

# Introduction to Business Analytics: Communicating with Data

**Module 1**

**Professor Kevin Hartman**

I

# Welcome!

# Syllabus Review

I

## Topics and materials for Module 1

Module	Discussion Topics	References and Readings
1	<b>Pictures You See with Your Brain</b> —Frameworks and approaches behind data collection and analysis methods used by business professionals	<p>Berinato: Introduction, Chapters 1, 2, and 9</p> <p>A word from a founding father of web analytics: <a href="http://marketinggeek.blogspot.com/2006/02/geeks-in-marketing.html">http://marketinggeek.blogspot.com/2006/02/geeks-in-marketing.html</a></p> <p>Comparing Data Visualization Software—Here Are the 7 Best Tools for 2018: <a href="https://www.forbes.com/sites/bernardmarr/2018/06/20/comparing-data-visualization-software-here-are-the-7-best-tools-for-2018/">https://www.forbes.com/sites/bernardmarr/2018/06/20/comparing-data-visualization-software-here-are-the-7-best-tools-for-2018/</a></p> <p>A Curated Guide to the Best Tools, Resources and Technologies for Data Visualization: <a href="https://datavis.tools/">https://datavis.tools/</a></p> <p>The "Door" Study—YouTube: <a href="https://www.youtube.com/watch?v=FWSzSQssiQ">https://www.youtube.com/watch?v=FWSzSQssiQ</a></p> <p>What Make a Good Visualization? <a href="http://www.informationisbeautiful.net/visualizations/what-makes-a-good-data-visualization/">http://www.informationisbeautiful.net/visualizations/what-makes-a-good-data-visualization/</a></p>

# Syllabus Review

I

## Topics and materials for Module 2

Module	Discussion Topics	References and Readings
2	<b>Working Fast and Thinking Slow</b> —Data collection, analysis methods, visualization techniques and tools used by business professionals based on data volume, velocity, and usability	<p>Roles Required for Successful Web Analytics Endeavors: <a href="http://marketinggeek.blogspot.com/2006/11/data-visualization-practice.html">http://marketinggeek.blogspot.com/2006/11/data-visualization-practice.html</a></p> <p>Big Data, Big Deal: <a href="http://www.stephen-few.com/blog/2012/09/19/big-data-big-deal/">http://www.stephen-few.com/blog/2012/09/19/big-data-big-deal/</a></p> <p>Big Data, Big Dupe—A Progress Report: <a href="http://www.stephen-few.com/blog/2018/02/23/big-data-big-dupe-a-progress-report/">http://www.stephen-few.com/blog/2018/02/23/big-data-big-dupe-a-progress-report/</a></p> <p>Different Tools for Different Tasks: <a href="http://www.stephen-few.com/blog/2018/02/19/different-tools-for-different-tasks/">http://www.stephen-few.com/blog/2018/02/19/different-tools-for-different-tasks/</a></p>

# Syllabus Review

I

## Topics and materials for Module 3

Module	Discussion Topics	References and Readings
3	<b>Finding Your Data Story</b> —Understanding visual form and how to begin the transition from work product to client-ready graphics	Berinato: Chapters 3 and 4  Wong: Introduction, Chapter 1  The World's Best Data Visualizations (2017): <a href="http://www.informationisbeautiful.net/2017/the-best-infographics-and-data-visualizations-2017/">http://www.informationisbeautiful.net/2017/the-best-infographics-and-data-visualizations-2017/</a>  FlowingData's R Visualization Tutorials: <a href="https://flowingdata.com/category/tutorials/">https://flowingdata.com/category/tutorials/</a>

# Syllabus Review

I

## Topics and materials for Module 4

Module	Discussion Topics	References and Readings
4	<b>Getting Your Story Across</b> —Distinctions between effective and poor data presentation	Berinato: Chapters 5–8  Wong: Chapters 2–5  Know Your Audience—Good Luck with That: <a href="https://www.stephenfew.com/blog/2018/03/15/know-your-audience-good-luck-with-that/">https://www.stephenfew.com/blog/2018/03/15/know-your-audience-good-luck-with-that/</a>  Don't Refrain from Using Uncommon Visualizations—Explain Them: <a href="http://www.thefunctionalart.com/2017/12/dont-refrain-from-using-uncommon.html">http://www.thefunctionalart.com/2017/12/dont-refrain-from-using-uncommon.html</a>  The "Action" Dashboard: <a href="http://www.kaushik.net/avinash/the-action-dashboard-an-alternative-to-crappy-dashboards/">http://www.kaushik.net/avinash/the-action-dashboard-an-alternative-to-crappy-dashboards/</a>

# Course Resources

## Required reading: Berninato's *Good Charts*



(Berinato, 2016)

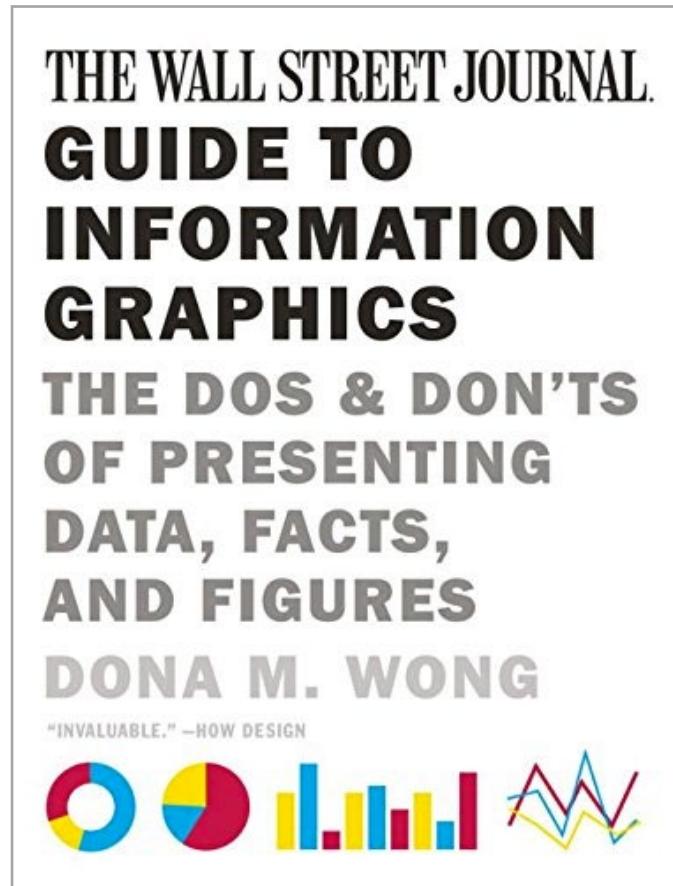
In *Good Charts*, Berinato provides an essential guide to how visualization works and how to use the new language of dataviz to impress and persuade. Dataviz today is where spreadsheets and word processors were in the early 1980s—on the cusp of changing how we work.

This book is much more than a set of static rules for making visualizations. It taps into both well-established and cutting-edge research in visual perception and neuroscience, as well as the emerging field of visualization science, to explore why good charts (and bad ones) create “feelings behind our eyes.”

*Good Charts* will help you turn plain, uninspiring charts that merely present information into smart, effective visualizations that powerfully convey ideas.

# Course Resources

Required reading: Wong's *The WSJ Guide*



(Wong, 2013)

The definitive guide to the graphic presentation of information.

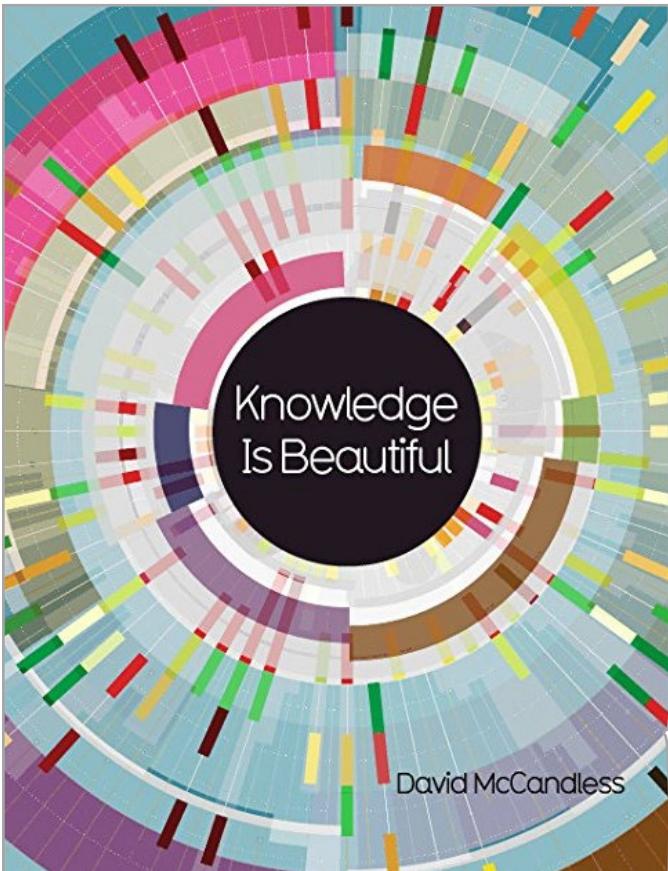
Information graphics is rarely taught in schools or is the focus of on-the-job training. Now, for the first time, Dona M. Wong, a student of the information graphics pioneer Edward Tufte, makes this material available to all of us. In this book, you will learn

- to choose the best chart that fits your data,
- the most effective way to communicate with decision makers when you have five minutes, and
- how to use color effectively.

Organized in a series of mini-workshops backed up with illustrated examples, this is an invaluable reference work for students and professionals in all fields.

# Course Resources

Recommended reading: McCandless's *Knowledge Is Beautiful*



(McCandless, 2014)

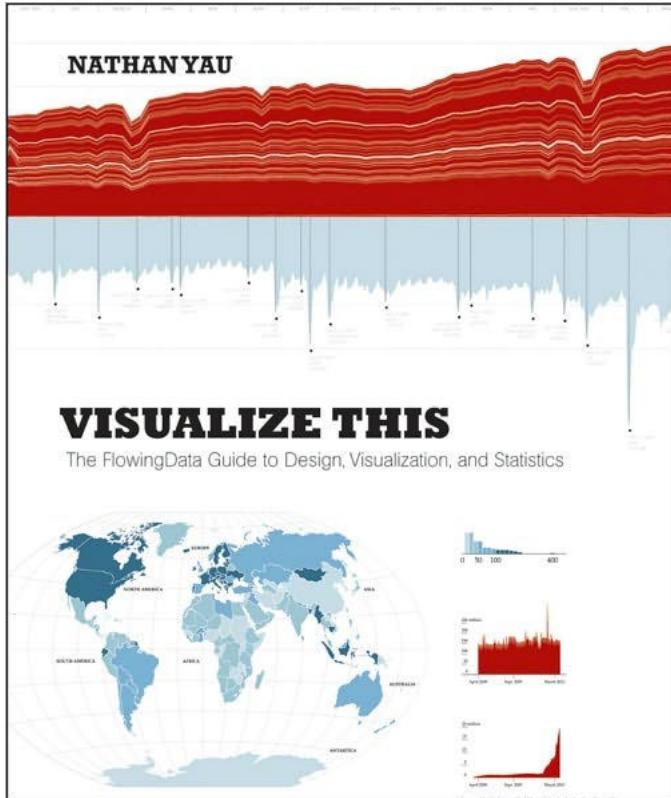
We are living in the Information Age, in which we are constantly bombarded with data on television, in print, and online. How can we relate to this mind-numbing overload?

Enter David McCandless and his amazing infographics: simple, elegant ways to understand information too complex or abstract to grasp in any way but visually. In his unique signature style, he creates dazzling displays that blend facts with their connections, contexts, and relationships, making information meaningful, entertaining, and beautiful.

Taking infographics to the next level, his book *Knowledge Is Beautiful* is an endlessly fascinating spin through the world of visualized data, which offers a deeper, more wide-ranging look at the world and its history. Covering everything from dog breeds and movie plots to the origins of life and a timeline of the far future, this stunning book is guaranteed to enrich your understanding of the world.

# Course Resources

Recommended reading: Yau's *Visualize This*



(Yau, 2011)

This book, by the data visualization expert and the creator of flowingdata.com Nathan Yau, presents a unique approach to visualizing and telling stories with data.

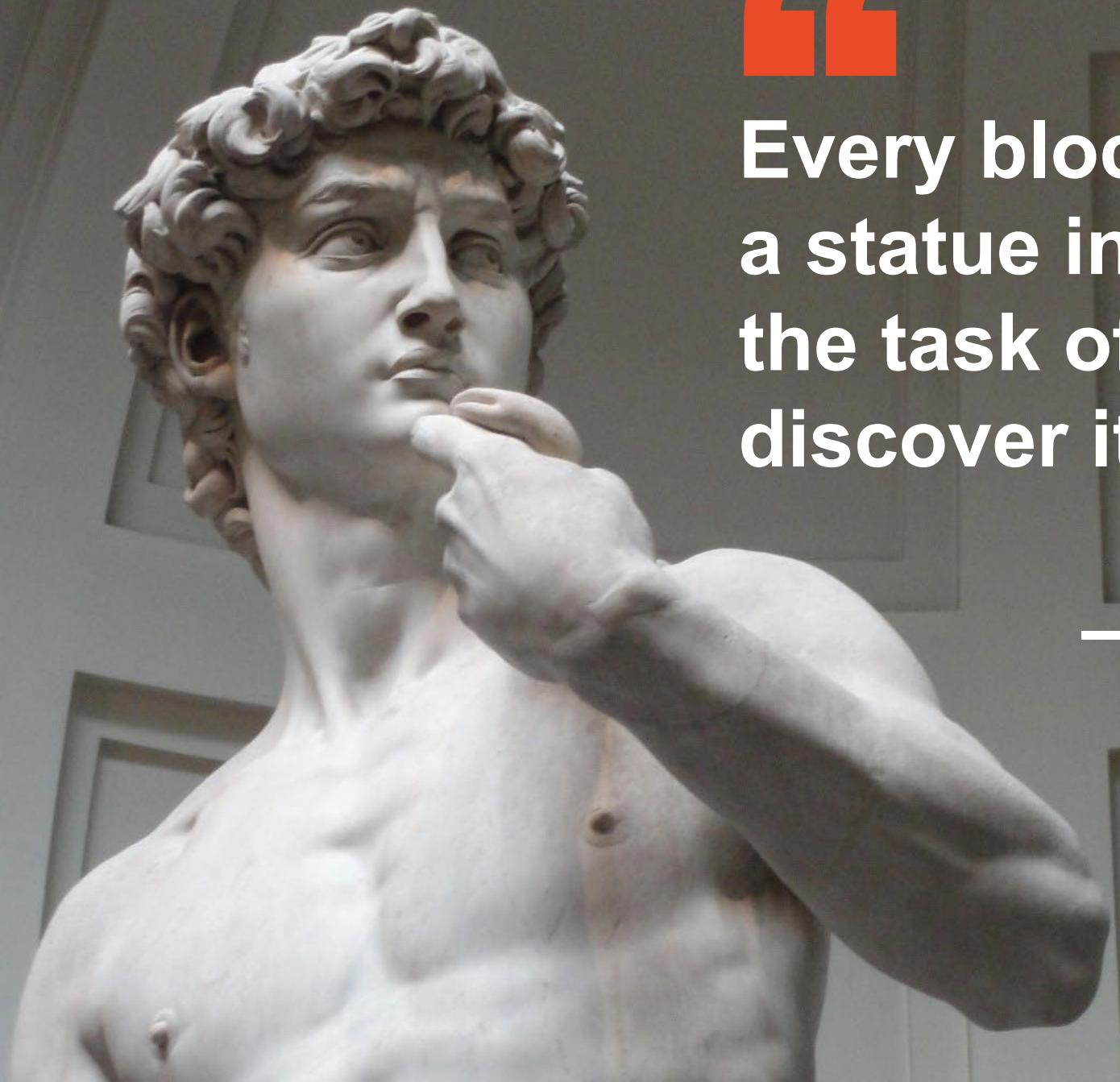
The book offers step-by-step tutorials and practical design tips for creating statistical graphics, geographical maps, and information design to find meaning in the numbers.

It details tools that can be used to visualize data-native graphics for the web—such as ActionScript, Flash libraries, PHP, and JavaScript—and tools to design graphics for print, such as R and Illustrator.

It contains numerous examples and descriptions of patterns and outliers and explains how to show them.

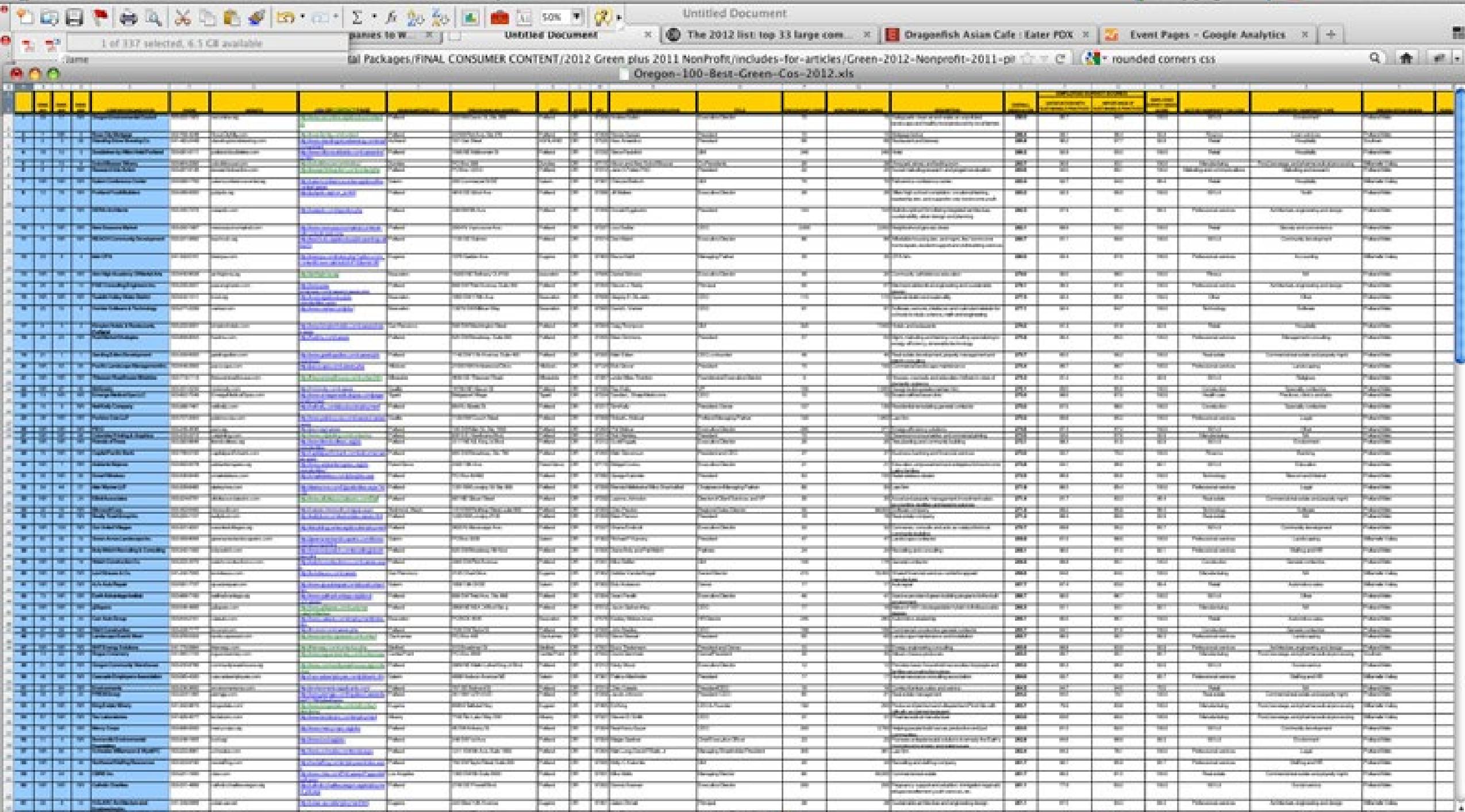


A Centering Thought ...



“  
Every block of stone has  
a statue inside it and it is  
the task of the sculptor to  
discover it.”

— Michaelangelo



“

Visualization is hard. I've written books, created software, directed films in my career, but visualization is by far the most challenging discipline I've ever engaged with. It's something about the precision needed at every level, I think. Concept, data, story, design, style—all are precision arts. In visualization, they're stacked one on top of the other. If one sags or slips, the entire edifice can collapse.

”

— David McCandless

I

**200 billion hours**  
per year spent watching TV by US adults

I

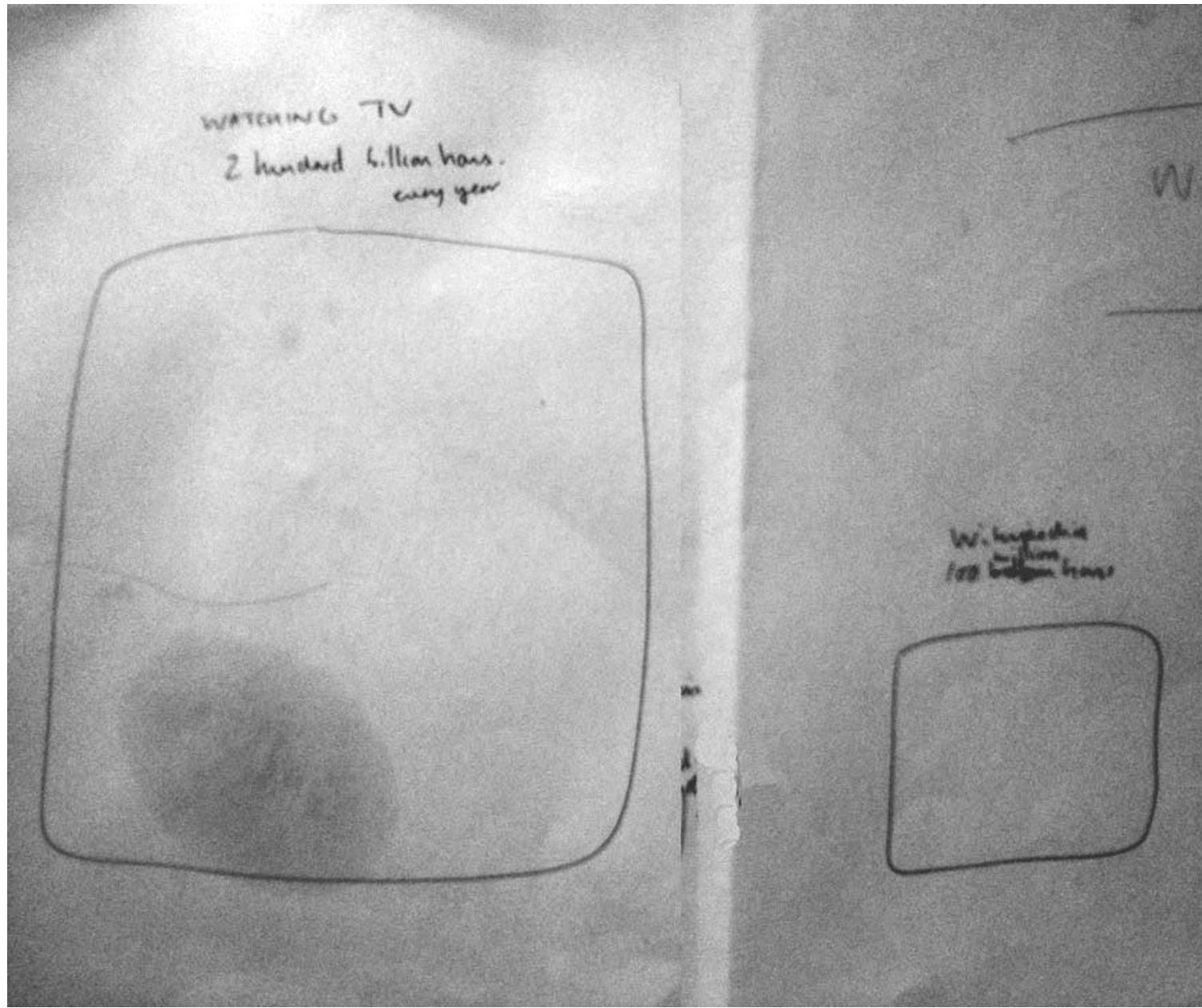
**200 billion hours**  
per year spent watching TV by US adults



**200 billion hours**  
per year spent watching TV by US adults

■ **100 million hours**  
to create Wikipedia

I



# In This Module

I

## Module 1: Pictures You See with Your Brain

### Key Concepts

Studying the history of dataviz

Understanding today's dataviz tools

Making connections with visuals

Evaluating the effectiveness of dataviz

# References

Berinato, S. (2016). *Good Charts: The HBR Guide to Making Smarter, More Persuasive Data Visualizations*. Brighton, MA: Harvard Business Review Press

McCandless, D. (2010). [Cognitive Surplus Visualized](#) [Online image]

McCandless, D. (2014). *Knowledge Is Beautiful: Impossible Ideas, Invisible Patterns, Hidden Connections—Visualized*. New York, NY: Harper Design

Wong, D. M. (2013). *The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts, and Figures*. New York, NY: W. W. Norton & Company

Yau, N. (2011). *Visualize This: The Flowing Data Guide to Design, Visualization, and Statistics*. Hoboken, NJ: Wiley

A photograph of three students sitting on a bench in a hallway. In the foreground, two female students are looking at a yellow notebook together; one is smiling. In the background, a male student in an orange shirt is looking down at his laptop. The background is slightly blurred.

