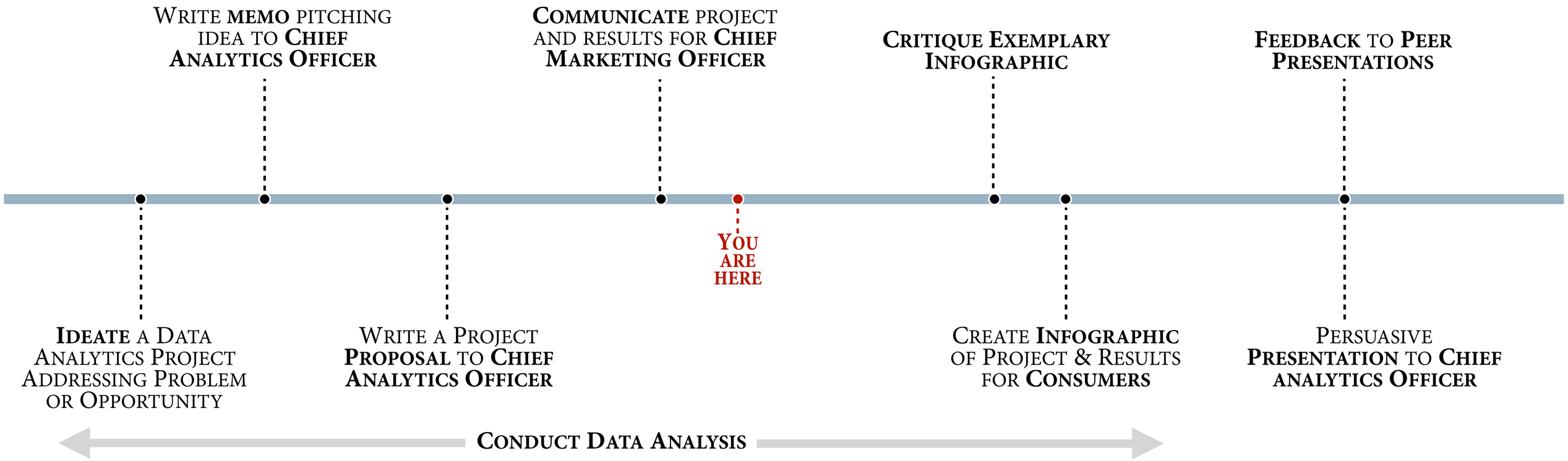
Storytelling With Data

Effective visuals — encoding data graphically: grammar, layers, color

Scott Spencer | Columbia University

Conceptual project timeline



the grammar of graphics

DATA : a set of data operations that create variables from datasets

TRANSFORMATIONS : variable transformations (*e.g., rank*)

SCALES : scale transformations (*e.g., log*)

COORDINATES : a coordinate system (*e.g., cartesian, polar*)

ELEMENTS : graphs (*e.g., points, lines*) and their aesthetic attributes (*e.g., color, opacity, shape, size, orientation*)

GUIDES : one or more guides (*axes, legends, etc.*)

Wilkinson, Leland

```
# load grammar of graphics
library(ggplot2)

p <-

  # functions for data ink

ggplot(data = <data>,
       mapping = aes(<aesthetic> = <variable>,
                     <aesthetic> = <variable>,
                     <...> = <...>) +
geom_<type>(<...>) +
scale_<mapping>_<type>(<...>) +
coord_<type>(<...>) +
facet_<type>(<...>) +
<...> +

  # functions for non-data ink

labs(<...>) +
theme(<...> = <...>) +
annotate(<...>) +
<...>
```

DATA
TRANSFORMATIONS

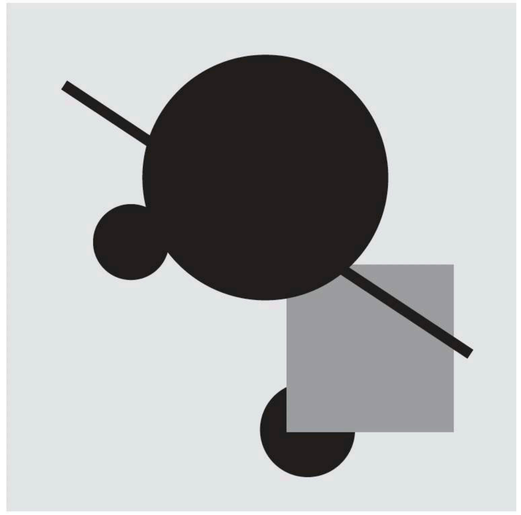
ELEMENTS
SCALES & GUIDES
COORDINATES

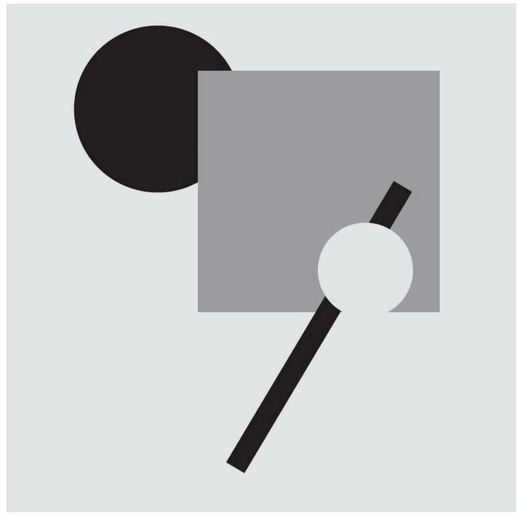
GUIDES

```
element_blank()
element_line(<...> = <...>)
element_rect(<...> = <...>)
element_text(<...> = <...>)
```

Wilkinson, Leland

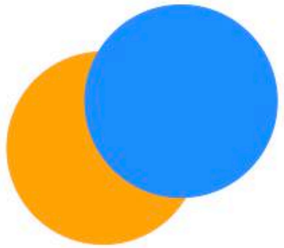
thinking about graphics as layers



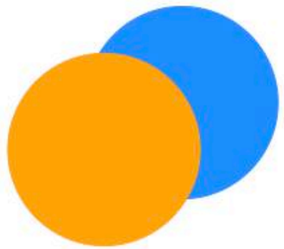


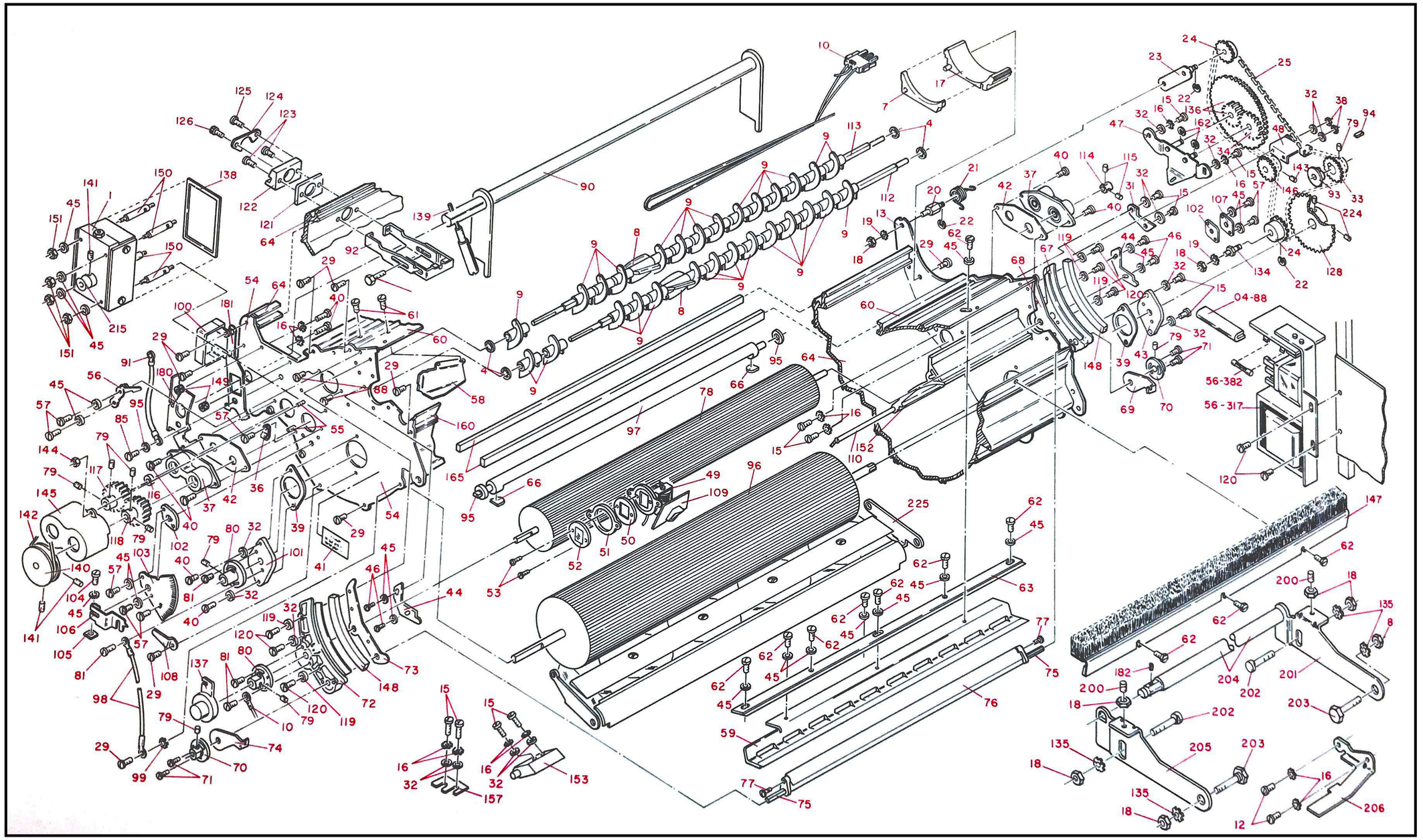
Samara, Timothy

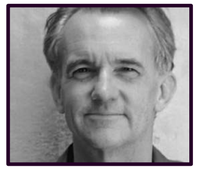
```
ggplot() +
  theme_void() +
  scale_x_continuous(limits = c(-5, 5)) +
  scale_y_continuous(limits = c(-5, 5)) +
  geom_point(aes(x = 0, y = 0),
             size = 50, color = "orange") +
  geom_point(aes(x = 1, y = 1),
             size = 50, color = "dodgerblue")
```



