



Data Visualization



Claire Babirye MS, Dorothy Kabarozi Bugingo
BS, Hewitt Tusiime BS



Definition

There's a story behind your numbers. Visualizing Data helps them come to life.....

Data Visualization is the communication of data in a visual manner or turning raw data into insights that can be easily interpreted by your readers.

Why visualize data?

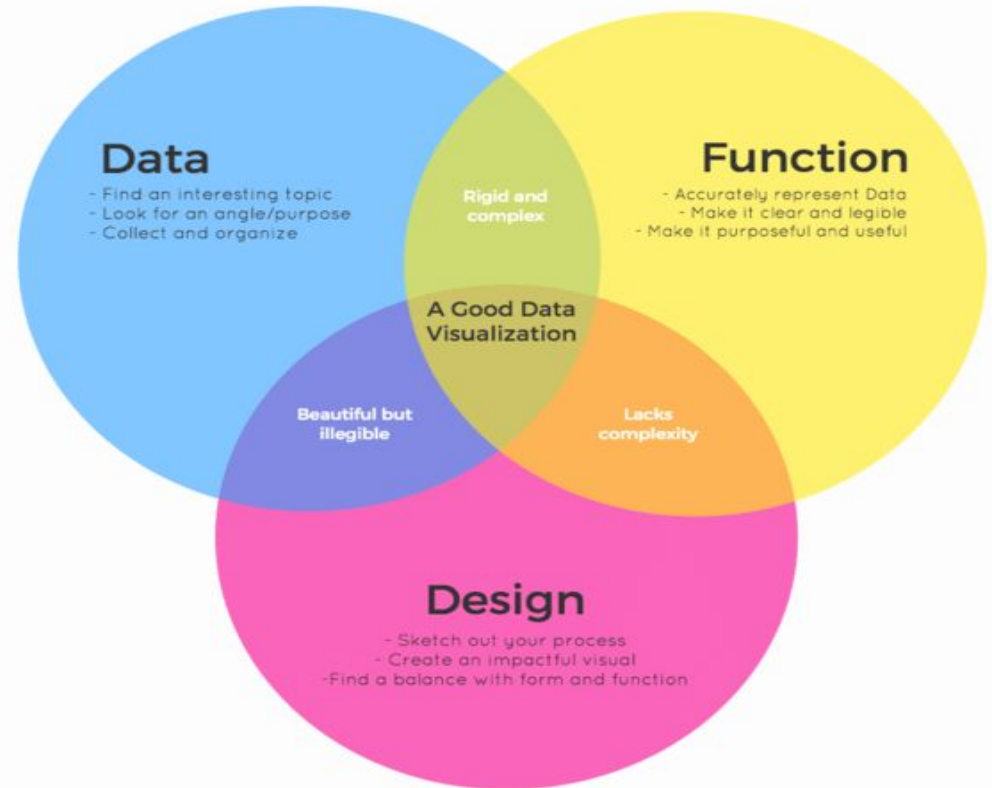
A picture is worth a thousand words

It is easier to remember pictures than text

Understanding data and interpret it to make decisions

Can summarize large amounts of complex data

What Makes A Good DATA VISUALIZATION





The purpose:

- Story telling
- Understand and Analyze data
- Communicate findings
- Quickly draw attention to key messages

How to use visualizations to
communicate effectively?

The background is a solid blue color. On the right side, there is a series of overlapping, semi-transparent geometric shapes. These include a large, light blue parallelogram and several smaller, darker blue rectangles and parallelograms. A single, bright green parallelogram is also visible, partially overlapping the blue shapes.

1

Decide on what your visualization should convey

FOCUS ON THE DATA

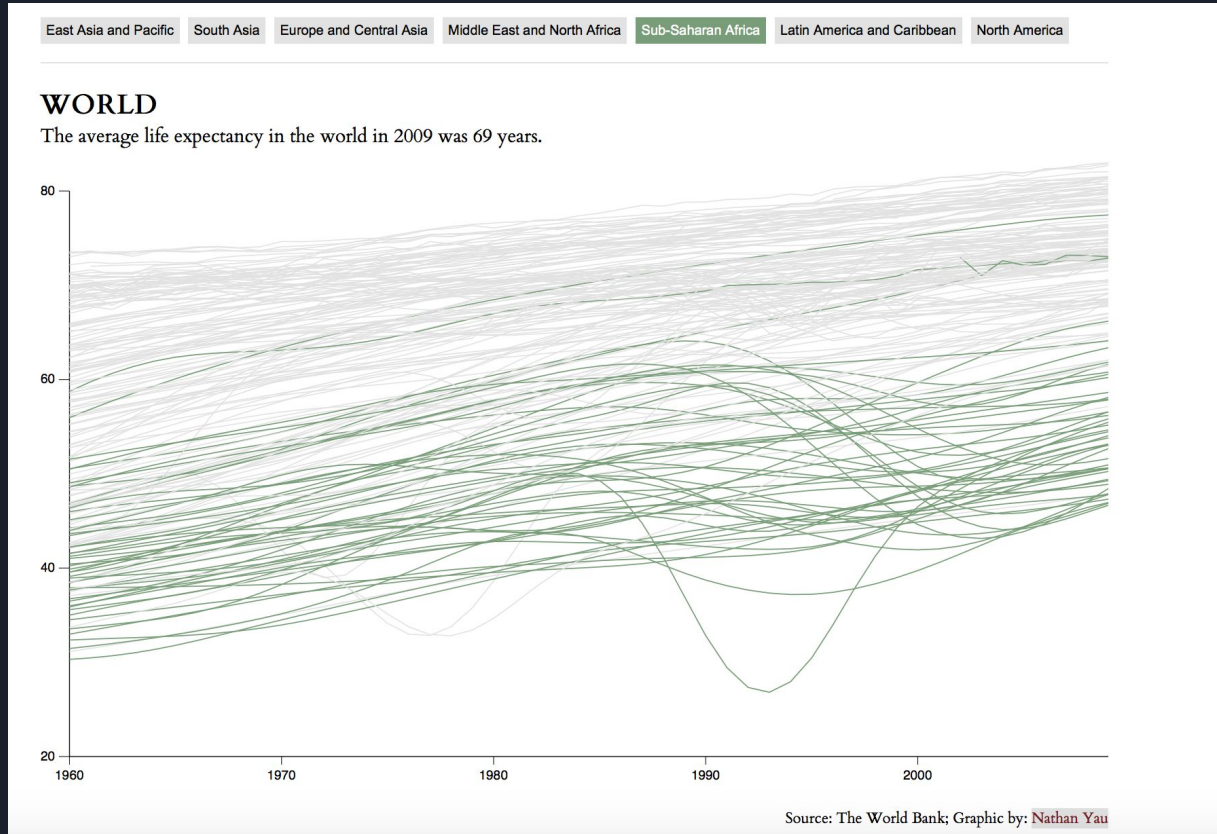
The style and structure of your visualization will depend on its **purpose**

