

A project by Yan Holtz and Conor Healy

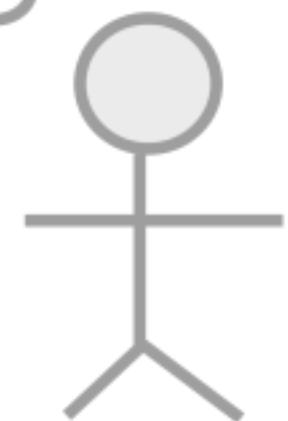
from Data to Viz

A classification of chart types based on input data format

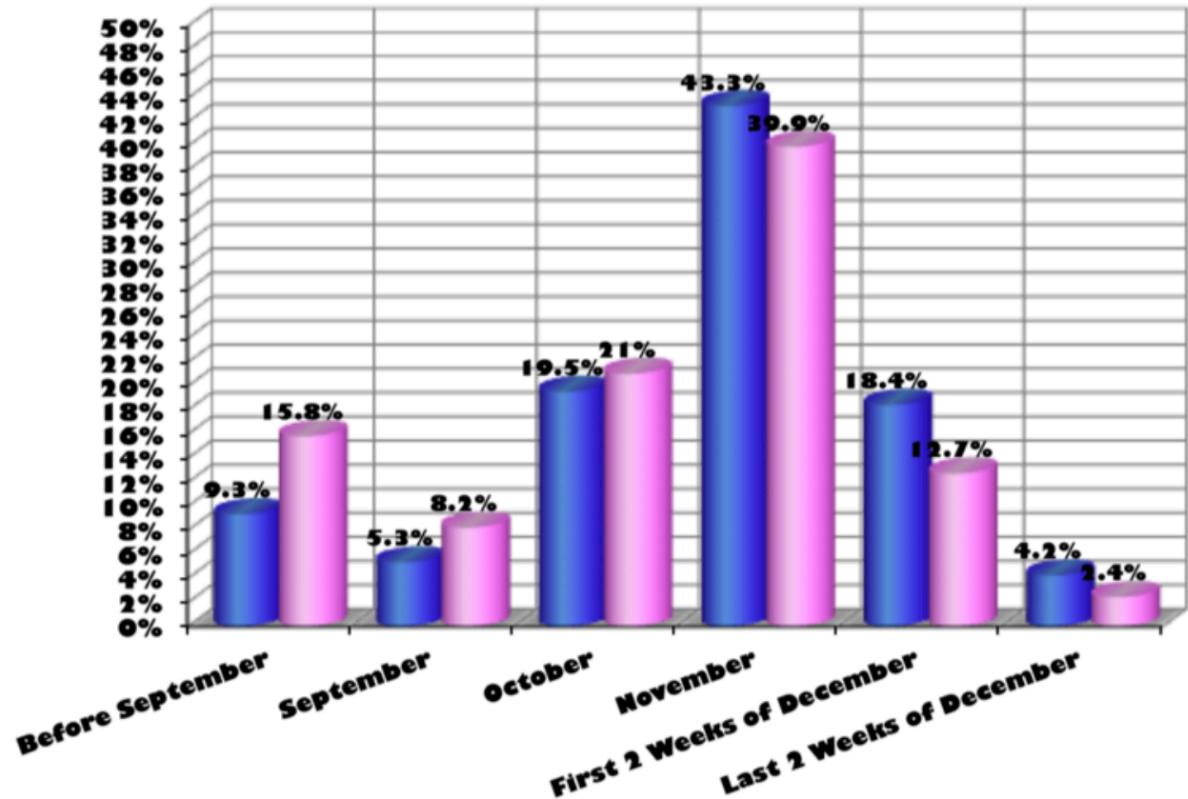
DataViz Meetup | Sydney | November 2018

Id	feature 1
A	10
B	12
C	15
...	...

What should I do with my
data ??



■ Men ■ Women



...

What about
that?

Source: [Storytelling with data](#) by Cole Nussbaumer Knaflic

Looking for a chart ?

What you
can do

What you
should do

Caveats to
avoid

How to
build it

WHAT YOU CAN DO

A classification of chart types based on data input format

Who sells more weapons ?

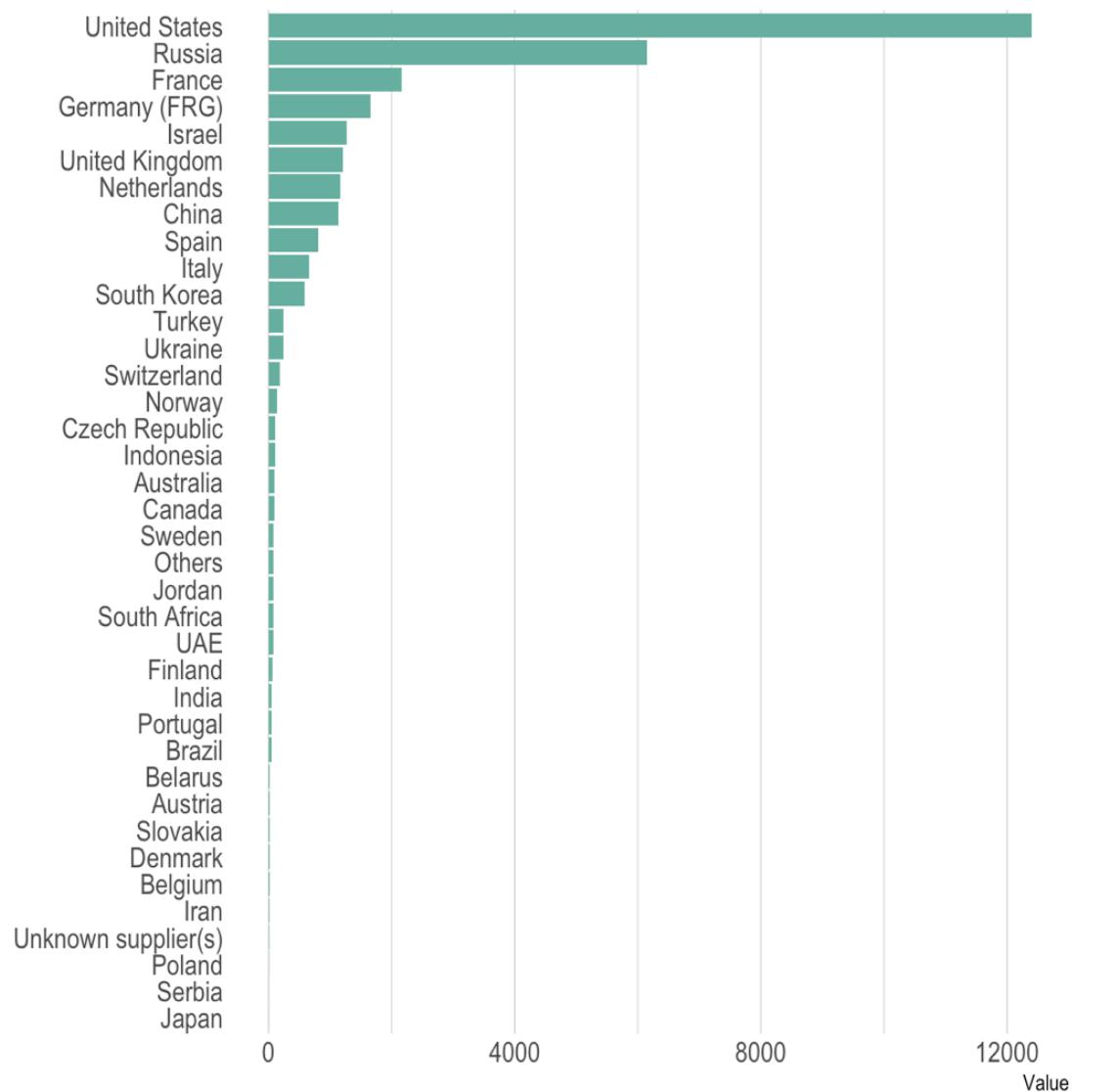
Country	Value
United States	12394
Russia	6148
Germany (FRG)	1653
France	2162
United Kingdom	1214
China	1131



?

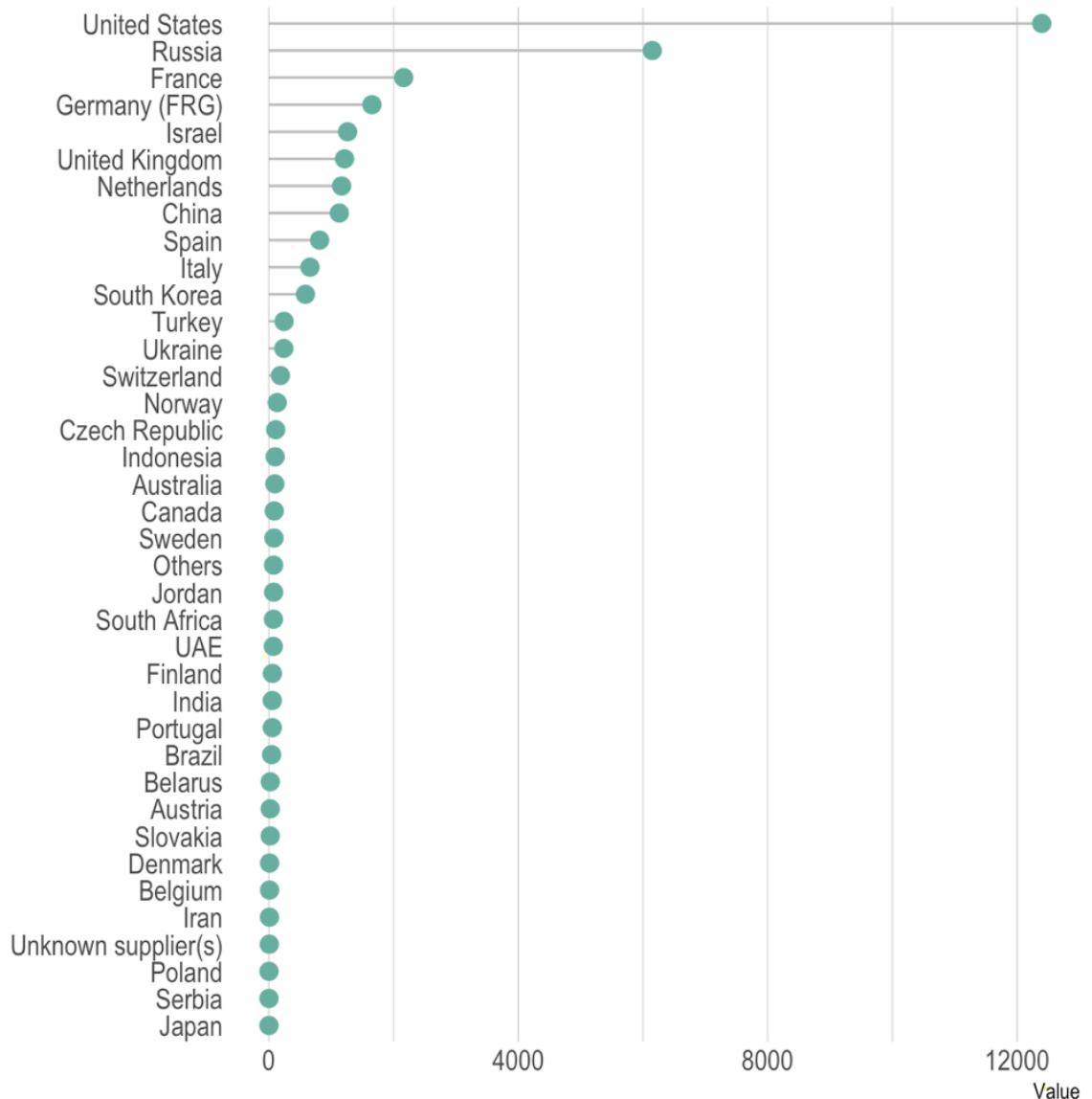
Who sells more weapons ?