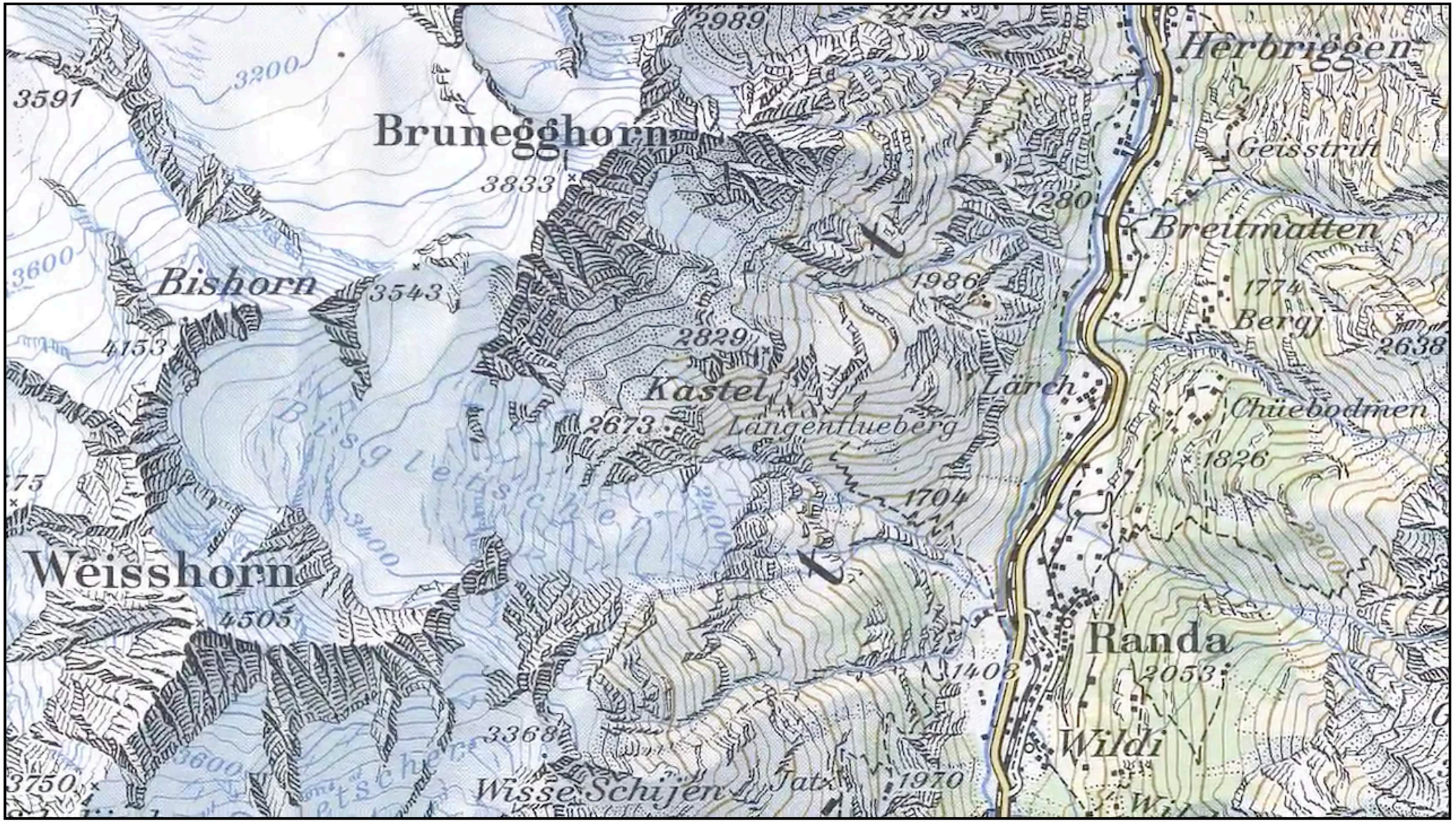
```
ggplot() +
  theme_void() +
  scale_x_continuous(limits = c(-5, 5)) +
  scale_y_continuous(limits = c(-5, 5)) +
  geom_point(aes(x = 1, y = 1),
             size = 50, color = "dodgerblue") +
  geom_point(aes(x = 0, y = 0),
             size = 50, color = "orange")
```

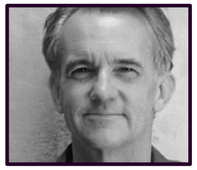






Tufte, Edward





Tufte, Edward

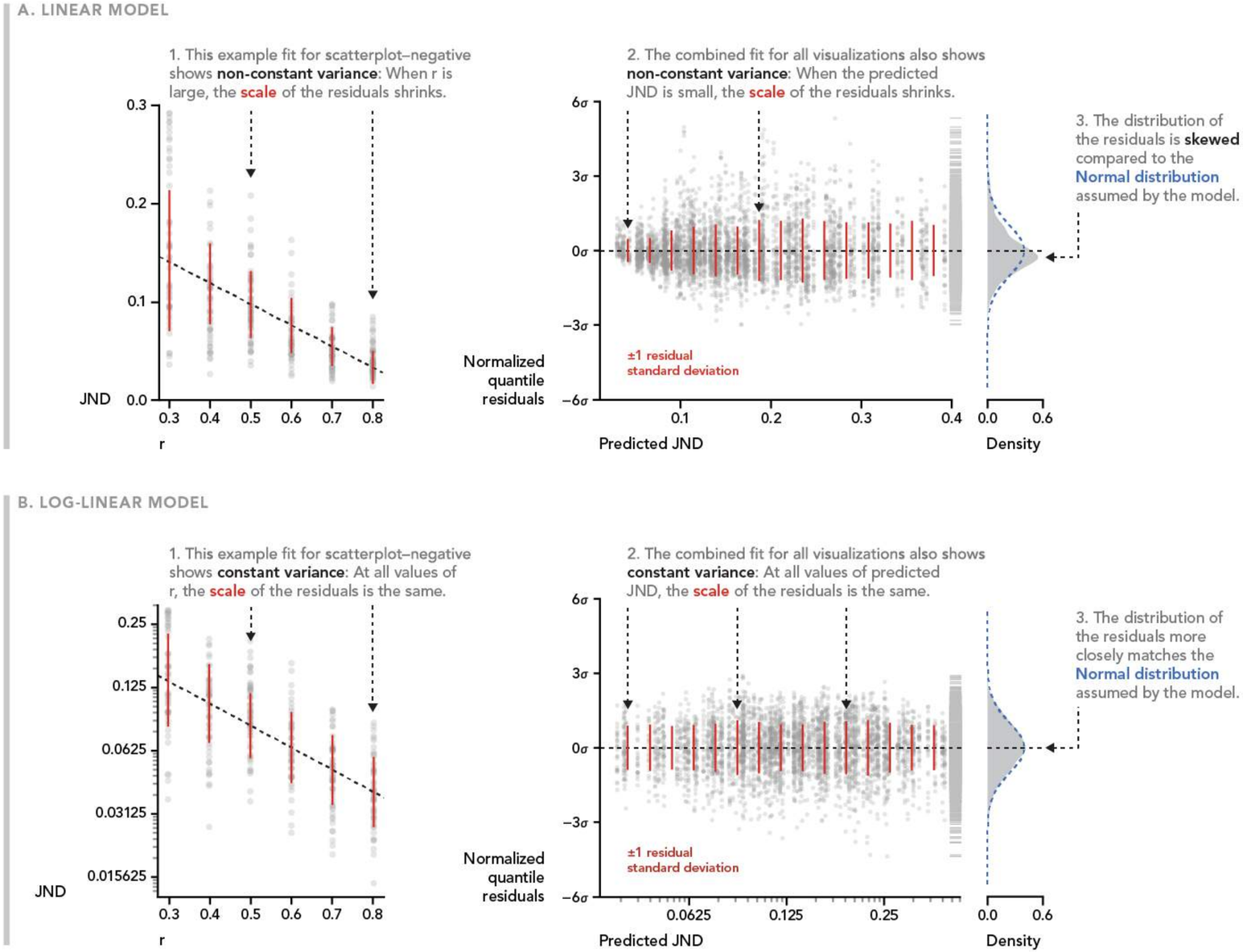
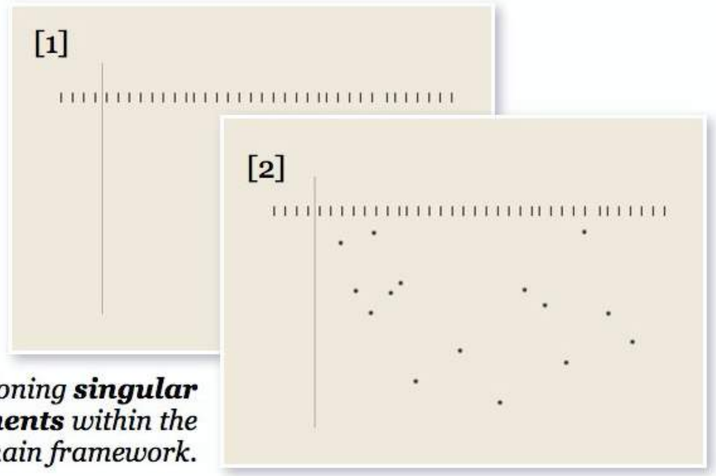


Fig. 3 Comparison of fits of the linear model (Section 3) and the log-linear model (Section 4). Example fits of each model to scatterplot–negative are shown in A.1 and B.1. Plots of normalized residuals for all visualization × direction pairs are shown in A.2 and B.2. Density plots of normalized residuals with comparison to the standard normal distribution are shown in A.3 and B.3.

