Introduction to Data Visualization: Questions, Encodings, and Comparisons

- This course equips you with a practical framework to evaluate and design visualizations that genuinely support understanding.
- Data visualization is a cognitive tool: its primary job is to help people perceive patterns and make accurate comparisons.
- Master three core lenses: (1) defining the analytical question, (2) selecting effective visual encodings, (3) designing meaningful comparisons.
- Workflow that structures this training: Question → Data → Encoding → Comparison → Insight.

Why Visualization is a Cognitive Tool, Not Just a Design Exercise

VS.

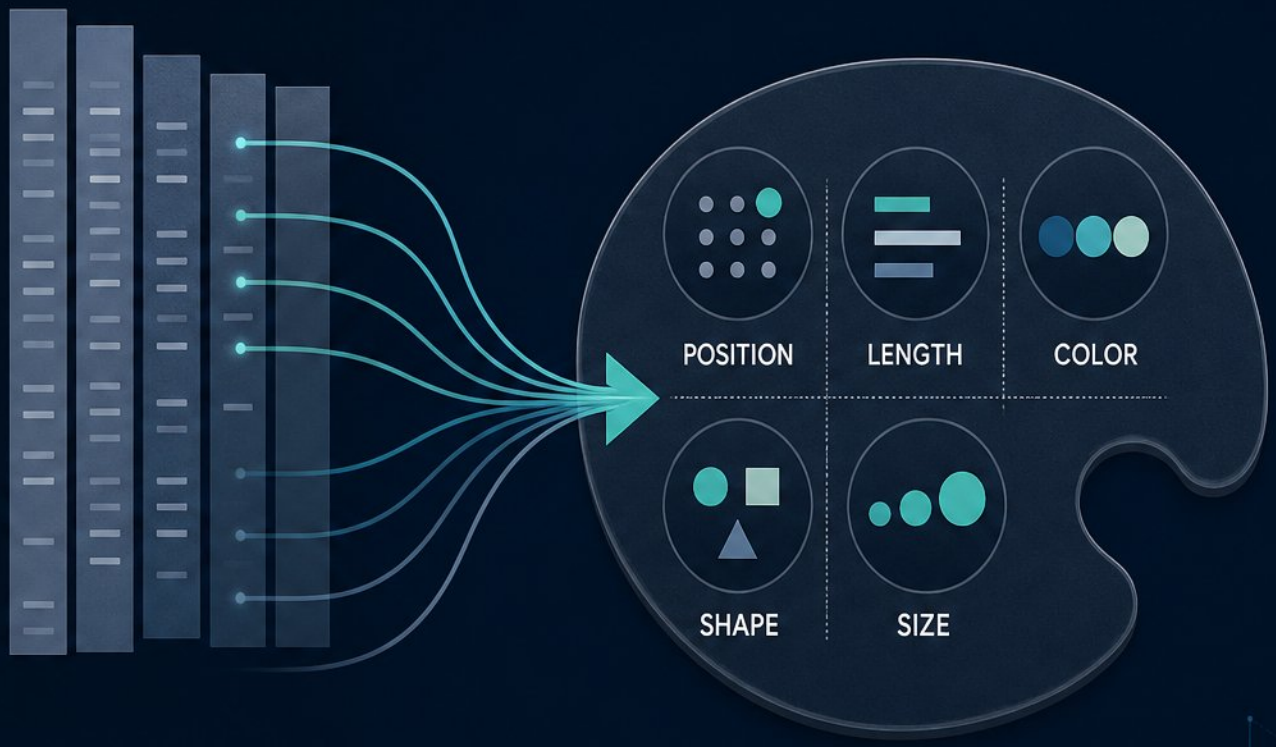
- Effective visualization aligns with human perception, not just aesthetics
- Attractive charts mislead when design conflicts with perceptual strengths
- Every design choice increases or decreases cognitive load
- A clean bar chart vs. an ambiguous but visually impressive dashboard

Lens One: Defining the Analytical Question

- A precise analytical question anchors every effective visualization
- Three families: exploratory, explanatory, and confirmatory
- Common types: composition, comparison, relationship, distribution, trend
- Shneiderman's Mantra: overview → zoom/filter → details-on-demand
- Exercise: reverse-engineer the question a sample chart answers

