

October 2026

TOOLKIT

THE EARLY LEARNING JOURNEY

Connecting Preschool Data to
Elementary Outcomes



Center for Education Policy Research
HARVARD UNIVERSITY



STRATEGIC DATA PROJECT

Table of Contents

Introduction	4
» Stage 1: <i>Understand Your Data Landscape</i>	5
» Stage 2: <i>Establish Data Access and Quality</i>	8
» Stage 3: <i>Link Data Across the PreK-2 Continuum</i>	12
» Stage 4: <i>Conduct Credible Analyses</i>	16
» Stage 5: <i>Create Useful Tools For Stakeholders</i>	19
Conclusion	23
Appendix: Tools	24

About the Toolkit

The **Early Learning Journey** is a collaboration between the Strategic Data Project and representatives from Boston Public Schools, the California Department of Education, Chicago Public Schools, the D.C. Office of the State Superintendent of Education, and the Washington State Education Research and Data Center. It is supported by the Gates Foundation.

Longitudinal data helps agencies understand how preschool experiences shape elementary outcomes. The shared goal: ensure that PreK investments translate into meaningful, measurable improvements in children's learning and development over time.

The Strategic Data Project (SDP), an initiative of the Center for Education Policy Research at Harvard University, trained the education leaders on data-driven decision-making in a two-year Strategic Data Project Fellowship. The SDP Fellows worked on data projects for their state and local education agencies and met regularly as a working group to solve problems and codify best practices in the toolkit.

The authors extend their gratitude to the organization above for their generous support.

Contributors to this paper include: Sarah Whitman Blank, Robert Carr, Miriam Greenberg, Alison Guerriero, Paige Kaliski, Lydia Killos, Liz McCrocklin, Kate McGovern, Melissa Nelson, Rachel Tropp, Melinda Wengrin, Yuzhu Xia

View the Full Toolkit at sdp.cepr.harvard.edu/early-learning-journey.

Introduction

Quality pre-kindergarten (PreK) can build a strong foundation for students' later learning. In 2024–25, state-supported preschool programs served nearly 1.8 million students nationally in a variety of settings, including public school classrooms, Head Start programs, and community-based organizations.

Most states and local education agencies (LEAs) collect PreK learning and development data and give kindergarten entry assessments, but few are able to track individual students over time. Due to fragmented data systems, it's difficult to follow students as they move between PreK programs and into later grades. Administrators don't always know how to compare assessments or identify which programs are effective, and families don't know when their children need more support.

This toolkit helps education agencies track longitudinal student outcomes: how children learn and develop as they move through the critical transition from preschool into the early elementary years. It connects information about the same children over time, such as the PreK programs they attended, how they performed at kindergarten entry, and how they progressed through second grade.

Organizations that systematically track PreK–2 longitudinal outcomes gain insights not possible from single-point assessments or program-level reports alone. With longitudinal data, agencies can:

- **Assess effectiveness:** Identify the program features and PreK assessment domains (e.g. literacy, math, social-emotional learning, and executive function) that predict success in later grades.
- **Prevent summer learning loss:** Measure how children's skills change from PreK to kindergarten and identify which populations are most likely to experience the "summer slide."
- **Intervene early:** Monitor achievement gaps and detect which children need additional support before they fall significantly behind, when interventions are most effective.
- **Align PreK and K–2:** Improve alignment between early childhood systems (including PreK assessments, standards, and curricula) and K–2 expectations and instruction.
- **Make strategic investments:** Make evidence-based decisions about PreK investments and target interventions to the populations and communities with the greatest need, particularly those who have historically experienced inequities in early education.

This toolkit provides resources and best practices for collecting and analyzing longitudinal data, and lessons learned from five state and local education agencies. It lays out a five-stage process, with practical data tools and implementation considerations for each stage:

- **Understand the data landscape:** Inventory data sources and align research questions.
- **Establish data access and quality:** Obtain available data and assess its quality.
- **Link data across the PreK–2 continuum:** Match student records across PreK and K–2 systems.
- **Conduct credible analyses:** Select appropriate methods and interpret findings in context.
- **Create useful tools for stakeholders:** Connect findings to critical decisions.

With thoughtful analysis of high-quality longitudinal data, agencies can ensure that all children are meeting developmental and learning milestones throughout their early education journey.

Stage 1: Understand Your Data Landscape

The first step is for agencies to understand what data exists and what constraints will shape analysis. The goal is to conduct a comprehensive landscape assessment that grounds all subsequent work in what is feasible and sets reasonable expectations with stakeholders. Understanding data limitations from the start prevents rework later if sample sizes are too small or identifiers do not align across data sets. Before linking any PreK–2 data, agencies must systematically inventory their data sources, identify focal populations, and align research questions with data availability.

