In recent years, state courts have become increasingly adept at using data to inform organizational decision making and system improvement. As courts have looked for ways to address racial disparities and promote equal justice, it has become clear that individual-level race and ethnicity information is a vital tool for these efforts. However, collecting race and ethnicity data from court users and court personnel presents a variety of challenges and requires careful consideration and planning.

In 2020, the National Center for State Courts (NCSC) released *Collecting Race & Ethnicity Data,* its first set of guidance on these issues. This document covered, at a high level, the reasons for collecting race and ethnicity data, a discussion of means to collect these data, and barriers to that collection.

In the four years since the publication of *Collecting Race & Ethnicity Data,* NCSC has been supporting state courts on their efforts to collect and use quality race and ethnicity data. Through the *Blueprint for Racial Justice Initiative*, NCSC developed the *Racial Justice Organizational Assessment Tool for Courts* and hosted a series of learning labs to help courts build their administrative capacity on race and ethnicity data.

This publication integrates the knowledge and experience gained from these collective efforts over the last four years. It lays out the fundamental practical considerations for effective race and ethnicity data collection in the state courts. This guide will help courts achieve race and ethnicity data collection that is both useful and accurately representative of the populations courts serve.

Why Collect Individual-level Race and Ethnicity Data

In 2020, the Conference of Chief Justices (CCJ) and the Conference of State Court Administrators (COSCA) released *Resolution 1: In Support of Racial Equality and Justice for All.* This resolution specifically recognizes efforts "to collect, maintain and report court data regarding race and ethnicity that enables courts to identify and remedy racial disparities" and the agreement of CCJ and COSCA "to continue and to intensify efforts to combat racial prejudice within the justice system, both explicit and implicit, and to recommit to examine what systemic change is needed to make equality under the law an enduring reality for all, so that justice is not only fair to all but also is recognized by all to be fair."

Resolution 1 reflected a shift in the way courts think about both the role of data and the impact of court users' identities on cases. In past years, many courts followed a "color-blind" approach to race and ethnicity information. This approach was generally well-intended, grounded in the underlying principle that race should not have an impact on court activities, experiences, and outcomes. Many people believed that not recording people's racial and other identities would prevent these identities from influencing cases. Unfortunately, research now makes clear that a color-blind approach does not eliminate racial disparities; it merely allows the systemic impacts of race to remain invisible[i]. Across the country, people of color and other minoritized communities have disproportionate contact with the courts, have different court experiences, on average, and have different court outcomes.

Repeated surveys by the NCSC reflect that voters of color rate state courts' job performance lower than do white voters.[iii] These patterns exist in spite of the fact that courts value equal justice and strive to give everyone a fair court experience. Courts that lack individual-level data on race and ethnicity have a difficult time identifying the sources of these disparities and avenues for system improvement.

Courts can reap tangible benefits from collecting race and ethnicity data and using these data to inform decision making:

- By proactively collecting and analyzing race and ethnicity data, courts can identify areas of disparate treatment and disparate impact under the law. Where marginalized groups experience more negative outcomes, courts can make informed decisions about how to improve the delivery of court services in pursuit of their mission of equal justice.
- Race and ethnicity data can be used to identify and communicate to the public when equal justice is being achieved. Courts can track improvements in outcomes over time, document the positive impacts of their projects and initiatives, and identify areas where court users are being well-served across racial and ethnic groups. These data allow courts to celebrate and communicate their successes.
- By taking a proactive stance on race and ethnicity data and assessment, courts build trust with the public and show that they are working to identify and address disparities. This proactive approach also reduces the likelihood that external entities will conduct their own investigations, using methods such as web scraping and FOIA requests, and publicize these issues. These data enable the courts to tell their own stories, to respond factually to external reports, and to promote public trust and confidence through transparency.

[i] Washington State Supreme Court Gender and Justice Commission. (2021). How Gender and Race Affect Justice Now: Final Report. Retrieved from: https://cdm16501.contentdm.oclc.org/digital/collection/accessfair/id/928/rec/1

[ii]Kentucky Court of Justice Response. (2020). A Guide for Identifying, Addressing, and Reducing Racial and Ethnic Disparities. Retrieved from: https://kycourts.gov/Court-Programs/Family-and-Juvenile-Services/Pages/Reducing-Racial-and-Ethnic-Disparities.aspx

[iii] GBAO Strategies. (2023). 2023 State of the State Courts – National Survey Analysis. Retrieved from: https://www.ncsc.org/__data/assets/pdf_file/0039/96879/2023-SoSC-Analysis-2023.pdf

For a more detailed discussion of how courts benefit from collecting and using race and ethnicity data, see the Blueprint for Racial Justice's *Data-Driven Decision Making Guide for Courts*. This guide also provides specific examples of court projects and initiatives.

