



# Data visualisation workshop

Lars Verspohl | [datamake.io](https://datamake.io) | @lars\_vers | lars@datamake.io

## The Plan...

- History and lots of charts...
- Why do we visualise? Visual perception
- How do we visualise? Visual encoding
- Design decisions Charts and pitfalls
- Exercises and discussions

# What is data visualisation?

**Data visualisation refers to the graphical representation of information**

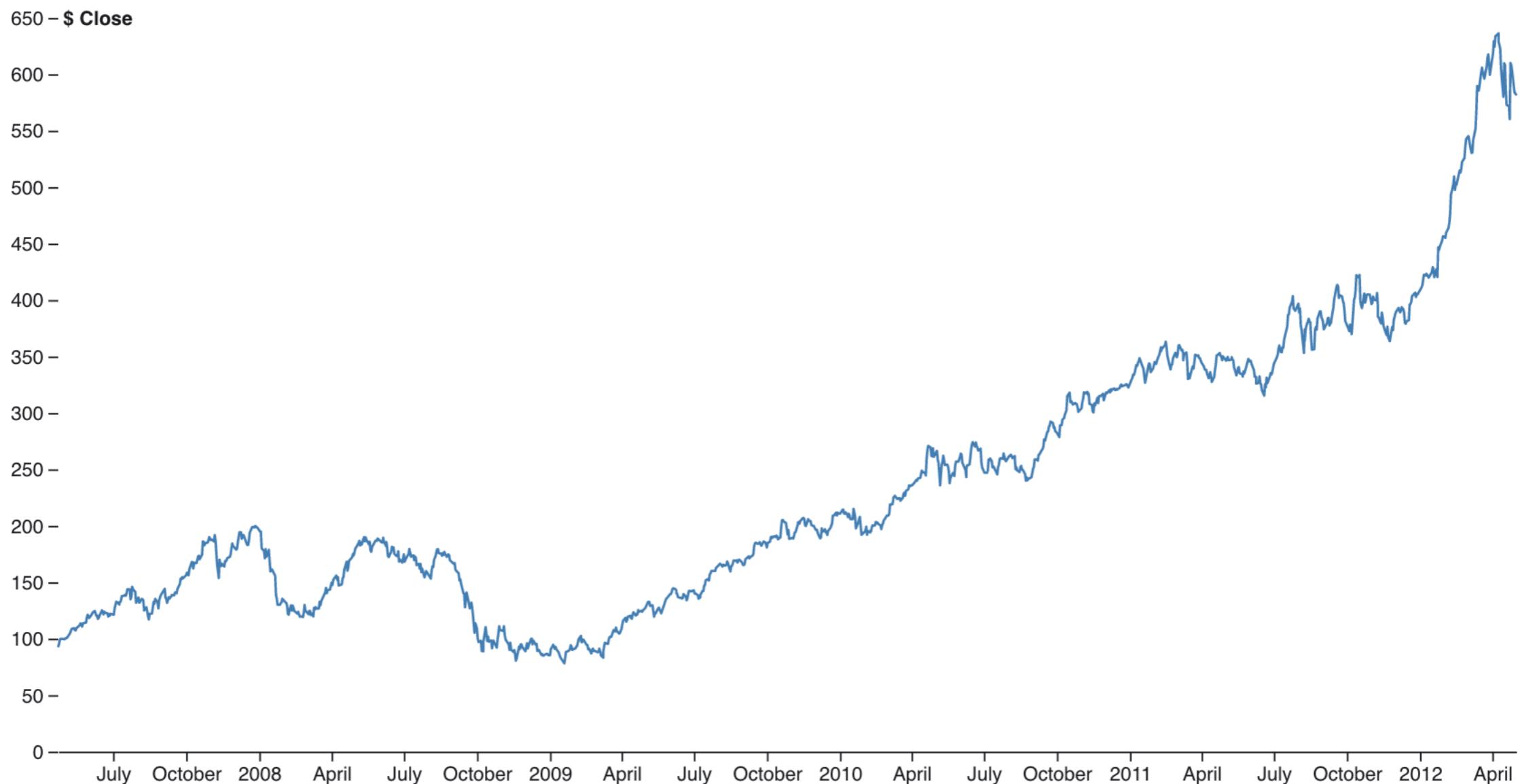
Data → Image





# Apple stock price

AAPL stock price from April 2007 to May 2012. Data: Yahoo Finance



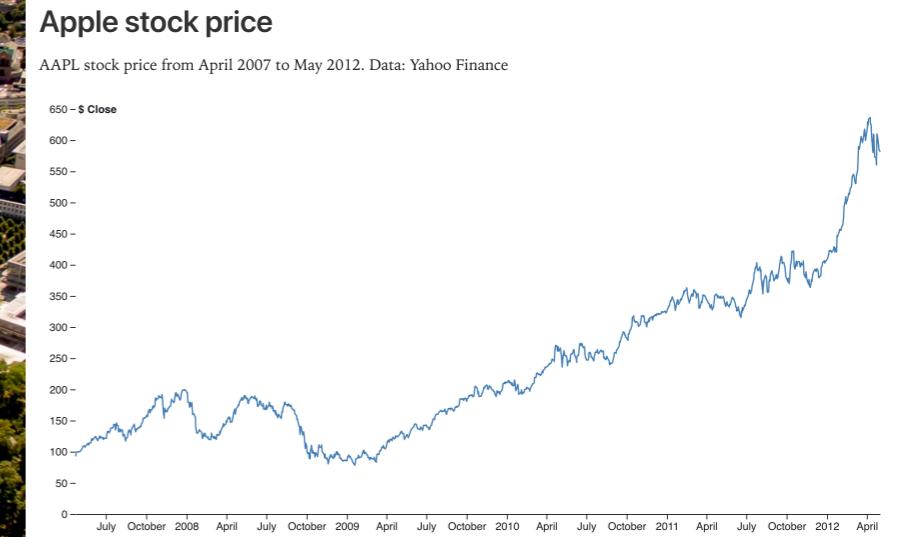
## Pansemic signs



## Polysemic signs



## Monosemic signs



### Abstract Art

No elements are defined

Reader creates meaning

### Photography

Elements are conventionally defined

Reader creates meaning within set conventions

### Data visualisation

Elements are precisely defined

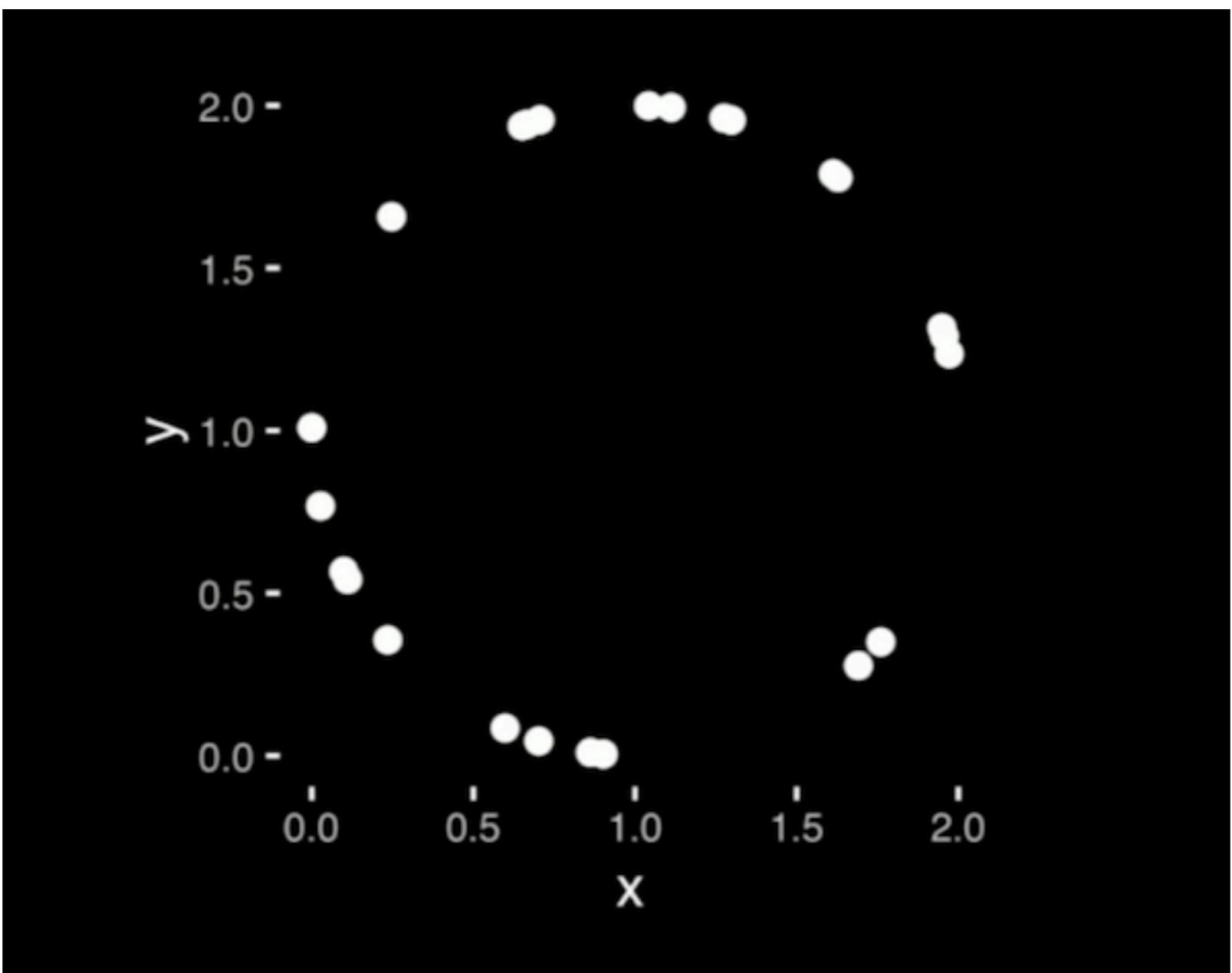
Author creates meaning

**The task of data visualisation is the maximum reduction of confusion**

Data visualisation refers to the graphical representation of precisely defined information

Why do we visualise?

x	y	x	y
1.972	1.236	0.111	0.542
1.112	1.994	0.902	0.005
0.000	1.009	0.598	0.085
0.665	1.942	1.613	1.790
0.235	0.356	1.298	1.955
0.247	1.658	0.651	1.937
1.275	1.961	1.949	1.316
0.702	0.045	0.099	0.567
1.760	0.350	0.862	0.010
1.691	0.277	0.027	0.768
1.628	1.778	0.706	1.956
1.957	1.290	1.042	1.999



# History

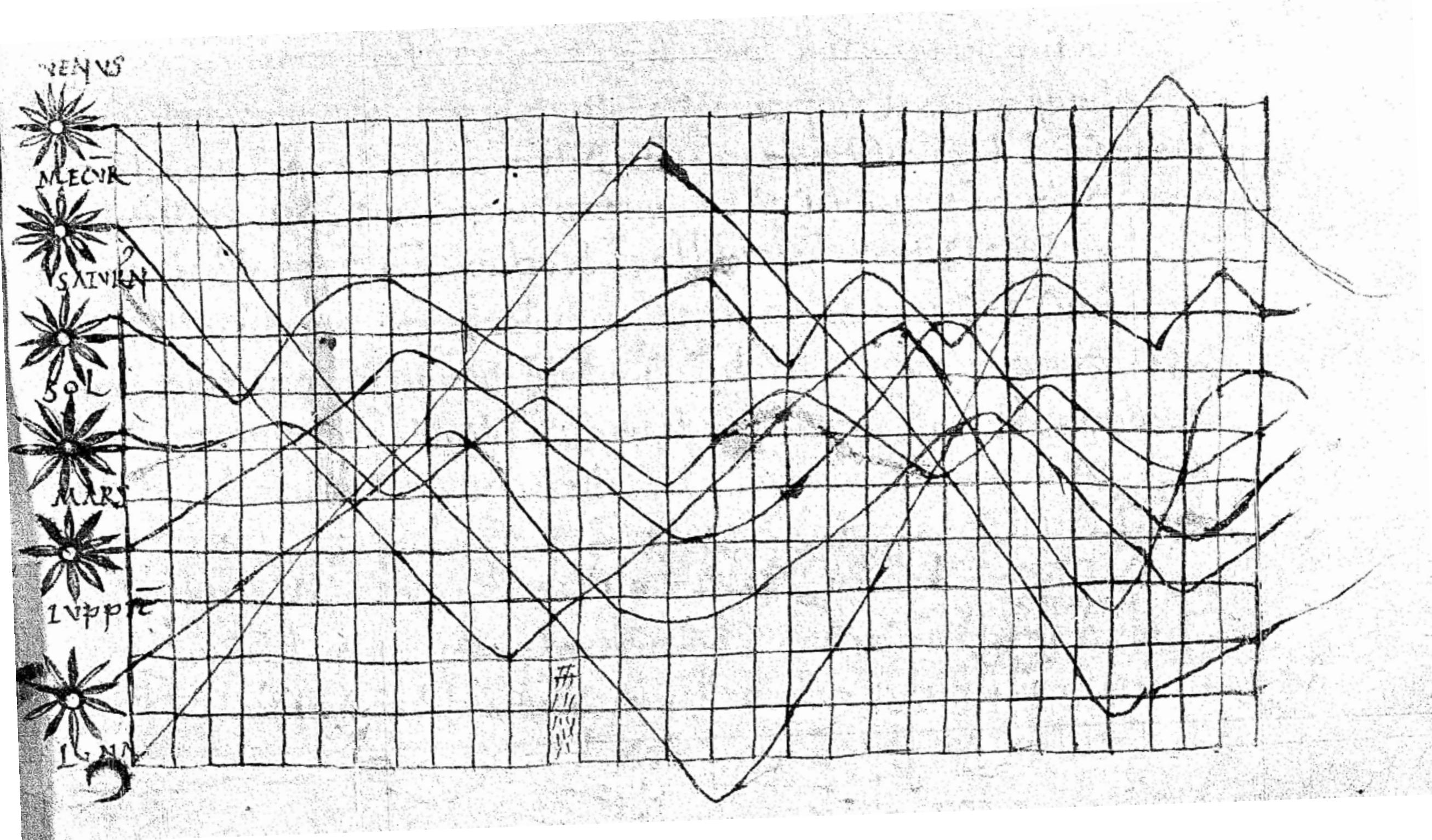
The dawn of science



6250 BC Konya town map, Turkey



150 AD Ptolemy, Geographica



950 AD Position of Sun, Moon, Planets

ratio reddit uniforme scilicet ad unitam.  
**C** Leam: unius ex excessu graduus  
 aportos a iusta in e proportionis  
 excessus graduus  
 aequaliter aportos e qualiter  
 distofis ut p; ex  
 unius secunde divisionis  
 a seruat tunc nulla  
 est in latitudine talis et  
 divisionem et difformis  
 afformiter difformis  
 idem eque distantiam  
 portionem sicut in se  
 latitudinem tamen est  
 futuris ubi logitur  
 se eque distantiam



