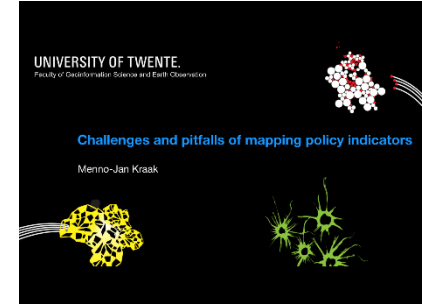




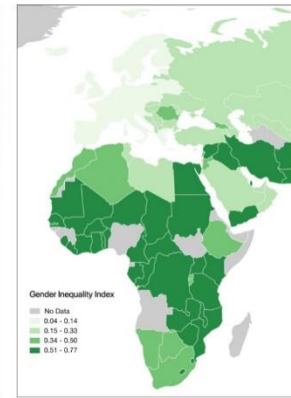
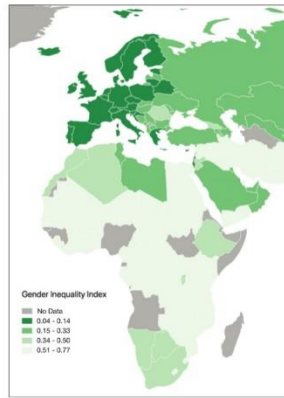
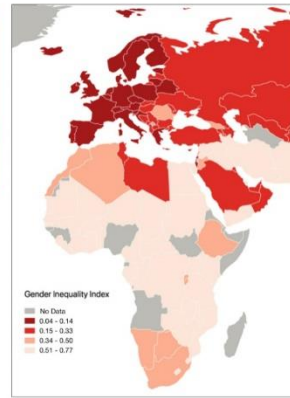
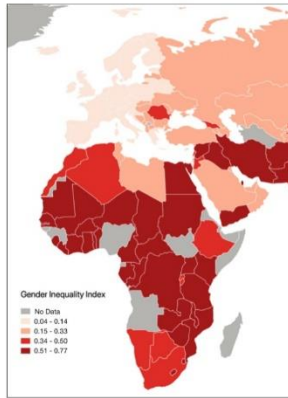
**Communiceren van statistische informatie (uit officiële statistieken en burgerbevragingen).**

**Reinoud Stoel en Martijn Tennekes**  
Centraal Bureau voor de Statistiek, Nederland

# De kleur maakt uit



## Choice of color - what do you want to tell?



# De context maakt uit



## Jeugdcriminaliteit daalt, maar niet alle vormen

31-5-2021 00:00



© Hollandse Hoogte / Robin Utrecht



NOS Nieuws • Maandag 31 mei 2021, 00:00 •  
Aangepast maandag 31 mei 2021, 08:39

**CBS: jeugdcriminaliteit neemt af, maar laatste jaren stijging ernstige delicten**



FINISHED FILES ARE THE  
RESULT OF YEARS OF SCIENTIFIC  
STUDY COMBINED WITH THE  
EXPERIENCE OF MANY YEARS.



M  
I  
N  
D  
F  
\*  
C  
K



M  
I  
N  
D  
F  
\*  
C  
K



ONS FEILBARE  
DENKEN

*Thinking, fast and slow*



DANIEL  
KAHNEMAN

NOBELPRIJSWINNAAR

*'Een zeer toegankelijk boek, dat leest als een trein'*

\*\*\*\*\* DE VOLKSKRANT

*'De belangrijkste psycholoog ter wereld'*

STEVEN PINKER

# Doel voor vandaag van mijn bijdrage

Creëren van bewustwording van het belang van cognitieve factoren bij het communiceren statistische informatie.



# Prof. Gerd Gigerenzer

Wat betekent '40%'?

- 1) Een kwart
- 2) 4 uit 10
- 3) elke 40<sup>ste</sup> persoon



40%; 0,4; 4/10 worden (door 'leken') niet altijd als gelijkwaardig beschouwd. Uit het werk van Gigerenzer volgt dat mensen **'natuurlijke frequenties'** het best begrijpen (4 uit 10).

Natuurlijke frequenties en onzekerheid:

4 uit 10, 40 uit 100, 400 uit 1000,...

# Natuurlijke frequenties en onzekerheid

Gevoelsmatig:

- Des te groter het aantal, des te kleiner de onzekerheid
- Des te kleiner de onzekerheid, des te smaller het betrouwbaarheidsinterval

95% betrouwbaarheidsinterval

4 uit 10  $\rightarrow$  {0,17;0,69}

40 uit 100  $\rightarrow$  {0,31;0,50}

400 uit 1000  $\rightarrow$  {0,37;0,43}

400000 uit 4000000  $\rightarrow$  {0,399; 0,401}

## Een DNA match...

DNA deskundige:

*'De random match probability is gelijk aan 1/10.000.  
Oftewel de kans op een match als de verdachte niet de  
donor is heel erg klein'.*

**P(match | verdachte is niet donor)**

Rechter:

*'De kans dat de verdachte niet de donor is heel erg klein'.*

**P(verdachte is niet donor | match)**



# Als cijfers al verkeerd geïnterpreteerd kunnen worden...

Hoe zit het dan bij visualisaties?

# Het brein: Je beste vriend, je ergste vijand.

Over de cognitieve aspecten van visualisatie



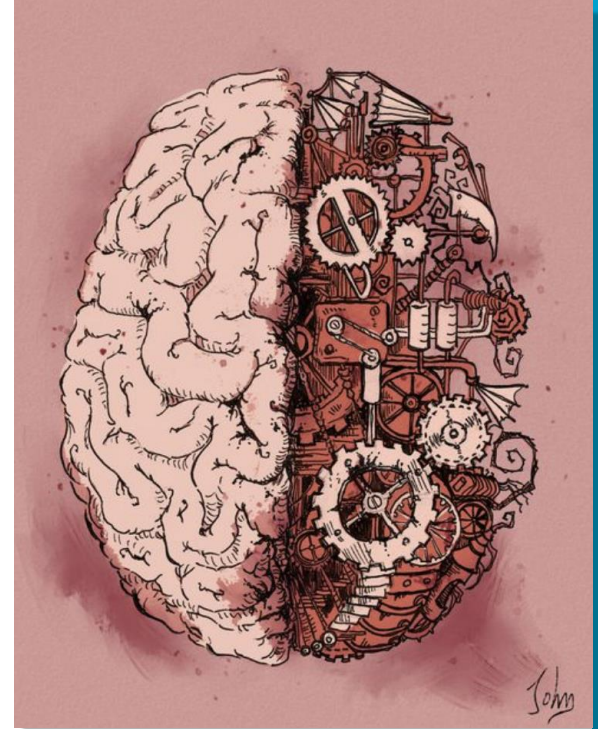


# Het brein

Jouw brein doet ongelooflijke dingen,  
vaak zonder dat je het door hebt.

Het brein ligt ten grondslag aan expertise  
en intelligentie.

Echter, juist de aspecten die jouw brein tot je grootste  
vriend maken, maken het ook tot je grootste vijand.



# Het brein

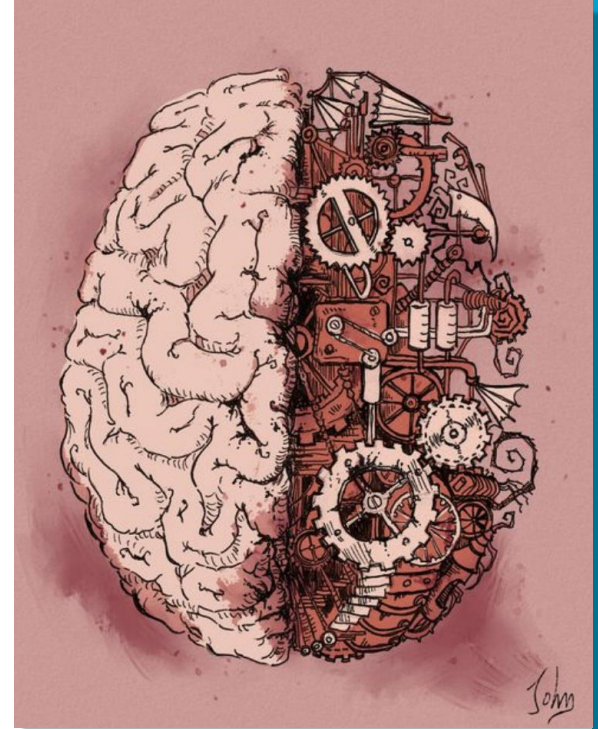
‘The mind is not a camera’

‘Machine’ met beperkte middelen.

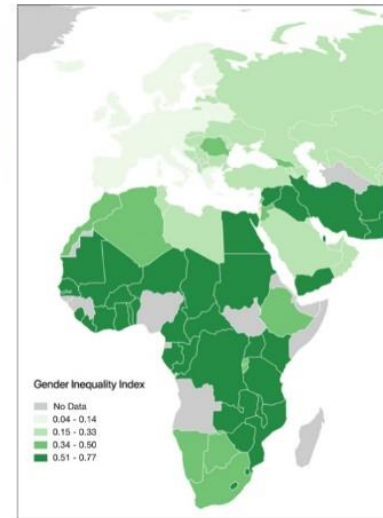
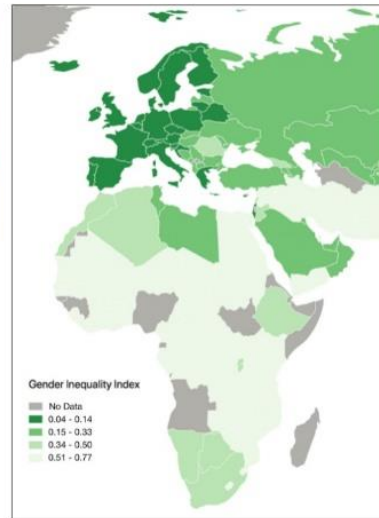
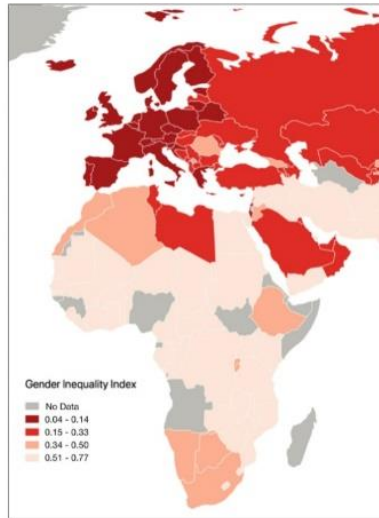
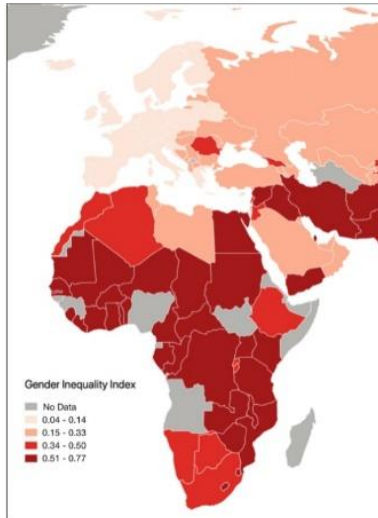
Actief i.p.v. passief: we organiseren de wereld,  
we maken veronderstellingen,  
en leggen structuur op.

We nemen de dingen niet waar van 'zoals ze zijn’.

Het is de basis van intelligentie en expertise....en leidt tot  
allerlei ‘bijverschijnselen’



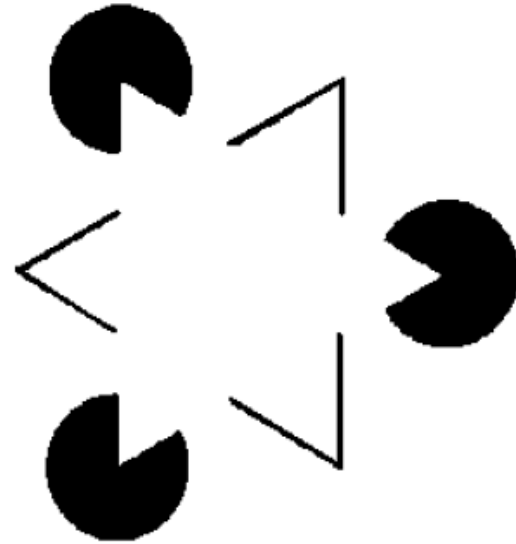
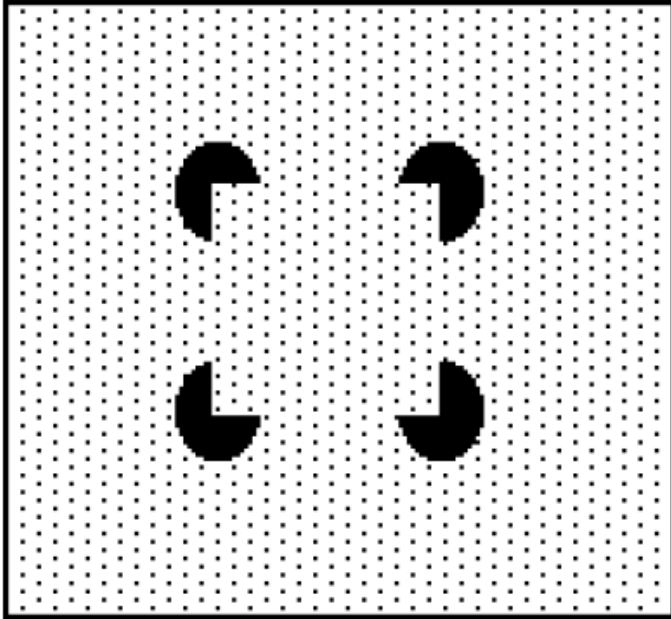
# Choice of color - what do you want to tell?



THE PWEOR  
OF  
CONGITVIE  
KONLWDEGE

aan de hand van 9 lessen

# Les 1: We zien dingen wel die er niet zijn



## Les 2: We zien dingen niet die er wel zijn

FINISHED FILES ARE THE  
RESULT OF YEARS OF SCIENTIFIC  
STUDY COMBINED WITH THE  
EXPERIENCE OF MANY YEARS.

Een zwarte man, met zwarte kleren, loopt op straat in het  
midden van  
de weg.

De straatverlichting is uit en opeens komt er een zwarte  
auto met  
gedoofde lichten recht op hem afgereden.

