

LEGENDE... Quantité et couleur pour chaque type de présentation					
	Étiquette	Boîte	Assiette	Boîte individuelle	Boîte empilée
à la cuillère.	11111	11111	11111	11111	11111
à la cuillère.	11111	11111	11111	11111	11111
à la cuillère.	11111	11111	11111	11111	11111

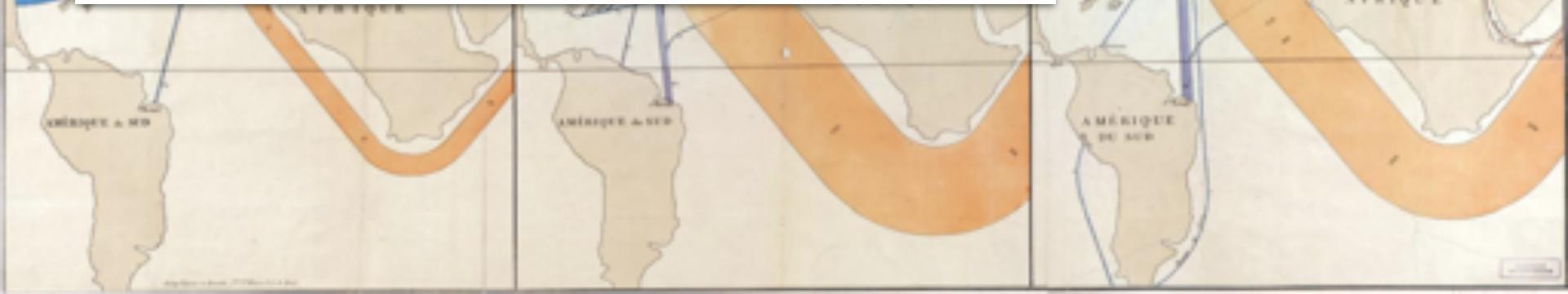
Carte figurative et approximative des quantités de COTON BRUT importé en France en 1858 et 1864 et en 1862.

Dessin par MC MINAHD, Imprimeur Général des Etats et Chambre des députés
Paris, le 14 Mai 1940.

Session 1.1



Introduction to Data Visualization



Online Course **Data Visualization for Professionals**

THE UNIVERSITY
of EDINBURGH

Benjamin Bach

June 2020

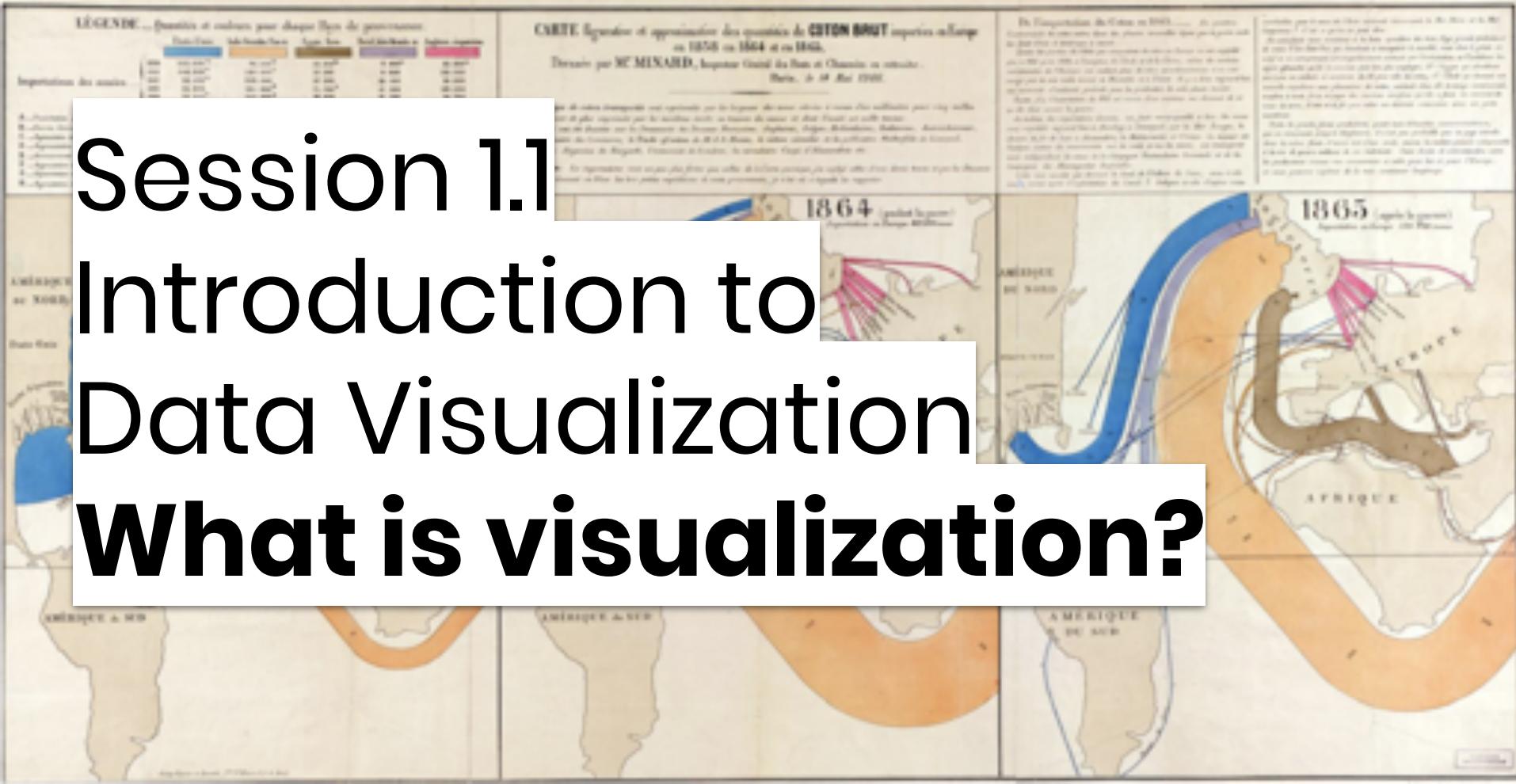
<http://benjbach.me>
<https://datavis-online.github.io>

-- Not for external use --

Outline

What are visualizations and what do we need to learn about them?

1. Visualization usage
2. Defining "*visualization*"
3. Visualization purposes
4. Deceptive visualizations (*how not to use visualization*)
5. Defining "*visualization literacy*"



Session 1.1

Introduction to Data Visualization

What is visualization?



Online Course

Data Visualization for Professionals

THE UNIVERSITY
OF EDINBURGH

Benjamin Bach

June 2020

-- Not for external use --

Visual representation of data:

Table - StateData ()

Load Snap

State	College Degree %	Per Capita Income	Median Household Income	Population
Alabama	20.6%	11486	30.4%	4389
Alaska	30.3%	17610	19.9%	9648
Arizona	27.1%	13461	22.3%	12989
Arkansas	17.0%	10520	25.4%	11213
California	31.3%	16409	26.0%	12452
Colorado	33.9%	14821	21.5%	15214
Connecticut	33.8%	20189	32.4%	15959
Delaware	27.9%	15854	30.1%	18714
District of Columbia	36.4%	18881	25.5%	11246
Florida	24.3%	14698	29.6%	16501
Georgia	24.3%	13631	24.2%	12885
Hawaii	31.2%	15770	28.1%	11051
Idaho	25.2%	11457	22.3%	13461
Illinois	26.8%	15201	22.8%	11893
Indiana	20.9%	13149	27.5%	13418
Iowa	24.5%	12422	23.2%	14068
Kansas	26.5%	13300	27.5%	14981
Kentucky	17.7%	11153	23.0%	11897
Louisiana	19.4%	10635	24.6%	10661
Maine	25.7%	12957	20.1%	12255
Maryland	31.7%	17730	25.5%	12904
Massachusetts	34.5%	17224	30.0%	11029
Michigan	24.1%	14154	31.5%	13527
Minnesota	30.4%	14389	30.0%	15713
			30.9%	14923
			16.1%	10520
			24.9%	13276
			25.7%	12311

Visual representation of data:

Does not show:

- Relations, patterns
- Distinguish values
- does not communicate
- Does not engage
- Does not reveal
- **Requires cognitive effort!**

Visual representation of data:

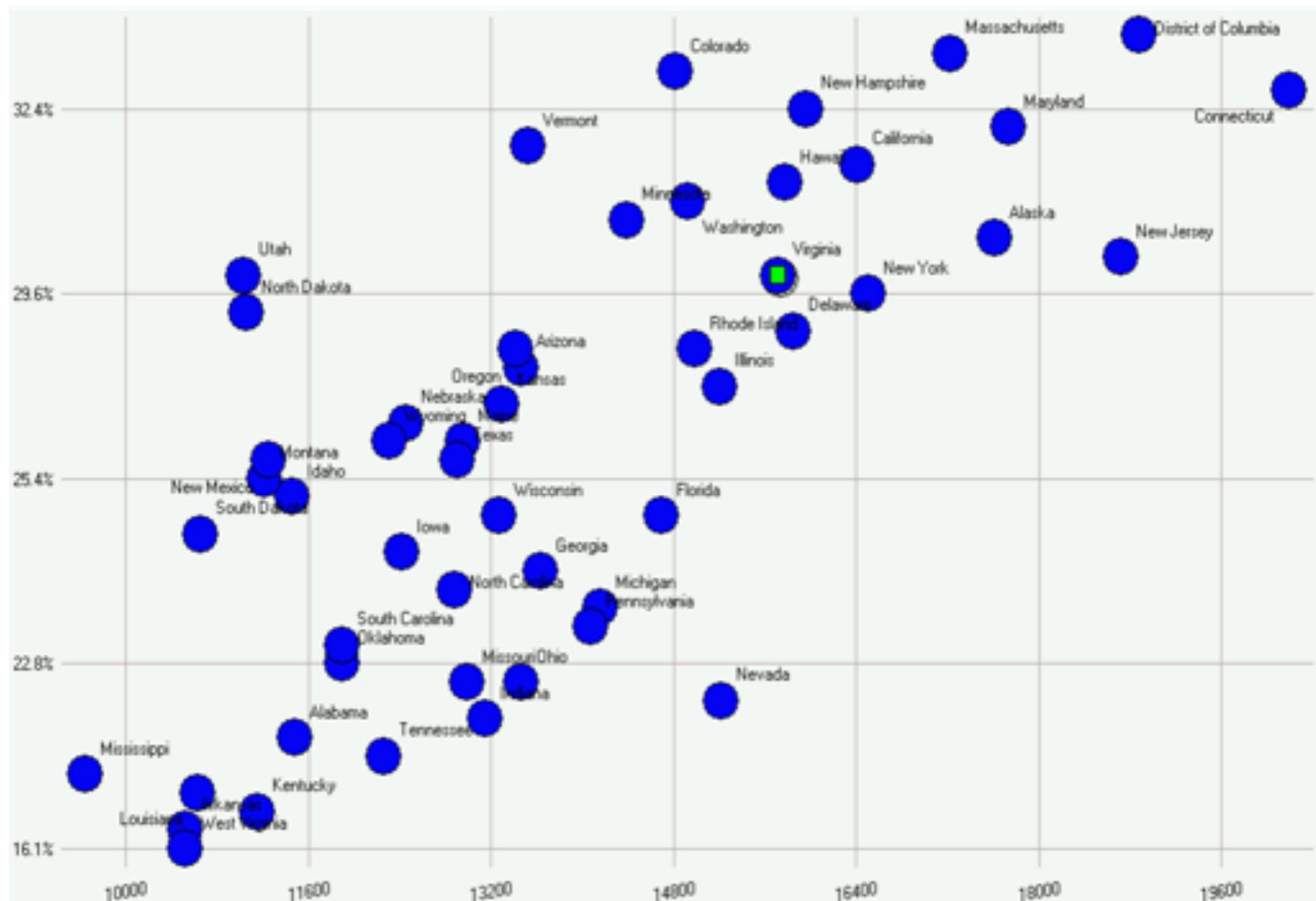
Does not show:

- Relations, patterns
- Distinguish values
- does not communicate
- Does not engage
- Does not reveal
- **Requires cognitive effort!**

Pre-attentiveness:

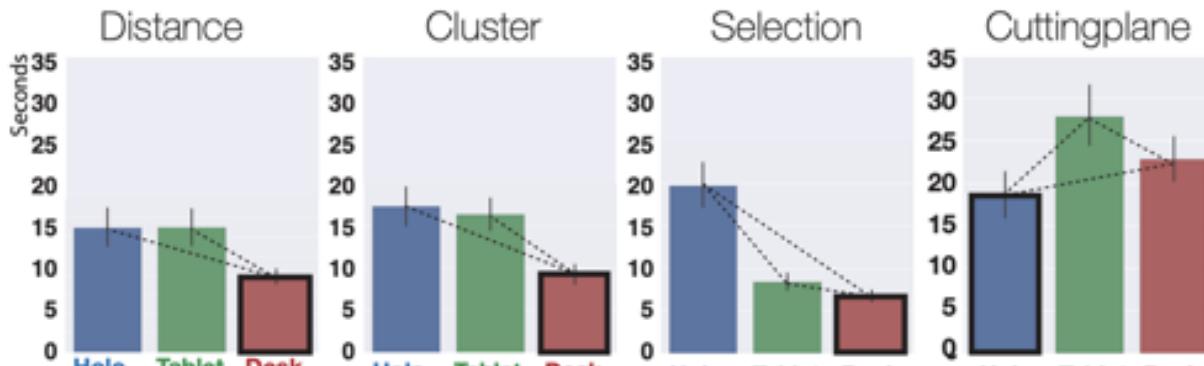
- *"is the subconscious accumulation of information from the environment"*
- realizing something before you think
- **Does not require cognitive effort.**

Visual representation of data:

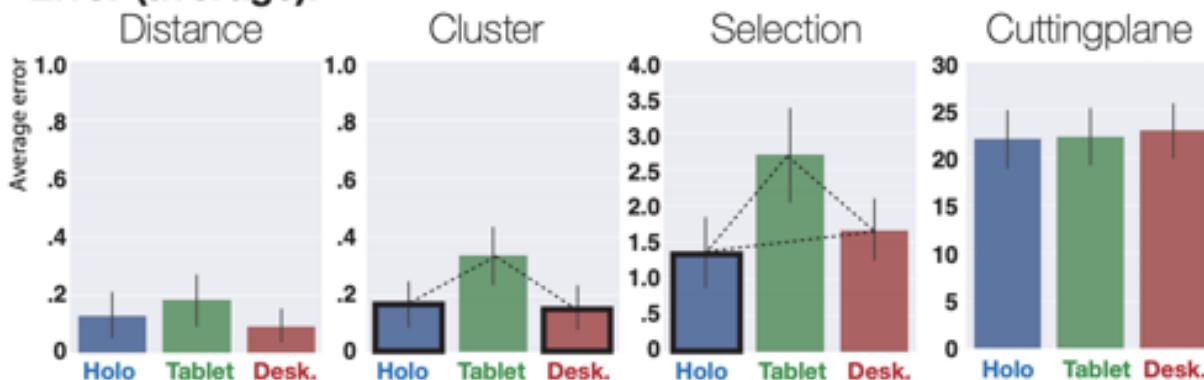


Statistical Reports

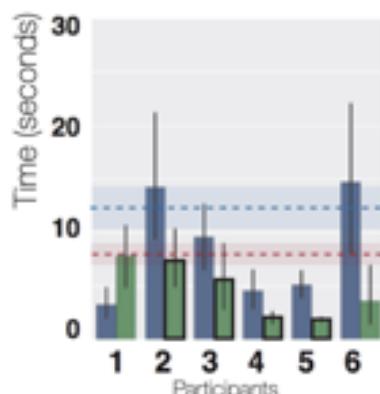
Time (seconds):



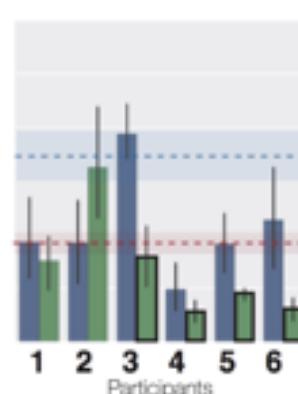
Error (average):



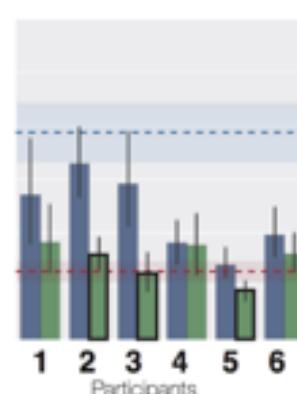
Distance



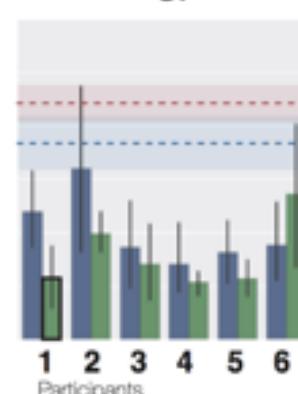
Cluster



Selection



Cuttingplane



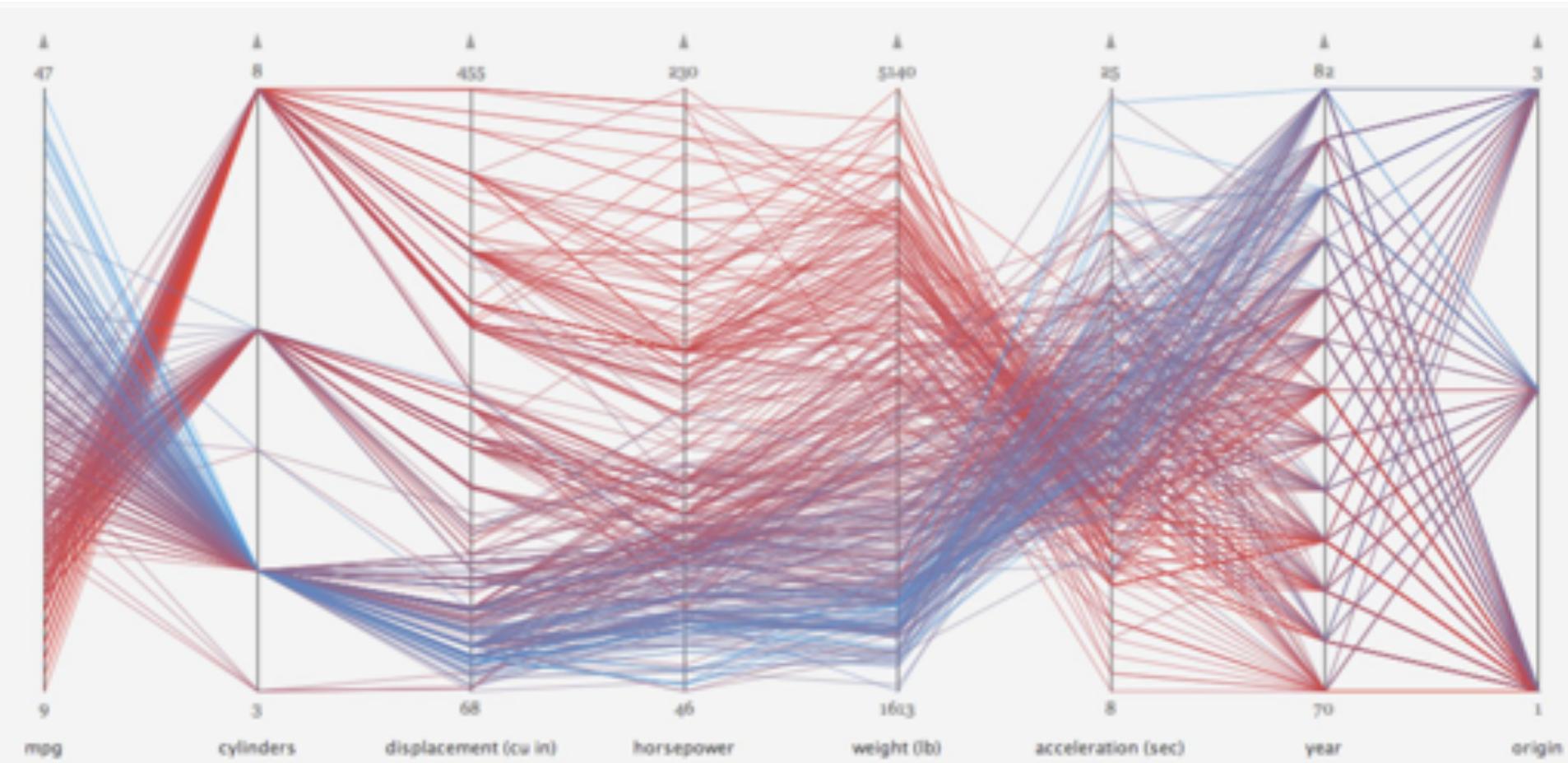
■ 1st session ■ 5th (last) session ■ HoloLens avg. 1st session ■ Desktop avg. 1st session

