

# Lecture 1: So It Begins!

CS 7250  
SPRING 2021  
*Prof. Cody Dunne*  
NORTHEASTERN UNIVERSITY

*Slides and inspiration from Michelle Borkin, Krzysztof Gajos, Hanspeter Pfister, Miriah Meyer, Jonathan Schwabish, and David Sprague*

# STAFF INTRODUCTIONS



Cody Dunne  
Assistant Professor  
Instructor



Girik Malik  
TA



David Saffo  
TA



Sara Di Bartolomeo  
TA



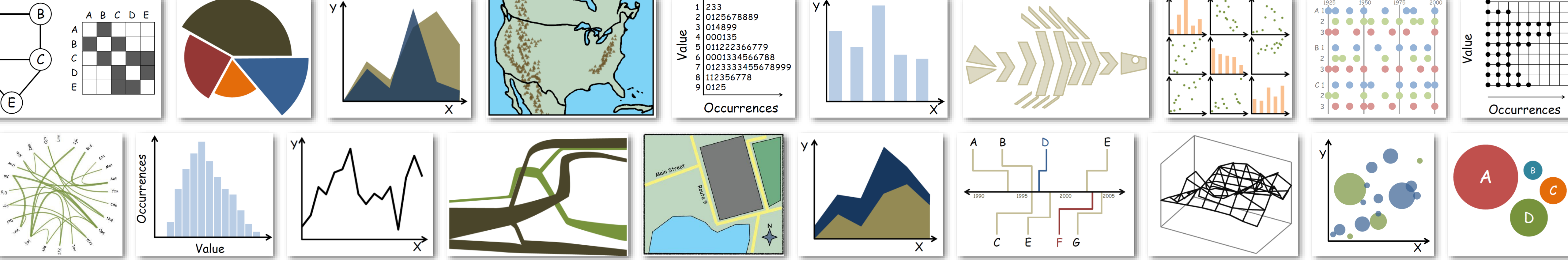
Sophia Gunzberg  
Service-Learning TA

# Course Homepage on Canvas

[c.dunne.dev/cs7250s21](https://c.dunne.dev/cs7250s21)

What is visualization  
anyway?

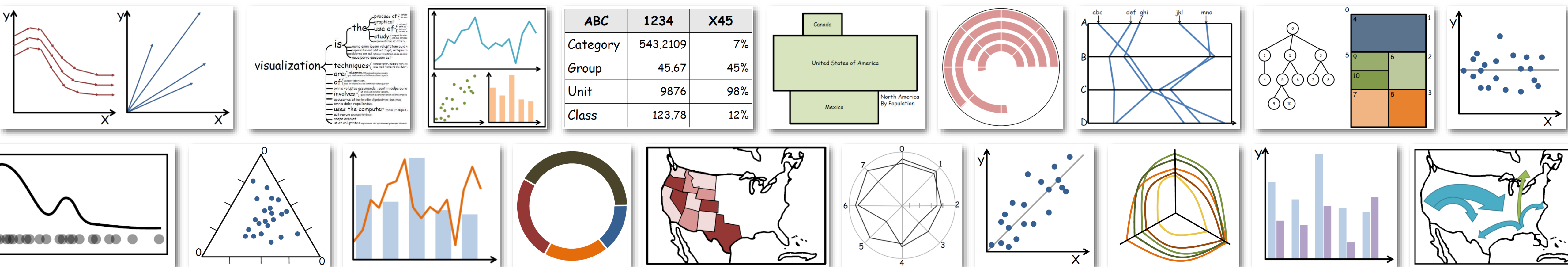




(static or interactive)

(abstract or spatial)

visualization: the visual representation of data to reinforce human cognition

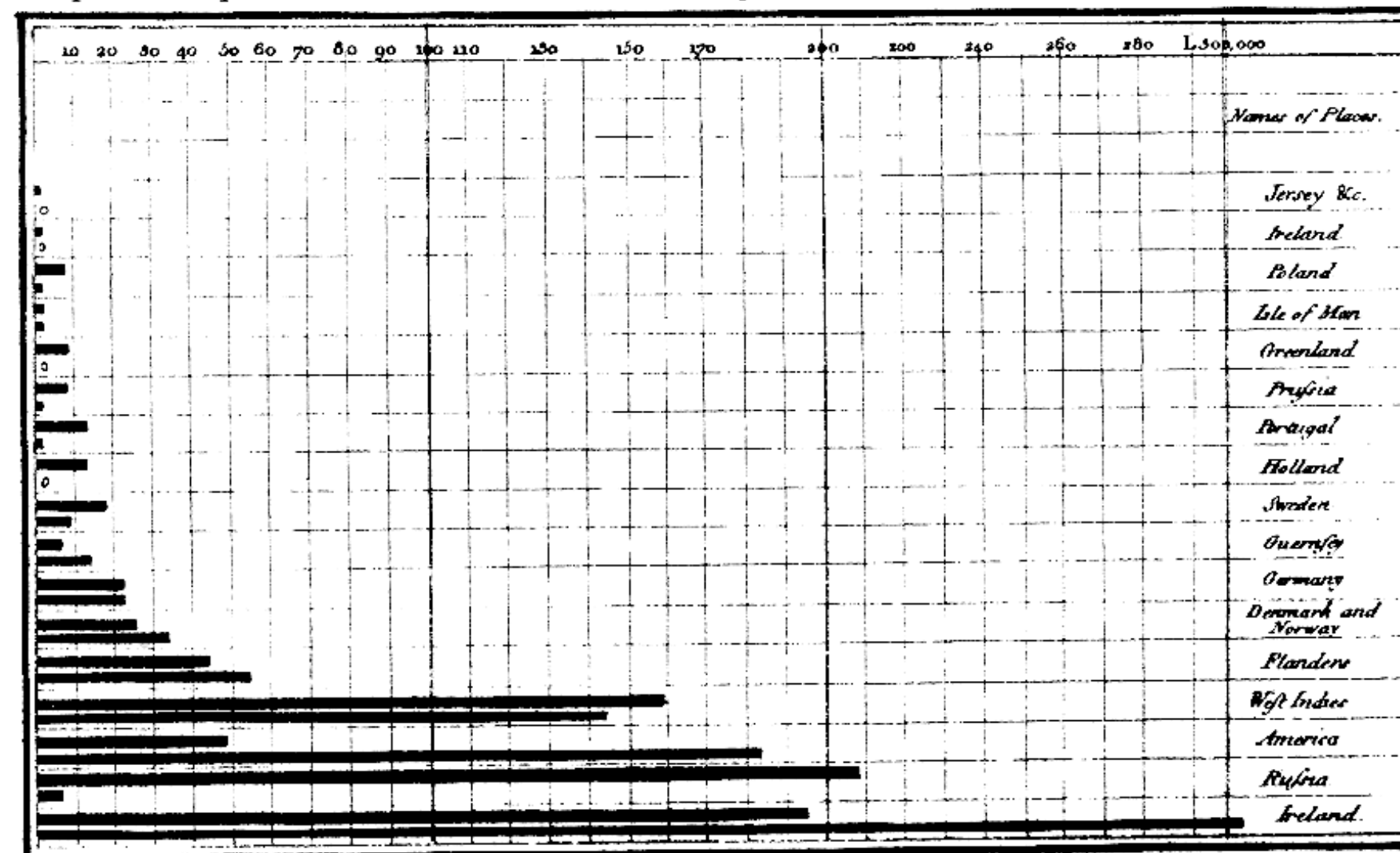






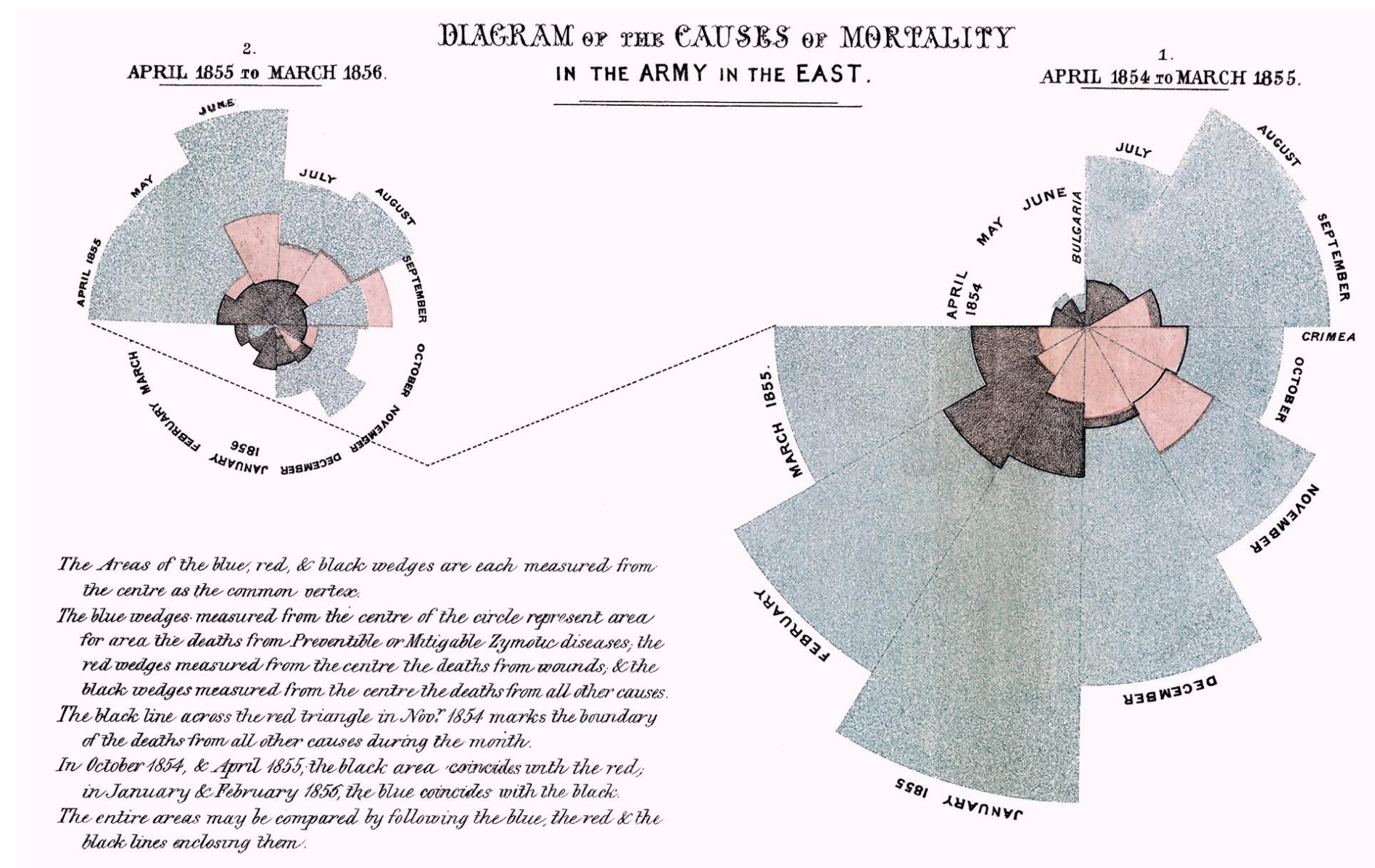
PTOLEMY (c. 150)

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



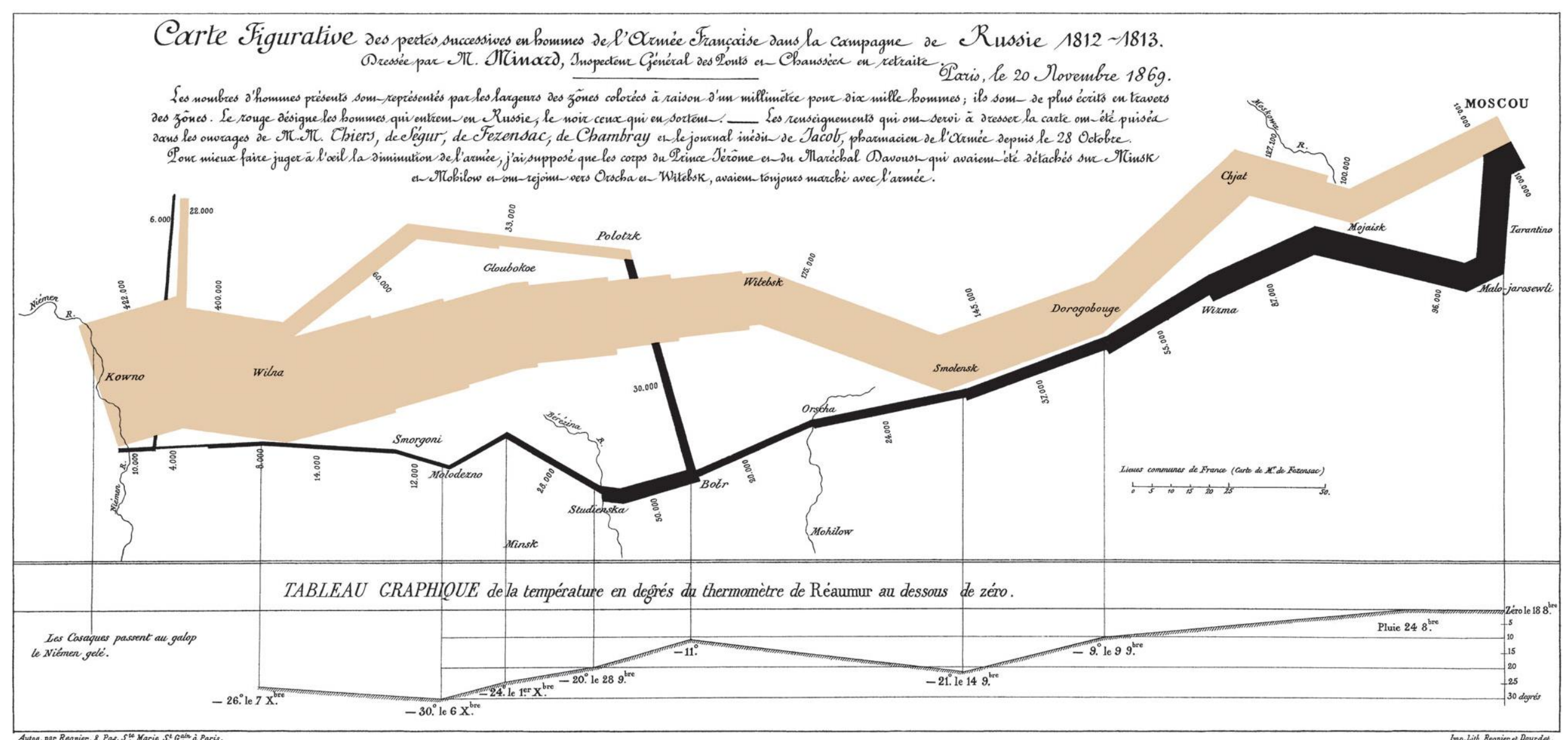
The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.  
 Published as the Act above June 7<sup>th</sup> 1788 by W<sup>m</sup> Playfair  
 Made comp<sup>t</sup> 352. 1/2 Lond<sup>n</sup>

WILLIAM PLAYFAIR (c. 1786)



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.

FLORENCE NIGHTINGALE (c. 1858)



Autog. par Regnier, 1. Pas. 5<sup>me</sup> Marie. 57 0<sup>me</sup> à Paris. Imp. Lit. Régis et Desrosiers.

CHARLES MINARD (c. 1869) 6



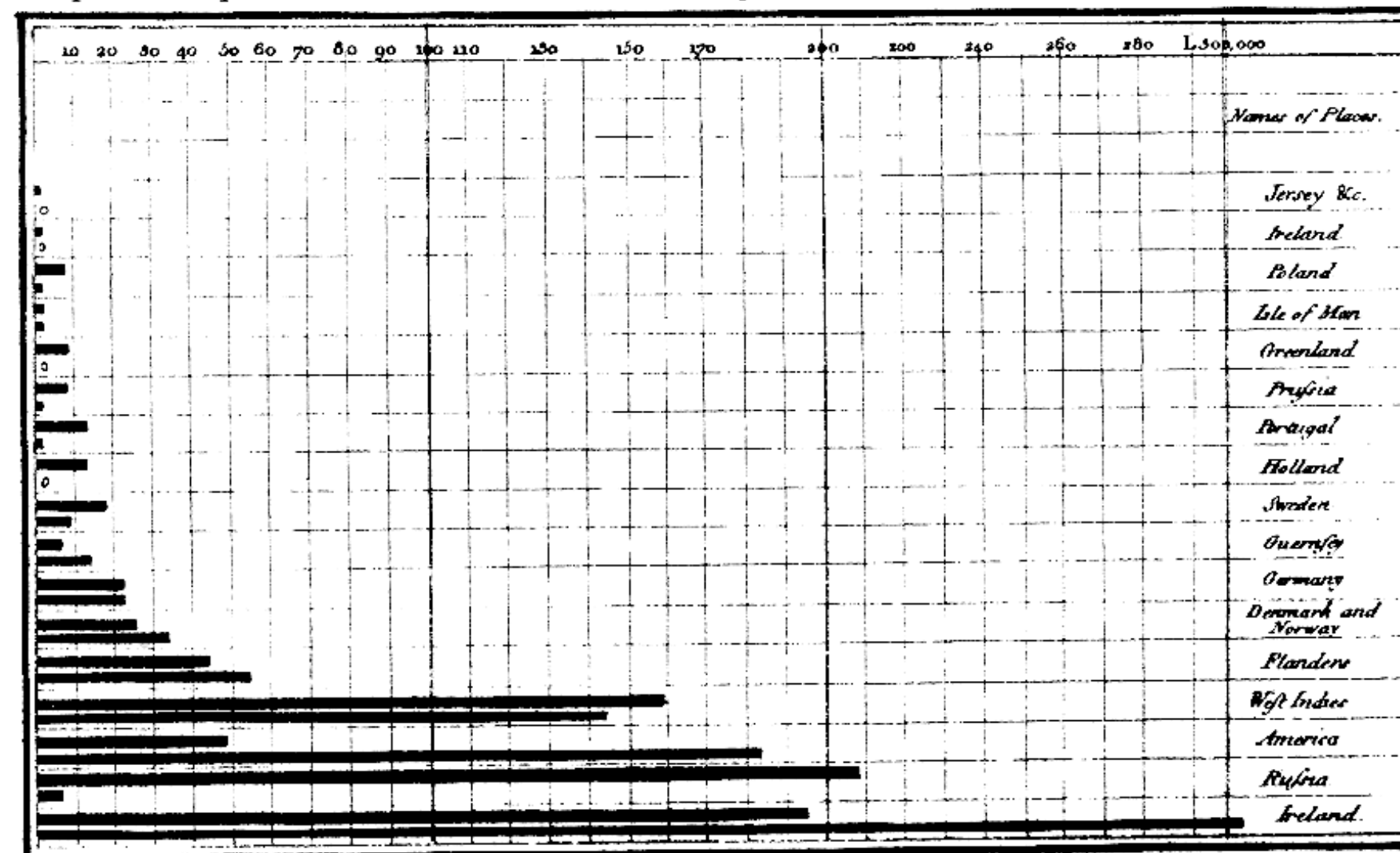






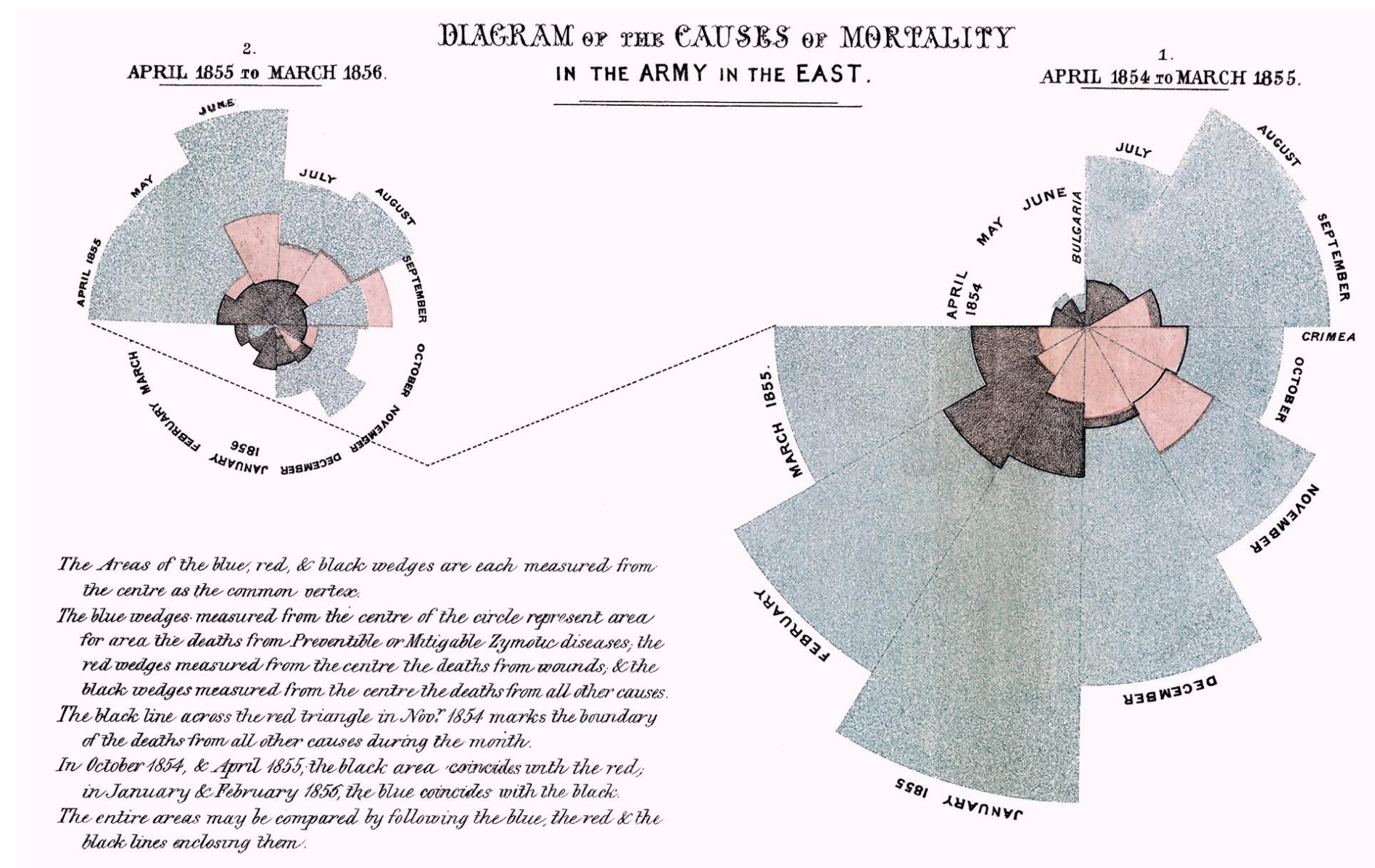
PTOLEMY (c. 150)

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

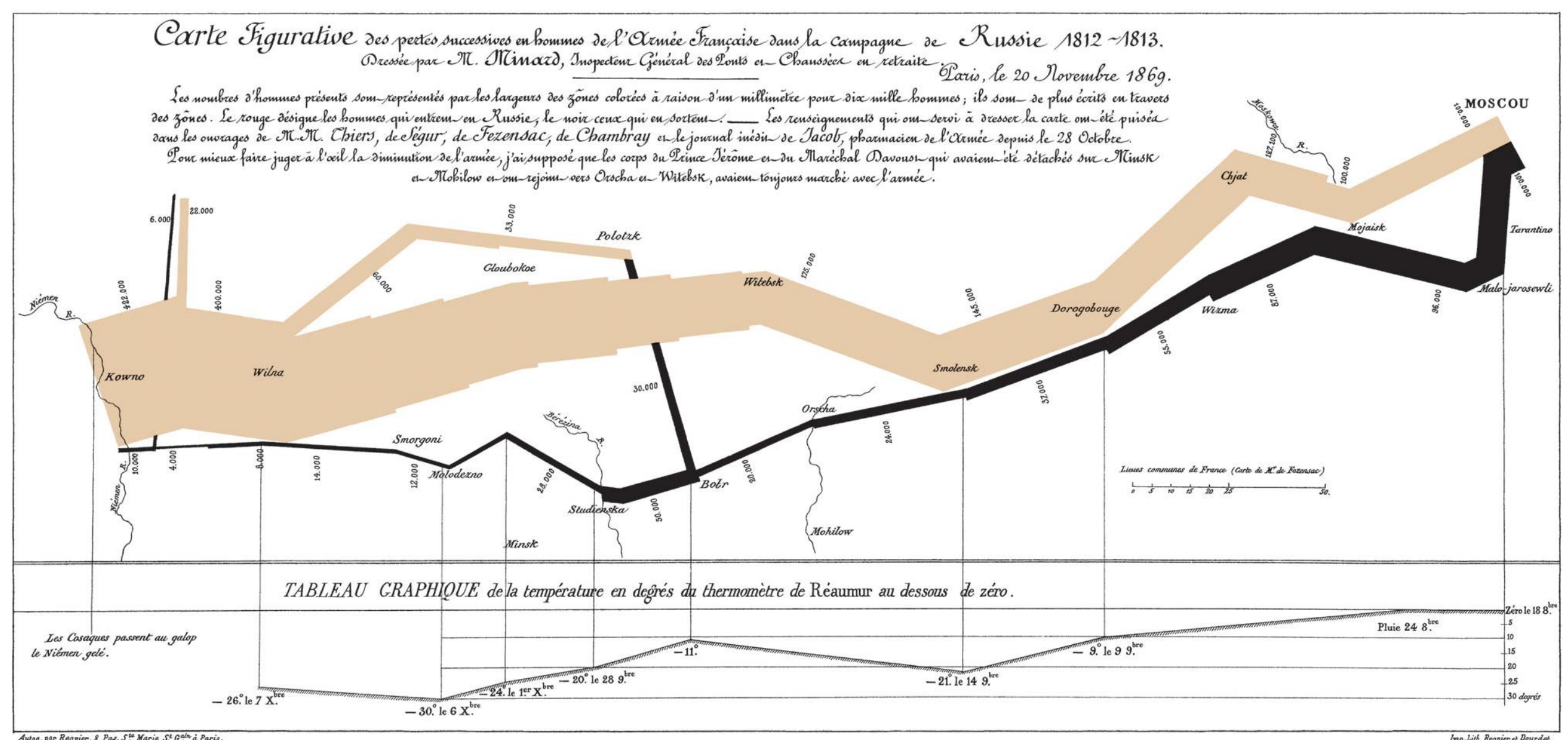


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.  
 Published in the Edinburgh Journal June 7<sup>th</sup> 1786 by W<sup>m</sup> Playfair  
 Made comp<sup>d</sup> 352. Strand. London.

