perceptual edge

Building Insight with Bricks

A New, More Perceptible Method for Encoding Quantitative Values in Geospatial Displays

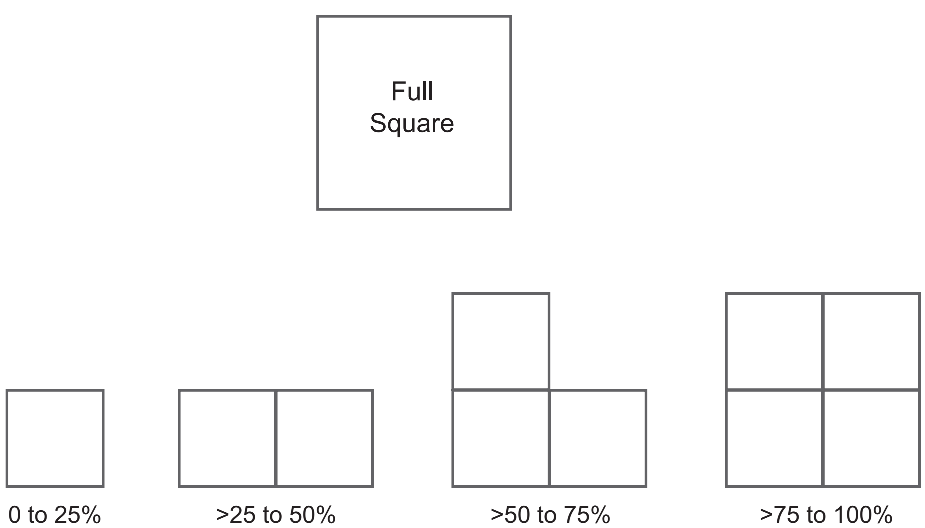
Stephen Few, Perceptual Edge

Visual Business Intelligence Newsletter

April/May/June 2013

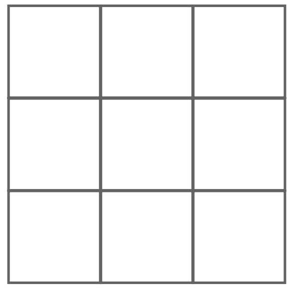
In the field of data visualization, we are always looking for better ways to encode quantitative information graphically—ways that can be easily, efficiently, and accurately decoded. Displaying quantitative values geospatially has always posed a particular challenge, because the most perceptually effective means of encoding values cannot be used: 2-D position is already used for the geospatial location of the value objects, and the lengths of value objects (e.g., bars) cannot share a common baseline because they must be positioned in specific geospatial locations. Consequently, we rely on inferior value encodings: primarily the sizes or color intensities of objects. Objects such as circles (a.k.a., bubbles) of varying sizes are commonly used, but our perception of area differences is poor. Coloring geographical regions in the form of a heat map is even worse, because we compare color intensities imprecisely and the varying sizes of the regions skews our perception of the colors that they contain. For example, if two regions are encoded with precisely the same dark gray, we tend to perceive the larger region as darker than the smaller.

While pondering this problem, I began to explore ways to improve the ease and precision of geospatially located values by adding two more preattentive visual attributes to objects that already vary in size: numerosity and shape. We can preattentively perceive quantities of objects up to a limit of three or four, which involves a perceptual process called *subitization*. To illustrate, we could subdivide a square into four smaller squares such that one of the smaller squares represents a range in value from 0 to 25%, two smaller squares combined represent values greater than 25% to 50%, three smaller squares combined represent values greater than 50% to 75%, and all four smaller squares combined represent values greater than 75% to 100%, as illustrated below.

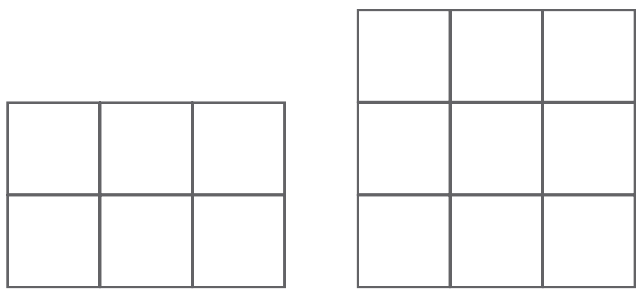


Now, when comparing these objects in a geospatial display, such as on a map, we would not just be comparing their size differences, but their count and shape differences as well.

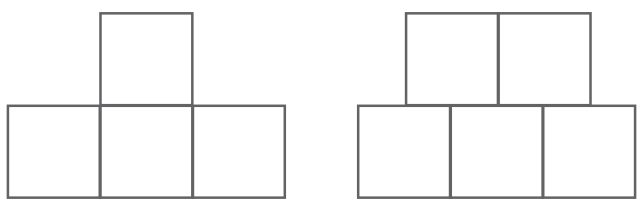
Four binned quantitative ranges are not enough, however, for effective geospatial quantitative displays. What if we expanded the number of ranges by arranging small squares into rows to combine the attributes of count and alignment together (i.e., the Gestalt principle of continuity) to form preattentively recognizable groups? In the following example, a large square is subdivided into nine smaller squares, arranged as a matrix of three rows.



