DSC 201: Data Analysis & Visualization

Interaction, Multiple Views, and Maps

Dr. David Koop
Quiz: What are the marks and channels?

Ice Cream
Pie
Cake
Cookies

Winter
Spring
Summer
Fall
Visual Encoding

• How do we encode data visually?
  - **Marks** are the basic graphical elements in a visualization
  - **Channels** are ways to control the appearance of the marks

• Marks classified by dimensionality:
  - Points
  - Lines
  - Areas
  - Also can have surfaces, volumes

• Think of marks as a mathematical definition, or if familiar with tools like Adobe Illustrator or Inkscape, the path & point definitions
Proper Use of Line and Bar Charts

[Zacks and Tversky, 1999, Munzner (ill. Maguire), 2014]
Scatterplot Matrices and Parallel Coordinates

Scatterplot Matrix

Math

Physics

Dance

Drama

Parallel Coordinates

Math

Physics

Dance

Drama

[Munzner (ill. Maguire), 2014]
Light Reflection & Absorption

[Image of fruits with arrows indicating light wavelengths]

Wavelength (nm)

400 500 600 700

[via M. Meyer]
Color != Wavelength

[Graph showing relative energy density vs. wavelength (nm) with peaks at 500 nm for yellow and 600 nm for brown.]

[via M. Meyer]
Color Blindness

[Diagram showing color盲ness for Normal, Protanopia, Deuteranopia, and Tritanopia]

[Image of colorful powders via M. Meyer]
Luminance

- HSL does not truly reflect the way we perceive color
- Even though colors have the same lightness, we perceive their luminance differently
- Our perception ($L^*$) is **nonlinear**

Corners of the RGB color cube

<table>
<thead>
<tr>
<th>Corner</th>
<th>Luminance</th>
<th>$L^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magenta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Green</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yellow</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$L$ from HSL

All the same

$L^*$

[Muizner (ill. Maguire), 2014 (based on Stone, 2006)]
Colormap

• A colormap specifies a mapping between colors and data values
• Colormap should follow the expressiveness principle
• Types of colormaps:

Binary

\[
\begin{align*}
\text{y} & : & 0 \\
\text{n} & : & 0 \\
\end{align*}
\]

Categorical

\[
\begin{align*}
\text{T} & : & \text{Green} \\
\text{F} & : & \text{Blue} \\
\text{A} & : & \text{Pink} \\
\end{align*}
\]

Diverging

\[
\begin{align*}
-1 & : & \text{Purple} \\
0 & : & \text{Gray} \\
+1 & : & \text{Red} \\
\end{align*}
\]

Sequential

\[
\begin{align*}
3 & : & \text{Gray} \\
2 & : & \text{Light Gray} \\
1 & : & \text{White} \\
\end{align*}
\]
Categorical vs. Ordered

- Hue has no implicit ordering: use for categorical data
- Saturation and luminance do: use for ordered data

[Luminance: black, dark gray, light gray, white]

[Saturation: gray, purple, magenta, pink]

[Hue: blue, red, magenta, green, cyan, yellow]

[Munzner (ill. Maguire), 2014]
Continuous Colormap

US EPA Regional Oxidant Model -- Midwest Ozone (ppbv): June 26, 1987, 18:00

[Bergman et al., 1995]
Segmented Colormap

US EPA Regional Oxidant Model -- Midwest
Ozone (ppbv): June 26, 1987, 18:00

[Bergman et al., 1995]
Issues with Rainbow Colormaps

[Bergman et al., 1995]
Two-Hue Colormap

[Bergman et al., 1995]
Bivariate Colormaps

Munzner (ill. Maguire), 2014

[Munzner (ill. Maguire), 2014]
Assignment 4

- [link]
- Visualization using Tableau
- Due Monday, Nov. 21
Interaction

• The view changes over time
• Changes can affect almost any aspect of the visualization
  - encoding
  - arrangement
  - ordering
  - viewpoint
  - attributes being shown
  - aggregation level
Animated Transitions

[http://bl.ocks.org/mbostock/3943967]
Animated Transitions

[http://bl.ocks.org/mbostock/3943967]
Selection

• Selection is often used to initiate other changes
• User needs to select something to drive the next change
• What can be a selection target?
  - Items, links, attributes, (views)
• How?
  - mouse click, mouse hover, touch
  - keyboard modifiers, right/left mouse click, force
• Selection modes:
  - Single, multiple
  - Contiguous?
Highlighting

- Selection is the user action
- Feedback is important!
- How? Change selected item's visual encoding
  - Change color: want to achieve visual popout
  - Add outline mark: allows original color to be preserved
  - Change size (line width)
  - Add motion: marching ants
Highlighting