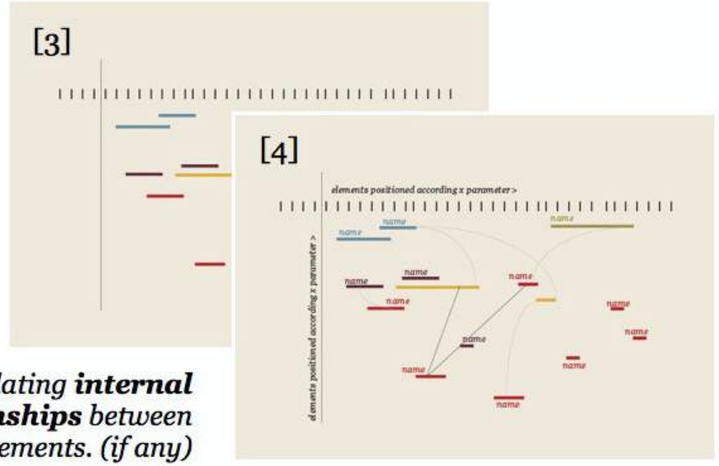
Lupi, Giorgia

*Composing the main **architecture** of the visualization*



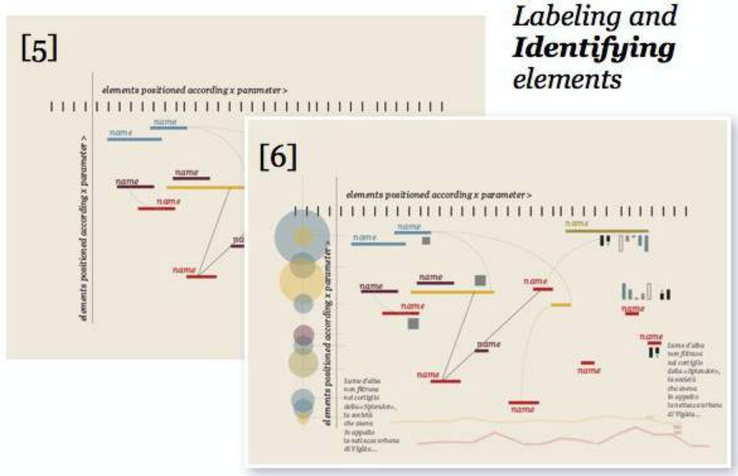
*Positioning **singular elements** within the main framework.*

*Constructing shaped elements of **dimensionality and form***



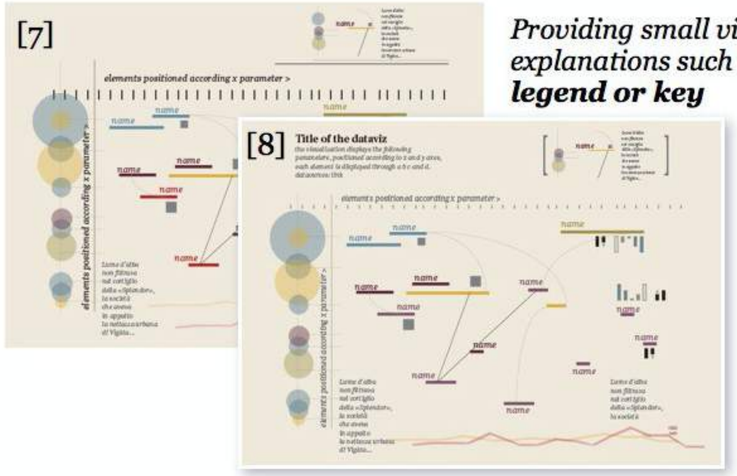
*Elucidating **internal relationships** between elements. (if any)*

*Labeling and **Identifying** elements*



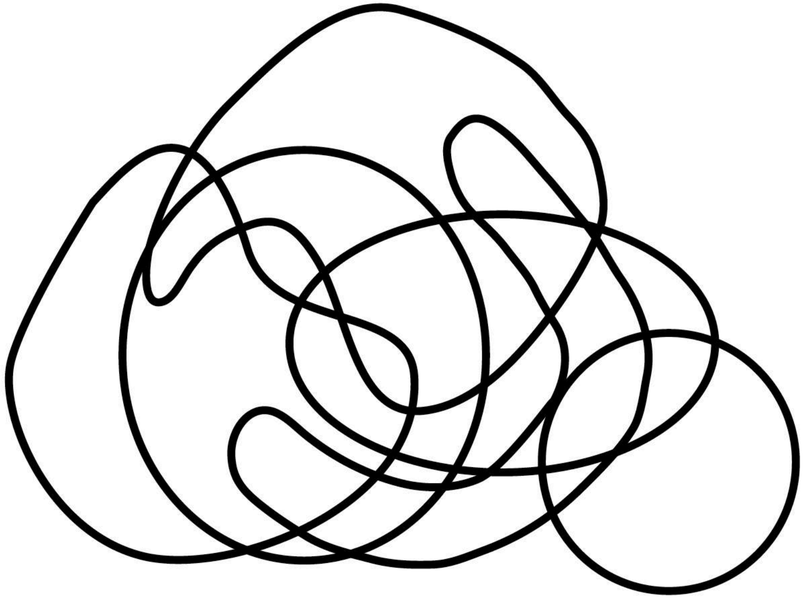
*Supplementing the greater story through the addition of "**minor or tangental tales**" elements*

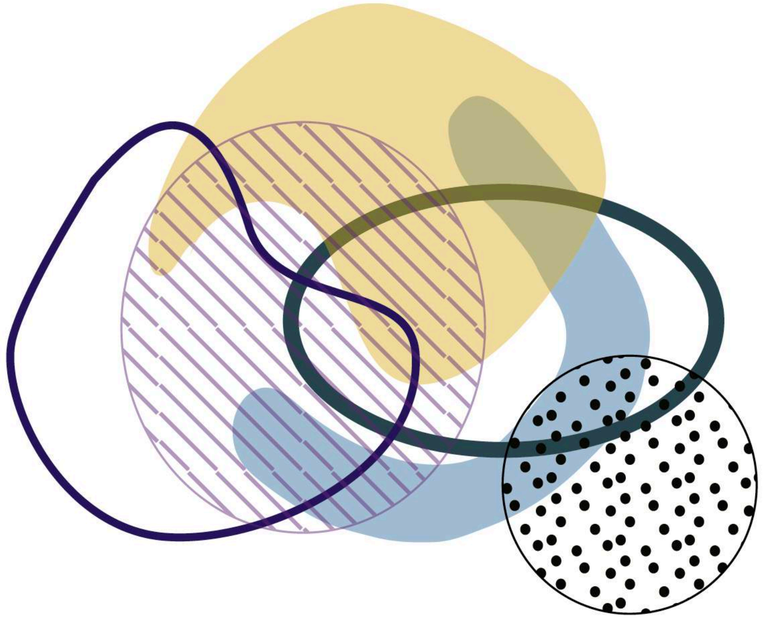
*Providing small visual explanations such as a **legend or key***

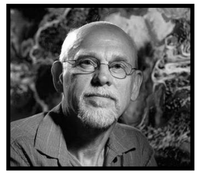


***Fine-tuning** and stylizing of elements shapes, colors, and weights to make hierarchies pop out.*

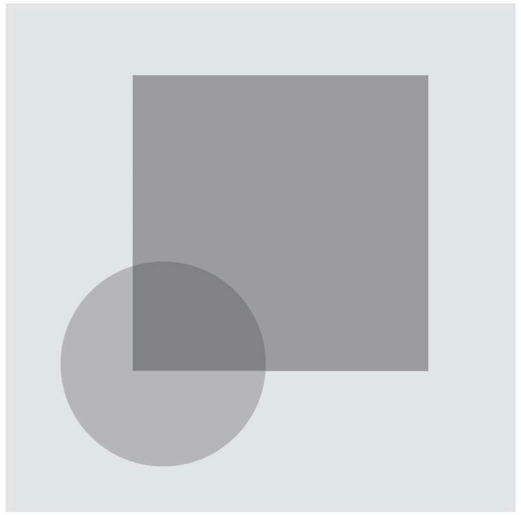
 in this Euler diagram, color and texture help to distinguish layers of information





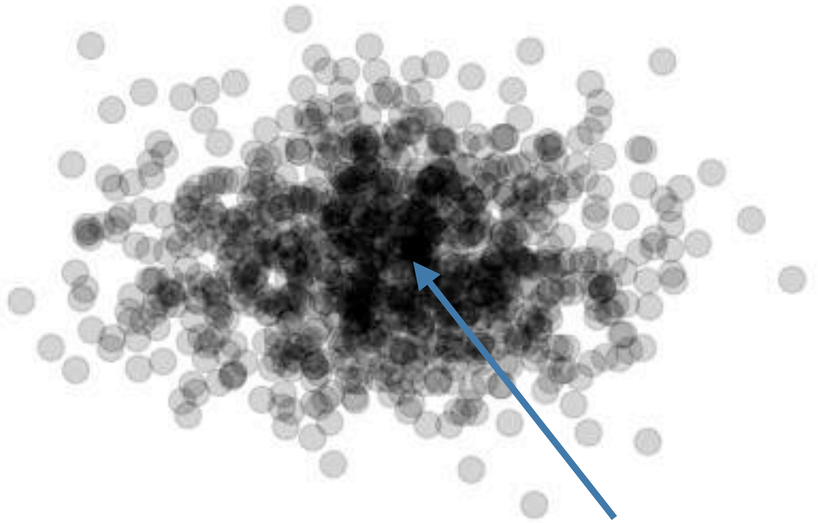


Ware, Colin



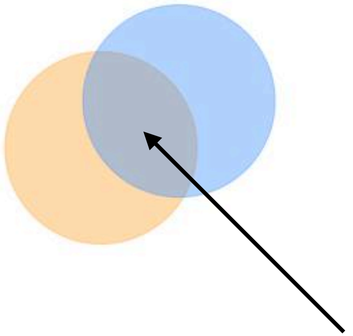
Samara, Timothy

```
x <- rnorm(1000)
y <- rnorm(1000)

ggplot() +
  theme_void() +
  scale_x_continuous(limits = c(-5, 5)) +
  scale_y_continuous(limits = c(-5, 5)) +
  geom_point(aes(x = x, y = y),
             size = 4, color = "black",
             alpha = 0.2)
```



