

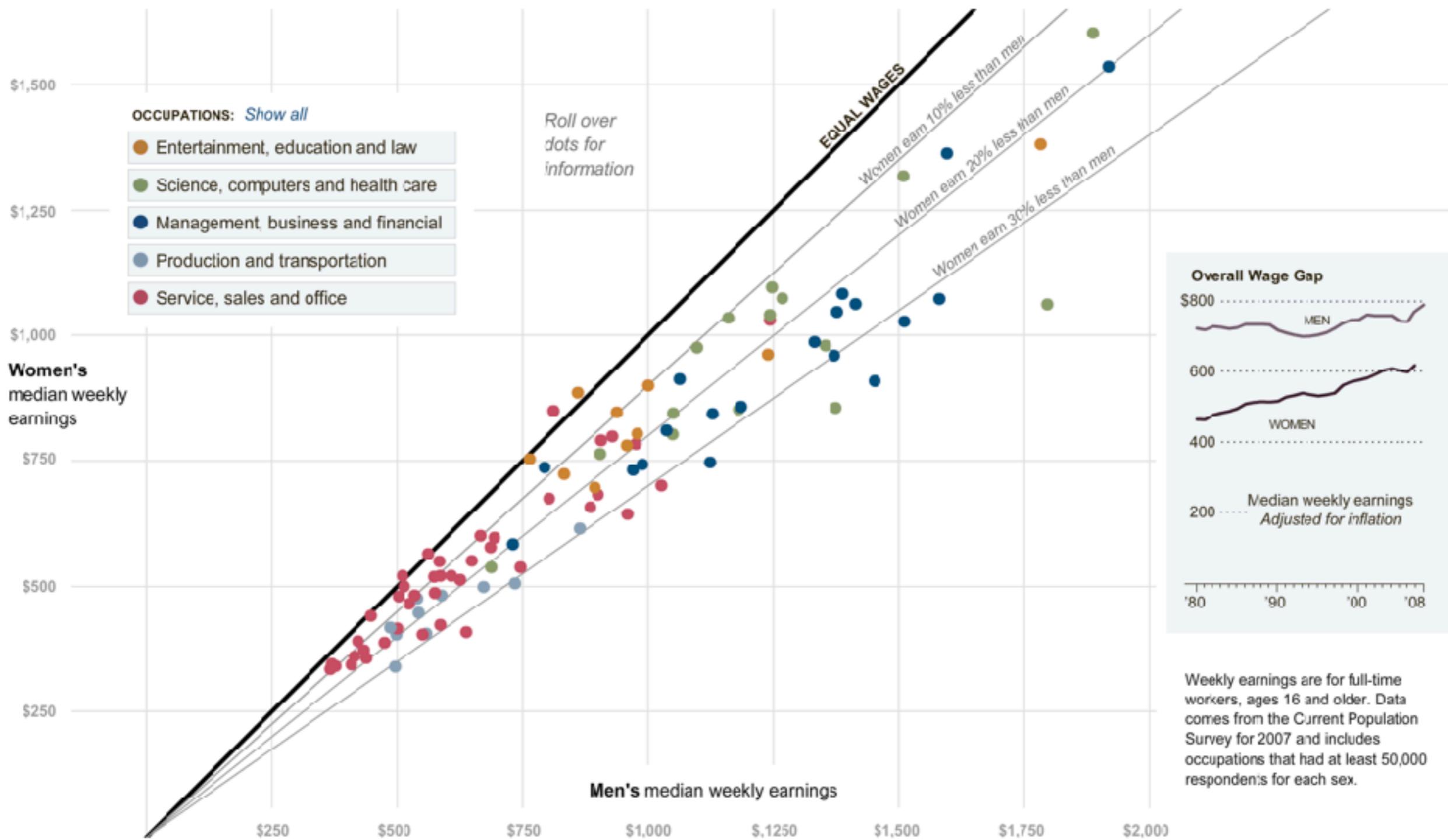
Visual Channels

Jeff Rzeszotarski
Assistant Prof, InfoSci



Why Is Her Paycheck Smaller?

Nearly every occupation has the gap — the seemingly unbridgeable chasm between the size of the paycheck brought home by a woman and the larger one earned by a man doing the same job. Economists cite a few reasons: discrimination as well as personal choices within occupations are two major factors, and part of the gap can be attributed to men having more years of experience and logging more hours.



Hannah Fairfield and Graham Roberts/The New York Times

[Send Feedback](#)

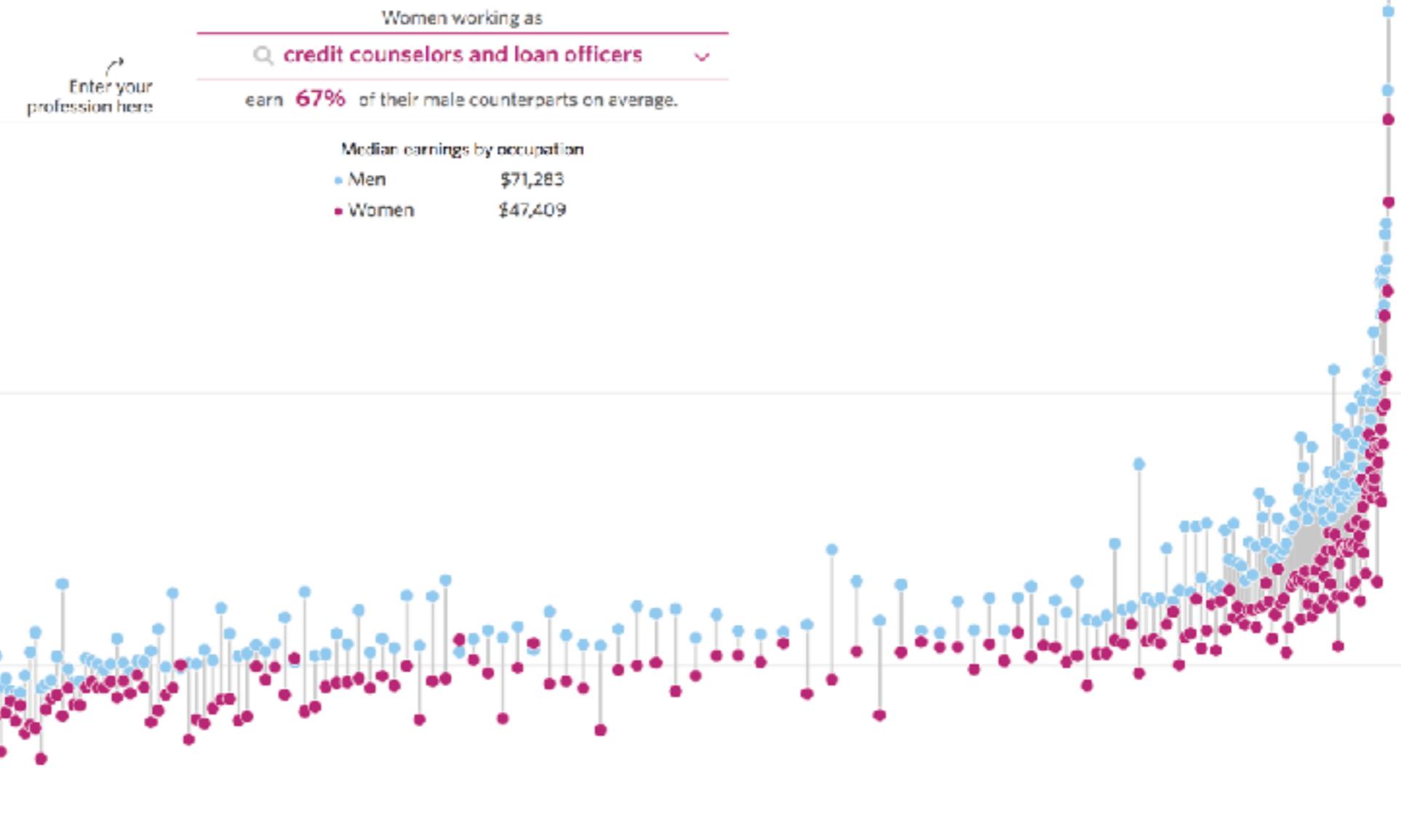
Sources: Bureau of Labor Statistics; Census Bureau

[TWITTER](#)[LINKEDIN](#)[SHARE](#)

What's Your Pay Gap?

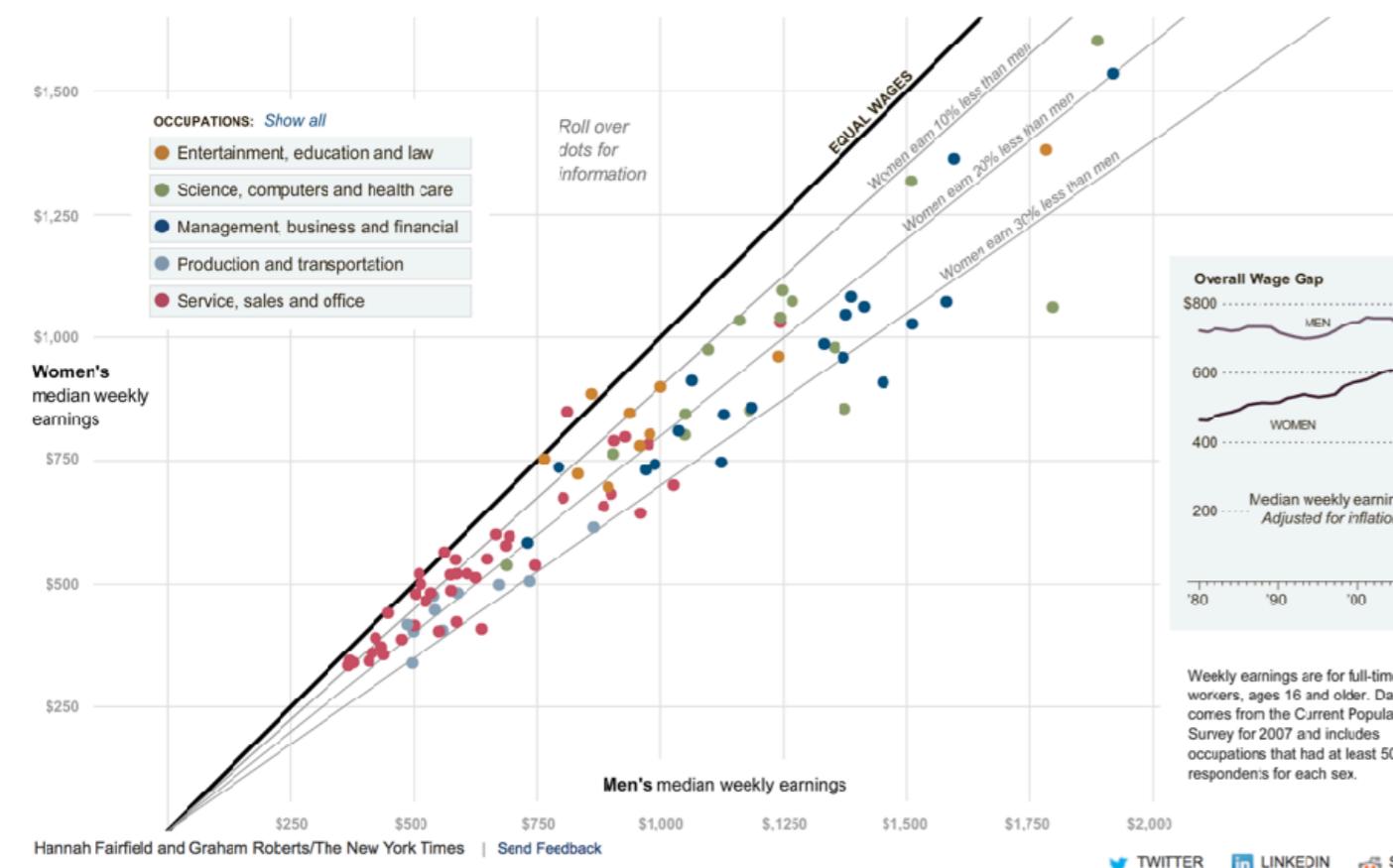
Women earn less than men in 439 of 446 major U.S. occupations, a Wall Street Journal examination of the gender pay gap found.

They earn more than men in seven occupations 

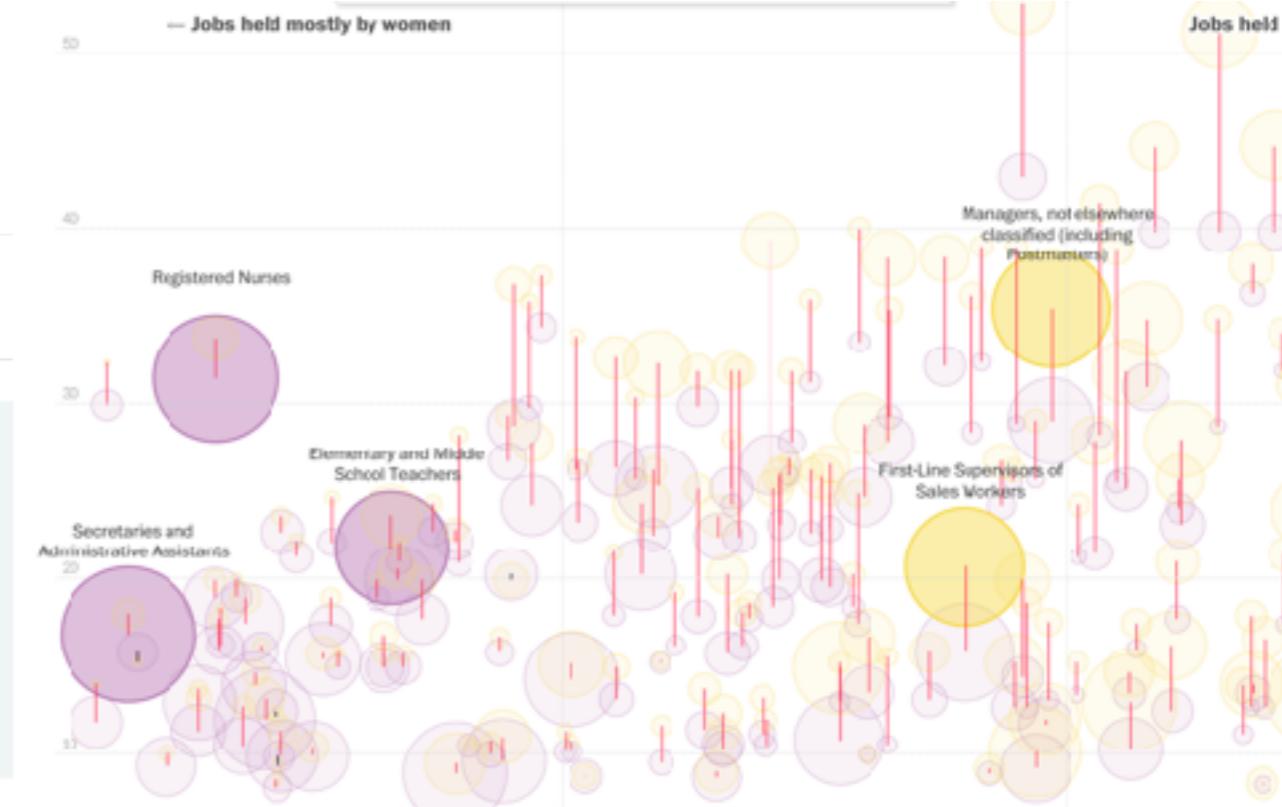
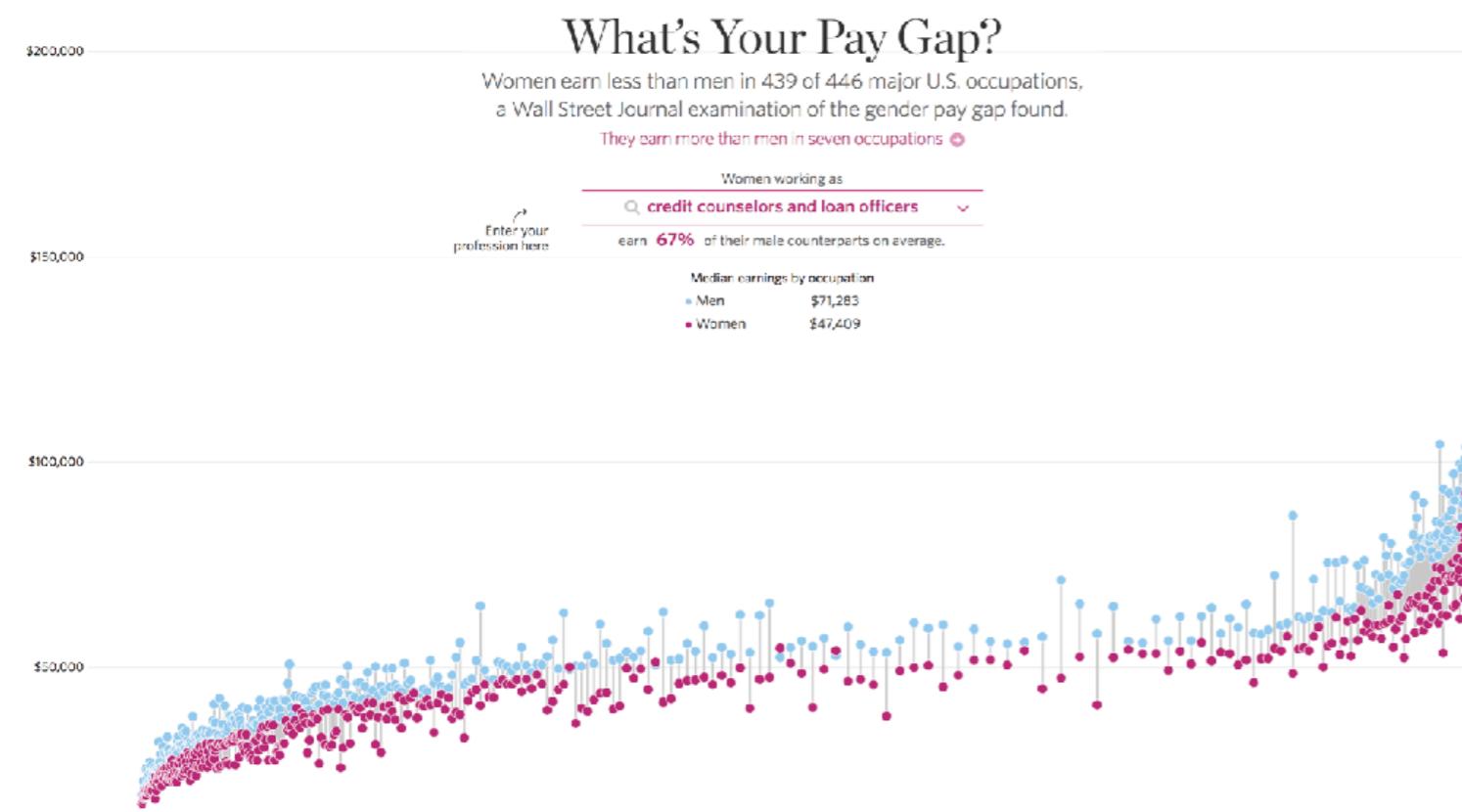


Why Is Her Paycheck Smaller?

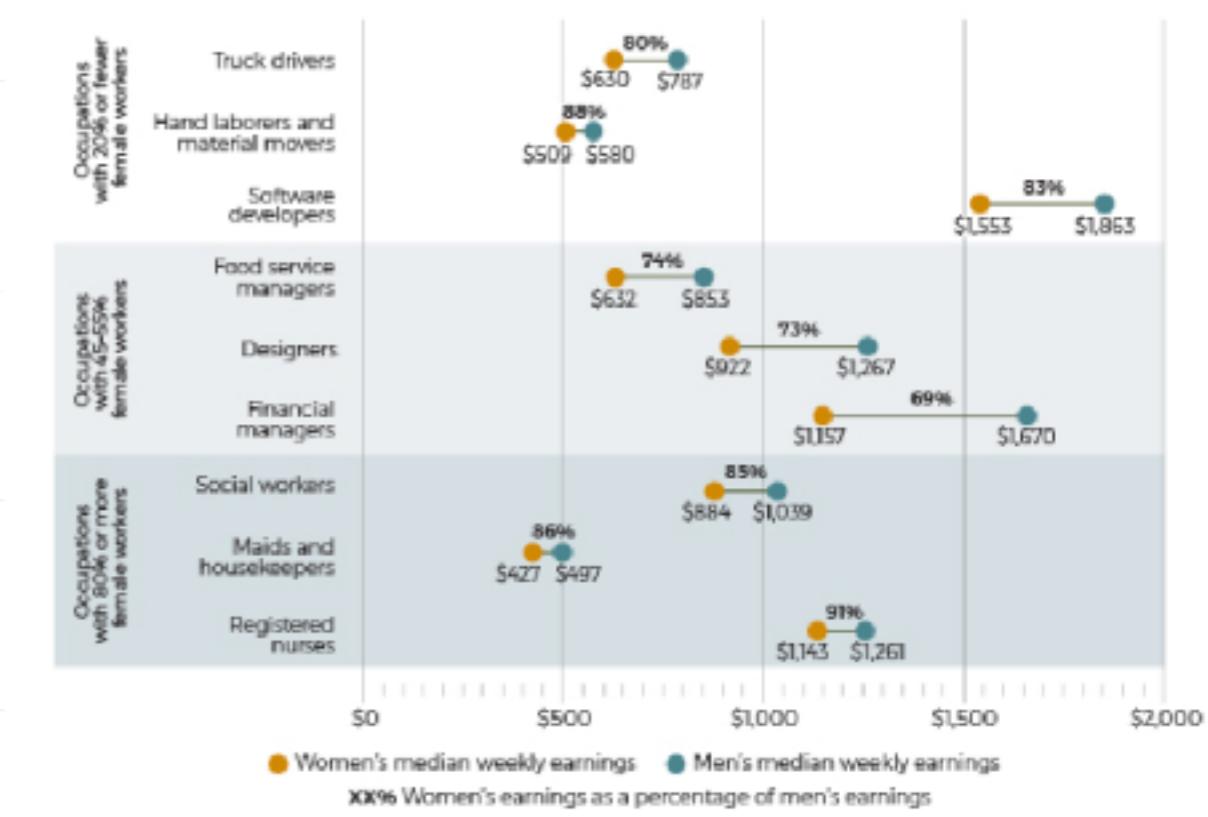
Nearly every occupation has the gap — the seemingly unbridgeable chasm between the size of the paycheck brought home by a woman and the larger one earned by a man doing the same job. Economists cite a few reasons: discrimination as well as personal choices within occupations are two major factors, and part of the gap can be attributed to men having more years of experience and logging more hours.



