

Lecture 15: Oct 8, 2018

Advanced Graphics

- *Designing a Graphic*
- *CRAP*
- *Chart Junk*
- *Modern Graphics*
 - *Animation and Interactive*

James Balamuta
STAT 385 @ UIUC



Announcements

- **hw06** is due **Friday, Oct 12th, 2018 at 6:00 PM**
- **Office Hour Changes**
 - **John Lee's** are now from **4 - 5 PM** on **WF**
 - **Hassan Kamil's** are now from **2:30 - 3:30 PM** on **TR**
- **Quiz 07** covers Week 6 contents @ [**CBTF**](#).
 - Window: Oct 9th - 11th
 - Sign up: [**https://cbtf.engr.illinois.edu/sched**](https://cbtf.engr.illinois.edu/sched)
- Want to review your homework or quiz grades?
Schedule an appointment.

Recap

- **Tidy Data**
 - Each **variable** must have its own **column**
 - Each **observation** must have its own **row**
 - Each **value** must have its own **cell**
- **Long vs. Wide**
 - **Long Data** has each row as one response per subject and any variables for the subject that do not change over time or treatment will have the same value in all the rows. Use **gather()**.
 - **Wide Data** has repeated responses or treatments of a subject in a single row with each response in its own column along with its properties. Use **spread()**.

Lecture Objectives

- Designing graphics to tell data narratives
- Critiquing graphs by using Gestalt tenets
- Differentiate between Static, Animated, and Interactive graphics.

Designing a Graphic

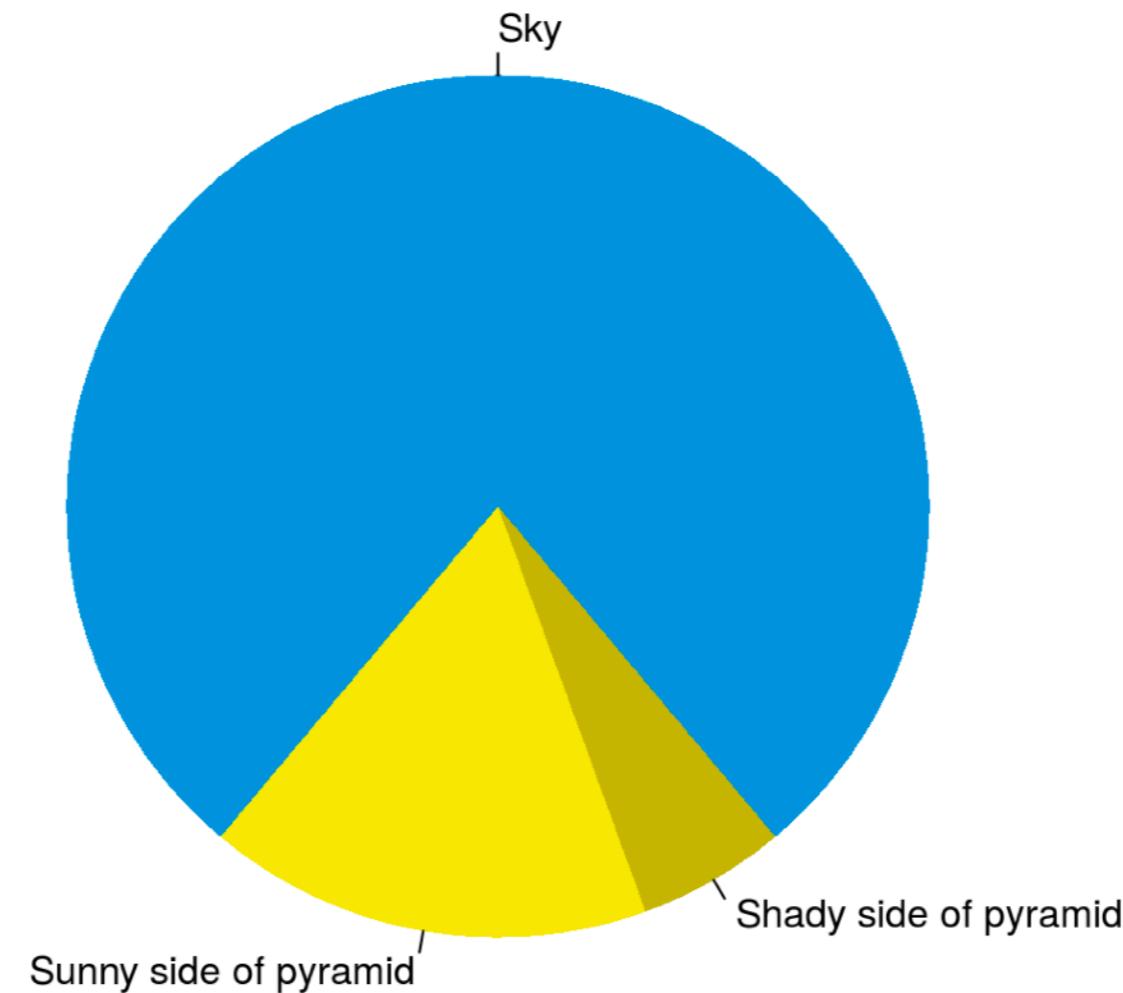
Career of David Bowie



[Source](#)

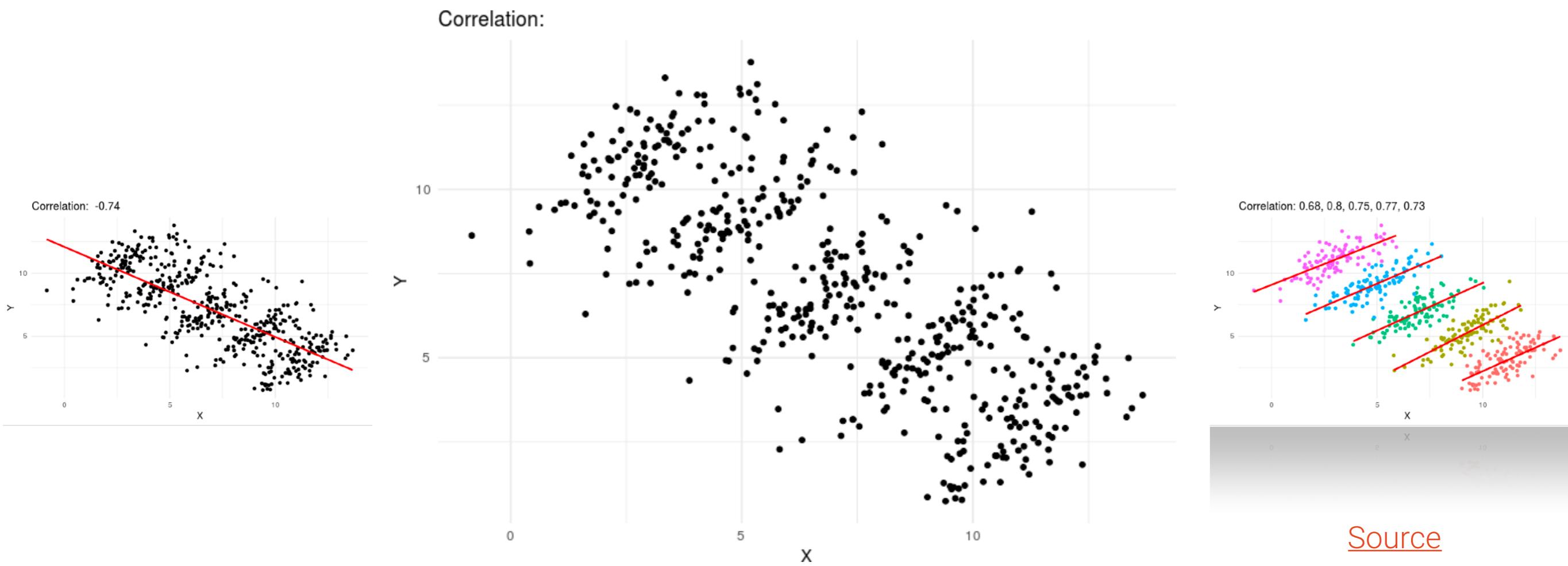
Graphs should...

... tell a **simple, truthful, insightful, beautiful, actionable** story ...

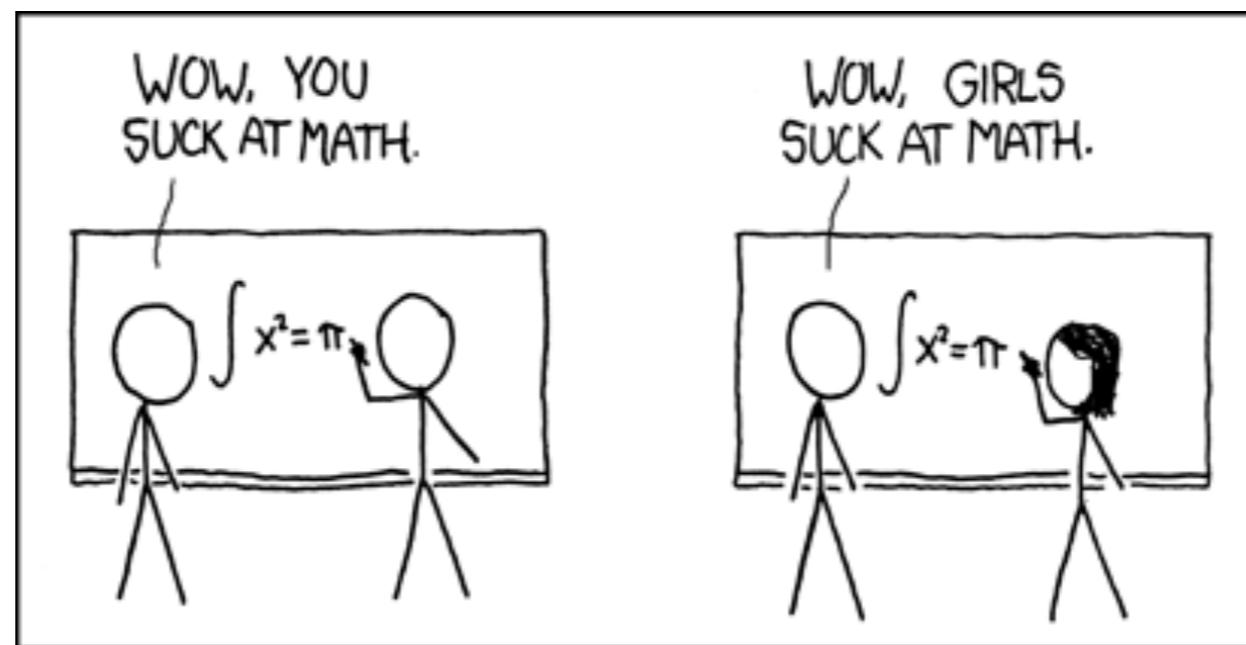


Definition:

Simpson's Paradox describes the phenomena of a pattern appearing in each group but disappearing or changing when groups are removed.