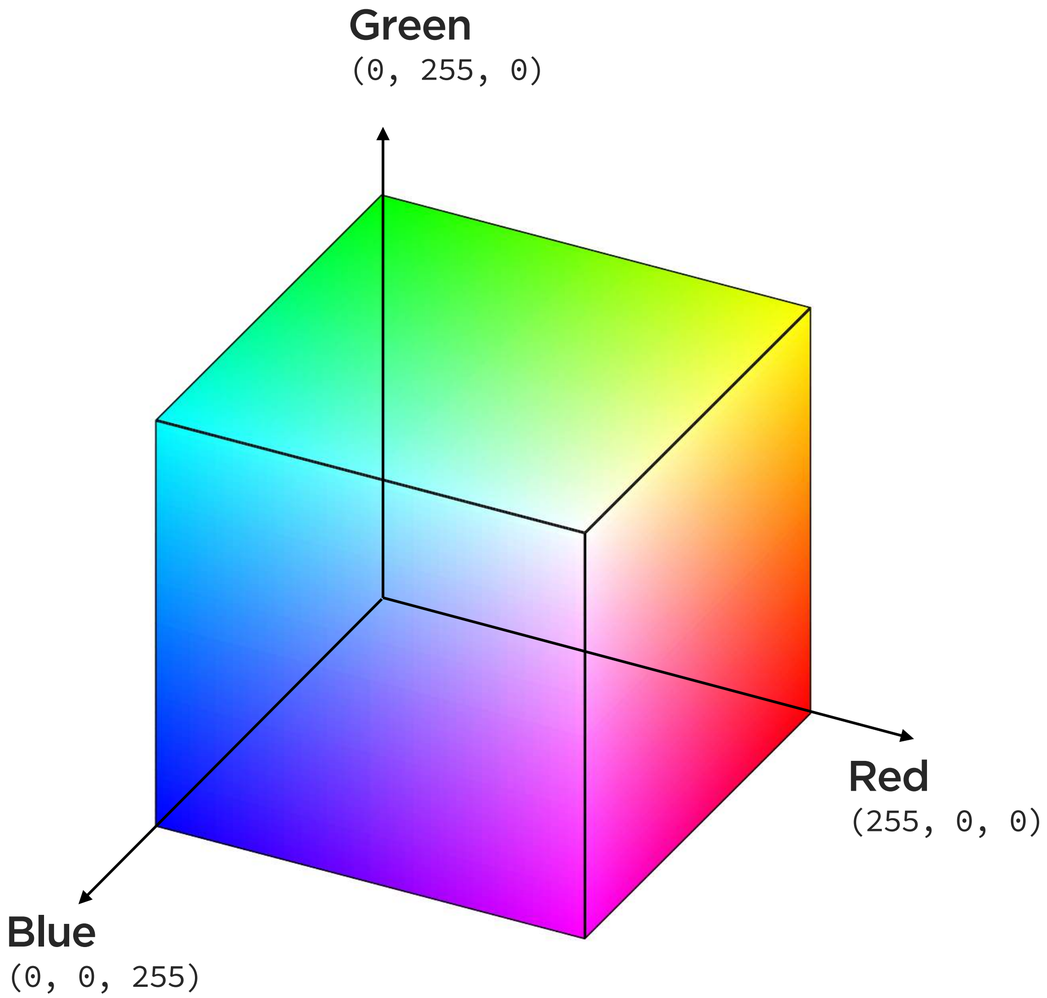
dense areas of data are darker

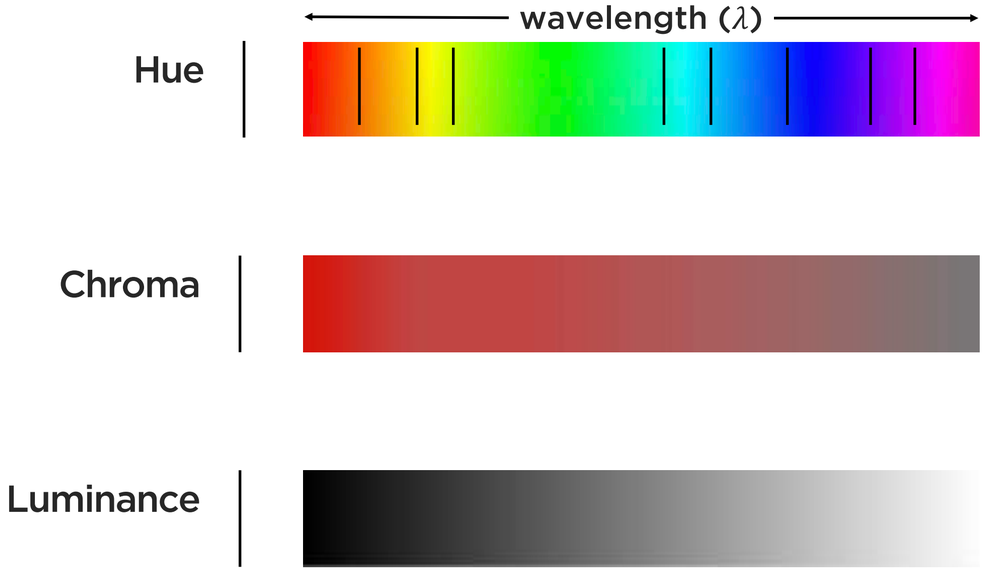
```
ggplot() +
  theme_void() +
  scale_x_continuous(limits = c(-5, 5)) +
  scale_y_continuous(limits = c(-5, 5)) +
  geom_point(aes(x = 0, y = 0),
             size = 50, color = "orange",
             alpha = 0.4) +
  geom_point(aes(x = 1, y = 1),
             size = 50, color = "dodgerblue",
             alpha = 0.4)
```



I didn't encode data with *this* color!?

encoding data as color

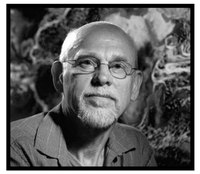

Green
(0, 255, 0)
Red
(255, 0, 0)
Blue
(0, 0, 255)


wavelength (λ)
Hue
Chroma
Luminance

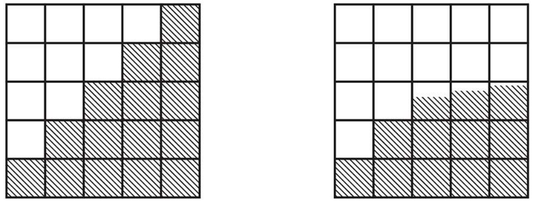

$f(x)$
75
50
25
66
50
51
94
29
12
33
56
40
66
12
64
52
17
32
35
84
60
32
64

LUMINANCE : the *measured* amount of light coming from some region of space.

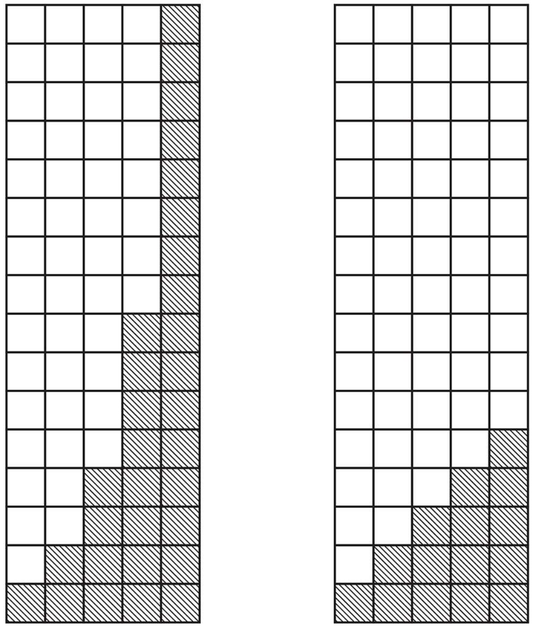
BRIGHTNESS : the *perceived* amount of light coming from that region of space.



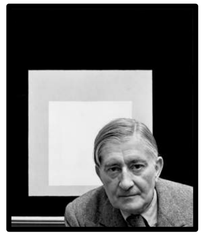
Ware, Colin



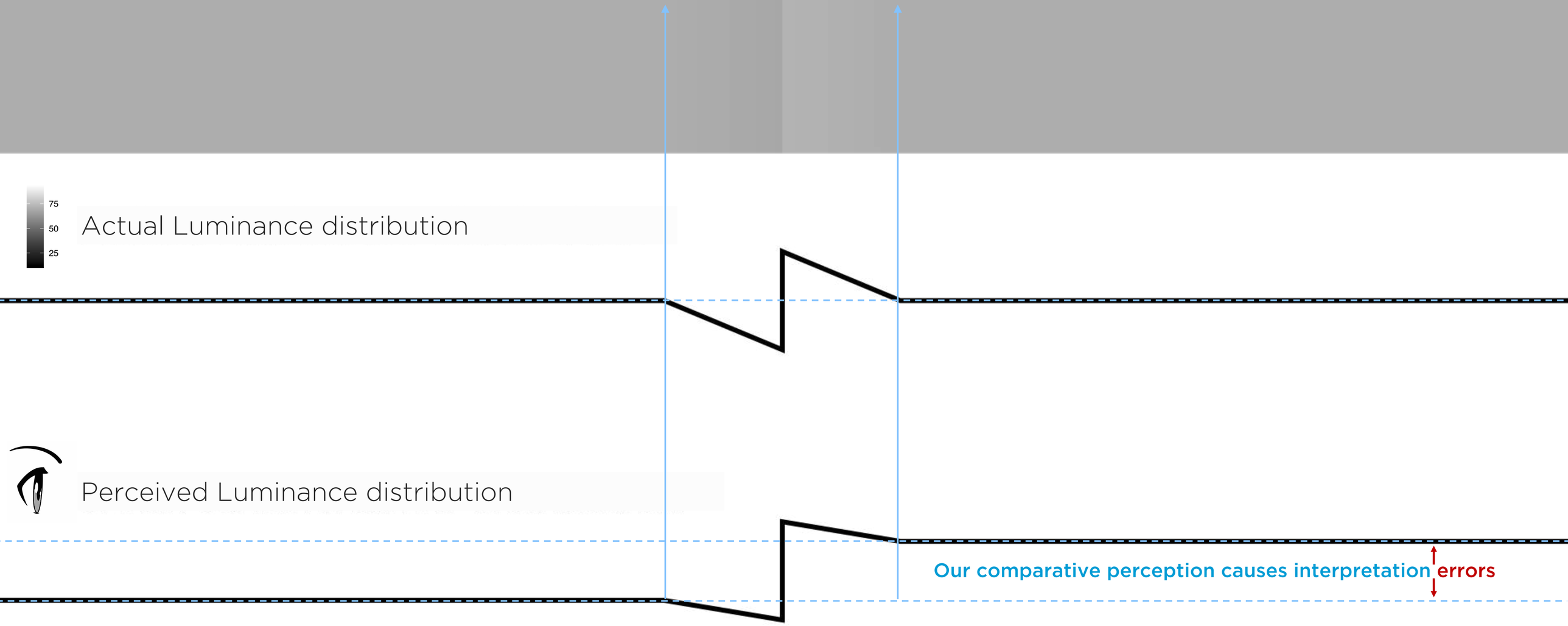
THIS PHYSICAL FACT REDUCES TO THIS PSYCHOLOGICAL EFFECT

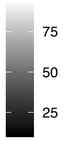


THIS PHYSICAL FACT PRODUCES THIS PSYCHOLOGICAL EFFECT



Albers, Josef


75
50
25
Actual Luminance distribution
Perceived Luminance distribution
Our comparative perception causes interpretation errors