Sources: Bureau of Labor Statistics; Census Bureau



The Earnings Ratio in Median Weekly Pay among Full-time Workers, Selected Occupations, 2016

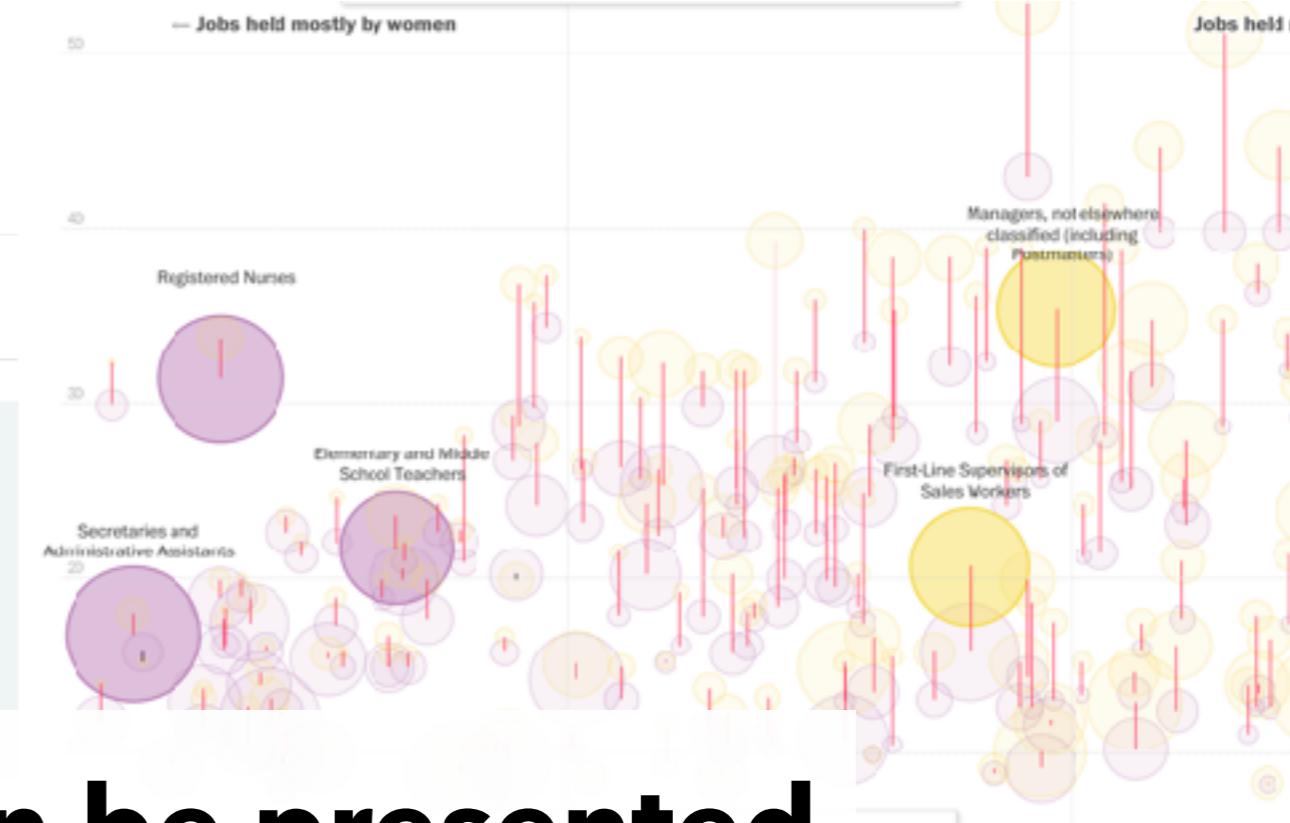


Related: Women in Elite Jobs Face Stubborn Pay Gap ↗

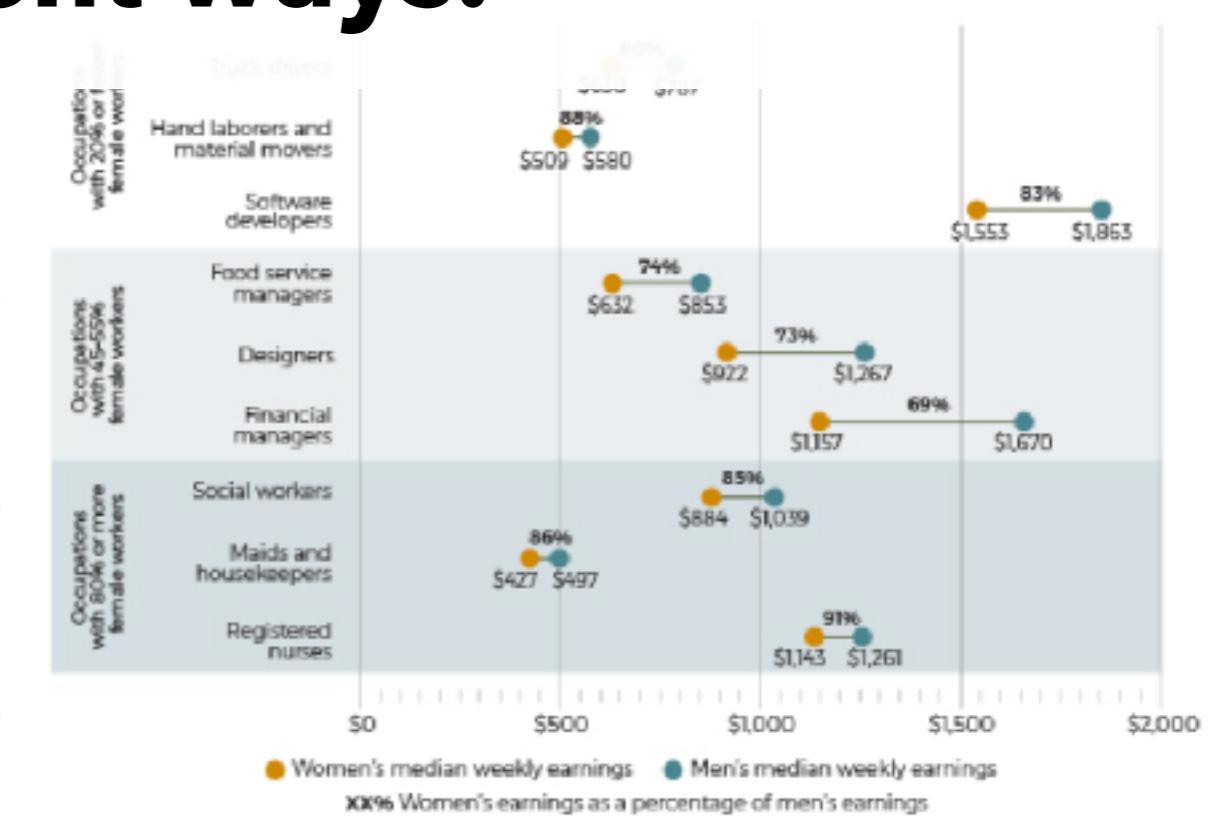
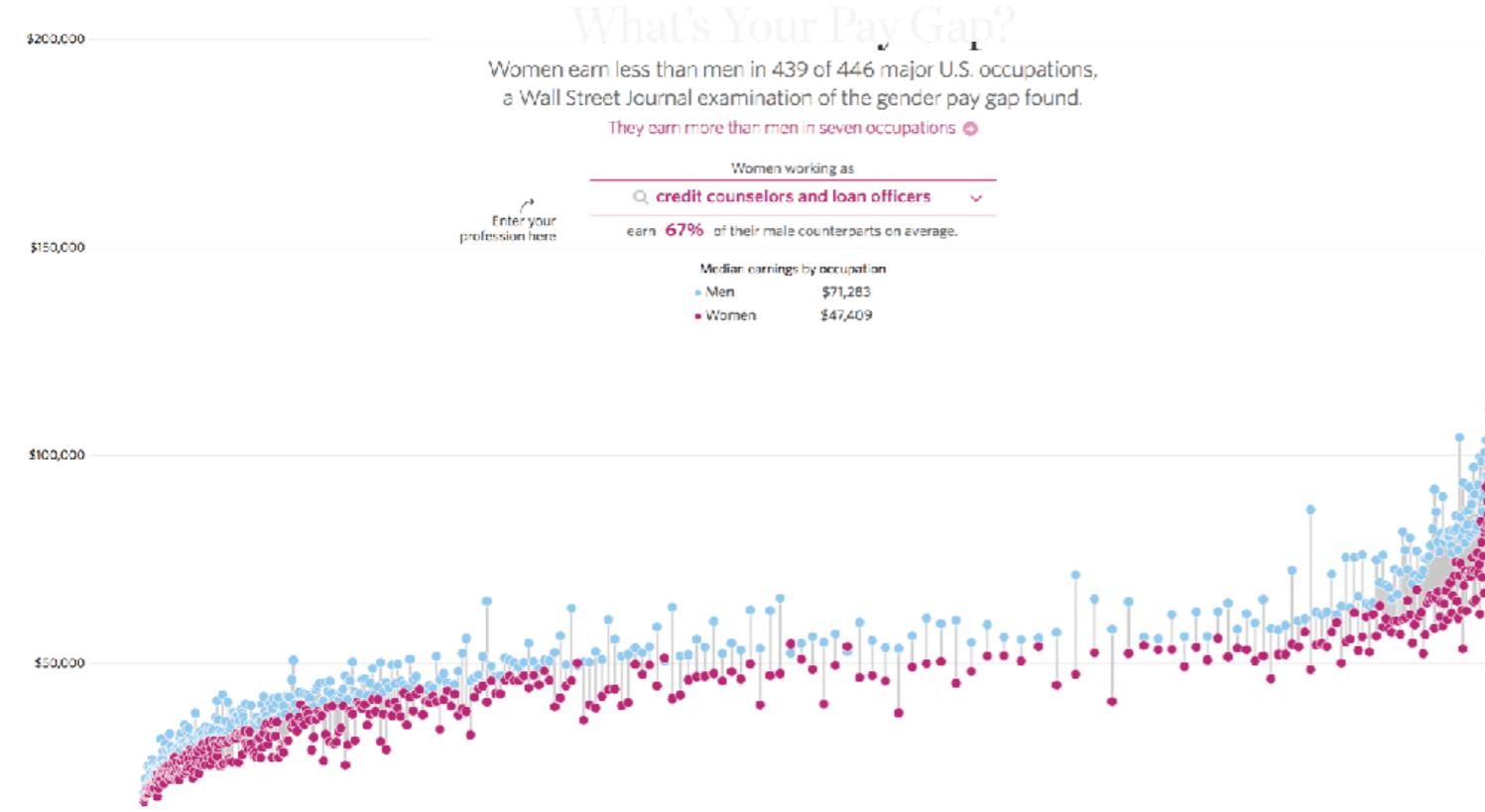


Why Is Her Paycheck Smaller?

Nearly every occupation has the gap — the seemingly unbridgeable chasm between the size of the paycheck brought home by a woman and the larger one earned by a man doing the same job. Economists cite a few reasons: discrimination as well as personal choices within occupations are two major factors, and part of the gap can be attributed to men having more years of experience and logging more hours.



The same data can be presented in many different ways.



Data

| f_1 | f_2 | f_3 |
|-------|-------|-------|
| 110 | 2 | 2 |
| 110 | 2 | 0 |
| 97.7 | 2.3 | 1.5 |
| 90 | 2 | 1 |
| 90 | 3 | 0 |
| 120 | 1 | 2 |
| 110 | 6 | 2 |
| 120 | 1 | 3 |

Data

- **Data Model - “raw” data**
 - Mathematical properties – sets, operators, etc.
 - e.g. 3-dimensional floating point vector

| f_1 | f_2 | f_3 |
|-------|-------|-------|
| 110 | 2 | 2 |
| 110 | 2 | 0 |
| 97.7 | 2.3 | 1.5 |
| 90 | 2 | 1 |
| 90 | 3 | 0 |
| 120 | 1 | 2 |
| 110 | 6 | 2 |
| 120 | 1 | 3 |

Data

- **Conceptual Model - constructs**
 - What the data *mean* (i.e. semanticity)
 - e.g. temperature, grade point, dollars

| Calories | Protein | Fat |
|----------|---------|-----|
| 70 | 4 | 1 |
| 120 | 3 | 5 |
| 70 | 4 | 1 |
| 50 | 4 | 0 |
| 110 | 2 | 2 |
| 110 | 2 | 0 |
| 97.7 | 2.3 | 1.5 |

Data - What they are

- **Data Model - “raw” data**
 - Mathematical properties – sets, operators, etc.
 - e.g. 3-dimensional floating point vector
- **Conceptual Model - constructs**
 - Semanticity
 - e.g. temperature, grade point, dollars



Data - Their use

- **Data Model - “raw” data**
 - Analyze, process, clean
 - Determine possible visual mappings
- **Conceptual Model - constructs**
 - Provide meaning, direct hypotheses
 - Visual alignment with semantic concepts
(e.g. “why are cold temperatures blue?”)



Kinds of Data

- **Nominal (labels, categories, groups)**

Dog breeds, ?????

- **Ordered (ordinal, rankings)**

Class grade, ?????

- **Quantitative**

- **Interval** – No fixed zero point

Dates, ?????

- **Ratio** – Fixed zero point

Integer numbers, ?????

S. S. Stevens, On the theory of scales of measurements, 1946



Kinds of Data

- **Nominal (labels, categories, groups)**

Dog breeds, party affiliation, wine varietals

- **Ordered (ordinal, rankings)**

Class grade, “size words”, “discretized” scales

- **Quantitative**

- **Interval** – No fixed zero point

Dates, SAT score, credit score – No universal “year 0”

- **Ratio** – Fixed zero point

Count, Things on a number line – Universal “0 years old”

S. S. Stevens, On the theory of scales of measurements, 1946



Example

- **Data**

44.0, 54.2, 78.4, 42.1, 102.3

- **Concept**

Temperature in Celsius

- **Kinds**

- Nominal

???

- Ordered

???

- Quantitative

(interval or ratio?)

S. S. Stevens, On the theory of scales of measurements, 1946



Example

- **Data**

44.0, 54.2, 78.4, 42.1, 102.3

- **Concept**

Temperature in Celsius

- **Kinds**

- Nominal

Water freezes / water doesn't freeze

- Ordered

Warm, cold, freezing

- Quantitative

Temperature measure

S. S. Stevens, On the theory of scales of measurements, 1946



| Name | Manufacturer | Calories | Protein | Fat | Sodium | Fiber |
|---------------------------|---------------------|-----------------|----------------|------------|---------------|--------------|
| 100% Bran | Nabisco | 70 | 4 | 1 | 130 | |
| 100% Natural Bran | Quaker Oats | 120 | 3 | 5 | 15 | |
| All-Bran | Kellogg's | 70 | 4 | 1 | 260 | |
| All-Bran with Extra Fiber | Kellogg's | 50 | 4 | 0 | 140 | |
| Apple Cinnamon Cheerios | General Mills | 110 | 2 | 2 | 180 | |
| Apple Jacks | Kellogg's | 110 | 2 | 0 | 125 | |
| Basic 4 | General Mills | 97.7 | 2.3 | 1.5 | 157.9 | |
| Bran Chex | Ralston Purina | 90 | 2 | 1 | 200 | |
| Bran Flakes | Post | 90 | 3 | 0 | 210 | |
| Cap'n'Crunch | Quaker Oats | 120 | 1 | 2 | 220 | |
| Cheerios | General Mills | 110 | 6 | 2 | 290 | |
| Cinnamon Toast Crunch | General Mills | 120 | 1 | 3 | 210 | |
| Clusters | General Mills | 110 | 3 | 2 | 140 | |
| Cocoa Puffs | General Mills | 110 | 1 | 1 | 180 | |
| Corn Chex | Ralston Purina | 110 | 2 | 0 | 280 | |
| Corn Flakes | Kellogg's | 100 | 2 | 0 | 290 | |
| Corn Pops | Kellogg's | 110 | 1 | 0 | 90 | |
| Count Chocula | General Mills | 110 | 1 | 1 | 180 | |
| Cracklin' Oat Bran | Kellogg's | 110 | 3 | 3 | 140 | |

Crispix

Crispy Wheat & Raisins

Nominal, Ordered, Quantitative (Interval, Ratio)



Operations

- **Nominal (labels, categories, groups)**

= ≠ ∈ ∉

- **Ordered (ordinal, rankings)**

= ≠ ∈ ∉ < >

- **Quantitative**

- **Interval** – only differences

= ≠ < > + -

- **Ratio** – proportions

= ≠ < > + - × ÷ %

S. S. Stevens, On the theory of scales of measurements, 1946



Common Conventions

Tableau

- **Dimensions**

“Categories” of data – N & O

Discrete, able to be binned

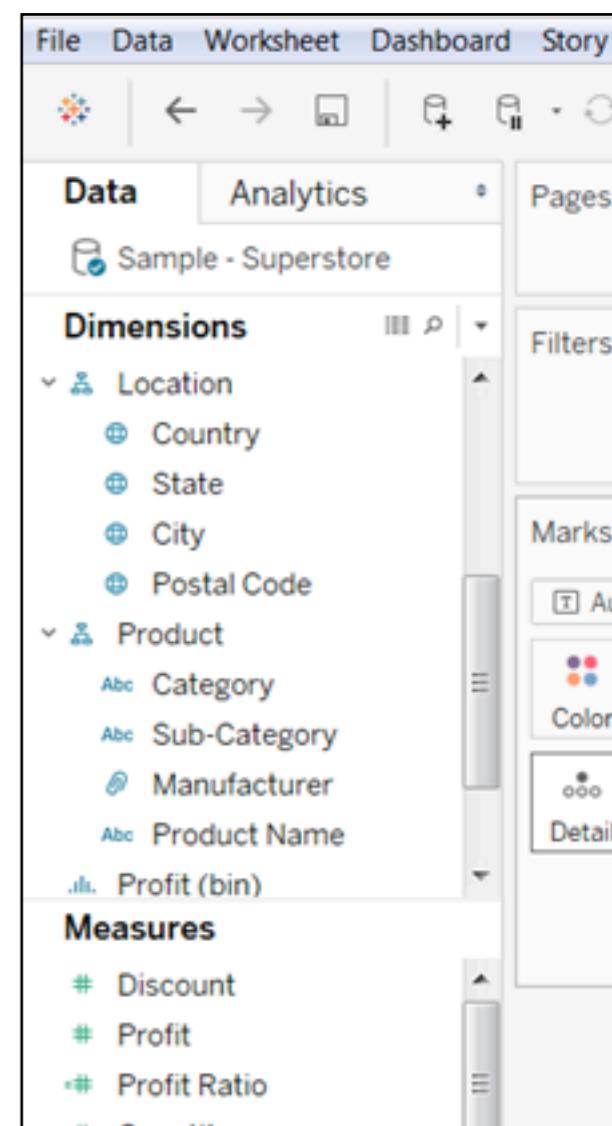
Species, year, grade, date

- **Measures**

“Mathematical” data - Q

Numeric, able to be aggregated w/functions

Petal width, height, temperature, grade point



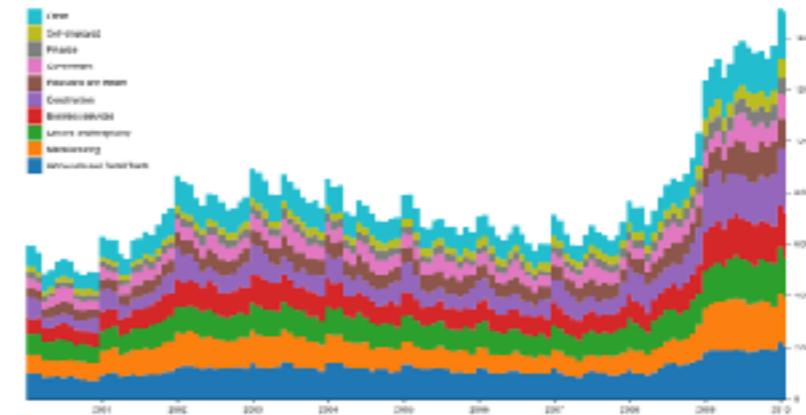
| name | Manufacturer | Calories | Protein | Fat | Sodium | Fib |
|---------------------------|---------------------|-----------------|----------------|------------|---------------|------------|
| 100% Bran | Nabisco | 70 | 4 | 1 | 130 | |
| 100% Natural Bran | Quaker Oats | 120 | 3 | 5 | 15 | |
| All-Bran | Kellogg's | 70 | 4 | 1 | 260 | |
| All-Bran with Extra Fiber | Kellogg's | 50 | 4 | 0 | 140 | |
| Apple Cinnamon Cheerios | General Mills | 110 | 2 | 2 | 180 | |
| Apple Jacks | Kellogg's | 110 | 2 | 0 | 125 | |
| Basic 4 | General Mills | 97.7 | 2.3 | 1.5 | 157.9 | |
| Bran Chex | Ralston Purina | 90 | 2 | 1 | 200 | |
| Bran Flakes | Post | 90 | 3 | 0 | 210 | |
| Cap'n'Crunch | Quaker Oats | 120 | 1 | 2 | 220 | |
| Cheerios | General Mills | 110 | 6 | 2 | 290 | |
| Cinnamon Toast Crunch | General Mills | 120 | 1 | 3 | 210 | |
| Clusters | General Mills | 110 | 3 | 2 | 140 | |
| Cocoa Puffs | General Mills | 110 | 1 | 1 | 180 | |
| Corn Chex | Ralston Purina | 110 | 2 | 0 | 280 | |
| Corn Flakes | Kellogg's | 100 | 2 | 0 | 290 | |
| Corn Pops | Kellogg's | 110 | 1 | 0 | 90 | |
| Count Chocula | General Mills | 110 | 1 | 1 | 180 | |
| Cracklin' Oat Bran | Kellogg's | 110 | 3 | 3 | 140 | |
| Crispix | Kellogg's | 11 | | | | |
| Crispy Wheat & Raisins | General Mills | 10 | | | | |

Dimensions, Measures



Visual Encodings

Data →



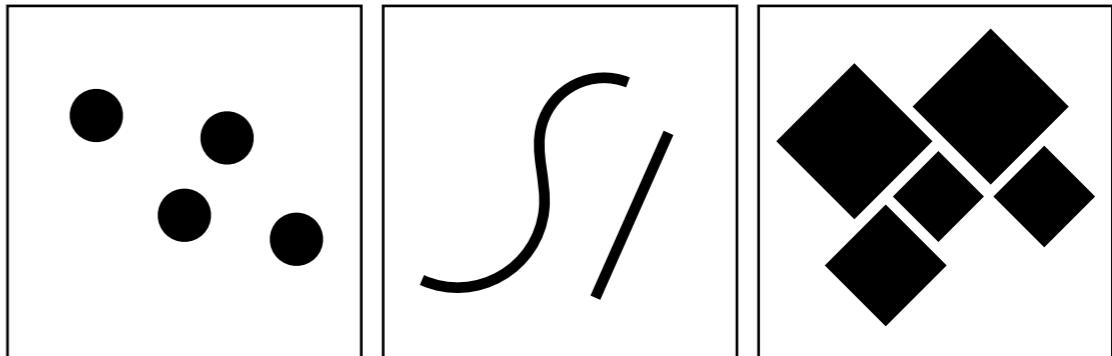
The process of mapping data attributes to visual features

Credit: Mike Bostock



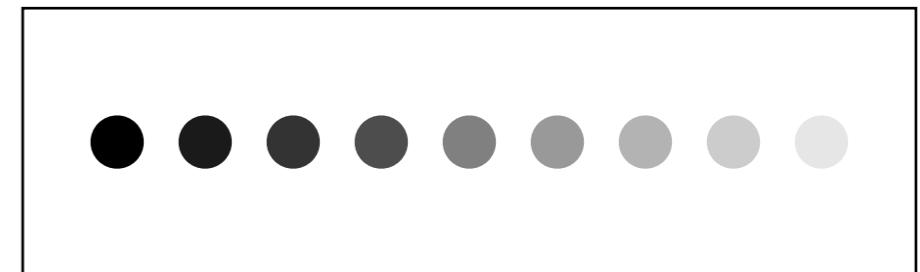
Marks

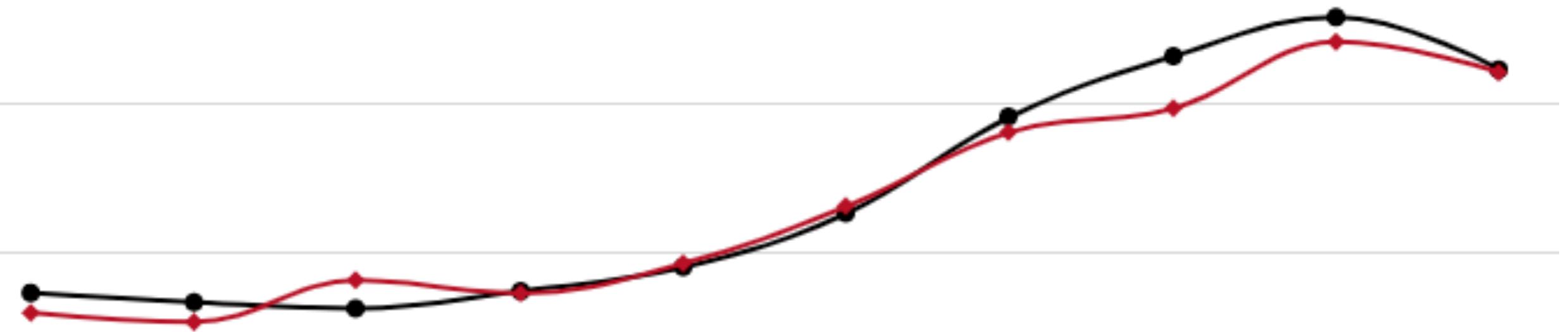
entities, links, objects

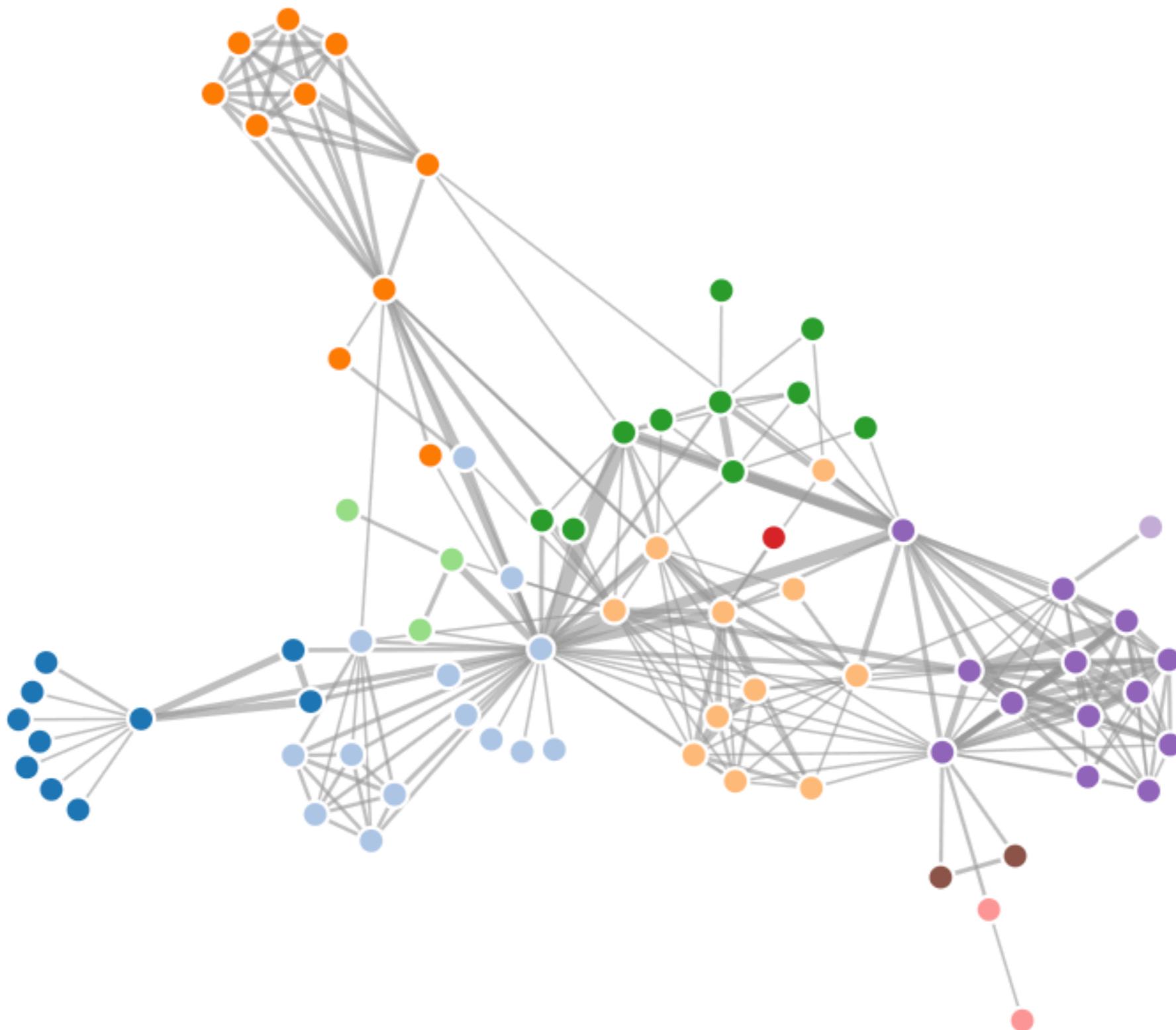


Channels

change based on data values

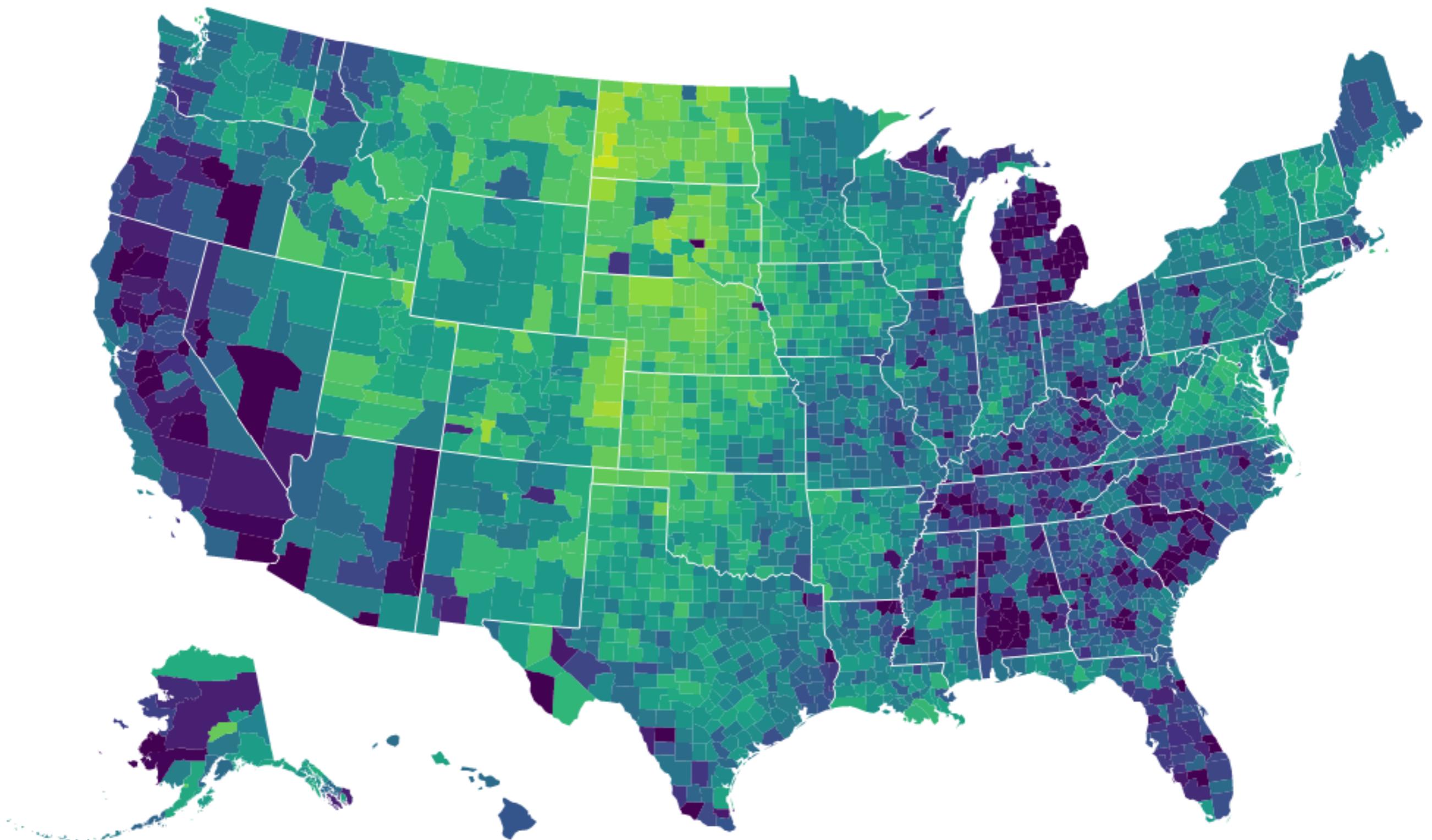






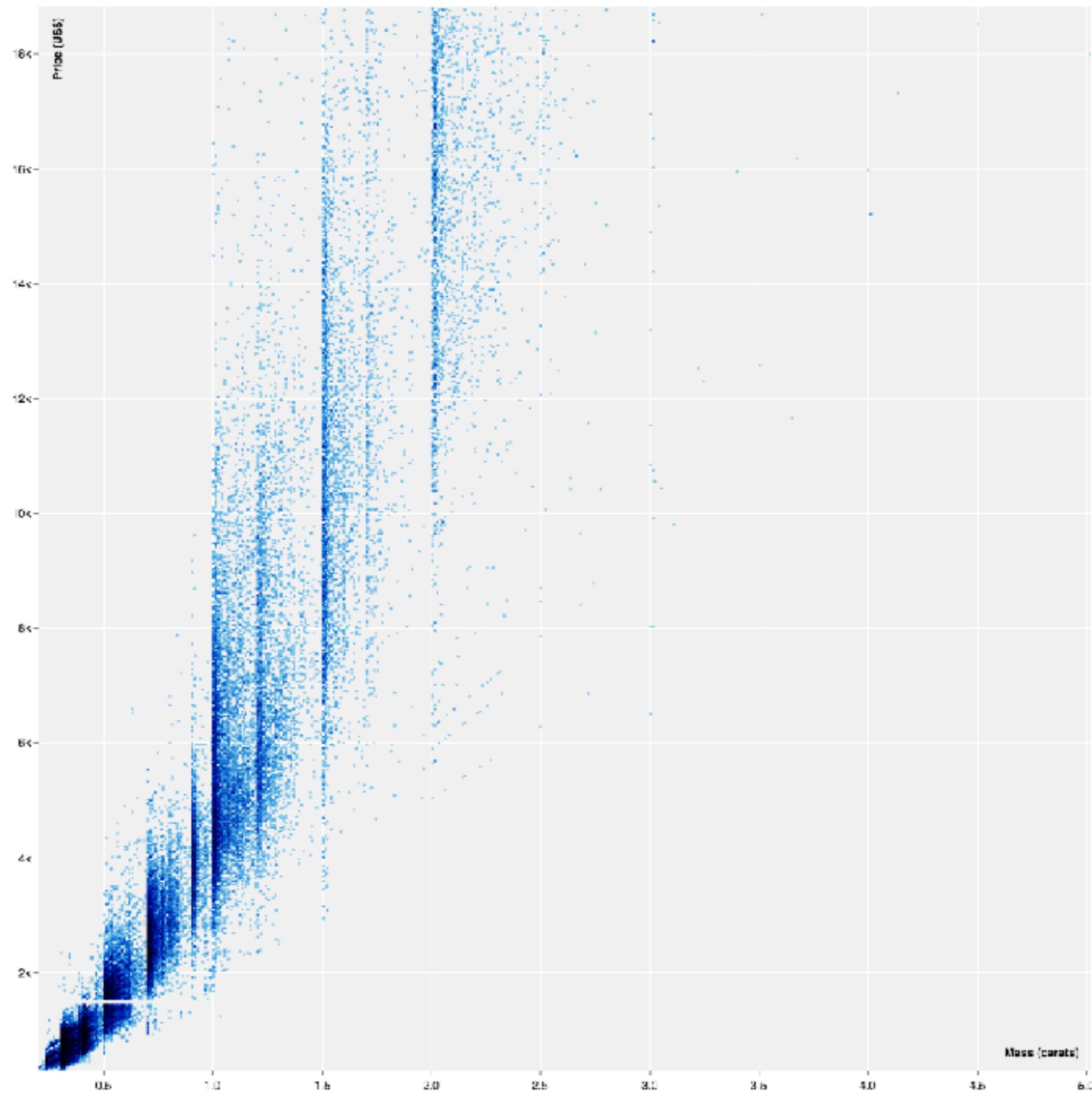
Credit: Mike Bostock





Credit: Mike Bostock

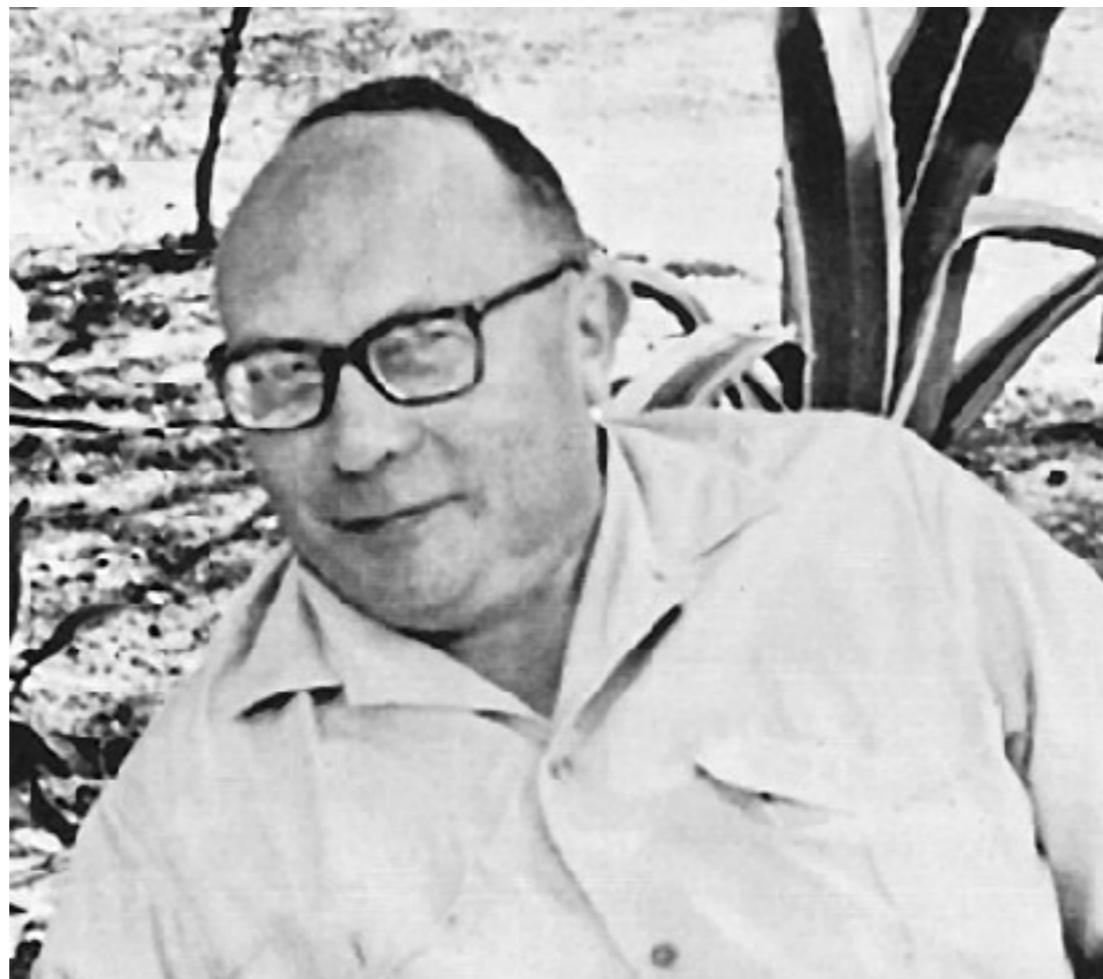




Credit:
Mike Bostock

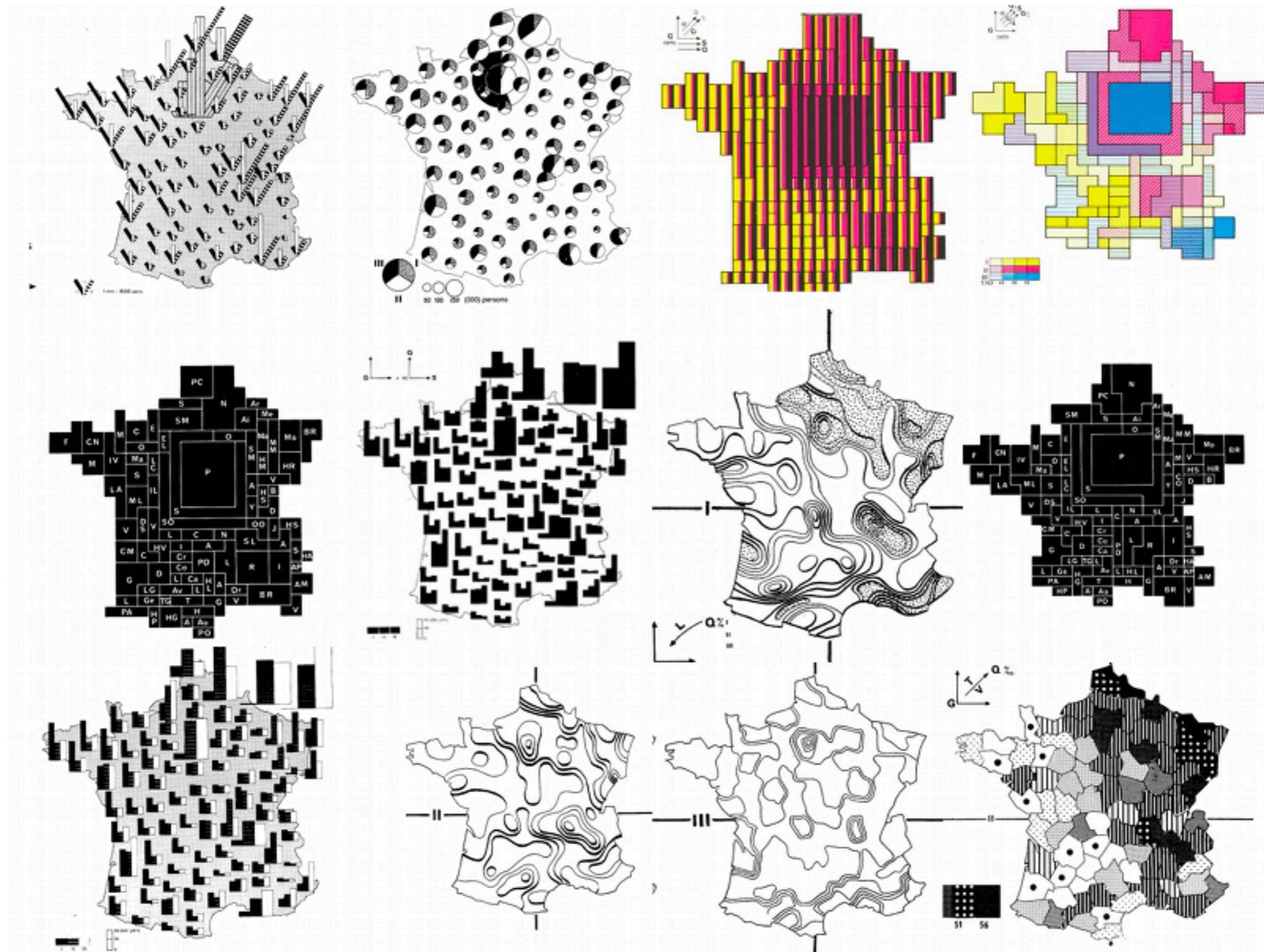


Jacques Bertin



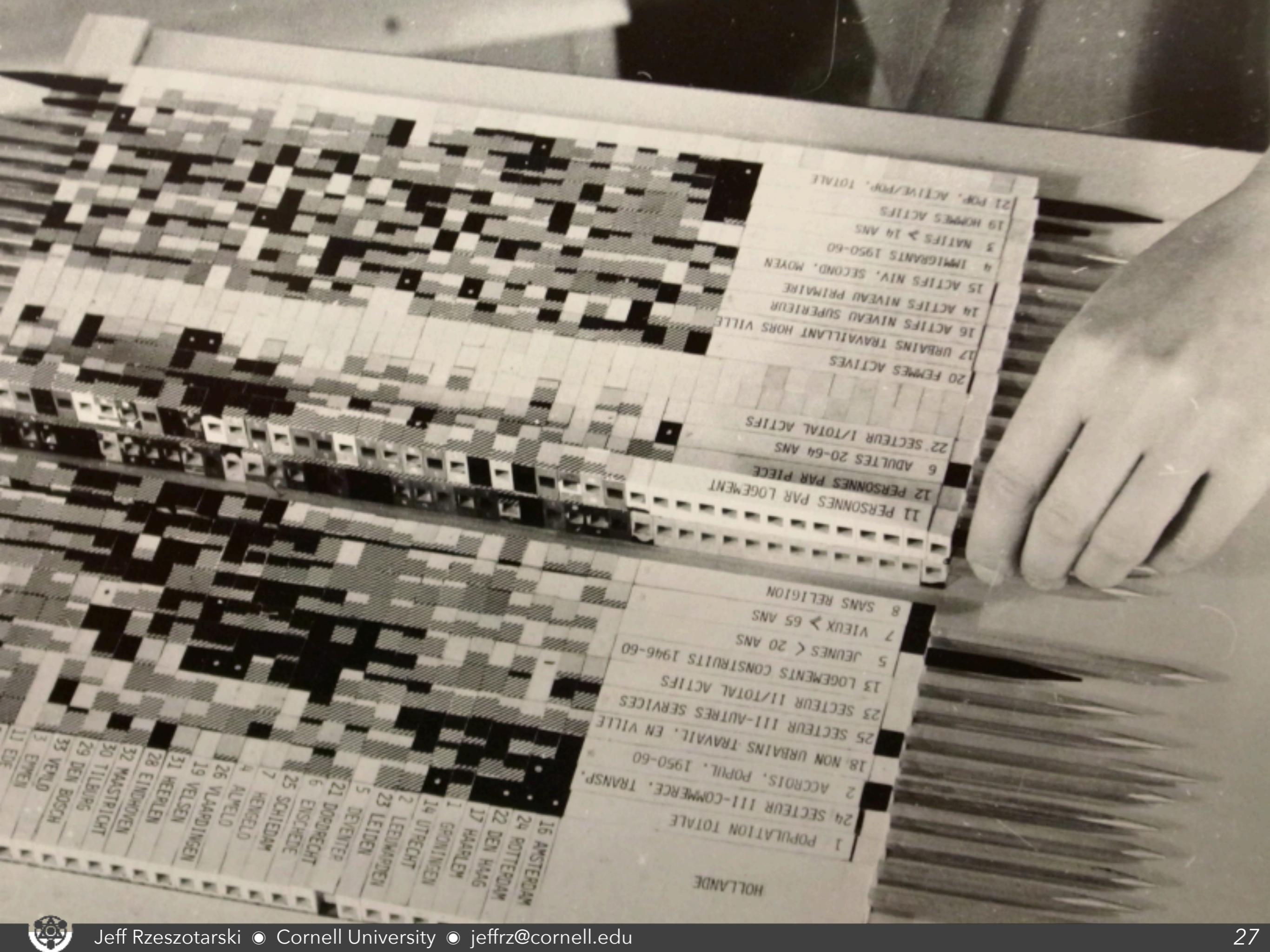
Thanks,
Wikipedia



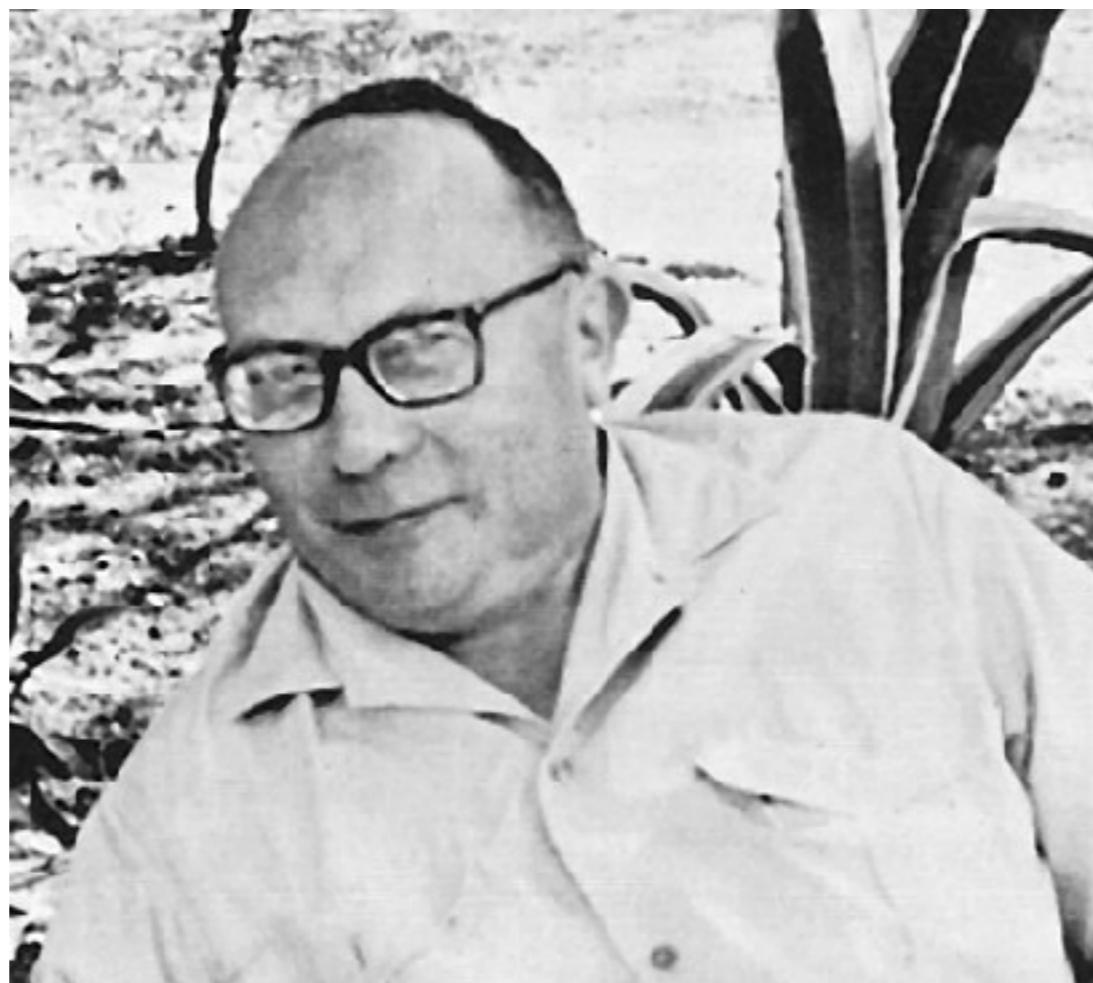


Bertin, Semiology of Graphics, 1983 ed





Jacques Bertin



Semiology of Graphics - 1967

Thanks,
Wikipedia



Semiotics

- “study of **meaning making**” Oxford English Dictionary
- “a general philosophical theory of **signs and symbols** that deals especially with their function in both **artificially constructed and natural languages** and comprises **syntactics, semantics, and pragmatics**”
Merriam Webster

(semiology is a tiny bit different, but that's a whole different course than this one)



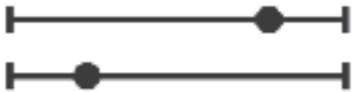
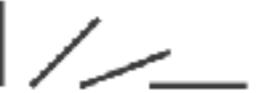
C h a n n e l s

LES VARIABLES DE L'IMAGE

| | POINTS | LIGNES | ZONES |
|--|-----------|-----------|------------------------------------|
| XY 2 DIMENSIONS DU PLAN | x x x | 3 3 1 | 15 9 14 1 18 21 2 14 15 1 |
| Z TAILLE | — — — | — — — | — — — |
| VALEUR | — — — | — — — | — — — |
| LES VARIABLES DE SÉPARATION DES IMAGES | | | |
| GRAIN | — — — | — — — | — — — |
| COULEUR | — — — | — — — | — — — |
| ORIENTATION | — — — | — — — | — — — |
| FORME | — — — | — — — | — — — |

Bertin, Semiology of Graphics, 1983 ed



| | | |
|-----------------------------|---|---|
| Position on common scale |  | |
| Position on unaligned scale |  | |
| Length (1D size) |  | |
| Tilt/angle |  | |
| Area (2D size) |  | |
| Depth (3D position) |  | |
| Color luminance |  | |
| Color saturation |  |  |
| Curvature |  |  |
| Volume (3D size) |  | |



Visual Channels



Position on common scale



Position on unaligned scale



Length (1D size)



Tilt/angle



Area (2D size)



Depth (3D position)



Color luminance



Color saturation



Curvature



Volume (3D size)



