Data Visualization (DSC 530/CIS 568)

Color

Dr. David Koop
Arrange Tables

Express Values

Separate, Order, Align Regions
- Separate
- Order
- Align

Axis Orientation
- Rectilinear
- Parallel
- Radial

Layout Density
- Dense
- Space-Filling

[Munzner (ill. Maguire), 2014]
Express Values: Scatterplots

- Data: two quantitative values
- Task: find trends, clusters, outliers
- How: marks at spatial position in horizontal and vertical directions

- Correlation: dependence between two attributes
  - Positive and negative correlation
  - Indicated by lines
- Coordinate system (axes) and labels are important!
List Alignment: Bar Charts

- Data: one quantitative attribute, one categorical attribute
- Task: lookup & compare values
- How: line marks, vertical position (quantitative), horizontal position (categorical)
- What about **length**?
- Ordering criteria: alphabetical or using quantitative attribute
- Scalability: distinguishability
  - bars at least one pixel wide
  - hundreds

[Munzner (ill. Maguire), 2014]
Streamgraphs

- Include a time attribute
- Data: multidimensional table, one quantitative attribute (count), one ordered key attribute (time), one categorical key attribute
- + derived attribute: layer ordering (quantitative)
- Task: analyze trends in time, find (maxmial) outliers
- How: derived position+geometry, length, color
- Scalability: more categories than stacked bar charts

[Byron and Wattenberg, 2012]
Dot and Line Charts

- Data: one quantitative attribute, one **ordered** attribute
- Task: lookup values, find outliers and trends
- How: point mark and positions

- Line Charts: add **connection mark** (line)
- Similar to scatterplots but allow ordered attribute

[Munzner (ill. Maguire), 2014]
Proper Use of Line and Bar Charts

- What does the line indicate?
- Does this make sense?

[Adapted from Zacks and Tversky, 1999, Munzner (ill. Maguire), 2014]
Heatmaps & Bertin Matrices

Heatmaps & Bertin Matrices

Scatterplot Matrix & Parallel Coordinates

![Scatterplot Matrix](image)

**Scatterplot Matrix**
- Math vs. Math
- Math vs. Physics
- Math vs. Dance
- Math vs. Drama
- Physics vs. Math
- Physics vs. Physics
- Physics vs. Dance
- Physics vs. Drama
- Dance vs. Math
- Dance vs. Physics
- Dance vs. Dance
- Dance vs. Drama
- Drama vs. Math
- Drama vs. Physics
- Drama vs. Dance
- Drama vs. Drama

**Parallel Coordinates**

![Parallel Coordinates](image)

Munzner (ill. Maguire), 2014
Pie Chart Studies

[R. Kosara and D. Skau, 2016]
Test 1

• Information

• Next Wednesday, March 6

• Format:
  - Multiple Choice
  - Free Response

• Questions will involve research papers discussed in class
  - List of papers on the web site

• Questions may involve interpreting code or writing pseudocode
  - Array functions
  - D3 selections
Project Proposal

- **Information**
- **Due Next Friday, March 8**
- **Choices:**
  - [CreateVis] Create a complex, interactive visualization of a dataset
  - [Research] Research project that involves visualization (new technique, evaluation, etc.)
- **Check with me if you're interested in the research project option**
- **Start:**
  - [CreateVis] Looking for a dataset: List of lists of datasets
  - [Research] Surveying literature, identifying goals
Color
Color
Color and Light

- Color is a **perceptive** property: color depends on the eyes and brain.

- Visible light is a small portion of the **electromagnetic spectrum** which is composed of waves that at various frequencies (wavelengths), all traveling at the speed of light.
# Electromagnetic Spectrum

