Lab #4: Thematic Mapping

OBJECTIVE: The primary objective of this lab is to understand some of the decisions cartographers make when creating dot density and choropleth maps.

SPECIFICS: Working with the state of Pennsylvania, first, you will create a dot density map to show the distribution of population, and second you will create a choropleth map to show the spatial distribution of income. In both cases your write-up will discuss the decisions you made to make the maps.

DELIVERABLES: One zip file (i.e., ajeu04.zip) containing the MS Word document for the write-up and one layer file from the dot density map. The choropleth maps will be embedded (included) in your write-up.

Part 1: Download and extract data

- Create a lab4 folder in U:\classes\gtech201\labs\n- Download data to your lab4 folder.
- Using WinZip, extract the data to your lab4 folder.

Part 2: Create a dot density map showing the distribution of population

For over 100 years, cartographers have used the dot mapping method to illustrate the spatial distribution of discrete geographic phenomena. With this technique, a point symbol of the same size and shape is repeated over an area to represent the value of an attribute. Any geometric figure may be used as the point symbol, but the round dot is the most common symbol used. It is through the repetition and placement of these dots, that the nature of the dot distribution emerges and the map reader can discern patterns. Thus, if the dot map is made well, the map reader’s eyes will not be attracted to individual dots, but to the changing spatial density across the map.

Two forms of dot mapping may occur, one-to-one or one-to-many mapping. In one-to-one mapping, one dot represents one feature instance. An example would be one dot represents one person. In a one-to-many mapping environment, one dot represents many feature
instances. For example, one dot represents 25 people. Usually in the mapping of population distributions, the later form of dot mapping is used. Rarely is it possible to represent every single person with a dot.

- Start ArcMap using a new empty map.
- Add the state.shp, county.shp, block_group.shp, and block_pop.shp data.
- Open the properties for the counties layer.
- In the Layer Properties window, click on the Symbology tab. Next, from the Show: list click on Quantities, and then select Dot density.

- In the Field Selection list, click on Population and then click the button. You can take a preliminary look at your dot density map by clicking on the Apply button at the bottom right of the window. Your map will update without closing the Layer Properties window. Click on the OK button to return to your map.

In creating a dot map, three important factors need to be considered. They are unit value, dot size, and dot location. Other relevant factors include determining a map scale, selecting a statistical unit to map the data at, and determining areas for dot placement. Producing a good dot map means striking a balance between the unit value and dot size. The challenge is finding
a unit value that will produce crude patterns of dot distributions in sparsely dense areas and does not result in an excessive number of dots in highly dense areas. Ultimately, the dot map is judged against what looks right and knowledge of the real distribution of the phenomenon.

- **Unit Value:** The unit value refers to the numerical value that each dot represents. A dot value of 30,000 means there will be one dot for every 30,000 people. You can adjust the dot value by either typing a new value or using the slider.

  If the unit value is too large, the dots appear to be spaced far part in dense areas and even more so in sparse areas. Widely spaced dots will produce an empty map. If the unit value is too small, the dots may unite to form a black area. The presence of a great many dots may suggest a much higher accuracy in their placement than actually exists. It has been suggested to choose a unit value in which two or three dots will be placed in the statistical unit that has the least mapped quantity. By having at least two or three dots for each statistical area, patterns of dot distributions will begin to form and with the variances in patterns across the map, the map’s message begins to appear.

- **Dot Size:** Dot size refers to how big or small the dot shall be. You can adjust the dot size by either typing a new value or by using the slider.
It is typical for the cartographer to experiment with various sized dots before a final decision is made. If the dots are too small, the distribution will appear sparse and insignificant, and patterns will not be visible; The map reader will not be able to discern any pattern by looking at the map. If the dots are too large, the map will give an impression of excessive density, which is erroneous. When the dots coalesce too much in the dark areas of the map, the result is an ugly map where the dark areas visually dominate the other information on the map. Therefore, dots should begin to coalesce or just touch with one another in statistical areas that are highly dense.

<table>
<thead>
<tr>
<th>(a)</th>
<th>(b)</th>
<th>(c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>. .</td>
<td>. .</td>
<td>. .</td>
</tr>
</tbody>
</table>

*Dot size.* (a) Dots appear too small and are hard to see. They fade into the background. (b) Dots appear too large and overlap with one another. (c) Dots are sized somewhere between those in a and b so that they are just touching one another.

