Using Advanced Analytics to Boost Student Success

Dr. Braden J. Hosch, Asst. Vice President for Institutional Research, Planning & Effectiveness, Nov. 11, 2018
Overview

• Institutional profile and grad rate improvements
• Initiatives
• “Traditional” IR / Analytics
• Predictive Analytics
• Takeaways
Stony Brook University Institutional Profile

Students:
26,254
Fall headcount

Graduate
33%
Undergraduate
67%

Institution:
Doctoral, Highest Research Activity
Founded 1957, joined AAU 2001

Undergraduate Profile
1323 avg. SAT

Pell Recipients
32%

7,272 Completions 2017-18

Program Profile

STEM 38%
Health 21%
Other 41%

Employees:
14,732 including hospital
2,704 faculty (FT & PT)

Finance:
2.5 billion USD annual budget
230 million USD research exp.
Freshman graduation rates increased fifteen percentage points in the last five years; equity gaps are largely closed.
## Improvements realized through multi-pronged approach

<table>
<thead>
<tr>
<th>Broad-based academic success team</th>
<th>“Traditional” institutional research</th>
<th>In-house analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd-party analytics</td>
<td>Policy and procedure reform</td>
<td>Mini-grants to seniors</td>
</tr>
<tr>
<td>Attention to special populations</td>
<td>Expanded advising</td>
<td>Class availability</td>
</tr>
</tbody>
</table>
Traditional IR - grad rates by DFW rates

Number of 1st Term Course Grades of D, F, W or U

4-Year Graduation Rate
- Zero: 12%
- One: 41%
- Two+ (red): 68%

Share of Cohort
- 85%
- 10%
- 5%

Source: IRPE Grad Rate data set v31; cohort entering 2014
Address Courses with Higher DFW Rates

Top 18 Fall 2010 courses
23.5%-37.9%

Top 18 Fall 2017 courses
18.1%-25.9%
Exploratory IR – number of course grades of A

Number of 1\textsuperscript{st} Term Course Grades of A or A-

4-Year Graduation Rate (pct)

- Two+: 40%
- One: 55%
- Zero: 73%

Share of Cohort

- Two+: 21%
- One: 22%
- Zero: 57%

Source: IRPE Grad Rate data set v31; cohort entering 2014
Method for local analytics: student-level predictions

Credit to: Nora Galambos, Ph.D., Senior Data Scientist

<table>
<thead>
<tr>
<th>Decision trees using SAS Enterprise Miner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification and Regression Trees (CART) method</td>
</tr>
</tbody>
</table>
## Data included in model

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Pre-college academic characteristics</th>
<th>College academic characteristics</th>
<th>Transactions, service utilization, activities.</th>
<th>Financial aid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>SAT scores</td>
<td>Credits accepted when admitted</td>
<td>Learning management system (LMS) logins</td>
<td>Expected family contribution AGI</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td>high school GPA</td>
<td>AP credits</td>
<td>advising visits</td>
<td>types and amounts of disbursed aid</td>
</tr>
<tr>
<td>geographic residence</td>
<td>average SAT scores of the high school</td>
<td>number of STEM and non-STEM</td>
<td>tutoring center utilization</td>
<td>Pell, Tuition Assistance Program</td>
</tr>
<tr>
<td>admitted.</td>
<td>GPA</td>
<td>courses current term</td>
<td>intramural and fitness class participation</td>
<td>(TAP).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>enrollment in high DFW courses</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>area of major.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
LMS Data Processing

- Count only one login per course per hour
  - A course can have up to 24 logins per day
  - Eliminates multiple logins just few minutes apart.