There is an understandable reluctance to collect race data in many courts, grounded in the fear that this information may reveal problems or undermine public trust. However, courts that undertake these initiatives generally find that they serve as evidence of the court's dedication to equity and transparency. When courts make it known that they are actively taking steps to address any disparities shown by the data, they *demonstrate* their commitment to equal justice. They become a leader and partner, rather than an obstacle, in efforts to achieve racial justice.
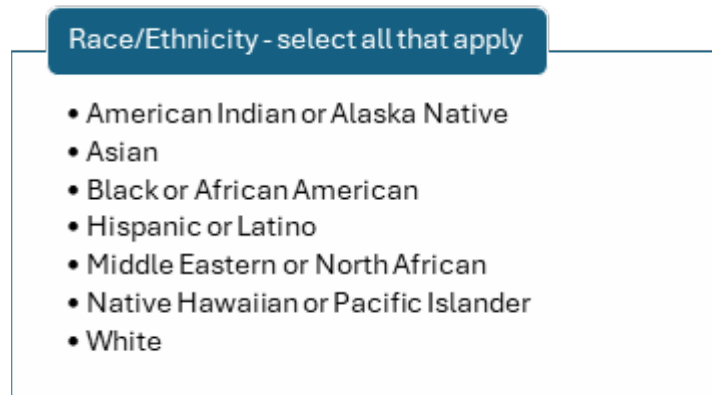
What Information to Collect

The categories that we use to define race and ethnicity in society are influenced by complex social, political, historical, and economic forces, and, as such, they evolve over time. This means that the decisions surrounding how to structure specific data elements in a court's case management system are complex. Courts are faced with the competing goals of using categories that accurately reflect the identities and experiences of the populations they serve, while also using data elements that are consistent across jurisdictions and over time. It also benefits courts if their data are compatible with the categories used by other major data sources, such as the U.S. Census and data exchange partners. These guidelines aim to balance all of these concerns.

Broad Racial and Ethnic Categories

Generally speaking, courts should consider two broad types of insights that they can draw from using their individual- and case-level data. The first type of insight looks at the court's overall caseload and makes comparisons over time or across jurisdictions. For examples of these types of data, see NCSC's Court Statistics Project. Comparing caseloads between courts requires that different jurisdictions can map their data onto a common set of broad categories. Similarly, examining how the population served by the court has changed over time requires that the court use compatible categories year after year.

The race and ethnicity data element in NCSC's *National Open Court Data Standards* is designed to serve this purpose. Courts that use NODS categories in their case management system (or categories that can be mapped onto the NODS categories) can use their data to get a snapshot of their caseload and court user population. and make use of external data sources like the U.S. Census. The NODS advisory committee will be considering the recent changes in how the Census defines race and ethnicity (see the *Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity*), while also maintaining categories that can be mapped onto those used by NODS in the past. This guide will be updated to reflect any changes to NODS.

Figure 1:

> ## Race/Ethnicity - select all that apply
>
> - American Indian or Alaska Native
> - Asian
> - Black or African American
> - Hispanic or Latino
> - Middle Eastern or North African
> - Native Hawaiian or Pacific Islander
> - White

There are several key features of this approach to measuring race and ethnicity:

- First, respondents are able to select multiple races. This allows people who identify as multiracial to accurately report their information.
- Second, race and ethnicity are combined in one item, allowing respondents to select all options that apply. This represents a shift from previous versions of the Census, which measured Hispanic ethnicity as a separate item. This change is based on the finding that some people who identify as Hispanic or Latine view this identity as an ethnicity and some view it as a race. In the past, not having the option to select Hispanic or Latine on the race item has led to large numbers of Census respondents identifying as "some other race."[iii] The goal of this new approach is to accommodate both those who identify racially as Hispanic or Latine and those who identify as Hispanic or Latine plus another racial category (e.g., Afrolatinas).
- Third, a category has been added for *Middle Eastern or North African (MENA)*. In previous versions of the Census, individuals who identify as MENA were included under *White*. This change reflects the recognition that MENA communities are minoritized and generally experience the barriers and stigmas faced by people of color.