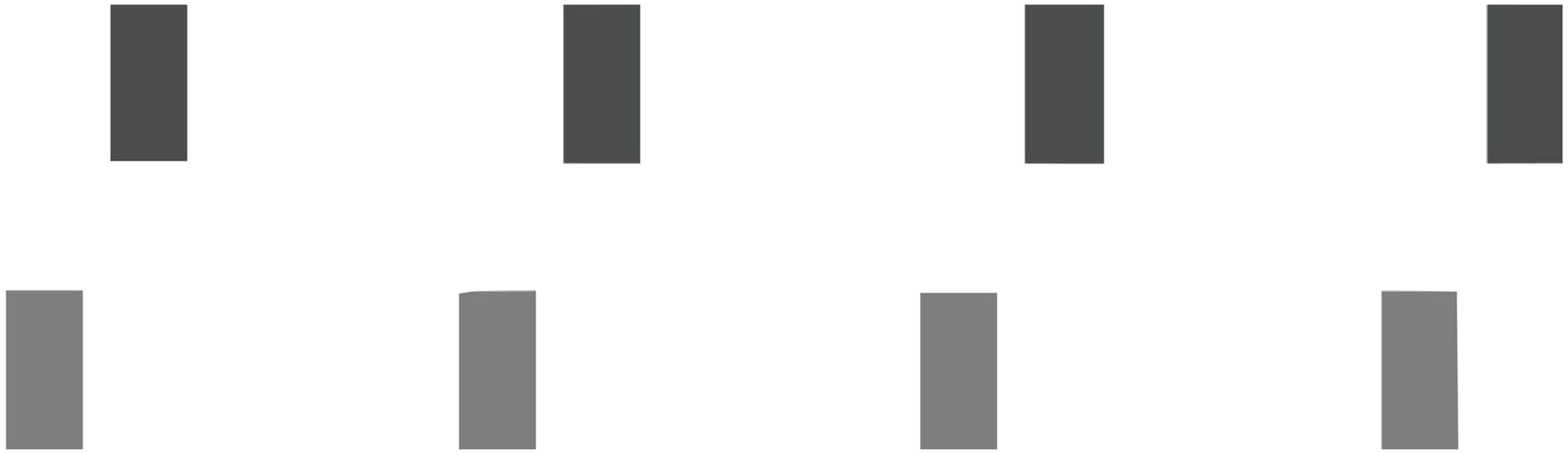
Actual Luminance distribution

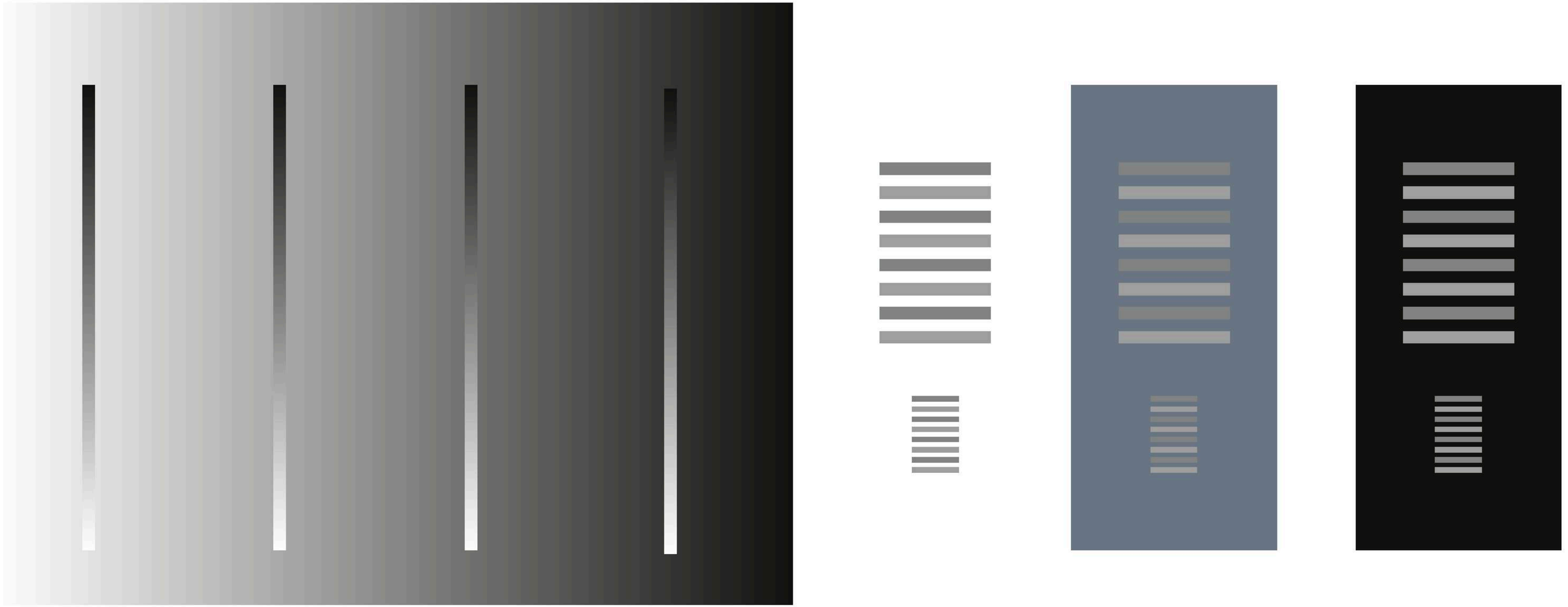
Perceived Luminance distribution

Our comparative perception causes interpretation errors

 background and adjacent luminance can interfere with our perception

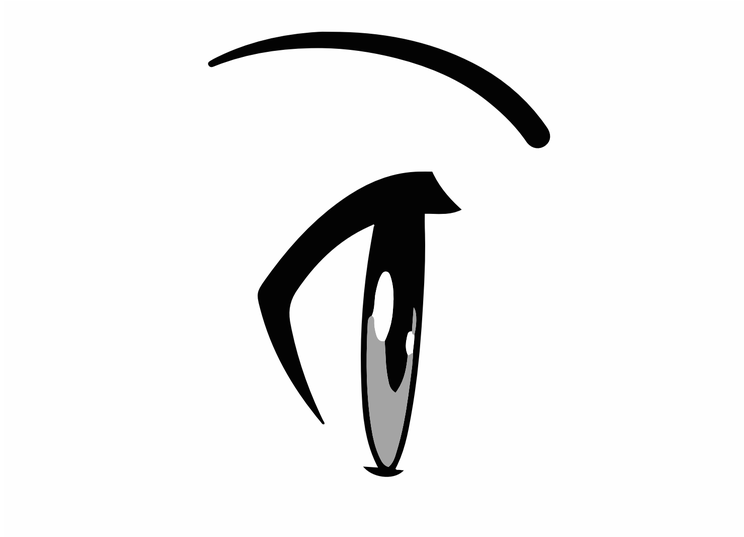
Ware, Colin

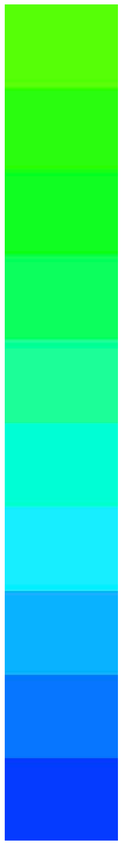




Ware, Colin

 as with luminance, hue values in the RGB color space fail to uniformly scale across values.





Same luminance or lightness?

HSL(250, 100%, 50%)
HSL(250, 100%, 50%)

HSL(60, 100%, 50%)
HSL(60, 100%, 50%)

Boronine, Alexei

Same saturation?

HSL(0, 90%, 40%)
HSL(0, 90%, 40%)

HSL(0, 90%, 80%)
HSL(0, 90%, 80%)

Boronine, Alexei

Equal difference between hues?

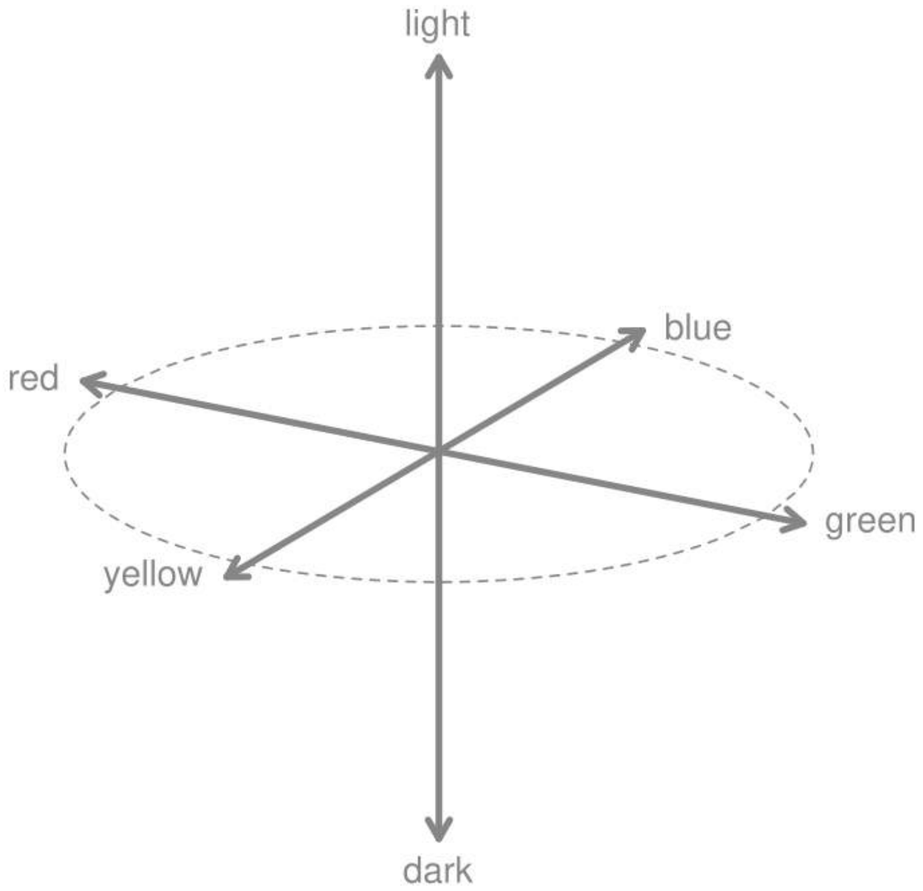
HSL(30, 100%, 50%)
HSL(30, 100%, 50%)

