

Questions about yesterday?

2

Tasks and Scenarios

Motivation

- Visualization as a usability problem
- Improve effectiveness of visualization
 - Task support,
 - perception,
 - interaction,
 - knowledge, ...
- Chose or build effective visualizations

Structuring and Processing Data

*External, raw,
noisy,
unstructured,
meaningless*

Data

Collection
Formatting
Cleaning
Mining
Exploration
Analysis

Visualization
+
Interaction

*Internal,
meaningful,
purposeful,
structured,
qualitative*

Insight + Knowledge

Understanding

Scenarios ~ Story

Describes:

- People
- Places
- Actions
- Artifacts

Does

- Explain a visualization use
- Identify challenges
- Support design process

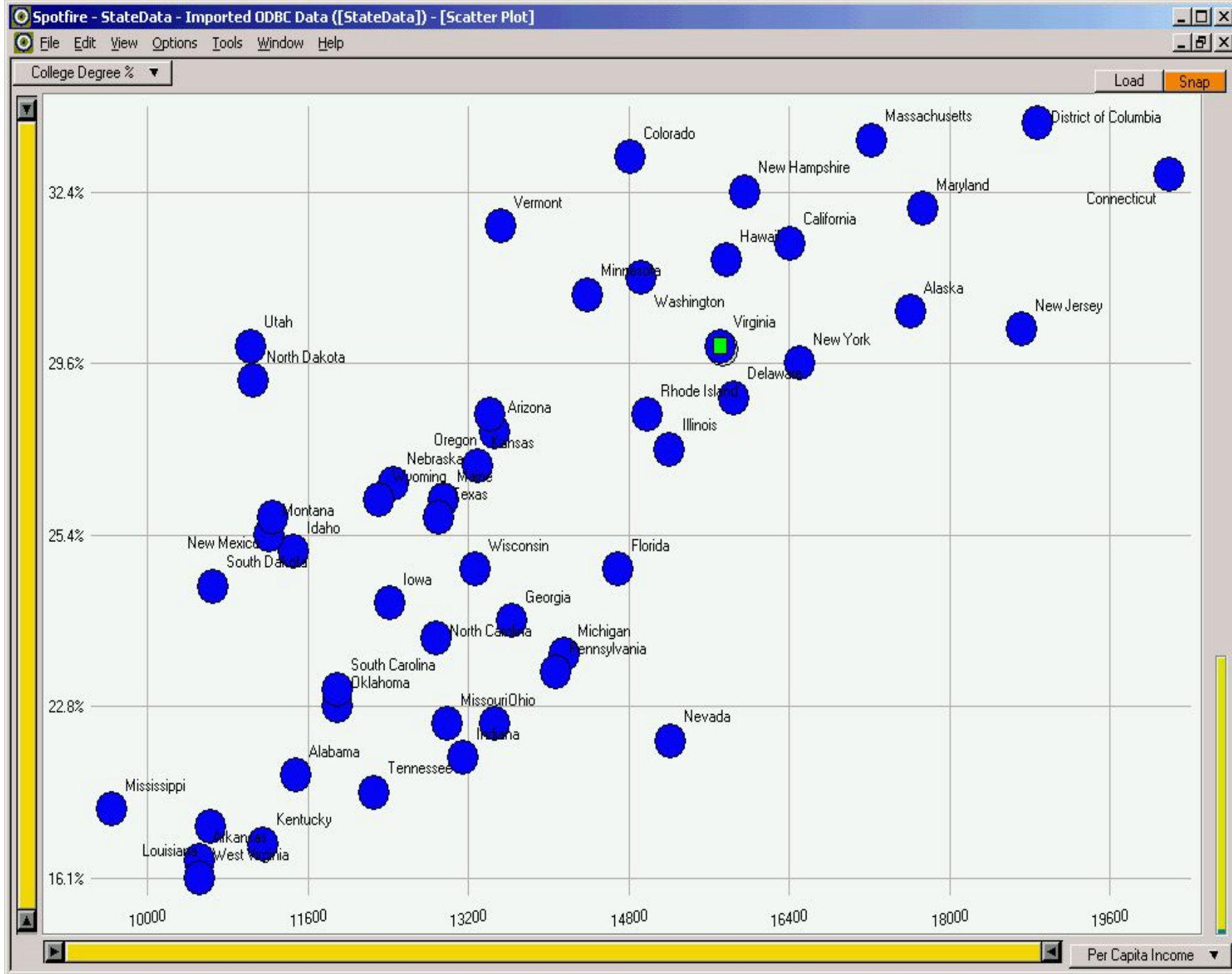
Data Challenge

1. Data
2. Data structures
3. Context
4. Audience
5. Questions / Tasks
6. Messages

AFFORDABLE CARE ACT ENROLLMENT



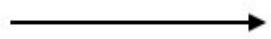
IE IN SECOND PLACE WITH \$26.5 MIL, WHILE "MUPP DOW FUT 16,325.00



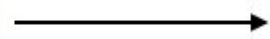
Exploratory

Explanatory

Data



Insight



Message

Exploratory

Explanatory



Data Centered

Domain Experts

Generating Insights

Lab Setting

Exploratory

Explanatory



Data Centered

Human Centered

Domain Experts

Non-expert Audience

Generating Insights

Conveying Messages

Lab Setting

In-the-wild

2.1

Data

Data

What is my data about?

- People, places, collection, context, certainty, storage..

2.2

Data Structures

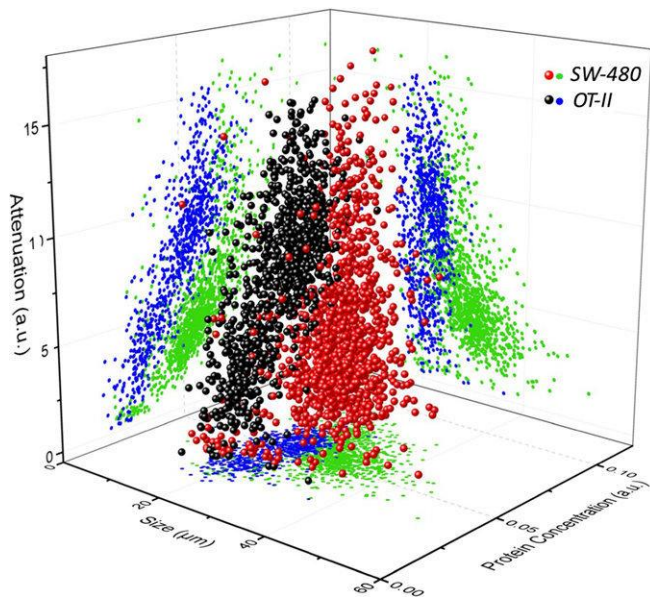
Data Structures: 1D

	A	B
1	Last Name	Sales
2	Smith	\$16,753.00
3	Johnson	\$14,808.00
4	Williams	\$10,644.00
5	Jones	\$1,390.00
6	Brown	\$4,865.00
7	Williams	\$12,438.00
8	Johnson	\$9,339.00
9	Smith	\$18,919.00
10	Jones	\$9,213.00
11	Jones	\$7,433.00
12	Brown	\$3,255.00
13	Williams	\$14,867.00
14	Williams	\$19,302.00
15	Smith	\$9,698.00
16		

Data Structures: 2D

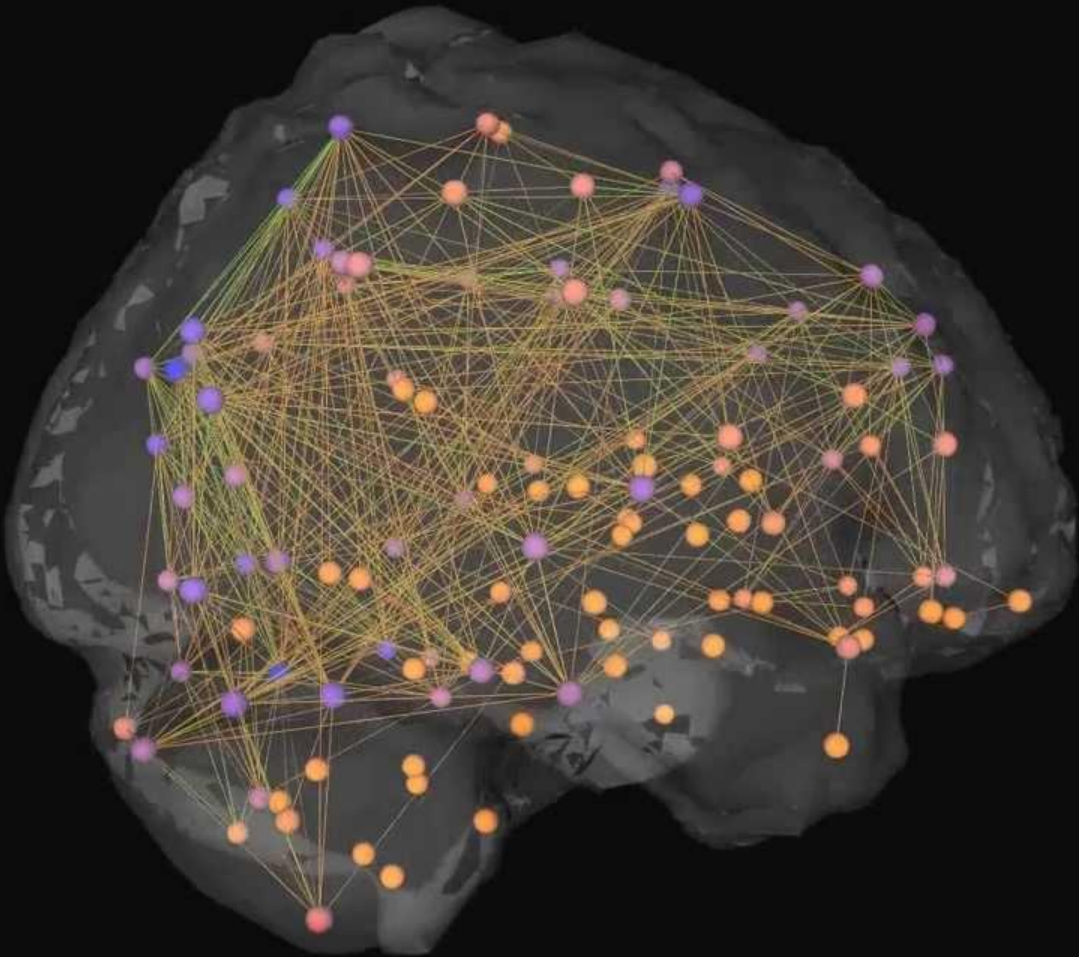
	A	B	C
1	Last Name	Sales	Country
2	Smith	\$16,753.00	UK
3	Johnson	\$14,808.00	USA
4	Williams	\$10,644.00	UK
5	Jones	\$1,390.00	USA
6	Brown	\$4,865.00	USA
7	Williams	\$12,438.00	UK
8	Johnson	\$9,339.00	UK
9	Smith	\$18,919.00	USA
10	Jones	\$9,213.00	USA
11	Jones	\$7,433.00	UK
12	Brown	\$3,255.00	USA
13	Williams	\$14,867.00	USA
14	Williams	\$19,302.00	UK
15	Smith	\$9,698.00	USA
16			

Data Structures: 3D

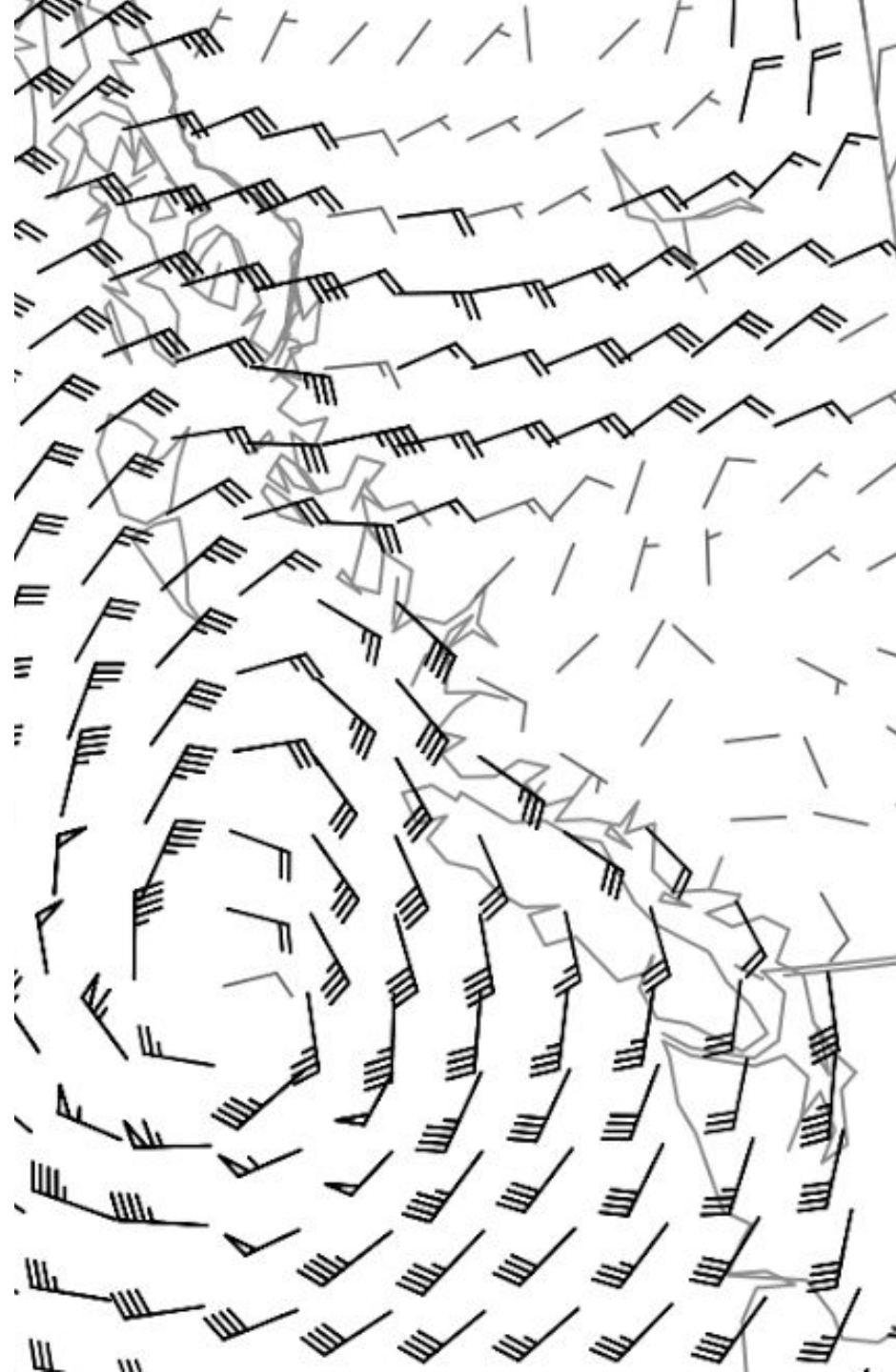


	A	B	C	D
1	Last Name	Sales	Country	Quarter
2	Smith	\$16,753.00	UK	Qtr 3
3	Johnson	\$14,808.00	USA	Qtr 4
4	Williams	\$10,644.00	UK	Qtr 2
5	Jones	\$1,390.00	USA	Qtr 3
6	Brown	\$4,865.00	USA	Qtr 4
7	Williams	\$12,438.00	UK	Qtr 1
8	Johnson	\$9,339.00	UK	Qtr 2
9	Smith	\$18,919.00	USA	Qtr 3
10	Jones	\$9,213.00	USA	Qtr 4
11	Jones	\$7,433.00	UK	Qtr 1
12	Brown	\$3,255.00	USA	Qtr 2
13	Williams	\$14,867.00	USA	Qtr 3
14	Williams	\$19,302.00	UK	Qtr 4
15	Smith	\$9,698.00	USA	Qtr 1
16				