Country	Value
United States	12394
Russia	6148
Germany (FRG)	1653
France	2162
United Kingdom	1214
China	1131



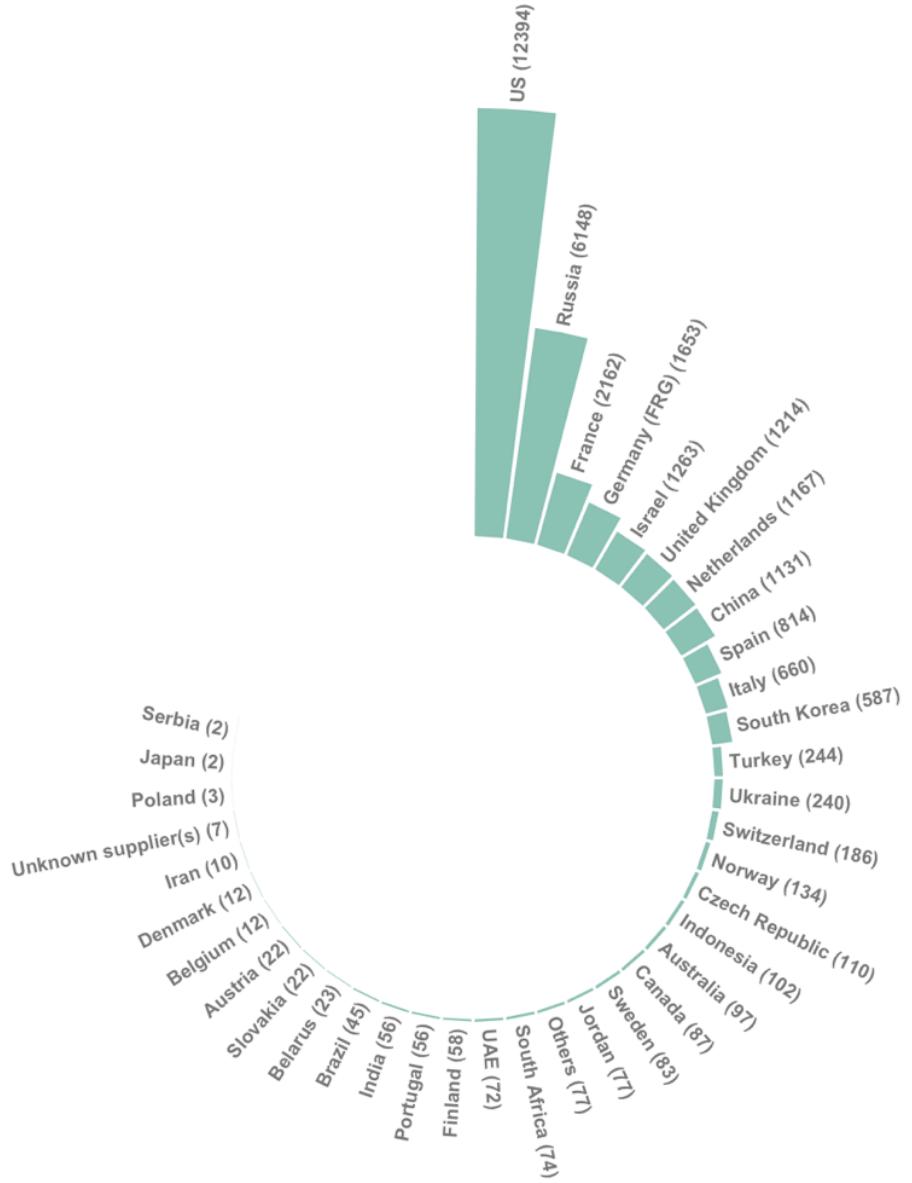
Who sells more weapons ?

Country	Value
United States	12394
Russia	6148
Germany (FRG)	1653
France	2162
United Kingdom	1214
China	1131



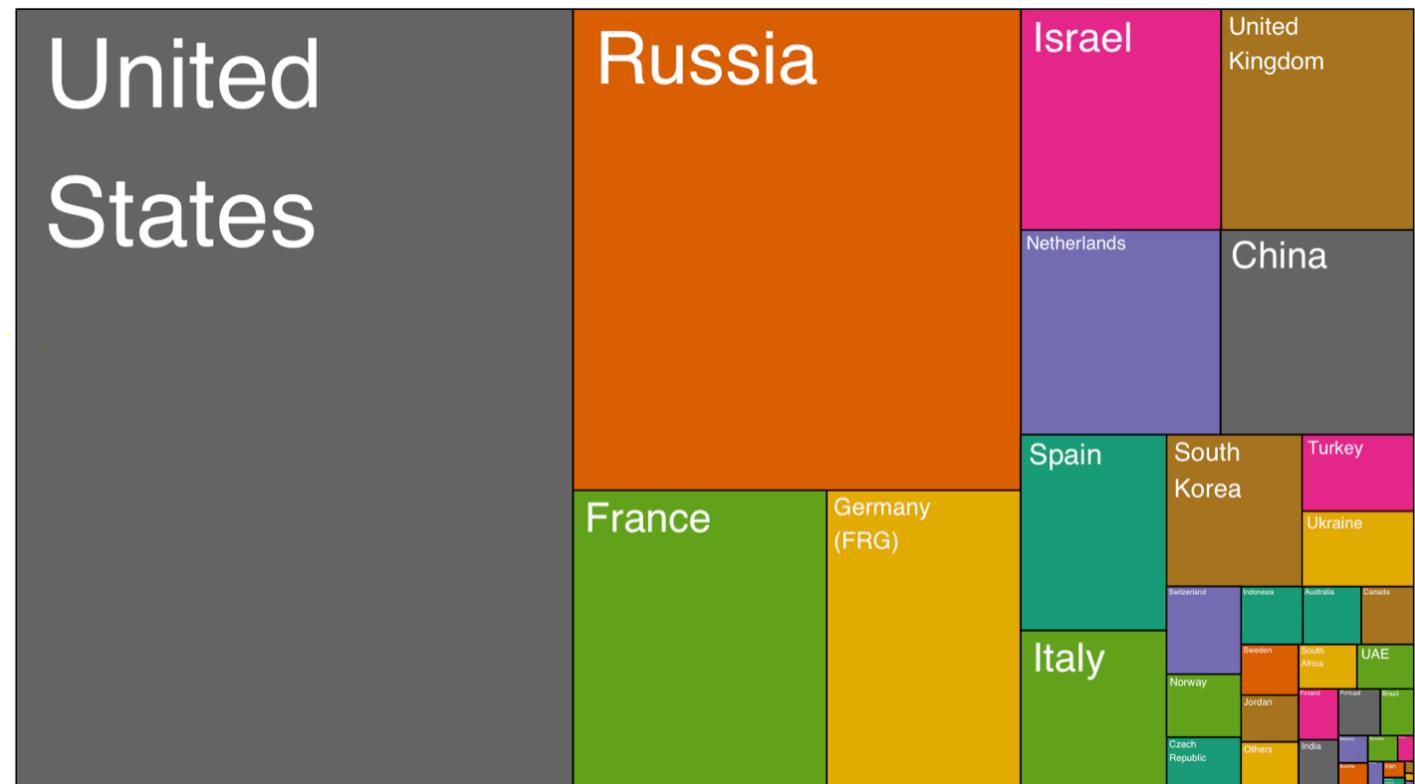
Who sells more weapons ?

Country	Value
United States	12394
Russia	6148
Germany (FRG)	1653
France	2162
United Kingdom	1214
China	1131



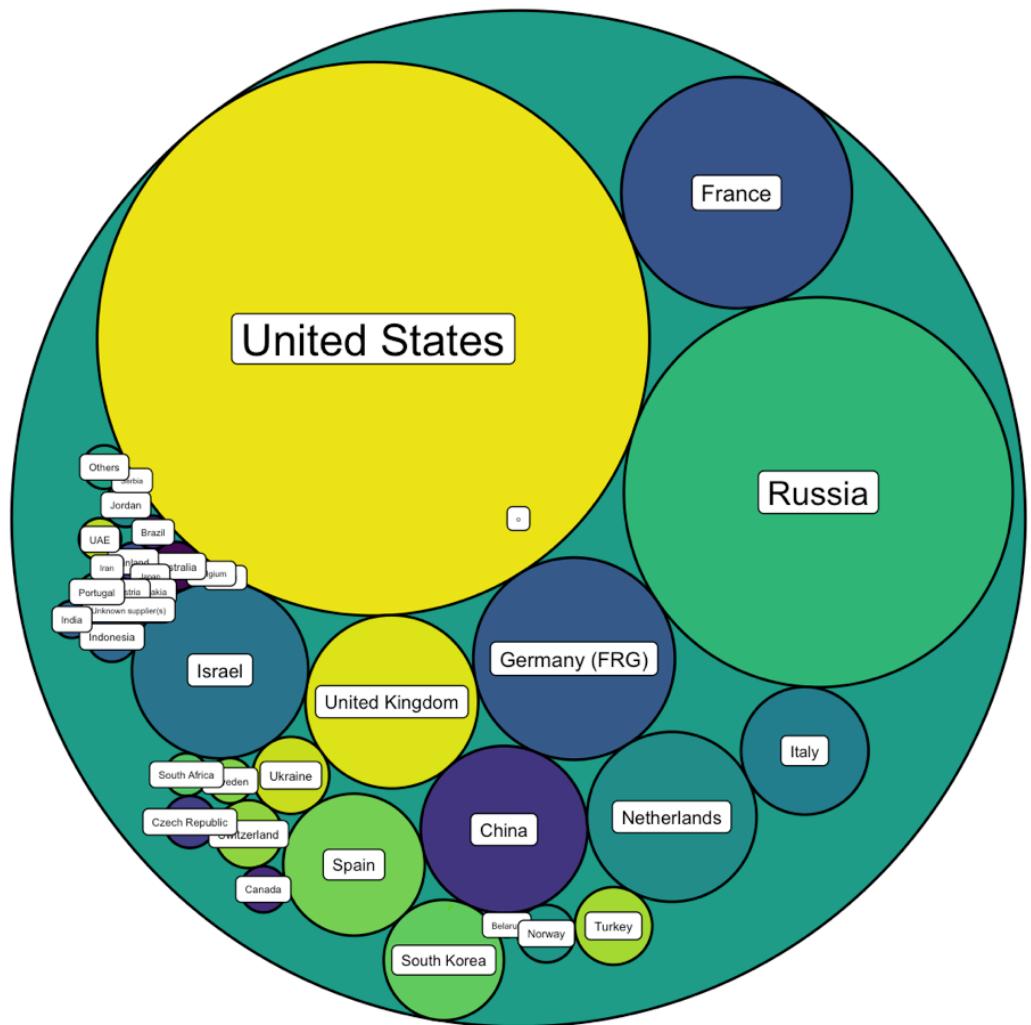
Who sells more weapons ?

Country	Value
United States	12394
Russia	6148
Germany (FRG)	1653
France	2162
United Kingdom	1214
China	1131



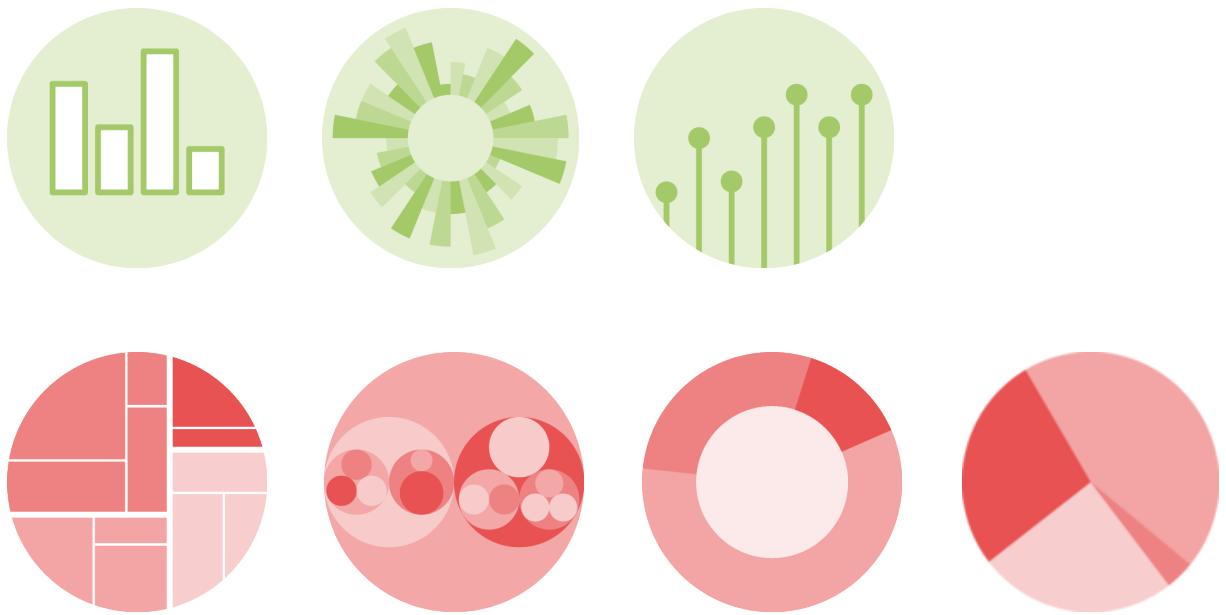
Who sells more weapons ?

