2021 – ArcGIS Pro

Symbology

Warthog Information Services

Created by Stefan Freelan

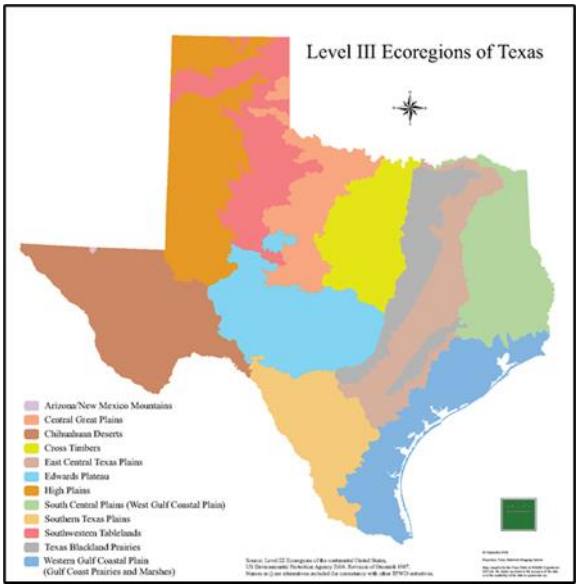
Horizontal units are in **meters** (UTM, 10-N, NAD-83).

WARTHOG
INFORMATION SERVICES

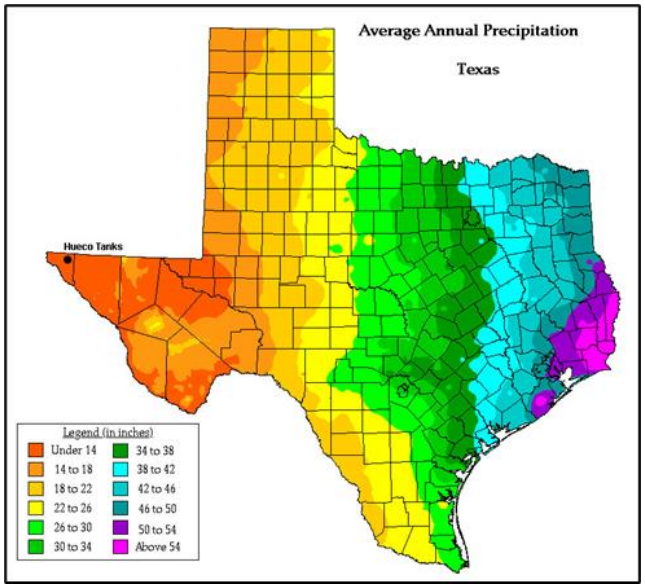
Exercise Introduction

A key element of a GIS is the ability to display features in different ways based upon their attributes. Symbology is often described as being either *qualitative* or *quantitative*:

- Qualitative symbology is based on some attribute quality (type of road, land use category, presence/absence of fish in streams, etc.).
- Quantitative symbology is based on numerical data (traffic volume, population, crop yield, etc.).



Ecoregions (qualitative)
Source: Texas Parks and Wildlife



Precipitation (quantitative)
Source: Texas Beyond History (texasbeyondhistory.net)

Different types of features have different aspects that can be modified to create a custom symbology:

- Points can be symbolized by changing the color, size, shape and/or rotation of the symbol – points can be abstract shapes (circles, triangles, etc.) or more representational (trees, airplanes, etc.).
- Lines can be differentiated by changing the color, thickness, or pattern (dashed vs. solid).
- Polygons can be symbolized by changing the color and/or pattern of the polygon's fill (solid vs. hatch vs. speckle) as well as by altering the polygon's outline.

Exercise Learning Objectives

- Qualitative vs Quantitative symbology for points, lines, and polygons
- Classification methods
- Density and normalizing data

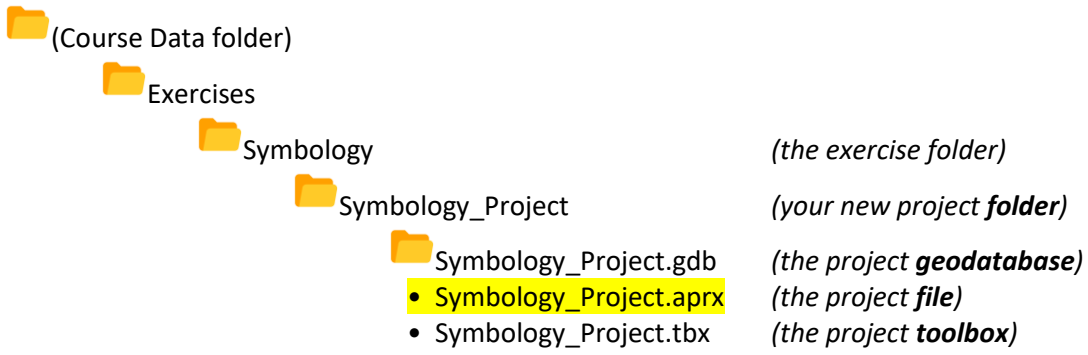
Step 1: Open the Exercise Project

Overview

For this exercise, a project with data and a map has already been created for you. It is stored in the Selections folder in the Exercises folder of your Course Data folder.

Instructions

- In Windows File Explorer, navigate to the **Exercises** folder in your Course Data folder.
- Open the **Symbology** folder.
- Open the **Symbology_Project** folder (a project folder).
- Locate the **Symbology_Project.aprx** ArcGIS Project file and double-click it to open the project.



The project should open with the **Symbology Map** view activated. If this map is not active, open it from the **Maps** folder in the Catalog pane. Depending on your software, you may have various panes open and/or available from previous work. You may close all panes but the Catalog pane and map Contents pane.

Step 2: Qualitative Symbology: Single Symbols

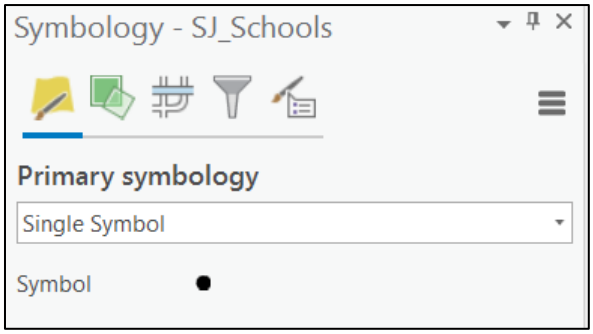
Overview

The most basic type of symbology (and the default when a new data layer is added to a map view) is a single symbol type for all of the features in the data layer. Technically, this is a qualitative symbology (based on the "quality" of being a feature). Having all features symbolized the same way is appropriate for many types of features, such as parks or schools. The choice of how this single symbol is displayed is controlled via the symbology properties.

Instructions

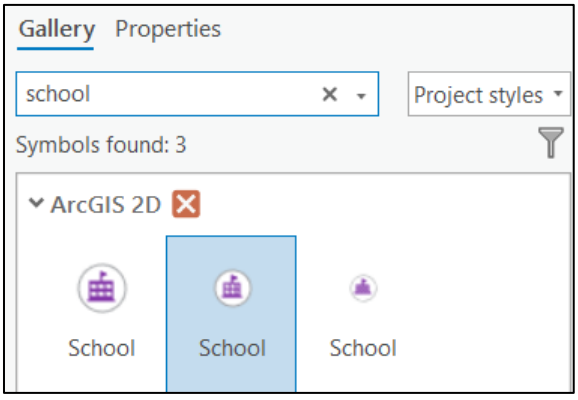
- In the map Content pane, turn On the **San_Juans** and **SJ_Schools** layers (if they are not already On).
- Right-click on the **San_Juans** layer and click *Zoom to Layer*.
- Right-click the **SJ_Schools** layer in the Contents pane and click *Symbology*. This opens the Symbology pane.

