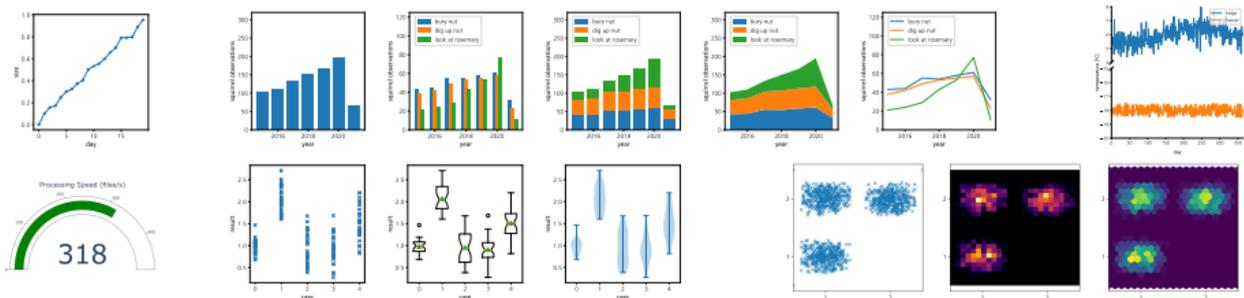


VISUALIZATION OF RESULTS

Lars Ole Schwen

MIDL 2021 Doctoral Symposium, 2021-07-02



1. Introduction

Purposes of Data Visualization

1. Introduction

Purposes of Data Visualization

- ✗ fill empty space on a page

1. Introduction

Purposes of Data Visualization

- ✘ fill empty space on a page
- ? graphic for the sake of a graphic

1. Introduction

Purposes of Data Visualization

- ✘ fill empty space on a page
- ? graphic for the sake of a graphic
- data exploration, checking for plausibility

1. Introduction

Purposes of Data Visualization

- ✘ fill empty space on a page
 - ? graphic for the sake of a graphic
 - data exploration, checking for plausibility
-
- show data in a presentation/poster

1. Introduction

Purposes of Data Visualization

- ✗ fill empty space on a page
- ? graphic for the sake of a graphic
- data exploration, checking for plausibility

- show data in a presentation/poster

- show data in a publication

1. Introduction

Purposes of Data Visualization

- ✘ fill empty space on a page
- ? graphic for the sake of a graphic
- data exploration, checking for plausibility
 - getting an overview for yourself (who hopefully knows what you are doing)
 - should be fast to change
 - does not need to be polished
- show data in a presentation/poster

- show data in a publication

1. Introduction

Purposes of Data Visualization

- ✘ fill empty space on a page
- ? graphic for the sake of a graphic
- data exploration, checking for plausibility
 - getting an overview for yourself (who hopefully knows what you are doing)
 - should be fast to change
 - does not need to be polished
- show data in a presentation/poster
 - the audience does not spend much time digesting the plot
 - should be simple without potential misunderstandings
 - should be automatically reproducible/adaptable
- show data in a publication

1. Introduction

Purposes of Data Visualization

- ✗ fill empty space on a page
- ? graphic for the sake of a graphic
- data exploration, checking for plausibility
 - getting an overview for yourself (who hopefully knows what you are doing)
 - should be fast to change
 - does not need to be polished
- show data in a presentation/poster
 - the audience does not spend much time digesting the plot
 - should be simple without potential misunderstandings
 - should be automatically reproducible/adaptable
- show data in a publication
 - the reader *may* spend more time digesting the plot
 - can be more complex, but is best kept as simple as possible

1. Introduction

Purposes of Data Visualization

- ✗ fill empty space on a page
- ? graphic for the sake of a graphic
- data exploration, checking for plausibility
 - getting an overview for yourself (who hopefully knows what you are doing)
 - should be fast to change
 - does not need to be polished
- show data in a presentation/poster
 - the audience does not spend much time digesting the plot
 - should be simple without potential misunderstandings
 - should be automatically reproducible/adaptable
- show data in a publication
 - the reader *may* spend more time digesting the plot
 - can be more complex, but is best kept as simple as possible
- support an interpretation, convey a message

1. Introduction

Purposes of Data Visualization

- ✗ fill empty space on a page
- ? graphic for the sake of a graphic
- data exploration, checking for plausibility
 - getting an overview for yourself (who hopefully knows what you are doing)
 - should be fast to change
 - does not need to be polished
- show data in a presentation/poster
 - the audience does not spend much time digesting the plot
 - should be simple without potential misunderstandings
 - should be automatically reproducible/adaptable
- show data in a publication
 - the reader *may* spend more time digesting the plot
 - can be more complex, but is best kept as simple as possible
- support an interpretation, convey a message
 - ✗ but don't mislead the audience/reader

1. Introduction

Outline

1. Introduction

2. Basic Types of Plots

3. Further Types of Plots and Visualizations

4. Tips & Tricks

2. Basic Types of Plots

Contents

1. Introduction

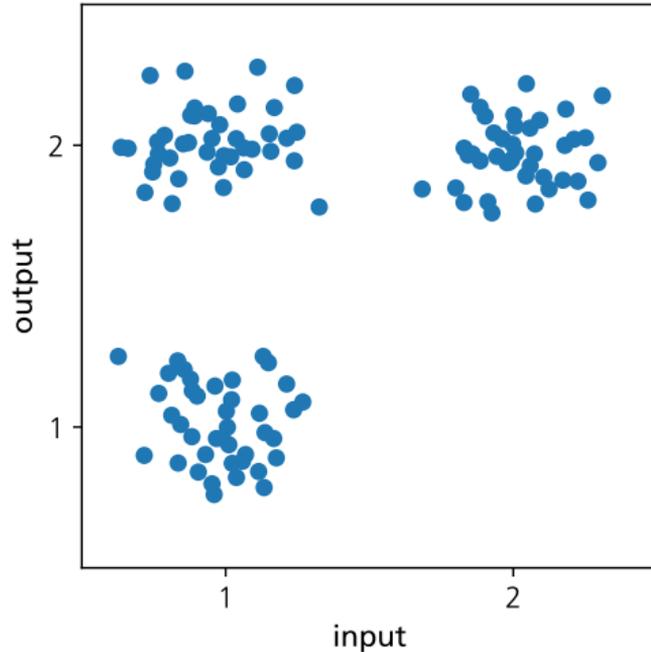
2. Basic Types of Plots

3. Further Types of Plots and Visualizations

4. Tips & Tricks

2. Basic Types of Plots

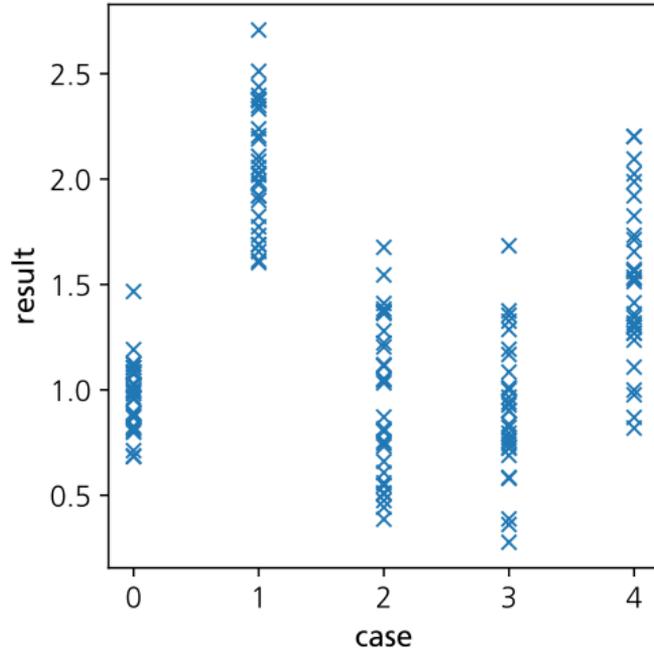
Scatter Plots



- show all data points for 2D data
- hint to clustering
- detailed overview of all data
- ✗ data points may overlap (⇒ mis-interpretations)

2. Basic Types of Plots

Scatter Plots

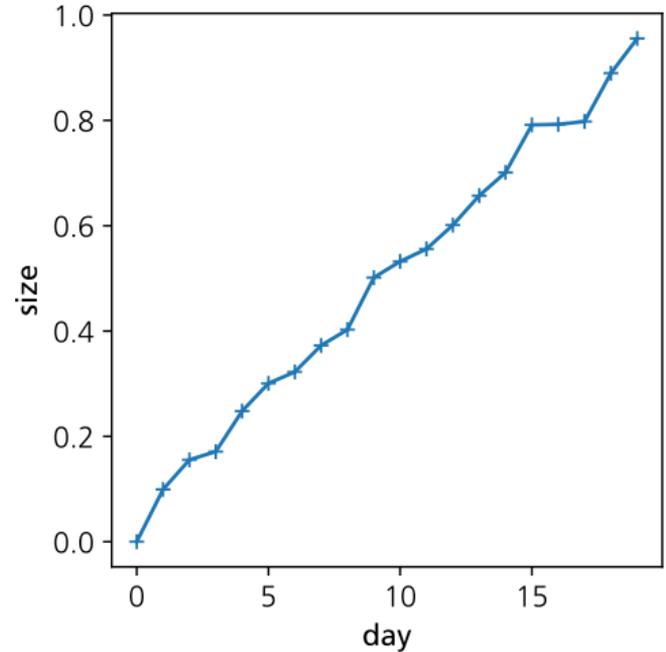


- show all data points for 2D data
- hint to clustering
- detailed overview of all data
- ✗ data points may overlap (⇒ mis-interpretations)