• Selection is the user action
• Feedback is important!
• How? Change selected item's visual encoding
  - Change color: want to achieve visual popout
  - Add outline mark: allows original color to be preserved
  - Change size (line width)
  - Add motion: marching ants

![Diagram of selected item with visual encoding changes]
Highlighting

# Multiple Views

<table>
<thead>
<tr>
<th>Encoding</th>
<th>Data</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Subset</td>
<td>None</td>
</tr>
<tr>
<td>Same</td>
<td>Redundant</td>
<td>Overview/Detail</td>
<td>Small Multiples</td>
</tr>
<tr>
<td>Different</td>
<td>Multiform</td>
<td>Multiform, Overview/Detail</td>
<td>No Linkage</td>
</tr>
</tbody>
</table>

[Muñzner (ill. Maguire), 2014]
Multiform

[Improvise, Weaver, 2004]
Multiform Views

• The same data visualized in different ways
• Does not need to be a totally different encoding (all choices need not be disjoint), e.g. horizontal positions could be the same
• One view becomes cluttered with too many attributes
• Consumes more screen space
• Allows greater separability between channels
Small Multiples & Brushing

- Same encoding, but different data in each view (e.g. SPLOM)

[http://bl.ocks.org/mbostock/4063663]
Overview-Detail View

[Image of a map with a detailed inset of a region.]

[Wikipedia]
Coordinated Views

Nasdaq 100 Index 1985/11/01-2012/06/29

Yearly Performance (radius: fluctuation/index ratio, color: gain/loss)

Days by Gain/Loss

Quarters

Day of Week

Days by Fluctuation(%)
Superimposed Layers

• Put different layers in the same spatial region, overlay information
• Usually each layer spans the entire view
• Must be **identifiable**: visually distinguishable
• Cartography has to deal with this a lot
• May be static or dynamic (user controls which layers are shown)
Superimposed Line Charts

[M. Bostock, http://bl.ocks.org/mbostock/3884955]
Aggregation and Filtering

• Too much data to display everything all the time
• Filtering: allow selectable subsets to be shown
• Aggregation: combine data items so that marks represent more than one item
Restaurant locations are derived from the New York City Department of Health and Mental Hygiene database. Due to the limitations of the Health Department’s database, some restaurants could not be placed.

By JEREMY WHITE

[Filtering NYC Health Dept. Restaurant Ratings]

[J. White, New York Times]
Aggregation: Histograms

- Very similar to bar charts
- Often shown without space between (continuity)
- Choice of number of bins
  - Important!
  - Viewers may infer different trends based on the layout

[Munzner (ill. Maguire), 2014]
Boxplots

- Show **distribution**
- Single value (e.g. mean, max, min, quartiles) doesn't convey everything
- Created by John Tukey who grew up in New Bedford!
- Show **spread** and **skew** of data
- Best for **unimodal** data
- Variations like vase plot for multimodal data
- Aggregation here involves many different marks
Boxplot Example

(a) Overall Activity
(b) Structural Activity
(c) Parameter Activity
(d) Layout Activity

Key:
- □: Task 1
- □: Task 2
- □: Task 3
- □: Task 4
- □: Task 5
- □: Task 6

Percentage (%)

[L. Lins et al., 2008]
Map Projection

[P. Foresman, Wikimedia]
Flattening the Sphere?

Central Meridian
(selected by mapmaker)

Great distortion at high latitudes
Examples of two rhumb lines (direction true between any two points)
Equator touches cylinder if cylinder is tangent
Reasonably true shapes and distances within 15 degrees of Equator

[USGS Map Projections]
Lambert Conformal Conic Projection

Two standard parallels (selected by mapmaker)

Large-scale map sheets can be joined at edges if they have the same standard parallels and scales

[USGS Map Projections]
Choropleth Map: What are Marks and Channels?

[M. Ericson, New York Times]
Choropleth Map

- Data: geographic geometry data & one quantitative attribute per region
- Tasks: trends, patterns, comparisons
- How: area marks from given geometry, color hue/saturation/luminance
- Scalability: thousands of regions

Design choices:
- Colormap
- Region boundaries (level of summarization)
Problem?

2008 Popular Vote

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Votes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obama</td>
<td>68 million</td>
</tr>
<tr>
<td>McCain</td>
<td>59 million</td>
</tr>
</tbody>
</table>

[M. Ericson, New York Times]
Problem?

2008 Popular Vote

- **Obama**: 68 million votes
- **McCain**: 59 million votes

Amount of red and blue shown on map

- **Obama**: 850,000 mi²
- **McCain**: 2,150,000 mi²

[Ericson, New York Times]
Adding Saturation

[M. Ericson, New York Times]
Two Variables

This map removes mostly uninhabited areas, revealing Mr. Bush's suburban and rural support in the East and South.

