How to Mine NCES Reports for Hidden Treasures

Dr. Barry Dupen, Purdue University, Fort Wayne

Dr. Dupen is an Associate Professor of Mechanical Engineering Technology at Purdue University Fort Wayne (PFW). He has nine years’ experience as a metallurgist, materials engineer, and materials laboratory manager in the automotive industry. His primary interests lie in materials engineering, mechanics, and engineering technology education. He is also an experienced contra dance caller.
How to Mine NCES Reports for Hidden Treasures

Abstract

The National Center for Education Statistics is a treasure trove of information on primary, secondary, and postsecondary education. NCES follows U.S. students from kindergarten through postgraduation employment, and reports the data in the annual Digest of Education Statistics, which is supplemented by detailed tables in spreadsheet form. The Digest discusses the ways in which many factors in higher education have changed over time; these factors include the popularity of different college majors, student diversity (e.g. age, gender, ethnicity & race, economic background), graduation rates, and employability of graduates by major. Some data, such as employability of 25-29 year olds with a bachelor's degree, have only been collected in very recent years, while other variables have been collected for decades. With the data, we can answer historical questions such as “how did the Great Recession affect college attendance and majors?” and student advising questions such as “given my interests and abilities, would an Engineering Technology degree be a good choice if I want to pay back my student loans by age 30?”

In addition, NCES publishes projected statistics for the next decade. Projections indicate a 13% increase in college enrollment in the next 11 years, with the highest growth rates amongst women, people of color, Hispanics, Asians, and students over 35. Professors and administrators in higher education can use this information for academic advising, for requesting resources, and for planning.

This paper will review some of the more interesting datasets available from NCES annual reports through the lens of the Baldrige Excellence Framework, which requires that data graphics show levels, trends, and comparisons (relevant benchmarks) over time. This paper will also discuss ways we can use NCES data to improve Engineering Technology education and advising.

Introduction

For a century and a half, the National Center for Education Statistics has been collecting, analyzing, and reporting data on the condition of education in the U.S. NCES produces several annual reports which are available at https://nces.ed.gov/surveys/annualreports/. These include:

- *Digest of Education Statistics* [1] is a huge report. The most recent edition is more than 1000 pages long, including text, graphs, and nearly 600 data tables. These tables are available online for every report going back to 1995, and recent tables are available in MS Excel format. Some data in the tables extends back to the middle of the 19th century, and a few tables include projections into the future. The Digest for a given calendar year is issued one to two years later. Tables are available online sooner than the Digest itself.

- *Projections of Educational Statistics* [2] contains about 35 pages of graphs and discussion, with more than 100 pages of appendices (data tables, statistical methodologies, etc.) The graphs typically report data from 2000 to 2014, with projections out to 2025. Data tables include additional historical data for previous decades.
The Condition of Education [3] contains more than 300 pages of discussion and graphs. Much of the data is drawn from the Digest. Many of the graphs show limited history (two years of data, either consecutive or separated by 5 or 10 years to show a two-point trendline). Reports are available online from 1989 to the present.

Trends in High School Dropout and Completion Rates in the United States [4] is a 50-page report (plus 30 pages of appendices and references) describing high school success and failure through the lenses of gender, race & ethnicity, income, disability, recency of immigration, nationality, and U.S. state. It addresses reasons why students drop out of high school. Most of the x-y graphs show history back to 1970. Reports are available online from 1983 to the present.

Status and Trends in the Education of Racial and Ethnic Groups [5] draws its data from the Digest. Many of the x-y scatter graphs show continuous curves without individual datapoints, while the bar graphs typically contain only two or three years' worth of data. Historical comparisons are complicated because the race categories changed in 2003. Also, prior to 2000, Asians and Pacific Islanders were counted together.

Indicators of School Crime and Safety [6] is a 130 page report (plus 105 pages of appendices and references) focusing on crimes in primary and secondary educational institutions. Previous editions of this report are available on the Bureau of Justice Statistics website [7]. Crime statistics for college campuses were added to this report starting in 2013; the current report includes college campus crime statistics from 2001 to 2014.

Some tables include continuous data ranging from the early 20th century to today, while other tables contain data on new measurables, such as unemployment rates and median incomes of 25 to 29 year-olds with bachelor's degrees [8] (reported in the Digest only since 2009). Another new measurable, “internet use of schoolchildren [9]” reported in Condition, had no meaning prior to 1990, when the World Wide Web was invented.