Spatial region



Color hue



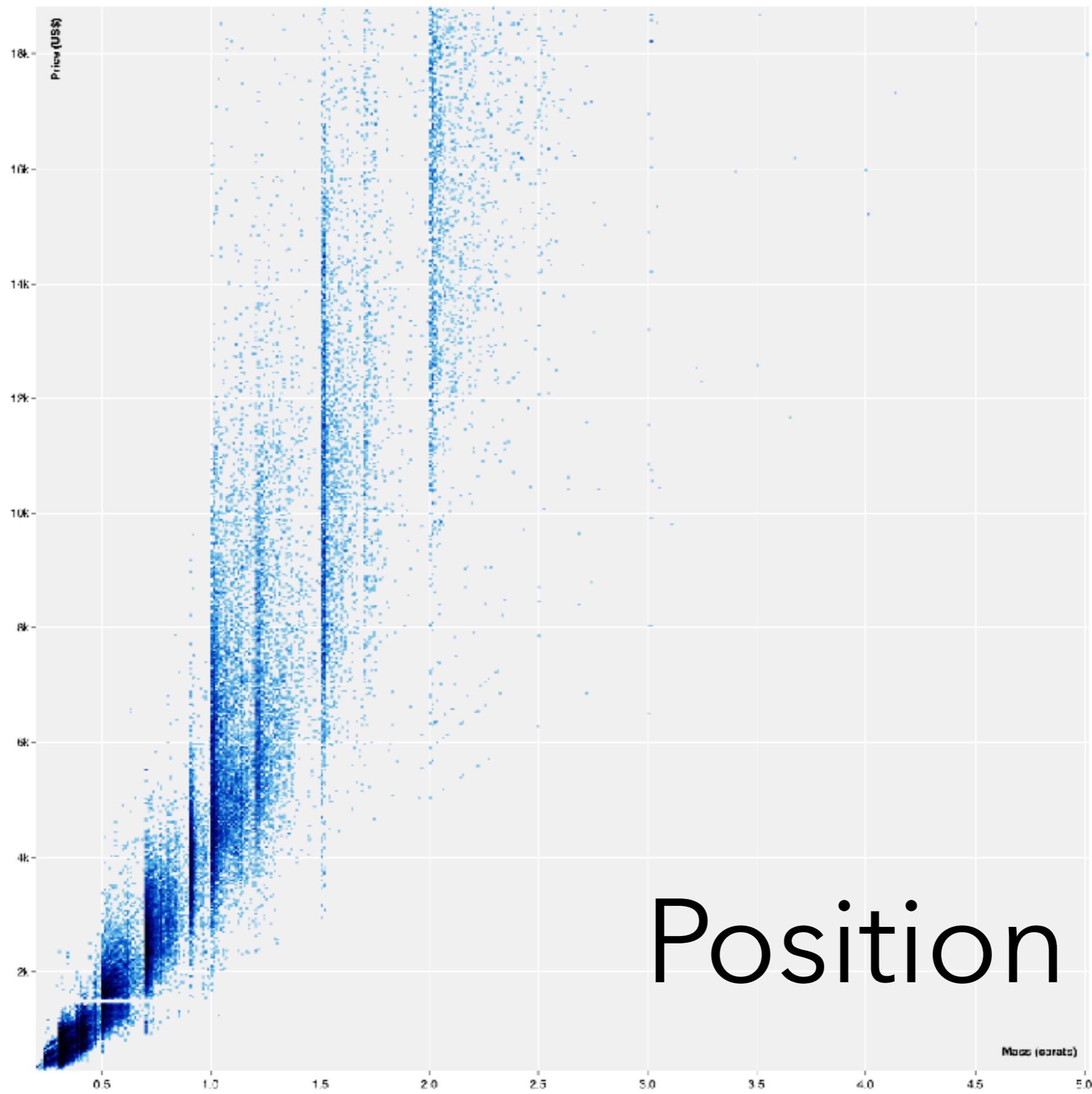
Motion



Shape



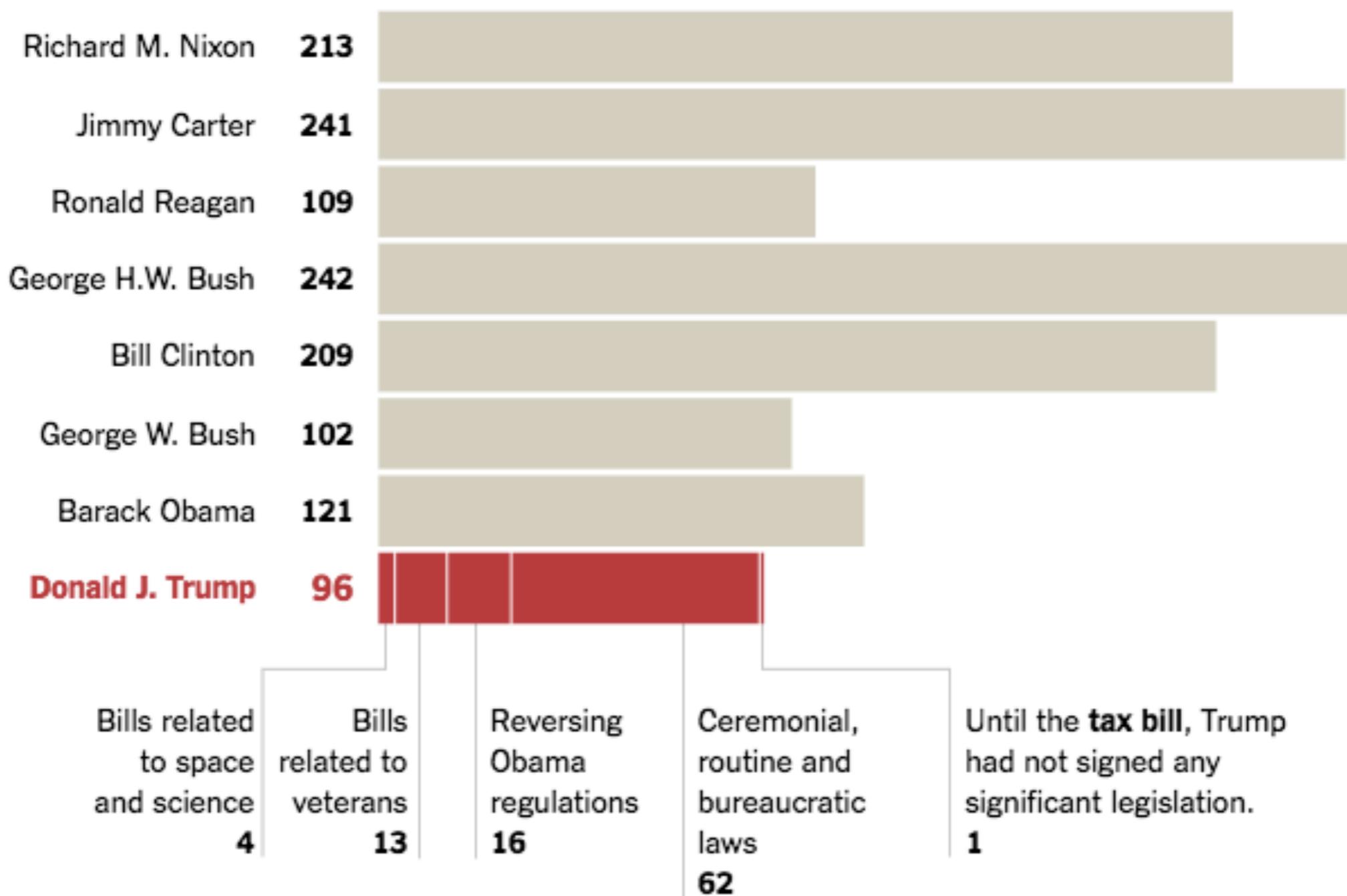
Visual Channels



Credit:
Mike Bostock



Length

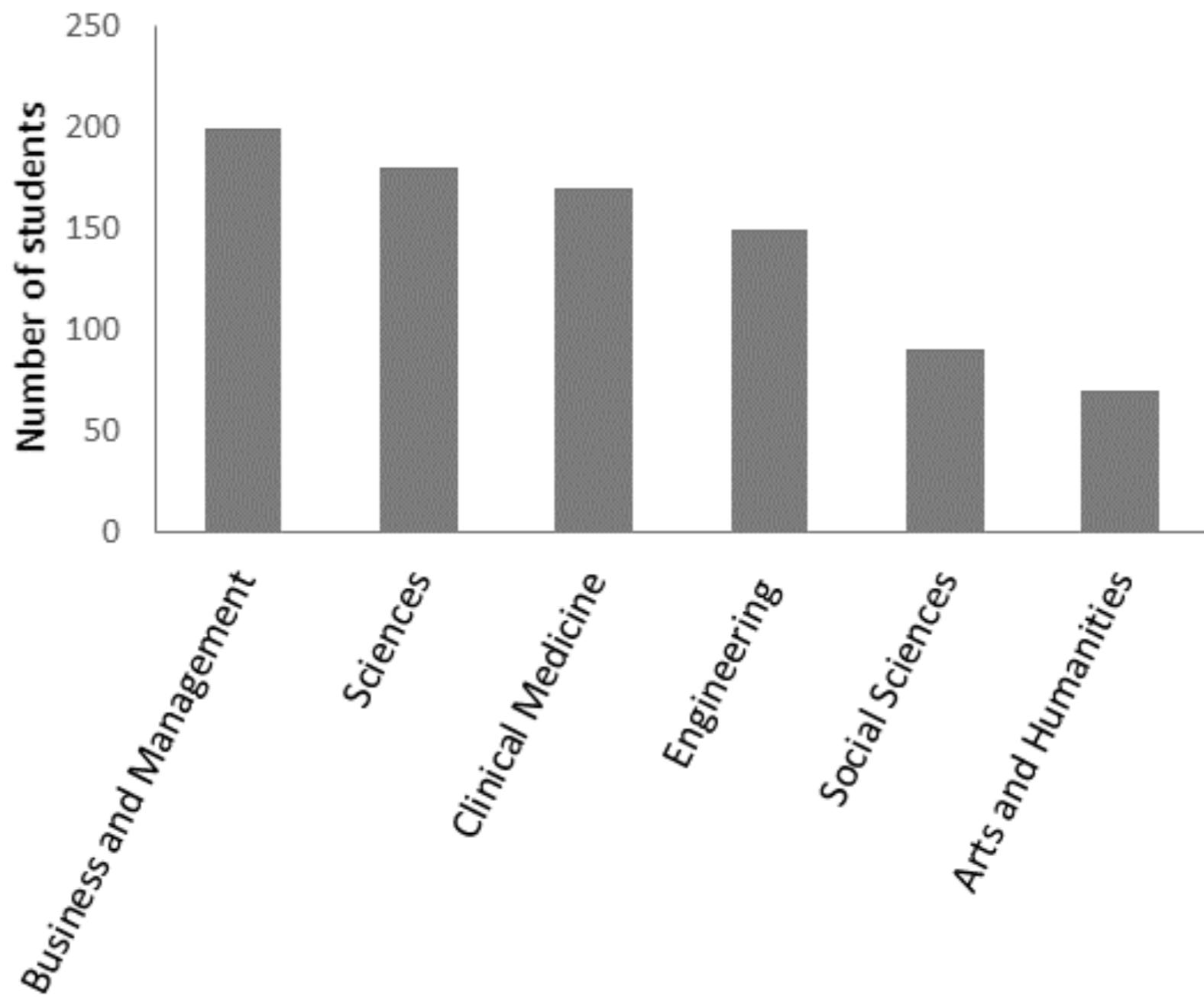


Trump figures as of Dec. 22. Sources: GovTrack; author's calculations

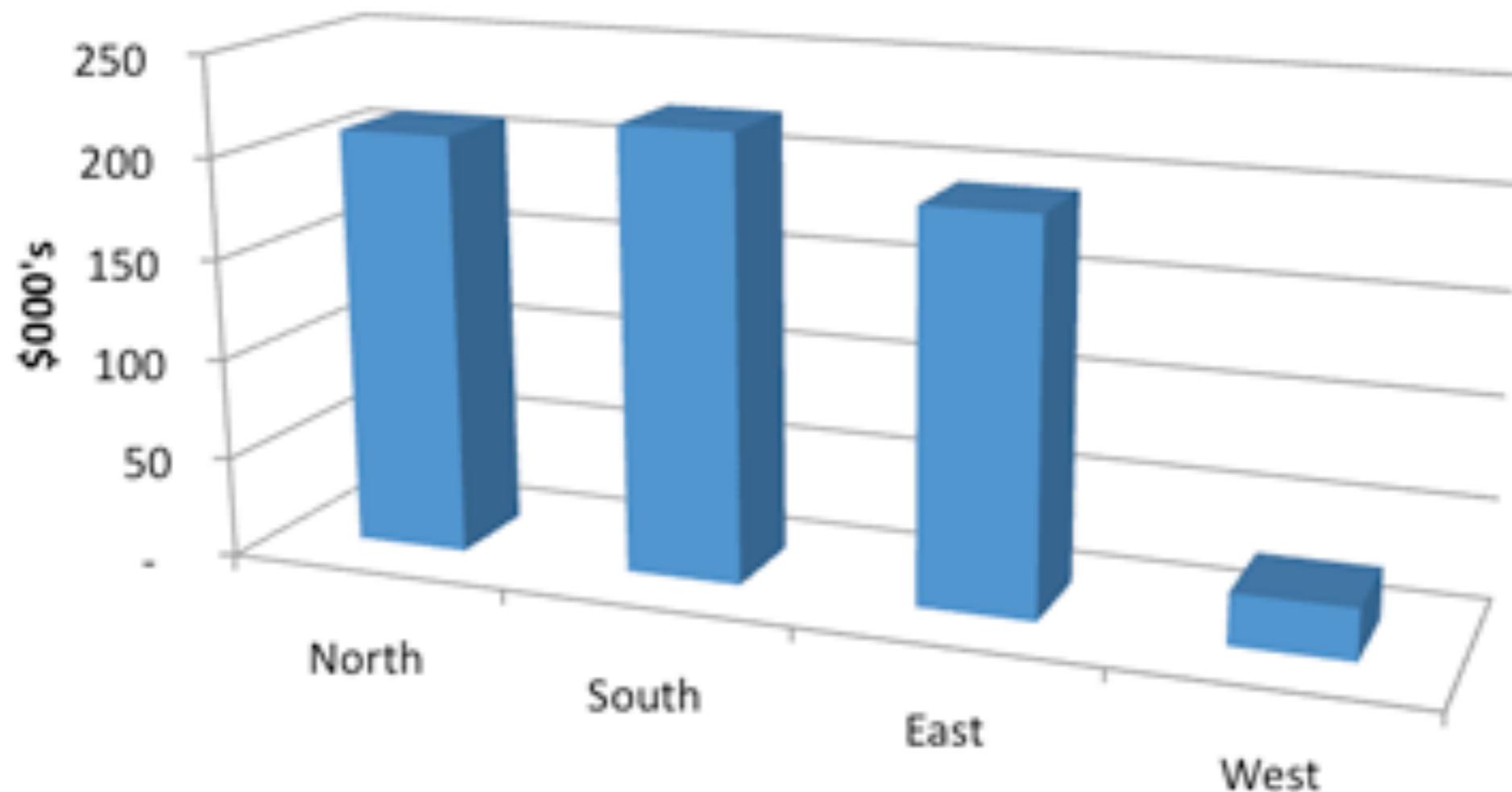
Credit:
NYT



Faculty choice by undergraduate students



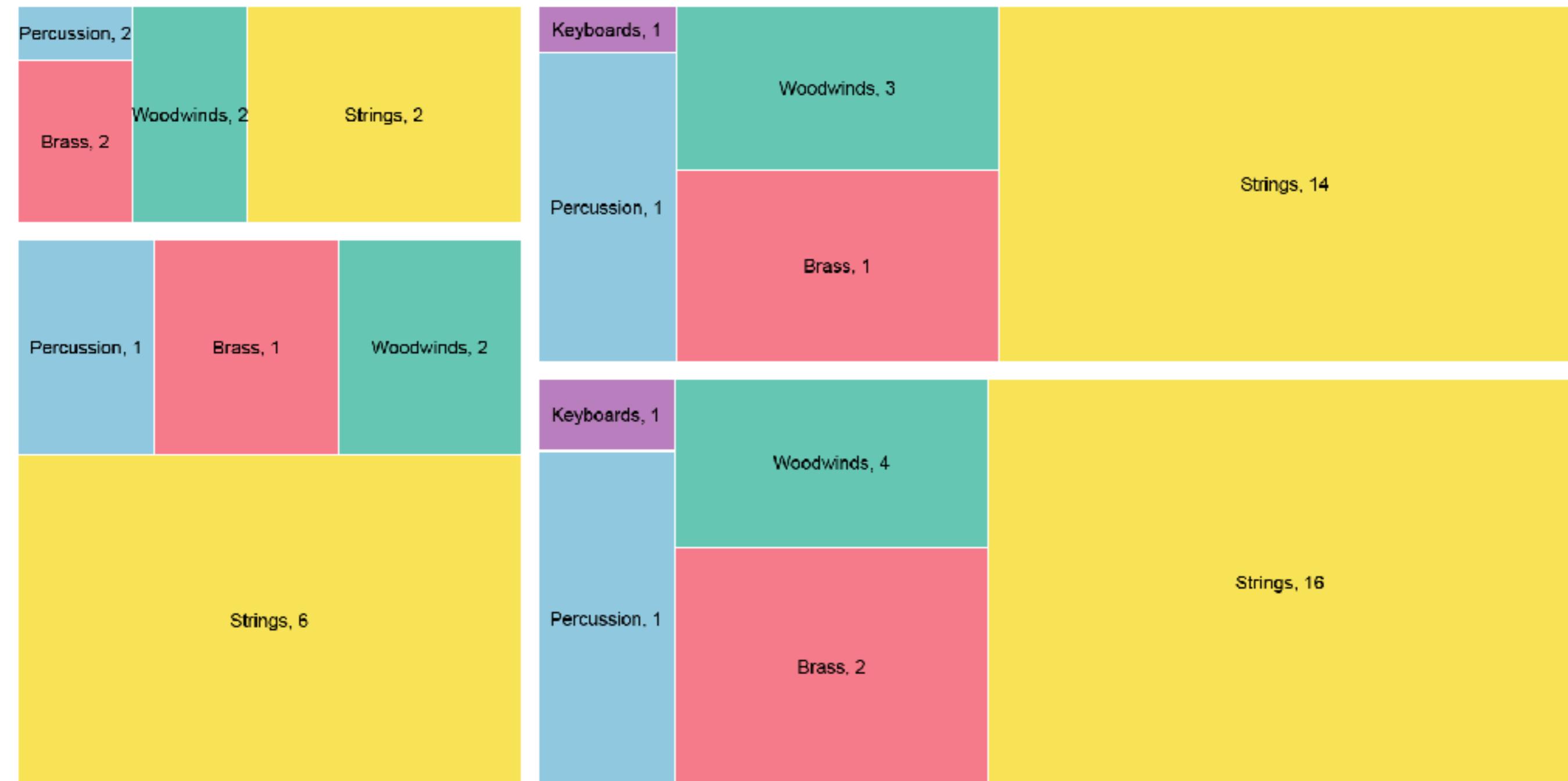
Total Sales by Division (Y/E 2013)



Tilt/Angle



Area

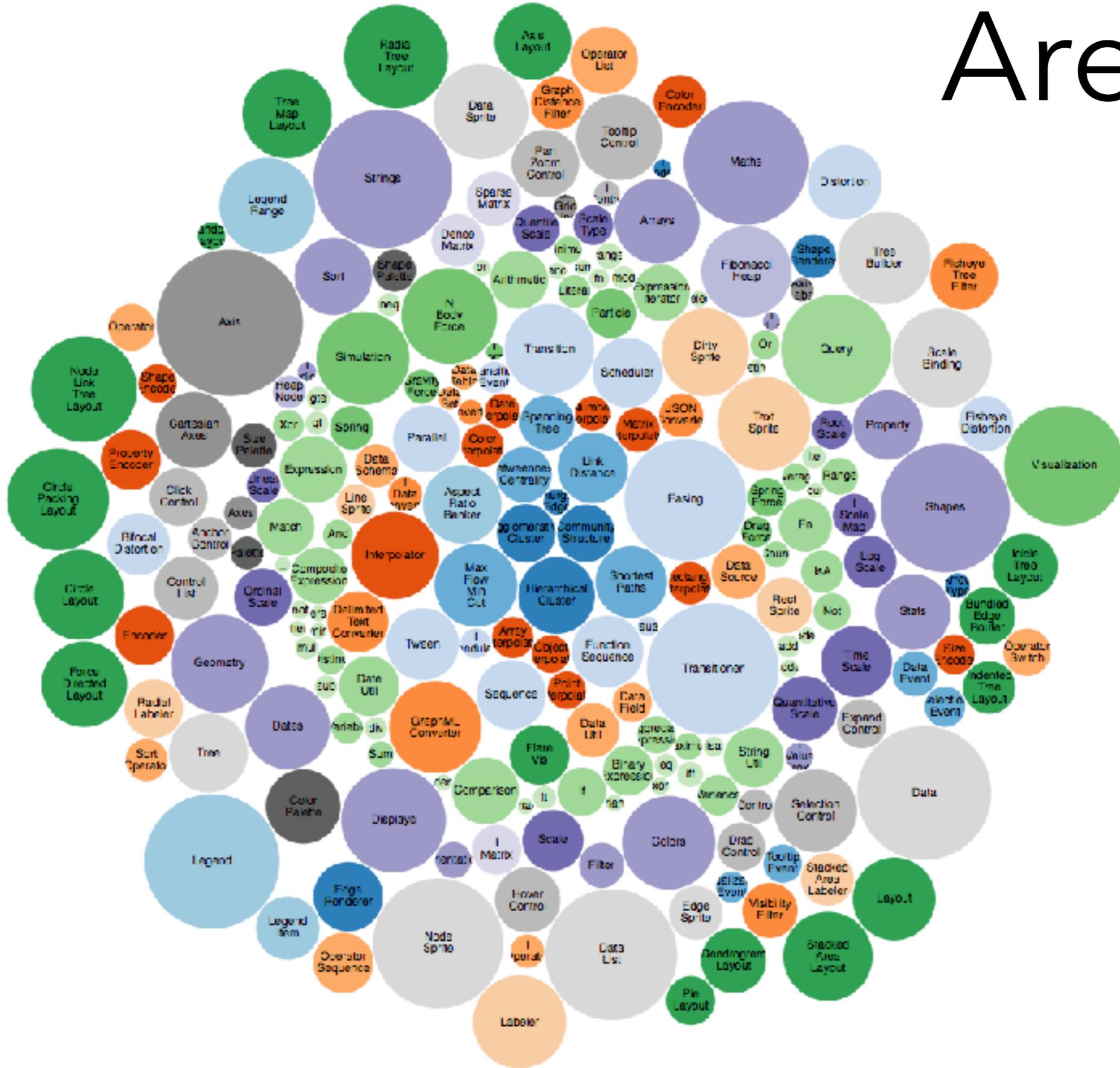


Credit:

RAWGraphs

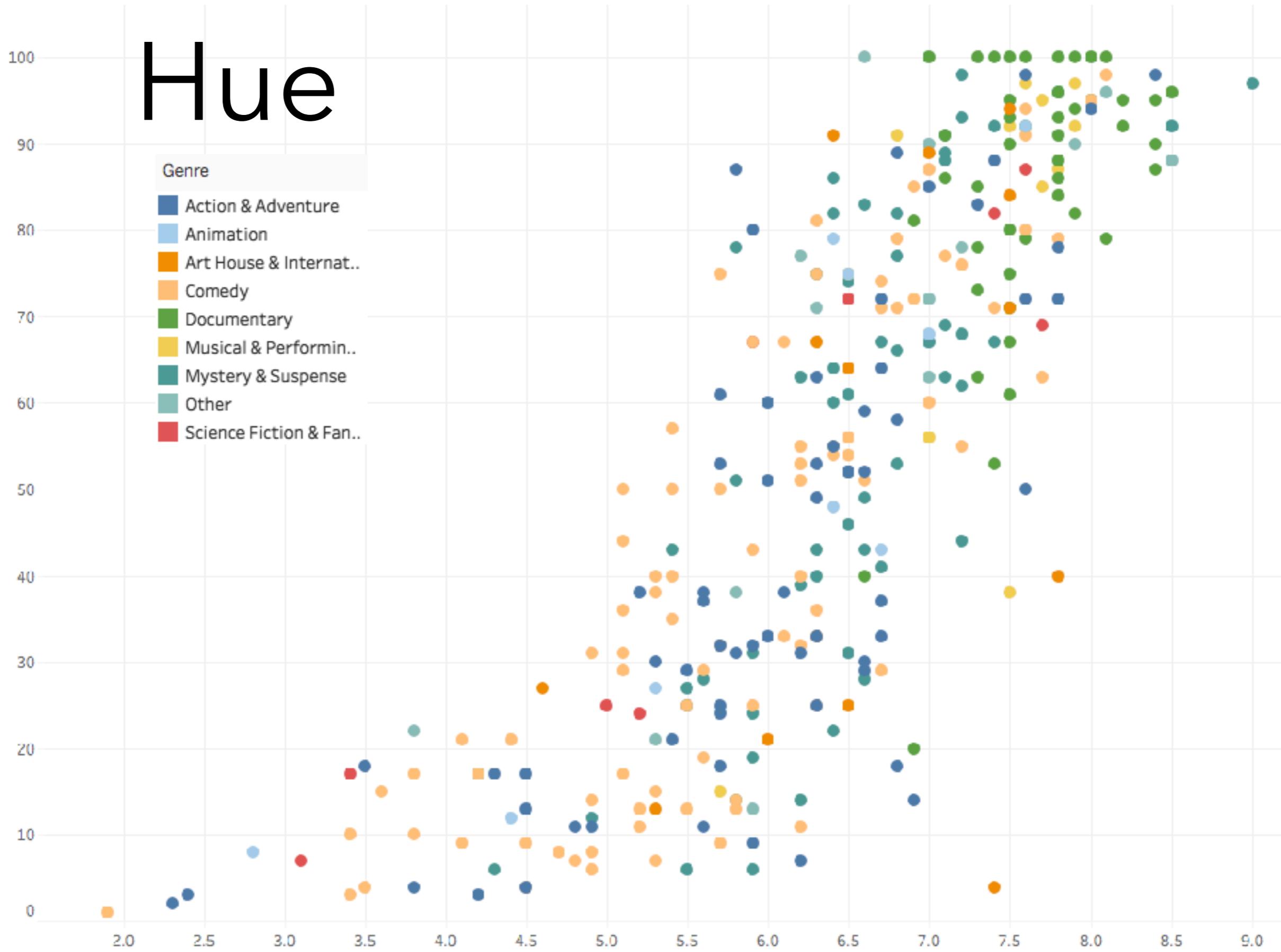


Area



Credit:
Bostock

Hue



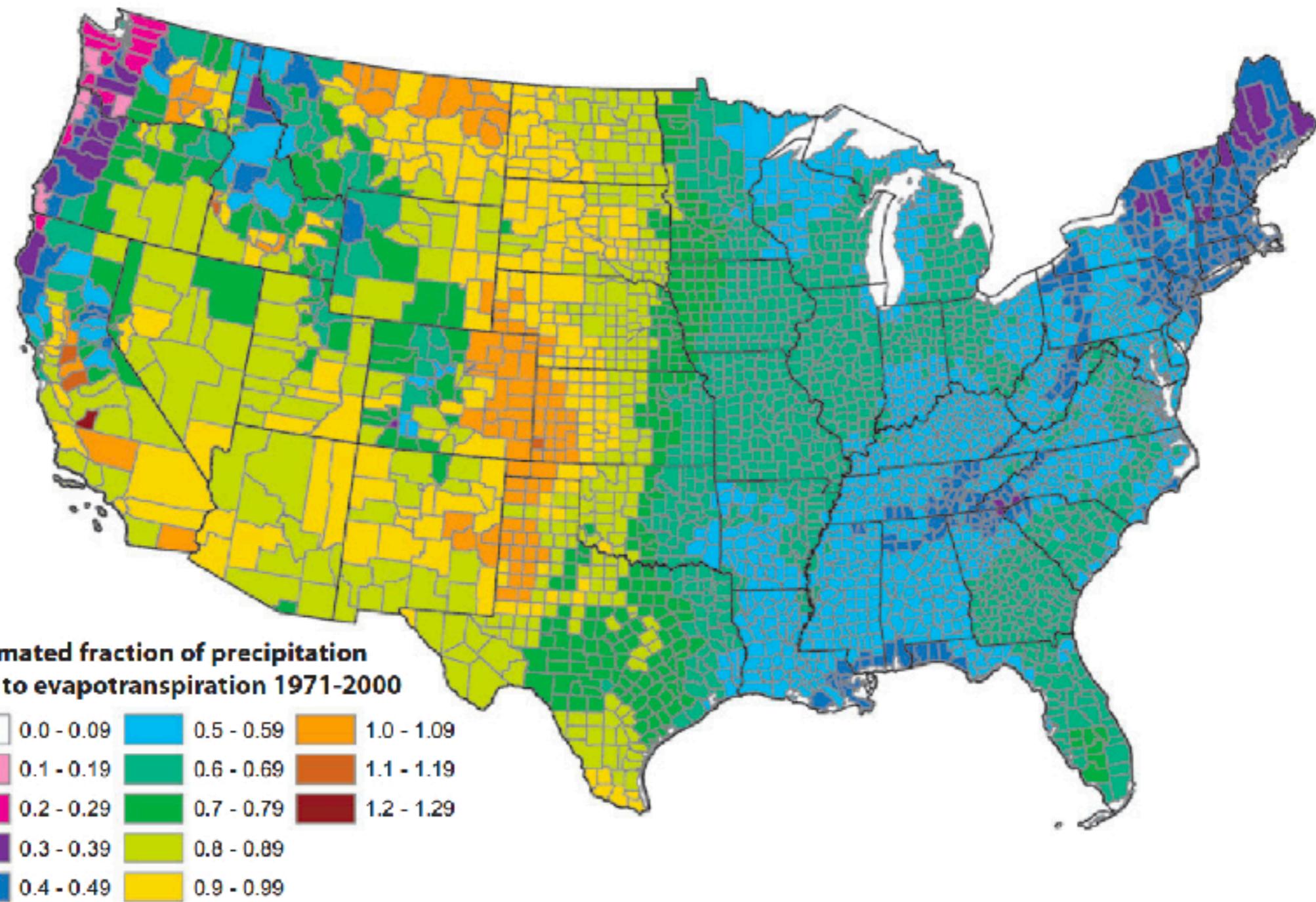
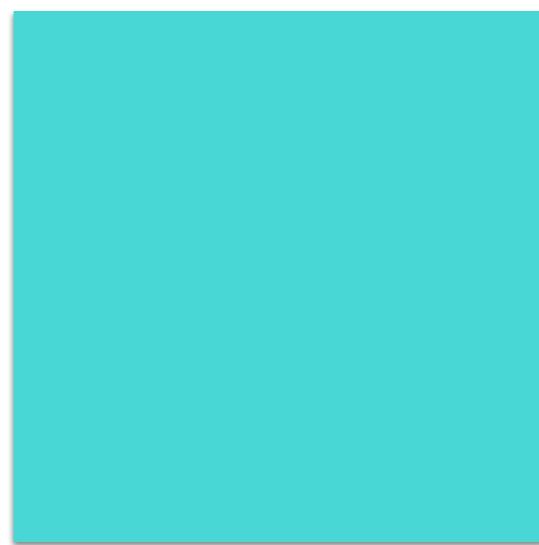


FIGURE 13. Estimated Mean Annual Ratio of Actual Evapotranspiration (ET) to Precipitation (P) for the Conterminous U.S. for the Period 1971-2000. Estimates are based on the regression equation in Table 1 that includes land cover. Calculations of ET/P were made first at the 800-m resolution of the PRISM climate data. The mean values for the counties (shown) were then calculated by averaging the 800-m values within each county. Areas with fractions >1 are agricultural counties that either import surface water or mine deep groundwater.

Hue



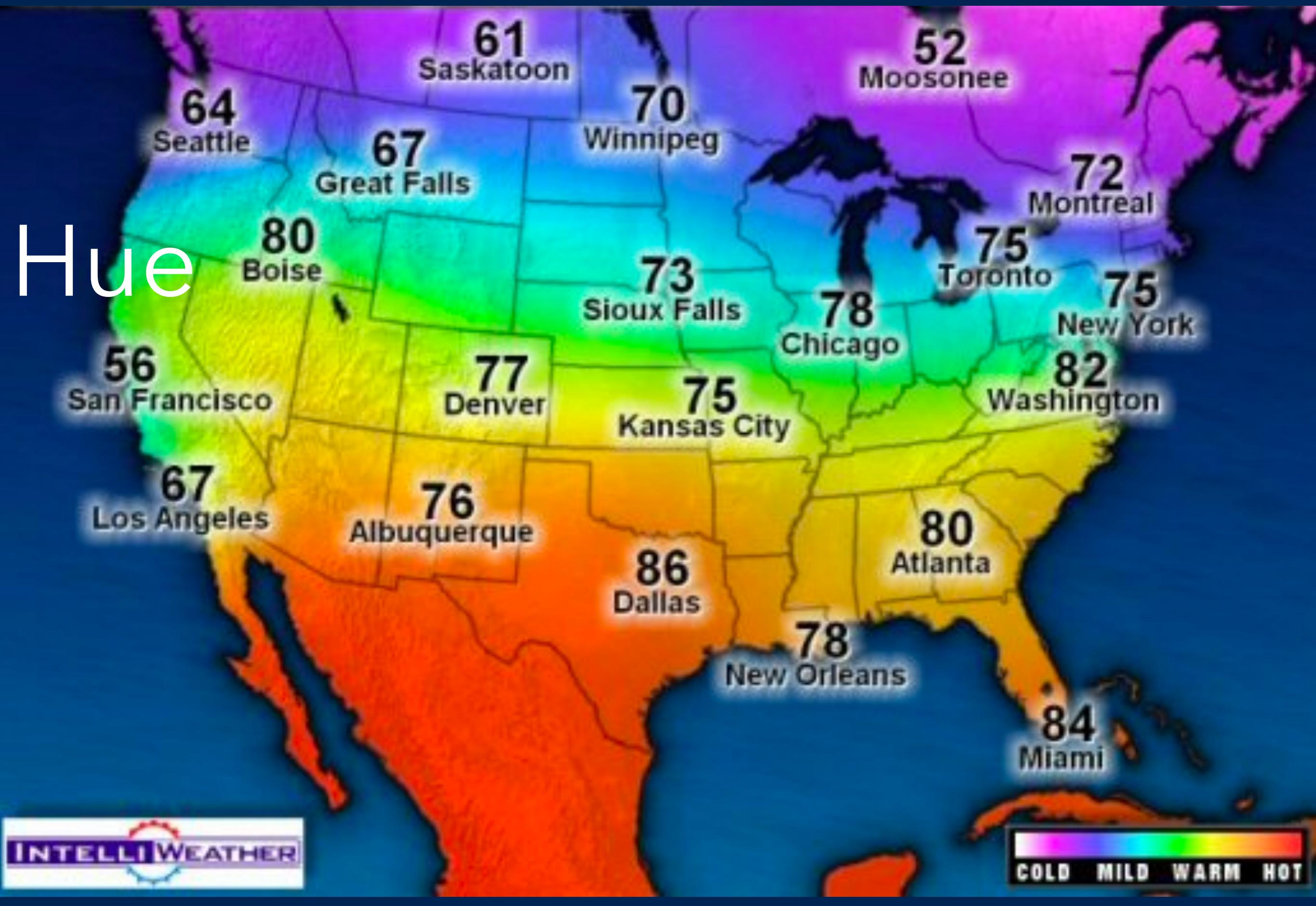
<



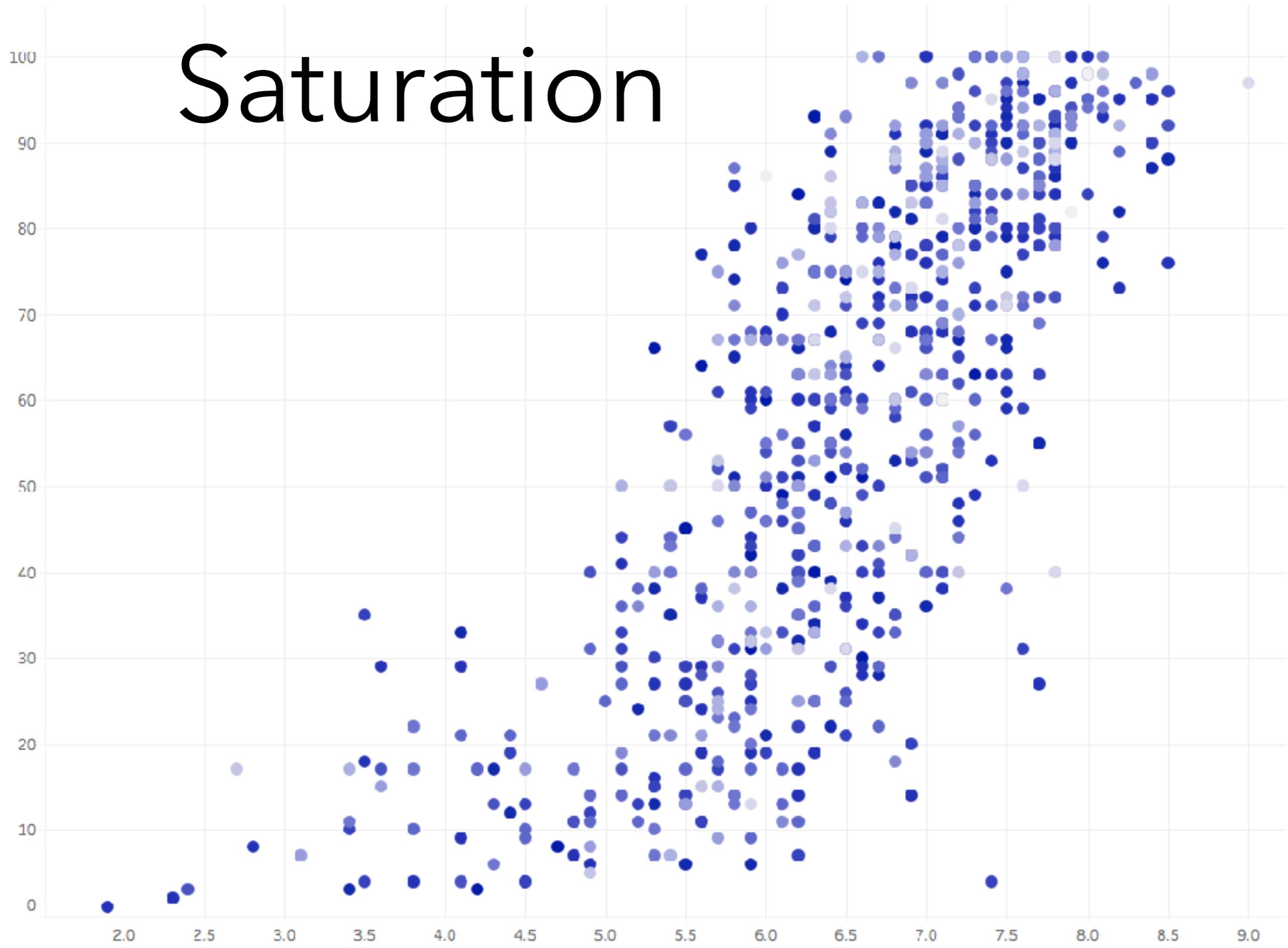
?

