Truth, beauty, and data

Session I

Data visualization workshop
Andrew Young School of Policy Studies
2024

Facts, truth, and beauty

What is truth?

Core principles of the universe?

Underlying trends in society?

Something transcendental?

Reality?

How do we find truth?

Science!

Neil deGrasse Tyson ✔
@neiltyson

The good thing about Science is that it's true whether or not you believe in it.

10:41 AM · Jun 14, 2013 · TweetDeck

14.3K Retweets   **8.3K** Likes

💬 761   ⟲ 14.3K   ♡ 8.3K   ↑

But wait!

Beware of scientism!

"... promotion of science as the best or only objective means by which society should determine normative and epistemological values"

Science is not the only way

Art

Music

Literature

Religion

Nature

Nothing here is factual...

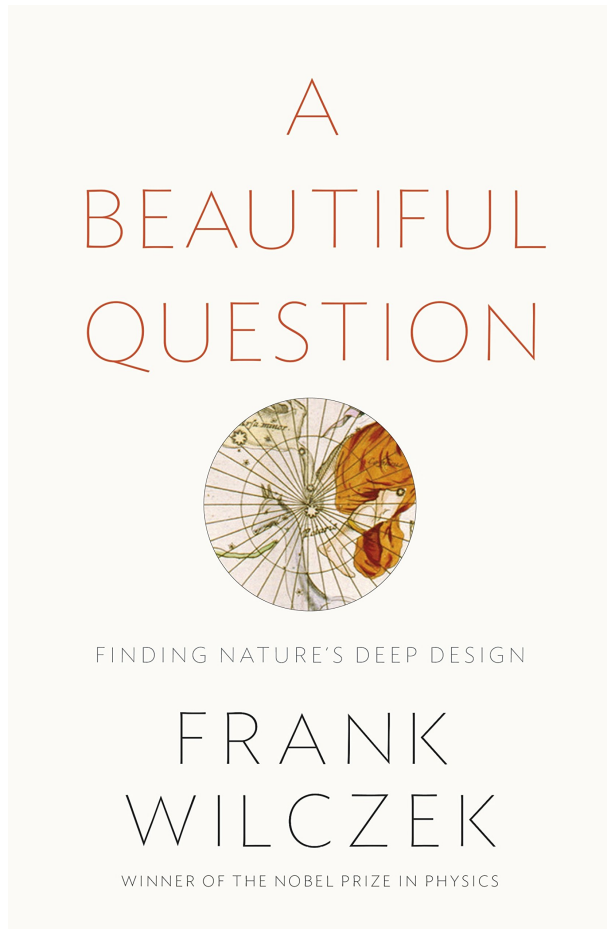## ...but it all reveals truth

Cosette

King Lear

Beethoven's 9th symphony

Facts ≠ truth

Where does truth come when there are no facts?

Beauty

Beauty in science

This is also true for science and math and other more factual realms

Rhetoric and beauty

λόγος • λέξις

Logos • Lexis

Res • Verba

Essence • Structure

Content • Form

Truth • Beauty

Content + form

Art is how we translate core, essential **content** (or truth!) to different **forms** for specific **audiences**.

