Predictive Analytics: Leveraging Powerful Data for Student Success

Jason E. Lane, PhD
Director, The Systems Center
Dean, School of Education
State University of New York at Albany
Predictive Analytics

A prominent barrier to student success, particularly among underrepresented, low-income, and first-generation students, is a lack of information. Predictive analytics and data-mining techniques have proven to be powerful methods of empowering and informing students. The focus of this initiative is on implementing predictive analytics across multiple systems, strengthening the data infrastructure needed to leverage these tools, and implementing the policy, curricular advances, and academic support programs needed to enable the successful use of predictive analytics.

http://ts3.nashonline.org/predictive-analytics/
Given that, the ongoing analytics revolution could not have come at a better time

Did you know . . .

• Since 2013, humans produce in two days as much data as the entire human race produced prior to 2003
• This means, over 90% of all data every created was created in the last two years.
• By 2015 the digital universe included 50 billion devices and over 180 billion zettabytes of data.
• It would take over 450 million years to transmit that data through a broadband internet connection
• The total amount of data being captured and stored by industry doubles every 1.2 years.
• Every minute, we send 204 million emails, generate 1.8 millions likes on Facebook, send 278,000 tweets, and up-load 200,000 photos to Facebook.
With data being produced at an ever faster rate, determining how to best use it can be a difficult undertaking.

| V | elocity | The creation of data and the appetite for insight continues to grow faster |
| V | olume | The amount of structured and unstructured data has grown to unimaginable levels |
| V | alue | Data must be used to create accurate insight that lead to organizational learning and continuous improvement |
| V | ariety | The nature of data is continuously changing, which has implications for the manner in which it is processed, stored, and analyzed |
| V | eracity | Data need to be accurate and trustworthy without both, as the old saying goes, “garbage in, garbage out” |
As colleges and universities become better at extracting the value of big data the manner in which it gets used will evolve.

- **Descriptive**: Reporting what happened
- **Diagnostic**: Analyzing why it happened
- **Real-Time**: Monitoring what is happening
- **Predictive**: Forecasting what will happen
- **Prescriptive**: Determining what should be done

Most colleges and universities are here. Many are trying to get here.
An Introduction to Data Analytics for Student Success
The primary goals of this course are to help you **understand the role of data** in supporting student success and provide you with knowledge about **how this information can be used** to enhance the ways in which you interact with students.
Introduction

Exploring Data: Definitions, Characteristics, and Possibilities

Supporting Student Success with Data Analytics

Finding Your Place in the Data Analytics Landscape

Ethics, Security, and Governance Issues

Integrating Data Analytics into Your Practice
“Data has the potential to enhance educational opportunities for all students, but data alone does not guarantee equity.”
The nature of the student body is ever-changing

• 58% of all undergraduates were post-traditional learners

• These students were more likely than other students to be female, non-white, and to live off campus

• They were also more likely to be seeking shorter term-credentials, and to already have a credential
Data presents a set of opportunities around equity

**Reflections on Data & Equity**

- Data can raise awareness of disparity or be used to justify discrimination.
- Data can be used for inclusion or exclusion.
- Data can be used to inform transformational educational intervention or to game the system.
- Interpretation of data can lead educators to provide extra support for students in need or support biases.
As students become more mobile we are uncovering several deficiencies in how transfer of credit is facilitated. A recent GAO analysis of credit transfer found that in the 2004 cohort, 35 percent of first-time students transferred credits between schools at least once during a six-year period. For these students who transfer, the GAO estimated a 43 percent loss of college credits, or an average of 13 credits lost per student. *This equates to essentially an entire semester of full time enrollment lost.*
Data can help institutions thread the needle

As colleges and universities face the need to respond to changing demographics and equity gaps, data can help guide the way.
Those who focus solely on the “predictive” without a focus on the “prescriptive” are sorely missing the mark.

Many are trying to get here.

- **Predictive**: Forecasting what will happen
- **Prescriptive**: Determining what should be done

Wizard

Doctor
A misunderstanding of the possible pros and cons of using predictive analytics can be very detrimental to student success and closing attainment gaps.