Hue only works for nominal values!

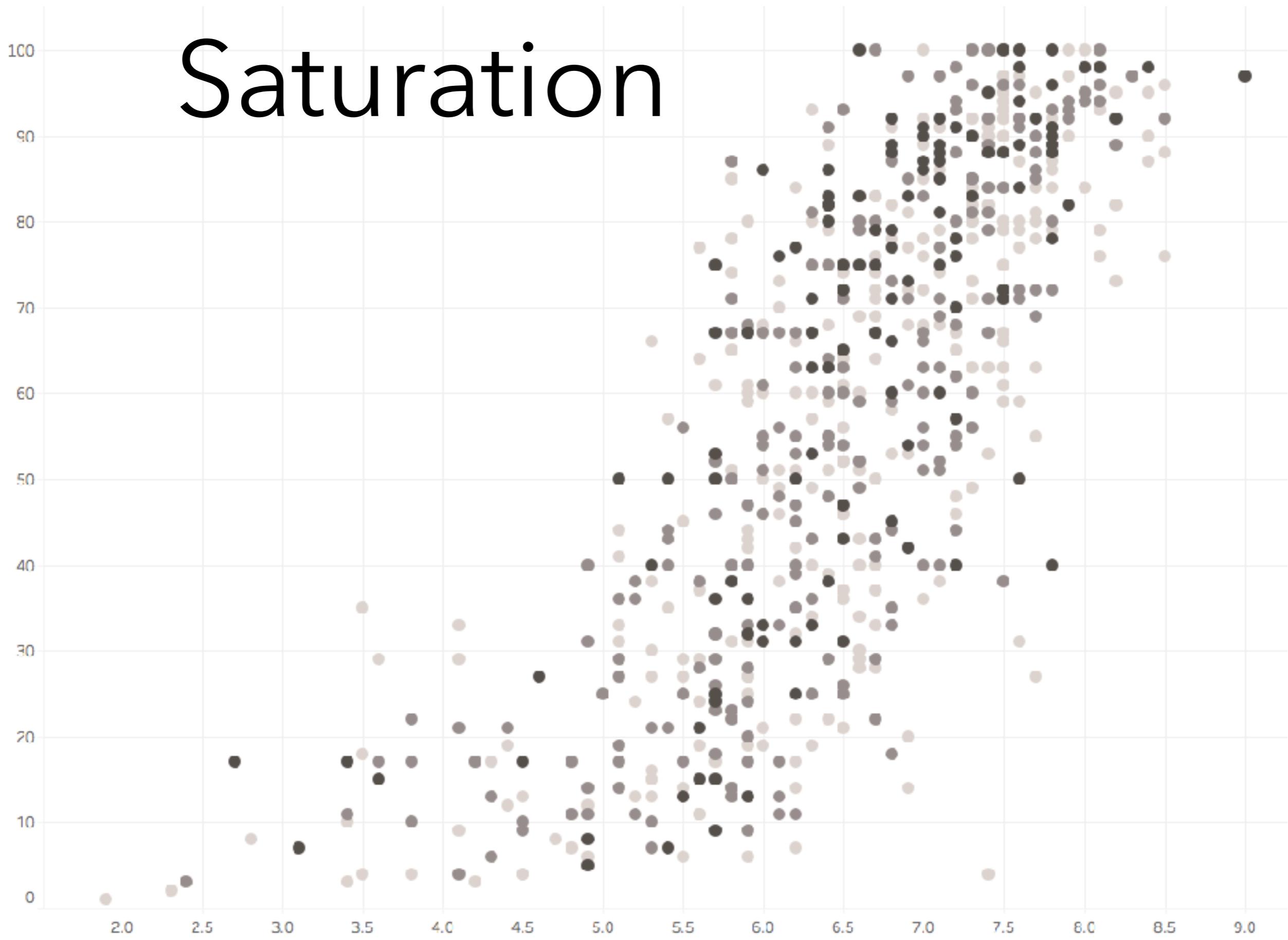
Hue



Saturation

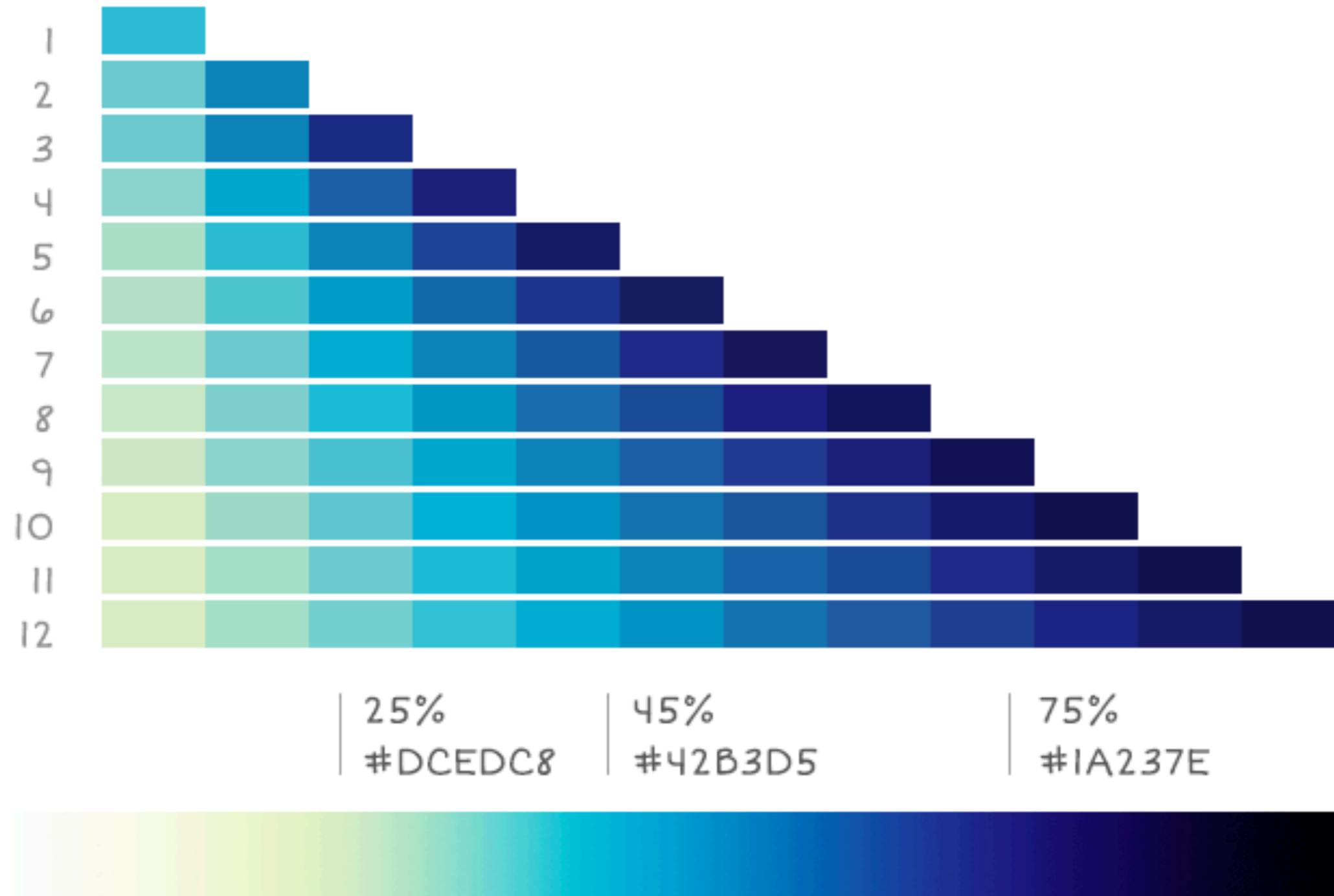


Saturation



Luminance+Saturation

OF COLORS

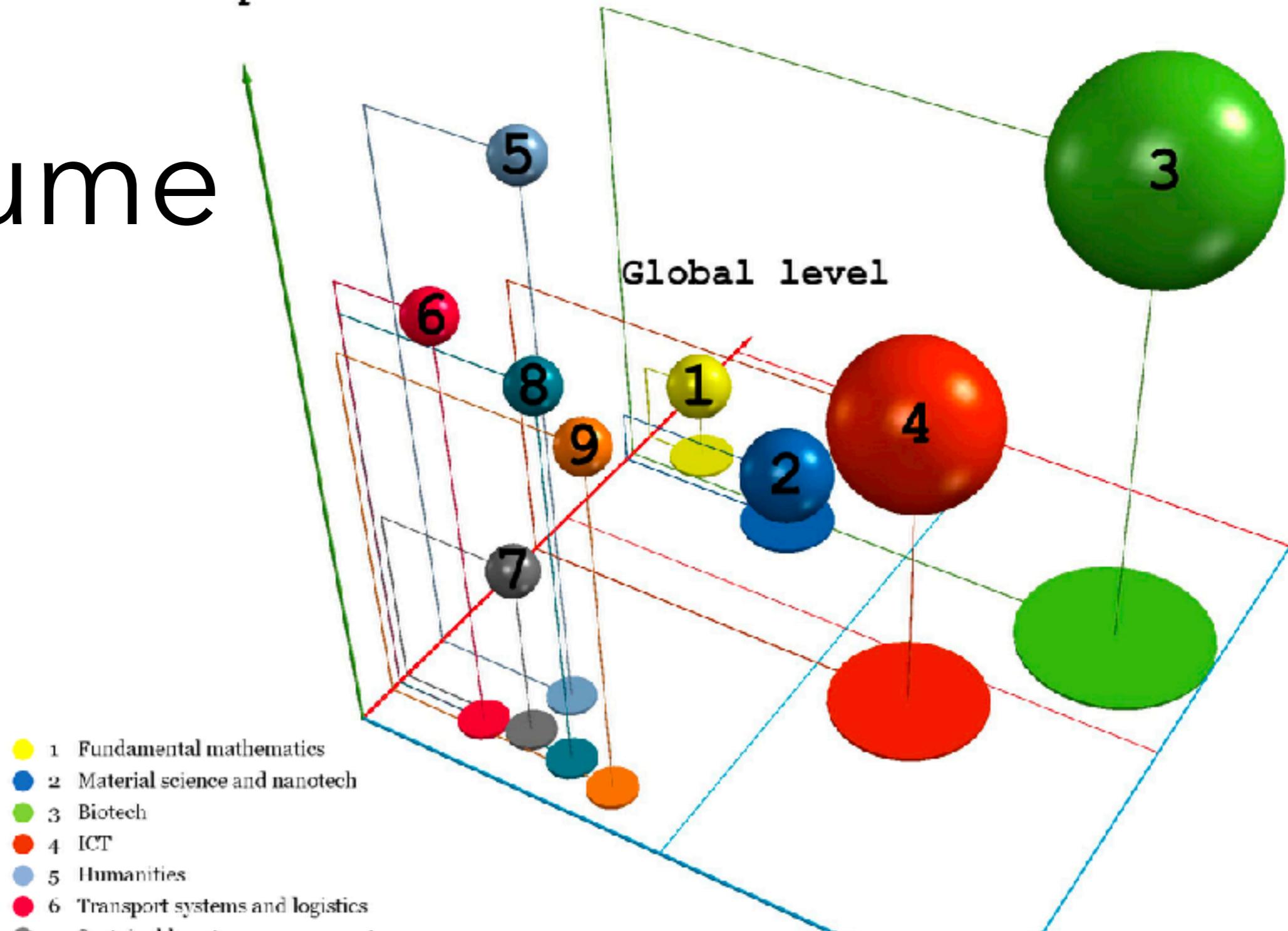


Credit:
Graphiq



Social impact

Volume



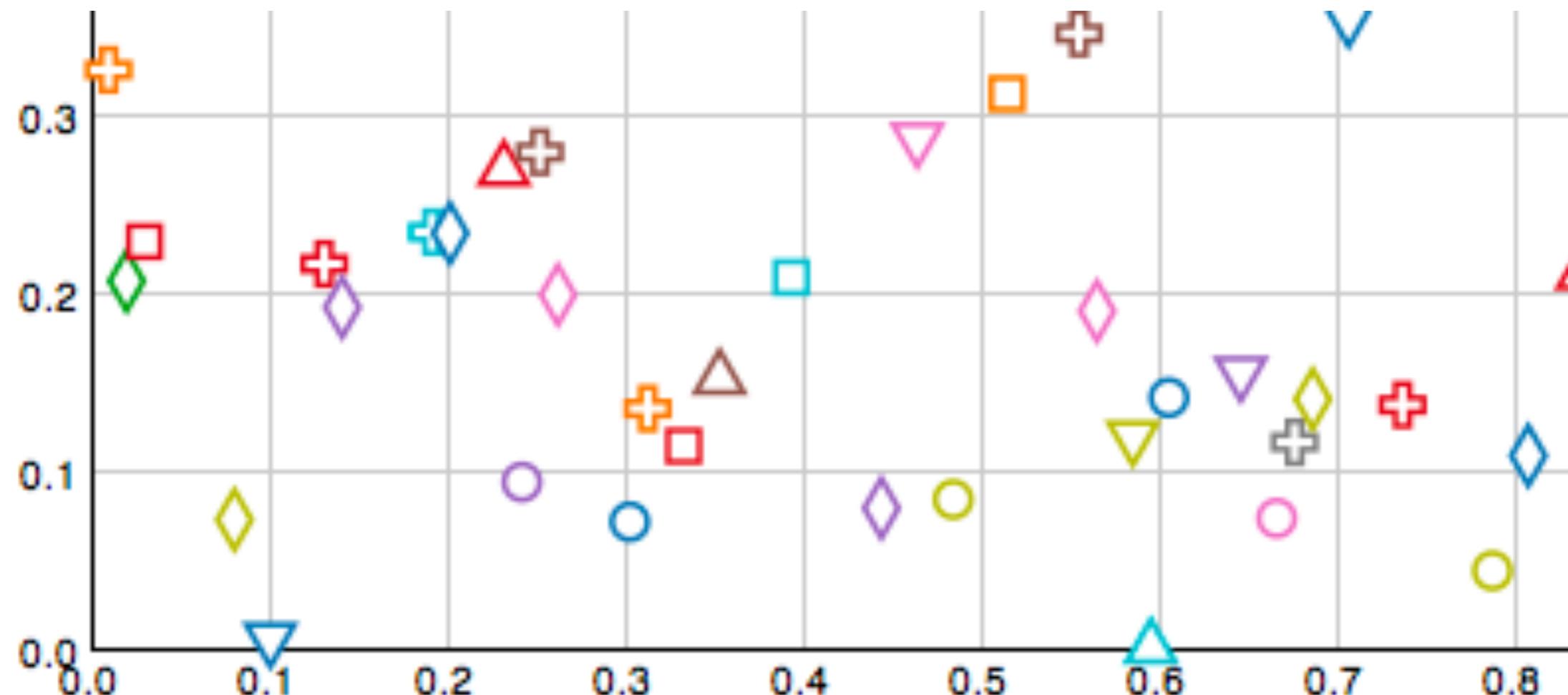
Credit:

Wikimedia

Regional priorities

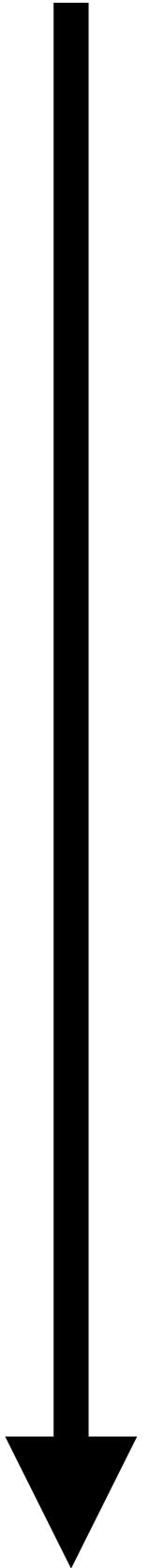


Shape / Symbols



Credit: d3

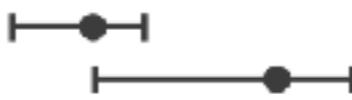




Position on common scale



Position on unaligned scale



Length (1D size)



Tilt/angle



Area (2D size)



Depth (3D position)



Color luminance



Same]

Color saturation



Same]

Curvature



Volume (3D size)



Spatial region



Color hue



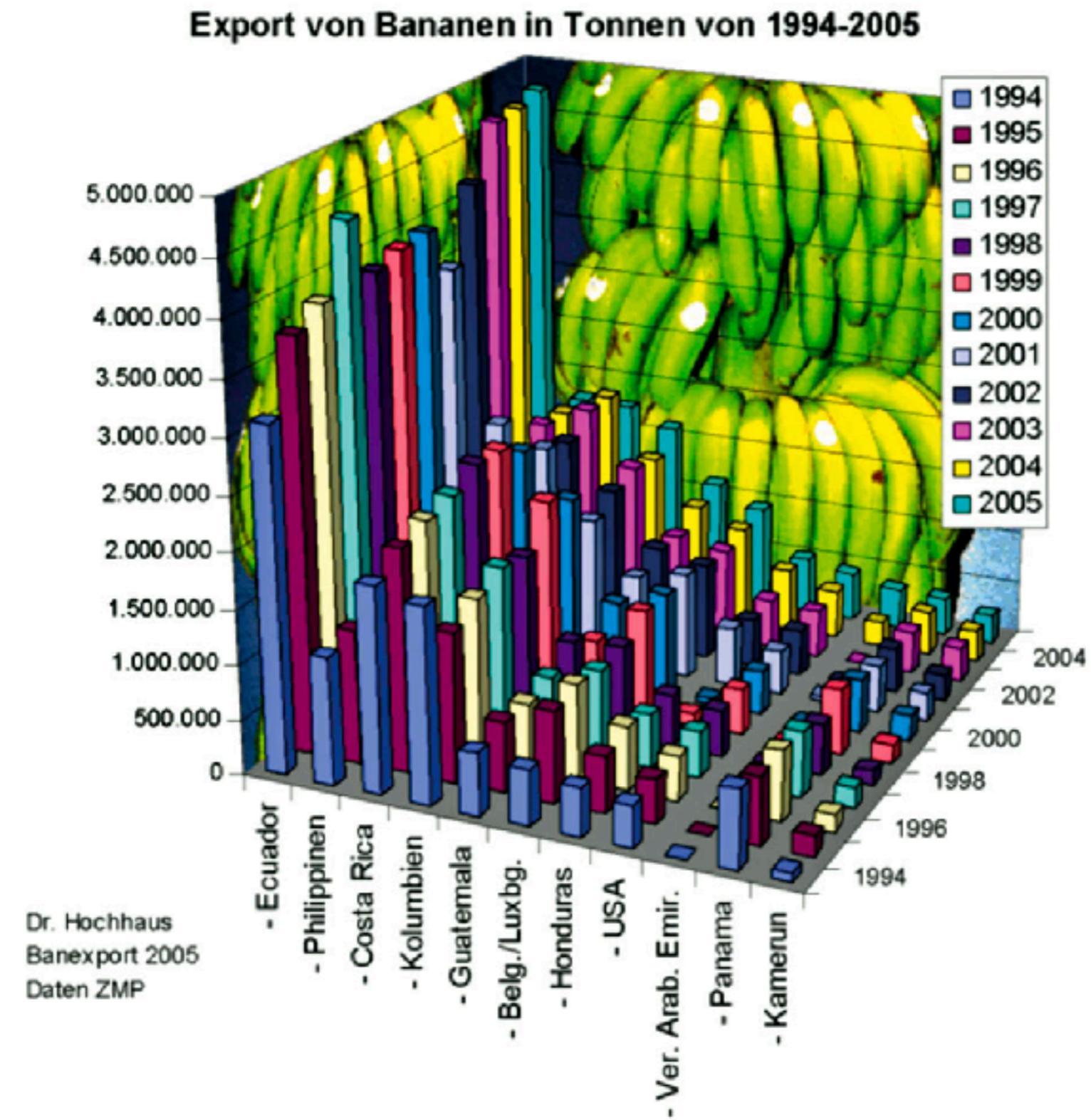
Motion



Shape



How do we effectively use marks/channels?