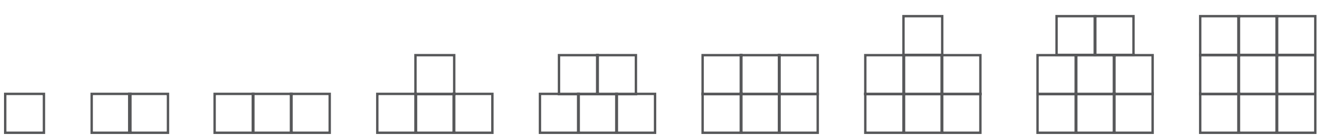
In theory, the rows could be subitized, from one through three, as well as the squares that make up those rows, also from one through three, simultaneously. In the following example, the difference between the object that is composed of six squares versus the one that's composed of nine squares can be preattentively perceived.



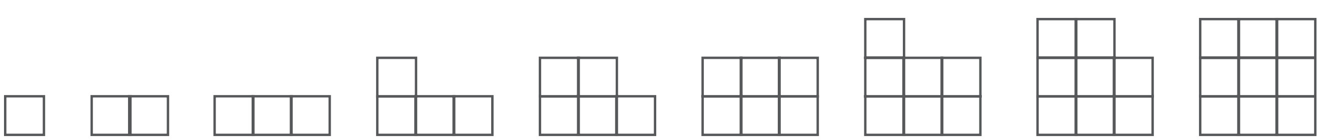
The next example is a bit more complex in that the top row is not complete in either of the two objects below. Even though they differ by one square only, this difference is easily detected. We can instantly see that one is greater than the other and also that it is greater by one unit.



Below is a set of nine combinations of squares that we can use to encode nine binned quantitative ranges:



Notice that, rather than adding squares to a new row aligned from left to right, I have positioned the first square of each new row in the horizontal center, and when two squares appear, I've positioned them in the center as well. I did this because, to my eyes, this arrangement produces shape differences with each new addition of a square that are more easily discriminated than the left-to-right arrangement, illustrated below:



To determine if my intuition is correct, however, this should be tested with eyes other than my own.

These combinations of squares are arranged much like a set of bricks that might be stacked and joined with mortar to form a wall. For this reason, I'm calling these quantitative value-encoding marks *bricks*. When displaying data geospatially, then, we can choose either bubbles or bricks.

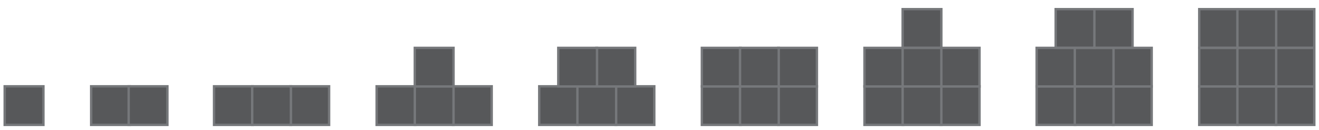
Design Considerations

Let's consider the various ways in which bricks can be designed to determine best practices.

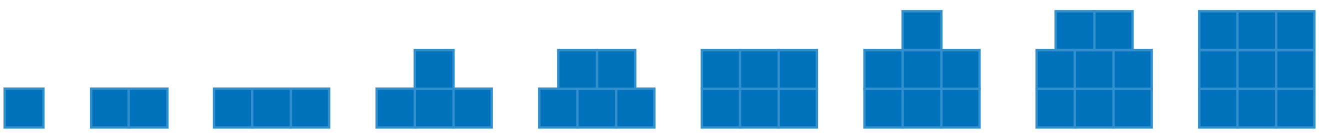
Borders and Fills

When using bricks without color fills as illustrated above, borders around each emphasize the count more than the shape of the overall object. Notice that in the next example, with bricks that are filled with color and lack delineated borders, the overall shape is emphasized, but the count becomes less obvious:

In the next example, I've added slightly lighters borders to the squares, which provides a nice balance between an emphasis on the overall shape and visibility of the count:



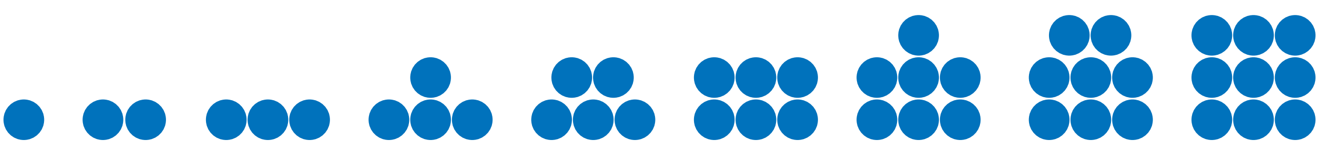
We are certainly not restricted to grayscale colors. Here's an example that uses two shades of blue:



Colors should be chosen to clearly discriminate bricks from the background, such as the geographical background of a map.