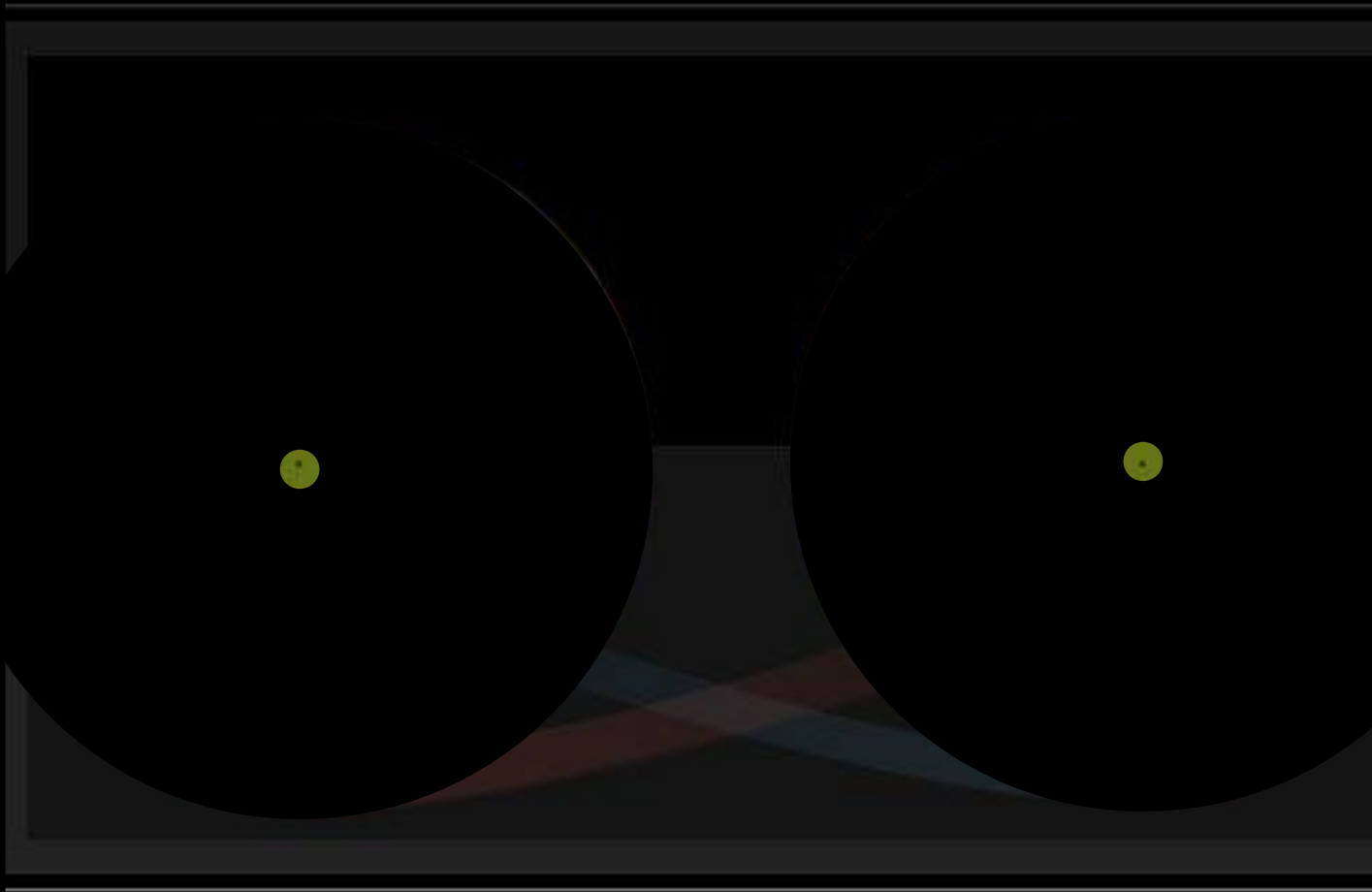
De bestuurder van de auto kan de man  
gelukkig nog op tijd ontwijken.

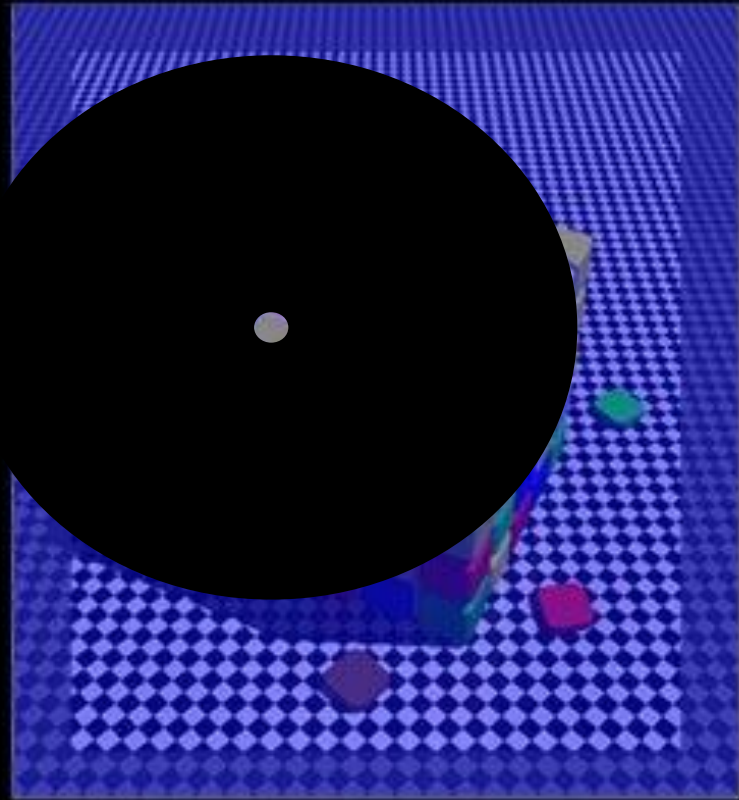
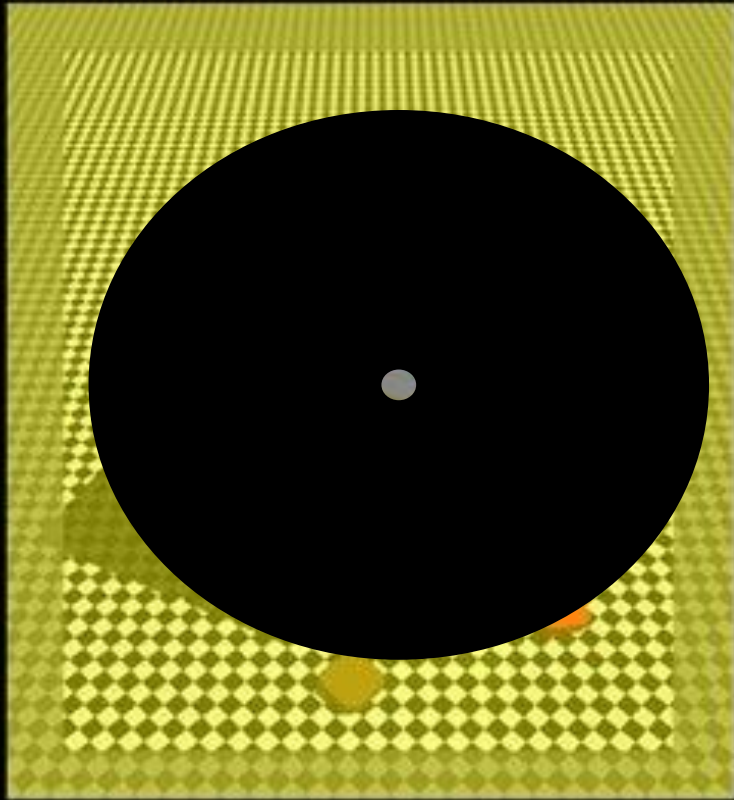
Hoe kan dit?

## Les 3: We maken aannames









# Les 4: We interpretern in de context

**12**  
**A B C**  
**14**

## Les 5: We geven betekenis (zelf als die er niet is)





# Les 6: Eenmaal gezien blijft gezien





# Les 7: De eerste indruk laat zich moeilijk veranderen





## Les 8: Ons brein doet dingen automatisch

**YELLOW**

**BLACK**

**PURPLE**

**ORANGE**

**BLUE**

**GREEN**

**BLUE**

**RED**

**YELLOW**

**GREEN**

**RED**

**BLUE**

**ORANGE**

**GREEN**

**RED**

**BLACK**

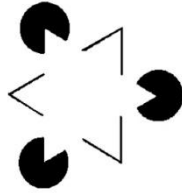
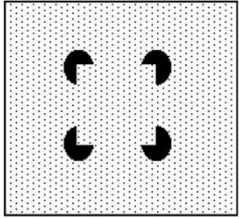
**PURPLE**

**ORANGE**

## Les 9: Ons brein zoekt naar bevestiging

2 4 6 8 ...

Les 1: We zien dingen wel die er niet zijn



Les 2: We zien dingen niet die er wel zijn

FINISHED FILES ARE THE RESULT OF YEARS OF SCIENTIFIC STUDY COMBINED WITH THE EXPERIENCE OF MANY YEARS.

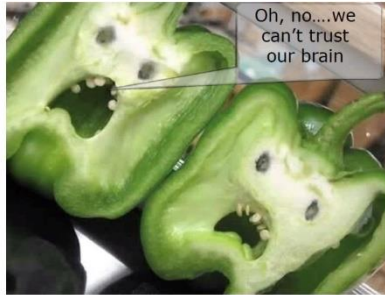
Les 3: We maken aannames



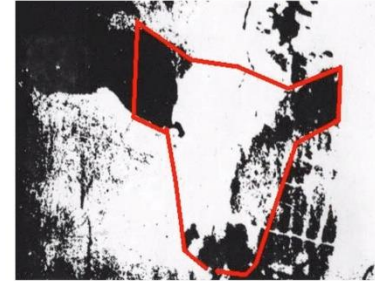
Les 4: We hebben moeite met context

12  
ABC  
14

Les 5: We geven betekenis (zelf als die er niet is)



Les 6: Eenmaal gezien blijft gezien



Les 7: De eerste indruk laat zich moeilijk veranderen



Les 8: Je brein doet dingen automatisch

YELLOW	BLUE	ORANGE
BLACK	RED	GREEN
PURPLE	YELLOW	RED
ORANGE	GREEN	BLACK
BLUE	RED	PURPLE
GREEN	BLUE	ORANGE

Les 9: Ons brein zoekt naar bevestiging

2 4 6 8 ...

# Spelen dit soort aspecten een rol bij de communicatie van statistische informatie?

Wat is de boodschap die je wilt overbrengen?

Wordt die boodschap zo ervaren?

Wordt die boodschap door iedereen zo ervaren?

Is er kennis uit de cognitieve wetenschappen die hierbij kan helpen? Is nader (experimenteel) onderzoek nodig?



