This section will focus on the use of color in map layouts and for representing data.

Module Topics

1. Color perception
2. Color models
3. Choropleth maps and normalization
4. Color ramp selection
5. Classification method selection
6. Selecting colors
7. Designing legends

2

These are the topics covered in this module.

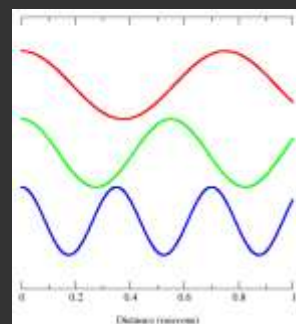
Light and Vision

3

Visible Spectrum

- A thin sliver of the electromagnetic spectrum that our eyes can sense
- Lots of information is available outside the visible spectrum, and we can build sensors to measure it

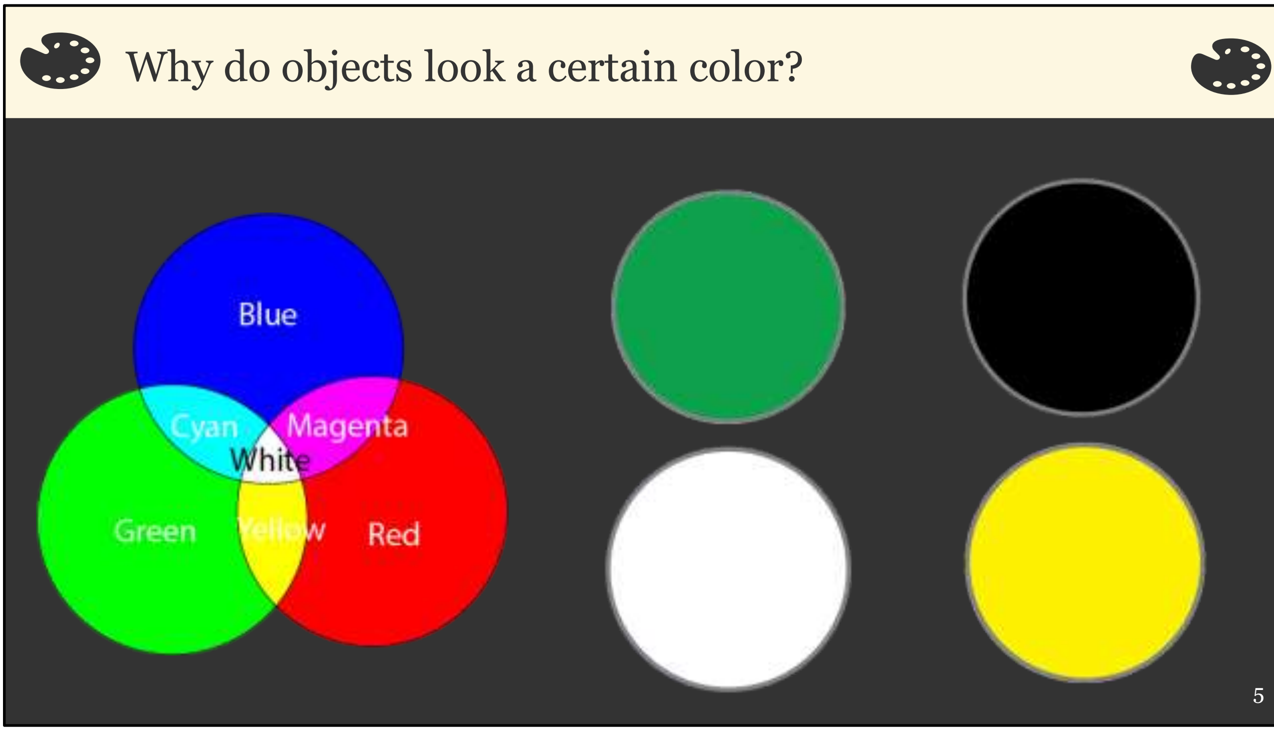
https://www.wnycstudios.org/story/211178-rip-rainbow/



https://commons.wikimedia.org/wiki/File:Visible_EM_modes.png

4

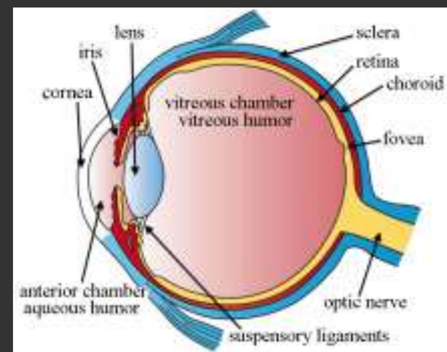
If you are interested in the physics of electromagnetic (EM) radiation and human vision, I would suggest checking out the RadioLab podcast linked here.



Objects take on different colors due to selective absorption of visible wavelengths. For example, a green object, like a leaf, looks green because it is absorbing the red and blue wavelengths while reflecting back the green. A yellow object would absorb the blue wavelengths but reflect back the green and red (red + green = yellow). White indicates that all wavelengths are reflected back (red + green + blue = white) while black indicates that all wavelengths are absorbed. Since our focus here is human vision, as opposed to printing, we will focus on additive color space as opposed to subtractive space. Additive space is used when explaining human vision and how images are generated by projected light sources, such as computer monitors, TVs, and projectors.

Human Vision

- ❖ Lens focuses light on retina
- ❖ Images are focused by changing the shape of the lens
- ❖ Fovea = portion of the retina where our visual acuity is the greatest
- ❖ Rods = black-and-white vision in low light
- ❖ Cones = color vision
- ❖ α, β, γ cones have peak sensitivity at 440 nm, 530 nm, and 570 nm *(blue, green, red)*



https://commons.wikimedia.org/wiki/File:Three_Main_Layers_of_the_Eye.png

6

Color vision is supported by our cones, of which we have more than 7 million. Some cones are sensitive to blue, while others are sensitive to green or red. See the peak sensitivities listed on the slide.

Our rods are more abundant and sensitive to light. They allow for low-light, monochromatic, or black-and-white, vision.

The existence of both rods and cones allows us to see in a variety of lighting conditions. In low light, colors cannot be differentiated.

To focus images, the shape of the lens is altered as opposed to changing the distance between the lens and focal plane, as is generally done with cameras.

Human Vision

- Bipolar and ganglion cells merge the input arriving from the rods and cones
- Each single ganglion cell corresponds to a group of rods or cones, a receptive field
- Optic nerve carries information from the retina to the brain and creates the blind spot

Within our eye, bipolar and ganglion cells merge the information received by groups of rods and cones. This merged information corresponds to a receptive field. This information is then transmitted to the brain along the optic nerves.

The brain then further processes this information.

Beyond the Eye

- Information from each eye travels separately until it reaches the optical chiasm
- Past the optical chiasm, pathways contain information from both eyes
- Information is then passed to the lateral geniculate nucleus (LGN)
- Interpretation of the visual information begins in the primary visual cortex

8

This slide lists some of the processing steps required after sensory information is collected by the eyes then transmitted to the brain along the optic nerves.

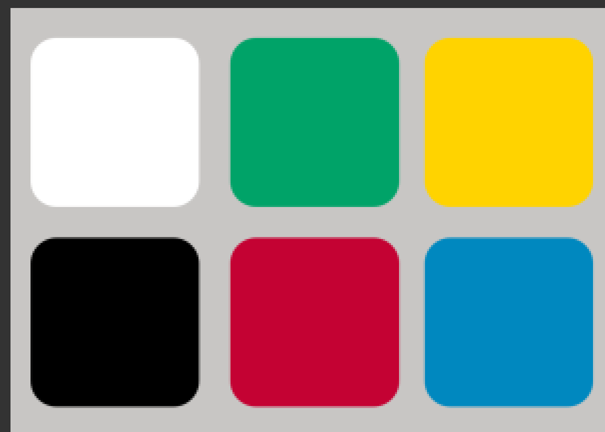
Theories of Color Perception

Trichromatic Theory

- Color perception is a function of the relative ratios of stimulation of different kinds of cones.

Opponent-Process Theory

- Based on a lightness-darkness channel and two opponent color channels: red-green and blue-yellow
- Colors within each opponent color channel are presumed to work in opposition to one another, meaning that we do not perceive mixtures of red and green or blue and yellow; rather, we see mixtures of pairs from each channel (red-blue, red-yellow, green-blue, green-yellow)



https://commons.wikimedia.org/wiki/File:Color_opponent_process_theory_contrast_of_complementary_colors.svg

There are two theories of human vision that each explain different aspects of color perception.

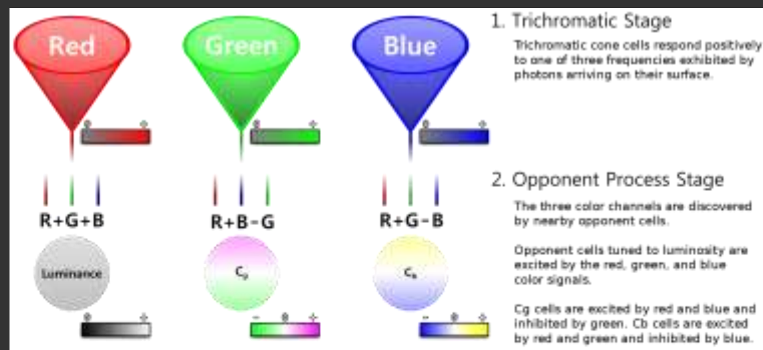
Trichromatic theory relates to the additive color space. Signals from cones that are sensitive to different wavelengths, red, green, or blue, are combined to yield a perception of color. With digital images, we mimic this process by storing true color images as separate red, green, and blue channels. Digital, true color sensors effectively mimic trichromatic human vision.

However, this theory does not explain all visual phenomenon. For example, if you stare at the images on the right side of the slide for an extended period, when you look away you should see an after image of the opposite color. Trichromatic theory cannot explain this. The Opponent Process Theory offers an alternative explanation. This theory suggests that color vision is created by processing pairs of opposite colors: white-and-black, red-and-green, and blue-and-yellow. After images are caused by an over stimulation of a certain color and a rebound effect that causes its opponent color to occur as an after image.

Which theory is correct? It turns out that both theories are necessary to explain human vision. So, both theories still play a role. Our understanding of human vision is incomplete.

Evidence for Opponent-Process Theory

- Opposite color afterimages
- Hue cancellation



https://commons.wikimedia.org/wiki/File:Modern_Color_Vision_Model.svg

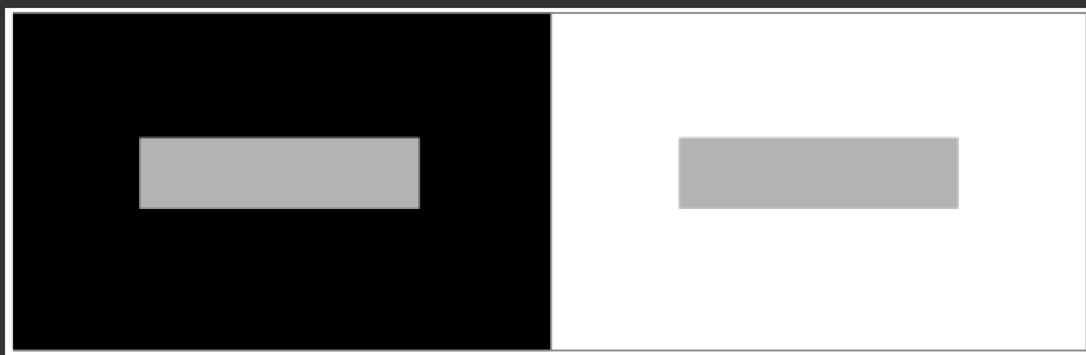
10

Evidence for the Opponent Processing Theory include opposite color afterimages and hue cancellation.

