

SCIENTIFIC FIGURE DESIGN

Guillaume Lobet

g.lobet@fz-juelich.de

<http://bit.ly/datavizualisation>



THE GOOD THE BAD AND THE UGLY

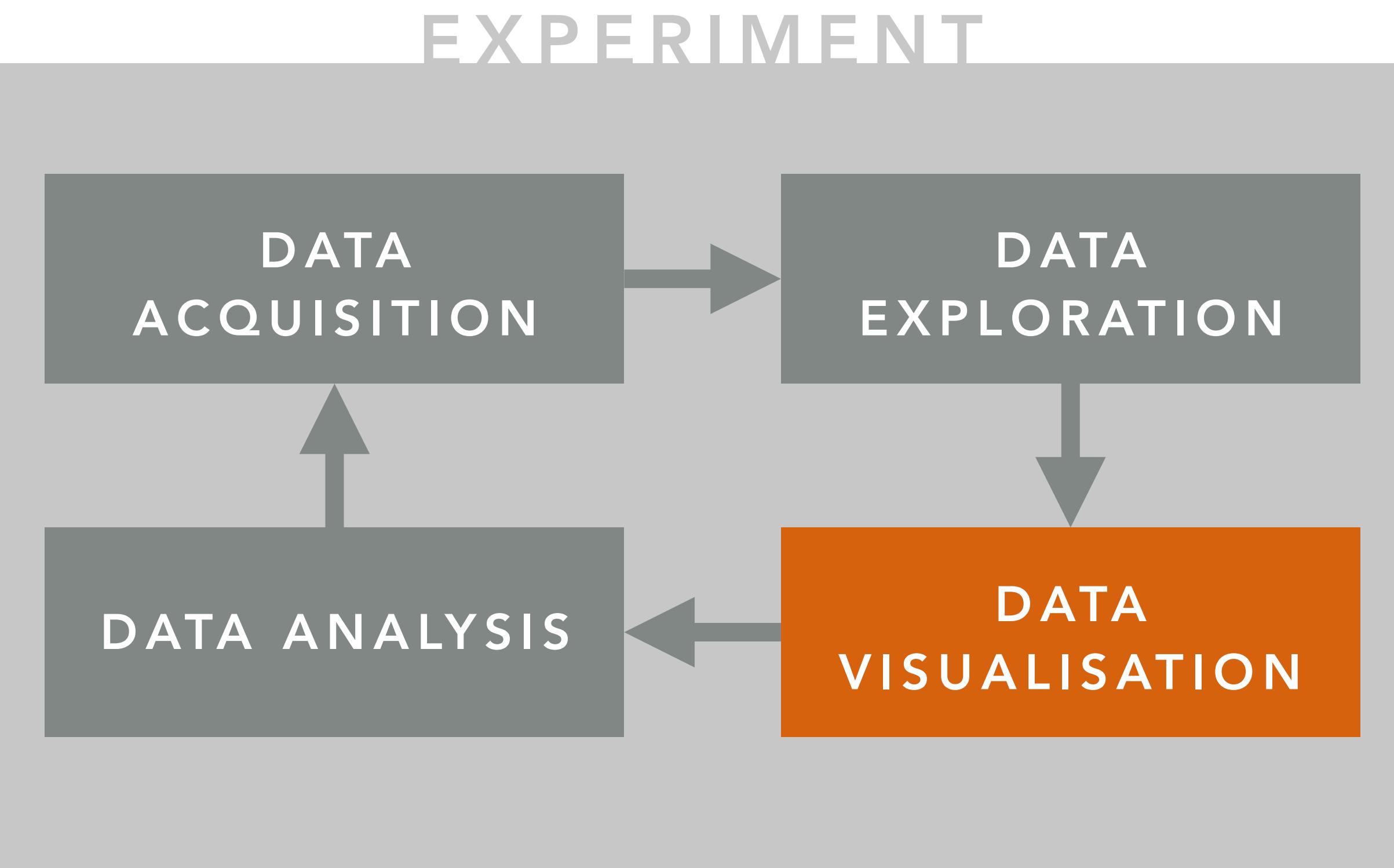
UCL
Université
catholique
de Louvain

 **JÜLICH**
FORSCHUNGSZENTRUM

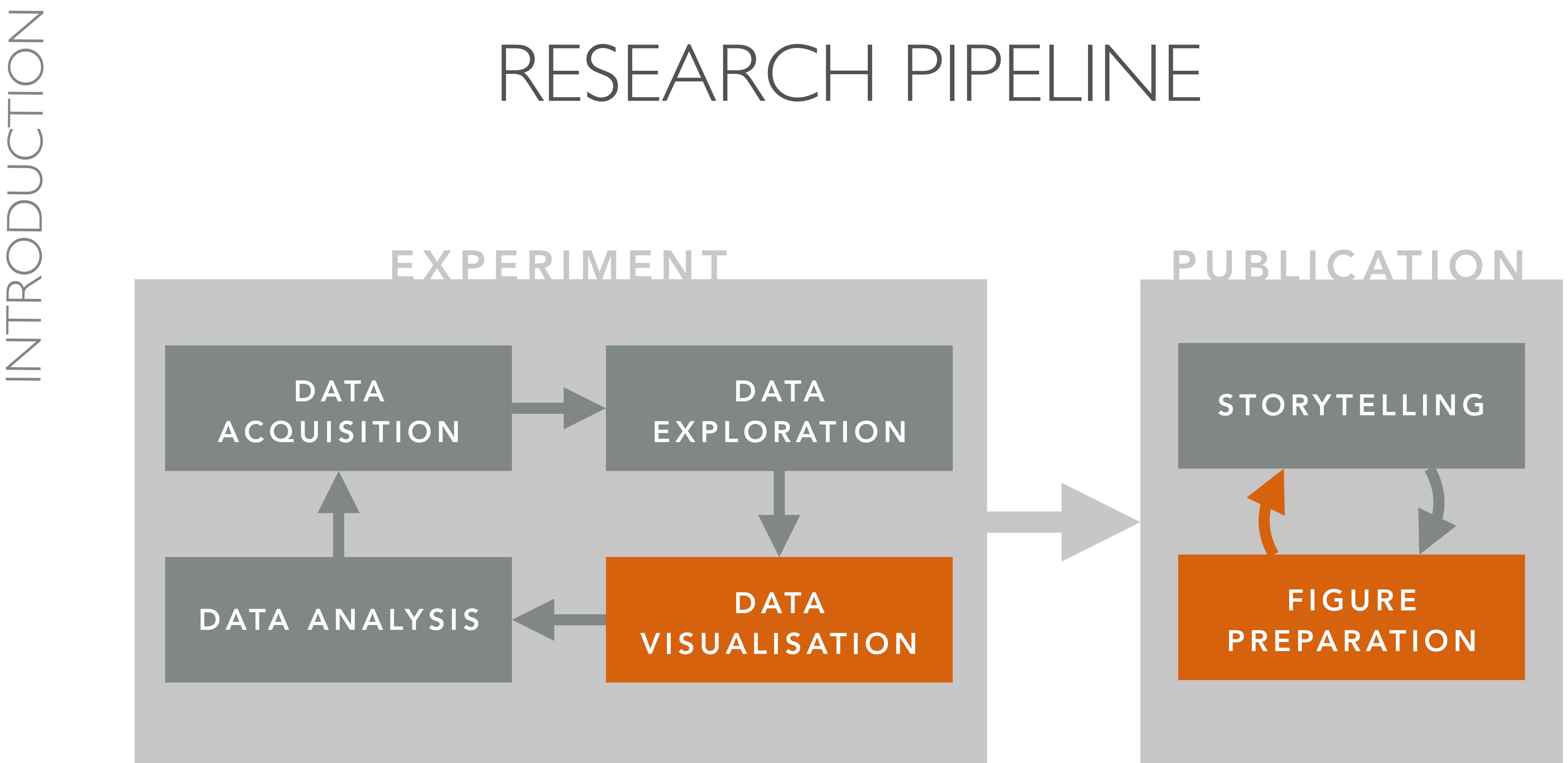
INTRODUCTION

RESEARCH PIPELINE

INTRODUCTION



RESEARCH PIPELINE



DATA GRAPHICS SHOULD:

- show the **data**
- induce viewer to think about **substance** rather than methodology
- encourage eye to **compare** different pieces of data
- avoid distorting **what the data represents**
- present **many numbers** in a small space
- make large data sets **coherent**
- reveal data at **several levels of detail** – broad overview and fine structure

INTRODUCTION

GRAPHICS REVEALS DATA AND PATTERNS

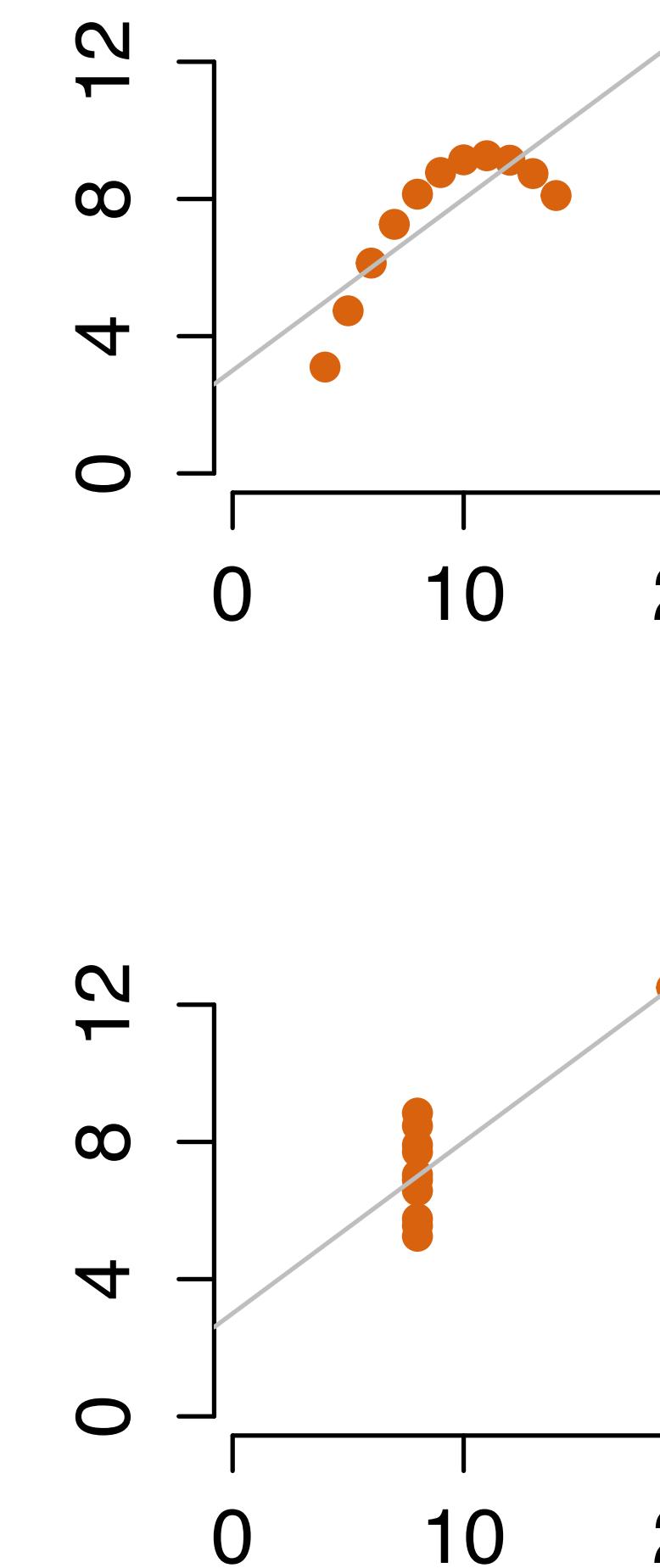
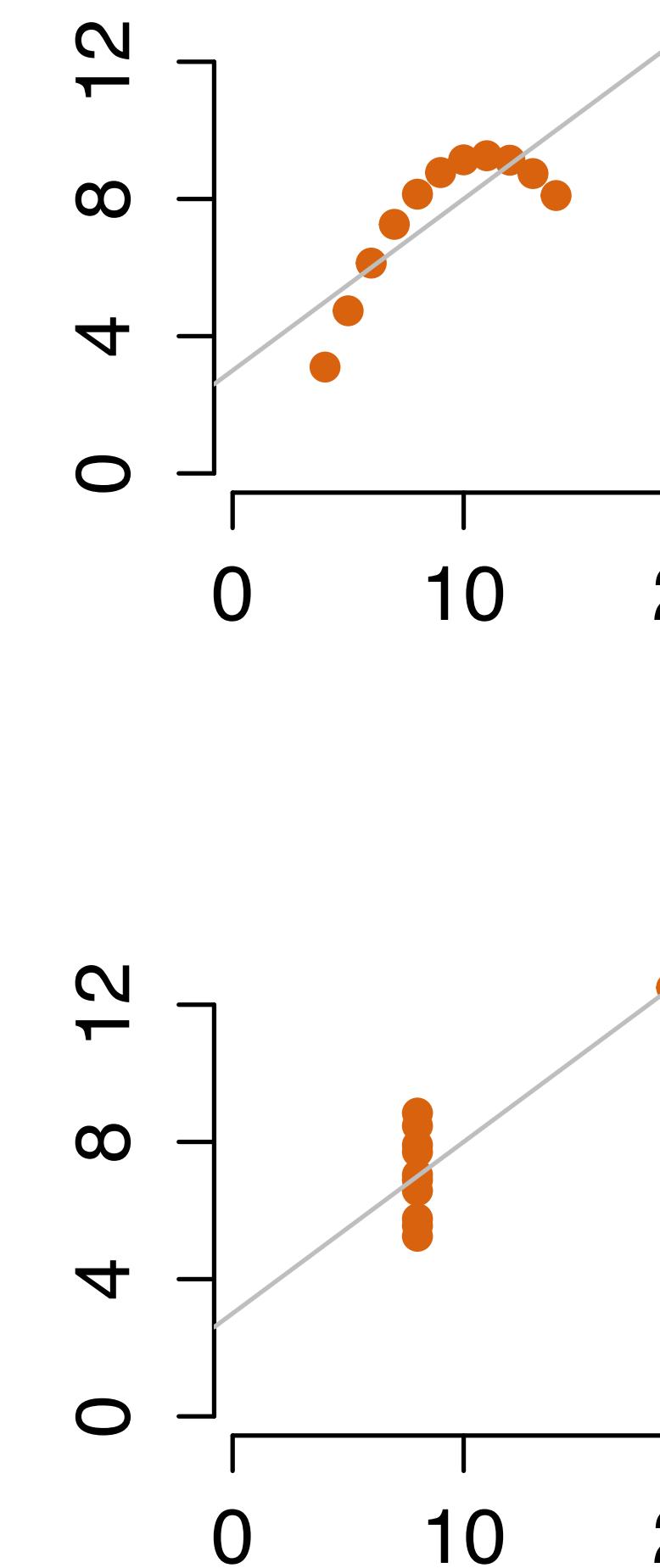
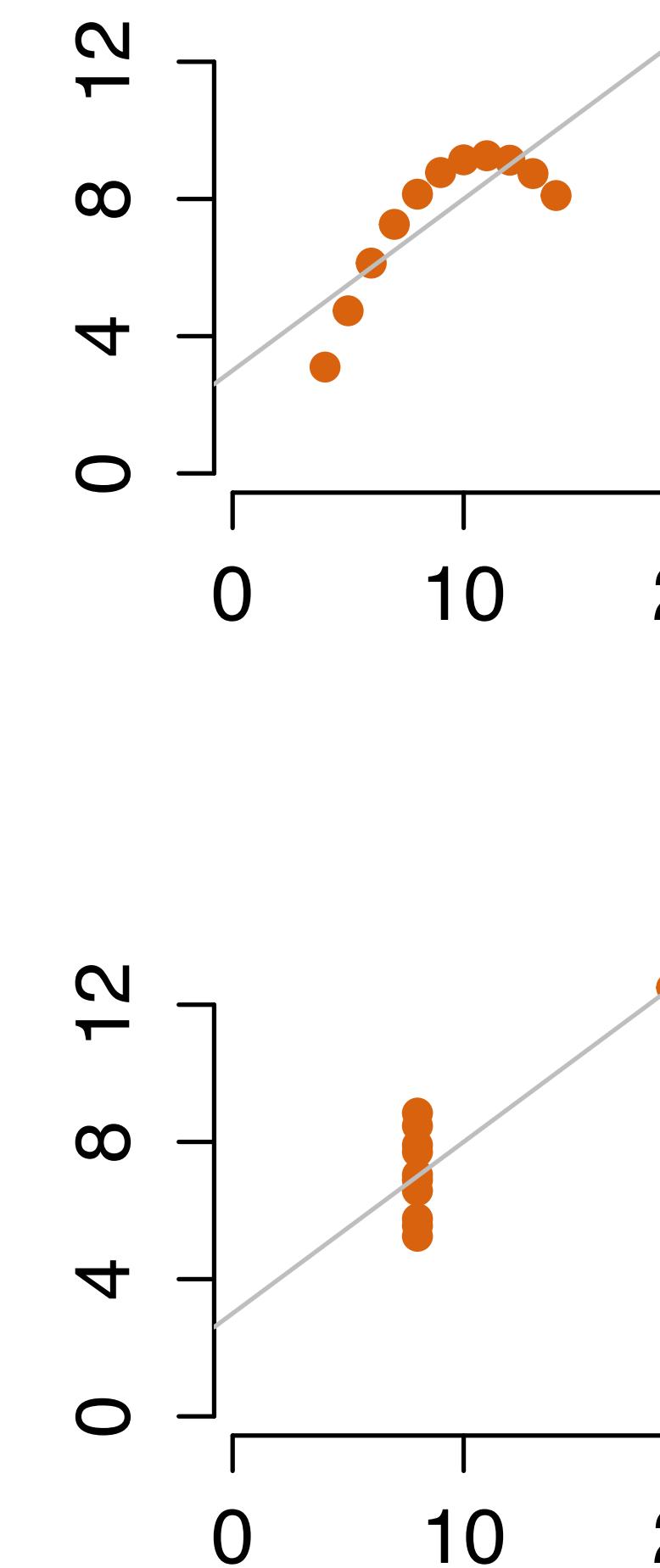
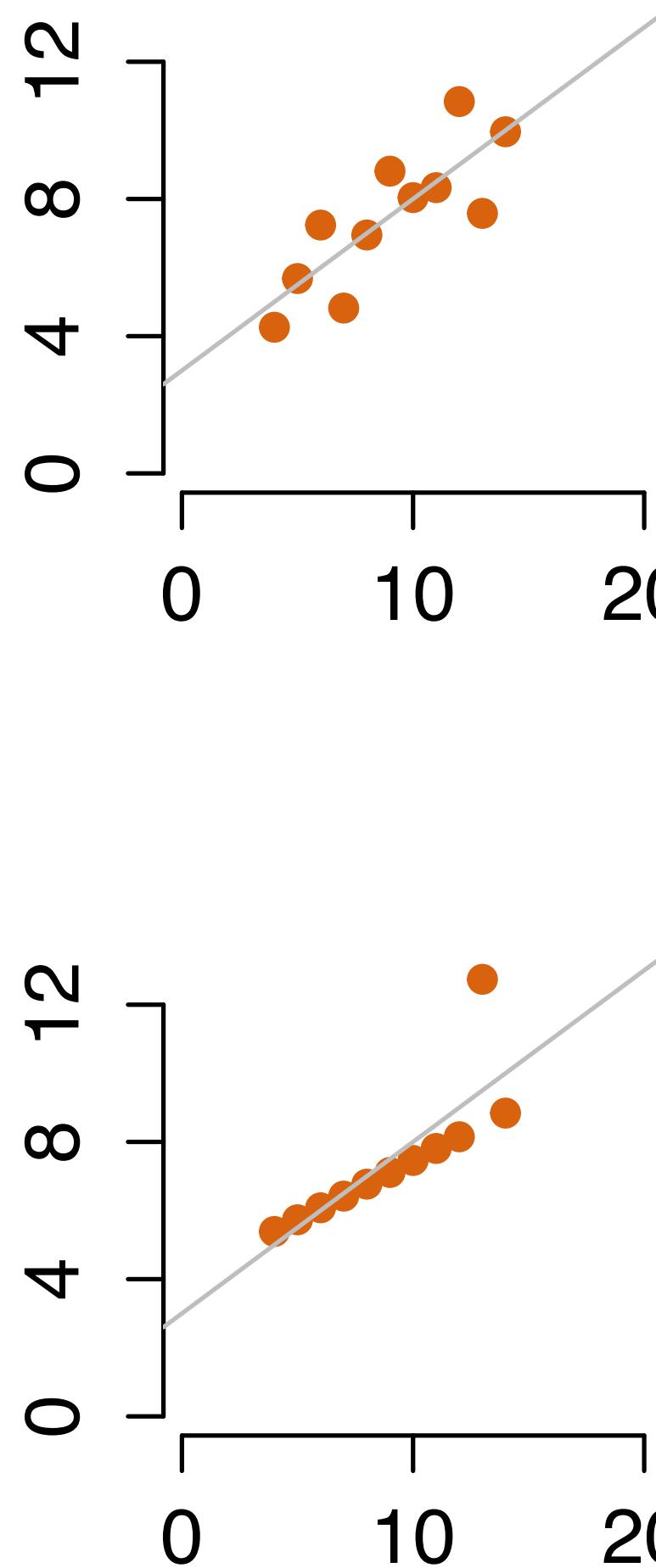
I		II		III		IV	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Anscombe quartet:

Property	Value
Mean of x	9
Sample variance of x	11
Mean of y	7.50
Sample variance of y	4.122 or 4.127
Correlation between x and y	0.816
Linear regression line	$y = 3.00 + 0.500x$

INTRODUCTION

GRAPHICS REVEALS DATA AND PATTERNS

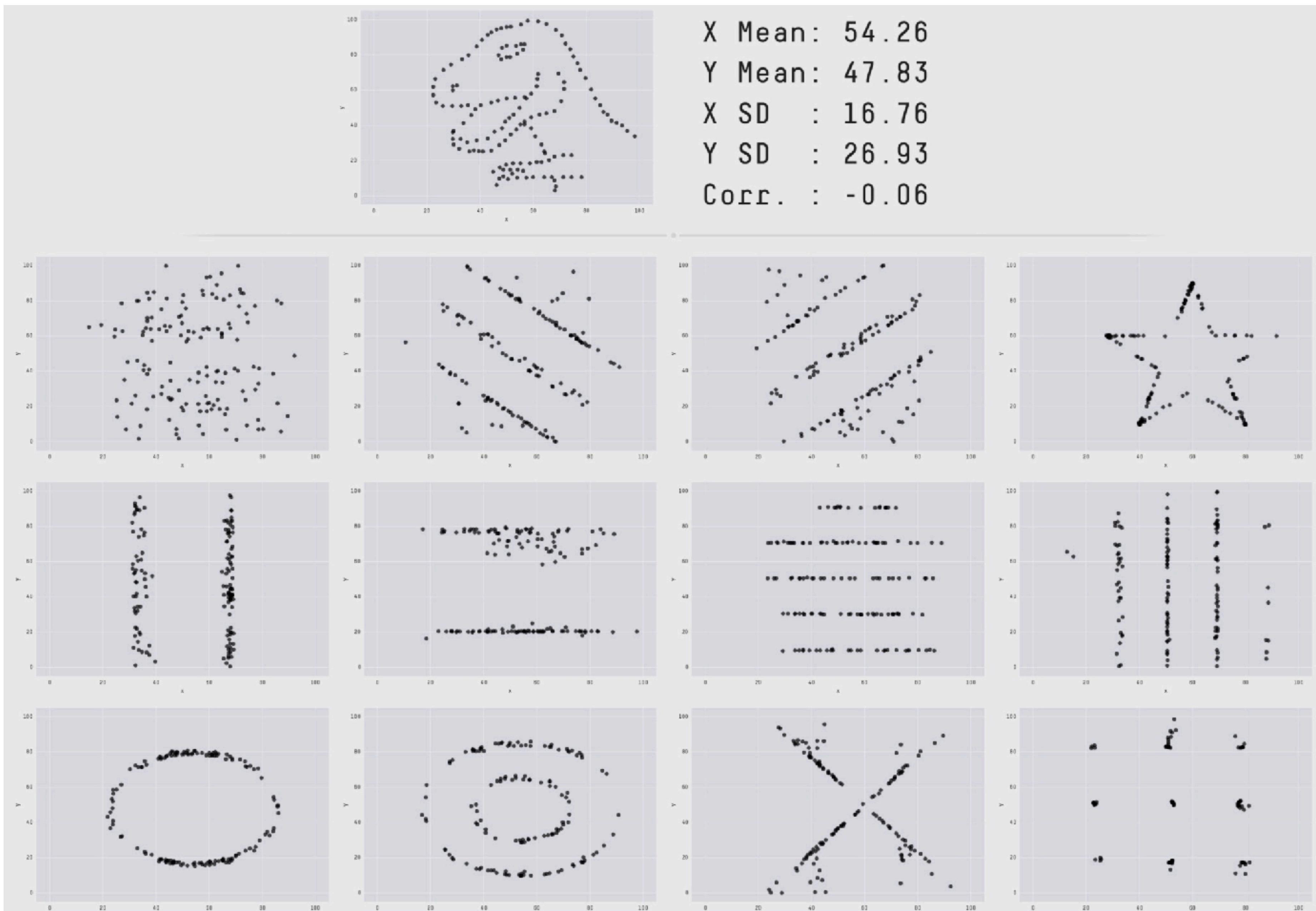


Anscombe quartet:

Property	Value
Mean of x	9
Sample variance of x	11
Mean of y	7.50
Sample variance of y	4.122 or 4.127
Correlation between x and y	0.816
Linear regression line	$y = 3.00 + 0.500x$