Difforme distofie

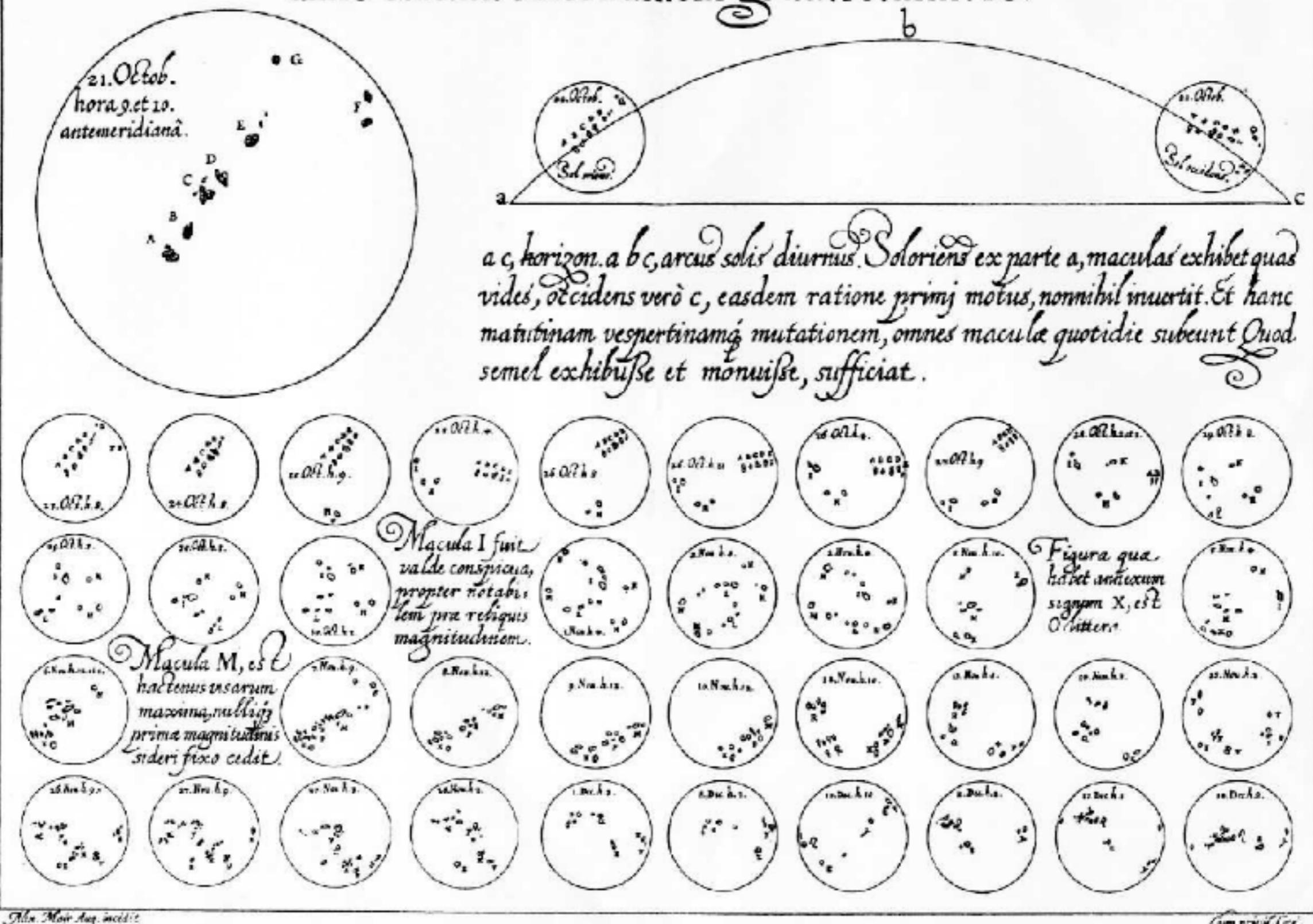


Df, df, distofie



# MACVLAE IN SOLE APPARENTES, OBSERVATAE

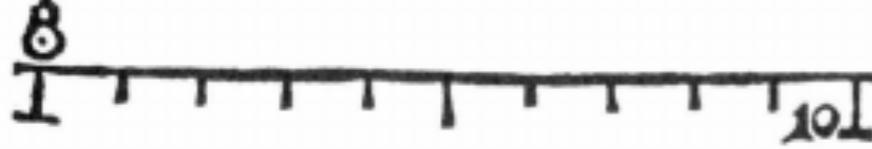
anno 1611. ad latitudinem grad. 48. min. 40.



1626 Sunspots over time, Christoph Scheiner

TOLEDO.

GRADOS DE LA LONGITUD.



G. Langrenius.

G. Mercator.  
I. Schoneraus.  
P. Lansbergius.  
T. Brabæc.

I. Regiomontanus.  
Orontius.  
C. Clavius.  
C. Ptolomæus.  
A. Argelius.  
A. Maginus.  
D. Origanius.

1644 Longitudinal distance between Toledo and Roma, van Langren

The Golden Age



## **William Playfair**

Scotland, 1759-1823

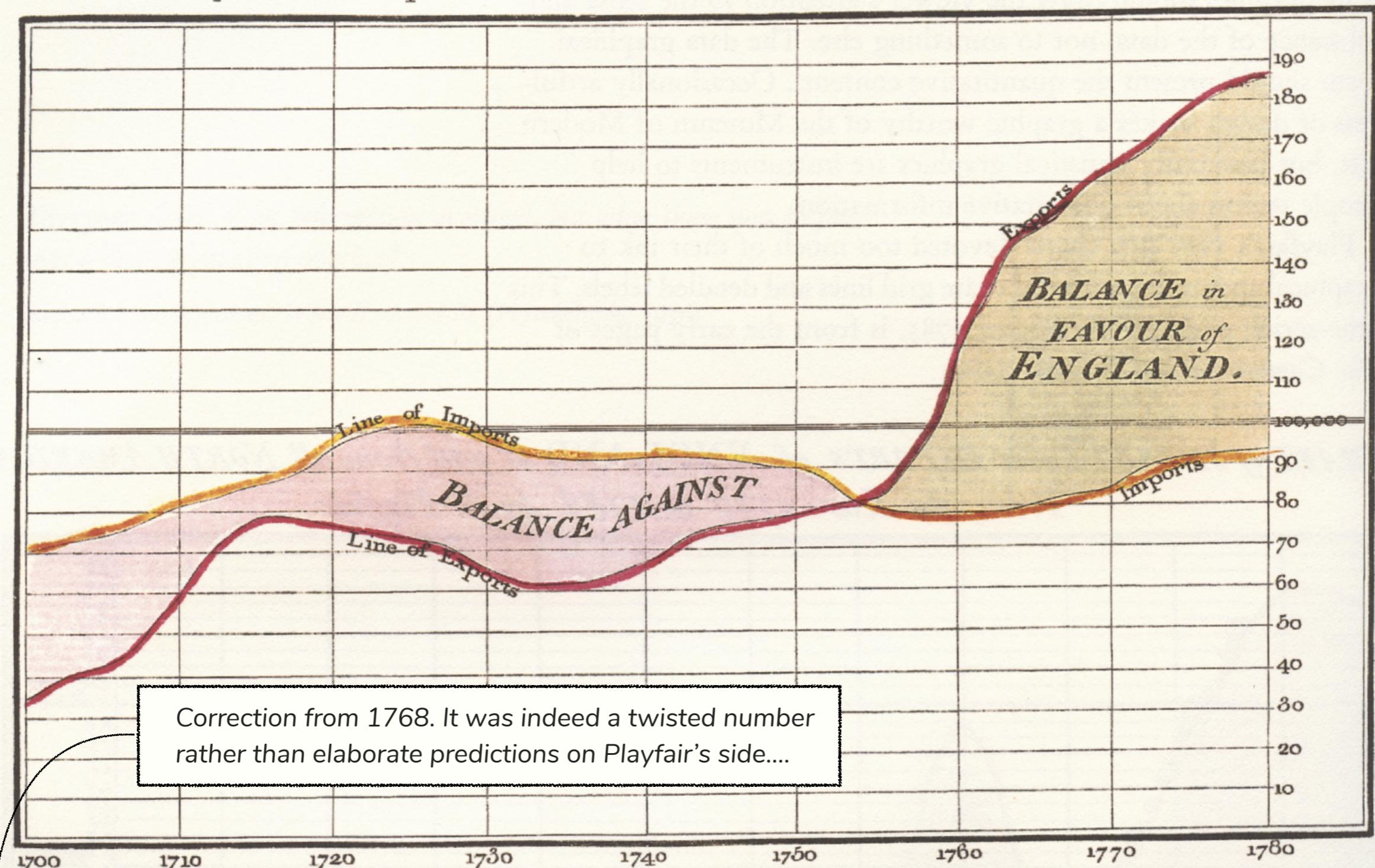
Engineer

Political Economist

Spy

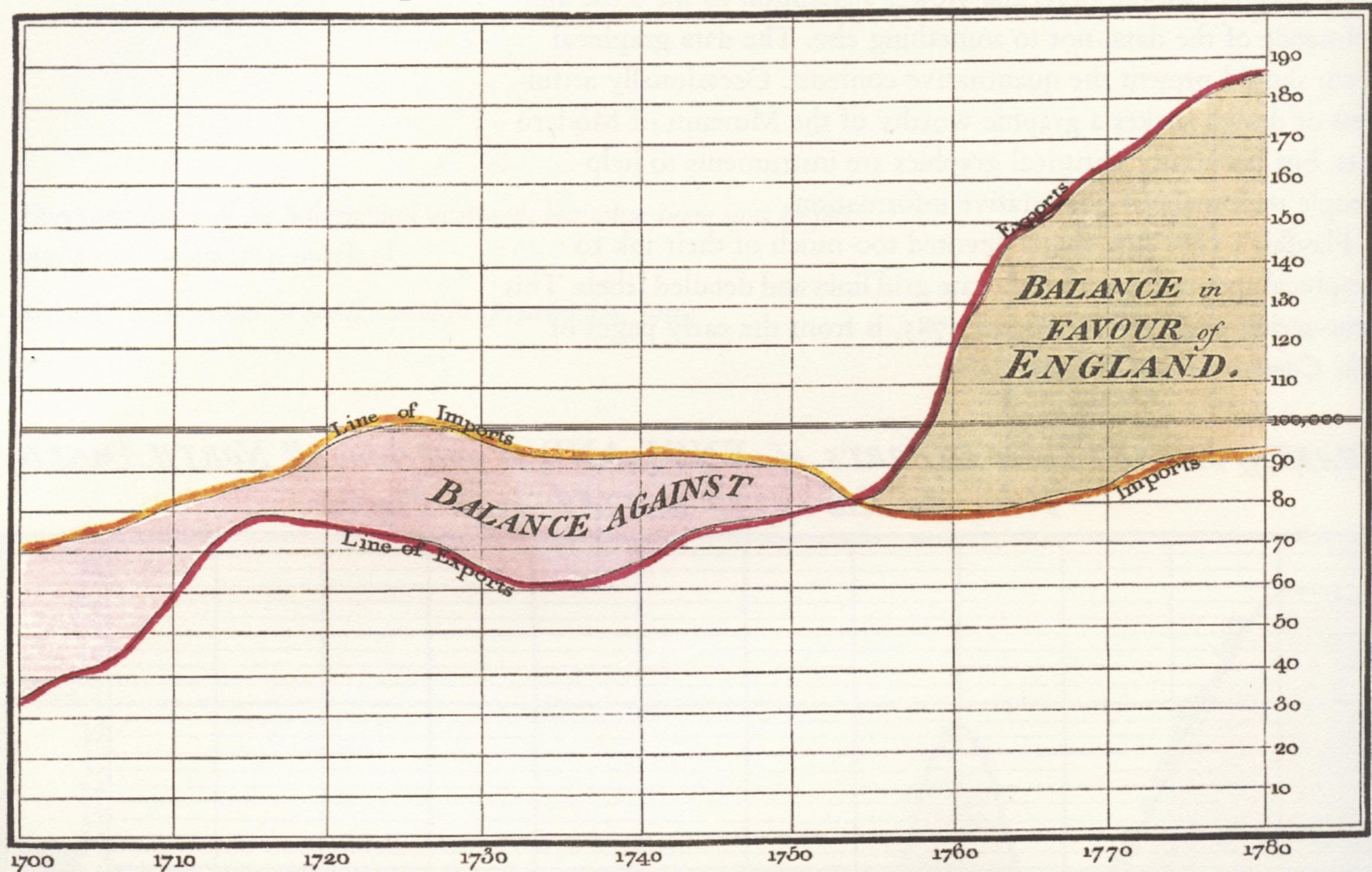
(Inventor) of  
the line chart,  
the area chart,  
the bar chart,  
the pie chart and  
the circle graph.

Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.



The Bottom line is divided into Years, the Right hand line into £10,000 each.  
Published as the Act directs, 1<sup>st</sup> May 1786, by W<sup>m</sup> Playfair  
Neale sculpt 352, Strand, London.