Opposite color afterimages are a phenomenon that consists of a color being perceived following overstimulation with its opponent.

Hue cancellation involves stimulation with a color (e.g. yellow) followed by an attempt to determine how much of the opponent color (e.g. blue) of one of the starting color's components must be added to eliminate any hint of that component from the starting color.

Simulated Contrast



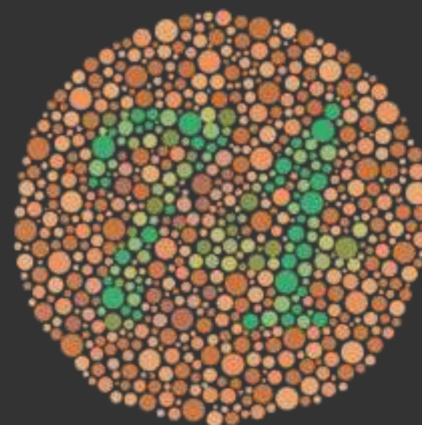
- The central gray strips are physically identical, but the one surrounded by black appears lighter

11

There are some additional visual phenomenon that are important to consider in map design. The gray rectangles on this slide are the same color. However, the one embedded within the larger, black rectangle appears lighter than the one surrounded by white. This phenomenon is known as simulated contrast. It is important for map design as this indicates that our perceptions of colors can be impacted by the surrounding colors. This is something that should be taken into consideration when designing map layouts and selecting color palettes.

Color Visual Impairments

- Anomalous Trichromats = some difficulty in distinguishing between red and green
- Dichromats = cannot distinguish between red and green



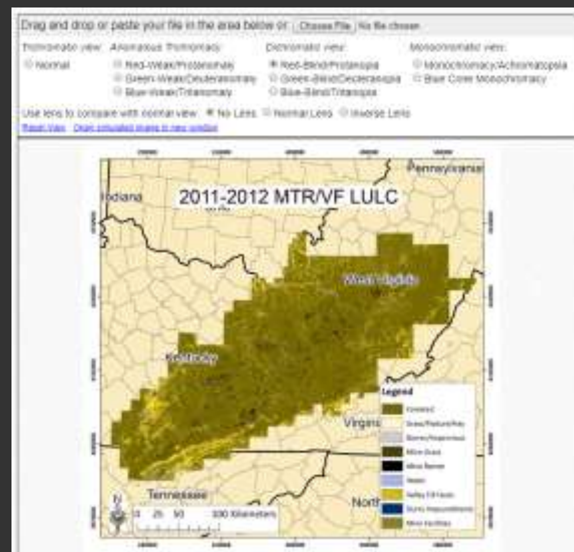
https://en.wikipedia.org/wiki/Color_blindness

12

Another factor to consider is how color perception is impacted by color vision impairments. Anomalous Trichromats have difficulties distinguishing between red and green while Dichromats cannot distinguish between these colors. There are more specific classifications and a wide variety of unique color vision impairments. Tools are available to select colors that are distinguishable, even if the viewer has a color vision impairment. For example, the ColorBrewer2 website allows you to filter out palettes that are not color blind accessible.

Color Visual Impairments

http://www.color-blindness.com/coblis-color-blindness-simulator/



13

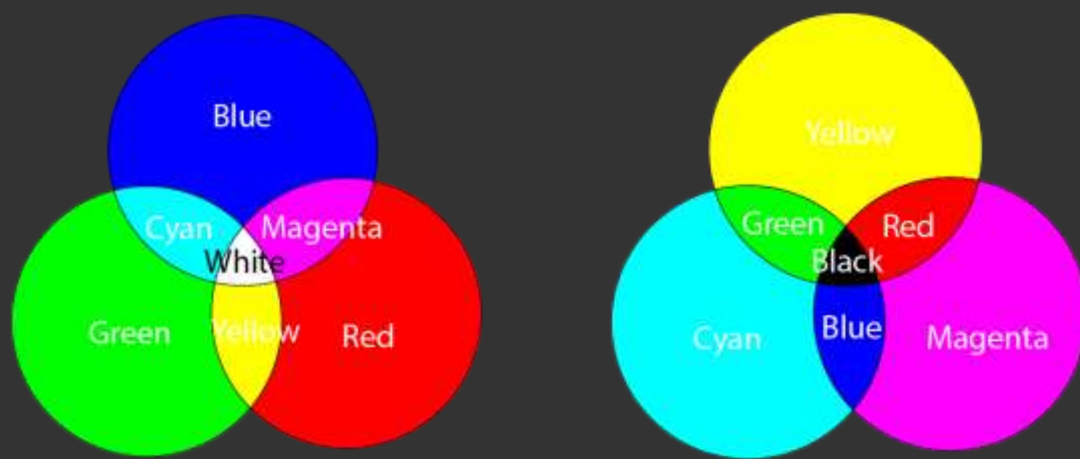
The link on this page allows you to simulate the perception of individuals with different color vision impairments. As you can see, there are a wide variety of classifications. This can be a valuable tool for assessing your map's accessibility.

In ArcGIS Pro, a Color Vision Simulator tool is available under the View Tab.

Color Models

14

Color Space



Additive Colors = used with computer monitors

Subtractive Colors = used in printing

15

Again, in this module we will focus on additive color space as opposed to subtractive space since we are interested in color perception.

The additive primaries (red, green, and blue) can be combined in different combinations to produce unique colors. This color space is generally used when dealing with projected light, such as reflected sunlight, computer monitors, TVs, and digital projectors. Subtractive color space is commonly used for printing. There is some value in understanding this color space, since it impacts how film is exposed and processed and how images are printed. However, we will not use it much in this class. The important concept here is that the subtractive primaries (Yellow, Cyan, and Magenta) sum to black and are used when dealing with dyes and printing. The dyes are effectively subtracting from the white paper or canvas. In contrast, the additive primaries sum to white.

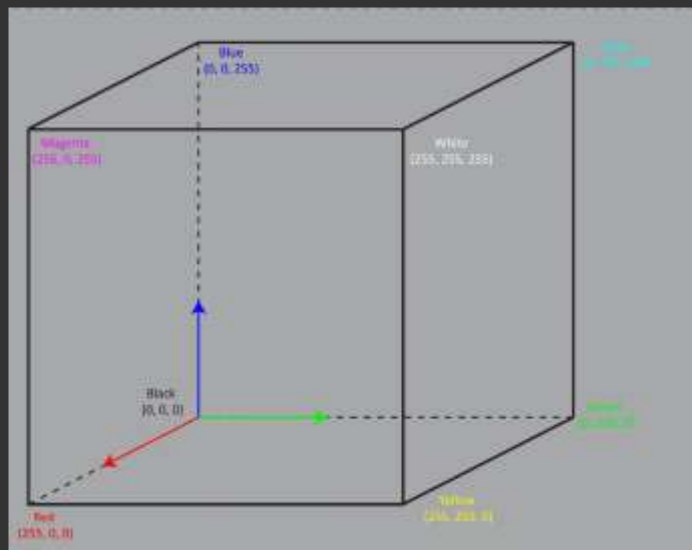
Color Wheel

https://commons.wikimedia.org/wiki/File:BYR_color_wheel.svg

16

The color wheel represents all the base colors or hues that can be perceived. How can we quantitatively define colors? In the next set of slides, we will discuss methods for defining colors.

RGB



17

Additive colors are generally defined based on relative proportions of the additive primaries: red, green, and blue. This space can be visualized as a cube. The corners of this cube represent the following:

Black = No Red + No Green + No Blue
White = Full Red + Full Green + Full Blue
Red = Full Red + No Green + No Blue
Green = No Red + Full Green + No Blue
Blue = No Red + No Green + Full Blue
Cyan = No Red + Full Green + Full Blue
Yellow = Full Red + Full Green + No Blue
Magenta = Full Red + No Green + Full Blue

Any other color would occur within the volume of the cube and could be represented as different proportions of the additive primaries red, green, and blue.

RGB

Example from Adobe Illustrator

Example from QGIS

Example from ArcGIS Pro

18

It is common to define colors using 8-bit color space where 256 gray or brightness levels are available for each primary color. By mixing relative proportions of these 0 to 255 values, many colors can be defined (256 X 256 X 256 = more than 16 million!).

Hex Codes

Integer	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Hexadecimal	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F

#ff0000

(15 * 16) + 15 = 255

(0 * 16) + 0 = 0

(0 * 16) + 0 = 0

(255, 0, 0) = Red

#ffff00

(15 * 16) + 15 = 255

(15 * 16) + 15 = 255

(0 * 16) + 0 = 0

(250, 255, 0) = Yellow

#B6AAC2

(11 * 16) + 6 = 182

(10 * 16) + 10 = 170

(12 * 16) + 2 = 194

(182, 170, 194)

Example from ArcGIS Pro

19

Instead of using RGB values as 0 to 255, colors can also be defined using hex codes.

As the slide describes, 0 through 9 indicate 0 through 9 while A through F indicate 10 through 15.

The first two characters define the red value, the middle two define the green value, and the last two define the blue value.

In each pair, the first value in multiplied by 16 and then the second value is added to get a 0 to 255 value for each primary color.

So, this offers a means to define an 8-bit RGB color using six characters.

Note that when you write hex codes you commonly start with a pound sign or hashtag.

Let’s go through the examples on the slide.

The first example represents red since the first two values yield 255 and the remaining pairs yield 0 and 0. So, full red, no blue or green.

The middle example represents yellow: full red plus full green yields yellow.

The last example is a mix of red, green, and blue and yields the color in which the resulting RGB values are displayed.

Note that many software packages and websites allow you to find hex codes from RGB values or RGB values from hex codes including ArcGIS Pro, Adobe Photoshop, and Adobe Illustrator.

RGBA

- R = Red (0-255)
- G = Green (0-255)
- B = Blue (0-255)
- A = Alpha (0-1 or 0 to 100%)

- A controls transparency/opacity

Example from ArcGIS Pro

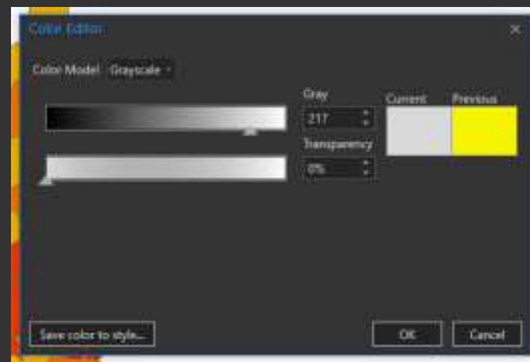
20

RGBA includes an alpha channel along with the RGB channels to control transparency.

The RGB values range from 0 to 255 while alpha varies from 0 to 1 or 0% to 100%, where 0 or 0% indicates fully transparent while 1 or 100% yields fully opaque.

