Session objectives

1. Define “analytics” and discuss challenges.

2. Examine several case studies that have used analytics to improve institutional effectiveness.

3. Examine the use of analytics in the following key areas:
   a) Program review and assessment
   b) Enrollment management
   c) Surveys
   d) ‘What-if’ decision-making
Two institutions, one mission

Activity

Think – Pair – Share Activity

3 minute questionnaire + 5 minute share

1. Area(s) of notable strength?
2. Area(s) for improvement?
3. If IR analytical capacity were enhanced, what campus/program issue could be better addressed or potentially solved?
Challenges for Institutional Research

- Compliance vs. Self-Improvement
- Developing a culture of evidence
- From reporting to analysis
- Converting data into ‘actionable’ information
- Follow highest standards, best practices
- Know your customers, mission
- Leverage technology, stay abreast of tech
- Empower staff, continuous honing of skills
- Effective senior-management support working with IR (and IT)

AIR Newsletter March 2016

“I've seen too many IR offices that operate like a reporting agency and focus IR analysis only on what has happened in the past. Decisions, however, are made about the future – specifically, about the expected outcomes of future events... For the future of IR, professionals should become active in helping to minimize the risks of a decision by providing insightful analysis about possible outcomes.”

– Bob Daly, eAIR Newsletter, March 2016

SOURCE: https://www.airweb.org/eAIR/specialfeatures/Pages/The-Future-Of-IR.aspx
What is analytics?

“Analytics is the use of data, statistical analysis, and explanatory and predictive models to gain insights and act on complex issues.”

- EDUCAUSE Center for Applied Research

EDUCAUSE Center for Applied Research Video: “What is Analytics?”
Downloaded from: http://www.educause.edu/ero/article/video-what-analytics

From hindsight to foresight

Gartner Analytic Ascendancy Model

- Most of IR is stuck here
- Why did it happen?
- What happened?
- What will happen?
- How can we make it happen?

Descriptive Analytics
Diagnostic Analytics
Predictive Analytics
Prescriptive Analytics

Optimization
Foresight
Information
Hindsight
Insight
Value
Difficulty
IR tasks applied to Gartner Ascendancy Model

### DESCRIPTIVE
1. Disaggregating student retention rates by gender, ethnicity, Pell, first generation status, etc.
2. Descriptive data in tables, charts, and graphs.
3. A cross-tabulation showing retention rates for students in learning communities versus non-learning community students.

### DIAGNOSTIC
1. Using inferential statistics to determine if there are statistically significant differences between groups and identify important drivers of behavior.
2. Interactive dashboards with slice and dice capability, drop downs.
3. A counter-factual analysis that controls for self-selection bias using student matching techniques.

### PREDICTIVE
1. Building a prediction model to identify which students are ‘at-risk’ of dropping out.
2. Interactive dashboards with ‘what-if’ capability for key decision-makers.
3. Using learning community as a variable in a prediction model for retention or time-to-degree.

### PRESCRIPTIVE
1. Delivering dropout risk assessment lists to student support services in order to provide actionable information.
2. Interactive dashboards used to push dropout risk data to academic advisors.
3. A more precise gauge of the impact of learning communities given to LC coordinators.

---

Analytics typology

Covered Today

- Descriptive
- Optimization
- Explanatory
- Action
- Diagnostic
- **Predictive**
- Business Intelligence
- A.I.
- Academic
- Prescriptive
- Visualization
- Big Data
- Exploratory
- Action
- Business
- Learning

Source: https://library.educause.edu/~/media/files/library/2012/1/el3026.pdf.pdf
Visualization

• Data visualization is used to communicate data or information by representing it as visual objects (e.g., heatmaps, chords, sankeys).

• The goal is to communicate information clearly and efficiently to help users make:
  – faster insights
  – clearer choices
  – faster decision making

Visualization with Microsoft Power BI (free)

Source: https://app.powerbi.com/visuals/

Gartner Magic Quadrant for BI (2017)

Visualization with Microsoft Power BI

Dashboards as a Collection of Tiles Providing Links to Underlying Reports

Public informative dashboards. In this example, Loma Linda University’s enrollment demographics are displayed on the University’s website.

Source: http://home.llu.edu/academics/academic-resources/educational-effectiveness/institutional-research
Visualization with Microsoft Power BI

Interactive charts and graphs with explanatory text. This example provides student success information as required by the regional accreditor.

Visualization with Microsoft Power BI

Interactive maps that allow disaggregating data by selecting areas of interest. In this example program enrollment is displayed by state selected.
Predictive Analytics

- Uses historical data to predict or forecast future behaviors, trends, or outcomes
  (i.e. enrollment likelihood, retention, course pass/fail, degree completion, gainful employment, etc.)
Benefits of predictive analytics

• Can generate “actionable” data (i.e., data used by academic support services to effectively assist students).

