Best Practices for Data Embroidery

Shemsi Alhaddad

Institutional Effectiveness and Research, University of South Carolina Lancaster, USA;
Alhaddad@email.sc.edu

Abstract

Data embroidery is a form of data art that is growing in popularity within and outside the scientific community. This paper formalizes some aspects of data embroidery that are applied intuitively by those outside the scientific community. After an introduction to data embroidery, the paper describes best-practice guidelines for data embroidery adapted from best-practice guidelines for print and digital data visualization. Next there are notes on embroidery-specific practices to consider when creating data embroidery. Finally, there are instructions for an introductory level data embroidery project that has been classroom tested.

Introduction

Data embroidery is an art form where data determines design elements, like color or motif. Data embroiderers may have one of three goals: to inform others, to journal, or to create a unique art piece [5]. To inform others of bee colony changes, I created a display relating stressors to the number of bee colonies in the U.S. over times. To evoke beehives, I adapted a digital hexagonal tile map of the U.S. [3]. I mounted the fabric in hive frames with bees I handstitched. The display and legend appear in Figure 1.

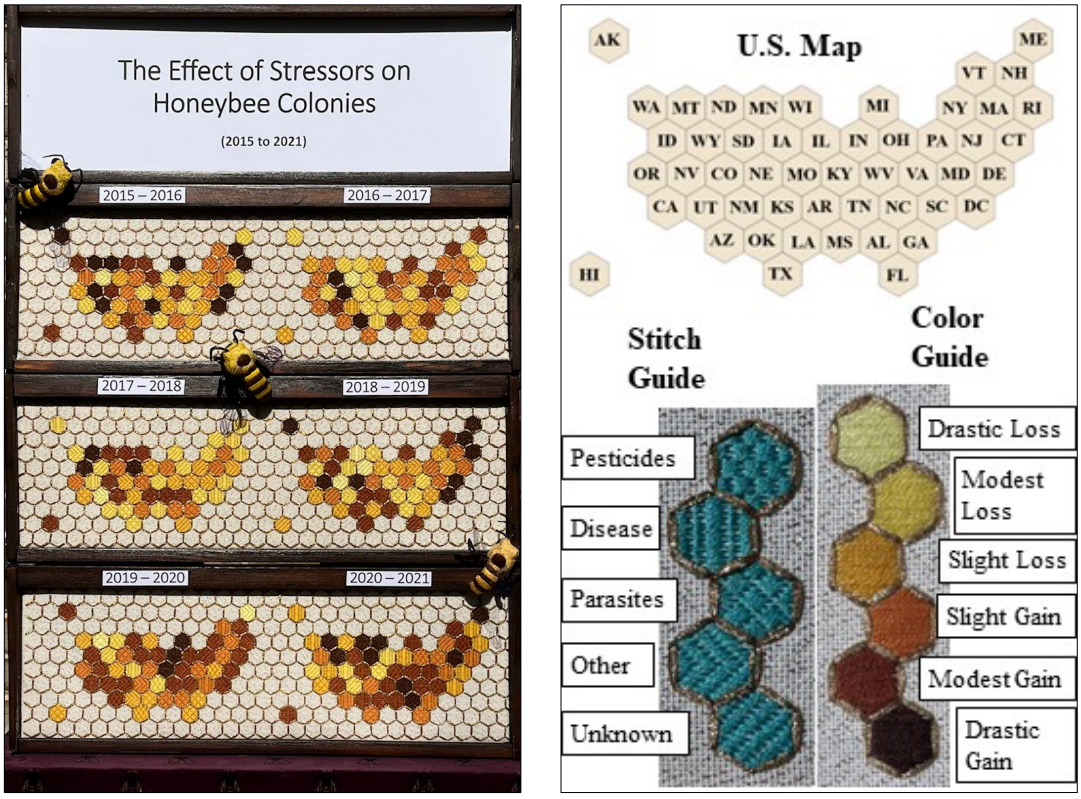


Figure 1: *Embroidered bee colony display with its legend.*

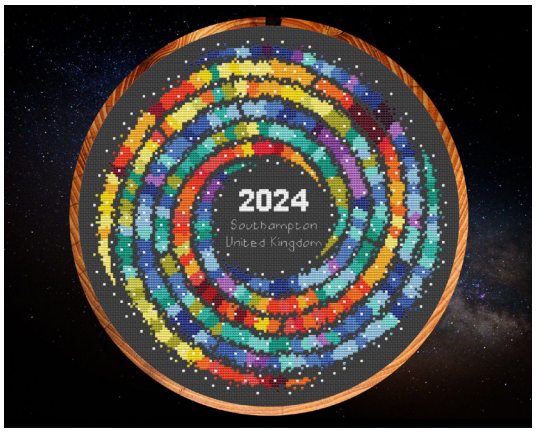
The display can be used to identify patterns in the data. Consider, for example, the closeup of the maps for 2015-2016 and 2016-2017 shown in Figure 2. Viewers can identify patterns such as clusters of stressors in neighboring states and changes in the number of colonies over time.

Figure 2: *Closeup of maps for years 2015-2016 and 2016-2017.*

Some people enjoy data embroidery as a form of journaling, not intended for external viewers. Sophie O'Neill of StirCrazyCrafterUK designs bookmarks with an image of books on shelves, shown in Figure 3(a). Stitchers use her design as an annual reading log, tracking a variable related to reading such as genre or book format. Absent a color key, outside viewers cannot interpret the underlying data. Stitchers can analyze their own data and identify trends in their reading habits [12].

When an artist's purpose is to create a unique piece of data art, the viewer may find pleasure in viewing emerging patterns with no knowledge of the underlying data [1]. Clare Bray of Climbing Goat Designs creates cross stitch patterns whose colors are determined by outdoor temperature [2]. Stitchers partition outdoor temperatures, assigning one color to each interval. The design contains 365 sections (366 for leap years), each stitched using the color of the day's high temperature, as in Figure 3(b). Viewers with no introductory information may enjoy the color clusters and fluctuations without associating them with temperatures.

