

DATA SCIENCE CENTER



University
of Bremen



DataNord

Data Visualization and Communication Using Python and Jupyter

CONTACT

DSC Data Scientists

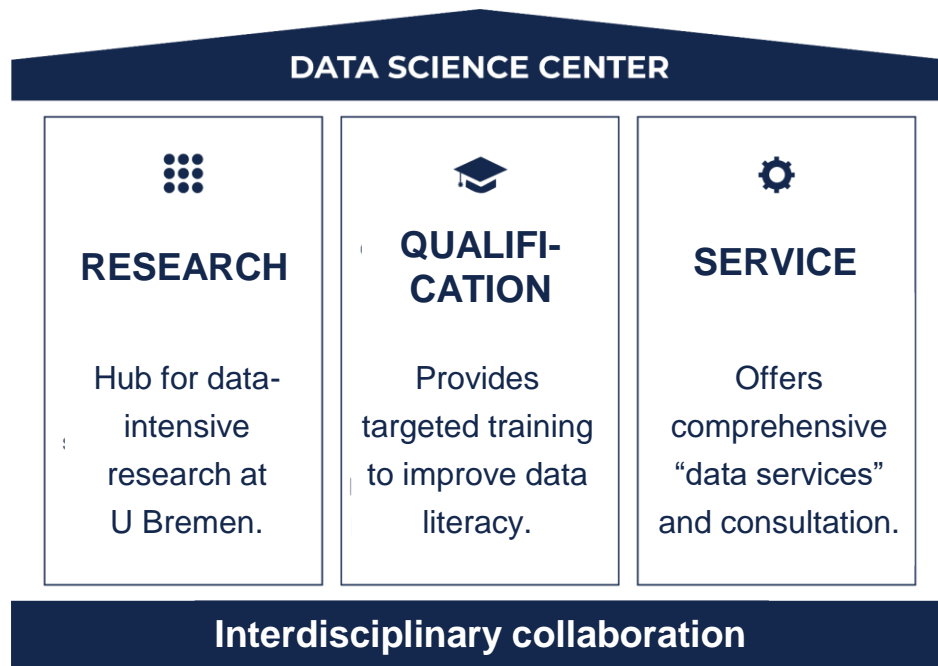
anolte@uni-bremen.de

NEWSLETTER

Sign Up!



DataNord Data Scientists for Training & Consultation



M.Sc. SARAH BÜKER
(Marine Environmental Sciences)



M.Sc. ANNIKA NOLTE
(Environmental Sciences)



Dr. MARYAM MOVAHEDIFAR
(Statistics)



Dr. SUSANNE DE VOGEL (Sociology)



M.A. NELE FUCHS
(Transcultural Studies, Philosophy)



Dipl.-Soz. HEIKE THÖRICHT (Sociology)

1. Motivation and elements behind effective data visualizations
2. “How to Python” in this workshop
3. Quickstart: Python programming (~30 min)
4. Basic plotting workflows and customization options
5. Plot types & plotting techniques
6. Mapping geospatial data
7. Animations
8. Interactive visualizations
9. Wrap-up

← Coffee Break (~15 min)

← Lunch Break (~1h)

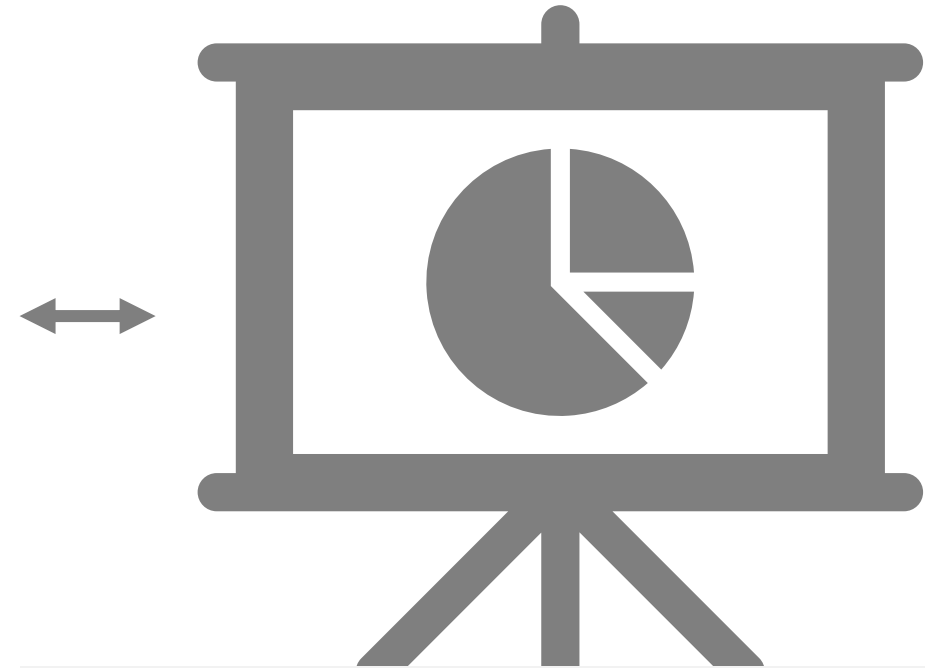
← Coffee Break (~15 min)



- ✓ Understand the role of data visualization in **storytelling**
- ✓ Recall the **key principles and best practices** that contribute to impactful data visualizations
- ✓ Acquire **skills for advanced data visualization** tailored to various data types and analytical objectives in the Earth Sciences using essential Python libraries
- ✓ Learn how to bring your data visualization **ideas to life**
- ✓ Understand the **interplay between data, plotting methods, and the messages** conveyed by visuals
- ✓ Promote **reproducibility and shareability** in data analysis
- ✓ (Getting started with Python via data visualization tasks – personal experience: Learning Python with visuals is fun!)
- × In-depth theoretical analysis of specific data types for visualization, such as time series, geospatial, or image data.

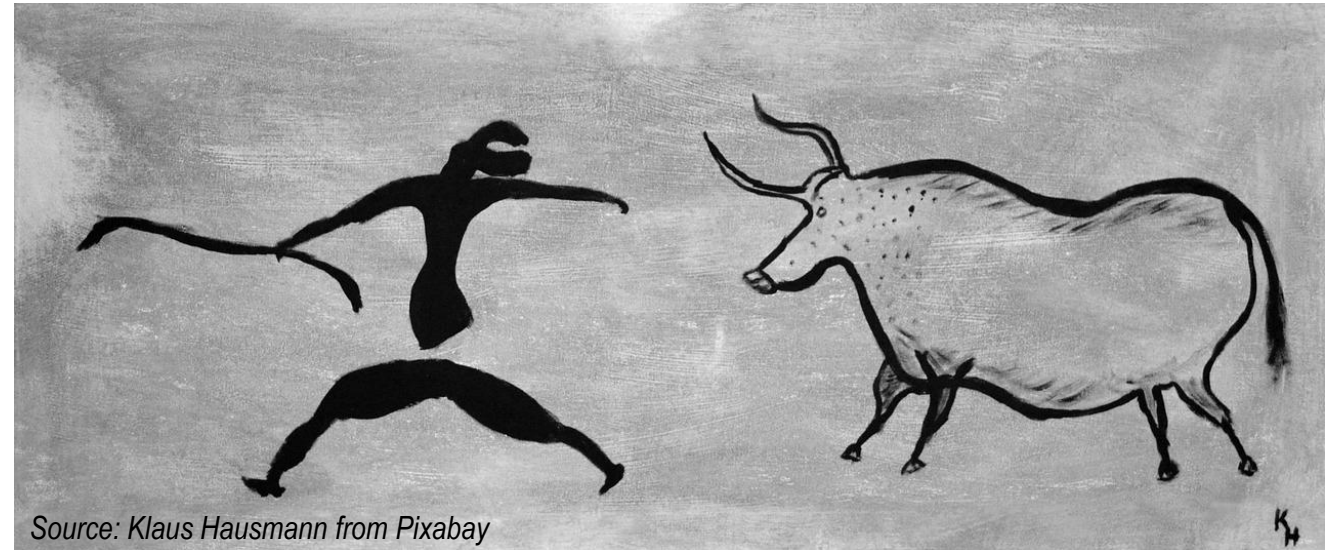


Overview of the Data Lifecycle. Adapted from RDMkit



Data visualization is the representation and presentation of information and data to make it more accessible, understandable, and actionable.

