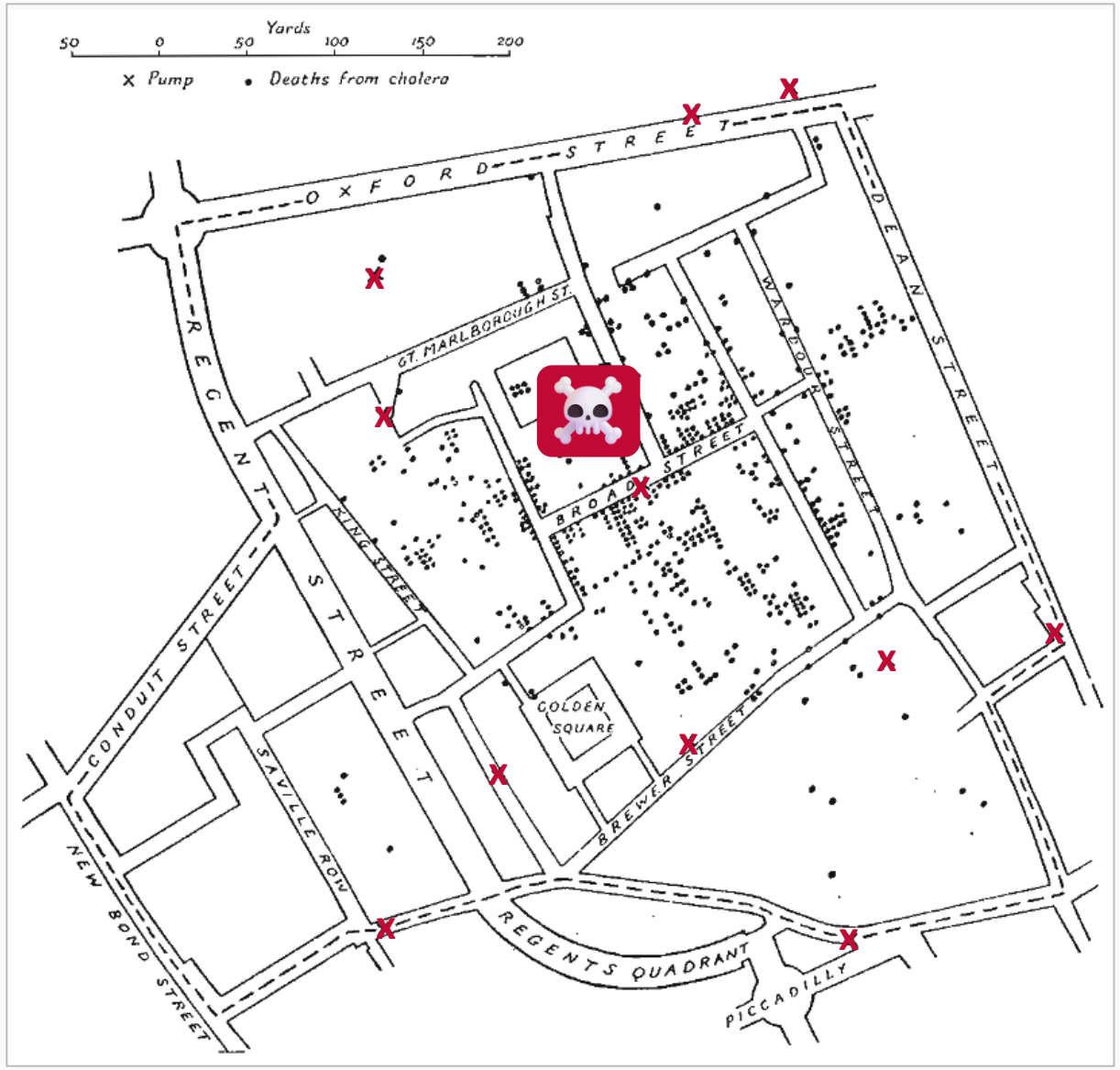
Wat is de langste bar?

Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press. p. 33

William Playfair (1786)
Introductie van area charts



Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press. p. 92



John Snow (1854)
Oorzaak cholera-uitbraak

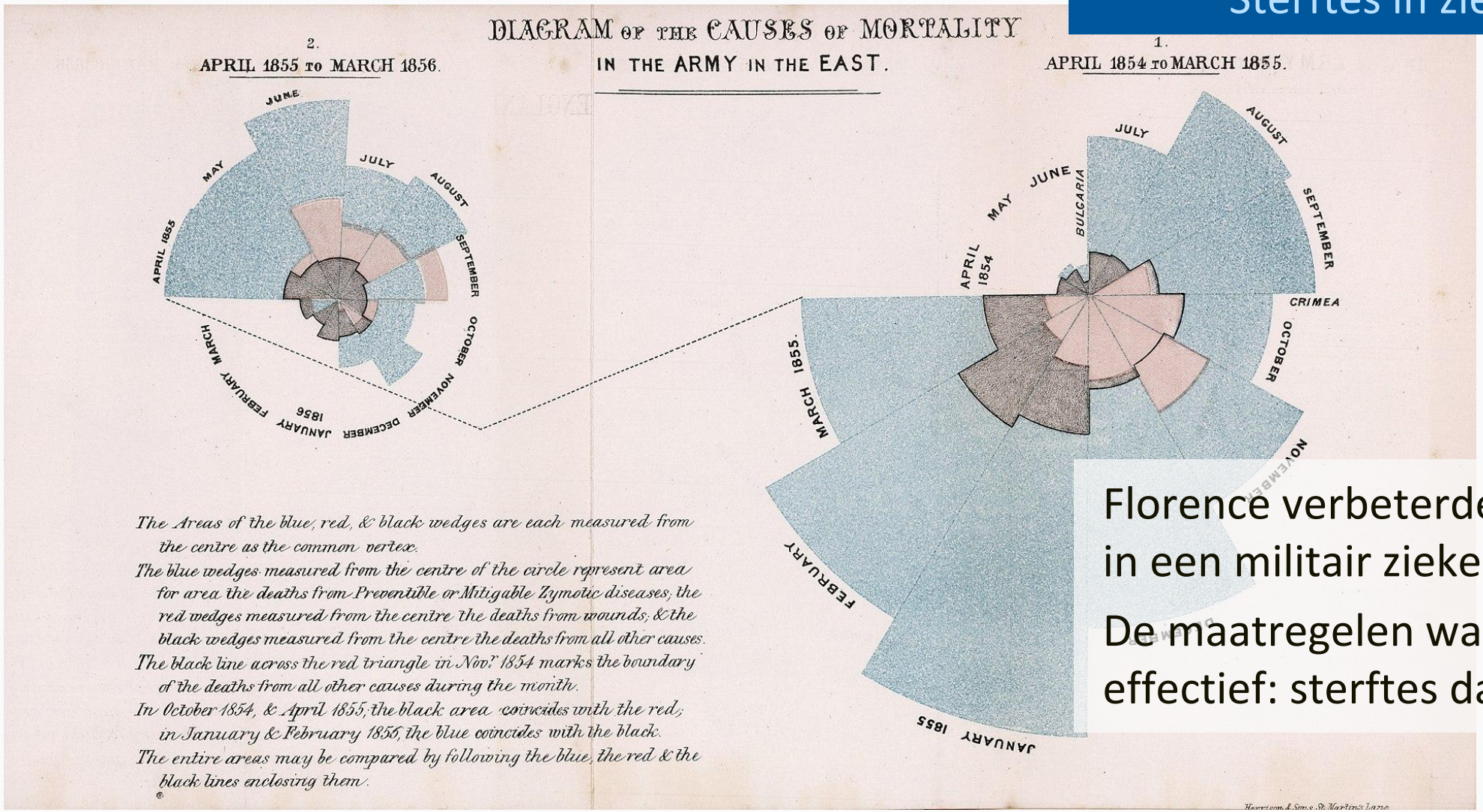
Link tussen • choleraslachtoffers en x waterpompen

Afsluiting van de pomp in Broad Street stopte de cholera-uitbraak

Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press. p. 24

Source: Information Visualization Course, Katy Borner

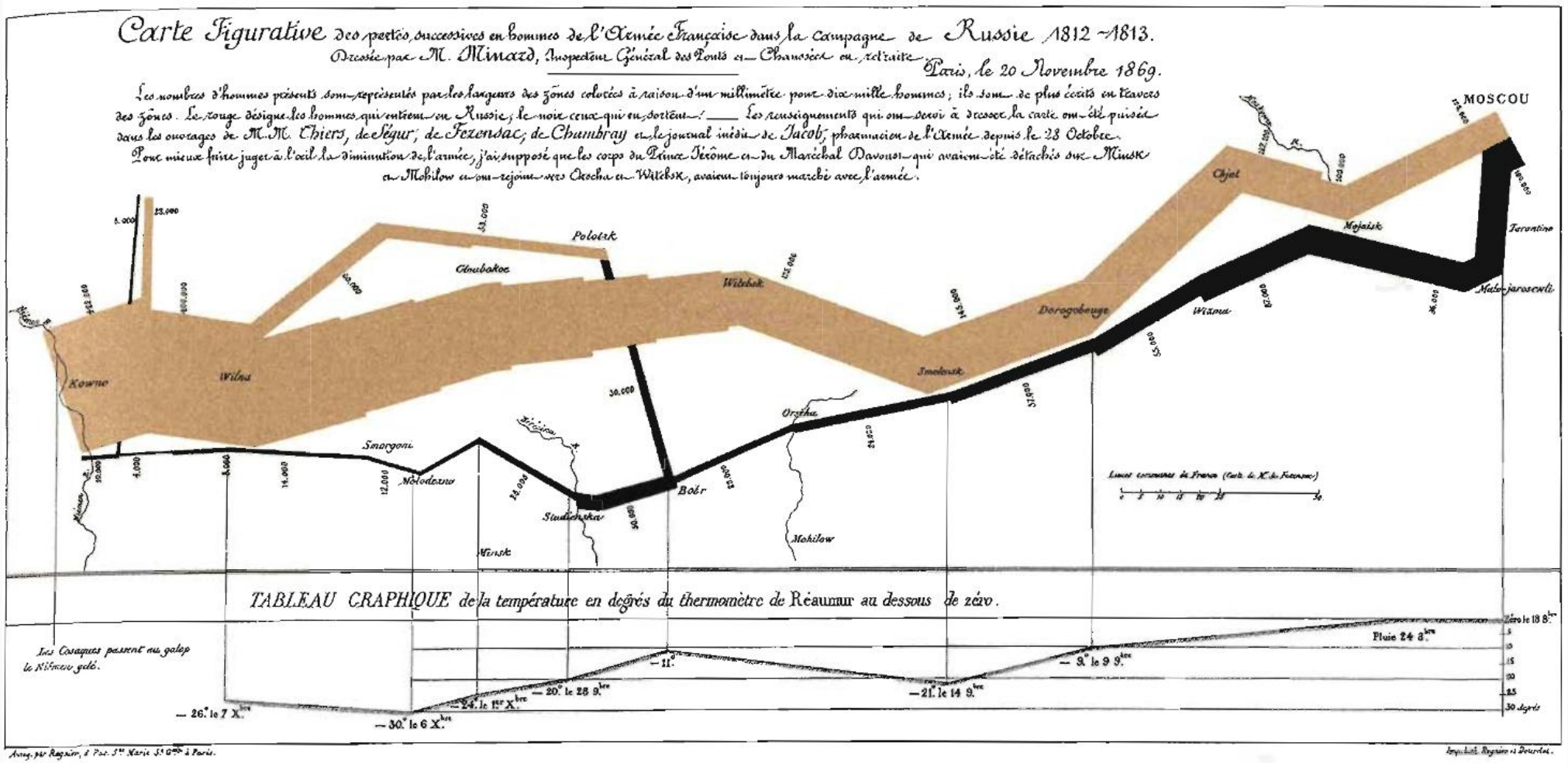
Florence Nightingale (1858)
Sterftes in ziekenhuis



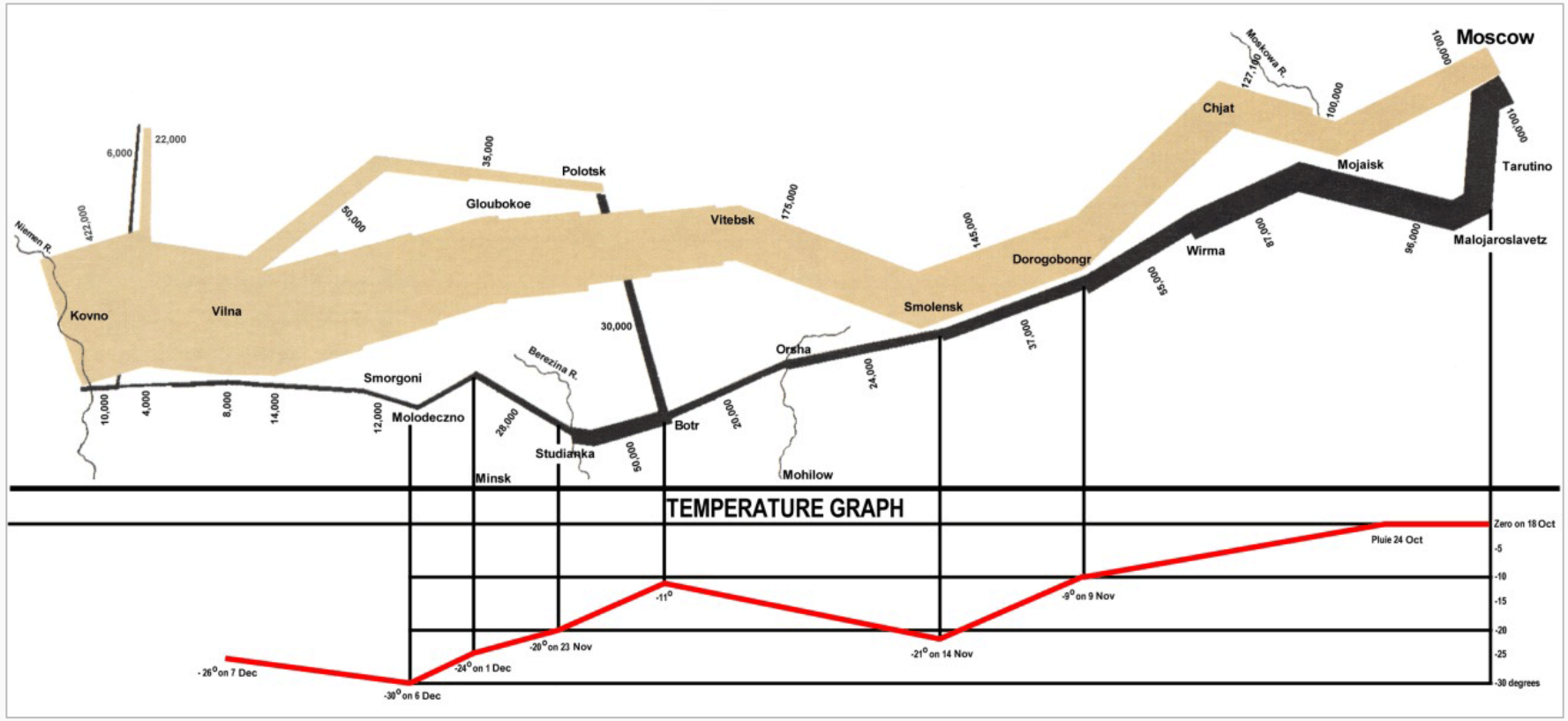
Florence verbeterde hygiëne in een militair ziekenhuis

De maatregelen waren effectief: sterftes daalden

Charles Minard (1869)
Napoleons leger



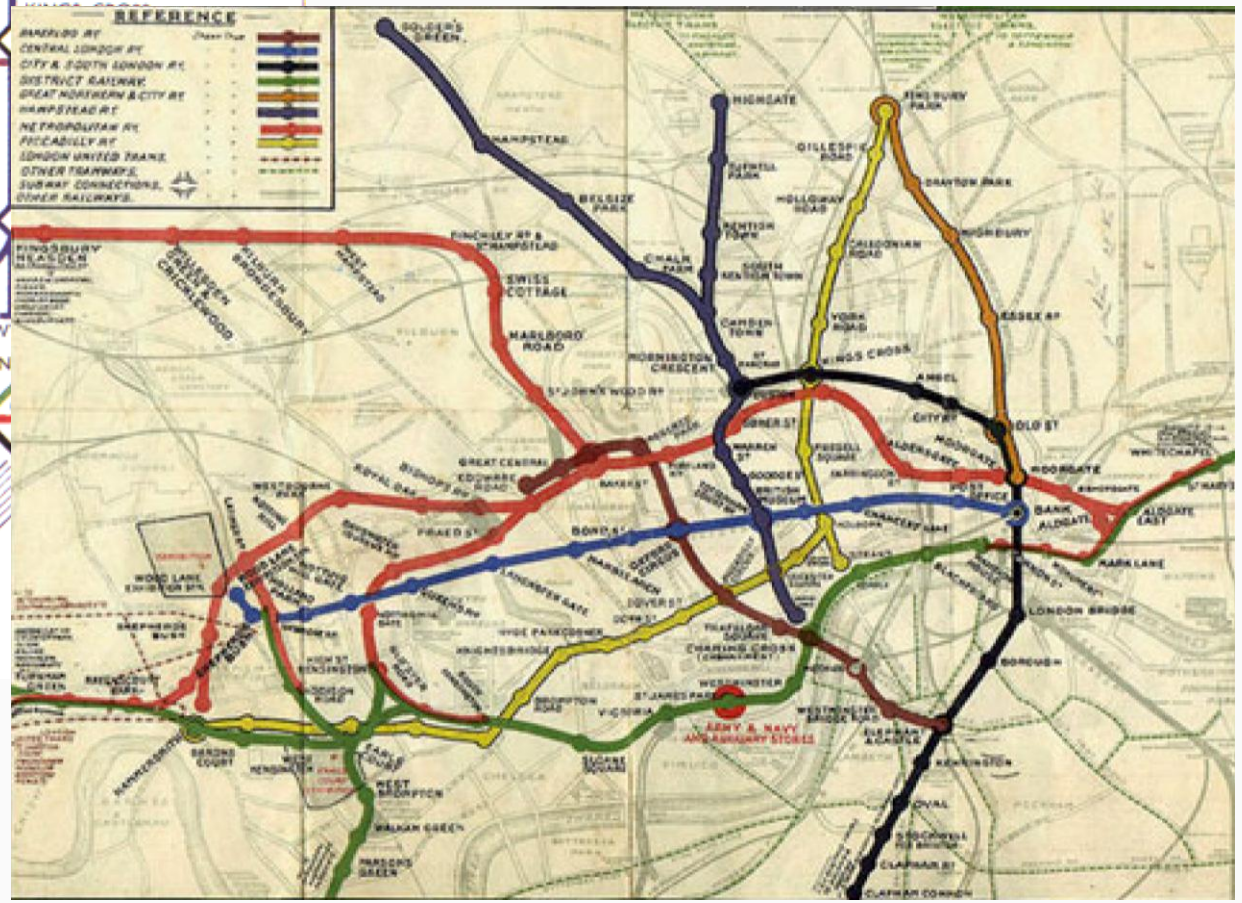
Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press. p. 41



Bruine lijn = soldaten bij de inval (422.000 vertrekken)
Zwarte lijn = soldaten bij terugtrekking (10.000 keren terug)
De vriestemperaturen (rode lijn) leidden tot veel doden

http://www.napoleanic-literature.com

Harry Beck (1933)
Metro in Londen



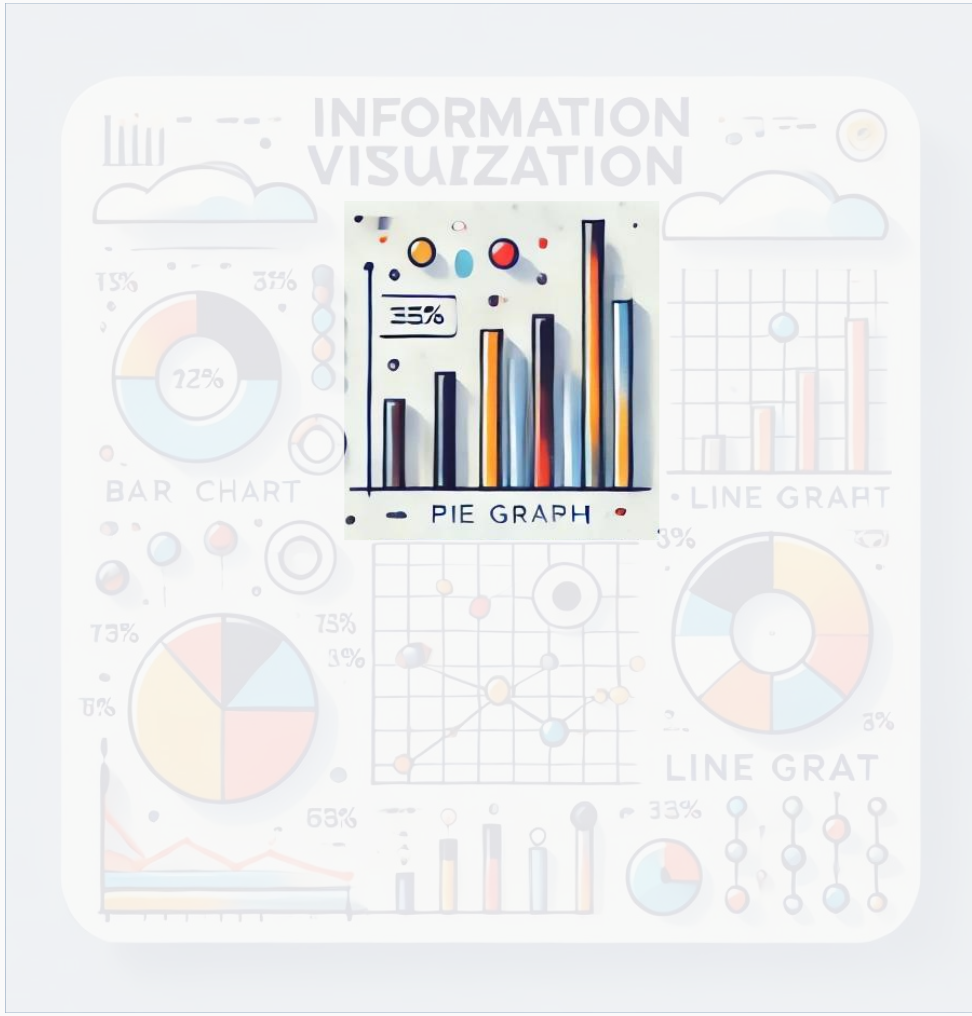
Robert Spence. 2014. *Information Visualization: An Introduction*. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-07341-5

Source: Information Visualization Course, Katy Börner

The beauty of data visualization. David McCandless. http://www.youtube.com/watch?v=5Zg-C8AAIGg