WILLIAM PLAYFAIR (c. 1786)



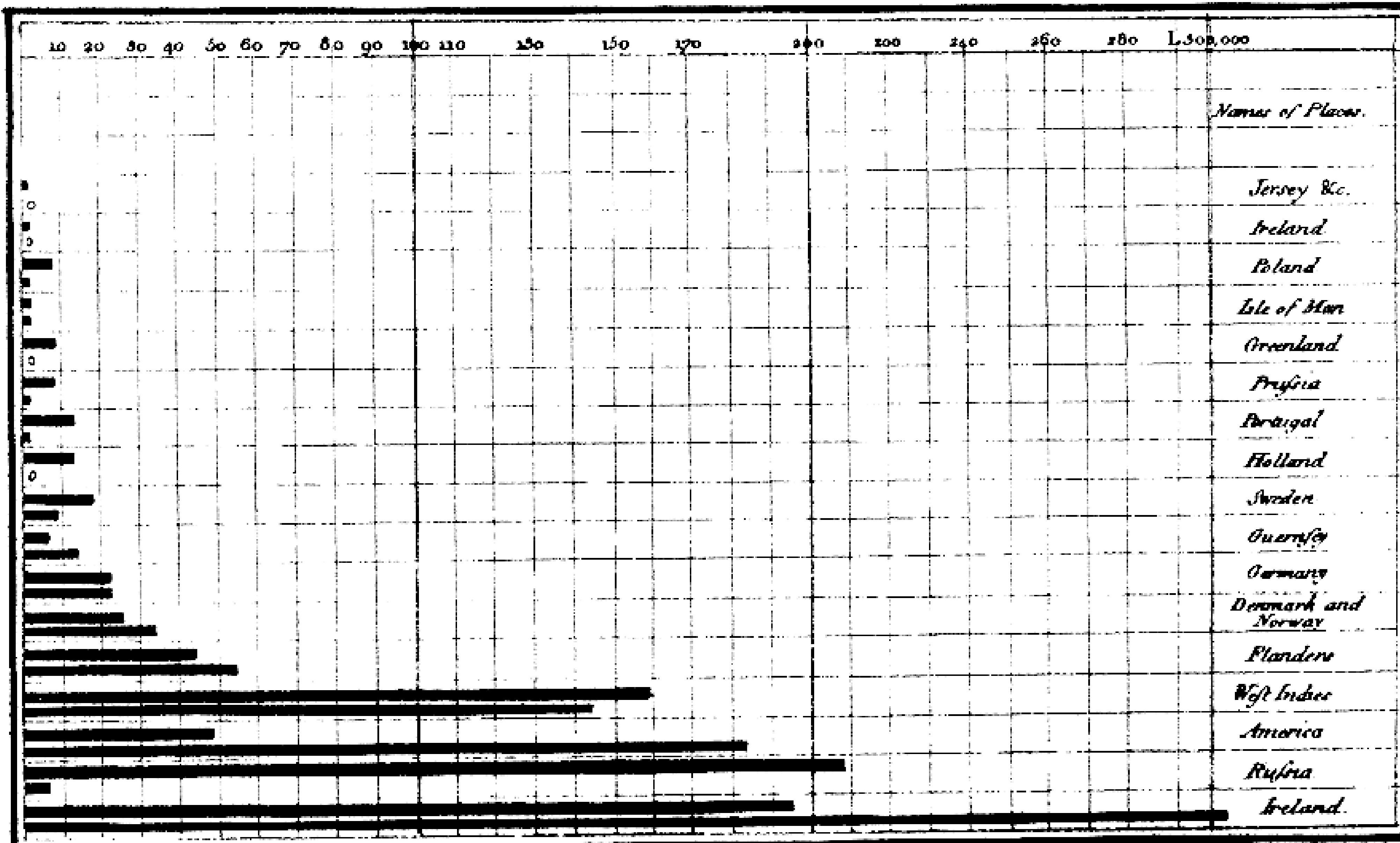
FLORENCE NIGHTINGALE (c. 1858)



CHARLES MINARD (c. 1869)



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



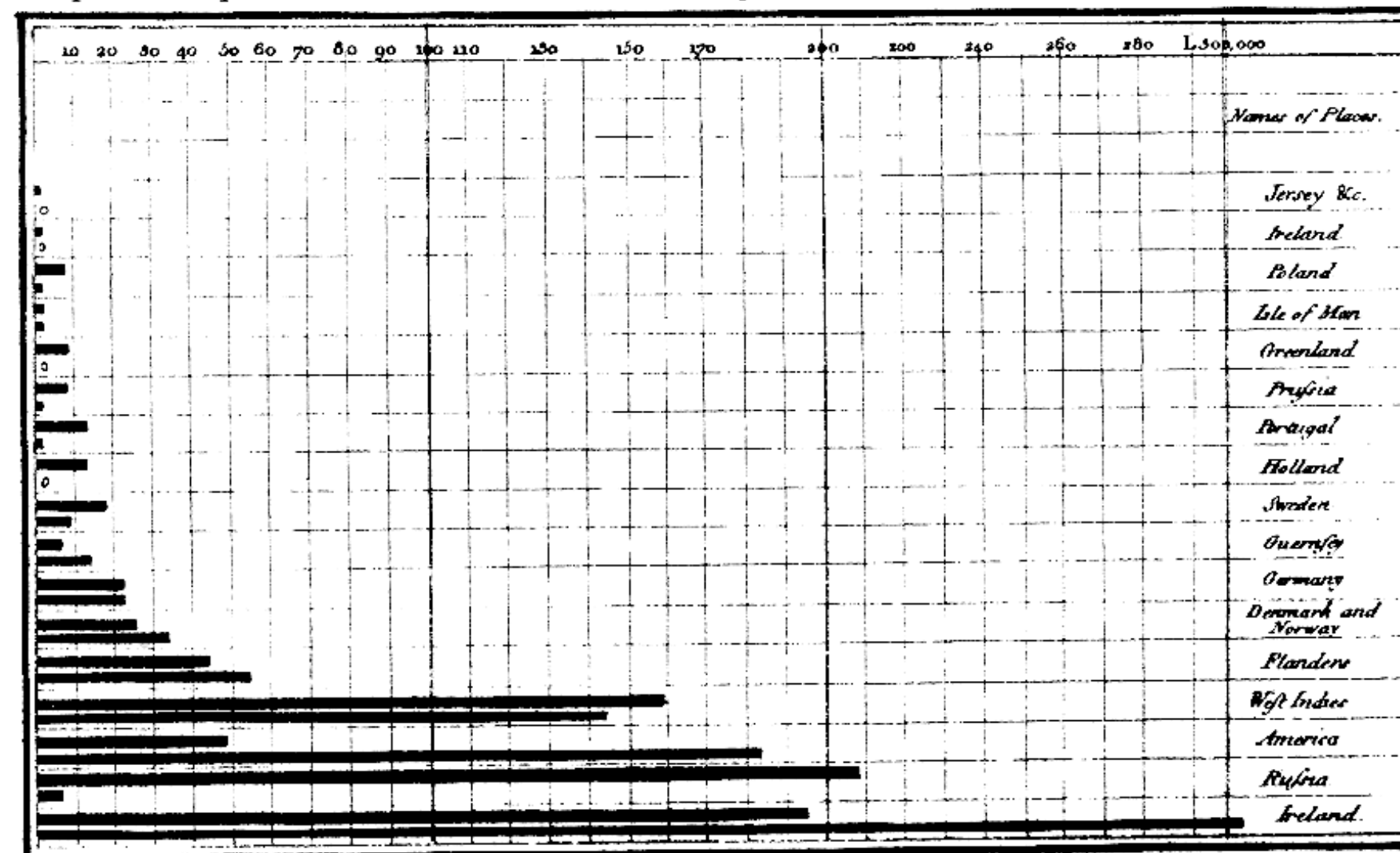
The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.





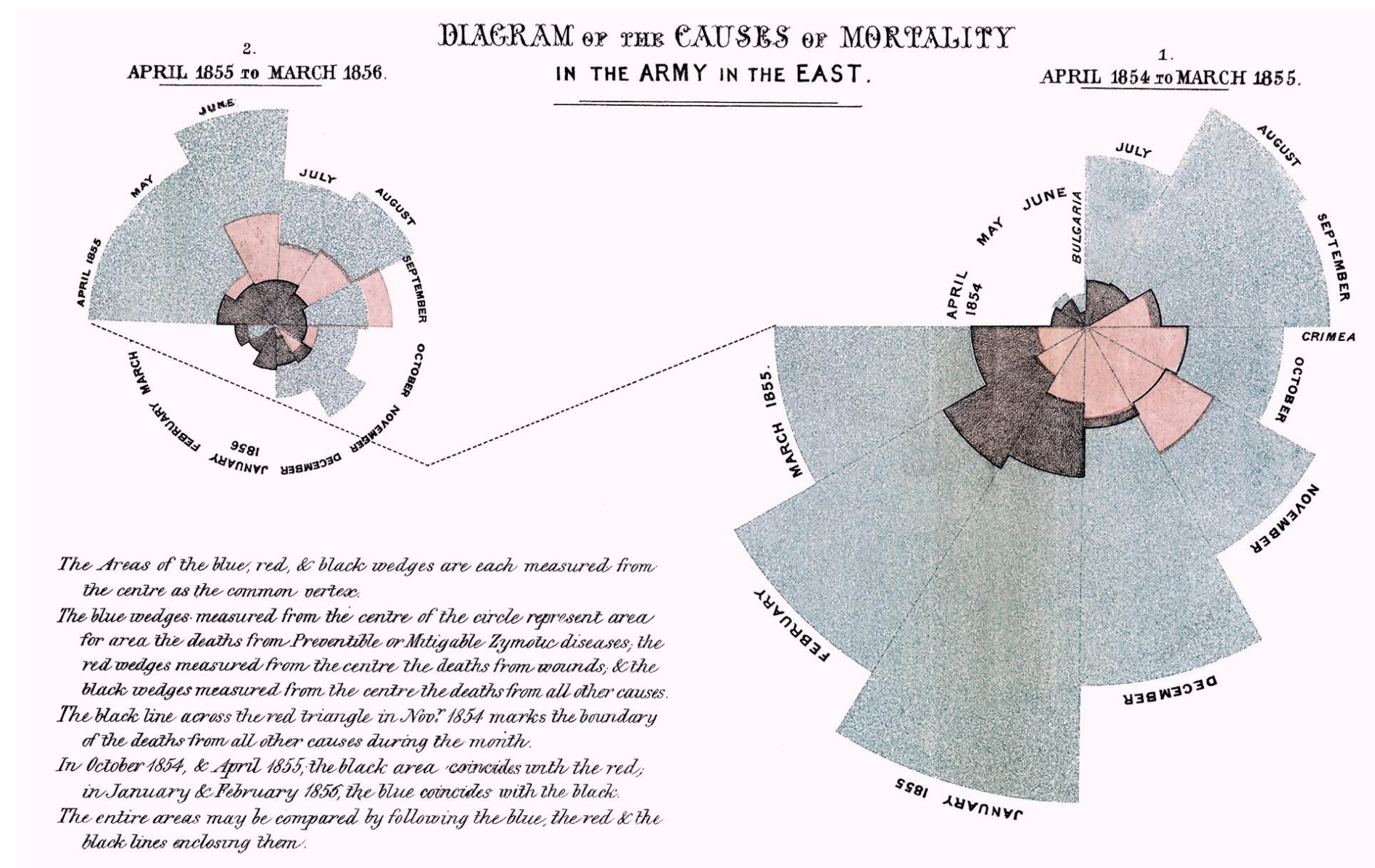
PTOLEMY (c. 150)

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

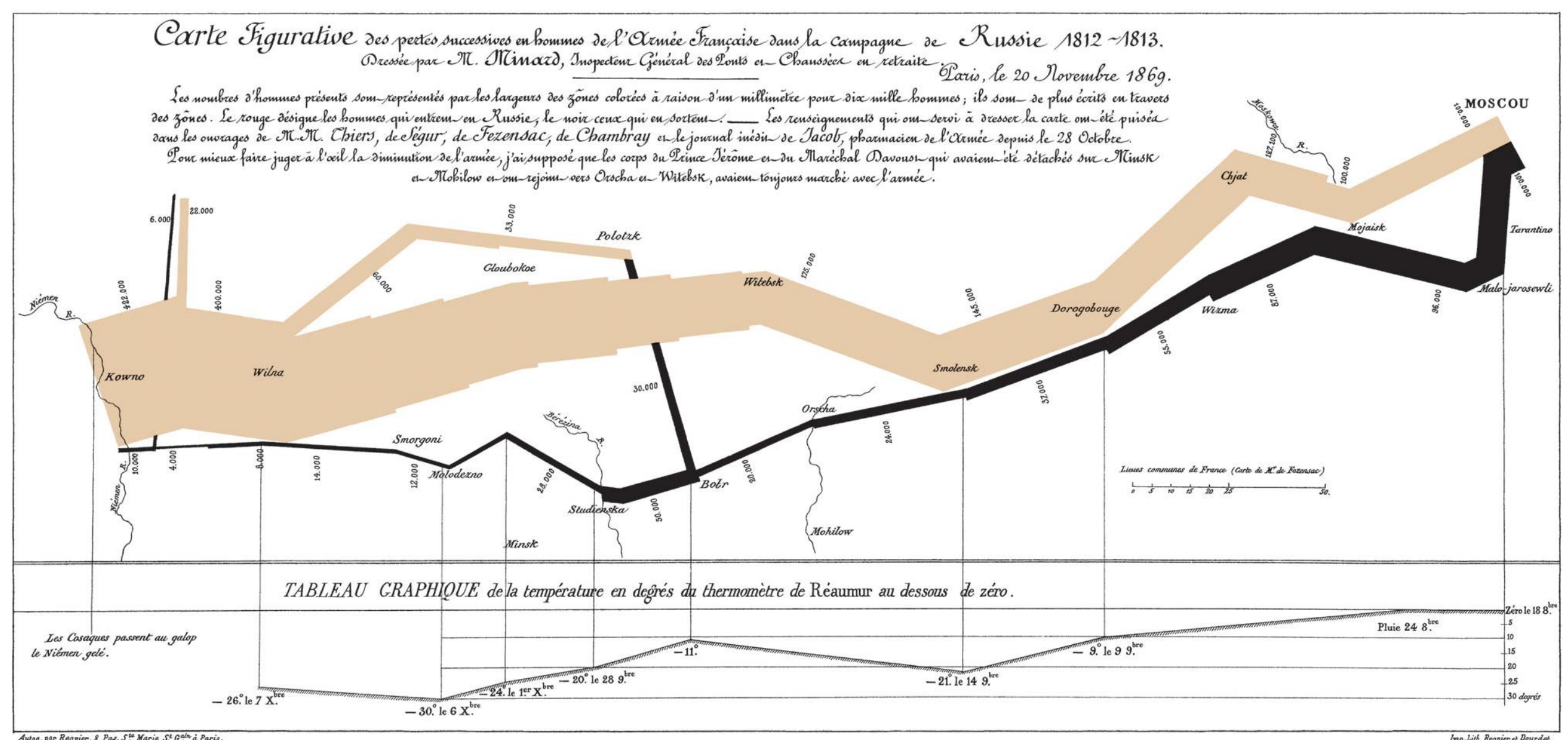


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.  
 Published as the Act above June 7<sup>th</sup> 1788 by W<sup>m</sup> Playfair  
 Made comp<sup>t</sup> 352. 1/2 Londn

WILLIAM PLAYFAIR (c. 1786)



FLORENCE NIGHTINGALE (c. 1858)



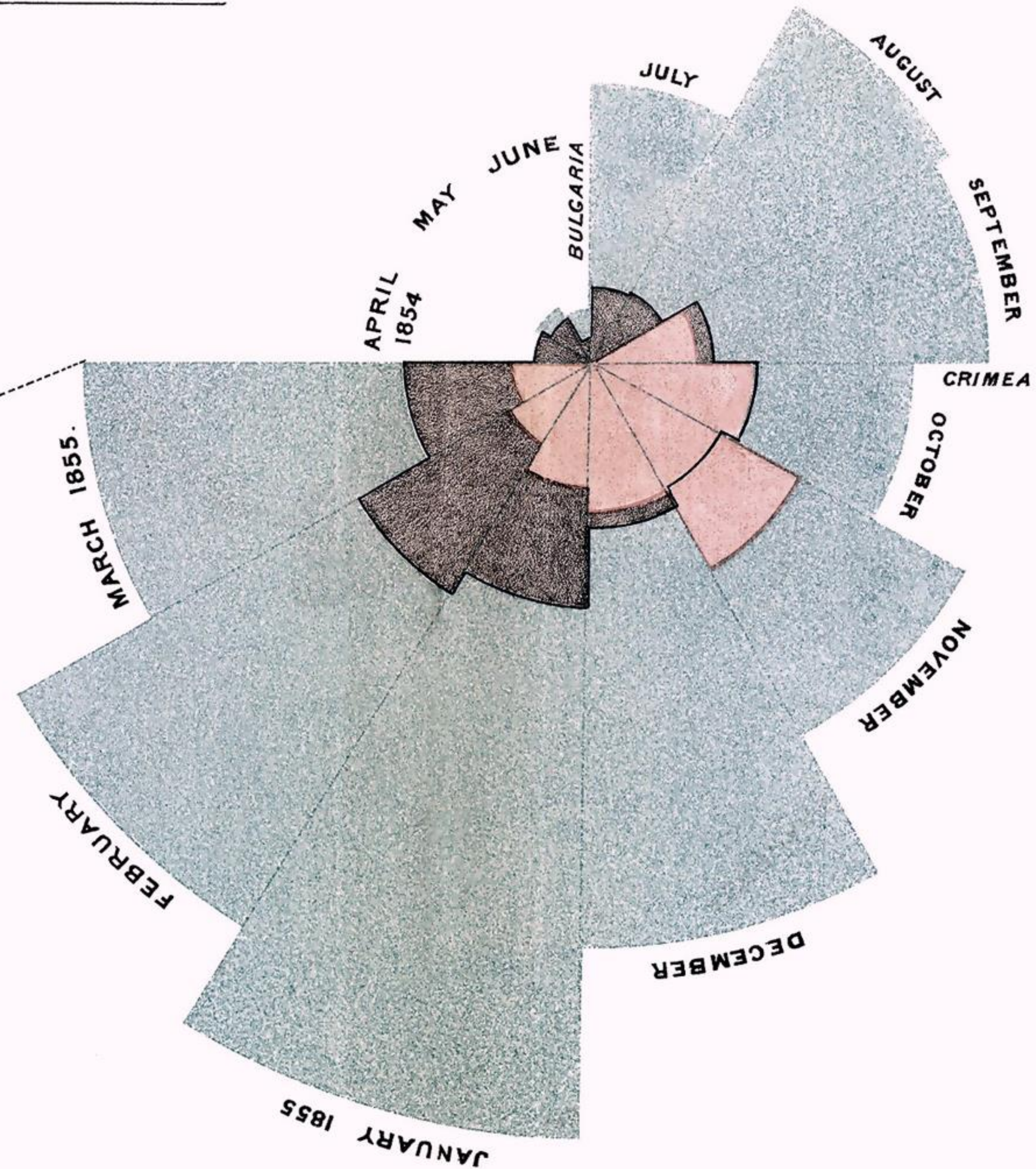
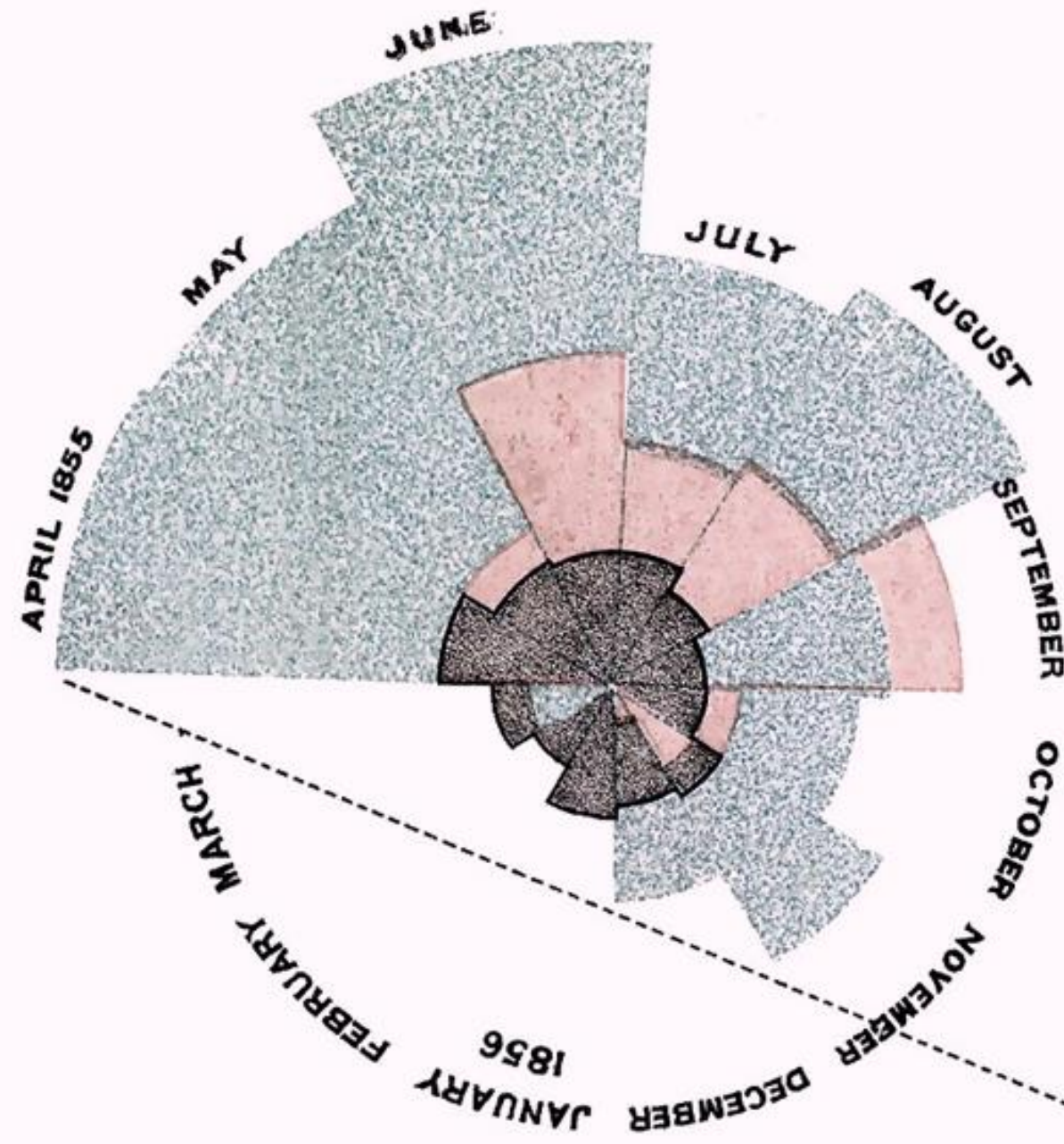
CHARLES MINARD (c. 1869)



