Arkansas Prekindergarten Research Alliance Data Inventory Overview

December 2015
# Contents

Introduction and purpose .................................................................................................................. 1  
Data inventory content ...................................................................................................................... 2  
How to use the data inventory ......................................................................................................... 2  
Data sources included in the inventory ............................................................................................. 3  
  Arkansas Department of Education Statewide Information System ........................................... 3  
  Qualls Early Learning Inventory .................................................................................................. 4  
  Arkansas Comprehensive Testing, Assessment, and Accountability Program  
    criterion- and norm-referenced tests ......................................................................................... 4  
  Head Start data .............................................................................................................................. 5  
    Head Start Enterprise System ..................................................................................................... 5  
    Arkansas Head Start Association and Arkansas Head Start State Collaboration Office .... 5  
    Arkansas Head Start Data and Statistics for the 2011–12 Program Year ............................... 5  
    Center for Law and Social Policy Head Start by the Numbers 2011 PIR Profile ............... 6  
  2008 Child Care Licensing Study ............................................................................................... 6  
  2011 Child Care Licensing Study .................................................................................................. 6  
  Kids Count Data Center ................................................................................................................ 6  
  Center for Law and Social Policy DataFinder .............................................................................. 7  
  Common Core of Data .................................................................................................................. 7  
  American Community Survey ....................................................................................................... 7  
Data sources partially inventoried or not included .......................................................................... 8  
  Arkansas Better Chance ................................................................................................................ 8  
  Better Beginnings and child care licensing data ......................................................................... 8  
  Traveling Arkansas Professional Pathways Professional Development Registry .................. 9  
  Child Care Assistance Program .................................................................................................... 9  
  Supplemental Nutrition Assistance Program .............................................................................. 10  
  Federal survey data collections .................................................................................................. 10  
References ........................................................................................................................................ 10  

## Appendixes

Appendix A: Data inventory population protocol guidance on searching  
  documentation manuals ............................................................................................................... 11  
Appendix B: Inventory fields ......................................................................................................... 14
Introduction and purpose

The Arkansas Prekindergarten Data Inventory, created by the Regional Educational Laboratory (REL) Southwest, identifies the data available at federal, state, and local levels to support research on prekindergarten (preK) care and education in Arkansas. Organizations at all levels collect and manage a wealth of data that can be used to conduct research and inform education policy and practice. Because so many organizations house data, however, it can be difficult to locate, evaluate, and obtain the data one needs. The Arkansas Prekindergarten Data Inventory addresses this issue, providing information on relevant data sources, the data elements they contain, and how those elements relate to the Arkansas Prekindergarten Research Alliance’s interests.¹

Using the inventory, alliance members and researchers can leverage the data available to inform and advance the alliance’s research agenda.

This document provides an overview of the Arkansas Prekindergarten Data Inventory, which is a separate Excel spreadsheet available upon request (see sidebar above). The inventory contains information about data collected and maintained by the Arkansas Department of Education (ADE), the Arkansas Department of Human Services (DHS), the U.S. Department of Education, and other organizations. Included data elements focus on children from birth through elementary school as well as parents, caregivers, teachers, preK programs, and elementary schools.

The data inventory is tailored to the needs of the Arkansas Prekindergarten Research Alliance. To create the inventory, a REL Southwest team, working in coordination with alliance members, identified potential data sources and key constructs that reflect the alliance’s research priorities and interests. For example, a key alliance interest is how preK program features and participation are related to elementary school readiness and performance. This alliance interest and others shaped the data sources included in the inventory.

Because the research agenda for the Arkansas Prekindergarten alliance continues to evolve and the data inventory may be used to inform multiple studies, the REL Southwest team erred on the side of inclusion when deciding which data elements, or variables, to capture. Thus, the team attempted to capture all variables related to preK care and preK and elementary education in Arkansas, including variables related to demographics, child health, and preK program participation and characteristics, as well as other covariates and outcomes. The team also determined the extent to which researchers may use unique identifiers to link school, student, and program data across multiple systems. The team did not inventory those data elements deemed irrelevant to the alliance’s research agenda, however, and was unable to inventory some relevant data sources (either in part or in full, as explained later in this document).

