(Mis-)use of colours in scientific visualizations

Guidance to avoid and improve your colour use (with some R tidyverse hints)

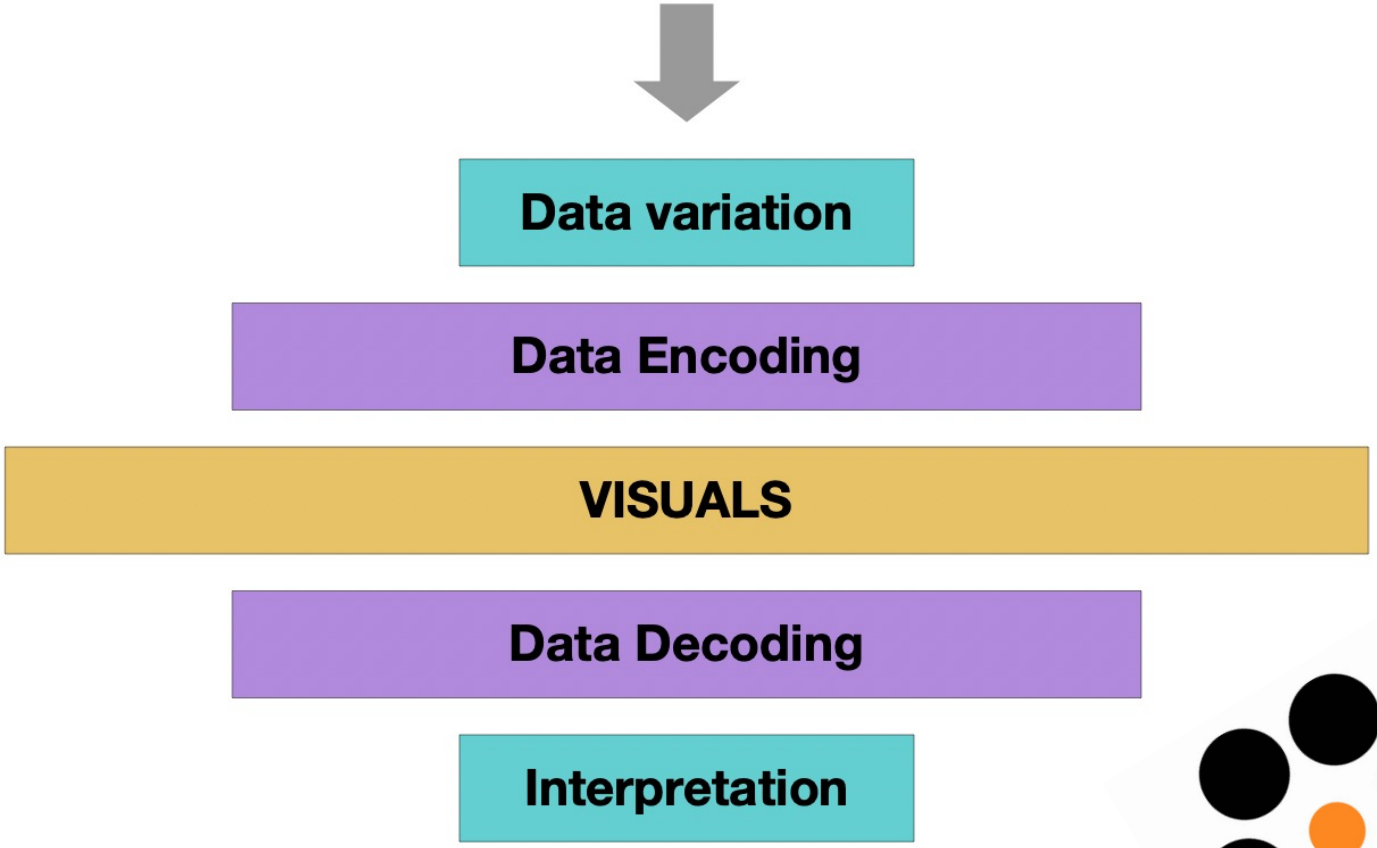
Michael Stölzle

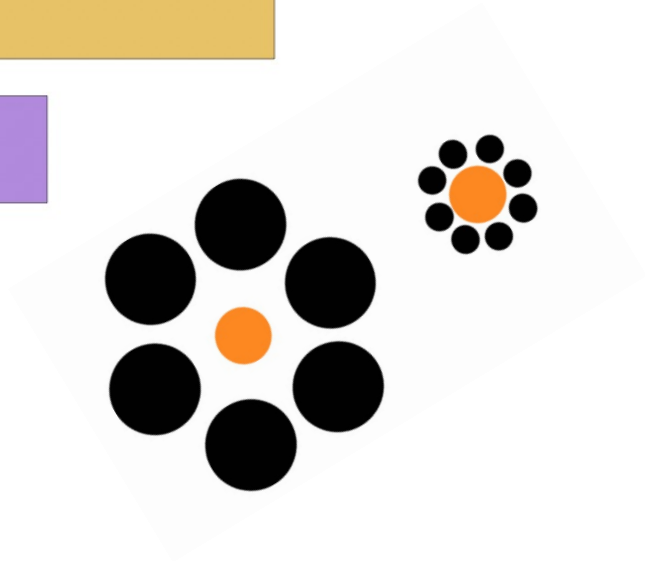
Faculty of Environment and Natural Resources, University Freiburg, Germany

E-Mail: michael.stoelzle@hydro.uni-freiburg.de
Twitter: @michistoelzle
Github: modche

26 May 2022 | Short Course| R in Hydrology

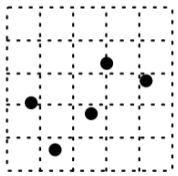
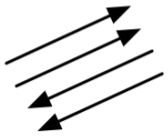
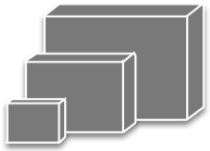
Why visualizations?