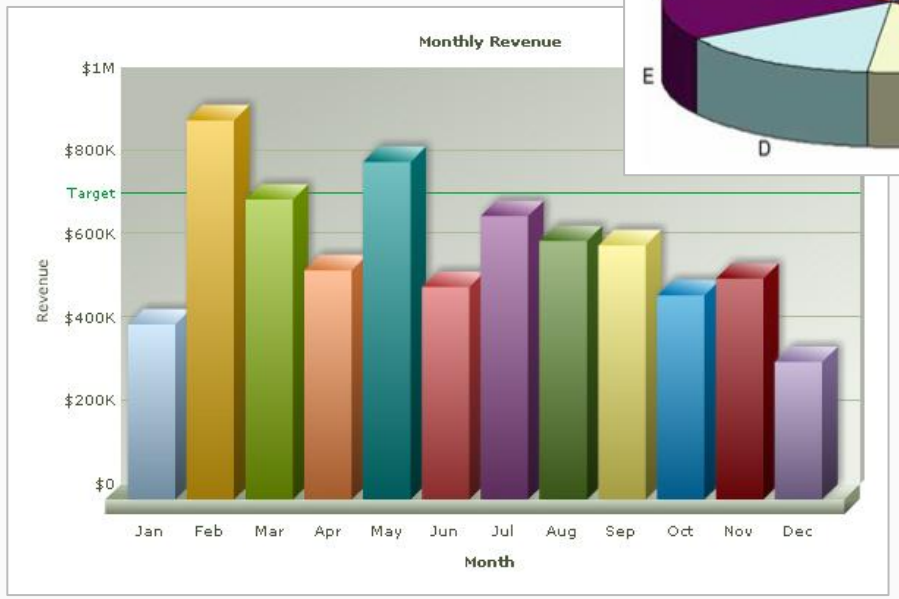
Algemene principes en visualisatietypes

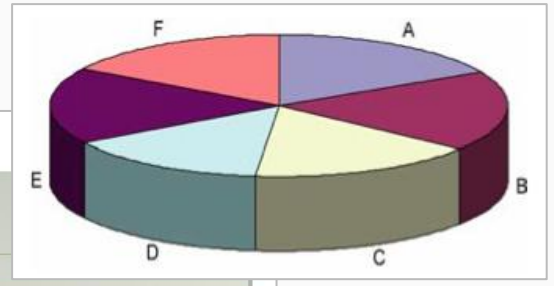
1. Voorbeelden en geschiedenis
2. Algemene vis-principes
3. Klassieke vis-types
4. Tools voor vis

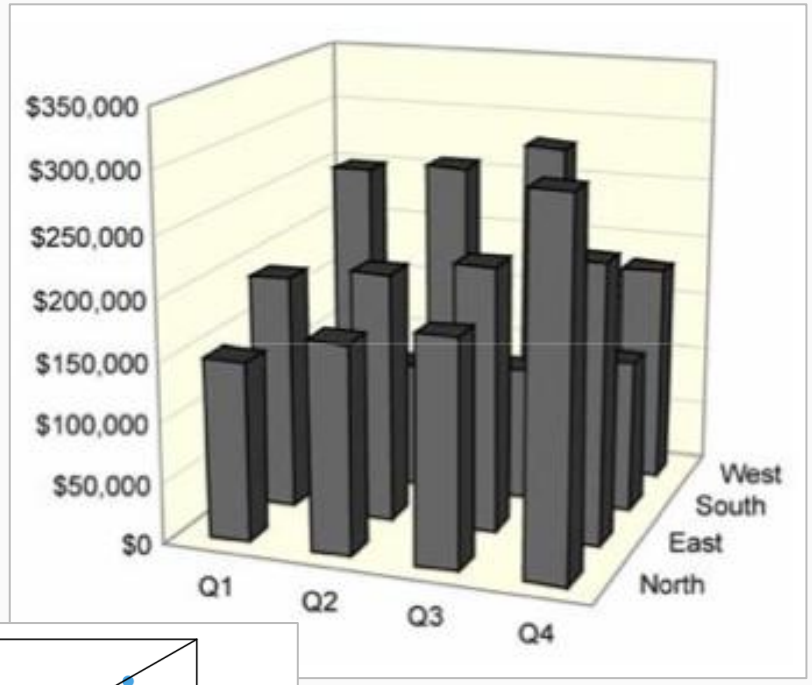

INFORMATION VISUIZATION
35%
PIE GRAPH
BAR CHART
LINE GRAPH
LINE GRAT

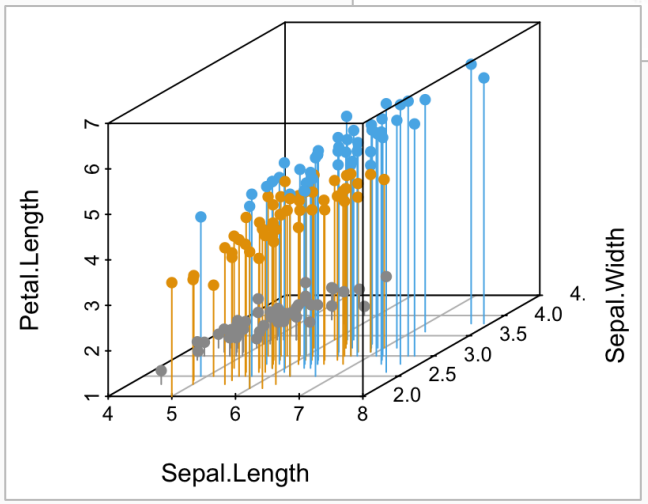
GENERAL PRINCIPLES FOR INFORMATION VISUIZATION
BAR CHARTS
PIE GRAPHS
BAR CARTS
SCATTER PLOTS

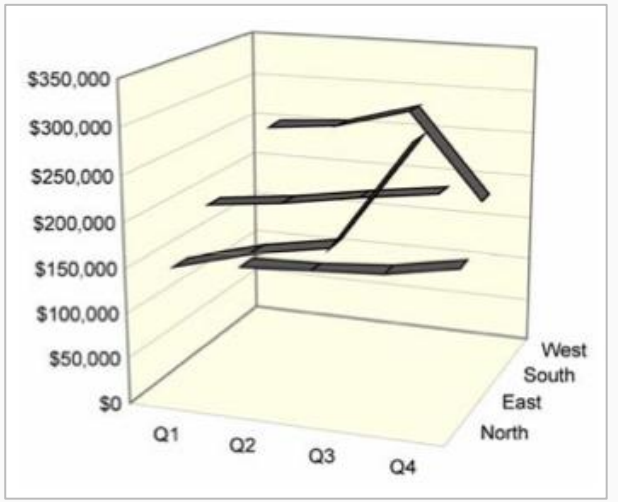
Vermijd 3D 🤮







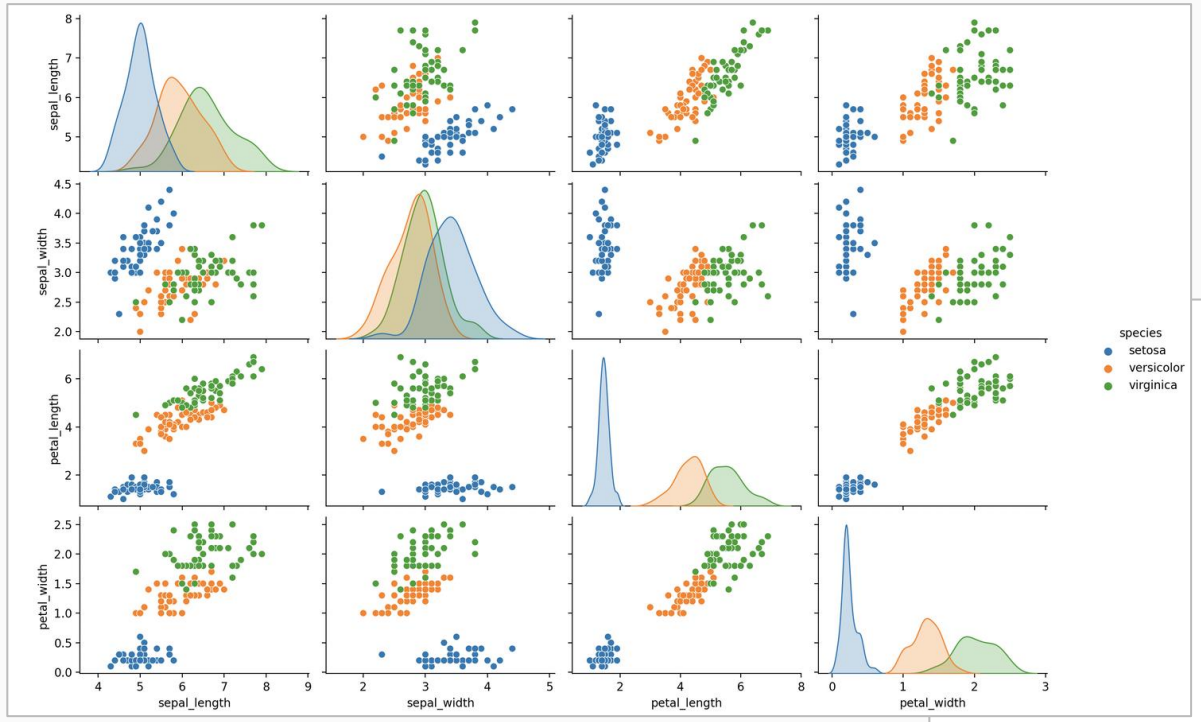




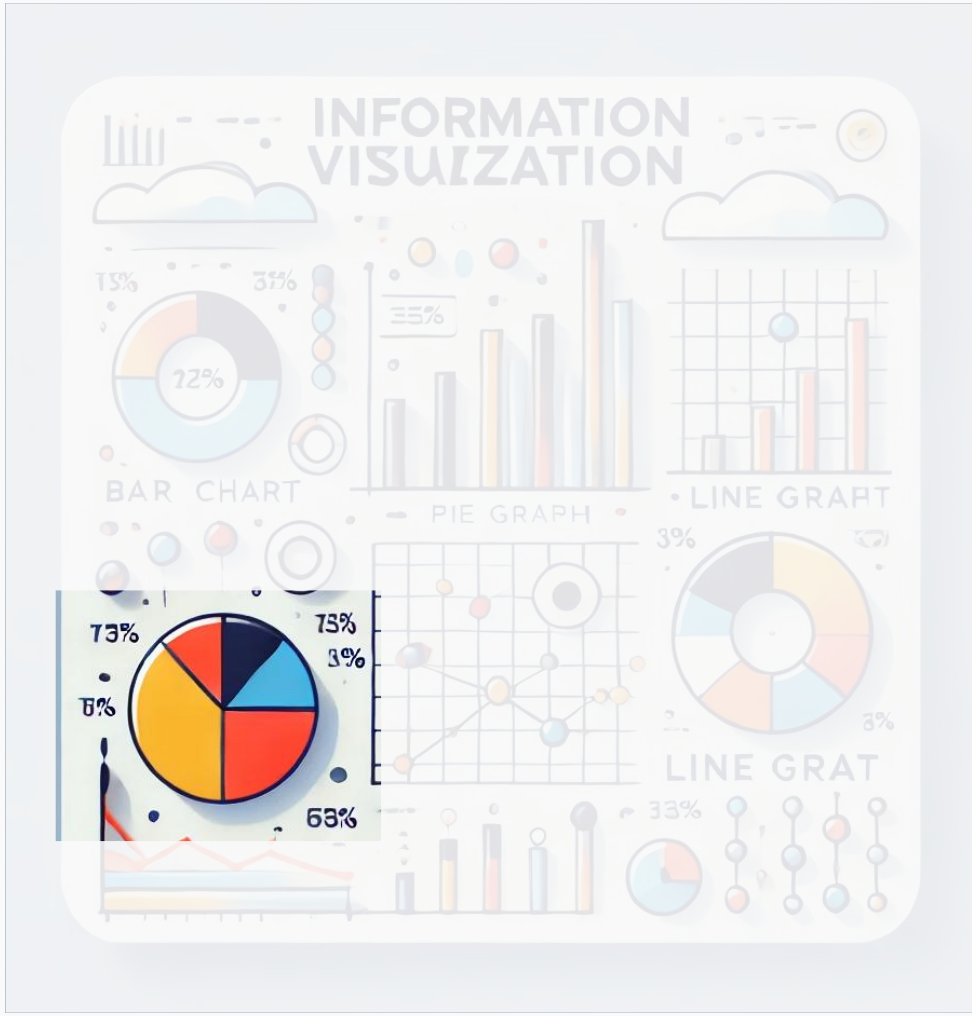
Perspectief vertekent lengtes en posities