Demo Demo Mode

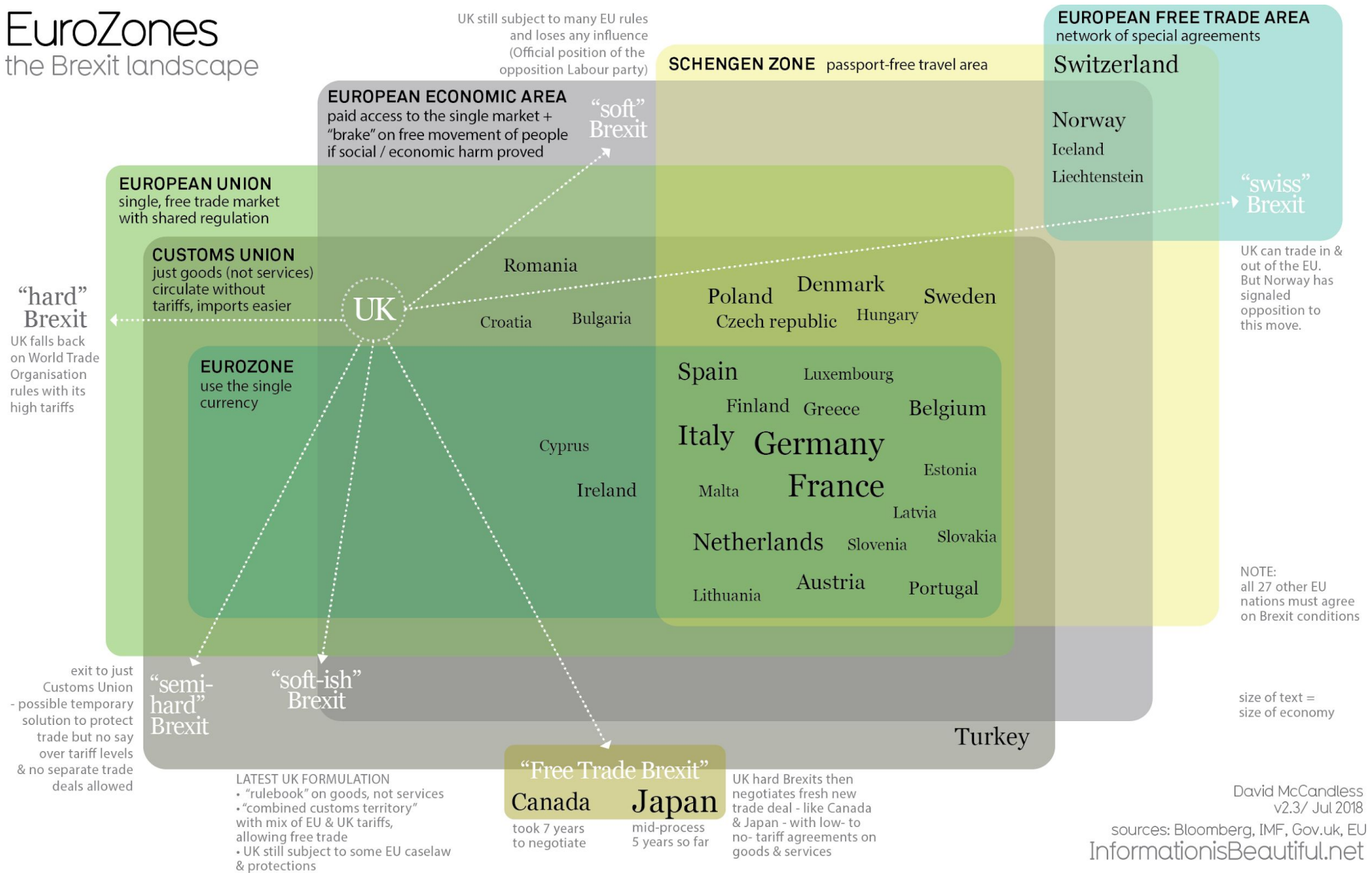


Multidimensional Data

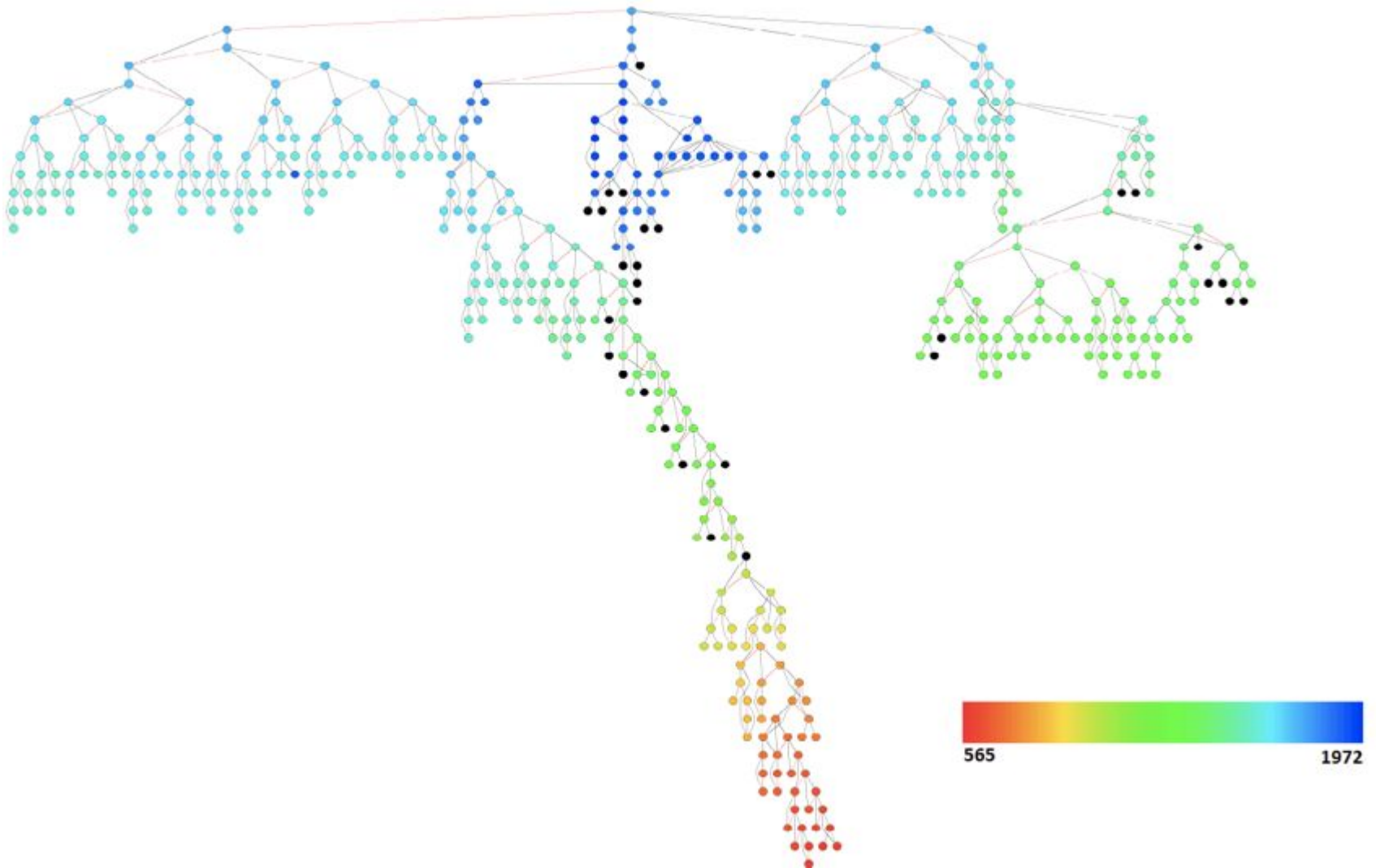


Sets

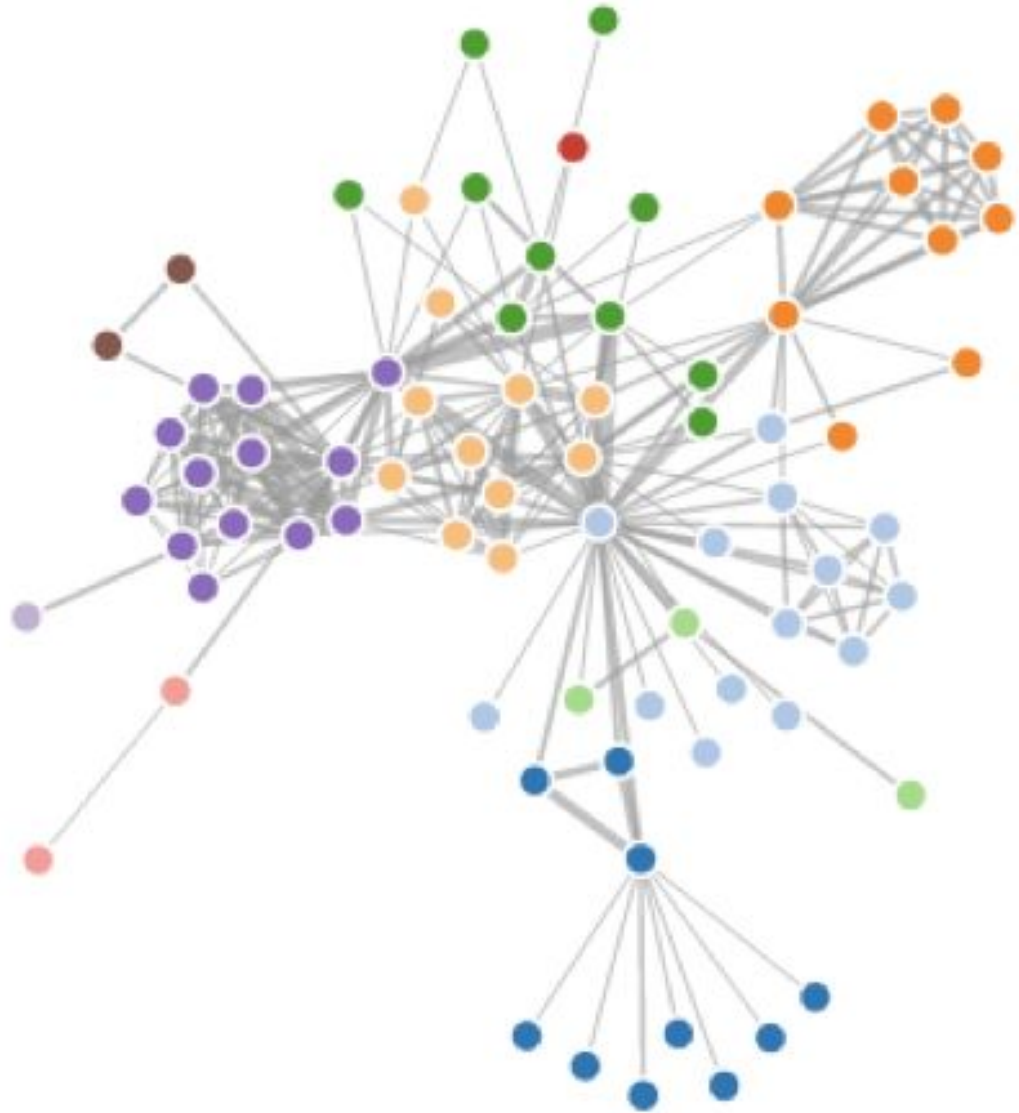
EuroZones the Brexit landscape



Trees



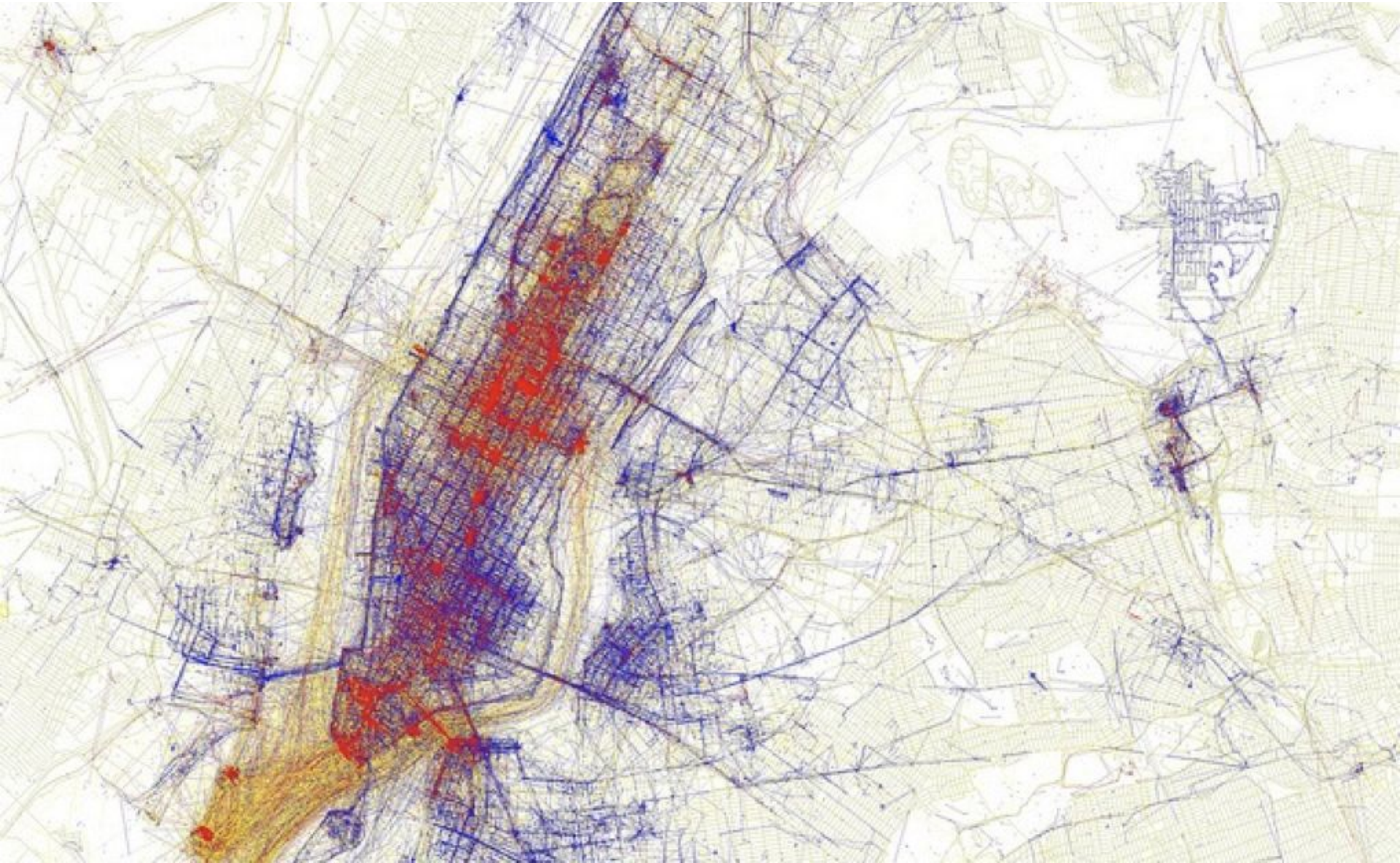
Networks



A heatmap titled 'BIRTHDAY RANK' showing the frequency of birthdays across 31 ranks (rows) and 12 months (columns). The color scale ranges from 'LESS COMMON' (light green) to 'MORE COMMON' (dark green/black). The x-axis is labeled 'DAY' with months JAN through DEC. The y-axis is labeled 'BIRTHDAY RANK' with numbers 1 through 31. The heatmap shows that lower ranks (1-10) generally have higher frequency (darker colors) across most months, with a notable peak in July for ranks 1-10. Higher ranks (20-31) show more variation, with some months being significantly less common (lighter green) than others.

