

The background of the slide features a close-up of an olive branch with several green olives. The branch is positioned diagonally across the frame. The background is a soft, out-of-focus sunset or sunrise sky with warm orange and yellow tones. A semi-transparent dark grey rectangular box is centered over the image, containing the main title and speaker information in white text.

# ADVANCING ANTI-RACISM, DIVERSITY, EQUITY AND INCLUSION IN DATA SCIENCE FOR SOCIAL IMPACT

Francisca García-Cobián Richter

March 8, 2023

Center on Poverty and Community Development

Jack, Joseph and Morton Mandel School of Applied Social Sciences

Case Western Reserve University

A photograph of a rocky path with a white text box overlay. The path is made of large, light-colored rocks and is surrounded by green grass and small plants. The text box is in the upper left corner and contains the title of the slide.

**SOCIAL JUSTICE IN AN ERA  
OF DATA ANALYTICS:**

**LOOKING TO THE PAST TO  
BUILD A BETTER FUTURE**

Decision making in health and human services is increasingly aided by data technologies.

“... an algorithm that screens for child neglect could harden racial disparities”

PBS News 04/29/2022

“Widely used health care algorithm has racial bias”

Harvard School of Public Health, 2019

“... police data [that] predictive policing programs rely on are skewed toward crimes committed by people of color”

William Isaac, Human Rights Data Analysis Group, 2016

**CERTIFICATE IN DATA  
SCIENCE FOR SOCIAL  
IMPACT AT CWRU  
(DSSI)**

Offered by the Jack, Joseph and Morton Mandel School of Applied Social Sciences and the Case School of Engineering, CWRU

12-credit, digitally certified

For students in graduate and professional programs at CWRU

Integrating community knowledge into data science to advance social justice within an ethical framework.