• Powerful and accurate predictive models can be constructed using matriculation data from your Student Information System (SIS).

Possible uses of predictive analytics

• Admissions recruitment
  – Predict which students are likely to enroll at your institution (Goenner & Pauls, 2006)

• Identifying at-risk students
  – Predict which students are likely to drop out or fall behind (Herzog, 2006; Sujitparapitaya, 2006)

• Students’ price responsiveness to tuition increases or financial aid incentives (Des Jardins, 2001; Herzog & Stanley, 2017)

• Other uses?
  – Student Learning
  – Strategic Planning
  – Finance
A few examples of colleges using predictive analytics...

University of Texas at Austin

Data Informs Every Decision

With more than 35,000 applicants to The University of Texas at Austin last year, it is critical we make data-informed decisions to ensure the enrollment of a high-quality class that has the resolve to graduate in four years.

David Laude, PhD
“Graduation Champion”
For the Class of 2017:

- 94.6 percent retention, up from 93.6 percent prior year, resulting in the highest one-year retention rate in the university’s history for returning freshmen.
- Average GPA of 3.28, up from 3.22 for the previous class.
- Students enrolled in and passed more SSH (average 13.32 hours passed) than any entering class in the past five years. Taking more credit hours each semester will help these students stay on track to graduate in four years.

Other Noteworthy Examples

1. Georgia State University (reduced achievement gaps, featured nationally)
2. University of Nevada, Reno (early pioneer, raised retention 4% pts, featured nationally)

...See syllabus for more references
Learning Analytics

What is Learning Analytics?

• Defined as “the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs”

• Data may be available in your institution’s Learning Management System (LMS).
Learning Analytics via Sakai LMS

**Performance/Activity in Class (LMS)**
- Number and frequency of LMS logins
- Amount of time spent on course website
- Number of discussion posts
- Responses to class polls
- Grades and formative quiz scores
- Percentage of points earned in course to date
- Change between past and current test/quiz scores

**Student In-Class Assignments (LMS)**
- Blogs, discussion forum posts,
- Essays, written assignments

**Student Learning Outcomes (LMS)**
- Measurements of student achievements in core competencies in class.

**Matriculation Predictors (SIS)**
- Demographics (age, gender & ethnicity), GPA, pre-collegiate HS GPA, standardized tests scores, first-generation, socio-economic & financial need
LA can help answer questions like...

- What is the likelihood that a particular student will pass the class?... or master a certain learning outcome?

- Are there dispositional characteristics that predict or explain performance in certain classes (i.e. Do males outperform females in STEM classes or vice versa)

- Can LMS data be used with SIS data to predict student persistence and degree completion?

LA at Purdue University

http://www.nbcnews.com/id/3032619/vp/32634348#32634348
Student Retention at Purdue

>3% point rise since 2009

<table>
<thead>
<tr>
<th>Year</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>89%</td>
</tr>
<tr>
<td>2010</td>
<td>90%</td>
</tr>
<tr>
<td>2011</td>
<td>91%</td>
</tr>
<tr>
<td>2012</td>
<td>91%</td>
</tr>
<tr>
<td>2013</td>
<td>93%</td>
</tr>
</tbody>
</table>

Other Noteworthy Examples

1. Rio Salado College - “RioPACE”
2. University of Michigan - “Ecoach”
3. University of Maryland, Baltimore – “Blackboard Learn”

...See syllabus for more references
Challenges to using analytics

Affordability
- Infrastructure
- Technology
- People/Expertise
- Opportunity Costs

Data availability
- Student Information System
- Learning Management System
- Budget/ Human Resource Silos

Predictive data
- Culture change
- Wary of misuse of data
- Questions about data used to generate scores
- Students’ access to risk scores
- Self-fulfilling prophecy

Combined Activity + Break

Think – Pair – Share Activity

5 minute discussion over break

1. What are some challenges at your institution to supporting a culture for analytics (i.e., affordability, data availability, expertise, etc)?

2. What strategies may be helpful for overcoming these challenges?
Program Review and Assessment

**Process and Purpose**

- Process of Program Review on campus
- Variable skillset of faculty and staff
- Curriculum mapping and rubric assessment
- Collecting assessment data
- Purpose
  - Bringing it all together
  - Using the data to make changes where needed - “closing the loop”
- Power BI examples
Rubric Element Scoring – Institution Example

• Challenge –
  – Assessment software couldn’t provide average rubric value
  – Assessment software didn’t allow complete data extract
  – Attempting to use the assessment software built-in analytics resulted in:

![Alert](image)

Oops. There is too much number crunching required to run your report. If you are using an “Excludes” or “Doesn’t Contain” filter, try removing or flipping them to “Include” or “Contains” filter. You can also try turning off subtotals and grand totals.