Core Questions

This stage helps agencies answer the following questions:

- **What PreK–2 data sources exist and who controls access to each?** In a mixed-delivery system, children’s data lives in many places: district PreK systems, Head Start, community-based programs, childcare subsidy systems, and K–12 databases. Determine who owns each data source and which policy, legislative, or funding conditions shape what can be accessed.
- **Which student populations are less-represented in our data and why?** Children who move between programs or are served outside state data systems may be missing. Identify coverage gaps, particularly for focal populations, to identify which questions the data can answer.

- **Who are the key stakeholders invested in PreK–2 outcomes and how should they be engaged?** Key stakeholders may include data owners (IT, assessment offices, program offices), policymakers and agency leadership, program administrators (district and community-based), researchers and evaluation teams, and families and educators. Identify them early and understand the decisions they need to make. This shapes which questions are prioritized, which outputs are used, and who is involved in data access and governance decisions.
- **How do assessment tools, timing, and constructs vary across the PreK–2 continuum?** PreK, kindergarten entry, and K–2 assessments rarely work together. Understand how tools differ in what they measure, when they are administered, and how scores are scaled. This determines whether longitudinal comparisons are defensible and if any analytic adjustments are needed.

How to Do It

“Data varies drastically across a mixed-delivery system. Allocate time to understanding the nuances of your data and ensure you talk with stakeholders in each of the different sectors of early childhood education.”

Melissa Nelson, Office of the State Superintendent of Education, Washington, D.C.

- **Create a comprehensive data source inventory** that lists all relevant PreK–2 data sources and documents the data owner, student identifiers used, timing of data collection, population coverage, and access requirements for each source.
- **Interview data owners** across agencies to understand what exists, how it is collected, how it is structured, and who has the authority to share it.
- **Identify focal populations** (e.g., children experiencing poverty, dual language learners, children with disabilities) and test sample sizes using preliminary data to confirm adequate sample sizes for disaggregated analysis.
- **Begin with the questions that matter most to stakeholders**, then assess honestly which are answerable with available data, which would require new data collection, and which need to be refined or set aside. Document the assumptions about data quality, completeness, and coverage that underpin each remaining question.
- **Identify all legal requirements governing data sharing**, including federal and state privacy laws and institutional review board (IRB) protocols, and document the approval processes and decision-makers for each data source. Assess whether existing data-sharing agreements cover the required access, or whether new agreements will be required before work can begin.

Implementation Tips

Common Challenges	How to Address Them
Data from private PreK providers and Head Start programs is inaccessible , leaving many students out of the analysis.	Estimate how many children are affected. If relationships exist with Head Start grantees or private providers, explore voluntary data sharing agreements.
There is no mandatory PreK assessment , and existing measures were often designed for program monitoring.	Document which assessments are in use and what each was designed to measure. Focus analysis on constructs that appear consistently across tools.
Focal population sample sizes are too small for meaningful analysis when data is disaggregated by subgroups.	Test sample sizes with preliminary data. Consider combining years of data, broader geographic aggregation, or acknowledging limitations upfront.
Children may attend multiple programs , move between providers mid-year, or transition across district boundaries.	Examine patterns of student mobility to set realistic expectations about which enrollment patterns may need to be excluded from analysis.
Stakeholders may want to answer questions that existing data cannot support.	Explain which data exist and what requires a new collection. Present multiple research options based on data availability and let stakeholders prioritize.

Boston Public Schools Rethinking the PreK Report Card

Boston Public Schools (BPS) provides PreK programming for around 5,000 children in both district and community settings, including community-based organizations, independent schools, and family child care providers. Leaders wanted a standardized data point to track kindergarten readiness across settings.

To this end, BPS scaled its existing PreK report card to community settings. However, leaders quickly confronted a core Stage 1 question: What are the limitations of this data? Upon examination, teachers were not grading consistently so the results were not comparable across schools.



Yuzhu Xia

Instead of forcing the tool to answer the wrong question, Strategic Data Project Fellow Yuzhu Xia helped BPS reassess the role of the report card within the broader PreK–2 system. The BPS team used family and educator feedback surveys, focus groups, and working group sessions to clarify how the report card was used. As part of this structured engagement effort, Xia developed a **Sample Family & Educator Feedback Survey** to translate stakeholder experiences into actionable evidence for redesign.

“It became clear that understanding the data landscape meant more than cataloging datasets,” says Xia. “It required mapping how tools, expectations and interpretations differed across district- and community-based settings. What appeared to be a shared report card functioned very differently in practice, surfacing the importance of engaging educators, coaches, and families to clarify purpose before defining analytic questions.”

Together, Xia and BPS reframed the report card’s primary function: to support families in understanding learning goals and ease the transition into kindergarten. This clearer, more realistic purpose became a catalyst for new work. BPS began developing a horizontally aligned, family-friendly report card that reflects the district’s core learning goals and can be used consistently in all PreK settings.

Tools and Resources

» [Sample Family and Educator Feedback Survey](#)

The survey helps districts capture how existing sources of data are interpreted and used by families, educators, and instructional staff. It supports informed decisions about whether—and how—existing tools are fit for purpose before they are used for system-level analysis.