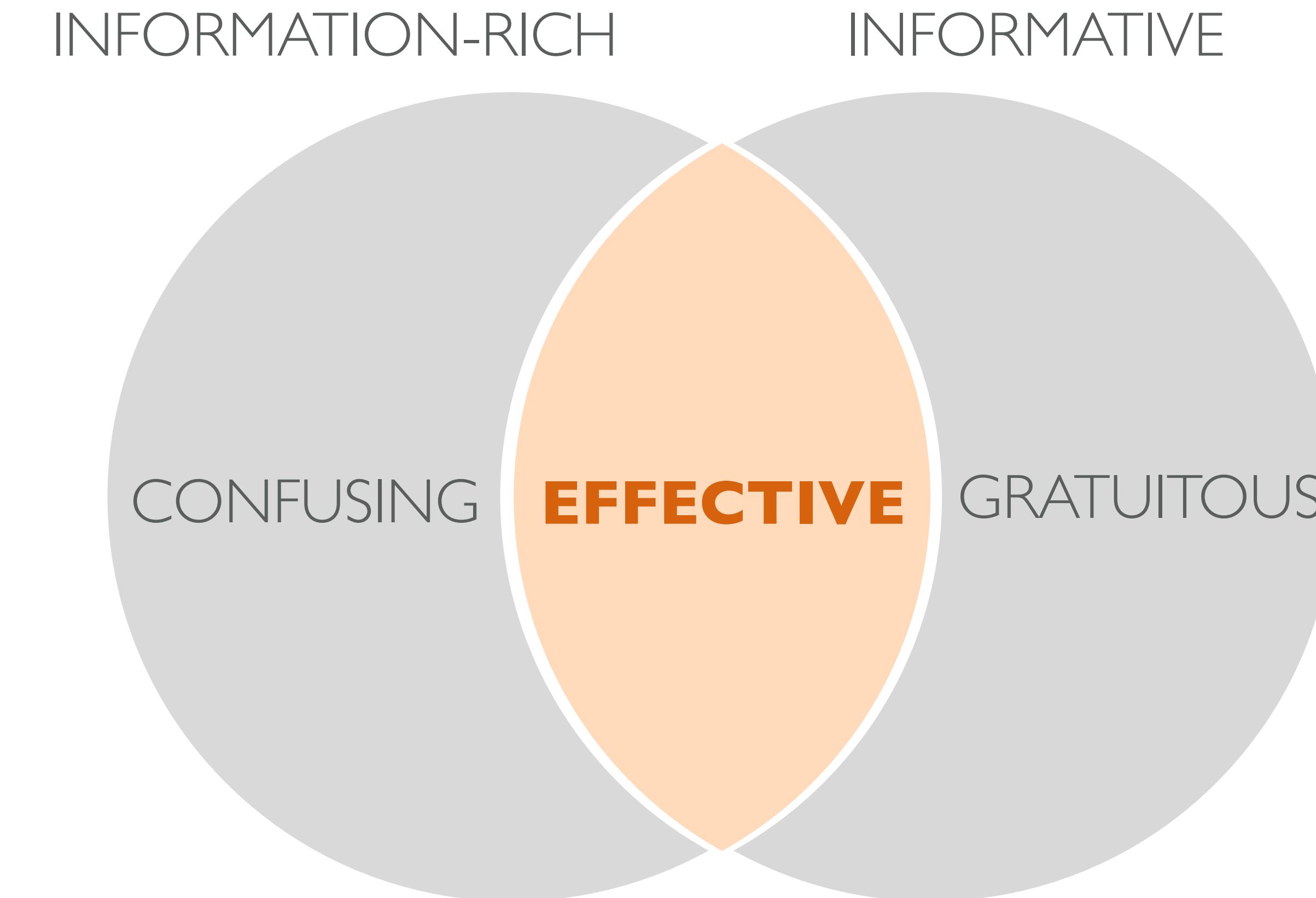
INTRODUCTION

DATASAURUS DOZEN



INTRODUCTION

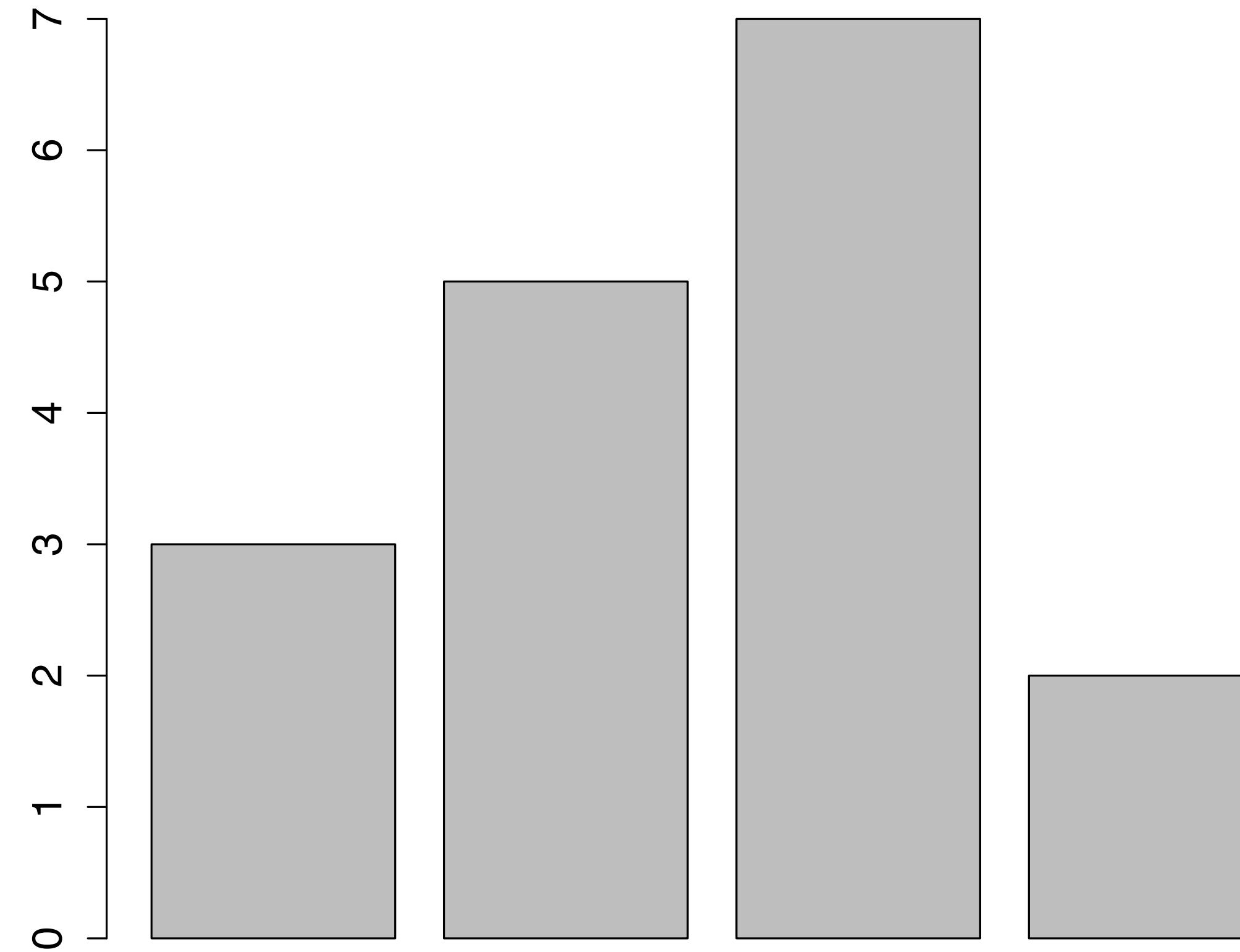
EFFECTIVE VISUALISATION



Krzywinsky, Behind every great visualization is a design principle

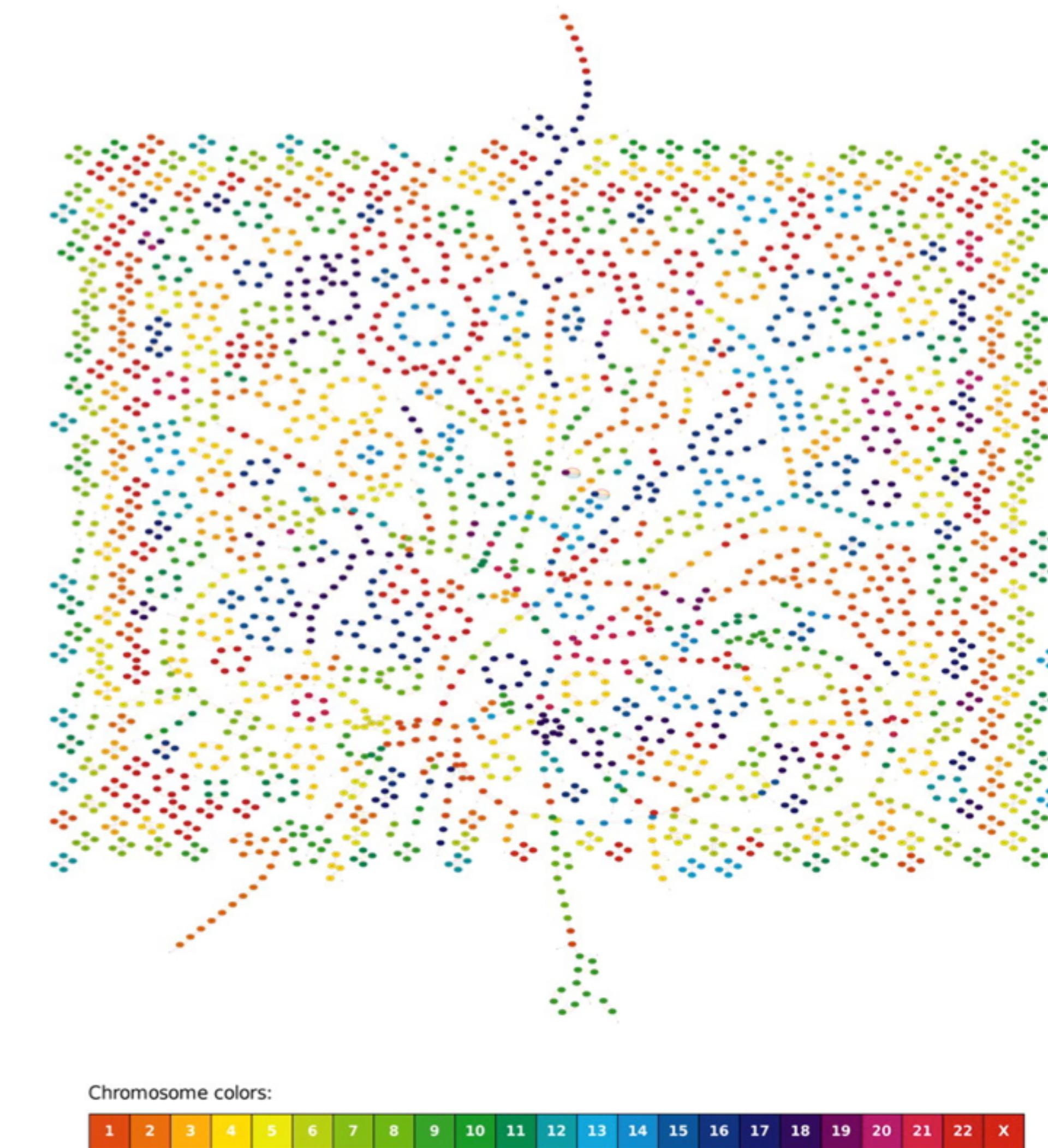
INTRODUCTION

INFORMATIVE, NOT INFORMATION-RICH



INTRODUCTION

INFORMATION-RICH, NOT INFORMATIVE



Alekseyev, M.A. and P.A. Pevzner, Breakpoint graphs and ancestral genome reconstructions. *Genome Res*, 2009. 19(5): p. 943-57.

INTRODUCTION

INFORMATION-RICH AND INFORMATIVE

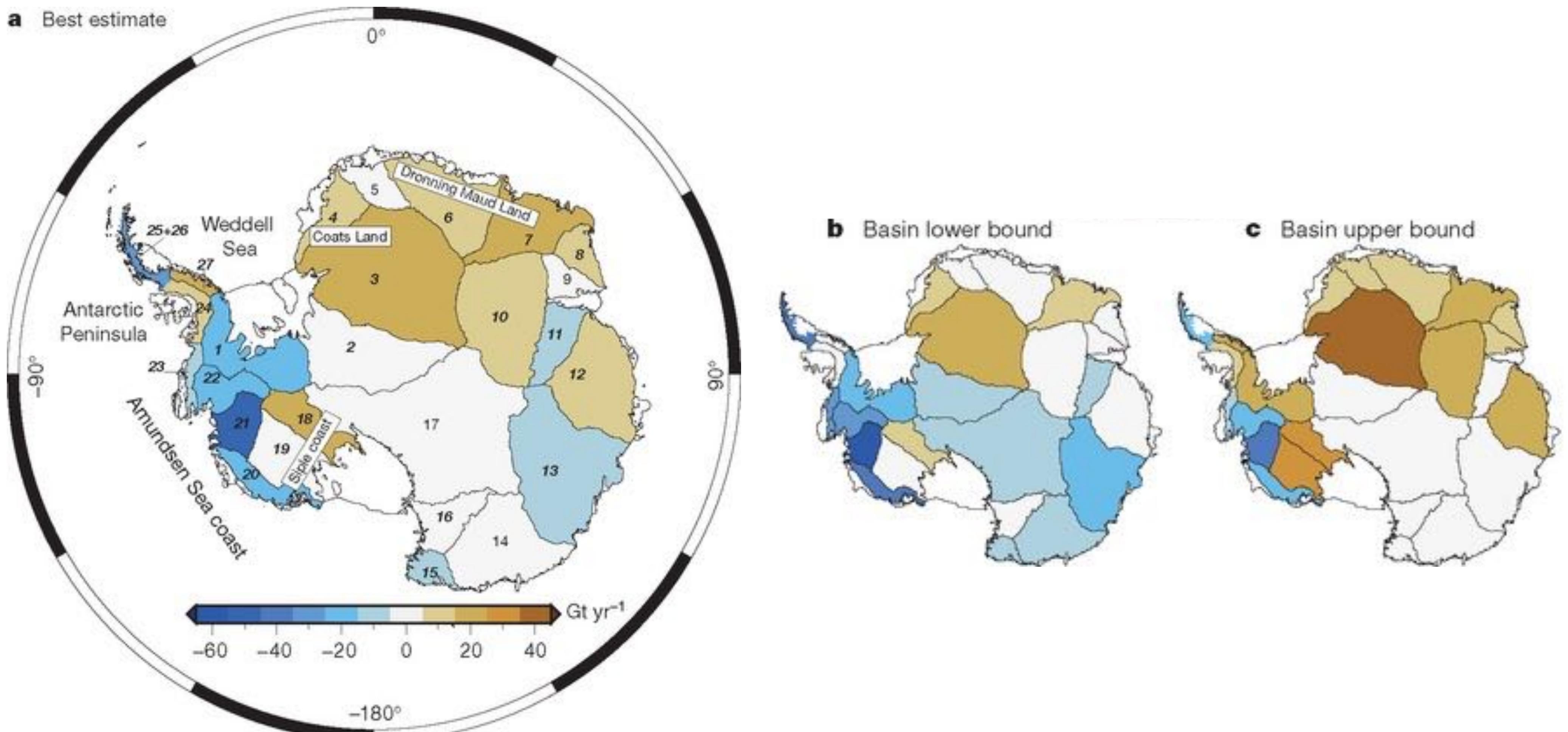
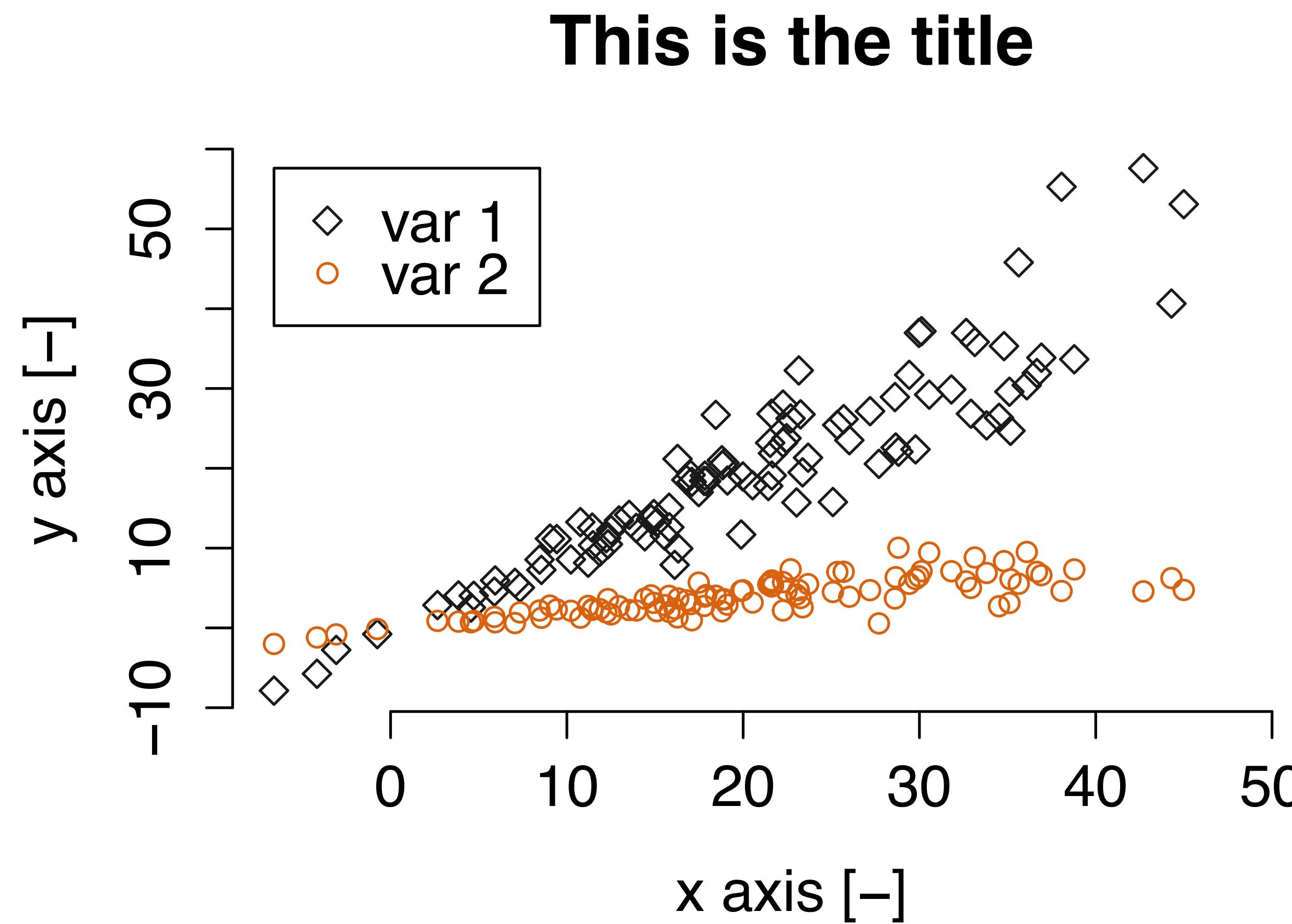