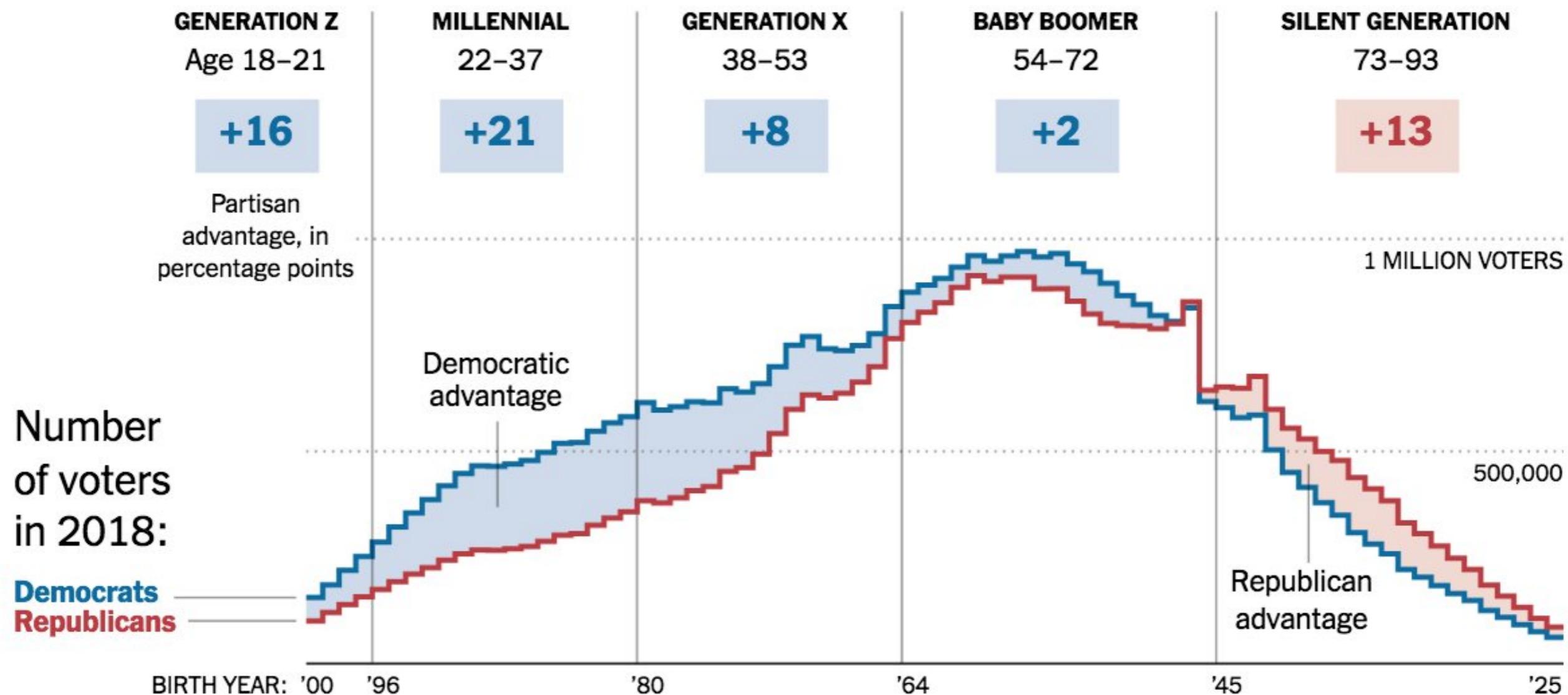
Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.



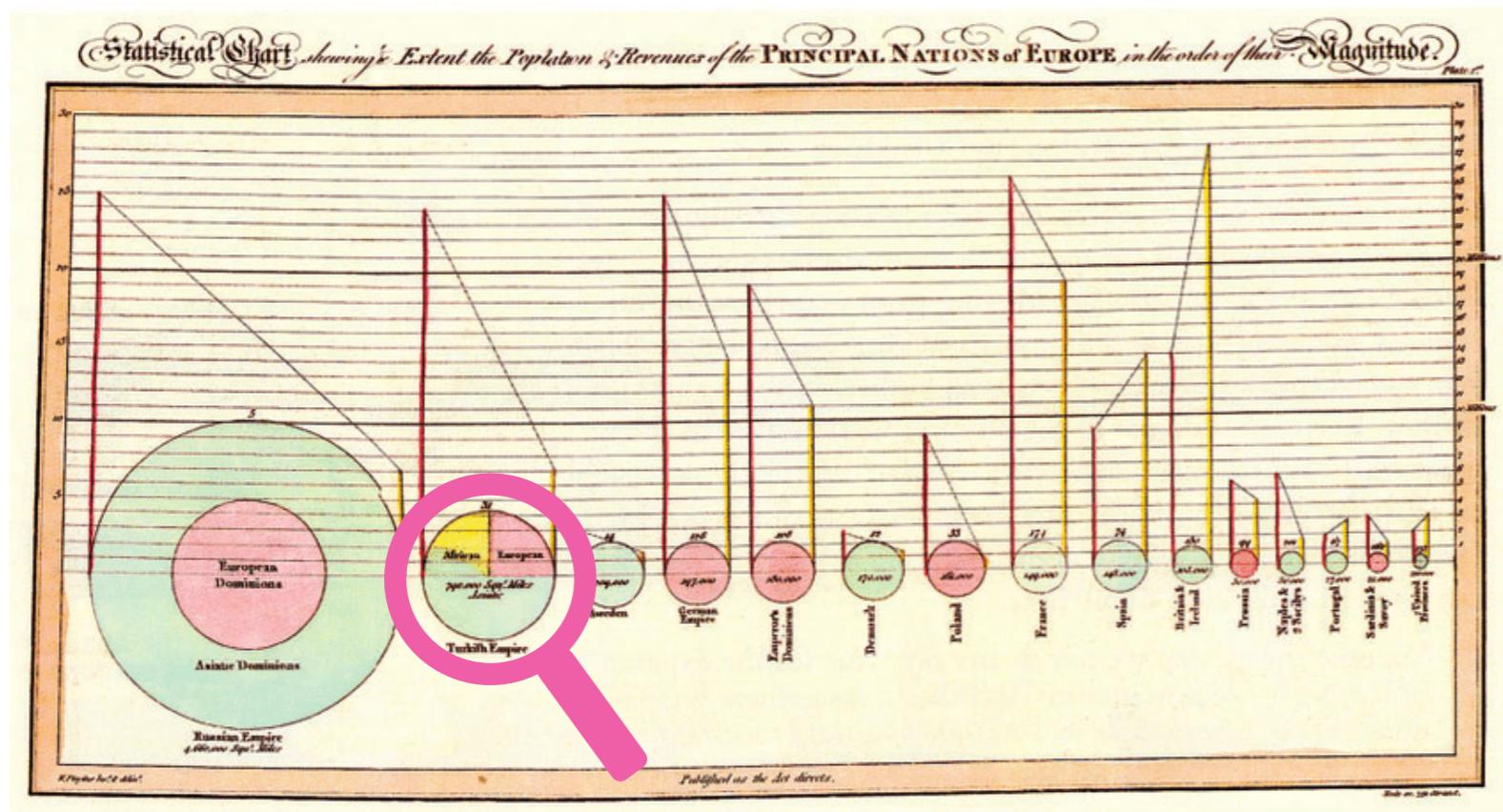
The Bottom line is divided into Years, the Right hand line into £10,000 each.  
Published as the Act directs, 1<sup>st</sup> May 1786, by W<sup>m</sup> Playfair  
Neale sculpt 352, Strand, London.

# A Democratic Edge, in Voters and Nonvoters

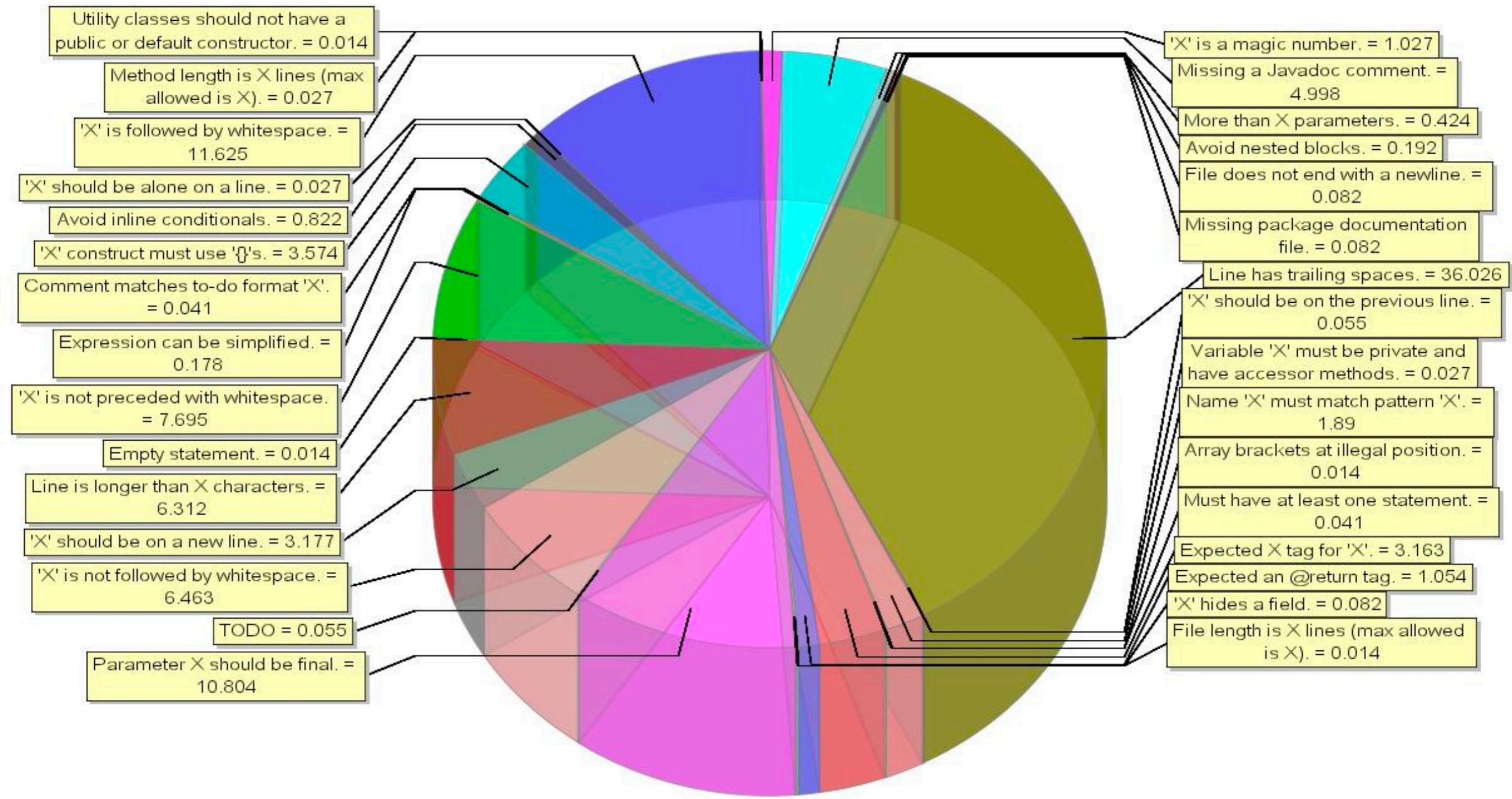
Democrats outnumber Republicans voters, except in the oldest generation, in this year's election. But even greater numbers of Democrats won't be voting — far more so than Republicans.



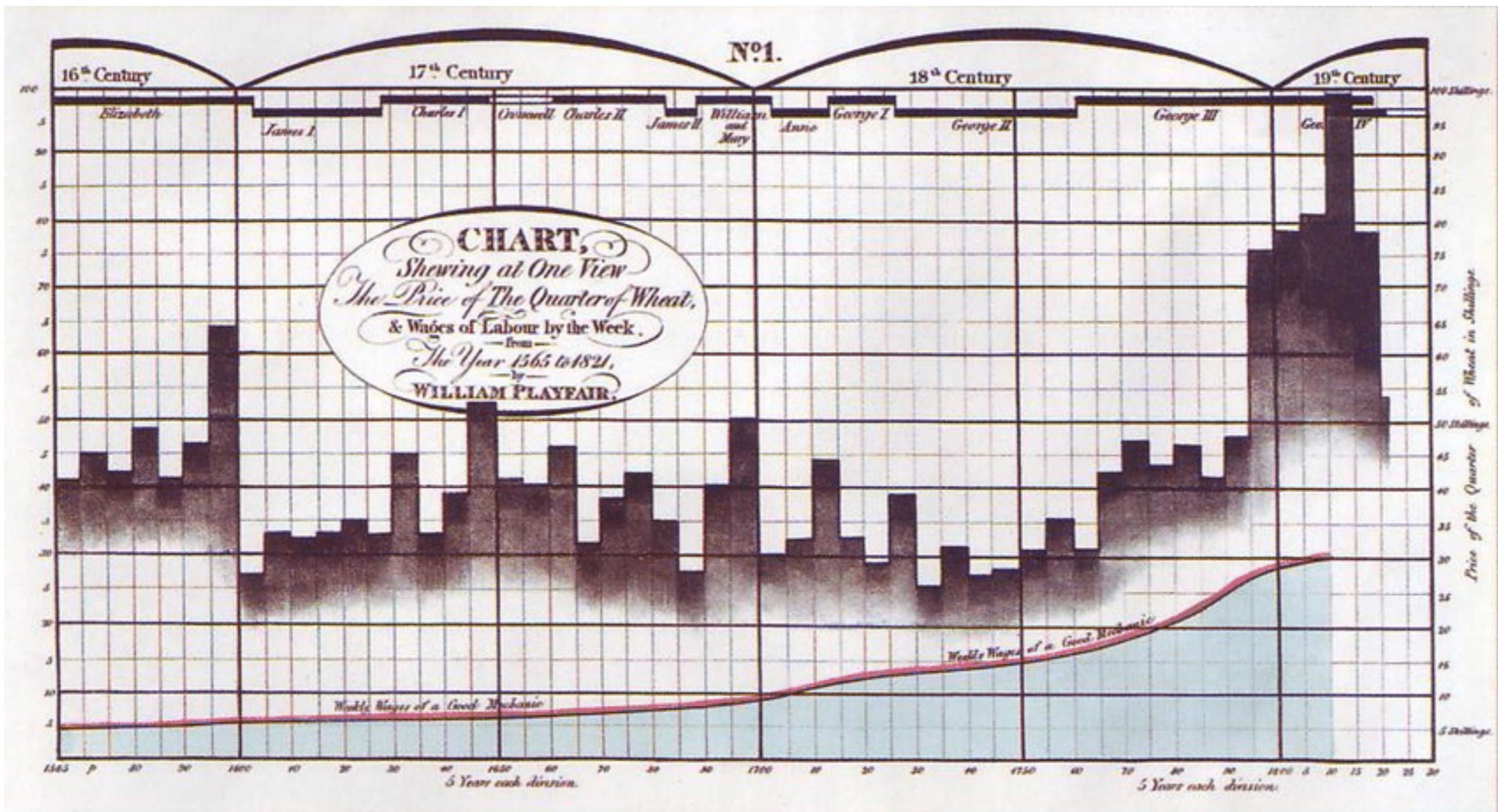
Last week Voters by party affiliation, Bill Marsh, NYT



