



/ ANITA  
B.ORG

20<sup>th</sup>

GRACE HOPPER CELEBRATION  
V I R T U A L



/ANITA  
B.ORG

20<sup>th</sup>

# GRACE HOPPER CELEBRATION

VIRTUAL

---

Beyond Imagination:  
The Secret to Designing Beautiful,  
Creative Custom Visuals

# WORKSHOP TEAM



AMBER

HOAK

[SHE/HER]

Presenter



KATE

LYTVYNETS

[SHE/HER]

Interaction  
Manager



DAVID

TITTSWORTH

[HE/HIM]

Q&A  
Manager



Fork of Data Visualization Workshop: Designing Data Visuals that Sparkle by

## D3 Workshop

Designing Data Visuals that Sparkle

Adding some  to your charts

<http://aka.ms/ghc19workshop>

Please use either FireFox, Edge, or Safari for this notebook. Chrome browsers don't work properly in Observable  (this includes the

Fork of Data Visualization Workshop: Designing Data Visuals that Sparkle by

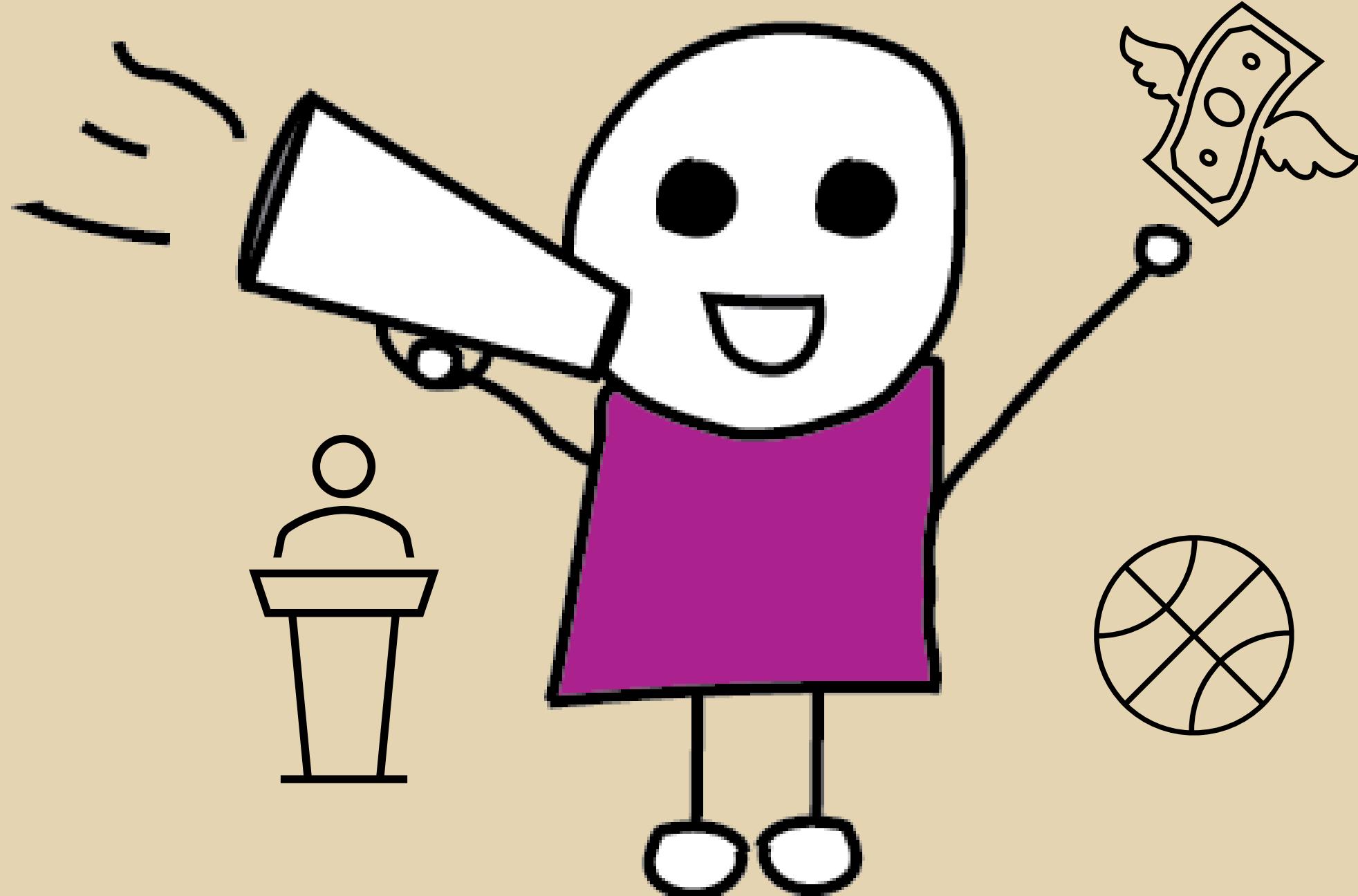
## D3 Workshop

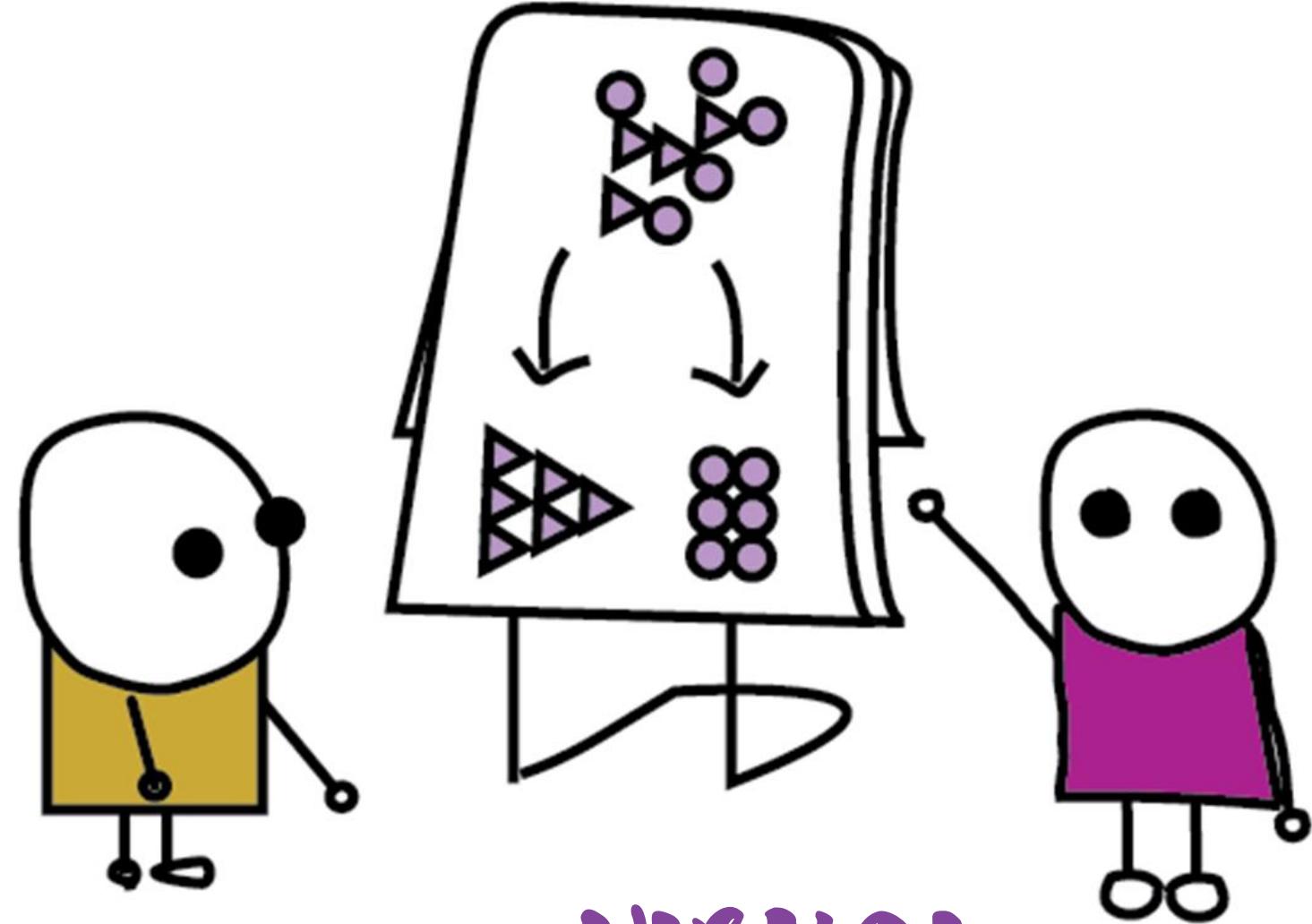
Designing Data Visuals that Sparkle

Adding some  to your charts

<http://aka.ms/ghc19workshop>

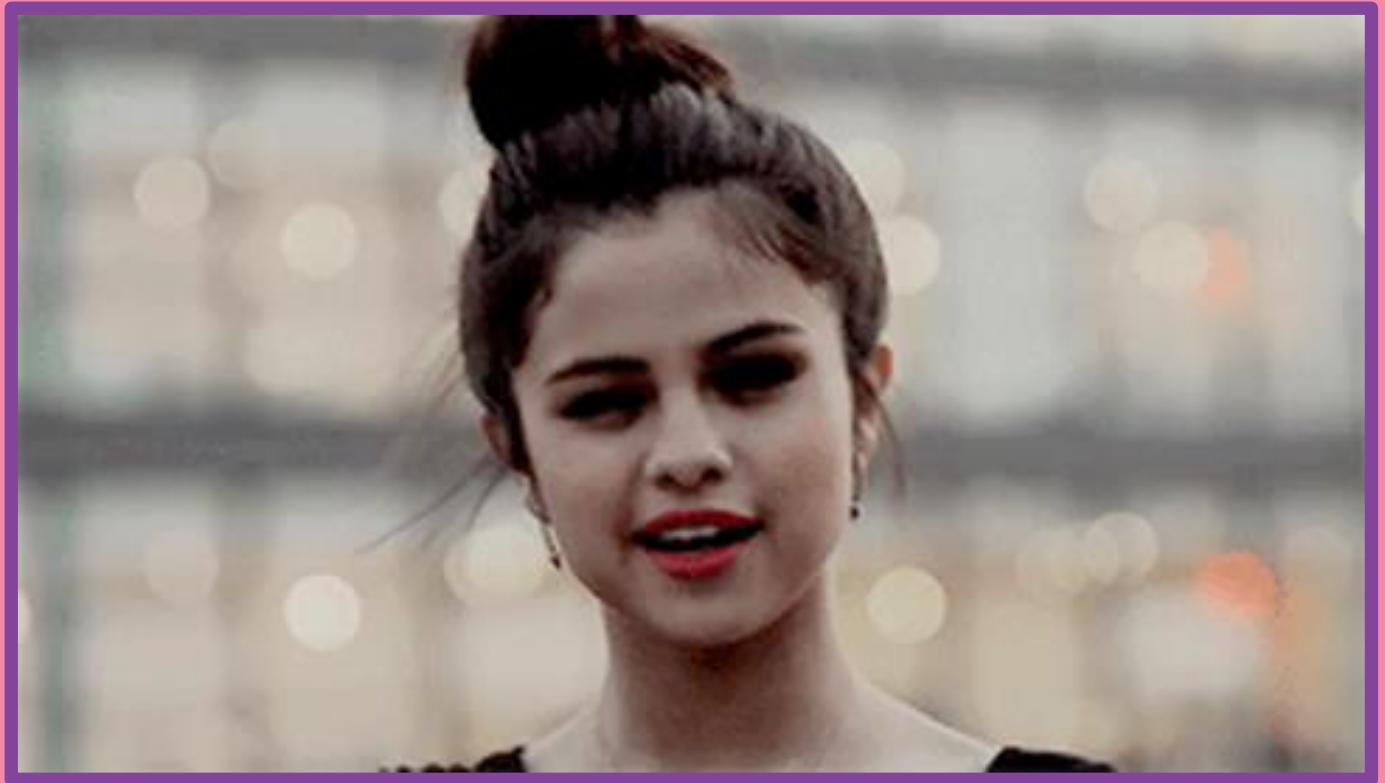
Please use either FireFox, Edge, or Safari for this notebook. Chrome browsers don't work properly in Observable  (this includes the





