



**Cody Dunne**  
Northeastern University

---

WHAT IS VISUALIZATION?



Feel free to interrupt with  
questions!

# Plan for Today

- Discuss what visualization is & why we care
- Get to know each other

# STAFF INTRODUCTIONS

## Instructors



[Cody Dunne](#)

[codydunne-and-tas@ccs.neu.edu](mailto:codydunne-and-tas@ccs.neu.edu), [c.dunne@northeastern.edu](mailto:c.dunne@northeastern.edu)

Office Hours: Tuesdays, 1–2pm on [Khoury Office Hours](#)

Fun Fact: I like pretty pictures that tell you something about data!

## Library Visualization Specialist



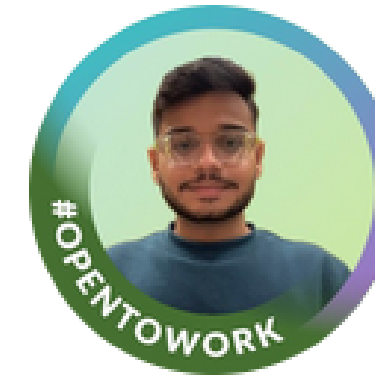
[Kate Kryder](#)

[codydunne-and-tas@ccs.neu.edu](mailto:codydunne-and-tas@ccs.neu.edu), [k.kryder@northeastern.edu](mailto:k.kryder@northeastern.edu)

Office Hours: By appointment on Zoom ([schedule here](#))

Fun Fact: Students who visit me often refer to the D3 Reusable Chart Model as the “Cody code.”

## Teaching Assistants



[Dhruv Miyani](#)

[codydunne-and-tas@ccs.neu.edu](mailto:codydunne-and-tas@ccs.neu.edu), [miyani.d@northeastern.edu](mailto:miyani.d@northeastern.edu)

Office Hours: Wednesdays 12–3:30pm, location TBD



[Evan Suslovich](#)

[codydunne-and-tas@ccs.neu.edu](mailto:codydunne-and-tas@ccs.neu.edu), [suslovich.e@northeastern.edu](mailto:suslovich.e@northeastern.edu)

Office Hours: Thursdays 1–4:30pm, location TBD

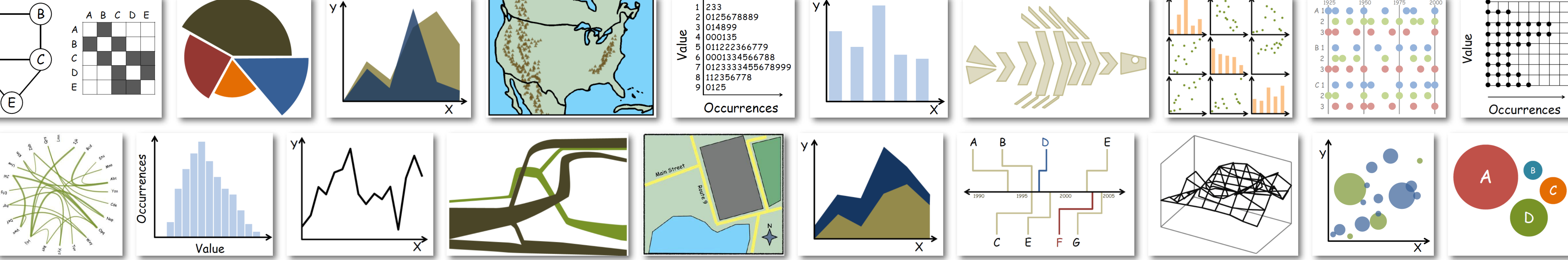


[Sibi Thirukonda](#)

[codydunne-and-tas@ccs.neu.edu](mailto:codydunne-and-tas@ccs.neu.edu), [thirukonda.s@northeastern.edu](mailto:thirukonda.s@northeastern.edu)

Office Hours: Fridays 6–8pm on [Khoury Office Hours](#)

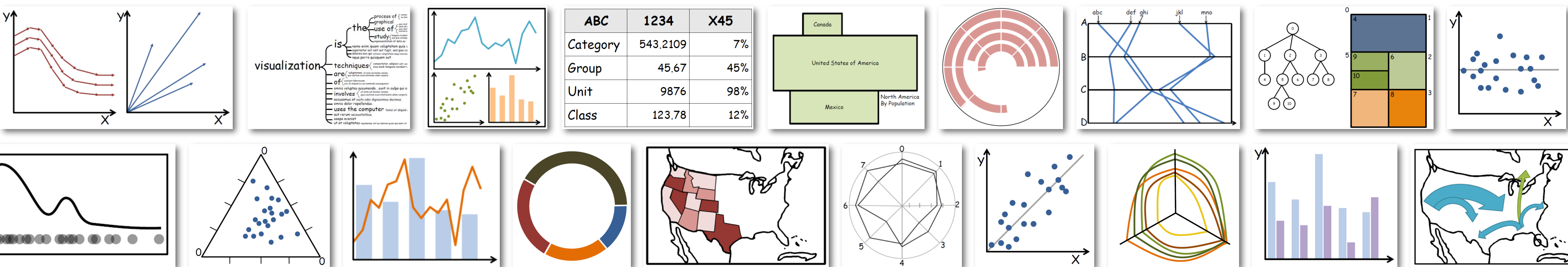
What is visualization  
anyway?



(static or interactive)

(abstract or spatial)

visualization: the visual representation of data to reinforce human cognition

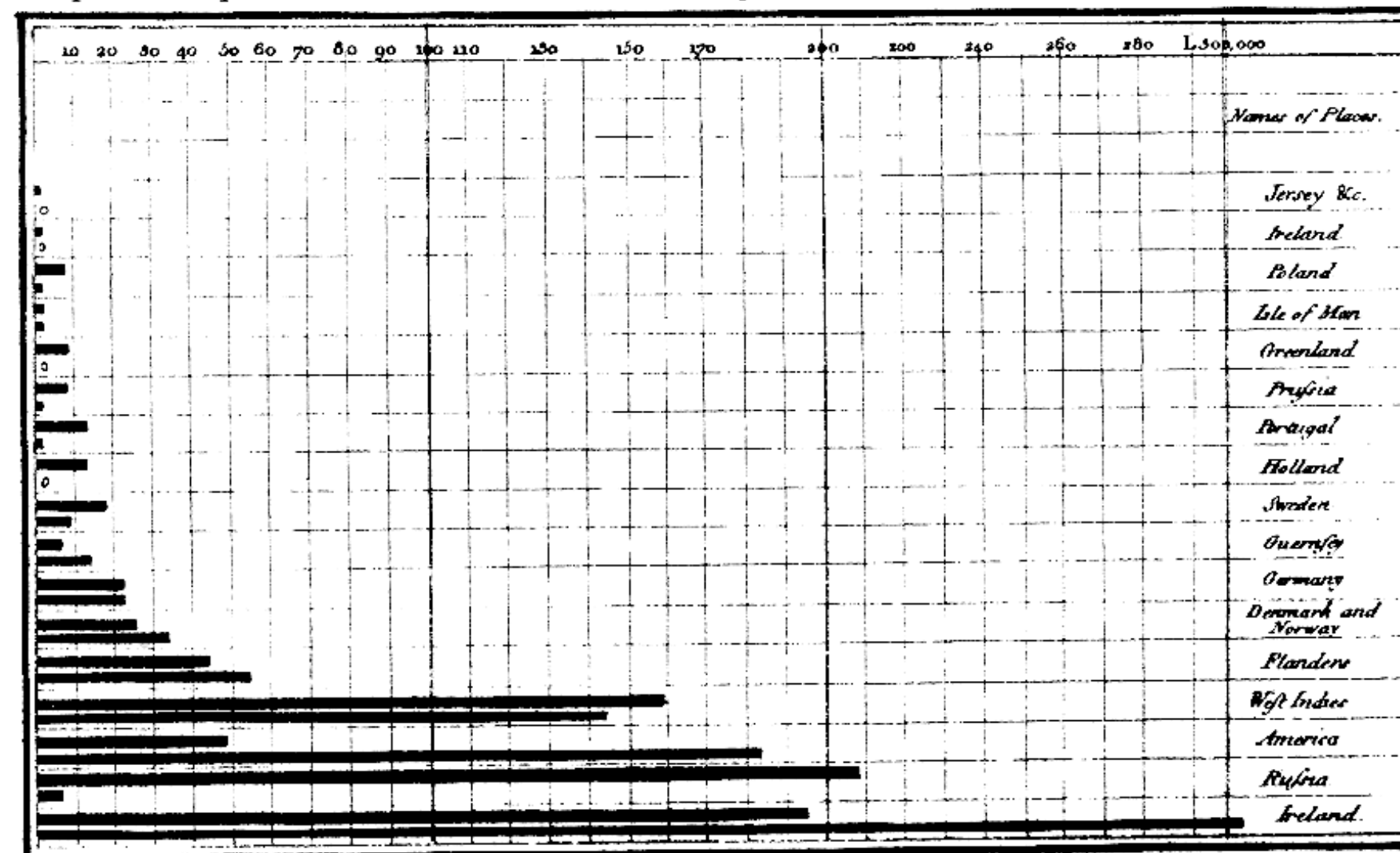






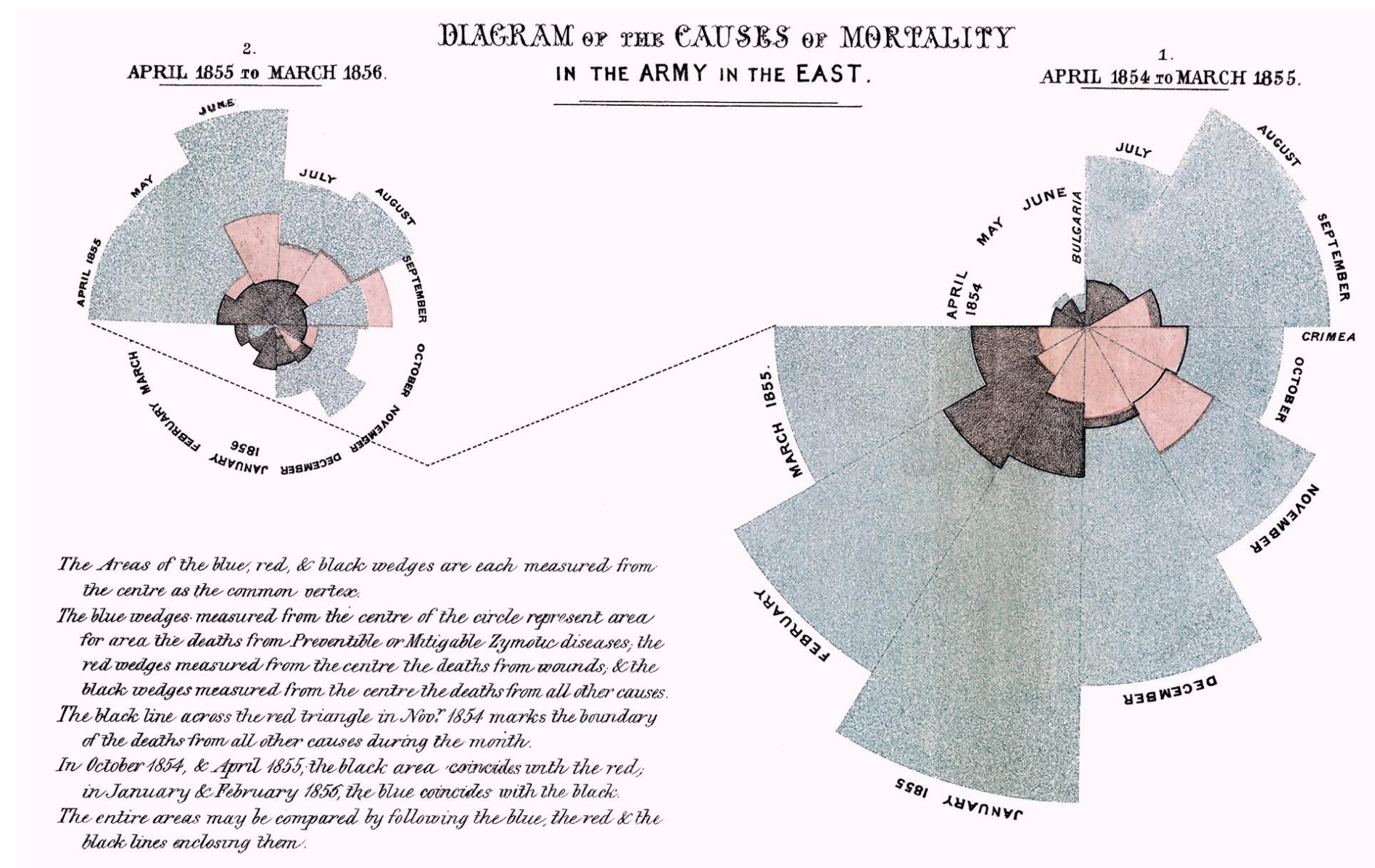
PTOLEMY (c. 150)

Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.

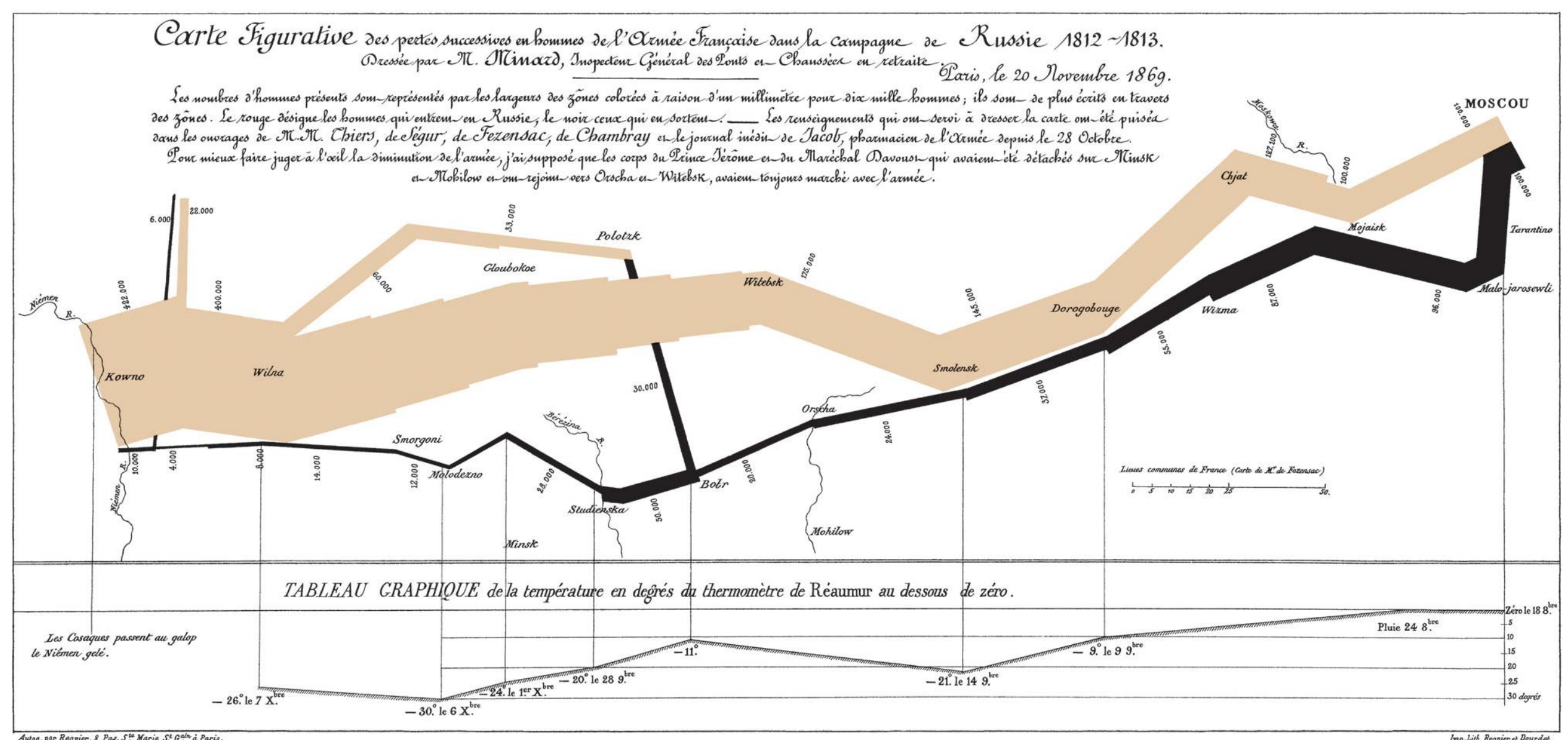


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.  
 Published as the Act above June 7<sup>th</sup> 1788 by W<sup>m</sup> Playfair  
 Made up by J. S. & Co. London

WILLIAM PLAYFAIR (c. 1786)



FLORENCE NIGHTINGALE (c. 1858)



CHARLES MINARD (c. 1869)



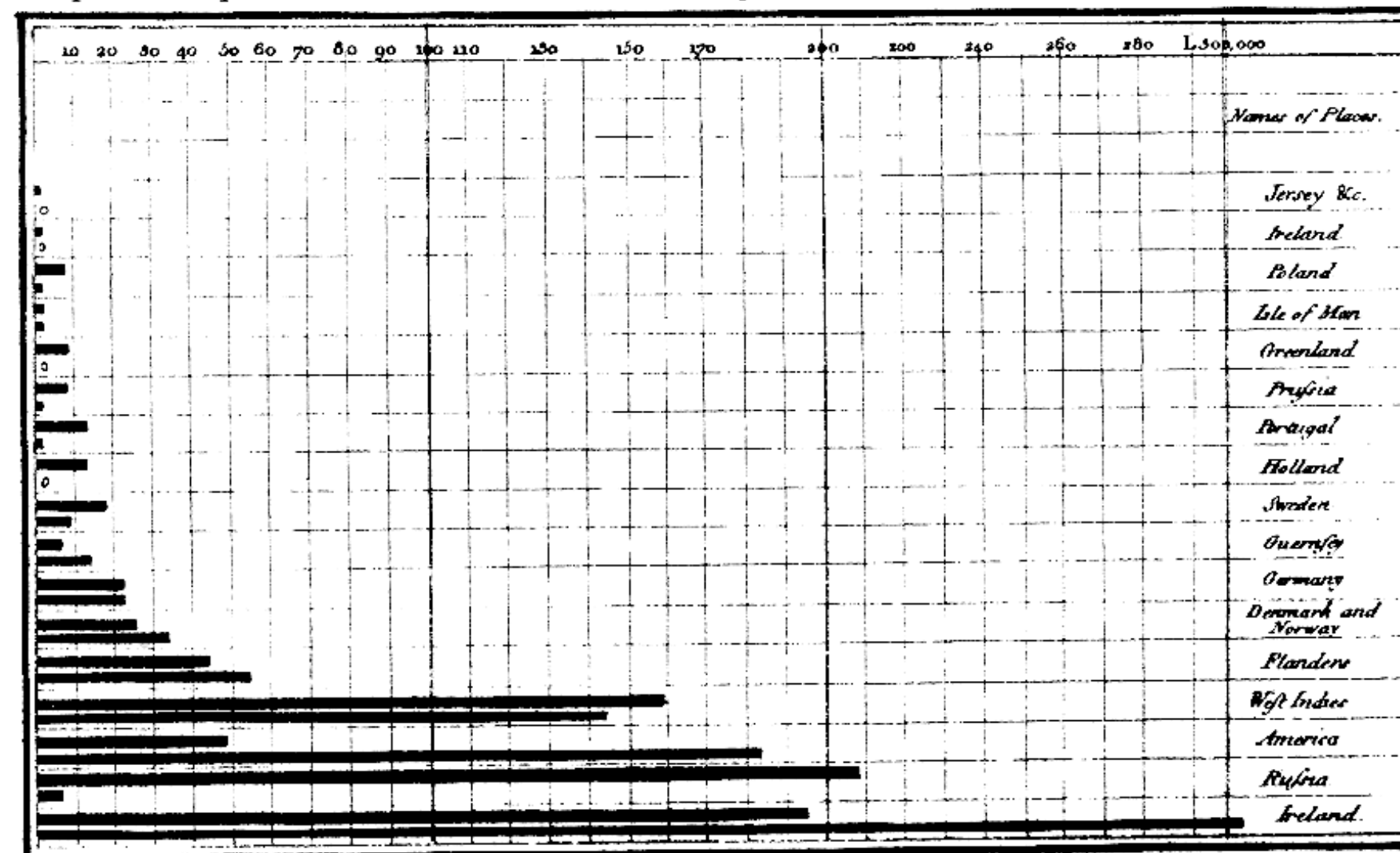






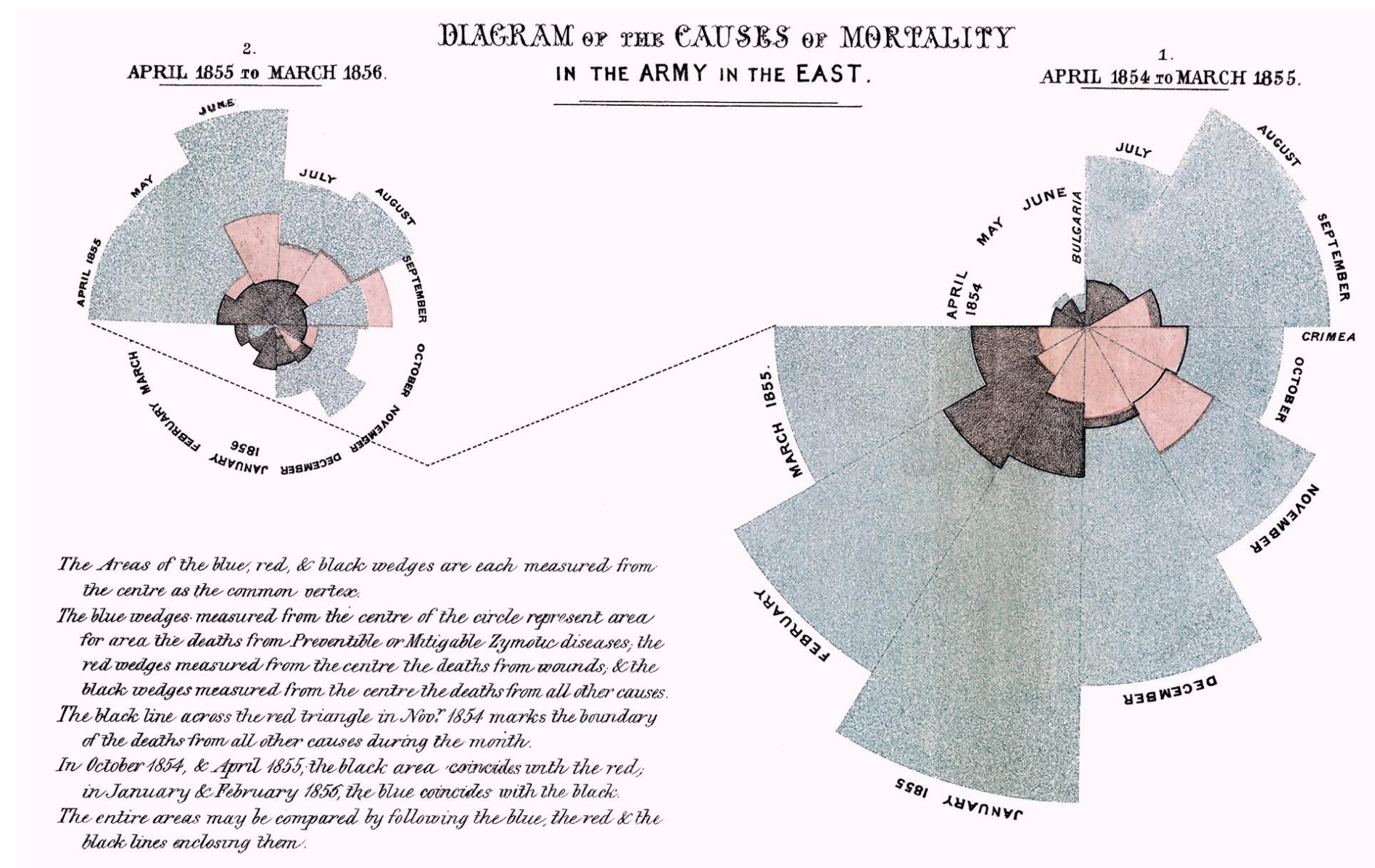
PTOLEMY (c. 150)

Exports and imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.

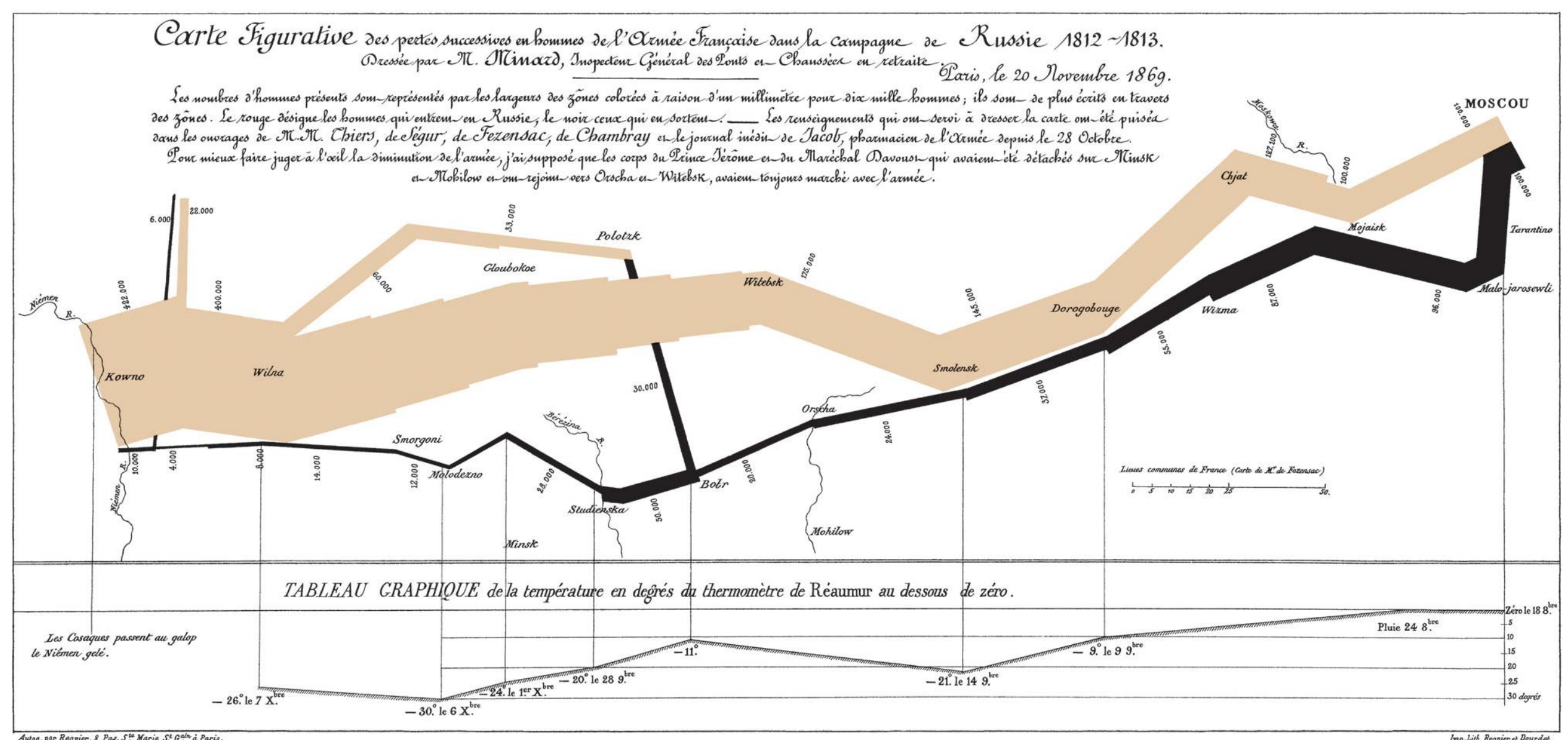


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.  
 Published as the Act above June 7<sup>th</sup> 1788 by W<sup>m</sup> Playfair  
 Made comp<sup>d</sup> 352. Strand. London.

WILLIAM PLAYFAIR (c. 1786)



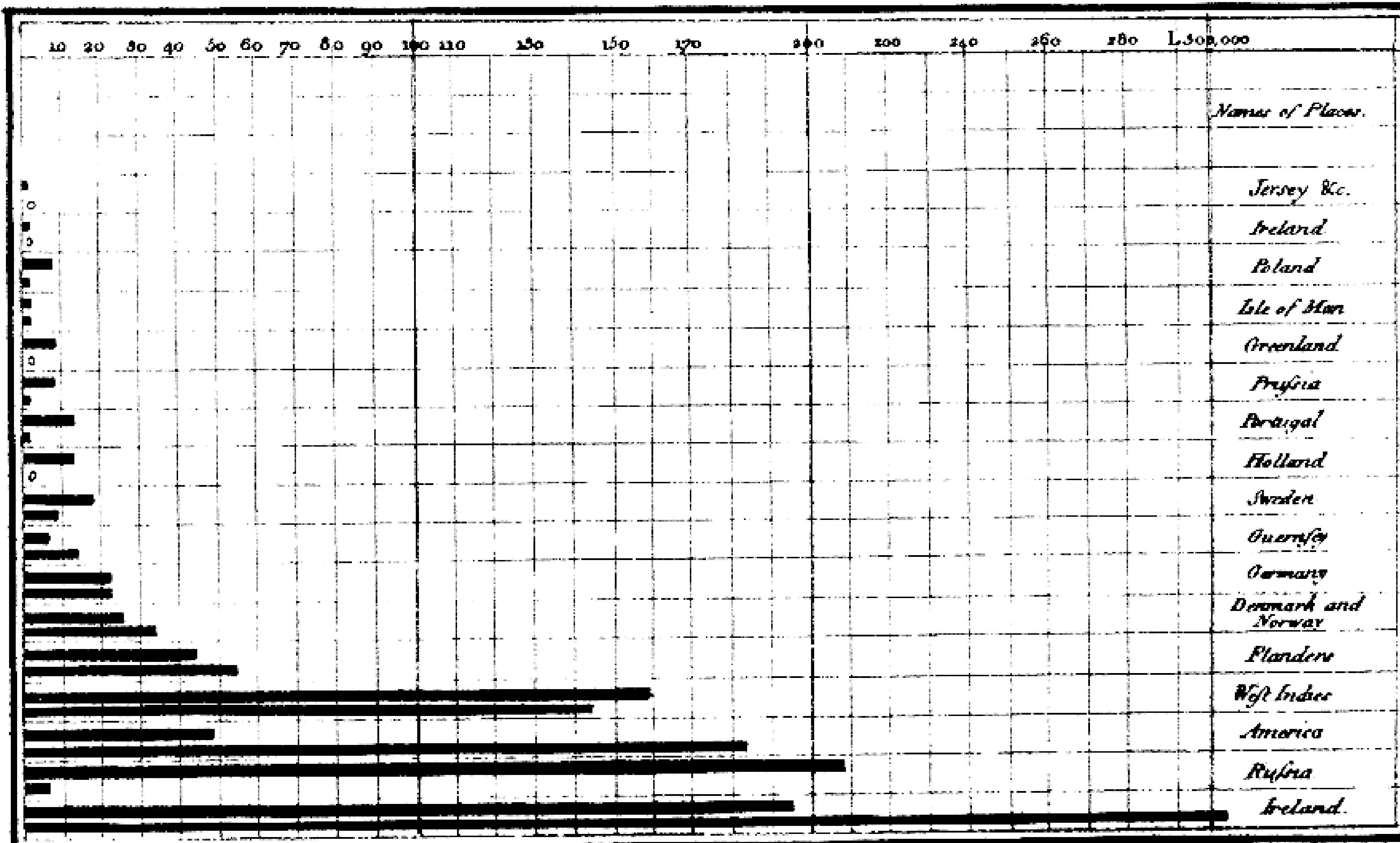
FLORENCE NIGHTINGALE (c. 1858)



CHARLES MINARD (c. 1869)



Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.



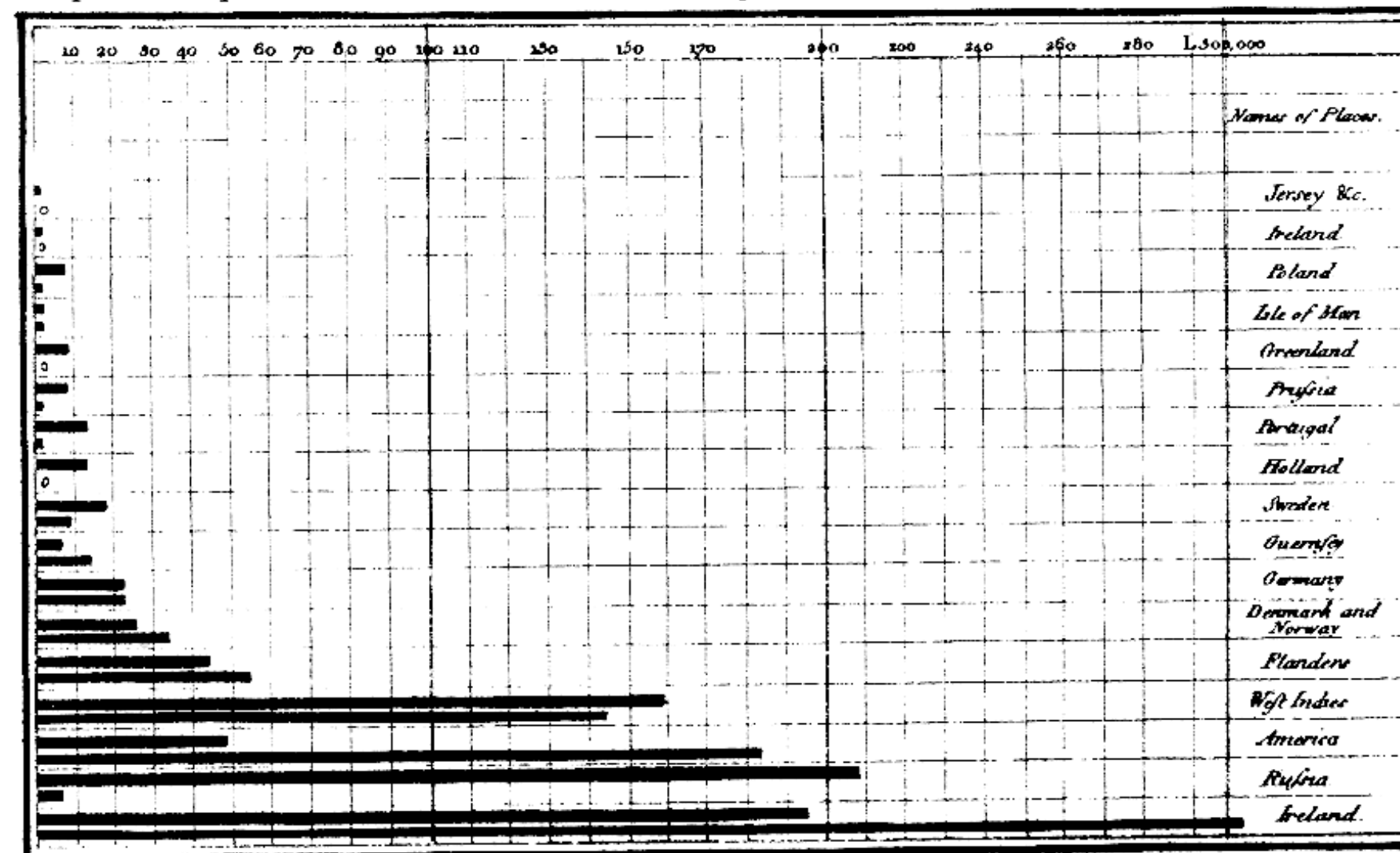
The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.





