

Storytelling with Data

Module 8: Effective visuals

Scott Spencer

Faculty and Lecturer
Columbia University

Agenda

Upcoming deliverable

Today's objectives

Graphics to narrative

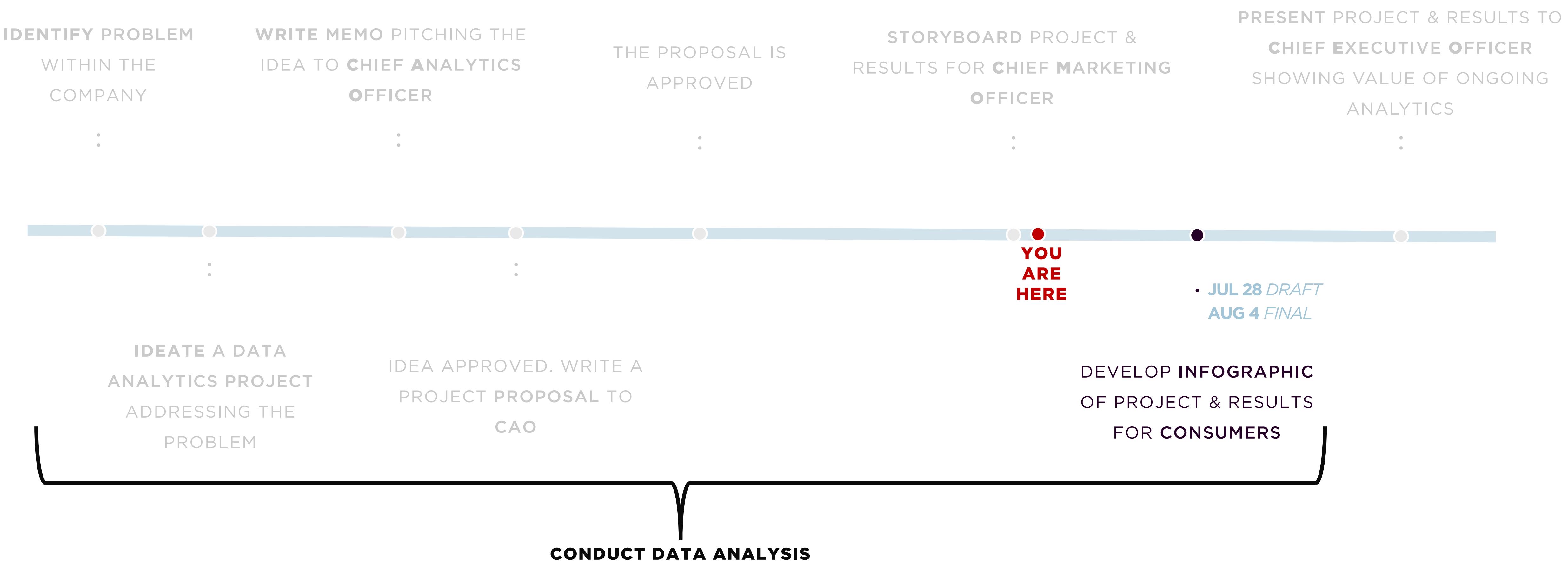
Complexity from common graphics

Questions or suggestions?

Upcoming deliverable

Upcoming deliverables

Information graphic – reframe your story, this time building off the messages you built for the marketing team in order to craft an infographic that displays the results of the analytic work in a way that is accessible, engaging, and exciting for a **general or consumer audience**.



Today's Objectives

Objectives

1

Creating a visual narrative
from common graphics

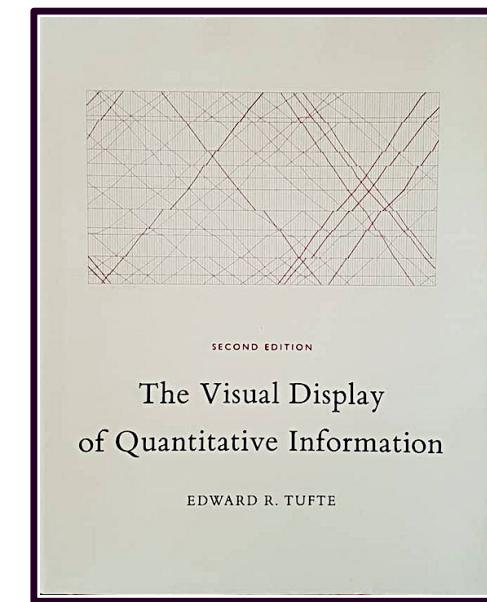
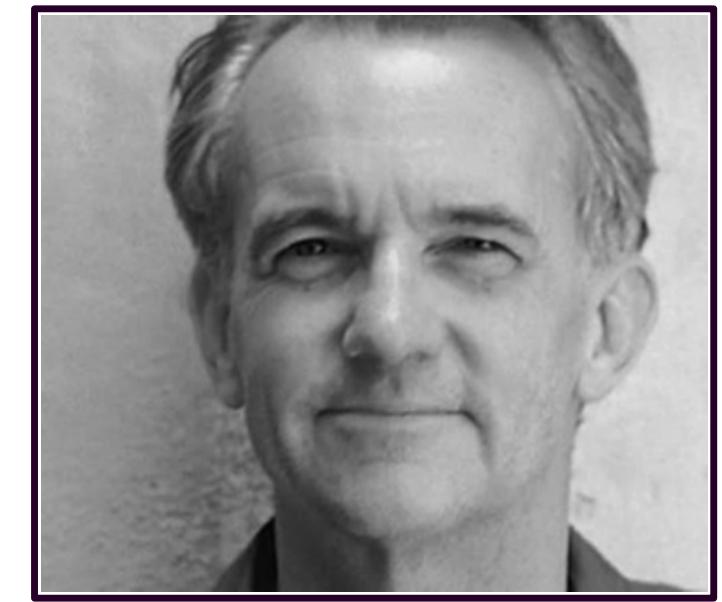
2

Consider the role of
layering and hierarchy in
graphic narrative

3

Building complexity from
common graphics, example

From default visual to visual narrative



The Visual Display of Quantitative Information

Tufte

Hailed "The Leonardo da Vinci of data" by the New York Times. He is professor emeritus of Political Science, Statistics, and Computer Science at Yale University.

Data-ink

The non-erasable core of a graphic, the non-redundant ink arranged in response to variation in the numbers represented. Annotations are part of the data-ink.

$$\text{data-ink ratio} = \frac{\text{data-ink}}{\text{total ink used to print the graphic}}$$

= proportion of a graphic's ink devoted to the non-redundant display of data-information

= $1.0 - \text{proportion of a graphic that can be erased without loss of data-information}$

Maximize within reason

Maximize the data-ink ratio, within reason.
Erase non-data-ink, within reason.
Erase redundant data-ink, within reason.

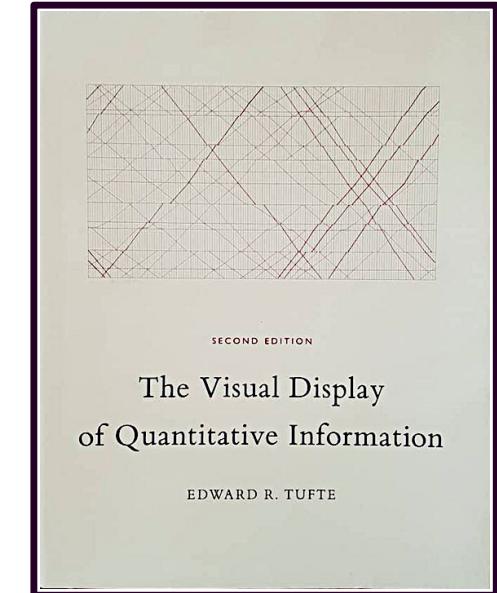
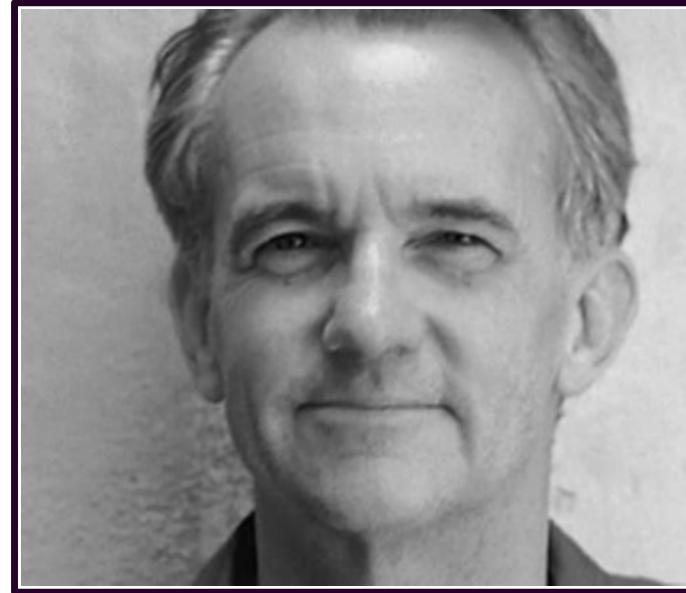
Size graphics for legibility

As the quantity of data increases, data measures must shrink. . . . The way to increase data density other than by enlarging the data matrix is to reduce the area of a graphic.
Graphics can be shrunk way down.

The Visual Display of Quantitative Information

Tufte

Hailed "The Leonardo da Vinci of data" by the New York Times. He is professor emeritus of Political Science, Statistics, and Computer Science at Yale University.



Simplicity of design, complexity of data

Words and pictures belong together

Proportion and scale: the shape of graphics

Graphical excellence is often found in simplicity of design and complexity of data.

Viewers need the help that words can provide. **Words on graphics are data-ink**, making effective use of the space freed up by erasing redundant and non-data-ink.

Note, the **size of type** on and around graphics can be quite **small**, since the phrases and sentences are usually not too long.

The principle of data/text integration is: data **graphics are paragraphs about data** and should be treated as such.

Our eye is naturally practiced in detecting **deviations from the horizon**.

Horizontally shaped plots tend to make it **easier to directly label** and explain the data.

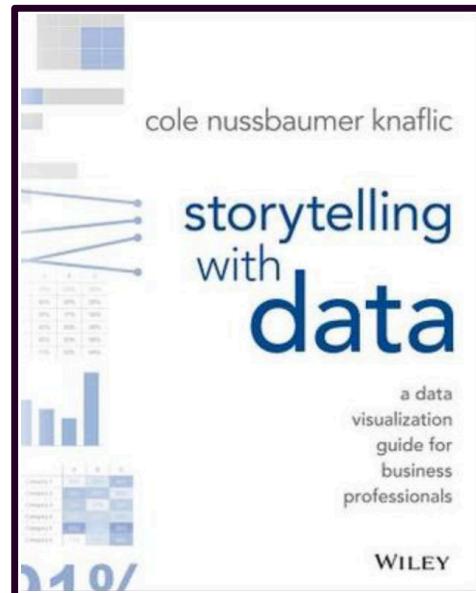
Tradition places the effect vertically, the cause horizontally.

The empirically studied **Golden Rectangle**, a 1.0×1.618 ratio is aesthetically pleasing.

Storytelling with data

Knaflic

The author is a consultant focused on visual displays. Her experience arose from human resources in Google where she applied theory learned as a student of Yale's Edward Tufte.



Please approve the hire of 2 FTEs

to backfill those who quit in the past year

Ticket volume over time



Data source: XYZ Dashboard, as of 12/31/2014 | A detailed analysis on tickets processed per person and time to resolve issues was undertaken to inform this request and can be provided if needed.

Discussion: what differences do you see? What advice has Knaflic applied?

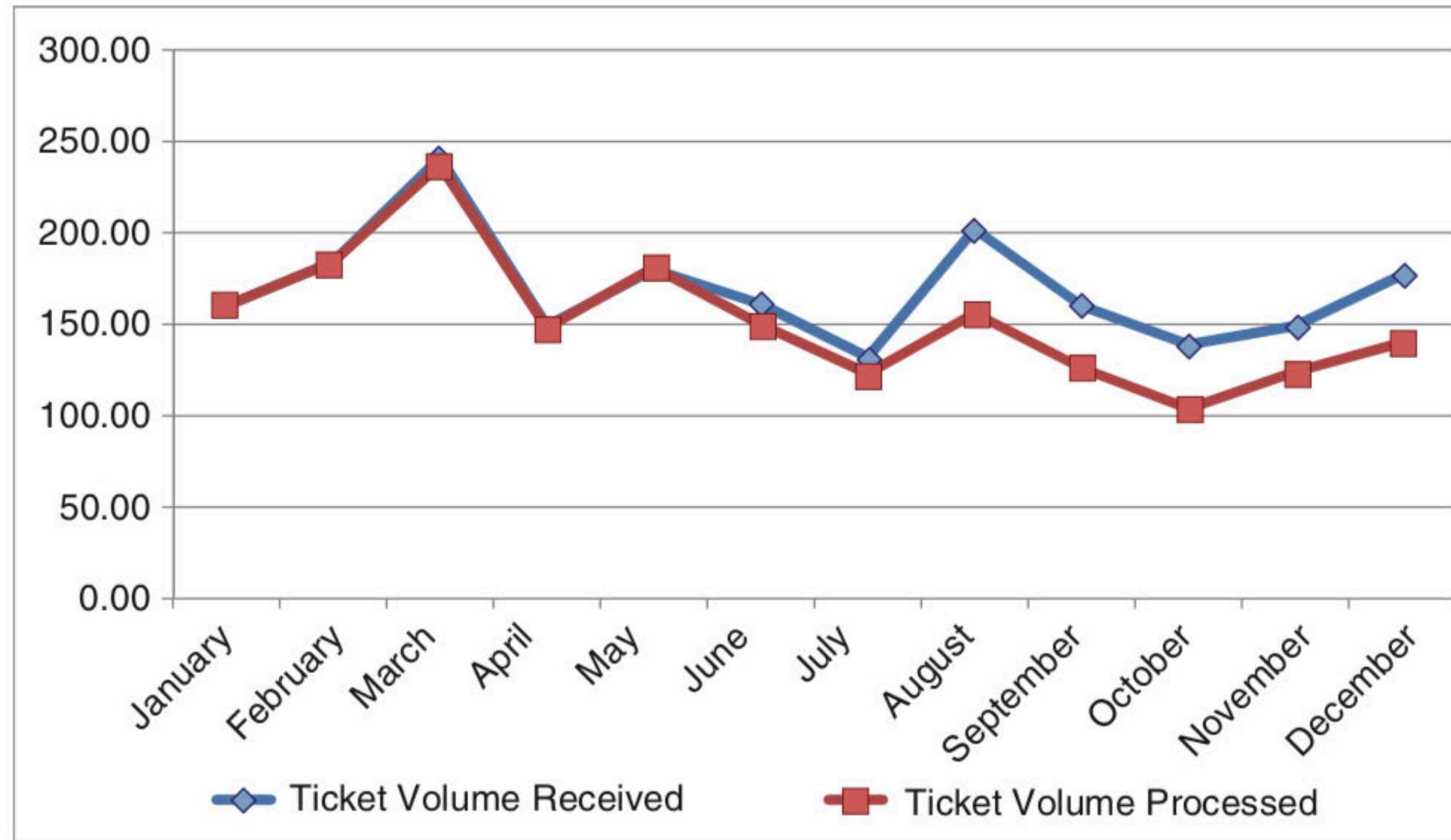


FIGURE 3.17 Original graph

Please approve the hire of 2 FTEs

to backfill those who quit in the past year

Ticket volume over time



Data source: XYZ Dashboard, as of 12/31/2014 | A detailed analysis on tickets processed per person and time to resolve issues was undertaken to inform this request and can be provided if needed.

From default visual to visual narrative – erase clutter, add narrative annotation ...