Grayscale

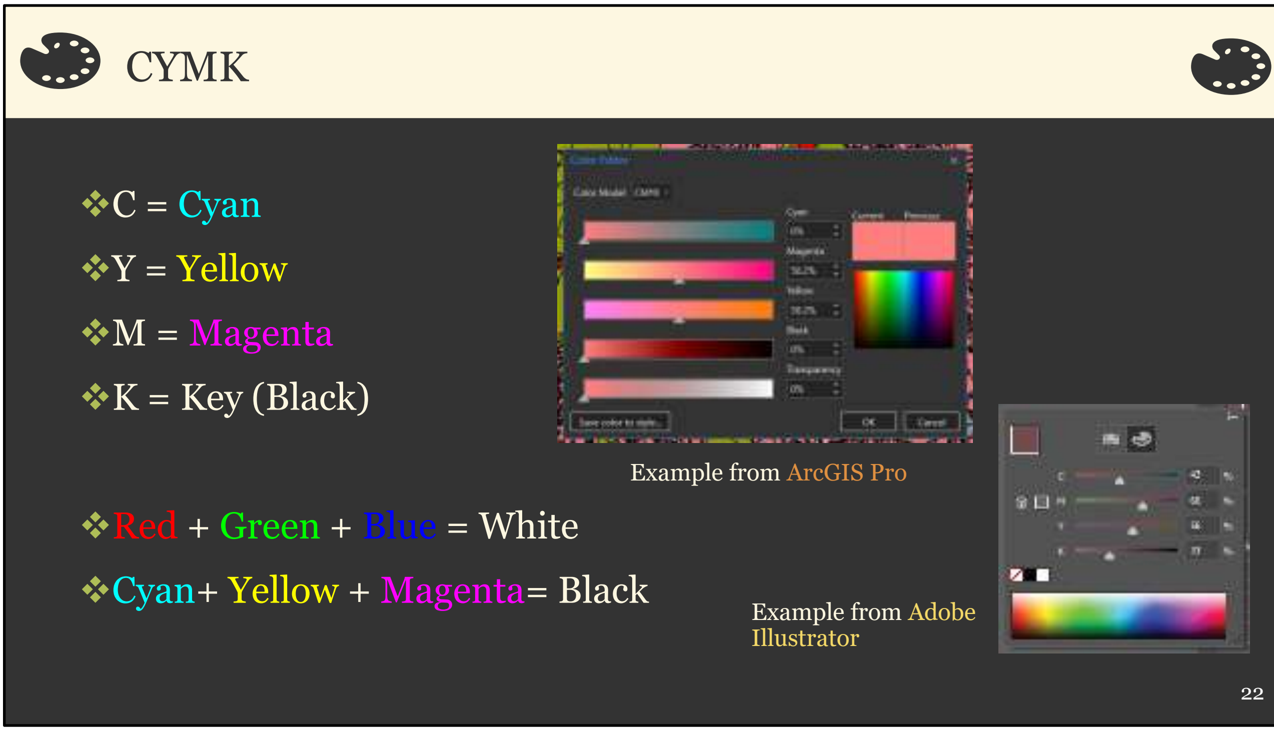
- For black-and-white production
- 0 = black
- 255 = White



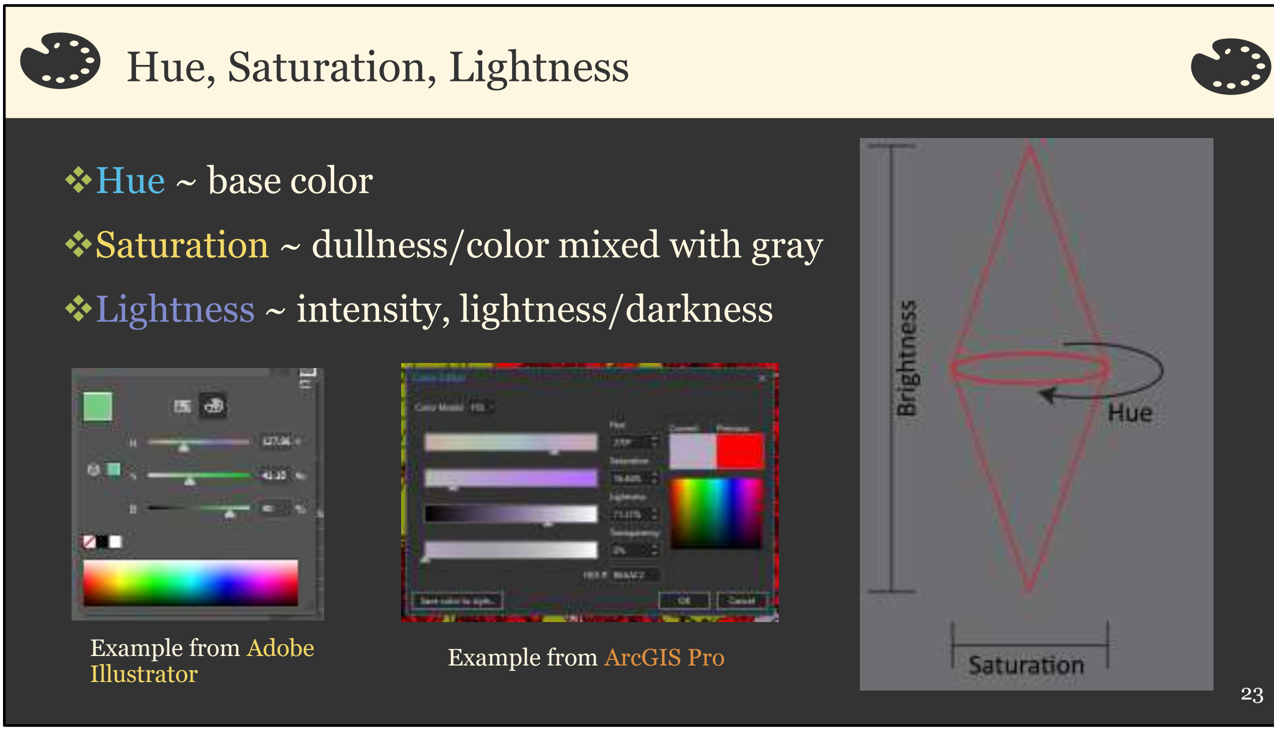
Example from ArcGIS Pro

21

Grayscale color space only allows for black-and-white renderings. 0 indicates black while 255 indicates white using an 8-bit space. Different gray levels would occur in between.



When using subtractive colors, such as for printing or when designing graphics that will be printed, colors are commonly defined using the subtractive primaries as percentages. This will consist of a percentage of cyan, yellow, magenta, and the key color (black). Sometimes, you can also define transparency or alpha, which is also generally expressed as a percentage.



An alternative means to represent color is using the hue, saturation, lightness, or HSL, space.

In this space, a color is defined using a base color (hue), a saturation (mixing the base color with gray), and a lightness (full brightness to full dark or white to black).

This space is often visualized as two cones that share a common base. The hue is represented as rotating around the cone, the saturation corresponds to moving toward or away from the center of the cone, and the brightness corresponds to moving up and down the cone.

Hue is commonly defined as an angle between 0 to 360 while saturation and lightness are scaled from 0 to 1 or 0% to 100%. It is also possible to include an alpha or transparency argument (HSLA).

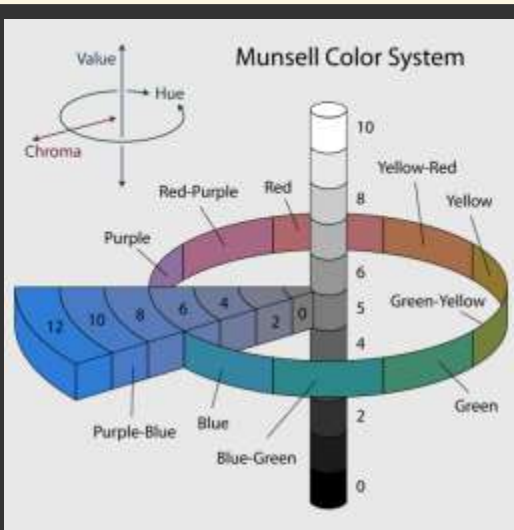
Disadvantage of RGB, CYMK, and HSB

- Equal steps in color model do not correspond to equal visual steps
- Eye discerns changes in red much more readily than changes in blue

24

There are some disadvantages with all of the color models that we have discussed to this point. Equal steps or change within the color models do not correspond to equal visual steps: our eyes can discern some changes more than others. Specifically, we can discern changes in red hues much more readily than changes in blue. Color models have been developed that take these differences in perception into account.

Munsell Model



https://commons.wikimedia.org/wiki/File:Munsell-system.svg

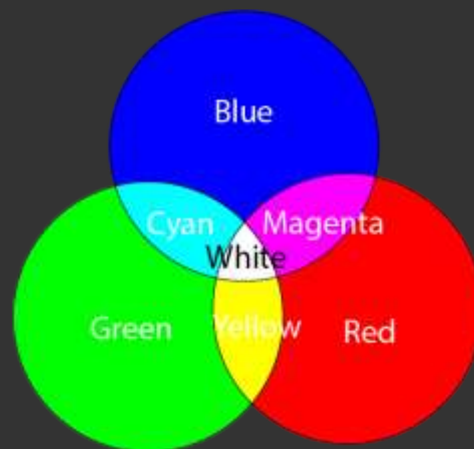
- Based on human perception of color
- Hue = color
- Value = perceived lightness
- Chroma = similar to saturation/gray tones of same hue

25

One example color model that attempts to approximate human color perception is the Munsell model. If you have ever collected or described soil samples, you may be familiar with this model, as the Munsell Soil Color Chart uses this system. This model has a similar structure to Hue, Saturation, Lightness. Hue is the same as hue and represents the base color. Value and chroma are similar to saturation and lightness, respectively; however, human color perception is accounted for.

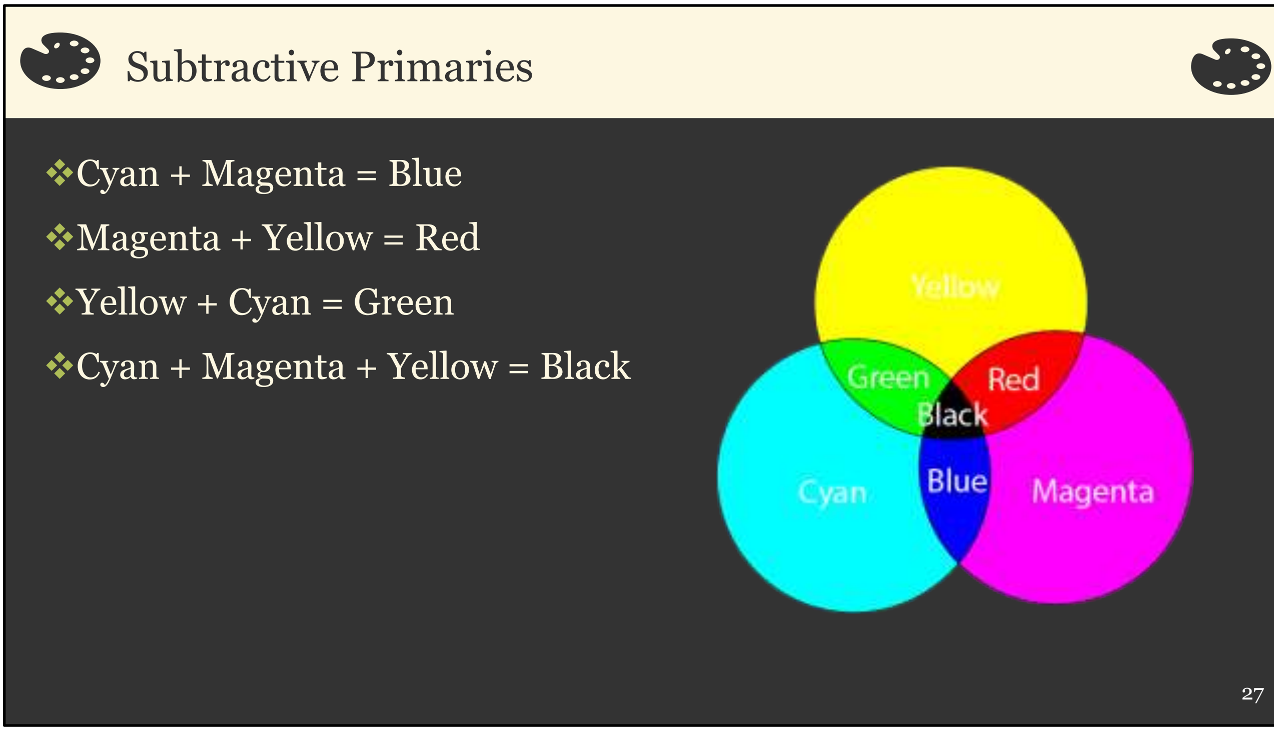
Additive Primaries

- Red + Blue = Magenta
- Red + Green = Yellow
- Blue + Green = Cyan
- Red + Blue + Green = White