Country	Value
United States	12394
Russia	6148
Germany (FRG)	1653
France	2162
United Kingdom	1214
China	1131



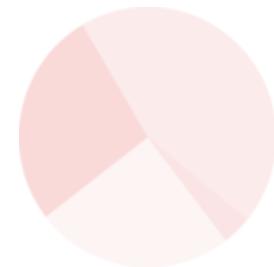
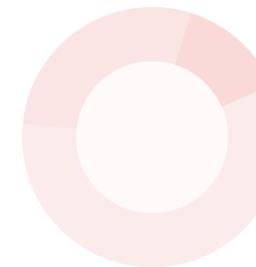
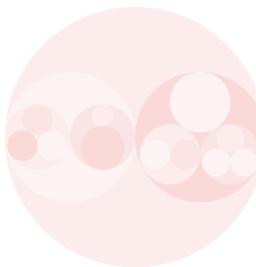
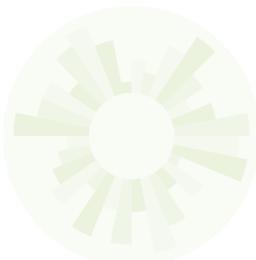
Who sells more weapons ?

Country	Value
United States	12394
Russia	6148
Germany (FRG)	1653
France	2162
United Kingdom	1214
China	1131



Who sells more weapons ?

Country	Value
United States	12394
Russia	6148
Germany (FRG)	1653
France	2162
United Kingdom	1214
China	1131



1 categorical variable

1 numerical variable

1 observation per group

Perception of probability

text	value
Improbable	33
Almost Certainly	98
Likely	60
Almost Certainly	98
Unlikely	10
Probably Not	25
About Even	50
Probably	75

Perception of probability

text	value
Improbable	33
Almost Certainly	98
Likely	60
Almost Certainly	98
Unlikely	10
Probably Not	25
About Even	50
Probably	75

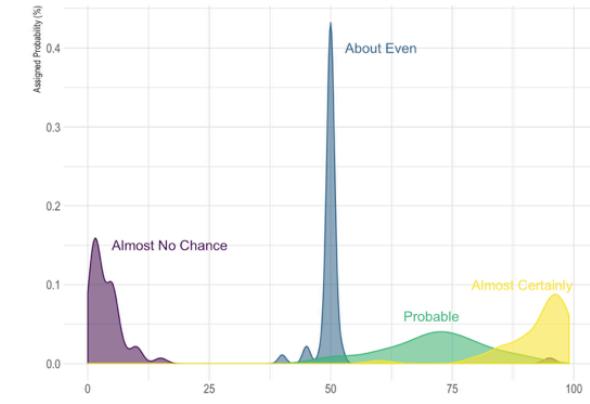
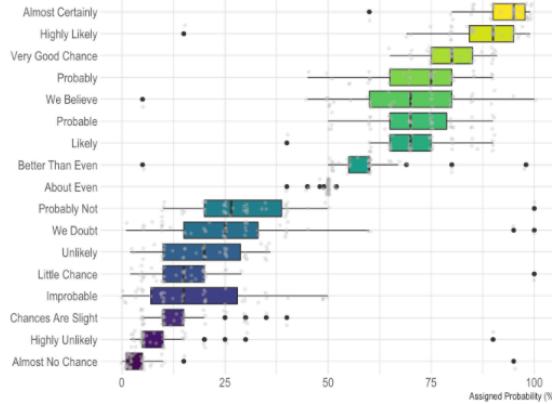
Several
observations
per group

1 categorical
variable

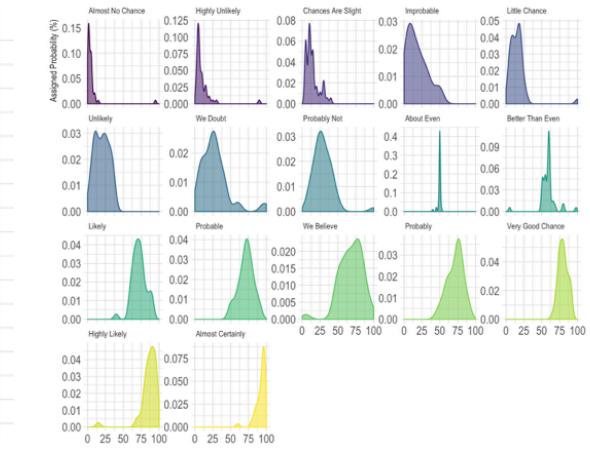
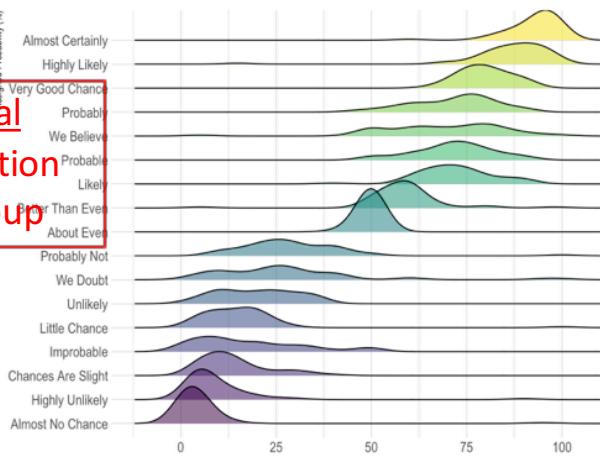
1 numerical
variable

Perception of probability

text	value
Improbable	33
Almost Certainly	98
Likely	60
Almost Certainly	98
Unlikely	10
Probably Not	25
About Even	50
Probably	75



Several
observation
per group



1 categorical
variable

1 numerical
variable

So...

- Knowing the **possibilities** is the **first step** in chart choice
- Not easy to know **all** the **chart types**
- Hard to figure out options **from a dataset**

→ Let's build a decision tree



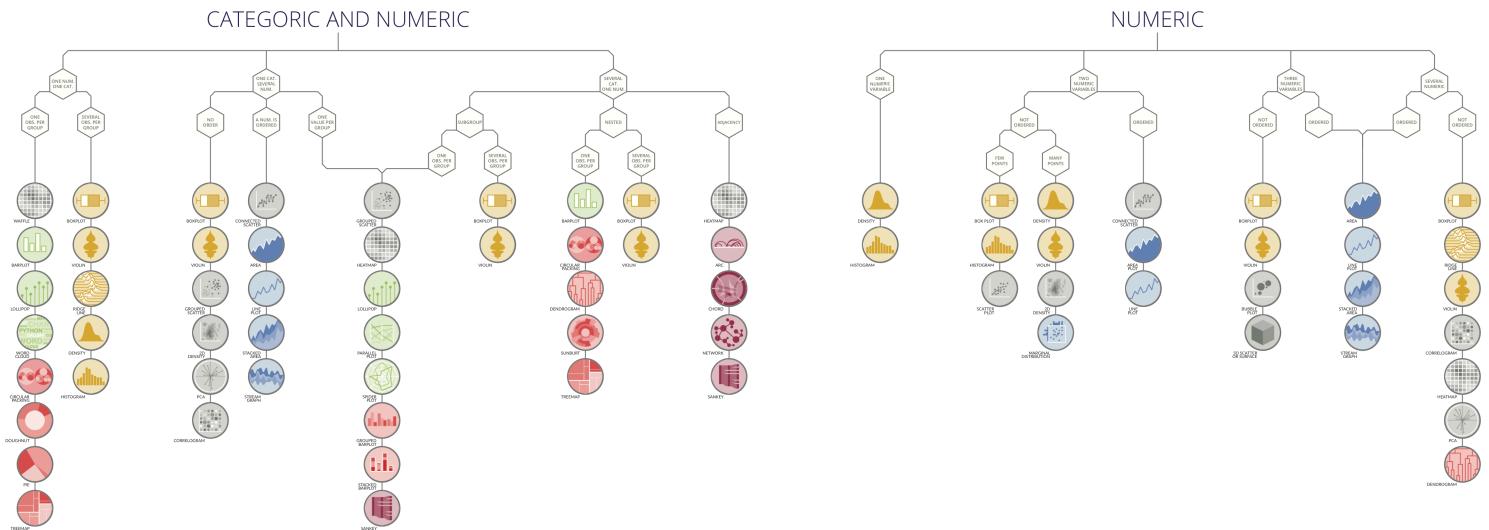
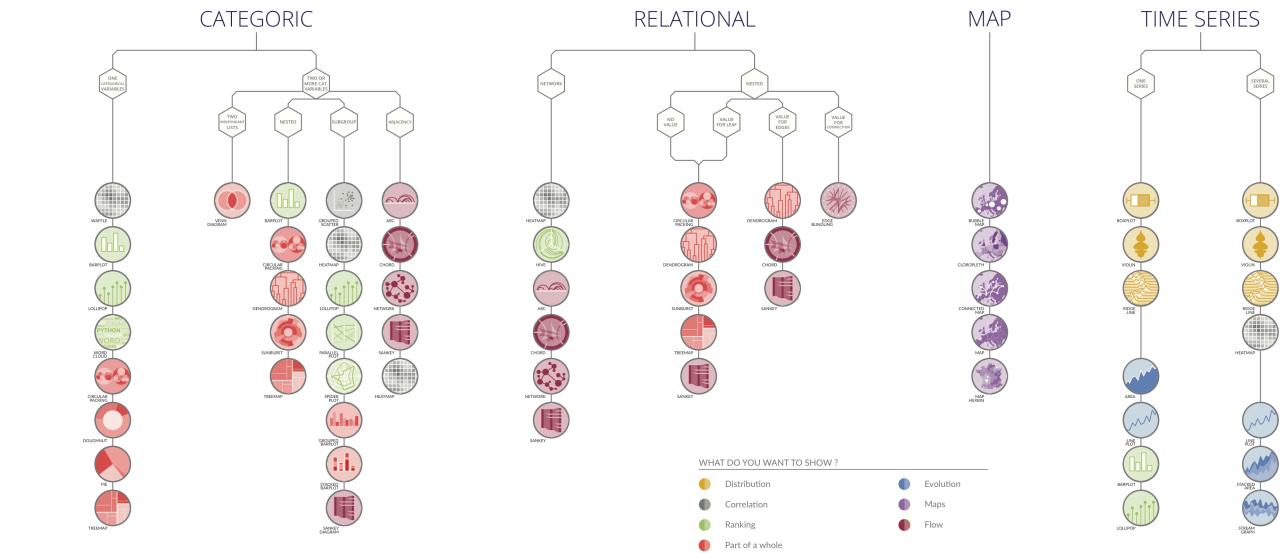
from Data to Viz

'From Data to Viz' is a classification of chart types based on input data format. It will help you find the perfect chart in three simple steps :

- 1 Identify what type of data you have.
 - 2 Go to the corresponding decision tree and follow it down to a set of possible charts.
 - 3 Choose the chart from the set that will suit your data and your needs best.