Alternate Shapes

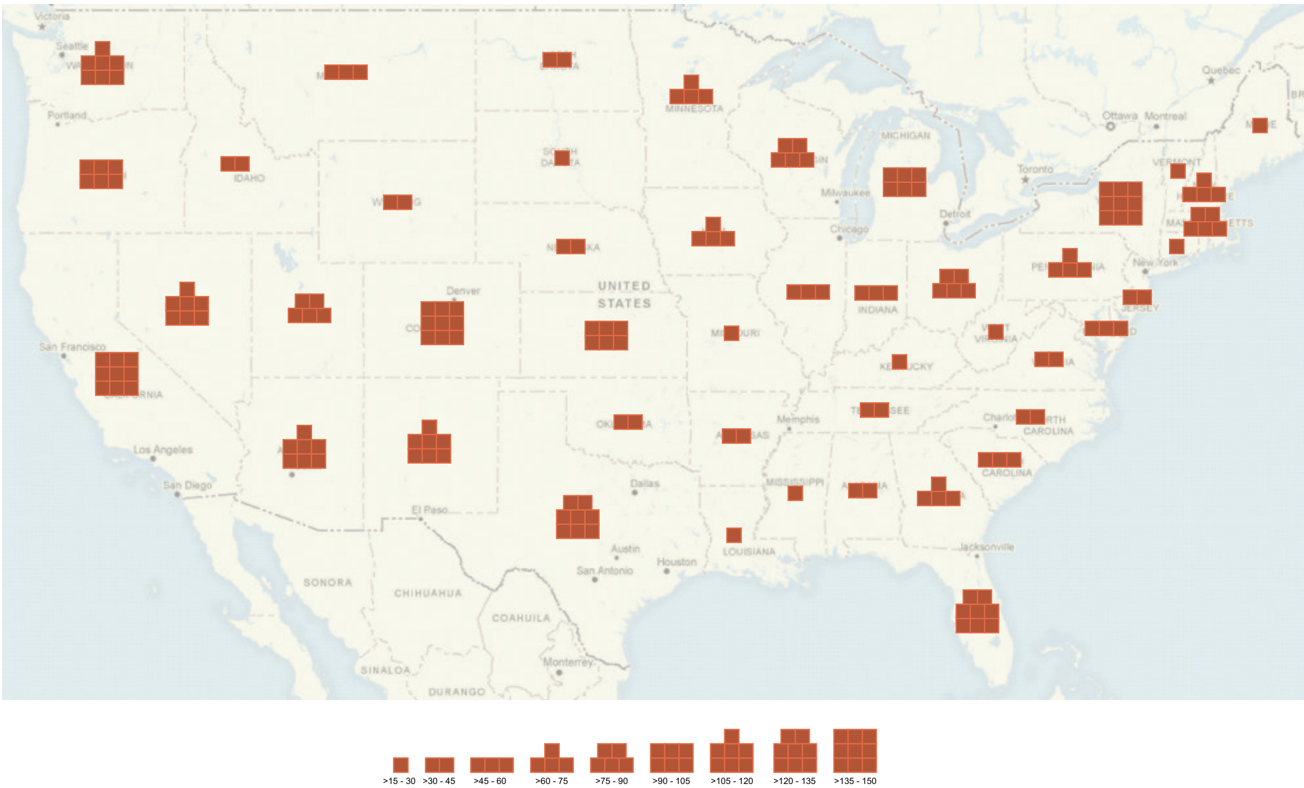
Bricks shaped as squares fit nicely together, but perhaps we don't need to restrict ourselves to this shape. Here is the same set of binned quantitative ranges, this time represented by combinations of circles:



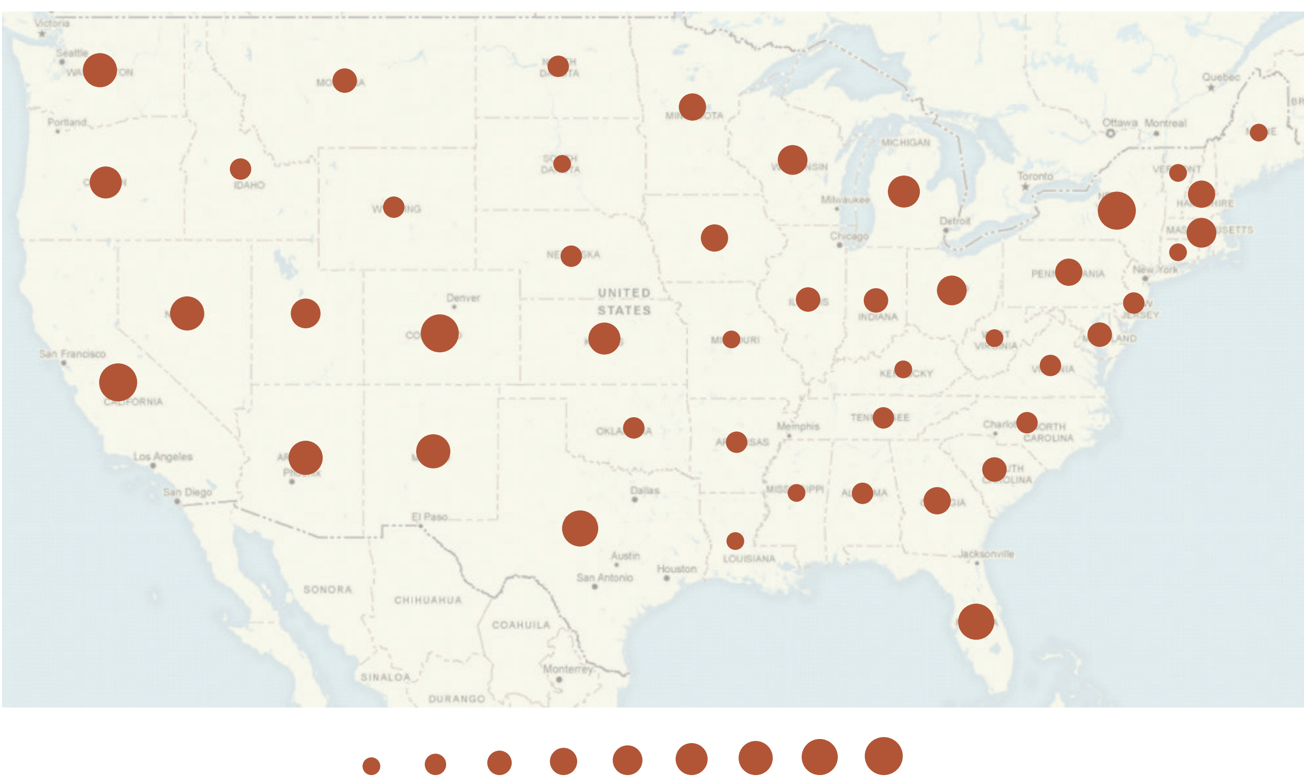
Because circles do not fit together neatly, borders are not needed to delineate them—white space between the circles does the job. The independence of each circle emphasizes the count, but appears to do so to an exaggerated degree. Also, notice the illusory visual artifacts where circles almost touch, which is visually annoying and distracting. It seems wise to stick with squares.