2. Basic Types of Plots

Line Plots

- may just 'connect the dots'
- may suggest interpolation along the axes



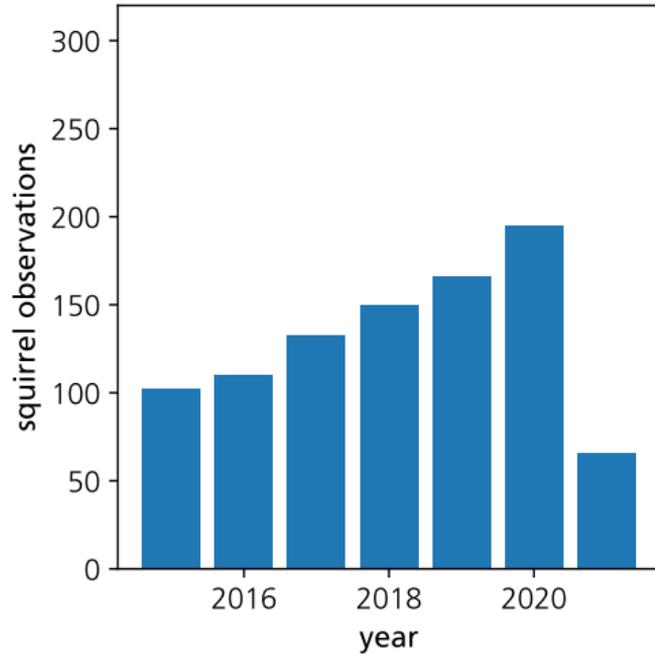
2. Basic Types of Plots

Bar Plots



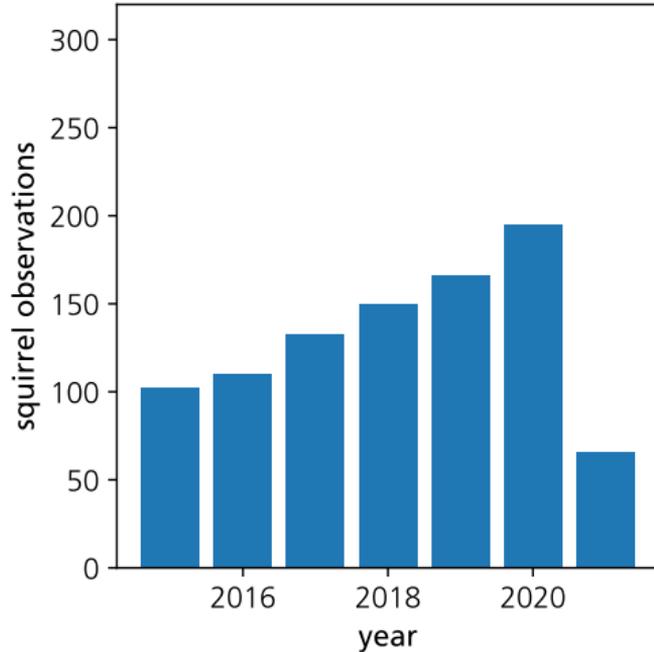
2. Basic Types of Plots

Bar Plots



2. Basic Types of Plots

Bar Plots

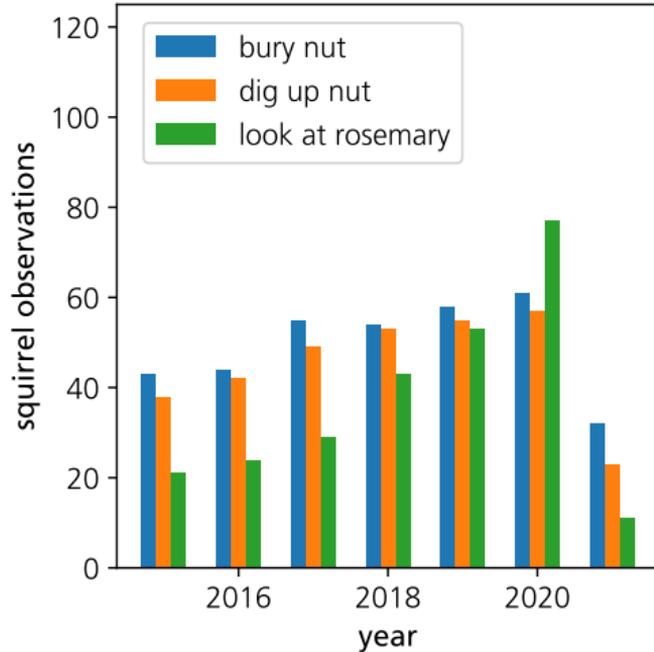


- numerical value for discrete categories
- may be horizontal or vertical



2. Basic Types of Plots

Bar Plots

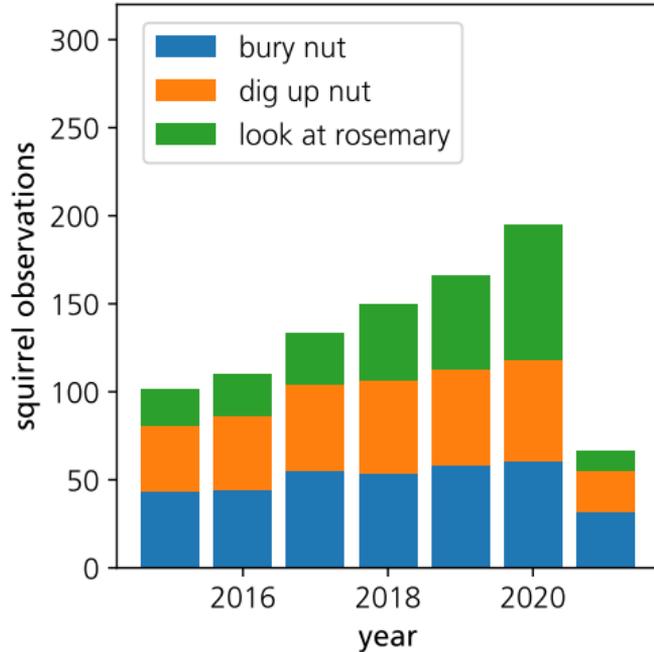


- numerical value for discrete categories
- may be horizontal or vertical
- may be grouped or stacked



2. Basic Types of Plots

Bar Plots

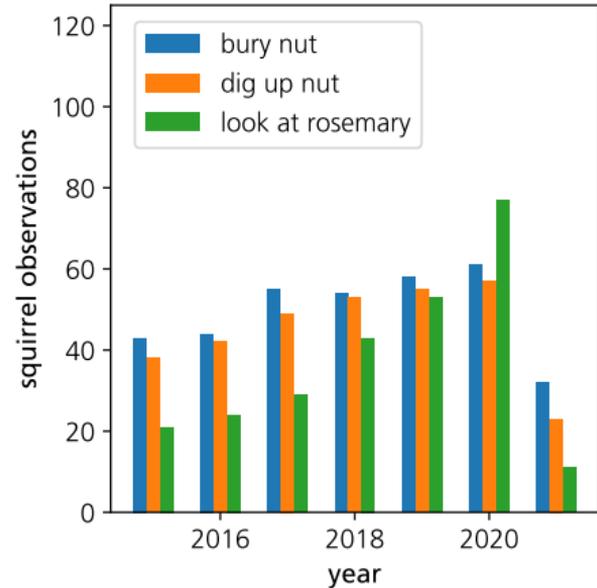
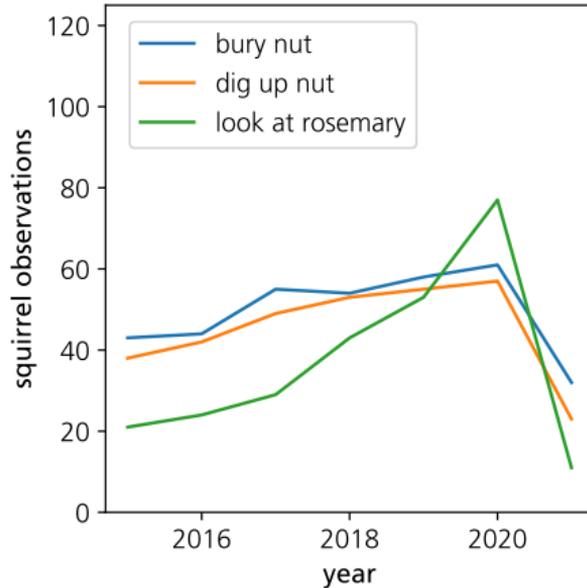


- numerical value for discrete categories
- may be horizontal or vertical
- may be grouped or stacked