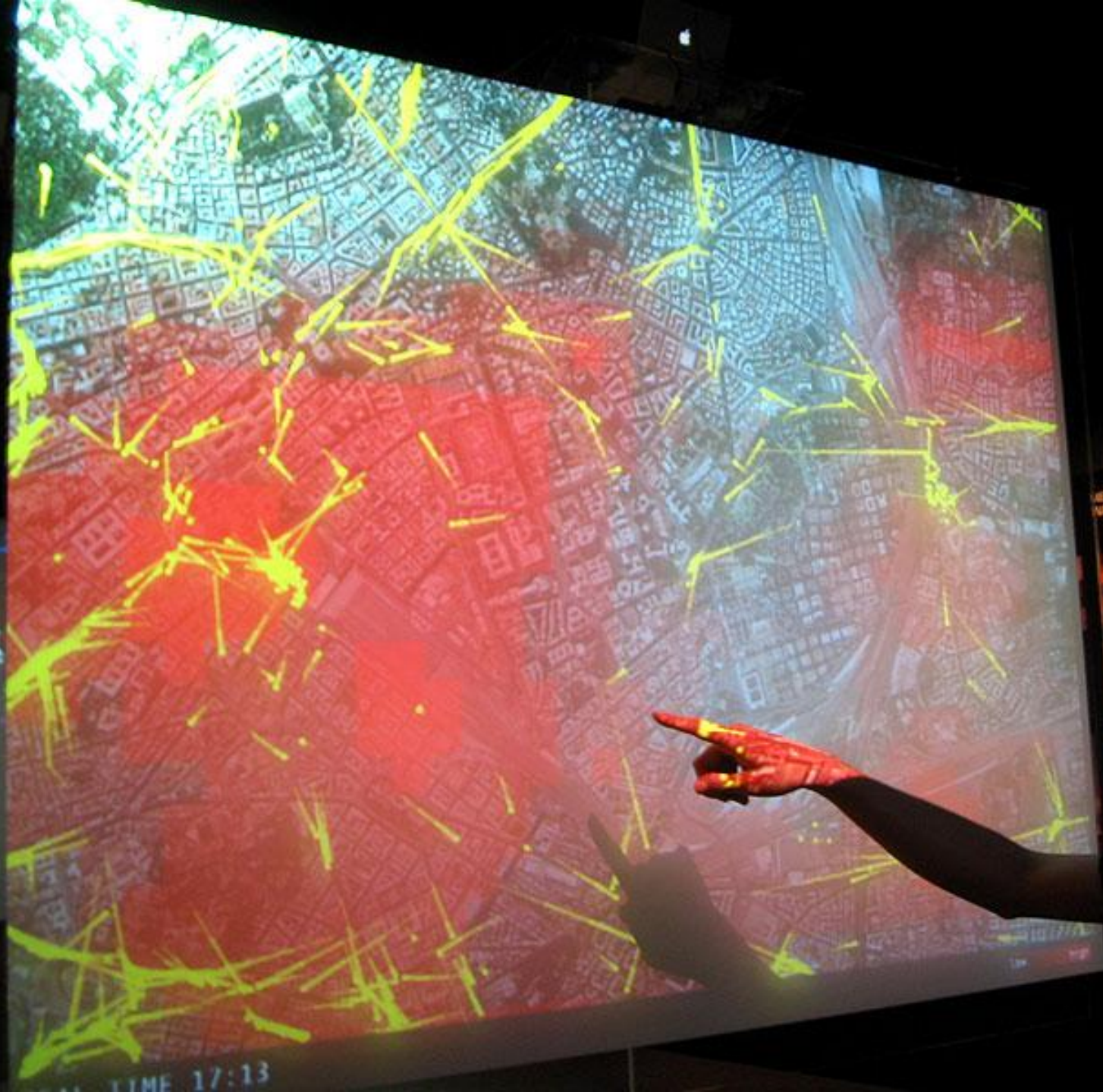
MORE COMMON

Geographical



2.3

Context



MEMORY
WALK

DIGITAL
ANNALS



TIME 17:13









Zählen und Prognostizieren

DEMOGRAPHISCHE SKULPTUR

Die begabte Skulptur erzählt durch die Umrisse der deutschen Bevölkerungspyramiden der Jahrgänge von 1900 bis 2010. Im Verlauf ihrer Jahrgangsscheitel bilden sich historische Ereignisse wie die zwei Weltkriege, das Wirtschaftswunder, aber auch grundlegende Änderungen unseres Lebensstiles durch Erfindungen wie die Anti-Baby-Pille ab. Seit geht es um 100 Jahren ist der Trend gleich: Dass wir immer länger leben und weniger Kinder bekommen als frühere Generationen. Dabei erreicht die absolute Bevölkerungsgröße in Deutschland heute ihren Zenit und wird trotz Zuneigung künftig rückläufig sein. Das Durchschnittsalter und der Anteil der Migranten an der Bevölkerung werden weiter steigen.



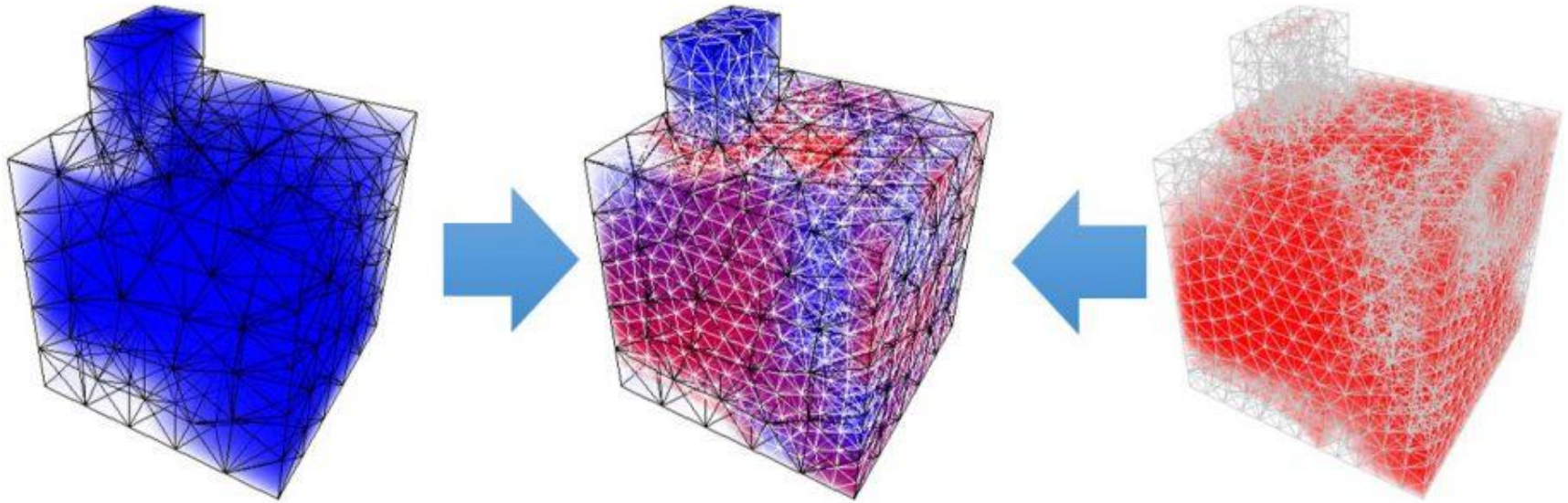


2.4

Audience

Fused visualization for causality discovery

JST A-STEP “Fused Visualization” (2012~2014)



Exploration of correlation between multiple variables

Kun Zhao, Naohisa Sakamoto, Koji Koyamada, Adaptive Fused Visualization for Large-scale Blood Flow Dataset with Particle-based Rendering, Journal of Visualization, doi:10.1007/s12650-014-0260-z, 2014,







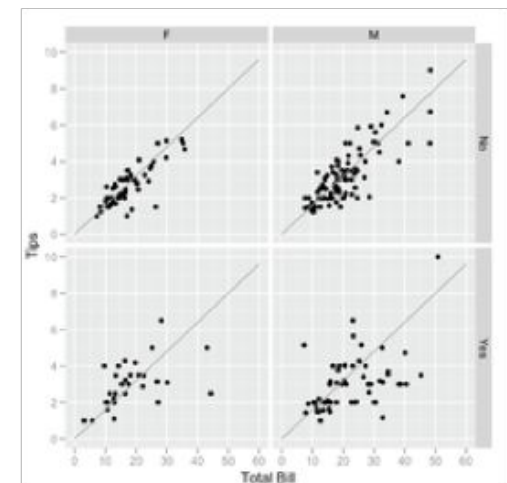
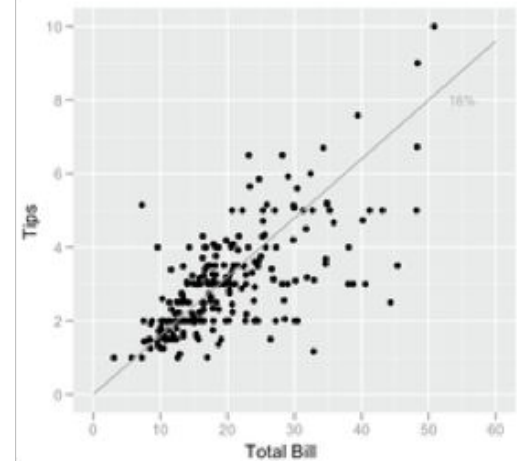
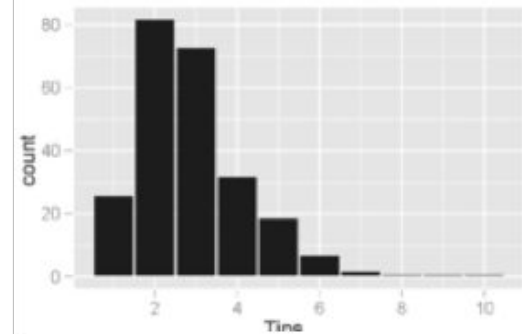
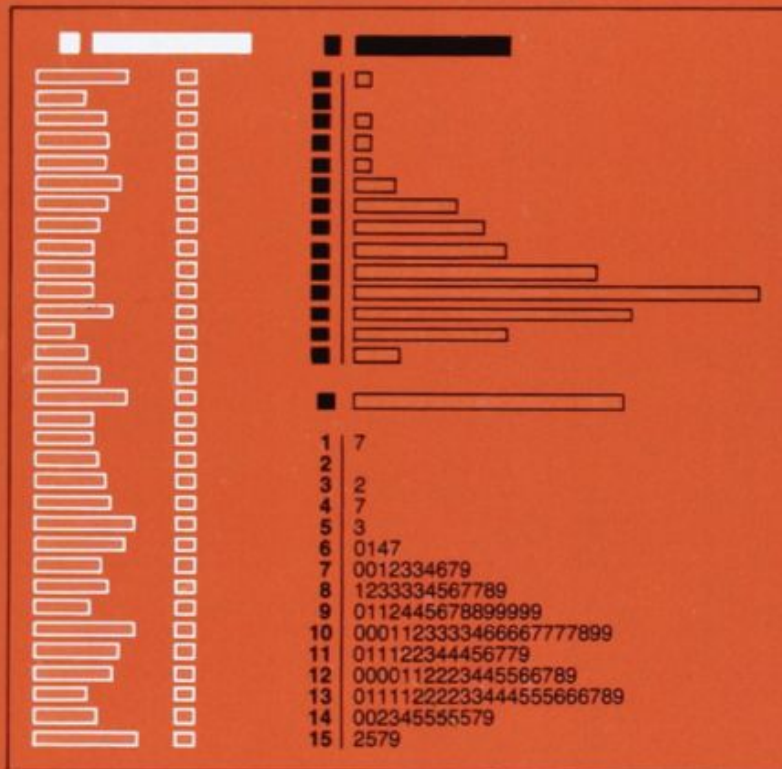


