CMPSC382
Lecture 32: Visualization Design Principles

Prof. John Wenskovitch

11/13/2015

(adapted from Harvard CS171, Hanspeter Pfister)
Last Time

• What is Visualization?
  – A technique for turning data into information.
  – Enhancing the visual system to solve problems that algorithms can’t solve efficiently.
  – An iterative visual process used to make hypotheses about data, and then to gain insight/knowledge about that data by exploring what the data is telling us.
Visualization Design

• “Design must be functional, and functionality must be translated into visual aesthetics, without any reliance on gimmicks that have to be explained.” –F.A. Porsche

• “Clutter and confusion are not attributes of information; they are failures of design.” –E. Tufte
Graphical Integrity

MOST DOWNLOADED BY CATEGORY

Entertainment

Books 11%

Games 12%

Productivity Tools 13%

17%
Graphical Integrity

When do people do their Holiday travel

91.8 Million Travelers

43.3 Million Travelers

End Of The Year (Christmas/New Years) (45%)
Travel for Thanksgiving (55%)

Top 5 Forgotten Items

22% Sun Cream
11% Hair Dryer
12% Bikes/razors/magnets
21% Sunglasses
21% Adaptors

SHOPPING WITH SOCIAL MEDIA

HAVE PURCHASED AN ITEM AFTER POSTING ABOUT IT
40%

FEEL CONNECTED TO BRANDS THEY SEE ON SOCIAL SITES
17%

OBTAIN INFO ABOUT OTHER CONSUMERS' EXPERIENCES
70%

POSITIVELY REVIEW BRANDS THEY LIKE
53%
Graphical Integrity
ALWAYS Start Your Bars at Zero
ALWAYS Start Your Bars at Zero
ALWAYS Start Your Bars at Zero

How 2012 Stacks Up
The Warmest Years on Record
Contiguous U.S.

- 1921: 53.81°
- 1999: 53.93°
- 1934: 54.14°
- 2006: 54.31°
- 1998: 54.32°
- 2012: 55.3°

Source: NOAA’s National Climatic Data Center - State of the Climate National Overview
Scale Distortions

Median household income in 2010 inflation adjusted dollars

Zero-Based
Non-Zero Based
Cherry-Picked/Incomplete Data

Temperature Anomaly -- Annual Mean (°C)
Cherry-Picked/Incomplete Data
Using Two Dimensions to Show One

THE SHRINKING FAMILY DOCTOR
In California
Percentage of Doctors Devoted Solely to Family Practice
1964 1975 1990
27% 16.0% 12.0%
1: 4,232 6,212
1: 3,167 8,594
1: 2,247 RATIO TO POPULATION
8,023 Doctors

IN THE BARREL...
Price per bbl. of light crude, leaving Saudi Arabia on Jan. 1
April 1 $14.55
$12.34
$11.51
$10.96
$10.55
$2.41
Pie Charts are Inherently Misleading

U.S. Smartphone Marketshare

- RIM: 21.2%
- Apple: 39.0%
- Palm: 3.1%
- Motorola: 9.8%
- Nokia: 7.4%
- Other: 19.5%
Pie Charts are Inherently Misleading
3D Pie Charts are Worse
Just No
Design Principle #1: Maximize Data-Ink Ratio

Data-Ink Ratio = \frac{\text{Data ink}}{\text{Total ink used in graphic}}
Design Principle #2: Avoid Chartjunk

Extraneous visual elements that distract from the message
Design Principle #3: Increase Data Density

Data density = \frac{\text{Number of data items}}{\text{Area of data in graphic}}
Design Principle #4: Layering and Separation

<table>
<thead>
<tr>
<th>Train No.</th>
<th>3701</th>
<th>3301</th>
<th>3801</th>
<th>3542</th>
<th>3765</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York</td>
<td>12:10</td>
<td>1:30</td>
<td>3:45</td>
<td>7:30</td>
<td>4:33</td>
</tr>
<tr>
<td>Newark, N. J.</td>
<td>1:43</td>
<td>10:30</td>
<td>5:21</td>
<td>8:50</td>
<td>11:45</td>
</tr>
<tr>
<td>North Elizabeth</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>6:45</td>
</tr>
<tr>
<td>Elizabeth</td>
<td>3:33</td>
<td>2:05</td>
<td>...</td>
<td>...</td>
<td>7:05</td>
</tr>
<tr>
<td>Peekskill</td>
<td>5:34</td>
<td>6:40</td>
<td>...</td>
<td>7:20</td>
<td>8:50</td>
</tr>
<tr>
<td>Ediison, N. J.</td>
<td>4:45</td>
<td>5:20</td>
<td>4:40</td>
<td>2:10</td>
<td>11:05</td>
</tr>
<tr>
<td>Princeton, N. J.</td>
<td>1:30</td>
<td>...</td>
<td>...</td>
<td>3:30</td>
<td>7:30</td>
</tr>
</tbody>
</table>

| New York  | 12:10 | 1:30 | 3:45 | 7:30 | 4:33 |
| Newark, N. J. | 1:43 | 10:30| 5:21 | 8:50 | 11:45|
| North Elizabeth | ... | ... | ... | ... | 6:45 |
| Elizabeth | 3:33 | 2:05 | ...  | ...  | 7:05 |
| Peekskill | 5:34 | 6:40 | ...  | 7:20 | 8:50 |
| Ediison, N. J. | 4:45 | 5:20 | 4:40 | 2:10 | 11:05|
| Princeton, N. J. | 1:30 | ...  | ...  | 3:30 | 7:30 |

Train No. | 3701 | 3301 | 3801 | 3542 | 3765
Questions to Ask

• Who is the audience? (Expert? Non-expert?)
• What questions does this visualization answer?
• What design principles best describe why it is good/bad?
• Why do you like/dislike this visualization?
• Can you suggest any improvements?
Evaluate This

Any Questions?