2. Basic Types of Plots

Line vs. Bar Plots I

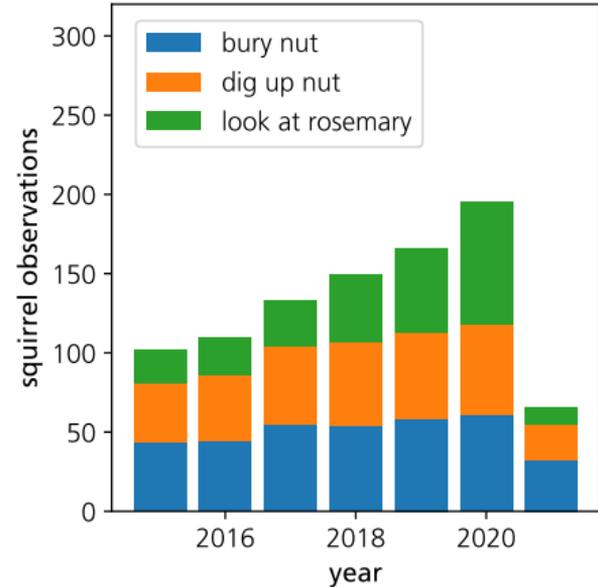
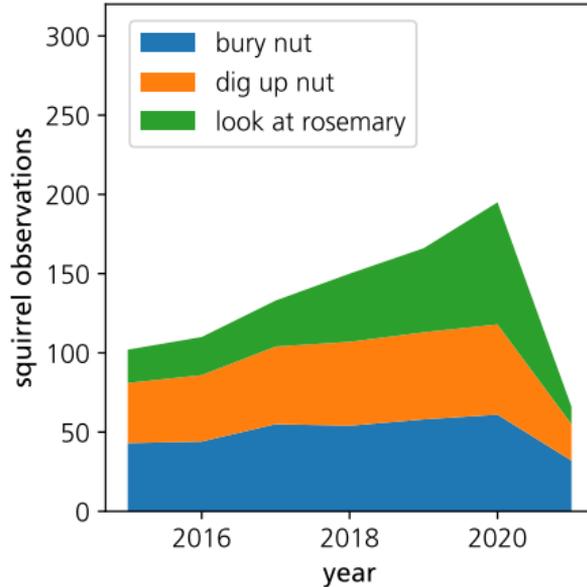


■ easier to read and compare

■ does not suggest interpolation

2. Basic Types of Plots

Line vs. Bar Plots II



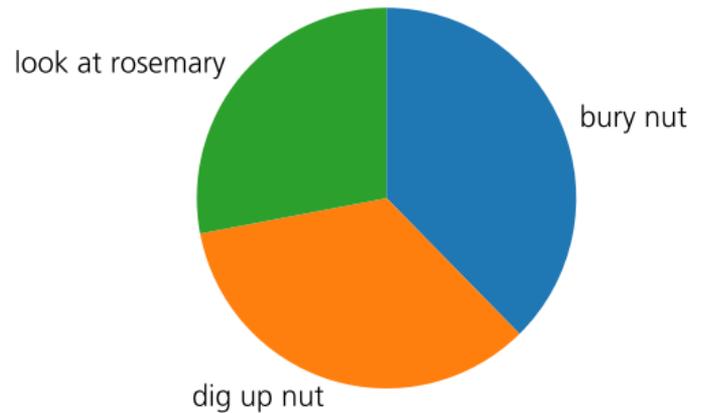
✗ areas visually misleading

✗ for both, comparison within category difficult

2. Basic Types of Plots

Pie Charts

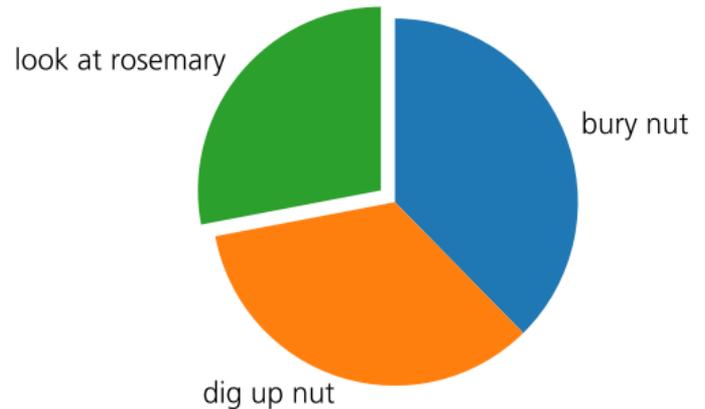
- illustrates proportion (i.e., mutually exclusive categories)
- rather uncommon in science
- difficult to compare absolute values



2. Basic Types of Plots

Pie Charts

- illustrates proportion (i.e., mutually exclusive categories)
- rather uncommon in science
- difficult to compare absolute values



2. Basic Types of Plots

Non-Plots

- plots are not always the best way to present numbers
- tables are useful for few or precise values

Country	Area in km ²	Continent
Afghanistan	652,230	Asia
Algeria	2,381,741	Africa
Andorra	468	Europe

3. Further Types of Plots and Visualizations

Contents

1. Introduction

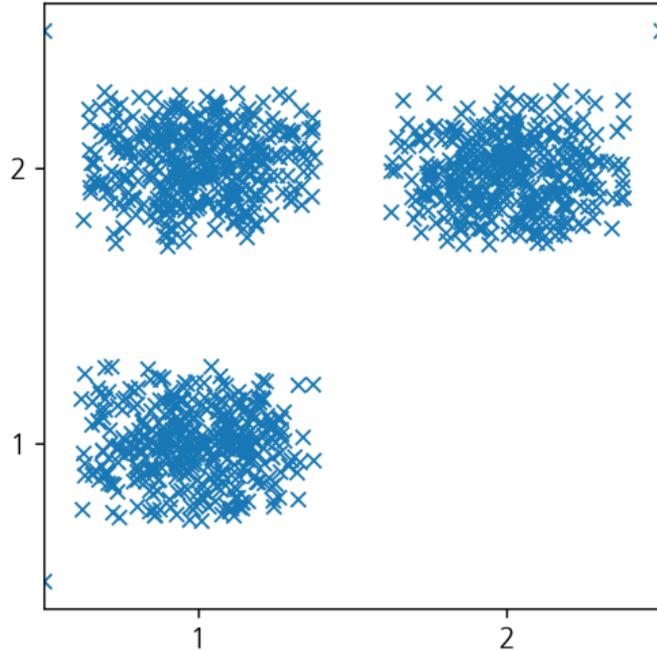
2. Basic Types of Plots

3. Further Types of Plots and Visualizations

4. Tips & Tricks

3. Further Types of Plots and Visualizations

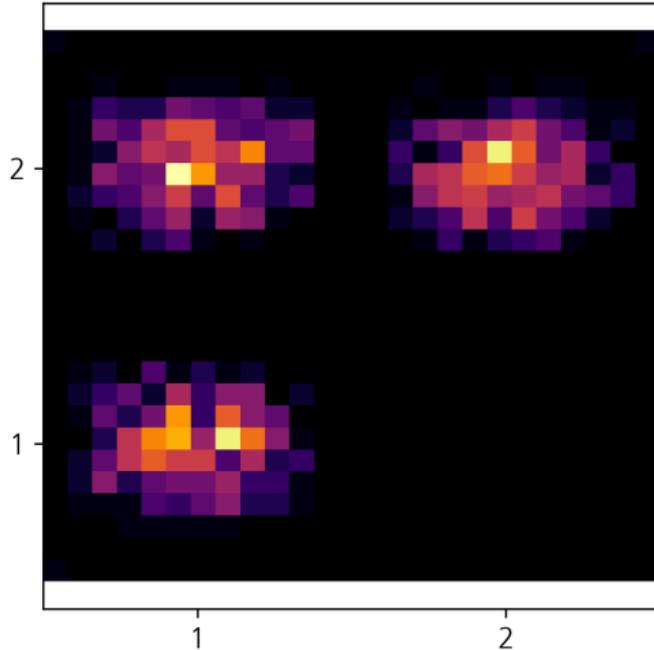
Heatmaps



✘ large number of data points: hard to read

3. Further Types of Plots and Visualizations

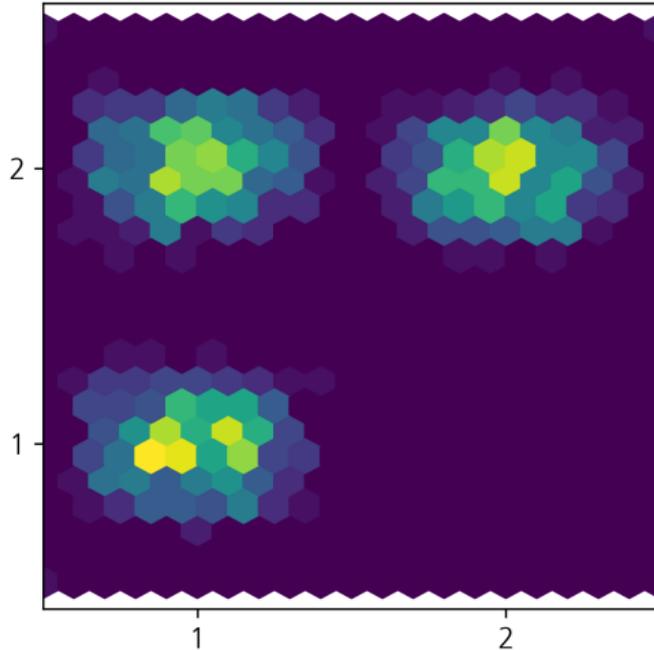
Heatmaps



- ✗ large number of data points: hard to read
- heatmap: color encodes value for 2D point

