Institute of Education Sciences
National Center for Education Statistics

NATIONAL INSTITUTE OF STATISTICAL SCIENCES
TECHNICAL EXPERT PANEL REPORT

COMPENDIUM ON GOOD PRACTICES FOR GRAPHICS AND MAPS

TABLE OF CONTENTS

NATIONAL INSTITUTE OF STATISTICAL SCIENCES

COMPENDIUM ON GOOD PRACTICES FOR GRAPHICS AND MAPS

EXECUTIVE SUMMARY

This compendium is an outgrowth of the Technical Expert Panel on Maps and Graphics organized in 2009 by the National Institute of Statistical Sciences (NISS) on behalf of the National Center for Education Statistics (NCES).

The purpose of the panel was to "assist the NCES [in] reviewing and revising, if necessary, maps, graphics and graphical displays in its publications and on its web site."

This document incorporates the findings and recommendations of the Technical Expert Panel, findings from a contemporaneous NISS study of NCES' methodology, reporting, and projections of education statistics[1]. and material from the section on features and feature extraction (§3) of a review undertaken by NISS in 2011[2]. This compendium includes details leading to defined good practices with extensive examples to illustrate.

NCES' current approach to graphics and maps in its hard-copy (downloadable or not) publications is very conservative. There are sound reasons for this:

- Consistency among publications,
- Heterogeneous readership,
- Cost.

However, the price of this conservatism is substantial: publications are sometimes both less informative than they should be and less exciting than they could be.

Good Practice: Always answer this question explicitly: *"Why is this information being presented graphically?"*

GRAPHICS

Good Practices

1. Employ horizontal bar charts as the default, especially when values are displayed adjacent to bars. Consider alternatives to tick marks.
2. Make clear which values in a graphic are derived from others. Be explicit about what items labeled "Total" are totals of.
3. Do not impose unnecessary constraints that obscure small data values in graphics. Consider selective use of nonlinear scales.
4. Maximize the content of graphics, measured using even simple metrics such as number of reported values per square inch occupied.

---

[1] Hussar and Bailey (2008)

[2] Karr et al., (2011). "Statistical Statements and Underlying Justifications in NCES Annual Reports."

5. In general, do not present discretely indexed data as if the index were continuous.
6. Especially when there is a single non-black color for a publication, it should be used consistently.
7. Unambiguously distinguish projected data values from actual data values.
8. Investigate new visualization methods.

MAPS

Good Practices

1. Use color as the preferred means of encoding numerical information in maps, paying attention to the need for grayscale reproduction.
2. Provide easy access to all data values underlying maps, either on the map itself or in associated tables.
3. When doing so is meaningful, include statistical significance in maps.

ALTERNATIVES TO TABLES

Some tables in NCES publications would be more effective as graphics or maps, provided that access to the data values is maintained.

Good Practice

1. Consider increased reliance on graphics and maps as substitutes for or complements to tables, but not to the point that data values are suppressed entirely.

UNCERTAINTY

Good Practices

1. Include uncertainties in graphics on a selective basis, especially when the "main message" is not diluted and the method used to encode uncertainty is well-established.
2. Pay continuing attention to ongoing research, as well as any broadly accepted practices that emerge.

INTERACTIVITY

Good Practice

1. Interactive sorting and linked views are mature technologies that NCES can employ immediately. Techniques, visual metaphors and software for manipulation of map break- points are still evolving. Giving them an opportunity to "crystallize" before adopting them seems prudent. Always ensure that interactive graphics and maps have a reset functionality.

PREFACE

The National Center for Education Statistics (NCES) commissioned the National Institute of Statistical Sciences (NISS) to assemble a panel of experts to provide guidance to NCES for the review and, where appropriate, revision of the graphics and maps and graphical displays on its website and in its publications. The perspective was to be user-centric; elicitation of user views was considered but not carried out. On 6-7 April, 2009 the Technical Expert Panel met in Washington, DC. The panel determined that the value of a Compendium relied on the accessibility of the information and adequate exemplars. Panel communications continued thereafter as examples were compiled and the report with clear, non-technical statements of good practices were drafted.

GOOD PRACTICES FOR GRAPHICS AND MAPS

I. INTRODUCTION

In general, the current NCES' approach to graphics and maps in its hard-copy (downloadable or not) publications is very conservative. There are sound reasons for this:

- Consistency with previous versions of the publications;
- The heterogeneous readership of the publications, and in particular the varying level of statistical—more generally, quantitative—sophistication among readers;
- Cost considerations, as evidenced by reluctance to employ multiple (non-black) ink colors in a single publication.

The price of this conservatism, however, is substantial: publications are sometimes both less informative than they should be and less exciting than they could be.

For some graphics in NCES publications, there does not seem to be a clear answer to the fundamental question of

"Why is this information being presented graphically?"

The question matters: in almost all cases, the data could instead be presented, sometimes more completely, in tabular form. Among common responses to "Why . . . graphically?" are:

Highlighting qualitative characteristics of data, for instance, that there is increasing trend over time, or that nearby states are similar with respect to some characteristic.

Facilitating comparisons among numerical values. Two values in a table are easy to compare. Three are not, and identification of similar elements - "clustering" - is nearly impossible.[3]

Extracting features from (especially, large or complex) data. Features include trend in single variables and relationships among variables (Karr et al., 2011).

Good Practice: Always answer this question explicitly.

II. GRAPHICS

Because it is intentionally concrete, this section contains only limited mention of the important but often somewhat philosophical tenets of Edward Tufte (Tufte, 1983, 1990, 1997). Wilkinson's Grammar of Graphics (Wilkinson, 2005) does underlie many of the comments.

[3] This may not be so in settings that allow tables to be sorted interactively.

2.1 Principal Items

Bar Charts. One significant improvement to NCES publications would be to replace vertical bar charts by horizontal bar charts. To illustrate, consider Appendix Figure A-1,[4] which is inefficient because the width of the bars is constrained by the need to put numbers above them. Consider instead the horizontal version in Figure 1. By comparison with Appendix Figure A-1, in Figure 1,

1. The expanded physical scale makes comparisons easier.
2. The horizontal layout reveals more about small values. See also discussion of **Disparate Values** below.
3. The horizontal layout includes both actual values and percentage changes from one time period to the next.
4. Year labels are explicit.

Note that Figure 1 contains low-key vertical lines, beneath all other graphical elements, that carry the numerical scale associated with the x -axis through the entire chart. These are much easier to follow than the tick marks in Appendix Figure A-1.

> *Good Practice*: Employ horizontal bar charts as the default, especially when values are displayed adjacent to bars. Consider alternatives to tick marks.

Totals. A widespread issue in NCES publications is the treatment of totals. In Figure 1, the "Total" bars are the sums of the "PK–8" and "9–12" bars. At some level, this is perfectly clear,

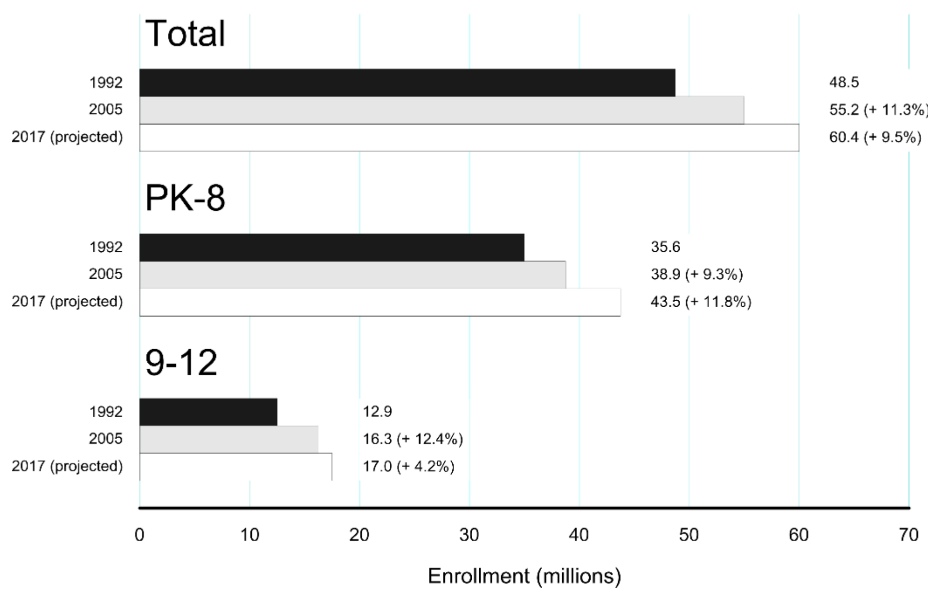


Figure 1: Alternative version of Appendix Figure A-1 in the form of a horizontal bar chart.

[4] This figure, as are all figures in the Appendix, is taken from Hussar and Bailey (2008), which was the basis of the comments and proposals in Karr (2009). It is not singled out here for criticism, but used only for illustrative purposes.

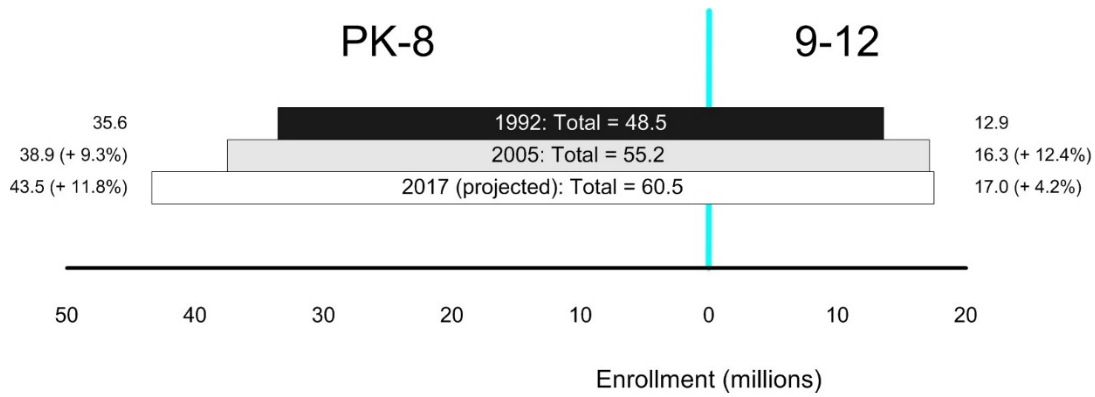


Figure 2: Alternative version of Appendix Figure A-1 in the form of a bi-directional horizontal bar chart.

but at the same time, the practice violates the principle that there be unambiguous indication when some elements of a graphic are derived from others. Figure 2 removes this problem by means of a bi-directional, sometimes called "back-to-back," bar chart. There, PK–8 enrollments are to the left of the light gray vertical line at "Enrollment = 0" and 9–12 enrollments to the right. Figure 2 makes explicit that "Total" is the sum of PK–8 and 9–12. It conveys the same information in approximately one-third less space. Also, it improves comparisons between PK–8 and 9–12. For instance, it is obvious in Figure 2 that the rate of PK–8 growth is increasing, but the rate of 9–12 growth is decreasing.

On the other hand, the capability for direct graphical comparisons between totals is reduced in Figure 2 as compared to Figure 1. For instance, it is apparent from the latter but not the former that the rate of growth of total enrollment is decreasing.

Graphics such as that in Figure 2 should not be employed in cases where the total of the "left" and "right" sides makes no sense (e.g., when one side is students and the other isteachers).

> *Good Practice*: Make clear which values in a graphic are derived from others. Be explicit about what items labeled "Total" are totals of.

Disparate Values. In some cases, NCES publications do not deal effectively with disparate numerical values. As a result, information about small values may be attenuated.

In some instances, the rationale appears to be to preserve a common scale across multiple panels in a figure, as in Appendix Figure A-2. The horizontal bar chart in Figure 3 removes the problem. While this alternative is problematic in other senses - there may be too much information for some users - it makes small values much more visible. Moreover, it permits comparisons - for instance, males to 18–24 year-olds - that are impossible in Appendix Figure A-2.

NCES' avoidance of nonlinear scale transformations - especially logarithms - is understandable, but may be excessively dogmatic. Logarithmic scales are ubiquitous in the scientific literature, and do appear in graphics in non-scientific publications.

> *Good Practice*: Do not impose unnecessary constraints that obscure small data values in graphics. Consider selective use of nonlinear scales.

Low Content Graphics. Some NCES publications contain figures that have relatively low content. Appendix Figure A-3, for instance, consumes approximately 10% of a page in order to display three values. The alternative in Figure 4 presents three times as much information: the coordinates of the endpoint of each line are the (Teacher, Pupil) numbers, and the slope of each line is the Pupil/Teacher ratio. The increasing rate of decline in the ratio is evident from the concavity in Figure 4, but hard to discern in Appendix Figure A-3. Figure 4 has deficiencies of its own, especially the skewed aspect ratio, but these do not interfere with its ability to communicate multiple pieces of information. In Figure 4, the two numerical values are encoded as lengths, and their ratio is encoded - mathematically consistently - as a slope.

Good Practice: Maximize the content of graphics, measured using even simple metrics such as number of reported values per square inch occupied.

Continuous vs. Discrete. The case for "continuous graphs" of discrete-time-indexed data (for example, annual data) is not always compelling. Figure 5 contains more information than either of Appendix Figures A-6 and A-7 - in fact more information than the two together - and at the same time it conveys at least as much visual gestalt. It also avoids any misimpression that the data are collected continuously in time. Trends are just as apparent as they would be with continuous lines.

Conversion of discrete to continuous creates other issues as well. In some NCES publications, this conversion is effected using piecewise linear functions, which do help convey changes in trend. In others, however, literal smoothing is involved, and while the details may be of interest to only a small number of readers, they may not be available in the publications - a lack of transparency that exposes NCES to external criticism. Excessive smoothing may also, of course, attenuate meaningful anomalies. For related discussion, see Karr et al. (2011).

Good Practice: In general, do not present discretely indexed data as if the index were continuous.