1801 Statistical Breviary, William Playfair

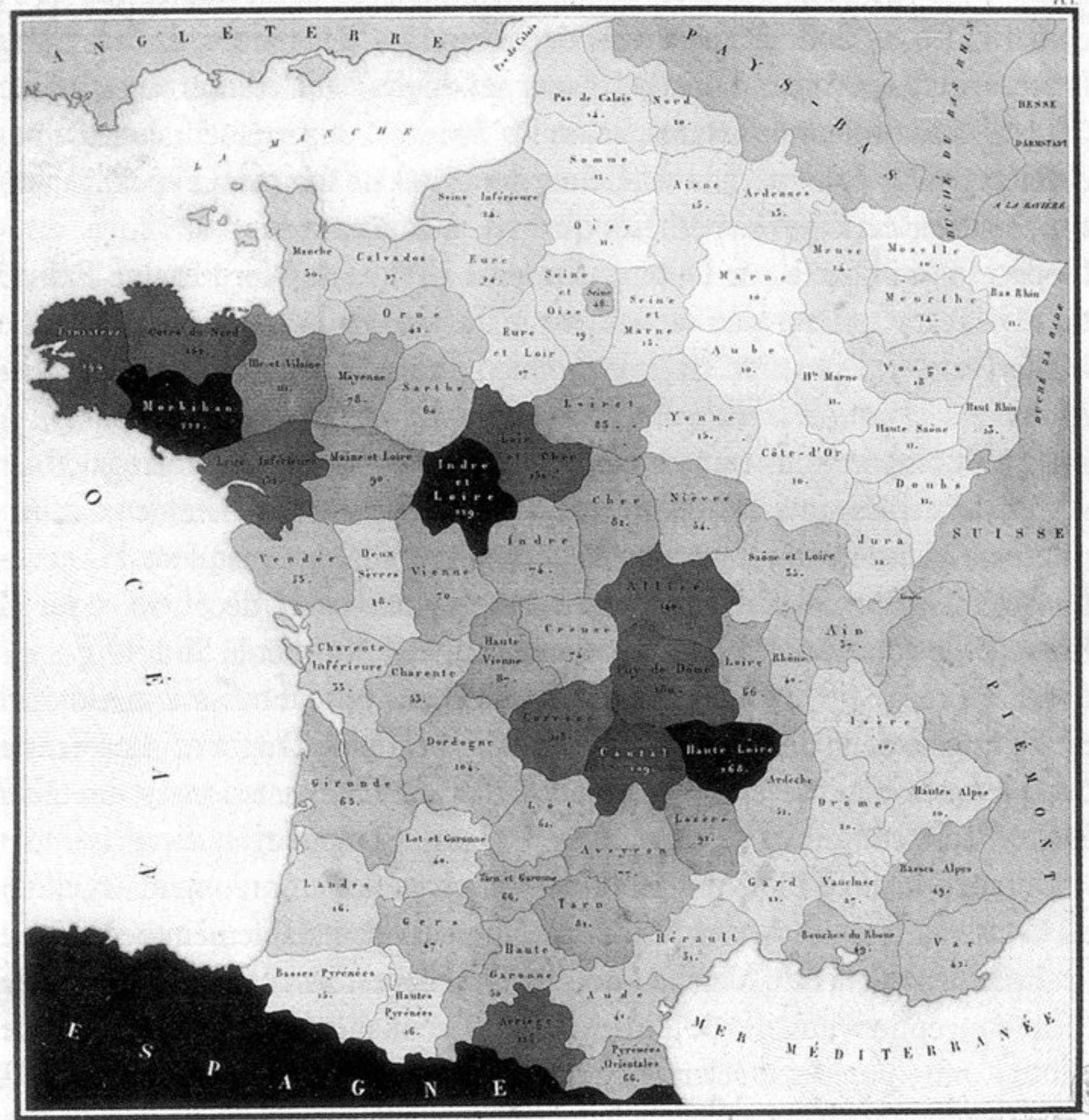


Recently Pie lunacy, Anonymous



1822 Price of Wheat and Wages of Labour, William Playfair

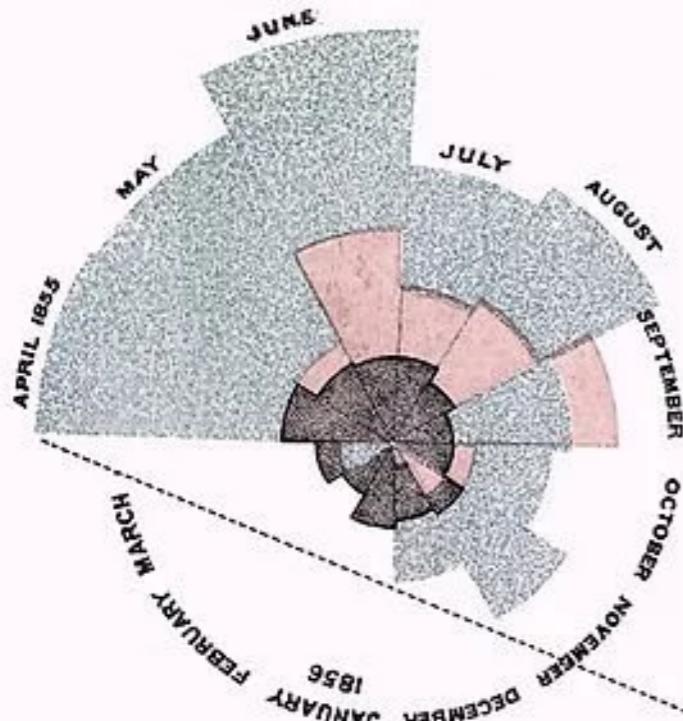
CARTE FIGURATIVE DE L'INSTRUCTION POPULAIRE DE LA FRANCE.



# 1826 Rates of illiteracy across France, Pierre Charles Dupin

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

APRIL 1855 TO MARCH 1856.



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

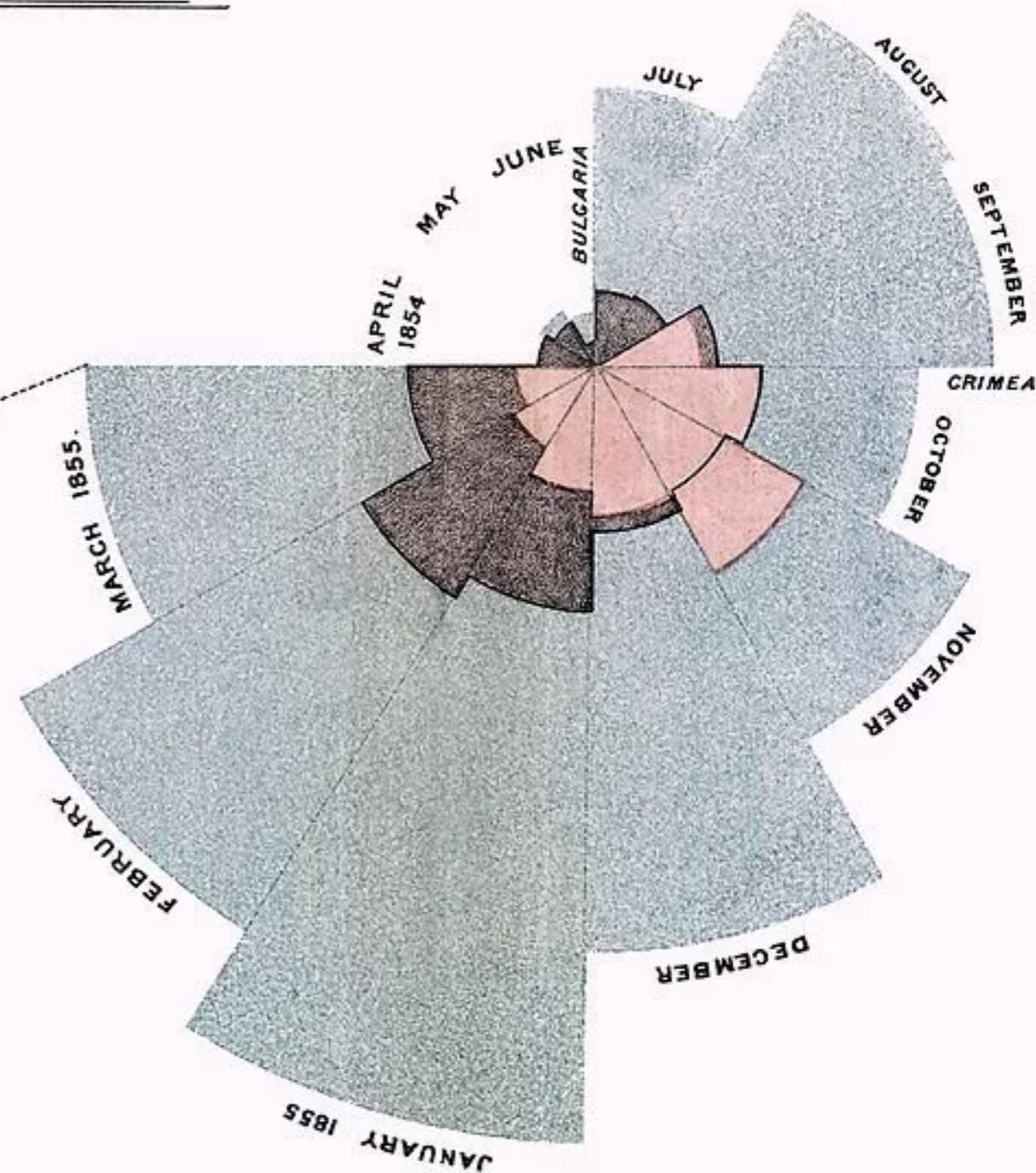
The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases; the red wedges measured from the centre the deaths from wounds; & the black wedges measured from the centre the deaths from all other causes.

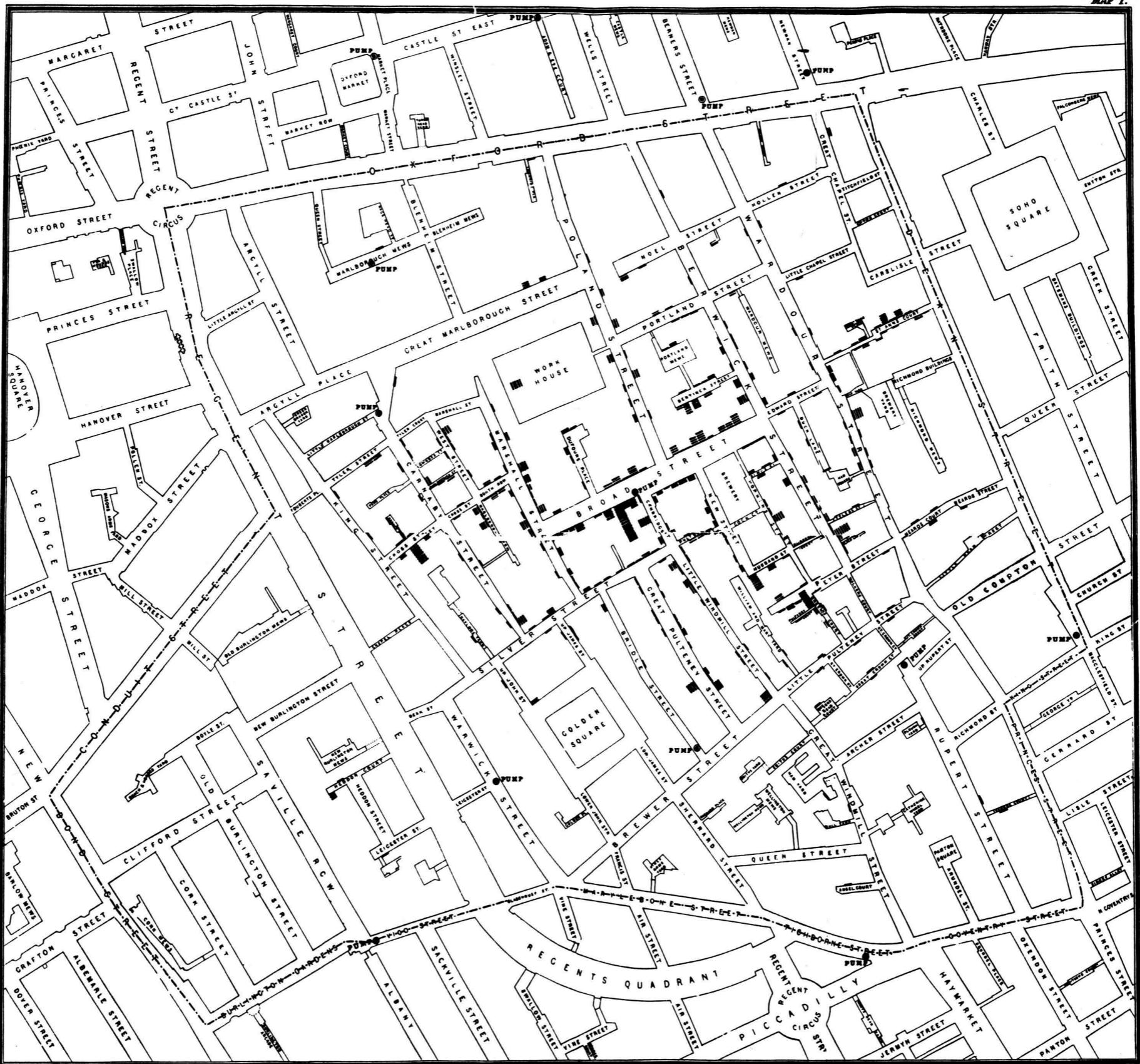
*The black line across the red triangle in Nov. 1854 marks the boundary of the deaths from all other causes during the month.*

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.

1.  
APRIL 1854 TO MARCH 1855.





1854 Cholera Map, John Snow





**Charles J. Minard**  
France, 1781-1870

Civil Engineer  
Interesting Haircut

Made one of the  
most acclaimed  
charts ever.

# 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 médical 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 en Molodexno et se rejoignirent vers Orscha et Witebsk, avaient toujours marché avec l'armée.

