Methodology

The primary data source for this report is a dataset provided by the Oklahoma Office of Juvenile Affairs (OJA), covering the years 2018 to 2023. This dataset offers detailed information on referrals, adjudicated offenses, and the final status of referrals within Oklahoma's youth justice system. We extend our sincere gratitude to our partners at OJA for their data sharing and the considerable time invested in explaining and clarifying our inquiries throughout the research process.

The OJA dataset is comprised of three main components:

- Referrals Data
- Adjudications Data
- Final Case Status Data

REFERRALS DATA

The referrals data, also referred to as the 'arrest_offense' table, contains information about offenses grouped as referrals to OJA. Key variables include:

- Unique identifiers for referrals and individuals
- Specific offense information
- Month, year, and county of referral
- Demographic information (age, gender, race)
- Referring agency

In a small fraction of cases (less than 0.3 percent), additional offenses are added to a referral at a later date. We used regex to group the referring agencies into broader categories.

ADJUDICATIONS DATA

The adjudications table provides information on court outcomes for adjudicated offenses. Significant variables include:

- Individual and referral identifiers
- Adjudicated offense details
- Court decisions and dates

An important consideration in this dataset is that for approximately 14 percent of cases, the count number (usually a unique identifier for each offense within a referral) is not unique for every referral ID. This required careful handling in our analysis.

FINAL CASE STATUS DATA

The case status table offers information on the final status for particular referrals. Key variables include:

- Time-related information (specifically, the year of the case status)
- Individual and referral identifiers
- Final case status

From the legal status field, we derived a 'category_legal' field using regex to group similar statuses for more effective analysis.

Our analytical approach involved several key steps:

We joined the referral and adjudication data using referral ID and count number. This integrated dataset formed the foundation for our analysis on adjudication practices and racial disparities within the system.

To examine demographic trends in the final assigned case status, we merged the referrals and final case status data using individual identifiers.

Throughout the report we rely on relative rate index (RRI) calculations in order to compare the involvement of historically marginalized groups to that of white youth at each stage of the youth justice system.[195] We calculate RRIs for each offense category and at various stages of the justice system (e.g., referral, custody, probation). This allows us to pinpoint where disparities occur within specific stages and how they may change as youth progress through the system.

Calculate the rate for each racial/ethnic group:

Rate = (Number of youth experiencing the event) / (Number of youth in the population)

Calculate the relative rate index:

RRI = (Rate for racial/ethnic group of interest) / (Rate for white youth)

The resulting RRI value is interpreted as follows[196]:

- RRI > 1: The racial/ethnic group of interest has a higher rate compared to white youth
- RRI = 1: The rates are equal
- RRI < 1: The racial/ethnic group of interest has a lower rate compared to white youth

For example, to calculate the relative rate of referrals for American Indian youth in Tulsa County:

Referral rate for American Indian youth in Tulsa = Number of youth American Indian youth referred/ Tulsa population of American Indian youth (individuals 18 and under)

Relative Rate Index = (Referral rate for American Indian youth in Tulsa) / (Referral rate for white youth in Tulsa)

For youth population data, we used the National Vital Statistics bridged-race estimates, which are based on U.S. Census Bureau population data.[197] These estimates provide a consistent way to categorize race and ethnicity across different data sources and over time.

These estimates have important limitations:

1. They do not provide information about what might be causing the observed disparities.
2. They have no predictive power regarding future trends or outcomes.
3. RRI calculations do not constitute a formal statistical analysis.

This report uses RRIs in conjunction with existing research findings and other contextual information to provide a detailed view of the disparities within the youth justice system while acknowledging the limitations of this metric.

Additional Considerations

While this dataset provides necessary insights into Oklahoma's youth justice system, there are notable limitations to our current analysis that we hope to address in future research:

- Some terminology used in court decisions and final case statuses is pending further clarification from OJA to ensure accurate representation in our findings.
- Establish more direct connections between referral offenses and their specific case outcomes.
- Examine the intersection of gender and race disparities more closely, with the aim of obtaining data that reflects gender identities beyond the binary categories presented here.

We remain committed to ongoing collaboration with OJA to refine our understanding of the data and to ensure that our analysis provides the most accurate and valuable insights possible for improving Oklahoma's youth justice system.

All of the analysis was done in R. For additional information on data cleaning, processing, and analysis please feel free to contact the Oklahoma Policy Institute's research team at research@okpolicy.org.