SESSION OBJECTIVES

• Participants will be able to:
  • Apply proper controls and create a dataset for a research study
  • Evaluate multiple statistical analyses, such as statistical and practical significance, logistic regression, and segmentation modeling, for appropriateness to the study
  • Facilitate conversations around findings to effect change at their college
MOTIVATION

• Communications professor observed 100% success rates in short-term courses
• Considerable literature on accelerated and compressed coursework
• Title III STEM Grant was considering compressed sequencing
CONTROLS

- Reviewed literature for methods
- Determined 8-week control would provide most data and usable results
- Other controls: college, course, course length, faculty, course type, and term (fall and spring only)
- Disaggregation: age, gender, and ethnicity
CREATING DATASET

- A total of 4,592 records (over 5 years) were identified for the treatment group
- Applying the controls, 11,002 records were identified for the control group
- Variables created: Course success, cumulative prior GPA (continuous), cumulative prior GPA (normalized), student’s age in term, first-time students
• Statistical significance: p-value
• Practical significance: effect size (Cohen’s d)
• Predictive analytics: logistic regression & segmentation analytics
STATISTICAL TECHNIQUES

• Statistical significance: p-value
• Practical significance: effect size (Cohen’s d)
• Predictive analytics: logistic regression & segmentation analytics
STATISTICAL SIGNIFICANCE

- P-value is a well known statistic that faculty and decision-makers understand (or at least think they do)
- Remember, the p-value is impacted by the sample size!

\[
t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}
\]

- Use responsibly
STATISTICAL TECHNIQUES

- Statistical significance: p-value
- **Practical significance: effect size** (Cohen’s d)
- Predictive analytics: logistic regression & segmentation analytics
PRACTICAL SIGNIFICANCE

• Effect size (Cohen’s $d$): difference of the two means divided by the pooled standard deviation

\[ d = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}} \]

• Using Cohen as a guide:
  • Small effect $\geq 0.20$
  • Medium effect $\geq 0.50$
  • Large effect $\geq 0.80$
STATISTICAL TECHNIQUES

- Statistical significance: p-value
- Practical significance: effect size (Cohen’s d)
- Predictive analytics: logistic regression & segmentation modeling
PREDICTIVE ANALYTICS

• Logistic Regression
  • Predicts binary outcome (i.e. success, no success)
  • Standard package in statistical software
  • Supports (A LOT) of continuous and dichotomous predictor variables
  • Assumes lack of relationship between predictor variables
PREDICTIVE ANALYTICS

- Logistic Regression
  - Dummy coding
  - Multicollinearity
  - Missing values
  - Overfitting the model
  - Selecting the cut off value
  - Choosing the best model
  - Goodness-of-fit test
  - Interpretation of results
PREDICTIVE ANALYTICS

- **Segmentation Modeling**
  - Classification tree (CRT) algorithm
  - Standard in some packages, add-on for others
  - Partitions cases along predictor variables where similar outcomes are grouped together
  - Visual output that is easy to describe and understand
RESULTS

Success Rates by Course Length

- Traditional courses: 69%
- Compressed courses: 75%
RESULTS

Success Rates by Course Length for Students with Higher and Lower than Average Prior GPAs

- Lower GPA
  - Traditional courses: 57%
  - Condensed courses: 77%

- Higher GPA
  - Traditional courses: 70%
  - Condensed courses: 85%
RESULTS

Success Rates by Course Length and Subject

- **English**
  - Traditional courses: 71%
  - Compressed courses: 88%

- **Reading**
  - Traditional courses: 85%
  - Compressed courses: 92%
RESULTS

- Course length has an odds ratio of 1.553 and a positive β coefficient meaning a student enrolled in a compressed course is one and a half times more likely to succeed than a student enrolled in a traditional-length course.
- Prior cumulative GPA has an odds ratio of 2.083 and a positive β coefficient meaning a student is two times more likely to succeed than a student with a GPA .73 points lower.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>β</th>
<th>p</th>
<th>Exp(β)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course Length</td>
<td>.44</td>
<td>&lt;.001</td>
<td>1.553</td>
</tr>
<tr>
<td>Prior GPA</td>
<td>.73</td>
<td>&lt;.001</td>
<td>2.083</td>
</tr>
</tbody>
</table>
RESULTS

Segmentation model best predicted success for:
1) continuing students with GPAs greater than 3.028 and
2) students with GPAs between 2.457 and 3.028 enrolled in condensed courses
CLOSING THE RESEARCH LOOP

- Publish and distribute a full report and dashboard
CLOSING THE RESEARCH LOOP

- **Presentation** to faculty chairs committee
- Left the statistical language out
- Focused on:
  - Correlated relationships
  - Variables that best predicted course success
  - **Implications on teaching and learning**
CLOSING THE RESEARCH LOOP

Fall 2014
- Math 090
- Math 942

Spring 2015
- Math 095
- Math 952

Fall 2015
- Transferable Math
- Math 090

Spring 2016
- Math 095
CLOSING THE RESEARCH LOOP

- Provided impetus to grant program to identify interested faculty for a pilot program
- Completed first term of Fast Track Math

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Non-Fast Track MATH-942</th>
<th></th>
<th>Fast Track MATH-942</th>
<th>Effect Size (d)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#</td>
<td>N</td>
<td>%</td>
<td>#</td>
<td>N</td>
</tr>
<tr>
<td>Success in MATH-942</td>
<td>22</td>
<td>50</td>
<td>44.0</td>
<td>41</td>
<td>61</td>
</tr>
<tr>
<td>Retain to MATH-952</td>
<td>7</td>
<td>22</td>
<td>31.2</td>
<td>41</td>
<td>41</td>
</tr>
<tr>
<td>Success in MATH-952</td>
<td>5</td>
<td>7</td>
<td>71.4</td>
<td>24</td>
<td>41</td>
</tr>
<tr>
<td>Success in Sequence</td>
<td>5</td>
<td>50</td>
<td>10.0</td>
<td>24</td>
<td>61</td>
</tr>
</tbody>
</table>
FINAL QUESTIONS AND THOUGHTS

I DON’T ALWAYS ANSWER QUESTIONS

BUT WHEN I DO, THEY’RE YOURS