Dataviz is a world with endless possibilities and this project does not claim to be exhaustive. However it should provide you with a good starting point. For an interactive version and much more, visit:

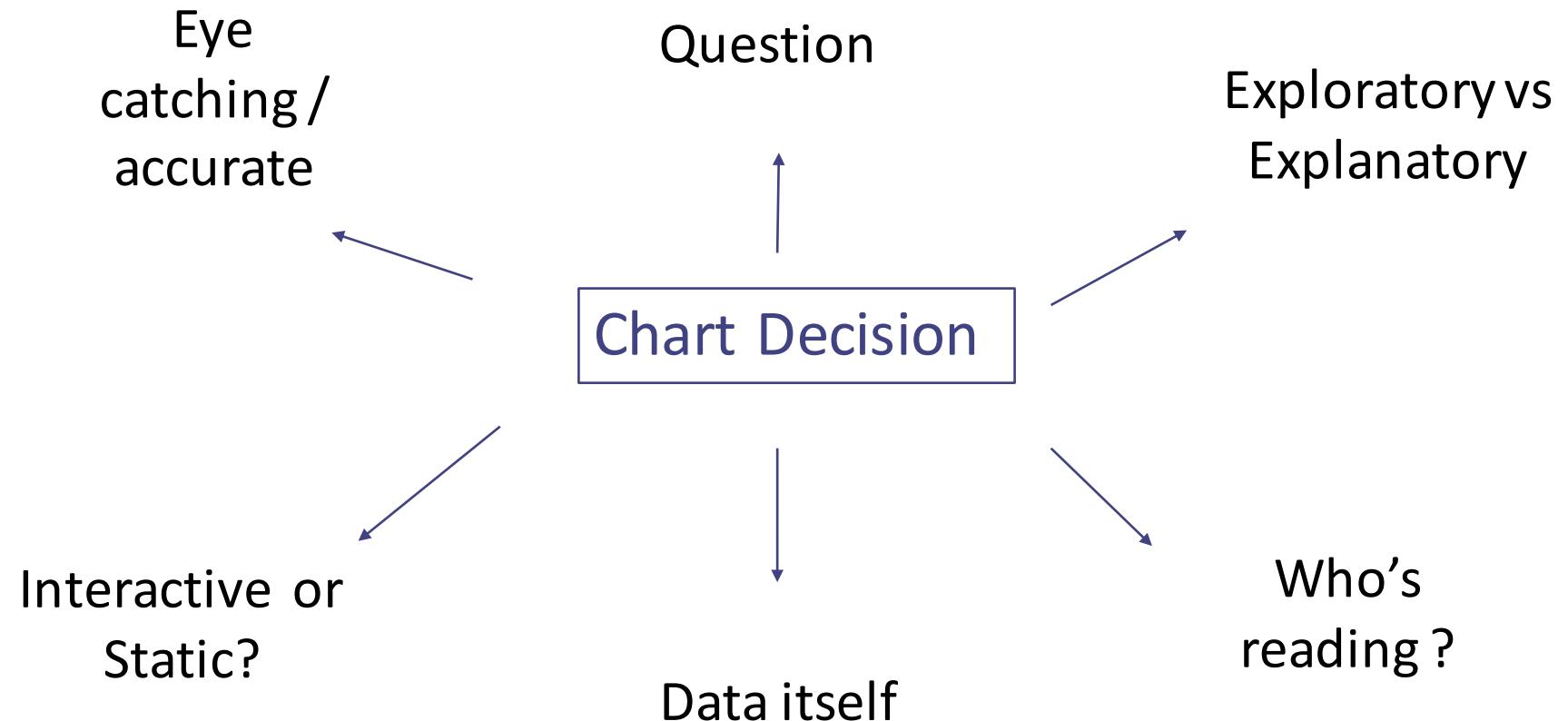
data-to-viz.com



WHAT YOU SHOULD DO

About 20 examples of storytelling with data

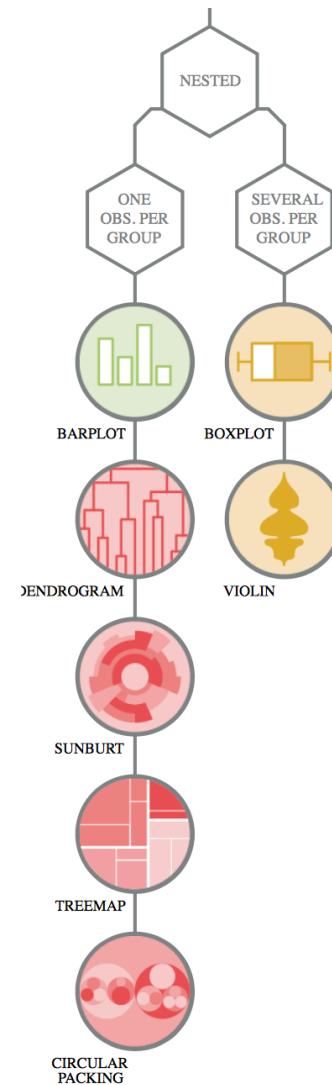
www.data-to-viz.com/#story



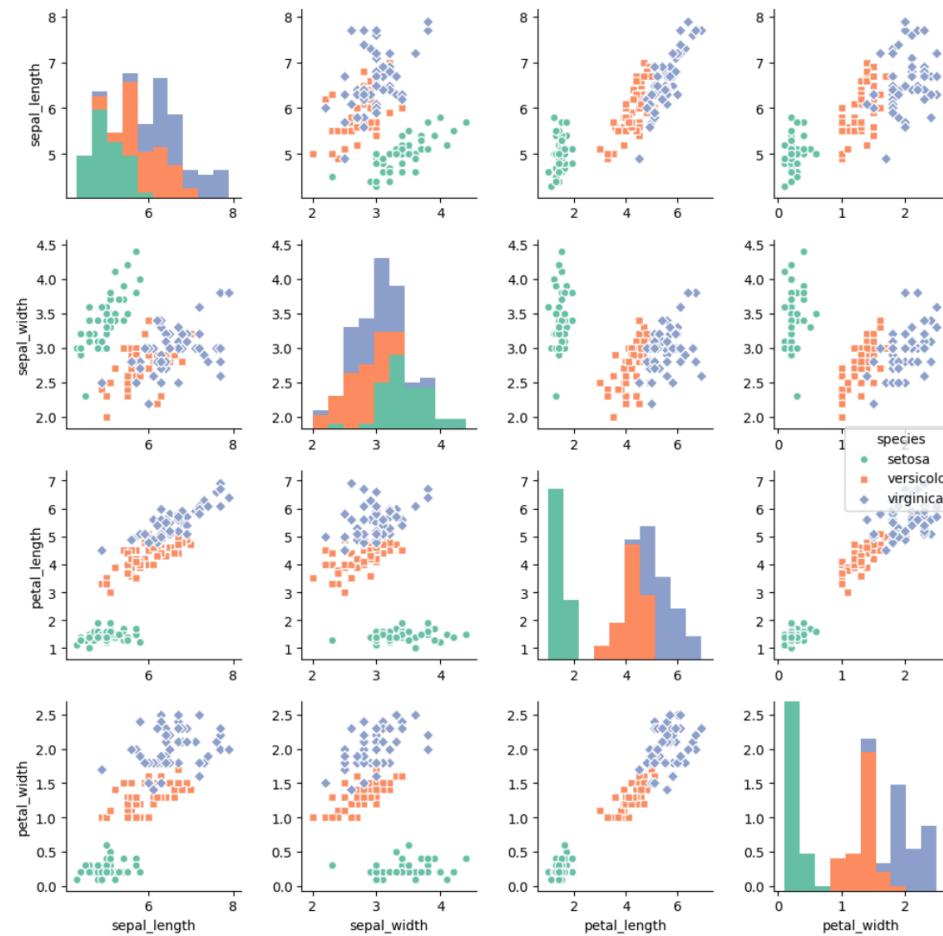
Question – Explo/Expl – Reader – Data – Interactivity – Eye catching

WHAT DO YOU WANT TO SHOW ?

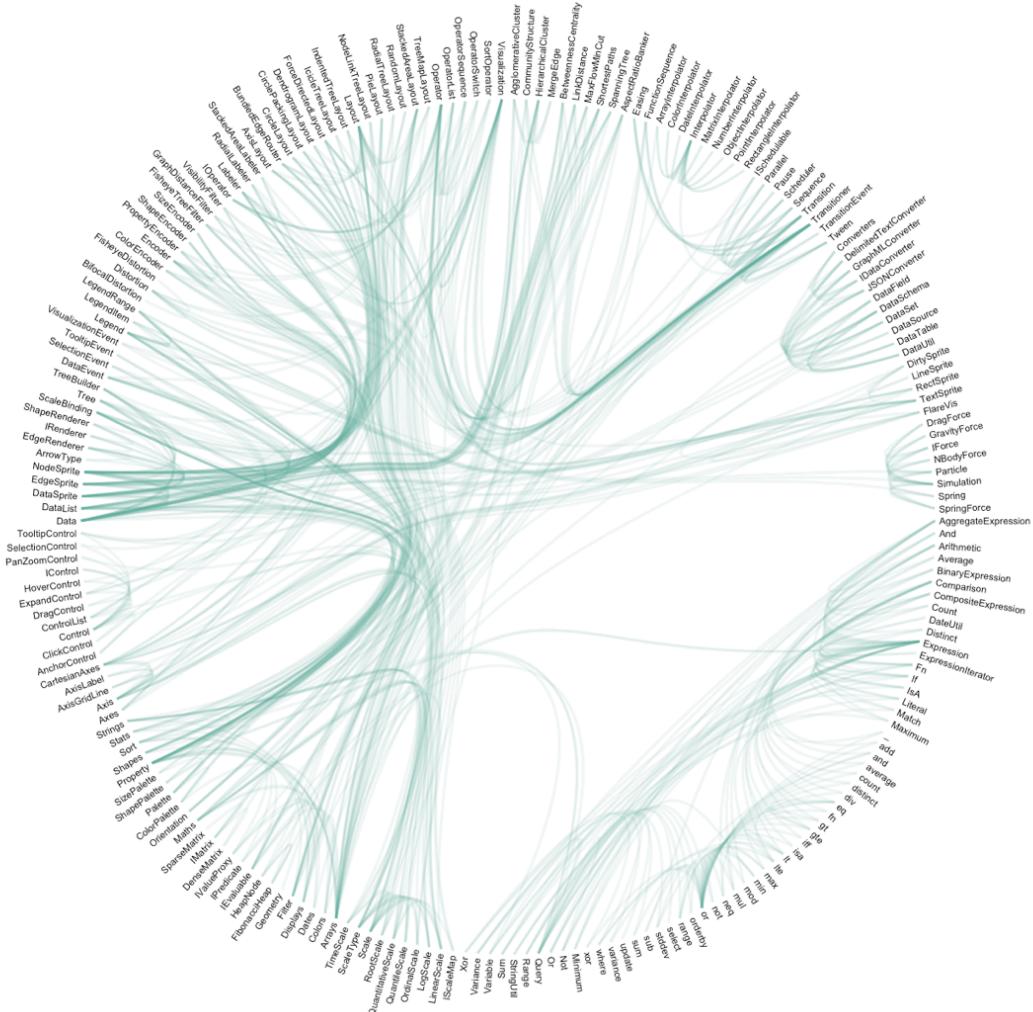
- Distribution
- Evolution
- Correlation
- Maps
- Ranking
- Flow
- Part of a whole



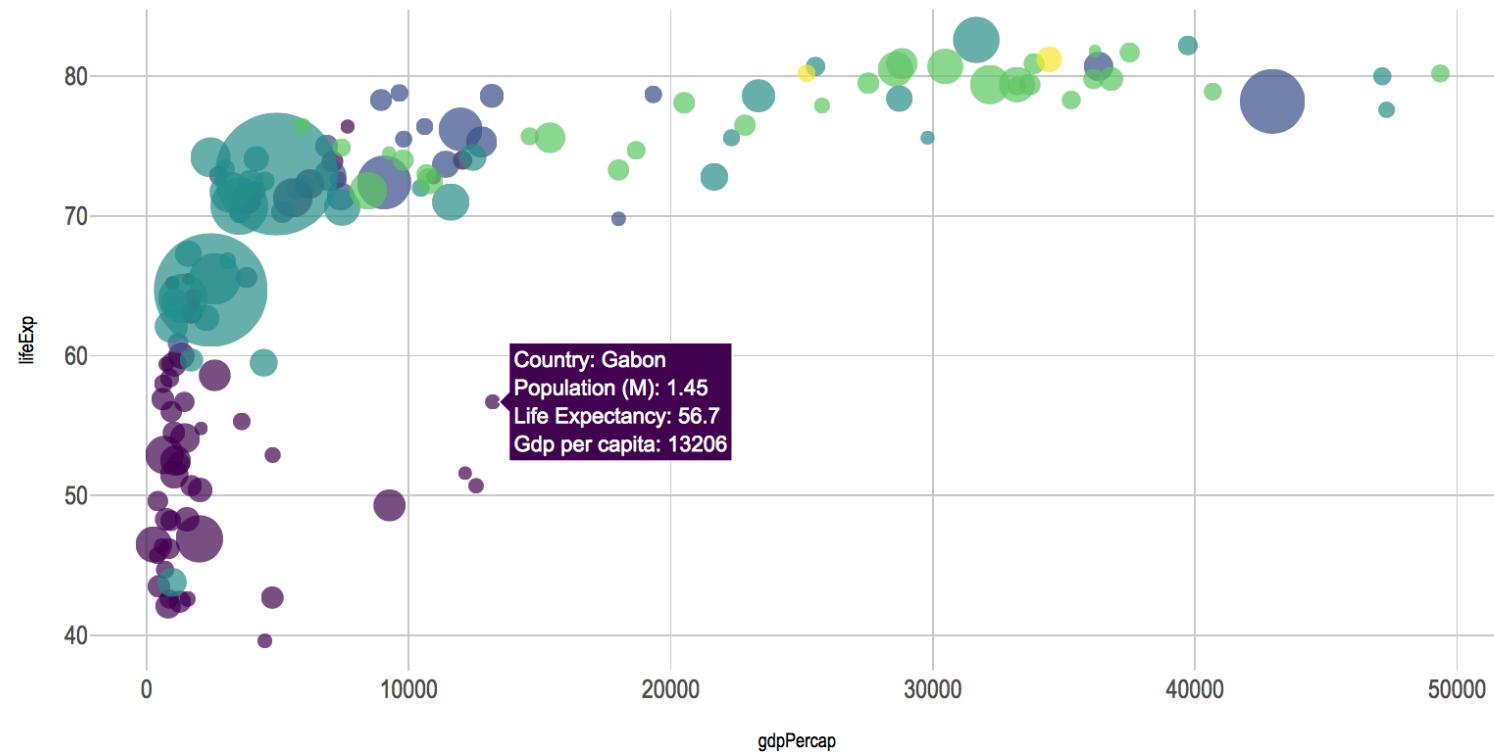
Question – **Exploratory** – Reader – Data – Interactivity – Eye catching