VISUAL  
COMMUNICATION

# PICTURE SUPERIORITY EFFECT







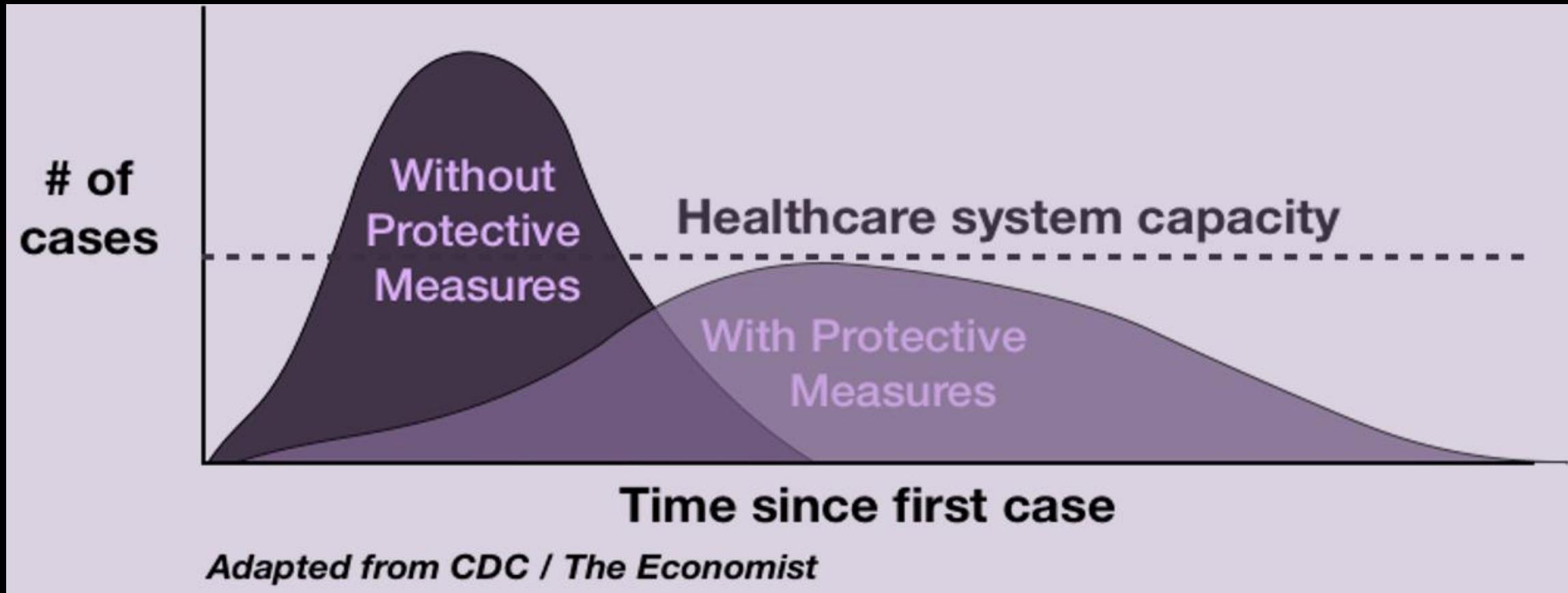
SO LONG, PARTNER

# Data Visualization

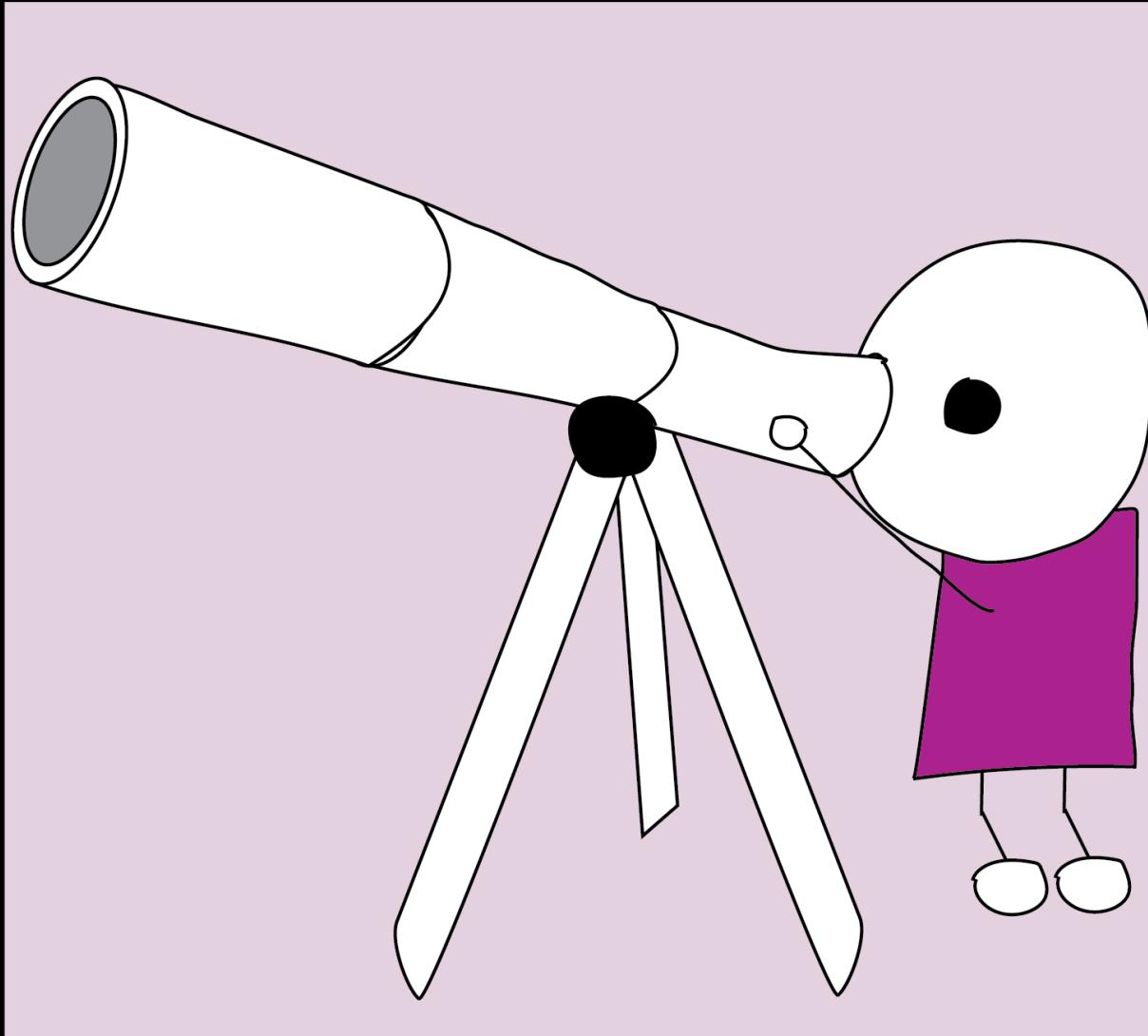


# Visual Communication

# FLATTEN THE CURVE

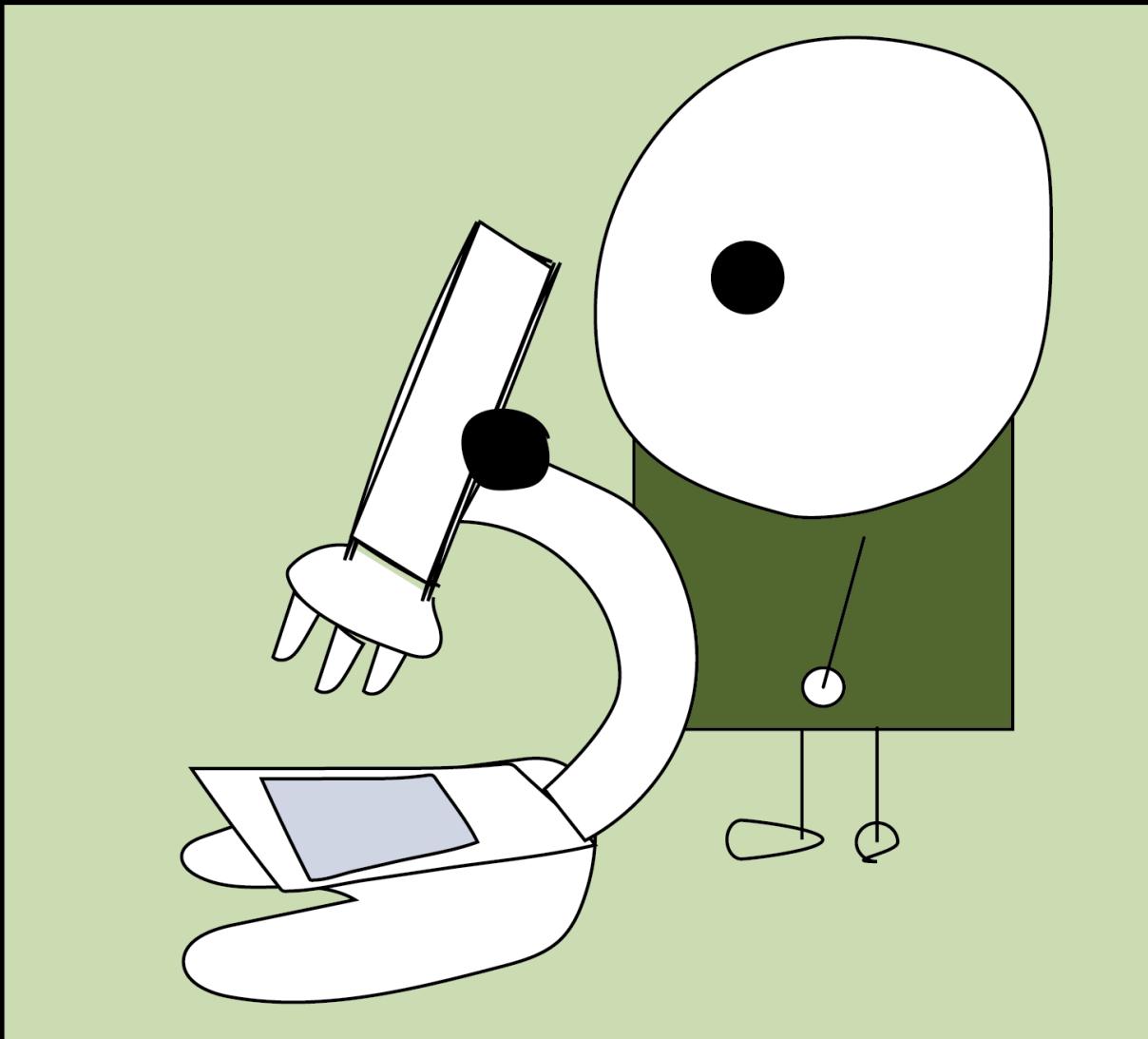


# Telescope

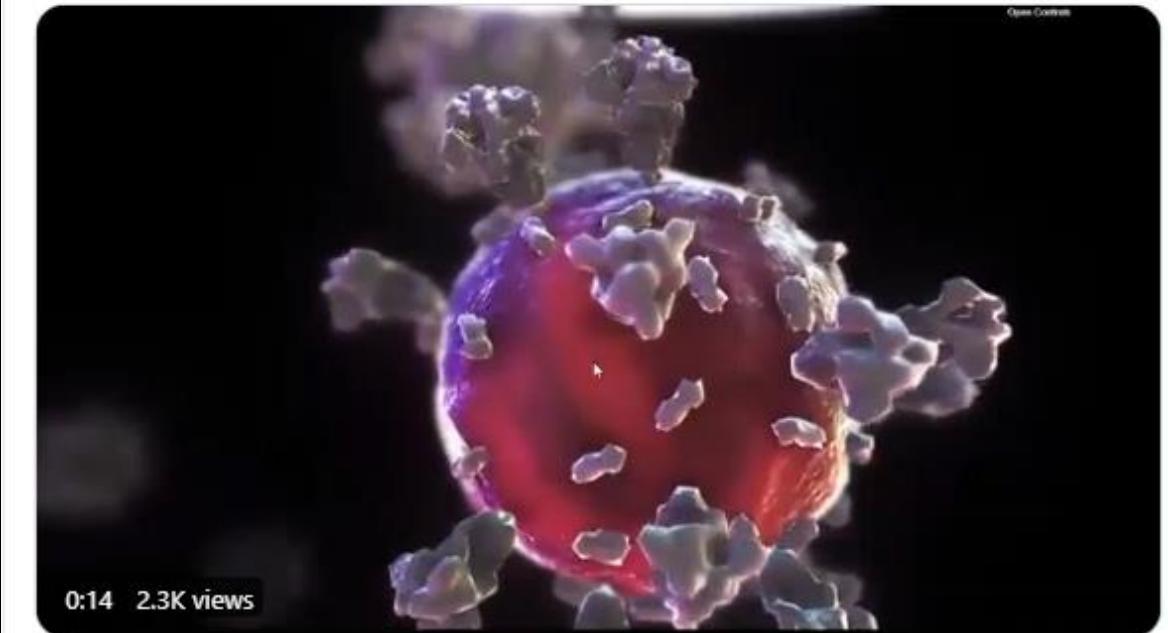


Source: <https://nbremer.github.io/planet-globe/>

# Microscope



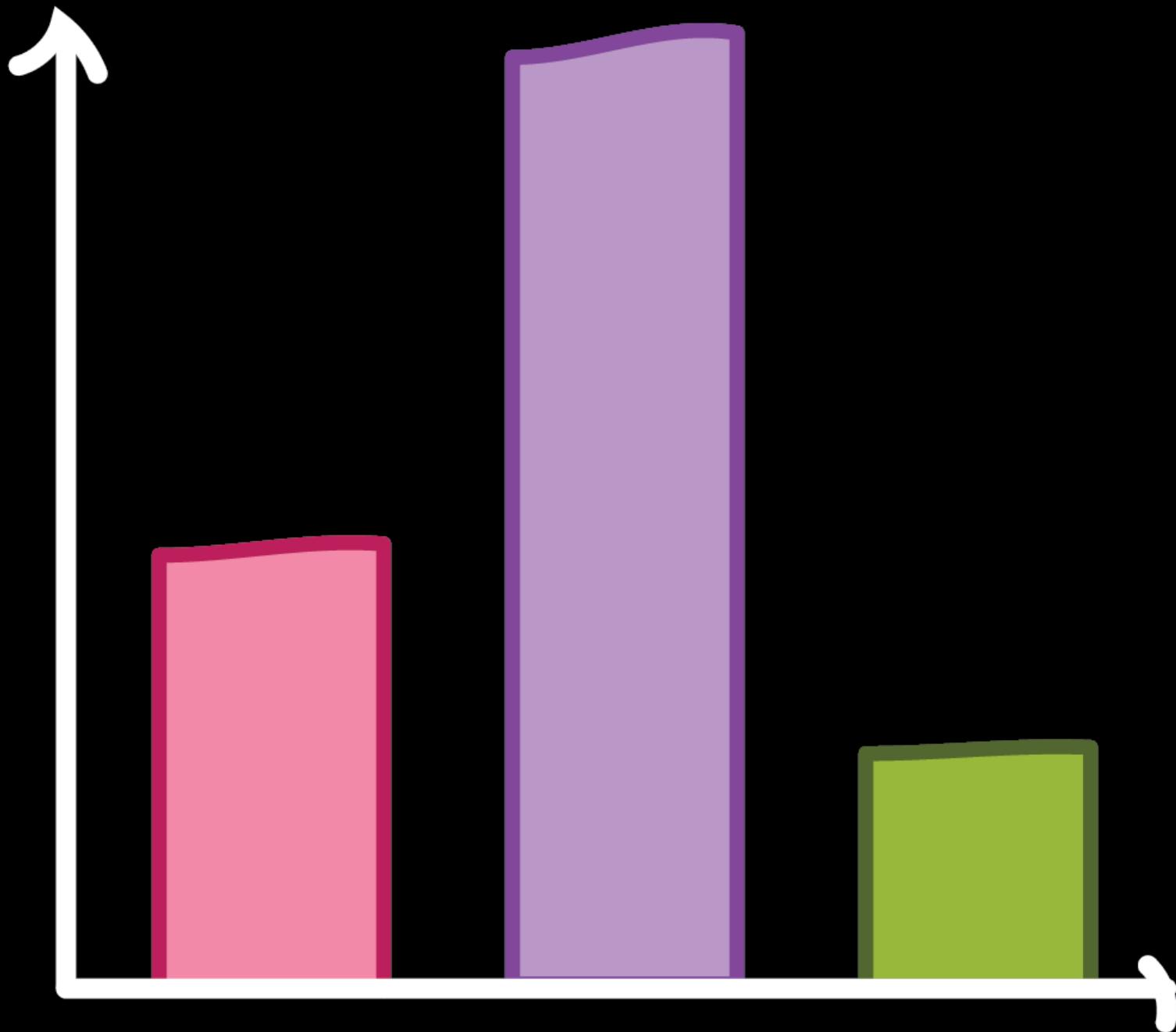
refraction look. The S-protein is "inspired" by the  
@jeroen\_claus's @entagma tutorial. #sciar #threejs  
#webgl #COVID19



6:11 AM · Jul 2, 2020 · Twitter Web App

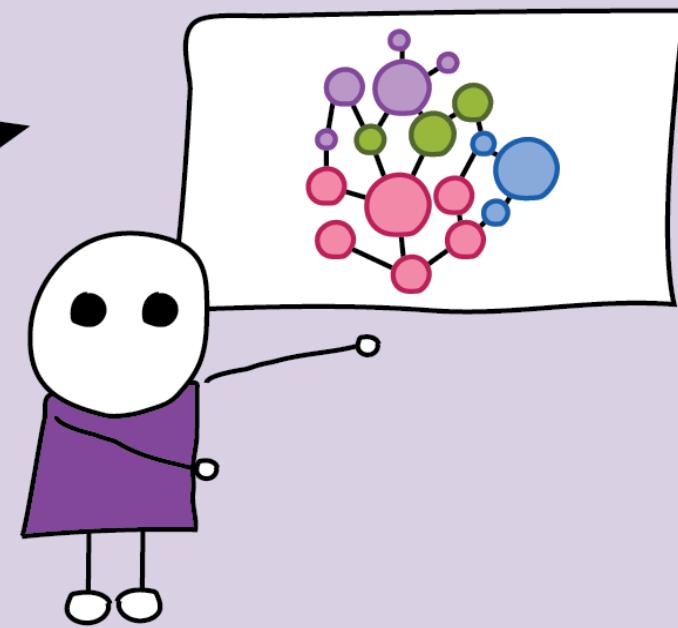
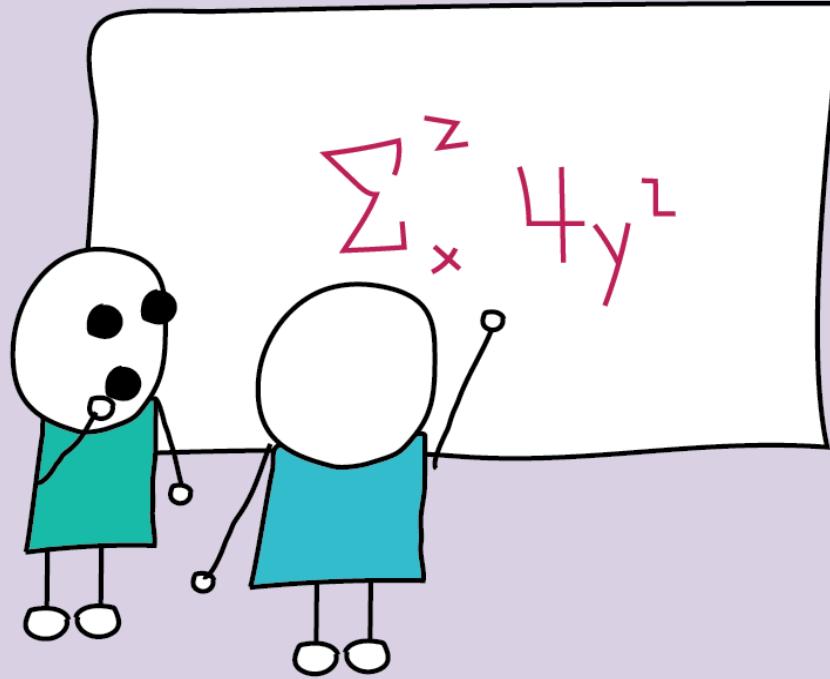
Source:

<https://twitter.com/edankwan/status/1278677798744403968>

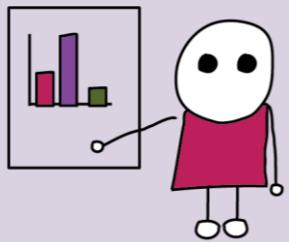


A network graph visualization on a dark background. The graph consists of numerous small, colorful circular nodes and a dense web of thin white lines representing connections between them. Superimposed on this network are the letters of the word "WORLD". Each letter is rendered in a 3D wireframe style, with vertical bars of varying heights. The letters are positioned along a diagonal path from the top-left towards the bottom-right. The "W" is at the top left, followed by "O", "R", "L", and "D" at the bottom right.

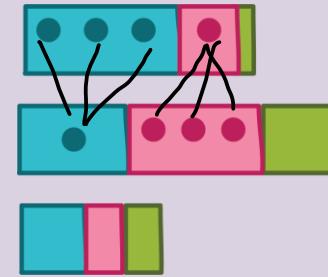
WORLD



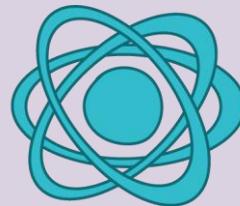
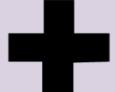
## OUT OF THE BOX SOLUTION



## CUSTOM SOLUTION



SVG



**DEFINE PURPOSE**

**DESIGN & DATA  
LITERACY**

**APPLY DESIGN  
CONCEPTS**

**DESIGN IN CODE**



Follow along  
with us!



Amber Hoak  · [linkedin.com/in/amberhoak](https://linkedin.com/in/amberhoak)

 Link shared Sep 22



# Custom Data Visualization Workshop

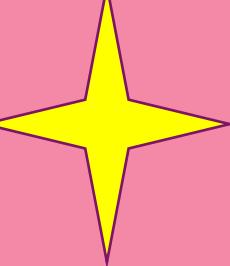
Hi! Welcome to our workshop. During the presentation, we will go through several little activities. These are broken up into several notebooks. We will guide you through each notebook but if you want to skip ahead, we wont stop you 😊

## Activity 1.

### Create a Monster Badge

Answer the mini questionnaire to make a custom visual that relates to you!

# Test it out!



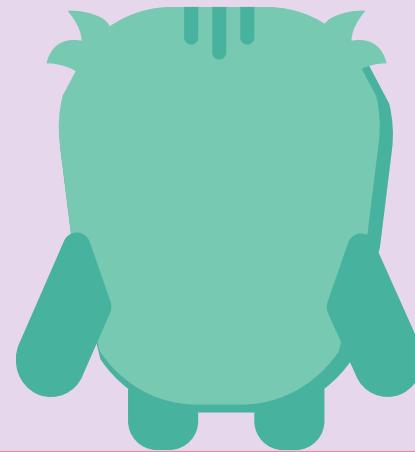
## ACTIVITY #1

<https://observablehq.com/@ahoak/create-a-monster-badge>

STUDENT



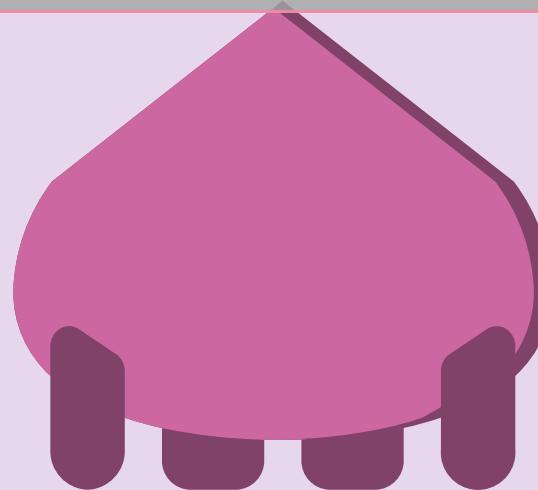
ENGINEER



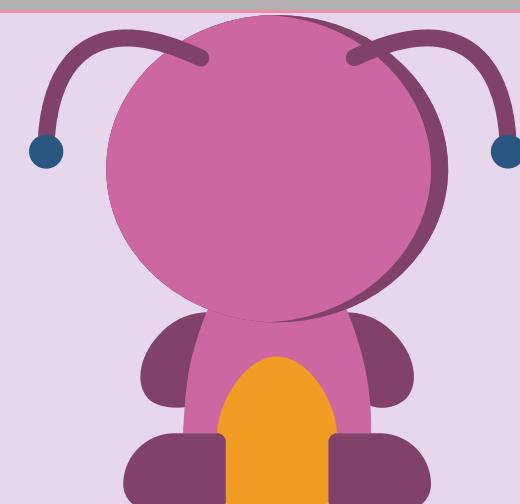
PM



BODY



DATA SCIENTIST

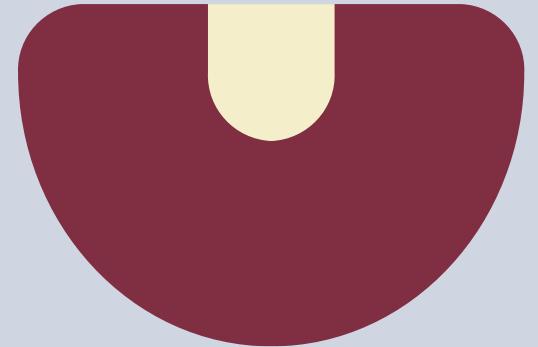
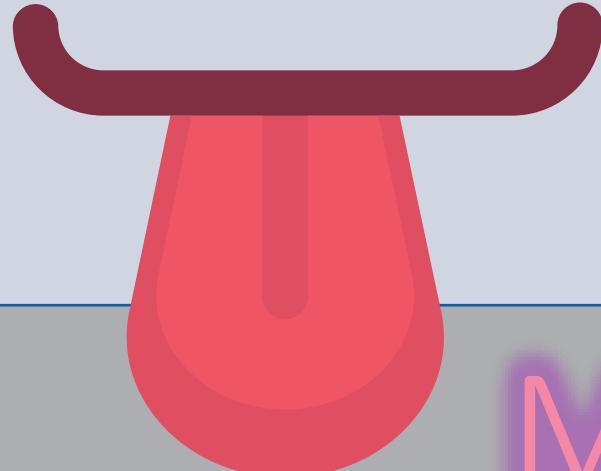


OTHER

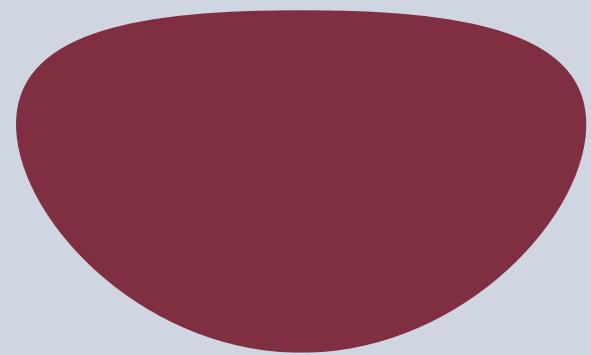
WEST COAST USA

SOUTHWEST USA

EAST COAST  
USA



MOUTH

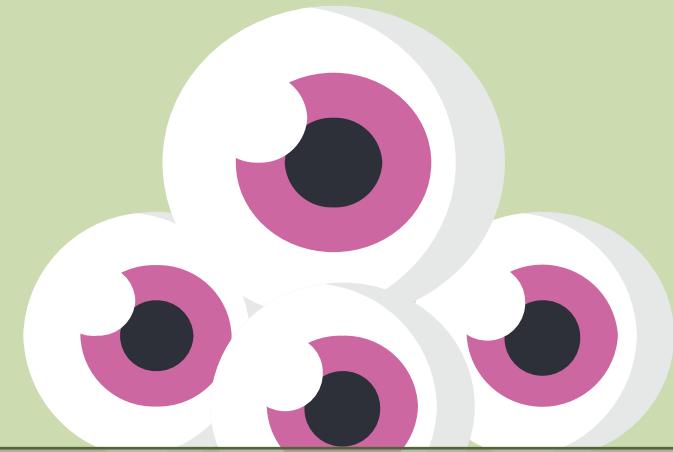
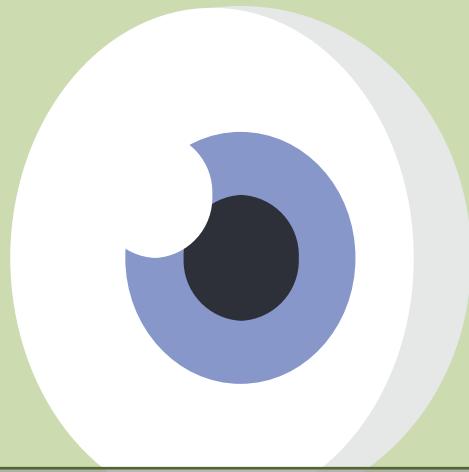


GLOBAL

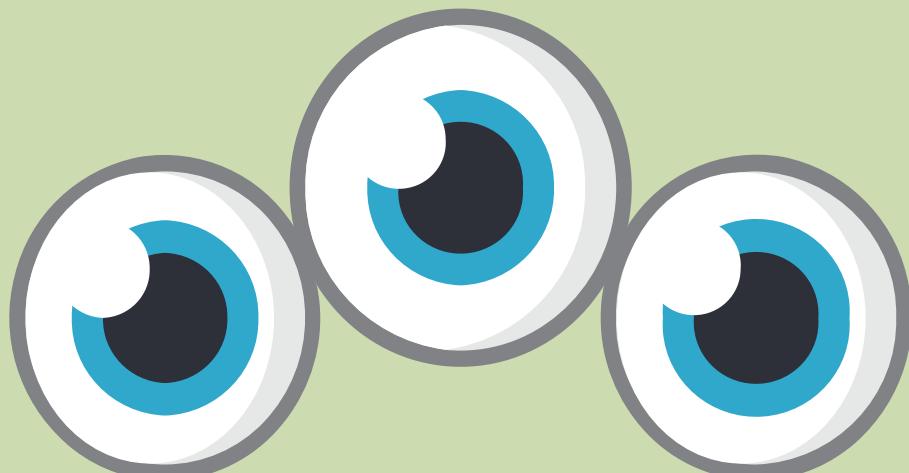
MIDWEST USA

JAVASCRIPT

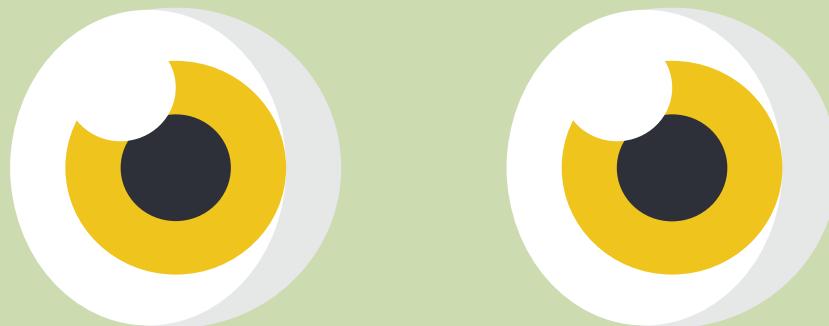
OTHER/NONE



EYES



JAVA/C++



PYTHON



Amber Hoak · [linkedin.com/in/amberhoak](https://linkedin.com/in/amberhoak)

Published Aug 28

# Create A Monster Badge



What is your occupation?

- Student
- Engineer
- PM
- Data Scientist
- Other



Where are you located?

- East Coast USA
- West Coast USA
- Midwest USA
- Southwest USA
- Global

What is your language of choice?

- JavaScript
- Python
- Java/C++
- Other/None

See your personal monster here:

Amber



Kate



David



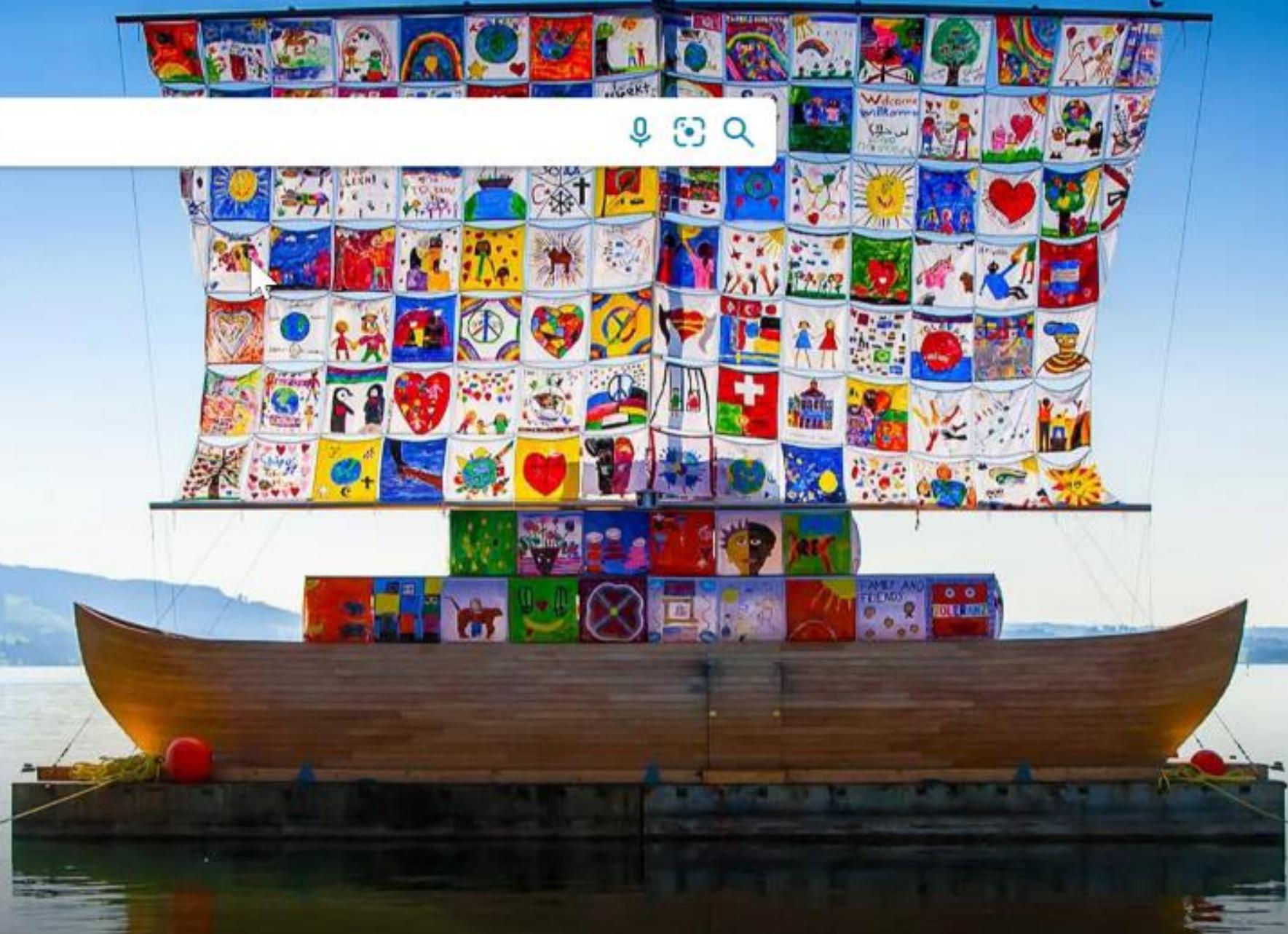
tidbits.fyi

# Yankee Candle: Making Scents of Emotions



Ever get a whiff of something that hits you with a powerful memory out of the blue? Maybe snickerdoodle cookies, fresh-cut grass, or even a scent you can't quite put a name to. No matter the specifics, something about a scent from the past seems to transport your mind back to your grandma's kitchen or childhood backyard more vividly than memories

**Goal:**  
**Impact of COVID-19 on**  
**American Unemployment**



# *What makes good data?*

- ▶ Is it trustworthy?
- ▶ How was this data collected?
- ▶ Is there enough data support your goal?
- ▶ Are there any inconsistencies?