The nearly 600 data tables of the Digest are listed online in this format [10]:

To use the website, click on the chapter, then the relevant set of tables, then the individual table.
Table 102.70, below, provides information on poverty rates of primary and secondary schoolchildren from 2000 to 2015. The “web-only table” disclaimer means that this table does not appear in the Digest itself. Click on the table title, and the table appears on-screen.

To obtain data prior to 2000, return to the Digest homepage and select an earlier year.

The Digest includes graphs for a small fraction of the data, but since the most tables are available in MS Excel format, it is easy to create graphs as needed.
Baldrige Excellence Framework

For more than 30 years, the Malcolm Baldrige National Quality Award has identified and recognized excellence in business, education, health care, and nonprofit organizations. Data graphs in Baldrige Award applications must meet certain standards, as illustrated in this graph from the Baldrige Award instruction booklet [11]:

The annotations associated with this graph discuss several features:

1. Both axes and their units are clearly labeled.
2. Levels and trends are reported for a key performance measure. In the example graph above, the performance measure is the increase in math achievement.
3. Results are presented over time. Baldrige Awards are not granted to institutions having a single year of high performance; instead, a record of high level performance and improvement are required. Five years of data appear on this graph (two years for School C), plus two years of projections. Text accompanying this graph would have to explain the reasoning behind these projections, especially the large increase projected for School B.
4. Arrows indicate which direction is “good”. In the example graph, “good” is upwards; in a graph showing a negative trait such as absenteeism or drop-out rates, “good” would be downwards.
5. Appropriate comparisons are shown clearly. In this graph, three schools in a school district are compared with each other, with the district as a whole, with a comparable school district, and with national trends.

This last point emphasizes the importance of benchmarking. Continual Improvement Processes focus on improving performance from one year to the next, but performance excellence also requires knowledge of best practices and performance levels of comparable institutions. The NCES practices benchmarking by comparing performance of educational systems within the US and in other countries.

Population Size

Table 101.10 of the Digest [12] estimates the size of the U.S. resident population as a whole from 1970 to 2017 in ten age brackets, including 5- & 6-year-olds and 18- & 19-year-olds. Most
“traditional” students start college at age 18 or 19, therefore changes in the size of this population affect the size of incoming college freshman cohorts.

The 1995 version of this table [13] includes population estimates back to 1960, providing a 57 year history. Since all 18- & 19-year-olds were either 5 or 6 years old 13 years in the past, we can estimate the population of 18- & 19-year-olds 13 years into the future using the current population of 5- & 6-year-olds. The number of 18- & 19-year-olds today is about 12% greater than the number of 5- & 6-year-olds 13 years ago, due to immigration and other factors. Using this 12% correction factor to project into the future, the population of 18- & 19-year-olds over a 70 year span looks like this:

![Graph showing population projections from 1960 to 2030.](image)

The large peak in 1980 represents the tail end of the post-World War II Baby Boom. The big dip in the center represents Generation X, while the second peak in 2009 represents the Echo Boom (also known as Millennials or Generation Y). College attendance rose during the Great Recession, and has declined in the years since, owing to changing economic conditions and an 8% drop in the population of traditional-age freshmen from 2009 to 2017.

**Students of the Future**

In the U.S., college students of the distant past were predominantly full-time, upper- or middle-class White males straight out of high school. In 1870, 21% of U.S. college students were female (and these women were all undergraduates); by 1980, female college students outnumbered male college students, and today, 56% of college students in the U.S. are female [14]. *Projections of Education Statistics to 2025* predicts enrollment of college students will increase 17% for females and 13% for males in the next decade [15], and the *Digest* projects 59% of college students will be female by 2025 [16].

In 1970, 28% of college students were nontraditional age (over 25); since 2000, they have comprised about 40% of the college student population, and this fraction is projected to remain unchanged through 2026 [17]. The number of nontraditional student population is growing, as shown in the *Digest* graph below [18].
The following graph from Chapter 4 of *Status and Trends* shows that high school graduation rates have improved since 1990 across all racial and ethnic groups, with the greatest improvement among Hispanic students [19].

Missing from the preceding graph are a descriptive label for the vertical axis and individual datapoints. Also missing are Asian & Pacific Islanders, a fourth major racial/ethnic group reported elsewhere in *Status and Trends*; as a consequence, it appears that White students are the highest-performing racial and ethnic group. By adding more history from the Digest [20] and by shifting the origin of the vertical axis, we can easily see that White and Hispanic students made little progress from 1972 to 1990, a period when Black students made significant progress. Asian & Pacific Islander students (denoted “Asian” on the graph) were not included in this dataset until
1990; they outperform White students in high school completion and in many other measures of academic success.

