

Visualization

Presenter: *Mojca Bavdaž (mojca.bavdaz@ef.uni-lj.si)*

Guest: *Jorge Camoes (excelcharts.com)*



EMOS Webinar, 17 May 2017

Outline



**Classic
examples**

**Visual
processing**

**Representation
& perception**

**Official
statistics**



16:30

Break

18:00



Classic examples of visualization



Bar chart & time-series line chart: Price of wheat & wages

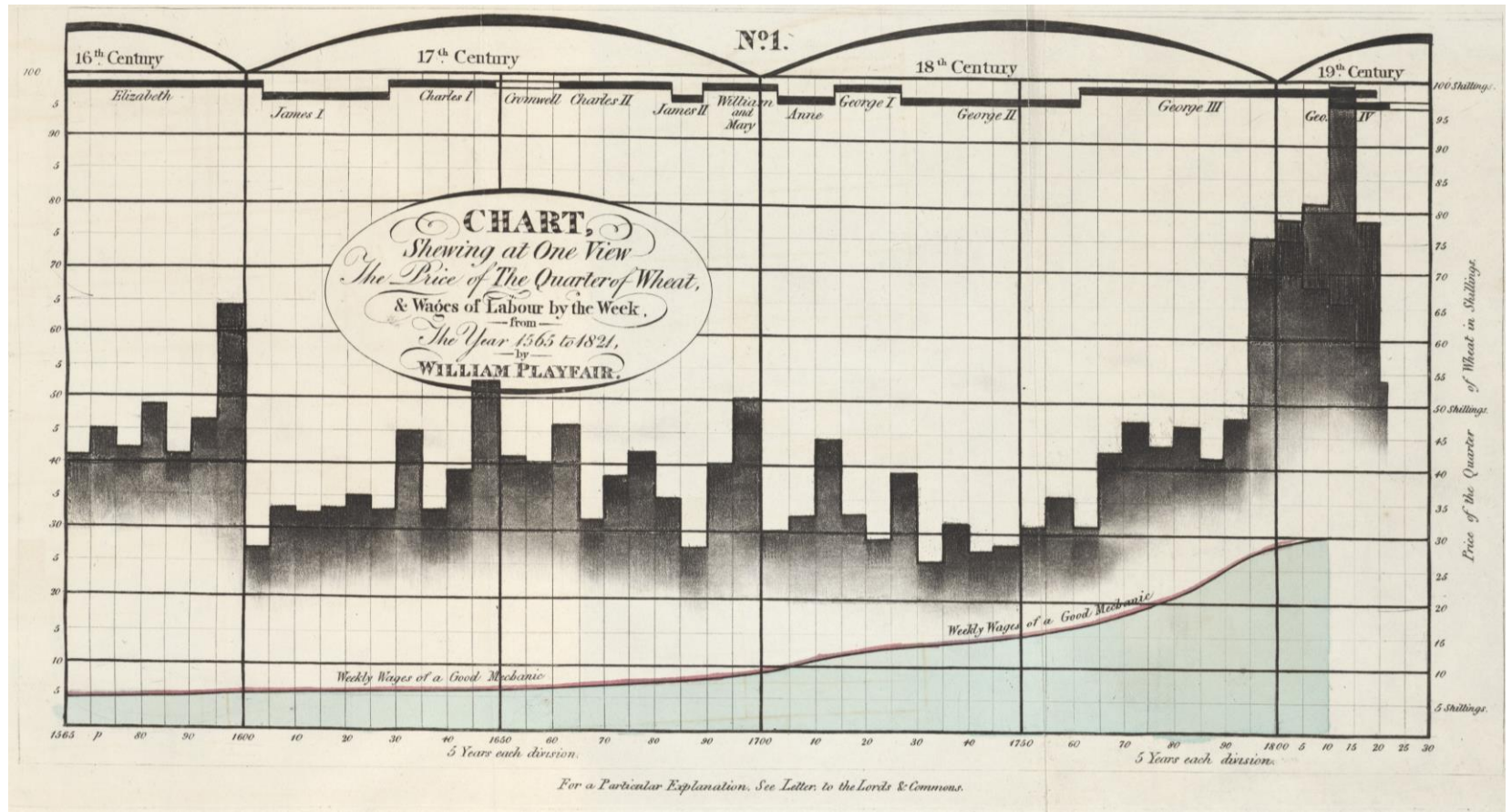
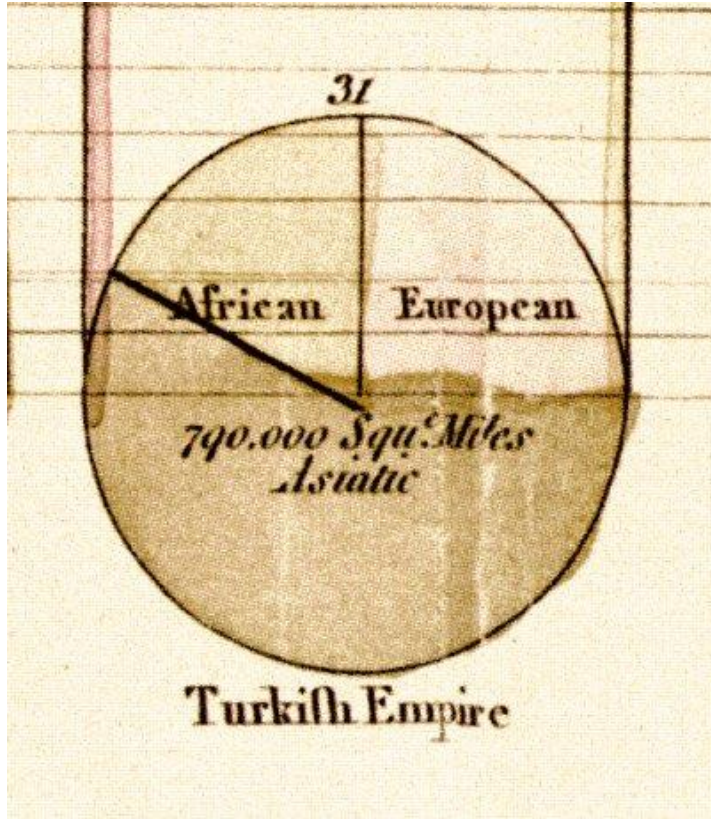


Chart shewing at one view the price of the quarter of wheat, & wages of labour by the week, from the year 1565 to 1821 by William Playfair (1822). (The second edition / with an additional chart.). Retrieved from <http://brbl-dl.library.yale.edu/vufind/Record/3566707>

Pie chart: Turkish Empire



The pie chart Turkish Empire by William Playfair (1801). In *Statistical Breviary*. Retrieved from <https://commons.wikimedia.org/wiki/File:Playfair-piechart.jpg>

Disease dot map: Cholera in London



Variant of cholera map by John Snow (1854). In *Wikimedia Commons*, Retrieved from <https://commons.wikimedia.org/wiki/File:Snow-cholera-map.jpg>

Polar area diagram: Causes of mortality in Crimean War

2.
APRIL 1855 TO MARCH 1856.

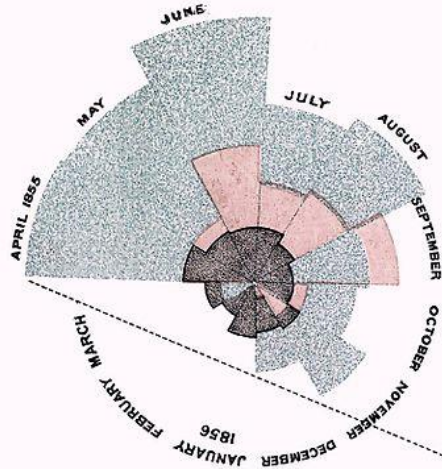
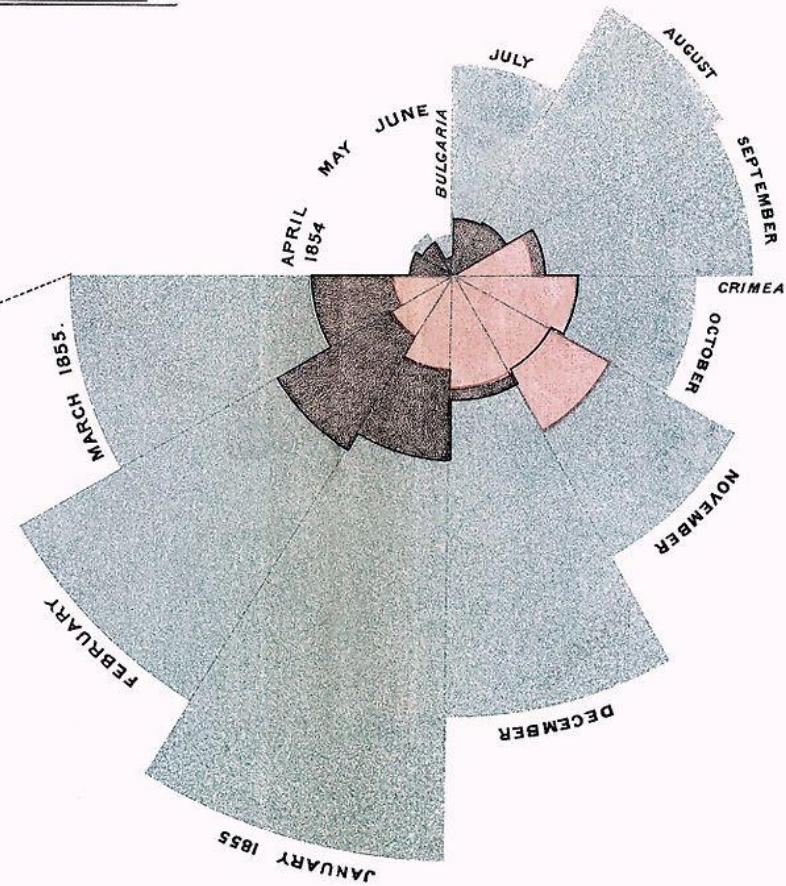


DIAGRAM OF THE CAUSES OF MORTALITY
IN THE ARMY IN THE EAST.

1.
APRIL 1854 TO MARCH 1855.



The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.

