Designing Data Visualizations
Part 2: Lab

Noah Iliinsky • @noahi

European Bioinformatics Institute • Cambridge • July 16, 2012
Two Phases

1. What to visualize

* Consider your inputs
  * your goals
  * their needs
  * shape of the data

* Write a spec for your visualization
  * select data to include

2. How to visualize

* Select axes for your most important data & relationship

* Consider & apply encodings for all other data dimensions

* Experiment, iterate, etc.
Three inputs.

- Reader
- Your visualization
- Data
- Designer
Define Desired Knowledge Before Structure
What do you want to show?
What questions are you trying to answer?

What do you want to show?
What actions/decisions are you trying to enable?

What questions are you trying to answer?

What do you want to show?
Who is consuming this data?
What are their needs?
Who is consuming this data?

If not you:

★ What are their priorities?
★ What are their biases?
★ What are their limitations?
★ What don’t you know about them?
What data dimensions do you have to play with?
Data has properties.

- Wheel size: numeric (actually categorical)
- Tire width: continuous
- Price: continuous
- Anti-puncture: binary
- Foldable: binary
What types of data do you have?

- categorical (grouped)
- ordinal (ranked, time, sequential)
- quantitative (numeric)
- relational (hierarchy, influence, etc.)
- location (... it’s complicated...)
What are the key relationships? (probably)
What are the key relationships?

What data is required to show them?
Define Desired Knowledge *Before* Structure
Statement of Goals
Show the relationship between A and B (and C...) across X (and Y) from m to n.
Show the relationship between A and B [and C...] across X [and Y] from m to n.

(“in order to determine our best and worst performing widgets.”)
What data are you actually going to use, based on that goal?

[A, B, X (from m to n)]
Now we start drawing.
Design strategies

- Limit the data & detail you include
- Use position for your most important relationship(s)
- Try different axes
- Consider default formats
- Use color for categories, not rank
- Encode other data and relationships with appropriate properties
Appropriate Encodings
### Properties and Best Uses of Visual Encodings

<table>
<thead>
<tr>
<th>Example</th>
<th>Encoding</th>
<th>Ordered</th>
<th>Useful values</th>
<th>Quantitative</th>
<th>Ordinal</th>
<th>Categorical</th>
<th>Relational</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>position, placement</td>
<td>yes</td>
<td>infinite</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>1, 2, 3; A, B, C</td>
<td>text labels (alphabetical or numbered)</td>
<td>optional</td>
<td>infinite</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td></td>
<td>length</td>
<td>yes</td>
<td>many</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>size, area</td>
<td>yes</td>
<td>many</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>angle</td>
<td>yes</td>
<td>medium/few</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pattern density</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>weight, boldness</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>saturation, brightness</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>color</td>
<td>no</td>
<td>few (&lt;20)</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>shape, icon</td>
<td>no</td>
<td>medium</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>pattern texture</td>
<td>no</td>
<td>medium</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>enclosure, connection</td>
<td>no</td>
<td>infinite</td>
<td>Good</td>
<td></td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td></td>
<td>line pattern</td>
<td>no</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>line endings</td>
<td>no</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>line weight</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Noah Iliinsky • ComplexDiagrams.com/properties • 2012-06

Position is Everything.
How can you use position to reveal your key relationship(s)?
Is there a good default format for this kind of relationship?
List (at least) three possible combinations of axes.
List (at least) three possible combinations of axes.

If you have a tool, try graphing each of these permutations.
Appropriate Encodings
What are good options for encoding your other data dimensions?
### Appropriate encodings

#### Properties and Best Uses of Visual Encodings

<table>
<thead>
<tr>
<th>Example</th>
<th>Encoding</th>
<th>Ordered</th>
<th>Useful values</th>
<th>Quantitative</th>
<th>Ordinal</th>
<th>Categorical</th>
<th>Relational</th>
</tr>
</thead>
<tbody>
<tr>
<td>position, placement</td>
<td>yes</td>
<td>infinite</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>1, 2, 3; A, B, C</td>
<td>text labels (alphabetical or numbered)</td>
<td>yes</td>
<td>infinite</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
</tr>
<tr>
<td>length</td>
<td>yes</td>
<td>many</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>size, area</td>
<td>yes</td>
<td>many</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>angle</td>
<td>yes</td>
<td>medium/few</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pattern density</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weight, boldness</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>saturation, brightness</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>color</td>
<td>no</td>
<td>few (&lt;20)</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shape, icon</td>
<td>no</td>
<td>medium</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pattern texture</td>
<td>no</td>
<td>medium</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>enclosure, connection</td>
<td>no</td>
<td>infinite</td>
<td>Good</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>line pattern</td>
<td>no</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>line endings</td>
<td>no</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>line weight</td>
<td>yes</td>
<td>few</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Noah Iliinsky • ComplexDiagrams.com/properties • 2012-06

http://ComplexDiagrams.com/properties
Pick encodings for your remaining data dimensions.

See how that looks.
Design strategies

- Limit the data & detail you include
- Use position for your most important relationship(s)
- Try different axes
- Consider default formats
- Use color for categories, not rank
- Encode other data and relationships with appropriate properties
Iterate, iterate, iterate.
Tools

* d3.js: structured frameworks
* processing: flexible, freeform
* R + ggplot2: stats & analytics
* Tableau: visual exploration & analytics
* colorbrewer2.org: palette generator
* colorlab.wickline.org: colorblind compatibility test
Thank you!

@noahi
gmail: iliinsky
ComplexDiagrams.com

How to do it
(this talk)

How they did it
(20 case studies)