Bricks in Context

Let’s look now at how this works on a map. Imagine that the following example displays sales revenues per state.



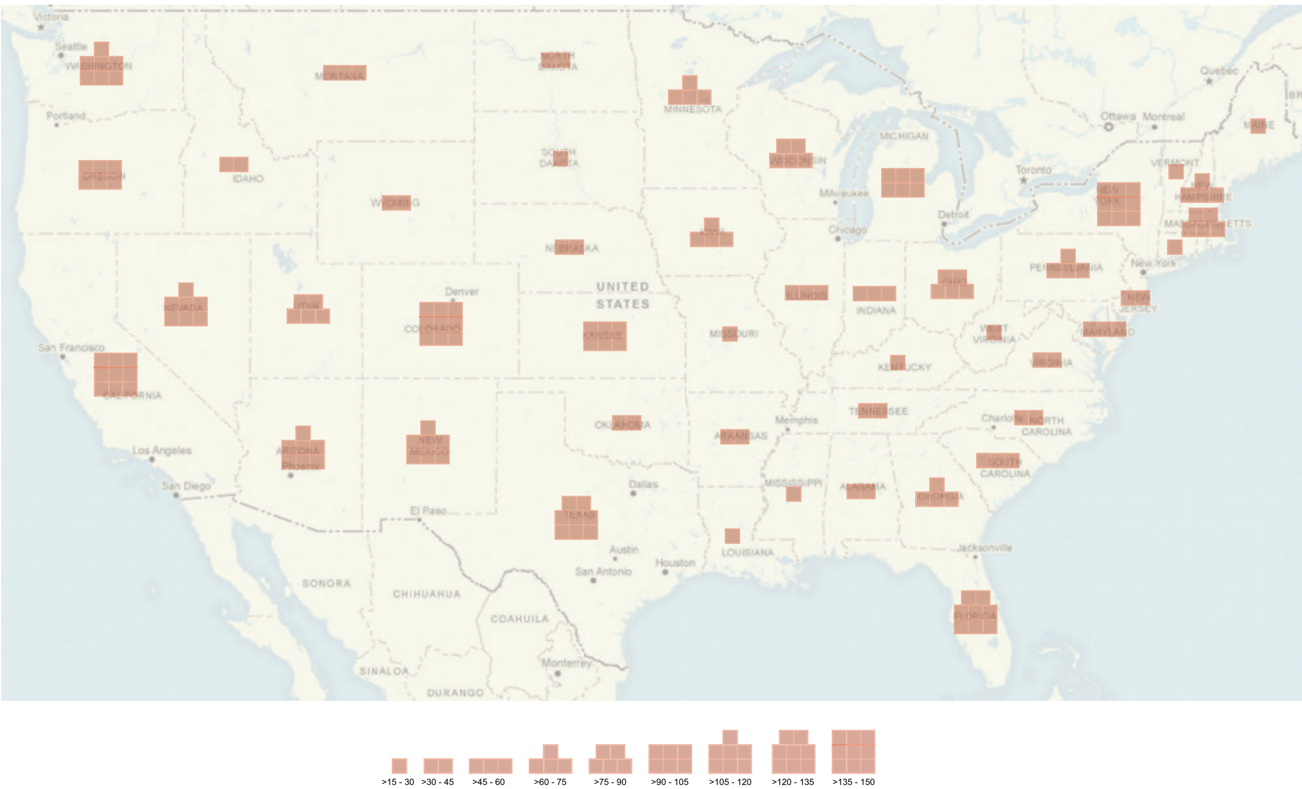
Now compare this bricks display to the conventional method using bubbles of varying sizes below:



Bubbles of graduated sizes proportional to the value ranges that they represent can be difficult to distinguish. Try finding all the bubbles of a particular size. This not only requires a conscious search, but one that is fairly slow. To make the bubbles of various sizes easy to distinguish, we would need to graduate their sizes to a degree that is greater than values that they represent and thus misrepresent them.