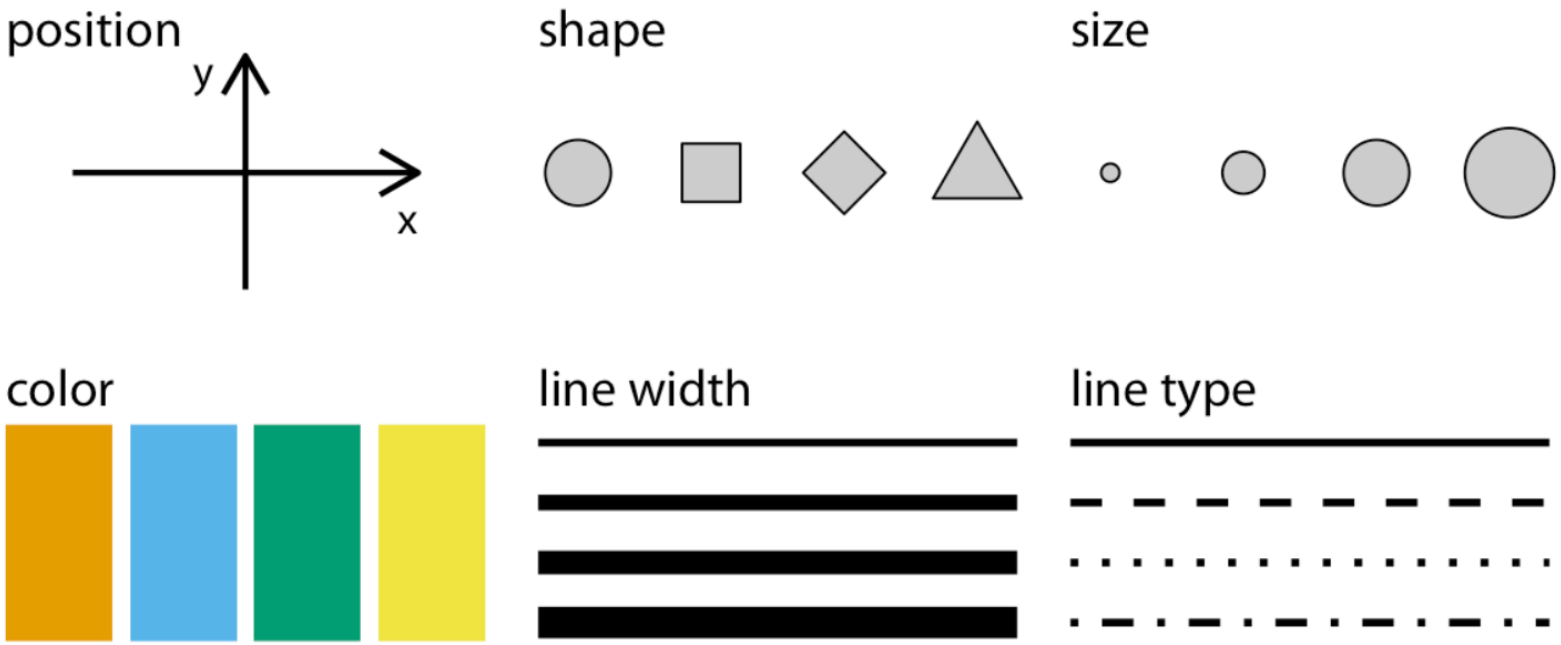


Visual cues, encodings and perception

◀ More accurate Less accurate ▶

Position	Length	Angle	Direction	Area	Volume	Saturation	Hue
							
Scatter plot	*Bar chart (stacked bar chart)*	*Pie chart*	*Slope chart*	*Bubble (Pie chart)*	*3D chart*	*Gradient shapes*	*Heatmap*

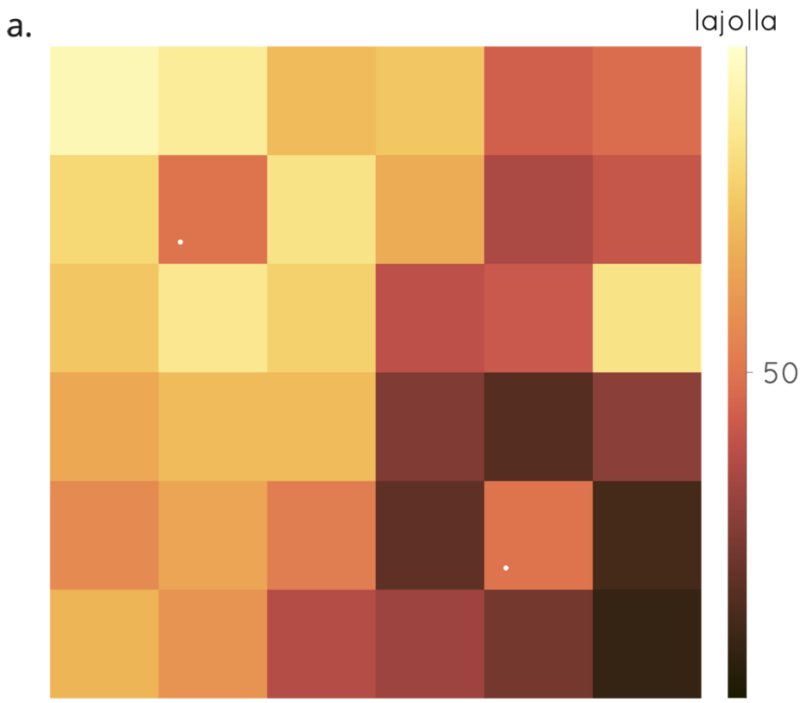
6 important encodings (in ggplot2)