How Scotland's political geography changed, seat by seat

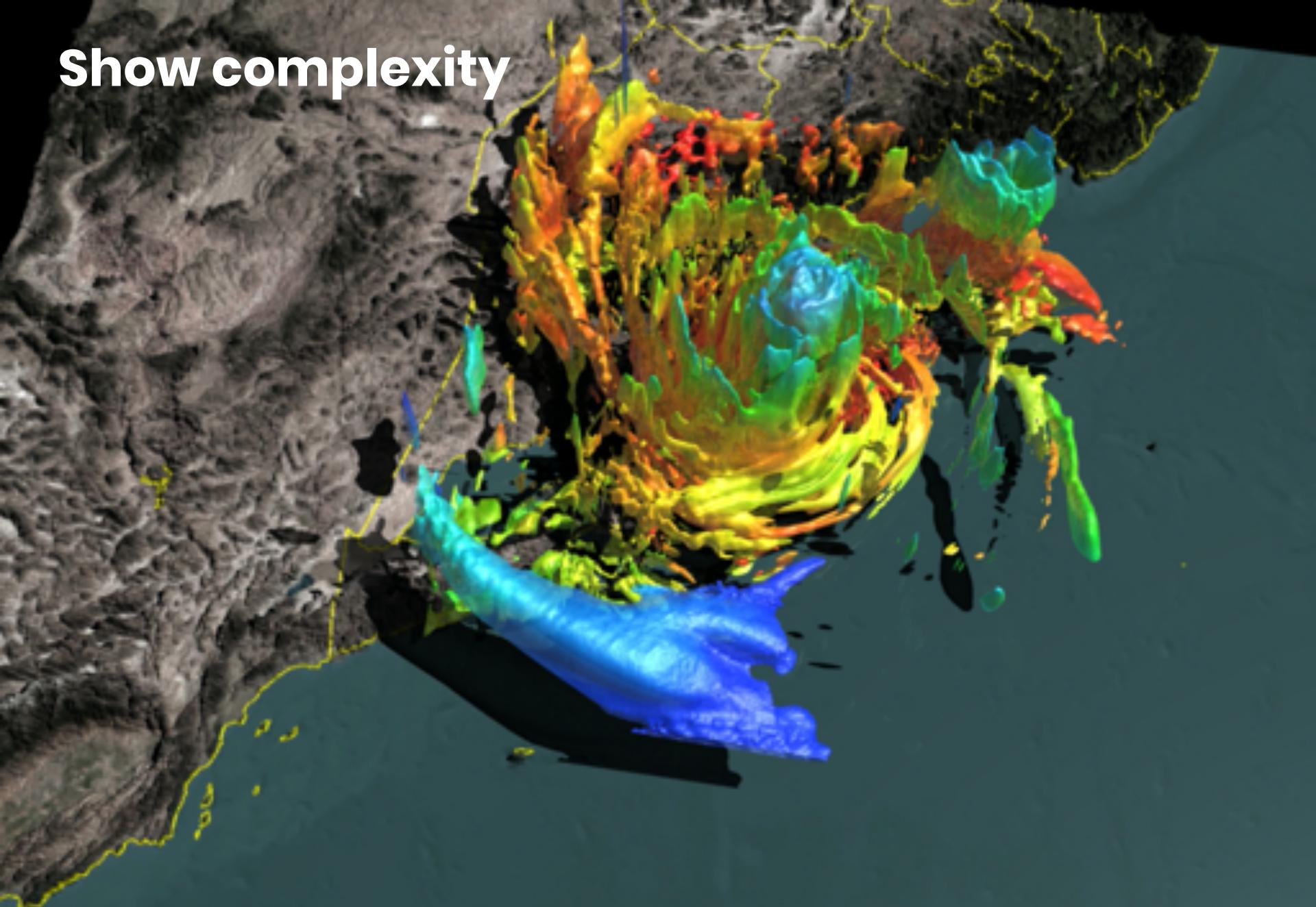
Overview



Exploration



Show complexity

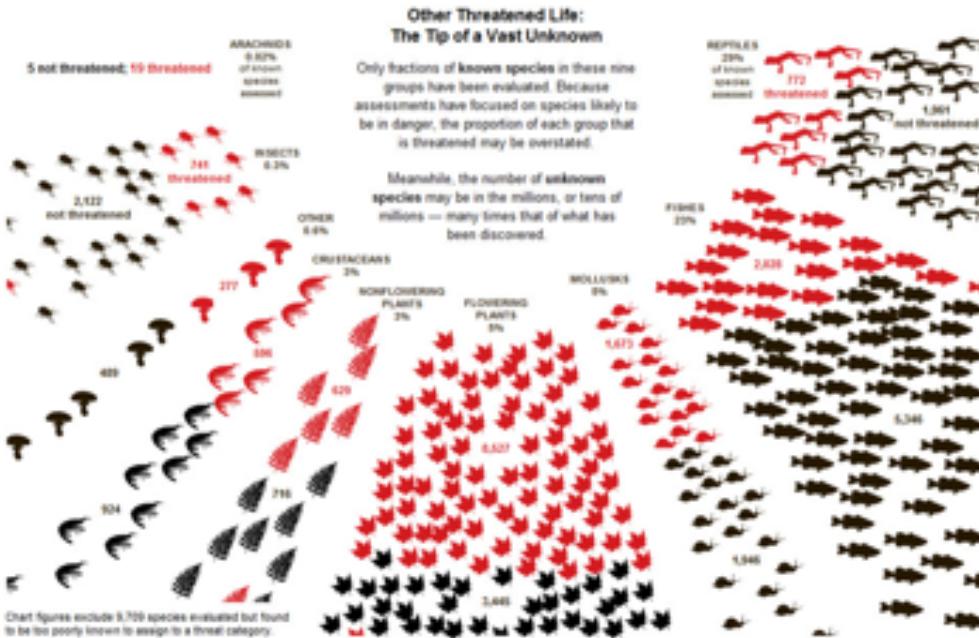


Engage

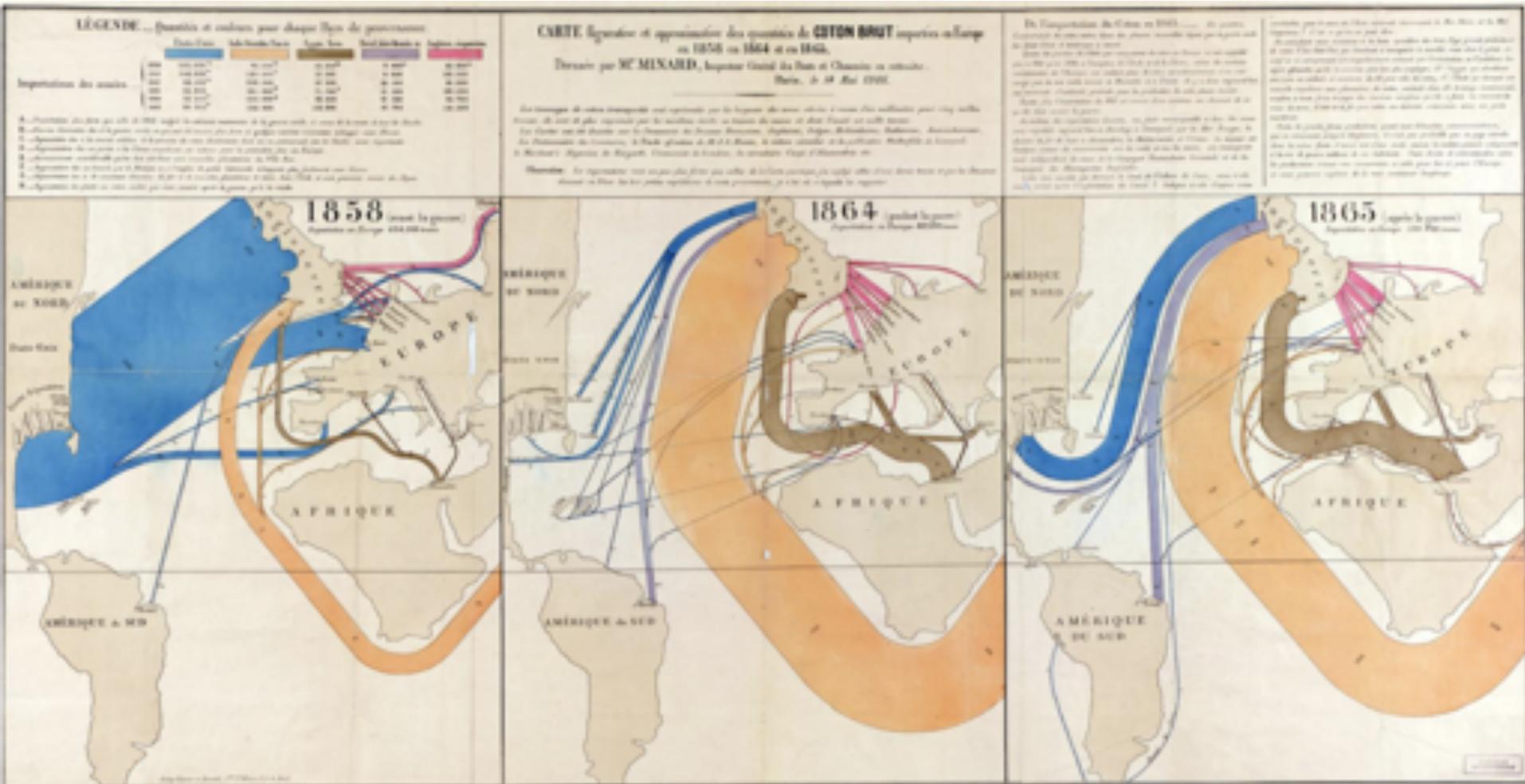
Published: June 1, 2012

OPINION

Are We in the Midst Of a Sixth Mass Extinction?



Explain



Storytelling

COMICS.



Illustration

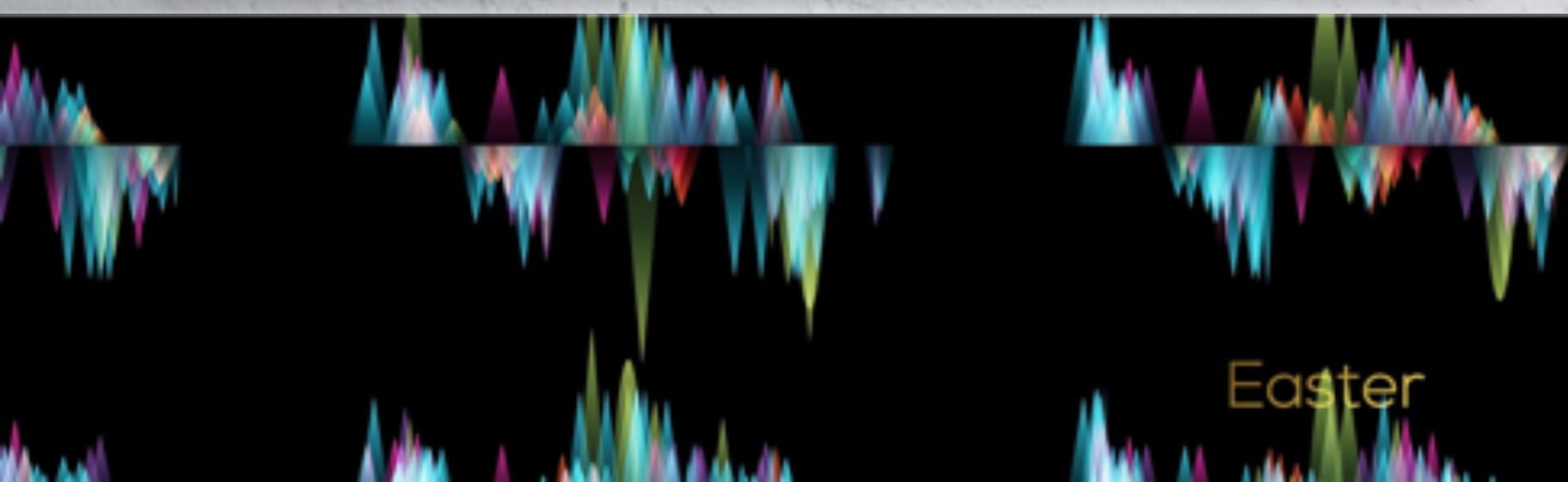


Art

a year in flights

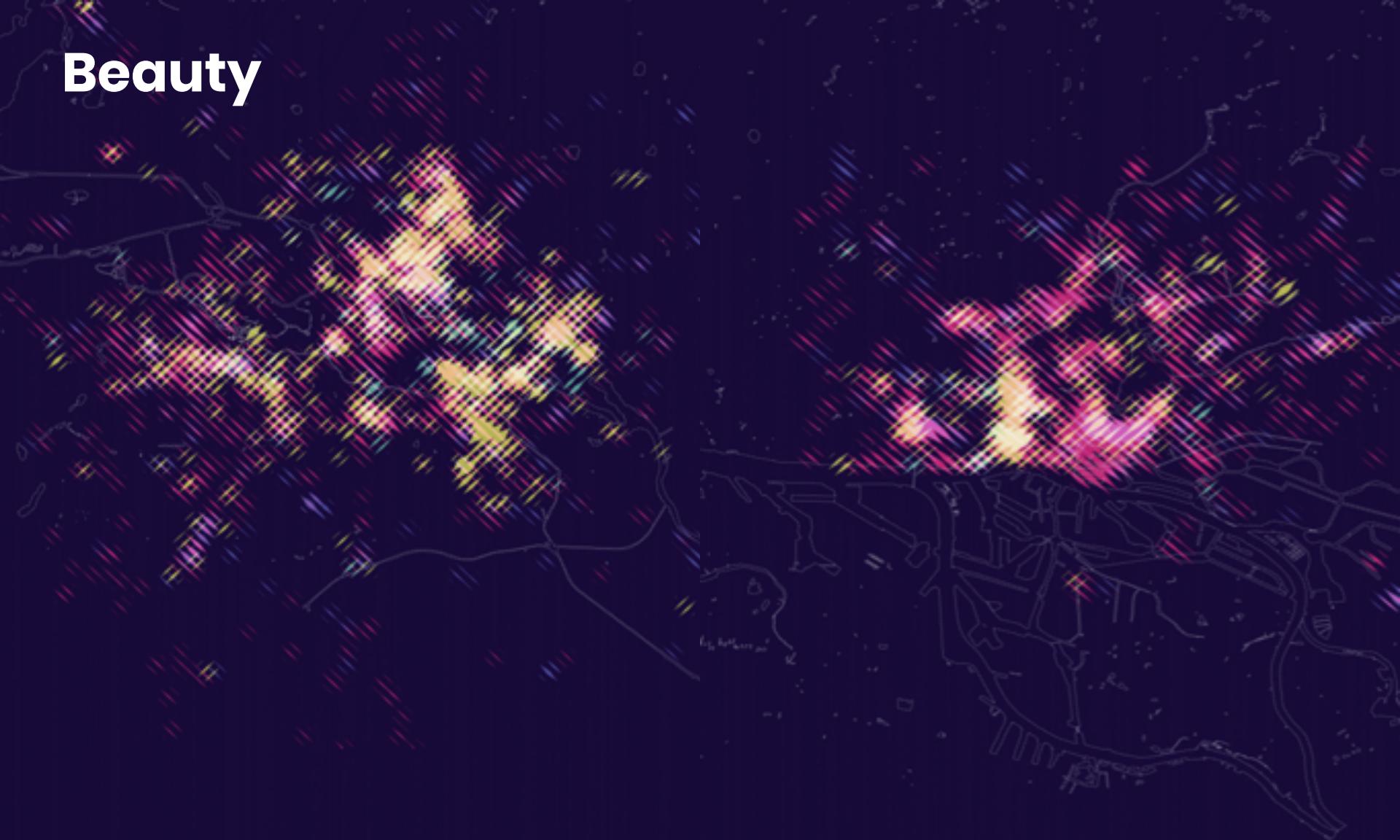


Nadie Brehmber: www.visualcinnamon.com



Easter

Beauty



128 • Music

590 • Shopping

654 • Nightlife

3024 • Food

Berlin Stadtbild

StadtBild is an attempt to map the beauty of cities as manifested in digital traces. While traditional maps show us buildings, roads and physical infrastructure, this map reveals where and in which form the city is alive.

Moritz Stefaner, 2013
<http://truth-and-beauty.net>

52 • Music

350 • Shopping

286 • Nightlife

1558 • Food

Hamburg Stadtbild

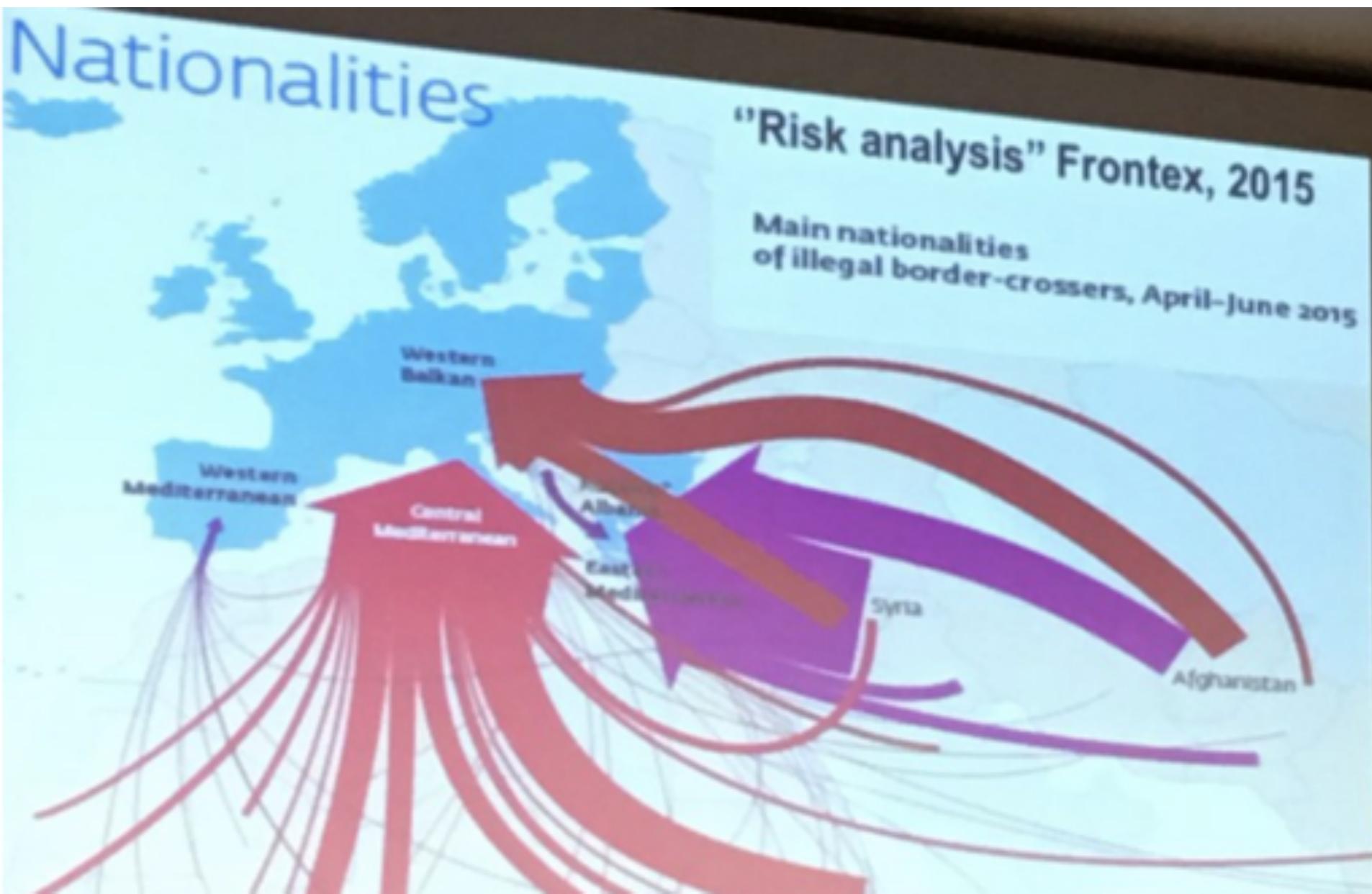
StadtBild is an attempt to map the beauty of cities as manifested in digital traces. While traditional maps show us buildings, roads and physical infrastructure, this map reveals where and in which form the city is alive.