¹ See http://relsouthwest.sedl.org/research-alliances/arkansas_prek.html for more information on the Arkansas Prekindergarten Research Alliance, including its goals, research agenda, and core members.
Data inventory content

The data inventory Excel spreadsheet is formatted according to the Version 3 Data Model Guide, established by the Common Education Data Standards (CEDS, 2013). To create the spreadsheet, the REL Southwest team used a protocol to guide the inventory population process and the systematic search of data sources (see appendix A).

The data inventory contains information about the specific data sources and systems where relevant data are housed. For each source, the inventory provides detailed information about each data element, including its level (for example, school-, course-, or student-level data) and whether it is an identifier (a variable that enables data to be linked across units and over time). When possible, the inventory also indicates the time of year each data element is collected and for which years data are available. In addition, the inventory categorizes each variable in terms of how it may be used in a rigorous research study. For example, the spreadsheet indicates whether a variable represents a child outcome measure, addresses a program or policy related to early childhood education in Arkansas, or could be used as a demographic variable or covariate (appendix B provides a list of the fields in the inventory). Last, to meet the Arkansas Prekindergarten Research Alliance’s need to identify data about preK programs, the inventory contains a field that indicates whether each data element is related to prekindergarten.

How to use the data inventory

The Arkansas Prekindergarten Data Inventory is intended to serve multiple purposes, including providing information on extant data that may be used to conduct rigorous research studies or to assist the alliance in developing and refining research questions. In addition, the inventory can be used to identify gaps in existing data systems, thereby suggesting potential areas for improvement in data collection and management to facilitate future research.

As mentioned, the data inventory is an Excel spreadsheet, or workbook, that consists of multiple worksheets. The first worksheet, “All Variables,” contains all inventoried variables across all data sources. This worksheet can be used to obtain a sense of all relevant variables. For example, a researcher may focus on the inventory field “Pre-K Program?” to determine which data sources contain information on preK program participation or characteristics. The researcher could then filter or sort the field to display only “Yes” entries.

The second worksheet, “Variable Linking Information,” contains information on variables that can be used to link inventoried data to other sources, including other data in the inventory. Starting with the third worksheet, the same information contained in the “All Variables” worksheet is provided separately for each data source. These individual data source worksheets can be used to focus on the relevant elements within a single data system.
Data sources included in the inventory

The following questions were used in determining which data sources to include in the data inventory:

- Are the data relevant to prekindergarten care and education experiences of Arkansas children or their subsequent elementary education experiences and outcomes?
- Are the data relevant to early childhood experiences, such as health care?
- Can the data be disaggregated by prekindergarten program?
- Are the data available for the state of Arkansas? Can the data be disaggregated further (by district, school, prekindergarten program, and so on)?

Using these guiding questions, the team chose to include the data sources described below. However, the inventory is intended to be a living document, and additional data sources may be added as they are identified and deemed relevant to the alliance’s needs.

Arkansas Department of Education Statewide Information System

The Arkansas Department of Education (ADE) Statewide Information System (SIS) is a comprehensive statewide database on students, programs, district personnel, and finances, as reported by school districts. The database is intended for use in planning and policymaking. The SIS website aggregates public data across a range of topics, including course enrollment, student demographics, teacher and staff counts, and school and district finances. Although individual-level data cannot be accessed directly from the SIS website, users can go to the SIS Reports area and either generate standard reports at the school, district, county, or state levels, or request custom reports. For more information, see https://adedata.arkansas.gov/sis/.

SIS data are available for the 2004/05 to 2014/15 school years, with the most recent documentation for the 2015/16 school year. REL Southwest used the 2013/14 documentation to inventory this data source, focusing on student, staff, and school variables. SIS financial and district data were considered less important to the alliance and were generally not inventoried.

SIS data can be linked to other data sources, including ADE-administered assessments, through the student identifier. In addition, school data include local education agency identifiers, enabling SIS data to be linked to the Common Core of Data to provide supplemental school information. Although not available to the public, SIS data are available for authorized research through the Arkansas Research Center (ARC).² REL Southwest is pursuing data use agreements to enable alliance members to access these data for approved activities. Alliance members interested in obtaining this dataset should contact REL Southwest for more information.

» In the Excel workbook, this data source is inventoried in the “SIS” worksheet.