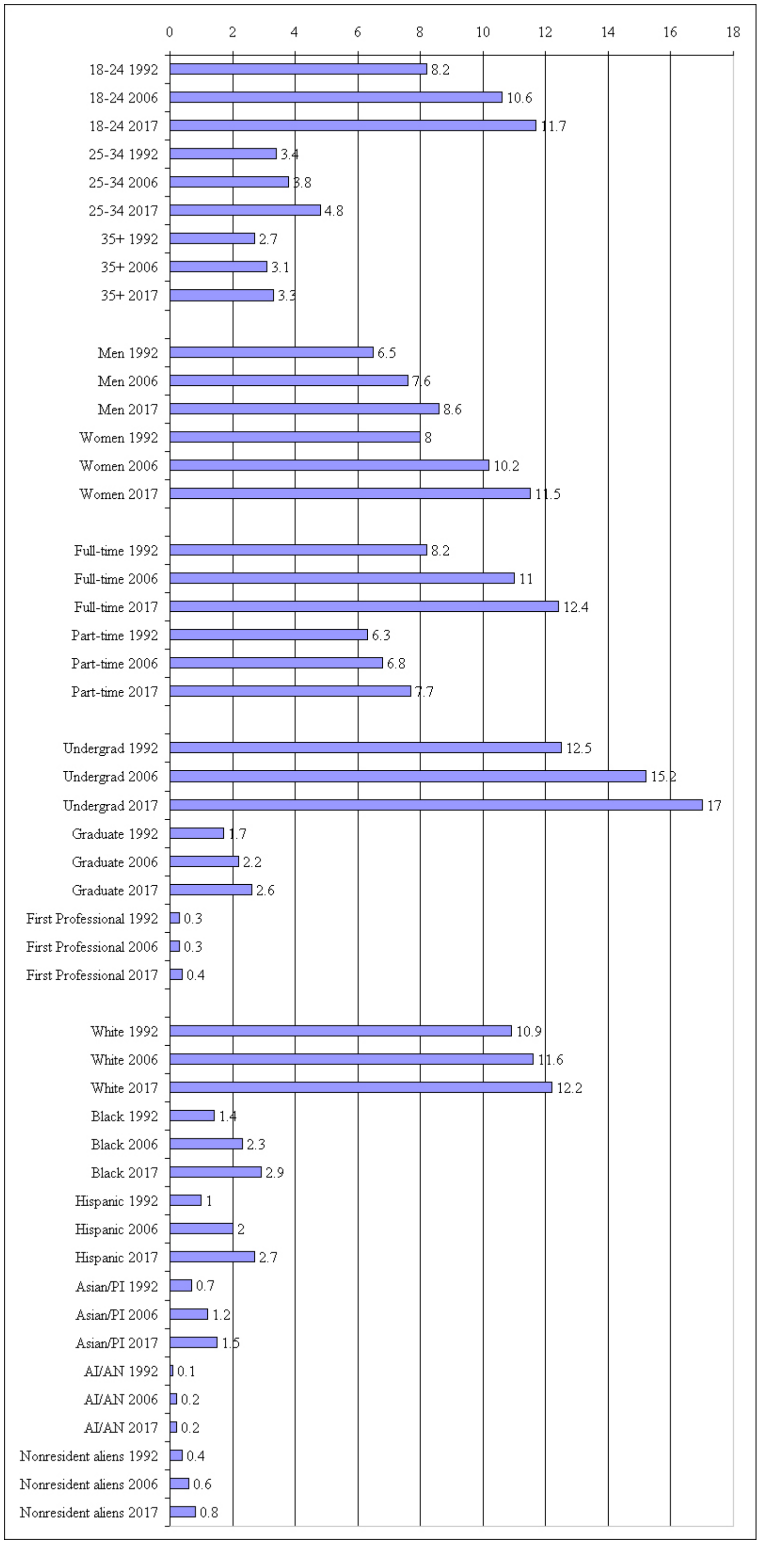


Figure 3: Alternative version of Appendix Figure A-2.

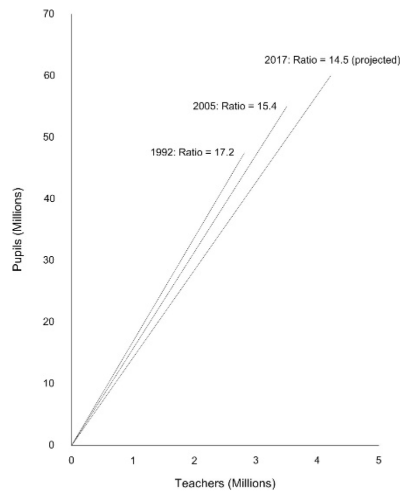


Figure 4: Alternative version of Appendix Figure A-3.

Color. As noted in §1, the limited use of color in NCES publications seems to be dictated by factors other than effectiveness. A secondary issue, of course, is grayscale printing. Two colors of approximately the same saturation and brightness (in a so-called H(ue)-S(aturation)-L(ightness) color model) may print identically on a gray-scale printer. There are, however, readily available color scales that do not have this problem.

The "black plus one other color" model used in NCES' *Projections of Education Statistics* is vulnerable to inconsistencies in the year-to-shade/color encoding across figures. For example, the predominant encoding in Hussar and Bailey (2008) is: for **1992**, light blue (RGB = (127,179,210)), for **2005**, mid blue (RGB = (0,85,165)), and for **2017** (projected), white (RBG = (255,255,255)). However, the scheme is not employed consistently: see Appendix Figure A-4. Theses color values do, though, translate well to grayscale hard copy.

Good Practice: Especially when there is a single non-black color for a publication, it should be used consistently.

Projections. The use of color and a virtually indiscernible increase in line thickness to distinguish actual from projected values is ineffective, particularly in grayscale hard copy. Figure 6 illustrates an alternative to Appendix Figure A-5.

Good Practice: Unambiguously distinguish projected data values from actual data values.

2.2 New Opportunities

Emerging, and in some cases emerged, visualization methods and tools are very powerful, especially for visualizing data points themselves, rather than statistical summaries of data, as well as for conveying relationships among variables. In this section, we present several of these methods, as a stimulus for NCES to explore them further.

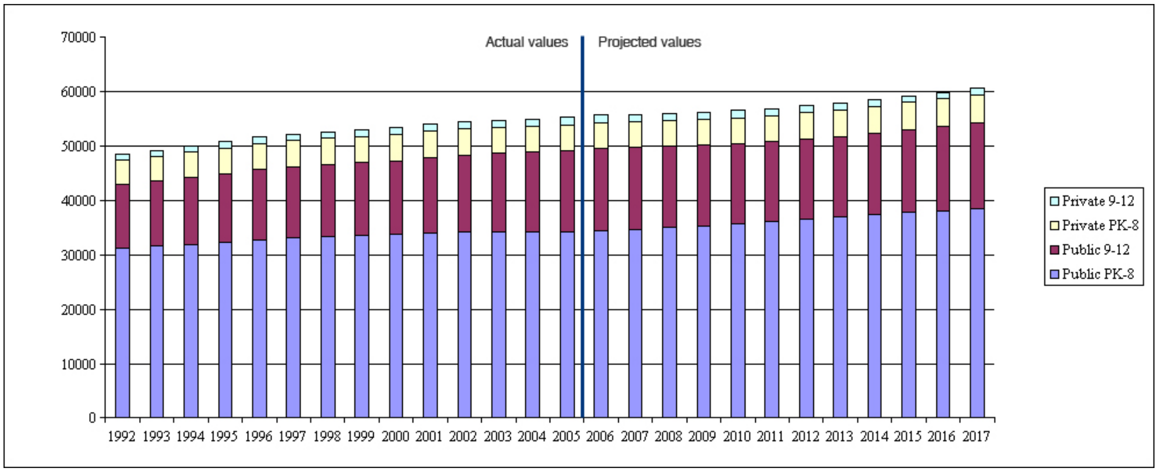


Figure 5: Alternative version of Appendix Figures A-6 and A-7 containing more information than those two figures combined.

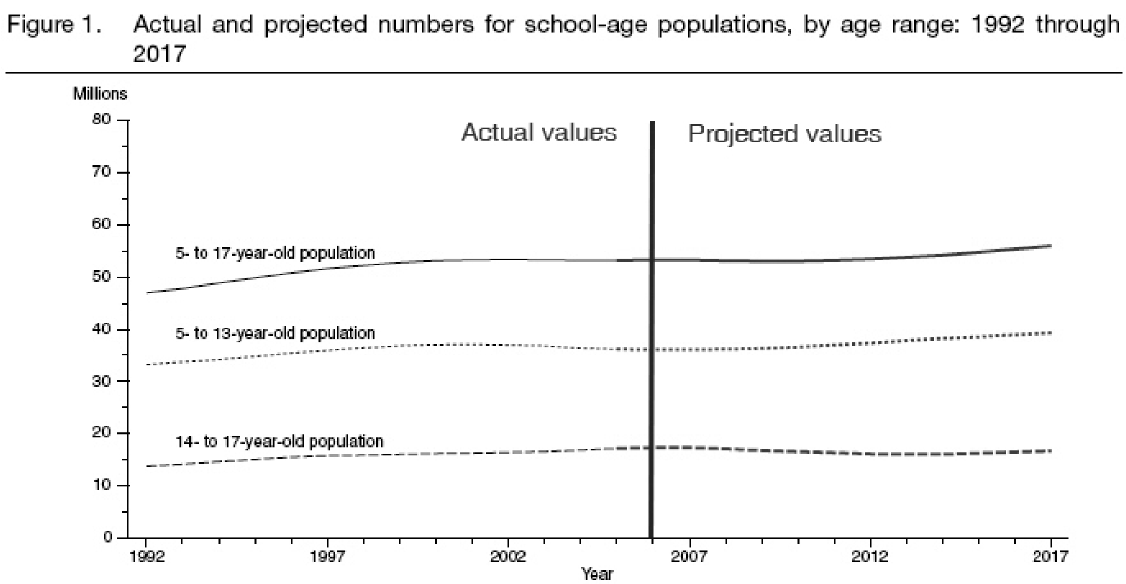


Figure 6: Alternative version of Appendix Figure A-5 in which actual and projected values are distinguished more clearly.

Most of the figures in this section arise from data employed in experiments conducted by NISS in support of the 2008 NCES/NISS Task Force on Nonresponse Bias Analysis (National Institute of Statistical Sciences, 2009). The underlying dataset consists of 29 variables for 1,255 schools, extracted from public release versions of the Common Core of Data (CCD) and Private School Survey (PSS). Variables that appear in the figures below and in §6.1 are:

Black: Total black enrollment

Charter: Charter school status

FreeLunch: Number of students eligible for free lunch FTE Teachers: Number of full-time equivalent teachers Hispanic: Total Hispanic enrollment

Locale: 8 categories

Male: Total number of male students

Reduced: Number of students eligible for reduced-price lunch

State in which the school is located

TotalStudents: Total number of students.

Most of the figures in this section were produced using JMP®.

Scatterplot Matrices. Scatterplots are venerable, of course, but also very useful. Figure 7 shows a scatterplot matrix, which comprises scatterplots for pair of seven numerical variables: **TotalStudents, Male, White, Black, Hispanic, FreeLunch** and **ReducedLunch**. In the scatterplots in the first column in Figure 7, each of the other variables is plotted on the y-axis against **TotalStudents** on the x -axis. There is redundancy in this figure, because every pair of variables is plotted twice, once with the first on the x -axis and once with the second on the y-axis. This redundancy is removed in Figure 8.

Mosaic Plots (Friendly, 1994) are graphical representations of two-way contingency tables, that is, of relationships between two categorical variables. Figure 9 contains an illustration, showing the distribution of **Charter** by **State**. The width of the bars for each state is proportional to the number of schools from that state in the data set. Each bar is split vertically according to **Charter**. The isolated bar at the right shows the nationwide proportions. The figure suggests, and analyses using JMP® confirm, that **Charter** and **State** are not independent.

Figure 10 is analogous, showing **Locale** by **State**. Independence fails in this case as well.

Bubble Plots are illustrated in Figures 11–13; see also Figure 16. At one level, they are scatterplots that can encode at least one additional characteristic, as well as handle both numerical and categorical variables, but they are in fact much more versatile.

Figure 11 is a scatterplot of two numerical variables - **FTETeachers** as a function of **TotalStudents -** but encodes **Title1**, a third, categorical characteristic, using color. It reveals an interesting

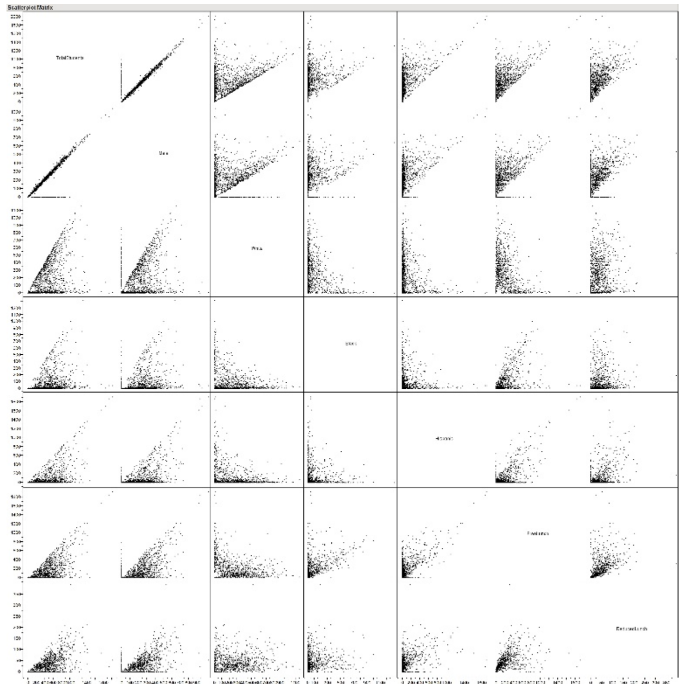


Figure 7: Scatterplot matrix for seven school-level demographic variables TotalStudents, Male, White, Black, Hispanic, FreeLunch and ReducedLunch.

property: schools with low student/teacher ratios, which lie in the upper left-hand portion of the graph, almost all have **Title1** = 1.

Figure 12 adds a second categorical variable—Male, which is encoded as the size of the bubbles.

Figure 13 shows one numerical variable, **FTETeachers**, as a function of the categorical variable State, with two additional variables. The numerical variable **TotalStudents** is encoded as bubble size, and the categorical variable **Charter** is encoded as color.

Multiple Characteristics. Many graphics in NCES publications depict only a single response. However, there are many cases where multiple responses are of interest, not only individually but also in terms of the relationships among them.[5]

Scatterplot matrices, discussed above, are one tool for this purpose. Mosaic plots also facilitate representation of multi-dimensional structure. Figure 14 illustrates for 8-dimensional data taken from the Current Population Survey (CPS). It reveals a multitude of relationships among four categorical variables: **race** - 2 categories, **salary** - 2 categories, **marital status** - 2 categories and **educational attainment** - 5 categories.

There are also direct methods. Figure 15 (Eick and Karr, 2002) simultaneously visualizes two

[5] Many discussions of visual scalability focus on the number rather than the dimension of data points, but some, notably Eick and Karr (2002), address both.