Source: Stephen Few

Oplossingen voor 3 of meer variabelen

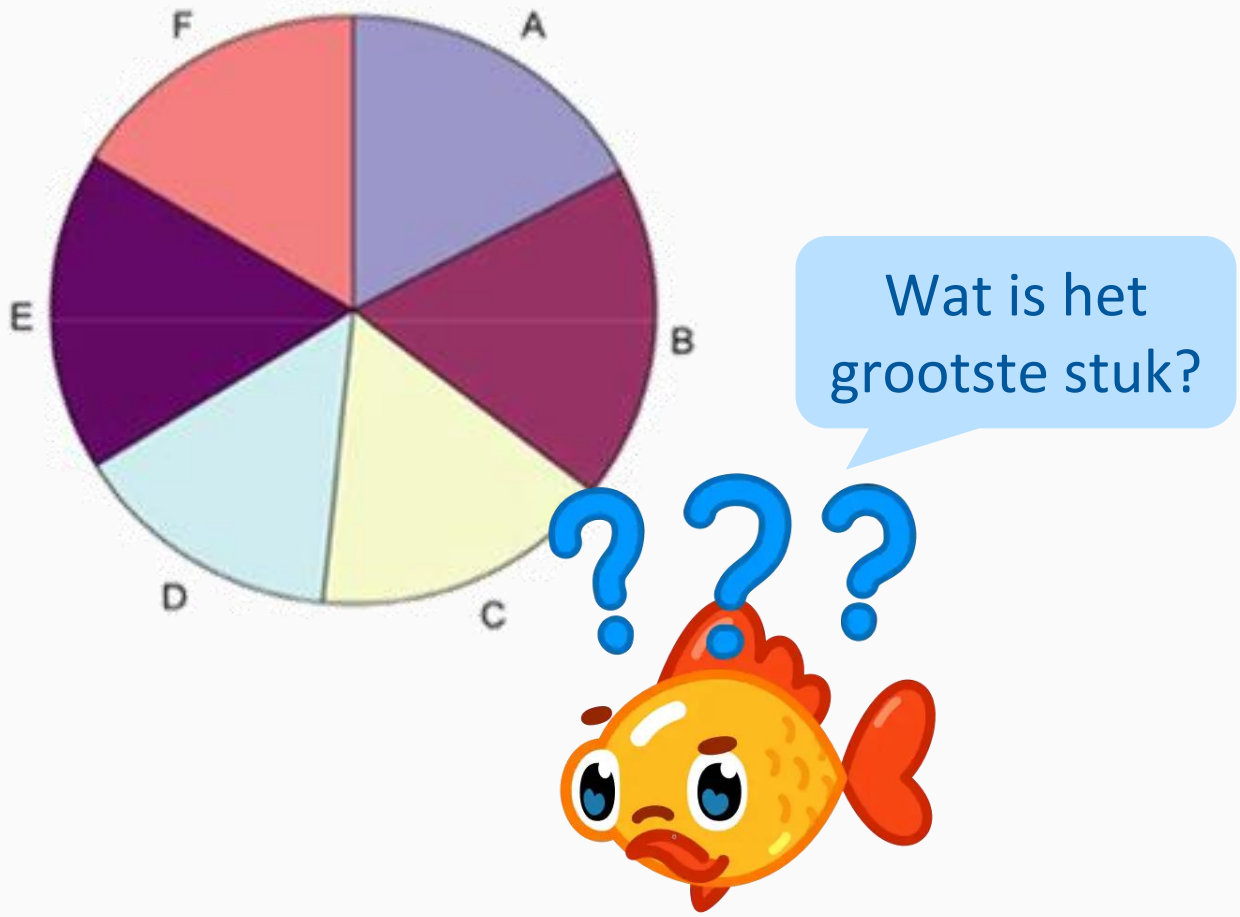


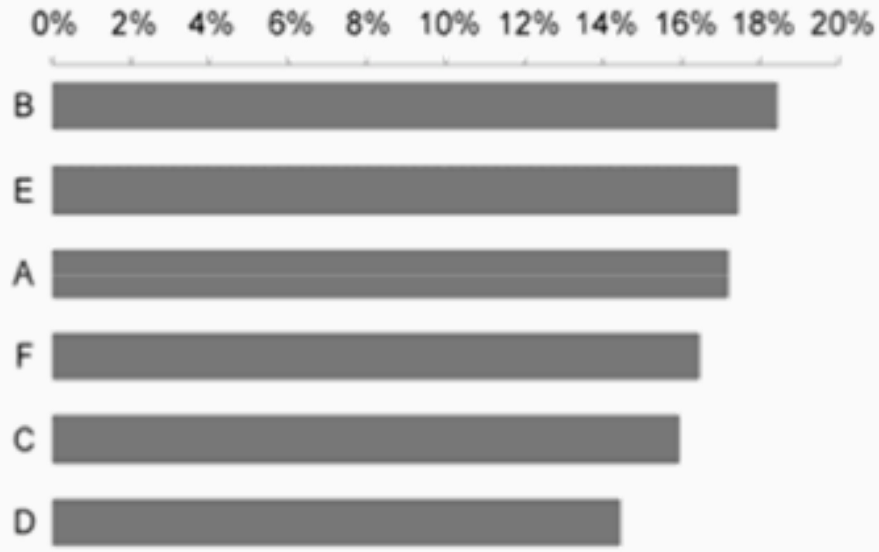
Source: Stephen Few


INFORMATION
VISUIZATION
BAR CHART
PIE GRAPH
LINE GRAPH
LINE GRAT

GENERAL PRINCIPLES FOR
INFORMATION VISUIZATION
BAR CHARTS
PIE GRAPHS
BAR CARTS

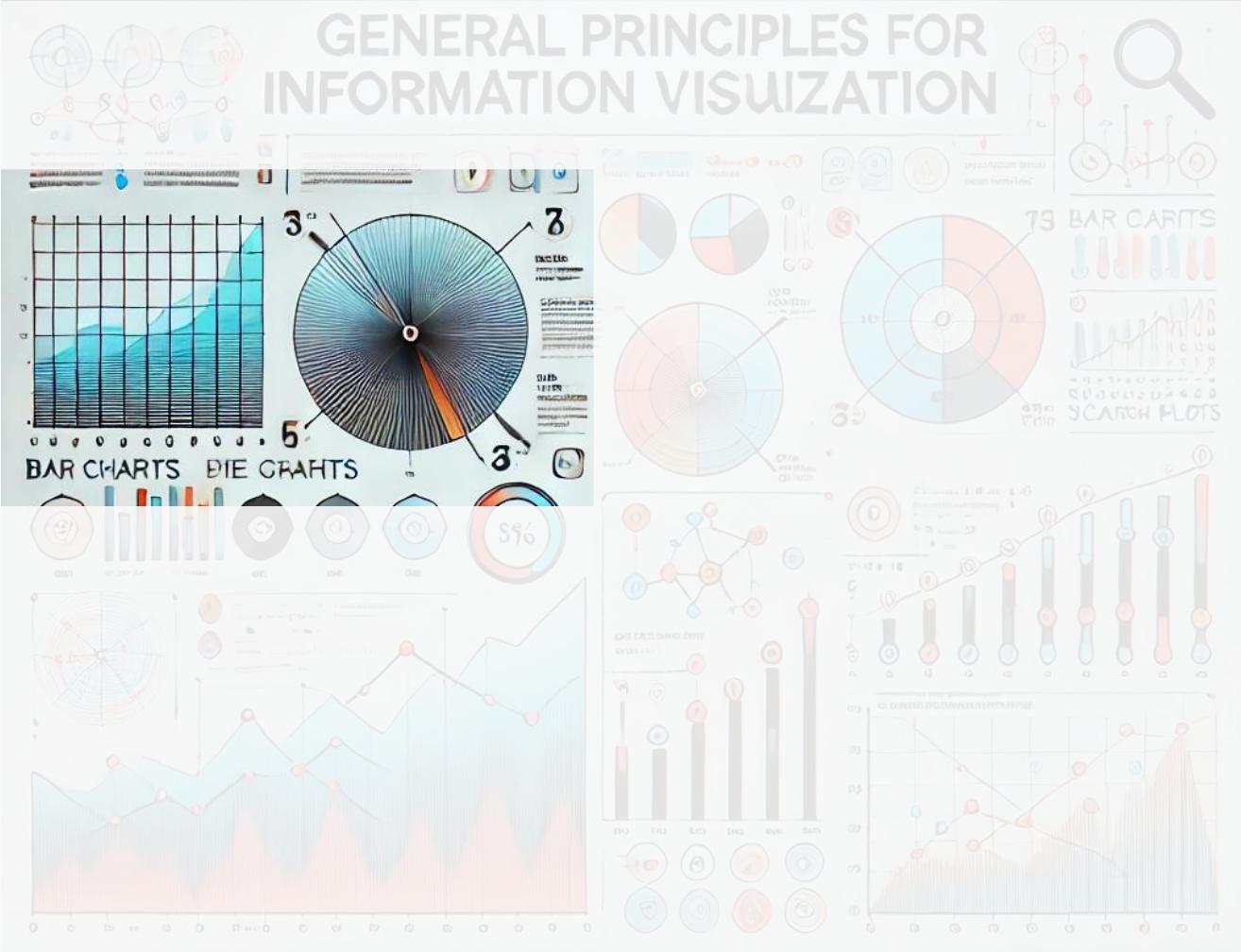
Vermijd (vaak) pie charts



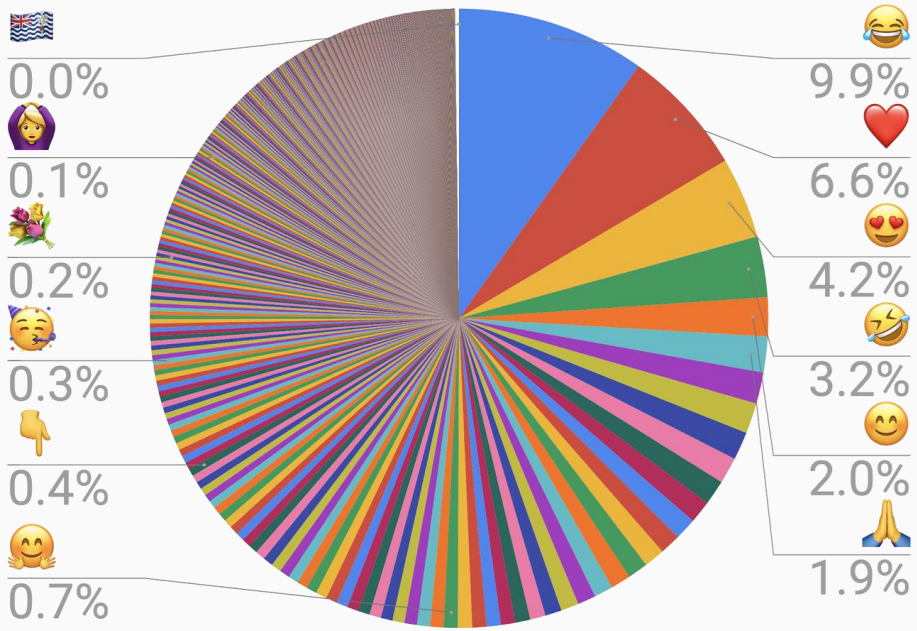


Source: Stephen Few

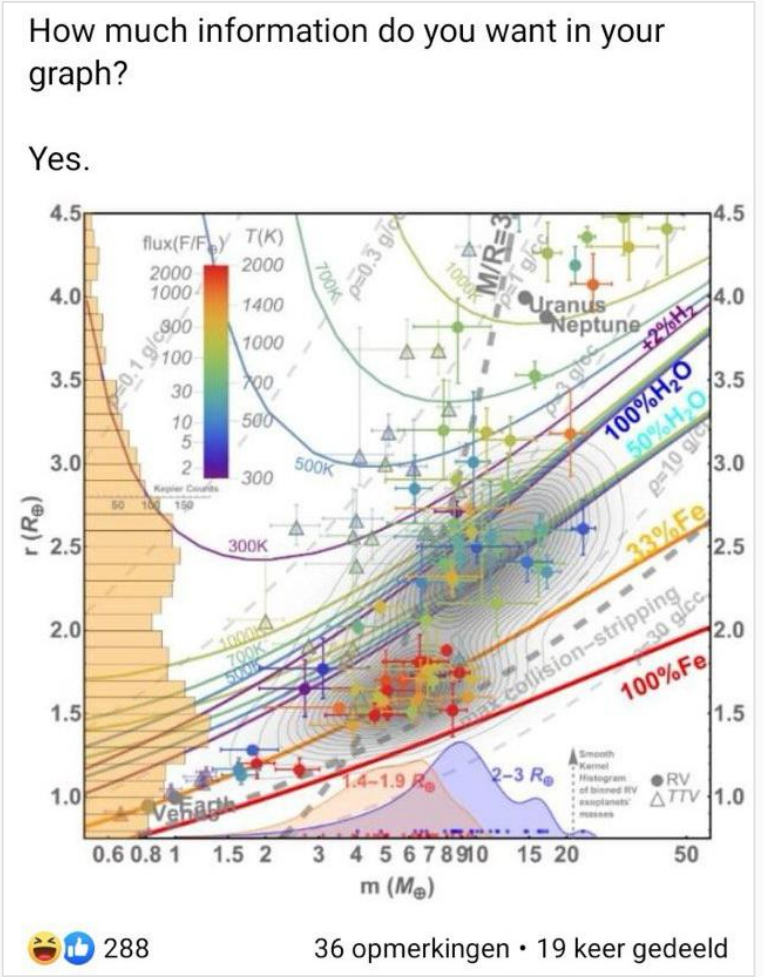
INFORMATION
VISUIZATION
BAR CHART
PIE GRAPH
LINE GRAPH
LINE GRAT


GENERAL PRINCIPLES FOR
INFORMATION VISUIZATION
BAR CHARTS
PIE CHARTS
BAR CHARTS

Zorg voor een duidelijke boodschap

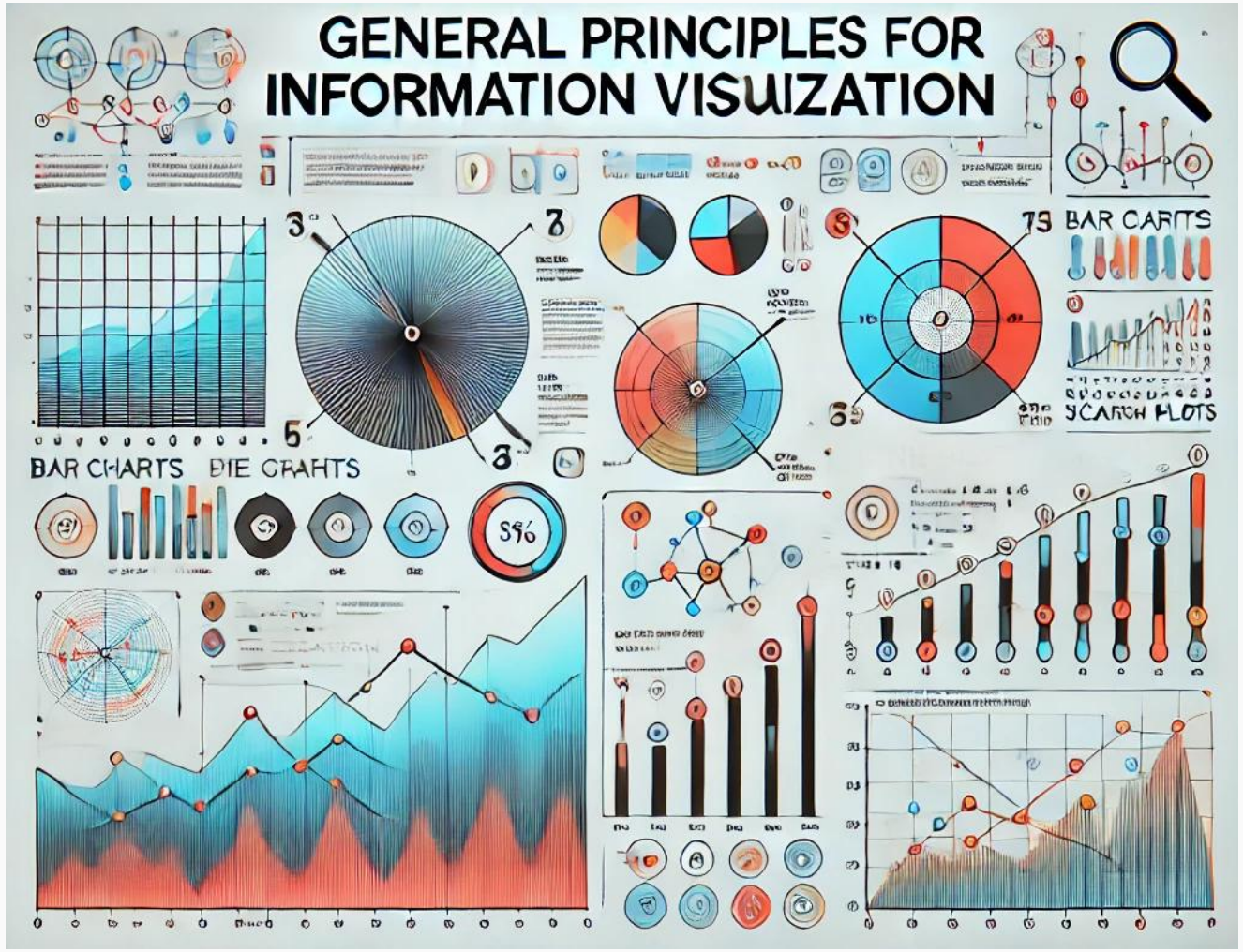


Pie charts met te veel categorieën of vis met te veel data zijn niet inzichtelijk



https://viz.wtf/post/189495004858/the-most-frequent-emoji-the-unicode-consortium


INFORMATION VISUIZATION
BAR CHART
PIE GRAPH
LINE GRAPH
LINE GRAT

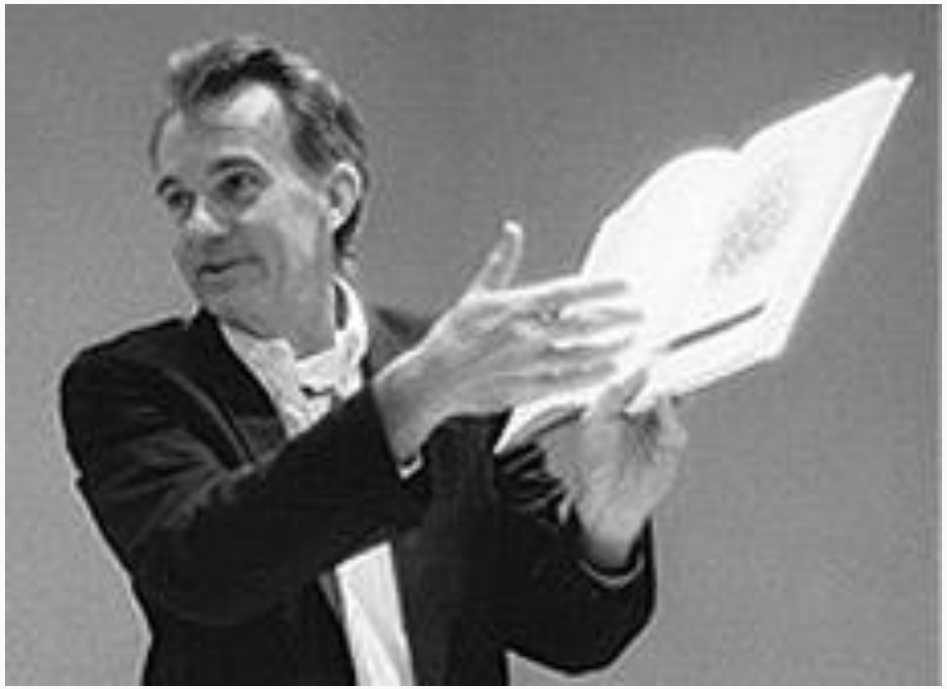

GENERAL PRINCIPLES FOR INFORMATION VISUIZATION
BAR CHARTS
PIE CHARTS
BAR CARTS
SCATCH PLOTS

Visualisatieprincipes van Tufte

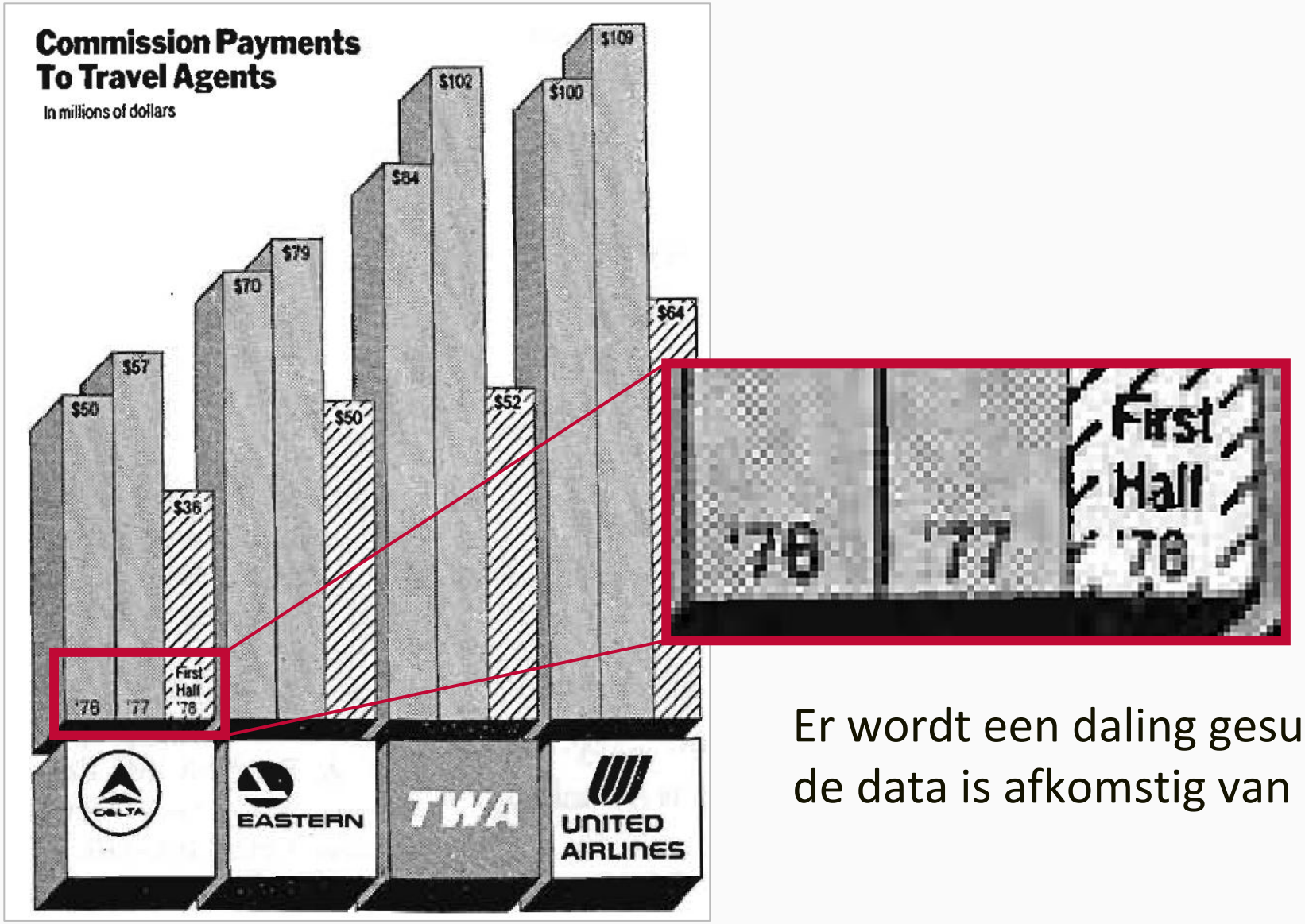
Grafische integriteit = vertel de waarheid met vis

Designesthetiek = maak vis duidelijk en precies

Hoorcollege 3 gaat uitgebreider in op grafische integriteit

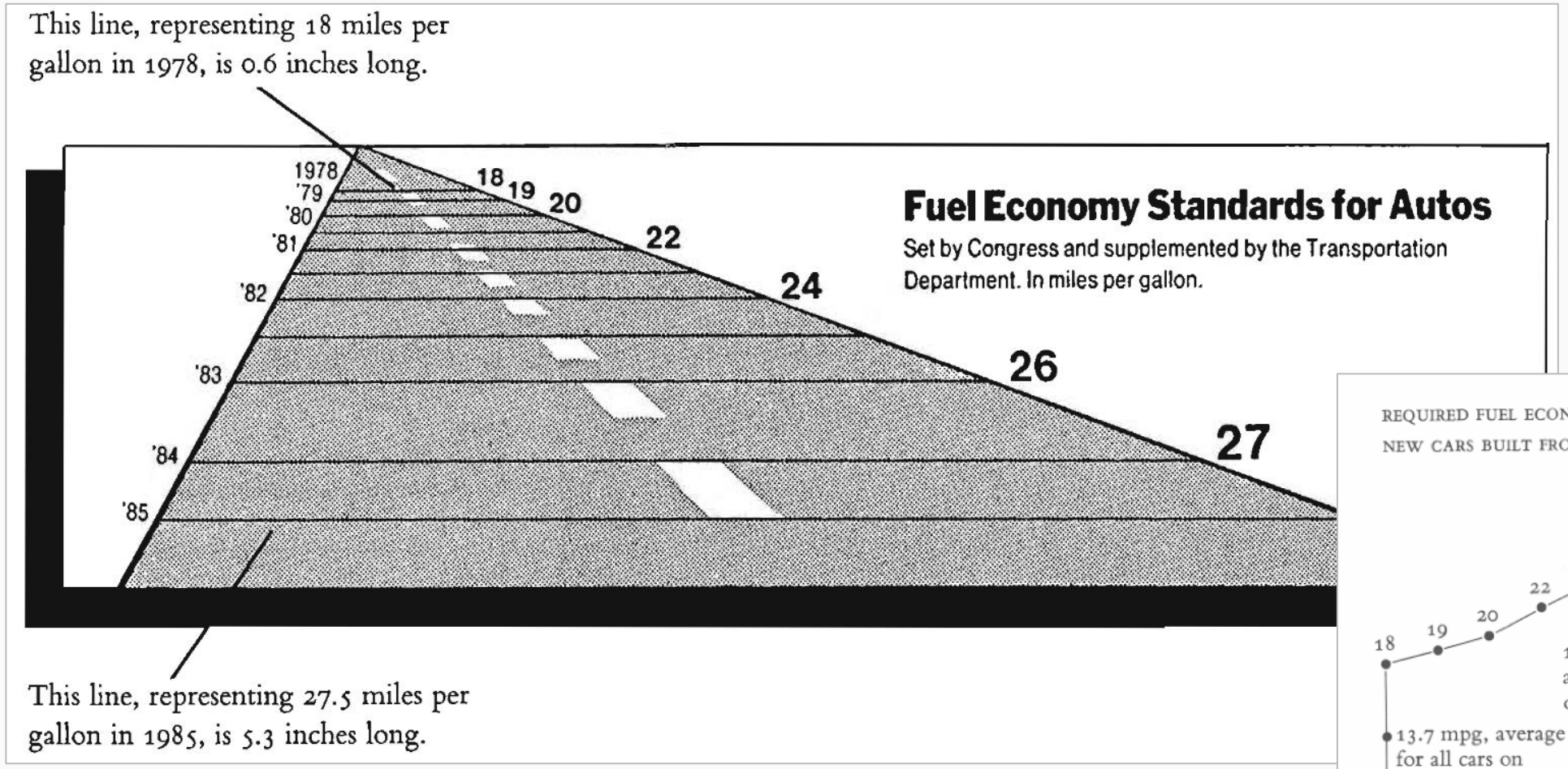


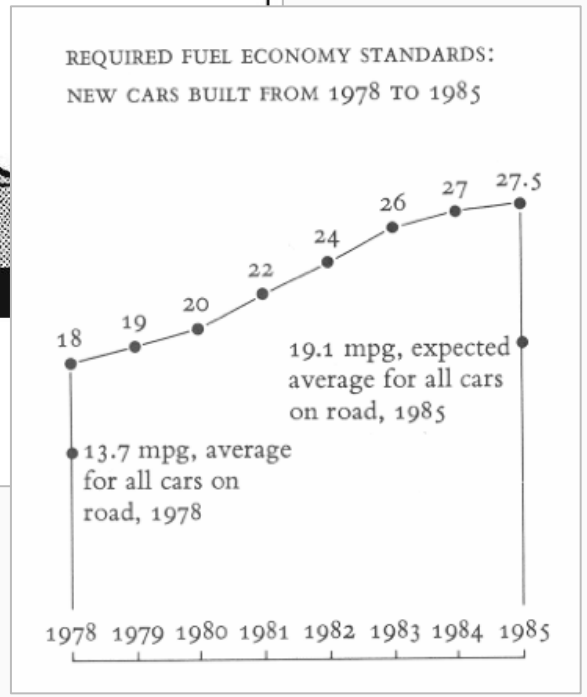
Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press.



Er wordt een daling gesuggereerd, maar de data is afkomstig van slechts 1/2 jaar

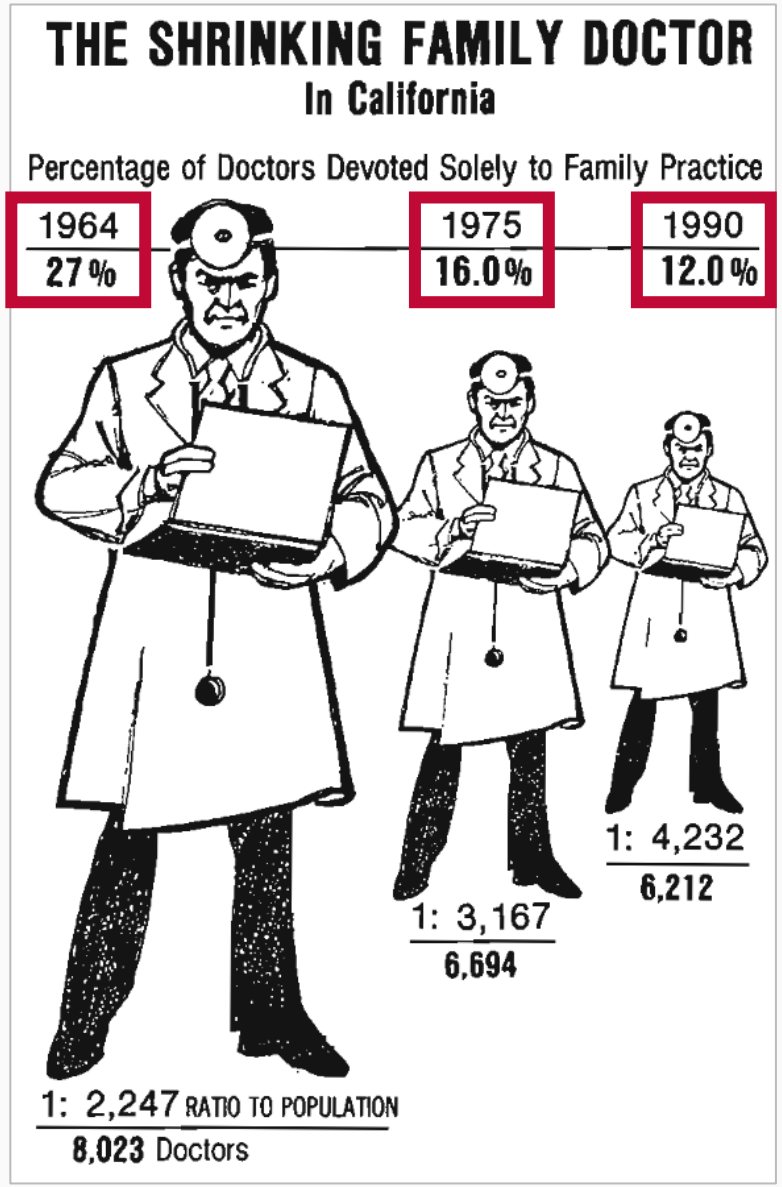
Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press. p. 54





Door het perspectief wordt de toename sterk overdreven

Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press. p. 57



Figuren zijn 2D, maar coderen 1D data

Perspectief overdrijft de afname

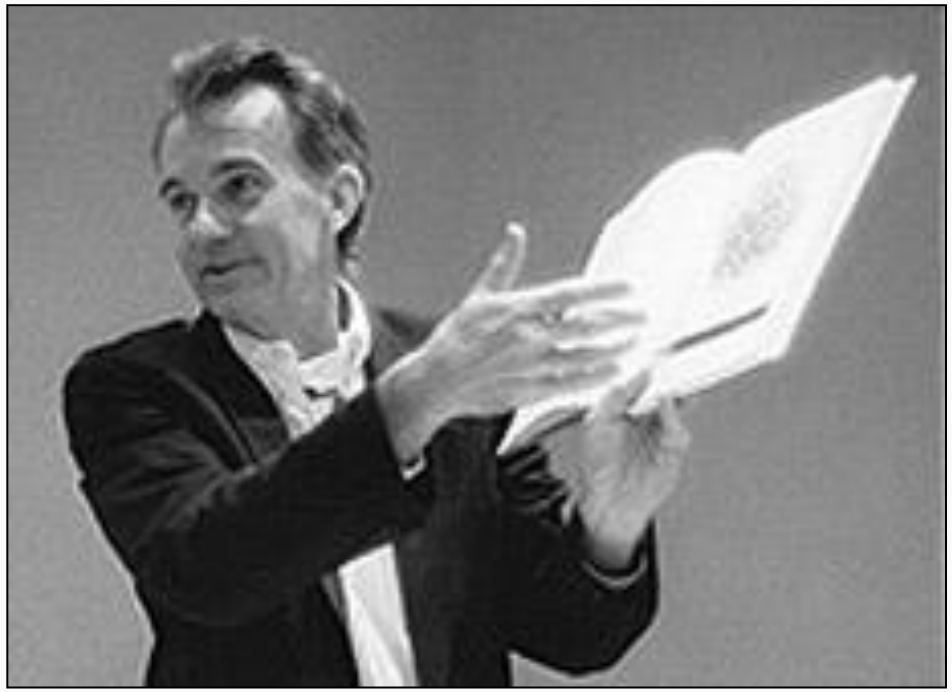
Foute horizontale spatiëring overdrijft de snelheid van de afname

Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press. p. 69

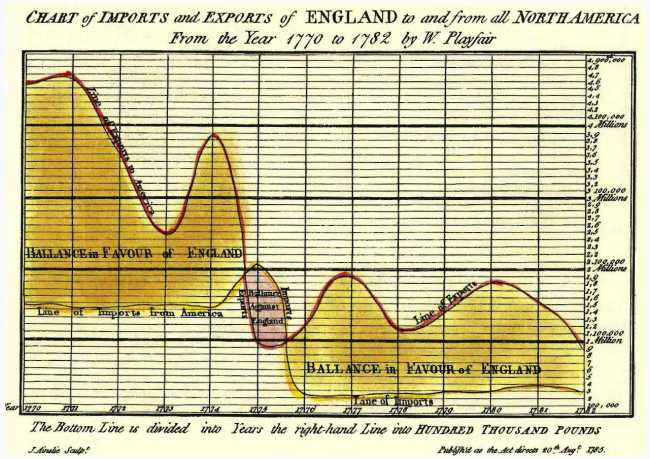
Visualisatieprincipes van Tufte

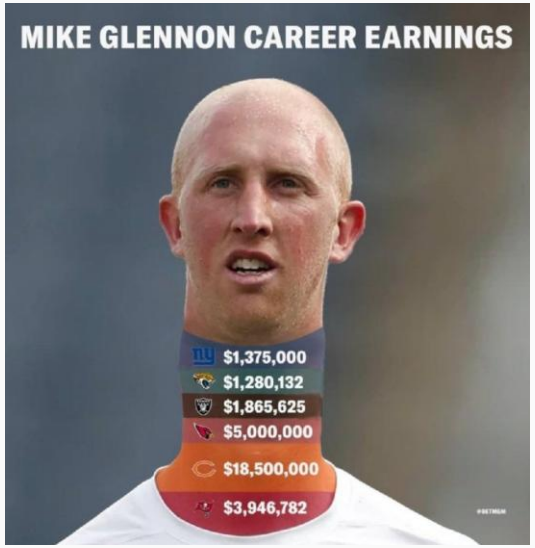
Grafische integriteit = vertel de waarheid met vis

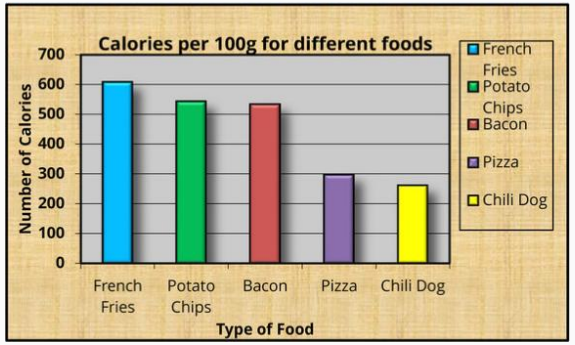
Designesthetiek = maak vis duidelijk en precies



Waarom zijn deze vis slecht?