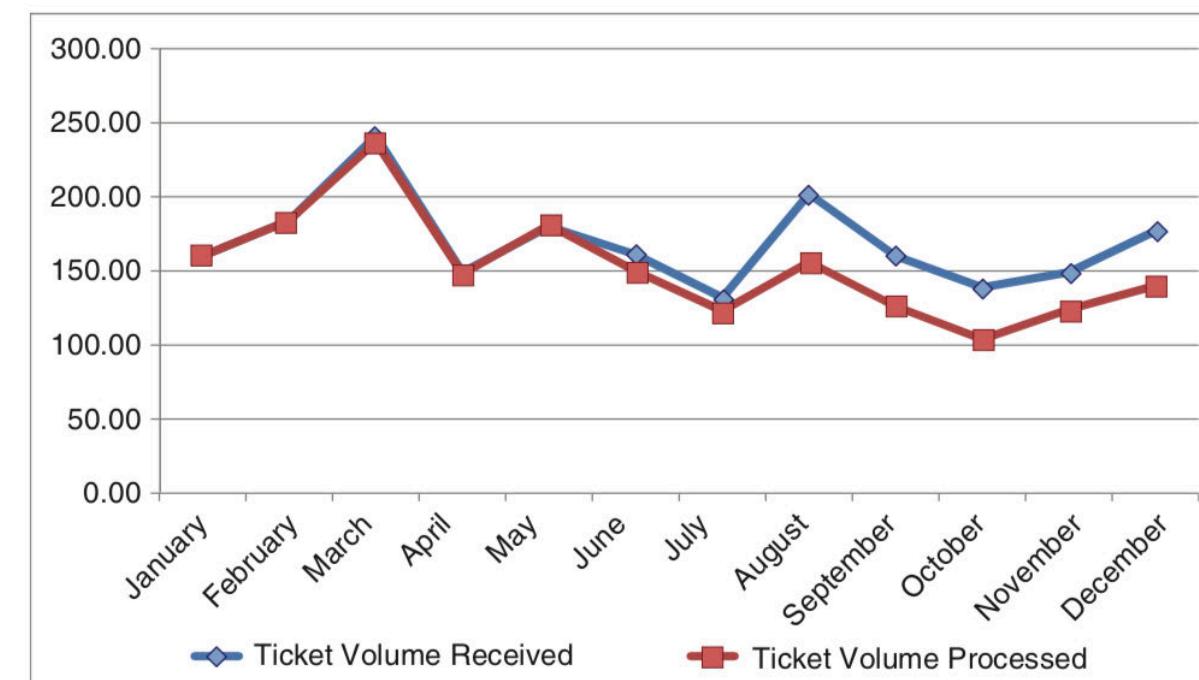


FIGURE 3.17 Original graph

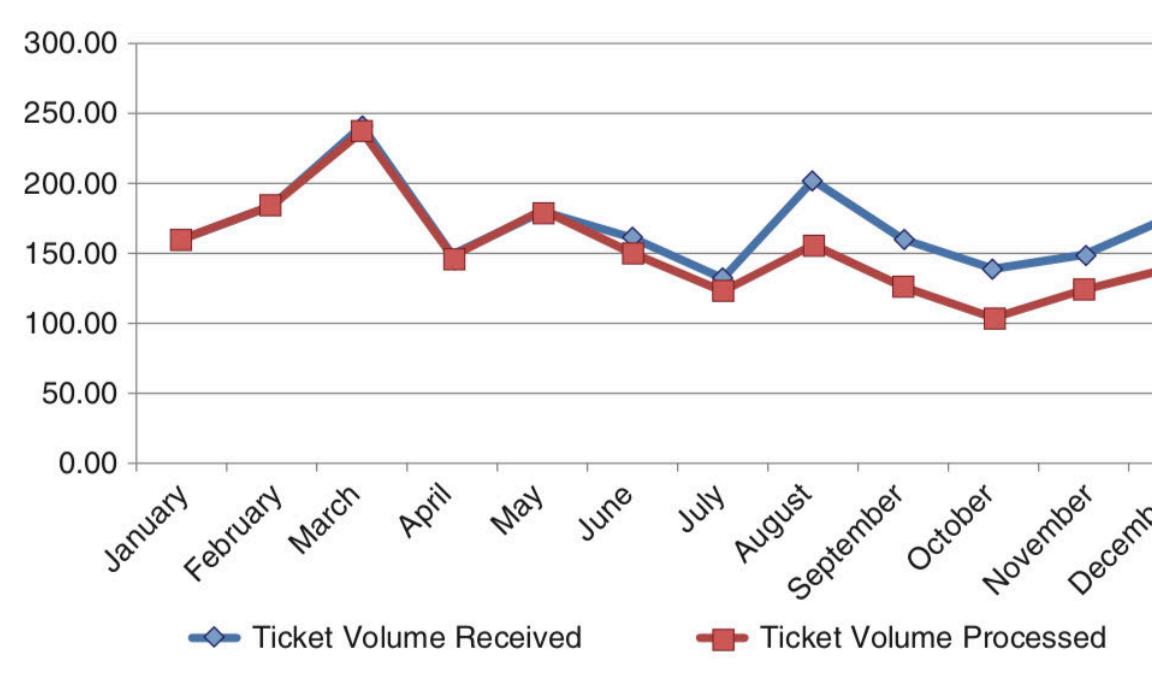


FIGURE 3.18 Remove chart border

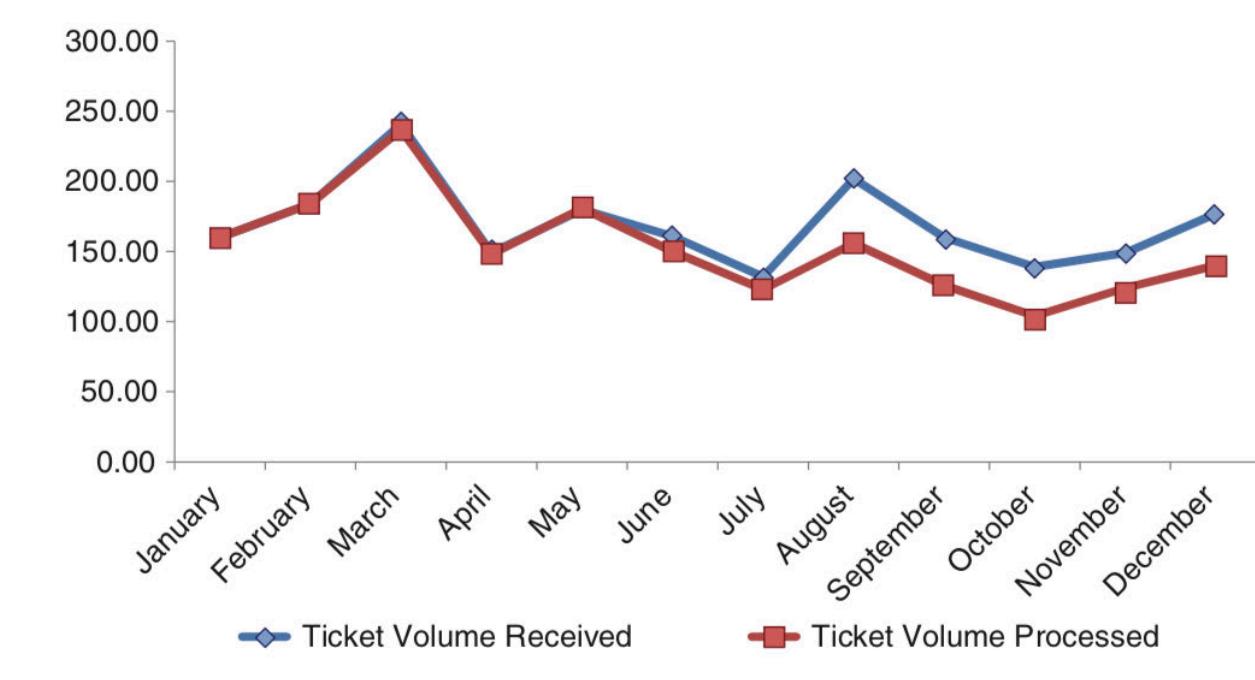


FIGURE 3.19 Remove gridlines

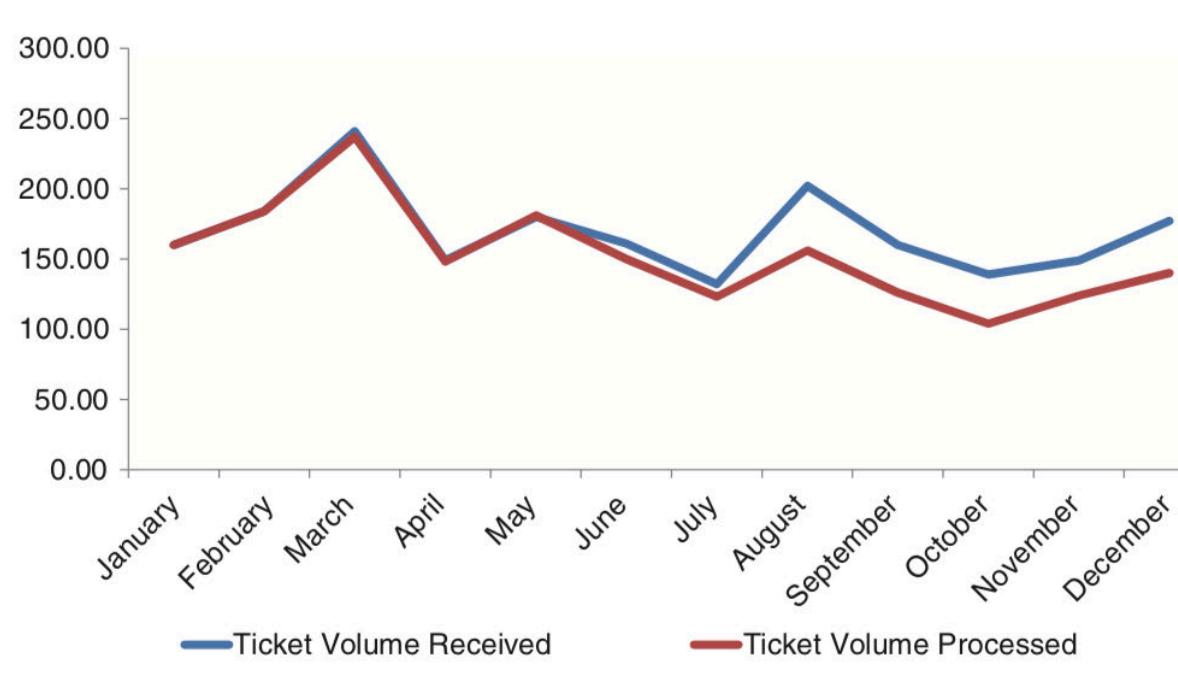


FIGURE 3.20 Remove data markers



FIGURE 3.21 Clean up axis labels

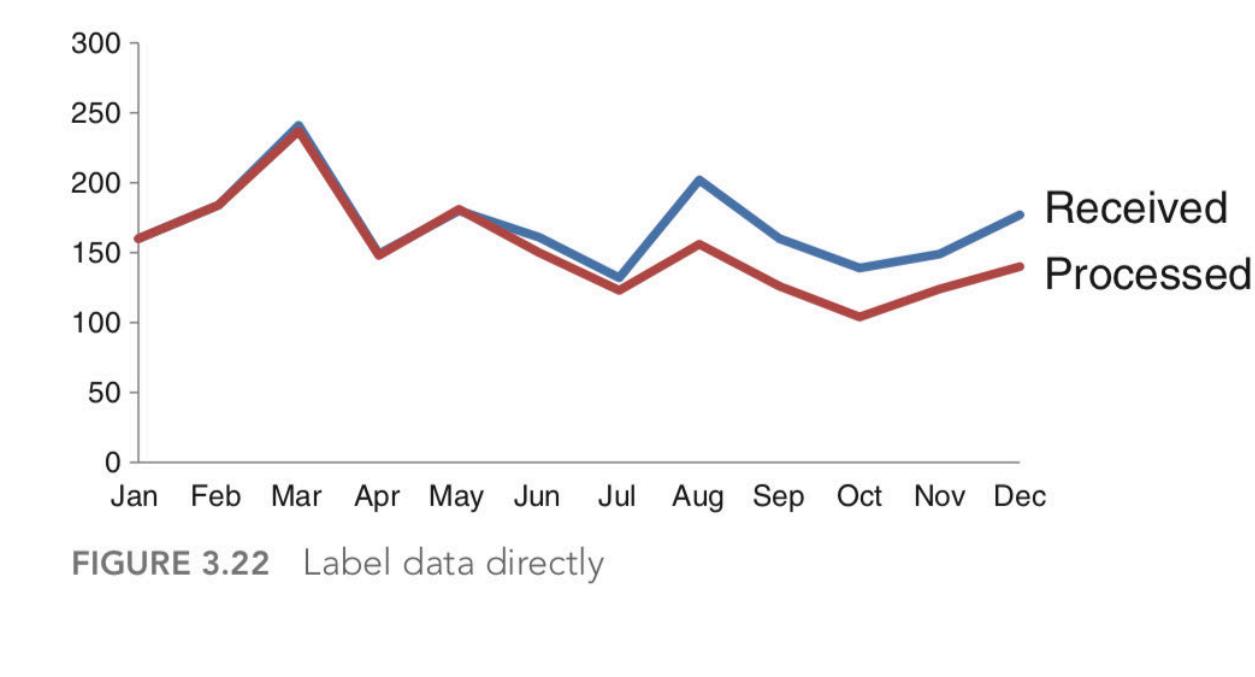


FIGURE 3.22 Label data directly

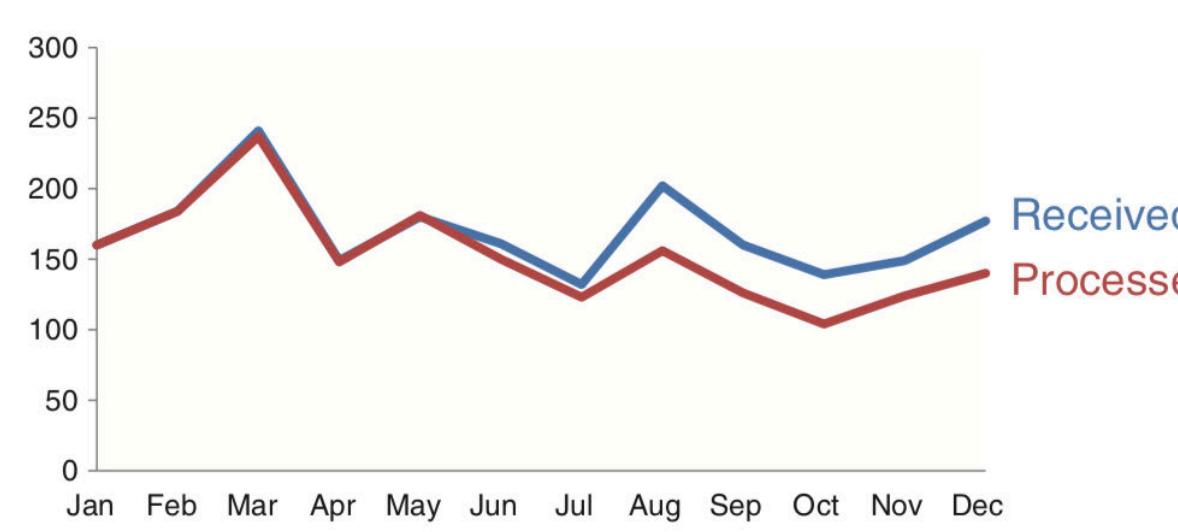


FIGURE 3.23 Leverage consistent color



FIGURE 4.11 First, push everything to the background

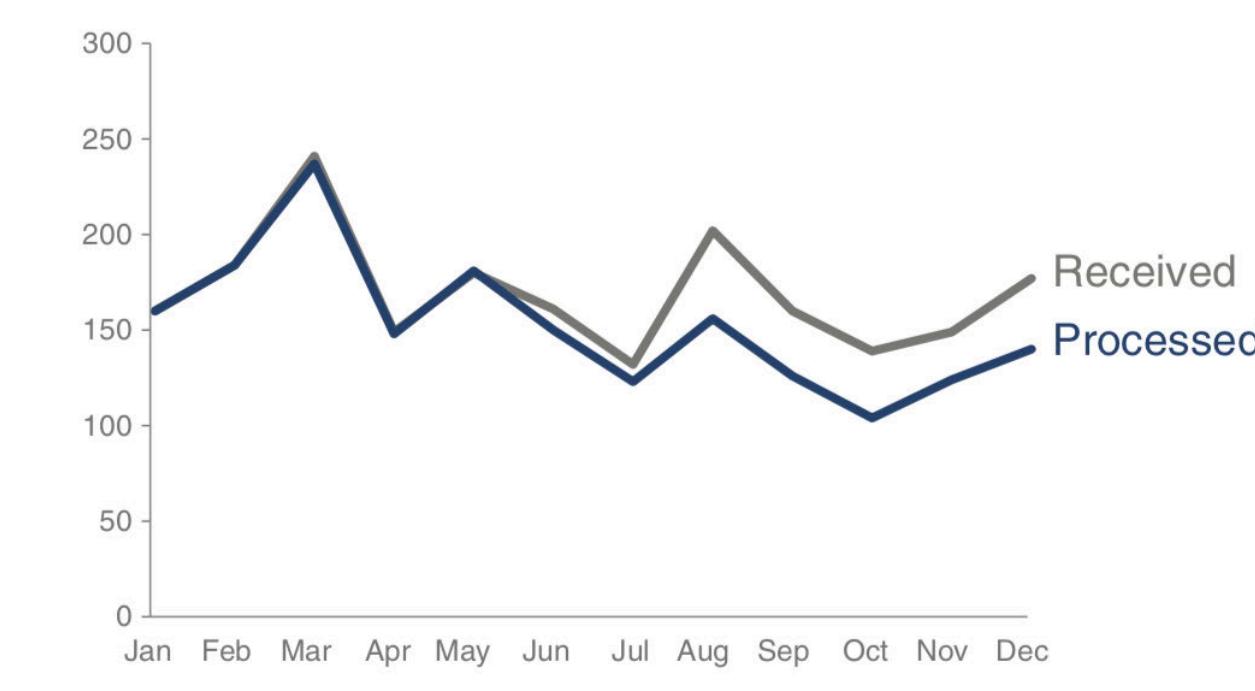


FIGURE 4.12 Make the data stand out

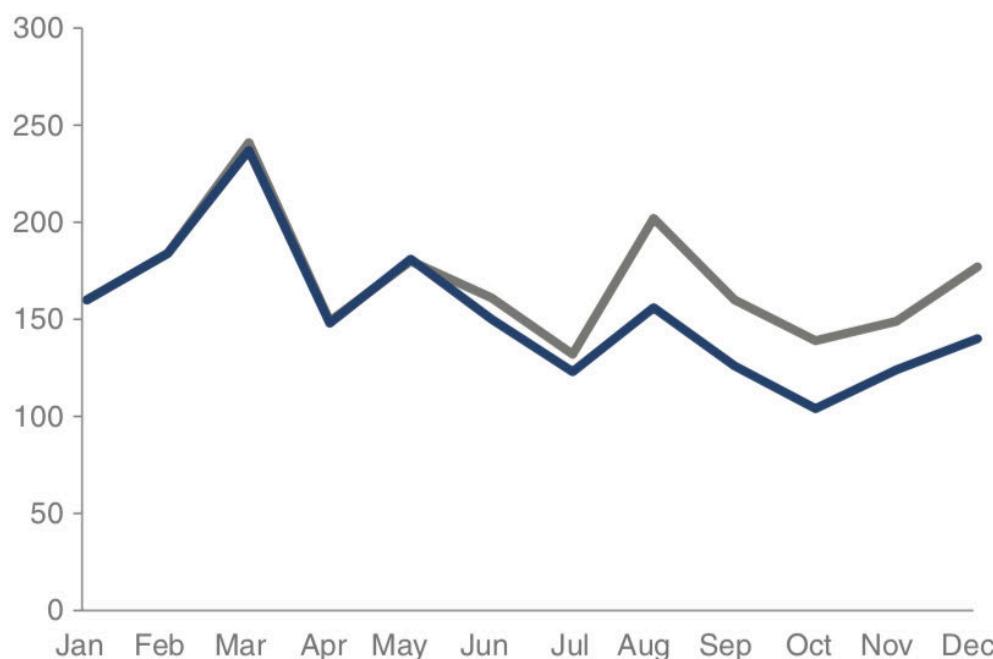


FIGURE 4.12 Make the data stand out



FIGURE 4.14 Data labels used sparingly help draw attention

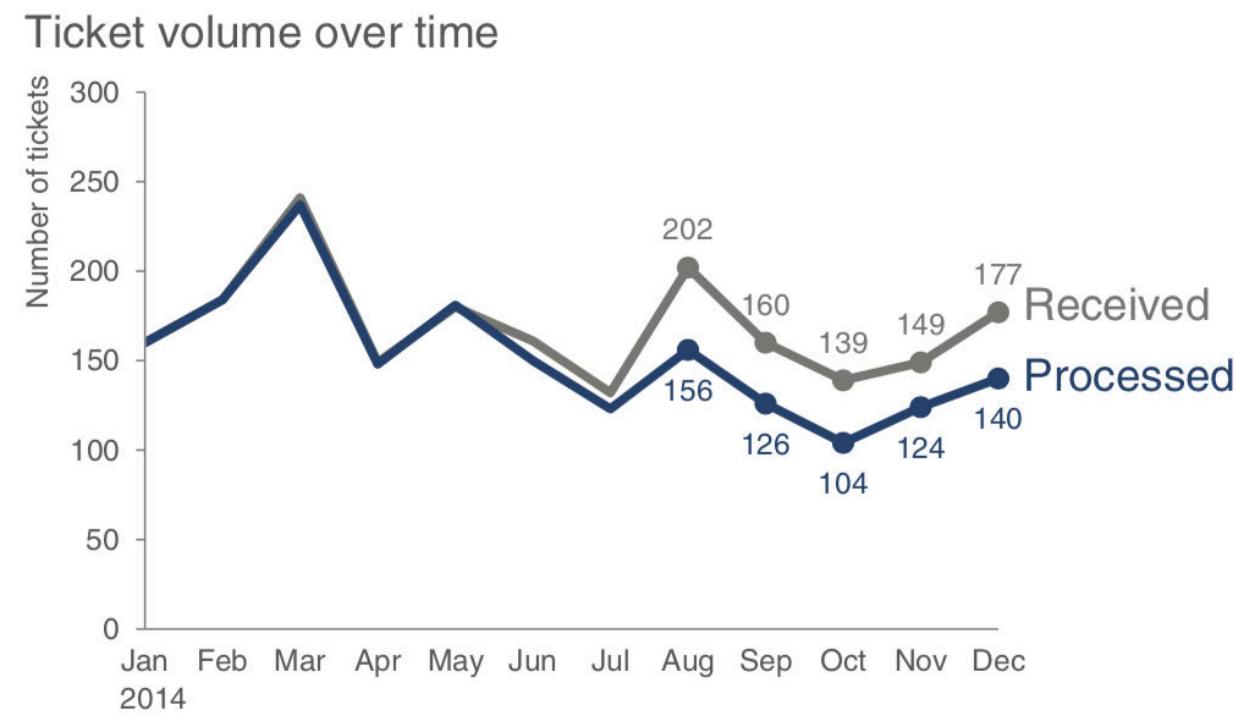
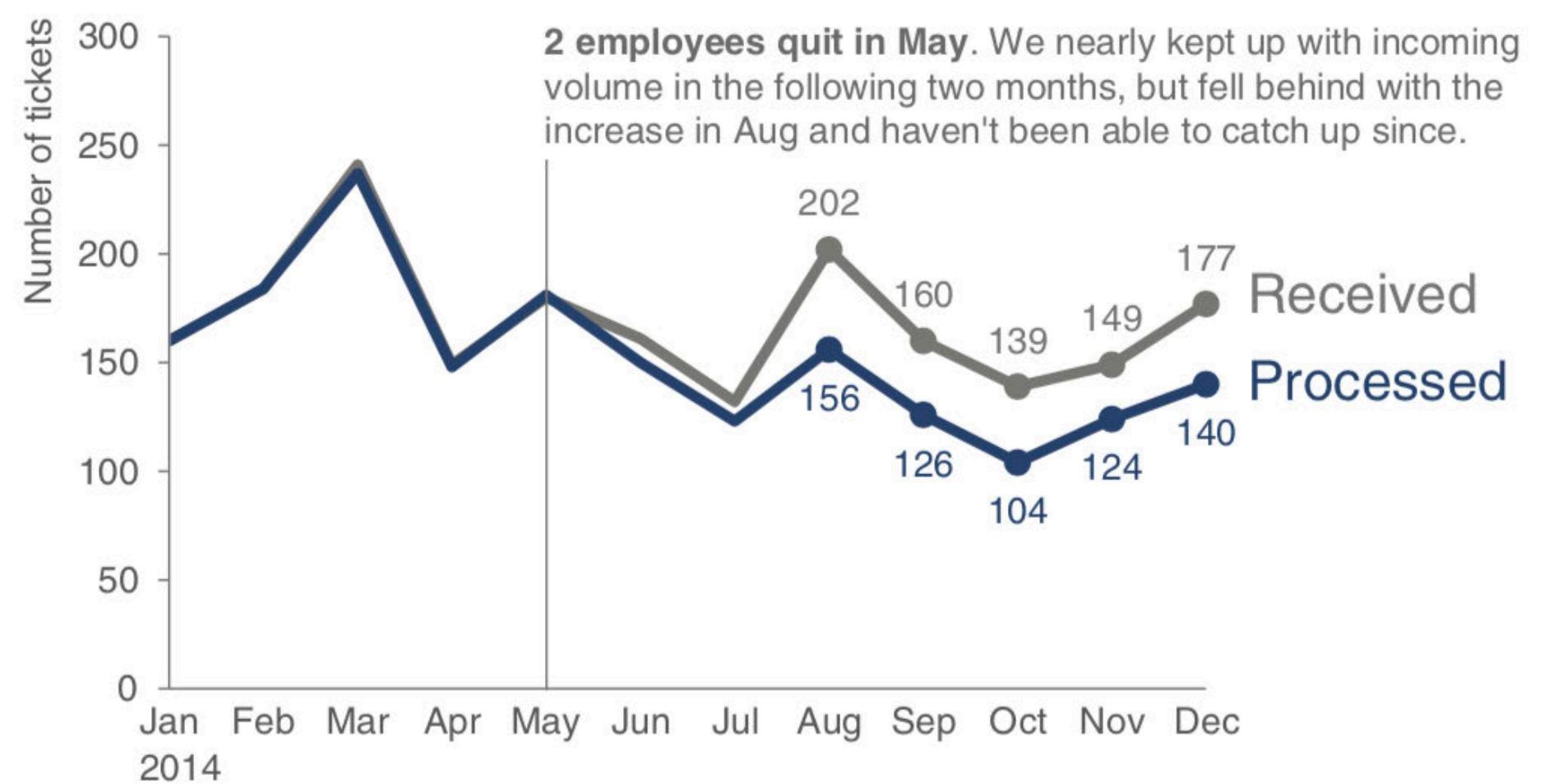


FIGURE 5.9 Use words to make the graph accessible

Please approve the hire of 2 FTEs

to backfill those who quit in the past year

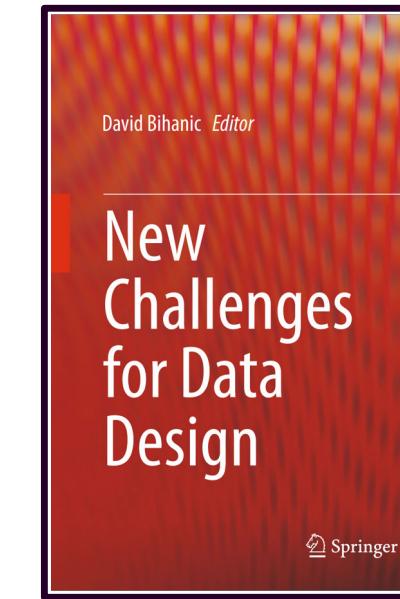
Ticket volume over time



Data source: XYZ Dashboard, as of 12/31/2014 | A detailed analysis on tickets processed per person and time to resolve issues was undertaken to inform this request and can be provided if needed.