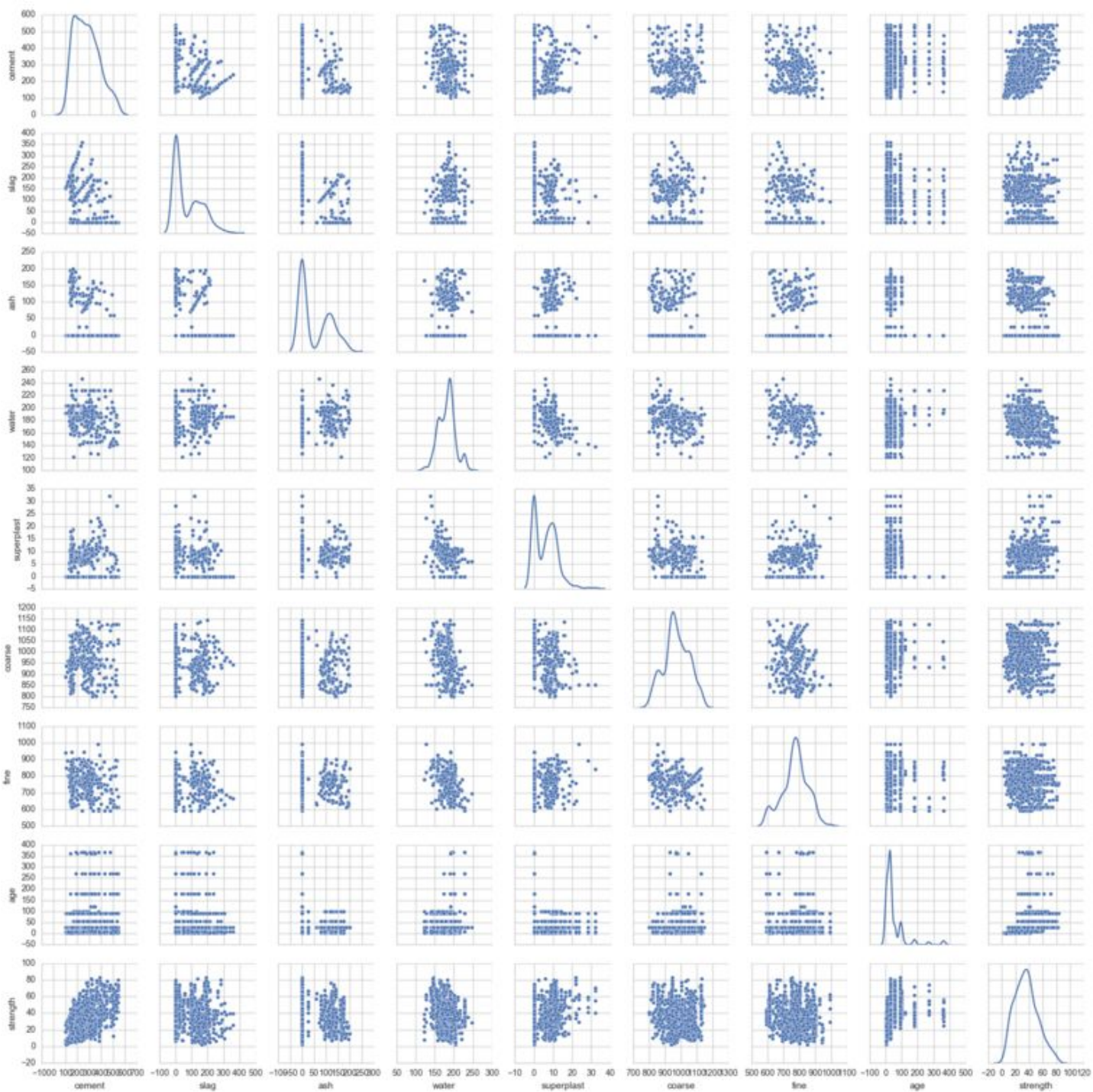
2.5

Tasks and Questions

John W. Tukey

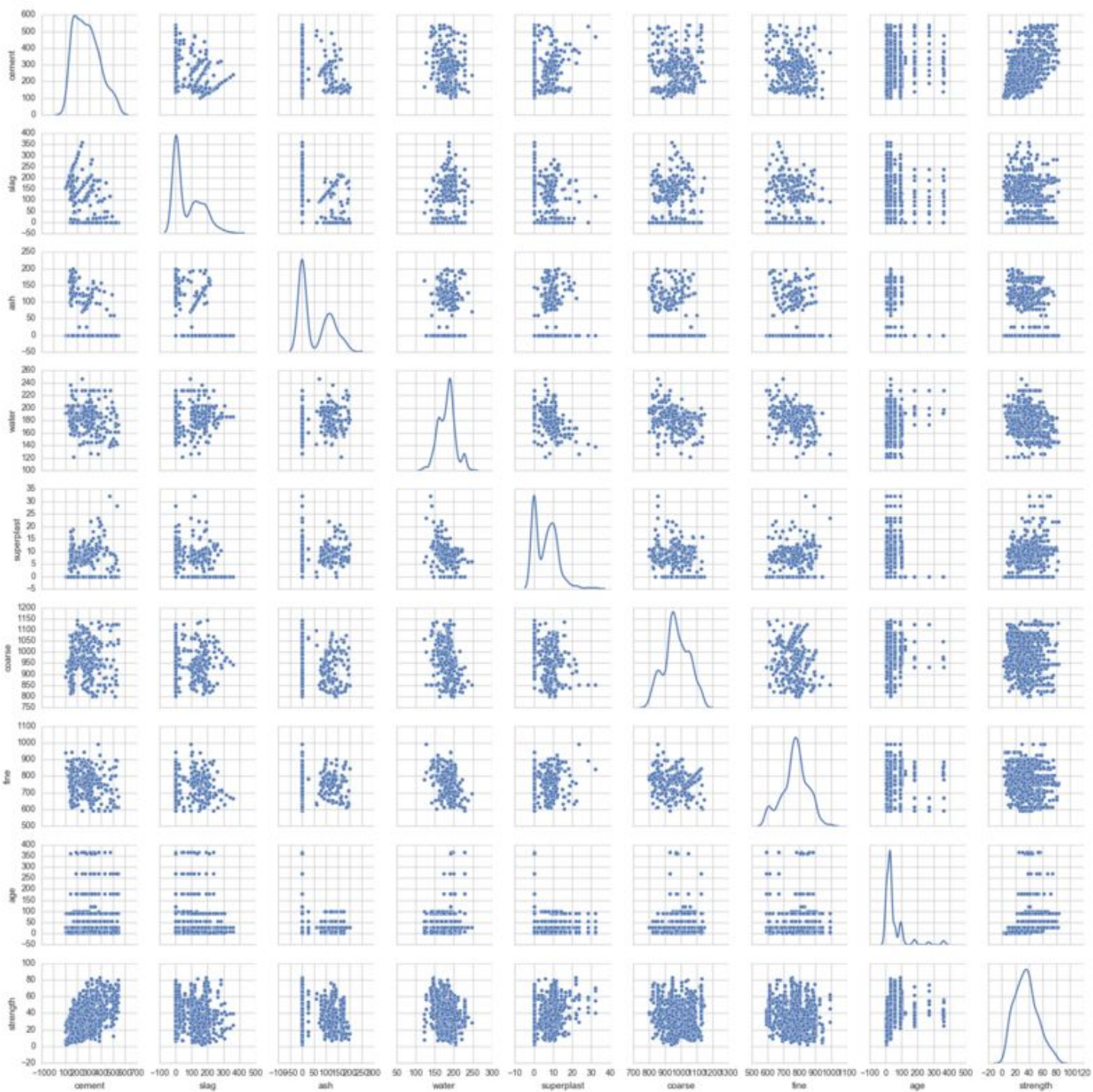
EXPLORATORY DATA ANALYSIS

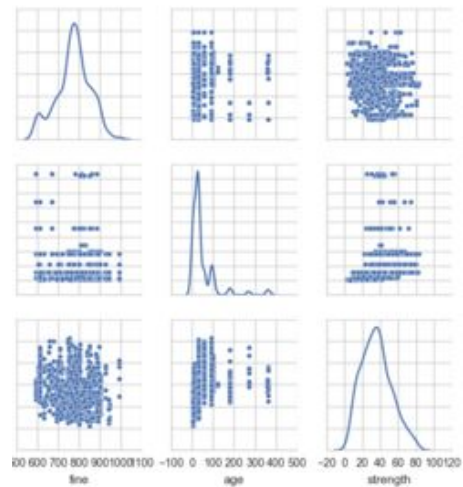
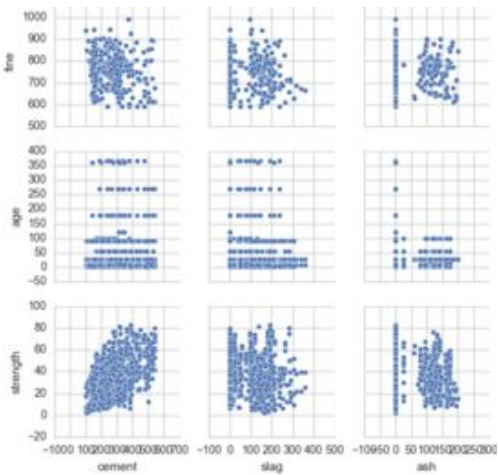
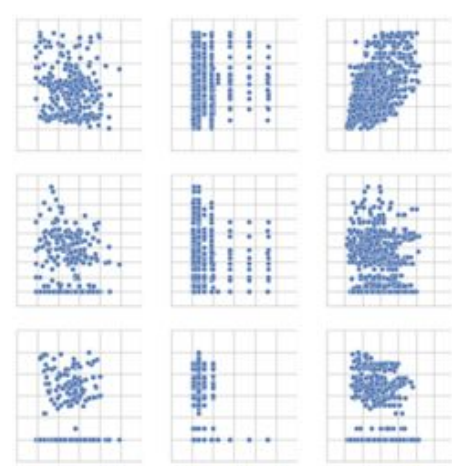
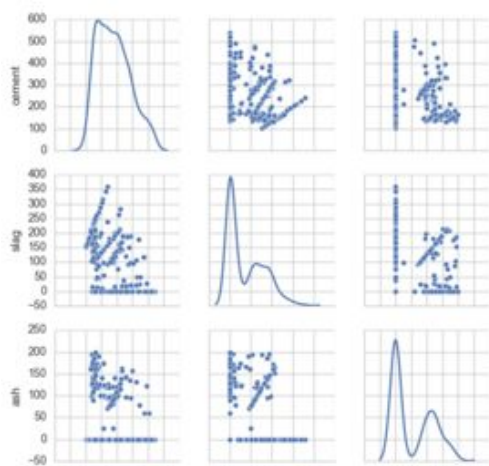


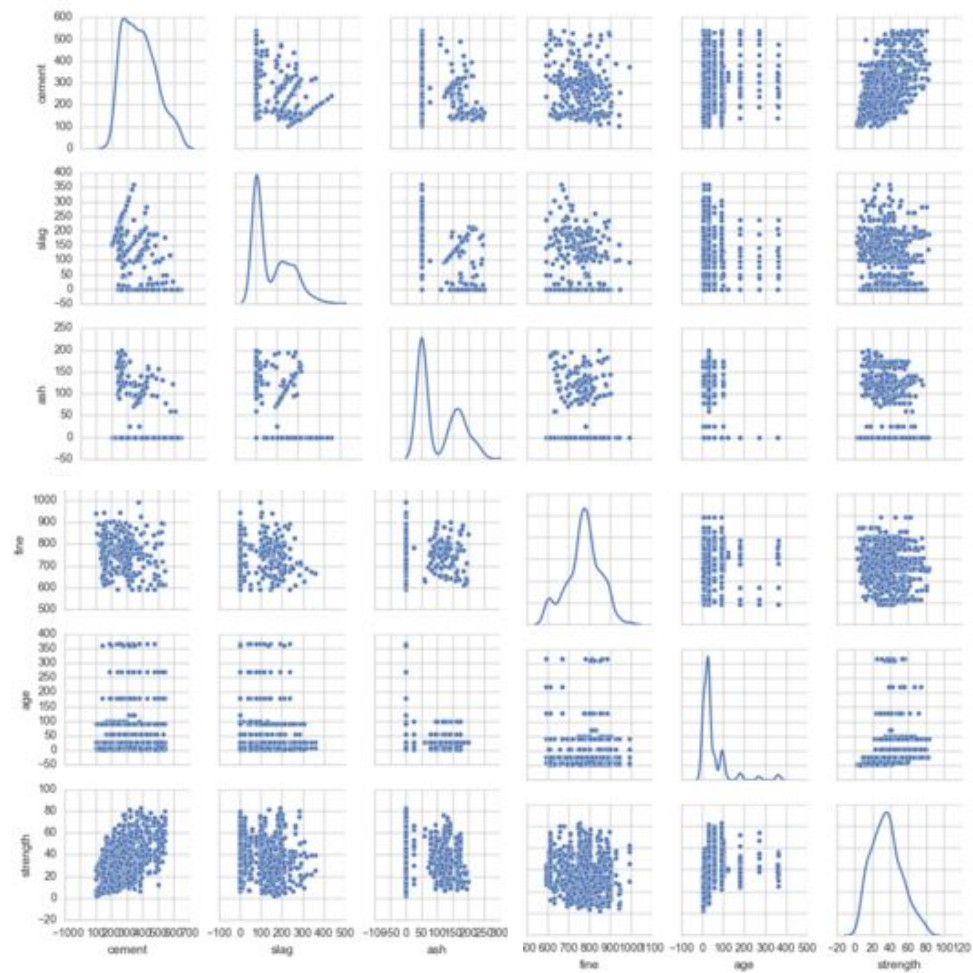


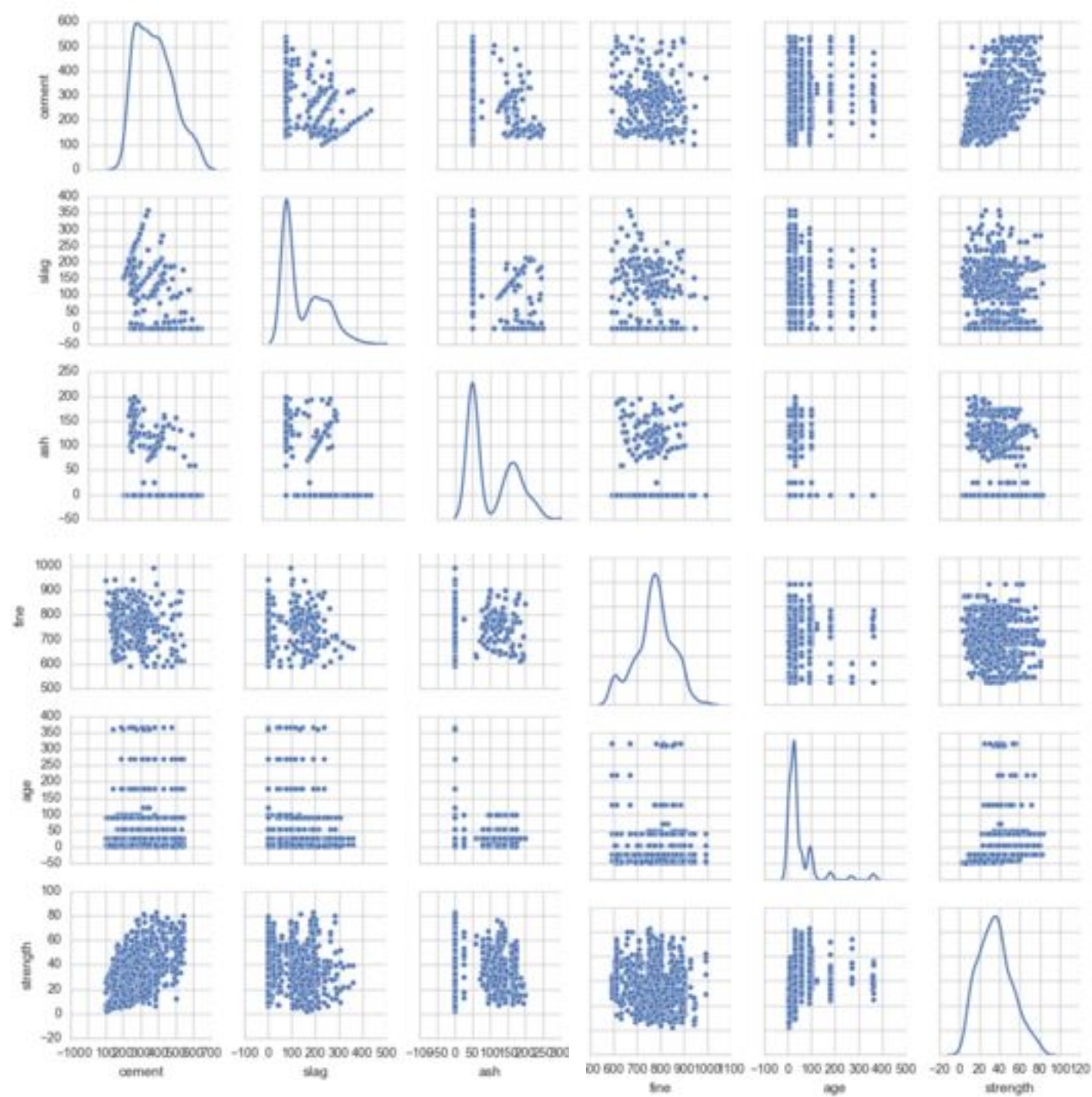
"Overview first,
zoom and filter,
detail on demand"