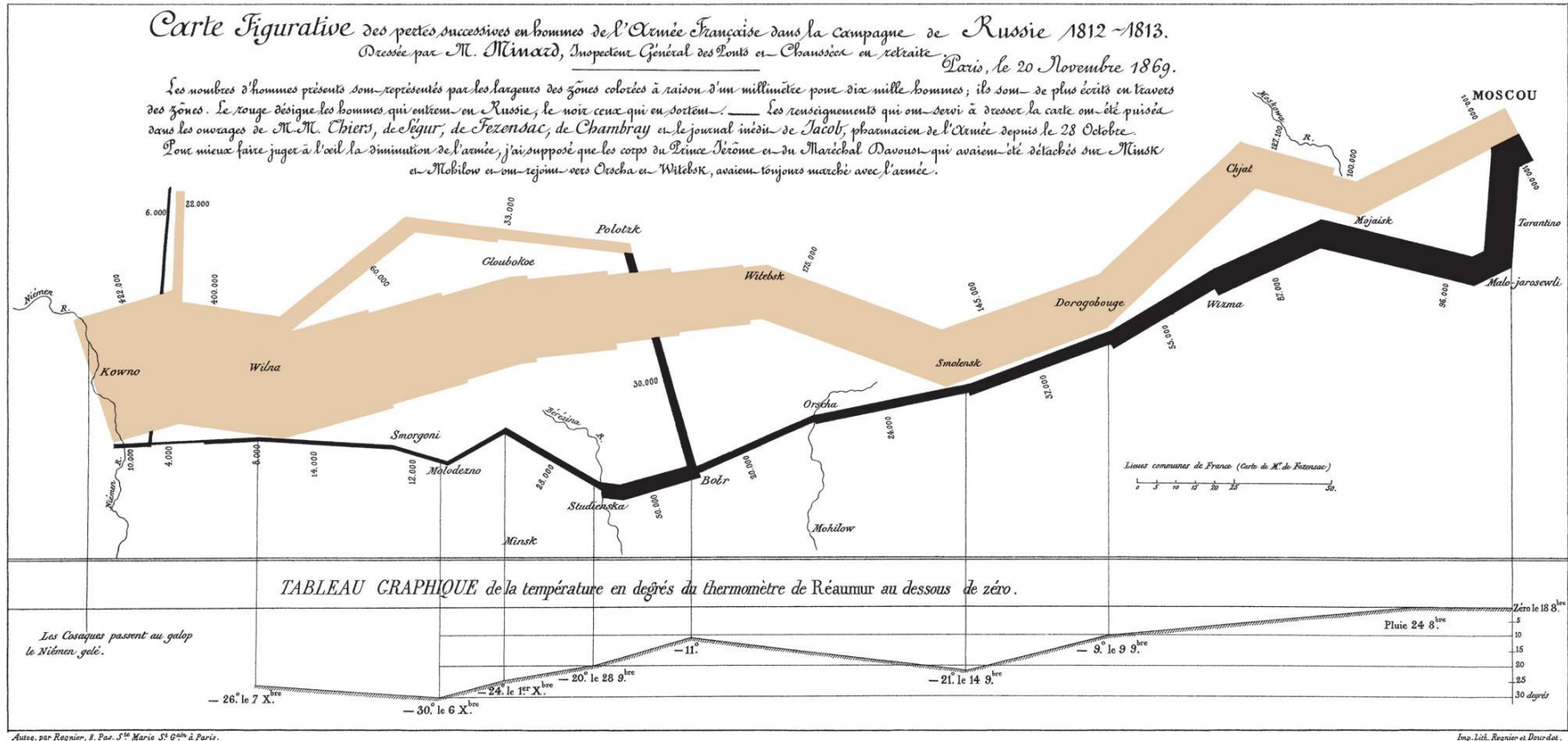
The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases, the red wedges measured from the centre the deaths from wounds, & the black wedges measured from the centre the deaths from all other causes.

The black line across the red triangle in Nov: 1854 marks the boundary of the deaths from all other causes during the month.

In October 1854, & April 1855, the black area coincides with the red; in January & February 1856, the blue coincides with the black.

The entire areas may be compared by following the blue, the red & the black lines enclosing them.

Flow map: Napoleon's March on Moscow



Role of design: London Underground map



By Henry C. Beck (1933). In *Wikipedia*, Retrieved from https://en.wikipedia.org/wiki/File:Beck_Map_1933.jpg

Long list of other contributors

Baron Pierre Charles Dupin (1826): Choropleth map

Charles Joseph Minard (1861): Map with diagrams

Francis Amasa Walker (1874): Population pyramid

Michael George Mulhall (1884): Pictograms

Karl Pearson (1892)(?): Histogram

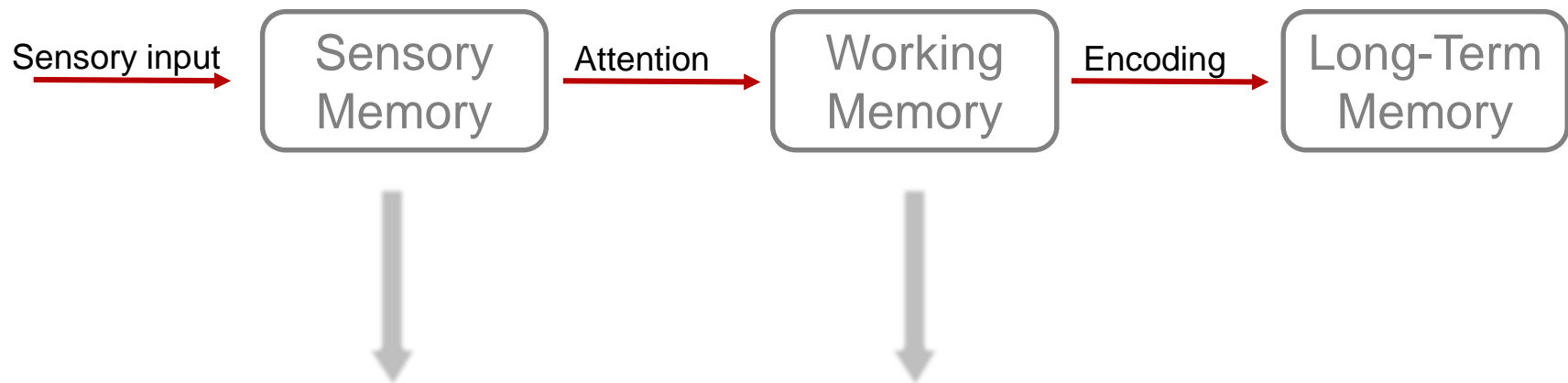
John W. Tukey (1969): Stem-and-leaf, box-and-whisker plots...



Basics of visual processing



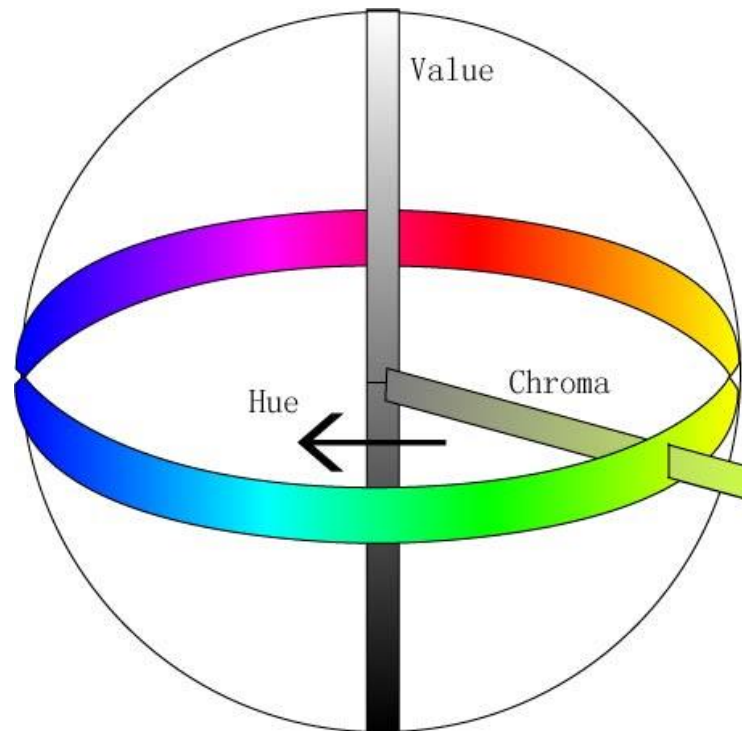
Role of memory



Preattentive attributes



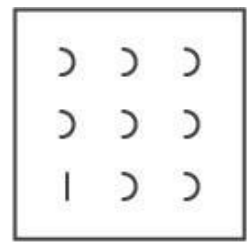
Preattentive attributes: Color



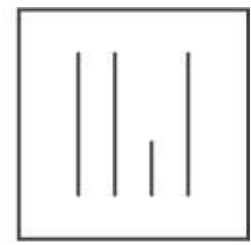
Preattentive attributes: Form



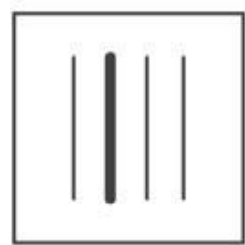
Orientation and collinearity



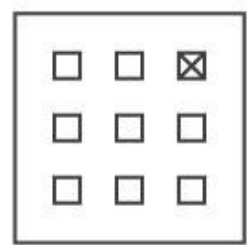
Curvature



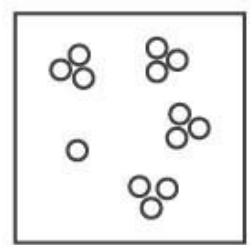
Length



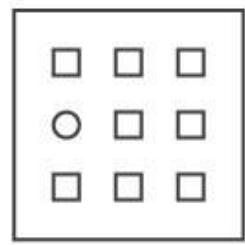
Width



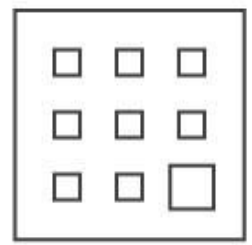
Added marks



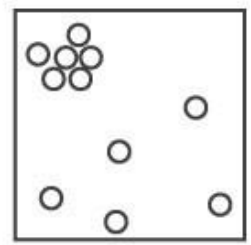
Numerosity



Shapes



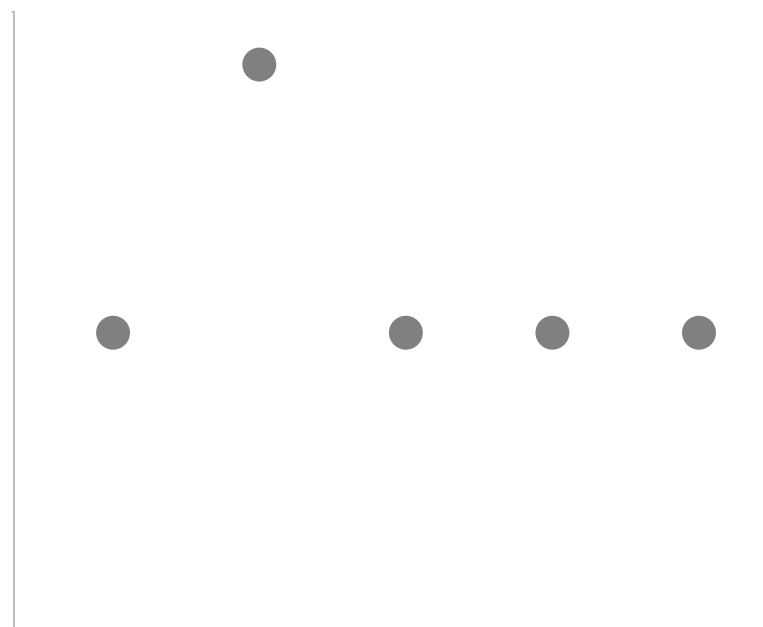
Size



Spatial grouping

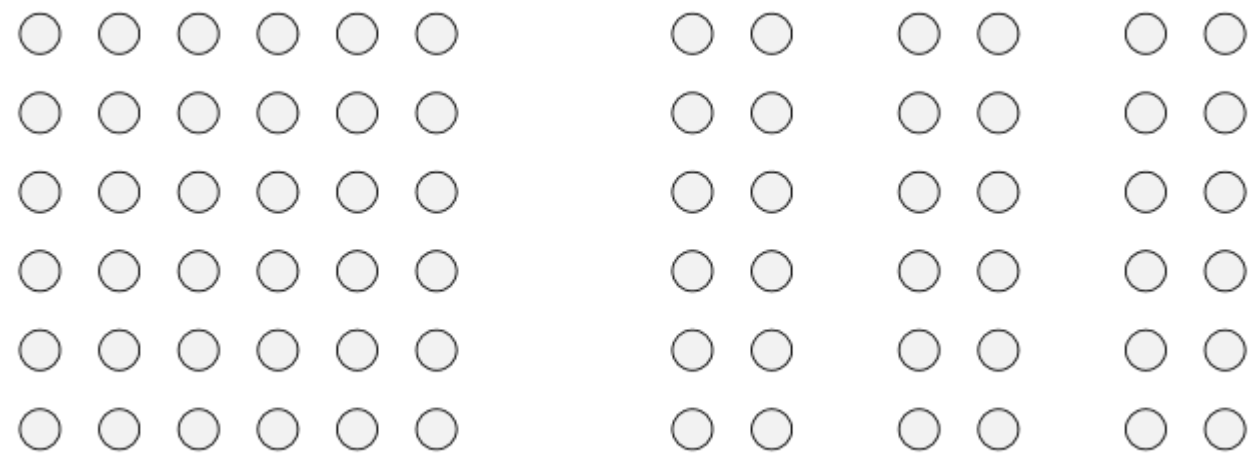


Preattentive attributes: Spatial positioning



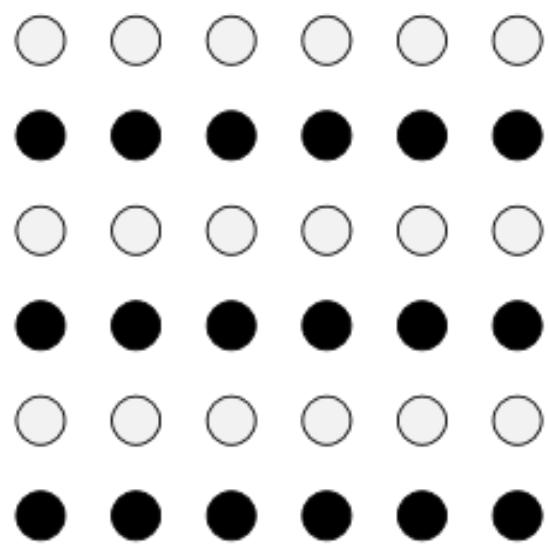
Gestalt principles

Gestalt principles: Proximity



Gestalt Proximity (2008). In *Wikimedia Commons*, Retrieved from https://commons.wikimedia.org/wiki/File:Gestalt_proximity.svg