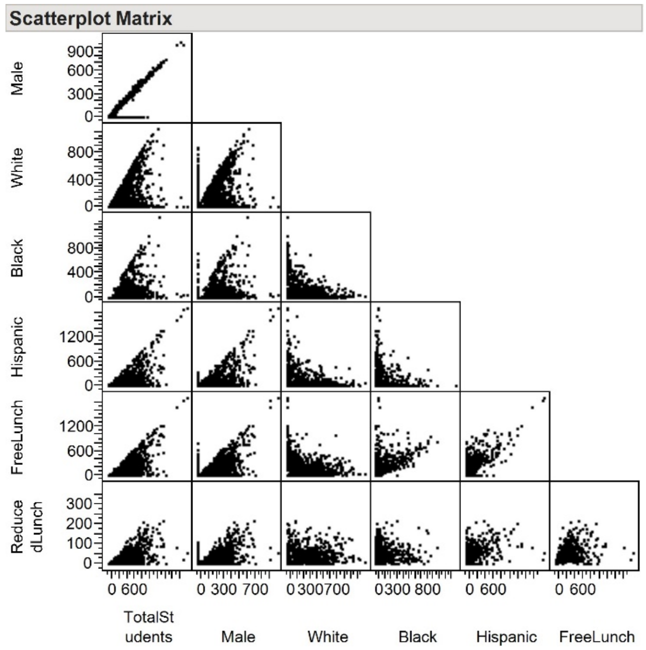


Figure 8: Scatterplot matrix of the same seven school-level demographic variables as in Figure 7, but with redundant plots removed.

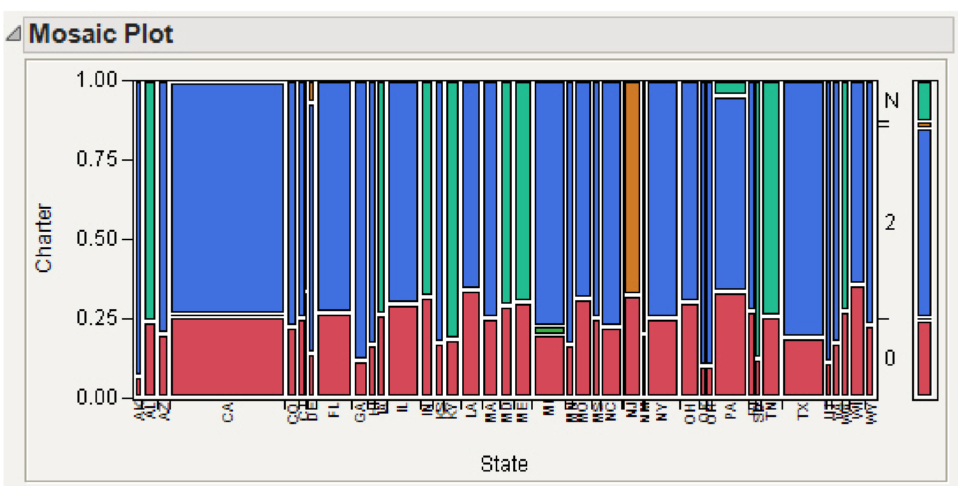


Figure 9: Mosaic plot for two-way contingency table showing the distribution of Charter by State

Figure 10: Mosaic plot for two-way contingency table showing the distribution of Locale by State

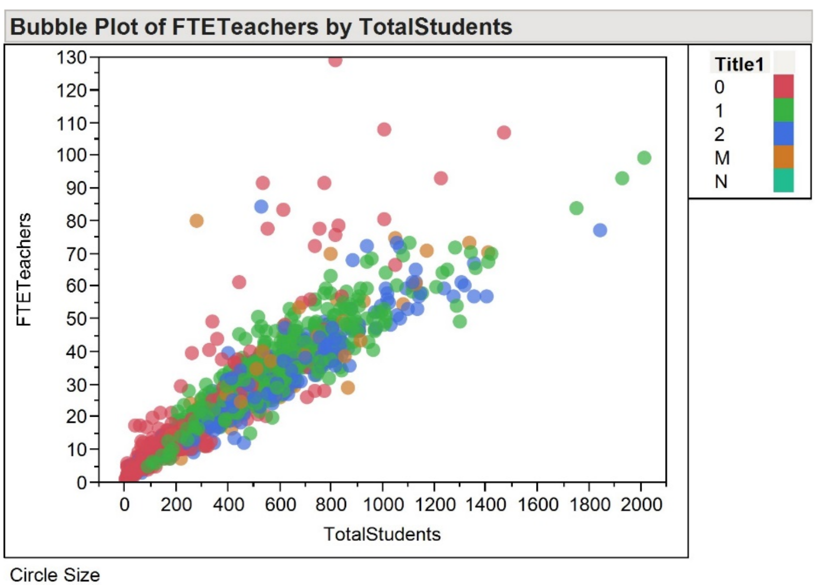


Figure 11: Bubble plot of FTETeachers vs. TotalStudents, with Title1 encoded by means of color

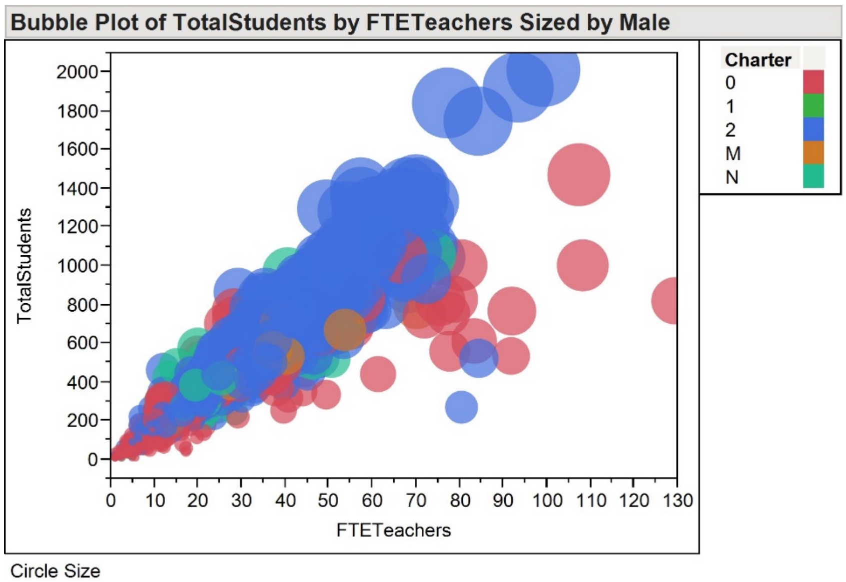


Figure 12: Bubble plot of **FTETeachers** vs. **TotalStudents**, with **Title1** encoded by means of color and **Male** encoded by size

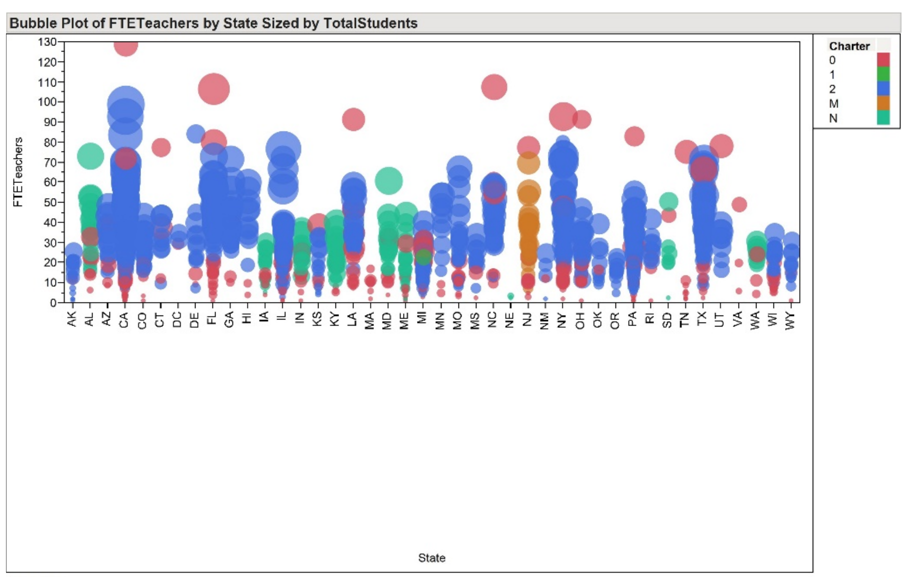


Figure 13: Bubble plot of FTETeachers as a function State, with TotalStudents encoded as bubble size and Charter encoded as color.

contingency tables indexed by the same rows and columns. The data arise from study of software developers; rows are developers and columns are modules of the code for a large system. One variable, the frequency with which each developer altered each module, is encoded as the width of the associated rectangle. A second characteristic, the organizational home of each developer, is encoded as color.

Figure 16, which is yet another bubble plot, also simultaneously shows two contingency tables indexed by **Charter** and **Locale**. One table contains **TotalStudents**, which is encoded in Figure 16 as bubble size, and the other contains **FTETeachers**, which (after being discretized) is encoded as color.

True three-dimensional (3-D) graphics[6] have legitimate uses. Figure 17 is a 3-way scatterplot of **FTETeachers**, **TotalStudents** and **White**, and is quite informative about the relationships among them. The JMP® software that produced it allows "live" rotation of the view.[7]

Parallel plots are not universally liked, but can be valuable if used selectively. Figure 18 shows the concept. It presents six school-level demographic variables - **TotalStudents, White, Black, Hispanic, FreeLunch** and **ReducedLunch** - for 17 schools in North Carolina for which Title1 = 1. There is one connected line for each school, which is further delineated by color. All six variables are plotted on a common vertical scale.

The illustration in Figure 19 conveys both the strengths and the weaknesses of parallel plots. It shows the same six school-level demographic variables as in Figure 18, but separately for those schools with **Title1** equal to 0, 1 and 2. There is still one line for each school, but now without color[8], and there is no longer a common vertical scale. Structure is apparent in Figure 19, for instance, that there are schools with high **Hispanic** enrollment, especially among schools with **Title1** = 0, that are low with respect to **FreeLunch**. On the other hand, the massive overplotting makes it impossible to track all six variables for any one school.

Good Practice: Investigate new visualization methods, including those described here.

Networks. Most NCES datasets do not have inherent network structure, so there may be limited need or opportunity to employ the wealth of available techniques for visualization of - in some cases, extremely large - networks. Figure 20 exemplifies their power. In it, nodes are individuals and edges are pairwise relationships between them. Color and size encode two additional node characteristics.

It is relatively easy to construct non-fanciful situations in which representing NCES data as networks might be valuable. One is labor force mobility, in which nodes would be sectors or even individual employers, and edges, through width or color, represent flows between them. A second example, which is more speculative, is a network visualization of student performance on the National Assessment of Educational Progress (NAEP), in which nodes are items and edges represent the frequency of correct answers to both items.

[6] As opposed to the "quasi-3-D" graphics such as bar charts with shadows that have become rampant in some Microsoft PowerPoint presentations.
[7] Not without problems: see §7.
[8] There are "not enough" colors to handle all 1255 schools.

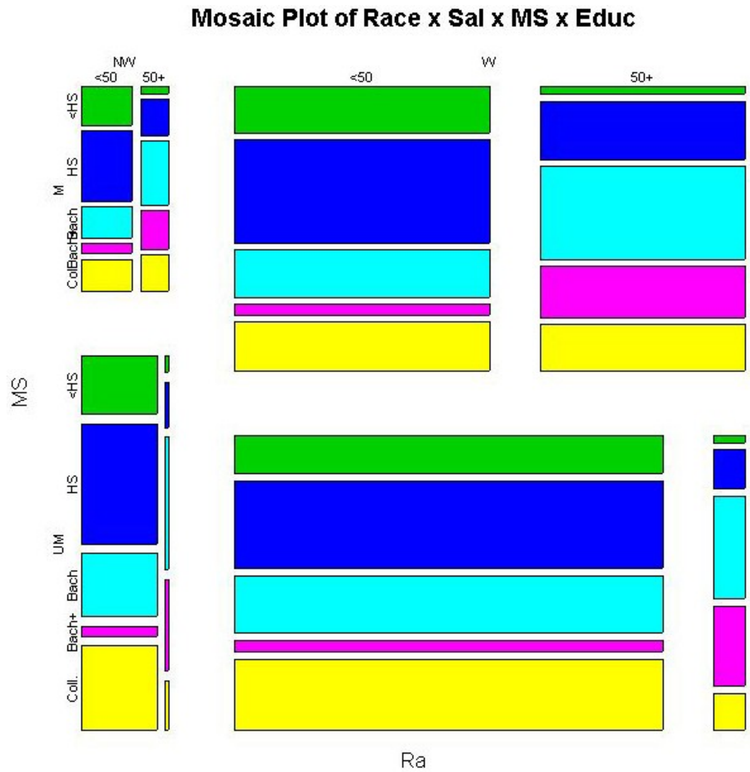


Figure 14: Mosaic plot containing four-dimensional data from the CPS.

Maintaining awareness of techniques for network visualization seems desirable

III. MAPS

For the purposes of this discussion, "maps" means maps of the US, at varying levels of geographical resolution. Maps constitute a major part of many NCES publications, and effectively facilitate state-to-state comparisons. Some maps in NCES publications convey less information than is possible, albeit not necessarily less than is optimal for some readers.[9]

Maps in publications serve three principal functions:

Access to (coarsened) data values, by means of encoding the values as shading or coloring. This function is illustrated in Appendix Figure A-8: the variable is the projected change in PK– 12 enrollment, and has been coarsened to 5 ranges. The user locates the state of on the map, notes the shading, and then reads the coarsened data value from the legend. In isolation, this function can be served equally effectively, and in higher resolution, by a table.

[9] Some of the material here is motivated by MacEachern (1995).