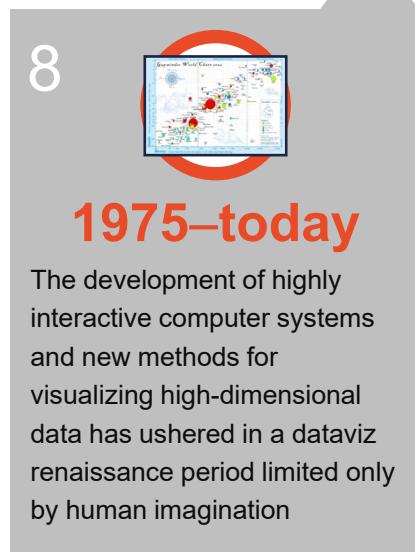
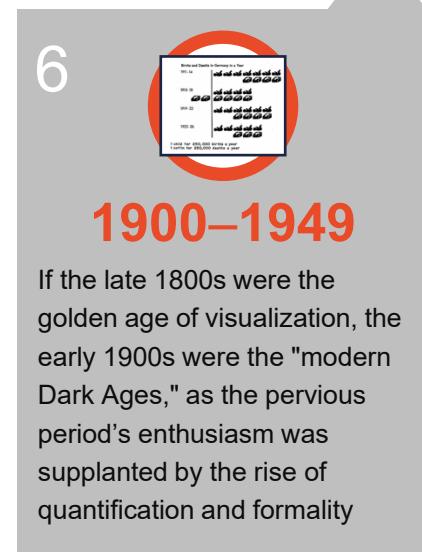
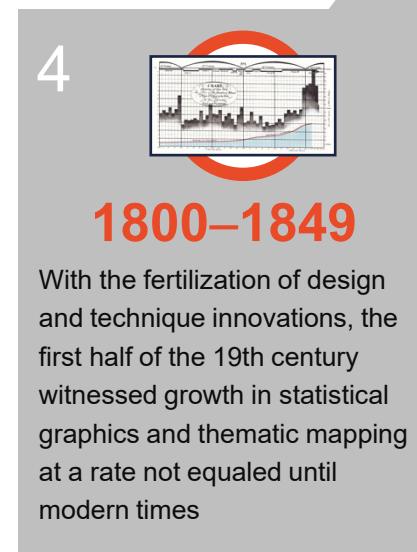
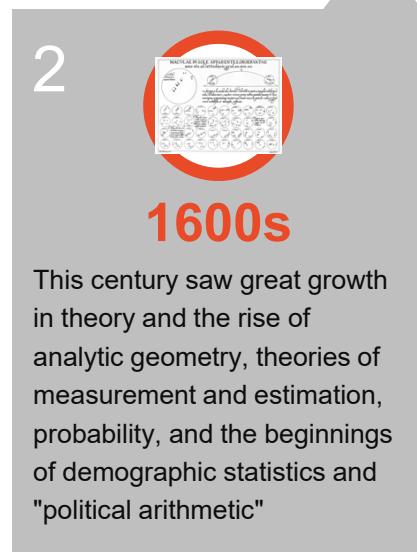
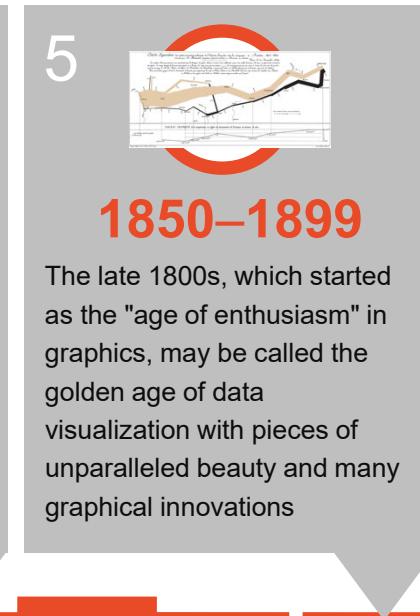
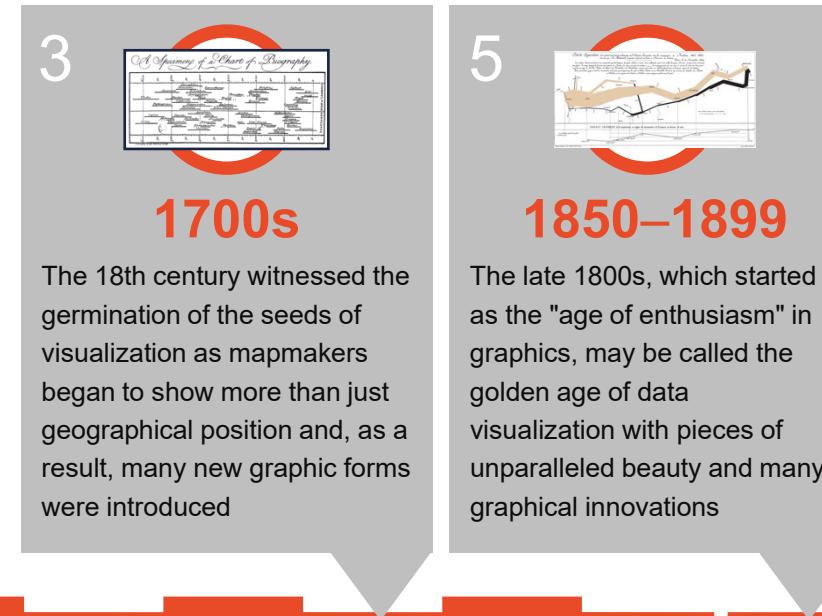
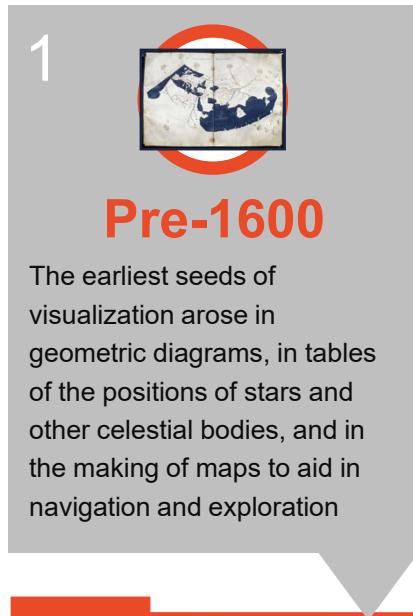
Module 1 Lesson 1

# Studying the History of Dataviz

# Studying the History of Dataviz

I

Dataviz has a long history, starting with mapmaking and fanning into statistics, politics, medicine, and other fields



**A** 6200 BC: Town map, with an erupting volcano (Hasan Daö?) and the Konya Plain, possibly the first map created

**B** 150: Ptolemy's world map

**C** 1375: Catalan Atlas

## Epoch I: Early Maps And Diagrams (Pre-1600)

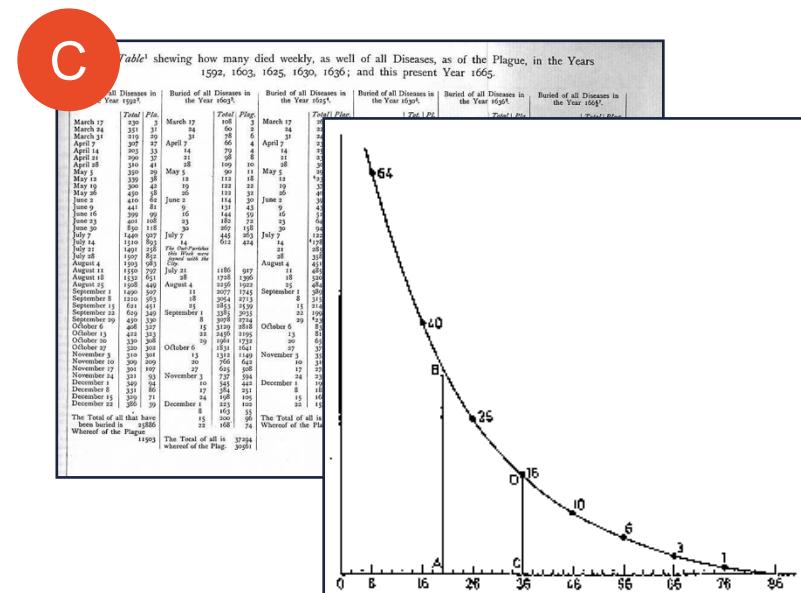
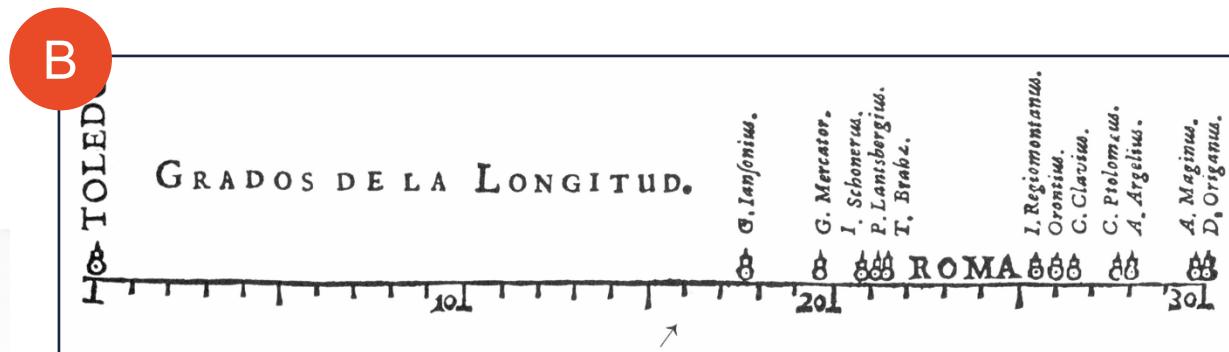
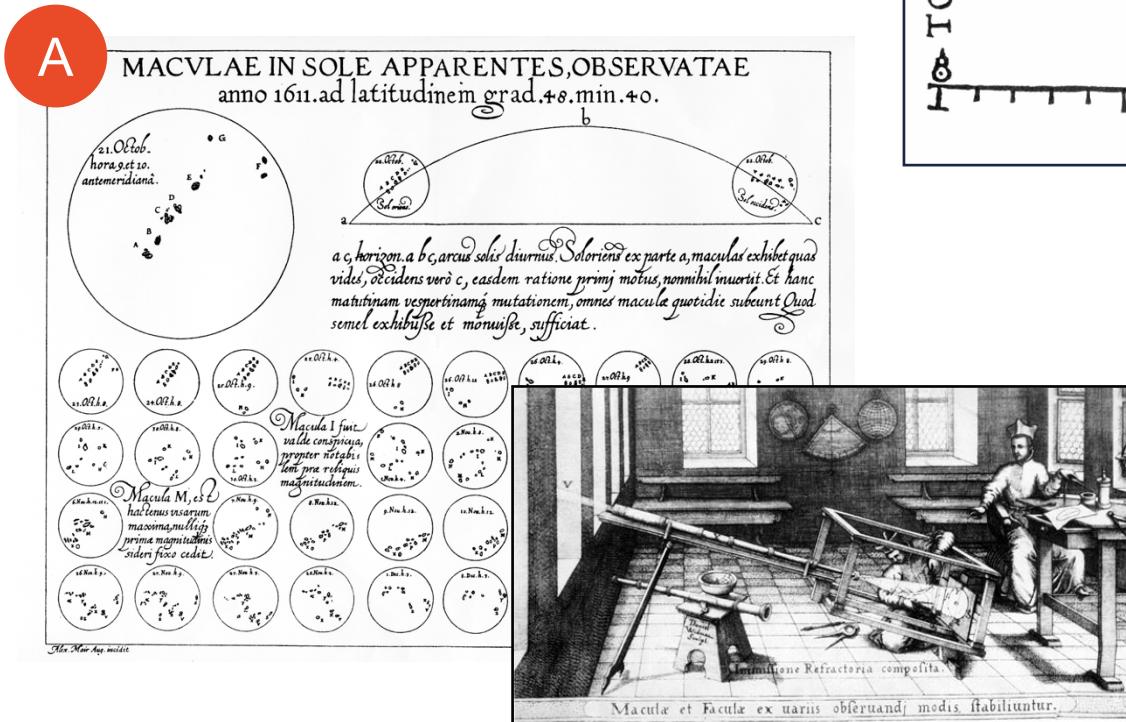


**A** 1626: Christopher Schooner's visual representations used to chart the changes in sunspots over time

1644: Michael F. van Langren;  
variations in determination of  
longitude between Toledo and  
Rome

1669: Christiaan Huygens; first graph of a continuous distribution function, a graph of Gaunt's life table

## Epoch II: Measurement and Theory (1600s)

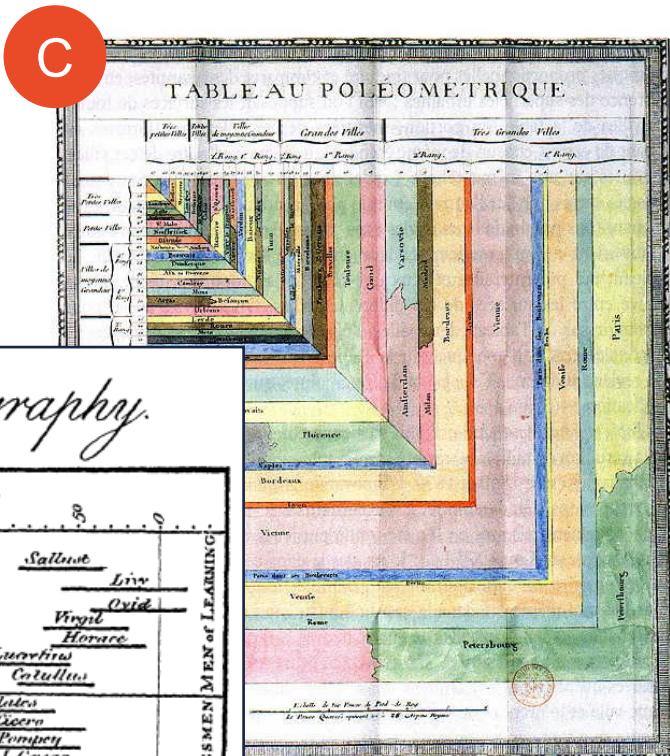
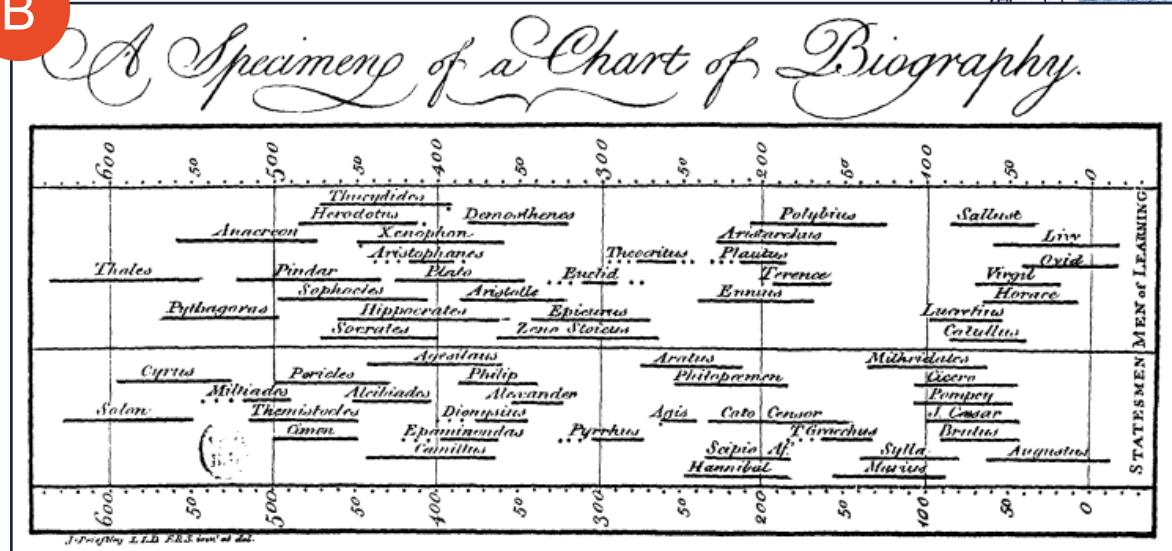
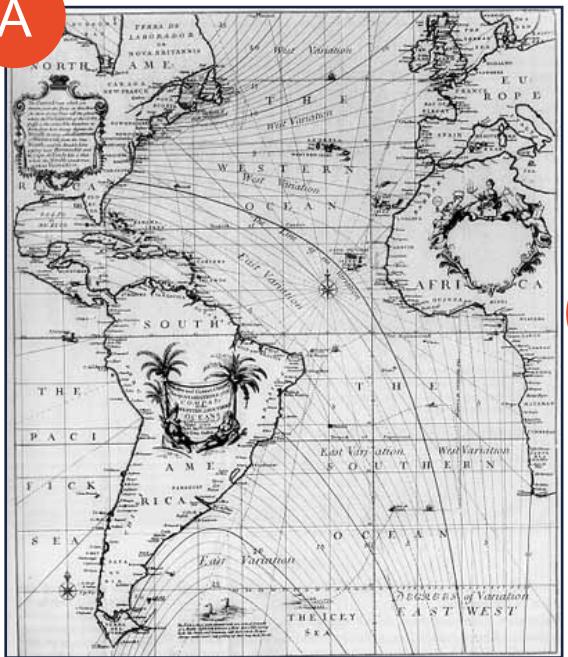


1701: Edmond Halley; contour maps showing curves of equal value (an isogonic map, lines of equal magnetic declination for the world, possibly the first contour map of a data-based variable)

**1765:** Joseph Priestley;  
life spans of 2,000  
famous people, 1200  
B.C. to 1750 A.D.,  
quantitative comparison  
by means of bars

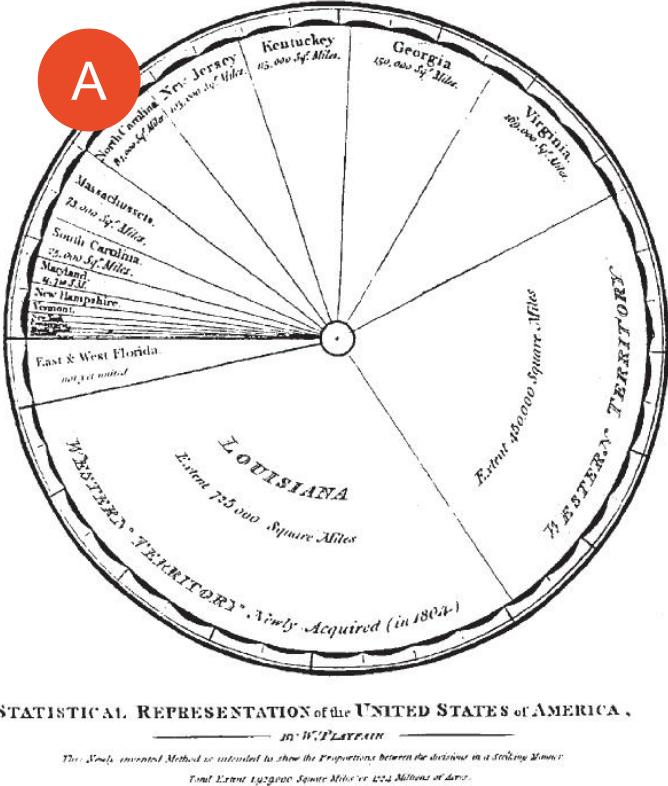
1782: Charles de Fourcroyde; use of geometric, proportional figures (squares) to compare demographic quantities by superposition, an early "tableau graphique"

