Getting the Right Data Visualization

Visualizations foster understanding. They weave data points together to tell stories, and people take action based on what they understand from those stories. So visualizations need to be clear and accurate, giving decision-makers confidence in their choices.

Insights for ArcGIS makes it easy to create visualizations. The data analytics workbench, now available in ArcGIS Online, allows users to explore both spatial and nonspatial data in several different formats—various maps, charts, and tables.

But with so many options for visualizing data, how can you be certain that you're selecting the right visualization? Start by asking three key questions, and you'll be on your way.

1 Is the data qualitative, quantitative, or temporal?

Answering this question will help determine which visualization works best for the data at hand.

Qualitative data is categorical, meaning that the data values are descriptive. Examples of qualitative data include county names, soil types, and animal species. In Insights for ArcGIS, these data types are often stored in string fields.

Quantitative data contains values that show measurement—for example, total sales, the average grade of students in a class, or pipe length. In Insights for ArcGIS, these data types are often stored in number or rate/ratio fields. Because the numbers are measurements, they are often continuous rather than discrete.

Temporal data is collected over time. For example, it can show the number of lightning strikes that occur during a storm or car counts from traffic sensors. In Insights for ArcGIS, these data types are stored in date/time fields.

2 What story needs to be told?

Determine the purpose of the visualization. What question needs to be answered? What message needs to be communicated? Take a look at these six groupings of visualizations to help figure that out.

Measure

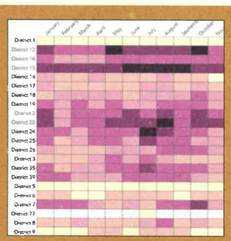
Measure visualizations are ideal when trying to ascertain the size, amount, or degree of a variable or phenomenon in an area, such as the amount of crime in a city or the value of sales. Charts are often a good choice for displaying this kind of data, and there are several types to choose from.

Bar graphs use either horizontal or vertical bars to show comparisons among categories. They make it easy to identify broad differences between groups at a glance.

Bubble charts represent numerical values by categories. Each bubble stands for a unique category, and the size of the bubble symbolizes the magnitude of the numerical variable, such as count or cost. These make it easy to compare categories.

Heat charts show total frequency in a matrix. Values in each cell of the rectangular grid are symbolized into classes, helping viewers see how the data is distributed into high and low values, trends, and more.

A crime analyst studying theft-related crimes in her city wants to find out which types of crime are happening most often and when. For this, she puts burglary data into a heat chart to visualize when these kinds of crimes occur most often. With this, she can see that District 12 has the most burglaries in May and October, while District 15 tends to have a large number of burglaries over the summer.



Treemaps also make it easy to see the different sizes of various categories compared to the whole. (See the "Part to Whole" section for more information.)

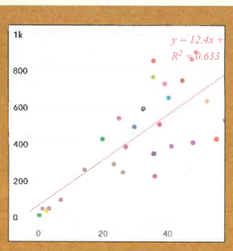
Relationship

When demonstrating a link or similarity between two or more things—such as how much money a company spends on advertising versus the amount of revenue it brings in—relational visualizations work best. There are a number of great options for this.

Choropleth maps allow quantitative values to be mapped by area, enabling viewers to see the highs and lows. These maps use normalized or proportional values, not counts collected over unequal areas or populations.

Scatterplots enable viewers to look at the relationship between two numeric values with both axes showing quantitative variables. The level of correlation among values can also be quantified, making it easy to identify causation.

A public works department has noticed an increase in water main leaks. Employees suspect that the number of leaks is directly affected by the total length of the pipes. But in some districts, this correlation appears to be true to a lesser degree. For those, the employees use a scatterplot to look at the total number of leaks versus the total length of pipes in each water district. Given that the best-fit line goes from the lower-left corner to the upper-right corner of the chart, there is a positive linear relationship between the total length of pipe and the number of service requests called in for those pipes. Thus, pipe length is likely a significant factor in all the leakages.



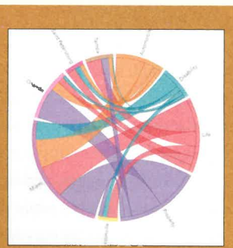
Chord diagrams and **spider**, or **desire, lines** work well, too.

Interaction

Visualizing interactions helps viewers understand the flow of information or goods. Diagrams and charts in this grouping might show migration patterns or trade among different countries. Ideal visualizations for qualitative data include the following:

Chord diagrams best reveal the interrelationships between categories—that is, how each of several categories is related to the others. With chord diagrams, viewers can easily compare similarities within a dataset or among different groups of data.

An insurance agent is reviewing the types of policies her company offers. One step in the review process is to compare the total insured value (TIV) in each policy class across cities. Using a chord diagram, she can display the policy class values (property, life, disability, and automobile) and city values (Jacksonville, Miami, Orlando, Saint Petersburg, and Tampa) as differently colored arcs around a circle. The length of the arc and the thickness of the chords are determined by the sum of the TIV. The agent can see not only which city or policy class records the highest and lowest values, but she can also determine the policy classes that contribute to the TIV for each city. From this diagram, it is clear that Miami has the largest TIV in each policy class, while Saint Petersburg and Jacksonville only have policies in three out of the four policy classes.



Spider lines, also called **desire lines**, show paths between origins and destinations, illustrating connections between places.