**ONE YEAR PROGRAM  
IN PERSON OR REMOTE**

**SASS 471: Introduction to Data Science  
for Social Impact**

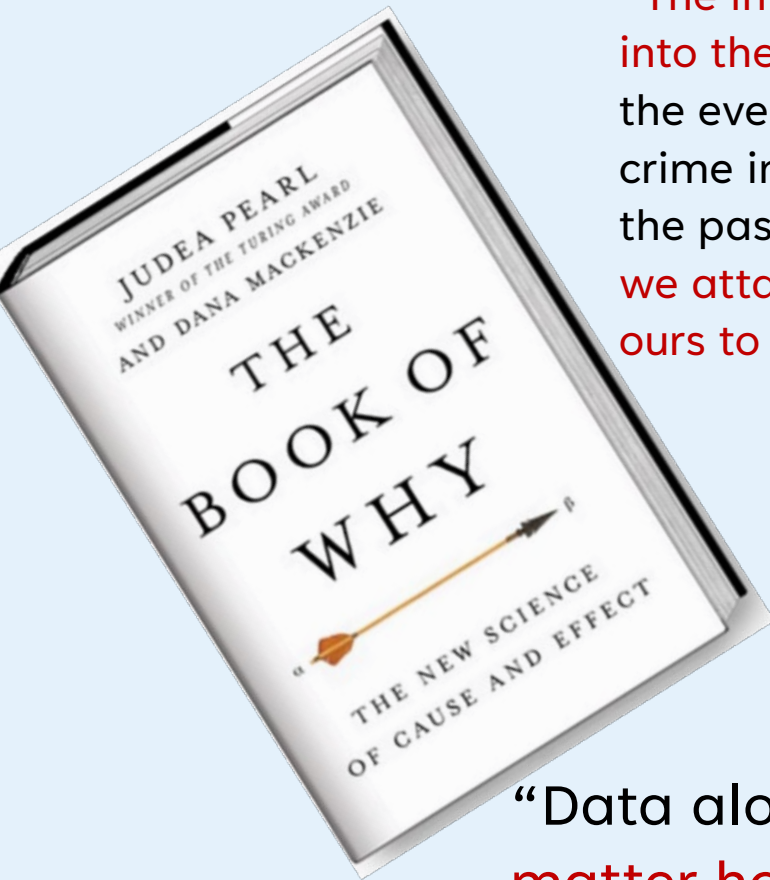
**DSCI 451: Exploratory Data Science**

**DSCI 453: Modeling and Prediction**

**SASS 472: Semester Research Project in  
Data Science for Social Impact**

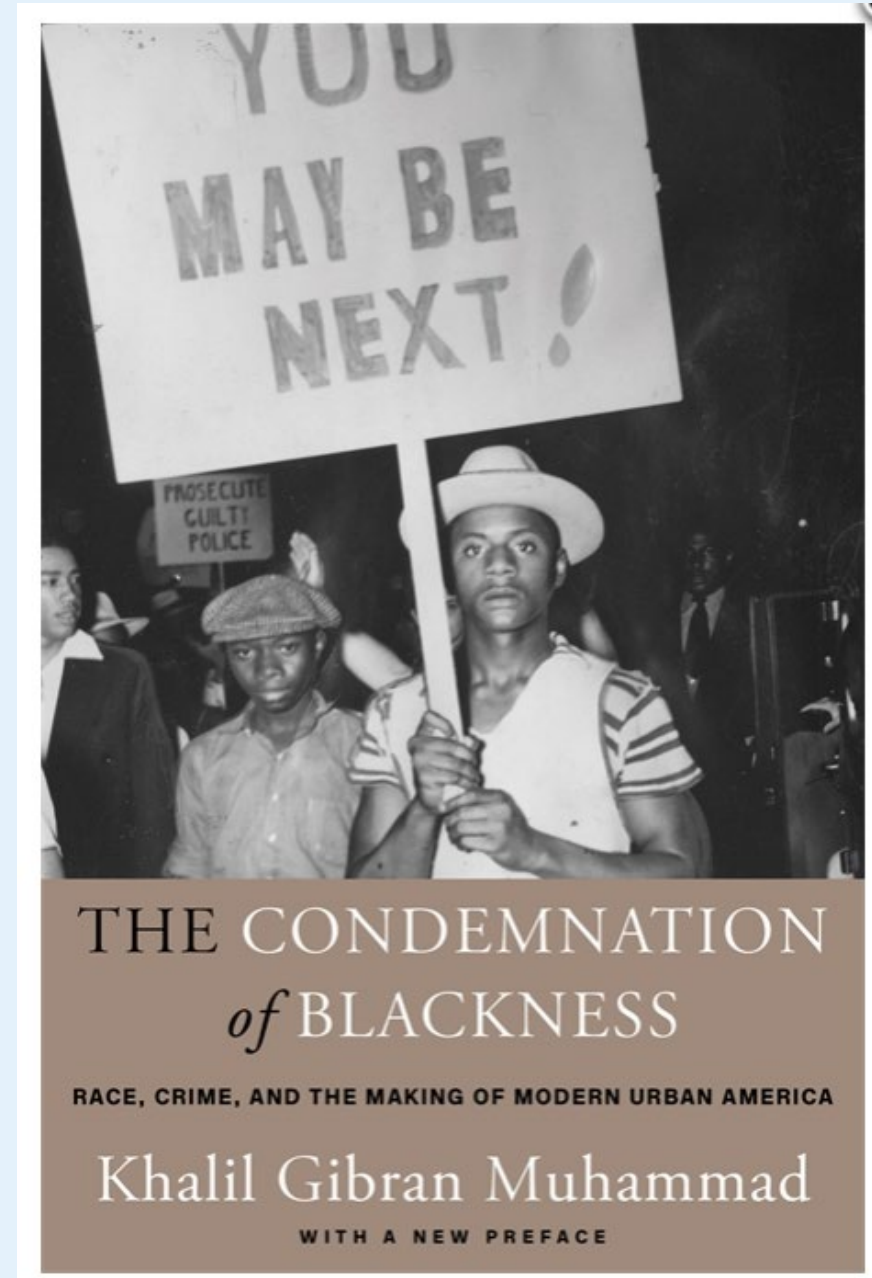


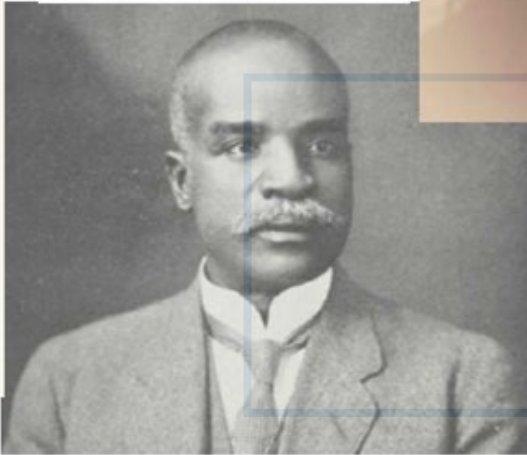
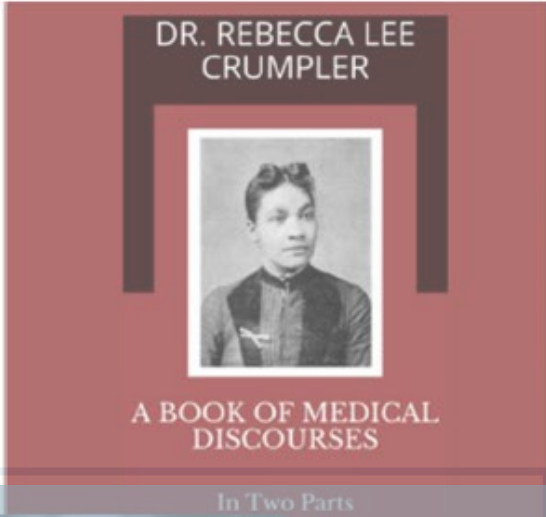
# SASS 500: SOCIAL DATA ANALYSIS AND RACISM, THEN AND NOW



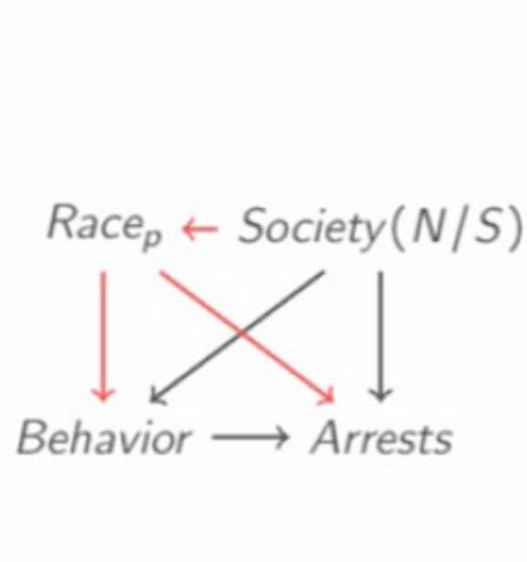
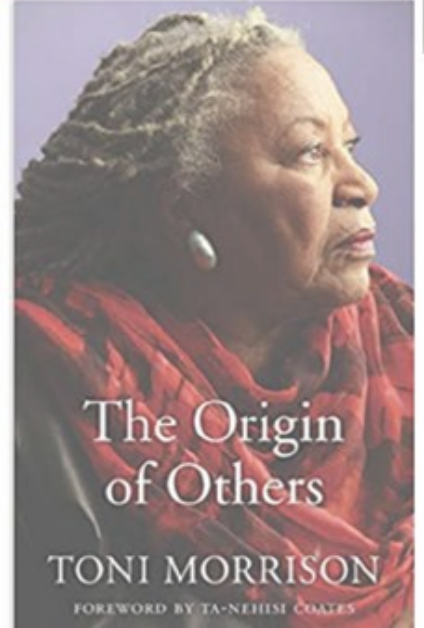
“The invisible layers of racial ideology packed into the statistics, sociological theories, and the everyday stories we continue to tell about crime in modern urban America are a legacy of the past. The choice about which narratives we attach to the data in the future, however, is ours to make.” - K. G. Muhammad, TCOB.

“Data alone are hardly a science, no matter how ‘big’ they get and how skillfully they are manipulated.” - J. Pearl, 2019 interview.





# Learning from Our Teachers



**FACING  
RACISM IN  
ACADEMIC  
RESEARCH**

**06/2020**

**AMERICAN ECONOMIC  
ASSOCIATION**

“We recognize that we have only begun to understand racism and its impact on our profession and our discipline.”

**06/2021**

**ASSOCIATION OF SOCIAL  
WORKERS**

NASW apologizes for racist practices in American social work. “... our profession and this association have not always lived up to our mission of pursuing social justice for all”

**10/2021**

**AMERICAN PSYCHOLOGY  
ASSOCIATION**

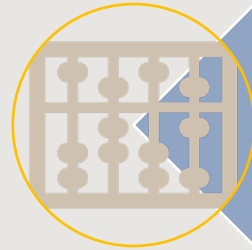
“Apology to People of Color for APA’s Role in Promoting, Perpetuating, and Failing to Challenge Racism, Racial Discrimination, and Human Hierarchy in U.S.”

**01/2023**

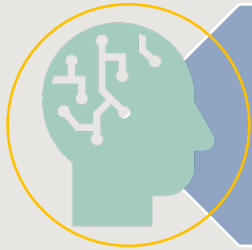
**AMERICAN SOCIOLOGICAL  
ASSOCIATION**

Statement on the Importance of Teaching and Learning About Race and Racism

# PIT-UN Grant to build resources for DSSI certificate

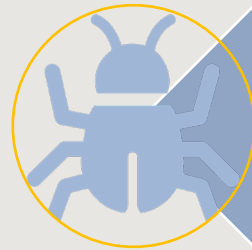


Using Administrative/Public Data



Guide to integrating community knowledge into data science projects.

Help students critically examine/challenge assumptions



Identifying/addressing bias in data analysis – C & P modeling.



Ethics and fairness considerations



# A DIVERSE TEAM

**12 Community Collaborators**

**Community partners**

**Alice Jackson, Founding Director of ThinkingCapp.**  
Claire Gauntner, Formerly at Enterprise Community Partners.  
Chris Knestrick, NorthEast Ohio Coalition for the Homeless  
Catherine Tkachyk, Chief Innovation & Performance Officer, Cuyahoga County  
Aysha Wilburn, Director of 2Gen Strategy – The Centers for Families and Children

**Student collaborators**

**Nicole Coury – MSASS and WSOM**  
**Kelly Orr, Economics**

**Faculty/Staff Partners:**

**Laura Bruckman, Engineering**  
**Shanina Knighton, Nursing**  
**Francisca G.-C. Richter, MSASS**  
**Emily Nelson, Swetland Center**  
Brian Gran, Sociology  
Claudia Coulton, Poverty Center  
Darcy Freedman & Rachael Sommer, Swetland Center

# FAIR2: A FRAMEWORK FOR ADDRESSING DISCRIMINATION BIAS IN SOCIAL DATA SCIENCE



## FRAME

metadata and data with **historical context and experiential knowledge** of those represented in the data



## ARTICULATE

model as a **causal graph to explicitly state assumptions** and hypotheses about the role of discrimination in the social phenomenon being addressed.



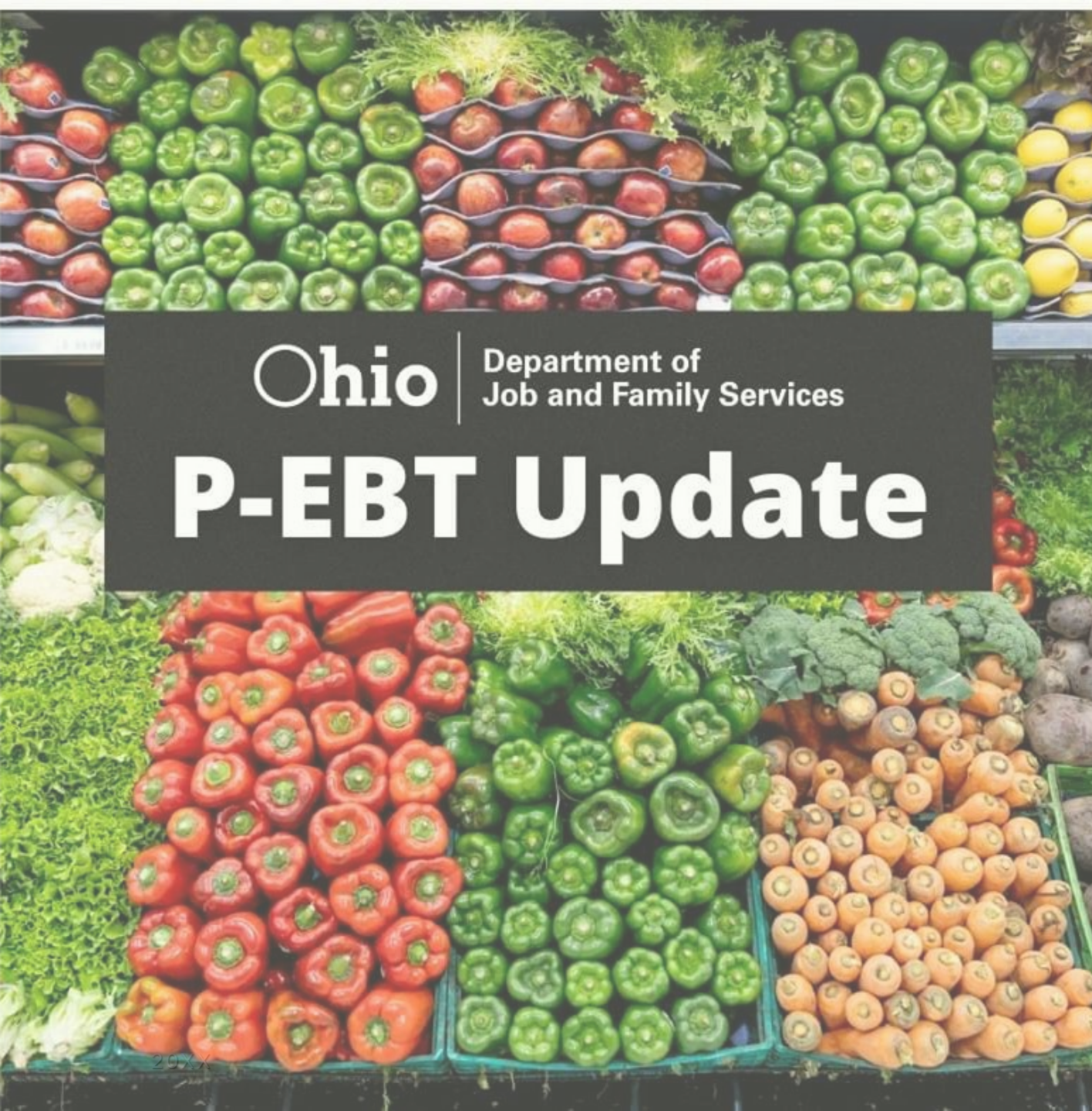
## IDENTIFY

**bias embedded in the data and variables** of interest to guide estimation and interpretation of results



## REPORT

findings and **seek feedback from members or agencies in the community** who have experiential knowledge of the social issue analyzed.



Ohio

Department of  
Job and Family Services

# P-EBT Update

## ANALYSIS OF RECERTIFICATION IN THE SUPPLEMENTAL NUTRITION ASSISTANCE PROGRAM (SNAP)

Using administrative data

### POPULATION

SNAP enrollees subject to  
recertification in locality L

### VARIABLE OF INTEREST

Failure to recertify: Who fails to  
recertify when eligible and why?

# FRAME : DATA CHAT ON SNAP RECERTIFICATION

## Data Intake Experience

Based on your experience:

Who entered the information you provided in the intake form or computer?

Were you offered an explanation for the need to ask these questions?

Do you think this data necessary is to provide the appropriate services?

If you have enrolled in other assistance programs, do you find that the data intake process is more or less demanding?

## Your Advice

Based on your experience:

What people or resources made the SNAP recertification process easier?

What were the most difficult parts of the SNAP recertification process?

In a perfect world what would the SNAP recertification process look like?

What could help reduce SNAP churn?



## EXPERIENTIAL KNOWLEDGE

### OBSOLETE COMMUNICATION

Hours long wait times on the phone;  
letters requiring additional information that come close to the interview day

### VARIATION IN ADMINISTRATION QUALITY

Perception that higher income localities provide better service than those with higher need.

### STRESSFUL VERIFICATION PROCESS

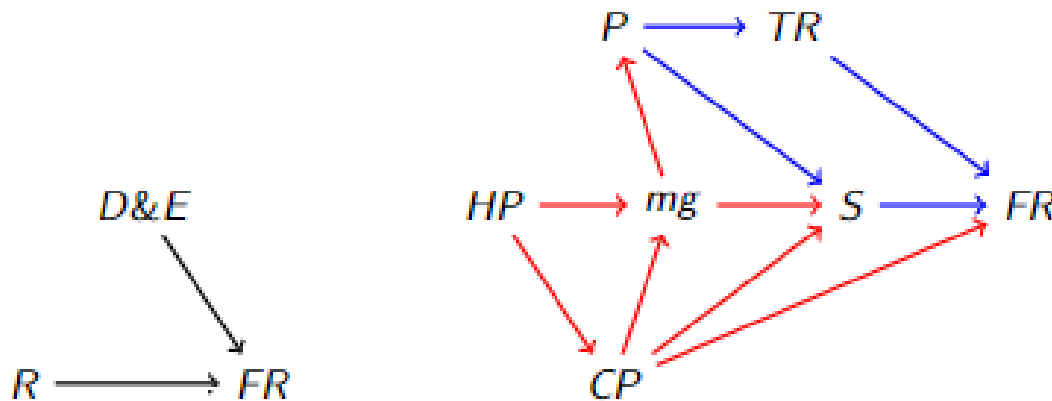
Seems confrontational.  
You need to “suck up your pride”