FIGURE 5.10 Add action title and annotation

Graphics with hierarchies, layers



The New Aesthetic of Data Narrative

Lupi

Her work is part of the permanent collection of the Museum of Modern Art. She is co-founder and Design Director of Accurat, a data-driven design firm. After receiving her master's degree in Architecture, she earned her PhD in Design at Politecnico di Milano.

Create displays with layers, hierarchies

Her process involves layering multiple sub-narratives over a main construct. To achieve this multilayered storytelling with data ... everything depends on the concept of layering, establishing hierarchies, and making them clear.

DATA HUMANISM

SMALL ~~big~~ data
data bandwidth ~~quality~~

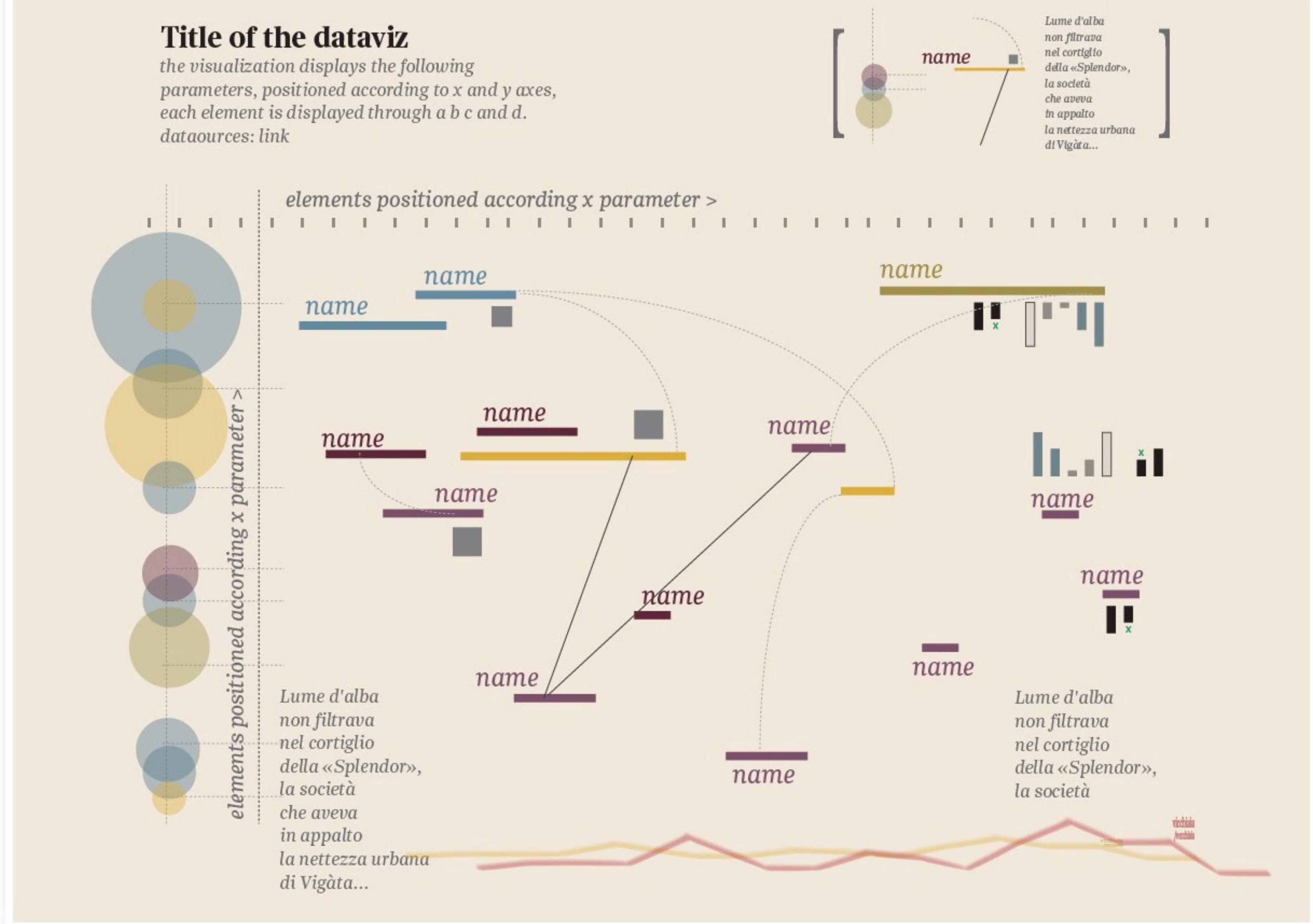
imperfect ~~infallible~~ data
subjective ~~impartial~~ data
inspiring ~~descriptive~~ data
serendipitous ~~predictive~~ data

data conventions ~~possibilities~~
data to simplify complexity ~~/ depict~~
data processing ~~drawing~~
data driven design

SPEND ~~save~~ time with
data is numbers ~~people~~
data will make us more efficient ~~human~~.

