

Distance Learning Dataset Training Glossary of Terms

Academic press: Academic press is a construct that encompasses information regarding expectations for student achievement, clear achievement-oriented goals, amount of time spent on instruction, and the amount of homework assigned to students.

Advanced Placement (AP): Advanced Placement (AP) courses are designed to provide opportunities beyond what the high school curriculum traditionally offers. The AP program, which is structured around a curriculum representing introductory college courses, allows students to enroll only in those courses in which they exhibit both ability and interest.

Adjusted Cohort Graduation Rate (ACGR): The Adjusted Cohort Graduation Rate (ACGR) is calculated based on the number of students who graduate in 4 years or less with a regular high school diploma divided by the number of students who form the adjusted cohort for the graduating class.

Averaged Freshman Graduation Rate (AFGR): The Averaged Freshman Graduation Rate (AFGR) provides an estimate of the percentage of public high school students who graduate on time—that is, 4 years after starting 9th grade—with a regular diploma.

Balanced Repeated Replication (BRR): See Replication technique.

Base weight: The product of the reciprocals of the probabilities of inclusion for all stages of sampling.

Bias (due to nonresponse): The difference that occurs when respondents differ as a group from nonrespondents on a characteristic being studied.

Bias (of an estimate): The difference between the expected value of a sample estimate and the corresponding true value for the population.

Bootstrap: See Replication techniques.

CAPI: Computer Assisted Personal Interviewing enables data collection staff to use portable microcomputers to administer a data collection form while viewing the form on the computer screen. As responses are entered directly into the computer, they are used to guide the interview and are automatically checked for specified range, format, and consistency.

Carnegie Units: The Carnegie Unit is a strictly time-based reference for measuring educational attainment used by American universities and colleges; the Carnegie Unit assesses secondary school attainment.

CATI: Computer Assisted Telephone Interviewing uses a computer system that allows a telephone interviewer to administer a data collection form over the phone while viewing the form on a computer screen. As the interviewer enters responses directly into the computer, the responses are used to guide the interview and are automatically checked for specified range, format, and consistency edits.

Distance Learning Dataset Training Glossary of Terms

Chi-squared Automatic Interaction Detector (CHAID) Analysis: This technique divides the respondent data into segments which differ with respect to the item being imputed. This segmentation process first divides the data into groups based on categories of the most significant predictors. It then splits each of these groups into smaller groups based on other predictor variables and merges categories of a variable found insignificant (by χ^2 test). This splitting and merging progress continues until no more statistically significant predictors are found. The imputation classes form the final CHAID segments.

Cohort: A group of individuals who have a statistical factor in common (e.g., year of birth, grade in school, year of high school graduation).

“Cold-deck” imputation: See Imputation.

Community college: typically only refers to 2-year public, where 2-year allows for any type of control.

Composite variables: Variable constructed through the combination of two or more variables (e.g., socioeconomic status) or through calculation by applying a mathematical function to a variable. Composite variable is also referred to as derived, constructed, or classification variable.

Computer Assisted Personal Interviewing: See CAPI.

Computer Assisted Telephone Interviewing: See CATI.

Consistency edit: See Edit.

Coverage error: Coverage error in an estimate results from the omission of part of the target population (undercoverage) or the inclusion of units from outside the target population (overcoverage).

Core courses: Core courses are those that typically satisfy requirements for a diploma (e.g., math, English, social studies, and science), whereas non-core courses do not.

Critical items or key items: Items deemed crucial to the methodological or analytical objectives of the study.

Design effect: The cumulative effect of the various design factors affecting the precision of statistics is often modeled as the sample design effect. The design effect, DEFF, is defined as the ratio of the sampling variance of the statistic (e.g., a mean or a proportion) under the actual sampling design divided by the variance that would be expected for a simple random sample of the same size. Hence, the design effect is equal to one, by definition, for simple random samples. For clustered multistage sampling designs, the design effect is greater than unity, reflecting that the precision is less than could be achieved with a simple random sample of the same size (if that were the sampling design). The size of the design effect depends largely on the intracluster correlation of the survey observations within the primary sampling units. Hence, statistics that are based on observations that are highly correlated within units will have higher design effects.

Distance Learning Dataset Training Glossary of Terms

DU: Housing or dwelling units.

EDFacts: EDFacts is a U. S. Department of Education initiative to put performance data at the center of policy, management and budget decisions for all K-12 educational programs. EDFacts centralizes performance data supplied by K-12 state education agencies (SEAs) with other data assets, such as financial grant information, within the Department to enable better analysis and use in policy development, planning and management.

