A Multi-Year Analysis (of NSSE Data) Using ANOVA

John Stanley, Institutional Analyst
Yang Zhang, Director of Institutional Research
University of Hawaii at Mānoa
miro@hawaii.edu

CAIR Conference Napa, CA, Nov. 20-22, 2013
Presentation Outline

• What is a multi-year data analysis and why is it useful?

• UH Mānoa example using NSSE data
  • Brief background on NSSE
  • Research question
  • Methodology
  • Step-by-Step data analysis
  • Results

• Limitations and future studies
What is a Multi-Year Data Analysis?

In multi-year designs, we are interested in measuring whether there has been a meaningful change in outcomes over time.
Research Questions that can be Addressed by ‘Multi-Year’ Analysis

• Cross-Sectional (A): “Are 2005 and 2008 first year students’ experiences about advising significantly different?”

• Longitudinal (B): “Did 2005 first year students’ experiences change in 2008 when they became seniors?”

• (C): “Are 2005 first year students’ perceptions of a supportive campus environment significantly different from 2005 seniors?”

From NSSE Multi-Year Data Analysis Guide 2012
Why NSSE?

• NSSE is popular. As of 11/18/2013, 721 institutions will participate in 2014. Two million students have participated NSSE since 2000.

• NSSE Schedule: Administered in spring, reports is available in summer, and student responses are available in early fall.

• NSSE focuses on student engagement. It indirectly measures the extent to which first-year and senior students participate in educational practices that contribute to their academic and personal growth.

• Survey questions and benchmarks
  • 5 benchmarks: level of academic challenge; active and collaborative learning; student-faculty interaction; supportive campus environment, enriching educational experiences.

• Who will be interested in findings: leadership, faculty, committees (retention and graduation, strategic planning, initiatives, etc.)
UH Manoa Example

• Context: UH Mānoa’s Strategic Plan for 2002-2010 called for increasing support for academic advising.

• Approximately 40 new academic advising positions were created:
  • Faculty involvement in advising and mentoring increased considerably between 2002 and 2010.
  • Manoa Advising Center (MAC) formed in January 2007.

• Can NSSE data be used to indirectly measure the effectiveness of these strategic planning decisions?

i.e. Given the implementation of additional advising services in 2006-07, how much did student perceptions of the quality of academic advising increase from 2005 to 2009?
Research Question

Given the implementation of additional advising services at UHM in 2006-07, how much did student perceptions in the quality of academic advising increase from 2005 to 2009?

Today’s Example is designed like this

From NSSE Multi-Year Data Analysis Guide 2012
Useful documentation provided by NSSE to get you started

- NSSE Multi-Year Data Analysis Guide
- Contextualizing NSSE effect sizes
- SPSS syntax for merging multiple years of data
- Excel codeset for tracking year-to-year variable changes
- Working with NSSE Data: A Facilitator’s Guide

Link to NSSE Data Analysis Resources:
http://nsse.iub.edu/html/analysis_resources.cfm
Relevant Literature


Data Analysis

Four Steps:

1. Merge Multiple Years of NSSE Data

2. Perform Data Management Tasks
   a) Check Data Quality (sampling error, changes over time)
   b) Random Sampling
   c) Weighting

3. Choose Statistical Method

4. Data Analysis
   a) One-way ANOVA
   b) Post-hoc tests (Tukey Test)
   c) Cohen’s d
Step 1: Merge Multiple Years of NSSE Data

Sub CombineDataFiles (strPath As String, strFilemask As String)
    Dim strFileName As String
    Dim strCmd As String
    Dim intFileNb

    If InStr(strPath, "/") = 0 Then 'no path given, use current folder
        strPath = objSpssApp.GetSPSSPath
    End If
    If Len(strFilemask) = 0 Then 'no file mask given, combine all files
        strFilemask = ".sav"
    End If

    'Get the first file name and load file
    strCmd = "GET FILE=", & strPath & strFilemask & "." & vbCr
    objSpssApp.ExecuteCommands strCmd

