Information Visualization

Uncertainty Visualization

Dr. David Koop
Uncertainty

- Uncertainty shows up in science all the time
  - Measuring
  - Modeling
  - Forecasting
- People know there is uncertainty in data analysis, but don't actually understand most ways of communicating the amount of uncertainty
How easy is it to ignore the uncertainty?

This contributes to **dichotomania**.

People Ignore Uncertainty

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient (Standard Error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>.41 (.93)</td>
</tr>
<tr>
<td>Countries</td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>1.31 (.33)**B,M</td>
</tr>
<tr>
<td>Chile</td>
<td>.93 (.32)**B,M</td>
</tr>
<tr>
<td>Colombia</td>
<td>1.46 (.32)**B,M</td>
</tr>
<tr>
<td>Mexico</td>
<td>.97 (.32)**CH,CO,V</td>
</tr>
<tr>
<td>Venezuela</td>
<td>.96 (.37)**B,M</td>
</tr>
</tbody>
</table>

[M. Kay]

D. Koop, CSCI 628, Fall 2021
People Ignore Uncertainty

<table>
<thead>
<tr>
<th></th>
<th>FiveThirtyEight: Trump's Chances</th>
<th>NYT Upshot: Trump's Chances</th>
<th>HuffPo Pollster: Trump's Chances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trump's Chances</td>
<td>28.6%</td>
<td>15.0%</td>
<td>2.0%</td>
</tr>
</tbody>
</table>

[M. Kay]
Better Ways to Present Uncertainty

FiveThirtyEight: Trump's Chances

NYT Upshot: Trump's Chances

HuffPo Pollster: Trump's Chances

286 cases in 1,000

150 cases in 1,000

20 cases in 1,000

[J. H. Gross, Washington Post, 2016]
Graphical Annotations of Distributional Properties

Intervals and Ratios

- error bars

Distributions

- violin plot
- gradient plot

Time

- hypothetical outcome plot
- quantile dot plot
- ensemble plot

[L. Padilla et al.]
## Uncertainty Visualization Theories

<table>
<thead>
<tr>
<th>Theory</th>
<th>Summary</th>
<th>Visualization Techniques</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency Framing [30] (Section 1.2)</td>
<td>Uncertainty is more intuitively understood in a frequency framing (1 out of 10) than in a probabilistic framing (10%)</td>
<td>icon array [13], quantile dotplot [11], hypothetical outcome plots [10]</td>
</tr>
<tr>
<td>Attribute Substitution [31] - Deterministic Construal Error [32] (Section 1.2)</td>
<td>If given the opportunity, viewers will mentally substitute uncertainty information for data that are easier to understand</td>
<td>hypothetical outcome plots [10]</td>
</tr>
<tr>
<td>Visual Boundaries = Cognitive Categories [21] (Section 1.2)</td>
<td>Ranges that are represented by boundaries lead people to believe that data inside and outside the boundary are categorically different</td>
<td>ensemble display [12], error bar alternatives [7, 9]</td>
</tr>
<tr>
<td>Visual Semiotics [14] (Section 1.2)</td>
<td>Some encoding techniques naturally map onto uncertainty</td>
<td>fuzziness, transparency, location, etc. [14], value-suppressing color pallet [25]</td>
</tr>
</tbody>
</table>

---

Frequency Framing hypothesis was initially proposed by Gerd Gigerenzer [30] in response to popular theories, which argued that human reasoning systematically deviates from rational choice according to mathematical rules (e.g., [33]). Gigerenzer hypothesized that our decisions seem flawed when we are provided with confusing information, such as probabilities communicated as percentiles (e.g., 10% chance). However, individuals can make rational choices if provided with information in a format they can understand easily, such as in frequencies or ratios (e.g., 1 out of 10). Gigerenzer argued that percentiles do not match the way people encounter probability in the world, and therefore lead to errors. Instead, it is more intuitive to depict probability as a frequency, as we have more exposure to these types of ratios (e.g., I hit traffic on this road 7 out of 10 times. I will take a different route tomorrow.) The frequentist framing hypothesis has substantial support from studies that find we can relatively automatically and accurately understand frequency formats, whereas probabilities are time consuming and highly error prone (for review and caveats, see [34]).