Truth is beautiful

Truth ≠ facts

Truth comes from aesthetic combination of content and form

Facts require beauty to be true

Data, truth, and beauty

Just show me the data!

```
head(my_data, 10)
```

```
## # A tibble: 10 × 2
##        x      y
##    <dbl>  <dbl>
##  1  55.4   97.2
##  2  51.5   96.0
##  3  46.2   94.5
##  4  42.8   91.4
##  5  40.8   88.3
##  6  38.7   84.9
##  7  35.6   79.9
##  8  33.1   77.6
##  9  29.0   74.5
## 10  26.2   71.4
```

```
mean(my_data$x)
```

```
## [1] 54.26327
```

```
mean(my_data$y)
```

```
## [1] 47.83225
```
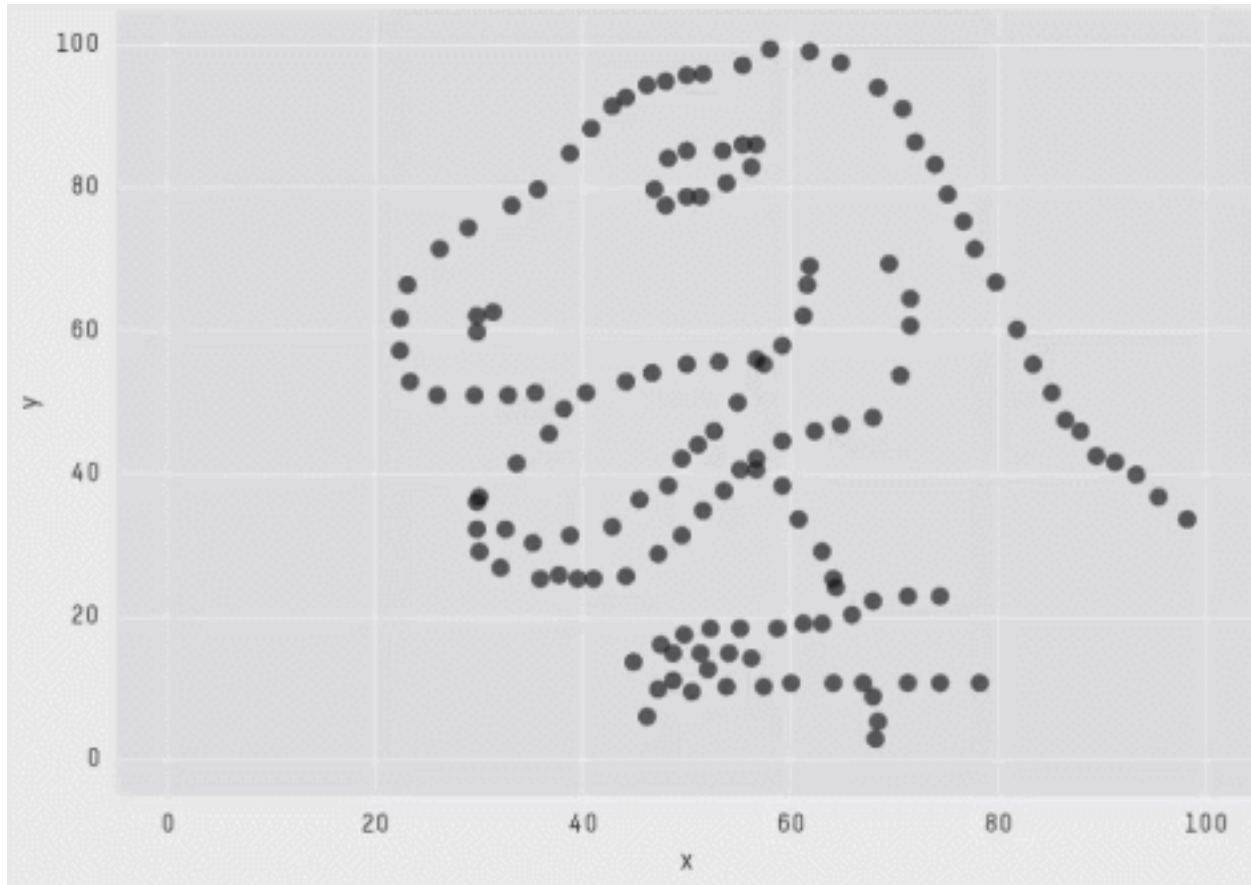
```
cor(my_data$x, my_data$y)
```

```
## [1] -0.06447185
```