<https://www.youtube.com/watch?v=ubNF9QNEQLA>



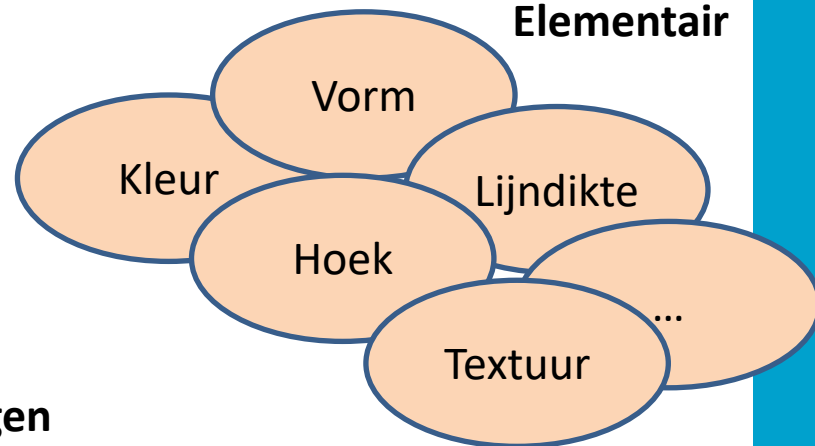
# Inzicht in datavisualisatie

Communicatie van statistische informatie

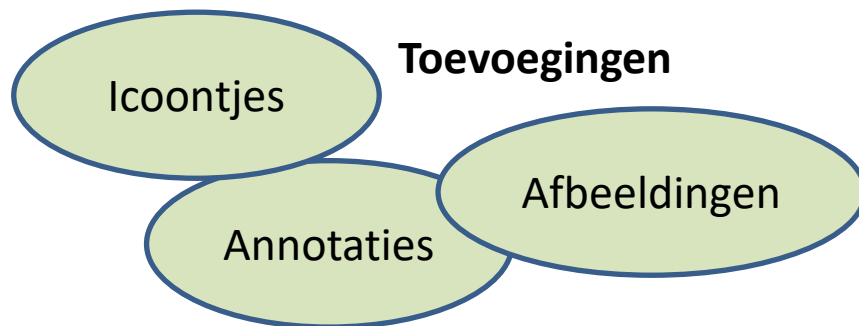


# Bouwstenen van grafieken

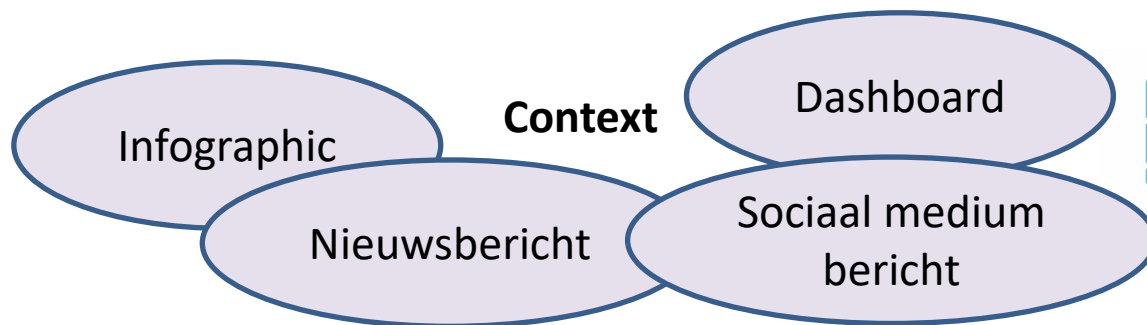
Elementair



Toevoegingen



Context

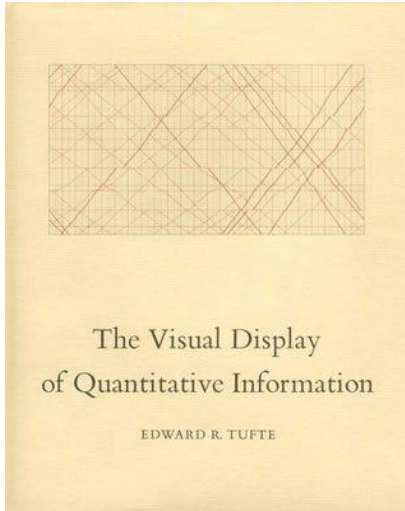


# Wat is een goede visualisatie?

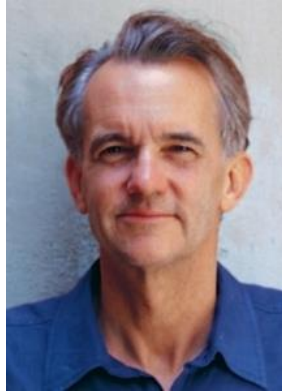




# Volgens de functionalisten



De bijbel voor informatie visualisatie, 1983



Edward Tufte



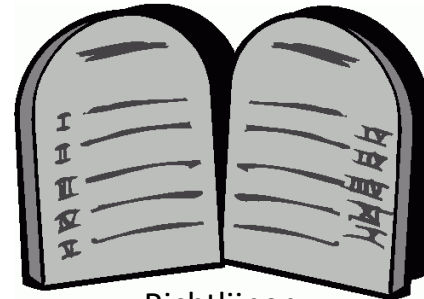
Stephen Few



Alberto Cairo

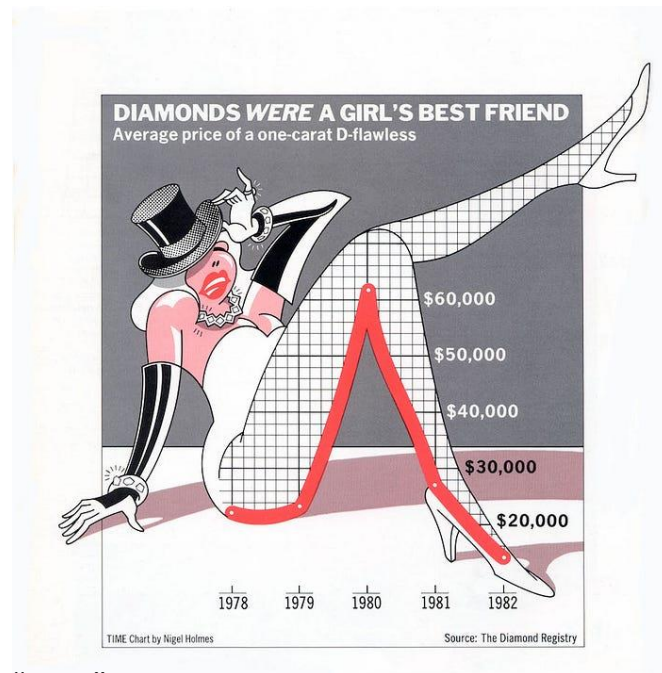


Jelke Bethlehem (oud CBSer)  
... en vele anderen



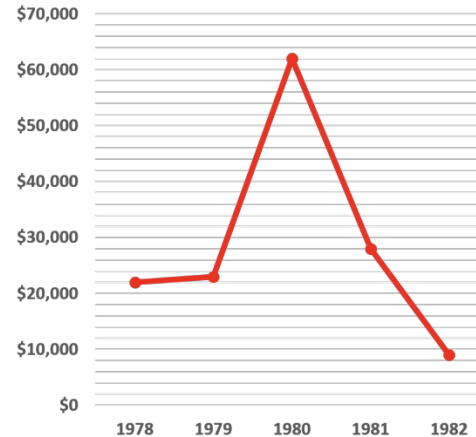
Richtlijnen

# Data-inkt verhouding: gebruik zo weinig mogelijk inkt



“Fout”

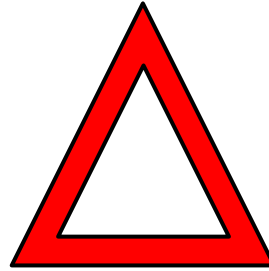
**DIAMONDS WERE A GIRL'S BEST FRIEND**  
Average price of a one-carat D-flawless



“Goed”

# Andere (onderbelichte) invalshoeken:

1. Waarnemingspsychologie
2. Cognitieve psychologie
3. Informatietheorie
4. Mens-computer-interactie
5. Vragenlijstontwikkeling



Theorieën/richtlijnen  
kunnen tegenstrijdig zijn!

Data-inkt verhouding



Gebruik textuur en 3D effecter  
herkenbaarheid van objecten t

Afbeeldingen/iconen kunnen helpen de  
context te begrijpen en de boodschap te  
onthouden.





# Gebruikersonderzoek!

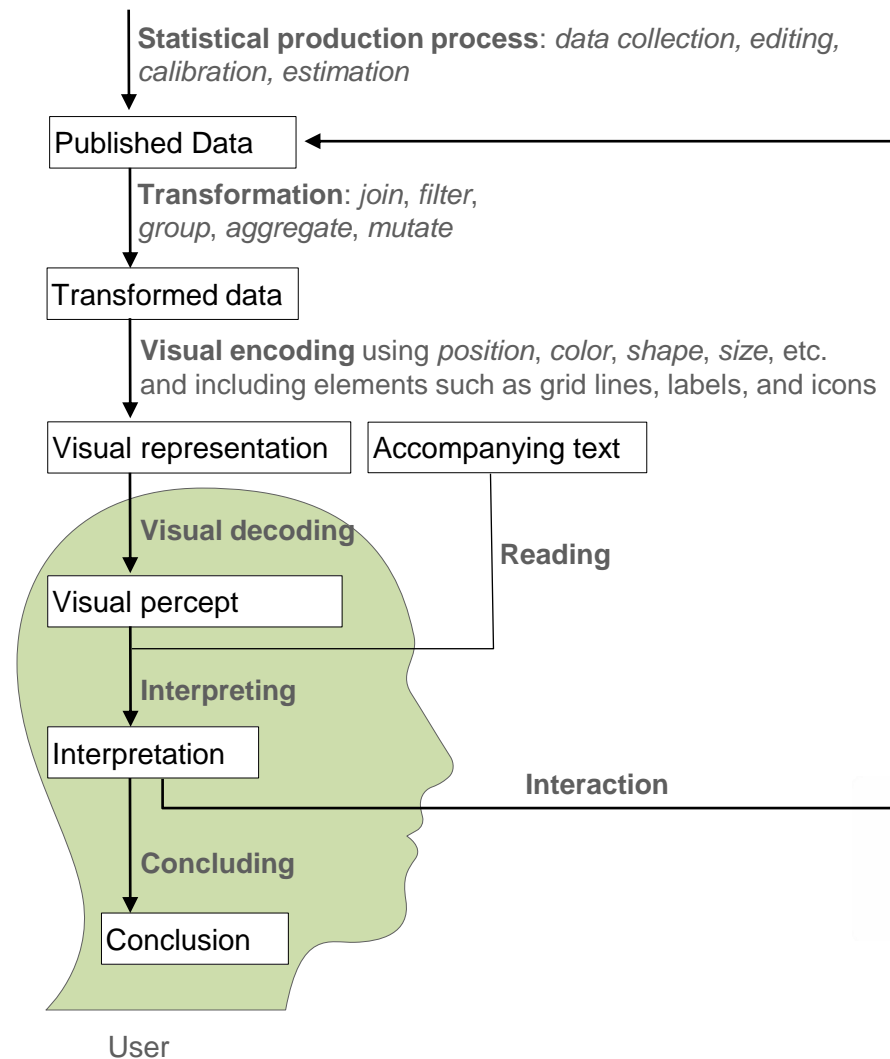
# De gebruiker

## Gebaseerd op deze bronnen

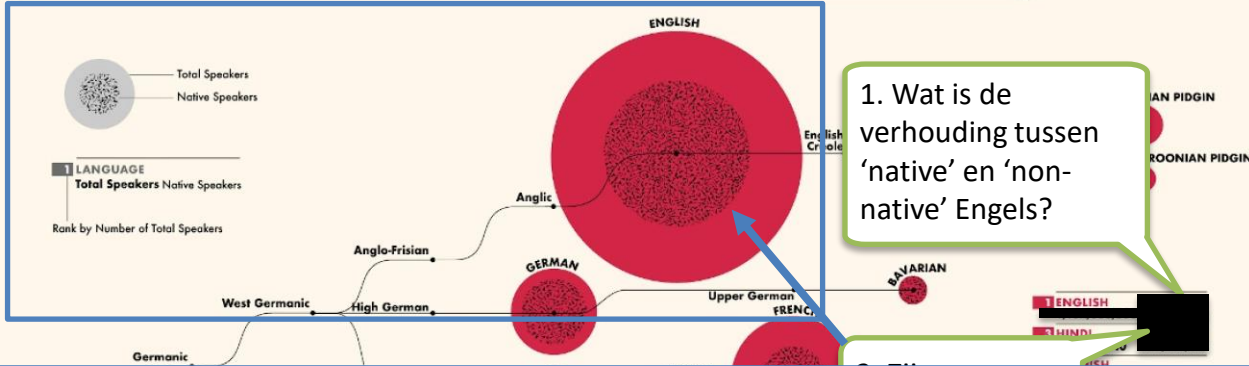
Chen, M. & Golan, A (2016) What may visualization processes optimize? IEEE Transactions on Visualization and Computer Graphics 22:12, 2619–2632.

Rensink, Ronald. (2014). On the Prospects for a Science of Visualization. 10.1007/978-1-4614-7485-2\_6.

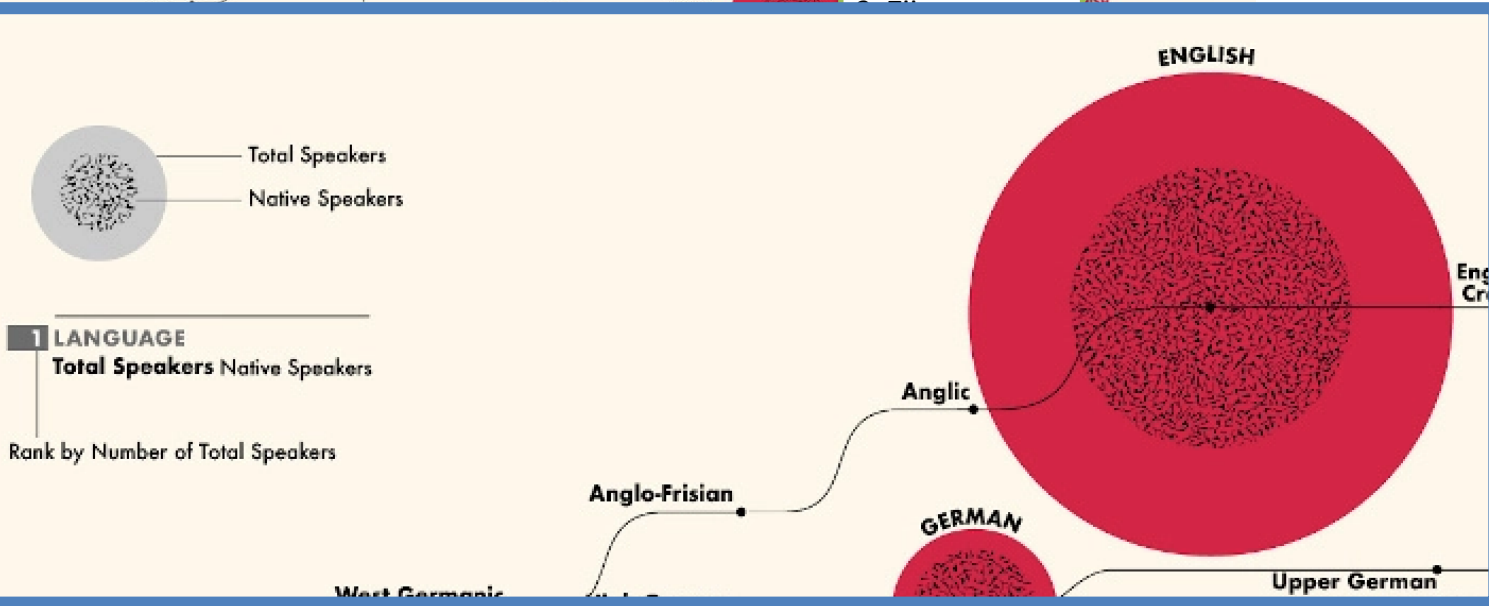
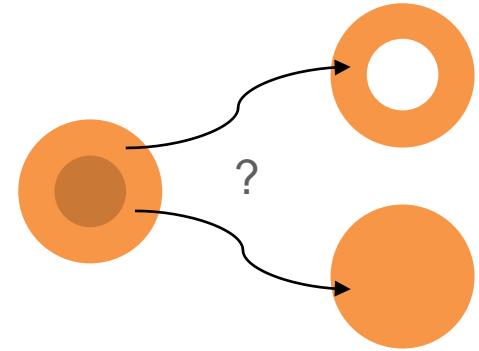
Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. New York: Cambridge University Press.



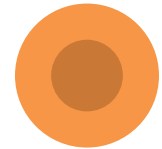
# The 100 Most-Spoken Languages in the World



Onduidelijkheid



of illusion

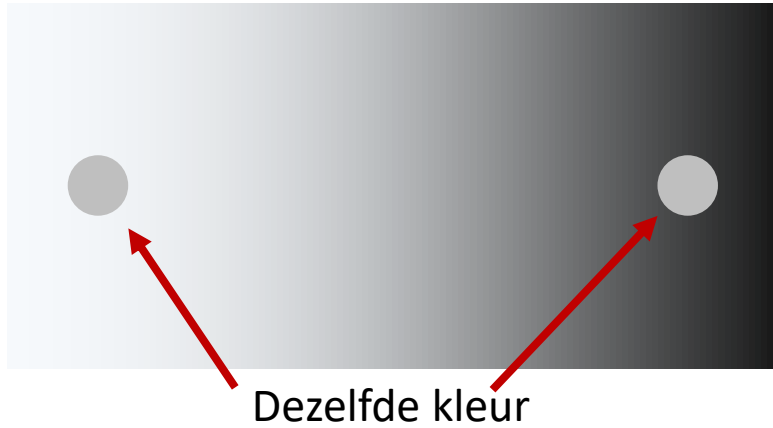


enste cirkel is

# Welk bord pak je (als je trek hebt)?

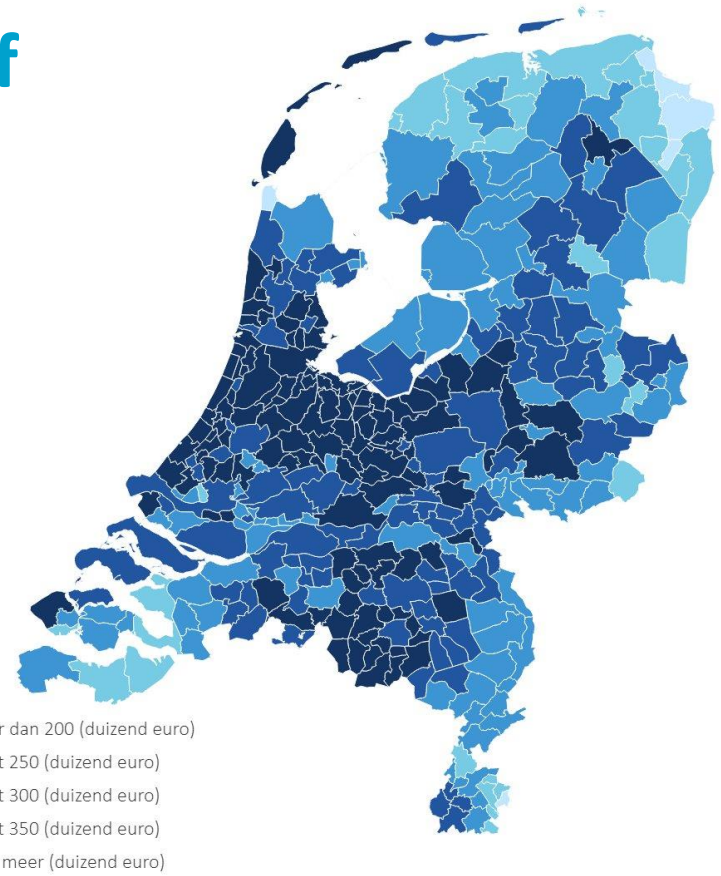
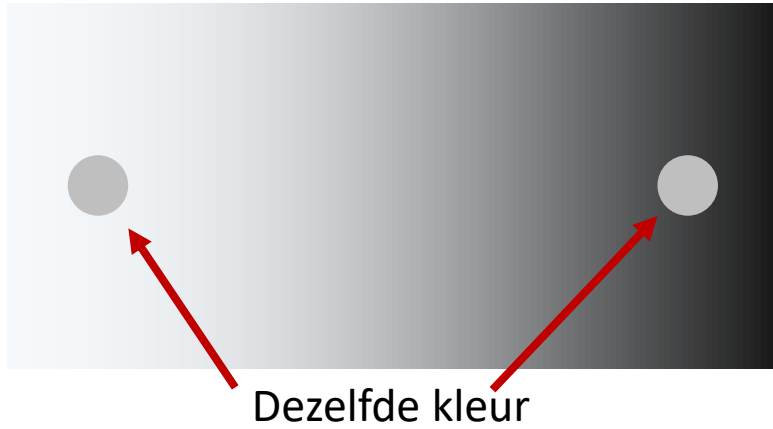