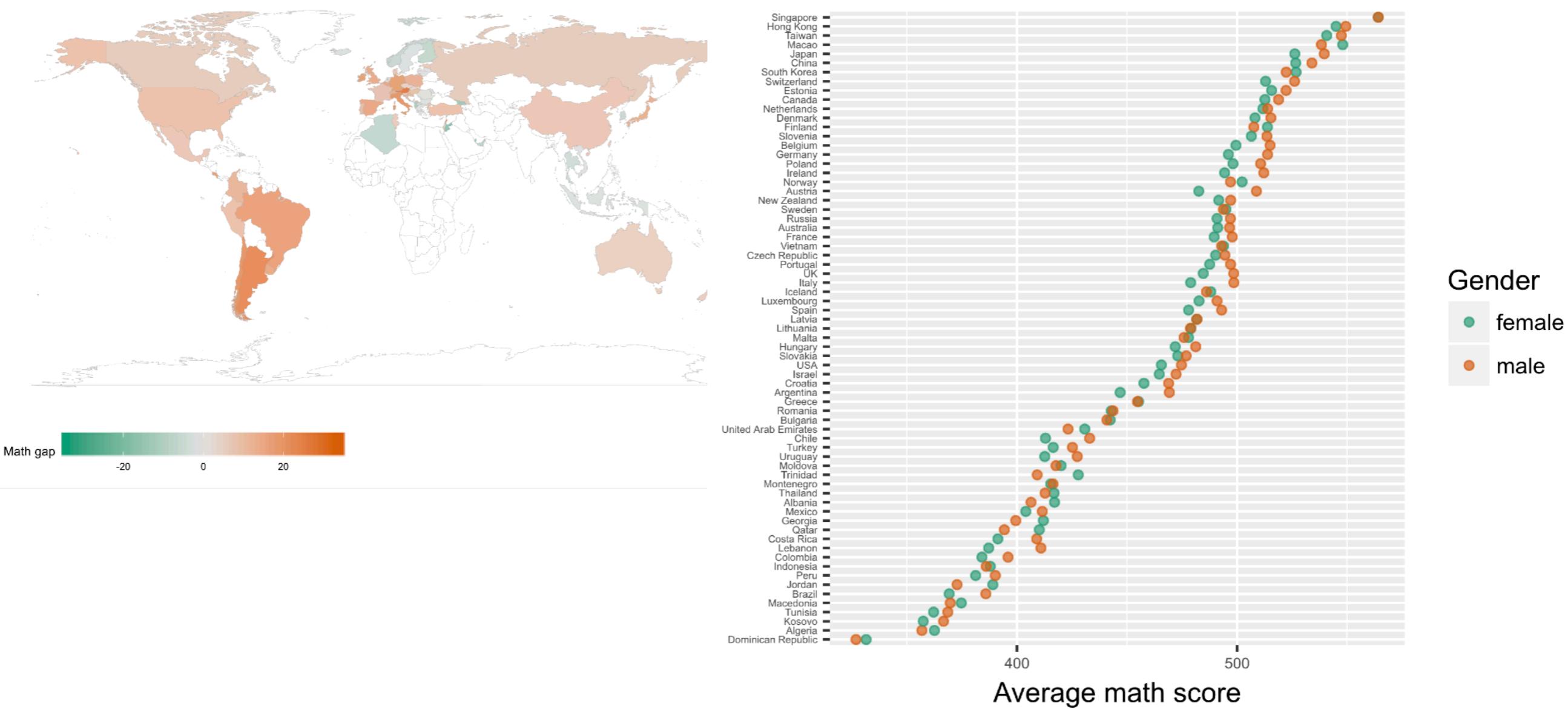
How it Works



<https://xkcd.com/385/>

Math

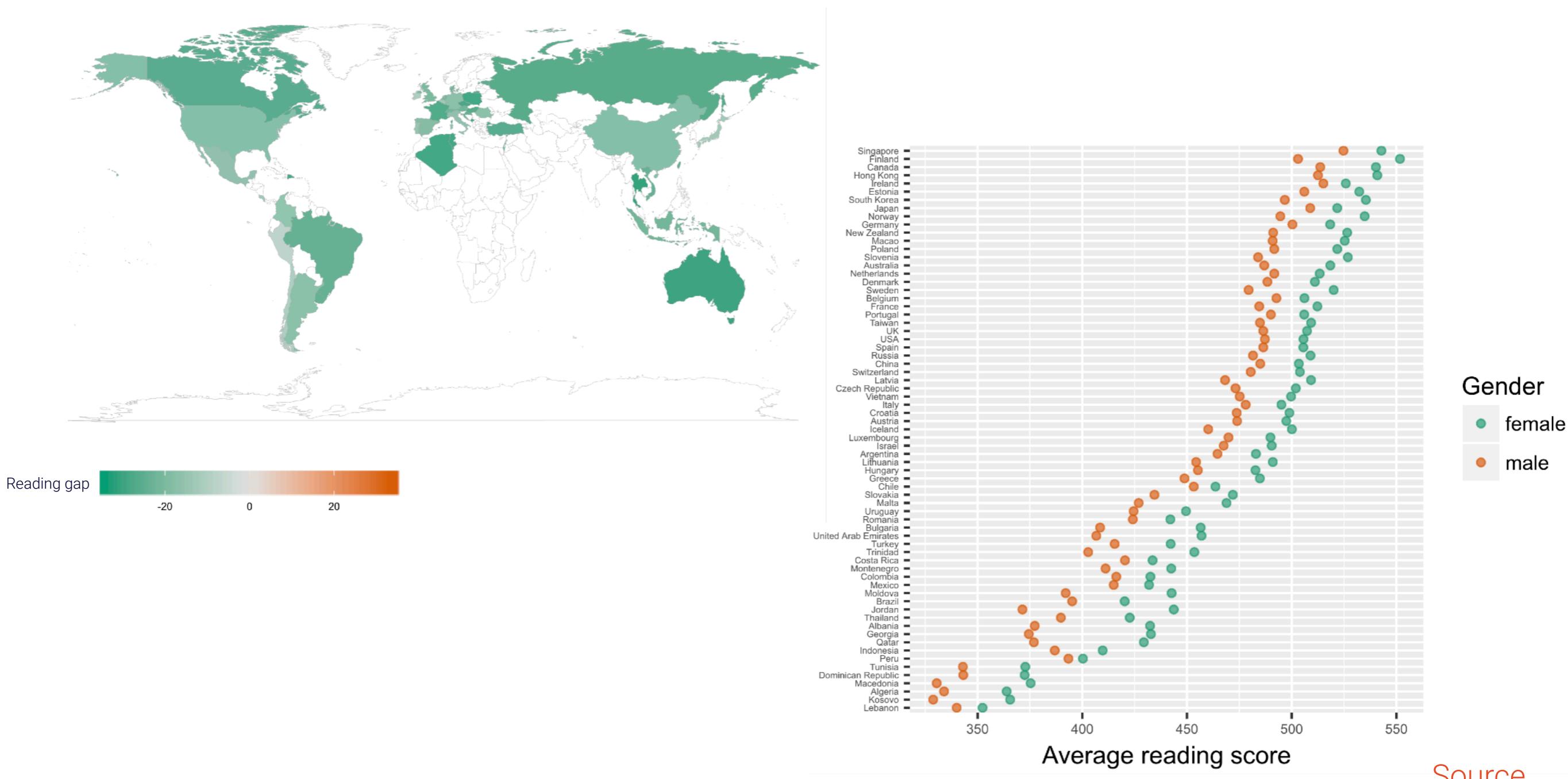
... are there really gender differences ???



[Source](#)

Reading

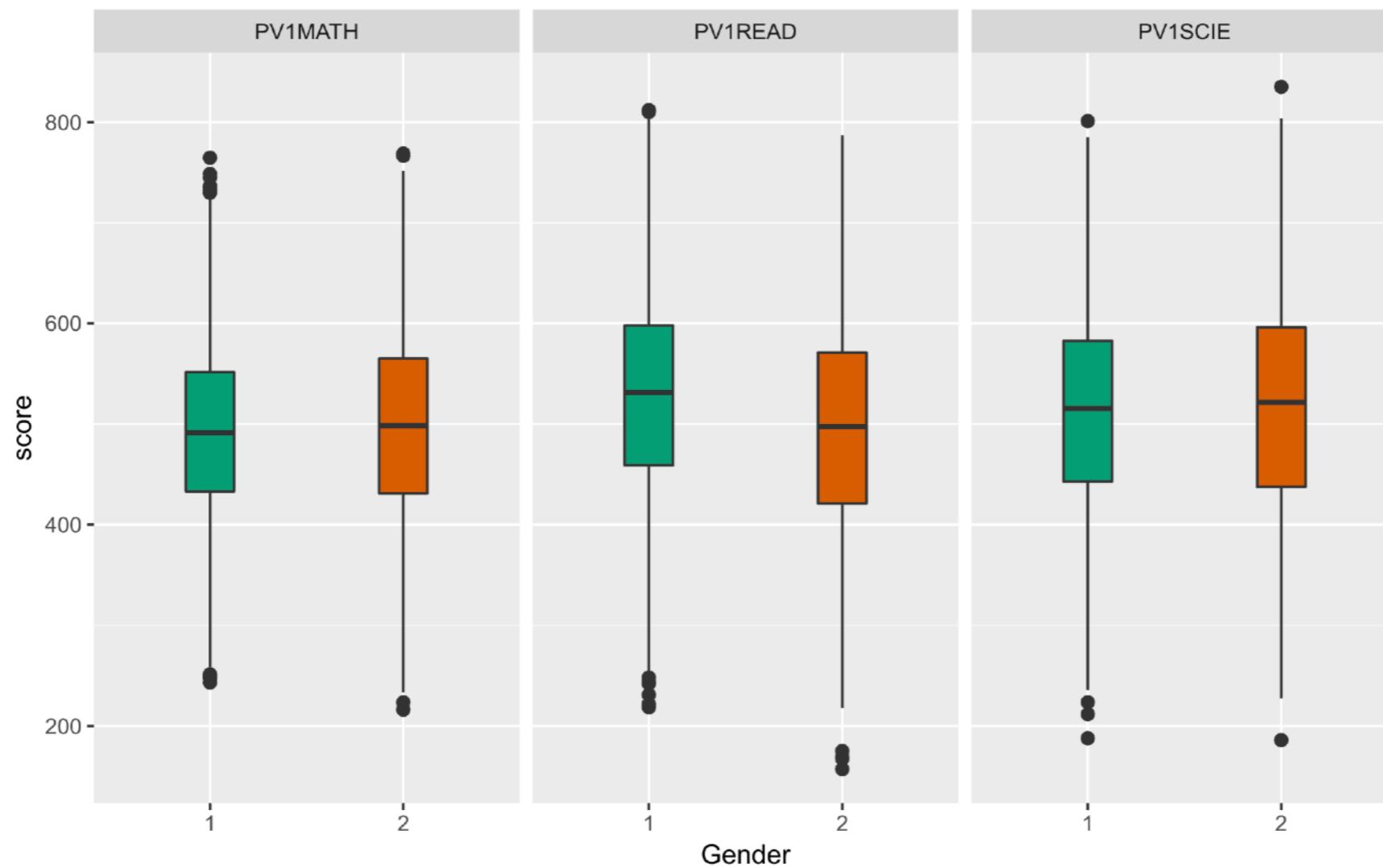
... are there gender differences ???



[Source](#)

Overall

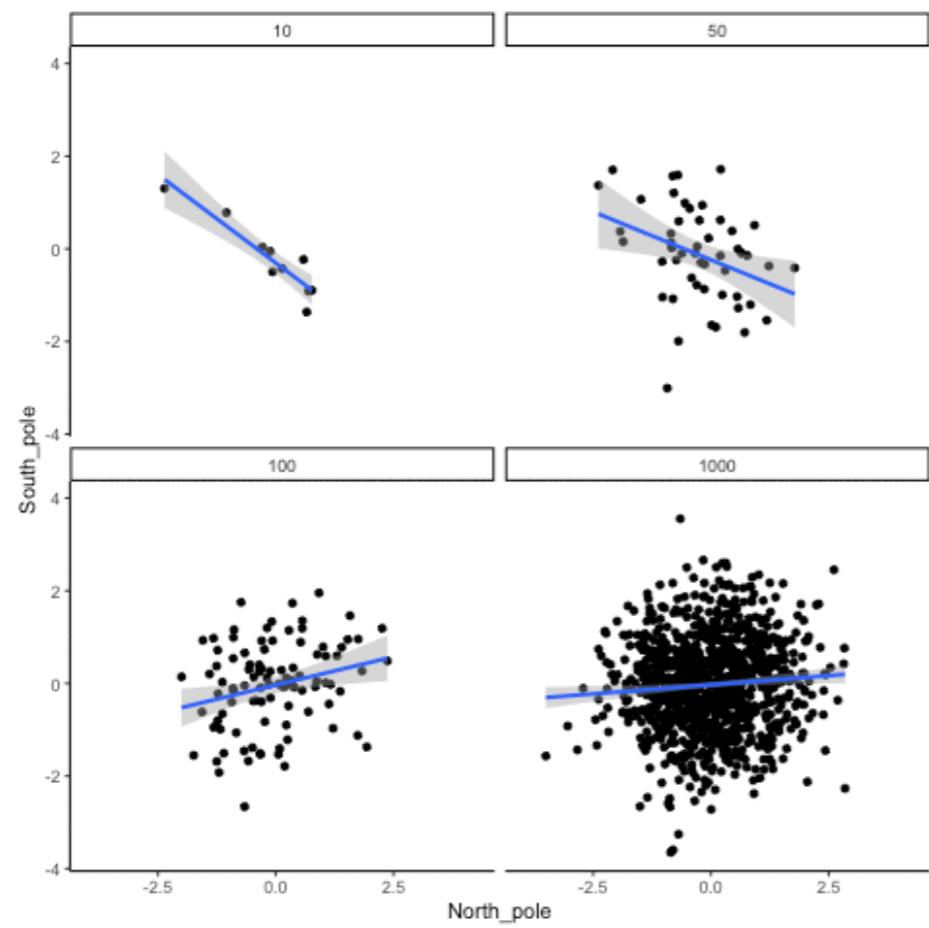
... not a lot of differences ...



[Source](#)

Definition:

Apophenia is the act of perceiving connections and patterns in meaningless data.



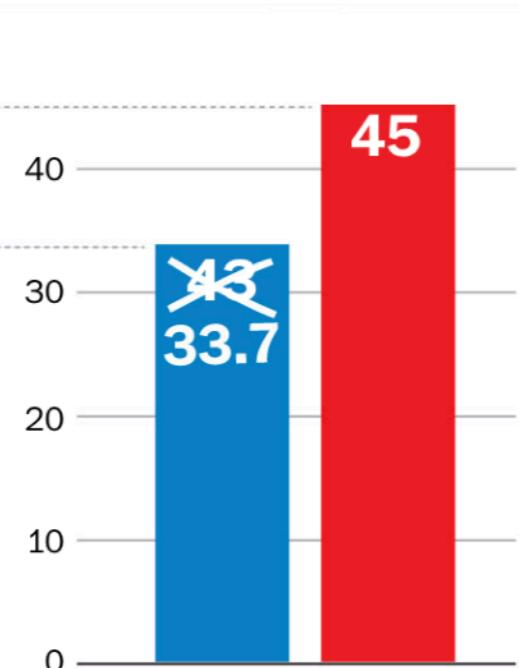
[Source](#)



[Source](#)

Lying with Statistics

... how to misrepresent a bar plot ...