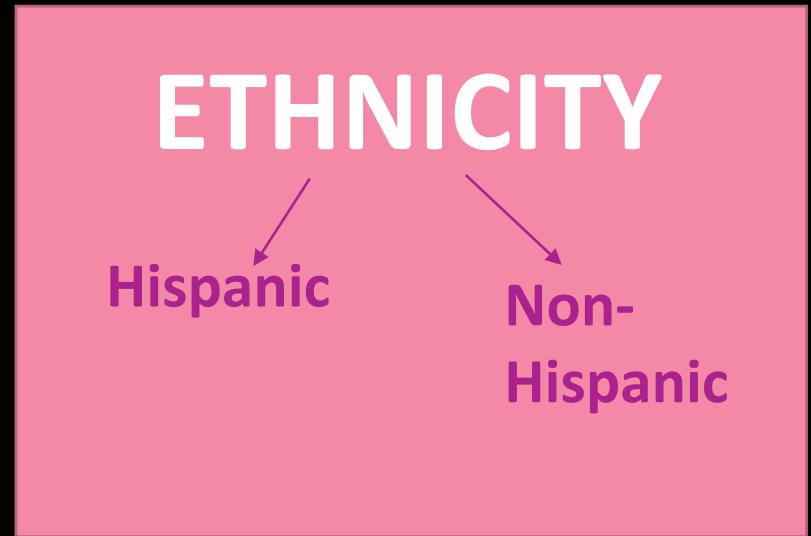
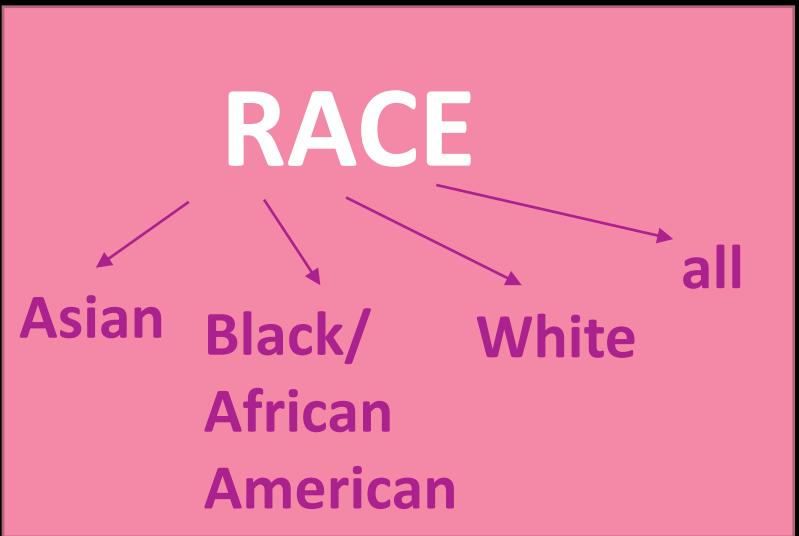
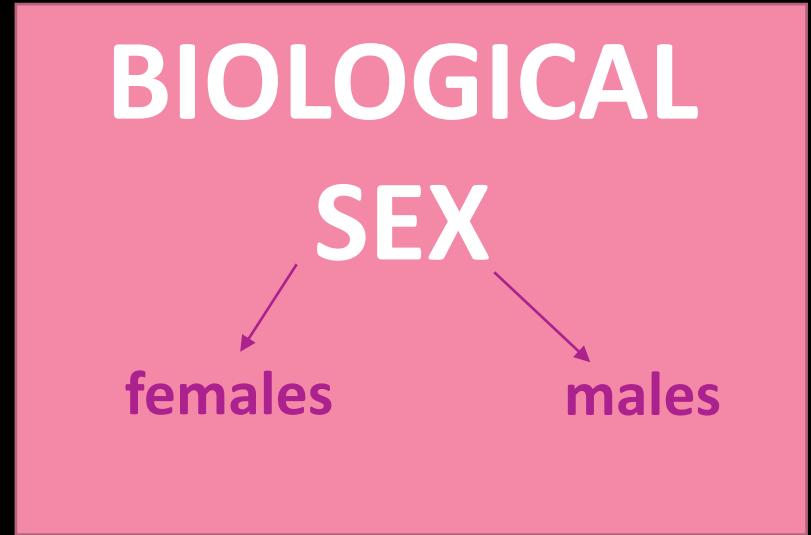
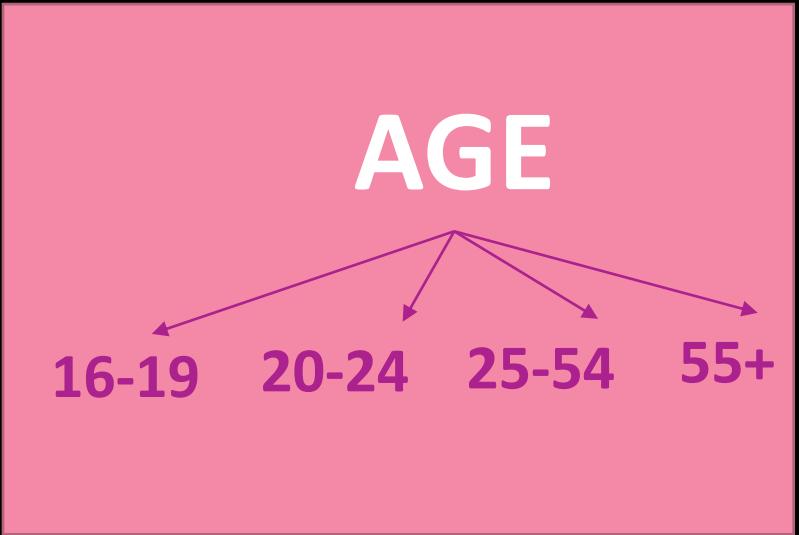
Employment-  
Population Rate

Unemployment Rate

Looking for Full-Time  
Employment

Looking for Part-Time  
Employment

Not in the Labor Force

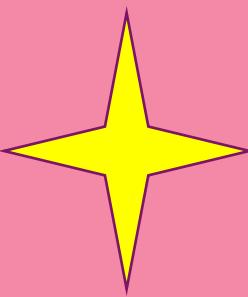


We want our  
visual...

LOOK AT THE WHOLE  
PICTURE

VIEW SIMPLE AND  
COMPLEX DATA

SEE ABNORMALITIES,  
TRENDS,  
DEPENDENCIES



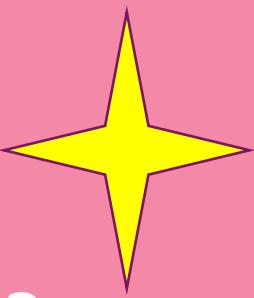
# Define the Purpose

Activity #2

# United States Unemployment Data Exploration

- All data collected from the US Bureau of Labor Statistics from the Current Population Survey and all are non-seasonally adjusted. Choose a dataset to explore facets within the data. Datasets collected include:
  - Employment-population ratio:** The employment-population ratio represents the number of employed people as a percentage of the civilian noninstitutional population. In other words, it is the percentage of the population that is currently working. The employment-population ratio is calculated as:  $(\text{Employed} \div \text{Civilian Noninstitutional Population}) \times 100$ .
  - Not in the Labor Force:** Persons who are neither employed nor unemployed are not in the labor force. This category includes retired persons, students, those taking care of children or other family members, and others who are neither working nor seeking work.
  - Unemployed Looking for Full-Time Work:** Persons who are unemployed looking for full time employment (35+ hours a week).

# *Define the Purpose*



- ▶ What is being represented in the data?  
What is not?
- ▶ Who is your target audience?
- ▶ What would you like to show?  
What are facets of interest?
- ▶ Do we have enough data?

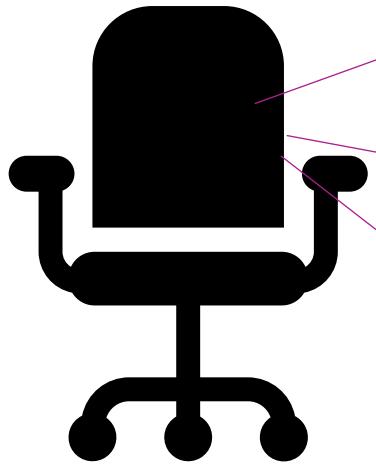
Post in the chat a question about the data you would like answered!

# *Questions*

- ▶ Are people retiring early?
- ▶ How is young population affected?
- ▶ Are there any unexpected trends with unemployment and race?
- ▶ How has the pandemic affected men versus women?

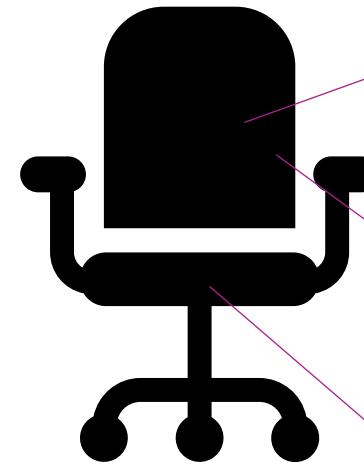
What type of data do we  
have?

# Quantitative



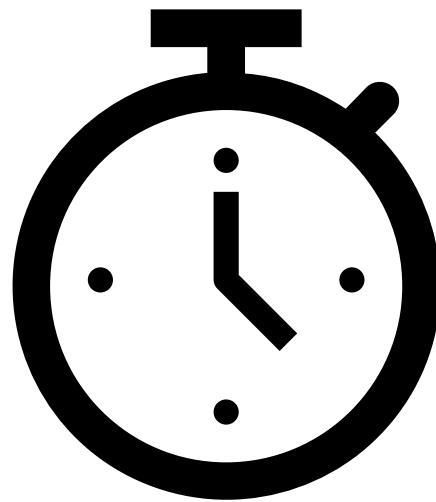
50 lbs.  
3 ft tall  
Sells for  
\$100  
Has 3  
wheels

# Qualitative



Made of  
Cloth  
Built in  
Italy  
Is colored  
black

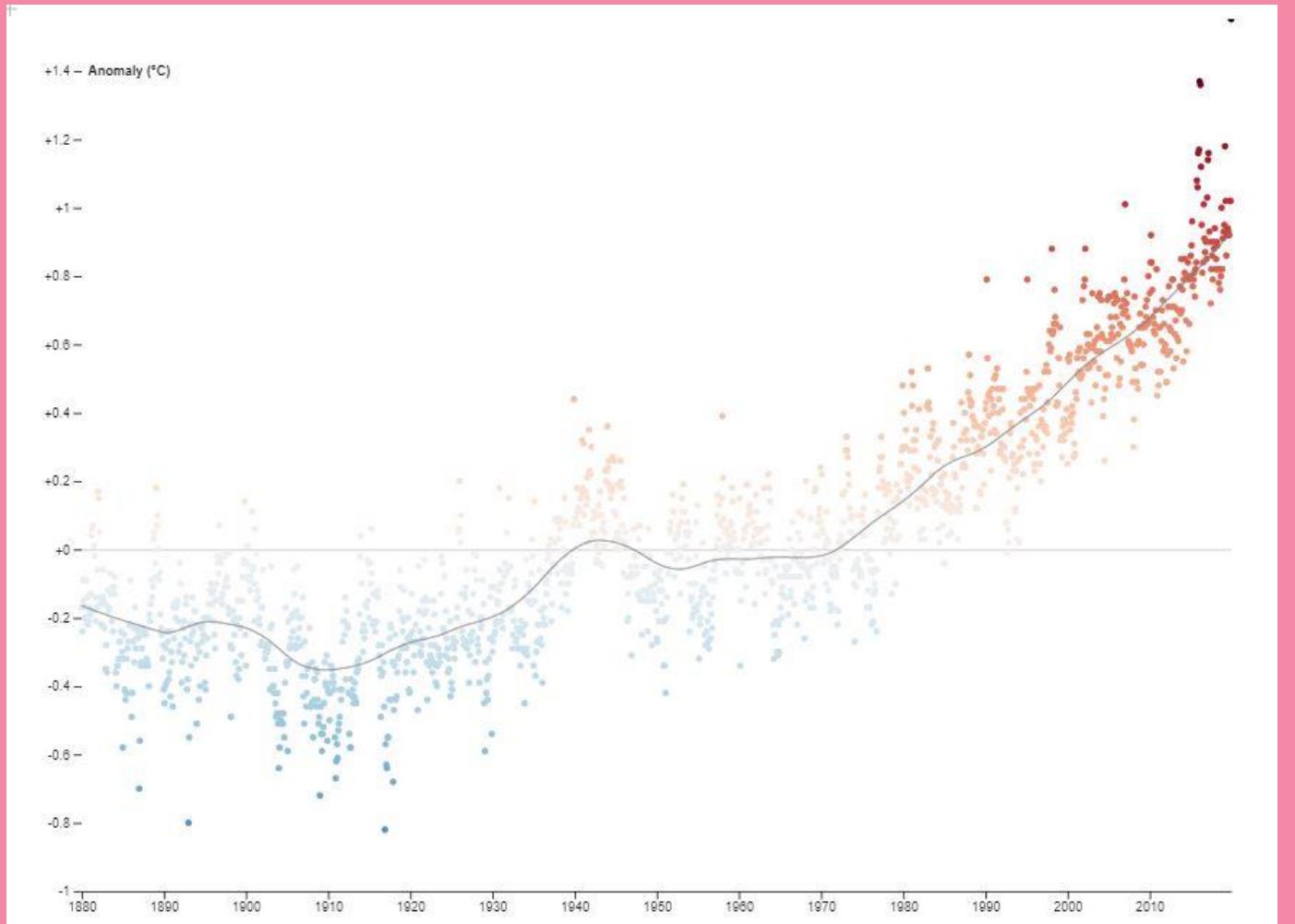
# Continuous



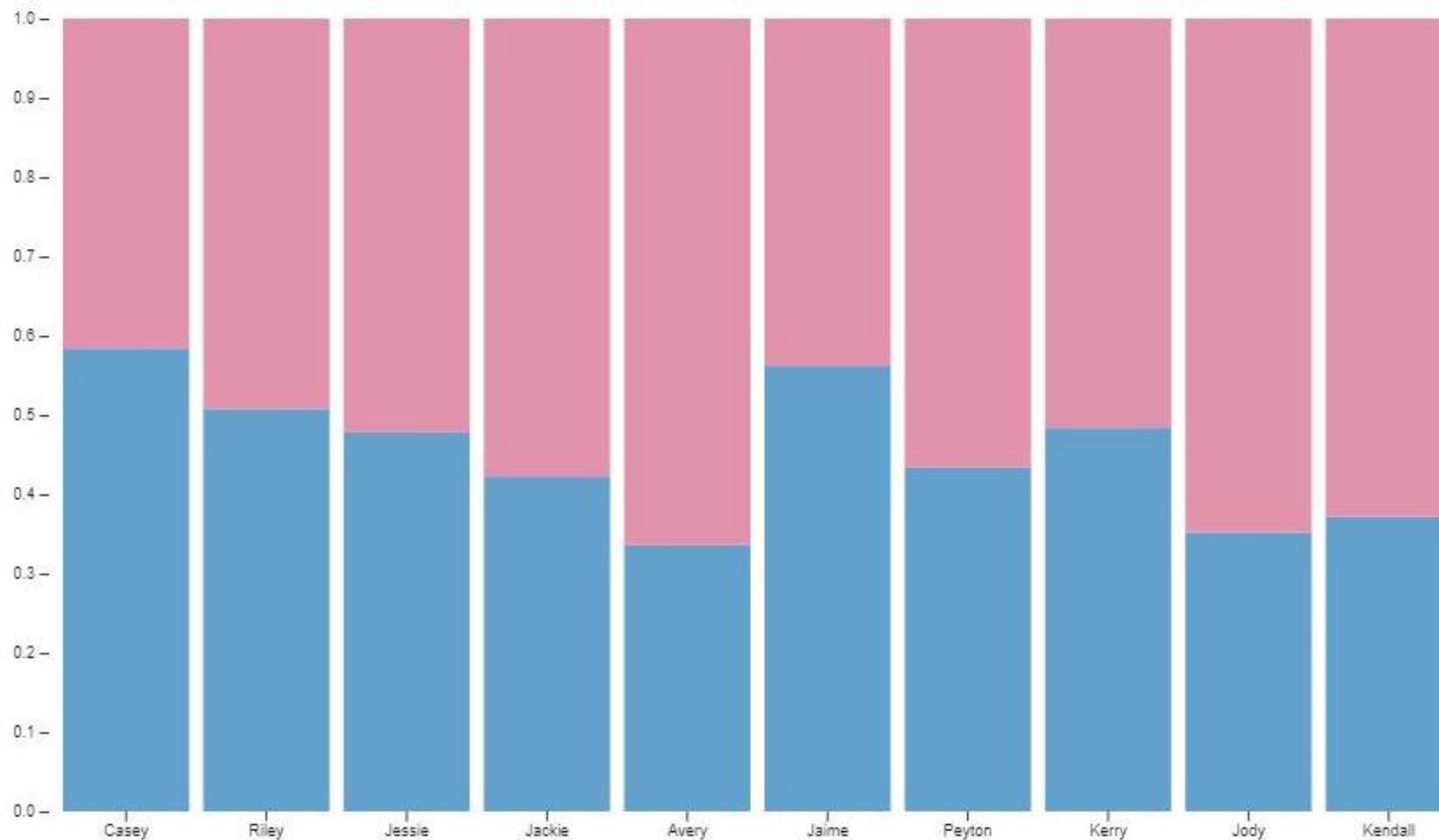
# Discrete

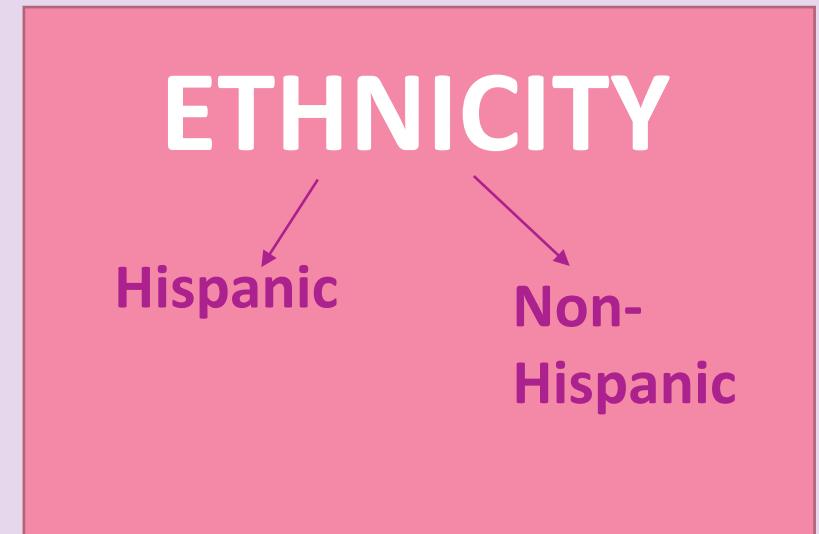
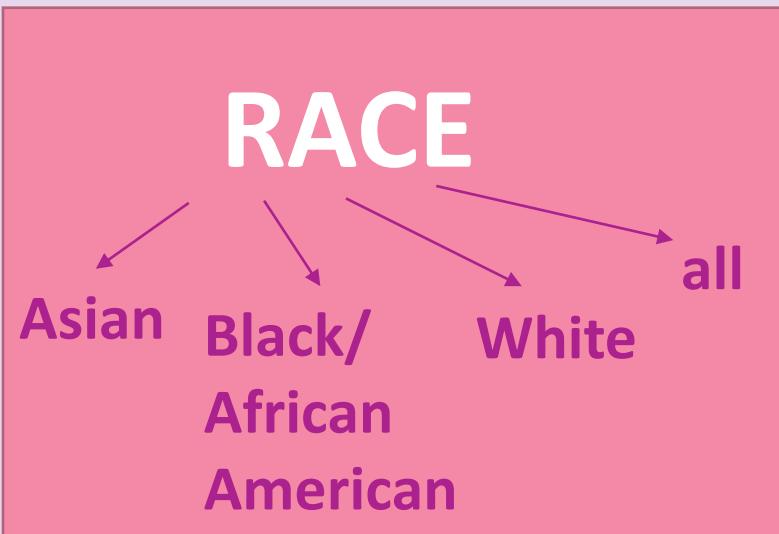
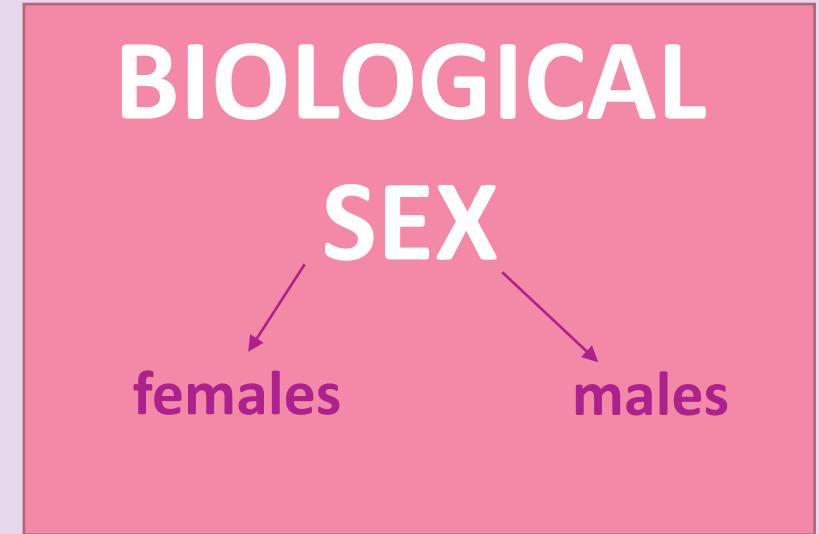
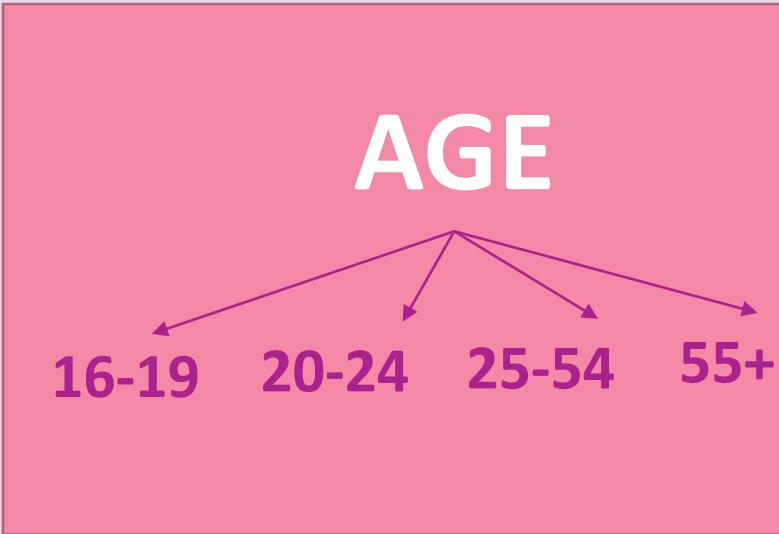