---

² Established in 2008 with the support of a grant from the Department of Education’s Institute for Education Sciences, ARC serves as a research hub for collaboration across several Arkansas state agencies. Research conducted through ARC combines data across preK, elementary, secondary, and postsecondary education and the workforce. For more information on ARC, see https://arc.arkansas.gov/. For examples of relevant ARC research, see Argue and Holland (n.d.) at https://arc.arkansas.gov/arc_web/resources/publications/ABC_Shrinks_Gap.pdf.
Qualls Early Learning Inventory
At the start of the school year, the ADE administers the Qualls Early Learning Inventory (QELI) to all Arkansas kindergarteners and to any grade 1 students who did not attend kindergarten. The QELI is used to assess student development in six areas: general knowledge, oral communication, written language, mathematical concepts, work habits, and attentive behavior. For general information on the QELI, see http://www.riversidepublishing.com/products/qeli/index.html. Information is also available about the QELI’s use in Arkansas.

QELI data include a student identifier that enables linking to SIS data and other student assessments, such as the Augmented Benchmark. These data are available for authorized research through ARC but are not publicly available. Alliance members interested in obtaining these data for research may submit a request at https://arc.arkansas.gov/arc_web/datarequest.php.

» In the Excel workbook, this data source is inventoried in the “ADE Qualls” and “Augmented Benchmark” worksheets.

Arkansas Comprehensive Testing, Assessment, and Accountability Program criterion- and norm-referenced tests
As part of the Arkansas Comprehensive Testing, Assessment, and Accountability Program, students across a range of grades are administered a set of criterion- and norm-referenced assessments, which make up the Augmented Benchmark Examinations. Students in grades 1, 2, and 9 are administered norm-referenced Iowa Tests—the Iowa Tests of Basic Skills in grades 1 and 2, and the Iowa Tests of Educational Development in grade 9. These tests measure the performance of students in comparison to a national norm and are used to identify students in need of remediation in language arts, mathematics, and science. Students in grades 3–8 receive criterion-referenced tests based on the Arkansas Mathematics, Science, and English Language Arts Curriculum Frameworks. These tests consist of multiple-choice and open-response questions in mathematics, reading, and writing as well as a writing component. In addition, in grades 5 and 7, the tests include multiple-choice and open-response questions in science.

To inventory this student-level data source, REL Southwest staff reviewed file documentation (Excel files) provided by ADE. The data source contains information on each student, including name, state student ID, demographics, characteristics such as gifted and talented and limited English proficiency (LEP) status and whether testing included LEP accommodations, and detailed information on the Augmented Benchmark Examinations. The assessment data also include a student identifier that enables linking to SIS data and other assessments, such as QELI.

Augmented Benchmark Examination data are available for authorized research through ARC but are not publicly available. Alliance members interested in obtaining this dataset should contact REL Southwest for more information. The Augmented Benchmark Examinations data are available through the 2013/14 school year. In 2014/15, Arkansas replaced the assessments with the Partnership for Assessment of Readiness for College and Careers (PARCC) assessments, for which data are not yet available. In the 2015/16 school year, the PARCC assessments were replaced with the ACT Aspire summative assessment.

» In the Excel workbook, this data source is inventoried in the “ADE NRT” and “Augmented Benchmark” worksheets.
Head Start data

Head Start data are available from multiple sources, including federal and state agencies as well as nongovernment research organizations. Because different sources have different foci (for example, program, state, and regional levels), various sources are included in the inventory.

Head Start Enterprise System

Head Start is a federal program managed by the Administration for Children and Families in the U.S. Department of Health and Human Services. Head Start and Early Head Start provide preschool-age children from low-income households, as well as their families, with services to improve cognitive, social, and emotional well-being. The programs may be based at centers or schools, family child care homes, or children’s own homes. All programs that receive Head Start funds are required to submit Program Information Reports (PIR), which provide information on services, staff, and the children and families served. PIR data, available on request, include program-level information, such as program name and address, program type, total number of funded Head Start or Early Head Start children enrolled, child characteristics, information about staff and funding, and the use of evaluative program assessments. To inventory Head Start PIR data, REL Southwest used the 2012/13 PIR Form (a 2015/16 form is available as well). Summaries of changes up through 2014/15 and 2015/16 are also available online.

In the Excel workbook, this data source is inventoried in the “Federal PIR” worksheet.