(a)



(b)

Figure 3: *Stitched examples: (a) a reading log [12], and (b) a temperature embroidery [2].*

Data embroidery best-practice guidelines described in this paper are adapted from recognized print and digital data visualization guidelines. Designers whose purpose is to create a unique piece of art are not expected to use data visualization best-practices, and so they are not considered in this paper. However, designers whose goal is to inform others or themselves should apply relevant guidelines to ensure that the viewer accurately understands the displayed information [5].

Types of Data and Indices

In this paper, variable data is classified as nominal, ordinal, interval, or ratio. *Nominal* data, like type of pet, is named and doesn't have a natural or culturally accepted order. *Ordinal* data, like satisfaction level, is named and has a natural or culturally accepted order. *Interval* data, like temperature, is numerical but zero does not indicate an absence of the variable. *Ratio* data, like height, is numerical with zero indicating an absence of the variable. Data is organized using an *index*, such as the date or geographic location. Indices are used to analyze trends or draw comparisons.

Hand Embroidery

Hand, or manual, embroidery is a traditional art form. Practitioners use a handheld needle to run decorative fibers, such as thread or floss, through fabric to create aesthetically pleasing designs. Common fabrics for embroidery have an even weave. This means the interwoven threads form evenly spaced horizontal and vertical lines, resembling graph paper. Although modern embroidery can also be achieved using a machine, the scope of this paper is limited to hand embroidery.

The colors of embroidery fibers are organized on a value scale, where higher values reflect more light and are referred to as *lighter*. Lower values are called *darker*. Saturated hues have different values. For example, yellow is lighter than red. For a fixed hue of embroidery fiber, the value can be altered by changing the dye composition; adding white (tinting) makes it lighter, while adding black (shading) makes it darker. Embroiderers can also alter the perceived value by changing the density of fiber on the fabric. For example, using fewer strands of fiber on white fabric results in a perceived lighter color [14].