# Global Temperature Trends



# Top Unisex Names



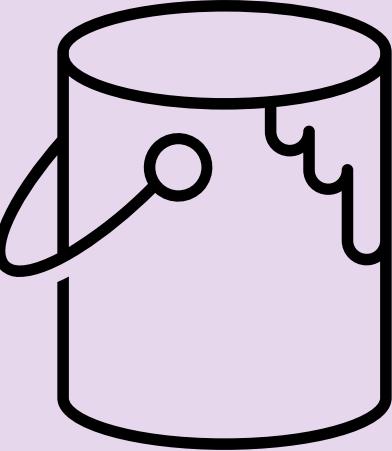


**AGE**

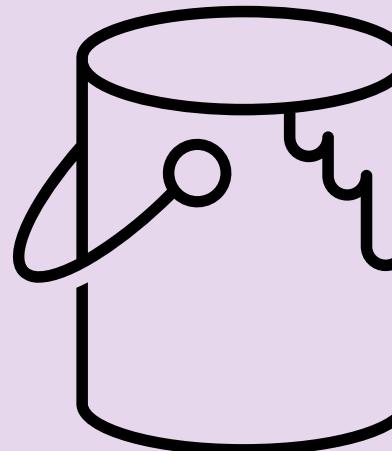
**SEX**

**RACE**

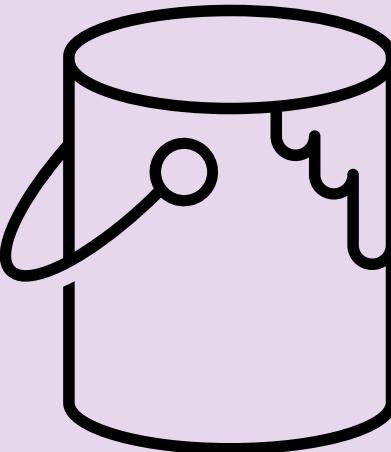
**ETHNICITY**



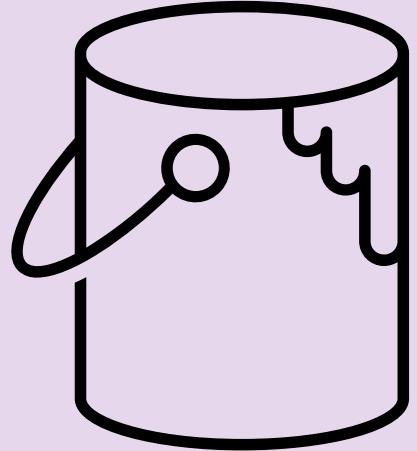
**SEX**



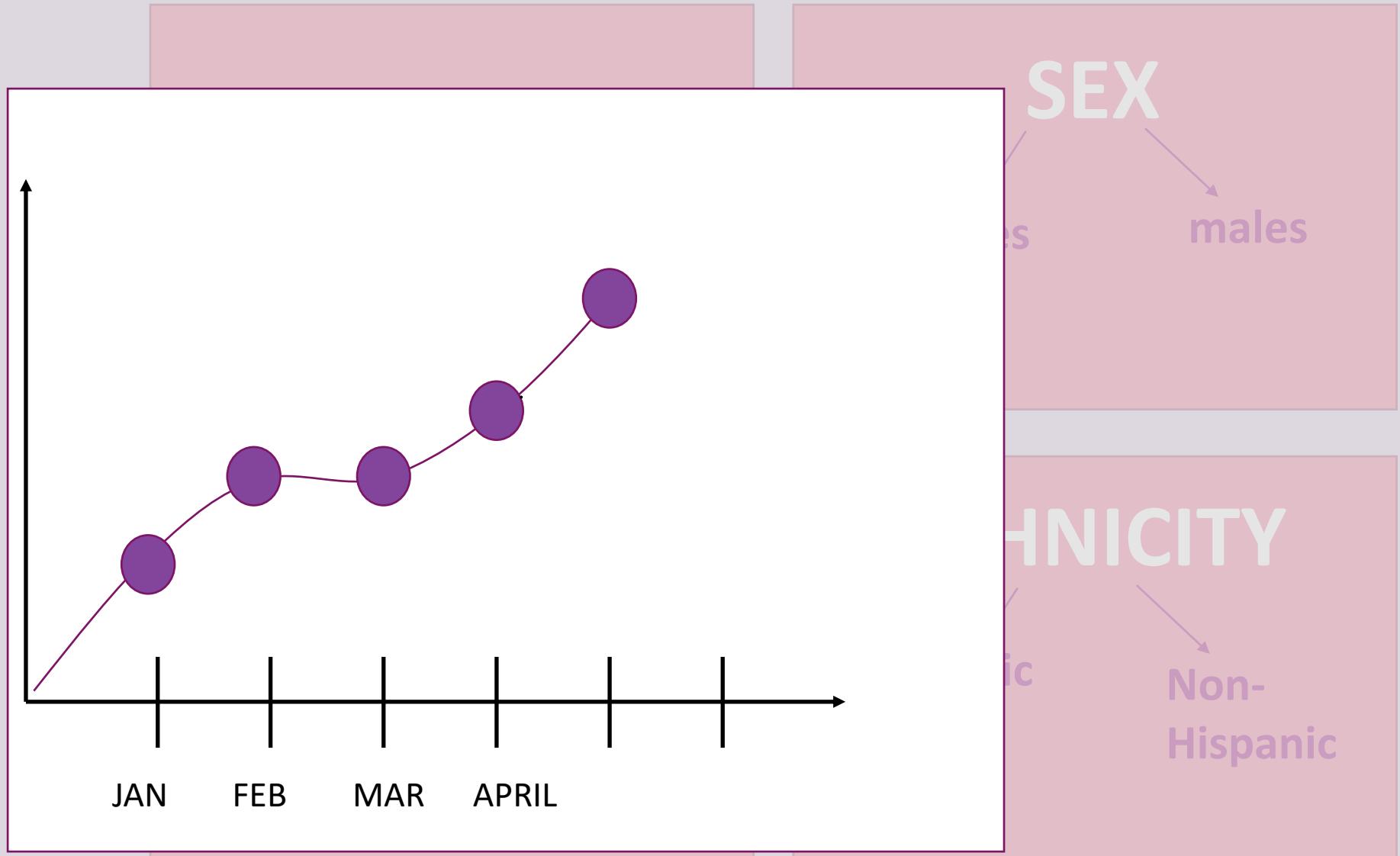
**AGE**



**ETHNICITY**

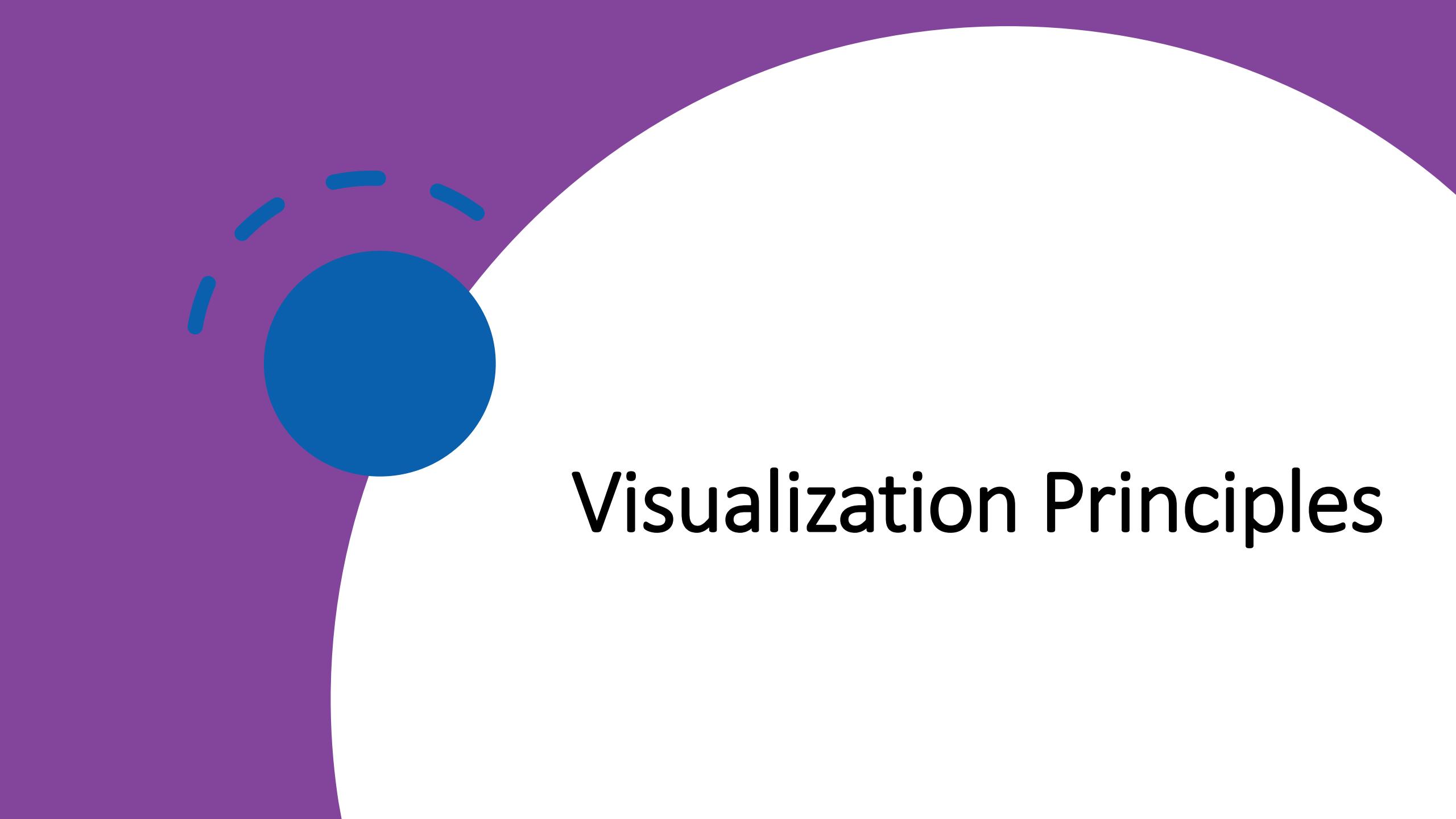


**RACE**

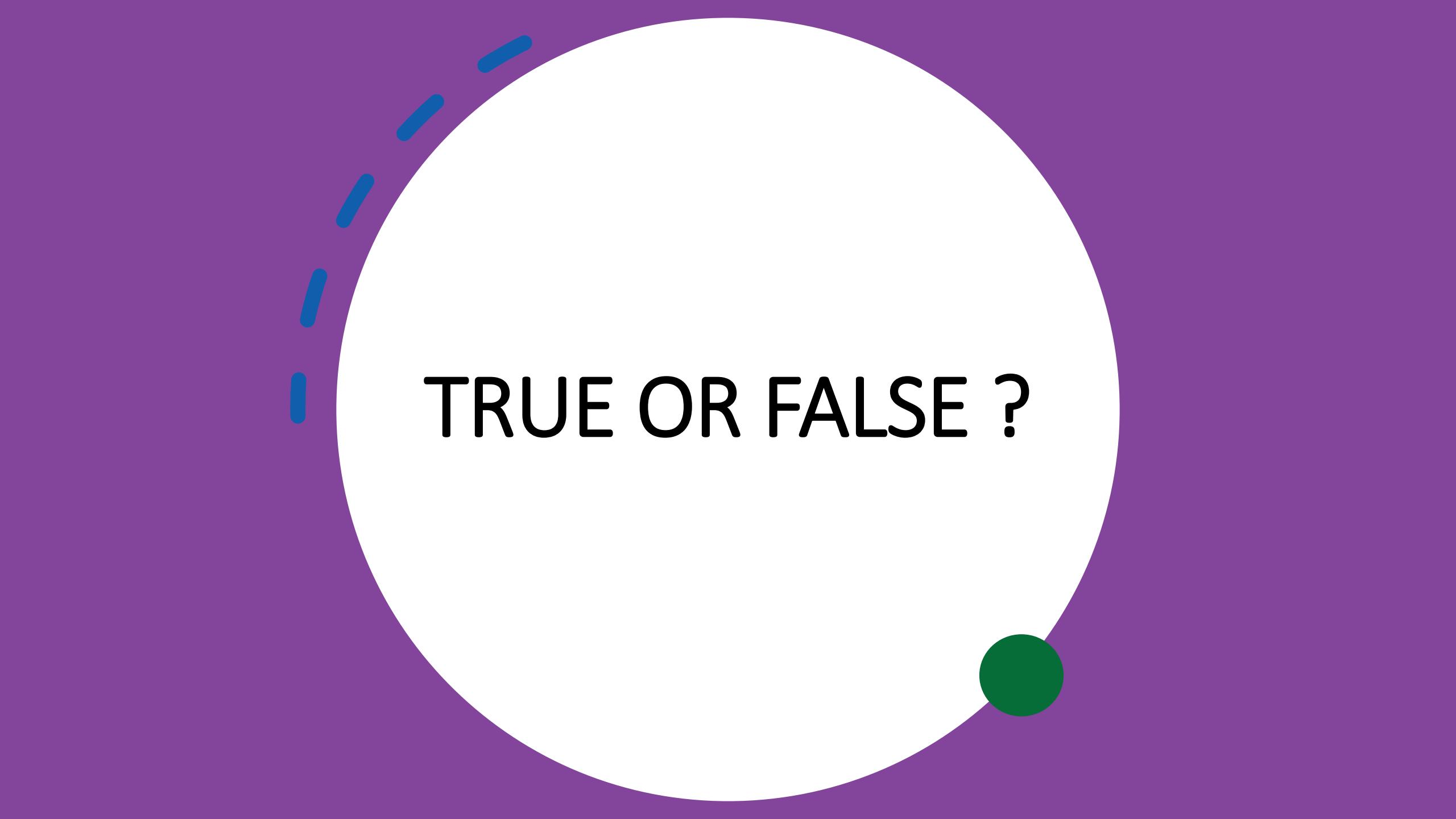


The background features abstract graphic elements: a large purple circle on the right, a smaller green circle above it, a blue circle at the top right, a brown square on the left, and several blue dashed lines radiating from the bottom left.

How do we avoid  
information overload?



# Visualization Principles



**TRUE OR FALSE ?**

CHART 1

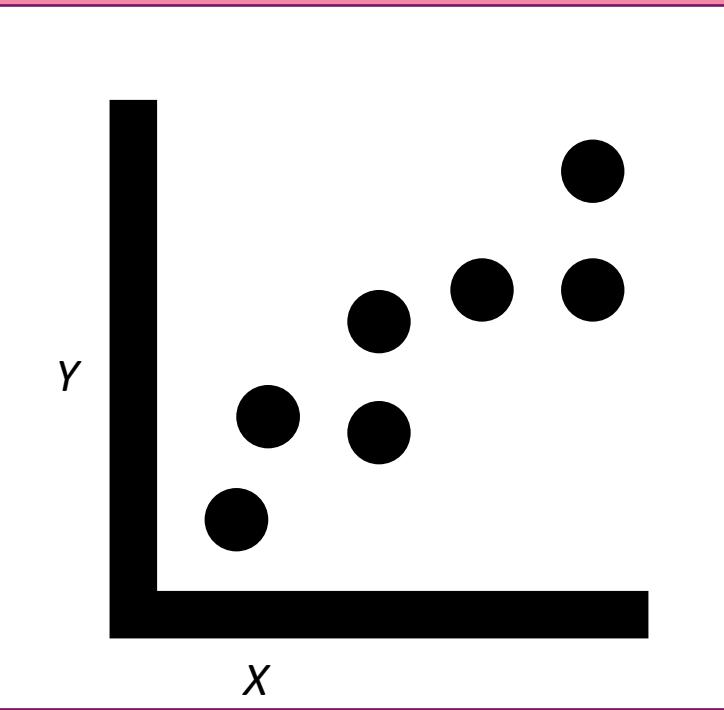
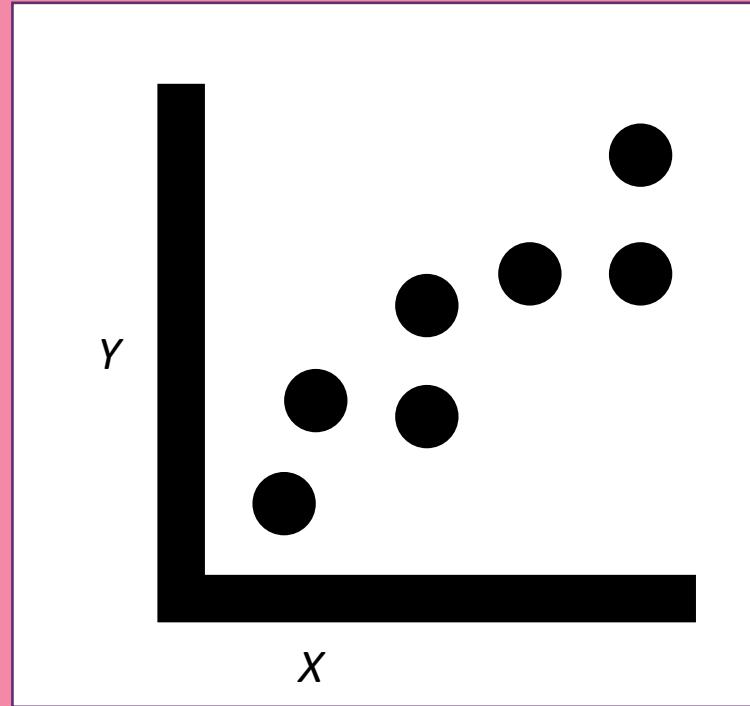


CHART 2



?

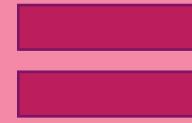


CHART 1

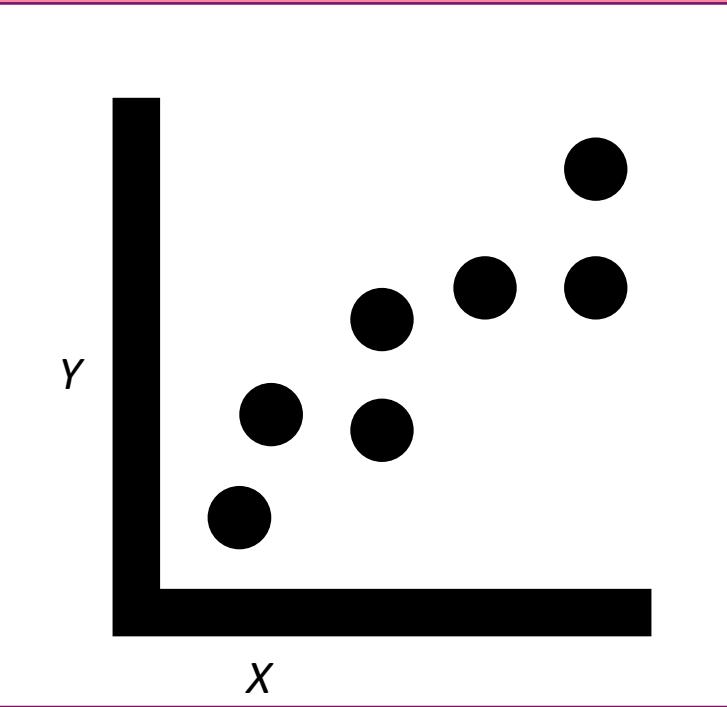
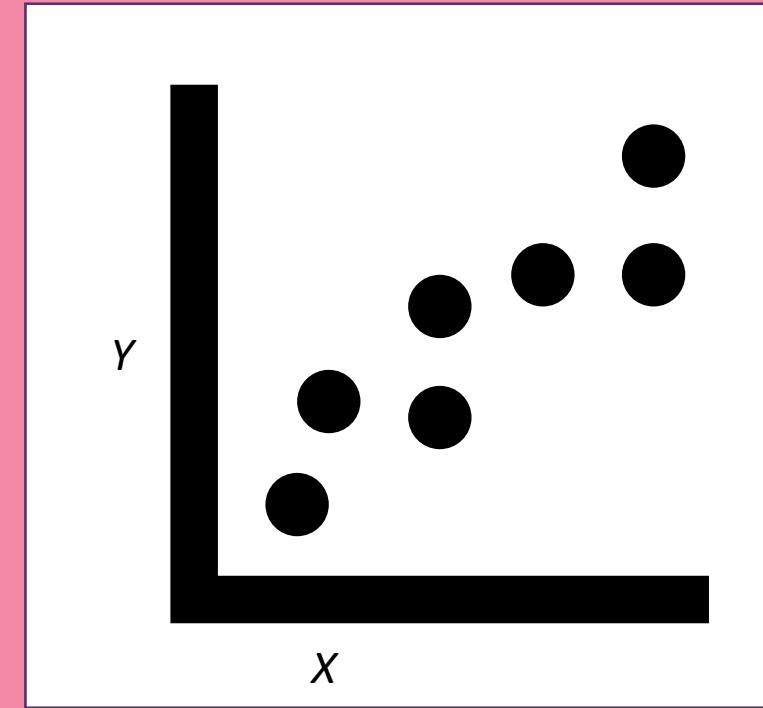


CHART 2



FALSE



Mean ( $x_1$ ) = Mean ( $x_2$ )

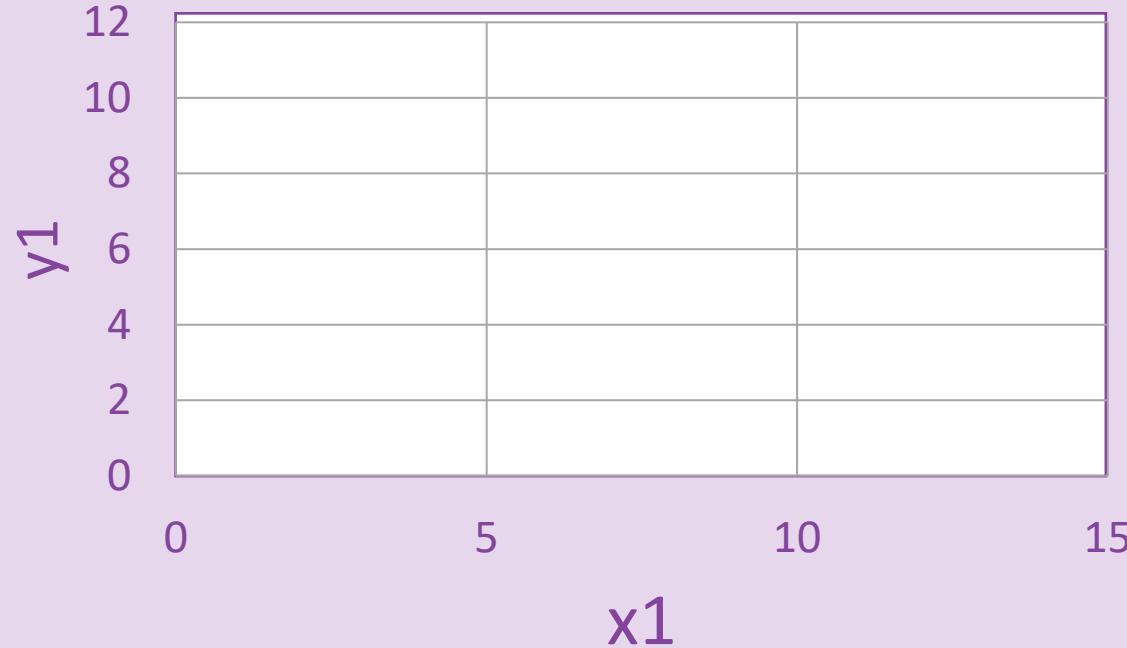
Mean ( $y_1$ ) = Mean( $y_2$ )

Variance ( $x_1$ ) = Variance ( $x_2$ )

Variance ( $y_1$ ) = Variance ( $y_2$ )