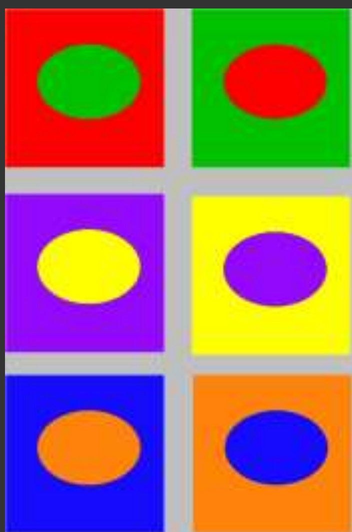
26

When discussing projected colors, such as human color perception or how colors are generated by a computer monitor, you can think of a specific mixture of the additive primary colors.



In contrast, for printing, colors are generated using the subtractive primaries.

Complimentary Colors



- Colors on opposite side of color wheel
- Make each other appear brighter

https://commons.wikimedia.org/wiki/File:Contrast_of_complementary_colors.jpg

28

In the next set of slides, we will discuss choosing colors to use in your map space. When designing maps, figures, websites, or other graphics, it is generally recommended to use a small set of colors to define your palette. One option is to make use of complimentary colors that are directly opposite from one another on the color wheel and tend to make each other appear brighter.

Selecting Color Palettes

- Choose subset of colors to use in your artwork/map

- Analogous = adjacent colors on color wheel
- Monochromatic = same base hue
- Complimentary = use complimentary colors
- Triad = three colors equally spaced around color wheel complimentary colors
- Split Complimentary = color and colors symmetrical to it on the color wheel
- Square = four colors equally spaced around color wheel complimentary colors
- Shades = same base hue with black added

https://color.adobe.com/create/color-wheel

29

To select more than two colors, different methods are available that tend to yield a set of colors that are visually pleasing. Take some time to read through these examples. The link on this page is for an Adobe website that can assist in the selection of a color palette that is visually pleasing using these different methods.

Selecting Color Palettes

https://color.adobe.com/create/color-wheel
https://coolors.co/
https://www.canva.com/colors/color-palette-generator/
http://colormind.io/
https://htmlcolorcodes.com/resources/best-color-palette-generators/

30

This slide provides some additional links that are useful for selecting color palettes.

Using Colors

31

Applications

Color ramps of a single hue can vary by:

1. Lightness/Brightness/Intensity
2. Saturation
3. Both

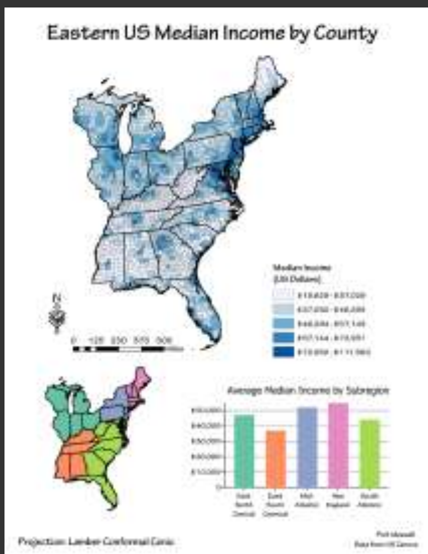
Color ramps of multiple hues can span:

1. The entire color wheel
2. A portion of the color wheel

32

We will now explore the use of color for representing our data using choropleth thematic maps. Different color sets will be used for different purposes. When using a single base hue, you can select multiple colors by changing the lightness, saturation, or both. If you plan to use multiple hues, the selected colors can span the entire color wheel or just a portion. Different color sets will have different valid use cases.

Choropleth Maps

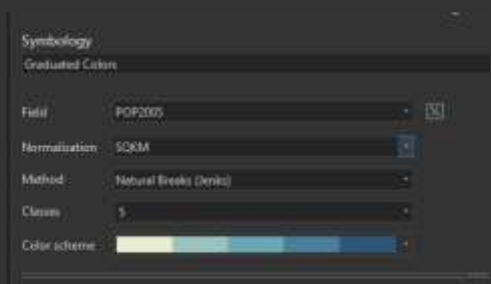


Can be used to display categorical or normalized continuous data

33

Remember that choropleth maps can be used to show categorical variables and normalized continuous variables. For example, median income, a continuous variable, is shown in this example map using different shades of blue while regions of the eastern United States are shown using unordered colors. Median income is normalized in this case since it represents the median, a measure of central tendency, calculated for individuals living in the county.

Normalization

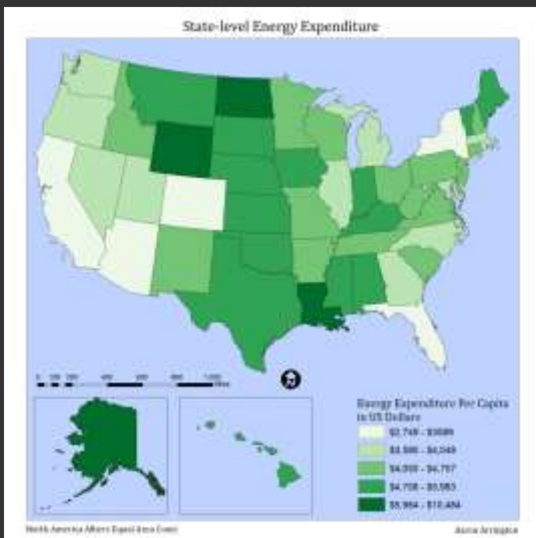


- Divide value by another value to normalize the measurements
- For example: population vs. population per square kilometers
- Provides for a fairer comparison
- Should normalize when producing choropleth maps

34

Again, color should only be used to visualize continuous variables that have been normalized. It is considered inappropriate to use color to visualize continuous data that have not been normalized. Examples of normalized data include median income, mean income, population density, all percentages and proportions, cases per 1000,000 people, etc. Examples of variables that are not normalized include total area, number of people, number of cases, etc. ArcGIS Pro allows for normalization to be applied relative to another provided attribute. If your data cannot be normalized, it would be best to use a different visualization, like a proportional or graduated symbol method.

Normalization



Example of normalized continuous data

Map by Aaron Arrington

This is an example of normalized data represented using a choropleth symbology. Energy use is normalized by reporting energy use per person (i.e., per capita).

Normalization (from Slocum et al.)

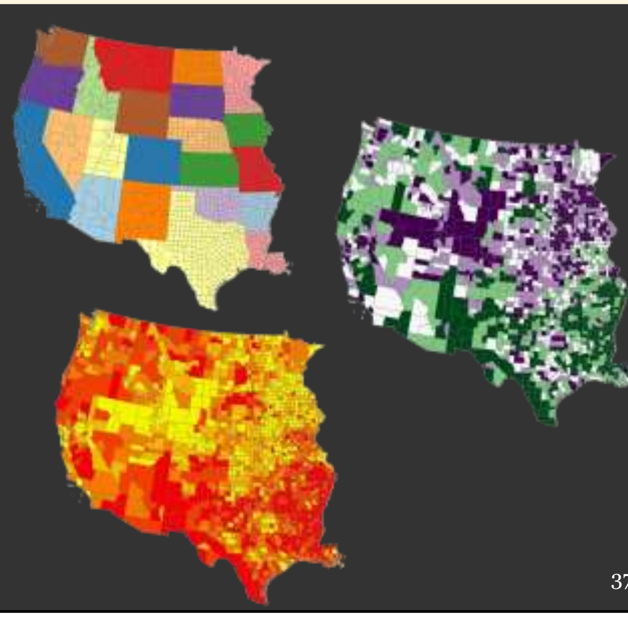
1. Choropleth technique is most appropriate for a phenomenon that is uniformly distributed within each mapping unit
2. Works best when there is little variation in size and shape of the map units
3. Largest features have the largest visual impact, but also have potentially the largest error of representation
4. Should only use normalized data
5. Can normalize by area to obtain a proportion or percentage
6. Can normalize not using area (for example, normalize by population)

36

Why is it that you should only used choropleth maps to visualize continuous data? Slocum et al. does a great job of explaining this, and I have summarized their explanation here. The main issue here is that larger map features have the most visual impact. However, the largest features tend to have the most uncertainty due to spatial aggregation. So, you are giving the largest visual weight to the map features with the most uncertainty. Normalization allows for a reduction in this disparate impact.

Color Ramp

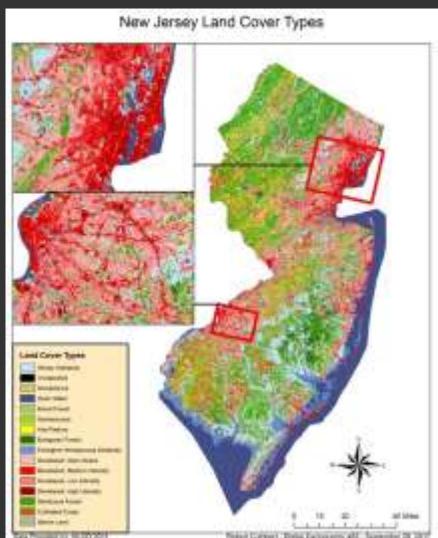
- You can change the colors used to symbolize your data
- You can also change the number of categories
- Nominal Color Scheme = used for categorical data (not ordered)
- Sequential Color Scheme = used for continuous data that increase in one direction
- Diverging Color Scheme = used for continuous data that diverge



37

Generally, nominal, unorder colors are used to represent categorical data. In contrast, sequential or diverging color schemes are used to represent quantitative or continuous data. Sequential schemes make the most sense when the values increase in one direction. Diverging color schemes make the most sense when values diverge from a central value, such as population density increase or decrease over time.