As a greater fraction of Hispanic students successfully completed high school in the past 15 years, they became eligible for college. The following graph from Chapter 5 of *Status and Trends* shows a growth in undergraduate college enrollment for all groups from 2000 to 2010, followed by a decline in the numbers of Black and White college students, and an increase in the numbers of Hispanic college students [21].
Again, no datapoints are marked, so it is not clear whether the graph uses annual data. A new census category, “two or more races,” was included in the database starting in 2010.

We can improve on the previous Status and Trends graph by using data from five editions of the Digest. Each Digest includes annual data for the most recent years, while older data is provided only once or twice per decade [22]. By using multiple overlapping data sets, annual data can be plotted back to 1988, with four additional datapoints in 1976, 1980, 1984, and 1988. We can also extend the graph into the future, using predicted undergraduate enrollments for the next decade (solid lines to the right). Projections predicts the following fall enrollment for undergraduate and graduate programs combined: a 30% increase in total enrollment of Hispanic college students from 2015 to 2025, a 20% increase in Black students, a 19% increase in Asian students, and a 3% increase in White students, for an overall increase of 12% [23]. These predictions are based on the expected change in the population of people aged 18-29, and do not include the costs of a college education or future economic conditions. Assuming the overall college enrollment increases are reflected in undergraduate programs, data and projections for undergraduate programs appear as follows:

Projections indicate that college enrollment will continue to grow more diverse in the next decade.

**Women in Engineering and Engineering Technology**

Scholars have noted a decline in the ratio of females to males in U.S. engineering and engineering technology majors in recent years [24], [25], [26]. Table 325.45 [27] of the Digest confirms this decline in the portion of women earning B.S. degrees in these fields in the first decade of the 21st century; however, the following graph offers hope. Since 2009, the percentage of female B.S. graduates has increased from 16.5% to 18.7%. The percentage of M.S. and Ph.D.
degrees earned by women in these fields has grown steadily since the early 1970s. In 2015, 25.2% of M.S. degrees and 23.2% of Ph.D. degrees in engineering and engineering technology were earned by women.

In the spirit of the Baldrige criteria, this graph should be compared with benchmarking data. For example, we could compare the U.S. to similar nations. Comparable data for other countries is not easy to obtain. Table 603.60 [28] of the Digest lists the percentage of postsecondary degrees awarded to women in eight fields for the 33 OECD nations plus 5 additional nations. “Postsecondary” lumps A.S., B.S., M.S., and Ph.D. degrees together, and the relevant category is “Engineering, Manufacturing, & Construction.”

Data for gender ratios of engineering graduates in Canada is available online from 2008 to 2015 [29], [30]. The following graph shows similar trends between U.S. and Canadian gender ratios over recent years. The Canadian data fluctuates more than the U.S. data because of the smaller absolute number of engineering graduates in Canada. (Although Canada's population is 11% of the U.S. population, Canada's annual output of B.S. engineering majors is 15% of the U.S. number [31].)
Another type of benchmark compares similar academic fields. The graph below compares engineering and engineering technology with two other historically male-dominated, mathematics-based fields of study. Computer and information sciences saw great progress in its percentage of female B.S. graduates up to 1985, then began a quarter century of decline, mirrored by an increase in physical sciences and science technologies. Since 2009, physical sciences has seen a slight decline, computer and information sciences has been steady, while engineering and engineering technology has been growing.
At the master's degree level, engineering and engineering technology lag behind the other two fields. Computer and information sciences saw a peak in the percentage of female M.S. graduates around 2000, then settled back to prior levels.

All three fields have seen steady progress at the Ph.D. level, but physical sciences and science technologies have seen the greatest growth.

**Employment Prospects**

In addition to teaching responsibilities, I serve as an advisor to incoming freshmen into the Mechanical Engineering Technology program at Indiana University – Purdue University Fort Wayne. Prior to the economic recession of 2008, new students and their parents rarely asked me about employment and salary levels after graduation, but during and since the recession, this question has become common. Fortunately, employment prospects in northern Indiana are very
good for Mechanical Engineering Technology graduates. Chapter 5 of the *Digest* discusses two major education outcomes: participation in the workforce and annual earnings [32]. Contributing factors include gender, race and ethnicity, age, and level of education. There are no great surprises in the five graphs presented in Chapter 5; here are some broad conclusions:

- People with more formal education earn more and have lower rates of unemployment than people with less formal education.
- Men are paid more than women in the same field with the same level of education.
- Some college majors (engineering, information systems, health professions) earn significantly more than others (humanities, social work, theology).
- From 1995 to 2014, the earnings of college graduates were largely unchanged when measured in constant dollars, while high school dropouts saw a 9% reduction in real pay.