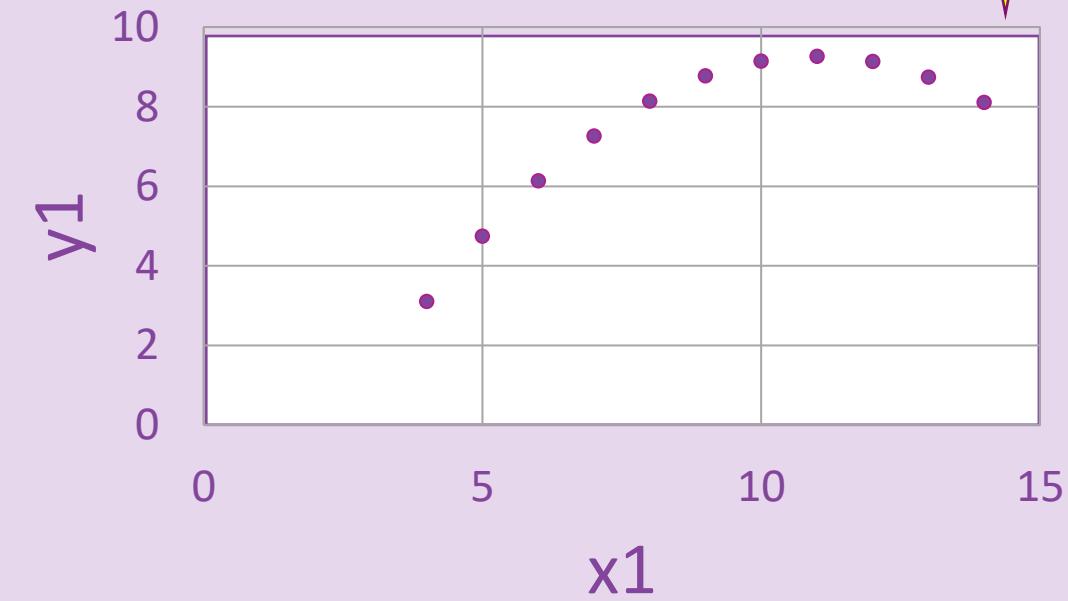
Linear regression = Linear Regression

# DATASET 1



VS

# DATASET 2



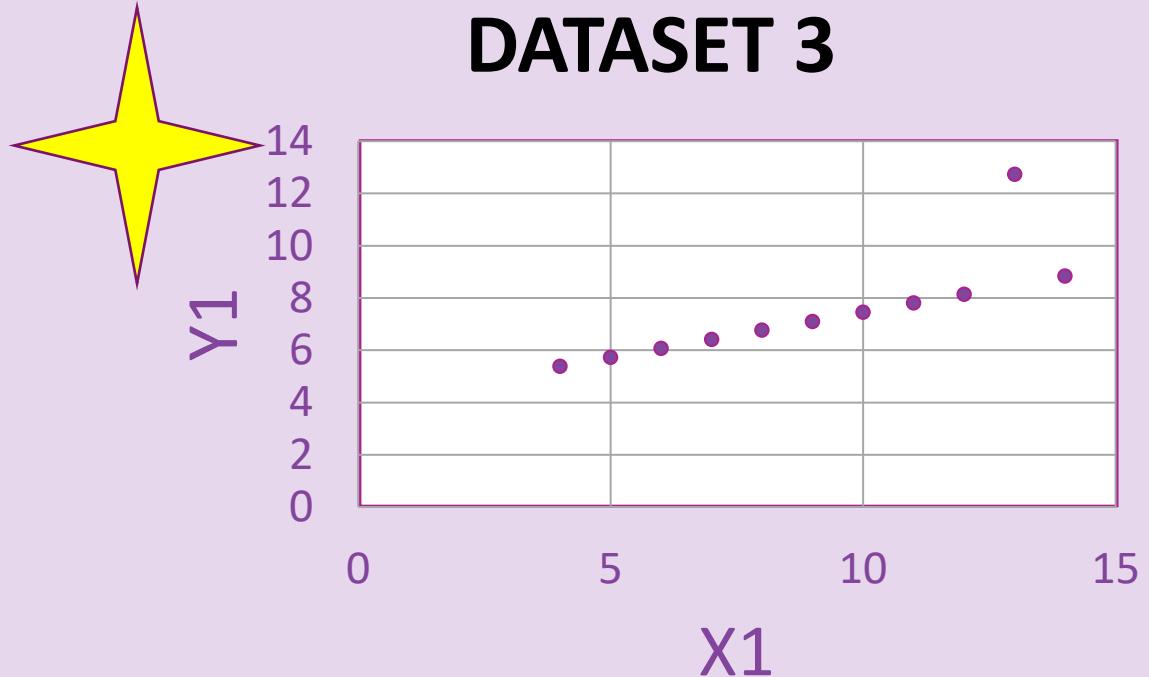
<http://etc.ch/FhT2>



TRUE

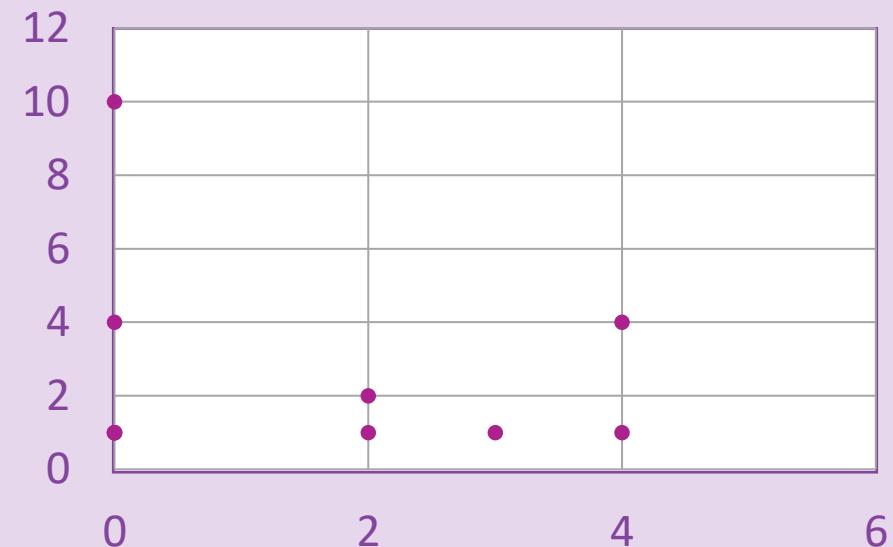


# DATASET 3



# DATASET 4

VS



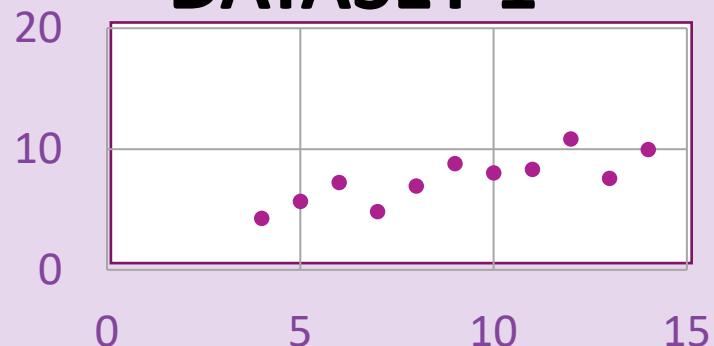
<http://etc.ch/FhT2>



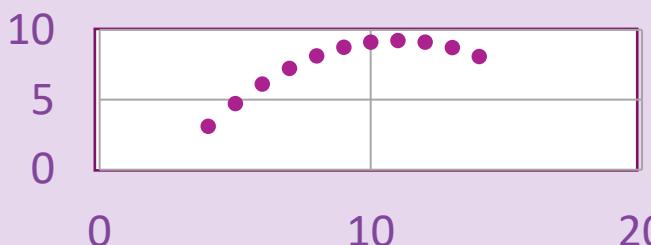
FALSE



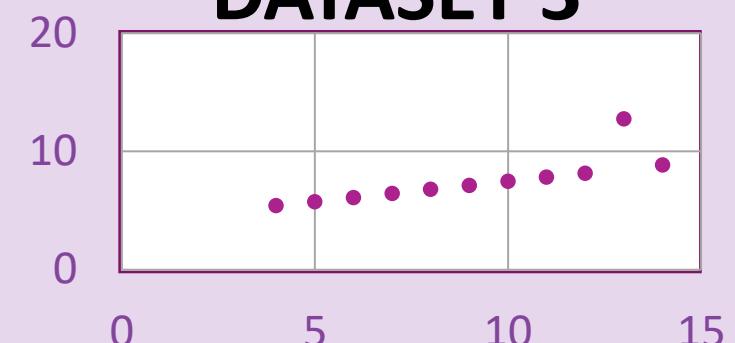
# **DATASET 1**



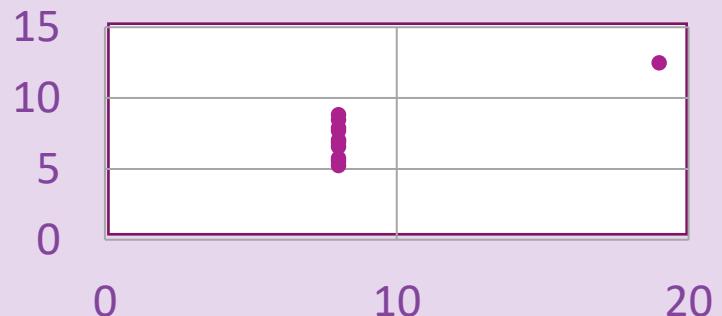
# **DATASET 2**



# **DATASET 3**



# **DATASET 4**



<http://etc.ch/FhT2>



TRUE

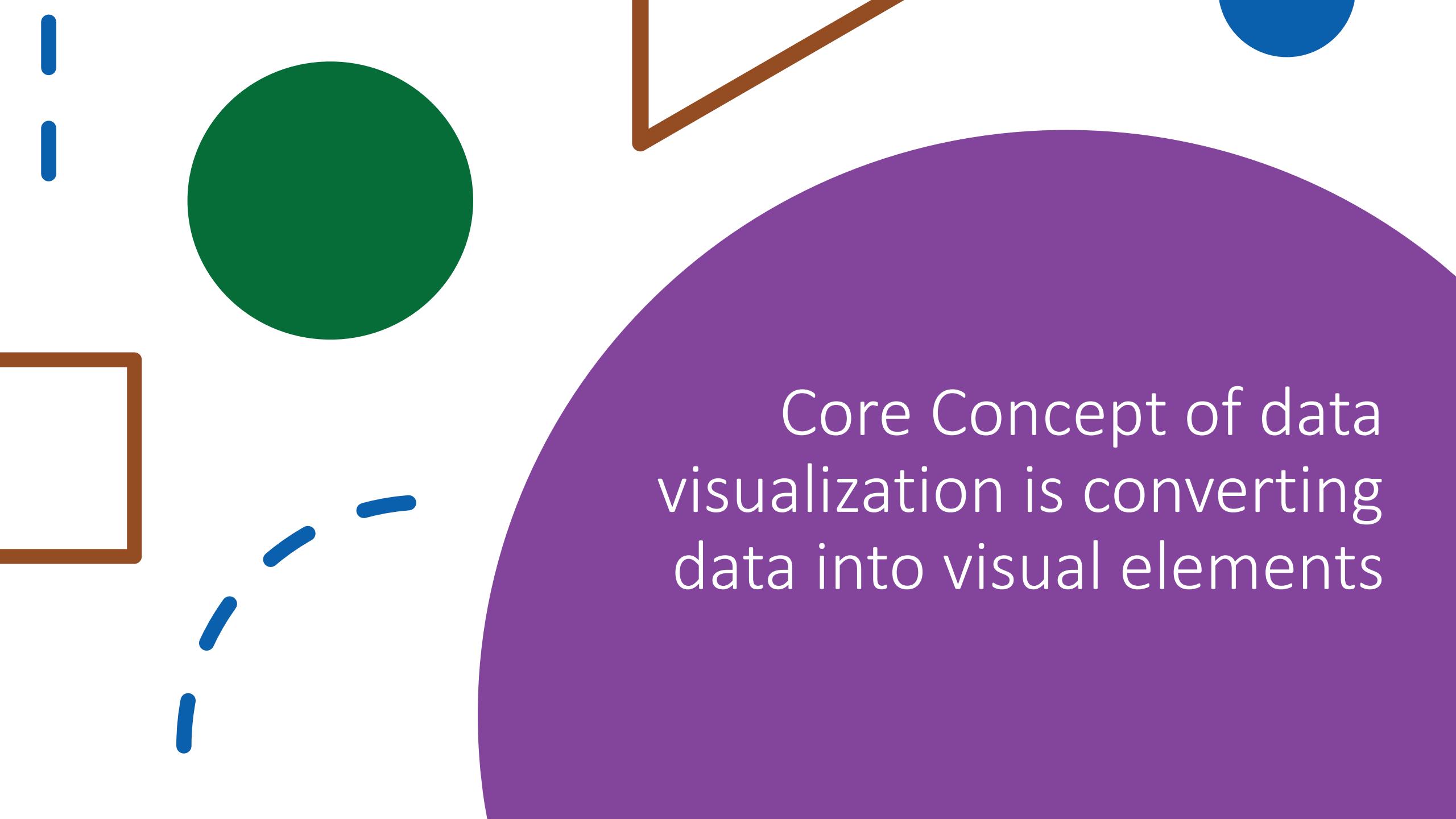


# Why Visualizations?

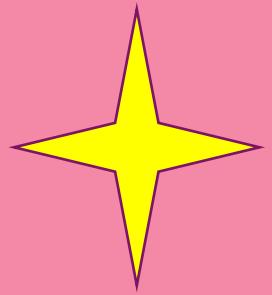
1		2		3		4	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Source: Anscombe's Quartet

Property	Value	Accuracy
Mean ( $x$ )	9	exact
Sample variance ( $x$ )	11	exact
Mean ( $y$ )	7.50	to 2 decimal places
Sample variance ( $y$ )	4.125	$\pm 0.003$
Correlation between $x$ and $y$	0.816	to 3 decimal palaces
Linear regression line	$y = 3.00 + 0.500x$	to 2 and 3 decimal places, respectively
Coefficient of determination of the linear regression	0.67	to 2 decimal places

The background features abstract geometric shapes in blue and brown on a white surface. A large green circle is positioned in the upper left, a brown square with a white center is in the lower left, and a blue circle is in the top right. A brown line extends from the top center towards the right.

Core Concept of data  
visualization is converting  
data into visual elements



# Visual Encodings

- Colors
- Shapes
- Position

**Activity #4** Order the visual encodings

Example	Encoding
	position, placement
1, 2, 3; A, B, C	text labels
	length
	size, area
	angle
	pattern density
	weight, boldness
	saturation, brightness
	color
	shape, icon
	pattern texture
	enclosure, connection
	line pattern
	line endings
	line weight



David Tittsworth



Link shared Sep 22

7 Files

# Accuracy of visual encodings

Color Shading



Position



Color  
Saturation



Direction



Length



Area

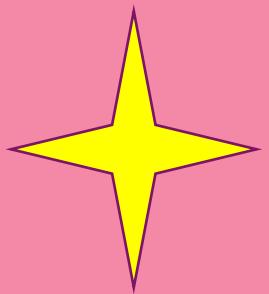


Volume

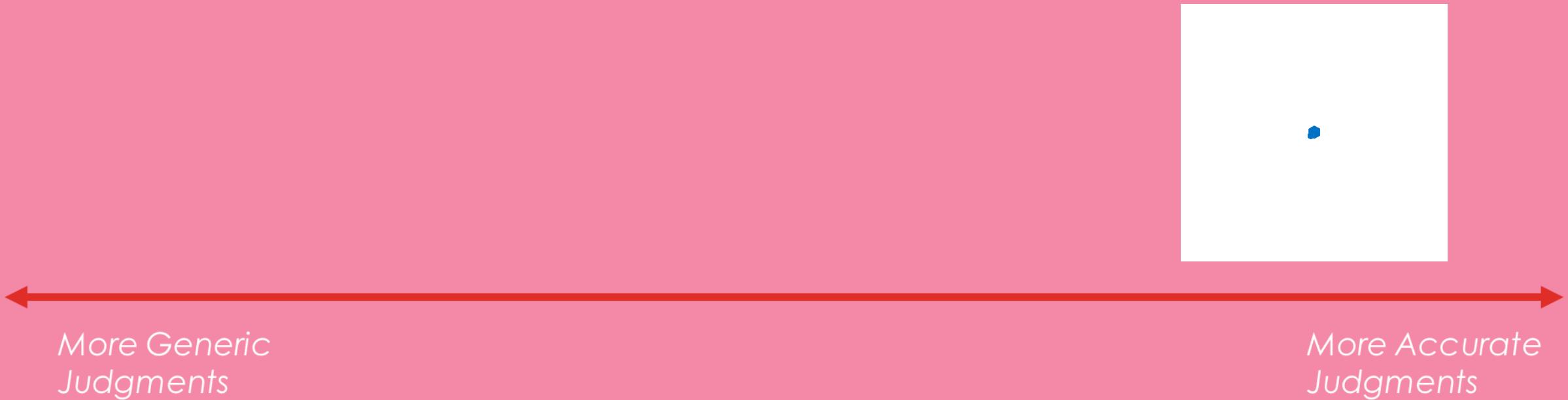


More Generic Judgements

More Accurate Judgements

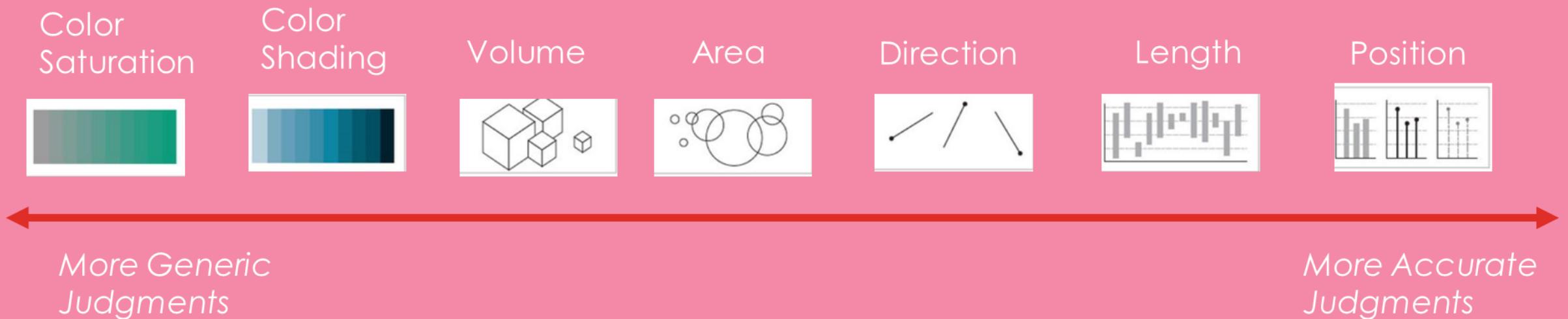


# ACCURACY OF VISUAL ENCODINGS



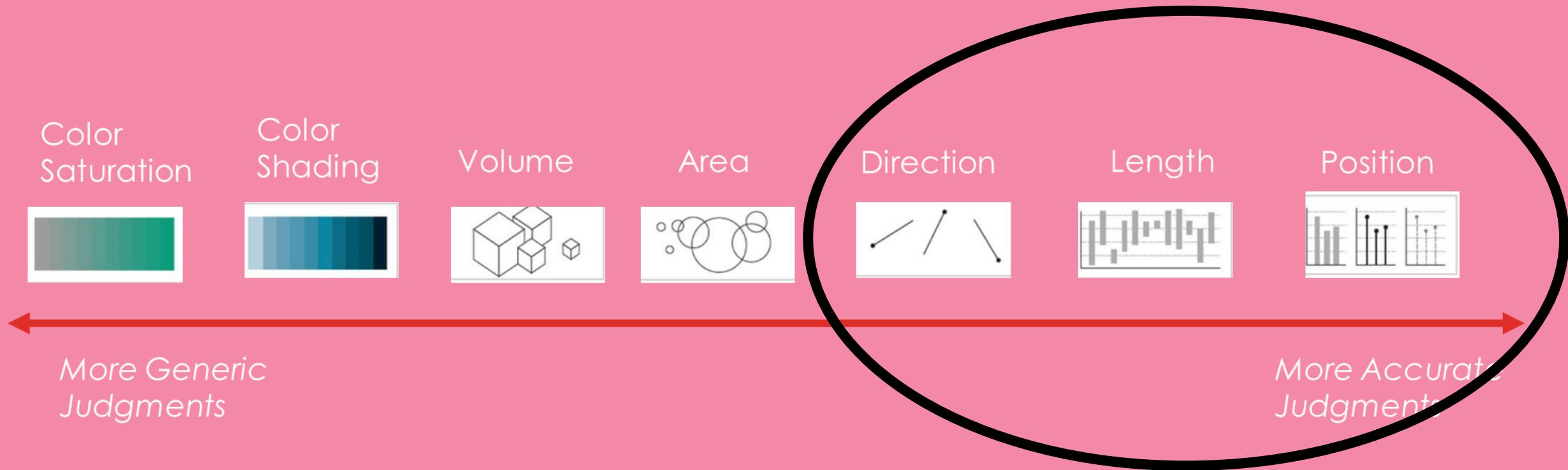
Source: Mackinlay, Jock. "Automating the Design of Graphical Presentations of Relational Information." ACM Trans. Graph. 5, no. 2 (1986): 110–41. doi:10.1145/22949.22950.

# ACCURACY OF VISUAL ENCODINGS



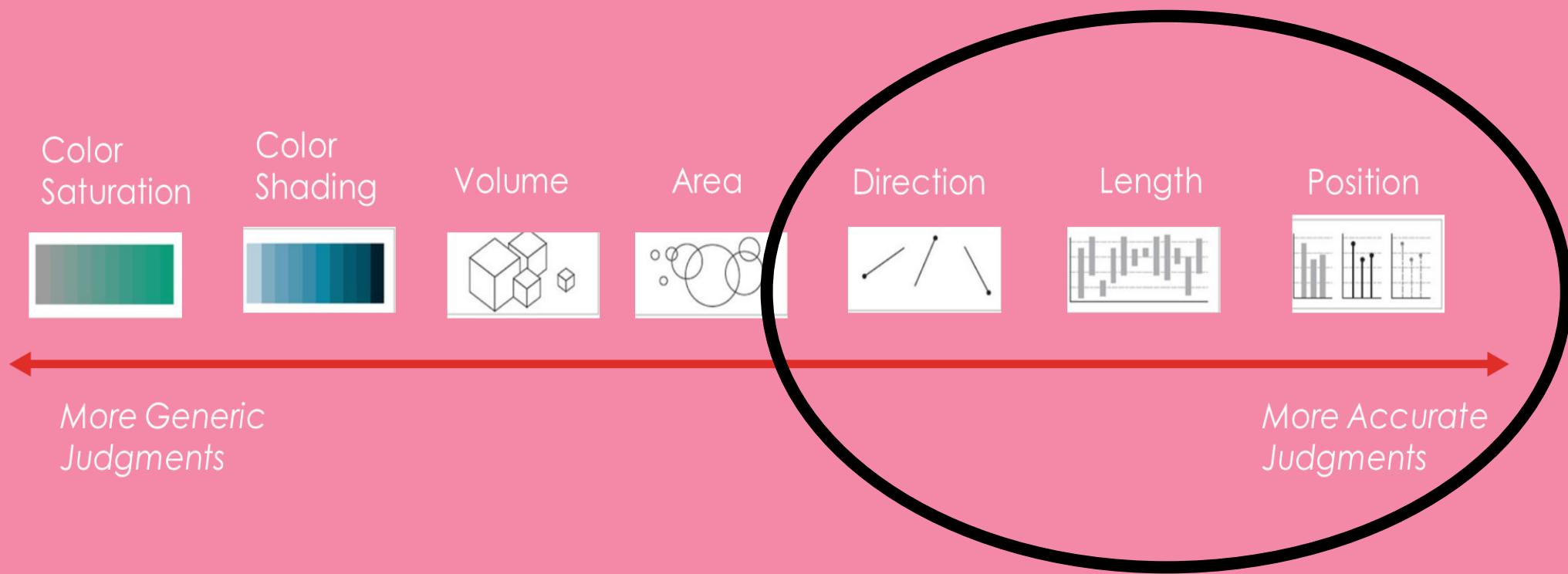
Source: Mackinlay, Jock. "Automating the Design of Graphical Presentations of Relational Information." ACM Trans. Graph. 5, no. 2 (1986): 110–41. doi:10.1145/22949.22950.

# ACCURACY OF VISUAL ENCODINGS



**Easily Shows Comparisons and exact numbers**

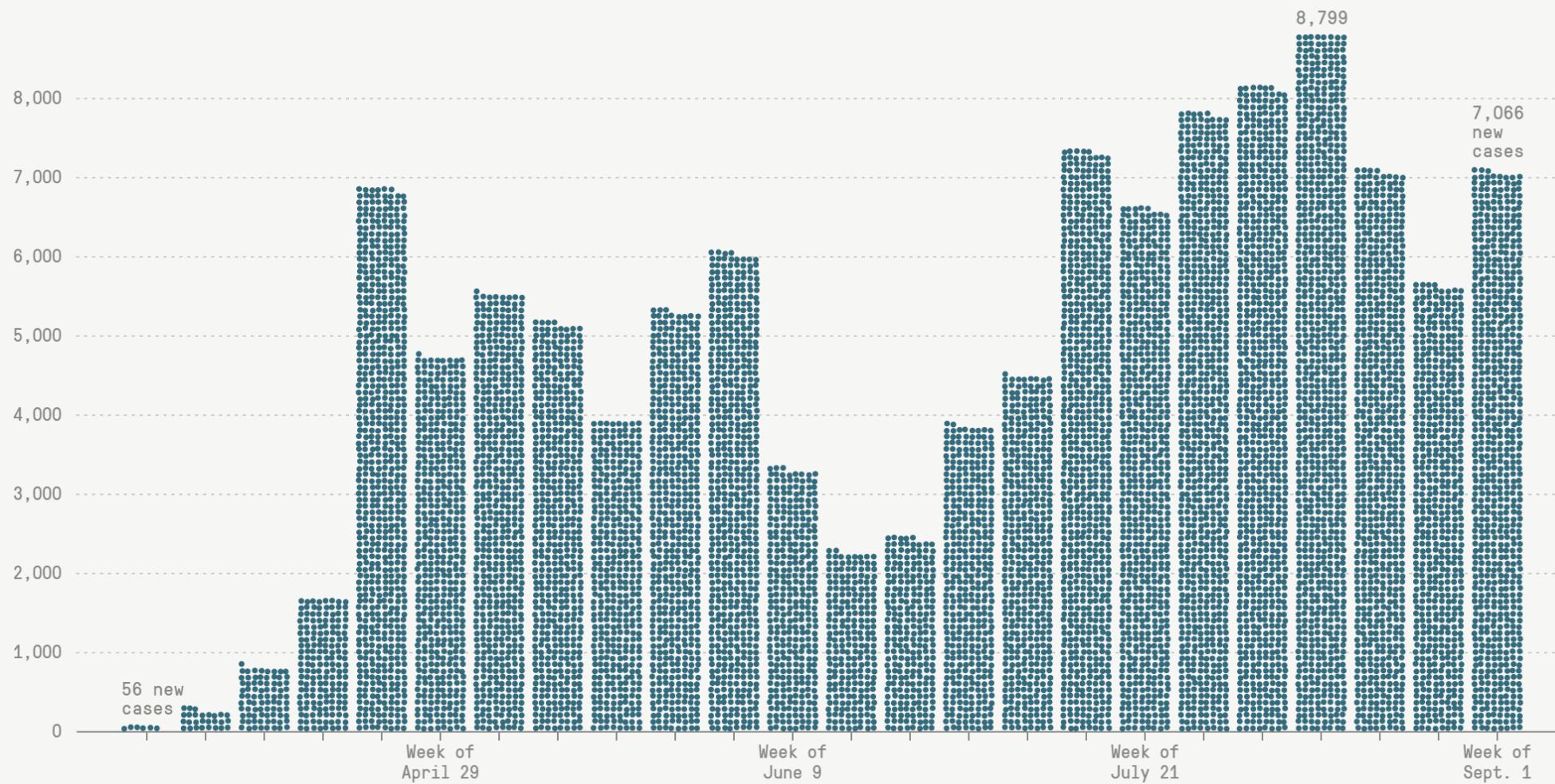
# ACCURACY OF VISUAL ENCODINGS



**Easily Shows Comparisons and  
exact numbers**

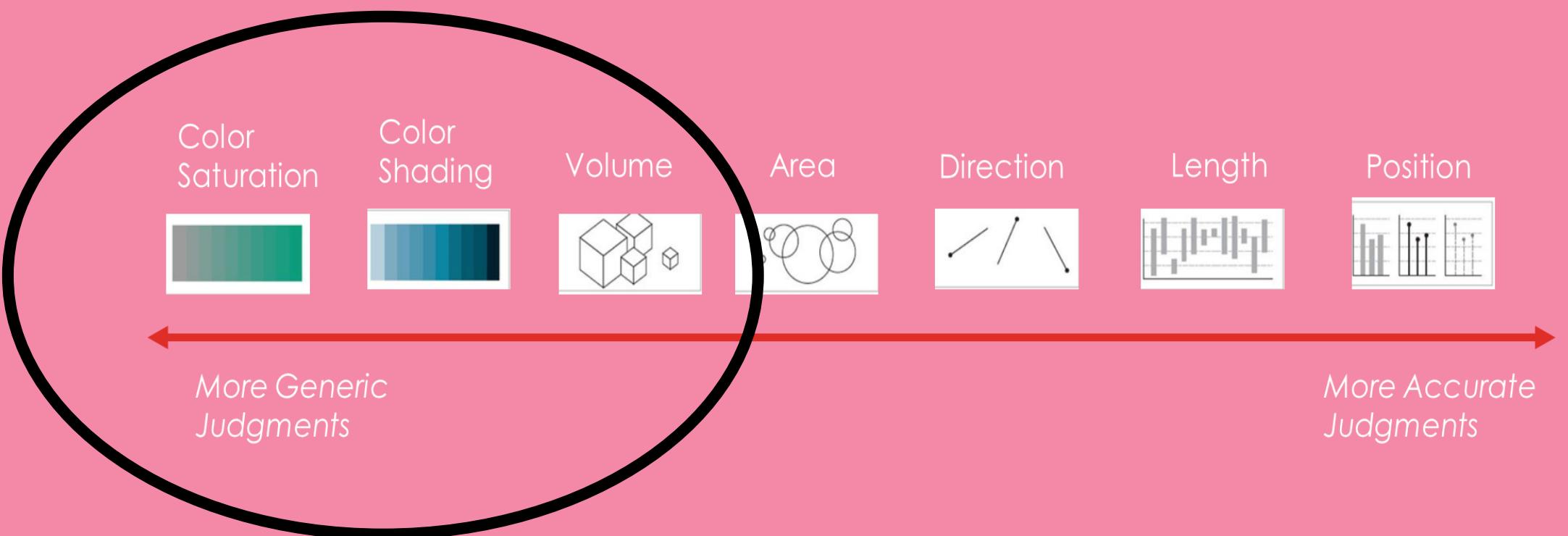
# Coronavirus Reported among Prisoners

Each ● represents 10 new cases



Source: <https://www.themarshallproject.org/2020/05/01/a-state-by-state-look-at-coronavirus-in-prisons>