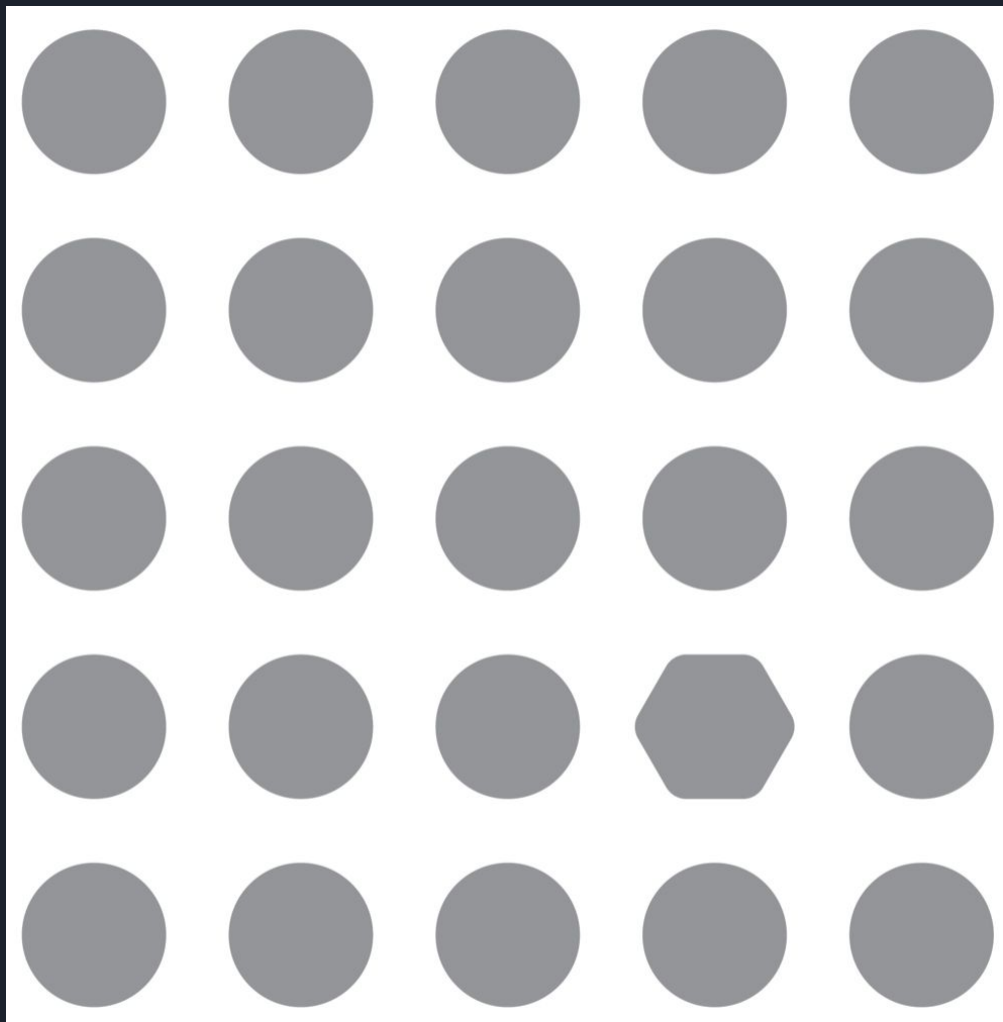
Understand your Audience

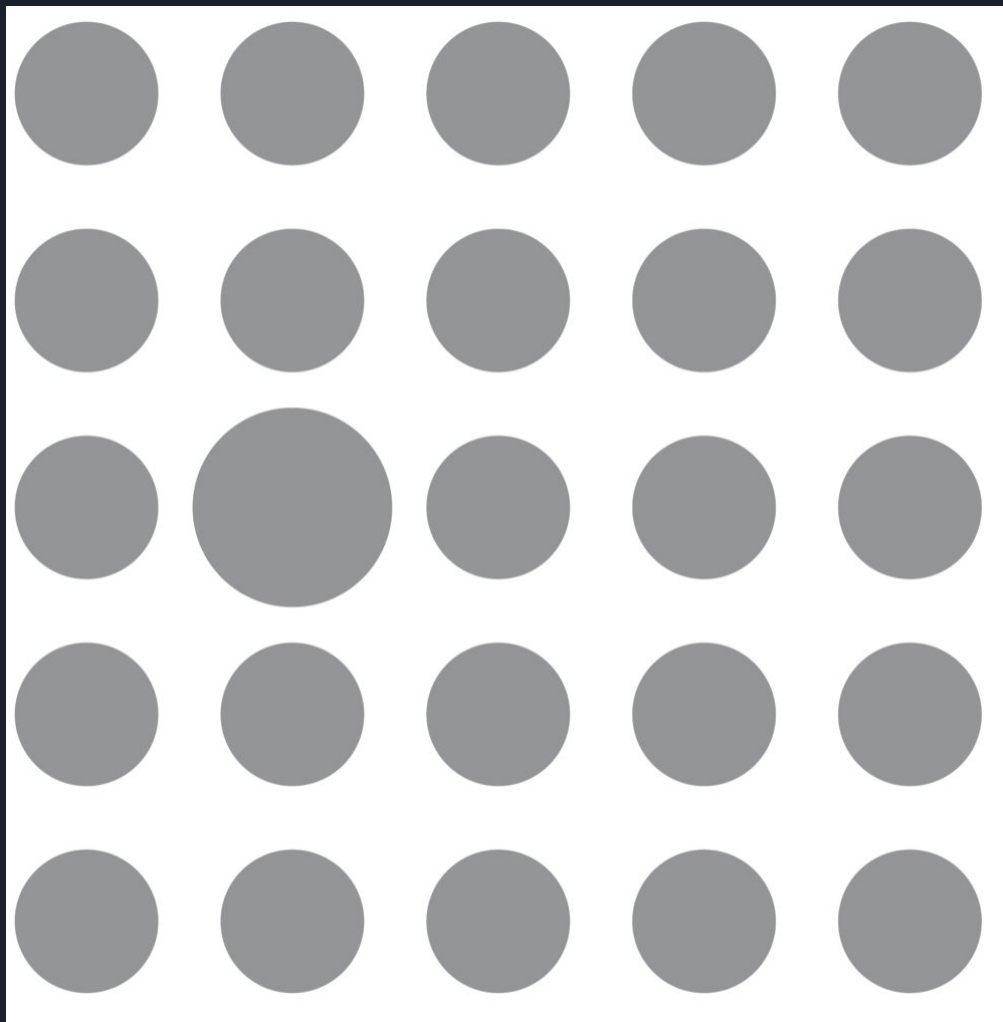
Tell a good story and don't visualise what does not matter