How do we effectively use marks/ channels?

Export von Bananen in Tonnen von 1994-2005



Effective Marks and Channels

Discernibility

Recognizability

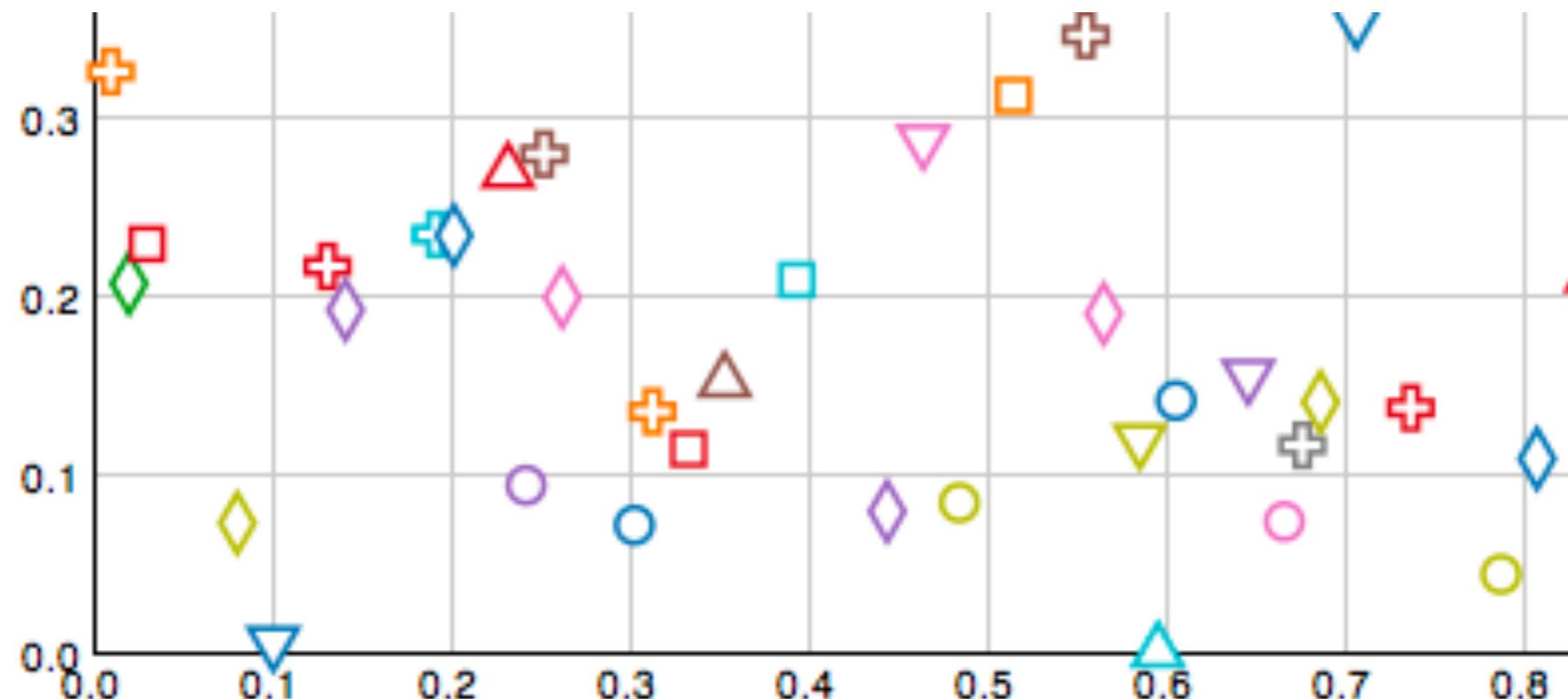
Proportionality

Redundancy / conflict



Discernibility

Marks should be recognizable and distinct

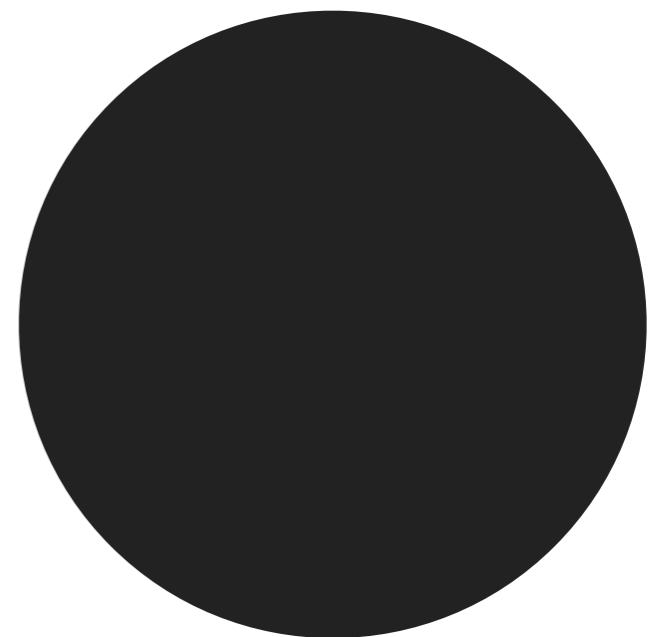
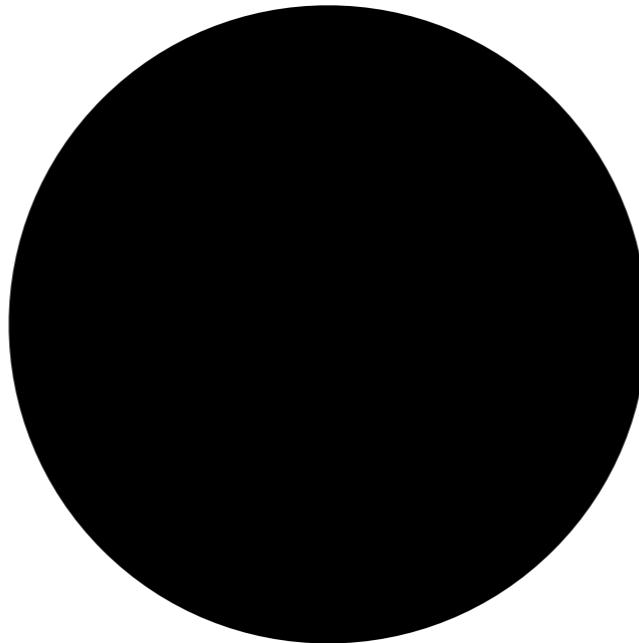
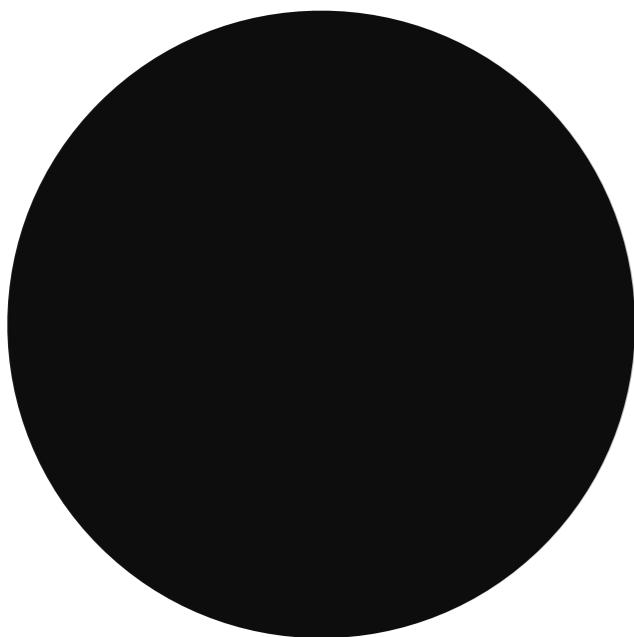


Credit: d3



Recognizability

Differences in channels should be noticeable

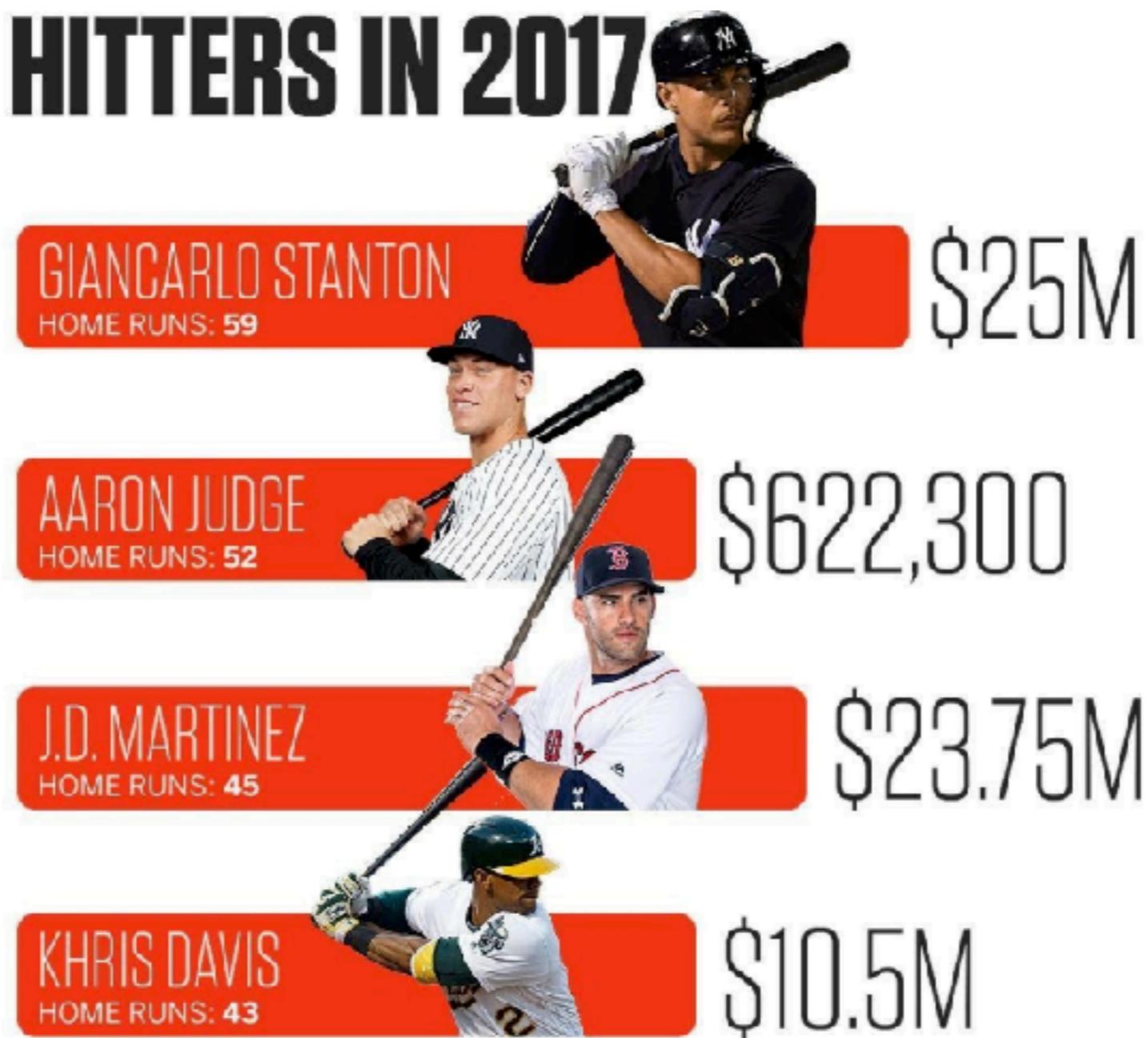


(Which of these is different from the others?)

(more on this later!)