  In the Symbology pane, the primary type of symbology is listed as *Single Symbol*.



  We will explore symbolizing using the gallery options.

- In the **SJ_Schools** Symbology pane, click the black circle symbol icon to open the *Gallery*.
- In the search box, type “school” (no quotes) and press Enter.
- Select the medium-sized school icon.

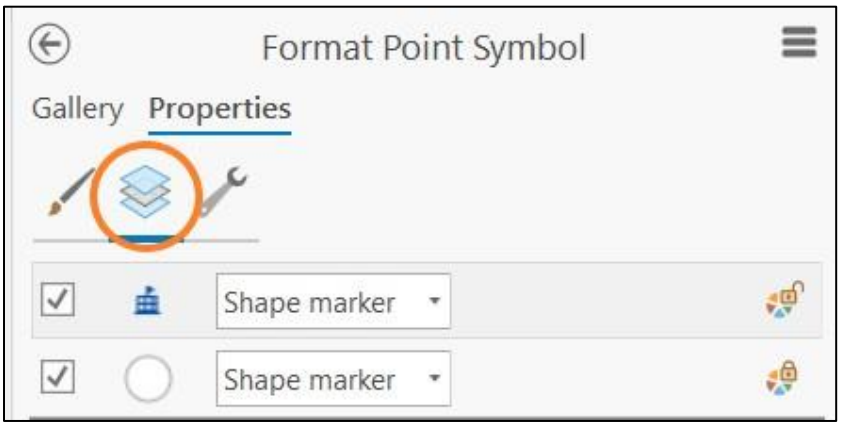


Next, we will format the appearance of this symbol.

- Click the *Properties* tab at the top of the Symbology pane.
- If it is not already expanded, expand the *Appearance* section.
- Change the color to a shade of dark blue.
- Click *Apply* at the bottom of the Symbology pane to see the results.

- Next, click the *Layers* sub-tab in the Symbology pane (the three layers icon).

The school symbol is actually made up of two separate symbols: One is a building icon (now dark blue). The second is a white circle with a gray outline.



- Click on the second symbol layer (the white circle).
- In the Appearance section, change the color to a shade of light blue.
- For the outline color, choose a shade of dark gray, and change the outline width to 1 pt.
- Click *Apply* to display the new changes for the school points.

- Save your project.

Step 3: Unique Values

Overview

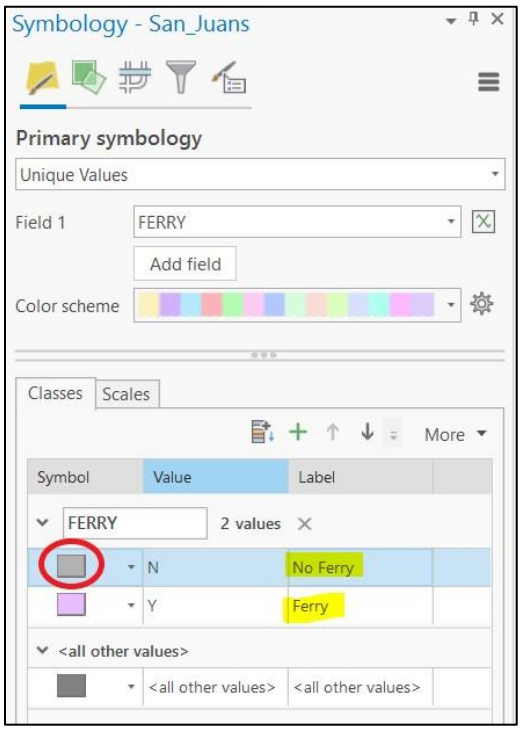
Qualitative symbology is typically used to show which categories different features belong to. The Unique Values type of symbology applies colors based on categories of features. In general, categories tend to be based on attributes that describe a feature's type, name or condition. For example, if the attribute data for a schools layer contained a field that specified each school as an elementary, high school, or university, you could symbolize the layer based on the values of that attribute, or field. In this step, we will use the San Juans layer as an example.

Instructions

- In the map Contents pane, right-click the **San_Juans** layer and click *Symbology* to open the Symbology pane.
- In the Primary symbology drop-down list, change the symbology from *Single Symbol* to *Unique Values*.
- For the "Field 1" drop-down, click FERRY.

  All of the possible values for the FERRY field should be added to the Symbol list, each with a unique color.

- From the *Color scheme* drop-down menu, choose a color pallet to your liking.



- You can also choose specific colors for the individual categories by clicking the colored boxes beside the individual symbol values to open the Format Polygon Symbol pane.
  - Use the back arrow at the top of the pane to return to the layer's symbology pane.

- Below *Label*, change the text for the symbol Labels from "N" and "Y" to "No Ferry" and "Ferry."

By default, a third category of symbol is added for "<all other values>". This is a catchall category in case there are some records that are neither "Y" or "N". If you don't want to include this in the map view Contents pane you can turn it off:

- Click the *More* drop-down menu and uncheck the box for *Show all other values*.

Step 4: Quantitative Symbology: Graduated Colors and Symbols

Overview

Quantitative symbology requires a numerical field. Quantitative data generally describes counts or amounts, ratios, or ranked values. One of the more common quantitative methods is the use of graduated colors to indicate numerical classifications.

Graduated colors can be used to show different classes of values for an attribute (such as population or rainfall). Features are grouped (classified) and each class is assigned a lighter or darker shade of color (typically darker colors indicate a larger value). Graduated symbols simply use the size of the symbol to depict the numerical data, as opposed to the color using graduated colors.

Instructions

- Turn On the **Wa_Counties_Census** layer via the map Contents pane.
- In the map Contents pane, right-click on the **Wa_Counties_Census** layer and click *Zoom to Layer*.

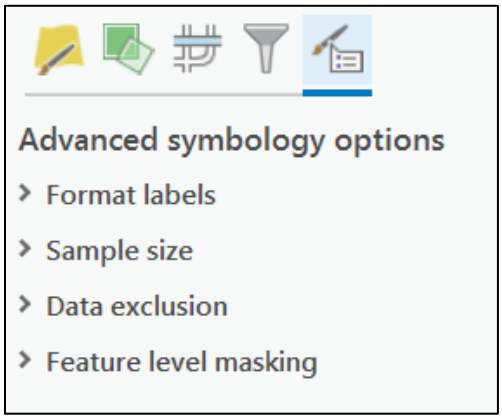
  Recommended: Turn Off all the other layers so we can focus on the counties.

- Right-click on **Wa_Counties_Census** and open its symbology.
- In the Symbology pane, choose *Graduated Colors* from the Primary symbology drop-down list.
- Then, specify the following:

Field:	**POP2000**	(2000 census population)
Normalization:	**None**	
Method:	**Natural Breaks (Jenks)**	
Classes:	**4**	
Color scheme:	**Choose a color ramp to your liking**	

