Big Data and Learning Analytics in Education

Marcus D. Childress, Ph.D.
Dean, School of Education
Professor of Instructional Design and Performance Technology

Harvey C. Foyle, Ph.D.
EDD IDPT Coordinator
Professor of Instructional Design and Performance Technology

Baker University
Overview

• Big Data - Definition
• Who’s collecting data?
• How Target discovered that a teen girl was pregnant (before her father)
• Analytics – Definition
• Learning Analytics - Definition
• Civitas Learning *Community Insights* report
• The Future of Learning Analytics
Big Data – A Definition

• Extremely large data sets that have been analyzed computationally to reveal patterns, trends and associations, especially relating to human behavior and interactions.
Lots of Infographics
Who’s collecting data?

Social media and networks
(all of us are generating data)

Mobile devices
(tracking all objects all the time)

Scientific instruments
(collecting all sorts of data)

Sensor technology and networks
(measuring all kinds of data)
Who’s collecting data?

- Web data, e-commerce
- Purchases at department/grocery stores
- Bank/Credit Card transactions
- Social Networks
- Universities/Colleges
You

Social Media

Banking/Finance

Your Known History

Gaming

Entertainment

Purchases

You

BIG DATA

facebook

VISA

CHASE

AMERICAN EXPRESS

SAP

IBM

NETFLIX

hibu

NFL NETWORK

XBOX 360

zynga

PINTEREST

NORDSTROM

AMAZON

LOWE'S
How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

Charles Duhigg outlines in the New York Times how Target tries to

• Target assigns every customer a Guest ID number, tied to their credit card, name, or email address that becomes a bucket that stores a history of everything they’ve bought and any demographic information Target has collected from them or bought from other sources. Using that, Target looked at historical buying data for all the women who had signed up for Target baby registries in the past.
Analyzed the data and patterns emerged.

- **Lotions** - Women on the baby registry were buying larger quantities of unscented lotion around the beginning of their second trimester.

- **Supplements** – Sometime during the first 20 weeks, pregnant women buy supplements like calcium, magnesium and zinc.

- **Soap and Cotton Balls** – Pregnant women buy lots of scent-free soap and extra-big bags of cotton balls, in addition to hand sanitizers and washcloths; signals they could be getting close to their delivery date.
• Able to identify about 25 products that, when analyzed together, allows Target to assign each shopper a “pregnancy prediction” score.
• Could also estimate due date to within a small window.
• Target could send coupons timed to very specific stages of her pregnancy.
Fictional Target Shopper

- Jane Smith, who is 24, lives in Kansas City and in April, among other things, bought unscented lotion, calcium and zinc supplements. There’s an 87 percent chance that she’s pregnant and that her delivery date is sometime in early September.
Big Data in Higher Education

• Use data to improve education?
Student Data

Instructional
- Grades
- Quizzes
- Assignments
- Participation/attendance
- LMS Engagement Data
  - Time online
  - First/last accessed
  - # Online sessions
  - Discussions read/posted
  - Interactions with other users
  - Content viewed
  - Interface usage (ex view grades)

Other
- Audience
- Response Systems
- Where do they sit

Institutional
- Entry Information
  - Demographics
  - High school ePortfolio
- Previous Instructional Records
  - Past grades
  - Year of study
  - Program info
  - Teaching eval.
- Support Services Accessed
  - Library
  - Help desk
  - Counseling
## Analytics

<table>
<thead>
<tr>
<th>Academic/Institutional Analytics</th>
<th>Learning/Instructional Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td>A process for providing higher education institutions with the data necessary to support <strong>operational</strong> and <strong>financial</strong> decision making (adapted from Goldstein and Katz).</td>
<td>The use of analytic techniques to help target <strong>instructional, curricular,</strong> and <strong>support resources</strong> to support the achievement of specific learning goals (adapted from Bach).</td>
</tr>
<tr>
<td>Focus is on the <strong>business</strong> of the institution</td>
<td>Focus is on the <strong>learners</strong> and their behaviors</td>
</tr>
<tr>
<td>Primary audience – <strong>management</strong> and <strong>administrators</strong></td>
<td>Primary audience – <strong>learners</strong> and <strong>instructors</strong></td>
</tr>
</tbody>
</table>

Learning Analytics

• Learning analytics is the use of intelligent data, learner-produced data, and analysis models to discover information and social connections, and to predict and advise on learning. - George Siemens

http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/
• Learning analytics is the use of data and models to predict student progress and performance, and the ability to act on that information. - EDUCAUSE
Learning Analytics

- Reporting Data
  (Summarize historical data)

- Analyzing Trends
  (Identify historical trends and correlations)