FIGURE BASICS

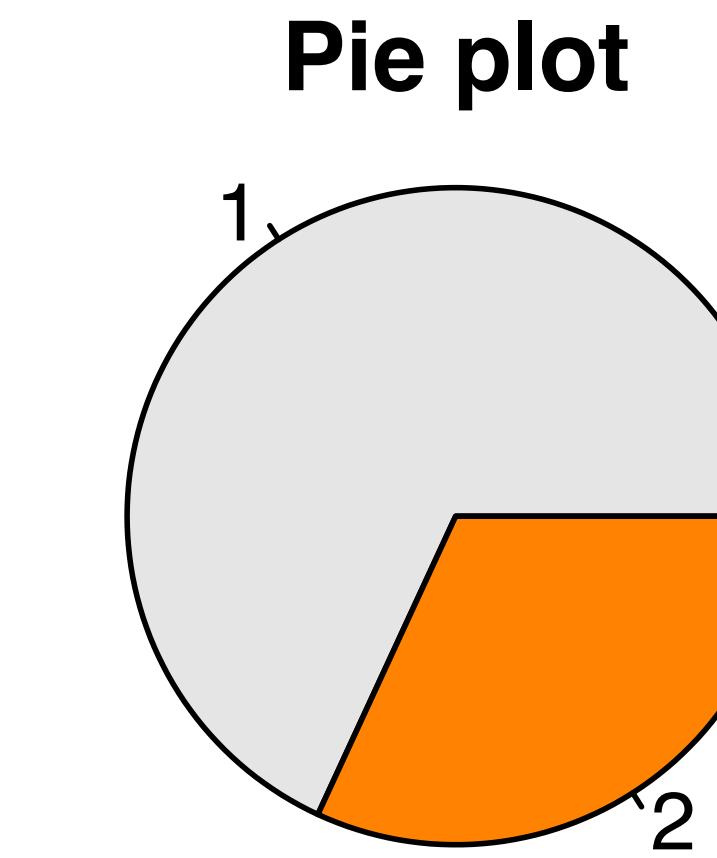
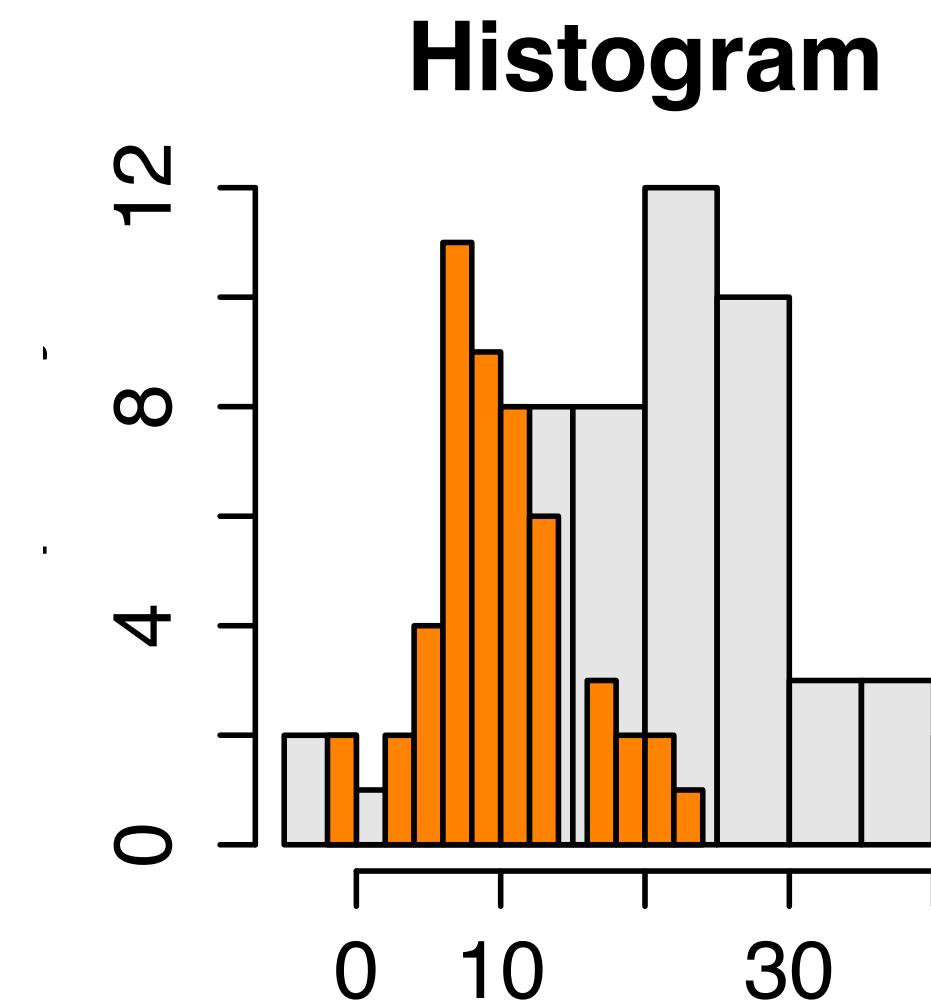
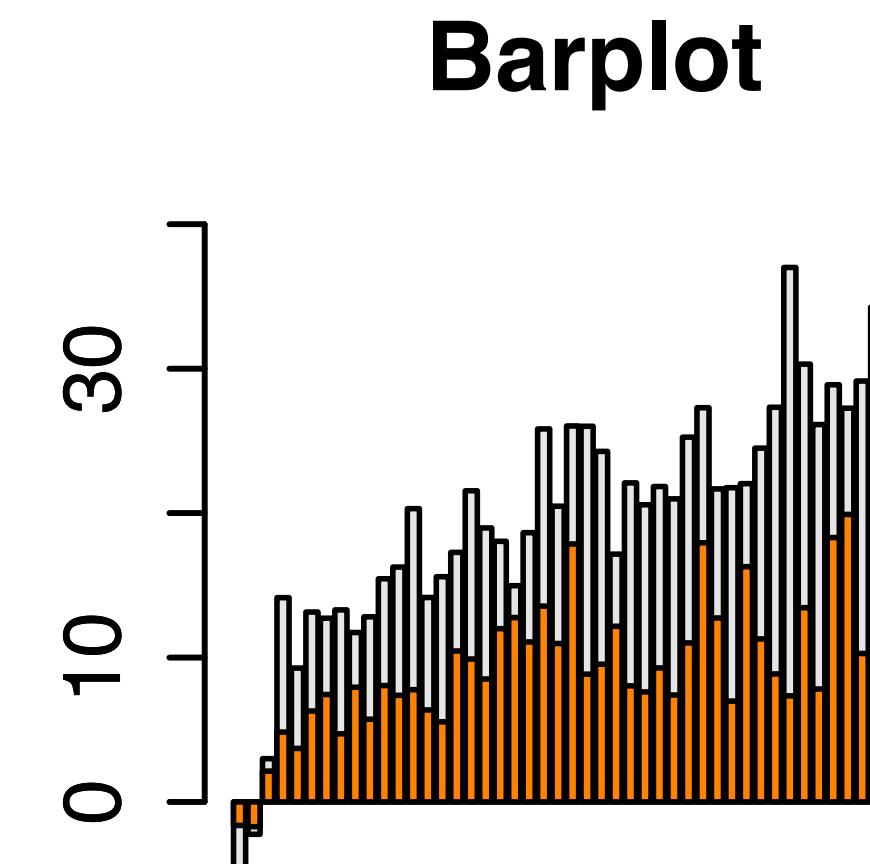
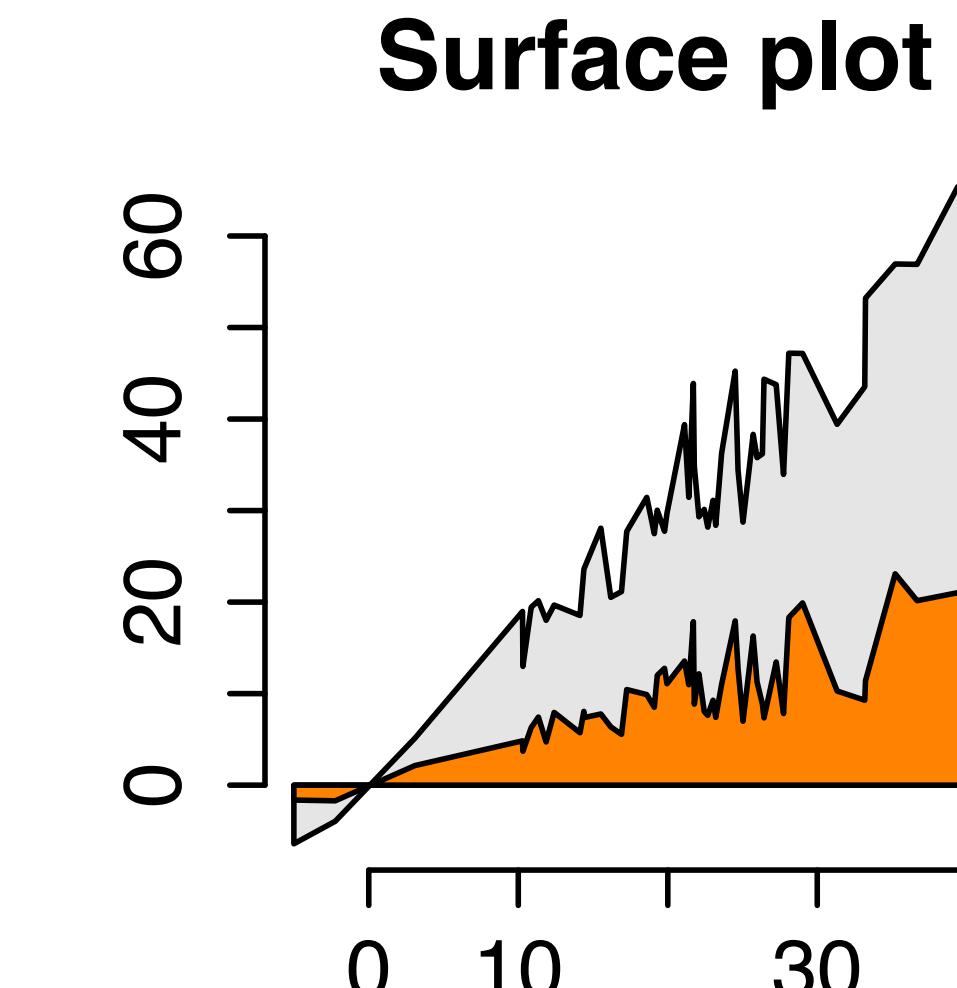
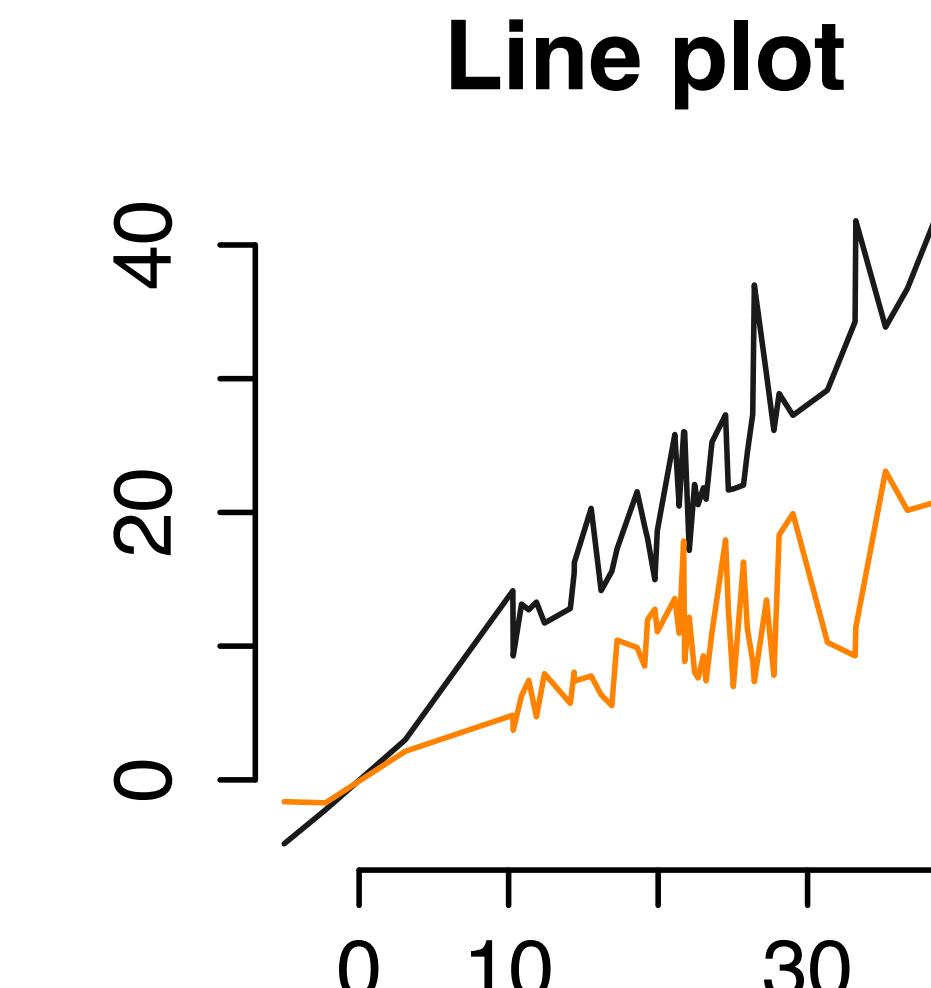
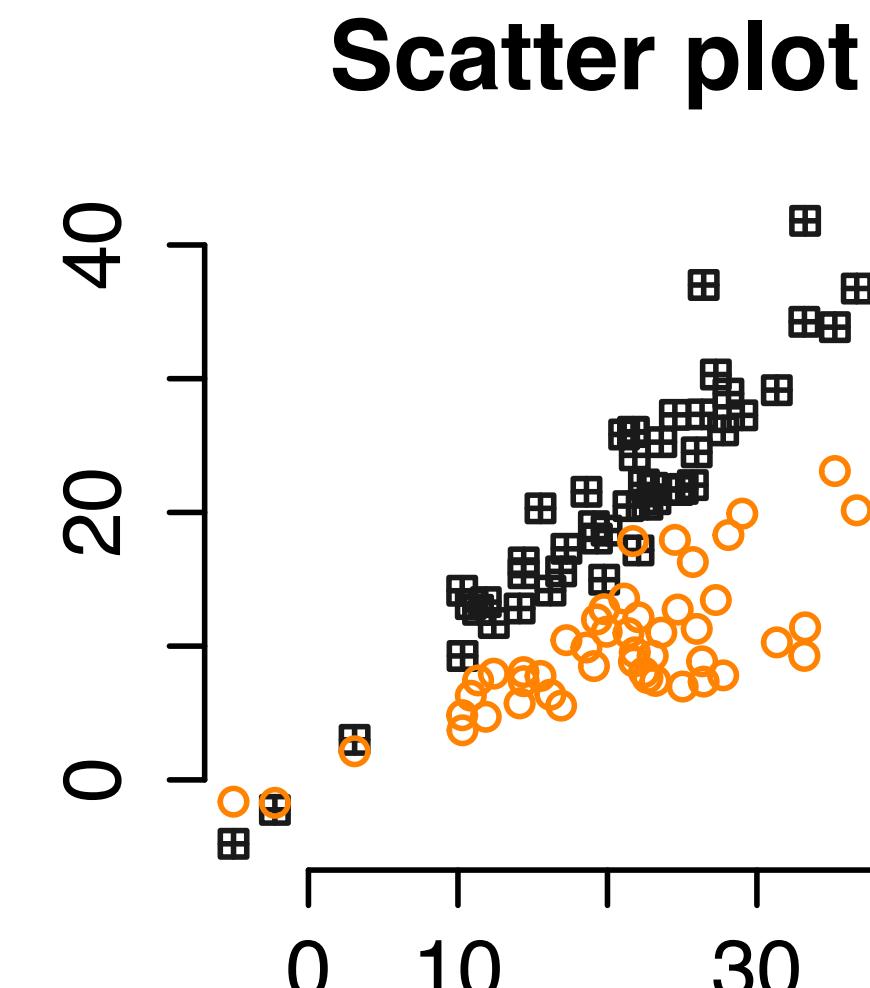
FIGURE ELEMENTS

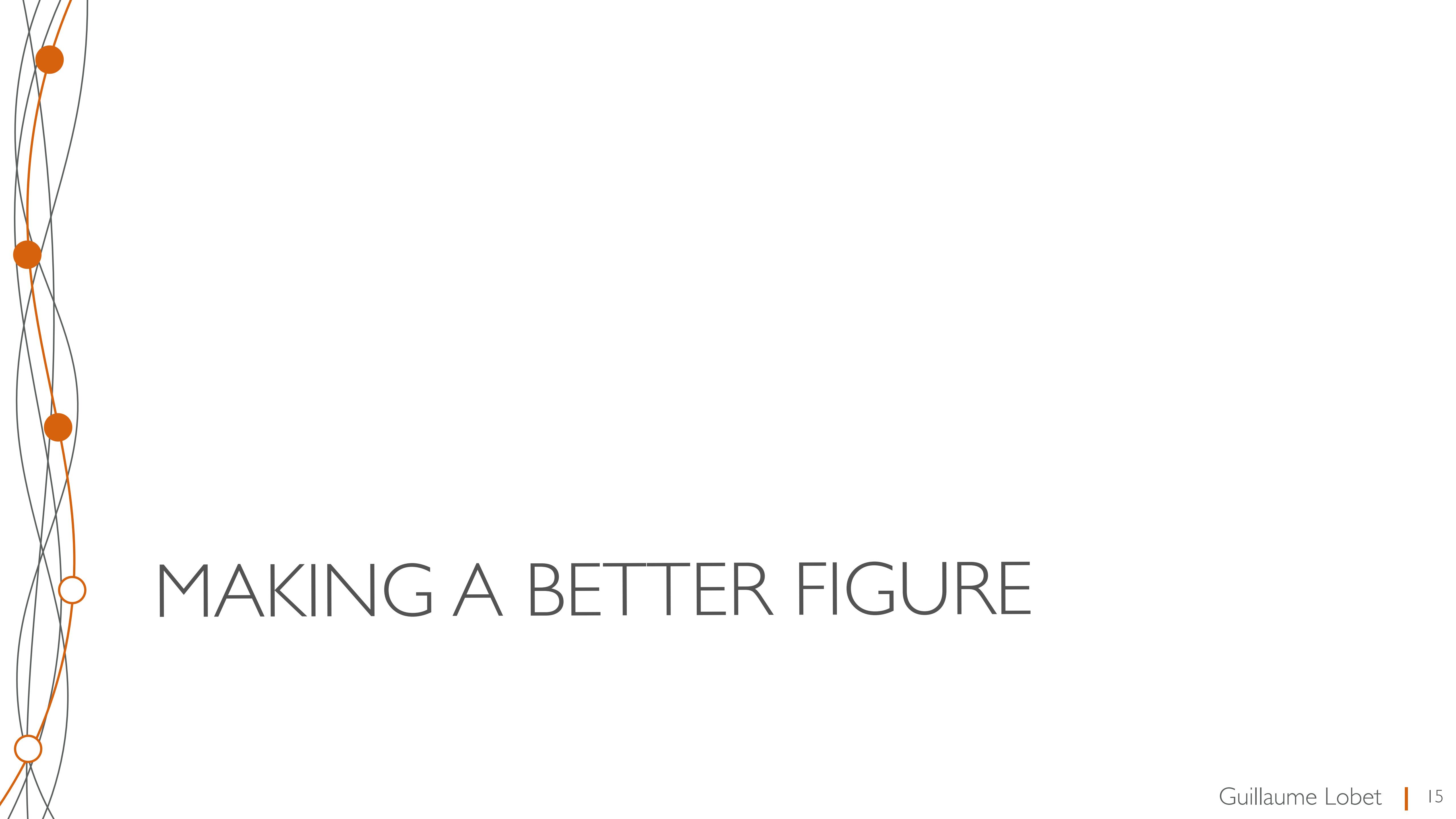


- data
- coordinate system
- scale
- context (legend)

FIGURE BASICS

USE THE RIGHT REPRESENTATION FOR YOUR DATA





MAKING A BETTER FIGURE

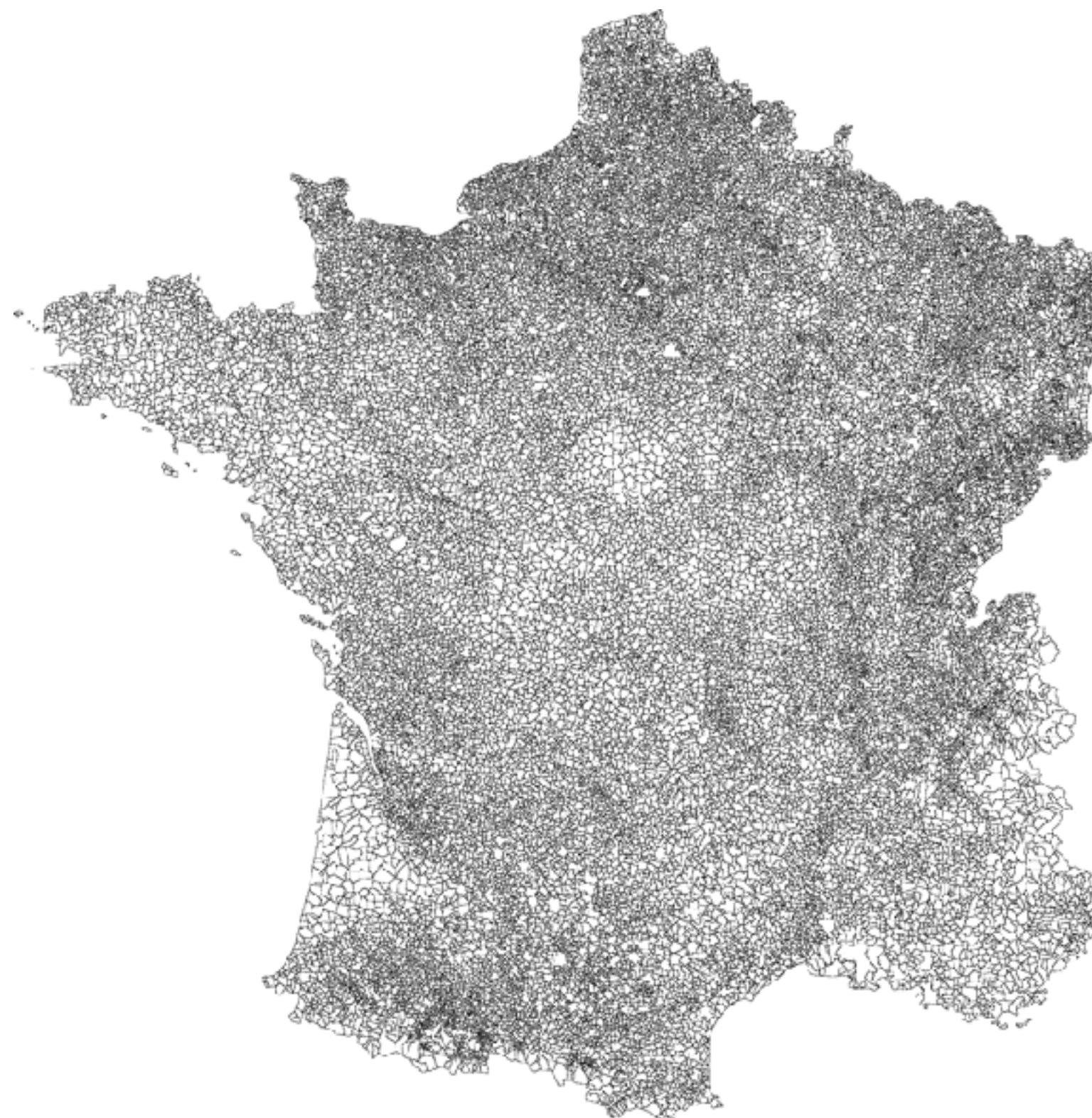
DATA / INK RATIO

- proportion of graphic's ink devoted to the **non-redundant display of data information**
- $1.0 -$ proportion of a graphic that can be **erased without loss of data information**
- data-to-ink ratio should always be **maximized**, within reason

