

Project Proposal: Leadership Age and Nonprofit Service Priorities

1. Research Question and Background

This project will analyze the association between leadership age and nonprofit service priorities through the dataset **"National Survey of Nonprofit Trends and Impacts Spring 2021"** from Urban Data Catalog. This dataset provides detailed information on organizational operations, leadership demographics, and programmatic services from approximately 2,300 respondents.

The central research question for this analysis is: **Controlling for covariates, is there an association between the age of a nonprofit's leadership (Chief Executive and Board Chair) and the organization's programmatic focus on serving LGBTQ+ community?**

This proposal hypothesizes that nonprofits with younger leaders (defined as under 45 years old) are more likely to formally designate the LGBTQ+ community as a primary or secondary service group.

2. Data Preparation and Methodology

The analysis requires a systematic data pipeline to manage the dataset's complexity. The following steps will be taken to prepare the variables for modeling:

- **Dependent Variable (DV):** The outcome variable will be a binary indicator of LGBTQ+ programmatic focus. This will be derived from the variable `PROGDEM_19` ("Do some or all of your programs focus on serving any of the specific groups listed below? - Individuals identifying as LGBTQ"). Responses of "This is a primary group we serve" (1) and "This is a secondary group we serve" (2) will be combined into a single "Focus" category (coded as 1). This will be compared against "We do not specifically focus programs on this group" (coded as 0).
- **Independent Variable (IV):** The primary predictor will be **"Younger Leadership."** This will be a binary variable created from the categorical age variables for the Chief Executive (`CEOBCDEM_3_1`) and Board Chair (`CEOBCDEM_3_2`). An organization will be coded as 1 ("Younger Leadership") if its CEO or Board Chair is in the '18-24', '25-34', or '35-44' age brackets.
- **Control Variables (Covariates):** A key challenge for this project is selecting appropriate covariates to control. This selection process is complex because the dataset contains nearly 1,000 variables, which must be narrowed down to isolate the true effect of leadership age.

3. Analysis Plan

The analysis will proceed in two stages:

1. **Exploratory Data Analysis (EDA):** We will first generate weighted descriptive statistics for all variables to understand the sample's composition. We will then create weighted bar charts and cross-tabulations to visualize the raw relationship between our "Younger Leadership" variable and the "LGBTQ+ Focus" outcome.
2. **Statistical Modeling:** I will test several models appropriate for our binary dependent variable. The final selected model will simultaneously incorporate the independent variable and all control variables.