Source: Washington Post

CRAP

CRAP

Contrast, **R**epetition, **A**lignment, **P**roximity

[Chapter six of Presentation Zen: Simple Ideas on Presentation Design and Delivery by Garr Rey](#)

1

Contrast

Contrasting Space

Contrasting Color

Serif

Times serif font

Sans Serif

Helvetica sans serif font

Script

Savoye script font

Monospaced

Courier monospaced font

* serif

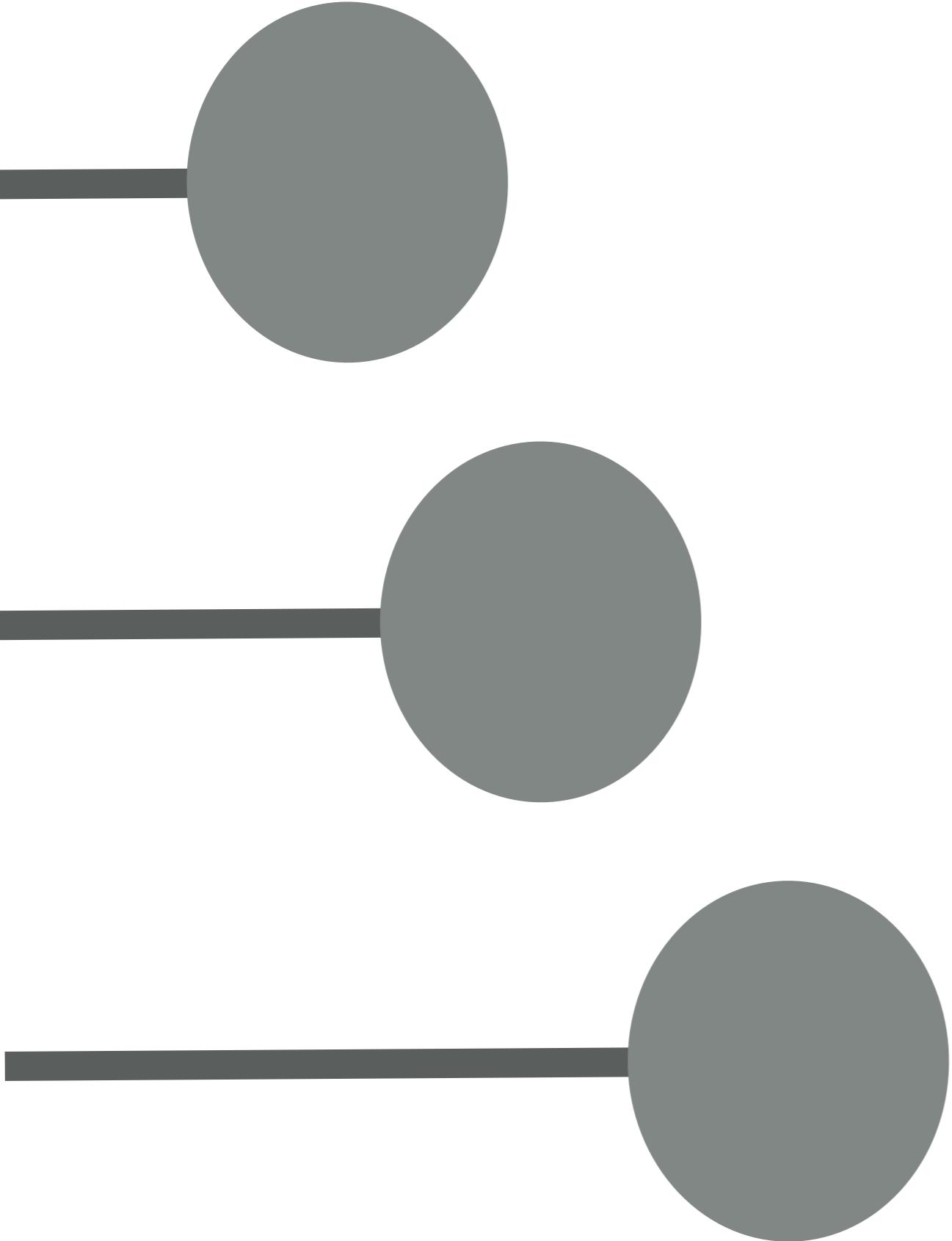
Contrasting Type

2

Repetition



Repetition: Shape



Repetition: Color and Shape

Do you know the muffin man?

Repetition: Text

3

Alignment

Did you remember to pick up a bag of flour?

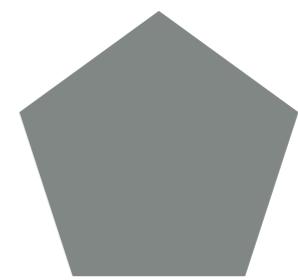
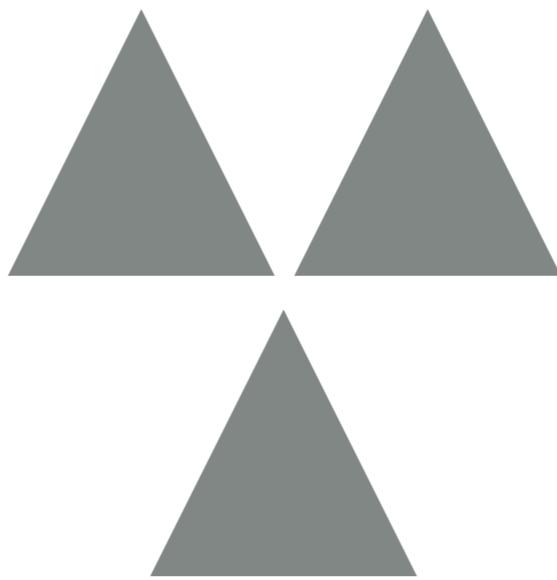
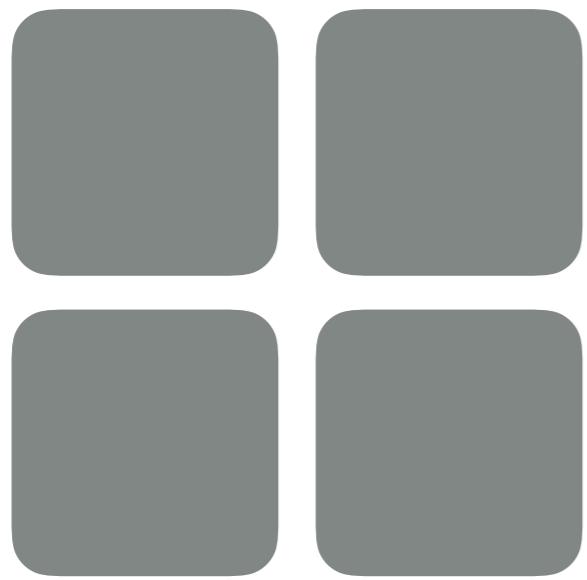
No. I'll swing by the store after work.

Thanks! I'm trying to make a cake.

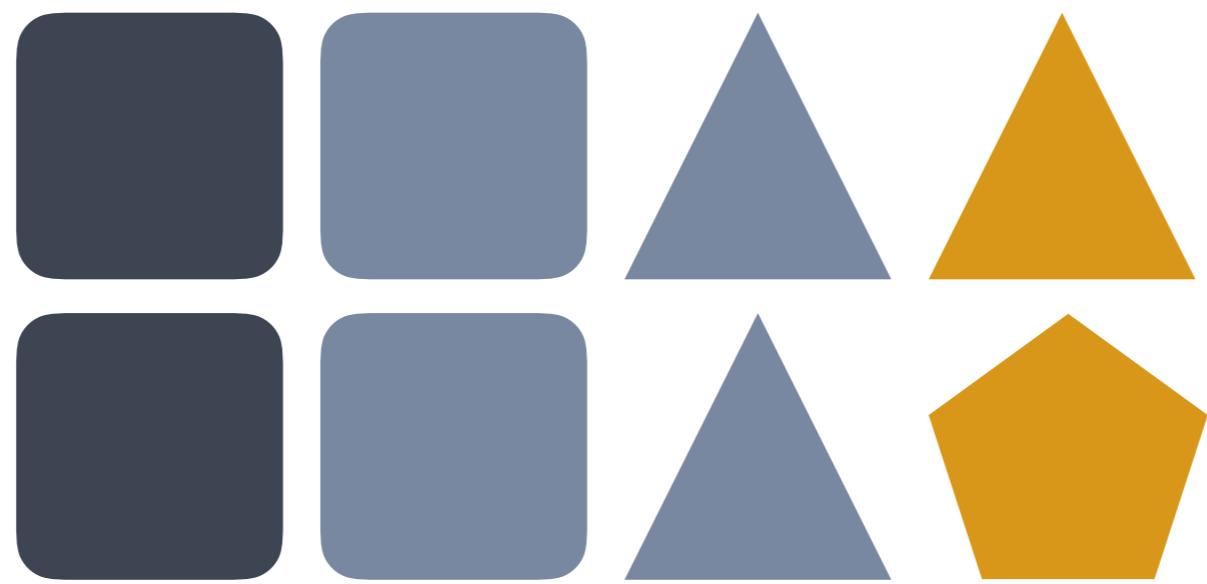
Alignment: Text

4

Proximity

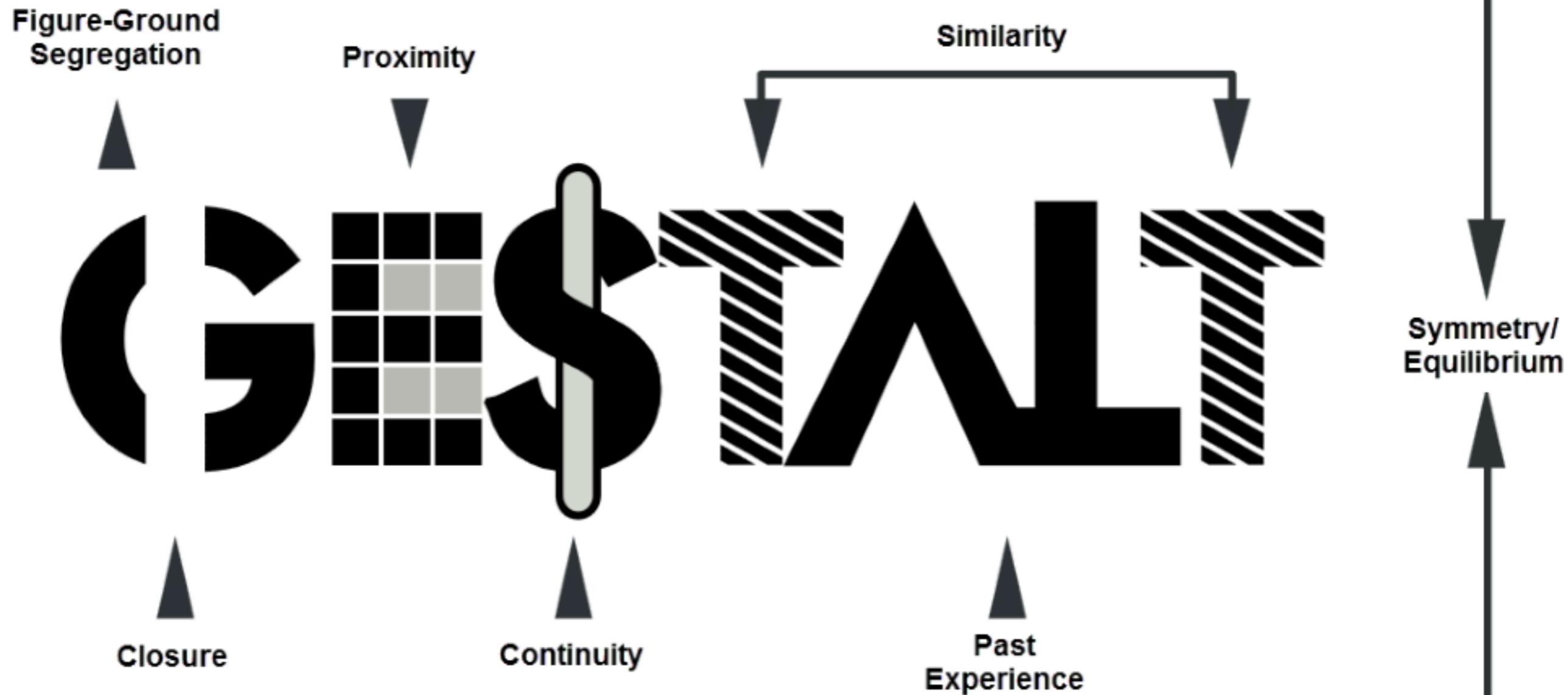


Proximity: Same Objects



Proximity: Color

CRAP emphasizes GESALT principles

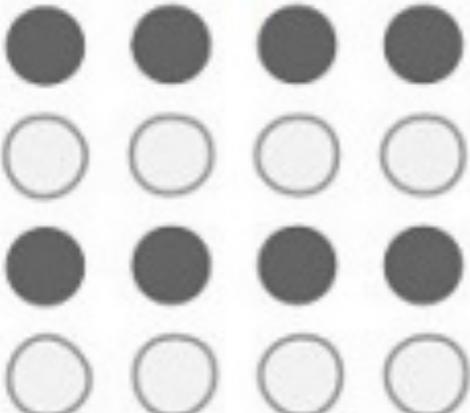


[Source](#)