- **Dot location:** Dot location refers to where the dot will be placed on the map. Ideally, the center of the dot would be placed at the location of the feature instance. On a one-to-one dot map, this is not difficult. However, on a one-to-many dot map, placing the dot becomes problematic because it must be decided where to place a dot that represents several differently located feature instances. Subjective placement of dots within an area introduces personal judgments and produces a biased map. Three proposed solutions to dot placement are possible: (a) uniformly distribute dots; (b) randomly distribute dots; or (c) apply the center of gravity principle.

The first option is to spread the dots uniformly across the statistical area. This approach is not recommended unless the real distribution occurs in such a manner. Rarely does geographic phenomena occur uniformly across space, so this approach would not be appropriate to use. Mapping the phenomenon in such a manner introduces a series of problems: the resulting map suggests no variation in density; produces a misleading ‘average’ picture; gives an impression of a stepped distribution; and does not reflect any abrupt changes present in the data.

The second option is to randomly distribute the dots within the statistical area. Dots are placed in the statistical unit of enumeration regardless of where the actual location of
the feature instance occurred. As a consequence of this technique, accurate location control is lost. This technique is acceptable for small scale mapping when the statistical unit is small. If applied to large-scale maps, unrealistic clusters and gaps will result in the dot pattern.

The third option is to apply the center of gravity principle. According to this principle, the dot would be placed in the statistical unit to best represent all the individual feature instances in the unit.

![Center of gravity principle. The red dots show the actual location of a feature instance. The black dot represents a position to place the dot. The position of the black dot is drawn towards the three red dots clustered together. The remaining dots produce a pull effect that alters the final position of the black dot.](image)

Another issue to be aware of in dot placement is political boundaries of statistical units. These boundaries may have been devised only for enumeration purposes, thus they should not be used to dictate the location of the dots. For example, population settlements occur irrespective of political boundaries, and under these circumstances, dots should be placed on or very near the boundary line. Doing otherwise would result in a map with areas that are clearly separated by white lines.

- **Statistical Unit:** Dot mapping usually involves the presentation of quantities by a statistical unit. In most circumstances, the smaller the statistical unit, the greater the accuracy of the final dot distribution. Smaller statistical units mean a smaller area for each dot, thus reducing the chance of locational error. For the sake of accuracy, large scale maps require statistical units smaller than blocks. Use of county subdivisions, tracts, or blocks are suitable for intermediate to small scaled dot maps.
- **Dot Color:** The dot color represents the color of each dot. Double-click on the colored dot or “Population” to open the Symbol Selector window. You can also change the symbol, size, and angle of the dot here.
- **Background Color:** In the Background section, click the button. A menu will appear with many color selections. You will select the color you want to use as the background color for each county.
- **Background Style:** In the **Background** section, click the button. The **Symbol Selector** window will appear allowing you to change the style, color, and width for each county’s border.

- Check **Maintain Density** to preserve the dot density. When checked, as you zoom in, the dot size will increase so a given area will visually appear as dense. Otherwise, the dot size will remain constant.

- Click the **Properties** button to set the dot placement options and use masking.

- The **Min, Mean, and Max windows** give an example of what the most populous (max), least populous (min), and average (mean) counties will look like on your map.

- Remember to click on the **Apply** or **OK** button to view the changes on your map.

---

**Pennsylvania’s Population Density**

Your Assignment: Experiment with various dot sizes, unit values, dot colors and statistical units to produce a legal-sized (8.5” x 14”) population dot density map. Your map scale may range from 1:1,500,000 to 1:2,000,000, but try to use rounded numbers. Once you are satisfied with your choices, save your symbology information in a layer file. In the table of contents, right click on the layer and select **Save As Layer File...** Use the following naming convention: 
**ajeu_dotsize_unitvalue.lyr**, where ajeu is your name, dotsize is a number, and unitvalue is a number. Your layer file will be named something like **ajeu_10_30000.lyr**.
**Write-up:** What were some of the combinations you tried for your dot density map? How did those combinations change your perception of the distribution of Pennsylvania’s population? What combination did you choose for the final map? Why did you make those choices? Include a discussion on the accuracy of your map.

**Part 3: Create a choropleth map showing the spatial distribution of income**

A choropleth map is a thematic map in which areas are shaded or patterned in proportion to the measurement of the statistical variable being displayed on the map.