Moritz Stefaner, 2013
<http://truth-and-beauty.net>

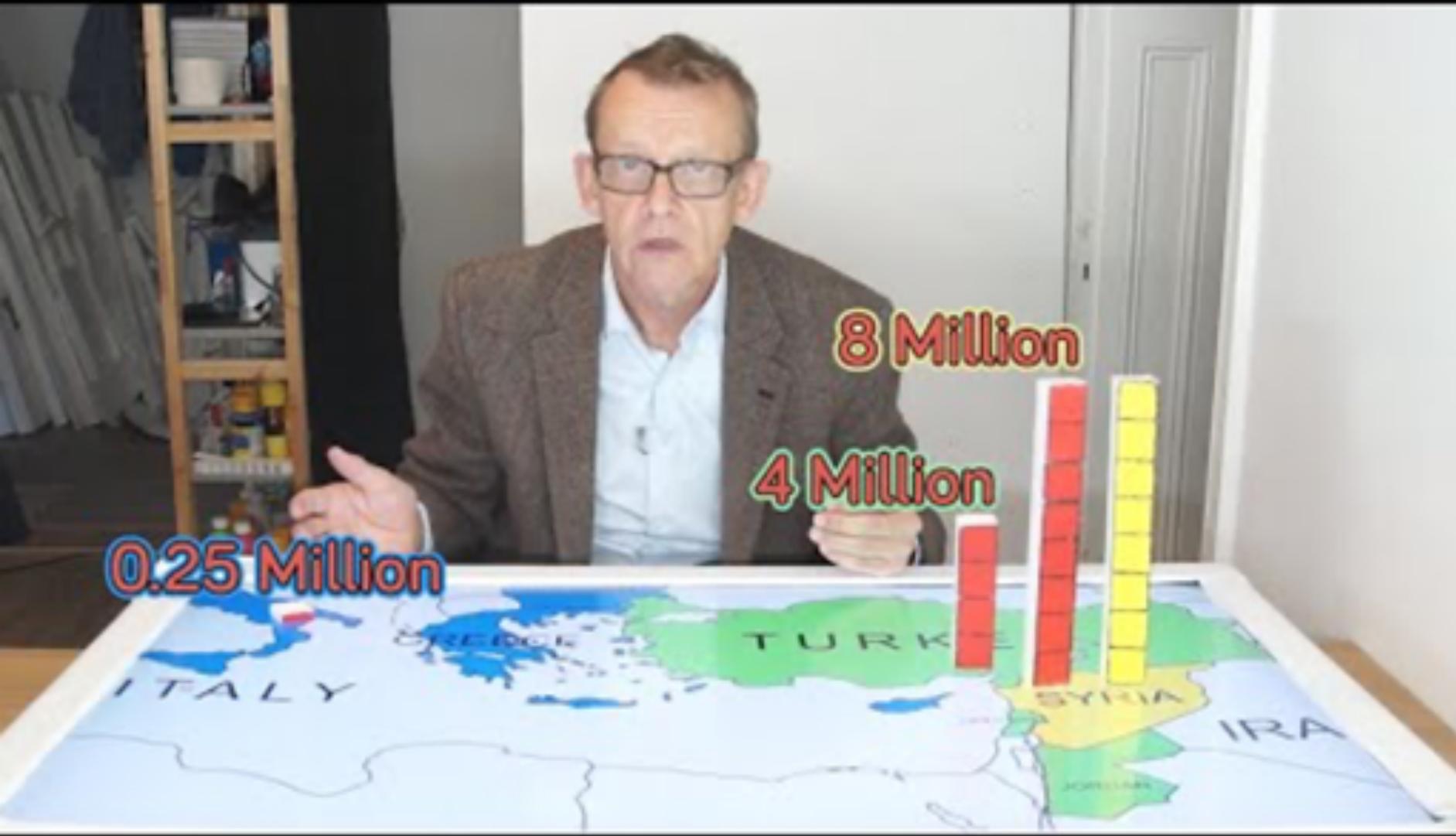
Experiencing

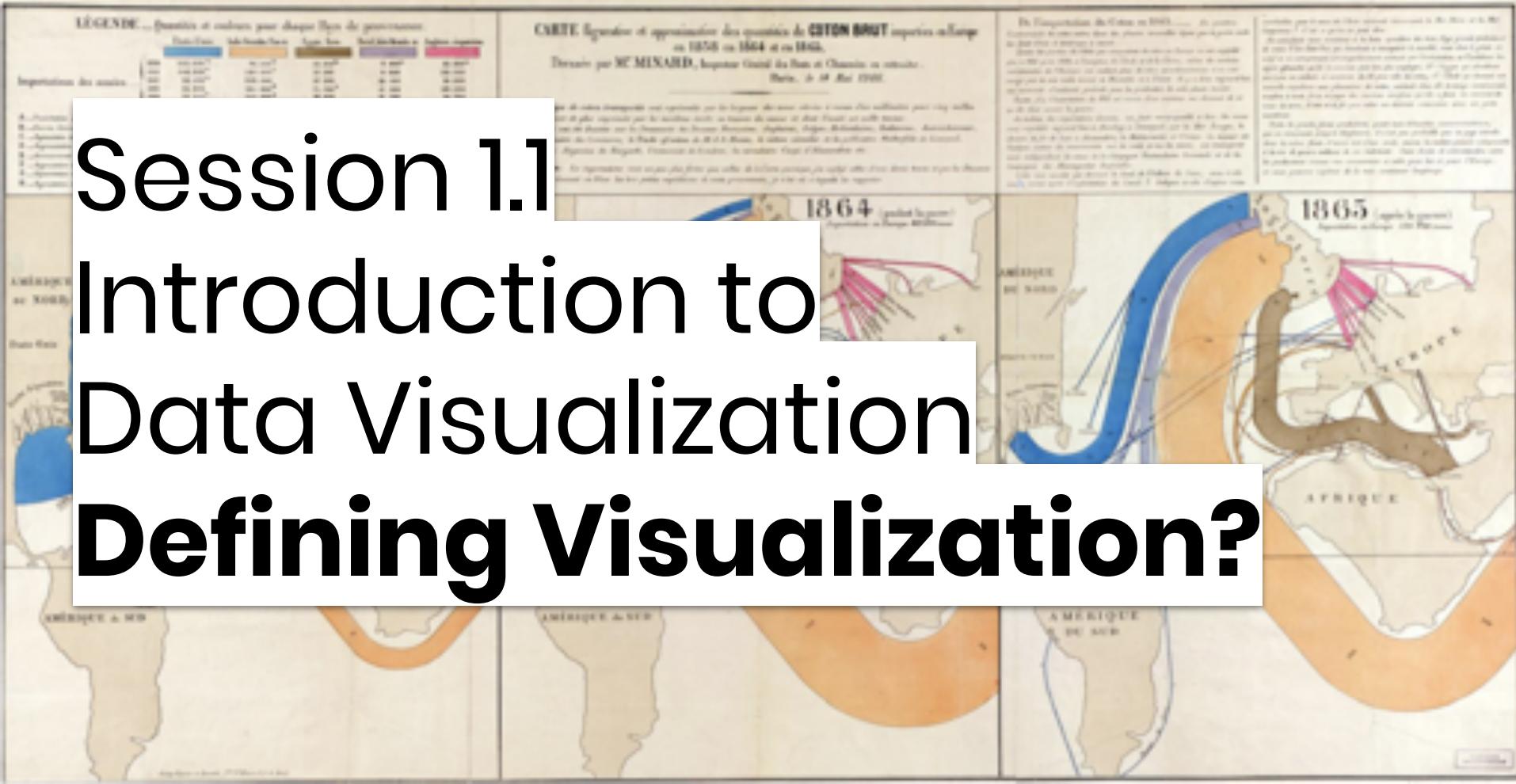


Propaganda



Explanation





Session 1.1

Introduction to Data Visualization

Defining Visualization?



Online Course **Data Visualization for Professionals**

THE UNIVERSITY
of EDINBURGH

Benjamin Bach

June 2020

-- Not for external use --

Definitions

"The use of computer-supported, interactive, visual representations of abstract data to amplify cognition" — Stu Card

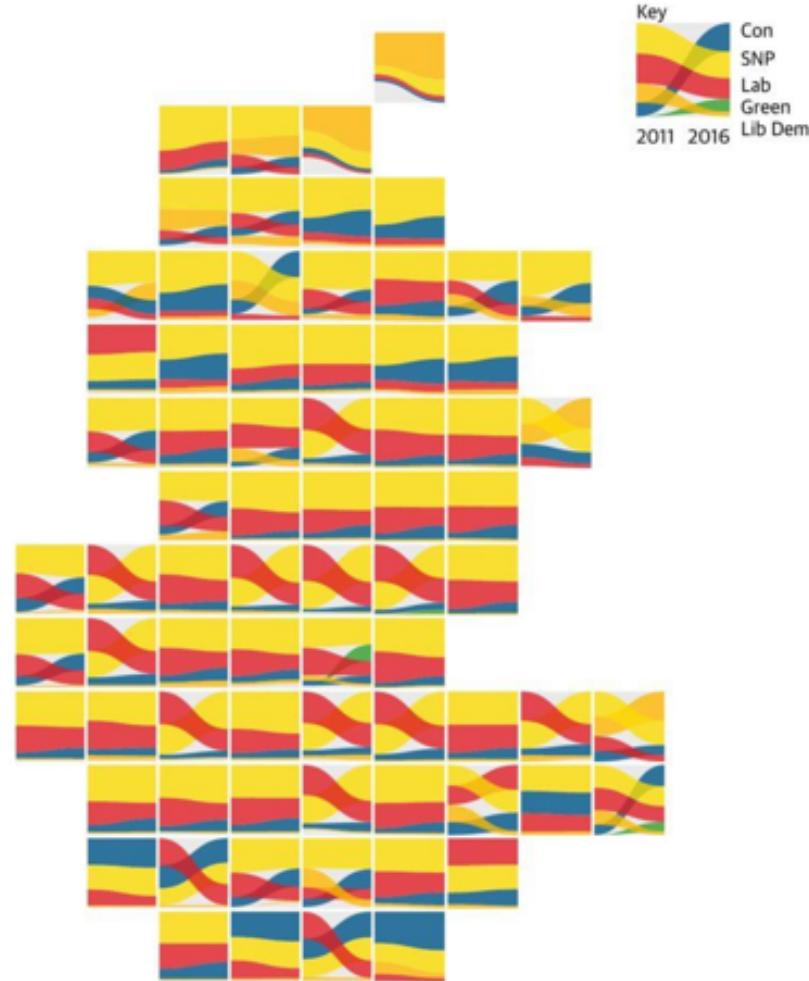
"... an accessible way to see and understand trends, outliers, and patterns in data." — Tableau

". to help people carry out tasks more effectively" — Tamara Munzner

Defining Concepts

How Scotland's political geography changed, seat by seat

- computer supported
- Interactive visual representations
- for abstract data
- helping people
- to see and understand
- trends, outliers, and patterns in data,
- and carry out tasks
- more effectively
- through amplifying cognition
- ...



But also...

Record information

- Tables, blueprints, photographs, diaries

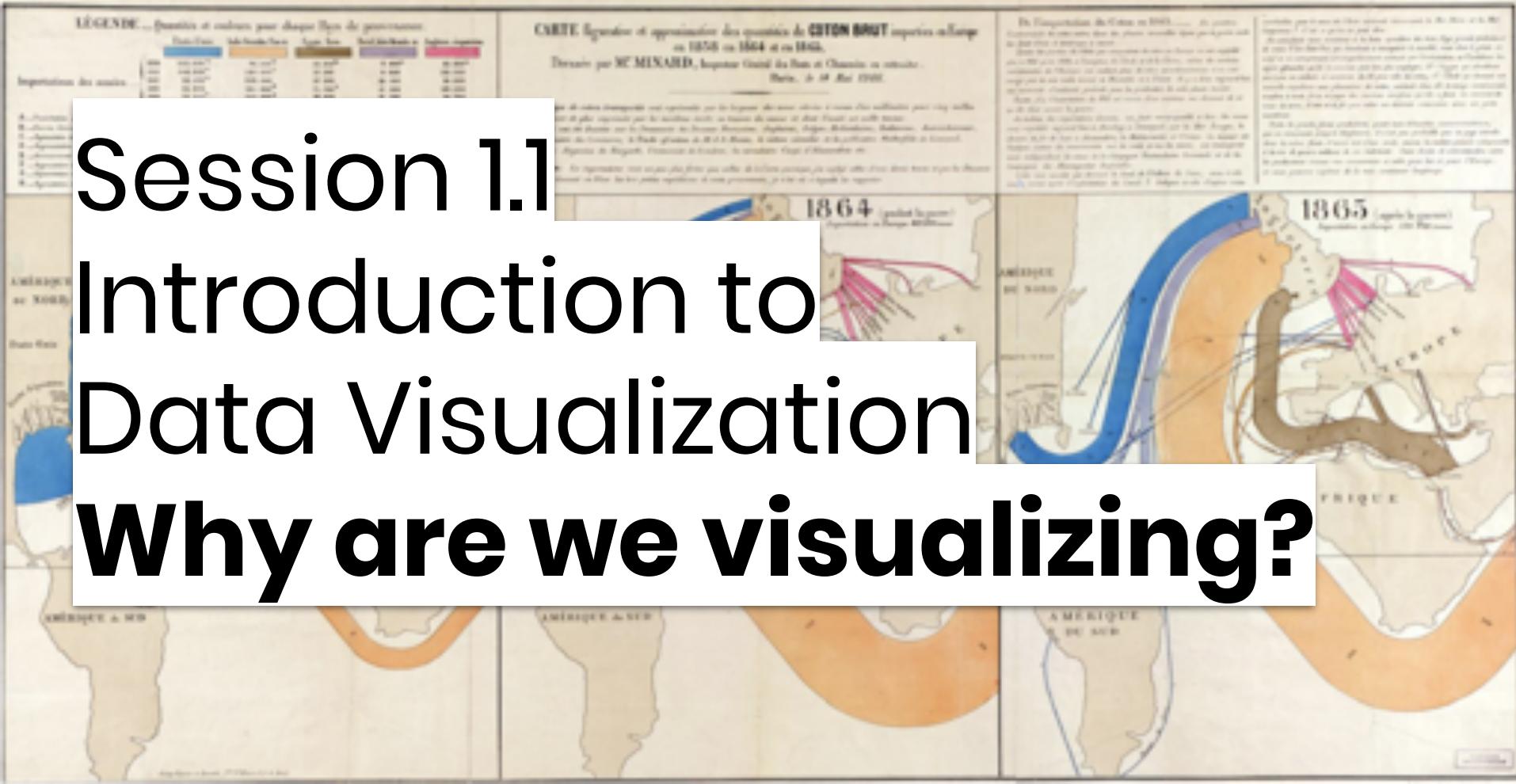
Externalize representation

- Support thinking and reasoning

Presentation and communication

Discovery, serendipity

...