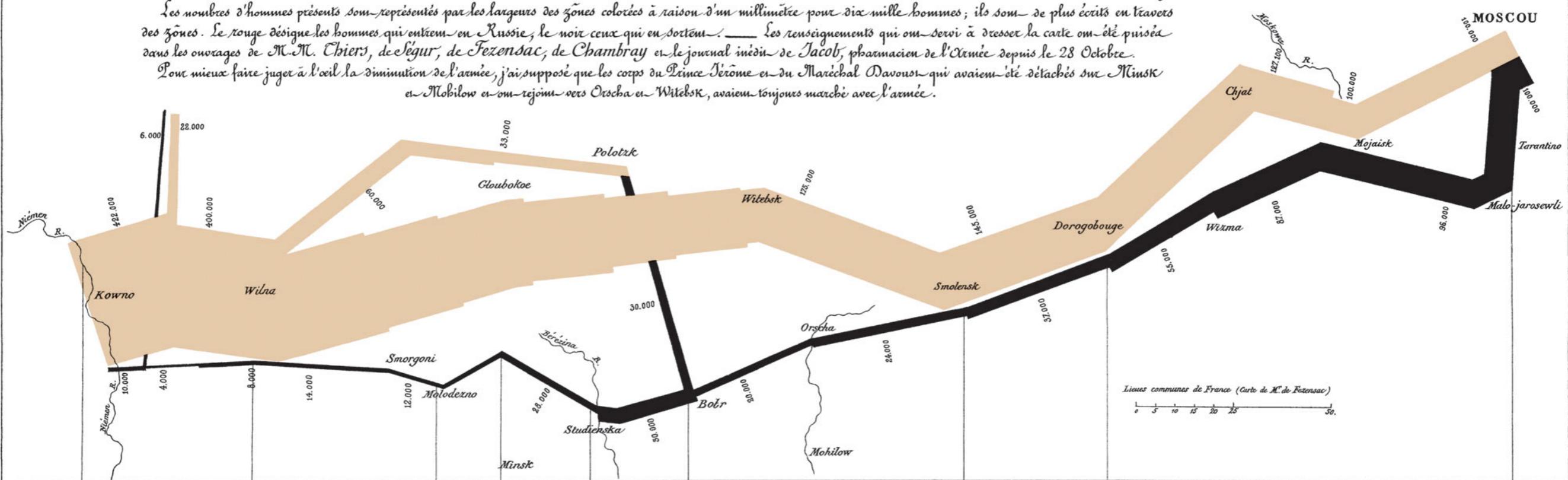
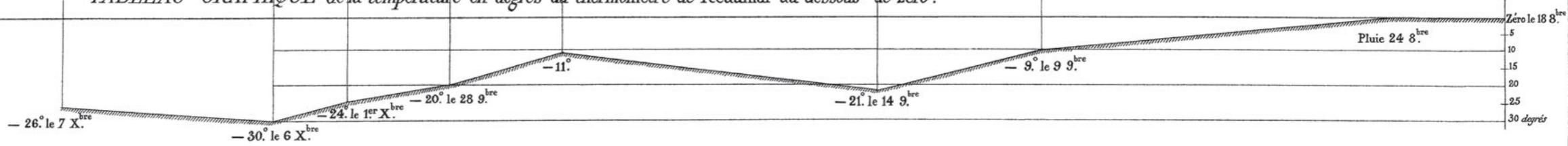


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Cosaques passent au galop  
le Niémen gelé.



Autog. par Regnier, 8. Pas. S<sup>e</sup> Marie S<sup>t</sup> G<sup>e</sup>ne à Paris.

Imp. Lith. Regnier et Dourdet.

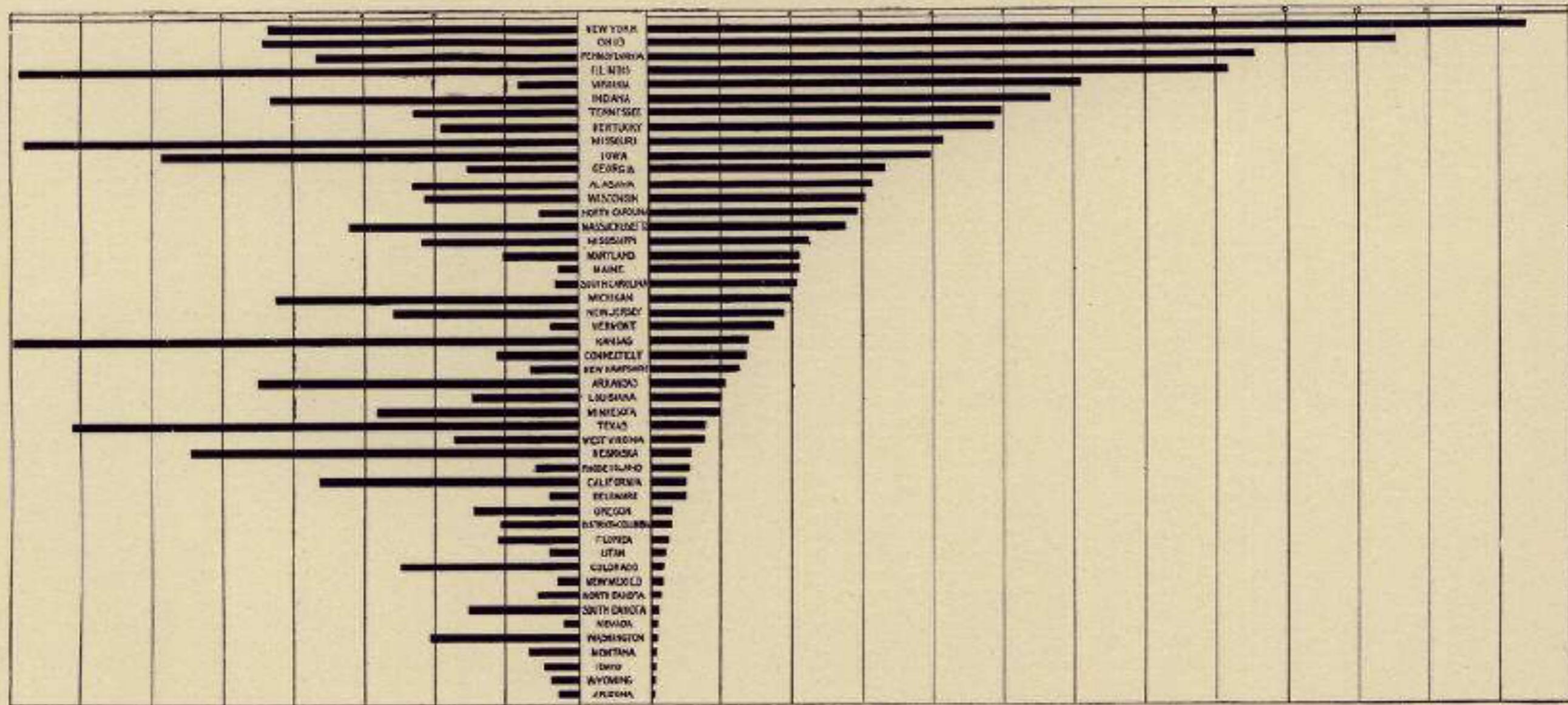
1869 Napoleons Russia campaign 1812-1813, Charles Minard

66. INTERSTATE MIGRATION—NUMBER OF NATIVE IMMIGRANTS AND NATIVE EMIGRANTS, BY STATES AND TERRITORIES: 1890.

Native immigrants.

(Thousands of thousands.)

Native emigrants.



1890 Interstate migration, Statistical Atlas of the 11th US Census

The Dark Ages

1900-1950

The modern era

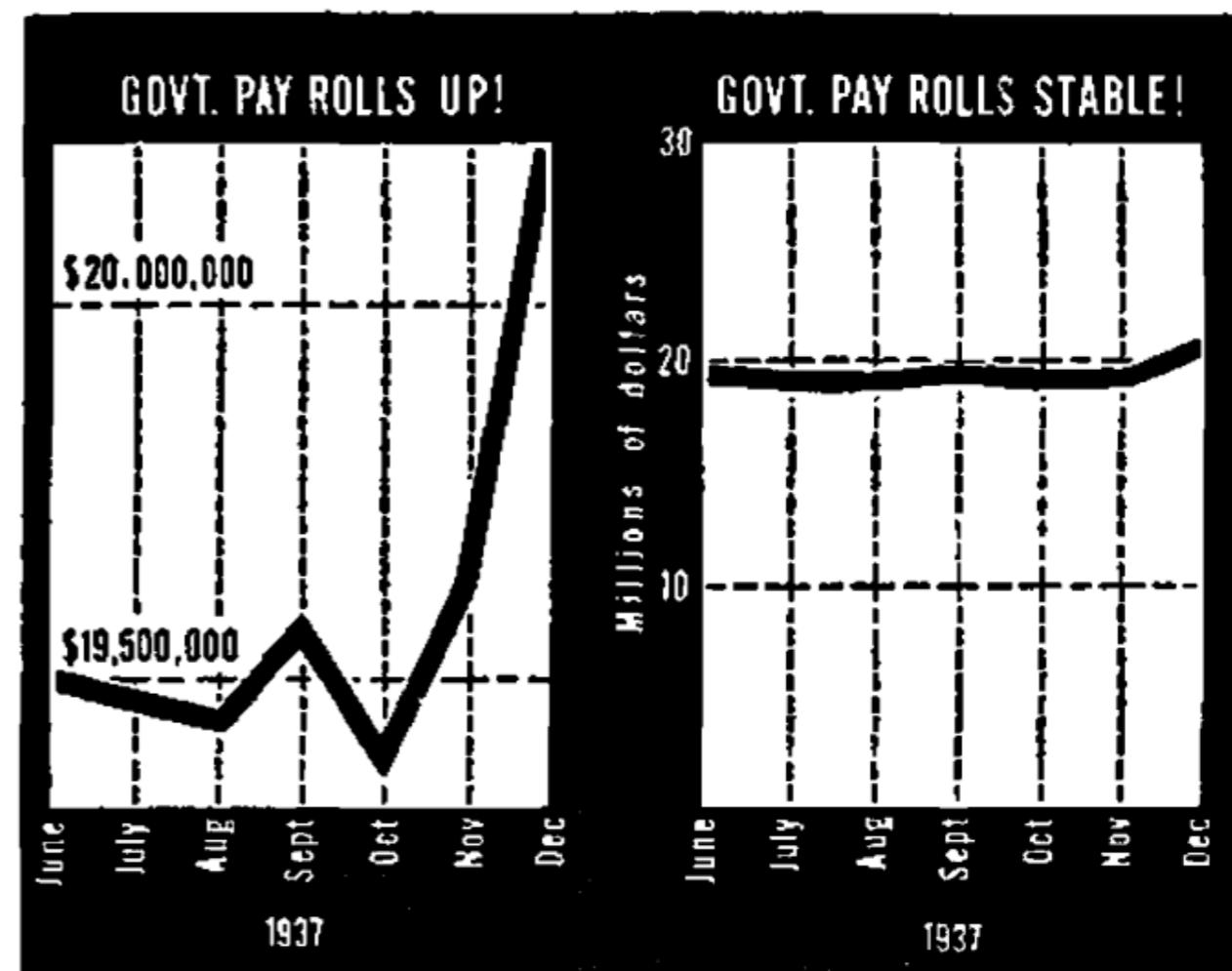
# How to Lie with Statistics



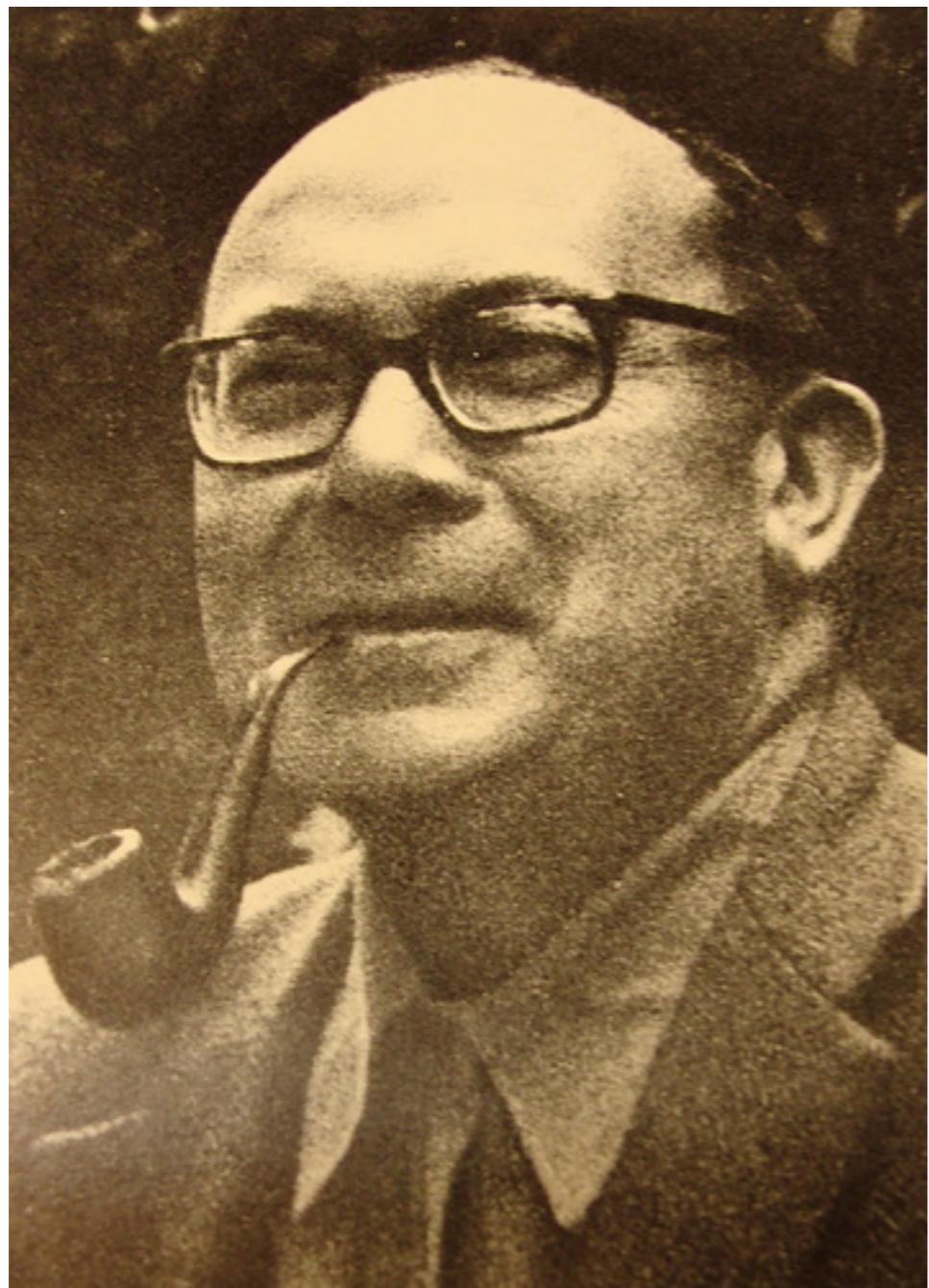
By  
DARRELL HUFF

Pictures by IRVING GELS

*31st printing*



**Jacques Bertin**



**John W. Tukey**



<http://stat-graphics.org/movies/prim9.html>

<https://bl.ocks.org/mbostock/4063663>

<http://www.datamake.io/project/tree-of-charts>



from Madison Ave, New York, NY 10016 to Greenwich Village, New York, NY

[Search Maps](#)

[Show search options](#)

Find businesses, addresses and places of interest. [Learn more](#)

[Get Directions](#) | [My Maps](#)

[Print](#) | [Send](#) | [Link](#)

Madison Ave, New York, NY 10016



Greenwich Village, New York, NY



[Add Destination](#) - [Show options](#)