Edits: Procedures for checking and modifying response in a survey.

Blanking edit: Deletes extraneous entries and assigns the —not answered code to items that should have been answered but were not.

Consistency edit: Identifies inconsistent entries within each record and, whenever possible, corrects them. If they cannot be corrected, the entries are deleted. Inconsistencies can be (1) within items or (2) between items. The consistency edit also fills some items where data are missing or incomplete by using other information on the data record.

Logic edit: Checks made of the data to ensure logical consistency among the responses from a data provider.

Range check: Determines whether responses fall within a predetermined set of acceptable values.

Relational edit check: Compares data entries from one section of the questionnaire for consistency with data entries from another section of the questionnaire.

Skip pattern check: Checks if responses correctly followed skip pattern instructions.

Summation check: Compares reported totals with the sums of the constituent data items.

Education system: Within the International Assessment Program studies, participating entities are referred to as education systems. It is important to note that these definitions are different across IAP studies.

PIRLS and TIMSS: Education systems within PIRLS and TIMSS include countries (complete, independent political entities) and “other education systems” which include a portion of a country, nation, kingdom, or emirate, or are other non-national entities.

PISA: Education systems within PISA include OECD countries, some non-OECD countries, and other sub-national systems including some U.S. States.

Estimate: A numerical value obtained from a statistical sample and assigned to a population parameter. The particular value yielded by an estimator in a given set of circumstances or the rule by which such particular values are calculated.

Distance Learning Dataset Training Glossary of Terms

Estimation: Estimation is concerned with inference about the numerical value of unknown population values from incomplete data, such as a sample. If a single figure is calculated for each unknown parameter, the process is called point estimation. If an interval is calculated within which the parameter is likely, in some sense, to lie, the process is called interval estimation.

Field test: The study of a data collection activity in the setting where it is to be conducted.

“Hot-deck” imputation: See Imputation.

Imputation: Imputation (for item or survey nonresponse) involves supplying a value if an item response is missing. The items may be missing because the respondent was careless, refused to provide an answer, or could not obtain the requested information. Since extensive amounts of missing data can seriously bias sample-based estimates, procedures for imputing missing values are often developed. Imputation is used to reduce nonresponse bias in survey estimates, simplify analyses, and improve the consistency of results across analyses.

Depending on the type of data to be imputed and the extent of missing values, a number of alternative techniques can be employed. These techniques include: logical imputation, the use of poststratum averages, “hot deck” imputation, and regression and other “modeling” techniques.

“Cold-deck” imputation: A process that imputes missing data with values observed from a past survey.

“Hot-deck” imputation: Hot deck refers to a general class of procedures for which cases with missing items are assigned the corresponding value of a “similar” respondent in the sample.

Random within class: This method divides the total sample into imputation classes according to the values of the auxiliary variables. Each nonrespondent is assigned a value randomly selected from the same imputation class.

Sequential (also known as traditional): The records of the survey are treated sequentially in the same imputation class and for each class a single value is stored to provide a starting point for a single pass through the data file. If a record has a response, that value replaces the previous value. If the record is missing, the currently stored value is assigned to that unit.

Logical imputation: Logical imputation can be applied in situations where a missing response can be inferred with certainty (or high degree of probability) from other information in the data record.

Poststratum averages: In the use of poststratum average, a record with missing data is assigned the mean value of those cases in the same “poststratum” for which information on the item is available.

Proc Impute: This is an advanced software package that performs three steps for each target variable to be imputed:

Distance Learning Dataset Training Glossary of Terms

- 1) Uses stepwise regression analysis to find the best combination of predictors among all variables included in the imputation model;
- 2) Creates homogeneous cells of records which have close predicted regression values; and
- 3) Imputes each missing record in a given cell with a weighted average of two donors, one from its own cell and the other from an adjacent cell.

Regression and other modeling techniques: These techniques operate by modeling the variable to be imputed, Y as a function of related independent variables, X_1, X_2, \dots, X_p . To preserve the variability of the Y 's at specific values of X_1, X_2, \dots, X_p , a residual, $\hat{\epsilon}$, is sometimes added to the predicted value determined from the model.

Independent variable: In regression analysis, when a random variable, Y is expressed as a function of variables X_1, X_2, \dots, X_p , plus a stochastic term, the X 's are known as "independent variables."