Choropleth Maps



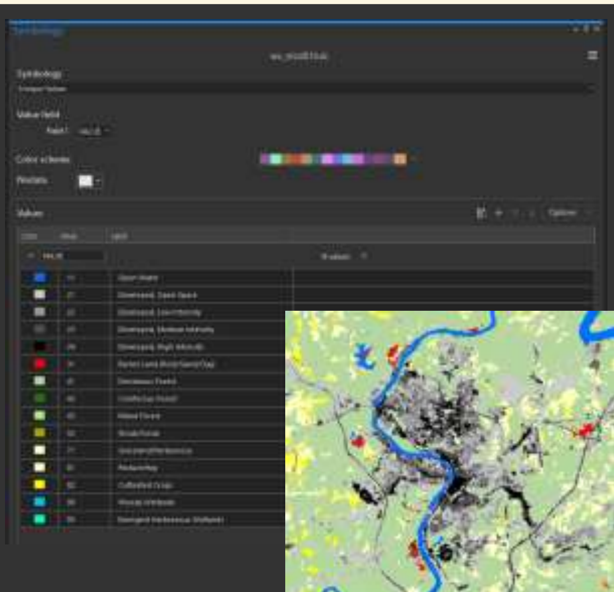
- Example for categorical data

Map by Robert Cuthbert

38

This is an example of maps that use color to show a categorical or qualitative variable with unordered colors. For some categorical data, certain colors may be associated with certain categories. For example, it makes sense to map water using a blue hue and forests using a green hue. Other categories may not have a color associated with them, such as rock types or soil units. In such cases, we generally choose colors that can be easily distinguished.

Symbolizing Categories



- Land cover type, private vs. public, urban vs. rural, soil types, etc.
- Use colors that make sense for the category (if appropriate)
- Colors should not be ordered

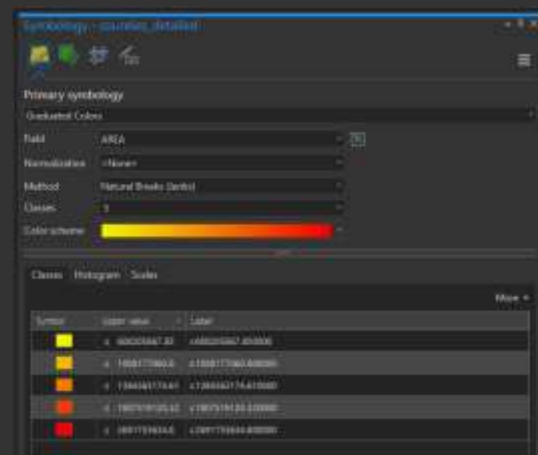
39

When choosing colors to represent categories, you may choose to use colors associated with the categories where appropriate. If no colors are associated, then we generally choose colors that are well differentiated, especially for neighboring features. For example, you wouldn’t want to have two rock units that are adjacent to each other represented using vary similar colors. Regardless of the colors used, they should not be ordered in anyway, as this implies an ordering to the data values.

Symbolizing Continuous Data

When classifying based on quantities you must choose:

1. The number of classes (5 – 7 is a good rule)
2. If normalization will be applied
3. Classification method
4. Colors used (we will discuss this in a later section)

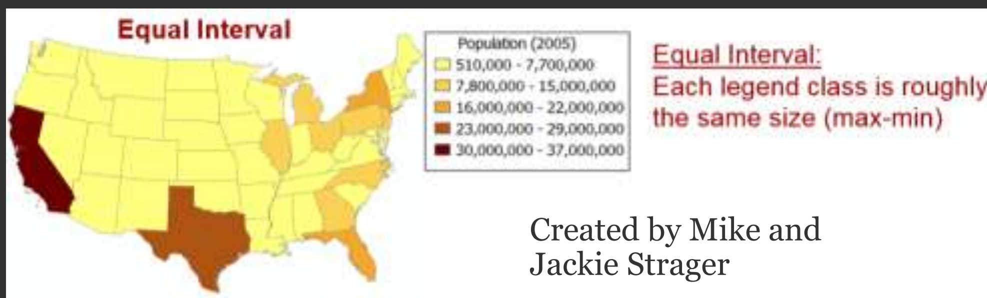


40

When symbolizing continuous data or quantities, a continuous ramp can be used or a finite number of colors. If a finite number of colors is used, then the data will have to be classified or binned in some way. There are a variety of methods for selecting class breaks, which will be discussed in the following slides. Also, when using bins, it is generally recommended to use 5 to 7 classes. This is so that differences are highlighted; however, the number of classes are limited so that the colors and meaning can be clearly distinguished.

Equal Interval

- Breaks data range into equal intervals
- Does not take into account natural breaks in the data



Created by Mike and Jackie Strager

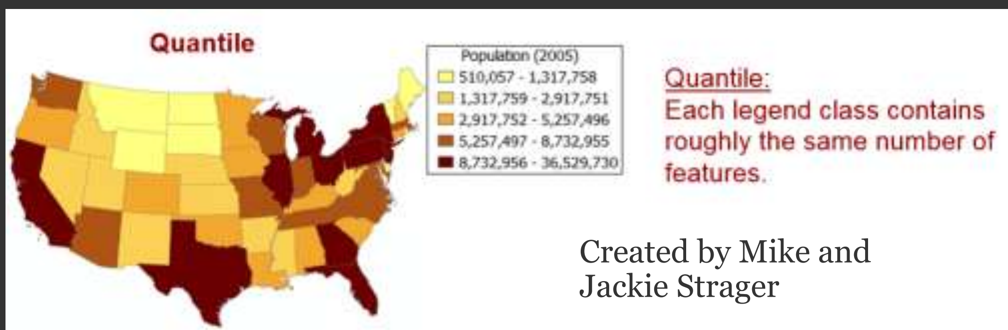
41

One common binning or classification method is equal interval. In this case, the same range of values occur in each bin.

For example, if you had a range of values from 0% to 100% and five bins, then the data would be divided into 0-20%, 20-40%, 40-60%, 60-80%, and 80-100%.

Quantiles

- Same number of features in each bin
- Does not take into account natural breaks in the data



Created by Mike and Jackie Strager

42

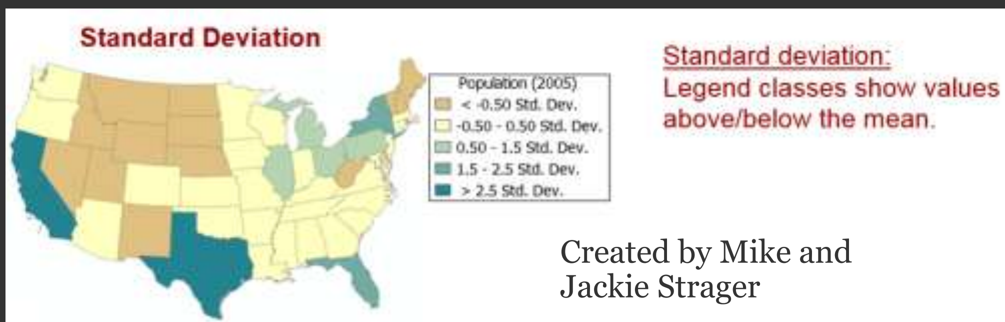
The quantiles classification method is commonly confused with equal interval. Using this method, an equal number of map features is placed in each bin as opposed to an equal range of values.

For example, if all 50 states were being symbolized using 5 classification bins, then 10 states would be placed in each bin.

This is generally the preferred classification method for ordinal data.

Mean-Standard Deviation

- Bins defined based on standard deviation of data subtracted or added to mean



Created by Mike and Jackie Strager

43

The mean-standard deviation method makes use of the central tendency and variability of the data.

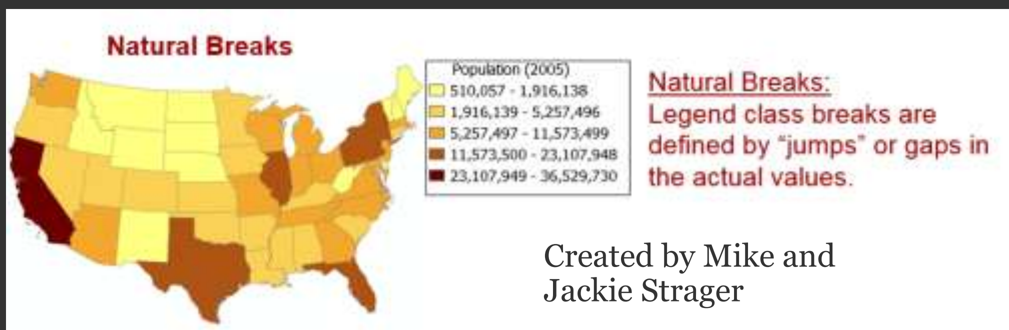
The data are symbolized based on their standard deviation from the mean. Values that are much lower than the mean will show a negative standard deviation, values close to the mean will have a standard deviation near zero, and values much higher than the mean will have a positive standard deviation.

This method can be useful if you want to compare how a map feature deviates from the average or mean conditions of all the map features.

This method generally should not be used if the data have a distribution that is far from normal, such as multimodal or skewed data.

Jenks-Caspell/Fisher-Jenks

- Algorithm that attempts to find "natural breaks" in the data
- ArcGIS uses an implementation of Fisher-Jenks

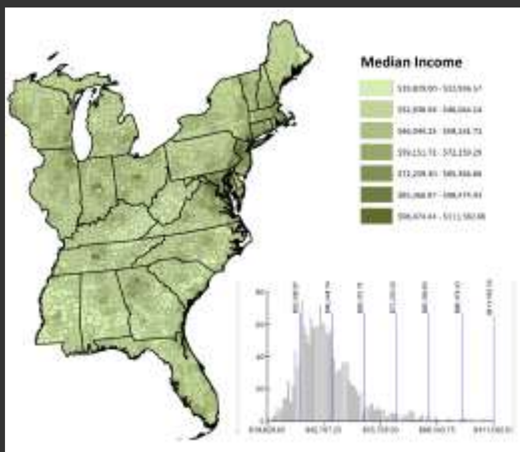


44

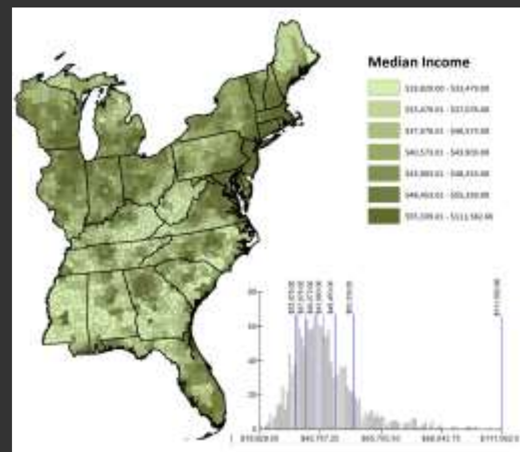
The natural breaks method makes use of an algorithm to find "natural breaks" in the data. Class breaks are then positioned relative to these natural data breaks. This is a data-driven classification scheme.

This method can be misleading if there are no breaks, jumps, or gaps in the data distribution. In such cases, it would be finding breaks that don't really exist or are just noise.

Classification Methods Compared



Equal Interval



Quantile

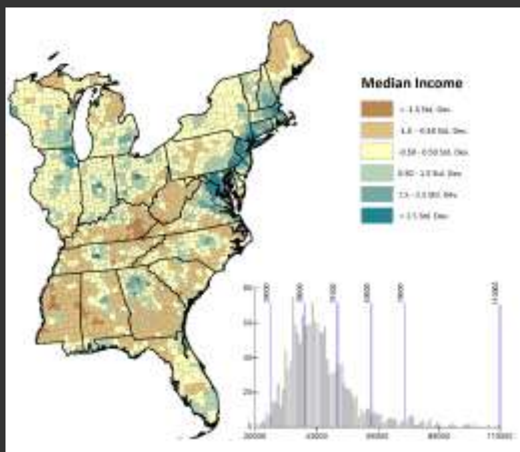
45

Here are some comparisons.