Tuftes principes voor designesthetiek

1. Laat vooral de data zien
2. Maximaliseer redelijkerwijs de **data-inkt ratio**
3. Wis redelijkerwijs inkt die geen data codeert
4. Verwijder overbodige data-inkt
5. Herwerk iteratief

"Perfectie is niet bereikt wanneer er niets meer is om toe te voegen, maar wanneer er niets meer is om weg te nemen."

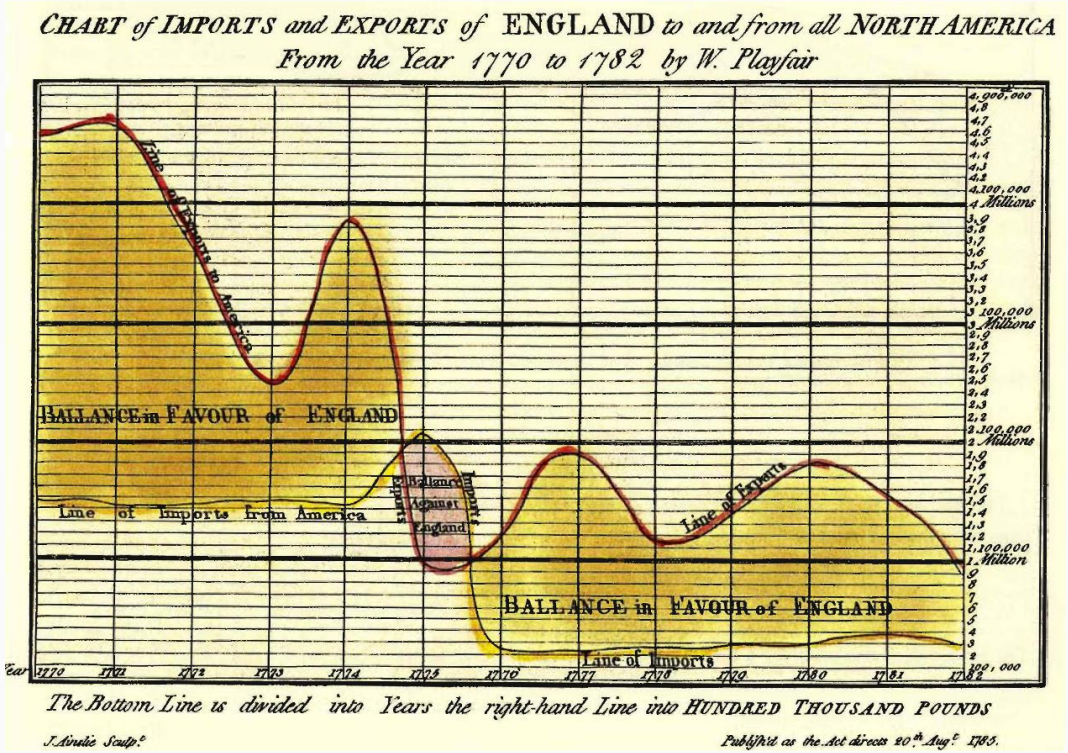
Antoine de Saint-Exupery (Franse schrijver)

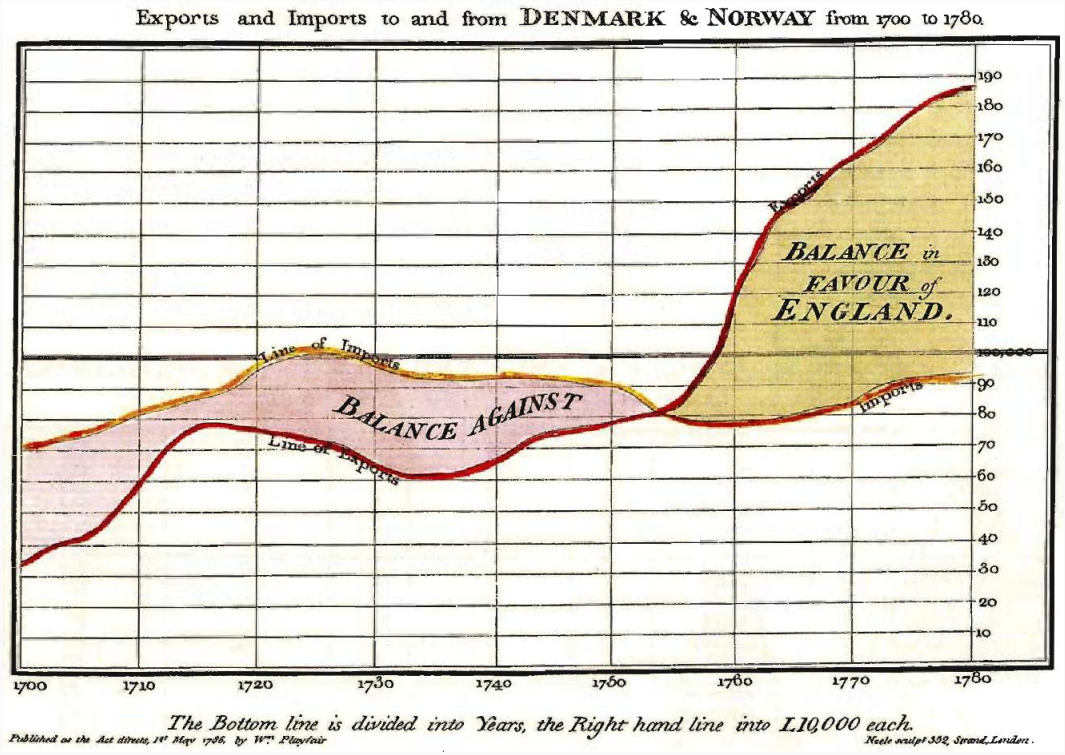
$$\text{Data-inkt ratio} = \frac{\text{data-inkt}}{\text{totale hoeveelheid inkt in de vis}}$$

$=$ proportie van de inkt in de vis gebruikt voor niet-redundante data

$= 1 -$ proportie van de vis die verwijderd kan worden zonder info te verliezen

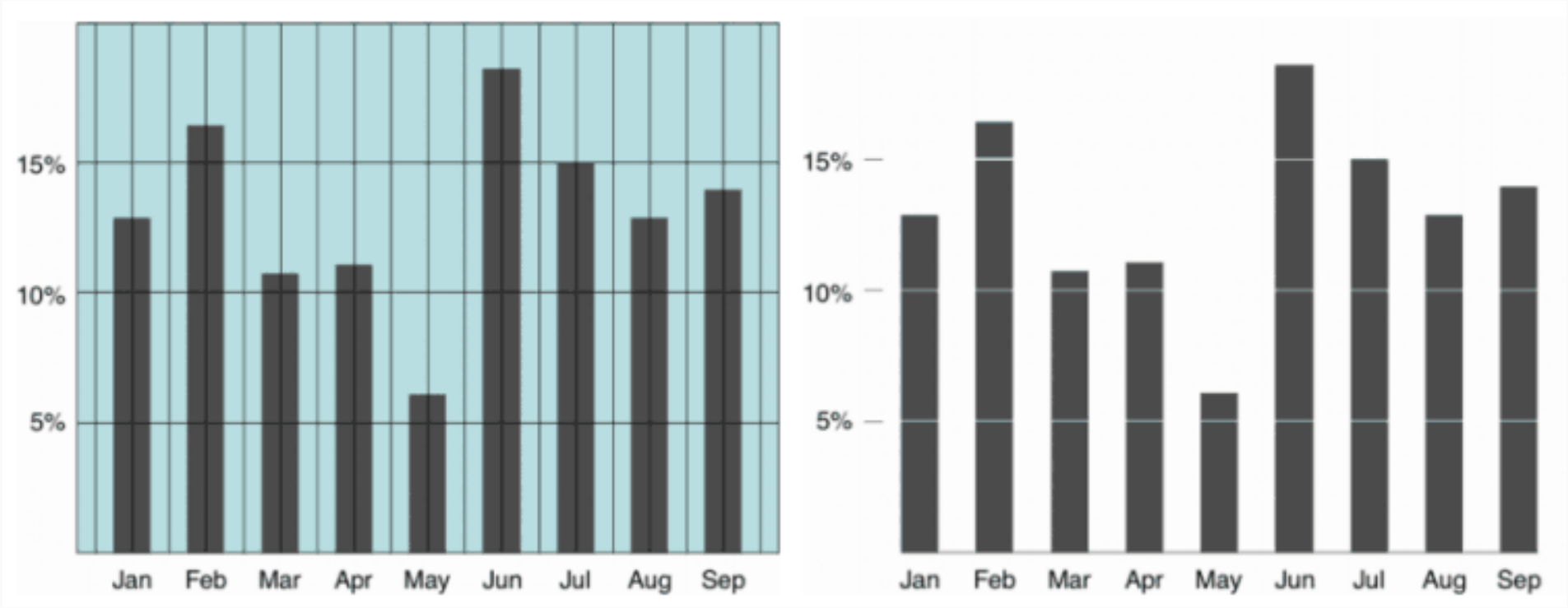
Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press. Chapter 4





Minder nadruk op roosterlijnen en minder gedetailleerde labels

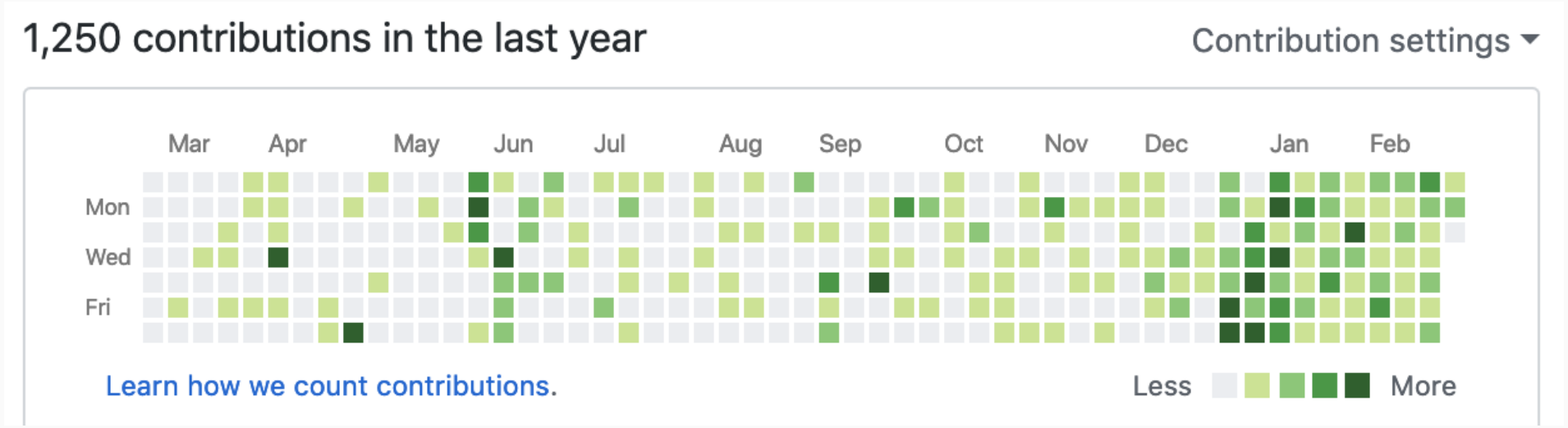
Tufte, E. R., & Graves-Morris, P. R. (1983). The visual display of quantitative information (Vol. 2, No. 9). Cheshire, CT: Graphics press. p. 91-92



Betere data-inkt ratio

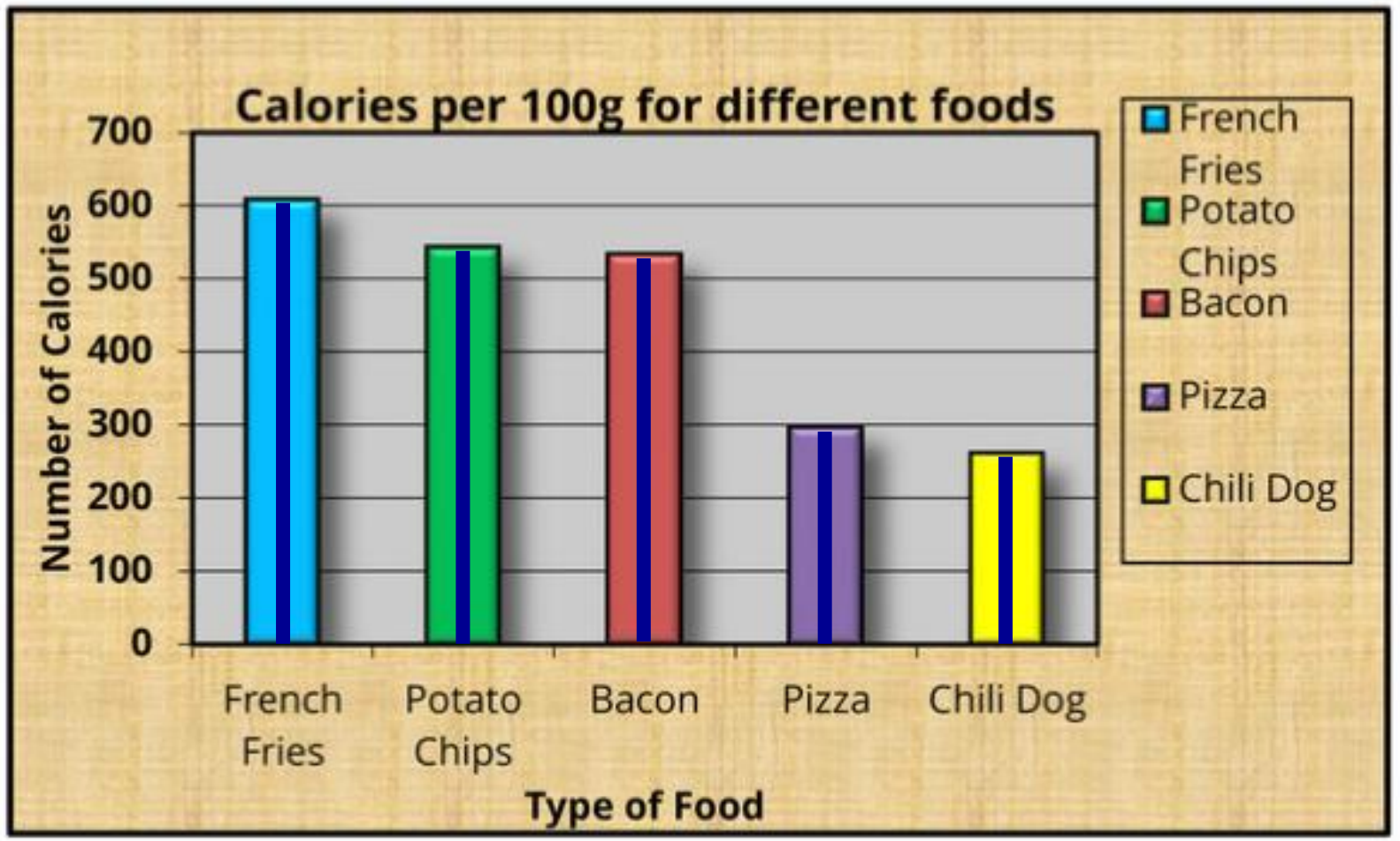
Lage data-inkt ratio: de enige relevante data zijn de percentages

Hoge data-inkt ratio: elke pixel is een datapunt



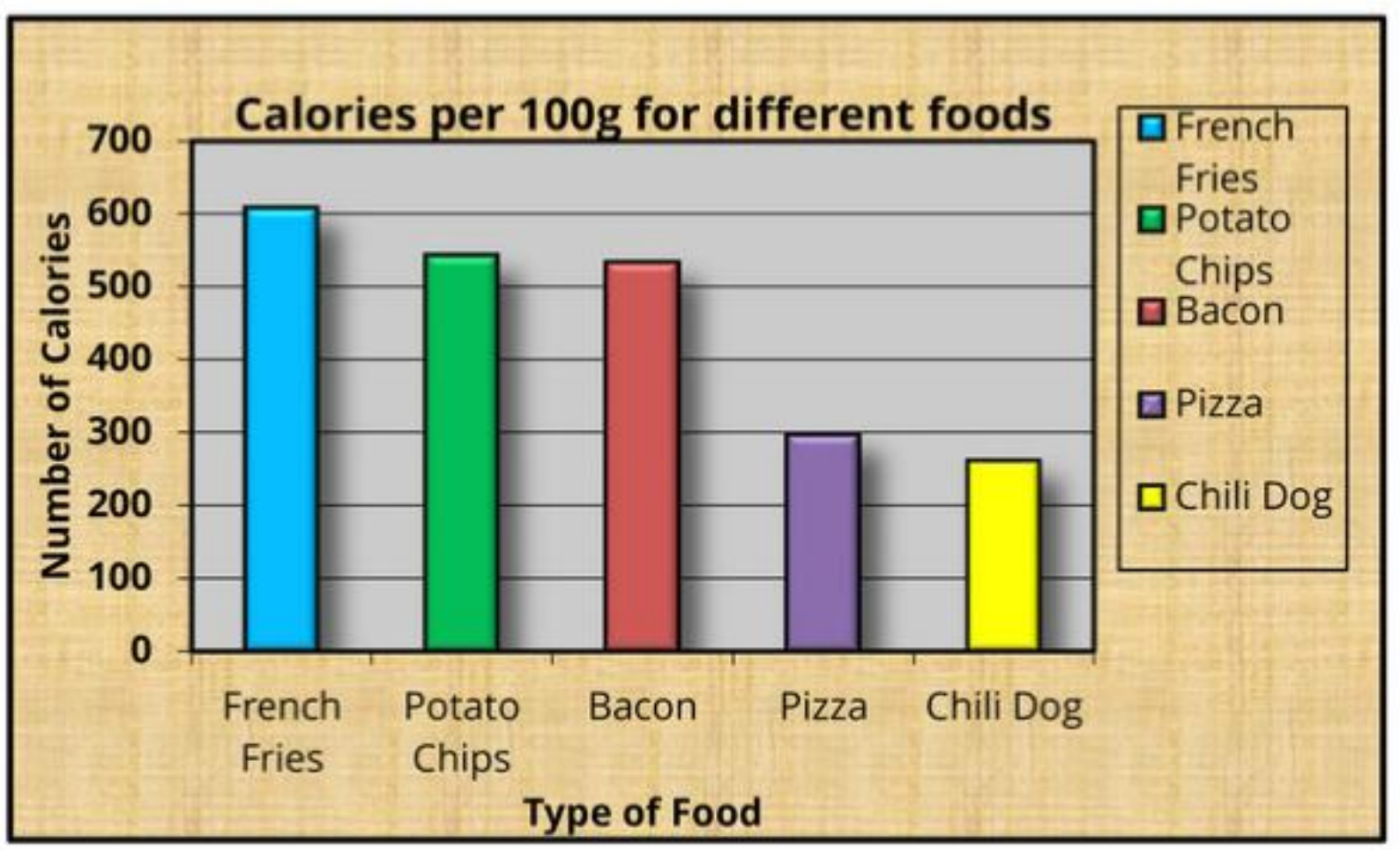
https://www.bencodezen.io/blog/github-contribution-graph-legend

Wat is de data-inkt ratio?