Best-Practice Guidelines

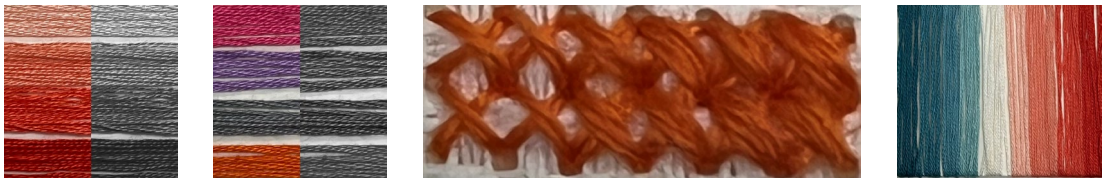
Data visualizations are used to identify trends and patterns rather than specific data values. They display some aspects of the data rather than the entire data set. Designers select a type of visualization based on the information they wish to convey [7]. Designers should be aware of the data visualization guidelines that are applicable to the type of visualization they choose [8]. This section contains adaptations of print and digital best-practice guidelines to embroidery.

Kosslyn's Principle of Appropriate Knowledge suggests that viewers better understand the message of a data visualization when they are familiar with the conventions of that visualization [9]. Some forms of data visualization are ubiquitous, such as maps and bar graphs. To facilitate comprehension, data embroidery designers should base their designs on existing print and digital visualization formats. Using a visualization format that is entirely new may make the information inaccessible to the viewer [4].

Visualizations used for data analysis differ from those intended to convey information to others [4]. In the latter case several data visualization guidelines support the use of a simple design. Kosslyn's Principle of Capacity Limitation indicates that viewers will not understand the message of a visualization if an excessive amount of information must be retained or processed [9]. To mitigate this issue, designers should define their message and create a display using relevant data. Tufte recommends maximizing data ink. This means most of the ink (or embroidery fiber) should be used to represent data [15]. Empirical studies show that viewers' comprehension is improved by removing decorative non-data ink, such as background images, while including some explanatory non-data ink, such as axes and titles [7]. This principle is applied to embroidery by avoiding decorative elements that distract from the data, even if they

enhance the aesthetic beauty of the work. Some decorative elements, such as an embroidered title, can be used to draw attention to the data or enhance the viewer's understanding of the data.

There are several guidelines that affect color choices. Tversky's Apprehension Principle says a data visualization is effective when the viewer can readily accurately understand the conveyed information [16]. Kosslyn's Principle of Discriminability indicates that representations of two different data values should be sufficiently different to be perceived as different [9]. Everyone perceives color differently [11]. To accurately convey stitched information, embroiderers digitally photograph their chosen colors and use a monochrome filter to determine if the shades are distinct. This universal design practice benefits all viewers, including those experiencing color vision deficiencies. Figure 4(a) shows colors with similar hues and different shades, in color and monochrome. Figure 4(b) shows colors with different hues but similar shades, in color and monochrome.



(a) (b) (c) (d)

Figure 4: *Color usage: (a) distinct shades of like hues in color and monochrome, (b) likes shades of distinct hues in color and monochrome, (c) increasing density, and (d) a divergent color scheme.*

Kosslyn's Principle of Compatibility indicates that a data visualization should be compatible with its meaning [9]. For ordered data, people associate lighter colors with smaller data values [7]. Using Figure 2 as an example, lighter colors indicate fewer bee colonies. Stitchers can also alter the density of stitches to indicate changes in data value, as in Figure 4(c). For ordered data that has a central value, one color can be used for data values on each side of the center, with darker shades indicating a greater distance from the center [7]. This type of divergent color scheme is shown in Figure 4(d).

Tufte's Congruence Principle is related to the Principle of Compatibility. For nominal data this principle implies that colors and motifs should evoke the data they represent [15]. Designers should be aware of cultural connotations of colors, such as red representing heat and blue representing cold [6]. Besides color, motifs or stitches can represent values of nominal data. For example, for data related to pets, furry stitches or motifs can represent cats and dogs, feathered stitches or motifs for birds, and scaly stitches or motifs for fish. Figure 5 shows examples I created of these types of stitches and motifs.

Figure 5: *A furry velvet stitch, a feather stitch, and a fish scale motif in short-and-long stitch.*

The Principle of Compatibility also impacts the relative position of motifs in a design. Relative positions of objects in the display should represent relative positions of data values in the real world [9]. If two

objects are close to each other in the display, viewers expect them to be related in the real world [7]. For data indexed in a linear manner, such as over time, motifs should be organized linearly, with appropriate spacing. For data indexed geographically, the relative positions of motifs should match their geographic locations. As an example, the honeybee colony display in Figure 1 shows the US states in geographic order, and the annual maps are shown in chronological order.

