

2014 Science and Engineering Indicators State Data Tool

April 2014

Nirmala Kannankutty
Senior Advisor, Office of the Director
National Center for Science and Engineering Statistics
National Science Foundation
nkannank@nsf.gov

Technical Notes

The 2014 Science and Engineering Indicators (SEI) State Data Tool contains trend data for most indicators. These data are available for download within the data tool and from the State Data Tool download page (<http://www.nsf.gov/statistics/seind14/index.cfm/state-data/download.htm>).

1. Standard Errors

The SEI State Data Tool contains data compiled from a large number of sources, which can be categorized as follows:

- *Data based on censuses.* These are complete population counts; therefore, there is no standard error associated with the estimate. Data or tables where standard errors are not applicable are labeled “na”.
- *Data based on samples.* Standard errors for estimates, where available, are provided by the source. The National Assessment of Education Progress (NAEP) data sets are the only data sets with complete standard errors for the time periods included in the state data tool and the chapter. Standard errors are incomplete in other data set based on samples: for example, standard errors are provided only for the past 3–6 years of the Local Area Unemployment Statistics (LAUS) survey or the Occupational Employment Statistics (OES) survey. The Business R&D Survey (BRDIS) data set has associated standard errors for its values, but some historical values of standard errors are not available due to updates to the estimates but not to the standard errors.
- *Data based on statistical models.* Standard errors cannot be provided for some estimates due to the estimating techniques of the data source (for example, gross domestic product [GDP] data and Census population estimates). Data or tables where standard errors are not available are labeled “NA.”
- *Data derived directly from the source data set.* (e.g., S&E doctorate holders as a percentage of the workforce). For data series where the standard error information for the source data is available, approximation formulas for combining sampling errors were used. Because the source data used to derive these estimates are from different independent samples, there is no covariance term included in the formulas.

Standard error tables are provided for download for all state data tool data where the standard errors are appropriate and available (<http://www.nsf.gov/statistics/seind14/index.cfm/state-data/download.htm>). In some cases, standard error information was not available for a data series. This is noted in the tables or on the website.

The following formulas were used to estimate standard errors for derived data series.

Sums and aggregates

Where available for aggregate estimates, such as the total for the United States, sampling errors were collected for the aggregate estimate as provided by the source.

In a few cases, aggregate estimates were calculated from individual parts of the aggregate and therefore, sampling errors also had to be calculated based on the individual parts of the aggregate. It was assumed that the covariance between the individual parts was negligible.

This formula was used where applicable for such roll-ups as national values or occupation categories (e.g., computer science).

$$SE (X + Y + Z \dots) = \sqrt{(SE(X))^2 + (SE(Y))^2 + (SE(Z))^2 \dots}$$

Difference

This formula was used when comparing the difference between two sets of estimates.

$$SE (A - B) = \sqrt{(SE(A))^2 + (SE(B))^2}$$

Quotient

(assuming X and Y are uncorrelated, using the first order Taylor series expansion, which is an approximate but widely used and accepted approach)

$$SE \left(\frac{\bar{X}}{\bar{Y}} \right) = \sqrt{\frac{(SE(\bar{X}))^2}{\bar{Y}^2} + \frac{\bar{X}^2}{\bar{Y}^4} (SE(\bar{Y}))^2}$$

Other formulas used

Data from the Occupational Employment Statistics Survey and some Business Research and Development and Innovation Survey was only available as the relative standard error (RSE)

$$RSE = \frac{\text{standard error}}{\text{estimate}} \times 100\%$$

Therefore, to transform the RSE to standard error, the following equation was used:

$$\text{standard error} = \frac{\text{estimate} \times RSE}{100\%}$$

2. Constant Dollar Data

The SEI State Data Tool presents data as current dollars. To facilitate comparisons over time, the data tool also has an option for presentation of the information as constant dollars in the table and chart views. The data tool uses constant 2005 dollars based on the Gross Domestic Product (GDP). The specific values are from June 2013 adjusted to a calendar-year basis, as prepared by the Bureau of Economic Analysis. The constant dollar adjustment is available in the state data tool for all financial indicators, except for Indicator 8-9 (Public school teacher salaries).

Table 1 provides the GDP price deflators used in the state data tool. These price indices are for the national GDP, and are not adjusted for states. The state data tool tables that are available for download present information as current dollars only. The data in Table 1 can be used to replicate the constant dollar information in the state data tool. It may also be applied to the standard error tables, as applicable.

Table 1
Calendar-year price deflators for the 2014 State Data Tool

Year	GDP price deflator (chained) 2005 dollars
1990	0.7226
1991	0.7482
1992	0.7660
1993	0.7829
1994	0.7994
1995	0.8161
1996	0.8316
1997	0.8463
1998	0.8558
1999	0.8684
2000	0.8872
2001	0.9073
2002	0.9220
2003	0.9414
2004	0.9679
2005	1.0000
2006	1.0323
2007	1.0623
2008	1.0858
2009	1.0953
2010	1.1099
2011	1.1336
2012	1.1539