- Predictive Analytics
  (Predict future outcomes and behaviors)
COMMUNITY INSIGHTS

Emerging Benchmarks & Student Success
Trends From Across The Civitas

VOLUME 1 | ISSUE 2 | DECEMBER 2016
What the Data are Saying

Benchmark 1

• Four LMS Activities are Consistently Most Predictive (academic performance)
  – Attendance - unique days visiting the online course (not count of logins)
  – LMS Grades - interim grades in the LMS
  – Course Material Engagement
  – Discussion Board Engagement
What the Data are Saying

Benchmark 2

• **Academic Performance is not the Primary Risk to Departure**

  – Across almost 4 million student records at 62 institutions, they found that 98.3% of these institutions are losing more students with 2.0 GPA or higher than below 2.0. A shocking 44% of the non-persisting students have a GPA of 3.0 to 4.0. **Most students who leave their college or university have GPAs at or over 2.0 and almost half have GPAs at or over 3.0.**
What the Data are Saying

Benchmark 3

• High School Performance May be a Better Indicator of Higher Ed Success than Norm Referenced College Admission Tests
  – They found no institutions for whom the SAT or ACT test scores were significantly more predictive than high school data.
  – The lowest predicted persistence rate is the group of students with above the median ACT scores but below the median High School GPA.
## FLY.COM FARE CALENDAR

<table>
<thead>
<tr>
<th>Su</th>
<th>Mo</th>
<th>Tu</th>
<th>We</th>
<th>Th</th>
<th>Fr</th>
<th>Sa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **March**
  - $166 (13, 15, 20, 22)
  - $166 (4, 5, 7, 11, 12, 19, 24)
  - $166 (20, 27)

- **April**
  - $166 (13, 15, 20, 22)
  - $166 (4, 5, 7, 11, 12, 19, 24)
  - $166 (20, 27)

- **May**
  - $166 (13, 15, 20, 22)
  - $166 (4, 5, 7, 11, 12, 19, 24)
  - $166 (20, 27)

- **June**
  - $181 (15, 22)
  - $198 (4, 5, 7, 11, 12, 19, 24)
  - $181 (20, 27)

- **July**
  - $181 (15, 22)
  - $198 (4, 5, 7, 11, 12, 19, 24)
  - $181 (20, 27)

- **August**
  - $198 (4, 5, 7, 11, 12, 19, 24)
  - $181 (20, 27)

- **September**
  - $198 (4, 5, 7, 11, 12, 19, 24)
  - $181 (20, 27)
A Decade Of U.S. Auto Sales By Vehicle Type

Cars
Pickup Trucks
SUVs & Crossovers
Vans


Sales (in units):
SECURITY & BIG DATA

• Yahoo email breach 2013-2014

• Automobiles hacked 2014

• Equifax breach 2017
SECURITY

• STRONG PASSWORDS

• SECURE ACCOUNTS

• CONSTANT VIGILANCE
The Future of Learning Analytics
Learning Analytics in Higher Education
The Future of Learning Analytics

• Increasing automation of learning analytics, data capture, reporting, and even solutions in response to problems identified by the data.
The Future of Learning Analytics

• Data from a greater variety of sources being used within learning analytics applications.
The Future of Learning Analytics

• More attention to the privacy and ethical aspects of learning analytics in terms of both evolution and practice; these considerations will be increasingly built into learning analytics technologies as a default.
The Future of Learning Analytics

- Will be shaped by battles about openness, especially about access to the core algorithms in predictive systems.
The Future of Learning Analytics

- Open analytics platforms that are modularized and extensible through open application programming interfaces (APIs) are likely to dominate the future of analytics.
The Future of Learning Analytics

- Will rely more heavily on and feed into cross-institutional repositories of analytics data in order to create a more robust benchmarking and predictive process.
The Future of Learning Analytics

• Will emphasize more advanced and personalized dashboards for students and instructors that will allow them to reflect on not just grades and other kinds of raw data.
The Future of Learning Analytics

• Will focus on more qualitative insights such as how the content of their work (or in the case of instructors, on the work of the class) scores on competencies such as verbal communication, teamwork, critical thinking, or creativity.
The Future of Learning Analytics

• Will show knowledge or content relationships between different courses or parts of courses and use social network analysis to show students’ performance on a range of different measures, not just grades, relative to the rest of the class or group.
Doctor of Education in Instructional Design & Performance Technology

BEGINS SPRING 2018
IDPT EDD DEGREE

• Orientation - Wednesday, April 11, 2018

• Classes Start – Monday, April 23, 2018

• RSVP to lheinen@bakerU.edu