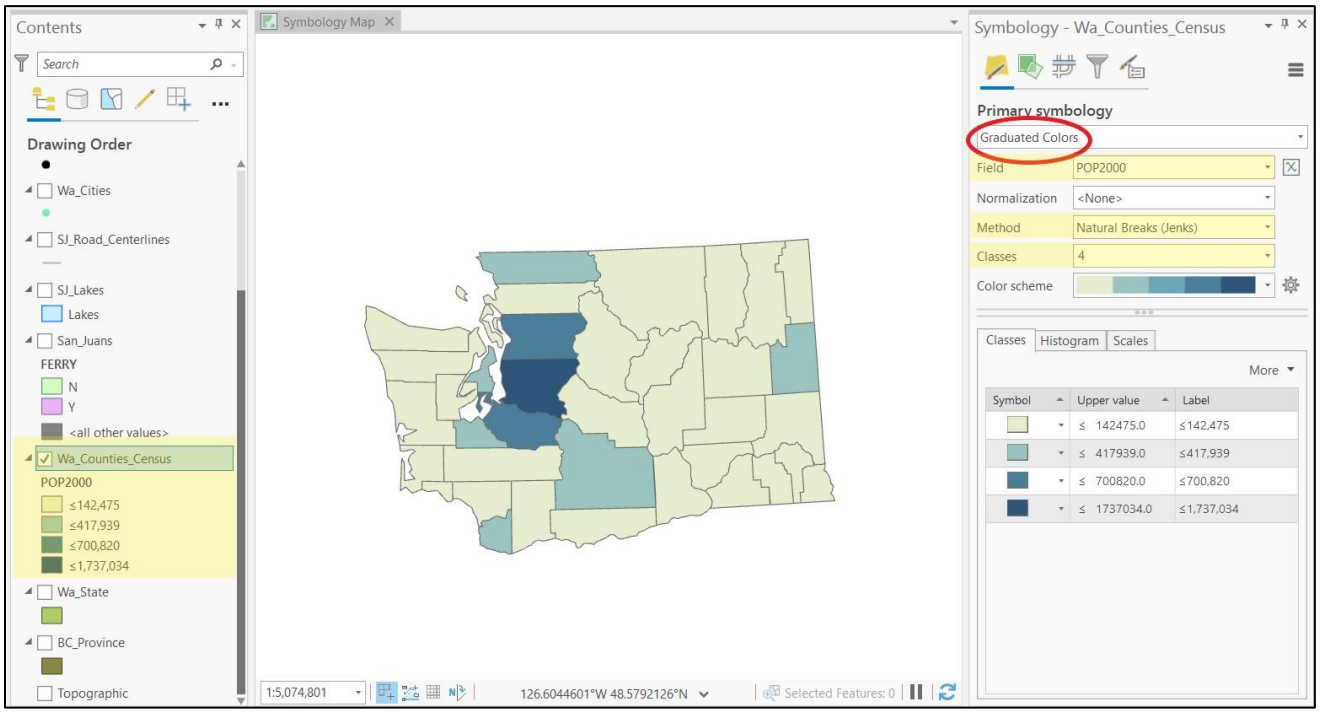
- At the top of the Symbology pane, click the *Advanced symbology options* tab.

  Remember that you can hover over a button or command to view the name and/or a screen hint.



- Expand the *Format labels* section and check the box for *Show thousands separators*. This adds commas to the symbol values/labels in the Contents pane.
- Return to the Primary symbology tab in the symbology pane (the left-hand tab)

You should now have 4 classes of population, based on the number of people in each county (in 2000), with the labels shown using commas.



Now, let's symbolize data using graduated symbols. Recall that graduated symbols use the size of the symbol to depict the numerical data, as opposed to the color using graduated colors symbology.

- In the map Contents pane, turn the **Wa_Counties_Census** layer Off.
- Turn the **Wa_State** and **Wa_Cities** layers On.
- Right-click the **Wa_State** layer and *Zoom to Layer*.
- Right-click the **Wa_Cities** layer and open its symbology.
- In the Symbology pane, choose *Graduated Symbols* as the Primary symbology to specify:

Primary Symbology:	**Graduated Symbols**
Field:	**HOUSEHOLDS** *(the number of households in the year 2000)*
Normalization:	**None**
Method:	**Natural Breaks (Jenks)**
Classes:	**3**

Optional, click the Template symbol to format the color or style of the symbol.

The cities should now be shown as small, medium, and large symbols, based on the number of households in each city.

Step 5: Normalization and Density

Overview

Showing the number of people in each county is one option for population symbology. Another is to show population density: the number of people per some unit of area (like population per square mile).

This is accomplished by "normalizing" the data by an area field. Normalizing the data creates a ratio (created by dividing the attribute values from two different fields). Normalizing by an area field (i.e., creating a density calculation) can be used to minimize apparent differences between large and small features. Any numerical data field can be used to normalize another numerical field (e.g., you could compare the current year's rainfall to the average rainfall). Normalizing simply divides the values of one field by the values of another.

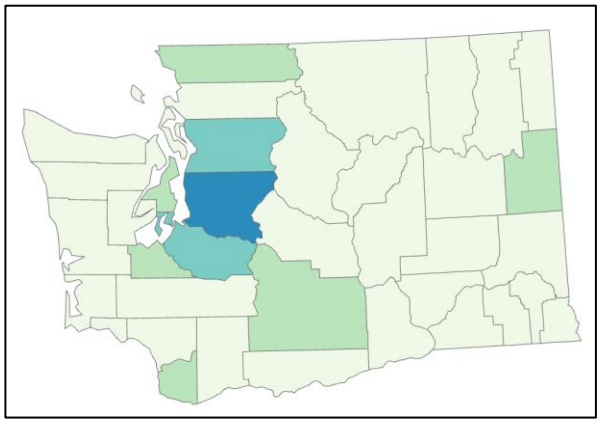
Displaying symbology based on density (a population count normalized by an area) is such a common symbolization method that some census data layers have a density field included.

Instructions

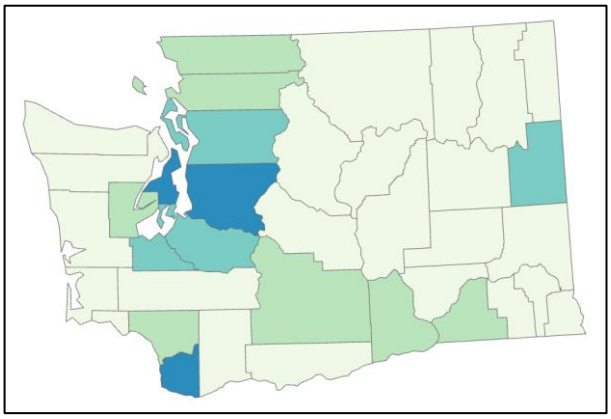
- Turn the **Wa_Cities** and **Wa_State** layers Off.
- Turn the **Wa_Counties_Census** layer back On.
- Open the **Wa_Counties_Census** layer Symbology pane (still using Graduated Colors).
- Change the Field to POP00_SQMI (Population 2000/Square Mile). Leave the other parameters unchanged from before, so the only new symbology setting is:

Field: **POP00_SQMI** *(Population 2000/Square Mile)*

This should change the appearance of the map slightly. The biggest change is in the smaller counties in the southwest part of the state (outside Portland Oregon) that have a relatively large number of people given their small size. This level of density does not show up as much when symbolizing by population alone. Switch the Field back to POP2000 to compare the difference.



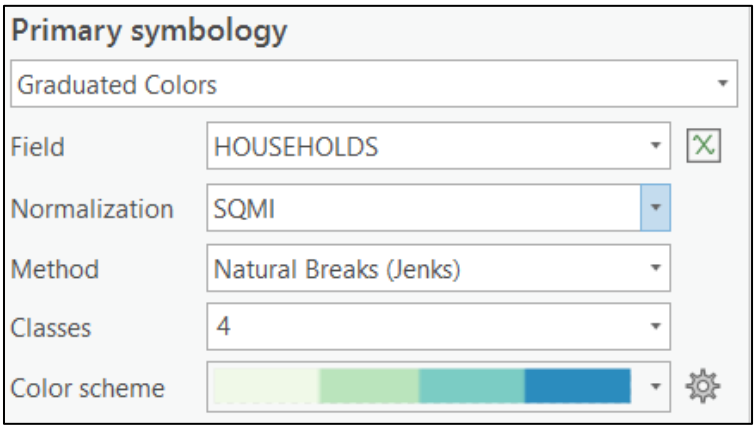
Graduated Colors, Field = POP2000



Graduated Colors, Field = POP00_SQMI

But what if you didn't have a density field in the attribute table? For example, what if we wanted to see the density of households. There is not a HOUSEHOLDS_SQMI field in the **Wa_Counties_Census** data, but we can still show the density of households per square mile:

- Change the Field to HOUSEHOLDS (the number of households in each county).
- From the *Normalization* drop-down menu, choose SQMI (square miles).
  (Do NOT use POP_SQMI).