3. Further Types of Plots and Visualizations

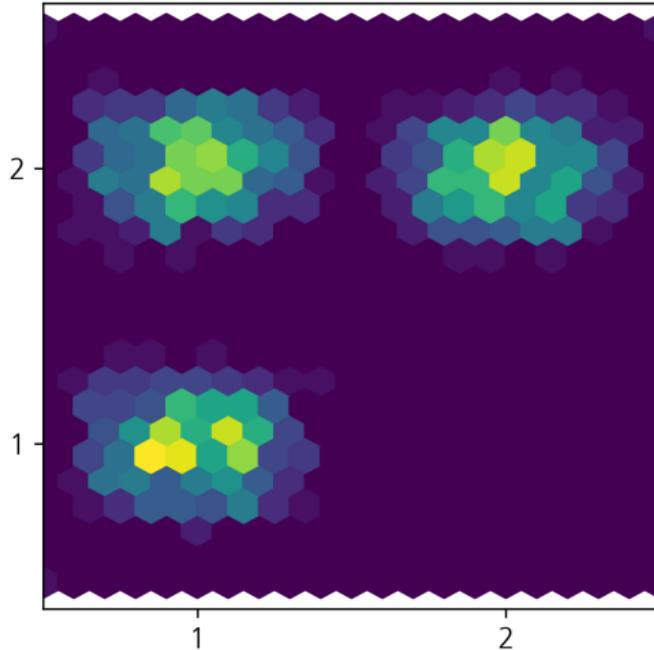
Heatmaps



- ✗ large number of data points: hard to read
- heatmap: color encodes value for 2D point

3. Further Types of Plots and Visualizations

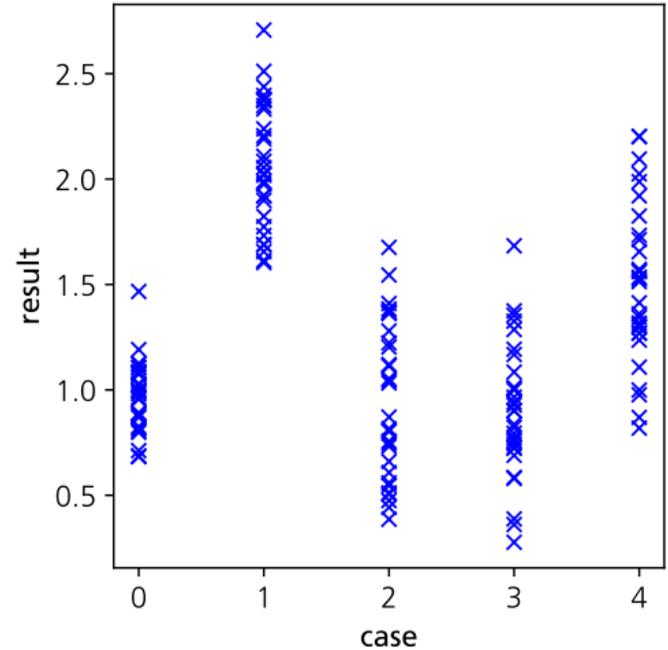
Heatmaps



- ✗ large number of data points: hard to read
- heatmap: color encodes value for 2D point
- we'll discuss color maps in a few minutes

3. Further Types of Plots and Visualizations

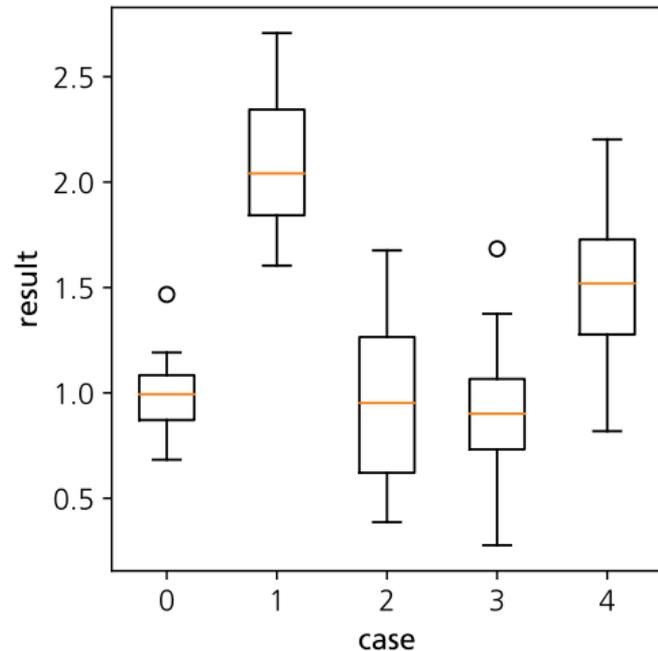
Box Plots



3. Further Types of Plots and Visualizations

Box Plots

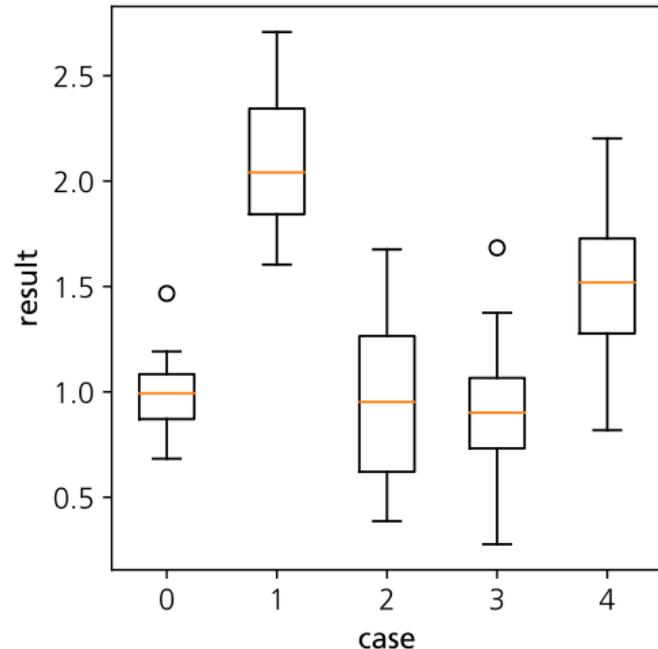
- provides compact overview over dataset
- ✗ no single boxplot standard



3. Further Types of Plots and Visualizations

Box Plots

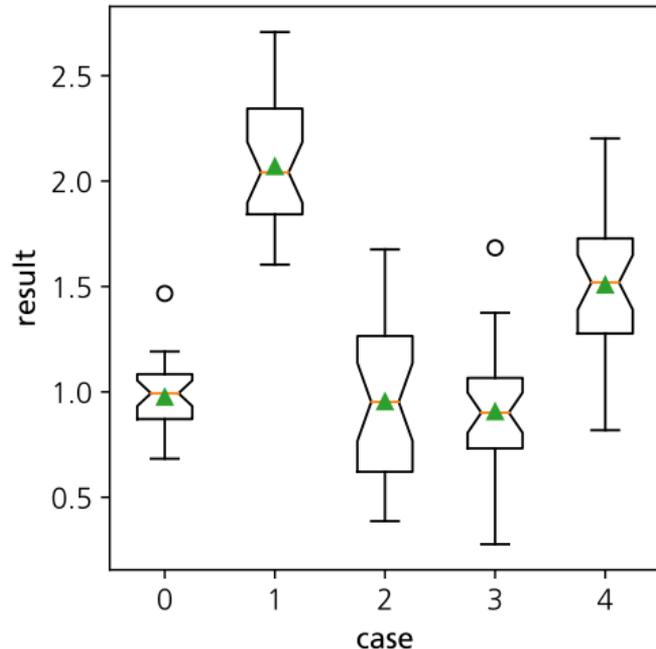
- provides compact overview over dataset
- ✗ no single boxplot standard
- here:
 - line: median
 - box: 1st to 3rd quartile (IQR)
 - whiskers: mean \pm 1.5 IQR
 - range of 'extremal values'
 - markers: points outside whiskers
 - 'outliers'



3. Further Types of Plots and Visualizations

Box Plots

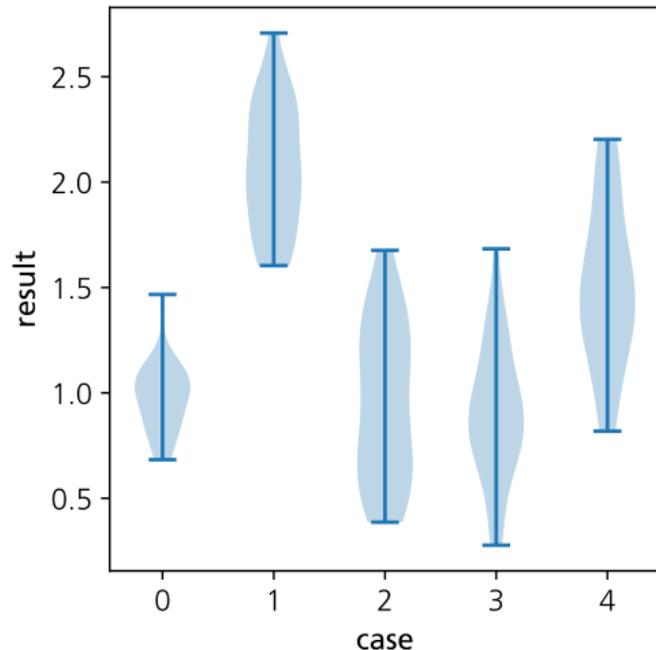
- provides compact overview over dataset
- ✗ no single boxplot standard
- here:
 - line: median
 - box: 1st to 3rd quartile (IQR)
 - whiskers: mean \pm 1.5 IQR
 - range of 'extremal values'
 - markers: points outside whiskers
 - 'outliers'
 - notches: confidence interval of median
 - triangle: mean