HSL(50, 100%, 50%)
HSL(50, 100%, 50%)

HSL(230, 100%, 50%)
HSL(230, 100%, 50%)

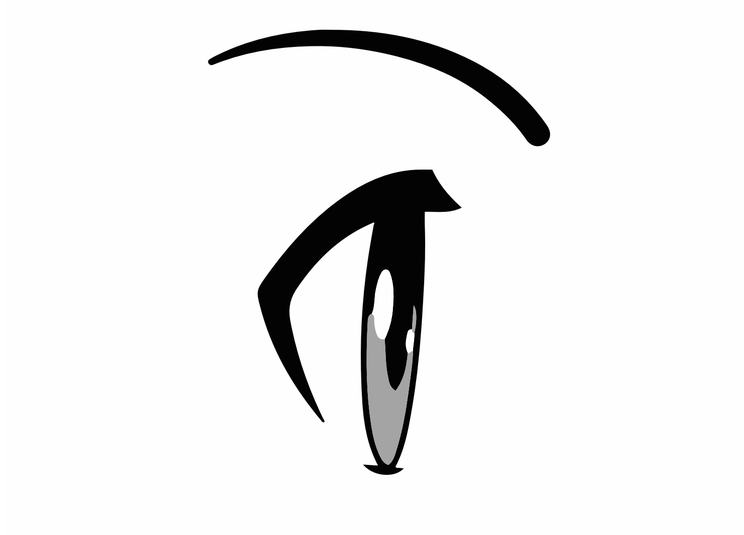
HSL(250, 100%, 50%)
HSL(250, 100%, 50%)

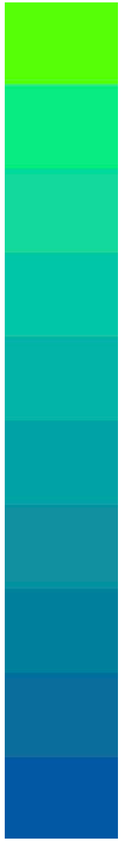
Boronine, Alexei



The International Commission on Illumination (CIE) studied human perception and re-mapped color into a space where we perceive color changes uniformly.

Their **CIELuv** color model has two dimensions — u and v — that represent color scales from red to green and yellow to blue.





Load functions for mapping color
https://www.hsluv.org/implementations/

```
url_hsluv_py <-
  paste0('https://raw.githubusercontent.com/',
         'hsluv/hsluv-python/master/hsluv.py')

reticulate::source_python(url_hsluv_py)
```

Create sample data encoded as hue, saturation, luminance

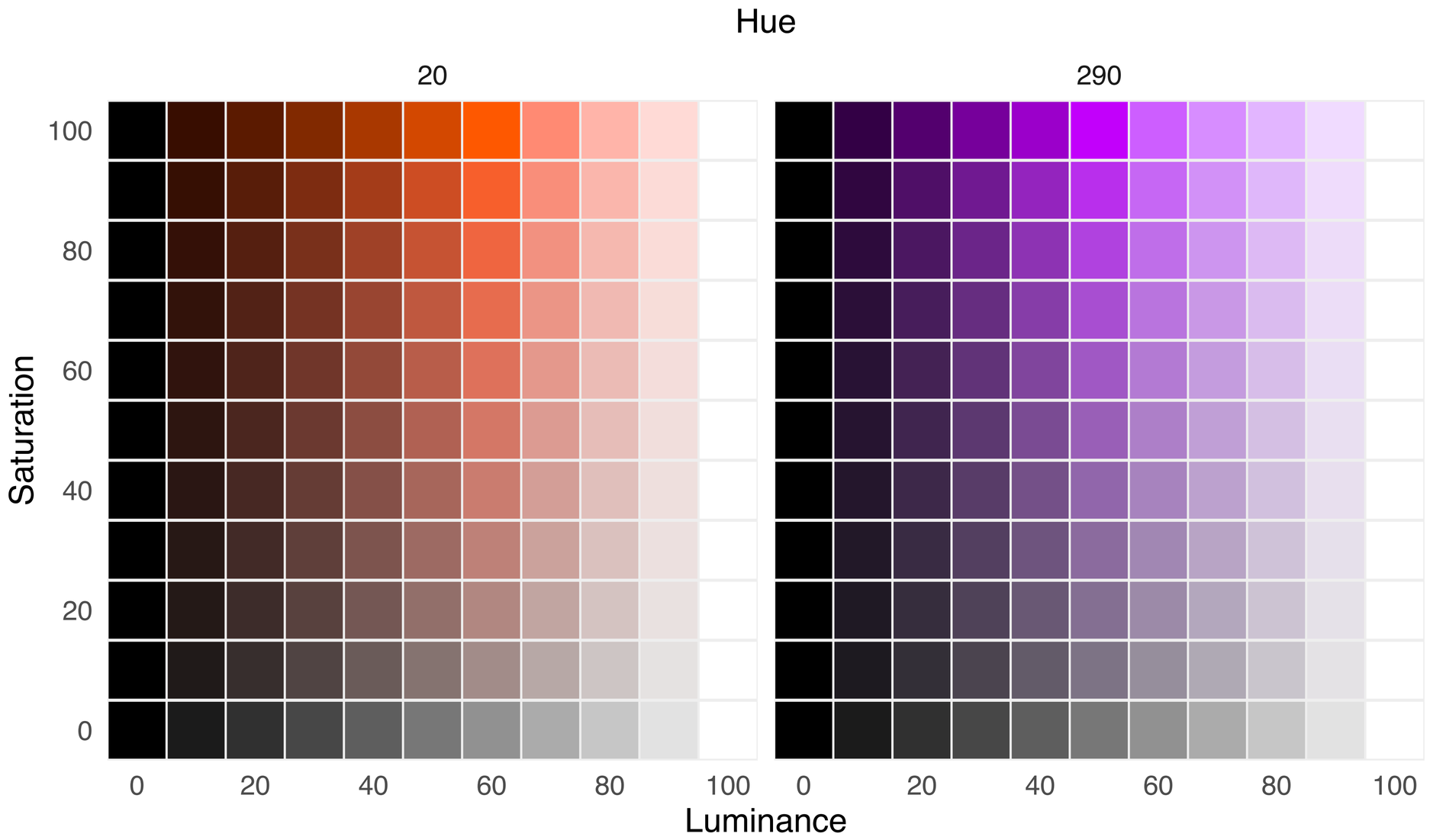
```
df <- expand.grid(H = c(20, 290),
                  S = seq(0, 100, by = 10),
                  L = seq(0, 100, by = 10))
```

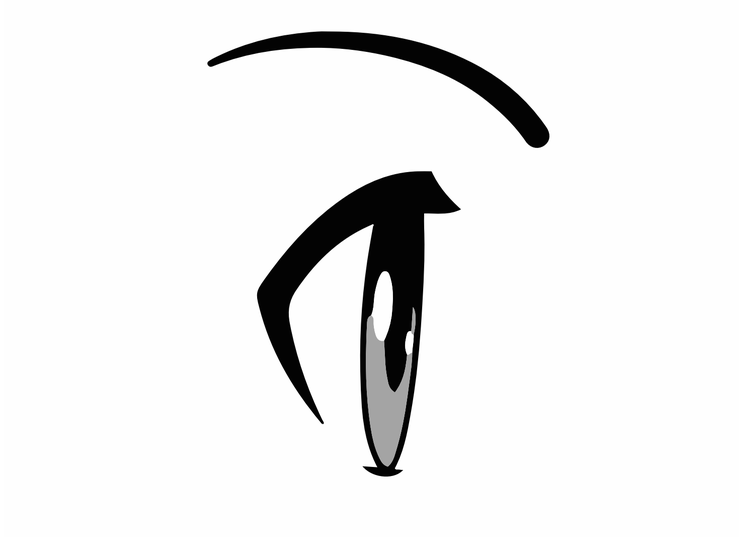
Convert HSLuv scaled values to RGB
color space as hex code #RRGGBB

```
for(i in seq(NROW(df)))
  df$colors <- with(df[i,], hsluv_to_hex(c(H, S, L)) )
```

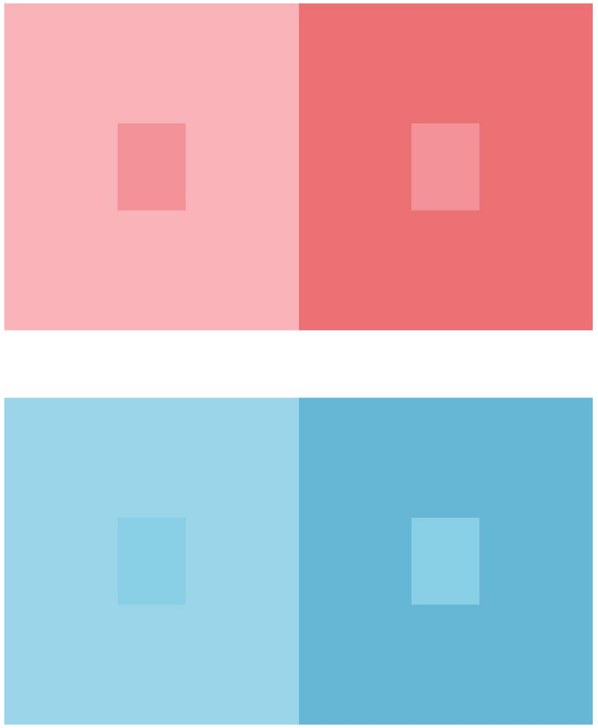
Plot data encoded as colors

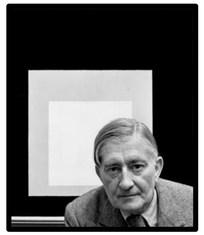
```
library(ggplot2)

ggplot(df) +
  theme_minimal() +
  theme(panel.grid = element_blank(),
        axis.text.x.top = element_blank()) +
  geom_point(aes(L, S),
             color = '#eeeeee',
             fill = df$colors,
             size = 10,
             shape = 22) +
  scale_x_continuous(breaks = seq(0, 100, by = 20),
                     sec.axis = sec_axis(~., name = 'Hue')) +
  scale_y_continuous(breaks = seq(0, 100, by = 20)) +
  facet_wrap(~H) +
  labs(x = 'Luminance',
       y = 'Saturation')
```