Seems reasonable

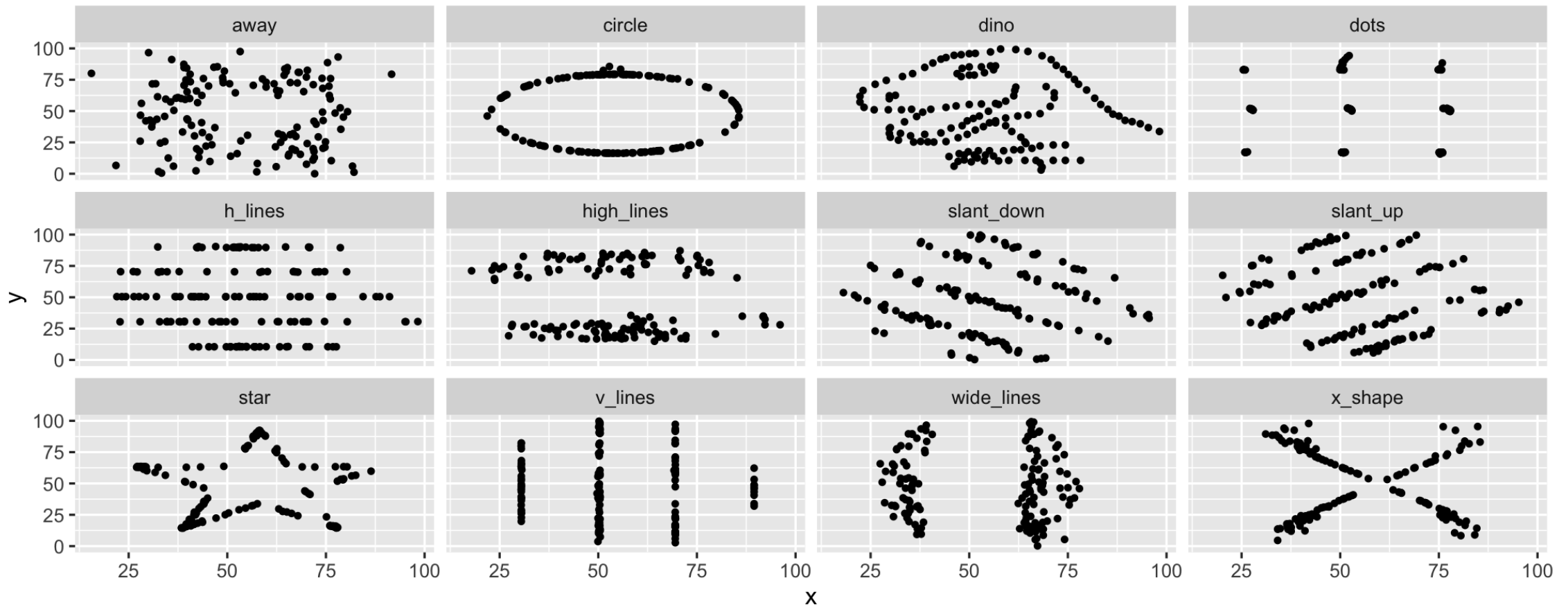**Seems reasonable**

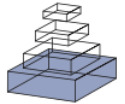**No correlation**

oh no

The Datasaurus Dozen

Raw data is not enough

Each of these has the same mean, standard deviation, variance, and correlation

Humans love patterns

Superior pattern processing is the essence of the evolved human brain

Mark P. Mattson [1,2]*

[1] Laboratory of Neurosciences, National Institute on Aging Intramural Research Program, Baltimore, MD, USA
[2] Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Humans have long pondered the nature of their mind/brain and, particularly why its capacities for reasoning, communication and abstract thought are far superior to other species, including closely related anthropoids. This article considers superior pattern

https://doi.org/10.3389/fnins.2014.00265
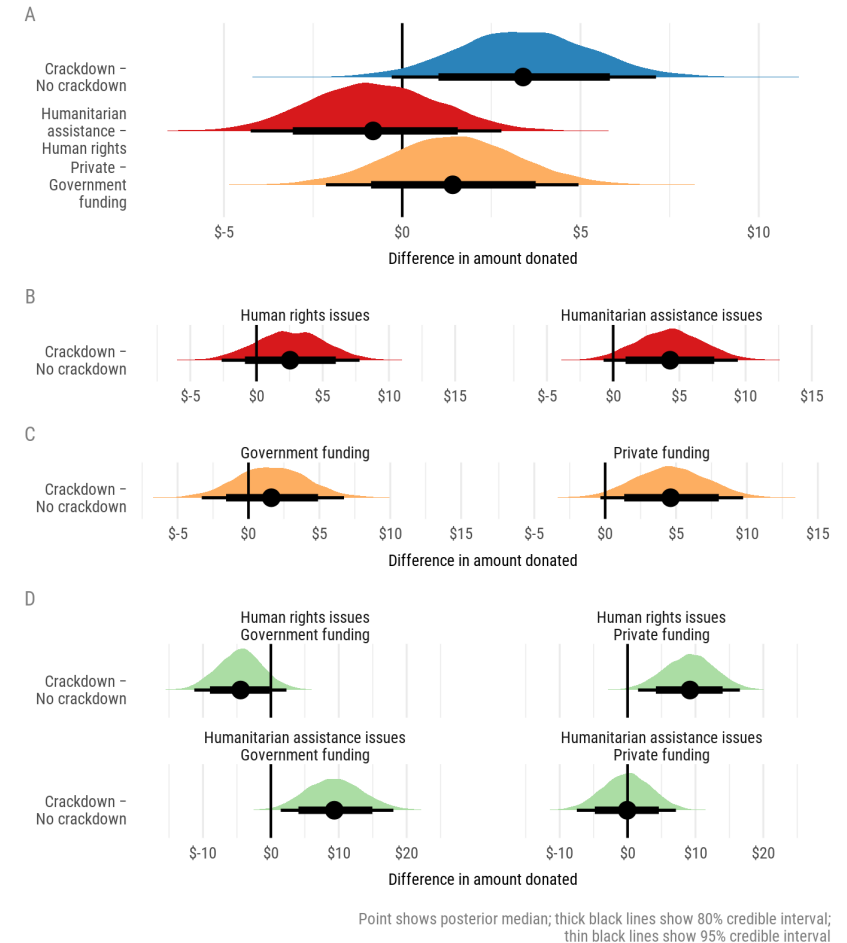
(Sometimes we love them too much)

Pareidolia: seeing patterns that aren't there.
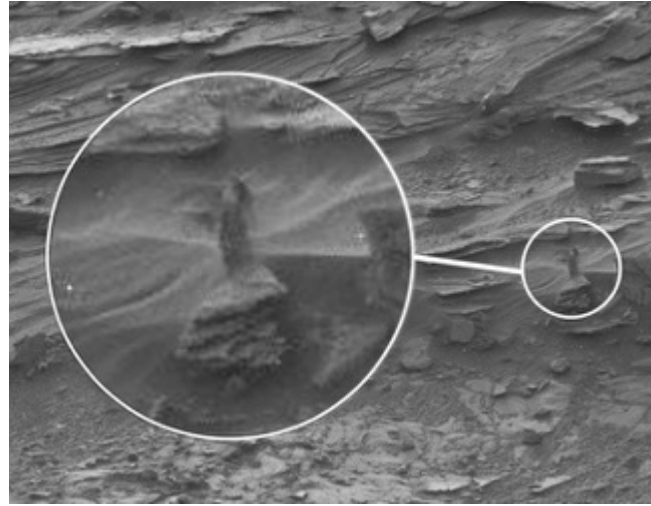
Beauty is necessary to see patterns

Table 2: Mean values and differences in means for amount donated in "crackdown" (treatment) and "no crackdown" (control) conditions; values represent posterior medians