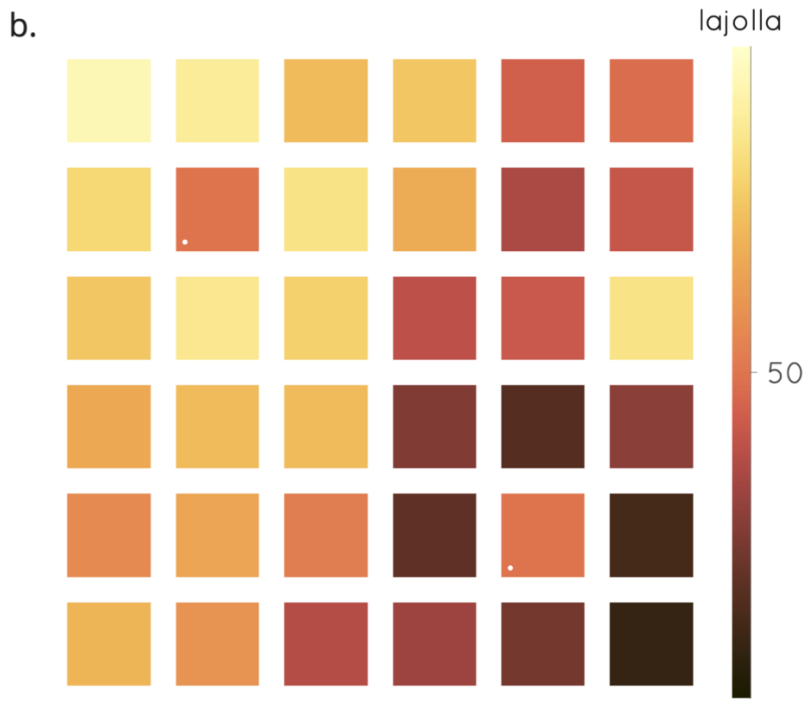
from: Claus O. Wilke

Colour perception

Heatmap without stroke lines



Heatmap with stroke lines



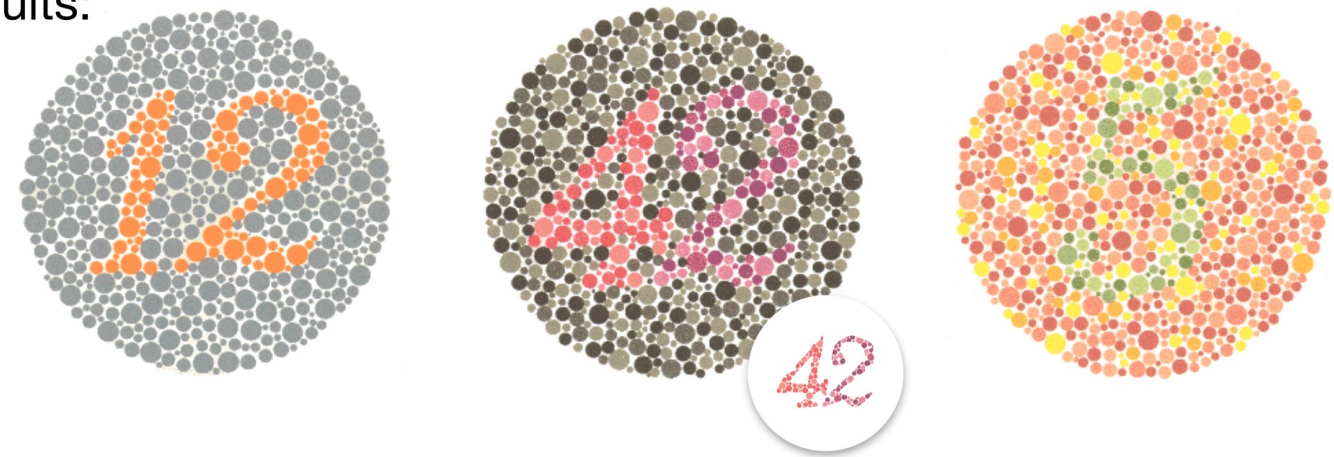
The misuse of colour in science communication

Fabio Crameri, Grace E. Shephard & Philip J. Heron

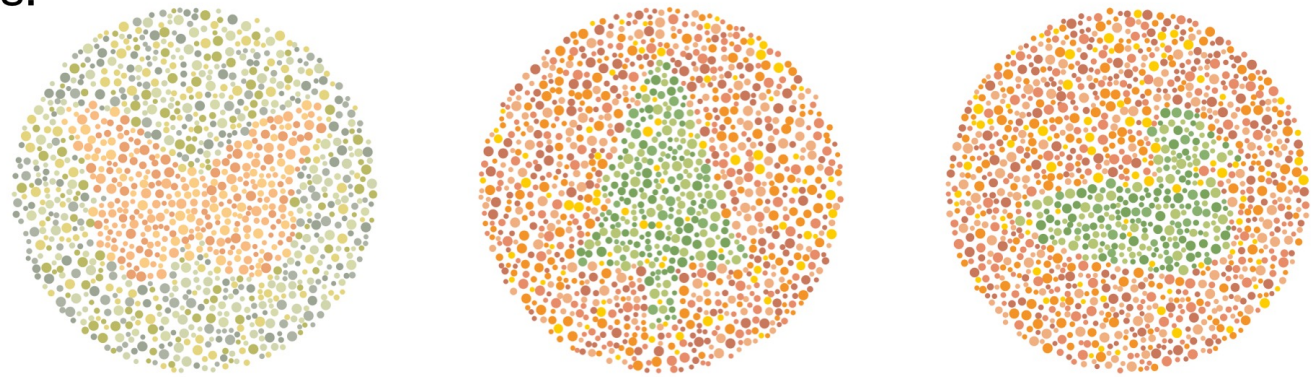
Nature Communications **11**, Article number: 5444 (2020) | Cite this article

Up 5-10% of people are colorblind.

For adults:



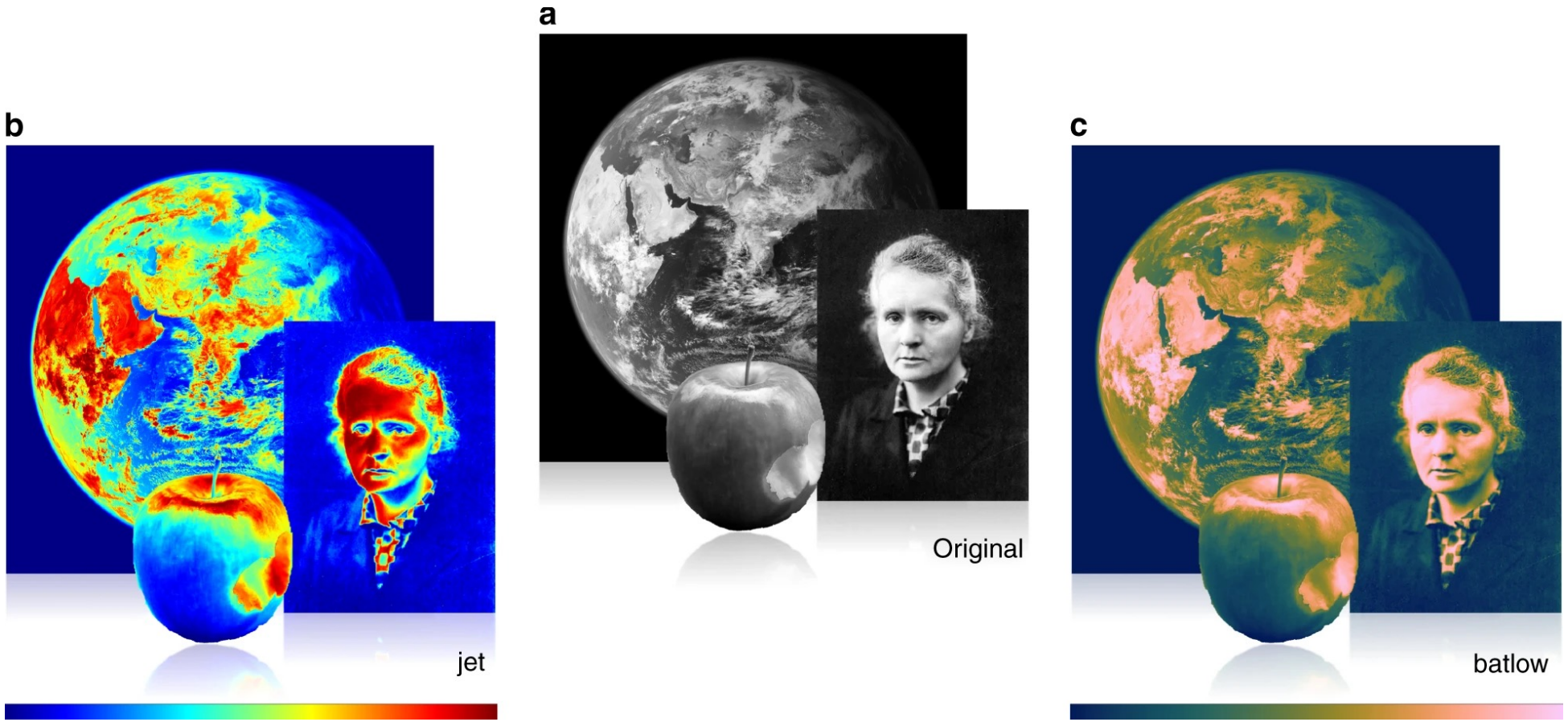
For kids:



The misuse of colour in science communication

Fabio Crameri, Grace E. Shephard & Philip J. Heron

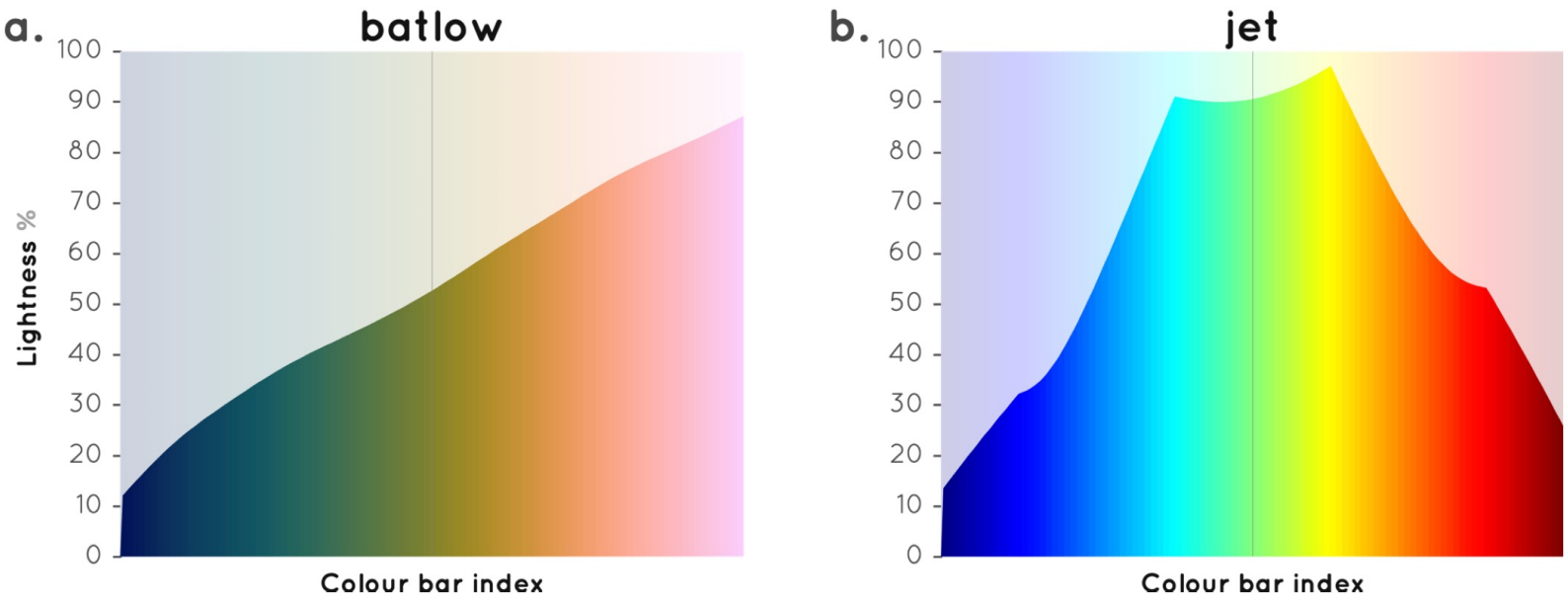
Nature Communications **11**, Article number: 5444 (2020) | Cite this article



The misuse of colour in science communication

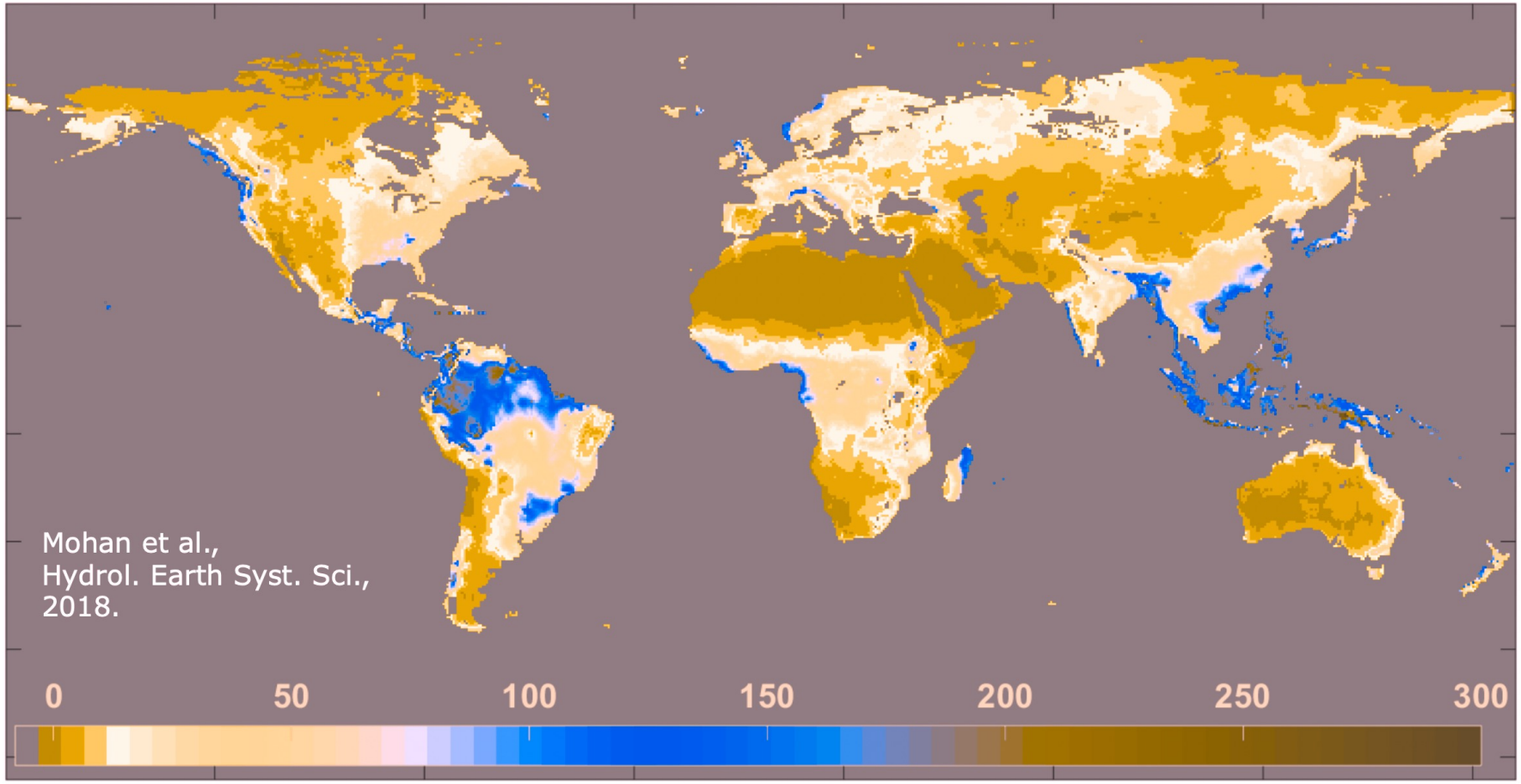
Fabio Crameri, Grace E. Shephard & Philip J. Heron