## Epoch III: New Graphic Forms (1700s)

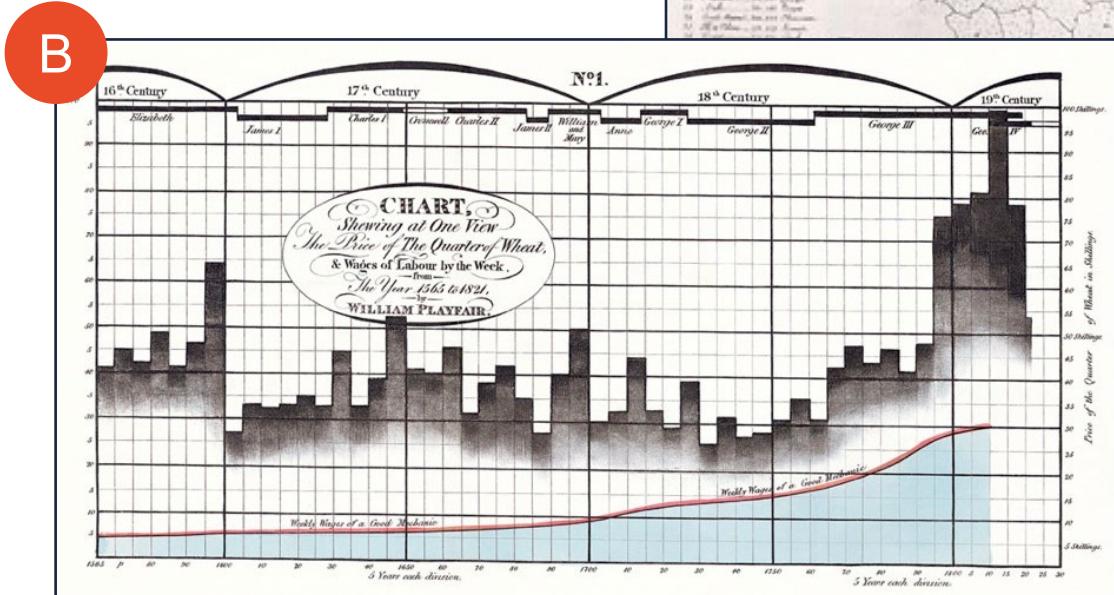


**A** 1801: William Playfair; invention of the pie chart and circle graph, used to show part-whole relations

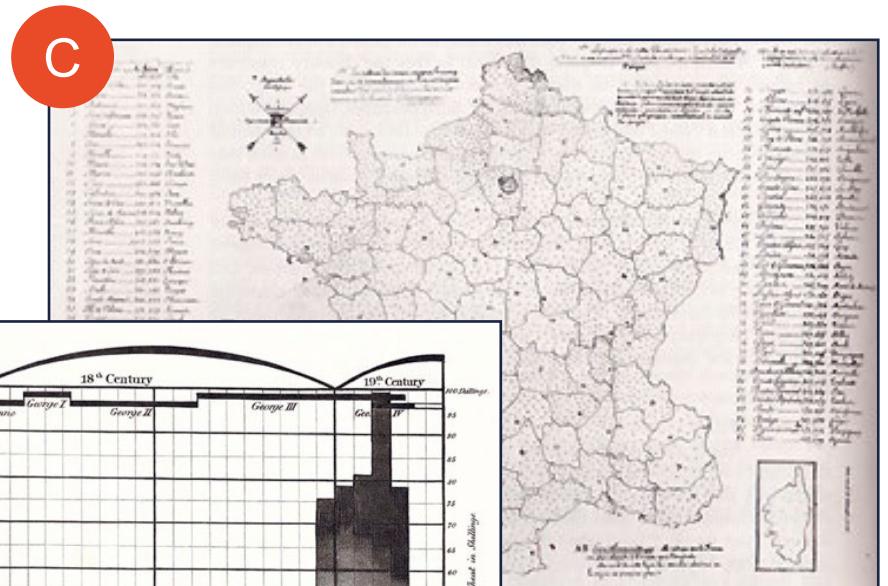
## Epoch IV: Beginnings of Modern Graphics (1800–1849)



**B** 1821: Playfair; time series graph of prices, wages, and ruling monarch



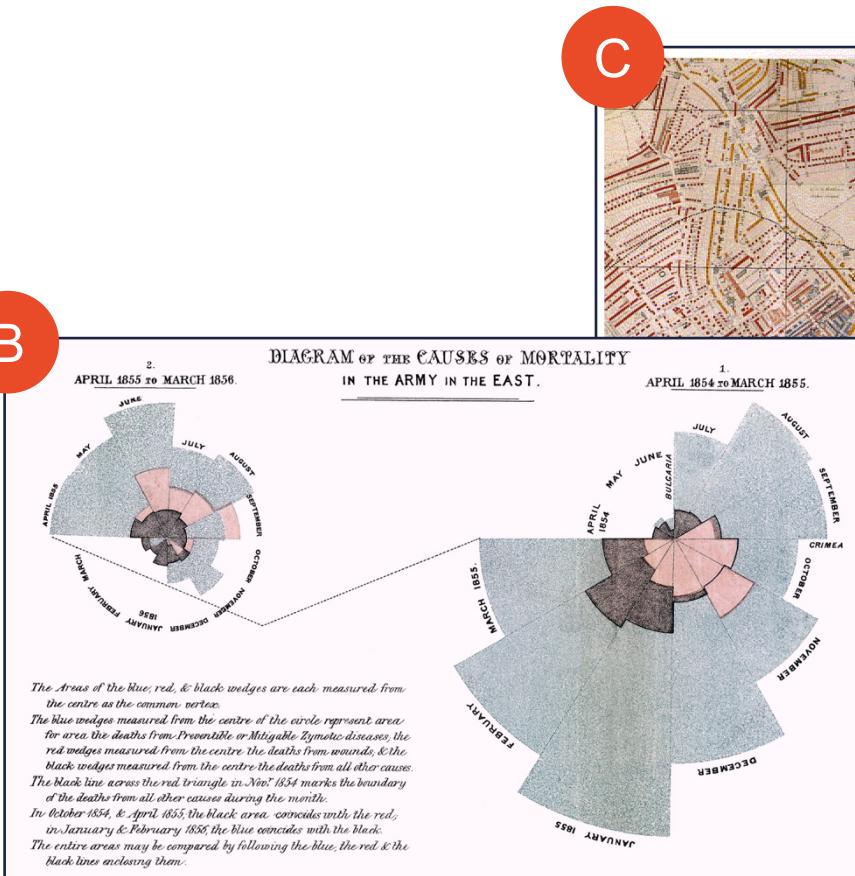
**C** 1830: Armand Joseph Frere de Montizon; first simple dot map of population, 1 dot = 10,000 people



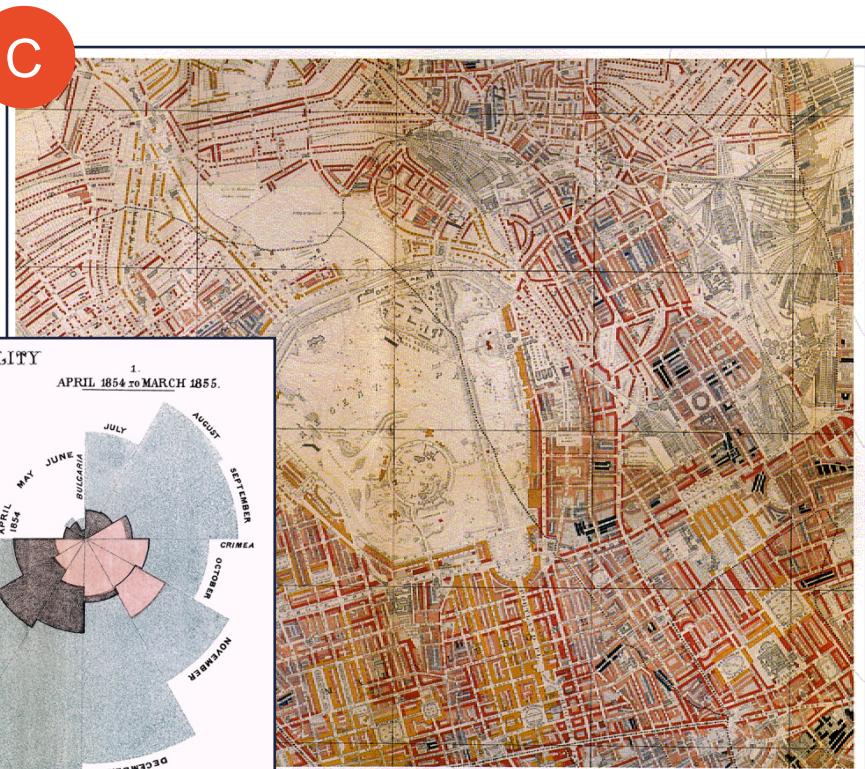
**A** 1855: John Snow; use of a dot map to display epidemiological data leads to discovery of the source of a cholera epidemic



**B** 1857: Florence Nightingale; polar area charts, known as "coxcombs," used in a campaign to improve sanitary conditions in the army

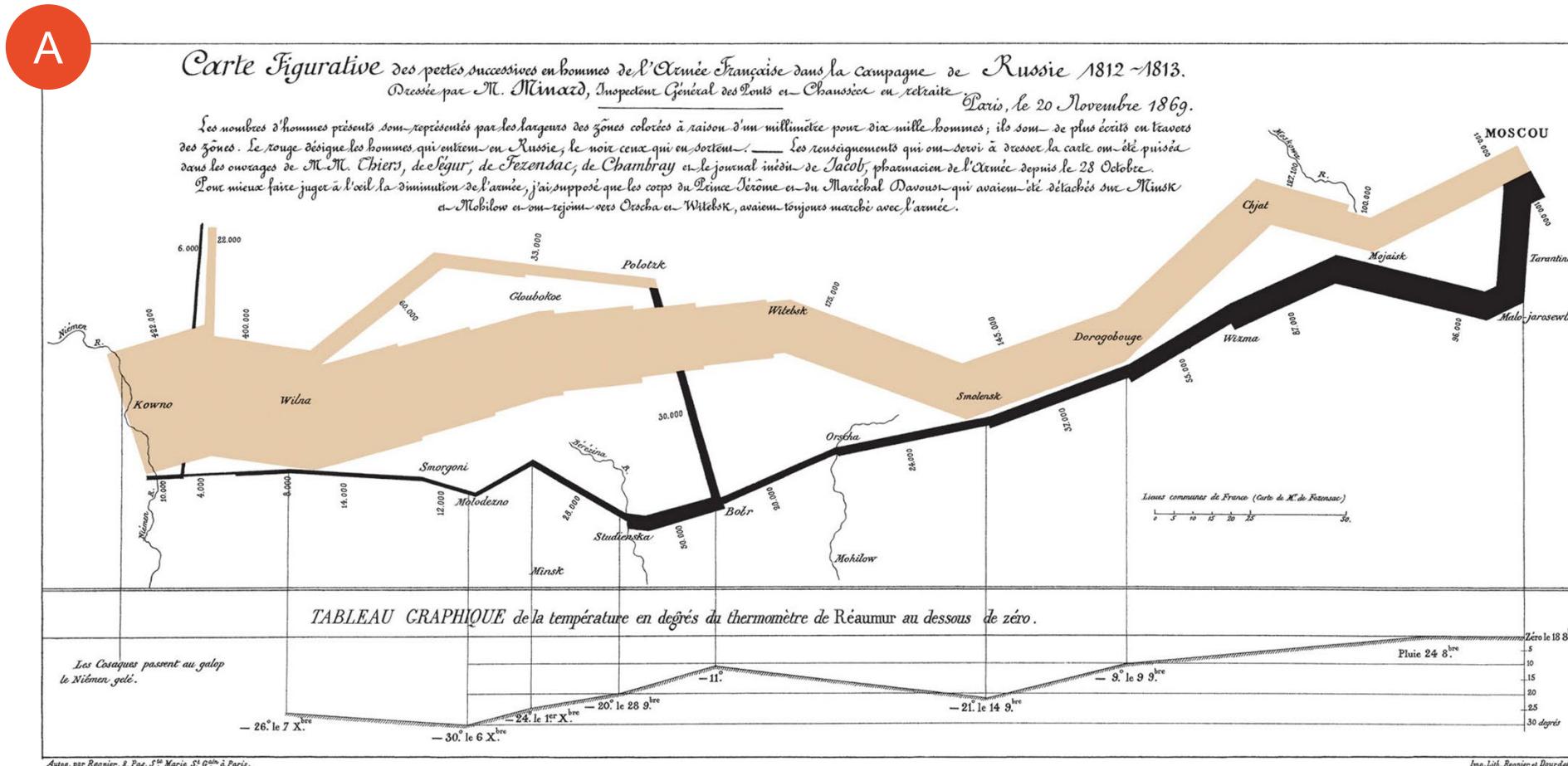


**C** 1889: Charles Booth; street maps of London showing poverty and wealth by color, transforming existing methods of social survey toward the end of the 19th century

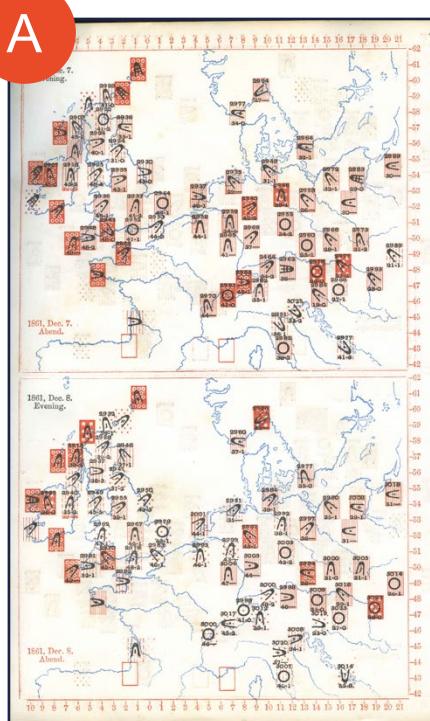


**A** 1869: Charles Joseph Minard; flow map of Napoleon's March on Moscow frequently called the best graphic ever produced. The illustration depicts Napoleon's army departing the Polish-Russian border bound for Moscow. A thick band illustrates the size of his army at specific geographic points during its advance and retreat. It displays six types of data in two dimensions: the number of Napoleon's troops, the distance traveled, temperature, latitude and longitude, direction of travel, and location relative to specific dates.

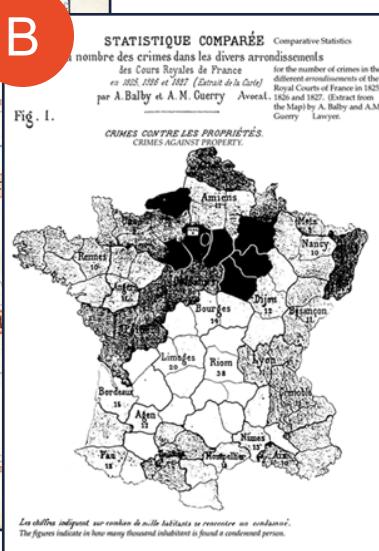
## Epoch V: The Golden Age of Data Analytics (1850–1899)



**A** 1861: Francis Galton; the modern weather map, a chart showing area of similar air pressure and barometric changes by means of glyphs displayed on a map

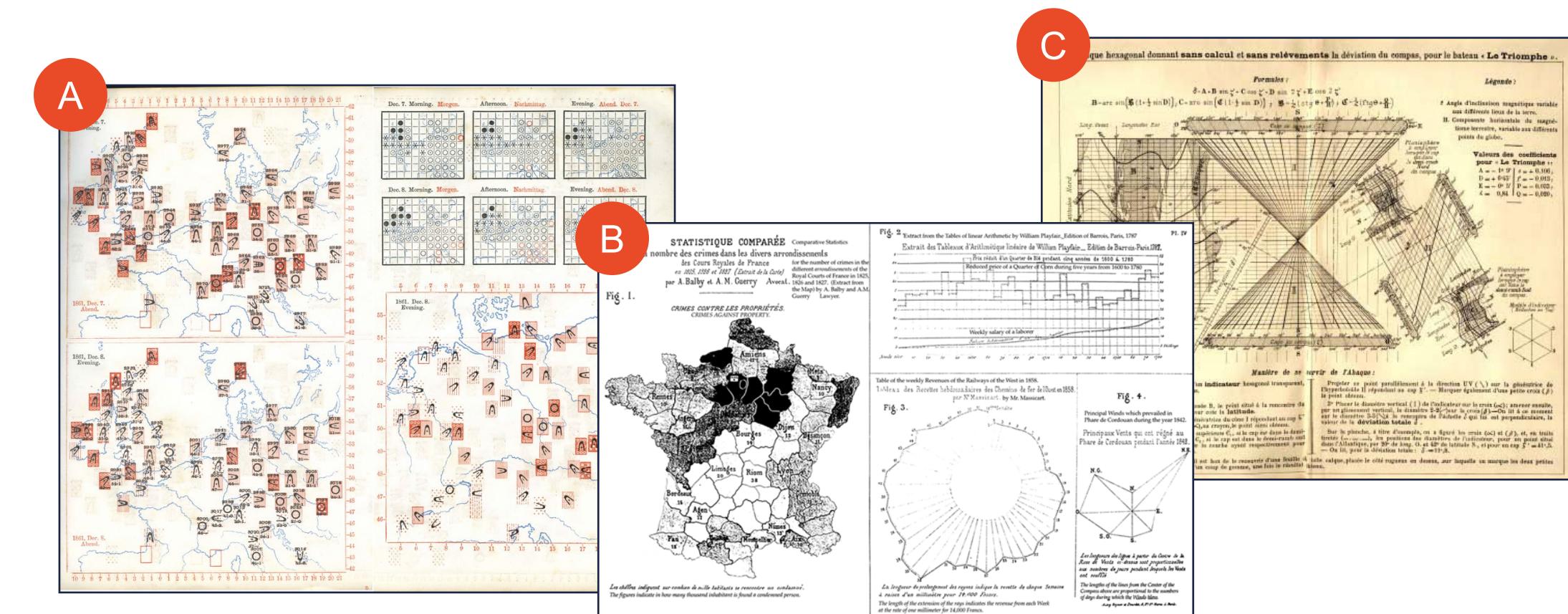


**B** 1885: Émile Levasseur; review of all available statistical graphics presented to the Statistical Society of London, classified as figures, maps, and solids



**C** 1885: Charles Lallemand; combination of many variables into multifunction nomograms using 3D, juxtaposition of maps, parallel coordinate and grids

## Epoch V: The Golden Age of Data Analytics (1850–1899)

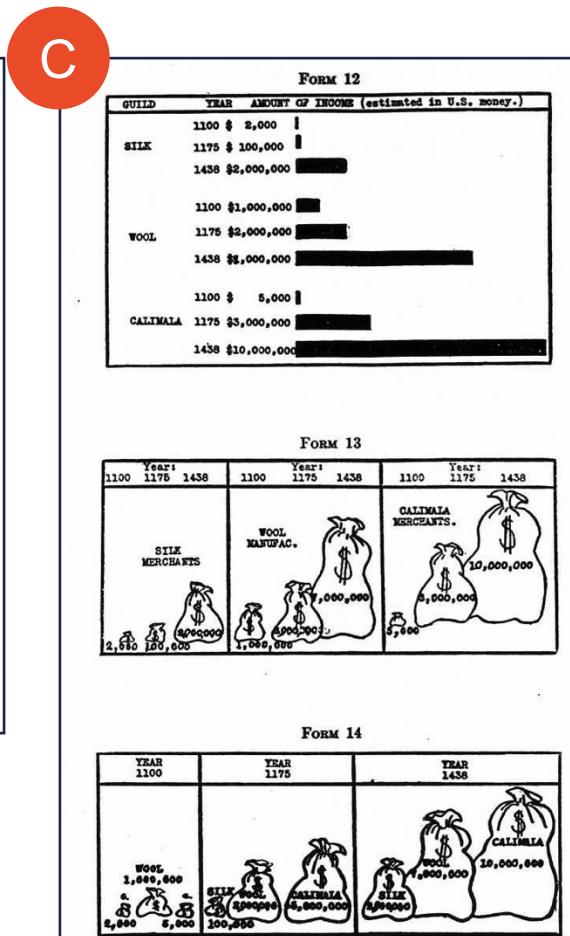
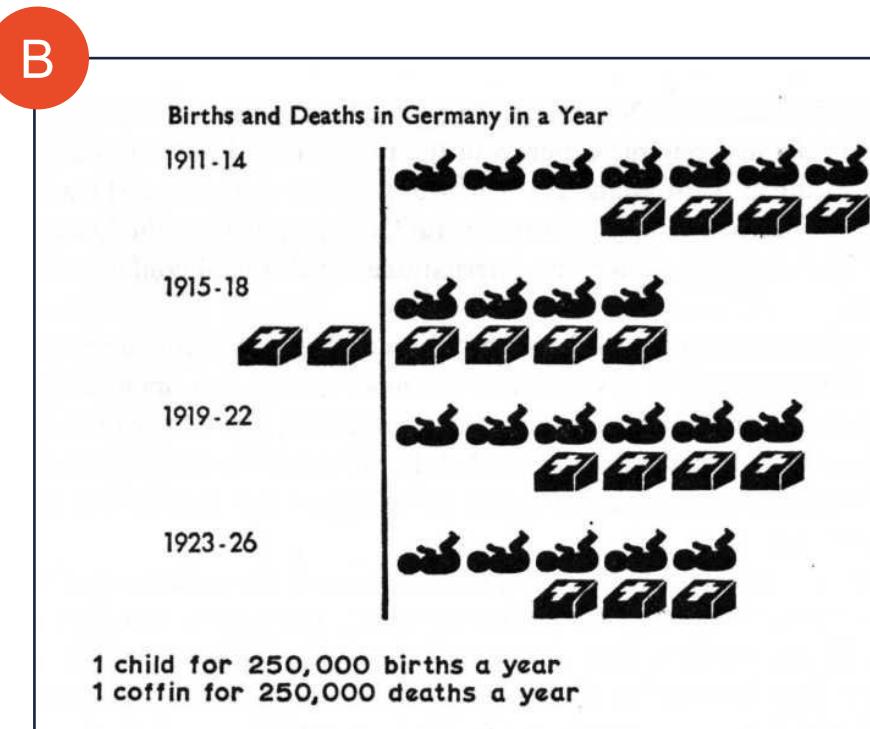
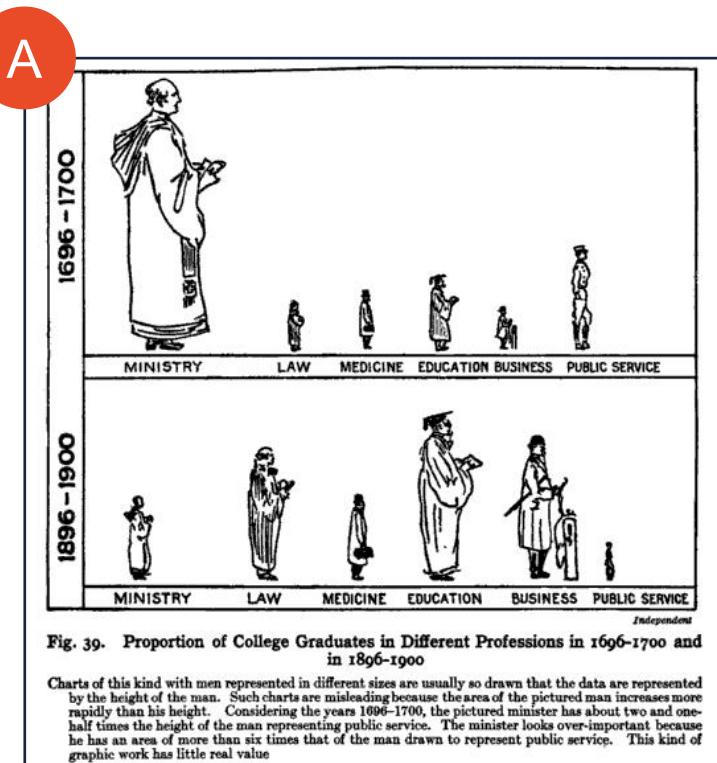


**A** 1914: Willard Cope Brinton; pictograms representing a series of numbers by icons (combining concepts of the bar graph and pictogram of varying sizes)

**B** 1924: Otto Neurath; introduction of the ISOTYPE standard (International System of Typographic Picture Education)

**C** 1932: N.J. Washburn; spate of articles on experimental tests of statistical graphical forms

## Epoch VI: The Modern Dark Age of Data Analytics (1900–1949)



**A** 1933: Henry Beck, an engineering draughtsman, designed a map like an electrical circuit board, using only vertical, horizontal and 45-degree angled lines. He located stations according to available space. The resulting map was geographically inaccurate but easier to use to determine how to get from point A to point B. Beck's idea was soon copied by most subway (and bus) companies around the world.

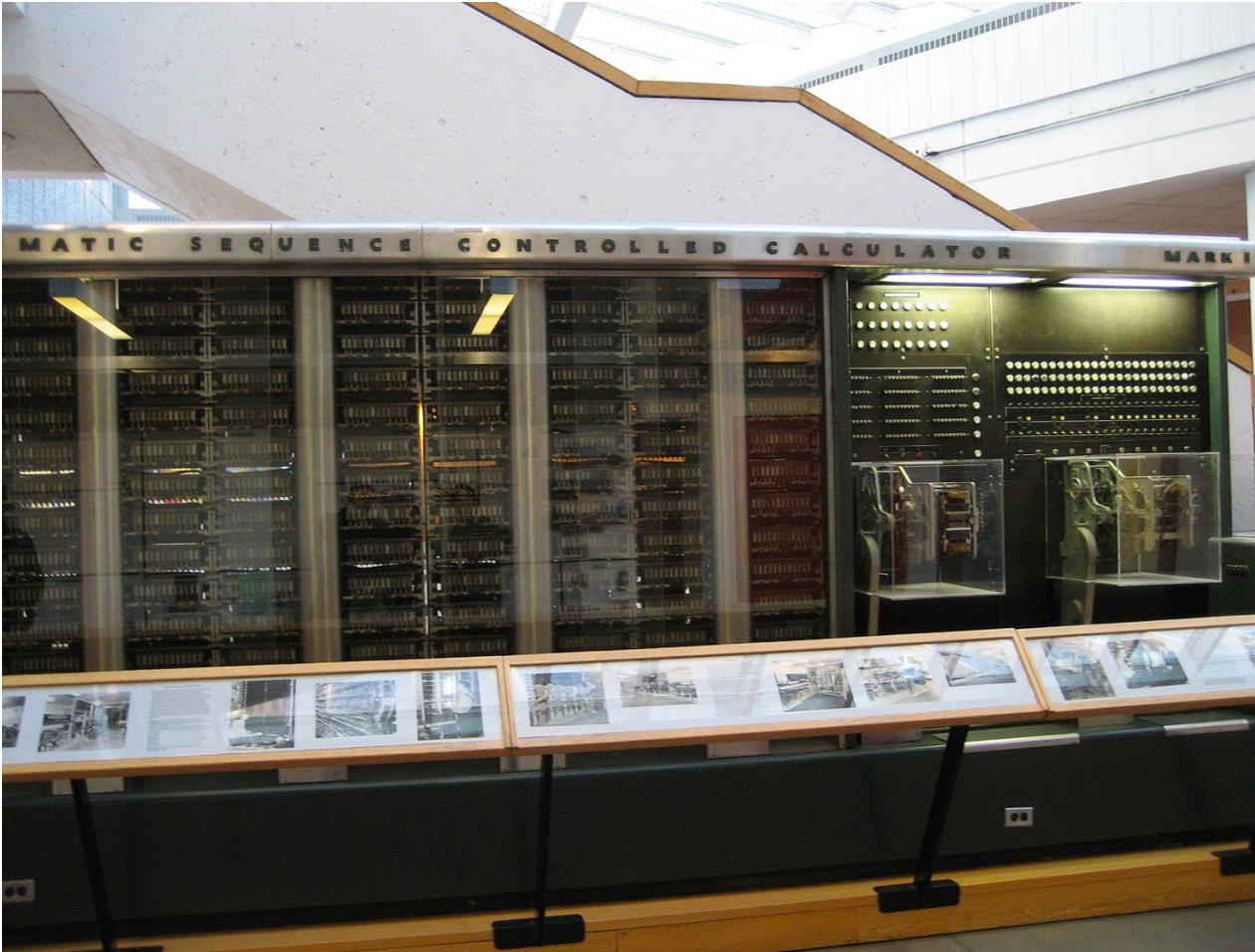
## Epoch VI: The Modern Dark Age of Data Analytics (1900–1949)



**A** **1944:** Harvard's Mark I, the first digital computer put in service. Known as the IBM Automatic Sequence Controlled Calculator' (ASCC), it was 50 feet long and weighed about 5 tons. The machine could perform all four arithmetic operations; moreover, it had special built-in programs to handle logarithms and trigonometric functions. It was slow, requiring 3 to 5 seconds for a multiplication, but it was fully automatic and could complete long computations without human intervention.

## Epoch VI: The Modern Dark Age of Data Analytics (1900–1949)

A

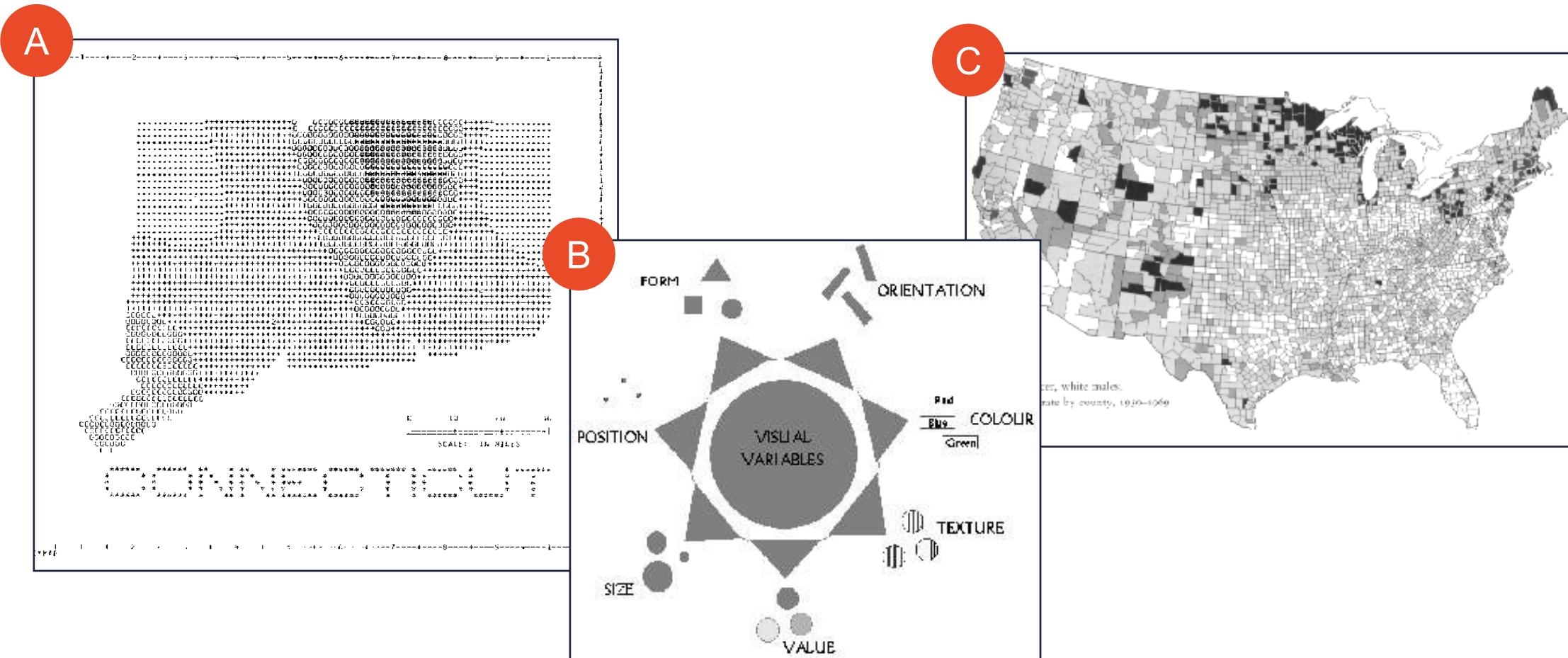


**A** 1960: Howard Taylor Fisher; initial development of geographic information systems, combining spatially-referenced data, spatial models, and map-based visualization

**B** 1967: Jacques Bertin; comprehensive theory of graphical symbols and modes of graphics representation

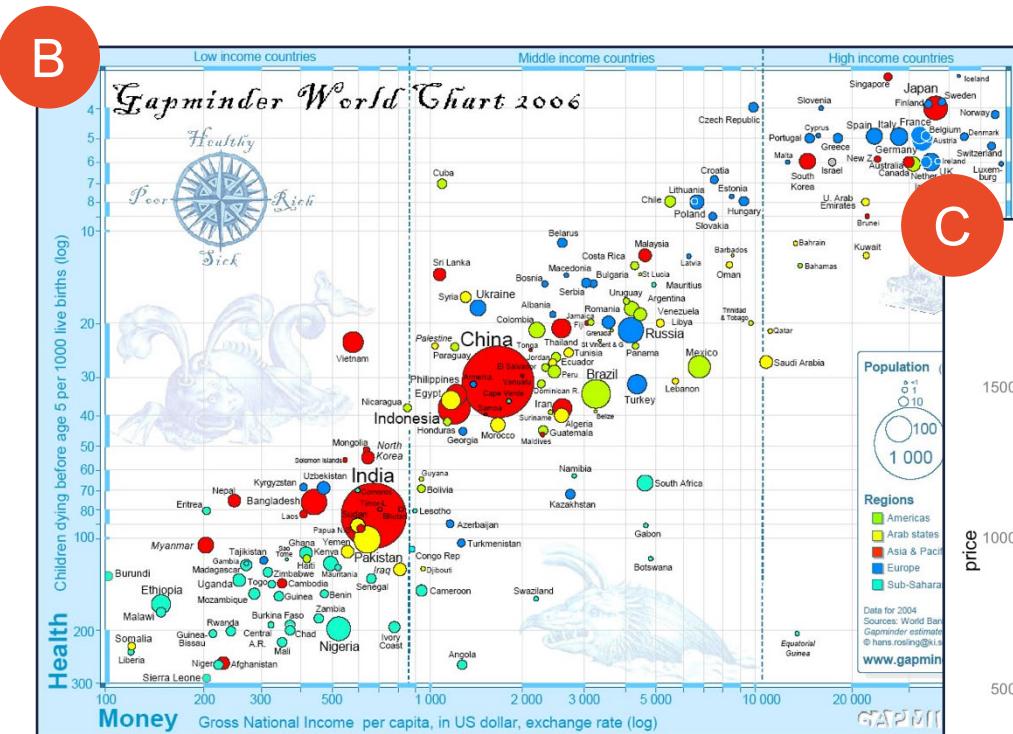
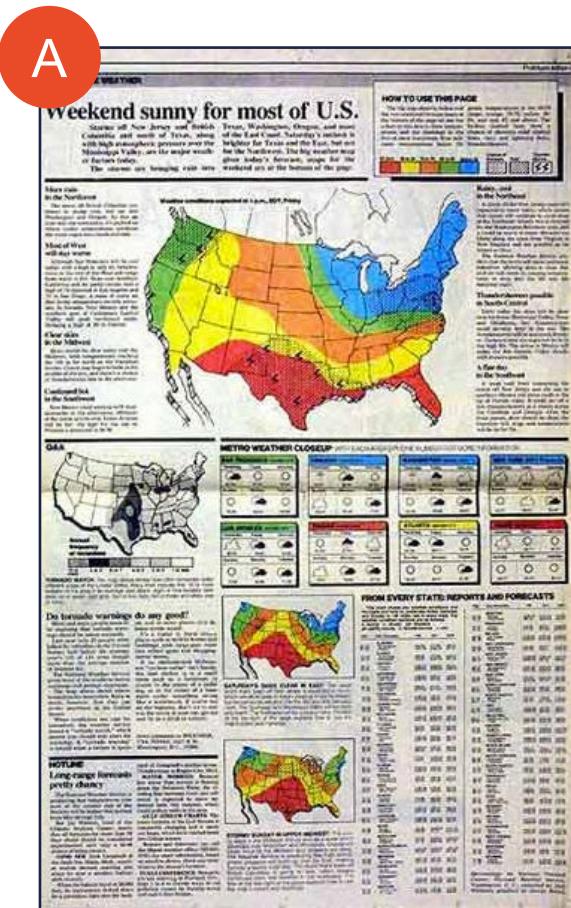
**C** 1974: US Census Bureau; color-coded bivariate matrix to represent two variables in a single map

## Epoch VII: The Rebirth of Data Analytics (1950–1974)

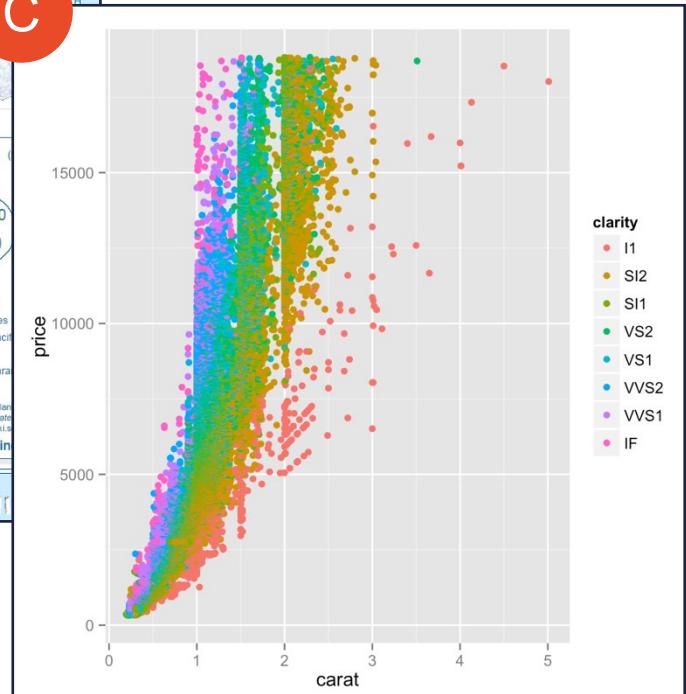


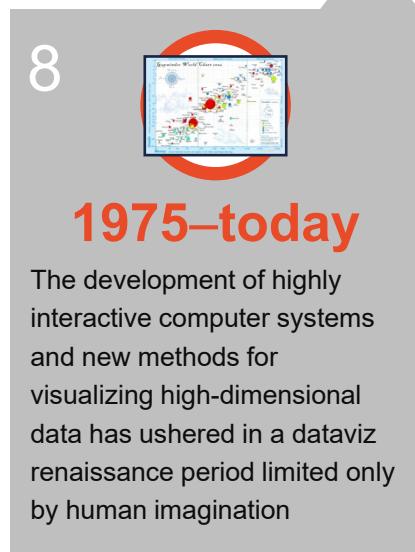
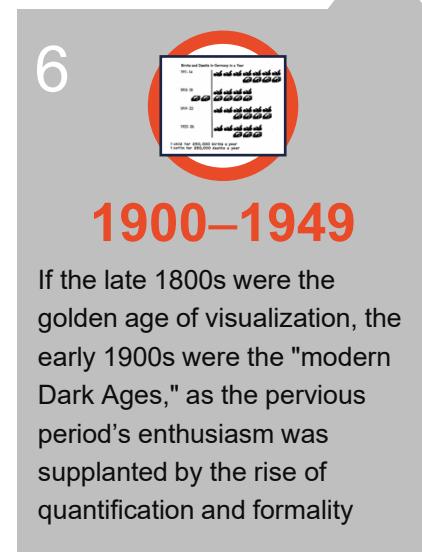
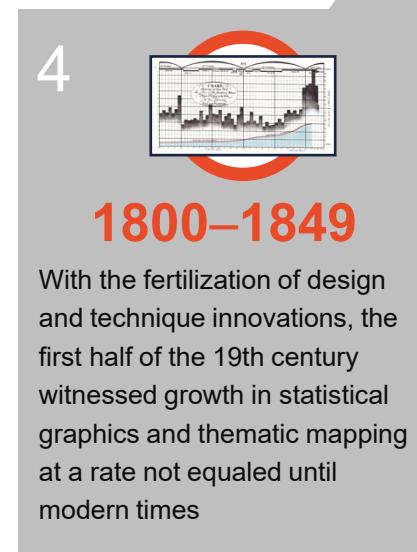
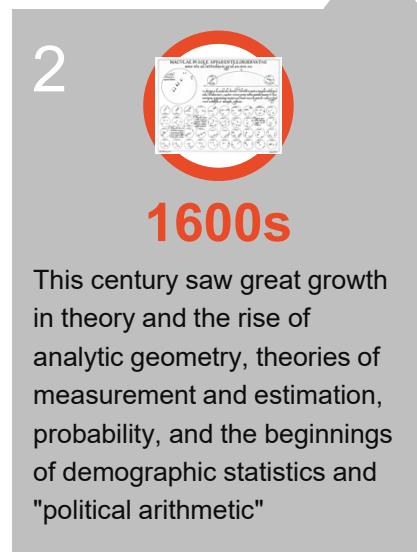
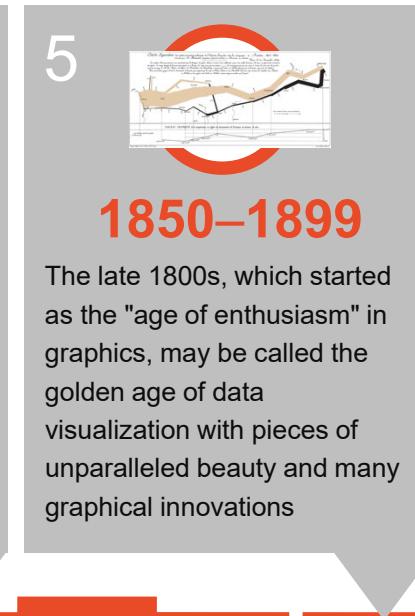
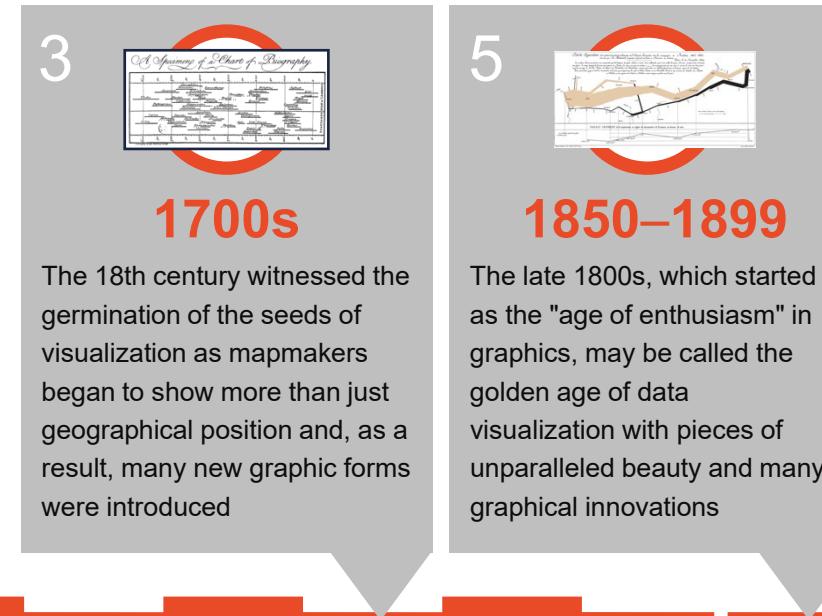
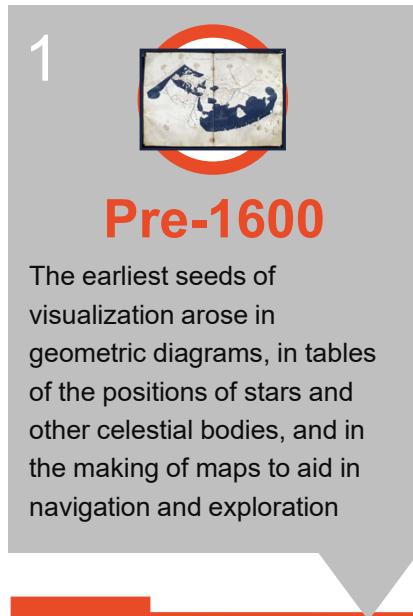
**A** 1982: The *USA Today* color weather map begins an era of color information graphics in newspapers

## Epoch VIII: High-Definition Data Analytics (1975–today)



**C** 2006: Hadley Wickham; an influential, open-source implementation of the Grammar of Graphics from Wilkinson in R (ggplot2)





# Studying the History of Dataviz Can Inspire and Inform

Eight distinct epochs trace the long, rich history of data visualization

Dataviz has evolved from hand-painted maps into sophisticated computer-generated visuals

Over time, the practice of data visualization has confronted issues of data scarcity, human politics, data misuse, and the scourge of clip art

Dataviz has emerged as a discipline that offers the ultimate blend of art and science

# References

Friendly, M., & Denis, D. J. (2001). *Milestones in the history of thematic cartography, statistical graphics, and data visualization* [Web document].



Module 1 Lesson 2

# Understanding Today's Dataviz Tools



# Understanding Today's Dataviz Tools

Dataviz tools represent an evolution in analytics that can best be described along two important dimensions

# Dataviz Tools Can Be Defined by Two Dimensions

I

**Ease of use**

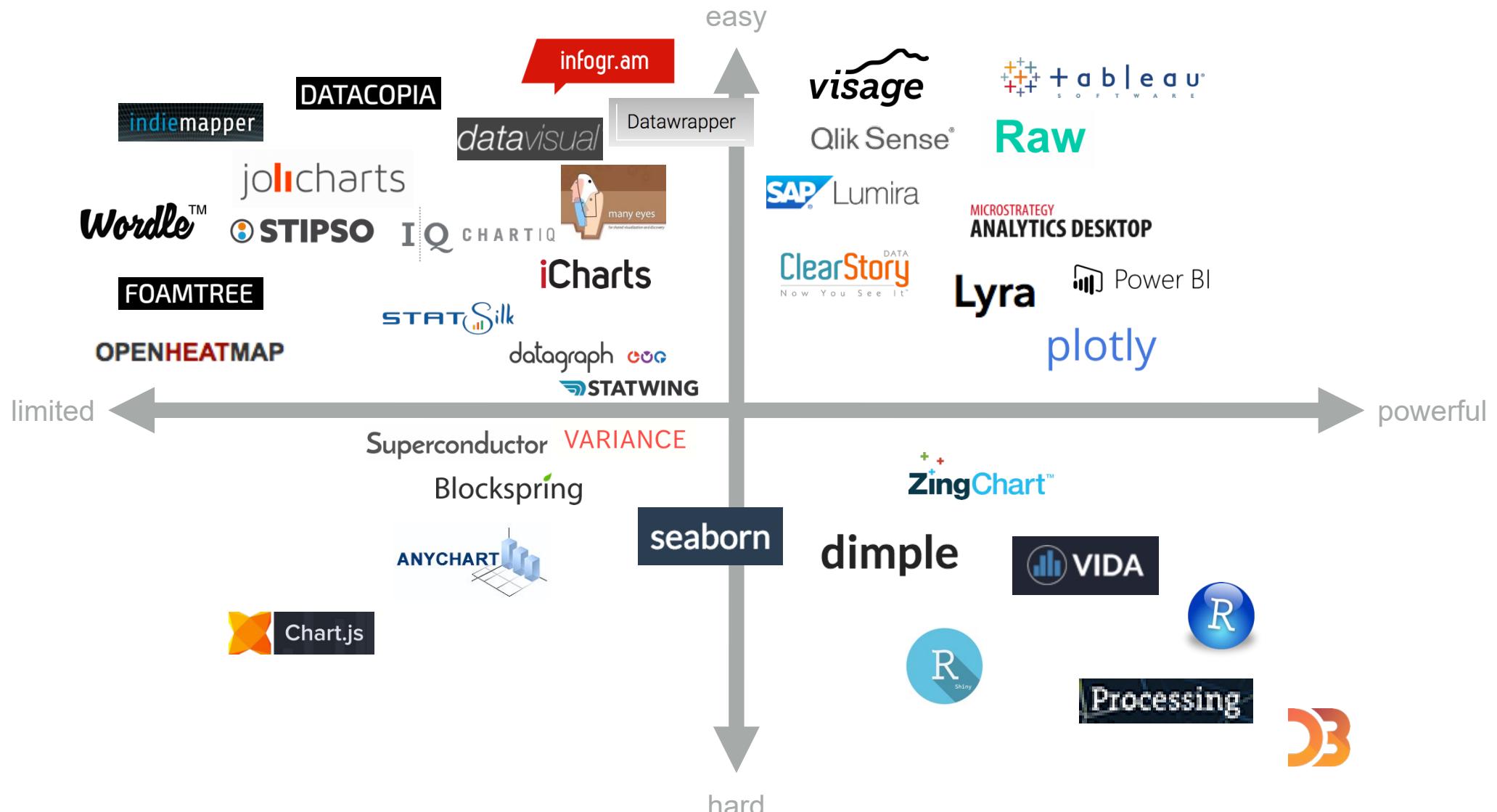
How quickly can one reasonably expect to learn the tool? How intuitive is it? What special skills does it require?

**Capability**

What is the breadth of the tool's functionality? How flexible is it? How much data can the tool effectively manage?

# THE DATA VISUALIZATION TOOL MARKET

An analysis of ease of use and capability



easy **I**

infogr.am

Datawrapper

DATACOPIA

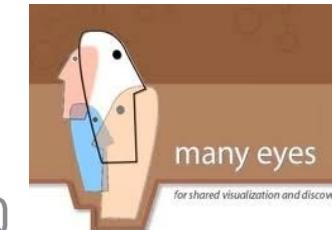
indiemapper

jolicharts

Wordle™

STIPSO

IQ CHARTIQ



iCharts

STATSilk

datagraph

STATWING

Superconductor VARIANCE

limited



Source: Adapted from David McCandless



easy **I**

infogr.am

Datawrapper

DATACOPIA

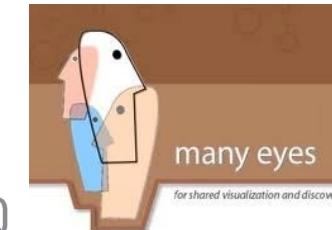
indiemapper

Wordle™

jolicharts

STIPSO

IQ CHARTIQ



iCharts

STATSilk

datagraph

STATWING

Superconductor VARIANCE

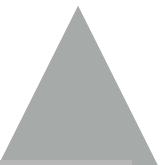
limited



Source: Adapted from David McCandless

I

easy



visage

wrapper

Qlik Sense®

tableau®  
S O F T W A R E

Raw

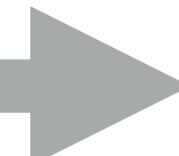
SAP Lumira

MICROSTRATEGY  
ANALYTICS DESKTOP

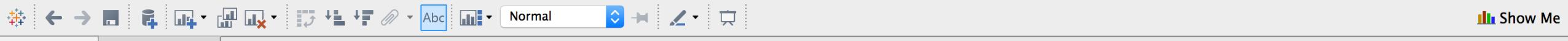
ClearStory DATA  
Now You See It™

Lyra Power BI

plotly



powerful



Data Analytics

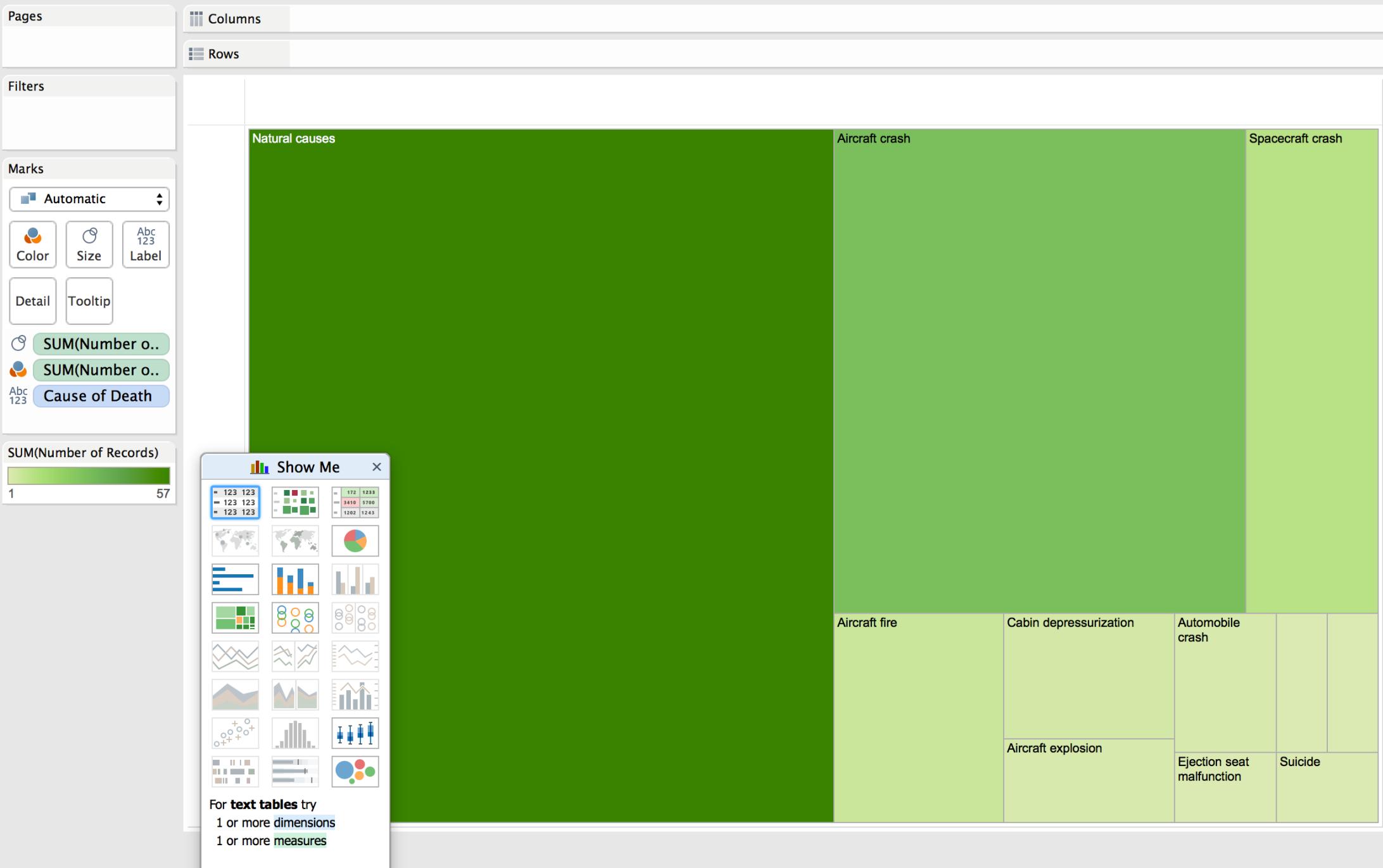
Astronautic-Deaths-19...

Dimensions

- Birth
- Cause of Death
- Cause of Death Detail
- Column
- Death
- Name
- Trained as
- Measure Names

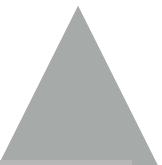
Measures

- # Age
- =# Number of Records
- # Measure Values



I

easy



visage

wrapper

Qlik Sense®

tableau®  
S O F T W A R E

Raw

SAP Lumira

MICROSTRATEGY  
ANALYTICS DESKTOP

ClearStory DATA  
Now You See It™

Lyra Power BI

plotly

ING

Source: Adapted from David McCandless

powerful

ING

CE

seaborn

I

powerful



dimple



hard

Source: Adapted from David McCandless

TOP

ING

CE

seaborn

I

powerful



dimple

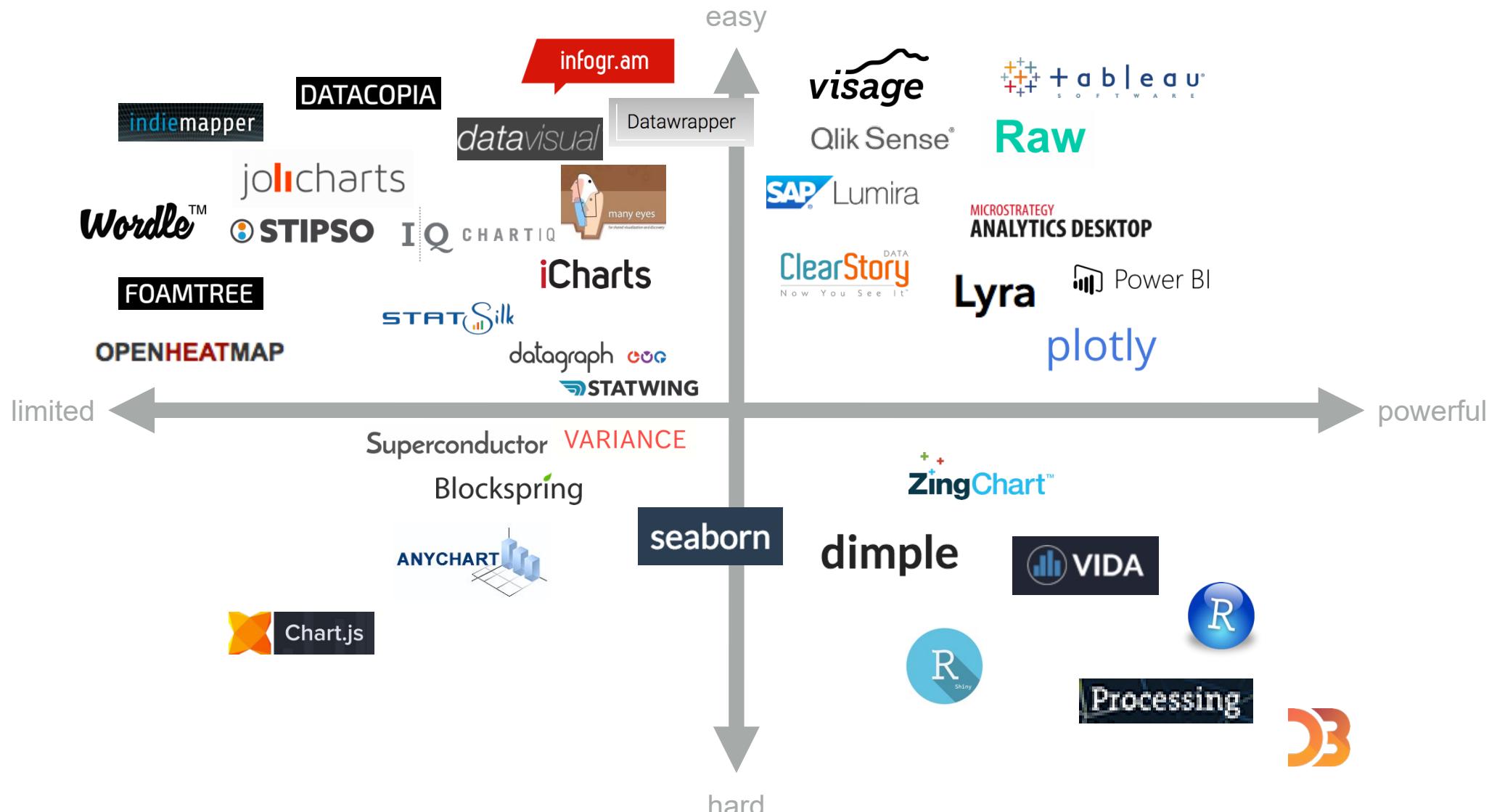


hard

Source: Adapted from David McCandless

# THE DATA VISUALIZATION TOOL MARKET

An analysis of ease of use and capability



# Keep a Number of Data Visualization Tools in Your “Toolbox”

Evaluate dataviz tools in their ability to offer efficiency: the ease with which you can access its capabilities

Immediately discard hard-to-learn tools that do not offer exceptional capabilities

Anchor on an easy-to-use tool with broad capabilities

Utilize other tools when they are the most efficient tools for the job

Constantly search for new dataviz tools in this dynamic, ever-changing market

A photograph showing several students in a hallway. In the foreground, two female students are sitting on a bench, looking at a yellow notebook together and smiling. Behind them, other students are sitting and working on laptops. The background shows a staircase and more hallway details.

Module 1 Lesson 3

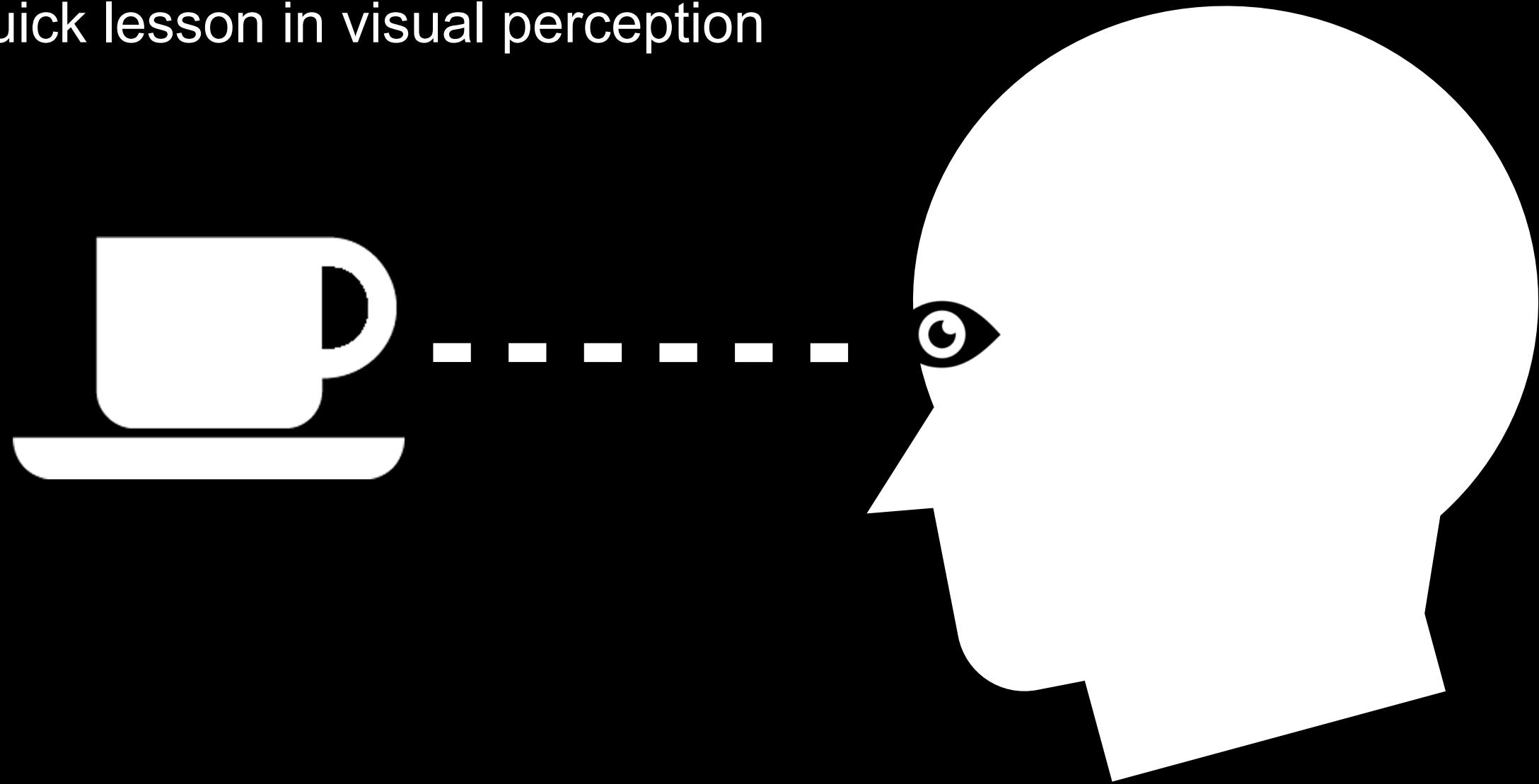
# Making Connections with Visuals

# Making Connections with Visuals

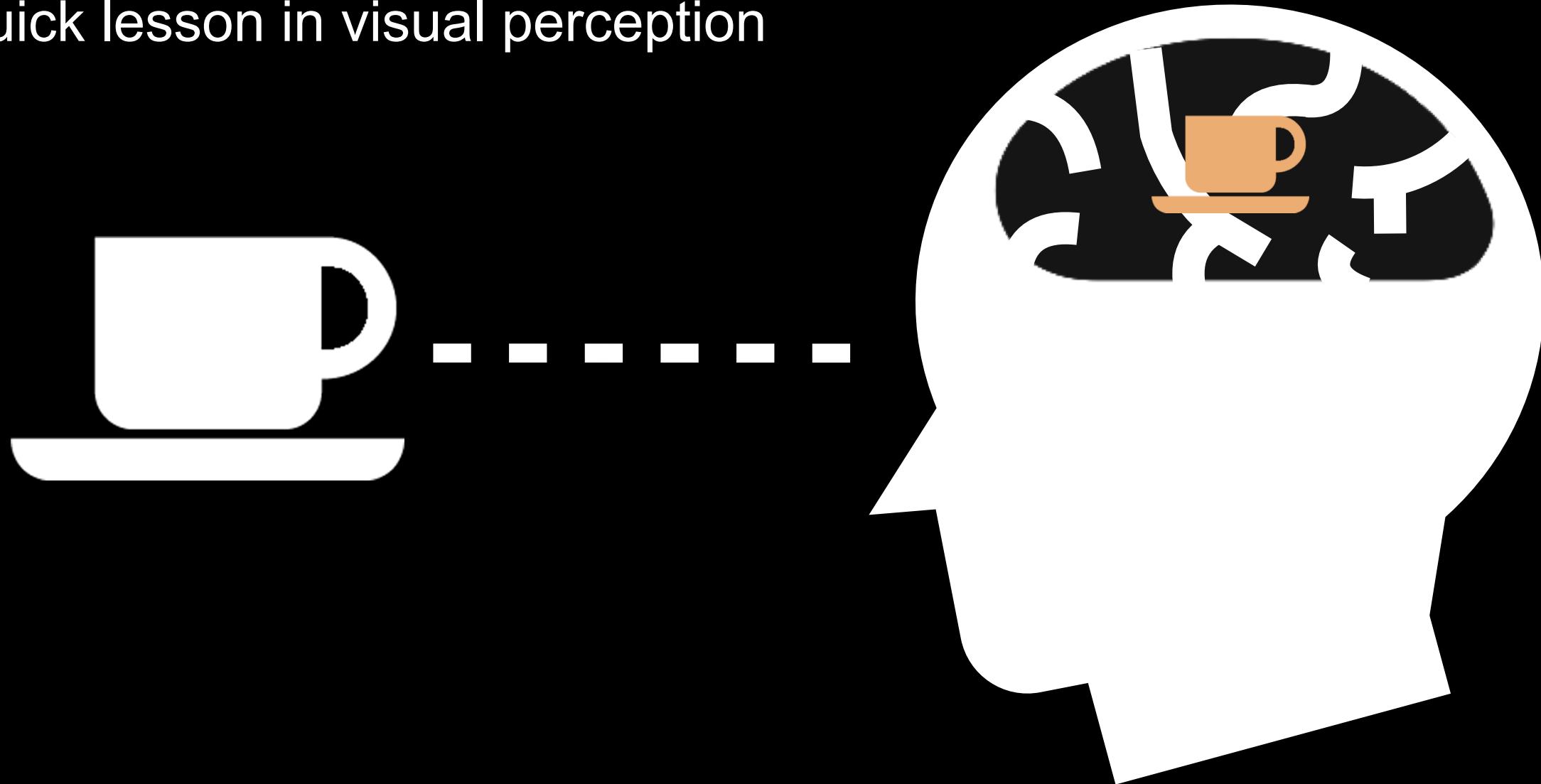
I

Visualization is rooted in brain science and is one of the most efficient ways to communicate ideas, insights, and concepts

# A quick lesson in visual perception



A quick lesson in visual perception

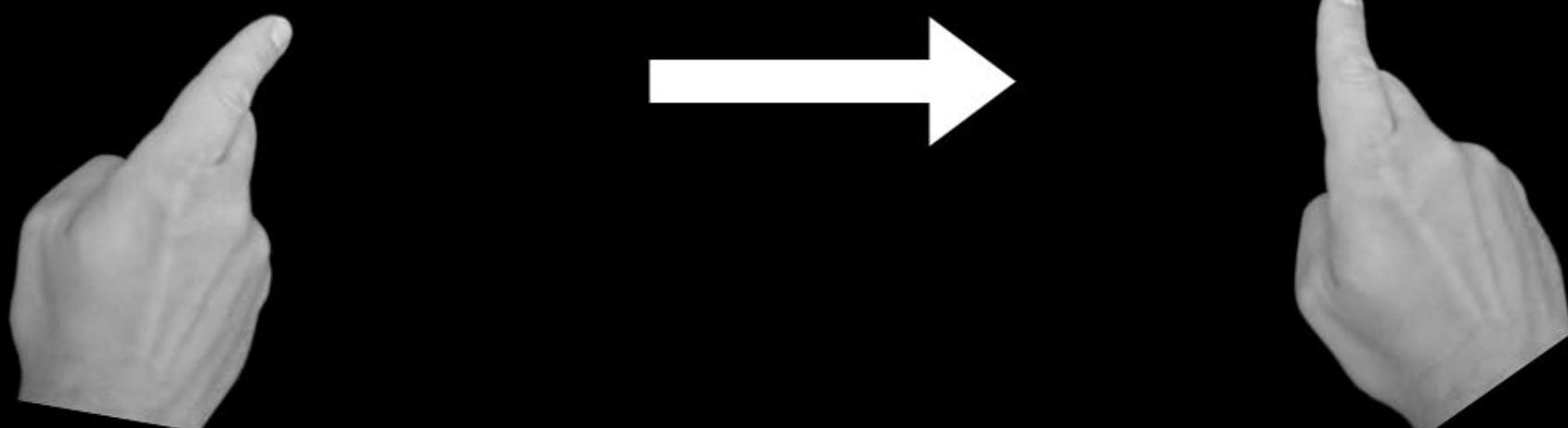




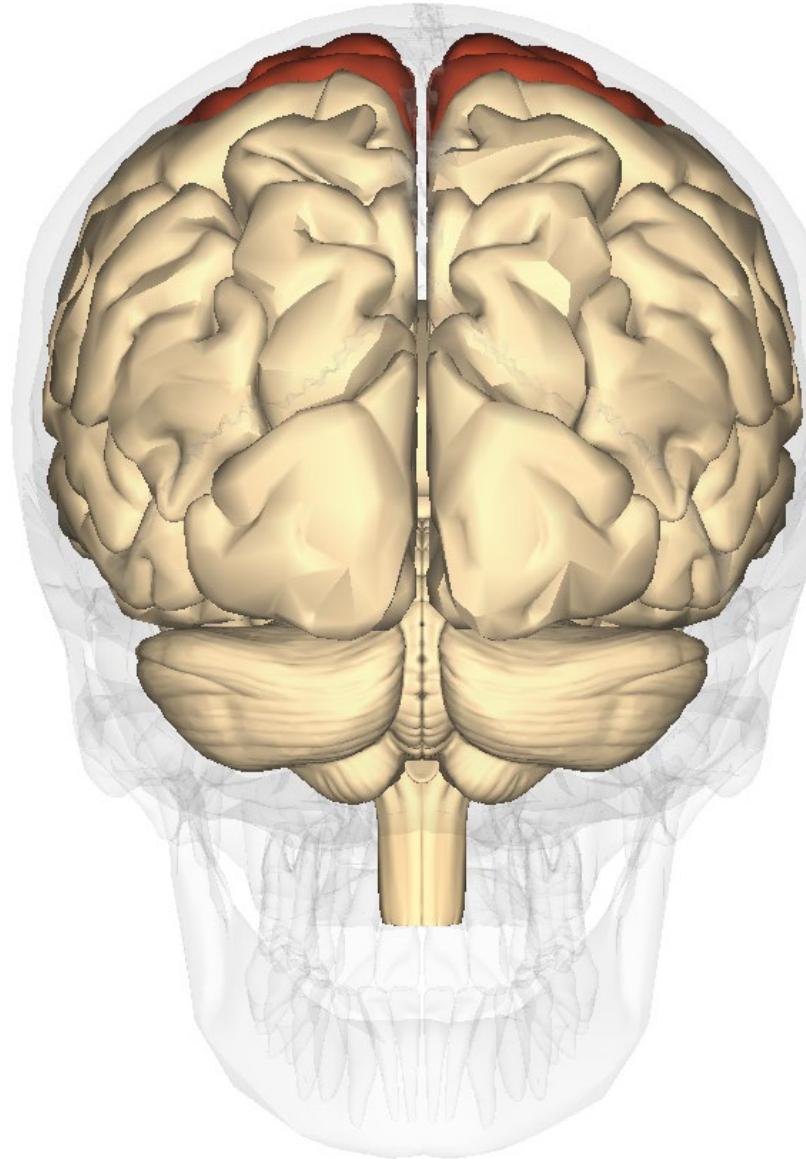
Source: Simons & Levin (1998)



?

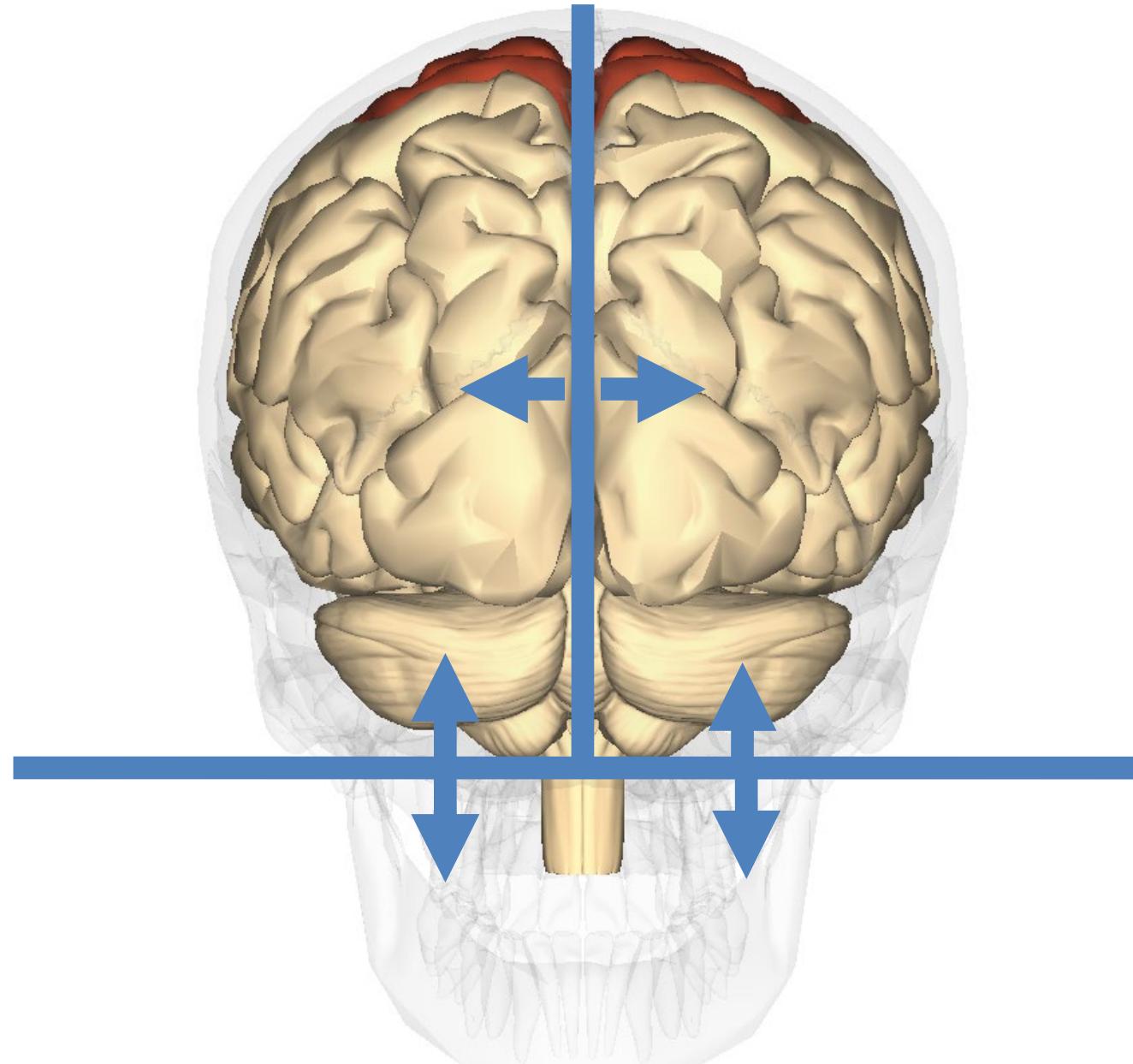


I



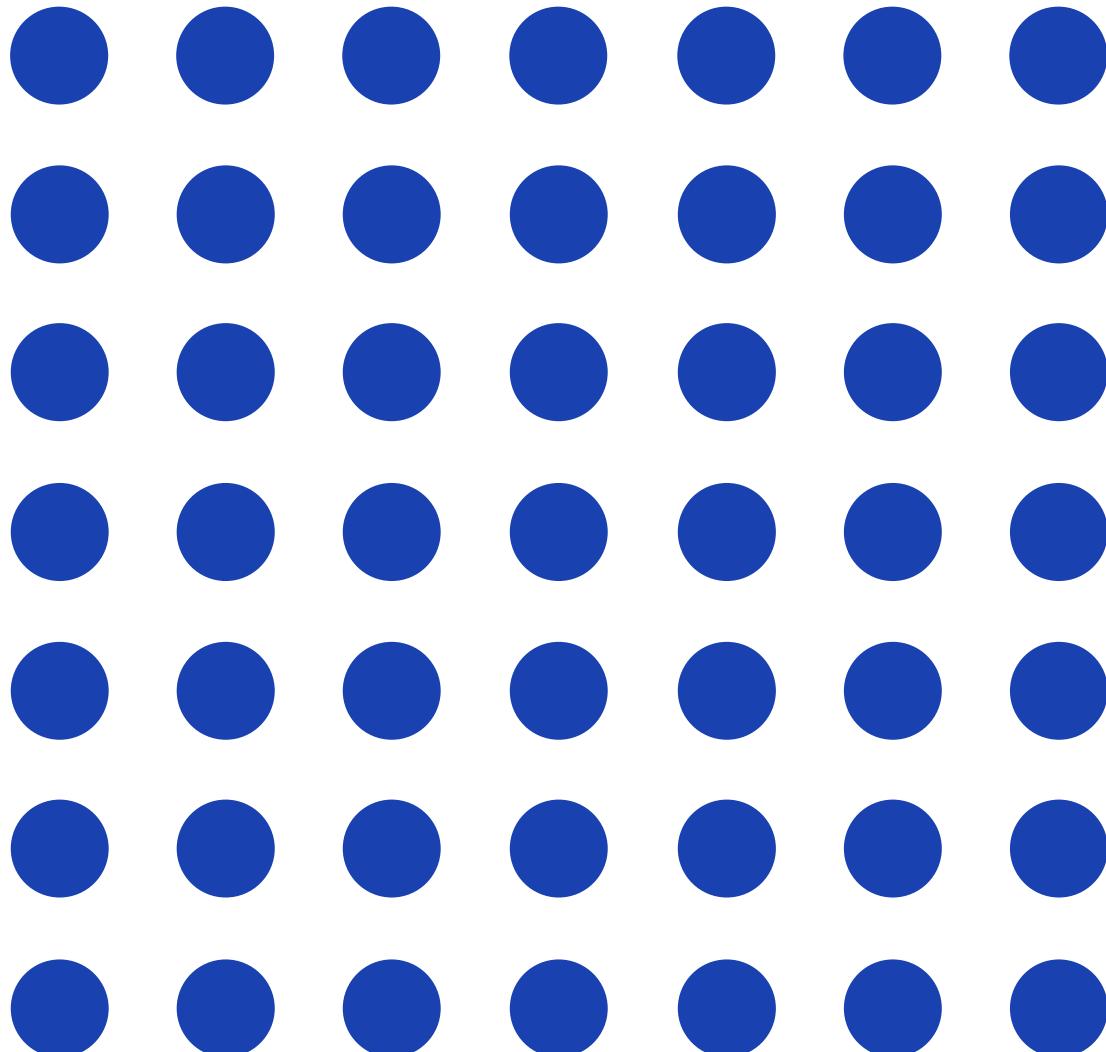
(Anatomography, 2012)

I

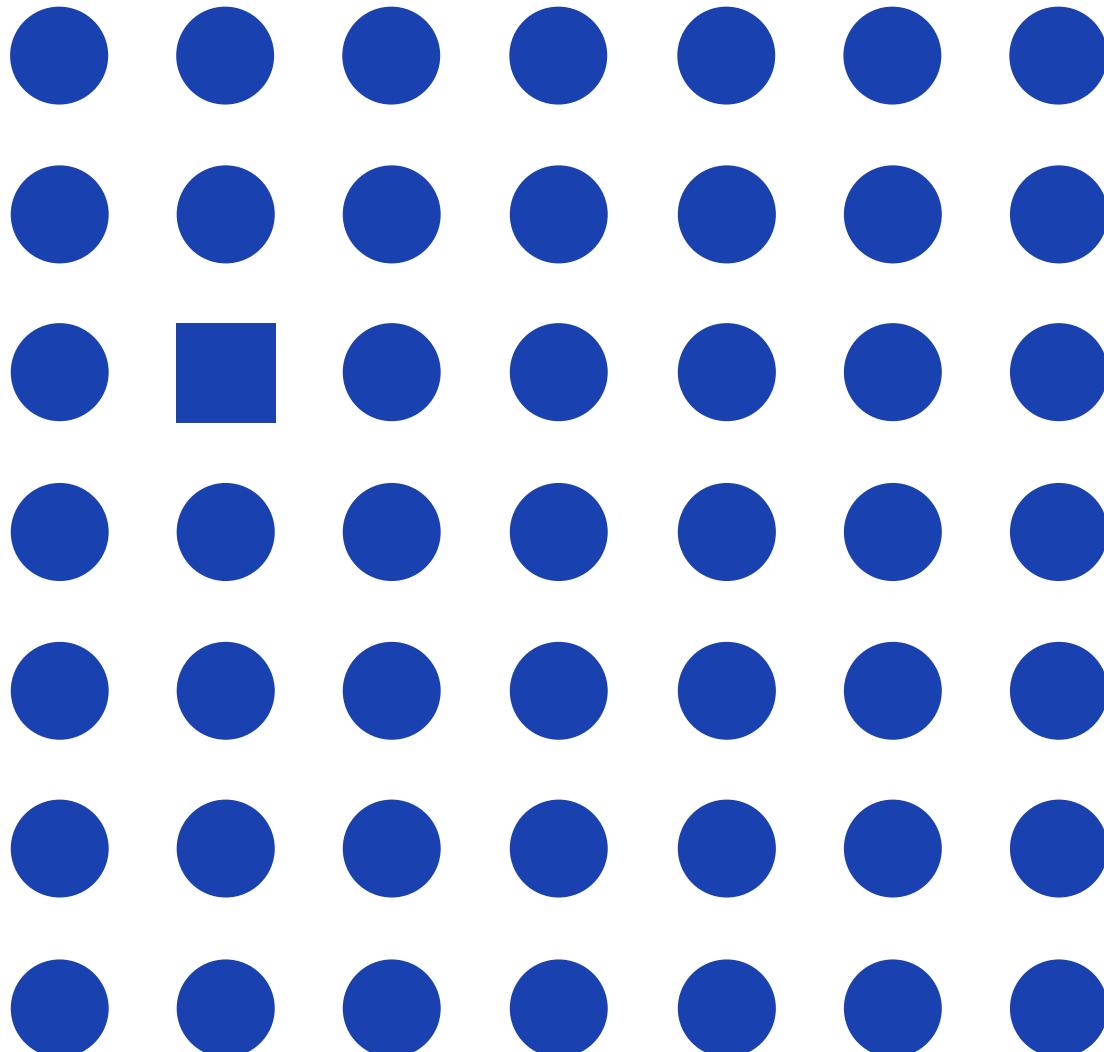


(Anatomography, 2012)

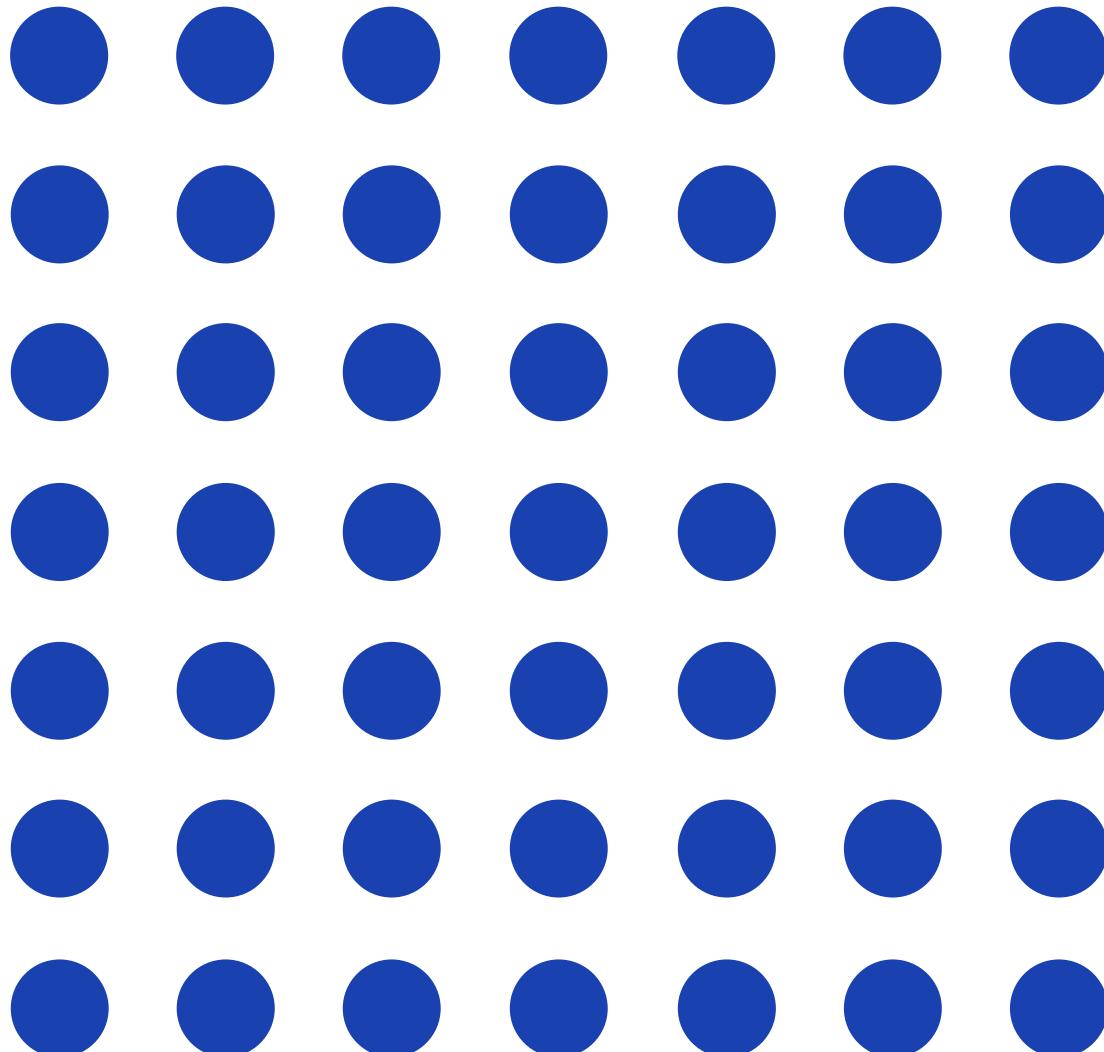
Our brains are  
programmed to  
detect differences



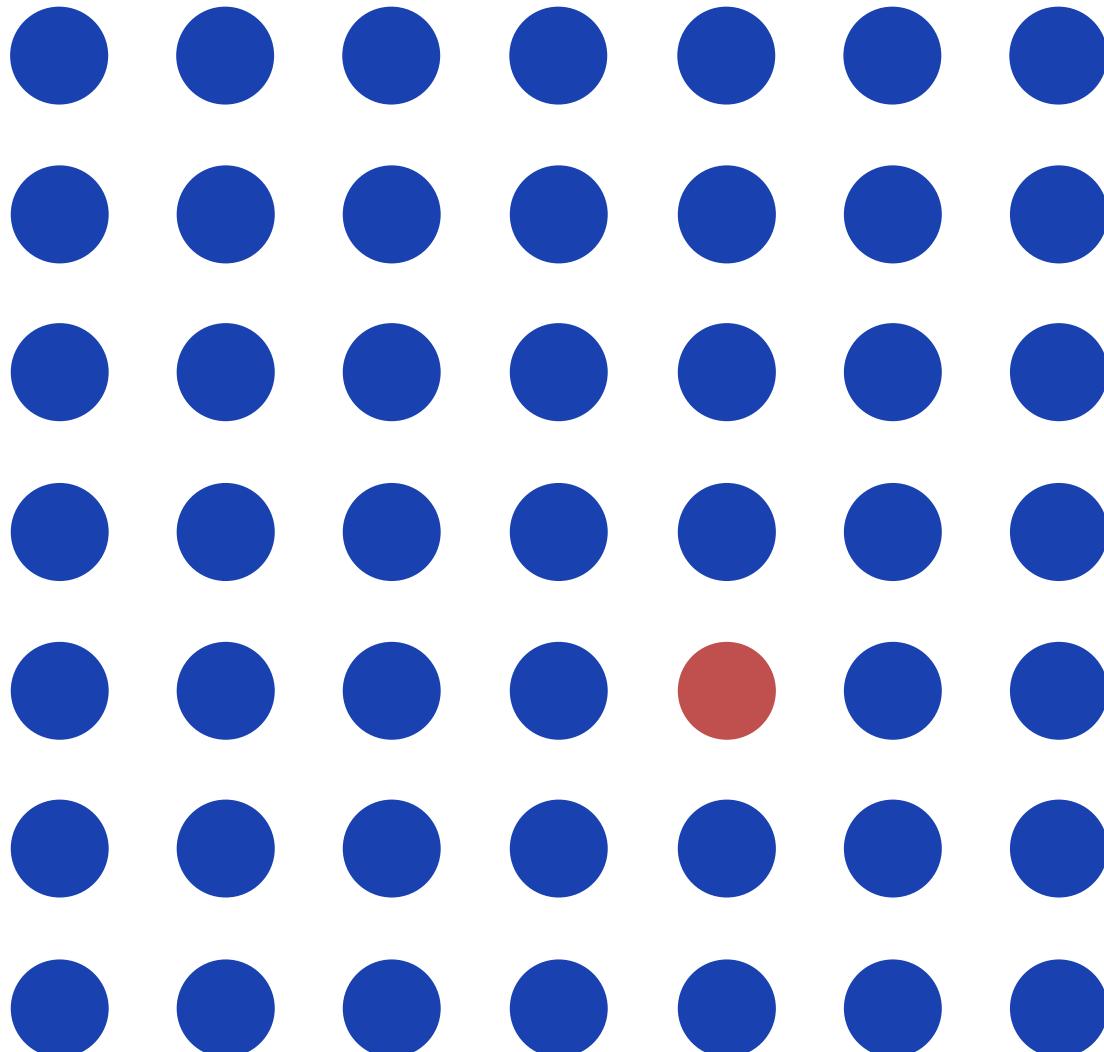
Our brains are  
programmed to  
detect differences



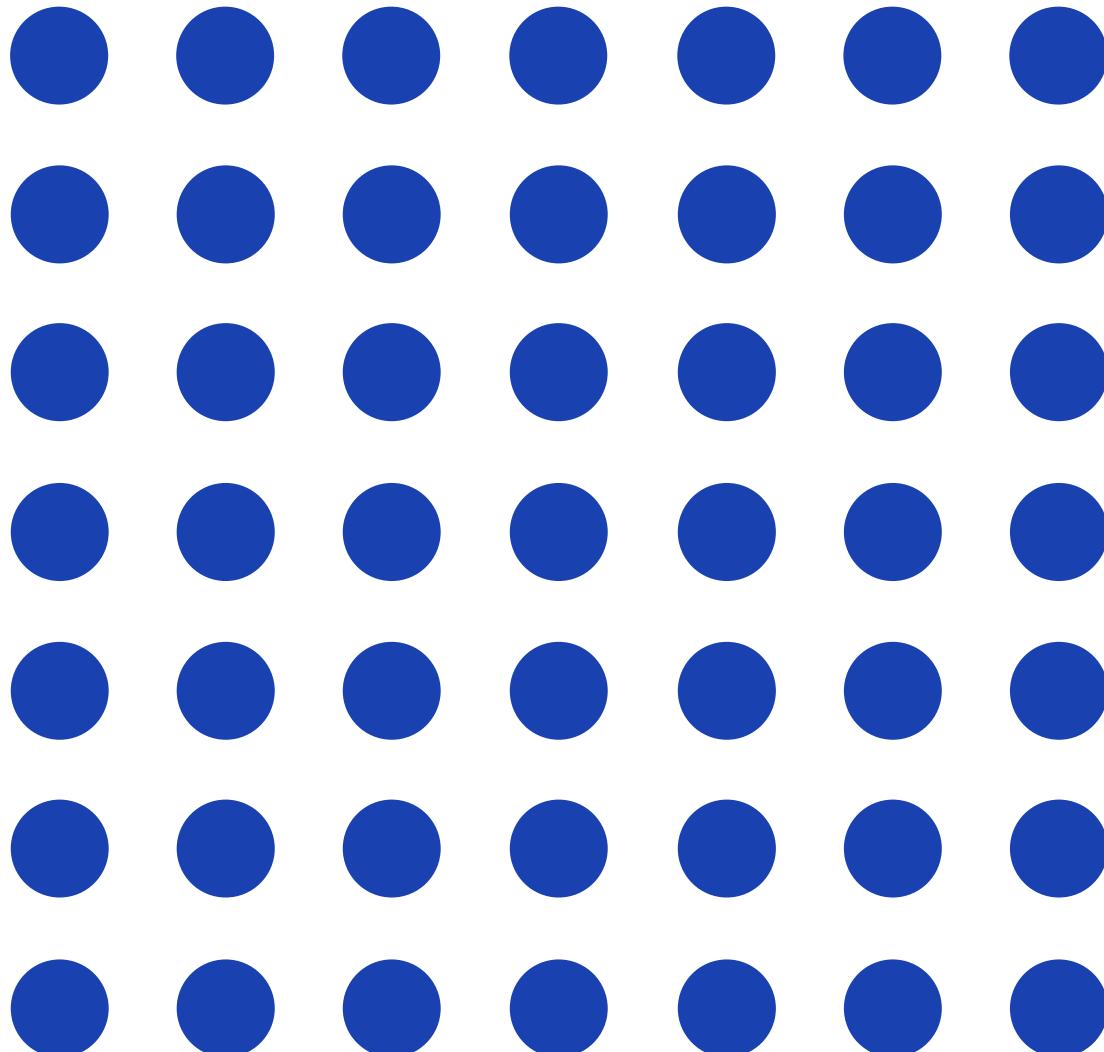
Our brains are  
programmed to  
detect differences