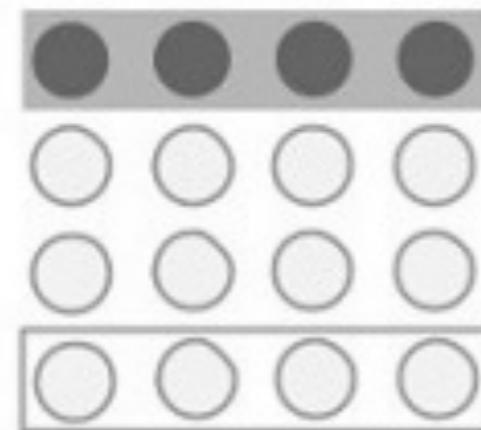
Proximity



Similarity



Enclosure



Symmetry



Closure



Continuity

Connection

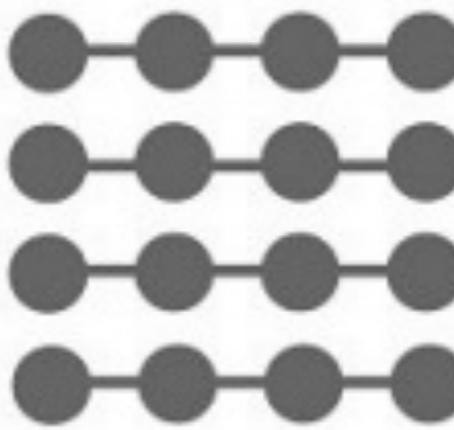


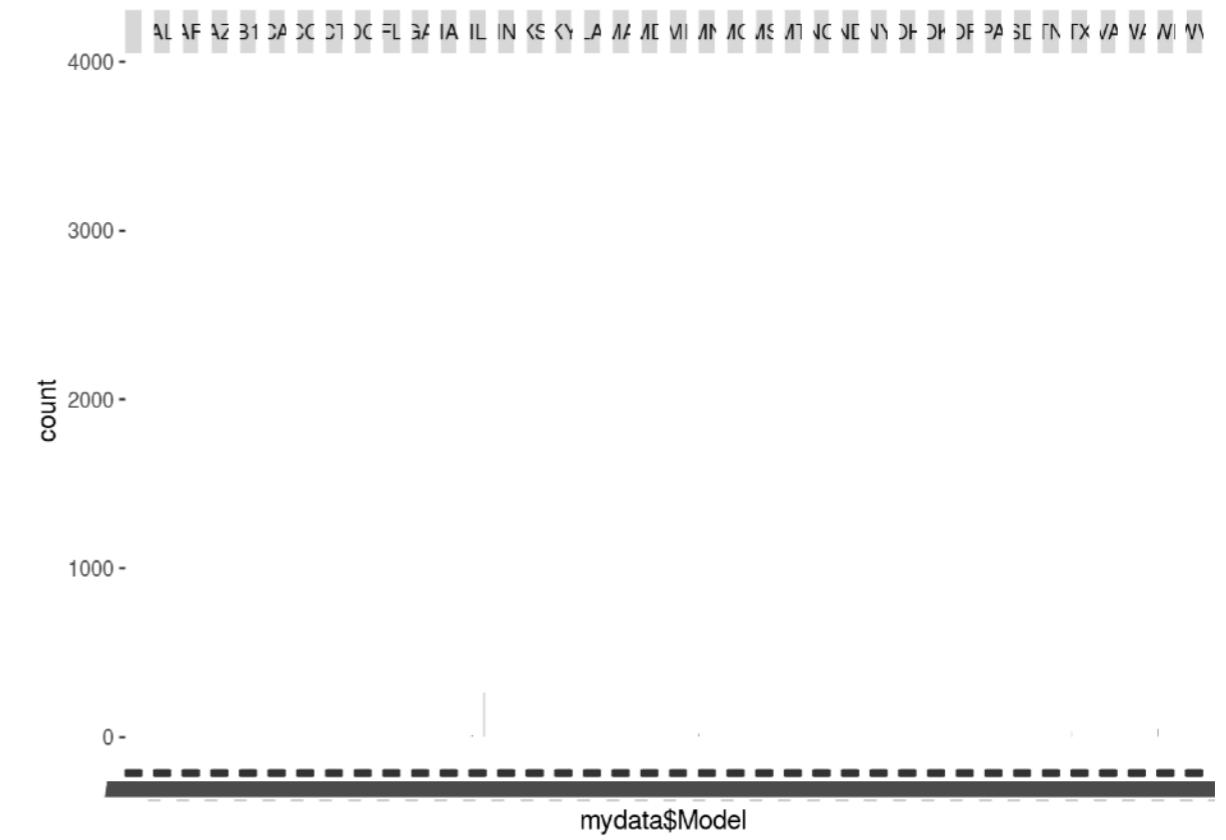
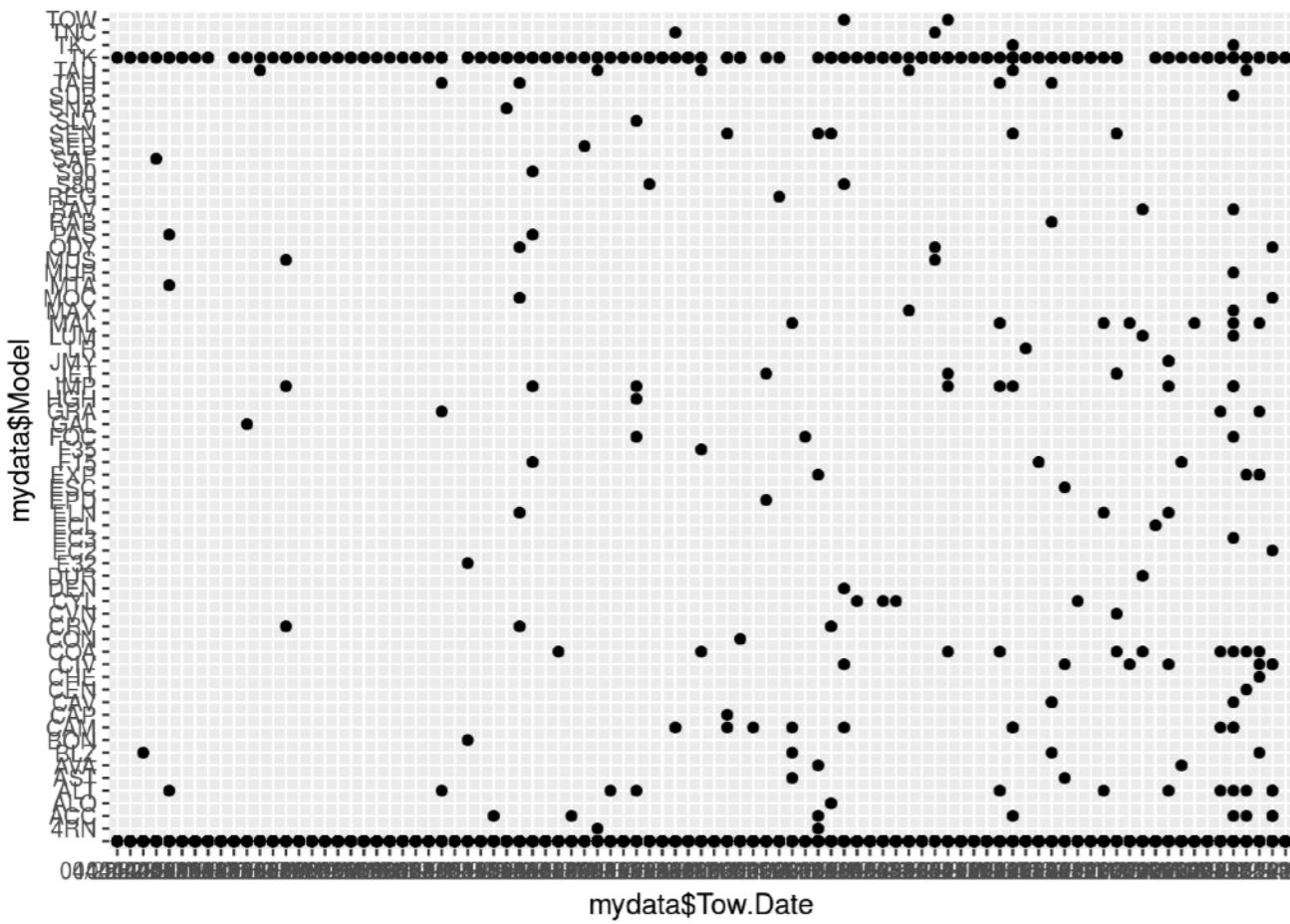
Figure & ground



[Source](#)

Questionable Graphs

... what's being shown here ???



Design Hiccups

... how could we improve these fliers ??

The document is a Microsoft Word file with the following content:

Sponsored by: abbvie, astellas, CITADEL | Securities

Hosted by: NIC-ASA, ICSA Midwest

JOIN US!

2018 ICSA Midwest & NIC-ASA Joint Fall Meeting

A Great Event for Networking and Updating Quantitative Skills

When: October 11- 12, 2018

Where: Delta by Marriott Chicago Northshore, Glenview, IL

Who: Statisticians, Business Analysts, Data Scientists and All Quantitative Scientists!

<http://midwest-icsa.org/>

Day 1:

- Day 1 short courses on historical data borrowing in clinical trials and master protocols and Day 2 parallel sessions
- Breakfast and lunch are included
- Generous student poster awards
- Career services for students available (including Citadel!)
- Free** registration for student attendees
- Free** cocktail reception with hors d'oeuvres for everyone!!
- New!!! Free** one night hotel stay for students outside of Chicago area who submitted posters (space limited, in the order of poster submission date)

Day 2:

- Day 2 parallel sessions
- Breakfast and lunch are included
- Generous student poster awards
- Career services for students available (including Citadel!)
- Free** registration for student attendees
- Free** cocktail reception with hors d'oeuvres for everyone!!
- New!!! Free** one night hotel stay for students outside of Chicago area who submitted posters (space limited, in the order of poster submission date)

Page 1 of 1 130 words English (United States) Focus 115%

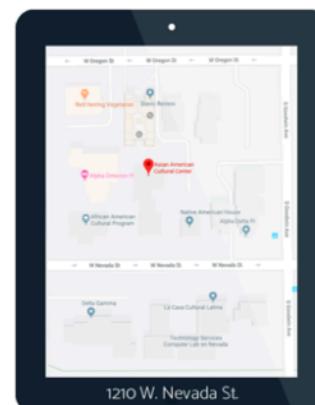
GLOBAL ENGAGEMENT LOUNGE

Fall 2018 Schedule of Topics

- 9/7 - Intro to Registered Student Organizations
- 9/14 - The Career Center & Work Authorization
- 9/21 - Taste of UIUC with McKinley Health
- 9/28 - Culture Shock and De-stress
- 10/5 - Wellness with Campus Rec
- 10/12 - Self-care & Flu Shot Station
- 10/19 - Taste of UIUC and Student Assistance Center
- 10/26 - Student Insurance
- 11/2 - Illinois Leadership Center
- 11/9 - Ten Myths about Traffic Law with Student Legal Services
- 11/16 - Taste of UIUC with McKinley Health
- 11/30 - RSO Representatives
- 12/7 - End of Fall Celebration



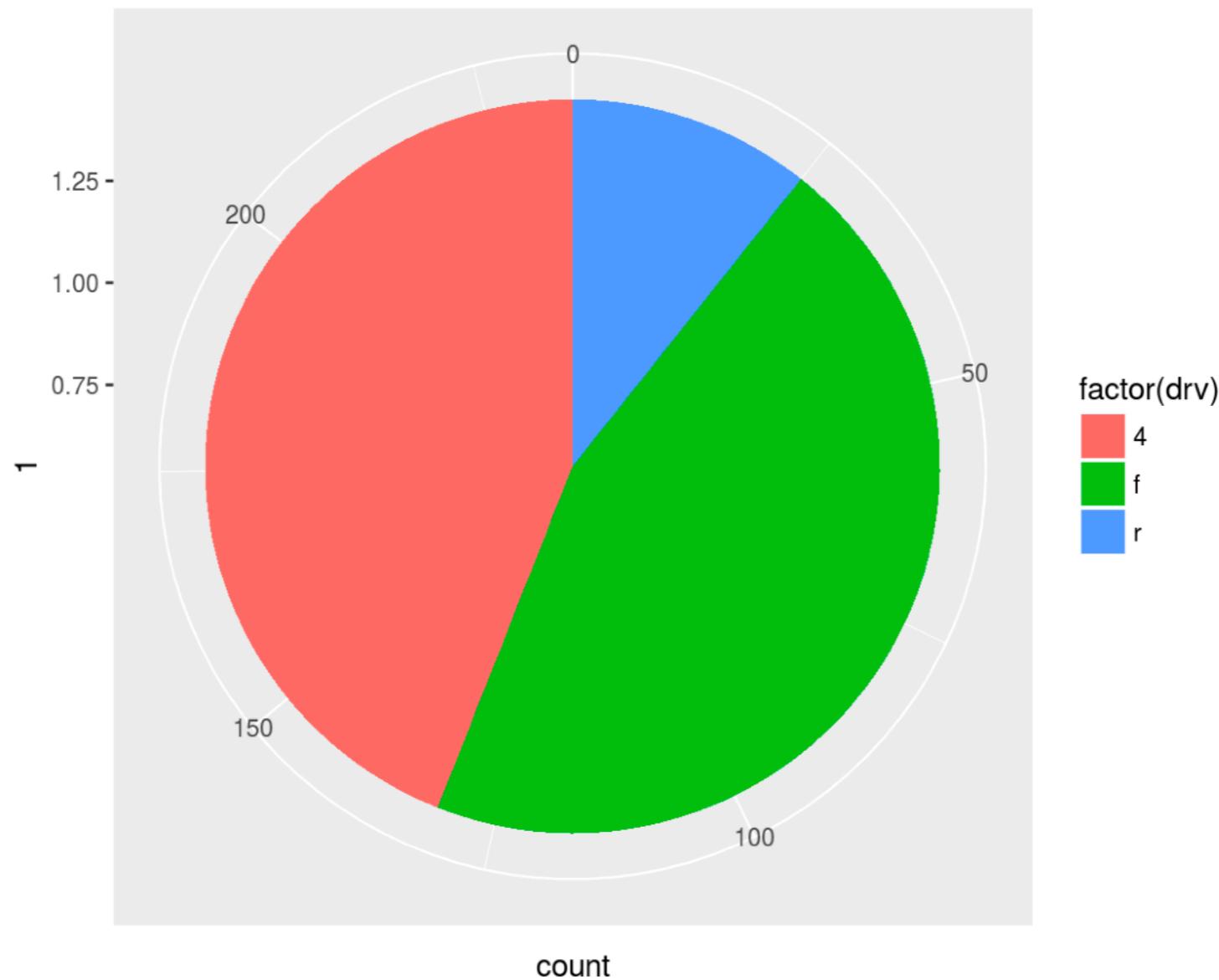
Join us for the Global Engagement Lounge at the Asian American Cultural Center! This is an informal place to hang out, meet new friends, and chat with staff from various campus units. Bring your friends! All are welcome!



Sponsored by: Strategic Initiatives Grant Program



Is a pie chart ideal?