# Kleurwaarneming is relatief

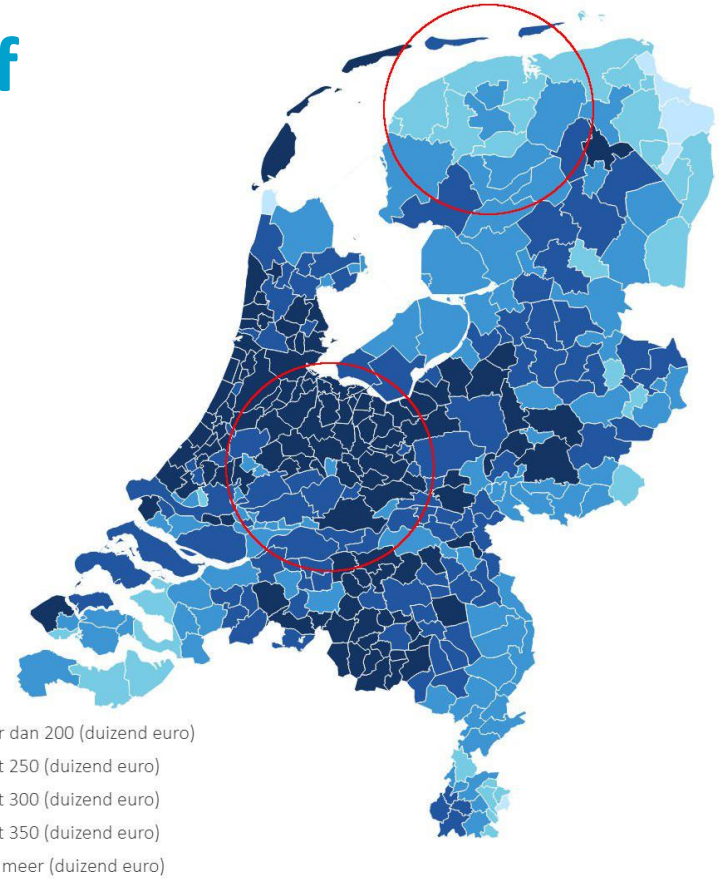
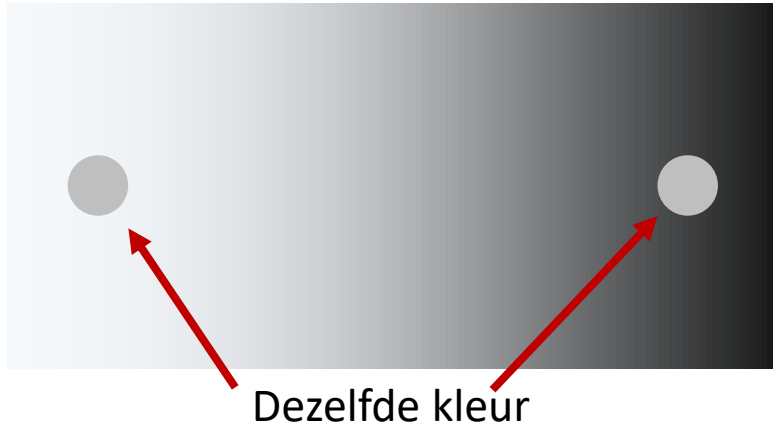




# Kleurwaarneming is relatief



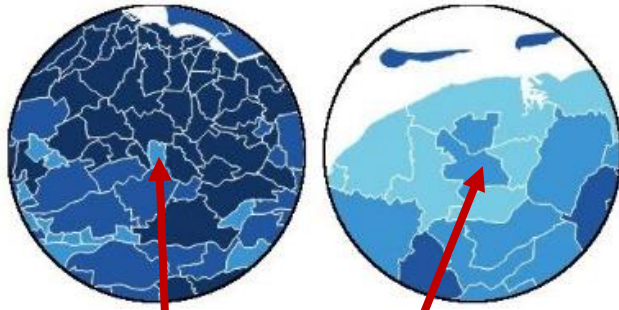
# Kleurwaarneming is relatief



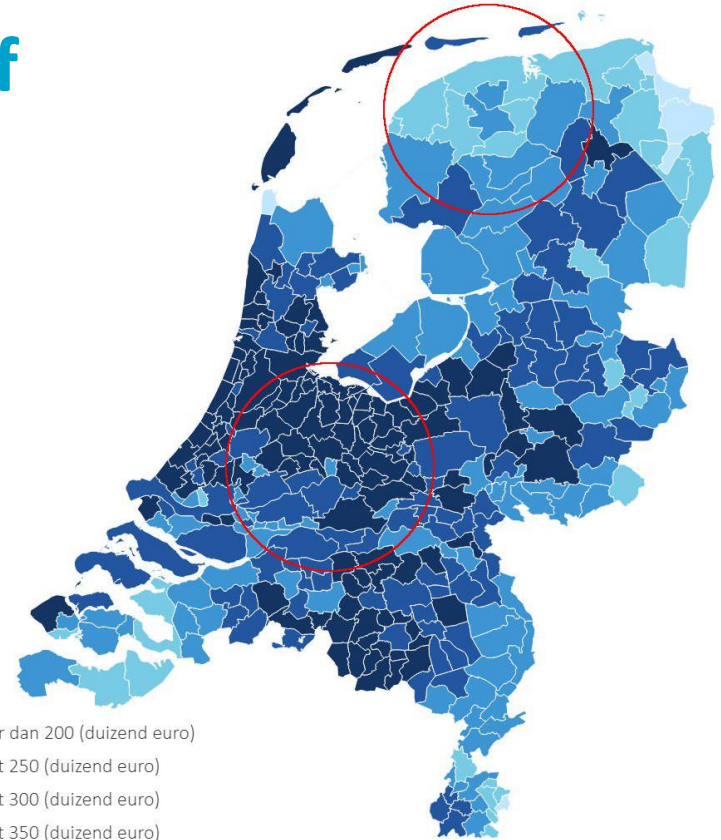
# Kleurwaarneming is relatief



Dezelfde kleur



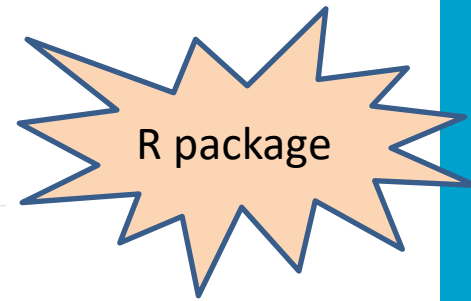
Dezelfde kleur



- Minder dan 200 (duizend euro)
- 200 tot 250 (duizend euro)
- 250 tot 300 (duizend euro)
- 300 tot 350 (duizend euro)
- 350 of meer (duizend euro)



# Analyse van kleuren



**Palette type**

- Categorical
- Sequential
- Diverging
- Bivariate

**Number of colors**

2 7 36

Color for missing values

**Color vision**

- Normal
- Deutan (red-green blind)
- Protan (also red-green blind)
- Tritan (blue-yellow)

**Palette series**

- brewer
- c4a
- carto
- hcl
- kovesi
- met
- misc
- parks
- poly
- scico
- seaborn
- stevens
- tableau
- tol
- viridis
- wes

**Text format**

- Hex
- RGB
- HCL

**Text color**

- Hide text
- Black
- White
- Automatic

**Sort**

Name

- Reverse sorting
- Show scores

All None Overview

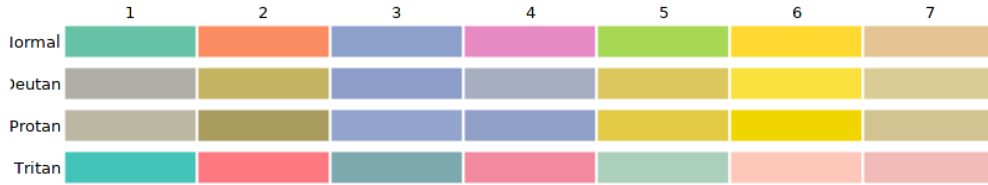
Series	Name	Max number	Colorblind-friendly	Vivid	Fair	Low contrast	3D Blues	Blues							References	
								1	2	3	4	5	6	7		
brewer	accent	7	👁️	🕶️	✖️	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
brewer	dark2	7	👁️	🕶️		🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
brewer	paired	12		🕶️	✖️	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
brewer	pastel1	8	👁️	✳️	👉	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
brewer	pastel2	7	👁️	✳️	👉	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
brewer	set1	8		🕶️	✖️	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
brewer	set2	7	👁️			🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
brewer	set3	11	👁️			🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
carto	antique	11		✳️		🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
carto	bold	12		🕶️	✖️	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
carto	pastel	11	👁️	👉	👉	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
carto	prism	11		🕶️		🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
carto	safe	11	😊		✖️	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
carto	vivid	12		🕶️	✖️	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
hcl	cold	36	👁️	✳️	👉	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
hcl	dark2	36	👁️	✳️	👉	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
hcl	dark3	36			👉	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R
hcl	dynamic	36	👁️	✳️	👉	🏠		🟢	🟡	🟠	🟤	🟡	🟢	🟠	🟤	📄 Bib JS R



Palette

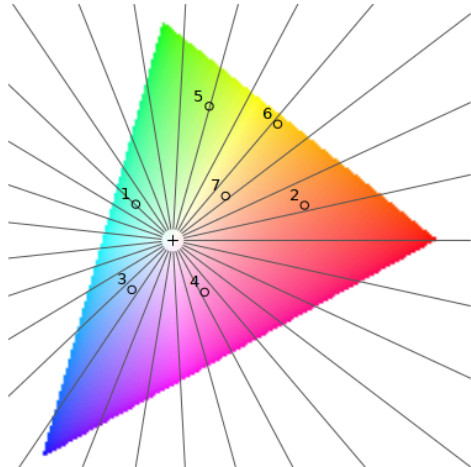
brewer.set2

## Color blindness simulation i

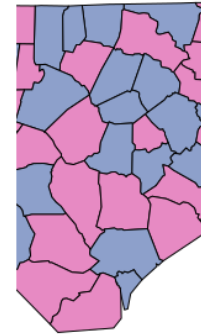
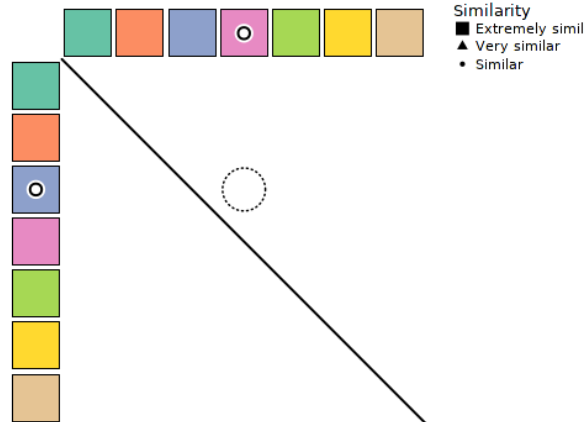


## Hue lines i

Normal color vision

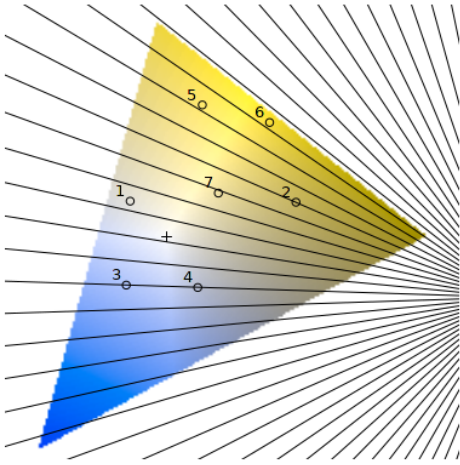


## Similarity matrix i

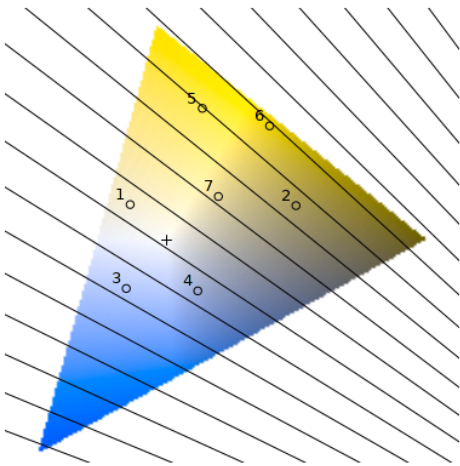


## Confusion lines

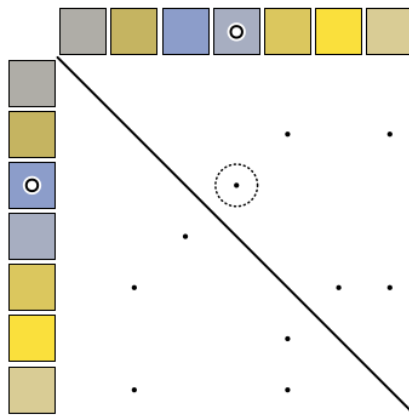
Deutan (red-green blind)



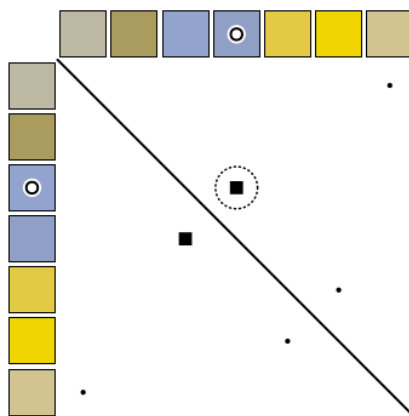
Protan (also red-green blind)



