NSSE Multi-Year Data Analysis Guide

About This Guide

Questions from NSSE users about the best approach to using results from multiple administrations are increasingly common. More than three quarters of NSSE participating institutions have administered the survey more than once. Some look for changes in the way their current students are engaged, some track possible trends, and others evaluate specific campus initiatives. Regardless of the reason, a variety of questions quickly emerges, such as:

- Can existing reports be used to evaluate changes from year to year, or should multi-year data sets be merged for separate analyses?
- I know the NSSE survey has changed over time, so how can I quickly identify comparable survey items and benchmarks?
- What is an appropriate method for determining if there has been a meaningful change between years?

The answers to these questions may be complicated because—especially in the early years—NSSE survey items as well as the construction of institutional reports have undergone several improvements and adjustments. Thus, the purpose of this document is to provide researchers with helpful resources, information, and suggestions for suitable approaches to NSSE multi-year analysis and to strengthen the validity of any final conclusions drawn.

We recommend the following tasks for any multi-year analyses of NSSE data:

1. Identify and focus on specific questions. Are you looking for general shifts in engagement, or something more specific to assess a particular campus initiative?

2. Select and employ appropriate methods of analysis. Determine your analytic approach and the statistical tests that would best identify changes from year-to-year, and the criteria that might be used for evaluating the magnitude of the change.

3. Attend to data quality for each year in the analysis. You should consider, for example, the response rate, total count, sampling error, representativeness, and circumstances surrounding the administration for each year included in your analysis.

4. Take into account changes in NSSE items and reports across years. Determine which NSSE items changed over the years and how your reports might differ in terms of weighting, construction of benchmarks, etc.

5. Merge multiple years of data. We provide a link to an Excel workbook to assist with identifying comparable NSSE variables and a SPSS syntax file and script to merge data sets.

The five tasks, explained in more detail in the following pages, are presented in this order as our recommendation, although it is probably not necessary to accomplish them sequentially. For example, the first thing you might do is merge your data before diving into the analysis.
Task 1: Identify and Focus on Specific Questions

Your research questions should be specific, answerable, and relevant to campus priorities. Tying them to current campus issues, initiatives to be evaluated, or policy decisions increases the study’s chances of being well-received and actually used. Below are three possible research questions for the fictional “NSSEville State University” that represent typical approaches to multi-year analysis.

1. “How stable was the level of academic challenge at NSSEville State University from 2007 to 2008?”

    In the absence of a major campus initiative, one would expect NSSE results to be relatively stable from one year to the next. Working from the assumption that little has changed on campus, institutions may conduct multi-year analyses to confirm stability, an indicator of their data’s reliability. When results are consistent between years, a campus can be more confident in the accuracy of population estimates found in the NSSE reports.

2. “Given the implementation of initiative X in 2006-07, how much did the level of active and collaborative learning at NSSEville State University increase from 2006 to 2008?”

    With budgets tightening and calls for increased accountability from both internal and external stakeholders, many schools use NSSE to gauge whether new initiatives are associated with higher levels of student engagement. With data collected prior to and after the implementation of an initiative or program on campus, a multi-year analysis can mirror a pre-test/post-test research design by comparing engagement levels before and after implementation.

3. “Given several campus initiatives aimed at increasing contact between faculty and students over the past decade, what trends are apparent in student-faculty interaction at NSSEville State University from 2001 to 2008?”

    Institutions with more than two years of data may be interested to see if results indicate any trends (e.g., trending up or staying consistently high). Keep in mind that the number of data points needed for a trend analysis is subjective and, in part, relies on institutional context and the particular questions being investigated.

Some researchers may be tempted to examine NSSE data for any type of results or findings without determining specific questions up front. Though such exploratory studies are possible, we generally advise against this approach. If you do proceed with this type of exploration, we suggest that you set limits to focus the effort, such as organizing the analysis by content area (e.g., student-faculty interaction or high-impact educational experiences).