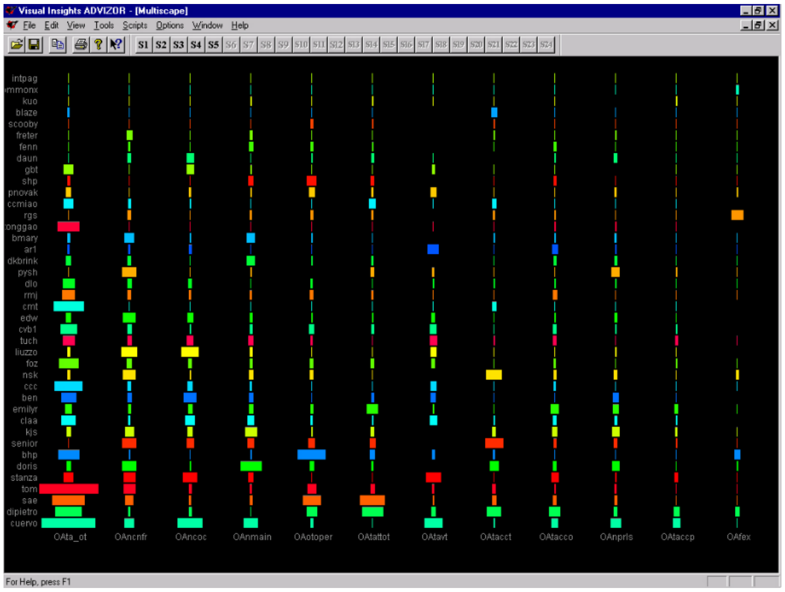


Figure 15: Simultaneous representation of two two-way contingency tables, one coded as symbol size and the other as symbol color (Eick and Karr, 2002).

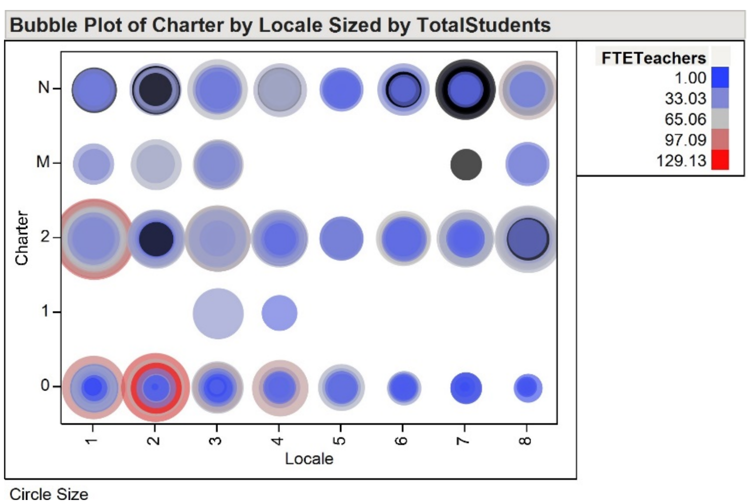


Figure 16: Bubble plot showing two contingency tables indexed by Charter and Locale: Total- Students is encoded as bubble size, while FTETeachers is encoded as color.

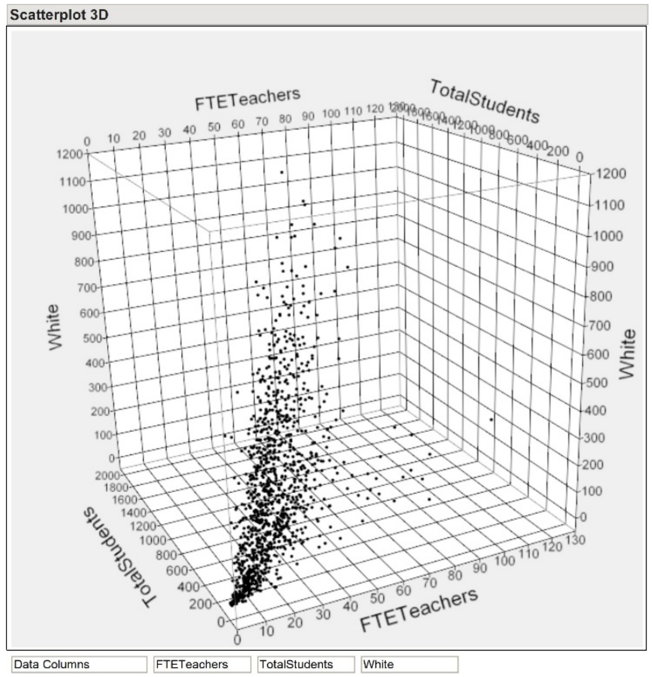


Figure 17: Three-dimensional scatterplot of FTETeachers, TotalStudents and White.

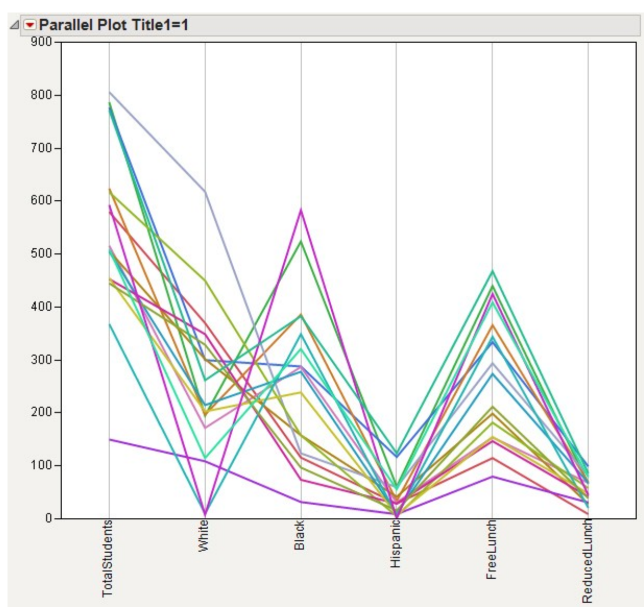


Figure 18: Parallel coordinates plot of six school-level demographic variables for 17 schools in North Carolina having Title1 = 1.

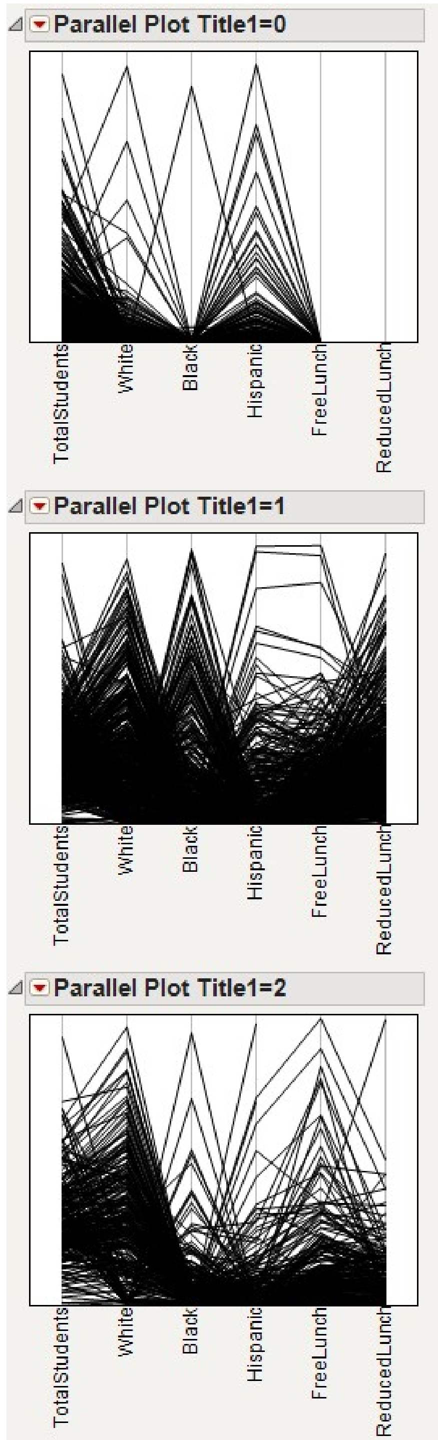


Figure 19: Parallel coordinates plot of six school-level demographic variables for 1255 schools, but separately for each value of Title1.

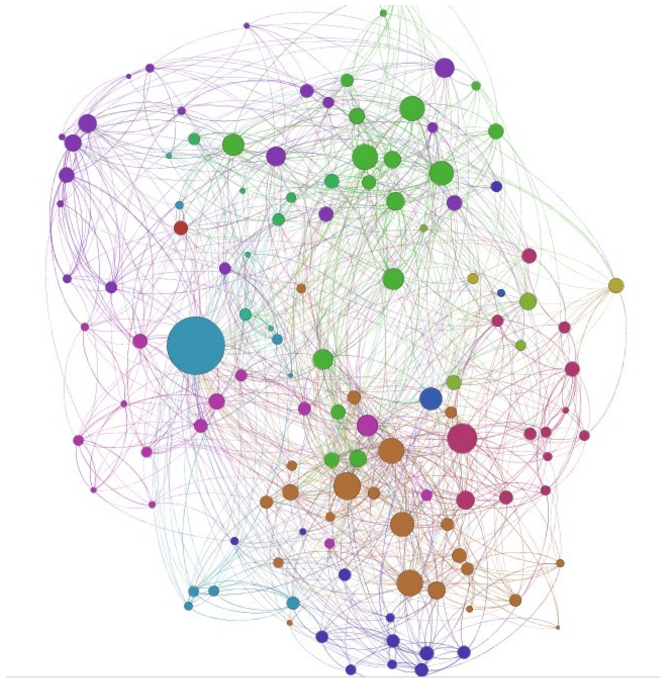


Figure 20: Network visualization, in which nodes are individuals and edges are pairwise relationships between them. Two additional node characteristics are encoded as size and color.

Elucidation of regional effects, which are manifested qualitatively in similarity of color or shading of adjacent states. This also is illustrated in Appendix Figure A-8.

Visual gestalt, that is, high-level, qualitative structure that may be difficult to define precisely, but is not difficult to recognize. In Appendix Figure A-8, there is a clear difference between the "Rust Belt states" and the "Sun Belt states." That Louisiana is an "outlier" is evident.

In contexts such as Appendix Figure A-8, the latter two func.tions may lead to misapprehension. For instance, the numerical values of two states with different shadings may be closer to each other than the values of two other states with the same shading. Also, some readers may, even if only subconsciously, interpret all map differences as statistically significant.

3.1 Principal Items

Value-to-Map Encoding. Many maps in NCES publications that employ the "black plus one color" model display numerical values at the state level by encoding them as some form of shading of the states. Often, the shading scheme is linked only loosely (and sometimes, non-intuitively) to the data values, usually in the sense that higher values correspond to darker or more complete shading. Sometimes the shading scheme fails to differentiate effectively between increases and decreases.

Encoding numerical values as shading patterns is well-known to be problematic. By contrast, the yellow–to–green heat scale in Pickle et al. (1997) is more effective, as illustrated in Figure 21.[10] Although there may be

[10] Nearly every person involved in creation of maps has a favorite color scale. No scale is perfect, and those that are used widely are actually quite similar.

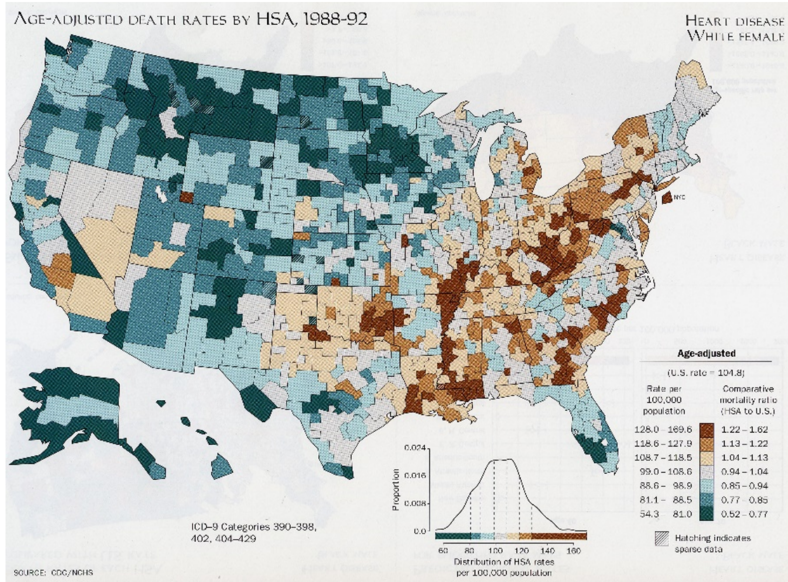


Figure 21: Map of county-level rates of mortality from cancer with a heat-scale index, reproduced from Pickle et al. (1997).

little need for county-level information[11] in NCES publications, the maps in Pickle et al. (1997) display county-level information easily, whereas the cross–hatching in Appendix Figure A-8 would fail badly at higher geographical resolution than states. If NCES were ever to reach the point of publishing district-level information, it would have to confront the "shading vs. color" issue more directly.

Moreover, by using color for values, Pickle et al. (1997) is able to use cross-hatching for two other important purposes: to differentiate maps for males from those for females, and to identify counties with unstable estimates because of low sample sizes (see also §6.1).

The color scheme in Pickle et al. (1997) does not reproduce well in grayscale, as demonstrated in Figure 22.

> *Good Practice:* Use color as the preferred means of encoding numerical information in maps, paying attention to the need for grayscale reproduction.

Access to Data Values. As noted previously, the extreme visual power of maps may be problematic in terms of the underlying data, by creating seeming differences that are not real as well as obscuring differences that are real. The best response is ready access to the data values underlying a map. This is often done in NCES publications. For instance, Appendix Figure A-8 is accompanied in Hussar and Bailey (2008) by the tables in Appendix Figure A-9.

[11] Which might threaten confidentiality.

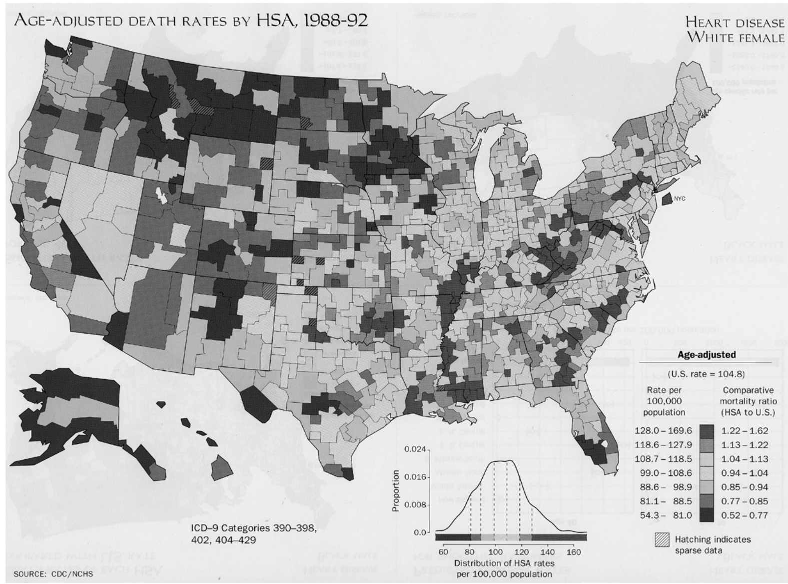


Figure 22: Grayscale version of the map in Figure 21.