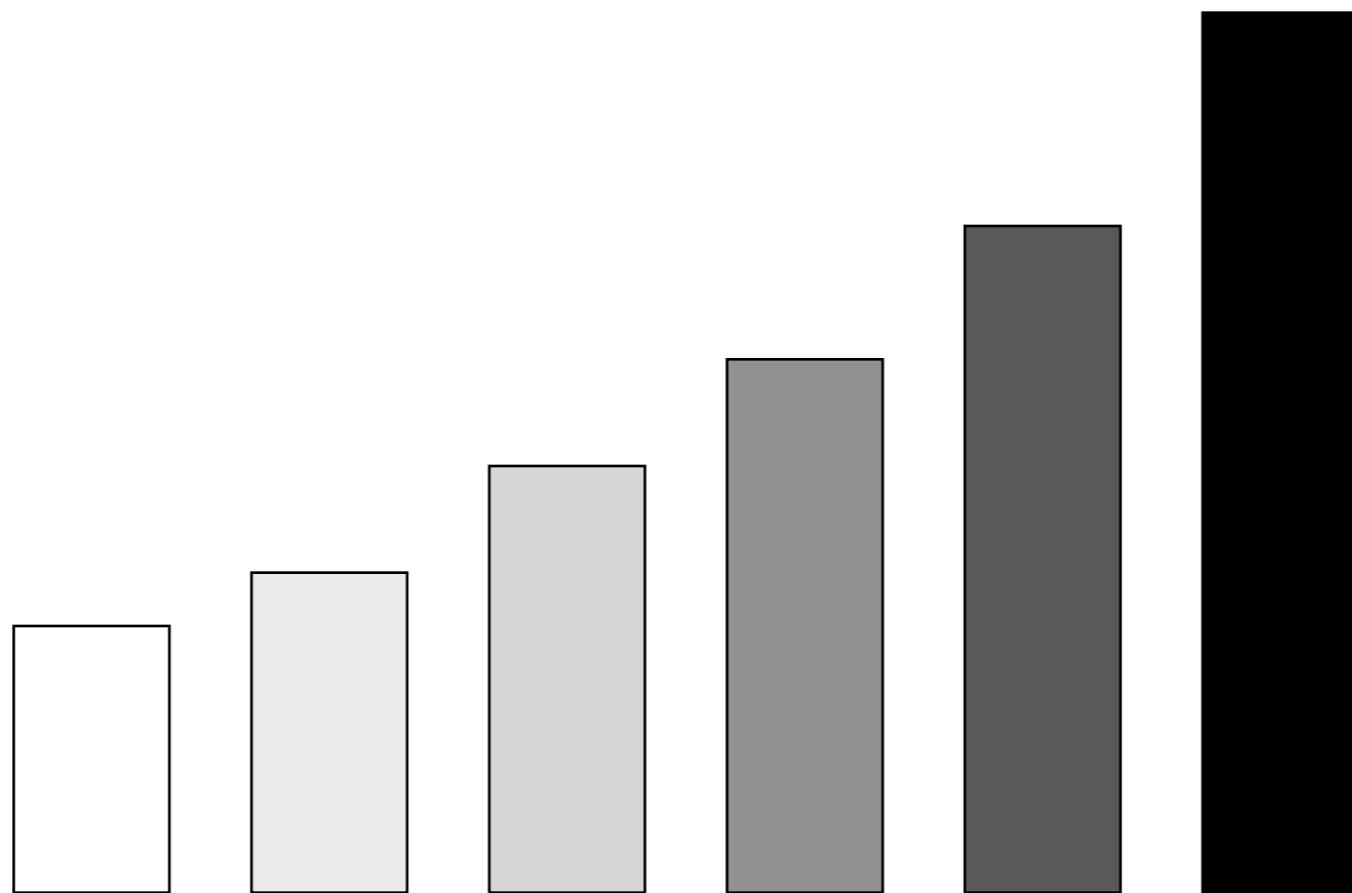
Proportionality

Channels should map accurately to data



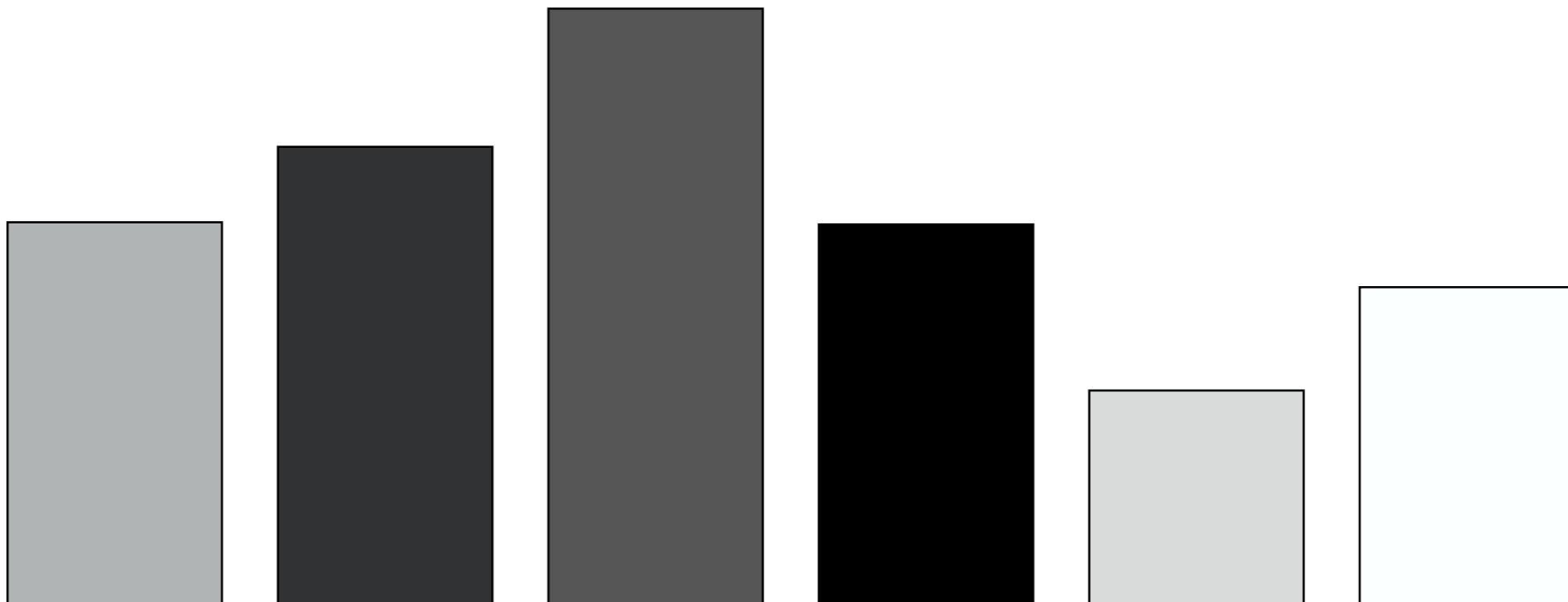
Redundancy / conflict

Channels influence each other



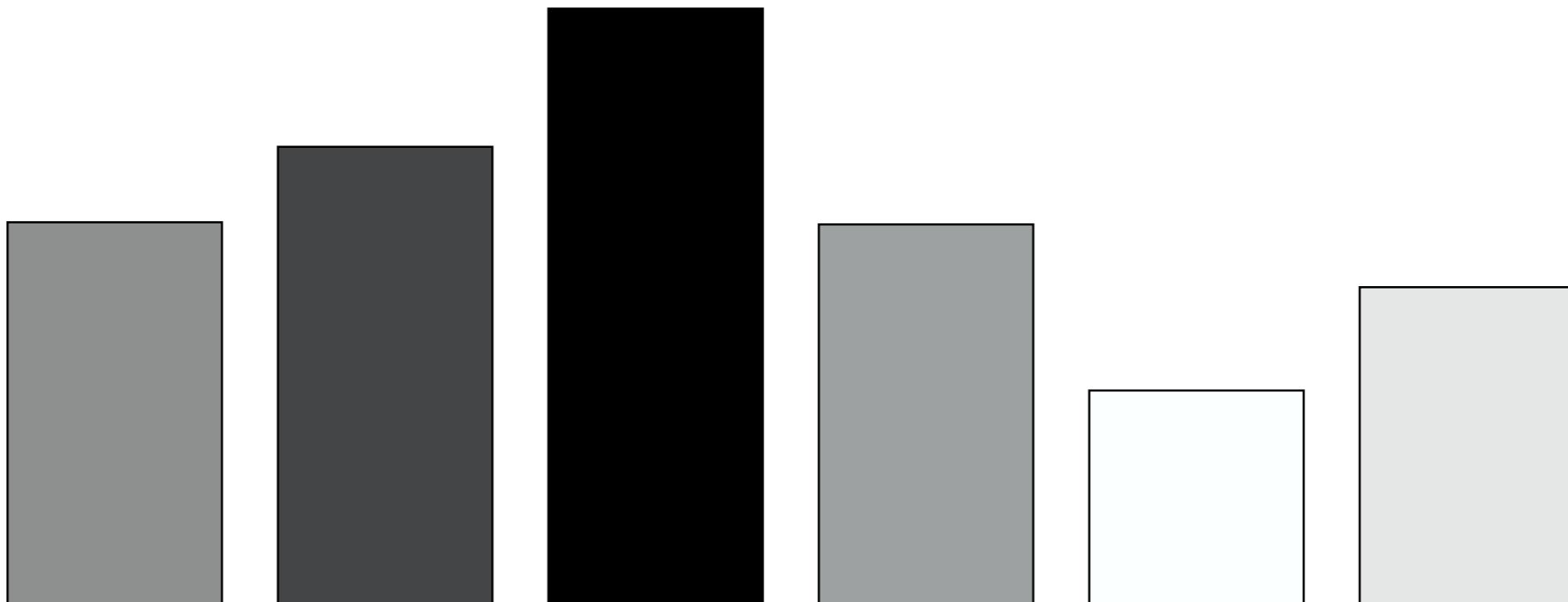
Redundancy / conflict

Channels influence each other



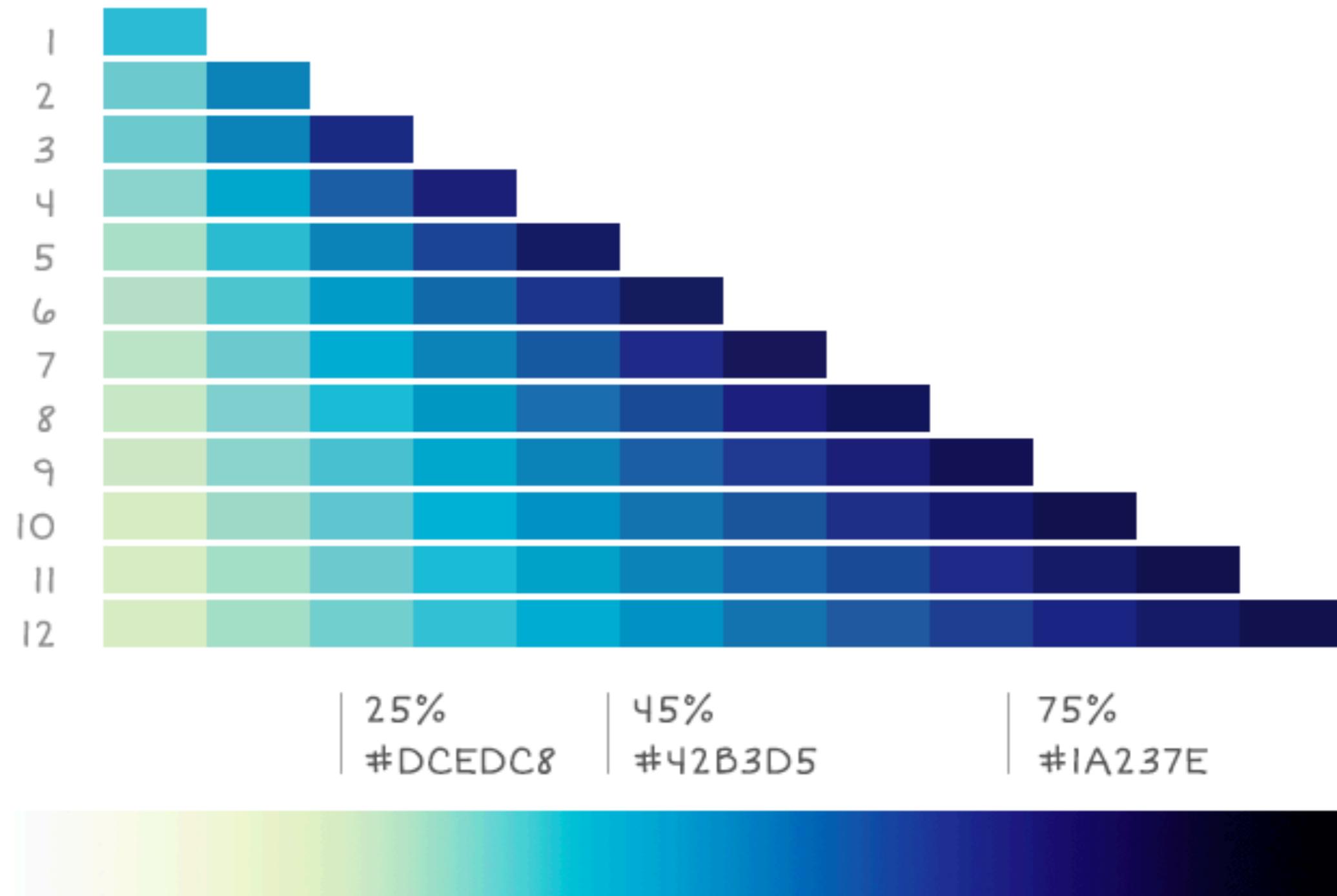
Redundancy / conflict

Maximize redundancy, minimize conflict



Luminance+Saturation

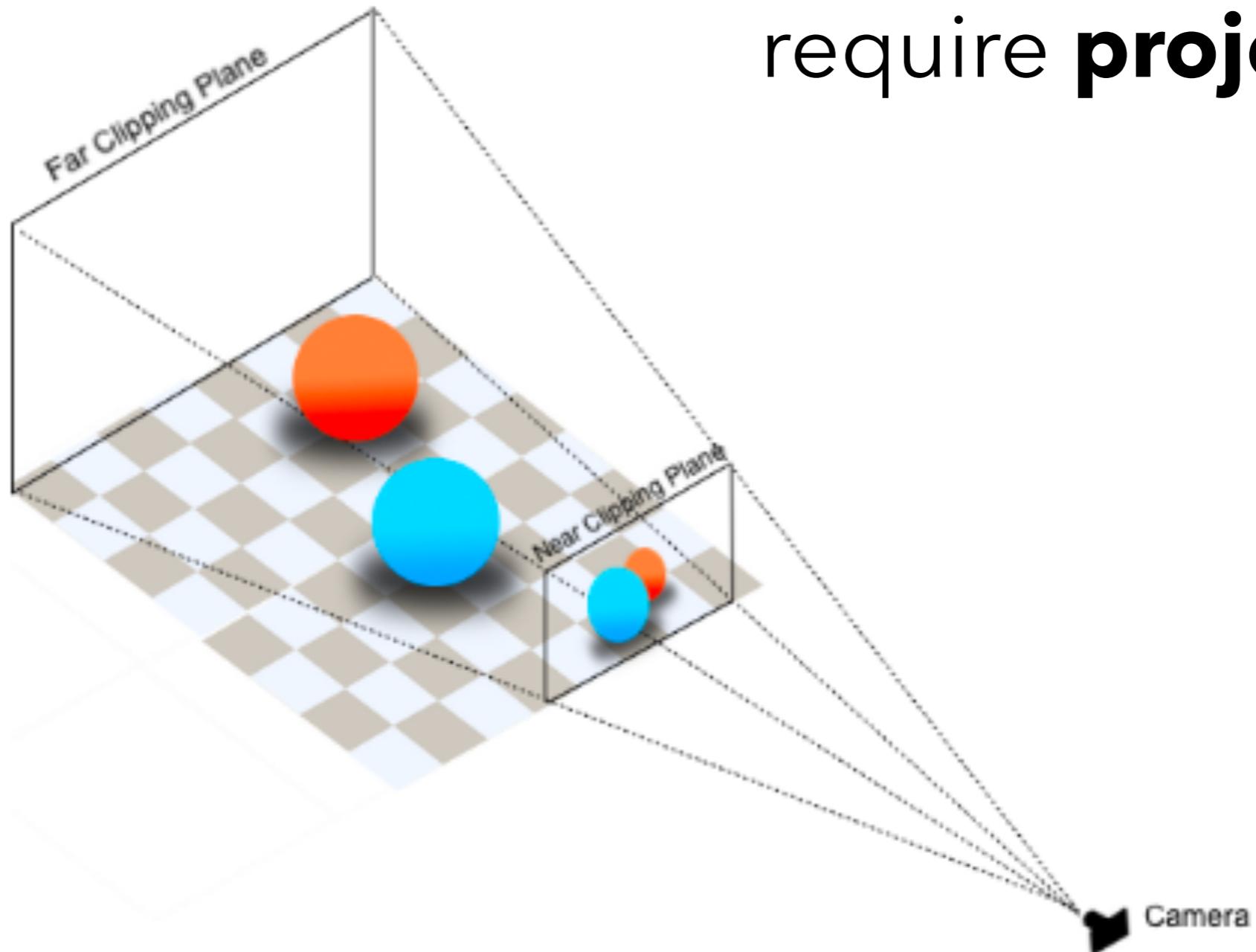
OF COLORS

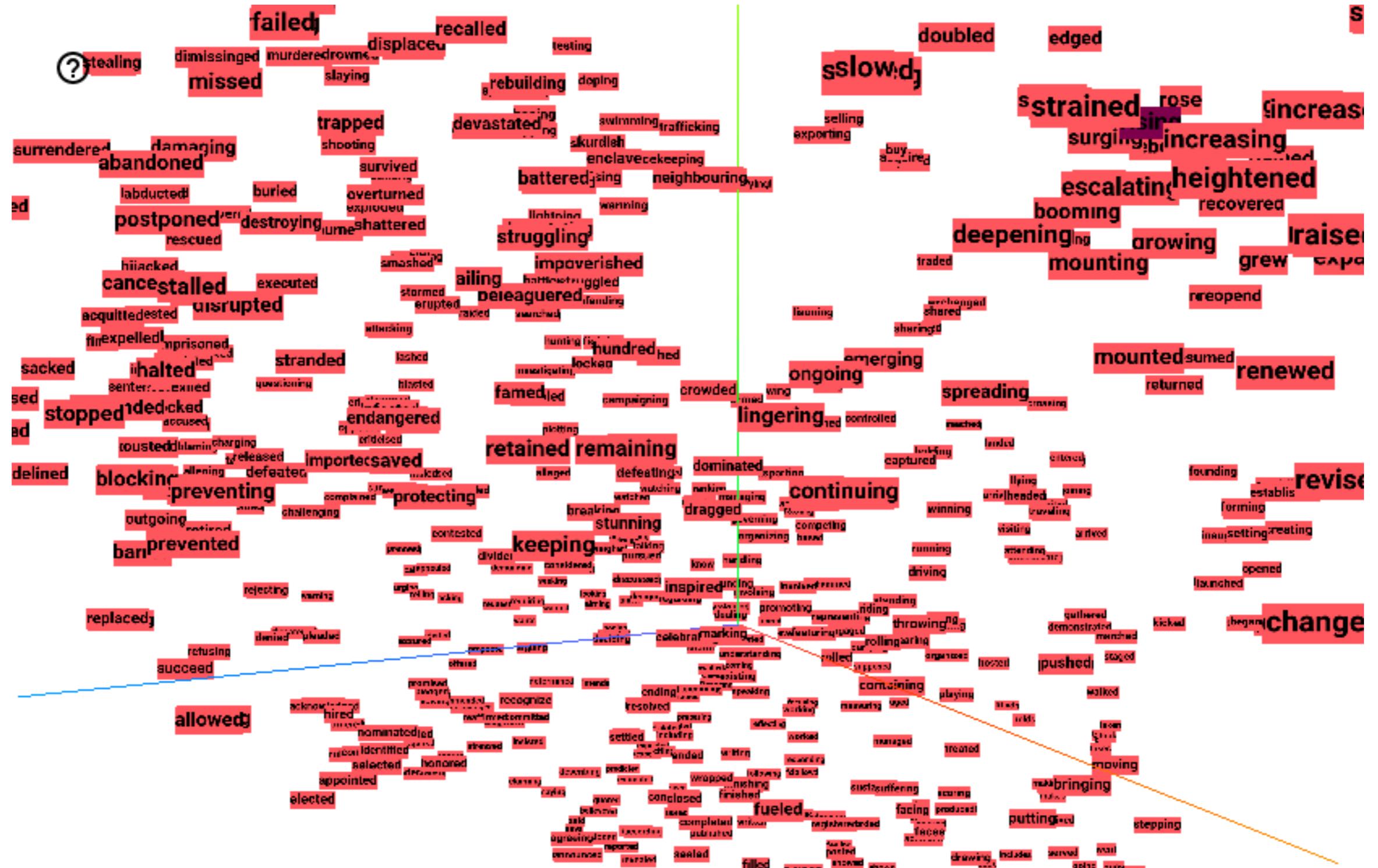


Credit:
Graphiq



3-D visualizations require **projection**

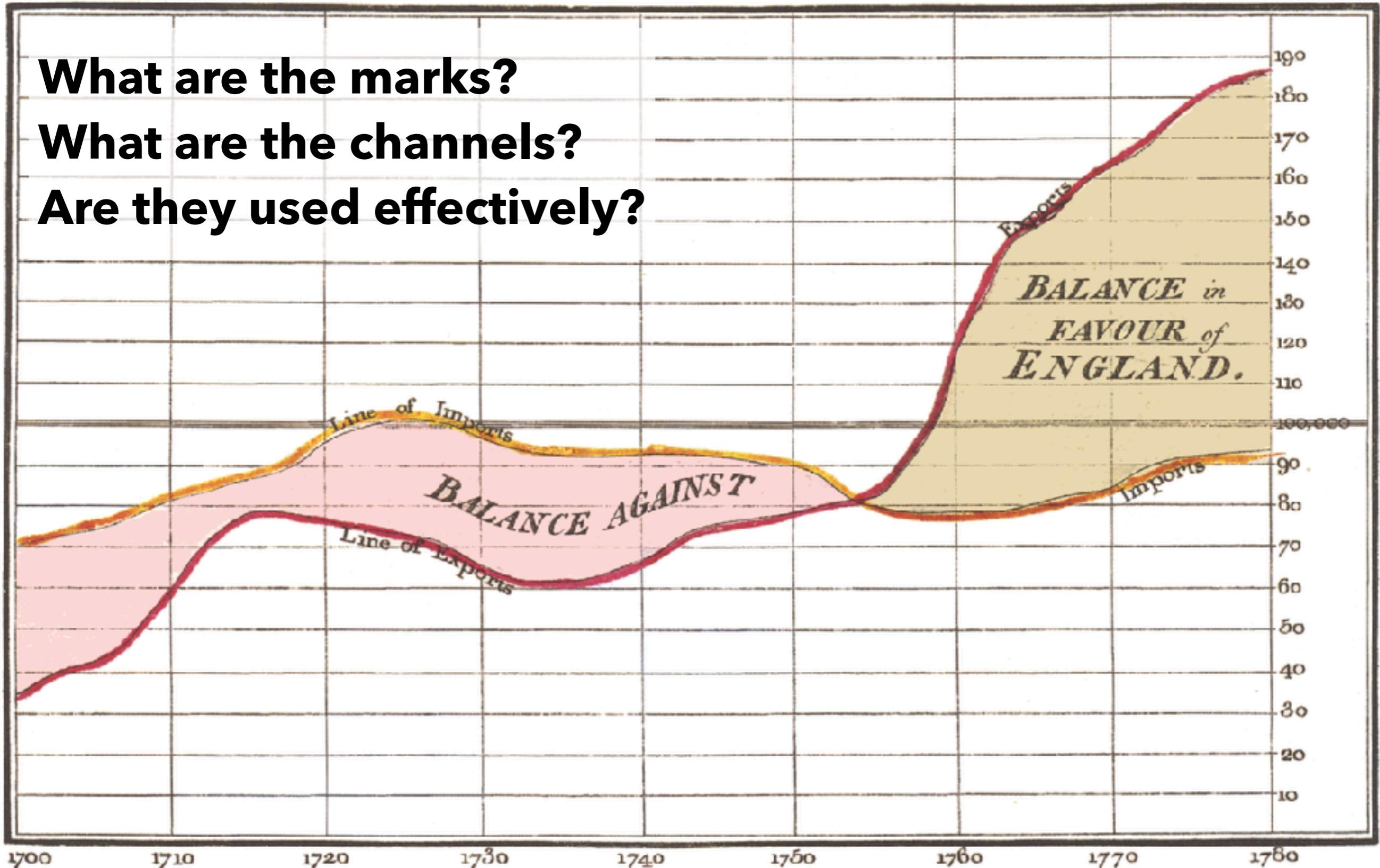




Which can turn reliable channels into very
unreliable ones!

(more later in term)

**What are the marks?
What are the channels?
Are they used effectively?**



The Bottom line is divided into Years, the Right hand line into £10,000 each.

Published as the Act directs, 1st May 1786, by W^m Playfair

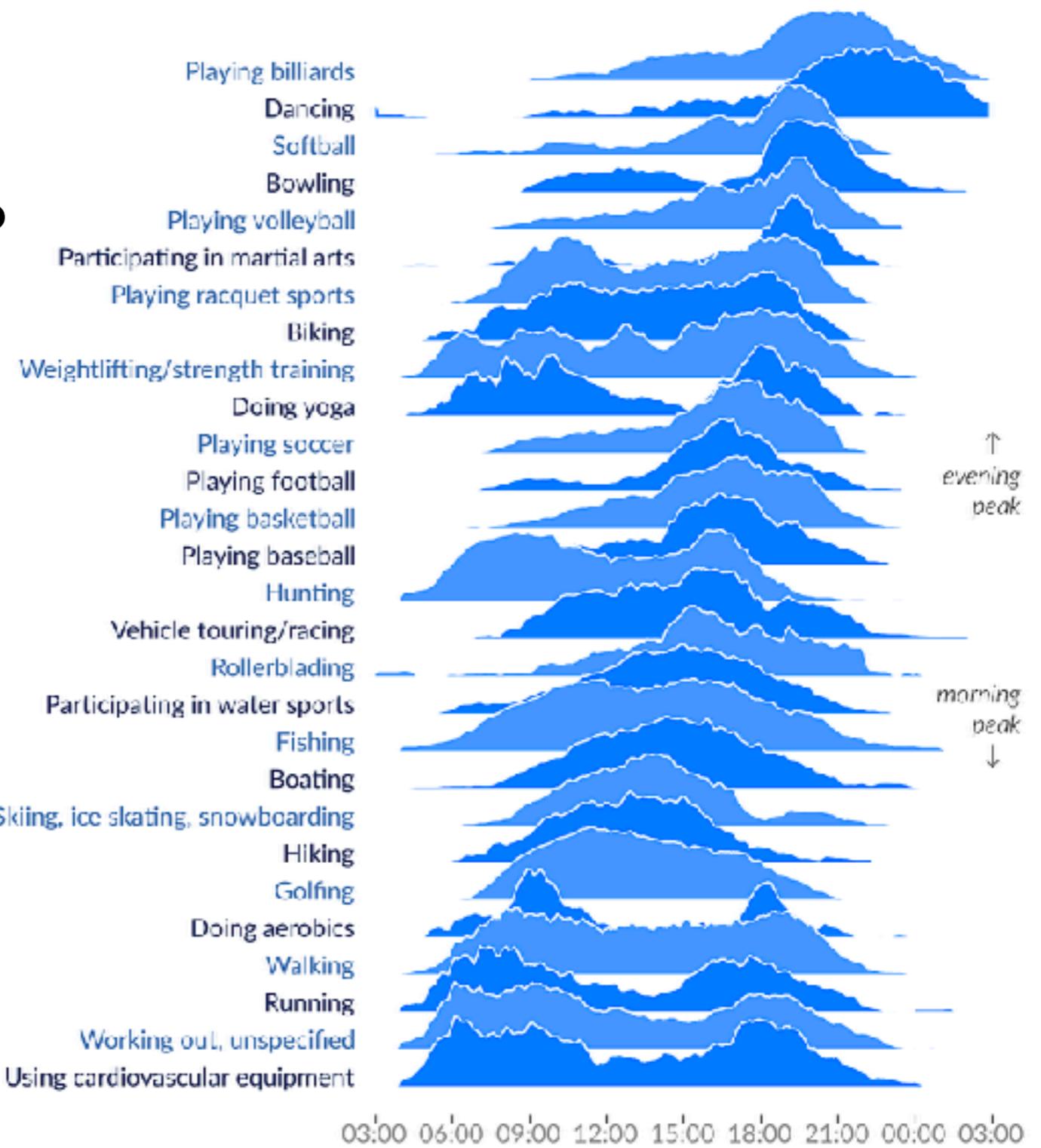
No. 62, Strand, London.



Peak time of day for sports and leisure

Number of participants throughout the day compared to peak popularity.
Note the morning-and-evening everyday workouts, the midday hobbies,
and the evenings/late nights out.

What are the marks? What are the channels? Are they used effectively?

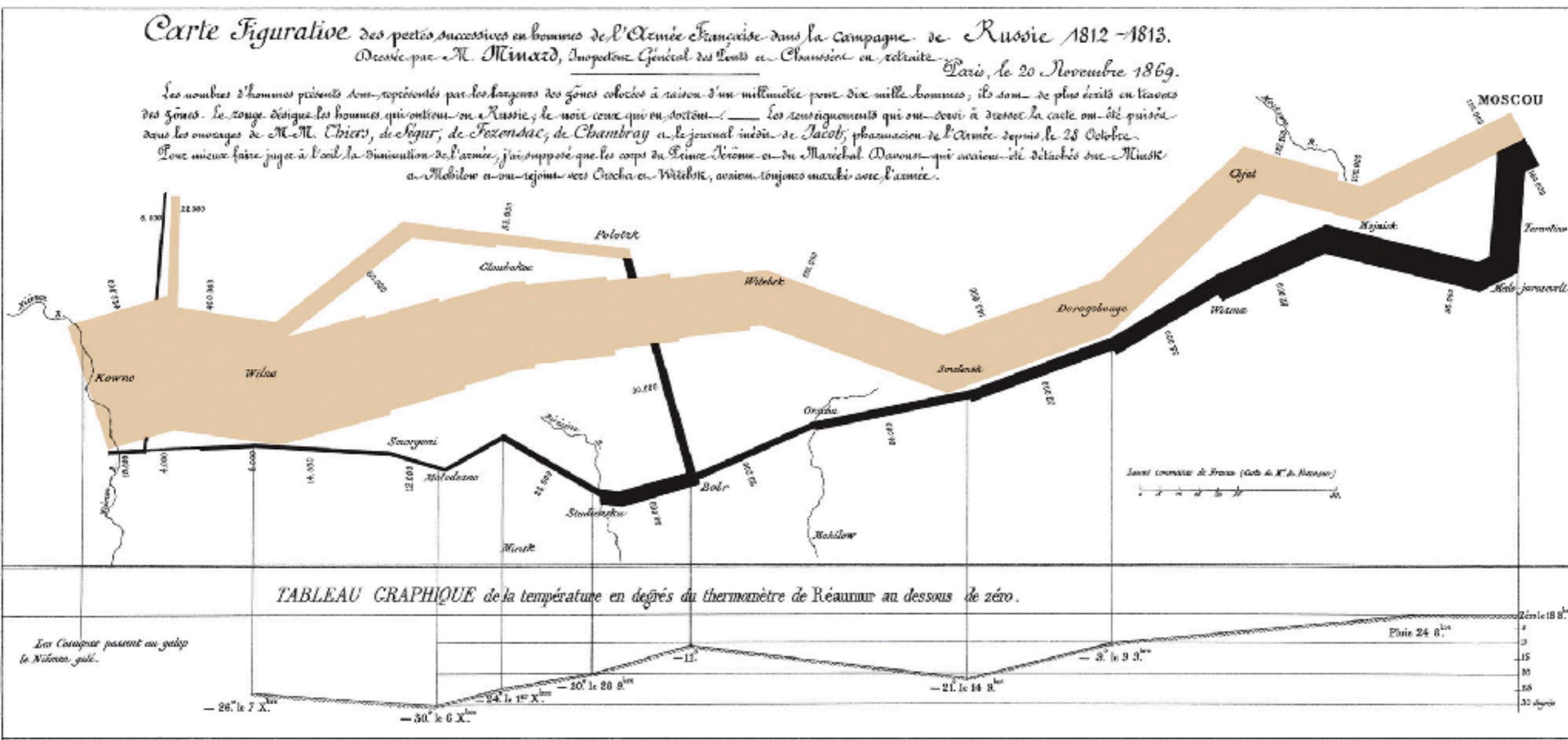


@hnrikhndborg | Source: American Time Use Survey



What are the marks? What are the channels?

Are they used effectively?



(maps out Napoleon's army going to Moscow (tan) and then retreating back (black), while diminishing from freezing winter temperatures (aligned line chart on bottom))