Predictive + Prescriptive = Witch doctor?
Whether or not data is big misses the important point . . . using it to help students

“Big data is bullshit . . . We’ve all been doing data for years. If anyone says big in front of it you should look at them very skeptically. It’s about finding big answers to big questions.”
Some issues to consider

- Implicit Bias
- Personal Agency
- Preventing Self-Fulfilling Prophecies
- Balancing our Attentions
- Developing Limited Pathways (probability v. actuality)
- Oversimplifying underlying issues
Might a self-fulfilling prophecy be created when staff, faculty, or students are made aware of predicted risk? Or when students are targeted for specific interventions based on past performance or demographics associated with attrition?

Laura Jenson (The Ethical Use of Student Data and Analytics)
Now, more than ever, our ethics are coming to light in the actions and/or interventions we take resulting from what we learn from data analytics.

Many of us are used to **making decisions** on how to **intervene** (or not to) based on our **instincts** about what a **student needs**, rather than looking at **student-generated data** to **guide** us.

**Avoiding Bias and Student Profiling**

Ethical practice dictates that we do not use the data we have access to, or our analysis and interpretation of that data, to define or stereotype our students.
What You Can Do

Addressing implicit bias can be difficult because of its nature. It is important to think about each and every student as an individual, rather than as part of a group.

When working with data detailing a student from a particular “group” -- particularly if they have been identified as potentially “at risk” -- remember that this student is an individual. Avoid responding from your impressions. Instead, seek out and discuss specific problems, and explore positive actions to remediate those problems.

Be self-aware of the potential for implicit bias to influence your thinking and work to keep it in check. Consider initiating dialogue among your colleagues and peers related to how implicit and/or unconscious bias impacts their interactions with students. Developing an inclusive approach to teaching and learning in your practice will mitigate potential biases, and support the appropriate use of data analytics in your work.
As a leader, you are responsible for guiding the establishment of clear ethical expectations related to data use, and to inform faculty and staff about those expectations.

Explore the following questions and scenarios related to ethics, security, and governance of data at your institution.

1. How are faculty and staff made aware of any relevant laws pertaining to the use and sharing of data?
2. How do you ensure that the data being used informs student decisions rather than dictates them?
3. Are students aware of how their data is being used and for what purpose?
SUNY seeks to measure and validate their national leadership in curricular, co-curricular and extracurricular innovations that advance student success, forward the SUNY Completion Agenda, and ultimately contribute to the long-term economic sustainability of NY.
Prediction-based Propensity Score Matching: Ensures apples-to-apples comparisons in a computationally efficient manner (example below)

Derive student variables

- High school GPA
- Pell grant eligibility
- Credits earned/term
- Cum GPA
- Grade consistency
- Enrollment behavior

Build two models

- Student success predictive model
  
  \[ y = f(x) \]
  
  \[ z = g(x) \]

Match in \( y \) & \( z \)