Title of the dataviz

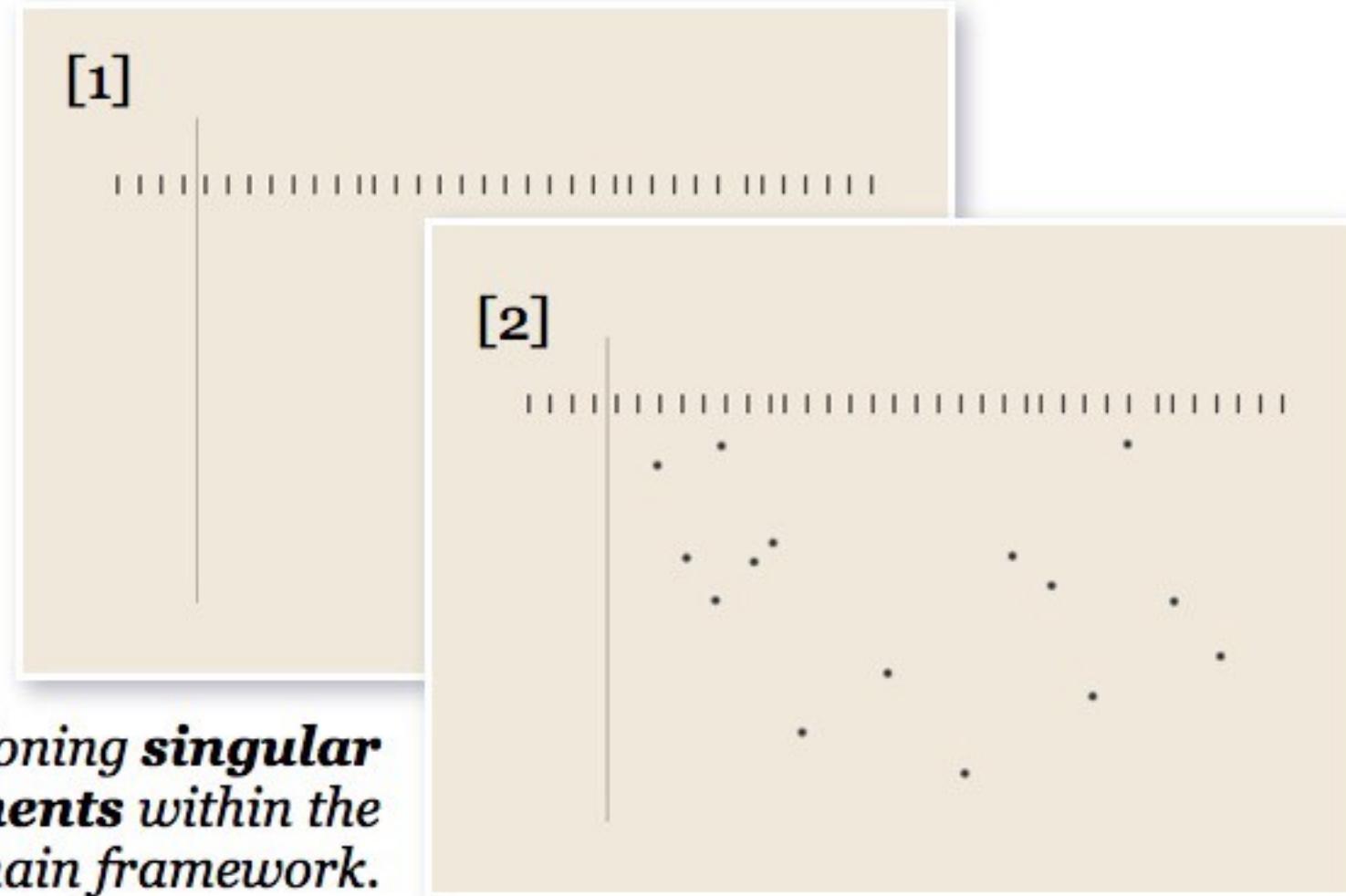
the visualization displays the following parameters, positioned according to x and y axes, each element is displayed through a b c and d. datasources: link



Layering, and making hierarchies clear

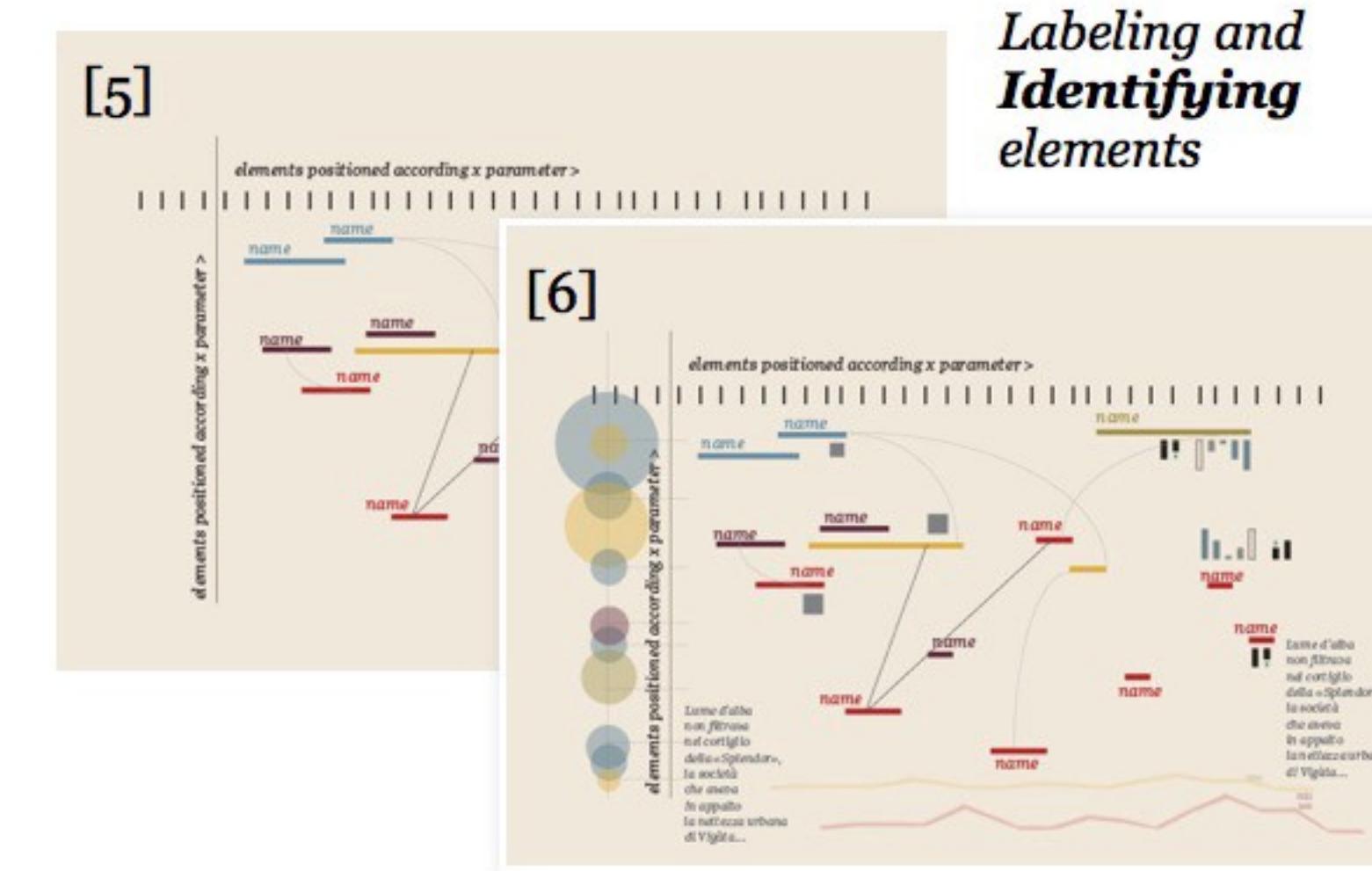
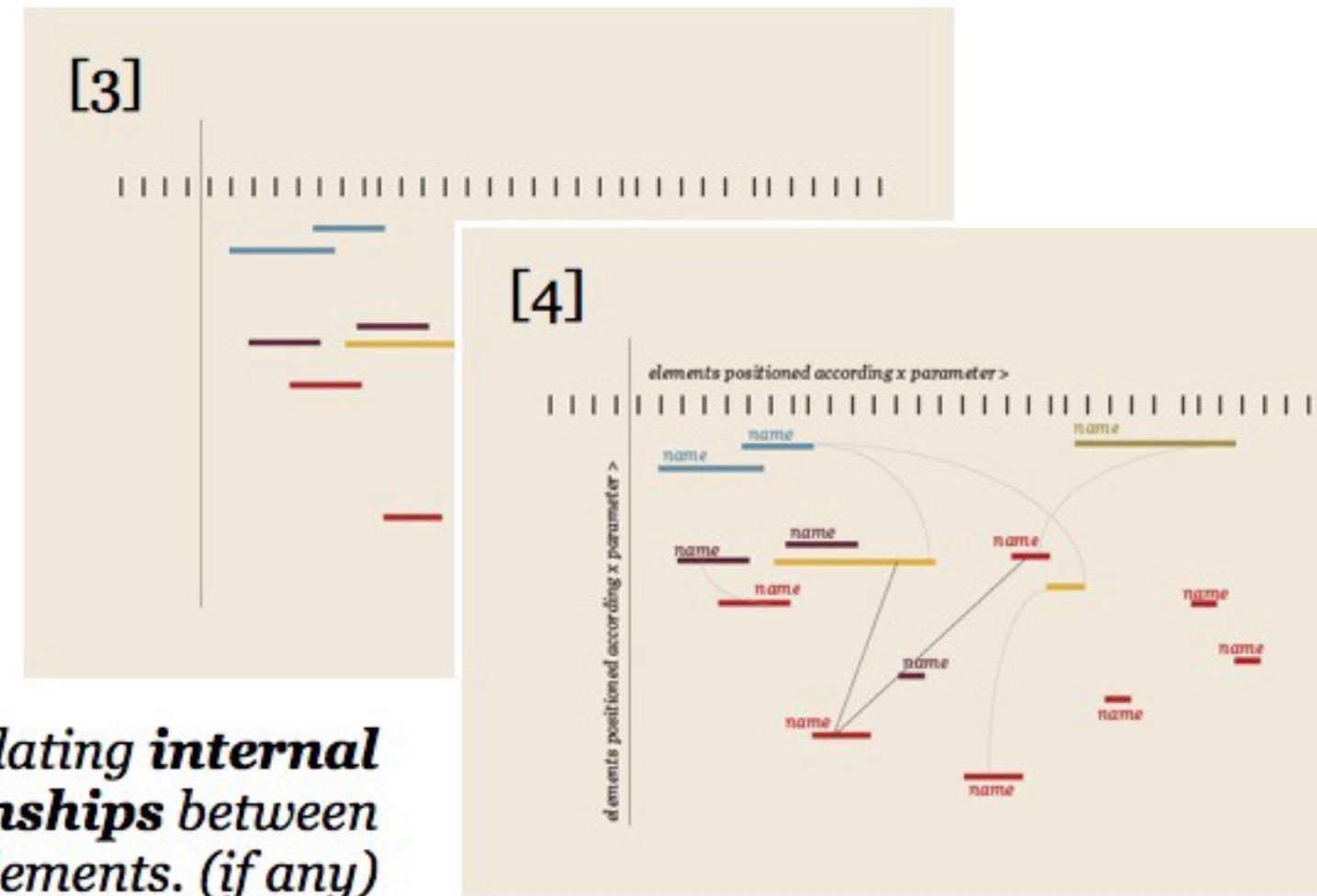
Composing the main architecture of the visualization

Positioning singular elements within the main framework.



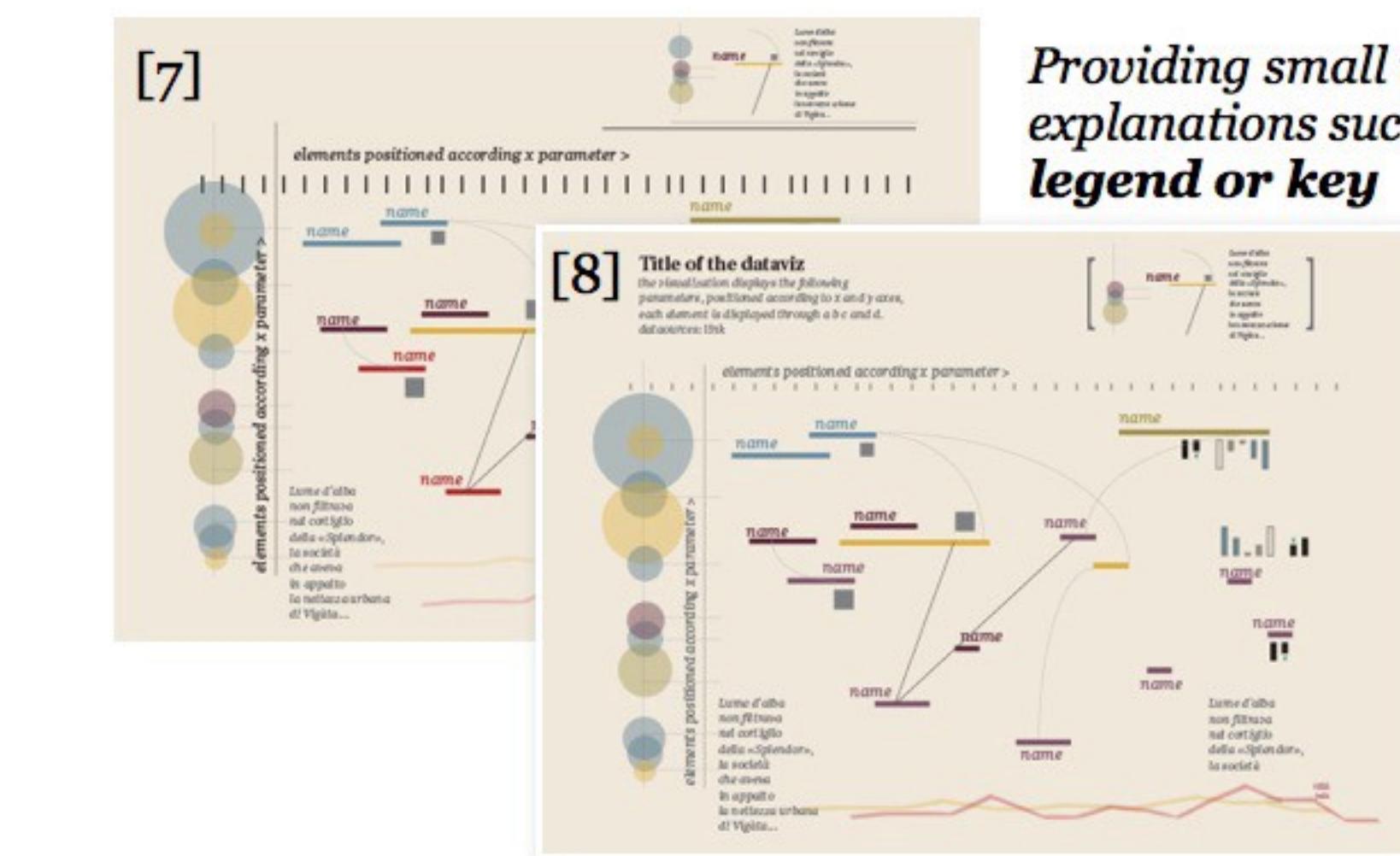
Constructing shaped elements of dimensionality and form