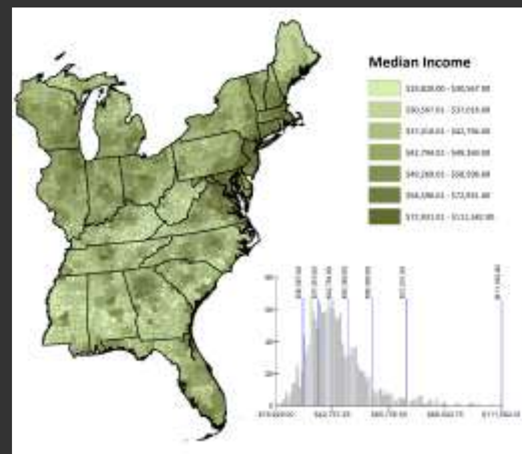
For equal interval, the class breaks, represented by the blue lines in the distribution, are equally spaced, or there is an equal range of values in each classification bin.

In contrast, the breaks are not evenly spaced for the quantile classification. Instead, a roughly equal number of features will be placed in each bin.

Classification Methods Compared



Mean-Standard Deviation



Natural Breaks

46

Using the mean-standard deviation classification, we can represent deviation from the mean median income by county. Red counties are lower than the mean while blue counties are higher than the mean.

The last example represents natural breaks in which an algorithm selects the breaks based on jumps or gaps in the data distribution.

There is not necessarily a best classification method, as this is case specific and there are pros and cons for each.

When deciding on a classification method, you should think about the nature of the data and how best to present it.

Maximum Distance

- Raw data are ordered from low to high
- Differences between adjacent values are computed
- Largest differences define breaks
- Tries to take into account the clustering of the data

47

Another, less commonly used classification method is maximum distance. The raw data values are ordered from low to high and differences between adjacent values are compared. The largest gaps are used to define the breaks. The major benefit of this method is that it takes into account clustering in the data.

Manual Classification

❖ Define your own breaks

48

It is also possible to define your own breaks or classification bins. When you do this, it is important to not manipulate the data to tell the story you want to tell. There is a whole book on this topic called "How to Lie with Maps."

If you decide to use manual breaks it is generally best if the breaks have some meaning. For example, if a program that awards grant money for economic development to counties makes its decision based on certain percent poverty thresholds, it would make since to symbolize using these thresholds to see which counties are candidates for a certain dollar amount from the program.

Method Comparison

Consideration	Equal Interval	Quantiles	Mean SD	Maximum Distance	Natural Breaks
Considers distribution of data along a number line	P	P	G (if normal)	G	VG
Ease of understanding concept	VG	VG	VG	VG	G
Ease of computation	VG	VG	VG	VG	VG
Ease of understanding legend	VG	P	G	P	P
Legend values match range of data in a class	P	VG	P	VG	VG
Acceptable for ordinal data	U	A	U	U	U
Assists in selecting number of classes	P	P	P	P	G

P = Poor
G = Good
VG = Very Good
A = Acceptable
U = Unacceptable

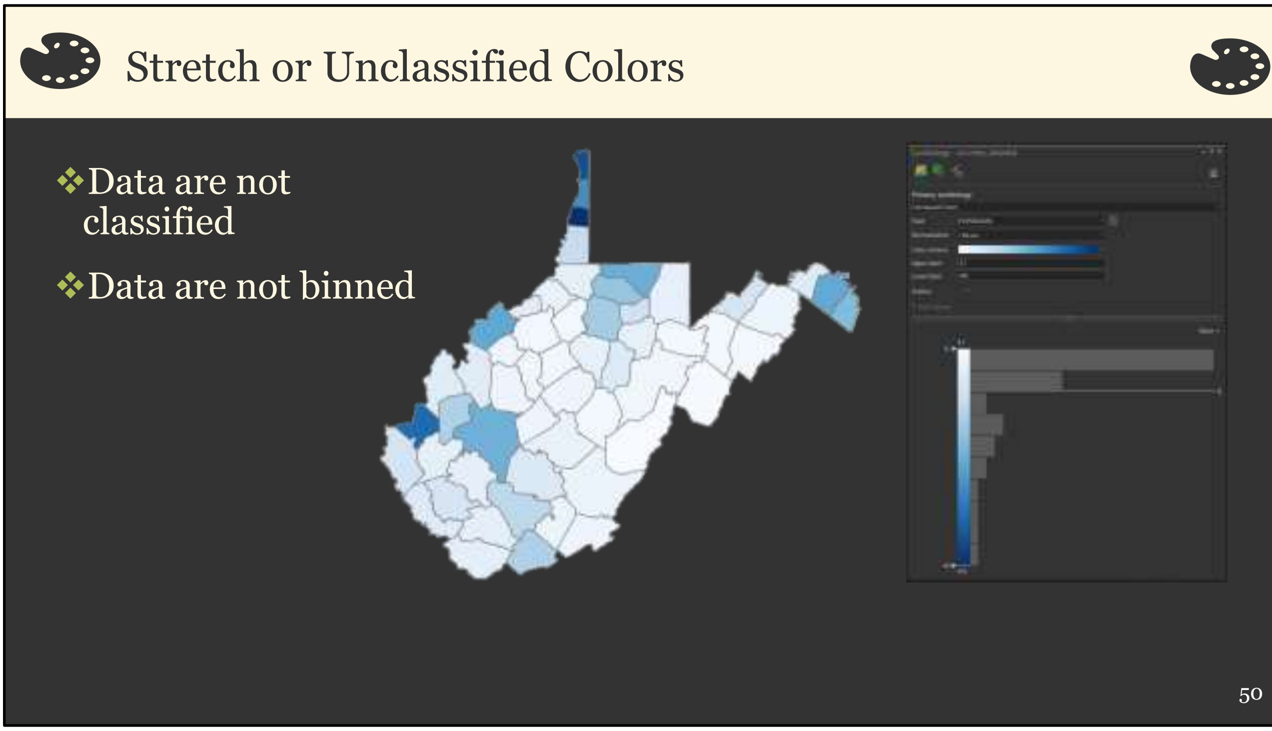
Slocum et al. Figure 4.5

49

This table from Slocum et al. compares the classification methods discussed above. Equal interval and quantiles do not consider the distribution of the data. Mean SD does; however, since it relies on the mean and standard deviation, it can be misleading if the data are not normally distributed. All the methods are generally easy to understand and easy to compute, especially with modern computers. The legend is most easily interpreted for equal interval and mean SD since equal interval uses the same range per bin and mean SD uses variance relative to the mean. The equal interval and mean SD methods have a poor match between the legend values and data range in the associated class since the breaks are defined mathematically and may not correspond well with the highest and lowest values in each class.

When visualizing ordinal data, only quantiles is an acceptable classification method as it relies only on the ranked order of the data. All other methods treat the data as numeric values to determine breaks based on data range, data central tendency and dispersion, and/or gaps in the distribution. These measures or considerations are not appropriate for ordinal data.

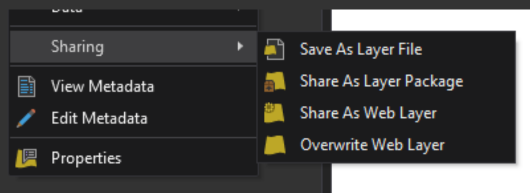
Lastly, only natural breaks assists the user in determining the number of classes to use since it attempts to locate breaks in the data distribution.



Instead of applying a classification to group the data and map the bins to a finite number of colors, you can also choose to simply use a color stretch. One drawback is that it may be difficult to correlate a map feature’s associated quantity with the legend color ramp.

Layer Files (ArcGIS)

❖Can be used to save layer symbology for later use



51

If you generate a legend and symbology for a layer, it can be saved in ArcGIS Pro as a layer file. This layer file can then be reapplied to the data or used on another dataset. Layer files are not copies of the data. Instead, they simply save the symbology for later use. I sometimes produce layer files to include with data layers so that the user can easily apply my preferred symbology.

Selecting Colors

52

Considerations

- Human perception of colors
- What you are mapping
- Categorical vs. numeric
- Nominal vs. ordinal
- Interval vs. ratio
- Sequential vs. divergent
- Color vision impairments
- Medium of presentation

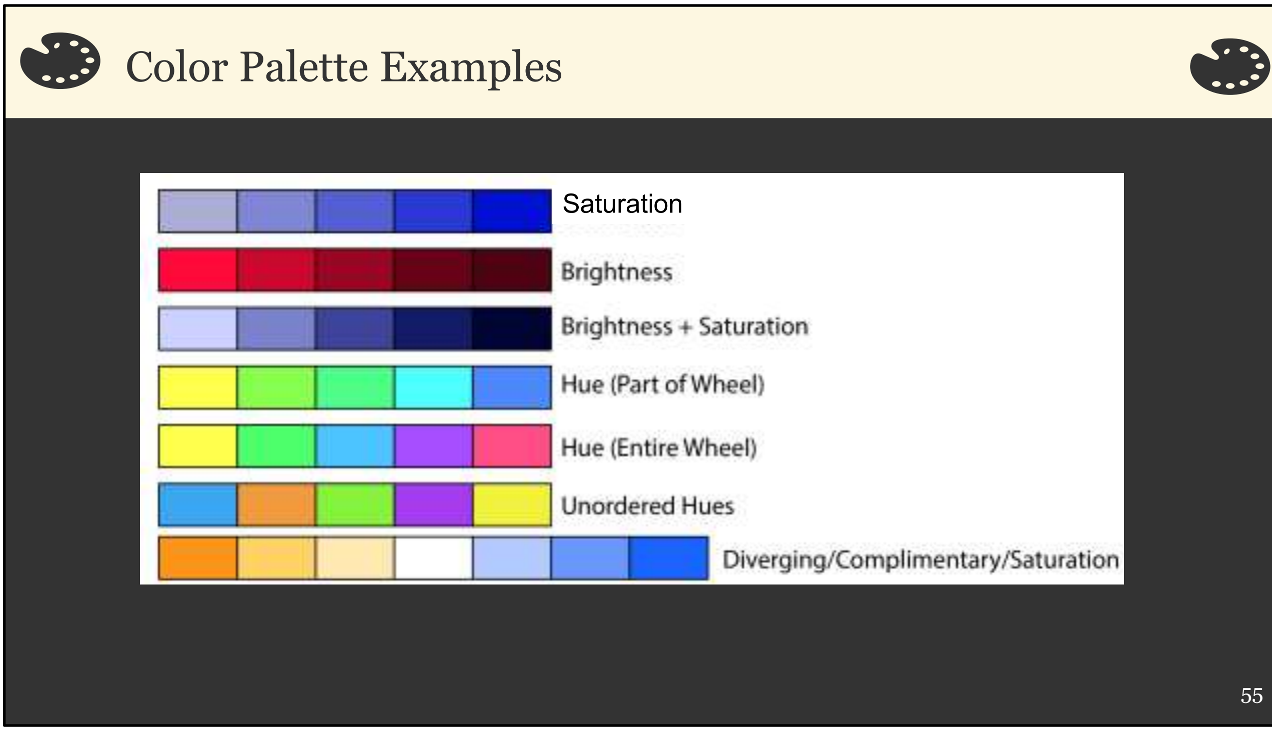
53

There are many considerations when selecting colors to map to data values or categories. Human perception should be taken into account. For example, in western societies, red tends to be associated with danger or alerts. So, red is generally not a good choice to represent a positive outcome. Green may be associated with nature or money. Blue tends to be associated with measures that are calm or cool. It can also be associated with water. You should also take into consideration color vision impairments and the medium of presentation. For example, if a map will be viewed from a greater distance, you may want to use colors that contrast well to ease interpretation. In the following slides, we will discuss issues associated with choosing colors in more detail.