Task 2: Select and Employ Appropriate Methods of Analysis

Approaches to Multi-Year Analysis

Figure 1 illustrates comparisons that can be made with multiple years of NSSE data. The paths labeled A or B require multiple years of data. The paths labeled C require a single year of data but involve comparing students from different class years. We generally recommend cohort comparisons (A), but under the correct conditions longitudinal comparisons (B) may be appropriate. We discourage comparisons between first-year students and seniors from the same year (C). In addition, with normative comparisons appearing in NSSE reports, some analysts may attempt to track and analyze changes or stability in these comparisons. We also discourage this type of approach because peer group composition often changes from year-to-year and peer group population estimates have not been consistently derived.

1 Test-retest analyses using 2002-2003 and 2004-2005 NSSE benchmarks indicate a high level of reliability. Correlations ranged between .78 and .92.

2 The first year and the senior year are different educational contexts, with different engagement patterns. A student’s engagement should not necessarily be expected to increase over the years, but rather can be high or low depending on the circumstances of the student’s activities. An equally important consideration when comparing differences between first-year and senior students is whether they are, in fact, consistent populations. That is, at some institutions as many as a quarter to a third of all first-year students will not persist to the second year, and even fewer make it to the senior year. The senior cohort includes only those students who “made it” plus, potentially, a significant proportion of transfer students and other part-time students who may have been studying for six or more years before making it to senior status. For these reasons, making comparisons between first-year students and seniors is challenging.
Cohort comparisons (recommended)

This approach compares your estimate of first-year student engagement in one year with your estimate of first-year student engagement in another year (likewise for seniors). In Figure 1, this relationship is represented by the A arrows. For example, you can examine how much first-year student engagement has changed over three years (e.g., 2005, 2006 and 2007). You may want to compare at least two years of data with a “base” year before concluding a systematic change has occurred.

Longitudinal (first-year to senior) analyses

This approach tracks the engagement of a panel of first-year respondents as they become seniors three years later, represented in Figure 1 by the B arrow. The objective is to examine how the senior-year experience compares with the first-year experience and, to the extent possible, estimate changes over time in the nature and frequency of student engagement with the institution’s resources for learning. Schools may be particularly interested in assessing whether or not a program initiated after the first year of data collection has had any observable effects in the students who are now seniors.

The panel data have some advantages. Individual characteristics such as sex, race, family background, ability, and pre-college experiences are held constant. Thus, interpretation of results allows you to focus on environmental factors that may influence the nature and frequency of student engagement in various areas.

However, several caveats apply with the longitudinal approach. First, like most panel studies, attrition from the original cohort is a concern. For various reasons, only a fraction of first-year respondents will become senior respondents for many institutions. Second, transfer students are excluded from the analysis, although they may constitute a large portion of seniors. Additionally, a notable proportion of first-year students do not persist to the senior year at some schools or take differing amounts of time to do so. As noted with the single year cross-sectional model (C), the longitudinal panel design faces the same problem with comparing NSSE results between two different populations and educational contexts.

Methods for Identifying Significant and Meaningful Change

Once you have decided upon your research questions and your approach, you need to choose analytic methods that will help identify both statistically significant and meaningful changes from year-to-year. You may find many statistically significant differences (especially if your sample is large), but it is important to determine if the differences are large enough to warrant serious attention. A number of approaches help in this respect. Although not a step-by-step guide, the following presents several useful and statistically straightforward ideas for looking at change, stability, or trends. Keep in mind that these methods are not mutually exclusive and can be used together to test your year-to-year changes in results. Refer to Table 1 for a summary of each method’s characteristics.

Statistical Analyses

- **t-test.** To determine if a statistically significant difference exists on a single measure between two years, employ t-tests. This test can tell you the likelihood that any differences between two years of data occurred by chance. Unfortunately, it does not speak to the magnitude of the difference. Use this test as an initial step to determine what items or benchmarks should be reviewed for change. Although best used with interval-level data, as with NSSE benchmark scores, many use this robust test (and other similar tests) with ordinal-level data (e.g., means of individual survey items that use a four category ordered response set as seen on NSSE).