By car

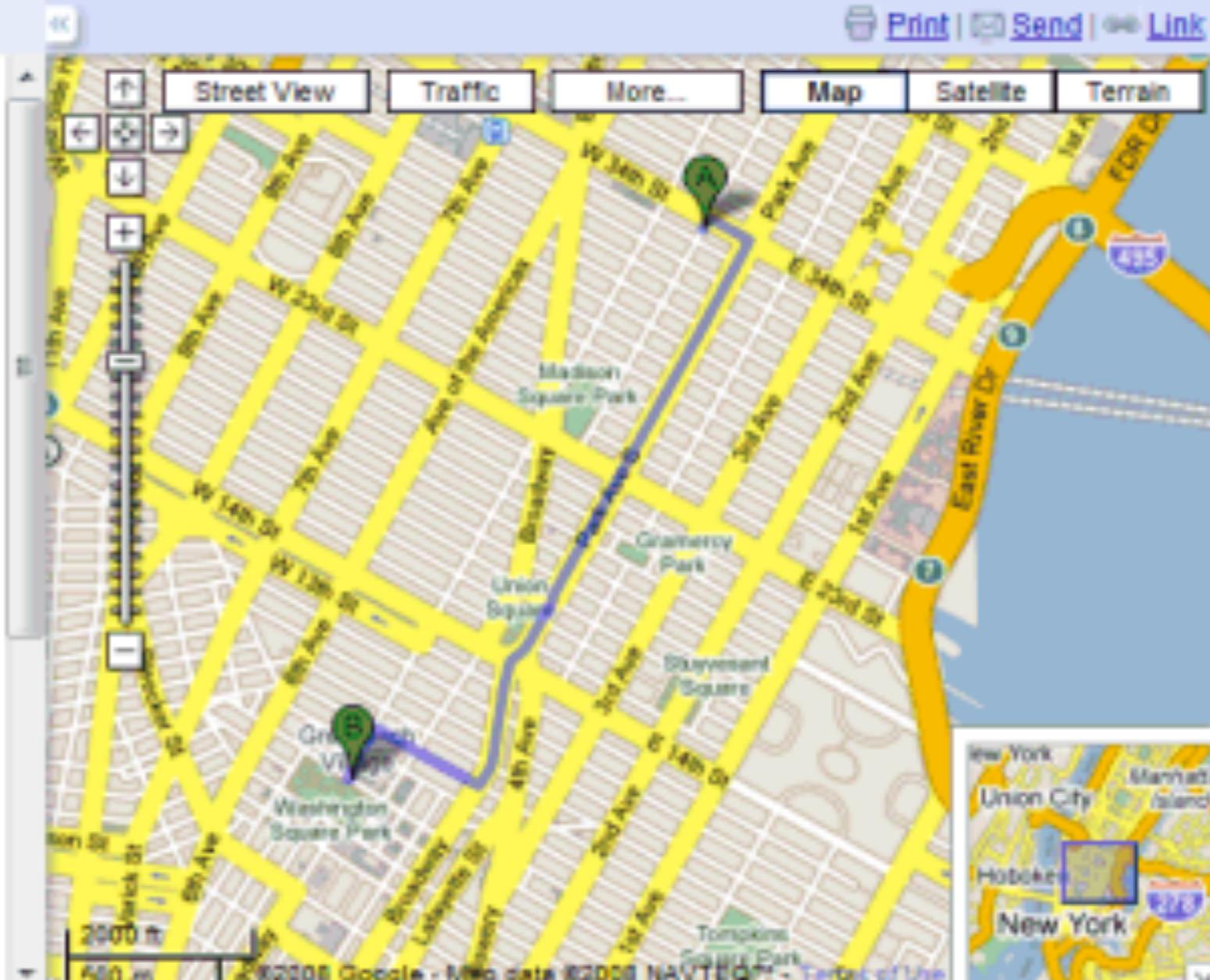
[Get Directions](#)

## Driving directions to Greenwich Village, New York, NY

1.7 mi – about 7 mins

Madison Ave  
New York, NY 10016

1. Head northeast on Madison Ave toward E 34th St 138 ft
2. Turn right at E 34th St 482 ft
3. Turn right at Park Ave 0.8 mi
4. Continue on Union Square E 0.2 mi
5. Union Square E turns slightly left and becomes Broadway 0.3 mi



# Examples

Uses	Types
Bread & Butter	Single Views Dashboards
Storytelling	Data journalism
Exploration	Knowledge bases
Novel uses	Augmented Reality

# Single charts

Market Summary > Apple Inc.

NASDAQ: AAPL

+ Follow

207.48 USD -14.74 (6.63%) ↓

Closed: 2 Nov, 19:59 GMT-4 · Disclaimer

After hours 206.50 -0.98 (0.47%)

1 day

5 days

1 month

6 months

YTD

1 year

5 years

Max



Open

209.55

Div yield

1.41%

High

213.65

Prev close

222.22

Low

205.43

52-wk high

233.47

Mkt cap

1.00T

52-wk low

150.24

P/E ratio

17.48

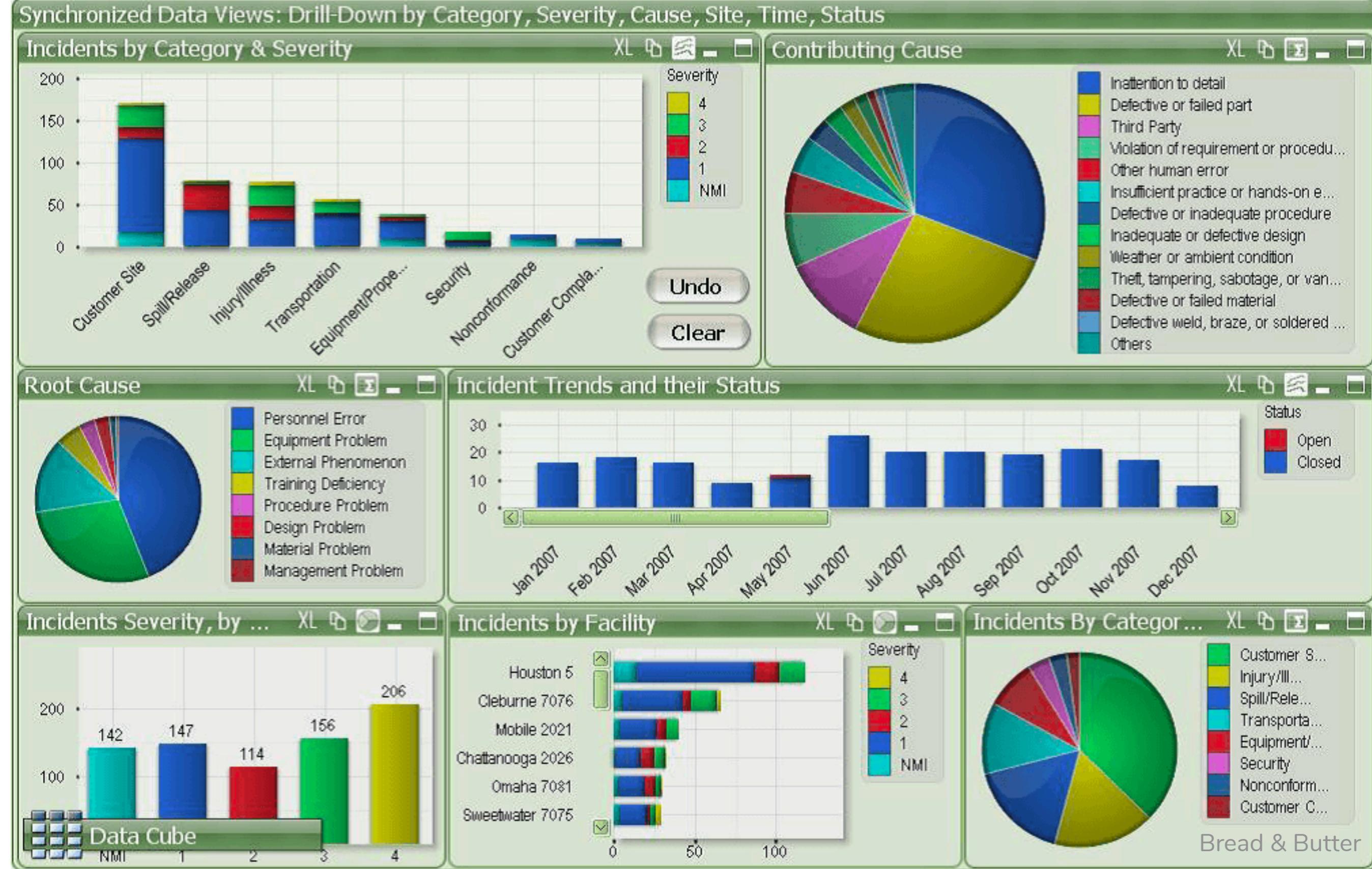
Google Finance

Bread & Butter

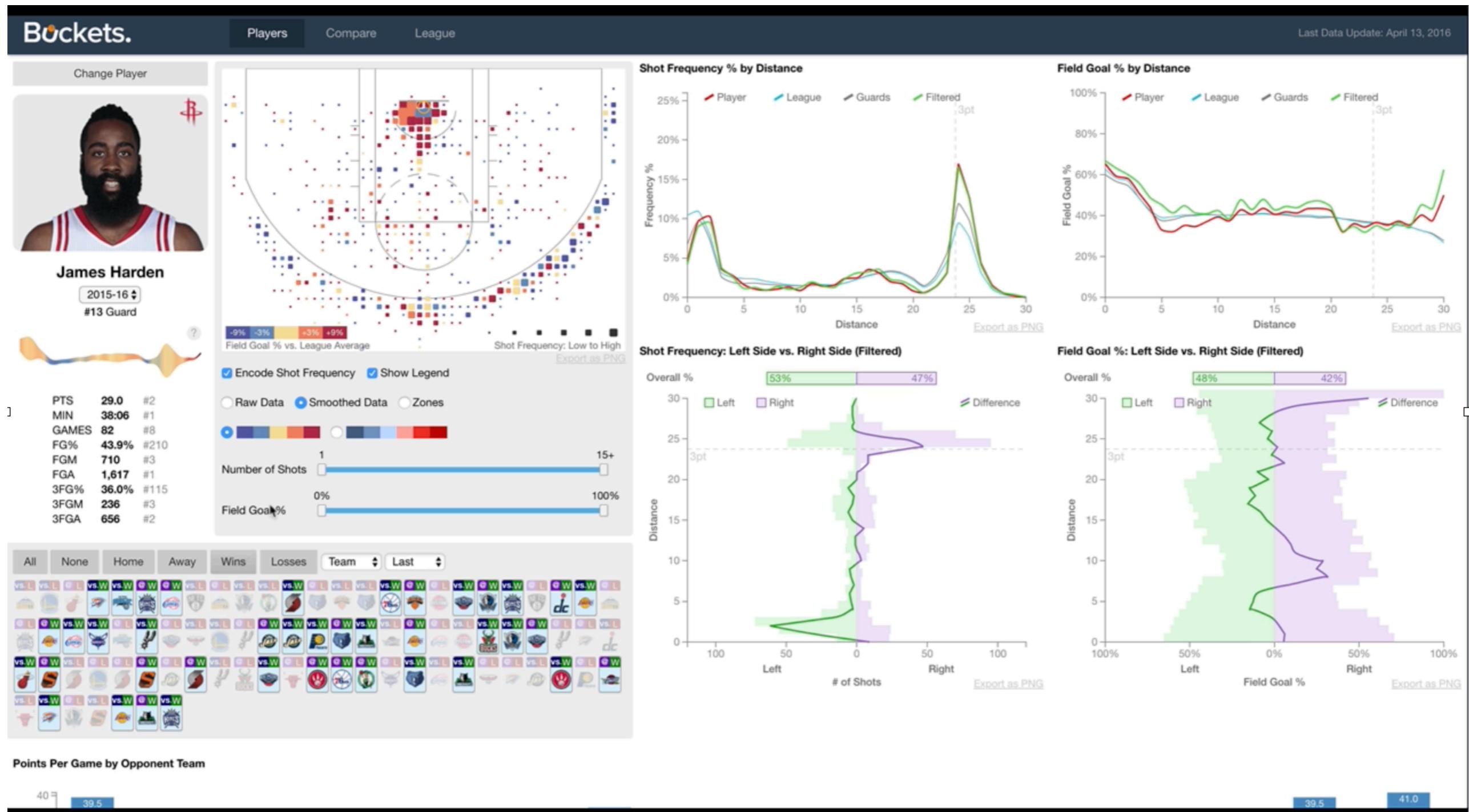
# Good dashboards

First Period: Algebra 1		As of May 1, 2012 (80% complete)											Help													
Student		Overall Course Grade					Assignments					Assessments		Attendance					Behavior							
		F	D	C	B	A	YTD	Spread	50%	Last	100%	Late	Last 5	Last	Jan 9	Feb 1	Mar 1	Apr 9	May 1	May 25	Ref	Det				
Frederick Chandler		F	.	X		53				.		2		41							2	2	1	1		
Bae Kim	E	F	.	X		59				.		3		39								3	2			
Fiona Reeves	S	D	X	.		65				.		3		47								1	1	1		
Brian Francis		D	.	X		65				.		2		67												
Anthony Harper	S	D	.	X		69				.	.	1		62												
Christopher Murphy		C	X	.		70				.	.	1		55												
Kirsten Holmes		C	.	X		72				.				67								1				
Roshawn Dawson		C	.	X		72				.				78								1	1			
Nikolas Mikhailovich	E	C	X	.		72				.	.			63												
James Martin		C	.	X		74				.		1		71												
Blaine Harper		C	X	.		75				.				71								1	1			
George Smith		C	.	X		76				.	.	1		76								1	1	1		
Regan Potrero		C	.	X		79				.		1		67												
Britta Jones		B	.	X		81				.				85												
Scott Ortiz		B	X	.		82				.		1		82												
Xu Mei		B	.	X		83				.				83												
Jaime Goss		B	.	X		84				.				82								1	1	1		
Samuel Miller		B	X	.		84				.				81												
Maria Garcia		B	.	X		86				.				72								2	1			
Jose Domingo		B	.	X		86				.				84												
Lawrence Parker		B	.	X		87				.				80								1	1			
Fariah Jackson		B	.	X		88				.				84												
Sarah Jameson		B	.	X		88				.		1		78								1	1			
David Chenowith		B	X	.		88				.				80								1				
Alison Perry		B	.	X		89				.				85												
Amala Singh		A	.	X		92				.				91								1				
Hannah Li		A	.	X		93				.				94												
James Snow		A	.	X	.	94				.				91												
Donald Chase		A	.		.	94				.				92								1	1			
Holly Norton		A	.	X	.	98				.				98												
E = English language deficiency S = Special Ed student																						13	11	6	6	
																						Last term	This term			
																						Tardies				