Elucidating internal relationships between elements. (if any)



Labeling and Identifying elements

Supplementing the greater story through the addition of “minor or tangential tales” elements

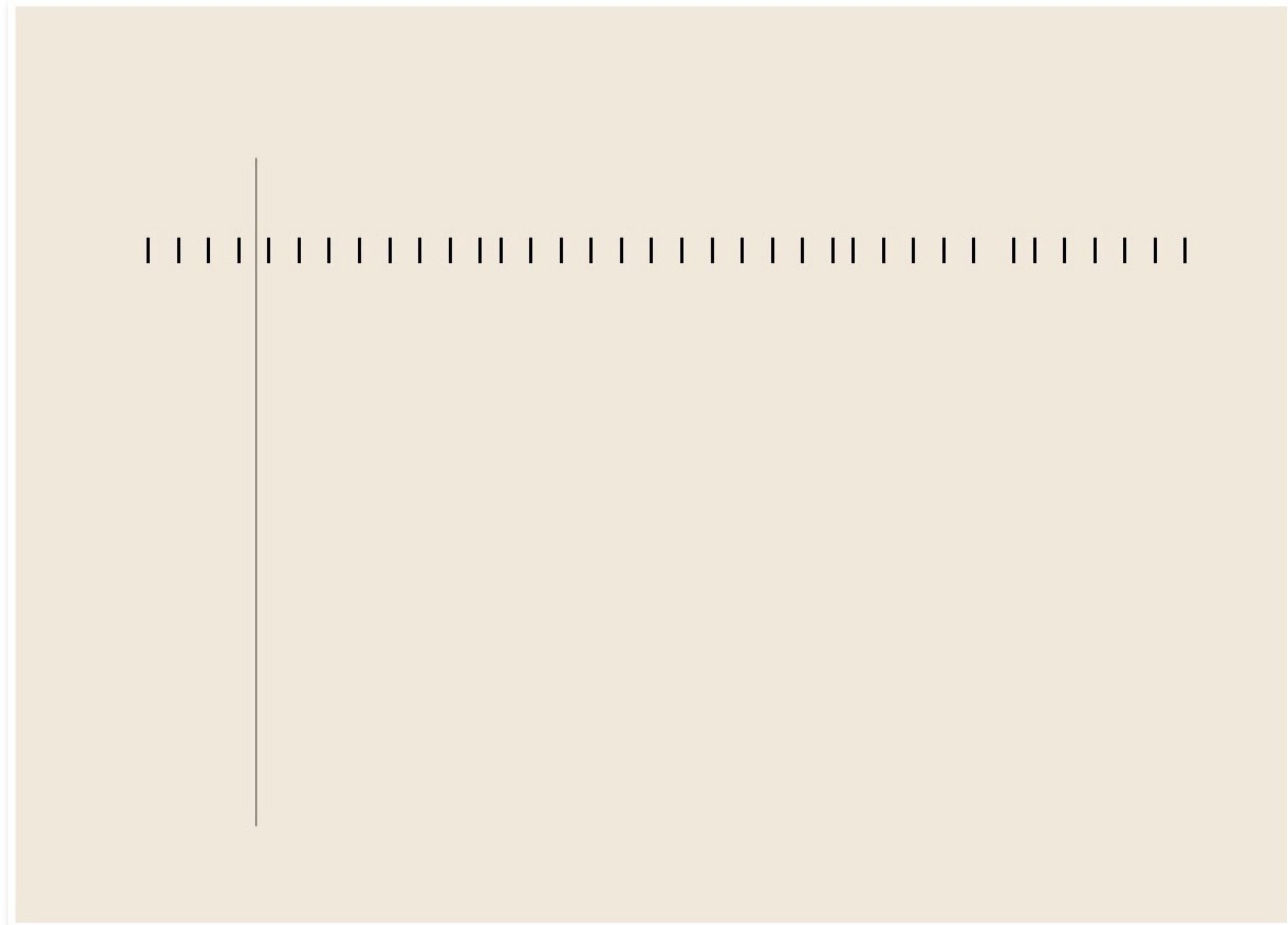


Providing small visual explanations such as a legend or key

Fine-tuning and stylizing of elements shapes, colors, and weights to make hierarchies pop out.

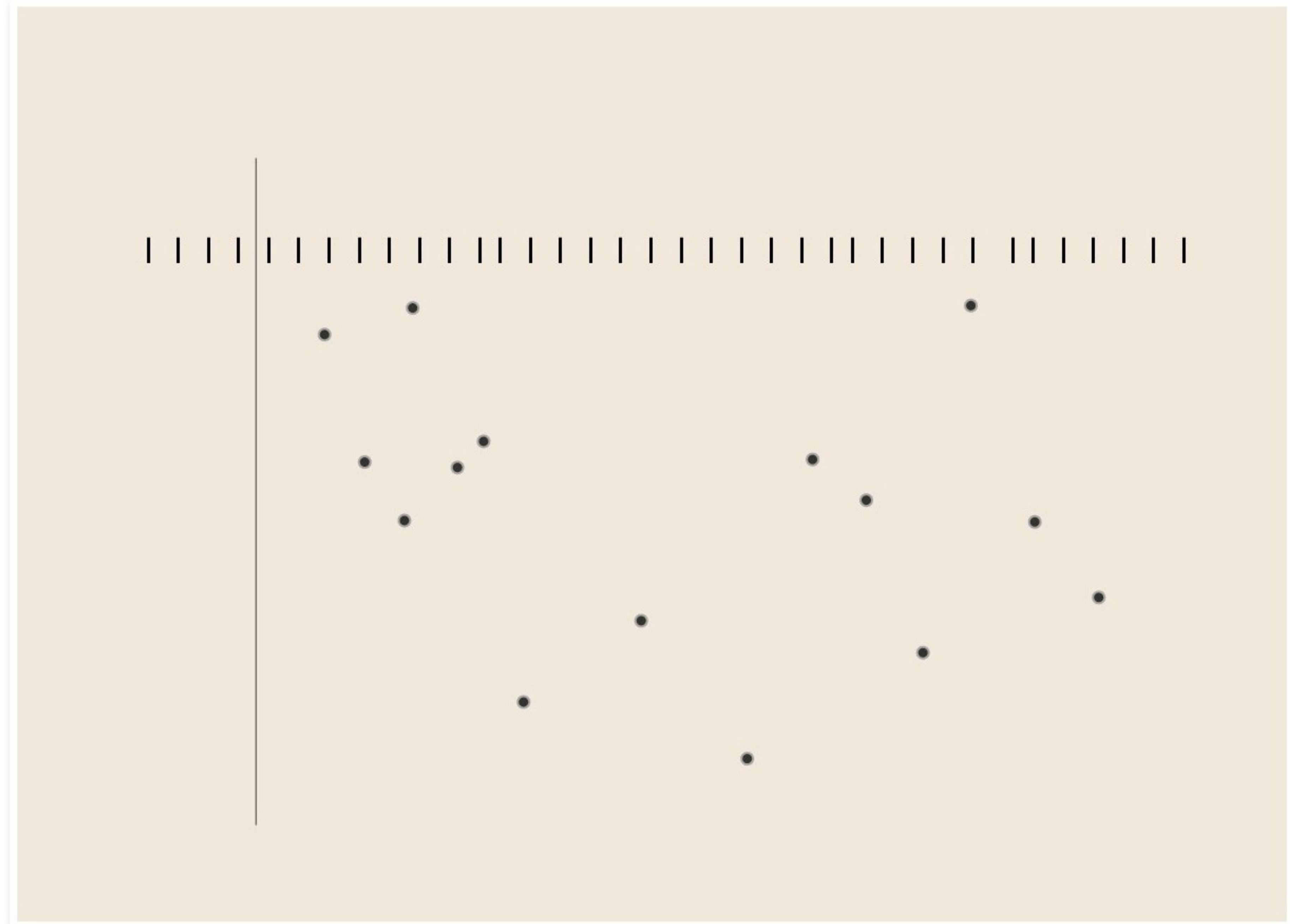
[1] Composing the main architecture of the visualization

Composing the main architecture: this acts as the formalized base through which the main story will be mapped and displayed, upon this, one will see the most relevant patterns emerging from the story: the essential “map” that conceptually identifies where we are. This base is essentially a matrix or pattern that will serve as our organizer. It may be composed of cells, or distances, or other interrelated multiples.



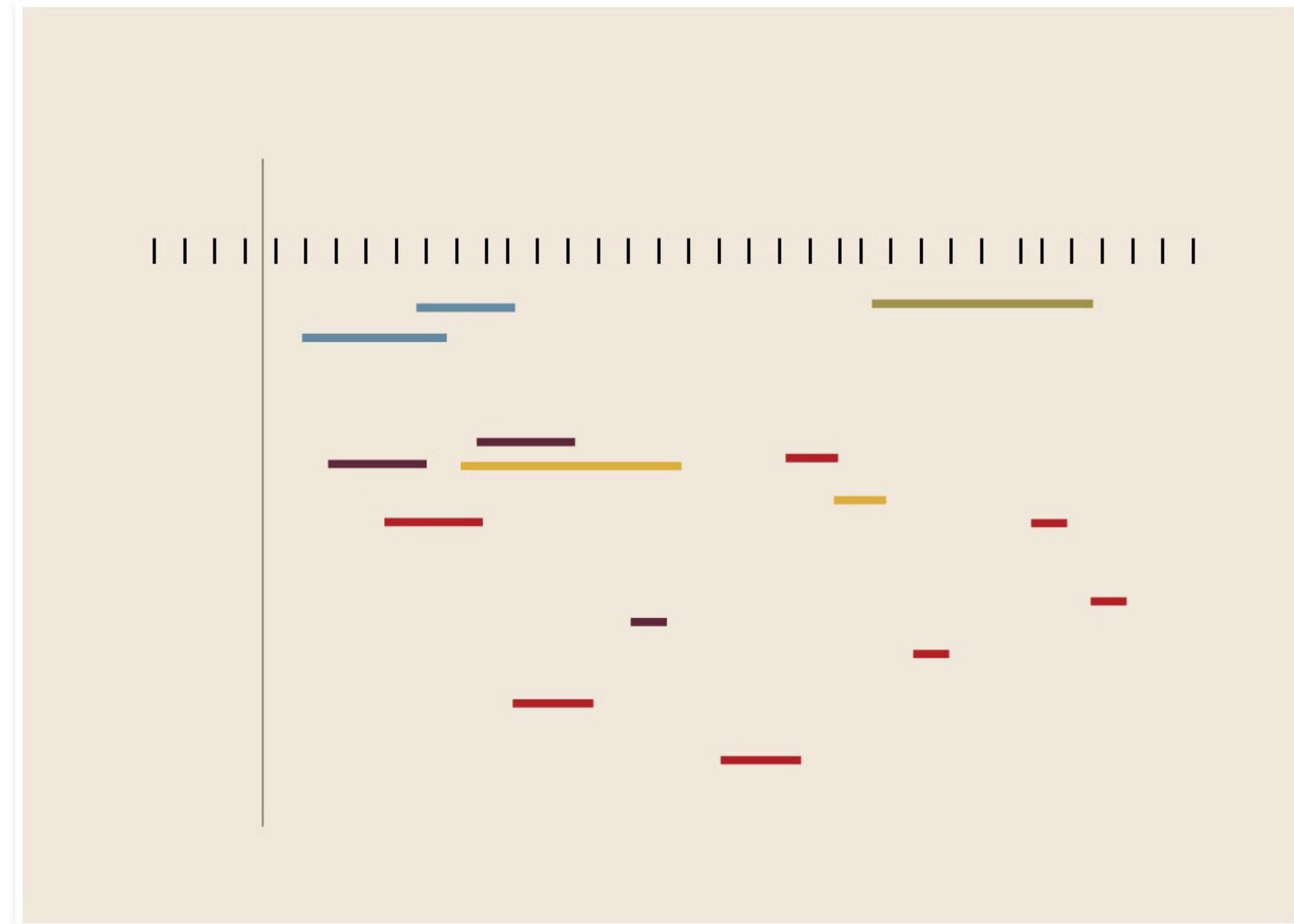
[2] Positioning singular elements within the main framework.

This process will test the effectiveness of the main architecture; the placement of elements reveals or confirms weaknesses and strengths, which may lead to modification of the main architecture.



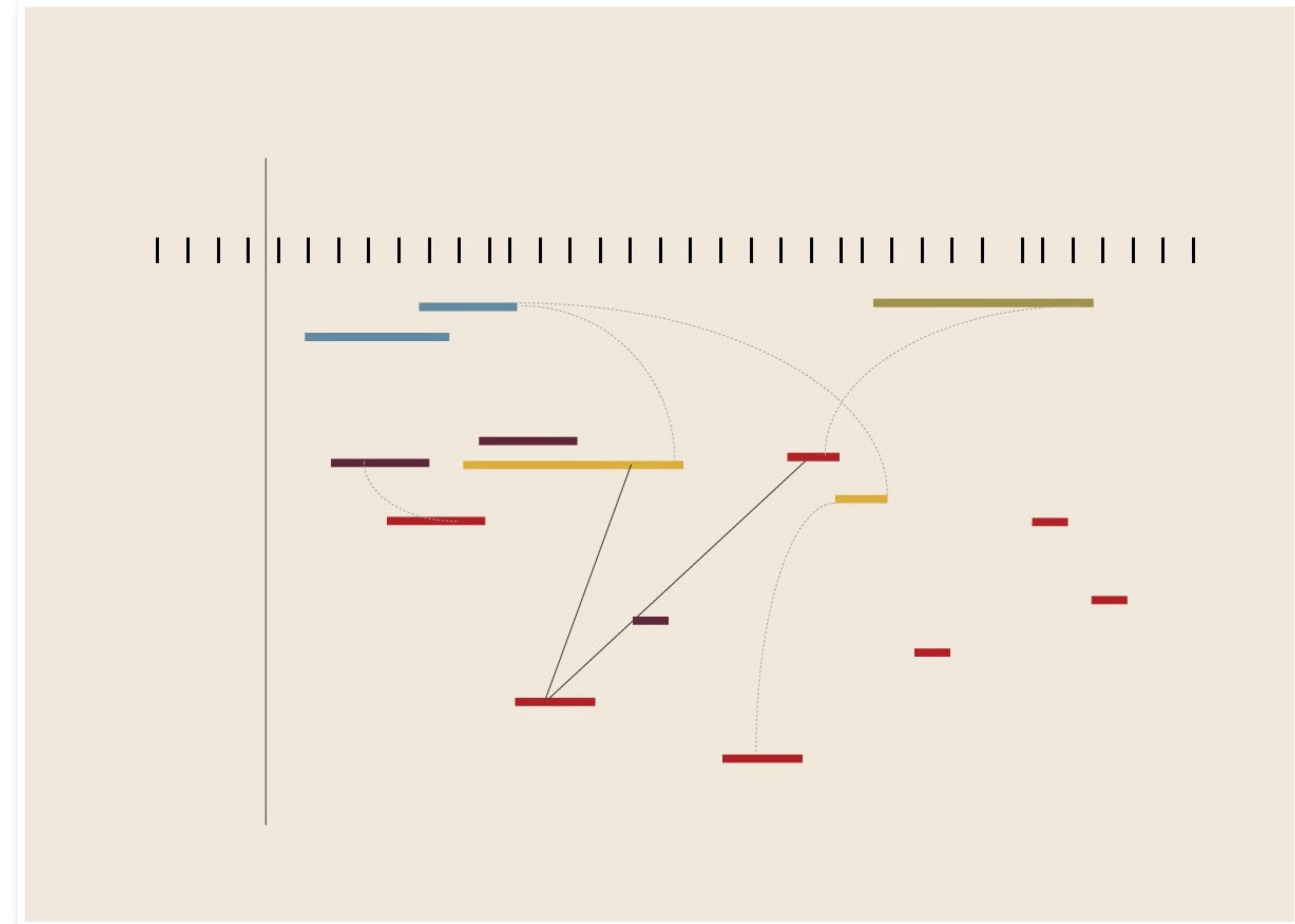
[3] Constructing shaped elements of dimensionality and form

Constructing shaped elements of dimensionality and form (essentially polygons) with quantitative and qualitative parameters and positioning these within the main architecture. As these elements have form they must also be identified through colors according to opportunities to establish categorizations, thus advancing clarity and relationships that serve to enhance the story.