The 31 tables in this chapter contain more detailed data. Table 501.80 [33] lists the unemployment rate for the past 40 years for people in various age categories. In the graph below, the category “some college” includes people who earned an Associate's degree and people who started but did not finish a Bachelor's degree.* The peak at the right shows that the Great Recession of 2008 was particularly hard on high school drop-outs. This graph can be a useful tool for persuading high school and college students to complete their education.

Table 505.10 [34] of the *Digest* lists employment statistics for people who are 25 to 29 years old and hold bachelor's degrees, ranked by 55 college majors, in 2010 and 2015. For each major it provides the numbers of people, percentage distribution, unemployment rate, and median annual income. This table was first reported in the 2010 *Digest*, covering data from calendar year 2009.

Prior to the 2015 *Digest*, median annual earnings were reported in current dollars only, so there was no adjustment for inflation, making year-to-year comparisons difficult. Starting in 2015, the *Digest* also includes a comparison between STEM and non-STEM fields.

---

* This term is used differently from one table to another within any given NCES report. “Some college” can either mean “college drop-outs only” or “college dropouts plus people who have earned an Associate's degree.” When reading a *Digest* table, be sure to check the definitions of any terms in the annotations below the table.
Plotting the data from Table 505.10 of the Digest shows us that 25 to 29 year olds holding bachelor's degrees in STEM fields and professions generally have lower unemployment, higher salaries, or both, than people holding bachelor's degrees in non-STEM fields. Although their salaries are low, education majors have the lowest unemployment rate of all majors, perhaps because teaching jobs have the built-in job security of an academic year. Liberal arts & humanities majors suffer the worst unemployment rate of all disciplines, equal to the unemployment rate of college dropouts.

The previous graph shows a snapshot in time: income and unemployment in 2015. In the spirit of the Baldrige criteria, the following graph shows how income and unemployment rate have changed over time for eight of the majors, from 2009 to 2015. The paths indicate changes from one calendar year to the next; the arrowheads point to 2015. Income data from Table 505.10 and its predecessors was adjusted for inflation using the CPI-U [35] published by the Bureau of Labor Statistics, and is reported in 2015 constant dollars. All eight majors showed an
improvement in the unemployment rate and a decrease in income as the nation emerged from its deepest economic downturn since the 1930s. Health professions saw a large reduction in pay. Fine and commercial art majors were more vulnerable to the Great Recession than any other major in 2009, with twice the median unemployment rate of college graduates. However, by 2015 they were more employable than liberal arts & humanities majors.

When incoming freshmen are unsure between two or more majors, I share these graphs with students and their parents to help guide the discussion.

**Conclusions**

The six NCES annual reports and their many tables contain information which can help professors and their employers to advise students and to plan for the future, in terms of
demographics, demand for degree programs, funding patterns, expenditures, and so forth. Some of the key points are:

- Most of the data used in the various NCES annual reports can be found in Digest tables.
- Digest table numbers often change from one edition to the next, and sometimes tables move from one chapter to another with edition changes. It is useful to investigate older editions for more complete historical data.
- Many of the tables containing expenditures, prices, or earnings provide numbers in current dollars. In order to compare money over a long timespan, these numbers should be adjusted for inflation using the CPI-U from the Bureau of Labor Statistics.
- Age groupings vary from one table to the next. For example, some tables categorize ages 25-29 and ages 30-34 separately, whereas other tables combine them as ages 25-34.
- Ethnic and race categories have changed over time, making historical comparisons difficult. Sometimes groups are separated (e.g. Asian is now distinct from Pacific Islander; these groups were once categorized together).
- Digest tables are available online 1 to 2 years before the full report is completed.

The graphs in NCES reports tend to lack critical elements, such as clearly-labeled axes (with the “Good” direction marked), data points, and relevant comparisons. Fortunately, the data files are freely available in Excel format, and with a little manipulation, they can tell a great story in graphical form.
References


[16] 2015 Digest, p.461, Table 303.10.


[23] Projections, Table 19, p.66.


[31] 2015 Digest, Table 322.10 p.634, and Canadian Engineers 2011-2015, Table UD.1.1.
[33] 2015 Digest, p.798.
[34] 2015 Digest, p.816-817.