Student Data Tracking

Founded in 2005, College Access Foundation of California has provided more than $50 million in need-based scholarships to low-income students through partnerships with over 100 college access programs (grantees) throughout the state. In 2008, the Foundation launched a student data collection project to better understand the composition of its scholarship recipients and to better assess the impact of its scholarship investments. The main objective was to determine whether the Foundation had accomplished its mission: to help motivated, financially needy California students attend – and complete – college.

Today, the student data collection project is an integral part of the Foundation's grant-making strategy. It enables the Foundation and its grantees to gauge progress towards shared goals of college access and completion and plays an important role in building grantee knowledge and expertise in: 1) optimally structuring scholarships to meet unmet financial need and to improve persistence and completion; and 2) assisting students to access public financial aid by leveraging the Foundation's scholarship dollars.

Prior to launching any successful student data collection project, scholarship providers should consider the following important issues:

1. Determine the purpose of the project: What is the purpose of the data collection effort? What outcomes are being measured? What kind of information would be most germane to accomplishing this purpose?
2. Protect student information: This is often the primary concern for both students and grantees and a main reason for their reluctance to share data. Protecting student privacy should drive decisions about the kinds of systems and tools to use, the personnel involved in the project, and protocols for reporting, storing, and sharing student data.
3. Consider the value of matching student information and gaining consent: Providers may have a need to match scholar data with other databases that neither the provider nor the grantees can collect themselves (e.g., National Student Clearinghouse [NSC]). The Foundation's grantees include consent forms as part of the scholarship application process to inform students of this intention, the process, and protection of their personal information.
4. Be selective about data fields: Each data field should be essential to answer strategic questions related to the data collection's purpose. The Foundation is interested in increasing student access and success and understanding the role of its scholarship dollars as part of the students' financial aid package. Thus, the Foundation's main data fields include the student's demographics, high school and postsecondary institution information, and financial aid amounts.

   - If a particular data field is sensitive or burdensome to obtain, consider using proxies. For example, instead of income data, the Foundation asks for expected family contribution (EFC) to measure financial need.

- Existing databases can enrich the dataset without burdening grantees to collect and report on the data. For example, when Foundation's grantees enter a high school name, its county, region and academic performance data are automatically populated from the state department of education's database.

5. Communicate findings: The Foundation holds two grantee conferences each year, which is an opportunity to share aggregate findings about the scholarship students with grantees. Seeing the analysis of the student data and comparing it with relevant benchmarks helps grantees understand and value the information they provide. This will ultimately strengthen the quality of data collected in future rounds and empower grantees to use data to improve their own programming. The Foundation's online database allows grantees to extract and analyze their own data.

Like any major project, there are cost implications. For a project with 5,000 students per year, the annual cost is roughly $35,000 plus the cost of a full-time analyst on staff. This includes the cost of an online database with individual user accounts for each grantee, enrollment information from NSC, and a database consultant.

The data collection period starts during the summer/fall and extends through the following spring term of the academic year on which it is being reported.

Even under the best circumstances, data collection is time consuming and requires a significant commitment to supporting and monitoring grantee participation. But with thoughtful implementation, student-level data can provide timely and comprehensive feedback on the impact of scholarship dollar investments.

Source: Jumin Kim, Research and Data Analyst, and Katie Tremper, Program Officer, College Access Foundation