Applying Predictive Analytics to Educational Challenges: *Ethical Issues*

Kim Darnell Aidan Feay Jessica Hays Fisher

Final Project W231 Fall 2018

Berkeley SCHOOL OF INFORMATION

Educational Quality and Efficacy Matter









What are the benefits and consequences of attempting to improve educational outcomes using the tools of *data science*?



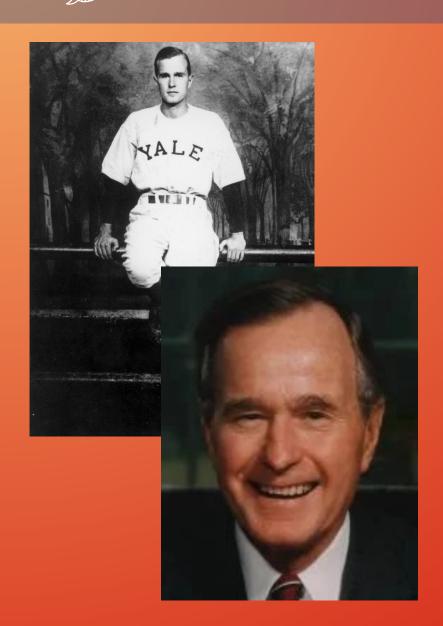
Machine Learning

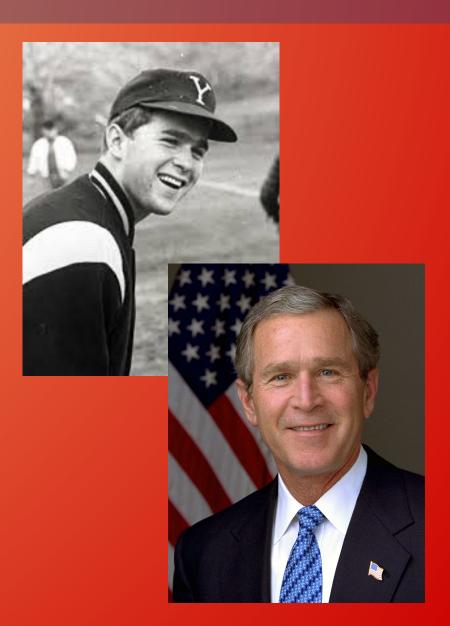


The Data Bear the Burden of *History*



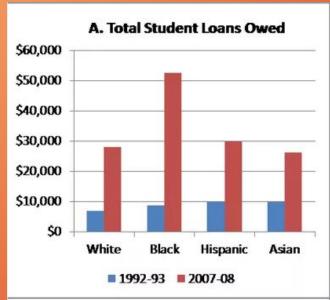
Privilege Perpetuates Itself

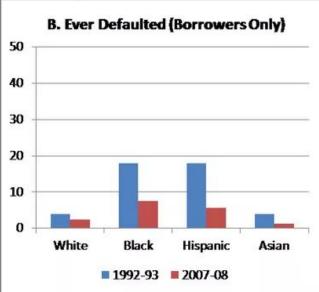


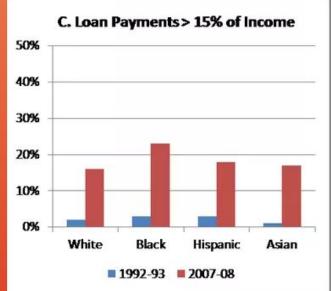


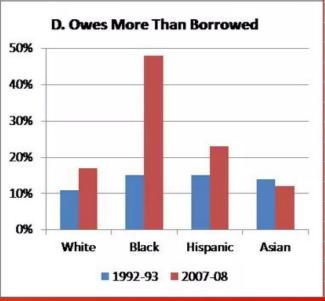


Growing Privilege is Increasingly Expensive

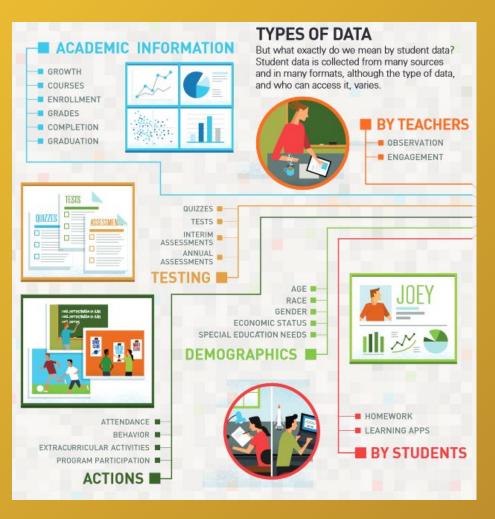








Data Science in K-12



- Standardized testing
 - More data, more granular tracking
- Approaches vary
 - Many public school systems dedicating few/no resources at district level
- Rise of Ed-tech companies
 - Outsourcing of supplemental data gathering, analysis, sharing of insights

Case Study: Tacoma Public Schools

- Dashboards inform planning for teachers and principals
- Shared data to drive transparency for students, parents, community
- Early interventions enabled and encouraged



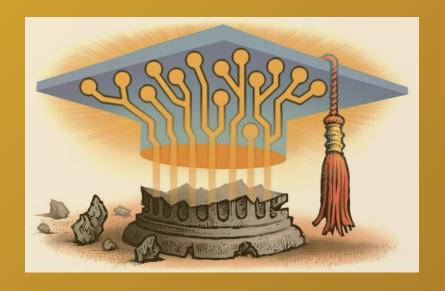
Graduation rates have risen by 28%; individual schools have seen even more drastic results



Data Science in Higher Education

Most schools are trying to do more with fewer public resources and greater public expectations

- The traditional stuff
 - Application information
 - School-related activities
 - ► Financial details
- ▷ Plus...
 - Online activities
 - Microassessment of progress
 - All on-campus interactions with technology



Case Study: Georgia State University



- Tracking students' progress through major programs
- Encouraging tutoring and advising through alerts
- Targeting high risk students for microgrants
- Applying adaptive learning tools across the curriculum

Georgia State University (cont.)

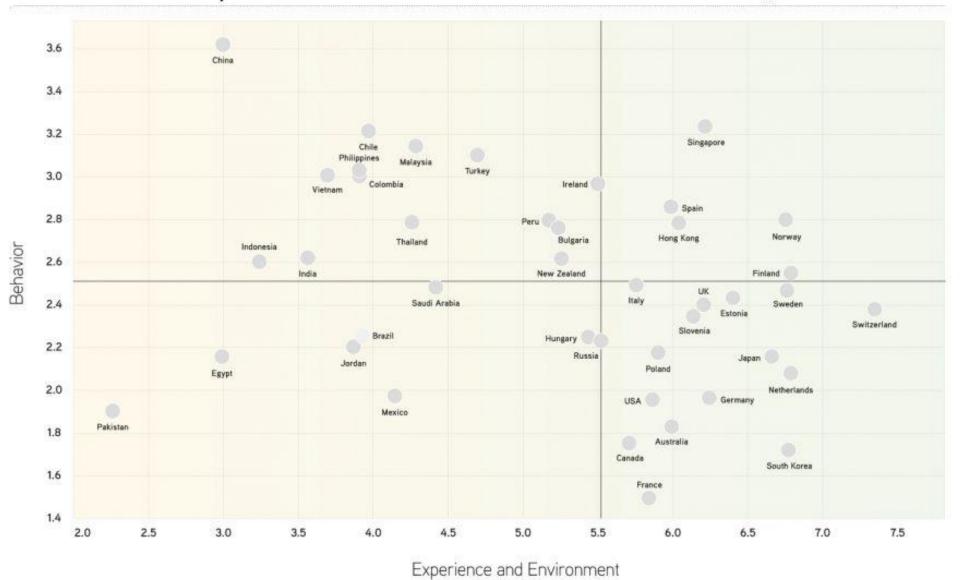
- Graduation rates up 2%per year since 2013
- Graduates finish sooner, saving millions in tuition
- The graduation gap has been closed!
- Students from all backgrounds are graduating at higher rates in high demand majors



Motivations of Participants: Qualitative

- Why are students studying?
 - Learning
 - Performance
- Do they trust technology?
 - Trust deficit





Motivations of Participants: Quantitative

- How do engagement metrics in education differ from traditional media consumption?
 - YouTube views vs lecture views
 - Students revisit sections of lectures in a way inconsistent with other media viewing habits
- "Measure YouTube views? Your employees will strive for more and more views."
 - Jeff Bladt and Bob Filbin



Motivations of Practitioners: Conscious

- Limited School Transparency
- Trade-offs
 - Altruism vs Profit
 - Difficulty of Measurement
 - ► Re-identification
 - ► FERPA

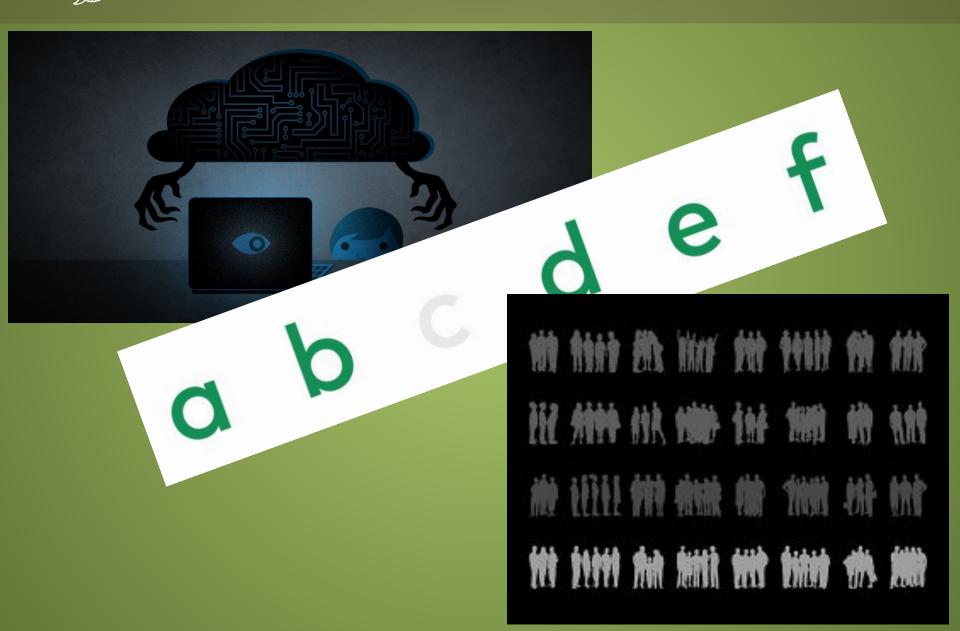


Motivations of Practitioners: Subconscious

- Lack of Domain Knowledge
- Hacking
 - Loss of trust
 - Public image
 - Ransoms
- Training
- Data Classification



Questions for Consideration: *Individuals*

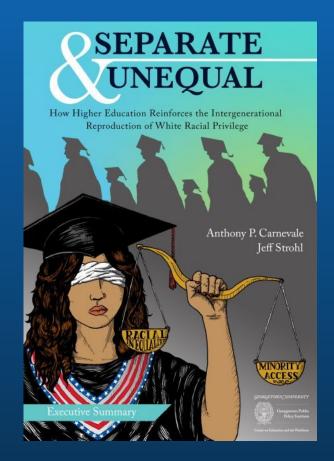




Data and Their Applications are Not Unbiased

Using big data to solve our educational problems runs the risk of perpetuating or exacerbating inequity

- Women and minorities are directed to lower paying majors, because that's what they were permitted to study in the past
- Colleges and universities often favor recruiting students from higher socioeconomic groups, because these students present the least risk in economic and academic terms
- Higher prestige schools show bias toward recruiting white students



How Can We Limit Harm?

- Diversity
 - Programmers
 - Decision-makers
 - Reviewers
 - ► Interviewers/Interviewees
- Clarity
 - Rationales
 - Metrics
 - Algorithms
 - Informed consent
- Opt-out



Questions? Comments?