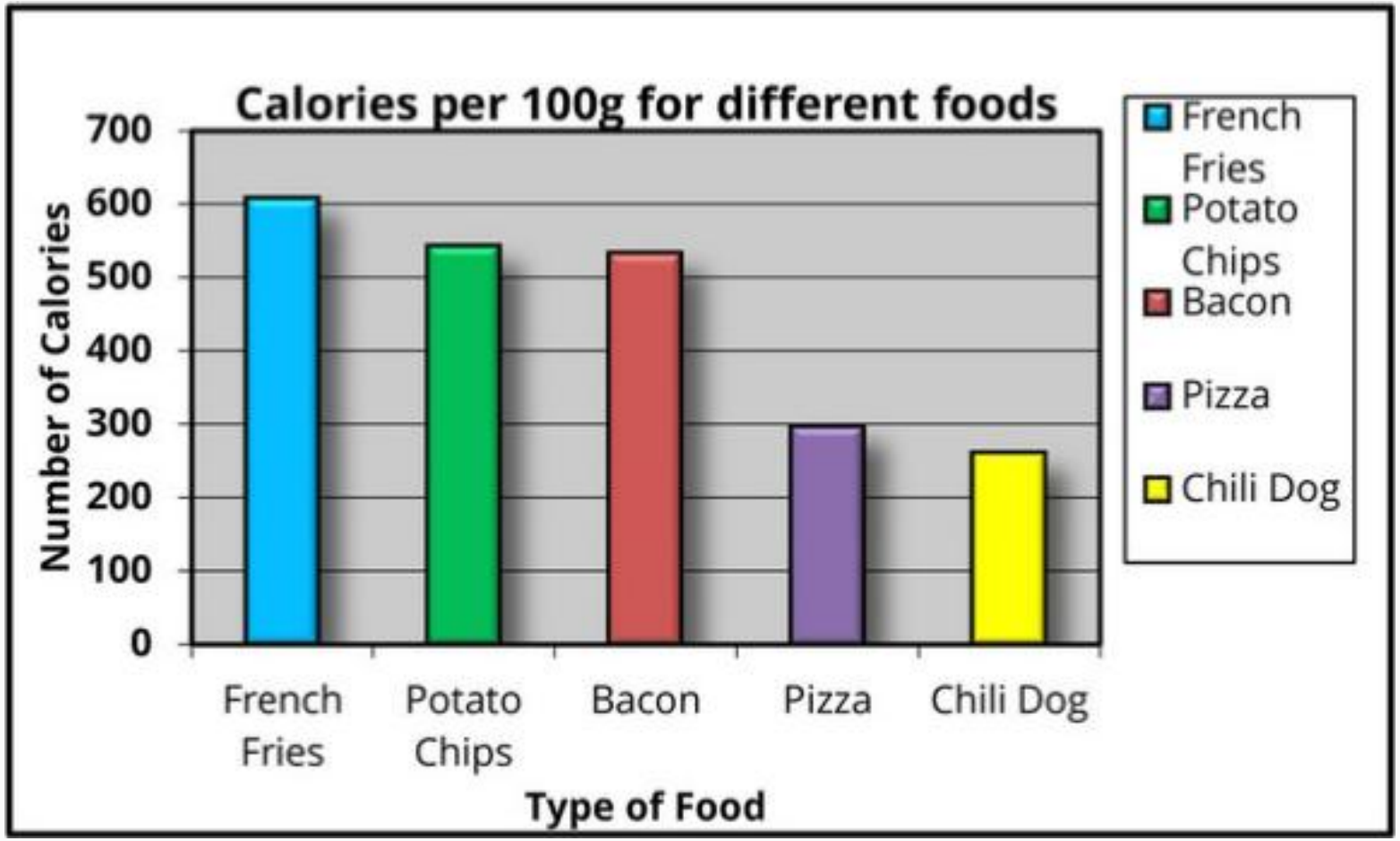
< 0.05

Oefening: verbeter de data-inkt ratio

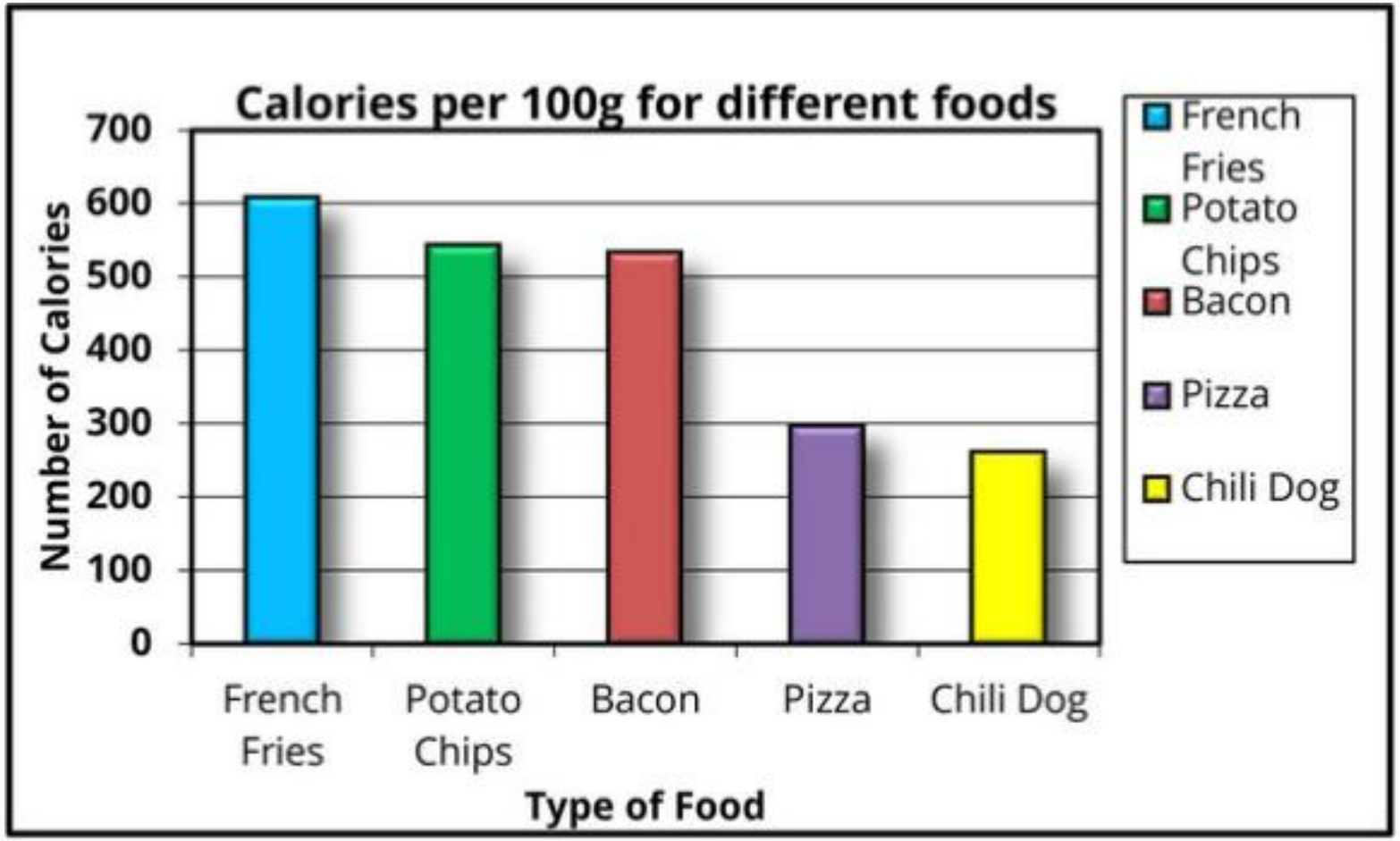


Er zijn een 10-tal verbeterpunten

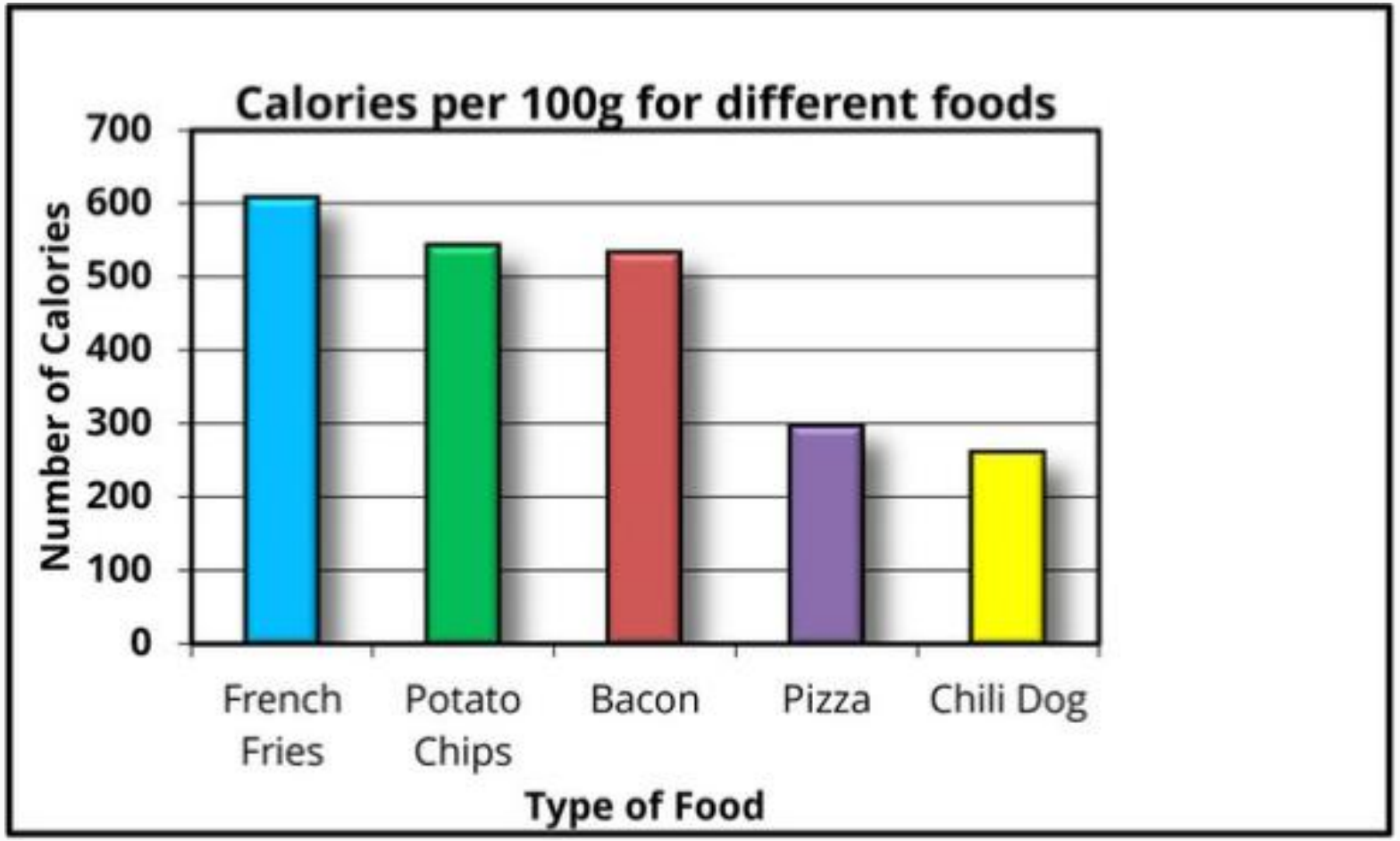
Verwijder achtergrond



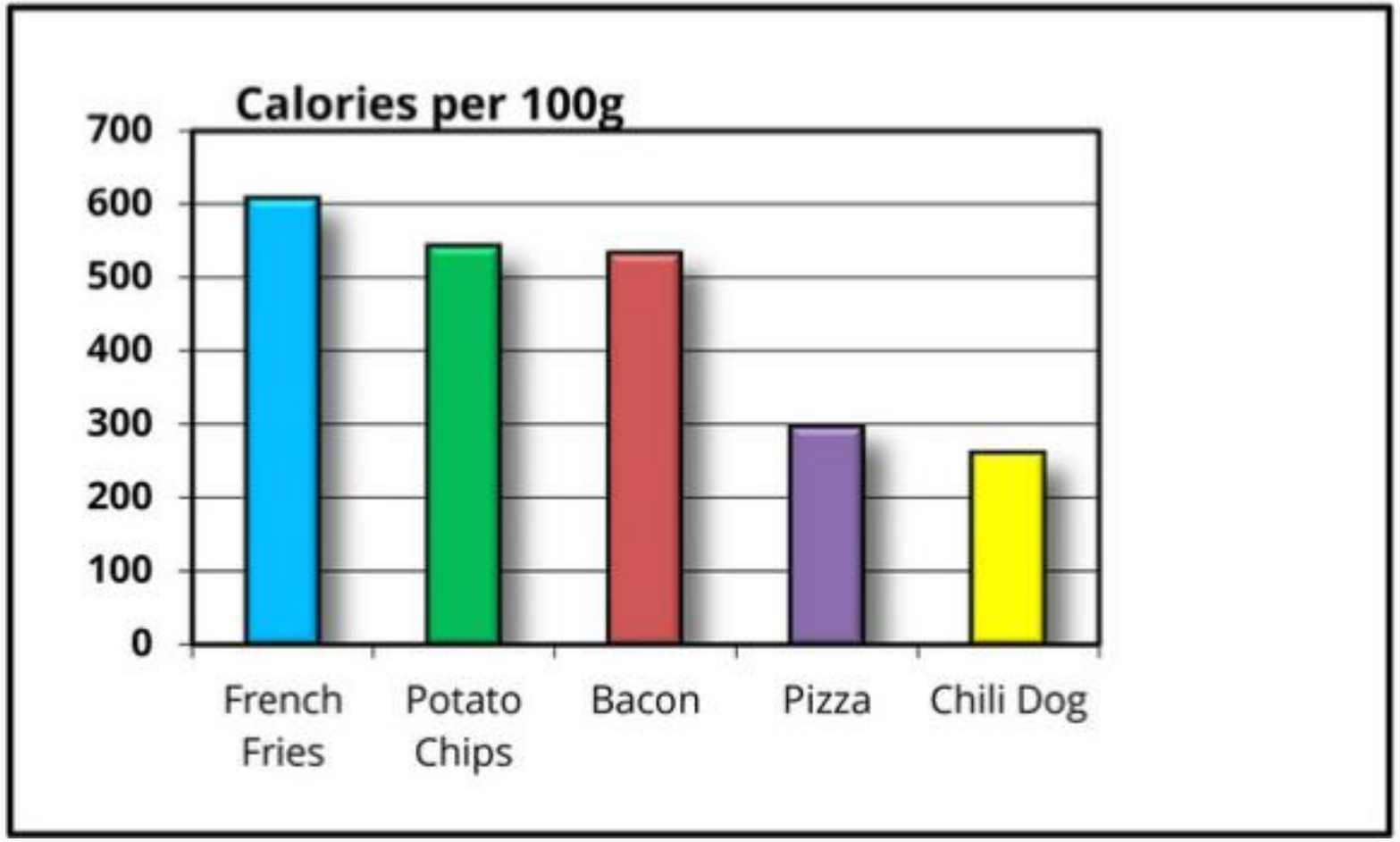
Verwijder achtergrondkleur



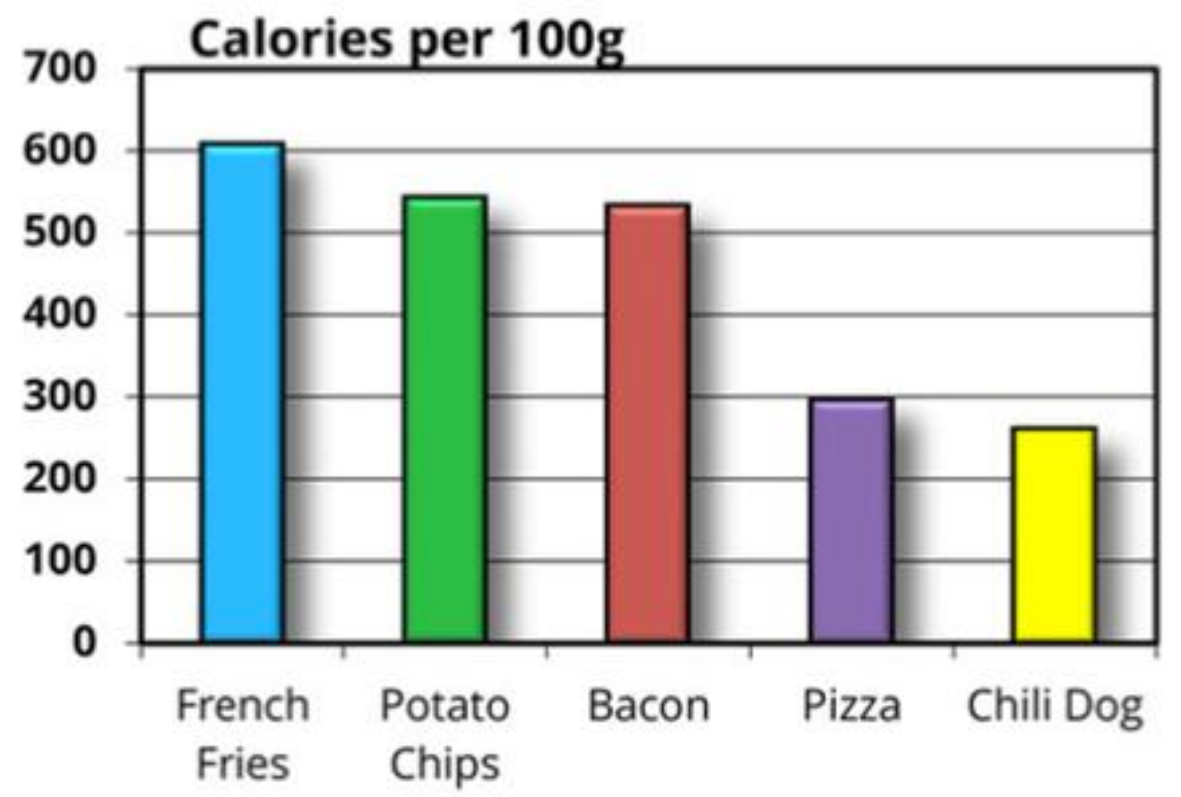
Verwijder de overbodige legende



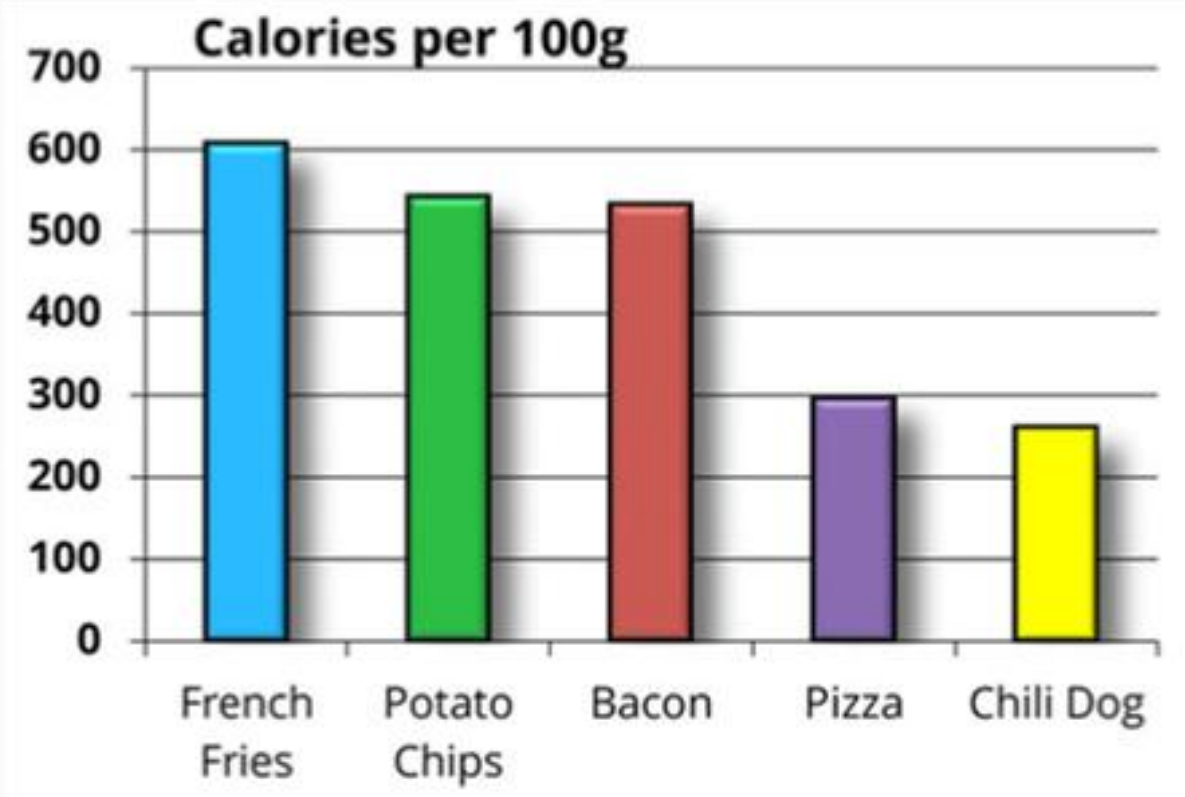
Verwijder overbodige labels



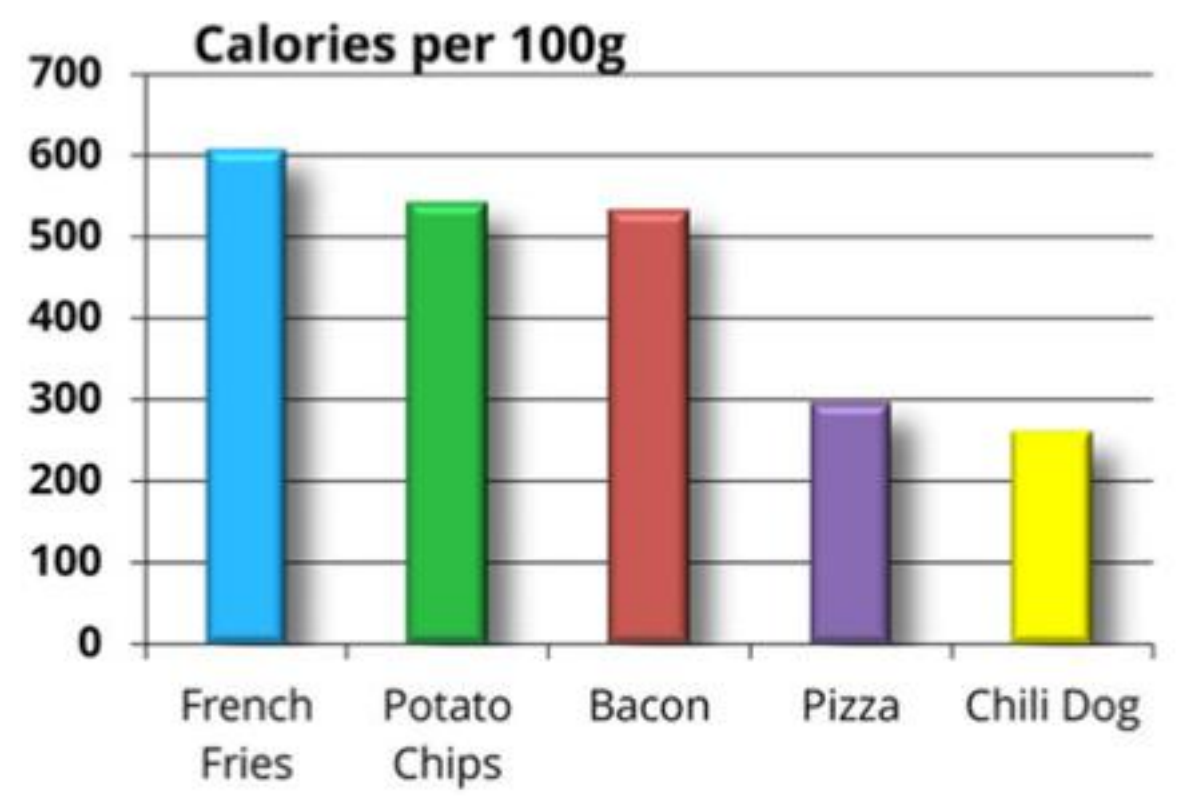
Verwijder de rand van de vis



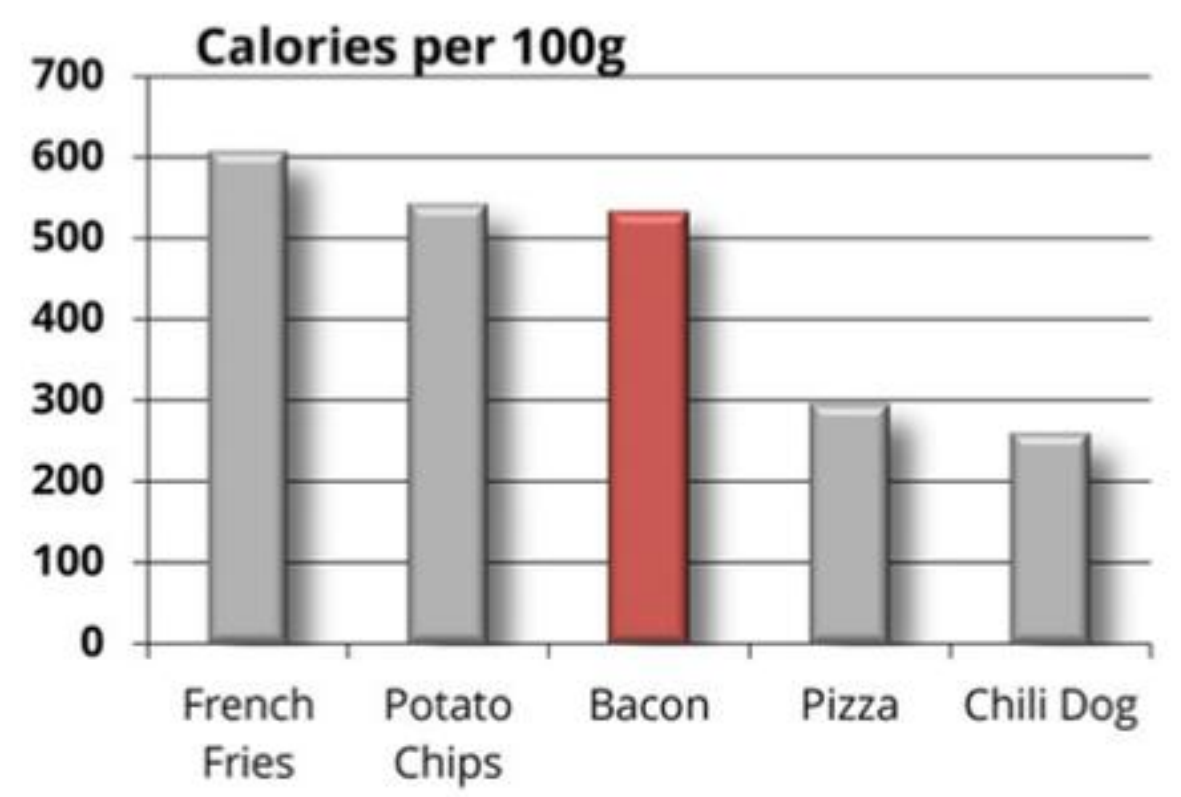
Verwijder de rand van het assenstelsel



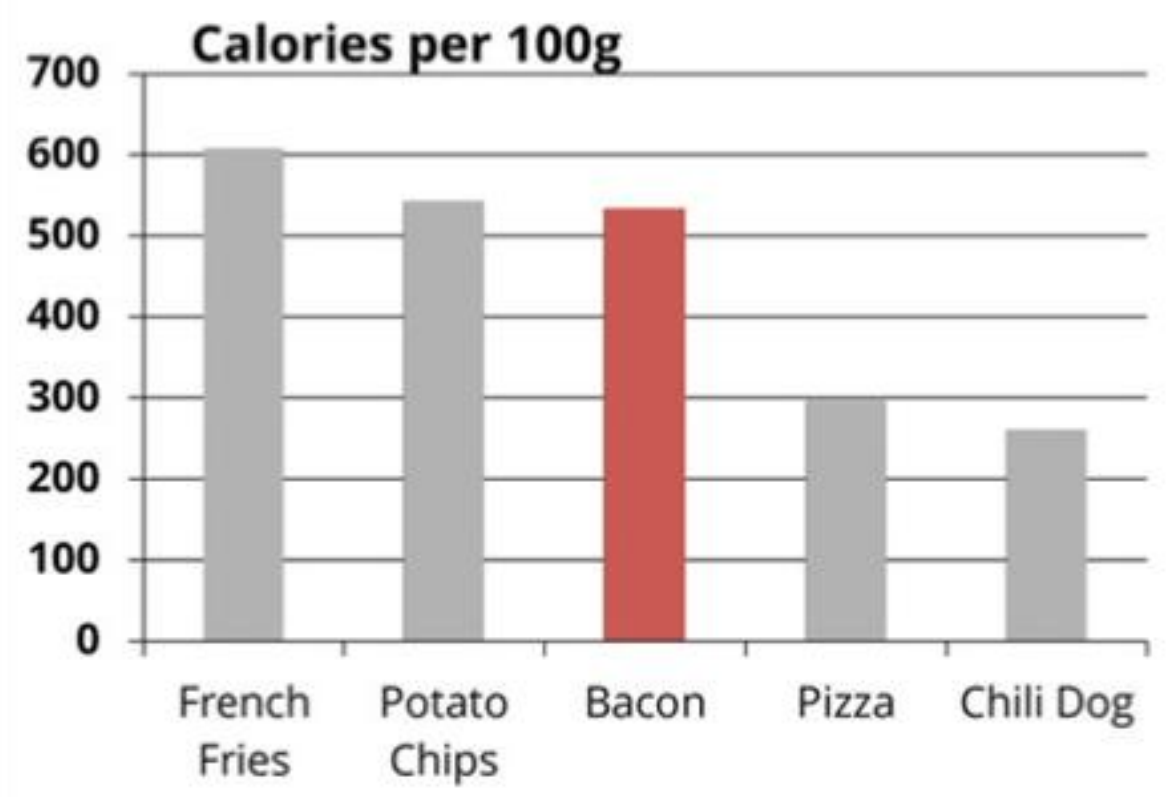
Verwijder randen van de bars



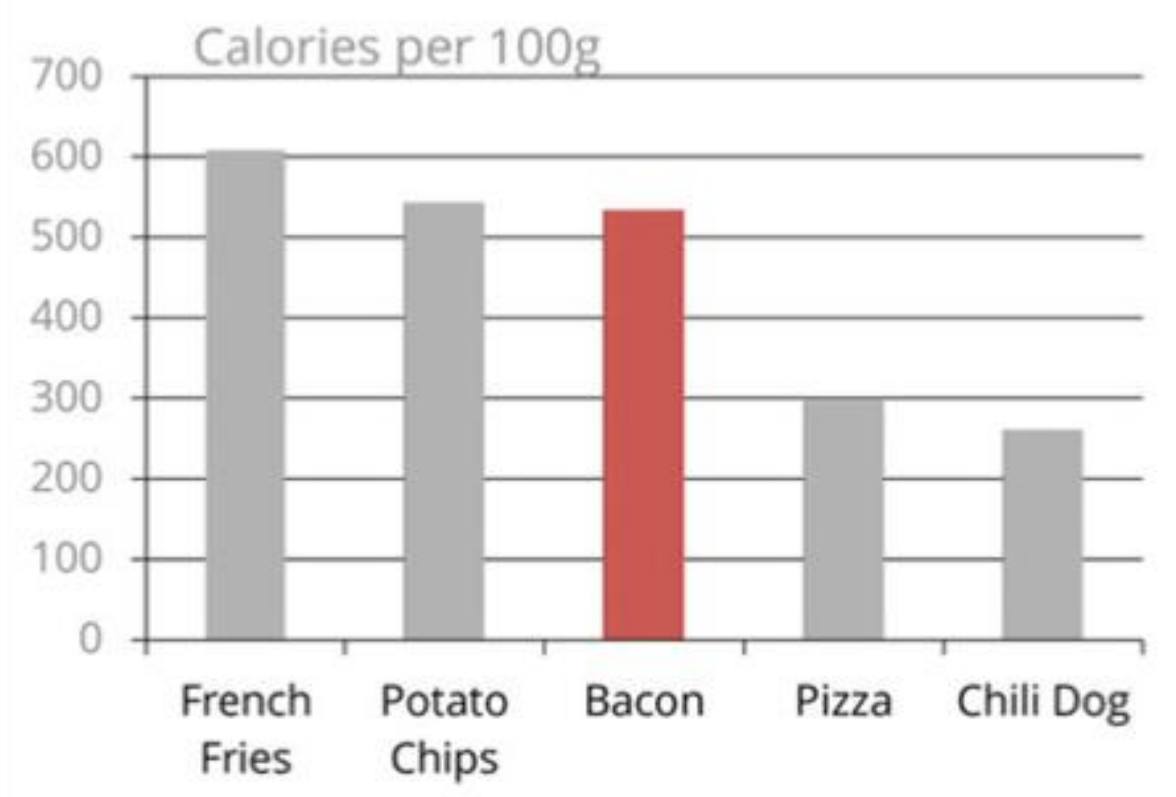
Reduceer betekenisloze kleuren



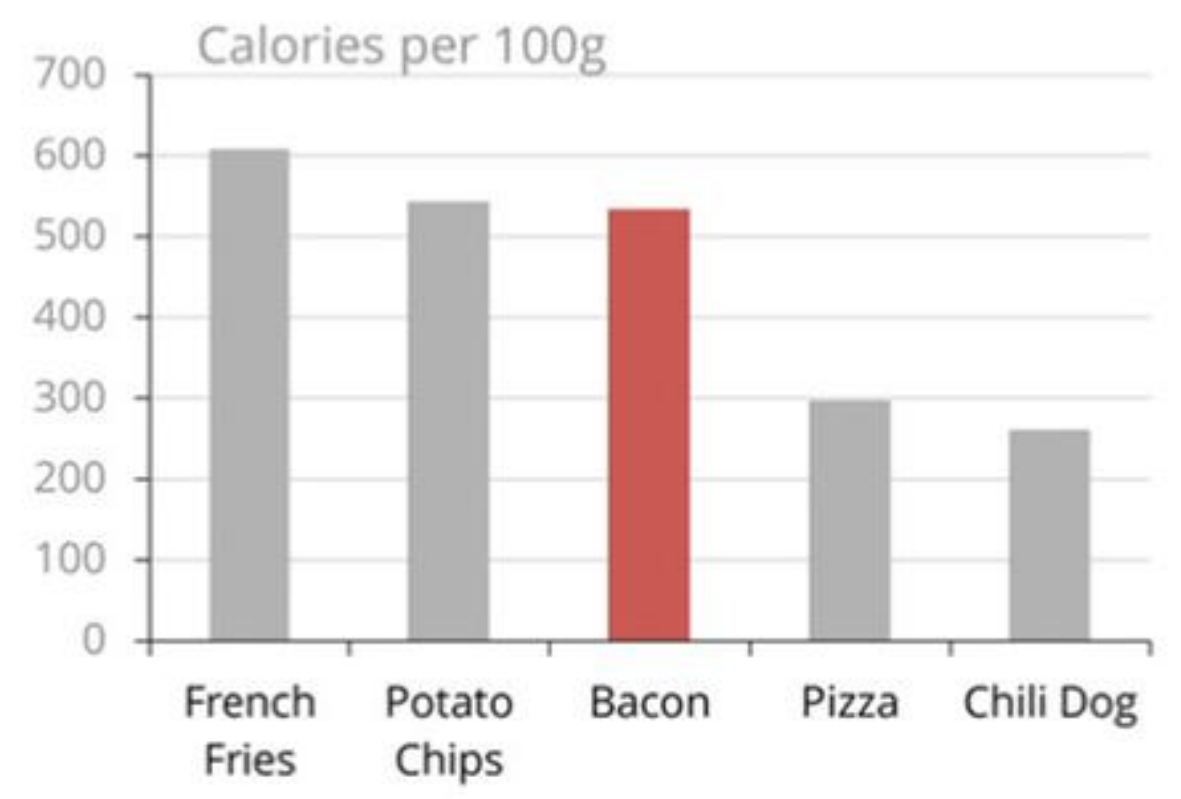
Verwijder schaduwen en highlights van bars