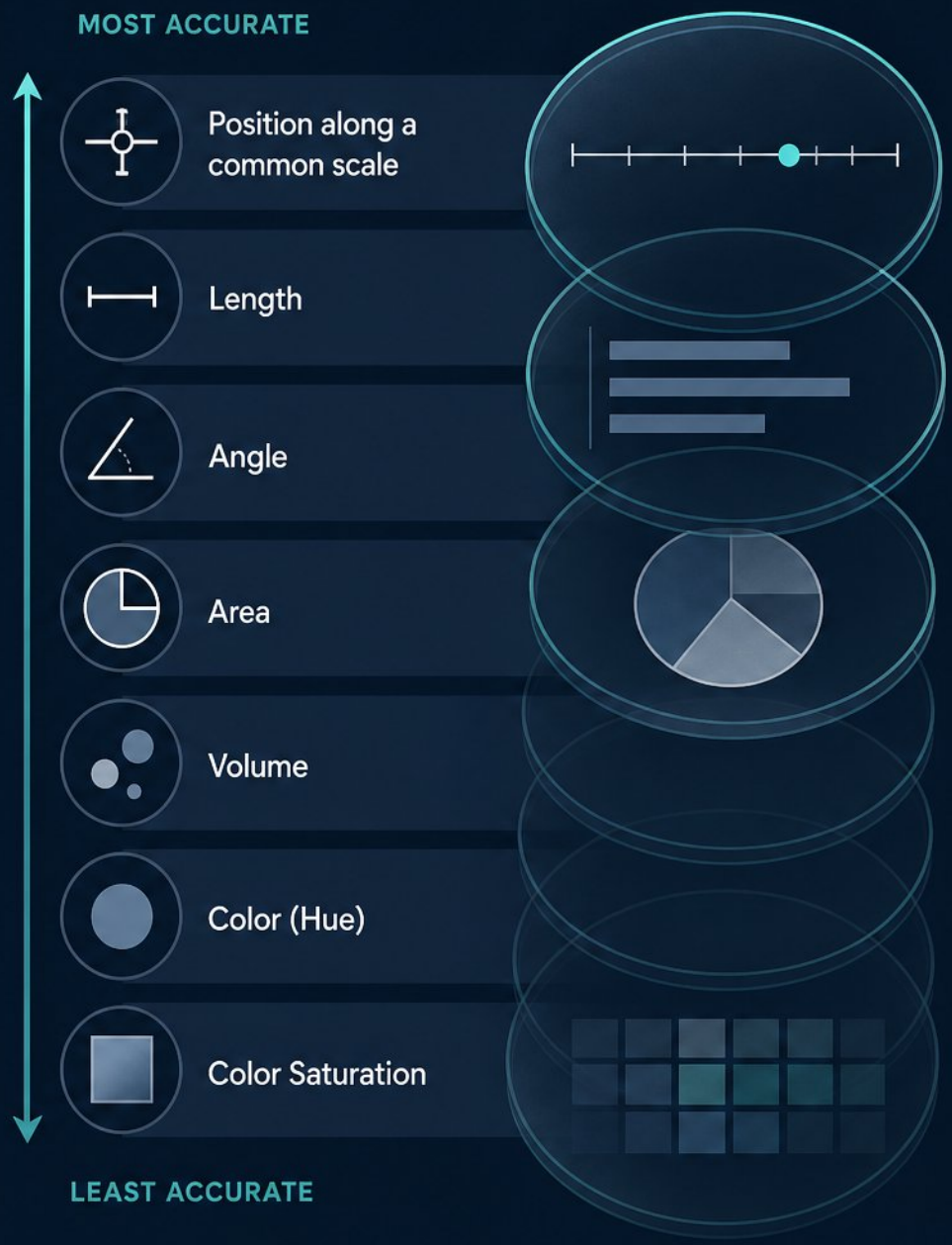
Lens Two: Visual Encodings – Mapping Data to Perception

- Visual encoding maps data fields to position, length, color, shape, or size
- Bertin identified core visual variables; modern research refines how they perform
- Expressiveness principle: encoding type must match data type (quantitative, ordinal, categorical)
- Avoid misleading encodings: area for 1D comparisons, poor color scales, 3D distortion



The Perceptual Hierarchy: Which Encodings Work Best

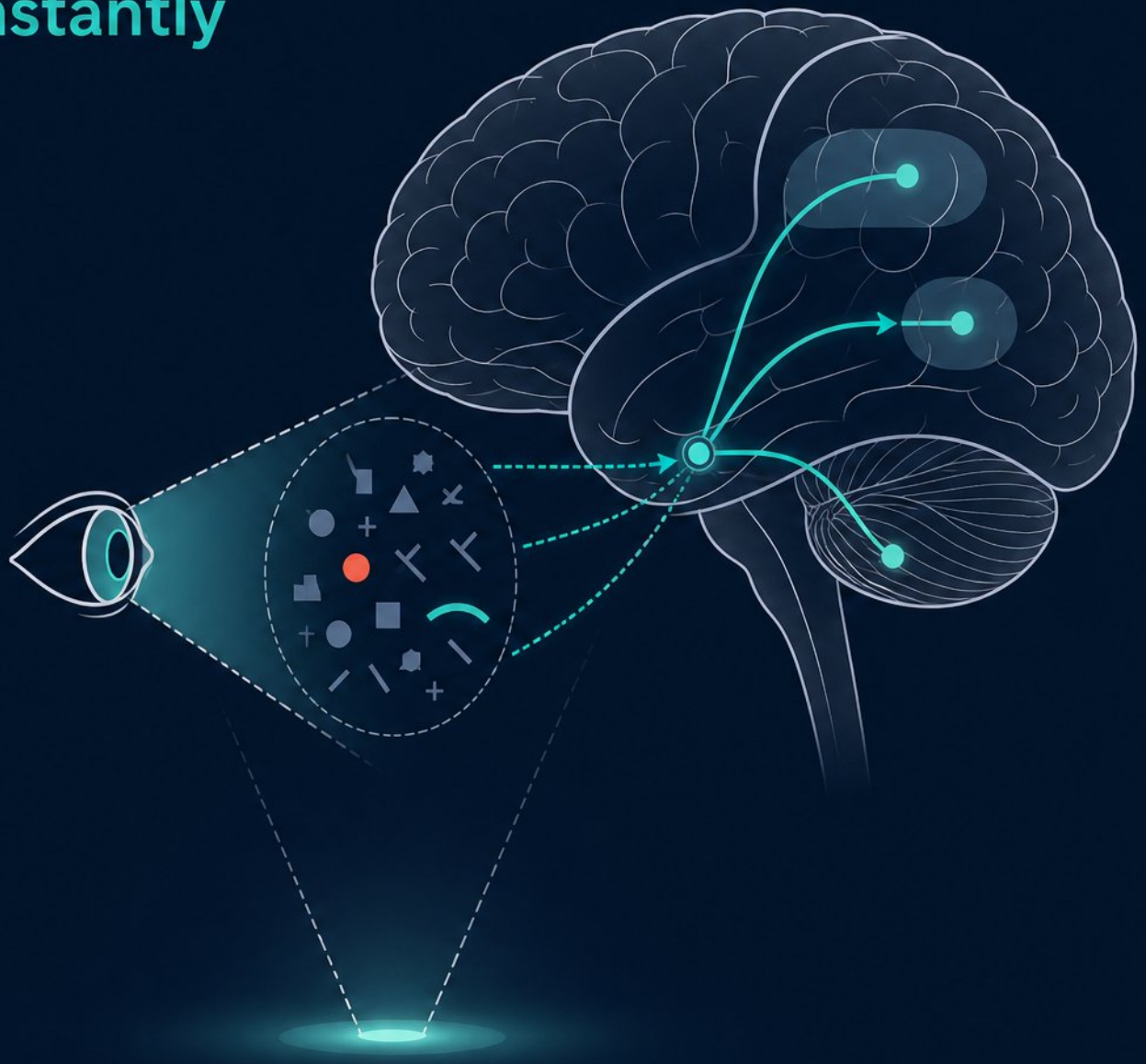
- Position along a common scale is the most accurate encoding
- Length and angle judgments are 1.4–2.5× less precise than position
- Bar charts (position/length) outperform pie charts (angle/area)
- Always choose the highest-ranked encoding your data allows