Session 1.1

Introduction to Data Visualization

Why are we visualizing?



Online Course **Data Visualization for Professionals**

 THE UNIVERSITY
of EDINBURGH

Benjamin Bach

June 2020

-- Not for external use --

Exploration

Insights

Explanation

Data centered

Human centered

Domain experts

Non-experts

Generating
Insights

Conveying
messages

Lab Setting

In-the-wild /
diverse

Visualizations goals

- Understanding
- Purposeful
- Overview
- Hypothesis generation
- Trends and outlier
- Uncertainty
- Communication
- Discussion
- Facilitate access

How Scotland's political geography changed, seat by seat

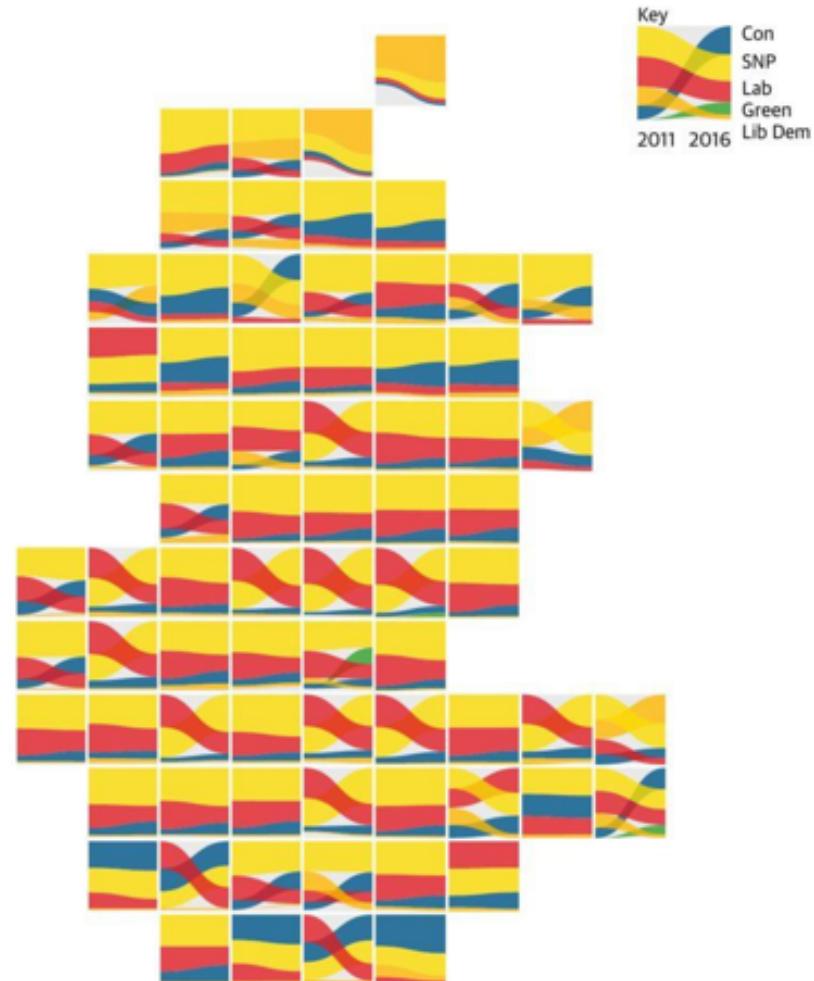
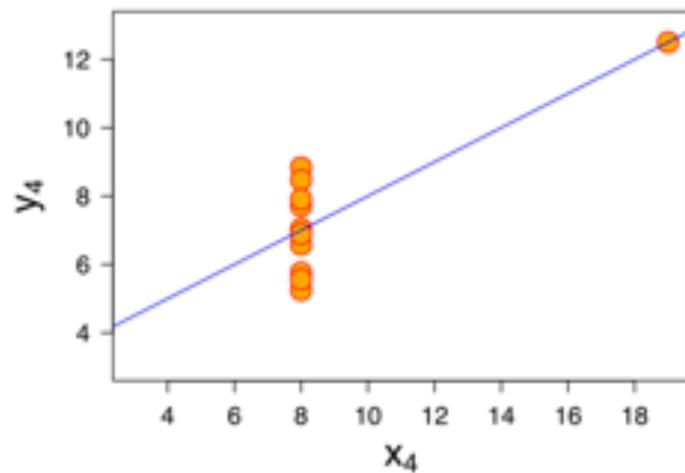
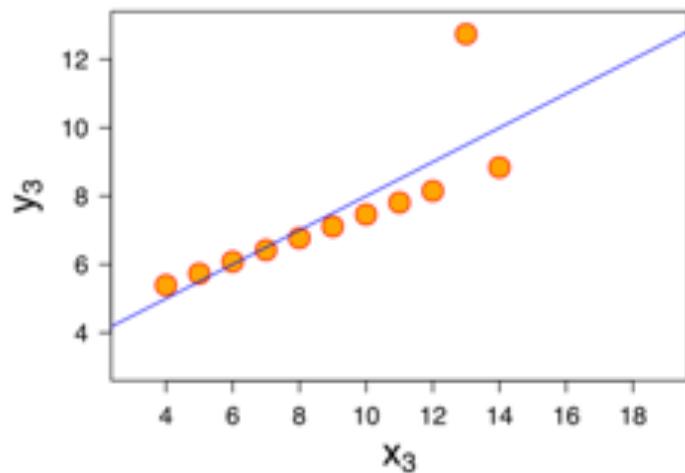
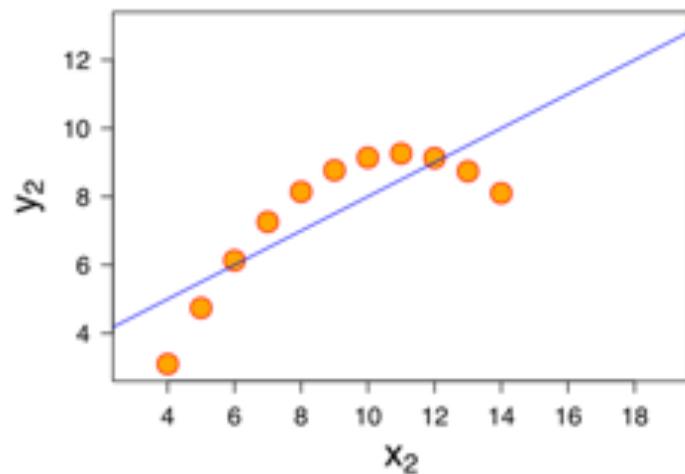
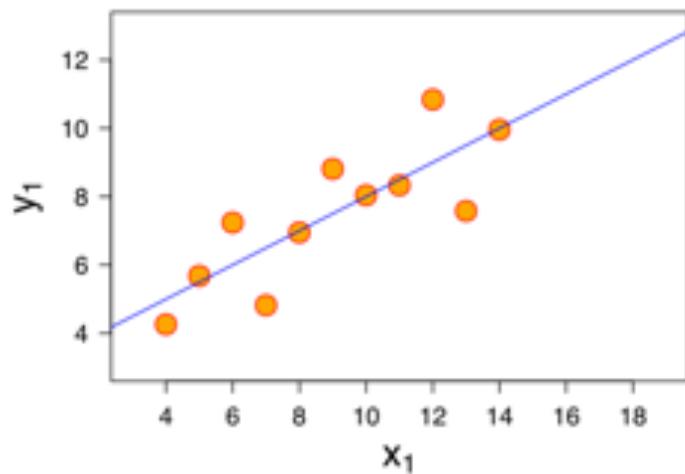


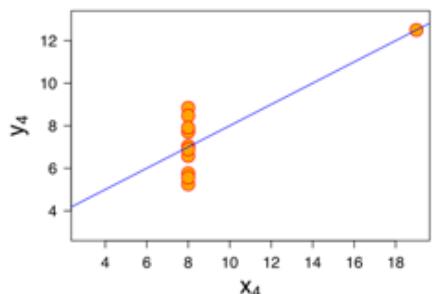
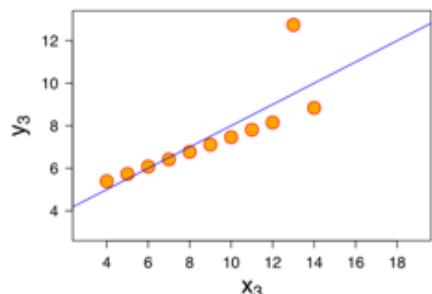
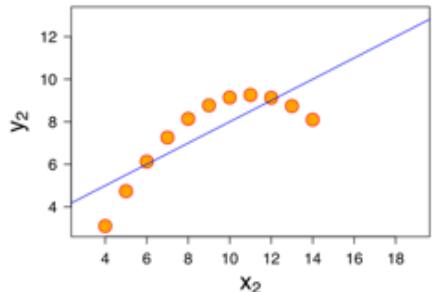
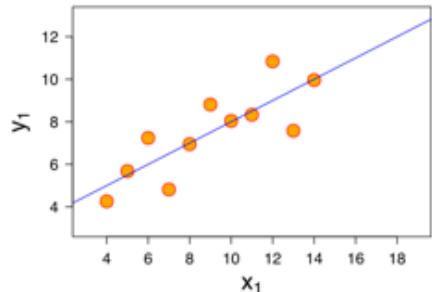
Table - StateData()

	College Degree %	Per Capita Income	Load	Snap	Population	Median Income
State	College Degree %	Per Capita Income				
Alabama	20.6%	11486				
Alaska	30.3%	17610				
Arizona	27.1%	13461				
Arkansas	17.0%	10520				
California	31.3%	16409				
Colorado	33.9%	14821				
Connecticut	33.8%	20189				
Delaware	27.9%	15854				
District of Columbia	36.4%	18881				
Florida	24.9%	14698				
Georgia	24.3%	13631				
Hawaii	31.2%	15770				
Idaho	25.2%	11457				
Illinois	26.8%	15201				
Indiana	20.9%	13149				
Iowa	24.5%	12422				
Kansas	26.5%	13300				
Kentucky	17.7%	11153				
Louisiana	19.4%	10635				
Maine	25.7%	12957				
Maryland	31.7%	17730				
Massachusetts	34.5%	17224				
Michigan	24.1%	14154				
Minnesota	30.4%	14389				
Mississippi	19.9%				9648	
Missouri	22.3%				12989	
Montana	25.4%				11213	
Nebraska	26.0%				12452	
Nevada	21.5%				15214	
New Hampshire	32.4%				15959	
New Jersey	30.1%				18714	
New Mexico	25.5%				11246	
New York	29.6%				16501	
North Carolina	24.2%				12885	
North Dakota	28.1%				11051	
Ohio	22.3%				13461	
Oklahoma	22.8%				11893	
Oregon	27.5%				13418	
Pennsylvania	23.2%				14068	
Rhode Island	27.5%				14981	
South Carolina	23.0%				11897	
South Dakota	24.6%				10661	
Tennessee	20.1%				12255	
Texas	25.5%				12904	
Utah	30.0%				11029	
Vermont	31.5%				13527	
Virginia	30.0%				15713	
Washington	30.9%				14923	
West Virginia	16.1%				10520	
Wisconsin	24.9%				13276	
Wyoming	25.7%				42311	

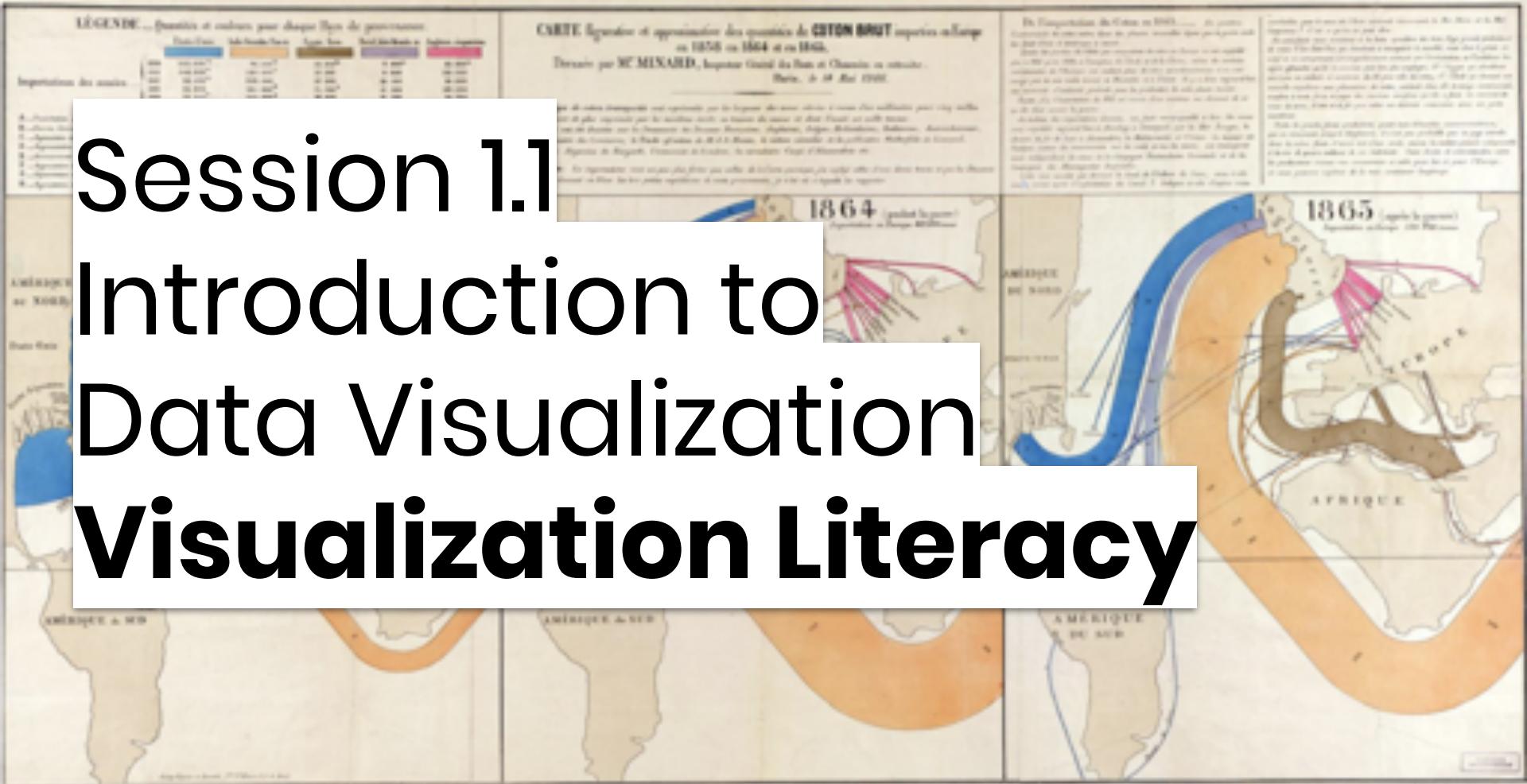
Anscombe's Quartet



Anscombe's Quartet



Property	Value
Mean of x	9
Sample variance of x	11
Mean of y	7.50
Sample variance of y	4.125
Correlation between x and y	0.816
Linear regression line	$y = 3.00 + 0.500x$
Coefficient of determination of the linear regression	0.67