## Perceived similarity matrices

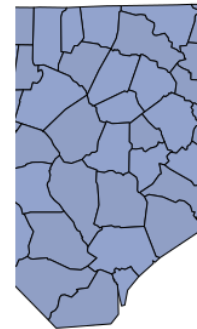
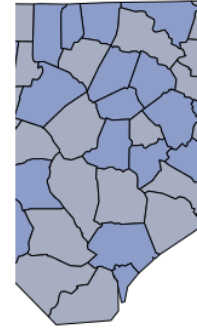


Similarity  
■ Extremely simil  
▲ Very similar  
• Similar

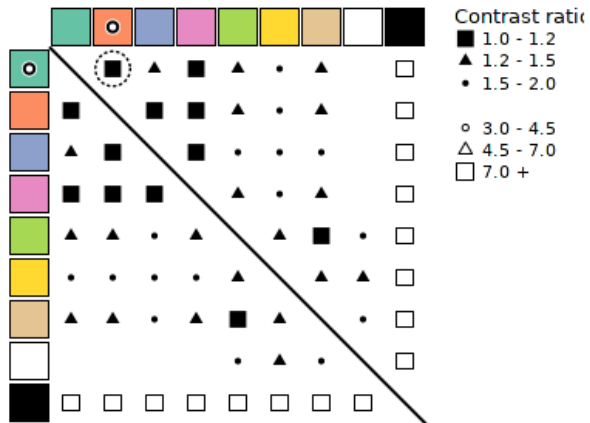


Similarity  
■ Extremely simil  
▲ Very similar  
• Similar

# Kleurenblindheid



## Contrast ratio



# Contrast

## Text readability

Contrast ratio (1.10): not safe to print text according to the WCAG 2.1

Font size 14pt

Font size 18pt

Font size 14pt

Font size 18pt

## Border lines needed?

Contrast ratio (1.10): strongly recommended to use border lines when plotting these colors next to each other

### Example chart

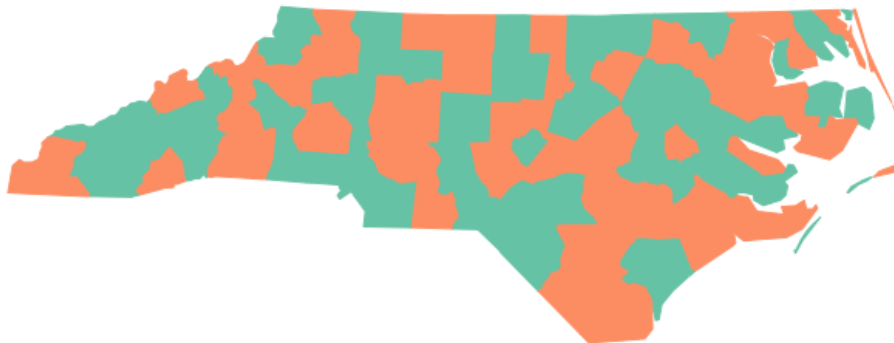
- Choropleth
- Barchart

### Line Width

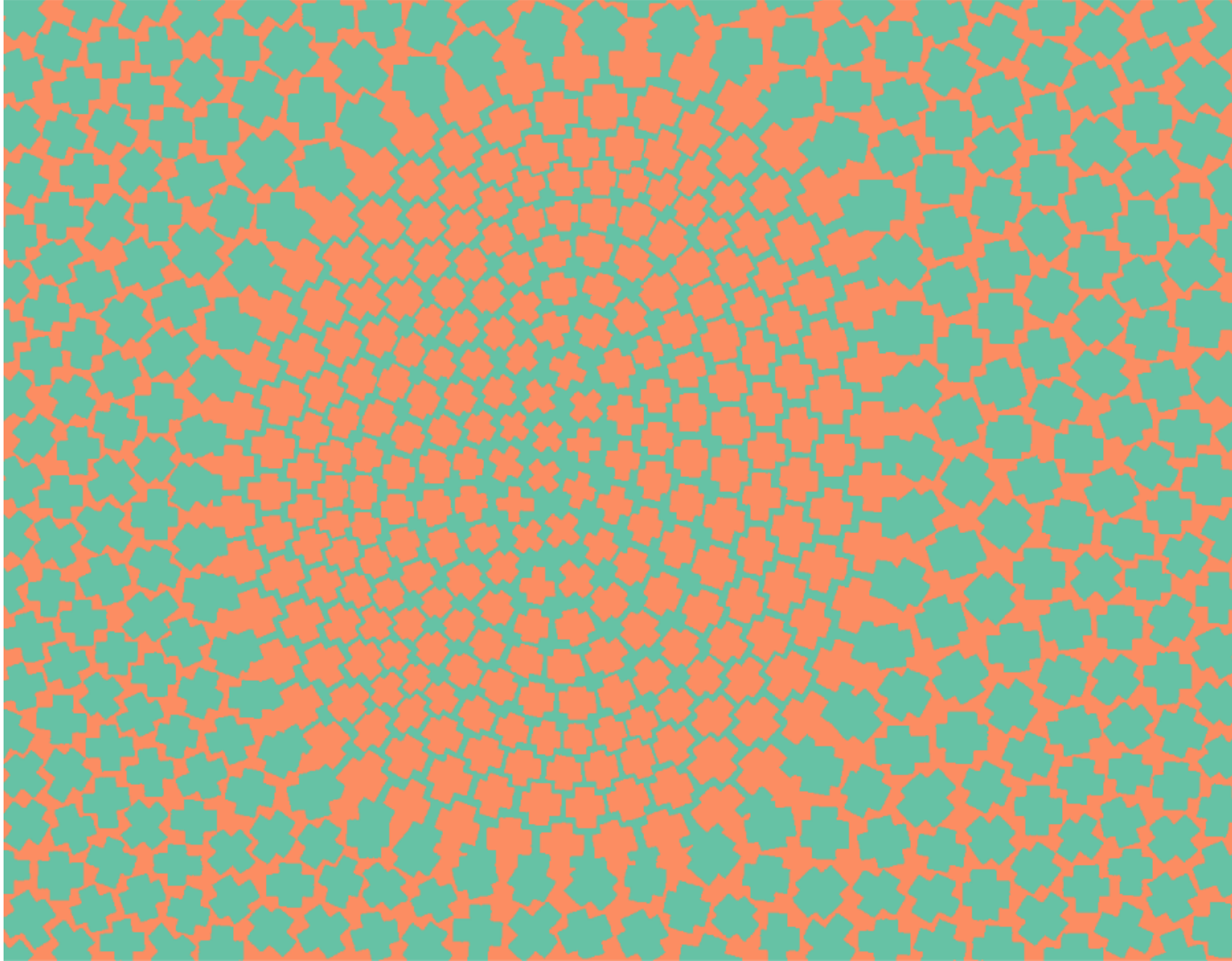


### Borders

black









Example chart

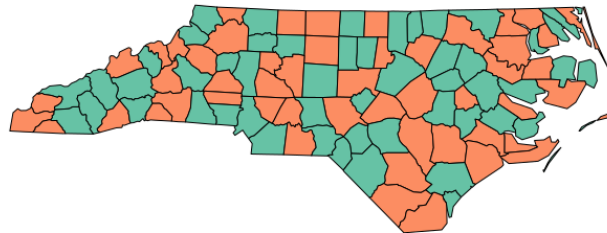
- Choropleth
- Bar chart

Line Width

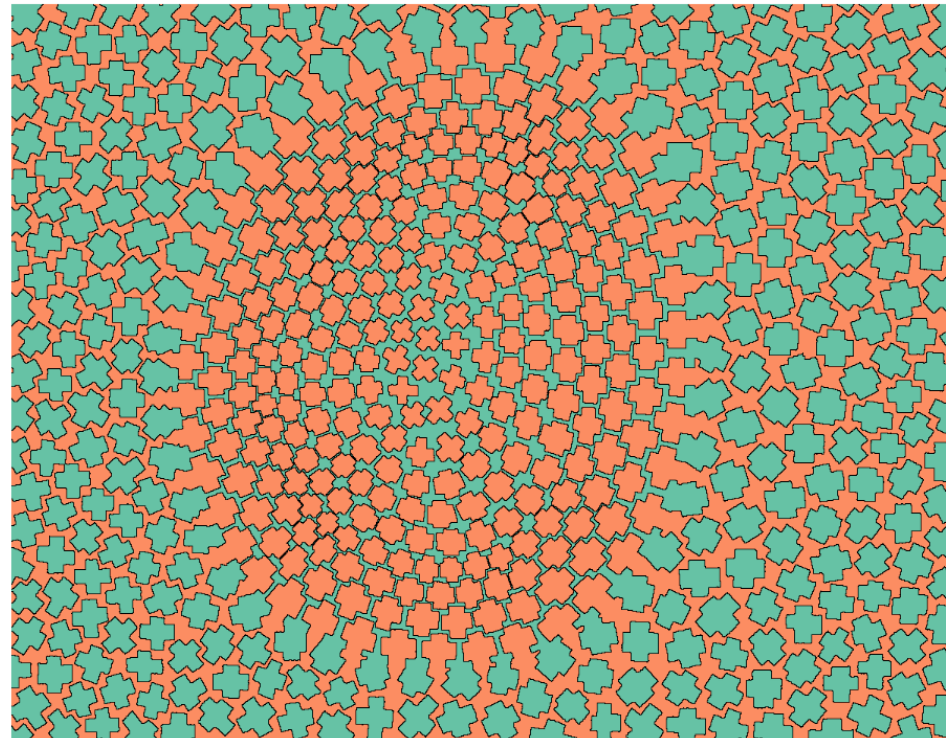


Borders

black



Optical Illusion Art



# Infographics



## **Nederland in Cijfers**

Een boekje met infographics over onze statistieken



# Grafiekpolitie | CBS gebruikt kromme grafiek over online gedrag 65-plussers



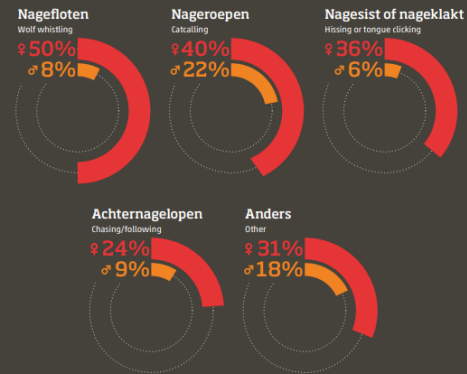
Om het online gedrag van ouderen weer te geven gebruikt het CBS in plaats van een staafdiagram een weergave in de vorm van concentrische cirkels. Dat oogt speels en decoratief, maar de grafiek verliest aan duidelijkheid: de lengte van de staven staat namelijk niet meer in verhouding tot de waarde die zij weergeven.



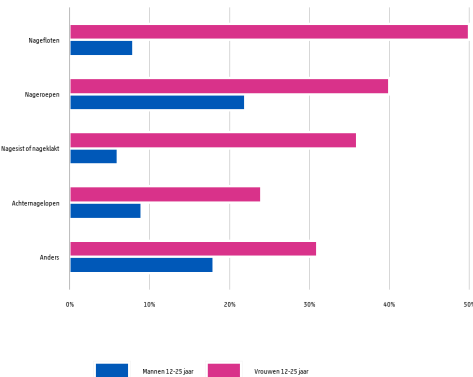
# Gebruikersstudie



## Kunnen jonge vrouwen rustig over straat? How many young women encounter street harassment?

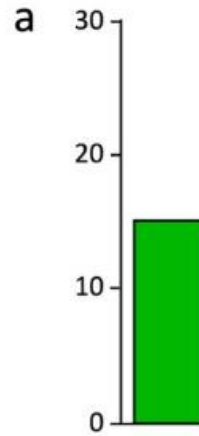


## Straatintimidatie zoals ervaren door jongeren

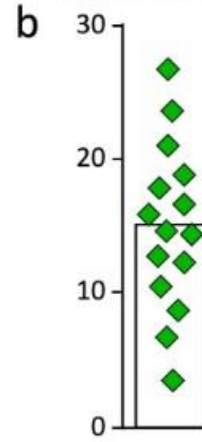


# De ene staafdiagram is de andere niet...

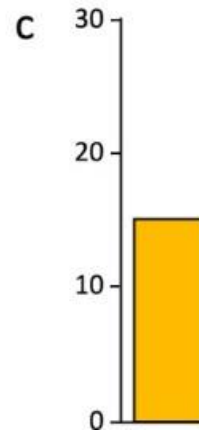
Mean bar graph



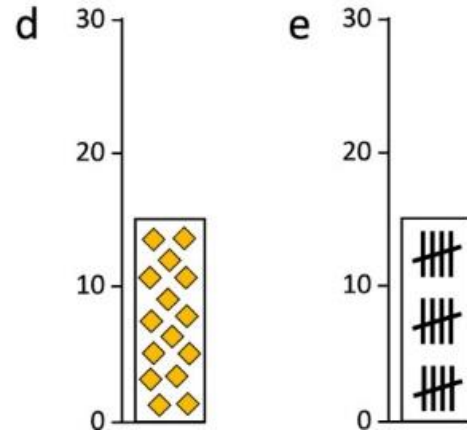
Mean bar graph data



Count bar graph



Count bar graph data



*Two graph walk into a bar*, Kerns et al, 2021, Journal of Vision vol 21

# Informatietheorie

*Visual Design*  
is to find the  
best way to  
**depict** some  
information

*Visual Design*  
is to find the  
best way to  
**lose** some  
information



# London Underground



**TRANSPORT FOR LONDON**  
EVERY JOURNEY MATTERS



**Check before you travel**

1. **Heathrow**  
TFL Rail customers should change at Terminal 2 & 3 for free rail transfer to Terminal 5

2. **Roundabout West**  
Step-free access for manual wheelchairs only

3. **Services or access at these stations are subject to variation.**  
Please search 'TFL Stations' for full details.