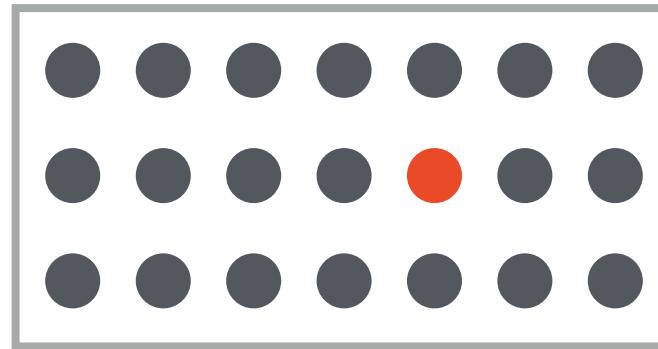
Our brains are  
programmed to  
detect differences



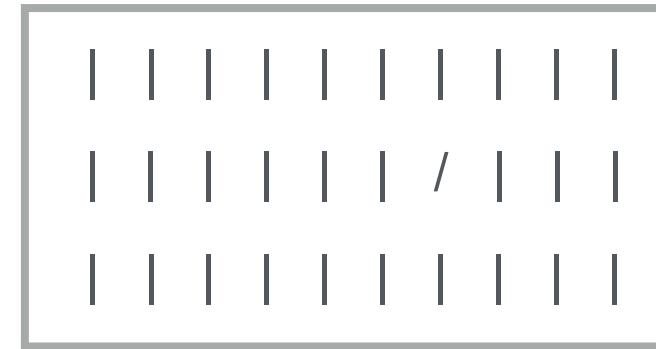
Our brains are  
programmed to  
detect differences



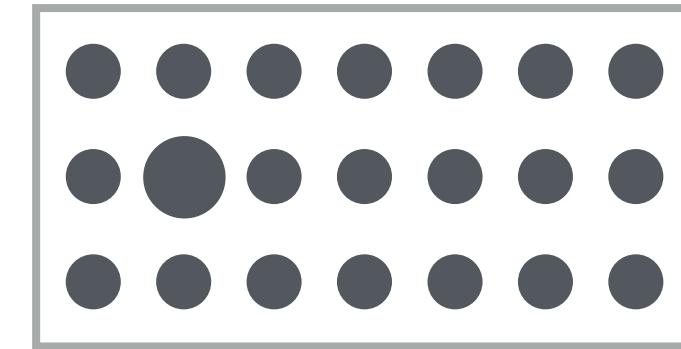
There are a host of pre-attentive attributes available



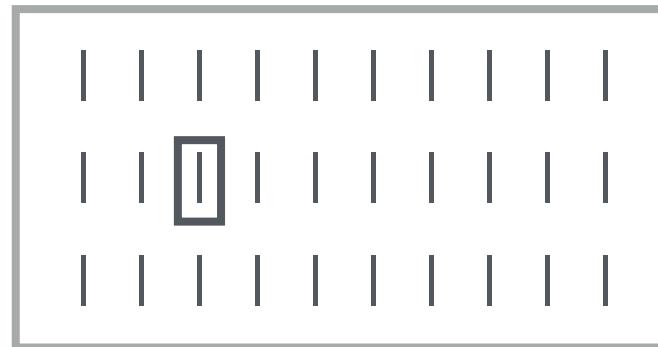
Hue



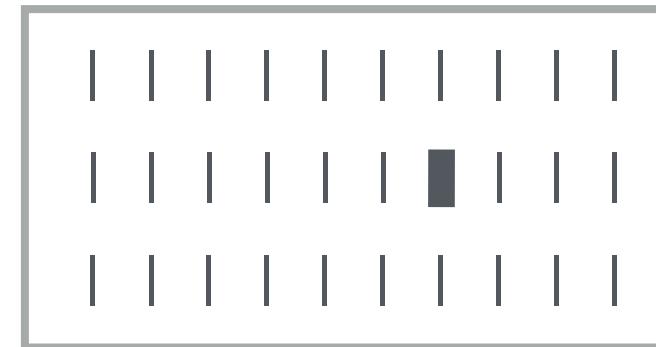
Orientation



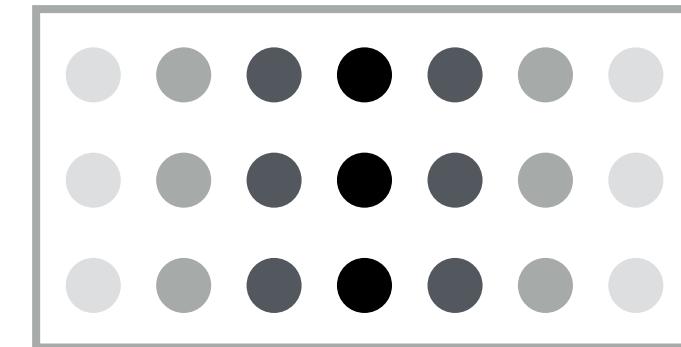
Size



Enclosure

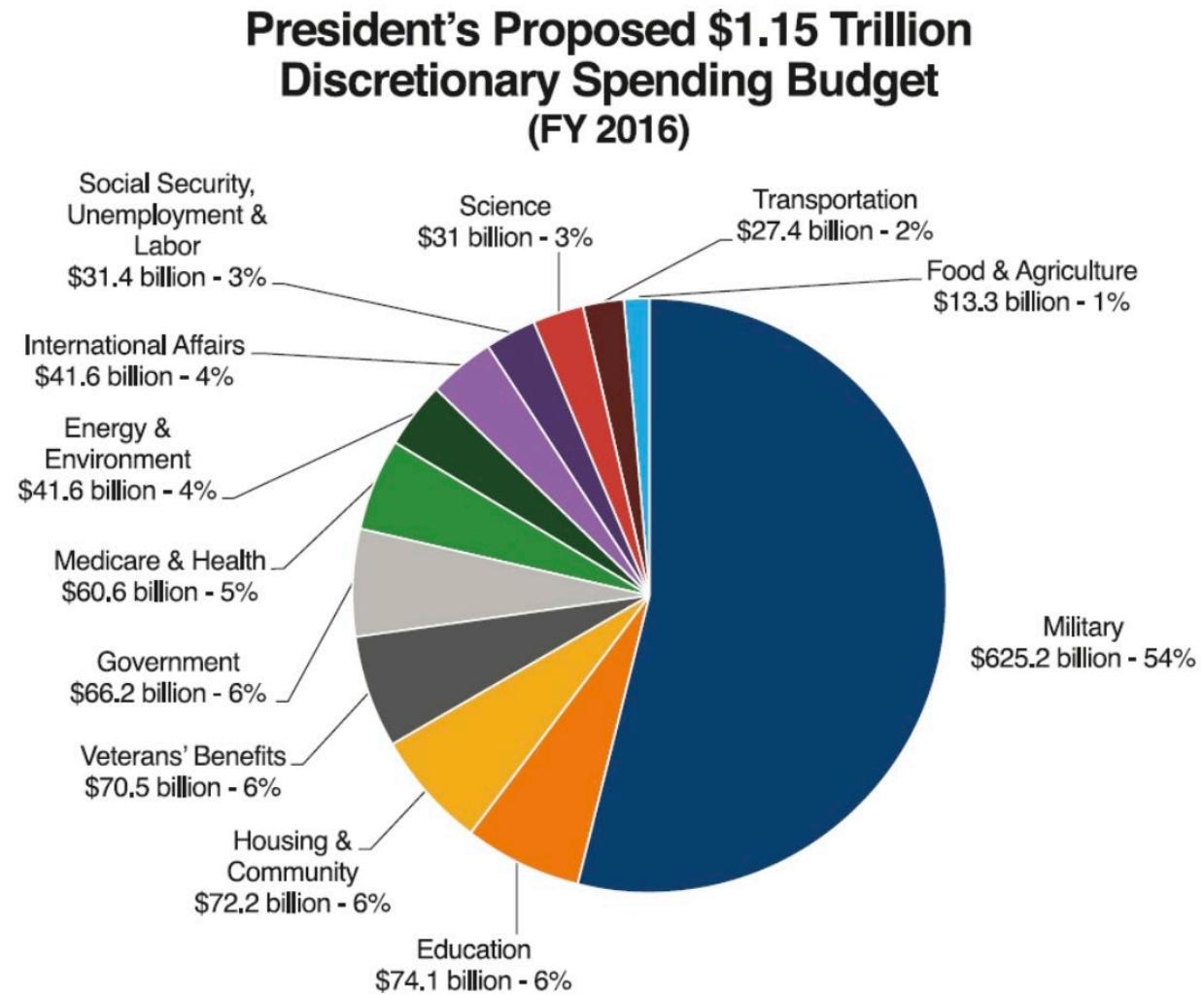


Width

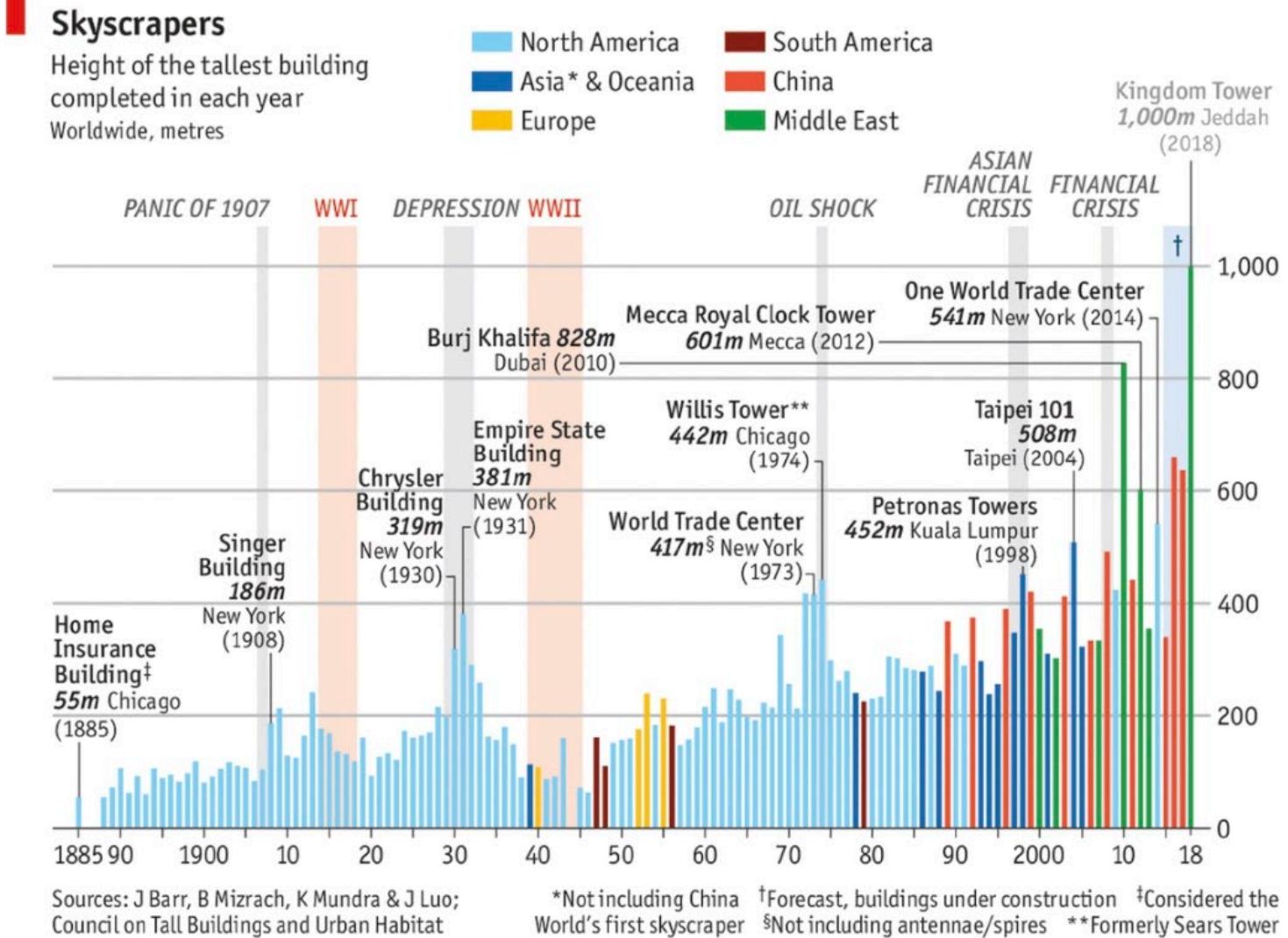


Intensity

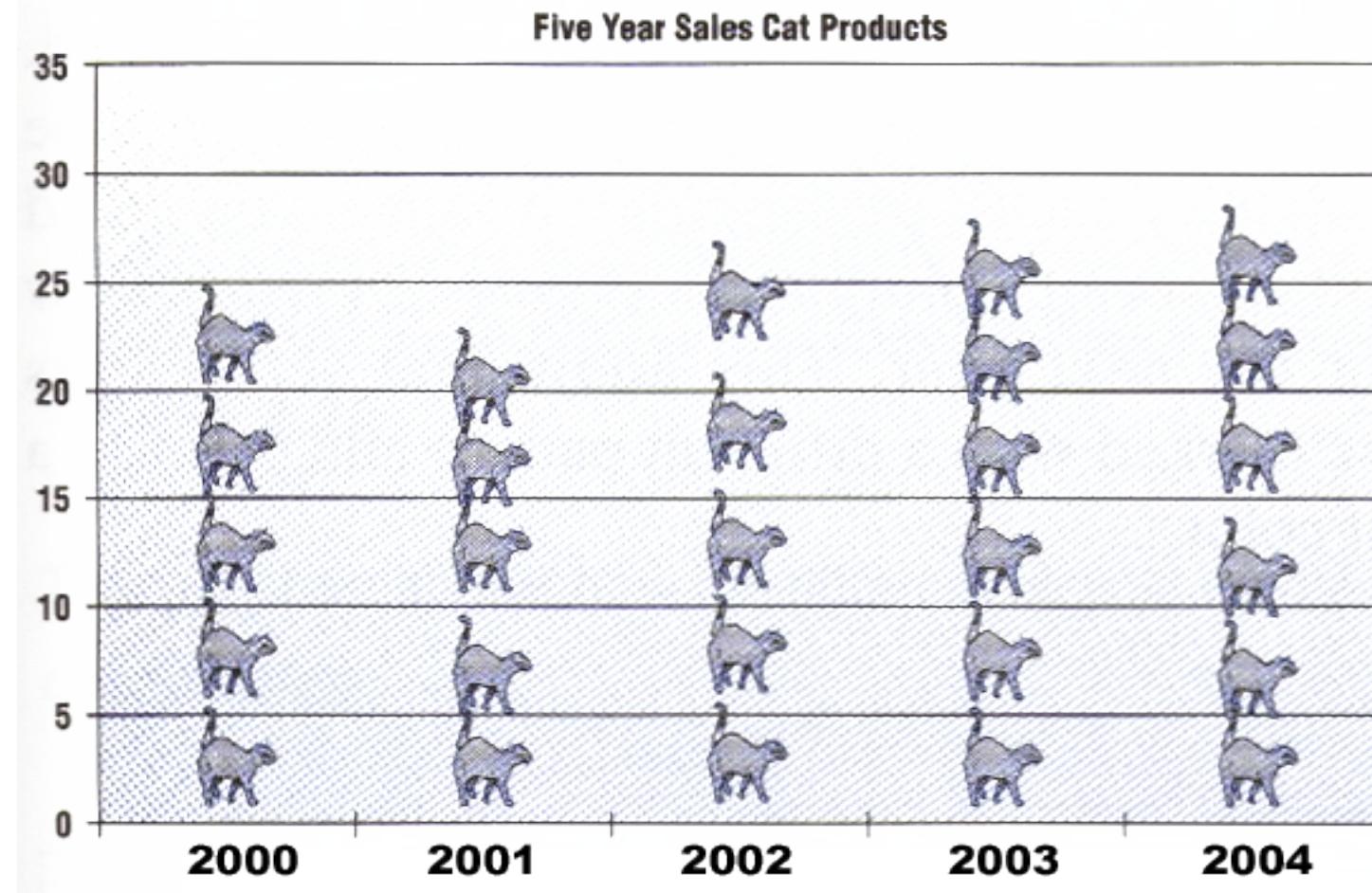
This seems so obvious.  
But then why do we still  
make graphics like this?



Or like this?



Or, for God's sake,  
like this?



# Use the Workings of Visual Perception to I Your Advantage

Design dataviz with the message you want communicated in mind

Devote yourself to keeping your audience's advanced brain (i.e., prefrontal cortex) as still and quiet as possible

Build pre-attentive attributes into your visualization to efficiently connect your message to your audience

Keep in mind that successful dataviz results from more than the use of color, shading, and other pre-attentive attributes

# References

Anatomography. (2012). [Frontal lobe—Posterior view](#) [Online image]. CC BY-SA 2.1

Simons, D. J., & Levin, D. T. (1998). [Real-world change blindness from Simons and Levin](#) [Video file]

A photograph of several students sitting on a long bench in a hallway. In the foreground, two female students are smiling and looking at a yellow notebook. Behind them, other students are seated, one reading a book and another looking down. The background shows a blurred view of the hallway and some architectural details.

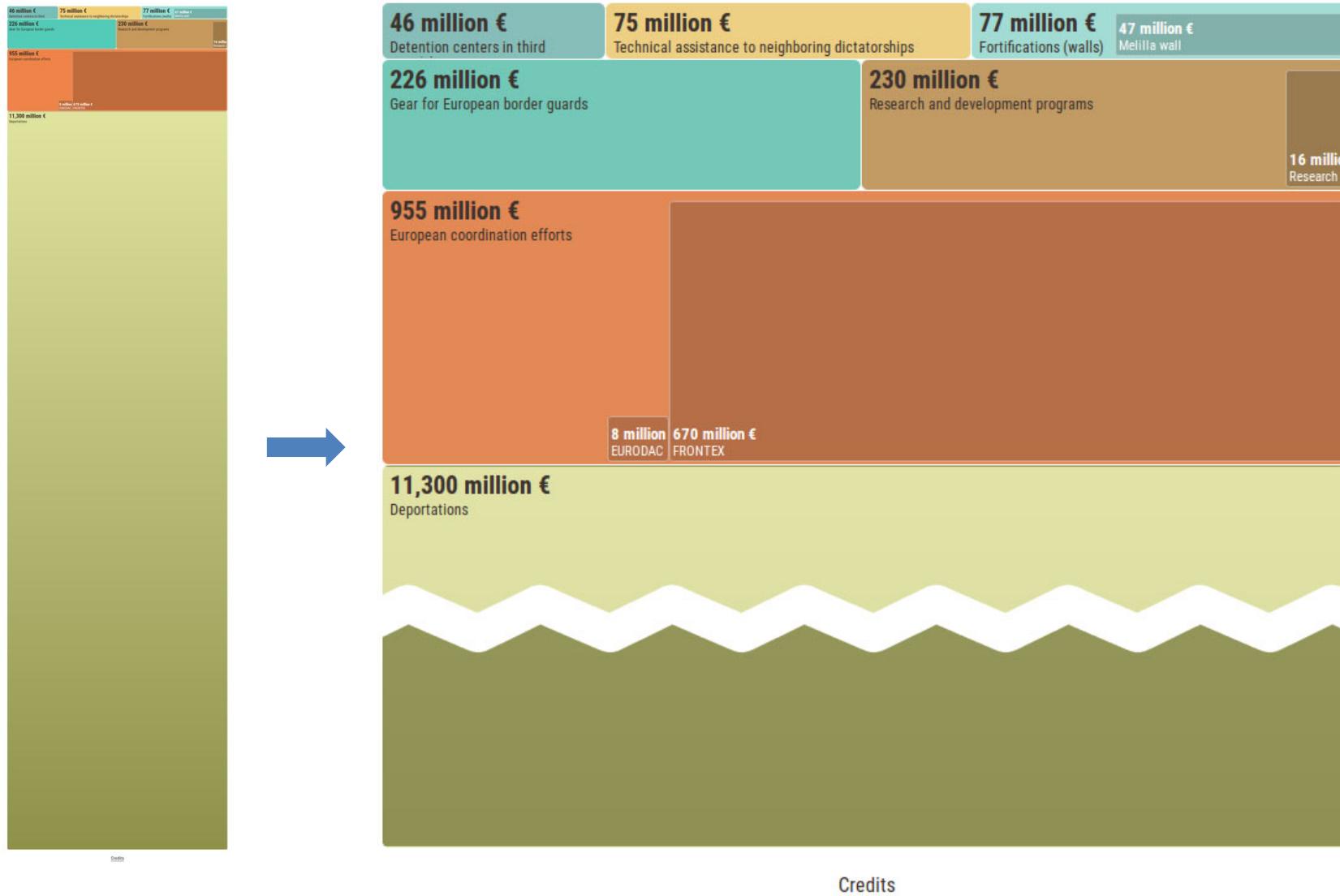
Module 1 Lesson 4

# Evaluating the Effectiveness of Dataviz

# Evaluating the Effectiveness of Dataviz

Creating effective, impactful, and good-looking dataviz requires a successful blend of several key elements