» [PreK–2 Assessment Inventory Template](#)

This inventory provides a structured way to document every assessment used across all settings, including district PreK, community-based programs, and Head Start, in one place.

» [Child Trends’ System STEP Forward with Data Framework and Implementation Guide](#)

These resources from Child Trends help agencies identify focal populations, assess data gaps, establish consistent data collection and use practices, and identify action steps toward a more equitable system.

Stage 2: Establish Data Access and Quality

The next step is to obtain the data and to assess its quality. This stage benefits from realistic timelines, standardized data-sharing agreements, and a thoughtful approach to managing missing data. Projects that rush these steps may stall mid-stream or work

with incomplete data. By taking a systematic approach and allowing enough lead time, agencies can access the fullest possible data set and lay the foundation for a credible analysis.

Core Questions

This stage helps agencies answer the following questions:

- **What does it take to establish data-sharing agreements, or are existing agreements sufficient?** Map the timeline, decision-makers, and friction points in the data access process to make a realistic plan and identify where new or updated agreements are required.
- **How accurate and consistent is the data?** Data rarely arrives clean. Assess whether records are accurate, consistently formatted, and that the assessments are administered in ways that support valid comparisons across programs and time points.
- **How complete is the assessment data across PreK–2, and are patterns of missingness random or systematic?** Incomplete data is to be expected. What matters is understanding who is missing and why, so agencies can assess whether findings will generalize to priority populations.

How to Do It

“Connect with your data privacy and security team as well as your internal data stewards early to best understand the policies and rules in place for data sharing in your context, both with internal teams and external stakeholders.”

Melissa Nelson, Office of the State Superintendent of Education, Washington, D.C.

- **Map legal requirements and identify the right decision-makers.** Identify all legal requirements governing data sharing, including the Family Educational Rights and Privacy Act (FERPA), state privacy laws, independent review board (IRB) protocols, and organizational policies. Determine who has authority to approve sharing for each data source and whether existing data sharing agreements cover the access you need, or whether new agreements are required. Build in a three- to six-month buffer for legal review, IT coordination, stakeholder sign-off, and approval.
- **Develop data sharing agreements with clear, specific terms.** Use standardized templates where possible to streamline negotiations across multiple partners. Agreements should specify the requested data elements, the purpose and scope of use, security protocols, and the timeline for data provision. Explain to data owners exactly how the data will be used.
- **Submit targeted data requests and establish secure transfer protocols.** Define exactly what data is needed and in what format before submitting requests. Establish

encryption standards, access controls, and data retention and destruction protocols that comply with your agreements and applicable regulations. Prioritize high-enrollment programs or critical data sources first to build momentum and surface access challenges early.

- **Assess data quality systematically upon receipt.** Profile data as soon as it arrives: check completeness, examine distributions, identify outliers, and compare against known benchmarks such as published enrollment counts or vendor assessment norms. Compare across years to detect inconsistencies in collection or coding practices. Document all quality issues, the decisions made in response, and any caveats that will apply to the analysis.
- **Analyze and document patterns of missingness.** Determine whether missing data is random or concentrated among specific populations, settings, or time periods. A server glitch that drops a few records is different from a pattern in which English language learners are missing from assessment data because of different testing protocols. Document who is missing and why.

Implementation Tips

Common Challenges	How to Address Them
Internal requests can also face delays when data needs to be shared across offices or divisions.	Follow the same process outlined above. Identify an executive sponsor with authority in both divisions.
Data owners are understandably protective of their data.	Build trust by explaining your purpose, involving them in the interpretation of unusual patterns, sharing findings before public release, and acknowledging their contributions publicly.
COVID-19 disrupted data collection from 2020 to 2022.	Exclude specific cohorts, adjust analyses, or acknowledge that some questions cannot be answered for pandemic years.
When data arrives, critical variables are incomplete, miscoded, or collected differently than planned.	Build a data quality review into your timeline immediately after receipt. Flag anything that doesn't match expectations and return to the data owner with specific questions.
Assessment administration practices vary across programs.	Interview program administrators to document variation and decide whether it warrants separate analyses by program type, or exclusion of certain programs or data sources.

DC Office of the State Superintendent of Education

Building Data Partnerships Across Local Education Agencies

The District of Columbia (D.C.) was one of the first U.S. jurisdictions to offer universal PreK for three- and four-year-olds and it has achieved some of the highest participation rates in the nation. OSSE's Fiscal Year 2024 PreK Report notes that 82 percent of three-year-olds and 95 percent of four-year-olds in D.C. were enrolled in publicly funded PreK in the 2023–24 school year.

Strategic Data Project Fellow Melissa Nelson wanted to understand how PreK programs supported students' learning as they transitioned from PreK into early elementary school. As D.C.'s state education agency, the Office of the State Superintendent of Education (OSSE) collects data on PreK program quality. Local education agencies (LEAs) administer interim reading and math assessments to students in kindergarten through second grade (K–2), and those results are not routinely reported to OSSE.