OK

Rubric Element Scoring – Institution Example

• Challenge –
  – Extract needed student IDs
  – Rudimentary analytics provided by assessment software

![Bar Charts](image)
Rubric Element Scoring – Institution Example

• Solution –
  – Create small reports in the assessment software
  – Download csv of created report
  – Recombine using Power BI query editor
Analytics Application #2

Enrollment Management

Predictive analytics for the enrollment funnel

- **Prospect Scoring**
  - Outcome: Likelihood of Accepting Admissions Offer
- **Retention Scoring**
  - Outcome: Likelihood of fall-to-spring retention, fall-to-fall retention
- **Graduation Scoring**
  - Outcome: Likelihood of graduation, time-to-degree
- **Persistence Scoring**
  - Outcome: Re-enrollment
How does predictive modeling work?

Student Dropout Risk Example

1. Student data collected from SIS
2. Build prediction model using standard regression techniques
3. Prediction estimates are calculated per individual student

Student John Doe at the beginning of the semester:
- First Generation Student
- < Average high school GPA (3.00)
- Attempting 12 credits (12)
- Low % of financial need met (65%)
- Undeclared major
- Not enrolled in a campus learning community
- No educational goals in survey
- Not working on campus

Probability of Dropping Out: 75%

A taxonomy of SIS data available for prediction

- **DEMographics**: Gender, Age, Ethnicity, Residency, Geographic Origin
- **Campus Experience**: On Campus Employment, Housing, Student Life Activities, Athletics, STAR Usage, Average Class Size
- **Academic**: High School GPA & Rank, SAT, AP CLEP, Educational Goals, Transfer GPA, # Transfer Credits
- **Pre-College**: Major, Credit Load, Credits Earned, First Term GPA, Distance Education, Dual Enrollment, High Failure Rate Courses, Courses Taken (including English)
- **Financial Need**: Need Based Aid, Non-need Based Aid, Pell Grant, Work Study, % of Aid Met
- **Milestones**: Credits earned, Credits attempted, Credit Completion Ratio, Math/English Enrollment/Completion, Continuous Enrollment, Milestone metrics
- **Interactions**: Ethnicity by Geographic Origin, Employment by Housing, High School GPA by First Term GPA, Residency by Need Based Aid, Ratio of SuccessfulAdds to Drops
Sample data for advisors/success coaches

<table>
<thead>
<tr>
<th>ID</th>
<th>LAST NAME</th>
<th>FIRST NAME</th>
<th>EMAIL</th>
<th>CURRENT CREDITS</th>
<th>RESIDENT</th>
<th>AP/CLEP</th>
<th>HS GPA</th>
<th>WORK ON CAMP</th>
<th>1st YR EXP</th>
<th>CLASS</th>
<th>% FIN NEED MET</th>
<th>STAR LOGINS</th>
<th>ADVISOR PREVIOUS CONTACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td></td>
<td></td>
<td></td>
<td>15</td>
<td>HI</td>
<td>6</td>
<td>3.80</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>77%</td>
<td>5</td>
<td>Y</td>
</tr>
<tr>
<td>002</td>
<td></td>
<td></td>
<td></td>
<td>14</td>
<td>HI</td>
<td>0</td>
<td>3.33</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>63%</td>
<td>3</td>
<td>N</td>
</tr>
<tr>
<td>003</td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>CA</td>
<td>6</td>
<td>3.00</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>45%</td>
<td>0</td>
<td>N</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ID</th>
<th>AGE</th>
<th>GENDER</th>
<th>ETHNICITY</th>
<th>COLLEGE</th>
<th>MAJOR</th>
<th>DEGREE</th>
<th>Ed Goal Specified</th>
<th>Relative Risk Value</th>
<th>Risk Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>18</td>
<td>F</td>
<td>CH</td>
<td>CA&amp;H</td>
<td>ART</td>
<td>BA</td>
<td>Yes</td>
<td>14.92</td>
<td>LOW</td>
</tr>
<tr>
<td>002</td>
<td>18</td>
<td>F</td>
<td>HW</td>
<td>CSS</td>
<td>SOC</td>
<td>BA</td>
<td>Yes</td>
<td>36.88</td>
<td>MEDIUM</td>
</tr>
<tr>
<td>003</td>
<td>18</td>
<td>M</td>
<td>UNDEC</td>
<td>UNDEC</td>
<td>UNDEC</td>
<td>UNDEC</td>
<td>No</td>
<td>89.18</td>
<td>HIGH</td>
</tr>
</tbody>
</table>