Using a Normalization field creates a new value for each county. In each case, the number of households (from the attribute table) is divided by the number of square miles (for the particular county). These new values (households divided by square miles) are used for the symbology classes. Mathematically, this is just dividing the value of one field by the value of another field. This math is done on-the-fly (no new fields are created). Normalizing data can also be done for other operations besides density.

We can do a similar density calculation/normalization using graduated *symbols*:

- In the Contents pane, turn the **Wa_Counties_Census** layer Off.
- Turn the **Wa_Cities** and **Wa_State** layers On.
- Right-click and open the Symbology for **Wa_Cities**.

  The symbology should be the same as you left it, but this time, we will add a normalization field:

- Using the drop-down list for Normalization, select AREALAND (the area of land for each city in square miles). The symbology settings for the **Wa_Cities** layer should look like this:

Field:	**HOUSEHOLDS**
Normalization:	**AREALAND**
Method:	**Natural Breaks (Jenks)**
Classes:	**3**

Now there should be three different classes of city, shown as small, medium, and large symbols based on the number of households per square mile. You may want to zoom in to the Puget Sound Area to have less overlap of symbols.

As with graduated colors, you can set the number of classes and choose from different classification methods and/or can choose to normalize the data when using graduated symbols.

Step 5: Classification Methods

Overview

When symbolizing quantitatively, features are grouped, or classified based on ranges of values, creating classes. How features are classified depends on the *classification method* used. Classification methods are used for graduated symbology (graduated colors and graduated symbols).

When classifying data there are different classification methods that can be used:

- **Natural breaks** (class breaks placed at the larger gaps that exist in the data)
- **Quantiles** (each class contains an equal number of features)
- **Equal interval** (all classes have an equal range of values)
- **Defined interval** (all classes use a fixed size range of values)
- **Manual interval** (class breaks can be entered manually)
- **Geometric interval** (mathematically defined class ranges based on range and frequency)
- **Standard deviation** (classes based on standard deviations)

Having a basic understanding of statistics and the different types of classification methods is both helpful and important to ensure that you are creating maps that accurately depict the data being shown. In this step, we will explore just a few of these classification methods (Natural breaks, Quantile, and Equal Interval), though you are encouraged to learn about the other methods. ESRI gives a review of each data classification method on its ArcGIS Pro website Help section.

Instructions

The default classification method is natural breaks (jenks). Natural breaks places "breaks" at larger gaps that exist in the data. It classifies, or groups, features based on these “natural” breaks or gaps in the data values.

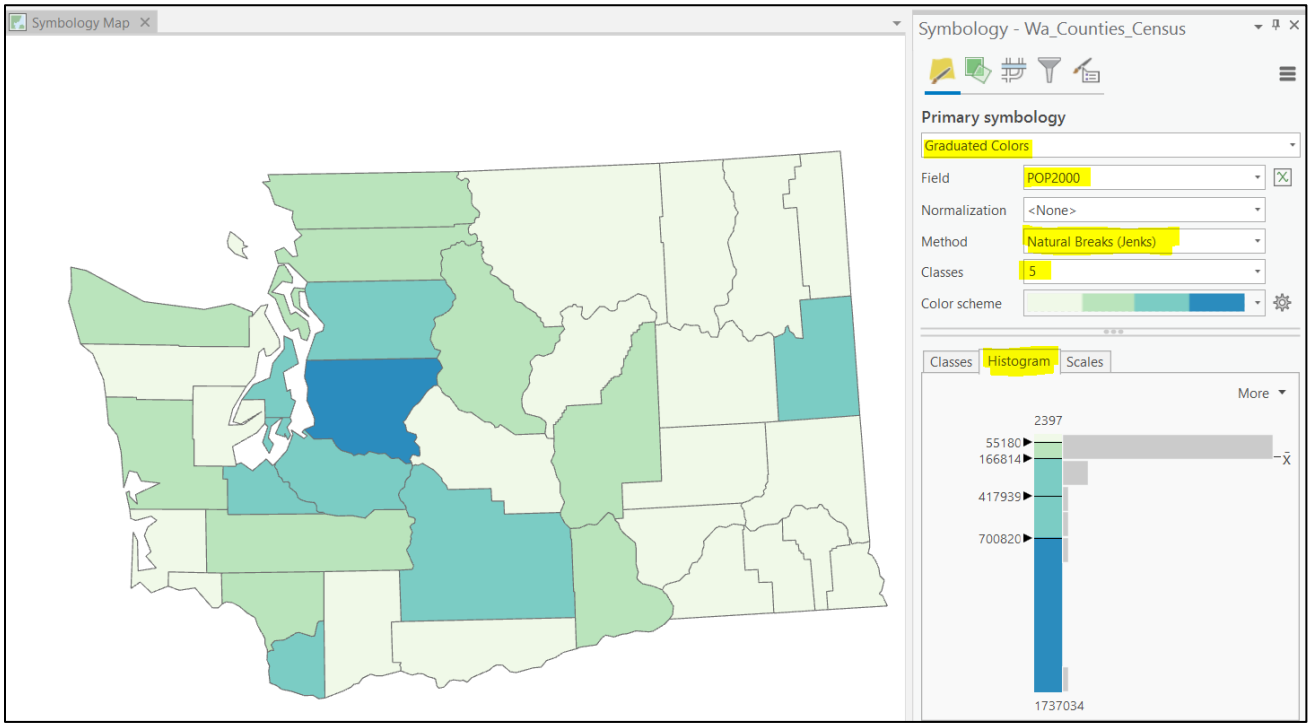
- Turn the **Wa_Cities** layer and **Wa_State** layers Off.
- Turn the **Wa_Counties_Census** layer On.
- Right-click the **Wa_Counties_Census** layer and click *Zoom to Layer*.
- Open the Symbology pane for the **Wa_Counties_Census** layer and specify the following settings:

Primary symbology:	**Graduated Colors**
Field:	**POP2000**
Normalization:	**<None>**
Method:	**Natural Breaks (Jenks)**
Classes:	**5**
Color scheme**:**	**<choose a color scheme>**

- Click to open the drop-down menu for *Method* to show the different classification methods that are possible.

