

“From analysis to communication”

Jo Wood

Professor of Visual Analytics

 @jwolondon.bsky.social

“From analysis to communication”

Jo Wood

Professor of Visual Analytics

 @jwolondon.bsky.social

Communication

is

Design

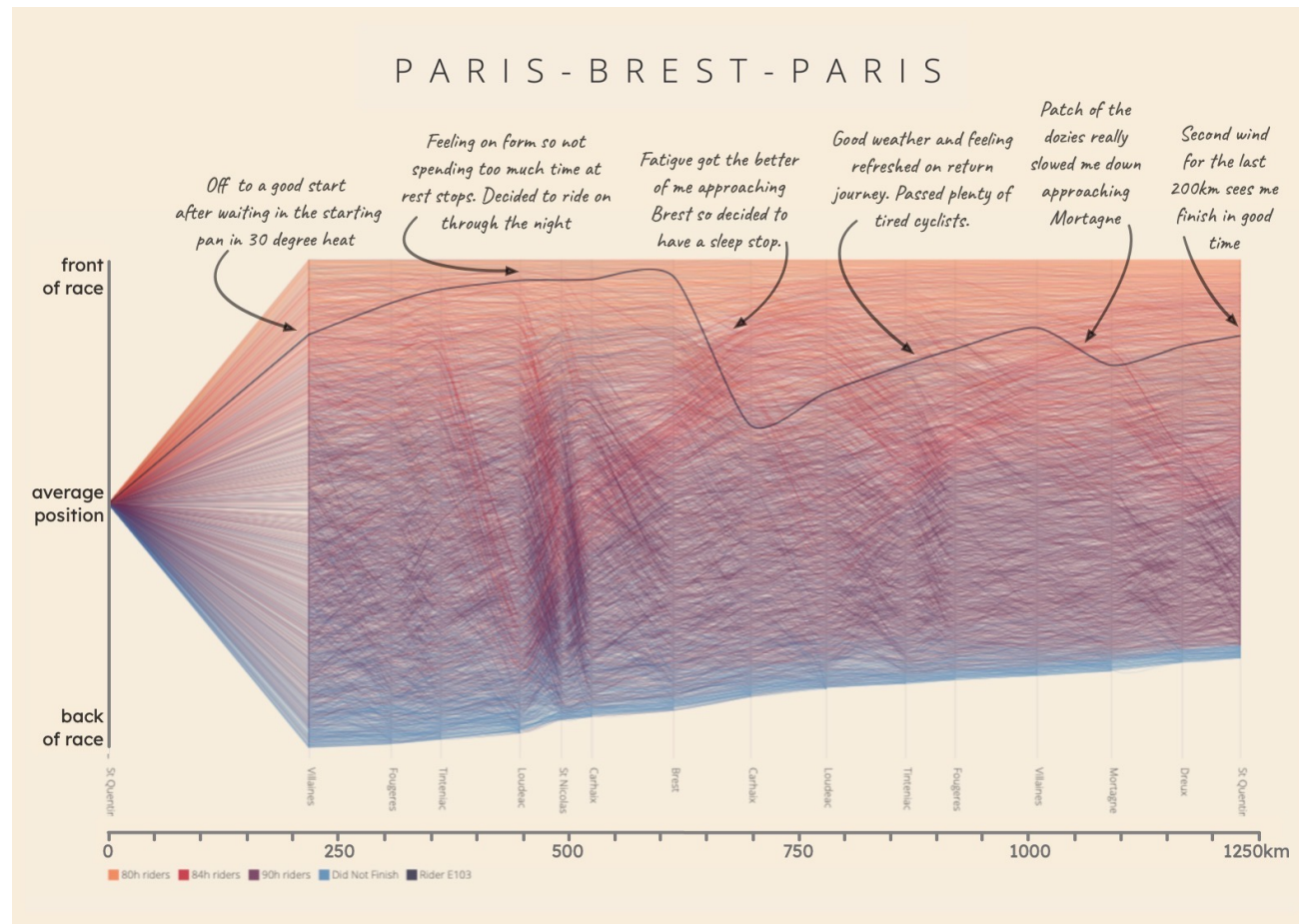
is

Analysis

is

Communication

Visualization is subjective



But so is this...

$$\Pr(y_i = 1) = \text{logit}^{-1} \left(X_i \beta + \sum_{k=1}^K \alpha_{j[i]}^k \right), \text{ for } i = 1, \dots, n$$

$$\alpha_j^k \mid \sigma^k \stackrel{\text{ind.}}{\sim} \text{N}(0, (\sigma^k)^2), \text{ for } k = 1, \dots, K, j = 1, \dots, J_k$$