- Student propensity score model

Match students such that they are equally likely to succeed and participate
<table>
<thead>
<tr>
<th>Program</th>
<th>Overall Effect Size</th>
<th>p-value</th>
<th># of Matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Civic Engagement</td>
<td>NSS</td>
<td>0.401</td>
<td>4056</td>
</tr>
<tr>
<td>Community Service</td>
<td>NSS</td>
<td>0.310</td>
<td>2588</td>
</tr>
<tr>
<td>Creative Works</td>
<td>-177</td>
<td>0.006</td>
<td>4125</td>
</tr>
<tr>
<td>EOP</td>
<td>NSS</td>
<td>0.248</td>
<td>397</td>
</tr>
<tr>
<td>Honors</td>
<td>NSS</td>
<td>0.428</td>
<td>323</td>
</tr>
<tr>
<td>Internship</td>
<td>388</td>
<td>0.000</td>
<td>2413</td>
</tr>
<tr>
<td>Online</td>
<td>NSS</td>
<td>0.099</td>
<td>1834</td>
</tr>
<tr>
<td>Research</td>
<td>-233</td>
<td>0.001</td>
<td>3867</td>
</tr>
<tr>
<td>Service Learning</td>
<td>NSS</td>
<td>0.386</td>
<td>3426</td>
</tr>
<tr>
<td>Writing Intensive</td>
<td>NSS</td>
<td>0.155</td>
<td>1230</td>
</tr>
<tr>
<td>Program</td>
<td># of positive segments</td>
<td>Examples of positive Segments</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Civic Engagement</td>
<td>4</td>
<td>Degree-BA ($161), Persistence Score - Top Quartile ($4124)</td>
<td></td>
</tr>
<tr>
<td>Community Service</td>
<td>3</td>
<td>Degree-BA ($175), Persistence Score - Top Quartile ($3786)</td>
<td></td>
</tr>
<tr>
<td>Creative Works</td>
<td>2</td>
<td>Persistence Score - Top Quartile ($4348)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Persistence Score - Third Quartile ($1565)</td>
<td></td>
</tr>
<tr>
<td>EOP</td>
<td>0</td>
<td>Persistence Score - Top/Third Quartile (+ve, p-value &lt; 0.05, N &lt; 200)</td>
<td></td>
</tr>
<tr>
<td>Honors</td>
<td>0</td>
<td>Persistence Score - Top/Third Quartile (+ve, p-value &lt; 0.05, N &lt; 200)</td>
<td></td>
</tr>
<tr>
<td>Internship</td>
<td>29</td>
<td>Non-STEM ($415), Persistence Score - Top Quartile ($4301)</td>
<td></td>
</tr>
<tr>
<td>Online</td>
<td>4</td>
<td>Tap Grant - N ($226), Persistence Score - Top Quartile</td>
<td></td>
</tr>
<tr>
<td>Program</td>
<td># of negative Segments</td>
<td>Example negative Segments</td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Civic Engagement</td>
<td>4</td>
<td>Degree-BS ($-496), Persistence Score - Bottom Quartile ($-3035)</td>
<td></td>
</tr>
<tr>
<td>Community Service</td>
<td>4</td>
<td>Degree-BS ($-856), Persistence Score - Bottom Quartile ($-3420)</td>
<td></td>
</tr>
<tr>
<td>Creative Works</td>
<td>26</td>
<td>Persistence Score - Bottom Quartile ($-3164)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Part-time ($-429), Age &lt;= 24 ($-126)</td>
<td></td>
</tr>
<tr>
<td>EOP</td>
<td>0</td>
<td>Persistence Score - Bottom/Second Quartile (-ve, p-value &lt; 0.05, N &lt; 200)</td>
<td></td>
</tr>
<tr>
<td>Honors</td>
<td>0</td>
<td>Persistence Score - Bottom/Second Quartile (-ve, p-value &lt; 0.05, N &lt; 200)</td>
<td></td>
</tr>
<tr>
<td>Internship</td>
<td>1</td>
<td>Persistence Score - Bottom Quartile ($-2836)</td>
<td></td>
</tr>
<tr>
<td>Online</td>
<td>2</td>
<td>Persistence Score - Bottom Quartile ($-3188)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Persistence Score - Second Quartile ($-255)</td>
<td></td>
</tr>
</tbody>
</table>
How Can We Close the Equity Gap in Prediction n-tile?

- Soft skills teaching (U. of Illinois at Urbana-Champaign example)
- Situational learning with realistic expectation setting to help deal with challenging work-related situations
- Workshop on lifelong learning skills? Many learn better with context from work projects.
- More tailored programs based on a student’s level of preparedness -- shallow staircase?
- Embedded purposeful reflection activities increase connection of experience to learning
- Techniques of highly effective health coaches in improving patient activation measure (PAM)
- Micro-credentialing
- Life coaching

Supporting Patient Behavior Change: Approaches Used by Primary Care Clinicians Whose Patients Have an Increase in Activation Levels

Jessica Greene, PhD
Judith H. Hibbard, DsPh
Carmen Alvarez, PhD, RN, NP-C, CNM
Valerie Overton, DNP
School of Nursing, George Washington University, Washington, DC
ISE Health Policy Group, University of Oregon, Eugene, Oregon
School of Nursing, Department of Comm-

**ABSTRACT**

**PURPOSE** We aimed to identify the strategies used to support patient behavior change by clinicians whose patients had an increase in patient activation.

**METHODS** This mixed methods study was conducted in collaboration with Fairview Health Services, a Pioneer Accountable Care Organization. We aggregated data on the change in patient activation measure (PAM) score for 7,144 patients to the primary care clinician level. We conducted in-depth interviews with 10 clinicians whose patients’ score increases were among the highest and 10 whose patients’ score changes were among the lowest. Transcripts of the interviews were analyzed to identify key strategies that differentiated the clinicians whose patients had top PAM change scores.
Questions?