[quote]
*Areas with less than three people per square mile.

[M. Ericson, New York Times]
Size Encoding

[M. Ericson, New York Times]
Cartograms: Distort Spatial Regions

[Election Results by Population, M. Newman, 2012]
House Races: More Geographic Data?

House Races: Maps Aren't Always Best

Show results for: All Districts

<table>
<thead>
<tr>
<th>Democrats expected to win easily</th>
<th>Democrats expected to win narrowly</th>
<th>Tossup seats</th>
<th>Republicans expected to win narrowly</th>
<th>Republicans expected to win easily</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>District</strong></td>
<td><strong>Dem.</strong></td>
<td><strong>Rep.</strong></td>
<td><strong>% Rpt.</strong></td>
<td><strong>District</strong></td>
</tr>
<tr>
<td>Ala. 7</td>
<td>72%</td>
<td>28%</td>
<td>100%</td>
<td>Ala. 2</td>
</tr>
<tr>
<td>Ariz. 4</td>
<td>67%</td>
<td>33%</td>
<td>100%</td>
<td>Calif. 18</td>
</tr>
<tr>
<td>Calif. 1</td>
<td>63%</td>
<td>37%</td>
<td>100%</td>
<td>Calif. 16</td>
</tr>
<tr>
<td>Calif. 3</td>
<td>63%</td>
<td>37%</td>
<td>100%</td>
<td>Calif. 14</td>
</tr>
<tr>
<td>Calif. 5</td>
<td>63%</td>
<td>37%</td>
<td>100%</td>
<td>Calif. 12</td>
</tr>
<tr>
<td>Calif. 6</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 10</td>
</tr>
<tr>
<td>Calif. 7</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 8</td>
</tr>
<tr>
<td>Calif. 8</td>
<td>80%</td>
<td>20%</td>
<td>100%</td>
<td>Calif. 4</td>
</tr>
<tr>
<td>Calif. 9</td>
<td>84%</td>
<td>16%</td>
<td>100%</td>
<td>Calif. 2</td>
</tr>
<tr>
<td>Calif. 10</td>
<td>59%</td>
<td>41%</td>
<td>100%</td>
<td>Calif. 1</td>
</tr>
<tr>
<td>Calif. 12</td>
<td>76%</td>
<td>24%</td>
<td>100%</td>
<td>Calif. 14</td>
</tr>
<tr>
<td>Calif. 13</td>
<td>72%</td>
<td>28%</td>
<td>100%</td>
<td>Calif. 10</td>
</tr>
<tr>
<td>Calif. 14</td>
<td>69%</td>
<td>31%</td>
<td>100%</td>
<td>Calif. 8</td>
</tr>
<tr>
<td>Calif. 15</td>
<td>69%</td>
<td>31%</td>
<td>100%</td>
<td>Calif. 4</td>
</tr>
<tr>
<td>Calif. 16</td>
<td>69%</td>
<td>31%</td>
<td>100%</td>
<td>Calif. 2</td>
</tr>
<tr>
<td>Calif. 17</td>
<td>69%</td>
<td>31%</td>
<td>100%</td>
<td>Calif. 16</td>
</tr>
<tr>
<td>Calif. 18</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 14</td>
</tr>
<tr>
<td>Calif. 19</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 12</td>
</tr>
<tr>
<td>Calif. 20</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 10</td>
</tr>
<tr>
<td>Calif. 21</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 8</td>
</tr>
<tr>
<td>Calif. 22</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 4</td>
</tr>
<tr>
<td>Calif. 23</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 2</td>
</tr>
<tr>
<td>Calif. 24</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 16</td>
</tr>
<tr>
<td>Calif. 25</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 14</td>
</tr>
<tr>
<td>Calif. 26</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 12</td>
</tr>
<tr>
<td>Calif. 27</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 10</td>
</tr>
<tr>
<td>Calif. 28</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 8</td>
</tr>
<tr>
<td>Calif. 29</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 4</td>
</tr>
<tr>
<td>Calif. 30</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 2</td>
</tr>
<tr>
<td>Calif. 31</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 16</td>
</tr>
<tr>
<td>Calif. 32</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 14</td>
</tr>
<tr>
<td>Calif. 33</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 12</td>
</tr>
<tr>
<td>Calif. 34</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 10</td>
</tr>
<tr>
<td>Calif. 35</td>
<td>66%</td>
<td>34%</td>
<td>100%</td>
<td>Calif. 8</td>
</tr>
</tbody>
</table>

[NYTimes]