# ACCURACY OF VISUAL ENCODINGS

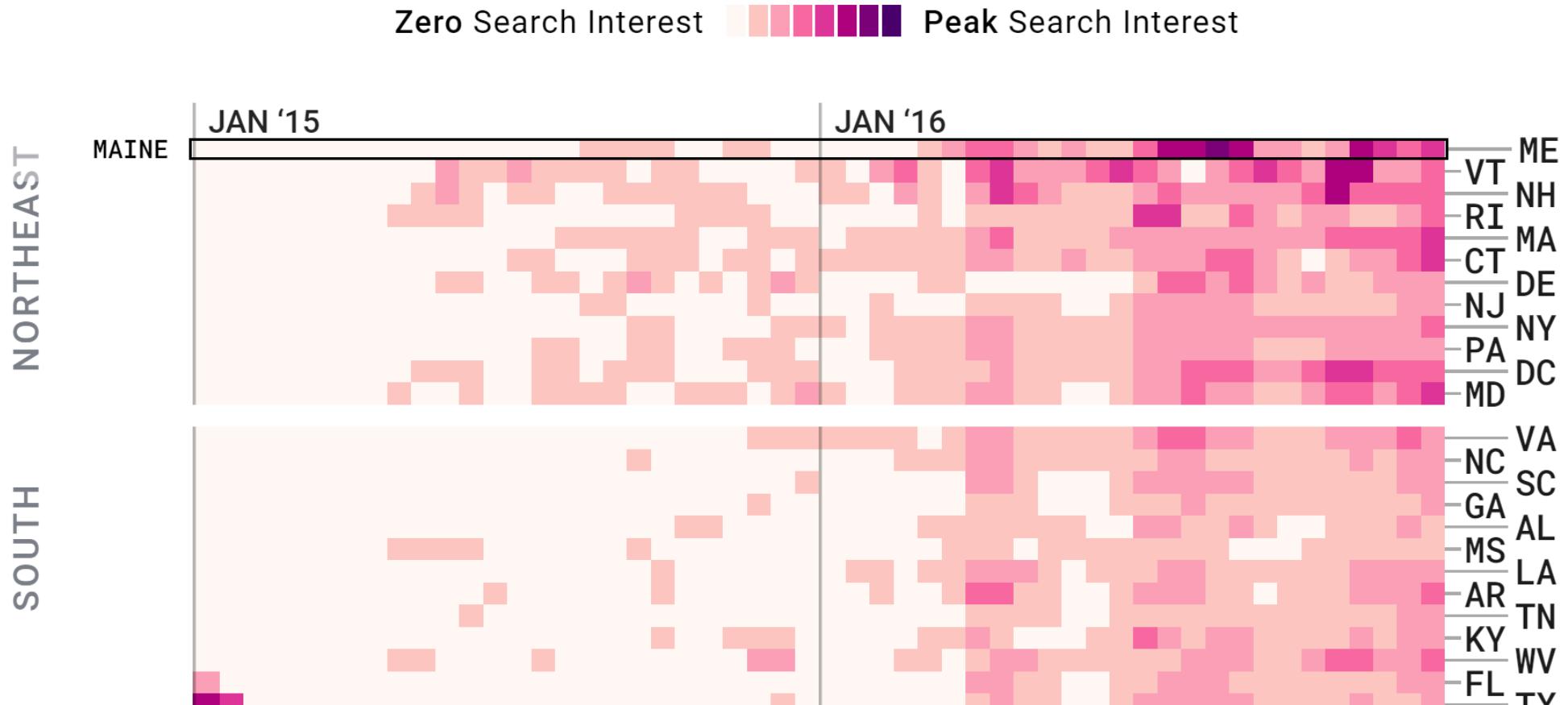


**Better for showing Trends**

# Where Words Come From

Search interest<sup>1</sup> for top rising definitions by state, 2015 - 2016

## GASLIGHTING by State

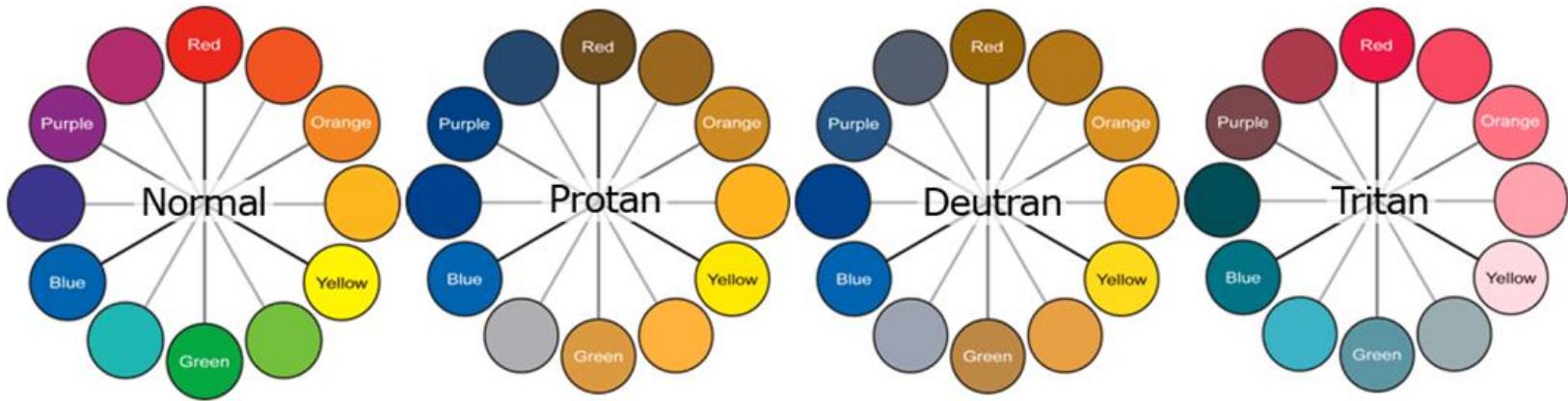


Source: <https://pudding.cool/2017/02/new-slang/>

# Accessibility Concerns



# Considerations with Color



- Color-Blind Accessible
- Perpetually Balanced
- 10 colors max



# Considerations with Design

- ARIA Tags
  - Enabling and disabling animation
  - Increasing letter kerning and stroke-width for low-vision users
  - Replacing complex visual with table or summary
-



# Animation

- Use it wisely!
  - Show context
  - Spotlight
  - Time
-

# Four Ways to Slice Obama's 2013 Budget Proposal

Explore every nook and cranny of President Obama's federal budget proposal.

[All Spending](#)[Types of Spending](#)[Changes](#)[Department Totals](#)

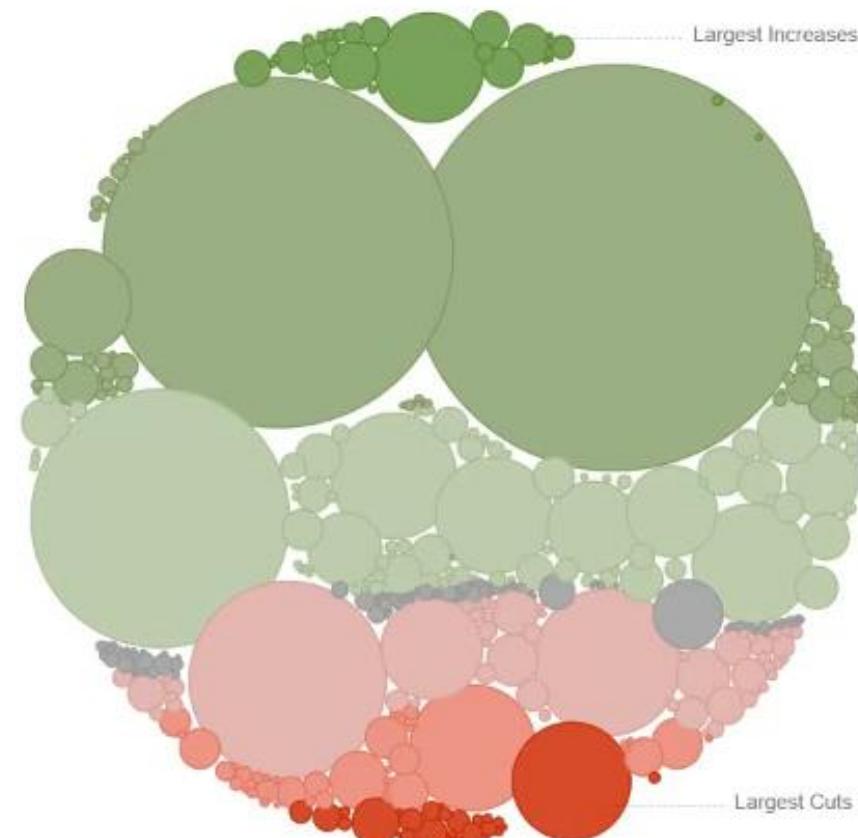
## How \$3.7 Trillion Is Spent

Mr. Obama's budget proposal includes \$3.7 trillion in spending in 2013, and forecasts a \$901 billion deficit.

Circles are sized according to the proposed spending.



Color shows amount of cut or increase from 2012.



The proposal forecasts  
a \$901 billion deficit.



## Demo for Role of Isolation in Context-Shifting

Sarah Drasner

PRO

+ Follow



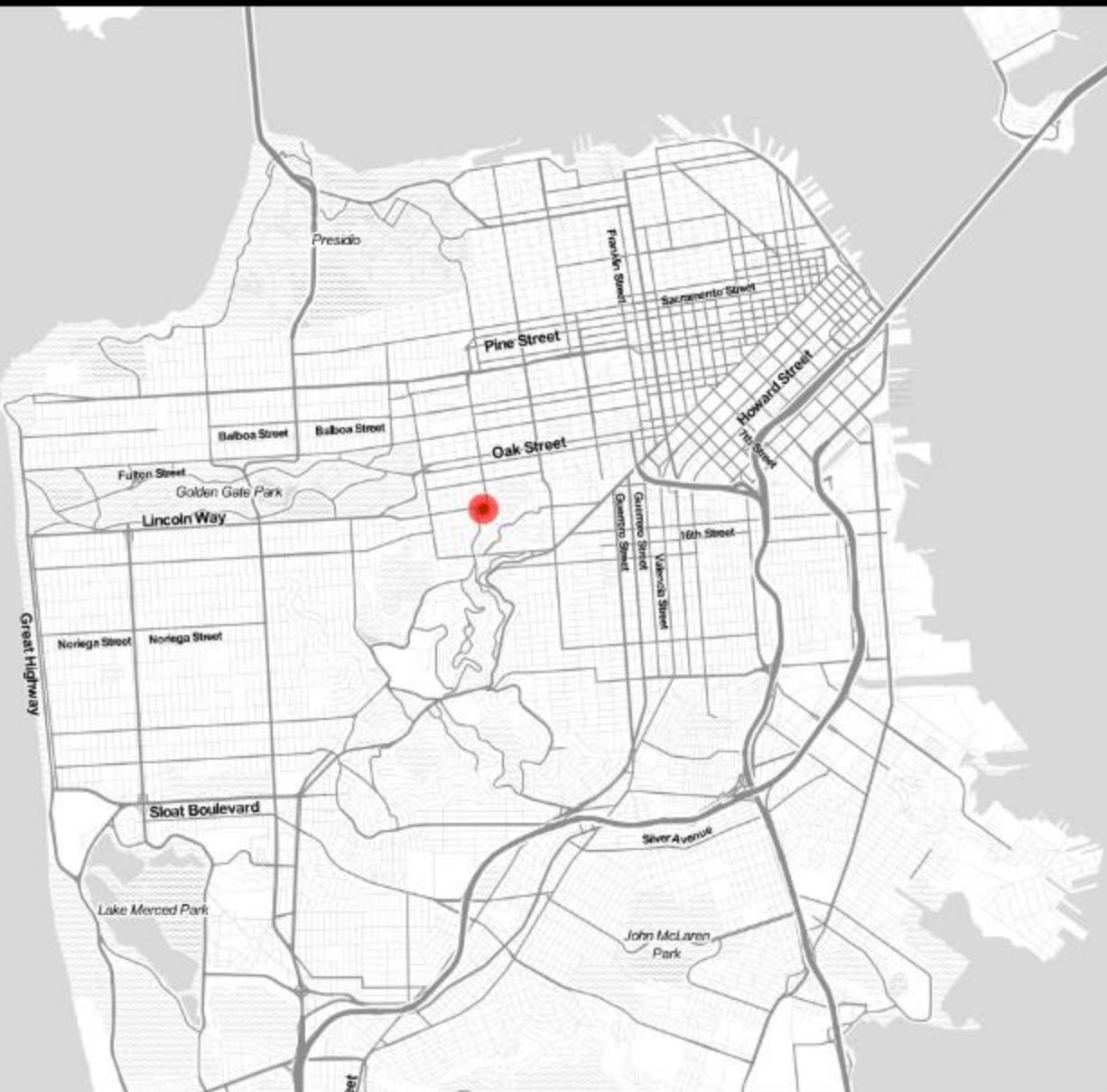
Change View

Sign Up

Log In



Source: <https://codepen.io/sdras/full/qOdWEP>



# ELLIS ACT EVICTIONS

0

SAN FRANCISCO HOUSEHOLDS FORCED OUT OF THEIR HOMES.  
1/1/1994 - 7/11/1994



The Ellis Act is a state law which says that landlords have the right to evict tenants in order to "go out of business." All units in the building must be cleared of all tenants- no one can be singled out. Most often it is used to convert to condos or group-owned tenancy-in-common flats. Once a building becomes a condo it is exempt from Rent Control, regardless of the age of the building, and even if a unit owner subsequently rents to a long-term tenant.

**There is no limit to the number of times a building owner can "go out of business".** Rent Board data shows some owners buying and Ellis-ing multiple buildings over time. If these buyers do not want to be landlords, why are they buying buildings full of rental units? These Ellis-ed buildings - now "out of business"- are also showing up for rent as illegal vacation rentals on sites like Airbnb and VRBO.

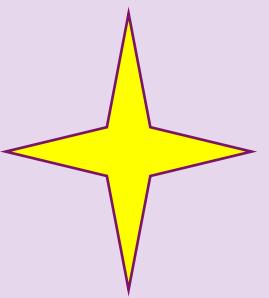
With landlords looking for ways to avoid renting to long-term tenants, the housing crisis in San Francisco will only be exacerbated. See our chart of no-fault evictions [here](#).

If you have been evicted, please fill out [our survey](#) to add your story to a comprehensive map in the making!

Also, please take our pledge to boycott renting or buying from a landlord who has profited by displacing tenants [here!](#) You can also look up an address to determine its eviction history.

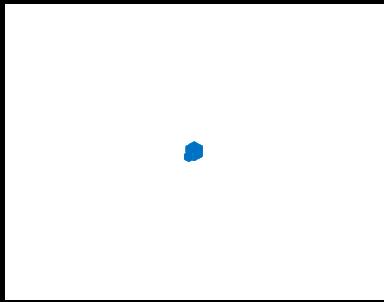
Map created by Anti-Eviction Mapping Project

Powered by D3, Leaflet and Carto

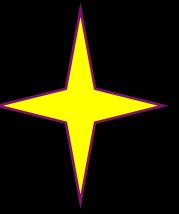


Let's play a game...

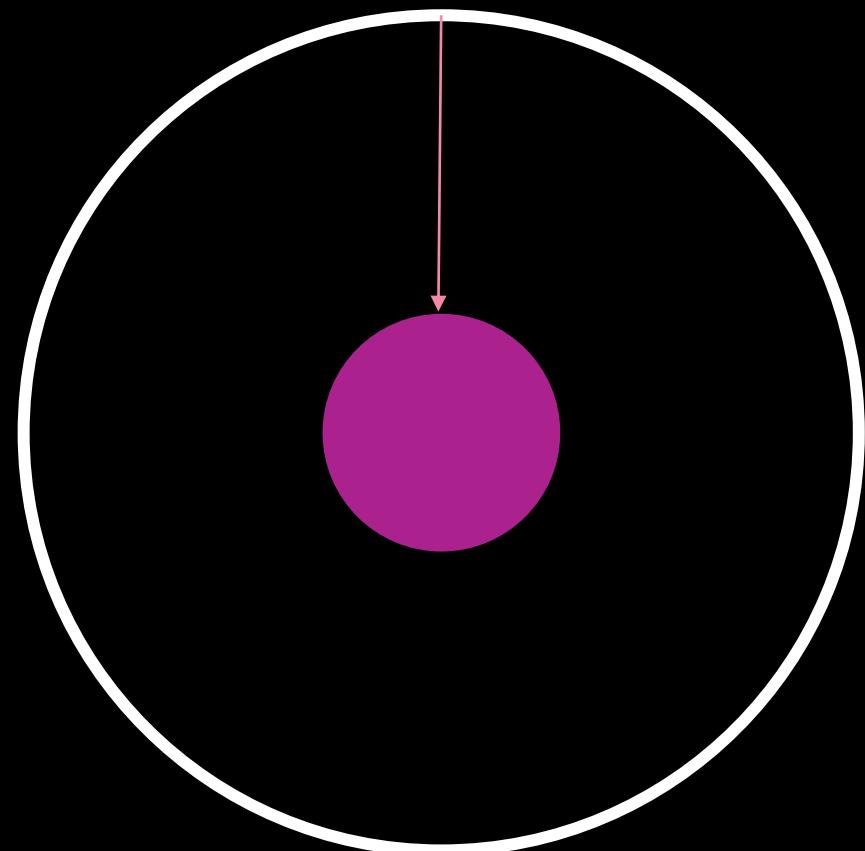
Activity #5



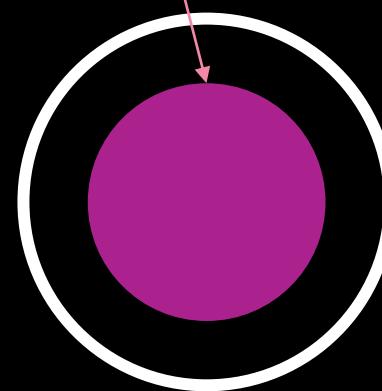
Choose the larger circle? \*



Circle #1



Circle #2

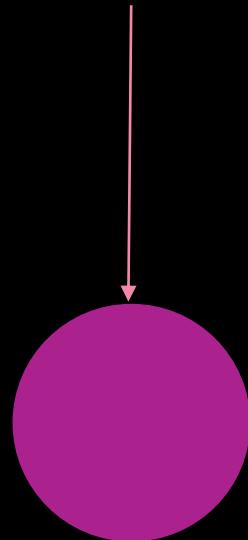


<http://etc.ch/FhT2>

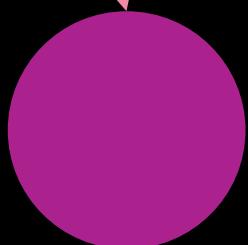


# Choose the larger circle? \*

Circle #1



Circle #2



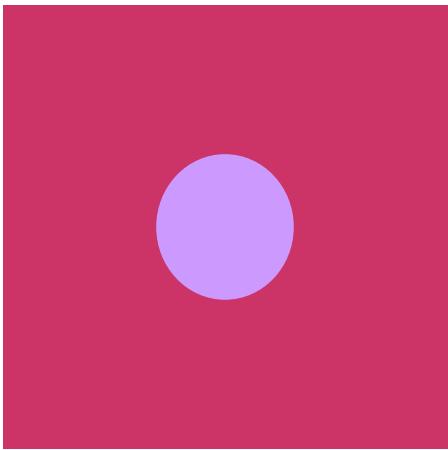
<http://etc.ch/FhT2>



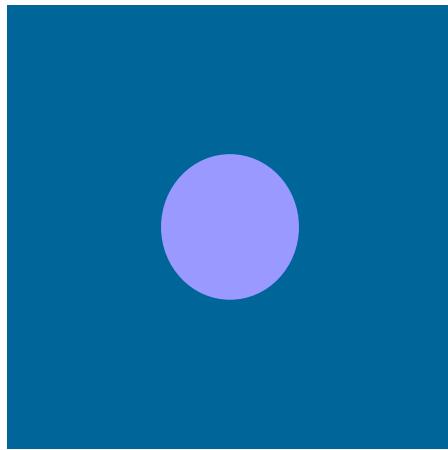
# Which circle is colored differently?



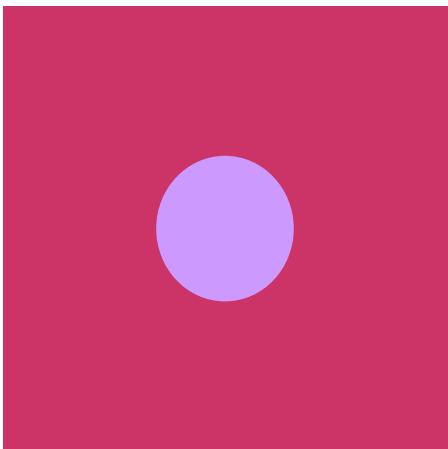
•



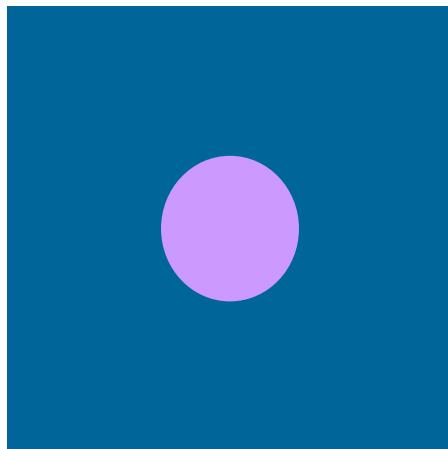
#1



#2

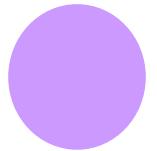


#3

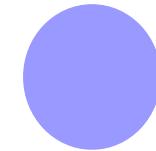


#4

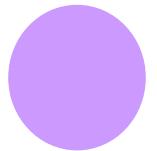
# Which circle is colored differently?