Kosslyn's Principle of Informative Changes indicates that people expect changes in properties of representations to carry information [9]. On a basic level, this means one color, motif, or stitch should be assigned to each data value. It further implies that combinations of stitch, motif, and color should be tested for consistent perception when color is used for one variable of data and stitch or motif for a different variable. Changes in texture can affect the perception of color [17]. One way to minimize misapprehension is to use stitches that have the same orientation. Figure 6 shows five needlepoint stitches with a vertical orientation. The Royal School of Needlework's Stitch Bank is an excellent tool for finding appropriate stitches [13].

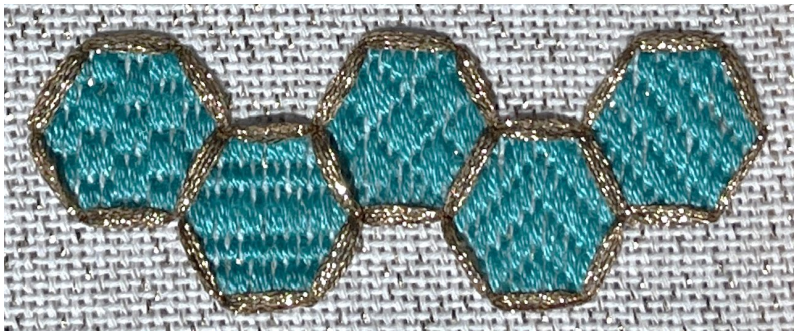


Figure 6: *Five needlepoint stitches with a vertical orientation.*

Tufte also asserted that interval and ratio data should be represented in a quantitative manner that is not misleading [15]. In print and online, bar graphs increase in one dimension, making them more accurately perceived than pictograms that increase in two dimensions. Embroidered visualizations should follow the same principle. A repeated motif is more accurately understood than a change to the area of the motif. For example, cross stitching ×××× to indicate triple the value of one × is easy to understand. However, it is difficult to determine precise ratios when comparing different sizes of a single motif, like × and ✕.

Ideally, the viewer should be able to see the entire data visualization without moving their eyes. Tufte refers to this as maximizing data density [15]. Embroidered work may be a small handheld piece, a large wall hanging, or any size in between. The designer should take into consideration the final presentation of the visualization when developing the design. Tufte asserted that using a repeated image, called a small multiple, is an efficient way to draw comparisons [15]. This can be interpreted in embroidery as a repeated motif whose color varies based on the data values. In Figure 1(a), each of the six US maps is a small multiple. Viewers use these maps to compare annual changes in the number of colonies.

After carefully developing a design, using appropriate guidelines, designers should be aware that their intuition or expertise may not be enough to create an effective visualization [7]. The designer should be mindful of their intended audience's perspective. Cognitive scientists use recognized scientific techniques to determine the effectiveness of data visualization. These include determining the speed with which subjects interpret displayed data, determining the accuracy of their interpretations, and using technology such as eye trackers to determine how subjects examine displays [6][7]. Although data embroidery is typically designed in a less rigorous manner, data embroidery designers attempt to reduce misunderstanding by using online and in-person social groups to test their designs [2][12].

Additional Considerations for Hand Embroidery

In addition to the guidelines listed in the previous section, the following guidelines are unique to data embroidery.

- Hand embroidery is a slow process. It is appropriate for visualizing historical data such as global temperatures, or slowly produced data such as books read over the course of a year. However, it isn't appropriate for visualizing time-sensitive data.
- The fabric used for data embroidery should be a firm even-weave, such as Aida cloth. This type of fabric keeps its shape and prevents skewing the shape or size of motifs.
- Different dye lots of fibers may have color variations. Designers should purchase sufficient fibers from one dye lot at the beginning of the project.
- Some fibers are only available in a limited number of colors. Six-stranded floss is commonly used for data embroidery because it is available in a wide range of colors. This makes it easy to create monochrome scales for ordinal data, and to find congruent colors for nominal data.

Applying Best-Practices to a Small Project

I use the following cross stitch project as optional extra credit in an introductory statistics course for high school and college students. Students define one numerical variable that can be analyzed on a ratio scale with a maximum expected value of ten. They then define two other variables, of any type, that are related to the numerical variable. Many of my past students analyzed how their sleep was affected by pairs of variables like stress level, work schedules, ambient light, and noise. They defined a unit of the numerical variable to be one hour of sleep. Some other students compared their study time to their stress and energy levels. They defined one unit of the numerical variable to be one half hour of study time. Students record seven days of data in a table and then create a cross stitched bookmark.

