Part III Visual Encoding

Colors

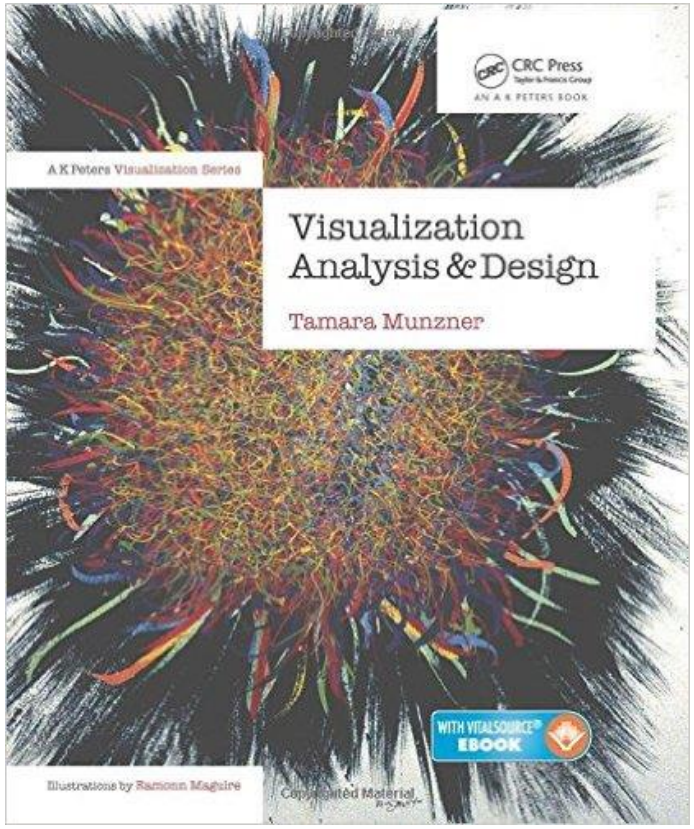
Computers are useless. They can only give answers.

Pablo Picasso

Fall 2020

Source of This Unit

- **Material of this unit is based on Chapter 10 of Tamara Munzner, *Visualization Analysis and Design*, AK Peters/CRC Press, 2014.**



Color Spaces: 1/6

- The **color space** of what colors the human visual system can detect is three dimensional.
- Some of these color spaces are convenient for computer manipulation, while others are a better match with the characteristics of human vision.
- The most common color space is the **RGB** system. Each color in this system is represented by the weights assigned to the **red**, **green** and **blue** channels.

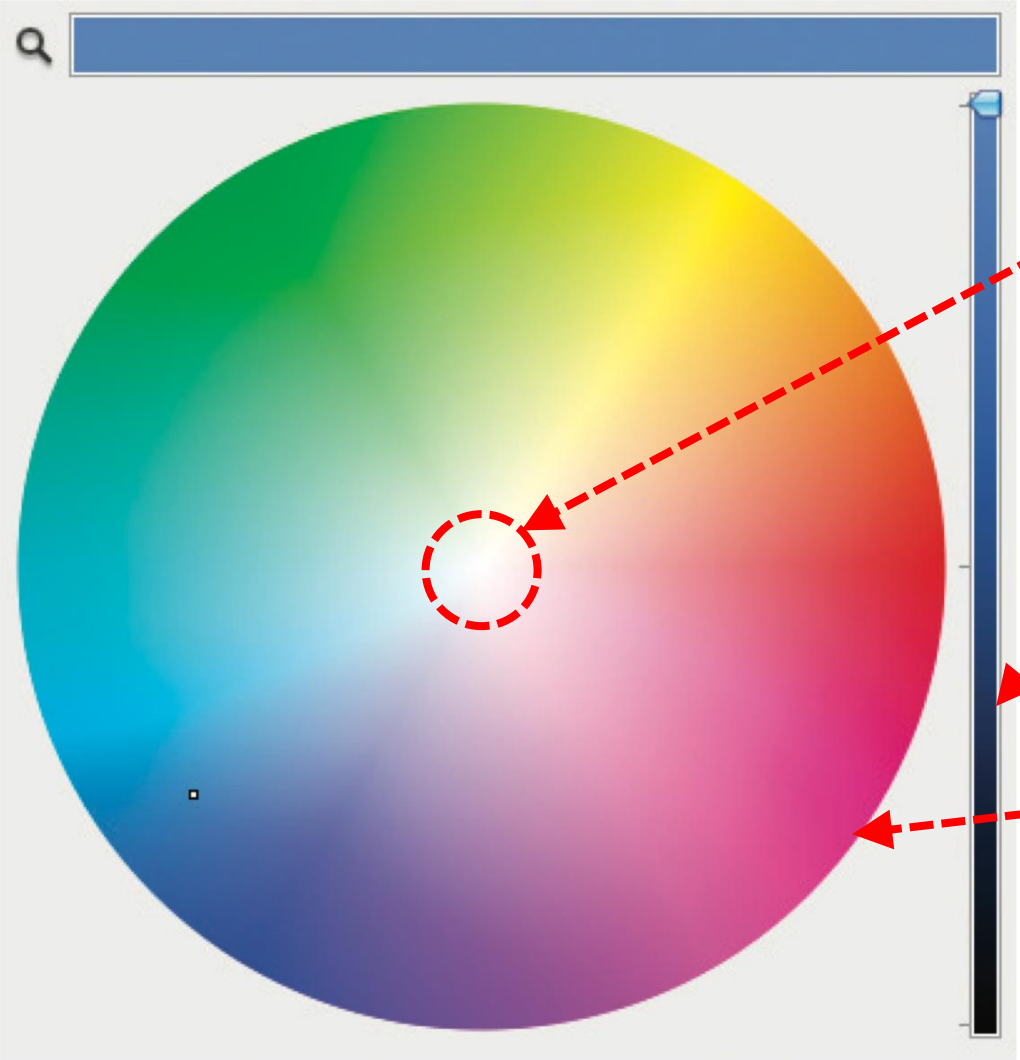
Color Spaces: 2/6

- The **HSL** system is more intuitive and is being used by artists and designers.
- Here, the HSL system requires **H**ue, **S**aturation and **L**ightness axes, and each color is defined by the weights assigned to the hue, saturation and lightness axes.
- The **hue** axis captures what we normally think of as pure colors (e.g., red, green, blue, yellow, purple, etc.) that are not mixed with pure white and pure black.

Color Spaces: 3/6

- The **saturation** axis is the amount of white mixed with a pure color.
- For example, pink is a partially desaturated red.
- The **lightness** axis is the amount of black mixed with a color.
- A common design for color pickers is a disk with white at the center and the hue axis wrapped abound the outside, with separate linear control for the amount of darkness vs. lightness.
- The **HSV** space is similar, where *V* is for grayscale and is linear to lightness *L*.