International Baccalaureate (IB): The International Baccalaureate (IB) program is designed to provide opportunities beyond what the high school curriculum traditionally offers. The IB program combines advanced content knowledge with a focus on the development of critical thinking and an appreciation of global issues. Earning an IB diploma requires a multiyear commitment across disciplines.

Item nonresponse: An item on a data collection form that is missing when a response was expected.

Item Response Theory (IRT): See Scaling.

Jackknife method: See Replication techniques.

Key items or critical items: Items deemed crucial to the methodological or analytical objectives of the study.

Likert scales: A psychometric scale commonly involved in research that employs questionnaires. It is the most widely used approach to scaling responses in survey research, such that the term is often used interchangeably with rating scale, or more accurately the Likert-type scale, even though the two are not synonymous.

Logic edit: See Edit.

Logical imputation: See Imputation.

Measurement error: Measurement error refers to errors in estimates resulting from incorrect responses gathered during the data collection phase of a survey. Measurement errors result, for instance, when the respondent gives (intentionally or unintentionally) incorrect answers, the

Distance Learning Dataset Training Glossary of Terms

interviewer misunderstands or records answers incorrectly, the interviewer influences the responses, the questionnaire is misinterpreted, etc.

Nonresponse: Cases in data collection activities in which potential data providers are contacted but refuse to reply or are unable to do so for reasons such as deafness or illness.

Nonresponse bias: This occurs when respondents as a group differ from nonrespondents in their answers to questions on a data collection form.

Nonsampling error: This term is used to describe variations in the estimates that may be caused by population coverage limitations, as well as data collection, processing, and reporting procedures. The sources of nonsampling errors are typically problems like unit and item nonresponse, the differences in respondents' interpretations of the meaning of the questions, response differences related to the particular time the survey was conducted, and mistakes in data preparation.

Non-core courses: See Core-courses.

Open-ended: A type of interview question that does not limit the potential response to predetermined alternatives.

Ordinary least squares (OLS): The estimator that minimizes the sum of squared residuals.

Out-of-range response: A response that is outside of the predetermined range of values considered acceptable for a particular item.

Oversampling: Deliberately sampling a portion of the population at a higher rate than the remainder of the population.

Plausible value: Proficiency value drawn at random from a conditional distribution of a survey respondent, given his or her response to cognitive exercises and a specified subset of background variables.

Plausible values methodology: Plausible values methodology represents what the true performance of an individual might have been, had it been observed, using a small number of random draws from an empirically derived distribution of score values based on the student's observed responses to assessment items and on background variables. Each random draw from the distribution is considered a representative value from the distribution of potential scale scores for all students in the sample who have similar characteristics and identical patterns of item responses. The draws from the distribution are different from one another to quantify the degree of precision (the width of the spread) in the underlying distribution of possible scale scores that could have caused the observed performances.

Population: All individuals in the group to which conclusions from a data collection activity are to be applied.

Population variance: A measure of dispersion defined as the average of the squared deviations between the observed values of the elements of a population and the corresponding mean of

Distance Learning Dataset Training Glossary of Terms

those values. MS Prefers use of “target population” – do not use “sample population.”

Poststratification: An estimation method that adjusts the sampling weights so that they sum to specified population totals corresponding to the levels of a particular response variable.

Poststratification adjustment: A weight adjustment that forces survey estimates to match independent population totals within selected poststrata (adjustment cells).

Precision: The difference between a sample-based estimate and its expected value. Precision is measured by the sampling error (or standard error) of an estimate.

Pretest: A test to determine performance prior to the administration of a data collection activity.

PSU: Primary sampling unit.

Probability sample: A sample selected by a method such that each unit has a fixed and determined probability of selection.

Proc Impute: See Imputation.

Processing: The manipulation of data.

Range check: See Edit.

Regression analysis: A statistical technique for investigating and modeling the relationship between variables.

Relational edit check: See Edit.

Replicate estimate: An estimate of the population quantity based on the replicate subsample using the same estimation methods used to compute the full sample estimate.

Replicate sample: One of a set of subsamples, each obtained by deleting a number of observations in the original sample for the purpose of computing the appropriate variance based on the complex design of the survey.

Replicate weight: The weight assigned to an observation for a particular replicate subsample.

Replicate: A term often used to refer to either the replicate sample or the replicate estimate, depending on context.

Replication technique: Method of estimating sampling errors that involve repeated estimation of the same statistic using various subsets of data providers. The major methods are balanced repeated replication (BRR), bootstrap, and the jackknife technique.

Balanced Repeated Replication (BRR): A method of replication that divides the sample into half-samples.