Another approach is shown in Figure 23: simply include the data values in the map. This map uses color to differentiate states with projected increases (black) in PK–12 enrollment from those with projected decreases (red). To some degree, Figure 23 lacks the visual gestalt present in Appendix Figure A-8, but that gestalt could easily be added by means of a pastel-like color scheme that does not obscure the data values.

Micromaps Carr and Pickle (2010) provide another way of linking the visual power of maps to the specificity of data values. The example in Figure 24 shows state-level mortality rates among white women for two forms of cancer. This figure contains a wealth of information, not all of which is readily apparent. The states are ranked by their county-level median values, and are displayed in groups of five,[12] which move successively farther from the national median as one goes up (above the median) or down (below the median) from the box labeled "Median." A per-county Box plot is given for each state.

The maps are cumulative moving from either the top or the bottom toward the median: the five "states of focus" are colored, which allows easy access to the Box plots, and states further from median are white.

> *Good Practice*: Provide easy access to all data values underlying maps, either on the map itself or in associated tables.

[12] A version of the "small multiples" approach to visual displays.

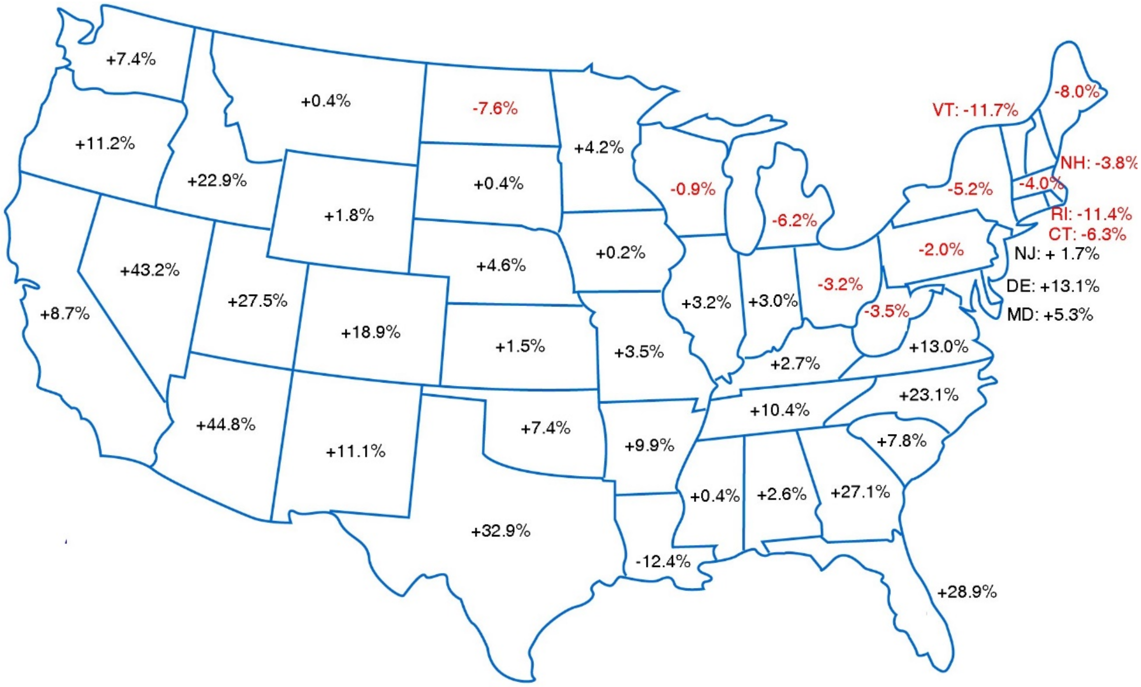


Figure 23: Alternative version of Appendix Figure A-8 that contains actual data values.

3.2 Additional Items

Statistical Significance. Also as noted previously, maps may both obscure and appear to create statistical significance. Figure 25 shows a simple and elegant solution. The underlying data values are state-level divorce rates.[13] States whose rates differ significantly from the national rate contain a red dot. One-sided tests could have been performed instead, and two colors of dots used, corresponding to statistically significantly greater than or less than the national rate.

Note that the map in Figure 25 redundantly but usefully includes the two-letter United States Postal Service (USPS) abbreviations for the state names. The placement of Alaska and Hawaii is more faithful to actual geography than the conventional placement used in most other maps here.

Good Practice: When doing so is meaningful, include statistical significance in maps.

Three-Dimensional Maps might be used in some circumstances as an alternative to encoding data values as colors, but seem hard to use effectively. Figure 26 depicts data that are "U.S. Robbery Statistics." The response is encoded categorically as color and - possibly - at higher resolution as bar height. That is, not all black bars appear to be of the same height, so based on bars, the rate is higher in New Mexico than in Oklahoma, which cannot be discerned from the

[13] The values are estimates derived from the American Community Survey (ACS).

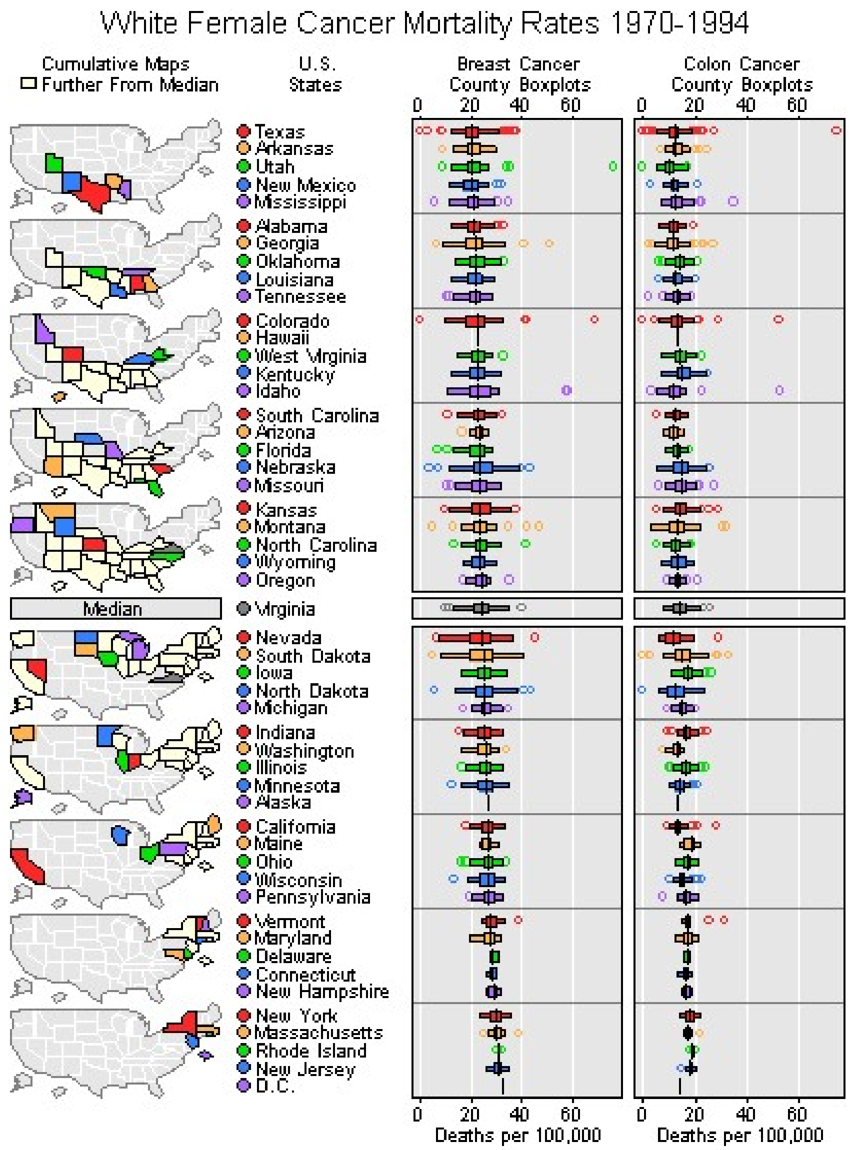


Figure 24: Example of a linked micromap that provides ready access to data values.

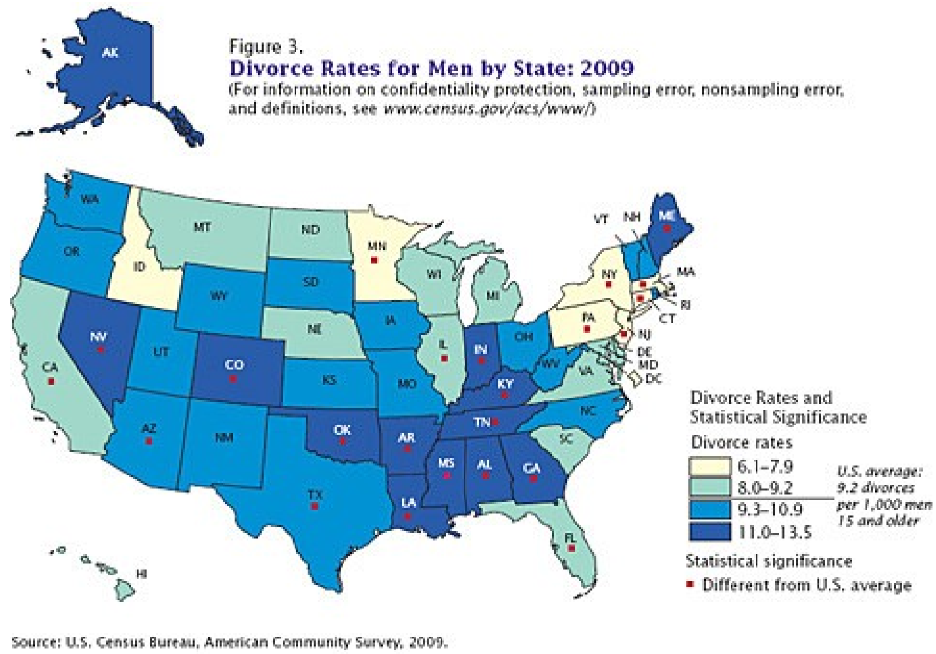


Figure 25: Map produced by the US Census Bureau showing divorce rates by state, with red dots indicating statistically significant differences from the national average.

colors. There are also problems with this map. There are no states in the category corresponding to cyan ("400"), and because of occlusion the viewer must guess that the magenta bar corresponds (presumably) to New York.

The *prism map* in Figure 27 is an alternative to Figure 26, showing the same data. In some ways it is more effective, and in others it is less so. In Figure 26, it appears as if black bars are not all of the same height, whereas in Figure 27, all states labeled in black clearly are at the same level. Figure 27 makes clear that Maryland is blue, which is not obvious in Figure 26. Neither map is able to show the data for either Delaware or the District of Columbia. The absence of many state boundaries in Figure 27 is likewise problematic.

Conditioned Choropleth Maps can convey relations among geography and as many as three other variables. The map in Figure 28 shows relationships among county-level geography, lung cancer mortality rates (deaths per 10,000 population), percentage of the population below the poverty level and annual precipitation. The response, which is lung cancer mortality rates for white males aged 65–74, is encoded as color—the blue, gray and red categories as shown at the top. **Annual Precipitation** and **Percent Below Poverty Level** are likewise discretized to three levels each, so that in some ways Figure 28 is a contingency table of maps.

To illustrate the insights arising from the map, it is clear that the highest rates are in the region stretching from east Texas to West Virginia (simplistically, the "Appalachians") in counties with

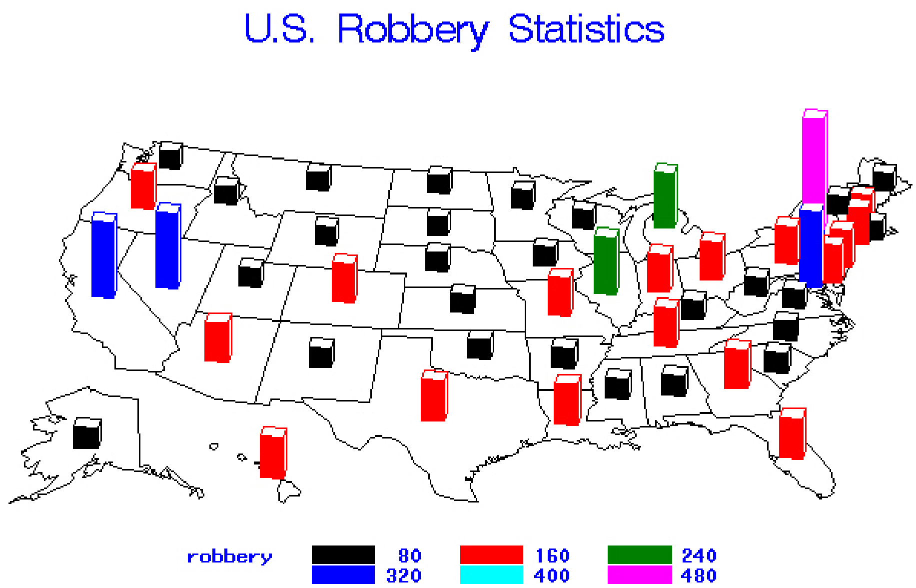


Figure 26: Three-dimensional block map in which a numerical variable is encoded as both color and (to an unclear extent) bar height.

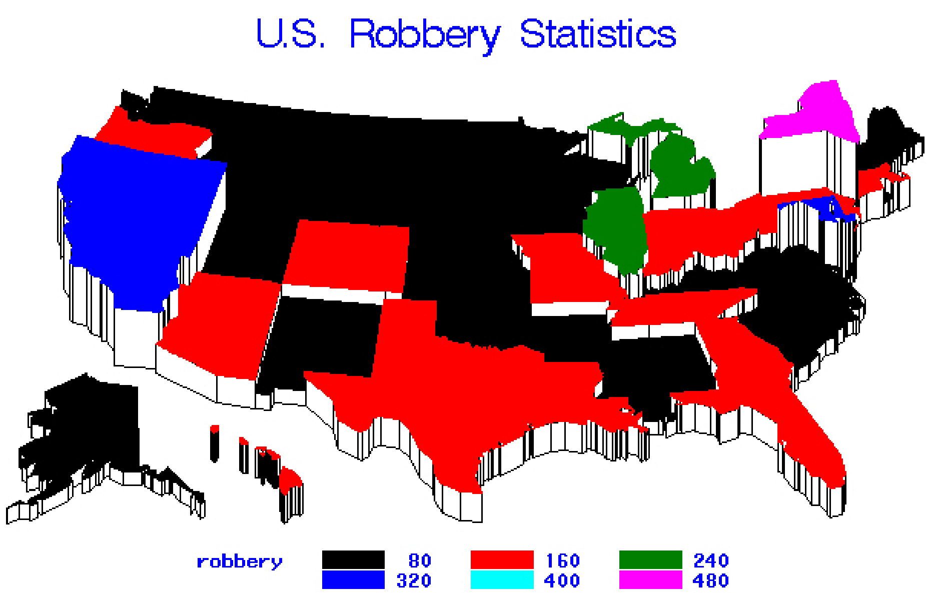


Figure 27: Three-dimensional prism map in which a numerical variable is encoded as "state height."