- **ANOVA.** Analysis of Variance provides a very good approach for reviewing a large number of survey items or benchmarks for which three or more years of results exist. A finding of no significant difference between years can be used to conclude that the data are relatively stable. Alternatively, if this test identifies significant differences between years, run post hoc tests such as Tukey or Scheffé to help identify where the differences exist.

- **Regression.** Multivariate statistical models can also address questions regarding significant changes between years, either using individual survey items or benchmarks. By creating a set of dummy variables for the administration year in your regression and using the base year as the reference group, you can determine if there is any statistically significant difference between the base year and...
subsequent results. If no statistically significant coefficients emerge, you can conclude results have remained stable relative to the base year. In addition, assess trends by analyzing the size of “year” coefficients in sequential order; coefficients may consistently increase from year to year, decrease, or stay stable. Furthermore, using regression analysis with student characteristics as controls provides more sophisticated information for evaluating the source of any changes from year to year. With controls in place, you will have more evidence that changes may be attributed to such things as campus initiatives and not changing student demographics.

Descriptive Analyses

- **Effect Size.** After finding statistically significant differences between years with various benchmarks, you will need to assess the meaningfulness of any apparent changes. Effect size statistics are often used to assess the magnitude of any change. Cohen’s $d$, the effect size calculation most often used at NSSE, is based on the mean difference between two years divided by the pooled standard deviation. An alternative way to estimate effect size would be to run regression models with standardized dependent variables and include administration year as an independent variable (as described above in the section on regression analysis). Regardless of your approach, changes between years can be interpreted in terms of standard deviation units. Using effect size statistics requires the analyst to establish criteria for determining whether a meaningful change has occurred. For additional information about interpreting effect sizes, see *Contextualizing NSSE Effect Sizes: Empirical Analysis and Interpretation of Benchmark Comparisons*, at nsse.iub.edu/pdf/effect_size_guide.pdf.

- **Percentage Change Using Dichotomized Variables.** Having multiple response options and a large number of survey items presents significant challenges for analysis and reporting. One way to handle this is to dichotomize survey variables by combining, for instance, “very often” and “often” into one category (“frequently”) and “sometimes” and “never” into another (“rarely”). If pertinent to your particular research question, you might also create a variable that specifically identifies those who say they “never” did something. SPSS syntax is available on our Web site to assist you with either approach. After you recreate variables, develop a table showing the percentage of respondents by year reporting they “frequently” or “never” do something and establish criteria for meaningful change. For example, you can use a minimum change of four percentage points between two years before declaring an upward or downward change. Because this approach does not test for statistical significance and uses a more subjective evaluation, you may want to take into account sampling error statistics to establish your criteria. The greater the sampling error, the more conservative you will want to be with establishing criteria for change.

Regardless of the method, as with any statistical analysis, there will be uncertainty with your findings. Given this uncertainty, we recommend triangulating all available data sources before implementing any policy change based on a NSSE multi-year analysis. If possible, NSSE should be one of several tools employed to assess change on campus.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Considerations for Multi-Year Analytical Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>t-test</td>
</tr>
<tr>
<td>How many years of data required?</td>
<td>2</td>
</tr>
<tr>
<td>Determines statistical significant difference?</td>
<td>Yes</td>
</tr>
<tr>
<td>Determines magnitude of difference?</td>
<td>No*</td>
</tr>
<tr>
<td>Are statistical controls possible?</td>
<td>No</td>
</tr>
</tbody>
</table>

* Determining the magnitude of the difference is possible but requires additional manipulations to variables of interest, inclusion of ad hoc software commands, or additional calculations.
Identifying Trends

If you plan to identify trends, you will have to decide what type of trend is most interesting to your audience. Typically, upward and downward trends interest audiences most, however consistently high or low results may warrant identification as well. As with other methods, establishing specific criteria for “high” and “low” will be important and depends in large part on campus context.