# 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.*

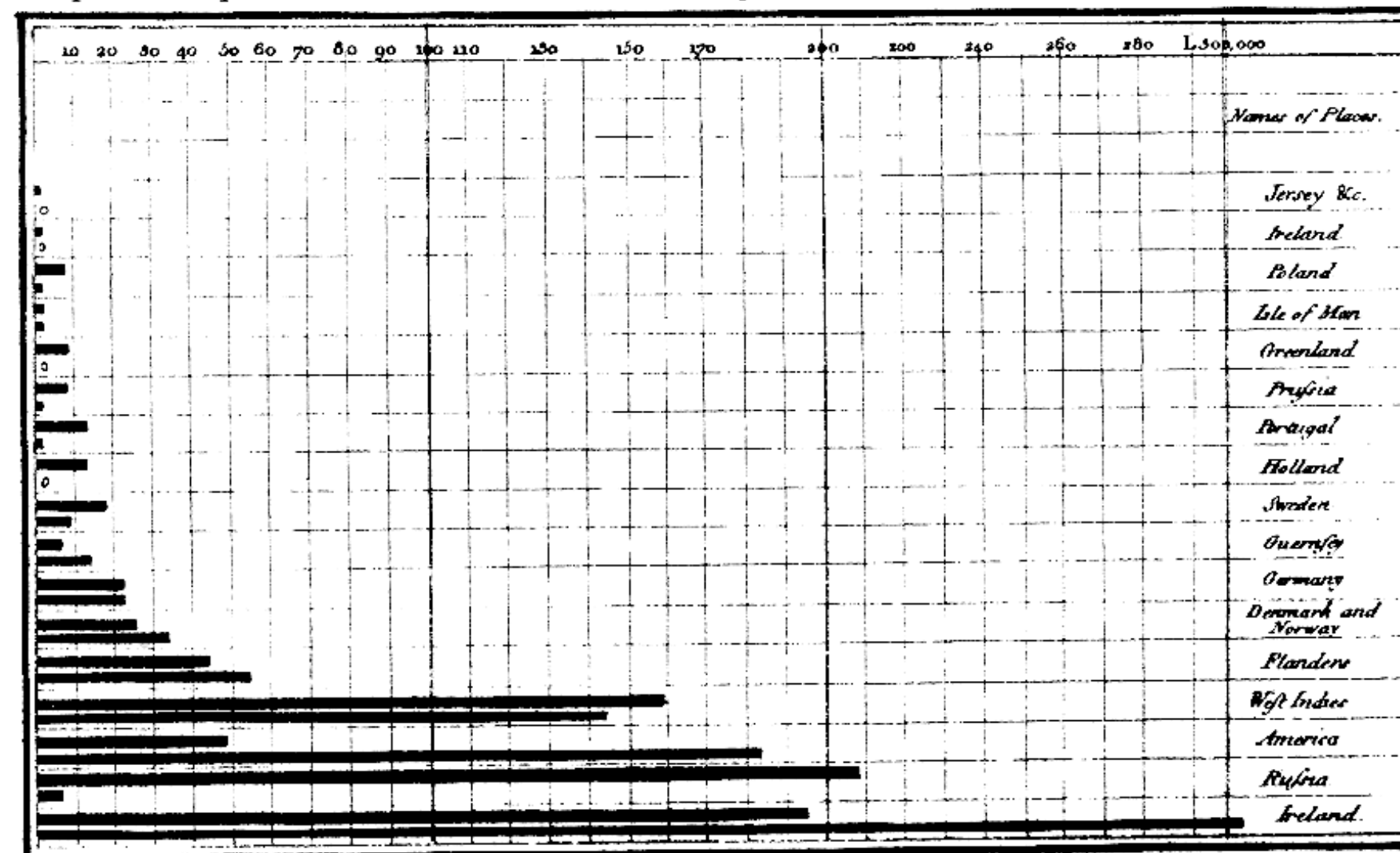
*The entire areas may be compared by following the blue, the red & the black lines enclosing them.*





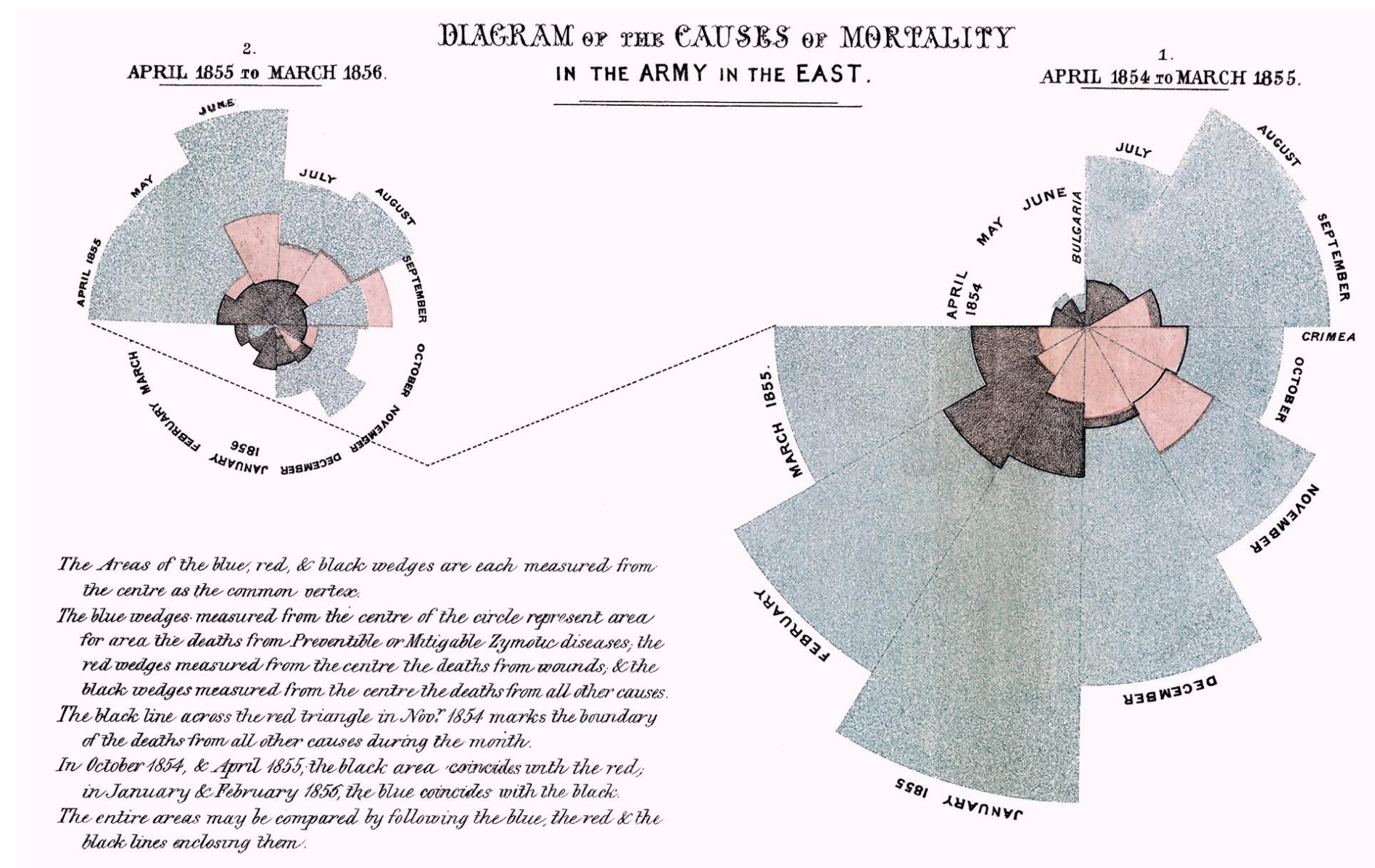
PTOLEMY (c. 150)

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

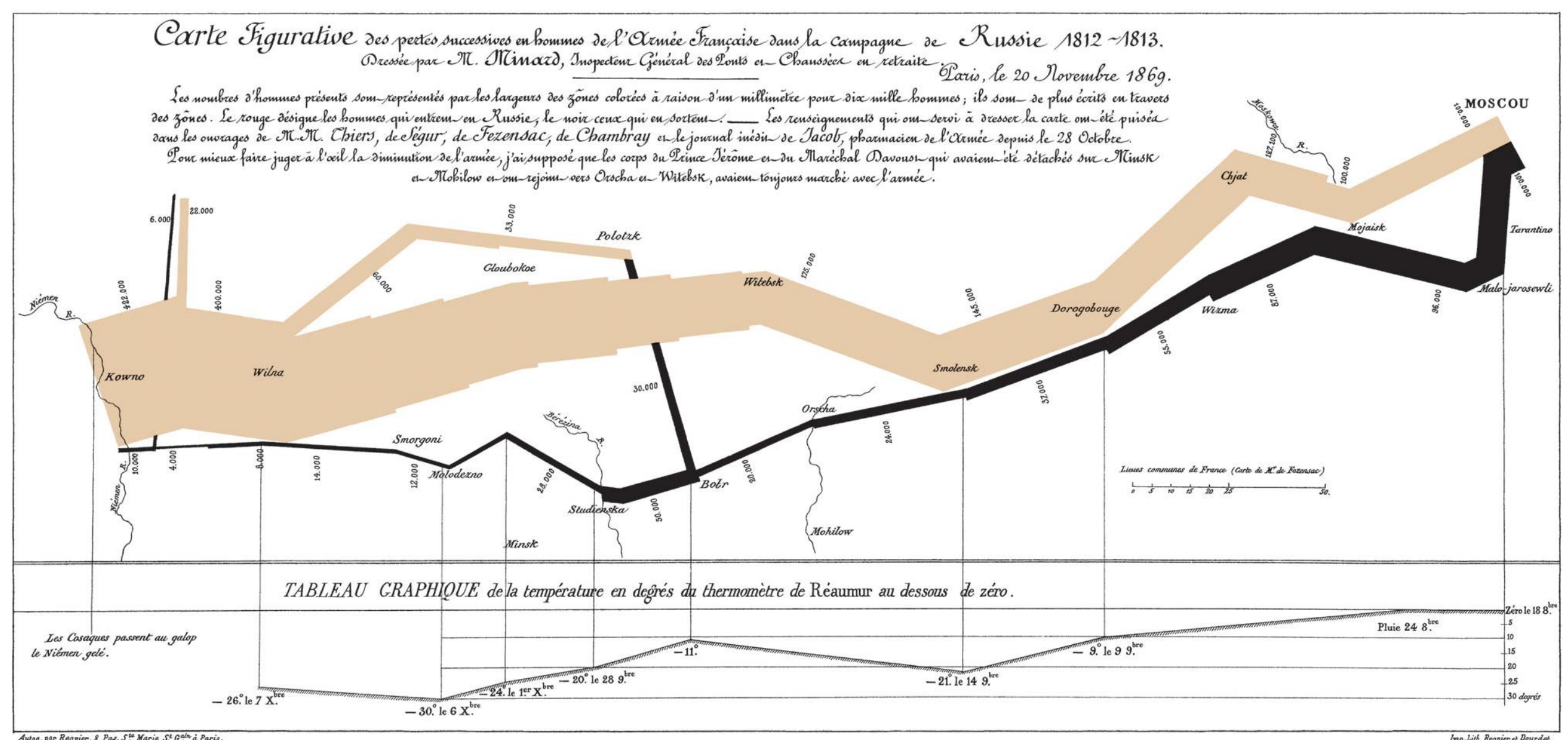


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.  
 Published as the Act above June 7<sup>th</sup> 1788 by W<sup>m</sup> Playfair  
 Made comp<sup>t</sup> 352. 1/2 Londn

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.

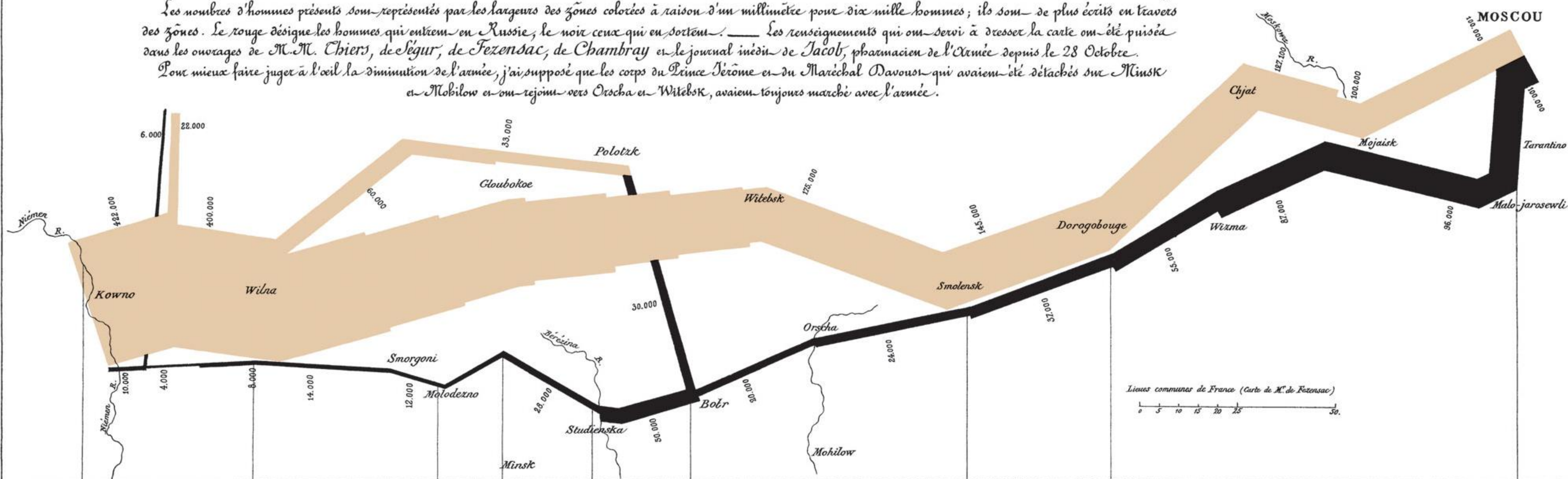
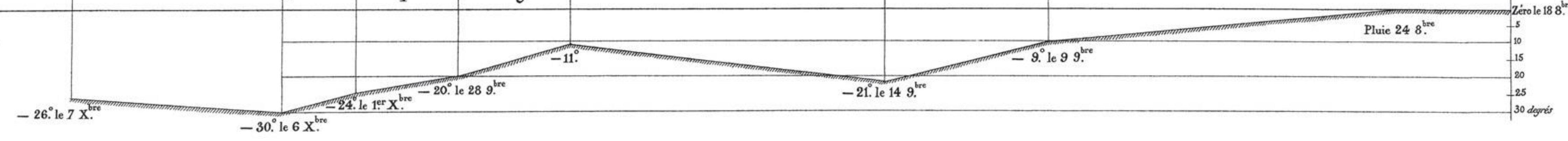


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

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.

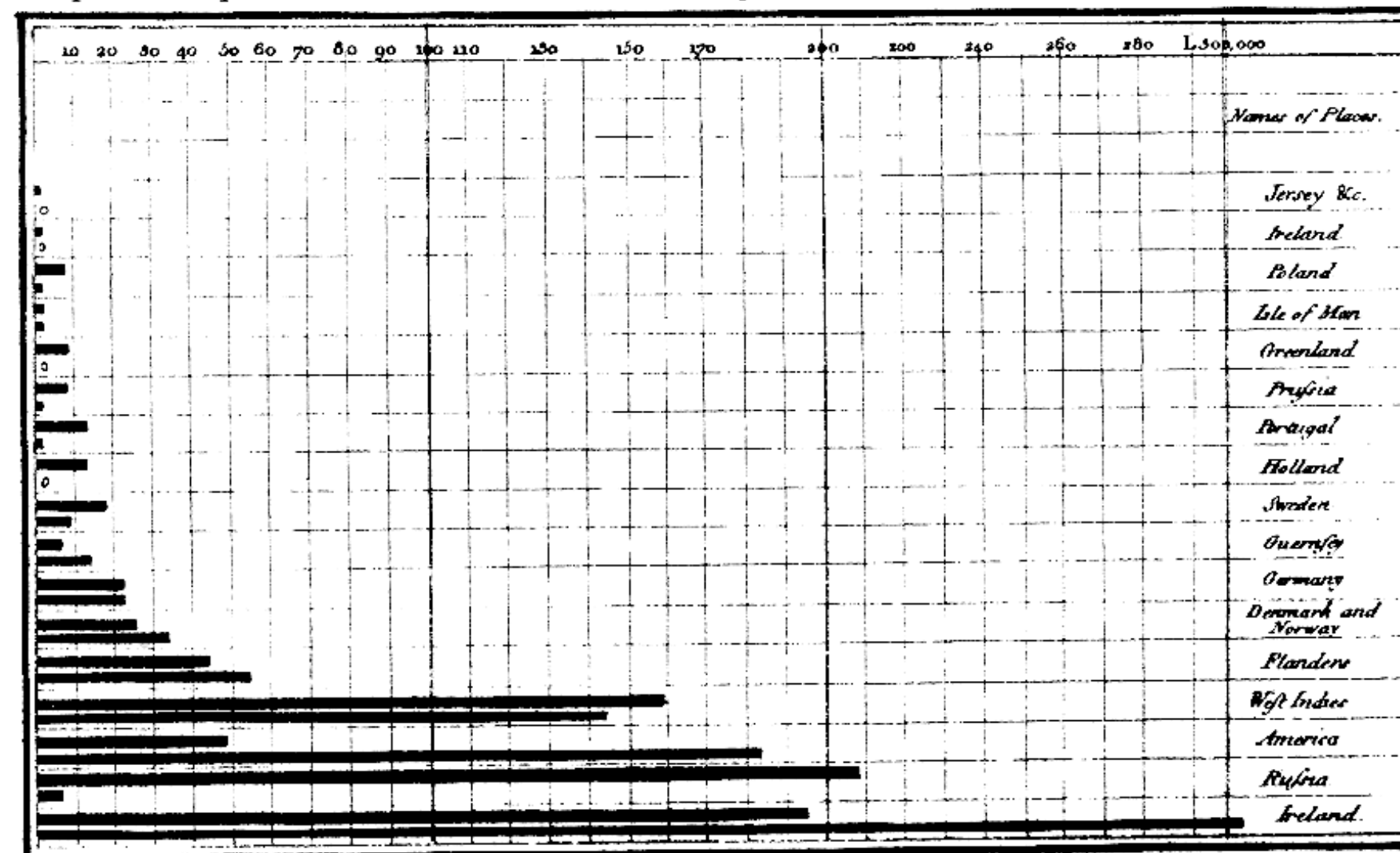
Imp. Lith. Regnier et Dourdet.





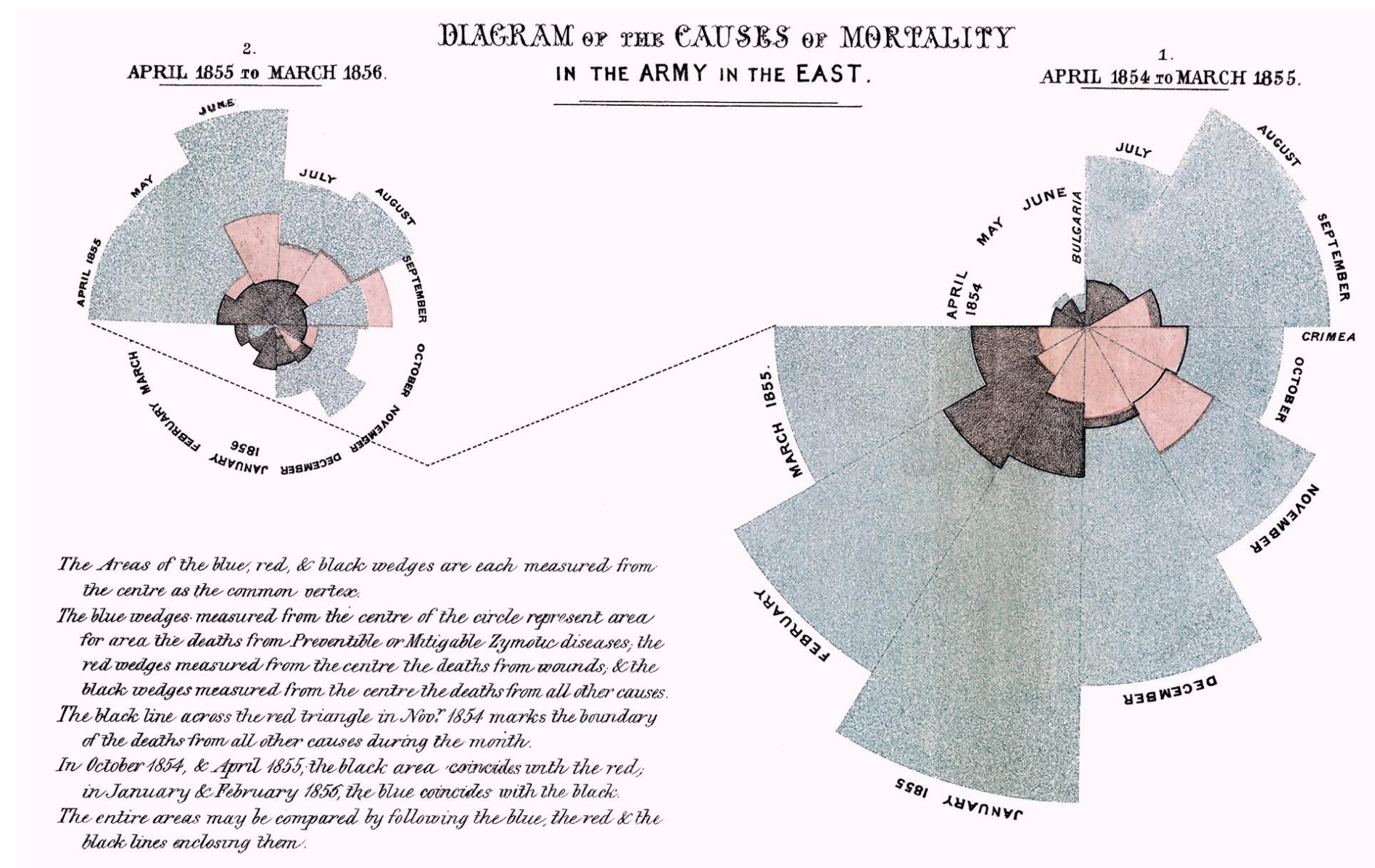
PTOLEMY (c. 150)