Gestalt principles: Similarity



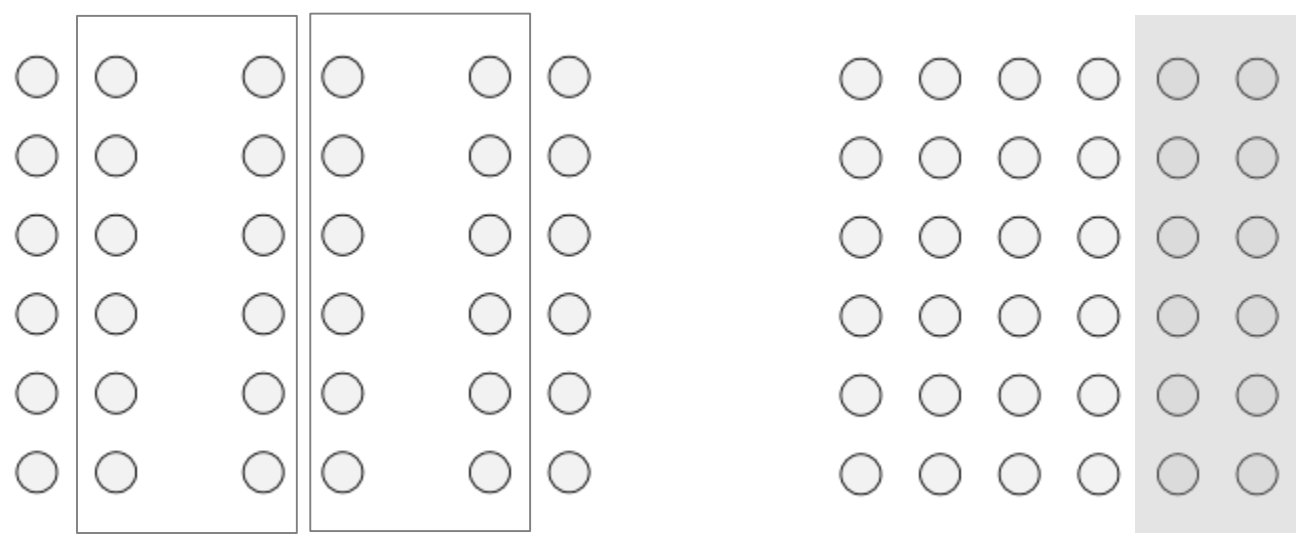
Gestalt Similarity (2008). In *Wikimedia Commons*, Retrieved from https://commons.wikimedia.org/wiki/File:Gestalt_similarity.svg

Gestalt principles: Closure

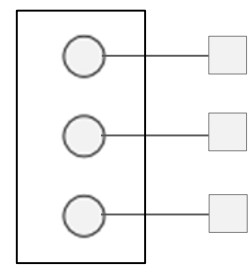
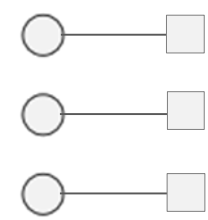
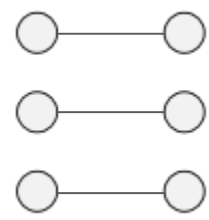
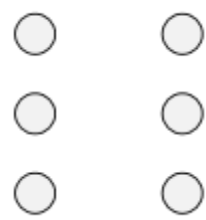


Gestalt Closure (2008). In *Wikimedia Commons*, Retrieved from https://commons.wikimedia.org/wiki/File:Gestalt_closure.svg

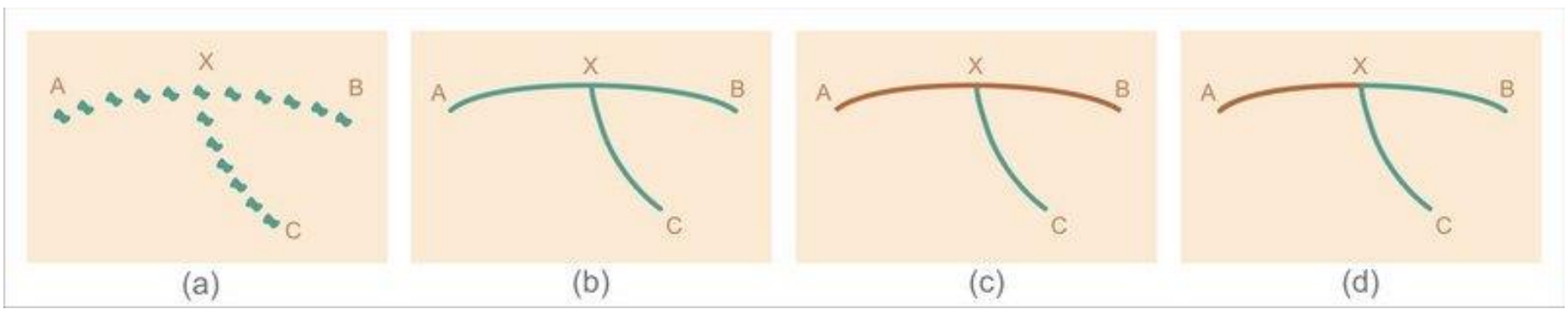
Gestalt principles: Common region/Enclosure



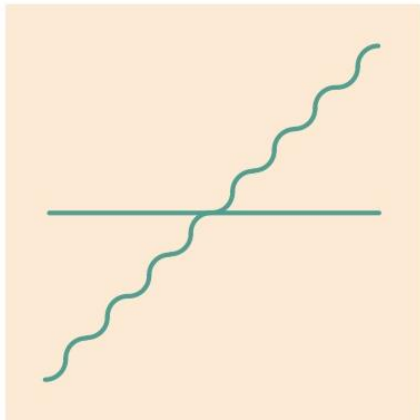
Gestalt principles: Connectedness



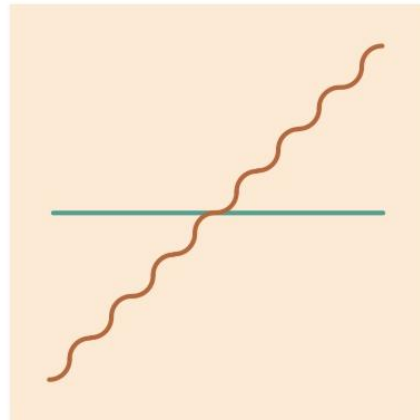
Gestalt principles: Continuity



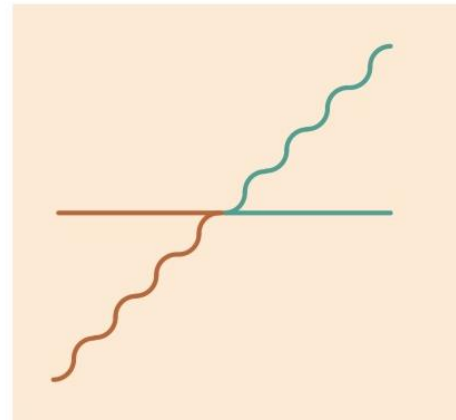
Gestalt principles: Good Gestalt



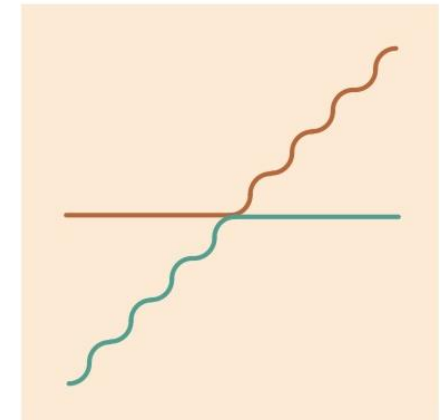
(a)



(b)

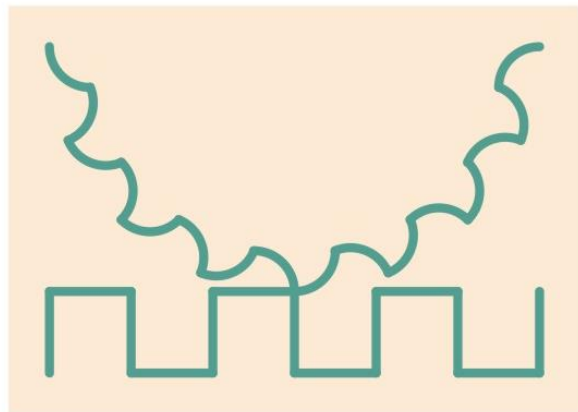


(c)

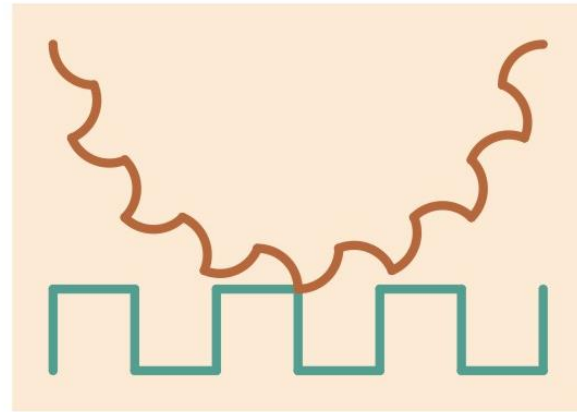


(d)

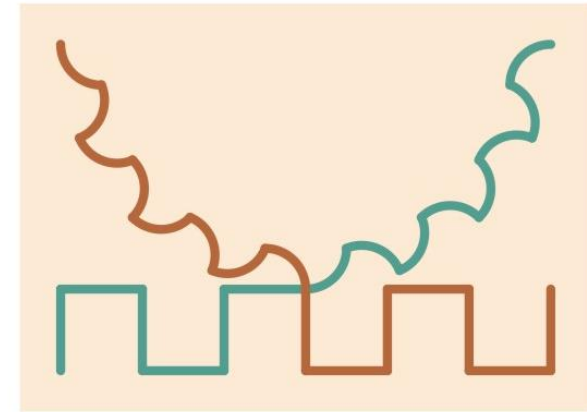
Gestalt principles: Good Gestalt



(a)



(b)



(c)



Visual representation of statistical data and perception

Theory of graphic symbols by Jacques Bertin



Basic visual units (marks): point, line, area.

