



# Effective Use of Demographic Information

Mona Khalil

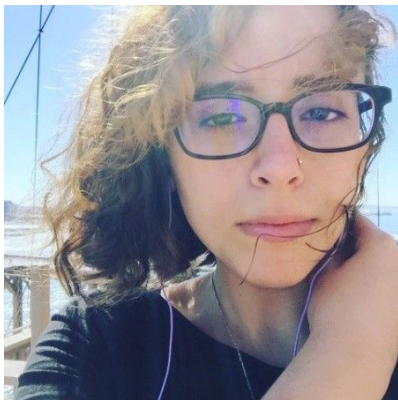
Devin Johnson

# Overview

- What is demographic information?
- Issues with research around demographic information
- Determining if demographic information is necessary
  - Weighing risks & benefits
  - Communicating analytical needs to stakeholders
  - Collecting better information when you need it
- Activity -- design a study without touching demographic information



# About Us



Mona Khalil

- Data Scientist at Greenhouse Software
- Master's in Developmental Psych
- Worked in education, government, academia, and ed-tech



Devin Johnson

- PhD Candidate at McMaster University
- Master's in Experimental Psych
- Worked in education, and market research as a Data Analyst

# What is demographic information?

Demographics are *characteristics* of a population.

- Typically measured in categories (*i.e.*, *race/ethnicity*) or on a scale (*age*)
- Characteristics usually agreed upon by the broader population and/or governing bodies
- Used to draw conclusions about behaviors, processes, and outcomes associated with subsets of the population



# What is demographic information?

- Types of categories used are nearly ubiquitous
  - I.e., race, gender, age
- Chosen categories for each characteristic differ across the world
  - Race/ethnicity is measured *relative* to the specific country or region
- Measured across a wide variety of activities
  - Schools, workplaces, surveys, research studies

9. What is Person 1's race? Mark ☒ one or more boxes.

- ☐ White  
☐ Black, African Am., or Negro  
☐ American Indian or Alaska Native — *Print name of enrolled or principal tribe.* ↗

- ☐ Asian Indian    ☐ Japanese    ☐ Native Hawaiian  
☐ Chinese    ☐ Korean    ☐ Guamanian or Chamorro  
☐ Filipino    ☐ Vietnamese    ☐ Samoan  
☐ Other Asian — *Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on.* ↗    ☐ Other Pacific Islander — *Print race, for example, Fijian, Tongan, and so on.* ↗

- ☐ Some other race — *Print race.* ↗

19 Is this person:

Mark ☒ more than one or specify, if applicable.

*This information is collected to support programs that promote equal opportunity for everyone to share in the social, cultural and economic life of Canada.*

- ☐ White  
☐ Chinese  
☐ South Asian (e.g., East Indian, Pakistani, Sri Lankan, etc.)  
☐ Black  
☐ Filipino  
☐ Latin American  
☐ Southeast Asian (e.g., Vietnamese, Cambodian, Malaysian, Laotian, etc.)  
☐ Arab  
☐ West Asian (e.g., Iranian, Afghan, etc.)  
☐ Korean  
☐ Japanese  
Other — Specify



# Problems with Demographic Information

- Demographic information is often measured in oversimplified ways
  - “Gender” measuring binary sex
  - Race/ethnicity measuring observable characteristics/ differences from a dominant group

**1. Which race/ethnicity best describes you?  
(Please choose only one.)**

- ☐ American Indian or Alaskan Native
- ☐ Asian / Pacific Islander
- ☐ Black or African American
- ☐ Hispanic
- ☐ White / Caucasian
- ☐ Multiple ethnicity / Other (please specify)

**What is your gender? \***

- ☐ Male
- ☐ Female

**What is your age category? \***

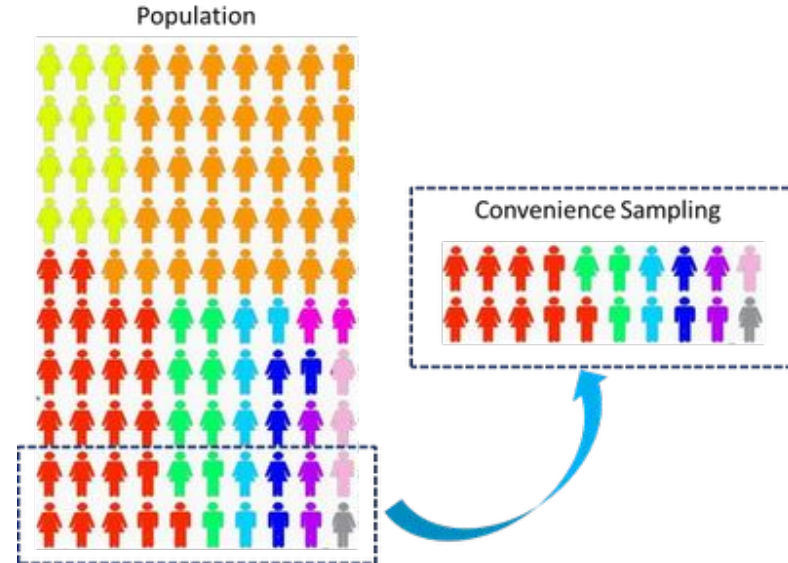
- ☐ Under 18
- ☐ 18-24
- ☐ 25-34
- ☐ 35-44
- ☐ 45-54
- ☐ 55-64
- ☐ 65 or Above