Question – Explo/Expla – Reader – Data – Interactivity – Eye catching

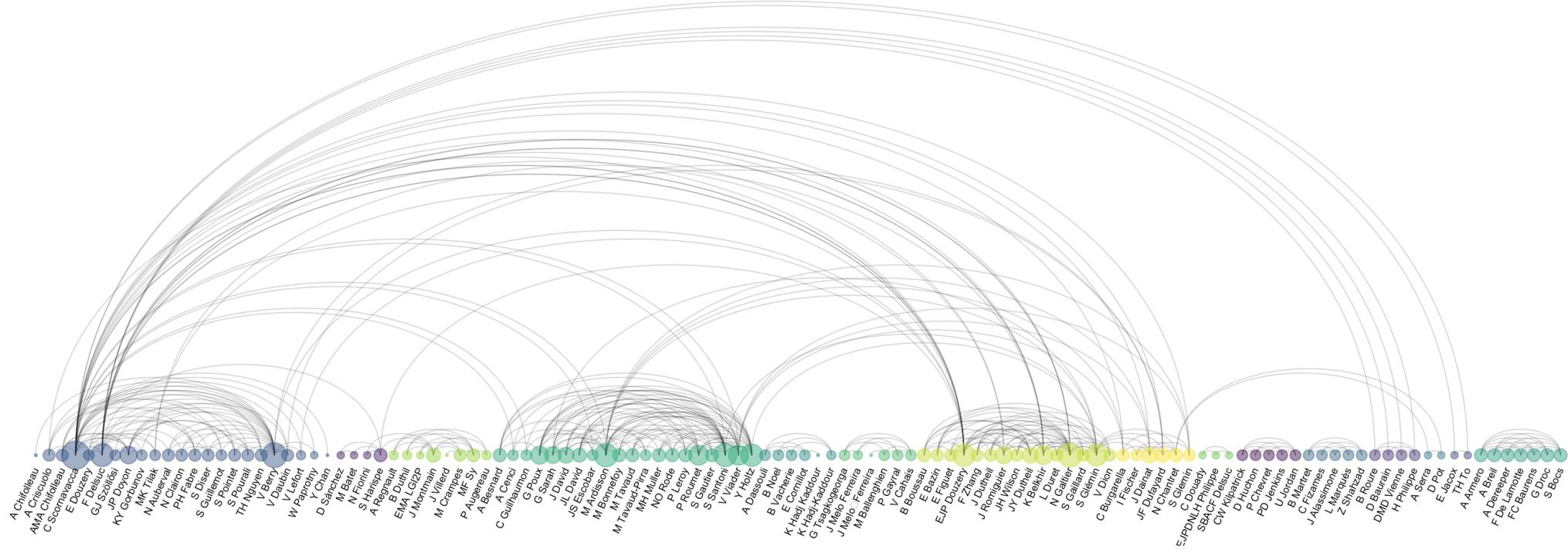


Question – Explo/Expla – Reader – Data – **Interactivity** – Eye catching



Question – Explo/Expla – Reader – Data – Interactivity – Eye catching





Co-authorship network of a researcher



WHAT YOU SHOULD NOT DO

A gallery of most common caveats

www.data-to-viz.com/caveats



Order your data

When displaying the value of several entities, ordering them makes the graph much more insightful.



To cut or not to cut?

Cutting the Y-axis is one of the most controversial practice in data viz. See why.



The spaghetti chart

A line graph with too many lines becomes unreadable: it is called a spaghetti graph.



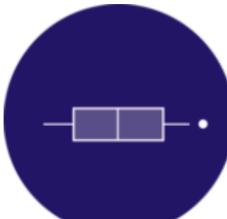
Pie chart

The human eye is bad at reading angles. See how to replace the most criticized chart ever.



Play with histogram bin size

Always try different bin sizes when you build a histogram, it can lead to different insights.



Do boxplots hide information?

Boxplots are a great way to summarize a distribution but hide the sample size and their distribution.



The problem with error bars

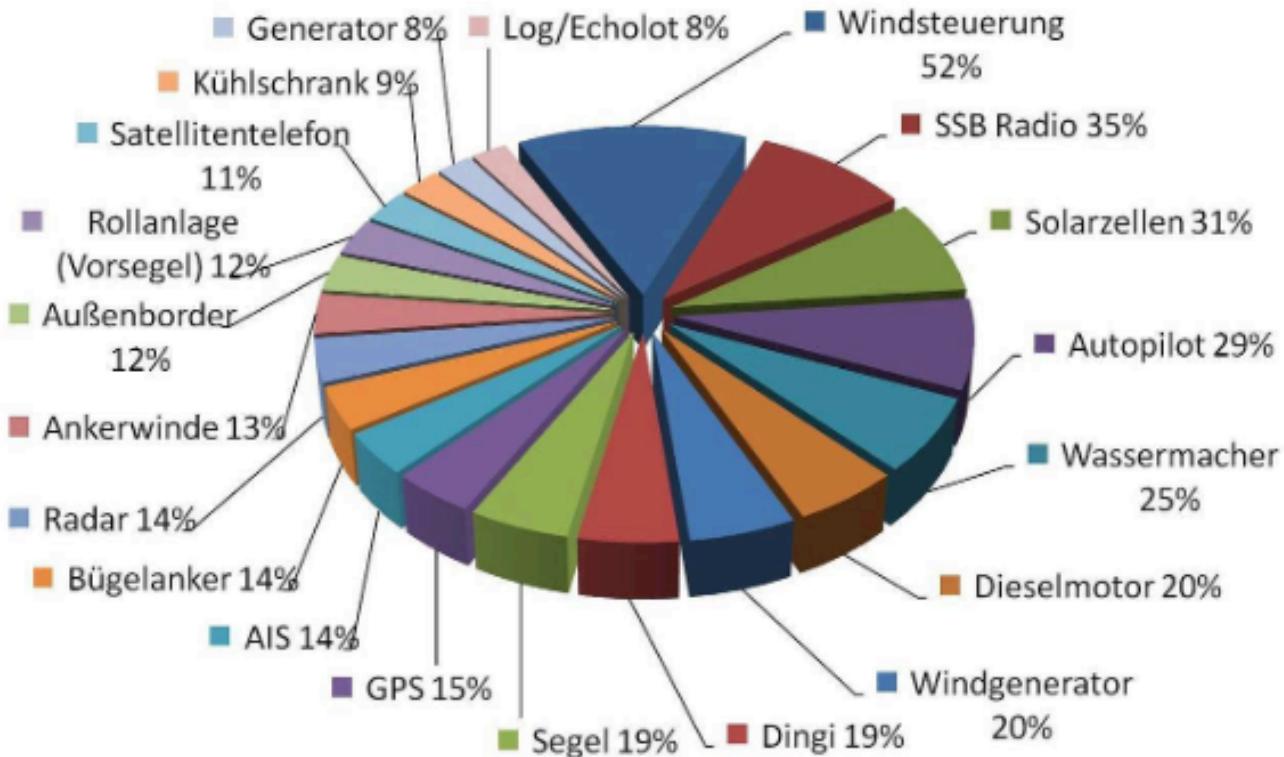
Barplots with error bars must be used with great care. See why and how to replace them.



Too many distributions.

If you need to compare the distributions of many variables, don't clutter your graphic.

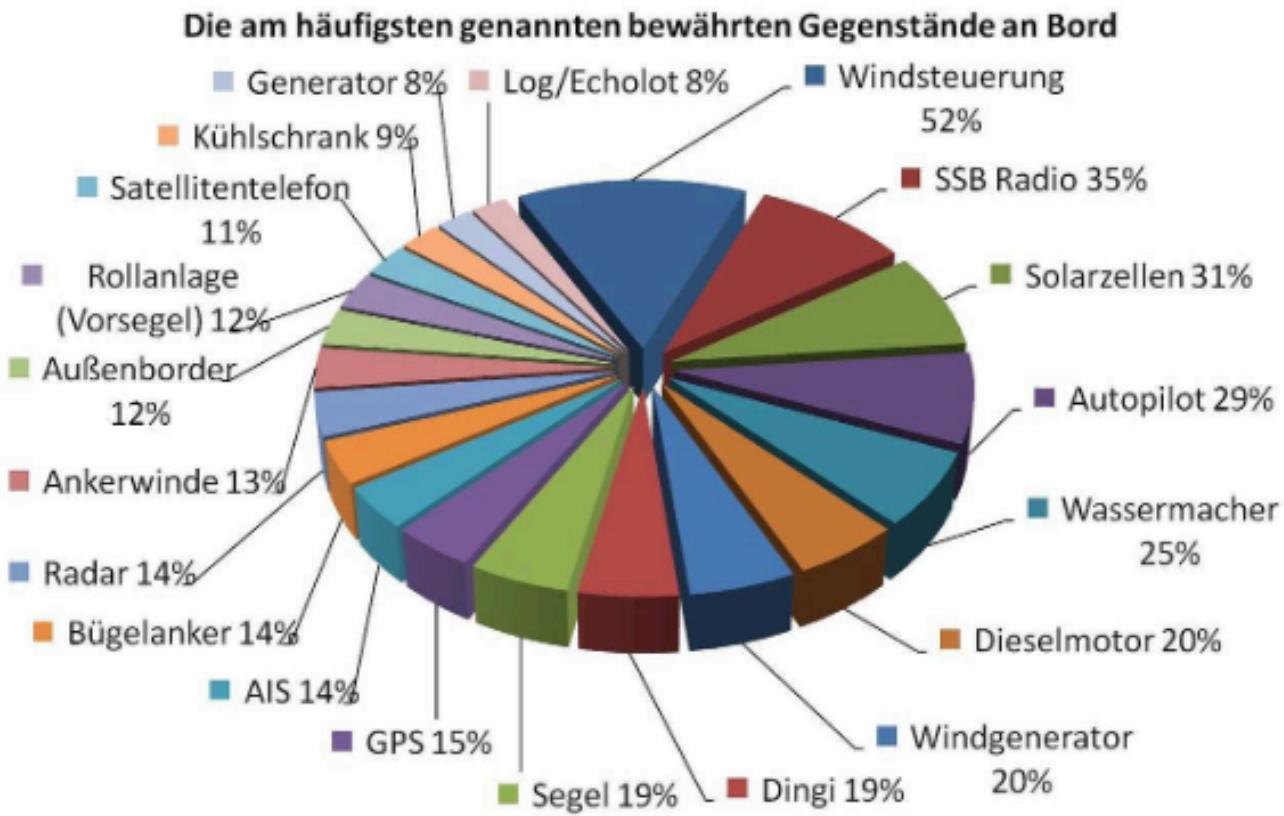
Die am häufigsten genannten bewährten Gegenstände an Bord



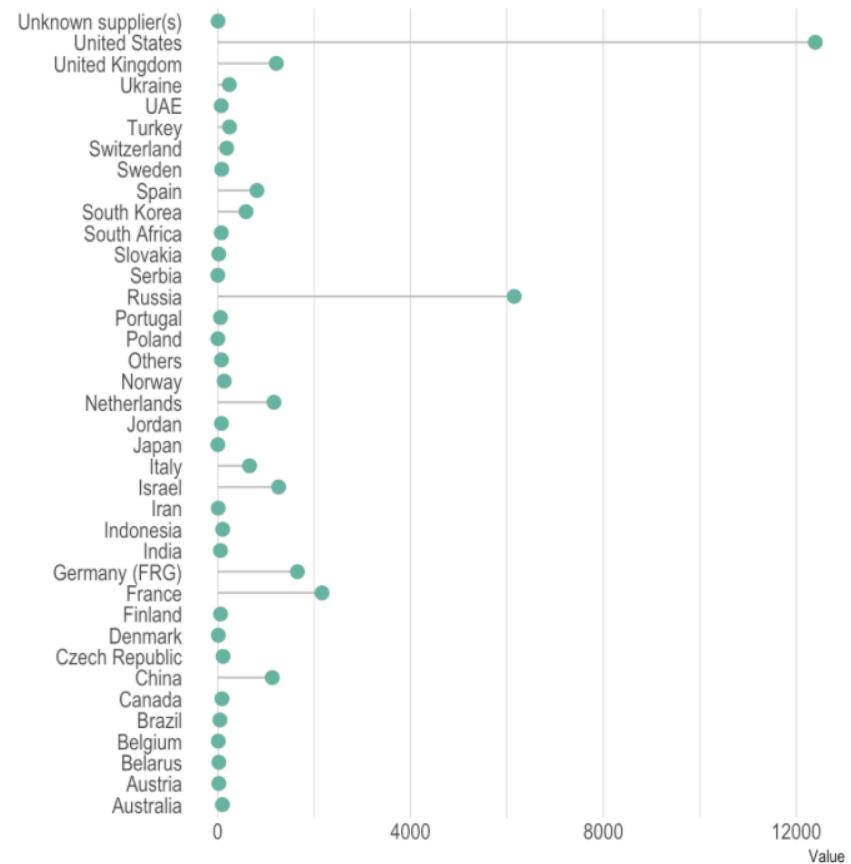
Source: [WTF Visualizations](#)

