Glossary

**Acquiescence Bias**: A tendency of survey respondents to “agree” more than they “disagree.” They tend to say “yes” more than they say “no.” This is typically seen with Likert scale questions that ask respondents the extent to which they agree or disagree with something.

**American Fact Finder**: A web portal that provides access to existing data collected by the U.S. Census Bureau.

**Annotated Bibliography**: Provides an organized list of resources pertaining to a given research topic.

**Anonymity**: The absence of identifying information that could link a human subject to his or her responses, opinions, behaviors, etc.

**Assessment of Program Process**: Centers on the extent to which a program’s services are reaching its targets. Program process assessment is synonymous with implementation.

**Assessment of Program Theory**: Refers to the conceptual design of a program. It centers on the way in which a program is supposed to operate in a perfect world. There are three essential elements that comprise a program theory. They include: the program’s impact theory, organizational plan, and service delivery plan.

**Assistant Moderator**: In the context of a focus group, the assistant moderator records the focus group session, takes notes, and observes.

**Attrition**: When research participants drop out of an experiment.

**Bivariate Analysis**: The analysis of two variables simultaneously. This is where you begin to look for relationships – or loose associations – between two variables. Two common bivariate techniques are correlations and cross-tabulations.

**Census**: When all units of analysis of a broader population are included in a study.

**Classic Experiment**: It involves at least two groups: one group gets exposed to the stimuli (i.e., the treatment group) and the other group does not (i.e., the control group). The classic experimental design further requires that the research subjects are randomly selected into the treatment and control groups. Finally, the dependent variable must be measured before the experiment begins and after the experiment concludes (the so-called pre-test post-test).

**Cluster Sampling**: Cluster sampling involves conducting multiple random samples, and it is most appropriate when there is no easy way of obtaining a sampling frame or putting together a sampling frame for your given population.
**Glossary**

**Coding:** The process of making data ready for statistical analysis. Coding involves taking any qualitative label and transforming it into a numerical value.

**Coding Schema:** A way of classifying or categorizing content within the text transcripts of interviews, focus groups, or field observation notes.

**Coefficient:** In an OLS regression equation, it represents the value of each of the explanatory variables relative to the dependent variable. The coefficients are numbers that show both the direction and the strength of the relationship between each explanatory variable and the dependent variable.

**Cohort Study:** Entail data collected at multiple intervals using a specific type of population – one that shares a particular life circumstance during a given time period.

**Composite Measure:** Combining multiple metrics that measure one variable forms a composite measure.

**Computer Assisted Telephone Interviewing (CATI):** A survey delivery system where the interviewer sits in front of a computer, delivers the questionnaire items to a randomly selected telephone number selected by the computer from a broader telephone number database, and records the answers that are immediately put into a database that can be analyzed at that moment.

**Confidence Level:** An indicator of “how sure you can be” that the sample, within the specified margin of error based on the sample size, reflects the population. This indicator is reported as a percentage and the default or standard confidence level is 95 percent.

**Confidentiality:** Keeping the information linked to human subjects protected from any person or persons that are not part of the research study.

**Constant:** It represents the expected value for the dependent variable if all of the explanatory variables are zero – that is, if all of the explanatory variables have no value. The constant is also known as the Y intercept.

**Content Analysis:** With a content analysis, a researcher – or a team of researches – combs through interview, focus group, and field observation transcripts in a systematic way so as to identify meaningful patterns within the text.

**Control Group:** The group that does not receive the stimulus in an experimental design.

**Control Group Matching:** A means of ensuring that the control group matches the treatment group without using random selection.

**Control Variable:** See extraneous variable.

**Controlling For:** Negating the effect of an explanatory variable’s influence on the dependent variable through statistics.

**Convenience Sampling:** A non-probability sampling technique that involves choosing research subjects that happen to cross the researcher’s path. In other words, a researcher might go to a public place, like a supermarket, and try to get people to complete his or her survey. Internet polls on news webpages are classic examples of convenience samples, as anyone who visits the webpage can respond.

**Core Articles:** Articles within a literature that focus explicitly on your research question.
**Correlation:** A measure of association between two variables where a numerical score is given that measures the strength and direction of the correlation. Correlations are scored from negative one to positive one.

**Cost Effectiveness:** Entails estimating the costs of achieving a specific outcome.

**Cost-Benefit Analysis:** Program costs are compared to tangible and intangible program benefits, which are expressed in monetary terms. Results are typically expressed in terms of a benefit-cost ratio, which is equal to the benefits of the program divided by its costs.

**Covert Research:** When a researcher disguises his or her identity and intentions so as to gain access to research subjects.

**Cross Sectional Design:** A survey design where a sample of research subjects is surveyed at one point in time.

**Cross-Tabulation:** A means of presenting the relationship between two variables in table format. These tables present frequency / percentage distributions.

**Curvilinear Relationship:** A relationship between two variables where if the value of one variables increases, the other increases to a point and then decreases. Or, if one variable decreases, the other decreases to a point and then increases. On a graph a curvilinear relationship looks like an inverted “U” or a “U.”

**Data Cleaning:** The process of ensuring that the data has been accurately entered into a spreadsheet per the coding scheme that was developed.

**Debriefing:** Involves asking each pre-test subject about his or her experiences taking the survey. It is essentially an opportunity to ask pre-test subjects why they skipped a question, or why they provided two answers to a question that only called for one.

**Debriefing Session:** The purpose of these debriefing sessions is to explain that deception was used and why it was used in a research study. This provides participants with an opportunity to discuss and hopefully reconcile any possible ramifications or issues that may stem from the deception used within the research.

**Decennial Census:** Mandated by the U.S. Constitution, it is a census of population and housing characteristics that is completed every ten years for each year ending in “0.”

**Deception:** Misrepresenting oneself or the intentions of a research project in order to gain necessary access to research subjects that otherwise would not be cooperative.

**Deduction:** The process of testing theories or hypotheses using quantitative research methods. It is the opposite of induction.

**Deduction to Induction:** A process of mixed methods research whereby deductive methods are first used to test a hypothesis and then inductive methods are used to further investigate why the hypothesis was accepted or rejected.

**Deductive Measurement:** The process of measuring variables using the deductive empirical tools. These include surveys, experimentation, and existing data.

**Deductive Research Question:** A research question where there is an identifiable relationship between and x factor and a y factor whereby x presumably impacts y.

**Dependent Variable:** The variable that a researcher wants to explain.
Glossary

**Descriptive Statistics:** The process of quantitatively describing the main features of your data. Also referred to as “summary statistics” because you are essentially summarizing your information. These types of statistics include the mean, standard deviation, and frequency and percentage distributions.

**Double Barreled Questions:** A double-barreled question expresses more than one thought yet only allows for one answer.

**Ecological Fallacy:** Faulty logic whereby group characteristics are applied to an individual.

**Empirical Tools:** The six methods by which information can be collected social scientifically. These tools include surveys, in-depth interviews, focus groups, experimentation, field observation, and existing data. These tools can be classified as either deductive or inductive. Deductive empirical tools tend to be quantitative in nature – quantitative in the sense that the information collected represents numbers or can be easily converted to numbers. Inductive empirical tools tend to be qualitative in nature – qualitative in the sense that the information collected using interviews, focus groups, and field observation tend to yield narratives and stories as opposed to numbers – or data that can be easily converted to numbers.

**Empiricism:** Observable through the senses.

**Equivalence:** By randomly selecting individuals for inclusion into the treatment and control groups, all of the outside variables that might influence the experimental outcome are theoretically spread equally between the two groups. All variables that cannot be measured, or even identified, that might impact the dependent variable, are spread equally among the treatment and control groups.

**Error Term:** In an OLS regression equation, the error term is comprised of something known as “residuals.” Observed values (i.e., the data collected) for the dependent variable and the explanatory variables are used to construct a regression equation that predicts values for the dependent variable. The predicted values rarely match perfectly with the observed values. These differences between the observed and predicted values are the residuals.

**Evaluation Sponsor:** The stakeholder responsible for getting the “things” that the evaluator needs to conduct the evaluation.

**Exhaustive Option Choices:** Survey option choices where the respondent has a legitimate choice.

**Existing Data:** Sources of data collected by “someone else” that researchers can use to test hypotheses, conduct program evaluations, or conduct policy analyses. Sometimes existing data is referred to as official statistics.

**Experimental Process:** The process of taking x number of research subjects and splitting them up into at least two groups. The treatment group, sometimes referred to as the experimental group, will receive the stimuli and the control group will not. This process measures the independent variable.

**Experimentation:** The process by which research subjects are split into two groups. One group known as the treatment group is exposed to some kind of external stimuli, while the other group, the control group, is not. The two groups are compared to see if the stimuli had any impact.

**External Stimuli:** A necessary precondition for the use of experimentation as an empirical tool.
**Extraneous Variable:** Other independent variables that help explain the dependent variable. Also referred to as control variables.

**Face-to-Face Survey:** A survey delivery method where the instrument is administered face-to-face.

**Field Observation:** The process by which researchers document what they observe in a given research setting. There are two primary types of field observations – participant observations and non-participant observations. With participant observations, researchers immerse themselves into an environment, documenting what they see or hear. They are part of the group. Non-participant observers keep their distance and observe in a detached fashion. They are not part of the group.

**Focus Groups:** Focus groups are small groups, usually consisting of six to ten participants. Much like in-depth interviews, focus groups rely on open-ended questions and broad topics to drive a dialogue. The purpose is to establish a dialogue about a topic.

**F-Statistic:** The statistic used to determine if an OLS regression model as a whole is statistically significant.

**Generalizability:** Because probability sampling relies on mathematics to select which units of analysis will be included in your study, what is learned about the sample can be applied to the larger population. This is called generalizability.

**Hard Refusal:** A non-response for a telephone survey marked by a stern refusal.

**Heterogeneous Population:** A diverse population demographically – e.g., income, education, race and ethnicity, age, etc…

**History:** History refers to an event that occurs between the pre-test and post-test that could impact the dependent variable.

**Homoscedasticity:** The concept of homoscedasticity is based on the premise that the residuals have relatively equal variance. In other words, the variance of the error term is constant.

**Hypothesis:** A tentative answer – or educated guess – that is paired with a deductive research question. A hypothesis answers the research question before the research takes place. Carrying out the research either confirms or rejects the researcher’s hypothesis.

**Hypothesis Testing:** See deduction.

**Independent Variable:** The primary explanatory variable that is embodied within a deductive research question. Also, it is any variable that impacts a dependent variable.

**In-Depth Interviews:** The process by which broad and open-ended questions are presented to research subjects.
Index: A composite measure where points are assigned to behaviors and/or feelings that represent the underlying concept generally speaking.

Induction: The process of generating theories or hypotheses using qualitative research methods. It is the opposite of deduction.

Induction to Deduction: A process of mixed methods research whereby inductive methods are first used to generate a hypothesis and then deductive methods are used to test that same hypothesis.

Inferences: When assumptions are made about a broader population based on a random sampling of that population.

Inferential Statistics: The process of drawing conclusions or making inferences about a sample that represents a larger population. Regression analysis is one example of, and a commonly used, inferential statistical technique.

Informed Consent: When obtaining informed consent, an evaluator must: (1) Explain the nature of the program evaluation, what it entails, and what the potential implications may be; (2) Inform human subjects that their participation is completely voluntary (At least from the point of view of the evaluator. In some instances, workers may be required by their superiors to participate); (3) Inform human subjects that every effort will be made to maintain the confidentiality of all information collected.

Informed Consent: The process of getting permission from each and every research subject that participates in any study. In addition to getting permission from the subject, the researcher should inform each research subject of the following: the nature of the research; all known risks of participating; the potential benefits of participating; that participation is completely voluntary; that the research subject may quit at any point; and that all information obtained that can linked to a specific research subject will remain confidential.

Inputs: Refer to a program’s resources (such as money, staff, supplies, buildings, and the like).

Instrumentation: Refers to potential changes with how the dependent variable is measured over time. Measures of variables do change as new factors are incorporated or better measures are developed through a trial and error process.

Internal Validity: The concept of internal validity refers to the degree of “certainty” that there is a relationship between the independent and dependent variables. Greater internal validity implies a greater likelihood that the observed (or unobserved) relationship between the independent and dependent variable is genuine.

Internet Survey: A survey delivery method where a survey is administered using the Internet.

Interrupted Time Series: This quasi-experimental technique involves examining a period of time before a policy was implemented and a period of time after that same policy was implemented. Then, you compare outcome indicators of the two time periods.

Interval Measure: Are rank ordered numerical measures that have an unnatural zero (unnatural meaning arbitrarily constructed).

Interview Protocol: An interview guide that lists the questions to be asked, as well as likely follow-up questions.
**Interviews:** The process by which broad and open-ended questions are presented to research subjects.

**IRBs (Institutional Review Boards):** Are chiefly responsible within U.S. colleges and universities for ensuring that research protocols are in place to protect human subjects. Any individual affiliated with a college or university that is conducting research using human subjects must have their research protocol reviewed by the IRB. The IRB process, rules, and regulations that apply to human subjects researchers are derived from the Protection of Human Subjects provision in the Code of Federal Regulations.

**Key Informants:** Individuals who have an intimate knowledge regarding the needs of the targets.

**Known Risks:** The potential burdens and/or hazards of participating in a given research study.

**Levels of Measurement:** The four levels of measurement include nominal, ordinal, interval, and ratio measures.

**Likert Scale:** An ordinal measure that is marked by the use of strongly agree, agree, disagree, and strongly disagree – or some variation thereof.

**Linearity:** A mathematical relationship represented graphically as a straight line.

**Literature Review:** A summary and synthesis of scholar works pertaining to a given research question.

**Logic Model:** A conceptual representation of a program’s theory that includes information regarding a program’s resources, the services it provides, and the outcomes it hopes to produce.

A logic model is organized in terms of program inputs, activities, outputs, and outcomes.

**Logistic Regression:** A multivariate technique that should be used when the dependent variable is an ordinal or nominal level measure. It is similar to probit regression.

**Longitudinal Design:** A survey design where a sample of research subjects is surveyed at multiple points in time.

**Mail Survey:** A survey delivery method where a survey is administered using mail.

**Margin of Error / Confidence Interval:** The estimated difference between the sample and the population, which is based on the sample size – NOT the number sampled. As a rule, the larger the sample size, the lower the margin of error.

**Maturation:** Refers to any changes to participants over a certain period of time. These changes can be physical, psychological, cognitive, behavioral, etc.

**Mean:** The arithmetic average of a given variable. To calculate this, you take all values for a given variable, add them together, and then divide by the total number of values.

**Mean:** The arithmetic average of a given variable.

**Measurement Validity:** The extent to which the operational definition of a variable is reflected by the measure.

**Measures of Central Tendency:** Univariate techniques that include mean, median, and mode.

**Measures of Dispersion:** Univariate techniques that include range and standard deviation.
Glossary

**Median**: The mid-point of all values for a given variable.

**Milgram Experiments**: Arguably the most infamous research study that brought research ethics to the forefront, Stanley Milgram was a Yale University social-psychologist who wanted to study obedience to authority.

**Mixed Methods**: The simultaneous use of deductive and inductive research methods.

**Mode**: The value that occurs most frequently for a given variable.

**Moderator**: The person responsible for running as focus group session.

**Multicollinearity**: Occurs when two or more explanatory variables are too strongly correlated.

**Multiple Regression Analysis**: An inferential statistical method that allows one to observe the relationship between a dependent variable and multiple independent variables. It allows inferences to drawn and predictions to be made about a larger population stemming from the results of a randomly selected subset of that population – otherwise known as a sample.

**Multivariate Analysis**: The analysis of many variables simultaneously.

**Mutually Exclusive Option Choices**: Mutually exclusive means that the respondent should have only one option choice.

**Natural Experiment**: The natural experiment is one where the researcher does not plan the manipulation of the independent variable. Some event beyond the control of the researcher occurs which provides an opportunity for the research to examine the impact of an event on a specific outcome (i.e. the IV = the time period before the event and the time period after the event and the outcome = the DV). Disaster scenarios typically provide opportunities for natural experiments.

**Needs Assessment**: When conducting a needs assessment, an evaluator tries to determine if, and to what extent, a social condition or problem exists. In other words, is there a “need” for a program? Conducting a needs assessment is a fundamental step in developing a new program or restructuring an existing one.

**Negative Linear Correlation**: Negatively correlated variables have an inverse association. An inverse association means that as one variable increases in value, the other decreases – or vice versa.

**Negative Linear Relationship**: A relationship between two variables where if the value of one variable increases, the other decreases. Or, if the value of one variable decreases, the other increases. Also known as an inverse relationship.

**Nominal Measure**: A measure that is a name or label that has no rank order.

**Non-Equivalent Two Group Design**: This quasi-experimental design is ideal for studies where you have intact groups. School classes are good examples of this. Non-equivalent, two group experiments are exactly like the classic experiment but with one difference – participants are not randomly selected into the treatment and control groups.

**Non-Probability Sampling**: The selection of participants or units to be included in one’s study is not done mathematically. Types of non-probability sampling methods include: convenience sampling, quota sampling, snowball sampling, and purposive sampling.
Non-Response Bias: Nonresponse bias is based upon the understanding that the opinions of non-respondents differ, to some measure, from those who have completed the survey.

Normality: A normal distribution of anything takes on the classic bell curve shape.

Null Hypothesis: The default position that there is no relationship between two variables.

Number Sampled: The number of units of analysis within a sampling frame that have been chosen for inclusion on the study.

Official Statistics: A term synonymous with existing data.

OLS Regression: A multivariate technique that allows for the analysis of multiple explanatory variables on a single dependent variable.

Operationalize: To make measureable.

Ordinal Measure: Are names or labels where there is rank ordering.

Outcomes: Refer to the anticipated short-term and long-term impacts of a program.

Outlier: A data point – or observation – that falls well outside the other data points.

Outputs: Refer to the “amount” of services delivered. Outputs are workload measures.

Panel Study: Panel studies entail the use of multiple cross sectional surveys using the same sample of individuals. The same people are used each time.

Peer Reviewed Articles: Articles that follow the social scientific principles of research, and have been evaluated for their quality by academics that have published scholarly papers and/or books that deal with the topic. Peer-reviewed is sometimes called “refereed.”

Periphery Articles: Articles that pertain to the “y” portion of the deductive research question.

Plagiarism: Taking credit for someone else’s ideas, thoughts, language, or work.

Policy Sciences: Throughout the late 1950s and early 1960s, the policy sciences emerged as a means of studying and addressing some of the most pressing societal problems through the use of highly quantitative and quasi-scientific approaches to social problem solving. Examples of these approaches include operations research and planning programming budgeting systems (PPBS).

Policymakers: Typically represent elected officials or high-level governmental appointees who determine whether a program is created.

Population: A “collection” or theoretical universe of whatever your unit of analysis is.

Positive Linear Correlation: A positive correlation implies that as one variable increases in value, the other variable increases in value. A positive correlation also means that as one variable decreases in value, the other decreases in value.

Positive Linear Relationship: A relationship between two variables where if the value of one variable increases, the other increases. Alternatively, it is a relationship between two variables where if the value of one variable decreases, the other decreases.

Glossary
Glossary

Post-Test: Measurement of the dependent variable after the experimental process.

Pre-Interview Preparation: In the context of telephone survey research, it entails arranging for a controlled environment with which to conduct telephone interviews and rehearsal of the survey instrument.

Pre-Test: Measurement of the dependent variable before the experimental process.

Pre-Test/Post-Test One Group Design: This quasi-experimental design is frequently used for program evaluations where program participants go through the program and outcomes are measures when they enter the program and at some point in time after they finish the program.

Probit Regression: A multivariate technique that should be used when the dependent variable is an ordinal or nominal level measure. It is similar to logistic regression.

Program Efficiency Assessment: Centers on whether the money and resources put toward a program are “well spent.” The two primary approaches to determining whether money and resources are well spent include cost-benefit analysis and cost-effectiveness analysis.

Program Impact Assessment: Centers on determining if a program has met its goals and objectives. Impact assessments are conducted using some form of experimental design.

Program Managers and Practitioners: Individuals responsible for directing or supervising some aspect of the program’s day-to-day operations, while practitioners implement policies and administer a program’s services.

Program Sponsors: Individuals or entities that are responsible for a program’s funding.

Program Targets: Those individuals that ideally benefit from the program’s services, or are eligible to receive the program’s services.

Purposive Sampling: A non-probability sampling technique where the researcher personally (and with purpose) selects which units of analysis will be included in the research. This is used primarily in instances where there are a limited number of people that have expertise in a given research area. Purposive sampling is sometimes referred to as “self-selection.”

P-Value: The p-value evaluates the statistical significance of each variable coefficient.

Qualitative Methods: Inductive research methods – i.e., in-depth interviews, focus groups, and field observation.

Qualitative Research: Research that is inductive in nature, relying primarily on in-depth interviews, focus groups, and/or field observation.

Qualitative Research Question: While deductive research questions are structured more rigidly with a clearly defined relationship in mind – that is, a factor “x” presumably impacts another factor “y” – inductive research questions tend to be more organic in their structure.

Quantitative Research: Research that is deductive in nature, relying primarily on surveys, experimentation and/or existing data.

Quasi-Experiment: An experimental design that does not follow the tenets of the classic experiment.

Quota Sampling: A non-probability sampling technique where a researcher tries to choose subjects that fit certain socio-demographic characteristics.
Random Digit Dialing (RDD): A means of randomly selecting telephone numbers to facilitate a telephone survey.

Random Probability Sampling: Units of analysis are chosen for inclusion mathematically. Mathematical means that each unit of analysis within a broader population has the same mathematical chance of being chosen – or sampled – as every other unit of analysis within the same population. Types of random probability sampling methods include: simple random sampling (SRS), systematic sampling, stratified sampling, and cluster sampling.

Random Selection: Using mathematics to determine which research subjects are chosen to be part of the treatment group or the control group.

Range: The range represents the difference between a variable’s highest value and its lowest value.

Ratio Measure: Numerical measures that have a natural zero.

Recall Bias: An inability of survey respondents to recall events altogether or recall events accurately and objectively.

Reduced Channel Capacity: The assumption that an individual’s cognitive abilities may be somewhat compromised given that survey questions and answer choices are presented through via telephone.

Refereed Articles: See peer-reviewed articles.

Regression Model: The regression equation components – i.e., the dependent, independent, extraneous variables, the error term, and the Y intercept (or constant) – are referred to as the regression “model.”

Regression to the Mean: The premise that individuals that “score” extremely low on a pre-test are more likely to score higher on the post-test, irrespective of the external stimuli. Conversely, those that score very high on the pre-test are much more likely to score lower on the post-test.

Reliability: The ability to replicate research. Or, the extent to which a method of measurement is repeatable.

Representativeness: The ability to generalize research findings to a broader population from a smaller subset of that population (i.e., a sample).

“Researchable” Research Question: See deductive research question.

Research Randomizer: A random numbers generator that can be used for random selection.

Response Fatigue: The degree and rapidity with which a survey respondent tires.

Response Rate: It is a comparison of the number selected to complete a survey versus those that actually complete the survey. The number of individuals that complete the survey divided by the number of individuals selected to complete the survey multiplied by 100 gives the response rate percentage.

R-Squared: The R-squared measures the predictive power of the model. The value of the R-squared ranges from zero to one. The higher the value, the better the predictive power of the model. The higher the R-squared the better the model is.

Sample: A smaller representation of a larger population.

Sample Size: Represents the number of units that are actually “part of the study.” Part of the
study means this: data were collected to measure each variable pertaining to a chosen UoA.

**Sampling Frame:** It is a list of units of analysis that represent your population. It is from this list that a random sample is drawn.

**Satisficing:** Occurs when a respondent answers survey questions without “expending substantial effort.” Respondents that are satisficing mechanically complete the task with as little thinking as possible.

**Scale:** A composite measure where the dimensions allow for a logical progression that shows the intensity of a concept.

**Scarcity Principle:** Informing reluctant respondents that they have been carefully chosen and that their views represent thousands of people in order to persuade them to complete the survey.

**Scatterplot:** A way of graphically showing the relationship between two variables. Each dot or data point represents one variable’s position along the x axis relative the second variables position along the y axis.

**Secondary Articles:** Articles within a literature that focus implicitly on your research question.

**Section Bias:** When there is an error the selection of individuals to participate in a research study, the consequences of which may include invalid or skewed results.

**Simple Random Sampling:** A random probability sampling technique that involves choosing “x” number units of analysis from a sampling frame using a random numbers generator or a random numbers table.

**Snowball Sampling:** A non-probability sampling technique where the researcher recruits participants through a referral system.

**Social Desirability Bias:** The premise of social desirability bias is that individuals will answer survey questions in a manner they think will be viewed favorably by others. Social desirability bias is particularly harmful for research topics that deal with deviant behavior or criminal behavior.

**Soft Refusal:** A soft refusal can be characterized as someone who refuses in a relatively polite or semi-polite manner, usually citing a lack of time or interest. These individuals are characterized as soft refusals simply because there is the possibility that they could be persuaded to participate at a later date.

**Stakeholders:** Individuals that have an “interest” in how well a program performs.

**Standard Deviation:** The degree to which a variable’s values are clustered around the mean. A high standard deviation indicates that the values are less clustered around the mean. There is more data variability, more data dispersion relative to the mean if a variable has a high standard deviation. Conversely, a low standard deviation indicates that a variable’s values are clustered more tightly along mean. A low standard deviation implies lower data variability.

**Stakeholders:** A univariate technique that measures the degree to which a variable’s values are clustered around the mean.

**Statistical Significance:** A statistically significant result does not occur by chance – i.e., a statistically significant result is one where the researcher can be reasonably sure (within some acceptable margin of error) that the result is
accurate. For example, if a sample of research methods students shows that there is “statistically significant” positive relationship between attendance and class performance, then we can be fairly certain, within some acceptable margin of error, that this result is trustworthy. Results that are not-statistically significant are given little credibility in social science research.

**Stratified Sampling:** A random probability sampling technique that should be used when there is an under-represented group within a given population. This technique requires that you split the sampling frame into two mutual exclusive groups and then draw two random samples.

**Survey Pre-Test:** A survey pre-test involves taking a small, non-random but representative subset of your target population and having those individuals complete your newly crafted survey questions. The purpose is to identify any problems with the survey instrument.

**Survey Question Bias:** Any word or phrase within the body or option choices of a survey question that might influence a respondent to choose one answer over another.

**Surveys:** The process by which closed ended questions, with either predetermined option choices or a finite number of possible choices, are presented to research subjects.

**SWOT Analysis:** A structured method used to evaluate the strengths, weaknesses, opportunities and threats of a program.

**Systematic Sampling:** A random probability sampling technique that involves choosing every “kth” unit of analysis from a sampling frame.

**Telephone Survey:** A survey delivery method where a survey is administered using the telephone.

**Testing Effect:** Refers to changes in the subject over time simply because the subject is being tested.

**Treatment Group:** The group that receives the stimulus in an experimental design.

**Trend Study:** Consist of multiple cross sectional surveys whereby the data collected for each survey uses the same variables but a different sample of survey respondents is used each time. The multiple samples represent the same larger population. Trend studies are used to identify changes to a population over time.

**T-Statistic:** Measures the likelihood that the actual value of the coefficient is not zero. The larger the t-statistic’s value, the lower the likelihood that the coefficient is actually zero. The lower bound threshold for assuming that the value of a coefficient is not zero is + or – 1.96.

**Tuskegee Experiments:** Conducted by the U.S. government in the early 1930s, the experiment wanted to test the impacts of untreated syphilis. There was a treatment group – which consisted of 399 African-American males who from Tuskegee, Alabama. These males had advanced syphilis. The control group consisted of 201 uninfected males. Even though the infected males were told they would be treated, they were not. The purpose of the study, unbeknownst to the infected treatment group participants, was to observe the effects of the disease pre and postmortem.

**Type I Error:** Incorrectly rejecting the null hypothesis. In other words, it is concluding that
Glossary

A relationship between two variables exists when in fact it does not.

Type II Error: Incorrectly accepting the null hypothesis. In other words, it is concluding that a relationship between two variables does not exist when in fact it does.

U.S. Census Bureau’s Quick Facts: An existing data source that provides easy access to population, housing, socio-economic, geographic and business data for the U.S. as a whole, states, counties, and municipalities.


Unit of Analysis: It is the “who” or “what” of the research project. It is the entity that you are studying. This entity can range from individuals, to collections of individuals, to organizations, to geographic locations.

Univariate Analysis: The analysis of one variable at a time.

Validity: The extent to which a method of measurement actually measures what it is designed to measure.

Variable: Any factor that takes on a varying characteristic.

Variable Definition: The conceptual basis of what a variable means.