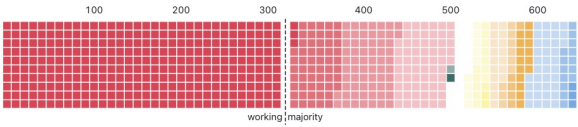
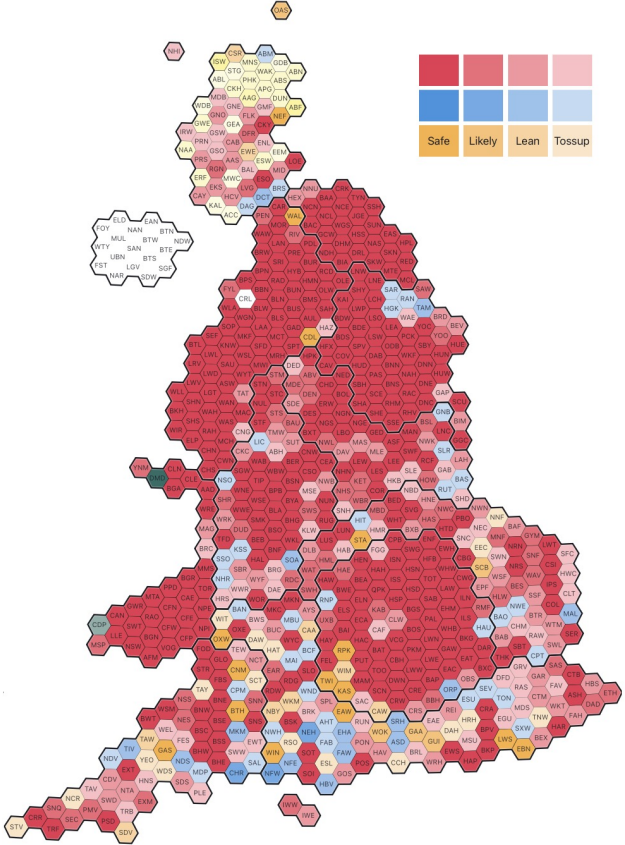
$$\sigma^k \sim \text{N}_+(0, 1), \text{ for } k = 1, \dots, K$$

$$\beta \sim \text{N}(0, 1),$$

(from Gao et al 2021, Improving multilevel regression and poststratification with structured priors)

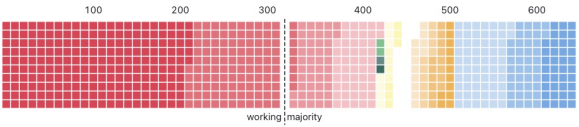
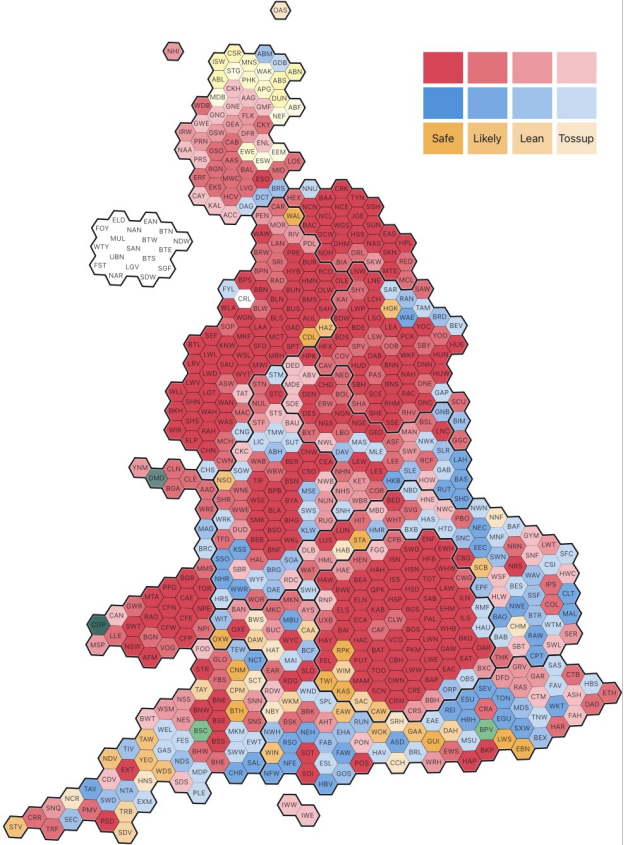
Survation MRP results

Based on 3 June 2024 MRP



YouGov MRP results

Based on 3 June 2024 MRP



Data science is a **process** resulting from **choices**.

We benefit from documenting that process and rationalising those choices.

But widespread uptake of process documentation requires it to be **low friction**.

