Uniting our Data to Inform our States

Jessica Cunningham, KYSTATS
Uniting our data to inform our states

States have made a great deal of progress in linking administrative data across agencies, building capacity for data analytics and literacy, and disseminating relevant and timely reports to inform evidence-based policymaking and performance-based programmatic and instructional decisions through funding from the State Longitudinal Data Systems (SLDS) and the Workforce Data Quality Initiative (WDQI) grant programs. Through the statute that created the Institute of Education Sciences (IES), the SLDS Grant Program has awarded competitive, cooperative agreement grants to states since 2005 to 55 states and territories for a total of $826 million dollars to date (Institute of Education Sciences, 2020). Over a decade later, states are at varying levels of linking administrative data ranging from early childhood through the workforce. This discussion will focus on an overview of the Kentucky experience in linking data, its value proposition, and potential for linking other state and federal data sources to inform policy.

Overview: The Kentucky Center for Statistics (KYSTATS)

Kentucky has discovered a great deal of value out of unifying its administrative data to inform the Commonwealth. KYSTATS houses the Kentucky Longitudinal Data System (KLDS) and the Labor Market Information Office (LMI). The KLDS is a centralized, comprehensive state longitudinal data system (SLDS), initially developed in 2009, but ratified into legislation in 2013. Since its inception, the system has expanded a great deal to link data from birth through the workforce. KYSTATS currently includes data from birth records, early childhood, public K-12 students and educators, adult education, public and independent postsecondary institutions, educator certifications, state financial aid, unemployment insurance wages and claims, Workforce Innovation and Opportunity Act (WIOA) training providers, apprenticeships, public assistance programs, driver licensing, corrections, revenue, State Wage Interchange System (SWIS), and death records.

Cross-agency data, including workforce data, are linked using in-house dynamic data matching algorithms based on personal identifiers, inclusive of Social Security number, first name, last name, date of birth, and other pertinent source elements. Confidential information needed for matching is not included in the de-identified data warehouse, instead they are replaced by a global unique identifier (GUID) that allows secure linkage of data across systems for cross-sector reporting by authorized KYSTATS research staff. The alignment of the education to workforce pipeline and additional data elements accrued across the past four years places Kentucky in the position to have sufficient statistical power to examine opportunities and outcomes for subgroups, such as public assistance recipients or justice-involved individuals. However, work is still needed to analyze the largest determinants of inequity as well as examine education to workforce trajectories for these groups across time to identify potential leverage points for intervention.
Because of the breadth of data received, student-level information is available across students’ entire timespan in the education system, allowing for the calculation of graduation rates for K-12 education, exit rates and the time at which a student exits the system, transfers between schools, student achievement growth over time, and drop-outs from the K-12 system. This detail of information allows for the identification of students considered high risk, enabling appropriate interventions and measures for their transitions. While individuals remain in the educational system, the following data are available: course enrollments, teachers of these courses, annual state assessment data, College and/or Career Readiness data, demographic information such as enrollment in the free and reduced price lunch program, and graduation data. These data enable the evaluation of student progress as well as teacher efficacy through the linkages between students and teachers recorded in the KLDS. With the KLDS data system expansion into birth through five, student-level information is also available for determining completion rates in preschool or head start programs, participation in home visitation programs, childcare facility, and facilitating assessment of early learning impact on school readiness for Kindergarten.

Upon completing K-12, the combination of the Kentucky Postsecondary Education Data System (KPEDS) data with the Unemployment Insurance (UI) wage data allows for assessment of transitions from secondary education to postsecondary education or the workforce. Transcript data from KPEDS allows for identification of hours attempted, completed, and earned during all semesters of enrollment for 2- or 4-year in-state public postsecondary institutions. Enrollment data are available at the Independent 4-year institutions as well; however, course grades are not received, limiting our ability to calculate hours earned. Kentucky is currently set up and utilizing the SWIS Data Sharing system to link interstate quarterly wage records with any other state who has also signed the Agreement to satisfy performance accountability requirements under WIOA.