Nature Communications **11**, Article number: 5444 (2020) | Cite this article



Supplementary Figure 1 Colour map lightness gradient. Lightness is a property of colour to which the human visual apparatus is most sensitive to and, as such, lightness is a pivotal property of a colour map. Shown are the lightness gradients across **a** *batlow*[1], for which it is monotonically increasing, and **b** *jet*, for which it is neither monotonically increasing, nor symmetric, nor has its maximum lightness in a sensible location.

(b) Standard deviation of groundwater recharge



colorblindr::

palette_OkabeIto {colorblindr} R Documentation

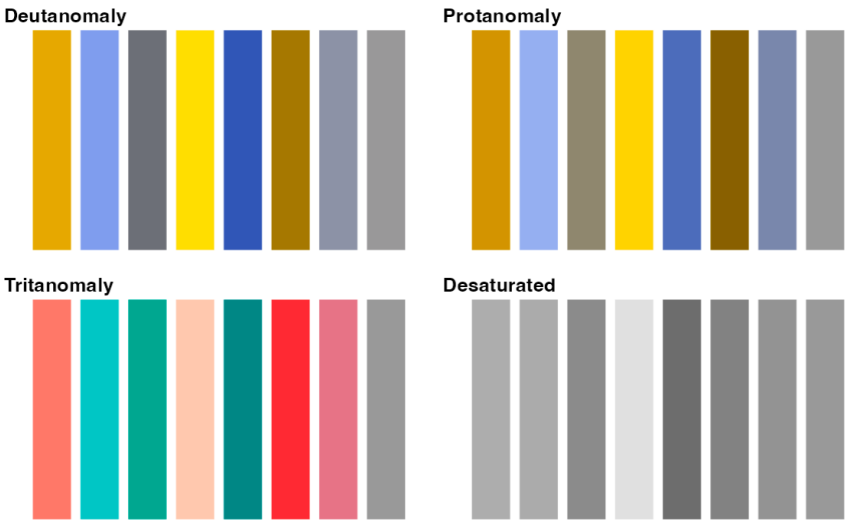
Color palette proposed by Okabe and Ito

Description

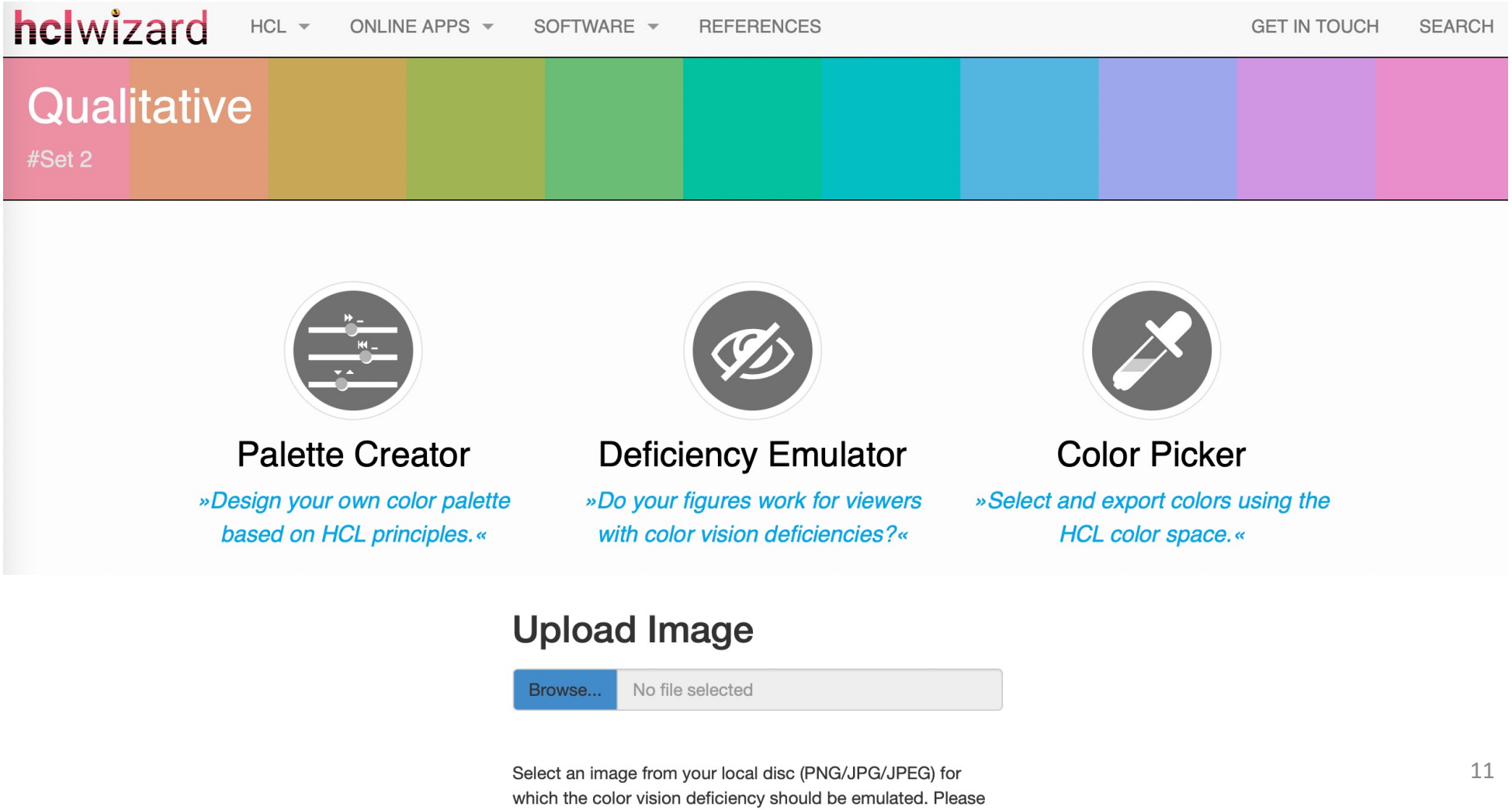
Two color palettes taken from the article "Color Universal Design" by Okabe and Ito, http://jfly.iam.u-tokyo.ac.jp/color/. The variant `palette_OkabeIto` contains a gray color, while `palette_OkabeIto_black` contains black instead.

```
ggplot(iris, aes(Sepal.Length, fill = Species)) +
  geom_density(alpha = 0.7) +
  scale_fill_OkabeIto()
```

```
colorblindr::palette_plot(palette_OkabeIto)
colorblindr::cvd_grid(palette_plot(palette_OkabeIto))
```