Color selection tips from Slocum et al.

1. Make use of unordered colors for nominal data
2. Make use of differences in lightness to produce sequential schemes for continuous data
3. Use lighter colors for low values and darker for higher values
4. You can increase visual contrast by also altering saturation
5. For diverging color schemes, vary lightness for two different hues that diverge from a common light hue or neutral gray
6. Consider color vision impairment
7. Consider simultaneous contrast
8. Consider color association and cultural bias
9. Consider aesthetics
10. Arithmetic increases in a color gun do not correspond to an arithmetic increase in perceived lightness, so the shades will not appear equally spaced to an observer

This slide summarizes tips for color selection presented in Slocum et al. Take some time to read through these suggestions. Note that when visualizing continuous data, it is generally recommended to use lighter colors for lower values and darker colors for higher values. Also, remember that human perception of color differences vary by hue, so arithmetic increases in a color generated by a color gun do not correspond to an arithmetic increase in perceived lightness.



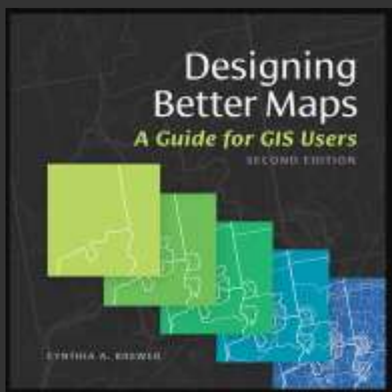
This slide shows some example color palettes. The first example uses a blue-purple base hue with adjustments to saturation (i.e., mixing with gray). In contrast, the second example uses a base red hue with changes to only brightness (i.e., lightness). The third example uses a combination of changes to brightness and saturation while maintaining the same base hue. All of these ramps would be appropriate for continuous, sequential data that increases in one direction.

The next two examples make use of changes to the base hue for a part of and the entire color wheel, respectively. Here, an order is implied, so this would not be appropriate for categorical data. Sometimes, ordered hues can be useful for continuous data; however, the order should be easily discerned.

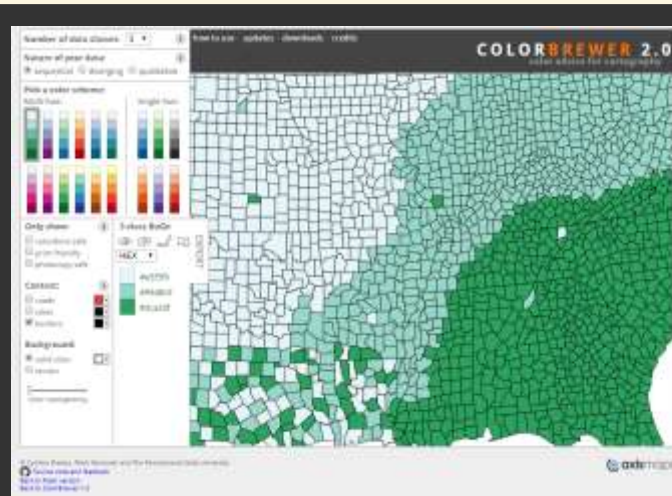
The second to last example represents unordered hues, which would be appropriate for categorical or qualitative data.

The final set is an example of a diverging color scheme. This palette specifically makes use of complimentary colors of base orange and blue hues. The middle color is a neutral white. On either side of the center, the orange or blue hues increase in saturation at the same rate. This provides a well contrasted, easy to interpret visualization for diverging data.

ColorBrewer 2.0



https://esripress.esri.com/display/index.cfm?fuseaction=display&websiteID=293&moduleID=0



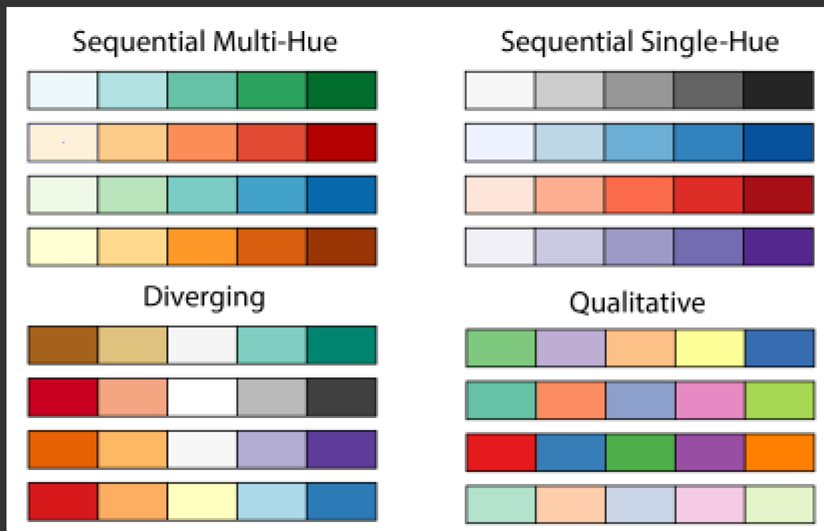
http://colorbrewer2.org/#type=sequential&scheme=BuGn&n=3

56

The ColorBrewer 2.0 website is a great resource for selecting color palettes. A variety of options are available for sequential, diverging, and qualitative color schemes. You can also select the number of colors desired and filter based on visual color impairment, print friendly, and/or photocopy safe options. Once you have decided on a color ramp, you can obtain the associated HEX codes, RGB values, or CYMK values.

Cynthia Brewer's text *Designing Better Maps* provides a detailed discussion of the use of color for visualizing data.

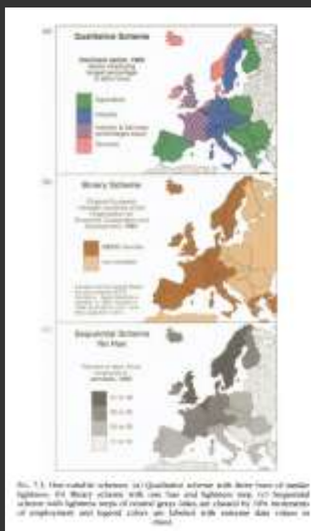
Color Palette Examples

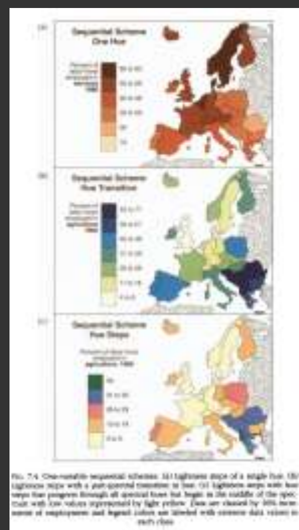


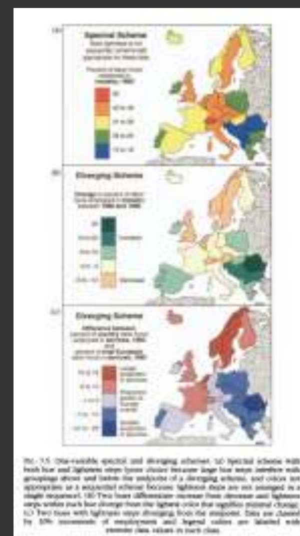
57

This slide provides some example color palettes from the ColorBrewer 2.0 website. The upper-left group represents sequential color schemes that use multiple hues. Note that even though multiple hues are used, the order is clear. The upper-right group are sequential color schemes that use a single hue with changes in only lightness and/or saturation. The lower-left group offer examples of diverging color schemes while the lower-right group shows unordered ramps appropriate for qualitative or categorical data.

Brewer, C.A., 1994. Color use guidelines for mapping. *Visualization in modern cartography*, *1994*, pp.123-148.

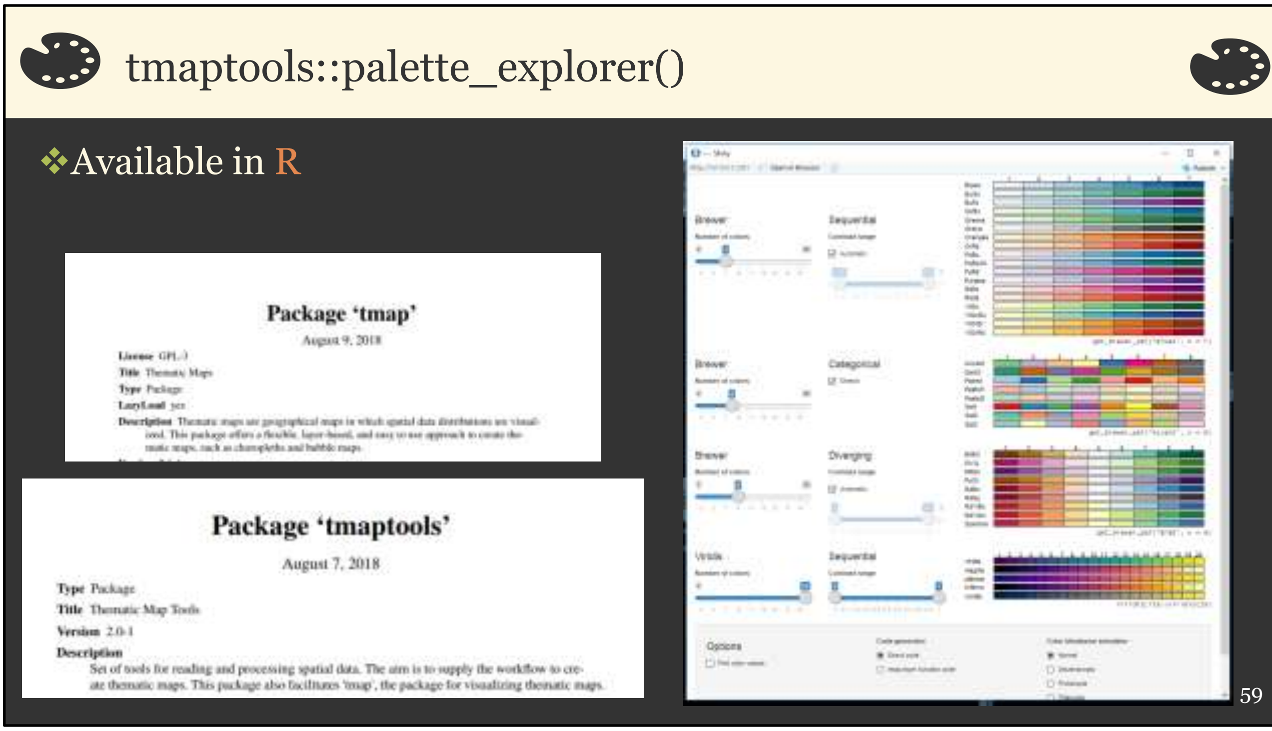






58

For a detailed discussion of the use of color for thematic visualization, I would recommend reviewing this 1994 publication by Cynthia Brewer. The figures on this slide are from this publication.