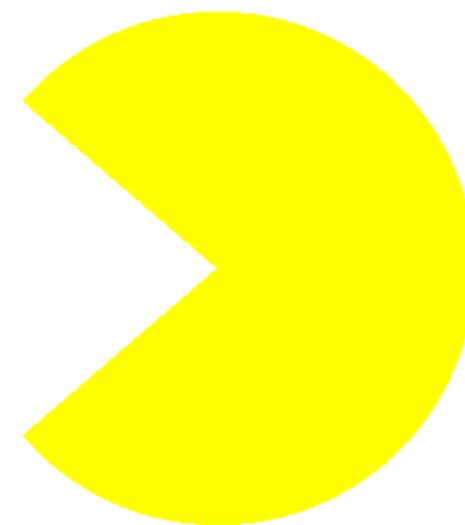


Is a pie chart ideal?

... yes, when talking about pie and pacman ...



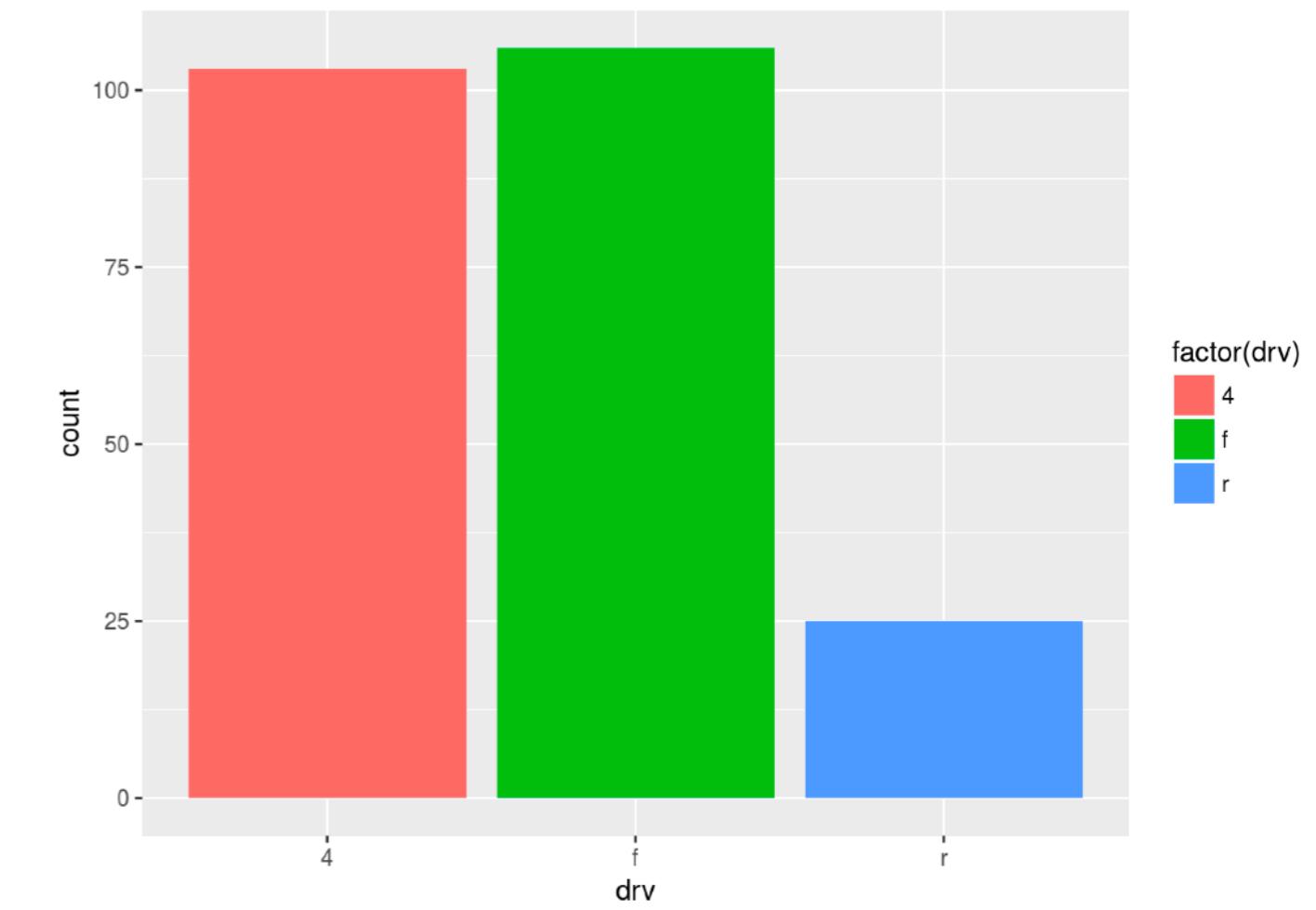
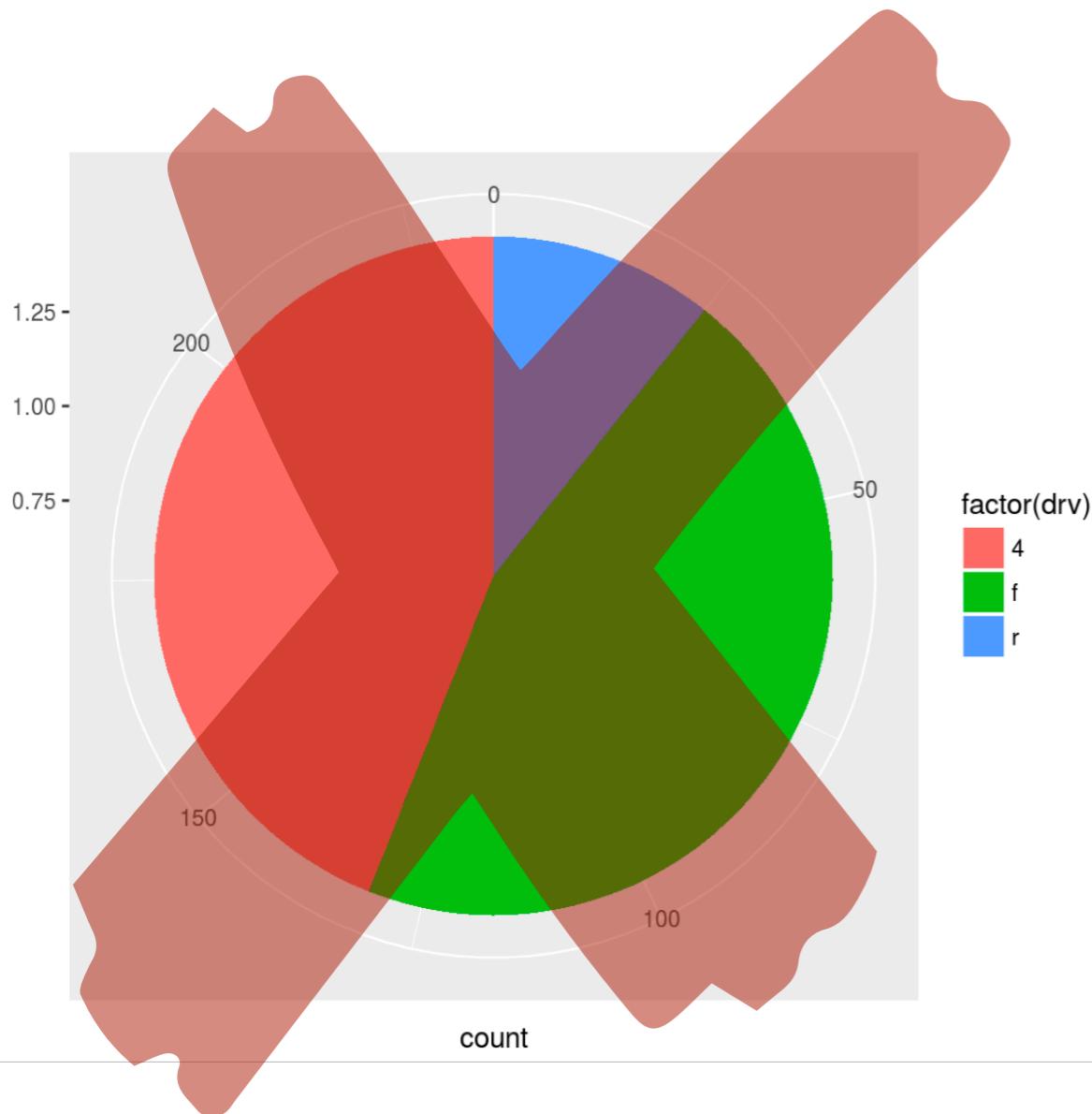
[Source](#)



[Source](#)

Is a pie chart ideal?

... no, a pie chart is not ideal when faced with > 2 categories ...



Spence, I. [No Humble Pie: The Origins and Usage of a Statistical Chart \(2005\)](#).

Journal of Educational and Behavioral Statistics. Vol 30, Issue 4, pp. 353 - 368

Cleveland, W. S. and McGill, R. [Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Method \(1983\)](#).

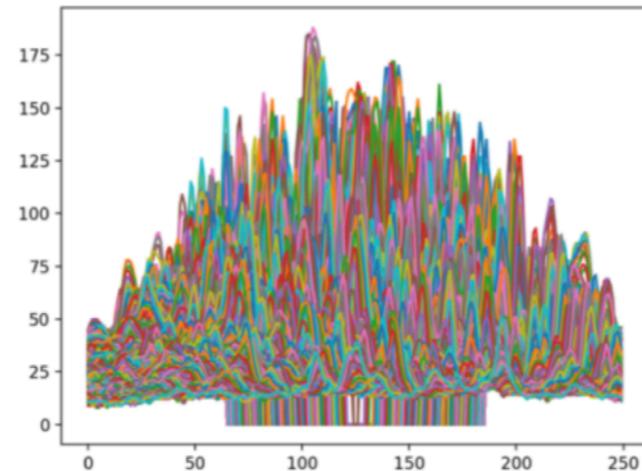
Journal of the American Statistical Association. Vol 79, Issue 387, pp. 531 - 554

When Graphing Goes Astray

accidental aRt

2018-08-15

14 notes



When data
visualization goes
beautifully wrong.
@accidental_aRt
Tweet Curated by
@kara_woo and
@ErikaMudrak

SUBMIT A POST

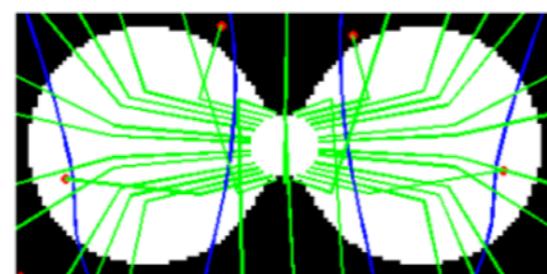
ARCHIVE / RSS

Image analysis gone horribly beautiful. via Ryan.

[Tweet](#)

2018-05-31

1 note

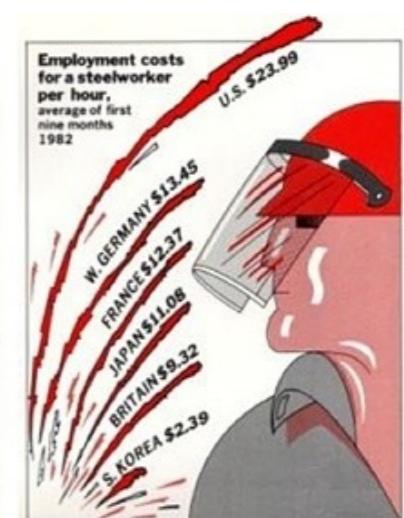
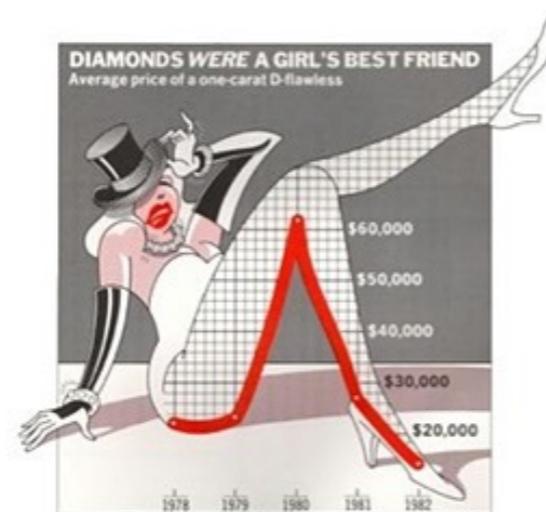


<http://accidental-art.tumblr.com/>

Chart Junk

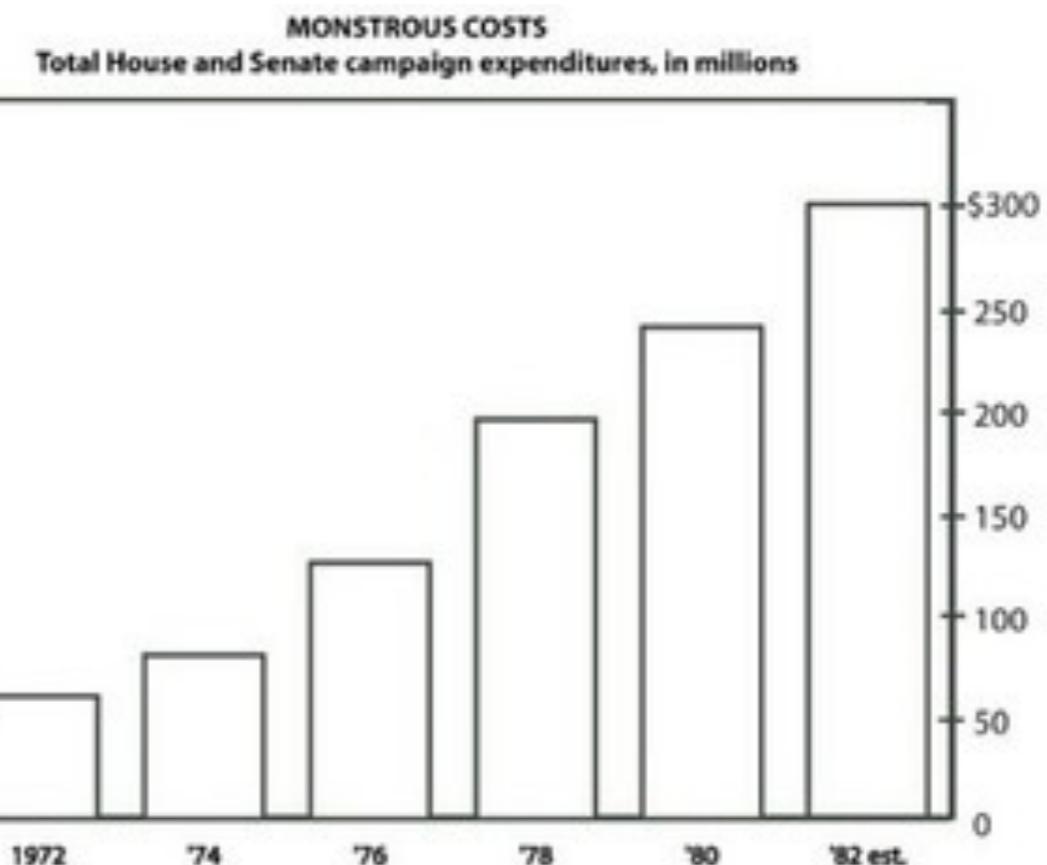
Definition:

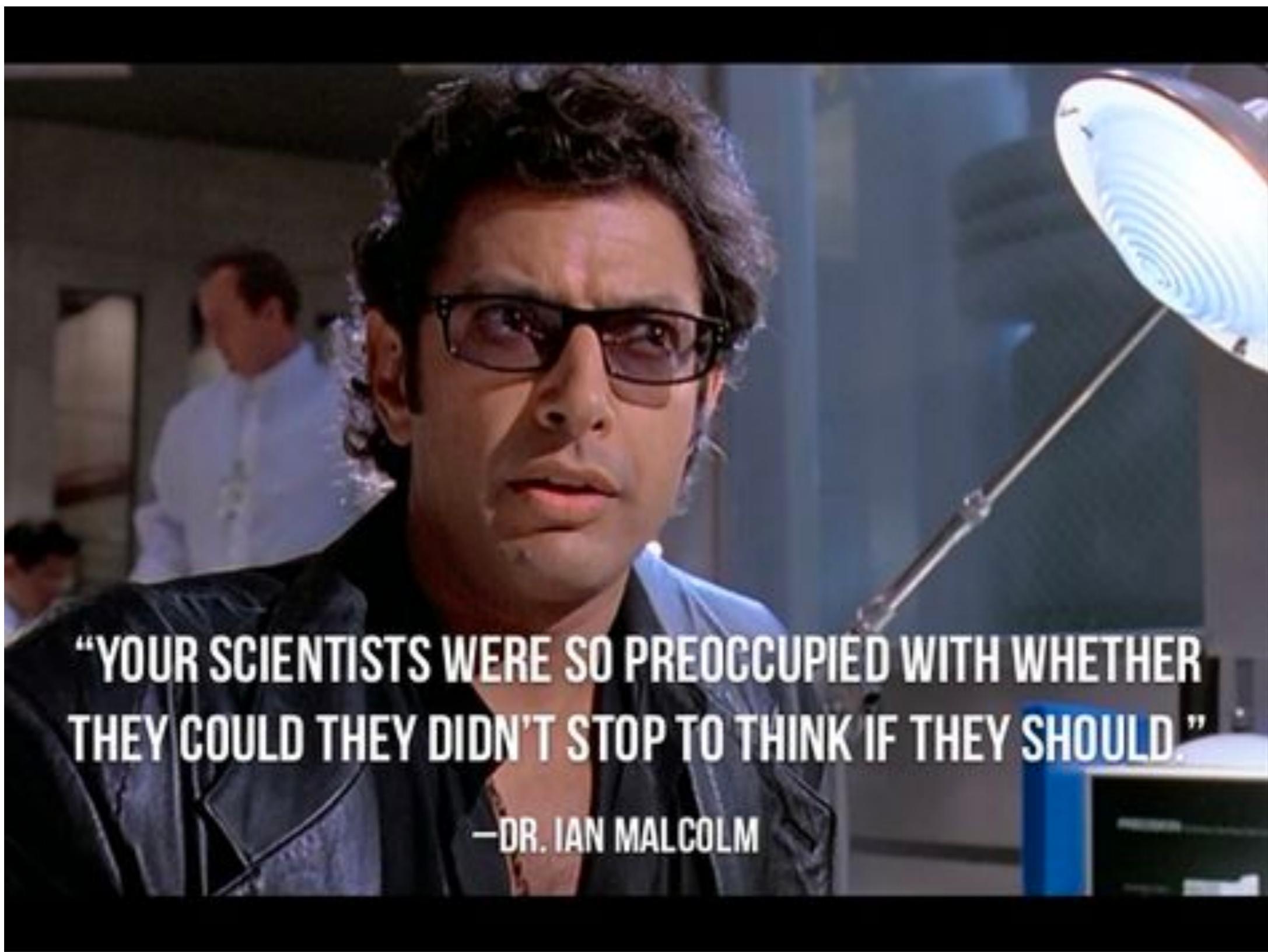
Chartjunk refers to incorporating visual elements that distract from the main informative purpose of the graphic.



Clarity of a Graphic

... visually appealing vs. descriptive ...





“YOUR SCIENTISTS WERE SO PREOCCUPIED WITH WHETHER
THEY COULD THEY DIDN'T STOP TO THINK IF THEY SHOULD.”

—DR. IAN MALCOLM

"Lie" Factor

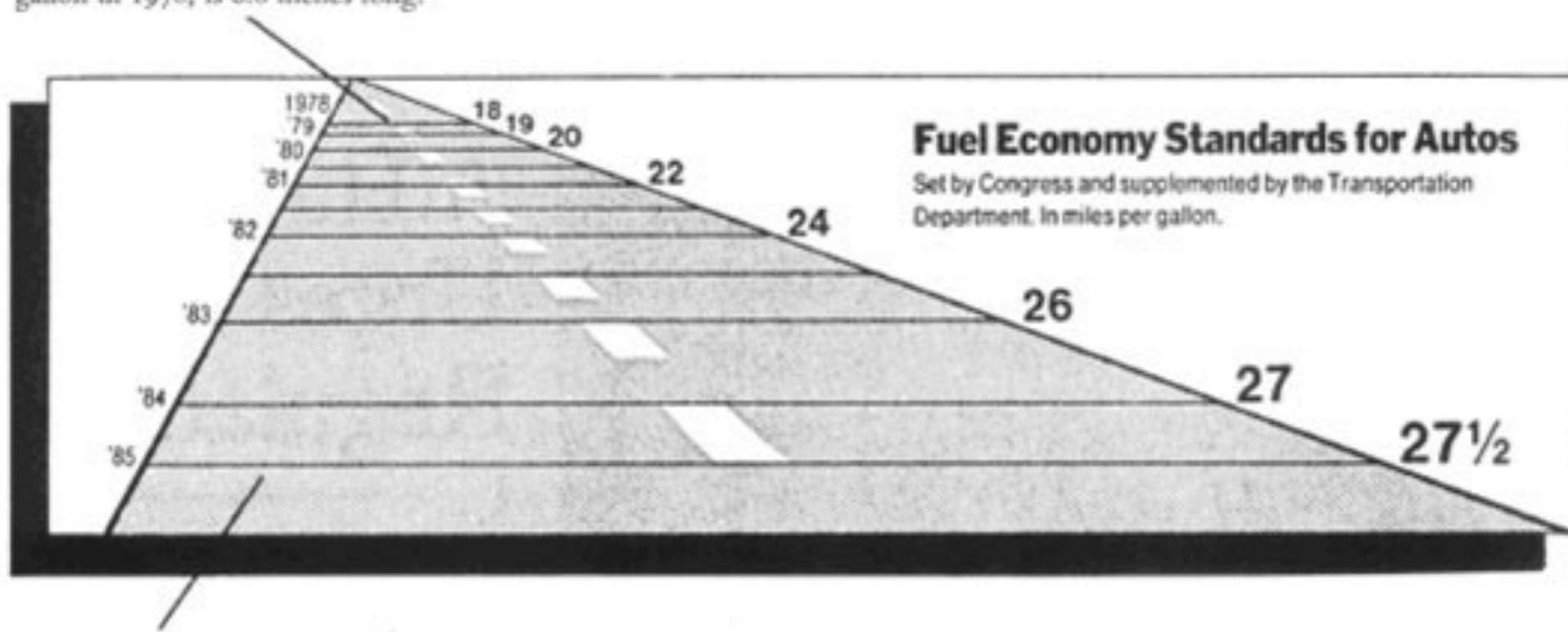
... tufte metric for assessing graphs ...

$$\text{Lie Factor} = \frac{\text{size of effect shown in } \mathbf{\text{graphic}}}{\text{size of effect shown in } \mathbf{\text{data}}}$$

$$\text{Size of Effect} = \frac{|\text{second value} - \text{first value}|}{\text{first value}}$$

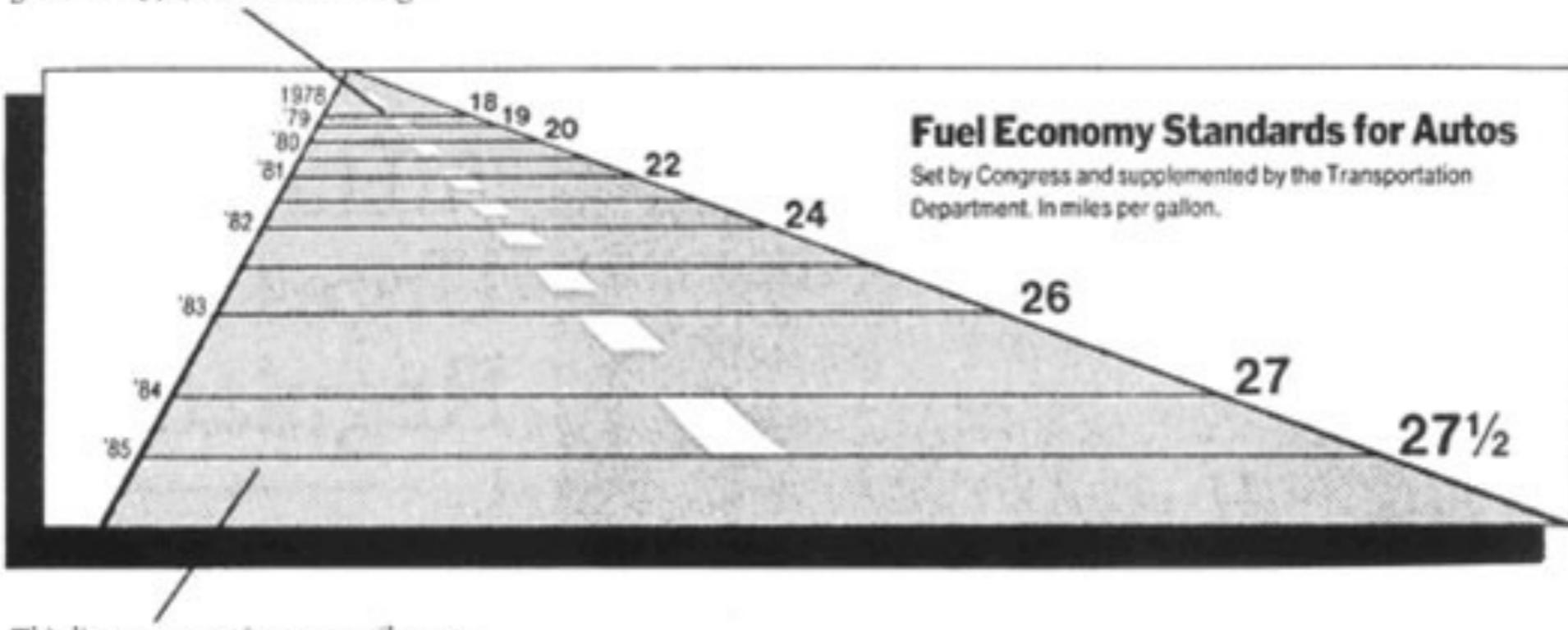
Lie Factor $> 1 \rightarrow$ exaggerated graph

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.



This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

This line, representing 18 miles per gallon in 1978, is 0.6 inches long.