[iv] United States Census Bureau. (2011). 2010 Census Shows America's Diversity. Retrieved from: https://www.census.gov/newsroom/releases/archives/2010_census/cb11-cn125.html

Pew Research Center. (2021). Majority of Latinos Say Skin Color Impacts Opportunities in America and Shapes Daily Life. Retrieved from: https://www.pewresearch.org/race-and-ethnicity/2021/11/04/measuring-the-racial-identity-of-latinos/

Locally Relevant Ethnic Communities

The second type of insight that courts can gain from their race and ethnicity data relates to the experiences and outcomes of court users who belong to different groups. In addition to collecting broad race and ethnicity categories that map onto the Census and other major data sources, courts should also consider what ethnic, tribal, immigrant, and cultural groups are represented in significant numbers in the court's jurisdiction. For example, if a court has a populous, diverse Asian population in its jurisdiction, the court may collect data on nationalities or cultural groups within this population. Whereas the high-level categories discussed above are useful for making comparisons across courts or agencies, this more granular data gives the court important, locally relevant information on experiences and outcomes among cultural groups that may be lost by using only the high-level categories. This information is vital for examining disproportionalities and disparities, as well as for ensuring that the court provides accessible, culturally responsive services.

Courts can identify these communities through multiple means. Census data provides population data that can be used as a starting point for identifying these groups. More in-depth discussions on this topic can be achieved through community engagement activities like town halls or focus groups with court users and local cultural organizations. The *Racial Justice Organizational Assessment Tool for Courts* includes a section on community engagement that will help in this effort.

Although more granular ethnic communities should be identified and included in equity analyses, courts must also take issues of confidentiality into account. If there are very few members of a specific community engaged with the court, there is a risk that these individuals will be identifiable in published information about this group. If a group has less than ten members represented in the data, take great care when publishing information about the group. If there are less than 6 members in the group, do not publish information about the group.

Sources of Race and Ethnicity Information

Finally, courts should also record the source of their race and ethnicity data. In order to make accurate interpretations of data, it is important for the court to know whether a person's race and ethnicity information was collected by the court or received from an outside agency. It is also important to know whether the data are self-reported or based on observation by another person.[v]

[v] United States Census Bureau. (2011). 2010 Census Shows America's Diversity. Retrieved from: https://www.census.gov/newsroom/releases/archives/2010_census/cb11-cn125.html
Pew Research Center. (2021). Majority of Latinos Say Skin Color Impacts Opportunities in America and Shapes Daily Life. Retrieved from: https://www.pewresearch.org/race-and-ethnicity/2021/11/04/measuring-the-racial-identity-of-latinos/

How to Collect Race and Ethnicity Information

Identifying methods of race and ethnicity data collection is also a complex task for courts. Methods of collection must be tailored to how the court operates and to the specific nature of different case types. For example, race and ethnicity data pertaining to the court workforce will come from a different source than data for criminal defendants or data for jurors. Race and ethnicity data collection for juvenile cases will be subject to different challenges than collection for civil cases. The rest of this section will review different data collection methods and their strengths and weaknesses.

Collection by the court (primary data collection)

Some race and ethnicity data can be collected by the court. For example, data on the court workforce can come from surveys of court staff or from data collected as part of the hiring process. Race and ethnicity data of limited samples of court users may come from court surveys on topics like procedural fairness or user satisfaction, while information on summoned jurors may be obtained from juror qualification questionnaires. Courts may collect more comprehensive race and ethnicity data linked to cases at the time of case filing, through eFiling or cover sheets.

Technology plays a role in allowing the court to implement primary race and ethnicity data collection. eFiling systems can prompt users to provide or review their racial and ethnic identity, as can court kiosks that are used to check-in court users.

Data collected by the court have the benefit of being structured in the way the court determines. The court can determine the categories to collect and the types of responses (e.g., providing an option to select all that apply) and provide individual-level data, which allows for more in-depth analyses than would aggregate data. Courts should partner with community members to develop and test questions about racial and ethnic identity in these contexts. As these options gain popularity, courts will be able to collect more data that adheres to a framework determined by the court.