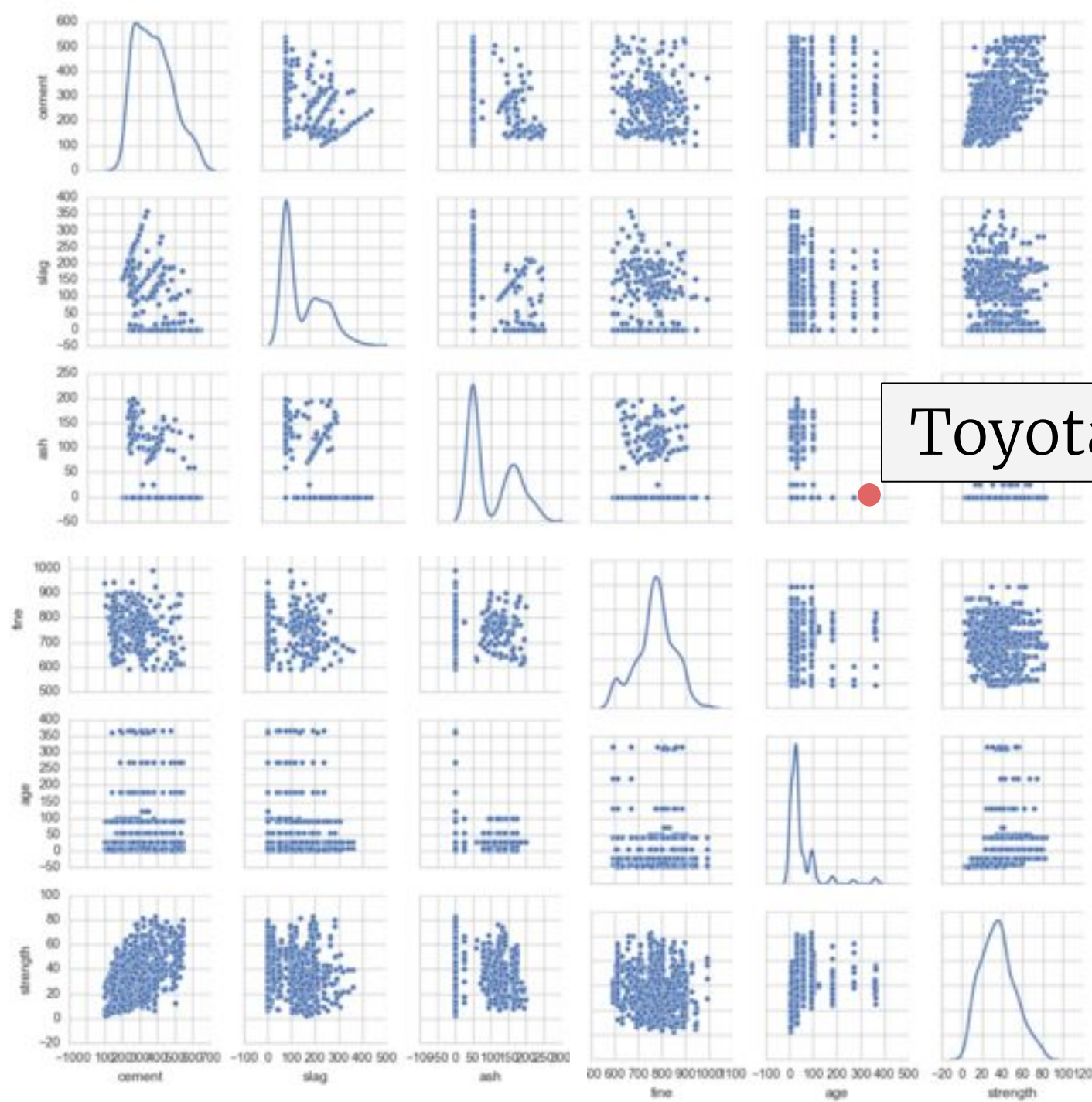
—*Ben Shneiderman*







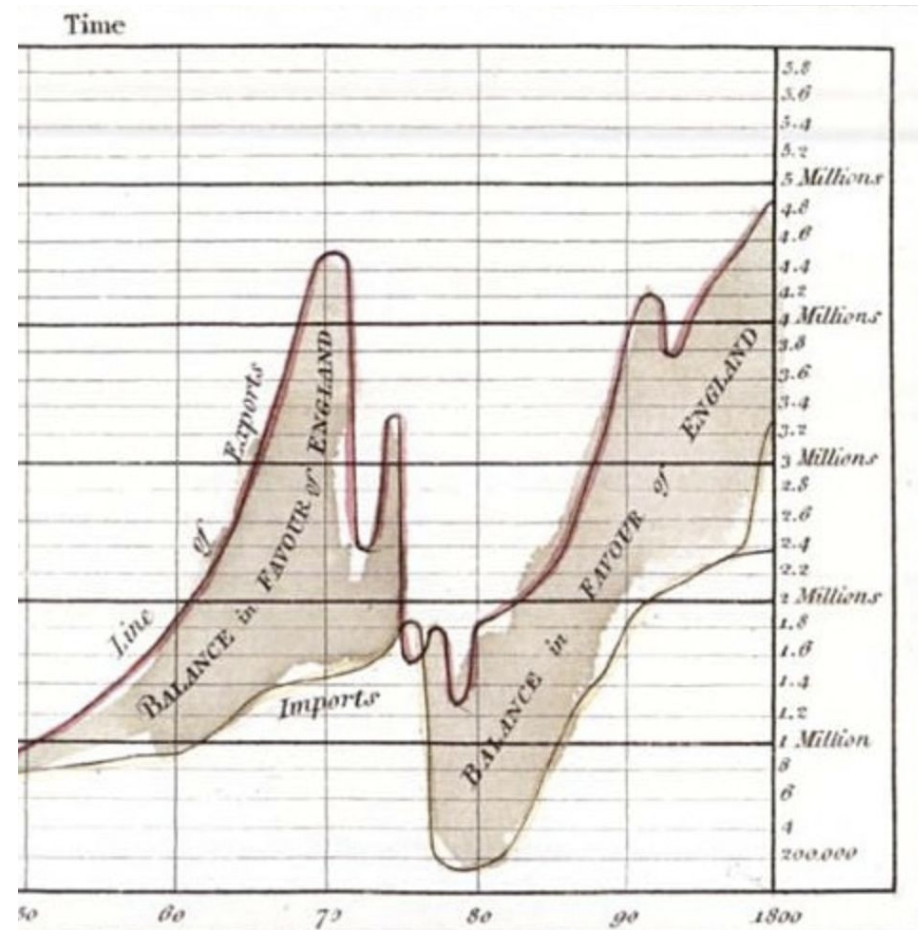




Toyota

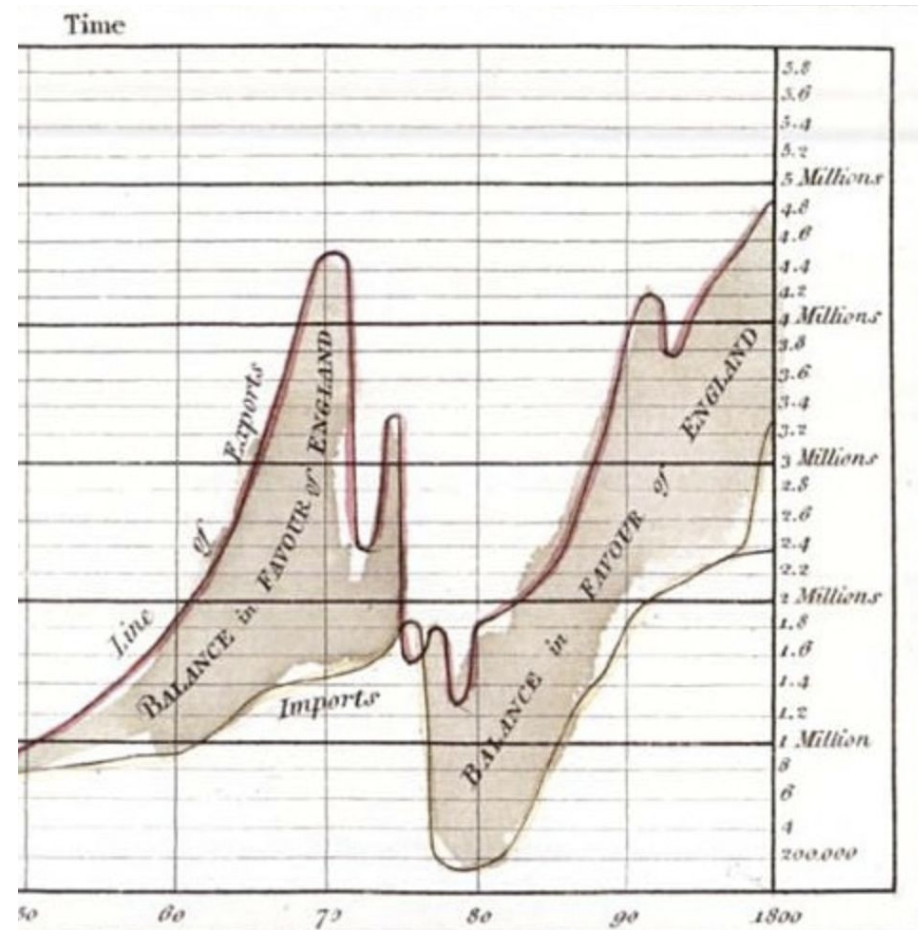
Jaques Bertin: Low level

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



Jaques Bertin: High-level

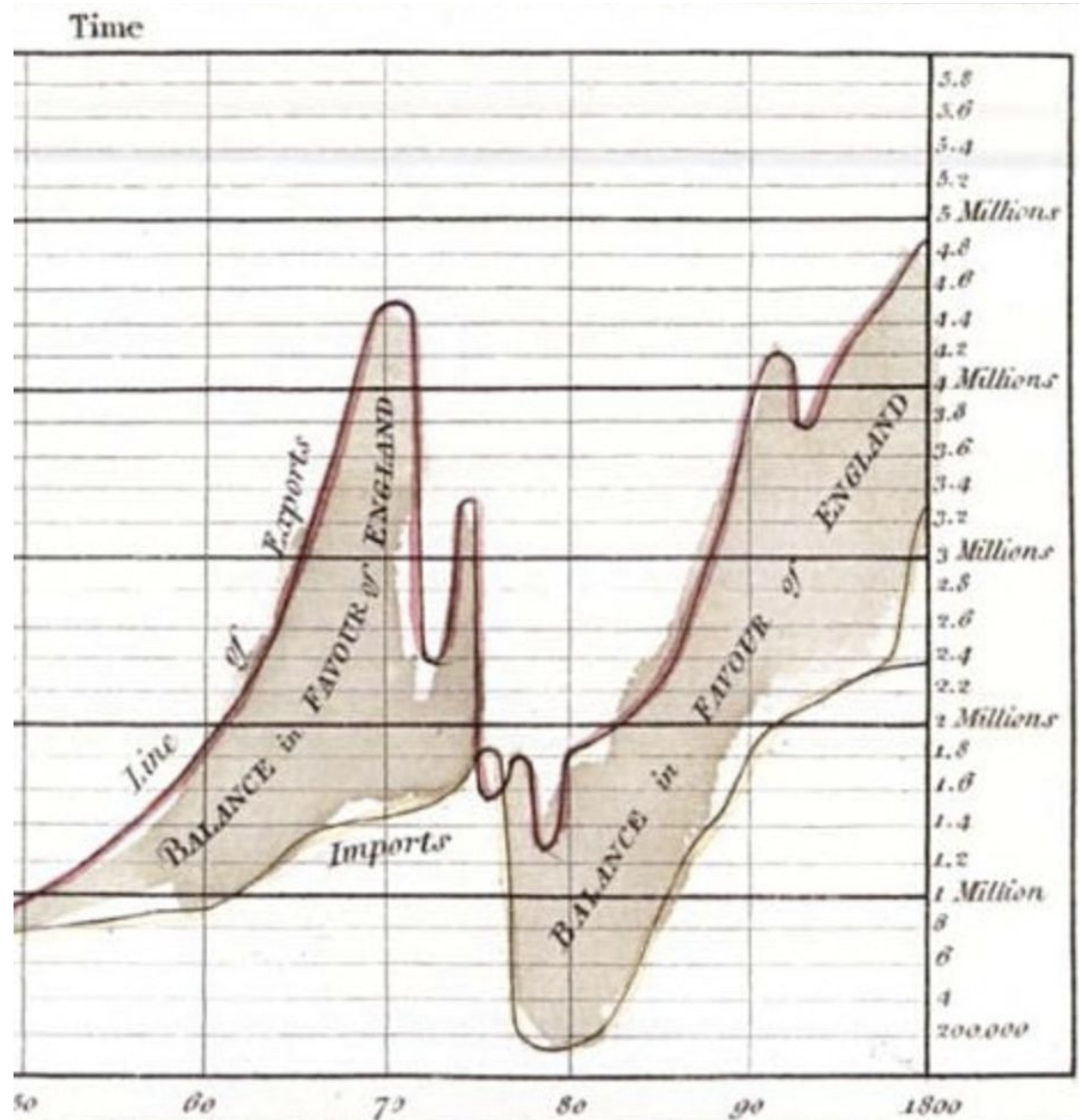
- Which trends
- Which outliers
- Correlations
- Anomalies
- Describe
- Compare
- Group
- ...



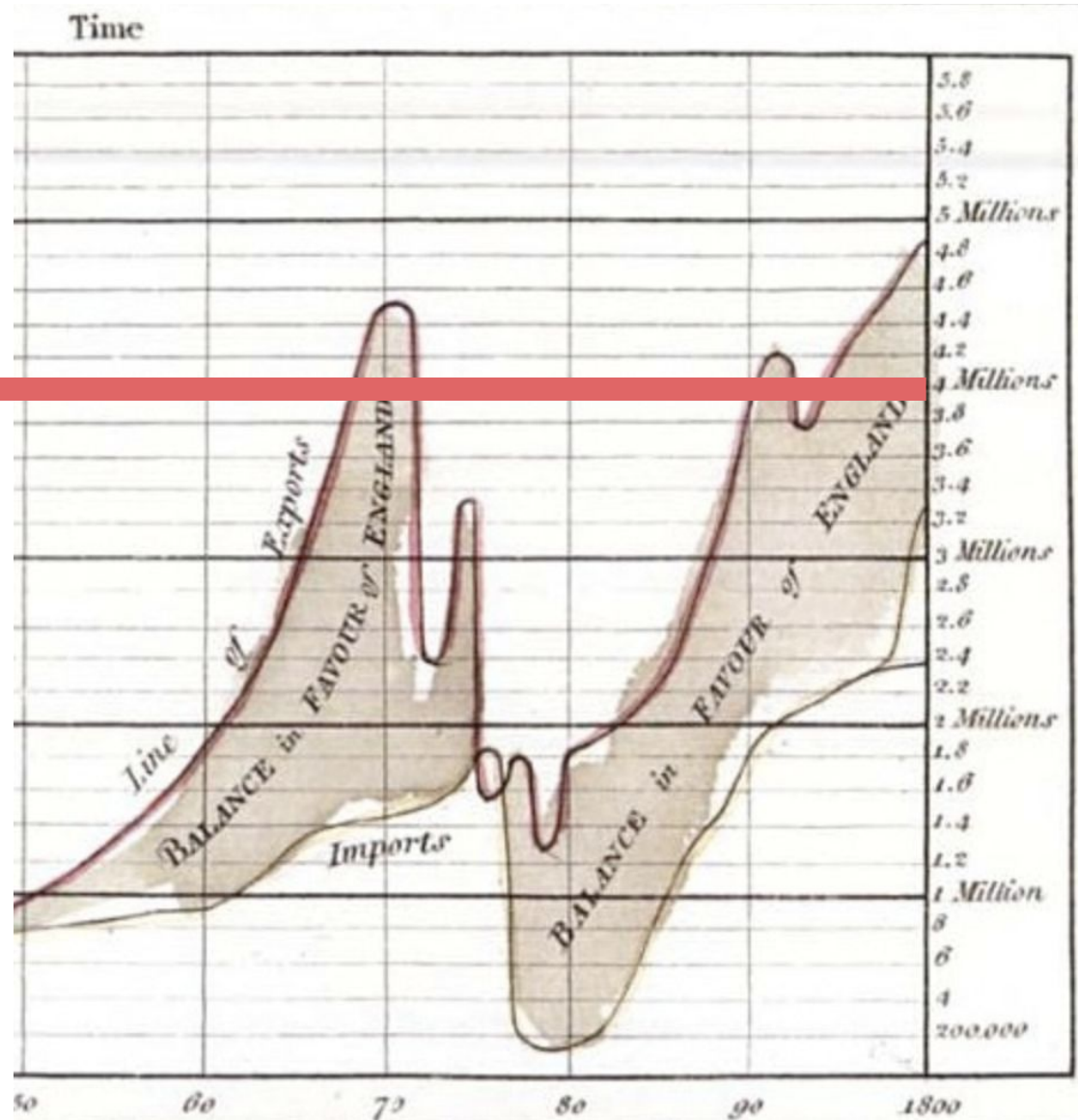
Amar et al.