Calculation

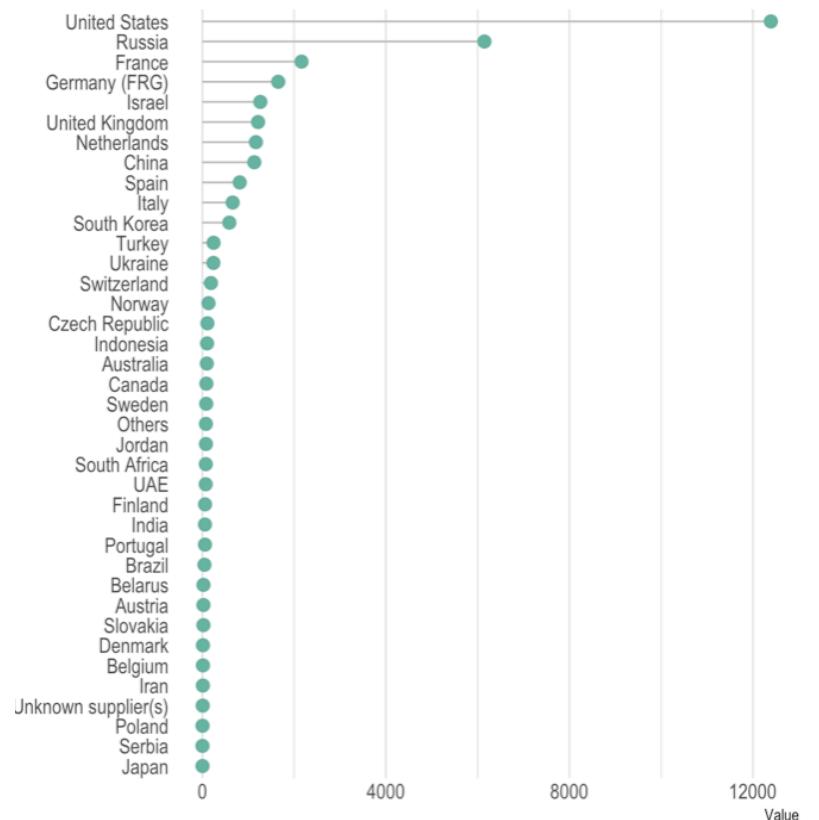
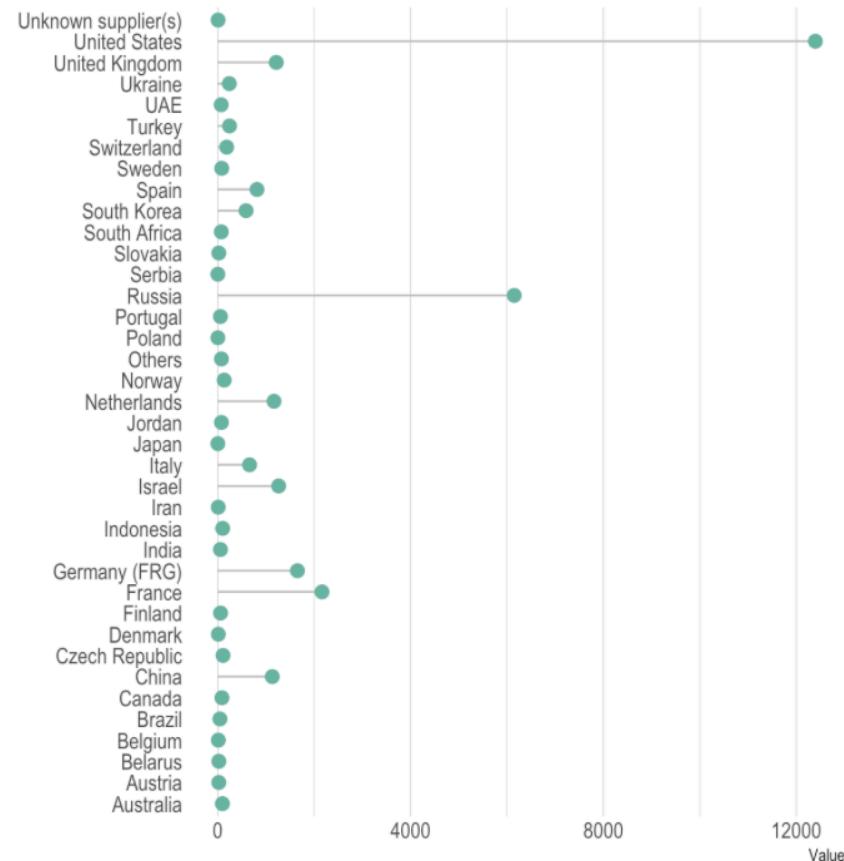
+
3D
+
Pie chart

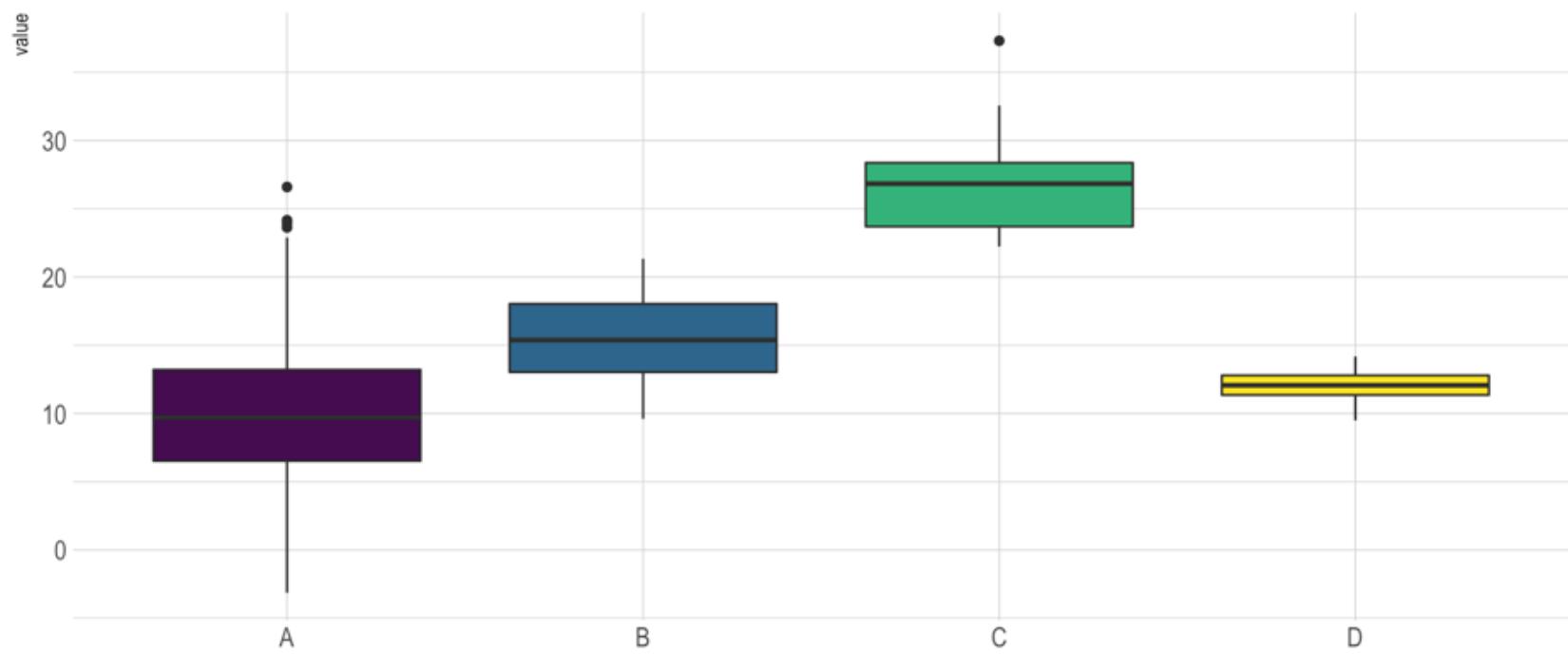


Source: [WTF Visualizations](#)



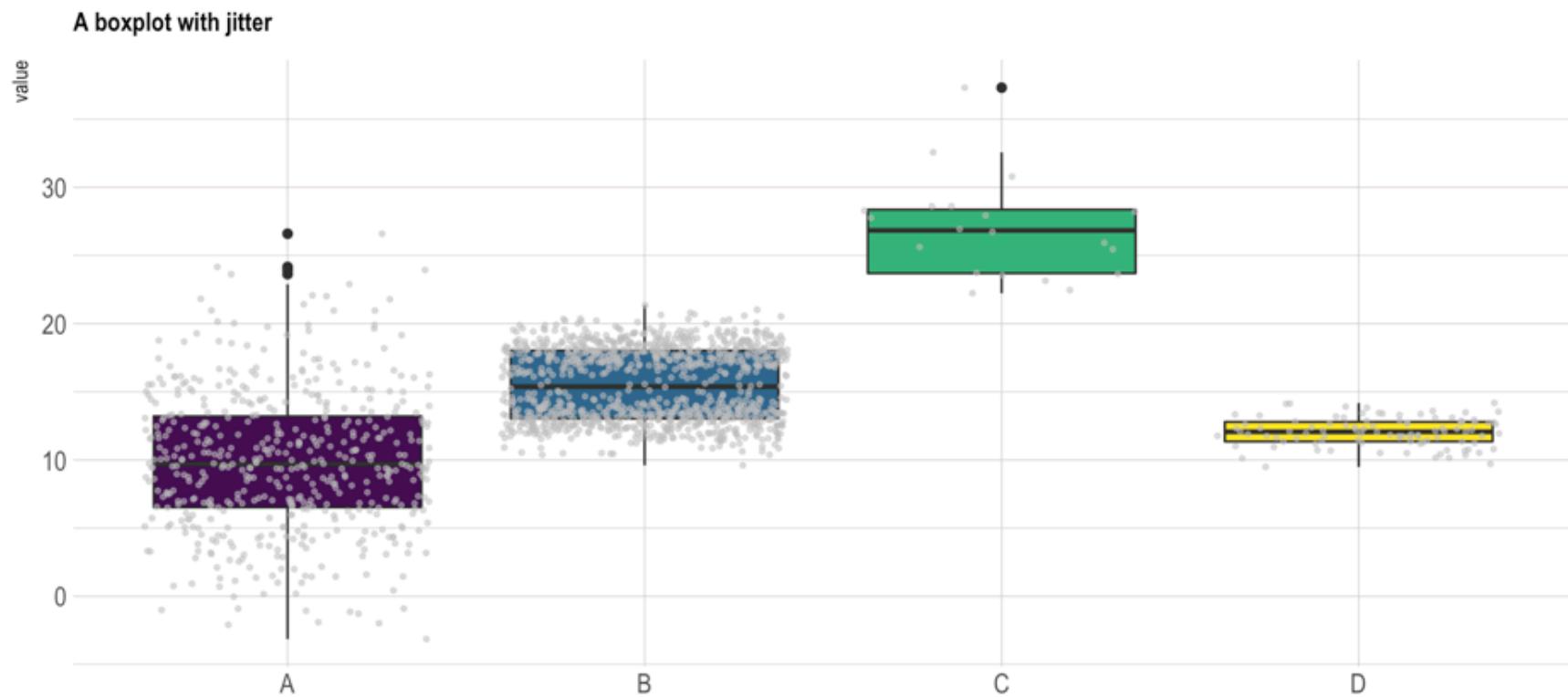
Order your data



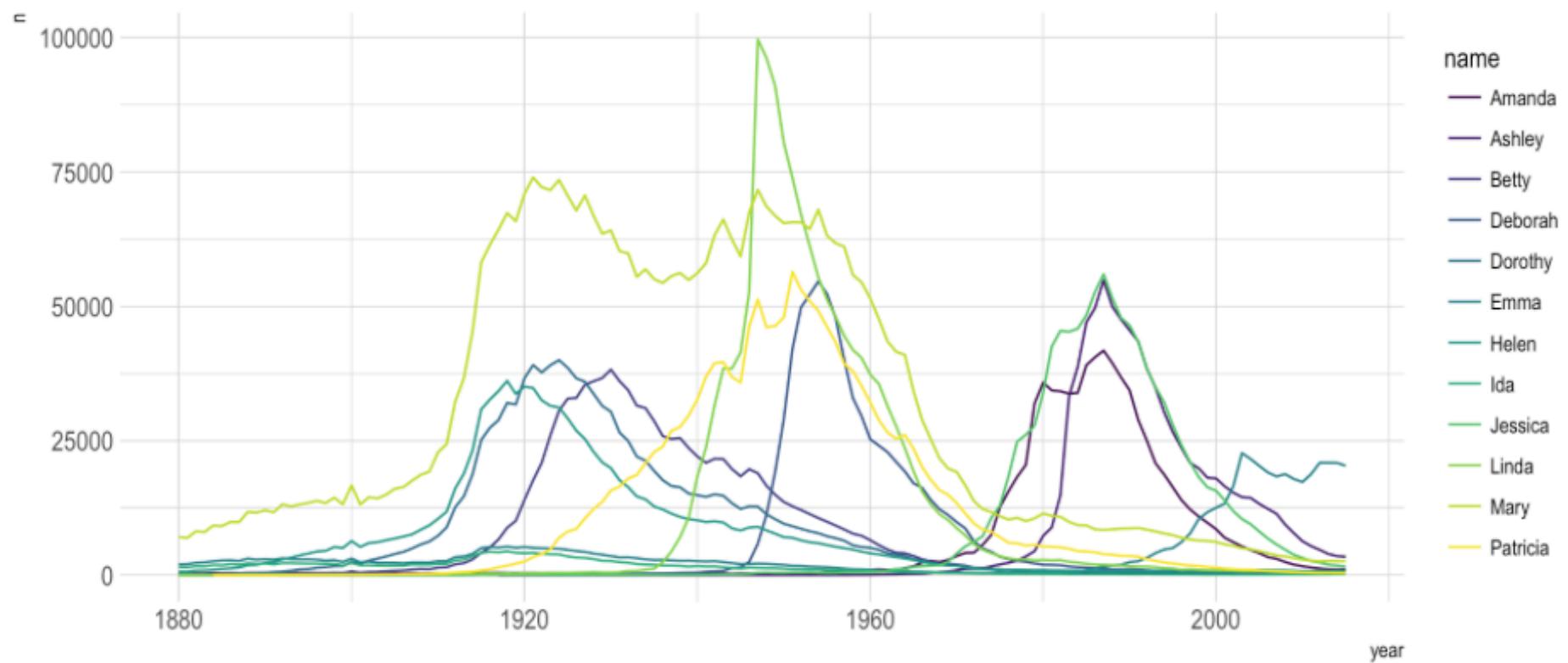


The issue with boxplot

[Read more](#)

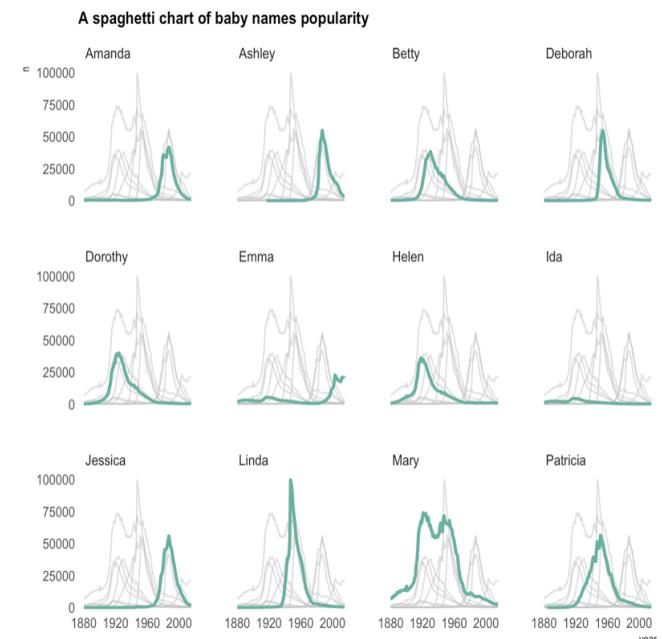
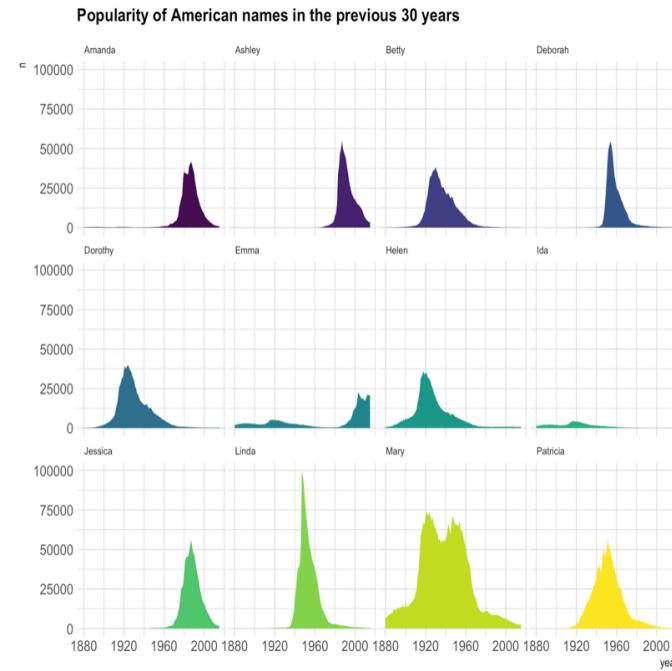
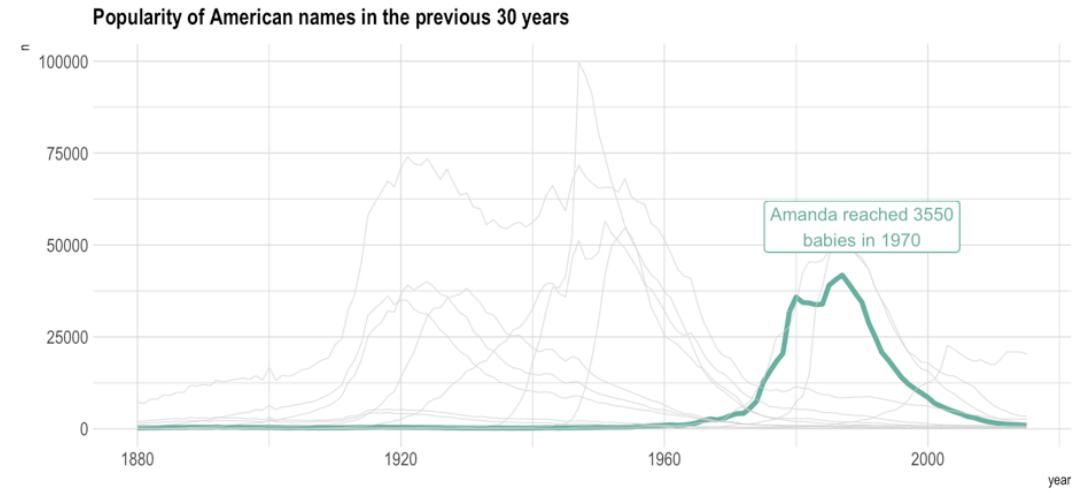


baby names popularity

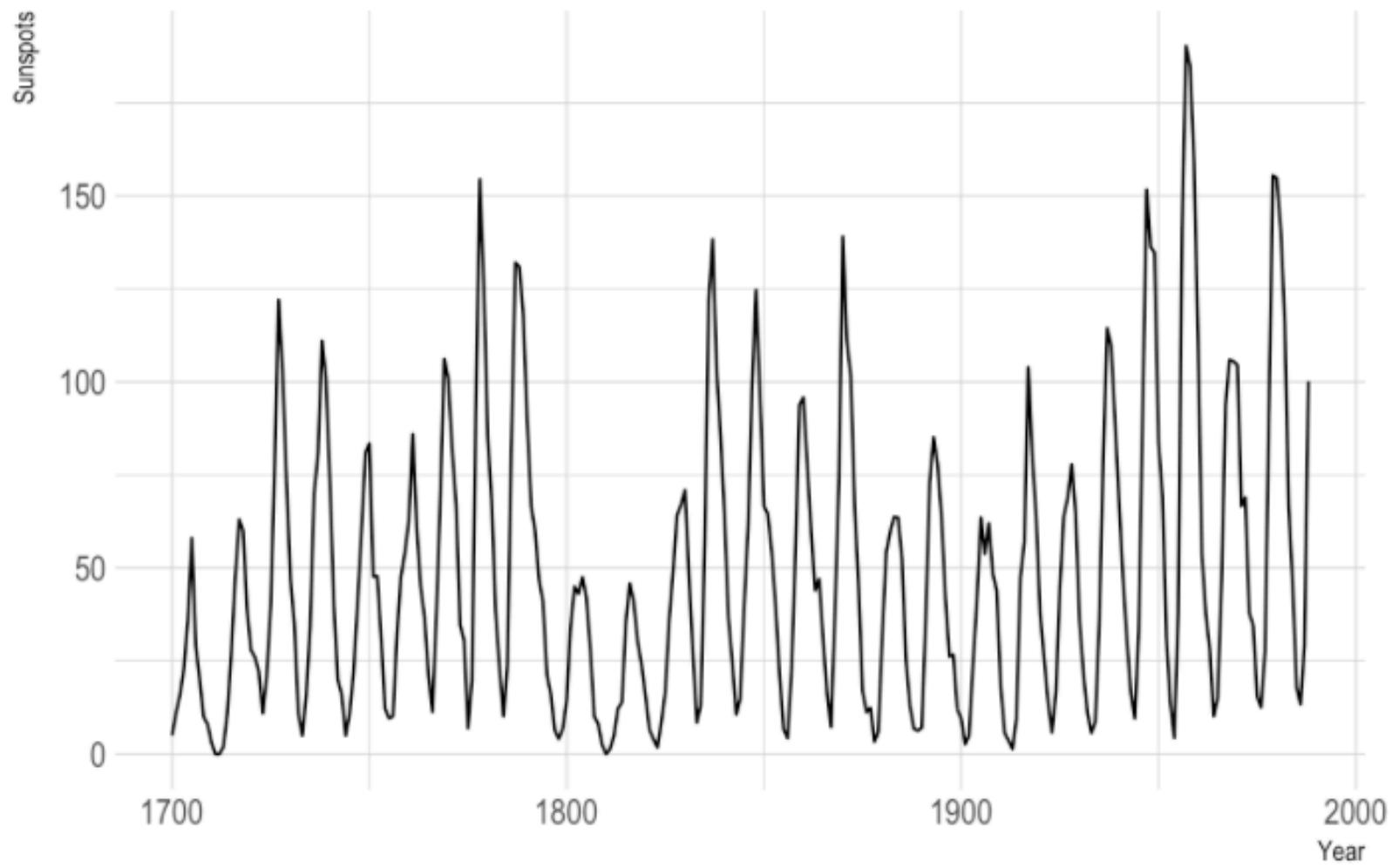


The Spaghetti chart

[Read more](#)