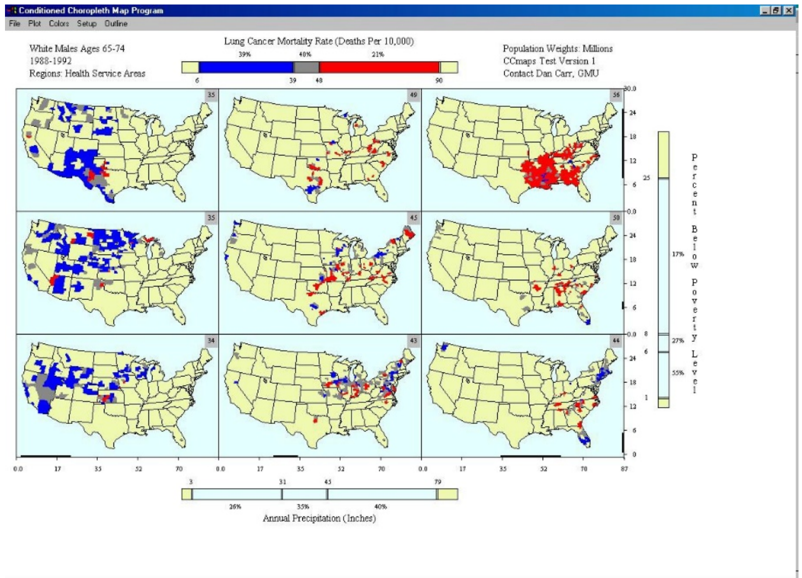


Figure 28: Conditioned choropleth map showing relationships among four variables. See text for details.

high poverty levels and high precipitation. The cross-shaped "hole" in the cell (New Mexico, low precipitation, high poverty) at the upper left consists of counties with medium poverty rate, in the middle left cell. There is a gestalt that low precipitation is associated with low rates—the leftmost column.

Anamorphic Maps. Maps that encode numerical characteristics by means of color are, in general, limited to showing only a single characteristic, which prevents their being used to display relationships between or among variables. Anamorphic maps encode two characteristics, one as color, and the other as "(distorted) size." The distortion is accomplished by sophisticated mathematical algorithms that transform sizes and shapes in a way that preserves adjacency of geographical regions.

Figures 29 and 30 illustrate. In Figure 29, the number of seats each state holds in the Electoral College is encoded as size—the more electoral votes, the larger the area of the state. The vote (Democrat or Republican) in the 2004 election is encoded as color. The adjacency structure of the US is maintained. In Figure 30, populations of countries in the world are encoded as areas. Color encodes geographical regions, but not especially usefully; it could have been used to encode another characteristic of interest, such as per capita income. Adjacency is preserved at the expense of contortions: Russia (unsaturated green) is nearly invisible compared to China (saturated green).

Anamorphic maps are not yet widely enough known and implemented for NCES to use them within the near-term future. Continuing attention seems prudent, however.

Good Practice: Selectively investigate use of mapping tools that convey more information that is contained in many of NCES' maps. Use of color is central to implementing new tools.

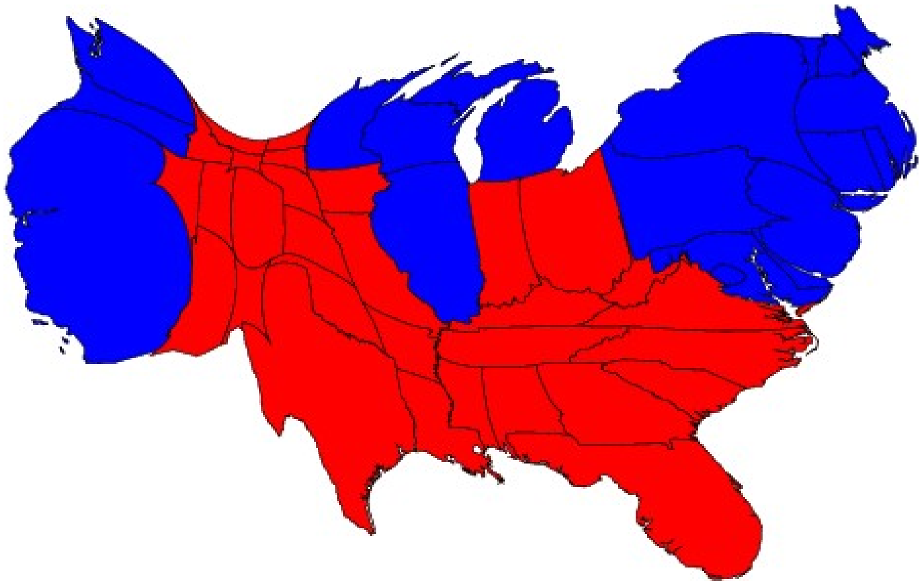


Figure 29: Anamorphic map of the US encoding electoral votes as size and 2004 election outcome as color, from http://www.datavis.ca/.

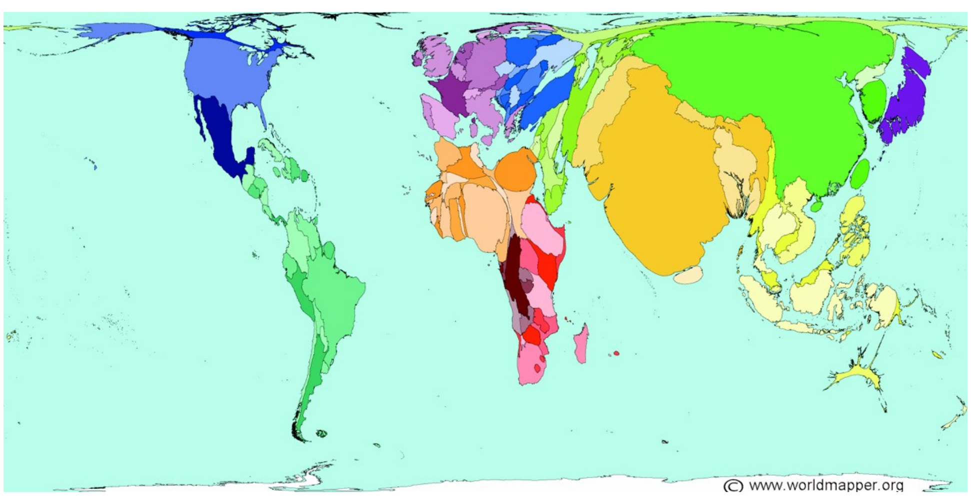


Figure 30: Anamorphic world map encoding population as size, from www.worldmapper.org.

IV. GRAPHICS AND MAPS AS ALTERNATIVES TO TABLES

Some tables in NCES publications would be more effective as graphics or maps, provided that access to the data values is maintained.

To illustrate, consider the two tables in Appendix Figure A-9.[14] One alternative is the horizontal bar chart in Figure 31. It contains the same detailed information as Appendix Figure A-9, but has several advantages. First, sorting by state name[15] facilitates finding specific states. Second, a visual sort by the magnitude of increase or decrease is possible, at least for the largest magnitudes. Third, the relative numbers of states with increases and those with decreases are clear. Finally, it is apparent that in general projected increases exceed projected decreases. Sorting Figure 31 by percentage increase reproduces the rankings in Appendix Figure A-9.

The map in Figure 23 is another alternative. Accessibility of information for specific states is as high as in Figure 31, and higher than in Appendix Figure A-9. The geographical structure of the decreases in projected enrollment "jumps out" of Figure 23. However, Figure 23 is weaker than Figure 31 with respect to visual comparison of the magnitudes of increases and decreases. A final alternative is an interactively sortable version of Figure 31; see §7 for further discussion.

> *Good Practice*: Consider increased reliance on graphics and maps as substitutes for or complements to tables, but not to the point that data values are suppressed entirely.

V. VISUALIZATION OF UNCERTAINTY

Many graphics and maps employed by NCES contain variables that are statistical estimates, which therefore carry associated uncertainties. However, inclusion of uncertainties in graphics and maps is relatively rare in NCES publications, even by means of traditional devices such as "error bars." Often, the reason appears to be to avoid user overload. On the other hand, there are also negative consequences. In particular, users may assume that all visual differences are significant, which is especially problematic for maps with coarse color scales. Patterns of uncertainty - for instance, does degree of uncertainty for one variable depend on the value of another variable, or is there spatial structure to the uncertainties? - may also be important, but are invisible.

Visualization of uncertainty is a difficult problem. It has spurred an active field of research, in which there are many promising ideas but few well-established best practices.

5.1 Visualizing Uncertainty in Graphics

How uncertainty can be visualized is inherently graphic-specific. In some cases, there are established paradigms. For example, Figure 32 shows a nonparametrically estimated distribution function of **TotalStudents** (see §3.2) together with confidence bands for the estimates. Few people have intuitive difficulty with this scheme, even if most do not grasp it in detail.

[14] Karr (2009) contains discussion of whether it makes sense to place states with increases in projected enrollment and those with decreases in separate tables, and concludes that it does not.

[15] This is a side comment, but sorting by state name and sorting by USPS 2-letter state code are not equivalent, which is a problem to which no consistently applicable solution seems to exist.

Box plots are increasingly familiar. Figure 33 uses them to convey uncertainties associated with the distribution of **TotalStudents** as a function of **Locale**.

Some researchers have proposed encoding uncertainty as color, but the suggested mappings are ambiguous and subject to misinterpretation.

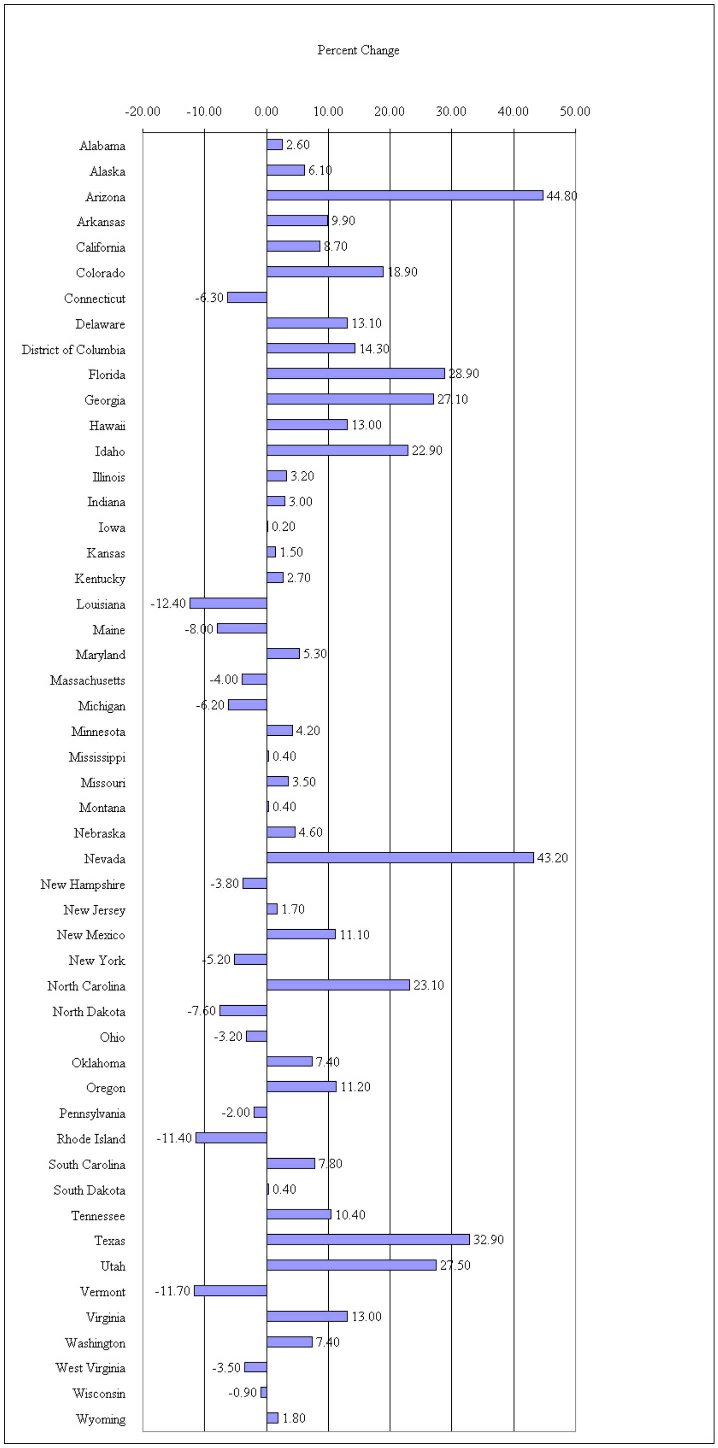


Figure 31: Graphical version of Appendix Figure A-9.

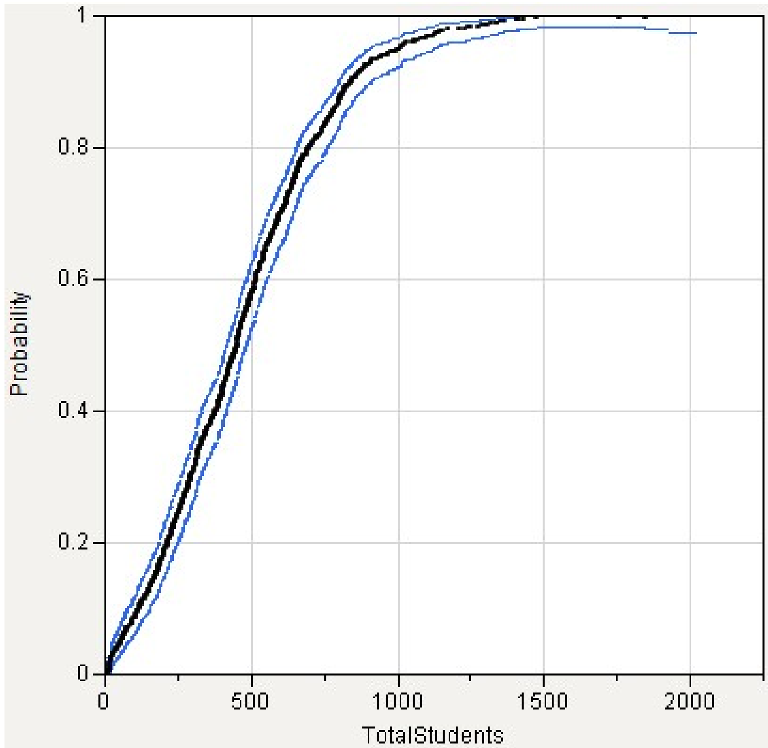


Figure 32: Estimated distribution function of TotalStudents with uncertainties visualized as confidence bands.