Session 1.1

Introduction to Data Visualization

Visualization Literacy



Online Course

Data Visualization for Professionals

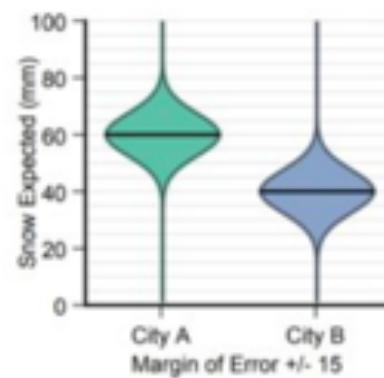
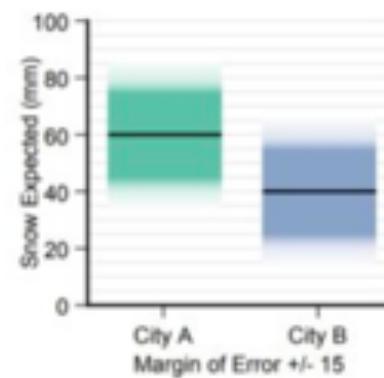
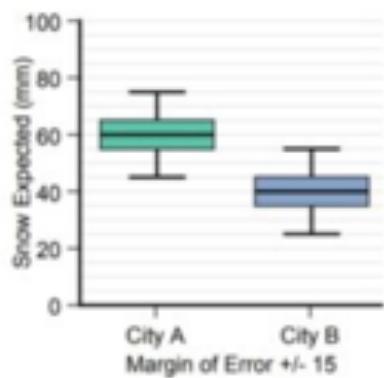
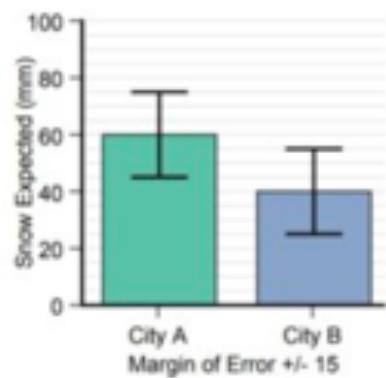
THE UNIVERSITY
of EDINBURGH

Benjamin Bach

June 2020

-- Not for external use --

What are the differences?





Boxplot

False-Friends

Boxplots



No outliers

No horizontal stroke

No horizontal stroke
inside the box

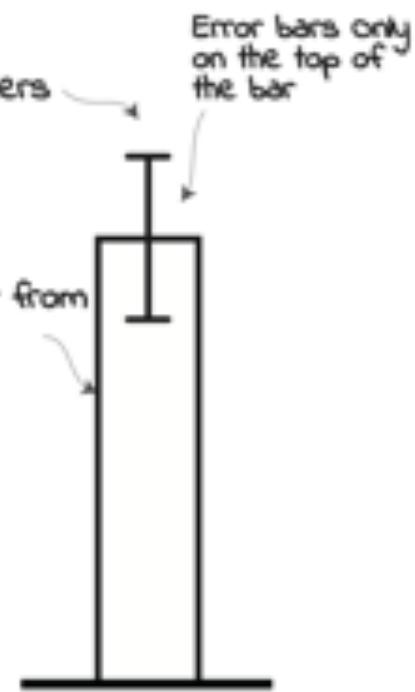
No horizontal stroke

A candlestick represents the price activity of an asset during a specified timeframe through the use of four main components: the open, close, high and low.

Candlesticks



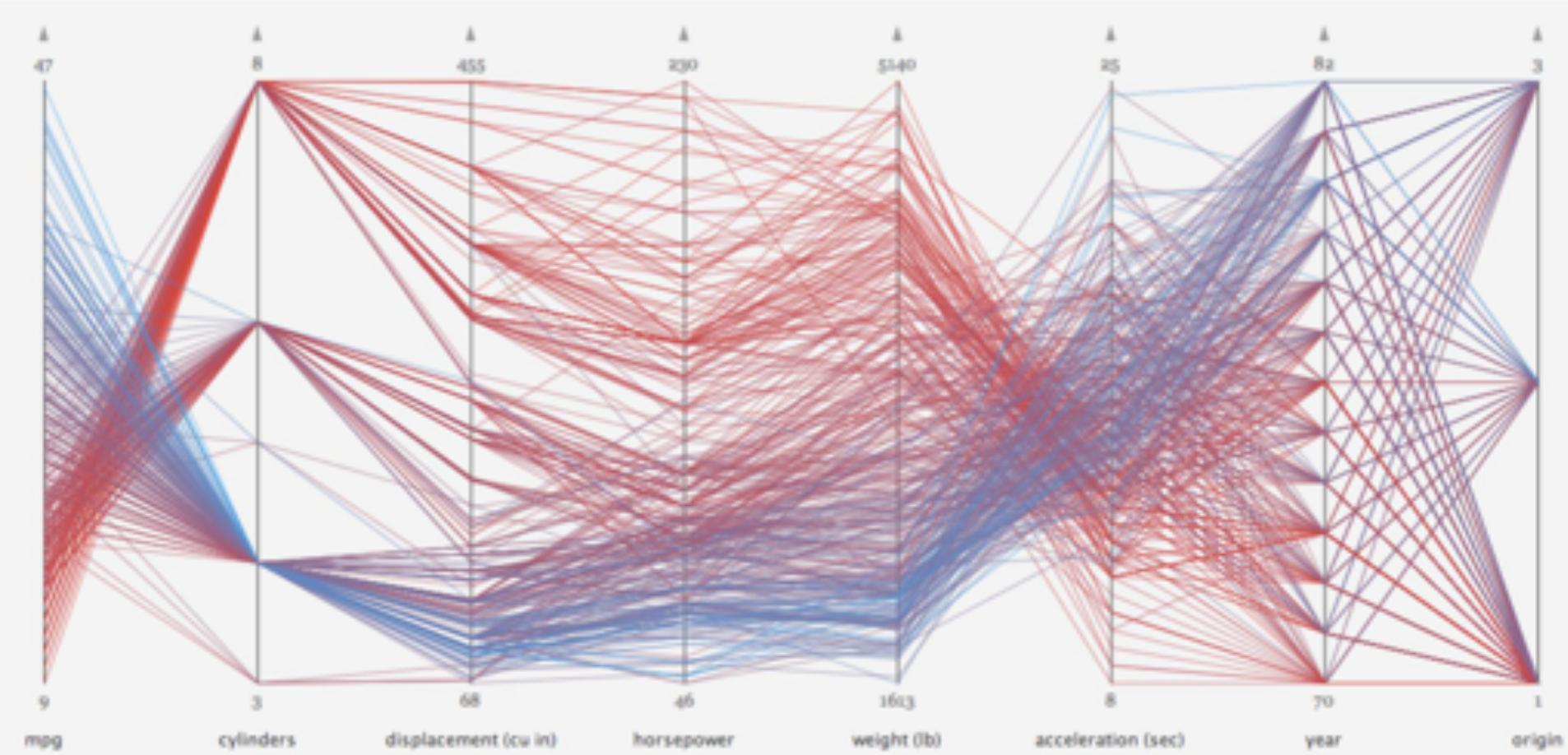
Error bars



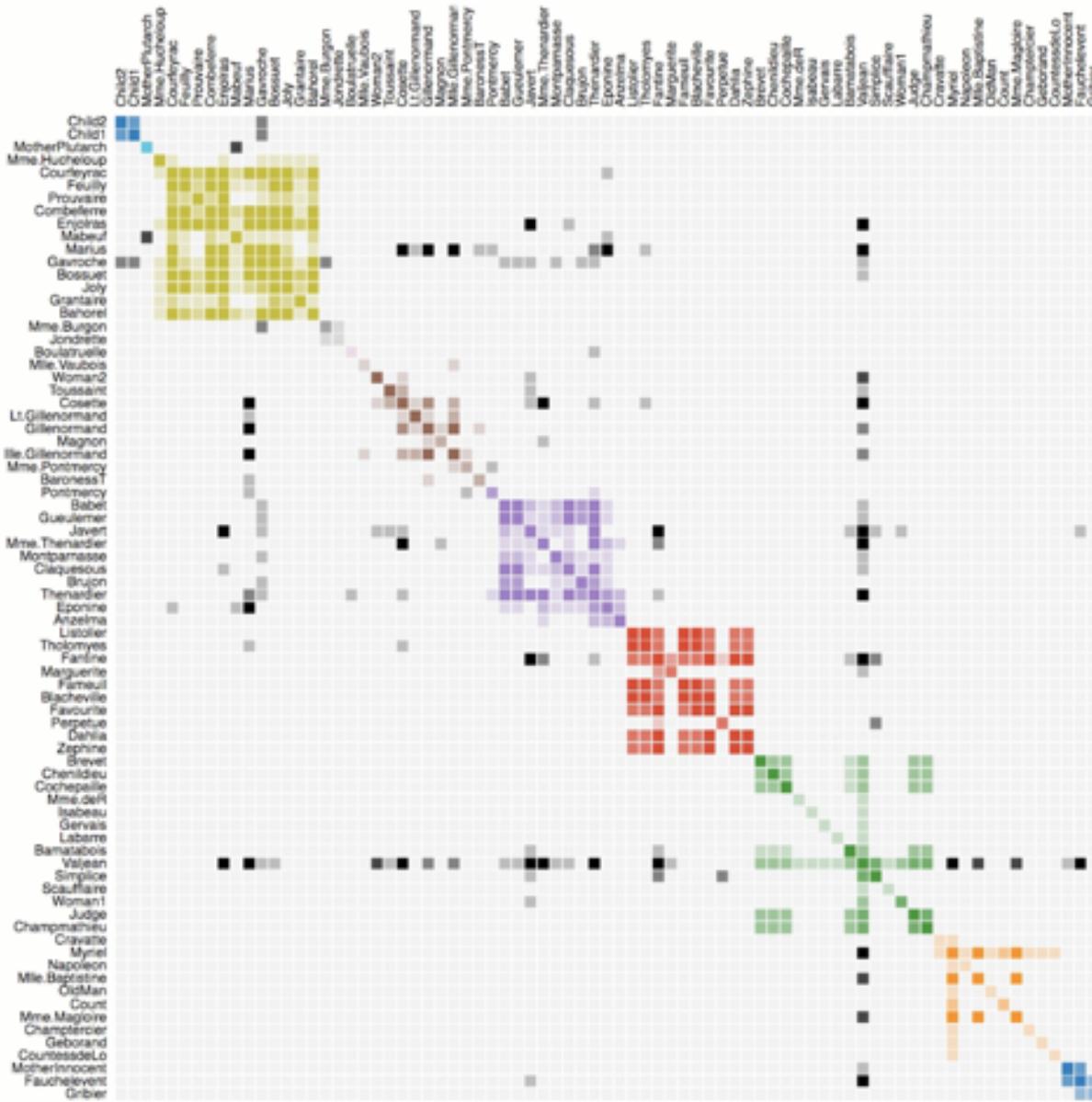
Error bars are graphical representations of the variability of data and used on graphs to indicate the error or uncertainty in a reported measurement.



Which visualization is this?



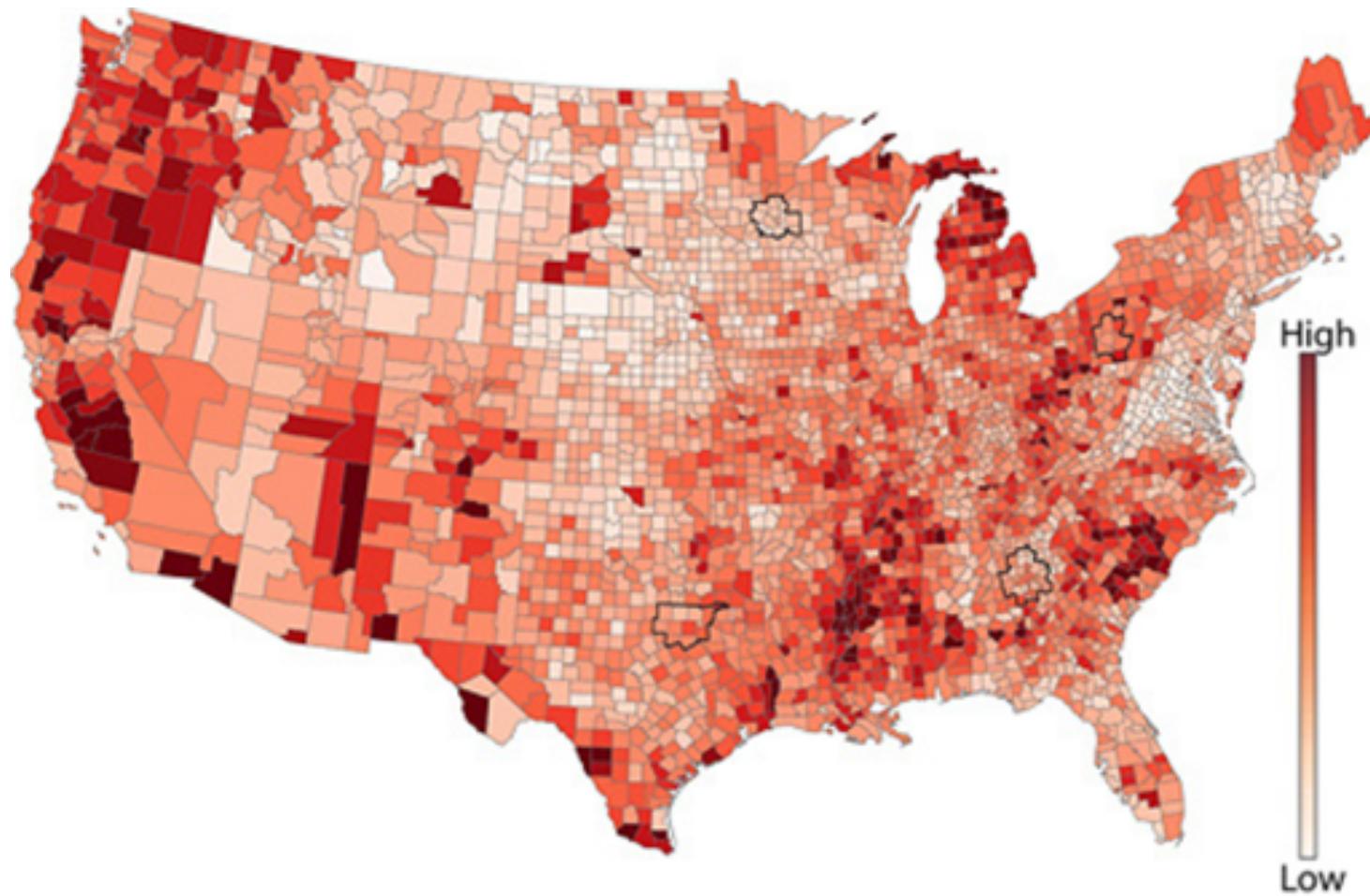
Which visualization is this?



What is the problem?

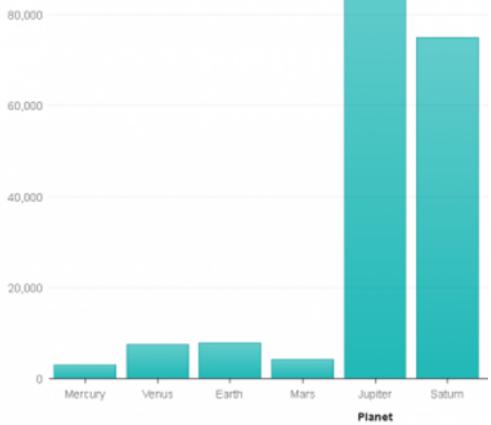


What is the problem?

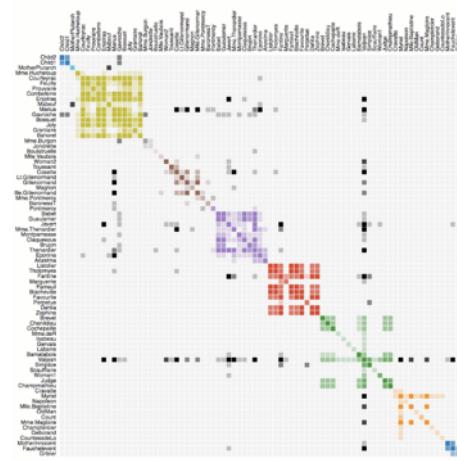


Simplicity vs. expressiveness

Simple



Complex



Limited

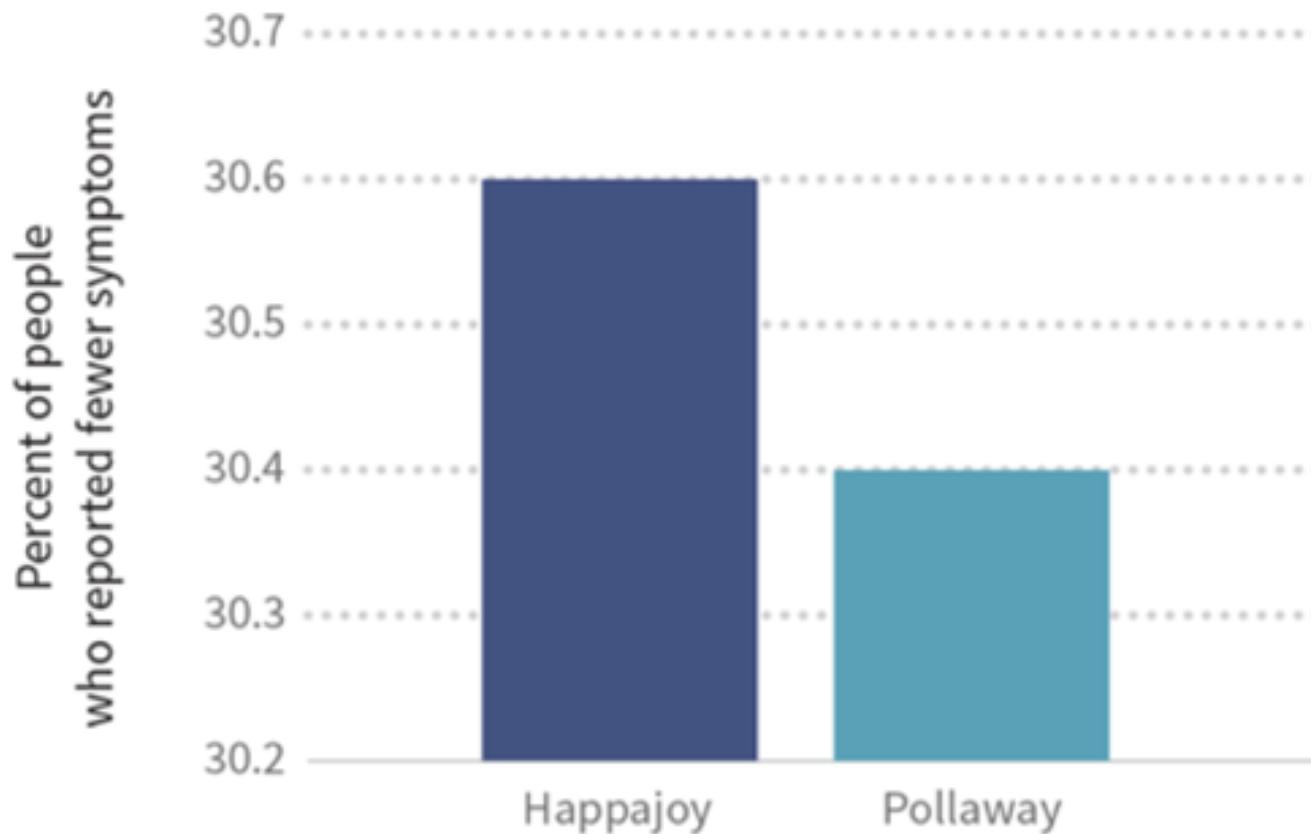
- Mono-dimensional data
- A "few numbers"
- quantities

Expressive

- Multi-dimensional data
- "Many numbers"
- relationships

How do you decide?

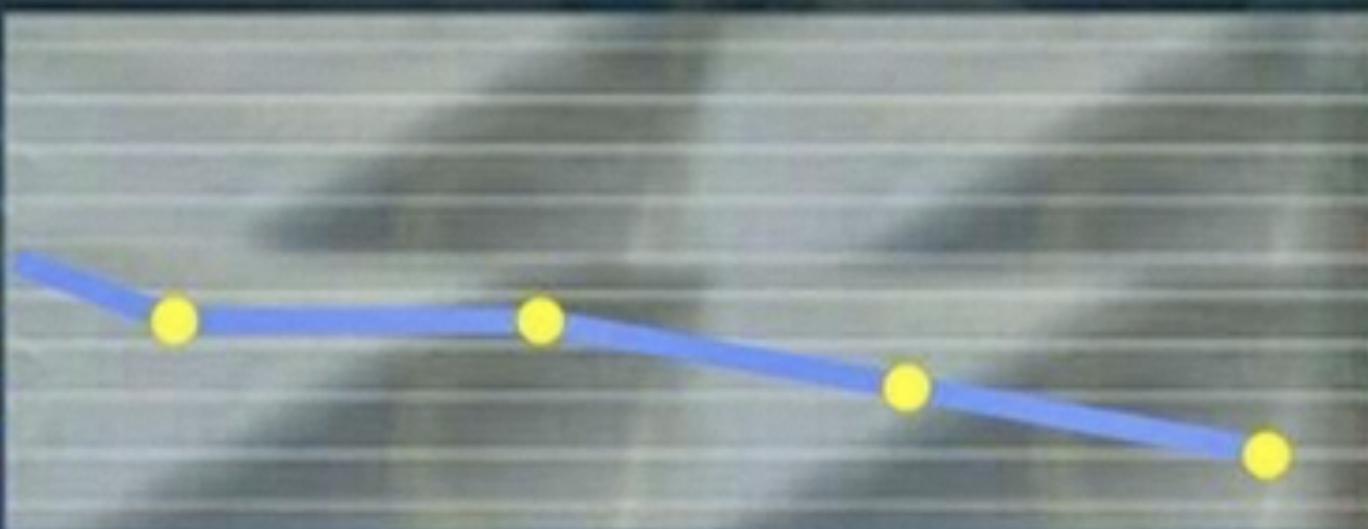
Effectiveness of Allergy Medicines



What's the take home message?

AFFORDABLE CARE ACT ENROLLMENT

4,000,000
5,000,000
6,000,000
7,000,000
8,000,000



FOX NEWS channel

As of
March 27th

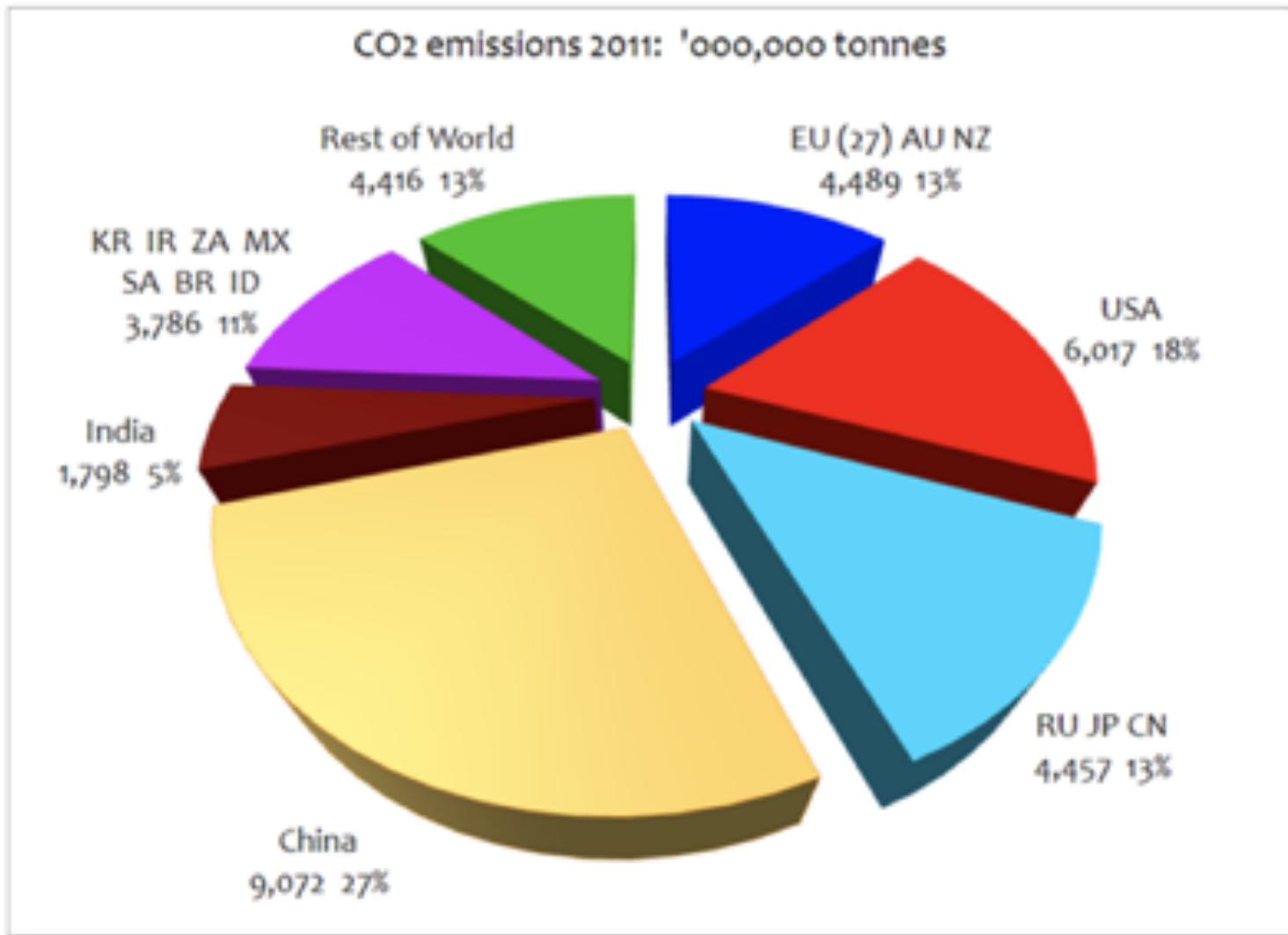
CBO
revised
estimate

CBO
original
estimate

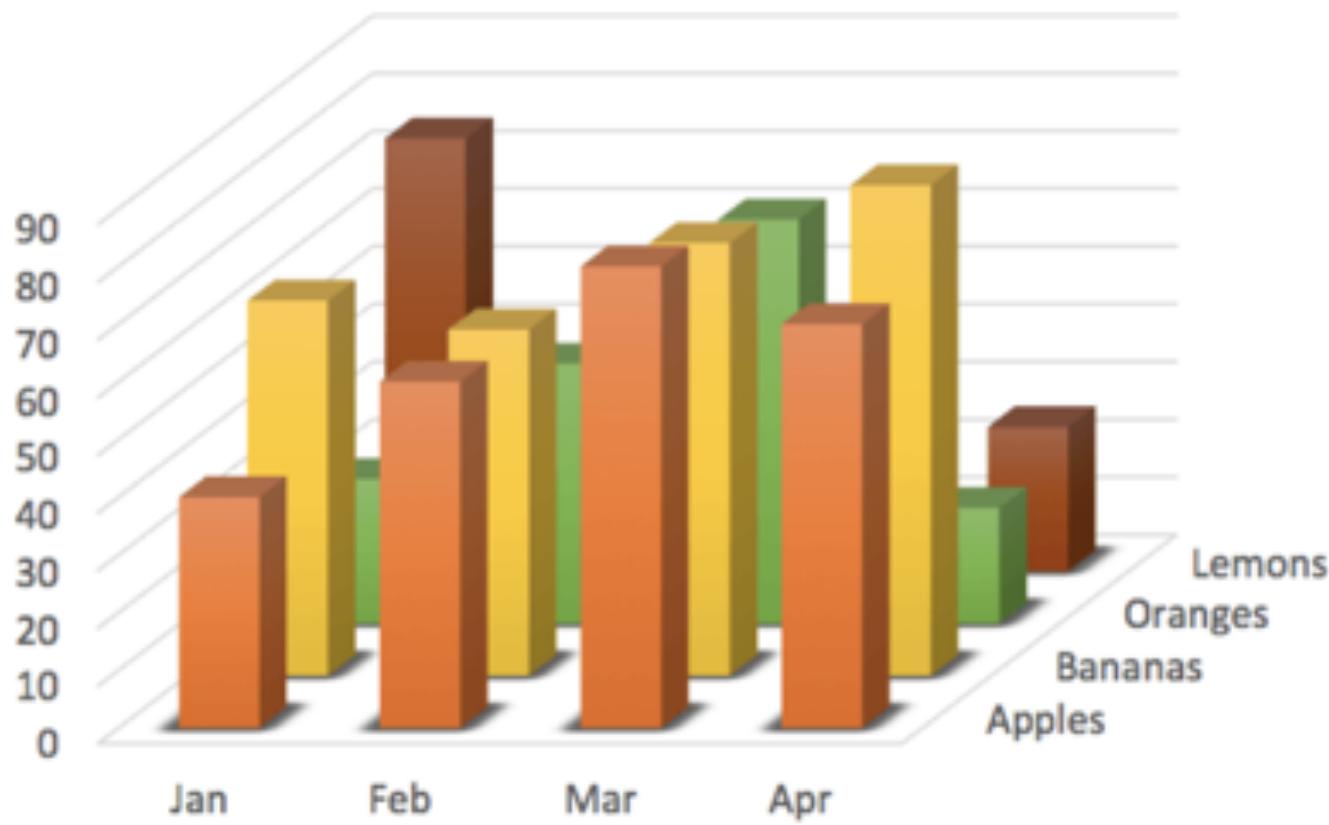
Enrollment
as of
Today

IE IN SECOND PLACE WITH \$26.5 MIL, WHILE "MUPP DOW FUT 16,325.00

Which slice is larder—cyan or green?



How many lemons?



How to decode circles?

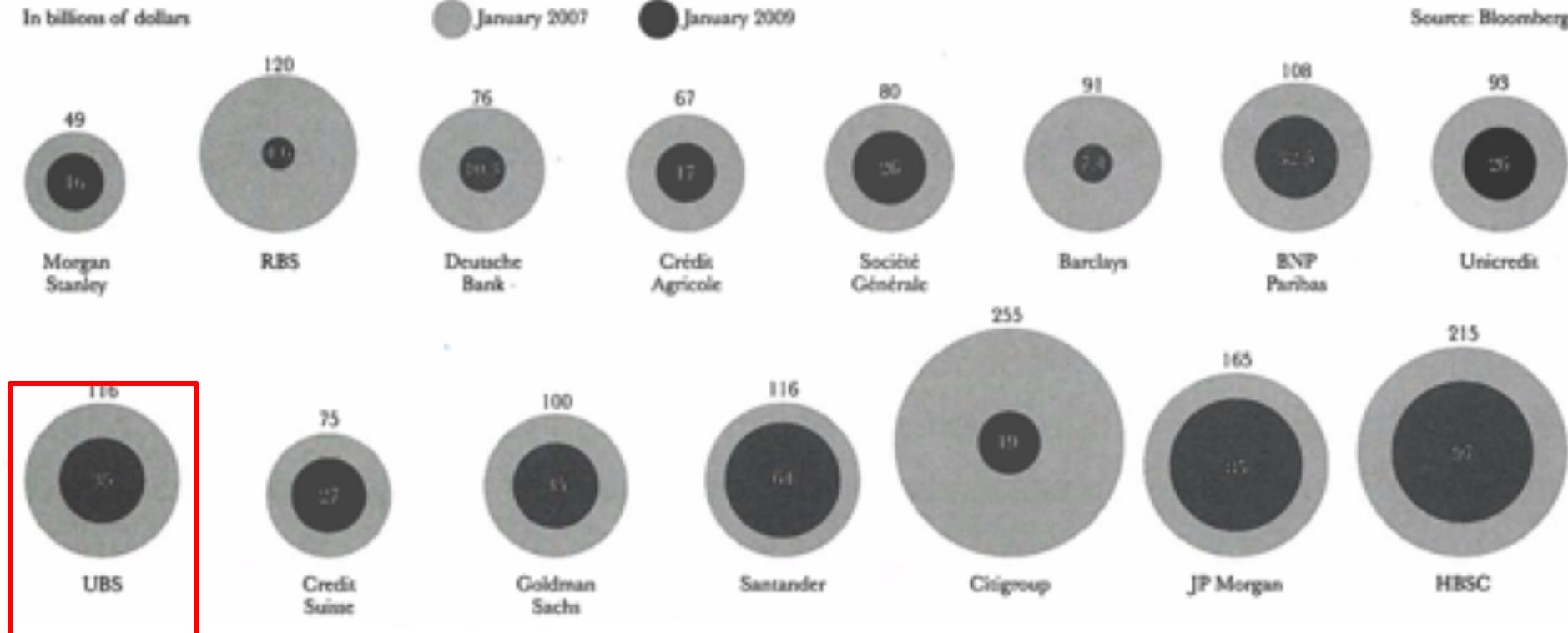
Market Capitalization of the World's Biggest Banks

In billions of dollars

January 2007

January 2009

Source: Bloomberg



a)



b)



c)



New Literacies

Data Literacy:

Understand a wide variety of tools for ***accessing, converting and manipulating data.***

Visual literacy:

Ability to read, evaluate, use, and create ***images***

Graphicacy:

Ability to understand and ***present information*** in the form of sketches, photographs, diagrams, maps, plans, charts, graphs and other non-textual, two-dimensional formats

Visualization Literacy

- the ability to **confidently use** a given data visualization,
- to translate questions specified in the data domain into **visual queries** in the visual domain,
- **interpreting visual patterns** in the visual domain as properties in the data domain

Visualization Literacy

Reading:

- Correctly decode (simple & complex) visual representations
- Know pitfalls and deceptions
- Think critically 'beyond' and see 'through' the visualization

Design:

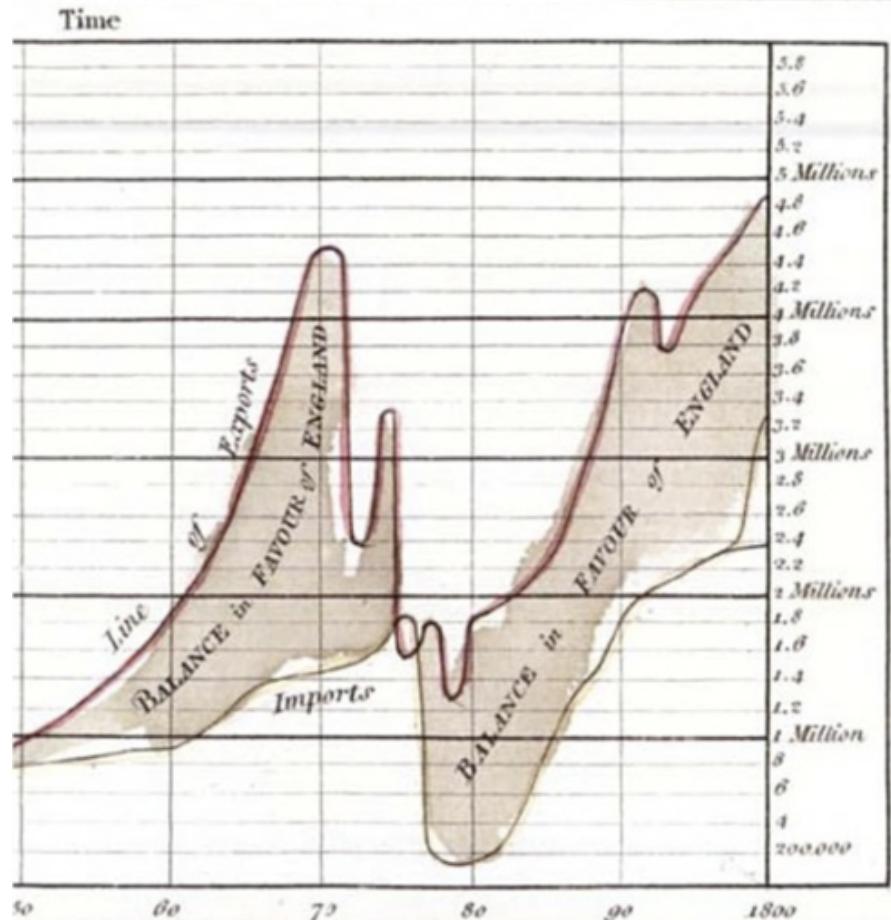
- Create efficient and effective visualizations
- Design efficient and effective visualizations

Explore:

- perform tasks: ask and answer questions
- Interact with visualizations

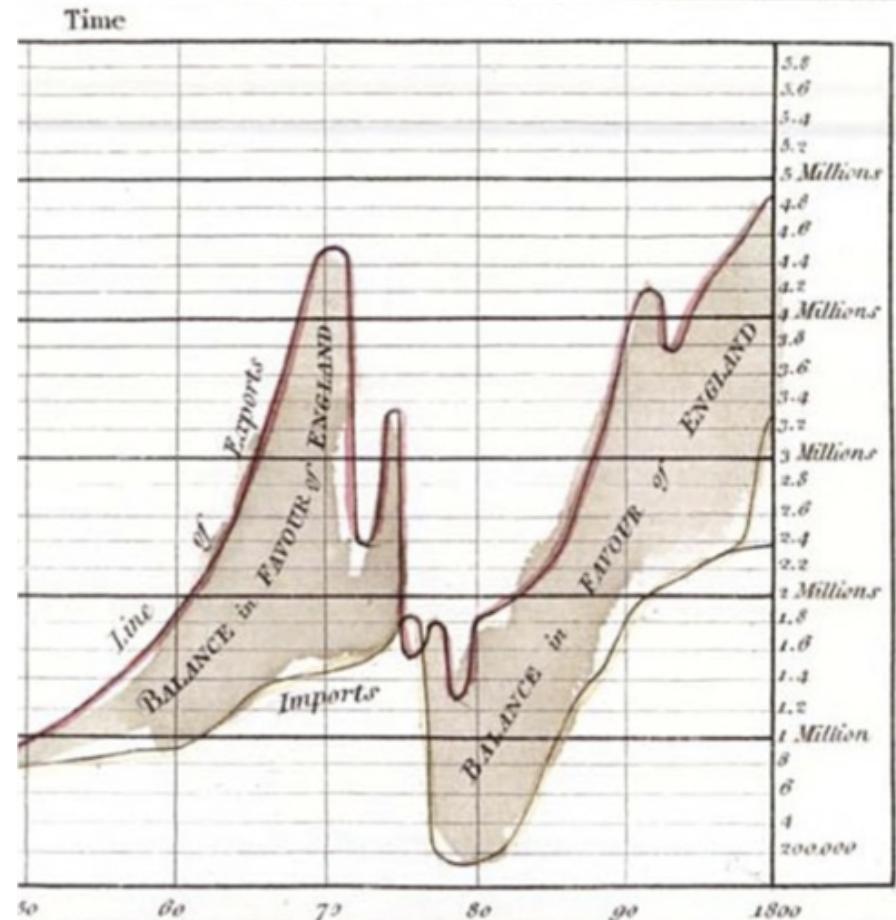
Low-level tasks

- Largest / smallest ...
- How many ...
- Which elements ...
- Where ...
- When ...
- What ...
- ...



High-level tasks

- Which trends
 - Which outliers
 - Correlations
 - Anomalies
 - Describe
 - Compare
 - Group
 - ...



CHARTE figurative et approximative des quantités de COTON BRUT importé en Angleterre en 1852-3, en 1854 et en 1855.

Dessin par MC MINAHD, inspecteur général des Bains et Chaussées au ministère
des Travaux Publics.

Session 1.1



Introduction to Data Visualization

Wrapping up





Online Course **Data Visualization for Professionals**

THE UNIVERSITY
of EDINBURGH

Benjamin Bach

June 2020

-- Not for external use --

Wrap-up

1. Visualizations come in **many forms**
2. Visualizations have **many goals**
3. **Visualizations are**
 - *external representations*
 - *Visually showing patterns in data*
 - *to amplify cognition*
 - *and support user tasks*
1. **Visualization literacy** is the ability to
 - *Correctly interpret a visual representation*
 - *to map domain tasks to visualization tasks*
1. Visualizations can support a range of **tasks at different levels**