PhenX Toolkit

Data Collection Worksheet

Please Note: The Data Collection Worksheet (DCW) is a tool to aid integration of a PhenX protocol into a study. The PhenX DCW is not designed to be a data collection instrument. Investigators will need to decide the best way to collect data for the PhenX protocol in their study. Variables captured in the DCW, along with variable names and unique PhenX variable identifiers, are included in the PhenX Data Dictionary (DD) files.

The Index of Concentration at the Extremes is based on U.S. Census Bureau data. This protocol describes how to make calculations using 5-year American Community Survey (ACS) estimates.

The ACS data used in this protocol can be accessed by using Excel to read the Summary Files at the U.S. Census Bureau's data.census.gov website (https://data.census.gov) or by using SAS programs to read the files. Users can find additional information on these tools at the following locations:

Using Excel to Access Summary Files: https://www2.census.gov/programs-surveys/acs/summary_file/2020/documentation/tech_docs/ACS_SF_Excel_Import_Tool.pdf

Using SAS to Access Summary Files: https://www.census.gov/content/dam/Census/library/publications/2019/acs/acs_summary-file_handbook_2019_ch04.pdf

The technical documentation for the American Community Survey (ACS) summary files is available online at https://www.census.gov/programs-surveys/acs/data/summary-file.html. Select "What ACS Summary File Data Users Need to Know" for an overview of the ACS Summary file and how it can be used to access data. Users not familiar with Census data should consult the technical materials.

To compute ICE, the following formula is used:

$ICE_i = (A_i - P_i)/T_i$

Where

A_i is equal to the number of affluent persons in neighborhood i (e.g., in the 80th income percentile)

P_i is equal to the number of poor persons in neighborhood i (e.g., in the 20th income percentile)

T_i is equal to the total population with known income level in neighborhood i

ICE is a single metric that simultaneously quantifies concentrated extremes of both privilege and deprivation, whereby a value of 1 connotes that all residents are in the privileged group and a value of –1 denotes that all residents are in the most deprived group.

ICE can meaningfully be computed for both smaller and larger geographic units (e.g., block group, census tract (CT), community district, city, county).

It is suggested to use 5-year annual average values for each variable because there are no public-use single-year CT-level ACS estimates available, and such estimates can both be imprecise and vary widely across years (because of both changing sampling frames and sample sizes).

Protocol source: https://www.phenxtoolkit.org/protocols/view/290801