1. Retrieve Value
2. Filter
3. Compute Derived Value
4. Find Extremum
5. Sort
6. Determine Range
7. Characterize Distribution
8. Find Anomalies
9. Cluster
10. Correlate

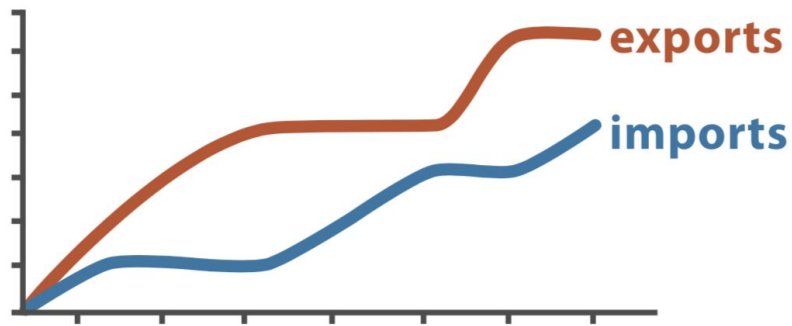
1. Retrieve Value



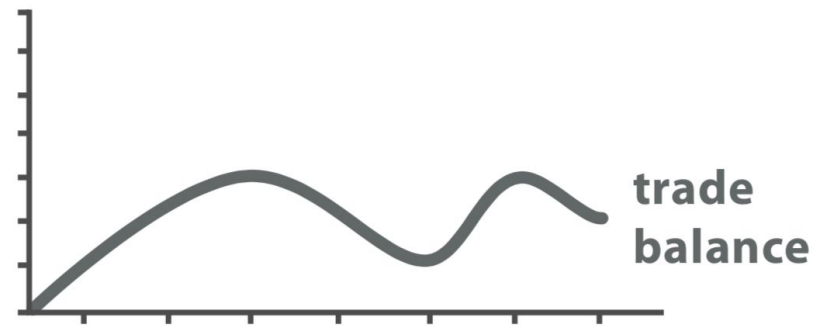
2. Filter Value



3. Compute Derived Value



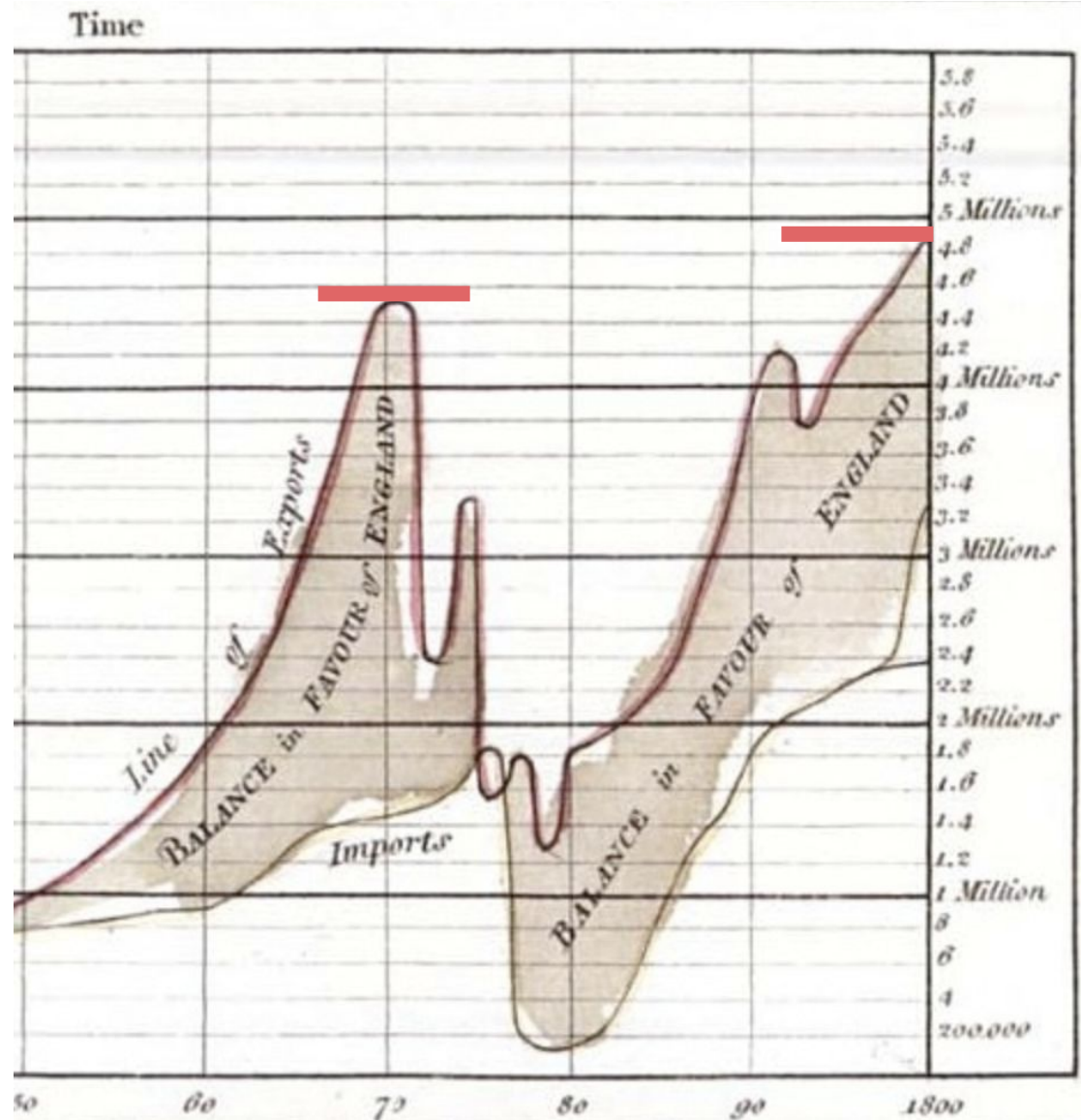
Original Data



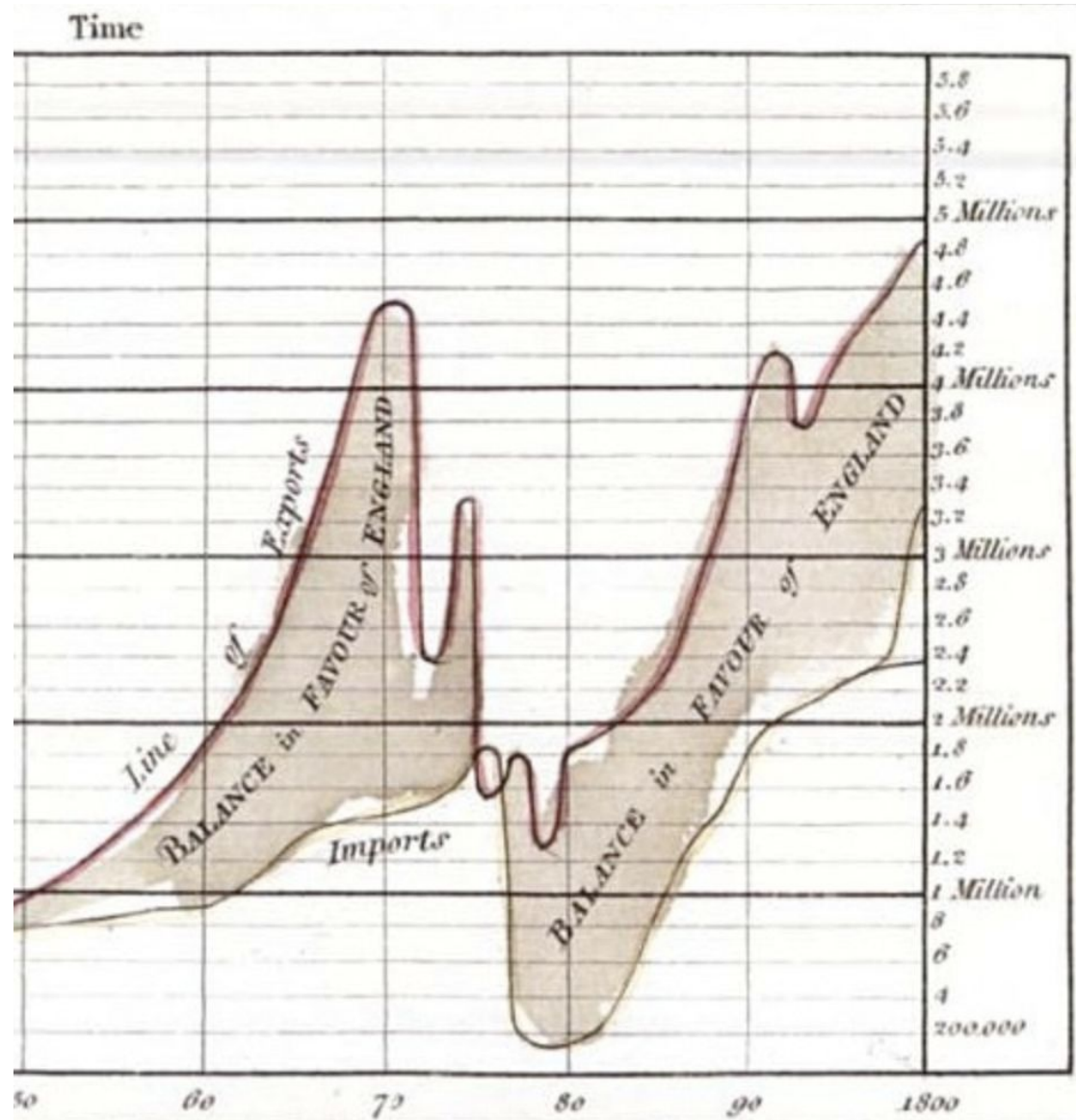
$$\text{trade balance} = \text{exports} - \text{imports}$$