One of the most effective ways to implement frequency framing of uncertainty information is with visualizations, and in this section we will detail two promising frequency-framing techniques. Researchers, predominantly in healthcare communication, have extensively studied the use of icon arrays (see Figure 1.1) to display ratios and have found strong evidence that they are useful for communicating forecasted probabilities of event outcomes. The second notable use of frequency formats in visualization is within the emerging study of quantile dotplots (see Figure 1.1). Whereas quantile dotplots are relatively new and have not received as much examination as icon arrays, they capitalize on the theoretical benefits of frequency framing and have demonstrated positive results in laboratory studies.

Icon arrays

A substantial body of research demonstrates that icon arrays are one of the most effective ways to communicate a single probabilistic value and can outperform textual descriptions of probabilities and frequencies [27, 35–42]. One of the key benefits of icon arrays is that they offload cognition by allowing a viewers visual system to compare the denominator and the numerator in a frequency probability format. Visual comparisons of this nature are easier and faster than numerical calculations.

The difficulty in comparing ratios can produce common errors, such as individuals focusing on the numerator of each ratio and neglecting the denominator, called denominator neglect (for review see [43]). For example, when comparing a cancer with a mortality rate of 1,286 of 10,000 people to a cancer with a mortality rate of 24 of 100 people, participants in a laboratory study incorrectly reported that the former cancer was riskier [44]. Researchers propose that individuals pay more attention to the relative differences in numerators (in this case, 1,286 vs. 24 deaths), even though they should consider the relative ratios (12.86% vs. 24% mortality) (e.g., [43, 44]). Several studies...
5. RESULTS

Figure 10 shows each of the six cases presented to the experiment participants, in their order of presentation, with the top row of each case showing the error cone view, and the bottom row showing our method. These examples were:

- Case 1
- Case 2
- Case 3
- Case 4
- Case 5
- Case 6

FIG. 10: The six cases as shown to experiment participants.

Hurricane Error Cones vs. Ensembles

[Cox et al.]
Ensembles not perfect either

[Padilla et al.]
Spaghetti Plot vs. HOP

[M. Kay]
Spaghetti Plot vs. HOP

[M. Kay]
Dithering to show uncertainty

Kansas City
Columbia
St. Louis
Springfield

Poverty Rate (%)
25
20
15
10
5

[D. Koop, CSCI 628, Fall 2021]

[Lucchesi and Wikle, 2017]
Bivariate Colormap (Uncertainty → Saturation)

[Correll et al., 2018]
Value-Suppressing Uncertainty Palette

[Correll et al., 2018]
Schedule

• Thursday: Progress Reports & Uncertainty
• Next Tuesday: Surveys Due & Presentations
• Tuesday, Oct. 26: No Class
• Thursday, Oct. 28: High-Dimensional Data Critique Due
Today's Paper: Critique Due

When (ish) is My Bus? User-centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems

Matthew Kay  
CSE | dub  
University of Washington  
mjskay@uw.edu

Tara Kola  
Computer Science  
Tufts University  
tara.kola@tufts.edu

Jessica R. Hullman  
iSchool | dub  
University of Washington  
jhullman@uw.edu

Sean A. Munson  
HCDE | dub  
University of Washington  
smunson@uw.edu

Bus Timeline

Route Timeline

<table>
<thead>
<tr>
<th>Density</th>
<th>Bus Timeline</th>
<th>Route Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dotplot (20)</th>
<th>Bus Timeline</th>
<th>Route Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dotplot (100)</th>
<th>Bus Timeline</th>
<th>Route Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Stripeplot (50)</th>
<th>Bus Timeline</th>
<th>Route Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td></td>
<td>10</td>
</tr>
</tbody>
</table>

Copyright is held by the owner/author(s). Publication rights limited to reproduction by the author(s) or the author's assignee for the personal use of the author(s) or the personal use of their assignee. Further reproduction, or any commercial act, without permission may be unlawful. The appearance of the US Department of Defense seal indicates that this publication is being made available in fulfillment of a contractual obligation. Authorization to photocopy portions of this material for internal or personal use is granted to academic libraries and other users registered with the Copyright Clearance Center. Copyright clearance code: 0001-0782/00/

D. Koop, CSCI 628, Fall 2021
Paper Presentation
Critique
Questions & Discussion