Visual variables:

- position (position),
- size (taille),
- shape (forme),
- color (couleur),
- brightness (valeur),
- orientation (orientation),
- granularity (grain).

Other visual variables:

- hue
- saturation
- resolution
- crispness
- transparency
- motion



Hans Rosling: Gapminder.org & Moving bubble chart



Hans Rosling:
The best stats you've ever seen

TED2006 - 19:50 - Filmed Feb 2006

48 subtitle languages

View interactive transcript

11,701,358 Total views



Perception of graphical elements



Weber. If difference between two lines is relatively

- large \Rightarrow easy to detect
- small \Rightarrow hard to detect

Stevens. Bias in judgments

length < area < volume

Cleveland. Perception tasks of decoding quantitative variables from most to least accurate:

Position

Length

Angle, Slope

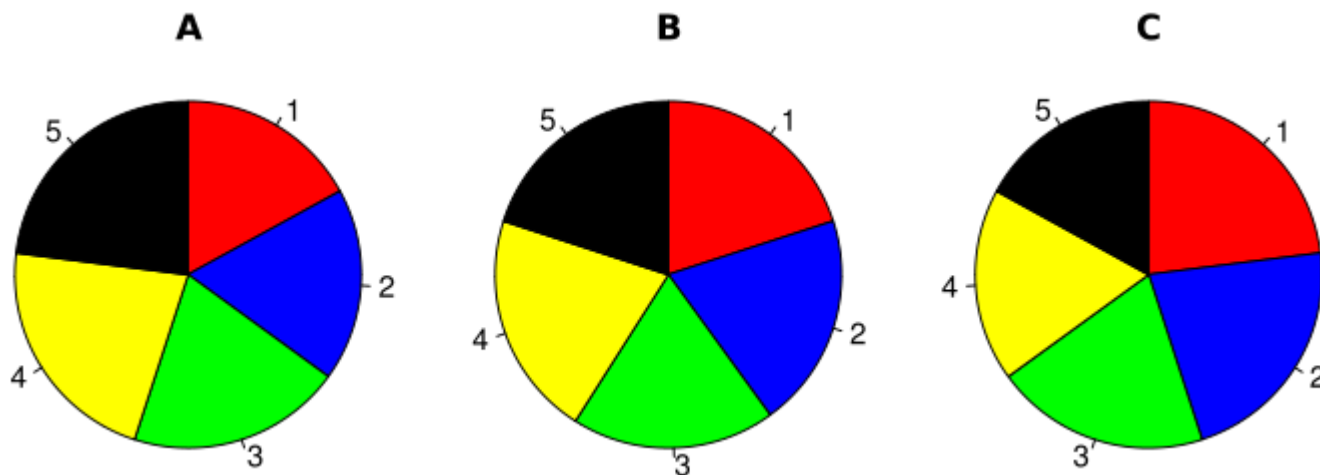
Area

Volume

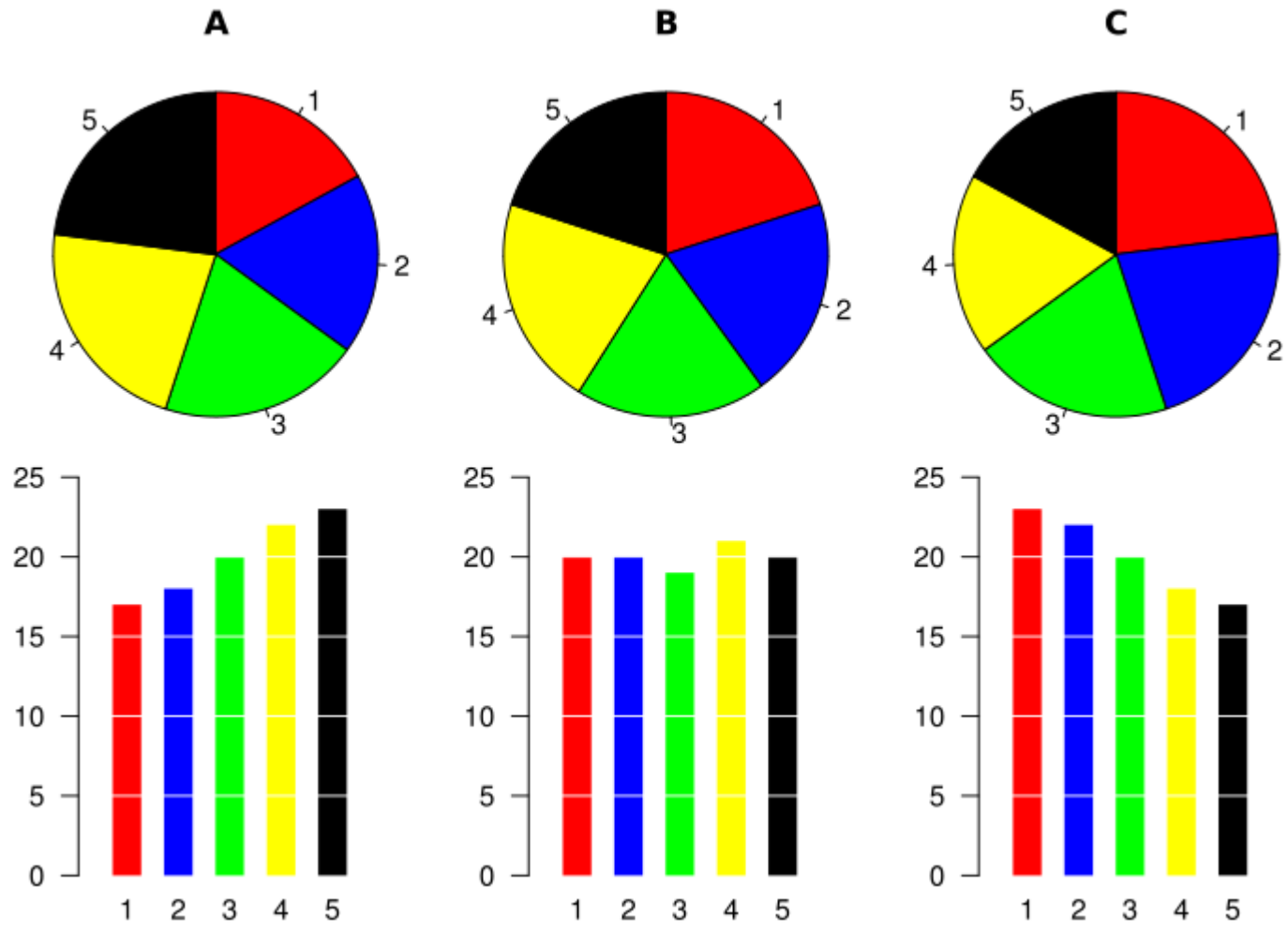
Color

A test 😊

I'll show you three pie charts.
Order the slices of each pie chart
by size from the largest to the smallest.



A test 😊



Pie charts (2007). In *Wikipedia*. Retrieved from http://en.wikipedia.org/wiki/Pie_chart

Theory of data graphics by Edward Tufte



- Above all else show the data.
- Maximize the data-ink ratio.

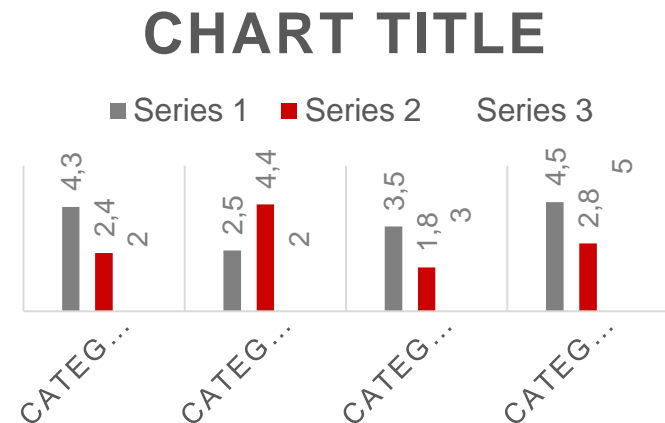
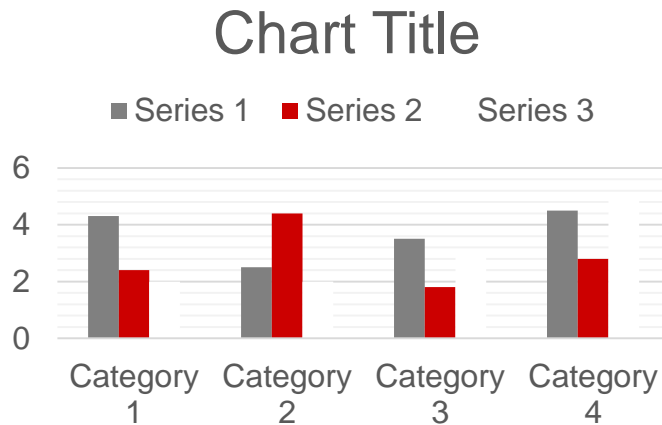
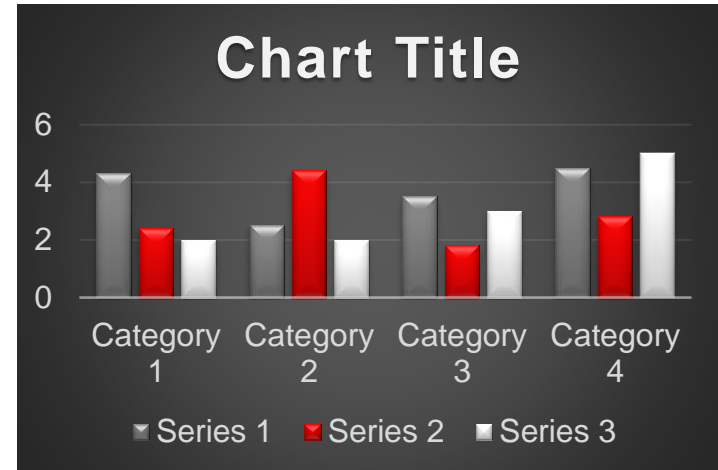
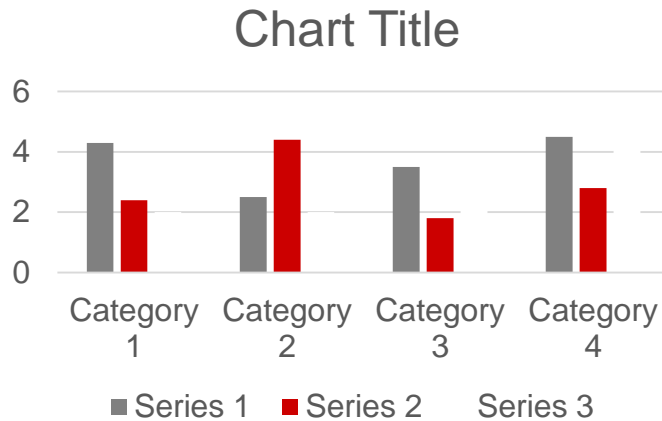
Data-ink ratio = Data ink / Total ink

= Share of ink devoted to non-redundant display of data-information
= 1 – Share of ink that can be erased without loss of data-information

- Erase non-data ink.
- Erase redundant data ink.
- Revise and edit.