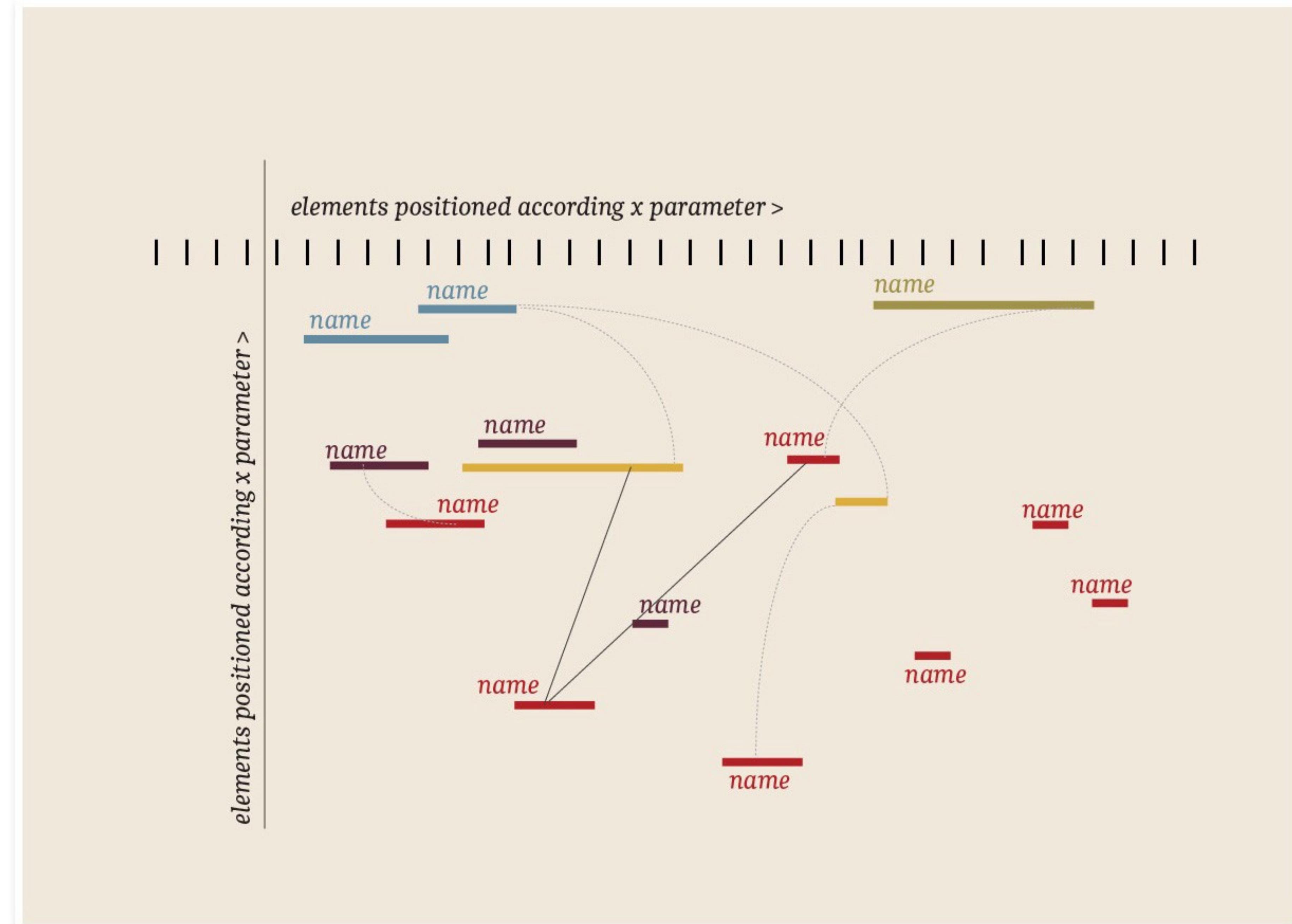
[4] Elucidating internal relationships between elements (if any)

These links, directives, and qualifiers serve to give the story a comprehensive texture and correlate dependencies within the story.



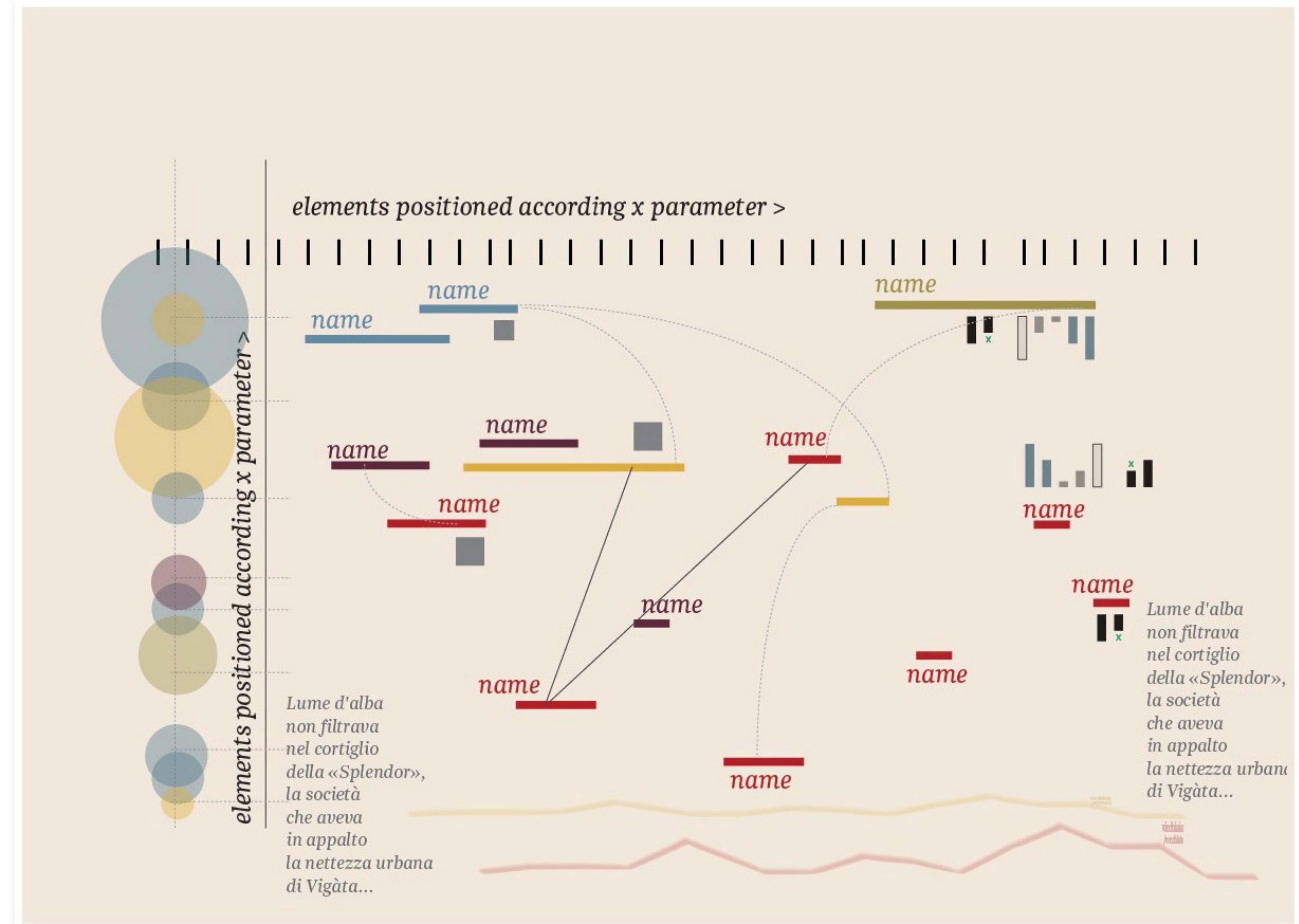
[5] Labeling and identifying

Through the addition of explanatory labels and short texts we provide requisite last mile clarity throughout the presentation.



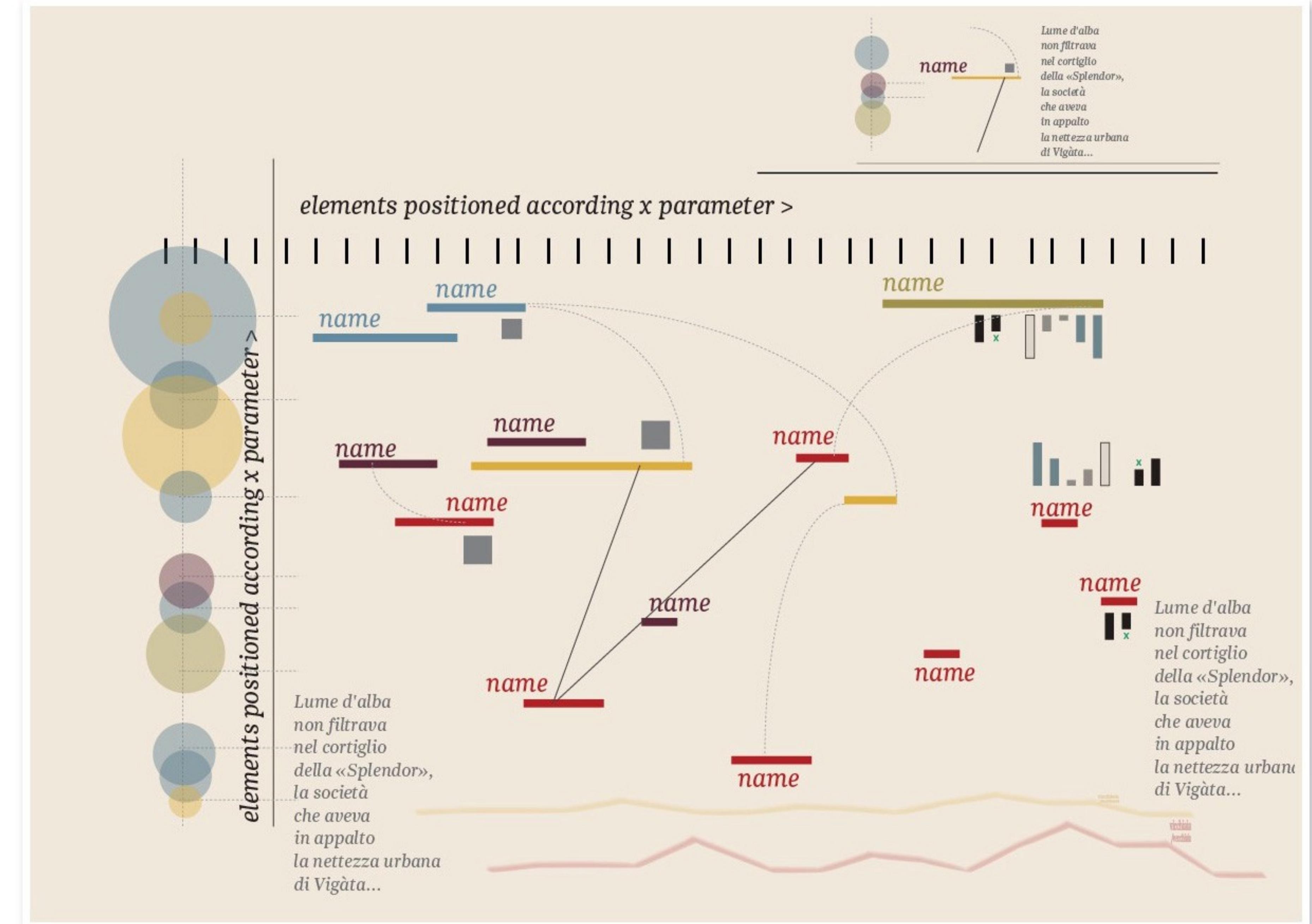
[6] Supplementing the greater story through the addition of minor or tangential tales elements.

We consider this a very important step to contextualize the phenomena in a wider world. These components link the story to external ideas, other times, or other places. Elements that are rendered here may come from very diverse sources—analysis that is undertaken once we have strongly established the core story. These elements, which may take the form of small images, textual components, graphic symbols, etc., are to be located where they best help to enrich the overall comprehension: they must not distract from the main story.



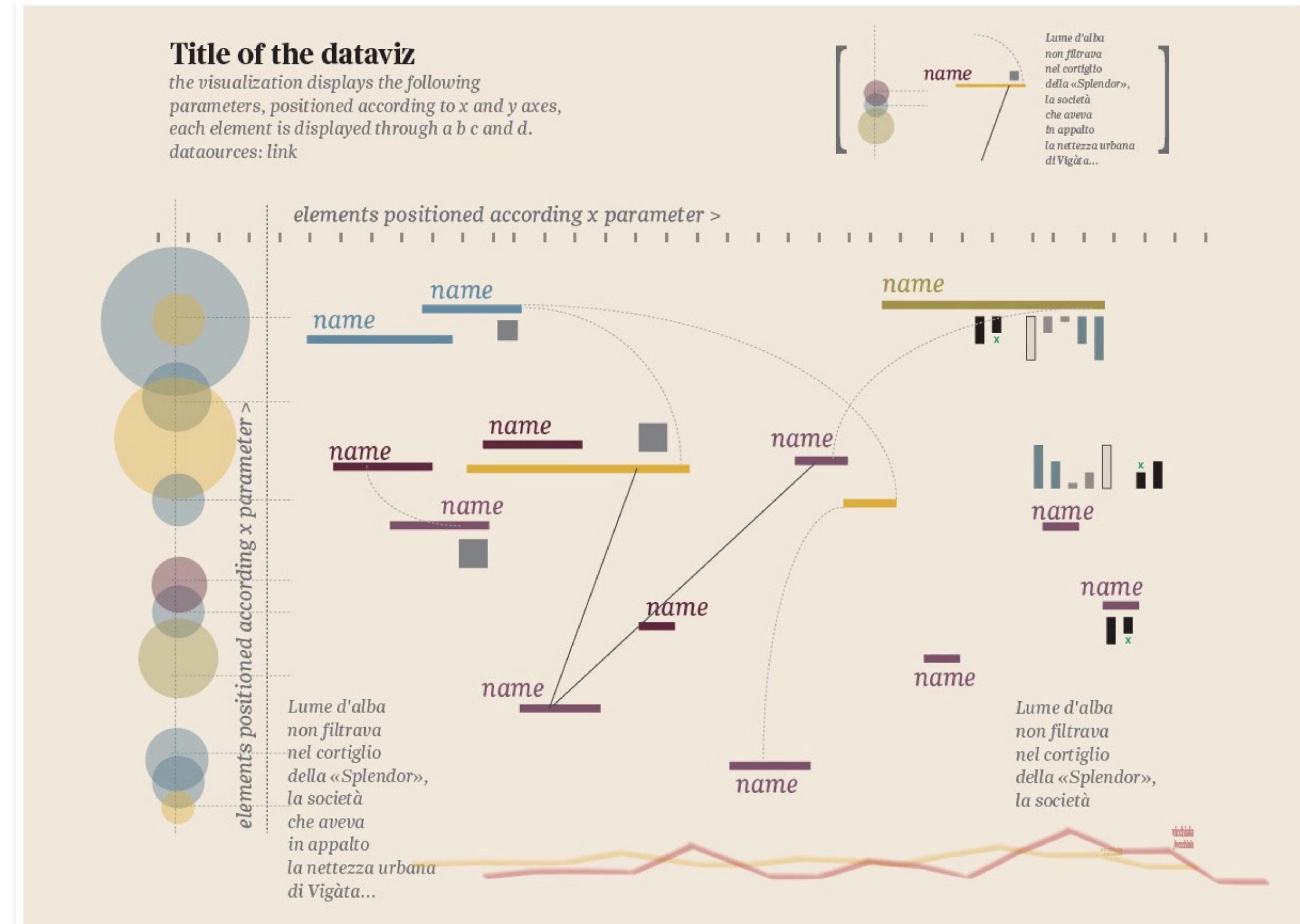
[7] Providing small visual explanations such as a legend or key that assists readers and the general public who may not be familiar with norms of data visualization.