Color Spaces: 4/6



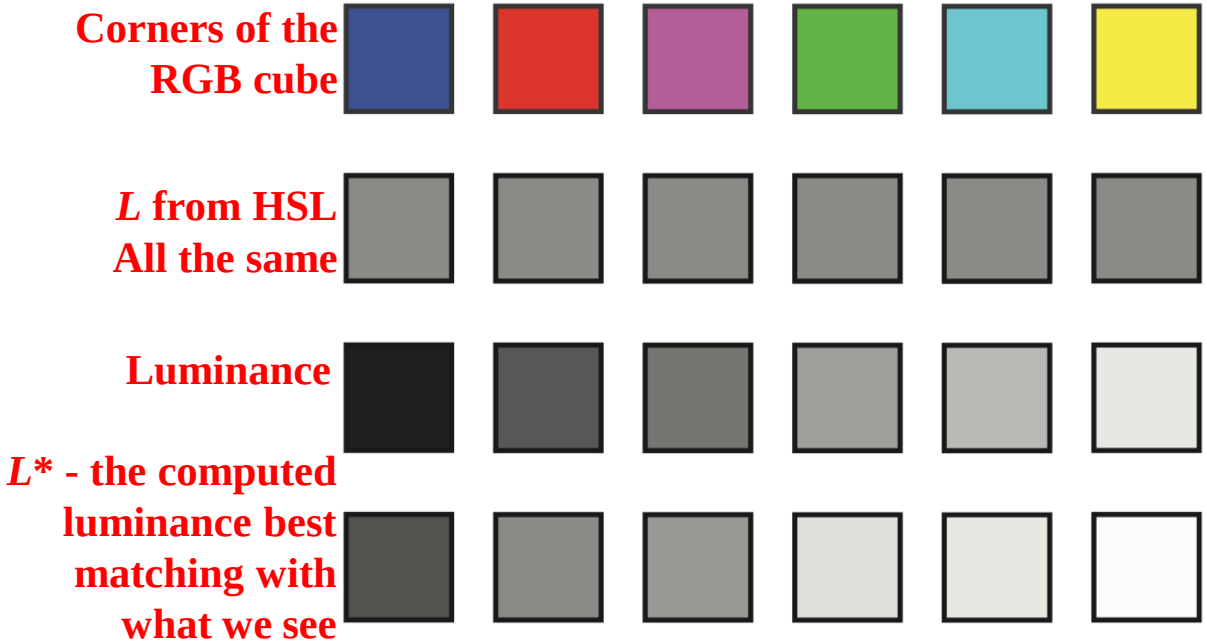
Color white is at the center

Lightness vs. darkness is selected via a linear slider

Most saturated colors are around the perimeter of the disk

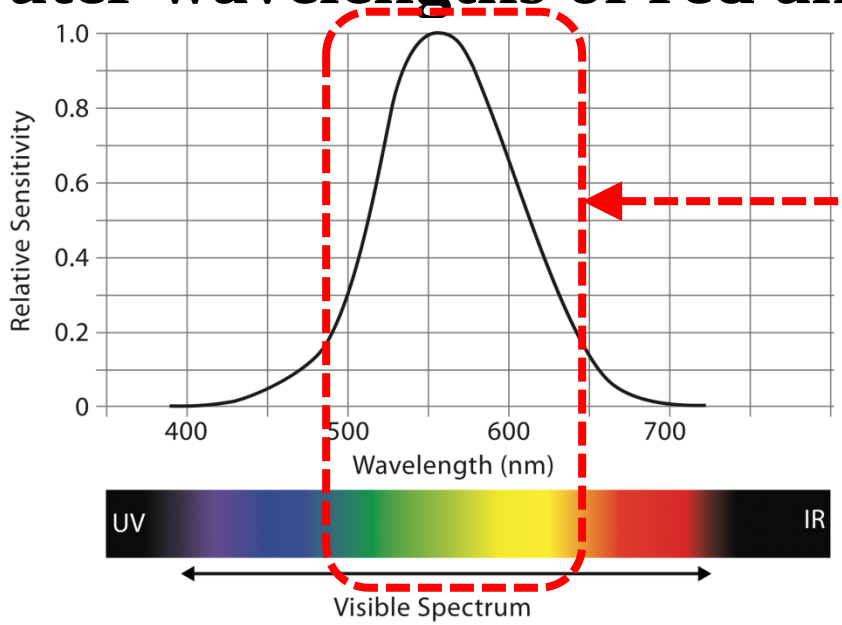
Color Spaces: 5/6

- **Note that the HSL space does not truly reflect how we perceive color.**
- **In particular, the lightness *L* is wildly different from how we perceive luminance.**



Color Spaces: 6/6

- The amount of luminance that human perceive depends on the wavelength.
- Human are much more sensitive to middle wavelengths of green and yellow than to the outer wavelengths of red and blue.



We are more sensitive to middle wavelengths of green and yellow.

Luminance, Saturation and Hue Channels: 1/4

- ***Luminance***, ***hue*** **and** ***saturation*** **can be used as a magnitude channel and as an identity channel.**
- **The *magnitude luminance* channel is suitable for ordered data types.**
- **However, *luminance* is a low accuracy channel in terms of perceiving whether noncontiguous regions have the same luminance because of contrast effects.**
- **Limit grayscale bins to no more than 4.**

Luminance, Saturation and Hue Channels: 2/4

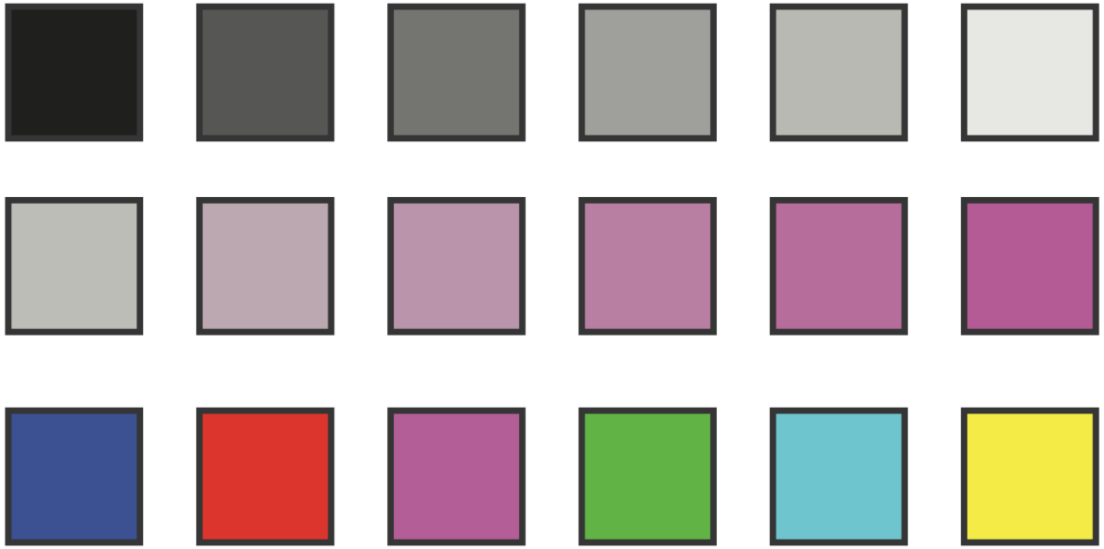
- **The *magnitude saturation* channel is also suitable for ordered data types.**
- ***Saturation* shares the same problem of low accuracy for noncontiguous regions. The number of discriminable steps for saturation is low: around 3 bins.**
- ***Saturation* may interact with the size channel: it is difficult to perceive in small region than in large ones.**
- **Use bright, highly saturated colors so that small regions are more distinguishable.**

Luminance, Saturation and Hue Channels: 3/4

- **The *identity* channel of *hue* is very effective for categorical data and showing groups. It is the highest ranked channel for categorical data after spatial position.**
- **However, *hue* and *size* can interact: *hue* is harder to distinguish in small regions.**
- **We can make fine distinctions in *hue* for contiguous regions, but we have very limited discriminability between separate regions.**

Luminance, Saturation and Hue Channels: 4/4

- *Hue* **does not** have an implicit perceptual ordering.
- But, people can and do learn from conventions, such as green-yellow-red traffic lights or the order or colors in the rainbow.



Luminance is ordered from dark to bright

Saturation is also ordered from lightly saturated to highly saturated

Hue does not have a similar perceived ordering

Transparency

- ***Transparency*** **is a channel related to the other three color channels, but cannot be used alone.**
- ***Transparency*** **coding interacts strongly with** ***luminance*** **and** ***saturation*** **and should not be used with them.**
- **It can be used with hue encoding with a very small number of discriminable steps, most frequently just two.**
- **Most often** ***transparency*** **is used with superimposed layers.**

