

B+

Matt Petrie  
Ethics overview  
4/29/2009

You have a few rough spots, but I like several of your ideas for enhancing the quality of the out-of-wedlock data.

This situation deals with the representation of data on teenage out-of-wedlock pregnancy in the US for 1993. The data is assumed to be of a larger set for all pregnancies in the country. The focus, however, is choosing how to represent this data in a map. The decision is between mathematical scaling of the visualization, which is accurate to scale, and arbitrary scaling, which is inaccurate but designed to be more identifiable by the viewer. Two ethical issues are central to this choice: 1) is the use of 1993 data from the largest 50 cities representative of the current issue, and 2) does the use of either scaling method bias the data or potentially mislead the average non-expert viewer? The best representation of the data is important; the researcher, journalist, and publisher are bound to accurate and correct presentation of the issue, and a good deal of this lies in map formulation. In addition, the value of journalistic integrity demands that all players maintain high standards for their products or be reprimanded by an informed public or, in this case, the original data publisher.

A number of parties have an interest in the product or could be affected by it. The New York Times and its shareholders could be accountable for a poor representation, as could the researcher and journalist. Additionally, other news agencies who pick up the story have an interest in its quality. The publishing government agency or NGO has an interest in the use of their data. Other government groups may wish to use the study for funding allocation. Local governments could be implicated for having a problem with unwed teenage pregnancy and their constituents may be influenced by the article, whether they read it or not. Even future unwed couples could find value in the study. Across all of these groups, the way the data is represented, the information presented or left out, and explanation will have consequences for how the information is processed and shapes perceptions and actions regarding unwed teenage pregnancy. A poor map, even when presenting sound data, could have

undesirable consequences.

Choices beyond map scaling could be problematic regardless of the chosen technique. First, data from 1993 on a current social issue is not relevant unless it is included in context. I don't believe this was the intention of the researcher and, therefore, the map would be inaccurate because it would represent a current issue using outdated information that could be misinterpreted by a casual viewer. Also, the use of only 50 cities is arbitrary and produces bias towards viewing the issue as both based in cities and spatially located in higher-density regions of the US. The included map realizations do not explain these features and are ethically unacceptable.

The proposed use of arbitrary scaling is also problematic because, without proper explanation, the map could both over and under-emphasize the degree of unwed teenage pregnancies across cities in the US. This problem is increased by the newsprint and internet format; readers are less likely to read the entire article or even a map caption before being influenced by the map. Even a reader of the article may not discern the intricacies of the map, and including them may reduce the effectiveness of the article.

A number of things could be done to improve map quality, but the overall goal is to represent unwed teenage pregnancy in its most current and reasonable format by minimizing potential misunderstandings conveying the scope and limitations of the data within the framework of the article. The map needs to be explained so it cannot be misunderstood by a casual observer. This may require an in-depth heading, not having a heading at all, or including the map within the context of the article and off the front page. This map should not be used to "sell" the article. The 1993 data is unacceptable as well. Also, the use of this article in a non-expert setting makes arbitrary scaling a poor choice because it is easily misleading, even when explained in context. Mathematical scaling is a better choice, but it still suffers from sampling issues and spatial distribution of cities. If cities are the focus, this must be

What about perceptual scaling?

okay for areas, but not points

explicit. Even better would be to use more recent county data and create a choropleth map to explain the distribution, although this might conflict with the article scope. Regardless of the needs of the article, the overriding framework of the map should be to accurately describe the data and minimize problematic misinterpretations.

Choosing the best format for the map is an easy process. Collusion between the author and researcher, both experts in their field, would frame the needs and limitations of the map. As a second step, having a number of non-experts, such as other employees of the Times, explain the map in its different forms before and after reading the article could be a good way to judge its effectiveness. In this case, I believe the process of analyzing map effectiveness by using personal experience and honest, expert knowledge is enough to produce a useful final result and provide a rationale for why specific details were included or not. While the issue is more than “mathematical versus arbitrary,” other needs can be addressed through the same framework of identifying needs, problems, and the best way to promote acceptable interpretation.

intensity  
idea