Preattentive Attributes:

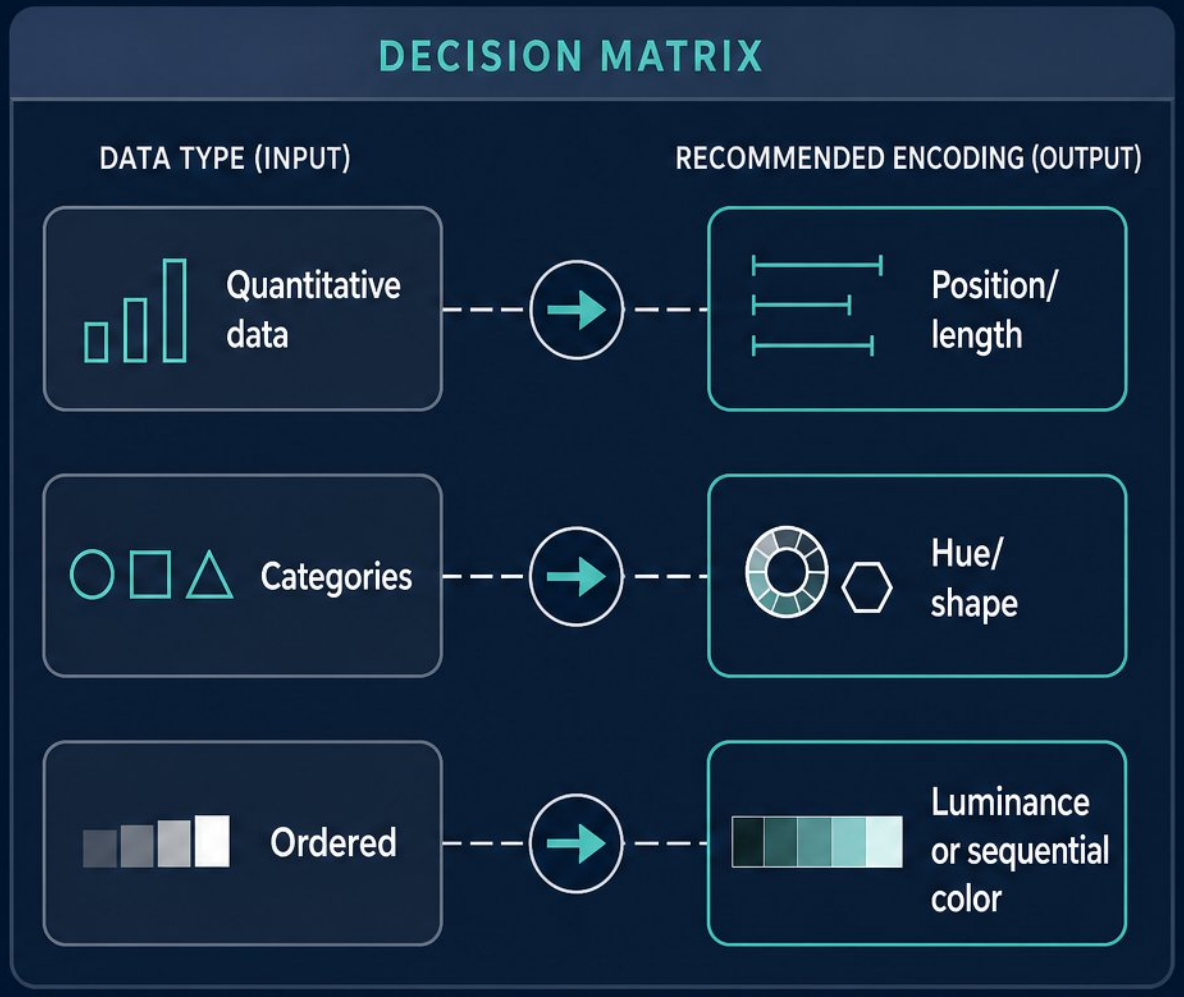
Guiding Attention Instantly

- Preattentive attributes are processed in under 250 ms, without focused attention
- Colin Ware defines four categories: color, form, movement, and spatial position
- Pop-out effects: a red object among blue, a curved line among straight lines
- Use to highlight key data; over-encoding with multiple signals creates confusion
- Strong features (luminance, hue) can mask weaker features (shape, curvature)