Another issue that emerges soon after you begin reviewing trend lines is whether or not you require successive year changes to label something as trending upward or downward. Create general rules for what constitutes a trend, taking into account such things as data plateaus. Establishing a clear picture from the start about what a trend line looks like will help ensure consistency throughout your analysis.

Task 3: Attend to Data Quality for Each Year in the Analysis

Evaluating Your Data

Some NSSE administrations yield more precise population estimates than others. Before embarking on your analysis, review the quality of your data for both first-year and senior respondents in each year. NSSE provides various pieces of information to help with this evaluation. The Respondent Characteristics report shows response rates and sampling errors, both of which are based on sample size. If one or more of your administrations showed very low response rates, you may want to consider excluding it from your analysis. Sampling error is an estimate of how much your respondents could differ on survey measures from the entire population of students at your institution. Smaller sampling errors (e.g., +/- 3% or 5%) are preferred. Data with larger sampling errors (such as +/-10%) need not be dismissed off hand, but any results using them should probably be interpreted more conservatively.

Each administration’s Respondent Characteristics report also details the student characteristics of your sample. Examine the data to see if there are major differences between your sample and the population. If you see significant differences, conduct additional investigations into whether NSSE results vary by these same characteristics. Even though NSSE provided weights based on gender and enrollment status in your data sets starting in 2005 (and can supply weights from earlier years upon request), there may be grounds to develop your own weights based on other student characteristics that appear to have a greater impact. Taking this additional step will help you determine whether or not changes in your data can be linked to such things as changing student demographics or campus initiatives. If you are interested in comparing particular subgroups over time, be aware that a small number of respondents in a given subgroup can prevent you from making reliable conclusions.

Choosing a Base Year

Determining if your data have changed between years or if trends are present requires that you identify a stable base year, typically the first year used in your study. For this reason, double check the data quality of your base year. However, we do not recommend using the 2000 version of NSSE as a base year due to the number of changes made since that administration. Lastly, you may find it beneficial to combine administrations to form a more reliable baseline.

Task 4: Take into Account Changes in NSSE Items and Reports Across Years

Should you analyze reports or data sets?

Some have asked if it would be appropriate to simply track results that appear in your standard NSSE reports, in place of the potentially arduous task of working with your raw data. Currently, given changes over time in how NSSE reports have been developed, the cleanest approach would be to merge the data sets and generate comparable statistics across multiple administrations. As shown in Table 2, changes to how weights have been applied and the types of respondents used to develop institutional estimates are two primary sources of variation. To produce comparable statistics across several years, we recommend using all randomly sampled respondents and adjusting for non-response with ‘weight1’ (or ‘stuwt2’ in 2003 and prior years). Finally, a major benefit of working with the data includes the flexibility to conduct tests to identify statistically significant differences between years.

Working with Benchmark Scores

In 2004, NSSE began calculating benchmark scores at the student level before aggregating them to the institution level. Thus, to compare benchmark results across years, be advised that benchmarks appearing on NSSE reports during the years 2000 through 2003 are not comparable with those produced since 2004.

---

3 To select on randomly sampled respondents, use the variables ‘smpl01’ (2004 and earlier) and ‘smpl05’ (starting in 2005). Values of ‘1’ and ‘2’ should be used with ‘smpl01’ and ‘1’, ‘2’, and ‘3’ with ‘smpl05’.
Change to Enriching Items Response Set in 2004

The response set for several items used in the creation of the Enriching Educational Experiences benchmark, such as participation in a learning community or completion of a study abroad program, also changed significantly in 2004. The result is that Enriching Educational Experiences scores through 2003 cannot be compared with 2004 and more recent data. For similar reasons, the only way to compare the early years of Student-Faculty Interaction with scores since 2004 is to create a new variable (named ‘SFC’ in data sets) by removing one of the survey items (research with faculty).

As shown in Table 3, the following benchmarks can be analyzed regardless of administration: ‘ACa’ (Level of Academic Challenge), ‘ACL’ (Active & Collaborative Learning), ‘SFC’ (Student Faculty Interaction—excluding the research with faculty item), and ‘SCE’ (Supportive Campus Environment). As previously mentioned, starting in 2005, benchmark scores and weights were included in your data file; institutions can obtain prior year scores and weights by contacting NSSE.