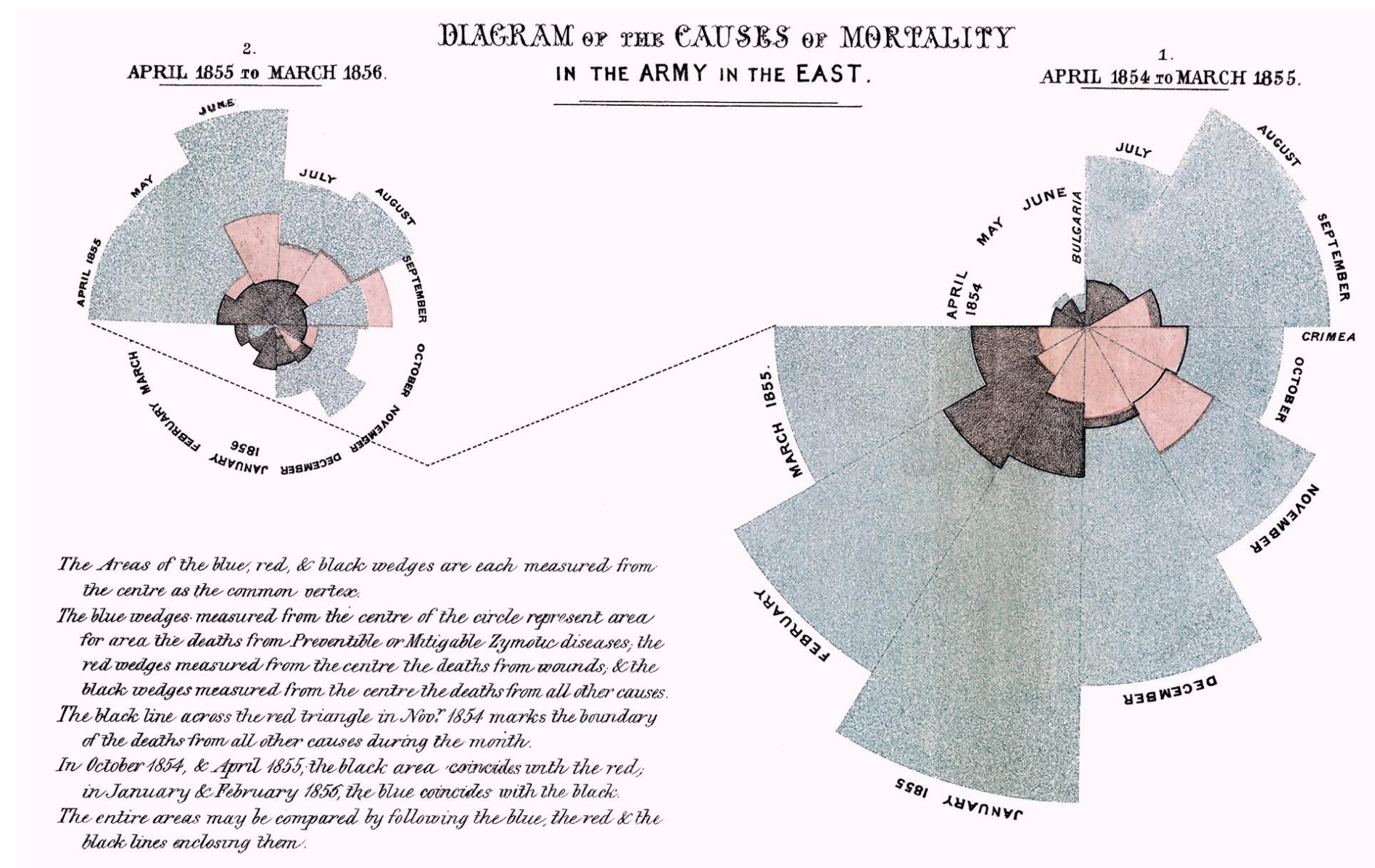
PTOLEMY (c. 150)

Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.

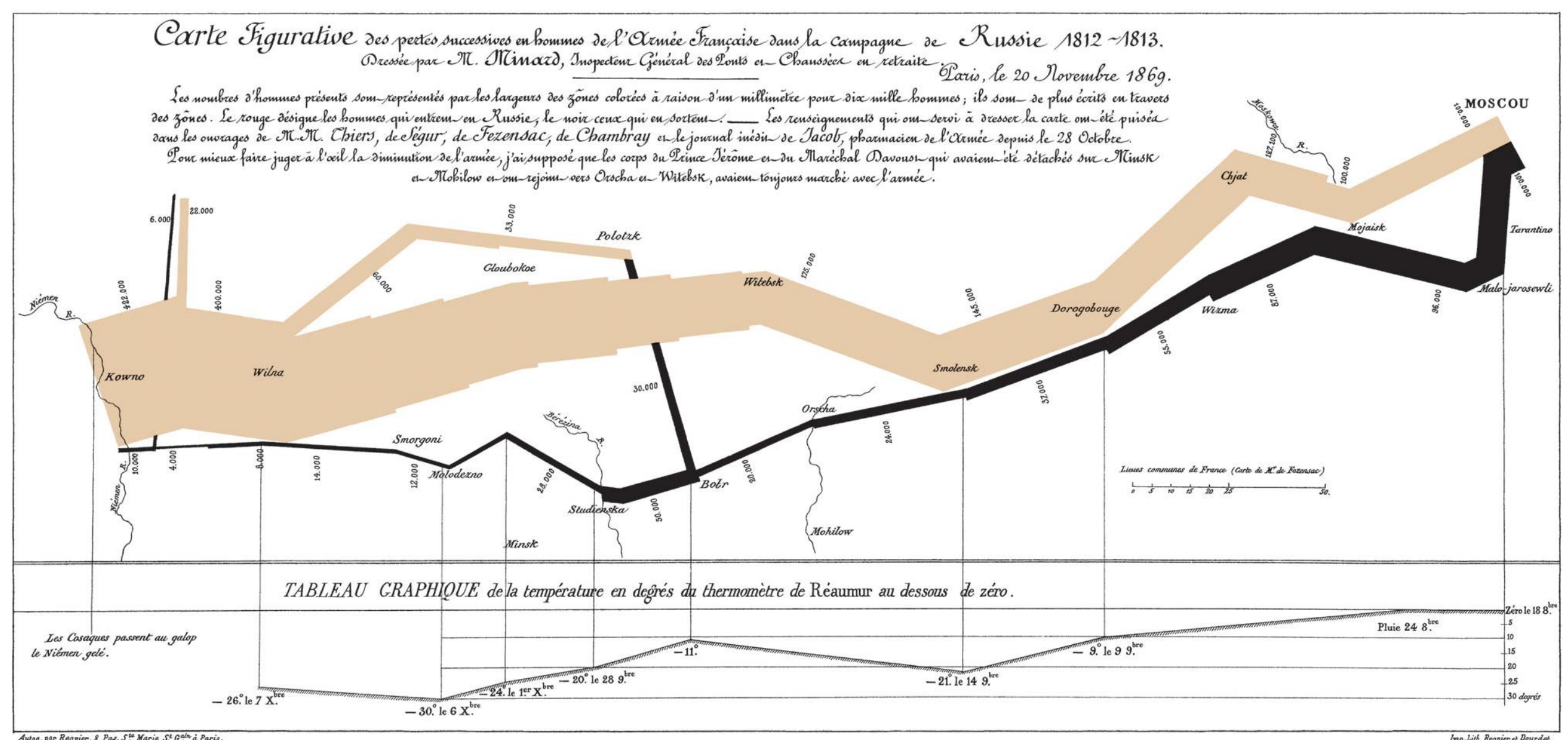


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.  
 Published in the Edinburgh Journal June 7<sup>th</sup> 1782 by W. Playfair  
 Made comp<sup>t</sup> 352. 1/2 Lond<sup>n</sup>

WILLIAM PLAYFAIR (c. 1786)



FLORENCE NIGHTINGALE (c. 1858)



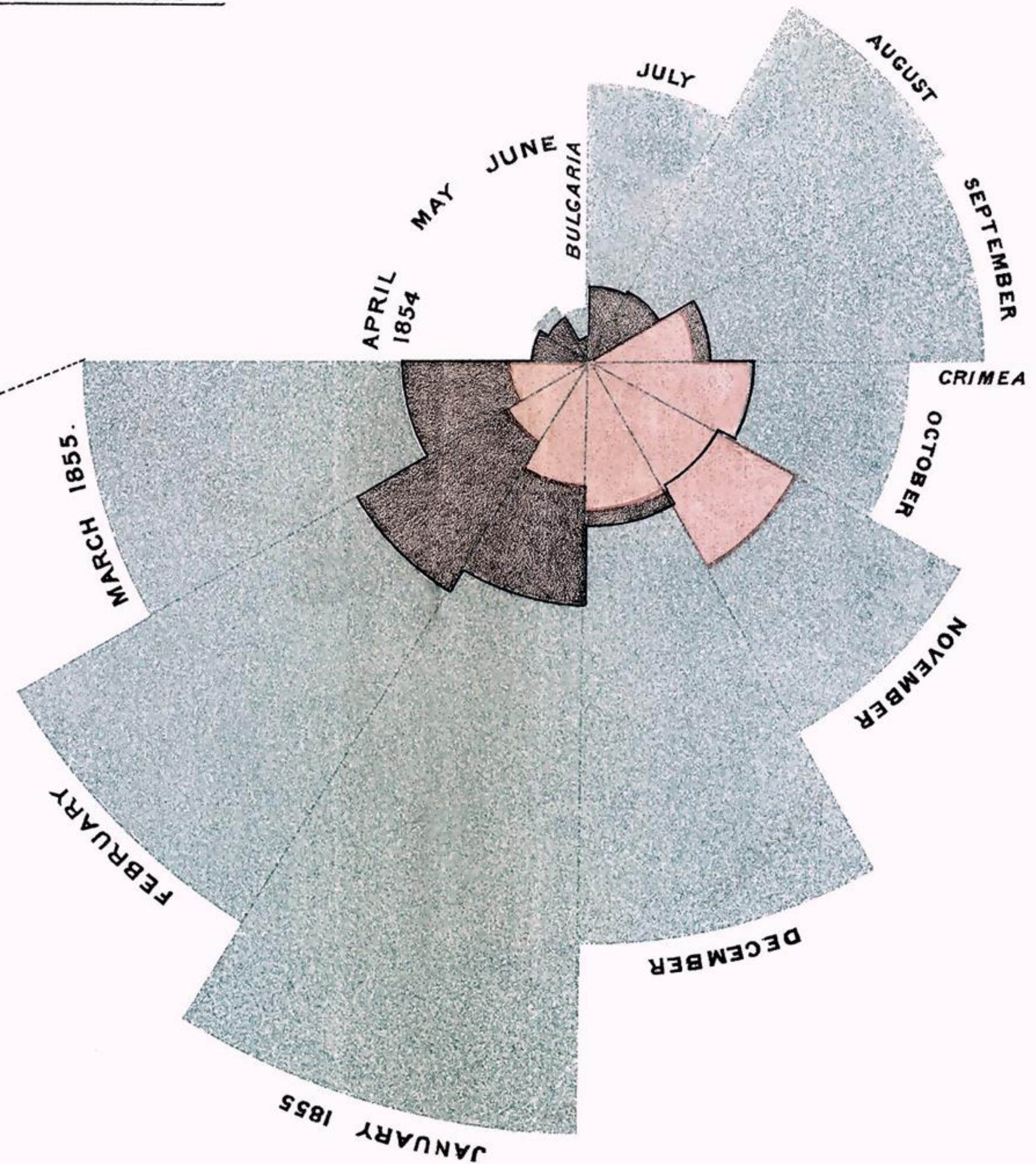
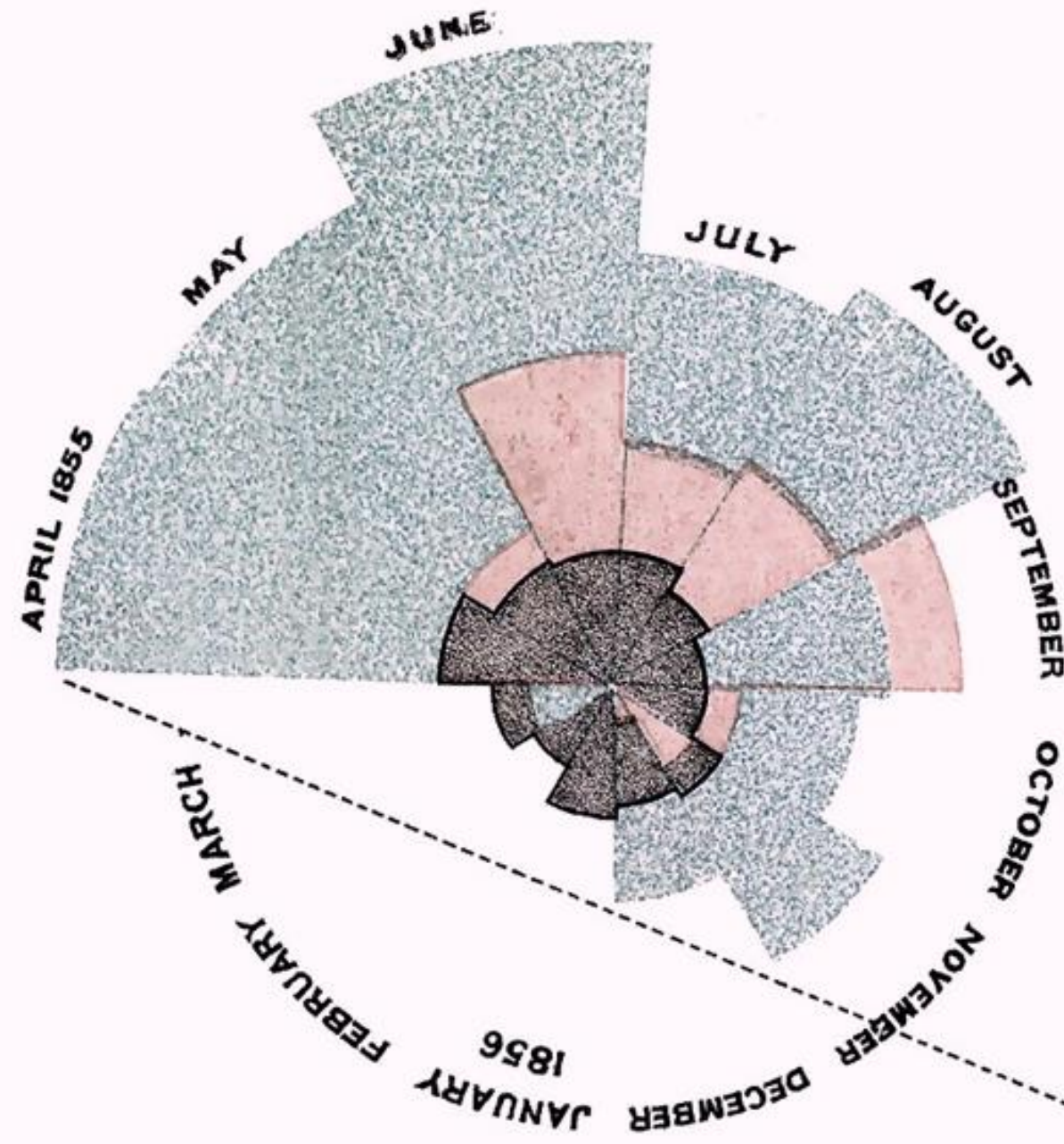
CHARLES MINARD (c. 1869)



# DIAGRAM OF THE CAUSES OF MORTALITY IN THE ARMY IN THE EAST.

1.  
APRIL 1854 TO MARCH 1855.

2.  
APRIL 1855 TO MARCH 1856.



*The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.*

*The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases; the red wedges measured from the centre the deaths from wounds; & the black wedges measured from the centre the deaths from all other causes.*

*The black line across the red triangle in Nov<sup>r</sup> 1854 marks the boundary of the deaths from all other causes during the month.*

*In October 1854, & April 1855, the black area coincides with the red; in January & February 1856, the blue coincides with the black.*

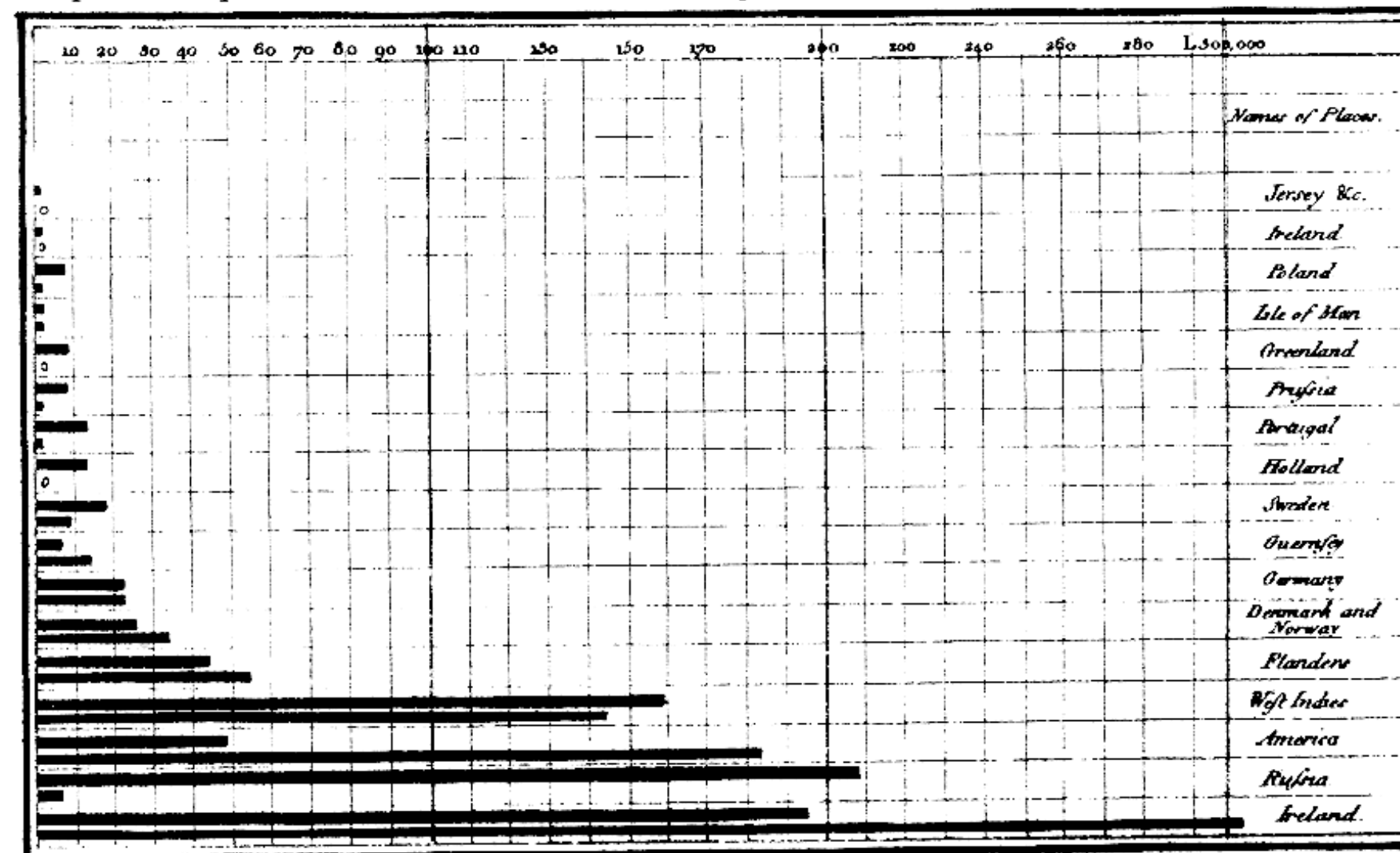
*The entire areas may be compared by following the blue, the red & the black lines enclosing them.*





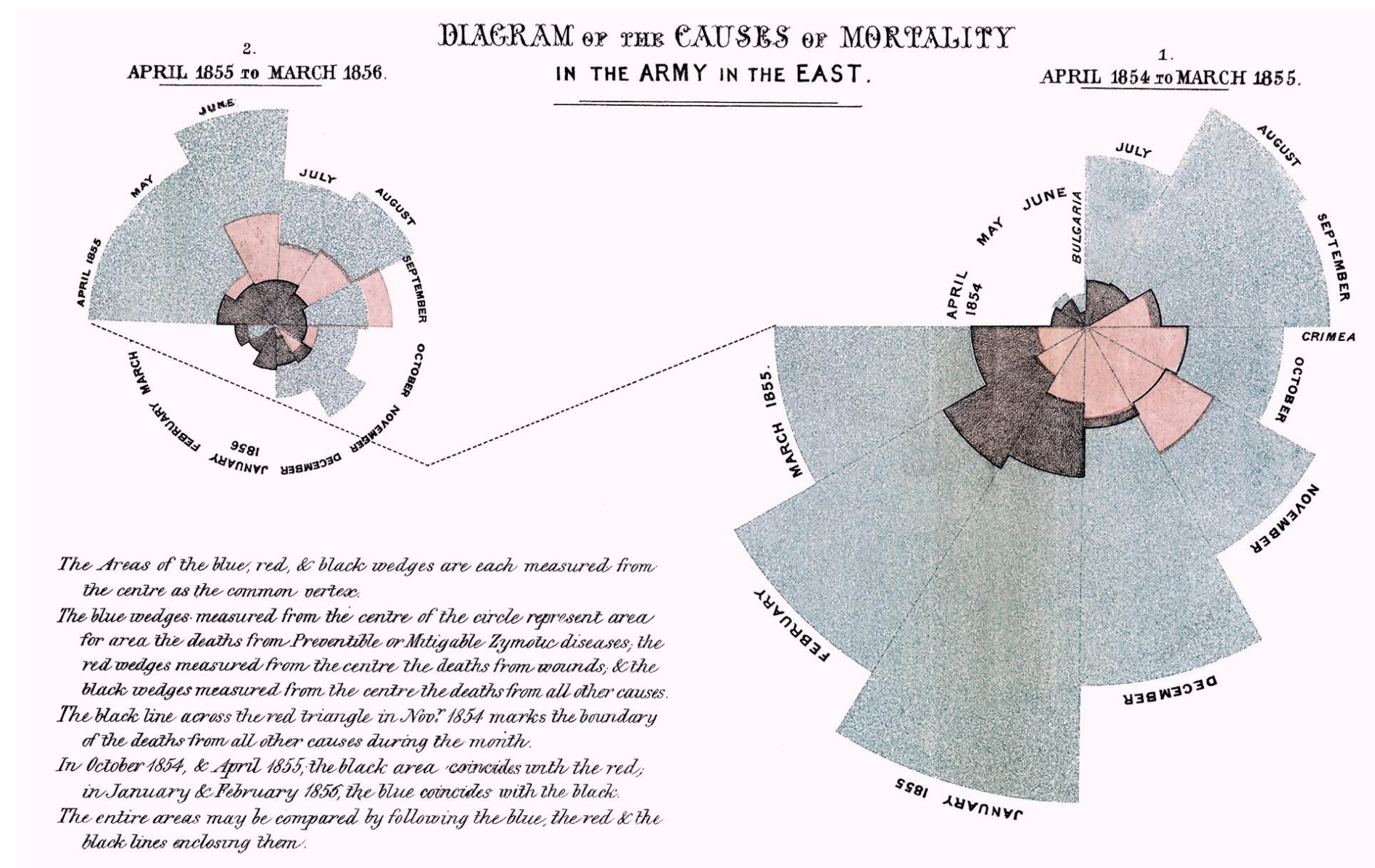
PTOLEMY (c. 150)

Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.

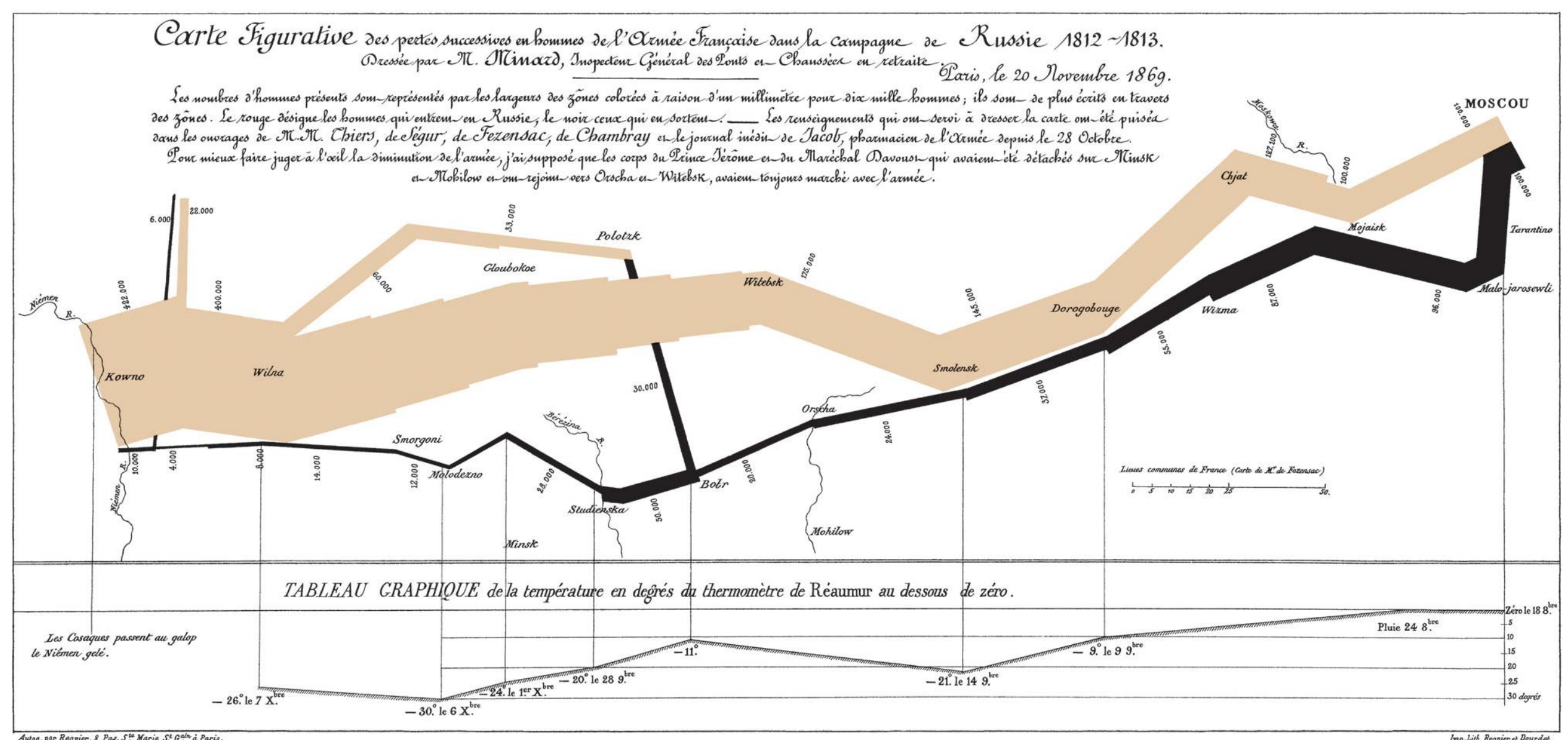


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.  
 Published as the Act above June 7<sup>th</sup> 1788 by W<sup>m</sup> Playfair  
 Made comp<sup>t</sup> 352. 1/2 Londn

WILLIAM PLAYFAIR (c. 1786)



FLORENCE NIGHTINGALE (c. 1858)



CHARLES MINARD (c. 1869)



# Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre. Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Mohilow et ont rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.

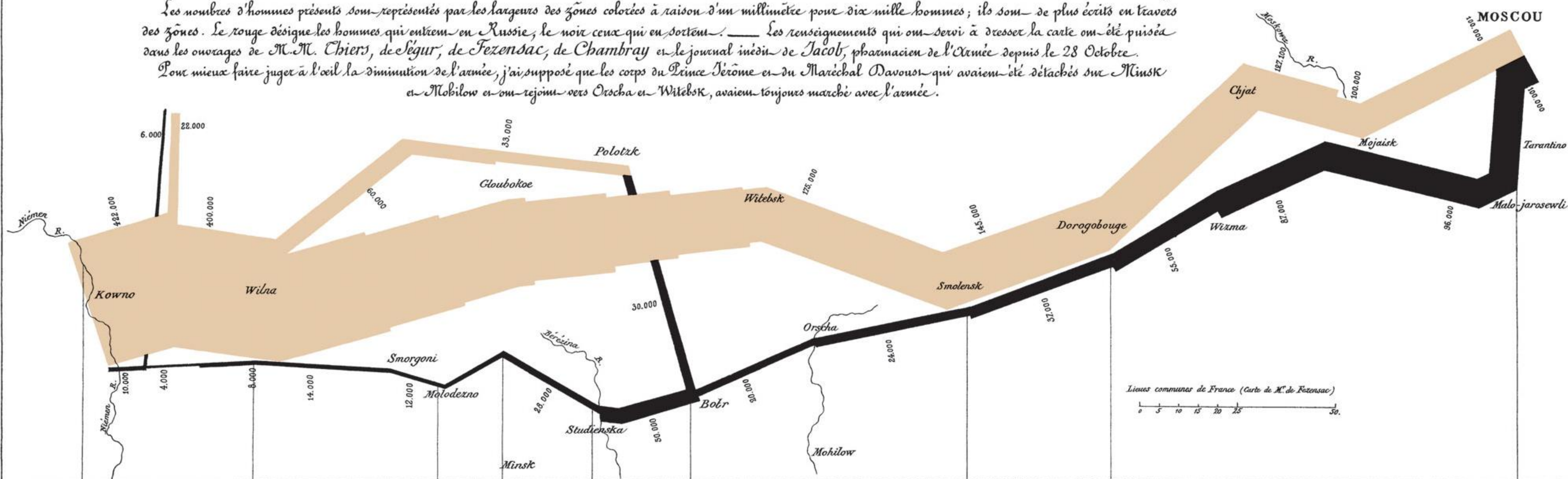
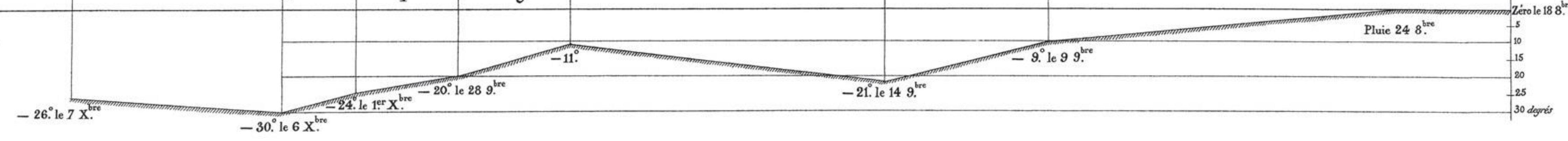


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Cosaques passent au galop le Niémen gelé.



Autog. par Regnier, 8. Pas. 5<sup>me</sup> Marie St G<sup>de</sup> à Paris.

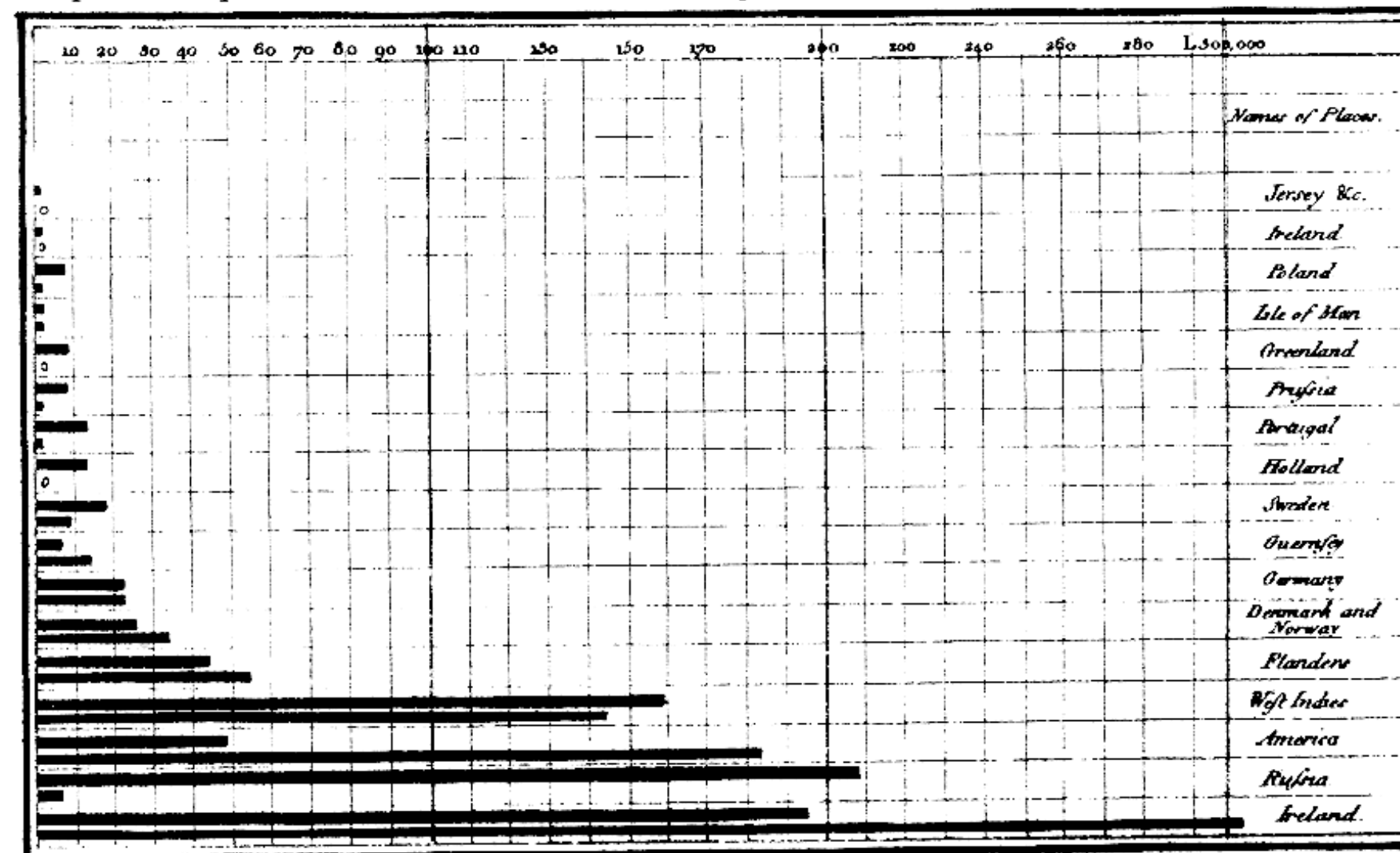
Imp. Lith. Regnier et Dourdet.