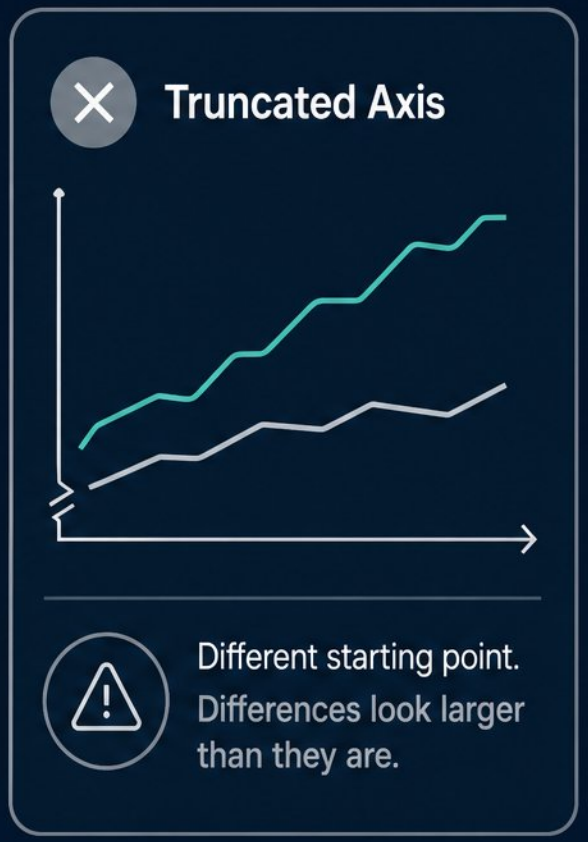
Selecting the Right Encoding: A Practical Checklist

- Start from your analytical question and data type, then apply the perceptual hierarchy.
- Deploy preattentive attributes strategically to guide the viewer's eye, not as decoration.
- Decision matrix: quantitative data → position/length; categories → hue/shape; ordered → luminance or sequential color.
- Avoid pitfalls: over-encoding, or encoding critical data in low-accuracy channels like color saturation.



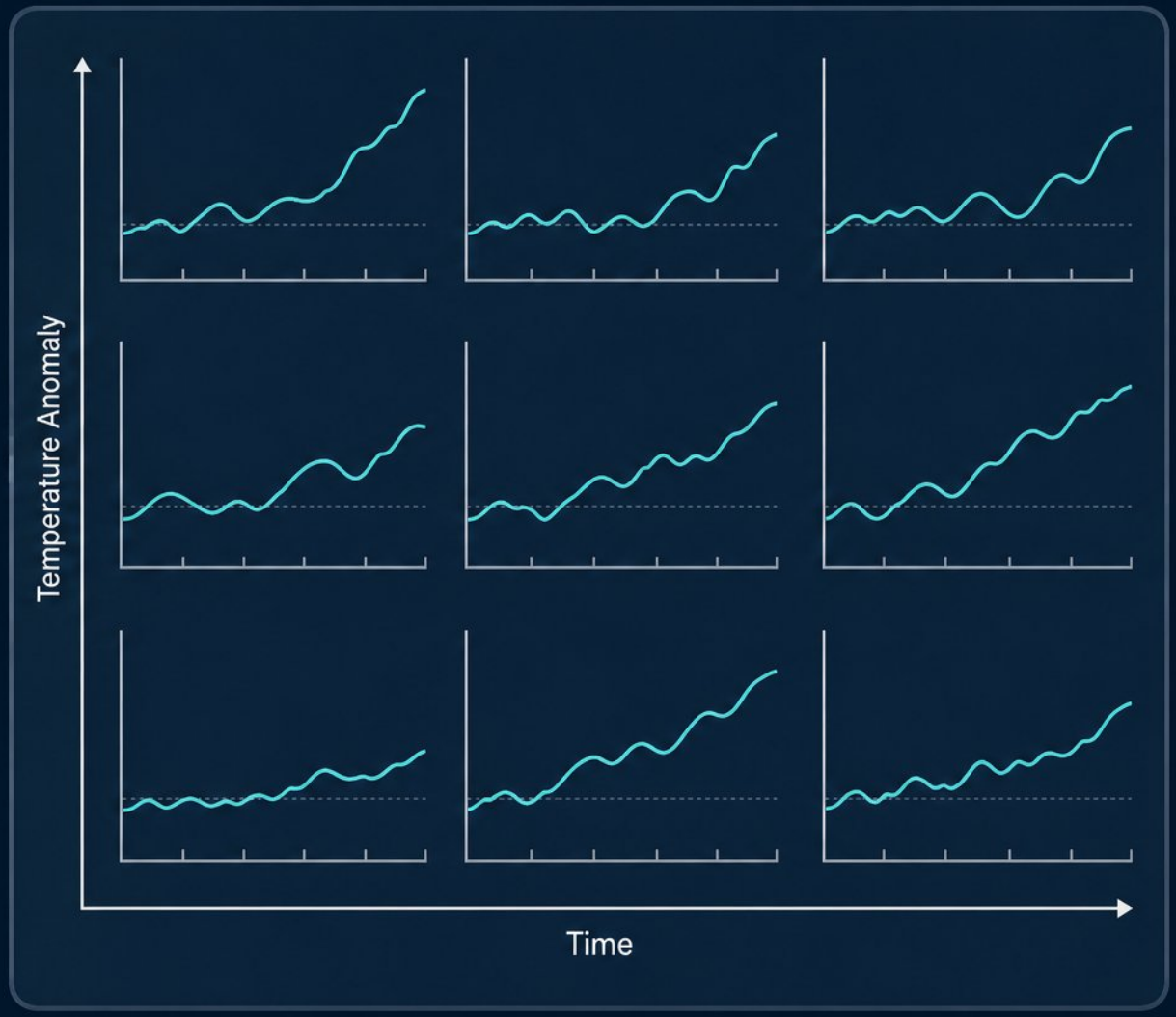
Lens Three: Designing Comparisons That Work