Since 2008 each participating institution receives a Multi-Year Benchmark Report within their NSSE Institutional Report to facilitate within-institution benchmark comparisons across years. The multi-year report presents comparable benchmark scores and relevant statistical information for each year the institution participated in NSSE. Using statistics on the multi-year report, institutions can compute simple $t$-tests and effect sizes across years without using raw data files from different years. To learn more, see the NSSE Institutional Report Web site at: nsse.iub.edu/links/institutional_reporting.

Task 5: Merge Multiple Years of Data

Preparing a NSSE multi-year data set for analysis involves identifying variables that have not changed over the years and then merging the cases from all years into a single file. We have developed a couple of tools to assist you with this process.

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean &amp; Frequency Reports</th>
<th>Benchmark Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Respondent Type</td>
<td>Weight Used</td>
</tr>
<tr>
<td>2003 and earlier</td>
<td>All randomly selected plus targeted</td>
<td>None</td>
</tr>
<tr>
<td>2004</td>
<td>All randomly selected plus targeted</td>
<td>None</td>
</tr>
<tr>
<td>2005</td>
<td>All randomly selected plus targeted, if requested</td>
<td>None</td>
</tr>
<tr>
<td>2006</td>
<td>All randomly selected</td>
<td>Weight3</td>
</tr>
<tr>
<td>2007- Present</td>
<td>All randomly selected</td>
<td>Weight1</td>
</tr>
</tbody>
</table>

* This information reflects how NSSE developed population estimates for institutions, not peer groups.
** Starting in 2004, NSSE used a new calculation for institution-level benchmark statistics.
NSSE began including benchmark scores and weights in data sets in 2005. Contact NSSE if you need benchmark scores from earlier years.

**SPSS syntax and script to merge data sets**

Once you identify a list of variables to use in your analysis, go to the SPSS command syntax file to merge your data sets in an automated fashion. The syntax file not only provides step-by-step instructions on how to run the script to merge your data, but it also helps with finalizing some important details of your multi-year data set.

The *NSSE Year-to-Year Survey Variable Tracking Sheet* and SPSS files are located on the NSSE Web site at nsse.iub.edu/html/analysis_resources.cfm.

**Case Study: A Multi-Year Analysis at NSSEville State University**

This case study from the fictional “NSSEville State University” is based on our observations of how institutions typically analyze their multi-year data and illustrates some of the recommendations in this guide. So, consider the following situation:

> In 2003, NSSEville State University found that their students interacted with faculty significantly less often than students attending NSSEville’s peer institutions. Given these results, the provost initiated first-year learning communities with seminars taught by senior faculty, a faculty fellows program in the residence halls, and a new student-faculty research program for upper-division students. Following NSSEville’s 2007 NSSE administration, the provost asked the director of Institutional Research to analyze the student-faculty interaction benchmark to test whether the campus initiatives had a substantial effect.

### Table 3

<table>
<thead>
<tr>
<th>NSSE Benchmark Variables by Year</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004 to present</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Academic Challenge</td>
<td>ACa</td>
<td>ACa</td>
<td>ACa</td>
<td>ACa</td>
</tr>
<tr>
<td>Active &amp; Collaborative Learning</td>
<td>ACL</td>
<td>ACL</td>
<td>ACL</td>
<td>ACL</td>
</tr>
<tr>
<td>Supportive Campus Environment</td>
<td>SCE</td>
<td>SCE</td>
<td>SCE</td>
<td>SCE</td>
</tr>
<tr>
<td>Student-Faculty Interaction</td>
<td>S01</td>
<td>S01</td>
<td>S01</td>
<td>≠ SFI</td>
</tr>
<tr>
<td>Student-Faculty Interaction</td>
<td>SFC</td>
<td>SFC</td>
<td>SFC</td>
<td>SFC</td>
</tr>
<tr>
<td>(without Research with Faculty item)</td>
<td>E01</td>
<td>≠ E02</td>
<td>E02</td>
<td>≠ EEE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Enriching Educational Experiences</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004 to present</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Tips on Developing Comparable NSSE Statistics**