Arkansas Head Start Association and Arkansas Head Start State Collaboration Office

The Arkansas Head Start Association and the Arkansas Head Start State Collaboration Office provide tabular summary data about Head Start, Early Head Start, and Migrant/Seasonal Head Start programs and the children served for each of the 22 agencies providing such services in Arkansas. The data are publicly available on the website at http://arheadstart.org/, which provides detailed information for each agency (listed as programs; see an example). REL Southwest inventoried this source from the tabular data provided. Much of the data available at this site are obtained from Head Start PIR data and aggregated to the agency level; however, information such as total staffing is not drawn from the PIR. Data from this source can provide information on Head Start services at the agency level without obtaining and aggregating program-level data.

In the Excel workbook, this data source is inventoried in the “ARHS PD” worksheet.

Arkansas Head Start Data and Statistics for the 2011–12 Program Year

An additional source of information on Arkansas Head Start programs and participation is the Arkansas Head Start Data and Statistics report, compiled by the Arkansas Head Start Association (www.arheadstart.org). Data included in the 2011–12 report, which was used to develop the inventory, were largely based on PIR data aggregated to the state level with additional data (for example, the percentage of state Head Start programs that have Better Beginnings ratings of two or three stars). These data can be used to evaluate statewide characteristics of Head Start, Early Head Start, and Migrant/Seasonal Head Start programs and the children they serve. The 2011–12 report, as well as the 2012–13 version, are available online.

In the Excel workbook, this data source is inventoried in the “ARHS DS” worksheet.
Arkansas Prekindergarten Research Alliance
Data Inventory Overview

**Center for Law and Social Policy Head Start by the Numbers 2011 PIR Profile**
The Center for Law and Social Policy (CLASP) analyzes PIR data for Head Start, Early Head Start, Migrant/Seasonal Head Start, and American Indian and Alaskan Native programs to create annual state fact sheets, viewable from the CLASP website. Each fact sheet is created using PIR data aggregated to the state level. The 2011 Arkansas state fact sheet was used to inventory this data source. The 2012 and 2013 Arkansas state fact sheets are also available online.

» In the Excel workbook, this data source is inventoried in the “ARHS PIR” worksheet.

**2008 Child Care Licensing Study**
The 2008 Child Care Licensing Study (CCLS) contains detailed information on the processes and policies for licensing child care centers and family child care programs in all 50 states and the District of Columbia. The data are collected by the National Association for Regulatory Administration (NARA) and the National Child Care Information and Technical Assistance Center (NCCIC), a service of the Child Care Bureau, the Administration for Children and Families, and the U.S. Department of Health and Human Services. The 2008 survey data codebook, available from the Inter-university Consortium for Political and Social Research (ICPSR), was used to inventory this source. The data include information on licensing of facilities and staff, inspections and monitoring, complaint investigations, enforcement actions, and role of licensing in quality initiatives. The NARA website provides information about the 2008 CCLS, as well as the entire CCLS series, and data and documentation files are available through the ICPSR website.

» In the Excel workbook, this data source is inventoried in the “CCLS 2008” worksheet.

**2011 Child Care Licensing Study**
In 2011, NARA and NCCIC partnered again to update the 2008 CCLS. The 2011 CCLS contains more than 3,000 variables and provides information on licensing practices, regulatory requirements, and operations by state and family child day care homes, group day care homes, and child day care centers. As with the 2008 CCLS, data and documentation files are available on the ICPSR website.3

» In the Excel workbook, this data source is inventoried in the “CCLS 2011” worksheet.

**Kids Count Data Center**
Data from the Kids Count Data Center, sponsored by the Annie E. Casey Foundation, have been published annually since 1989. The most recent data release (2013) focuses on 16-child level indicators—for the nation and by state—across four domains: Economic Well-Being, Education, Health, and Family and Community. Information about Kids Count data can be found at http://datacenter.kidscount.org/about. Although the data are not available for download, the Kids Count Data Center enables users to generate custom reports and data tables.

» In the Excel workbook, this data source is inventoried in the “Kids Count” worksheet.

---

3 The CCLS was repeated in 2013, but data documentation is not available through ICPSR and information about the study from NARA is currently unavailable.
Center for Law and Social Policy DataFinder
The Center for Law and Social Policy (CLASP) DataFinder is an online table-creation tool that aggregates data from multiple sources, such as the American Community Survey (ACS), the U.S. Department of Health and Human Services for Child Care and Development Fund statistics, and the Administration of Children and Families for Head Start statistics. The DataFinder tool can be used to generate customized tables of demographic and administrative data for programs for low-income individuals and families; tables can be generated for the nation as well as by state and for selected communities (no Arkansas communities are included). The most recent data are for 2014. The tables can be downloaded, but the data are not linkable to other data sources. For more information on the DataFinder tool, see http://www.clasp.org/data.

» In the Excel workbook, this data source is inventoried in the “CLASP” worksheet.

Common Core of Data
The Common Core of Data (CCD), an annual data collection sponsored by the National Center for Education Statistics, provides state-, district-, and school-level data, including information on education agencies and the students enrolled in them. CCD files are publicly available, and the CCD website provides an analytic “build-a-table” tool to create tabulations from the CCD. For more information on the CCD, see http://nces.ed.gov/ccd/ccddata.asp.

The CCD contains fiscal and nonfiscal data. REL Southwest determined that the fiscal variables were unrelated to the alliance’s research agenda and did not inventory them. The majority of nonfiscal variables were inventoried, however, with the exception of information on library and administrative support staff, district and school addresses, and imputation flags. The 2010/11 documentation was used to inventory CCD variables, including student enrollment counts by race, gender, and grade level as well as school-level staffing. The most recent CCD documentation is for 2013/14. Data from the CCD, which are available for download online, may be useful for creating baseline reference information about preK and early elementary school and district student enrollment in Arkansas, for the state or by region.

» In the Excel workbook, this data source is inventoried in the “CCD” worksheet.

American Community Survey
The American Community Survey (ACS), an ongoing survey conducted by the U.S. Census Bureau, collects information on individuals, including demographic information such as age, sex, race/ethnicity, family composition, and school and preschool enrollment. The ACS can be analyzed at the national, state, and regional levels. Its potential utility for this alliance includes the ability to examine national, state, and regional trends in preschool enrollment using the five-year Public Use Microdata Samples (PUMS). ACS data are not linkable to other sources but are available for research. PUMS data can be requested or analyzed using web tools maintained by the Census Bureau. For more information on the ACS, see http://www.census.gov/acs/www/.

» In the Excel workbook, this data source is inventoried in the “ACS” worksheet.

The ACS and CCDF data used by CLASP to generate custom tables are available from other sources as well. ACS data are available from ICPSR and the U.S. Census Bureau; CCDF data are available from ICPSR.

4 The ACS and CCDF data used by CLASP to generate custom tables are available from other sources as well. ACS data are available from ICPSR and the U.S. Census Bureau; CCDF data are available from ICPSR.
Data sources partially inventoried or not included

The following data sources were either partially inventoried or not included in the inventory because they were redundant, did not support the alliance’s current research agenda, or did not provide sufficient documentation to inventory (or fully inventory) them.

Arkansas Better Chance

The Arkansas Better Chance (ABC) Program is a state-funded preK program that serves children from birth up to age 5 or kindergarten entry. Arkansas first began offering the program in 1991, with eligibility and enrollment expanding over time (Barnett, Carolan, Fitzgerald, & Squires, 2011). ABC funding requires each program to enter information into the state’s Child Outcome, Planning and Administration (COPA) database, including data about staff and the children and families served. Mandatory information includes student eligibility information, student growth assessment, student developmental screening, student attendance, and staff information. ABC forms are available on the Arkansas Department of Human Services (DHS) website. Child forms indicate that the COPA database also enables the collection of information on child health and well-being, including health status, health screening results, immunization status, and social and behavioral indicators, although none of these data elements are mandatory. The Child Application form also has many required fields, including detailed caregiver (usually parent or guardian) information, such as name, contact information, and employment information; family roster information; child name and Social Security Number; and basic demographics.

Although data are entered into the COPA system by staff at eligible child care programs, the data are processed and maintained by the software provider. A data dictionary or other documentation on the ABC fields within the COPA system, needed to populate the inventory, is not available from the Arkansas DHS. Because of the importance of ABC data to the alliance’s research agenda, however, information from the ABC forms was entered into a worksheet in the inventory.

In the Excel workbook, this data source is partially inventoried in the “ABC” worksheet.