3. Further Types of Plots and Visualizations

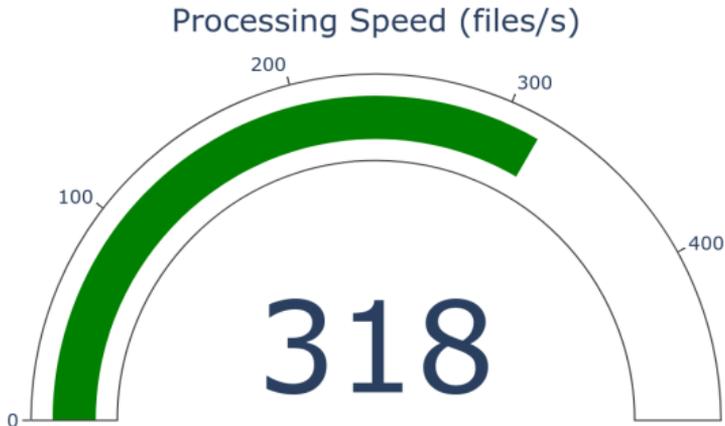
Violin Plots

- provides histogram (kernel density estimate)
- not standardized, requires explanation



3. Further Types of Plots and Visualizations

Gauge Plots



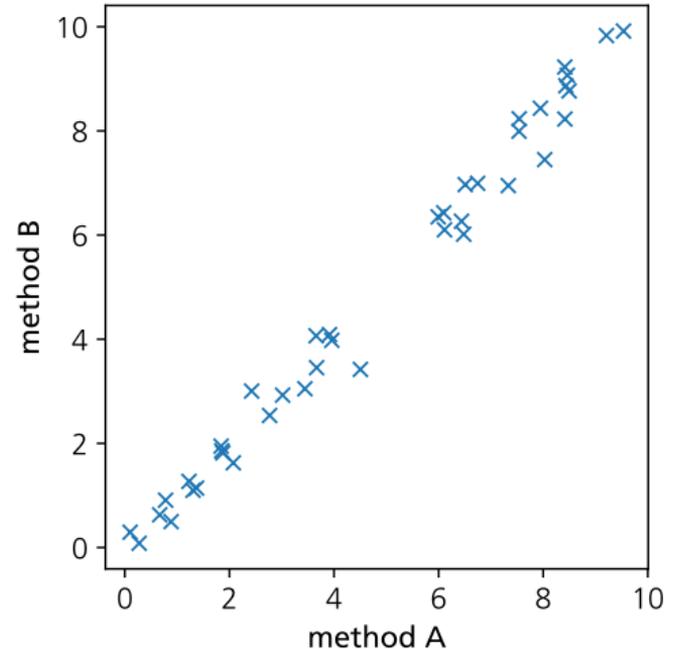
useful for

- performance measures
- concept graphics, dashboards

3. Further Types of Plots and Visualizations

Bland–Altman Plots

- aka Tukey mean-difference plot
- purpose: comparison of two values that “should be the same”
 - e.g., when comparing two algorithmic approaches



3. Further Types of Plots and Visualizations

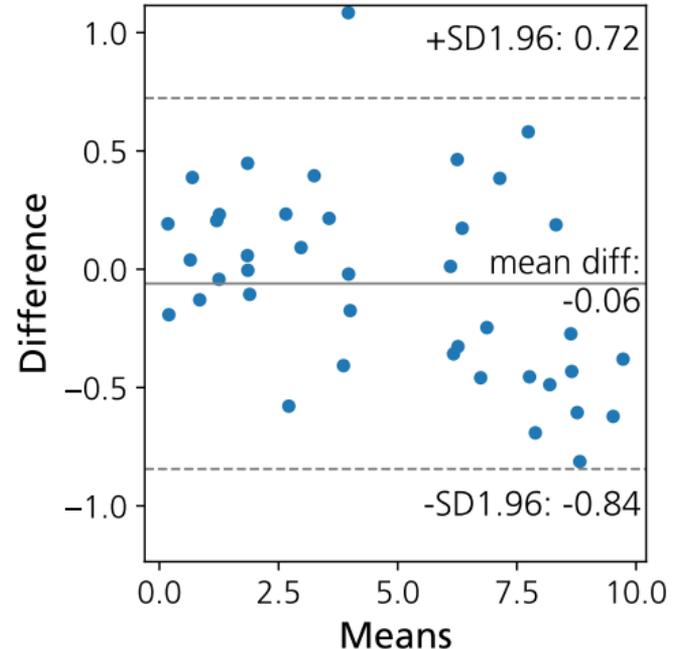
Bland–Altman Plots

- aka Tukey mean-difference plot
- purpose: comparison of two values that “should be the same”
 - e.g., when comparing two algorithmic approaches
- horizontal axis: mean of the two
- vertical axis: difference of the two: mean and ± 1.96 standard deviations (95 % limits of agreement)

3. Further Types of Plots and Visualizations

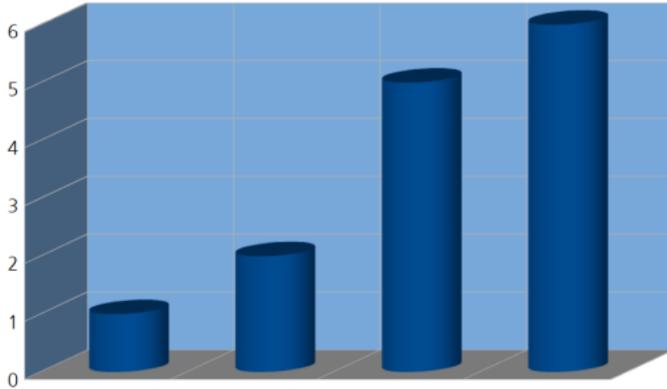
Bland–Altman Plots

- aka Tukey mean-difference plot
- purpose: comparison of two values that “should be the same”
 - e.g., when comparing two algorithmic approaches
- horizontal axis: mean of the two
- vertical axis: difference of the two: mean and ± 1.96 standard deviations (95 % limits of agreement)
- example of designing a plot for a purpose



3. Further Types of Plots and Visualizations

3D Plots: Best Avoided



- ✗ 3rd dimension: no additional information
- ✗ distracting and hard to read
- ✗ frequently used to show off or gloss over poor-quality data/results

4. Tips & Tricks

Contents

1. Introduction

2. Basic Types of Plots

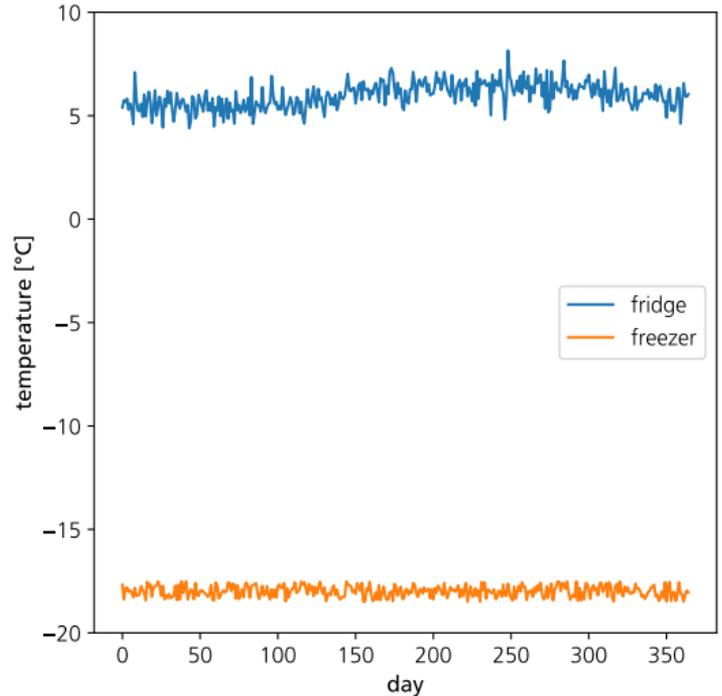
3. Further Types of Plots and Visualizations