Many webtools… (like http://hclwizard.org)



Ok, no rainbow color map…

… but how to choose the right color map?

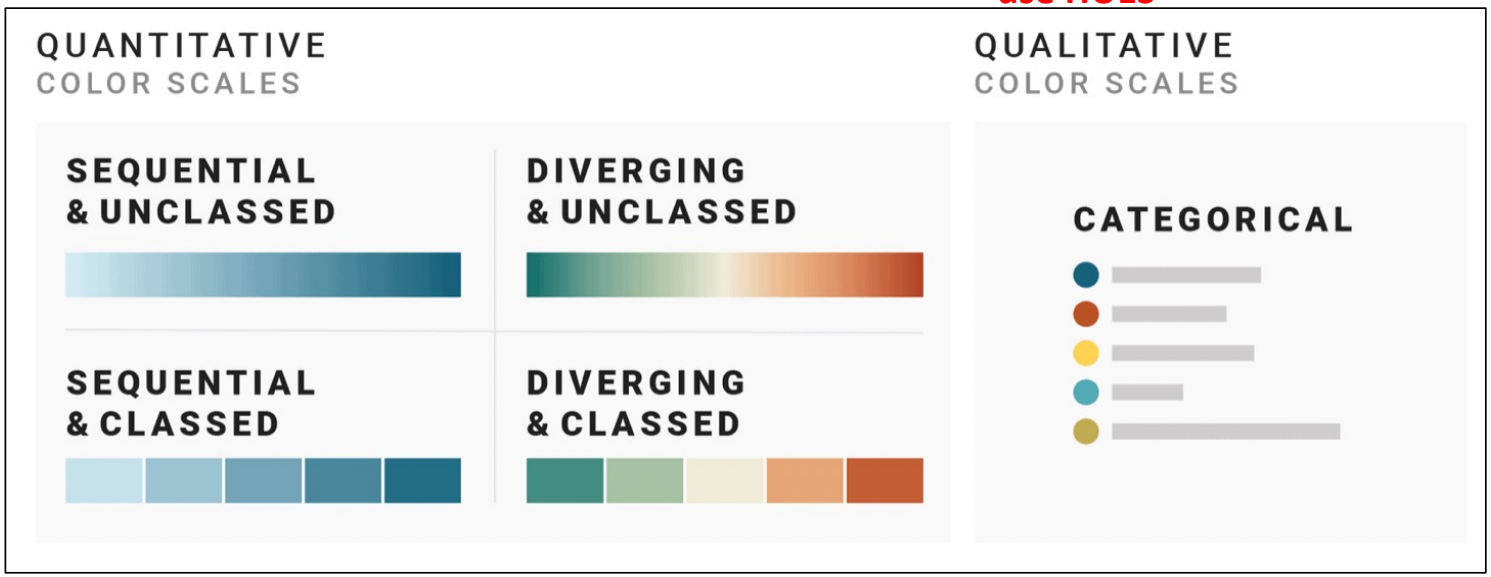
Just right! Your data determines your colour map!

numeric values with one direction
there is a order
= use SHADES

countable or categories
no order, same perceptual weight
= use HUES

intuitive judgement, most exact color map

give statistical guidance or data is classed



meaningful middle value:

- **zero, 50%, mean**
- **threshold, target value**

any scale, highlight is possible:

Source: https://blog.datawrapper.de/which-color-scale-to-use-in-data-vis/

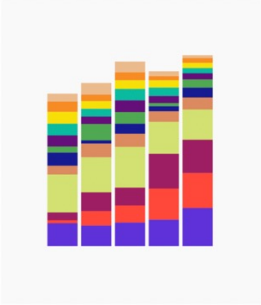
Avoid colour(s) – or less is sometime more…

1. Highlight with colour

NOT IDEAL

BETTER

2. Show shades, not hues



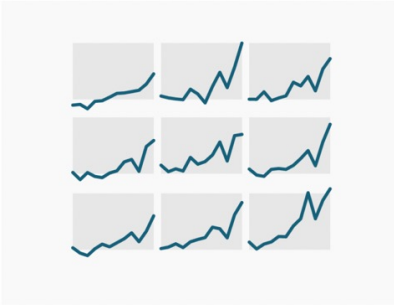
NOT IDEAL

NOT IDEAL

BETTER

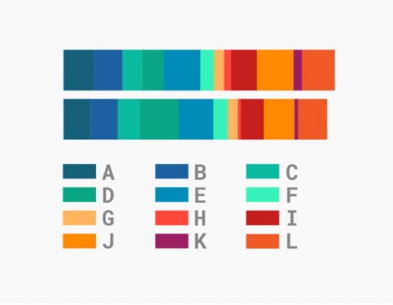
3. Go for small multiples (`facet_wrap`)

NOT IDEAL



BETTER

4. Doesn't work? Change graph type!



NOT IDEAL

BETTER

https://blog.datawrapper.de/10-ways-to-use-fewer-colors-in-your-data-visualizations/

Avoid colour (as primary encoding)