There is also the issue of how visually to deal with cases where the uncertainty is so high that associated estimates are deemed unreliable (Karr and Kinney, 2011). Some graphics and maps do not readily accommodate this problem, others do. To illustrate, in the map in Figure 23 one could simply omit the unreliable estimates, albeit with appropriate explanation. There is no correspondingly obvious solution for Appendix Figure A-8. Coloring using white is not possible, and black might be ambiguous, as well as lead to a map that is dominated visually by "non-values." The map in Figure 21 (§4.1) uses cross-hatching to denote unreliable estimates.

Good Practice: Include uncertainties in graphics on a selective basis, especially when the "main message" is not diluted and the method used to encode uncertainty is well-established.

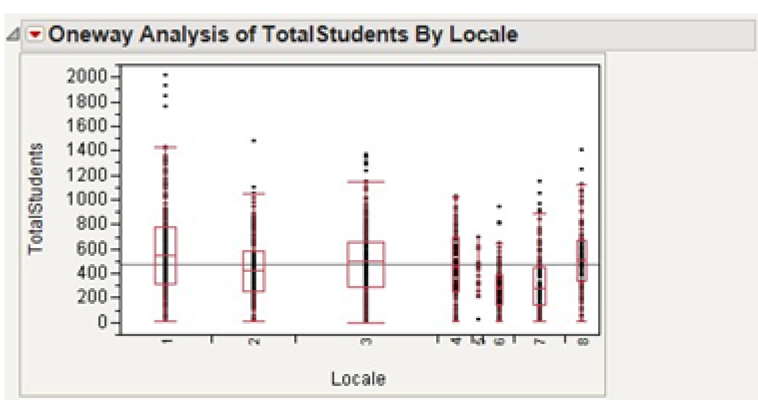


Figure 33: Box plots used to visualize uncertainty in TotalStudents as a function of Locale.

5.2 Visualizing Uncertainty in Maps

None of the map models discussed in §4 is especially amenable to inclusion of uncertainties. As noted in §6.1, color is not appropriate in most cases, although this might be possible in the context of 3-dimensional maps (Figures 26 and 27). Ideas worthy of continuing attention are:

Saturation, in combination with color, with more saturated colors being "more certain." Nearly no users of NCES publications will know of the "hue/saturation/brightness" color model,[16] but the visual interpretation is not difficult. Figure 34 shows three blue colors that differ only with respect to saturation. Interpreting movement from left to right there as increasing uncertainty ("haziness") is straightforward.

Glyphs such as those used to denote statistical significance in Figure 25. Glyphs whose size (or even color) encodes uncertainty may work in some situations. Figure 35 is a prototype version: it is the same map as in Figure 23, but with traffic-signal-colored glyphs representing (artificially generated) uncertainties: green = low uncertainty, yellow = intermediate uncertainty, red = high uncertainty.

Cross-hatching, which is often used in maps for other purposes, may be able to convey uncertainty.

Good Practice: Pay continuing attention to ongoing research, as well as any broadly accepted practices that emerge.

Figure 34: Three blue colors of the same hue and brightness, but different saturations.

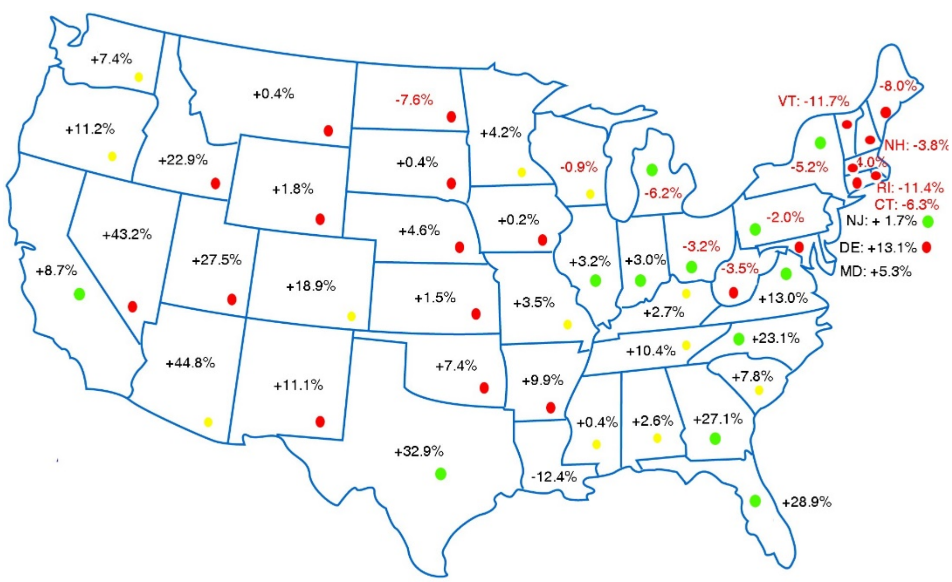


Figure 35: Version of the map in Figure 23 with glyphs representing uncertainties.

[16] For a non-authoritative but useful introduction, see http://en.wikipedia.org/wiki/HSL_and_HSV.

VI. INTERACTIVITY

This compendium is focused largely on "hard copy" publications, which are distributed both physically and as PDF files available from the NCES web site. At some point, however, NCES may also provide interactive web versions of its static graphics, and here we note some functionalities that are inherently useful, as well as address issues raised earlier.[17]

The most important of these functionalities are the following.

Sorting. As noted in §3.1, the data in Appendix Figure A-9 and Figure 31 can be sorted either by state or by magnitude of change. Both sort orders are informative. Capabilities for interactive sorting are well-developed and easily applied. Figure 36 shows an example containing data from the *Health Data for All Ages* section of the web site of the National Center for Health Statistics (NCHS). The capability to sort by multiple columns is present in almost every table published by NCHS. Nearly every software package used by analysts can sort data by one or more columns; users surely will come to expect this functionality in on-line tables.

Linked Views, where the linkage allows selections in one view - via "brushing" with a computer mouse - to propagate to other views of the same data, is a central strength of interactive displays Eick and Karr (2002). To illustrate, consider Appendix Figure A-2, which contains six (year, age of student, sex of student, attendance status of student, degree level, and race/ethnicity of student) distinct categorizations of population of students enrolled in degree-granting postsecondary institutions. In effect, the components of that figure are five two-way marginals of the underlying 6-dimensional contingency table.[18] Linked views are one means for exploring higher-dimensional structure of the data. For example, selection, using a mouse, of the 18–24 category in the first panel in Appendix Figure A-2 would split each bar in every other panel into "18-24" and "other." See also Figure 14, which is a screen shot of an interactive display.

Figure 37 shows the notion of "propagation of selection" more explicitly. The data are measurements emissions of CO (carbon monoxide) and NOx (oxides of nitrogen) by a single automobile under varying driving conditions. The scatterplot is of the **emissions**; the histograms are of **speed** (sp) and **acceleration** (acc). The selection was made in the **speed** histogram, and comprises data points with speed greater than 38 miles/hour. Its effect has been propagated automatically to the two other views. Color encodes CO emissions. Some information about omitted data points is retained in the gray portions of the views.

Interactive mosaic plots (§3.2) are especially effective in exploration of high-dimensional structure. Micromaps (§4.2) provide similar functionality for geographically indexed data.

User-Set Breakpoints for Maps. Technologies are available that allow users interactively to manipulate category boundaries for maps such as those in Figure 28 or Appendix Figure A-8, ideally allowing more detailed understanding of the underlying data.

Reset Capability. The need to provide users a means to "go back to the beginning" is often essential. To illustrate, Figure 38 shows the initial and one rotated view of the same variables as in Figure 17, namely, **FTETeachers**, **TotalStudents** and **White**. The latter seems not especially illuminating, but there is no way

[17] This discussion is dissociated entirely with website requirements in Section 508 of the US Rehabilitation Act.
[18] year × age, year × sex, year × attendance status, year × degree level and year × race/ethnicity.

other than trial-and-error to return to the starting point. Providing this capability requires only a “Click to Reset” button.

Good Practice: Interactive sorting and linked views are mature technologies that NCES can employ immediately. Techniques, visual metaphors and software for manipulation of map break- points are still evolving. Giving them an opportunity to “crystallize” before adopting them seems prudent. Always ensure that interactive graphics and maps have a reset functionality.

APPENDICES

Appendix A: References

Carr, D. B. and Pickle, L. W. (2010). *Visualizing Data Patterns with Micromaps.* Taylor & Francis, London.

Eick, S. G. and Karr, A. F. (2002). Visual scalability. *J. Computational and Graphical Statist.*, 11(1):22–43.

Friendly, M. (1994). Mosaic displays for multi-way contingency tables. *J. Amer. Statist. Assoc.*, 89:190–200.

Hussar, W. J. and Bailey, T. M. (2008). Projections of Education Statistics to 2017. Technical report, National Center for Education Statistics. Available on-line at http://nces.ed.gov/programs/projections/projections2017/.

Karr, A. F. (2009). White paper on NCES projections of education statistics: Presentation and methodology. Final version, October 26, 2009.

Karr, A. F. and Kinney, S. K. (2011). NISS comments on presentation of “unstable estimates”. Draft report submitted to NCES on July 15, 2011.Karr, A. F., Kinney, S. K., and Nunnelly, J. N. (2011). NISS review of statistical statements and underlying justifications in NCES Annual Reports. Draft Final Report: July 15, 2011.

MacEachern, A. (1995). *How Maps Work*. The Guilford Press, New York.

National Institute of Statistical Sciences (2009). NCES/NISS task force on nonresponse bias analysis. Final report, July 7, 2009.

Pickle, L. W., Mungiole, M., Jones, G. K., and White, A. A. (1997). *Atlas of United States Mortality*. U. S. Department of Health and Human Services, Hyattsville, MD.

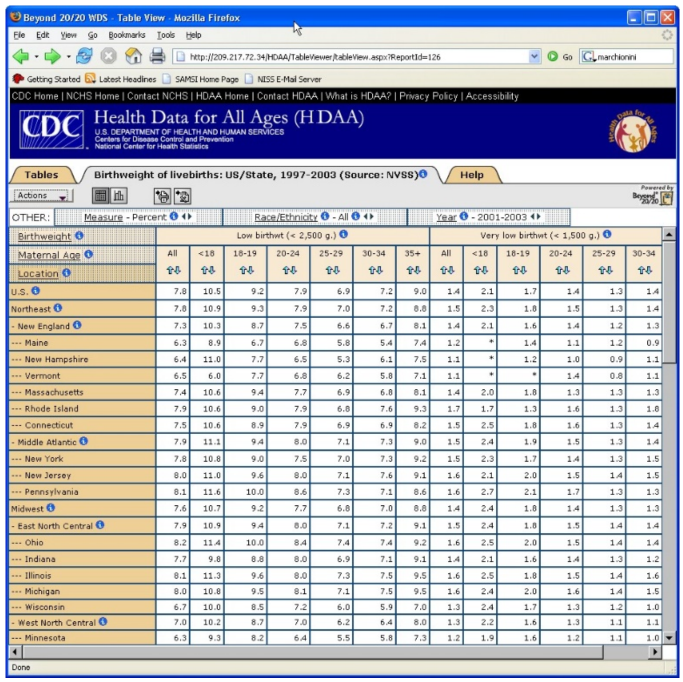
Tufte, E. (1983). *The Visual Display of Quantitative Information*. Graphics Press, Chesire, CT.

Tufte, E. (1990). *Envisioning Information*. Graphics Press, Chesire, CT. Tufte, E. (1997) *Visual Explanations*. Graphics Press, Chesire, CT.

Wilkinson, L. (2005). *The Grammar of Graphics*. Springer–Verlag, New York. 2nd. edition.

Appendix B: Figures from Carr, D. B. and Pickle, L. W.

Carr, D. B. and Pickle, L. W. (2010). *Visualizing Data Patterns with Micromaps.* Taylor & Francis, London.



Birthweight / Maternal Age / Location	Low birthwt (< 2,500 g.)							Very low birthwt (< 1,500 g.)					
	All	<18	18-19	20-24	25-29	30-34	35+	All	<18	18-19	20-24	25-29	30-34
U.S.	7.8	10.5	9.2	7.9	6.9	7.2	9.0	1.4	2.1	1.7	1.4	1.3	1.4
Northeast	7.8	10.9	9.3	7.9	7.0	7.2	8.8	1.5	2.3	1.8	1.5	1.3	1.4
- New England	7.3	10.3	8.7	7.5	6.6	6.7	8.1	1.4	2.1	1.6	1.4	1.2	1.3
--- Maine	6.3	8.9	6.7	6.8	5.8	5.4	7.4	1.2	*	1.4	1.1	1.2	0.9
--- New Hampshire	6.4	11.0	7.7	6.5	5.3	6.1	7.5	1.1	*	1.2	1.0	0.9	1.1
--- Vermont	6.5	6.0	7.7	6.8	6.2	5.8	7.1	1.1	*	*	1.4	0.8	1.1
--- Massachusetts	7.4	10.6	9.4	7.7	6.9	6.8	8.1	1.4	2.0	1.8	1.3	1.3	1.3
--- Rhode Island	7.9	10.6	9.0	7.9	6.8	7.6	9.3	1.7	1.7	1.3	1.6	1.3	1.8
--- Connecticut	7.5	10.6	8.9	7.9	6.9	6.9	8.2	1.5	2.5	1.8	1.6	1.3	1.4
- Middle Atlantic	7.9	11.1	9.4	8.0	7.1	7.3	9.0	1.5	2.4	1.9	1.5	1.3	1.4
--- New York	7.8	10.8	9.0	7.5	7.0	7.3	9.2	1.5	2.3	1.7	1.4	1.3	1.5
--- New Jersey	8.0	11.0	9.6	8.0	7.1	7.6	9.1	1.6	2.1	2.0	1.5	1.4	1.5
--- Pennsylvania	8.1	11.6	10.0	8.6	7.3	7.1	8.6	1.6	2.7	2.1	1.7	1.3	1.3
Midwest	7.6	10.7	9.2	7.7	6.8	7.0	8.8	1.4	2.4	1.8	1.4	1.3	1.3
- East North Central	7.9	10.9	9.4	8.0	7.1	7.2	9.1	1.5	2.4	1.8	1.5	1.4	1.4
--- Ohio	8.2	11.4	10.0	8.4	7.4	7.4	9.2	1.6	2.5	2.0	1.5	1.4	1.4
--- Indiana	7.7	9.8	8.8	8.0	6.9	7.1	9.1	1.4	2.1	1.6	1.4	1.3	1.2
--- Illinois	8.1	11.3	9.6	8.0	7.3	7.5	9.5	1.6	2.5	1.8	1.5	1.4	1.6
--- Michigan	8.0	10.8	9.5	8.1	7.1	7.5	9.5	1.6	2.4	2.0	1.6	1.4	1.5
--- Wisconsin	6.7	10.0	8.5	7.2	6.0	5.9	7.0	1.3	2.4	1.7	1.3	1.2	1.0
- West North Central	7.0	10.2	8.7	7.0	6.2	6.4	8.0	1.3	2.2	1.6	1.3	1.1	1.1
--- Minnesota	6.3	9.3	8.2	6.4	5.5	5.8	7.3	1.2	1.9	1.6	1.2	1.1	1.0

Figure 36: Example of an interactively sortable table, taken from the web site of the NCHS.