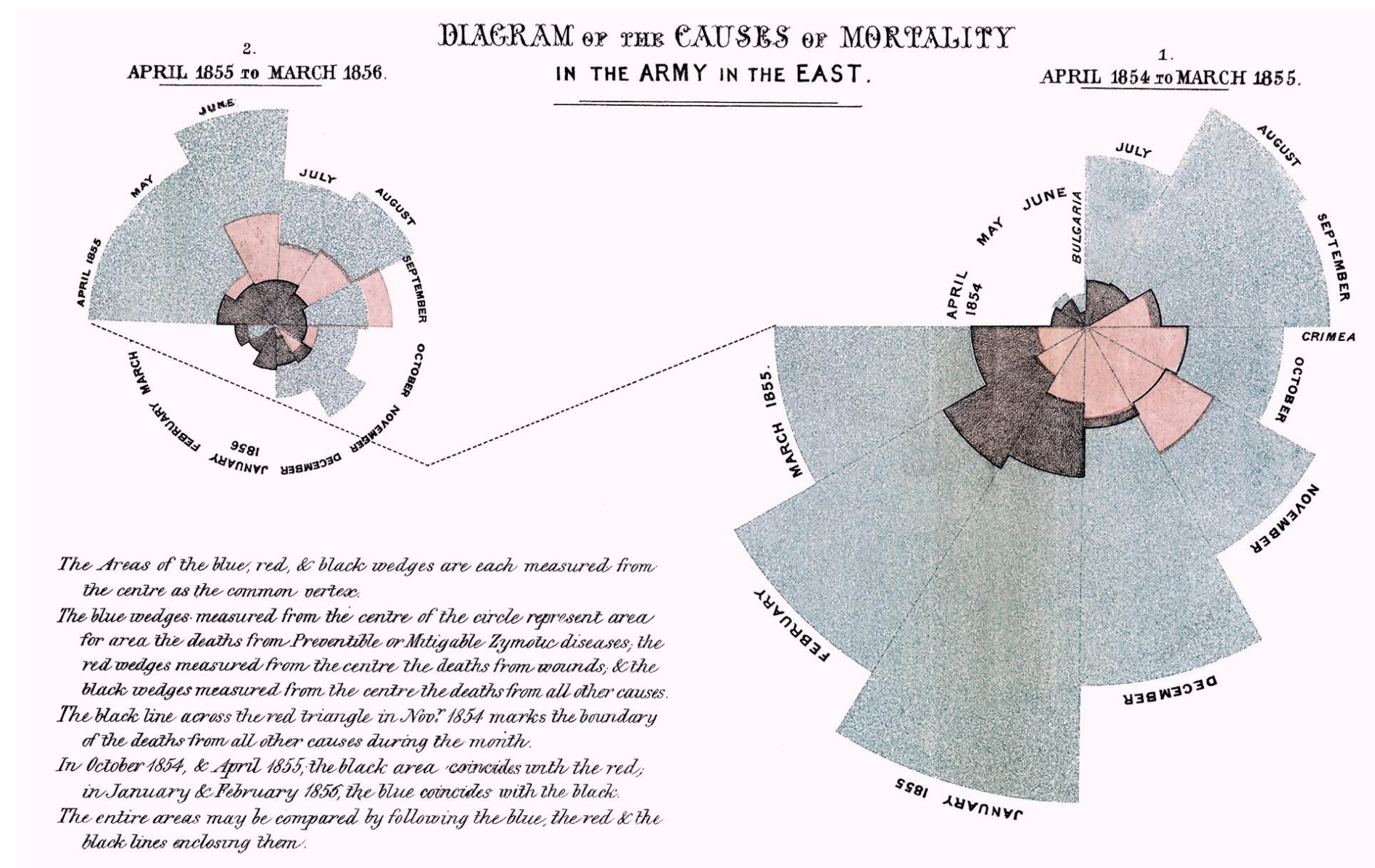
PTOLEMY (c. 150)

Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.

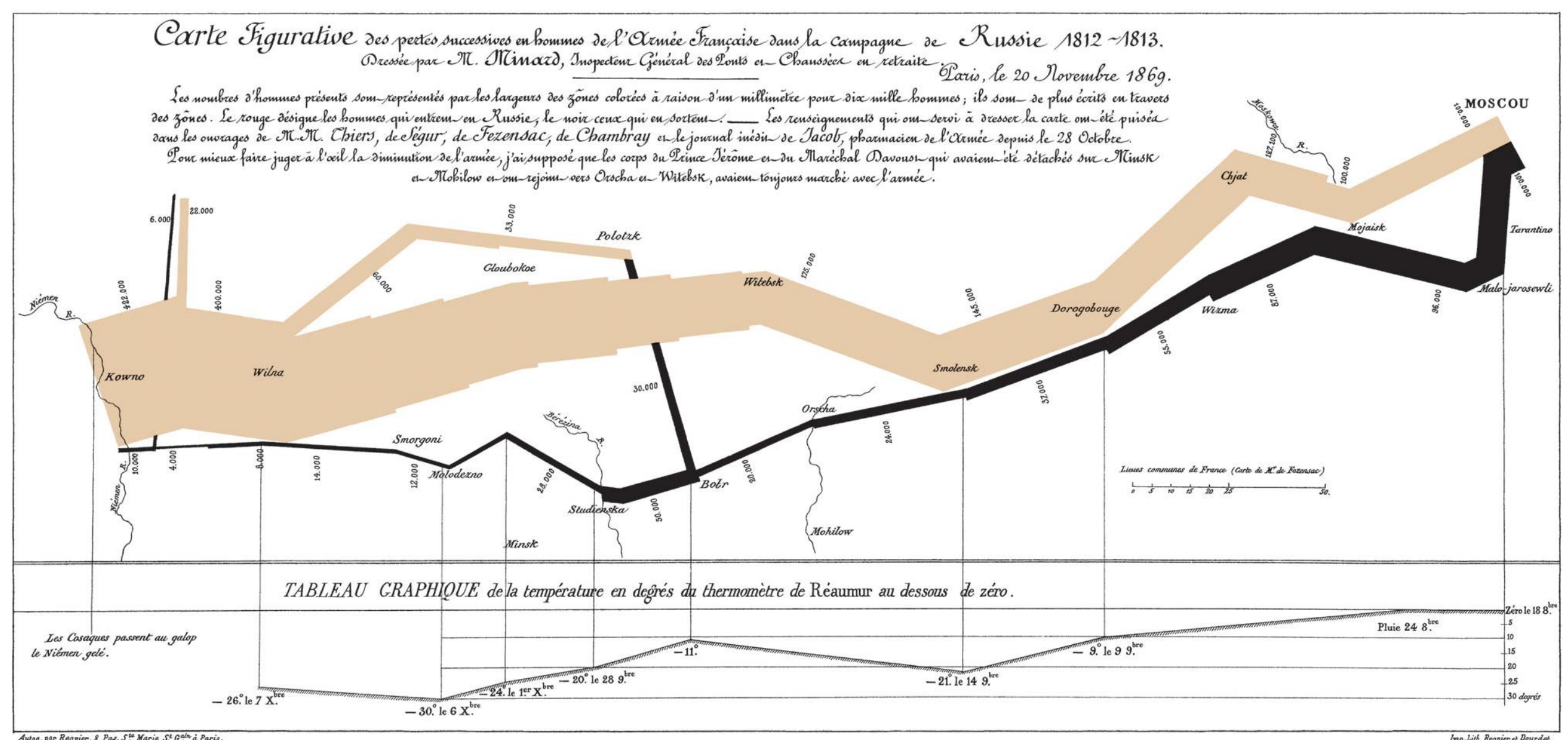


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.  
 Published as the Act above June 7<sup>th</sup> 1788 by W<sup>m</sup> Playfair  
 Made comp<sup>t</sup> 352. 1/2 Lond<sup>n</sup>

WILLIAM PLAYFAIR (c. 1786)

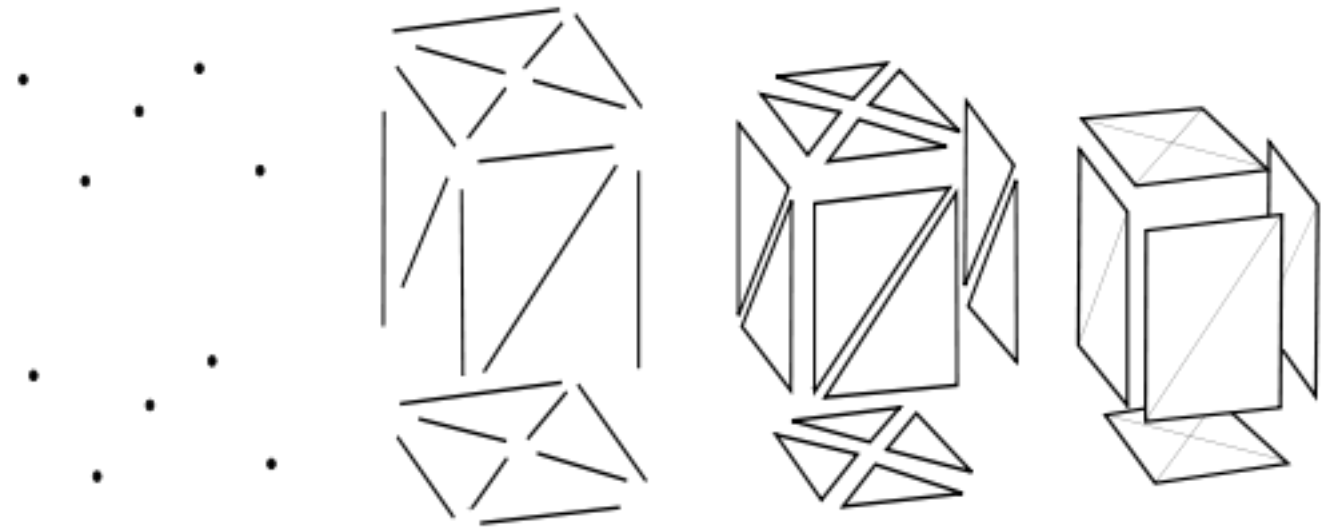


FLORENCE NIGHTINGALE (c. 1858)

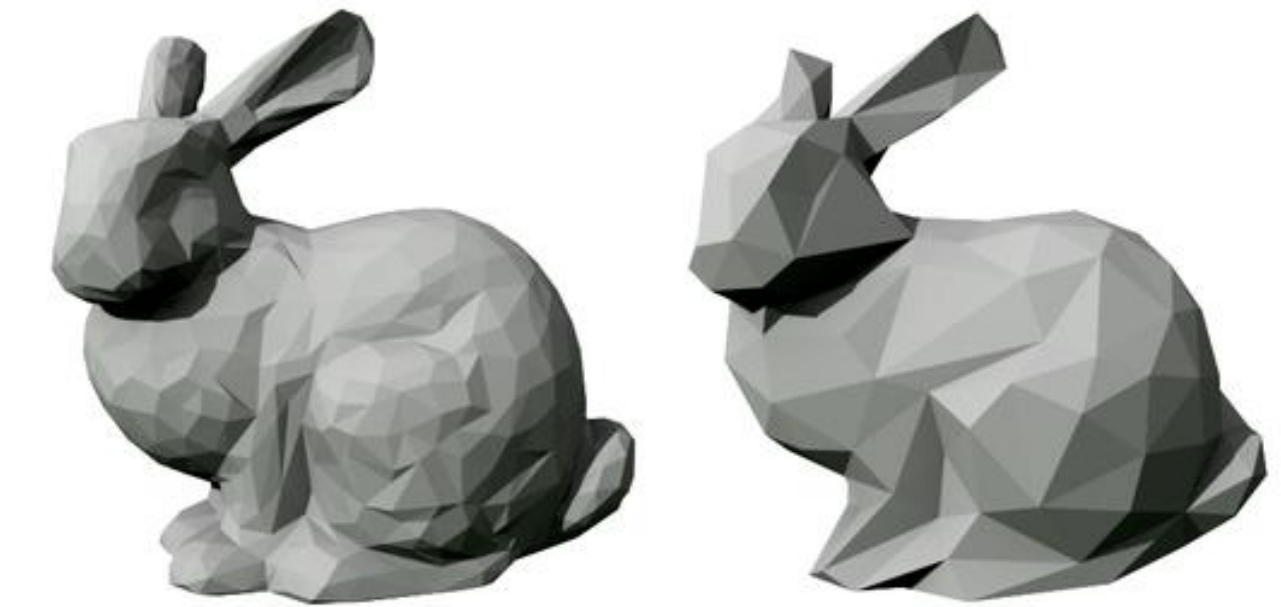


CHARLES MINARD (c. 1869)

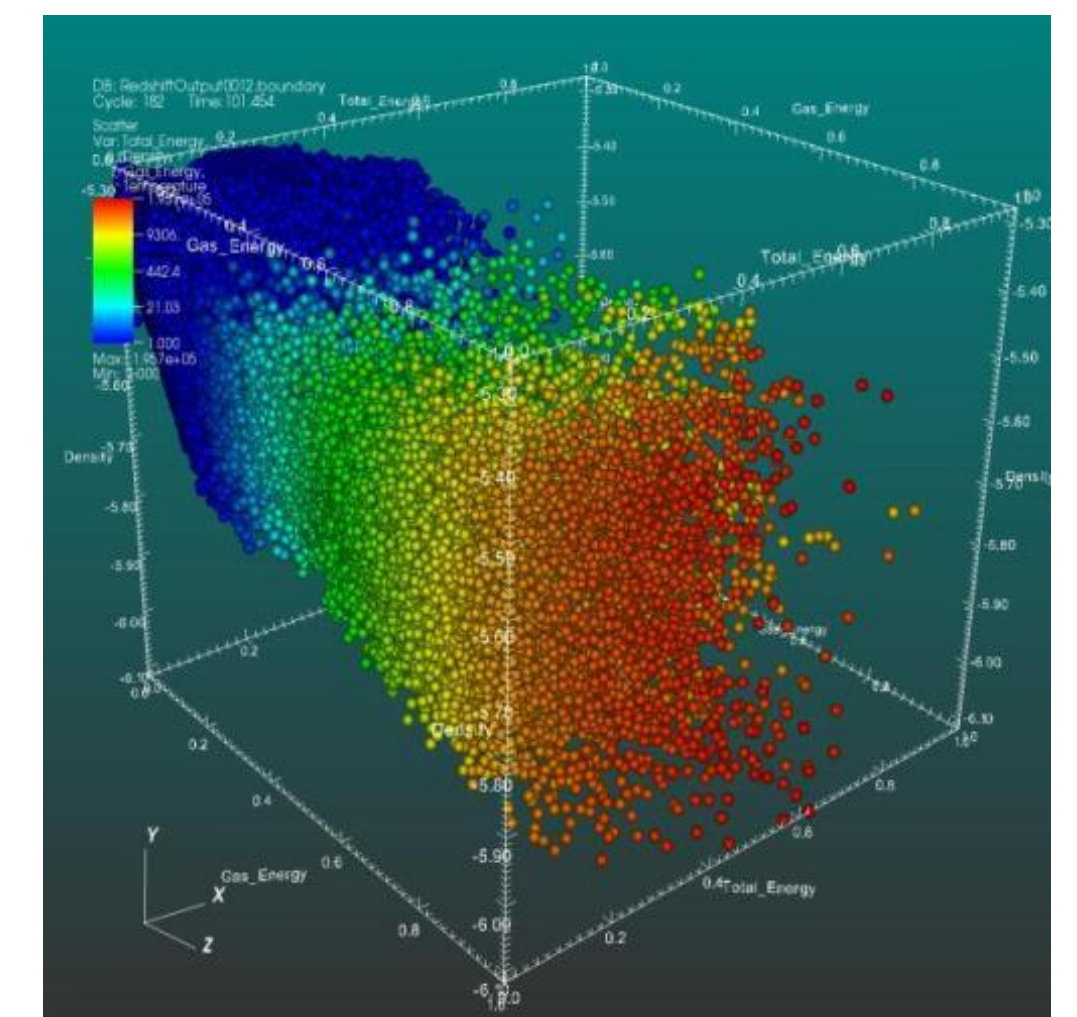
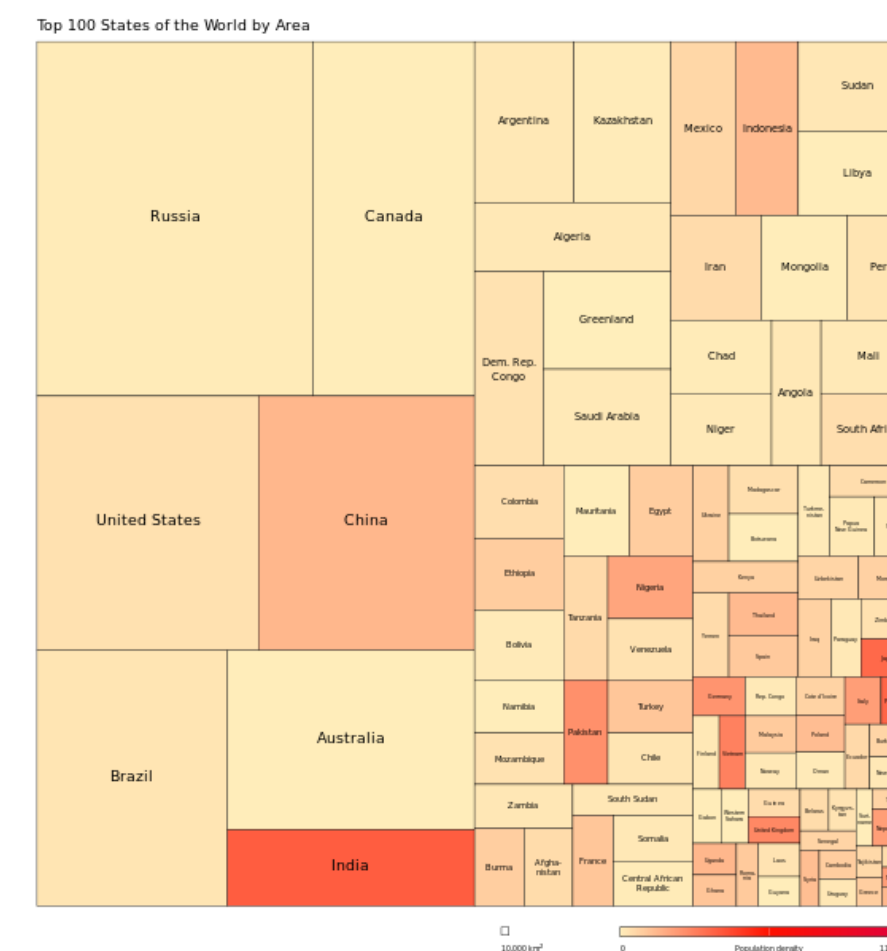
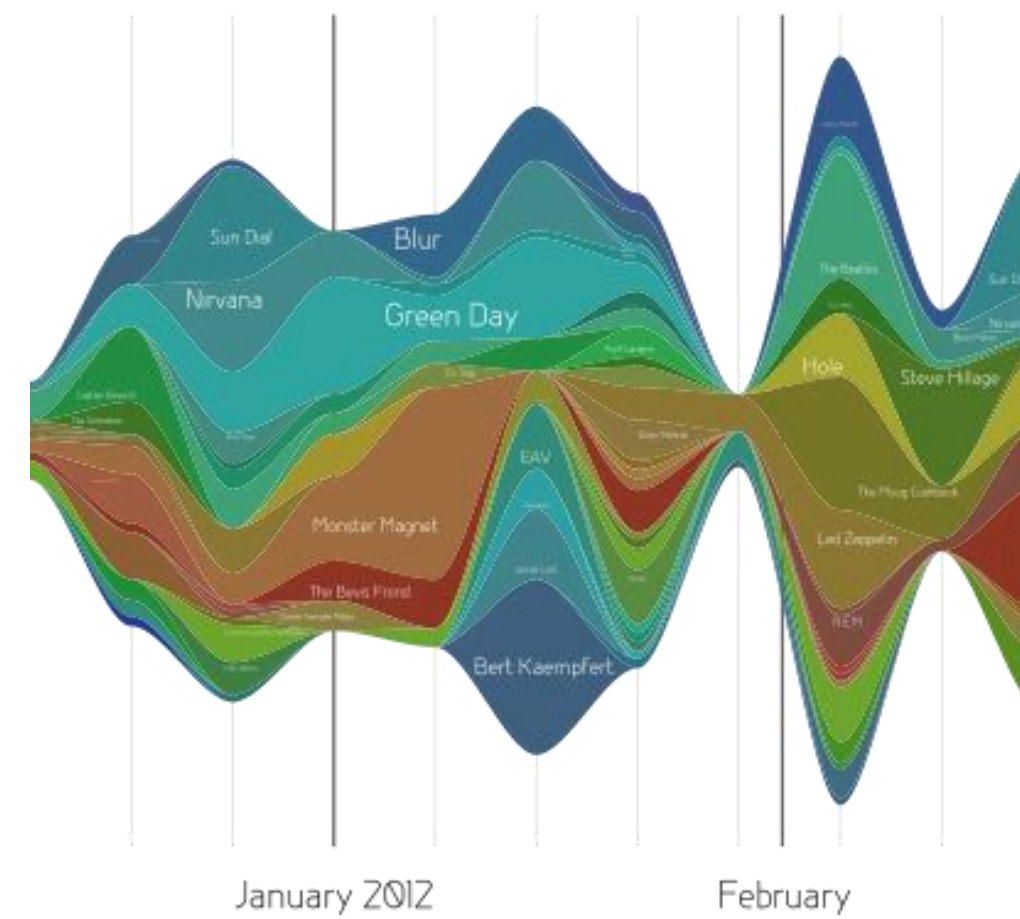
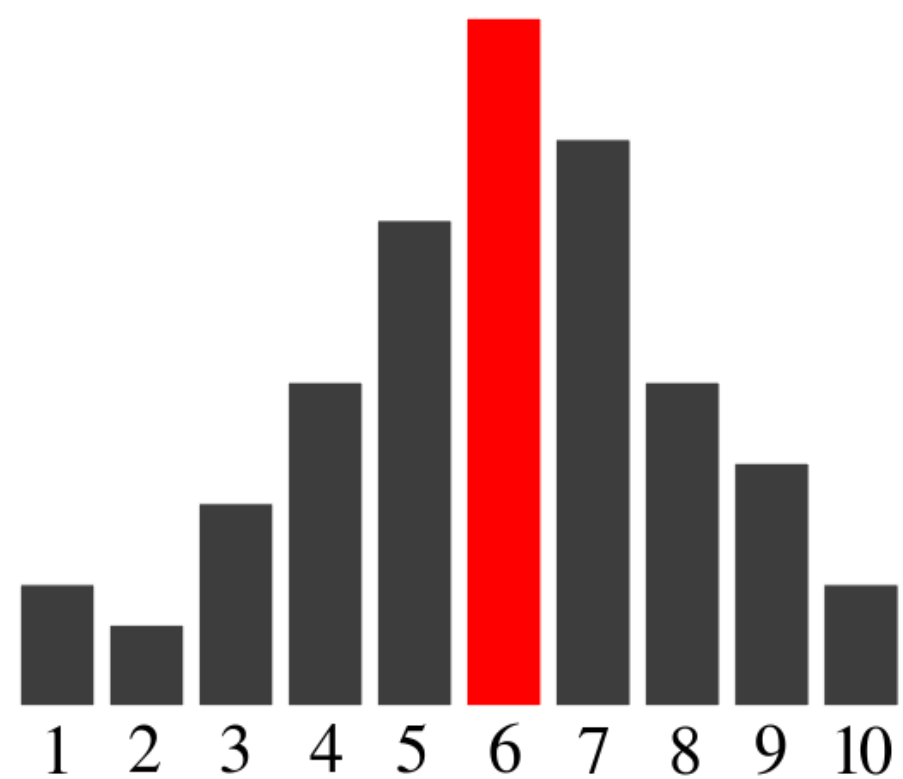




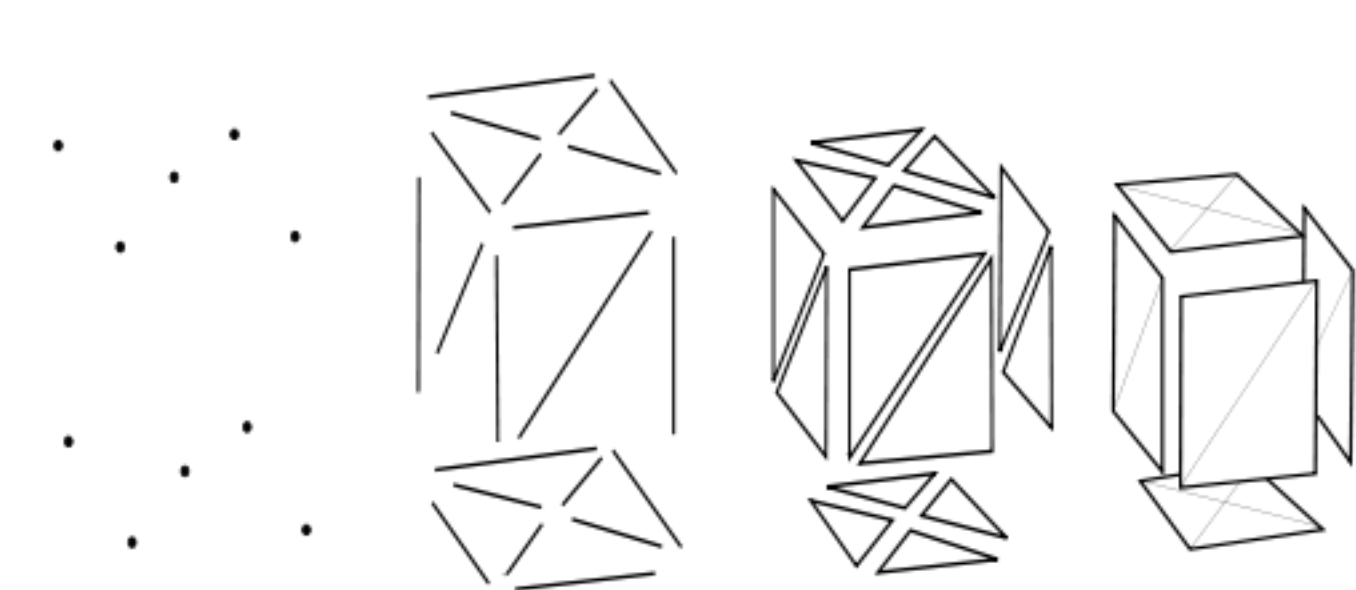
# computer graphics



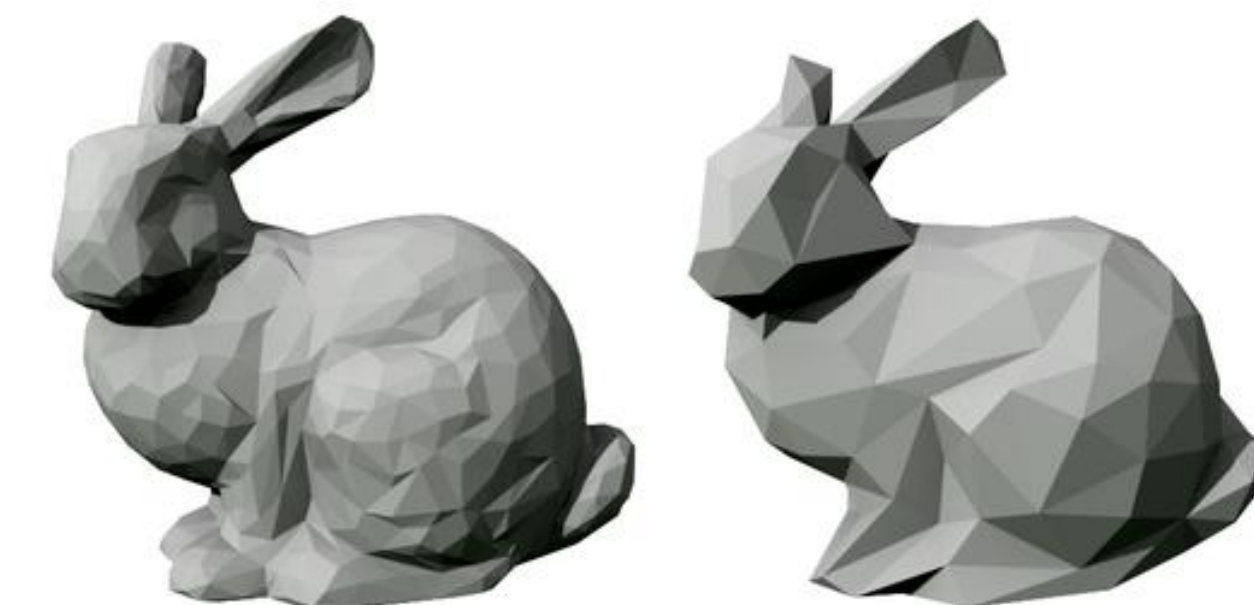
# visualization



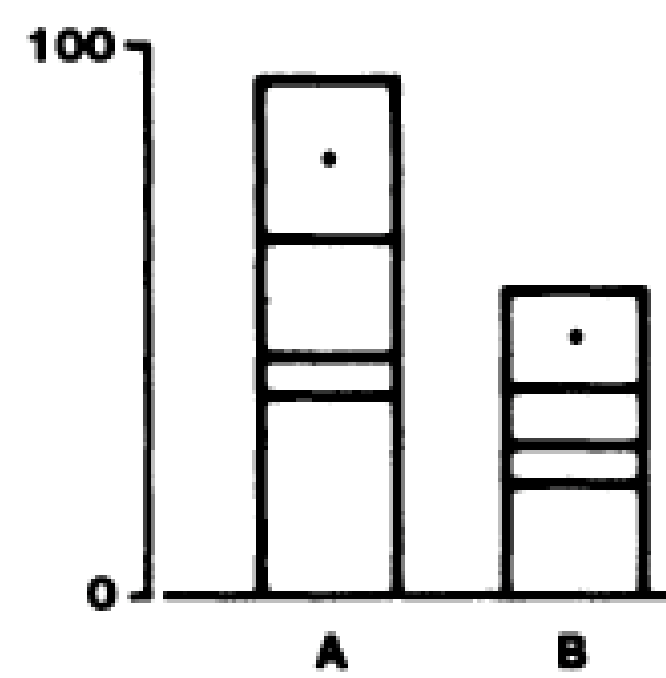
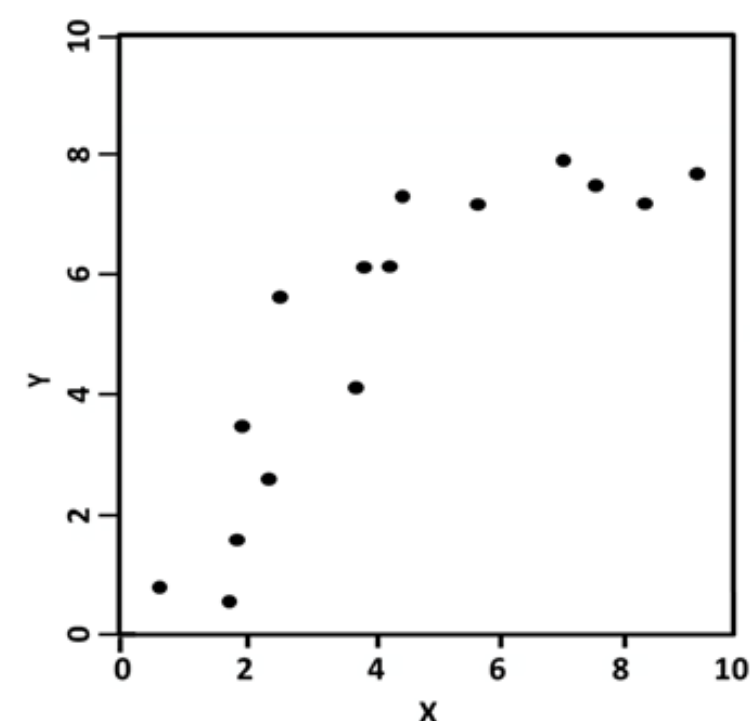