    'Combine the other files
    intFileNb = 2
    While strFileName <> ""
        strFileName = Dir$(strPath)
        If strFileName <> "" Then
            strCmd = "ADD FILES /FILE=", & strPath & strFilemask & "." & vbCr
            strCmd = strCmd & "IF MISSING(source) source=", & intFileNb & "," & vbCr
            strCmd = strCmd & "ADD VALUE LABEL source ", & intFileNb & "," & strFilemask & "." & vbCr
            strCmd = strCmd & "EXECUTE."
            Debug.Print strCmd
            objSpssApp.ExecuteCommands strCmd, True
            intFileNb = intFileNb + 1
        End If
    Wend

    'Save the combined file
    strCmd = "SAVE OUTFILE=", & strPath & "combinedfile.sav".
    objSpssApp.ExecuteCommands strCmd, True
End Sub

NSSE provides a SPSS script that can merge multiple years of data in to one combined file called “combinedfile.sav”.

http://nsse.iub.edu/html/analysesSyntax_original.cfm
Step 2: Review Respondent Characteristics

Respondent Characteristics

The adjacent table displays your number of respondents, response rate, and sampling error by class. Sampling error is an estimate of the margin by which the true percentage of your students may differ from the reported percentage on a given item (because not all of your students completed surveys).

<table>
<thead>
<tr>
<th>Class</th>
<th>N</th>
<th>Resp. Rate</th>
<th>Sampling Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>First-Year Students</td>
<td>613</td>
<td>26%</td>
<td>+/-3.4%</td>
</tr>
<tr>
<td>Seniors</td>
<td>685</td>
<td>26%</td>
<td>+/-3.3%</td>
</tr>
</tbody>
</table>

From NSSE Executive Snapshot 2008 (University of Hawaii at Mānoa)
Step 2: Randomization and Weighting

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.

From NSSE Multi-Year Data Analysis Guide 2012
Step 3: Choosing a Method

Table 1
Considerations for Multi-Year Analytical Methods

<table>
<thead>
<tr>
<th></th>
<th>t-test</th>
<th>ANOVA</th>
<th>Regression</th>
<th>Effect Size</th>
<th>Percentage Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>How many years of data required?</td>
<td>2</td>
<td>3 or more</td>
<td>2 or more</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Determines statistical significant difference?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Determines magnitude of difference?</td>
<td>No*</td>
<td>No*</td>
<td>No*</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Are statistical controls possible?</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</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.

From NSSE Multi-Year Data Analysis Guide 2012
Step 4: One-way ANOVA in SPSS
One-way ANOVA: Options

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Width</th>
<th>Decimals</th>
<th>Label</th>
<th>Values</th>
<th>Missing</th>
<th>Columns</th>
<th>Align</th>
<th>Measure</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC</td>
<td>Numeric</td>
<td>8</td>
<td>2</td>
<td>Academic Challenge</td>
<td>None</td>
<td>None</td>
<td>6</td>
<td>Right</td>
<td>Scale</td>
<td>Input</td>
</tr>
<tr>
<td>ACA</td>
<td>Numeric</td>
<td>8</td>
<td>2</td>
<td>Academic Challenge</td>
<td>None</td>
<td>None</td>
<td>6</td>
<td>Right</td>
<td>Scale</td>
<td>Input</td>
</tr>
<tr>
<td>acadpt01</td>
<td>Numeric</td>
<td>8</td>
<td>2</td>
<td>Hours per 7-day</td>
<td>[1, 0]</td>
<td>None</td>
<td>8</td>
<td>Right</td>
<td>Scale</td>
<td>Input</td>
</tr>
<tr>
<td>ACL</td>
<td>Numeric</td>
<td>8</td>
<td>2</td>
<td>Active and Collaborative Learning</td>
<td>None</td>
<td>None</td>
<td>6</td>
<td>Right</td>
<td>Scale</td>
<td>Input</td>
</tr>
<tr>
<td>acct</td>
<td>Numeric</td>
<td>8</td>
<td>0</td>
<td>ACT total score</td>
<td>None</td>
<td>None</td>
<td>6</td>
<td>Right</td>
<td>Scale</td>
<td>Input</td>
</tr>
<tr>
<td>advise</td>
<td>Numeric</td>
<td>8</td>
<td>0</td>
<td>Overall, how would you evaluate the quality of academic advising you have received?</td>
<td>[1, Poor]</td>
<td>None</td>
<td>6</td>
<td>Right</td>
<td>Scale</td>
<td>Input</td>
</tr>
</tbody>
</table>

**Dependent List**

- Overall, how would you evaluate the quality of academic advising you have received?