Literate visualization structures the design process through design exposition

Wood, Kachkaev and Dykes (2019) Design exposition with literate visualization,
IEEE transactions on visualization and computer graphics 25.1 (2018): 759-768.

Jupyter (Python)
Observable (JavaScript)
Pluto (Julia)
R-Markdown (R)
Litvis (Elm)

The image displays three overlapping browser windows from the ObservableHQ platform, each showing a different narrative schema for evaluating visualizations. The windows are titled 'Socratic Dialogue', 'Visualization Algebra Evaluation', and 'New Idiom'. Each window includes a sidebar with a list of templates, a main content area with text and diagrams, and a right-hand sidebar with additional information.

Socratic Dialogue
INM402 Data Visualization 2024
Students taking INM402 Data Visualization at City, University of...
Narrative schema for evaluating a visualization through Socratic dialogue. Modify and add questions and answers below.

Naive questioner:
What are you trying to achieve with this visualization?

I want my visualization to...

Naive questioner:
Why have you chosen this data source and sample?

Because I...

Naive questioner:
Why have you made these visual mark design choices?

Because I...

Naive questioner:
To what extent does your visualization meet your original objectives?

While it does X, it doesn't quite do Y...

Visualization Algebra Evaluation
INM402 Data Visualization 2024
Students taking INM402 Data Visualization at City, University of...
Narrative schema for evaluating a visualization design and dataset using the framework proposed by Kindmann & Scheidegger (2014).

Principle of Representation Invariance
 $\alpha = 1_D \Rightarrow \omega = 1_V$. Confirm that non-meaningful changes in data representation, such as table row order have no discernable effect on visualization. If they do, the visualization will contain **hallucinators**.

Principle of Unambiguous Data Depiction
 $\omega = 1_V \Rightarrow \alpha = 1_D$. What are the smallest meaningful changes in data that should result in identifiable changes in the visualization? If visualization changes are not detectable, the visualization will contain **confusers**.

New Idiom
INM402 Data Visualization 2024
Students taking INM402 Data Visualization at City, University of...
Narrative schema for evaluating a new visualization idiom (new form of visualization design).

Aims
This idiom aims to...

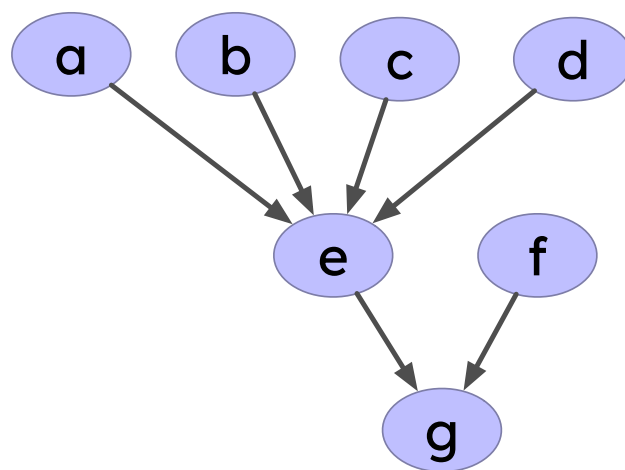
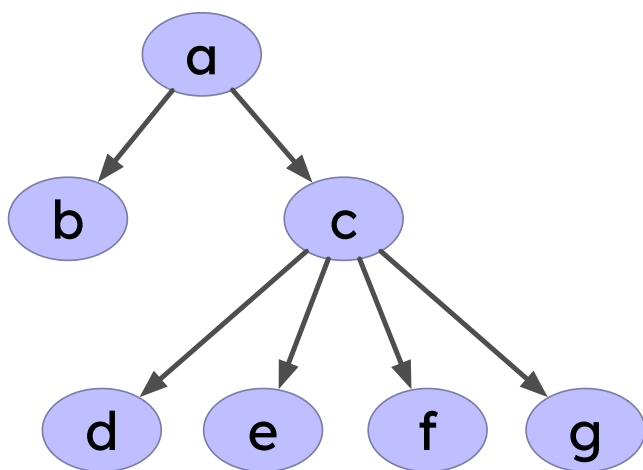
- aim one
- aim two
- etc.

Description
A description of how the idiom works.

Architypes
Provide some architypes that illustrate how the idiom would depict some simple, knowable patterns. This helps to learn how to 'read' and make sense of visualizations created with this idiom.

Limitations
What are the limitations of the idiom? What can it not do? What problems might a reader face when interpreting the idiom?

Narrative schemas



We could also apply **narrative schemas** to **data selection, shaping** and **analysis**

Prompts for low-friction structured **exposition** of **rationale**

Data science is not about **product** but
encoding and **communicating** thought.