MS PowerPoint 2016: Insert Chart: Clustered Column



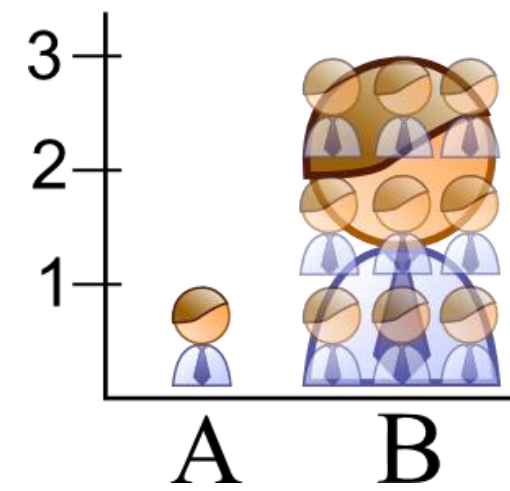
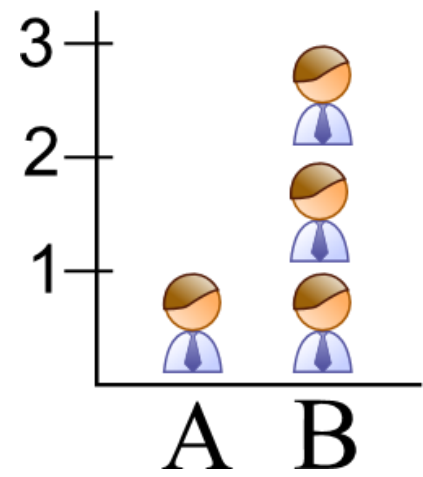
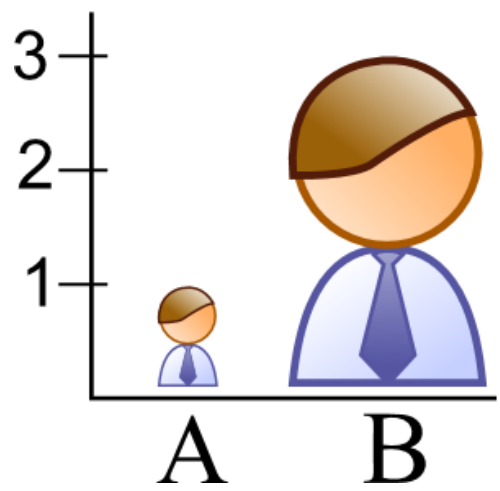
Graphical integrity (Edward Tufte)



- Clear, detailed, thorough labeling; explanations of the data on the graphic itself etc.
- Lie factor:

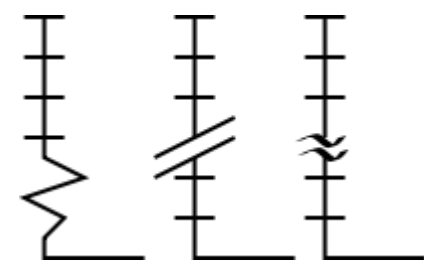
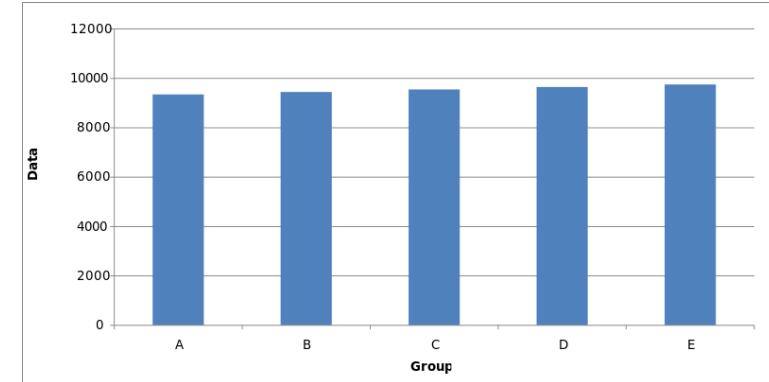
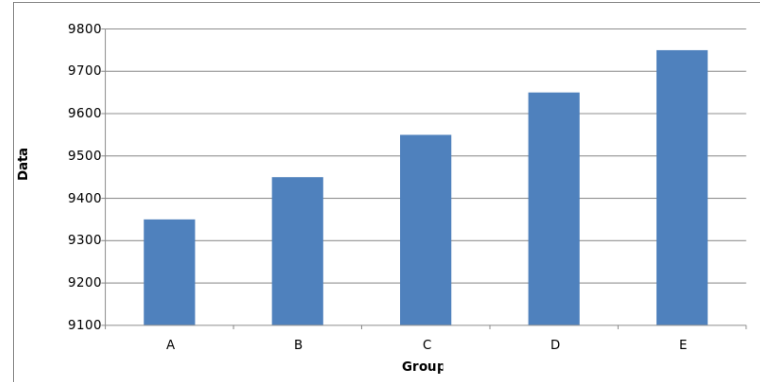
= Size of effect shown in graphics / Size of effect in data

Pictograms



Misleading graphs (2012). In *Wikimedia Commons*. Retrieved from https://commons.wikimedia.org/wiki/File:Improperly_scaled_picture_graph.svg
https://commons.wikimedia.org/wiki/File:Picture_Graph.svg
https://commons.wikimedia.org/wiki/File:Comparison_of_properly_and_improperly_scaled_picture_graph.svg

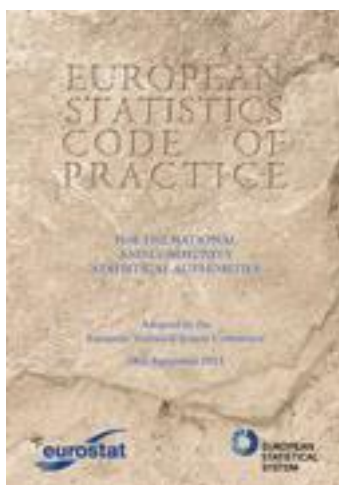
Truncated axes



Misleading graphs (2012). In *Wikimedia Commons*. Retrieved from https://commons.wikimedia.org/wiki/File:Truncated_Bar_Graph.svg
https://commons.wikimedia.org/wiki/File:Bar_graph.svg
https://commons.wikimedia.org/wiki/File:Y-axis_break.svg

Role of visualization in official statistics

European Statistics Code of Practice

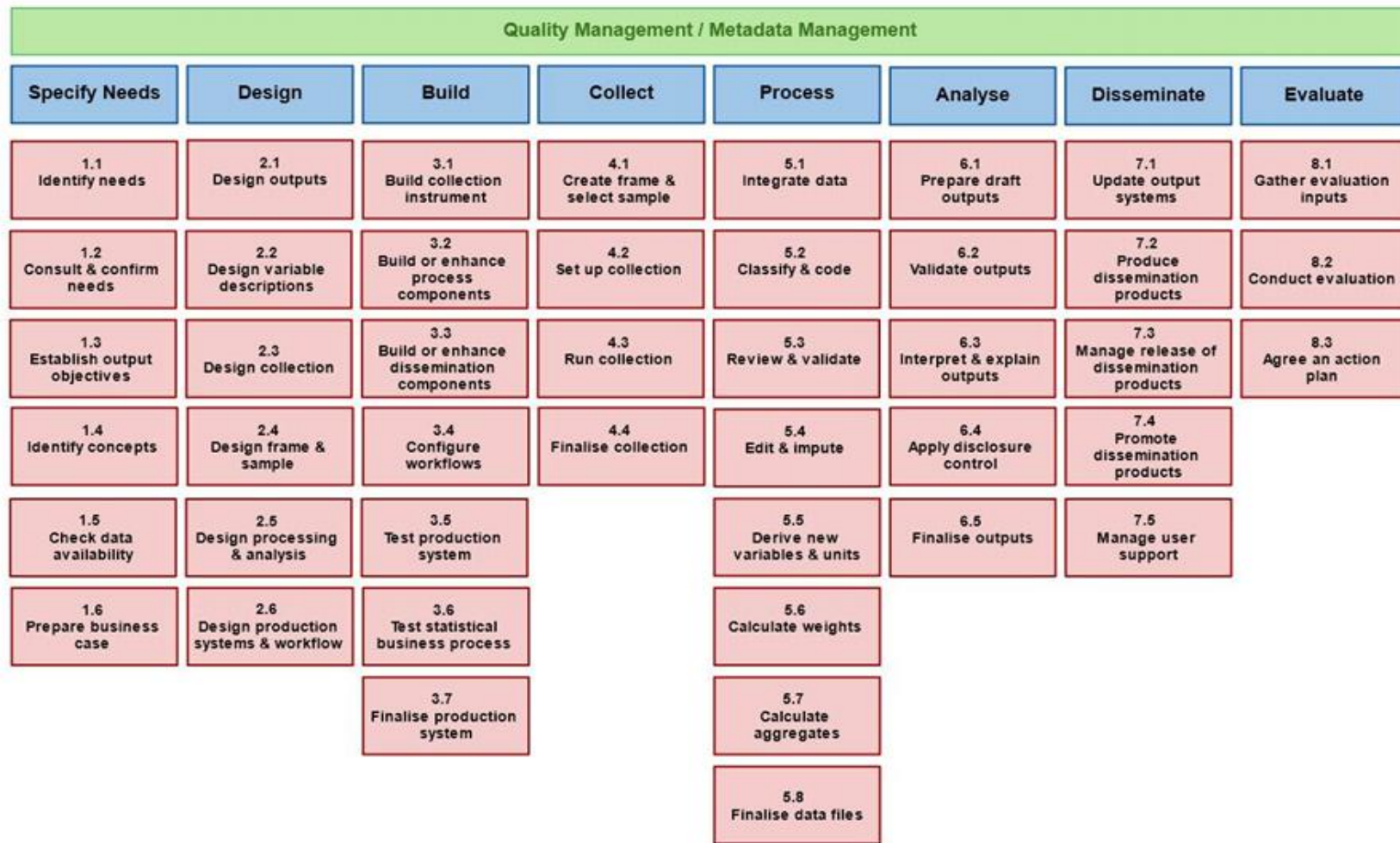


Principle 15: Accessibility and Clarity

European Statistics are **presented in a clear and understandable form**, released in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance.



Generic Statistical Business Process Model (GSBPM)