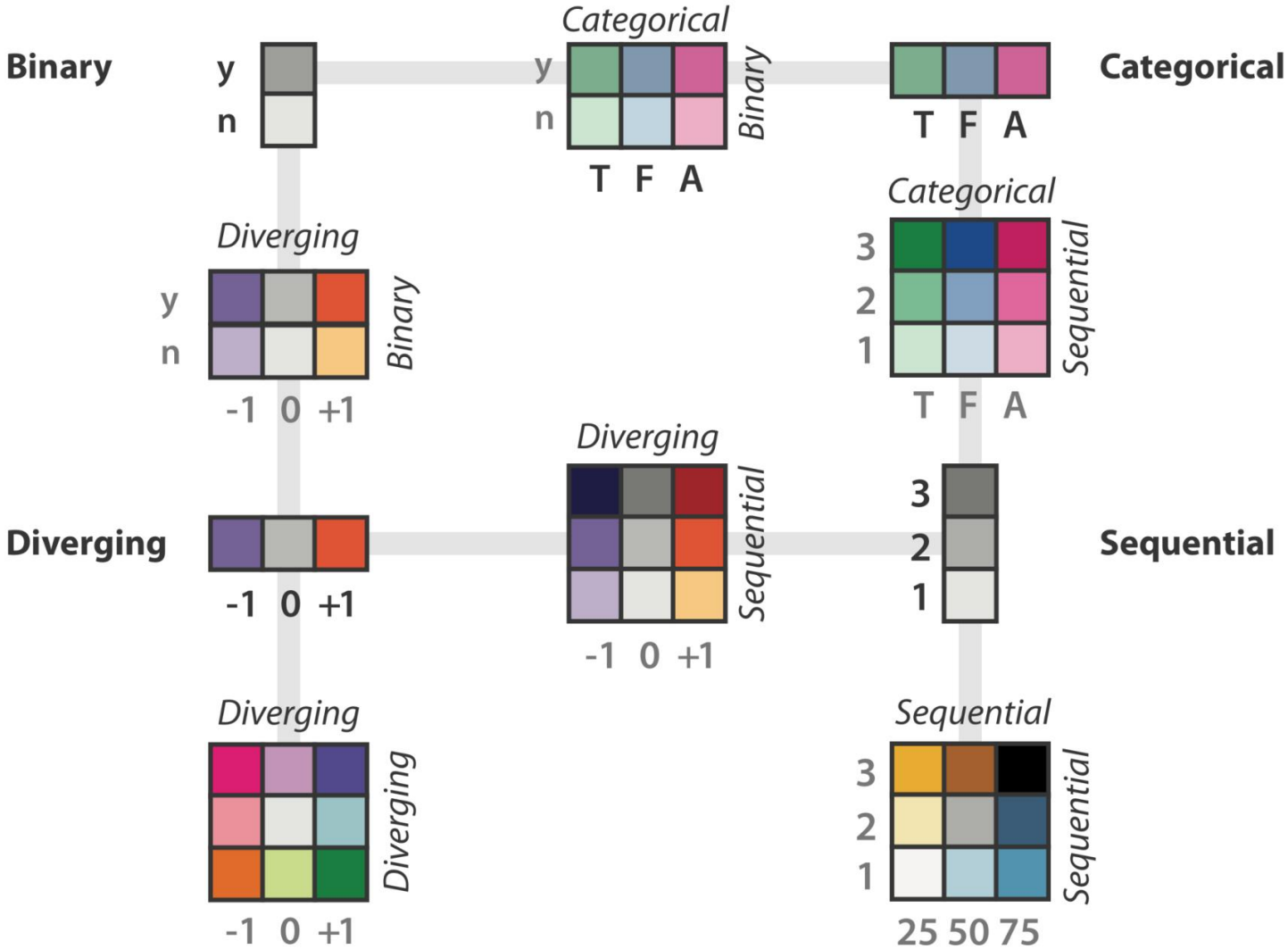
Colormaps: 1/3

- A **colormap** defines a mapping between colors and data values, a visual encoding with colors.
- **Colormap**s can be categorical or ordered, and ordered colormaps can be sequential or diverging.
- **Colormap**s can also be a continuous range of values, or segmented into discrete bins of colors.
- Continuous **colormap**s are used for showing quantitative attributes, especially those associated with inherently spatial fields.

Colormaps: 2/3

- Segmented **colormap**s are suitable for categorical data.
- For ordinal data, segmented **colormap**s will emphasize its discrete nature, while continuous **colormap**s will emphasize its ordered nature.
- Bivariate **colormap**s encode two attributes at the same time.
- It can be more difficult for people to interpret when both attributes have multiple levels.

Colormaps: 3/3



Categorical Colormaps: 1/11

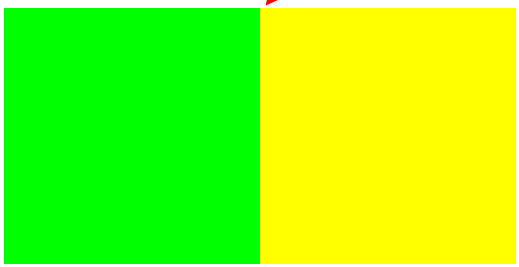
- A categorical **colormap** uses color to encode categories and groupings.
- Categorical **colormap**s are typically designed by using color as an integral identity channel to encode a single attribute, rather than to encode three completely separate attributes with the three channels of hue, saturation and luminance.
- The number of discriminable colors for coding small separated regions is limited to between 6 and 12 bins.

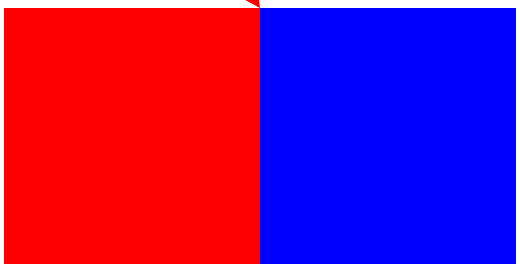
Categorical Colormaps: 2/11

- **A good set of initial choices are the fully saturated and easily nameable colors, which are also the opponent color axes: red, blue, green and yellow.**
- **When more colors are needed, there are orange, brown, pink, magenta, purple and cyan.**
- **Be careful with luminance contrast: for some uses, the colors should be close in luminance to avoid major differences in salience and to ensure that all can be seen against the background.**

Categorical Colormaps: 3/11

- **Example**: Fully saturated green and yellow will have much less luminance contrast against a white background than red and blue.