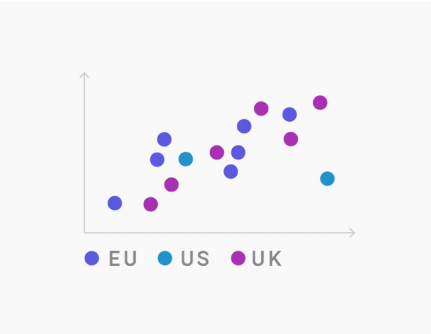
1. Pie charts

NOT IDEAL

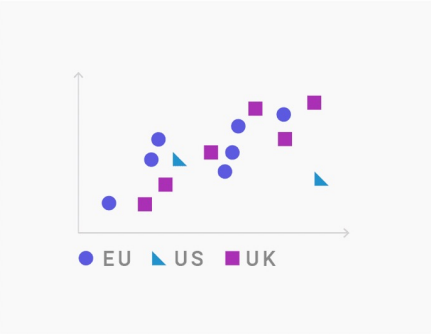
BETTER

From colour to position

2. Scatter plots



NOT IDEAL

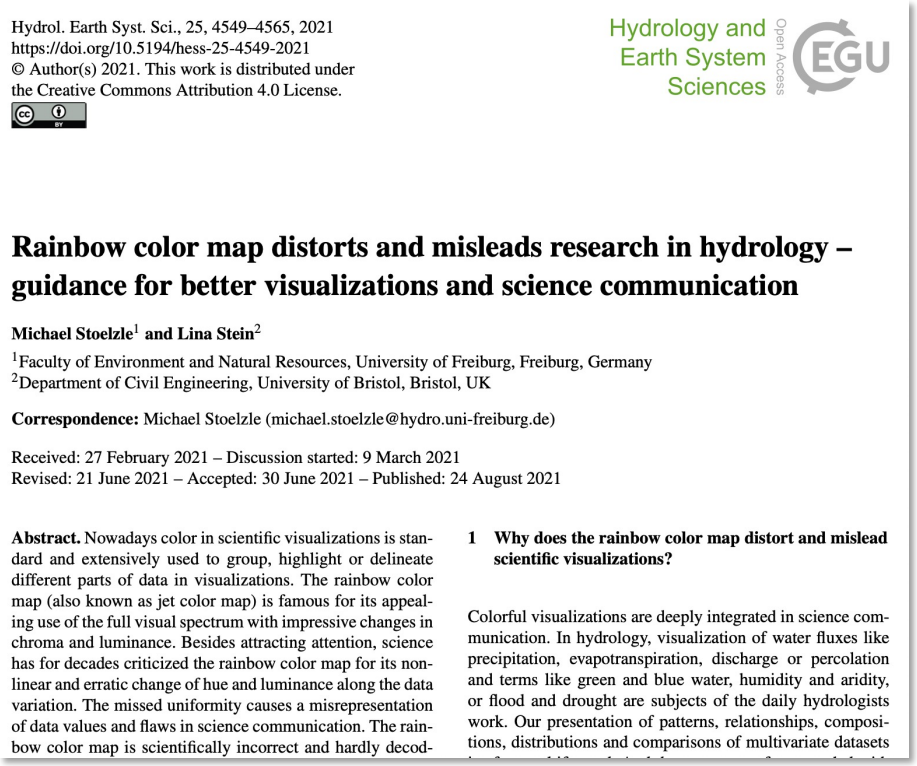


BETTER

From colour to shape

https://blog.datawrapper.de/colorblindness-part2/

Techniques to improve line charts



Hydrol. Earth Syst. Sci., 25, 4549–4565, 2021
https://doi.org/10.5194/hess-25-4549-2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Hydrology and Earth System Sciences
Open Access
EGU

Rainbow color map distorts and misleads research in hydrology – guidance for better visualizations and science communication

Michael Stoelzle[1] **and Lina Stein**[2]

[1]Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany
[2]Department of Civil Engineering, University of Bristol, Bristol, UK

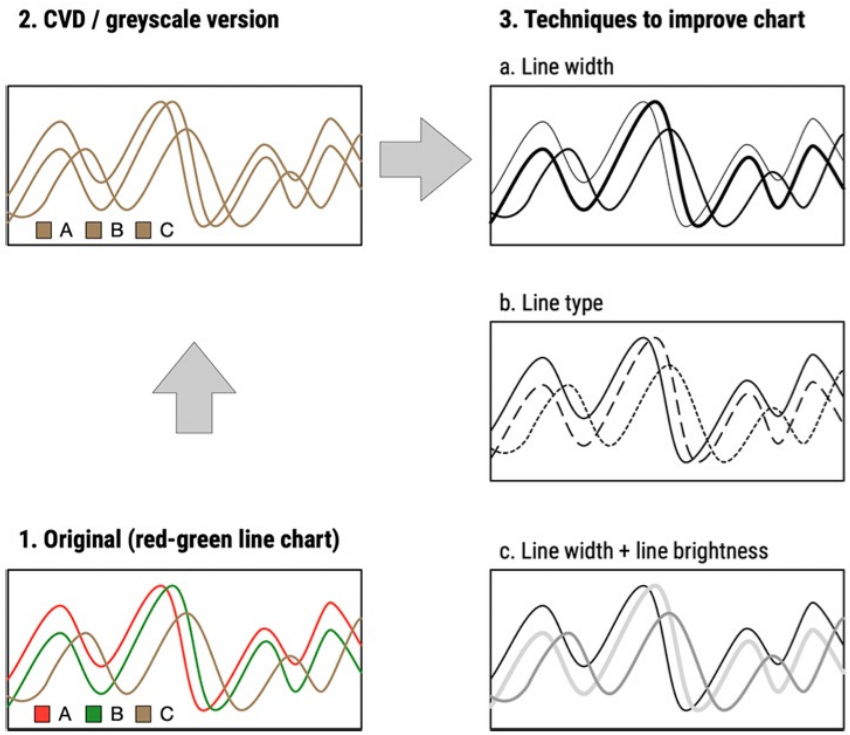
Correspondence: Michael Stoelzle (michael.stoelzle@hydro.uni-freiburg.de)

Received: 27 February 2021 – Discussion started: 9 March 2021
Revised: 21 June 2021 – Accepted: 30 June 2021 – Published: 24 August 2021

Abstract. Nowadays color in scientific visualizations is standard and extensively used to group, highlight or delineate different parts of data in visualizations. The rainbow color map (also known as jet color map) is famous for its appealing use of the full visual spectrum with impressive changes in chroma and luminance. Besides attracting attention, science has for decades criticized the rainbow color map for its non-linear and erratic change of hue and luminance along the data variation. The missed uniformity causes a misrepresentation of data values and flaws in science communication. The rainbow color map is scientifically incorrect and hardly decod-