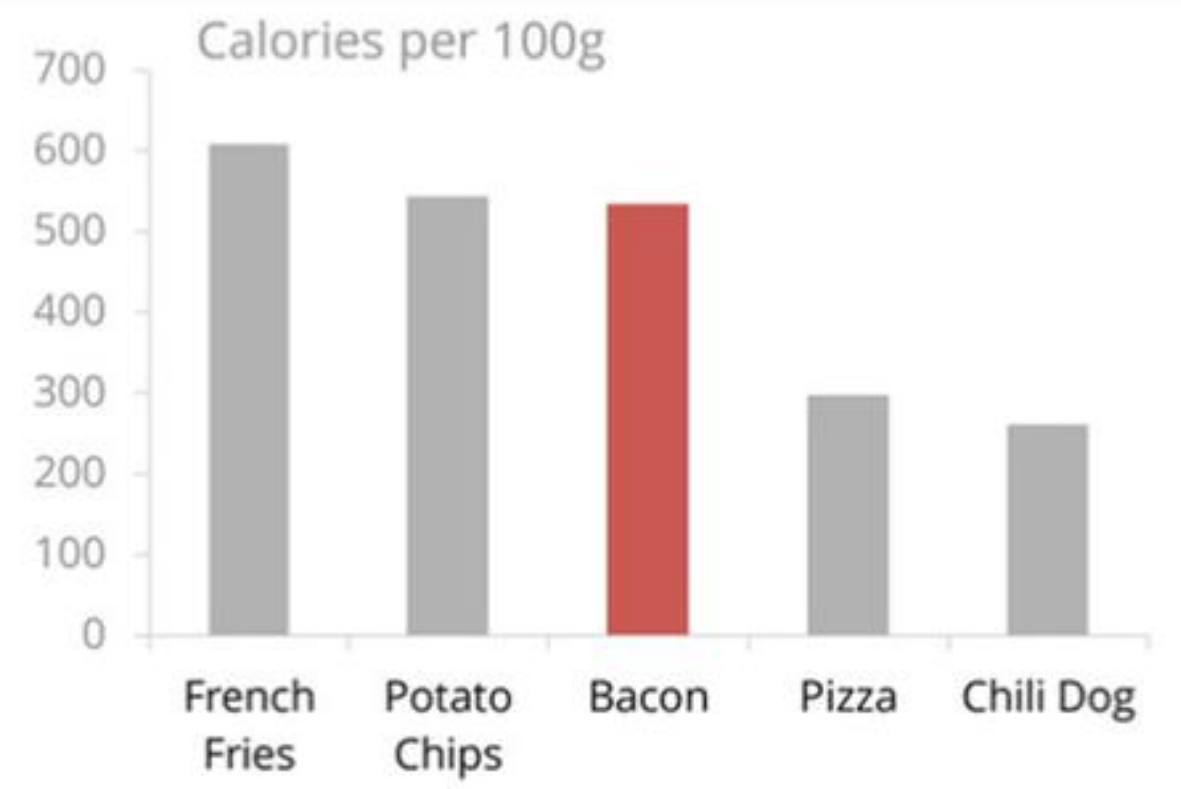
Maak details minder prominent



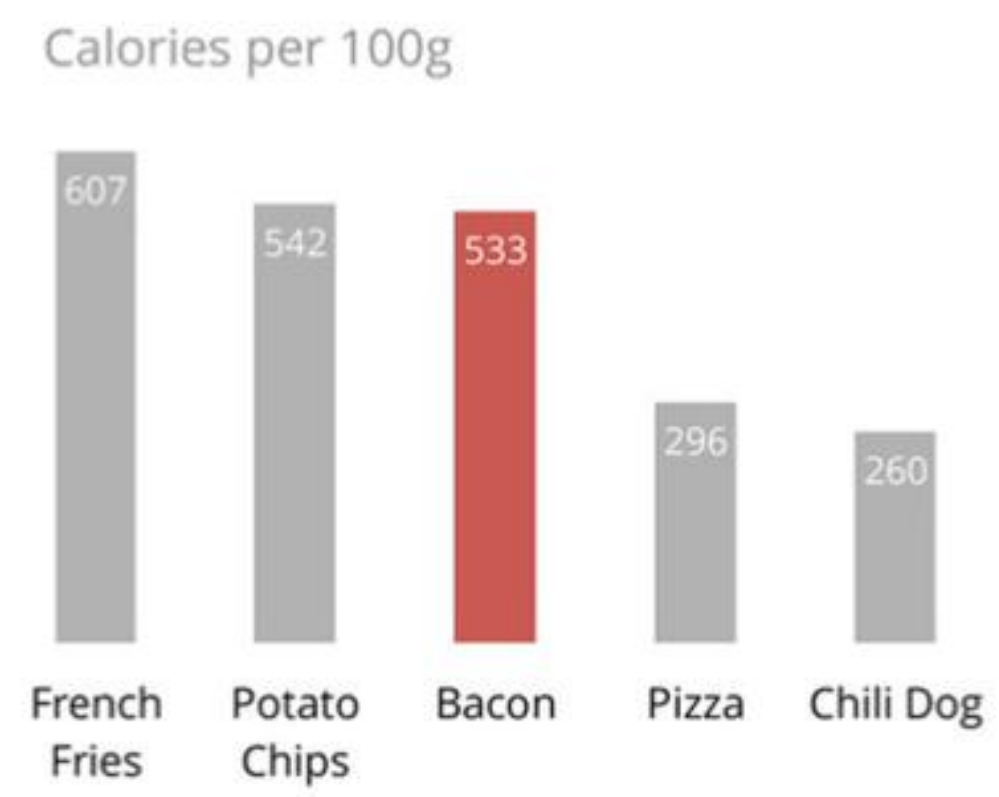
Maak rasterlijnen minder prominent



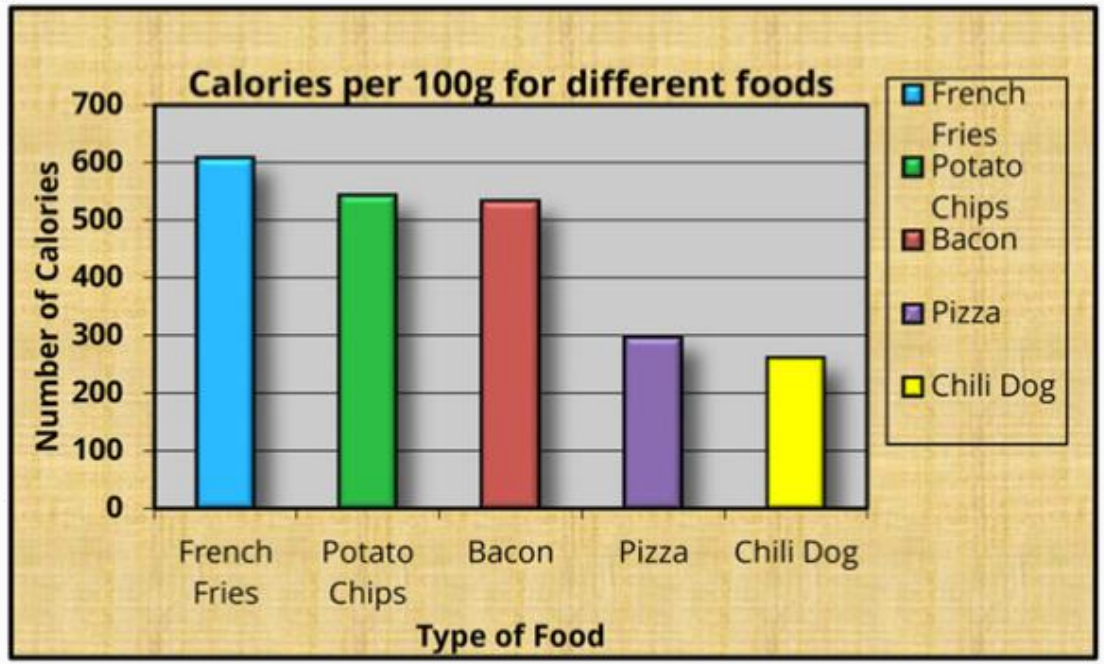
Beter: verwijder rasterlijnen



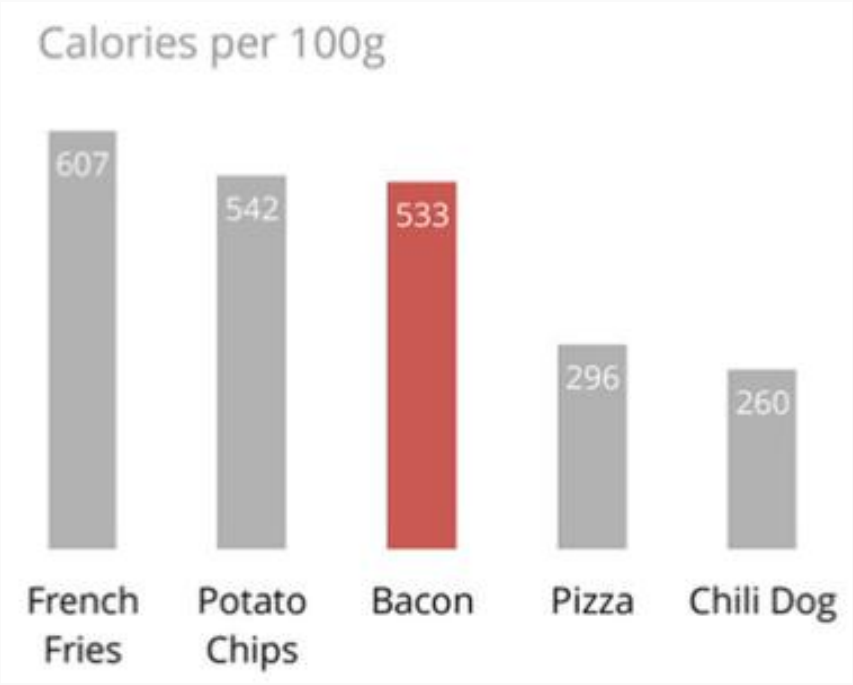
Beter: labels de bars direct



Conclusie: less is more



Voor

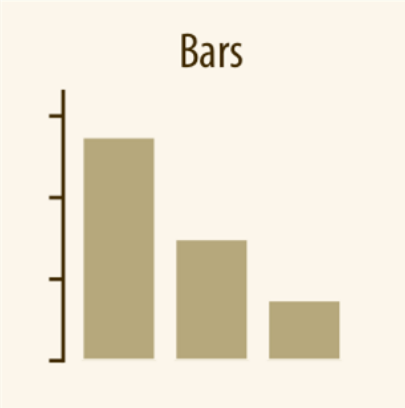


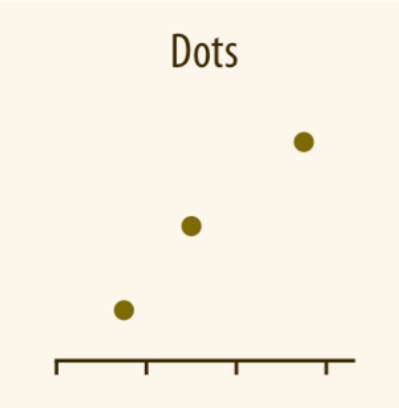
Na

Algemene principes en visualisatietypes

1. Voorbeelden en geschiedenis
2. Algemene vis-principes
3. Klassieke vis-types
4. Tools voor vis

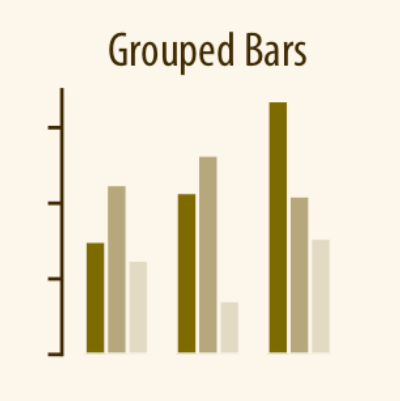
Aantallen

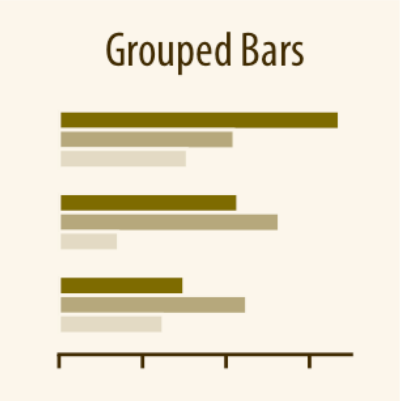


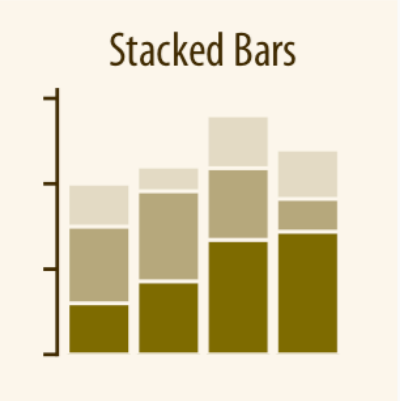


1 categorie

meerdere categorieën

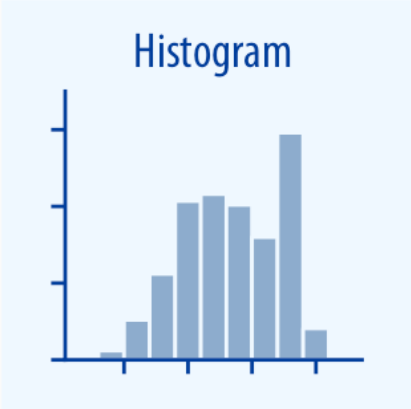


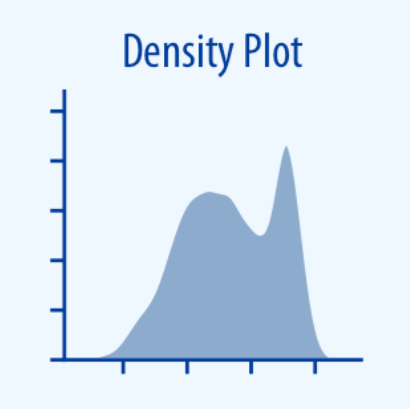


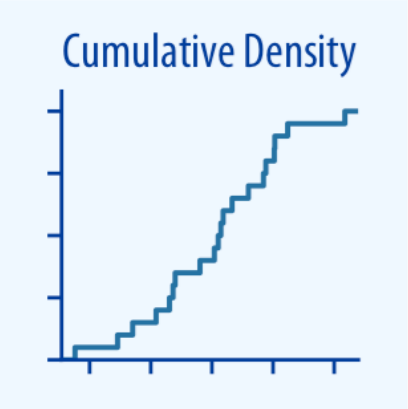


https://clauswilke.com/dataviz/directory-of-visualizations.html

Distributies



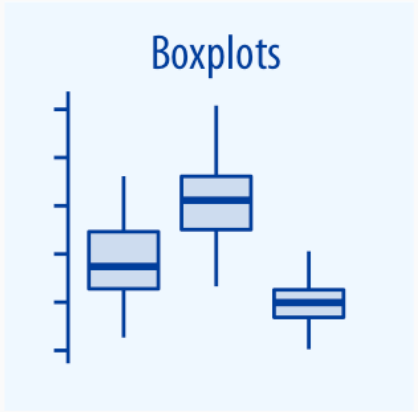


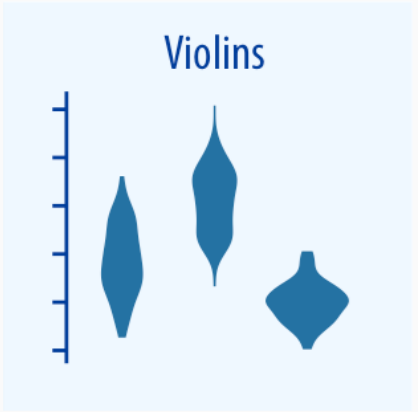


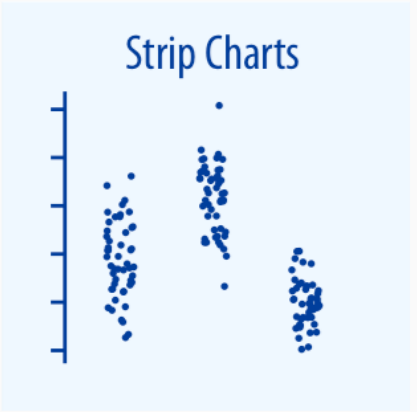
1 distributie

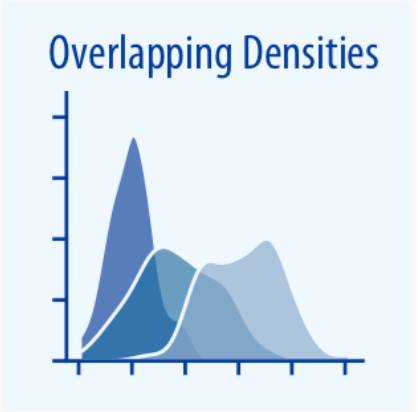
https://clauswilke.com/dataviz/directory-of-visualizations.html

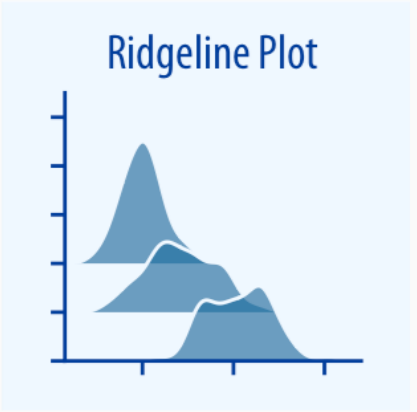
Distributies







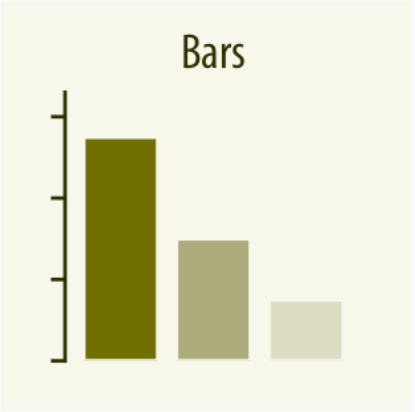


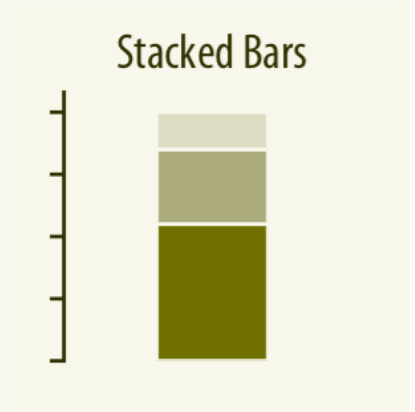


meerdere distributies

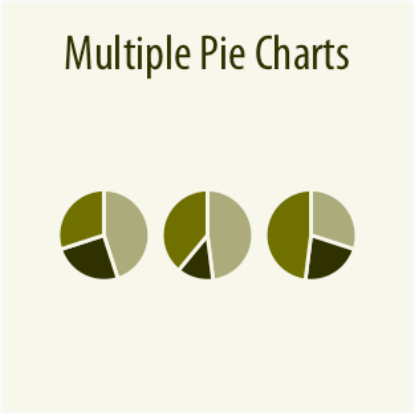
https://clauswilke.com/dataviz/directory-of-visualizations.html

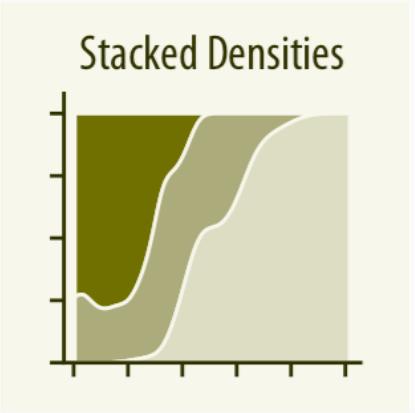
Proporties





1 conditie





meerdere condities

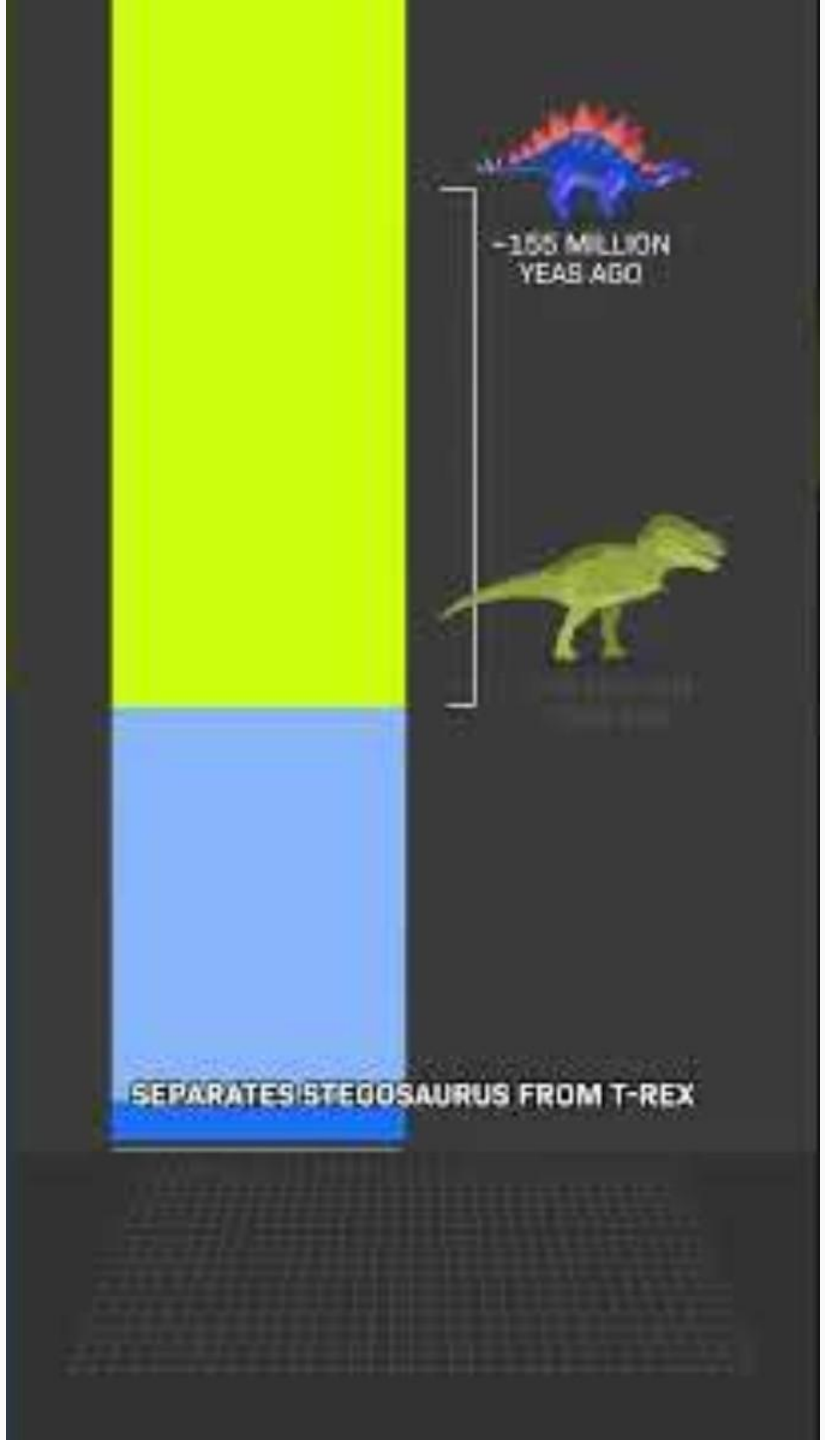
https://clauswilke.com/dataviz/directory-of-visualizations.html

Proporties

meerdere
grouping variables

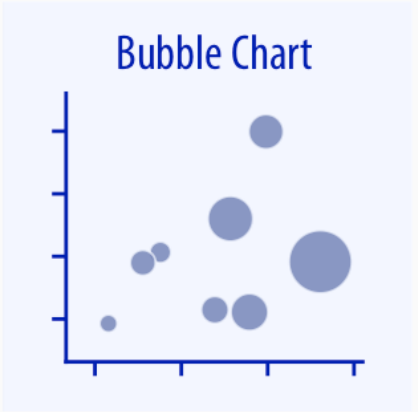
https://clauswilke.com/dataviz/directory-of-visualizations.html

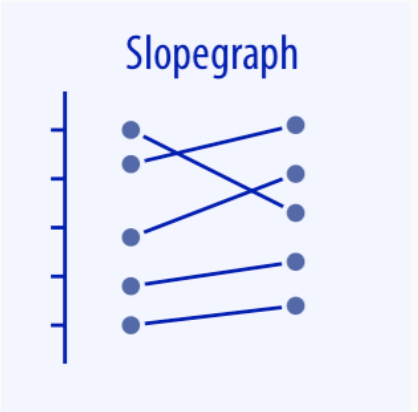
Voorbeeld van een inzichtelijke stacked bar 🦖