Transparency

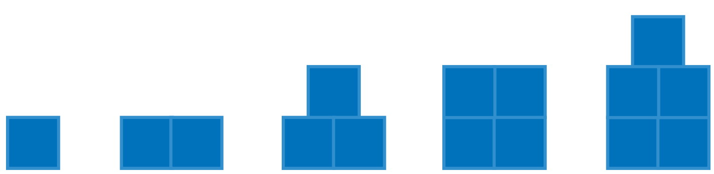
The problem of bricks occluding useful geographical details, such as the names of the states in the examples above, can be handled just as we would when using bubbles, by making the objects transparent. The following example illustrates a transparency level of 50%:



Although it may not be apparent when the map is small, as shown above, at full size the state names are legible when viewed through the data objects.

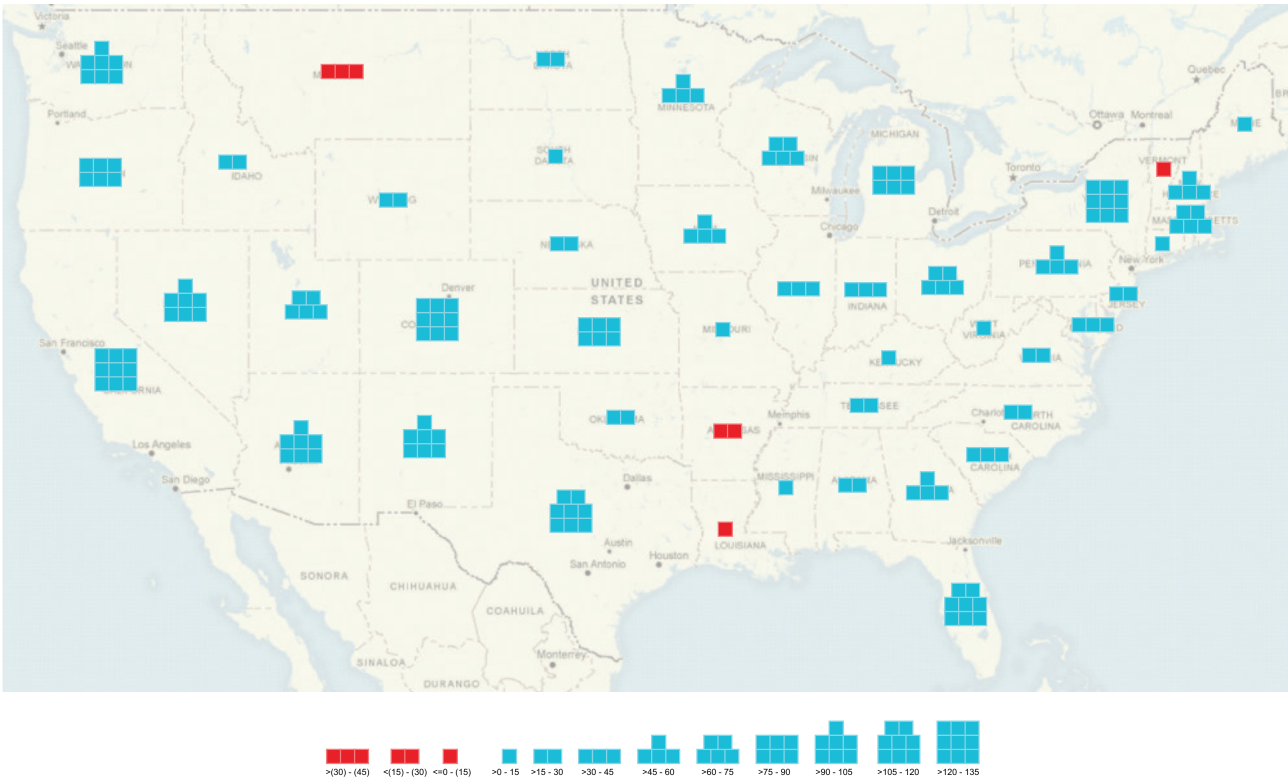
Alternate Numbers of Intervals

In all of the examples so far, I’ve shown ranges that consist of nine quantitative intervals (i.e., bins), but fewer can be used. Here’s a set of bricks that encodes five intervals:



Diverging Scales

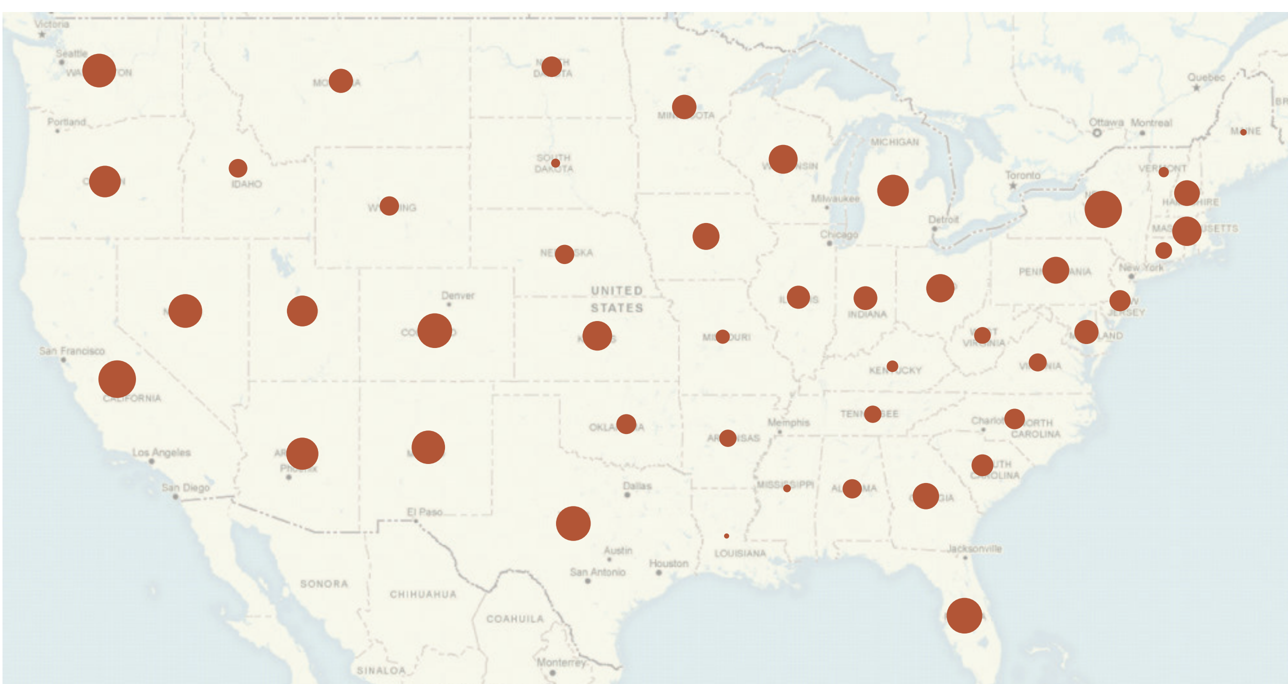
In addition to sequential quantitative scales that run in one direction, bricks can also be used to display diverging scales that run in two directions (e.g., positive and negative values), by using a different color for each direction. In the following example, bricks are used to display profits and losses:



Binned Ranges

Bricks are designed to represent binned ranges of values, not continuous variation. Is this a problem? Slight differences in the sizes of objects cannot be detected when they're positioned on a map, except when near one another. It is actually easier for most analytical purposes to detect quantitative differences when they are binned into ranges and represented by objects that vary as easily detectable visual differences.

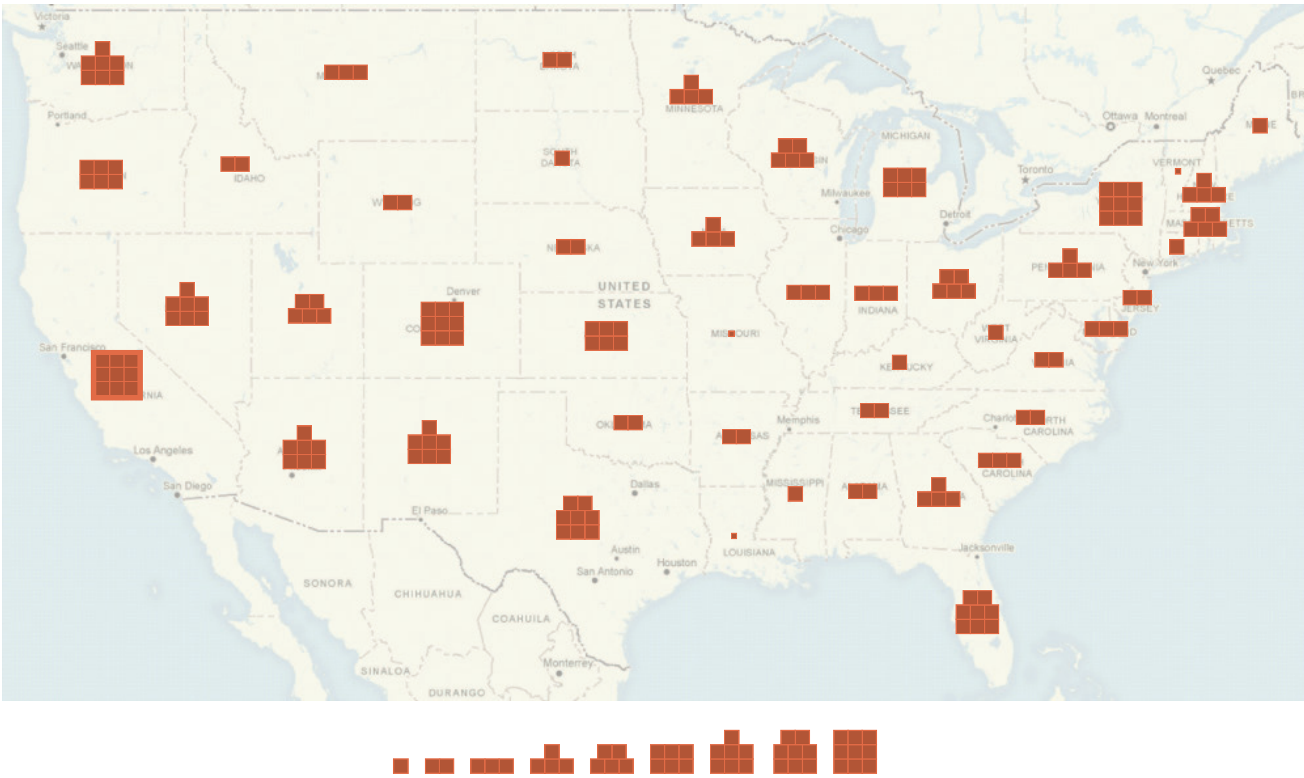
What we do lose when using binned ranges, however, is a full sense of variation within the data set, which can be seen when using bubbles of continuous rather than binned sizes, such as in the example below:



It would be useful, however, to find a way to reveal a greater sense of variation using bricks.