Considerations

Primary data collection, or original data collection, requires more time and resources from the court than secondary data (data collected for another purpose). Primary data collection often requires the use of cover sheets or eFiling, which may not be part of the business process for all case types. It also relies on the willingness of court users to provide their race and ethnicity information to the court. Court users may be concerned that providing race and ethnicity data will be detrimental to their case outcome or other decisions or responses made by the courts. Although NCSC is developing guidance on *Data Use Statements,* which will increase responses by explaining how information will be used to help ensure equal treatment under the law, non-response remains a concern. Primary data collection will also be difficult, if not impossible, in cases where the court never makes contact with the user. This occurs when the person named as a defendant or respondent does not communicate with or appear in court (e.g., in a default judgment).

Data Exchanges

Information exchanges with other agencies are an efficient strategy for obtaining race and ethnicity data which eliminate the need for courts to undertake the effort of primary data collection. These agencies may include the State Drivers' License Agency, the Statewide Automated Child Welfare Information System (SACWIS), corrections, or law enforcement agencies. A significant benefit of this method of data collection, particularly for criminal cases, is that data imported from criminal justice agencies may already be in a format compatible with the National Crime Information Center's (NCIC) race and ethnicity data framework.

A data exchange typically requires a memorandum of understanding (MOU) with the agency providing or receiving the data.

According to the *Data Governance Policy Guide*, each MOU should include at a minimum:
• What data will be exchanged;
• Format of data to be exchanged;
• Frequency of data exchange (e.g., real time, hourly, daily, monthly);
• How to handle inaccurate data;
• How each agency will be notified if the data sent are changed in format or content;
• Who the primary contact persons are at each agency;
• Expiration date of the agreement.

Courts must also determine what data are automatically added to the court data management system, and what data requires additional review before acceptance.

Considerations

Data from outside agencies will be structured in a way determined by those agencies, and therefore, may not adhere to current best practices in race and ethnicity data collection. Courts may not be able to determine, for example, whether data were collected through self-reporting by the person or by another person's observation (e.g., law enforcement officer).

When possible, courts can map external agency data to the preferred framework. However, if, for example, the agency only allows for the single selection of a race field and uses a "more than one race" or "multi-racial" category, multi-race people will not be represented accurately by their actual racial heritage in the data and cannot be mapped to a framework that allows for the selection of multiple races. Similarly, if the agency data combines racial categories (like Asian/Pacific Islander), courts will be unable to disaggregate those data into individual categories.

Courts should map external agency data to the recommended categories to the extent they are able. NCSC further recommends, through NODS, that courts record the sources of the data to help courts identify how the data were collected, and any framework issues with the data. This practice can help to address the issue with exporting data to the NCIC framework, which does not conform with best practices since it combines several individual categories (White and Middle Eastern; Asian and Pacific Islander) and allows for only a single selection for race. Rather than conforming to the categories used by the NCIC, courts should maintain more granular, multi-select data and map those data to the NCIC's categories for that data exchange. Multi-racial individuals may have to be categorized as "unknown" in that mapping due to the constraints of the NCIC system. If courts maintain a source field for data, they can use the race/ethnicity data linked to law enforcement collection for the purposes of reporting judgments in criminal cases.

Figure 2:

> ### Valid NCIC Data Values for Race as of June 2024
>
> - A – ASIAN - ASIAN OR PACIFIC ISLANDER - A PERSON HAVING ORIGINS IN ANY OF THE ORIGINAL PEOPLES OF THE FAR EAST, SOUTHEAST ASIA, THE INDIAN SUBCONTINENT OR THE PACIFIC ISLANDS.
> - B – BLACK - A PERSON HAVING ORIGINS IN ANY OF THE BLACK RACIAL GROUPS OF AFRICA.
> - I – AMERICAN INDIAN - AMERICAN INDIAN OR ALASKAN NATIVE - A PERSON HAVING ORIGINS IN ANY OF THE ORIGINAL PEOPLES OF THE AMERICAS AND MAINTAINING CULTURAL IDENTIFICATION THROUGH TRIBAL AFFILIATIONS ORCOMMUNITY RECOGNITION.
> - U – UNKNOWN - UNKNOWN.
> - W – WHITE - WHITE - A PERSON HAVING ORIGINS IN ANY OF THE ORIGINAL PEOPLES OF EUROPE, NORTH AFRICA, OR MIDDLE EAST.)