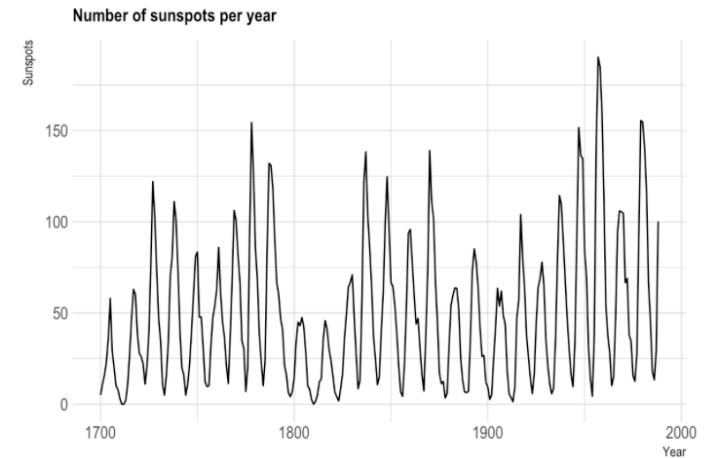


Number of sunspots per year

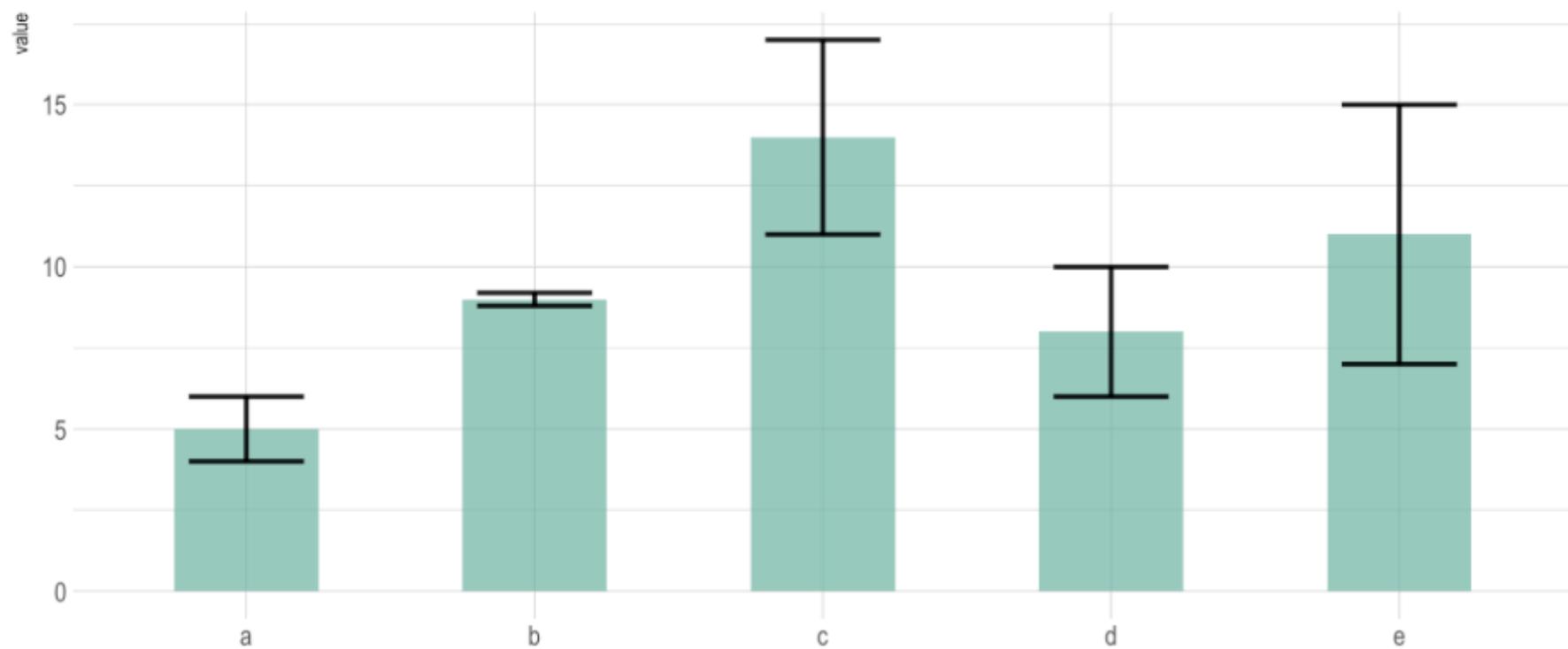


Mind the aspect ratio

[Read more](#)

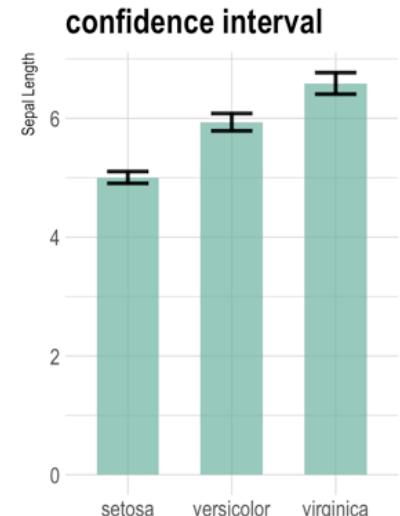
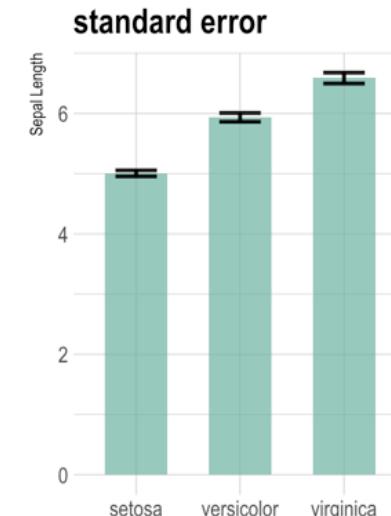
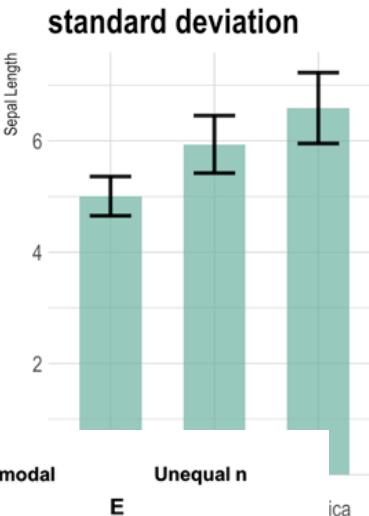
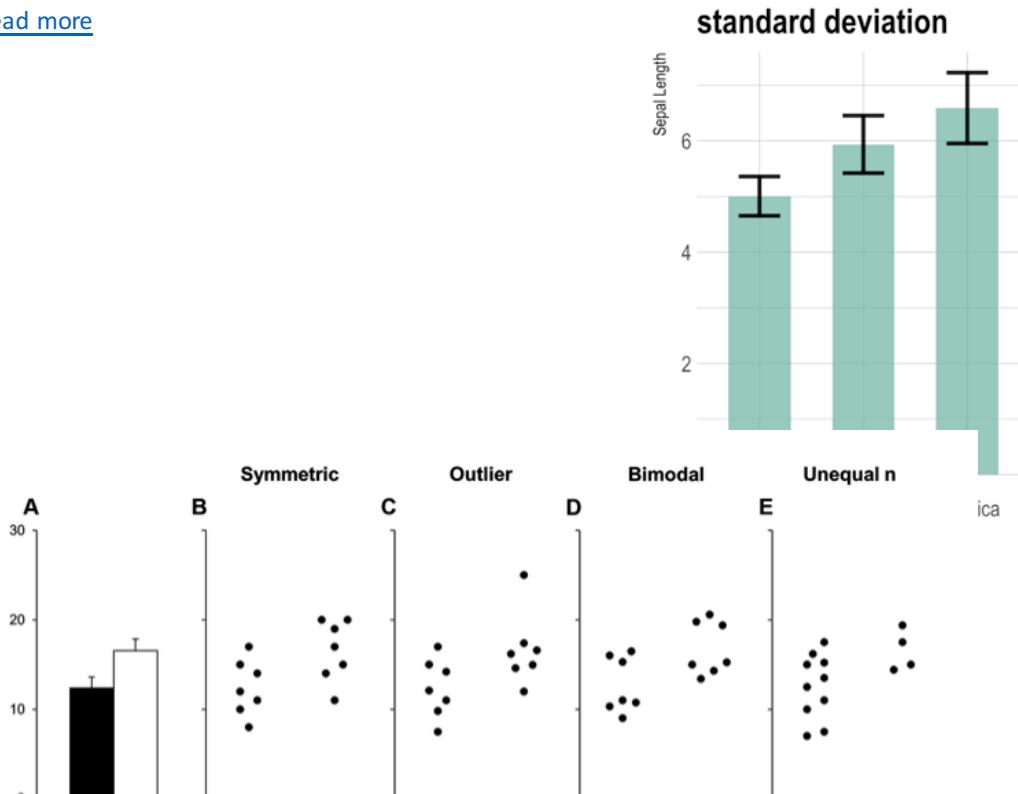


A barplot with error bar



The issue with error bars

[Read more](#)



Test	p value			
T-test: Equal var.	0.035	0.050	0.026	0.063
T-test: Unequal var.	0.035	0.050	0.026	0.035
Wilcoxon	0.054	0.073	0.128	0.103

Weissgerber et al. 2015

HOW TO DO IT

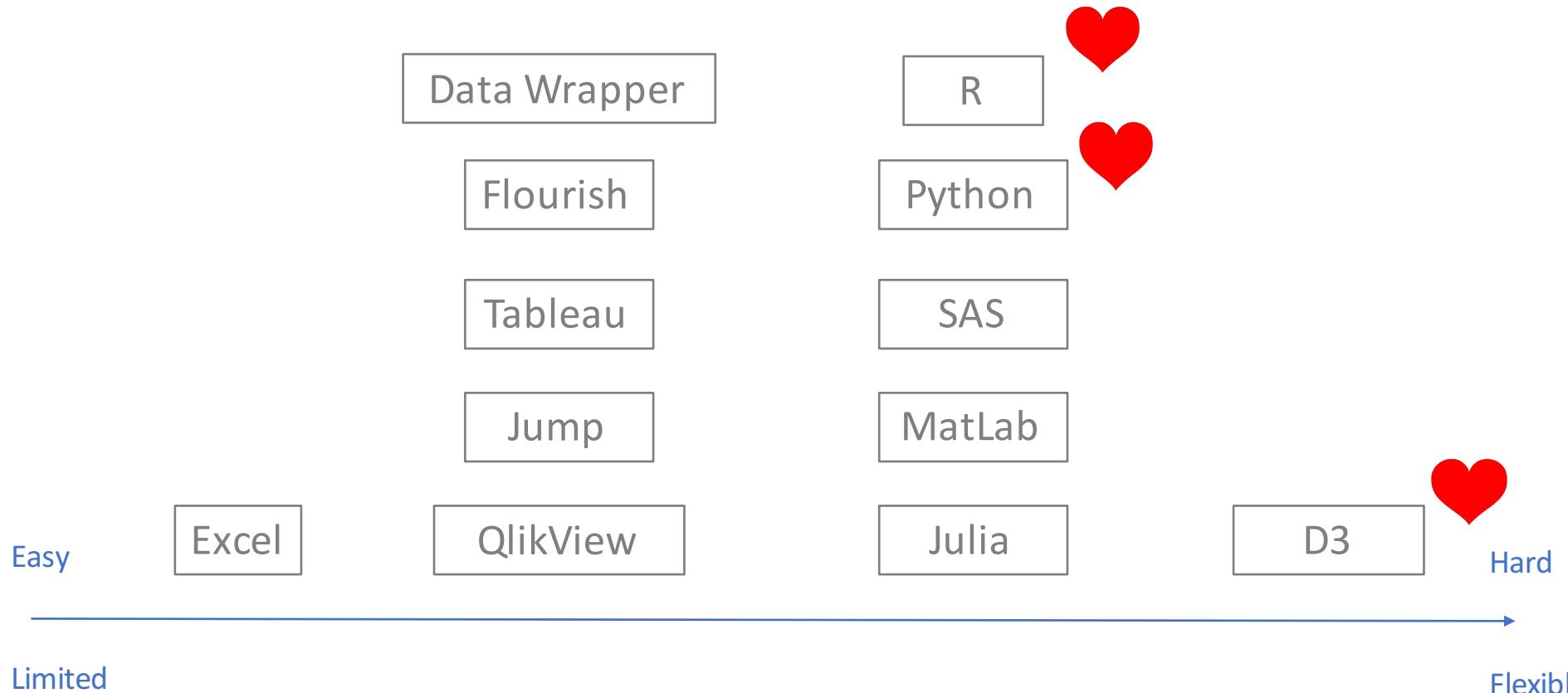
The R and Python graph galleries

See the [chart maker directory](#) by visualisingdata.com



** Far from being exhaustive*

See the [chart maker directory](#) by visualisingdata.com



* Far from being exhaustive



from Data to Viz



[**R-graph-gallery.com**](#)

[**Python-graph-gallery.com**](#)

[**D3-graph-gallery.com**](#)



Data-to-viz.com



@R_Graph_Gallery



github.com/holtzy/Talk



Yan.holtz.data@gmail.com



www.yan-holtz.com

Data-to-viz.com



@R_Graph_Gallery



github.com/holtzy/Talk



Yan.holtz.data@gmail.com



www.yan-holtz.com



KANTAR
Information is Beautiful
Awards

