The Vision of Visual Analytics

Zachary Skrivanek / Eli Lilly & Company

Introduction
Role of the Statistician
Technological Change
Drug Development in the 21st Century
Visual Analytics
Bivariate Example #1

Anscombe's Quartet

1. $r = 0.82$
   
   $y = 0.50x + 3.00$

2. $r = 0.82$
   
   $y = 0.50x + 3.00$

3. $r = 0.82$
   
   $y = 0.50x + 3.00$

4. $r = 0.82$
   
   $y = 0.50x + 3.00$
Bivariate Example #2

X Mean: 54.2635812
Y Mean: 47.8330711
X SD : 16.7676255
Y SD : 26.9309892
Corr. : -0.0600966
Bivariate Example #2

X Mean: 54.2635812
Y Mean: 47.8330711
X SD: 16.7676255
Y SD: 26.9309892
Corr.: -0.0600966
Univariate Example #1
Univariate Example #2
Table vs Figure

<table>
<thead>
<tr>
<th>Type</th>
<th>Incidence</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate</td>
<td>238,590</td>
<td>18.7%</td>
</tr>
<tr>
<td>Breast</td>
<td>232,340</td>
<td>18.2%</td>
</tr>
<tr>
<td>Lung</td>
<td>228,190</td>
<td>17.9%</td>
</tr>
<tr>
<td>Colon</td>
<td>142,820</td>
<td>11.2%</td>
</tr>
<tr>
<td>Melanoma</td>
<td>76,690</td>
<td>6.0%</td>
</tr>
<tr>
<td>Bladder</td>
<td>72,570</td>
<td>5.7%</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>69,740</td>
<td>5.5%</td>
</tr>
<tr>
<td>Thyroid</td>
<td>60,220</td>
<td>4.7%</td>
</tr>
<tr>
<td>Kidney</td>
<td>59,938</td>
<td>4.7%</td>
</tr>
<tr>
<td>Leukemia</td>
<td>48,610</td>
<td>3.8%</td>
</tr>
<tr>
<td>Pancreas</td>
<td>45,220</td>
<td>3.5%</td>
</tr>
</tbody>
</table>
N = 1,274,928 cancer cases
Insights from many disciplines

- Alberto Cairo

more complex and deeper

more intelligible and shallower
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

more complex and deeper

more intelligible and shallower
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

more complex and deeper

more intelligible and shallower

redundancy
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

more complex and deeper

novelty

more intelligible and shallower

redundancy
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

more complex and deeper

more intelligible and shallower

novelty

redundancy

familiarity
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

more complex and deeper

more intelligible and shallower

originality

novelty

redundancy

familiarity
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

more complex and deeper

more intelligible and shallower
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

multidimensionality
originality
novelty

more complex
and deeper

more intelligible
and shallower

unidimensionality
familiarity
redundancy
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

more complex and deeper

more intelligible and shallower

density  multidimensionality  originality

novelty

redundancy

familiarity

unidimensionality  lightness
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

density  multidimensionality
originality
novelty

more complex and deeper

redundancy
familiarity
unidimensionality
lightness
decoration

more intelligible and shallower
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

more complex and deeper

more intelligible and shallower
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo
Insights from many disciplines

- graphic design
- computer science
- cognitive psychology
- journalism
- statistics

- Alberto Cairo

density multidimensionality
functionality originality
abstraction novelty
redundancy figuration
familiarity decoration
unidimensionality lightness

more complex and deeper
more intelligible and shallower
History of Visualization

• 1637 - Descartes first uses 2d grids to visually encode numbers
• 1786 - William Playfair's "The Commercial and Political Atlas"
• 1855 - John Snow uses maps to link the 1854 London cholera epidemic to contaminated drinking water.
• 1857 - Florence Nightingale uses stacked bar and pie charts to persuade Queen Victoria to improve conditions on British military hospitals.
• 1869 - Charles Minard's "Napolean's Last March"
• 1954 - Darrel Huff's "How to Lie with Statistics"
• 1977 - John Tukey introduces boxplots.
• 1983 - Edward Tufte's "Visual Displays of Quantitative Information"
• 1994 - William Cleveland's "The Elements of Graphing Data".
• 2004 - Stephen Few "Show me the Numbers".
• 2014 - Tamara Munzner's "Visualization Analysis & Design"
• Nowadays dominated by computer scientists (on the technical side) and business analytics (on the more applied side)
Hyperbilirubinemia Secondary Analysis workflow

• Task 1: Evaluate the proportion of cases due to conjugated bilirubin elevation from cases due to unconjugated bilirubin elevation
• Task 2: Evaluate concurrent changes in hemoglobin
• Task 3: Evaluate time to onset of bilirubin elevations
• Task 4: Evaluate time to resolution of bilirubin elevations (on drug and after drug)
• Task 5: Evaluate dose-response relationship
• Task 6: Evaluate the presence of concurrent adverse events
• Task 7: Evaluate risk on the basis of subgroups; i.e., do the cases of interest differ from the “non-cases” on the basis of age, gender, ethnicity, body mass index, country of origin
Hepatotoxicity Graphics Workflow

Bilirubin vs ALT
Change from Baseline

Hy’s Law/ eDish/mDish

Upper right quadrant

By-patient Line Plots

Patient Profiles

Boxplots

Kaplan-Meier
The Vision of Visual Analytics
Zachary Skrivanek / Eli Lilly & Company

Introduction
Role of the Statistician
Technological Change
Drug Development in the 21st Century
Descriptive
Diagnostic
Predictive
Prescriptive
Highly Interactive

Automated Analyses
Highly Interactive

Automated Analyses
Communication

data visualization
Hans Rosling
Presentation Zen
The Vision of Visual Analytics

Zachary Skrivanek / Eli Lilly & Company

Drug Development in the 21st Century
Skeuomorphism

representing real world objects
data temp;
infile DATALINES dsd missover;
input a b c d;
CARDS;
1, 2, 3, 4
, 3, , 5
, 3
;
run;
QWERTY
The Vision of Visual Analytics
Zachary Skrivanek / Eli Lilly & Company

- Introduction
- Role of the Statistician
- Technological Change
- Drug Development in the 21st Century
electronic submission
hyperlinks?
Web based solutions
s://github.com/d3/d3/wiki/Gallery
The Vision of Visual Analytics

Zachary Skrivanek / Eli Lilly & Company

- Introduction
- Role of the Statistician
- Technological Change
- Drug Development in the 21st Century