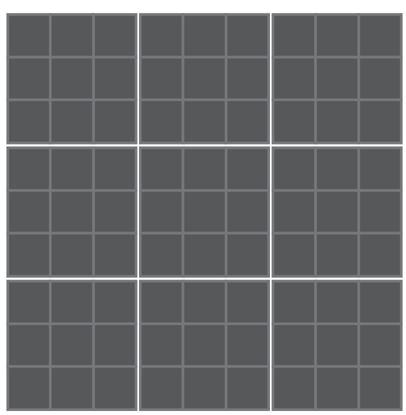
Enhanced View of Variation

In many cases the subtler sense of variation that we need exists primarily at the ends of the range to detect outliers. In this case, bricks can be slightly enhanced to distinctly display outliers. Low outliers can be displayed as sub-bricks—squares that are much smaller than a regular brick—and high outliers can be displayed as by increasing the stroke weight of the border around the maximum brick structure, as illustrated in the following example.

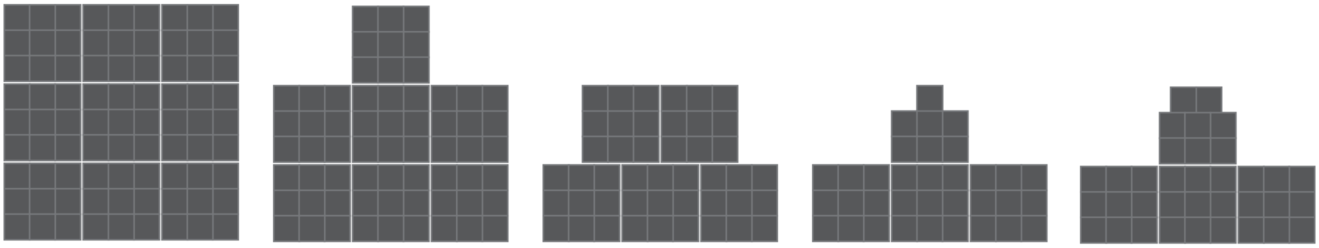


Greater Precision

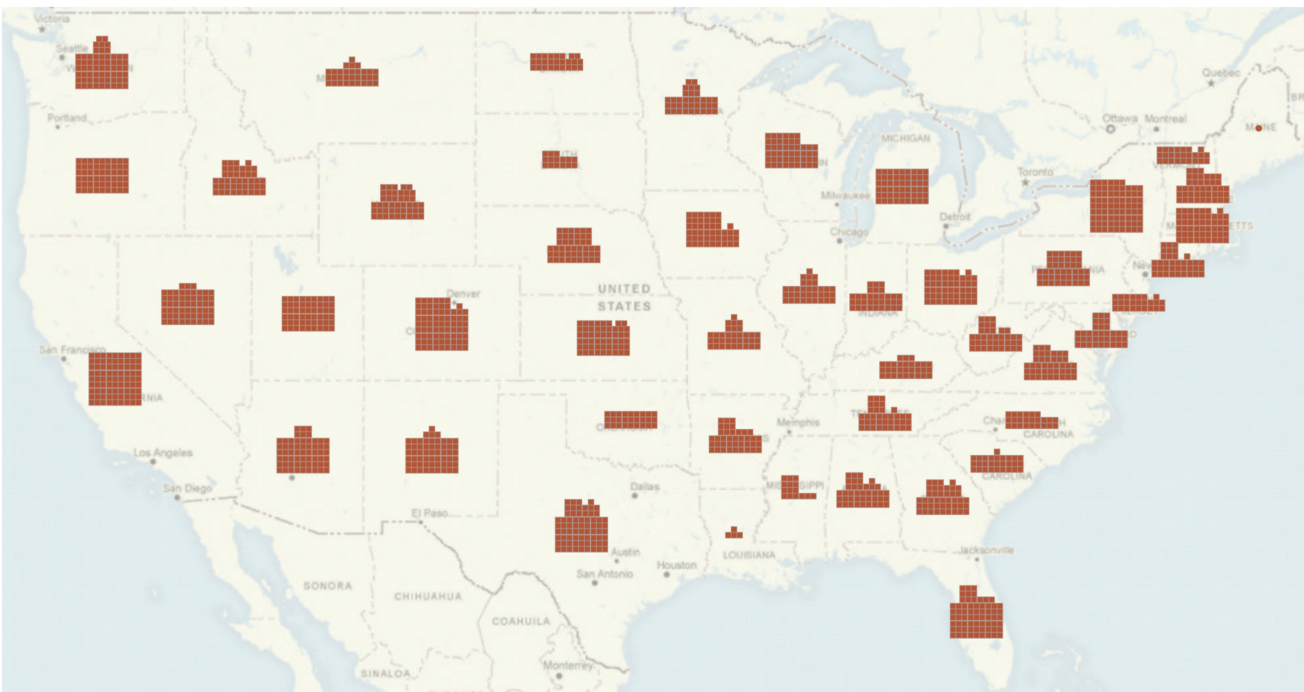
If needed, greater precision within all of the intervals could be conveyed through more complex arrangements of bricks, such as the following, which consists of 81 intervals (one large square subdivided into nine medium-sized squares, each in turn subdivided into nine small squares).