*It needs to come from a sense of mutual trust. The natural state of humans is cooperation. If you're being combative, everyone reacts viscerally. They don't have enough people to intake, so you're discouraging people from applying.*

*You know when you set a goal, and you try to make it, and you fail?  
That's how I feel. It's a defeated feeling. There's nothing I can do,  
nobody who can help you.*

# ARTICULATE

Directed acyclic graphs modeling failure to recertify for food assistance (FR).



Left: Naive model that aims to explain differences in FR with demographic and economic characteristics (D&E) and race (R).

Right: Model includes historical and experiential knowledge.

*HP* = historical processes that shaped discrimination in US society (like Jim Crow or redlining);

*CP* = contemporary policies that allocate scarce resources to the social safety net (weak transportation system, understaffed public assistance administration);

*mg* = classification into marginalized grouping;

*P* = poverty;

*TR* = time and resource scarcity;

*S* = stigma.

## IDENTIFY

### Potential biases in data

#### CHURN

flag FR in administrative data  
IF an individual drops from  
SNAP and **re-enters within 1-3**  
months.

#### RACE AS A DEMOGRAPHIC VAR?

To assess inequities

#### MISSING DATA PATTERNS

Seemingly non-random.  
Places with higher need seemed  
less resourced – hours long wait  
on phone.

#### WHAT ARE WE IGNORING?

underlying mechanisms that  
lead to inequities have strong  
implications for setting up  
research designs and identifying  
policy solutions.

#### WHO COULD WE BE MISSING?

Folks with more than 3 months  
lapses

Or in places with data entry  
issues due to low-resources.

**EMPOWER  
GRADUATES TO  
PLAY A CRITICAL  
ROLE IN  
HARNESSING  
TECHNOLOGY FOR  
SOCIAL IMPACT**

## **USE OF FORCE BY POLICE**

## **INCARCERATION & BAIL**

Chanelle J Howe, Zinzi D Bailey, Julia R Raifman, John W Jackson, [Recommendations for Using Causal Diagrams to Study Racial Health Disparities](#), American Journal of Epidemiology, Volume 191, Issue 12, December 2022, Pages 1981–1989, <https://doi.org/10.1093/aje/kwac140>

## **CHILD WELFARE**

## **HOMELESS SERVICES**

Whitney R Robinson, Audrey Renson, Ashley I Naimi, [Teaching yourself about structural racism will improve your machine learning](#), Biostatistics, Volume 21, Issue 2, April 2020, Pages 339–344, <https://doi.org/10.1093/biostatistics/kxz040>

## **HEALTH SERVICES**

## **HOUSING & LEAD EFFECTS RESEARCH**



# RACISM VS RACE IN DATA

System vs Individual traits

Wynton Marsalis, Musician, Educator,  
Leading Advocate of American Culture



"Well, race is a...

...Race is like, for this country, it's like, the thing in the story, in the mythology, that you have to do for the kingdom to be well.

And it is always something you don't want to do. And it is always that thing that is so much about you confronting yourself, that is tailor-made for you to fail dealing with it.

And the question of your heroism, and of your courage, and of your success at dealing with this trial is: can you confront it with honesty, and do you confront it, and do you have the energy to sustain an attack on it? "

**THANK YOU**

For more information contact

Mandel:

Francisca García-Cobián Richter

[fxr58@case.edu](mailto:fxr58@case.edu)

Richard Sigg, [ras300@case.edu](mailto:ras300@case.edu)

Engineering:

Laura Bruckman, [lsh41@case.edu](mailto:lsh41@case.edu)

This work was generously funded by  
grant #015865 from the Public Interest  
Technology University Network – New  
America Foundation