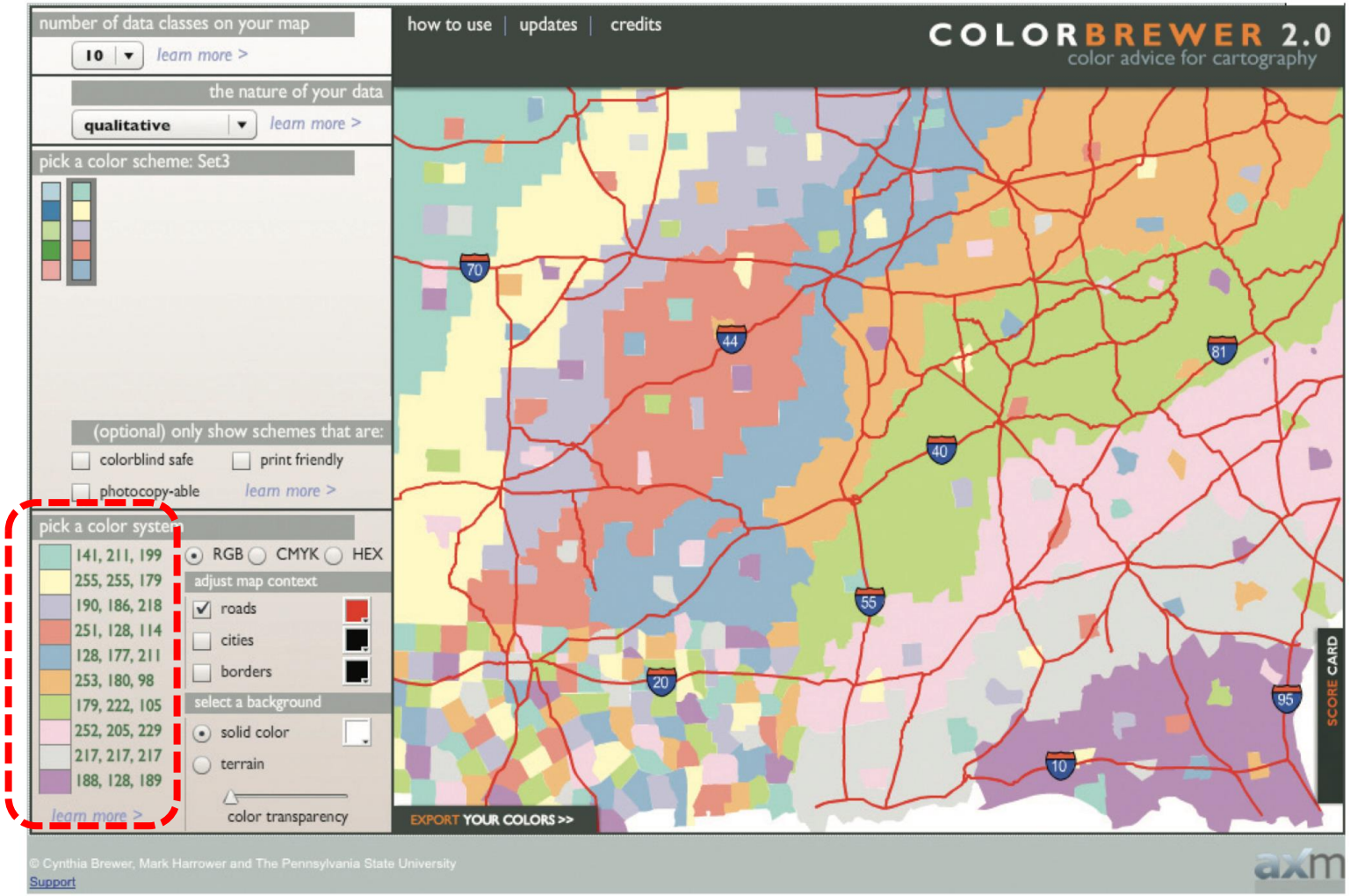




Categorical Colormaps: 4/11

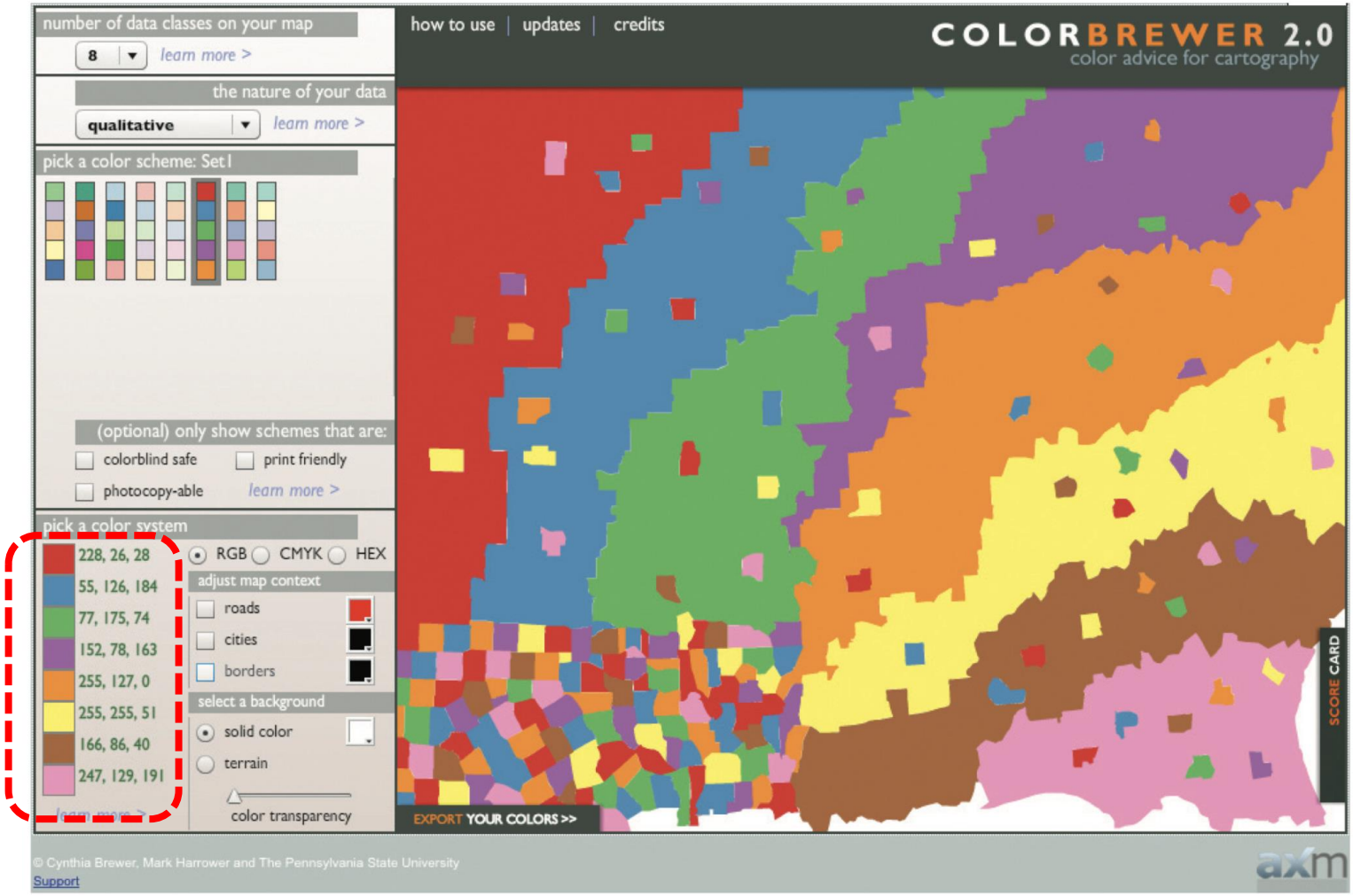
- For other uses, colors should be sufficiently different in luminance that they can be distinguished even in black and white.
- **Colormap**s for small regions such as lines should be highly saturated, but large regions such as areas should have low saturation.
- An appropriate **colormap** may depend on the mark type being used.

Categorical Colormaps: 5/11



This 10-color low saturation map works well with large areas.

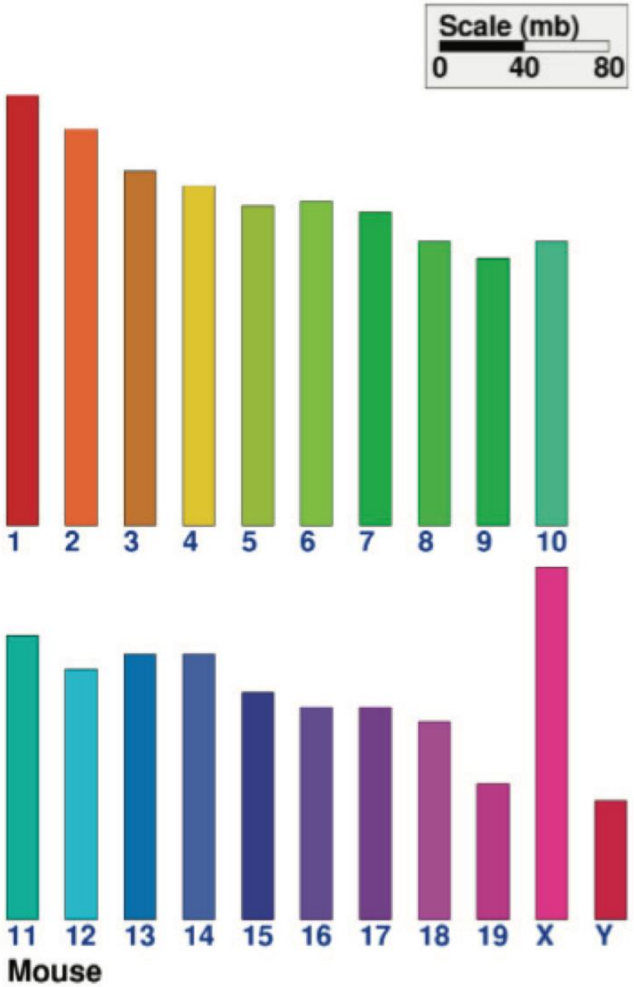
Categorical Colormaps: 6/11



This 8-color high saturation map works poorly with large areas. It would be better suited for small regions.

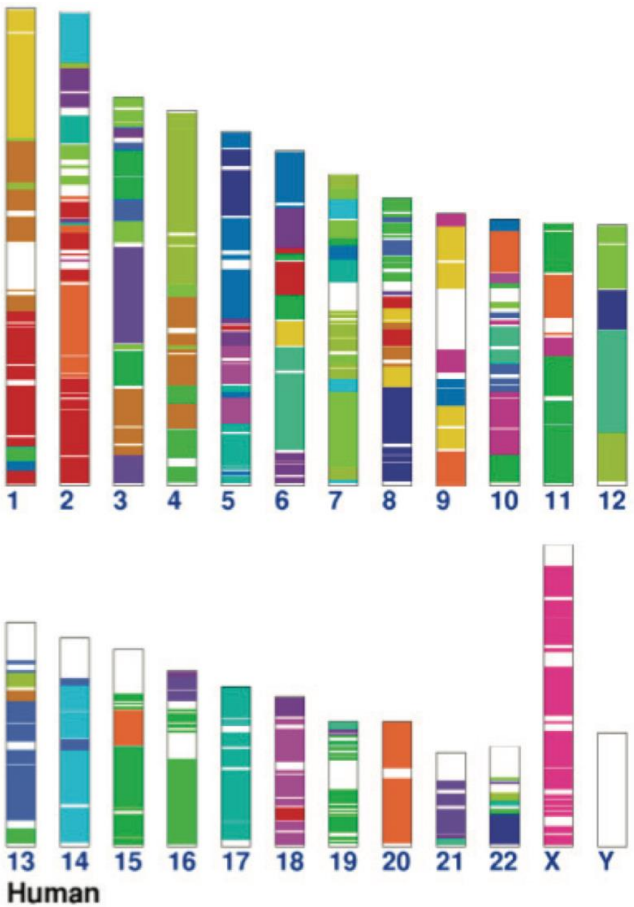
Categorical Colormaps: 7/11

- The 21 colors used as an index for each mouse chromosome can indeed be distinguished in large regions next to each other.
- These colors act as a legend and an index, because regions are large and the most subtle differences are between regions that are right next to each other.



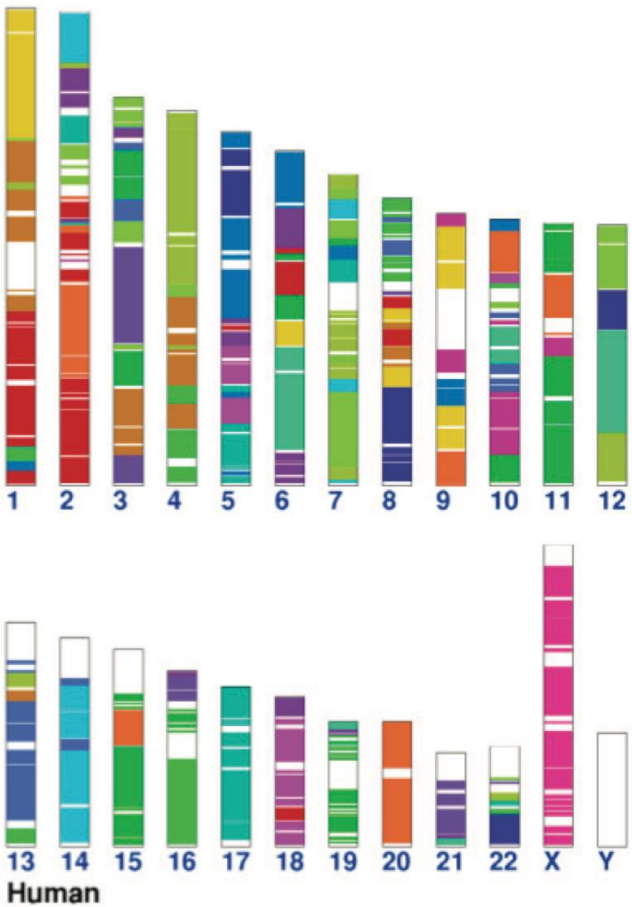
Categorical Colormaps: 8/11

- The regions of the human chromosomes that correspond to those in the mouse chromosomes have been colored to illustrate how genomic regions have moved around as the species evolved independently from each other after diverging from a common ancestor.

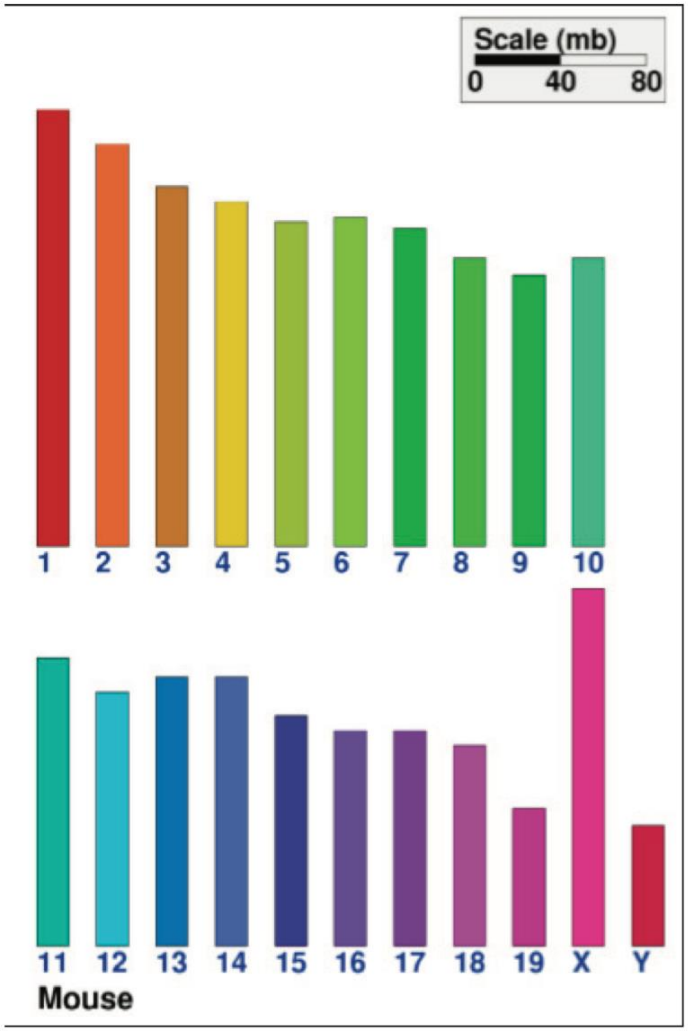


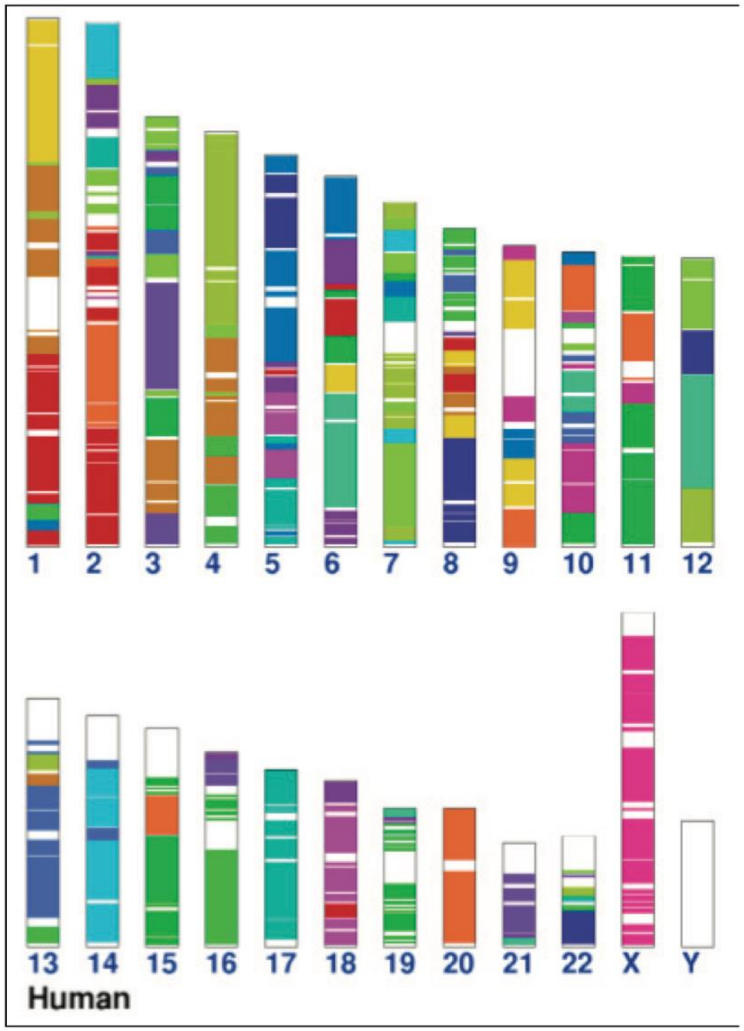
Categorical Colormaps: 9/11

- The colored regions are much smaller and noncontiguous.
- These 21 colors are not distinguishable from each other.
- Only 3 (resp., 3) bins of green (resp., pinks and purples) in the human (resp., mouse) view rather than 7 (resp., 5) in the mouse view.



Categorical Colormaps: 10/11





Categorical Colormaps: 11/11

- **How to solve this problem?**
  - **Transform the data that takes into account of the nature of the data and task, so that each bin can be encoded with a distinguishable color.**
    - **A. Find possible hierarchical structure that can be used to derive meaningful aggregated groups.**
    - **B. Filter the attributes to only encode a small set of the most important ones with color, and aggregate all of the reset into a new category.**
  - **Use a different encoding idiom that uses other visual channels instead of, or in addition to, the color channel alone.**

Ordered Colormaps: 1/13

- An **ordered** colormap is appropriate for encoding ordinal or quantitative attributes.
- The two major variants (e.g., **sequential** and **diverging**) of continuous colormaps for ordered data have expressiveness characteristics that should match up with the attribute type.

Ordered Colormaps: 2/13

- A **sequential** colormap ranges from a minimum value to a maximum value.
- If only the luminance channel is used, the result is a grayscale ramp.
- When hue is added, one end of the map is a specific hue at full saturation and brightness.
- If saturation is a variable, the other end is pale or white. If luminance is the varying quantity, the other end is dark or black.

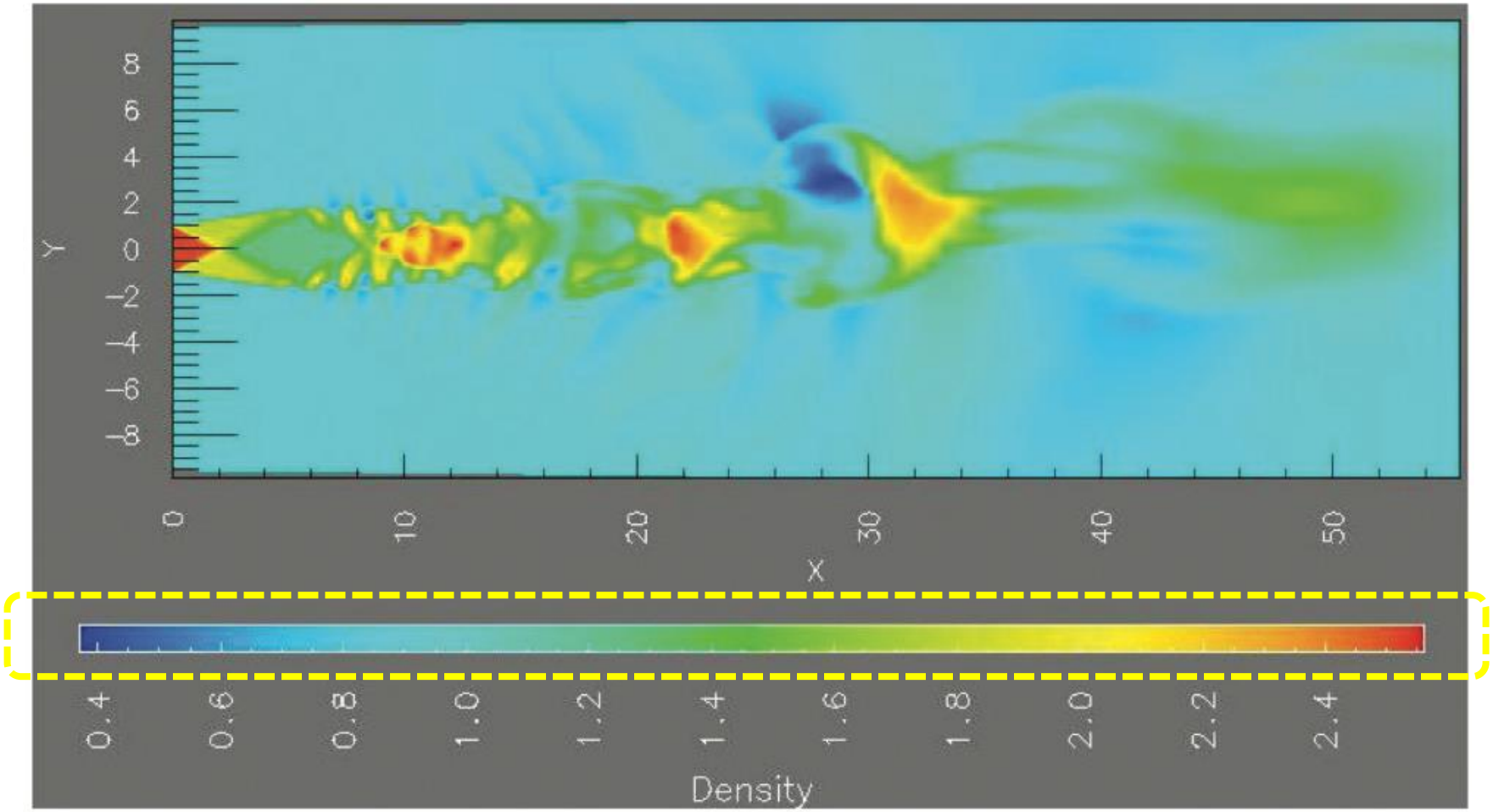
Ordered Colormaps: 3/13

- A **diverging** colormap has two hues at the endpoints and a neutral color as a midpoint.
- This neutral color may be white, gray or black, or a high-luminance color such as yellow.

Ordered Colormaps: 4/13

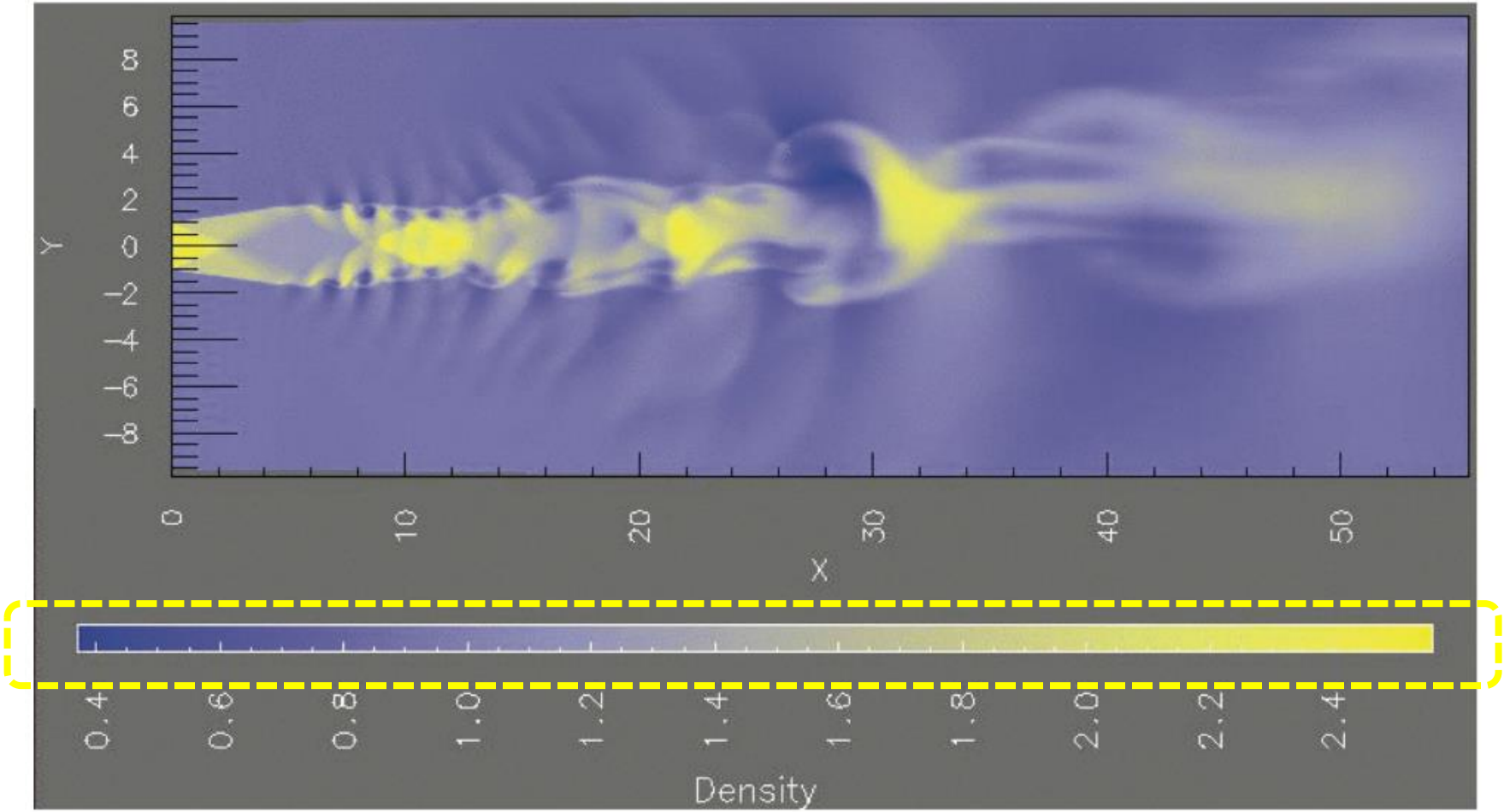
- **The number of unique hues to be used in continuous colormaps depends on what level of structure should be emphasized: the *high-level* structure, the *middle range* of local neighborhoods, or *fine-grained* details.**

Ordered Colormaps: 5/13



Use many hues as in a rainbow colormap to emphasize mid-scale neighborhood

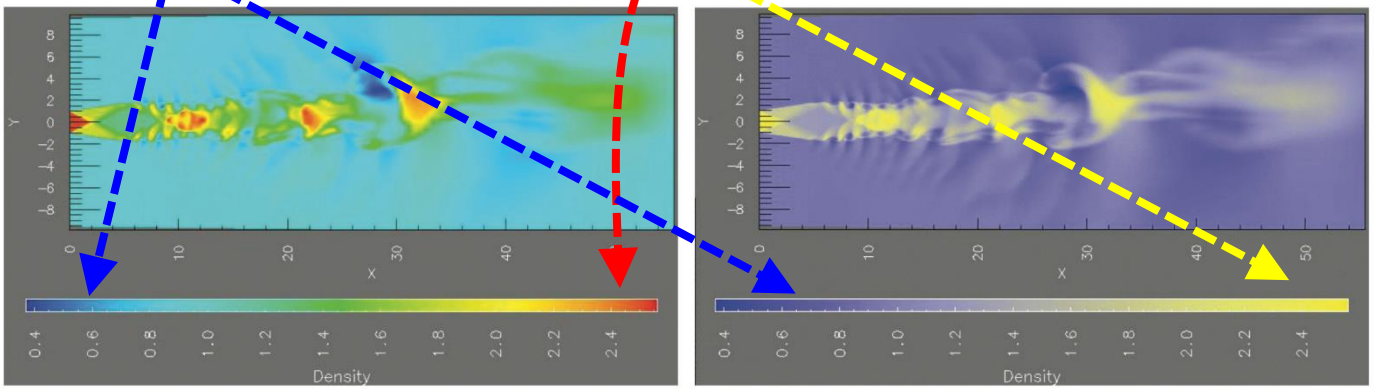
Ordered Colormaps: 6/13



Use a 2-hue blue-yellow colormap to emphasize large-scale

Ordered Colormaps: 7/13

- With a rainbow colormap, people can easily discuss specific subranges because the differences are easily nameable: “the red part vs. the purple part”.
- With a colormap only using saturation or luminance changes, people can only talk about the blue side and yellow side.

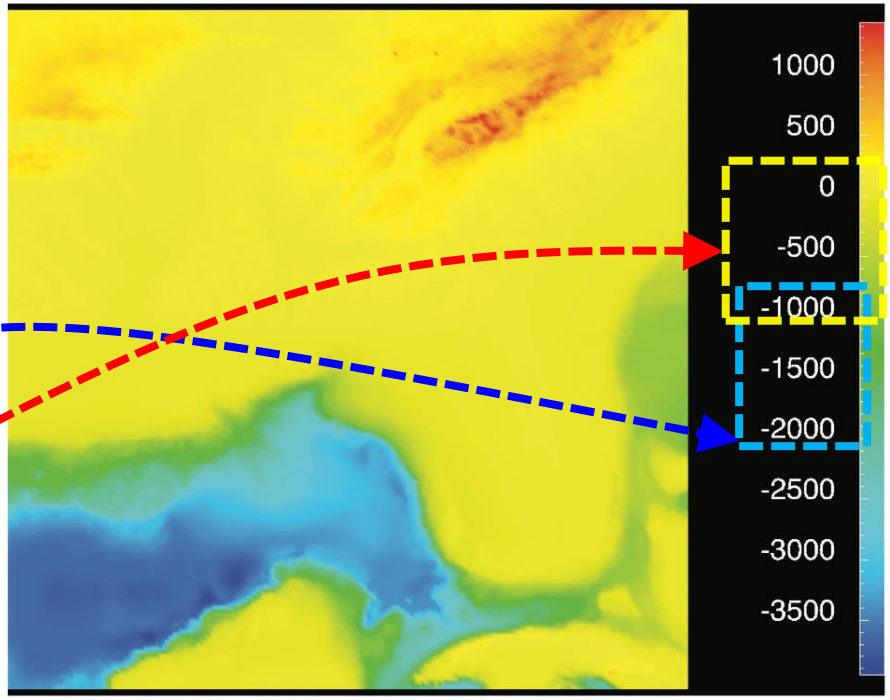


Ordered Colormaps: 8/13

- **Rainbow colormaps, which are usually a default choice in many software packages, suffer from three serious problems.**
  1. **Hue is used to indicate order, despite being an identity channel without an implicit perceptual ordering.**
  2. **The scale is not perceptually linear: steps of the same size at different points in the colormap range are not perceived equally by our eyes. Expressivity mismatch.**
  3. **Fine details cannot be perceived with the hue channel; the luminance channel would be a better choice. Accuracy mismatch.**

Ordered Colormaps: 9/13

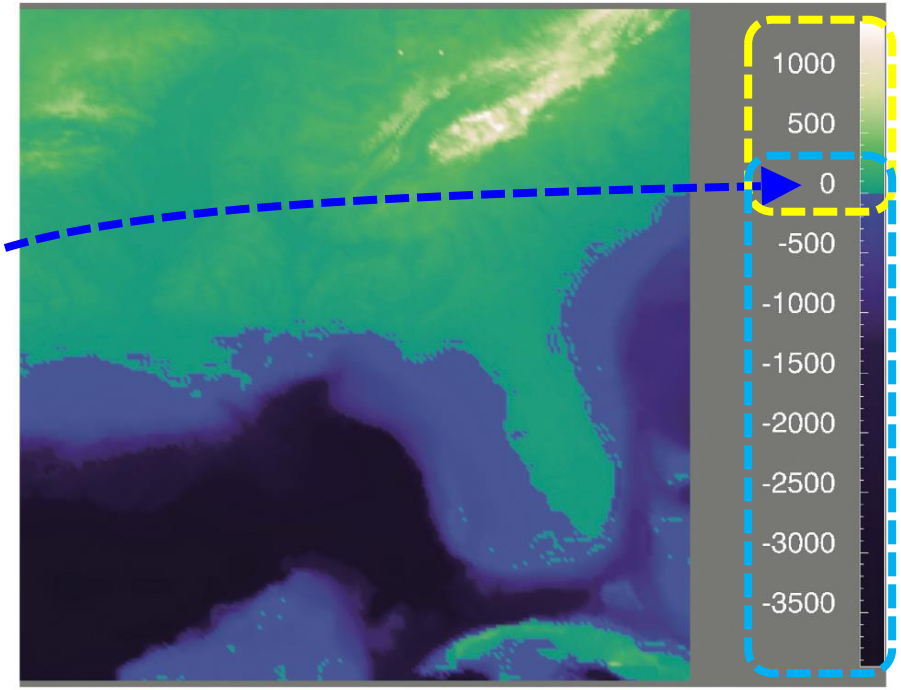
- A range of 1000 units has different characteristics depending on where within the colormap it falls.
- The range [-2000, -1000] has three distinct colors (i.e., cyan, green and yellow).
- The range of the same size [-1000, 0] looks all yellow!



What and where is it?

Ordered Colormaps: 10/13

- **A color map that combines monotonically increasing luminance with multiple hues for semantic categories with a clear segmentation at the 0 point.**
- **In doing so, it shows high-level, mid-level, and low-level structure clearly.**



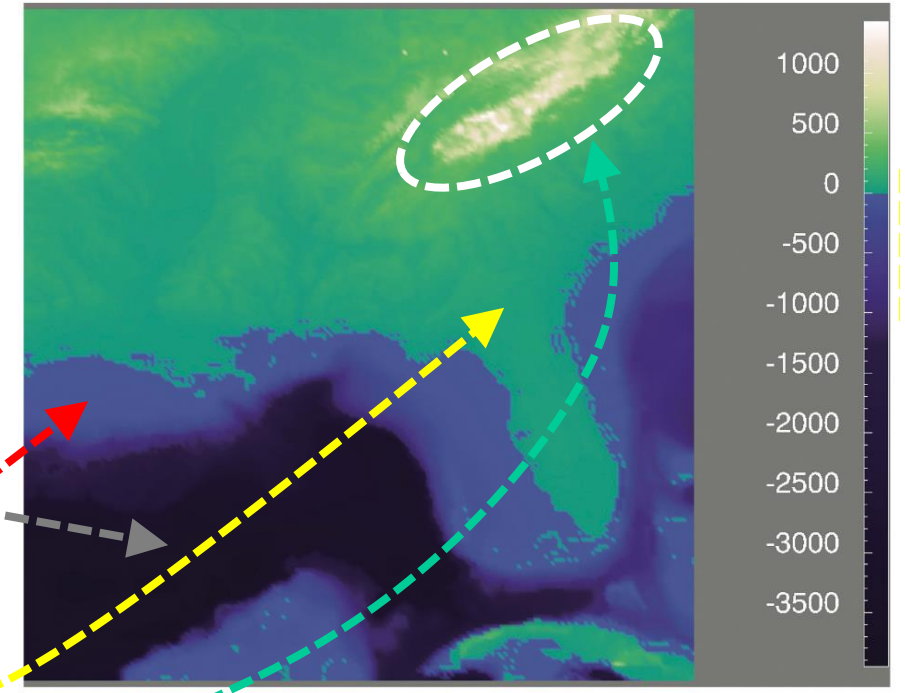
South-East of the US!

Ordered Colormaps: 11/13

- One way to address all three problems is to design **monotonically increasing luminance** colormaps.
- More precisely, the multiple hues being used are ordered according to their luminance from lowest to highest.
- The varying hues allow easy segmentation into categorical regions.
- Luminance is a magnitude channel, providing perceptual ordering.

Ordered Colormaps: 12/13

- **Hue is used to create a semantically meaningful categorization:**
  - **the 0 point matches with sea level,**
  - **the dark blue sea,**
  - **the cyan continental shelf,**
  - **the green lowlands,**
  - **the white mountains.**



South-East of the US!

Ordered Colormaps: 13/13

- **It is possible to create a perceptually linear rainbow colormap, but at the cost of losing part of the dynamic range because the fully saturated colors are not available for use.**

The standard rainbow colormap is perceptually nonlinear

Perceptually linear rainbow colormap is possible, but they are less bright with a decreased dynamic range

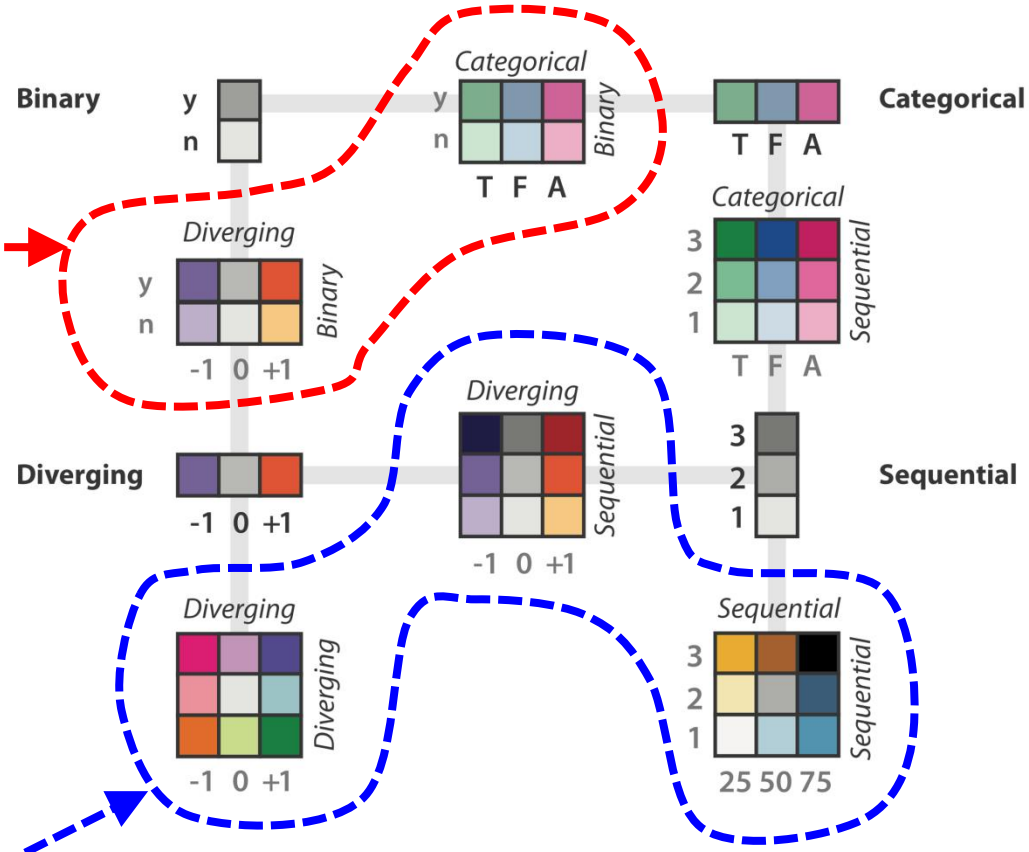
Segmented rainbows work well for categorical data when the number of categories is small

Bivariate Colormaps: 1/2

- **The safest use of the color channel is to visually encode a single attribute: these colormaps are known as univariate.**
- **Colormaps that encode two separate attributes are bivariate.**
- **When one of the two attributes is binary, meaning it has only two levels, then it is straightforward to create a comprehensive bivariate colormap with two families of colors.**

Bivariate Colormaps: 2/2

- Fixing a base set of hues and varying their saturation.
- When both attributes are categorical with multiple levels, results will be poor.
- Combinations of two **sequential** or **diverging** attributes with multiple levels is a middle ground.



Colorblind-Safe Colormaps: 1/2

- **Color blindness** is a sex-linked inherited trait that affects 8% of males and 0.5% females.
- In the common forms of color blindness the ability to sense along the red-green color axis is limited or absent.
- The problem is not limited to simply telling red apart from green; many pairs that are discriminable to people with normal color vision are confused, including red-black, blue-purple, light green-white, and brown-green.

Colorblind-Safe Colormaps: 2/2

- **Avoid the use of only the hue channel to encode information: in addition to hue design, categorical colormaps that vary in luminance or saturation.**
- **If possible, avoid colormaps that emphasize red-green, especially in divergent red-green ramps.**
- **Check http://www.rehue.net/, http://www.color-blindness.com/ and http://www.etre.com/tools/colourblindsimulator/ for color blindness simulators.**

The End