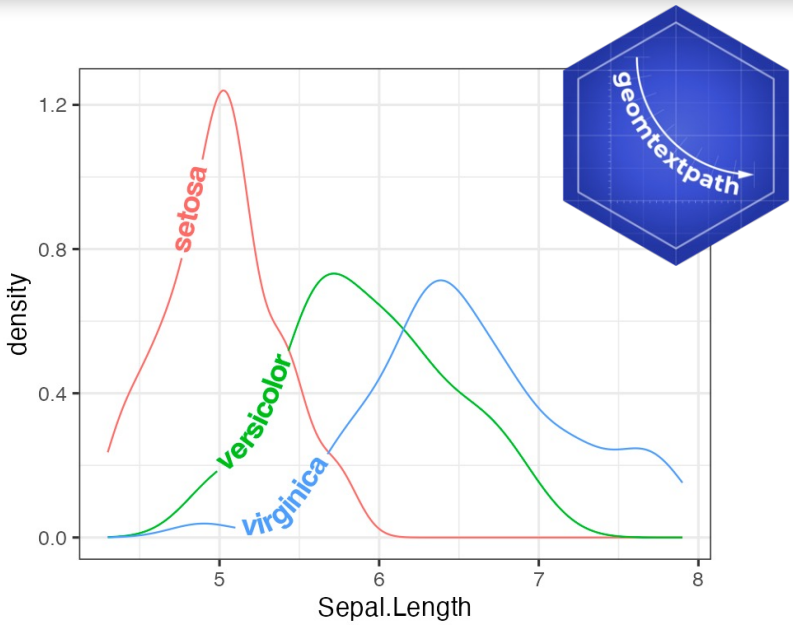
1 Why does the rainbow color map distort and mislead scientific visualizations?

Colorful visualizations are deeply integrated in science communication. In hydrology, visualization of water fluxes like precipitation, evapotranspiration, discharge or percolation and terms like green and blue water, humidity and aridity, or flood and drought are subjects of the daily hydrologists work. Our presentation of patterns, relationships, compositions, distributions and comparisons of multivariate datasets



Direct labeling instead of colours and legends

```
ggplot(iris, aes(x = Sepal.Length,
                 colour = Species,
                 label  = Species)) +
  geom_textdensity(size = 6,
                   fontface = 2,
                   hjust = 0.2, vjust = 0.3)
```

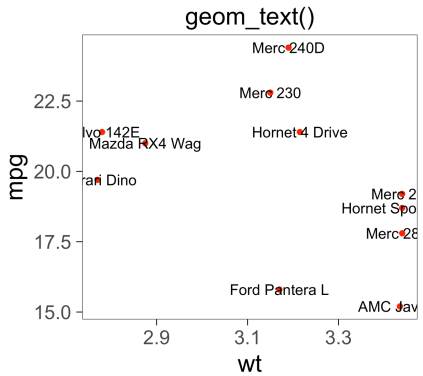


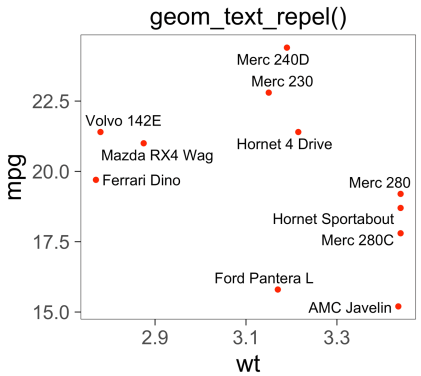

```
p <- ggplot(data, aes(x,y)) + geom_point()

# Labels with geom_text
p +  geom_text(aes(label = name),
               check_overlap = TRUE)

# Avoid overlaps by repelling text labels
p + geom_text_repel()

# Labels with background
p + geom_label_repel()
```





A checklist (n=10)

- ❑ Colour is not the primary encoding?
- ❑ Data determines color map?
- ❑ No rainbow color map? Figure works as black-white-print out?
- ❑ Colours are colorblind-safe?
- ❑ Multi-hue color maps should be linear in lightness gradient (e.g. viridis)
- ❑ A legend needed? Go for direct labeling!
- ❑ Consider small multiples to reduce data dimensions
- ❑ Encoding time is appropriate for audience (paper vs. presentation)?
- ❑ Binned color map for more guidance (emphasize statistical values like quartiles)
- ❑ Convert to diverging color map to emphasize extremes.

Collection of gg-packages

padlet.com/michistoelzle/ggplotextensions

Overview: 100+ ggplot extensions

erikgahner/**awesome-ggplot2**

A curated list of awesome ggplot2 tutorials, packages etc.

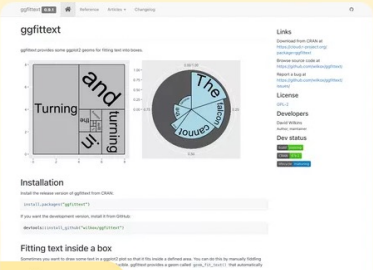
600 Stars 76 Forks

github.com

0 0 0

Add comment

Fit text in Boxes.



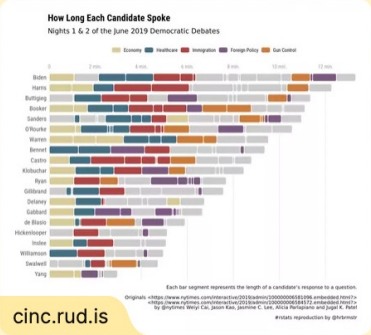
wilkox.org

Fit Text Inside a Box in ggplot2

1 0 0

Add comment

ggchicklet - improved bar charts!



cinc.rud.is

Create Chicklet (Rounded Segmented Column) Charts

Playful, but sometimes useful. Best of 'white stroke lines'

https://cinc.rud.is/web/packages/ggchicklet/

0 0 0

Add comment

Overplotting? gghdr!

github.com

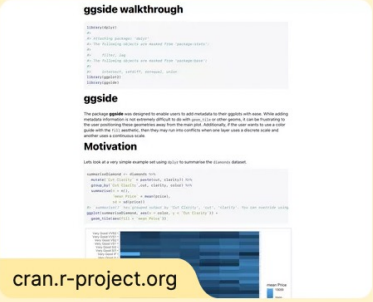
GitHub - Sayani07/gghdr: Plots of highest density regions (HDR) for ggplot2

0 0 0

Add comment

3 graphs in 1: patchwork

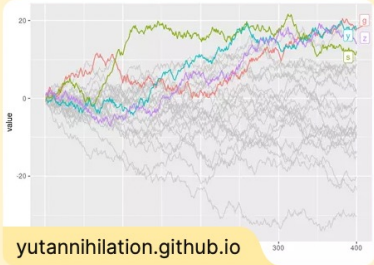
ggside - add marginal distributions



cran.r-project.org

ggside walkthrough

{gghighlight} Focus on one line!



yutannihilation.github.io

Highlight Lines and Points in ggplot2

0 0 0

Add comment