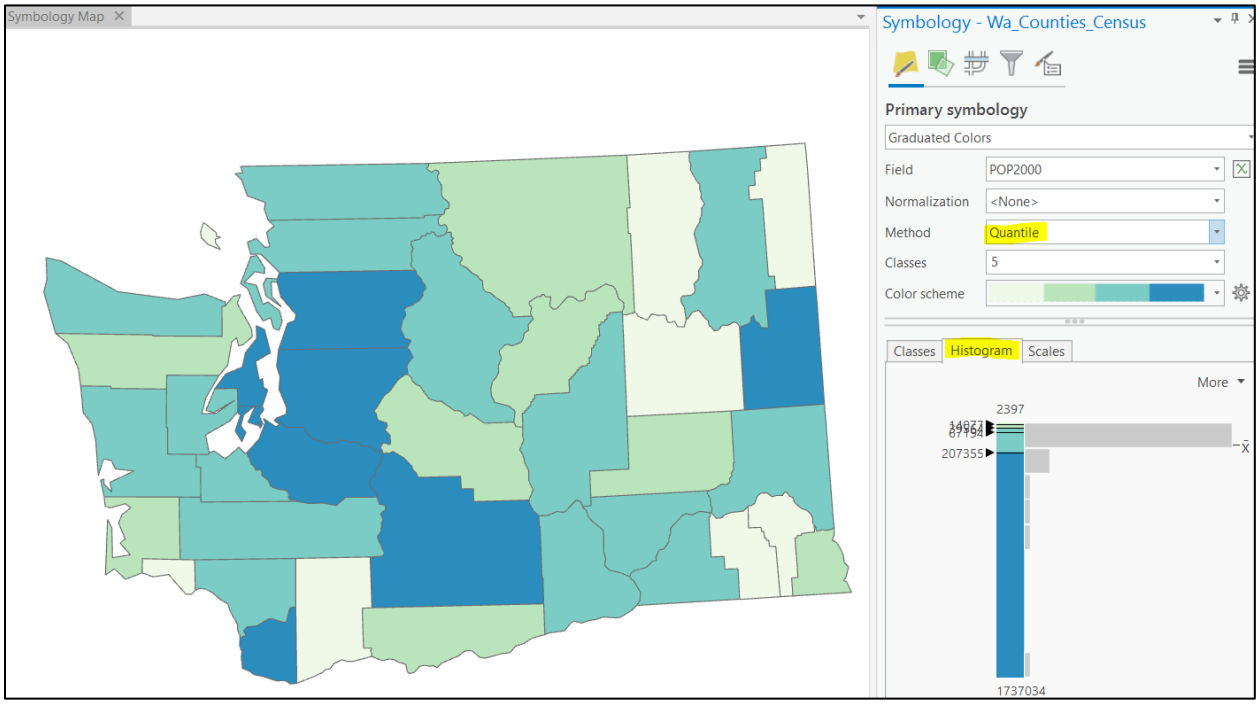
  Read the text under the name of the method in the drop-down list to see a quick summary of the method.

- Ensure the Method is *Natural Breaks (Jenks).*
- In the lower section of the Primary symbology pane click the *Histogram* tab.



The symbology histogram displays the colors used for each class, with the class break values, on the left. On the right, the gray bars on the right show the frequency distribution for the data (the frequency of values based on the chosen Field). Note that with Natural Breaks there a more counties assigned to the lower population classes, and fewer high-population counties. Analyzing the histogram can help you understand how the classification method is classifying this data.

- Still in the Symbology pane for the **Wa_Counties_Census** layer, change the classification Method from Natural Breaks to *Quantile* (still using Graduated Colors and 5 classes).



The quantile classification method creates classes by applying the same number of records, or features, to each class, rather than dividing classes based on value ranges.

With Natural Breaks there were more counties in the lowest category than any other. The Quantile method forces an equal number of features in each class. Thus, there are now fewer counties shown with the lightest color (i.e., the counties in the lowest-population category). Because the population values for the counties are clustered towards the lower-population values (with only a few high-population counties) the Quantile method distributes more of the lower-population counties into the "medium" classes.

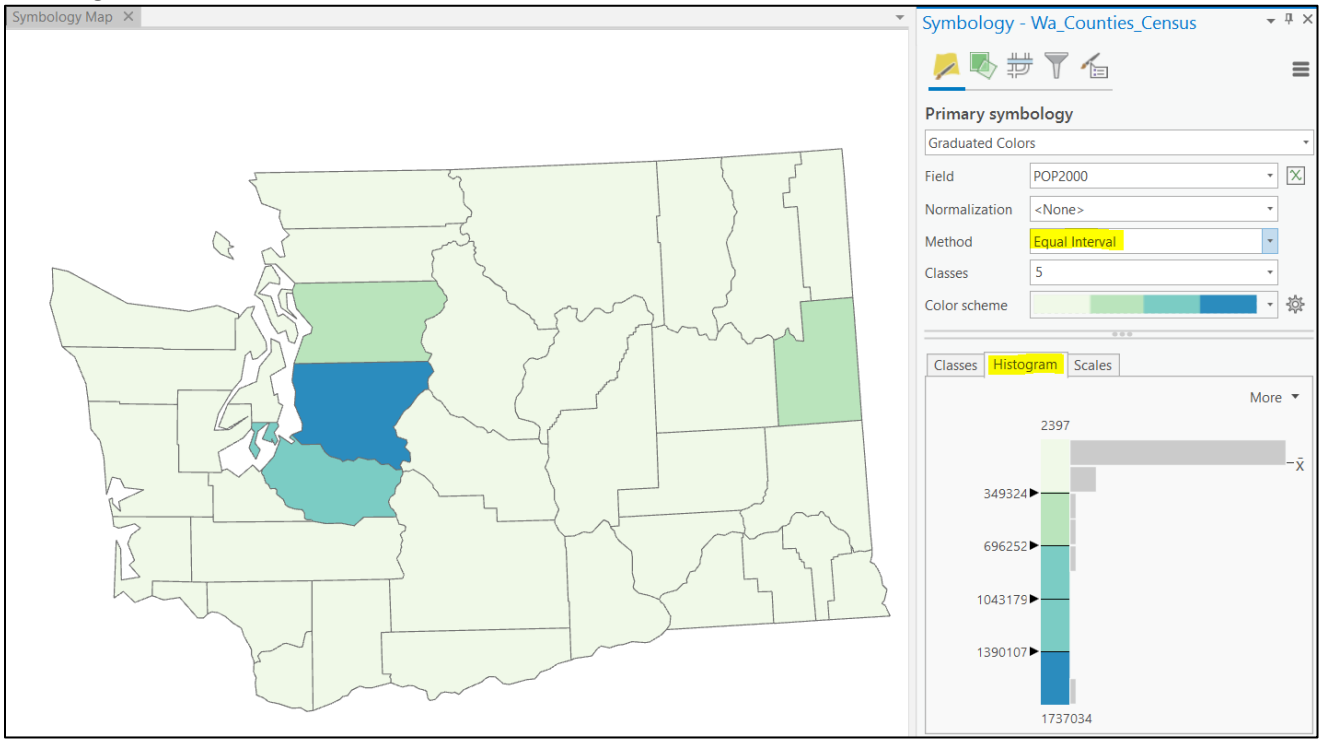
Look at the Histogram to view the class break values as compared to the frequency distribution. Since we are using the same POP2000 field, the frequency distribution (the gray bars) should be unchanged. The only difference between the Natural Breaks histogram and the Quantile histogram is the location of the class breaks on the left.

- Next, change the classification Method from Quantile to *Equal Interval*.

The equal interval classification method simply divides the range equally by however many desired classes. This method would work well with data that is evenly distributed but can underrepresent certain classes when the data is not evenly distributed.

For the counties of Washington state, almost all of the counties end up in the lowest class (the lightest color). This is because King County (Seattle) has a population much greater than any other county. This maximum value is used to determine the total range of values. This range is then divided into 5 equal class intervals. Some classes have only one county, and indeed, the 4th class does not have any counties in it at all. The Equal Interval establishes class breaks based on the total range be divided into equal ranges of values, regardless of the actual data distribution.

Look at the Histogram with Equal Intervals. In this case, a county has to have a population of more than 349,324 to appear in the second class. This map makes it look like almost all of the people who live in Washington live near Seattle.



As we can see, the same data can be represented in very different ways, and thus perceived very differently by the audience. This is why it is important to understand how different classification methods work so that you create maps that accurately depict data.

Experiment with different data classification methods, different numbers of classes and optionally normalization to learn more about how the data can be displayed...

Step 6: Unclassed Colors

Overview

Another type of quantitative symbology is an unclassified ramp of colors. Rather than breaking the data into a specified number of classes, the colors are graduated from low to high. With unclassed colors, every single feature could potentially have a unique color.