- "Compared to what?" is the core question; most charts fail by omitting it
- Use small multiples, superposition, and aligned baselines for fair comparison
- Chart junk and inconsistent scales distort differences and increase cognitive load
- Maximize the data-ink ratio: erase everything that does not carry information



Small Multiples:
The Power of Repeated, Comparable Views

- Use the same chart type, axes, and scales across all panels
- Let readers focus on data changes, not design variations
- Order panels logically to reveal a pattern or narrative
- Replace cluttered "spaghetti charts" with clean, readable panels
- Climate, census, and survey data are classic use cases



Common Comparison Mistakes in Practice

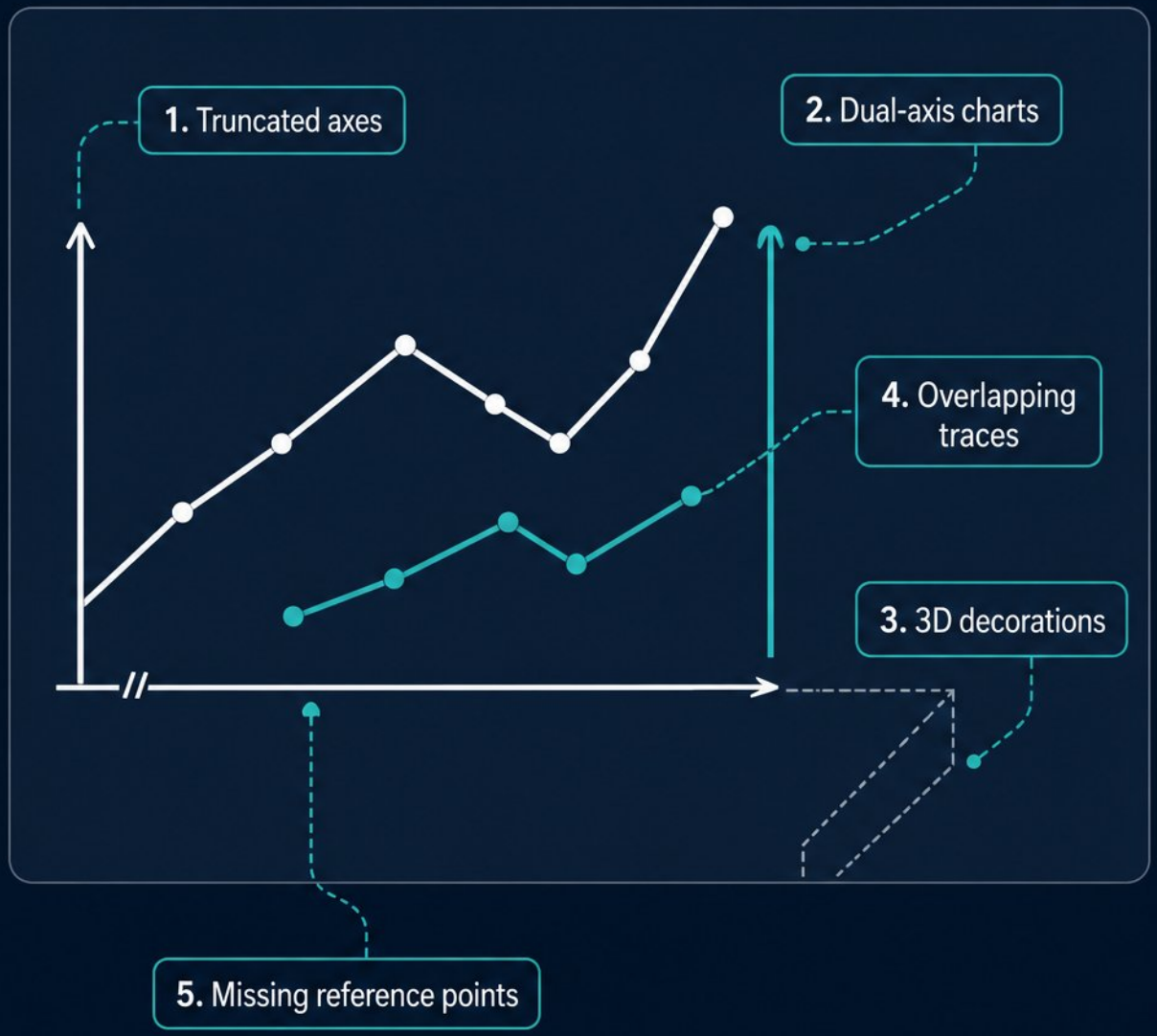
- **Truncated axes:** always check baseline alignment to avoid exaggerated differences