Race and ethnicity data will be maintained according to the standards of the data source agencies, which may differ from other agencies or the court's own practices. Data users should have an understanding about the background of the received data to minimize inaccuracies, especially when integrating these data with court-collected data or other sources. The NCSC recommends obtaining sufficient documentation on how the agency maintains its database, including details such as time stamps. When collecting data from multiple sources, courts need to know which record is the most recent for an individual. In addition to knowing how the agency collected the data, it is beneficial to know what data processing occurred, if and how often records have been updated in the past, how data were standardized to the final categories, and how missing data were treated.

Statistical methods to impute race and ethnicity

At times, a court may wish to implement equity analyses when neither original nor secondary data collection is feasible. Although data collected at the individual level are preferable, racial imputation offers a solution when such data are needed to conduct certain analyses but are not available. Racial imputation, first developed by researchers in the health policy field, has become an important methodology that provides probabilities for an individual's race or ethnicity when that information is missing or incomplete. While there are various approaches, the Bayesian Improved Surname Geocoding (BISG) method has emerged as a leading and validated technique. This method combines surname information and an individual's geocoded address to calculate probabilities. When the address and the name are known, researchers can estimate the probability that an individual is, for example, Black, given that they live in a certain location, based on that location's Census Bureau demographics. A researcher then updates this initial probability with the probability that the individual is Black (or another race) given their surname, using surname lists compiled by the Census Bureau. All required data are publicly available from the Census Bureau and can be readily accessed via download or API calls.

It is important to note that racial imputation is only appropriate for analysis at an aggregate level. This method allows for an estimated probability that someone is of a certain race, but it cannot truly predict race for any individual. However, the probabilities for each racial category assigned to an individual can be used to aggregate the records and describe the demographic composition of a given sample. Racial probabilities (rather than estimated definite categorizations) are sufficient for conducting disparity or outcome analyses, or any analysis where race is used as an explanatory variable. The method enables courts to detect patterns of disparity and while disparate treatment or outcomes cannot be discerned for a given individual, the aggregate results allow for the identification of broader inequities. Courts can also use these results to prioritize areas for original data collection, enabling further analyses with actual data in areas where estimated results had indicated potential issues.

Considerations

There is variation in how well the technique works for different racial groups and across jurisdictions. The lingering effects of racial segregation both geographically and in terms of marriage make this a generally effective and accurate method in the United States. However, the method is less effective in highly heterogeneous locations. Similarly, the assumption that certain surnames are more common in specific racial or ethnic groups applies to certain groups (e.g. Hispanic and Asian) more than others and does not apply easily to multiracial individuals – a growing group - or when names are adopted via marriage. Additionally, the Census Bureau's surname dictionary does not include all surnames, and any data quality issues with the court's original records (name spelling mistakes, incomplete addresses) will cause issues when attempting to merge records with the Census information. These caveats must be clear to all who consider insights gained through this method.

While the method's basic principle is commonly agreed upon, research on improving the method's accuracy is fast evolving, and various tools build on the basic technique, each offering different improvements for various needs. Available open-source tools differ in several aspects, including the number and type of racial categories utilized and the reference data chosen from the Census Bureau (populations and time frames). Therefore, a researcher utilizing this method must consider which, if any, of the readily available tools will fit the given analysis framework. As these limitations might interfere with a court researcher's ability to conduct accurate analyses, including the ideal level of detail and data from the correct time period, NCSC is developing an imputation tool specifically designed for court purposes.

Additional Practical Considerations

In addition to determining the data format that the court will collect and the methods for collecting it, there are a number of practical issues courts should consider.

Case Management Systems and their Vendors

Technology plays a key role in determining what data the court can collect and store. In some cases, courts may aim to collect data in a manner consistent with NCSC guidelines but be unable to use or store the data in that format due to constraints imposed by a case management system (CMS). Courts that want to follow best practices in race and ethnicity data collection (such as ensuring that mapping to the categories contained in the OMB guidelines is possible; allowing for a multi-select option; collecting more granular data based on the needs of the court) need to ensure that their CMS allows for these practices.