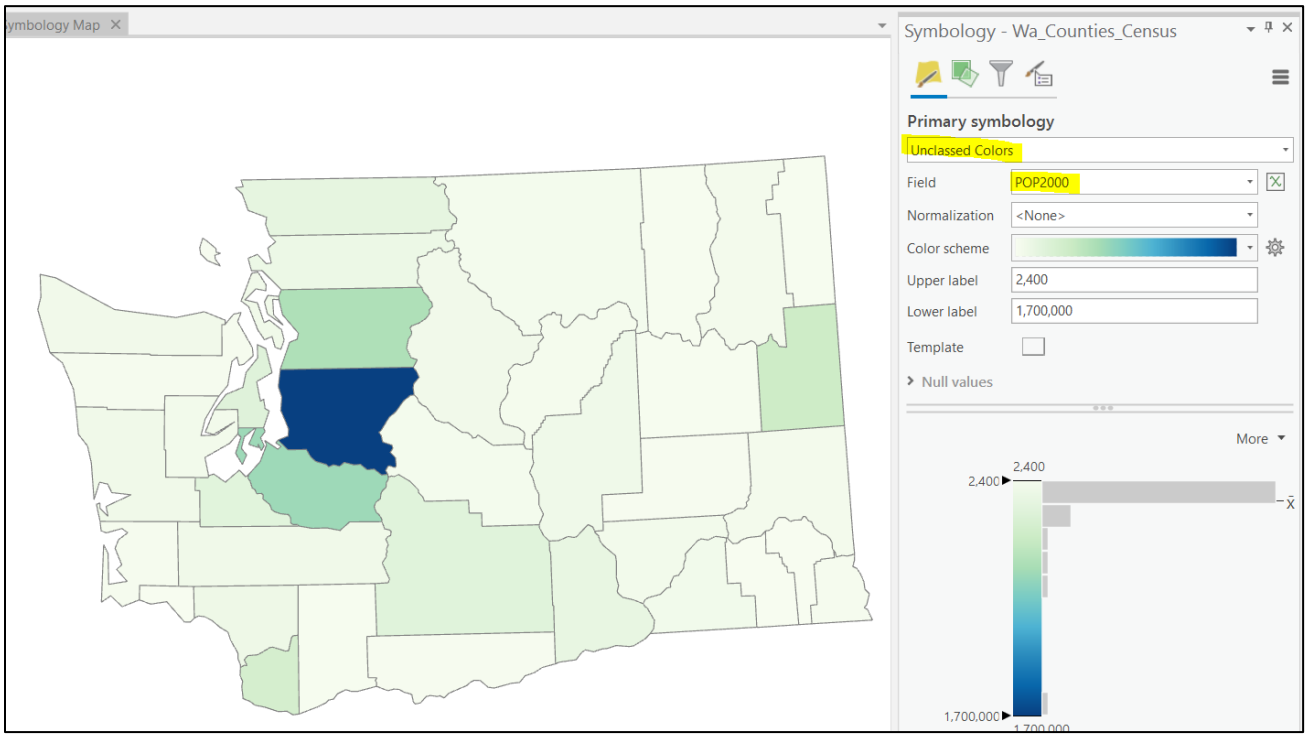
Instructions

- Still in the **Wa_Counties_Census** Symbology pane, choose *Unclassed Colors*:

Primary symbology:	**Unclassed Colors**
Field:	**POP2000**
Normalization:	**<None>**
Color scheme**:**	**<optionally choose a color scheme>**

Note that Unclassed Colors does not used classes, so there is not an option for the number of classes or the Classification method.

Where graduated colors would classify, for example, all of the counties within a given population range and assign them the same color, unclassed colors distributes the color scheme evenly across the range of values for the features. Unclassed colors use a range of colors, with potentially no two features having the exact same color. The counties are symbolized based on their relative location along the color ramp (based on the highest and lowest values of the chosen field - population in this case).



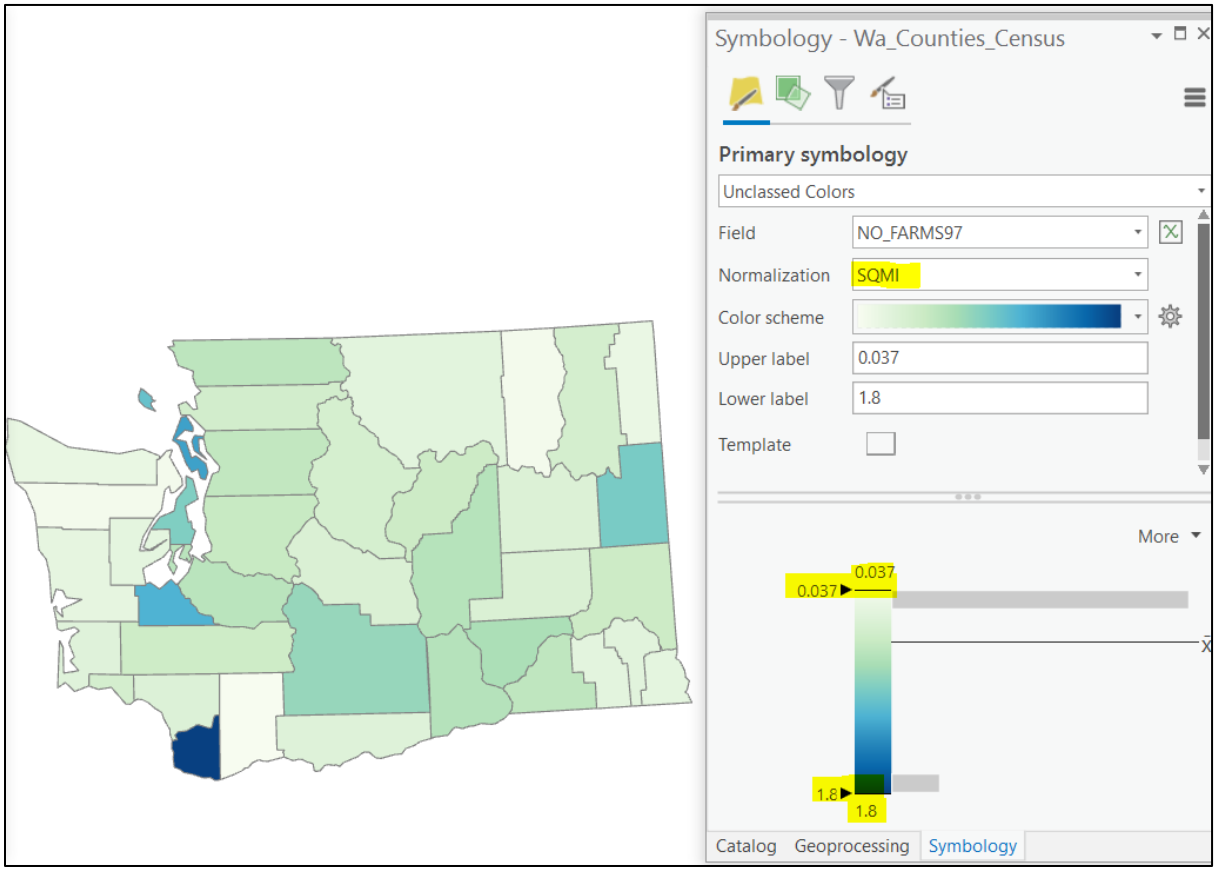
Now lets try a normalization method with unclassed colors, using a different field.

- Change the Field to NO_FARMS97 (the number of farms in 1997)

  Observe the range of values, the lowest being 63 farms for Skamania County and the highest being 3,400 farms for Yakima County (for 1997).