- **Dual-axis charts:** use only when values share a meaningful relationship

- **3D decorations:** distorts angles and areas, making accurate comparison impossible

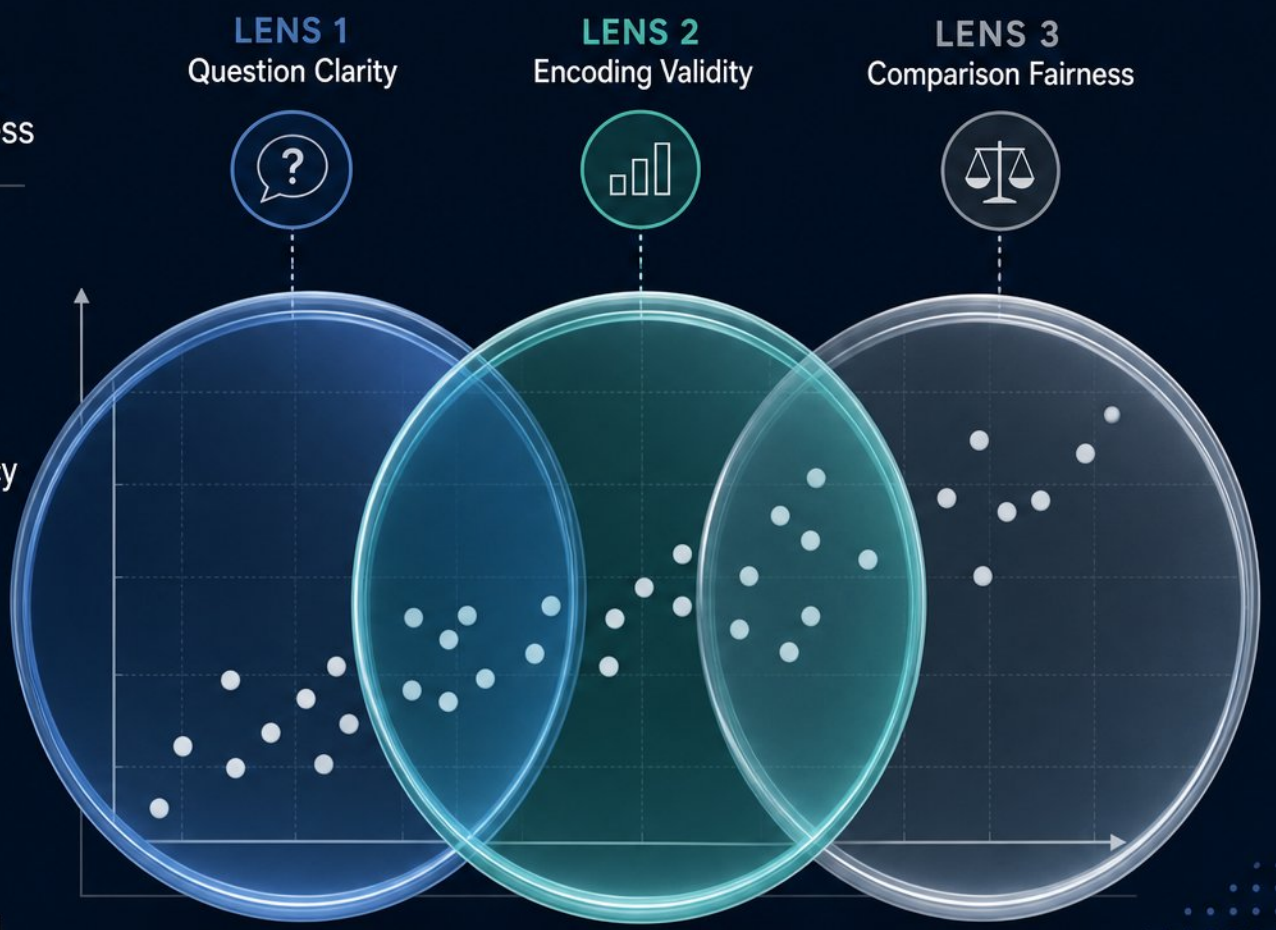
- **Overlapping traces:** consider small multiples instead of tangled lines

- **Missing reference points:** every chart needs a "compared to what?" benchmark



Putting It All Together: A Three-Lens Evaluation Rubric

- **Audit systematically:** Question clarity → Encoding validity → Comparison fairness
- Can you state the intended analytical question the chart answers?
- Does the chart use high-accuracy encodings (position, length) for key data?
- Is the comparison explicit, aligned, and fairly scaled?
- Red flags: pie charts with many slices, dual axes, missing baselines, over-encoding