**3. What is your sexual orientation?**

- ☐ Asexual
- ☐ Bisexual
- ☐ Gay
- ☐ Heterosexual or straight
- ☐ Lesbian
- ☐ Pansexual
- ☐ Queer
- ☐ None of the above, please specify

# Problems with Demographic Information

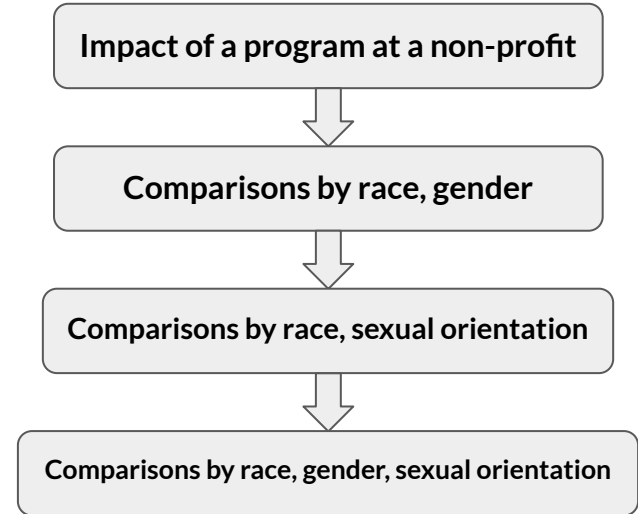
- Demographic characteristics are a “measurement of convenience”
  - Importance is inferred, with little justification
  - Used as a proxy for *representative* samples
  - Used as a “fallback” in research when other hypothesis cannot be supported





# Problems with Demographic Information

- Analyzing demographic information can easily become invasive, or draw attention away from the broader context of a research project
  - Requires disclosure of personal information
  - Decreased anonymity of vulnerable populations
  - Requires significant time investment for research findings that are often *less actionable*







# Do you *need* demographic information?

1. If you are unable to collect any kind of demographic information, what questions would you ask, and hypotheses would you have?

- What questions can you come up with?
- What other types of data would you collect?

# Do you *need* demographic information?

2. Does the *benefit* associated with collecting demographic information outweigh the *risks*?

- Potential risks:
  - Alienating research participants
  - Collecting incorrect information
  - Drawing inaccurate conclusions that negatively impact vulnerable populations

Risks and benefits in research:

<https://research.ucdavis.edu/policiescompliance/irb-admin/researchers/project-guidance/risks-and-benefits/>





## Do you *need* demographic information?

3. If you find statistically significant differences between 1 or more demographic group, will you be able to take concrete actions to address the needs of those groups?
  - I.e., if non-white students perform worse in your class, what can and will you do about it?
  - Will the actions you recommend be taken?



# Collecting Better Demographic Data

If you've answered yes to the previous questions, we recommend the following criteria for producing actionable insights:

- Expand question choices to include groups more representative of people's lived experiences
  - non-binary and gender-nonconforming options
  - race/ethnicity categories representative of the non-white and immigrant groups in a country/society
  - (see [Respectful Collection of Demographic Data](#))

Gender: ?

☐ female

☐ male

☐ nonbinary

☐ prefer not to say

☐

Would you describe yourself as transgender? ?

☐ yes

☐ no

☐ prefer not to say



# Collecting Better Demographic Data

- Think closely on the intersections of privilege and power when comparing groups of individuals
  - Preserves the anonymity of participants from vulnerable groups
  - Provides insight into whether there are broad issues of power dynamic and privilege being captured in your work
  - Example:
    - white vs. non-white respondents
    - male vs. non-male respondents
    - cisgender vs. transgender respondents



# Collecting Better Demographic Data

- Further divide non-dominant groups only when necessary
  - You suspect issues of bias or inequity impacting a specific group (*and can recommend actions to remedy it*)
  - Your research involves **specific** minority groups
    - Healthcare needs of elderly New Yorkers of Latin American descent
    - School adjustment among Syrian refugee children
  - Further dividing into subgroups does not risk the identification of specific individuals
    - (i.e., a study with only 2 transgender respondents)

---

# Demographic Data & Internal Stakeholders

- Internal stakeholders (management, external teams, etc) in many cases may request analytics projects that heavily rely on demographic data
- Communicating when and where this data is necessary and when it is not can be a challenge



# Improving reporting with little external “buy-in”

- Start within your own team
  - You'll need support and it's exhausting doing this individually
- Bring all aspects to the table
  - Ethics, quality, liability
- Progress can be frustratingly slow





# Data without (or justified) demographic info

- In small groups you'll be presented with a snapshot of a real data set
- Work together to devise research questions that don't rely on the demographic data available
- If you feel you must use demographic information, provide a rationale for its use



shutterstock.com • 229793107



# Resources

- [Respectful Collection of Demographic Data](#)
- [Risks and Benefits in Research](#)



# Thank you!

Mona: [twitter.com/mona\\_kay](https://twitter.com/mona_kay)

Devin: [twitter.com/Devin\\_li](https://twitter.com/Devin_li)

We have a podcast!

[anchor.fm/badmethods](https://anchor.fm/badmethods)