H_{1b}	Amount$_{Treatment}$	Amount$_{Control}$	Δ	$\%\Delta$	$p(\Delta \neq 0)$
Crackdown – No crackdown	16.34	12.93	3.39	26.3%	0.97
Humanitarian assistance – Human rights	14.06	14.85	-0.82	-5.5%	0.67
Private – Government funding	15.13	13.71	1.42	10.4%	0.79
H_{2b} and H_{3b}	Amount$_{Crackdown}$	Amount$_{No\ crackdown}$	Δ	$\%\Delta$	$p(\Delta \neq 0)$
Human rights issues	17.4	14.86	2.54	17.2%	0.83
Humanitarian assistance issues	15.91	11.68	4.3	36.9%	0.95
Government funding	13.83	12.24	1.61	13.1%	0.74
Private funding	18.95	14.23	4.62	32.4%	0.97
H_{2b} and H_{3b} (nested)	Amount$_{Crackdown}$	Amount$_{No\ crackdown}$	Δ	$\%\Delta$	$p(\Delta \neq 0)$
Human rights issues, Government funding	10.56	15.15	-4.46	-29.5%	0.91
Human rights issues, Private funding	23.76	14.5	9.19	63.8%	0.99
Humanitarian assistance issues, Government funding	21.42	11.89	9.35	77.9%	0.99
Humanitarian assistance issues, Private funding	15.69	15.72	-0.05	-0.3%	0.51

A

Crackdown – No crackdown

Humanitarian assistance – Human rights

Private – Government funding

Difference in amount donated

B

Human rights issues — Humanitarian assistance issues

Crackdown – No crackdown

C

Government funding — Private funding

Crackdown – No crackdown

Difference in amount donated

D

Human rights issues Government funding — Human rights issues Private funding

Crackdown – No crackdown

Humanitarian assistance issues Government funding — Humanitarian assistance issues Private funding

Crackdown – No crackdown

Difference in amount donated

Point shows posterior median; thick black lines show 80% credible interval; thin black lines show 95% credible interval

Beautiful visualizations

What makes a great visualization?

Truthful

Functional

Beautiful

Insightful

Enlightening

Alberto Cairo, *The Truthful Art*

What makes a great visualization?

"Graphical excellence is the **well-designed presentation of interesting data**—a matter of substance, of statistics, and of design ... [It] consists of complex ideas communicated with clarity, precision, and efficiency. ... [It] is that which **gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space** ... [It] is nearly always multivariate ... And graphical excellence requires **telling the truth about the data**."

Edward Tufte, *The Visual Display of Quantitative Information*, p. 51

What makes a great visualization?

Good aesthetics

No substantive issues
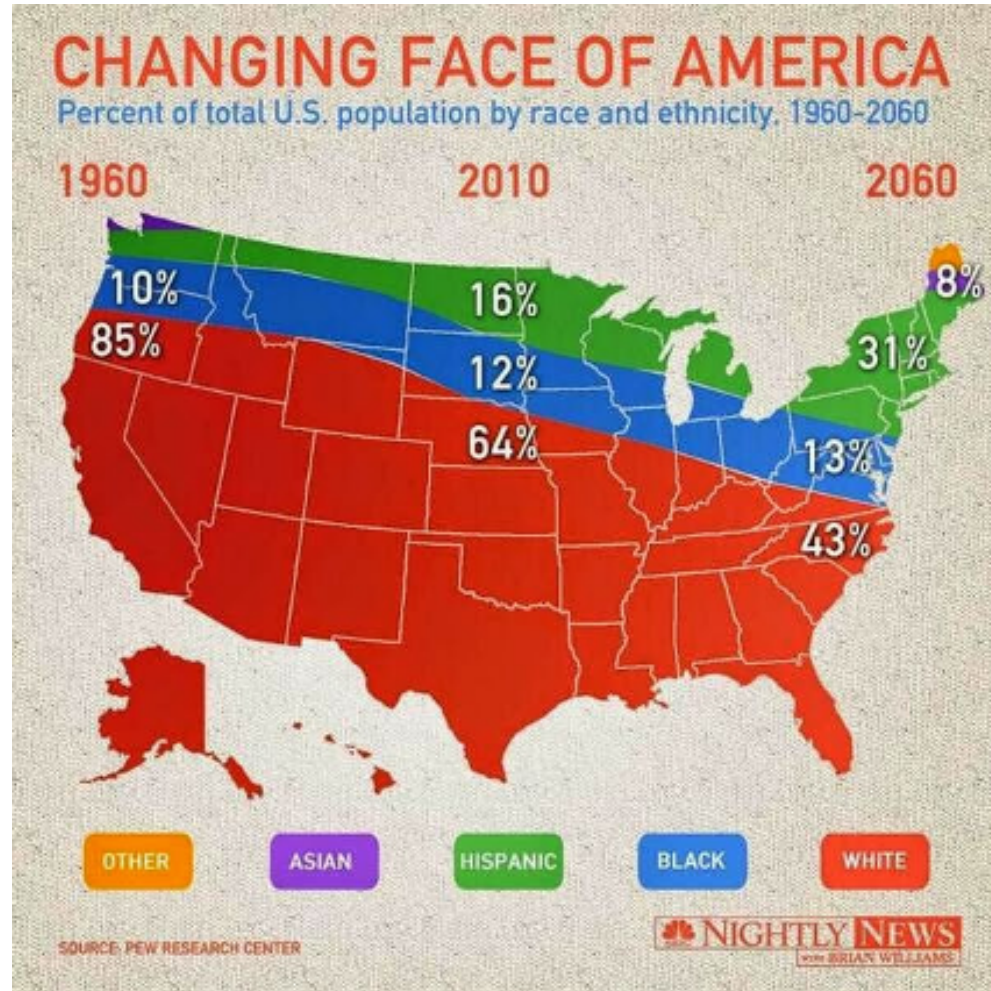
No perceptual issues

Honesty + good judgment

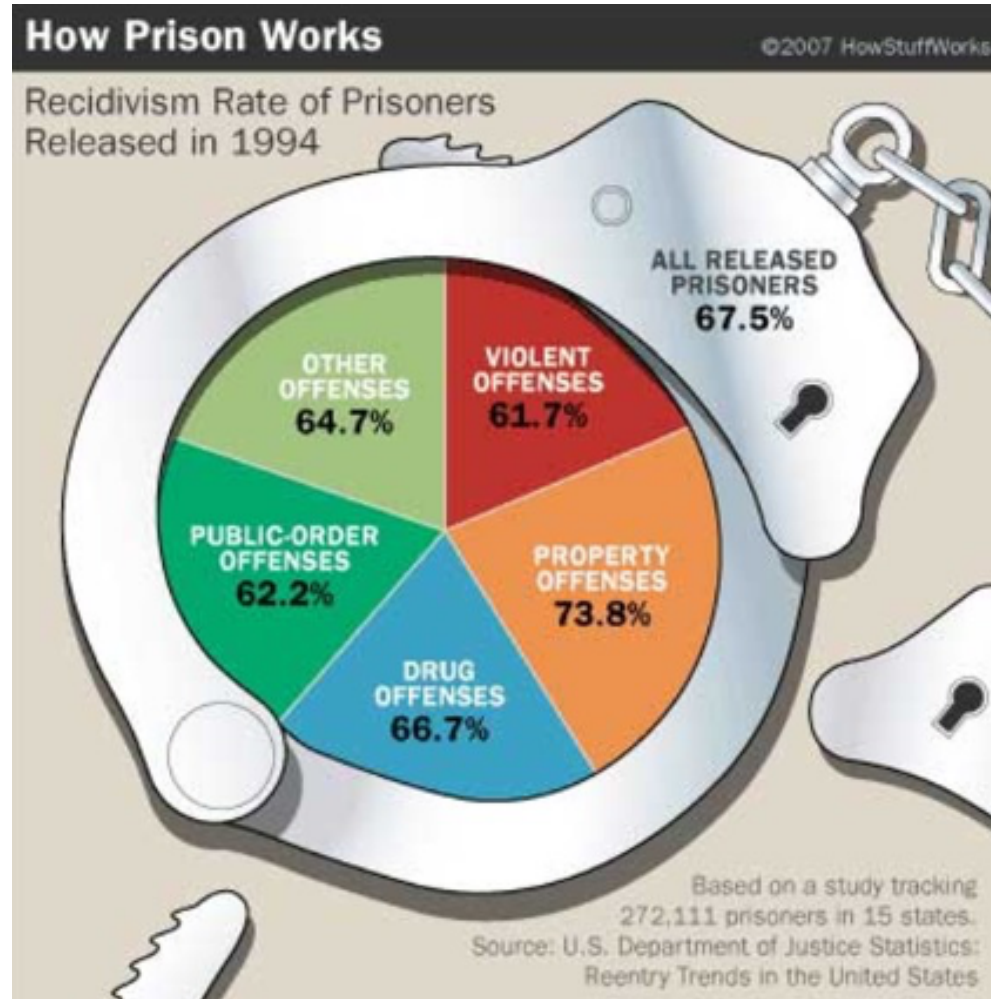Kieran Healy, *Data Visualization: A Practical Introduction*

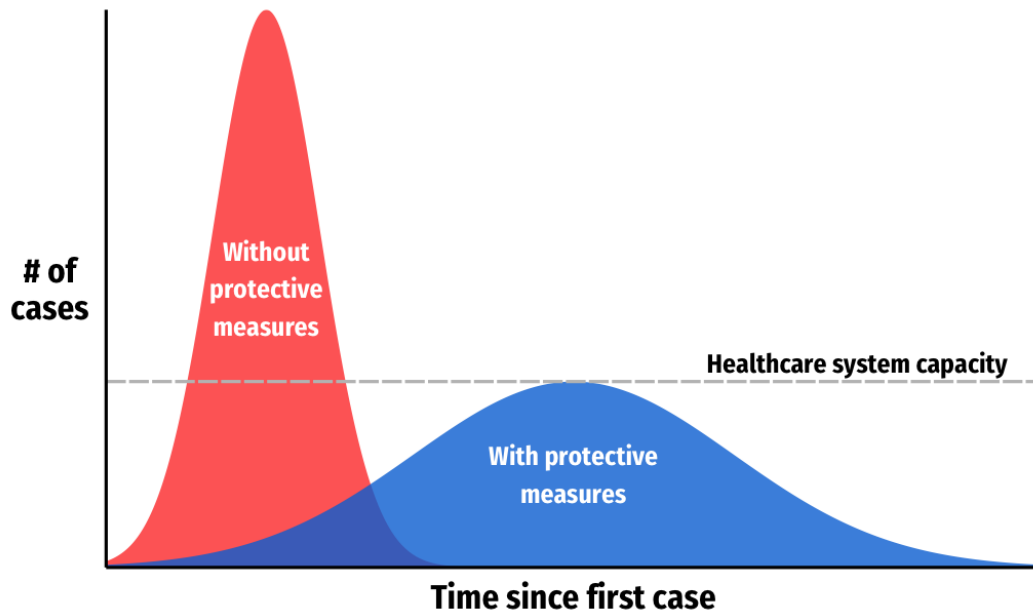# What's wrong?

(f) Distribution of Genus
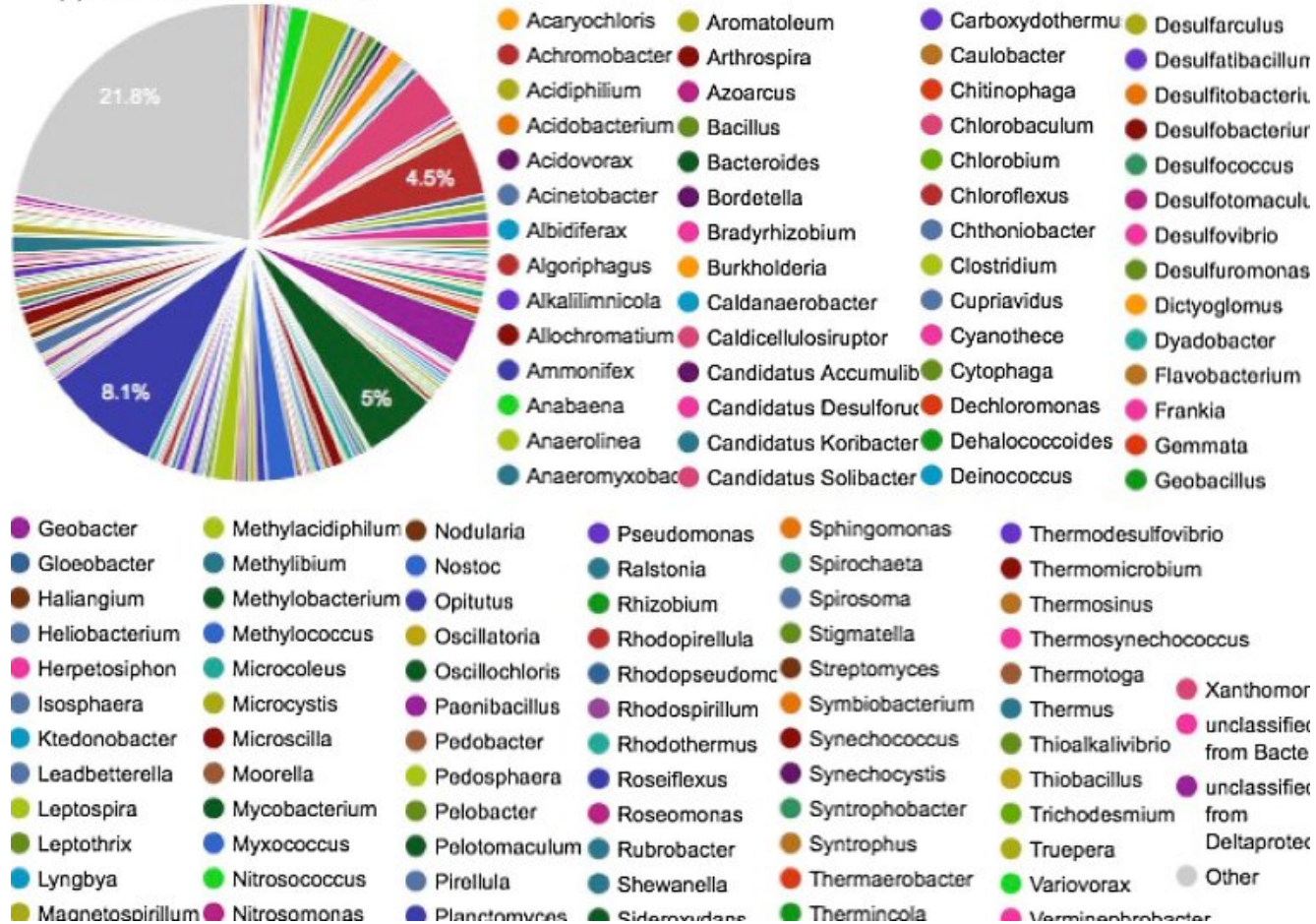
What's wrong?

What's wrong?

What's right?

Flatten the curve!
Slow down community spread by social distancing

# of cases

Without protective measures

Healthcare system capacity

With protective measures

Time since first case

Adapted from the CDC and The Economist
Visit flattenthecurve.com

Carl T. Bergstrom ✔ @CT_Bergstrom · Mar 6
3. There is a lot of complicated epidemiological modeling behind this idea, but this graphic strips all of that away, and discards irrelevant details to provide a straightforward story that people find easy to grasp at a glance.

It *simplifies* and *highlights* what matters.

💬 6   ⟲ 198   ♡ 1.8K

Show replies

Carl T. Bergstrom ✔ @CT_Bergstrom · Mar 6
4. I've seldom seen a piece of sci-comm matter so much. We have an opportunity to flatten the #COVID19 #coronavirus epidemic curve by aggressive social distancing and other measures.

But people don't understand what the point is, if the virus is going to circulate broadly.

💬 8   ⟲ 313   ♡ 2K

Show replies

Carl T. Bergstrom ✔ @CT_Bergstrom · Mar 6
5. This graph provides the answer, powerfully and concisely.

And because of that, it has exploded across twitter and other media. I've used it myself a number of times. This graph is changing minds, and by changing minds, it is saving lives.
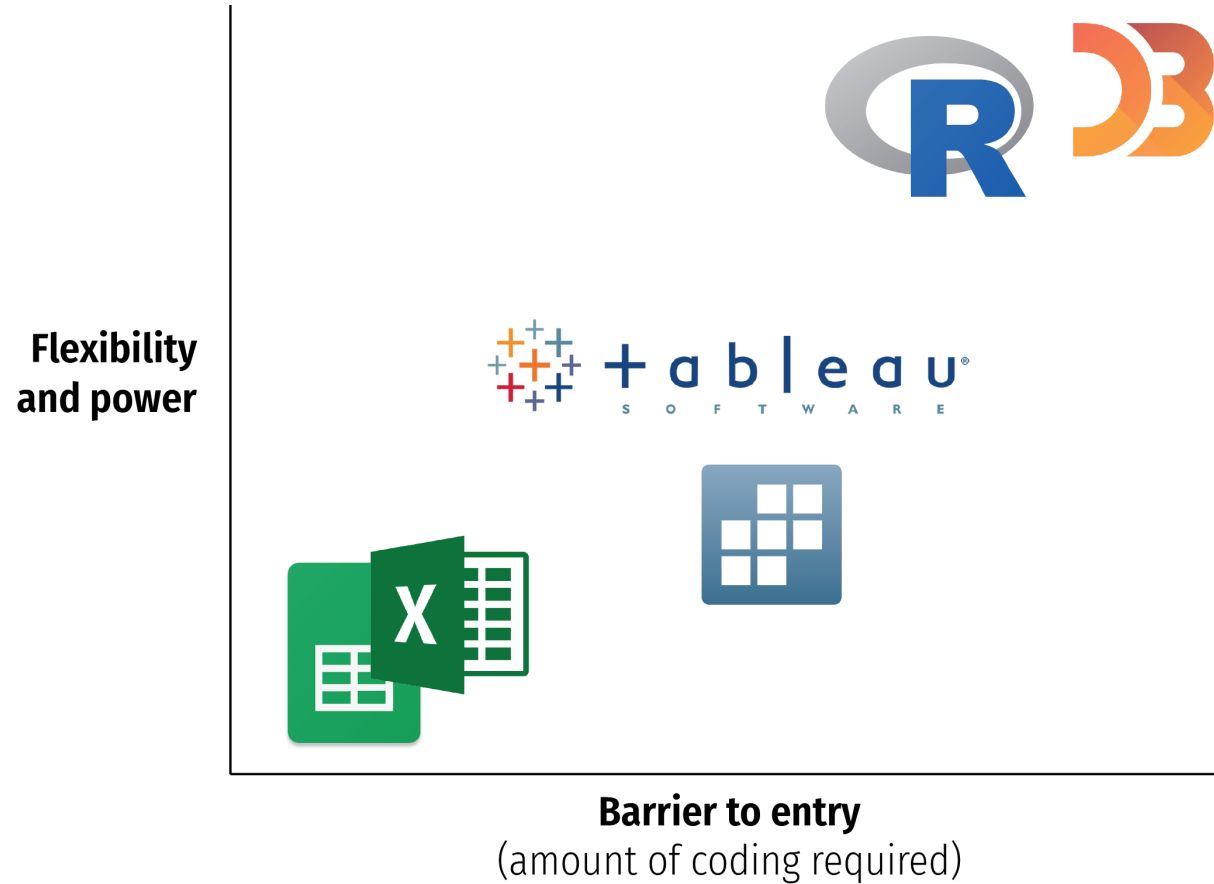
💬 6   ⟲ 196   ♡ 1.5K

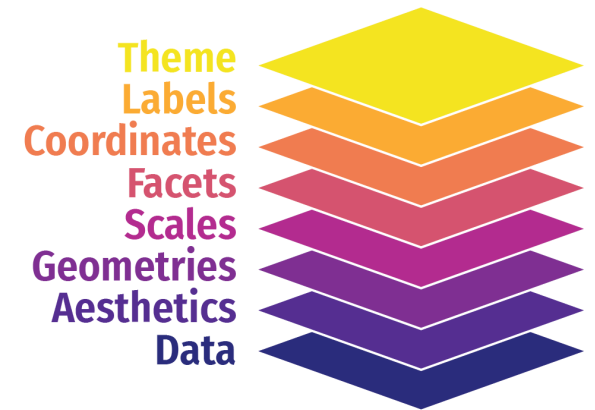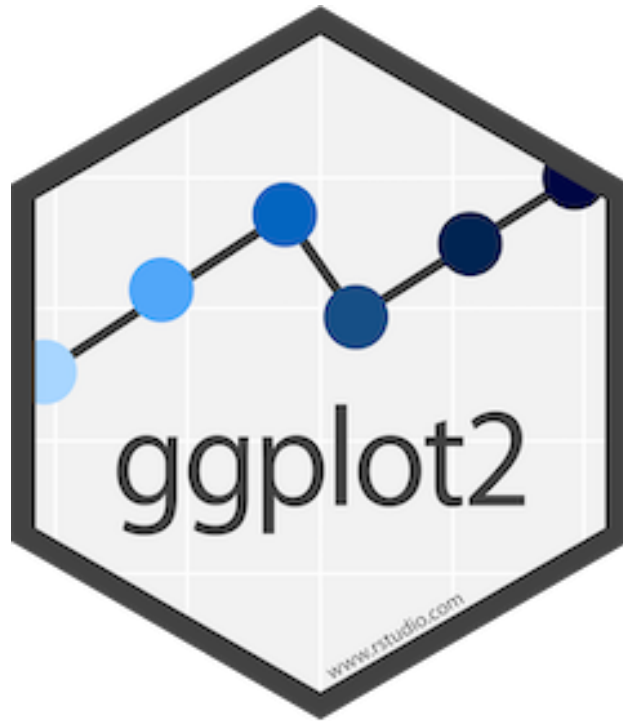Thread by Carl T. Bergstrom

Workshop details

Recognize and create beautiful and truthful visualizations with real world data

Technology

Flexibility and power (vertical axis)
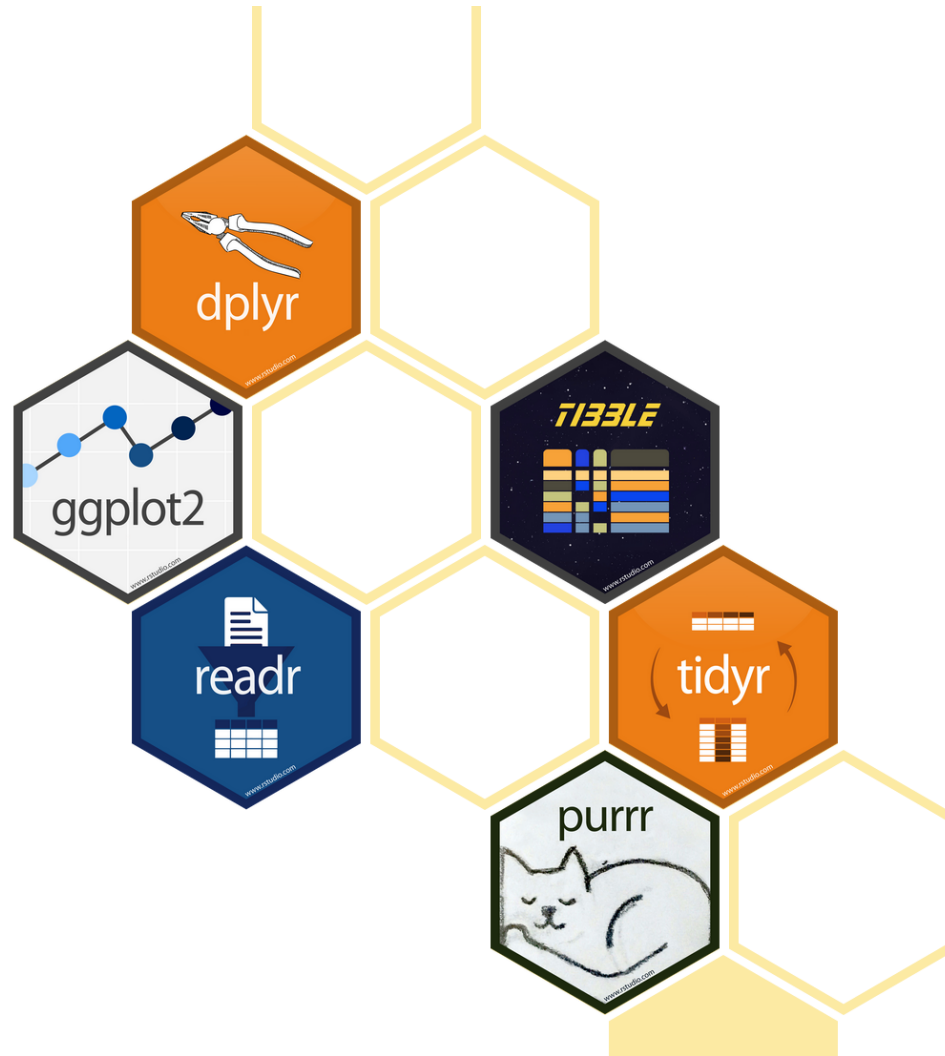
Barrier to entry
(amount of coding required)

Technology

The tidyverse

Sucking

"There is no way of knowing nothing about a subject to knowing something about a subject without going through a period of much frustration and suckiness."

"Push through. You'll suck less."

Hadley Wickham, author of {ggplot2}

Sucking

Karen Rinaldi, "(It's Great to) Suck at Something"

Workshop plan

Truth, beauty, and data

Getting started with R and RStudio

Visualize data with {ggplot2}

Transform data with {dplyr}

Principles of graphic design