4. Tips & Tricks

4. Tips & Tricks

Scaling: Data Ranges

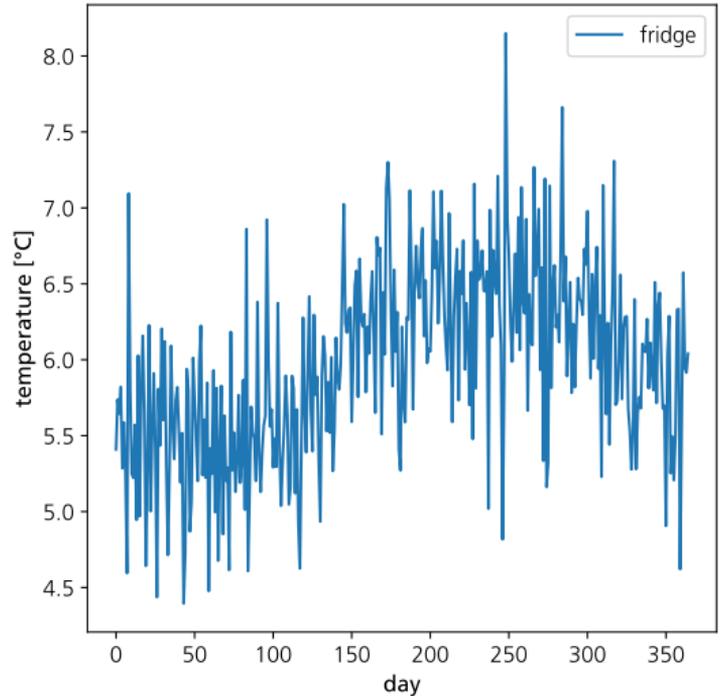
- full range shows all data, but may hide relevant variation



4. Tips & Tricks

Scaling: Data Ranges

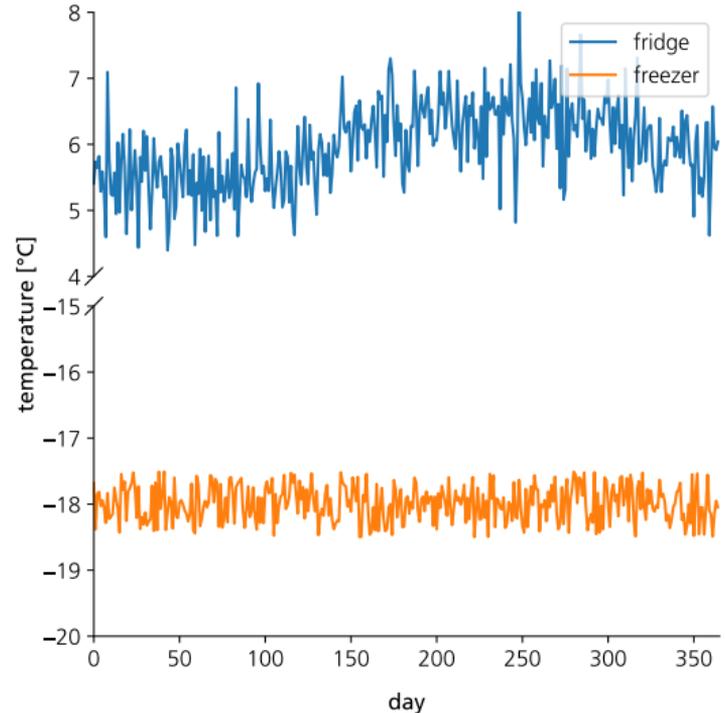
- full range shows all data, but may hide relevant variation
- limited range shows relevant variation, but may hide important outliers



4. Tips & Tricks

Scaling: Data Ranges

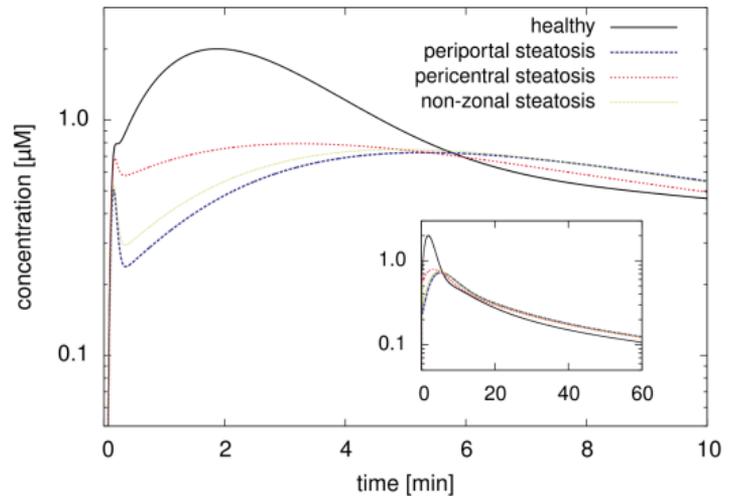
- full range shows all data, but may hide relevant variation
- limited range shows relevant variation, but may hide important outliers
- possibilities:
 - interrupted axes



4. Tips & Tricks

Scaling: Data Ranges

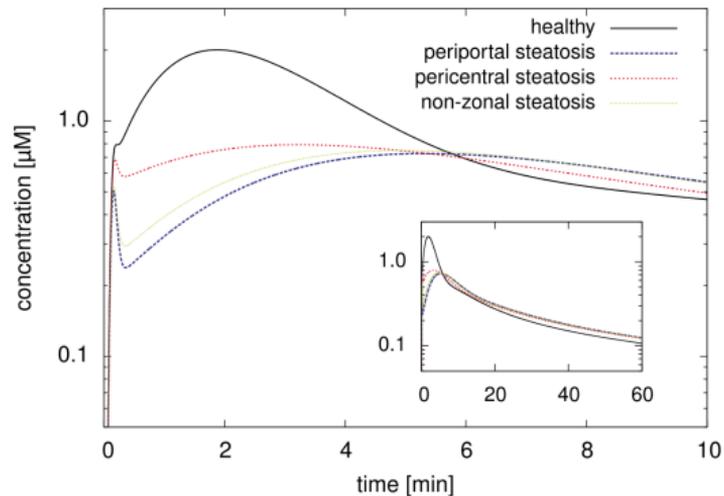
- full range shows all data, but may hide relevant variation
- limited range shows relevant variation, but may hide important outliers
- possibilities:
 - interrupted axes
 - inset plots



4. Tips & Tricks

Scaling: Data Ranges

- full range shows all data, but may hide relevant variation
- limited range shows relevant variation, but may hide important outliers
- possibilities:
 - interrupted axes
 - inset plots
 - logarithmic scaling



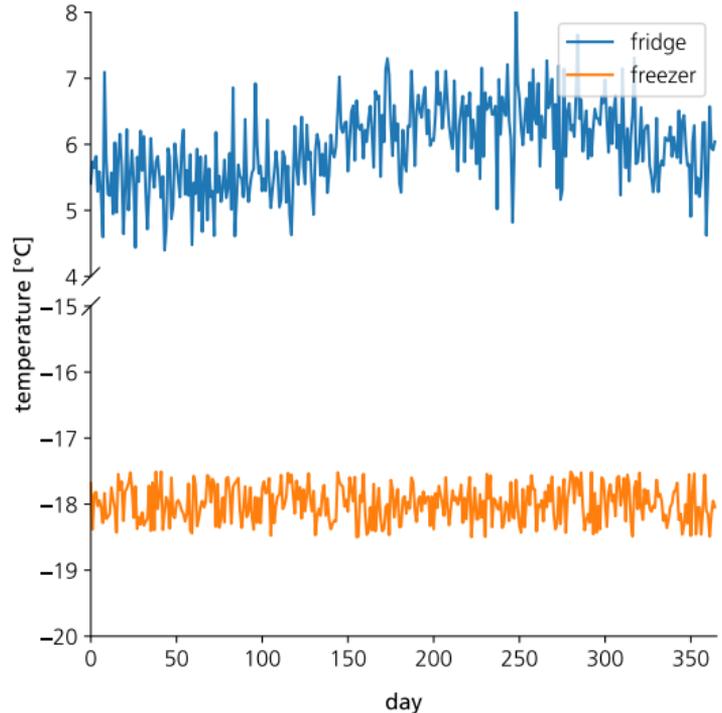
4. Tips & Tricks

Scaling: Data Ranges

- full range shows all data, but may hide relevant variation
- limited range shows relevant variation, but may hide important outliers
- possibilities:
 - interrupted axes
 - inset plots
 - logarithmic scaling

Units

- consider using intuitive rather than SI (base) units in plots
- may need to adapt default tics (e.g., 12, 24, 36, ... hours)



4. Tips & Tricks

Color

- always expect
 - a bad projector/printer
 - poor lighting
 - a color-blind audience

4. Tips & Tricks

Color

- always expect
 - a bad projector/printer
 - poor lighting
 - a color-blind audience
- primary way for distinguishing lines and symbols should be style, color can be used *in addition*

4. Tips & Tricks

Color

- always expect
 - a bad projector/printer
 - poor lighting
 - a color-blind audience
- primary way for distinguishing lines and symbols should be style, color can be used *in addition*
- in particular for articles etc. printed by other people, assume a poor-quality, black-and-white, 2-on-1 printout