# Bad dashboards

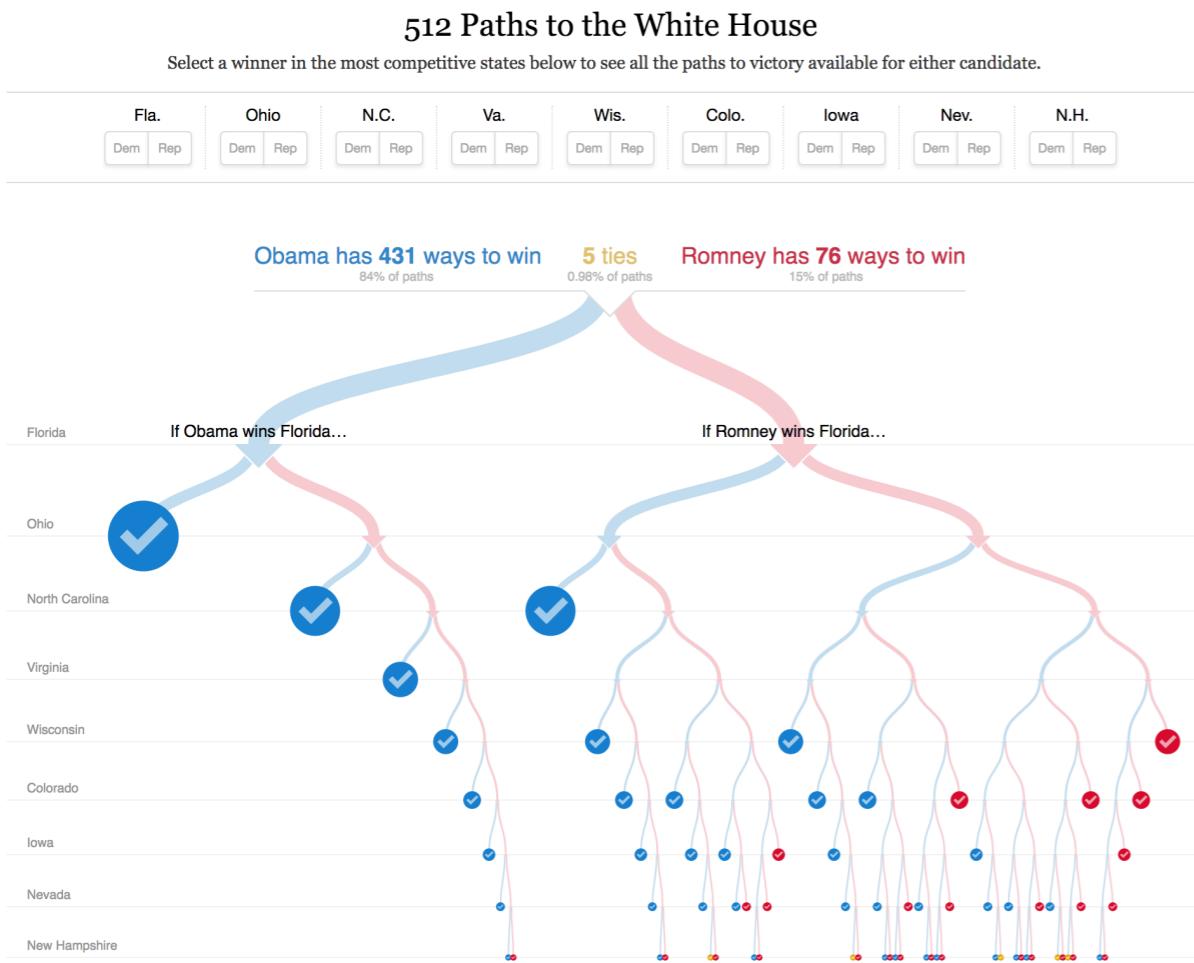


# Interactive dashboards



Bread & Butter

# Interactive data journalism



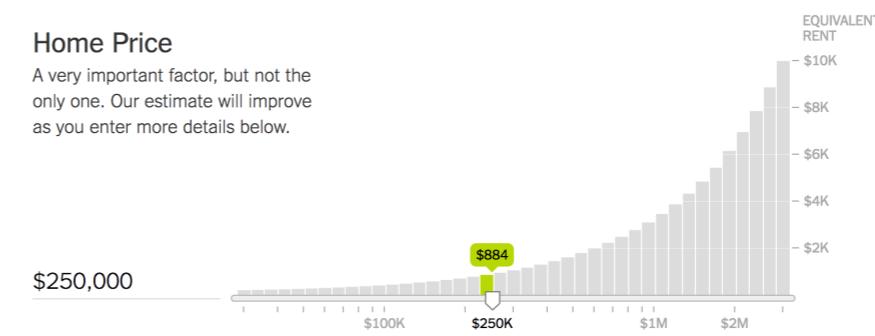
### Is It Better to Rent or Buy?

By MIKE BOSTOCK, SHAN CARTER and ARCHIE TSE

The choice between buying a home and renting one is among the biggest financial decisions that many adults make. But the costs of buying are more varied and complicated than for renting, making it hard to tell which is a better deal. To help you answer this question, our calculator takes the most important costs associated with buying a house and computes the equivalent monthly rent. [RELATED ARTICLE](#)

#### Home Price

A very important factor, but not the only one. Our estimate will improve as you enter more details below.



If you can rent a similar home for less than ...

**\$884 PER MONTH**

... then renting is better.

	Rent	Buy
Initial costs	\$884	\$60,000
Recurring costs	\$106,941	\$163,398
Opportunity costs	\$15,396	\$44,587
Net proceeds	-\$884	-\$145,649
Total	\$122,337	\$122,337

#### How Long Do You Plan to Stay?

Buying tends to be better the longer you stay because the upfront fees are spread out over many years.



How to Read the Charts Charts that are relatively flat indicate factors that are not particularly important to the outcome. Conversely, the factors that have steep slopes have a large impact.