<table>
<thead>
<tr>
<th>Radiation Type</th>
<th>Wavelength (m)</th>
<th>Approximate Scale of Wavelength</th>
<th>Frequency (Hz)</th>
<th>Temperature of objects at which this radiation is the most intense wavelength emitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radio</td>
<td>$10^3$</td>
<td>Buildings</td>
<td>$10^4$</td>
<td>~272 °C</td>
</tr>
<tr>
<td>Microwave</td>
<td>$10^{-2}$</td>
<td>Humans</td>
<td>$10^8$</td>
<td>~173 °C</td>
</tr>
<tr>
<td>Infrared</td>
<td>$10^{-5}$</td>
<td>Butterflies</td>
<td>$10^{12}$</td>
<td>9,727 °C</td>
</tr>
<tr>
<td>Visible</td>
<td>$0.5 \times 10^{-6}$</td>
<td>Needle Point</td>
<td>$10^{15}$</td>
<td>~10,000 K</td>
</tr>
<tr>
<td>Ultraviolet</td>
<td>$10^{-8}$</td>
<td>Protozoans</td>
<td>$10^{16}$</td>
<td>10,000,000 K</td>
</tr>
<tr>
<td>X-ray</td>
<td>$10^{-10}$</td>
<td>Molecules</td>
<td>$10^{18}$</td>
<td>~10,000,000 K</td>
</tr>
<tr>
<td>Gamma ray</td>
<td>$10^{-12}$</td>
<td>Atoms</td>
<td>$10^{20}$</td>
<td></td>
</tr>
</tbody>
</table>

Penetrates Earth's Atmosphere?
- Y: Yes
- N: No

Radiation Type

- Radio
- Microwave
- Infrared
- Visible
- Ultraviolet
- X-ray
- Gamma ray

Frequency (Hz)

- $10^4$
- $10^8$
- $10^{12}$
- $10^{15}$
- $10^{16}$
- $10^{18}$
- $10^{20}$

Temperature of objects at which this radiation is the most intense wavelength emitted

- 1 K: ~272 °C
- 100 K: ~173 °C
- 10,000 K: 9,727 °C
- 10,000,000 K: ~10,000,000 °C

[Wikimedia, NASA]
Light Reflection & Absorption

How does light work?

[via M. Meyer]
Color ≠ Wavelength

But rather, a combination of wavelengths and energy [via M. Meyer]
Human Color Perception

- Humans do not detect individual wavelengths of light
- Use rods and cones to detect light
  - rods capture intensity
  - cones capture color
Human Color Perception

- Humans are **trichromatic**—we have three different types of cones:
  - S (430nm): blue
  - M (540nm): green
  - L (570nm): "red"
- Note that the response curves **overlap**
- Spectra of visible light are "covered" by these responses
- Three numbers -> color

[Vanessaezekowitz at en.wikipedia]
Human Color Perception

Metamerism: same three responses == same color
Metamerism

[Diagram showing relative power vs. wavelength (nm) with colors representing different wavelengths with peaks at various intervals.]

[via M. Meyer]
Color

• Cones respond to different areas of the visible light spectrum
• Cover all wavelengths but certain wavelengths generate greater responses
• Color is determined by calculations based on the responses from the different cones
• Opponent Process Theory: three "opponent" channels
  - Light/Dark
  - Blue/Yellow
  - Red/Green
• Opposite colors are not perceived together
Opponent Process Theory

[Machado et. al, 2009]
Color Blindness

[Ishihara (Plate 9) via Wikipedia]
Color Blindness

- Sex-linked: 8% of males and 0.4% of females of N. European ancestry
- Abnormal distribution of cones (e.g. missing the S, M, or L types)
- Either dichromatic (only two types of cones) or anomalous trichromatic (one type of cones has a defect)
  - Protanopia (L missing), Protanomaly (L defect)
  - Deuteranopia (M missing), Deuteranomaly (M defect) [Most Common]
  - Tritanopia (S missing), Tritanomaly (S defect) [Rare]
- Dichromacy is rarer than anomalous trichromacy
- Opponent process model explains why colors cannot be differentiated
Color Blindness

Normal

Protanopia

Deuteranopia

Tritanopia

[via M. Meyer]
Simulating Color Blindness

<table>
<thead>
<tr>
<th>Protagania</th>
<th>Photop. Subst.</th>
<th>Scale Ratio</th>
<th>0.96*Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Image 1]</td>
<td>[Image 2]</td>
<td>[Image 3]</td>
<td>[Image 4]</td>
</tr>
</tbody>
</table>

[Image 5]  

[D. Koop, DSC 530, Spring 2019]

[Machado et. al, 2009]
Simulating Color Blindness

(Machado et. al, 2009)
Simulating Color Blindness

[Machado et. al, 2009]
# Simulating Deuteranopia (Colormaps)