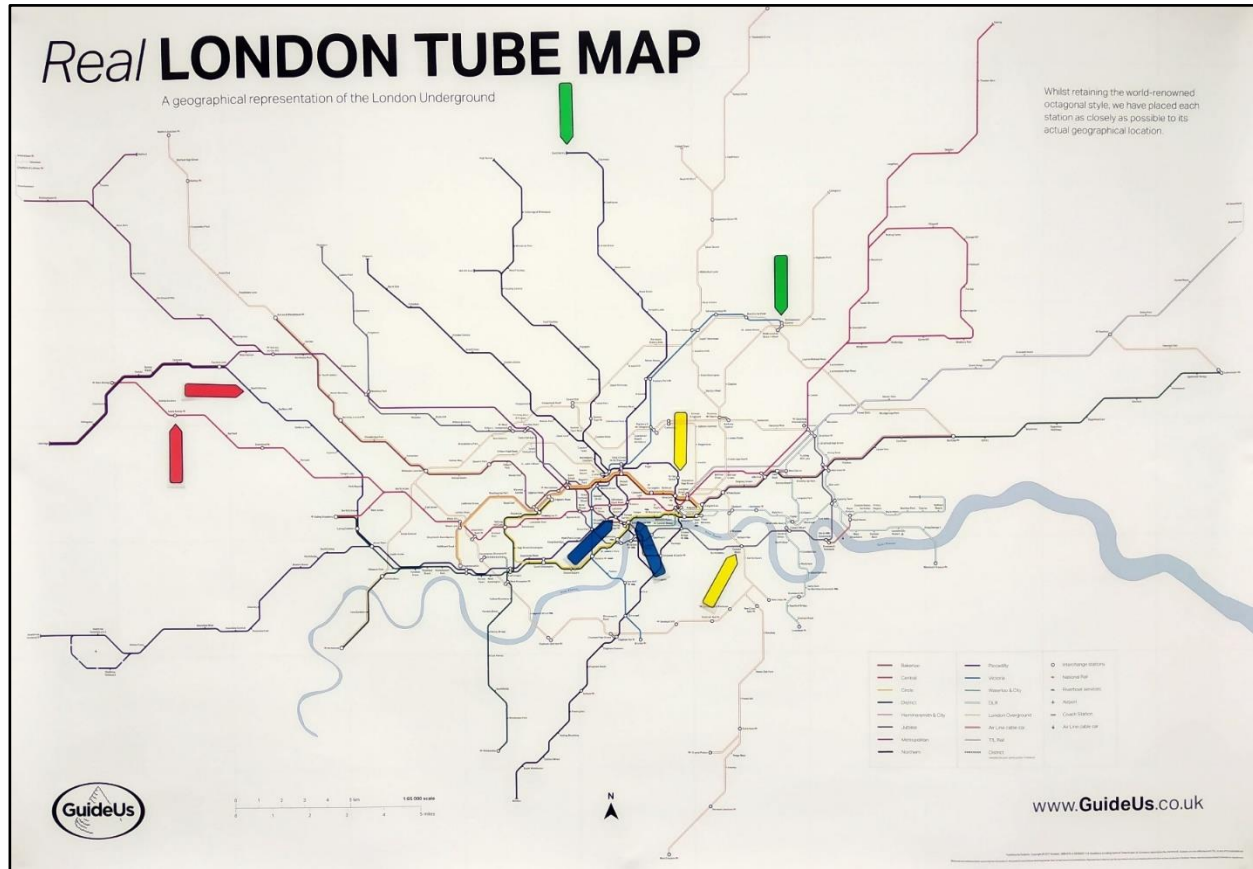
- Key to lines**
- Bakerloo
  - Central
  - Circle
  - District
  - Hammersmith & City
  - Jubilee
  - Metropolitan
  - Northern
  - Piccadilly
  - Victoria
  - Waterloo & City
  - DLR
  - Emirates Air Line cable car (special fares apply)
  - London Overground
  - TFL Rail
  - London Trams
  - District (see website and/or local media website)

- Key to symbols**
- Interchange stations
  - Step-free access from street to train
  - Step-free access from street to platform
  - National Rail
  - Airport
  - Riverboat services
  - Victoria Coach Station
  - Emirates Air Line cable car

- Explanation of zones**
- 1. Zone 1
  - 2. Zone 2
  - 3. Zone 3
  - 4. Zone 4
  - 5. Zone 5
  - 6. Zone 6
  - 7. Zone 7
  - 8. Zone 8
  - 9. Zone 9
  - 10. Zone 10
  - 11. Zone 11
  - 12. Zone 12
  - 13. Zone 13
  - 14. Zone 14
  - 15. Zone 15
  - 16. Zone 16
  - 17. Zone 17
  - 18. Zone 18
  - 19. Zone 19
  - 20. Zone 20
  - 21. Zone 21
  - 22. Zone 22
  - 23. Zone 23
  - 24. Zone 24
  - 25. Zone 25
  - 26. Zone 26
  - 27. Zone 27
  - 28. Zone 28
  - 29. Zone 29
  - 30. Zone 30
  - 31. Zone 31
  - 32. Zone 32
  - 33. Zone 33
  - 34. Zone 34
  - 35. Zone 35
  - 36. Zone 36
  - 37. Zone 37
  - 38. Zone 38
  - 39. Zone 39
  - 40. Zone 40
  - 41. Zone 41
  - 42. Zone 42
  - 43. Zone 43
  - 44. Zone 44
  - 45. Zone 45
  - 46. Zone 46
  - 47. Zone 47
  - 48. Zone 48
  - 49. Zone 49
  - 50. Zone 50
  - 51. Zone 51
  - 52. Zone 52
  - 53. Zone 53
  - 54. Zone 54
  - 55. Zone 55
  - 56. Zone 56
  - 57. Zone 57
  - 58. Zone 58
  - 59. Zone 59
  - 60. Zone 60
  - 61. Zone 61
  - 62. Zone 62
  - 63. Zone 63
  - 64. Zone 64
  - 65. Zone 65
  - 66. Zone 66
  - 67. Zone 67
  - 68. Zone 68
  - 69. Zone 69
  - 70. Zone 70
  - 71. Zone 71
  - 72. Zone 72
  - 73. Zone 73
  - 74. Zone 74
  - 75. Zone 75
  - 76. Zone 76
  - 77. Zone 77
  - 78. Zone 78
  - 79. Zone 79
  - 80. Zone 80
  - 81. Zone 81
  - 82. Zone 82
  - 83. Zone 83
  - 84. Zone 84
  - 85. Zone 85
  - 86. Zone 86
  - 87. Zone 87
  - 88. Zone 88
  - 89. Zone 89
  - 90. Zone 90
  - 91. Zone 91
  - 92. Zone 92
  - 93. Zone 93
  - 94. Zone 94
  - 95. Zone 95
  - 96. Zone 96
  - 97. Zone 97
  - 98. Zone 98
  - 99. Zone 99
  - 100. Zone 100



# London Underground





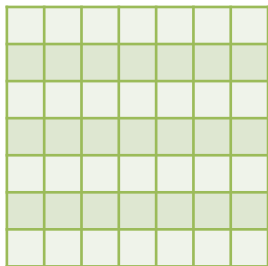
# Flow map



Werk  
gemeente (355)



Thuis  
gemeente (355)

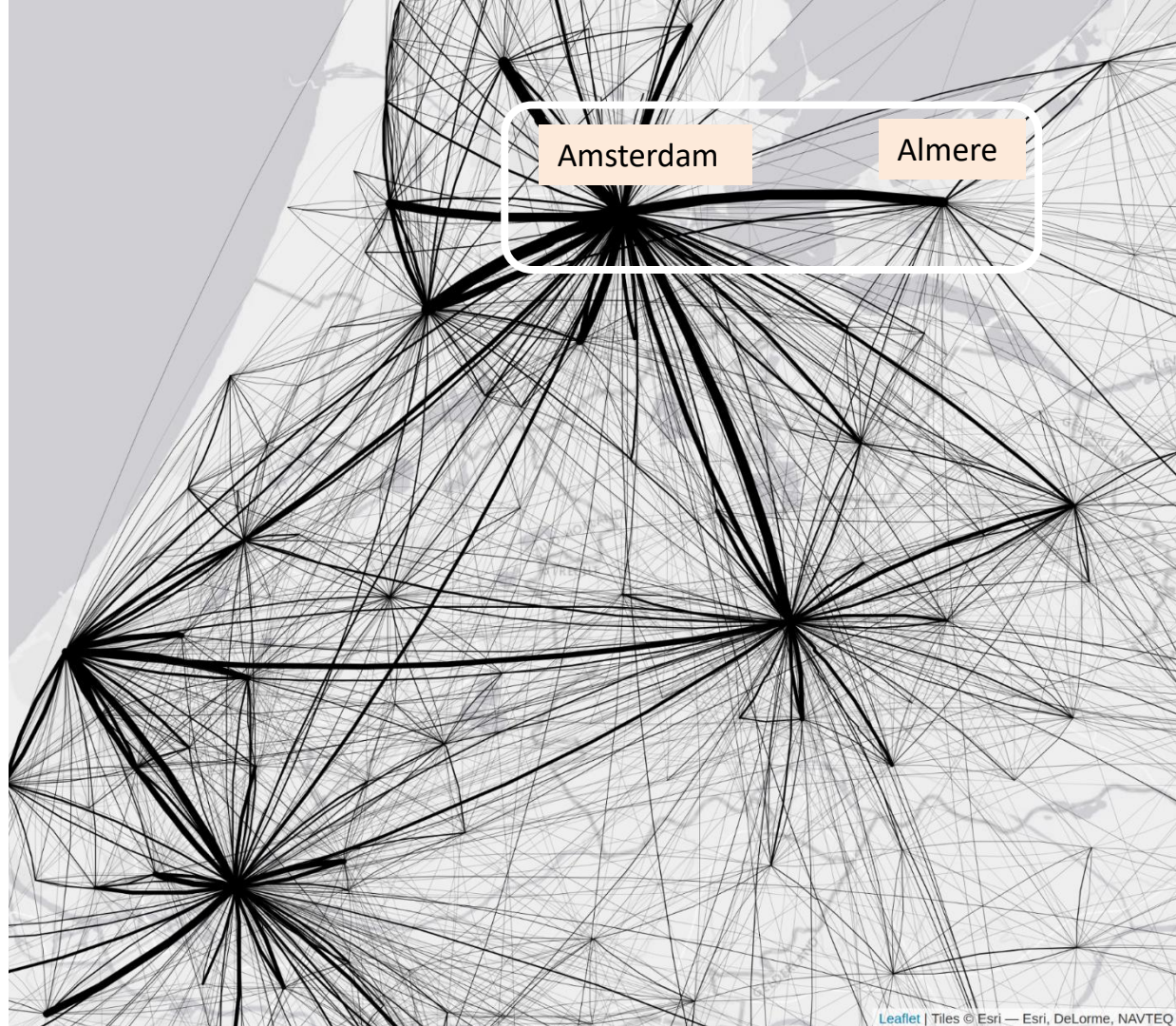


Richting correspondeert met “naar het werk gaan”

Amsterdam



Almere





# Donutkaart



Commuting

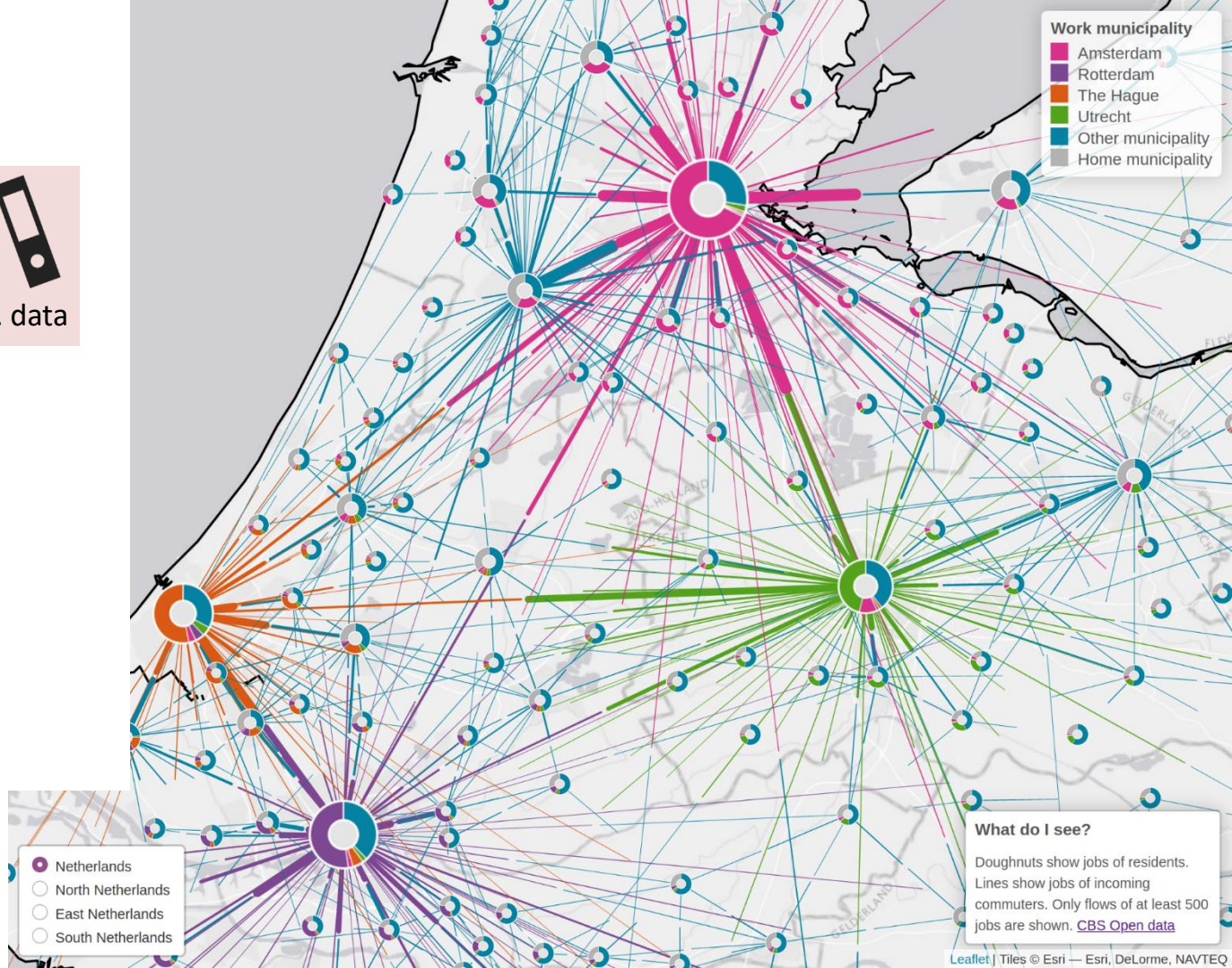


Admin. data

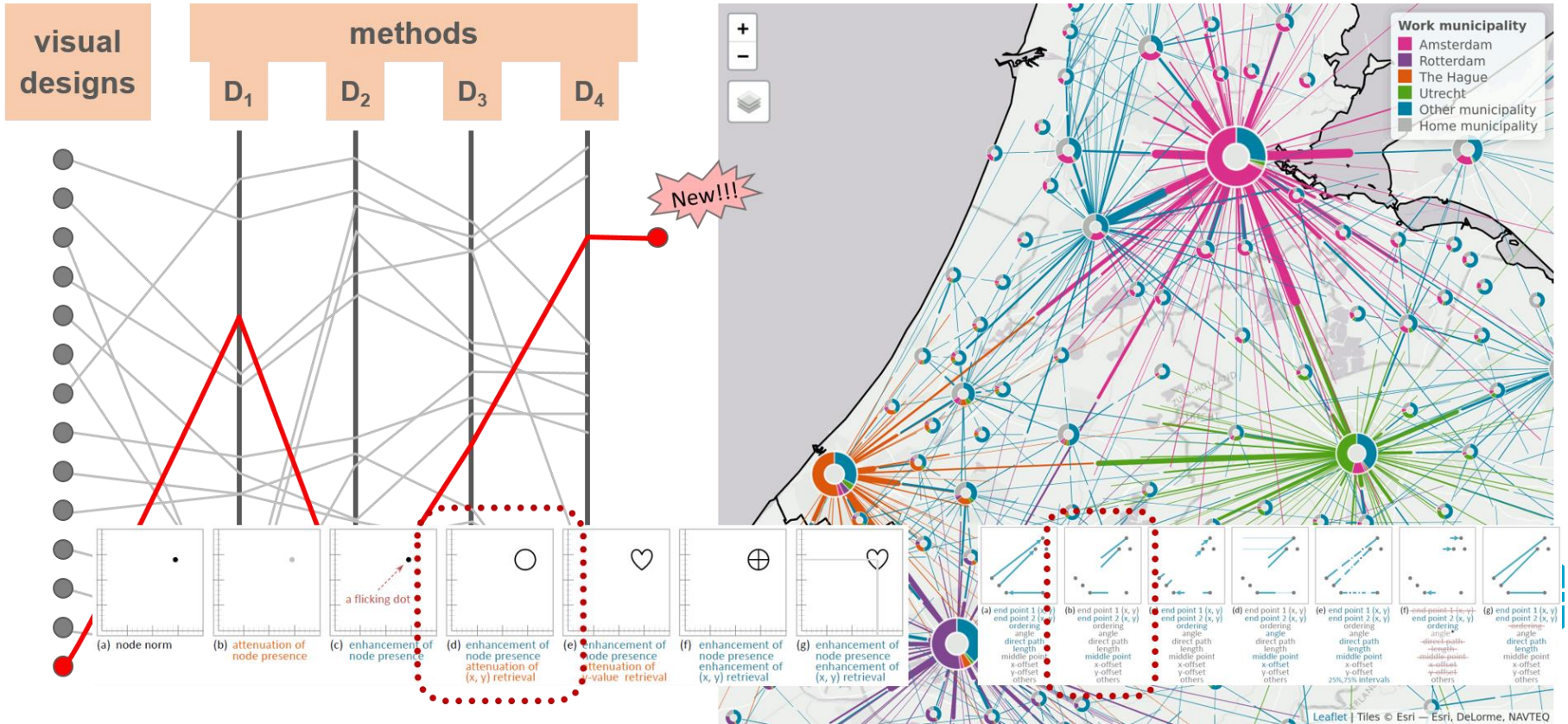


[donutkaart](#)

“Less is More” principe  
Door *minder* informatie te  
weer te geven, wordt er *meer*  
informatie waargenomen

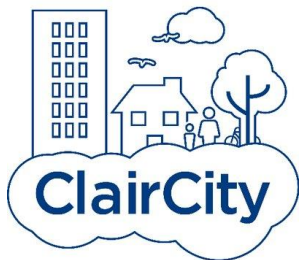


# Informatietheorie toegepast

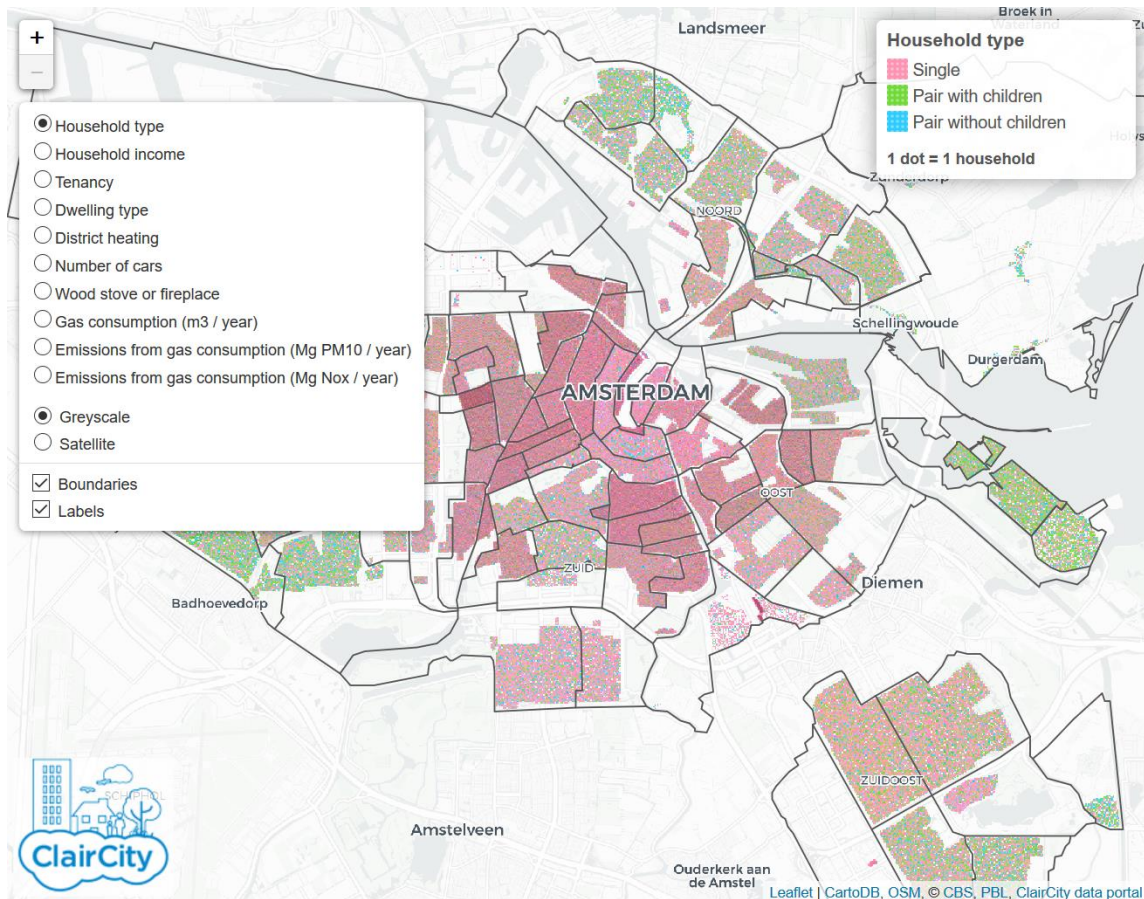




# Dot map



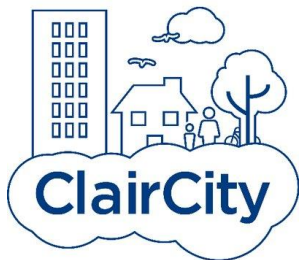
- Gesimuleerde data op wijkniveau voor Amsterdam
- Elke stip representeert een huishouden
- Stippen zijn in residentiële gebieden (OpenStreetMap) geplaatst



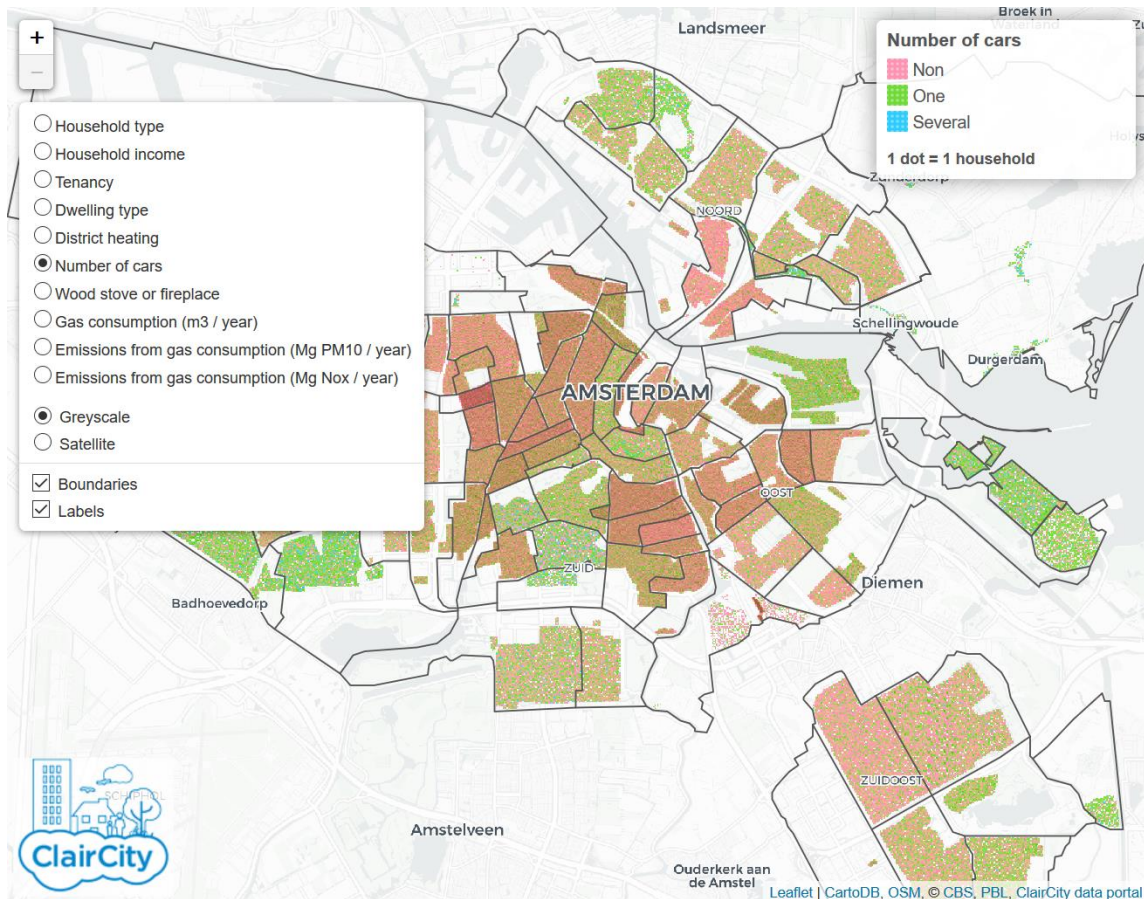
<https://claircitydata.cbs.nl/pages/dotmaps>



# Dot map

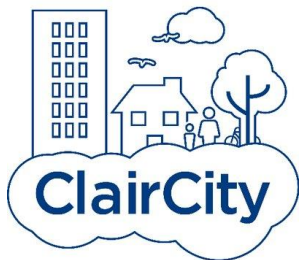


- Gesimuleerde data op wijkniveau voor Amsterdam
- Elke stip representeert een huishouden
- Stippen zijn in residentiële gebieden (OpenStreetMap) geplaatst

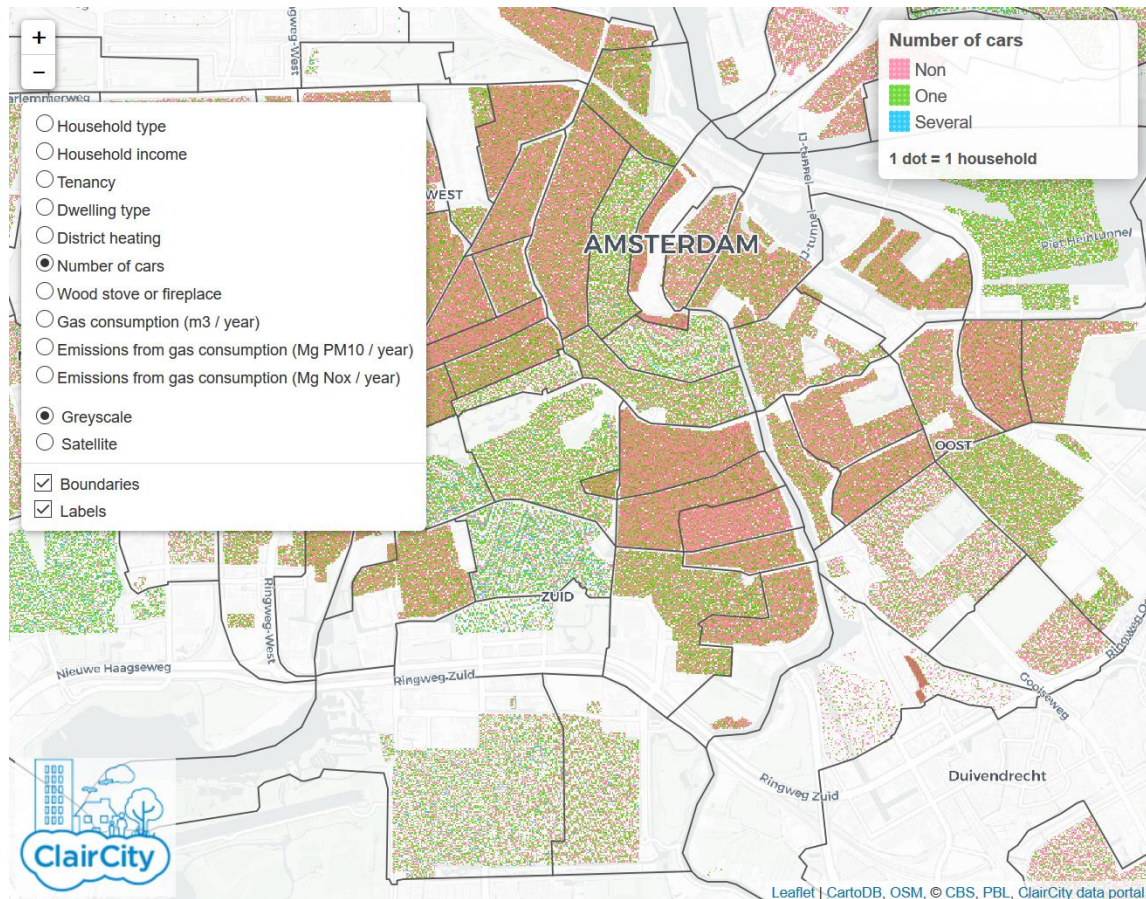




# Dot map



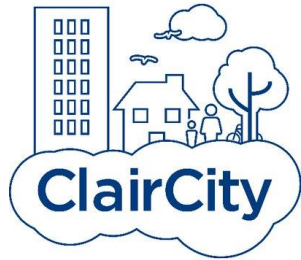
- Gesimuleerde data op wijkniveau voor Amsterdam
- Elke stip representeert een huishouden
- Stippen zijn in residentiële gebieden (OpenStreetMap) geplaatst