As a side note, the ColorBrewer 2.0 palettes have been integrated into other tools, such as the open-source data science software tool R.

Choropleth Legends

60

Choropleth legend tips from Slocum et al.

1. Vertical legends: high values at top, low at bottom
2. Horizontal legends: values should increase from left to right
3. Numeric values to right of vertical legend color boxes
4. Numeric values beneath color boxes for horizontal legend
5. Color boxes should not have gaps between them
6. Can include graphs along with the legend
7. Consider the desired and appropriate precision of the measurements (how many decimal places)

We will now turn our attention to choropleth map legend design. These tips are summarized from Slocum et al. Take the time to review these suggestions.

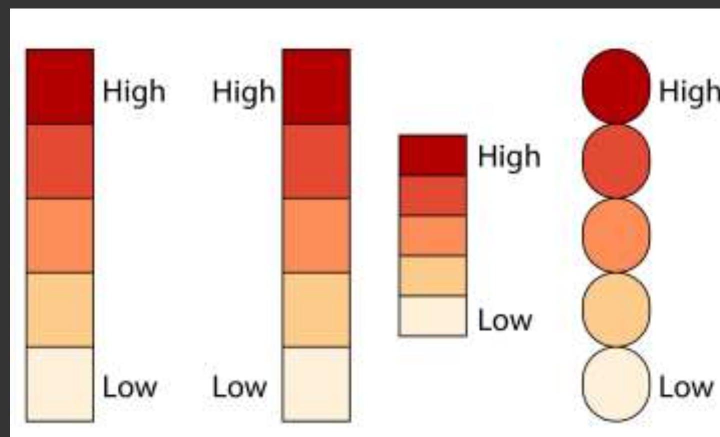
If a legend is designed vertically, high values should be at the top and low values at the bottom. Horizontal legends should increase from left to right.

For vertical legends, labels should be placed to the right of the color boxes or symbols. For horizontal legends, the numeric values should be placed beneath the color boxes or symbols.

The individual legend components showing the colors used should not have gaps between them. However, sometimes it may be appropriate to include graphs with the legend to further explain the data distribution. Lastly, the precision or number of significant digits shown should be intentionally chosen.

Legends

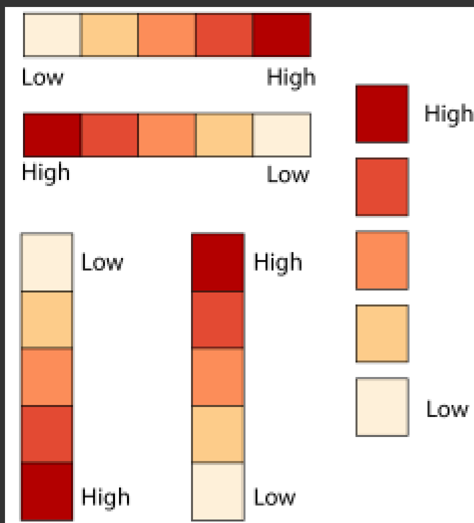
- Placement of text labels
- Size/aspect ratio of symbols
- Shape of symbols



62

Again, text should be placed to the right of vertical legends. You can also make adjustment to the size and aspect ratio of the boxes or use a non-rectangular symbol.

Legends

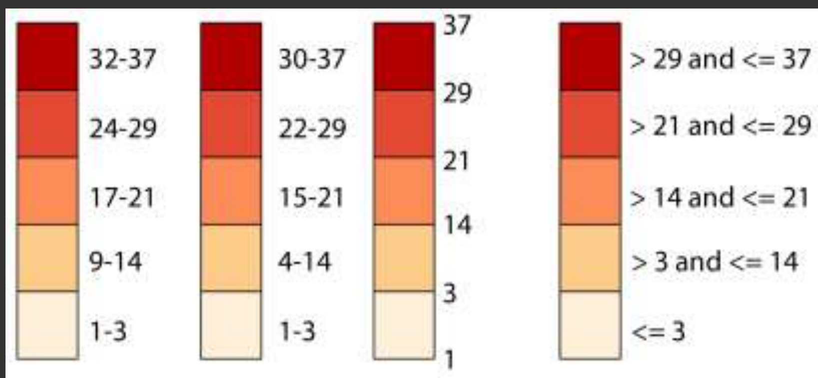


- Low-to-High vs. High-to-Low
- Vertical vs. Horizontal
- Separated vs. Not Separated

63

For horizontal legends, the values should increase from right to left. For vertical legends, values should increase from bottom to top. Labels should be to the right of a vertical legend or beneath a horizontal legend. Generally, it is not advised to separate the individual symbols making up the legend.

Legends



- Number of decimal places
- Gaps or no gaps
- Class breaks
- Ranges

64

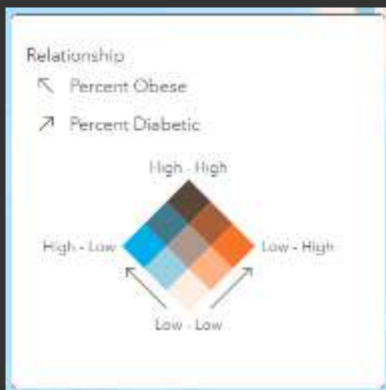
Another consideration is how best to define the labels. In the first example on the left, the range of values is defined for each bin with data gaps where no values are present. In the second example, the ranges are expanded so that no data gaps are included. I generally recommend not including gaps, since this may be confusing to the reader. However, this can be misleading if there are large gaps in the distribution.

In the third example, the upper and lower values are labelled for each break. This can be effective and is easy to interpret.

Lastly, to be more explicit, you can include less than or equal to, greater than or equal to, and greater than signs to make the class breaks very precise.

Again, there is not necessarily a correct labeling method. Instead, this will depend on the data and intent of the map. However, you should always consider how easily legends can be interpreted by the viewer and intended audience.

Bivariate Choropleth Maps



Example from ArcGIS Online using County Health Rankings 2018

65

It is possible to map two continuous variables using a bivariate choropleth symbology. When this method is used, legends are commonly provided as a diamond representing the different color combinations and how they relate to the two variables being mapped. This will also require that the axes are clearly defined.

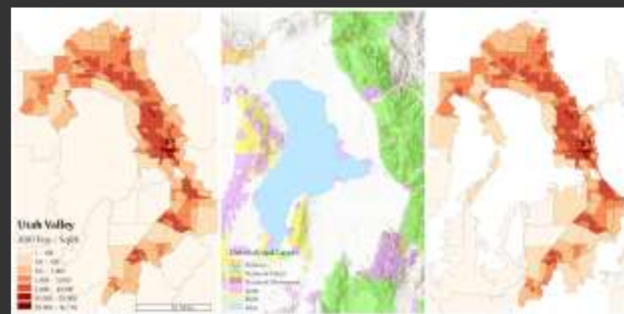
In the provided example, percent obese is symbolized with the blue ramp while percent diabetic is symbolized with the orange ramp. The color at the top of the diamond corresponds to both high percent obese and high percent diabetic. The color at the bottom of the diamond corresponds to low for each while the left and right positions are high for one variable and low for the other.

Dasymetric Mapping

66

Dasymetric Maps

- Uses color
- Makes use of ancillary data to redistribute raw totals
- Colors do not have to align with mapping units/aggregating units.
- Used for standardized or normalized data



https://commons.wikimedia.org/wiki/File:Utah_Valley_dasymetric_map.png

67

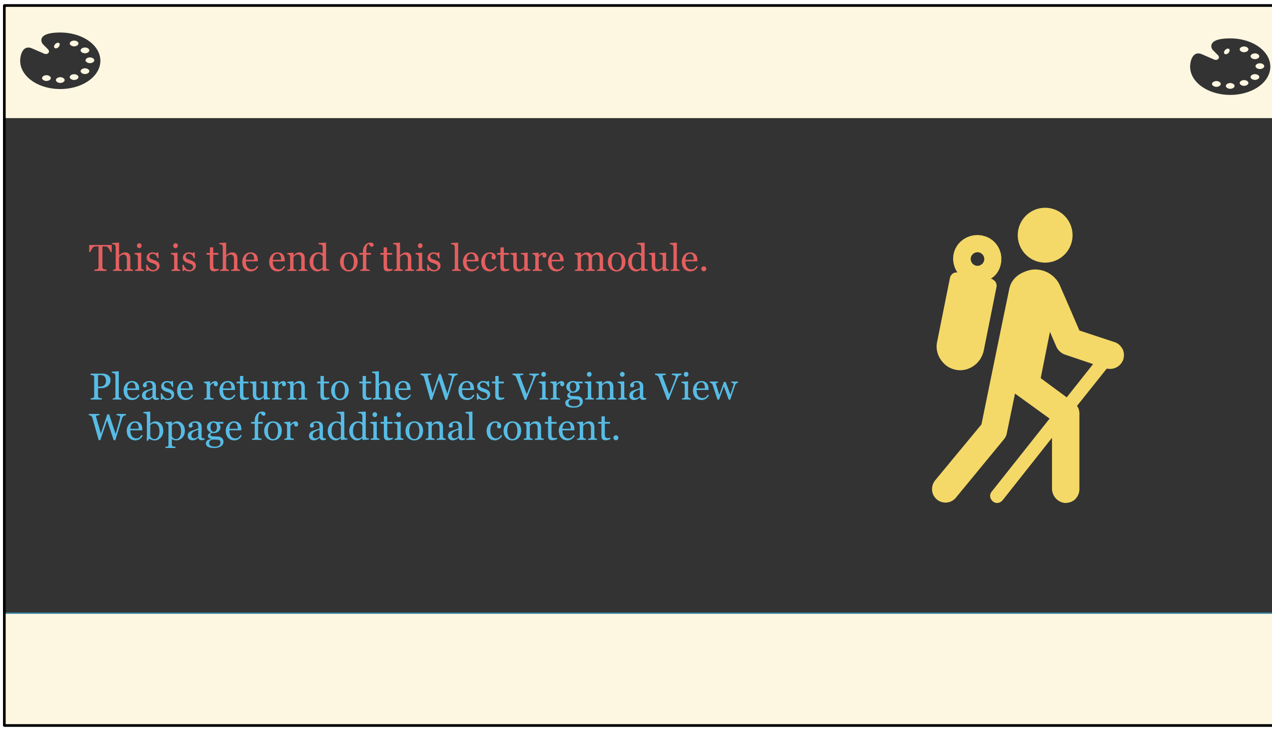
We will not discuss dasymetric mapping in detail in this course, but I did want to mention it. Similar to choropleth maps, dasymetric maps use color. However, in this case ancillary data are used to redistribute the raw totals so that they do not have to align to the aggregating units. Similar to choropleth maps, these are only appropriate for normalized data.

Ancillary Data

- Land use/land cover maps
- Night light imagery
- Remotely sensed imagery
- Additional boundaries
- Census data
- Etc.

68

The ancillary data used will depend on the data being mapped. For example, a land cover dataset could be used to redistribute percent forest or percent cropland estimates. Night light imagery could be used to redistribute population density data. Remotely sensed data and derived products have been shown to be useful for generating these types of maps.



Thanks! Hope you found this useful.