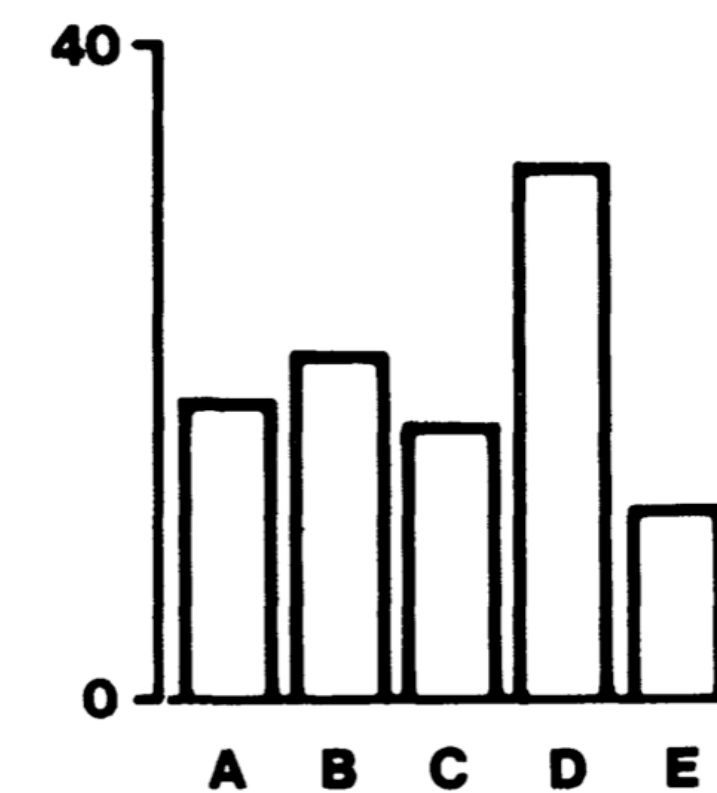
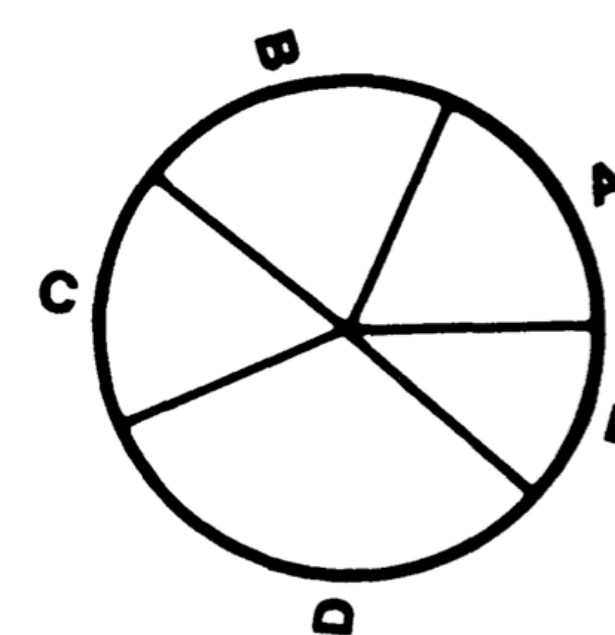
# computer graphics



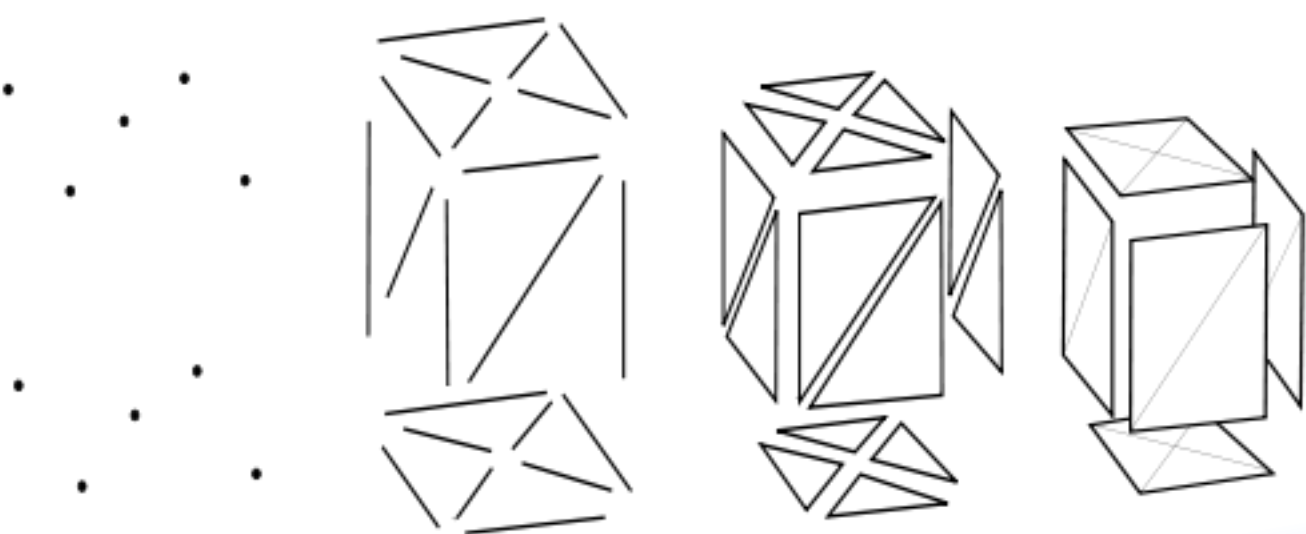
# visualization



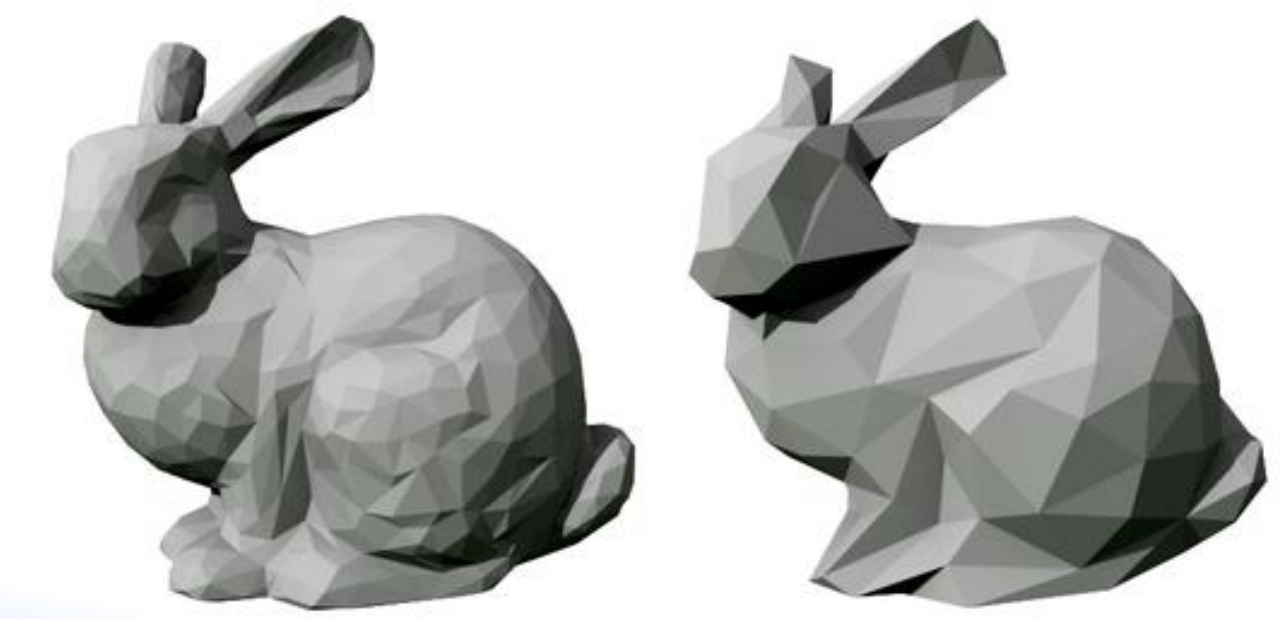
# statistics







computer graphics



HCI

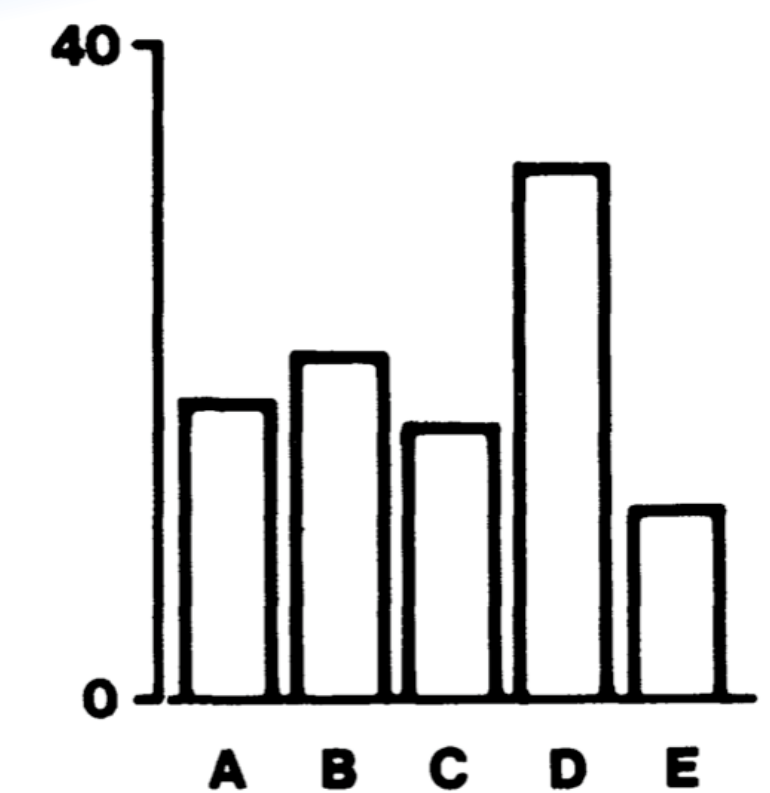
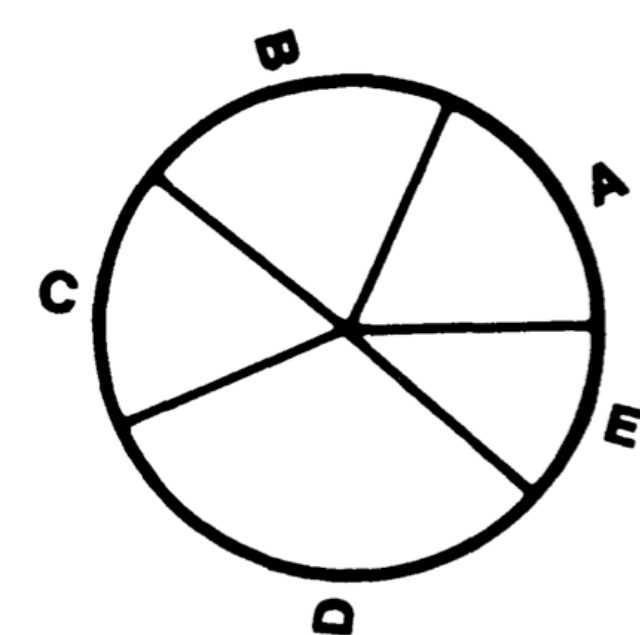
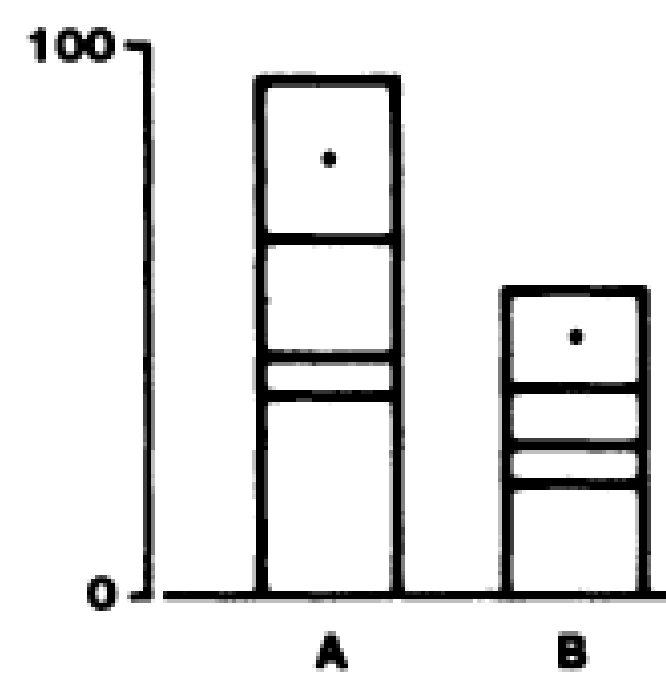
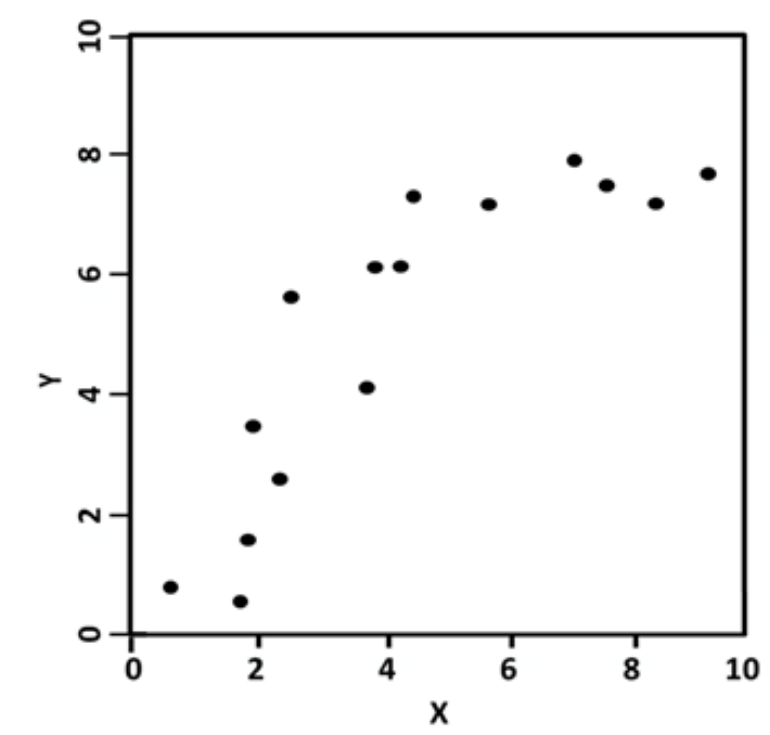
design

visualization

psychology

art

statistics





Ok, but why do we need  
visualization?



LSST



Financial Markets



# CHALLENGES:

- Scalability
- Complexity reduction
- Humans in-the-loop  
→ human cognition and memory is limited

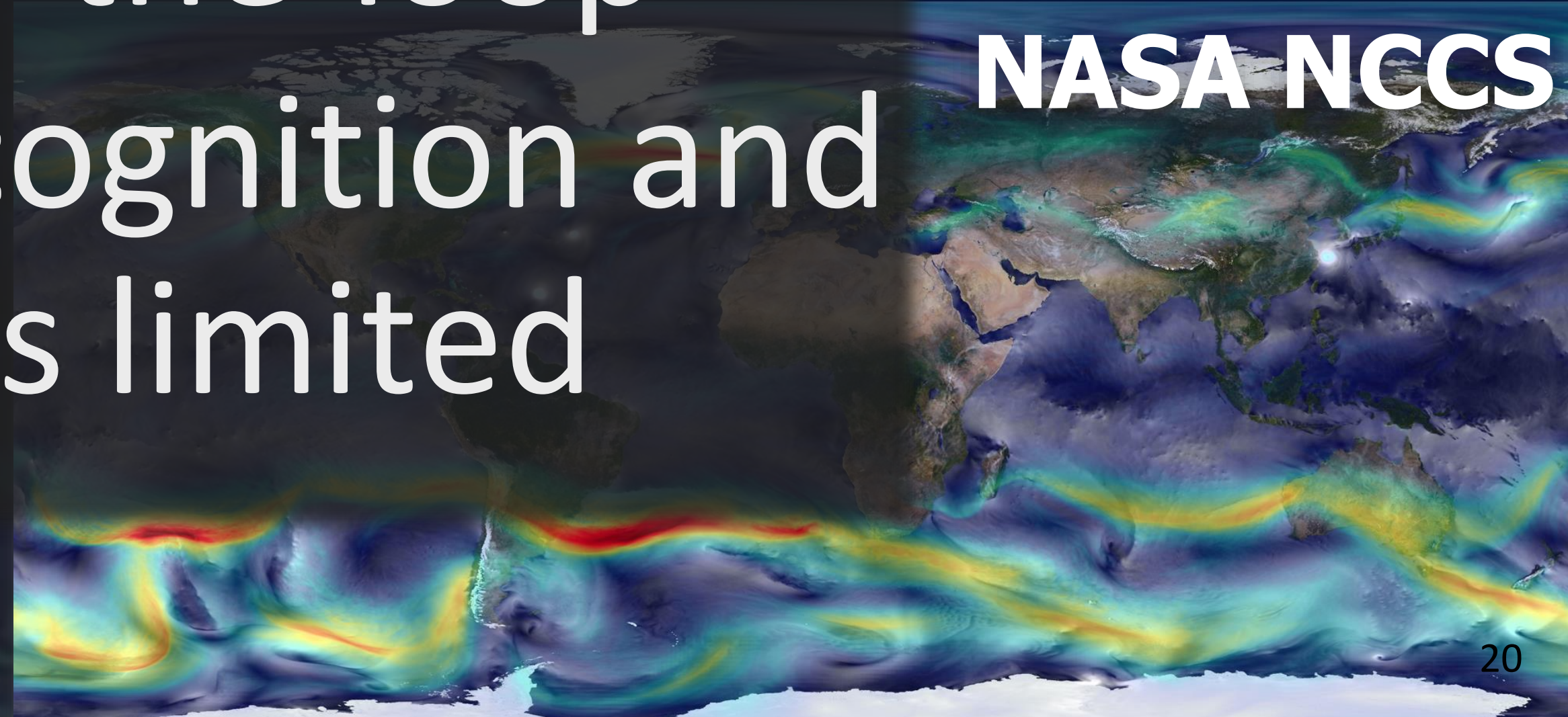


LHC

GenBank



NASA NCCS



ngrams





q t f j n i x i g j u n a s b b t g r  
k c l b v t x j x z x m x r g k l x  
q h m z y w t e y j w n o



q t f j n i x i g j u n a s b b t g r  
k c l b v t x j x z x m x r g k l x  
q h m z y w t e y j w n o



q t f j n i x i g j u n a s b b t g r  
k c l b v t x j x z x m x r g k l x  
q h m z y w t e y j w n o



q t f j n i x i g j u n a s b b t g r  
k c l b v t x j x z x m x r g k l x  
q h m z y w t e y j w n o



# “change blindness”



<https://www.youtube.com/watch?v=FW5xSQsspiQ>



# The "Door" Study

from Simons & Levin (1998)



# “change blindness”



<https://www.youtube.com/watch?v=FW5xSQsspiQ>



Again, why do we need  
visualization?

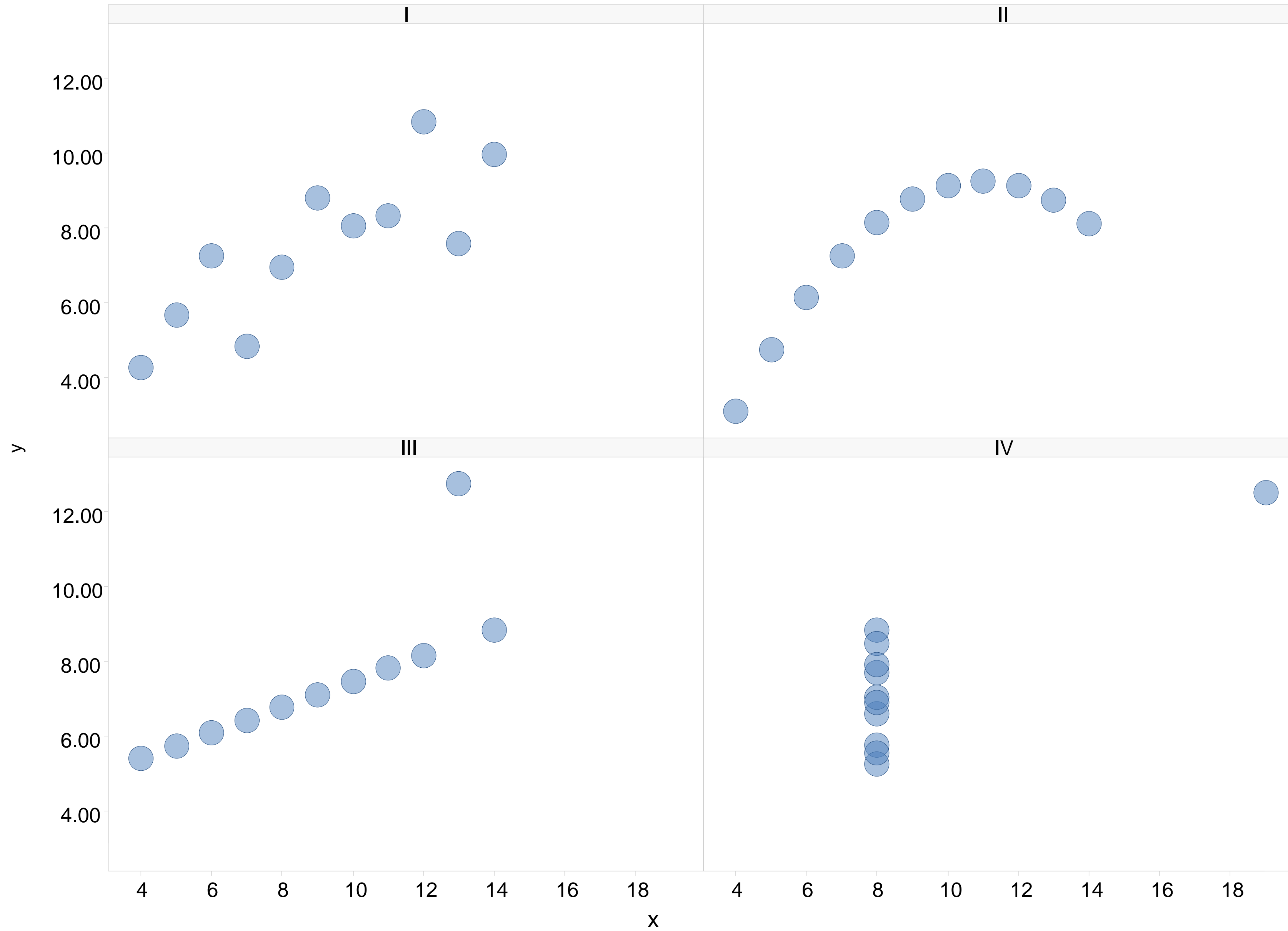


I		II		III		IV	
x	y	x	y	x	y	x	y
10.00	8.04	10.00	9.14	10.00	7.46	8.00	6.58
8.00	6.95	8.00	8.14	8.00	6.77	8.00	5.76
13.00	7.58	13.00	8.74	13.00	12.74	8.00	7.71
9.00	8.81	9.00	8.77	9.00	7.11	8.00	8.84
11.00	8.33	11.00	9.26	11.00	7.81	8.00	8.47
14.00	9.96	14.00	8.10	14.00	8.84	8.00	7.04
6.00	7.24	6.00	6.13	6.00	6.08	8.00	5.25
4.00	4.26	4.00	3.10	4.00	5.39	19.00	12.50
12.00	10.84	12.00	9.13	12.00	8.15	8.00	5.56
7.00	4.82	7.00	7.26	7.00	6.42	8.00	7.91
5.00	5.68	5.00	4.74	5.00	5.73	8.00	6.89



	<b>Value</b>	<b>Equality</b>
<b>X Mean</b>	9	=
<b>Y Mean</b>	7.50	.00
<b>X Variance</b>	11	=
<b>Y Variance</b>	4.12	.00
<b>Correlation</b>	0.816	.000
<b>Linear regression line</b>	$y = 3.00 + 0.500x$	.00 and .000



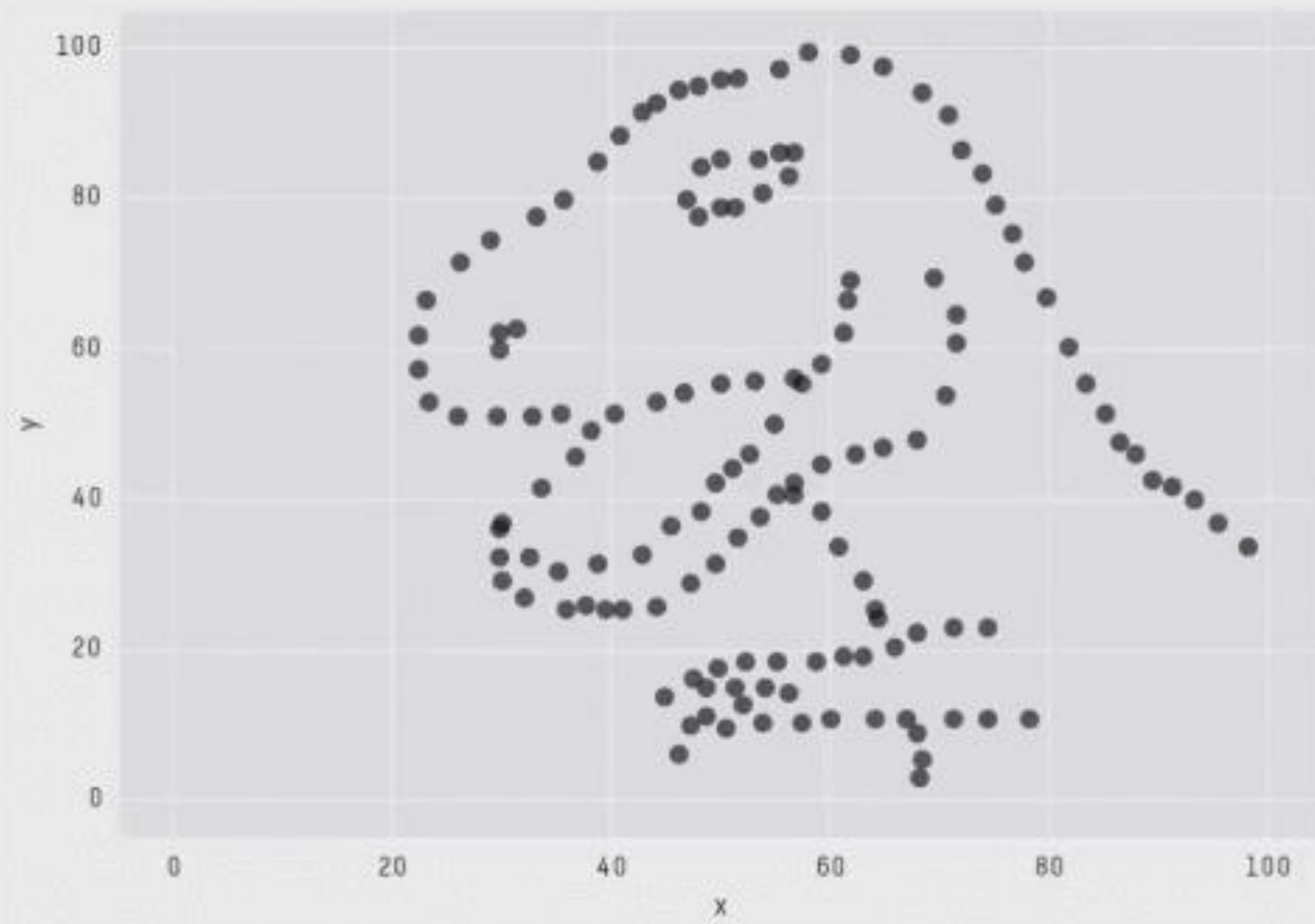




There are three types of lies: lies,  
damned lies, and statistics

Unknown author, popularized by Mark Twain





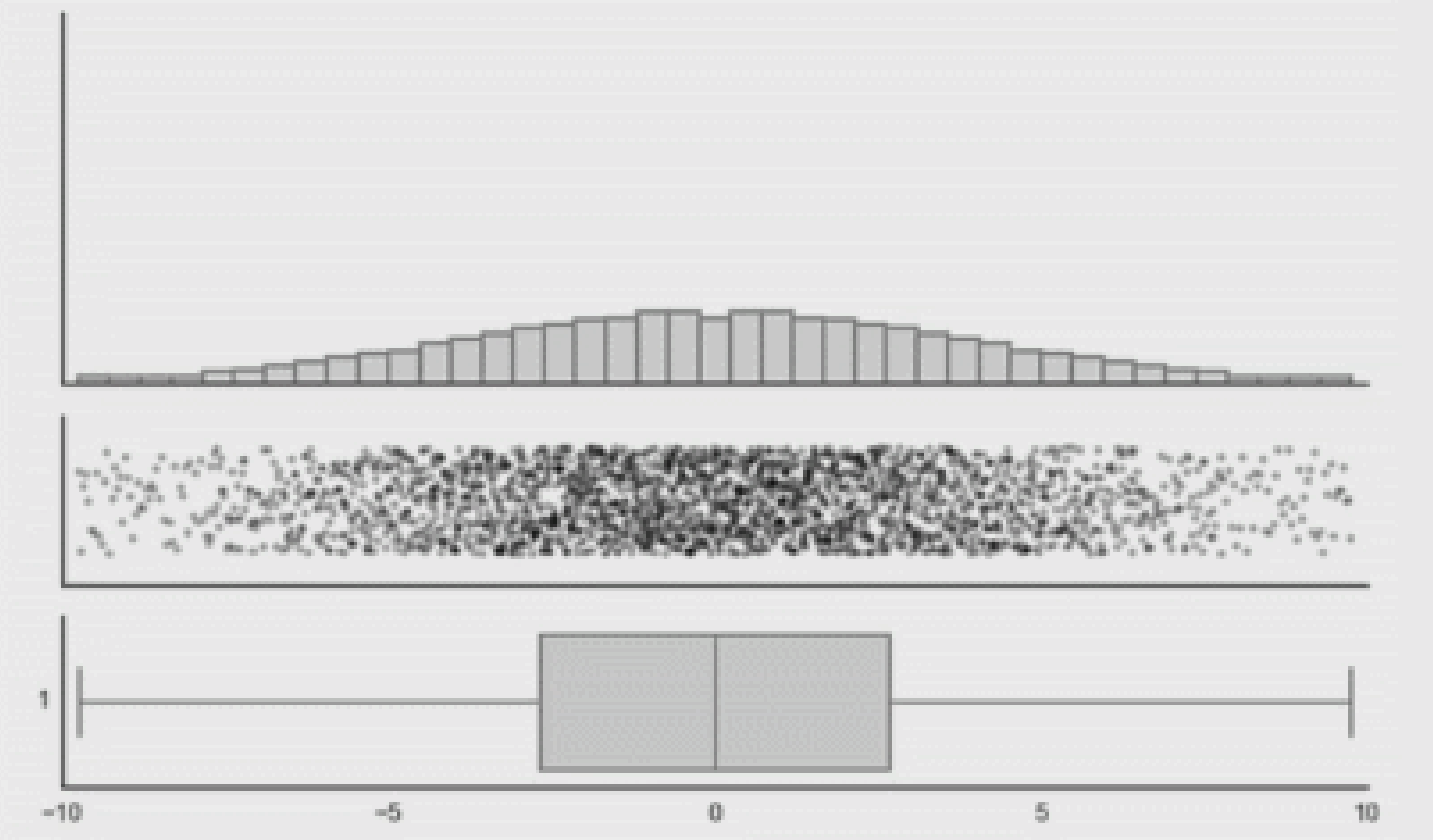
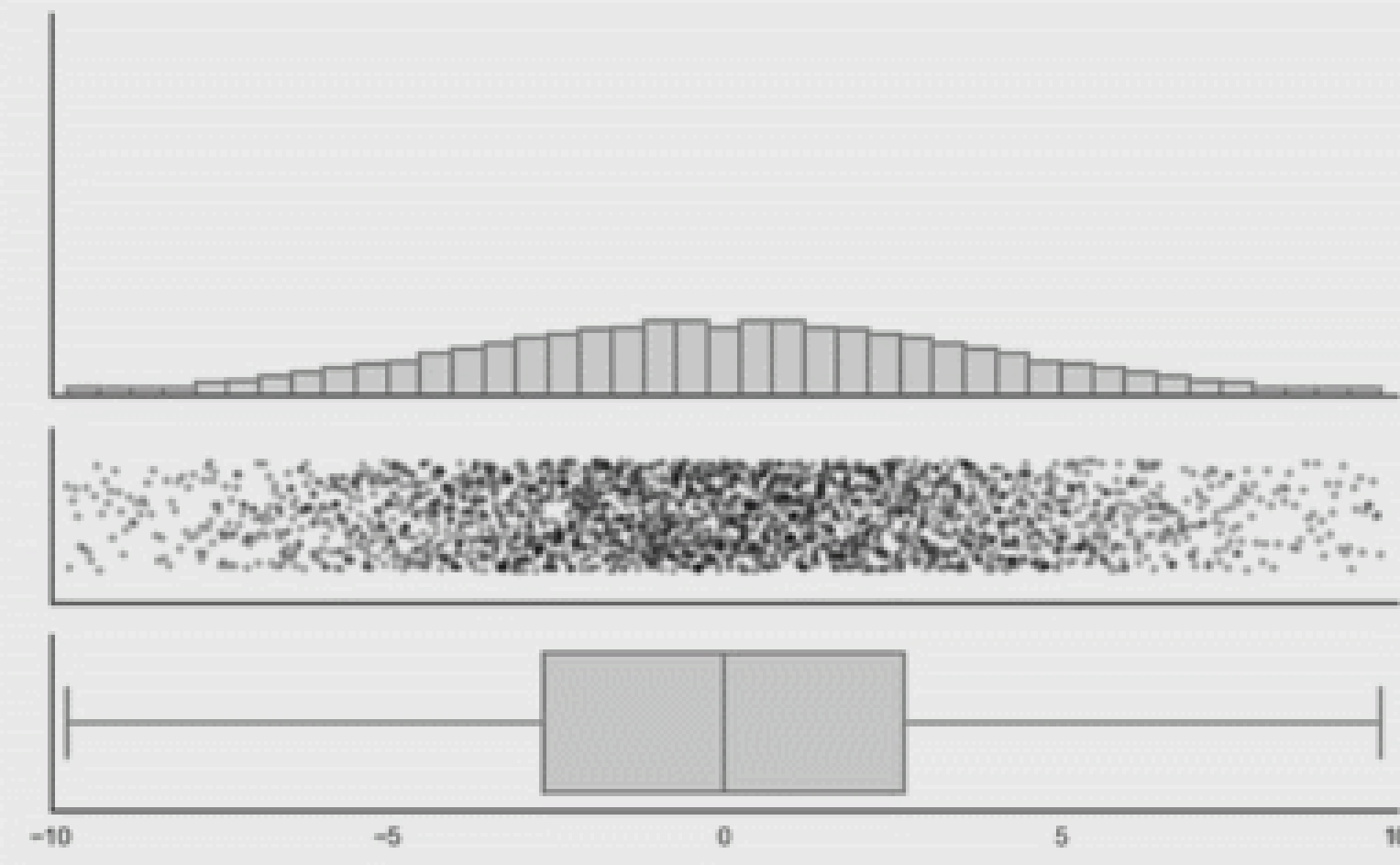
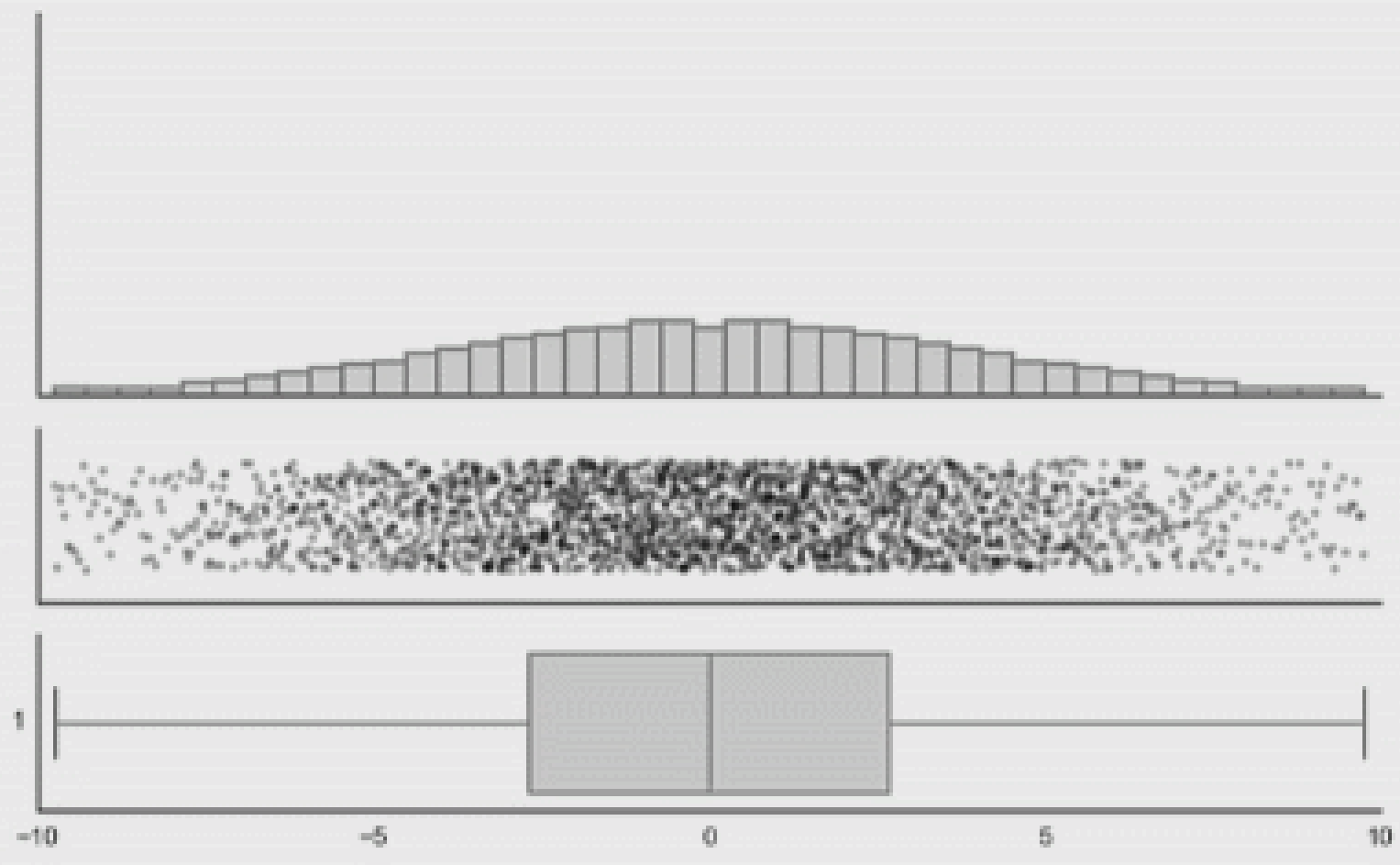
X Mean: 54.2659224  
Y Mean: 47.8313999  
X SD : 16.7649829  
Y SD : 26.9342120  
Corr. : -0.0642526