These are composed to enlighten the layered idea of the visualization, often constructed as miniatures of the layers themselves. The process usually involves simplification of the general architecture (e.g. the x and y axes, base timelines, or map components) as well as minimal explicit shapes, colors, and dimensions of singular elements. These explanations also provide units of measurement for distances and volumes.



[8] Fine-tuning and stylizing of elements' shapes, colors, and weights to make hierarchies pop out.

By visually highlighting the most relevant elements and lightening the other background layers of information, we should be able to allow information to be selectively and sequentially revealed, helping readers discover stories by themselves and recognizing the patterns or interrelationships from one element within the story to another.



Building information density from common business graphics

Winner, Information is Beautiful Award

Lupi & Fragapane, et al

Lupi heads a design firm and this visualization is a representative example from her series of exploratory data-visualizations originally published for La Lettura, their Sunday cultural supplement, spanning two years.

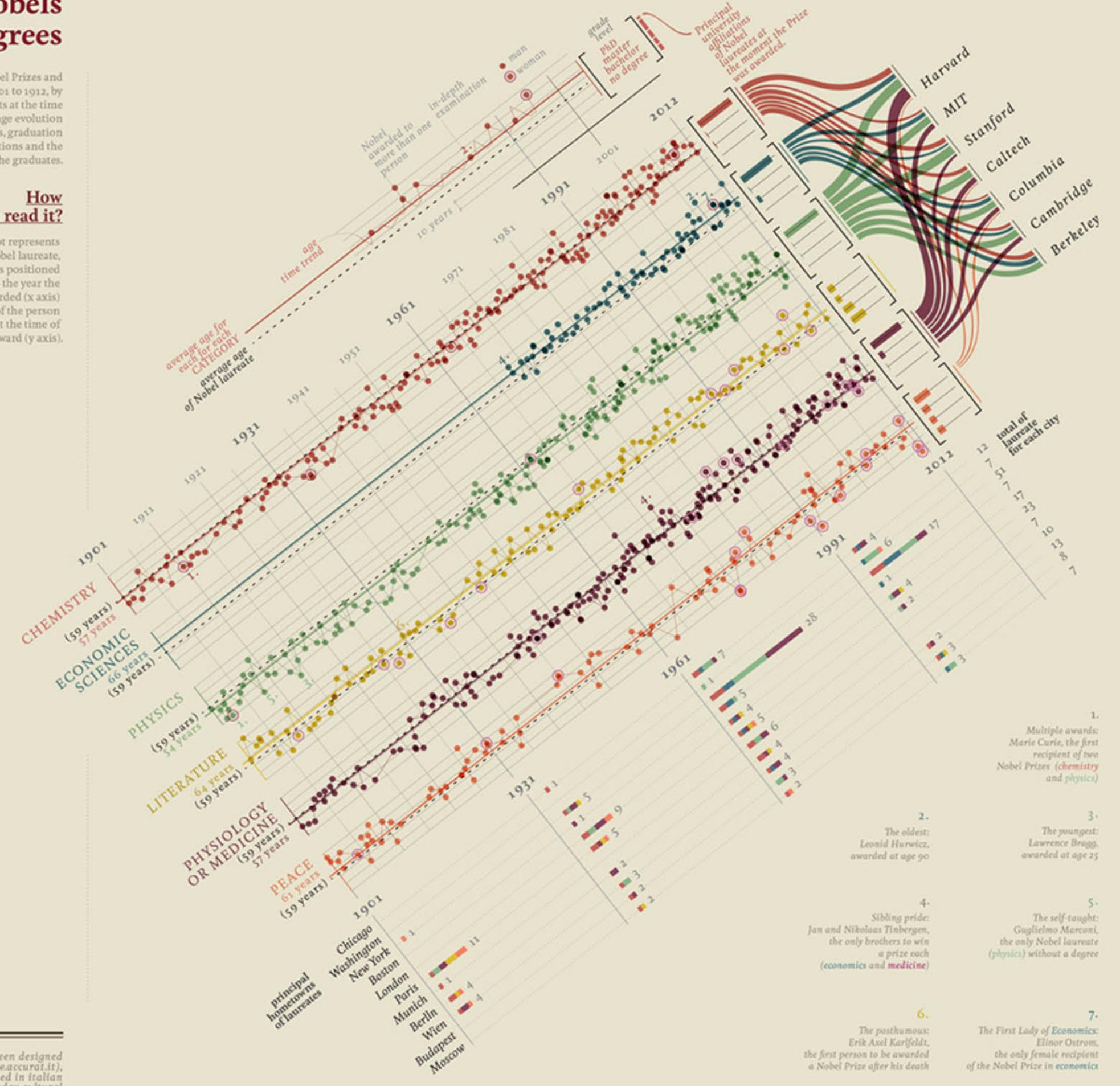


Nobels no degrees

This visualization explores Nobel Prizes and graduate qualifications from 1901 to 2012, by analysing the age of recipients at the time prizes were awarded, average age evolution through time and among categories, graduation grades, main university affiliations and the principal hometowns of the graduates.

How to read it?

Each dot represents a Nobel laureate, each recipient is positioned according to the year the prize was awarded (x axis) and age of the person at the time of the award (y axis).



Components are integrated with color, font—placement—lines, and annotation.

All four displays share data information. These relationships are integrated together using consistent color, in the case of prize categories,



and integrated using the relative placement of each display for both categories and other shared data. Placement of the histograms and Sankey diagram alongside the same categories in the central display reinforces the connection.

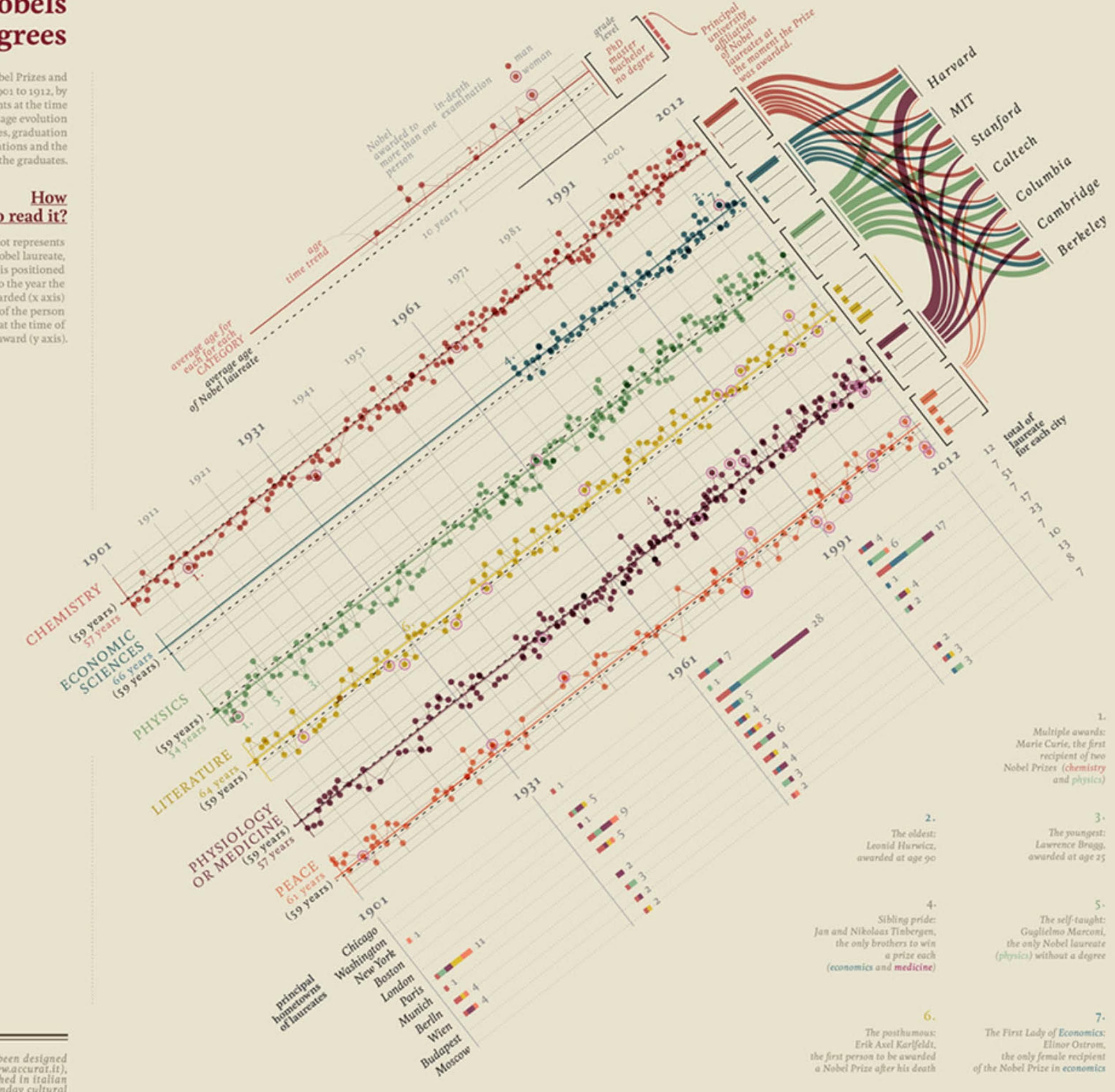
And notice, for example, that the per-era stacked bars share dates by aligning a common baseline with the x-axis of the central display.

Nobels no degrees

This visualization explores Nobel Prizes and graduate qualifications from 1901 to 1912, by analysing the age of recipients at the time prizes were awarded, average age evolution through time and among categories, graduation grades, main university affiliations and the principal hometowns of the graduates.

How to read it?

Each dot represents a Nobel laureate, each recipient is positioned according to the year the prize was awarded (x axis) and age of the person at the time of the award (y axis).



The visualization has been designed and produced by Accurat (www.accurat.it), and was originally published in Italian on *La Lettura* the Sunday cultural supplement of *Corriere della Sera*.