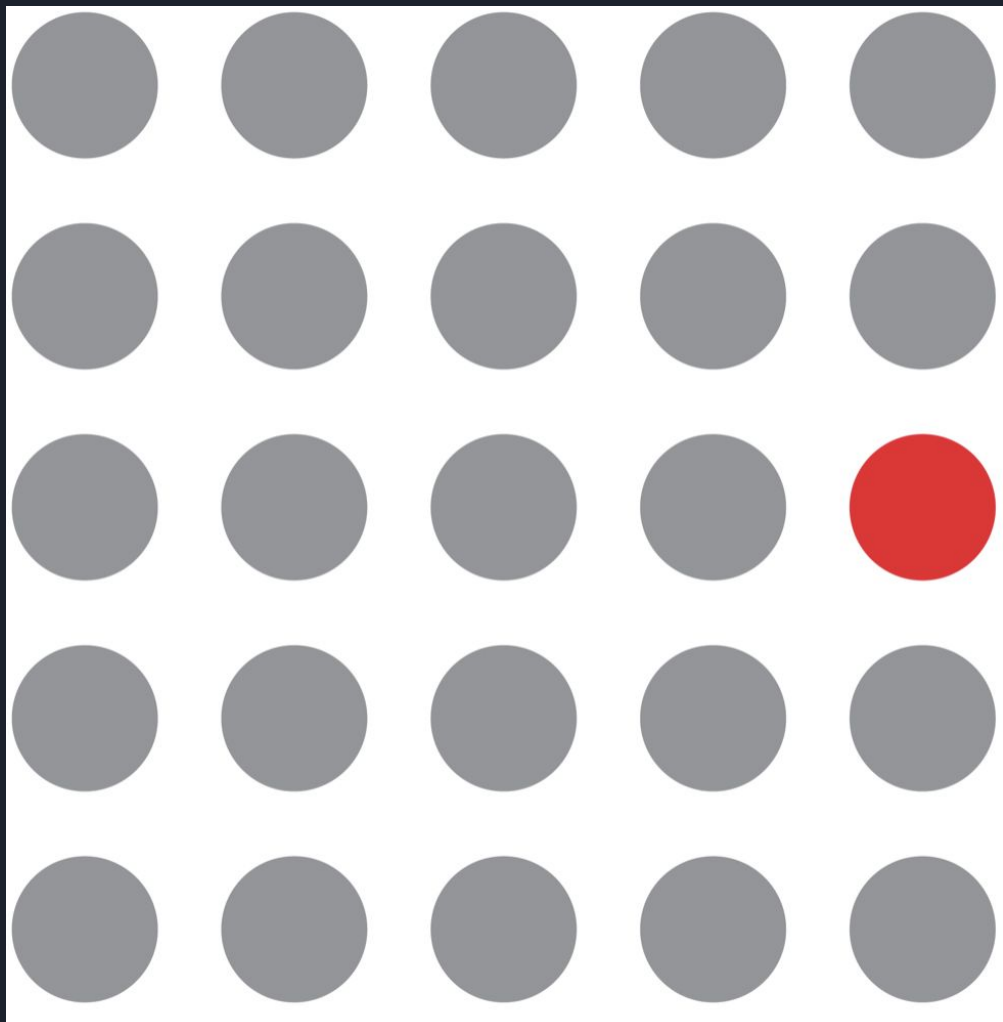
2

Use color and size to highlight and suppress information





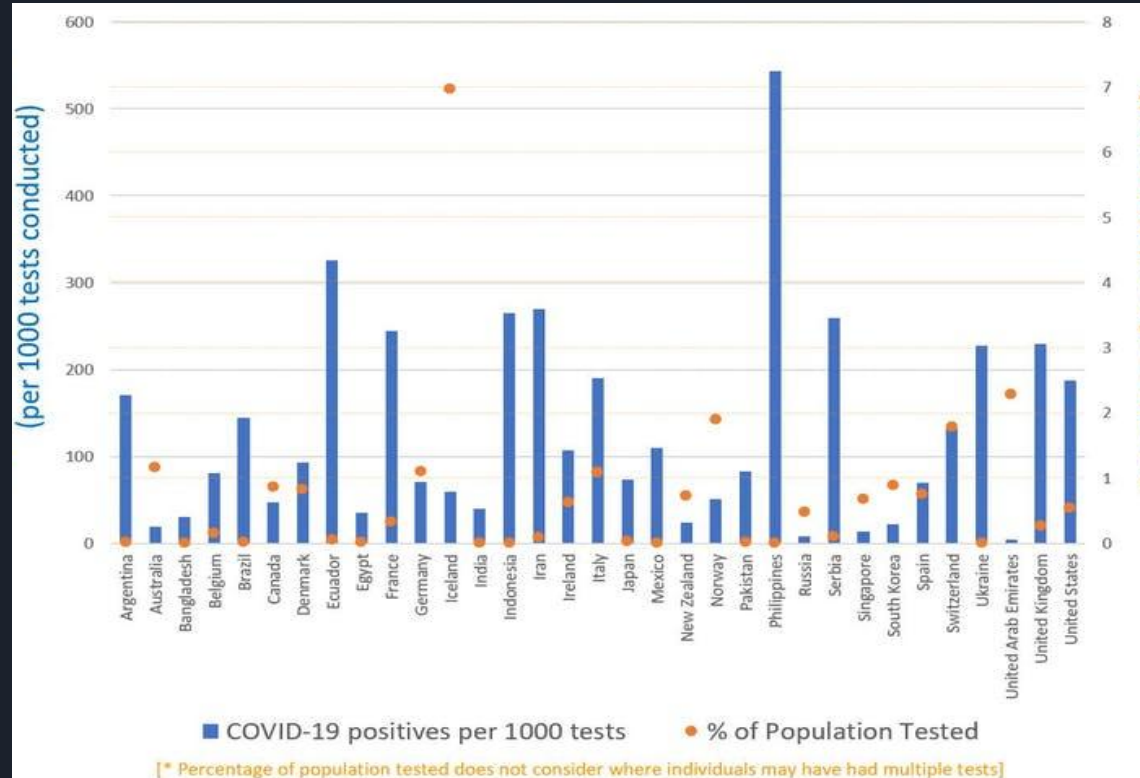




3

Use length and position to express quantitative information. Use color for categorical information

Column and bar charts allow for **more accurate comparison of information over time** compared to pie charts



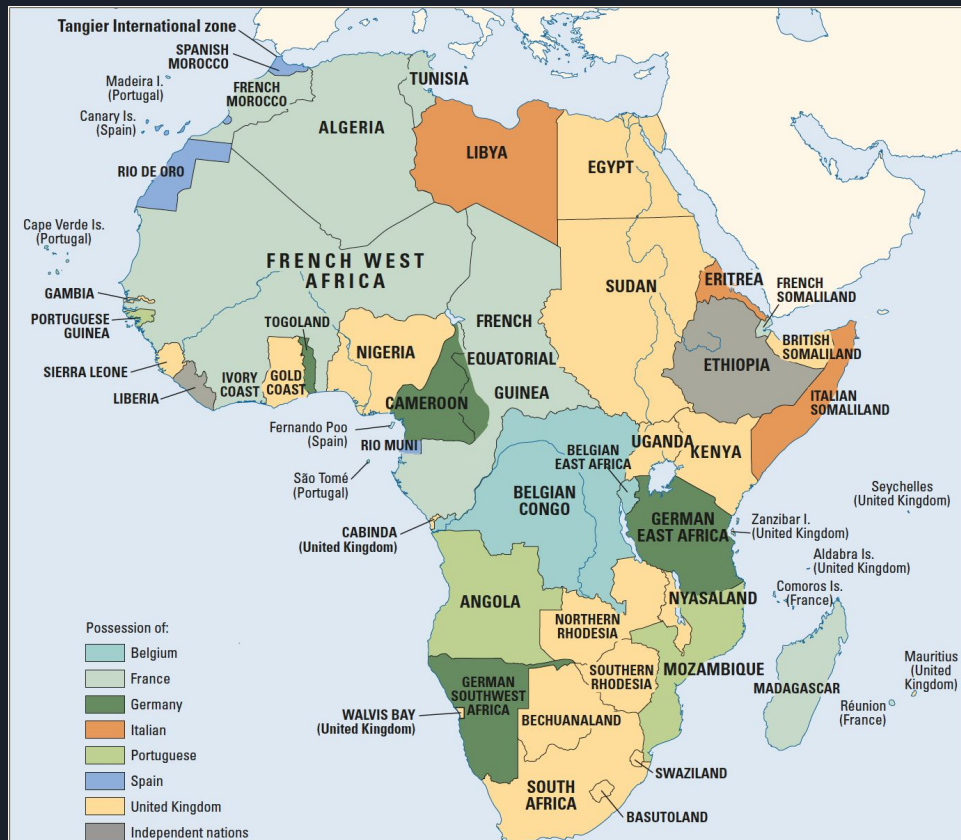
4

Think carefully about color selection and usage

Use color to create groupings

Add a **single color** to a black and white image

Use black and white to add contrast to an image with a single color gradient



5

Think carefully about color selection and usage

Consider
those with
color
blindness

Red

Stop

Dangerous

Hot

Green

Moving

Money

Plants

Blue

Water

Cool

Safe

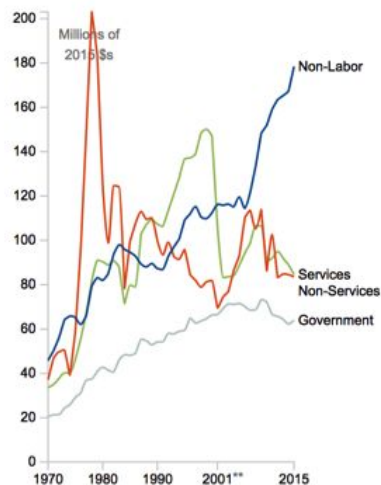
6

Use all available space and proper scales

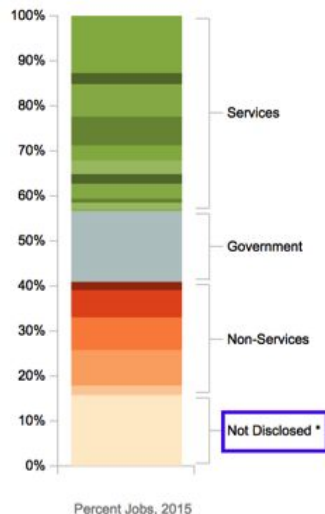
Moffat, CO

Western Region

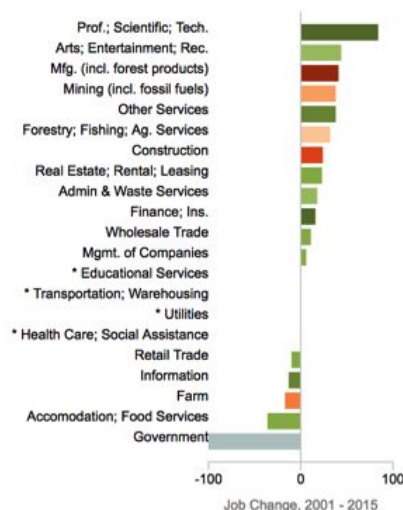
What are the sources of personal income growth?



What industries contribute the most jobs?



What industries are adding or losing jobs?