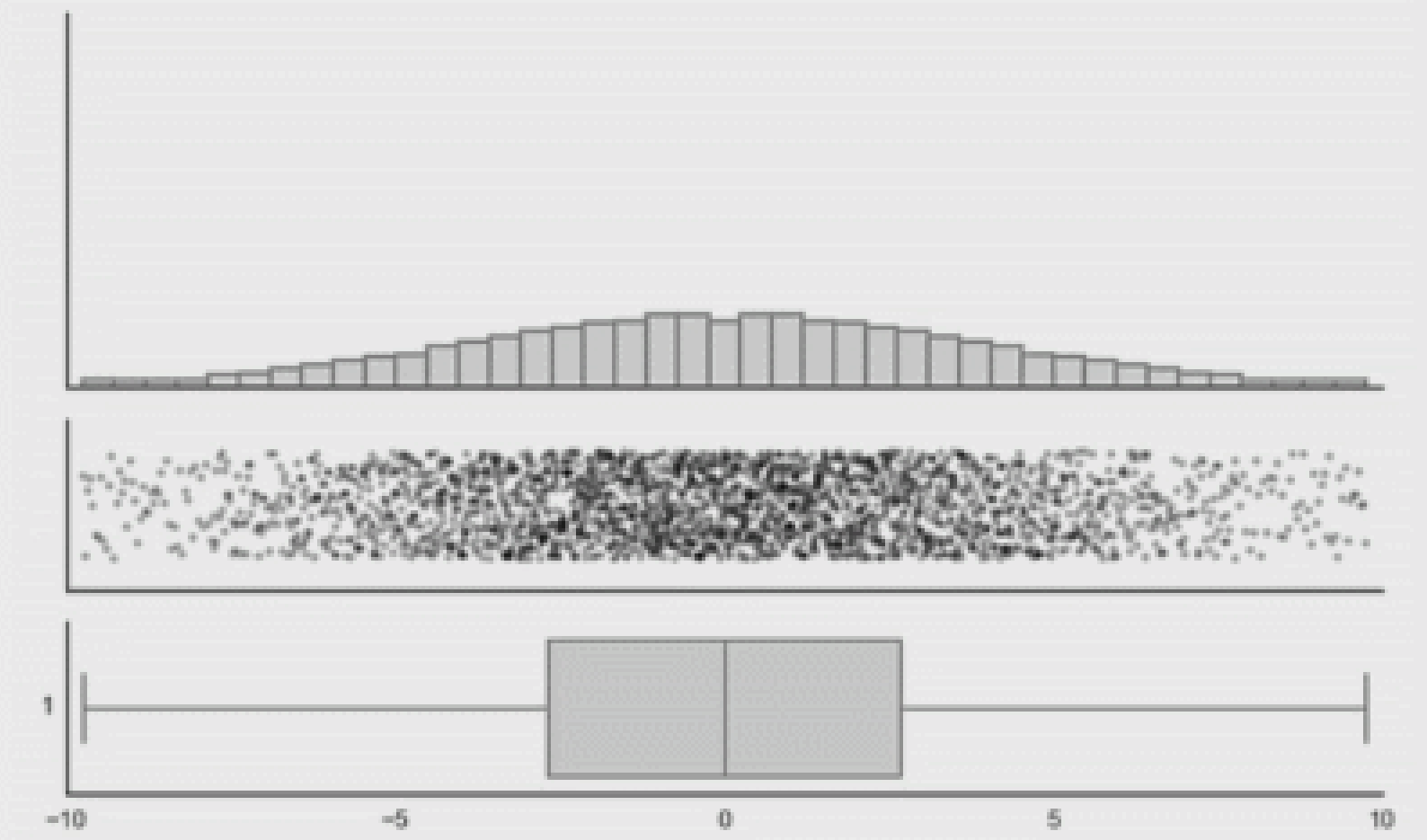
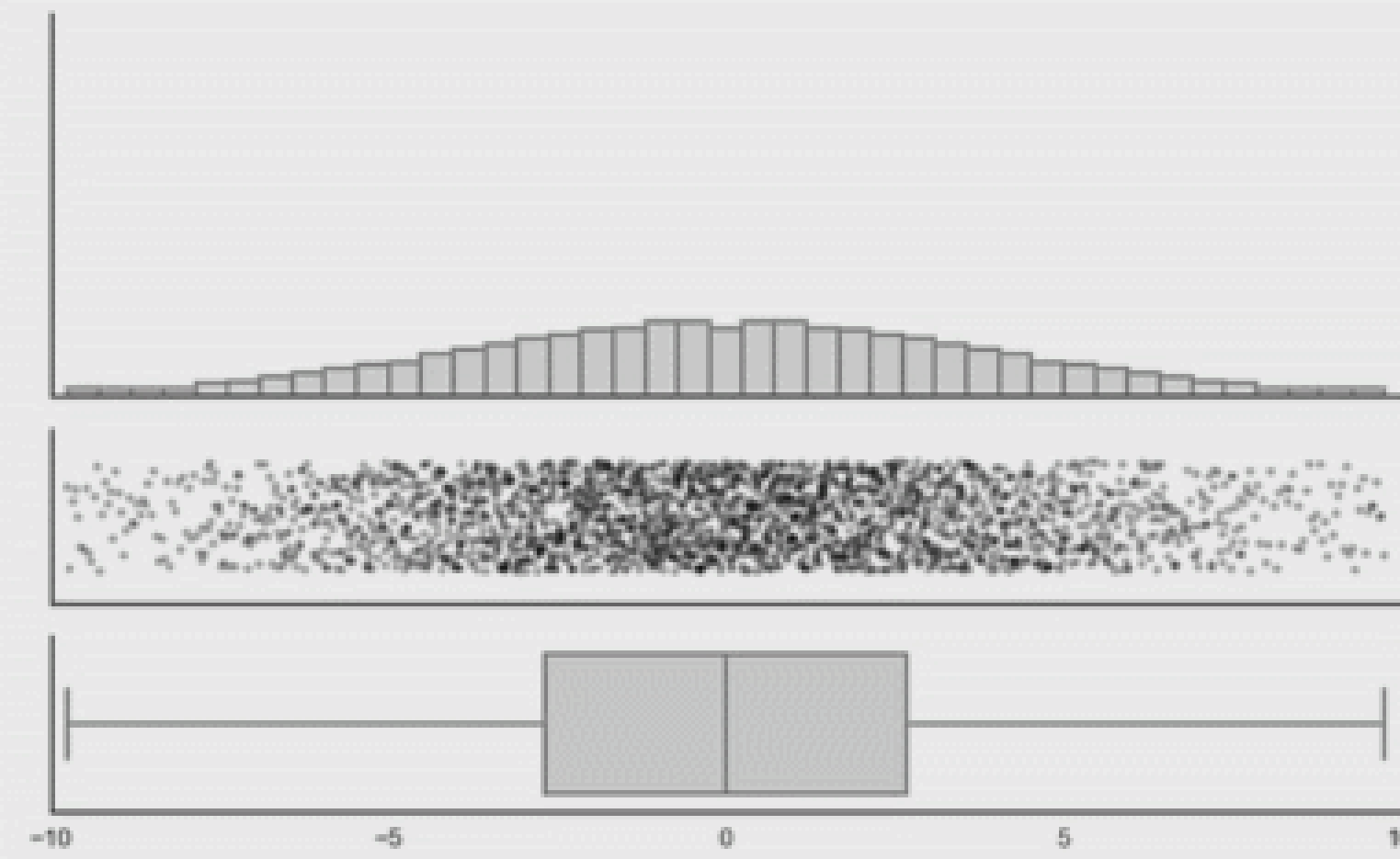
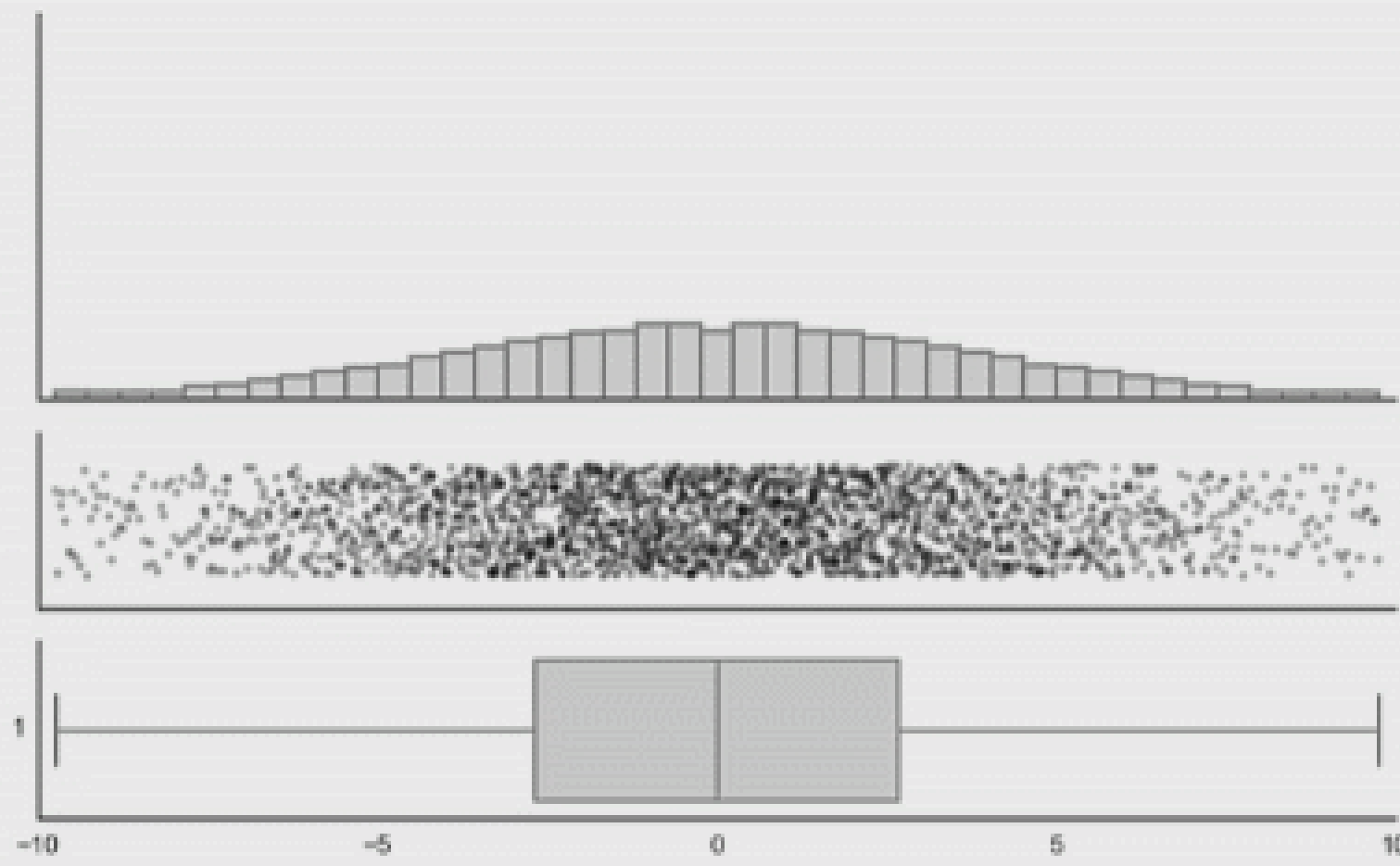
No catalogue of techniques can convey a willingness to look for what can be seen, whether or not anticipated. Yet this is at the heart of exploratory data analysis. ... the picture-examining eye is the best finder we have of the wholly unanticipated.

– Tukey, 1980

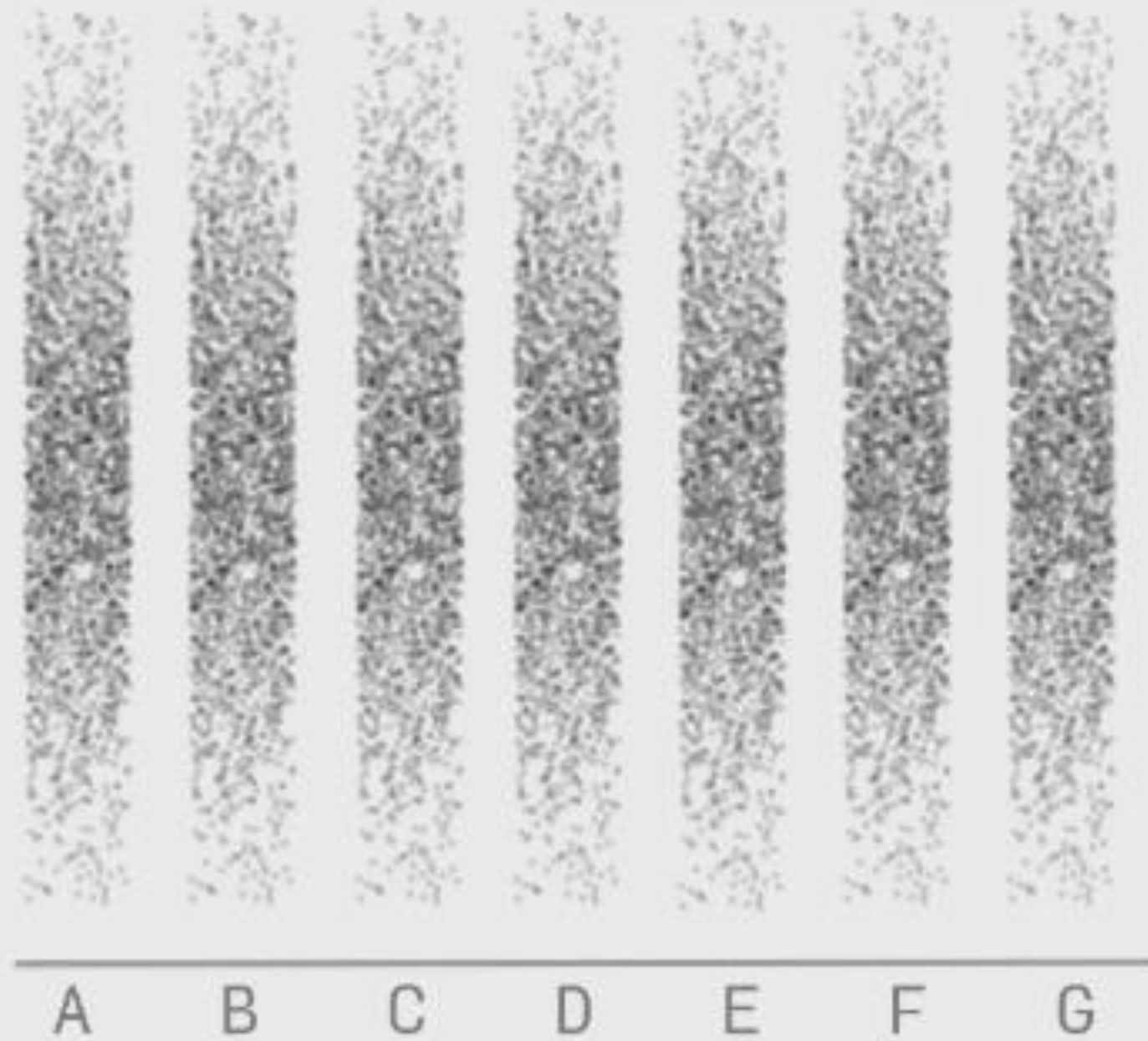




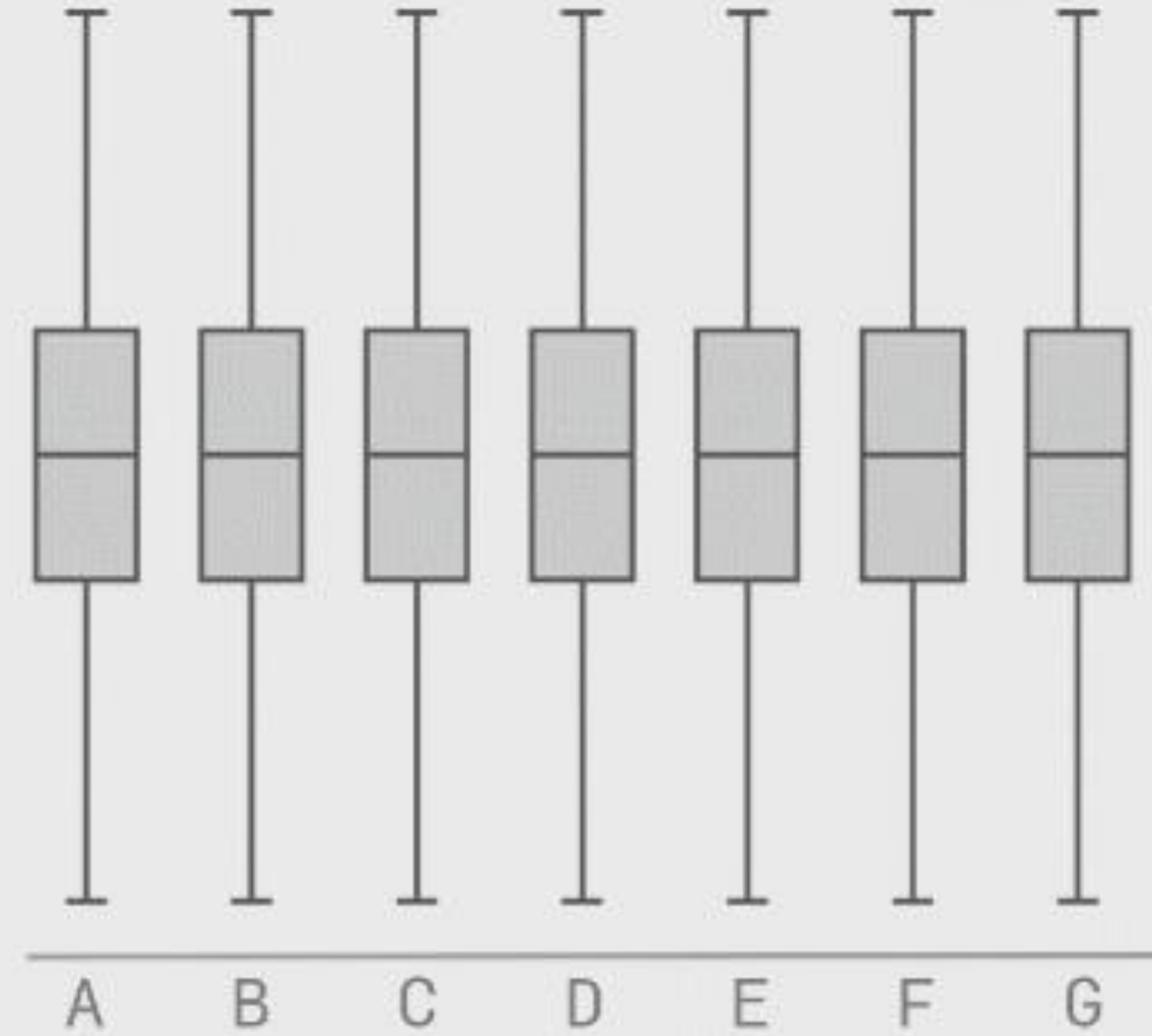




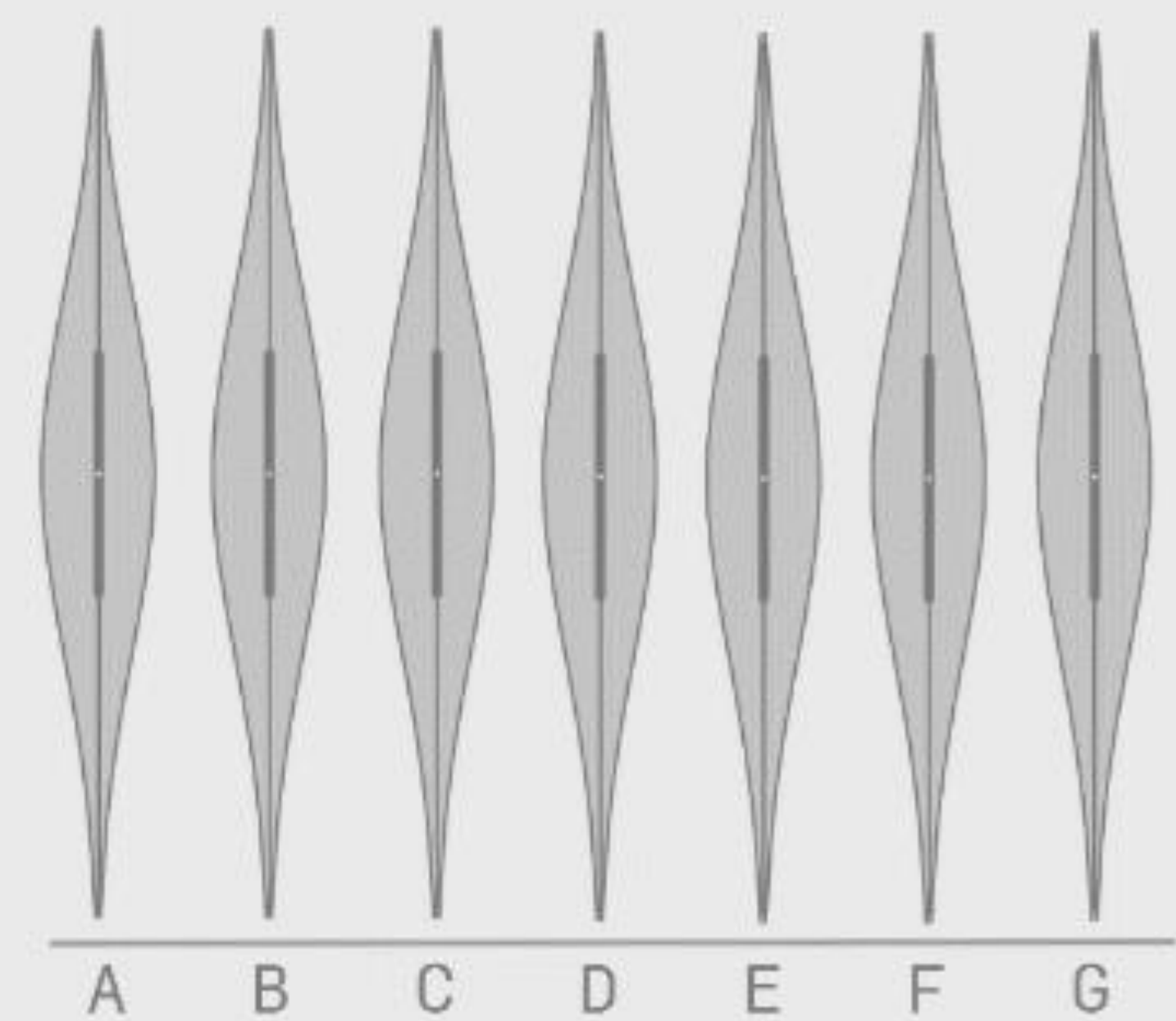
**Raw Data**



**Box-plot of the Data**



**Violin-plot of the Data**





Ok, but why do we need  
visualization?



# Why visualize your data?

- Help cognition
- Expand memory
- Generate hypotheses
- Answer questions
- Make decisions
- Find patterns
- Record
- Clarify
- Communicate
- Inspire



# In-class sketching — table tents

*20 min*

[neu-ds-4200-f23.github.io/in-class/table-tents/](https://neu-ds-4200-f23.github.io/in-class/table-tents/)



# DESIGN RULES OF THUMB



# Design Rules of Thumb

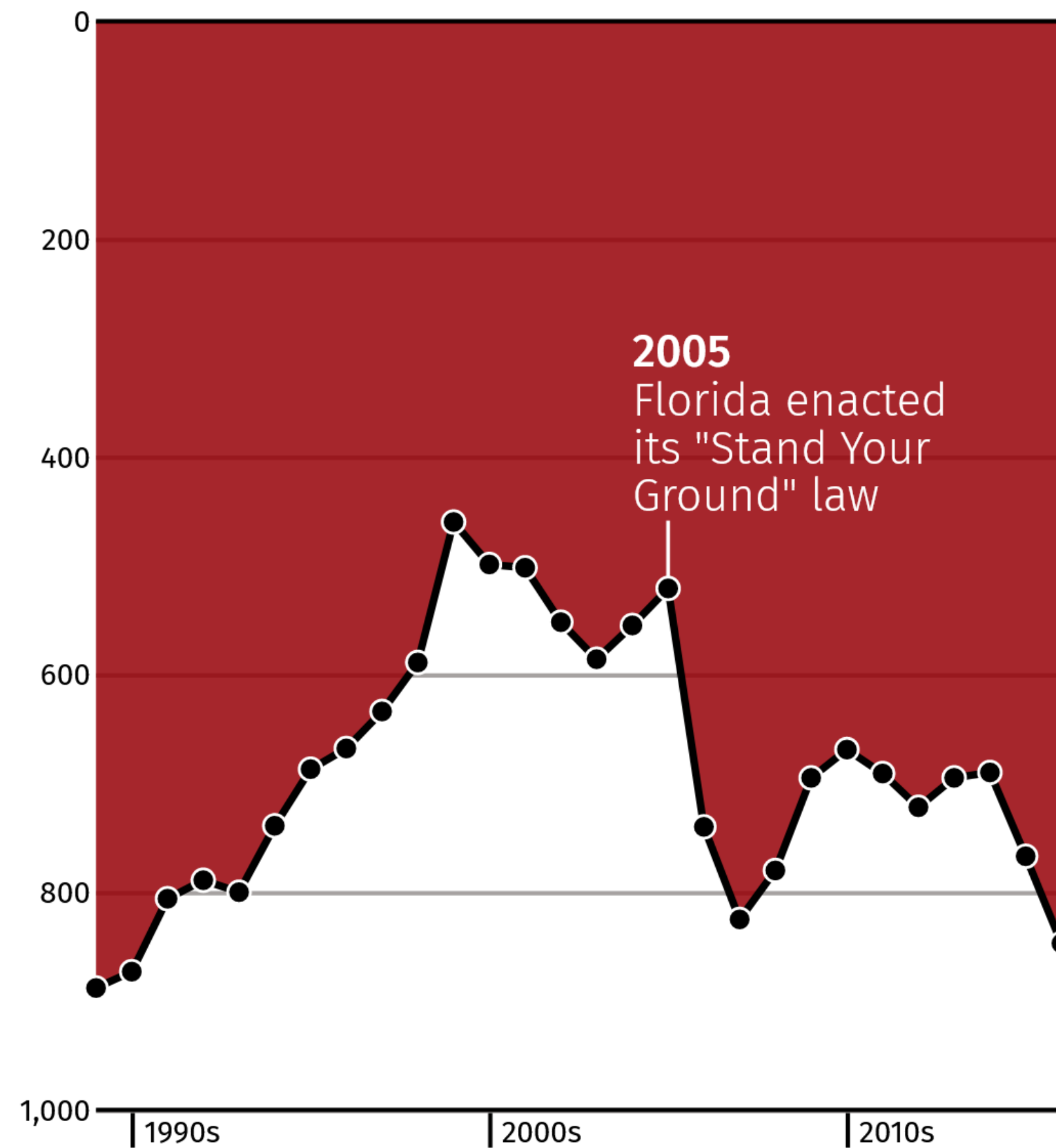
1. Function first, form next



# “Function first, form next”

## Gun deaths in Florida

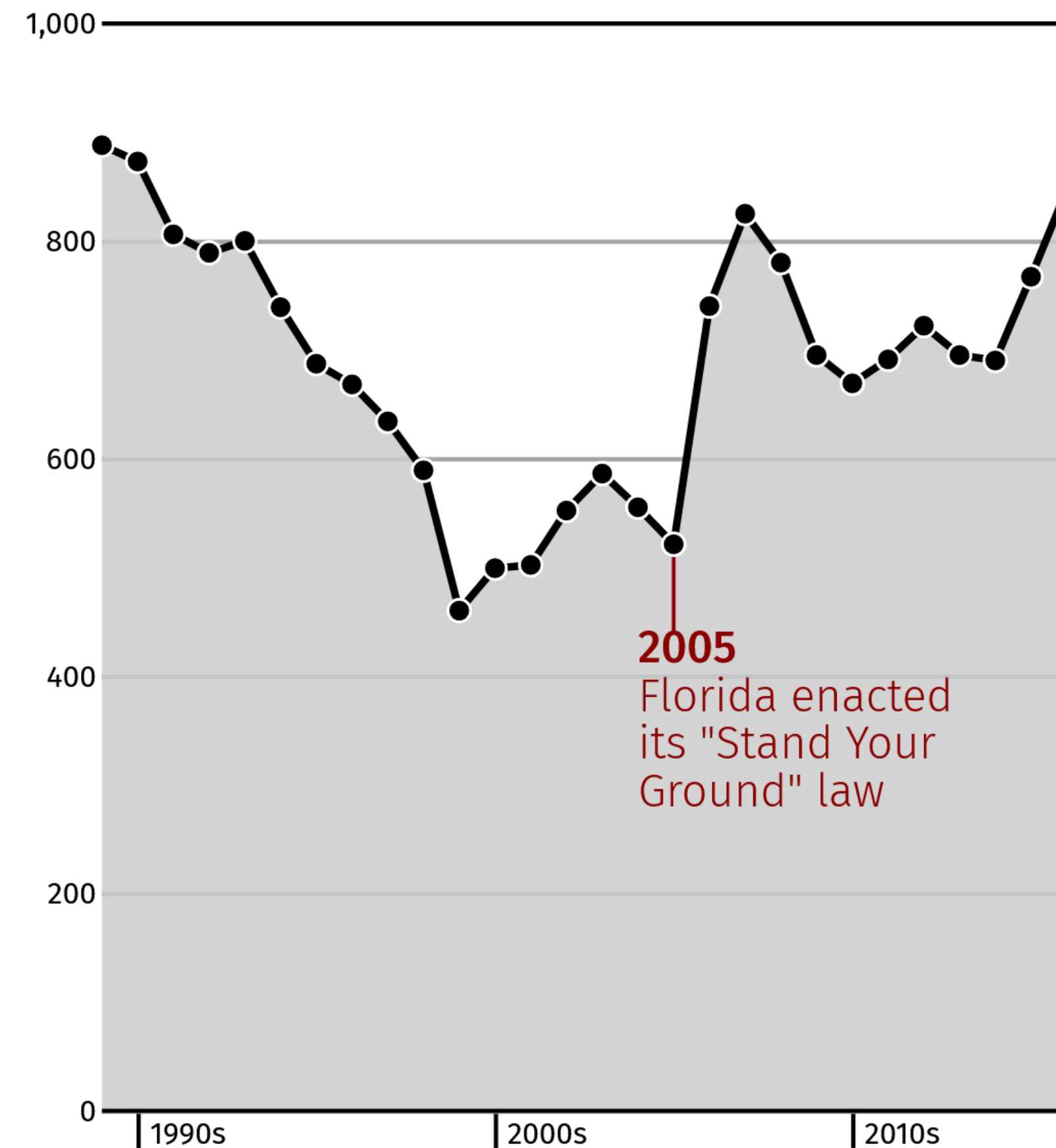
Number of murders committed using firearms