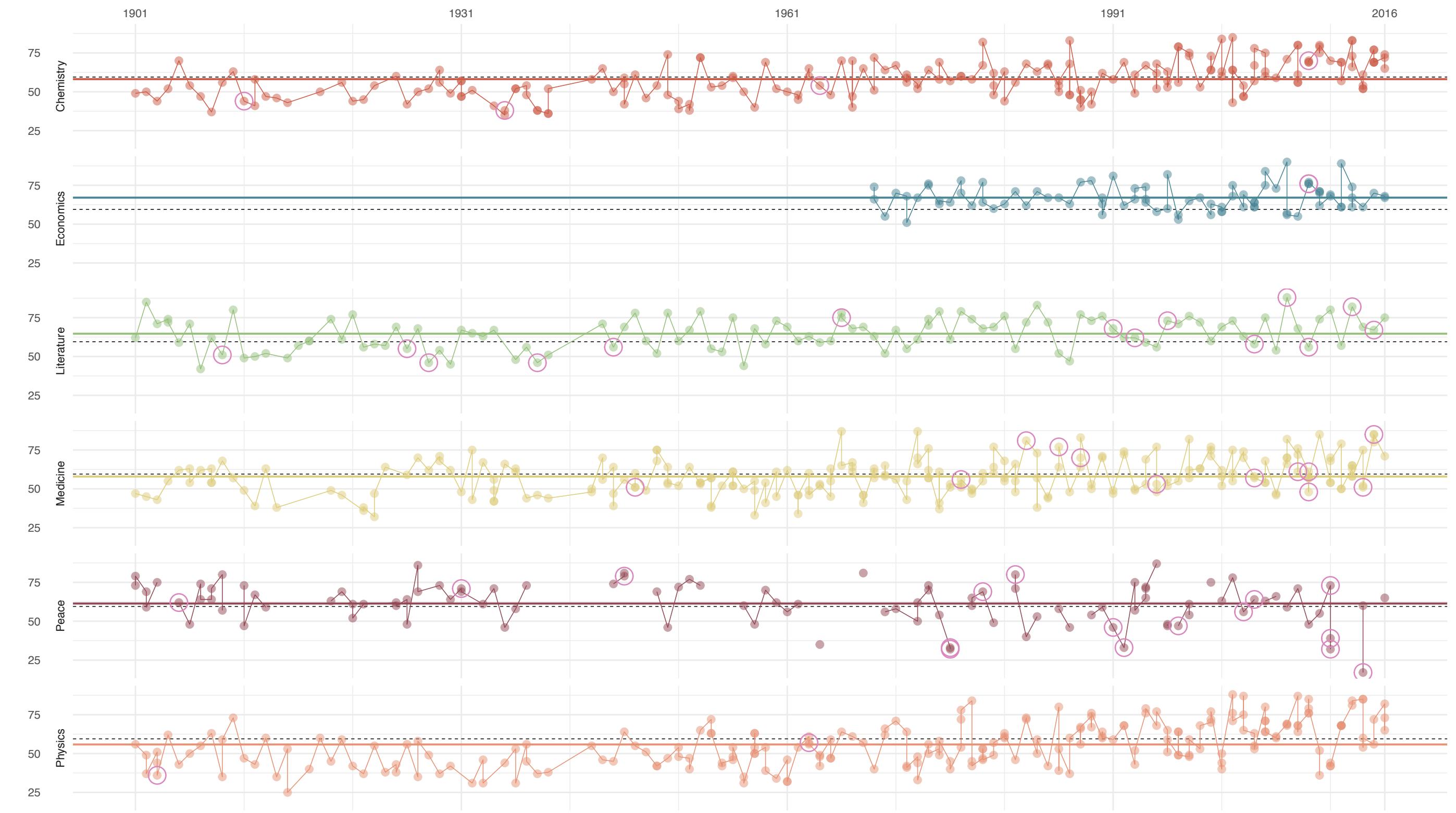
Approximating the components

Central display layering points, lines

Per-category histograms

Per-era stacked bar charts

Sankey diagram



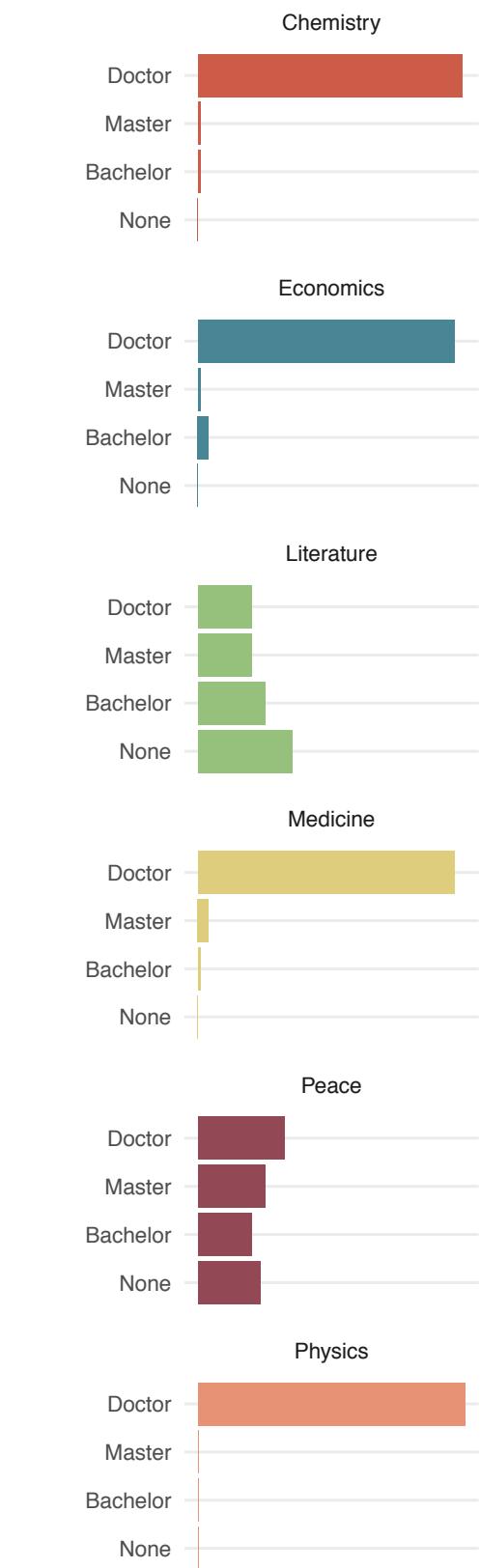
Approximating the components

Central display layering points, lines

Per-category histograms

Per-era stacked bar charts

Sankey diagram



Approximating the components

Central display layering points, lines

Per-category histograms

Per-era stacked bar charts

Sankey diagram



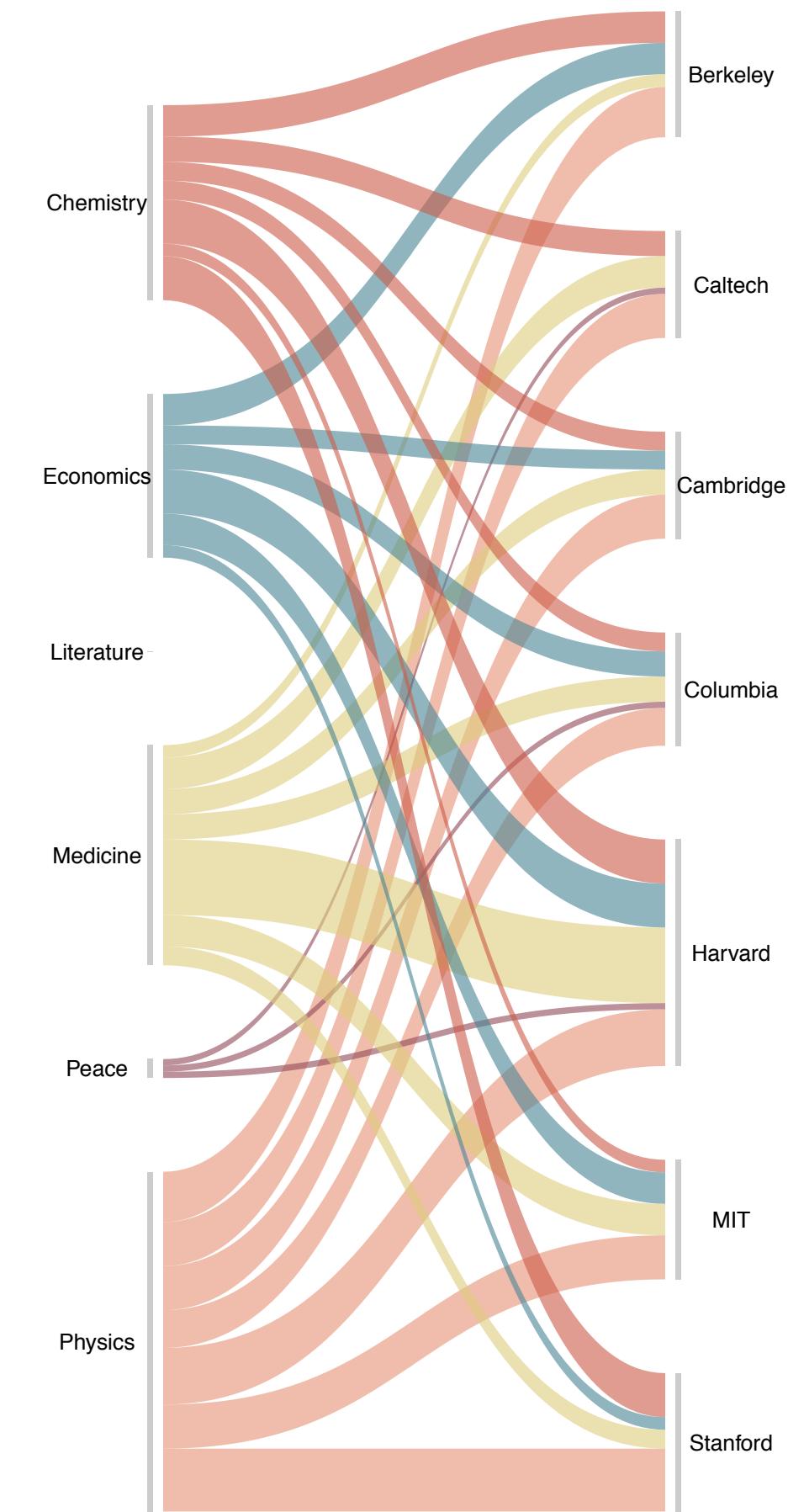
Approximating the components

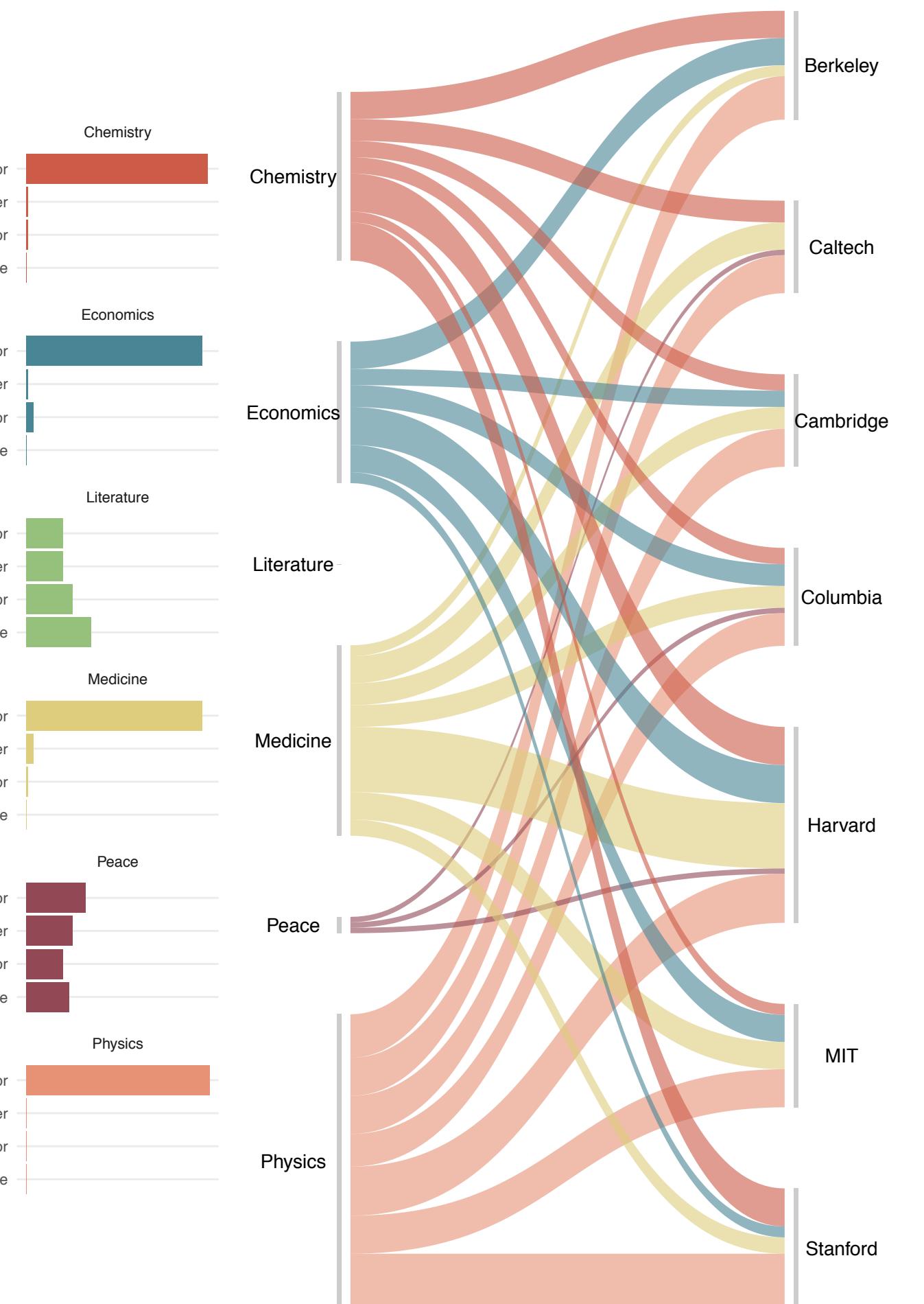
Central display layering points, lines

Per-category histograms

Per-era stacked bar charts

Sankey diagram



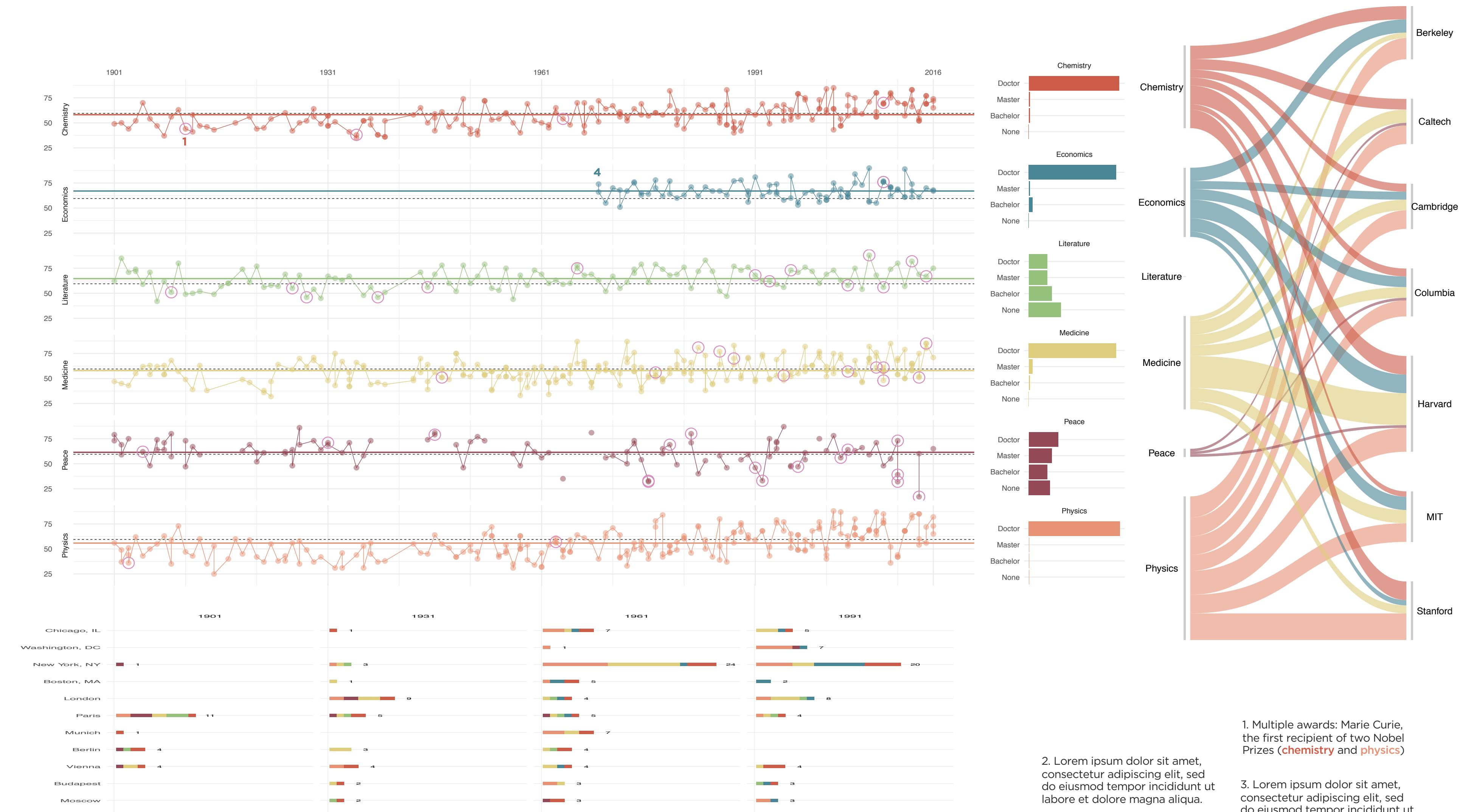


Nobels, no degrees

This visualization explores Nobel Prizes and graduate qualifications from 1901 to 2012, by analyzing the age of recipients at the time prizes were awarded, average age evolution through time and among categories, graduation grades, main university affiliations and the principal hometowns of the graduates.

How to read it?

Each dot represents a Nobel laureate, each recipient is positioned according to the year the prize was awarded (x axis) and age of the person at the time of the award (y axis).



1. Multiple awards: Marie Curie, the first recipient of two Nobel Prizes (**chemistry** and **physics**)

2. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

3. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

4. Sibling pride: Jan and Nikoalaas Tinbergen, the only brothers to win a prize each (**economics** and **medicine**)

6. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

7. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.

Nobel Laureates and Prizes

Chart type: *connected scatter plot, alluvial diagram, small multiples of stacked bar charts*

This example is based on the [award-winning design](#) by [Accurat](#).

[Download Dataset in CSV Format](#)



Follow us on Twitter 

Ask questions, report issues and request features 

© All Rights Reserved.

On the horizon

For Next Week, Module 9:

Agenda next week

The minimum

Next deliverable, *draft* information graphic

Perceptions of uncertainty, missing data

Information graphics

Andrews, R J. *Info We Trust: How to Inspire the World with Data.* Wiley, 2019. Print. Chp 17-18

While the whole book is an excellent read of plain language descriptions of what we've been covering, focus on chapters 17-18 for class discussion.

Liu, Zhicheng et al. "Data Illustrator." New York, New York, USA: ACM Press, 2018. 1-13. Web.

Read to consider how our process of designing infographics may be affected by the tools we use.

Feedback

Your turn

Ask me a question and I'll try to answer.

Your project

In creating your project, did you find data missing from the information you gathered? If so, how would you describe or visualize it? How did that affect your analysis?



See you
next week!