Scale does not always have to include zero unless there is some missing data

Don't ignore the nulls, show it and know how to visualise it

7

Use text and labels to improve interpretation



Use meaningful titles

Label axis, as needed

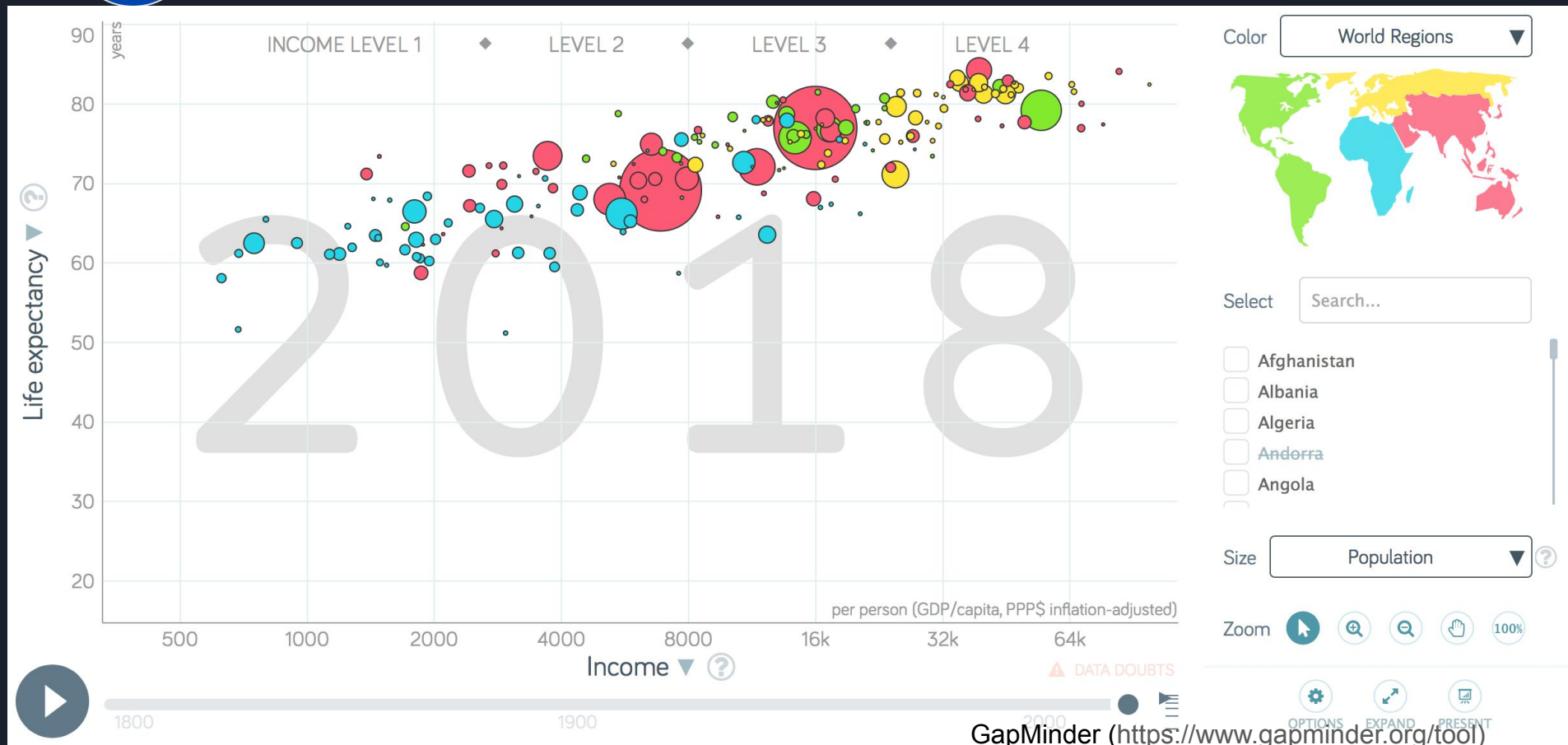
Add texts directly to the image - do not always rely on legends

Lines should not obstruct points

Use colors (e.g. light grey) and weight that lessen focus on tick marks and grids