Derived Data

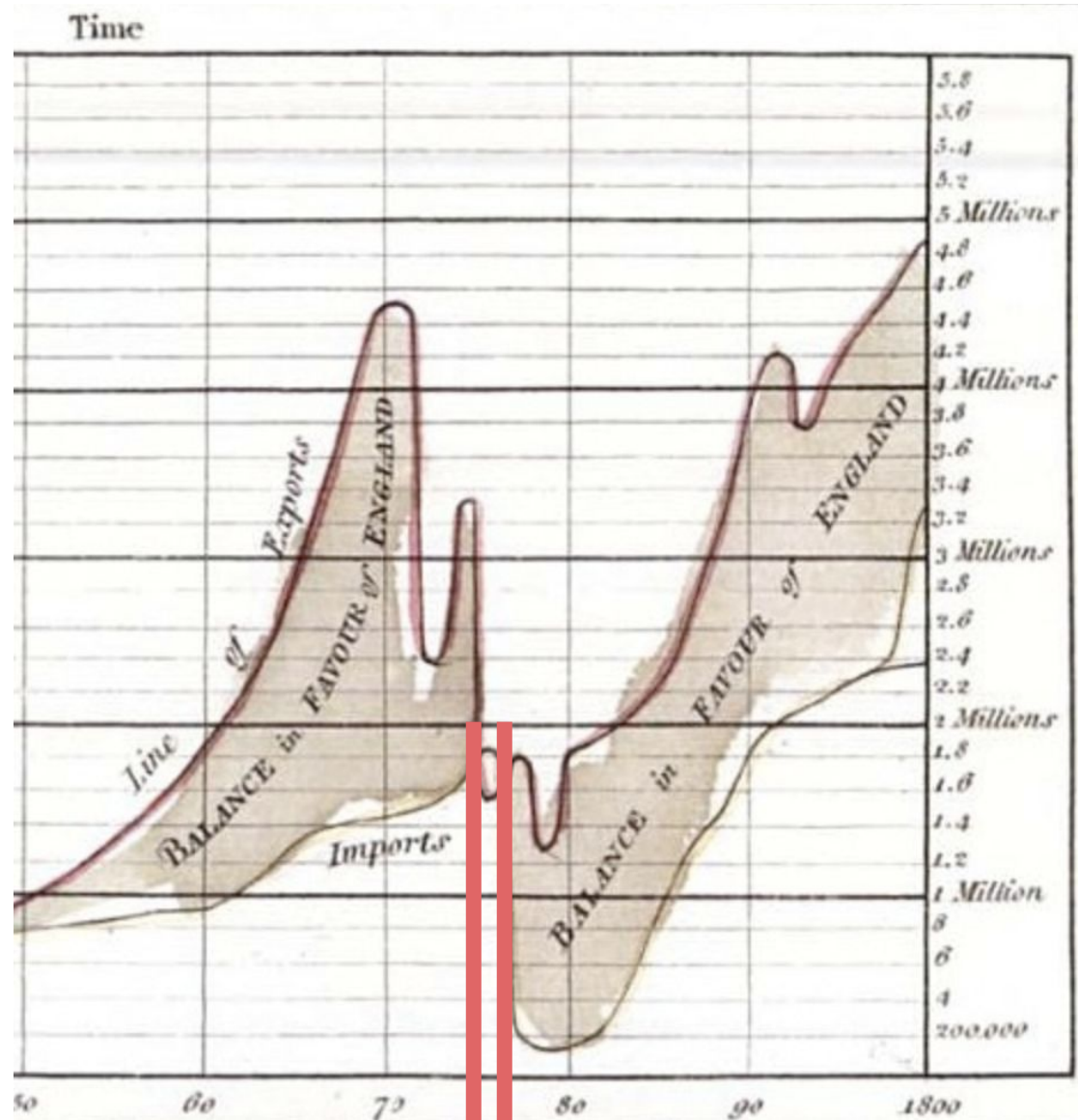
4. Find Extremum



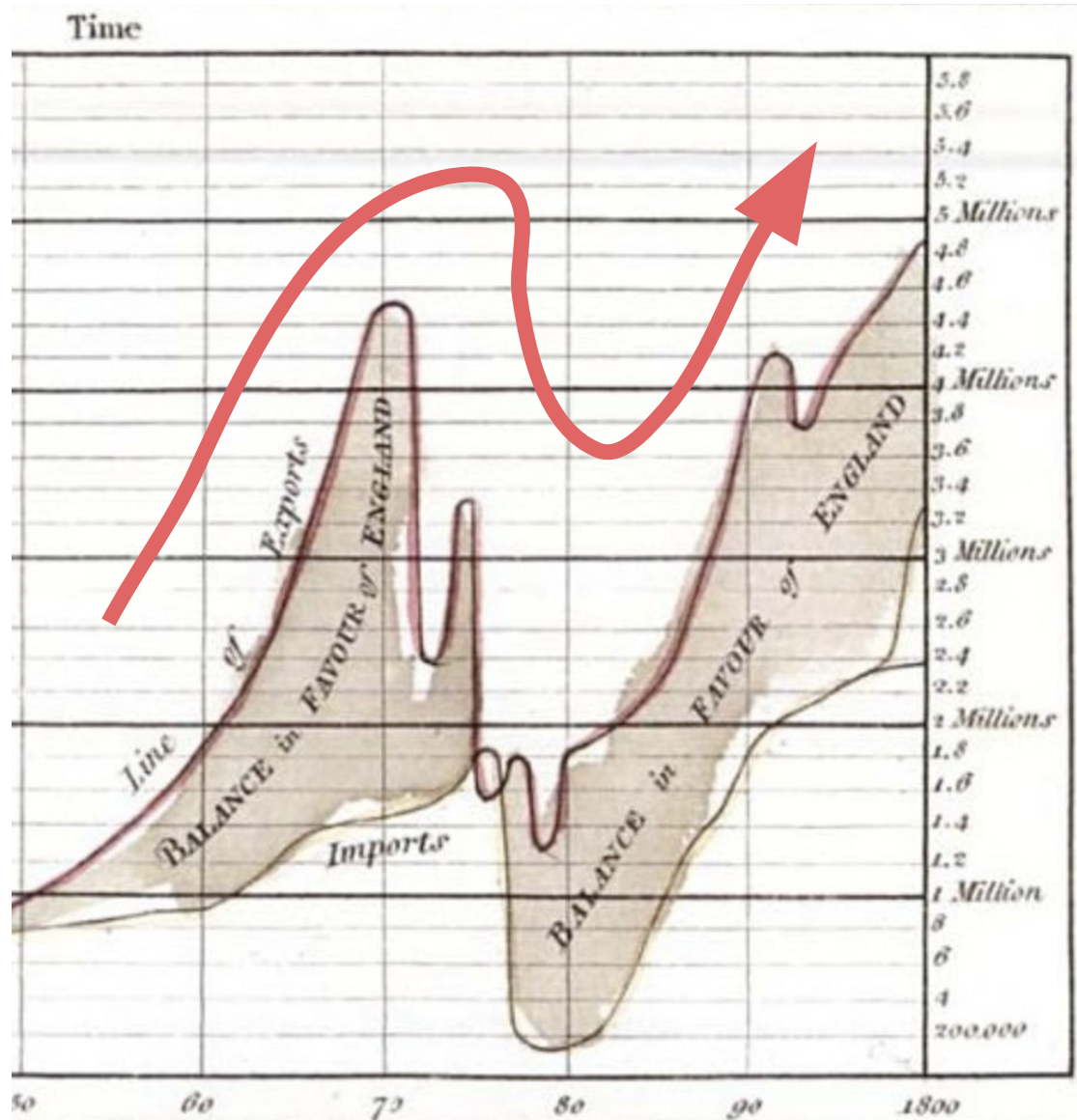
5. Sort



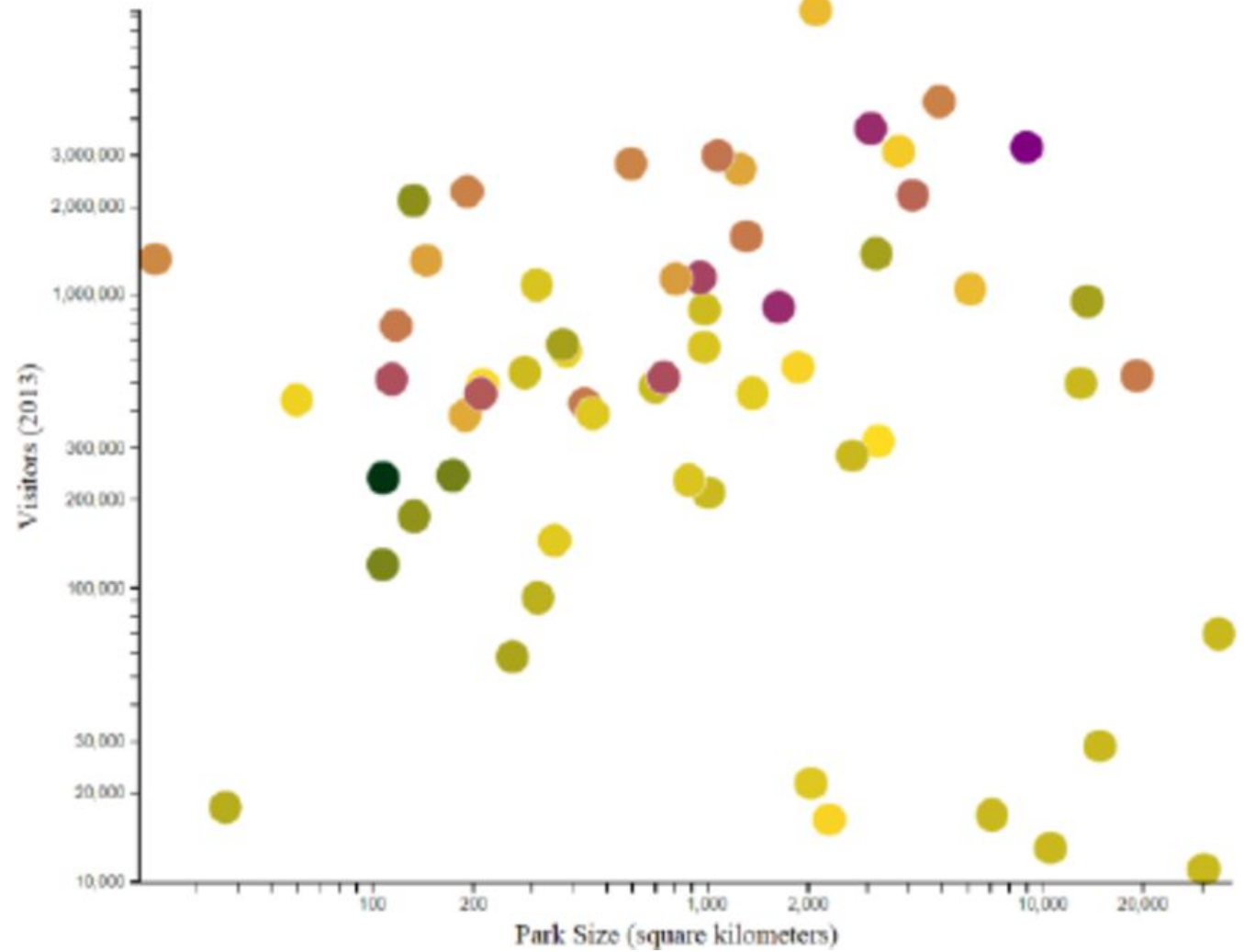
6. Determine Range



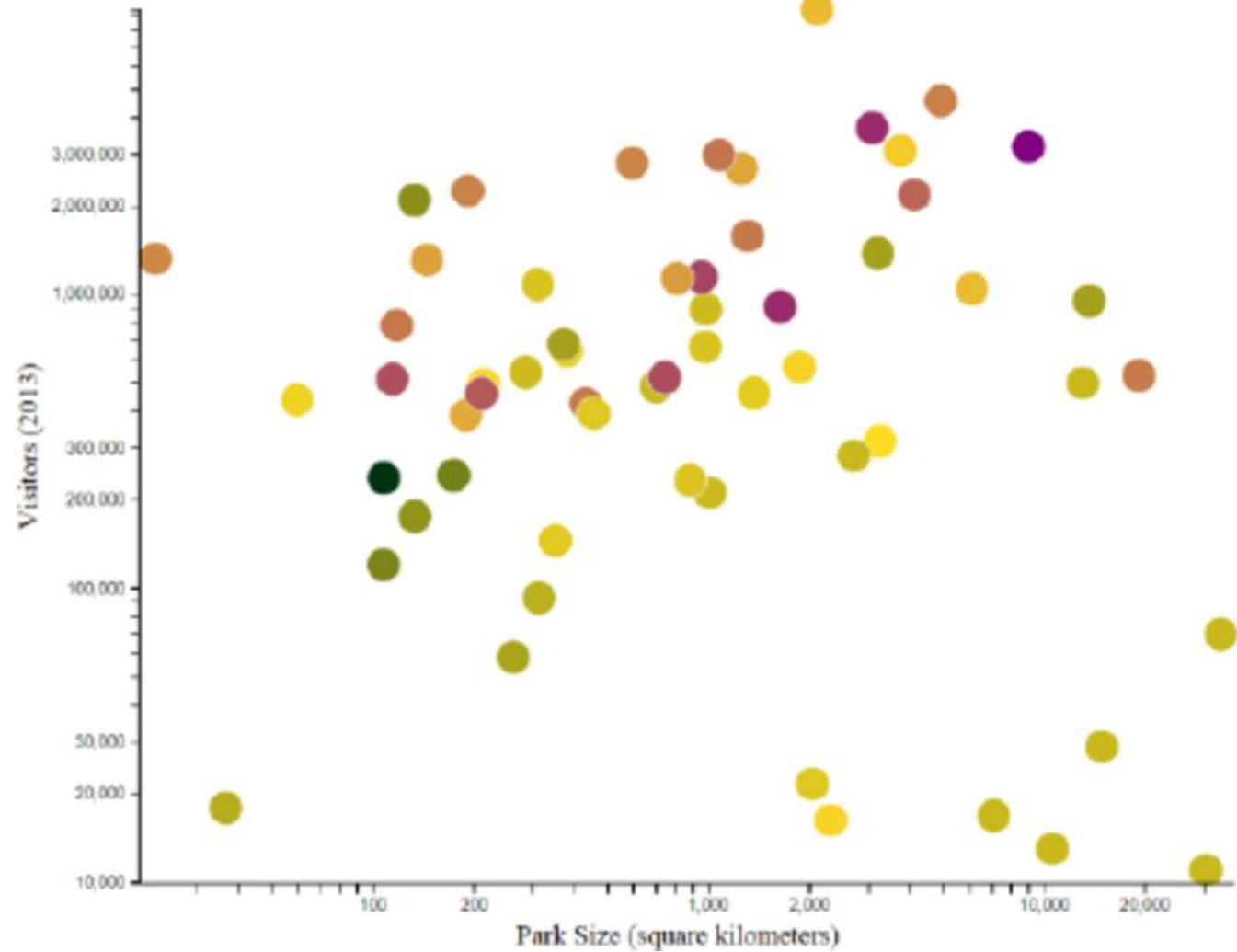
7. Characterize Distribution



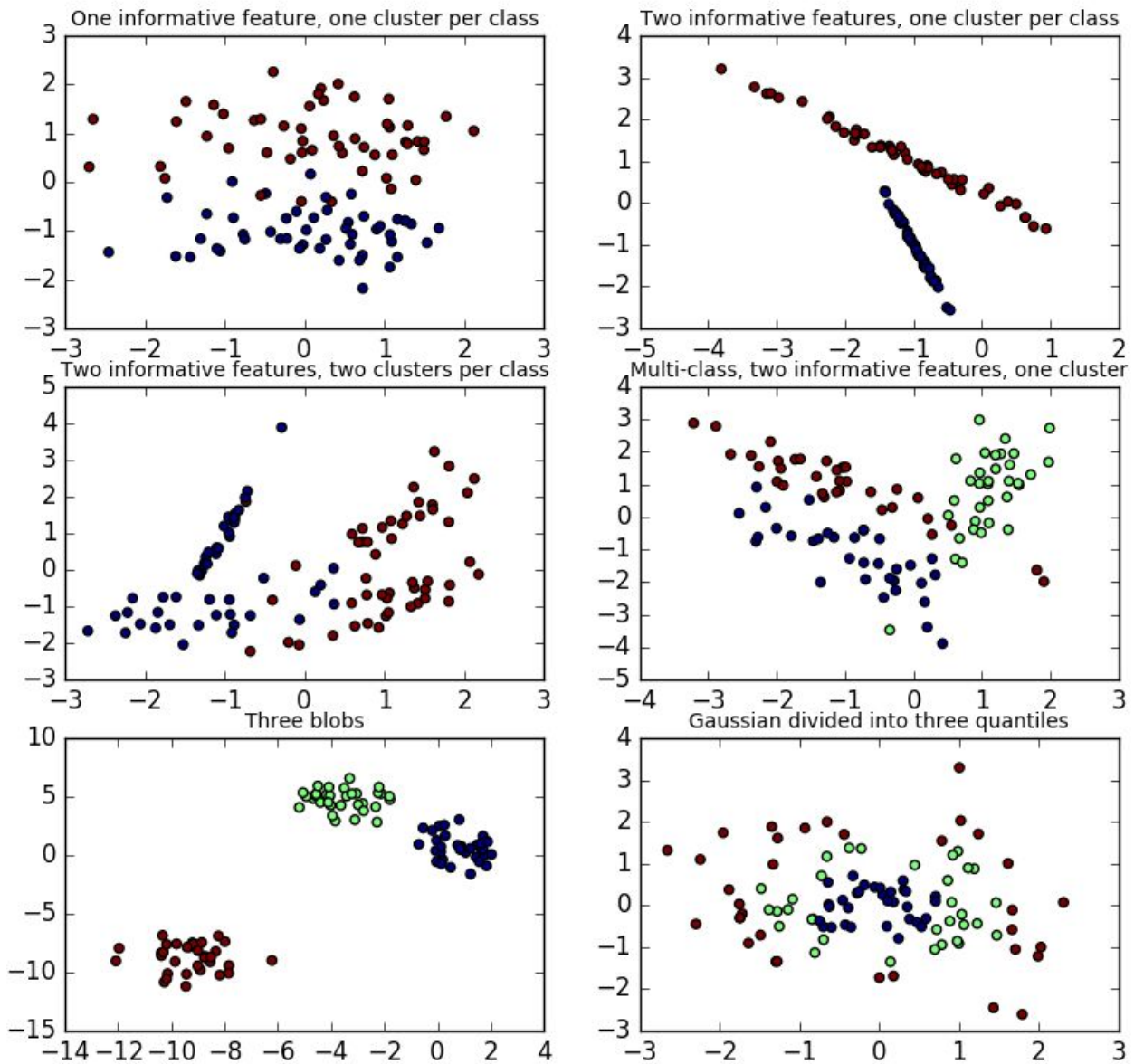
8. Cluster



9. Correlate

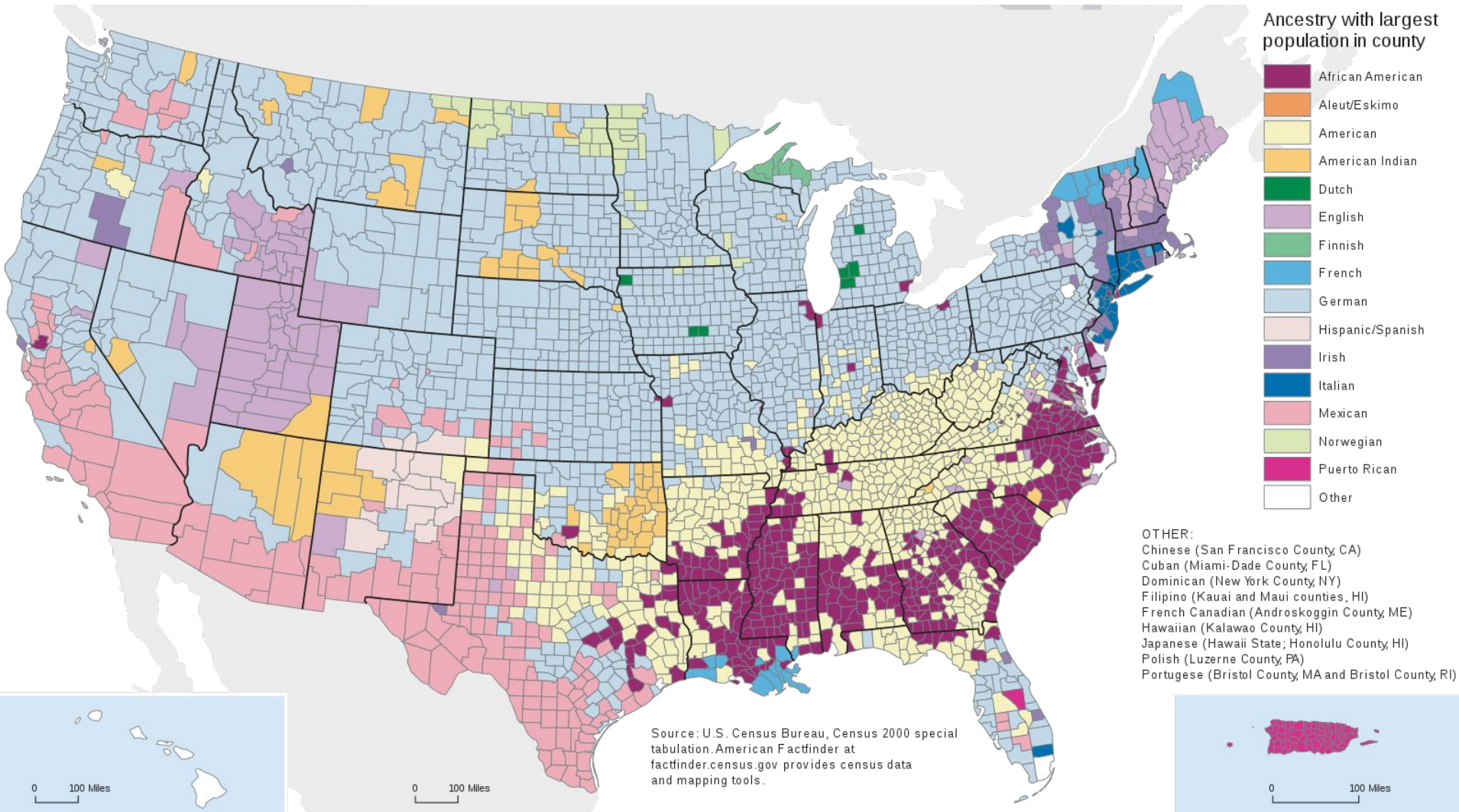


9. Correlate



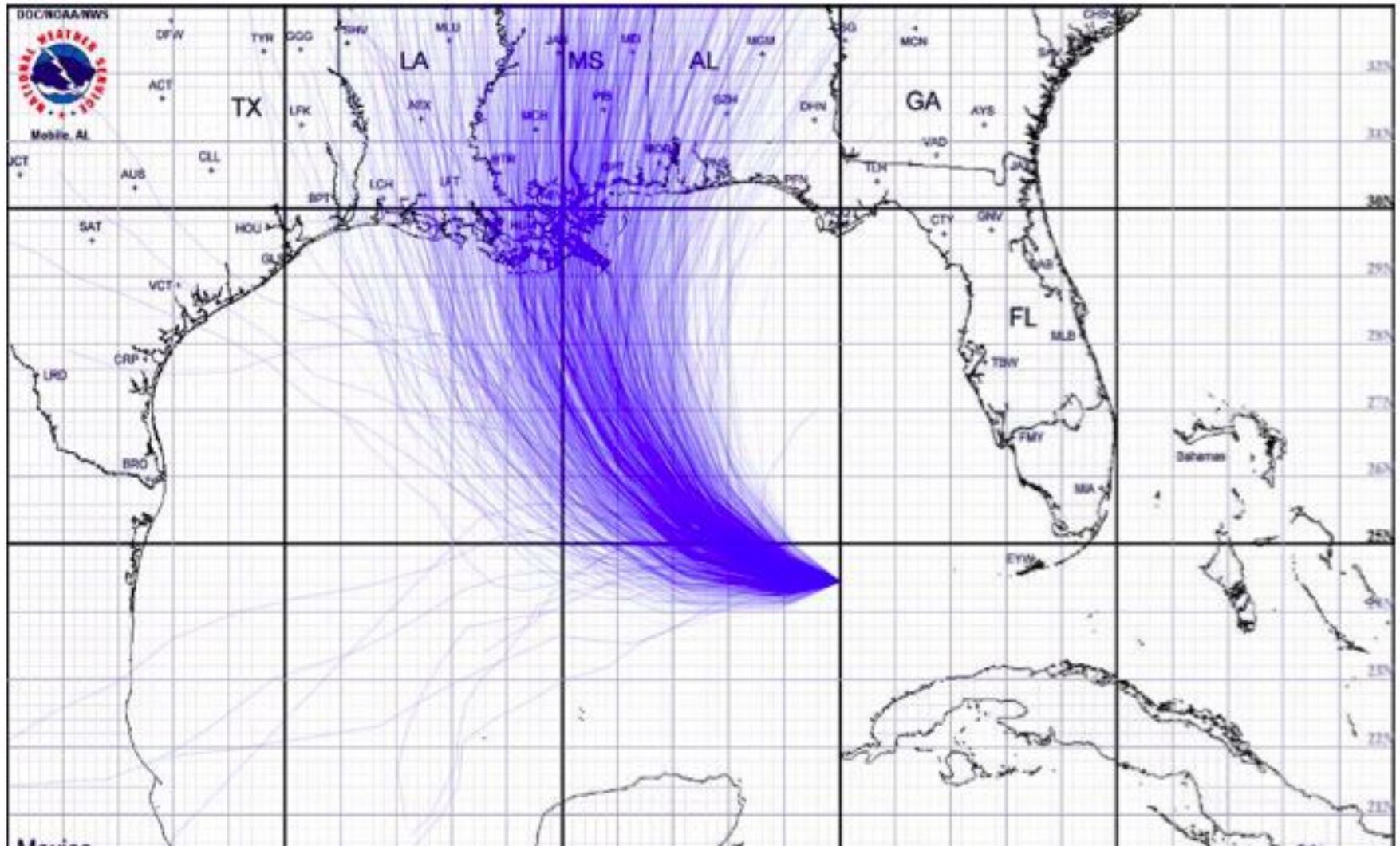
Discovery

Trends, outliers, clusters, anomalies...



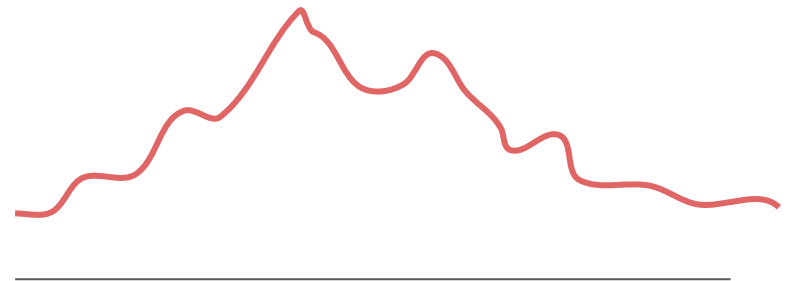
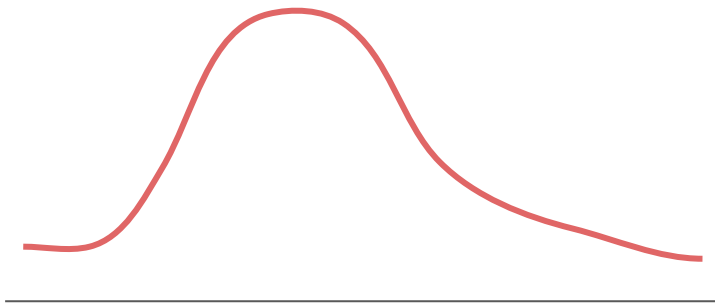
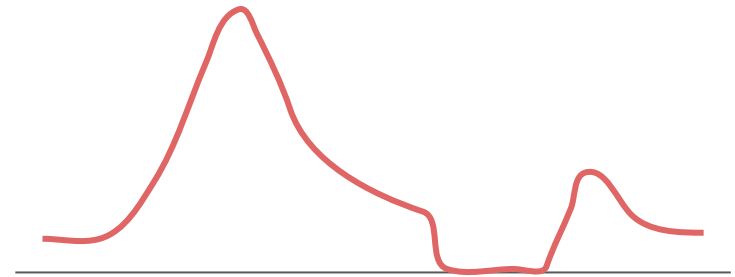
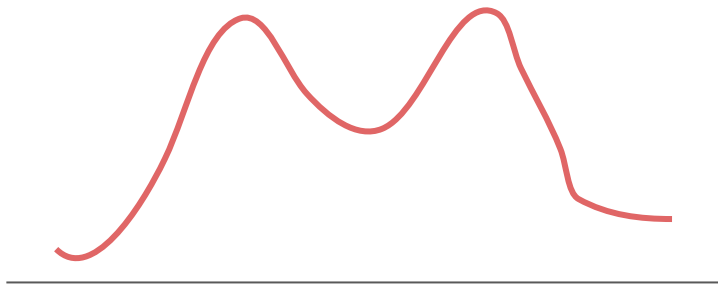
Data Quality

Noisiness, missing data, certainty, duplicates...



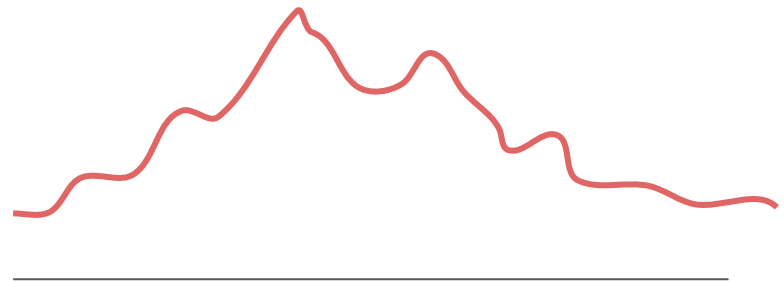
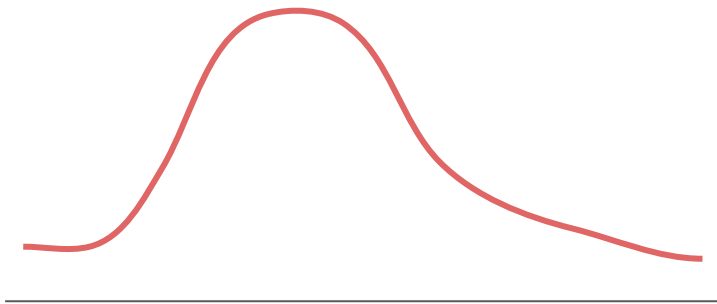
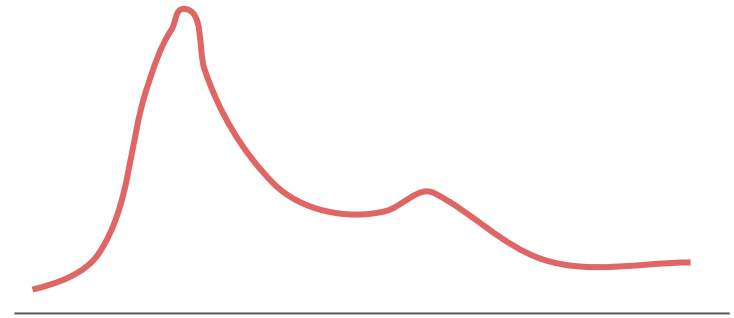
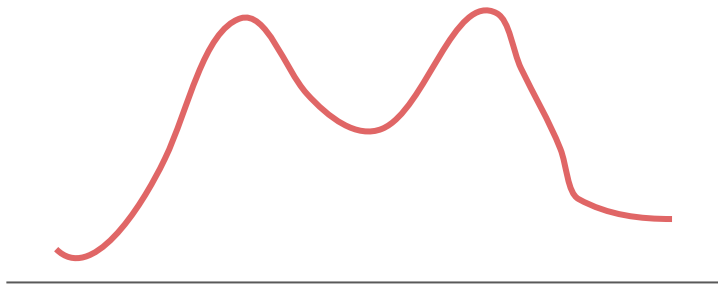
Data Quality

Noisiness, missing data, certainty, duplicates...



Inform Analysis