Source: Florida Department of Law Enforcement

## Gun deaths in Florida

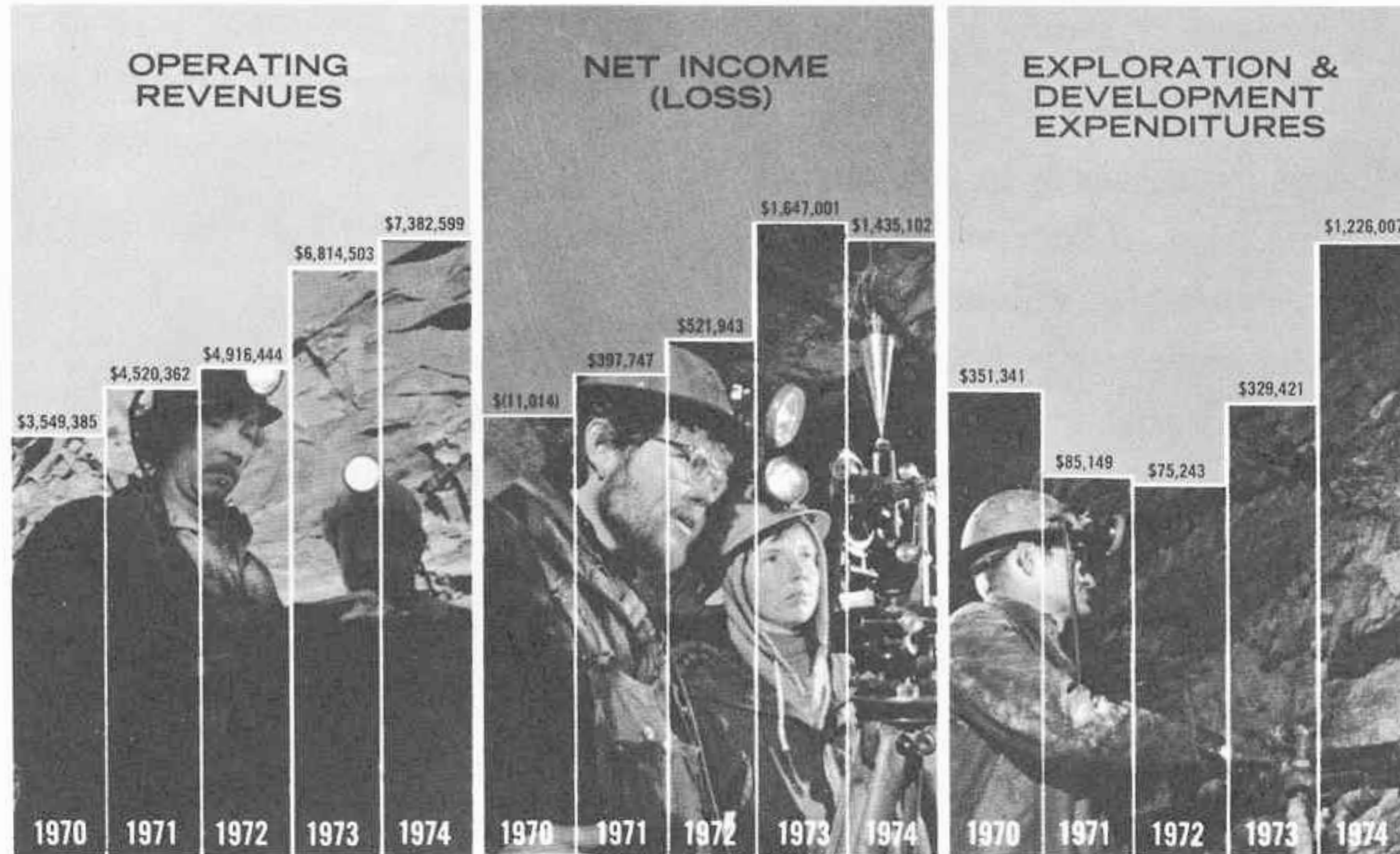
Number of murders committed using firearms



Source: Florida Department of Law Enforcement



# “Function first, form next



“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

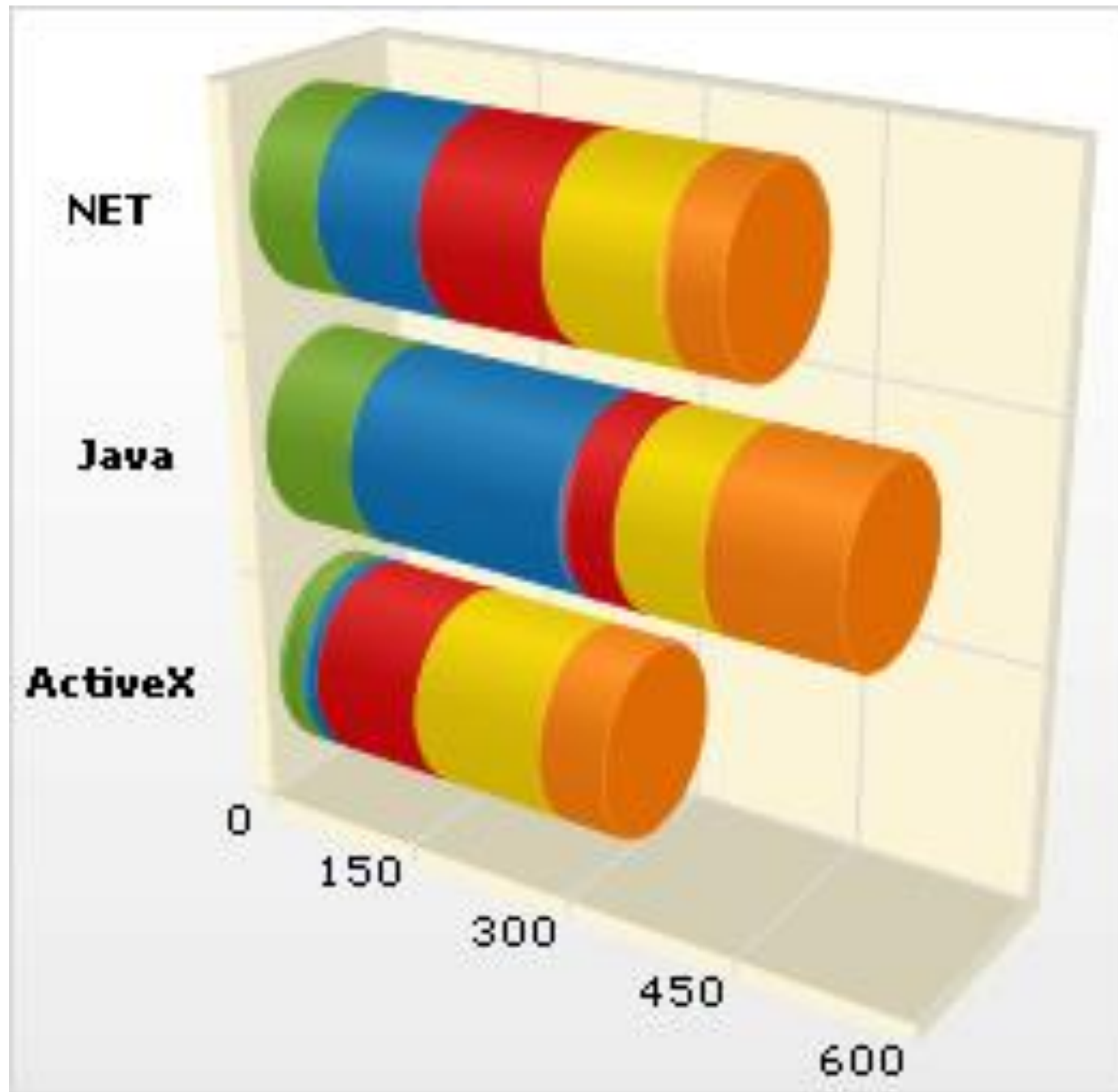


# Design Rules of Thumb

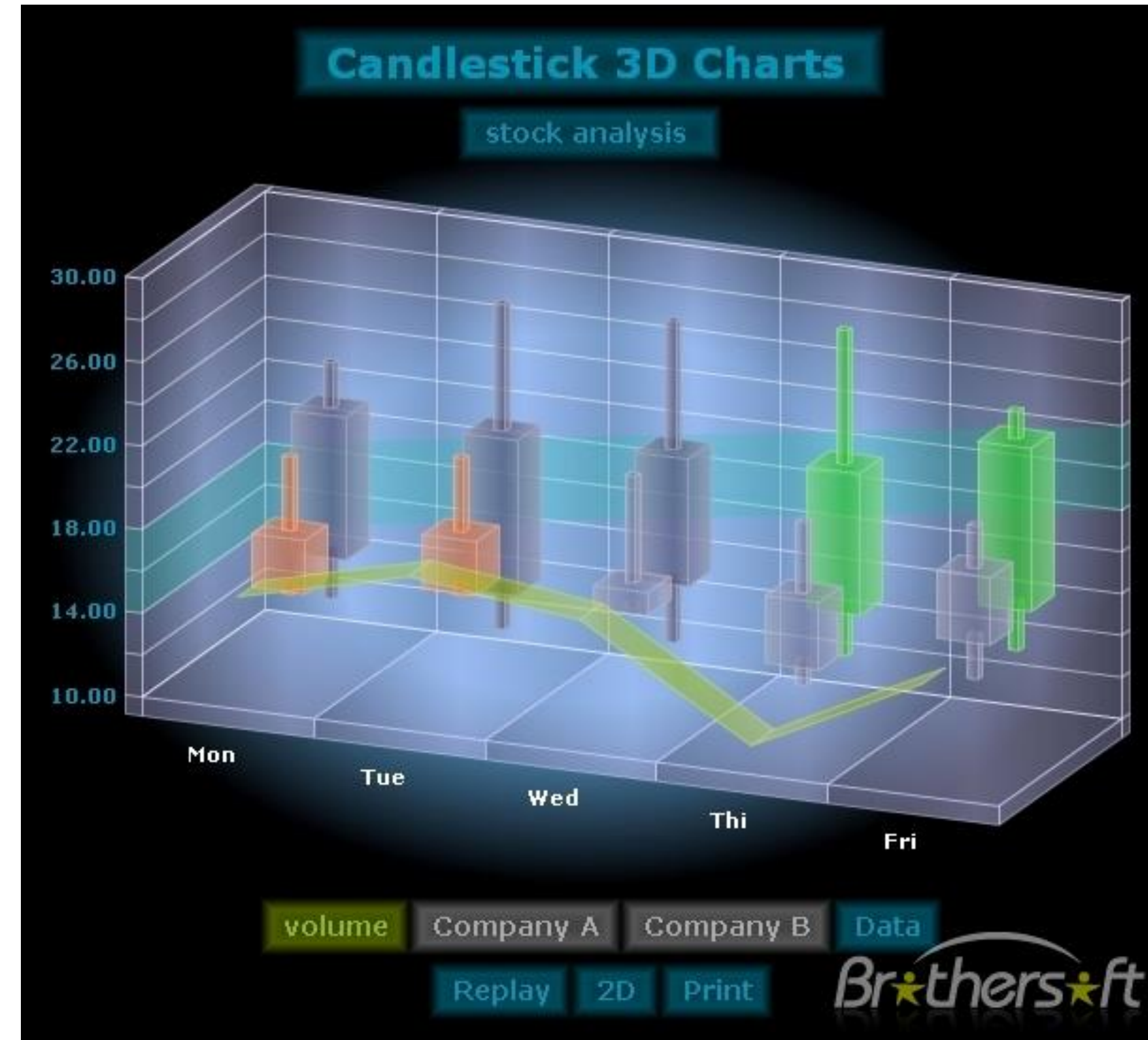
1. Function first, form next
2. No unjustified 3D



# “No Unjustified 3D”



<http://help.infragistics.com/Help/Doc/WinForms/2014.2/CLR4.0/html/Images/Chart Bar Chart 03.png>



[http://img.brothersoft.com/screenshots/softimage/0/3d\\_charts-171418-1269568478.jpeg](http://img.brothersoft.com/screenshots/softimage/0/3d_charts-171418-1269568478.jpeg)

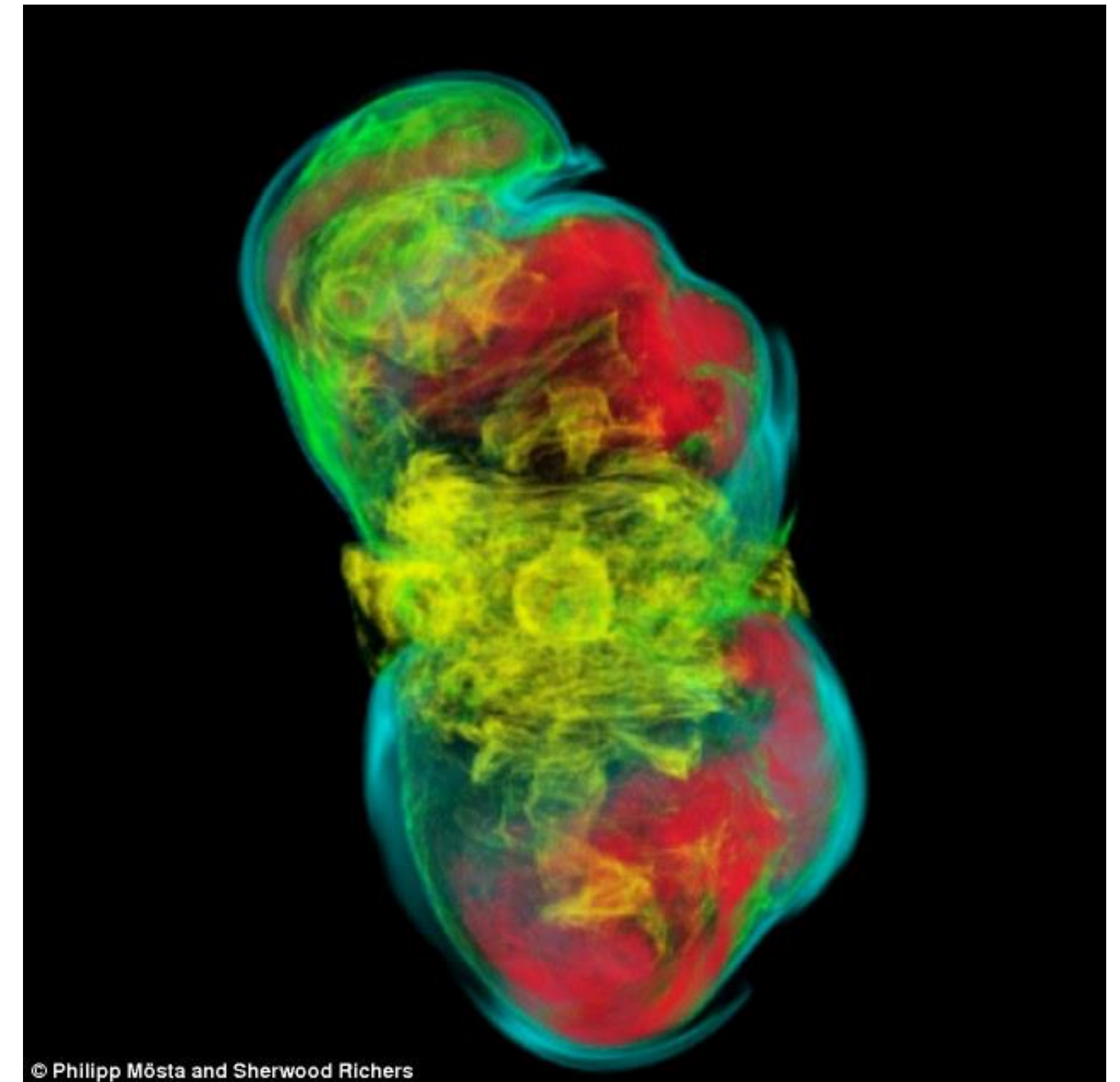
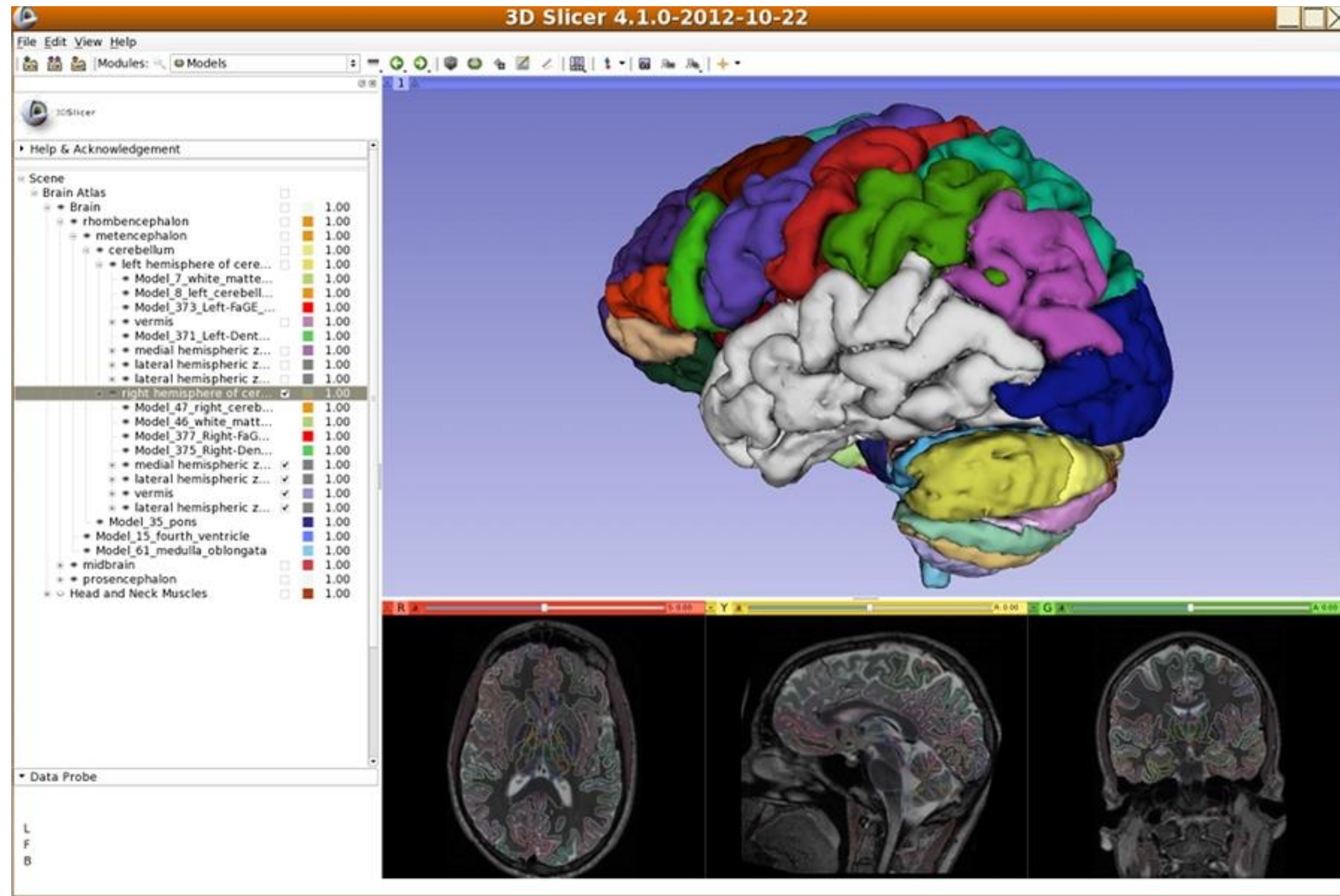


# “No Unjustified 3D”





# “No Unjustified 3D”



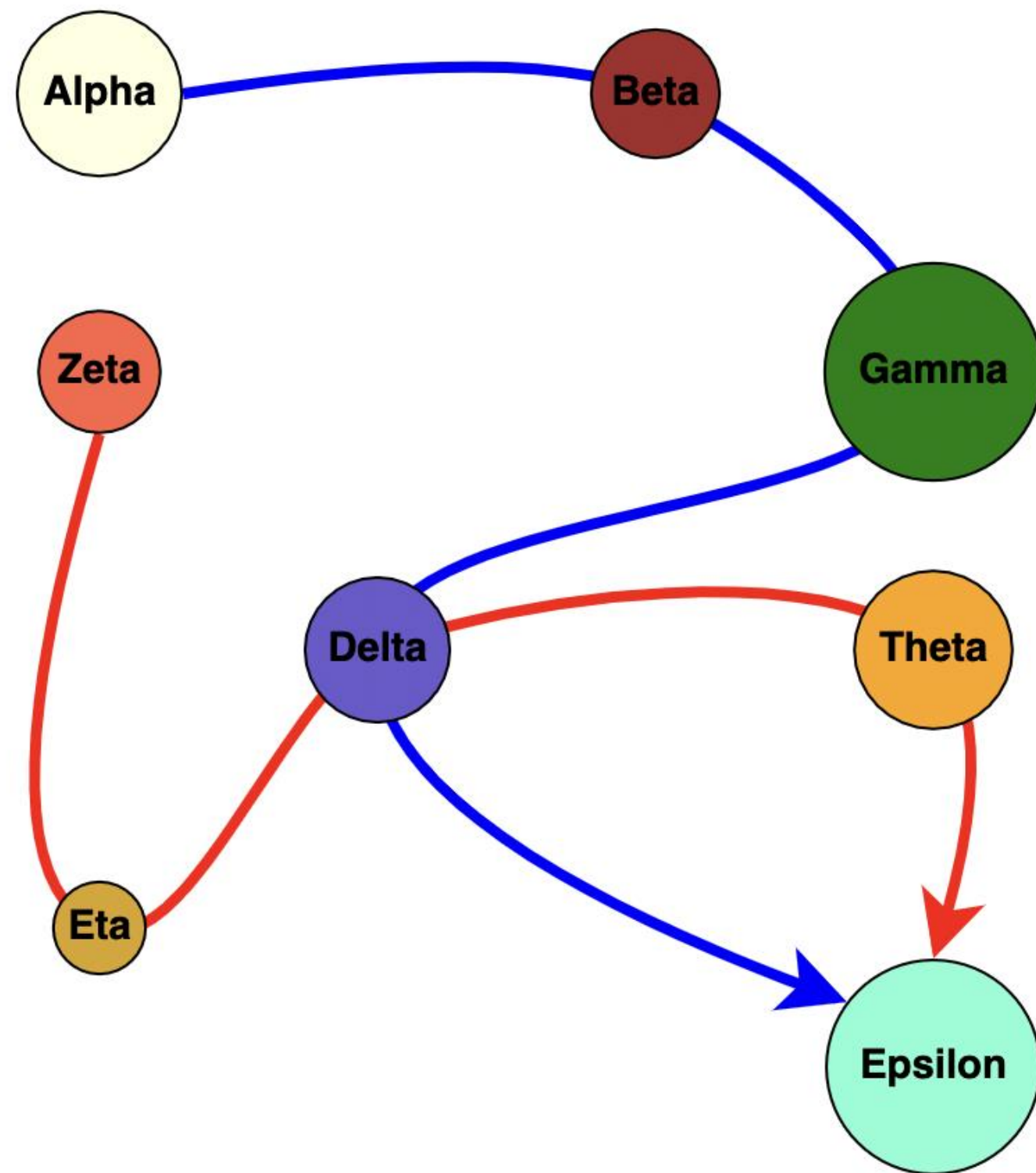


# Design Rules of Thumb

1. Function first, form next
2. No unjustified 3D
3. No unjustified 2D



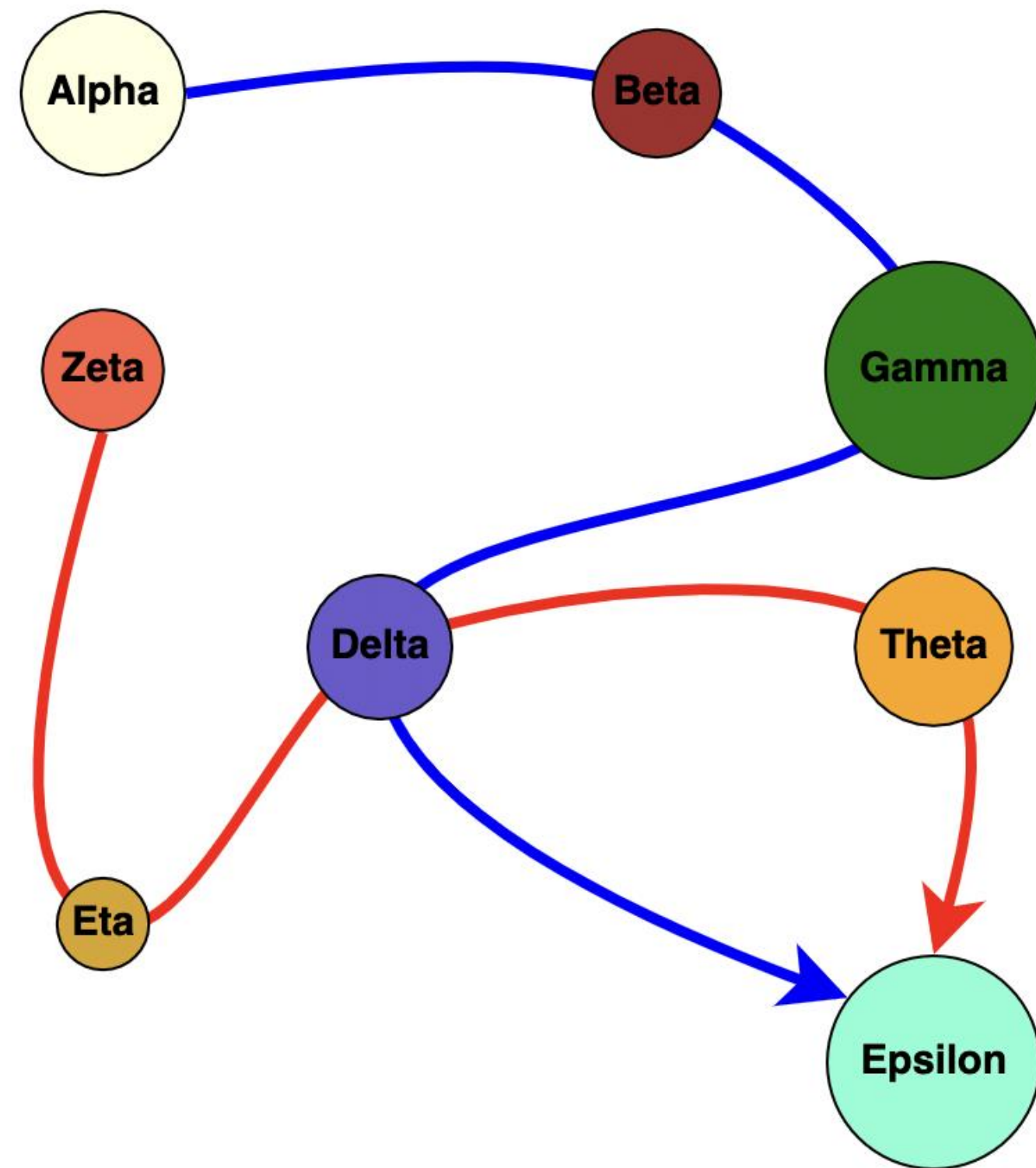
# “No Unjustified 2D”



Task: What color is Delta?



# “No Unjustified 2D”



Task: What color is Delta?

Node	Color
Alpha	White
Beta	Maroon
Delta	Purple
Epsilon	Teal
Eta	Mustard Yellow
Gamma	Green
Theta	Orange
Zeta	Pink

If the task doesn't need a 2D visualization, then don't use one.

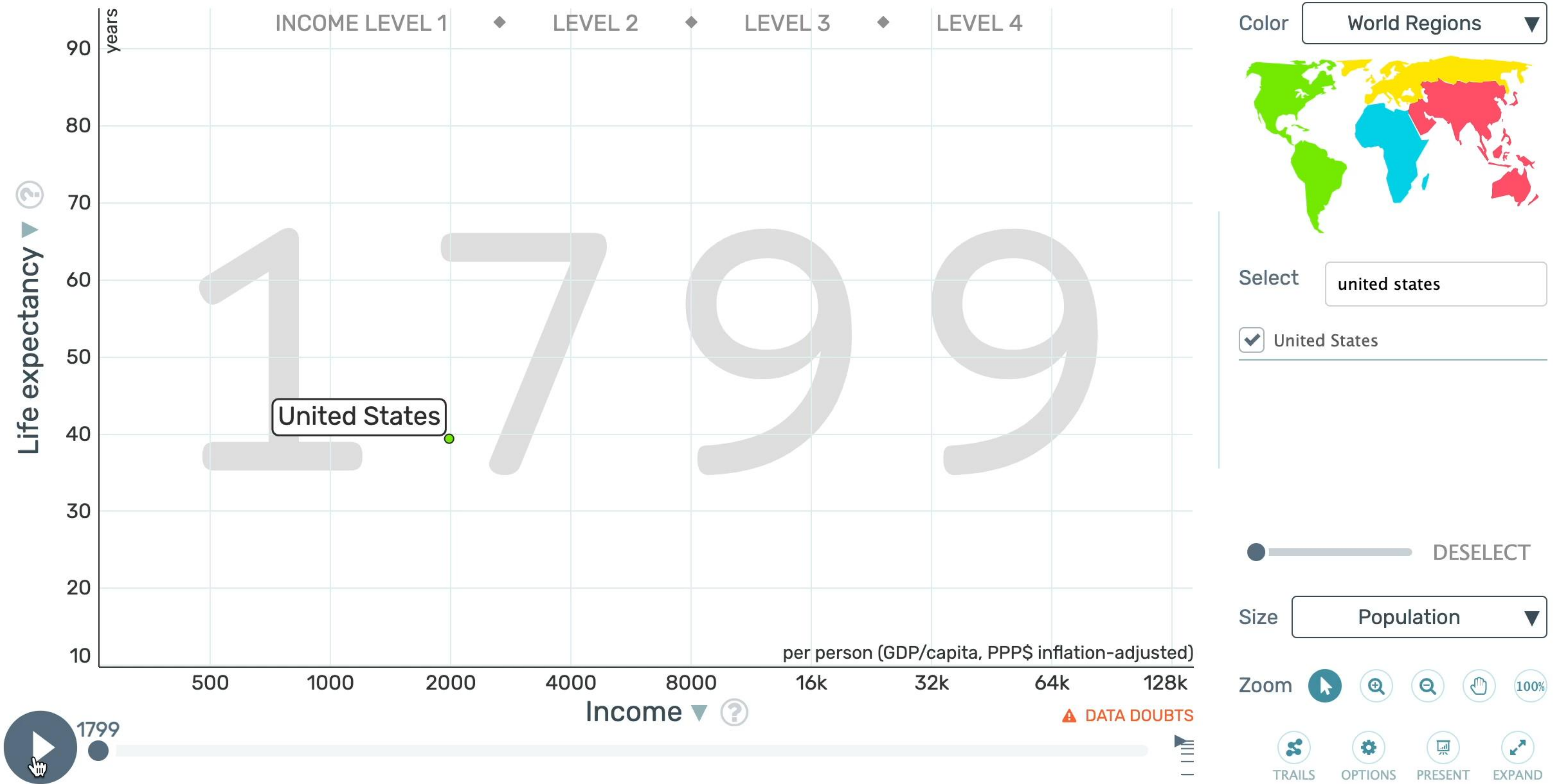


# Design Rules of Thumb

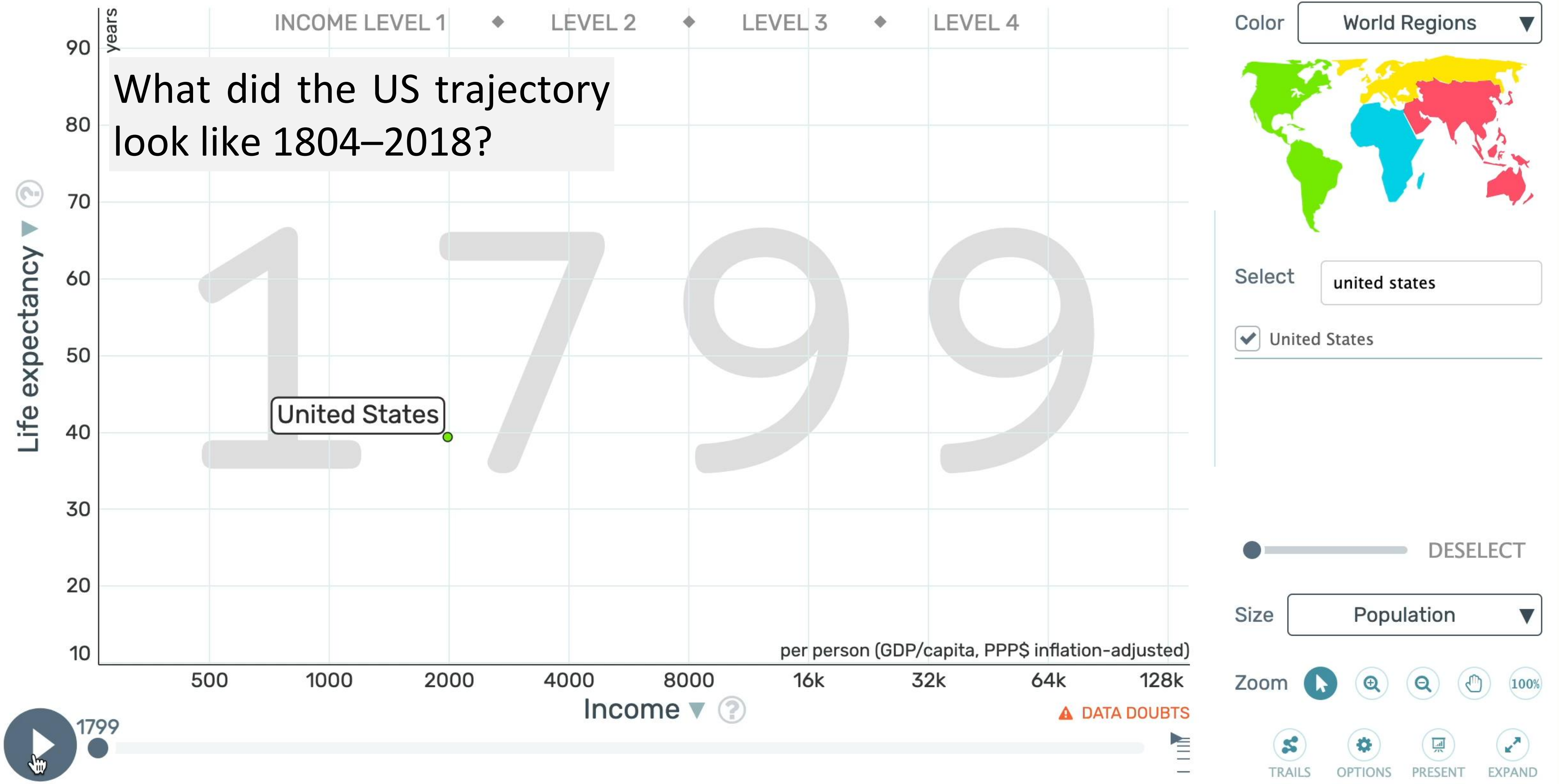
1. Function first, form next
2. No unjustified 3D
3. No unjustified 2D
4. Eyes beat memory