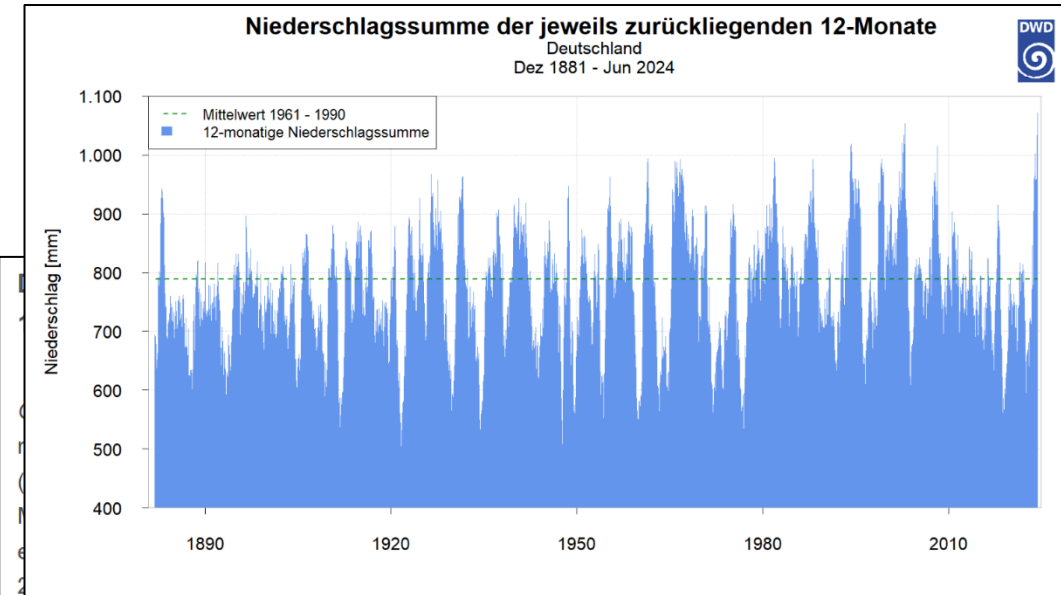
- ... is the most natural form of **education**
- ... help us **to make sense of things**, e.g. developments in time; human connections
- ... has changed with the **inventions** over the course of history:
 - printed stories in newspapers, magazines, and books
 - motion picture camera made movies possible
 - TV and streaming
 - data visualization tools



Data storytelling is communicating the meaning of a dataset with visuals and a story that are customized for each particular audience.

Motivation for Data Visualization

- Quickly and clearly **understand** scientific data:
 - reducing confusion, mistakes and time-effort
 - uncover trends, patterns, relationships, and outliers.



Litern pro Quadratmeter. Im Vergleich dazu beträgt der vieljährige Mittelwert der Referenzperiode 1961-1990 rund 789 Liter pro Quadratmeter im Jahr.

Ein Blick ins nationale Klimaarchiv des DWD zeigt: Während der vergangenen zehn Jahre wurden überwiegend eher unterdurchschnittliche jährliche Niederschlagssummen beobachtet. (Abb. 2) Betrachtet man die Zeitreihe seit Messbeginn 1881 zeigen die Daten aber eine leichte Zunahme der jährlichen Niederschlagsmengen in Deutschland. Allerdings kommt es dabei laut DWD immer wieder zu einem Wechsel zwischen trockenen und feuchten Perioden. Kaspar: "Der Niederschlag zeichnet sich durch eine hohe Variabilität sowohl von Jahr zu Jahr als auch über längere Zeiträume hinweg aus." (Abb. 3) Die Trockenheit der vergangenen Jahre wurde nun durch eine sehr feuchte 12-monatige Phase abgelöst, die das Niederschlagsdefizit mit jedem Monat weiter reduziert hat.

Source: https://www.dwd.de/DE/presse/pressemitteilungen/DE/2024/20240703_die-zwoelf-nassesten-monaten-seit-messbeginn_news.html?nn=16210

Motivation for Data Visualization

- Quickly and clearly **understand** scientific data:
 - reducing confusion, mistakes and time-effort
 - uncover trends, patterns, relationships, and outliers.
- Boosted **collaboration** by creating a shared narrative that fosters understanding and effective teamwork, especially in interdisciplinary teams – simplification of complex data makes it easier to work together.



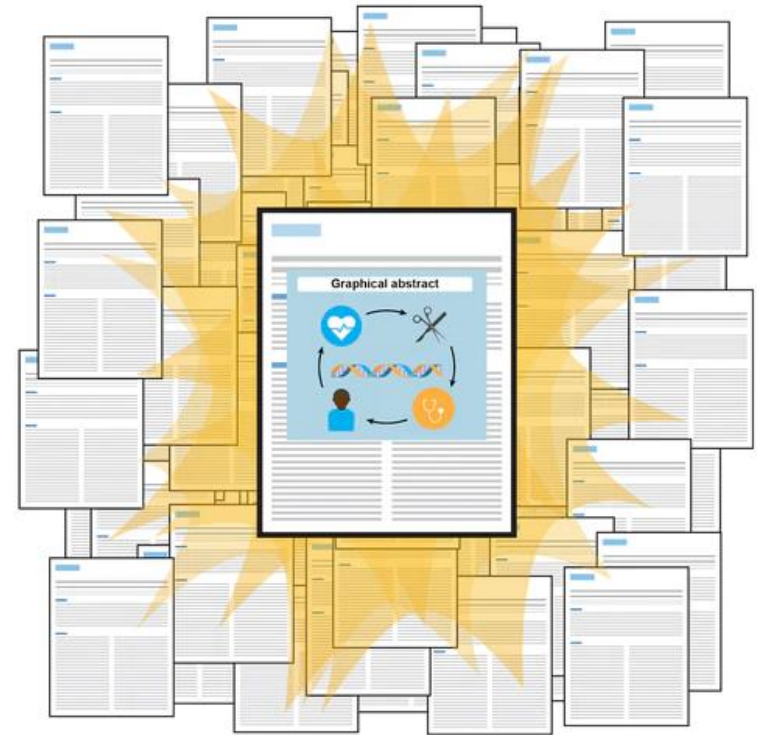
Source: rawpixel.com from Freepik

Motivation for Data Visualization

- Quickly and clearly **understand** scientific data:
 - reducing confusion, mistakes and time-effort
 - uncover trends, patterns, relationships, and outliers.
- Boosted **collaboration** by creating a shared narrative that fosters understanding and effective teamwork, especially in interdisciplinary teams.
- Improve grant proposals, streamline the peer review process (e.g., through visual abstracts), elevate presentations, and improve social media **outreach**.