Statistic tests, cleaning methods, clustering methods, ...

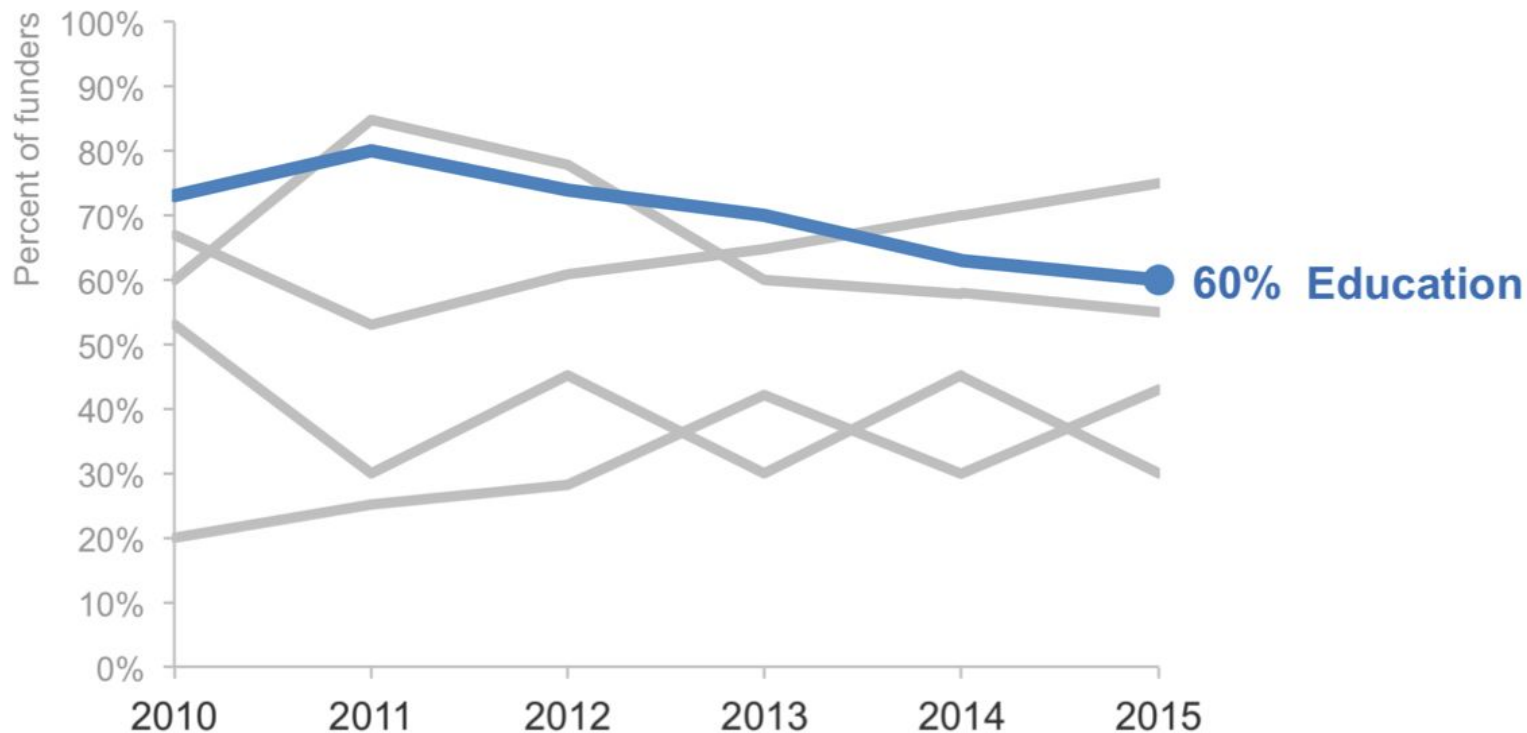


Presentation

Clarity, highlight elements, direct user attention, ..

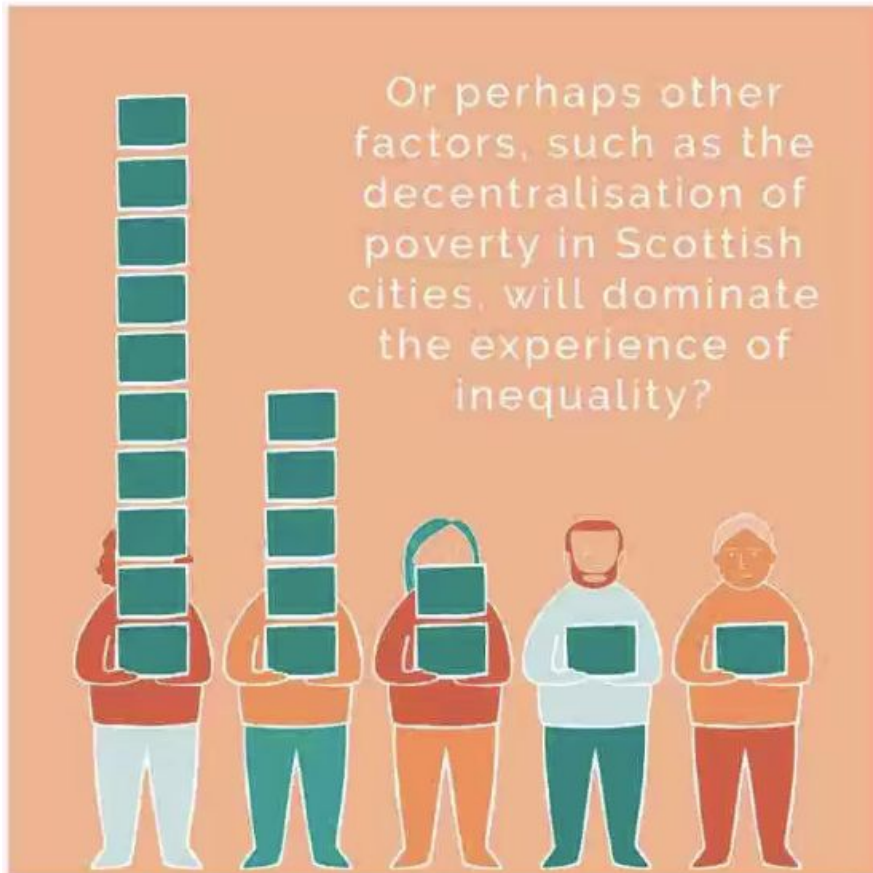
Support for education non-profits has decreased markedly over the past 5 years.

Types of non-profits supported by area funders



Storytelling

Attract attention, drag readers in, communicate message, engage, entertain ...



Data-specific Tasks

- Networks: Lee et al.
- Temporal networks: Ahn et al., Bach et al.
- Time: Aigner et al.
- Geography: Peuquet et al.
- Spatio-temporal: Andrienko et al.
- ...

2.6

Messages

Take-home messages

- Not vaccinating is bad
- 1 Flight US–Europe = 1 tons of CO₂
- Use/produce less plastic in your life
- Data is too uncertain to work with
- Interactive network exploration is helpful
- ...

Data Challenge Method Sheet

Data:

Structure:

Context:

Audience:

Questions:

Messages:

Questions?