Now, since we would expect larger counties to have more farms, this might be a good situation to normalize the data. Meaning, depending on the purposes of our analysis, it might be more helpful to compare farm *density* (the number of farms per square mile) rather than the raw number of farms in each county. A county with a large geographic area may have more farms simply because there is more land. A smaller county with much fewer farms, might have a higher farm density. It could be potentially misleading to compare the two counties based on raw numbers alone (unless the sole purpose of your map is to simply show the number of farms in each county). Let's normalize the data:

- In the **Wa_Counties_Census** Symbology pane, use the Normalization drop-down list to select SQMI (square miles). This normalizes the number of farms per county by square miles (density).



Density symbology typically makes some of the smaller features appear darker and some of the larger features appear lighter… Add / remove the Normalization field in the Symbology pane to compare the map with and without normalization (i.e. raw number of farms vs. farms per square mile).

Step 7: Proportional Symbols

Overview

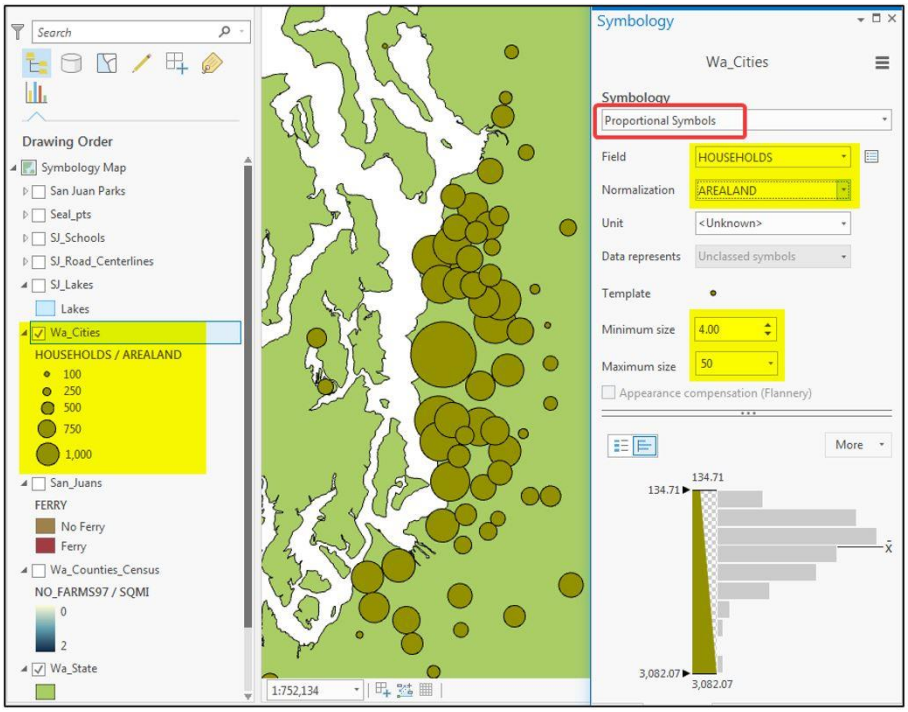
Another way to symbolize quantitatively is with proportional symbols. With proportional symbols, the sizes of the symbols are shown as a continuous range of sizes rather than being divided into classes (somewhat similar to the way unclassed colors displays a range of colors without breaking the data into classes – Proportional Symbols might be called 'Unclassed Symbols').

Instructions

- In the map Contents pane, turn Off the **Wa_Counties_Census** layer.
- Turn On the **Wa_State** and **Wa_Cities** layers.
- Change the **Wa_Cities** symbology to *Proportional Symbols*:

Primary Symbology:	**Proportional Symbols**
Field:	**HOUSEHOLDS**
Normalization:	**AREALAND**
Minimum size:	**4**
Maximum size:	**50**
Color scheme**:**	**<optionally choose a color scheme>**

Cities are now shown as a range of small to large symbols based on the number of households per square mile. Similar to graduated colors, proportional symbols are not grouped into classes. With proportional symbols every symbol could potentially be a unique size. You may want to zoom in.



Step 8: Heat Maps

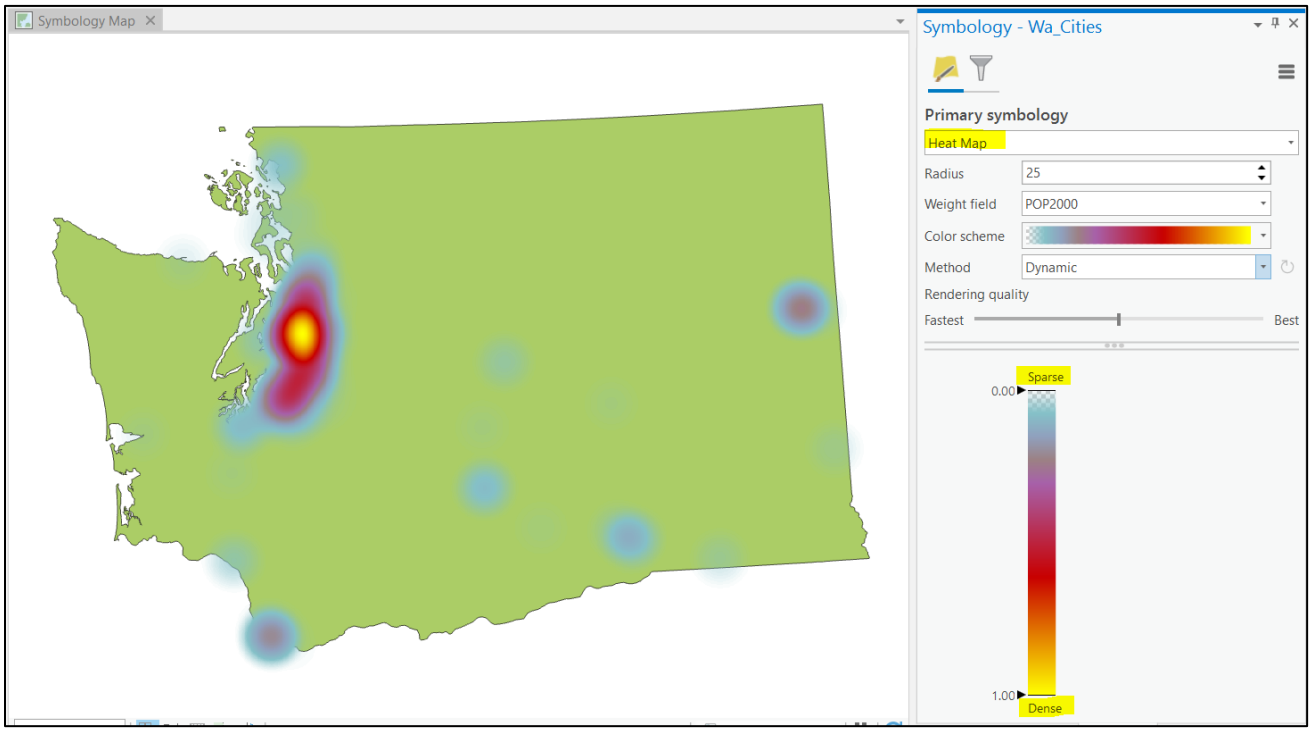
Overview

Point features can also be displayed using a Heat Map style of symbology. A "Heat Map" interpolates the areas between the points to show concentrations of data ("hot spots"). This can be great for consolidating large numbers of point features where it is otherwise difficult to visually difficult to differentiate a concentration of points from many points.

Instructions

- In the Contents pane, leave the **Wa_State** and **Wa_Cities** layers On.
- Right-click the **Wa_State** layer in the contents pane and *Zoom to Layer*.
- For the **Wa_Cities** layer symbology choose *Heat Map*:

Primary Symbology:
Radius: **25**
Weight field: **POP2000**
Method: **Dynamic** *(the density is recalculated as the scale changes)*



Since the radius of the heat map is dependent on scale (and is set to change dynamically), zoom in to the Puget Sound region for a more nuanced heat map. Experiment zooming in and out of the map in different areas of the state.