Hue
20
290
Saturation
100
80
60
40
20
0
0
20
40
60
80
100
Luminance

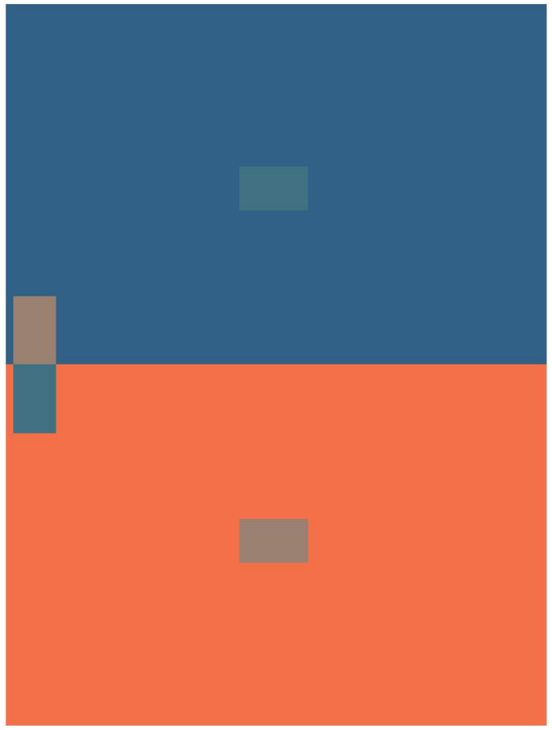


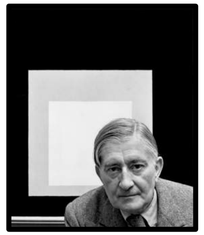
interaction of color



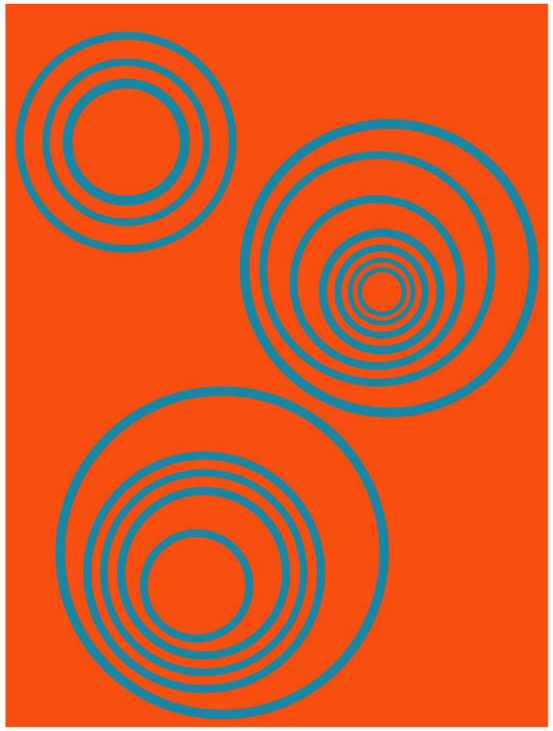


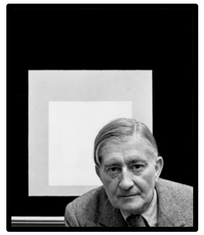
Albers, Josef





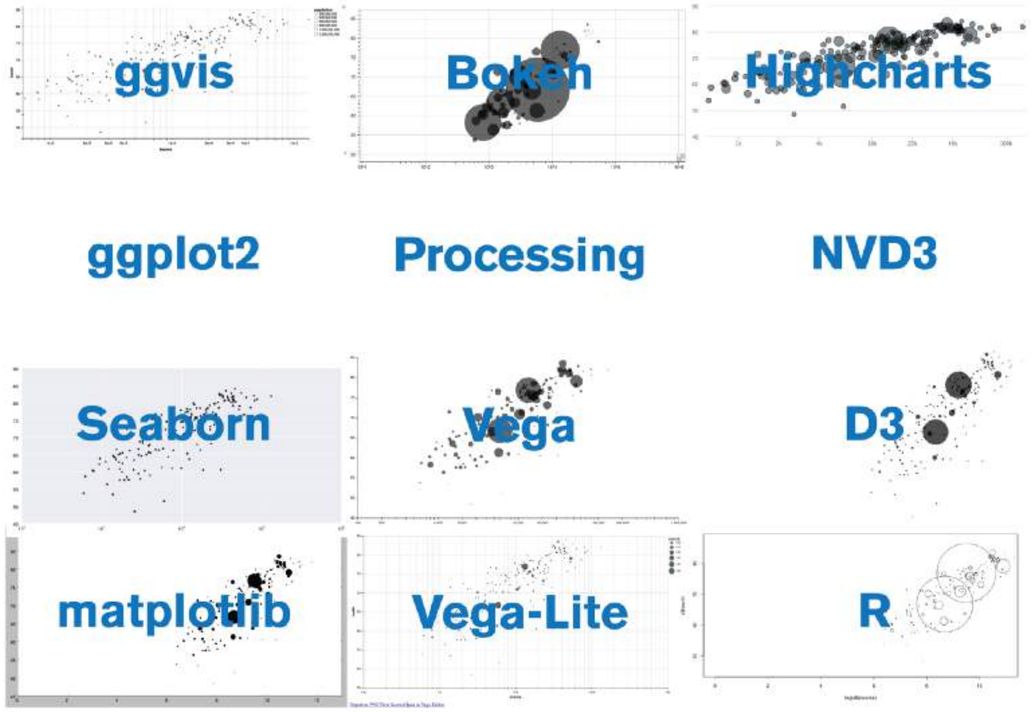
Albers, Josef

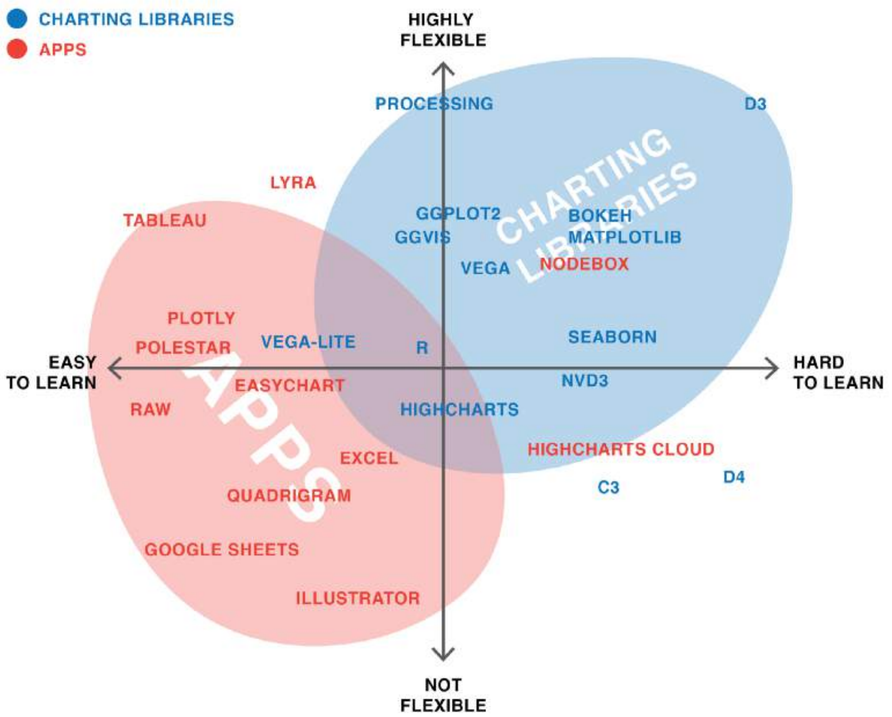




Albers, Josef

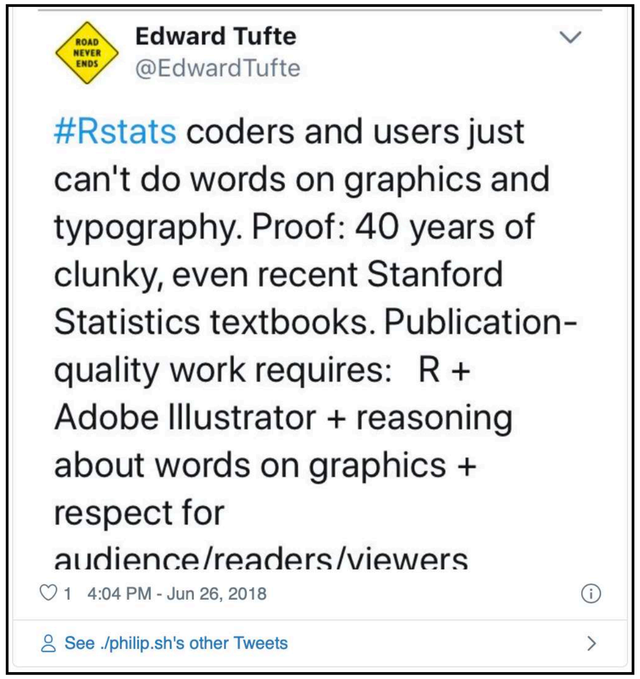
perspectives on tools

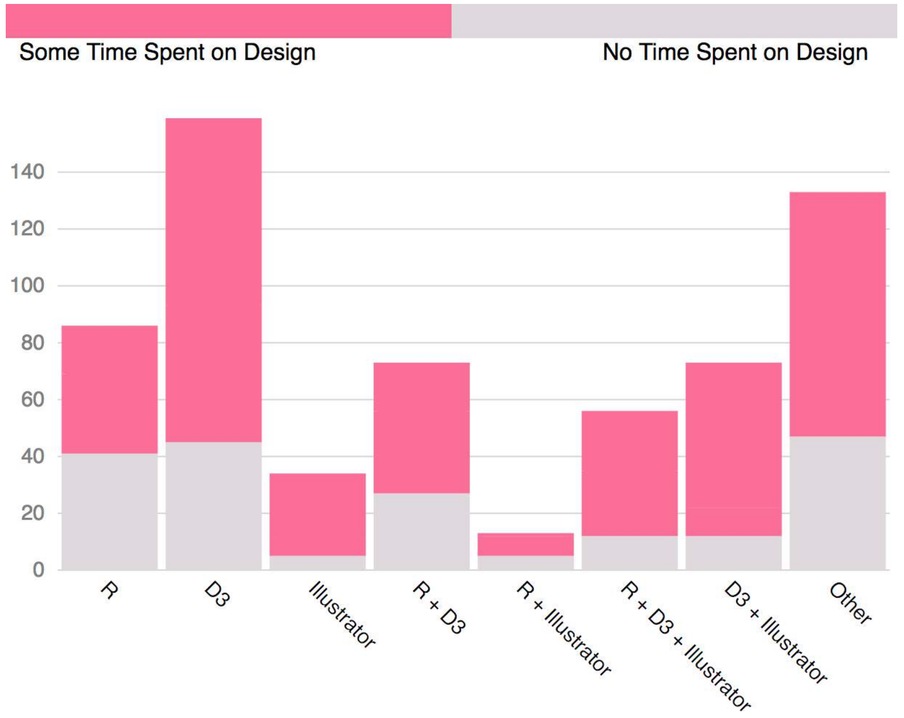




"There are no perfect tools, just good tools for people with certain goals."