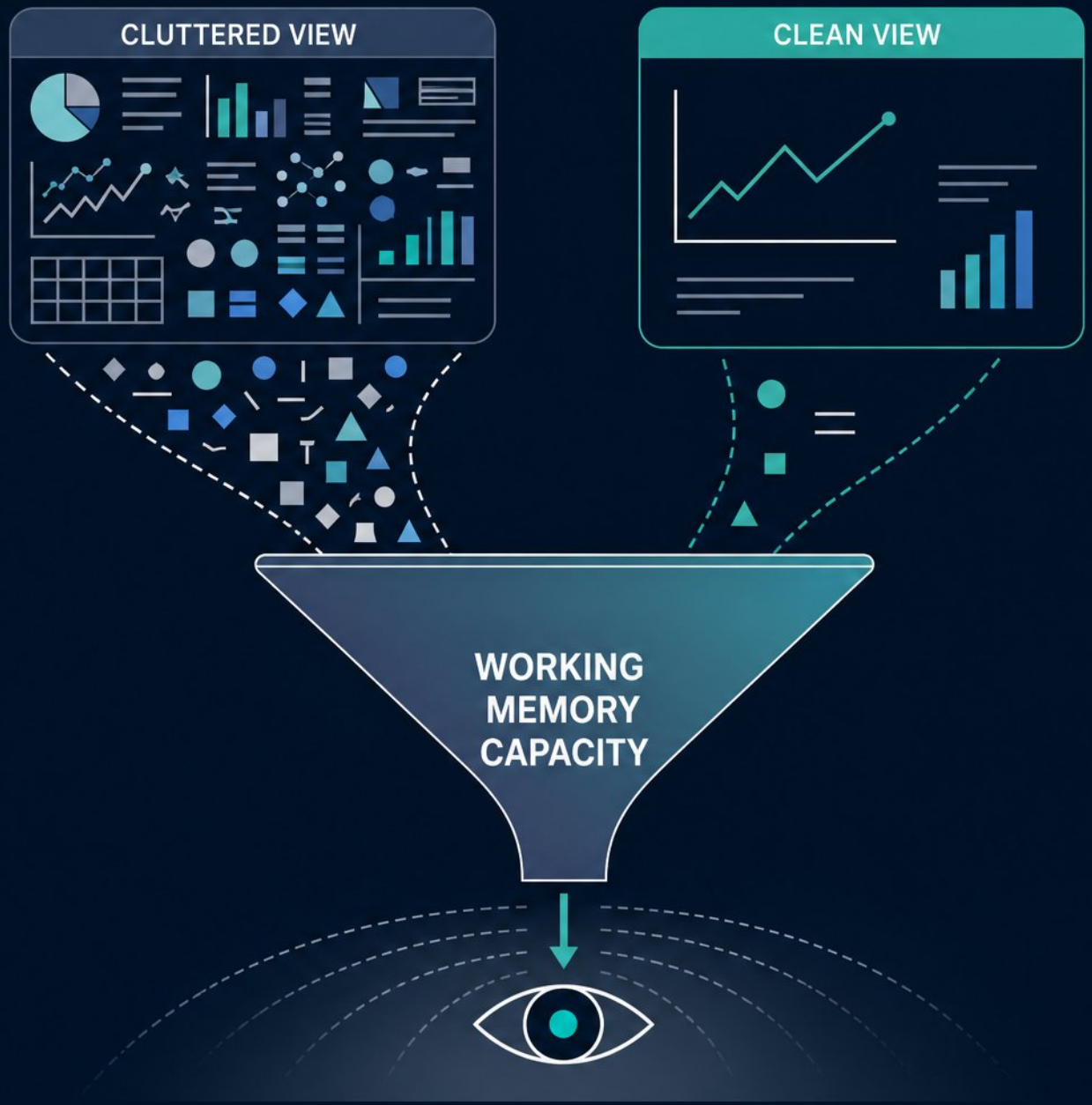


Group Diagnosis: Evaluating Real-World Charts

- Use the three lenses:
  Question → Encoding → Comparison
- For 'mystery charts,'
  reverse-engineer the intended question
- Spot subtle encoding flaws
  that distort perception
- Offer constructive critique
  grounded in perceptual evidence

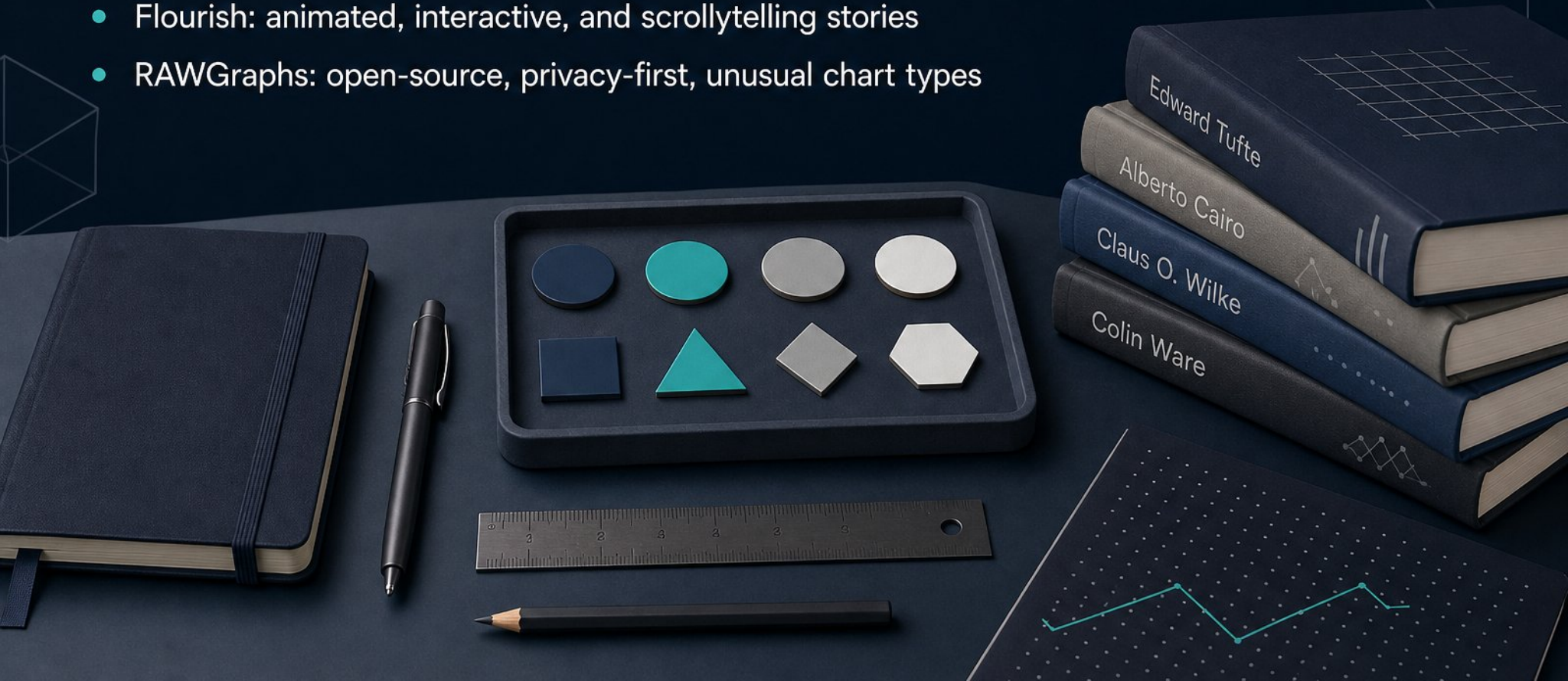
Cognitive Load and the 'Dashboard Fallacy'