<table>
<thead>
<tr>
<th>Colormap</th>
<th>Simulation of green deficient colour blindness (deuteranopia) at 0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>RdY1Gn</td>
<td><img src="image" alt="RdY1Gn Colormap" /></td>
</tr>
<tr>
<td>Dark2</td>
<td><img src="image" alt="Dark2 Colormap" /></td>
</tr>
<tr>
<td>jet</td>
<td><img src="image" alt="jet Colormap" /></td>
</tr>
<tr>
<td>Spectral</td>
<td><img src="image" alt="Spectral Colormap" /></td>
</tr>
<tr>
<td>Purples</td>
<td><img src="image" alt="Purples Colormap" /></td>
</tr>
<tr>
<td>RdY1Bu</td>
<td><img src="image" alt="RdY1Bu Colormap" /></td>
</tr>
<tr>
<td>Y1GnBu</td>
<td><img src="image" alt="Y1GnBu Colormap" /></td>
</tr>
<tr>
<td>linearL</td>
<td><img src="image" alt="linearL Colormap" /></td>
</tr>
<tr>
<td>parula</td>
<td><img src="image" alt="parula Colormap" /></td>
</tr>
</tbody>
</table>

[@neilrkaye, reddit]
Simulating Deuteranopia (Colormaps)

Simulation of green deficient colour blindness (deuteranopia) at 0%

- RdYlGn
- Dark2
- jet
- Spectral
- Purples
- RdYlBu
- YlGnBu
- linearL
- parula

@neilrkaye

[@neilrkaye, reddit]
Primary Colors?

- Red, Green, and Blue
- Red, Yellow, and Blue
- Orange, Green, and Violet
- Cyan, Magenta, and Yellow
Primary Colors?

- Red, Green, and Blue
- Red, Yellow, and Blue
- Orange, Green, and Violet
- Cyan, Magenta, and Yellow
- All of the above!
Color Addition and Subtraction

- RGB (Red, Green, Blue) for additive color mixing
- CMY (Cyan, Magenta, Yellow) for subtractive color mixing

- Colors combine in the additive model
- Colors subtract in the subtractive model
Color Spaces and Gamuts

- All colors visible to the average human eye are contained inside the diagram.
- The colors along any line between two points can be made by mixing the colors at the end points. In this case, Green + Red = Yellow. This line represents the most saturated colors.
- The edge of the diagram, called the spectral locus, represents pure, monochromatic light measured by wavelength in nanometers. These are the least saturated colors.
- The least saturated colors are at the center, emanating from white.
- Color gamut: subset of colors that can be represented by mixing the colors at its corners.
- "Line of purples": these colors are fully saturated but can only be made by mixing two colors (red and blue).

[Anatomy of a CIE Chromaticity Diagram]
Color Spaces and Gamuts

• **Color space**: the organization of all colors in space
  - Often human-specific, what we can see (e.g. CIELAB)

• **Color gamut**: a subset of colors
  - Defined by corners on in the color space
  - What can be produced on a monitor (e.g. using RGB)
  - What can be produced on a printer (e.g. using CMYK)
  - The gamut of your monitor != the gamut of someone else's != the gamut of a printer
Color Models

- **A color model** is a representation of color using some basis.
- RGB uses three numbers (red, blue, green) to represent color.
- Color space ~ color model, but there can be many color models used in the same color space (e.g. OGV).
- **Hue-Saturation-Lightness (HSL)** is more intuitive and useful:
  - Hue captures pure colors.
  - Saturation captures the amount of white mixed with the color.
  - Lightness captures the amount of black mixed with a color.
  - HSL color pickers are often circular.
- **Hue-Saturation-Value (HSV)** is similar (swap black with gray for the final value), linearly related.
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: blue, red, magenta, green, cyan, yellow]
[L from HSL: all the same]
[Luminance: black, dark gray, gray, light gray, white]
[L*]

[Munzner (ill. Maguire), 2014 (based on Stone, 2006)]
Perceptually Uniform Color Spaces

- \( L^*a^*b^* \) allows perceptually accurate comparison and calculations of colors

[J. Rus, CC-BY-SA (changed to horizontal layout)]
Luminance Perception (Spatial Adaption)

E. H. Adelson, 1995

Edward H. Adelson

[E. H. Adelson, 1995]
Luminance Perception (Spatial Adaption)

[Edward H. Adelson, 1995]
Simultaneous Contrast
Simultaneous Contrast
Simultaneous Contrast
Simultaneous Contrast
What colors?
What colors?