MAKING A BETTER FIGURE

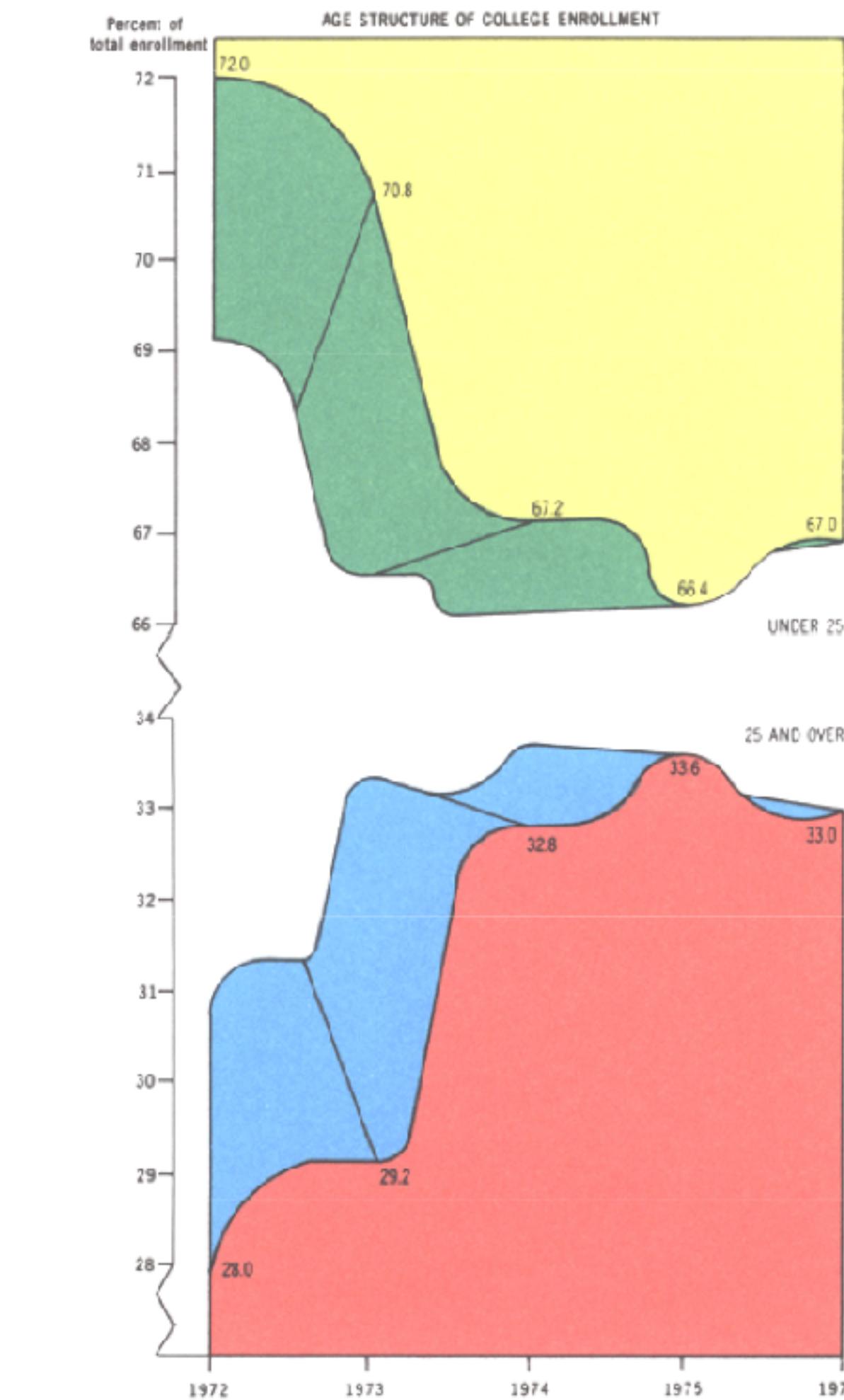
DATA / INK RATIO

High



240 000 numbers

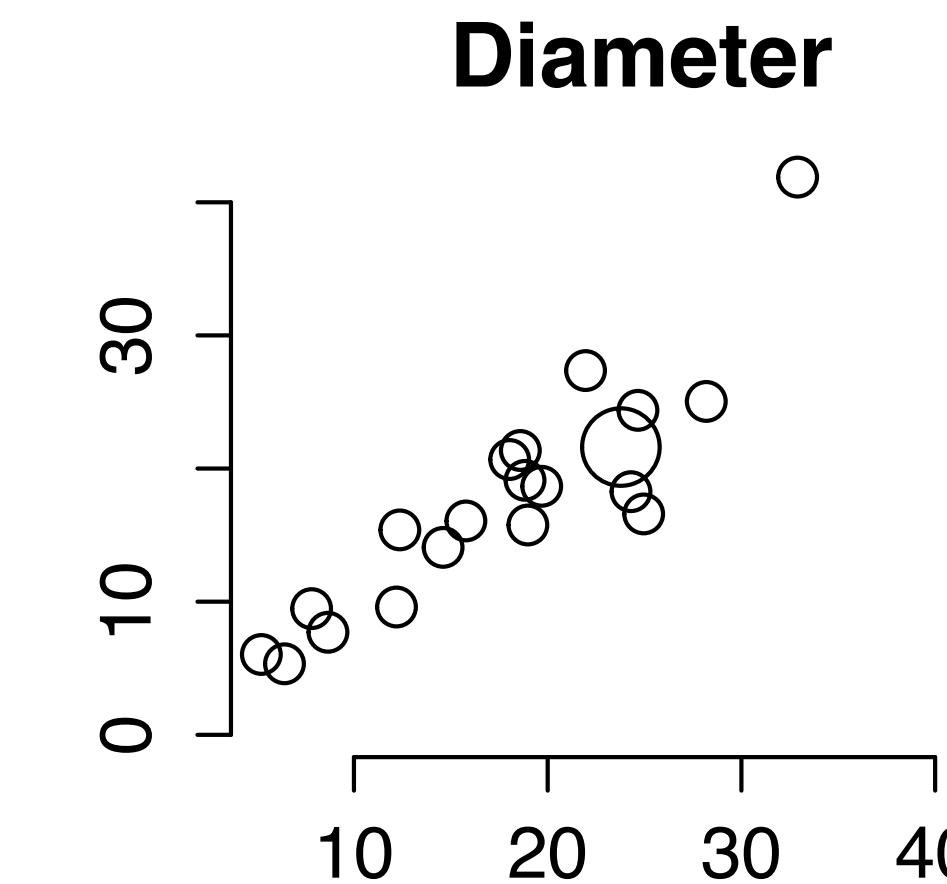
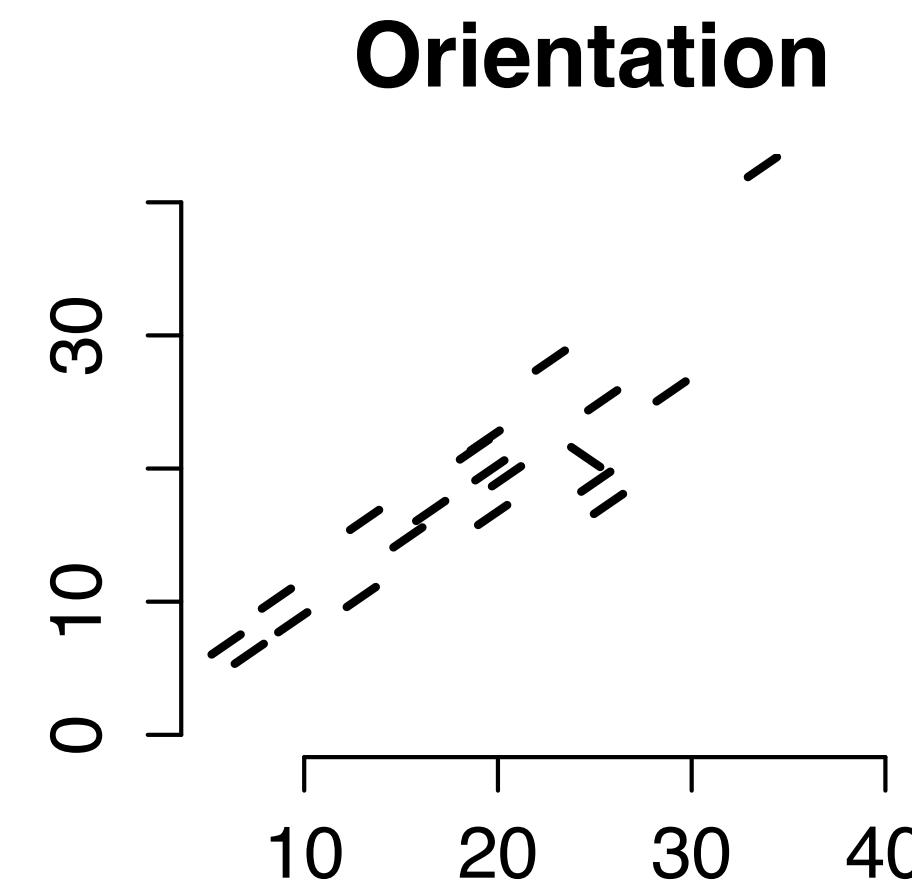
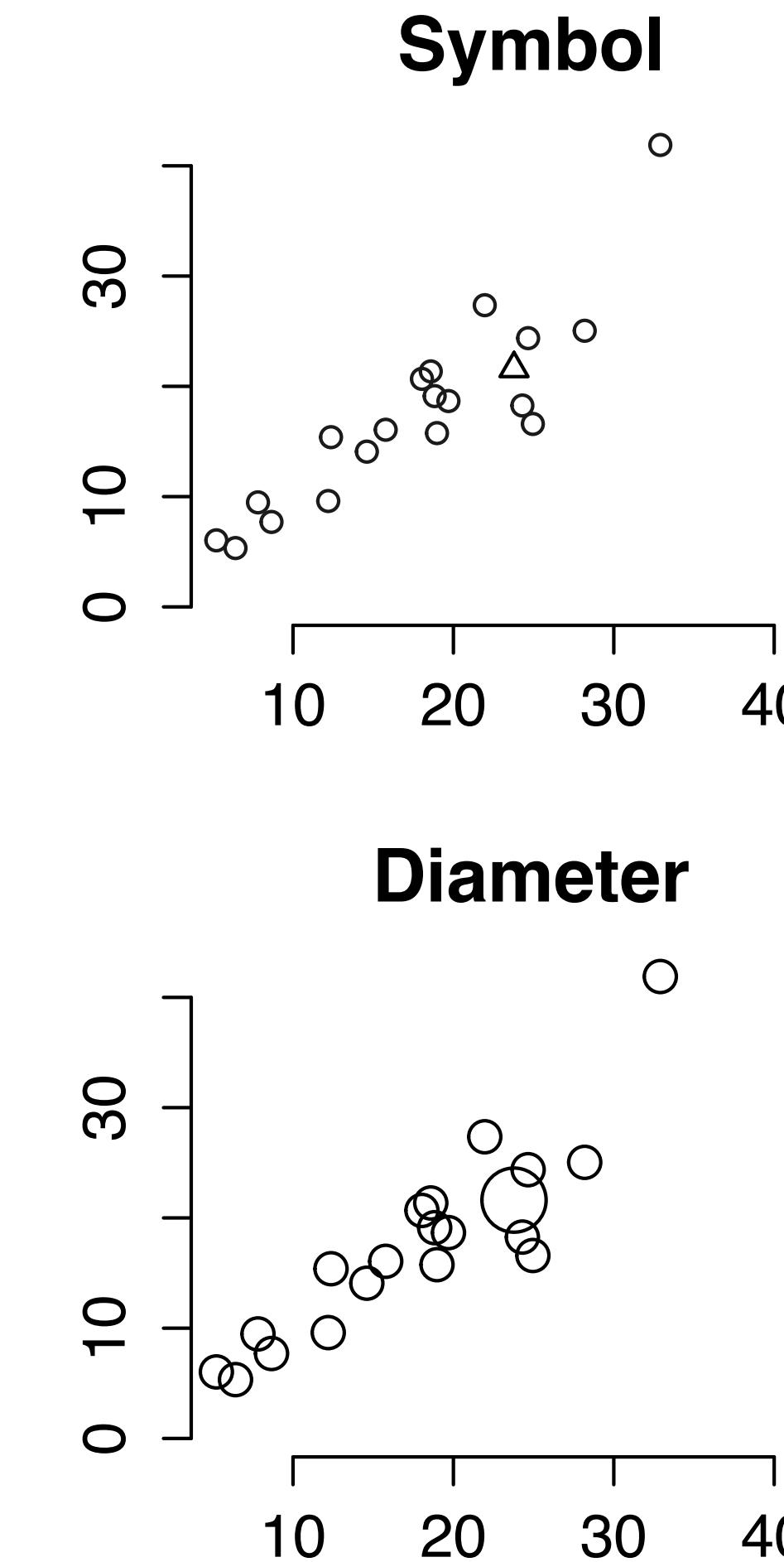
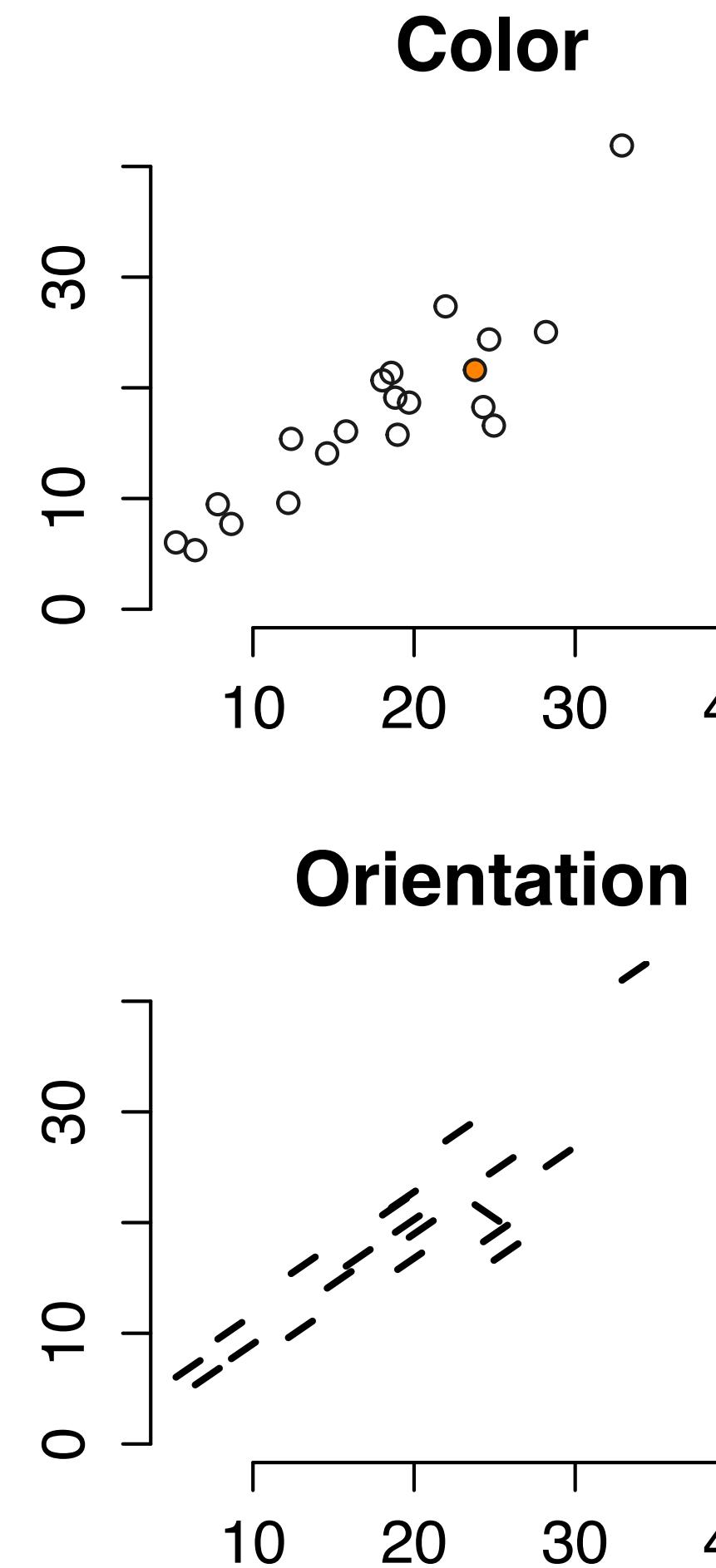
Shockingly low



5 numbers

MAKING A BETTER FIGURE

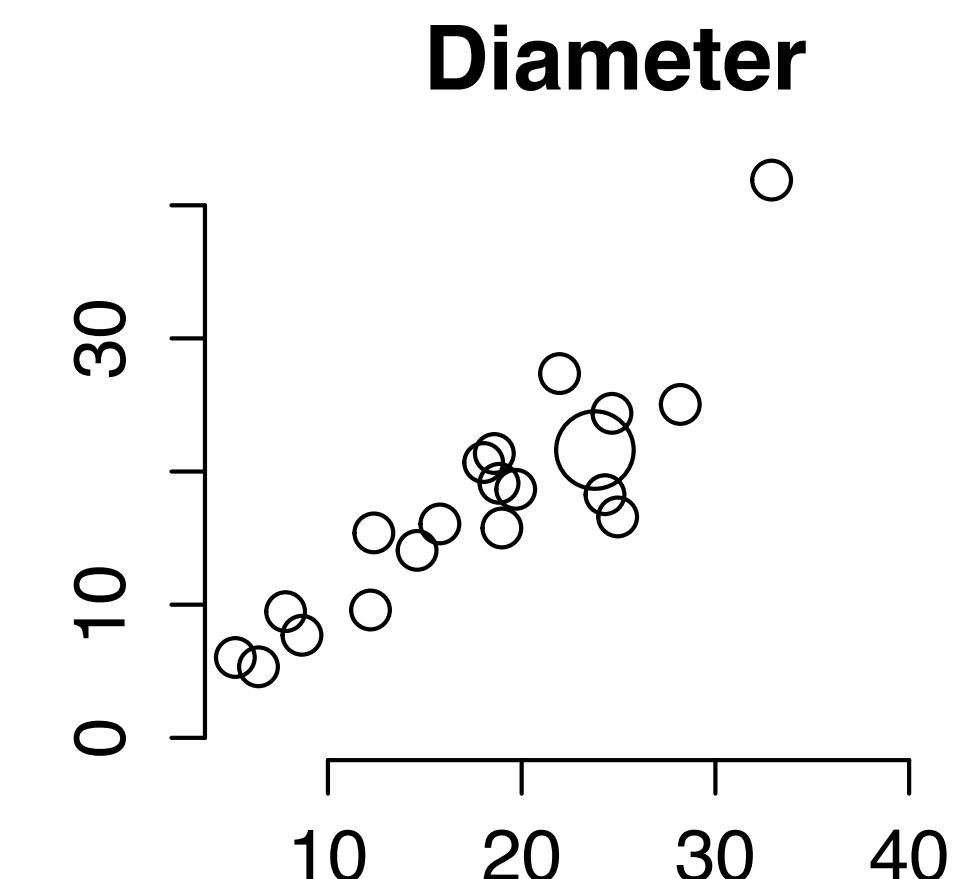
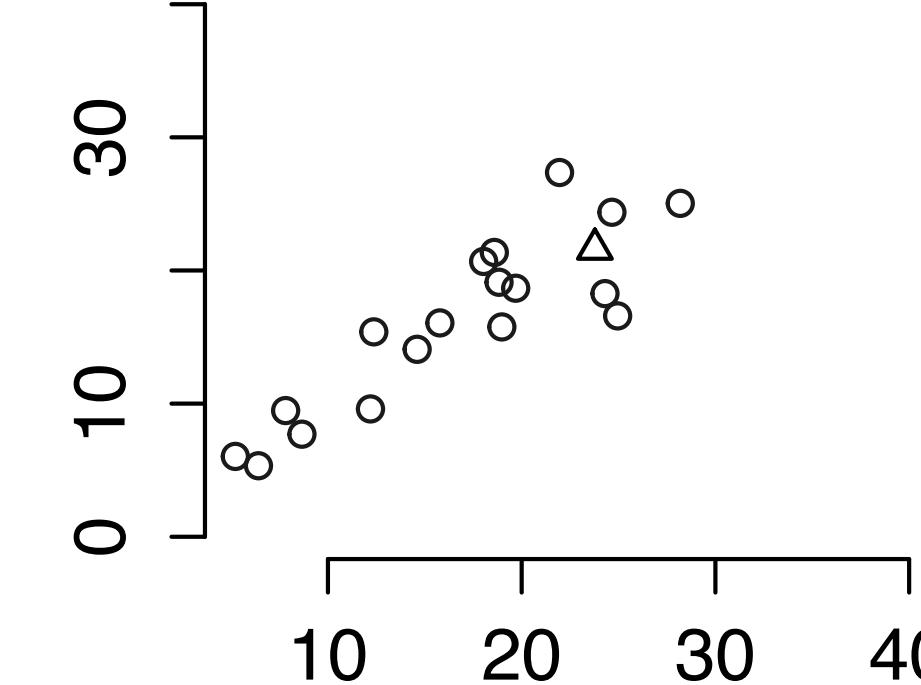
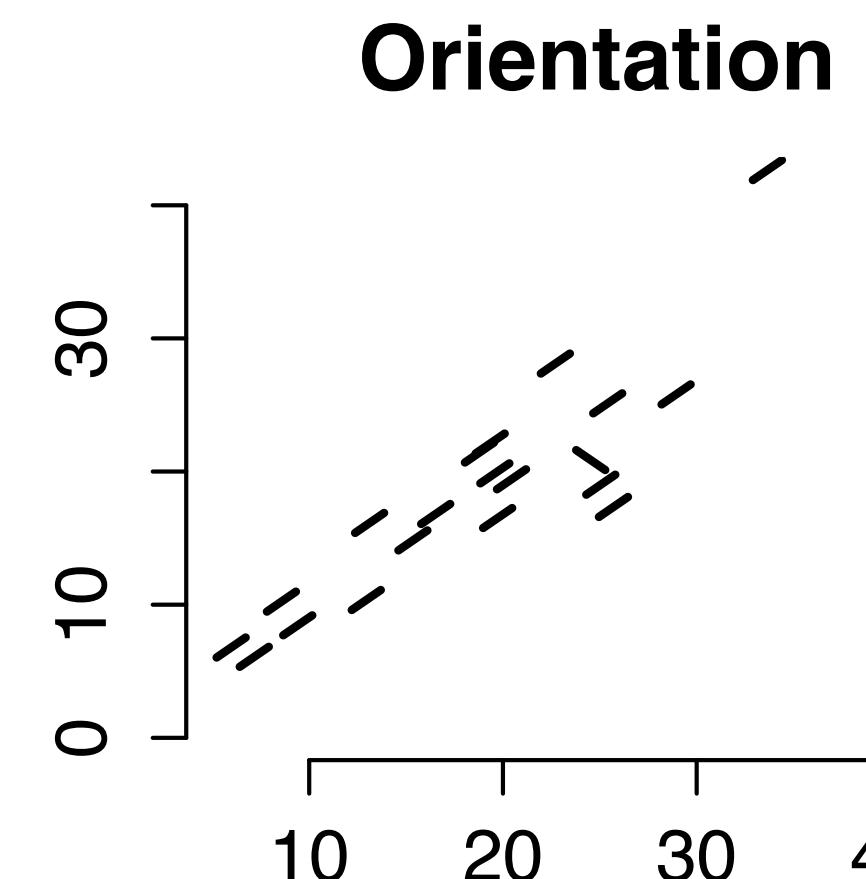
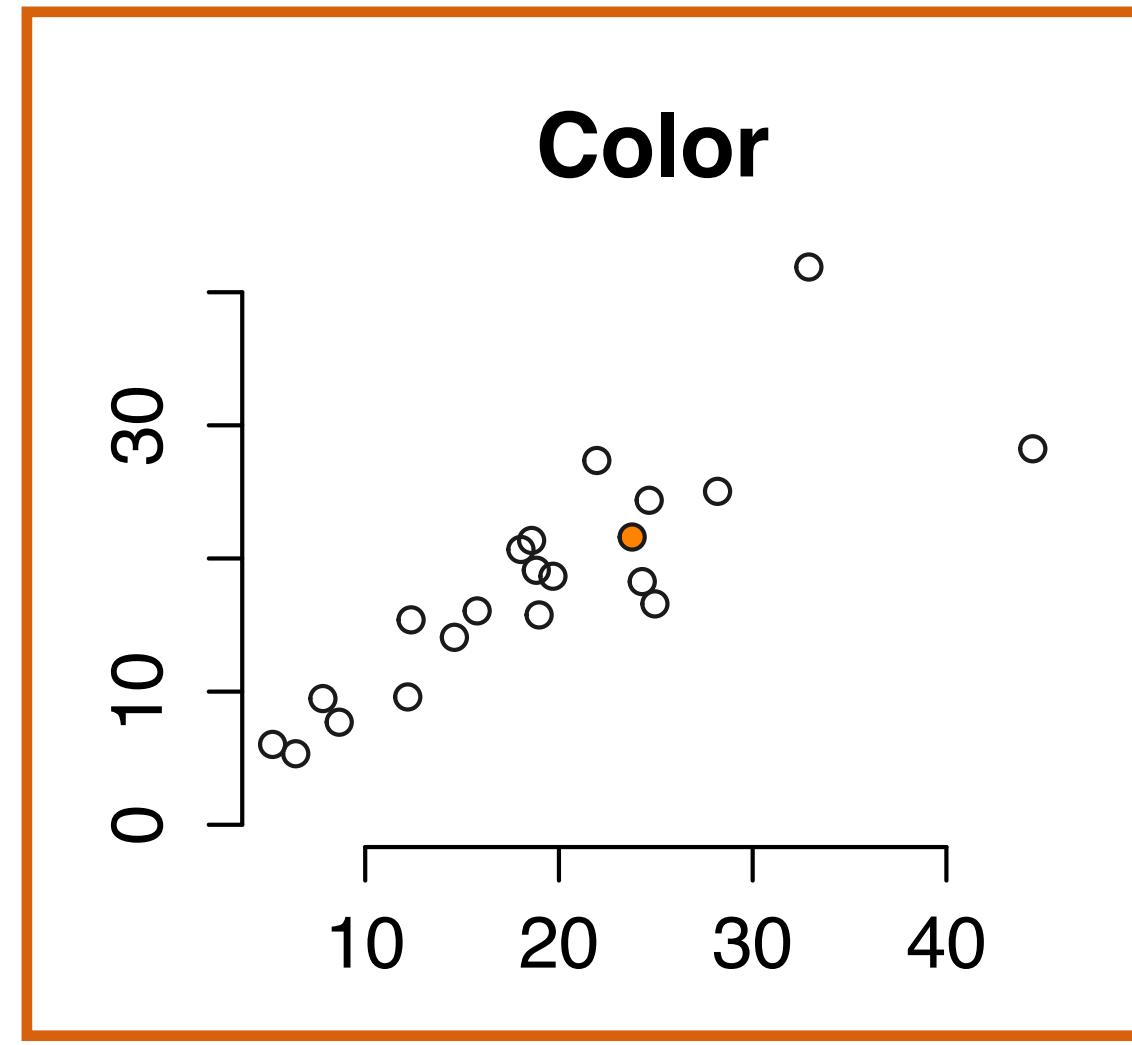
SALIENCE



- Set an object **apart** from its surrounding
- **Highlighting** the message

MAKING A BETTER FIGURE

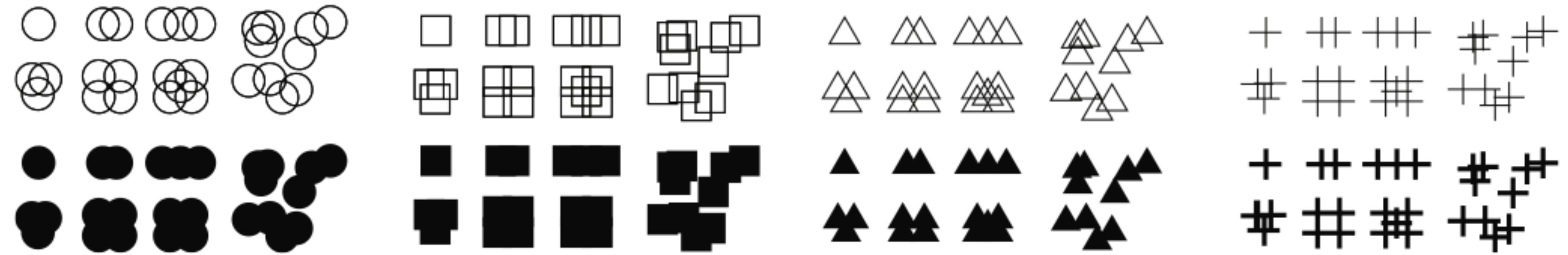
SALIENCE



- Set an object **apart** from its surrounding
- **Highlighting** the message

MAKING A BETTER FIGURE

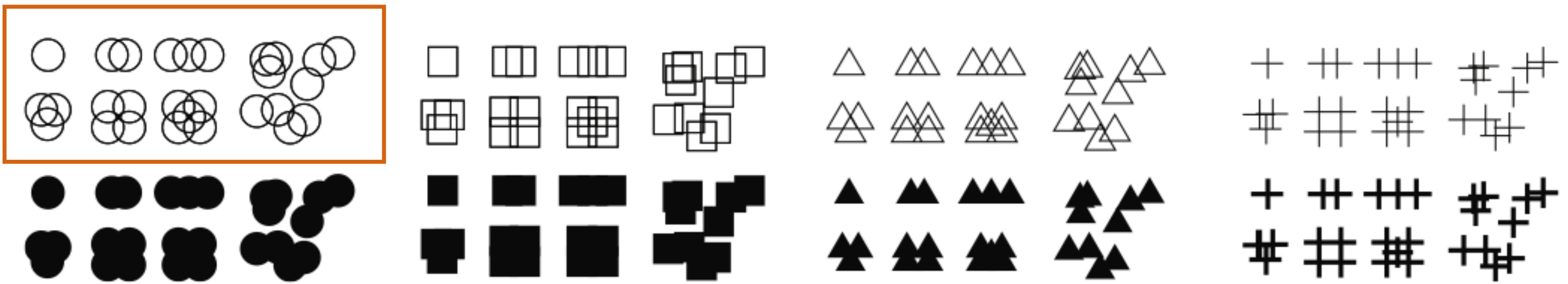
SYMBOLS



- Parsability is not the same for all symbols
- Prefer the empty circles if possible

MAKING A BETTER FIGURE

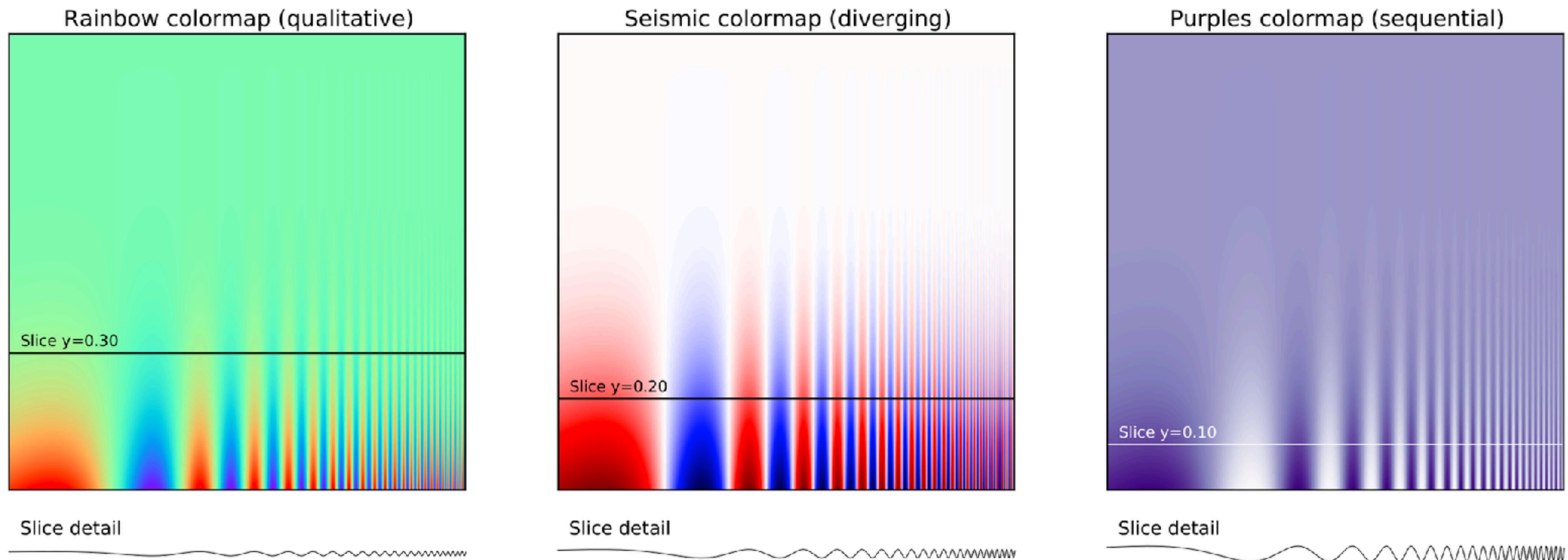
SYMBOLS



- Parsability is not the same for all symbols
- Prefer the empty circles if possible