**Tip #1: Use All Randomly Selected Cases**
- For 2004 and earlier, use values of ‘1’ and ‘2’ for “smpl01”
- For 2005 to the present, use values of ‘1’, ‘2’, and ‘3’ for “smpl05”

**Tip #2: Use Weights**
- “stuwt2” for 2001 to 2003 administrations
- “weight1” for 2004 to current administration

**Tip #3: Exclude Ineligibles**
- For 2005 to the present, use values of ‘1’ for “inelig” to ensure only eligible respondents are included
The following summarizes the approaches taken to address this problem, corresponding to the five tasks described above.

1. **Identify and focus on specific questions**
   - The analyst focused the research questions to ask, “Did the average student-faculty interaction score significantly and substantially increase between 2003 and 2007 for both first-year and senior students?”

2. **Select and employ appropriate methods of analysis**
   - It was determined that a cohort comparison approach, rather than a longitudinal analysis, would be appropriate to answer the research question.
   - After merging the two years of data (see 5. below), the analyst ran t-tests for statistically significant differences between 2003 and 2007 student-faculty interaction scores for both first-year and senior results.
   - Upon discovering a statistically significant increase in scores, the analyst calculated Cohen’s $d$ effect sizes to assess whether or not the magnitude of the change was meaningful. Reference material on NSSE’s Web site helped the analyst establish criteria for a meaningful change.
   - To be sure that differences in student demographics could not account for the shift in scores, the analyst then developed regression models that included a dummy variable for year and controls for various student characteristics. The dependent variable, student-faculty interaction, and all continuous independent variables were standardized so that the regression coefficients could be interpreted as effect sizes.
   - The analyst employed weights provided by NSSE to adjust for the disproportionate representations of gender and part-time status in the data set.
   - The analyst also gathered alternative sources of information (campus focus groups, course evaluations, etc.) for evaluating the research question at hand so the NSSE findings could be confirmed.

3. **Attend to data quality for each year in the analysis**
   - Before running the analysis, the analyst reviewed Respondent Characteristics reports from 2003 and 2007 to evaluate data quality.
   - Upon reviewing response rates, the number of respondents, and sampling error statistics, the analyst felt comfortable that population estimates using NSSE data were sufficiently precise.

4. **Take into account changes in NSSE items and reports across years**
   - Realizing that 2003 and 2007 student-faculty interaction scores (SFI) cannot be directly compared, the analyst correctly used the variable ‘SFC’ (see Table 3) to analyze the changes between 2003 and 2007.
   - The analyst identified that the weight and sample type variables were named differently in 2003 and 2007 (‘stuwt2’/’weight1’ and ‘smpl01’/’smpl05’, respectively).
   - In order to make the two years as comparable as possible, the analyst included only randomly selected cases in the analysis (i.e., respondents with ‘smpl01’ values of 1 and 2 in 2003 and ‘smpl05’ values of 1, 2, and 3 in 2007).

5. **Merge multiple years of data**
   - Unable to find ‘stuwt2’ or ‘SFC’ in their 2003 data file, the analyst contacted NSSE for the data.
   - Using the NSSE Year-to-Year Survey Variable Tracking Sheet, the analyst identified comparable variables between the two years. Anticipating other questions, however, the analyst created a data set not only with ‘SFC’ but with all potential variables of interest.
   - In order to seamlessly merge the data sets, the analyst found the SPSS script on the NSSE Web site to automatically merge the two data sets.

We want to hear from our users. Please send us descriptions of how NSSE multi-year results are being used on your campus. We also invite suggestions for improving the NSSE Multi-Year Data Analysis Guide. Please direct correspondence to the project staff at nsse@indiana.edu.
Additional Resources on Selected Topics

General Statistics
www.statsoft.com/textbook/stathome.html

Regression Analysis

Effect Size

Advanced Topics