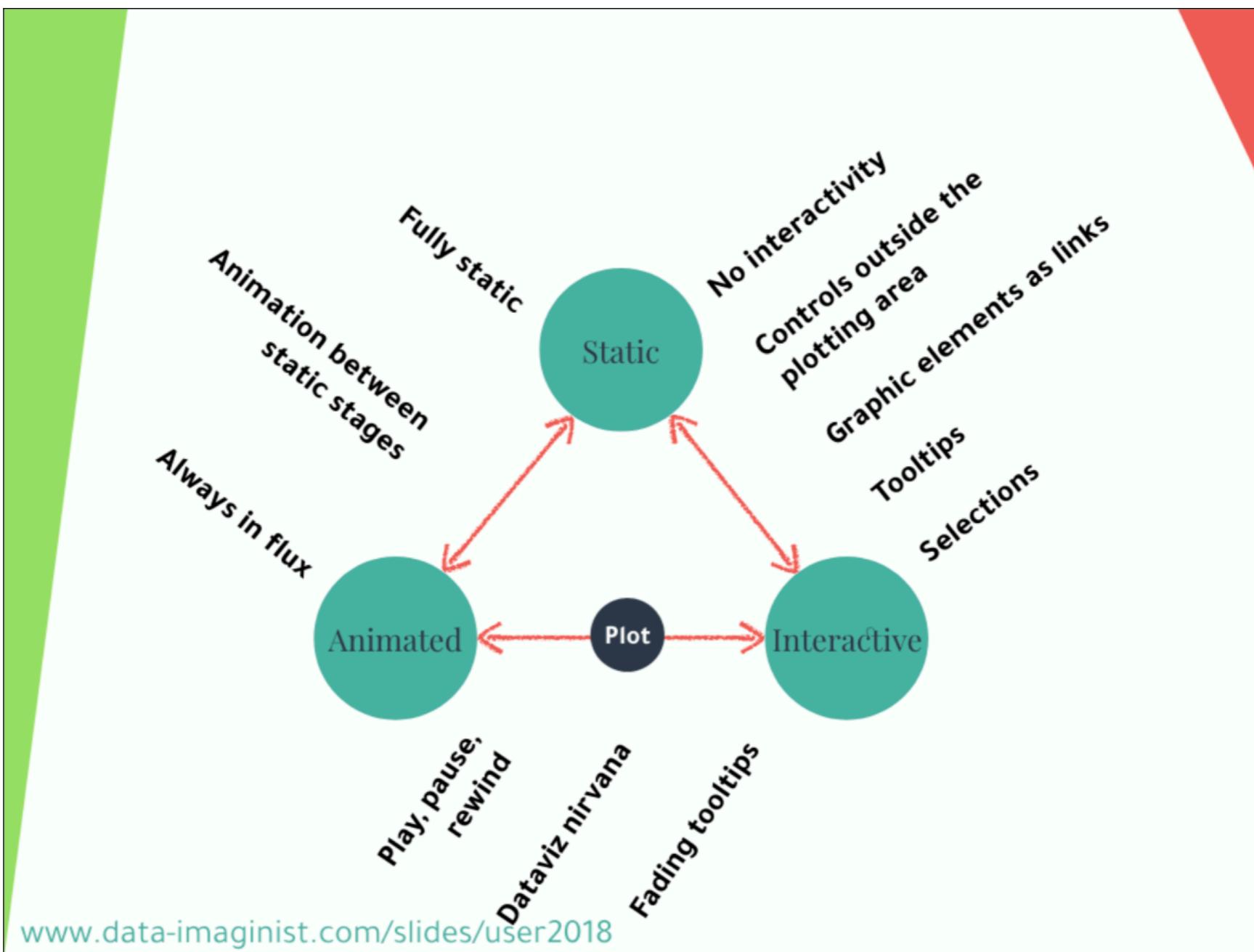
This line, representing 27.5 miles per gallon in 1985, is 5.3 inches long.

$$\text{Lie Factor} = \frac{5.3 - 0.6}{\frac{27.5 - 18}{18}}$$

Modern Graphics

Graph Relations

... comparing different information mediums ...

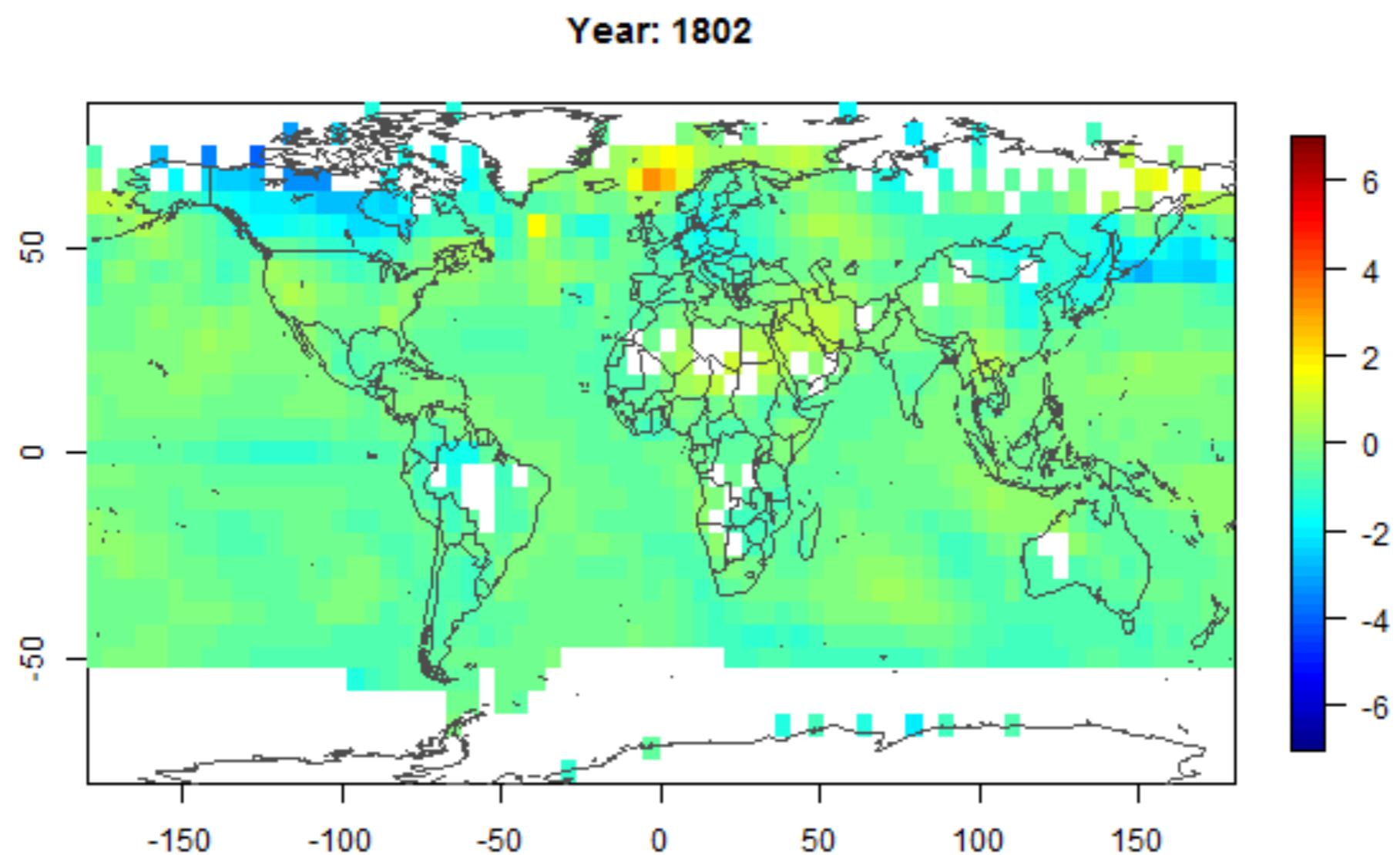


Source

Animated

Animated Time Lapse

... Temperature Change over Time and Space ...



Reconstructed Climate Data

Render Multiple Images

... stand alone plots ...

	heatmap1001.png [v2]	Apr 15, 2015 by James Jos...	13.7 KB
	heatmap1000.png [v2]	Apr 15, 2015 by James Jos...	13.8 KB
	heatmap0999.png [v2]	Apr 15, 2015 by James Jos...	13.8 KB
	heatmap0998.png [v2]	Apr 15, 2015 by James Jos...	13.8 KB
	heatmap0997.png [v2]	Apr 15, 2015 by James Jos...	13.8 KB

Combined into a **GIF**

... Graphic Interchange Format or Animated Movie ...

 heatmap1001.png V2	Apr 15, 2015 by James Jos...	13.7 KB
 heatmap1000.png V2	Apr 15, 2015 by James Jos...	13.8 KB
 heatmap0999.png V2	Apr 15, 2015 by James Jos...	13.8 KB
 heatmap0998.png V2	Apr 15, 2015 by James Jos...	13.8 KB
 heatmap0997.png V2 [redacted]	Apr 15, 2015 by James Jos...	13.8 KB [redacted]

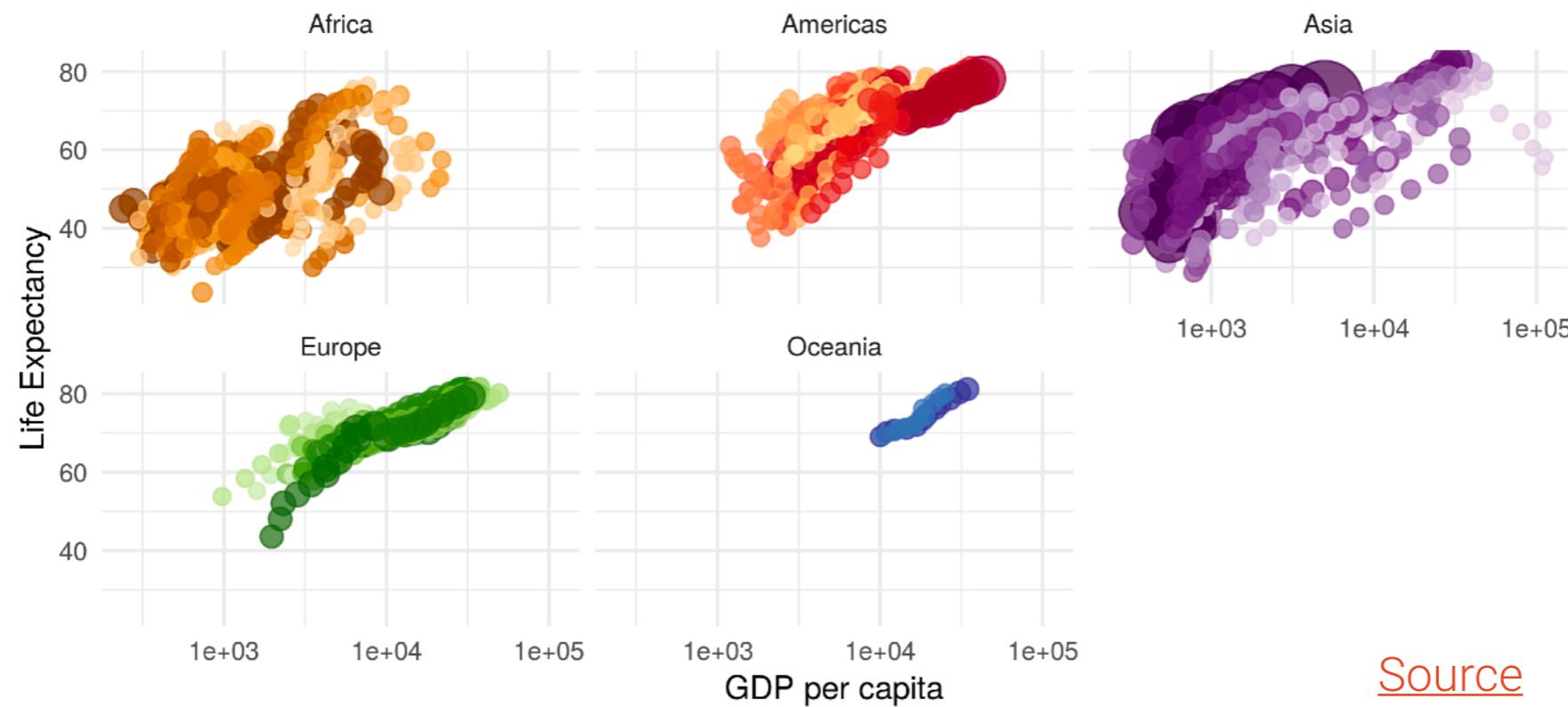


 hm_animated.gif V2 [redacted]	Apr 15, 2015 by James Jos...	24.1 MB [redacted]
---	------------------------------	--

Code Guide for animate

GDP per Capita vs. Life Expectancy

Gapminder data across all years



```
# Stagnate ggplot2
```

```
# Display all points in time ...
```

```
# install.packages("gapminder")
```

```
library("gapminder")
```

```
ggplot(gapminder, aes(gdpPercap, lifeExp, size = pop, colour = country)) +
```

```
  geom_point(alpha = 0.7, show.legend = FALSE) +
```

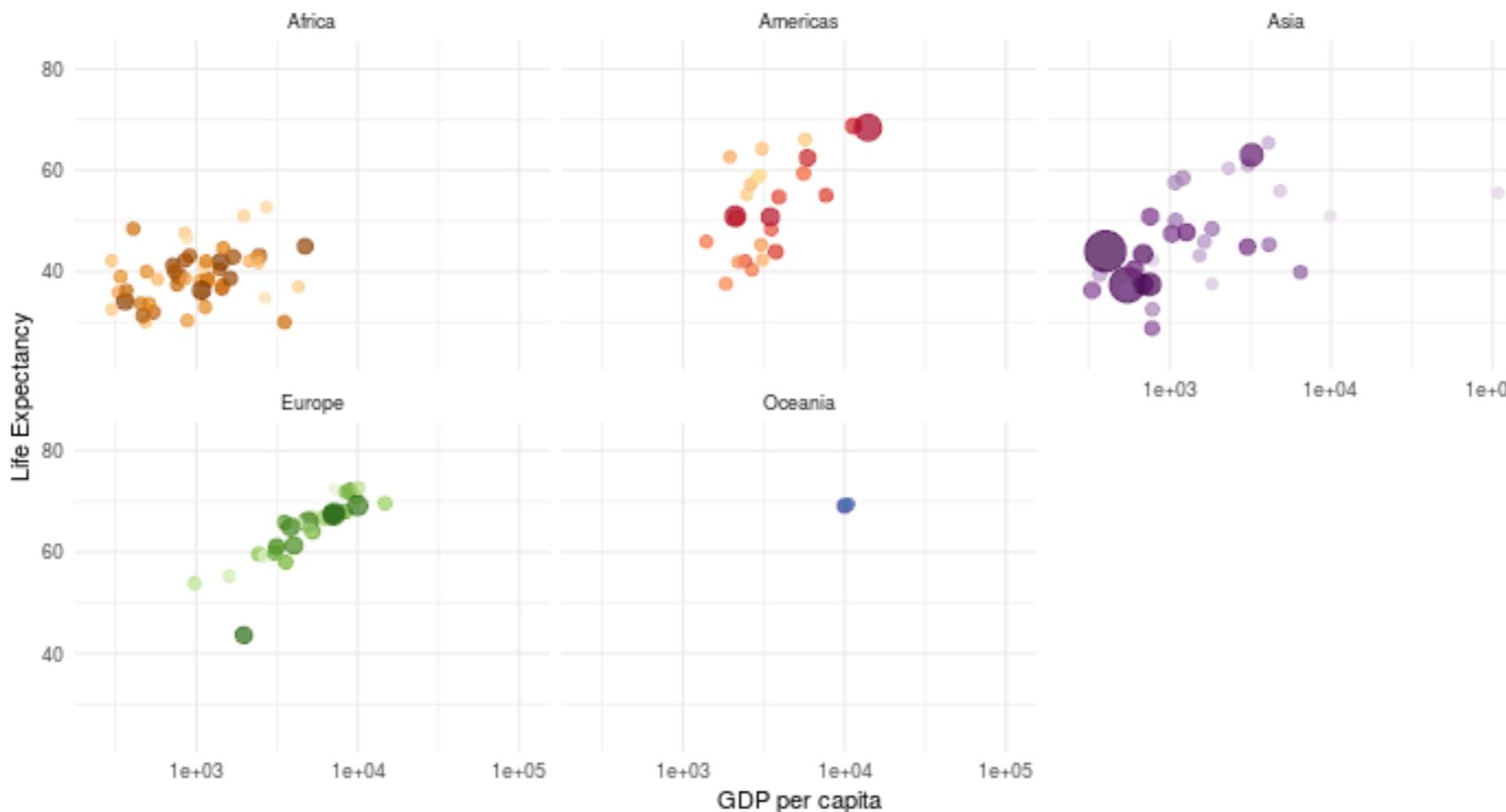
```
  scale_colour_manual(values = country_colors) +
```

```
  scale_size(range = c(2, 12)) +
```

```
  scale_x_log10() +
```

```
  facet_wrap(~continent)
```