As an example, I recorded the number of steps I walked each day, rounded to the nearest thousand. I also recorded the outdoor temperature and primary weather condition. I analyzed the number of steps on a ratio scale, with one unit representing 1000 steps. I treated temperature as an ordinal variable, using the following temperature groups: 77°F and less; 78-82°F; 82-85°F; 86-89°F; 90°F and higher. I used shades of peach to represent the temperatures, associating lighter colors with lower temperatures. I treated weather conditions as a nominal variable. During that week, the days were sunny, cloudy, and rainy, and I used the colors yellow, grey, and blue, respectively. I organized my data in Table 1.

Table 1: *Data used to create a bookmark.*

Day	**Temperature (Interior Color)**	**Weather Condition (Outer Color)**	**Number of Steps (Number of Interior Stitches)**
Monday	85°F (medium peach)	Sunny (yellow)	9,000 (9 stitches)
Tuesday	81°F (light peach)	Sunny (yellow)	10,000 (10 stitches)
Wednesday	78°F (lightest peach)	Cloudy (grey)	12,000 (12 stitches)
Thursday	82°F (light peach)	Rainy (blue)	4,000 (4 stitches)
Friday	85°F (medium peach)	Sunny (yellow)	8,000 (8 stitches)
Saturday	76°F (lightest peach)	Cloudy (grey)	14,000 (14 stitches)
Sunday	92°F (darkest peach)	Sunny (yellow)	6,000 (6 stitches)

The data is used to cross stitch a bookmark containing seven rectangles that measure four by seven stitches. Each rectangle shows one day of data. The ten interior stitches of a rectangle display the value of the ratio-scale variable. The color of these interior stitches is determined by one of the related variables. The border color for each rectangle is determined by the second related variable. For the data organized in Table 1, the final embroidered bookmark is shown in Figure 7. The leftmost rectangle represents Monday. On that day, nine medium peach stitches represent 9000 steps walked in 82- to 85-degree weather. The yellow border of that rectangle indicates it was a sunny day. On days when the number of interior stitches exceeded ten, I stitched the excess outside of the rectangle to emphasize the overflow.

Figure 7: *Cross-stitched bookmark using the data in Table 1*

I selected cross stitch for this project because it is one of the most common types of embroidery. In a recent survey of embroiderers, 88.5% of 2698 respondents stated that they practice cross stitch, a larger proportion than any other category. Of the cross stitchers, 63.6% identified as Advanced. By comparison, in each of the other categories of embroidery less than half selected Advanced [10]. Another advantage of cross stitch is that the supplies are inexpensive and widely available. Note that cross stitch requires fine motor skills and may not be accessible to some students. If assigning this project, it is best to offer an alternative assignment such as a short research paper or oral presentation.

The bookmark project follows the guidelines described in this paper. Each of the seven rectangles is considered a small multiple. The design resembles a bar chart, following the Principle of Appropriate Knowledge. The simple design arranges the numerical variable into two sets of five, allowing viewers to subitize the quantity; this follows the Principle of Capacity Limitation. Every stitch provides information related to the data, maximizing the data ink. Students apply the Principle of Compatibility by using one × for each value of the ratio-scale variable. Following this principle, students keep the linear order of the time index by stitching the first day's data on one end and using adjacent rectangles for consecutive days. They use both the Principle of Compatibility and the Principle of Discriminability when selecting colors. For ordinal data they use a monochrome scale with distinct color values. For nominal data they use distinct, culturally representative colors. The data density is maximized on the finished product, as the stitched area contains data with very little white space.