Melissa Nelson

Because OSSE did not routinely receive these data, it invited D.C.'s 37 LEAs offering public PreK to participate in a time-limited data collection. To establish clear expectations, OSSE developed a standard **Data Consent Form** outlining the purpose of the analysis, the process for exchanging data from LEAs to OSSE, and the exact variables requested. A data consent form is different from a data sharing agreement because OSSE is not sharing student-level data, but rather receiving data only. The data consent form was designed to establish a consistent process and document expectations across data received from LEAs and showed that OSSE had considered privacy, security, and appropriate data use.

OSSE invited LEAs to participate and offered to share findings from this project with participating LEAs. Two-thirds of the LEAs expressed interest in this research topic, and four completed the data-submission process during the project period to participate in the project. Though data represents a subset of D.C. students, analysis can provide insight into supports to help students continue to learn and grow across the PreK to elementary school transition. As a next step, the project team is considering opportunities to offer listening sessions for interested and participating LEAs to review initial aggregate-level findings, gather feedback, and build an ongoing data partnership.

Developing the consent form, communicating with the LEAs, and collecting the available data took approximately 10 months. The experience reinforced the importance of allocating enough time to building data partnerships, and the value of customizing a data consent form template rather than drafting one from scratch. Ultimately, the lessons learned were to begin work early and streamline forms to reduce delays.

Tools and Resources

» [Sample Data Consent Form Template](#)

This customizable template, adapted from the D.C. Office of the State Superintendent of Education, reduces the time spent drafting data agreements from scratch.

» [Data Quality Documentation Template](#)

This template provides a structured way for agencies to record data quality issues, decisions, and caveats for each data source used in longitudinal analysis.

» [Data Request Specification Template](#)

This template documents exactly what data is being requested from a data owner, and creates a running record of what was requested, when, and under what terms.

Stage 3: Link Data Across the PreK–2 Continuum

The next step is to match student records across PreK and K–2 systems in order to follow the same child across time and systems. Every decision made at this stage—such as which children are included and how assessments are compared—shapes what the data can and cannot show. The goal is to create longitudinal student records that enable analysis of developmental trajectories, program impacts, and transition effects while being explicit about what linkage decisions mean for interpretation.

Core Questions

This stage helps agencies answer the following questions:

- **What percentage of PreK students can we successfully match to kindergarten records?** Match rates reveal whether the linked dataset is representative of the entire PreK population. Rates that vary by program type or setting may signal that certain groups of children are systematically missing from longitudinal analyses.
- **Which assessment domains show meaningful alignment across PreK–2 tools?** Not all domains can be compared across tools. Identify where alignment exists—or not—to determine whether longitudinal comparisons are defensible and which analytical approaches are appropriate.
- **How does children’s age at assessment vary across the PreK–2 continuum, and how should this affect our analyses?** In early childhood, a few months of age can meaningfully affect developmental scores. Document variation in assessment timing and age at administration to avoid misleading comparisons across programs or times.

How to Do It

“Attempts to link information across the PreK–2 continuum revealed misalignment not only between systems, but between meanings attached to similar indicators. Linkage efforts require shared definitions, intentional sequencing, and cross-department coordination before technical matching can be effective.”

Yuzhu Xia, Boston Public Schools

- **Document identifier systems and determine your matching strategy.** Inventory all student identifiers across PreK and K–12 systems and select the appropriate matching approach: **exact matching** when a universal state ID exists (they are often not issued until kindergarten), **deterministic matching** when multiple ID systems require reconciliation using consistent identifiers (such as name, date of birth, or sometimes address), and **probabilistic matching** when no systematic identifiers are available, using statistical methods to estimate whether records belong to the same individual. Document the approach and its tradeoffs: exact matching produces higher-confidence matches, but other methods capture more students.
- **Implement matching and document match rates.** Run the matching protocol and generate match rates overall and by program type, district, and demographic group. Validate a sample of matches manually. Document the makeup of the unmatched population. Systematic differences between matched and unmatched children affect who the analysis represents.
- **Create enrollment histories and establish inclusion criteria.** Define what counts as a meaningful PreK-to-K trajectory, including how to handle gaps, children attending multiple programs, and mid-year transitions. Excluding students with complex patterns creates a cleaner sample but can systematically leave out the most mobile, most disadvantaged children. Document how inclusion decisions affect sample size and which populations are represented in the data set.
- **Document assessment timing and account for age at administration.** Record exact assessment windows and calculate time elapsed between linked assessments. Developmental scores are sensitive to age at testing. Document variation in children’s ages at assessment and decide in advance how much temporal variation is acceptable for linking records.
- **Assess assessment alignment and choose your approach accordingly.** Map domains and constructs across assessment tools before assuming comparability. Where alignment is genuine, domain-level comparisons may be defensible. Where it is not, analyze measures separately. Document decisions and be explicit with stakeholders about what is and is not being compared.