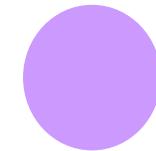
#1



#2



#3

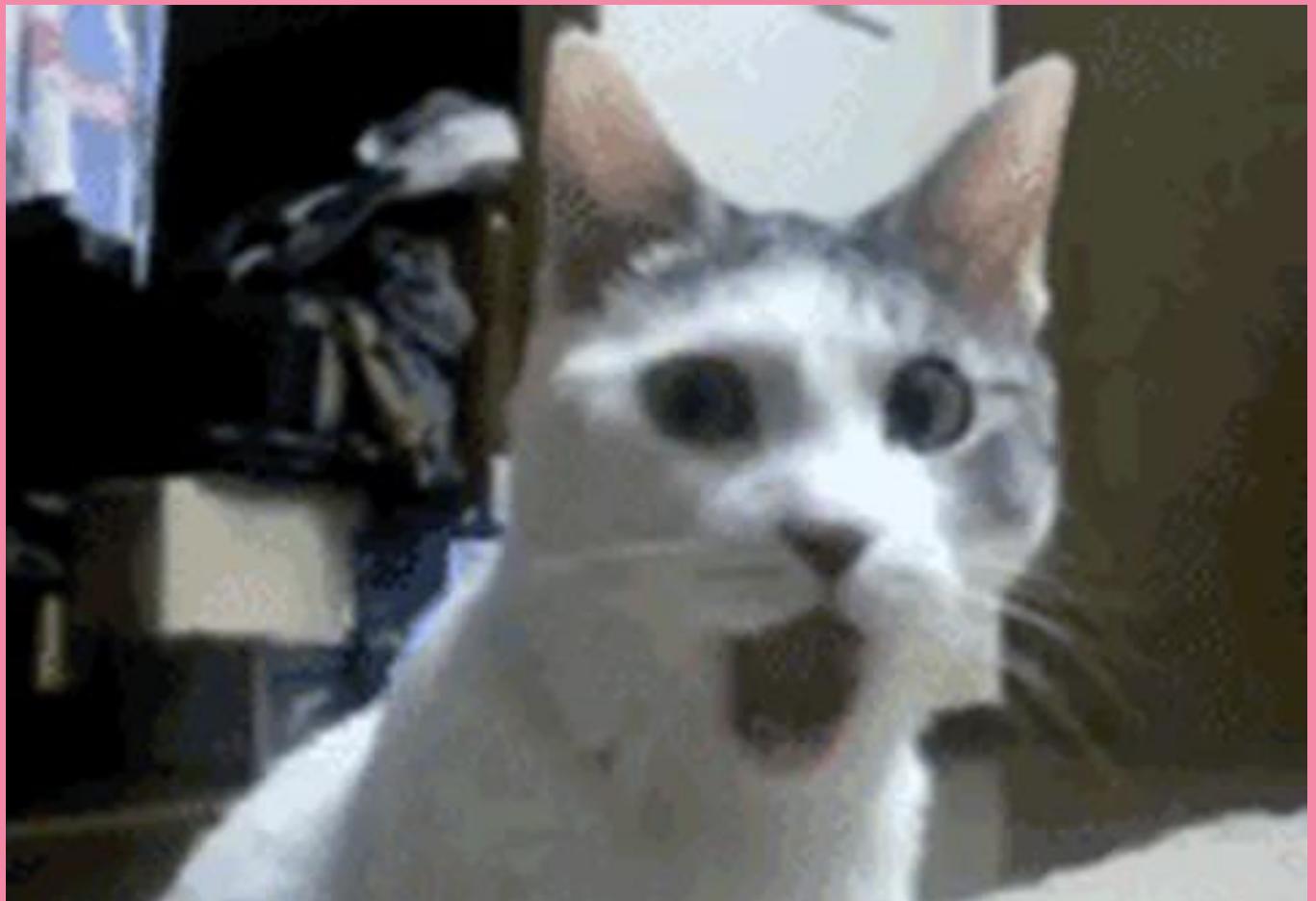


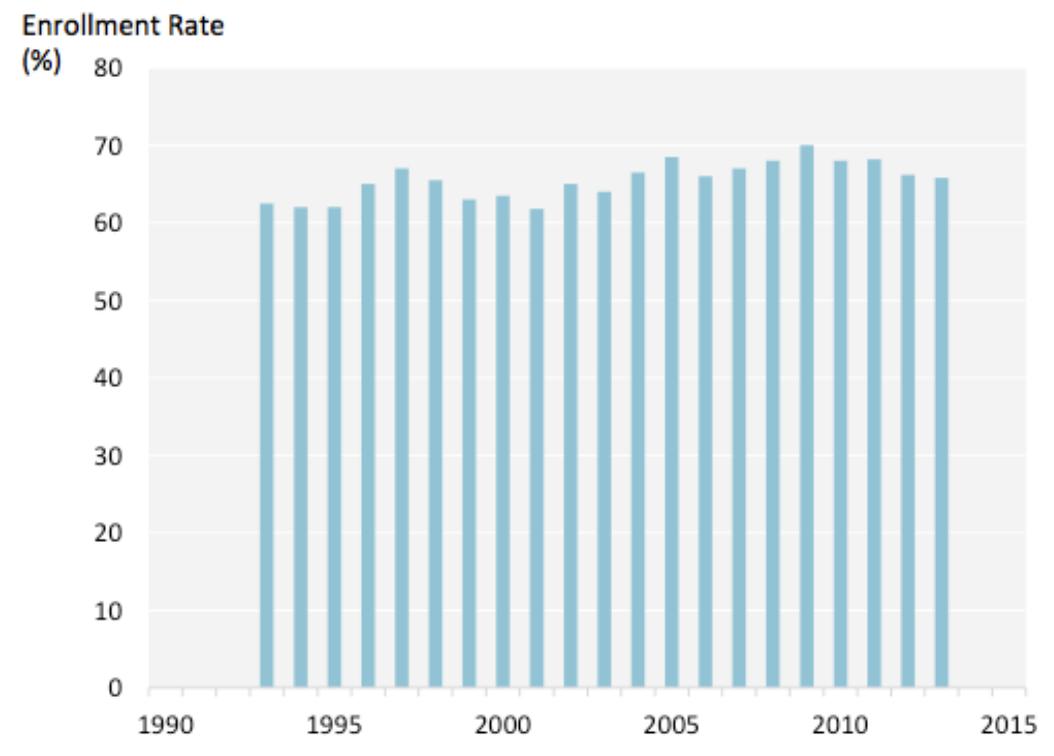
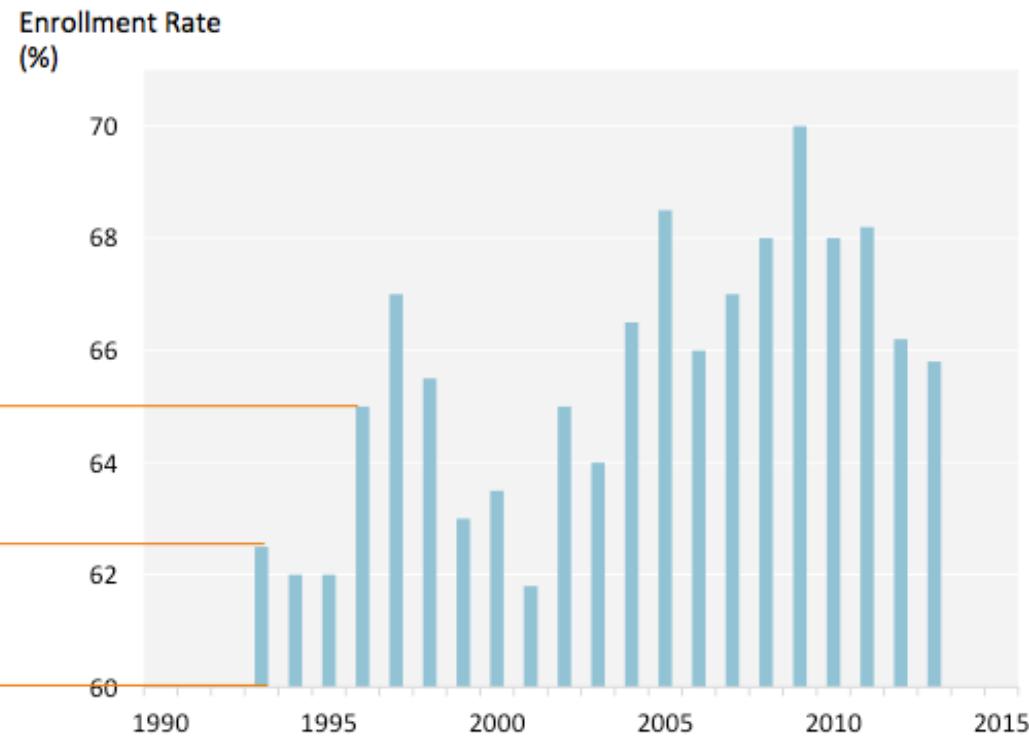
#4

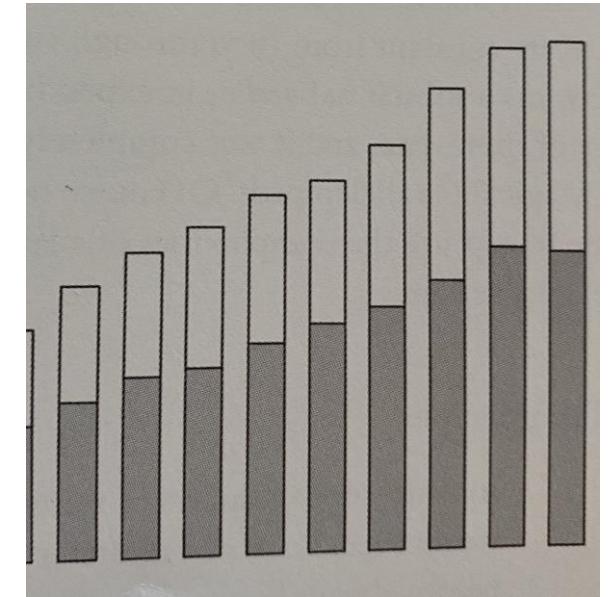
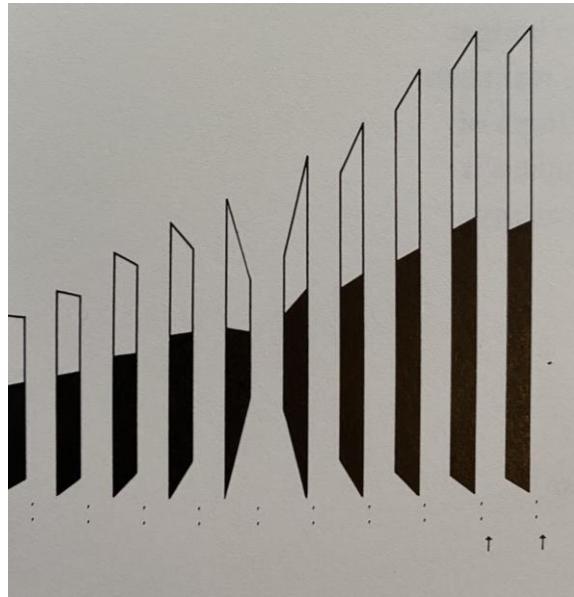
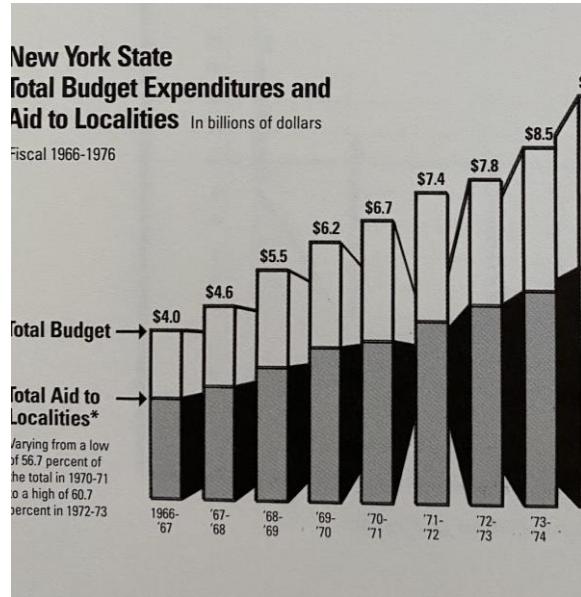


What we see can be  
misleading... choose  
encodings carefully!

*Gotchas!*

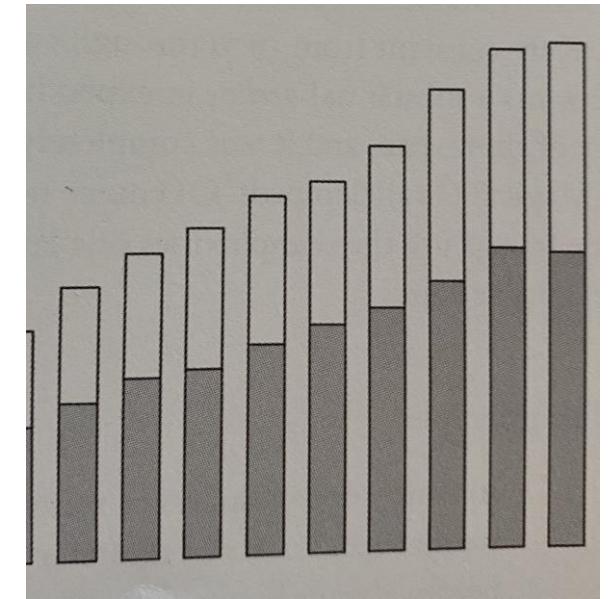
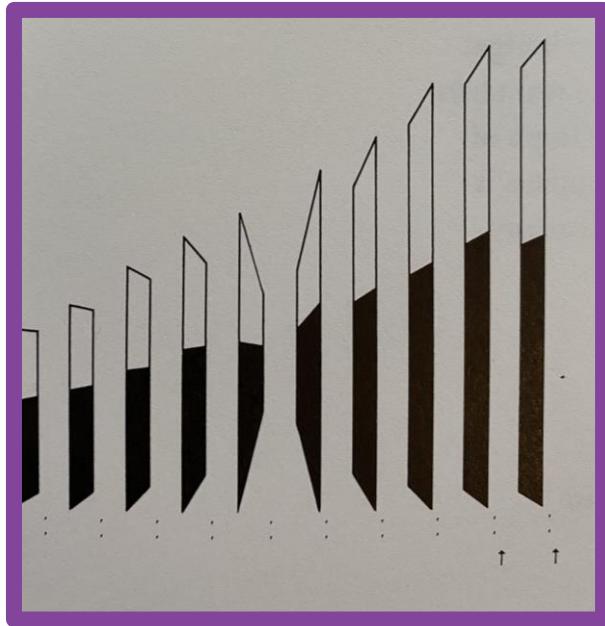
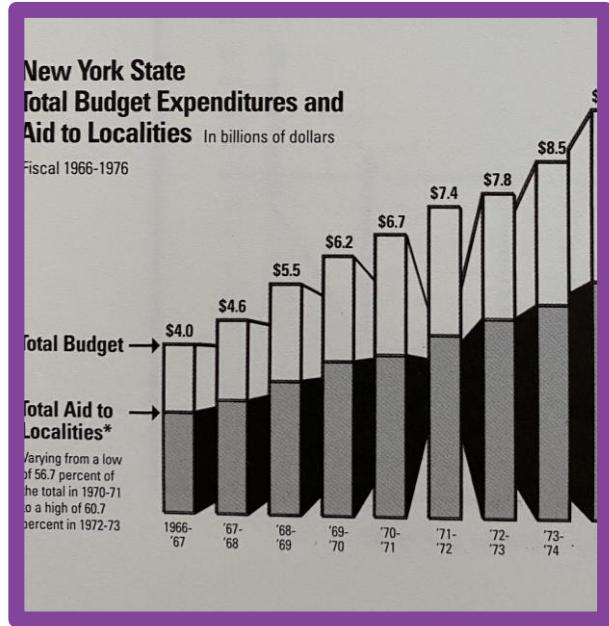






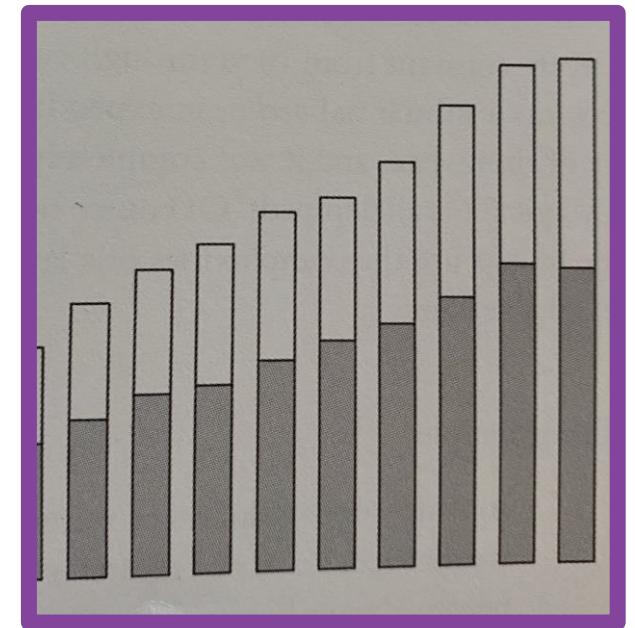
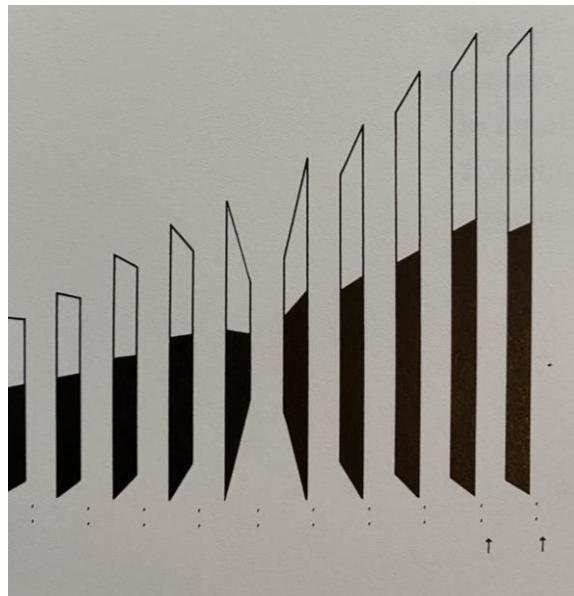
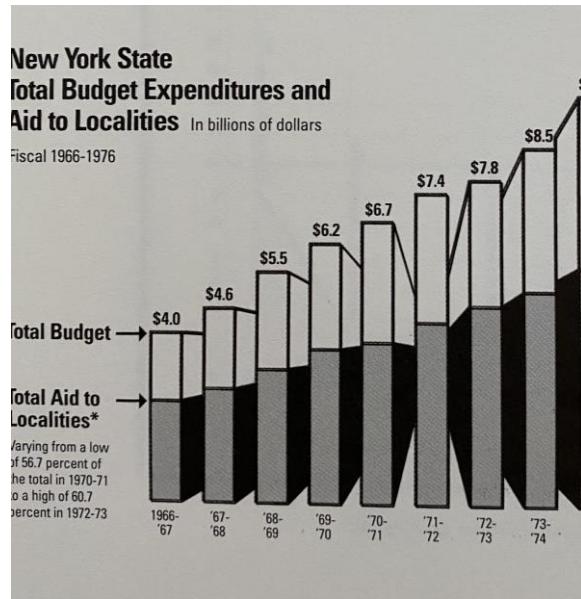
“Above all, show the data” - Tufte

Maximize data-to-ink ratio, minimize perceptual artifacts



“Above all, show the data”- Tufte

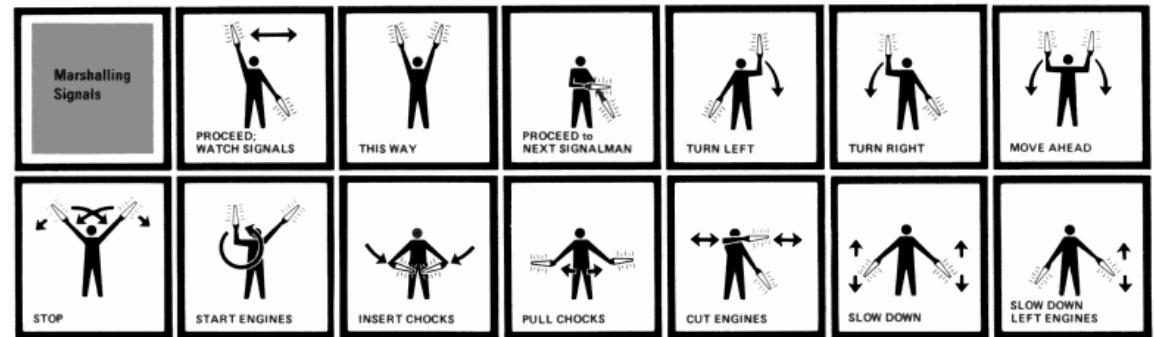
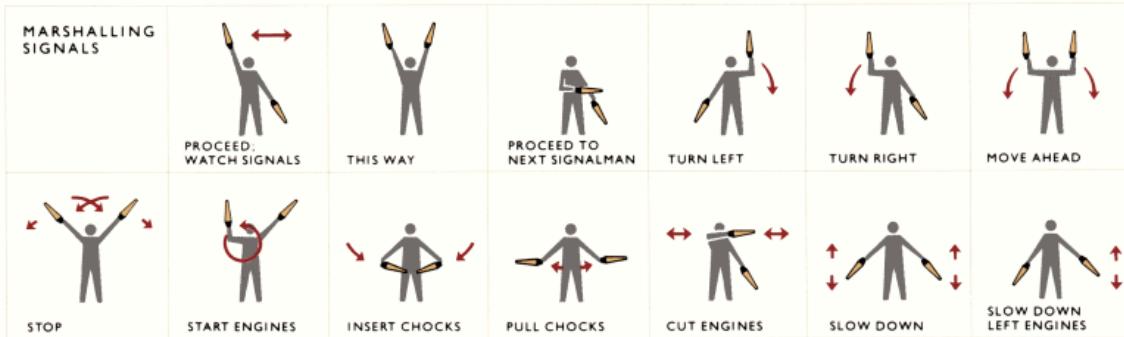
Maximize data-to-ink ratio, minimize perceptual artifacts



