Structured Purpose: Implementing Python in purposive sample selection for evaluation interviews

Valerie A. Lucas[1], Tara Carr[1], Gaurav Dave[1]
[1]Abacus Evaluation, Center for Health Equity Research at University of North Carolina at Chapel Hill

UNC SCHOOL OF MEDICINE Abacus Evaluation

RADXUP

Introduction

- ***Rapid Diagnostics for Underserved Populations (RADx-UP)*** is a grant funded by the National Institutes of Health (NIH) to improve COVID-19 testing uptake in community sites across the Unites States, territories, and Tribal Nations.
- A key evaluation objective was to ***conduct qualitative, semi-structured interviews*** to understand the context of project implementation at a subset of sites.
- We chose to use Python instead of traditional purposive sampling for a ***more systematic and structured approach*** to better account for the ***characteristics of the available sites*** and ***the aims of the research*** in our sampling.

Methodology

- We ***designed and implemented a sample selector in Python,*** a programming language, using the sample size and project characteristics as constraints. We aimed for a sample size of nine out of 69 sites.

Applying nine initial conditions to construct sets of nine sites and ***create computationally manageable site sets*** (ex. requiring the 1st site to be an LGBTQ+ community and the 2nd site to be a rural community).

All Possible Sets of Nine Sites
$N = \binom{69}{9} = 56{,}672{,}074{,}888$

Sets of Nine Sites with Initial Conditions
n = 32,366,880

Sequentially applying more conditions (ex. requiring ≥1 site to be a Pacific Islander community, then ≥2 sites to use an experimental methodology).

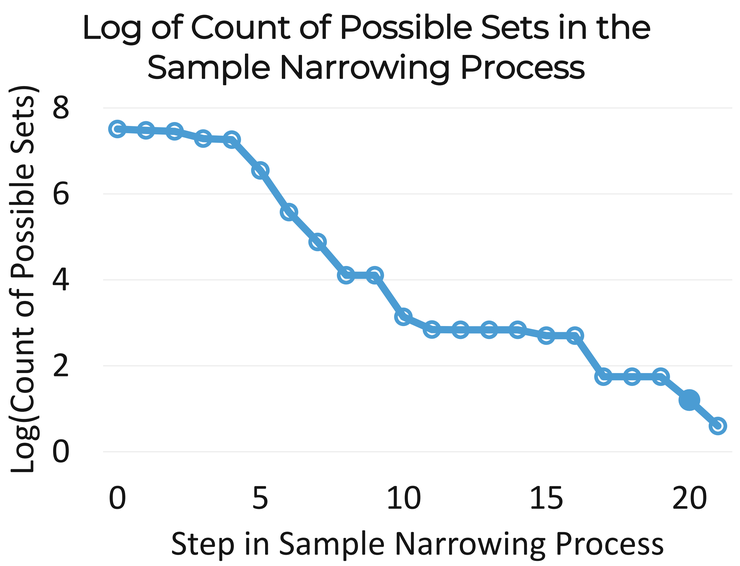
Final Set of Nine Sites
n = 1

Researchers and evaluators can make the subjective process of purposive sampling individuals or programs more systematic with a simple Python program.

View and download the Python file from the Carolina Digital Repository

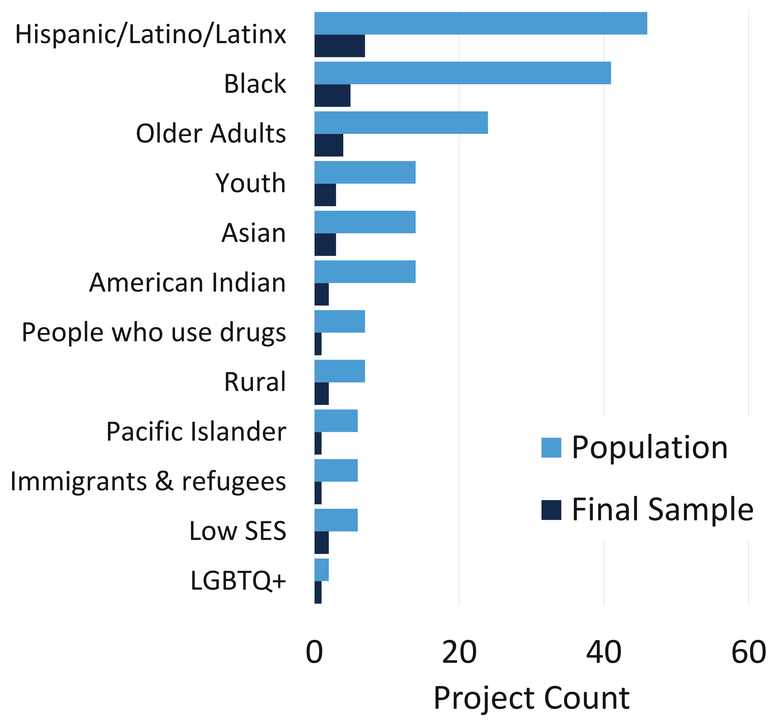
Methodology (cont.)

- Each project has ***characteristics like study design, target population, and region*** of the United States from which we created selection priorities and further narrowed the sample set.
- The program takes about ***30 minutes to run on an ordinary laptop computer***.



Results

- The process yielded a ***diverse sample*** that met 23 of our narrowing criteria.
- Selection design ***maximizes "intersectional" projects***, those that work with more than one possibly overlapping target population.



Conclusions

- Our Python code provides a more ***structured method of purposive sampling for interviews***, provided there are known characteristics of a population of evaluation subjects.
- Tracking & evaluation teams can ***rank conditions*** for a sample of evaluation projects from most important to least important.

Contact Information
Valerie Lucas, valerie_lucas@med.unc.edu

Acknowledgments
RADx-UP community & academic partners, RADx-UP CDCC, NIH 1U24MD016258-01