Implementation Tips

“Even without a universal student ID, data leaders can link records across early childhood and K–12 systems by using common identifiers, like names and dates of birth. Linking records in this manner requires setting clear thresholds for acceptable matches, and investigating if any limitations result when records cannot be matched reliably.”

Robert Carr, California Department of Education

Common Challenges	How to Address Them
High student mobility creates complex enrollment histories. Children move between programs mid-year, switch districts, or experience gaps in enrollment, making it unclear which PreK experience to attribute to their outcomes.	Create enrollment history variables that capture program counts, gaps, and mid-year transitions. Define inclusion criteria and document how different rules affect sample composition (e.g., excluding mobile students can cut out those experiencing housing instability). Test whether findings remain consistent when key inclusion criteria are changed.
Assessment timing varies. PreK assessments often occur between March and June; kindergarten entry between August and October. Gaps between linked assessments can range from 2 to 6 months, shifting development scores.	Document exact assessment windows for all measures and calculate the time elapsed between assessments for each child. Account for age at administration. Define the acceptable level of variation for linking, and clearly flag temporal gaps when reporting findings.
Assessment tools don’t align. PreK observational tools and kindergarten screeners measure different constructs on different scales.	Consider parallel presentation: show PreK growth on PreK measures and K–2 growth on K–2 measures. Where genuine alignment exists at the domain level, cross-tool comparisons may be defensible.
Matching validation surfaces data quality issues that weren’t visible in single-system checks, such as name-spelling variations, date-of-birth errors, duplicate records, and inconsistent demographic fields.	Review a sample of high-confidence matches, borderline matches, and non-matches before finalizing the linked file. Errors often cluster by data source and point to upstream entry issues. Refine your matching algorithm based on what you find.
Inclusion and matching decisions disproportionately exclude focal populations. Students who don’t match are more likely to be children experiencing poverty, children with disabilities, or dual language learners.	Compare the characteristics of included and excluded students before finalizing the sample. Report any systematic differences and consider adjusting inclusion criteria or conducting separate analyses for high-exclusion groups.

California Department of Education

Building a Reliable Link Between Preschool and Elementary School Records

The California Department of Education uses the Desired Results Developmental Profile (DRDP) to assess children on a comprehensive set of developmental measures spanning infancy to kindergarten. This includes students enrolled in the California State Preschool Program (CSPP).

As a Strategic Data Project Fellow, Dr. Robert Carr set out to examine if students' DRDP scores were related to their educational success in elementary school, including third-grade through fifth-grade reading and math achievement. While previous research has demonstrated a positive link between children's early learning skills and later school achievement, this link had not been established for the DRDP assessment system.



Robert Carr

To analyze this, Dr. Carr was able to work with data from two cohorts of CSPP students who received DRDP assessments in 2017–18 or 2018–19 and have now progressed through elementary school. Carr needed to match student records across three distinct data systems: student-level CSPP records, DRDP scores, and school-age assessment scores. He tackled a core Stage 3 question: What percentage of students can be matched across systems?

Dr. Carr worked with his colleagues to develop new procedures to link the records of more than 100,000 students. This represents more than half of the total population of CSPP students who were enrolled in the two cohorts. Student records might not have matched across the three data systems because some students were not required to have a DRDP assessment and/or some students did not continue on to enroll in California Public Schools through fifth grade. Carr used statistical analyses to confirm that matched students were comparable to the total population of enrolled students along various social, demographic, and programmatic characteristics. This indicates that any findings based on the data for the matched students could represent the population of CSPP students overall.

For the first time, CDE leaders can examine the elementary outcomes of students enrolled in the CSPP. Dr. Carr's systematic approach to linking data both produced a robust data set and provided a model for future longitudinal data projects.

Tools and Resources

» [PreK–2 Assessment Crosswalk Matrix](#)

Use this PreK–2 Assessment Crosswalk Matrix to determine what can and cannot be compared over time. This tool allows you to map assessments by developmental domain, document what they measure and how they are administered and scored, and judge whether measures are aligned, partially aligned, or not comparable. The completed crosswalk helps agencies select defensible analytical approaches, identify gaps in measurement, and clearly explain what the data can and cannot show.

Stage 4: Conduct Credible Analyses

Once data is linked, it's ready for analysis. This requires accounting for missing data, selecting appropriate methods, and interpreting results in context. The goal is to produce credible, actionable insights that acknowledge data limitations and provide multiple plausible explanations for findings, especially when they're surprising.