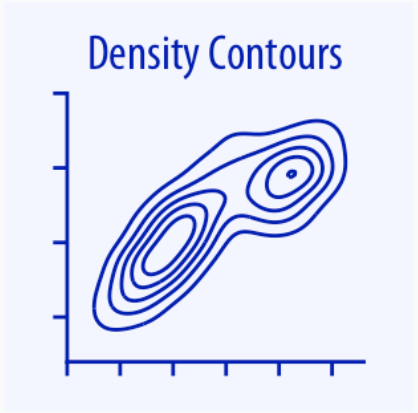
https://youtube.com/shorts/E6Zrm5Cj8M8?si=wmMcsf3vflMgJj-0

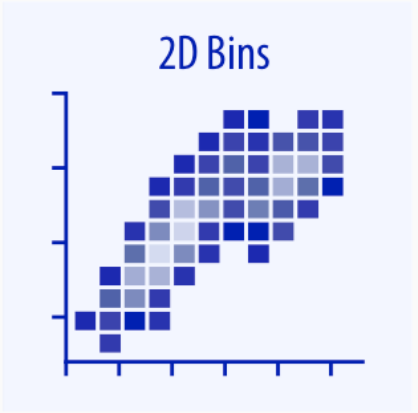
Verbanden tussen 2 variabelen

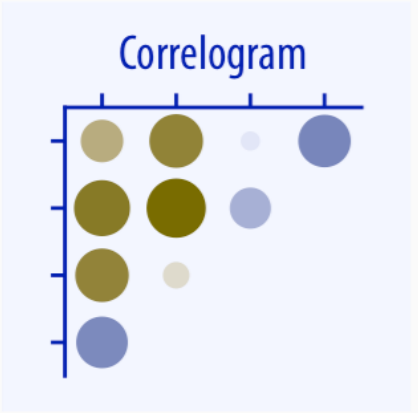




beperkt aantal datapunten



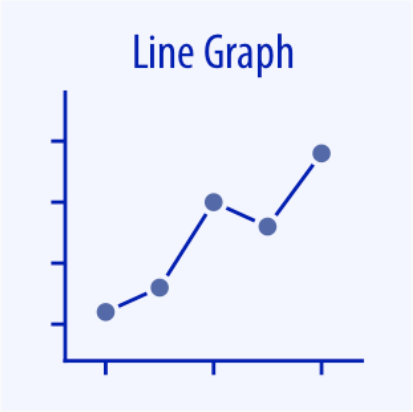


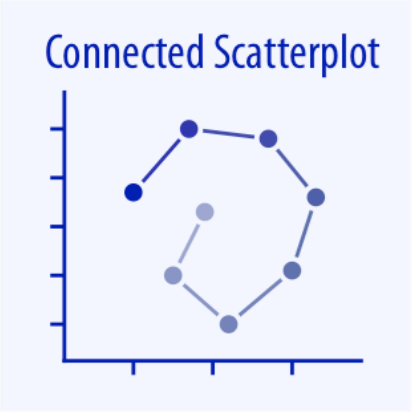


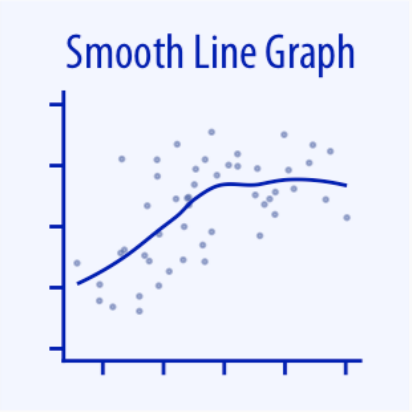
veel datapunten

https://clauswilke.com/dataviz/directory-of-visualizations.html

Verbanden tussen 2 variabelen



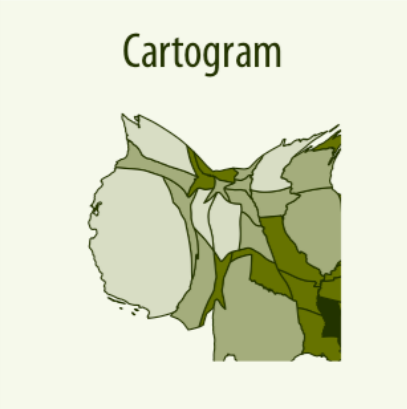


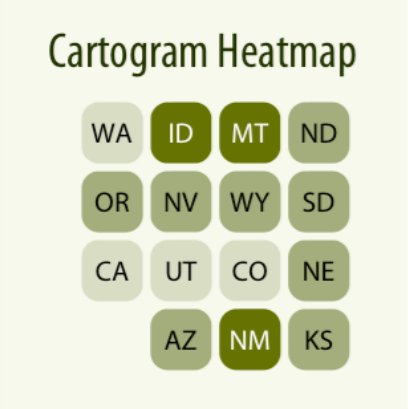


evoluties en trends

https://clauswilke.com/dataviz/directory-of-visualizations.html

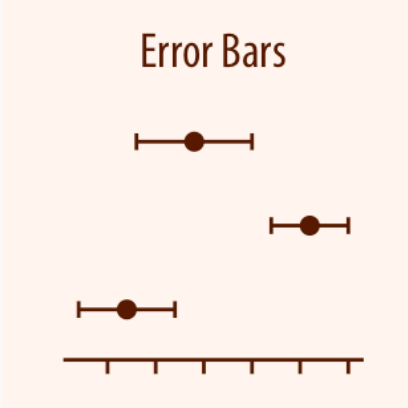
Geospatiale data

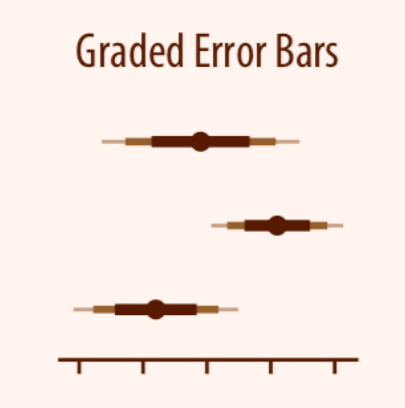




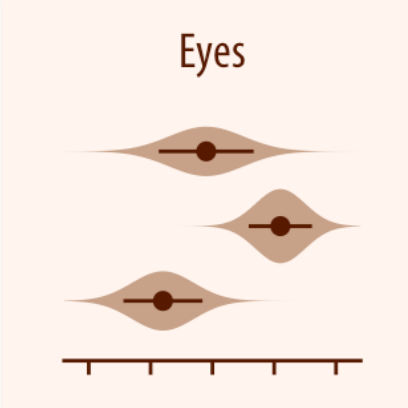
https://clauswilke.com/dataviz/directory-of-visualizations.html

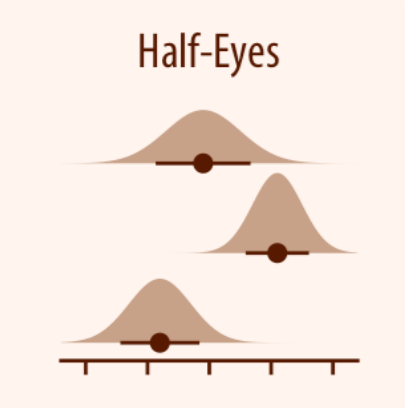
Onzekerheid

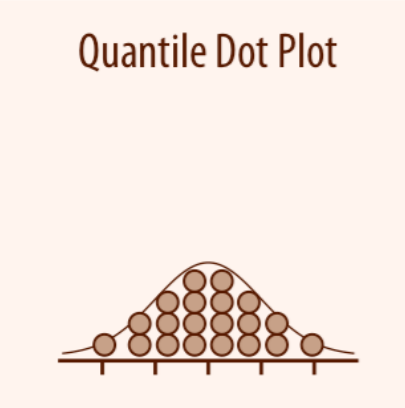




onzekere waardes



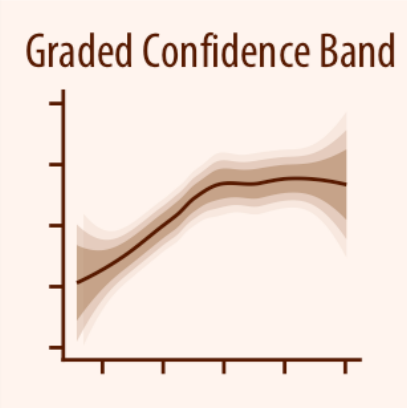


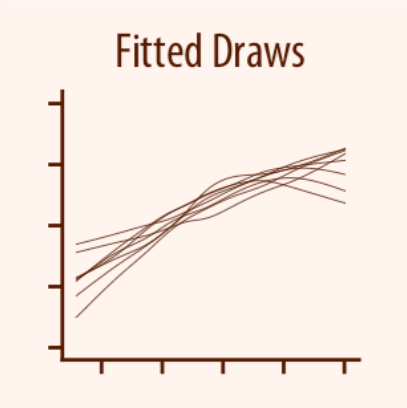


onzekere distributies

https://clauswilke.com/dataviz/directory-of-visualizations.html

Onzekerheid





onzekere evoluties

https://clauswilke.com/dataviz/directory-of-visualizations.html

Algemene principes en visualisatietypes

1. Voorbeelden en geschiedenis
2. Algemene vis-principes
3. Klassieke vis-types
4. Tools voor vis

Enkele bekende vis-tools

meer interactie,
meer flexibiliteit,
stijlere leercurve

Enkele bekende vis-tools

Pen en papier

Complete grafische vrijheid
Geen interactieve data mogelijk
Ideaal om snel ideeën over te brengen

Enkele bekende vis-tools

matplotlib

PowerPoint, Excel, ggplot2, matplotlib ...

Standaardvis met aanpassingsopties

Data aanpassen is mogelijk

Snel statische visualisaties bouwen

Enkele bekende vis-tools

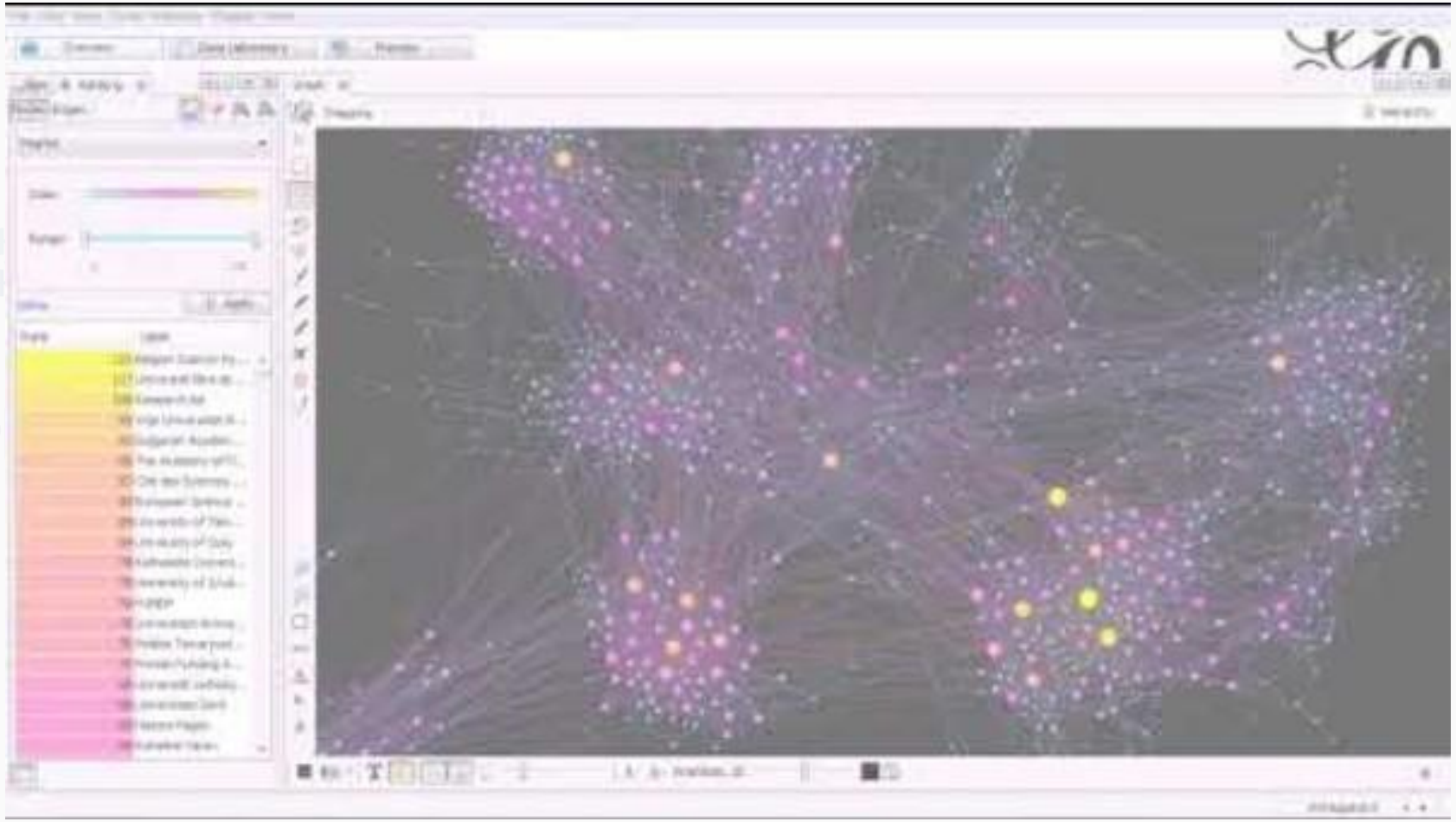
plotly, Gephi ...

Geavanceerdere interacties met data zijn mogelijk

Kan op specifieke datatypes toegespitst zijn

Enkele bekende vis-tools

Gephi



https://gephi.org/features

Enkele bekende vis-tools

Power BI

Power BI, tableau, Spotfire ...

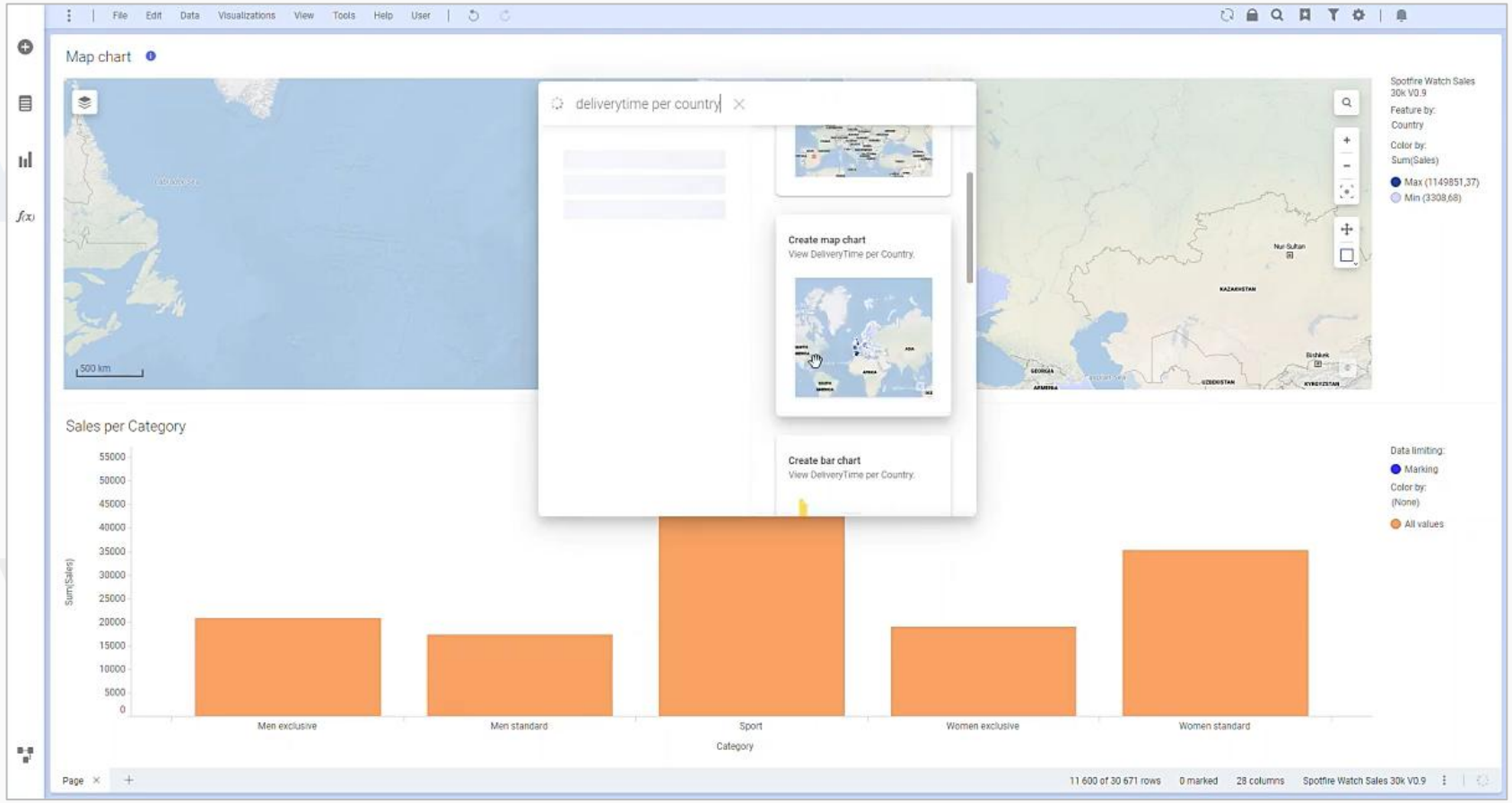
Geavanceerdere interacties met data zijn mogelijk

Ondersteunt veel datatypes

Typisch met *drag and drop* dashboards bouwen

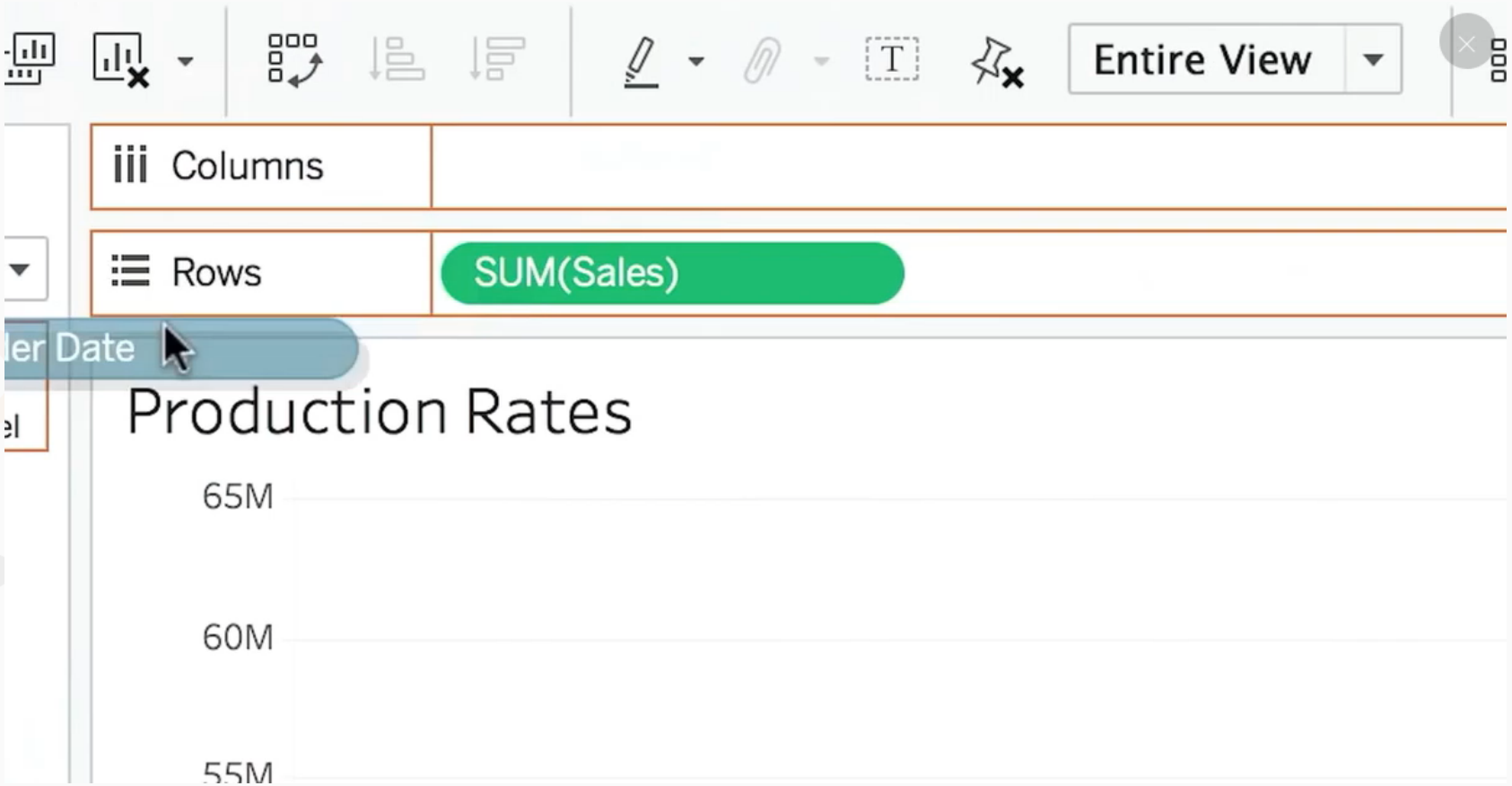
Enkele bekende vis-tools

Spotfire™



https://www.spotfire.com/overview

Enkele bekende vis-tools



https://www.tableau.com

Enkele bekende vis-tools

Observable

D3, Vega (Lite), Observable …

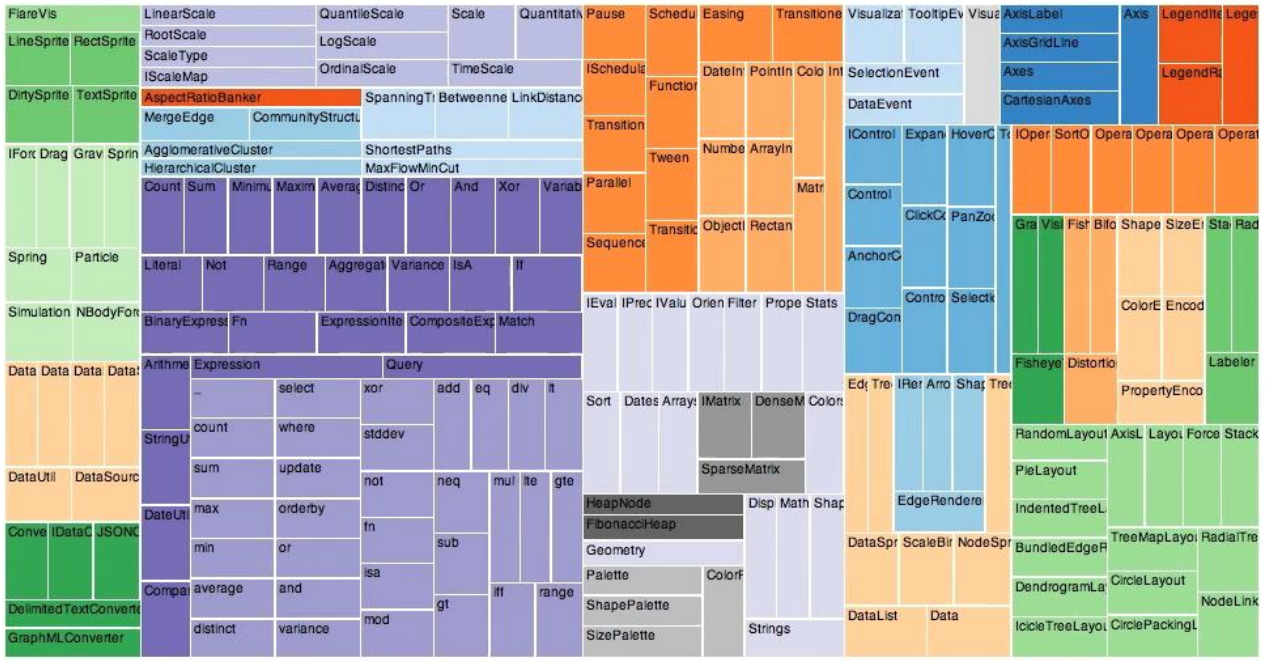
Geavanceerdere interacties met data zijn mogelijk

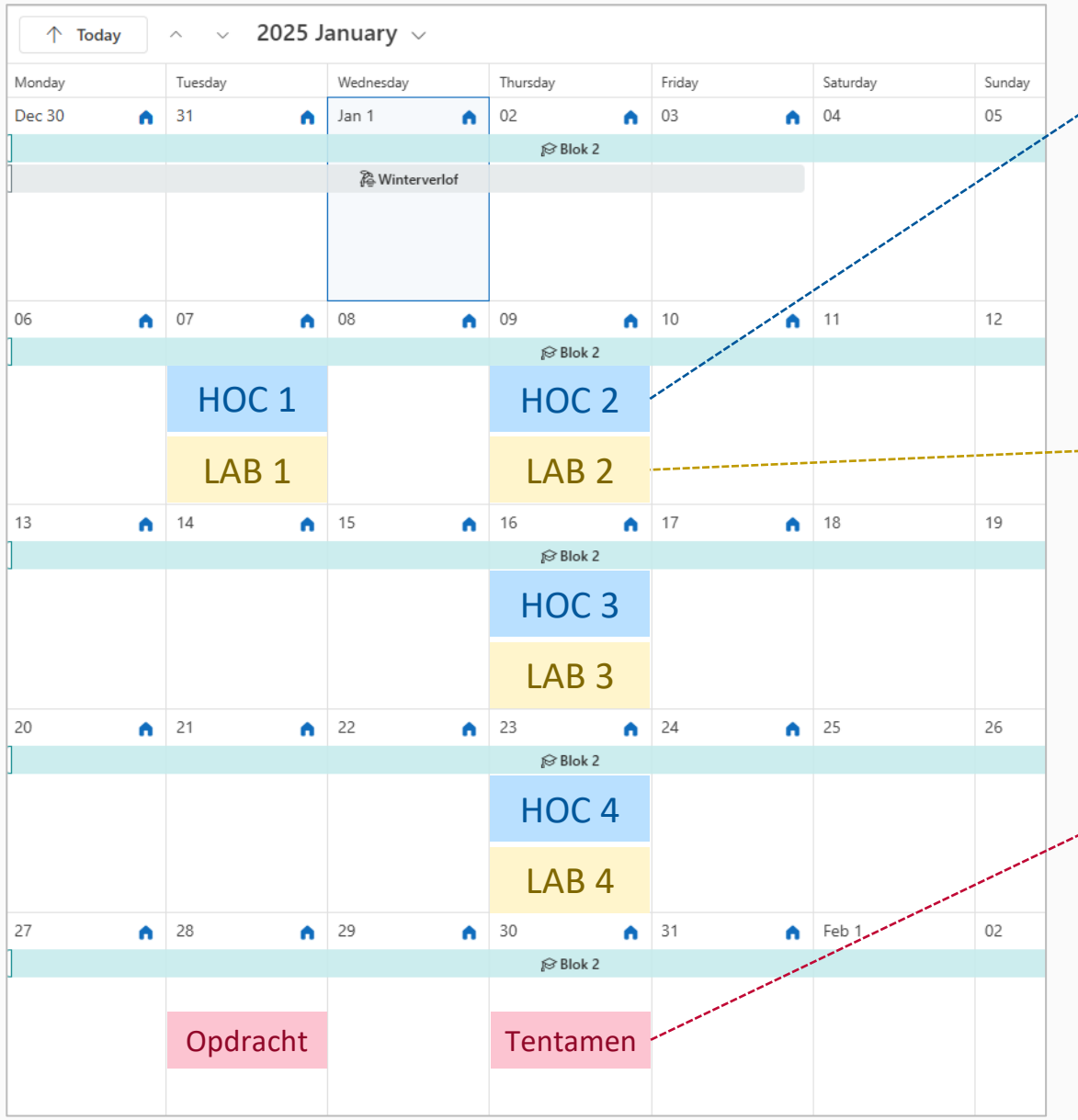
Grote grafische flexibiliteit

Typisch stijle leercurve

Enkele bekende vis-tools

D3 Treemap





👂 4 hoorcolleges

HOC 1: Algemene vis-principes en vis-types
HOC 2: Theoretische fundamenten van vis
HOC 3: Inleiding ggplot2 + hoe vis kan misleiden
HOC 4: Vis in rapporten + interactieve vis

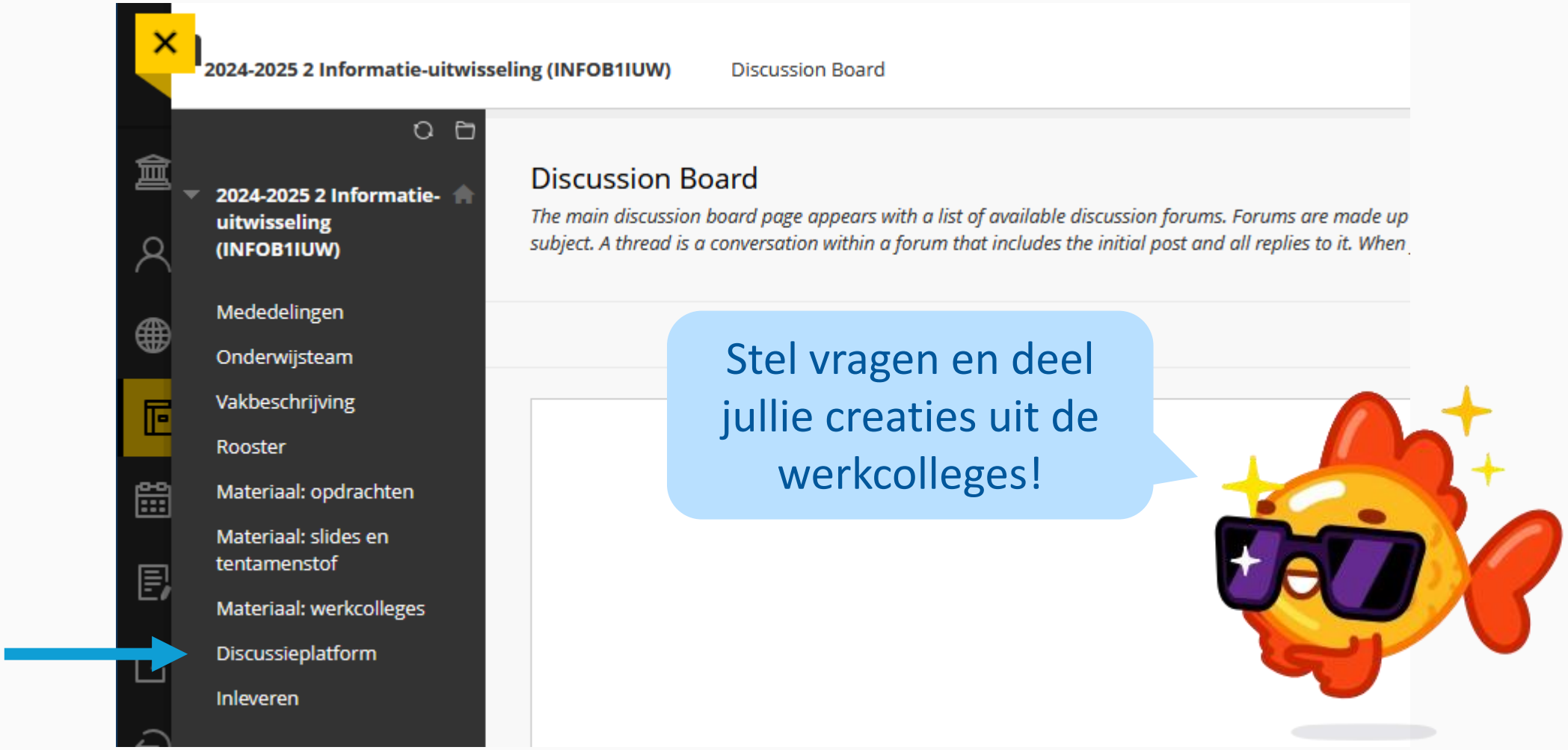
💪 4 werkcolleges

Hands-on ervaring met kleine datasets
Vis ontwerpen op papier
Vis implementeren in Excel en R
Bestaande (interactieve) vis analyseren

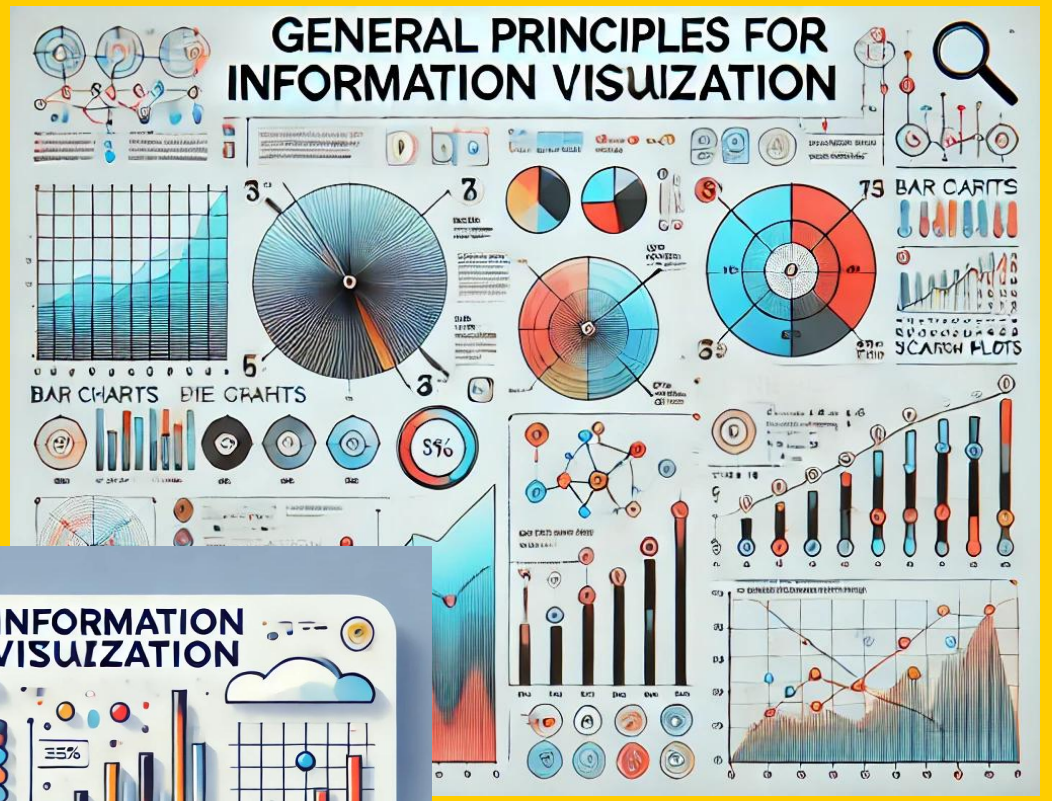
🎓 Leerdoelen

Fundamenten van vis kennen en toepassen
Vis kunnen ontwerpen en implementeren
Sterktes en zwaktes van concrete vis herkennen
Kritisch kunnen reflecteren over misleidende vis

Gebruik het discussieplatform op Blackboard



Hoorcollege 1
Algemene principes en visualisatietypes



Informatie-uitwisseling:
informatievisualisatie

Jeroen Ooge

Universiteit Utrecht