Red, yellow, blue

Purple, orange do not exist!
What does this mean for visualization?
What does this mean for visualization?

- We need to be aware of colorblindness when encoding via color
- Our brains may misinterpret color (surrounding colors matter!) even if we aren't colorblind
- Be careful! Don't assume that adding color always works the way you intended
- Use known colormaps when possible
Violations of CIELAB Assumptions

• Examine how visualization design affects color perception

• CIELAB:
  - Approximately perceptually linear
  - 1 unit of Euclidean distance = 1 Just Noticeable Difference (JND)
  - JND: people detect change at least 50% of the time

• Assumptions CIELAB makes:
  - Simple world
  - Isolation
  - Geometric

[D. Szafir, 2017]
Simple World Assumption

[D. Szafir, 2017]
Problems with Simple World Assumption

Viewing Distance

Environmental Surround

Ambient Illumination

Direct Illumination

Gamma, Whitepoint, Resolution, Peak Color Outputs

Viewing Population

Crowdsourced Sampling

Szafir, Stone, & Gleicher, 2014

Reinecke, Flatla, & Brooks, 2016

[D. Szafir, 2017]
Isolation Assumption
Problems with Isolation Assumption
Geometric Assumption

[Visualizations violate three CIELAB assumptions]

[D. Szafir, 2017]
Size Problem with Geometric Assumption

Visualizations violate three CIELAB assumptions:

- Size-based Sampling
- Carter & Silverstein, 2010
- Stone, Szafir, & Setlur, 2014

[D. Szafir, 2017]
Shape Problem with Geometric Assumption

[D. Szafir, 2017]
Types of Geometry

Diagonally Symmetric Marks

Elongated Marks

Asymmetric Marks

Area Marks

[ D. Szafir, 2017]
Run the tests!
6 (diameters, within) × 6 (color differences, within) × 3 (color axis, between)

81 participants on Mechanical Turk (5,668 trials)

[D. Szafir, 2017]
Point Size: consistent with previous results

50% JND for Scatterplot Points

Stone et al.  \( \Delta E = 1.0 \)

[D. Szafir, 2017]
Bar Thickness and Length: longer bars help

50% JND for Bars

Bar Length encoded as point size

JND in CIELAB

Bar Thickness

Visual Angle

Δb* for points

Δa* for points

ΔL* for points

[D. Szafir, 2017]
Line Thickness: better than points

50% JND for Lines

JND in CIELAB

Line Thickness

Visual Angle

Δb* for points

Δa* for points

ΔL* for points

R² = .97
R² = .93
R² = .90
Color perception in real-world visualizations is complicated
Akiyoshi Kitaoka's Illusion pages
Colormap

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

**Binary**

<table>
<thead>
<tr>
<th>y</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Diverging**

<table>
<thead>
<tr>
<th>-1</th>
<th>0</th>
<th>+1</th>
</tr>
</thead>
</table>

**Categorical**

<table>
<thead>
<tr>
<th>T</th>
<th>F</th>
<th>A</th>
</tr>
</thead>
</table>

**Sequential**

<table>
<thead>
<tr>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
</table>

[Munzner (ill. Maguire), 2014]
Categorical vs. Ordered

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

[Munzner (ill. Maguire), 2014]
Categorical Colormap Guidelines

• Don't use too many colors (~12)
• Remember your background has a color, too
• Nameable colors help
• Be aware of luminance (e.g. difference between blue and yellow)
• Think about other marks you might wish to use in the visualization
Categorical Colormaps
Number of distinguishable colors?

[Sinha & Meller, 2007]
Number of distinguishable colors?

[Sinha & Meller, 2007]
Discriminability

- Often, fewer colors are better
- Don't let viewers combine colors because they can't tell the difference
- Make the combinations yourself
- Also, can use the "Other" category to reduce the number of colors