Year: 1952



```
# Animate Plot!
# install.packages("devtools")
# devtools::install_github("thomasp85/gganimate")
library("gganimate")
ggplot(gapminder, aes(gdpPercap, lifeExp, size = pop, colour = country)) +
  geom_point(alpha = 0.7, show.legend = FALSE) +
  scale_colour_manual(values = country_colors) +
  scale_size(range = c(2, 12)) +
  scale_x_log10() +
  facet_wrap(~continent) +
  labs(title = "Year: {frame_time}", x = "GDP per capita", y = "Life Expectancy") +
  transition_time(year) +
  ease_aes("linear")
```

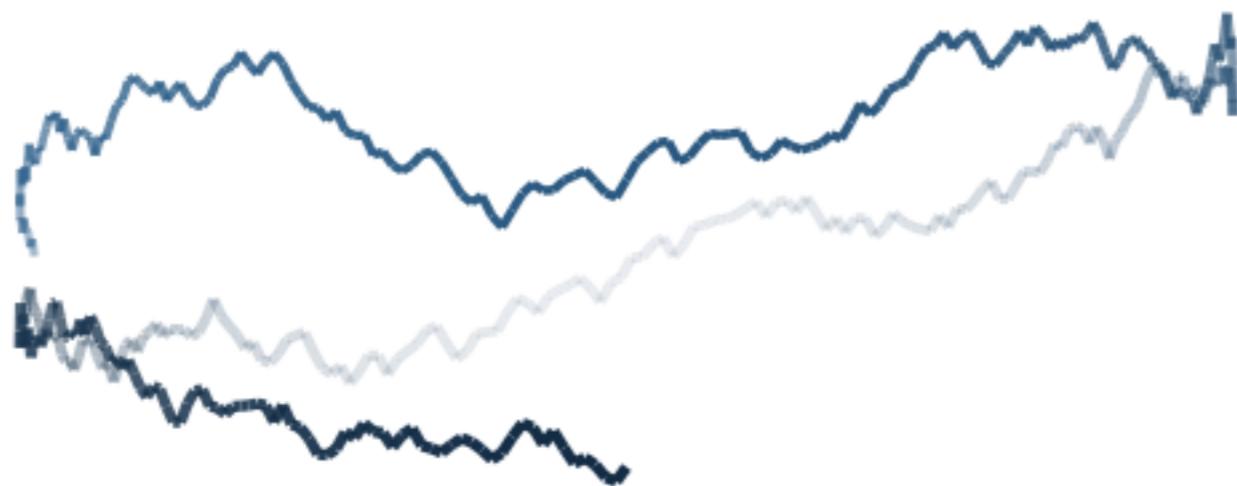
New Grammar

... gganimate focuses on solving animation ...

- **transition_***() defines how the data should be spread out and how it relates to itself across time.
- **view_***() defines how the positional scales should change along the animation.
- **shadow_***() defines how data from other points in time should be presented in the given point in time.
- **enter_*/exit_***() defines how new data should appear and how old data should disappear during the course of the animation.
- **ease_aes()** defines how different aesthetics should be eased during transitions.

Seasonality

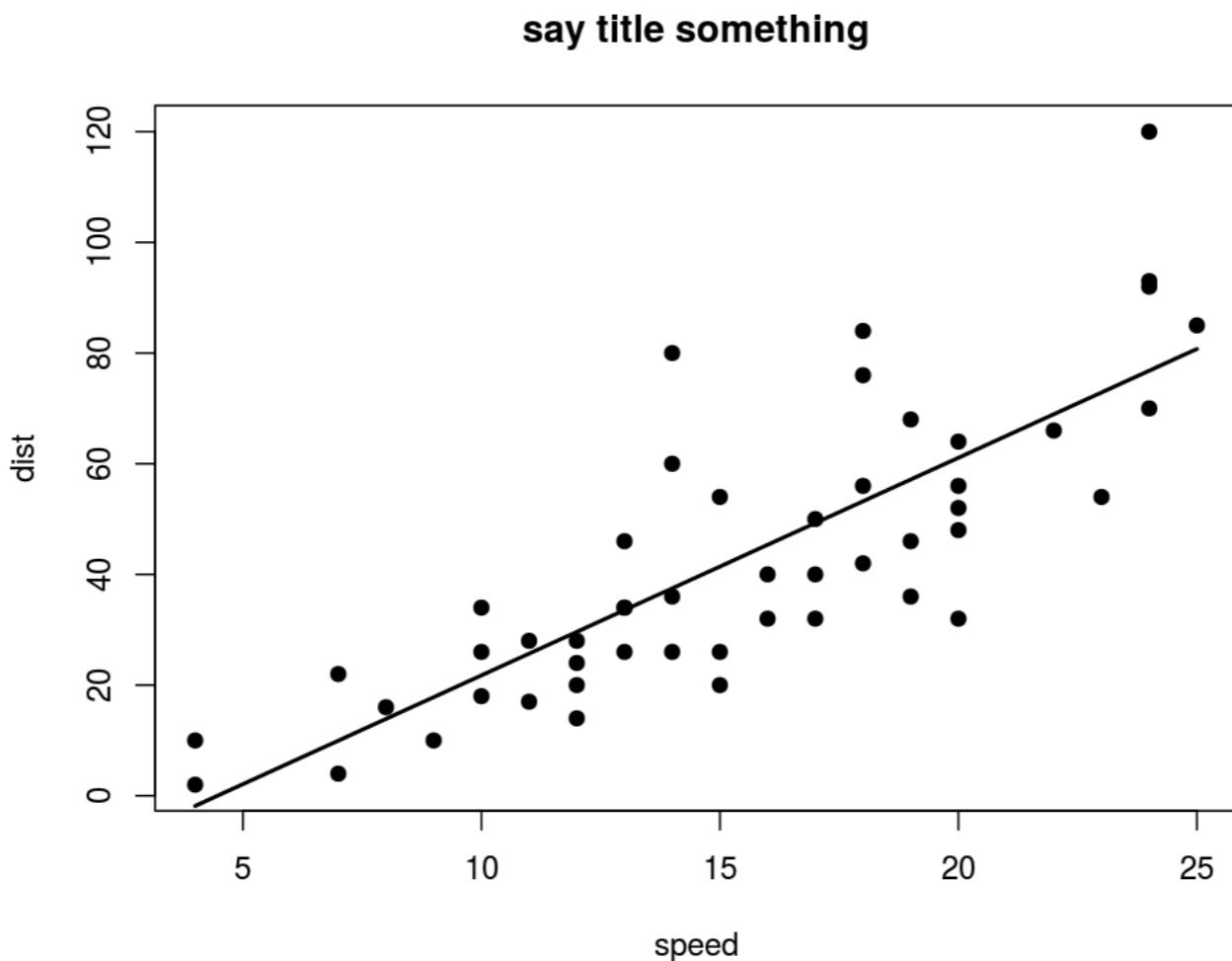
... viewing seasonal patterns ...



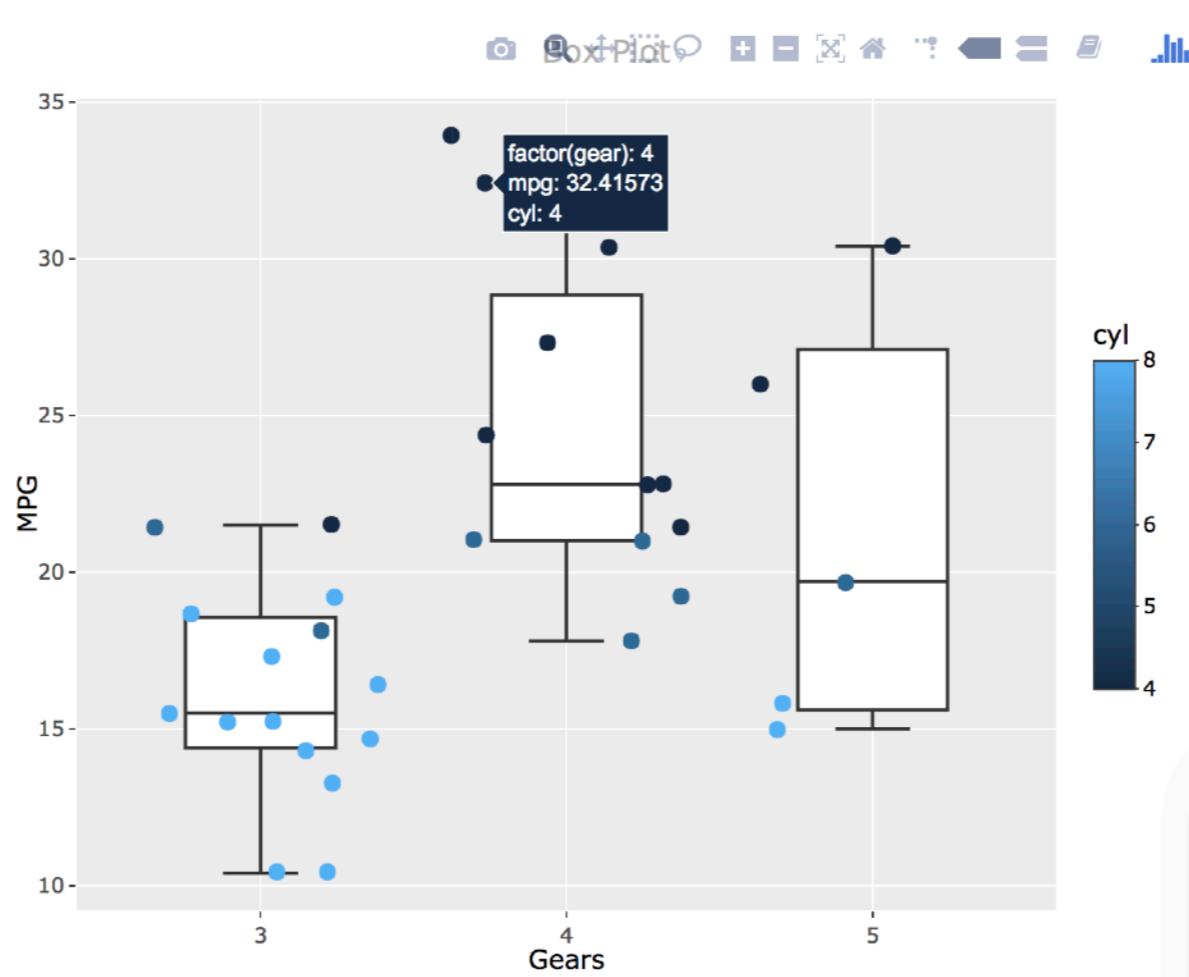
[Source](#)

Interactive

Voice Activation



You are recommended to use Google Chrome to play with this app. To change the title, say something that starts with "title", e.g. "title I love the R language", or "title Good Morning". To change the color of points, say something that starts with "color", e.g. color "blue", or color "green". When the app is unable to recognize the color, the points will turn gray. To add a regression line, say "regression". To make the points bigger or smaller, say "bigger" or "smaller".



Interactive Graphics
Searching the Data Live

```
# install.packages("plotly")
library("plotly")
p = ggplot(mtcars, aes(x = factor(gear), y = mpg, color = cyl)) +
  geom_boxplot() +
  geom_jitter(size = 2) +
  labs( title = "Box Plot", x = "Gears", y = "MPG" )
ggplotly( p )
```

Recap

- **Designing a Graphic**
 - Emphasis the data's narrative.
 - Be ware of *Simpson's paradox*, *Apophenia*, and lying with graphics.
- **CRAP**
 - **C**ontrast, **R**epetition, **A**lignment, **P**roximity
 - Tenets of Gestalt Design
- **Chart Junk**
 - Useless embellishment on the plot that impacts clarity of plot
- **Modern Graphics**
 - Ability to modify display or view data over time.

Acknowledgements

Acknowledgements

- Thomas Lin Pedersen's ganimate package talk at UseR! 2018
 - <https://www.youtube.com/watch?v=21ZWDrTukEs>
 - <https://github.com/thomasp85/ganimate>
- Di Cook's Ihaka Lecture
 - <http://www.dicook.org/files/ihaka/#1>

This work is licensed under the
Creative Commons
Attribution-NonCommercial-
ShareAlike 4.0 International
License