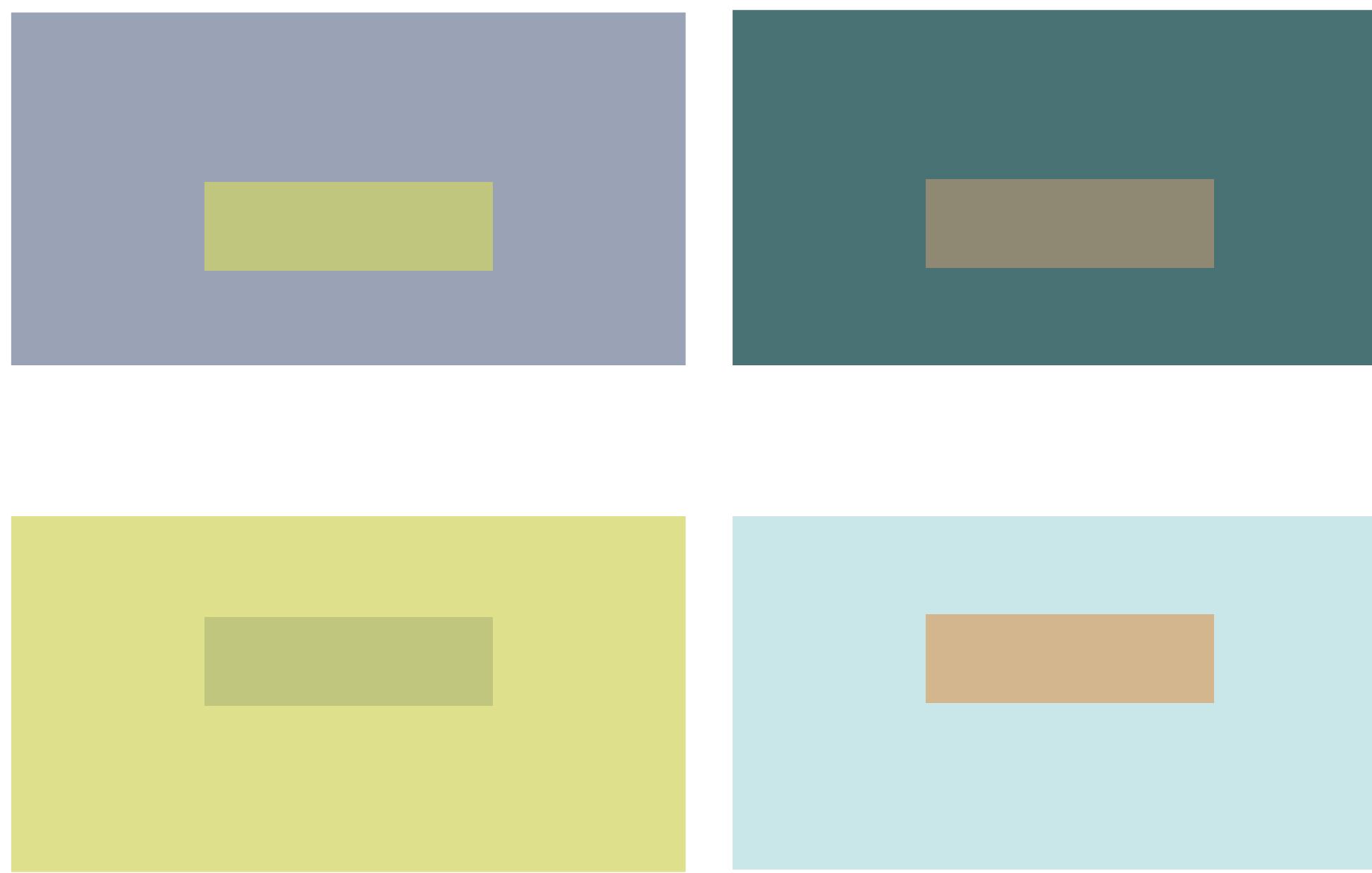
MAKING A BETTER FIGURE

USE COLORS EFFICIENTLY



When possible: avoid using colours

USING COLOURS

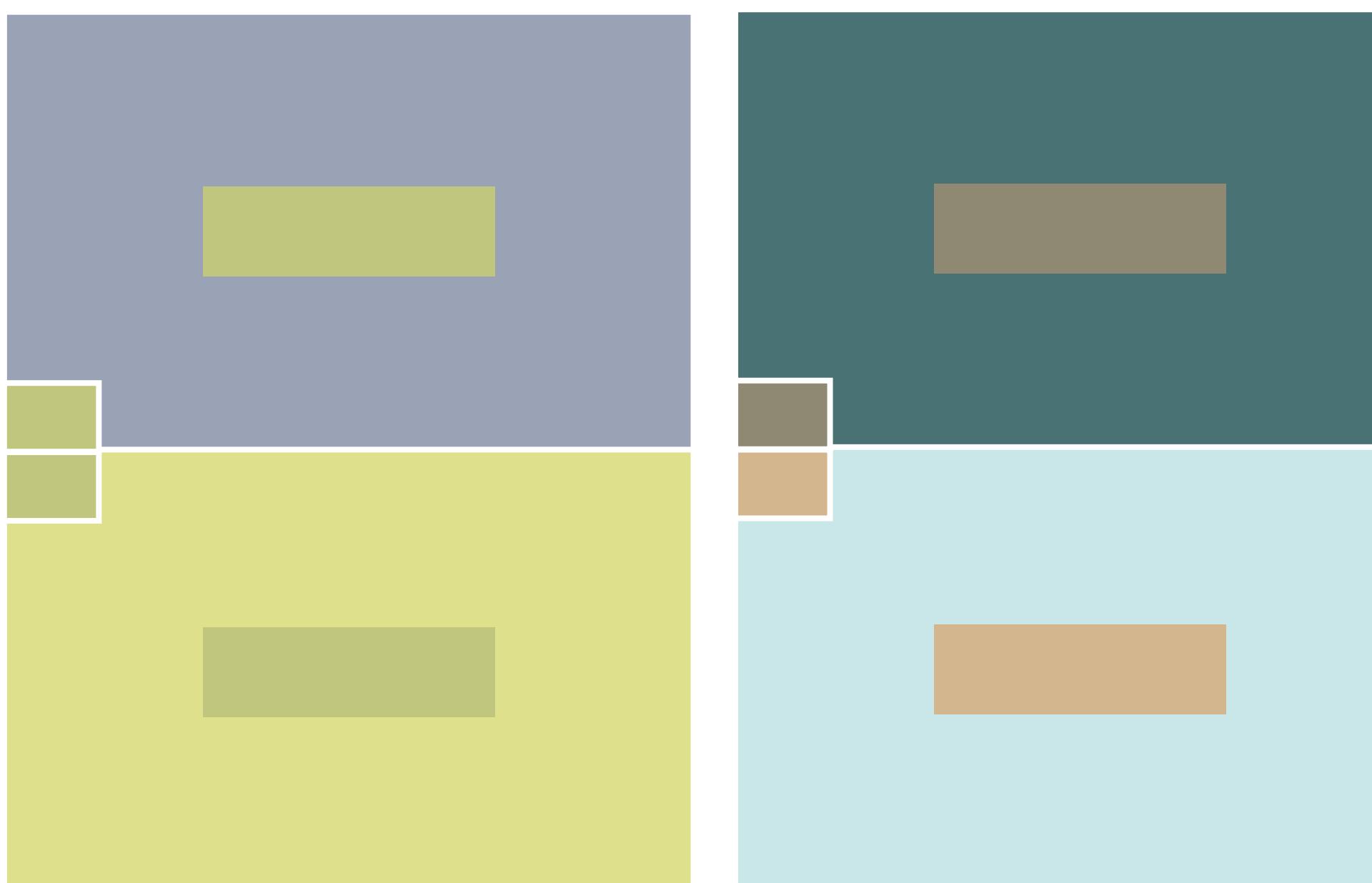


- low **parsability** by human eyes
- **relative** to its environment
- issue when **printing** in greyscale
- more expensive in printed publication
- color **blindness**

When possible: avoid using colours

MAKING A BETTER FIGURE

USING COLOURS



- low **parsability** by human eyes
- **relative** to its environment
- issue when **printing** in greyscale
- more expensive in printed publication
- color **blindness**

When possible: avoid using colours

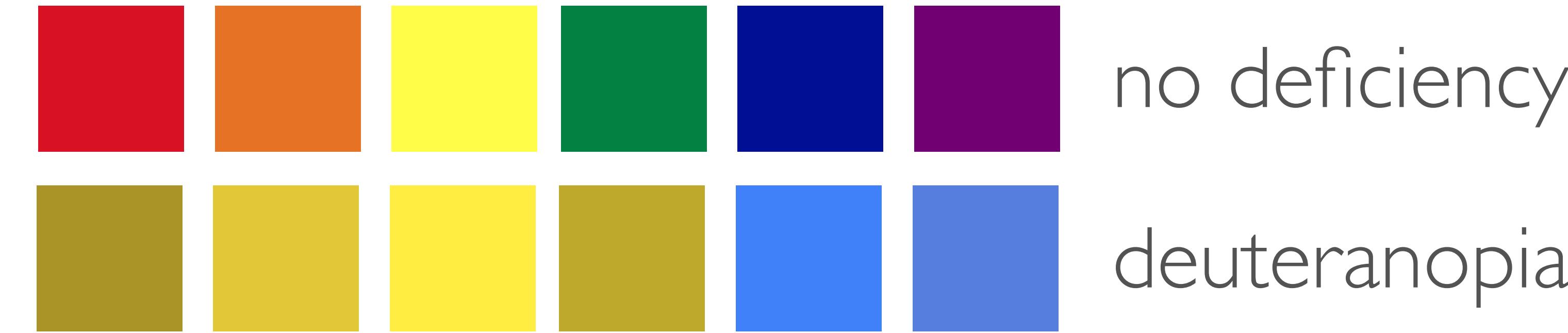
COLOR DEFICIENCY



no deficiency

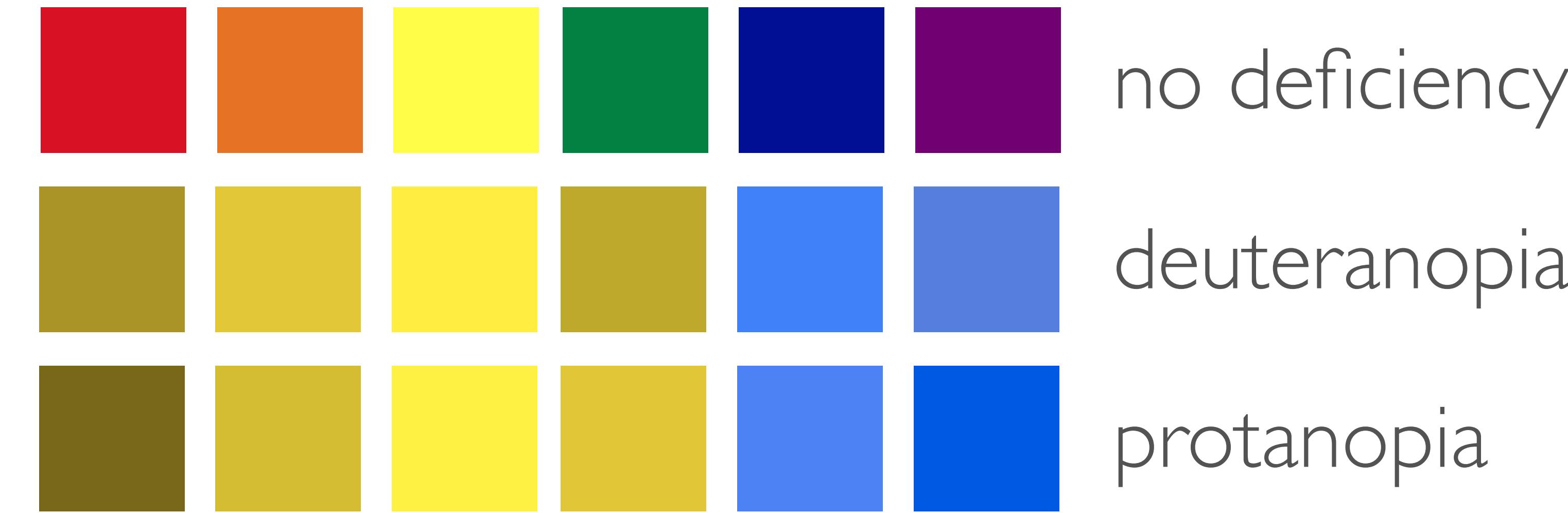
8 percent of men and **0.5** percent of women experience the common form of red-green color blindness

COLOR DEFICIENCY



8 percent of men and **0.5** percent of women experience the common form of red-green color blindness

COLOR DEFICIENCY



8 percent of men and **0.5** percent of women experience the common form of red-green color blindness

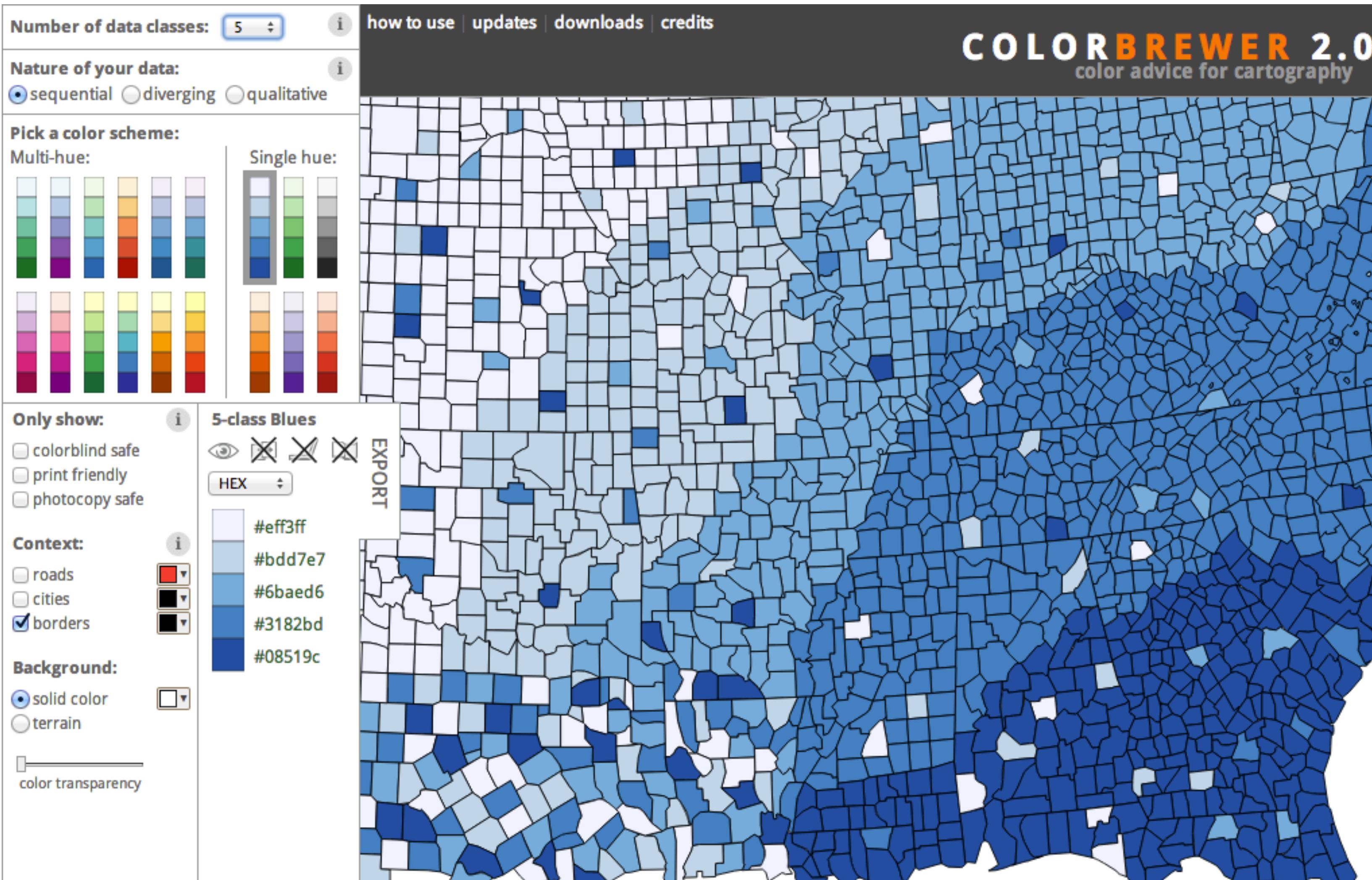
COLOR DEFICIENCY



8 percent of men and **0.5** percent of women experience the common form of red-green color blindness