4. Tips & Tricks

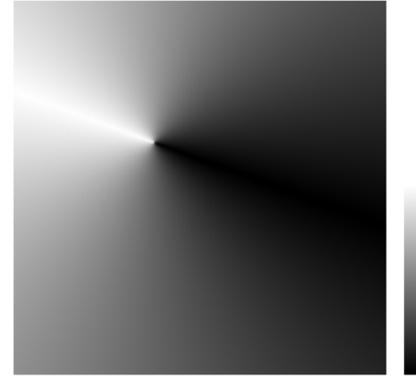
Color

- always expect
 - a bad projector/printer
 - poor lighting
 - a color-blind audience
- primary way for distinguishing lines and symbols should be style, color can be used *in addition*
- in particular for articles etc. printed by other people, assume a poor-quality, black-and-white, 2-on-1 printout (with your name still clearly legible)

4. Tips & Tricks

Color Schemes

- choose suitable color scheme



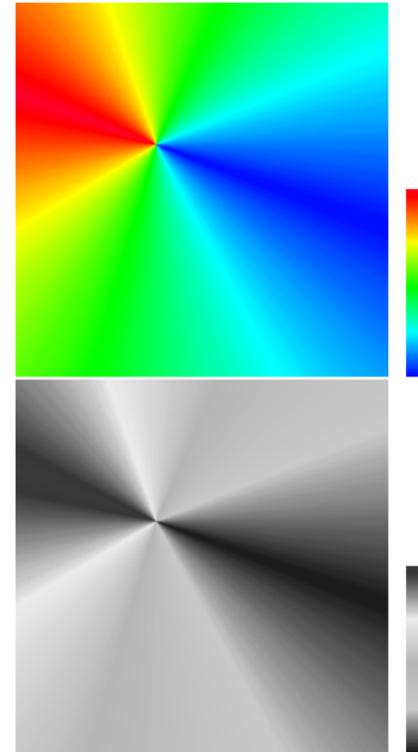
adapted from <https://doi.org/gk7m>

Raw Data

4. Tips & Tricks

Color Schemes

- choose suitable color scheme
- sequential for data ranging from low to high
 - rainbow (common, but questionable)

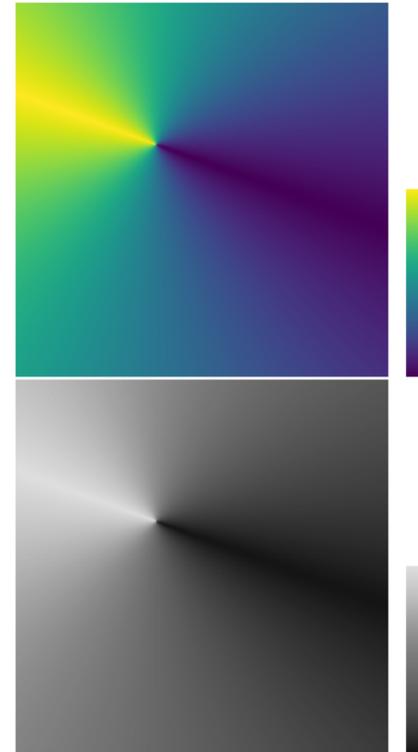


adapted from <https://doi.org/gk7m>

4. Tips & Tricks

Color Schemes

- choose suitable color scheme
- sequential for data ranging from low to high
 - rainbow (common, but questionable)
 - many other options

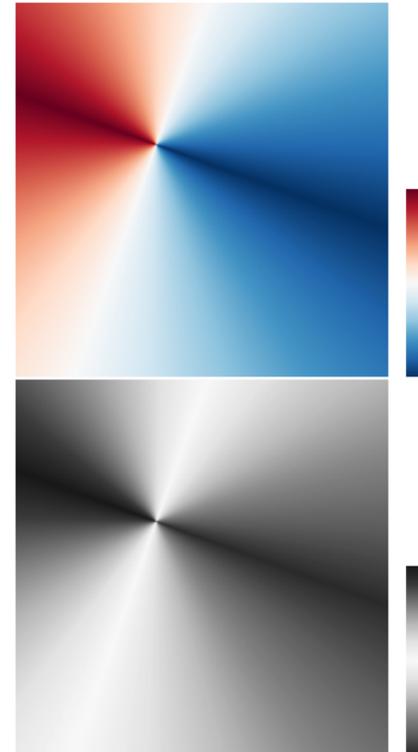


adapted from <https://doi.org/gk7m>

4. Tips & Tricks

Color Schemes

- choose suitable color scheme
- sequential for data ranging from low to high
 - rainbow (common, but questionable)
 - many other options
- divergent for data where middle range has special meaning



adapted from <https://doi.org/gk7m>

4. Tips & Tricks

Color Schemes

- choose suitable color scheme
- sequential for data ranging from low to high
 - rainbow (common, but questionable)
 - many other options
- divergent for data where middle range has special meaning
- qualitative for categorical data (should not be too many)

adapted from <https://doi.org/gk7m>

4. Tips & Tricks

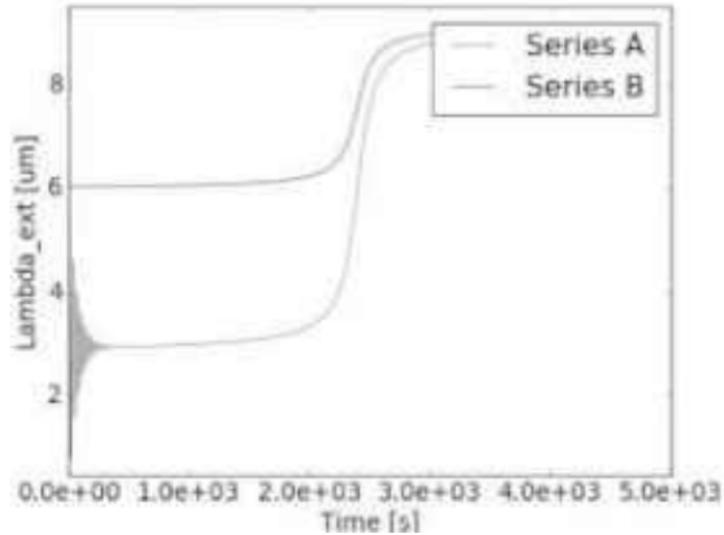
Pay Attention to Details

Why?

- Excellent results deserve an excellent presentation.
- Proper presentation proves you know what you are showing.

4. Tips & Tricks

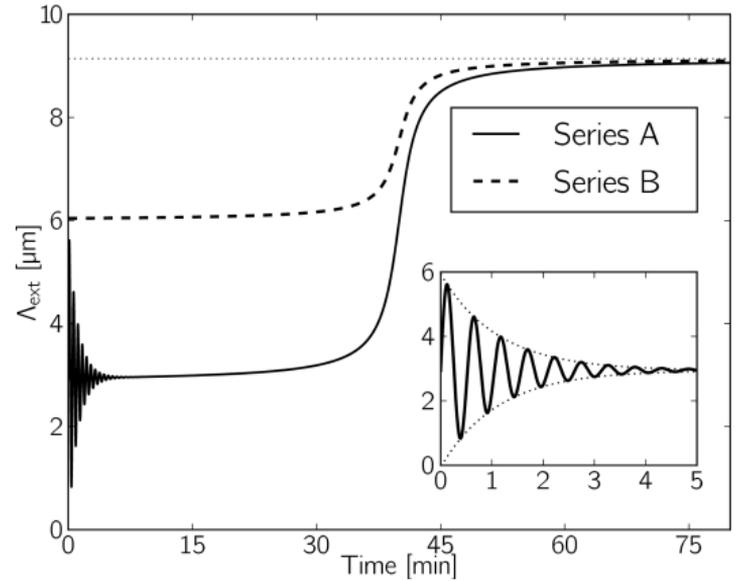
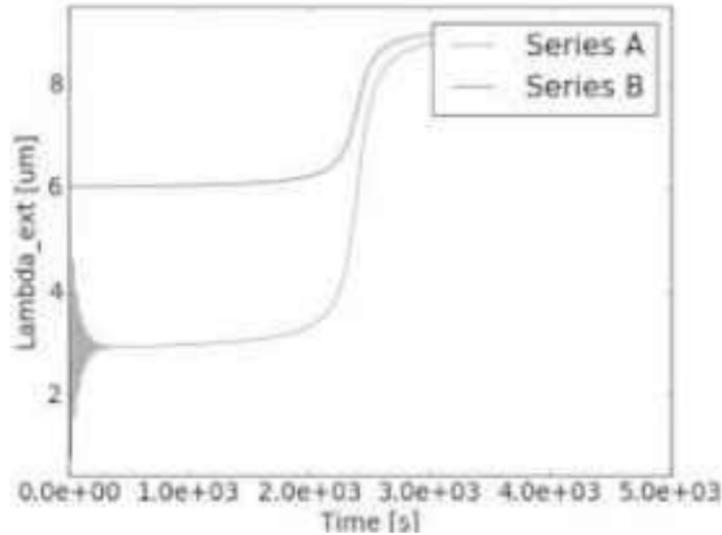
Bad vs. Good Plots



adapted from <https://doi.org/gk7n>

4. Tips & Tricks

Bad vs. Good Plots



adapted from <https://doi.org/gk7n>

4. Tips & Tricks

Labeling

- be precise and concise

4. Tips & Tricks

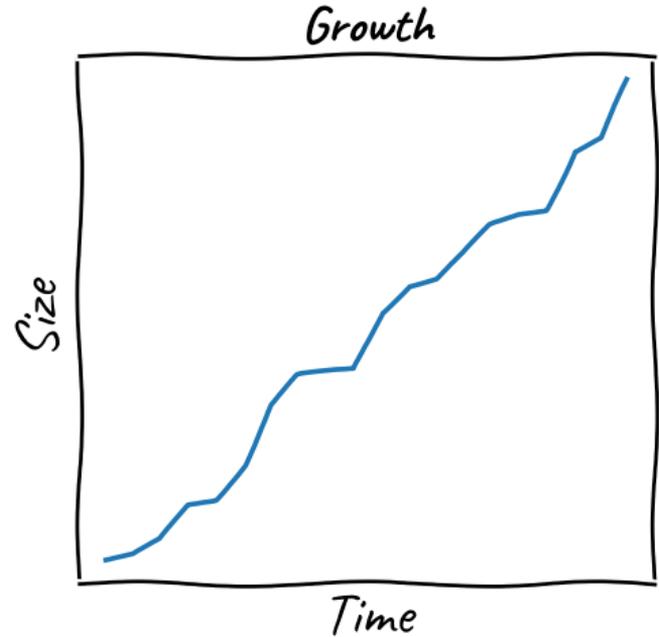
Labeling

- be precise and concise
- don't use too many tics
(or non-integer tics for integer data)

