

Organizational Capacity Analysis Methodology

*Inland Empire Vital Conditions Philanthropic Network Study
By Mapping Black California*

DATA SOURCE & SCOPE

Dataset Characteristics

- **Source Information:** IRS Form 990 data for organizations operating in Riverside and San Bernardino Counties
- **Observation Unit:** Individual 501(c) tax-exempt organizations
- **Total Organizations:** 4,344 active organizations
- **Time Period:** Most recent two reporting years (2023-2024 primary; 2022 data for older filers)
- **Geographic Scope:** Riverside and San Bernardino Counties, California
- **Total Revenue:** \$16.1 billion across most recent two years
- **Total Assets:** \$26.6 billion
- **Total Liabilities:** \$11.1 billion

DATA STRUCTURE & VARIABLES

Organization-Level Variables

Identification & Classification:

- Organization name
- EIN (Employer Identification Number)
- IRS NTEE code (primary and major category)
- NAICS code (2017 classification)
- 501(c) subsection classification
- Year formed (range: 1868-2025)

Geographic Variables:

- City
- County (Riverside or San Bernardino)
- State (California)

Financial Variables

- Total assets (USD)
- Total revenues (USD)
- Total expenses (USD)
- Total liabilities (USD)
- Program-related expenses (percent of total expenses)

- General management expenses (percent of total expenses)
- Personnel compensation (percent of total expenses)

Operational Variables

- Has full-time employees (binary)
- Has endowment support (binary)
- Grant activity (grantmaker, grant recipient, both, or neither)

ANALYTICAL FRAMEWORK

Analysis 1: Regional Context & Ecosystem Demographics

Objective

Characterize the size, composition, and temporal development of the regional nonprofit sector.

Methods

Sectoral Distribution Analysis:

- Grouped organizations by NAICS code (2017 classification)
- Grouped organizations by NTEE major category
- Calculated counts and percentages for each sector

Temporal Formation Analysis:

- Extracted year formed (range: 1868-2025)
- Created decade categories
- Analyzed formation patterns by era
- Identified peak formation periods

Analysis 2: Financial Structure & Institutional Concentration

Objective

Examine concentration of financial resources, identify mega-institutions, and assess county-level differences in organizational scale and financial health.

Methods

Asset Concentration Analysis:

- Ranked organizations by total assets
- Identified top 10 asset holders by county
- Calculated cumulative percentage of total assets
- Analyzed institutional composition (hospitals, universities, foundations)

County Comparative Analysis:

- Aggregated financial metrics by county
- Calculated per-organization averages
- Computed debt-to-asset ratios
- Calculated operating margins (net income / revenue)
- Compared institutional structures across counties

City-Level Geographic Concentration:

- Aggregated revenue by city within each county

- Identified top 10 cities by total revenue
- Calculated per-organization averages by city
- Analyzed operating margins by municipality

Analysis 3: Operational Efficiency & Financial Health

Objective

Examine expense allocation patterns across program services, management, and personnel to assess operational efficiency and identify concerning patterns.

Methods

Sample Construction:

- Filtered to organizations reporting $\geq \$1,000$ in expenses
- Final sample: 3,889 organizations

Expense Structure Analysis:

- Calculated program expense ratio (program / total expenses)
- Calculated management expense ratio (management / total expenses)
- Calculated personnel ratio (compensation / total expenses)
- Classified organizations by efficiency patterns

100% Management Expense Subanalysis:

- Identified 57 organizations with 0% program / 100% management
- Segmented by personnel cost ratio:
 - - Zero personnel (0%): 44 organizations
 - - Low-moderate personnel (1-49%): 8 organizations
 - - High personnel (50-96%): 5 organizations
- Analyzed organizational characteristics by segment
- Assessed legitimacy and intervention needs

DATA CLEANING & PREPARATION

Geographic Standardization

- Standardized city names (removed variations, corrected misspellings)
- Verified county assignments

NTEE Code Standardization

Extract major category (first letter of NTEE code):

```
df['NTEE.Major'] = df['NTEE code, primary'].str[0]
```

Map codes to sector names covering 24 total categories by IRS Publication 4838 - National Taxonomy of Exempt Entities (NTEE) Codes

<https://www.irs.gov/pub/irs-tege/p4838.pdf>

Financial Data Processing

- Calculated derived metrics (debt ratios, margins, per-org averages)
- Handled negative values (losses, deficit operations) appropriately

Temporal Data Handling

- Standardized year formed (range: 1868-2025)
- Created decade categories for formation analysis
- Calculated organizational age (2024 - year formed)

DATA QUALITY CONSIDERATIONS

Missing Data

Financial Metrics:

Revenue data: >95% completeness for organizations reporting any activity.

Treatment: Used available-case analysis; reported N for each metric.

Classification Data:

NTEE Unknown ('Z'): Present in dataset. Treatment: Manually reclassified major recipients; retained 'Z' with flag

Assumptions & Limitations

Assumption 1: Most recent two-year data represents current organizational state

Assumption 2: Foundation NTEE represents organizational mission (though grantmaking may be broader)

Assumption 3: Organizations report financial data accurately and consistently

Limitation 1: Analysis limited to 501(c) organizations with IRS filings

Excludes: Religious congregations below filing threshold, informal community groups, grassroots organizations without tax-exempt status, government agencies, for-profit social enterprises

Limitation 2: Data represents only foundation-to-recipient grants; excludes: individual donor contributions, government grants, corporate sponsorships, fee-for-service revenue.

Limitation 3: Network limited to intra-Inland Empire flows; excludes: IE foundations giving outside region, external foundations giving into region. Preserved for future analysis.

Limitation 4: Point-in-time snapshot (2023-2025).

VALIDATION & RELIABILITY

Total Dataset: (N = 4,344):

Fields with 100% Completeness (N = 4,344)

Organizational-Level:

- Organization name: N = 4,344
- EIN: N = 4,344
- County: N = 4,344

- City: N = 4,344
- NTEE code: N = 4,344
- NAICS code: N = 4,344

High Completeness (>95%):

Organizational-Level:

- Year formed: >97% complete
- Total revenue: >95% for active organizations
- Total assets: >95% for active organizations

Expense Analysis Sample (N = 3,889):

- Total expenses: N = 3,889 (filtered to $\geq \$1,000$)
- Program expense %: Variable completeness
- Personnel compensation %: Variable completeness