Core Questions

This stage helps agencies answer the following questions:

- **How do outcomes differ between children who attended PreK and those who did not, and does this vary by program type?** Compare children across program types, including private, state-funded, and Head Start, and against peers who did not attend PreK to understand whether public investments reach children who need them most and whether different program models produce different results.
- **Which PreK assessment domains are most predictive of kindergarten and early elementary success?** Not all early skills predict later outcomes equally. Explore which domains matter most, and where current assessments leave gaps, to prioritize what to measure and where to invest in more substantial data collection.
- **Do PreK program quality features relate to kindergarten readiness and K–2 outcomes?** Features like teacher qualifications, curriculum quality, and program dosage may matter as much as attendance alone. Understand which quality factors drive results to make more targeted investments in program improvement.
- **How do developmental trajectories differ across focal populations, and where do gaps emerge or narrow?** Tracking how different groups of children progress across the PreK–2 continuum reveals where disparities take root and where they persist. This is essential for designing interventions that reach the children who need the most support.

- **How much of the variation in K–2 outcomes is explained by PreK experience?** PreK is one of many influences on early elementary outcomes. Identify how much of the variation is attributable to PreK, relative to factors like family demographics or school quality. This helps agencies set realistic expectations about what PreK can and cannot accomplish on its own.

How to Do It

“Use stakeholder feedback to inform analysis. In D.C., stakeholders were clear that they wanted analysis focused on student growth over time, rather than using one assessment score as a point-in-time estimate. This can help guide the analysis chosen so that findings are useful.”

Melissa Nelson, Office of the State Superintendent of Education, Washington, D.C.

- **Start with descriptive analysis before modeling.** Describe the sample, assess data completeness, examine score distributions for outliers, and explore preliminary patterns. Descriptive analysis surfaces surprises and shapes every decision that follows. It helps you understand your data, catch quality issues, and generate hypotheses for modeling.
- **Analyze and document missing data patterns.** Calculate missingness rates by variable, subgroup, and program type, and determine whether patterns are random or concentrated among specific populations. Select an appropriate strategy, document the decision, and be explicit about who the analysis sample represents.
- **Match the analytical method to the research question.** Descriptive comparisons, regression, growth modeling, and quasi-experimental designs each answer different questions. When in doubt, start simple. Plan disaggregated analyses for focal populations early, prepare sensitivity analyses, and consider whether person-centered approaches might communicate findings more clearly than variable-centered ones.
- **Interpret findings in context.** Review results alongside descriptive statistics, consider alternative explanations, and account for assessment limitations, selection effects, and historical disruptions like COVID-19. Engage stakeholders in interpretation, distinguish correlation from causation, and present multiple plausible explanations for surprising findings rather than a single conclusion.

Implementation Tips

Common Challenge	How to Address Them
Selection bias threatens causal interpretations. Children who attend PreK differ from those who do not on many dimensions, and observed differences in outcomes may reflect pre-existing characteristics rather than program effects.	Be explicit about whether the analysis is descriptive or attempting causal inference. For causal questions, consider quasi-experimental designs if the data support them. If not, conduct descriptive analysis and frame findings accordingly.
Lack of independent review. Errors in code, data handling, or interpretation can go undetected. Even accurate findings can be misread: descriptive patterns get treated as causal evidence, and correlations get reported as predictions.	Build verification into the analytic process. A peer review from a colleague, a check-in with a research partner, or simply time away from the work before final review all help catch errors. Label descriptive findings as descriptive. Avoid language like "predicts" or "causes" unless the design supports it.

Washington State Education Research and Data Center Making Complex Data Analysis-Ready

Washington state has several programs geared towards its PreK students from low-income households, including a childcare subsidy for eligible families and a learning assistance program for students who need extra academic support. State leaders wanted to know how participation in both programs influenced kindergarten readiness and third-grade test performance.

Melinda Wengrin led the analysis as a Strategic Data Project Fellow at the Washington State Education Research and Data Center. First, she had to determine whether a complex, multi-source early learning dataset was fit for analysis. She used descriptive analysis to fully understand each of the data sources. The central challenge was assessing data quality, comparability, and representativeness across cohorts and years.



Melinda Wengrin

She used descriptive techniques to surface problems and guide next steps: summary statistics to detect shifts over time; checks of distributions (are they normal or skewed?); cross-tabulations by year, race, gender, and program; subgroup comparisons to detect population differences across programs; and checks for outliers (are extreme values valid or data entry errors?).

This process surfaced changes in student demographics and the provider landscape over time. It also surfaced an unusual jump in scores that was traced to a change in the test scoring methodology, so those scores needed to be rescaled. And it identified small but important gaps in data for Spanish-speaking students and English Language Learners, which had implications for an equity analysis.

Wengrin developed a **Data Dictionary** to document each of these variables and the implications for the analysis. Her thorough descriptive work laid the foundation for a comprehensive and credible analysis, which showed that early support for students from low-income households is correlated with higher assessment scores.

Tools and Resources

» [Sample Data Dictionary](#)

This exemplar provides a variable-by-variable reference for the linked longitudinal dataset used in Washington's PreK–2 analyses. It documents each variable's definition, data origin, and any modifications applied during preparation.