8

Balance complexity and clarity

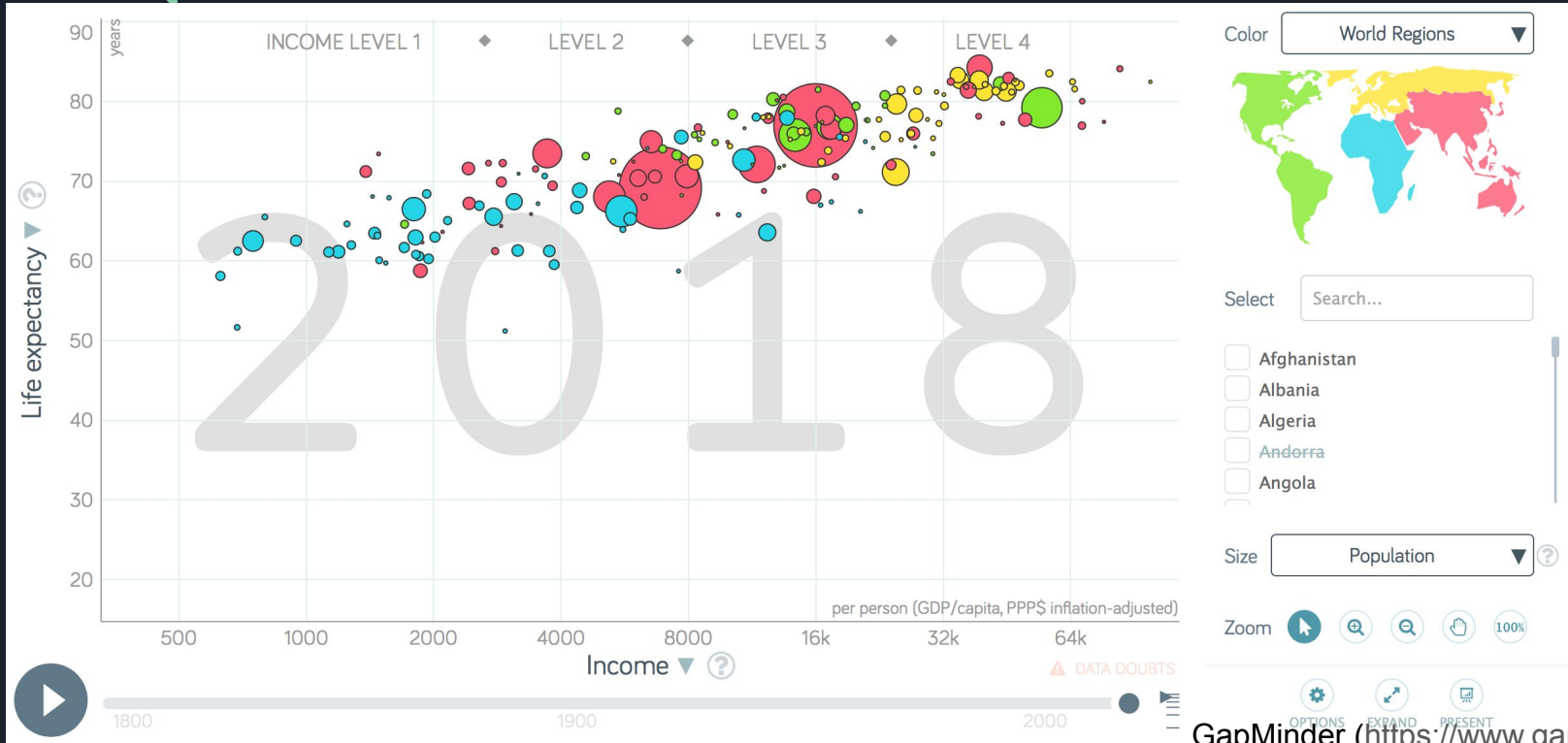


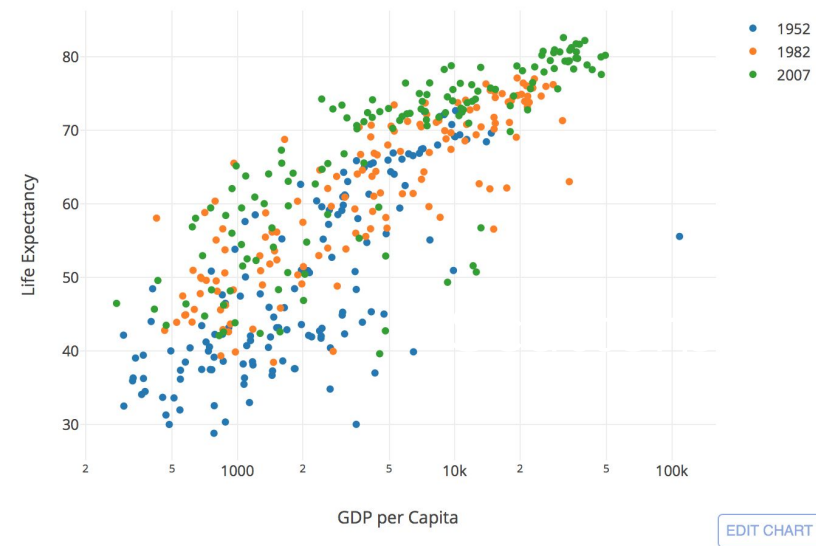
Examples



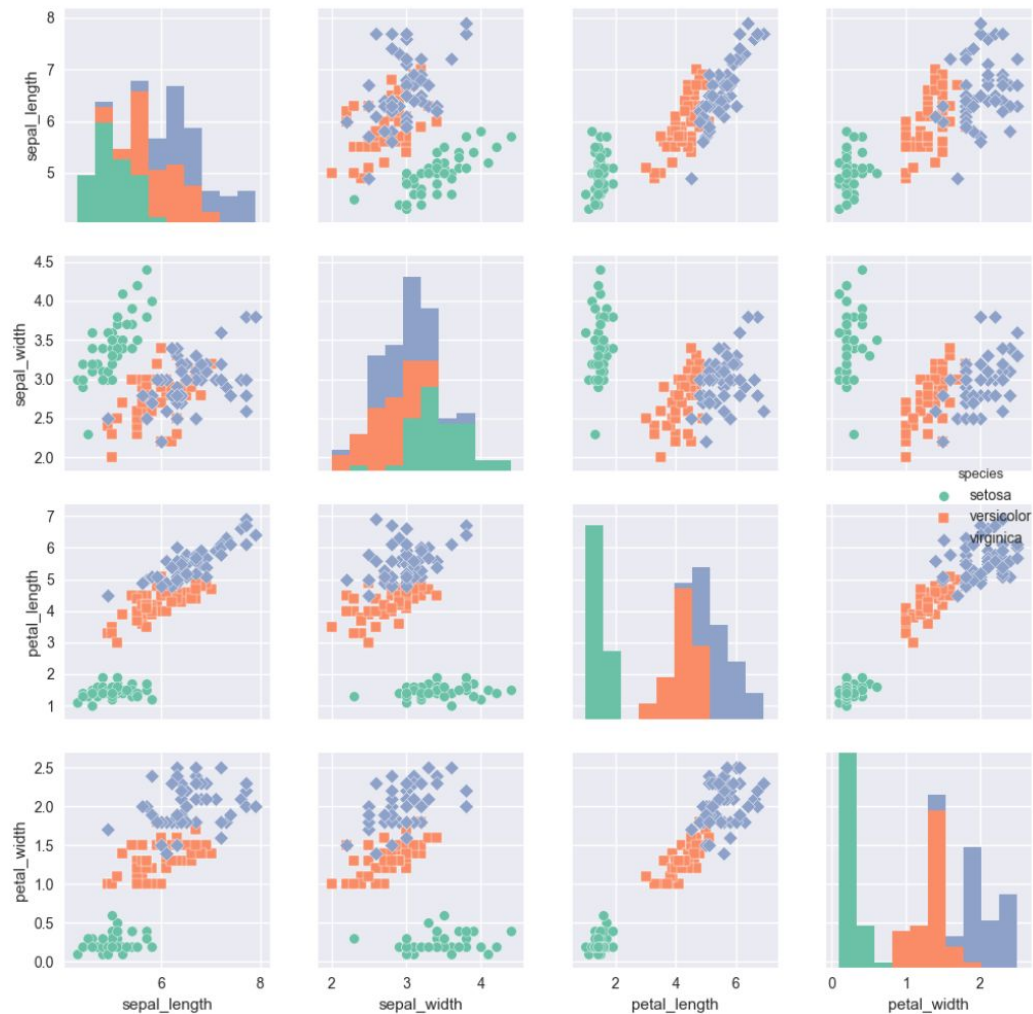
Bubbles

When to use?
Visualize correlation/association





- Connected scatter
- Correlogram
- Heatmap



When to use?

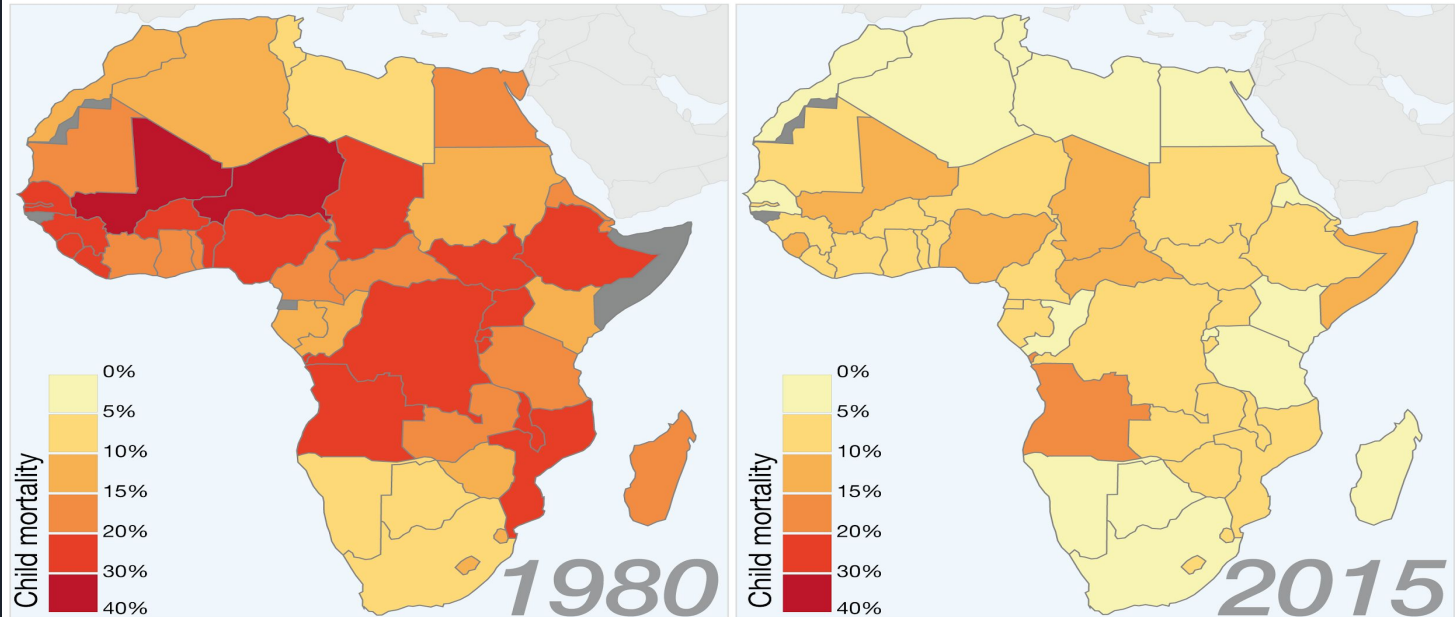
Useful for spatial visualizations

Maps

Child Mortality in 1980 and 2015

Child mortality is the probability that a newborn will die before reaching the age of 5.