A visual abstract is
shared 8 times more
often on social media

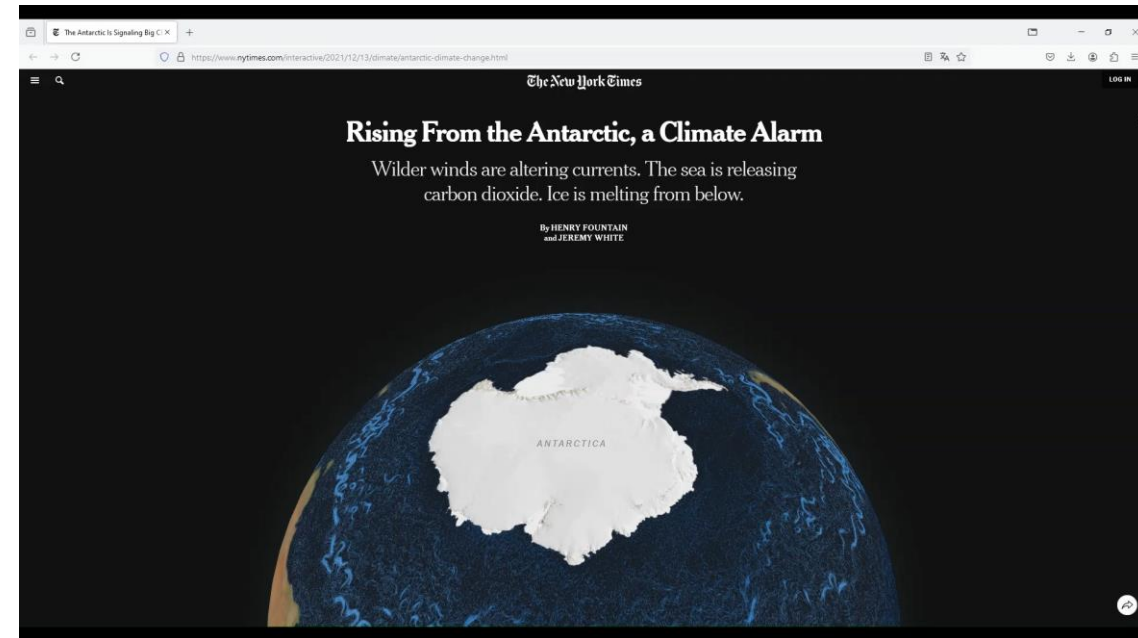
People remember visuals
6 times better than text



Source: <https://solutions.springernature.com/blogs/visibility/what-is-a-graphical-abstract-and-why-do-i-need-one-for-my-paper>

Motivation for Data Visualization

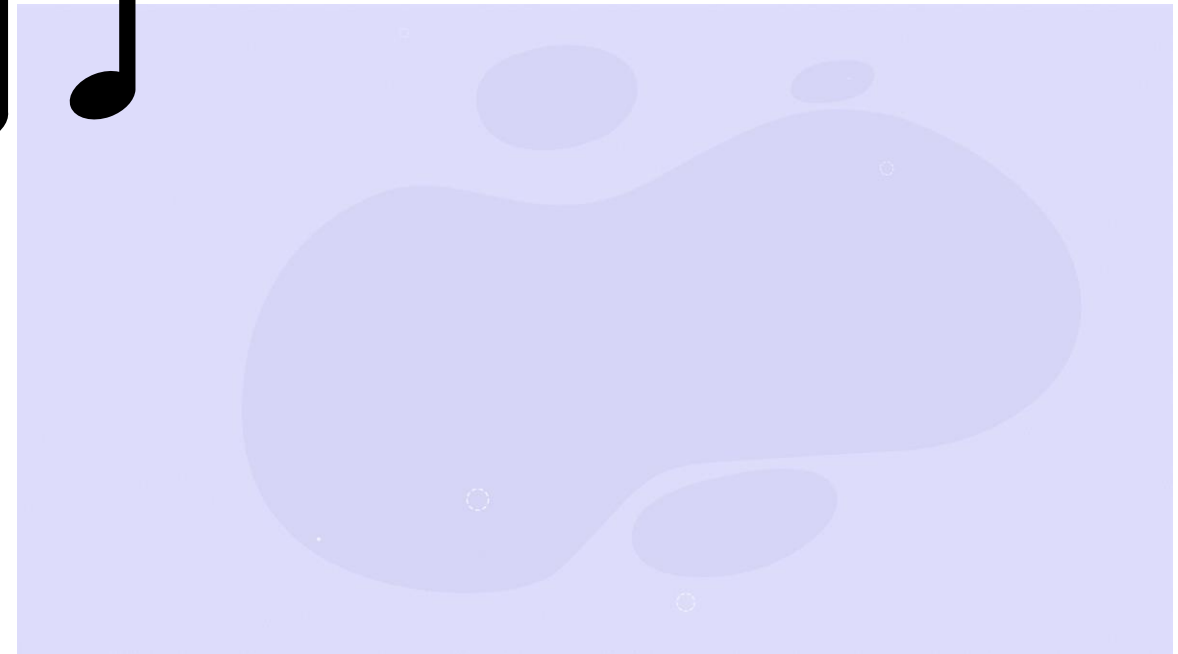
- Quickly and clearly **understand** scientific data:
 - reducing confusion, mistakes and time-effort
 - uncover trends, patterns, relationships, and outliers.
- Boosted **collaboration** by creating a shared narrative that fosters understanding and effective teamwork, especially in interdisciplinary teams.
- Improve grant proposals, streamline the peer review process (e.g., through visual abstracts), elevate presentations, and improve social media **outreach**.
- Making scientific information **more engaging and accessible** for everyone, including policymakers, students, and the general public.



Source: <https://www.nytimes.com/interactive/2021/12/13/climate/antarctic-climate-change.html>