- More data on screen often increases cognitive load beyond working memory capacity
- Intrinsic load (data complexity) cannot be eliminated, but extraneous load (poor design) can be minimized
- Over-encoding with color, size, and shape simultaneously splits attention and slows comprehension
- Declutter: remove redundant labels, lighten gridlines, use subtle color, focus on one comparison per view
- Match design to the viewer's chart literacy and domain familiarity



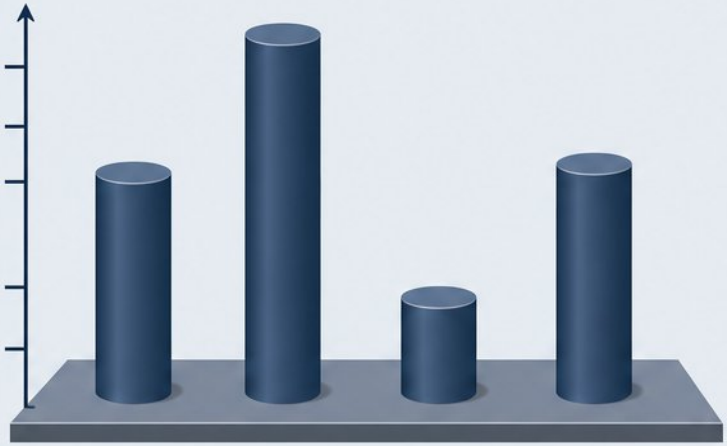
Your Toolkit:
Tools and Principles for Continued Practice

- Core workflow: Question → Encoding → Comparison → Insight
- Recommended reading: Few, Tufte, Cairo, Wilke, and Ware
- Datawrapper: fast, accessible charts and embeds for newsrooms
- Flourish: animated, interactive, and scrollytelling stories
- RAWGraphs: open-source, privacy-first, unusual chart types



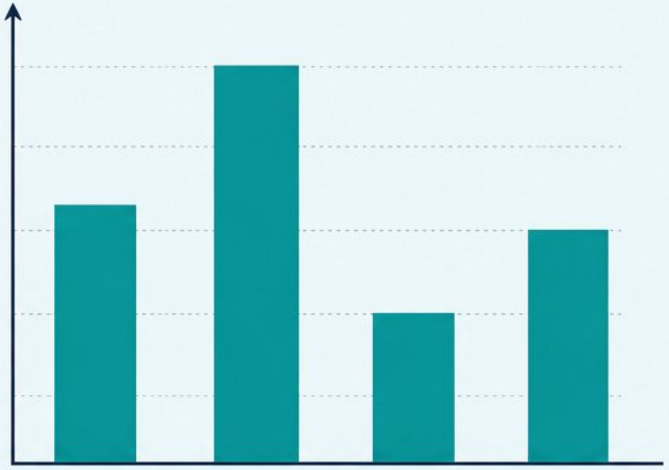
Final Exercise: Redesign a Flawed Chart

FLAWED CHART



- 3D distorts perception and exaggerates differences
- No clear question or takeaway
- Inconsistent spacing and no clear scale
- Hard to compare values accurately

REDESIGNED CHART



- 2D bars support accurate length comparisons
- Clear structure makes the takeaway obvious
- Consistent spacing with a common baseline
- Easier to compare and interpret quickly

ENCODING TOOLKIT

- POSITION (Along Scale)
- LENGTH (2D Bars)
- SHAPE (Simple)
- BASELINE (Common)
- GRIDLINES (Light)
- COLOR (Consistent)

1. Apply the Question → Encoding → Comparison lenses to a sample chart
2. Identify specific flaws: poor encoding, unclear question, or unfair scale
3. Sketch or specify a redesign that improves accuracy and clarity
4. Share and explain your changes, linking each to a perceptual principle
5. Reflect: what one technique will you immediately apply next time?

PersonWise.

PersonWise • AI digital-human course platform

Scan the code or click anywhere to watch this course live, with voice Q&A

personwise.ai/t/0ab7785b/public