Distance Learning Dataset Training Glossary of Terms

Bootstrap: A resampling technique of creating replicates by drawing random samples with replacement that mirror the original sampling plan for a pseudo-population constructed from the original sample.

Jackknife method: A method of replication that creates replicates (subsets) by excluding one unit at a time from the sample.

Rounding convention: typical system for rounding decimals.

Sample: A subgroup selected from the entire population.

Sampling error: When a sample rather than the entire population is surveyed, estimates can differ from the true population values that they represent. This difference, or sampling error, occurs by chance, and its variability is measured by the standard error of the estimate. Sample estimates from a given survey design are unbiased when an average of the estimates from all possible samples would yield, hypothetically, the true population value. In this case, the sample estimate and its standard error can be used to construct approximate confidence intervals, or ranges of values, that include the true population value with known probabilities.

Sampling variance: A measure of dispersion of values of a statistic that would occur if the survey were repeated a large number of times using the same sample design, instrument, and data collection methodology. The square root of the sampling variance is the standard error.

Sampling weights: See Weighted estimates.

Scaling (item response theory): Item response theory (IRT) scaling assumes some uniformity in response patterns when items require similar skills. Such uniformity can be used to characterize both examinees and items in terms of a common scale attached to the skills, even when all examinees do not take identical sets of items. Comparisons of items and examinees can then be made in reference to a scale, rather than to the percent correct. IRT scaling also allows the distributions of examinee groups to be compared.

This is accomplished by modeling the probability of answering a question in a certain ways as a mathematical function of proficiency or skill. The underlying principle of IRT is that, when a number of items require similar skills, the regularities observed across patterns of response can often be used to characterize both respondents and tasks in terms of a relatively small number of variables. When aggregated through appropriate mathematical formulas, these variables capture the dominant features of the data. IRT enables the assessment of a sample of students in a subject area or subarea on a common scale even if different students have been administered different exercises

SSU: Secondary sampling unit, sometimes referred to as segments.

Self-weighting: Each person (or household) in the population has an equal probability of selection and is given the same sampling weight for statistical calculations.

Distance Learning Dataset Training Glossary of Terms

Skip pattern check: See Edit.

Special population: A subset of the total population distinguishable by unique needs, characteristics, or interests (e.g., disadvantaged students, gifted students, physically or mentally handicapped students, vocational education students).

Standard deviation: The most widely used measure of dispersion of a set of values. It is equal to the positive square root of the population variance.

Standard error: The positive square root of the sampling variance. It is a measure of the dispersion of the sampling distribution of a statistic. Standard errors are used to establish confidence intervals for the statistics being analyzed.

Statistically significant: There is a low probability that the result is attributable to chance alone.

Summation check: See Edit.

Suppression: Data that do not meet NCES data quality standards are sometimes suppressed or imputed during data cleaning processes. NCES also suppresses unreliable results in published data or reports. Some small numbers may be suppressed in public data files and in published data or reports for confidentiality reasons. Also see Imputation.

STEM: Science, Technology, Engineering, and Mathematics.

Taylor Series: The Taylor Series variance estimation procedure is a technique to estimate the variances of nonlinear statistics. The procedure takes the first-order Taylor Series approximation of the nonlinear statistic and then substitutes the linear representation into the appropriate variance formula based on the sample design. For stratified multistage surveys such as the National Postsecondary Student Aid Study (NPSAS), the Taylor Series procedure requires analysis strata and analysis replicates defined from the sampling strata and primary sampling units (PSUs) used in the first stage of sampling.

Target population: See Population.

Text register: In linguistics, a register is a variety of a language used for a particular purpose or in a particular social setting (e.g., using features of prescribed grammar, choosing more formal words, refraining from using contractions, etc.).

Time series: A set of ordered observations on a quantitative characteristic of an individual or collective phenomenon taken at different points in time. Usually the observations are successive and equally spaced in time.

Ultimate sampling unit (USU): The smallest unit which is the subject of sample selection. Ultimate sampling units across NCES surveys vary and include school districts, schools, principals, classrooms, teachers, students, households, adults, and children.

Unit nonresponse: The failure of a survey respondent to provide any information.

Distance Learning Dataset Training Glossary of Terms

Variable: A quantity that may assume any one of a set of values.

Variance: See Population variance and Sampling variance.

Weighted estimate: Estimate from a sample survey in which the sample data are weighted (multiplied) by factors reflecting the sample design. The weights (referred to as sampling weights) are typically equal to the reciprocals of the overall selection probabilities, multiplied by a nonresponse or poststratification adjustment.