Because slight gaps of white space delineate the medium-sized squares, producing a 3x3 matrix, which subdivides further into 3x3 matrices of small squares, more precise differences in quantity can be seen fairly easily and efficiently. In the following example, even differences of a single square can be detected with ease.



This additional precision comes at a cost in speed of perception, however, especially with geospatial displays consisting of many values. Precise comparisons can't be made preattentively. This cost forces me to ask, "Do we really need greater precision than that which is available through a single 3x3 matrix of squares that can display a quantitative scale of nine intervals?" For visual data exploration and analysis purposes, we rarely do, but by expanding the number of squares we gain more than precision; we also gain a better view of variation. The example below provides a view of variation that is comparable to view that's provided by bubbles that vary in size continuously:

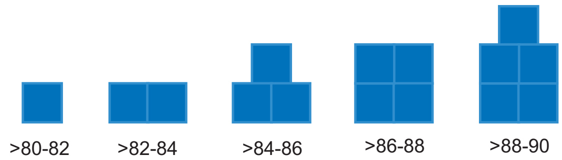


What's the Downside?

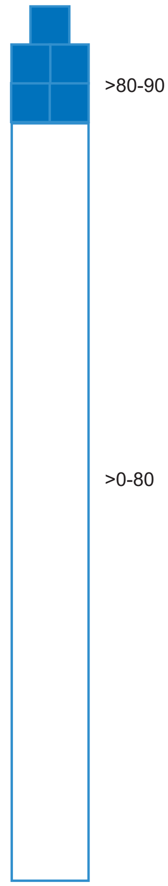
So far, I can think of two uses of bubbles that cannot be applied to bricks:

1. Bubbles can overlap to some degree and still be legible, especially those that consist only of borders without fill colors, but bricks cannot overlap and still work effectively.
2. Bubbles can be used to simultaneously encode two quantitative variables: one based on size and the other based on color intensity. While it might be possible to vary the color intensity of bricks to encode a second variable, it is more difficult to do because the border color must also vary in association with the fill color to keep them clearly visible. A good coloring algorithm could handle this, but it might not be possible to do this in a way that does not complicate rapid comparisons.

Both bubbles and bricks suffer from a problem that is seldom considered. When the quantitative range across which the values are spread begins at zero, it is possible for the sizes of bubbles and bricks to correspond to the relative sizes of the values that they represent. In this case, a bubble or brick that is twice the size of another would represent twice the value. When the quantitative range does not begin with zero—especially when it begins far from zero—this proportional correspondence in size is difficult to support. Imagine a range of values that extends from 80 to 100, encoded by the following set of bricks.



Although a two-brick mark accurately adds an incremental range of values to a one-brick mark, it does not represent a range of values that is twice that of the other even though it is twice the size. To keep the sizes of bricks equal in proportion to the value ranges that they represent (e.g., >0 – 10 displayed as one brick, >10-20 displayed as two bricks, and so on up to >80-90 displayed as nine bricks), all of the values that were shown on the map would be represented by the same nine-brick mark, which wouldn't allow us to see any differences in value between them. Bubbles as well would fail in the same manner. Although this problem isn't new or unique to bricks, it might be helpful to alert readers to it by providing a legend similar to the one that follows for the range of bricks that appears above.



A Final Word

This is only an introduction to bricks, a new and potentially useful way to encode quantitative values in geospatial displays. Time, research, and experience will demonstrate the effectiveness of this method (or lack thereof) and further hone the ways in which it should be designed. I invite you to explore the merits of bricks by experimenting with them, testing them in real-world scenarios, implementing them in products, and extending their design variations however you find useful. Share what you find with me and together we will improve bricks to extend the quality and reach of data visualization.

Try Bricks Yourself

You can experiment a bit with bricks on your own using a Web-based test bench that Hannes Reijner of Panopticon Software created to compare this method of geospatial display to bubbles. The test bench allows you to manipulate several aspects of bricks (size, color, borders, etc.) to explore variations in their design.

Discuss this Article

Share your thoughts about this article by visiting the Building Insight with Bricks thread in our discussion forum.

About the Author

Stephen Few has worked for nearly 30 years as an IT innovator, consultant, and teacher. Today, as Principal of the consultancy Perceptual Edge, Stephen focuses on data visualization for analyzing and communicating quantitative business information. He provides training and consulting services, writes the quarterly *Visual Business Intelligence Newsletter*, and speaks frequently at conferences. He is the author of three books: *Show Me the Numbers: Designing Tables and Graphs to Enlighten*, Second Edition, *Information Dashboard Design: The Effective Visual Communication of Data*, and *Now You See It: Simple Visualization Techniques for Quantitative Analysis*. You can learn more about Stephen's work and access an entire library of articles at www.perceptualedge.com. Between articles, you can read Stephen's thoughts on the industry in his blog.