- Logins totaled by week

- Per-course login rates calculated for STEM and non-STEM courses

- Class assignment grades not yet included
  - Timing and data processing issues
  - Completeness issues
  - Significant noise and false positives
**Decision Tree Model for Freshmen GPA: Part 1—HS GPA <= 92.0**

<table>
<thead>
<tr>
<th>HS GPA &lt;= 92.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>LMS logins per non-STEM crs, wk 2-6 &gt;=11.3 or missing</td>
</tr>
<tr>
<td>LMS logins per non-STEM crs, wks 2-6&lt;11.3</td>
</tr>
</tbody>
</table>

### Avg. HS SAT CR >570
- SAT Math CR >1360
  - AP STEM Crs >= 1
  - AP Stem Crs = 0
  - Highest DFW STEM Crs Rate >= 17%
  - Highest DFW STEM Crs Rate <17%
  - SAT Math CR <=1360
  - SAT Math CR >=1360
  - Logins per STEM crs, wk 2-6 >=32.2
  - Logs per STEM crs, wk 2-6 <32.2

### Avg. HS SAT CR <=570
- SAT Math <680 or miss.
- Logins per non-STEM crs, wk 2-6 >=11.3 or missing
- Logins per non-STEM crs, wk 2-6<11.3
- Non-STEM crs logs Wk 1 >= 3 or miss.
- Non-STEM crs logs <3

### Avg. HS SAT CR >=540
- AP STEM Crs. >=1
- AP STEM Crs = 0
- Logs per STEM crs, wk 2-6 >=5.3 or miss
- Logs per STEM crs, wk 2-6<5.3

### Avg. HS SAT CR < 540
- SAT Math <680 or miss.
- Logs per STEM crs, wk 2-6 >=1 or miss.
- Logs per STEM crs, wk 2-6 <1 or miss.

### SAT Math <680 or miss.
- SAT Math CR >1360
  - AP STEM Crs >= 1
  - AP Stem Crs = 0
  - Highest DFW STEM Crs Rate >= 17%
  - Highest DFW STEM Crs Rate <17%
  - SAT Math CR <=1360
  - SAT Math CR >=1360
  - Logins per STEM crs, wk 2-6 >=32.2
  - Logs per STEM crs, wk 2-6 <32.2

### Non-STEM crs logs Wk 1 >= 3 or miss.
- Non-STEM crs logs <3
- STEM crs logs Wk 1 >=5 or miss.
- STEM crs logs Wk 1 <5