- Start **ArcMap** using a **new empty map**.
- Add the **counties_income.shp** data.
- Open the **properties** for the counties layer.
- In the Layer Properties window, click on the **Symbology** tab. Next, from the **Show** list click on **Quantities**, and select **Graduated colors**.
- In the Fields area, under **Value**: select **Income**.

![Layer Properties dialog box](image)

In the **Color Ramps** area, you can see the classes and ranges for income.
- Open **ArcMap's Help** (i.e., ArcGIS Desktop Help). Click on the **Contents** tab and navigate to **Mapping and visualization -> Symbolizing data -> Applying symbology**. Read the sections on **Ways to map quantitative data** and **Standard classification schemes**.

Experiment with different classification schemes and number of classes to produce a choropleth map.

- **Classification method**: Click on the **Classify...** button to open the **Classification** window. Under **Classification Method**: click on the down arrow to view a list of classification schemes. Any changes you make will affect the histogram and the break values.

To decide which scheme to use, you need to know how the data values are distributed across their range. Looking at a histogram for your data, the horizontal axis represents the attribute values, while the vertical axis represents the number of features having a particular value. Here are some guidelines when choosing a classification scheme:
If your data is unevenly distributed (many features have the same or similar values, and there are gaps between groups of values), use natural breaks.

If your data is evenly distributed and you want to emphasize the difference between features, use equal interval or standard deviation.

If your data is evenly distributed and you want to emphasize the relative difference between features, use quantile.

- **Number of classes**: In the Classification field, click on the drop down arrow under the Classes field. Any changes you make will affect the histogram and the break values.

Most map readers can distinguish up to seven colors on a map, so using more than seven classes will make it hard to find features with similar values. Four or five classes will usually reveal patterns in the data, without confusing the map reader. Using fewer than three or four classes doesn’t show much variation between features, and hence shows no clear patterns.

If you’re exploring the data to see what kinds of feature groupings and patterns emerge, you may want to use more classes at first. Each feature will be in a narrower range, with values closer to its actual value.

- You can change the color scheme used to show your results by clicking on the Color Ramp.

Usually, you assign shades of one or two colors to the classes. You pick the colors representing the lowest and highest classes, then pick the intermediate colors or let ArcMap assign them for you.

If you have fewer than five or six classes, use one color and vary the shade. Most people can only distinguish up to seven colors. Most people also interpret darker colors to mean “more” or “greater,” so assign the darkest shade to the higthest class. If the color scheme is reverse, right click on one of the symbols and select Flip Symbols from the pop-up menu.
Different colors have different visual impacts. Red and oranges attract the most attention; blue-green, the least.

It is easier to distinguish between shades of purples and blues than shades of other colors, so you might use these if you have more than four or five classes. Keep in mind that certain colors have special meanings for some people. For example, red is often used to indicate hot spots, such as areas with many crimes, or areas unsuitable for a particular use, such as those too steep to build on.

If you have more than seven or eight classes, you may want to use a combination of colors and shades, using two or even three colors to help distinguish the classes. Warm colors (red, orange, or yellow) are a good choice for the classes representing higher values since they highlight these values; cool colors (green, blue, or purple) can be used for lower values.

Using two colors is also good for showing data with both positive and negative values, such as percentage above or below an average value: shades of one color (such as red) would show percentages above the average, while shades of another color (such as
blue) would show percentages below, with a neutral color showing the average. This approach is particularly good for maps using classes based on standard deviation.

- To change the appearance of the label, right click on one of the symbols and select **Format Labels**... You may also click on the values in the label field to manually change the values or to label the classes differently.
- Remember to click on the **Apply** or **OK** button to view the changes on your map.
- In the **Layout View**, create a printable map in landscape format. Include an appropriate title and legend on your map. Feel free to add in other map elements such as a north arrow, scale bar, neatline, text identifying the classification scheme.
- Export your map as a GIF.

**Your Assignment:** Create four or five choropleth maps based on income. All maps will use the same number of classes and color schemes, except the standard deviation classification scheme. Use the natural breaks, equal interval, quantile, and standard deviation classification schemes. You decide on the number of classes.

**Write-up:** Compare and contrast these four choropleth maps. How is your perception of the distribution of income changed (if at all) amongst these maps. Which classification scheme best represents your data? If using a defined interval, include that as a fifth map and explain why. Be sure to include the maps you created in ArcMap as part of your write-up. In MS Word, use the **Insert -> Picture** (from file) function.