Questionnaire design

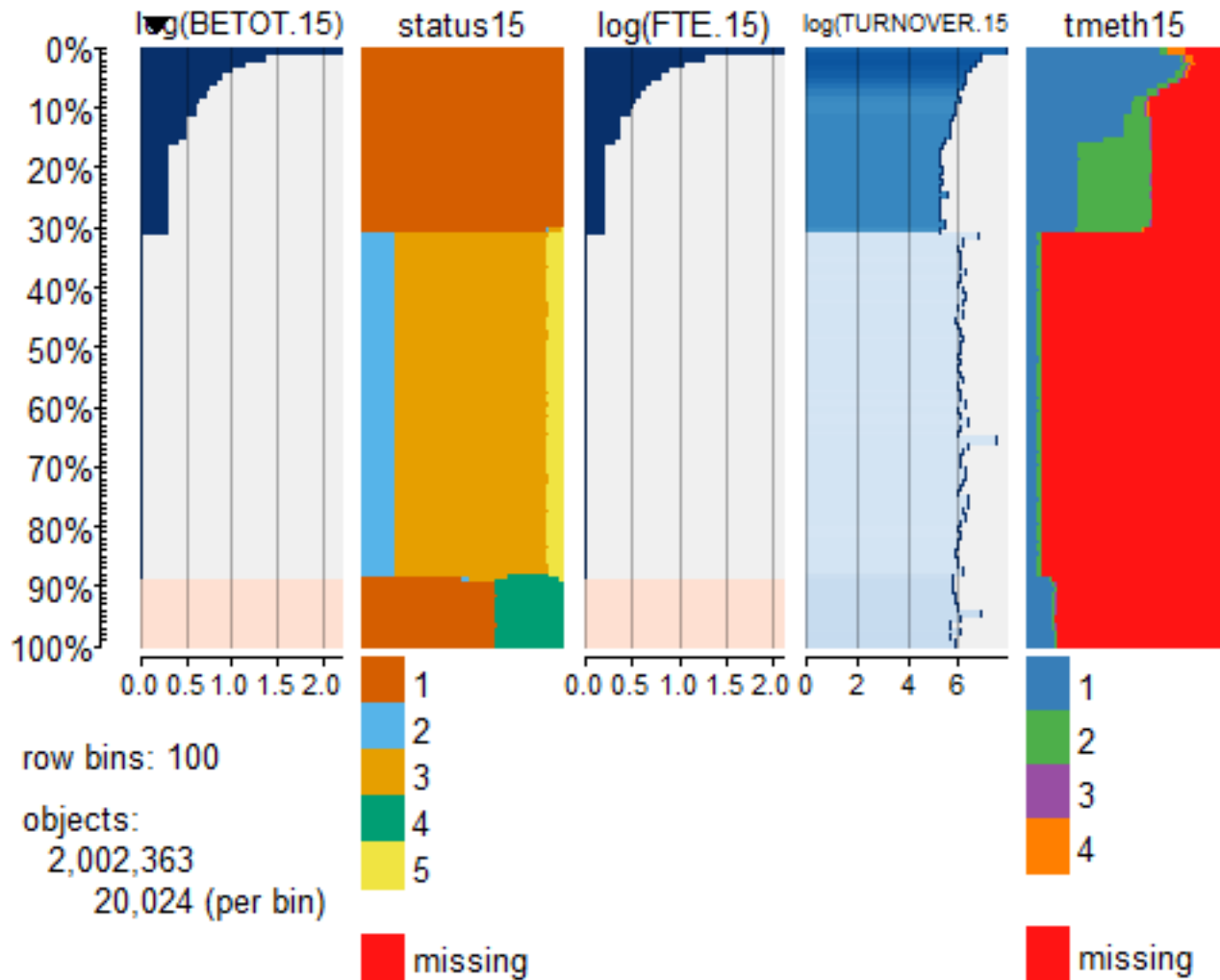
Jenkins & Dillman (1995) introduced into survey research design principles that emphasize visual presentation of information, e.g.:

- Present survey question + instruction in close proximity
- Deemphasize information that disturbs question-answering process (e.g. legal basis)
- Use visual elements consistently throughout the questionnaire

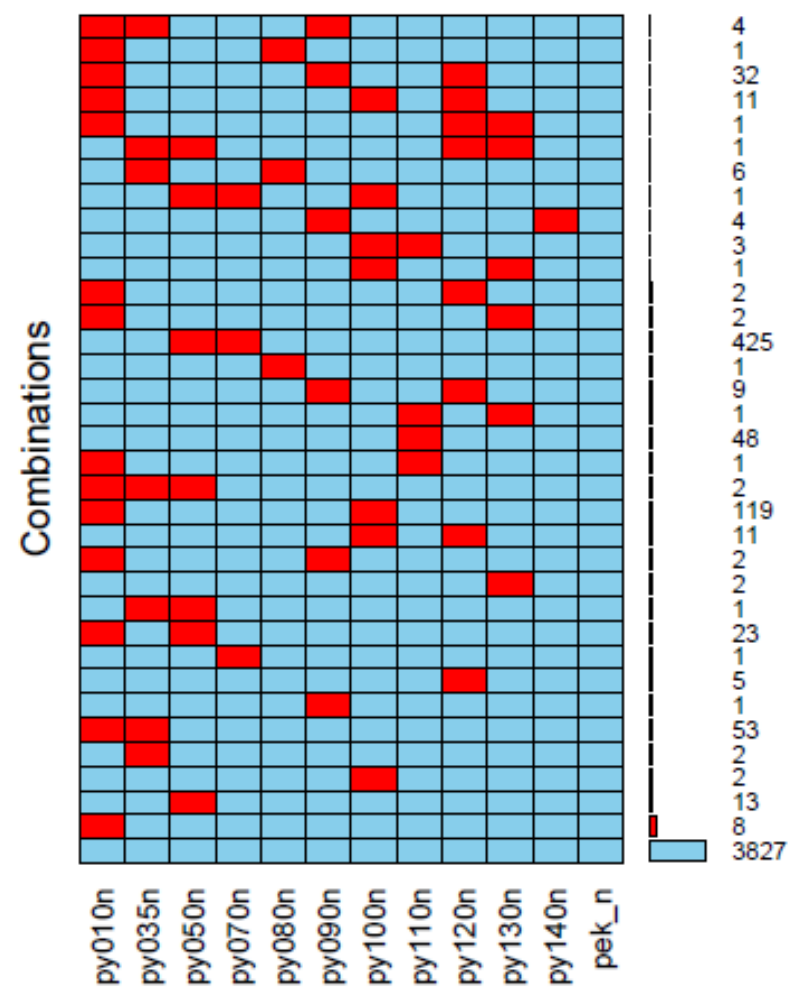
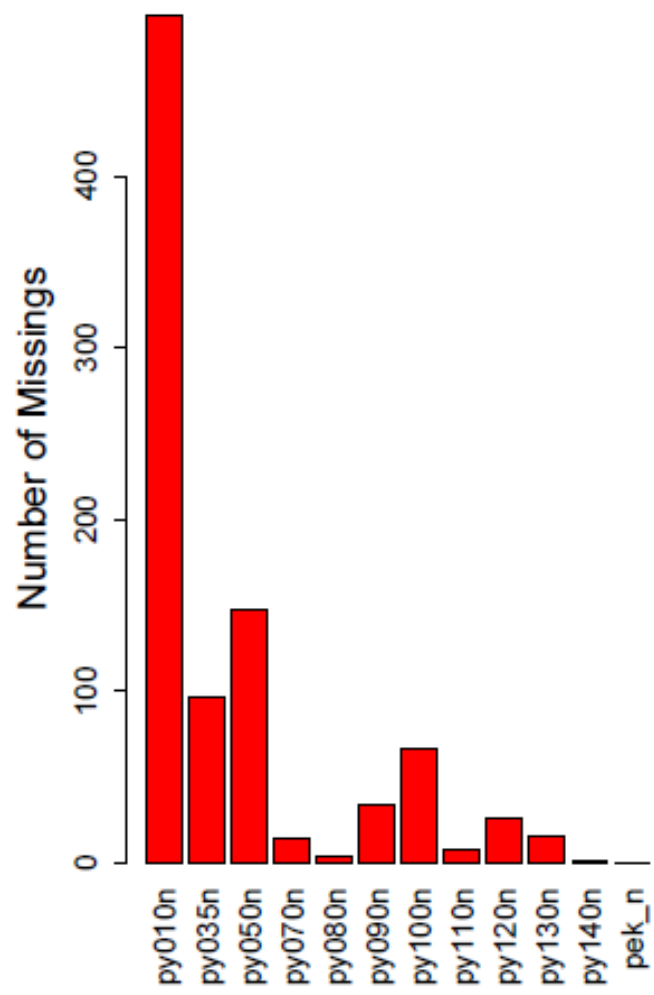


Jenkins, C., & Dillman, D. (1995). Towards a Theory of Self-Administered Questionnaire Design. In *Survey Measurement and Process Quality* by L. Lyberg et al., Wiley. Retrieved from <https://www.sesrc.wsu.edu/Dillman/papers/1997/A%20Theory%20of%20Self-Administered%20Questionnaire%20Design.pdf>

Processes of statistical production



Processes of statistical production



Processes of statistical production

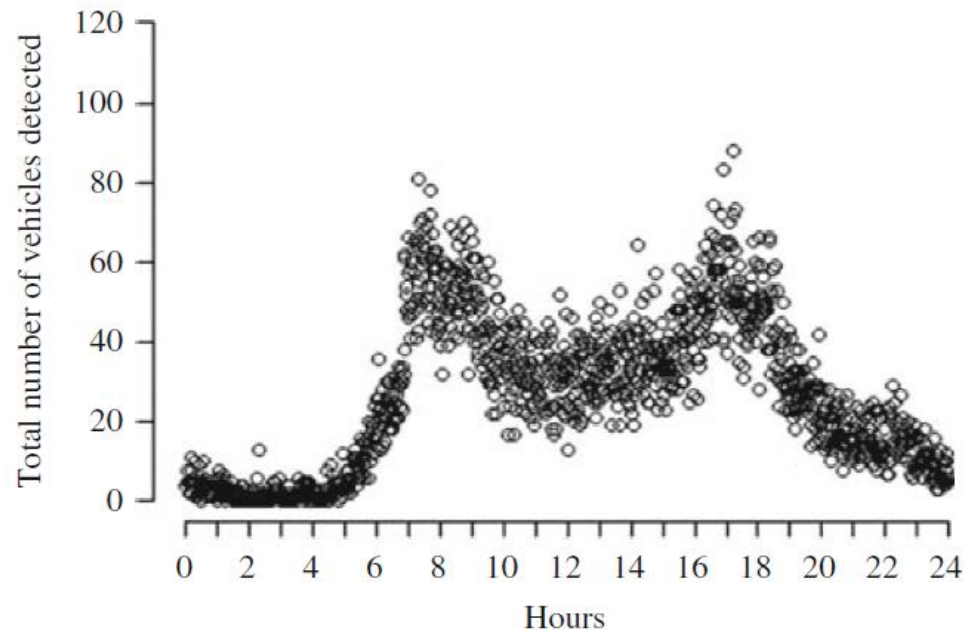


Fig. 3. Total number of vehicles counted by a detection location on highway A4 near Bergen op Zoom.

Data dissemination

Nature of data

User groups & Purpose

Medium



Jorge Camoes excelcharts.com



Color for non-designers: make it functional, not aesthetic



Manage color stimuli intensity: create layers of meaning with gray, pale colors and saturated colors

Minimize it, play with gray

Define functional tasks

Categorize

Group

Emphasize

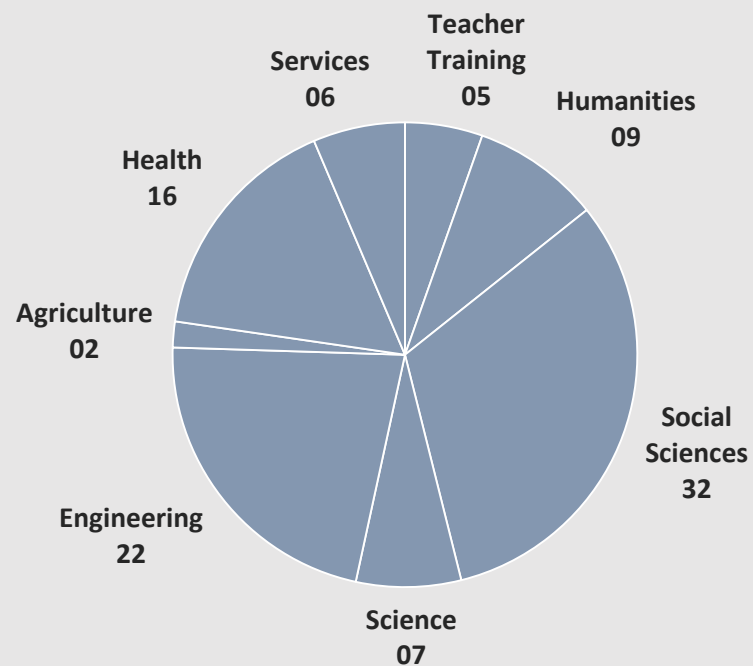
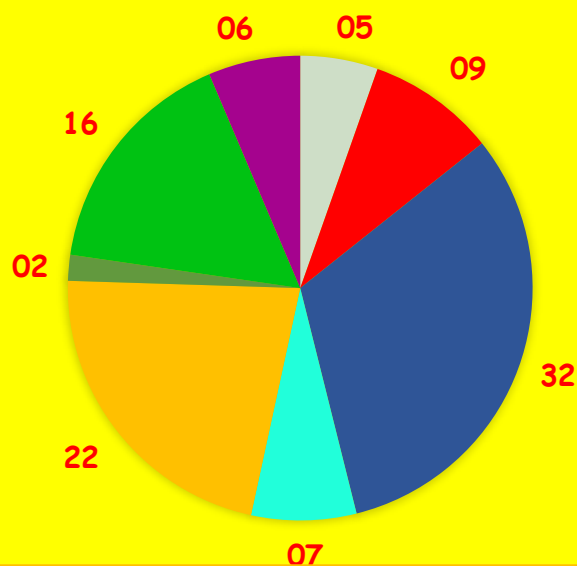
Sequence