Summary and Conclusions

Data embroidery is an emerging style of data art that is gaining popularity through online and in-person embroidery groups. Many designers, like those described in the introduction, provide clear instructions using applicable guidelines described in this paper. Stitchers have a wide range of mathematical abilities. Data embroidery allows these diverse groups to enhance their mathematical skills in a familiar, relaxed setting. During one international temperature stitch-along, I witnessed stitchers who self-identified as having weak math skills comfortably discuss these guidelines—without being aware of the formal guidelines. They discussed different approaches to partitioning their local temperatures in order to maximize color variation in their finished piece. They discussed color selection for the ordinal temperature data, all selecting hot colors for higher temperatures and cold colors for lower ones. They used darker shades of color to indicate more extreme temperatures, and lighter shades near the center of

the temperature range. Some chose to add weather condition as a nominal variable. They chose accessories that evoked the weather condition, such as metallic thread for lightning storms, and white beads for snow. They also discussed the advantages and disadvantages of using the local high, low, or average temperature each day, correctly noting that the average temperatures give a narrow range of values. Most opted to use the high. Some altered the design by creating a double motif for each day, one part for the high and one for the low temperature.

Mathematical educators can take advantage of the popularity of data embroidery by creating projects that engage the population in data visualization. Engaging stitchers through a medium they are familiar with is likely to lower inhibitions. Mathematicians can lead discussions with stitchers related to best-practices guidelines. Stitchers may be more comfortable discussing design elements related to embroidery, their area of expertise, than discussing computer generated graphs in a mathematics classroom. Using a familiar medium to engage them with mathematical thinking gives them a personal connection to mathematics that they may not attain in a typical classroom.

References

[1] I. Biederman and E. Vessel. "Perceptual Pleasure and the Brain." *American Scientist,* vol. 94, no. 3, 2006, pp. 247-253.

[2] C. Bray. Climbing Goat Designs. *Rainbow Temperature Galaxy Cross Stitch Pattern.* 2024. https://www.climbinggoat.co.uk/products/rainbow-temperature-galaxy-cross-stitch-pattern.

[3] D. DeBelius. *Let's Tesselate: Hexagons for Tile Grid Maps*. NPR. 2015. https://blog.apps.npr.org/2015/05/11/hex-tile-maps.html

[4] B. Dutrow. "Visual Communication: Do You See What I See?" *Elements*, vol. 2, no. 3, 2007, pp. 119-126.

[5] S. Few. *Data Art vs. Data Visualization: Why Does a Distinction Matter?* Visual Business Intelligence. 2012. https://www.perceptualedge.com/blog/?p=1245

[6] J. Harsh, M. Campillo, C. Murray, C. Myers, J. Nguyen, and A. Maltese. "'Seeing' Data like an Expert: An Eye-Tracking Study Using Graphical Data Representation." *CBE Life Sciences Education,* vol. 18, no. 3, 2019. https://pubmed.ncbi.nlm.nih.gov/31397653/

[7] M. Hegarty. "The Cognitive Science of Visual-Spatial Displays: Implications for Design." *Topics in Cognitive Science,* vol. 3, no. 3, 2011, pp. 446-474.

[8] C. Kelleher and T. Wagener. "Ten Guidelines for Effective Data Visualization in Scientific Publications." *Environmental Modelling & Software,* vol. 26, no. 6, 2011, pp. 822-827.

[9] S. Kosslyn. *Graph Design for the Eye and Mind.* New York: Oxford University Press, 2006.

[10] J. Miller and H. Drake. IDEAS xLab. *2021 EGA National Membership Survey Summary Report.* Embroiderers Guild of America. 2021. https://egausa.org/app/uploads/2022/01/EGA_Survey-demographics-summary-1-2022.pdf.

[11] National Eye Institute. *Color Blindness.* 2019. https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness.

[12] S. O'Neill. StirCrazyCrafterUK. *The Stir-Crazy Crafter*. 2024. https://www.etsy.com/shop/stircrazycrafteruk/

[13] Royal School of Needlework. *RSN Stitch Bank.* 2024. https://rsnstitchbank.org/.

[14] M. Shipp. *Color for Embroidery.* Stitches by Shipp, 1997.

[15] E. Tufte. *The Visual Display of Quantitative Information*. Graphics Press, 1983.

[16] B. Tversky, J. Morrison, and M. Betrancourt. "Animation: Can It Facilitate?" *International Journal of Human-Computer Studies,* vol. 57, no. 4, 2002, pp. 247-262.

[17] J. Xin, H. Shen, and C. Lam. "Investigation of texture effect on visual colour difference evaluation." *Color Research and Application*, vol. 30, no. 5, 2005, pp. 341-347.