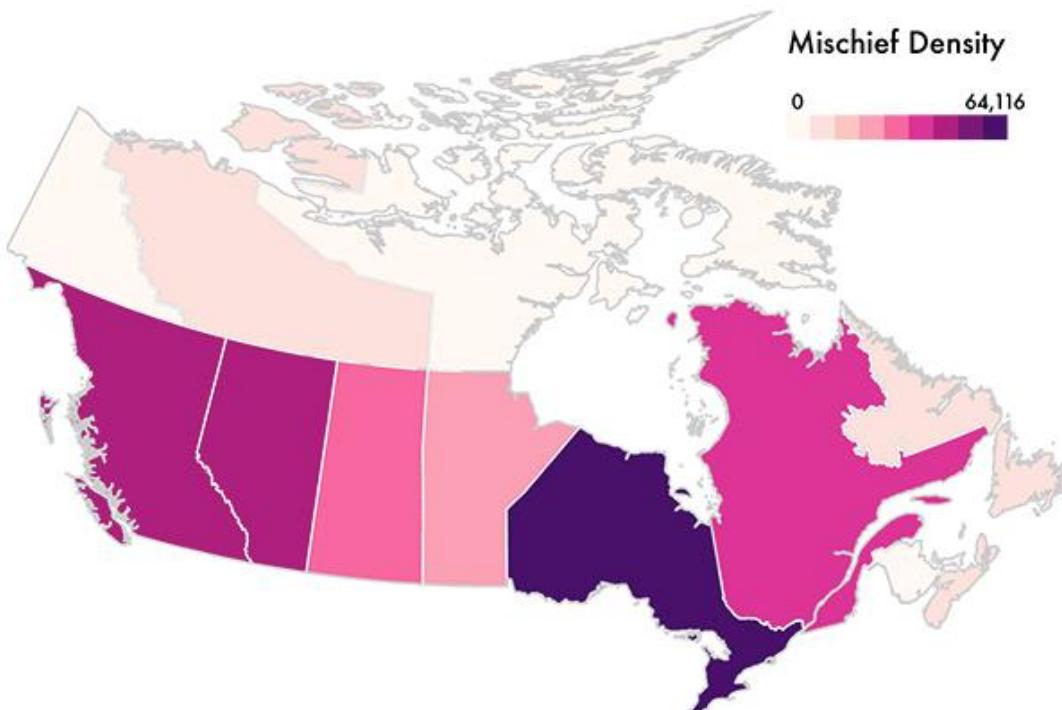
“Above all, show the data”- Tufte

Maximize data-to-ink ratio, minimize perceptual artifacts

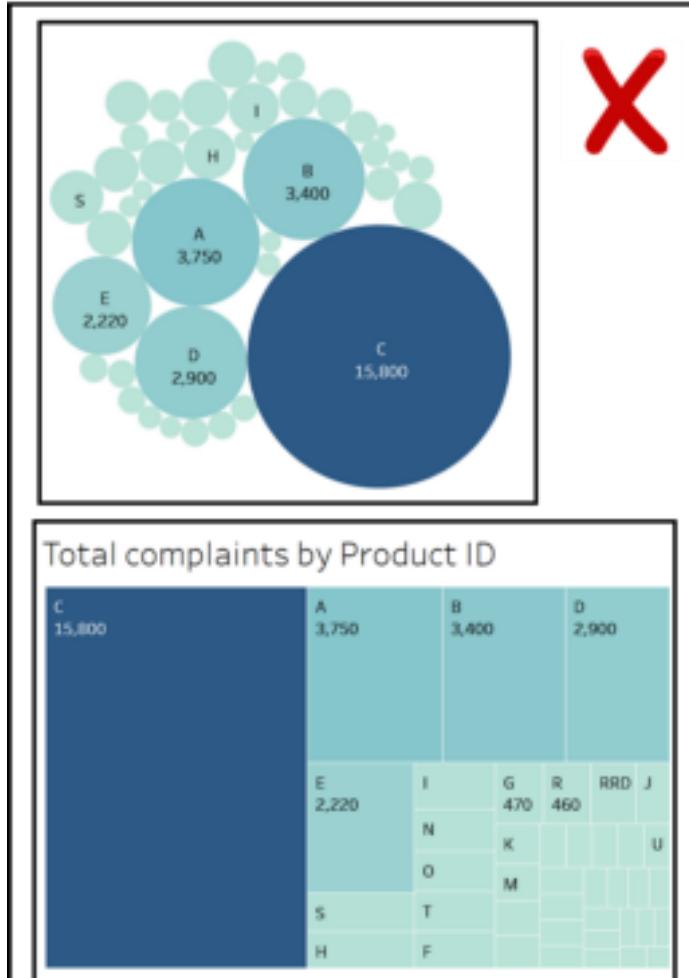
# Emphasize the Message



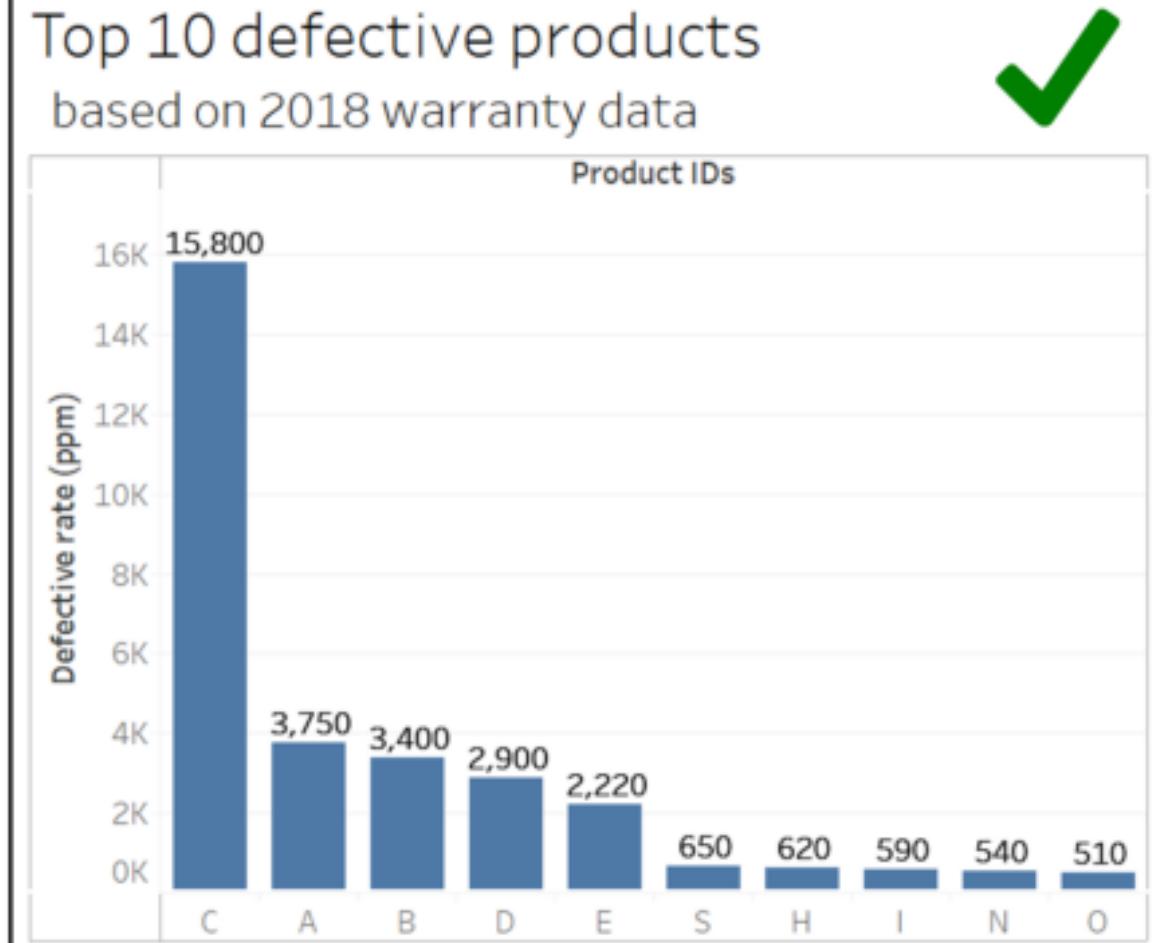
# Beware the choropleth population map



*Ontario has highest “mischief”, but also highest population*



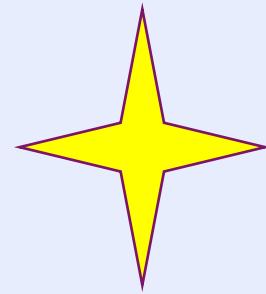
X



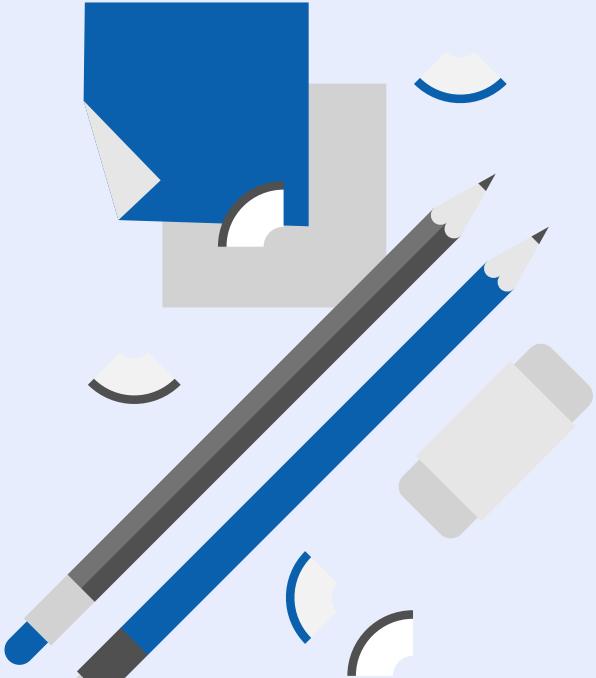
✓

Source: <https://towardsdatascience.com/tips-for-effective-data-visualization-d4b2af91db37>

# Design it!



## Activity #6



**Amber Hoak** • [linkedin.com/in/amberhoak](https://linkedin.com/in/amberhoak)

Published Sep 22

ISC

Fork of Draw Me by D3

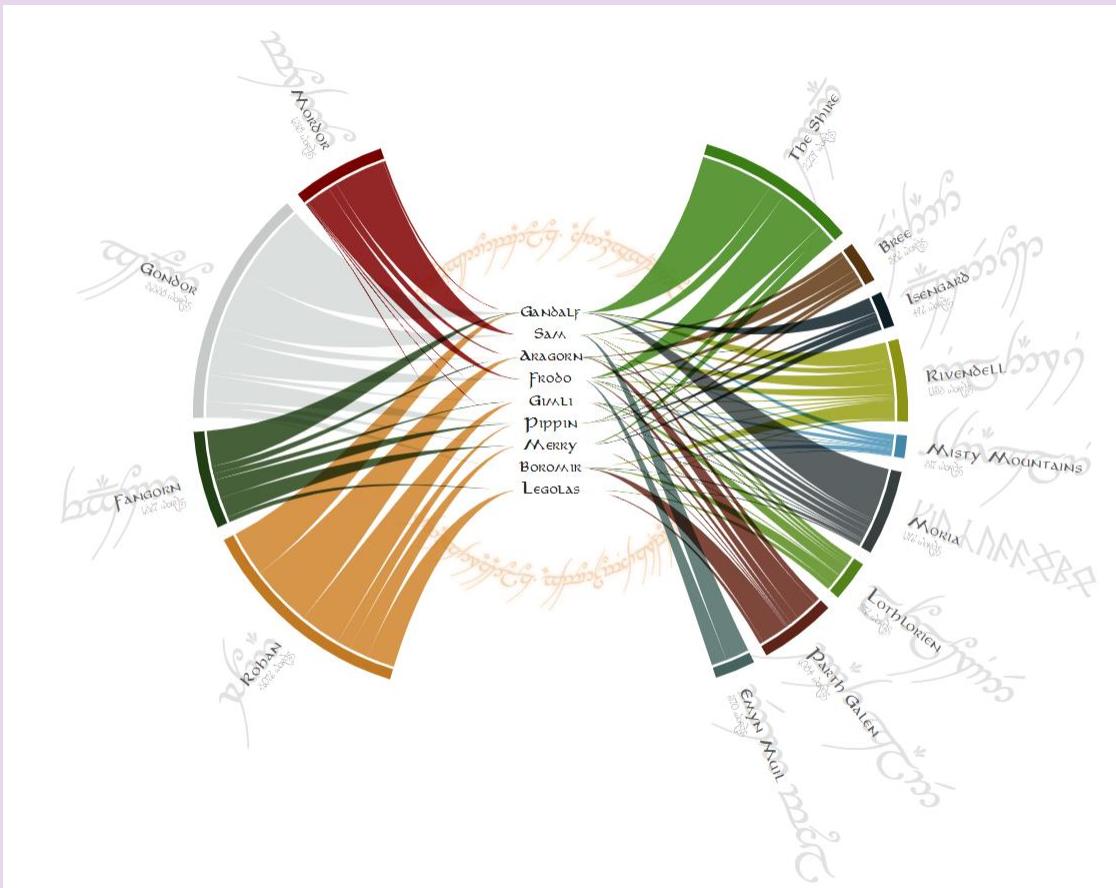
# Designing for Your Data

## Choosing a chart type

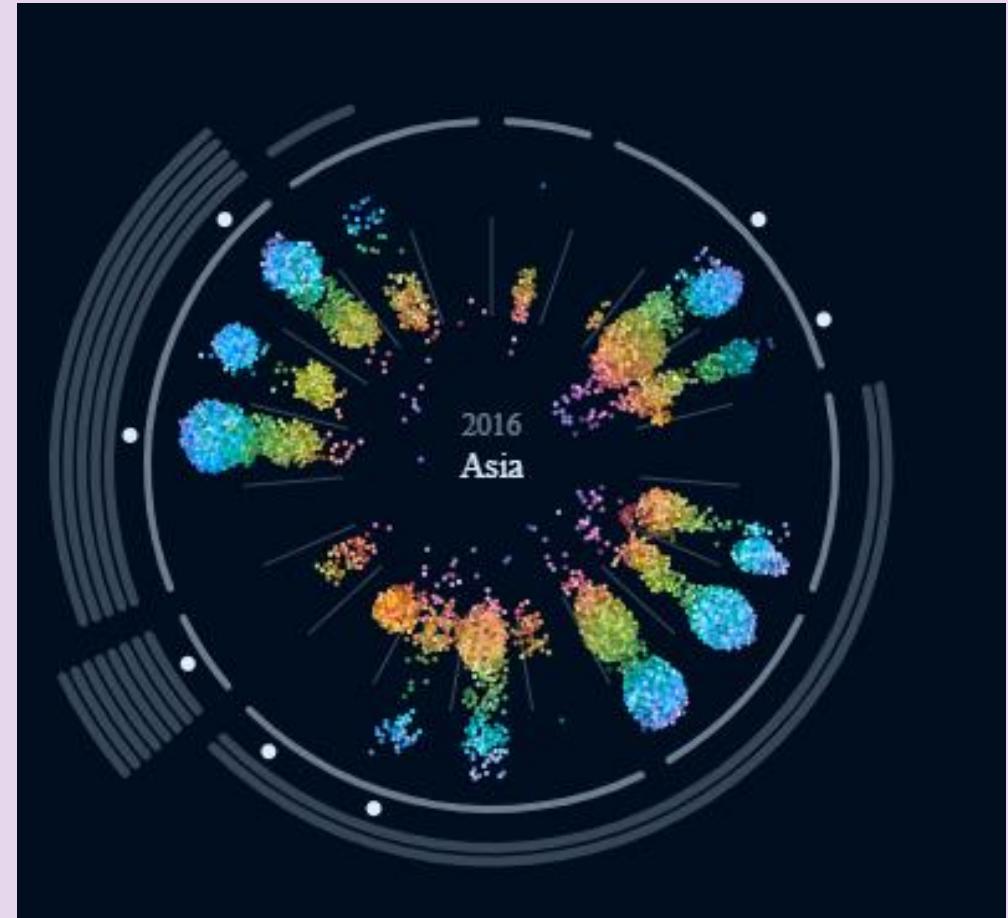
Pick a starting point for your visual based on your dataset. Choose a dataset from:

- Unemployment Rate
- Employment-Population Rate
- Not in the Labor Force
- Looking for Part-Time Employment
- Looking for Full-Time Employment

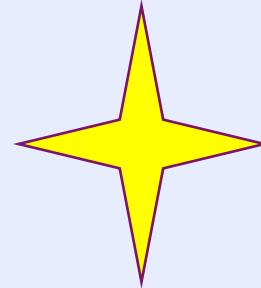
# Creative Approach: Combine Elements



Source: <http://www.datasketch.es/july/code/nadieh/>

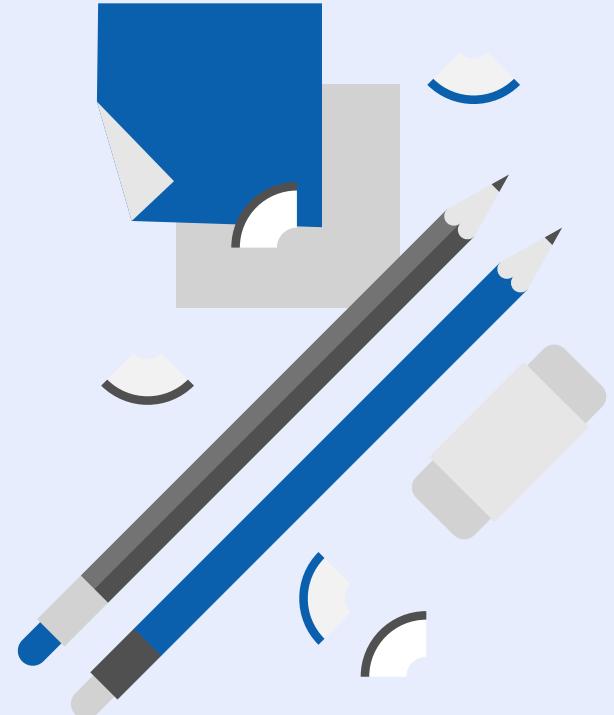
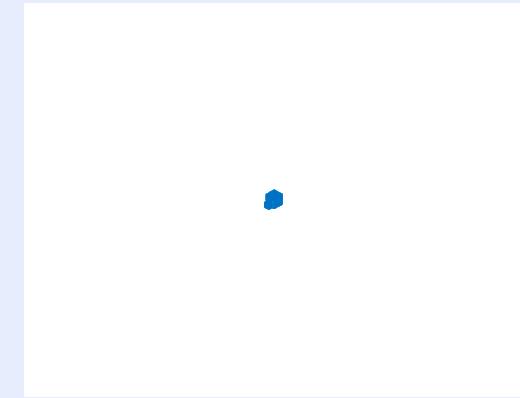


Source: <https://sxywu.com/travel/>

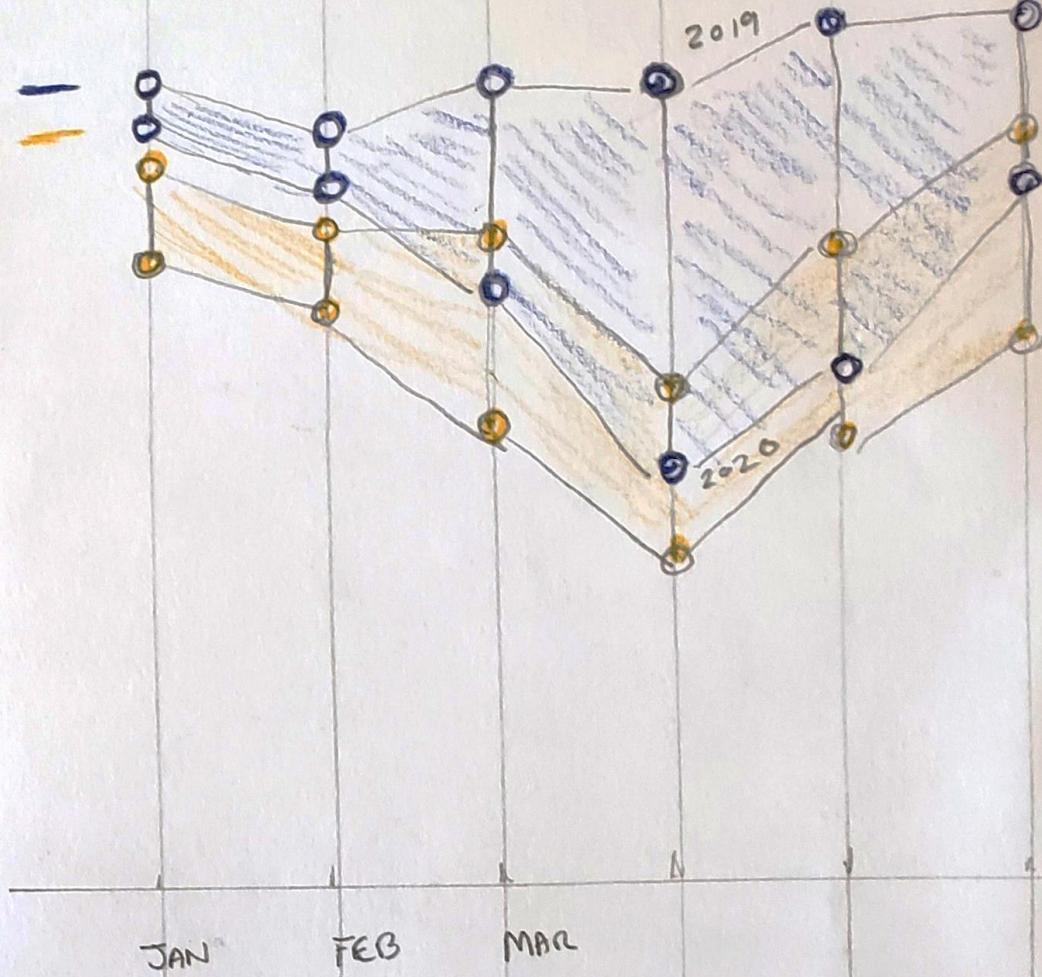
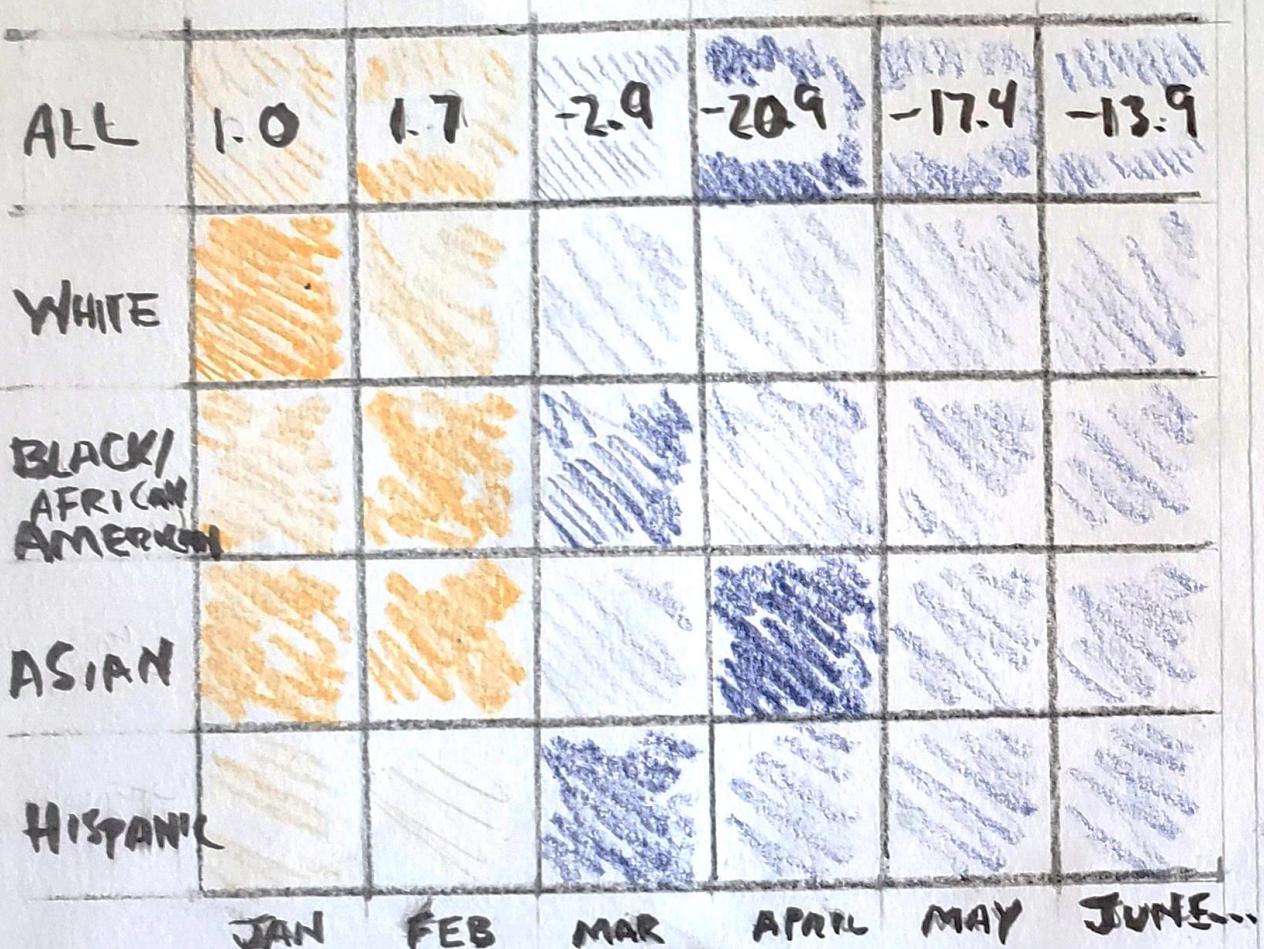


# Design it!

Post your dataset and chart type  
in the chat!



FEMALES 20-24





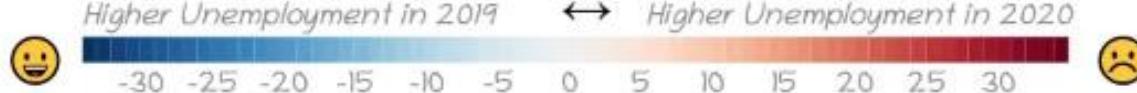
Amber Hoak · [linkedin.com/in/amberhoak](https://linkedin.com/in/amberhoak)

Link shared Sep 21

# How has COVID-19 Affected US Unemployment?

## Unemployment Rate (%) 12 Month Net Change

*Explore the unemployment dataset by selecting a heatmap tile.*



Explore the final visual.

Post an insight that you found in the chat!



## **Software:**

Power BI, Tableau, Datawrapper, Excel, Google Sheets/Studio

## **JavaScript:**

- Chart.js, Google Charts, Semiotic, react-vis, plotly.js
- chart-parts
- D3.js, Vega

## **Python:**

- Bokeh, matplotlib/Seaborn, Altair (Vega), ggplot, plotly.py

## **R:**

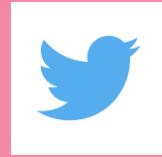
- ggplot2, plotly

# Where to find data?

- **Data.world** : <https://data.world/>
- **The Pudding**: <https://github.com/the-pudding/data>
- **Data.gov**: <https://www.data.gov/>
- **USAFACTS**: <https://usafacts.org/>
- **Our World In Data**:  
<https://ourworldindata.org/>



# Thank you!



@amber\_hoak



/amberhoak