Better Beginnings and child care licensing data

Better Beginnings, Arkansas’s voluntary quality rating and improvement system (QRIS) for licensed and registered child care providers, provides a “systemic approach to assess, improve, and communicate the level of quality in early care and education programs. Program types include center-based care, family child care, and school age care programs” (Arkansas Department of Human Services, Division of Child Care and Early Childhood Education, 2010, p. 6). Parents can use the Better Beginnings rating (Level 1, 2, or 3) to assist in their selection of child care.

Child care providers can apply for a Better Beginnings rating and provide information on specific program components: administration, administrator and staff qualifications, professional development, learning environment, facility assessment, and child health and development. For each component, the Better Beginnings rating system provides clear milestones that must be met. For example, the Environment Rating Scale (ERS) or Youth Program Quality Assessment (YPQA) is used to rate the quality of the environment in a center-based child care program. A Level 1

5 REL Southwest confirmed that the Arkansas Department of Human Services does not possess these materials.
rating requires that the facility complete a self-evaluation, a Level 2 rating requires that the facility score an average of 3.00 or higher on the ERS or YPQA for each classroom reviewed, and a Level 3 rating requires that the facility score an average of 4.00 or higher on the ERS or 3.75 or higher on the YPQA for each classroom reviewed. Similar criteria exist for other program types through the Family Child Care Environment Rating Scale Assessment and the School-Age Care Environment Rating Scale or YPQA. Better Beginnings QRIS documentation provides more information regarding eligibility, requirements, and the types of assessments and evaluations.

Better Beginnings data are maintained in the Arkansas Data Warehouse as part of the Child Care Licensing data system. REL Southwest confirmed that suitable documentation is not available to inventory this data source. However, the Arkansas DHS data warehouse provided limited variable information, which was used to partially populate the inventory.

» In the Excel workbook, this data source is partially inventoried in the “CCLAS” worksheet.

**Traveling Arkansas Professional Pathways Professional Development Registry**

The Traveling Arkansas Professional Pathways (TAPP) Professional Development Registry is the tracking component of the Arkansas Early Childhood Professional Development System. The TAPP registry is voluntary, and members have access to a listing of professional development opportunities, with the TAPP registry maintaining a transcript of trainings received over the past year. This data source is relevant to the Arkansas Prekindergarten alliance as a resource for understanding patterns of professional development activities among early childhood care and education providers. REL Southwest confirmed that suitable documentation is not available to inventory this data source. However, the Arkansas DHS data warehouse provided limited variable information, which was used to partially populate the inventory. For more information on the TAPP registry, see [http://professionalregistry.astate.edu/registry/whatis.asp](http://professionalregistry.astate.edu/registry/whatis.asp).

» In the Excel workbook, this data source is partially inventoried in the “TAPP” worksheet.

**Child Care Assistance Program**

The Arkansas DHS Division of Child Care and Early Childhood Education provides families from low-income households with financial assistance for quality child care through the Child Care Assistance Program and the Child Care Development Fund (CCDF). This program is under the administration of the U.S. Department of Health and Human Services. The Arkansas DHS maintains a database of information about participating children, their families, and the programs the children attend, as well as applicants who are awaiting support or those whose applications were denied or determined to be ineligible. REL Southwest confirmed that suitable documentation is not available to inventory the participant database (a related database that contains billing information for program participants was not considered). However, the Arkansas DHS data warehouse provided limited variable information, which was used to partially populate the inventory. For more information on the Child Care Assistance Program, see [http://humanservices.arkansas.gov/dccece/Pages/ChildCareAssistance.aspx](http://humanservices.arkansas.gov/dccece/Pages/ChildCareAssistance.aspx).

» In the Excel workbook, this data source is partially inventoried in the “KidCare” worksheet.
Supplemental Nutrition Assistance Program

The Supplemental Nutrition Assistance Program (SNAP; formerly known as the Food Stamp Program) provides food assistance to eligible households to cover a portion of a household’s food budget. In addition, the program provides SNAP recipients with nutrition education, employment and training, and in some cases, work experience. Data on SNAP providers is collected and maintained in the Arkansas DHS data warehouse.

REL Southwest confirmed that suitable documentation is not available to inventory the participant data. The Arkansas DHS data warehouse did provide limited variable information, but an examination of the data elements indicated that the data collected pertain to SNAP programs and facilities rather than to families and children. As a result, REL Southwest decided not to include this data source in the inventory. For more information on SNAP, see http://humanservices.arkansas.gov/dco/Pages/SupportServices.aspx.