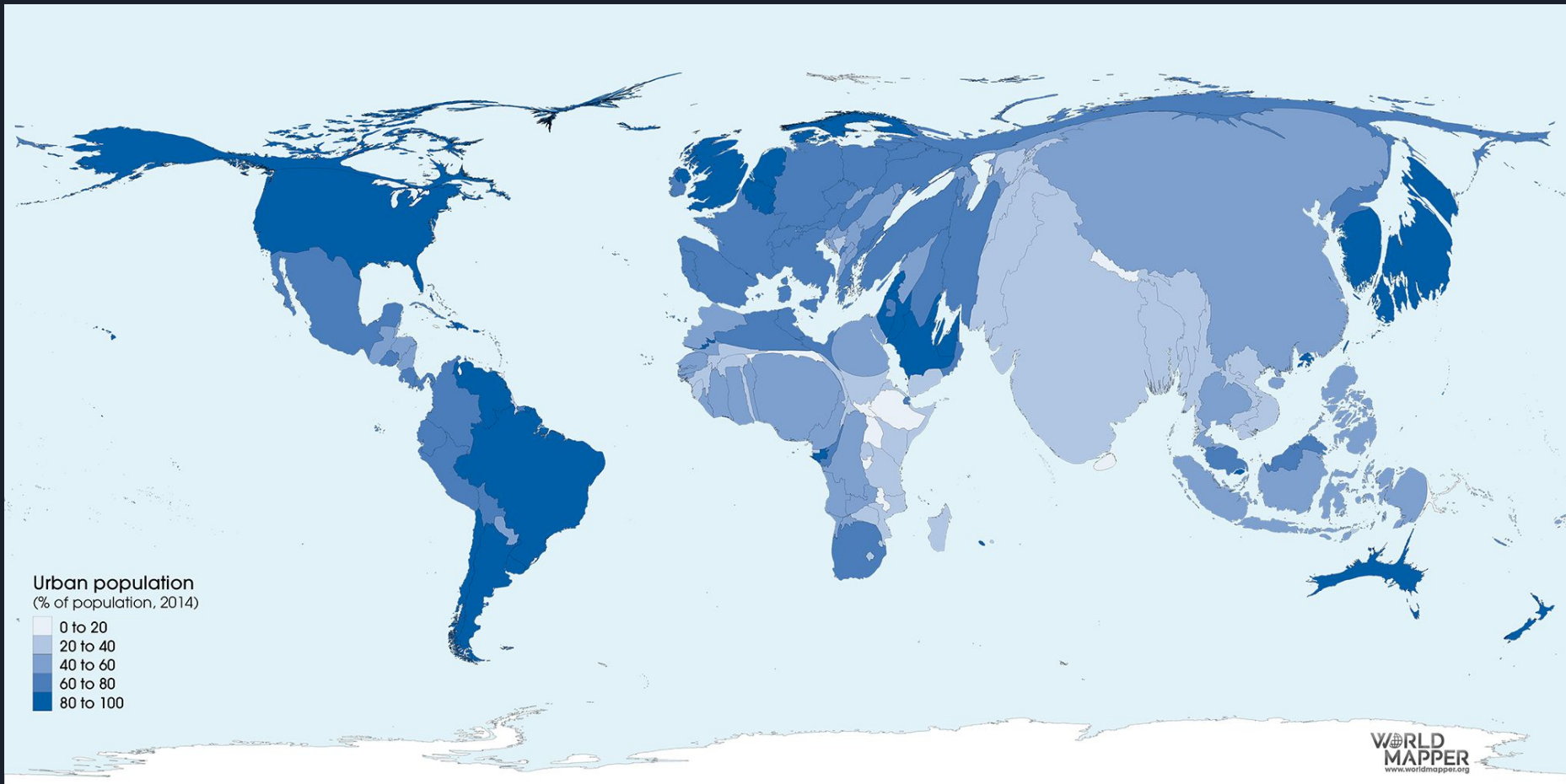
Our World
in Data



Data source: UN Child Mortality Estimates

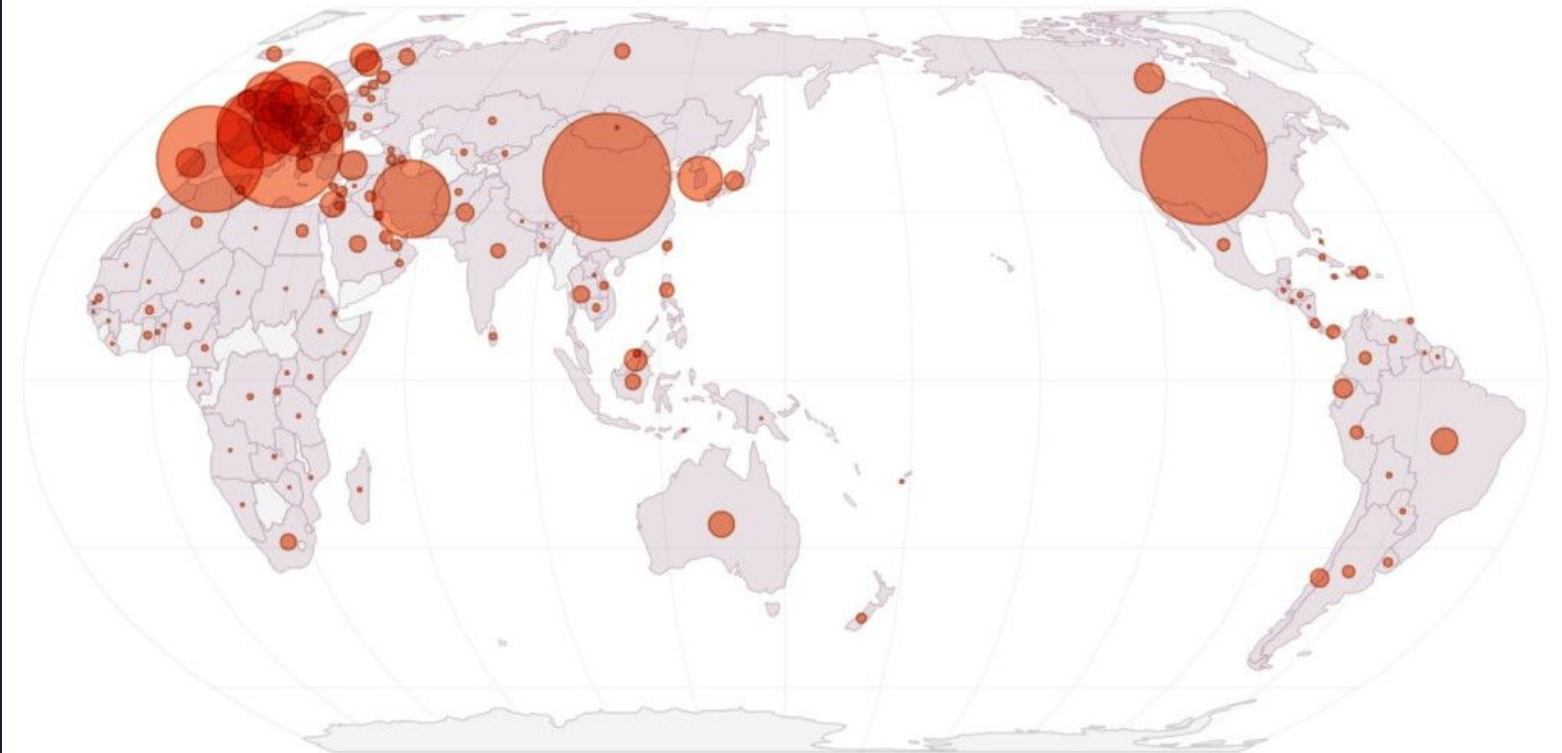
This data visualization is part of [AfricanData.org](https://africaindata.org) – an Our World in Data project.

Licensed under [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) by the author Max Roser.



Confirmed cases
523,163

Reported deaths
23,639



- Maps with bubbles
- Maps with pins

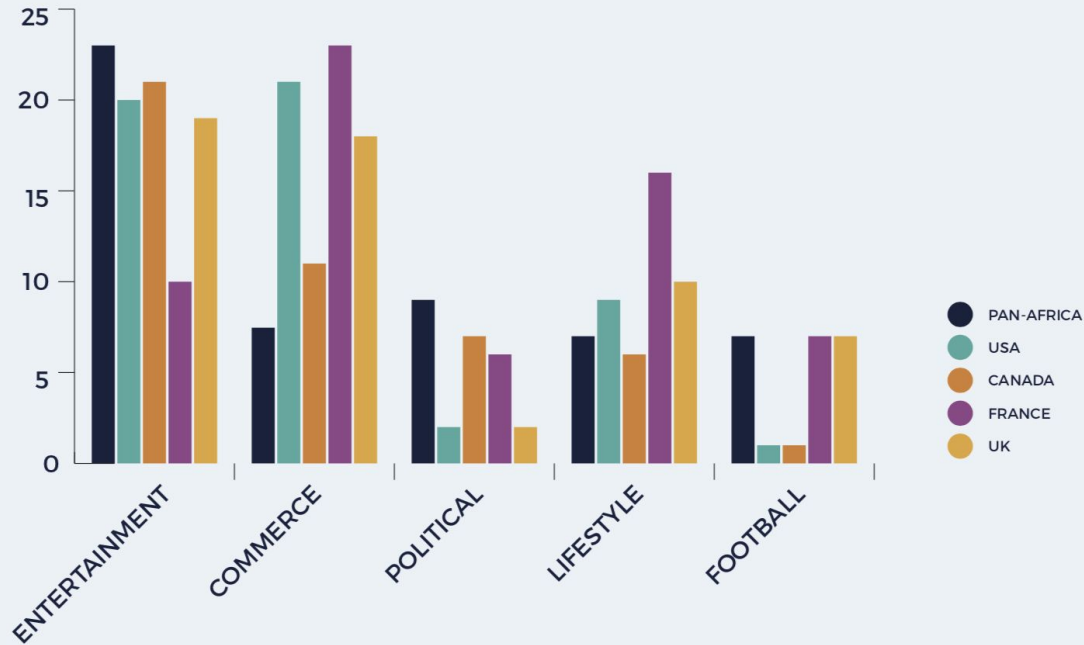
Map showing global spread
on coronavirus as of march
20 by the washington post

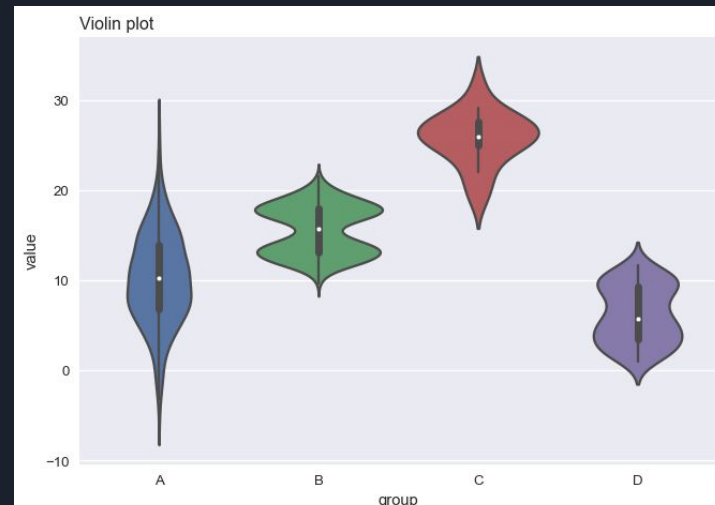
Bar plots

When to use?