**One-Way ANOVA: Options**

- **Statistics**
  - Descriptive
  - Fixed and random effects
  - Homogeneity of variance test
  - Brown-Forsythe
  - Welch

- **Means plot**

- **Missing Values**
  - Exclude cases analysis by analysis
  - Exclude cases listwise
One-way ANOVA: Post Hoc Test
### One-way ANOVA: Descriptive Statistics

#### Descriptives

Overall, how would you evaluate the quality of academic advising you have received at your institution?

<table>
<thead>
<tr>
<th>Group Description</th>
<th>N</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Std. Error</th>
<th>95% Confidence Interval for Mean</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>NISSE05 Data (UH Manoa) sav</td>
<td>93</td>
<td>2.62</td>
<td>0.76</td>
<td>0.090</td>
<td>2.48</td>
<td>2.70</td>
<td></td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>NISSE07 Data (UH Manoa) sav</td>
<td>98</td>
<td>2.60</td>
<td>0.90</td>
<td>0.090</td>
<td>2.32</td>
<td>2.68</td>
<td></td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>NISSE08 Data (UH Manoa) sav</td>
<td>106</td>
<td>2.81</td>
<td>0.78</td>
<td>0.074</td>
<td>2.68</td>
<td>2.95</td>
<td></td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>NISSE09 Data (UH Manoa) sav</td>
<td>105</td>
<td>2.75</td>
<td>0.81</td>
<td>0.080</td>
<td>2.63</td>
<td>2.87</td>
<td></td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>402</td>
<td>2.67</td>
<td>0.77</td>
<td>0.038</td>
<td>2.60</td>
<td>2.75</td>
<td></td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

#### Test of Homogeneity of Variances

Overall, how would you evaluate the quality of academic advising you have received at your institution?

<table>
<thead>
<tr>
<th>Levene Statistic</th>
<th>df1</th>
<th>df2</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.522</td>
<td>3</td>
<td>398</td>
<td>0.000</td>
</tr>
</tbody>
</table>

#### ANOVA

Overall, how would you evaluate the quality of academic advising you have received at your institution?

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>5677</td>
<td>3</td>
<td>1892</td>
<td>3.245</td>
<td>0.022</td>
</tr>
<tr>
<td>Within Groups</td>
<td>232143</td>
<td>398</td>
<td>583</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>237820</td>
<td>401</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
One-way ANOVA: Post Hoc Results

### Post Hoc Tests

**Dependent Variable:** Overall, how would you evaluate the quality of academic advising you have received at your institution?  
**Games-Howell**

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Group 2</th>
<th>Mean Difference</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>65% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSSE05 Data (UH Manoa) sav</td>
<td>NSSE07 Data (UH Manoa) sav</td>
<td>-119</td>
<td>.121</td>
<td>.758</td>
<td>-19</td>
</tr>
<tr>
<td>NSSE05 Data (UH Manoa) sav</td>
<td>NSSE06 Data (UH Manoa) sav</td>
<td>-188</td>
<td>.169</td>
<td>.314</td>
<td>-77</td>
</tr>
<tr>
<td>NSSE05 Data (UH Manoa) sav</td>
<td>NSSE09 Data (UH Manoa) sav</td>
<td>-129</td>
<td>.100</td>
<td>.574</td>
<td>-39</td>
</tr>
<tr>
<td>NSSE07 Data (UH Manoa) sav</td>
<td>NSSE05 Data (UH Manoa) sav</td>
<td>446</td>
<td>.134</td>
<td>.788</td>
<td>-43</td>
</tr>
<tr>
<td>NSSE05 Data (UH Manoa) sav</td>
<td>NSSE09 Data (UH Manoa) sav</td>
<td>-309*</td>
<td>.116</td>
<td>.045</td>
<td>-.61</td>
</tr>
<tr>
<td>NSSE05 Data (UH Manoa) sav</td>
<td>NSSE08 Data (UH Manoa) sav</td>
<td>249</td>
<td>.185</td>
<td>.192</td>
<td>-03</td>
</tr>
<tr>
<td>NSSE07 Data (UH Manoa) sav</td>
<td>NSSE05 Data (UH Manoa) sav</td>
<td>-188</td>
<td>.169</td>
<td>.314</td>
<td>-77</td>
</tr>
<tr>
<td>NSSE07 Data (UH Manoa) sav</td>
<td>NSSE09 Data (UH Manoa) sav</td>
<td>303*</td>
<td>.116</td>
<td>.045</td>
<td>.01</td>
</tr>
<tr>
<td>NSSE05 Data (UH Manoa) sav</td>
<td>NSSE09 Data (UH Manoa) sav</td>
<td>59</td>
<td>.095</td>
<td>.095</td>
<td>.31</td>
</tr>
</tbody>
</table>