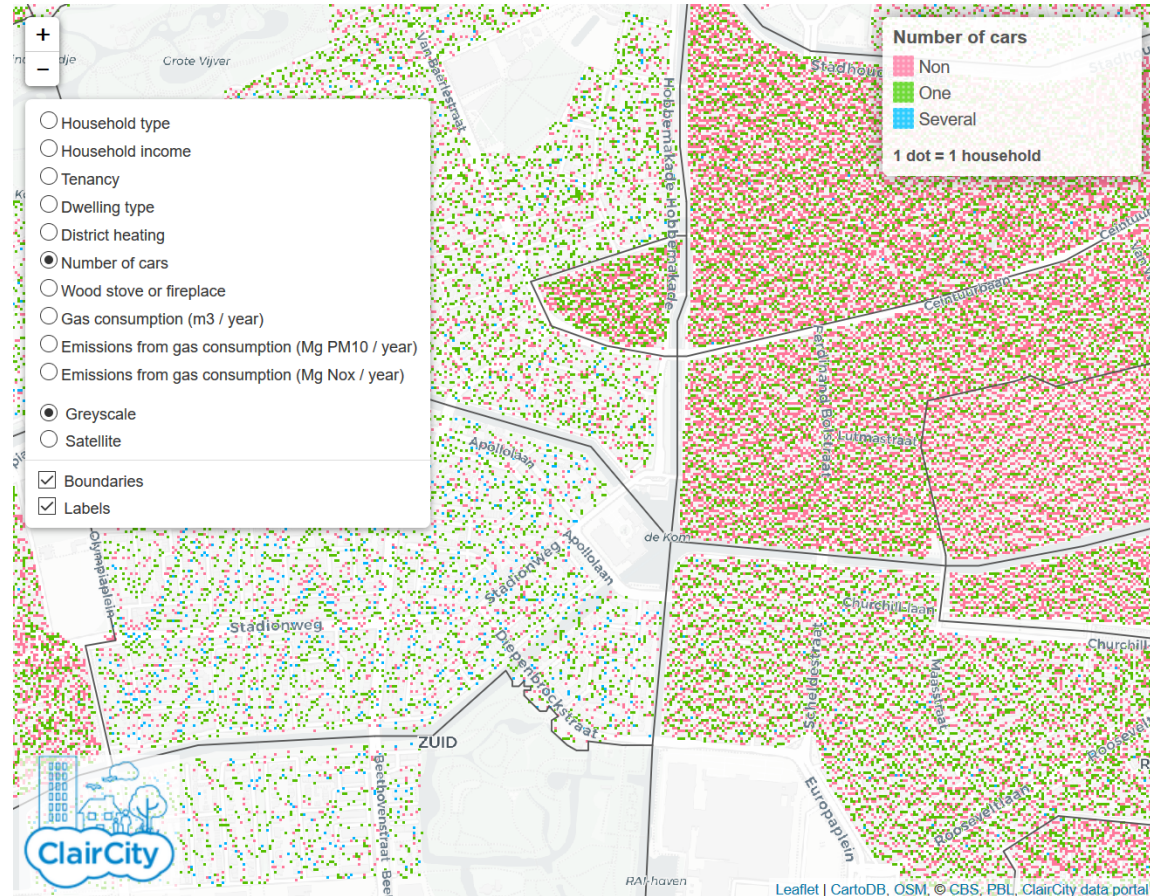
<https://claircitydata.cbs.nl/pages/dotmaps>



# Dot map



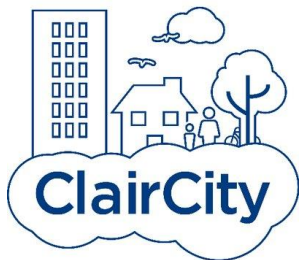
- Gesimuleerde data op wijkniveau voor Amsterdam
- Elke stip representeert een huishouden
- Stippen zijn in residentiële gebieden (OpenStreetMap) geplaatst



<https://claircitydata.cbs.nl/pages/dotmaps>



# Dot map



- Gesimuleerde data op wijkniveau voor Amsterdam
- Elke stip representeert een huishouden
- Stippen zijn in residentiële gebieden (OpenStreetMap) geplaatst



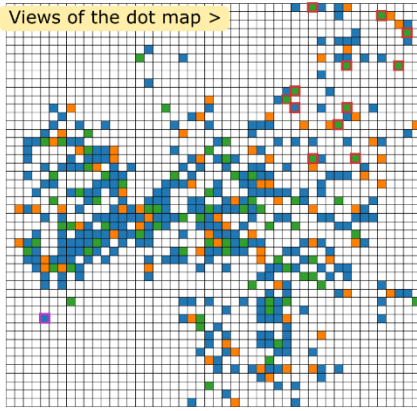
<https://claircitydata.cbs.nl/pages/dotmaps>





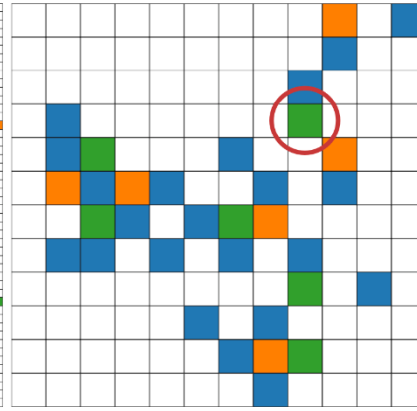
# Stippen aggregeren

Views of the dot map >



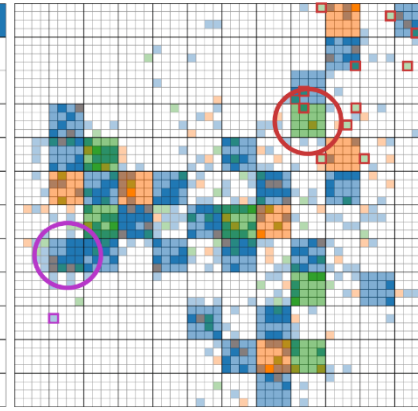
CLASS BALANCE :: 11684.0

OUTPUT



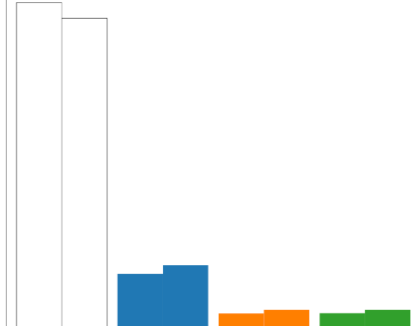
REPRESENTATION :: 7802.25

OVERLAY

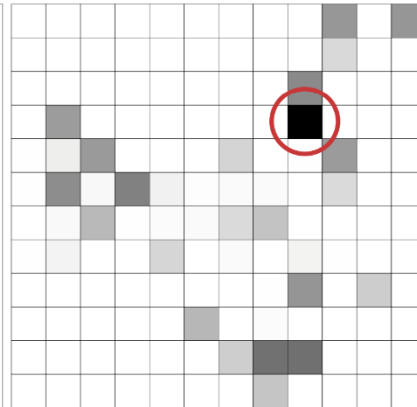


PRESENCE :: 325.0

Configurable small multiples to explore our three measures >



Class Balance

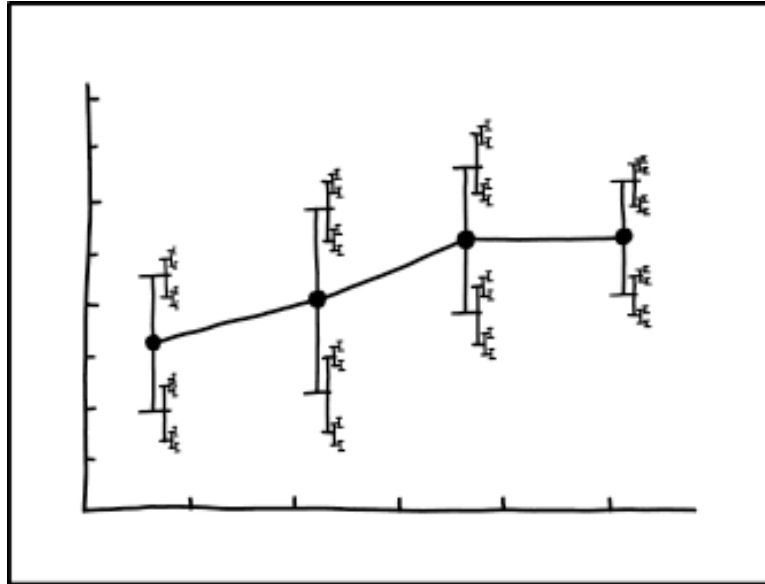


Representation



Presence

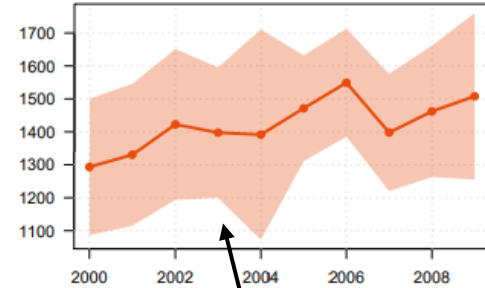
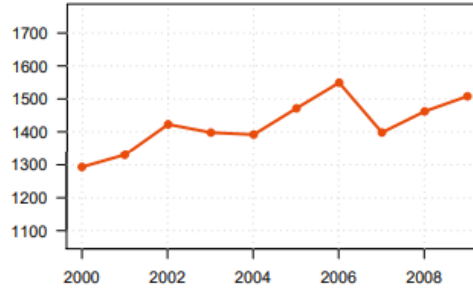
# Onzekerheid



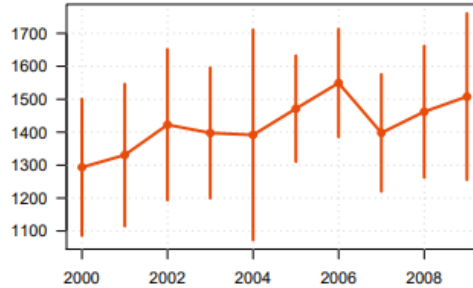
I DON'T KNOW HOW TO PROPAGATE  
ERROR CORRECTLY, SO I JUST PUT  
ERROR BARS ON ALL MY ERROR BARS.



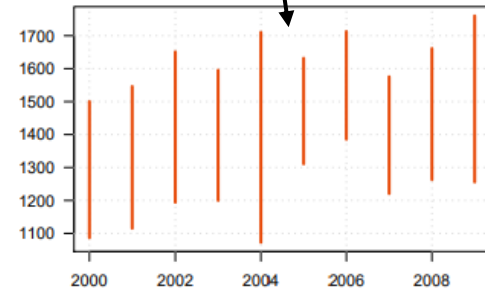
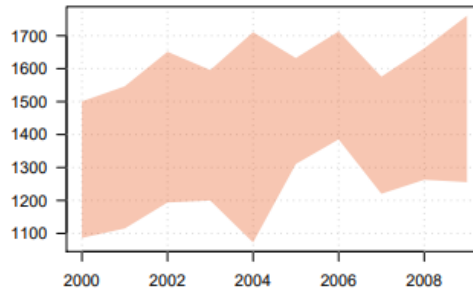
# Lijngrafieken met onzekerheid



Beste methode als  
puntschatting bekend is

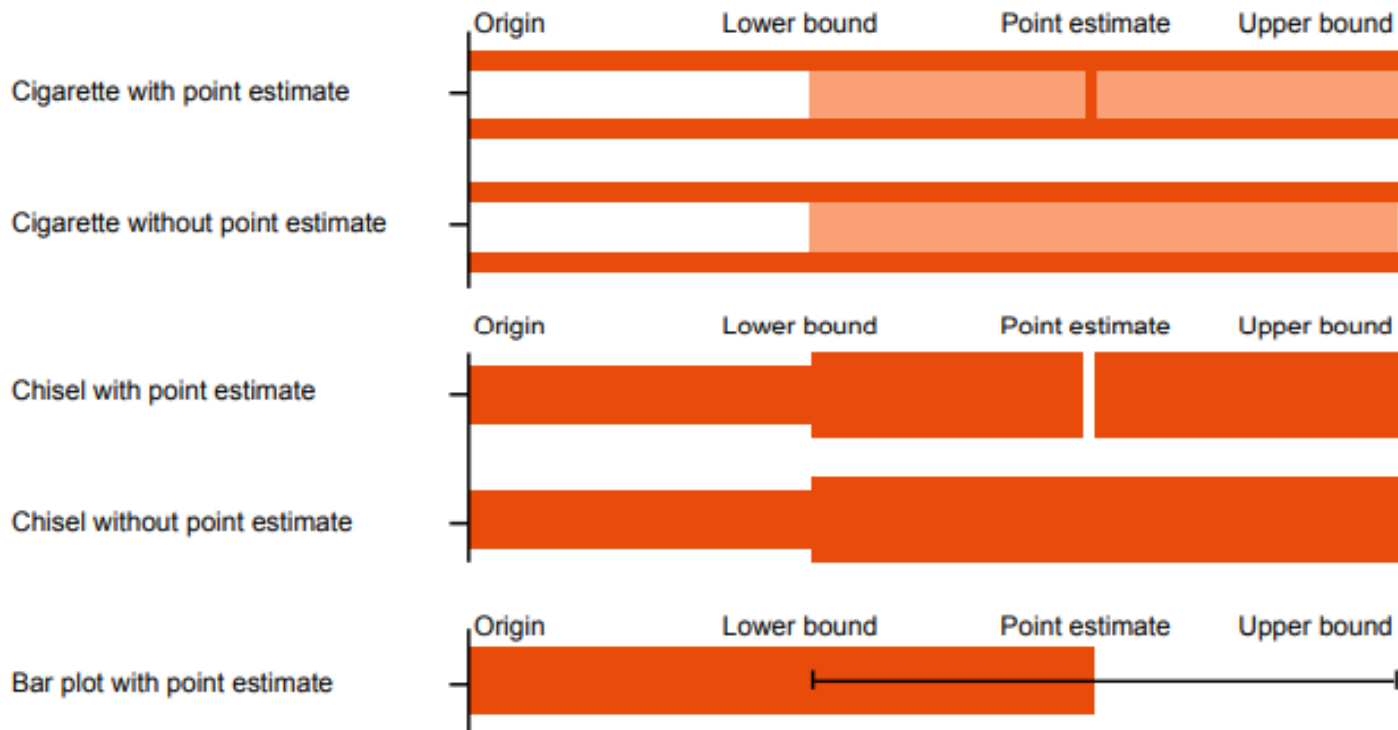


Beste methode als  
puntschatting onbekend is



Laan, D.J. & de Jonge, Edwin & Solcer, Jessica. (2015). Effect of Displaying Uncertainty in Line and Bar Charts - Presentation and Interpretation. IVAPP 2015 - 6th International Conference on Information Visualization Theory and Applications

# Staven met onzekerheid



- Geen duidelijke winnaar
- Gebruikers CBS data waarderen de toevoeging van onzekerheidsvisualisatie!

# Kleur voor onzekerheden



## Palette type

- Categorical
- Sequential
- Diverging
- Bivariate

## Subtype

Sequential x desaturated

## Palette series

- brewer
- c4a
- carto
- hcl
- kovesi
- met
- misc
- parks
- poly
- scico
- seaborn
- stevens
- tableau
- tol
- viridis
- wes

Color for missing values

Tritan (blue-yellow)

## Text format

- Hex
- RGB
- HCL

## Text color

- Hide text
- Black
- White
- Automatic

## Sort

Name

- Reverse sorting
- Show scores

All None Overview

Series	Name	Max number	Colorblind-friendly	Vivid	Fair	Hue type	Low contrast	3D Blues	1	2	3	4	5	References
c4a	br_bivg	Inf	☺	✳										
c4a	bu_bivg	Inf	☺											
c4a	gn_bivg	Inf	☺	✳										
c4a	pu_bivg	Inf	☺	✳										

# Conclusie

Datavisualisatie is een **interdisciplinair** vakgebied  
Veel **inzicht** wordt verkregen door vakgebieden te **combineren**

**Gebruikerstesten** zijn essentieel!