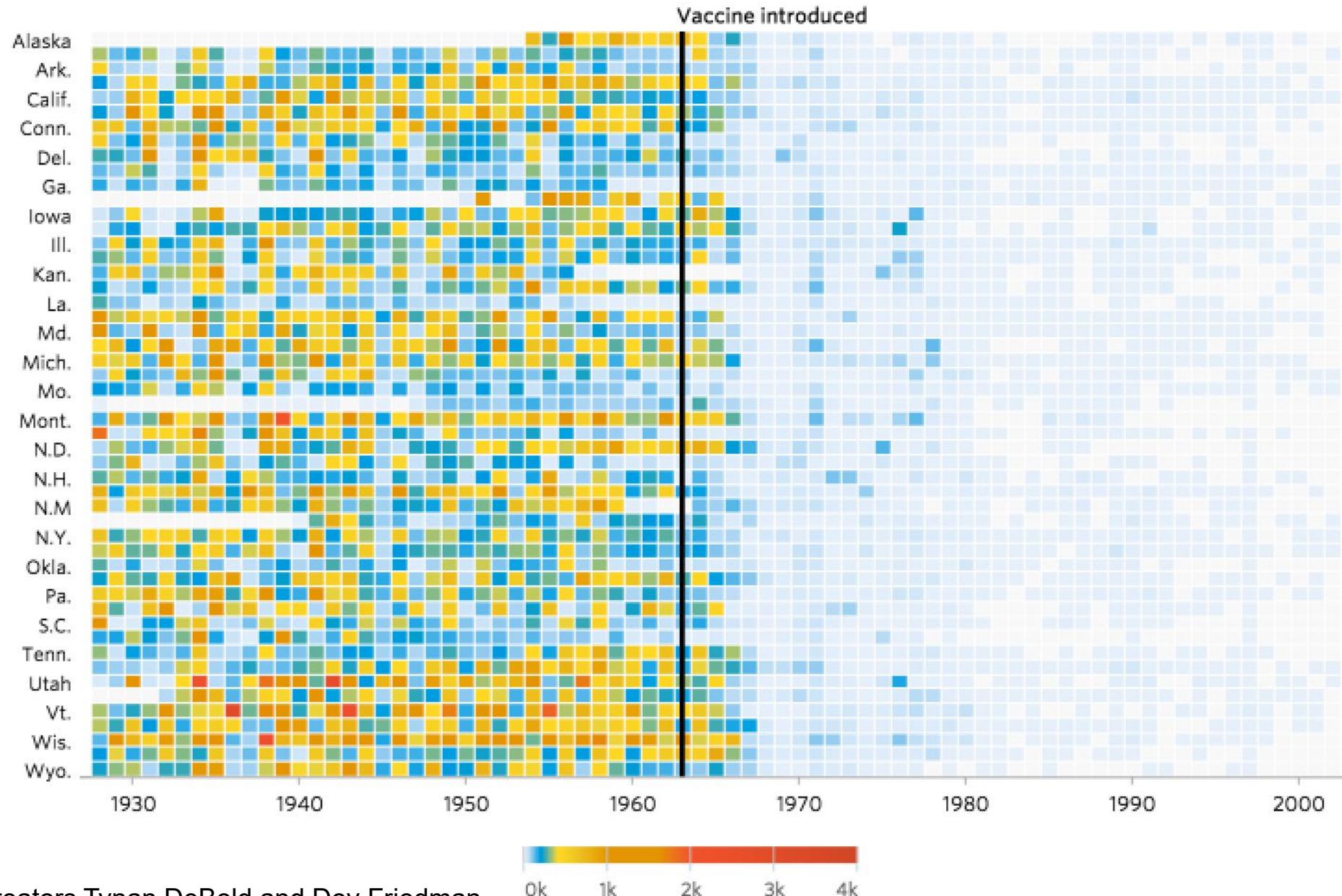
# Billion Dollar-o-Gram



Source: Open Budgets EU's "Billion Dollar-o-Gram" generator



# Measles



1.83 million  
square miles





**1.39 million  
square miles**

**2012**

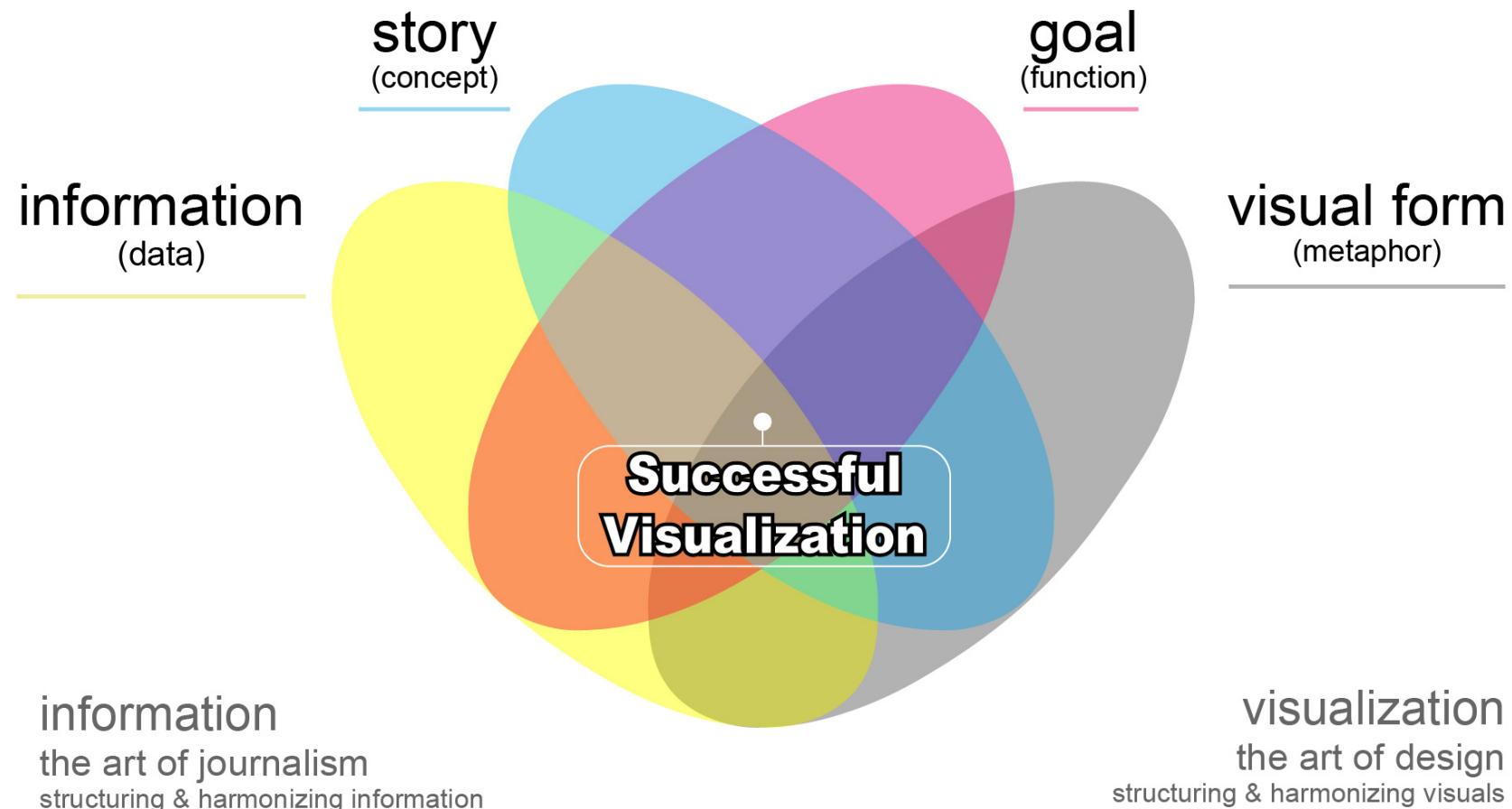
I



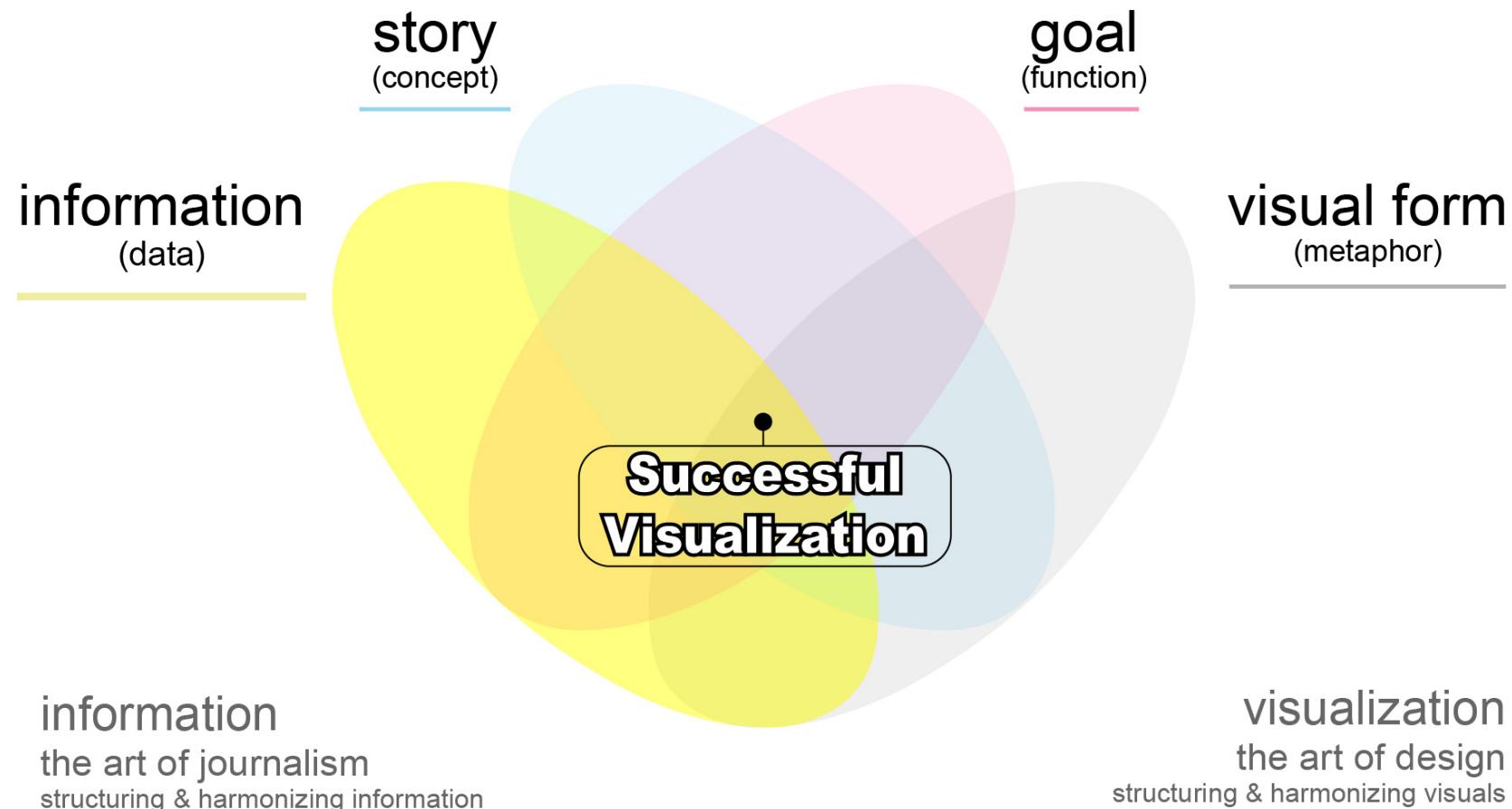
440,000 square miles are gone: That's bigger than Texas

Source: National Snow & Ice Data Center, U.S. EPA

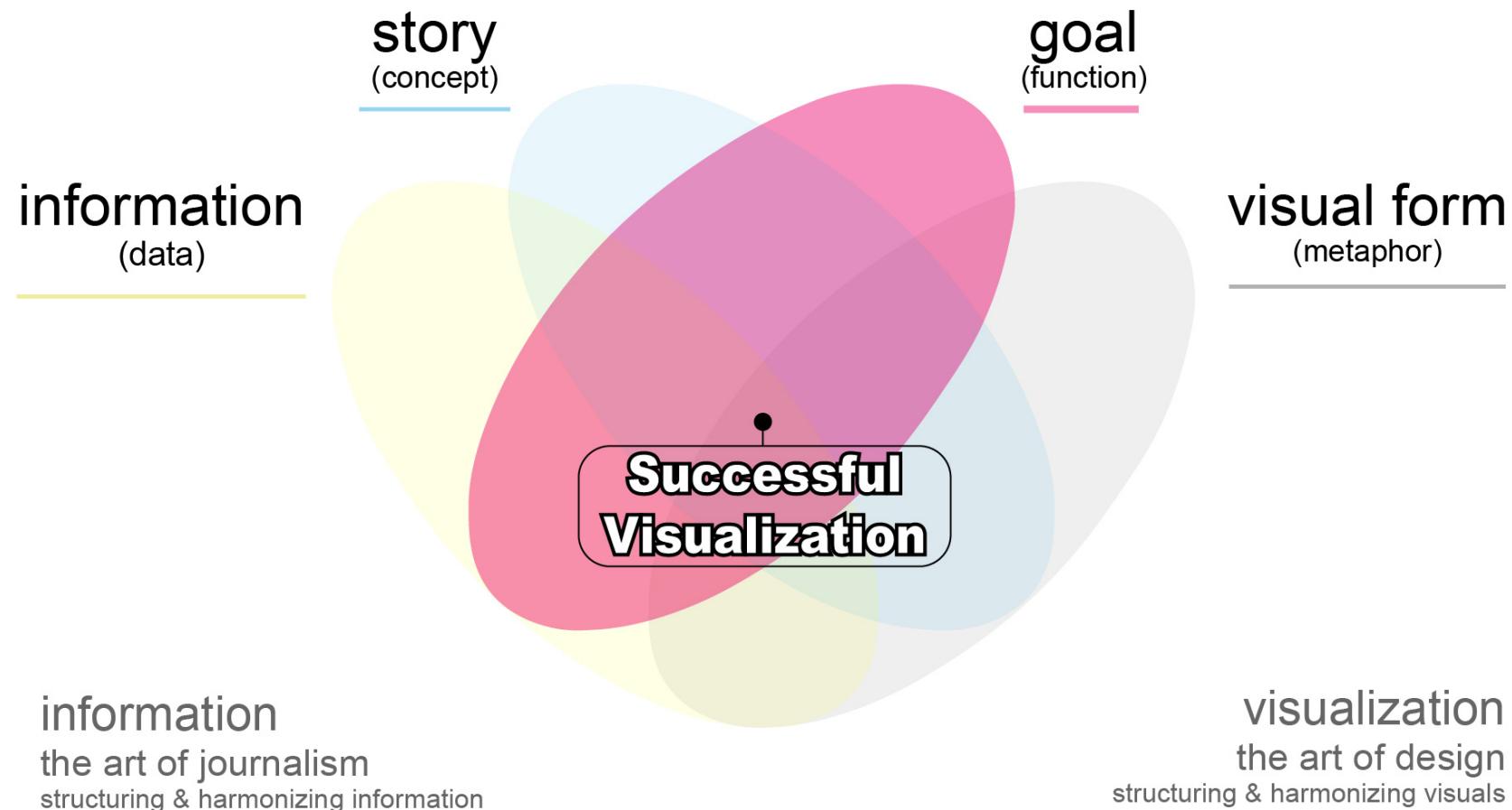
# McCandless Offers a Thorough Definition of Good Data Stories



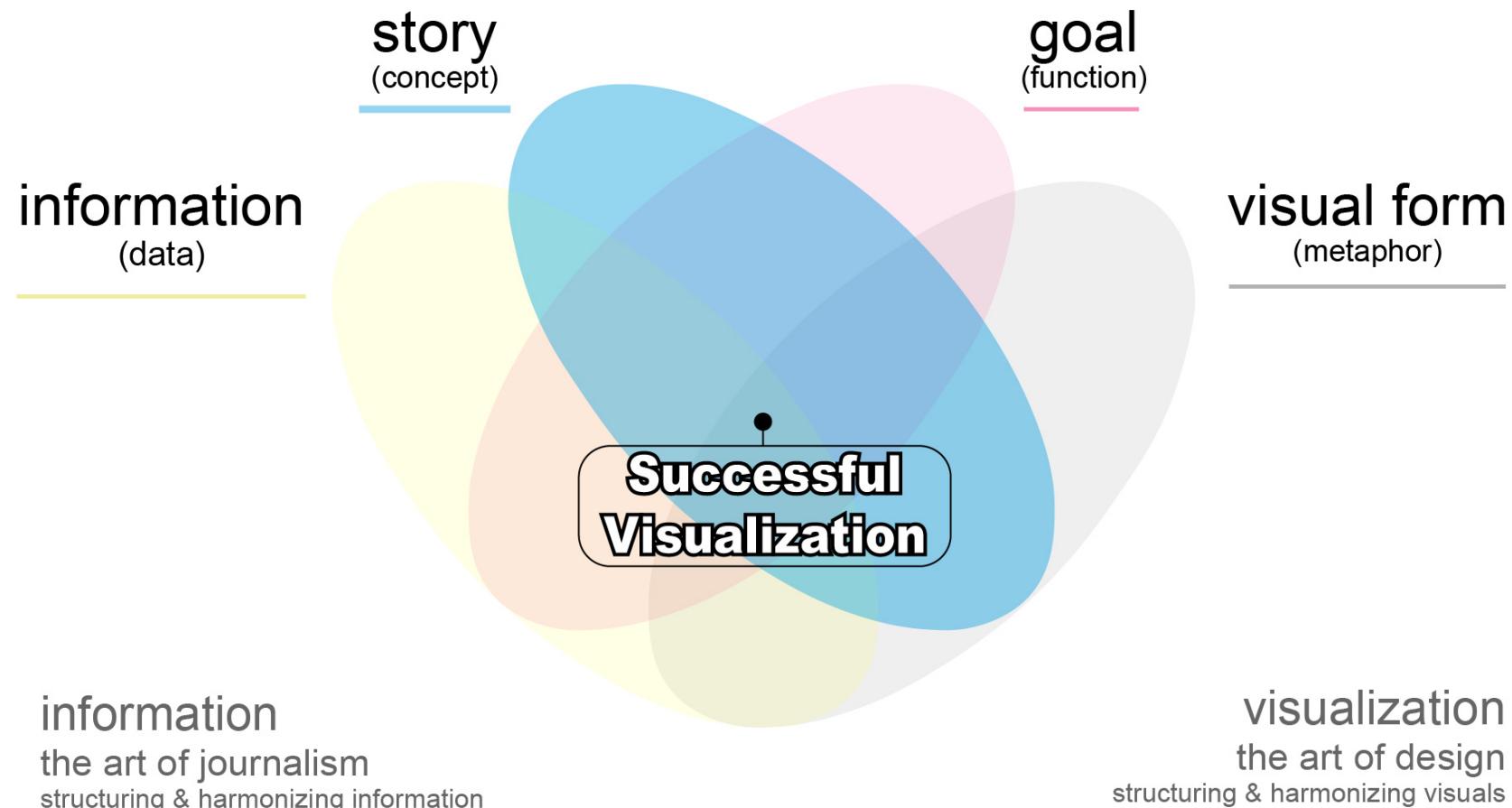
# McCandless Offers a Thorough Definition of Good Data Stories



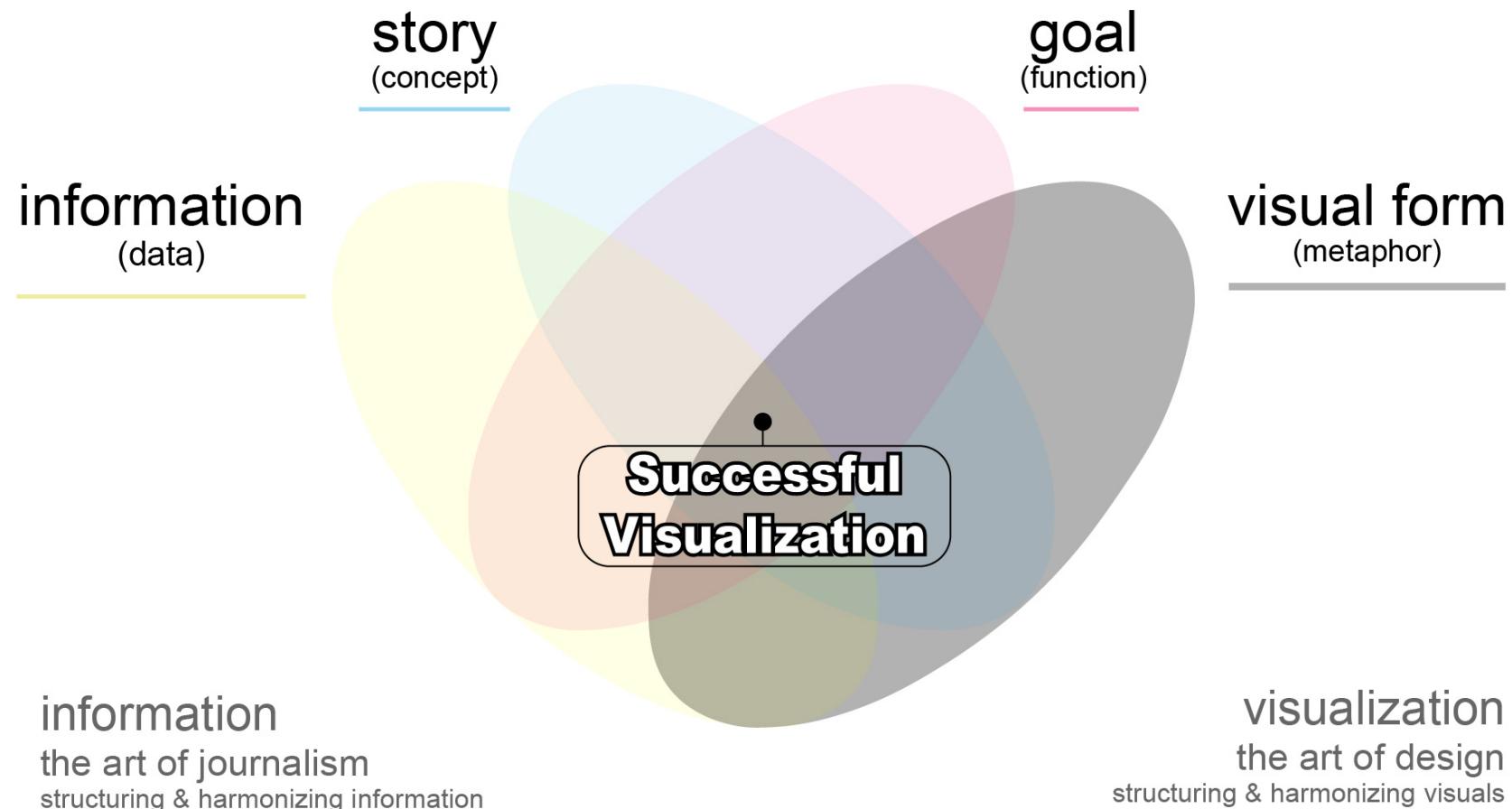
# McCandless Offers a Thorough Definition of Good Data Stories



# McCandless Offers a Thorough Definition of Good Data Stories



# McCandless Offers a Thorough Definition of Good Data Stories



# McCandless's Framework Brings Quality I to All Dataviz Elements

Provides a clear answer to the elusive question: “What makes a data visualization good?”

Goes beyond simply considering the image to evaluate the entire creative process

Compels the author to pay attention to each individual component of a data visualization

Ensures all the elements come together to form a single, harmonious dataviz

# In This Module

I

## Module 1: Pictures You See with Your Brain

### Key Concepts

Studying the history of dataviz

Understanding today's dataviz tools

Making connections with visuals

Evaluating the effectiveness of dataviz

# References

Github. (2012). [Open Budgets EU's "Billion-dollar-o-gram" generator](#) [Online image]