

Understanding Student Experiences on Campus: The Seduction of Big Data

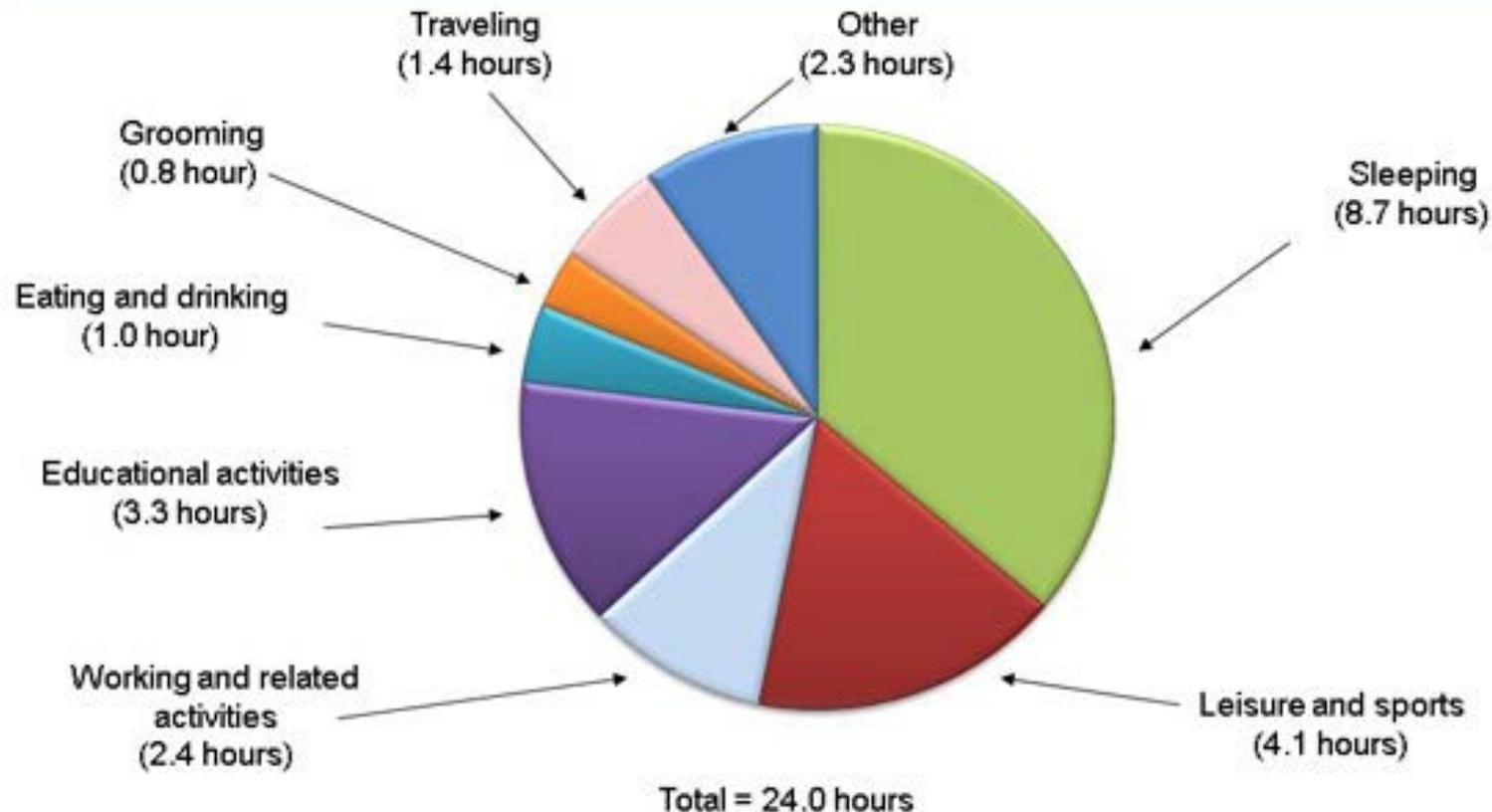
SONIA DELUCA FERNÁNDEZ, Ph.D.

sdf4@nyu.edu

NEW YORK UNIVERSITY



Time use on an average weekday for full-time university and college students



NOTE: Data include individuals, ages 15 to 49, who were enrolled full time at a university or college. Data include non-holiday weekdays and are averages for 2010-14.

SOURCE: Bureau of Labor Statistics, American Time Use Survey

Assessment process

- Mission
- Goals and outcomes
- Educational opportunities, activities
- Assessment
- Review

Headlines

Taming "Big Data": Using Data Analytics for Student Success and Institutional Intelligence

Big Data's Coming Of Age In Higher Education

Welcome to the Era of Big Data and Predictive Analytics in Higher Education

FOR BIG DATA EXPERTISE THIS IS THE BIG TIME.

What Universities Can Learn from Big Data

Improving Higher Education Performance with Big Data



What is big data?

- Process, approach, and product
- Massive amounts of information, point-in-time bits
- Real-time generated, and warehoused (pseudo-synchronous and asynchronous)
- May or may not be attached to a context
- Structured (e.g., minutes on webpage, clicks) or unstructured (e.g., a tweet)
- Undetermined consent, use

Big data usage in higher education

Enrollment management and marketing

- How does student behavior relate to bill paying?
- How can we get more of [these/those] students?

Finance and budgeting

- How do micro-adjustments affect revenue generation?

Persistence prediction, student tracking for “success”

- When or where are students at risk for dropping out?

Learning analytics – Leveraging LMS

[Assumes a positivistic approach]

Big data sources in higher education

- MOOCs
 - Time-on-task
 - Access
 - Patterns
- Social media
 - Text searches
 - Location and movement

“...data mining [students’] digital footprint...”

DATA COLLECTED IN 00:00:50 SECONDS

694,445
GOOGLE SEARCH QUERIES

20,000+
NEW POSTS
ON TUMBLR.

1,600+ READS ON
Scribd.

168 MILLION
EMAILS SENT

60+ NEW
VIDEOS
.....
25+ HOURS
TOTAL DURATION

1,500+
BLOG POSTS
.....
60+ NEW BLOGS

70+ DOMAINS
ARE REGISTERED

13,000+ iPhone
Applications Downloaded

695,000+ facebook
STATUS UPDATES

79,364
WALL POSTS

510,040
COMMENTS

320+ NEW
TWITTER
ACCOUNTS
.....

50+ WORDPRESS
DOWNLOADS
.....

12,000+ NEW ADS
POSTED ON CRAIGSLIST

98,000+ TWEETS

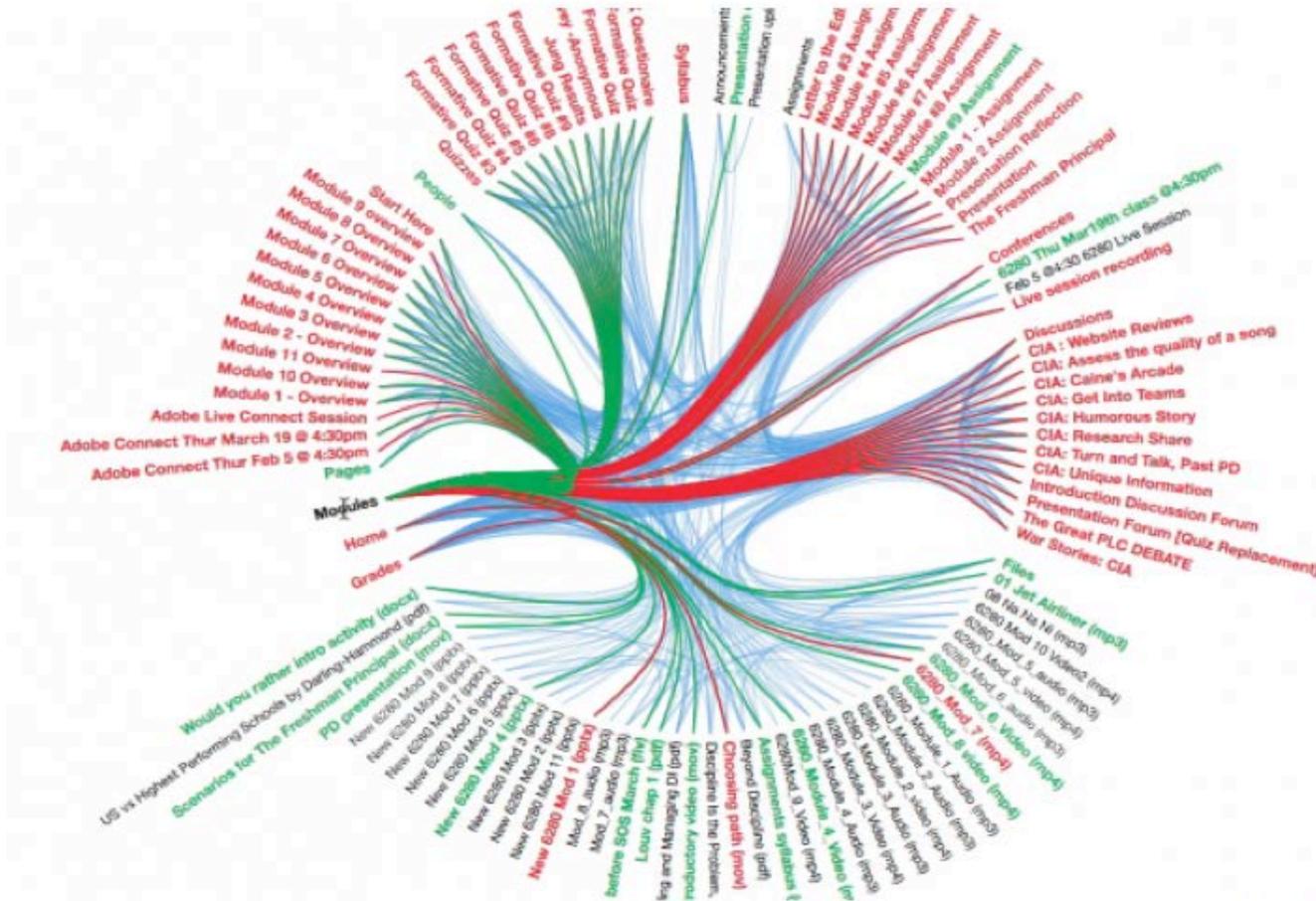
125+ PLUGIN DOWNLOADS

1700+ FIREFOX DOWNLOADS

Big data issues

- Data storage and security
- Challenges with alignment: inquiry and data generation
- Analysis and interpretation
- Distance from scientific methods

Learning analytics



Courtesy of Utah State U.

A data visualization produced by Utah State U. depicts student activity in one online course.

Pitfalls of big data

Misalignment between inquiry focus and data collection

- In a typical week, how do students spend their time?
- Data source: Twitter, Foursquare, and building swipes from the last five years

Under-examination of the inquiry assumptions

- Students engage social media and use Twitter, Foursquare
- Students use social media in similar quantities and qualities over a given time period
- All students engage social media similarly
- Students' social media usage reflects how they spend their time
- Students represent themselves accurately and truthfully on social media
- All students swipe into buildings, at all times

Pitfalls, continued

Inappropriate data analyses

- Factor analysis: means obscure group differences
- Large sample sizes can image random data points into patterns, spurious correlations
- Reliability tests can fail
- Can we tell which students in our inquiry have (some, all) missing data? How would unstructured data be imputed or extrapolated?
- How are outliers identified? How are outliers valued?

Errors in interpreting results

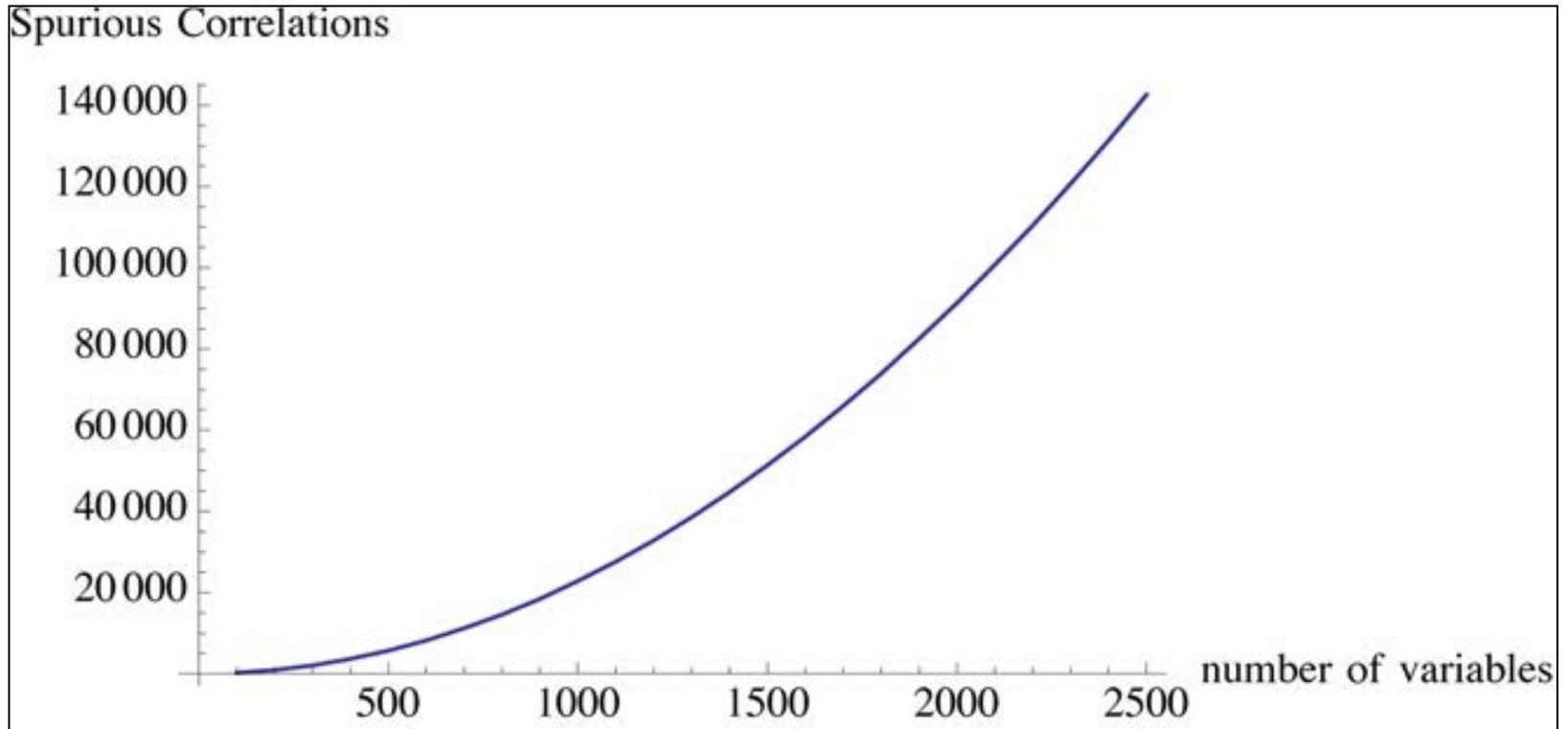
- Misapplication and generalizations made
- #allstudents
- Correlation equated with causation



Umm... is this a mess?

Fatal flaw?

...big data means anyone can find fake statistical relationships, since the spurious rises to the surface. This is because in large data sets, large deviations are vastly more attributable to variance (or noise) than to information (or signal). –*N. N. Taleb (2013)*



This is the tragedy of big data: The more variables, the more correlations that can show significance. Falsity also grows faster than information; it is nonlinear (convex) with respect to data (this convexity in fact resembles that of a financial option payoff). Noise is antifragile. *Source: N.N. Taleb*

The traditional approach to assessment

- Preoccupied with methods (not methodology) and something shiny (e.g., dashboards, KPIs, big data)
- Fails to provide challenge to how big data usage ignores questions of research theory, assumptions, interpretation, and ethics.
- Privileges particular types of data (i.e., experiences)
- Value-free, neutral, apolitical



This doesn't feel right.

Acceptance of traditional assessment

Reliance on methods and tools to generate a particular kind of data. This has created a...

- Rejection of underrepresented voices, experiences
- Limited and partial truths
- Over-valuation of existing systems, particular knowledge
- Reproduction of inequality and reification of white supremacy

Critical approach to assessment

- Inquiry is a social and political activity
- Make assumptions, hopes, and intentions, explicit
- Consider and assume agency of person behind data point, and role of researcher
- Expose and address power, privilege, and structures
- Consider thoughtfully histories and contexts
- Eschew colorblind and ideological neutral claims
- Apply knowledge to the improvement of the organization (e.g., institution), as opposed to the assimilation of the individual

Supporting equity?

- Increasing scale of information and analyses in PWIs and HWIs may further silence experiences of students of color
- Big data is unusually susceptible to reproducing discrimination and privilege through researcher bias, confirmation bias, and an emphasis on iterative model generation
- Unstructured data is analyzed in similar (if not the same) ways as structured data – challenges with identifying and interpreting distinctions (e.g., jargon)
- When context is generalized, high levels of inference are required (e.g., relationship between time on webpage and test score)

Understanding students' experiences

- How are the most vulnerable on our campuses being served by our assessment practice?
- Change the emphases:
 - **Statistics and sound bites** → descriptions and possibilities
 - **Educational outputs** → experiences and learning
 - **Student remediation** → institutional accountability
- For example: the TILT project (UNLV and AAC&U, Winkelmes, 2016)



Now what do we do?

Critical assessment

An equity-minded approach to understanding student experiences on campus requires answering...

- What is real? How do social, political, cultural, and economic lenses influence what we consider real and legitimate? (ontology)
- What (cultural) norms do we privilege? To what extent do we know and make explicit biases and values? (axiology)
- What constitutes knowledge, evidence? In assessment, which ways of knowing, and what kinds of evidence do we privilege? (epistemology)

Critical race theory

- Racism is real and pervasive. (ontology)
- Recognize the falsehood of dominant ideologies and “master narratives” (e.g., neutrality, colorblindness, meritocracy). (ontology)
- Interdisciplinarity is required. (epistemology)
- Privilege counternarratives and lived experiences. (epistemology and methodology)
- Make explicit transformation and liberation (a social justice project). (axiology)

We need ...

- The right people at the right table –expertise attached to appropriately identified issues
- The right questions with assumptions identified
- Good, authentic information
- To interpret information in context and with ample scrutiny – computation, social meaning making
- To balance between the general and the particular, erring on the side of (vulnerable) student interests.

Moving inquiry closer to student experiences

- Explore and engage your worldview.
- Investigate and discuss **who decides, who benefits.**
- Confront emphases on procedure and methods.
- Consider effects and justifications of “market” and economic drivers.
- Discuss and critique how meaning is attached to data, results. (Data does not own meaning.)
- Consider how value is attached to what is measured – how is work prioritized?
- Recognize and address the extent to which assessment work prevents social justice and structural transformations.

Resources

- Cabrera, N. L. (2012). Working through Whiteness: White male college students challenging racism. *The Review of Higher Education*, 35(3), 375-401.
- Gillborn, D. (2007). Education policy as an act of white supremacy: whiteness, critical race theory and education reform. *Journal of Education Policy*, 20(4), 485-505.
- Hardt, M. (26 Sept 2014). How big data is unfair: Understanding sources of unfairness in data driven decision making. *www.medium.com*
- Martinez-Aleman, A. M., Pusser, B., & Bensimon, E. M., Eds. (2015). *Critical approaches to the study of higher education: A practical introduction*. Baltimore: Johns Hopkins University Press.
- Stephens, N. M., et al. (2012). Unseen disadvantage: How American universities' focus on independence undermines the academic performance of first-generation college students. *Journal of Personality and Social Psychology*, 102 (6), 1178- 1197.
- Wall, A. F., Hursh, D., & Rodgers, J. W. (2014). Assessment for Whom: Repositioning Higher Education Assessment as an Ethical and Value-Focused Social Practice. *Research and Practice in Assessment*, 9 (Summer 2014), 5-17.

SONIA DELUCA FERNÁNDEZ, Ph.D.

NEW YORK UNIVERSITY

sdf4@nyu.edu