Data Storytelling in the Music Streaming Industry

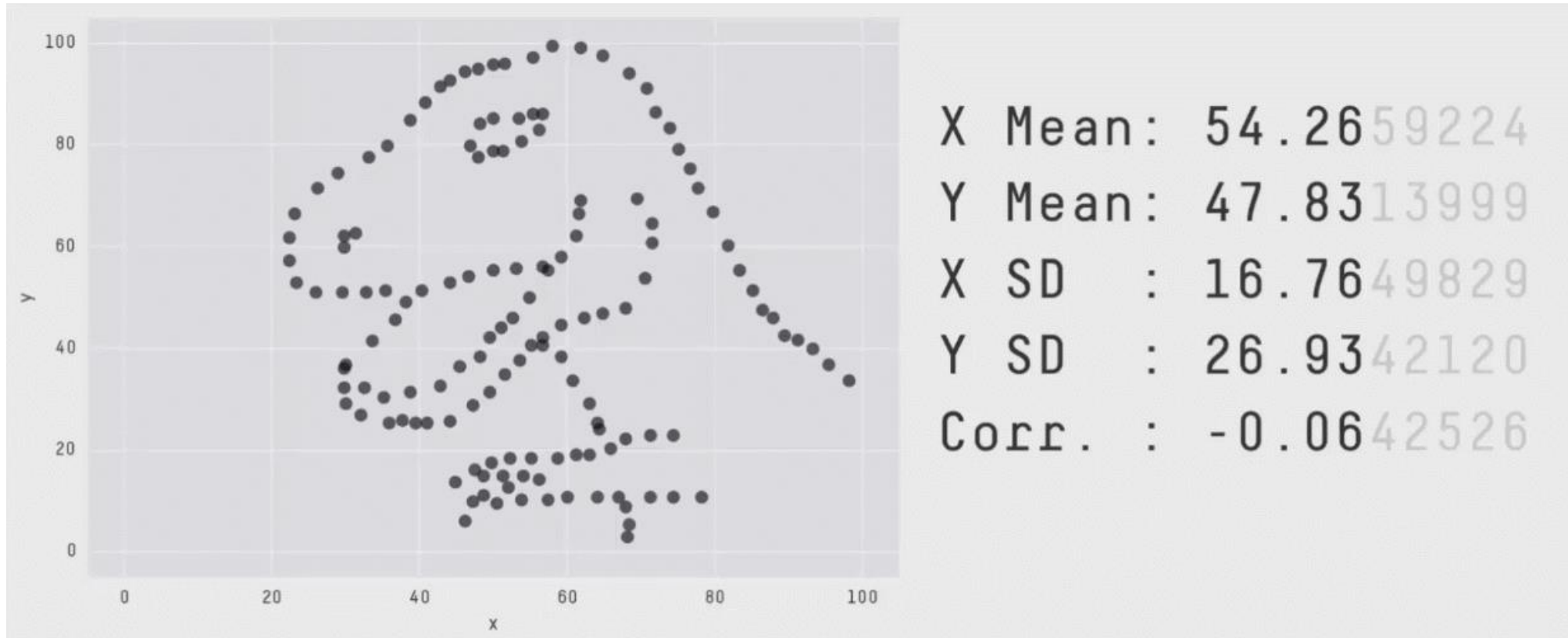
- Because data storytelling is so efficient, industry makes use of it a lot
- Example from the music streaming industry:
Annual reviews of the customers listening behavior with animated visuals
 - Identifies and highlights: Songs, artists, genres, etc.
 - Congratulations for being a top listener
 - How much time spend enjoying the service



Source: Freepik (www.freepik.com)

Anscombe's Quartet: Same statistics – different visual patterns

Never trust summary statistics alone. Always visualize your data to tell the whole story.

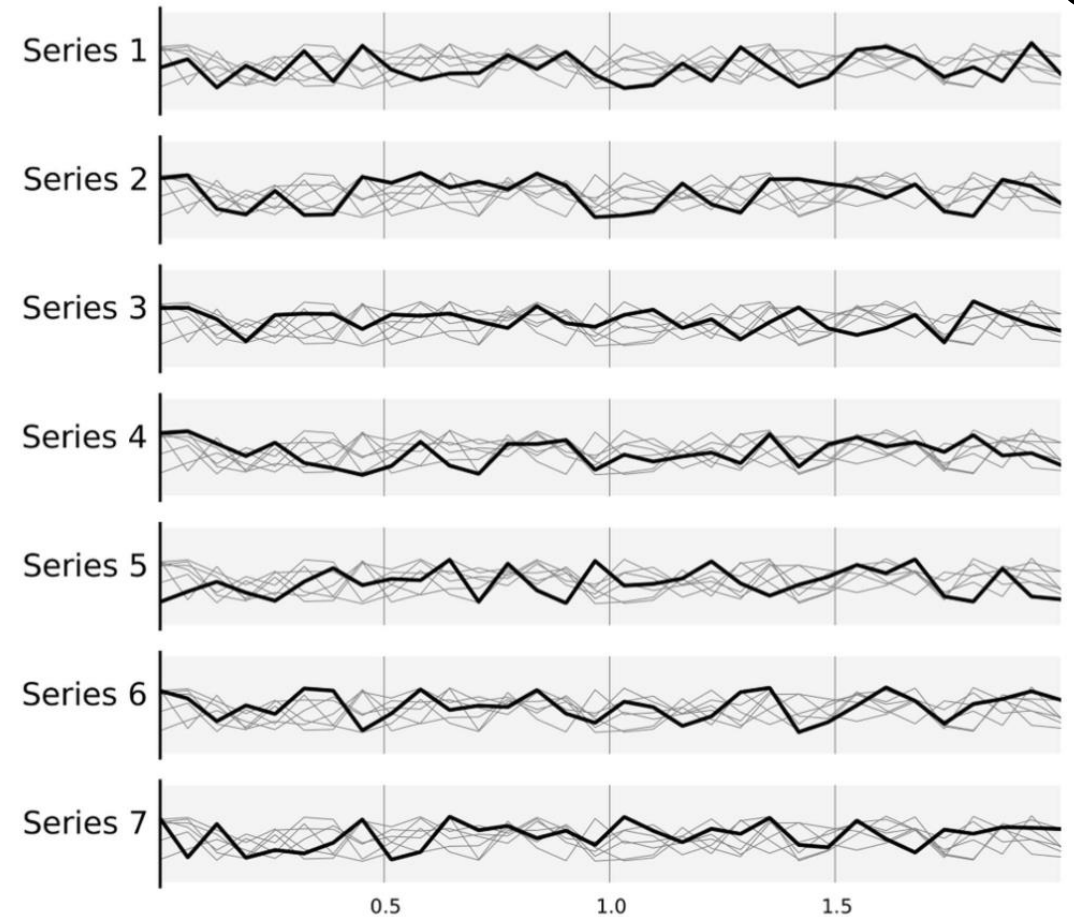
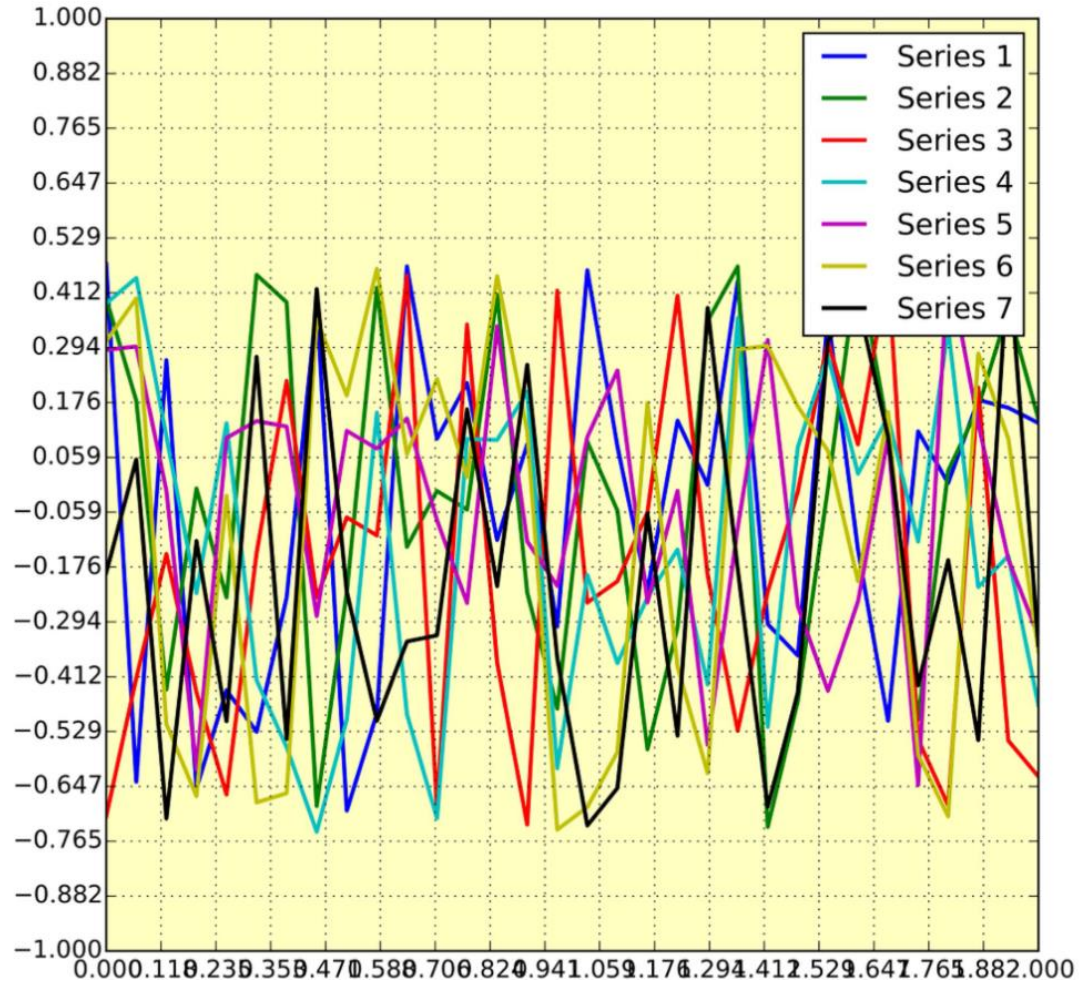


Sources: <https://www.research.autodesk.com/publications/same-stats-different-graphs/>; Albert Cairo
(<http://www.thefunctionalart.com/2016/08/download-datasaurus-never-trust-summary.html>)

Some Worst and Best Practices



“Avoid chartjunk”

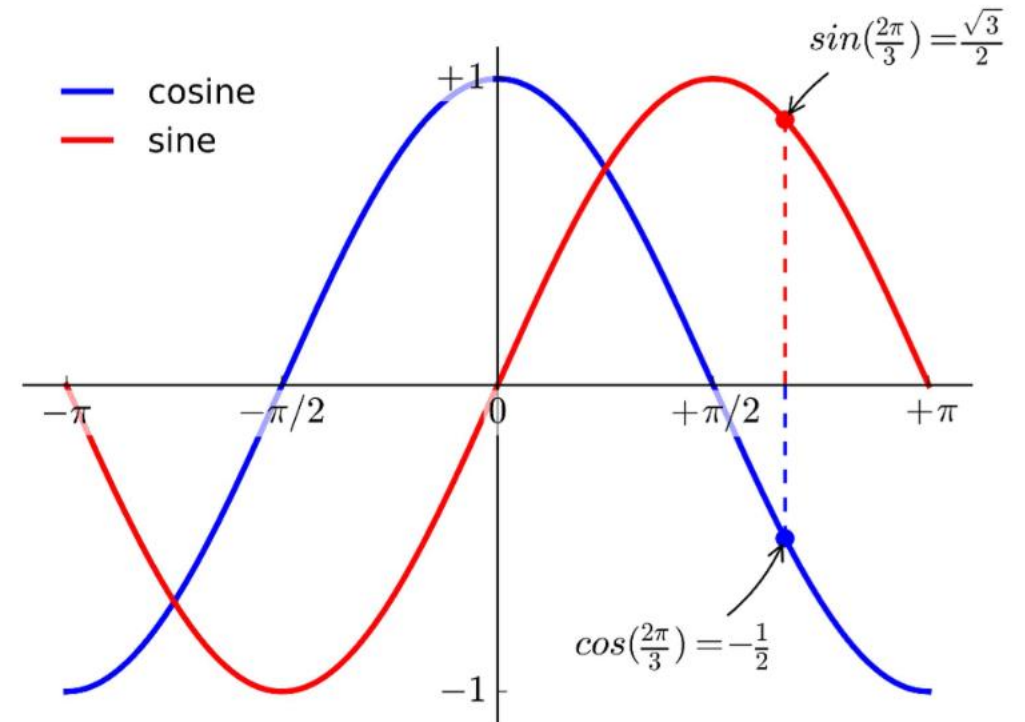
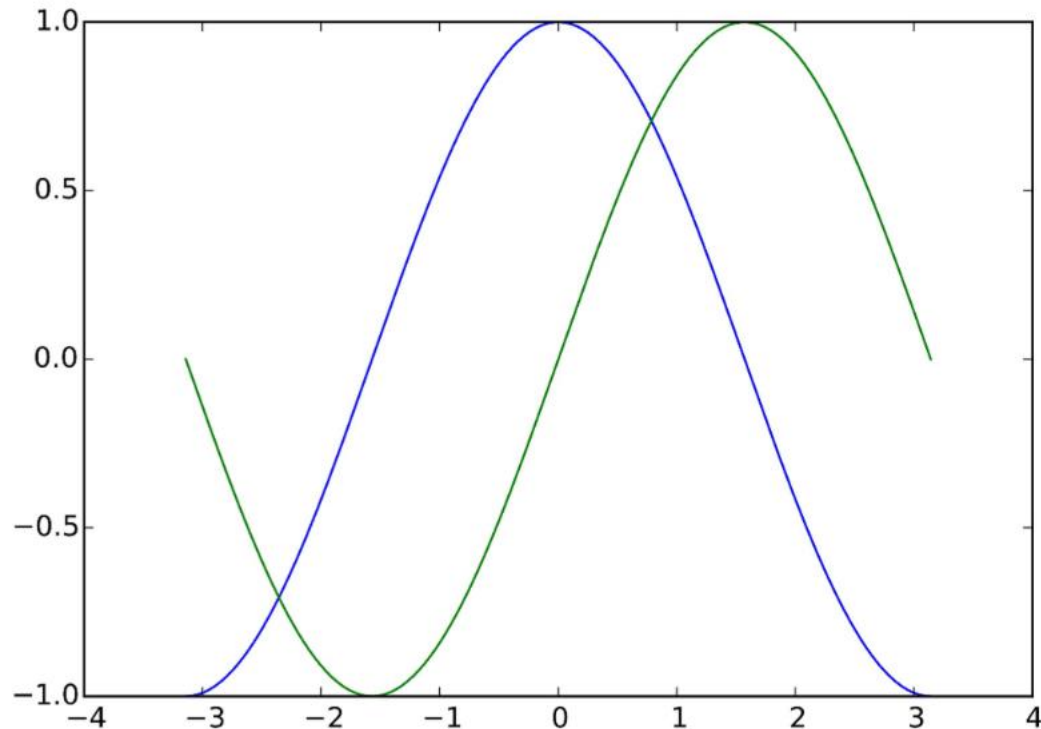


Source: Rougier, Nicolas P.; Droettboom, Michael; Bourne, Philip E. (2014): Ten simple rules for better figures: Public Library of Science (10) (9).

Some Worst and Best Practices



“Do go beyond the defaults”

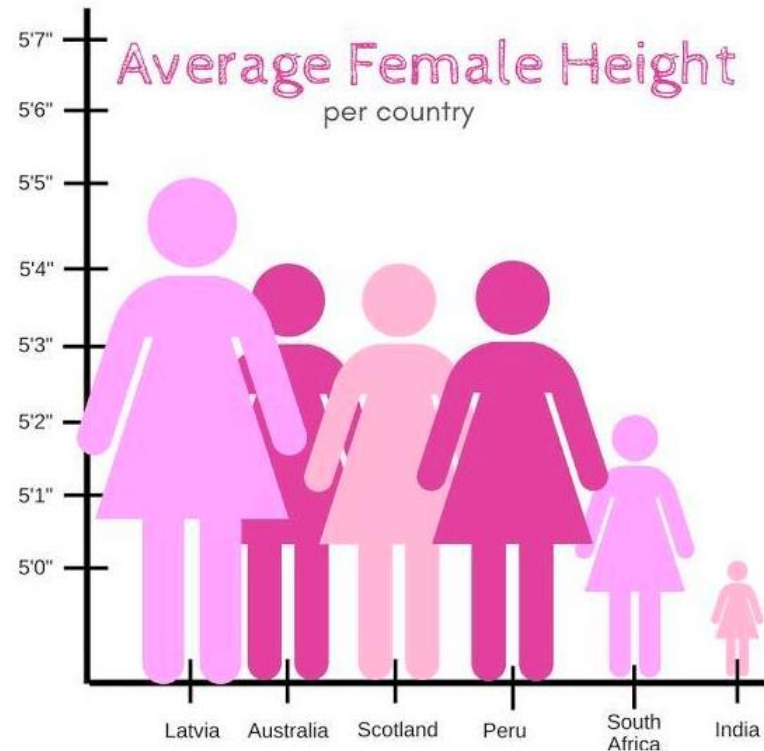


Source: Rougier, Nicolas P.; Droettboom, Michael; Bourne, Philip E. (2014): Ten simple rules for better figures: Public Library of Science (10) (9).

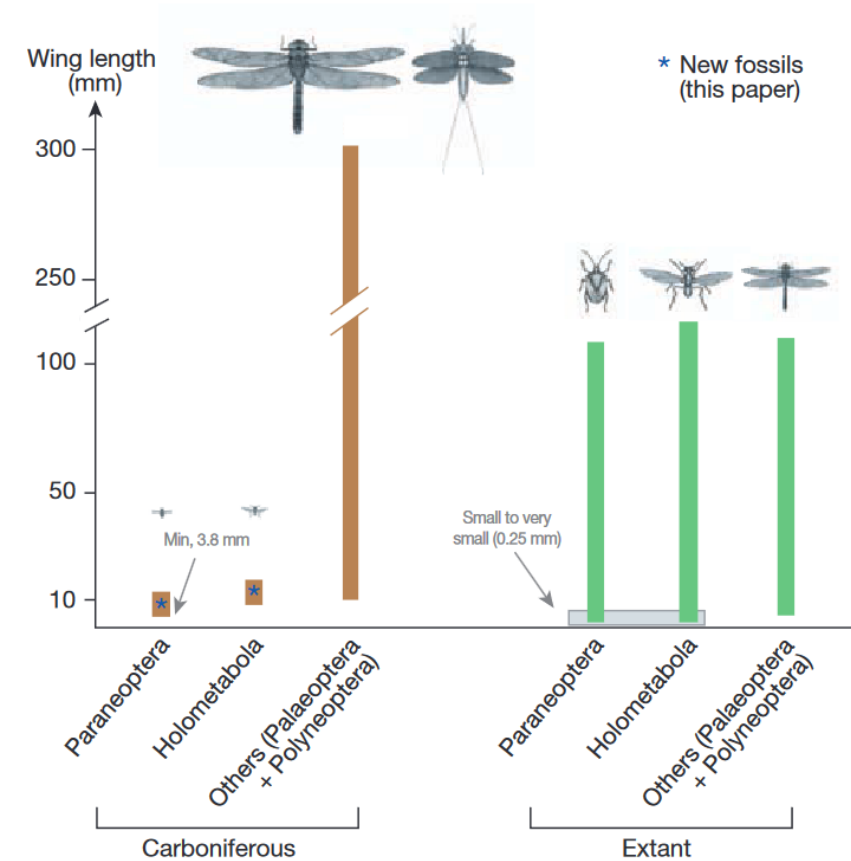
Some Worst and Best Practices



“Stick to universal rules”



Source: <https://ninjatables.com/bad-data-visualization-examples/>

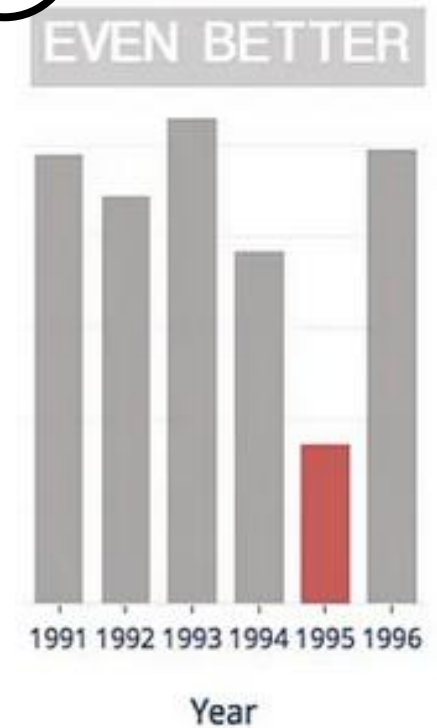
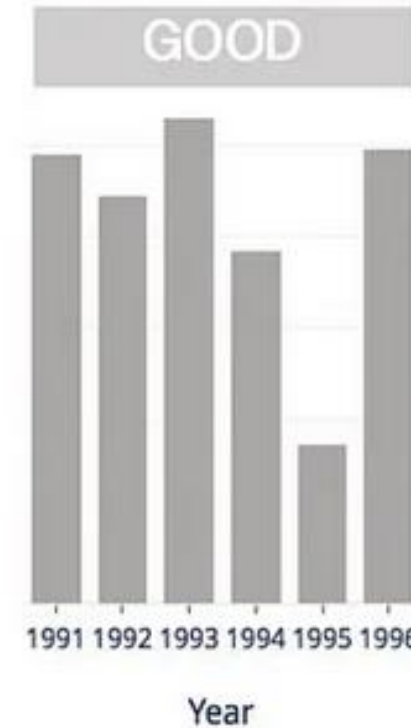
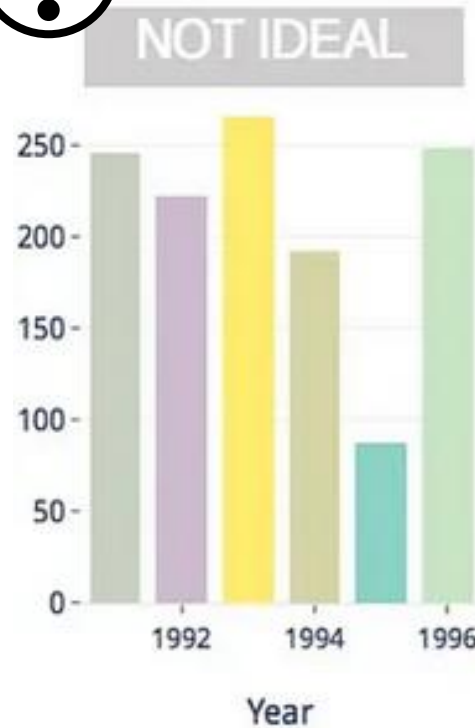


Source: Nel, André; Roques, Patrick; Nel, Patricia; Prokin, Alexander A.; Bourgoin, Thierry; Prokop, Jakub et al. (2013): The earliest known holometabolous insects. In *Nature* 503 (7475), pp. 257–261.

Some Worst and Best Practices

“Use colors wisely”

- Color where it makes sense
- Sensitive to cultural contexts
- Consistency across multiple visualizations
- Color as part of the storytelling
- Differentiability and accessibility (e.g. to people with color vision deficiencies)



Color Blindness Simulator

<https://www.color-blindness.com/coblis-color-blindness-simulator/>

feature" on Edge and Internet Explorer. All others should support everything just fine.