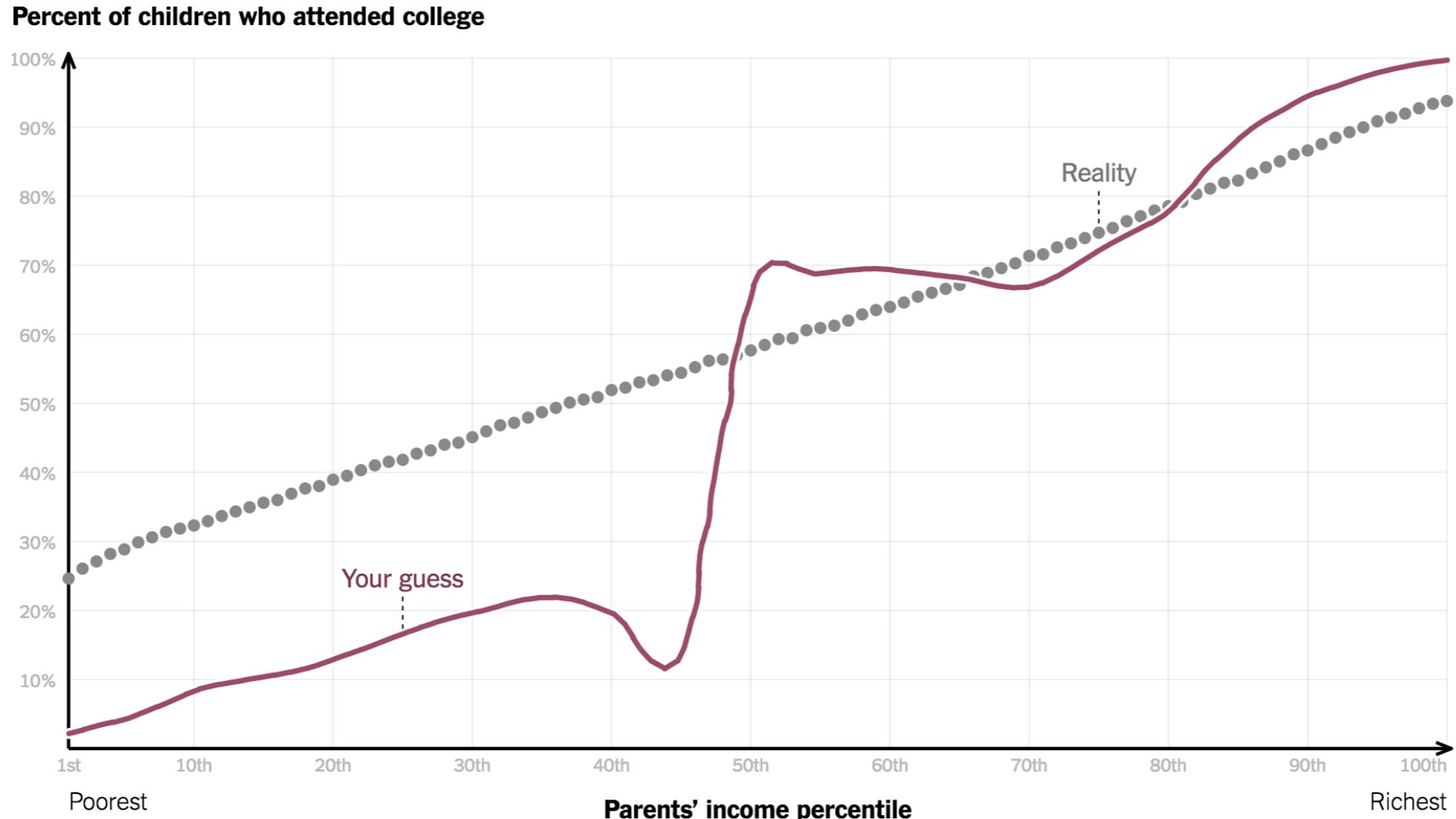
512 Paths to the White House

Rent Calculator

Storytelling

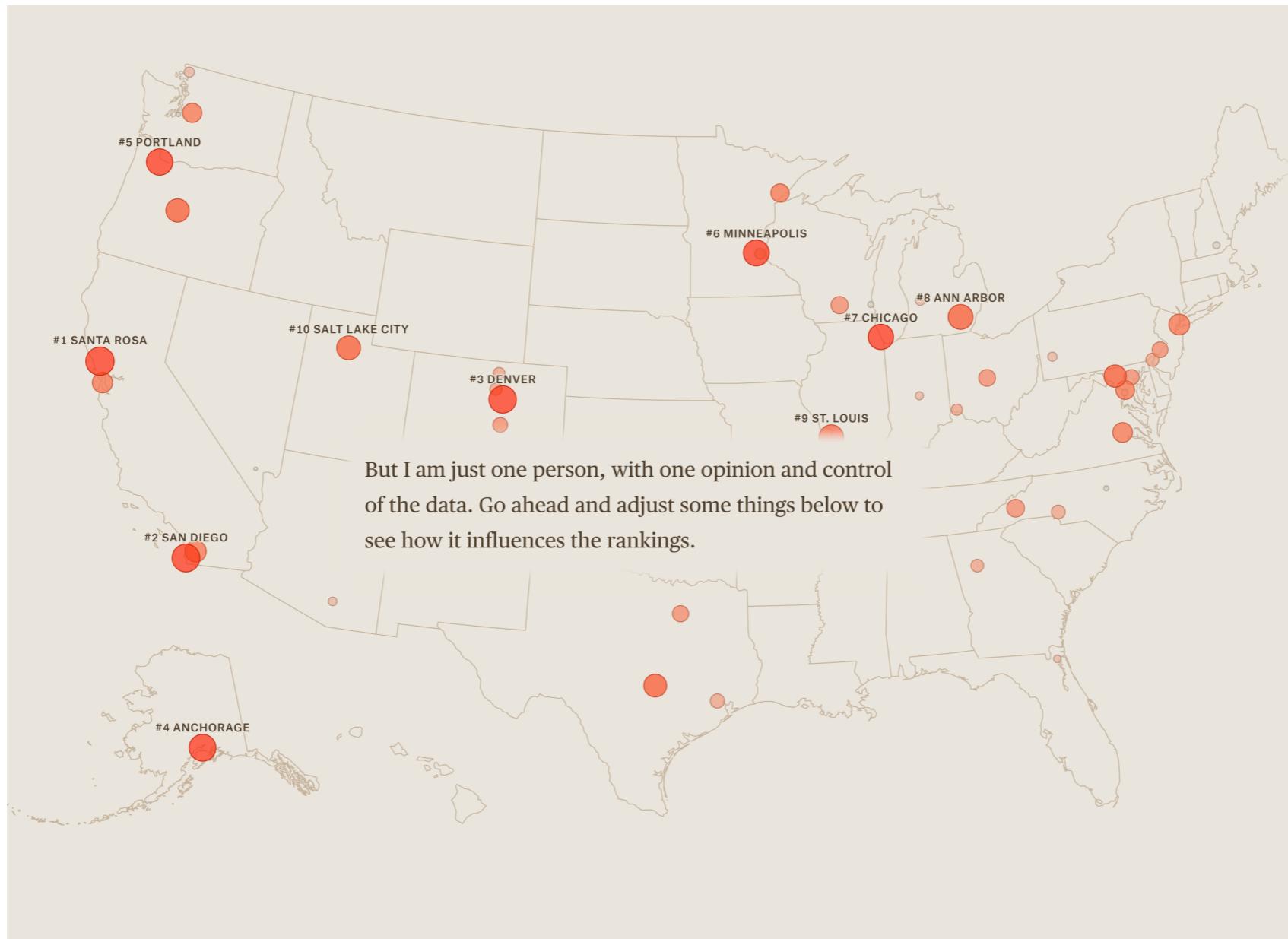
# Author - Reader cooperation

**Not the worst!**



You Draw It

# Scrollytelling



American Craft Beer

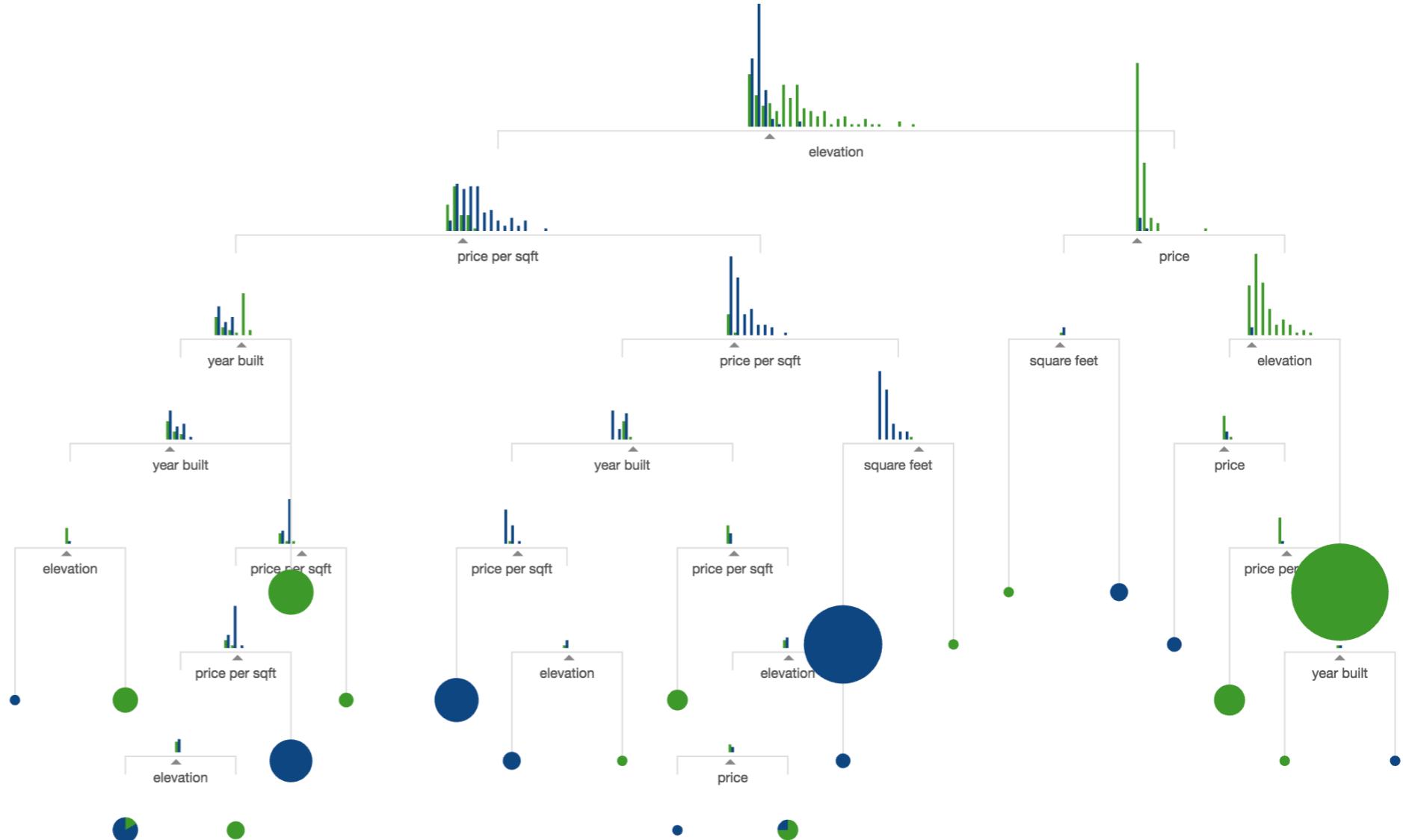
Storytelling

Graphics  
+  
Video  
+  
Story



# Explorative Explanations

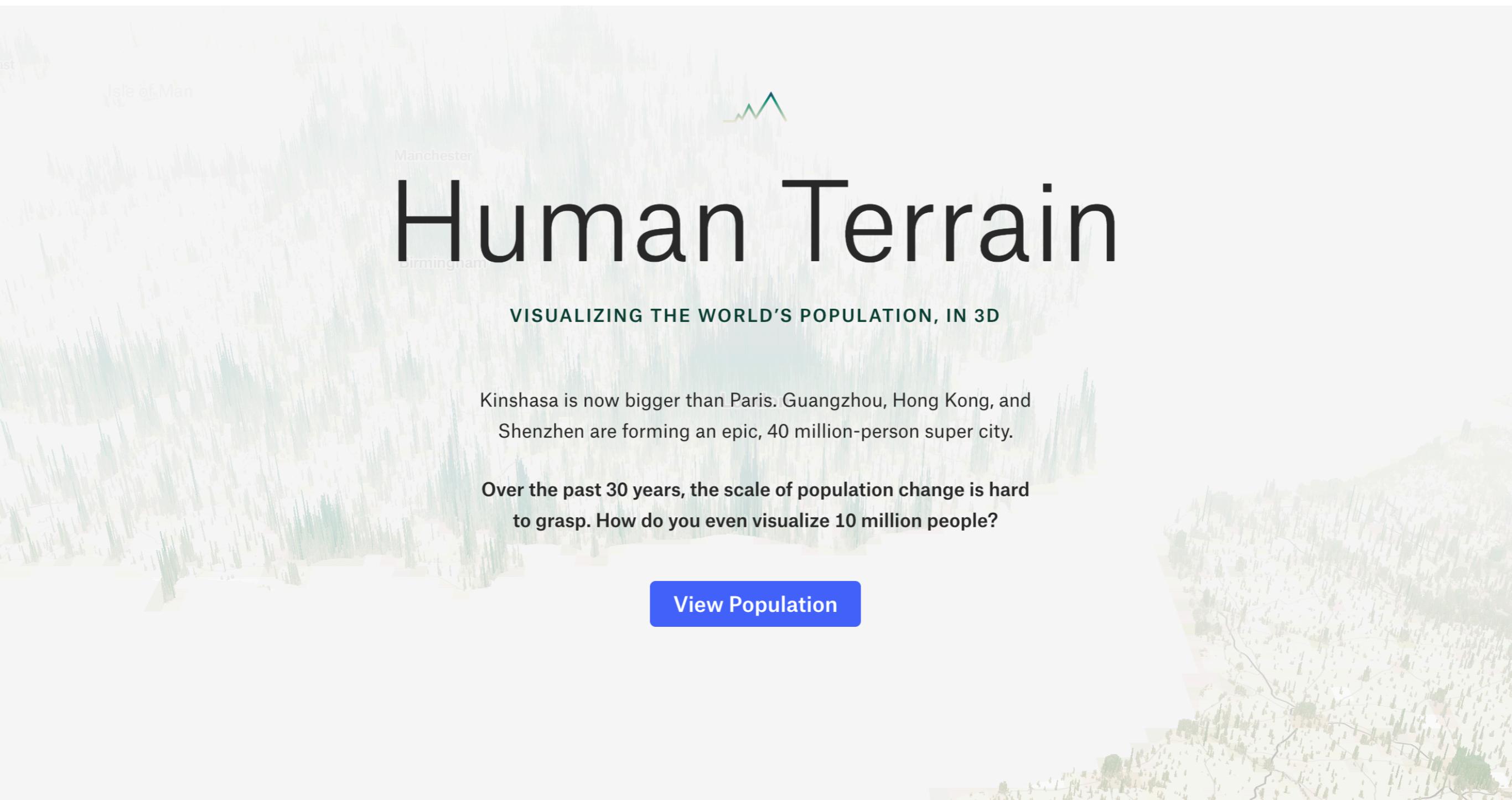
Adding several more layers, we get to  
**96%**.



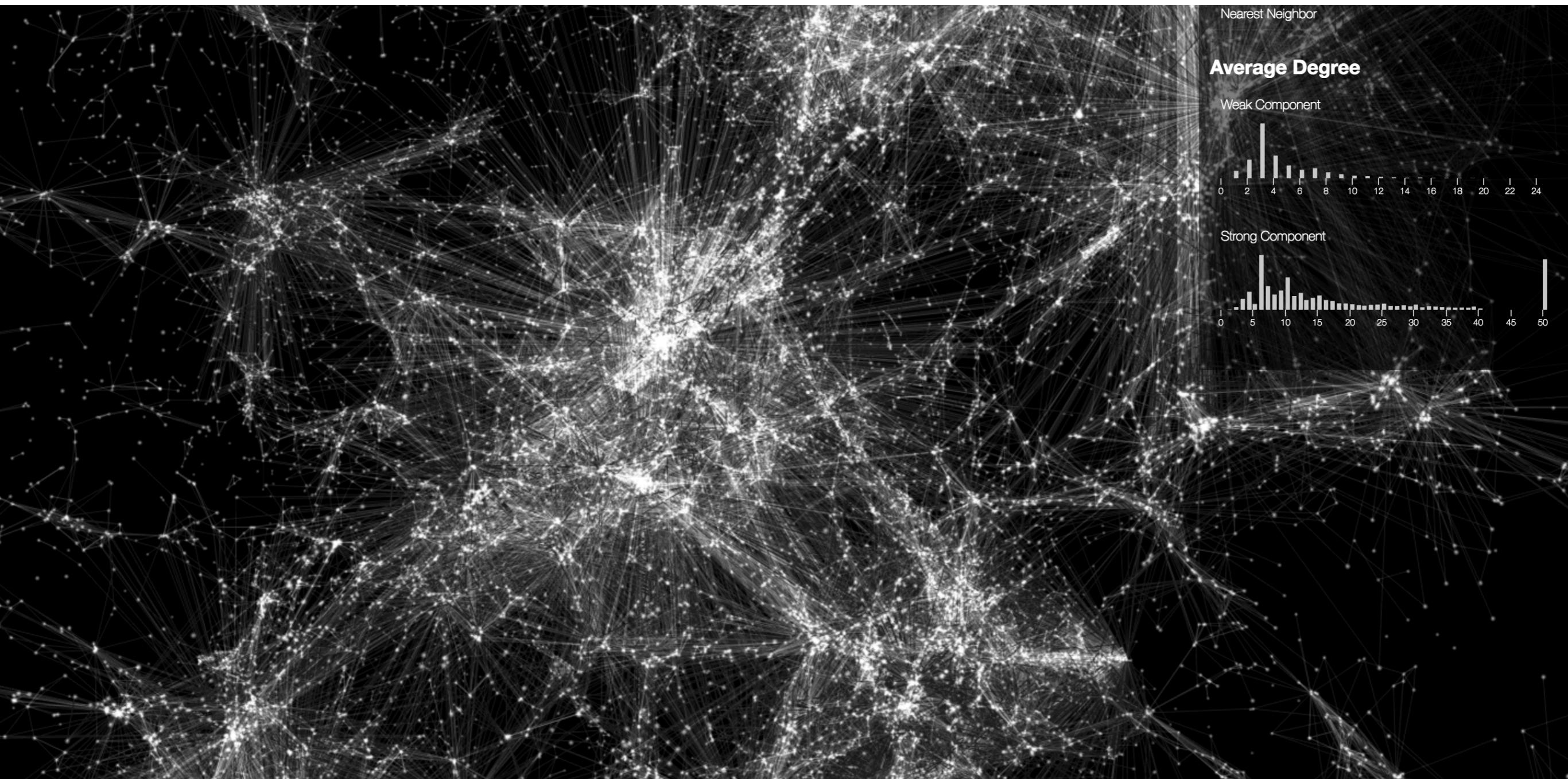
Introduction to Machine Learning

Storytelling

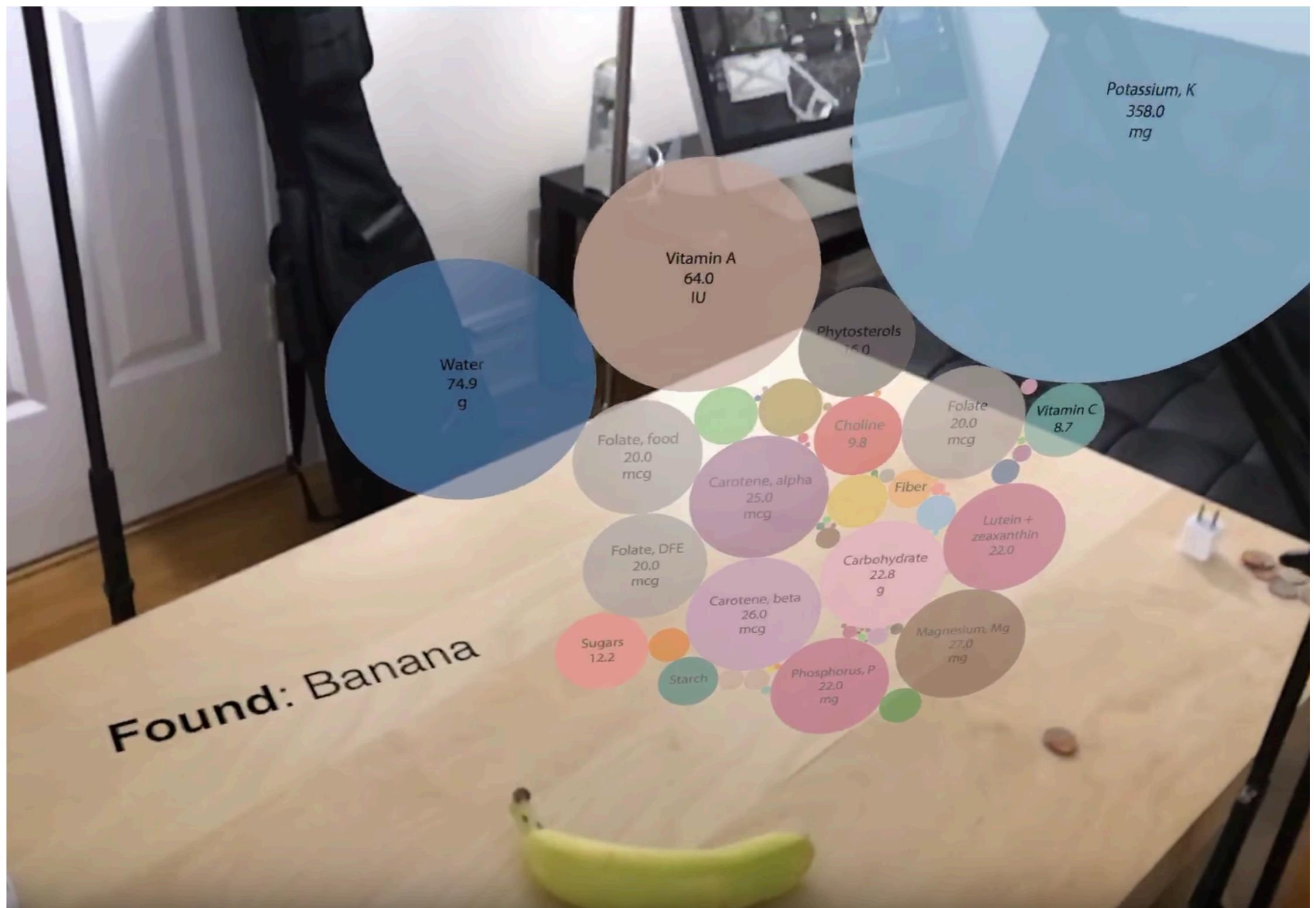
# Knowledge-bases



# Knowledge-bases



# Augmented Reality



Novel uses

# Summary

- We visualise for a long time
- 

- Stages of visualisation:

6000BC Scientific ivory tower

1700 The golden age of bureaucracy

1900 The dark ages of formal statistics

1950 The modern era of computation

- 
- Visualisation: data → precisely defined graphic
- 

- Author creates meaning by picking the message
- 

- Ways to bend the truth...