Diverge

Alert



Stimuli intensity

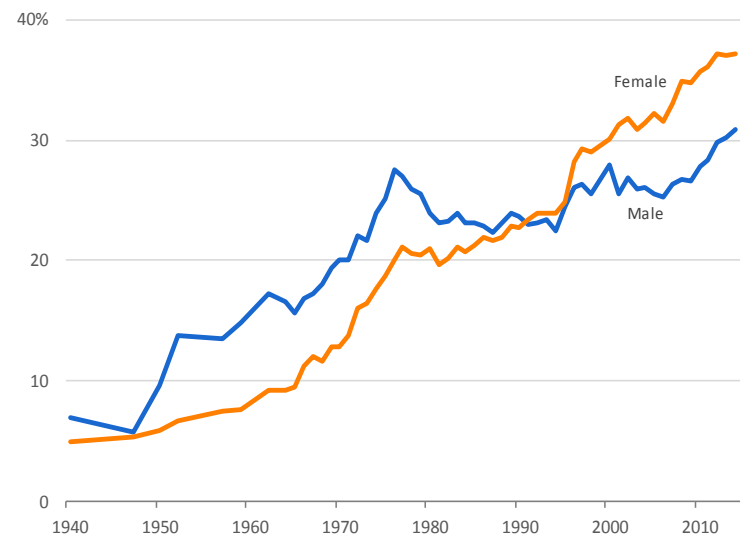


Categorize



AFTER MORE THAN 30 YEARS, ARE MEN PLAYING CATCH-UP IN EDUCATION?

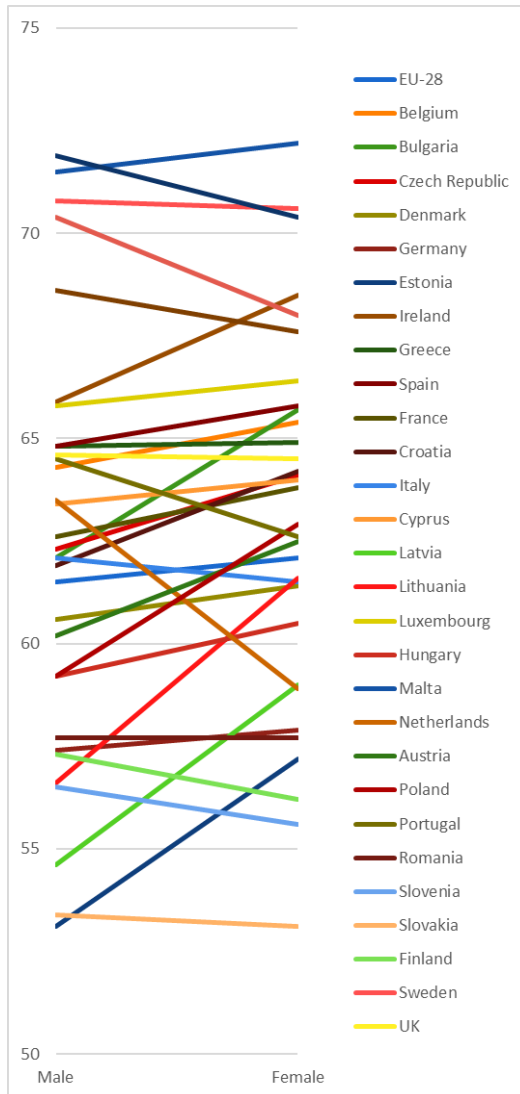
People aged 25 years and over who have completed college



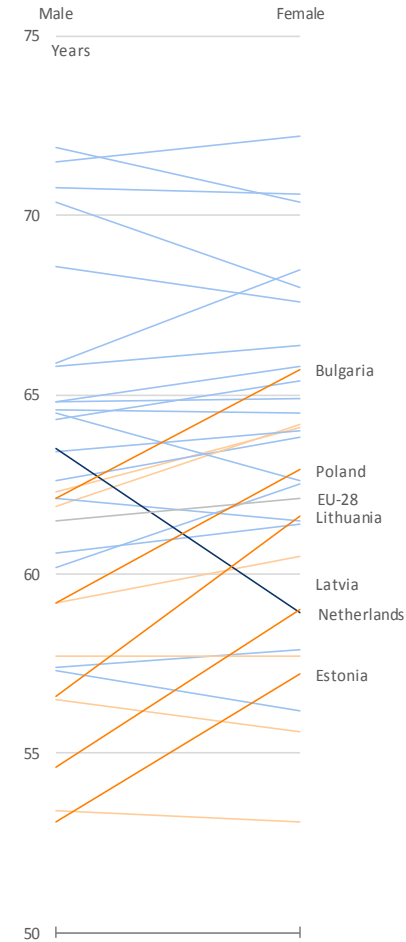
Source: U.S. Census Bureau



Emphasize



COUNTRIES WHERE THE GAP IN HEALTHY LIFE EXPECTANCY BETWEEN MEN AND WOMEN IS MORE THAN THREE YEARS
Healthy life expectancy by sex and country, in 2012



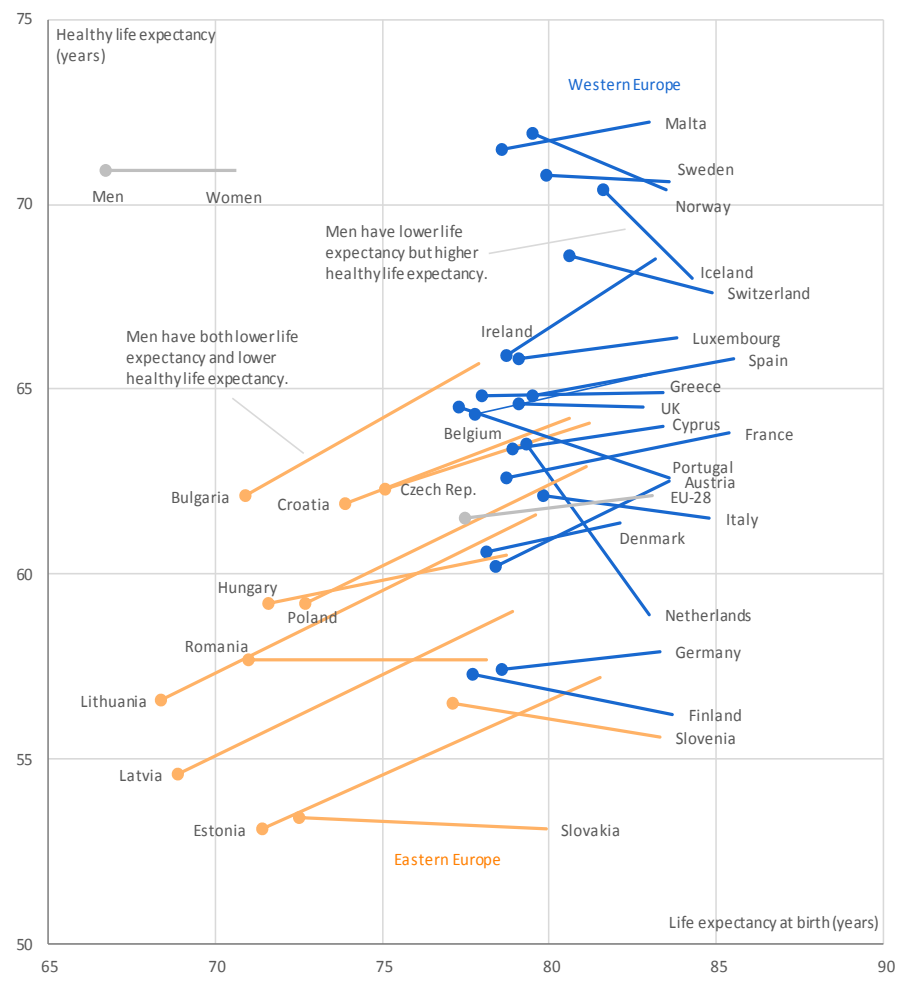
Source: Eurostat

Group



IN MOST COUNTRIES, WOMEN ENJOY A LONGER AND HEALTHIER LIFE THAN MEN

Large gap in life expectancy at birth in Eastern Europe



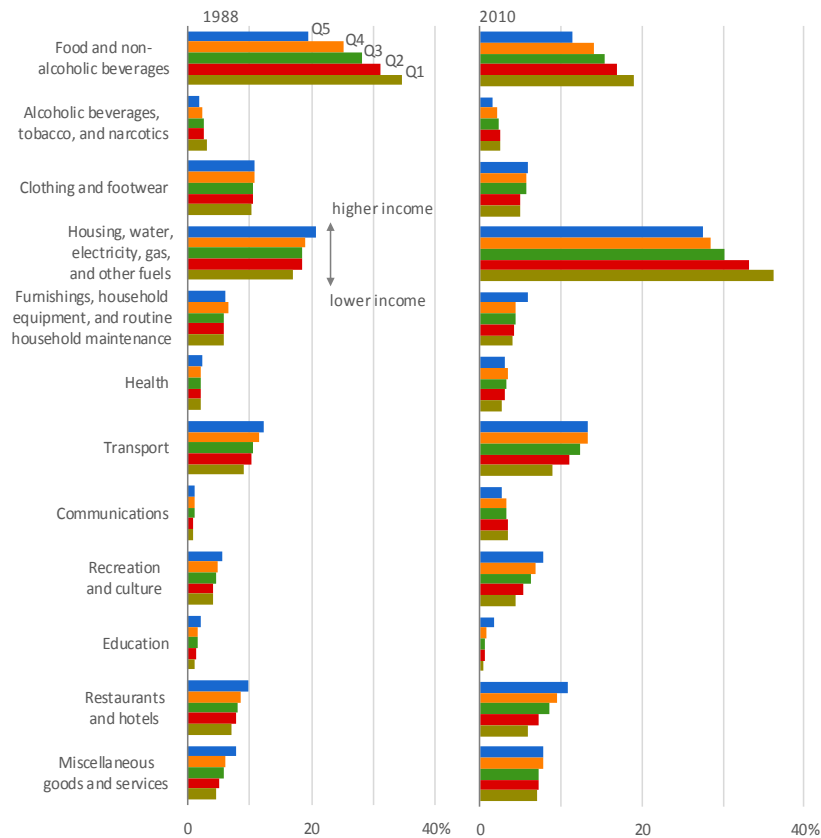
Source: Eurostat



Sequence

LESS IN FOOD, MORE IN HOUSING: CHANGES IN EXPENDITURE IN SPAIN

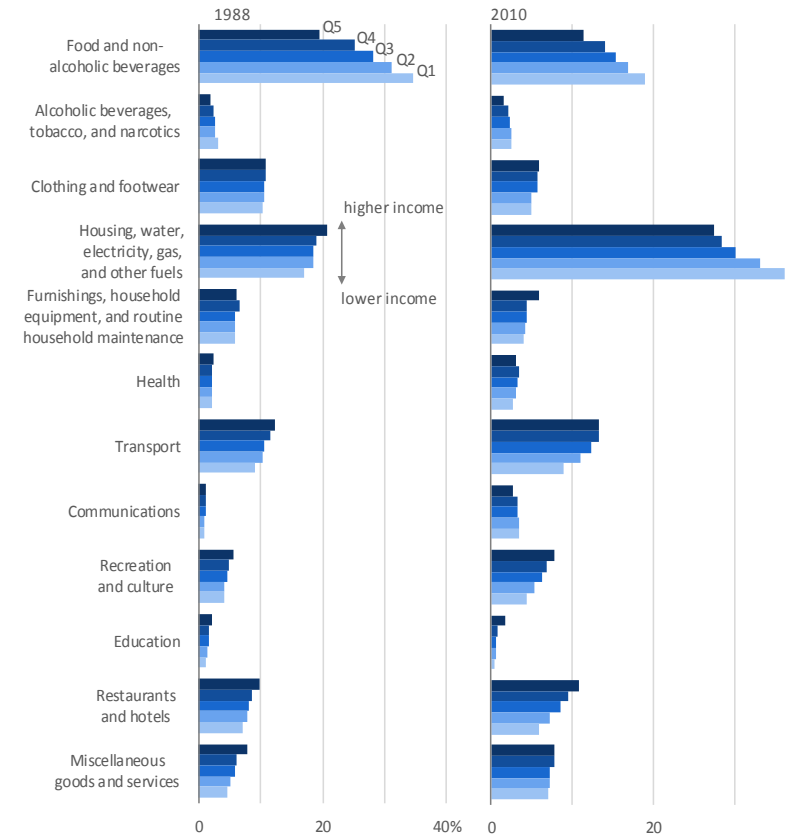
Proportion of household expenditure per category and income quintile