So go ahead, choose an image through the upload functionality or just drag and drop your image in the center of our Color **BL**indness **S**imulator. It is also possible to zoom and move your images around using your mouse – try it out, I hope you like it.

Drag and drop or paste your file in the area below or: Keine Datei ausgewählt.

Trichromatic view:	Anomalous Trichromacy:	Dichromatic view:	Monochromatic view:
<input checked="" type="radio"/> Normal	<input type="radio"/> Red-Weak/Protanomaly	<input type="radio"/> Red-Blind/Protanopia	<input type="radio"/> Monochromacy/Achromatopsia
	<input type="radio"/> Green-Weak/Deuteranomaly	<input type="radio"/> Green-Blind/Deuteranopia	<input type="radio"/> Blue Cone Monochromacy
	<input type="radio"/> Blue-Weak/Tritanomaly	<input type="radio"/> Blue-Blind/Tritanopia	

Use lens to compare with normal view: ☒ No Lens ☐ Normal Lens ☐ Inverse Lens

[Reset View](#)

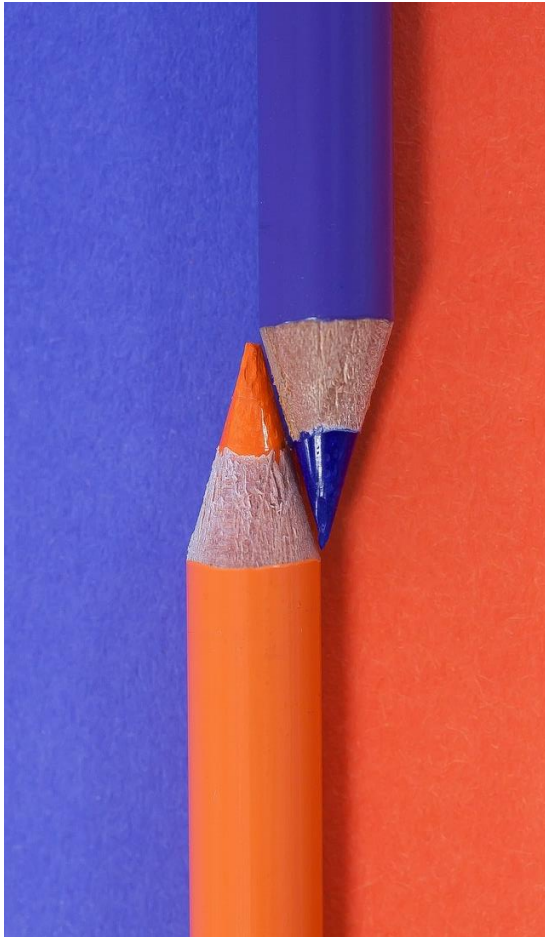


Zoom, move and lens functionality only with your own images available.

<https://www.color-blindness.com/coblis-color-blindness-simulator/>

The role of color in data visualizations

Differentiation



Source: stux from Pixabay

Emphasis



Source: schuetz-mediendesign from Pixabay

Aesthetics



Source: Alexas_Fotos from Pixabay

Emotions



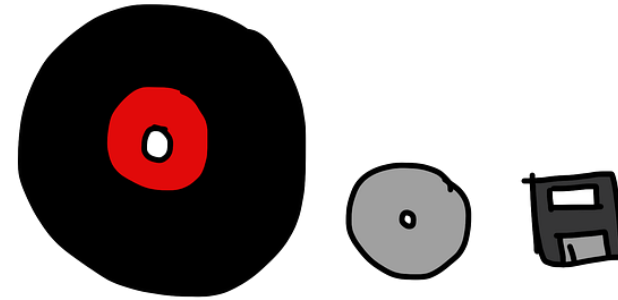
Source: Adam Filipowicz from Shutterstock

Summary: What Makes a Good Data Visualization?

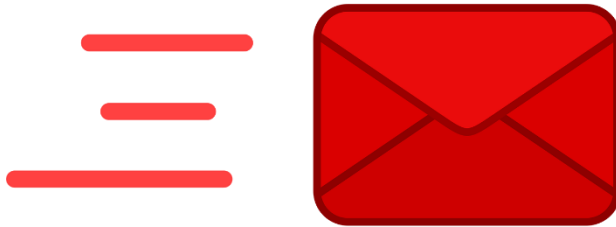
Best Practice Layout



Medium Adapted



Message Focused



Audience-Centric



The Python Universe for Data Visualization



Fundamental Plotting Library



Fundamental Libraries

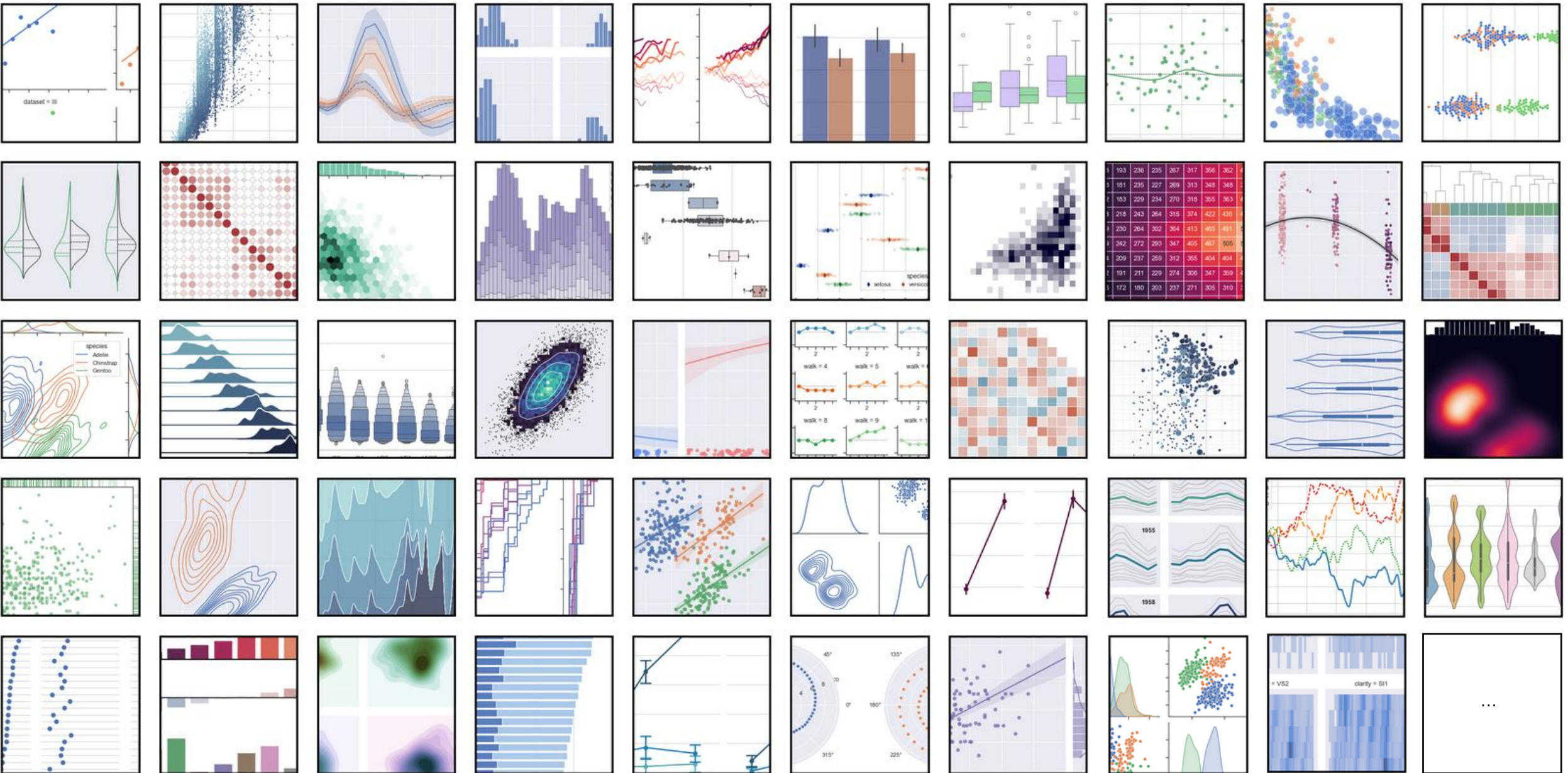


seaborn

...