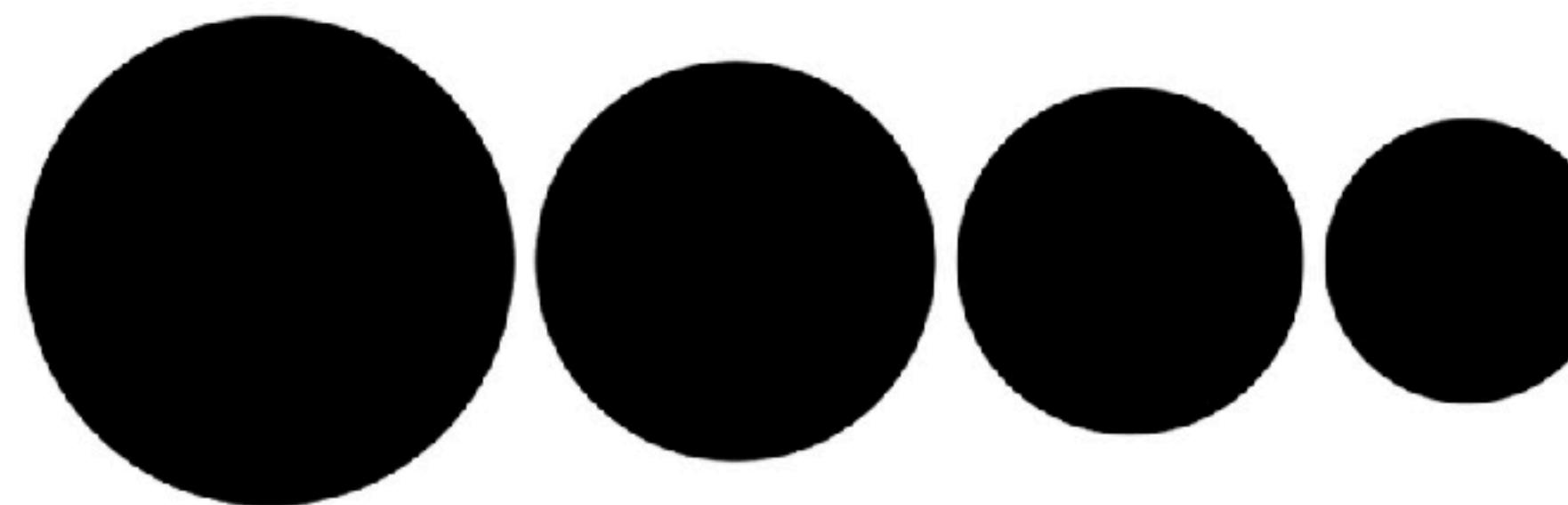
MAKING A BETTER FIGURE

COLOR BREWER 2



MAKING A BETTER FIGURE

DO NOT MISLEAD THE READER



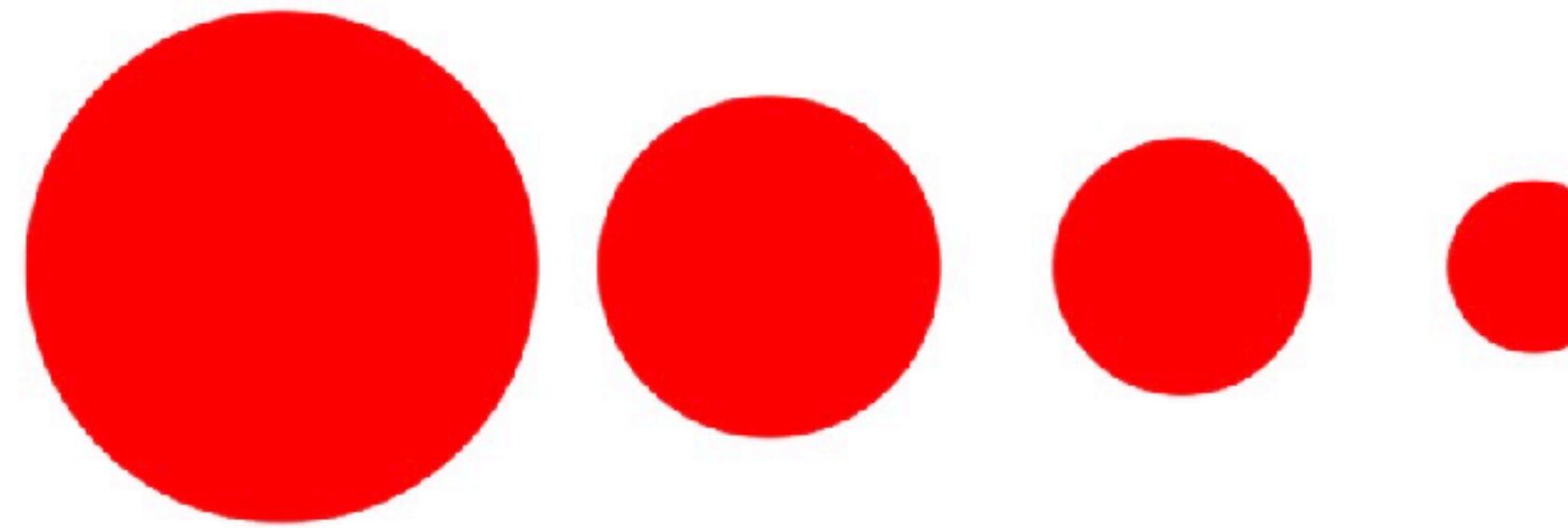
Relative size using disc area

Relative size using disc radius

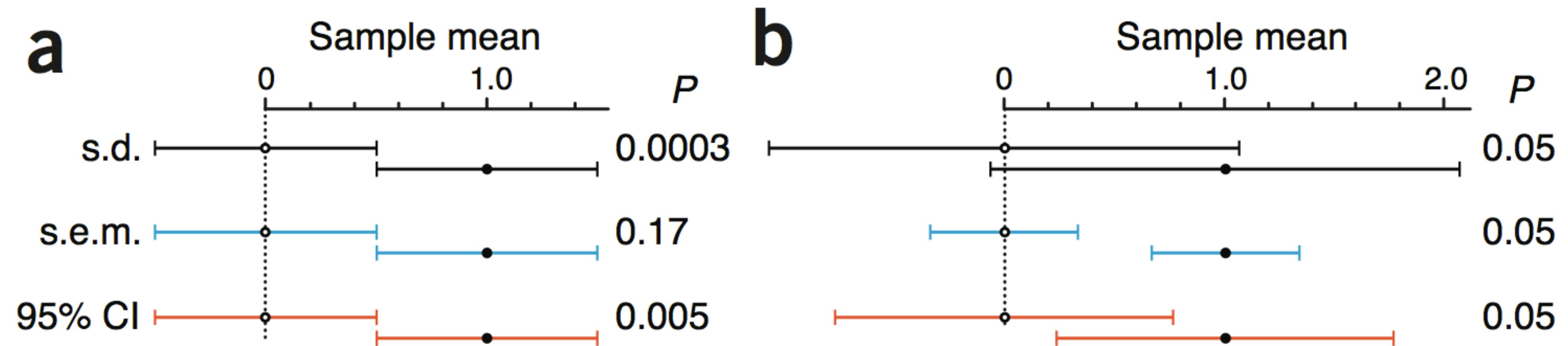


Relative size using full range

Relative size using partial range



ERROR BARS

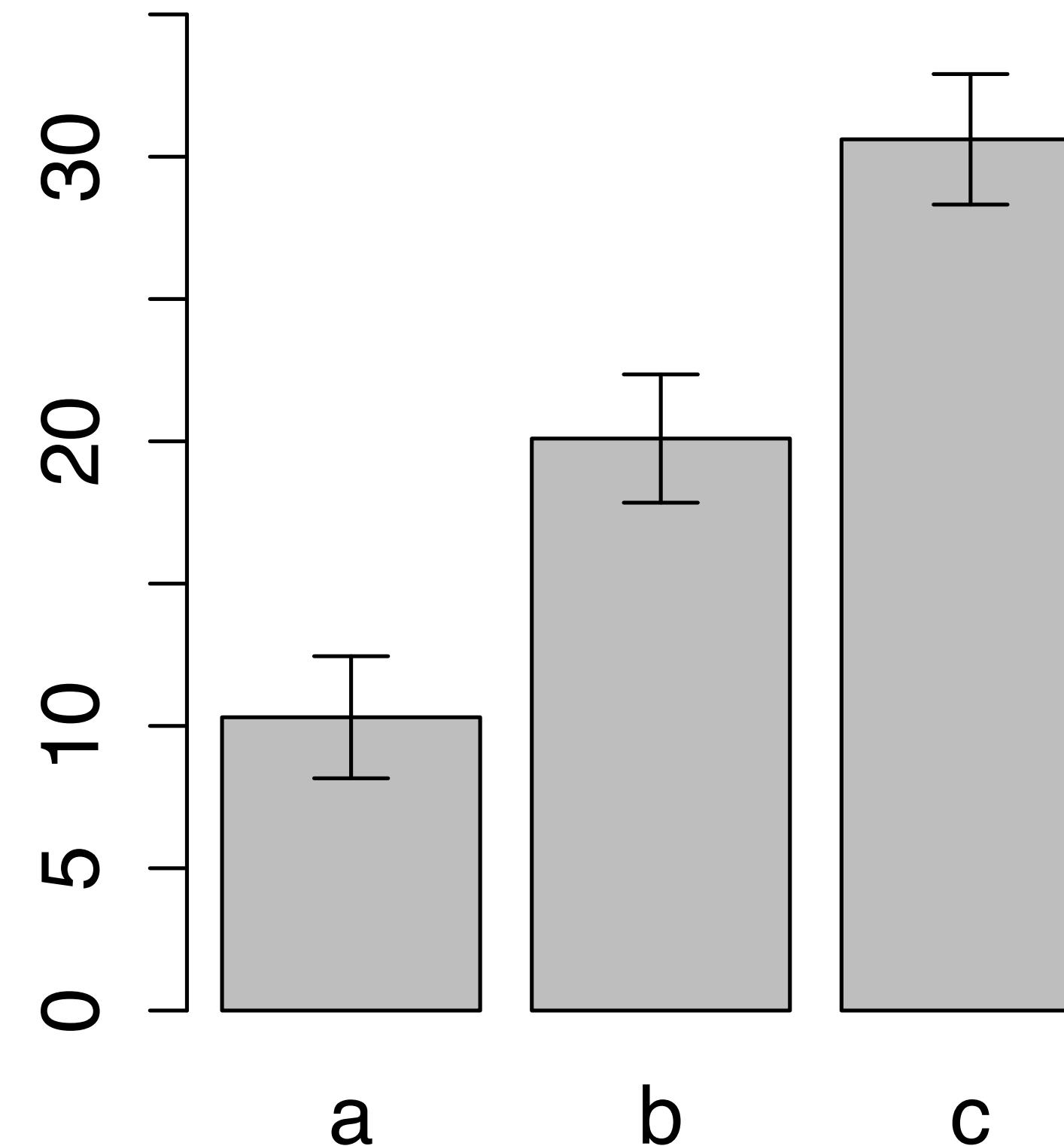


- standard deviation
- standard error of the mean
- confidence interval

always inform the reader about
which one you chooses

MAKING A BETTER FIGURE

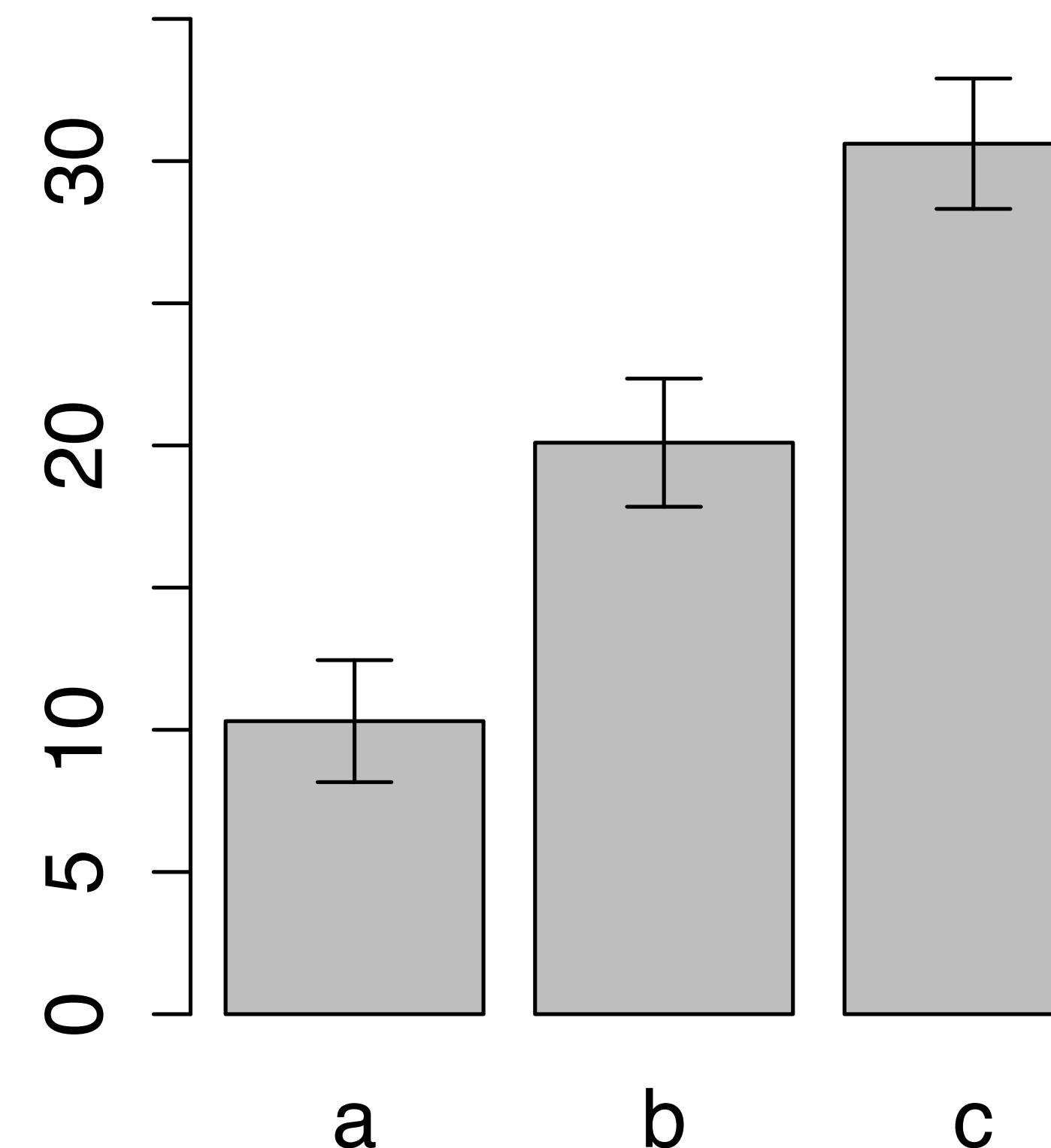
THE SHAMEFUL BARPLOT



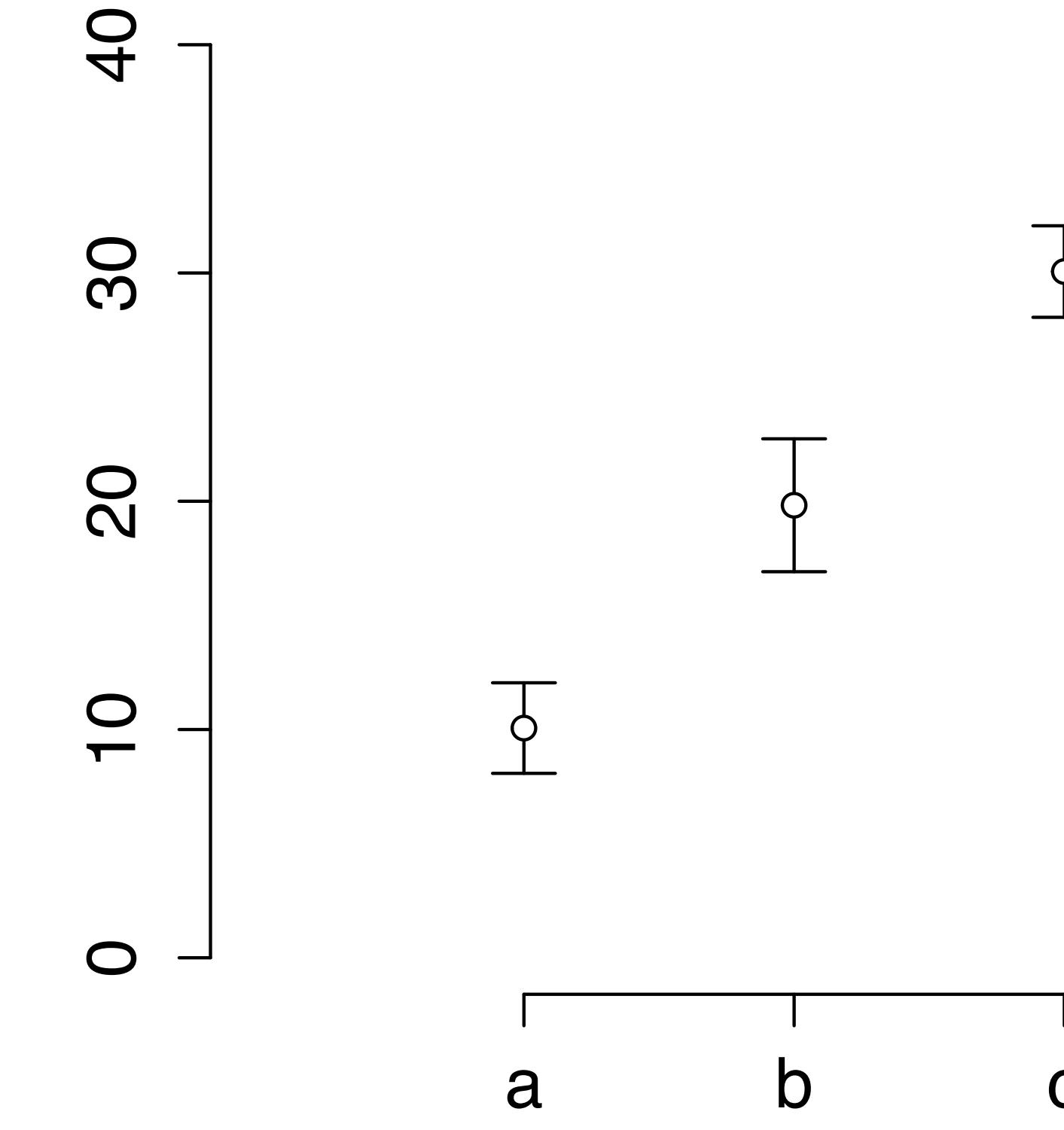
- Low data / ink ratio
- Informative but not information rich

MAKING A BETTER FIGURE

THE SHAMEFUL BARPLOT



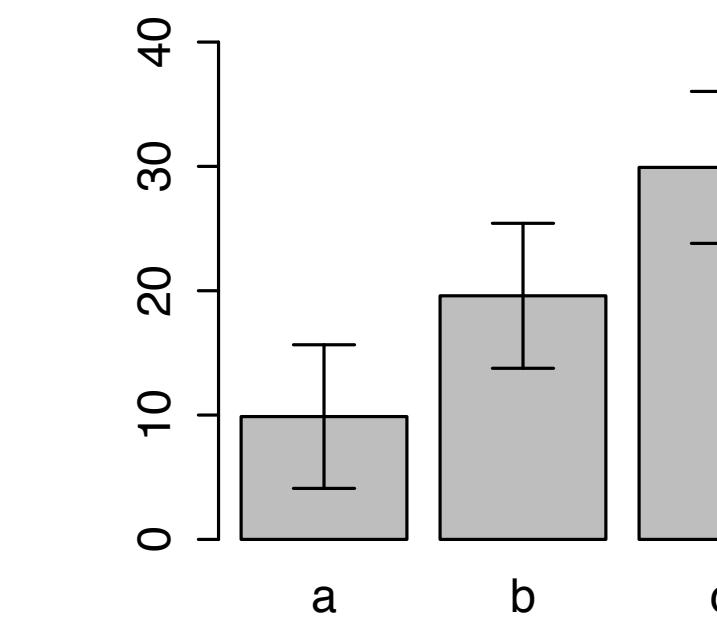
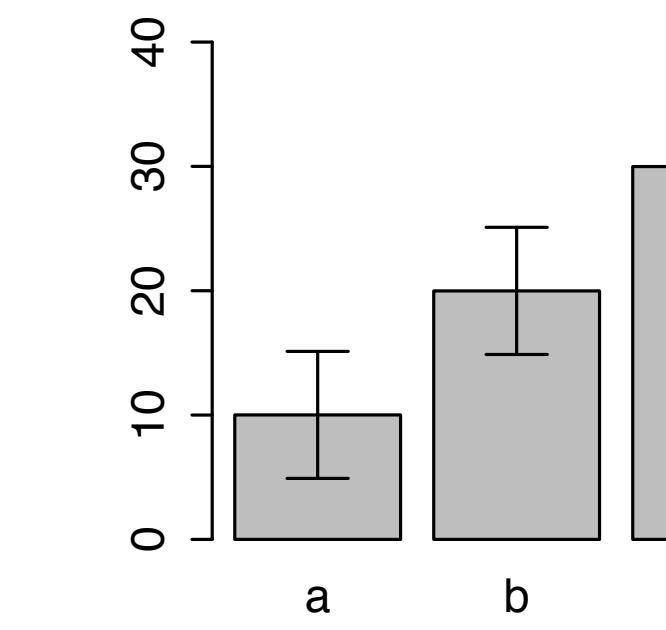
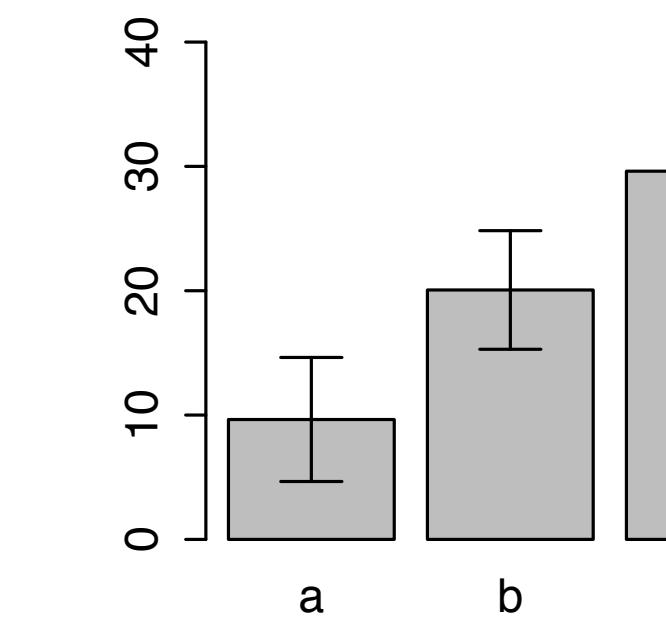
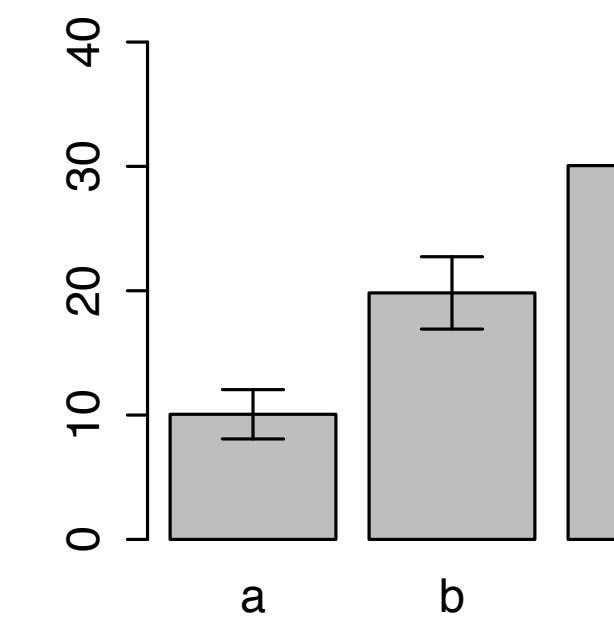
- Low data / ink ratio
- Informative but not information rich