# “Eyes Beat Memory”

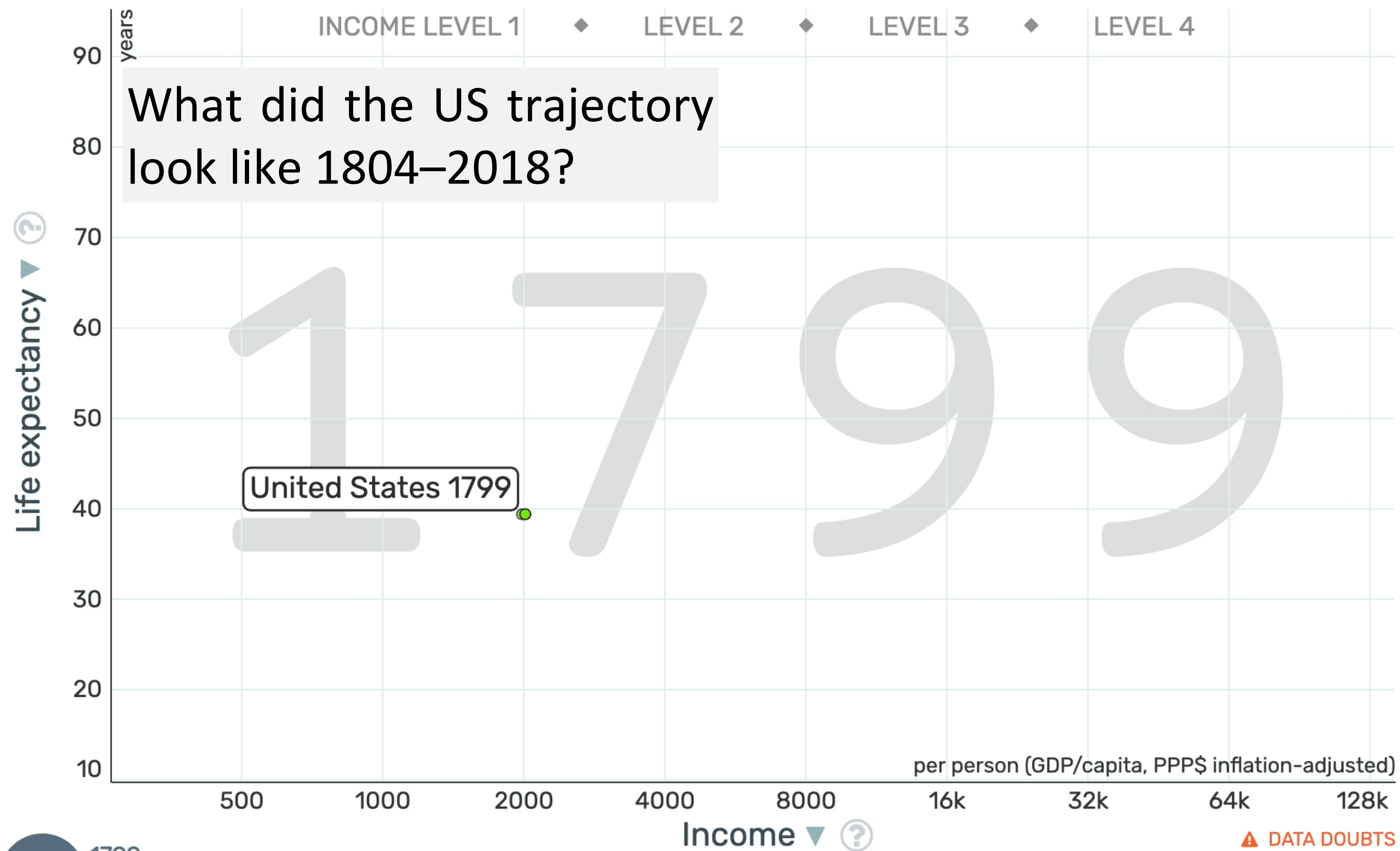


# “Eyes Beat Memory”





# “Eyes Beat Memory”



Color: World Regions

Select: united states

United States

DESELECT

Size: Population

Zoom: [Icons for zoom in, zoom out, pan, 100%]

TRAILS OPTIONS PRESENT EXPAND

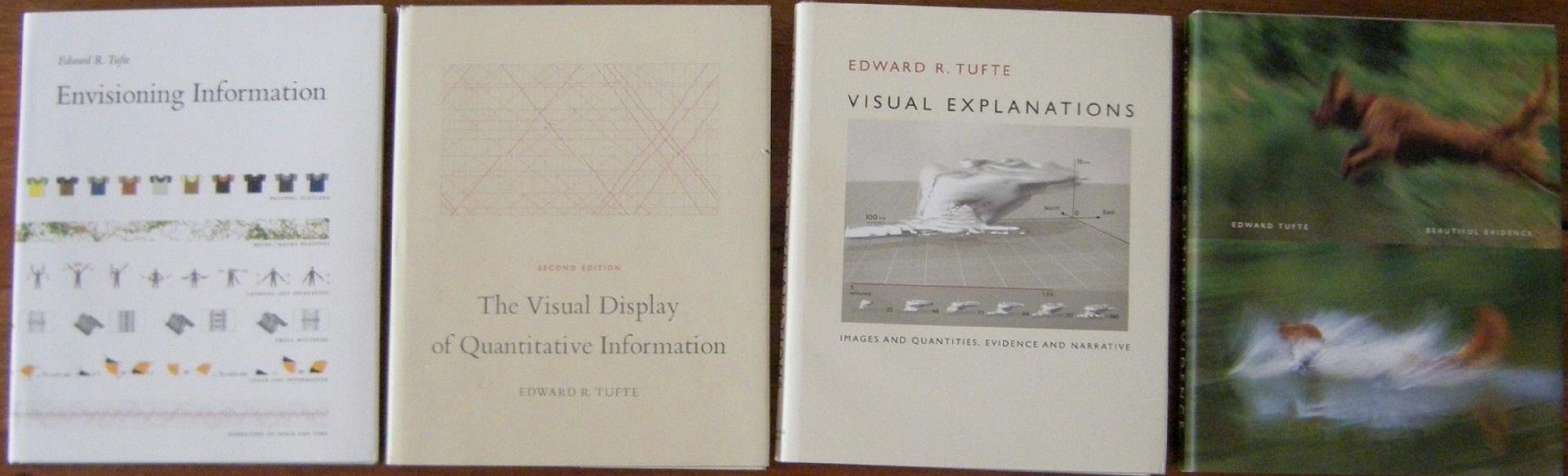
1799

# Design Rules of Thumb

1. Function first, form next
2. No unjustified 3D
3. No unjustified 2D
4. Eyes beat memory



# Edward Tufte



Tufte will be doing one of his one-day courses in Boston on Oct. 3, 4, 5 2023. \$240 for students includes these books. <https://www.edwardtufte.com/tufte/courses>



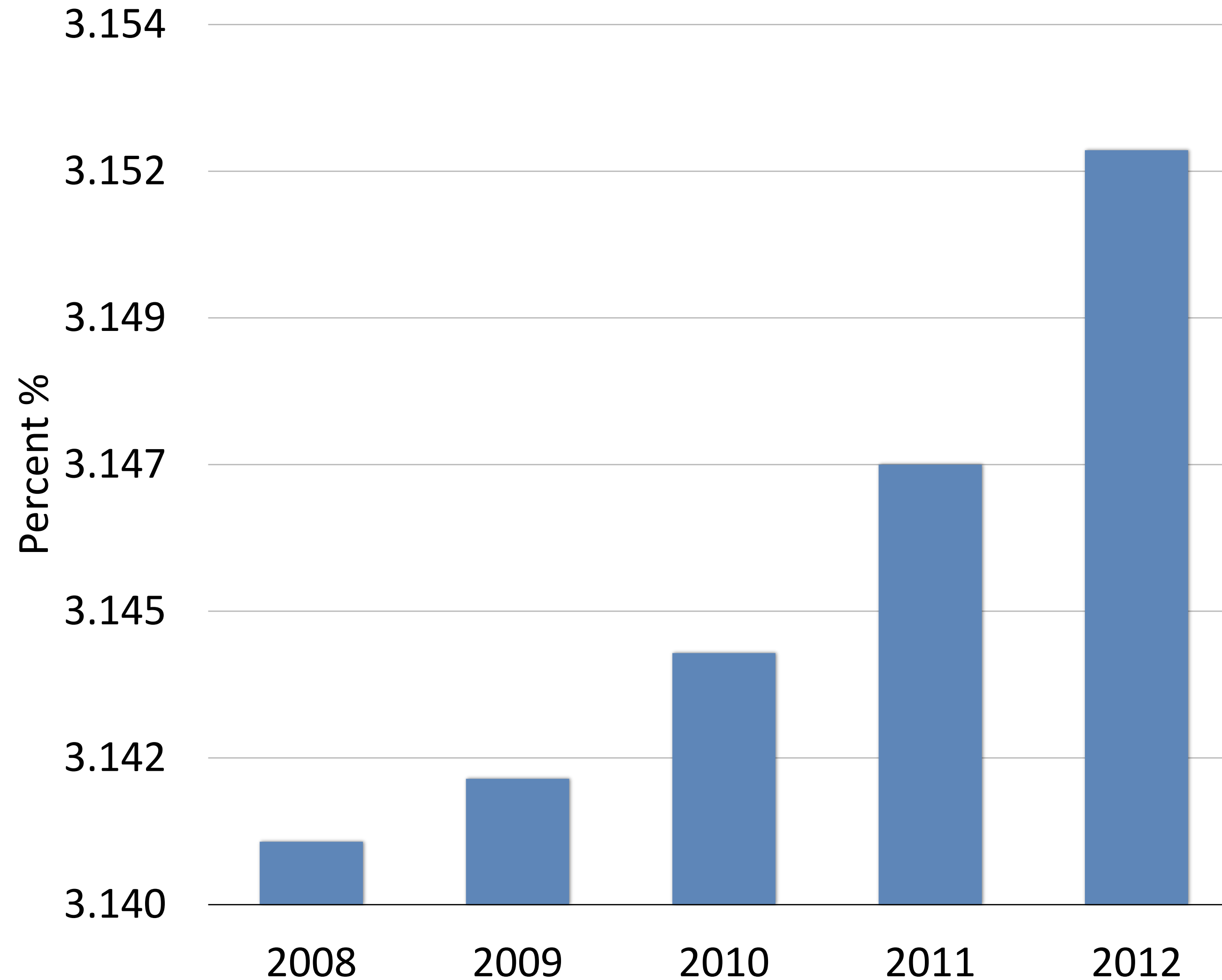
# “Graphical Integrity”

“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

*(Axes and axis labels, titles, annotations, legends, etc.)*



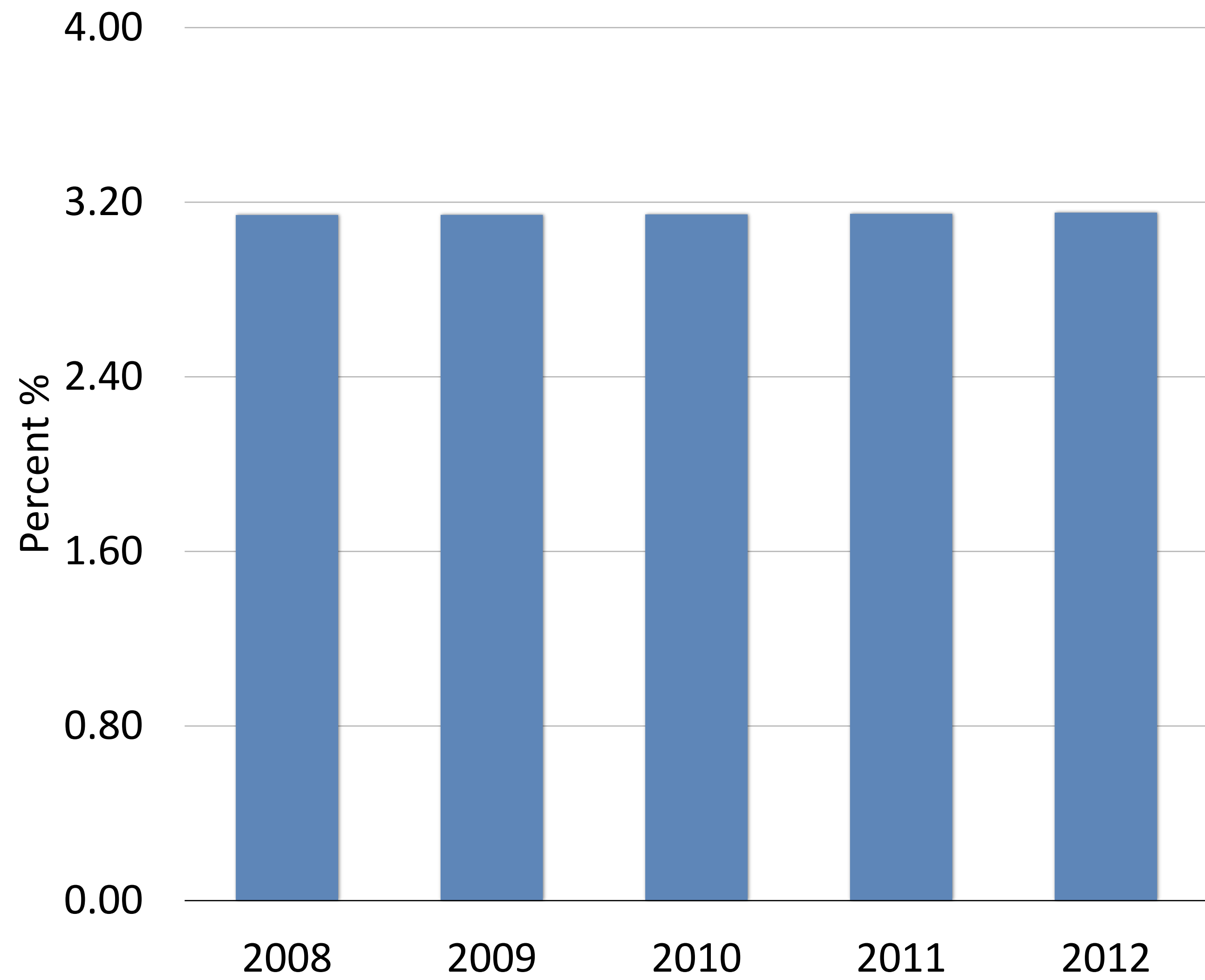
## Interest Rates



“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

Based on <http://data.heapanalytics.com/how-to-lie-with-data-visualization>

# Interest Rates



CONTEXT!

“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

Based on <http://data.heapanalytics.com/how-to-lie-with-data-visualization>



# “Double the axes, double the mischief”



“Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.”

<http://www.thefunctionalart.com/2015/10/double-axes-double-mischief.html>

# “Graphical Integrity”

“The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured.”



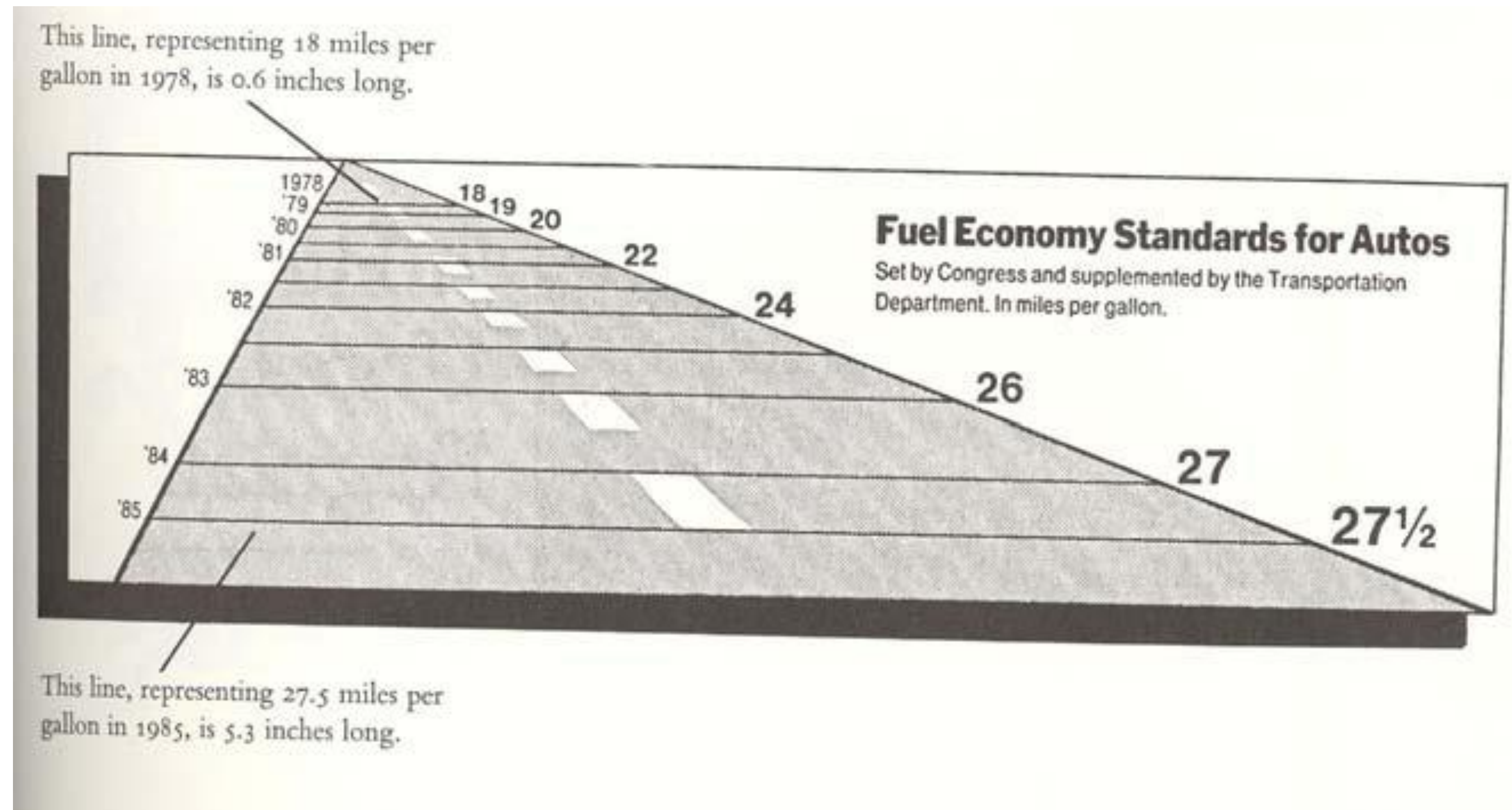
# Lie Factor

$$\text{Lie Factor} = \frac{\text{Size of effect in graphic}}{\text{Size of effect in data}}$$

Lie Factor = >1, overstating

Lie Factor = 1, accurate :-)

Lie Factor = <1, understating



“The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured.”

# Lie Factor

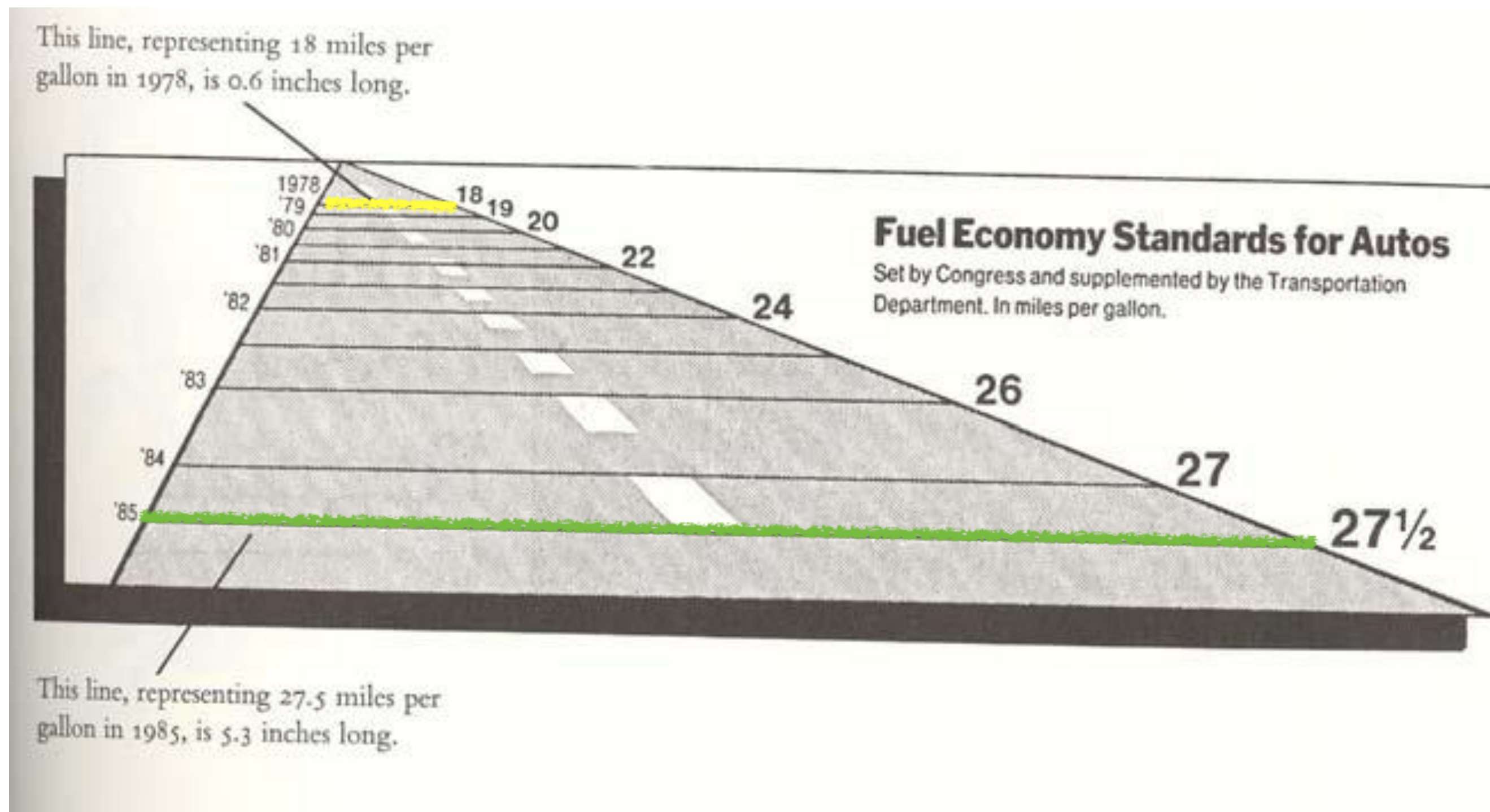
$$\text{Lie Factor} = \frac{\text{(Size of effect in graphic)}}{\text{(Size of effect in data)}}$$

$$\text{Image} = \frac{5.3'' - 0.6''}{0.6''} = 7.83 = 783\%$$

$$\text{Data} = \frac{27.5 - 18}{18} = 0.53 = 53\%$$

$$\text{Lie Factor} = \frac{783\%}{53\%} = 14.8$$

Lie Factor = >1, overstating



“The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured.”



# Lie Factor

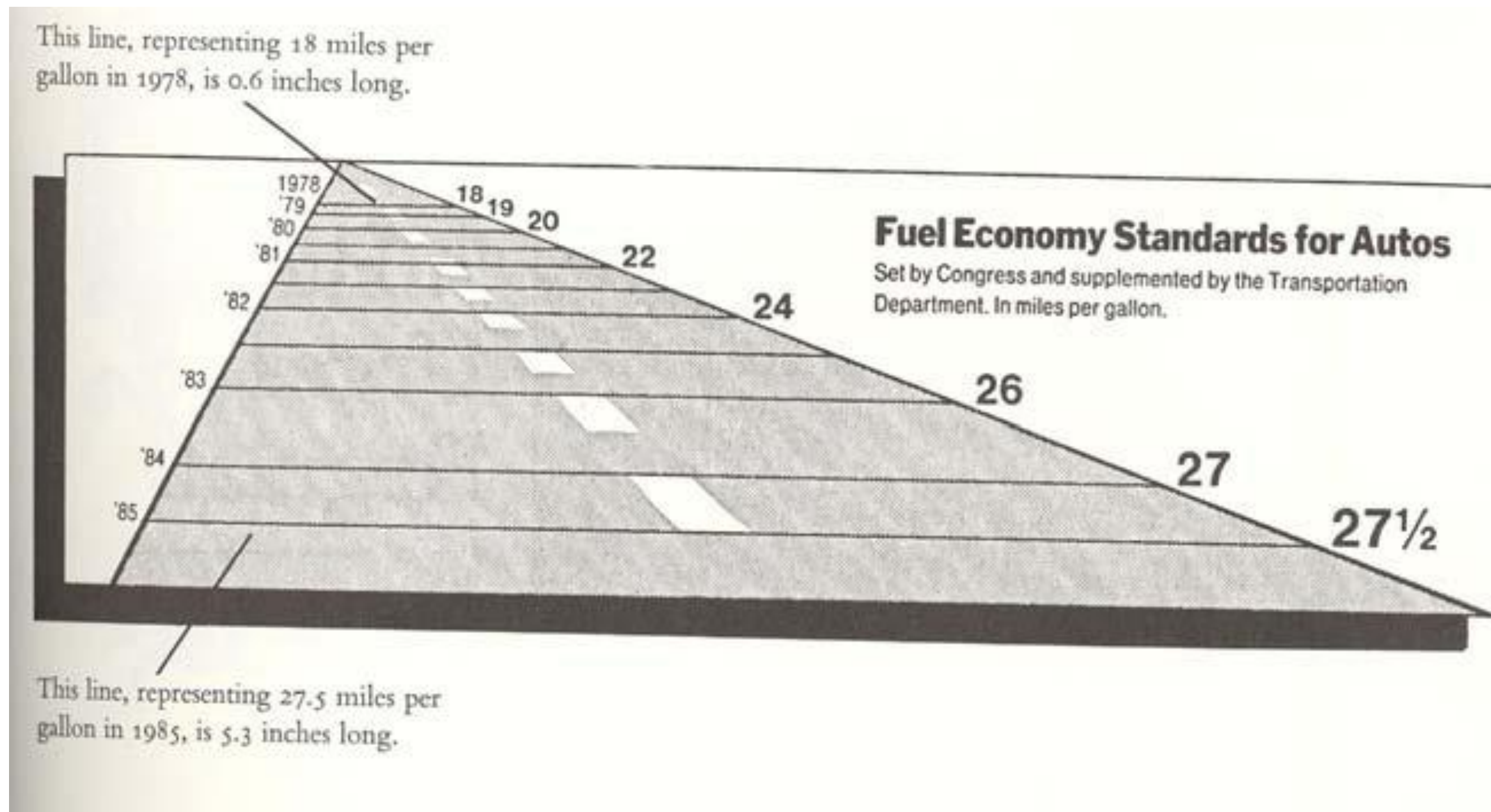
$$\text{Lie Factor} = \frac{\text{Size of effect in graphic}}{\text{Size of effect in data}}$$

$$\text{Image} = \frac{5.3'' - 0.6''}{0.6''} = 7.83 = 783\%$$

$$\text{Data} = \frac{27.5 - 18}{18} = 0.53 = 53\%$$

$$\text{Lie Factor} = \frac{783\%}{53\%} = 14.8$$

Lie Factor = >1, overstating



18  
27.5

“The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured.”

IN-CLASS ACTIVITY:  
Calculate for yourself!

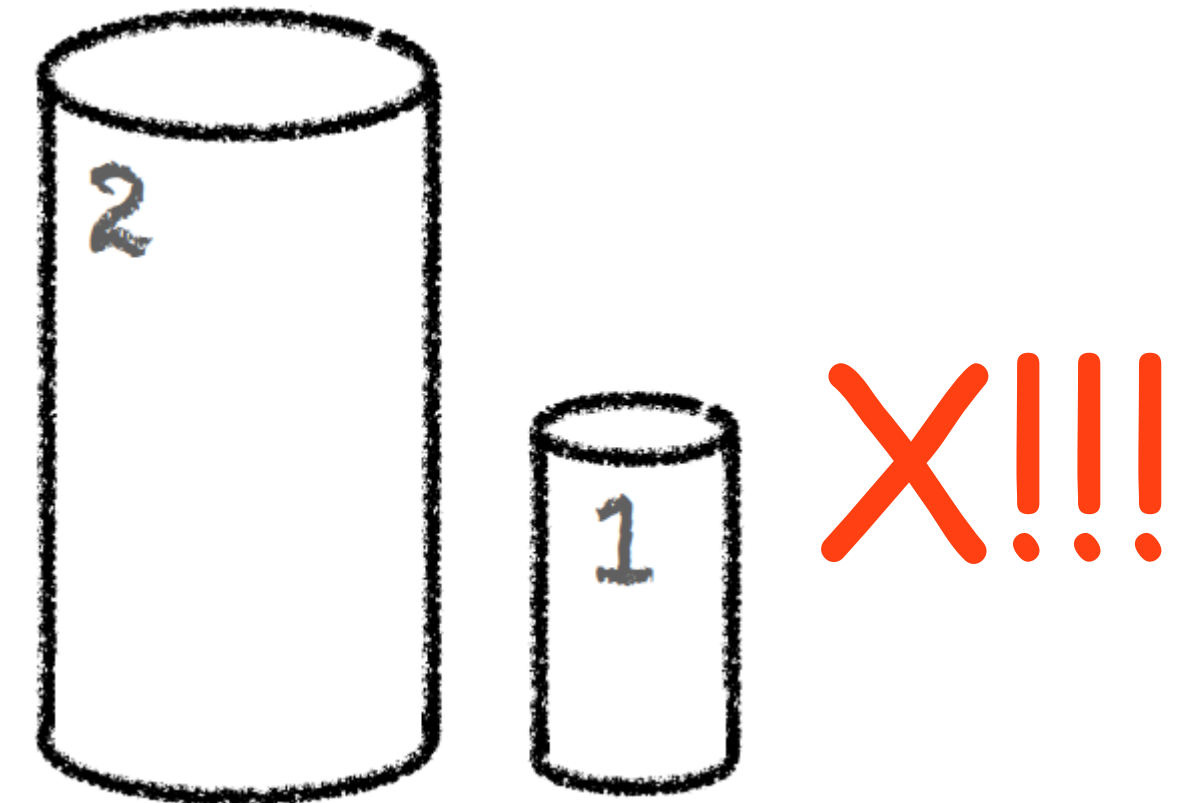
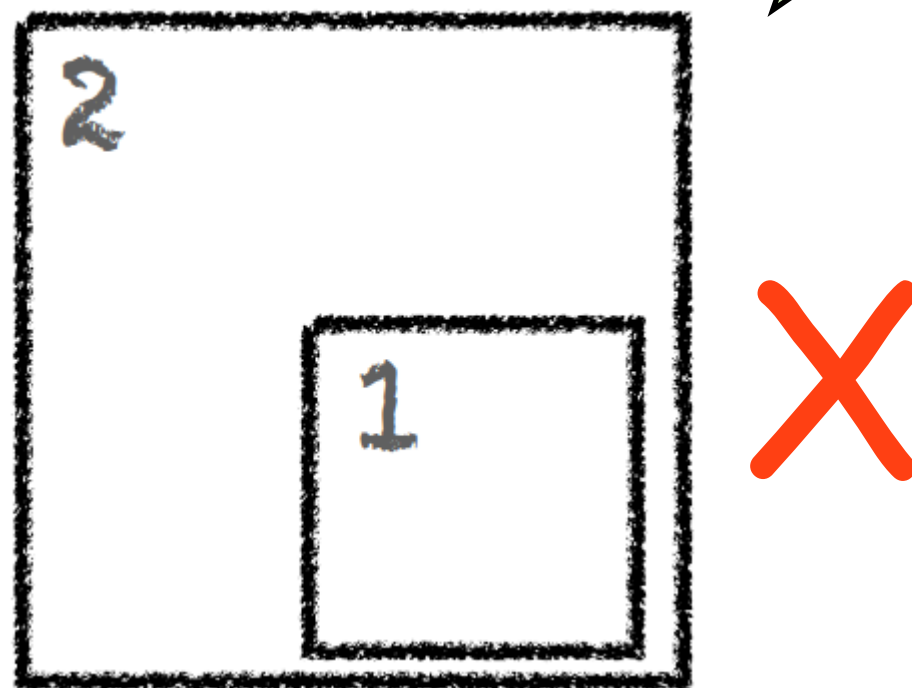
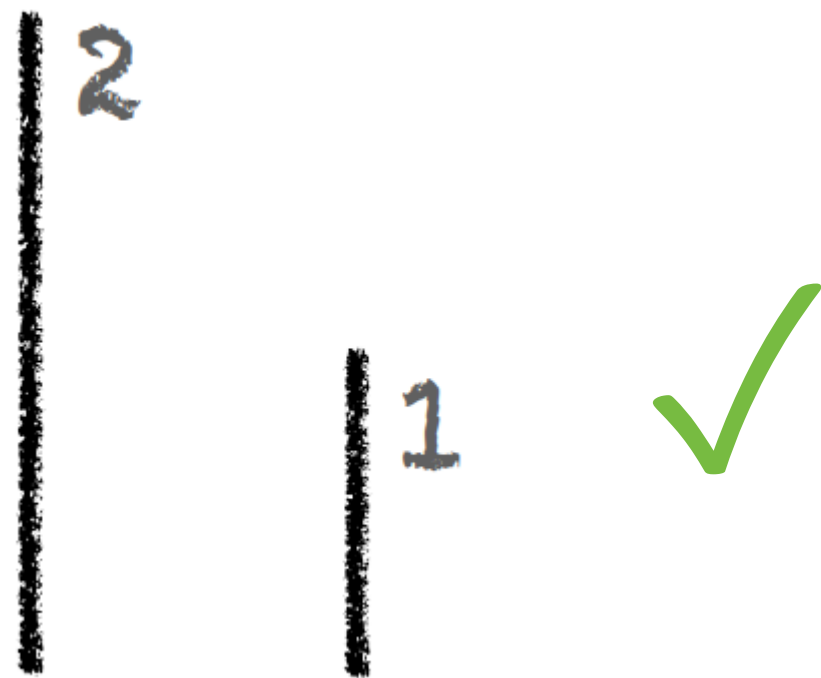
# Lie Factor

$$\text{Data} = \frac{2 - 1}{1} = 1 = 100\%$$

$$\text{Lie Factor} = \frac{\text{Size of effect in graphic}}{\text{Size of effect in data}}$$

Make sure *area* is proportional to data!