Choosing Types of Visualizations

Source: <https://seaborn.pydata.org/examples/index.html>



Python Packages Cheat Sheets

matplotlib

Cheat sheet Version 3.7.4

Quick start

```
import numpy as np
import matplotlib as mpl
import matplotlib.pyplot as plt

X = np.linspace(0, 2*np.pi, 100)
Y = np.cos(X)

fig, ax = plt.subplots()
ax.plot(X, Y, color='green')

fig.savefig("figure.pdf")
plt.show()
```

Anatomy of a figure

Basic plots

```
plot(X, Y, [fmt], ...)
scatter(X, Y, ...)
bar(h)(x, height, ...)
imshow(Z, ...)
contour[f](X, Y, Z, ...)
pcolormesh(X, Y, Z, ...)
quiver(X, Y, U, V, ...)
pie(X, ...)
text(x, y, text, ...)
fill_between(X, Y1, Y2, color, where)
```

Advanced plots

```
step(X, Y, [fmt], ...)
boxplot(X, ...)
errorbar(X, Y, xerr, yerr, ...)
hist(X, bins, ...)
violinplot(D, ...)
barbs(X, Y, U, V, ...)
eventplot(positions, ...)
hexbin(X, Y, C, ...)
```

Scales

```
ax.set_yscale(scale, ...)
linear
symlog
log
logit
```

Projections

```
subplot(..., projection=p)
p='polar'
p='3d'
p=ccrs.Orthographic()
import cartopy.crs as ccrs
```

Lines

```
linestyle or ls
capstyle or dash_capstyle
markers
```

Colors

```
plt.get_cmap(name)
Uniform
Sequential
Diverging
Qualitative
Cyclic
```

Colormaps

```
plt.get_cmap(name)
Uniform
Sequential
Diverging
Qualitative
Cyclic
```

Tick locators

```
from matplotlib.ticker import
ax.xaxis.set_major_locator(locator)
NullLocator()
MultipleLocator(8.5)
FixedLocator([0, 1, 5])
LinearLocator(numticks=3)
IndexLocator(bases=[5, 10, 100])
AutoLocator()
MaxNLocator(nbins=10)
LogLocator()
ScalarFormatter()
StrMethodFormatter()
PercentFormatter()
```

Tick form

```
from matplotlib.ticker import
ax.xaxis.set_major_formatter(formatter)
NullFormatter()
FixedFormatter(1000)
FuncFormatter(lambda x, pos: '%1.1f' % x)
FormatStrFormatter("%d %b %Y")
ScalarFormatter()
StrMethodFormatter("%s")
PercentFormatter()
```

Ornament

```
ax.legend(handles, labels)
```

Legend

```
ax.legend(handles, labels)
```

Keyboard shortcuts

Ctrl+S	Save	Ctrl+W	Close plot
Ctrl+R	Reset view	F	Fullscreen 0/1
F	View forward	B	View back
P	Pan view	O	Zoom to rect
X	X pan/zoom	Y	Y pan/zoom
G	Minor grid 0/1	M	Major grid 0/1
L	X axis log/linear	L	Y axis log/linear

Ten simple rules

1. Know your audience
2. Identify your message
3. Adapt the figure
4. Captions are not optional
5. Do not trust the defaults
6. Use color effectively
7. Do not mislead the reader
8. Avoid "chartjunk"
9. Message trumps beauty
10. Get the right tool

Getting help

- matplotlib.org
- github.com/matplotlib/matplotlib/issues
- discourse.matplotlib.org
- stackoverflow.com/questions/tagged/matplotlib
- https://gitter.im/matplotlib/matplotlib
- twitter.com/matplotlib
- Matplotlib users mailing list

matplotlib

Cheat sheet Version 3.7.4

Quick start

```
import numpy as np
import matplotlib as mpl
import matplotlib.pyplot as plt
```

X = np.linspace(0, 2*np.pi, 100)
Y = np.cos(X)

fig, ax = plt.subplots()
ax.plot(X, Y, color='green')

fig.savefig("Figure.pdf")
plt.show()

Anatomy of a figure

Figure Anatomy

Subplots layout

```
subplot[s](rows, cols, ...)
fig, axs = plt.subplots(3, 3)
```

Subplots

Getting help

- matplotlib.org
- github.com/matplotlib/matplotlib/issues
- discourse.matplotlib.org
- stackoverflow.com/questions/tagged/matplotlib
- https://gitter.im/matplotlib/matplotlib
- twitter.com/matplotlib
- Matplotlib users mailing list

Basic plots

```
plot([X], Y, [fmt], ...)
X, Y, fmt, color, marker, linestyle
```

```
scatter(X, Y, ...)
X, Y, [s]izes, [c]olors, marker, cmap
```

```
bar[h](x, height, ...)
x, height, width, bottom, align, color
```

```
imshow(Z, ...)
Z, cmap, interpolation, extent, origin
```

```
contour[f](X, Y, Z, ...)
X, Y, Z, levels, colors, extent, or
```

```
pcolormesh(X, Y, Z, ...)
X, Y, Z, vmin, vmax, cmap
```

```
quiver([X], [Y], [U, V, ...])
X, Y, U, V, C, units, angles
```

Advanced plots

```
step(X, Y, [fmt], ...)
X, Y, fmt, color, marker, where
```

```
boxplot(X, ...)
X, notch, sym, bootstrap, widths
```

```
errorbar(X, Y, xerr, yerr, ...)
X, Y, xerr, yerr, fmt
```

```
hist(X, bins, ...)
X, bins, range, density, weights
```

```
violinplot(D, ...)
D, positions, widths, vert
```

```
barbs([X], [Y], [U, V, ...])
X, Y, U, V, C, length, pivot, sizes
```

```
eventplot(positions, ...)
positions, orientation, lineoffsets
```

```
hexbin(X, Y, C, ...)
X, Y, C, gridsz, bins
```

Scales

```
ax.set_... linear any values
```

```
ax.set_... log values > 0
```

```
ax.set_... symlog
```

```
ax.set_... logit
```

Projections

```
ax.set_... p='polar'
```

```
ax.set_... p='3d'
```

Line & Marker Styles

Lines

```
linestyle or ls
```

```
capstyle or dash_capstyle
```

```
"butt" "round" "projecting"
```

Markers

```
marker
```

```
marker every
```

```
["0", "1", "2", "3", "4", "5", "6", "7", "8", "9", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "R", "S", "T", "U", "V", "W", "X", "Y", "Z", "a", "b", "c", "d", "e", "f", "g", "h", "i", "j", "k", "l", "m", "n", "o", "p", "q", "r", "s", "t", "u", "v", "w", "x", "y", "z", "A", "B", "C", "D", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "O", "P", "Q", "
```

Spatial Data Visualizations Using Maps

Interactive Visualizations