Courts should investigate the functionality of their CMS regarding race and ethnicity data. For example, at first glance, a system may appear to not allow for a "select all that apply" option for race. However, upon further investigation, the court may find that once the "multi-racial" option is selected, the user can then select all of the races that apply in a sub-menu. In other cases, (re)configuring the values in the race/ethnicity table may accomplish the change needed.

Courts should contact vendors to discuss the functionalities of their race and ethnicity data framework. The vendors will provide insight into functionality that already exists, or request if the vendor can implement the desired changes.

Public Communication about the Purposes and Uses of the Data

As described above, because some court users are unaccustomed to being asked for race or demographic information when they interact with the courts, users will likely have questions about why their race is being recorded and how it will be used. Courts should develop a communications strategy to provide clear and concise explanations to court users. NCSC is developing guidance on *Data Use Statements* to help the courts develop their messaging.

Personnel Training

In cases where court personnel are responsible for entering race and ethnicity data into the case management system, courts should provide training to ensure that the information is entered consistently. It is important to minimize the number of cases with missing data, as these limit the extent to which the court can use the data to answer important questions. It is also important to ensure that the categories are defined and selected consistently by individual staff members. Finally, in situations where court personnel are asking court users directly for their race (as opposed to court users completing a form themselves), courts should train staff on how to ask these questions in a consistent and respectful manner and to be provided resources to answer court users' questions.

Access to the Data

Courts that collect individual-level race and ethnicity data should have clear policies about who has access to the information. Courts should consider under what circumstances the data may be made available to external parties, such as researchers, policy makers, the public, or the media. Researchers should have protocols in place to protect the data and individual's privacy, which reduces the risk of data sharing. Information on race or ethnicity should not be publicly shared if someone could be identified as a result of that information. One common guideline is not to share race or ethnicity information if there are fewer than six individuals in a particular category. In that situation, data should be combined into a larger category and reported in aggregate format. Courts should also consider whether individuals may become identifiable if their race and ethnicity data is combined with additional information about them. Finally, courts should have a clear understanding of how the data fit into the court's obligations under local and federal Freedom of Information laws, and they should develop policies and procedures to protect court users' privacy, as needed.

Data Analysis Planning

Once a court collects or imputes race and ethnicity data, data should be regularly reviewed and analyzed for insights into court users' experiences. Some examples for courts include the forthcoming *Racial Equity Process Guide*, which provides information about how to analyze race and ethnicity data to identify disproportionalities and disparities. The *Equity and Inclusion Assessment Tool for Drug Treatment Courts* and the *Equity Analysis for Dependency Courts* provide examples of these methods in different court contexts.

Future Work

NCSC continues to work with various personnel within the state courts to achieve efficient and effective race and ethnicity data collection. NCSC, through discussions, convenings, learning labs, and educational programs, is dedicated to building and advancing the guidance on race and ethnicity data collection and analysis. The court community, including Court Statistics Project data specialists, human resource managers, state and local leaders, and diversity, equity, and inclusion professionals, has collectively shared successes on using race and ethnicity data to inform decisions to remedy inequities. The court community as a whole has demonstrated a willingness to ensure that "justice is not only fair to all, but also is recognized by all to be fair."[vi] Race and ethnicity data will enable courts to communicate their successes and build trust among those it serves. As we learn from each other and advance our understanding of the ways courts can collect and utilize race and ethnicity data, NCSC will continue to release updated guidance.

NCSC Resources

- *Collecting Race & Ethnicity Data*
- *Blueprint for Racial Justice Initiative*
- *Racial Justice Organizational Assessment Tool for Courts*
- *Data-Driven Decision Making Guide for Courts*
- *Data Governance Policy Guide*
- *Equity Analysis for Dependency Courts*

Forthcoming

- Racial Equity Process Guide
- Guidance on Data Use Statements
- Updated NODS race and ethnicity categories

The Court Statistics Project is the only source for comparable annual state court caseload data, with court data published at www.courtstatistics.org. For more information about court data, please see our website at www.courtstatistics.org/court-statistics/interactive-caseload-data-displays/csp-stat. NCSC is available to provide training or assistance! Contact CSP Staff at csp@ncsc.org for more information.

[vi] Conference of Chief Justices and Conference of State Court Administrators. (2020). Resolution 1: In support of racial equity and justice for all. Retrieved from: https://ccj.ncsc.org/__data/assets/pdf_file/0017/51191/Resolution-1-In-Support-of-RacialEquality-and-Justice-for-All.pdf