Don't use 3D bar charts!



$$\text{Image} = \frac{2 - 1}{1} = 1 = 100\%$$

$$\text{Image} = \frac{2^2 - 1^2}{1^2} = 3 = 300\%$$

$$\text{Image} = \frac{2 * \pi 1^2 - 1 * \pi 0.5^2}{1 * \pi 0.5^2} = 7 = 700\%$$

$$\text{Lie Factor} = \frac{100\%}{100\%} = 1$$

$$\text{Lie Factor} = \frac{300\%}{100\%} = 3$$

$$\text{Lie Factor} = \frac{700\%}{100\%} = 7$$

“The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities measured.”



# “Graphical Integrity”

“The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.”

**“CHART JUNK”**



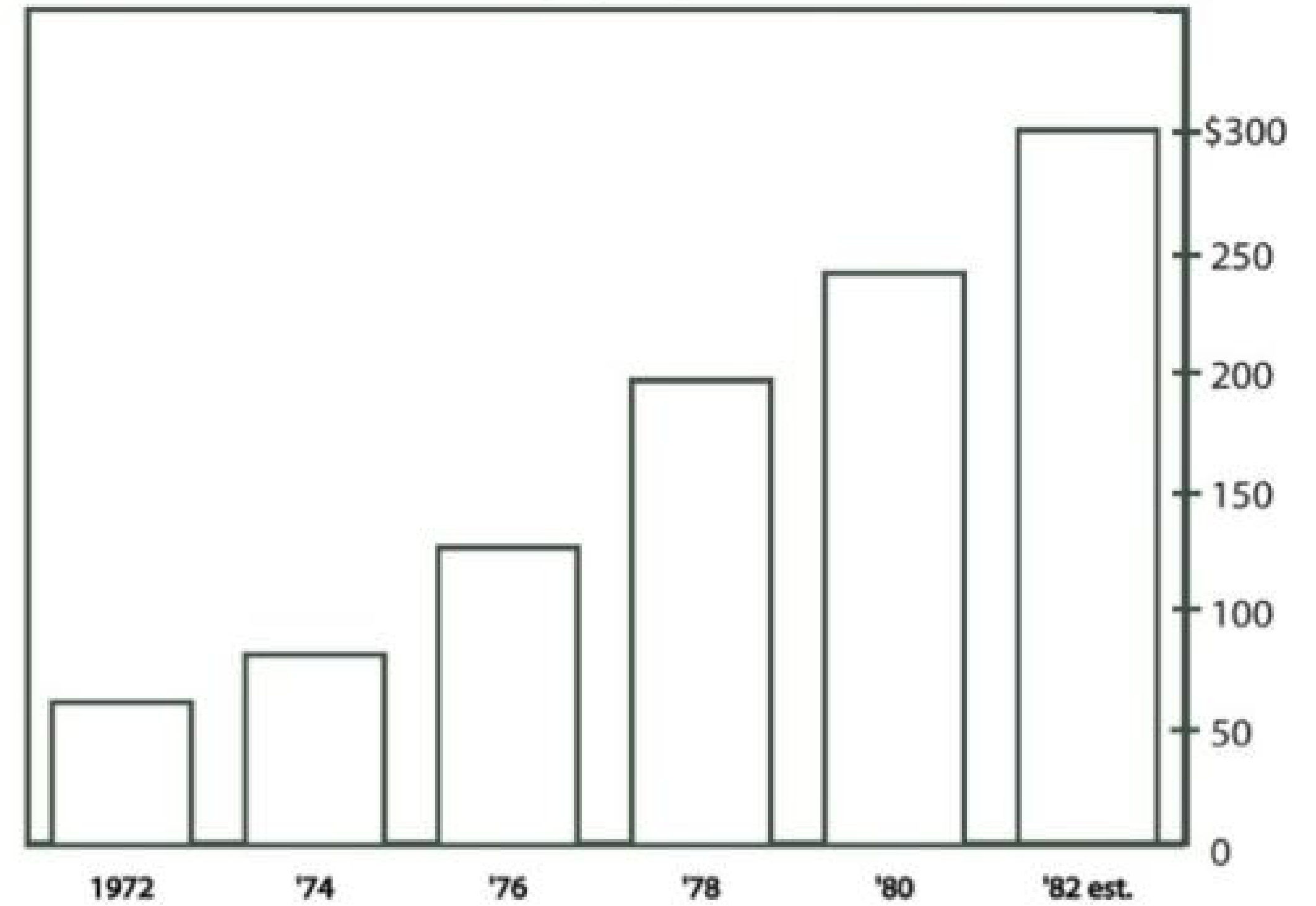
# “Chart Junk”

## MONSTROUS COSTS

Total House and Senate campaign expenditures, in millions

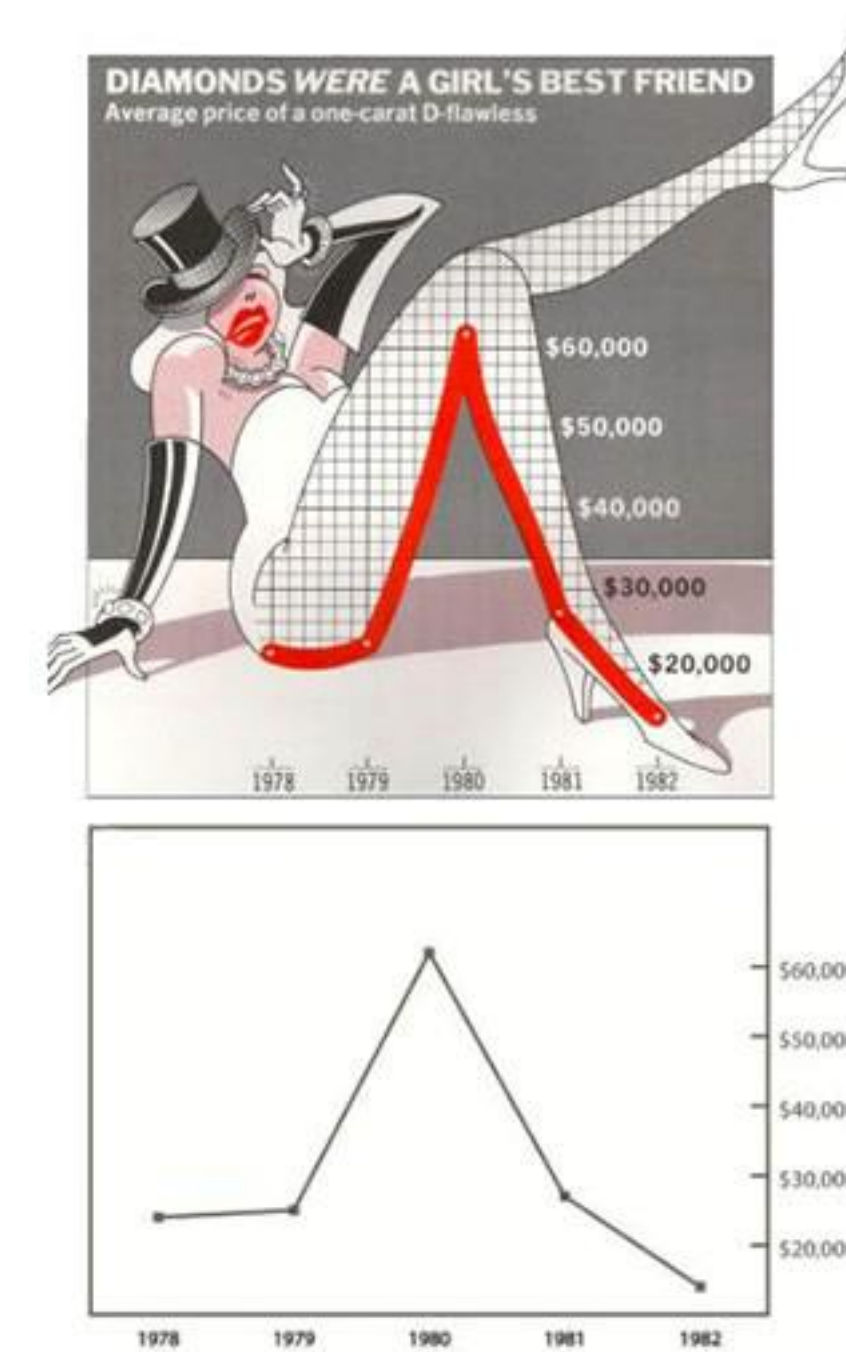
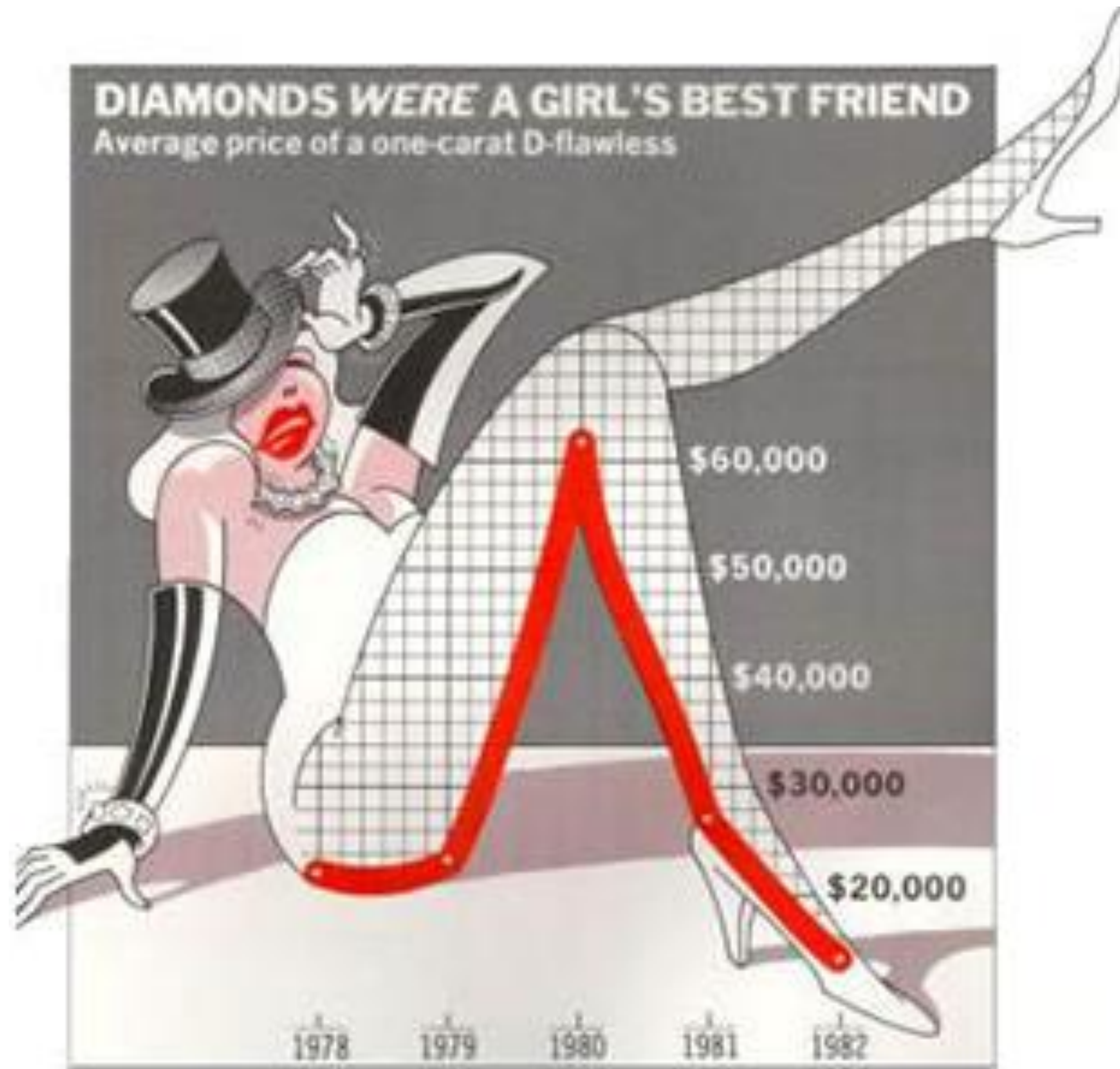


MONSTROUS COSTS  
Total House and Senate campaign expenditures, in millions





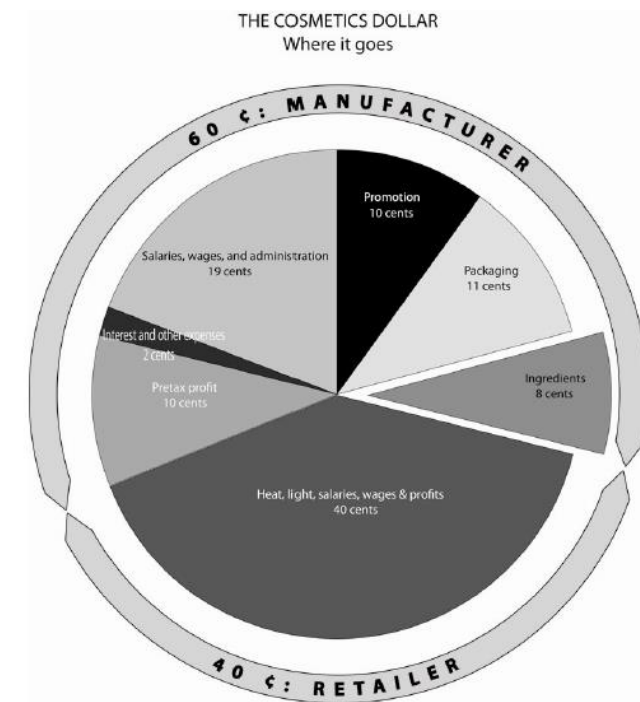
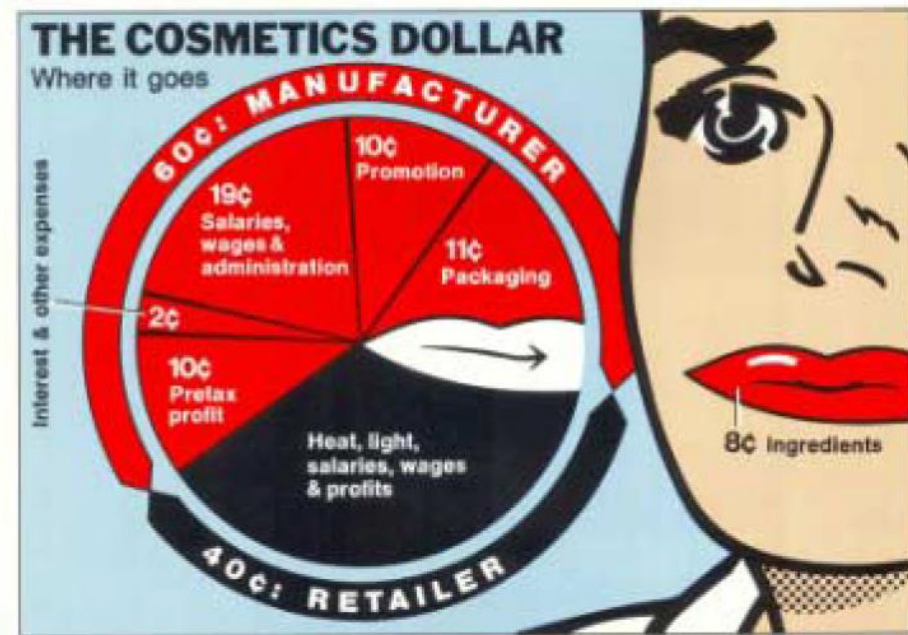
# “Chart Junk”





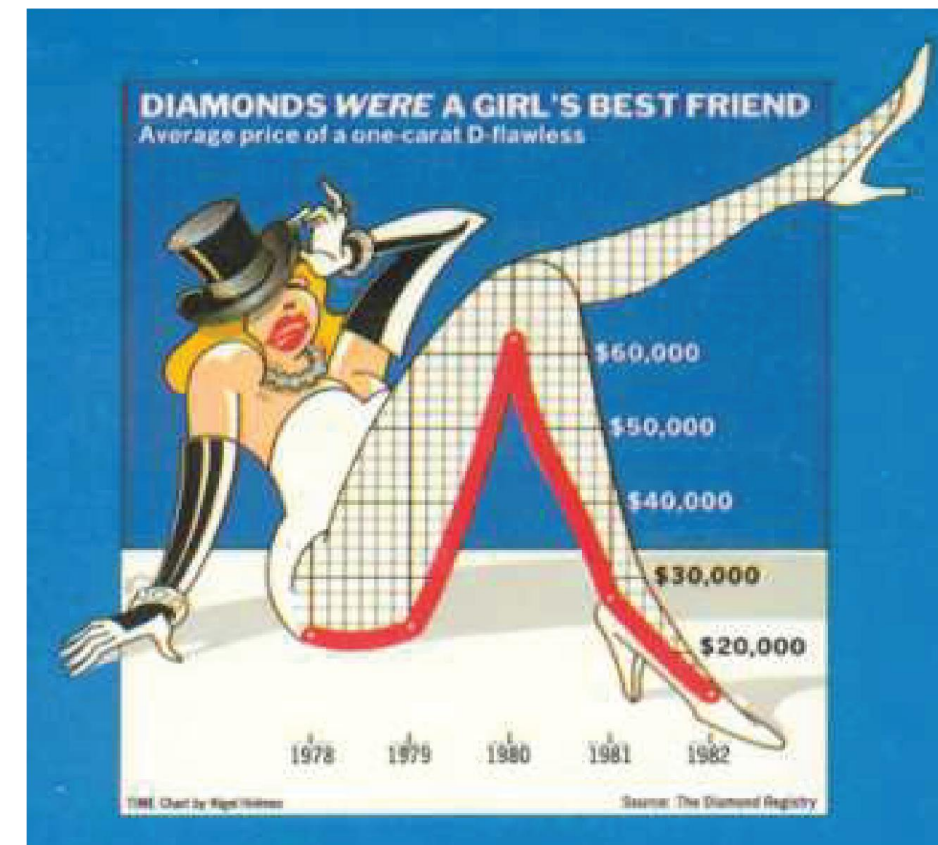
# “Chart Junk Debate”

## Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts



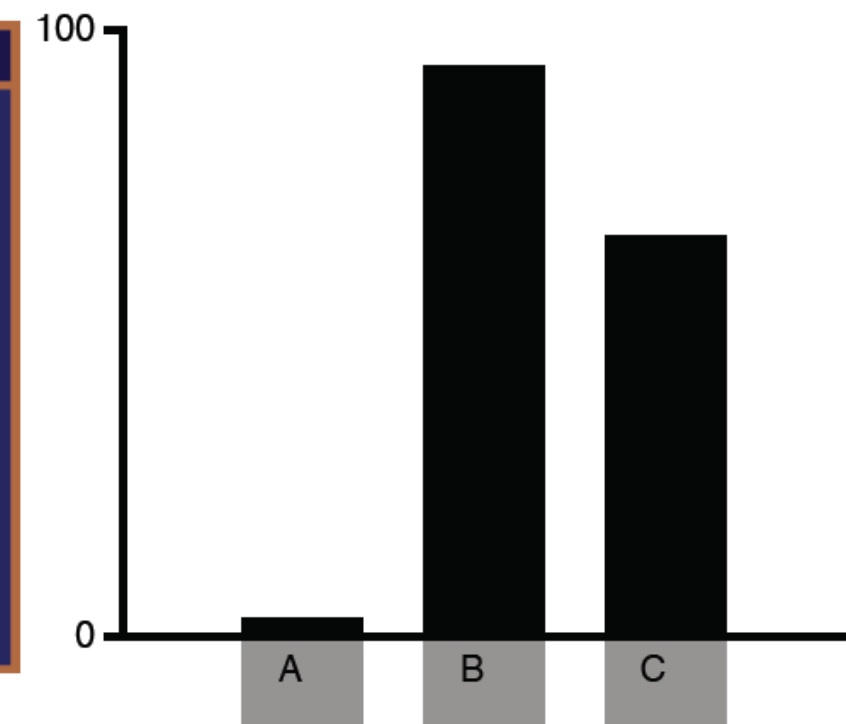
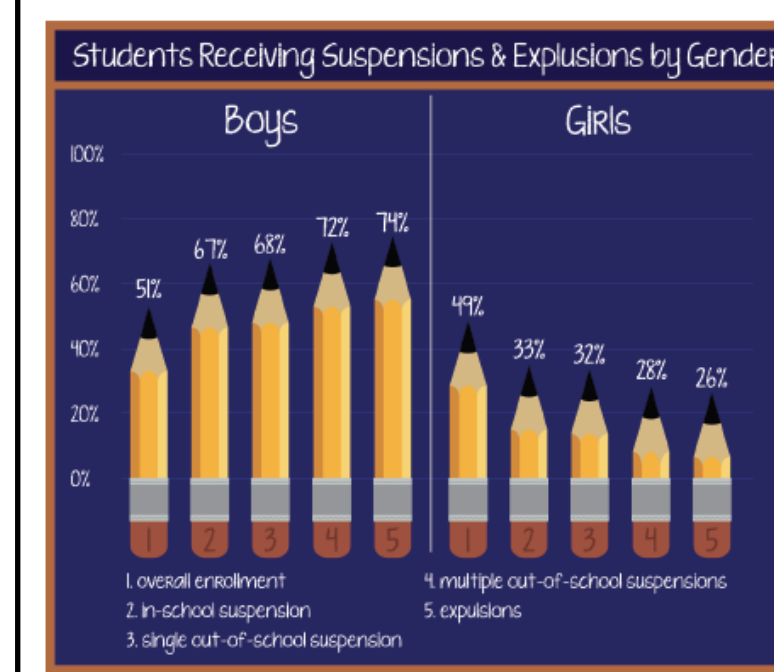
Bateman, et al. (2010)

## Benefitting InfoVis with Visual Difficulties



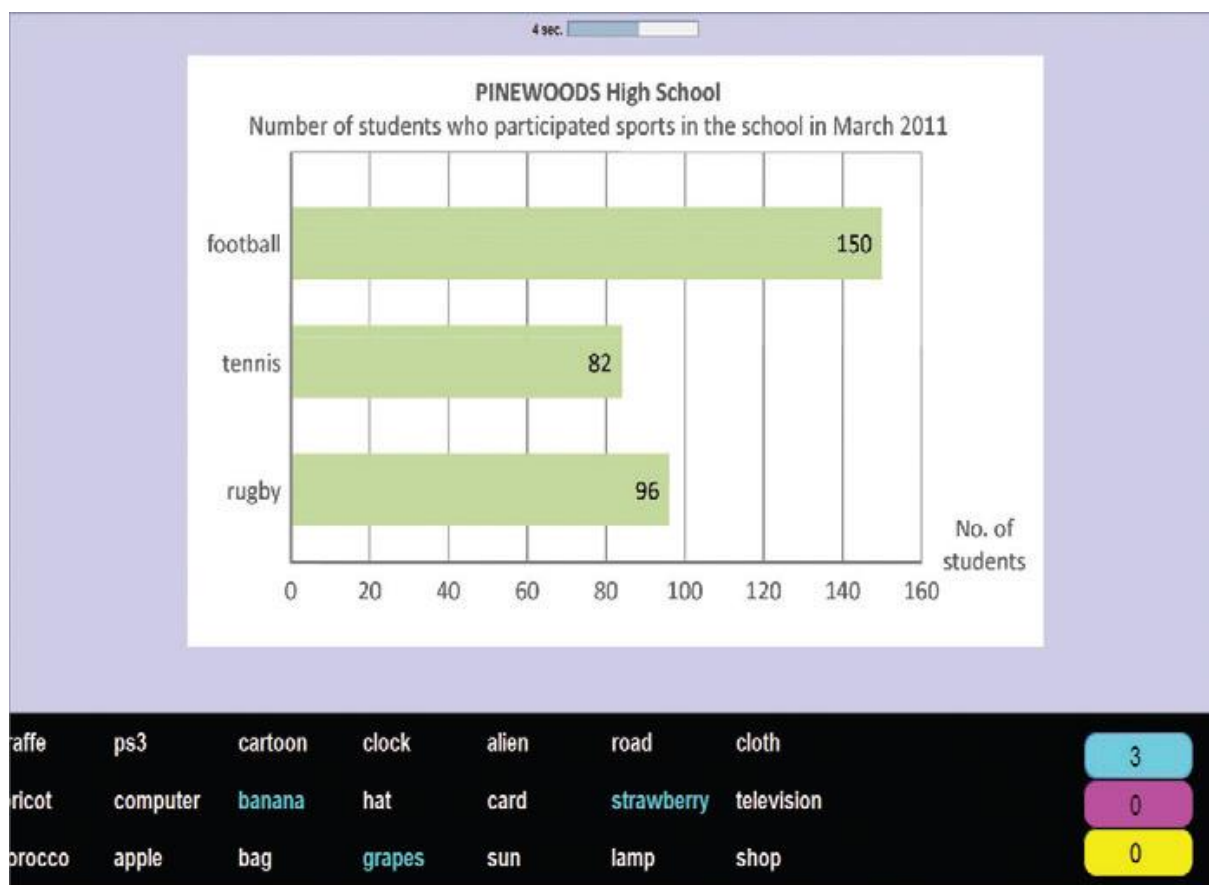
Hullman, et al. (2011)

## An Evaluation of the Impact of Visual Embellishments in Bar Charts



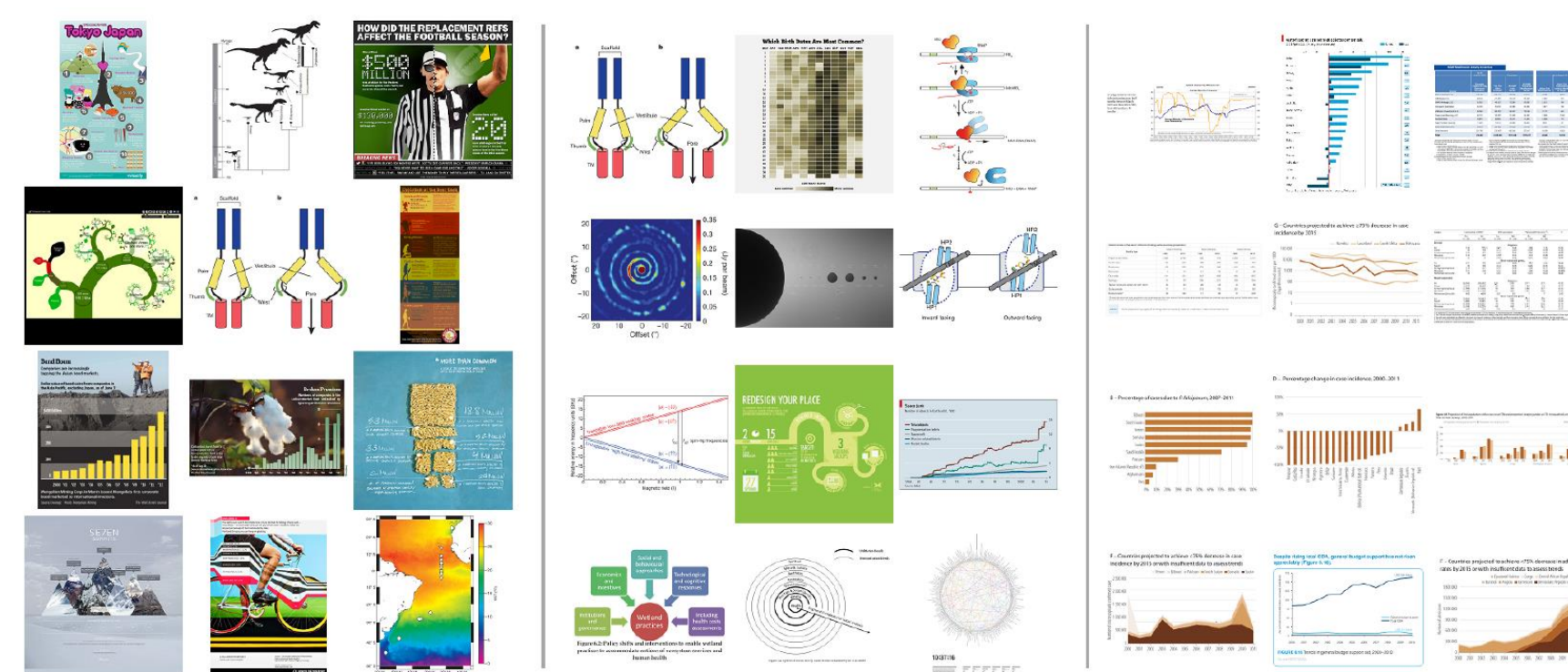
Skau, et al. (2015)

## An Empirical Study on Using Visual Embellishments in Visualization



Borgo, et al. (2012)

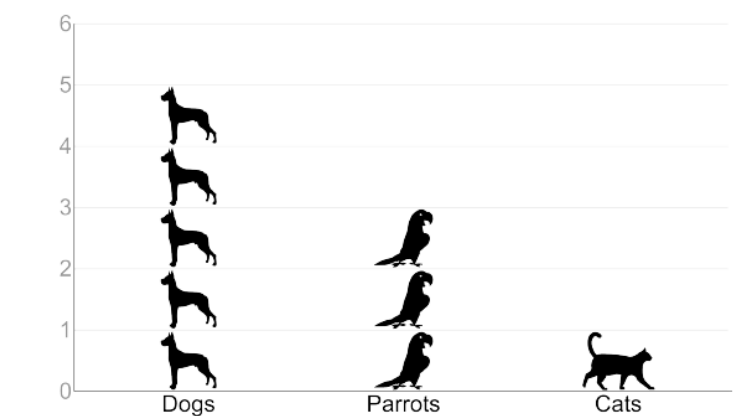
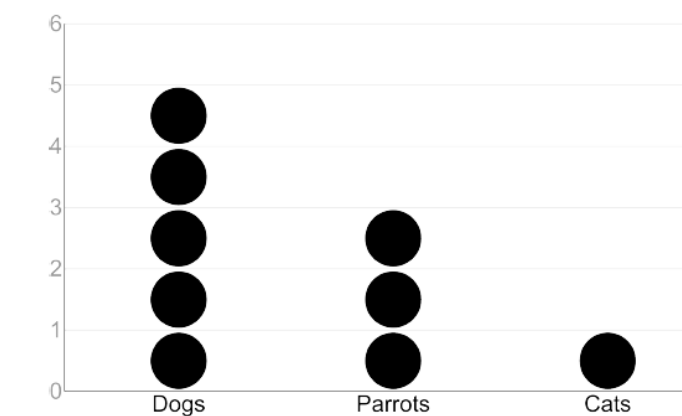
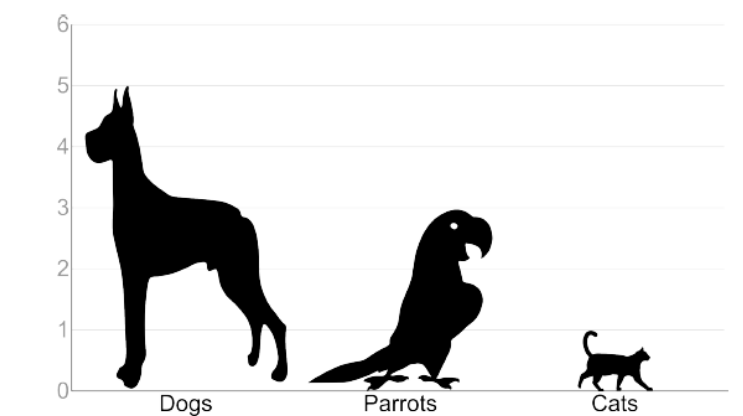
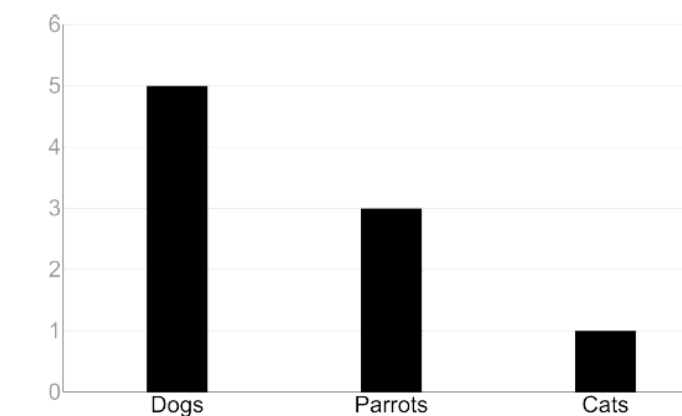
## What makes a visualization memorable?



Borkin, et al. (2013)

Borkin, et al. (2015)

## ISOTYPE Visualization – Working Memory, Performance, and Engagement with Pictographs



Haroz, et al. (2015)

# “Chart Junk”

Chart junk can... persuade, help with memorability, engage  
... bias, limit data-ink ratio, clutter, lower trust

Take-away: *it depends on your audience, task, and context...*



# For Next Time

[neu-ds-4200-f23.github.io/schedule/](https://neu-ds-4200-f23.github.io/schedule/)

Look at the upcoming assignments and deadlines

- Textbook, Readings, & Reading Quizzes—Variable days
- In-Class Activities—If due, they are due 11:59pm the same day as class

Everyday Required Supplies:

- 5+ colors of pen or marker
- White paper
- Laptop and charger

Use Slack for general questions, email [codydunne-and-tas@ccs.neu.edu](mailto:codydunne-and-tas@ccs.neu.edu) for questions specific to you.

## Schedule

Please be sure to read the assigned reading before class.

### Week 1: Introductions

**Fri, Sep 08**

*What is Visualization*

### Week 2: Design & Marks

**Tue, Sep 12**

*Design rules of thumb*

Required Readings:

- 1 VAD Chapter 1—What is Vis, and Why Do It?
- 2 VAD Chapter 6—Rules of Thumb

**Fri, Sep 15**

*Marks and channels*

Required Readings:

- 1 VAD Chapter 5—Marks and Channels
- 2 [A Tour through the Visualization Zoo by Heer, Bostock, and Ogievetsky \(2012\)](#)

A1—Getting started Due at 11:59pm