Kentucky’s existing partnership with the Coleridge Initiative enables a more granular linking of state administrative data across state lines to inform policy. Collaboration occurs in the Administrative Data Research Facility (ADRF), a FedRAMP moderate approved environment, and utilizes the five safes framework: safe people, safe products, safe settings, safe data, and safe outputs. Secure data transfer protocol and a common hashing method for personally identifiable information enables cross-state linkage of confidential microdata. Project spaces are restricted and contain the subsets of data needed for the approved research question with all data combination and manipulation occurring in this secure environment. Final outputs undergo a strict review of code and data to ensure that all federal and state confidentiality requirements are met or exceeded prior to export. No microdata are removed from this environment, as requested outputs are pre-aggregated and redacted. KYSTATS staff have attended trainings to facilitate usage of this environment; furthermore, staff have proven their capacity to leverage these resources through the publication of the pilot Multi-state Postsecondary Report, which explores education to workforce flows across four states. KYSTATS has also demonstrated capacity to secure research agreements and collaborate on common data models and methods with crucial partners in adjacent states.
Discussion: Experiential Value Proposition

The same responsiveness to change that has propelled KYSTATS and the KLDS to be a beacon to other states embarking on linking data across the early education through workforce pipeline, is reflected in the changing nature of the analyses with which the P-20 Collaborative, now known as KYSTATS to reflect the broader scope of data included within the KLDS, has been tasked by the Commonwealth. KYSTATS initially began its reporting with a focus on readily identifiable transition points, such as high school graduation to college or career. As the depth and breadth of the information integrated within the KLDS has expanded, the focus and role of KYSTATS has expanded in kind. KYSTATS has proven its utility for informing potential policy cost as well as in evaluating whether a program accomplishes its stated purpose. As a reflection of this, KYSTATS is being tasked as the third party evaluator for governmental policy outcome evaluation (Executive Order 2018-571). The increased scope of analytic responsibilities assigned to KYSTATS, such as policy evaluations and program impact reports, makes the need for KYSTATS to develop a comprehensive understanding of regional, inclusive of in-state and bordering states, equity issues. A regional understanding is critical to our ability to provide actionable and complete information to stakeholders.

With the extensive capacity of the KLDS, KYSTATS is uniquely poised to explore several problems that are broader in scope than our state’s borders. These explorations, code, technical documentation, and products can be used as a basis for regional collaboration and subsequent improvements focused on encapsulating the scope and mechanisms of equity issues. By starting from a single state and drawing in adjacent states to collaborate and expand data products, we ensure the preservation of local knowledge and appropriate handling of the within-state nuance intrinsic to administrative data, ultimately developing common methods and models. Only through a deep understanding of the data generating mechanisms and cross-state collaboration on critical metric definitions and needed data elements can we effectively understand the commonalities and differences in equity issues within and across state boundaries.

With more questions than answers, KYSTATS strives to focus on areas that we are uniquely positioned to address. The types of questions that we consider ourselves extraordinarily placed to tackle are those where the particular confluence of data sources integrated within the KLDS offer a unique problem scope. Additionally, KYSTATS prioritizes research dedicated to exploring and solving problems that will be confronted by all states at some point in SLDS development. When these two areas intertwine, you have the sweet spot of research.

Critical Issues:

Addressing Equity Challenges for Sensitive Subgroups: National foundations are building their vision around equity issues and barriers in education through the workforce for vulnerable, at-risk populations (Buteau & Orensten, 2020). Many states have not developed the capacity to explore these issues comprehensively due to data challenges. With burgeoning partnerships
between KYSTATS and both the Cabinet for Health and Family Services and the Justice and Public Safety Cabinet, Kentucky is poised to investigate these equity issues and barriers. KYSTATS will leverage this knowledge to inform our state as well as to start a national conversation centered on the role of the SLDS in supporting strategic planning to address these issues at the state and federal levels.

*Understanding Postsecondary Value:* The idea of postsecondary return on investment can be conceptualized at various levels, such as from the perspective of the individual pursuing the credential or from the perspective of the state providing financial aid. Not only has measuring the value of postsecondary credentials been a part of state and national conversations, but also the idea of measuring equitable value (Postsecondary Value Commission, 2021). Regardless of the unit of analysis for examining postsecondary value, programmatic and policy decision-making is impeded when the question of return on investment is left out of the equation. The additional understanding KYSTATS is seeking of an equity-lens into barriers to successful postsecondary or career transitions will only serve to amplify the utility of return on investment calculations.