4. Tips & Tricks

Labeling

- be precise and concise
- don't use too many tics (or non-integer tics for integer data)
- use numbers and units (unless the plot is merely qualitative)



4. Tips & Tricks

Labeling

- be precise and concise
- don't use too many tics
(or non-integer tics for integer data)
- use numbers and units (unless the plot is merely qualitative)
- check the resulting plot for plausibility,
in particular when converting units

4. Tips & Tricks

Multiple Plots Side-By-Side

- use the same or sufficiently clearly distinct ranges on the horizontal axis for plots aligned vertically and vice versa

4. Tips & Tricks

Multiple Plots Side-By-Side

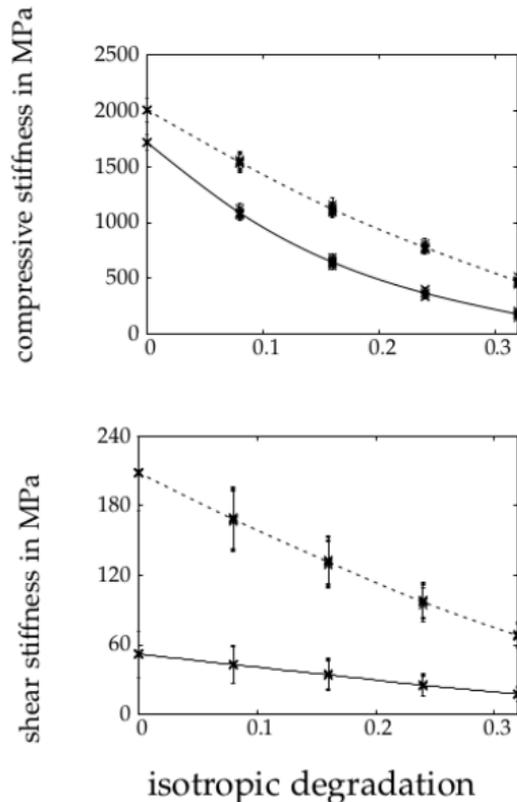
- use the same or sufficiently clearly distinct ranges on the horizontal axis for plots aligned vertically and vice versa
- use joint labels rather than redundancy

4. Tips & Tricks

Multiple Plots Side-By-Side

- use the same or sufficiently clearly distinct ranges on the horizontal axis for plots aligned vertically and vice versa
- use joint labels rather than redundancy
- make sure lines/boxes are *properly* aligned

Example from <https://doi.org/ggn3bx>, Fig. 7.23 (adapted)

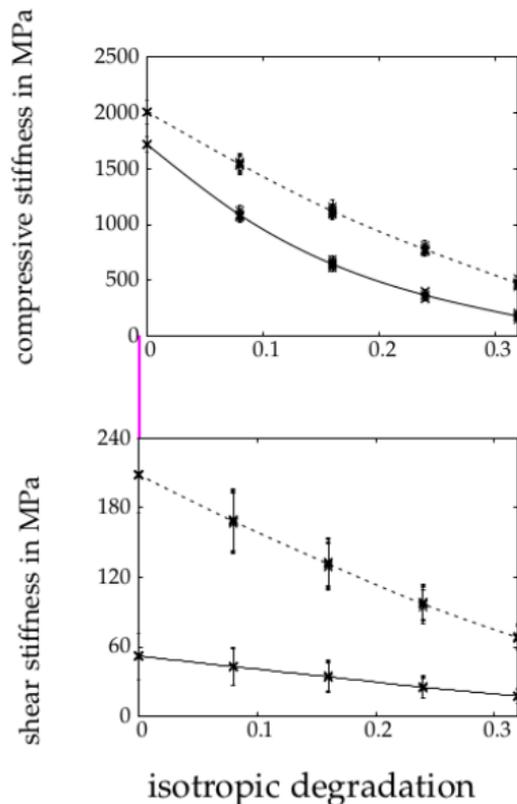


4. Tips & Tricks

Multiple Plots Side-By-Side

- use the same or sufficiently clearly distinct ranges on the horizontal axis for plots aligned vertically and vice versa
- use joint labels rather than redundancy
- make sure lines/boxes are *properly* aligned

Example from <https://doi.org/ggn3bx>, Fig. 7.23 (adapted)

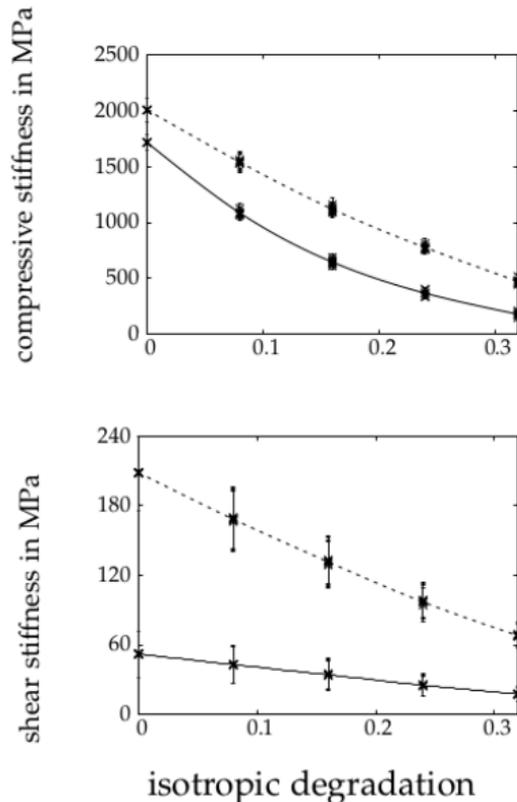


4. Tips & Tricks

Multiple Plots Side-By-Side

- use the same or sufficiently clearly distinct ranges on the horizontal axis for plots aligned vertically and vice versa
- use joint labels rather than redundancy
- make sure lines/boxes are *properly* aligned
- use consistent point and line style, size, thickness, color, ...

Example from <https://doi.org/ggn3bx>, Fig. 7.23 (adapted)



4. Tips & Tricks

But my #!\$\$%plot does not support this!?

4. Tips & Tricks

But my #!\$\$%plot does not support this!?



4. Tips & Tricks

Plotting Tools

- ✗ Spreadsheets (Excel etc.)
 - not made for scientific plots
 - no automation

4. Tips & Tricks

Plotting Tools

- ✗ Spreadsheets (Excel etc.)
 - not made for scientific plots
 - no automation
- gnuplot
 - long history
 - custom scripting language

4. Tips & Tricks

Plotting Tools

- ✗ Spreadsheets (Excel etc.)
 - not made for scientific plots
 - no automation
- gnuplot
 - long history
 - custom scripting language
- python
 - matplotlib, seaborn, plotly, etc.
 - integrates well with python-based data analysis

4. Tips & Tricks

Plotting Tools

- ✗ Spreadsheets (Excel etc.)
 - not made for scientific plots
 - no automation
- gnuplot
 - long history
 - custom scripting language
- python
 - matplotlib, seaborn, plotly, etc.
 - integrates well with python-based data analysis
- pgfplots
 - integrates well with \LaTeX document preparation

4. Tips & Tricks

Plotting Tools

- ✗ Spreadsheets (Excel etc.)
 - not made for scientific plots
 - no automation
- gnuplot
 - long history
 - custom scripting language
- python
 - matplotlib, seaborn, plotly, etc.
 - integrates well with python-based data analysis
- pgfplots
 - integrates well with \LaTeX document preparation
- R
 - integrates well with R-based data analysis

4. Tips & Tricks

Some literature:

- Fundamentals of Data Visualization, *O'Reilly* 2019, <https://clauswilke.com/dataviz/>
- ColorBrewer: Color Advice for Maps, <https://colorbrewer2.org>
 - categorical color maps and their properties
- Data visualization: A view of every Points of View column
<http://blogs.nature.com/methagora/2013/07/data-visualization-points-of-view.html>
 - links to a few *Nature Methods* articles on plotting and more general figure design

Contact: Ole Schwen <ole.schwen@mevis.fraunhofer.de>