Stage 5: Create Useful Tools for Stakeholders

This stage focuses on turning analytical results into actionable, accessible tools for diverse stakeholders. This may include interactive dashboards for continuous monitoring, research briefs for policymakers, and infographics for broader audiences. Each stakeholder tool should answer a core question: What should someone do differently because of this? The goal is to create tools that connect findings to decisions within stakeholders' control and empower them to act on the results.

Core Questions

This stage helps agencies answer the following questions:

- **Which findings translate into actionable steps for different audiences?** The same finding may mean something different to a state administrator, a district leader, and a program director. Identify what each audience can do in response to a finding.
- **How can we present findings to policymakers in ways that directly inform decisions?** Policymakers need findings framed around the choices in front of them, not methodological detail. Understand what decisions are on the table, and the deadlines for making them.
- **How should we communicate data limitations and assessment alignment challenges without undermining confidence in the findings?** Every dataset has limitations. Be transparent about what the data cannot show, while keeping the focus on what it can.

- **Which PreK–2 metrics should we monitor on an ongoing basis to track system performance?** Not every finding warrants regular monitoring. Identify a focused set of metrics that connect to decisions stakeholders can act on to build sustainable reporting routines.

How to Do It

- **Anticipate assessment skepticism.** Some stakeholders may already question the value, implementation, or validity of assessments, and limited or inconclusive findings can deepen that skepticism. Clearly distinguish limitations of the assessment itself from limitations in administration, data linkage, or analysis. Be transparent about what the evidence does and does not show, while guarding against broader conclusions the analysis cannot support.
- **Decide whether a dashboard is the correct output.** Dashboards require ongoing maintenance. Build one when there is a need for ongoing monitoring, multiple stakeholders will use it regularly, data will be updated consistently, and resources exist to maintain it over time. When those conditions are not met, a well-designed report or slide deck will serve stakeholders better.
- **Create layered outputs for different audiences.** Technical reports with full methodology serve researchers and evaluators. Two- to four-page briefs focused on key findings serve policymakers. Infographics and plain-language summaries reach broader audiences. Tailored slide decks work for presentations where findings need to land quickly. Plan the complete set of outputs from the beginning rather than retrofitting a technical report for broader audiences.
- **Connect findings to decisions stakeholders can make.** For each audience, identify what decisions are within their control and frame findings around those choices. Offer options rather than a single recommendation, consider what is feasible to implement, and suggest pilots or phased approaches where full implementation is uncertain.
- **Build in feedback loops and plan for sustainability.** Plan how findings will be revisited, who will own ongoing communication, and how stakeholder feedback will be incorporated into future analyses. Identify whether analysis and reporting can be embedded into existing agency processes, so it becomes routine rather than a one-time effort. Findings connected to continuous improvement cycles are far more likely to influence decisions over time.

Implementation Tips

Common Challenges	How to Address Them
Different stakeholders need different levels of detail. Legislators want summaries, administrators need operational specifics, and researchers want methodology. A single output rarely serves all three.	Identify two or three key stakeholder groups and prioritize those outputs first. A well-designed two-page brief for policymakers and a separate technical appendix for researchers serve everyone better than one document that tries to do both.
A dashboard exists, but nobody uses it. It shows every available metric, but is too complicated to use without a separate explainer document.	Design for specific user needs rather than showing everything at once. Start with a high-level view and allow users to drill down into detail when they need it. Include enough contextual information that numbers are interpretable without a separate document.
Agencies underestimate dashboard maintenance after launch. Dashboards require regular data updates, staff training, and technical support.	Before building, ask honestly whether the conditions for a sustainable dashboard exist. If not, consider a well-designed static report updated on a regular schedule. If yes, secure a long-term maintenance commitment before development begins, not after.
Findings sometimes reveal uncomfortable truths about inequities, program performance, or whether investments are paying off. Sharing these findings carries political and relational risk.	Anticipate sensitive findings and think through how they will land with each stakeholder group. Frame results around where support is needed rather than where programs have failed. Preview findings with stakeholders and give them time to prepare a response.
Translating findings into actionable recommendations is not straightforward, particularly when the data can describe patterns but cannot prescribe solutions.	Engage stakeholders in the translation process. Offer options with tradeoffs and acknowledge where the data can point a direction but cannot make the choice. Suggest pilots or a phased approach.

Chicago Public Schools

Listening to Stakeholders Before Building a Dashboard

The Chicago Public Schools' Office of Early Childhood Education (OECE) wanted more school and network leaders to prioritize learning in PreK–2. However, early childhood data was split across three assessments and difficult for leaders to access, so OECE aspired to build a dashboard that would unify the data and drive instructional changes to support the youngest learners in the district.

Strategic Data Project Fellow Paige Kaliski planned in-depth stakeholder engagement sessions to design the dashboard. To start, Kaliski interviewed stakeholders across multiple teams and roles, including classroom coaches, professional learning specialists, and OECE leaders. Then she facilitated a group “Dream Data Dashboard” activity, codified in the Stakeholder Engagement Protocol for Dashboard Design.