Academic advising intervention example

**John**
- First Generation
- SAT-M/SAT-R = 900
- H.S. GPA = 3.00
- 85% Unmet Financial Need
- Undeclared
- 12 credits registration
  - Dropout risk probability: **60%**
  - Risk group: **6 of 10**

**Intervention strategy:**
- Proactive advising
- Meta major pathway mapping
- Revisit financial aid support and other support
- Check for possible ill-advised registration choices

**Ken**
- SAT-M/SAT-R = 1100
- H.S. GPA = 3.50
- Declared major (Accounting)
- Local address within 5 miles
- 15 credits registration
- Educational Goals = “Earn B.A.”
  - Dropout risk probability: **15%**
  - Risk group: **1 of 10**

**Intervention strategy:**
- Monitor *Starfish* reporting
- Mid-semester check-in
- Re-assess dropout risk at end-of-semester
Analytics system for enrollment mgmt.

Dashboard tiles organized to ascend analytics from descriptive to prescriptive.

Predictive analytics

Top Predictors
- Resident
- High Yield H.S.
- Expressed Interest
- Gender
- Low Yield U.S. Region
- CTAIR SOEST App.

Strongest

Weakest

These variables account for approximately 30% of the variance in a student's likelihood of completing college (Pseudo R Square = 0.30).

Fall 2018 Holdout Results
University of Hawaii

Model Summary
- J-Log Likelihood: -250.4
- Cox & Snell R Square: 0.26
- Nagelkerke R Square: 0.26

Classification Table
- Did Not Enroll: 451, 784
- Enrolled: 908, 921
- Overall: 72.4%

Parameter Estimates for Fall 2018 Enrollment Yield
- Constant: -3.78, p = 0.12
- Resident: 2.75, p = 0.11
- Historically High Yield H.S.: -0.12, p = 0.03
- Expressed Interest: 0.01, p = 0.00
- Gender: 0.03, p = 0.00
- Low Yield U.S. Region: -0.03, p = 0.00
- CTAIR SOEST applicant: -0.03, p = 0.00
- Non Cal: -0.01, p = 0.00
- High yield H.S. region: 0.01, p = 0.00
- Application date 270 days or older: -0.01, p = 0.00
- Number of applications: 0.01, p = 0.00
- Miles from campus over 20: 0.01, p = 0.00
A prediction system for retention

Dashboard tiles organized to ascend analytics from descriptive to prescriptive.

Predictive analytics
Prescriptive analytics

Technical training available for predictive analytics

AIR Workshop: A Step-by-Step Introduction to Building a Student-at-Risk Prediction Model
Instructors: Serge Herzog (University of Nevada-Reno) and John Stanley (University of Hawai‘i – West Oahu)
Prescriptive analytics activity

Post hoc Activity

Read *Retention data use* . . . , see spreadsheet, and then
A. Generate list of reasons from multiple perspectives.
B. Draft techniques/strategies to address reasons.
C. Share with a neighboring table.

Analytics Application #3

Surveys
Power BI Solution for Qualtrics Survey Data

- Challenge – each row contained individual data elements
- Columns by individual numbered question or numbered answer
- Report needs to be updated as students continue to take survey
- Report needs to be disaggregated by school and program
- Comments need to be organized

Power BI Solution for Qualtrics Survey Data

- Dataset is duplicated for each question
- Extra columns (other answers) deleted
- New columns for filters are added
- Columns renamed for clarity
- Each step is recorded, can be “played back” and edited at any time
Power BI Solution for Qualtrics Survey Data

• Each copied table is then merged
• Power BI has options that will automate this each time source table is updated
• Resultant merged table now ready to be used for developing visualization

Analytics Application #4

What-if Analysis
Tuition revenue what-if planner

What-if activity

Final Activity

Read Cost Estimator Worksheet. . . , see dashboard, and then

A. Generate a research question to answer using the parameters in the dashboard.

B. Discuss which parameters need to be checked and unchecked in order to visualize the answers.

C. Share with a neighboring table.
Institutional benchmarking using what-if

New WASC comparative dashboard under development. Compare actual versus predicted rates and perform what-if analyses.

ARC Session: Thursday, April 20, 2:15-3:15

Thank You

John Stanley
Director, Institutional Research
University of Hawai‘i – West Oahu
jstanley@hawaii.edu

Kenneth Nelson, MD, MS
Associate Director, Educational Effectiveness
Loma Linda University
wknelson@llu.edu