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

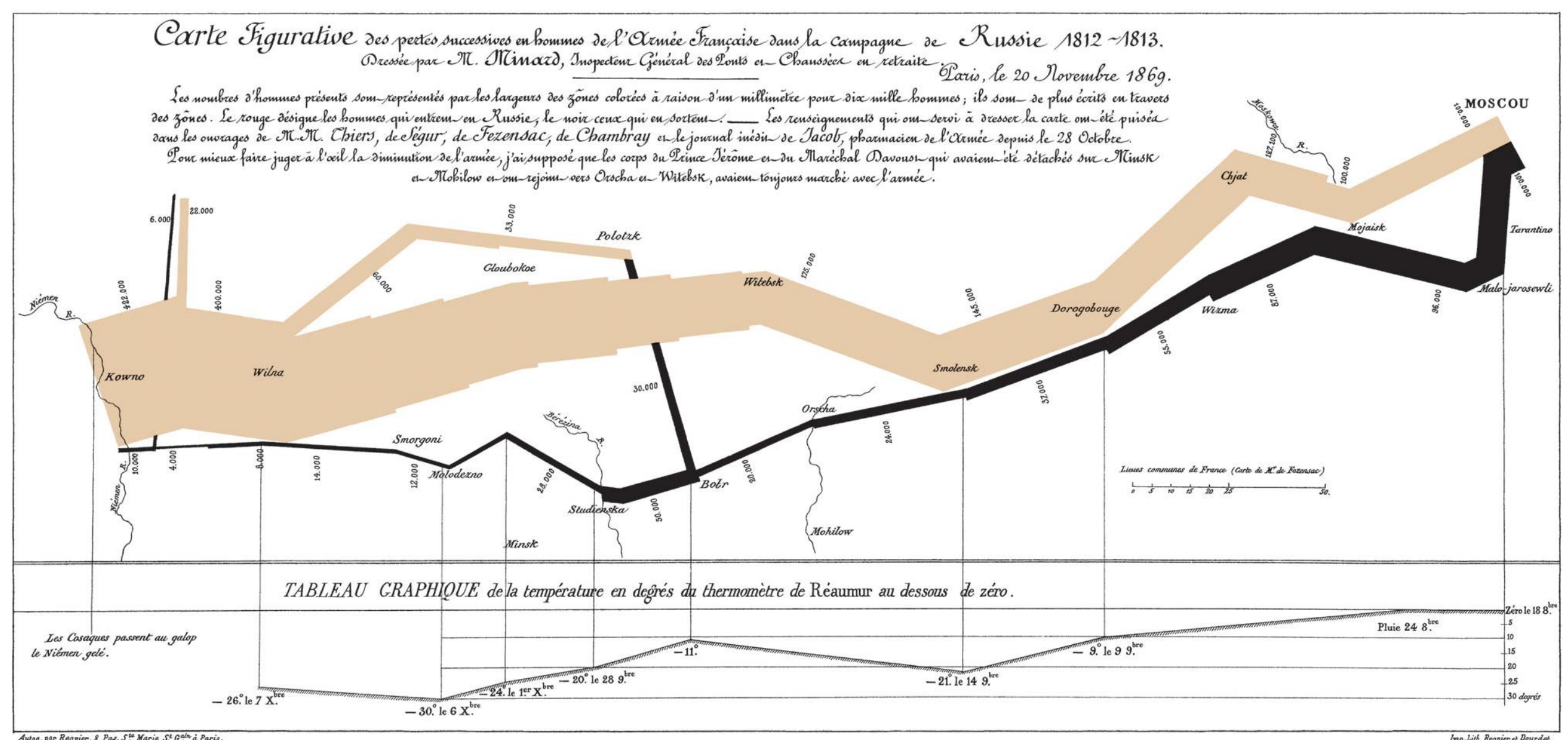


The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports.  
 Published as the Act above June 7<sup>th</sup> 1788 by W<sup>m</sup> Playfair  
 Made comp<sup>t</sup> 352. 1/2 Londn

WILLIAM PLAYFAIR (c. 1786)

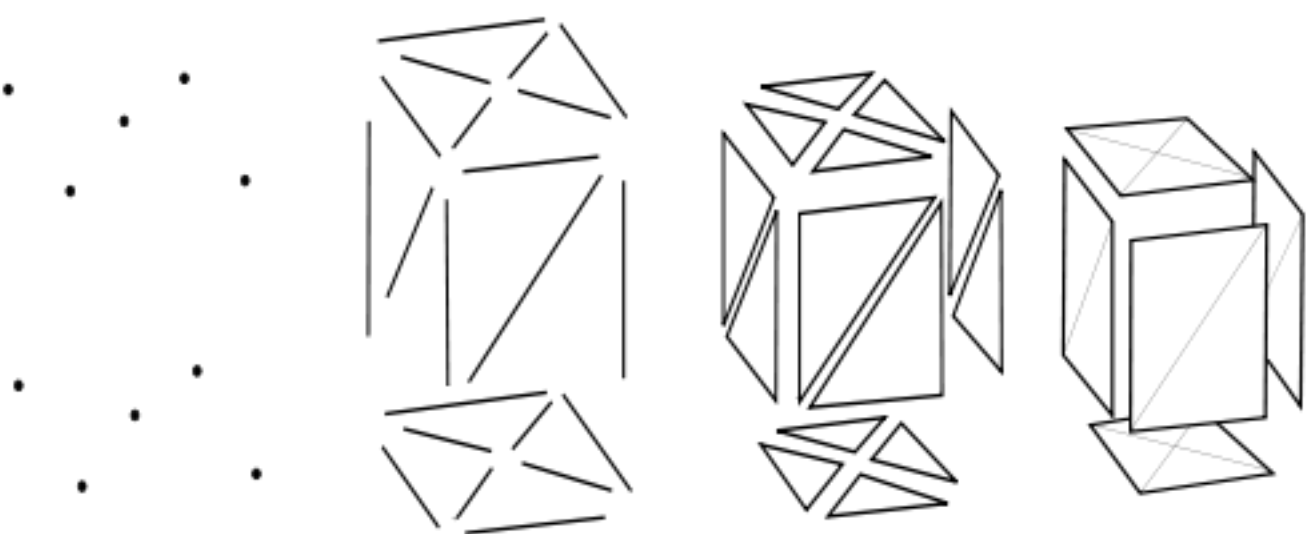


FLORENCE NIGHTINGALE (c. 1858)

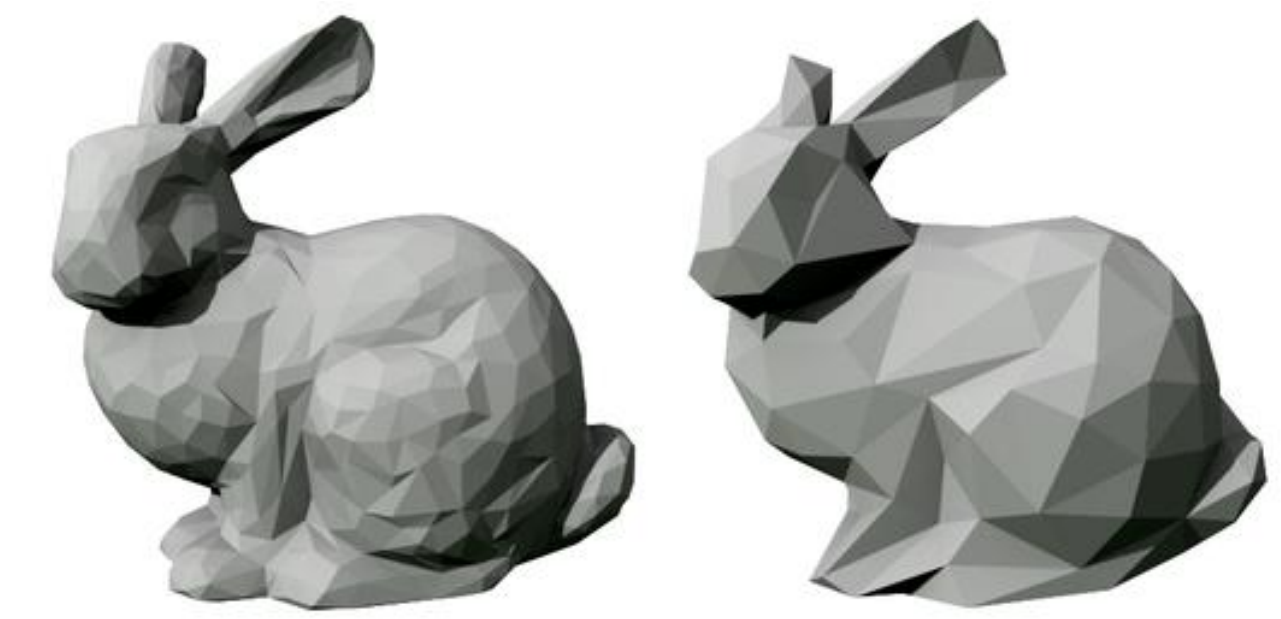


CHARLES MINARD (c. 1869) 14

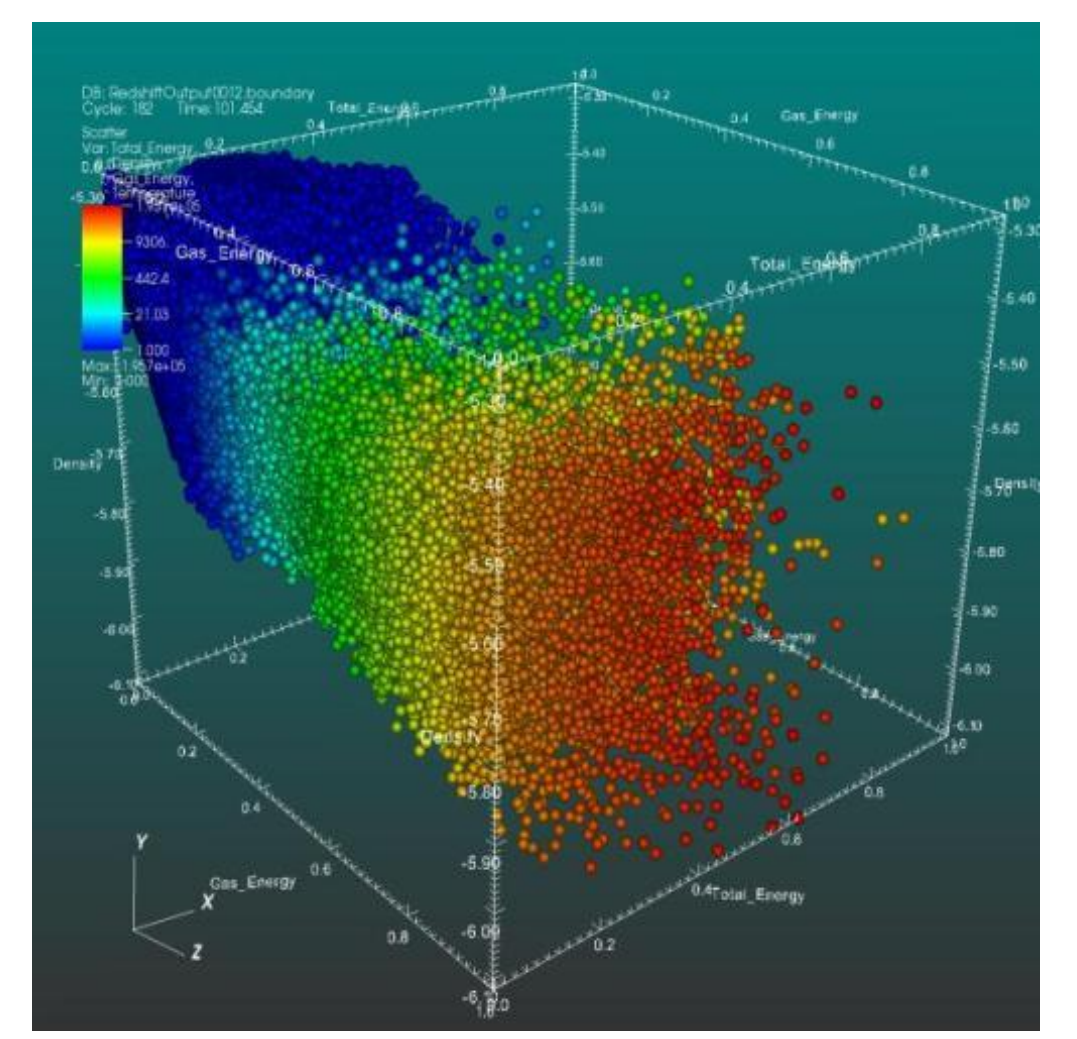
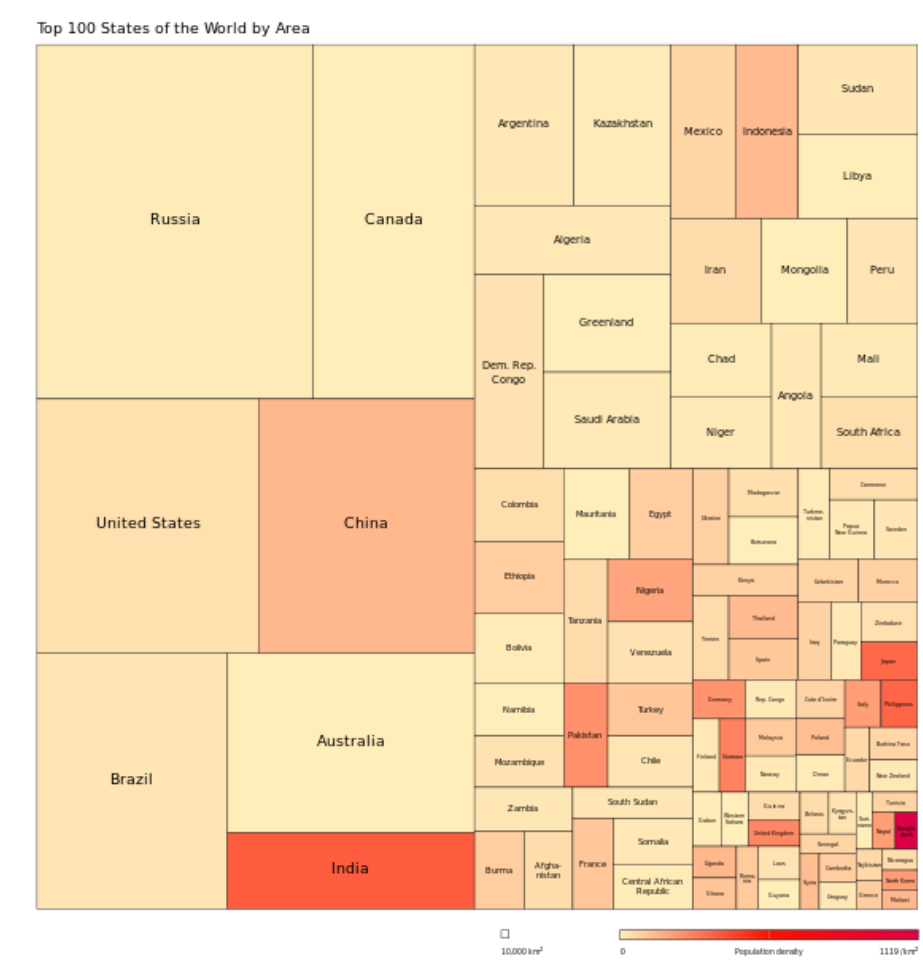
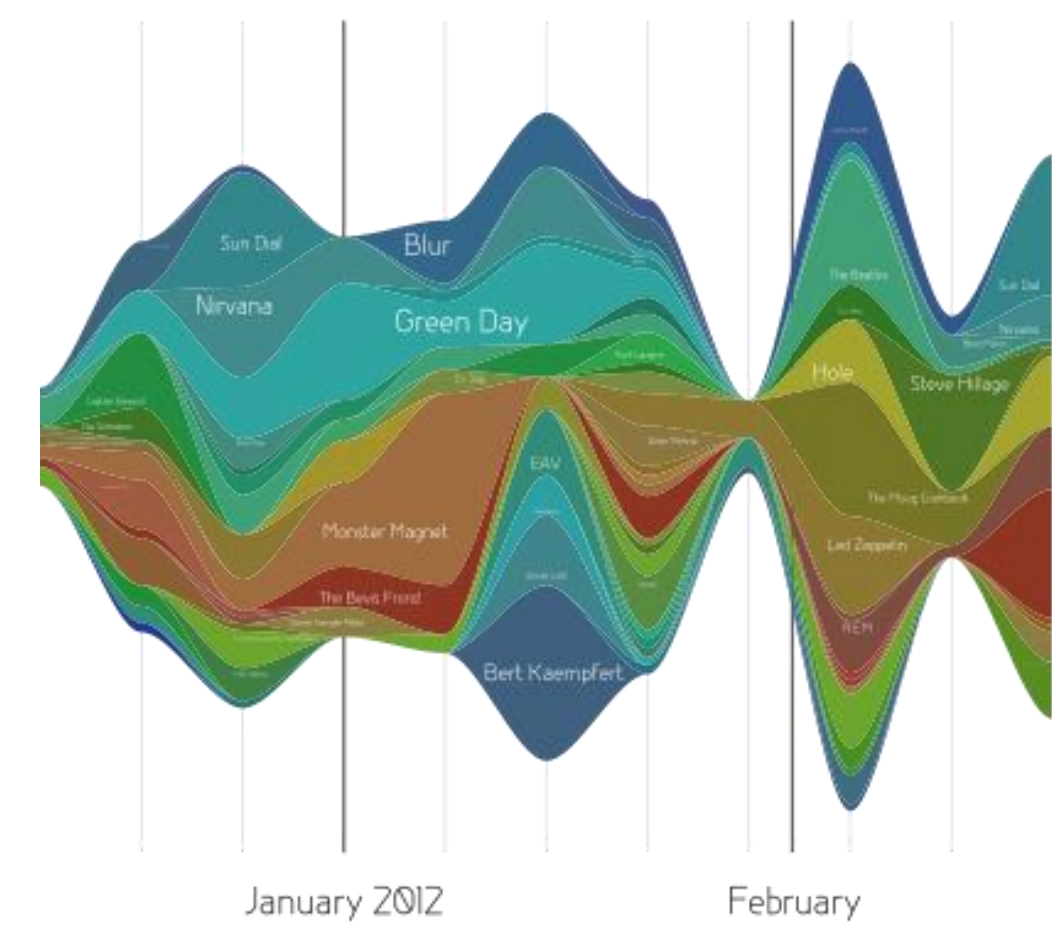
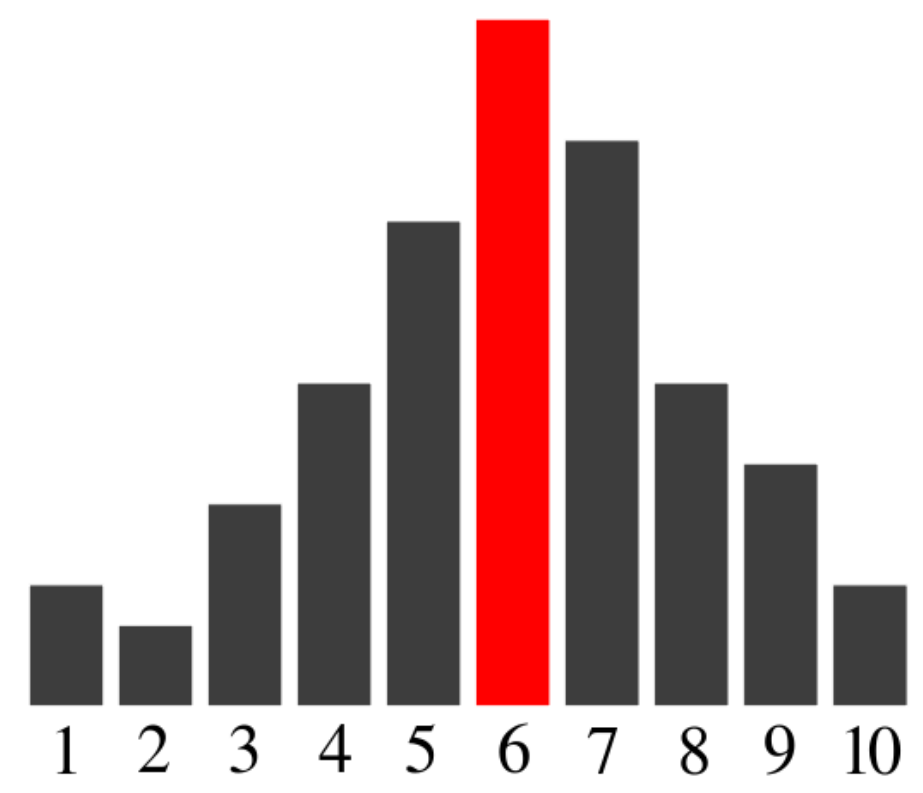




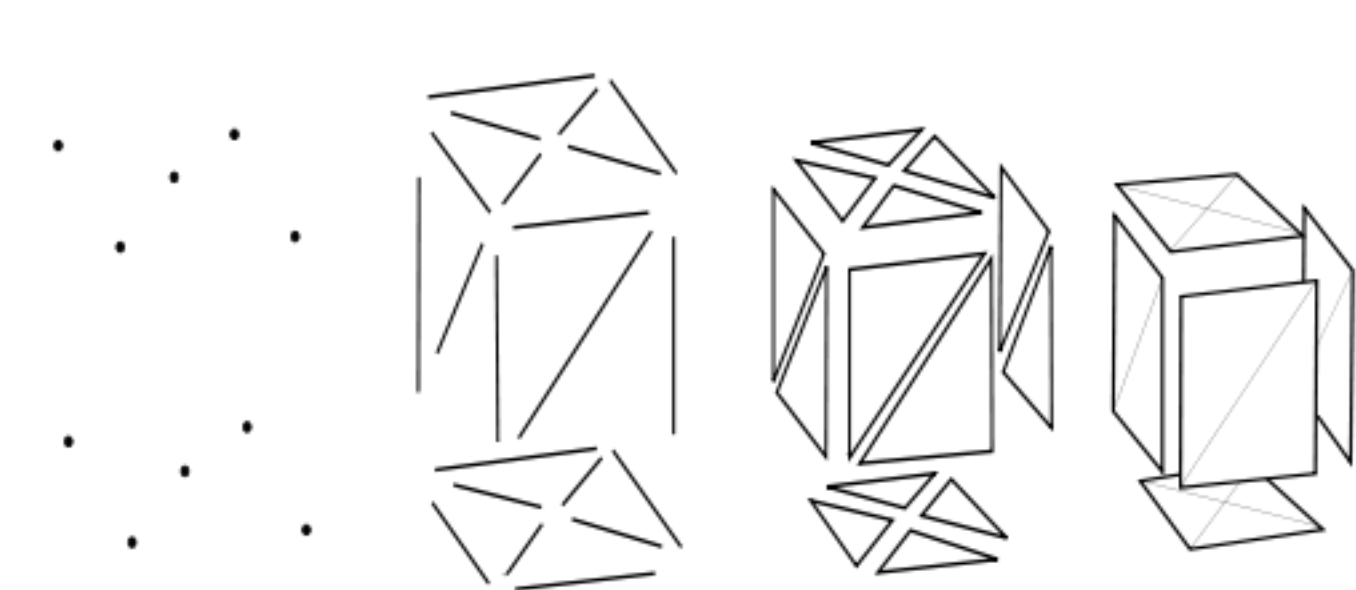
# computer graphics



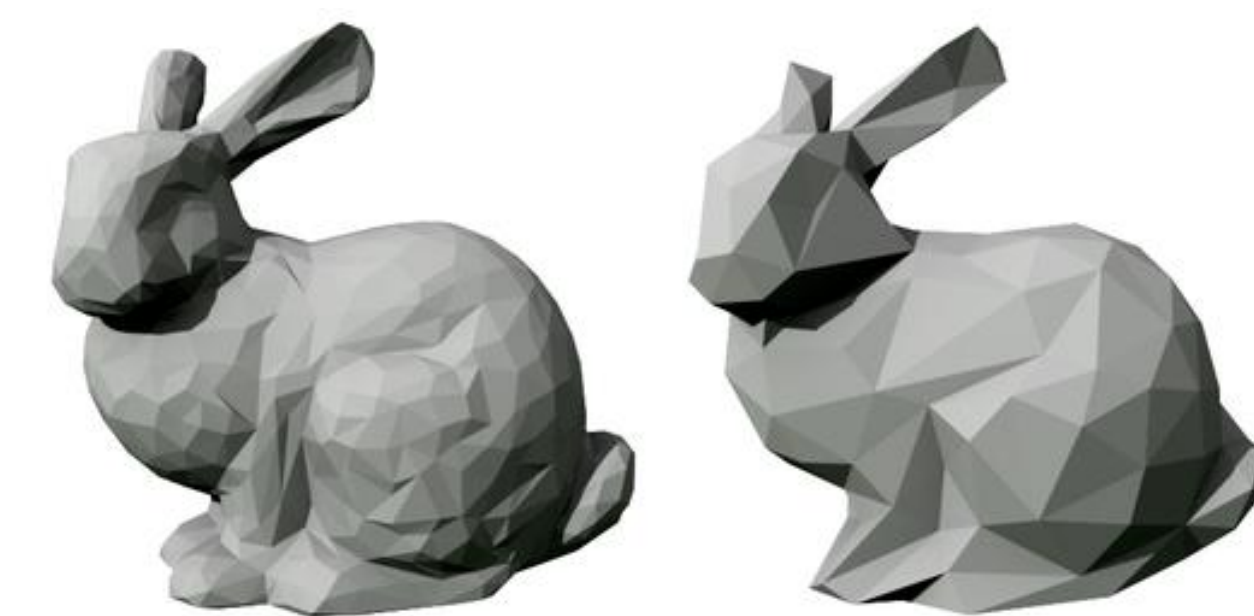
# visualization



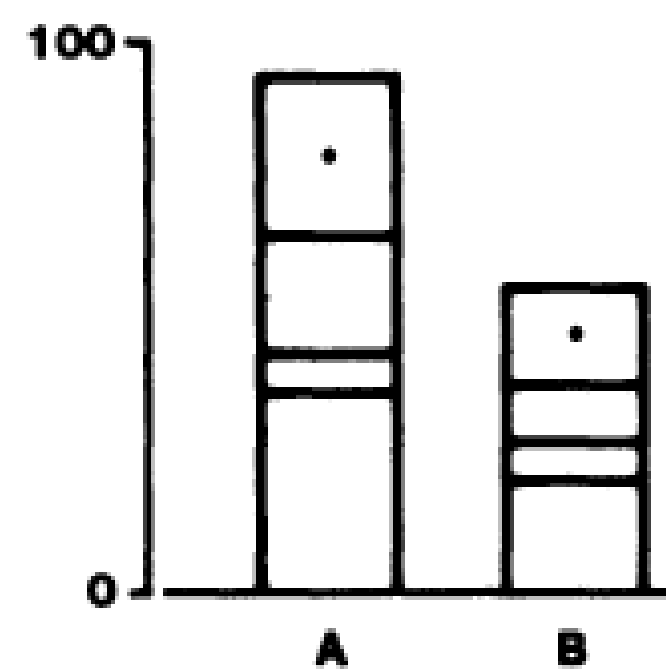
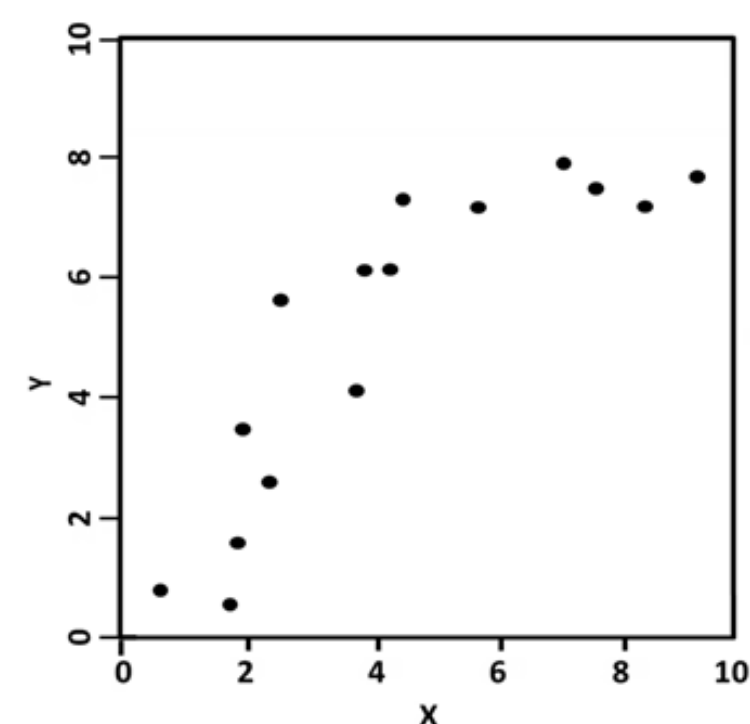




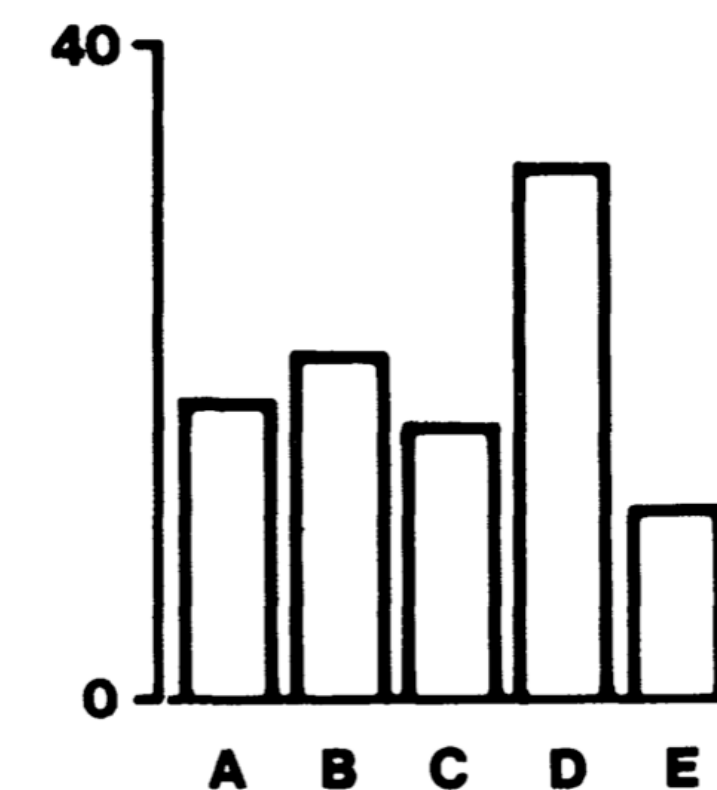
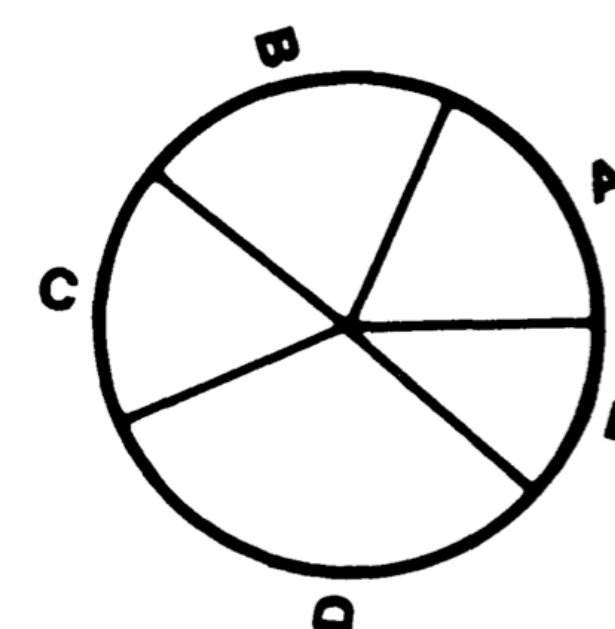
computer graphics



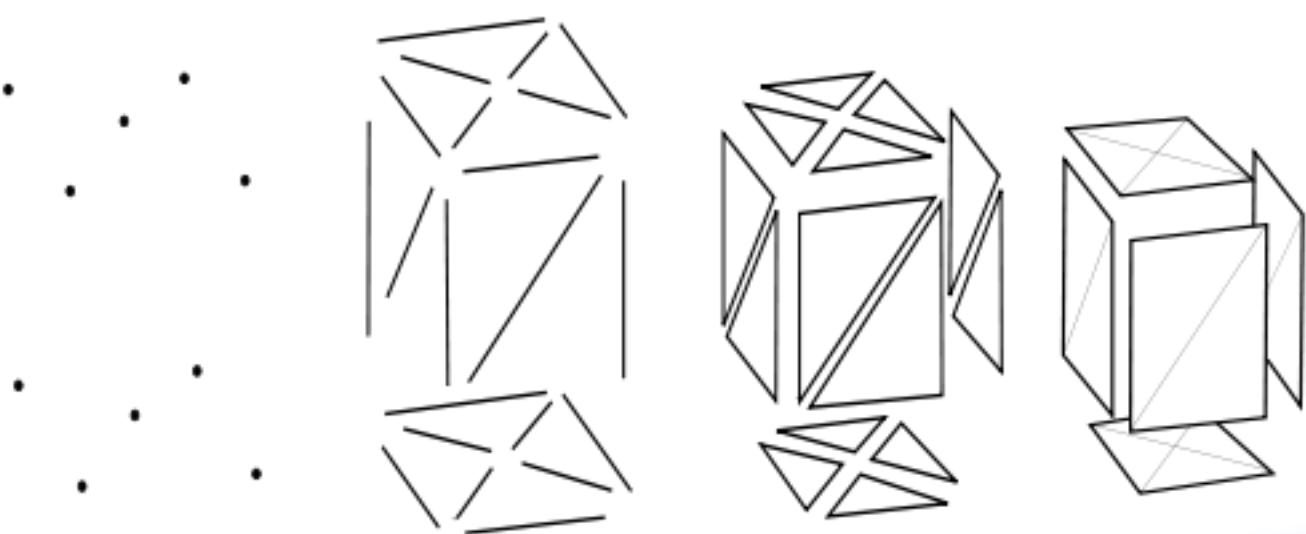
visualization



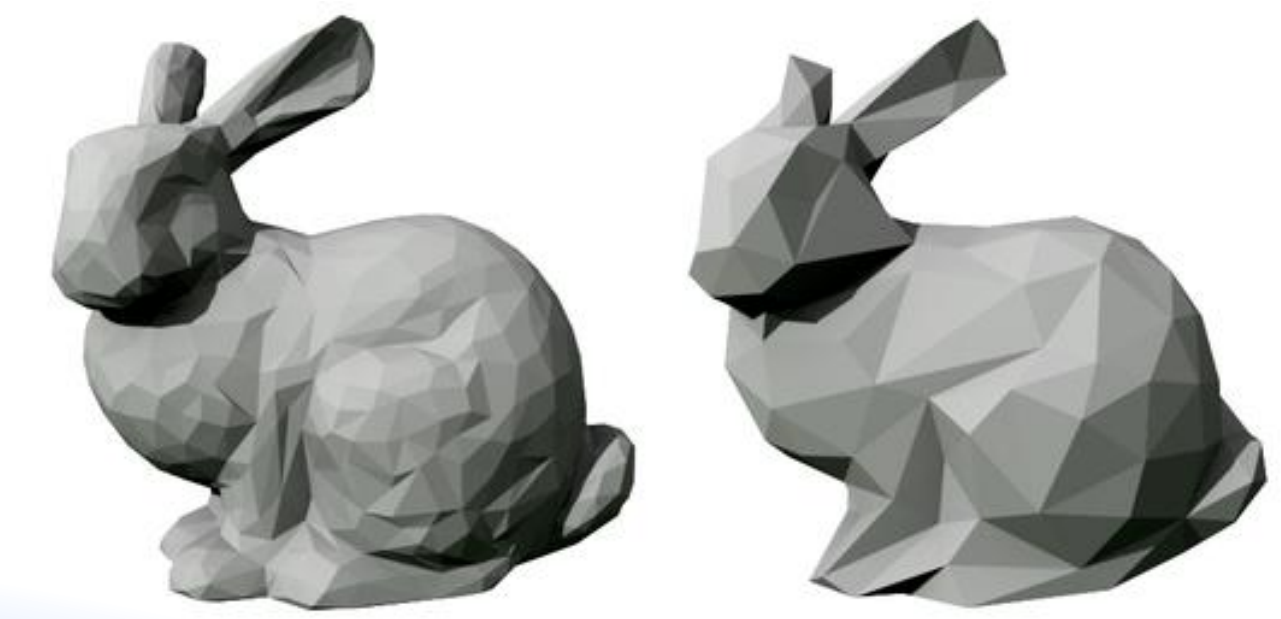
statistics







computer graphics



HCI

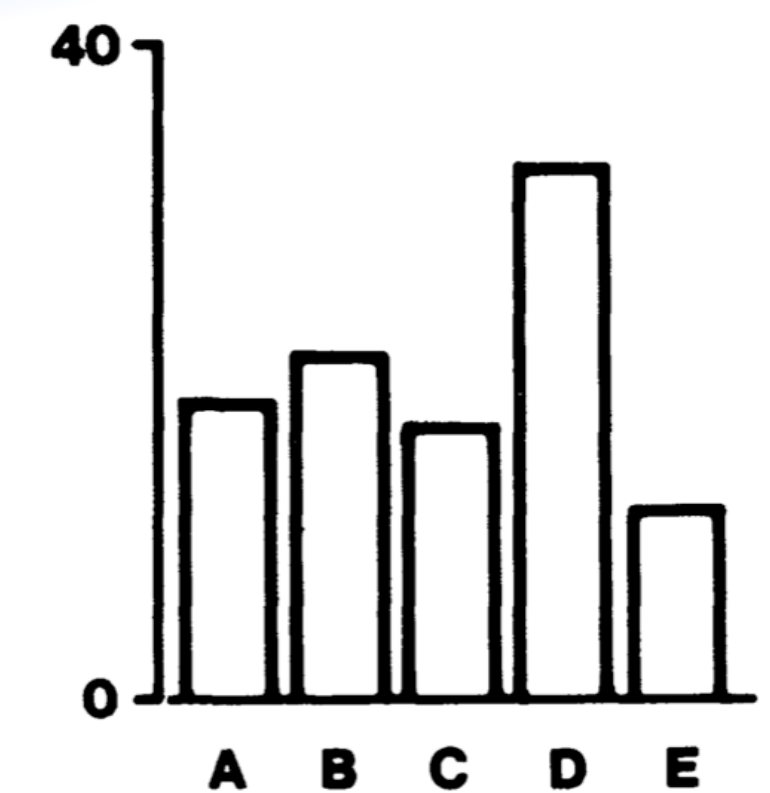
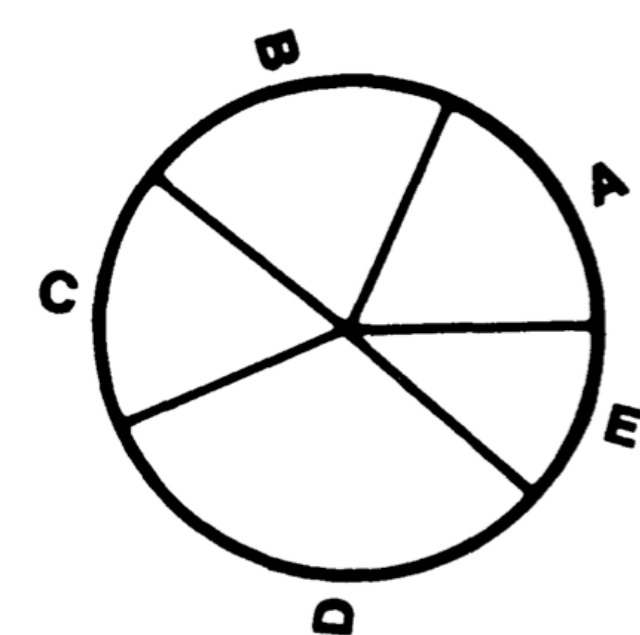
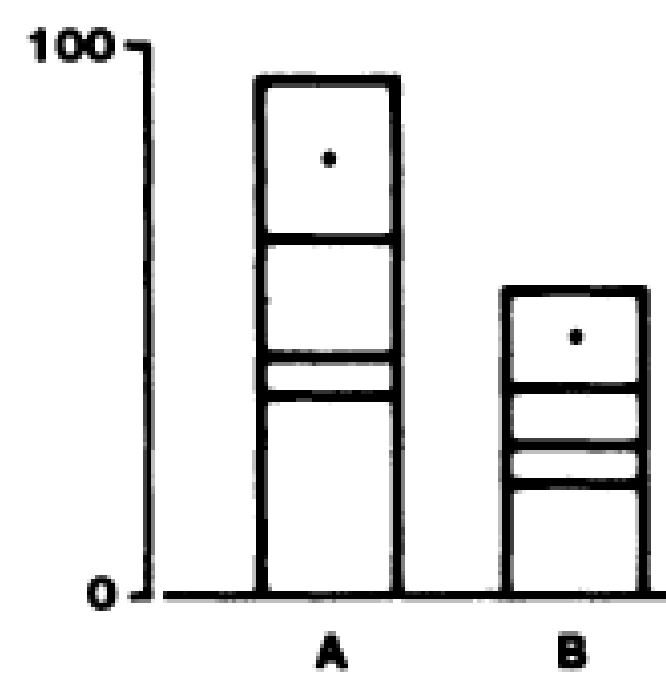
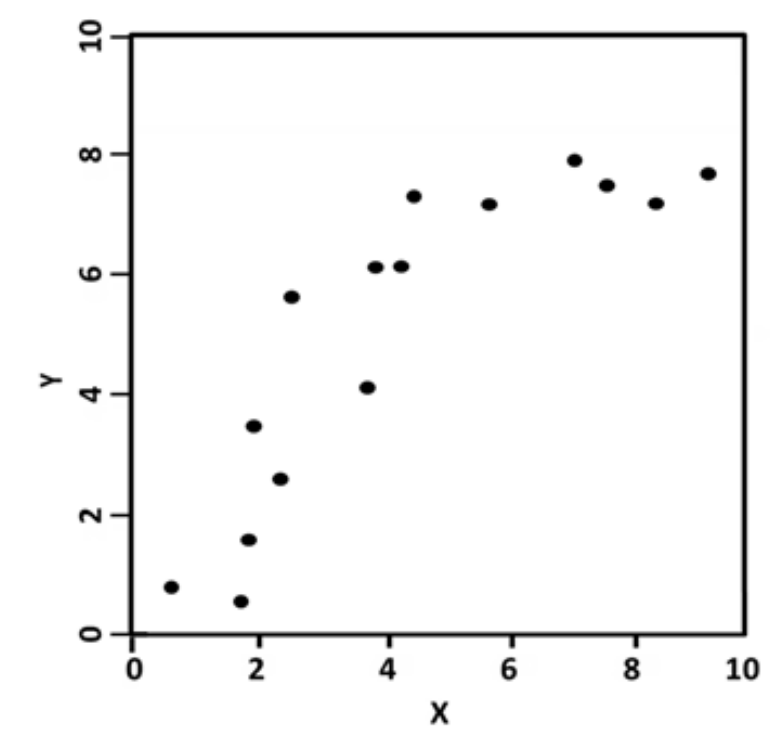
design

visualization

psychology

art

statistics

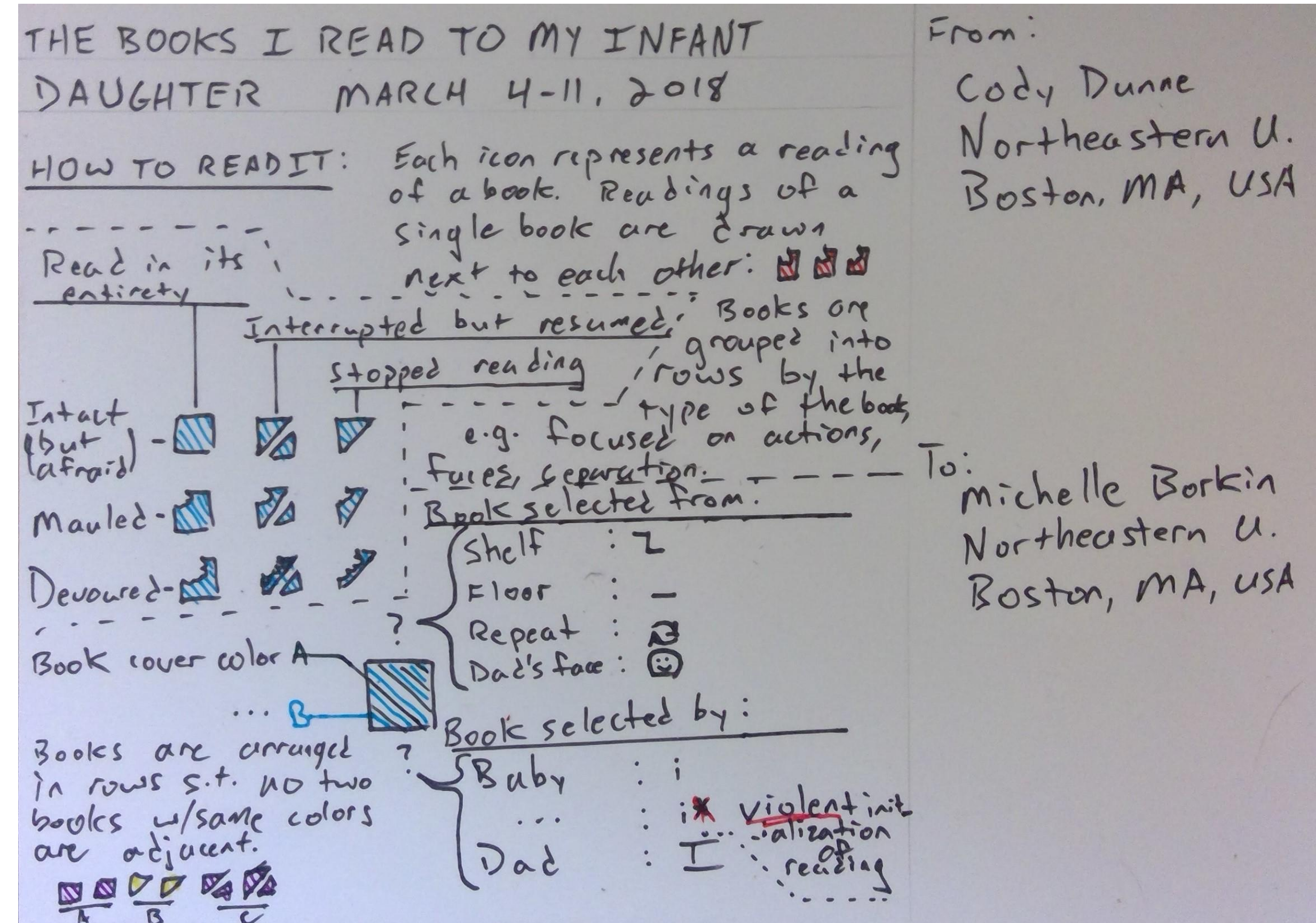
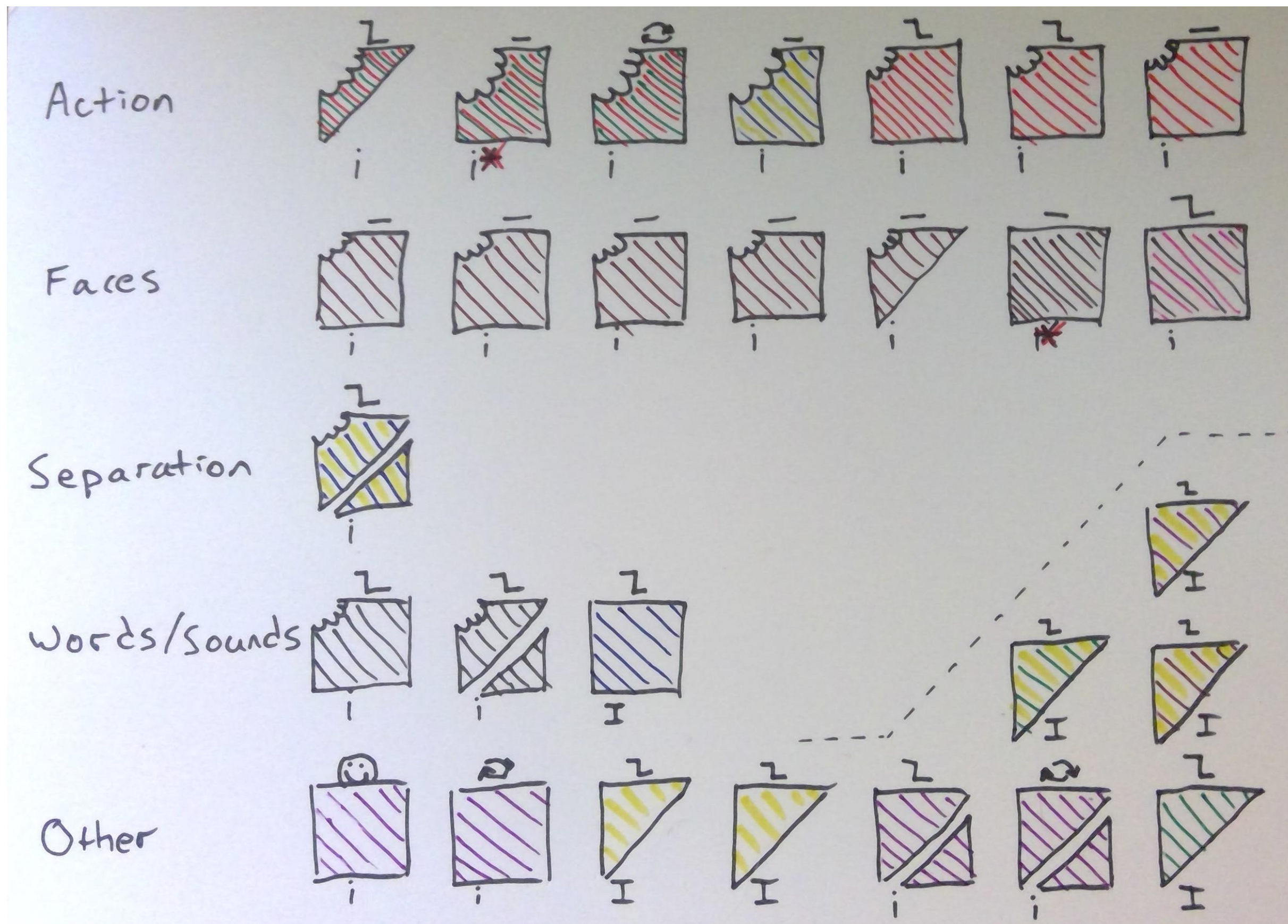




# In-Class Sketching — Visualize Me

15 min

<https://northeastern.instructure.com/courses/63405/assignments/874496>





Ok, but why do we need  
visualization?

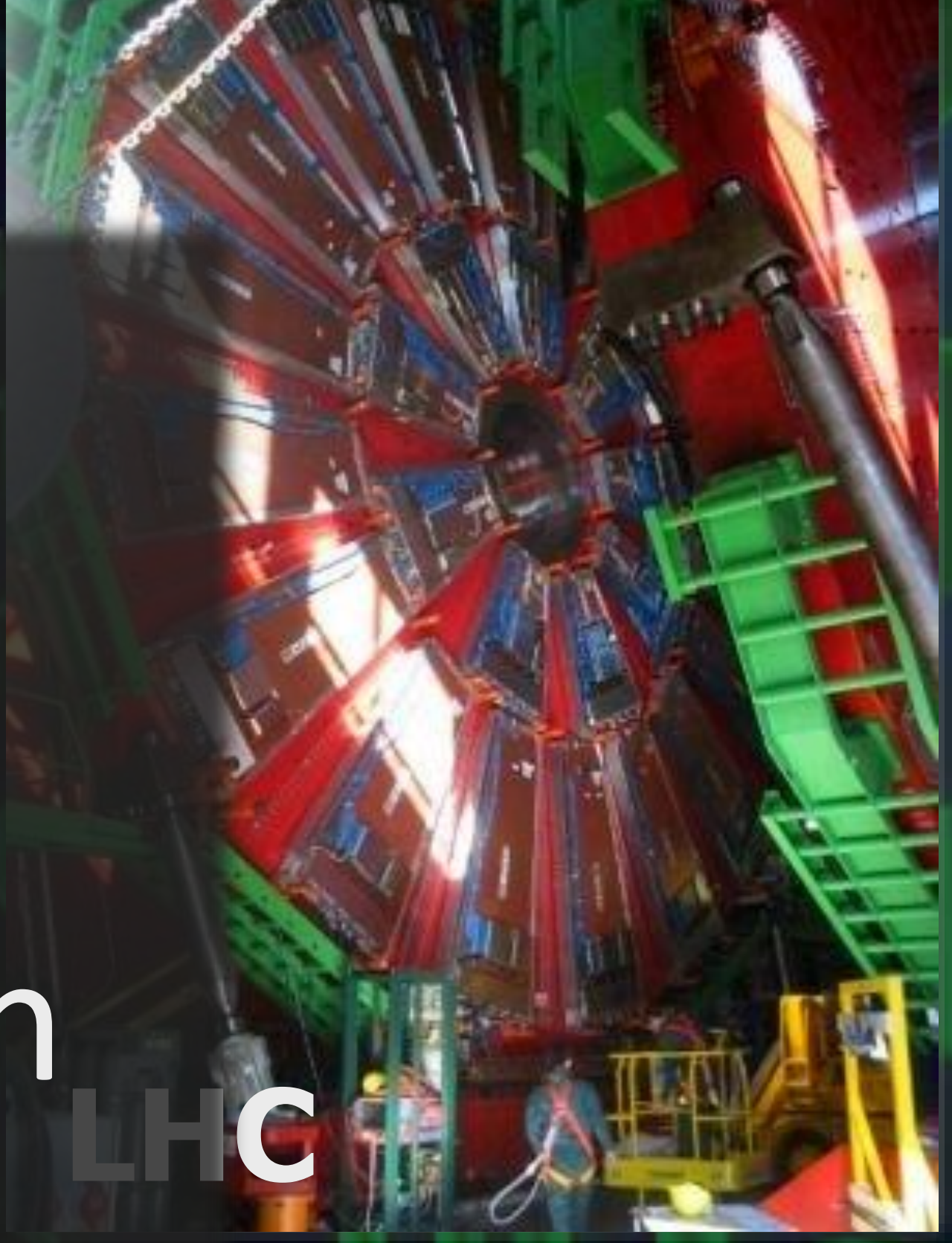




LSST



Financial Markets



LHC

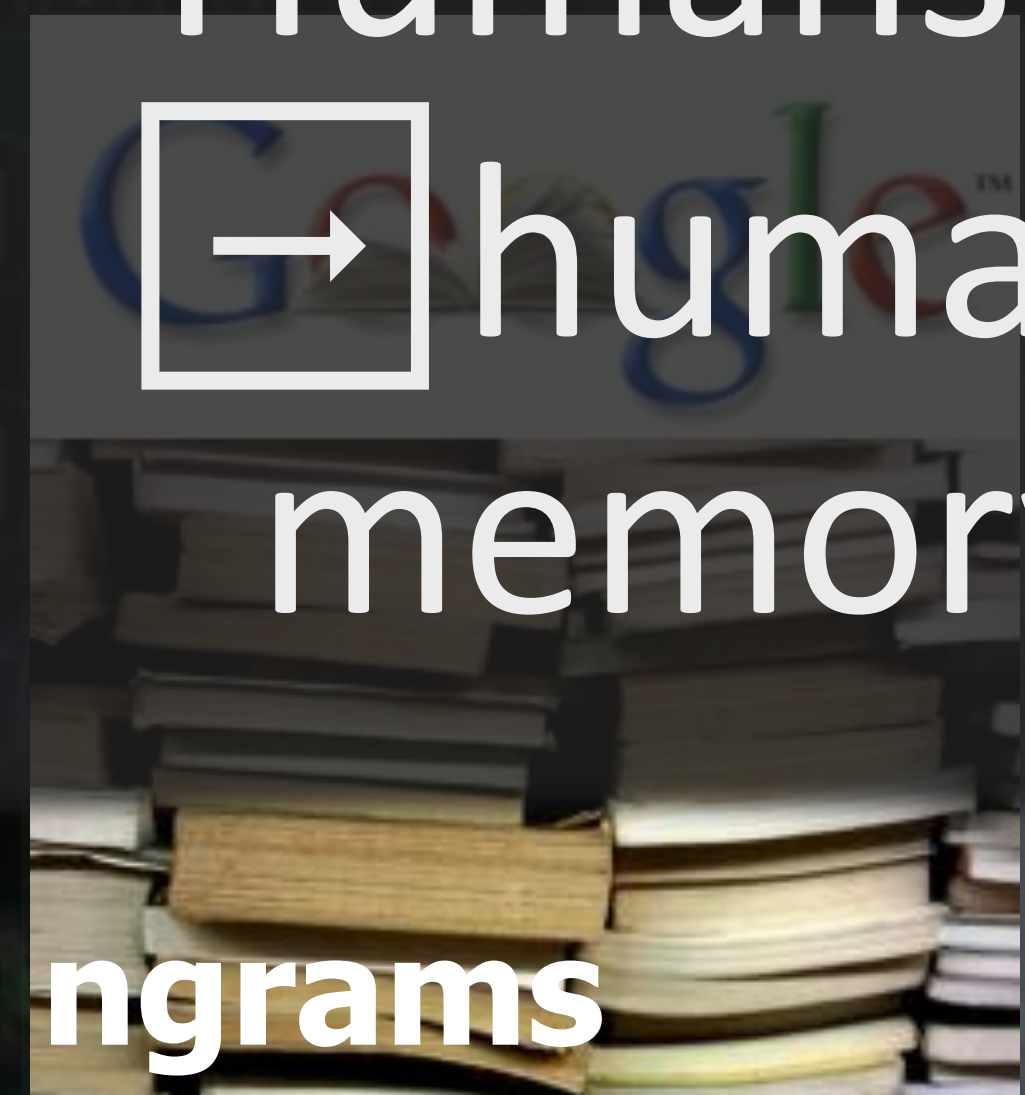


GenBank

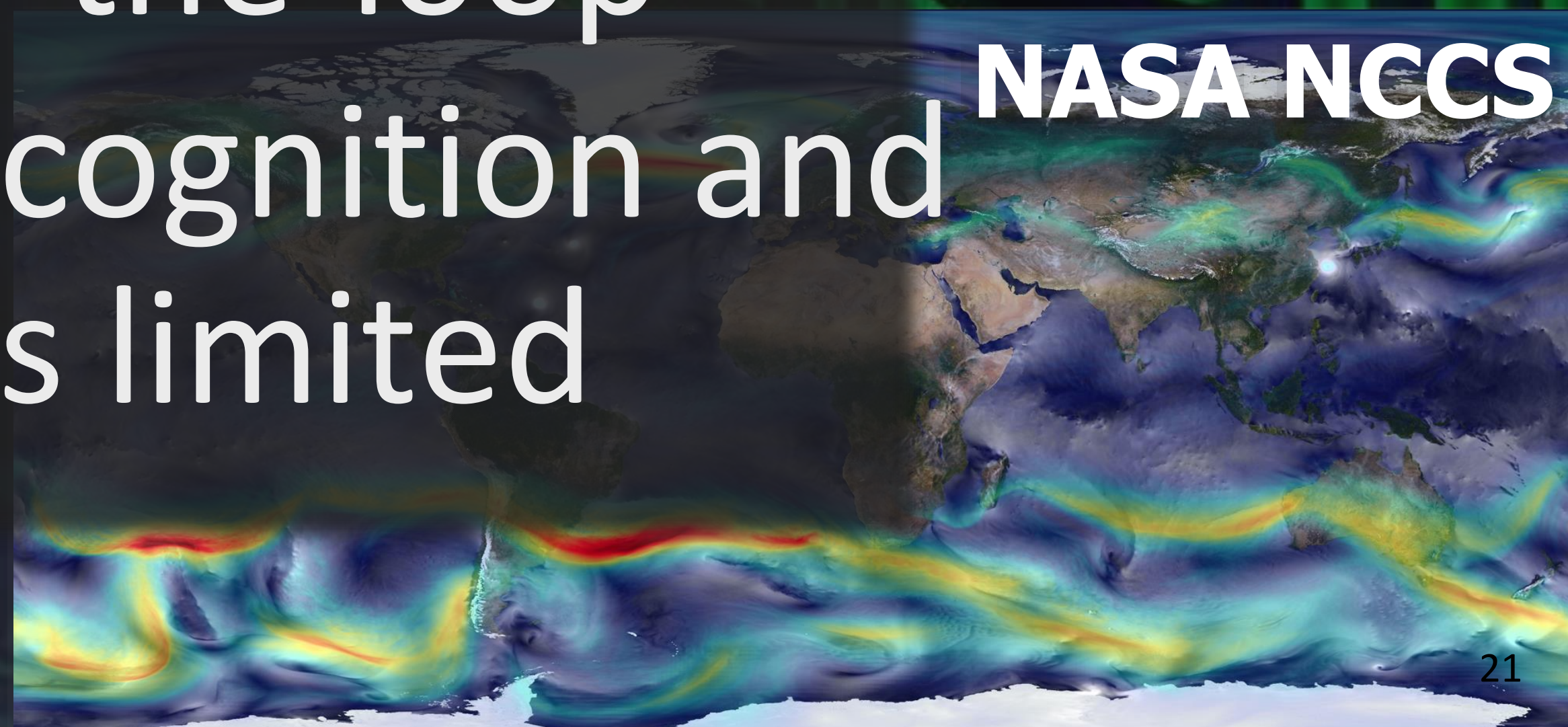
# CHALLENGES:

- Scalability
- Complexity reduction
- Humans in-the-loop

→ human cognition and memory is limited



ngrams



NASA NCCS



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>



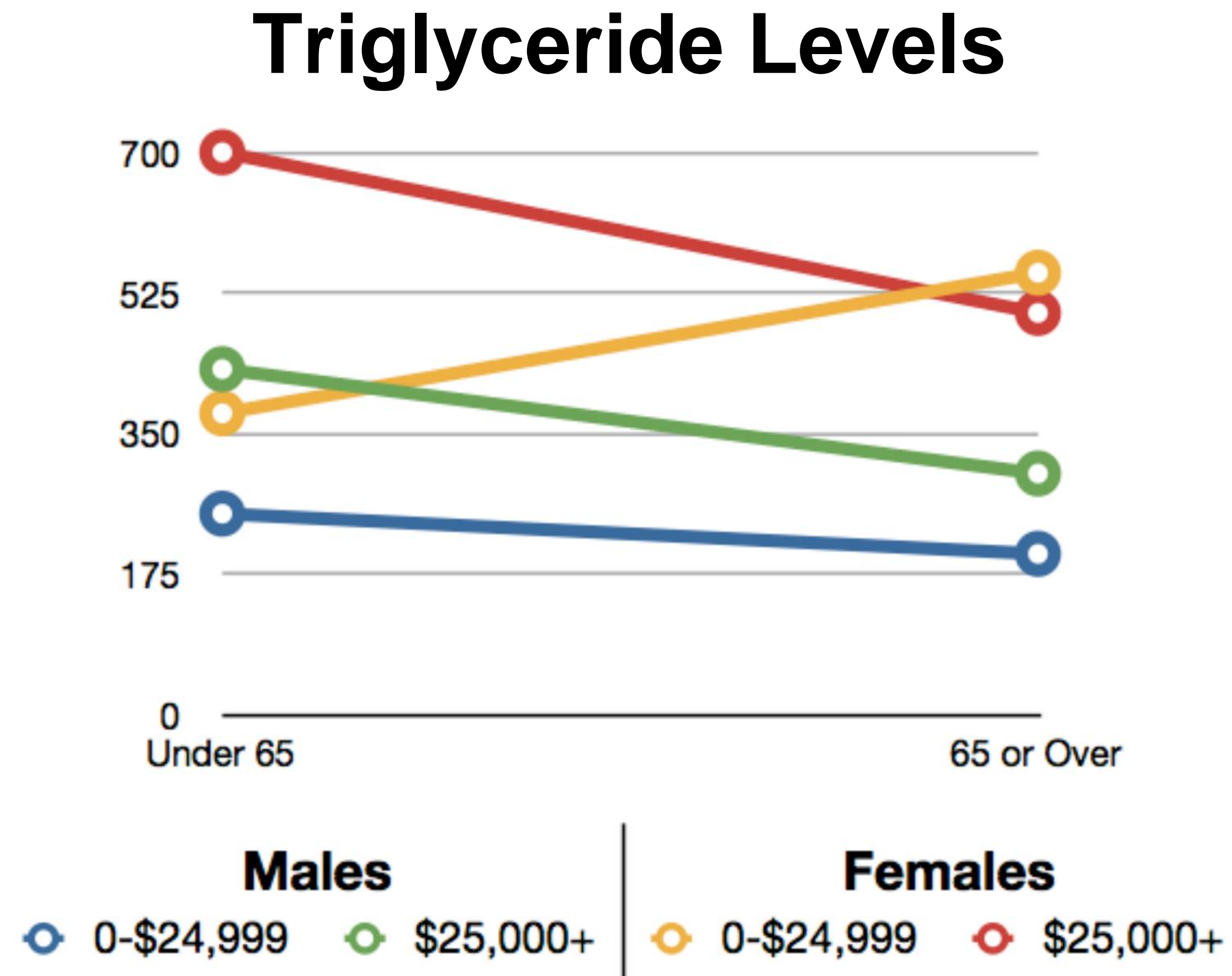
Which gender and income level shows a different effect of age on triglyceride levels?

## Triglyceride Levels

	Males		Females	
Income Group	Under 65	65 or Over	Under 65	65 or Over
0-\$24,999	250	200	375	550
\$25,000+	430	300	700	500



Which gender and income level shows a different effect of age on triglyceride levels?



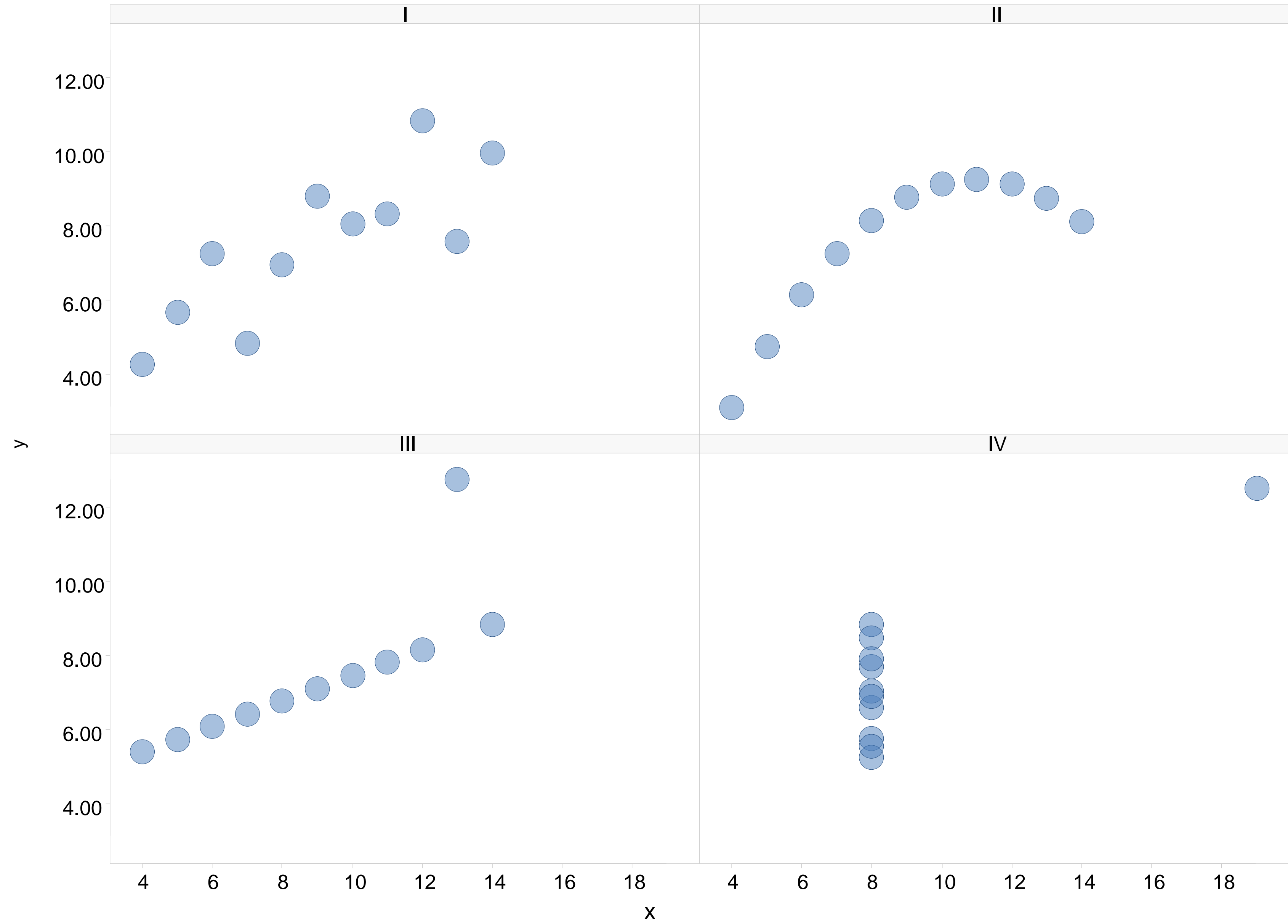


I		II		III		IV	
x	y	x	y	x	y	x	y
10.00	8.04	10.00	9.14	10.00	7.46	8.00	6.58
8.00	6.95	8.00	8.14	8.00	6.77	8.00	5.76
13.00	7.58	13.00	8.74	13.00	12.74	8.00	7.71
9.00	8.81	9.00	8.77	9.00	7.11	8.00	8.84
11.00	8.33	11.00	9.26	11.00	7.81	8.00	8.47
14.00	9.96	14.00	8.10	14.00	8.84	8.00	7.04
6.00	7.24	6.00	6.13	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



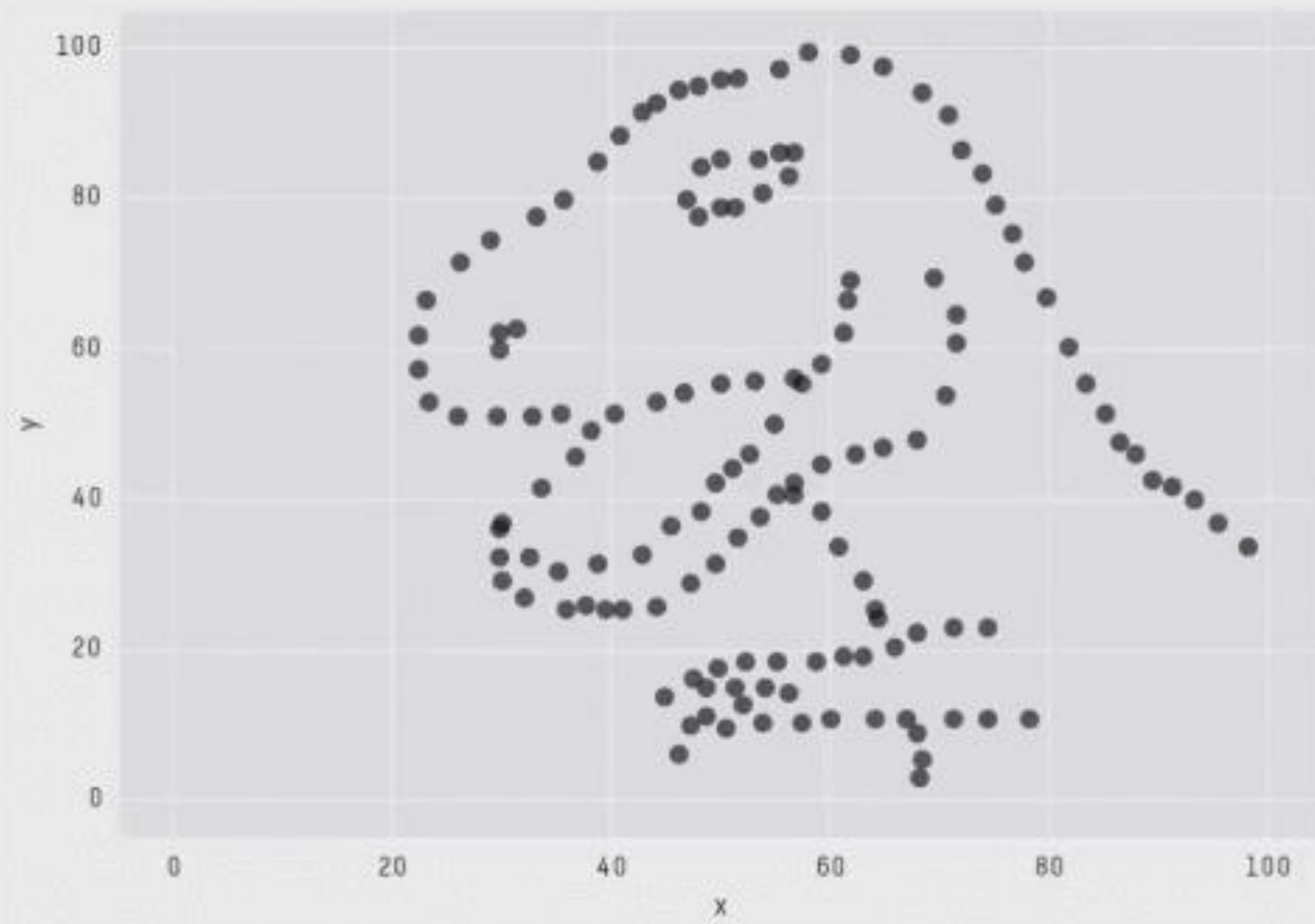




There are three types of lies: lies,  
damned lies, and statistics

Mark Twain (maybe)





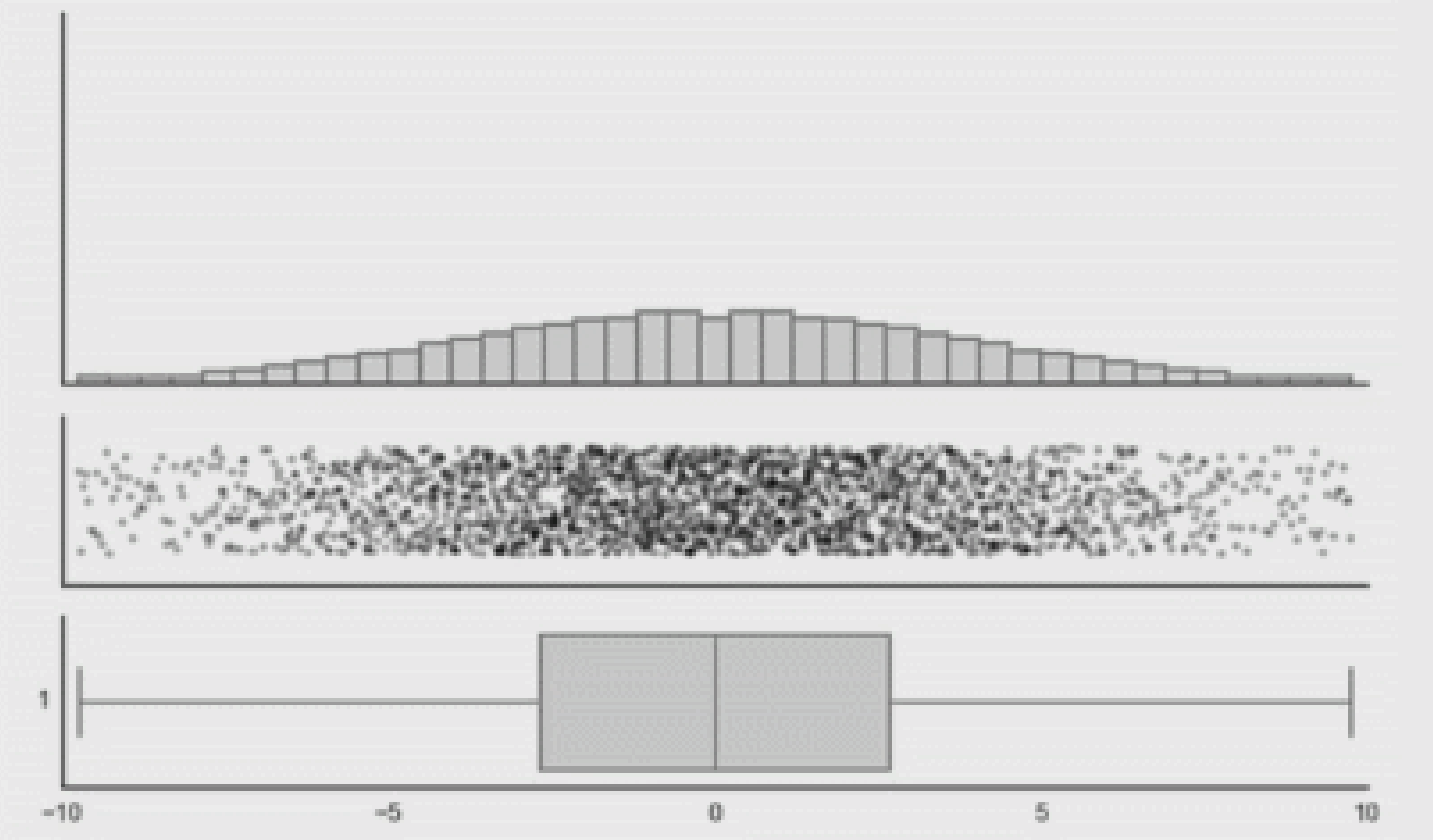
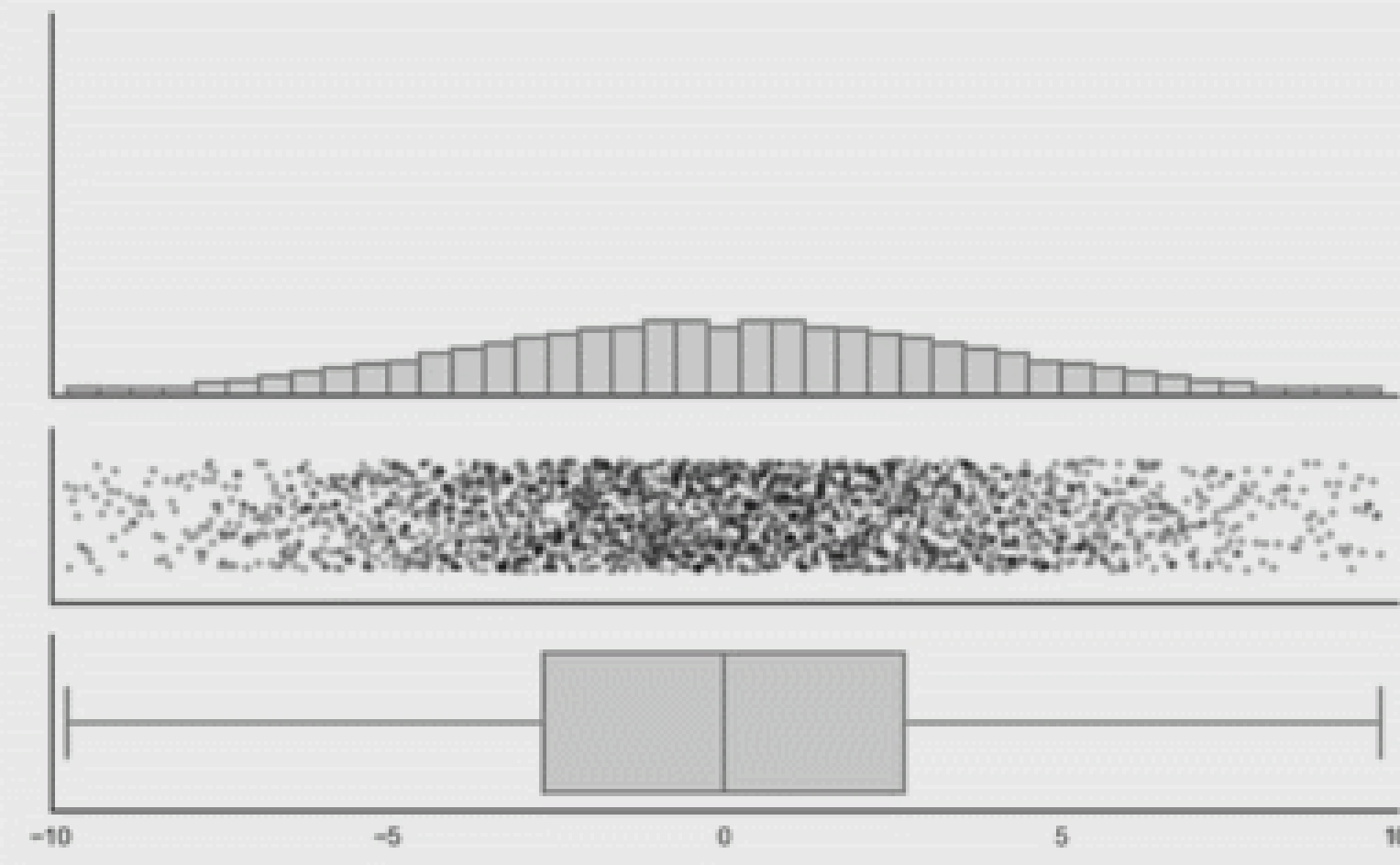
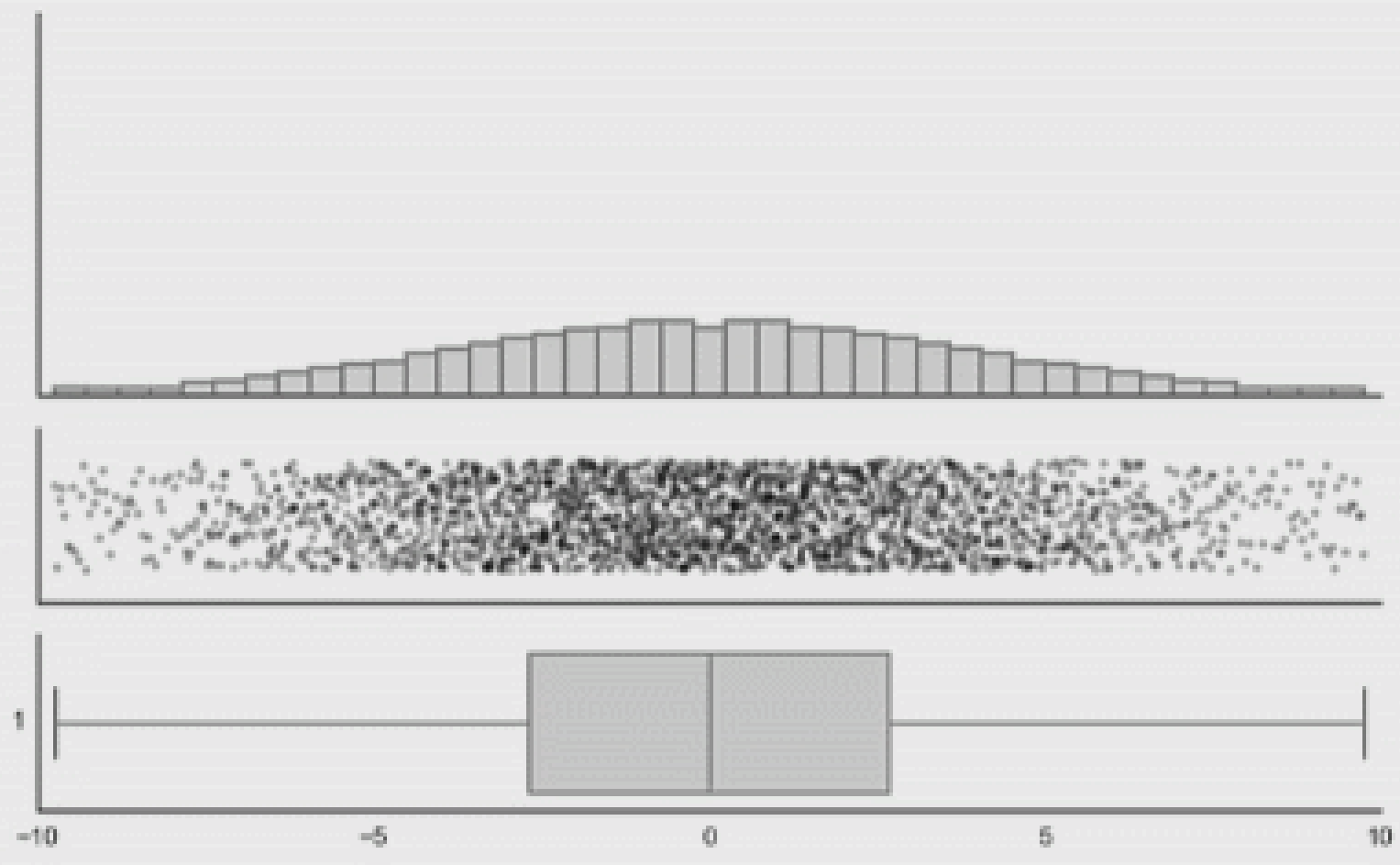
X Mean: 54.2659224  
Y Mean: 47.8313999  
X SD : 16.7649829  
Y SD : 26.9342120  
Corr. : -0.0642526



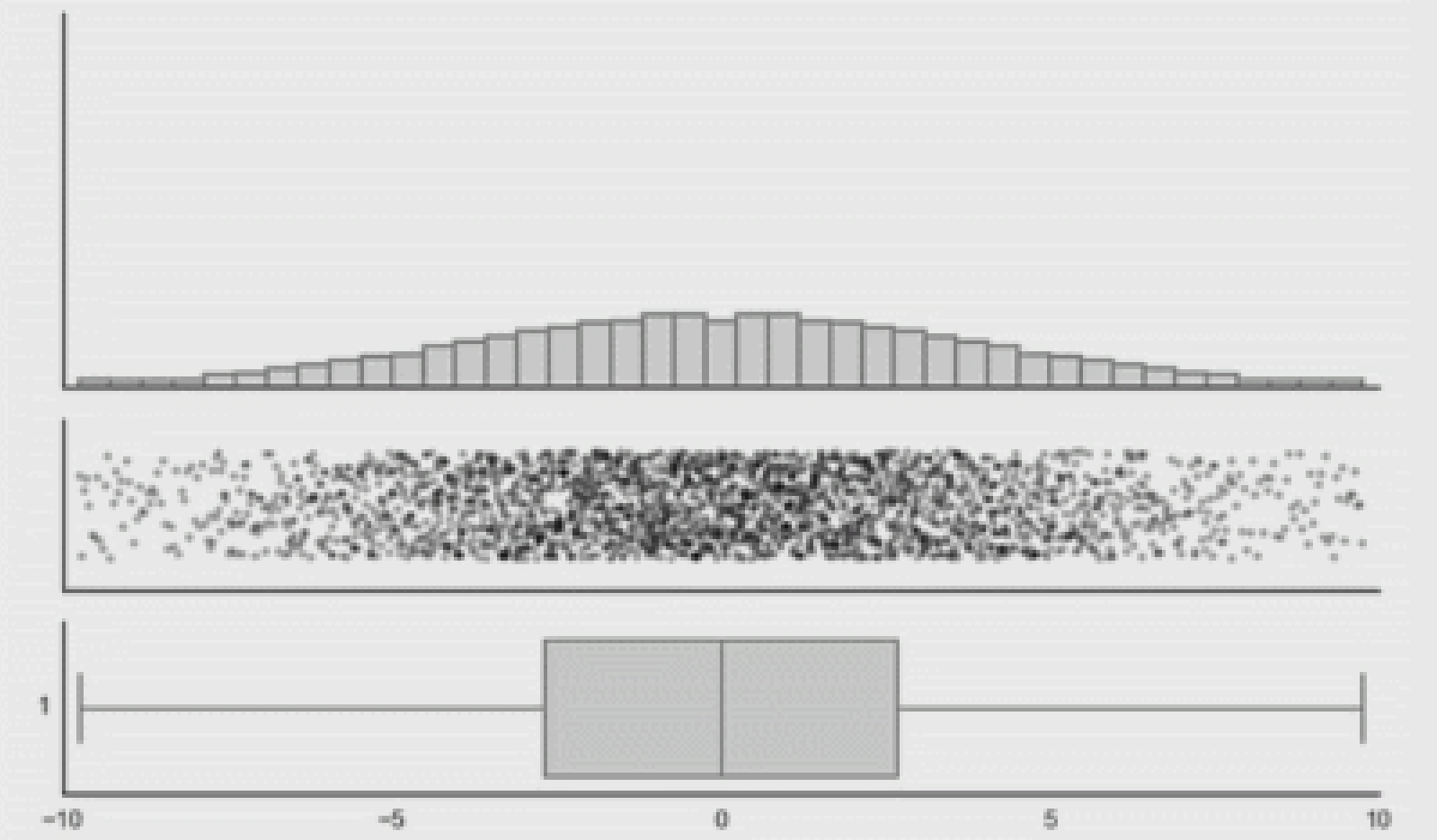
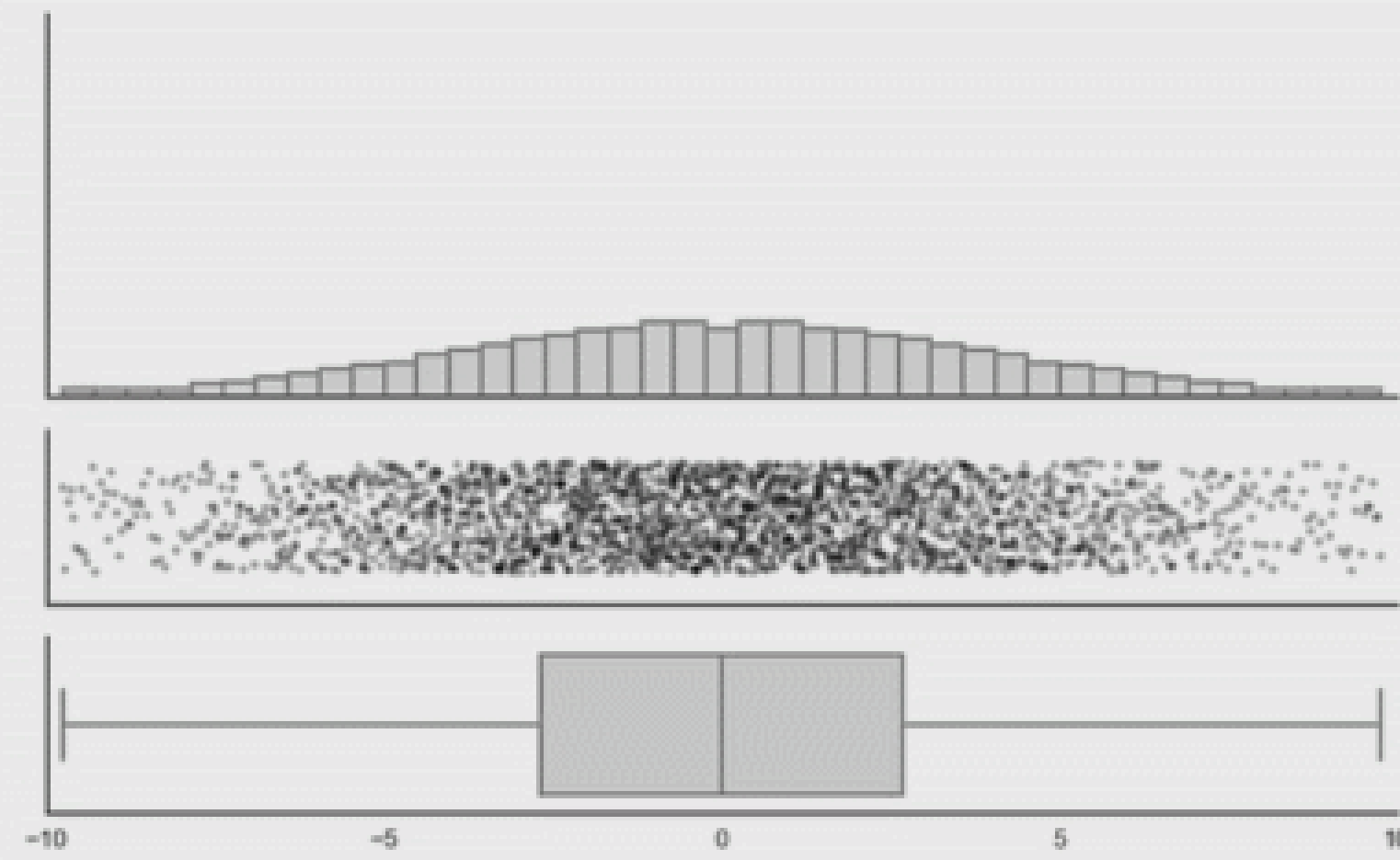
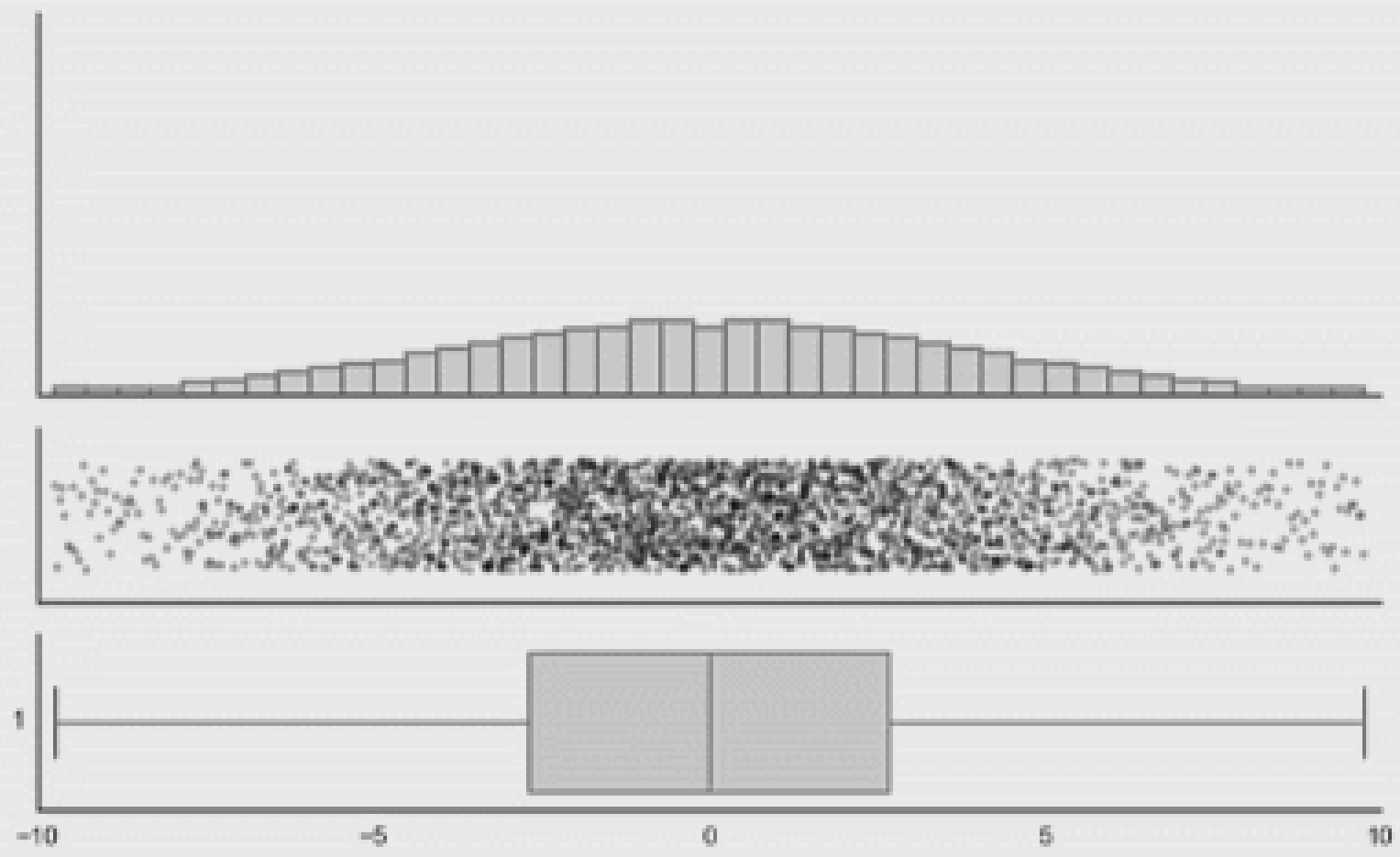
No catalogue of techniques can convey a willingness to look for what can be seen, whether or not anticipated. Yet this is at the heart of exploratory data analysis. ... the picture-examining eye is the best finder we have of the wholly unanticipated.

– Tukey, 1980

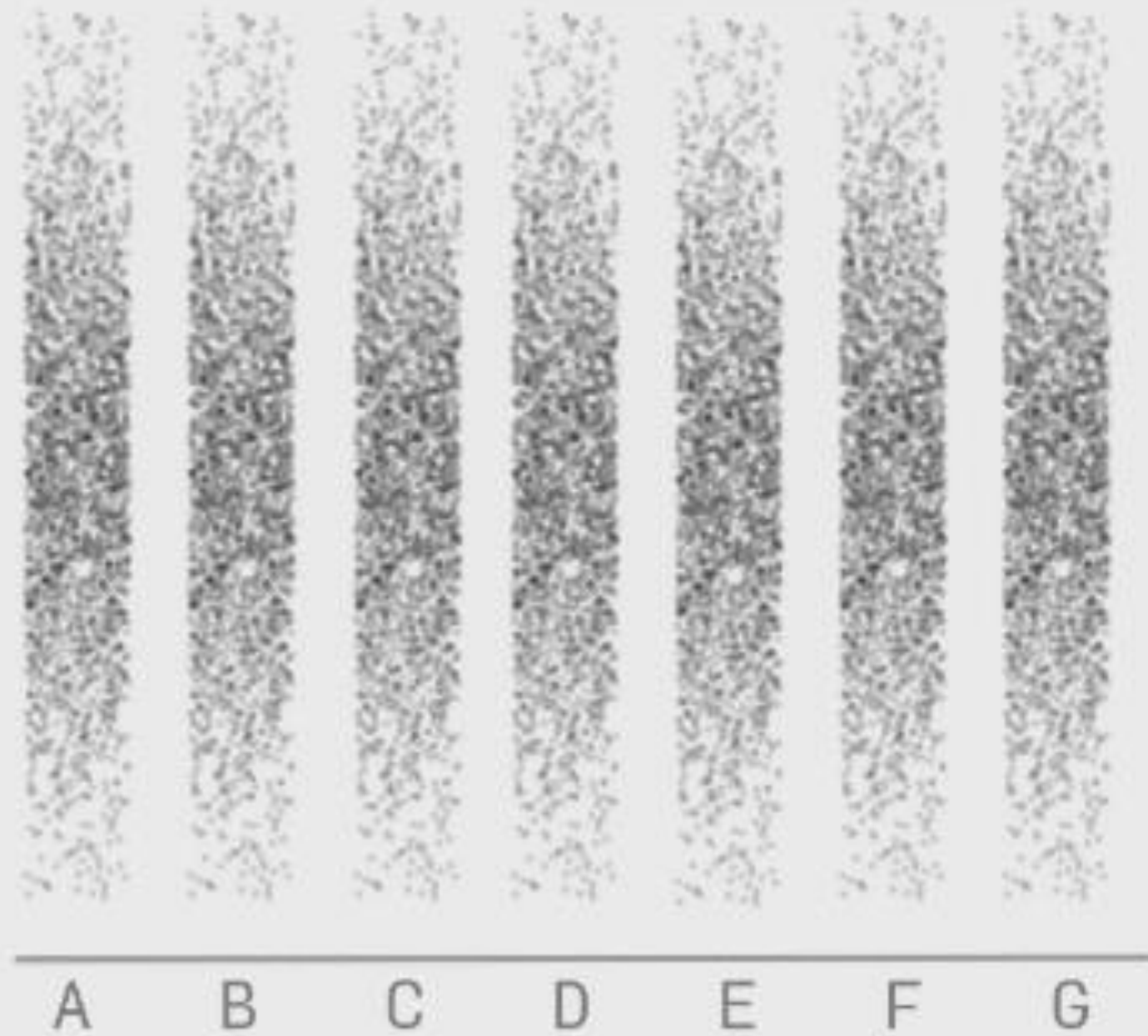




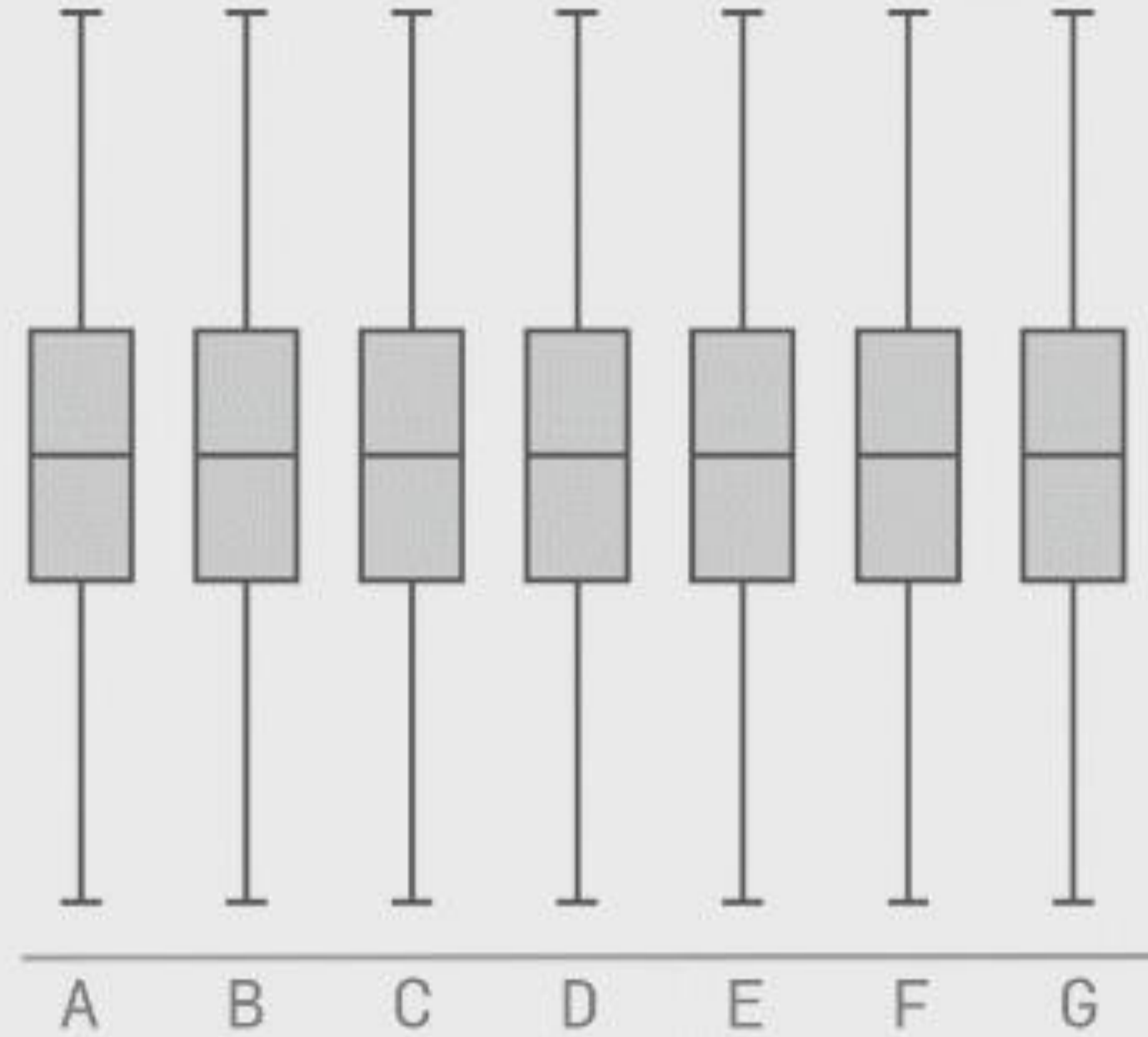




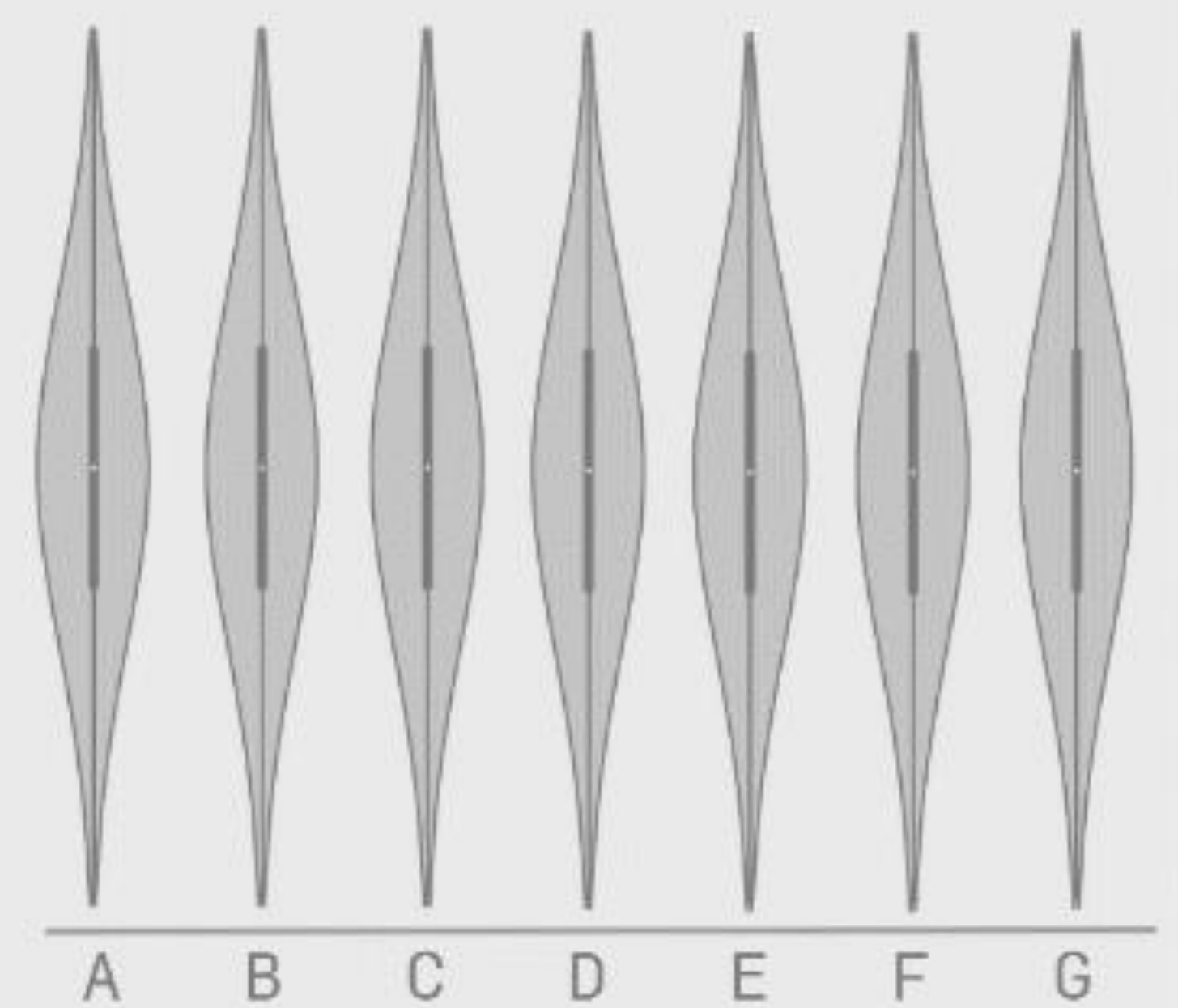
**Raw Data**



**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



# Why visualize your data?

Hans Rosling:

## The best stats you've ever seen

TED2006 · 19:50 · Filmed Feb 2006

48 subtitle languages ?

View interactive transcript









# Why visualize your data?



Generate hypotheses, Answer questions, Find patterns,  
Record, Clarify, Communicate, Inspire



# Hall of Fame or Hall of Shame



Prof. Krzysztof Gajos



# wind map

past patterns

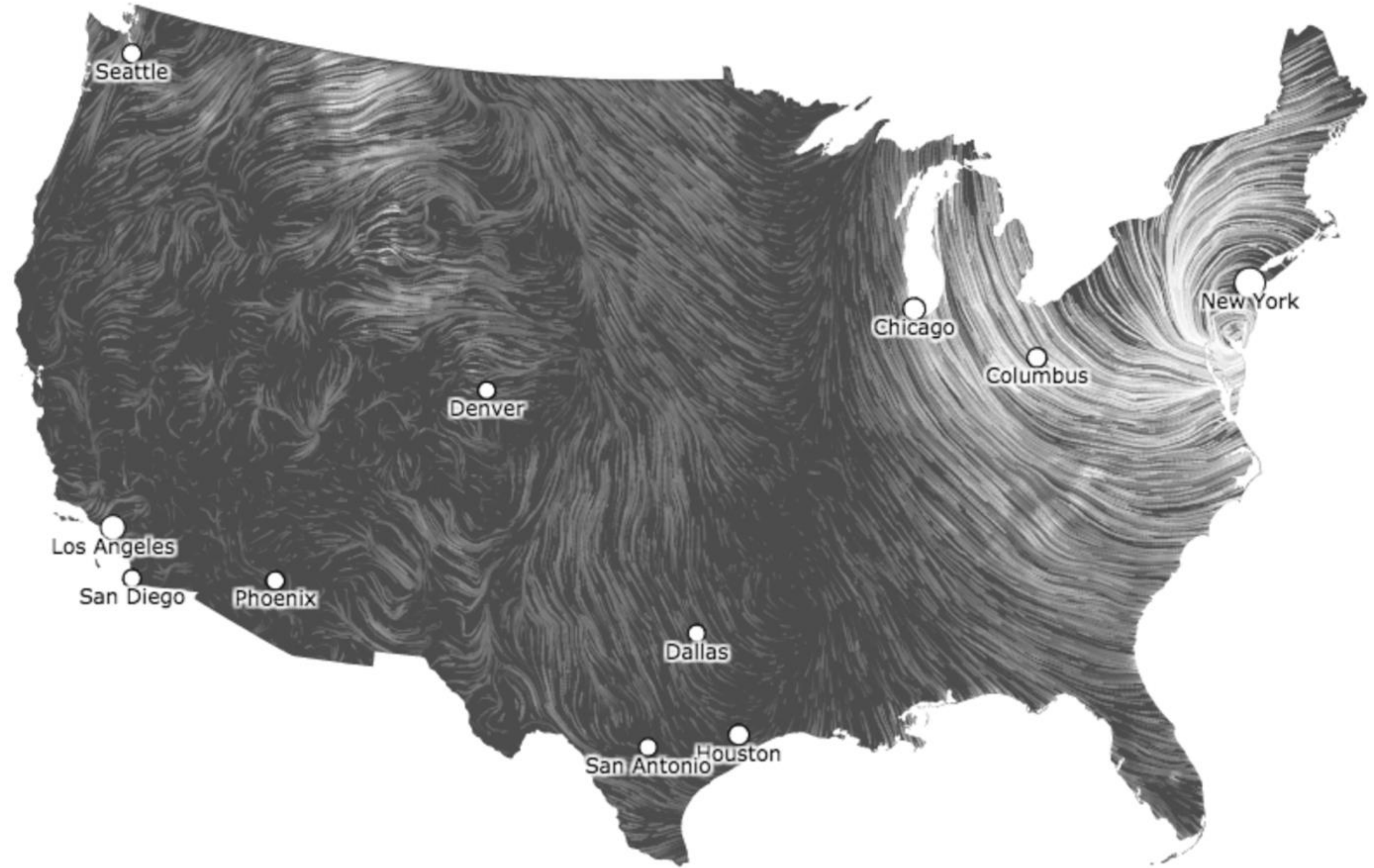
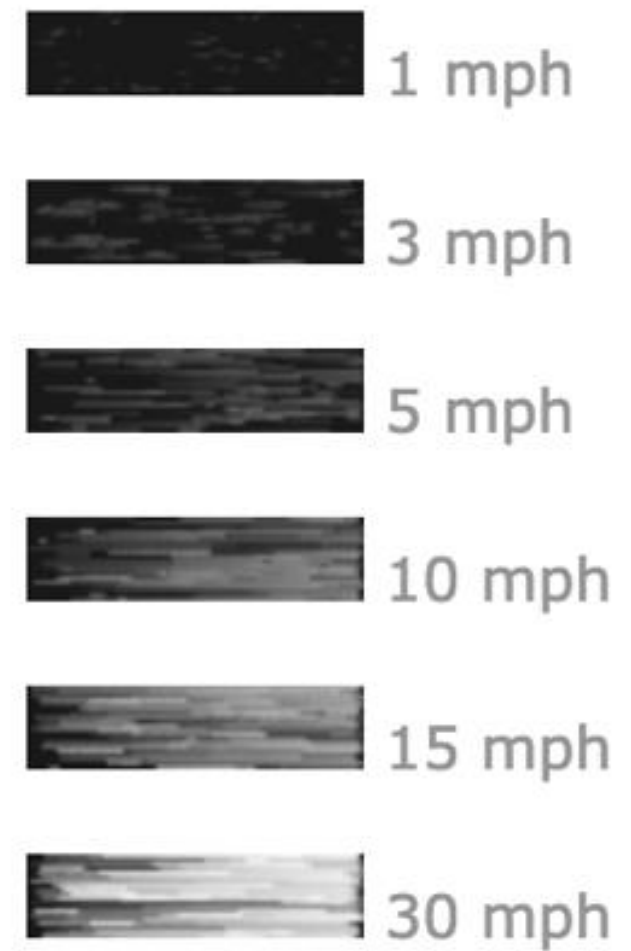
<previous next>

**October 29, 2012**

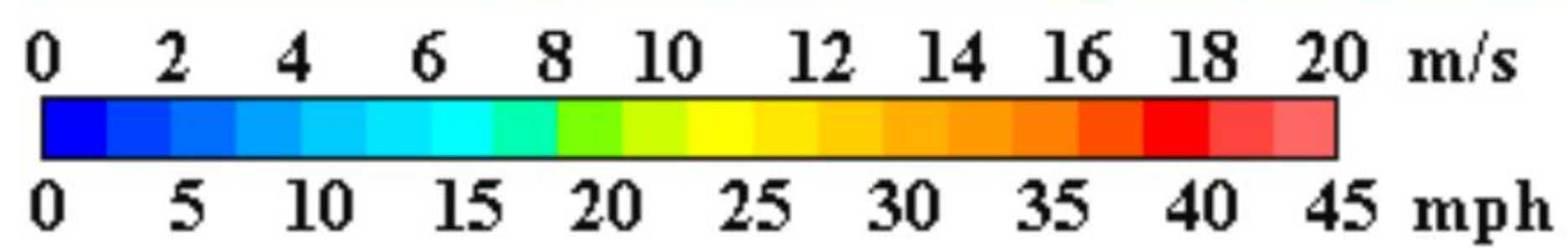
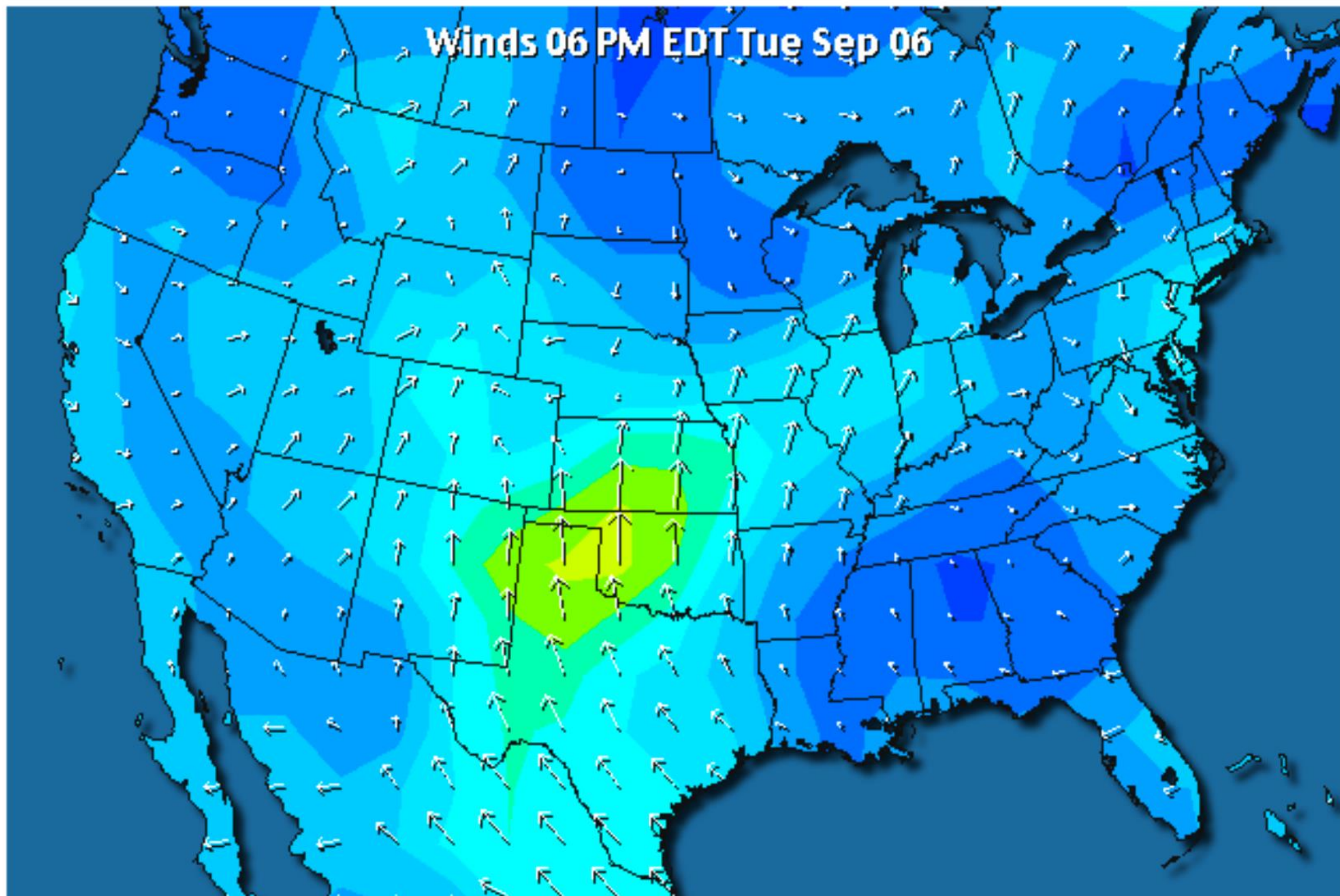
8:59 pm EST

(time of forecast download)

top speed: **45.1 mph**  
average: **9.4 mph**









# Upcoming Assignments & Communication

<https://northeastern.instructure.com/courses/63405/assignments/syllabus>

Look at the upcoming assignments and deadlines

- Textbook, Readings, & Reading Quizzes — Variable days
- In-Class Activities — 11:59pm same day as class
- Assignments & Projects— Generally due **R 11:59pm**  
Assignment 1a,b,c due **in two days**
- [Project Overview](#)

Everyday Required Supplies:

- 5+ colors of pen/pencil
- White paper
- Laptop and charger

Use Canvas Discussions for general questions, email the TAs/S-LTA/instructor for questions specific to you.