Federal survey data collections

Federal surveys, such as the Early Childhood Longitudinal Study, Birth Cohort (ECLS-B); the ECLS-Kindergarten Cohort (ECLS-K); and the National Household Education Surveys (NHES), can provide national-level estimates on children’s early care and education experiences. The ECLS-B is a longitudinal cohort study, and the NHES is a repeating cross-sectional survey, but neither can provide estimates at the state level for Arkansas.

References


Request the Arkansas Prekindergarten Data Inventory

This document describes the purpose, contents, and use of the Arkansas PreK Data Inventory. The data inventory is a separate Excel spreadsheet available upon request. Please contact:

Sarah Caverly, Ph.D.
Analytic Technical Support Lead
REL Southwest
800-476-6861, ext. 6509 | scaverly@air.org
Appendix A: Data inventory population protocol guidance on searching documentation manuals

This document provides guidance to REL Southwest team members on methods for extracting information relevant to the Arkansas Prekindergarten (PreK) Data Inventory. The goal of this process is to populate the inventory spreadsheet with information about the datasets—and the variables in those datasets—of potential interest to the Arkansas PreK Research Alliance.

Team members: When the task leader provides a documentation manual or codebook for a data source relevant to the Arkansas PreK Research Alliance, please follow this protocol to ensure maximum information is extracted. These guidelines are to assist with keyword searches in locating helpful information in documentation manuals and are not comprehensive. Please contact the task leader as questions arise that are not covered below. Please note, when reviewing a data source that may be relevant to both the Arkansas PreK and New Mexico PreK Research Alliances (for example, a national-level data source on preK education), make sure you discuss the source with the task leader in advance so that the inventory process is done efficiently.

Step 1: Perform a surface, or light read, of the entire documentation manual or codebook.
Upon receiving the manual/codebook, review it in its entirety to obtain a sense of the information provided.

Step 2: Determine whether the documentation manual/codebook provides a listing of the variables contained in the data source.
⇒ If yes, continue to Step 3.
⇒ If no, contact the task leader to determine a plan for obtaining this information.

Step 3: Extract the list of variables and determine to the extent possible whether each variable is of use to the data inventory.
- If possible, export the variable list into an Excel spreadsheet and create a column that indicates for each variable, whether it is relevant to the data inventory.
- Send this information to the task leader, who will determine the next steps in variable identification.

Step 4: Extract information from each relevant variable.
Upon approval of the list of relevant variables, populate this template:
- There should be one row for each relevant variable.
- You may find it helpful to start by entering the variable name and description (columns G and H), and then completing the remaining columns.
- Highlight cells where information is either not available or additional materials would be required to obtain information. Add explanatory information in the Comments cell.
Step 5: Conduct a search for additional information to ensure that the extracted information is comprehensive.

In addition to the basic variable information that has been extracted, information on such topics as data quality, methods required to produce estimates, ability to link within and among data sources, and definitions of subgroup variables may be of particular use to the Arkansas PreK Research Alliance. Please perform searches on each of the topics listed below and add any relevant information to the data inventory template. This list will be updated and expanded as necessary once work is underway.

a. Complex sample design

Often, and particularly for federal data sources, samples are drawn to be representative of a population of interest. Samples that are not Simple Random Samples (for example, that incorporate sample stratification or clustering) are called complex samples. To produce representative estimates from the sample, the survey design must be taken into account (often using replicate weights or PSU and STRATA). The documentation will provide information on what is needed to incorporate the complex sample design, and that basic information should be noted in the data inventory. It may be especially important for this alliance to capture survey design information that allows for longitudinal analysis. To locate this information in documentation manuals or codebooks, search for the following terms:

- Complex sample
- Complex survey
- Strata
- Stratum
- Cluster
- Weight
- Adjustment
- Jackknife
- BRR
- Taylor series
- Replicate weight
- PSU
- Longitudinal

b. Linking information

Because the Arkansas PreK alliance focuses on issues involving preK education, it may be of particular importance to this alliance to capture as much information on potential data linkages as possible, both over time and across levels and sources. For example, it is important to highlight the potential to link students over time, as well as to link students to preK programs. To ensure that all relevant information is extracted from data documentation manuals and codebooks, beyond the information you extracted by reading the documentation, search for the following terms:

- ID
- Code
- Identification
- Link
- Associated
- Program
- Student
- Longitudinal
- Association
c. **Data quality issues**

Sometimes documentation manuals will provide information on known data quality issues, including excess “missingness.” This type of data quality information will be useful for alliance members and should be captured, as possible, in the data inventory tool. To ensure that relevant information is extracted from data documentation manuals and codebooks, search for the following terms:

- Error
- Errata
- Data anomaly
- Data anomalies
- Accuracy
- Missing
- Unknown
- Not ascertained
- Out of range
- Implausible


d. **Program information**

Because this alliance may need detailed information on preK programs for members’ research, it is important to pay particular attention to any program information available in these data sources (in program-, school-, or student-level files) to obtain maximum utility from the data elements inventoried. To ensure that relevant information is extracted from data documentation manuals and codebooks, search for the following terms:

- Program
- Enrollment
- Financial
- Head Start
- License
- Licensing
- Better Beginnings
- Accredit
- Title 1
## Appendix B: Inventory fields

### Table B.1. Fields Included in the Arkansas Prekindergarten Data Inventory

<table>
<thead>
<tr>
<th>Field name</th>
<th>Field description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID</td>
<td>Randomly assigned ID number that can be used for reference.</td>
</tr>
<tr>
<td>Domain</td>
<td>Unit of variable (for example, student, school).</td>
</tr>
<tr>
<td>System</td>
<td>Data source from which the variable was extracted.</td>
</tr>
<tr>
<td>Entity</td>
<td>The entity to which the variable is relevant. For example, most variables in a student-level file will be relevant to the student, but the ID code for the district the student attends would be a “district” entity and the school code for the school the student attends would be a “school” entity.</td>
</tr>
<tr>
<td>Category</td>
<td>Is the variable an identifier, analysis variable, outcome variable, and so on?</td>
</tr>
<tr>
<td>Restricted or Public?</td>
<td>Is the variable publically available, or is it access restricted?</td>
</tr>
<tr>
<td>Variable</td>
<td>Short description of variable.</td>
</tr>
<tr>
<td>Variable Formal Name</td>
<td>Variable name as defined in dataset.</td>
</tr>
<tr>
<td>Definition</td>
<td>Definition of variable.</td>
</tr>
<tr>
<td>Option Set</td>
<td>Response options (valid for categorical data only).</td>
</tr>
<tr>
<td>Format</td>
<td>Format of variable (for example, 1–20 alpha numeric characters, 4 numeric characters, whole number).</td>
</tr>
<tr>
<td>Source</td>
<td>What entity provided the data, if known (for example, district, state)?</td>
</tr>
<tr>
<td>Variable ID</td>
<td>Variable ID from the data source, if applicable.</td>
</tr>
<tr>
<td>Reference</td>
<td>Reference source from which the information was extracted (for example, name of documentation; website).</td>
</tr>
<tr>
<td>Periodicity</td>
<td>Time(s) of the year that data are collected.</td>
</tr>
<tr>
<td>Reference Point</td>
<td>Are the data tied to a specific reference point, such as “in the last year”?</td>
</tr>
<tr>
<td>Year(s) Available</td>
<td>Data years available.</td>
</tr>
<tr>
<td>Pre-K program?</td>
<td>Yes, No, or Unknown field indicating whether the variable is relevant to a preK program (for example, a program code of association in a student-level file, enrollment information, program information).</td>
</tr>
<tr>
<td>Notable Definitions</td>
<td>Definitions that are likely to be of key interest when determining suitability for future research. For example, if the variable contains an identifier, note in this column whether the identifier can be used for linking purposes and at what levels.</td>
</tr>
<tr>
<td>Known Issues?</td>
<td>Yes, No, or Unknown field indicating whether the variable contains any known data quality issues or excess “missingness.”</td>
</tr>
<tr>
<td>Complex Sample Design</td>
<td>If complex sample design must be accounted for in order to produce representative estimates, note the type of incorporation needed (for example, BRR, Jackknife) in this column.</td>
</tr>
<tr>
<td>Comments</td>
<td>Notes about any relevant information that is not captured in the previous columns.</td>
</tr>
</tbody>
</table>