

IYS Sampling Design

Illinois Youth Survey (IYS) uses a two-stage stratified cluster sampling design with Probability Proportional to grade level enrollment Size (PPS) to select schools. PPS sampling is highly recommended when the size of the sampling unit (i.e., grade-level enrollment) varies greatly between schools.

All Illinois public schools with students enrolled in 8th, 10th, or 12th grades are initially included in the statewide sampling frame. Schools with less than 16 students in 8th, 10th, and 12th grades; schools with no grade configuration; and other special types of schools (e.g., vocational schools, special education schools, department of corrections, or other state-funded public schools) are initially excluded from the sampling frame.

Illinois has a very diverse student distribution across four geographic areas. More than 50% of Illinois students reside in Suburban Chicago counties, 20% live in Other Urban/Suburban counties, about 15% live in Chicago, and about 11% reside in Rural counties. The sampling design ensures that the sample reflects the true student distribution across all geographic areas and grade levels.

The sampling frame is therefore stratified by four geographic areas and three grade levels, resulting in 12 primary sampling units (PSUs). Separate school rosters are generated for each PSU but to make school recruitment easy, the same roster is used for both 10th and 12th grades. Some schools are assigned to more than one PSU, for example, an eligible school from Suburban Chicago with 8th, 10th, and 12th grades would be assigned to the Suburban Chicago 8th grade PSU and the Suburban Chicago 10th/12th grade PSU. When all eligible schools are assigned to one or more appropriate PSUs, schools are selected using PPS sampling. Typical enrollment size varies from 25 students per grade to 1,400 students per grade, according to ISBE Fall Housing Data. Because not all schools that are asked to participate in the IYS state sample agree to do so, schools are oversampled to reach the required number of schools.

After selecting, recruiting, and administering the survey, 50 surveys/students are randomly drawn from the valid surveys within each participating school per grade and geographic area. For those grades/schools in which the number of participating students is less than 50, all students are selected.

This student-level data constitutes the state sample. This sample is then weighted for sampling design and estimated student absenteeism in a given school year. To reduce any selection biases and to align between the sample and the reference population, the state sample is then brought back to the reference population based on race, gender, and income.