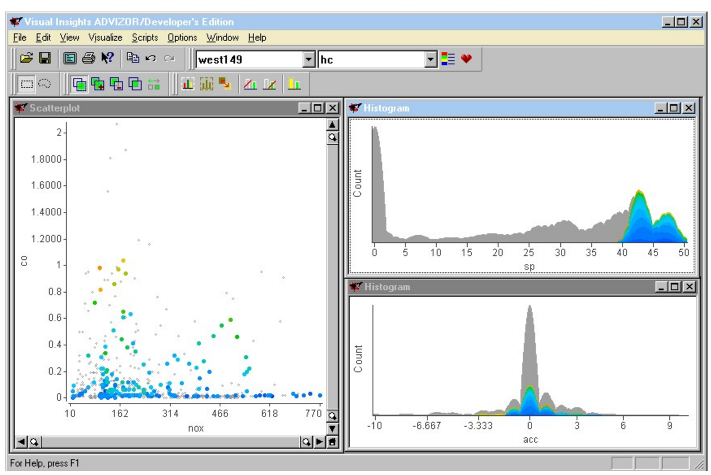


Figure 37: Example of linked views of the same data in which selection in one view propagates to the other views. See text for details.

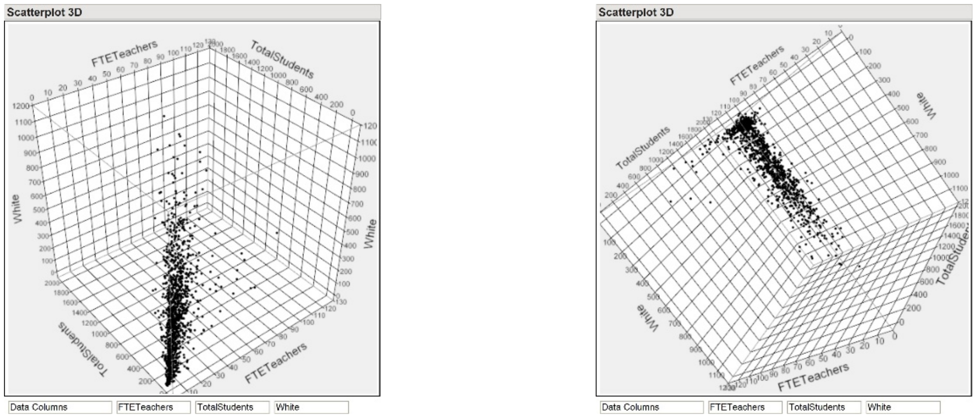


Figure 38: Initial (left) and rotated (right) 3-dimensional scatterplots of FTETeachers, TotalStudents and White

Appendix C: Figures from Hussar, W. J. and Bailey, T. M.

Hussar, W. J. and Bailey, T. M. (2008). Projections of Education Statistics to 2017. Technical report, National Center for Education Statistics. Available on-line at http://nces.ed.gov/programs/projections/projections2017/.

The figures in this Appendix were also discussed in Karr (2009) and were cited merely as exemplars in the body of this report with no specific intent to criticize them.

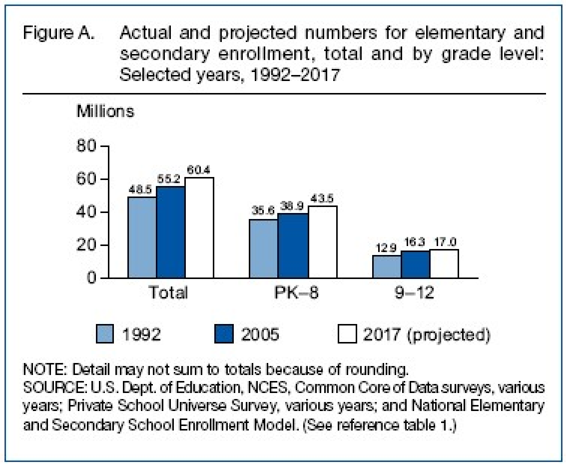


Figure A-1: Figure A of Hussar and Bailey (2008)

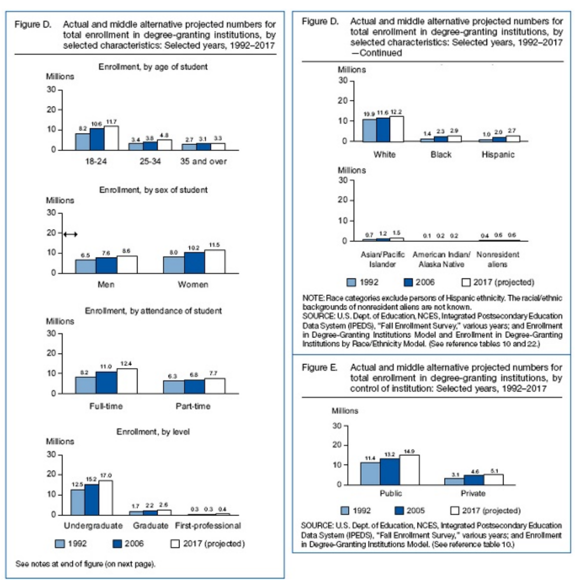


Figure A-2: Figure D of Hussar and Bailey (2008).

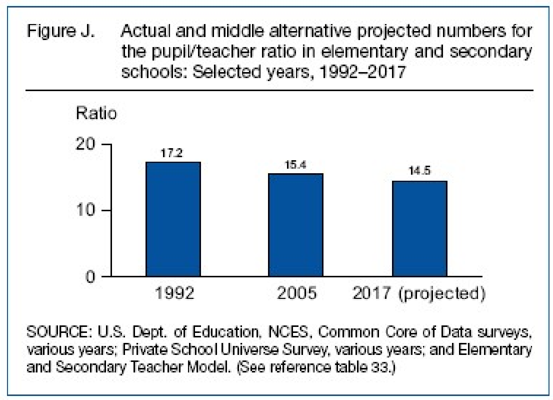


Figure A-3: Figure J of Hussar and Bailey (2008).

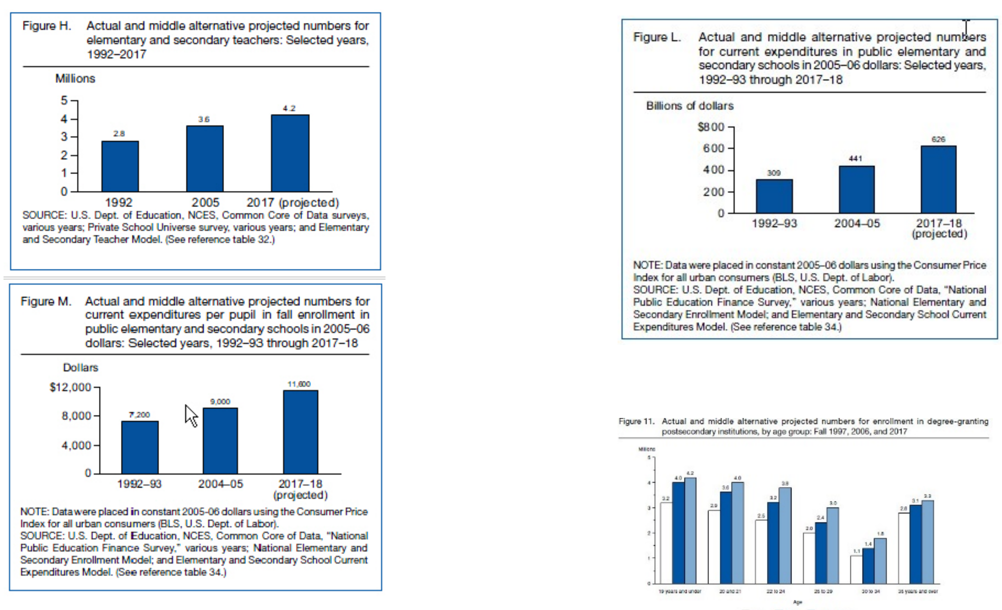


Figure A-4: Figures H, L, M and 11 of Hussar and Bailey (2008), which are not shaded consistently with other figures there.

Figure 1. Actual and projected numbers for school-age populations, by age range: 1992 through 2017

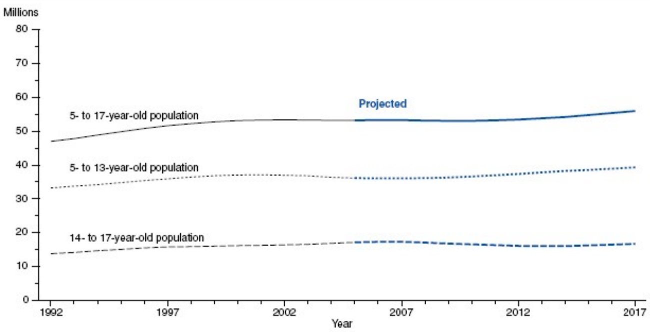


Figure A-5: Figure 1 of Hussar and Bailey (2008).

Figure 2. Actual and projected numbers for enrollment in elementary and secondary schools, by grade level: Fall 1992 through fall 2017

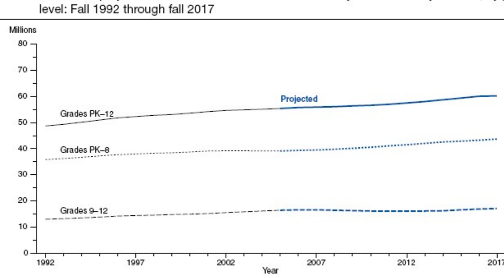


Figure A-6: Figure 2 of Hussar and Bailey (2008).

Figure 3. Actual and projected numbers for enrollment in elementary and secondary schools, by control of school: Fall 1992 through fall 2017

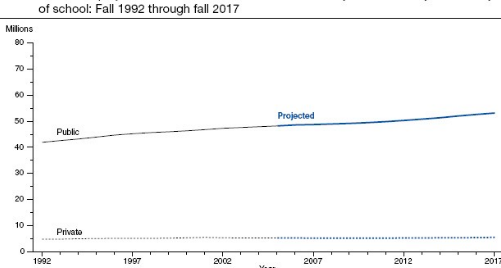


Figure A-7: Figure 3 of Hussar and Bailey (2008).

Figure 5. Projected percentage change in grades PK–12 enrollment in public schools, by state: Fall 2005 through fall 2017

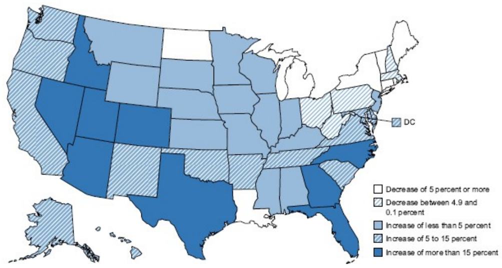


Figure A-8: Figure 5 of Hussar and Bailey (2008)

Table A. Projected percentage increases in public elementary and secondary school enrollment, by state: 2005 through 2017

State	Percent change	State	Percent change
Arizona	44.8	Washington	7.4
Nevada	43.2	Oklahoma	7.4
Texas	32.9	Alaska	6.1
Florida	28.9	Maryland	5.3
Utah	27.5	Nebraska	4.6
Georgia	27.1	Minnesota	4.2
North Carolina	23.1	Missouri	3.5
Idaho	22.9	Illinois	3.2
Colorado	18.9	Indiana	3.0
District of Columbia	14.3	Kentucky	2.7
Delaware	13.1	Alabama	2.6
Virginia	13.0	Wyoming	1.8
Hawaii	13.0	New Jersey	1.7
Oregon	11.2	Kansas	1.5
New Mexico	11.1	Wisconsin	0.9
Tennessee	10.4	Mississippi	0.4
Arkansas	9.9	Montana	0.4
California	8.7	South Dakota	0.4
South Carolina	7.8	Iowa	0.2

SOURCE: U.S. Dept. of Education, NCES, Common Core of Data surveys and State Public Elementary and Secondary Enrollment Model. (See reference table 5.)

Table B. Projected percentage decreases in public elementary and secondary school enrollment, by state: 2005 through 2017

State	Percent change	State	Percent change
Louisiana	-12.4	New York	-5.2
Vermont	-11.7	Massachusetts	-4.0
Rhode Island	-11.4	New Hampshire	-3.8
Maine	-8.0	West Virginia	-3.5
North Dakota	-7.6	Ohio	-3.2
Connecticut	-6.3	Pennsylvania	-2.0
Michigan	-6.2		

SOURCE: U.S. Dept. of Education, NCES, Common Core of Data surveys and State Public Elementary and Secondary Enrollment Model. (See reference table 5.)

Figure A-9: Tables A and B of Hussar and Bailey (2008).

Appendix D: Panelists and Contributors

Panelists (2010)

Daniel B. Carr, PhD
George Mason University

Linda W. Pickle, PhD
StatNet Consulting

Leland Wilkinson, PhD
Systat

Panel convened by National Institute of Statistical Sciences

Alan Karr, PhD, Chair

S.K. Kinney, PhD (Contributor)

Jamie Nunnelly, MA (Contributor)