Rost, Lisa Charlotte



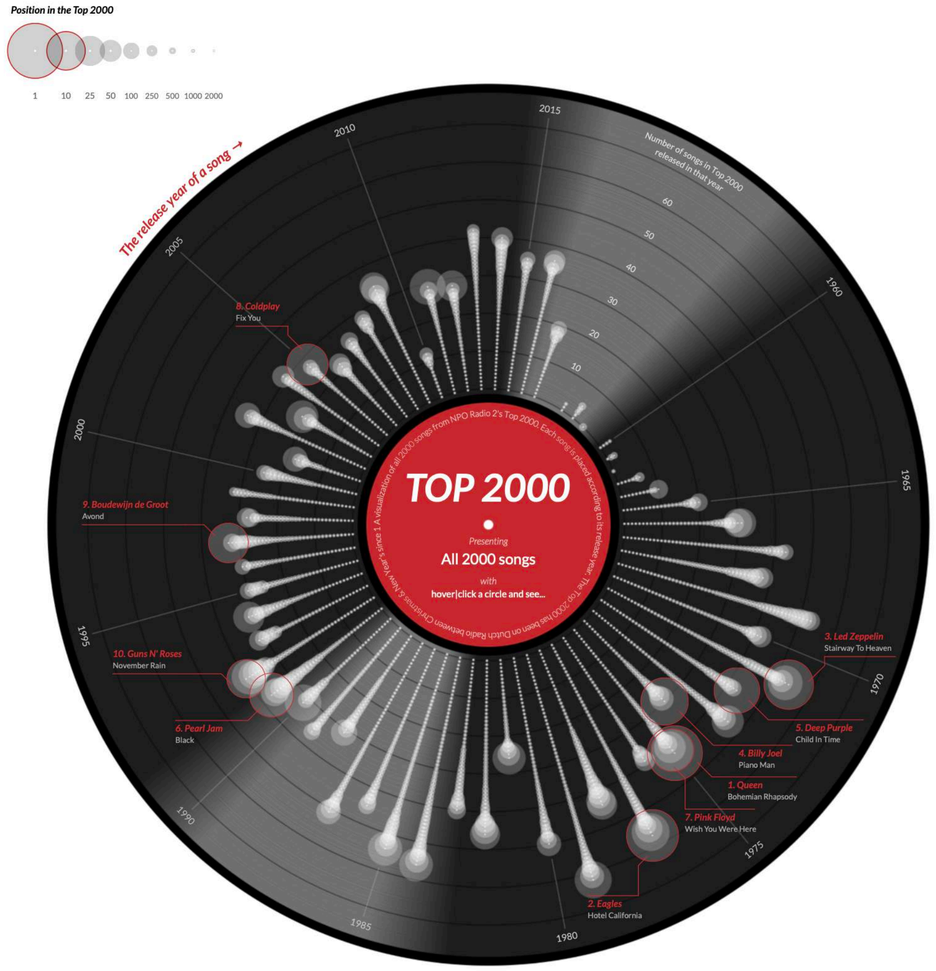


Meeks, Elijah

practice in the studio

Let The Music Play

All songs from the TOP 2000 of 2017 according to their release years



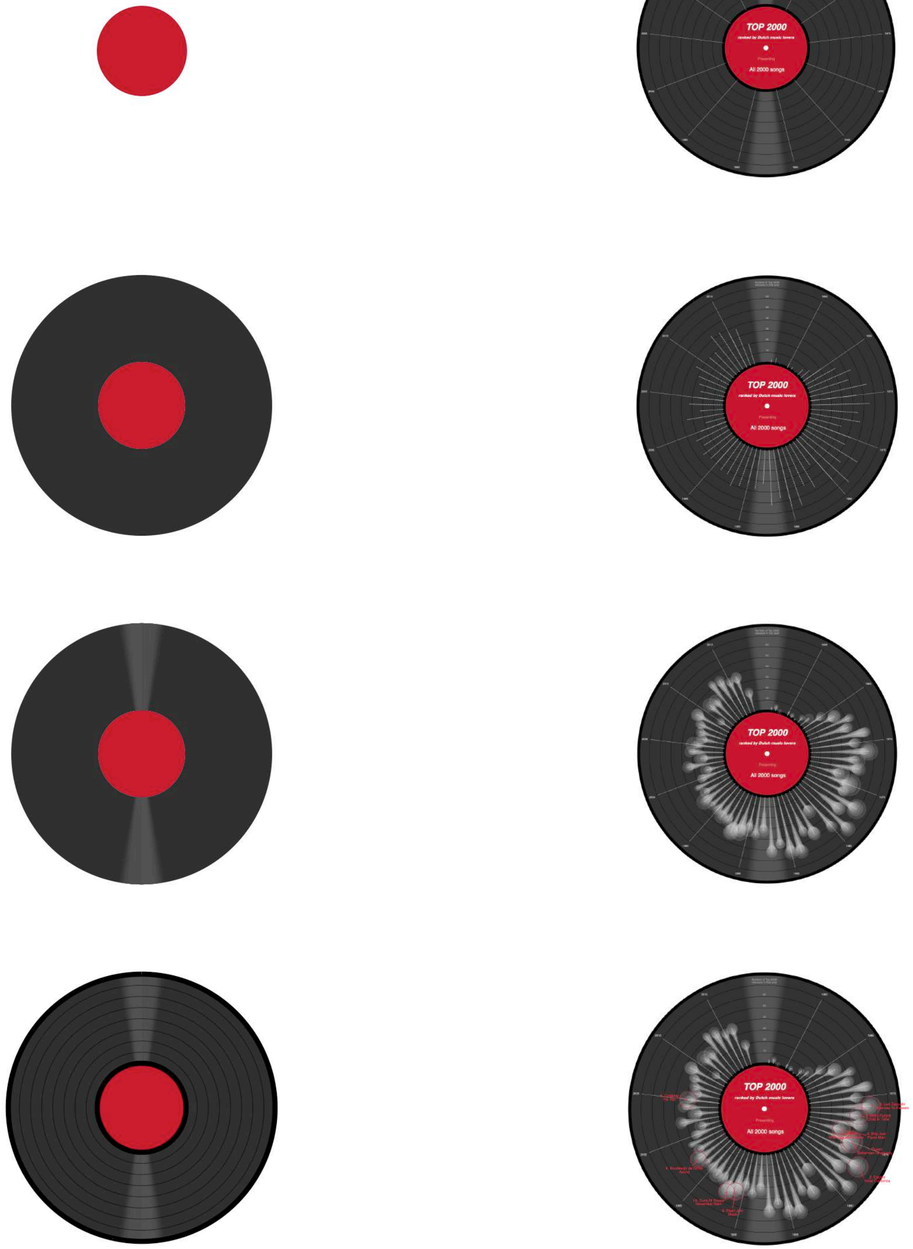
When hovering (or clicking on mobile) a circle all the songs of that artist are highlighted and connected by a line. When the song happens to be a collaboration between two artists, you'll see multiple lines, one for each artist.

Created by Nadieh Bremer | Visual Cinnamon

Data from NPO Radio 2

Ga naar de Nederlandse versie

Check out my Top 2000 visuals from 2015 & 2016



References

Albers, Josef. *Interaction of Color*. Yale University Press, 2006, and interactive app on iPad: http://interactionofcolor.com

Bremer, Nadieh. "*Let the Music Play*." Github pages. Accessed March 26, 2020. https://nbremer.github.io/top2000vinyl/.

Boronine, Alexei. "*Color Spaces for Human Beings*." HSLuv.org (blog), March 26, 2012. https://www.hsluv.org and https://www.boronine.com/2012/03/26/Color-Spaces-for-Human-Beings/.

Kay, Matthew, and Jeffrey Heer. "*Beyond Weber's Law: A Second Look at Ranking Visualizations of Correlation*." IEEE Transactions on Visualization and Computer Graphics 22, no. 1 (January 31, 2016): 469–78.

Lupi, Giorgia. "*The Architecture of a Data Visualization: Multilayered Storytelling through 'Info-Spatial' Compositions*," February 2015. https://medium.com/accurat-studio/the-architecture-of-a-data-visualization-470b807799b4.

Meeks, Elijah. "*You Can Design a Good Chart with R: But Do R Users Invest in Design?*," July 2, 2018. https://medium.com/nightingale/you-can-design-a-good-chart-with-r-5d00ed7dd18e.

———. "*Non-Linear Storytelling: Journalism through Info-Spatial Compositions*." Parsons Journal for Information Mapping IV, no. 4 (October 2012): 1–11.

Rost, Lisa Charlotte. "*One Chart, Nine Tools – Revisited*." Lisa Charlotte Rost (blog), October 18, 2018. https://lisacharlotterost.de/datavistools-revisited.

———. "*What I Learned Recreating One Chart Using 24 Tools*." Code in Journalism. Source (blog), December 8, 2016. https://source.opennews.org/articles/what-i-learned-recreating-one-chart-using-24-tools/.

Samara, Timothy. *Design Elements: A Graphic Style Manual. Understanding the Rules and Knowing When to Break Them*. Rockport, 2014.

Spencer, Scott. "*Demonstration of Layers in Graphics.*" Columbia University., March 6, 2020.

Tufte, Edward R. "*Layers and Separation*." In Envisioning Information. Graphics Press, 1990.

———. "*Color and Information*." In Envisioning Information. Graphics Press, 1990.

Ware, Colin. *Information Visualization: Perception for Design*. Fourth. Philadelphia: Elsevier, Inc, 2020.

Wickham, Hadley. "*A Layered Grammar of Graphics*." Journal of Computational and Graphical Statistics 19, no. 1 (January 2010): 3–28.

Wilkinson, Leland. *The Grammar of Graphics*. Second. Springer, 2005.