Useful for rankings

Top five themes of hashtags around the world

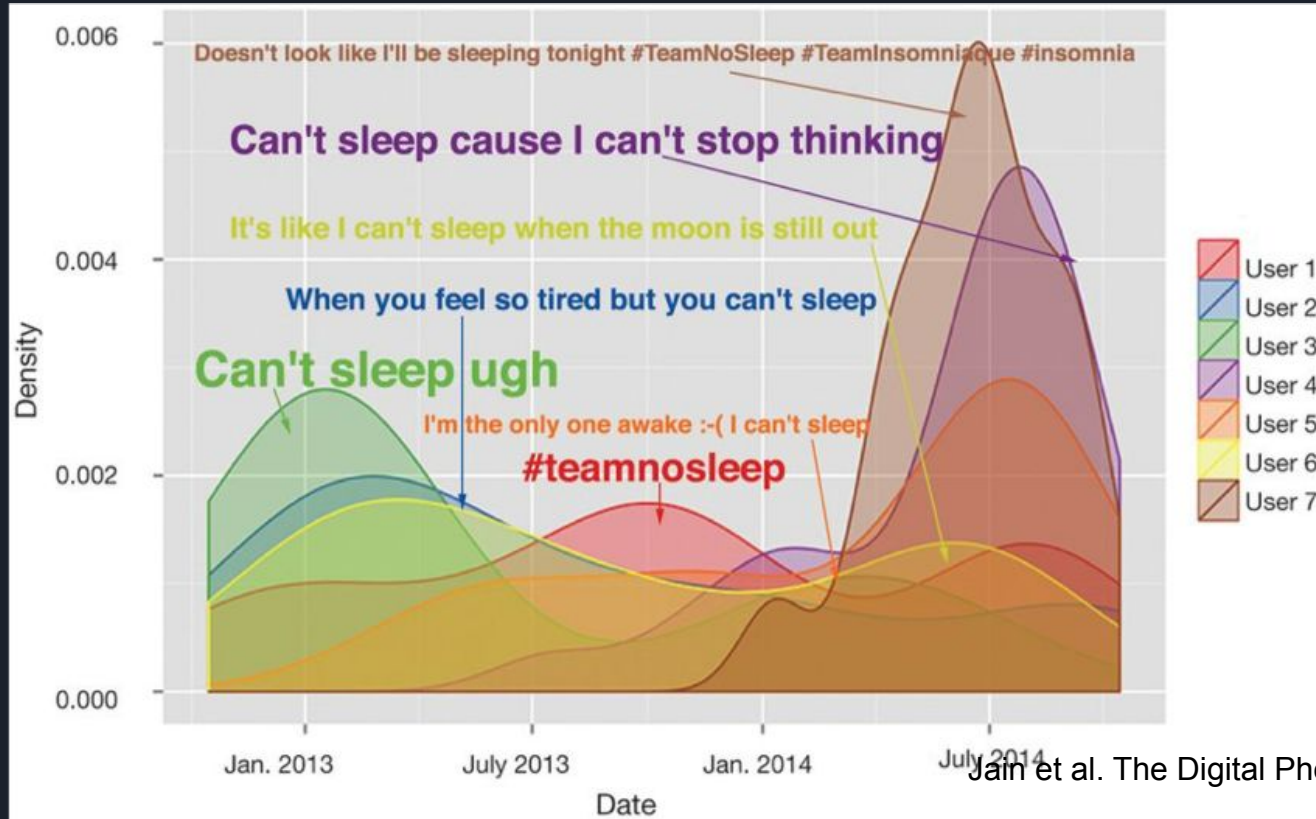


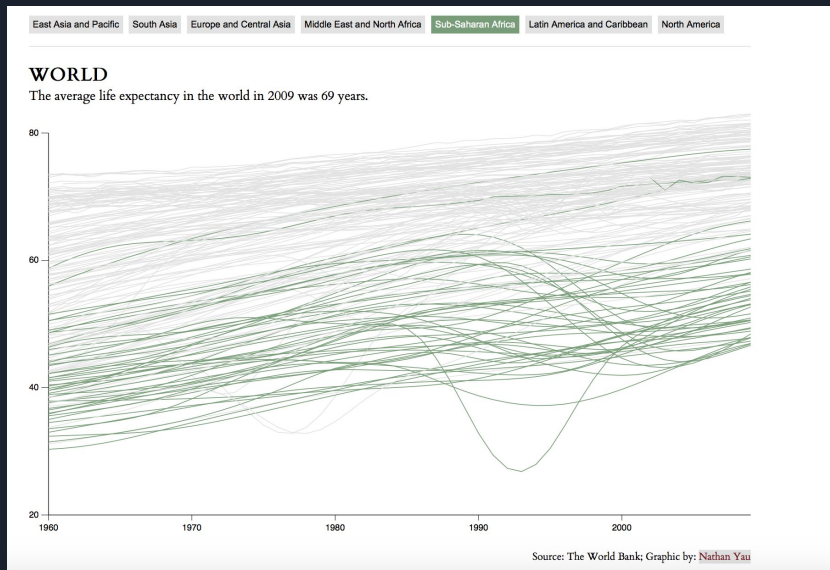


Area/density plots

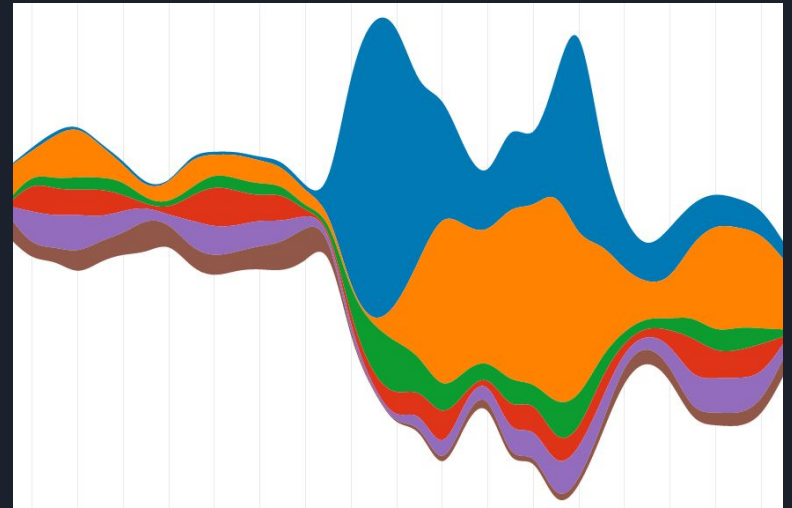
When to use?

Useful for showing evolution





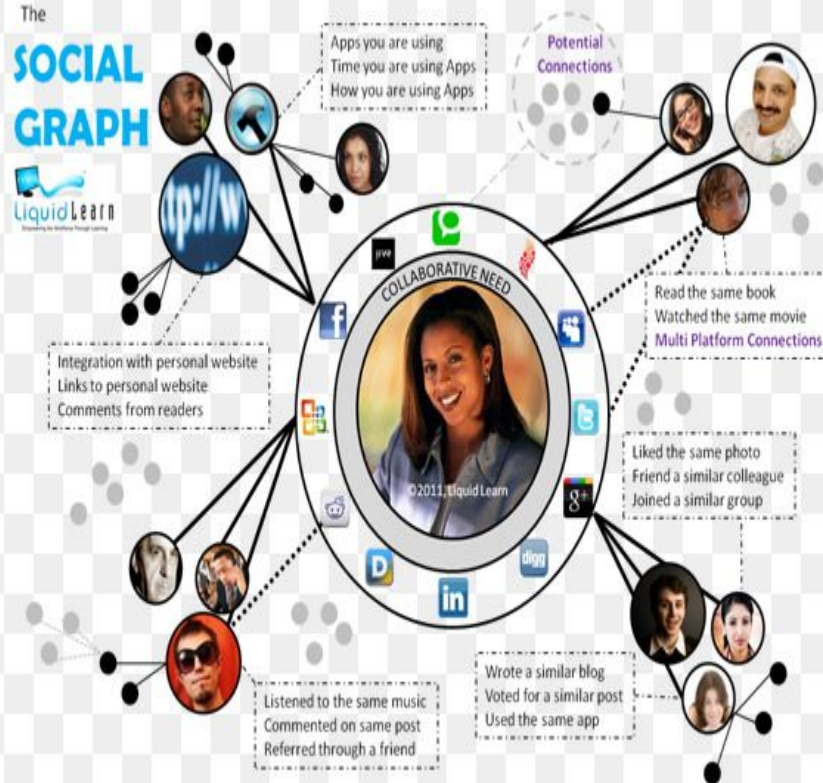
- Line plot
- (Stacked) area plot
- Stream chart