Source: Eurostat

LESS IN FOOD, MORE IN HOUSING: CHANGES IN EXPENDITURE IN SPAIN

Proportion of household expenditure per category and income quintile

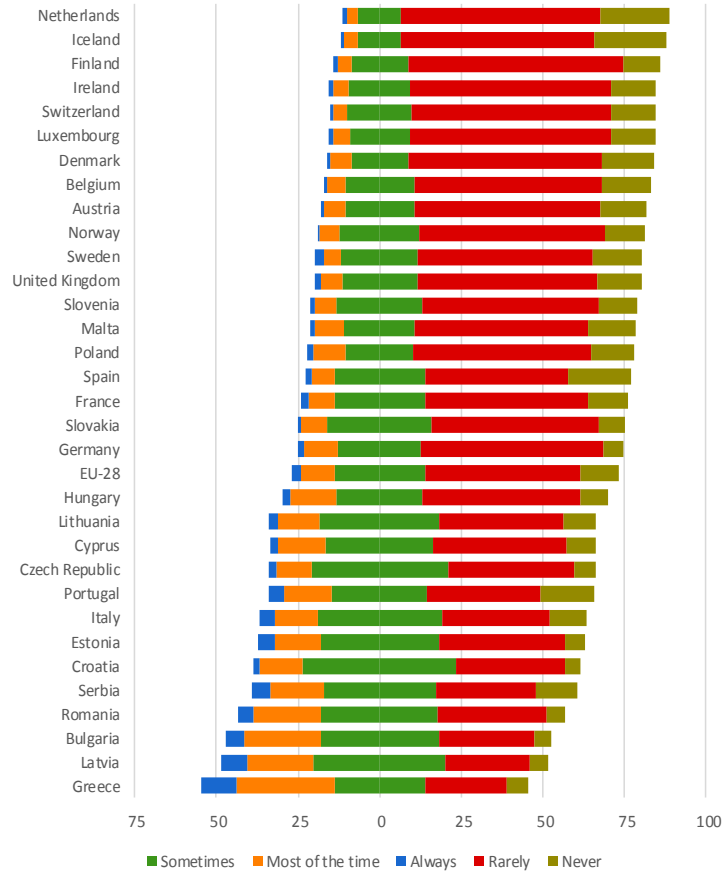


Source: Eurostat

Diverge

FREQUENCY OF BEING HAPPY IN THE LAST 4 WEEKS

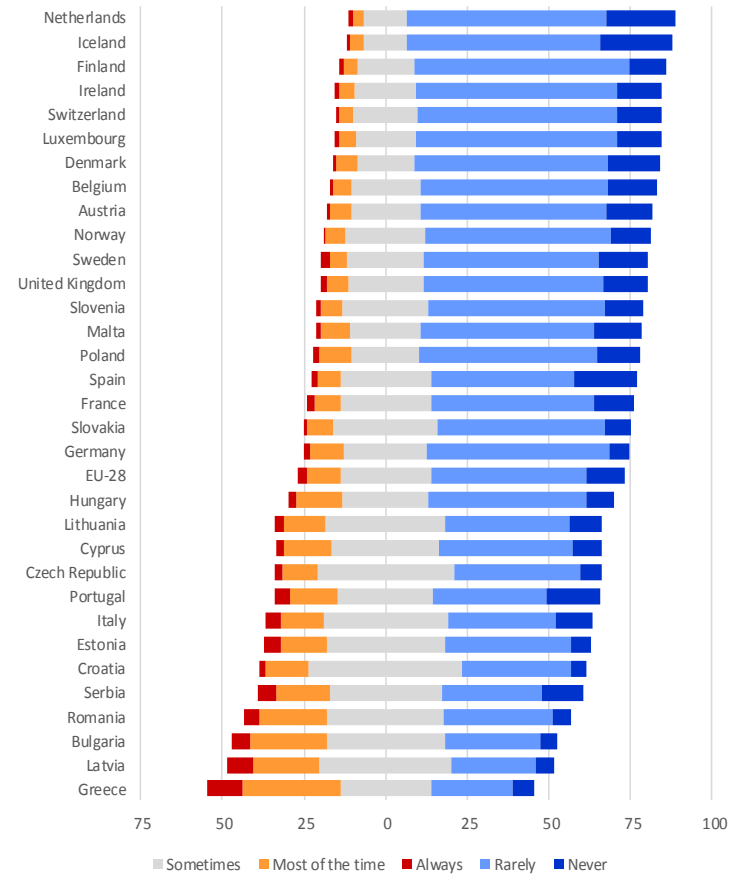
Population over 16 years old in 2013



Source: Eurostat

FREQUENCY OF BEING HAPPY IN THE LAST 4 WEEKS

Population over 16 years old in 2013

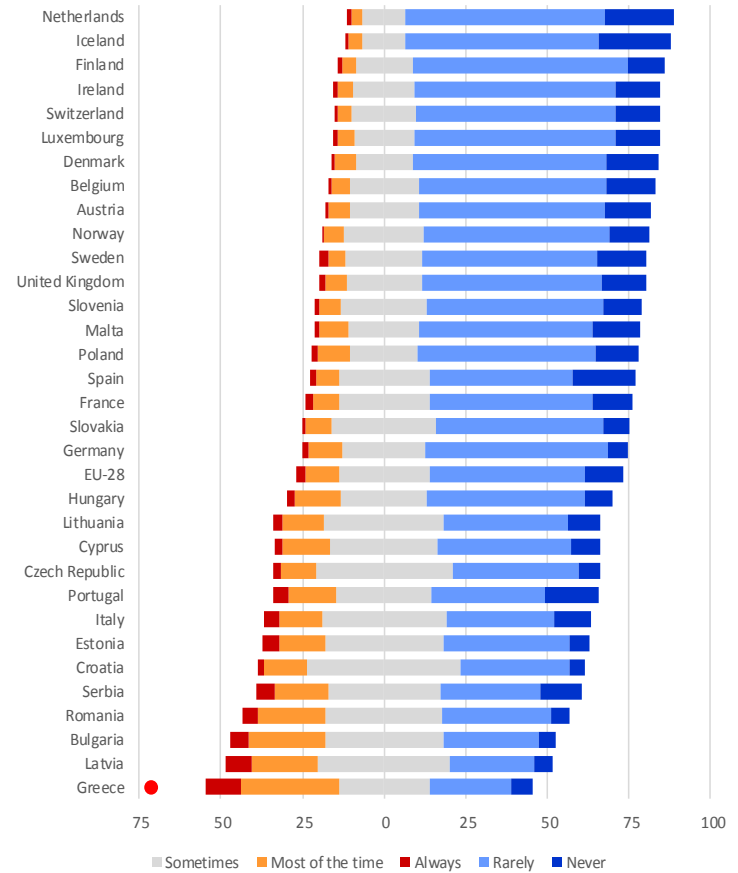


Source: Eurostat

Alert

FREQUENCY OF BEING HAPPY IN THE LAST 4 WEEKS

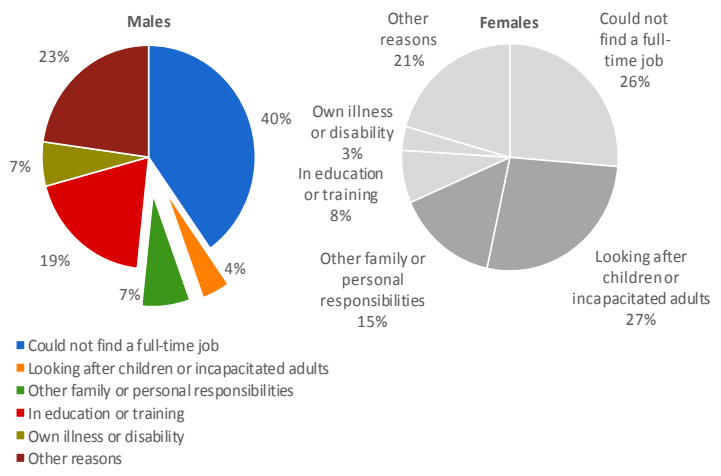
Population over 16 years old in 2013



Play with gray

MAIN REASONS FOR PART-TIME EMPLOYMENT BY SEX

From 15 to 64 years old, in the European Union (EU-28), in 2014

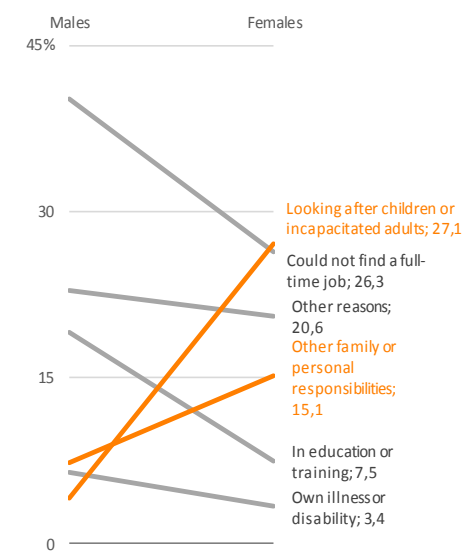


Source: Eurostat

No need for color

MAIN REASONS FOR PART-TIME EMPLOYMENT BY SEX

From 15 to 64 years old, in the European Union (EU-28), in 2014



Source: Eurostat

Use gray for context



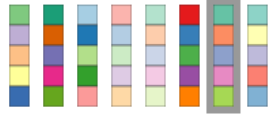
Color palettes: color brewer

COLORBREWER 2.0
color advice for cartography

Number of data classes: 6

Nature of your data: sequential diverging qualitative

Pick a color scheme:



Only show: colorblind safe print friendly photocopy safe

Context: roads cities borders

Background: solid color terrain

6-class Set2

	102,194,165
	252,141,98
	141,160,203
	231,138,195
	166,216,84
	255,217,47

EXPORT

RGB

© Cynthia Brewer, Mark Harrower and The Pennsylvania State University

[Source code and feedback](#)
[Back to Flash version](#)
[Back to ColorBrewer 1.0](#)

axismaps



<http://colorbrewer2.org>

Questions?



Further reading: Starting set 😊