Paige Kaliski

The session opened with the question: “What is the platonic ideal data dashboard?”, to highlight the connection between effective dashboards and clear use cases. Each participant then spent 10 minutes drawing their “dream data dashboard” if all PreK through 2nd grade data was available. Then each group collaboratively designed a dream dashboard on chart paper and shared their idea with the room, emphasizing why they chose what they did and how it would help them do their job better.

Each participant brought a different perspective to the table about dashboard use cases. Those directly supporting schools were particularly interested in rapid results after assessments and granular data (e.g., “phonological awareness” and not just “literacy”). They also wanted to see multiple data points beyond assessment, such as student attendance, mobility, and interventions.

Ultimately, Kaliski and the team decided that a dashboard was not the right approach to improving PreK–2 data use across the district at that moment. Many of the key data requirements—particularly the granularity and urgency—were not possible with the existing realities.

Instead, the OECE team decided to prioritize three intentional strategies to achieve their goal: building up PreK–2 content knowledge across new audiences, advocating to the state to improve the district’s PreK–2 data availability, and strategically incorporating PreK–2 data into places where similar elementary data already exists—dashboards, presentations, reports, etc.—so that it doesn’t live in an isolated, disconnected place with

limited usage. By engaging in a process to deeply understand user needs, Kaliski set a path forward for an intentional, strategic, and comprehensive approach to enhancing PreK-2 data use in CPS.

Tools and Resources

» [Stakeholder Engagement Protocol for Dashboard Design](#)

This is an example protocol for engaging stakeholders in collaboratively designing a data dashboard. Participants review the data, discuss their favorite dashboards, and draw their “Dream Dashboard.”

Conclusion

Linking data across systems and time will always be inherently messy. Perfection is not the point. The goal is to be deliberate in the choices, clear-eyed about tradeoffs, and explicit in naming what the data can and cannot show. When organizations systematically track PreK–2 longitudinal outcomes, they are better equipped to invest in effective programs, improve early childhood learning experiences, and intervene when students need support. The result is a preschool system that sets all students up to thrive in kindergarten and beyond.

Appendix: Tools Readiness Check

You are ready to move on if you meet the following criteria at each stage.

Stage 1: Understand Your Data Landscape

“We know what data exists, who owns it, and what questions it can answer.”

- We have documented all data sources with owners, identifiers, timing, and access requirements.
- We have tested sample sizes for focal populations and confirmed adequate numbers for analyses.
- Stakeholders are engaged and aligned on priority questions.
- We understand which questions can be answered with existing data and which require new data.
- We have mapped identifier systems and know the matching strategy needed.

Stage 2: Establish Data Access and Quality

“We have a realistic understanding of the data limitations that will shape analysis.”

- We have received all critical data sources identified in Stage 1.
- Data sharing agreements are in place with clear terms for use, security, and reporting.
- We have conducted a comprehensive data quality assessment and documented findings.
- We understand patterns of missingness and their implications for what results will represent.

Stage 3: Link Data Across the PreK–2 Continuum

“We have made deliberate linkage decisions and documented the tradeoffs made.”

- We have successfully matched PreK students to K–2 records with documented match rates.
- We have documented exact assessment timing and calculated intervals between assessments.
- We have made explicit decisions about which enrollment patterns to include vs. exclude.
- We understand what is and isn’t comparable across assessment tools.
- We have validated match quality through manual review.
- We can articulate who our linked sample represents and who is excluded.

Stage 4: Conduct Credible Analysis

“We started with description and let the research model drive question choice.”

- We have completed comprehensive descriptive analyses.
- We have analyzed missing data patterns.
- We have completed planned analyses and interpreted results in context.
- We have engaged stakeholders in the interpretation of those analyses.
- We have documented all decisions.

Stage 5: Create Useful Tools for Stakeholders

“If someone used these findings to make a real decision tomorrow, we would trust how they were interpreted, applied, and explained.”

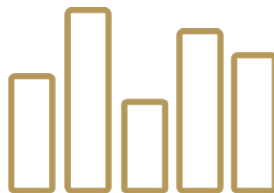
- We have identified which output formats are needed for which audiences.
- We have created outputs connecting findings to actionable decisions.
- We have built in the “so what” explicitly into your tools and outputs for stakeholders.
- We have a maintenance plan if we are creating a dashboard.
- We have communicated limitations honestly.
- We have a dissemination plan for our findings.



Center for Education Policy Research

HARVARD UNIVERSITY

The **Center for Education Policy Research (CEPR)** at Harvard University is committed to transforming education through the power of quality research and evidence. By generating relevant and accessible research, developing technical and organizational capacity to use evidence, and creating a community of schools and systems dedicated to using evidence for progress, CEPR brings together researchers and practitioners to uncover and share the best of what works—and what doesn't—in education.



STRATEGIC **DATA** PROJECT

© 2026 Center for Education Policy Research at Harvard University

sdp.cepr.harvard.edu
sdp@gse.harvard.edu