Networks

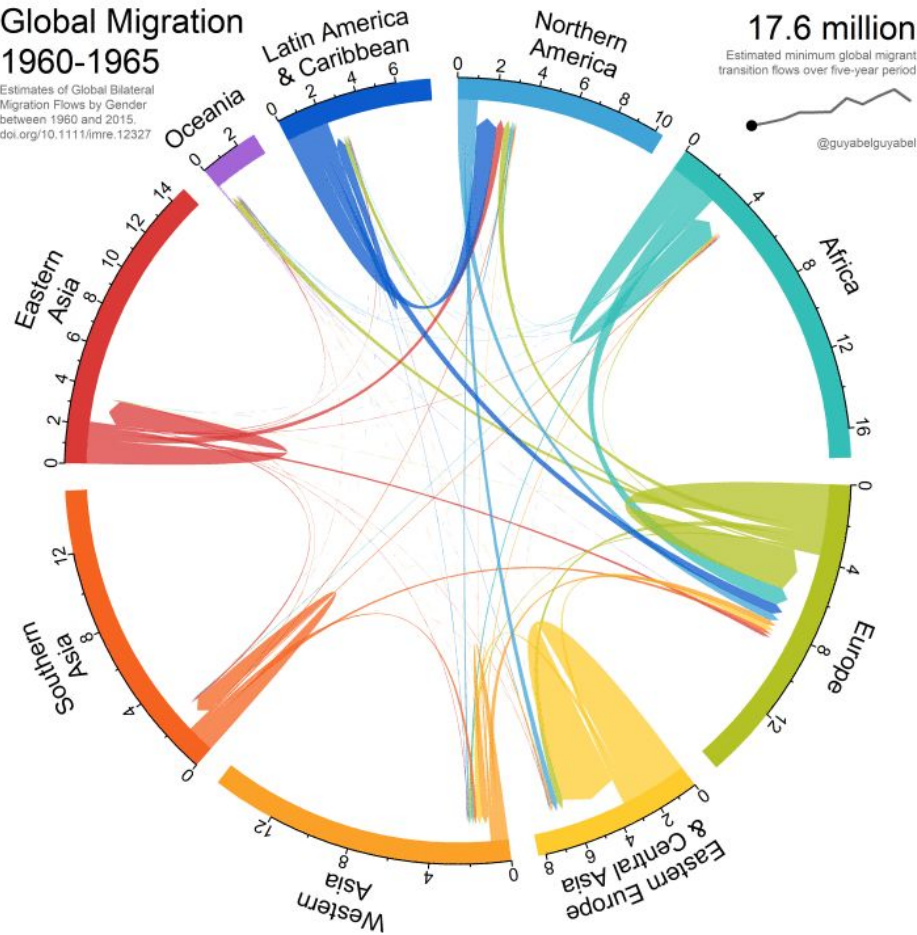
When to use?

Useful for information flow



Global Migration 1960-1965

Estimates of Global Bilateral
Migration Flows by Gender
between 1960 and 2015.
doi.org/10.1111/imre.12327



Code available from:

<https://guyabel.com/post/animated-directional-chord-diagrams/>

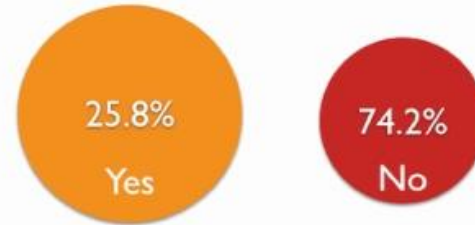
Chord diagram

Bad visualizations



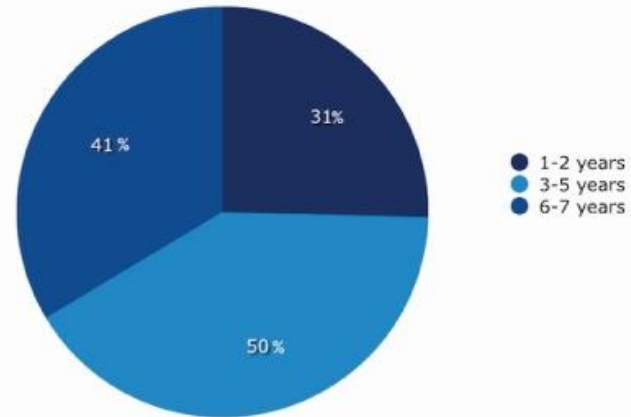
Which of these
images is quite
confusing?

Have You Ever Liked a Brand on Facebook?



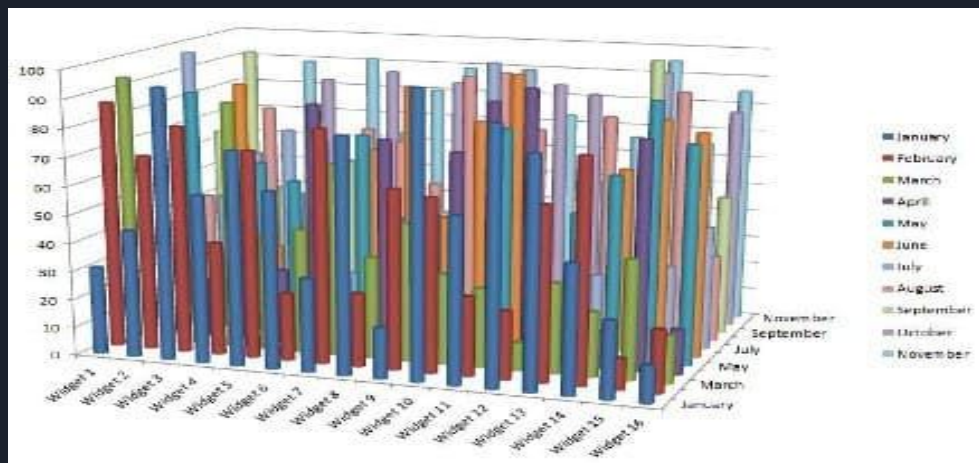
A

Years Experience Required by Employers

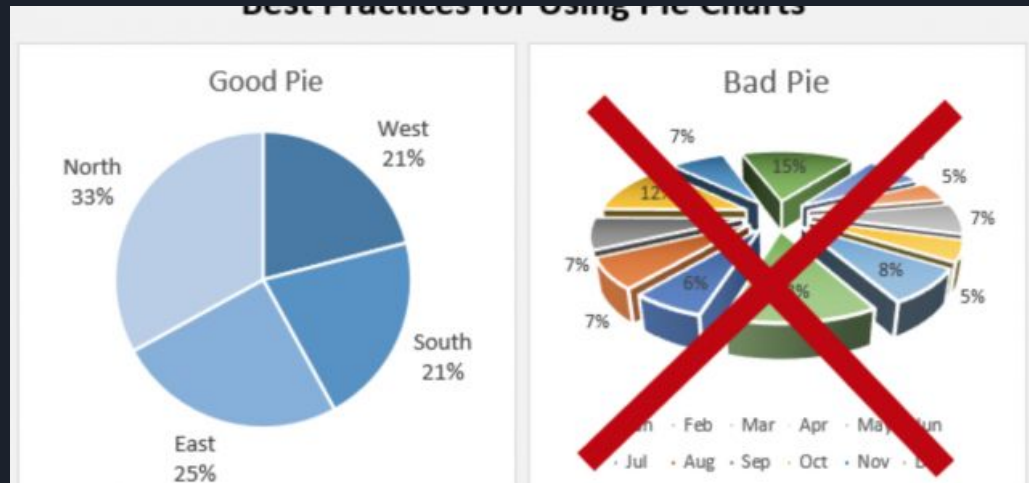


C

Bad visualizations



Best Practices for Using Pie Charts



Tools and Resources





Python libraries

- Matplotlib
- ggplot
- Seaborn
- Bokeh
- Pygal
- Plotly
- Geoplotlib
- Gleam
- Missingno
- Leather
- Pydot



Other tools

- Tableau
- R ggplot2 and others
- D3

Less is More

“Perfection is Achieved Not When There Is Nothing More to Add, But When There Is Nothing Left to Take Away” – Antoine de Saint-Exupery