- High data / ink ratio
- Still not information rich

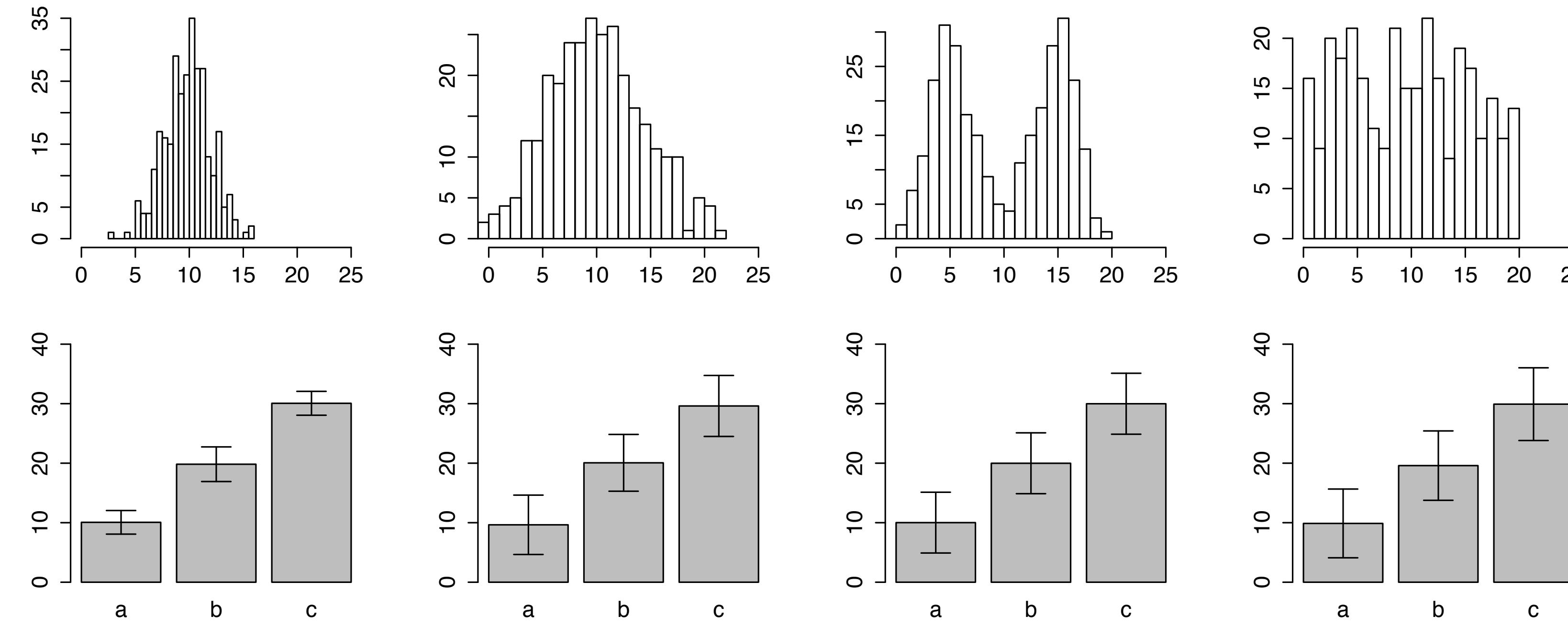
MAKING A BETTER FIGURE

THE SHAMEFUL BARPLOT



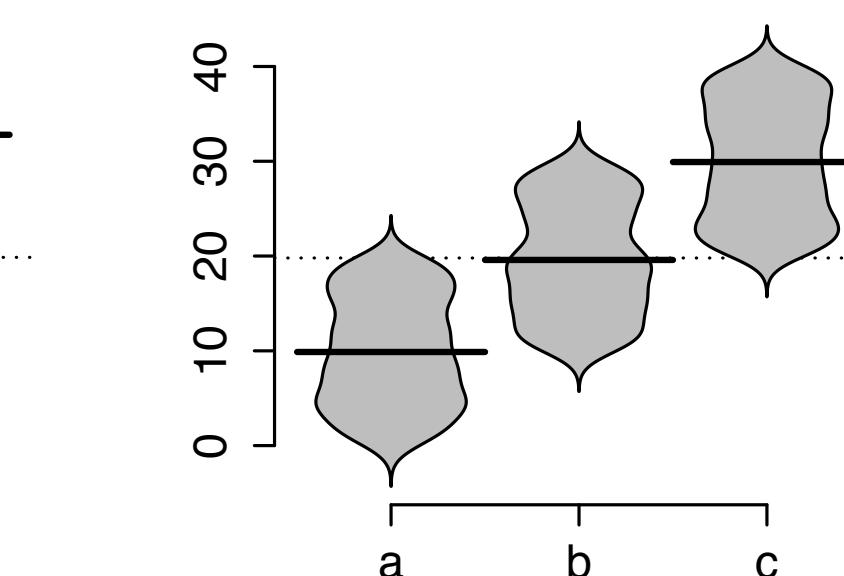
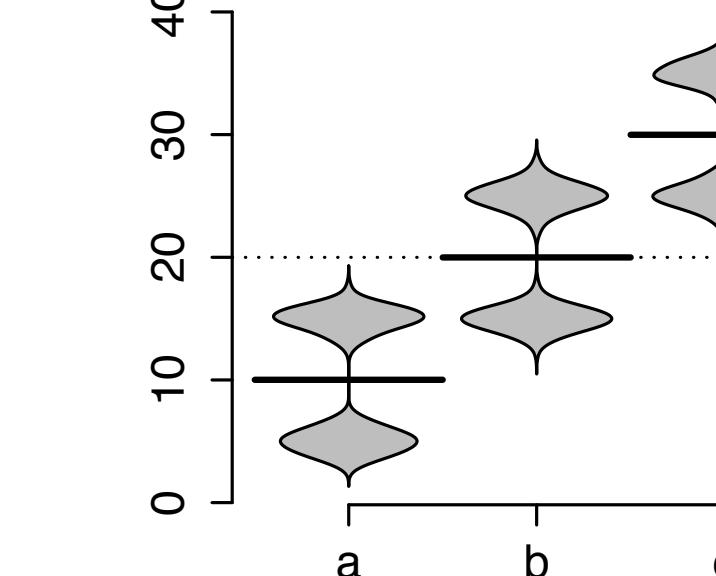
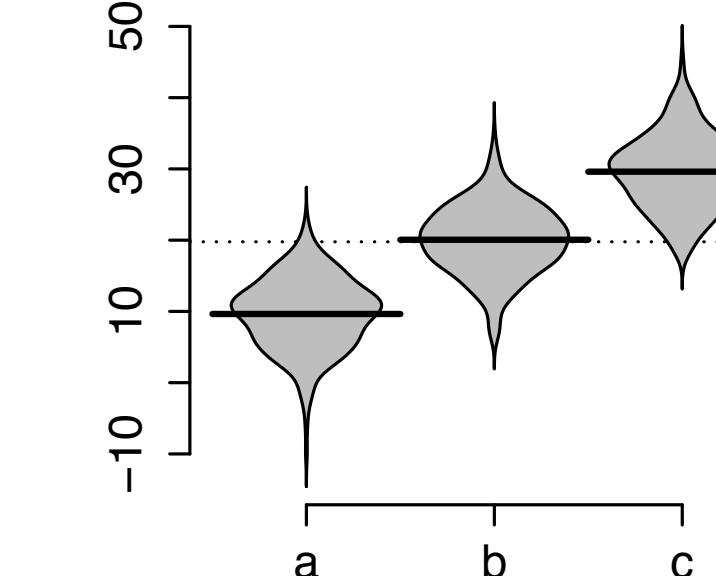
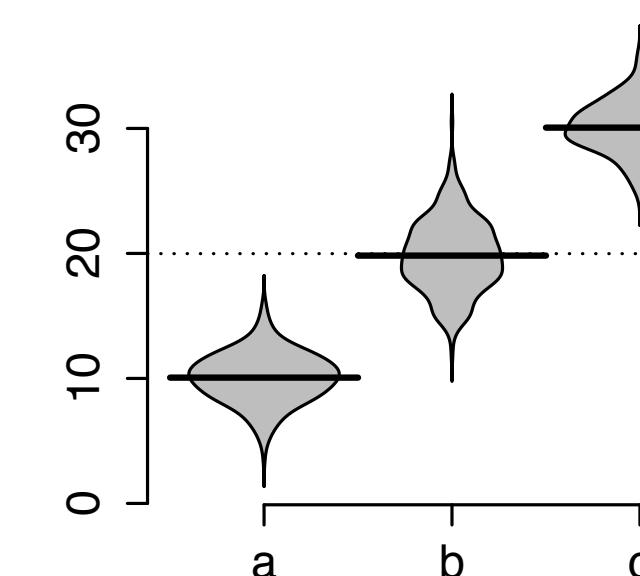
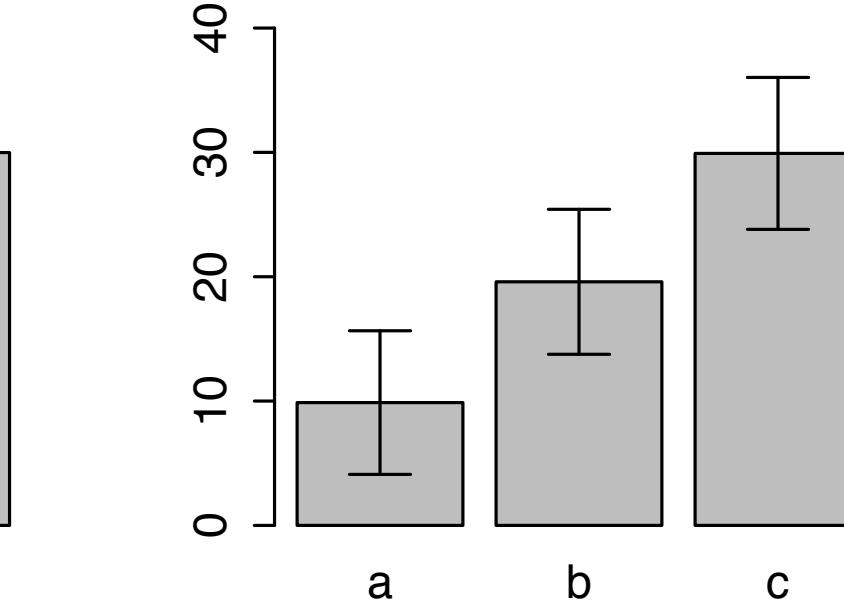
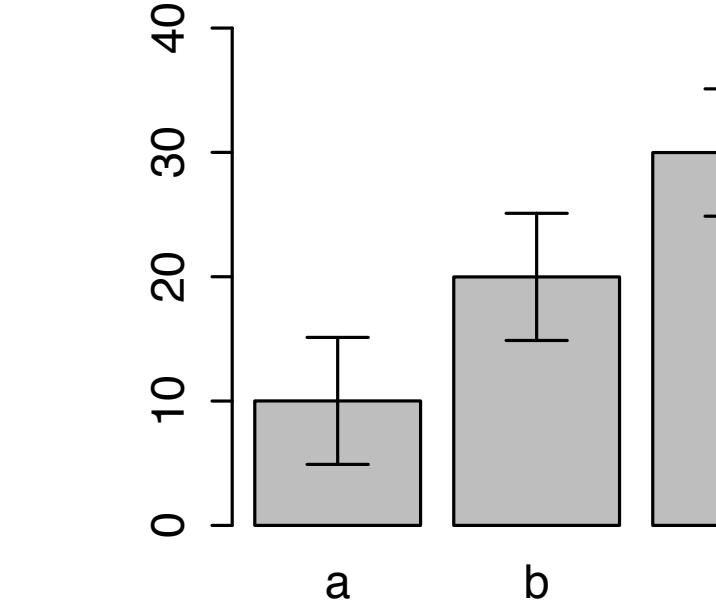
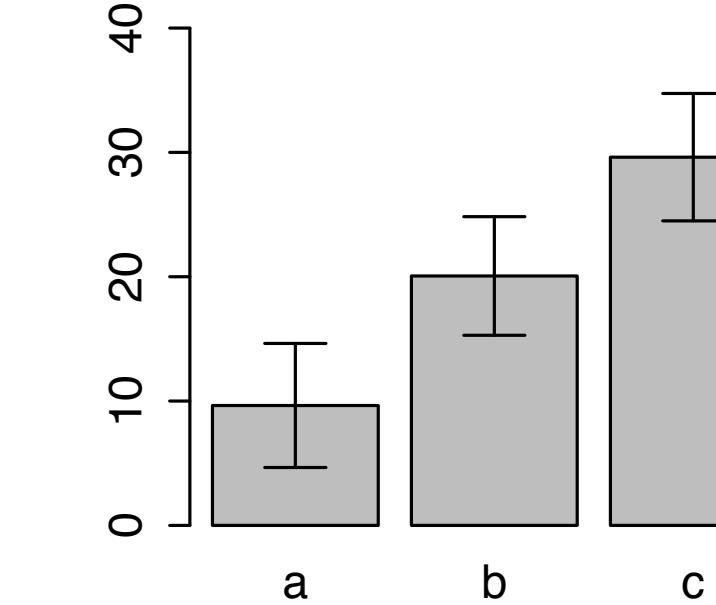
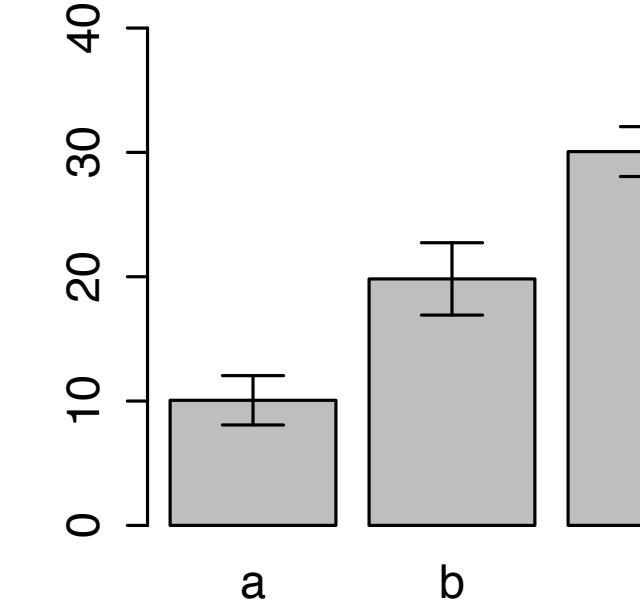
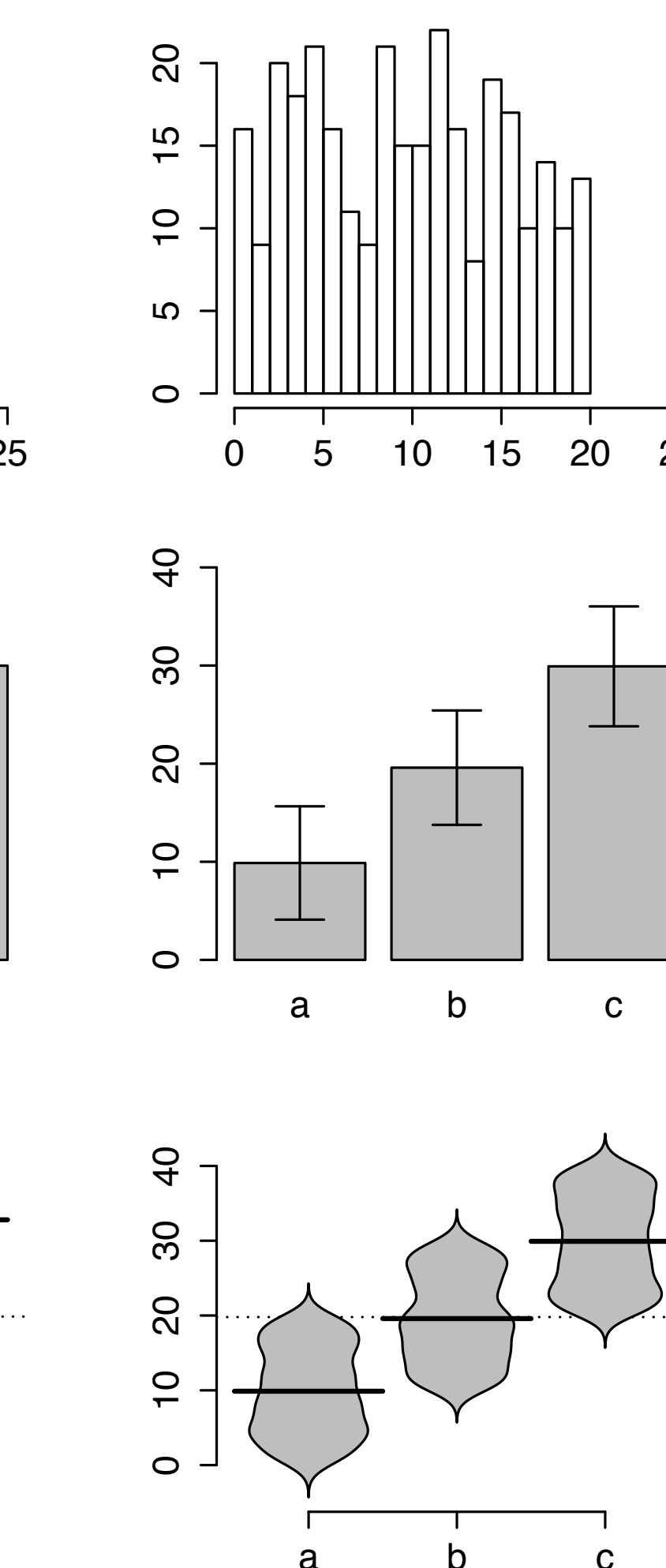
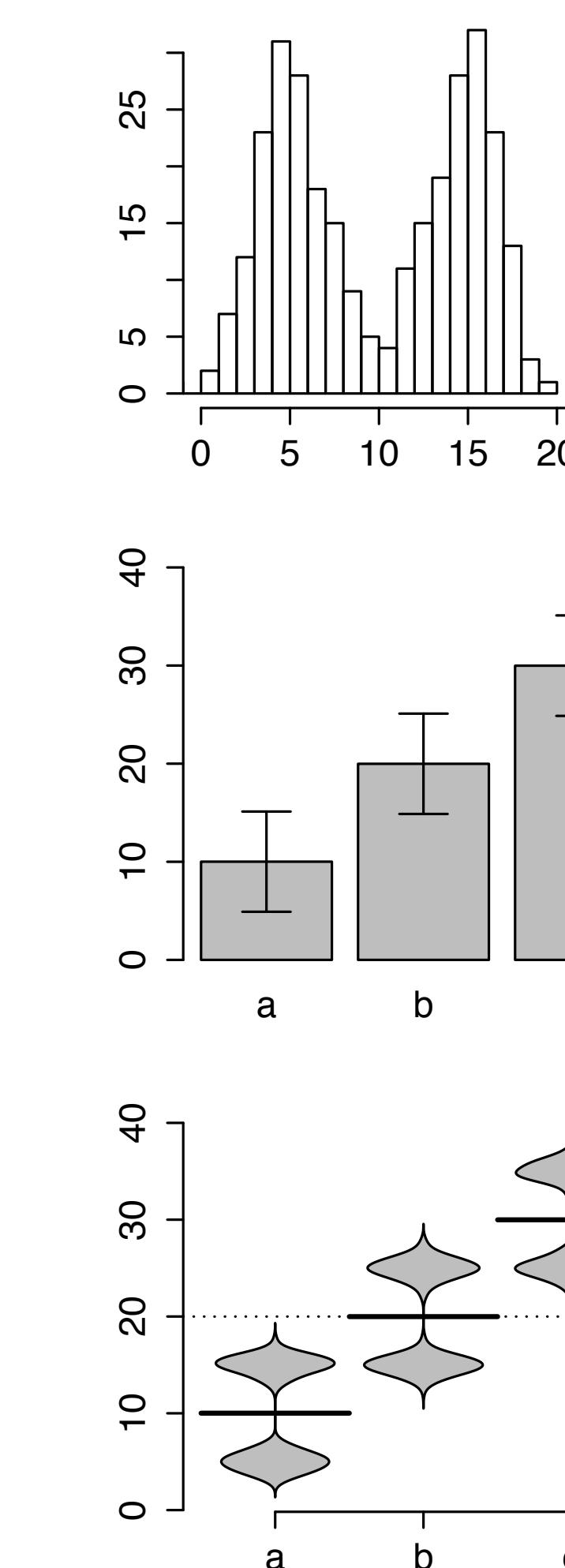
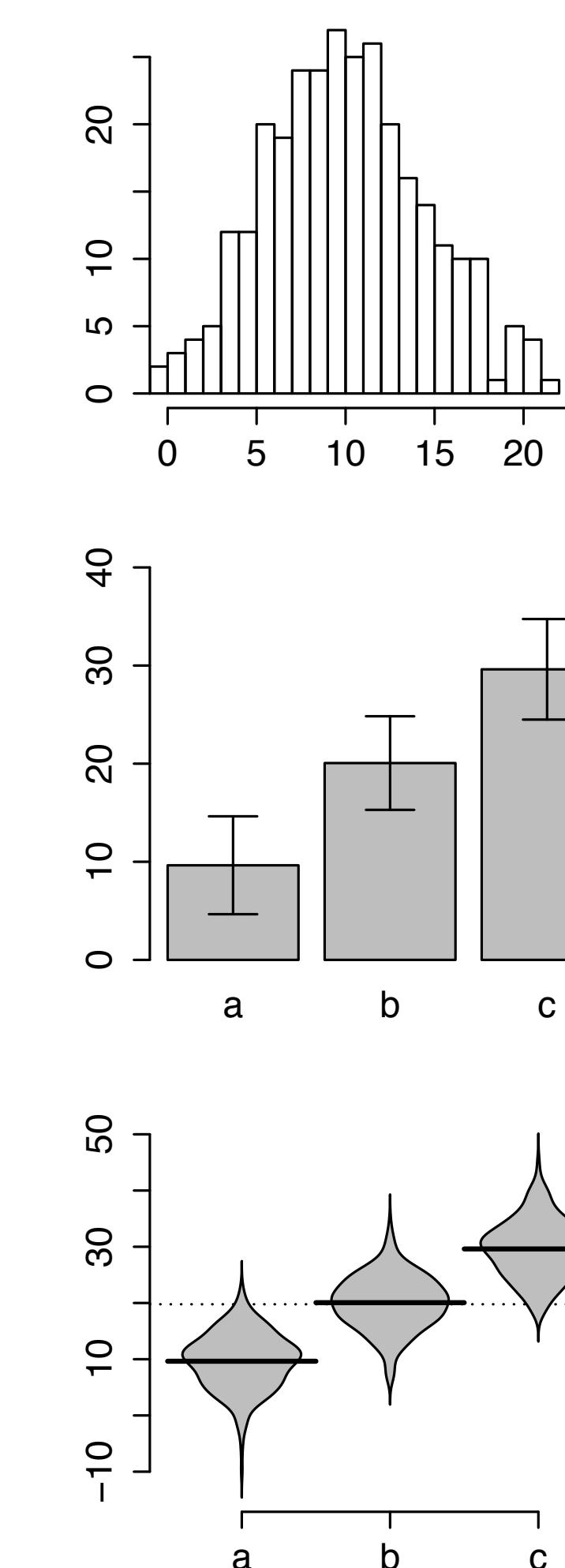
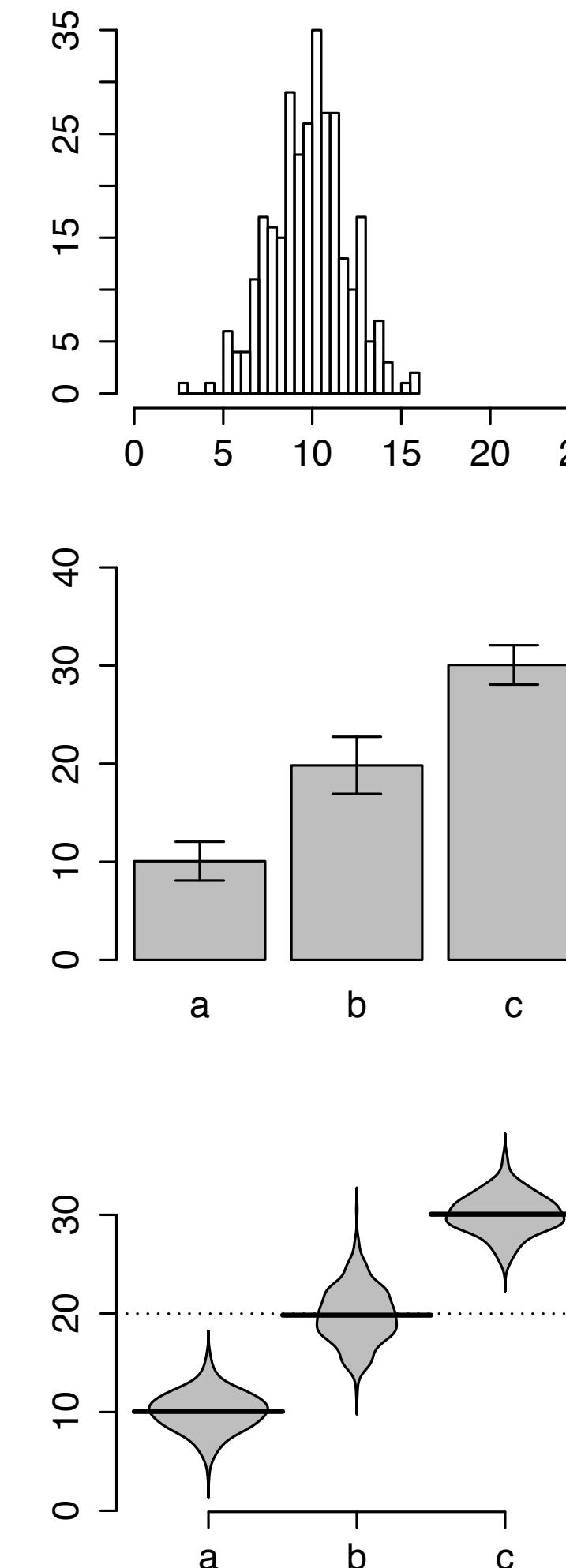
MAKING A BETTER FIGURE

THE SHAMEFUL BARPLOT



MAKING A BETTER FIGURE

THE SHAMEFUL BARPLOT

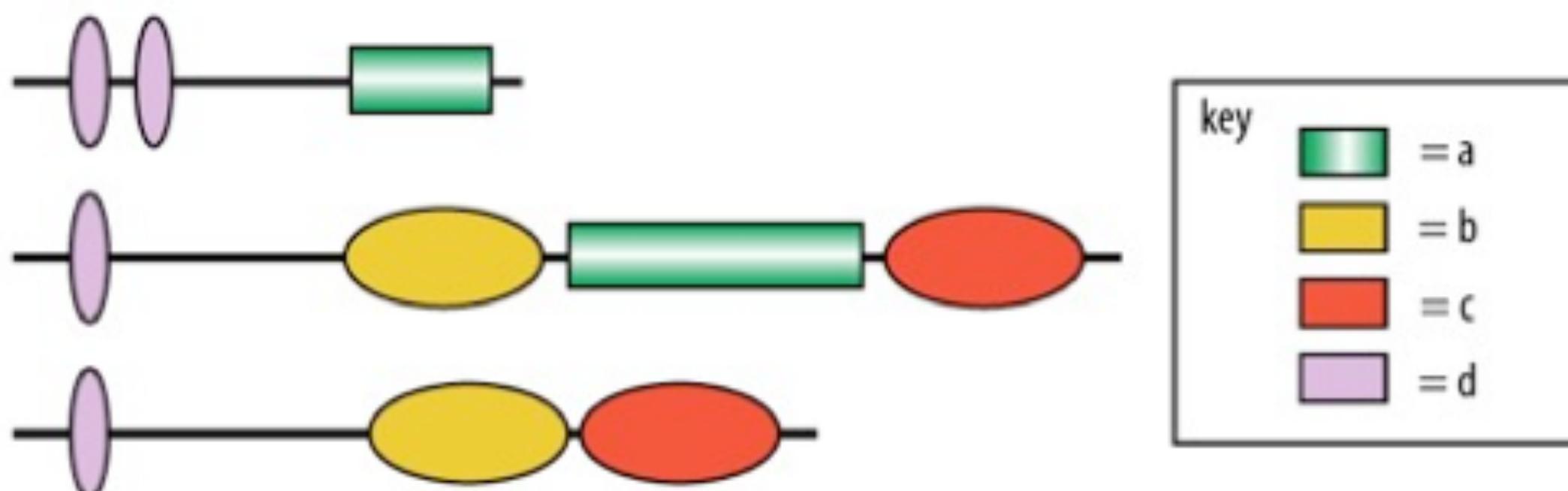
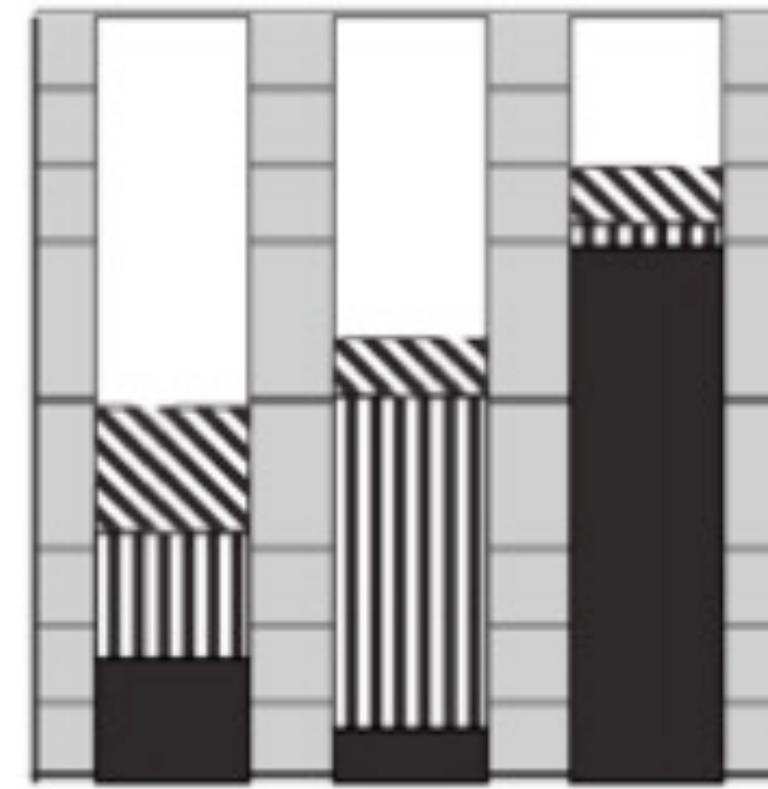


- Boxplot
- Violin plot
- Bean plot

MAKING A BETTER FIGURE

SIMPLIFYING FIGURES

chartjunk



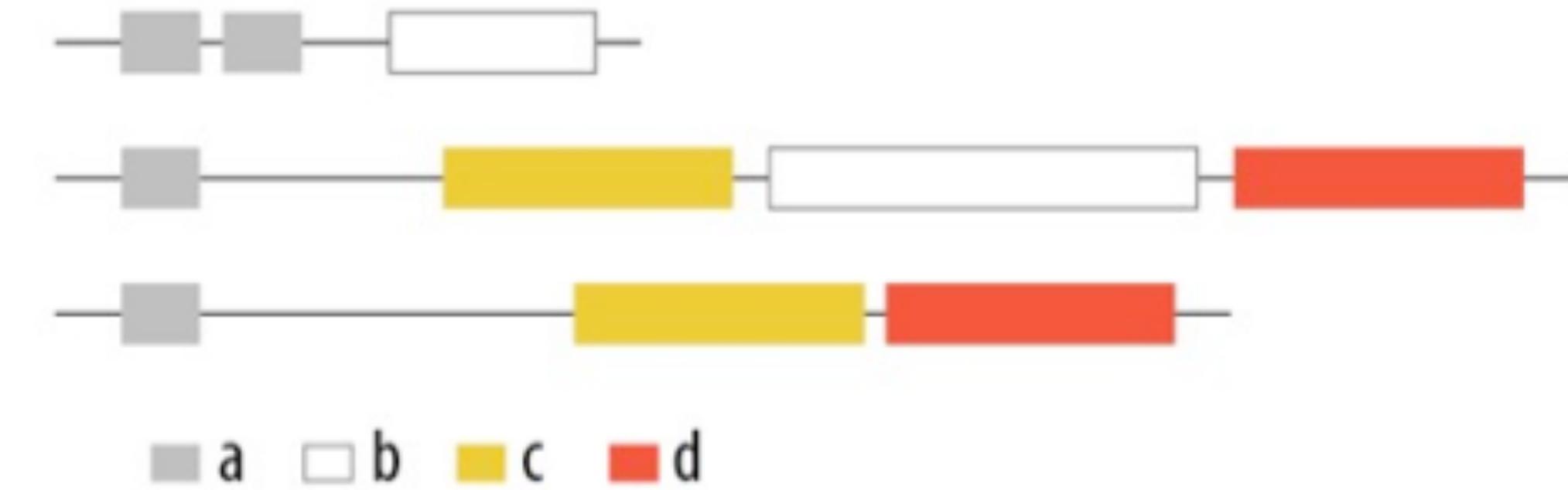
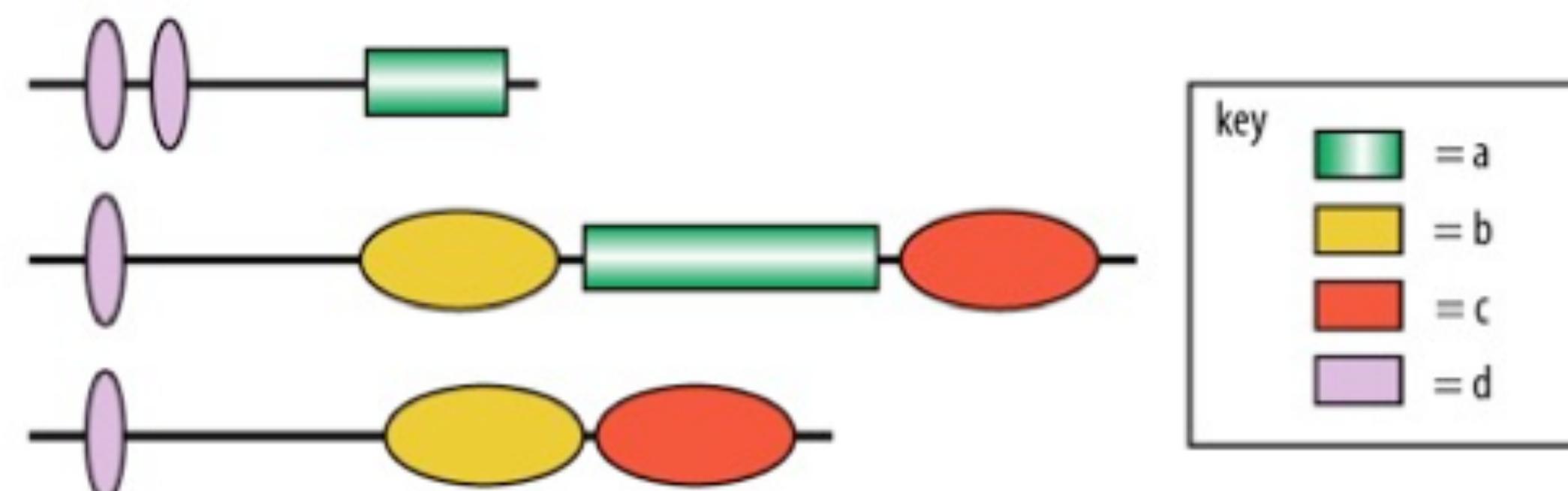
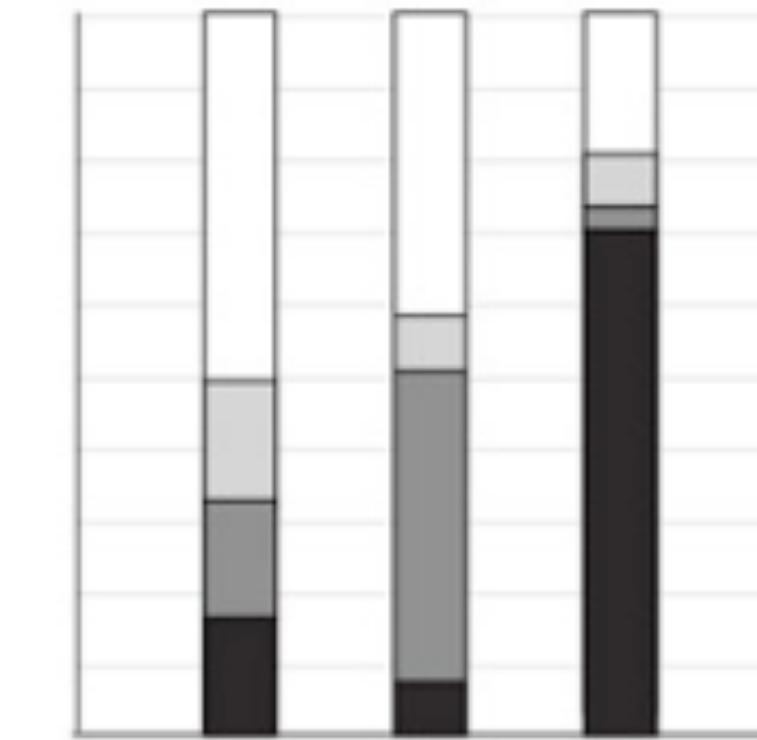
MAKING A BETTER FIGURE