*The mean difference is significant at the 0.05 level.
One-way ANOVA: Means Plot

First-Year Undergraduates

Mean of Overall, how would you evaluate the quality of academic advising you have received at your institution?

1 = “Poor”
2 = “Fair”
3 = “Good”
4 = “Excellent”

path=c:\temp\multiyear\
One-way ANOVA: Means Plot

First-Year Undergraduates

Mean of Overall, how would you evaluate the quality of academic advising you have received at your institution?

1 = “Poor”
2 = “Fair”
3 = “Good”
4 = “Excellent”

p = 0.023
Effect Size Calculators

Calculate Cohen's $d$ and the effect-size correlation, $r_{Y2}$, using --

- means and standard deviations
- independent groups $t$ test values and $df$

For a discussion of these effect size measures see Effect Size Lecture Notes

Calculate $d$ and $r$ using means and standard deviations

Calculate the value of Cohen's $d$ and the effect-size correlation, $r_{Y2}$, using the means and standard deviations of two groups (treatment and control).

Cohen's $d = M_1 - M_2 / s_{pooled}$

where $s_{pooled} = \sqrt{\left(\frac{s_1^2 + s_2^2}{2}\right)}$

$r_{Y2} = d / \sqrt{(d^{\sup2} + 4)}$

Note: $d$ and $r_{Y2}$ are positive if the mean difference is in the predicted direction.

Available at: http://www.uccs.edu/~lbecker/
One-way ANOVA: Means Plot

First-Year Undergraduates

Mean of Overall, how would you evaluate the quality of academic advising you have received at your institution?

p = 0.023; Cohen’s d = -0.375

1 = “Poor”
2 = “Fair”
3 = “Good”
4 = “Excellent”
## Proposed Reference Values for Effect Size Interpretation

(NSSE Benchmark Comparisons)

<table>
<thead>
<tr>
<th></th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>0.1</td>
</tr>
<tr>
<td>Medium</td>
<td>0.3</td>
</tr>
<tr>
<td>Large</td>
<td>0.5</td>
</tr>
<tr>
<td>Very Large</td>
<td>0.7</td>
</tr>
</tbody>
</table>
Summary

• Choose criteria to measure based on strategic planning priorities.
  • Identify and focus on specific questions. Are you looking for general shifts in engagement, or something more specific to assess a particular campus initiative?

• Follow steps to ensure data quality, randomization, and weighting.

• Post-hoc tests to see where the differences lies

• Check effect sizes
Limitations and Future Studies

Limitations/ Challenges:

-- Sample size difference in different years
-- Limited knowledge about college initiatives and strategies
-- Change of NSSE instruments over years, especially in 2013
  (Benchmarks to Engagement Indicators)

Future studies employing a multi-year analysis

-- analyzing other NSSE questions & benchmarks
-- using other external or home grown surveys results. Such as graduating seniors survey, first year experience survey, alumni survey, etc.

Questions and Suggestions?
John Stanley, Institutional Analyst
Yang Zhang, Director

University of Hawaii at Manoa
Manoa Institutional Research Office

Questions: miro@hawaii.edu

Link to this presentation: