(I) WHAT IS THIS TALK ABOUT
Data Science

Social Justice

Education

actually vast
WHAT IS THIS TALK ABOUT

NOT A
EDUCATING ABOUT SOCIAL JUSTICE ISSUES BY USING DATA
EDUCATING ABOUT SOCIAL JUSTICE PROBLEMS EXACERBATED BY DATA

For truly ethical AI, its research must be independent from big tech | Timnit Gebru

A year ago I found out, from one of my direct reports, that I had apparently resigned. I had just been fired from Google in one of the most disrespectful ways I could imagine.

My name is Joy Buolamwini. I’m a poet of code on a mission to stop an unseen force that’s rising. A force that I call the coded gaze — my term for algorithmic bias.
EDUCATING ABOUT DATA IN A WAY THAT IS SOCIALLY JUST

Data Science as a Foundation for Inclusive Learning

by Robert A. Lue

Published on Nov 19, 2019

Inclusivity in Statistics and Data Science Education

Jeff Witmer

To cite this article: Jeff Witmer, Inclusivity in Statistics and Data Science Education, Journal of Statistics and Data Science Education 29:1 (2021) 2—3
Using a Latent Class Forest to Identify At-Risk Students in Higher Education

Kevin Pelaez
San Diego State University

Richard A. Levine
San Diego State University

Maureen Guarcello
San Diego State University

Mark Laumakis
San Diego State University

Juanjuan Fan
San Diego State University

Geographical Network Analysis and Spatial Econometrics as Tools to Enhance our Understanding of Student Migration Patterns and Benefits in the U.S. Higher Education Network

Manuel S. González Canché

Black Economists on Race and Policy: Contributions to Education, Poverty and Mobility, and Public Finance

Dania V. Francis, Bradley L. Hardy, and Damon Jones
Critical race theory (CRT) in education centers, examines, and seeks to transform the relationship that undergirds race, racism, and power. CRT scholars have applied a critical race framework to advance research methodologies, namely qualitative interventions. Informed by this work, and 15 years later, this article reconsiders the possibilities of CRT applications to quantitative methodologies through QuantCrit… We argue that quantitative approaches cannot be adopted for racial justice aims without an ontological reckoning that considers historical, social, political, and economic power relations. Only then can quantitative approaches be re-imagined and rectified.
WHAT IS THIS TALK ABOUT
Now with FIVE SOCIETIES!!!

Mathematics majors need “a command of data analysis and statistical inference at a level equivalent to that attained in an applied data analysis course.”
DOES THE MATH MAJOR* REQUIRE AN APPLIED DATA ANALYSIS COURSE?

<table>
<thead>
<tr>
<th>Universities with largest enrollments:</th>
<th>“Top” liberal arts colleges</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Central Florida</td>
<td>X</td>
</tr>
<tr>
<td>2. Texas A&amp;M</td>
<td>X</td>
</tr>
<tr>
<td>3. Ohio State</td>
<td>X</td>
</tr>
<tr>
<td>4. Florida International</td>
<td>X</td>
</tr>
<tr>
<td>5. Florida</td>
<td>X</td>
</tr>
<tr>
<td>7. Arizona State</td>
<td>X</td>
</tr>
<tr>
<td>8. Illinois</td>
<td>X</td>
</tr>
<tr>
<td>10. Texas – Austin</td>
<td>X</td>
</tr>
</tbody>
</table>
WHAT IS THIS TALK ABOUT

inspiring you to incorporate data skills into undergraduate mathematics education in ways that involve social justice concepts*

*This is not by any stretch social justice activism, though arguably it contributes to supporting conditions
(II) TEACHING ABOUT/FOR JUSTICE
This course will help you understand the world through a mathematical lens and will develop your powers of argumentation and critical thinking. We will explore and utilize diverse areas of discrete mathematics including logic, set theory, functions and relations, combinatorics, probability, networks, and more. We also will discuss methods and styles of mathematical proofs in order to prepare you for more advanced math courses. Finally, while mathematical knowledge is often perceived as being “pure,” we will highlight some ways in which it is socially constructed and hence subject to human limitations and biases.
“Johnstone insisted that my duality theorem is ‘folklore,’ whilst admitting at the same time that no statement let alone any proof of it appears in the literature preceding my preprint, on the sole grounds that “all the elements in the proof were known long before you wrote them down”… and on the pretention that any expert in the subject could have put them together to prove the theorem.”
Epistemic injustice in mathematics

Colin Jakob Rittberg\textsuperscript{1} · Fenner Stanley Tanswell\textsuperscript{2} · Jean Paul Van Bendegem\textsuperscript{1}

Received: 31 January 2018 / Accepted: 8 October 2018
© Springer Nature B.V. 2018

Abstract
We investigate how epistemic injustice can manifest itself in mathematical practices.
Diversity in proof appraisal
M Inglis, A Aberdein - Mathematical cultures, 2016 - Springer
We investigated whether mathematicians typically agree about the qualities of mathematical proofs. Between-mathematician consensus in proof appraisals is an implicit assumption of …

Save  Cite  Cited by 26  Related articles  All 11 versions
1. The authors create social networks in which nodes are Twitter users. What two approaches do they use for assigning edges?

2. What are some measure of network structure that the authors use to quantify and study the networks?

3. What is the authors' goal in using those measures? What are they trying to identify?

4. What is the #AllMenCan network?

5. How is the #CrimingWhileWhite network different (structurally) from the #AllMenCan network?
Mathematical Modeling

Math 433
Prof. Chad Topaz

WHY TAKE THIS COURSE

Just as you can strive to understand the world through the lenses of art, economics, history, sociology, and more, you can similarly strive to understand it through the lens of mathematics. Mathematical modeling refers to translating real (or imagined!) phenomena into a mathematical object, using mathematics to study that object, and in turn, discovering things about the real (or imagined!) world.
Goals
- Art through a data-driven modeling lens
- Artists with marginalized identities
- Celebrate, not co-opt
- Consider artistic process at granular level

Process/Deliverables
- Select artist(s) and artwork(s)
- Take data (image acquisition, processing)
- Generate art (algorithms)
- Artists’ statement (writing)
Computational Models of Abstract Art
Mathematical Modeling

Math 433
Prof. Chad Topaz

COMPUTATIONAL MODELS
OF ABSTRACT ART

Mandy El-Sayegh, Net-Grid (2021)

Saisha Goboodun, Porter Johnson, Aayushi Pramanik (2021)
Affirmative action, critical mass, and a predictive model of undergraduate student body demographics

Daniel P. Maes, Julia Tucher, Chad M. Topaz

Published: May 12, 2021 · https://doi.org/10.1371/journal.pone.0250266
Markov Chain Model
RESULTS FOR UC BERKELEY

Validation

Prediction

Race/Ethnicity
- Black
- Asian American
- Hispanic/Latinx
- White

Year 2010 2012 2014 2016 2018

Year 2020 2022 2024 2026 2028

Demographic Proportion
New York City jails: COVID discharge policy, data transparency, and reform

Eli Miller, Bryan D. Martin, Chad M. Topaz

Published: January 19, 2022 · https://doi.org/10.1371/journal.pone.0262255
De Blasio said he has been working closely with NYPD Commissioner Dermot Shea, the Mayor’s Office of Criminal Justice, the Department of Correction and district attorneys to come up with a list of people who are at high risk of contracting coronavirus and are at low risk of reoffending.
RESULTS

“There was a dramatic and unprecedented reduction in the New York City jail population… In concordance with de Blasio’s stated policy, we find that being discharged during the focus week is associated with a lower probability of readmission as compared to being discharged during the same calendar week in previous years.”
Race- and gender-based under-representation of creative contributors: art, fashion, film, and music

Chad M. Topaz, Jude Higdon, Avriel Epps-Darling, Ethan Siau, Harper Kerkhoff, Shivani Mendiratta & Eric Young

Nature Humanities and Social Sciences Communications 9, Article number 221 (2022)
Undergraduate Research

Williams College

INDEPENDENT STUDY / RA-SHIP
Data ethics

“The American Statistical Association (ASA) provides ethical guidelines for statistical practice (American Statistical Association, Committee on Professional Ethics, 2018), addressing over 50 issues related to professional accountability, data and methodological integrity, public transparency, responsibility to research subjects, and more. We have considered all guidelines in detail, and we now mention a few that play a special role in our study.”
1 Introduction

Between 2010 and 2020, the U.S. multiracial population – individuals with more than one racial identity – grew from 9 million to 33.8 million (US Census, 2020). Multiracial children and adolescents compose an increasingly significant portion of the student population (McCubbin, 2013), yet their educational outcomes remain understudied (Charmaraman et al., 2014). The study of multiracial students poses complex ethical issues, not least because researchers remain undecided about the best way to identify multiracial populations. Most racial identification data is collected through surveys, but the way survey respondents are asked about their racial identity and ethnicity has a significant impact on their responses (Patten, 2015; Parker et al., 2015).
High School Longitudinal Study '09

Original:
1,941 (9%)
Multiracial

Adjusted:
5,316 (25%)
Multiracial
1 Introduction

We investigate the impact of legacy admissions—the practice at many highly selective colleges and universities of preferring to admit students who are related to alumni—on campus racial diversity. Legacy students have a 45% admissions advantage over non-legacy students at the most selective US universities. Recent lawsuits against Harvard have made previously inaccessible admissions data available, revealing the racial composition of legacy admits as 43% white and less than 16% African American, Asian American, and Hispanic. We use data from the Common Data Set for over 40 highly selective colleges and universities from 2016-20 to analyze the relation of stated admissions criteria to racial demographics on campus.
C7. Relative importance of each of the following academic and nonacademic factors in your first-time, first-year, degree-seeking (freshmen) admission decisions.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Very Important</th>
<th>Important</th>
<th>Considered</th>
<th>Not Considered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rigor of secondary school record</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class rank</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic GPA</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardized test scores</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Application essay</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendation(s)</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interview</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extracurricular activities</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Talent/ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Character/personal qualities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First generation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alumni/ae relation</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Geographical residence</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>State residency</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Religious affiliation/commitment</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Racial/ethnic status</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volunteer work</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work experience</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level of applicant interest</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
MODEL PREDICTIONS

- Observed
- Predicted for no-legacy
(III) BUILDING YOUR SKILLS*

*IF NECESSARY
SOME STARTING POINTS

Modern Data Science with R
Second Edition

Posit Cloud lets you access Posit's powerful set of data science tools right in your browser—no installation or complex configuration required.

GET STARTED  ALREADY A USER? LOG IN
SOME STARTING POINTS

Data Science Specialization
Launch Your Career in Data Science. A ten-course introduction to data science, developed and taught by leading professors.

4.5 38,574 ratings

Jeff Leek, PhD  2 more instructors

Enroll for Free  Starts Jul 17

Financial aid available

478,525 already enrolled
SOME STARTING POINTS

Your Campus
“The way to right wrongs is to turn the light of truth upon them.”

– Journalist, educator, and civil rights pioneer Ida B. Wells
QSIDE

- 501(c)3 tax-exempt nonprofit
- Established in 2019
- Research-to-Justice-to-Action projects
- Education and community building
EDUCATION: DATATHONS

HIGH SCHOOL DATATHON4JUSTICE

We are happy to be working with the following schools:
- Chesapeake Math and IT Academy (Laurel, Maryland)
- Glendale Unified School District (Glendale, California)
- Paramount School System (Paramount, California)
- Plainfield Community School Corporation (Plainfield, Indiana)

If you have any inquiries, feel free to reach us at:
- qside@qsideinstitute.org
- @qsideinstitute
Acknowledgments

Introduction

3 Understand the data and the research study
   3.1 Define the type of study
   3.2 Understand how the museums in the study were chosen
   3.3 Define our variables and observations
   3.4 Download data using RStudio
   3.5 Key concepts, commands, and techniques

4 Get data ready to analyze in RStudio
   4.1 Describe the structure of the data
   4.2 Describe the content of data
   4.3 Prepare data for analysis
   4.4 Key concepts, commands, and techniques

5 Explore, summarize, and visualize a single variable
   5.1 Create tabular displays of categorical data and summaries of numerical data
   5.2 Create plots to summarize and visualize categorical and numerical data
   5.3 Key concepts, commands, and techniques

6 Infer from data for a single variable
   6.1 Use statistics to make inferences about population parameters based on a random sample from that population
   6.2 Create confidence intervals for categorical and numerical data
   6.3 Key concepts, commands and techniques

7 Explore, summarize and visualize relationships between two variables
   7.1 Create tabular displays of two variables
   7.2 Create plots to visualize two variables
   7.3 Key concepts, commands and techniques

Appendix A: American Community Survey Data
Appendix B: Global World Population
EDUCATION: CONFERENCES

Data4Justice Conference
April 16, 2021 via Zoom

Plenary Lecture:
Heather McGhee
Former president of Demos, American political commentator, co-chair of Color of Change, and author of The Sum of Us: Why Racism Costs Everyone and How We Can Prosper.

Max Jordan Nguegni Tako
Yale University School of Medicine
"Victimization of Justice: Understanding of Justice and Violence as a Barrier Against Health Equity"

Dr. Jasmine Drake
Texas Southern University
"The Intersectionality of Forensic Science and the Criminal Justice System"

Dr. Diana Greenwald
Isabella Stienart Gardner Museum
"Why Have There Been No Great Women Artists? Artistic Labor and Time Constraint in Nineteenth-Century America"

Dr. Srivi Ramasubramanian
Texas A&M University
"Data for Social Justice in Media Content"

Dr. Lola Fajana
Duke University School of Medicine
"Diversity, Representation, and Access in the Practice of Inclusive Research"

Dr. Christian Smith
Biola College
"Identifying Racially Discriminatory Federal Judges"

Dr. Alisa Richardson
USC Annenberg School of Communication
"The Daring Archival Crisis: How Social Media Disappears Black Witnessing"

Dr. Anne-Marie Núñez
The Ohio State University
"Data for Equity in Education"

Lourdes Vera
Northeastern University
"A Community-Based Tool to Measure Hydrogen Sulfide Near Oil and Gas Development"

Dr. Alejandro Ponce
World Justice Project
"Environmental Governance Indicators"

Keynote Speaker
Danielle Wood
Scholar of societal development with a background in science design, earth science applications, systems engineering, and technology policy.

Data4Justice Conference
A virtual, all-day event for scholars, activists, policymakers, and allies to explore the ways we can power social justice together. Want to learn more about how data, research, and activism come together to create change? Join us!
April 22, 2022 via Zoom

Speakers and discussion surrounding five Research Focus Areas:

- Health Care Equity
- Criminal Justice
- Inclusion in Arts/Media
- Environmental Justice
- Education Equity
EDUCATION: SEMINARS (YOUTUBE!)

Autumn Asher BlackDeer
Decolonizing Data: A Quantitative Native Approach to Indigenous Mental Health and Higher Education
October 28, 2021
4-5pm EST | Zoom
QSIDE Colloquium

Samuel Sinyangwe
Using Data to Fight Police Violence
October 14, 2021
4-5pm EST | Zoom
QSIDE Colloquium

Molly Coleman
Unrigging the Law: Building a Civil Legal System that Works for the People
September 30, 2021
4-5pm EST | Zoom
QSIDE Colloquium 2021

Size Matters to Lesbians Too: Queer Trans Feminist Interventions into Data, Algorithms, and Visualizations
Jen Jack Giesekeing
January 27, 2022
4-5pm EST | Zoom
QSIDE Colloquium

Taj Mustapha
A Road to Inequity Paved with Good Intentions: Data Science and Health Care Delivery in the US
February 11, 2022
4-5pm EST | Zoom
QSIDE Colloquium 2021

Gaelan Smith
Data Visualization Engineer and Knowledge Manager
February 4, 2021 4:00pm EST
QSIDE Colloquium 2021
Daniel Bullman
Peer Recovery Specialist, graduate student
Project: Continuity of care for justice impacted individuals

Kaniqua Outlaw
Community health worker, graduate student
Project: Criminal sentencing disparities in North Carolina

Ryan Rowe
Undergraduate student
Project: Social media’s impact on anti-qualified immunity legislation

Christopher Dunstan
Graduate student, dancer
Project: Crediting dance pioneers
A FEW OF OUR PARTNERS

USC Race and Equity Center
Our mission is to ILLUMINATE, DISRUPT, and DISMANTLE racism in all its forms.

National Gallery of Art

Life Unseen

Responsible Business Initiative for Justice

Perseverer™
CONSORTIUM

- Bates College
- Bard College
- Clarkson University
- Colgate University
- Colorado State University
- University of Colorado – Boulder
- Duke University
- Hamilton College
- Lewis University
- Macalester College
- Michigan State University

- Northwestern University
- Rice University
- Seattle University
- Southwestern University
- Swarthmore College
- Texas Southern University
- University of the District of Columbia
- University of Utah
- Voorhees College
- Washington and Lee University
- Wofford College
GETTING INVOLVED
WWW.QSIDEINSTITUTE.ORG

• Sign up as an affiliate (announcements, newsletters, etc.)
• Have your department/school/institution join the consortium
• Use the Data4Justice case study (soon, “studies”)
• Participate in research labs
• Participate in data hackathon
• Attend Datathon4Justice
• Participate/support SQuARED Justice (student online conference)
• Watch recorded talks
• Reach out if you need something, have questions, etc.