SIMPLIFYING FIGURES

chartjunk

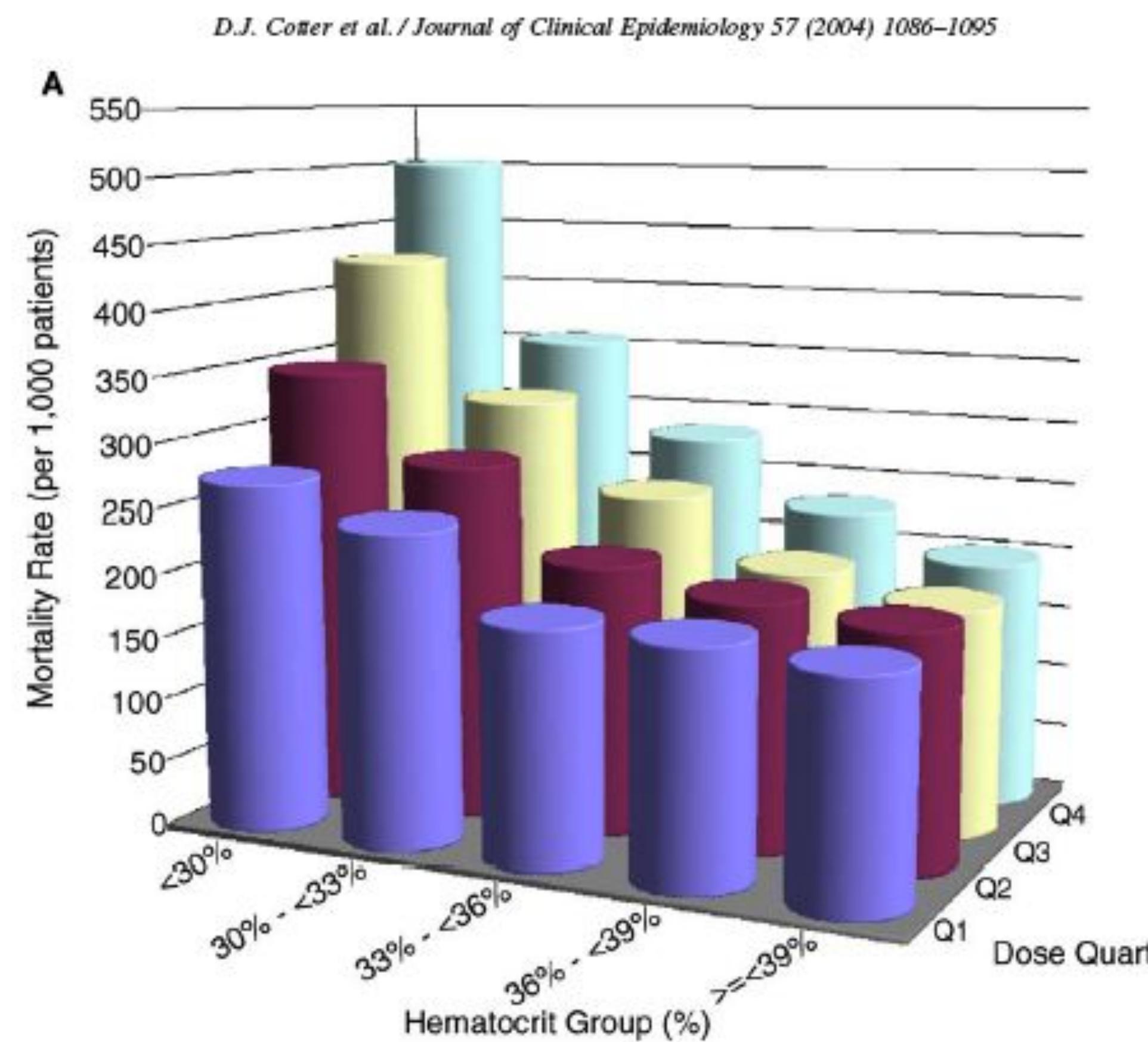


visually concise



MAKING A BETTER FIGURE

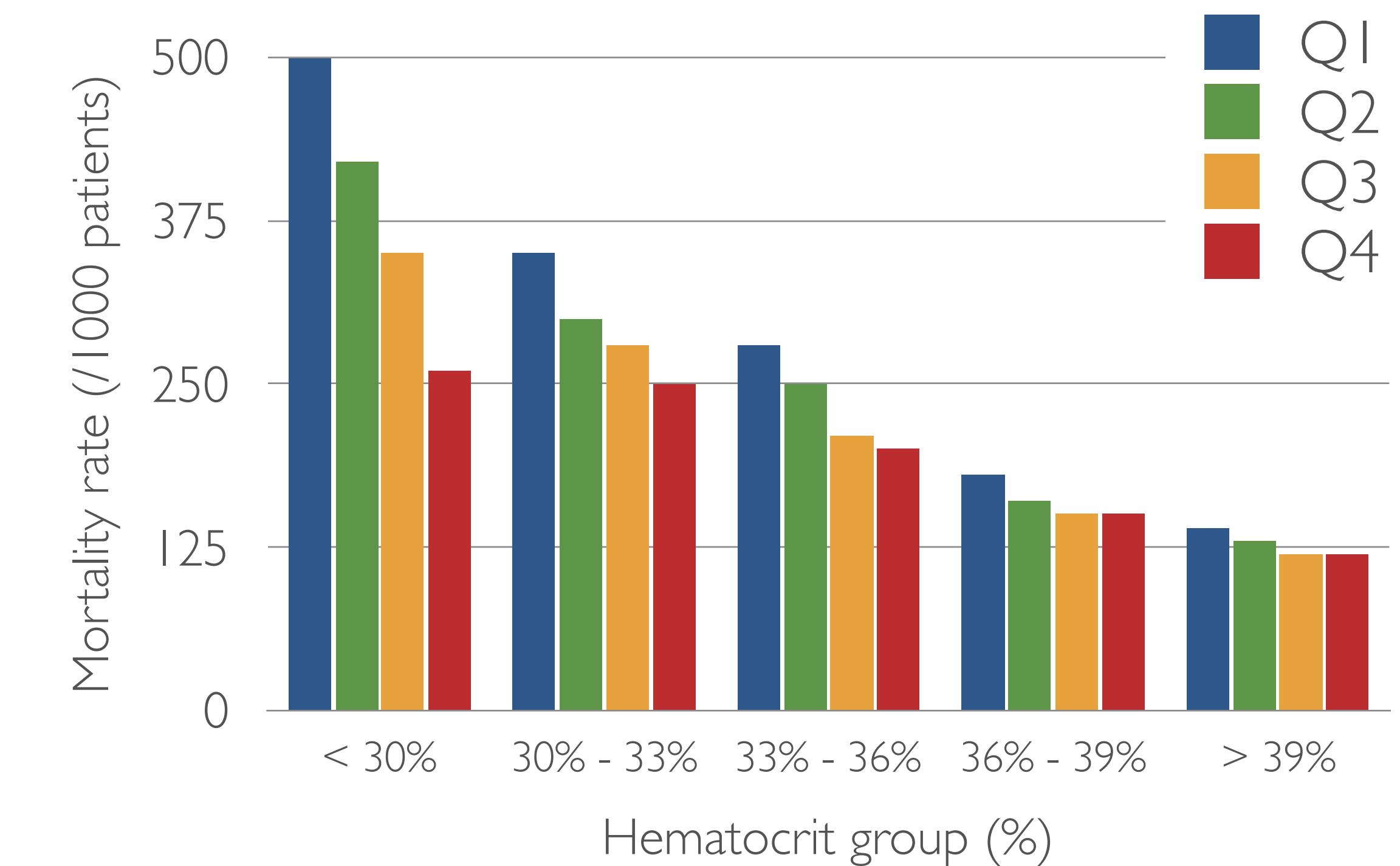
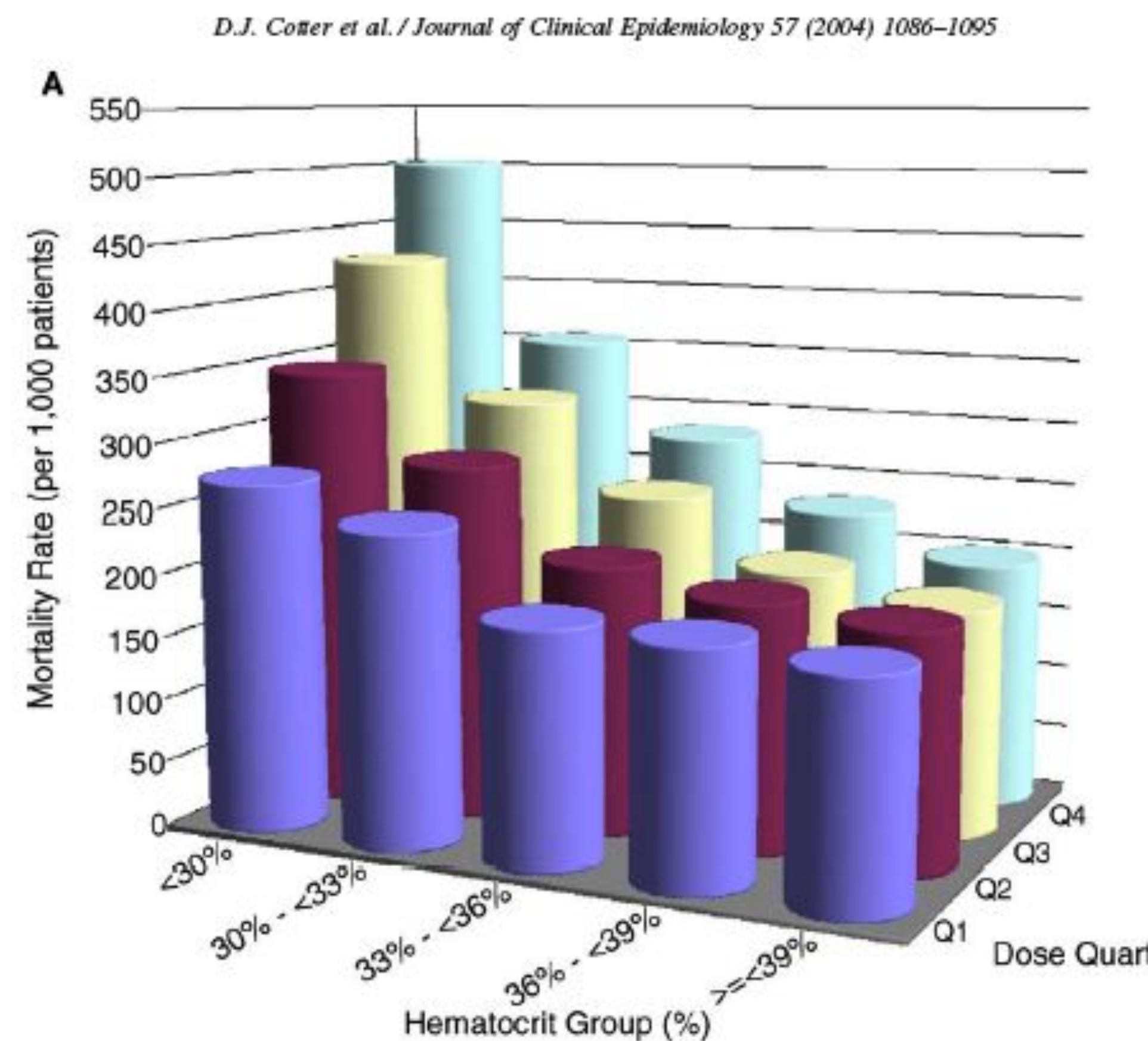
SIMPLIFYING FIGURES



Cotter DJ, et al. (2004) Hematocrit was not validated as a surrogate endpoint for survival among epoetin-treated hemodialysis patients. Journal of Clinical Epidemiology 57:1086–1095

MAKING A BETTER FIGURE

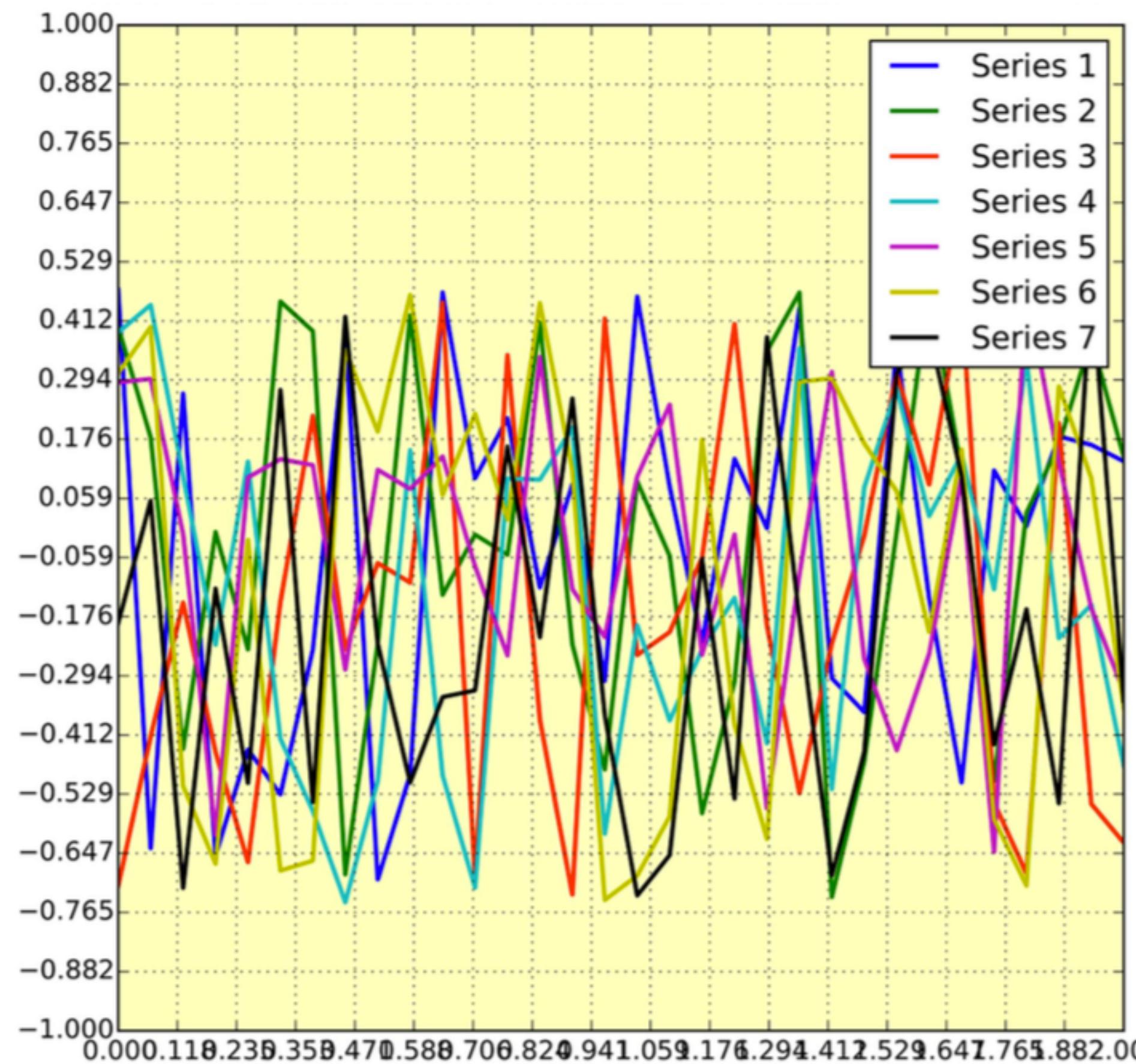
SIMPLIFYING FIGURES



Cotter DJ, et al. (2004) Hematocrit was not validated as a surrogate endpoint for survival among epoetin-treated hemodialysis patients. Journal of Clinical Epidemiology 57:1086–1095

MAKING A BETTER FIGURE

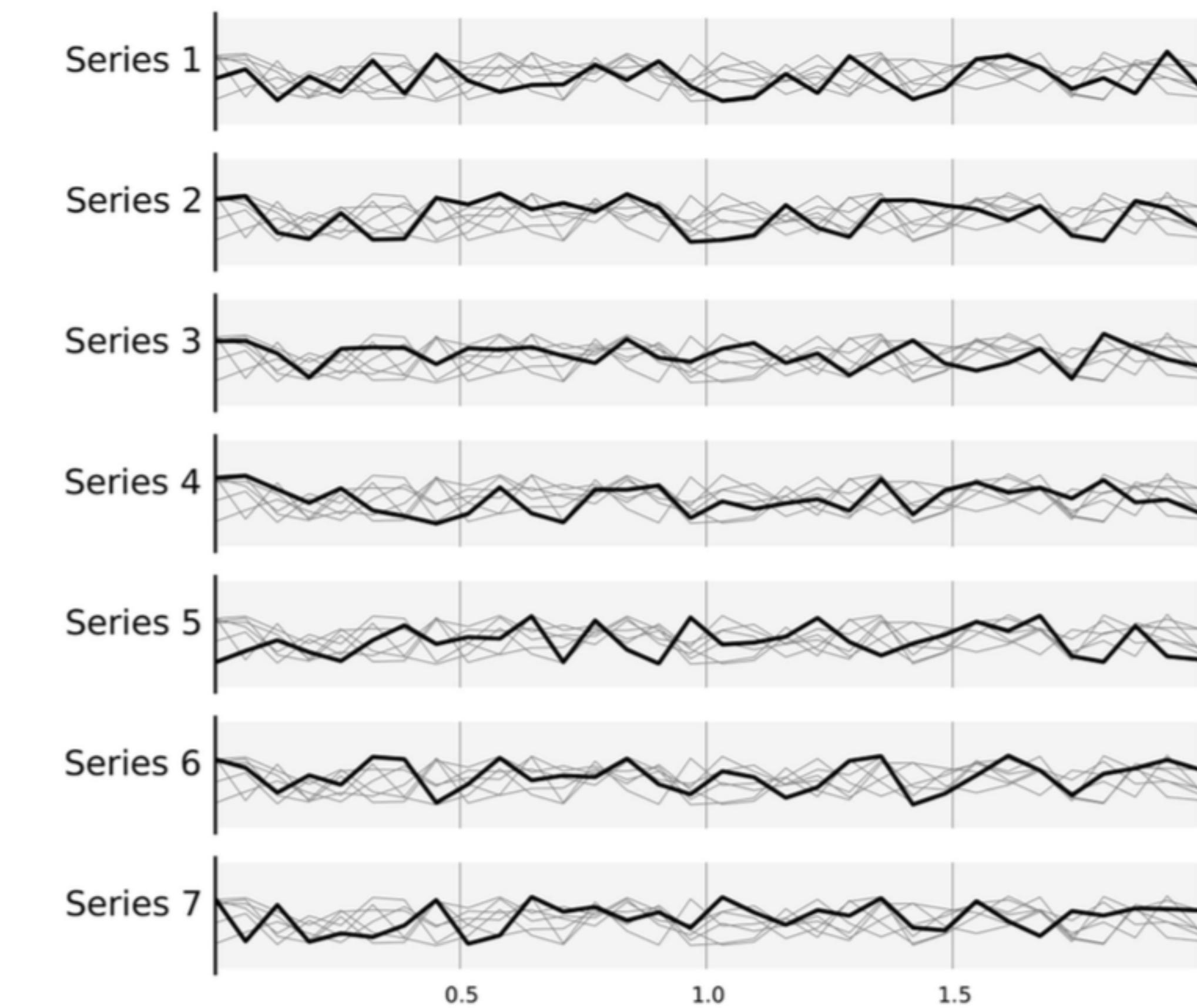
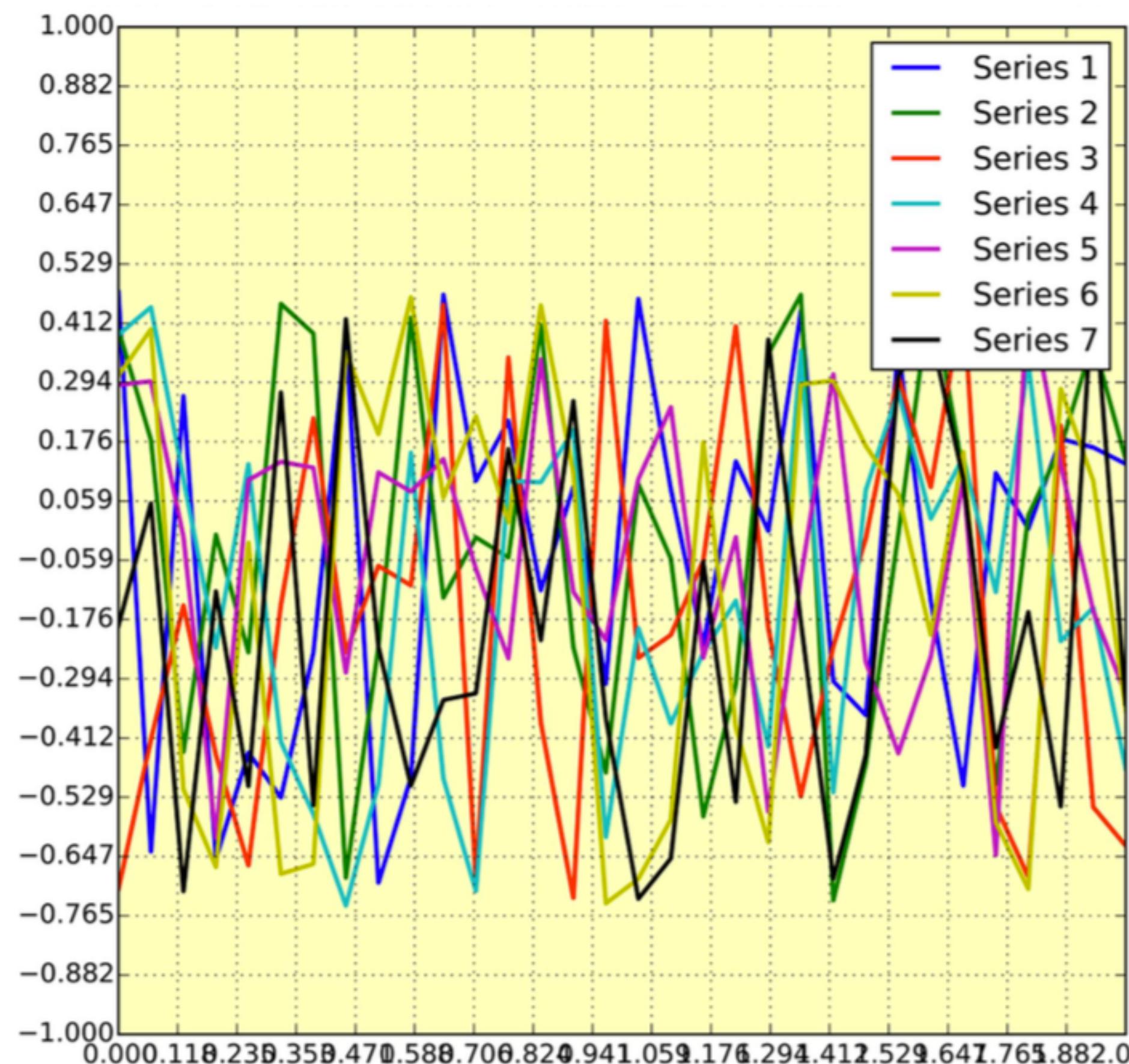
SIMPLIFYING FIGURES



Rougier, N. P., Droettboom, M., & Bourne, P. E. (2014). Ten Simple Rules for Better Figures. *PLoS Computational Biology*, 10(9), e1003833–7.

MAKING A BETTER FIGURE

SIMPLIFYING FIGURES

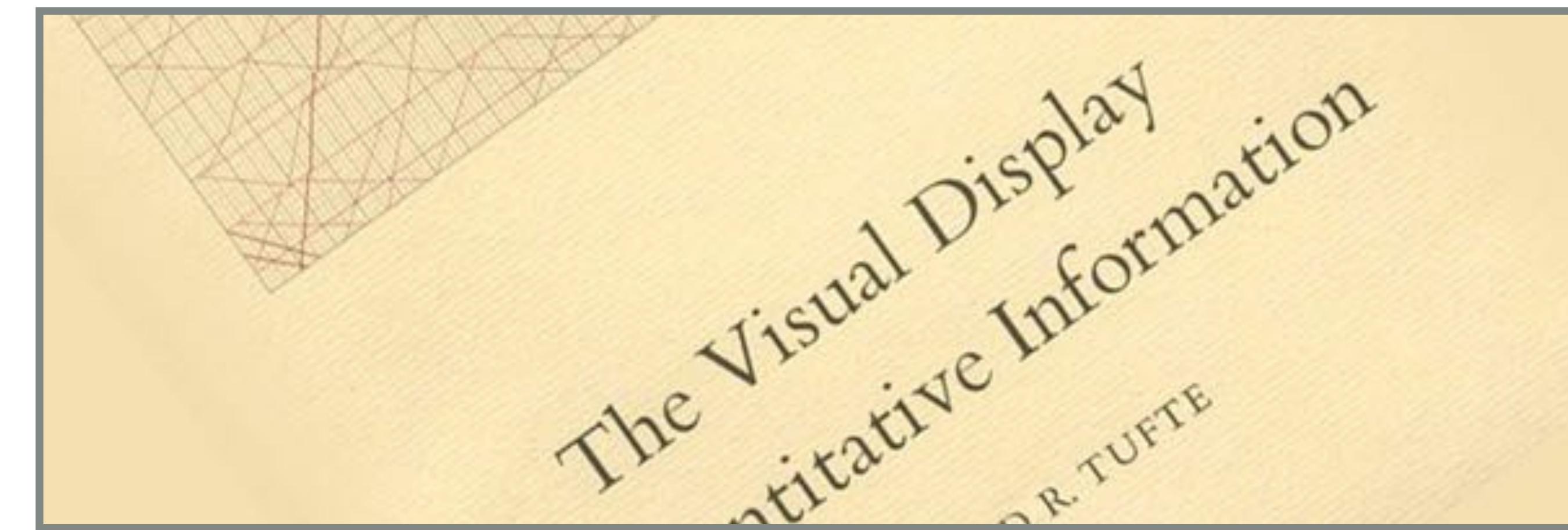


Rougier, N. P., Droettboom, M., & Bourne, P. E. (2014). Ten Simple Rules for Better Figures. *PLoS Computational Biology*, 10(9), e1003833–7.



GETTING FURTHER

ADDITIONAL RESOURCES



- Visual display of quantitative information - [Edward R Tufte](#)
- Point of view - Nature Methods - [Bang Wong](#)
- [Martin Krzywinski](#)
- Rougier et al. (2014). Ten Simple Rules for Better Figures. *PLoS Computational Biology*
- Weissgerber et al. (2015). Beyond Bar and Line Graphs. *PLoS Biology*

ADDITIONAL RESOURCES

Infographics creation

- <http://piktochart.com/>
- <https://infogram>
- <http://www.easel.ly/>



THANK YOU FOR YOUR
ATTENTION

ANY QUESTIONS?