Step 9: Dot Density

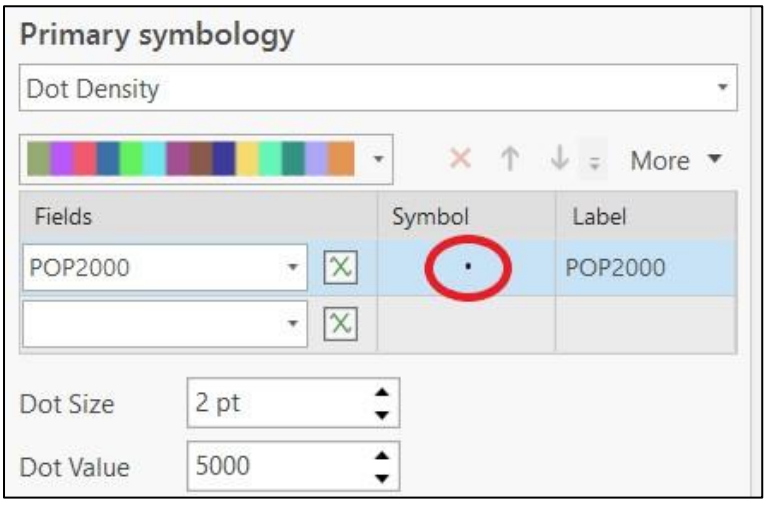
Overview

There is also a symbology option (for polygons) to create a dot density map. Dot density symbology distributes a random display of dots across each polygon, with each dot representing a certain number of cases or instances (i.e. one dot = 500 people, or 1 dot = 200 car accidents).

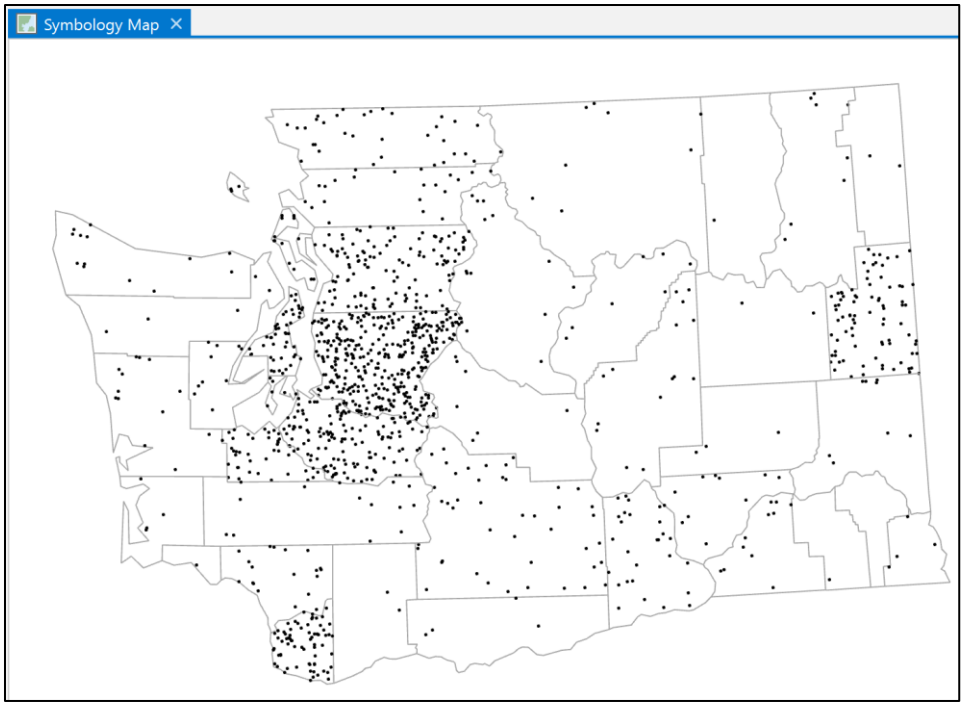
Instructions

- In the map Contents pane, turn the **Wa_Cities** layer Off.
- Turn On the **Wa_Counties_Census** layer.
- Right-click the **Wa_Counties_Census** layer and click *Zoom to Layer*.
- For the **Wa_Counties_Census** symbology, choose *Dot Density*:
  - In the Fields dropdown specify **POP2000**
  - Change the color of the dots by clicking on the dot symbol below Symbol (beside the POP2000 Field).
  - Choose a color and click *Apply*.
  - Use the back arrow in the upper left corner of the Symbology pane to go back to the Primary symbology pane.

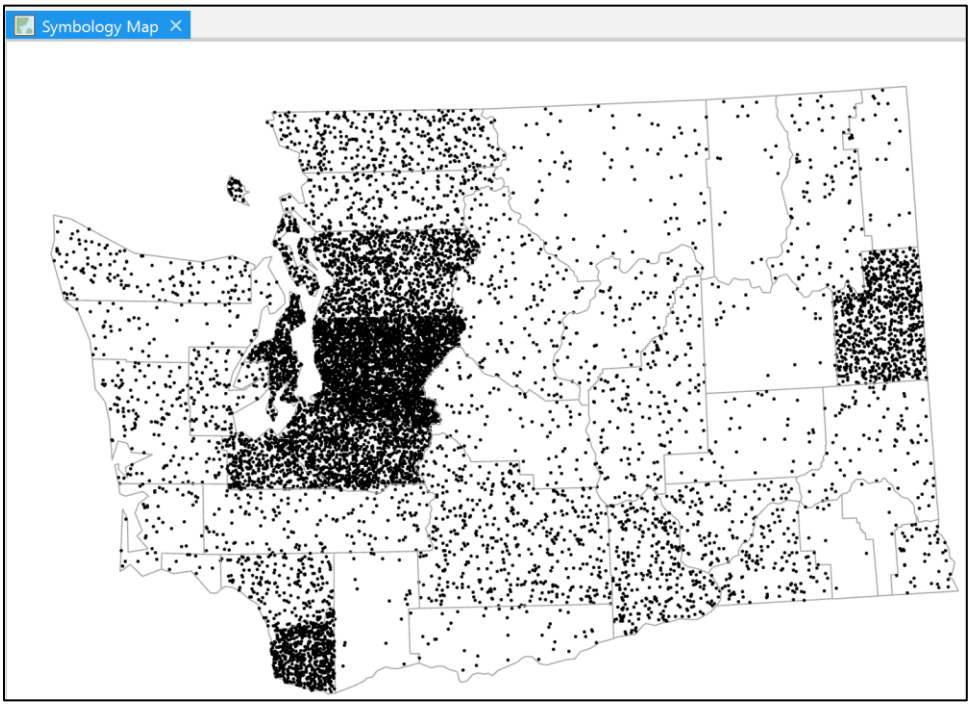
*You may want to turn Off the **Wa_State** layer Off (or change its color) to better see the dots. You can also change the color of the Dot Density Background by clicking on the Background symbol icon.*



In the example above that the Dot Value is 5,000, meaning that each dot represents 5,000 people. Dot Size is just the size of the symbol (it has nothing to do with the dot value). Experiment with the Dot Size and Dot Value to display more or fewer, larger or smaller, dots.



Dot Density symbology with one dot representing 5,000 people.



Dot Density symbology with one dot representing 500 people.

One of the drawbacks of using Dot Density symbology is that the dots are randomly and uniformly distributed across the feature. This can lead the viewer to assume that the population is likewise evenly distributed across the geography, while this is rarely the case. Humans tend to live near roads, waterways, and near other humans (i.e., in cities). In addition, low population counties may look entirely uninhabited and the breaks between county borders can appear overly dramatic.

End of Exercise

Symbology

Warthog Information Services