### STEM crs logs Wk 1 = 0
- Stem crs logs Wk 1 <1 or miss.
- Stem crs logs Wk 1 = 1 or miss.

### Logs per STEM crs, wk 2-6 >=5.3 or miss
- Logs per STEM crs, wk 2-6<5.3

### Logs per STEM crs, wk 2-6<1 or miss
- Logs per STEM crs, wk 2-6 >1 or miss.

### Avg. GPA = 3.63
N = 46

### Avg. GPA = 3.20
N = 23

### Avg. GPA = 2.92
N = 34

### Avg. GPA = 3.25
N = 94

### Avg. GPA = 3.09
N = 78

### Avg. GPA = 2.94
N = 121

### Avg. GPA = 2.75
N = 57

### Avg. GPA = 2.69
N = 64

### Avg. GPA = 2.75
N = 16

### Avg. GPA = 2.12
N = 18

### Avg. GPA = 2.62
N = 305

### Avg. GPA = 1.94
N = 25

### Avg. GPA = 1.59
N = 13
## Decision Tree Model for F14 Freshmen GPA: Part 2—HS GPA > 92.0

### HS GPA > 92.0 or Missing

<table>
<thead>
<tr>
<th>Scholarship = Yes</th>
<th>Scholarship = No</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HS GPA &gt;= 96.5 or missing</strong></td>
<td><strong>LMS logins per non-STEM crs. Wk 2-6 &gt;=10.4</strong></td>
</tr>
<tr>
<td><strong>Math Placement Exam &gt;= 5</strong></td>
<td><strong>AP STEM Crs. &gt;=1</strong></td>
</tr>
<tr>
<td><strong>Math Placement Exam &lt; 5</strong></td>
<td><strong>AP STEM Crs = 0</strong></td>
</tr>
<tr>
<td>Logs per STEM Crs, wks 2-6 &gt;=29.1</td>
<td>DFW STEM Crs Total &gt;=2</td>
</tr>
<tr>
<td>SAT Math &gt;=700 or miss.</td>
<td>SAT Math &gt;=760</td>
</tr>
<tr>
<td>Avg. HS CR, M, Wrt &gt;=1830</td>
<td>SAT Math &lt;760</td>
</tr>
<tr>
<td>Ethnic Group = Asian, Afr. Amer., Unk.</td>
<td>DFW non-STEM 1st yrs &lt;28%</td>
</tr>
<tr>
<td>Ethnic Group = White, Hispanic</td>
<td>DFW non-STEM 1st yrs &gt;=28%</td>
</tr>
<tr>
<td>SAT Math &lt;700 or miss.</td>
<td>STEM Crs logs Wk 1 &lt;=8</td>
</tr>
<tr>
<td>Logs per non-STEM crs, wks 2-6 &lt;29.1</td>
<td>Logs per STEM crs, wks 2 6 &lt;10.9</td>
</tr>
<tr>
<td>Avg. GPA = 3.63 N = 285</td>
<td>Avg. GPA = 2.52 N = 16</td>
</tr>
<tr>
<td>Avg. GPA = 3.40 N = 83</td>
<td>Avg. GPA = 1.30 N=11</td>
</tr>
<tr>
<td>Avg. GPA = 3.05 N=30</td>
<td>Avg. GPA = 2.90 N = 73</td>
</tr>
<tr>
<td>Avg. GPA = 3.76 N=26</td>
<td>Avg. GPA = 3.05 N=194</td>
</tr>
<tr>
<td>Avg. GPA = 3.52 N = 74</td>
<td>Avg. GPA = 3.76 N = 72</td>
</tr>
<tr>
<td>Avg. GPA = 3.59 N = 54</td>
<td>Avg. GPA = 3.03 N = 11</td>
</tr>
<tr>
<td>Avg. GPA = 3.13 N = 54</td>
<td>Avg. GPA = 3.49 N = 101</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Math Placement Exam &lt; 15.6</strong></th>
<th><strong>LMS logins per non-STEM crs. wk 2-6 &lt; 10.4</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>SAT Math &lt;700 or miss.</td>
<td>DFW STEM Crs Total &lt;2</td>
</tr>
<tr>
<td>Logs per STEM Crs, wks 2-6 &lt;29.1</td>
<td>SAT Math &lt;760</td>
</tr>
<tr>
<td>Avg. HS CR, M, Wrt&lt;1830</td>
<td>DFW non-STEM 1st yrs &lt;28%</td>
</tr>
<tr>
<td>Ethnic Group = White, Hispanic</td>
<td>DFW non-STEM 1st yrs &gt;=28%</td>
</tr>
<tr>
<td>SAT Math &lt;700 or miss.</td>
<td>STEM Crs logs Wk 1 &lt;=8</td>
</tr>
<tr>
<td>Logs per non-STEM crs, wks 2-6 &lt;29.1</td>
<td>Logs per STEM crs, wks 2 6 &lt;10.9</td>
</tr>
<tr>
<td>Avg. GPA = 3.63 N = 285</td>
<td>Avg. GPA = 2.52 N = 16</td>
</tr>
<tr>
<td>Avg. GPA = 3.40 N = 83</td>
<td>Avg. GPA = 1.30 N=11</td>
</tr>
<tr>
<td>Avg. GPA = 3.05 N=30</td>
<td>Avg. GPA = 2.90 N = 73</td>
</tr>
<tr>
<td>Avg. GPA = 3.76 N=26</td>
<td>Avg. GPA = 3.05 N=194</td>
</tr>
<tr>
<td>Avg. GPA = 3.52 N = 74</td>
<td>Avg. GPA = 3.76 N = 72</td>
</tr>
<tr>
<td>Avg. GPA = 3.59 N = 54</td>
<td>Avg. GPA = 3.03 N = 11</td>
</tr>
<tr>
<td>Avg. GPA = 3.13 N = 54</td>
<td>Avg. GPA = 3.49 N = 101</td>
</tr>
</tbody>
</table>

---

**Note:**
- SAT Math <700 or miss. indicates students who did not meet the Math Placement Exam requirement.
- STEM Crs logs Wk 1 < 8 indicates students who did not enroll in any STEM courses in the first week.
Analytics dashboard
Population monitoring and drill to detail

Double click on a bar to drill down to students records

Search for a student’s name or choose an ID

HS GPA and LMS login sliders
Final thoughts

Selected Technical considerations

- Information delivery
- Data quality/governance
- False positives/negatives

***Use of analytics is not just technical***

- Culture change
- Trust
- Ethics