*Leveraging Geocoding for Research and Reporting:* One recent data success for KYSTATS was geocoding addresses to the census tract for all individuals with a known address in the KLDS, opening the opportunity for a more nuanced understanding of location for use in both analysis and reporting. The importance of neighborhoods was highlighted in work by Chetty, Hendren, and Katz (2015), which reviewed the long-term educational and economic impact of the experimental Moving to Opportunity program for participants by age level at time of participation. Despite the known importance of this information, geocoding accurate addresses from data sources to the census tract level and incorporating it into education and workforce reporting is uncharted territory for most states. Both state and national conversations convey the need for exploration of a better school-level poverty measure beyond free and reduced lunch (Snyder & Musu-Gillette, 2015). KYSTATS is currently undertaking the next step - incorporating geocoding into reports in a meaningful and digestible way to inform education and workforce decisions as well as provide a broader understanding of equity issues across the state.

*Common Data Issues – Bridging State Data Silos:* KYSTATS is aware of issues that extend past state borders and can disproportionately affect postsecondary institutions located in and workforce planning for border regions, particularly border regions that are contained within a cross-state metropolitan statistical area (MSA). Kentucky cross-state MSAs have communicated a great need for linking education and workforce data across state lines to inform regional planning. An example of the critical need for regional workforce planning across state lines is the most recent partnership formally signed by leaders from KentuckianaWorks and Southern Indiana Works. These two local workforce areas have created the first of its kind Bi-State Plan for Advancing the Regional Workforce in an effort to be more responsive to the needs of the region’s employers, job seekers and students.
To address this need for incorporating multi-state performance measures, Kentucky began working with the Coleridge Initiative in 2019, formalizing this partnership in 2020, to utilize the Administrative Data Research Facility (ADRF) for secure cross-state collaborative efforts using confidential microdata. Using this environment, Kentucky partners with other states to develop a common understanding of the regional education to workforce pipeline. Individual-level microdata is critical to accurately representing education and workforce outcomes for various programs, particularly for cross-state MSAs. Inclusion of out-of-state outcomes is critical when attempting to understand trajectories for subpopulations over time. Consider employment and earnings for high school graduates seven years post-graduation. Restricting to within state unemployment insurance (UI) covered wages leaves us with a portion of graduates that may be unemployed, employed in state in a non-UI covered job, or employed out-of-state. With each additional state incorporated within the analysis, we become more confident in determining actual employment and real-world wages. Our focus on border states still leaves us with some unknowns when considering specific subgroups with more opportunity to migrate out of the state and region. Notably, individuals tend to have more opportunity and resources for out-state migration when they have a graduate credential in a specialized and in-demand field (Kodrzycki, 2001).

**Potential for Future Data Linkages with Value Proposition**

One of the greatest value proposition to linking employment and wage records with education data for states was replacing the self-reported wage with actual wages and having the capability of examining more robust employment and retention measures. There are limitless possibilities for linking data across state and federal sources to inform policy and programmatic decisions. For example, linking the Survey of Earned Doctorates (SED) with state employment and wage records would provide value to both the state and federal agencies. The SED is an annual census administered to all individuals receiving a research doctorate from an accredited U.S. institution in a given academic year. The SED collects information on the doctoral recipient's educational history, demographic characteristics, and post-graduation plans to assess characteristics of the doctoral population and trends in doctoral education and degrees. From the state side, Kentucky captures approximately 46% of Kentucky postsecondary graduates earning a doctoral degree as being employed in Kentucky the following year, which increases to roughly 60% when employment records from the border states of Ohio, Indiana and Tennessee are linked. Post-graduation plans from the SED survey could help provide additional context to this percentage in addition to provide valuable demographic information for those earning doctoral degrees in other states and migrating to Kentucky for work. Higher credentials tend to correspond with increased potential for mobility, so a survey that addresses this particular group will continue to add value even as we add additional states, especially for doctoral degree earners with other characteristics affiliated with increased mobility, such as male, single, and relatively young (Compton & Pollock, 2007; Corbett M., 2007). Conversely, administration of this survey captures self-reported employment and salary data for one year following graduation; however, linking this doctoral...
data to state employment and wage records could provide insight into actual employment and retention outcomes for doctoral earners opening the potential for postsecondary value studies at the doctoral level.
References:


