Spring 2018

Glossary of Business Evidence

Paul C. Boyd
Johnson & Wales University - Providence, Paul.Boyd@jwu.edu

Glossary of Business Evidence

Paul C. Boyd, Ph.D.

These definitions are presented to assist with the practical interpretation and application of the concepts (and not, necessarily, the strict statistical explanation). That is, they are intended to assist decision makers with the useful interpretation of the information provided without becoming entangled in the specific details of methodology or procedures.

Italicized terms have separate entries in this glossary.

Please note that this is a work-in-progress. Any and all edits, revisions, corrections and/or additions are most welcome @ Dr.Boyd@research-advisors.com.

Alpha (α) – For practical purposes, the acceptable *Probability* of making a mistake and concluding that a *Hypothesis* (alternate hypothesis) is true, when, in fact, it is not. Alpha is also known as *Statistical Significance*. As is evident, Alpha should be quite low; a small chance of making such a mistake.

Alpha (α) is the acceptable risk of making a *Type I Error* when drawing conclusions from research. That is, it is the probability of concluding that a relationship (or difference) exists, when, in fact, it does not due to random sampling error).

More formally, alpha is the acceptable probability of incorrectly rejecting a true *Null Hypothesis*.

Many statistical procedures promote that the acceptable alpha (α) be determined in advance. That is, researchers must determine an acceptable level of this risk prior to conducting the statistical analysis.

Then, if the *P-Value* associated with the *Test Statistic* is higher than the acceptable alpha, the null hypothesis cannot be rejected and the (alternative) hypothesis cannot be determined to be true. But, if the p-value is less than the acceptable alpha, then there is sufficient evidence to conclude that the (alternative) hypothesis is likely to be true.

In business research, the typical level of alpha that is generally deemed acceptable is .05. However, for more precise use, the acceptable level of alpha should be determined based upon the consequences of making a Type I error. If the potential consequences are dire, then the acceptable alpha should be quite small.

Alternative Hypothesis – See *Hypothesis*

Ambiguous (question) – An ambiguous question on a questionnaire is one that is written in such a way that the respondent is uncertain as to what their response communicates.

For example, a respondent may have a hard time answering a question that asks respondents to agree or disagree with the statement "I frequently shop at Target" since no hint as to what "frequently" means. In this case, the respondent would be uncertain if a negative response communicated that they didn't shop at Target at all, shopped at Target less than once a week, shopped at Target less than they shopped elsewhere, etc.

Analysis of Covariance (ANCOVA) – An *Analysis of Variance* (ANOVA) procedure that adjusts or *Controls* for differences in a third variable (called a *Covariate*) while assessing the relationship between the *Independent Variable* (the *Factor*) and the *Dependent Variable* (the *Response Variable*).

The covariate is typically a pre-existing condition that is measured either through a *Pre-Test* or by an *Extraneous* or *Background Variable*.

Analysis of Variance (ANOVA) – A statistical test to determine if there is evidence to support a *Hypothesis* that there is a relationship between a *Categorical* (*Nominal* or *Ordinal*) *Independent Variable* and a *Metric Dependent Variable*. Typically, the independent variable has more than two possible values (called "levels" or "treatments"). The *Means* of the various levels are compared to determine if at least two are different.

For example, ANOVA would be used to determine if there was a relationship between season (winter, spring, summer, fall) and absenteeism (measured in the number of missed days); if the mean number of absent days was different between any two of the four seasons.

The ANOVA statistic is the *F-ratio*. The probability (*P-Value*) associated with the F-ratio is compared to a previously determined acceptable level of *Alpha* (α) to determine whether or not there is evidence to support the hypothesis.

Note: ANOVA will determine whether a relationship (likely) exists, but it will not reveal the nature of the relationship. Other statistical tests (e.g., Sheffé's test or basic *T-Tests* between all of the groups) would be needed for conclusions about the differences between specific groups (although often a visual examination of the data – e.g., a comparative set of *Box Plots* – is often sufficient for drawing these conclusions).

A Posteriori – A statement based upon the analysis of data (and not a testable *Hypothesis*); inductive reasoning. Contrasts with *A Priori*.

Caution is warranted when using a posteriori approaches due to the tendency to try to fit the reasoning to the (random) data, rather than seeing if the data supports the reasoning. See, for example, *Spurious Relationship*, *Fishing Expedition*, *Data Mining, P-hacking* and *Post Hoc* assessments.

A Priori – A statement (i.e. a *hypothesis*) based upon theory or logic that is postulated prior to the collection and analysis of data; deductive reasoning. Contrasts with *A Posteriori*.

Applied Research – Research designed to address specific problems/decisions rather than to expand or refine a general theory. See *Basic Research*.

Argument – A set of reasons (*Premi*ses) posited as evidence for a conclusion. See *Deductive* and *Inductive* Reasoning.

ARIMA – Short for Auto-Regressive, Integrated, Moving Average models in *Time Series Analyses*. See, for example, McCleary and Hay (1980).

Arithmetic Mean – The formal term for the (*Central Tendency*) concept of the *Average*. Namely, the sum of the observations divided by the number of observations. Accurate for *Metric* data; vague and imprecise for *Ordinal* data; useless for *Nominal* data.

Assumption – An unstated *Premise* in an *Argument*. Assumptions may be problematic in logic and decision making when their veracity is difficult to demonstrate. See also: *Ockham's Razor*.

Asymmetric – A distribution that is not balanced across its *Median*. See *Skew*.

Average (arithmetic mean) – The *Quantitative* "middle" of a set of numerical values. The average is calculated by adding all of the values together and then dividing by the number of values.

μ is the symbol used to represent the average (mean) of a *Population*.

$\overline{X}$ is the symbol used to represent the average (mean) of a *Sample*.

Average deviation (D) – A measure of *Dispersion*; the average amount that the individual observations from a group are different than the *Arithmetic Mean*. The average deviation is calculated by summing together the absolute difference (that is, regardless of the direction of the difference) between each observation and the arithmetic mean and dividing by the number of observations.

$$D = \frac{\sum |X - \mu|}{N}$$

Note: While easily understood, the average deviation is only rarely reported. The most common measure of dispersion is the *Standard Deviation*, which has a wider range of uses (e.g. *Inferential Statistics*).

Background Variables – *Variables* that are extraneous to the research being conducted but may still influence variables within the study. For example, the air temperature may impact the amount of soda purchased from week to week, regardless of the amount of advertising. *Demographic Variables* are often background variables.

B2B Panel – A *Research Panel* that is used for business-to-business research studies. The members of such a panel are typically purchase decision makers who have consented to participate in various research studies.

As with all studies, care should be taken when utilizing B2B panels to ensure that the sample is truly representative of the target population. For example, even though a small business consultant is the sole decision maker for a computer purchase, you may not want to include him or her in a study of Chief Technology Officers.

Banner – In computer generated *Cross-Tabs*, the banner is the list of the measured *Variables* and their *Response Categories* that appear (in columns) across the top of the page for each variable analyzed (*Stub*). The banner categories represent sub-sets of the sample that may be compared to one another to assess similarities or differences. The variables in the banner are typically demographic or other descriptive measures whose response categories serve to delineate *Comparison Groups* (e.g., potential *Market Segments*).

Often, statistical tests are performed on the differences in responses between various pairs of response categories within the banner to determine if the groups are statistically different (using *t-tests* or *z tests*).

Basic Research – Research conducted to grow a general body of knowledge (e.g., to expand theory), rather than to answer a specific question.

Also known as 'Fundamental' or Pure' research. Basic research contrasts with *Applied Research*.

Beta (β) – The *Probability* of making a *Type II Error* when drawing the conclusion that the *Null Hypothesis* should be accepted from research. For practical purposes, Beta is the probability of making a mistake and concluding that no relationship (or difference) exists, when, in fact, one does.

More formally, Beta is the probability of accepting the *Null Hypothesis*, when, in fact, it is not true.

Note: *Statistical Power* is calculated as 1 - β. Power is the likelihood of not making a Type II error, of finding a relationship when one exists.

Beta Test – A *Pilot Study*; a 'field test.' Beta testing provides the actual product or service to a *Sample* of the target *Population* for 'real-life' use to see how it actually performs.

For example, new or upgraded software products are typically beta tested to get feedback as to usefulness (and to identify 'bugs') prior to broader release.

Beta weights – In *multiple regression* procedures, beta weights are the unique, relative importance of the *Normalized* (*standardized*) individual *Independent Variables* in the equation that explain or predict the *Dependent Variable*.

When assessing the beta weights of a multiple regression formula, those with the highest values have greater relative influence on the value of the dependent variable than do those with lower values. Note that the relative value of the beta weight for variables may be influenced by the order in which they are entered into the regression equation (due to *Collinearity* among the independent variables).

More formally, Beta weights are the standardized *Regression Coefficients* in multivariate analyses.

Bias (Research) – The difference between what something is and what it is reported to be. There are four primary types of bias in research and analysis:

Response Bias; where respondents answer untruthfully (either knowingly or unknowingly).

Sample Bias; wherein the sample observed (measured) does not accurately represent the population about which conclusions are to be drawn. This includes *Nonresponse Bias* and *Self-selection Bias*.

Estimation bias; in which a statistical *Estimator* (predictor variable) systematically under- or over-estimates the phenomenon in question.

Instrumentation bias; where the instrument used to measure a variable systematically under- or over-estimated the value of variable.

See also: *Cognitive Bias*.

Binary Variable – A *Dichotomous Variable* whose only 2 possible values are, typically, 0 and 1 (indicating the presence or absence of some attribute). See, for example, *Dummy Variable*.

Binomial Distribution (probability) – A *Distribution* that describes the probabilities all of the possible outcomes of a finite number of *Trials*, each of which has the same *Probability* of *Success*. A binomial probability refers to the probability of a specific number of successes from such a set of trails.

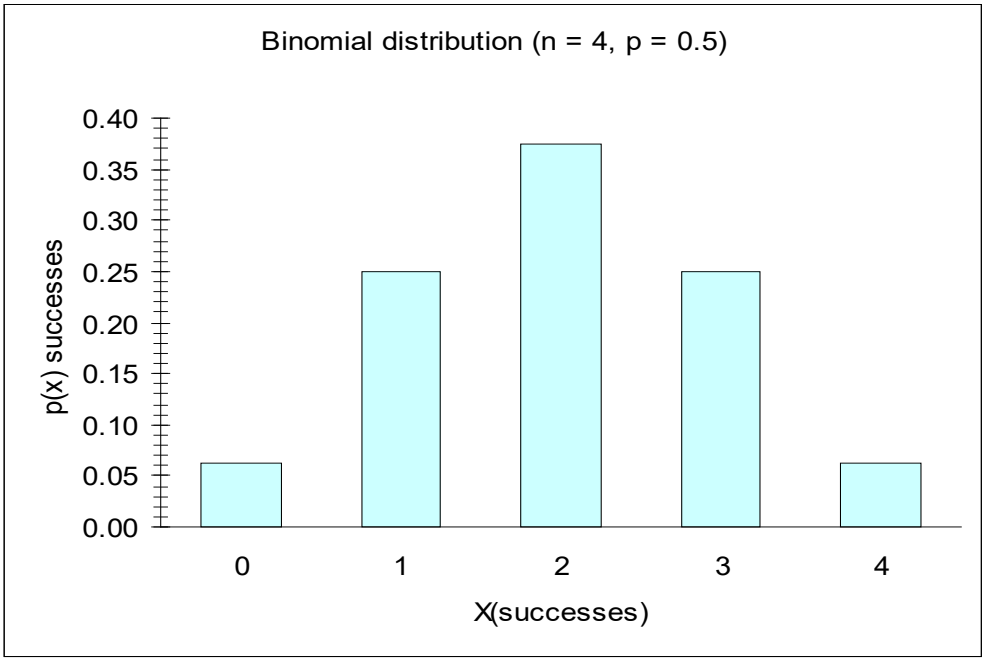
For example, the binomial distribution of the number of "heads" (successes) when four coins are tossed (4 trials) … and the probability that the coin will land heads is 50-50 (.50) would show that, randomly:

- ✓ No heads (0 successes) will come up 1 out of every 16 times (or 6.25% of the time) we toss 4 coins.
- ✓ One head (1 success) will come up 4 out of every 16 times (or 25% of the time) we toss 4 coins.

- ✓ Two heads (2 successes) will come up 6 out of every 16 times (or 37.5% of the time) we toss 4 coins.
- ✓ Three heads (3 successes) will come up 4 out of every 16 times (or 25% of the time) we toss 4 coins. And,
- ✓ Four heads (4 successes) will come up 1 out of every 16 times (or 6.25% of the time) we toss 4 coins.

See the graph below for a visual depiction of this binomial distribution.

Binomial distributions are typically used to describe the probabilities associated with *Categorical* (and not *Metric*) variables. Binomial distributions are used for many situations, for example when the samples sizes (e.g., the number of trials) are large and/or the probability of success is other than 50-50.



Hypothesis tests involving population *proportions* should employ binomial distributions as they are more accurate than using a *Z Approximation*.

When the number of trials is large, the binomial distribution begins to take on the shape of a *Normal Curve* — so much so that the normal curve may be used as an approximation of the binomial distribution. However, given the speed of binomial calculations on computers, the exact binomial calculations are preferred (over the use of the approximate normal distribution) as they are more precise.

Binomial Variable – A *Binary Variable*; a variable that has only two possible outcomes (or two response categories). The outcome of instant interest is typically called a "*success.*"

Binomial variables are characterized by the probability of success (which must remain constant) and the number trials (items selected or individuals asked). See *Binomial Distribution*.

A binary variable is an extreme and simple *Dichotomous Variable*.

Bivariate – Relating to two variables; such as the correlation between two variables. Simple regression is a bivariate procedure, as is a simple *Cross-tab* or *Contingency Table*. Compare with *Univariate* and *Multivariate*.

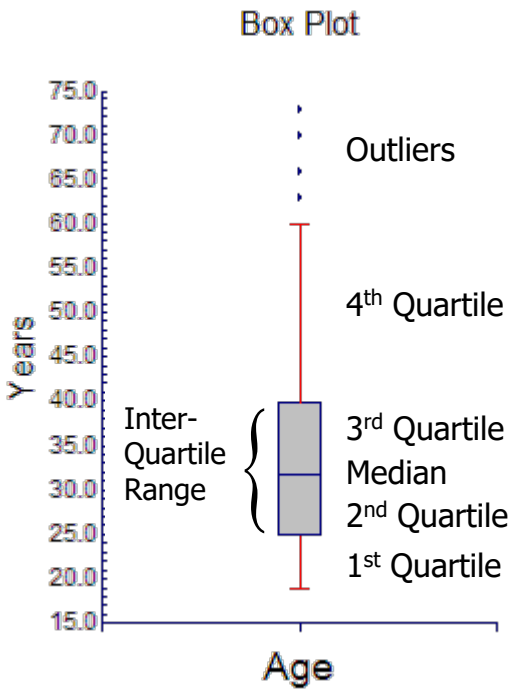
Blind Experiment – An experiment in which the test subjects do not know whether they are part of the *Control Group* or the *Experimental Group*. Blind experiments are conducted to minimize the impact of the *Placebo Effect* and, to some extent, *Self-Selection Bias*.

See also, *Double-Blind Experiment*.

Bounded Rationality – The proposition that, particularly in complex situations, people do not follow the Rational Model of decision making. Rather, that human rationality is limited (bounded) by processes of cognition and expedience (i.e., *Saticificing*) that lead to suboptimal results in decision making.

" … the assumptions of perfect rationality – are contrary to fact. It is not a question of approximation; they do not even remotely describe the processes that human beings use for making decisions in complex situations." (Simon, 1978, p. 366). Note that Simon was awarded the Nobel Prize in Economics in 1978 for his work developing this theory.

Box plot – A graphical representation of the four *Quartiles* of a *Metric* distribution in which the 1st and 4th quartiles are represented by single lines (along the relevant scale) and both of the middle two quartiles (the *Interquartile Range*) are represented by boxes on either side of the *Median*. *Outliers* are represented by dots away from the rest of the plot.



Box plots are often useful for comparing the differences in the distributions of separate populations when multiple plots are placed in the same chart. For example, box plots are often useful in comparing the distribution of values for the different levels during an *ANOVA* procedure. Box plots may be either vertical, such as this one, or horizontal.

Boyd, Laurie – The love of my life.

Branching Question – A survey question where the selection of different *Response Categories* results in respondents being directed to different subsequent questions. See *Skip Pattern*.

A branching question is of the type where respondents who select response categories A or B are directed to next answer question X, but those who answer C or D are directed to question Y.

Categorical Scale (or variable) – A scale or variable that may take on only a specific number of values (such as gender; "male" or "female"). For analytical purposes, the available categories for the scale or variable must be both *Mutually Exclusive* and *Exhaustive*. Categorical variables are

contrasted with *Continuous* variables; variable that may take on any value within a broad range (such as temperature or rent per square foot).

The only descriptive statistic for a categorical variable is the *proportion*; although the proportions for different groups may be combined (for example, the proportion of the sample that prefer either black or gray colored automobiles).

The *nominal* and, frequently, the *ordinal levels of measurement* are considered categorical variables.

Most survey questions (those with a *forced response*) have response categories that are set on a categorical scale. See, for example, the *Likert scale*.

Case – An individual or other entity (organization, branch, *Dyad*, etc.) from whom *Data* (observations) have been collected; for example, an individual survey respondent.

Causal research – The general name for research that is intended to determine whether or not there exists a relationship (*Causality*) between variables wherein the value of one in some way influences the value of the other (such as "advertising" and "sales"). Causing variables are called *Independent Variables* (or predictors), while the caused variable is called the *Dependent Variable*.

Causal Research generally attempts to determine if a change in the independent variable results in a change in the dependent variable.

Note that a relationship among variables is a necessary, but not sufficient condition of causality.

Causality – A rather amorphous concept that refers to nature of the relationship between two or more variables. One variable is said to "cause" another, if the value of the second variable is, in some way and to at least some extent, dependent upon the value of the first variable. *Independent Variables* are causing variables and *Dependent Variables* are "caused." Evidence of causality must include:

1) a theoretical (or at least logical) reason for the dependency,
2) the correct temporal order (change in the "caused" variable must occur after the change in the causing variable), and
3) *Concomitant Variation*; that is, some degree of association between differences in the first variable and differences in the second. Statistical tests such as *Correlation*, *Chi-square*, and *ANOVA* are all tests of levels of association between variables.

Evidence of concomitant variation alone does not constitute causality. There are too many potential alternative explanations (such as

relationships with *Intervening Variables* or random attributes of the sample) to safely draw such a conclusion.

To be established in a sound and objective manner, causality should be hypothesized first (*A Priori*) and then concomitant variation tested. If concomitant variation is detected first (e.g., through the construction of a *correlation matrix*) and then theory or logic constructed after-the-fact (*A Posteriori*), then the argument that causality exists is extremely weakened. See, for example, *Spurious* relationships.

For the most part, it is unlikely that one variable is the sole "cause" of another (e.g., it is unlikely that sales are caused only by advertising). Consequently, statistical studies such as *Regression Analysis* attempt to determine the degree of the relationship between the variables, typically through the analysis of the proportion of the *Variance* in the dependent variable that may be attributed to variance in the independent variable(s).

Census – A data collection process in which every member of a *Population* is *Measured*. In most cases a census is impractical due to time, costs and/or the inability to identify every member of the population.

A census permits exact knowledge of a population. That is, there is zero risk in making either a *Type I* or *Type II* error.

If a census is impractical, then an *Inductive* (*Inferential*) process using a *Sample* must be utilized (and the risks of Type I and Type II errors emerge).

Central Limit Theory – A theory that describes the *Sampling Distribution* of all (or an infinitely large set of) the same *Sample Statistic* taken from a *Population* – each *Statistic* (typically, the sample *Mean* or *Proportion*) must be calculated using the same *Sample Size (n)*.

The Central Limit Theory states that the distribution of all possible sample means ($\overline{X}$) [for example] with the same sample size (n) and taken from a single population, will take the (approximate) shape of a *Normal Curve*. The normal curve so described will have a mean that is equal to the population mean (μ) and a standard deviation that is equal to the population standard deviation (σ) divided by the square root of the sample size ($\sqrt{n}$). This standard deviation, when referring to distributions of sample means, is called the *Standard Error* of the sample mean (SE).

The Central Limit Theory may be used to describe the distribution of any sample statistic (e.g., proportion or variance) as a normal curve having a mean that is the value of the associated parameter.

Knowledge of the characteristics of a normal curve permit the calculation of the probability that the *Statistic* of any particular sample

will be within a certain proximity to the actual population *Parameter* (i.e., the population mean). For example, the characteristics of the normal curve are such that 95% of all the cases are within 1.96 standard deviations away from (either side of) the mean. Thus, using the Central Limit Theorem, there is a 95% probability that a sample taken at random from a population will be within 1.96 standard errors on either side of the true (but unknown) population mean.

Use of the Central Limit Theory is critical for both the logic and the calculations for *Confidence Intervals* and *Hypothesis Tests*. It is the foundation of most *Inferential Statistics*.

Central Tendency – Measures that identify the middle of a group (of numbers or observations). There may only be a "middle" if there is some way of ordering the observations (lowest to highest or through a *metric* scale). The common measures of central tendency are the *Mean* and the *Median* (and, occasionally, the *Mode*). See also, *Geometric Mean*.

Chi-square (X^2) – A *Nonparametric* statistic used to determine whether or not a relationship exists between two *Categorical Variables.*

Chi-square is used to test *Hypotheses* that a relationship exists between categorical (non-metric) variables. That is, it is a statistic utilized to establish whether or not the value of one categorical variable is, in some way, dependent upon the value of another categorical variable. For example, it could be utilized to establish whether or not a relationship exists between favorite televised sport and preferred brand of cola.

More formally, chi-square is a test to determine whether the *null hypothesis* – that a relationship does not exist between the two variables (that the two variables are *independent*) – should be rejected.

The probability (*P-Value*) associated with the X^2 determines the probability of making a *Type I error* if the researchers conclude that a relationship exists. The hypothesis should be accepted (and the null hypothesis rejected) if the p-value is less than the established acceptable *Alpha (α)*.

Classification Variable – Any variable that is used to segment the population into *Comparison Groups*. *Demographic* variables, technographic variables, and *Firmographic* variables are all examples of classification variables.

Closed-ended Question – A question on a survey or questionnaire that has a limited number of predetermined response options (*Response Categories*). Most of the questions used in *Quantitative Research* (that is, those that will be subject to statistical analysis) are close-ended questions.

Contrasts with *Open-Ended Question*.

Coefficient of Determination (r^2) – In *Correlation* procedures, the *Proportion* of the *Variance* in one *Variable* (typically the *Dependent Variable*) that is "explained" or accounted for by differences in another variable (the *Independent Variable*).

The coefficient of determination is calculated as the square of the *Correlation Coefficient* (r^2 = r * r).

In multiple regression situations, this coefficient is called the coefficient of multiple determination and is symbolized R^2. R^2 identifies the proportion of the variance in the dependent variable that is explained by the combined effects of all of the independent variables. This is not a simple additive process of the individual coefficients due to *Collinearity* between independent variables.

Higher r^2 values (those close to 1) indicate that a substantial proportion of the variance in the dependent variable may be explained by the values of the independent variable. (For example, how much of the variance in a respondent's "intent to purchase" could be reasonably attributed to seeing a commercial?) In real world situations, high r^2 values are rare and may signal the presence of intervening variables.

Coefficient of Variation – The ratio of the sample standard deviation to the sample mean, expressed as a percent.

$$CV = \frac{s}{\bar{x}} \times 100$$

The coefficient of variation is used to standardize variation across variables that are measured on different scales in order to make relative comparisons. That is, the coefficient of variation would be used to determine if one variable is relatively more spread out than another.

Cognitive Bias – Any mental process wherein an individual either misapprehends data or fails in a rational attempt to explain cause-and-effect. *Confirmation Bias*, described herein, is one of a multitude of cognitive biases. See, for example, the List of Cognitive Biases at Wikipedia.

Cohort – A group of individuals with the same or similar time-based characteristic. For example, all people born in 1993 or all employees who started working for the company in 2012.

Collinearity – The extent to which multiple independent variables are correlated with one another, and, thus, measure a common attribute.

Comparison Groups – Sub-groups (sub-populations) from within a population that are statistically compared to determine whether or not differences exist with regard to a second variable. Comparison groups are typically generated from *classification variables* (e.g., *Demographic* variables).

In *cross-tab* reports, comparison groups are represented by different columns in the *Banner*.

Comparative Rating Scale – A scale that requires respondents to compare numerous items to one another, rather than evaluate each item independently.

For example, a series of questions that asks respondents to identify "how much" they prefer one brand of cola to another uses a comparative rating scale.

Composite Variable – A *Variable* created by combining two or more variables; typically, by either averaging or summing the values of the component variables. Composite variables are often estimates of a *Concept*, with the component variables representing *Indicators*.

Note that summing the component variables may over weight some of the component variables if they are not all measured on the same scale.

Concept – An attribute or aspect for which there is no common measure (e.g., employee morale or customer satisfaction). Since a concept may play an important role in the development and testing of a *Theory*, it is often important to provide an *Operational Definition* (used for purpose of the specific research) constructed through one or more *Indicators* (measurable elements believed to be associated with the concept). See also, *Construct Validity*.

See, for example, Hayward and Hambrick's (1997) development of the concept of "CEO Hubris" as constructed through the indicators of: 1) recent organizational success, 2) media praise for the CEO, and 3) CEO's self-importance as reflected by their relative compensation.

Concomitant Variable – A *Covariate*; a variable whose value tends to coincide with the value of another variable.

For example, electricity usage is a concomitant variable with air temperature in Arizona (when the temperature goes up, so do electrical bills for air conditioning).

Concomitant Variation – Differences in one variable that tend to coincide with differences in another variable. Concomitant variation is a critical component in establishing *causality*, BUT NOT THE ONLY ONE.

Concomitant variation is typically measured through *Correlation* (R^2), the analysis of variance (ANOVA) *F-Ratio*, or *Chi-square* (X^2). Note that, while used generically to encapsulate this concept, 'correlation' is one of many types of concomitant variation.

Confidence Interval – Pragmatically, the *Range* within which it may be reasonably concluded falls the true (but unknown) population *Parameter* (e.g., the population mean). "Reasonably concluded" typically means "95% sure" (but other values may be presented). Smaller (narrower) confidence intervals are more precise, and, therefore, of greater value to decision making.

In general, the best way to narrow the range of the confidence interval is to increase the *sample size*.

The formal definition of a confidence interval implies some reservation about the accuracy of any given interval (because some values used in the calculations must also be estimated). Rather than concluding from a confidence interval …

> We are 95% confident that the true population parameter falls within the range from X (*LCL*) to Y (*UCL*).

It would be more precise to conclude that …

> 95% of the time we construct a confidence interval by this method, the range from X to Y will contain the true population parameter.

Confidence intervals may be constructed for any *parameter*, but the most common are for population *averages* (μ) or *proportions* (π), where it is typically referred to as the *Margin for Error*.

Note: To achieve greater confidence (e.g., say 99% rather than 95%), the confidence interval must be larger (broader); that is, it must have a larger range.

Confirmation Bias – The tendency to seek information (conduct research) in such a way as to only find results that verify a previously held point-of-view.

Conjoint Analysis – One of many research designs intended to determine the relative importance consumers place on the various attributes of a product or service.

The methodology employed during conjoint studies typically asks a large number of respondents to repeatedly select a preferred product from (numerous) sets of two or more products; each with a unique combination of multiple attributes (e.g., features, price, and brand).

Through the analysis of the relative frequency that each specific attribute is selected, it is possible to provide a somewhat more objective determination of the relative value that each respondent places upon each attribute. Combining the weights of numerous individuals, it is possible to establish population-wide preferences.

Constant Sum Scale – A form of *Comparative Rating Scale* that is constructed in such a manner that the responses to a series of questions (must) all add to a specific number; typically, 100.

Construct Validity – A type of instrumentation *Validity* that refers to the extent to which a question (or series of questions) measures a *Concept* that cannot be directly observed. The extent to which the *Indicators* capture the fundamental aspects of a *Concept*.

For example, "customer satisfaction" cannot be directly measured. Consequently, researchers measure what they believe to be the components of customer satisfaction (typically as defined through *Operational Definitions*).

Consumer Panel – a *Research Panel* type of *Sampling Frame* that includes members of the buying public. Consumer panels are widely used in on-line research. However, some caution should be employed when considering an on-line consumer panel to ensure that the population of interest may be found on-line (the elderly and lower income strata may be under-represented in on-line consumer panels).

Content Analysis – Any of a variety of methods for assessing qualitative (e.g., written, *open ended* or text) information that typically involve counting words or concepts of interest.

Content Validity – Instrumentation *Validity* that addresses the subjective assessment of how well a questionnaire captures the full extent of the *Concept* in question.

Contingency table – A descriptive method for showing the relationship between two *categorical variables*. The response categories for one variable (variable A) are listed across the top, while the response categories for the other variable (variable B) are listed down the left side. At the intersection of each column (values of variable A) and each row (values of variable B) appears either the absolute or relative frequency (*proportion*) of that combination of responses in the *database*. In the contingency table below, there are 40 Pandora users in suburban areas in the database.

A row is also provided to identify the total number or proportion of observations in each variable A category and a column provided for the totals of the variable B categories.

CONTINGENCY TABLE
Variable A: Digital Music Source

Variable B: Location	Spotify	Pandora	i♥Radio	Total
Urban	28	17	20	65
Suburban	12	40	18	70
Rural	8	12	25	45
Total	48	69	63	180

Variations of these tables include row, column and/or cell percentages. *Chi square tests* are often conducted on contingency tables to determine if there is a statistical relationship between the two variables.

Note: *Crosstabs* are a form of contingency tables.

Continuous Scale (or variable) – A *Metric Scale* or variable that may take on any value within a broad range (such as exact monthly salary in dollars or number of units sold). Continuous scales have an infinite number of possible values as there are an infinite number of possible values between ant 2 points on the scale (e.g., there are an infinite number of points between one and two meters in length).

Continuous variables are contrasted with *Discrete* metric variables that can take on only a limited set of values.

Control – The methodological or statistical removal of the effects of *extraneous variables* or *covariates*.

Control Group – In effective *Experimental Designs*, the portion of the sample for whom the *Independent Variable* has not been changed or manipulated.

The use of control groups in experiments allow researchers to account for many *Extraneous Variables* (threats to *Internal Validity*), ensuring that any changes reported for the *Dependent Variable* are truly the result of changes in the independent variable. Compare with *Experimental Group*.

For example, in a study to determine the effectiveness of an advertisement, two groups may be evaluated on their likelihood to purchase a product: one group who had seen the advertisement (the experimental group) and a second group who had not seen the advertisement (the control group). The difference in purchase likelihood between the two groups could then be measured in an attempt to assess the impact of the advertisement.

Convenience Sampling – A type of *Non-Random Sampling* in which potential respondents are selected on the basis of their accessibility (and not so-much on the extent to which they represent the broader population of interest).

Convenience sampling is what professors do when they use students as surrogates for a broader population during research studies. Studies based upon convenience sampling methods are susceptible to *threats to external validity*;

Correlation – A measure of *Concomitant Variation* that quantifies the nature and strength of the relationship between *Metric* or *Ordinal Variables.* [Using correlation procedures on *Categorical Variables* is not a valid statistical practice.]

The two most common correlations are Pearson (for *Metric* data) and Spearman (for ranked – *Ordinal* – data).

Both correlations are measured by the statistics r (the "*Correlation coefficient*") and r^2 (the "*Coefficient of Determination*"). R measures

the direction and magnitude of the relationship between the variables and may take values between -1 and 1.

A positive correlation (i.e., a correlation with an r between 0 and 1) indicates that the variables tend to move together; that larger values of one variable are typically associated with larger values of the other. A negative correlation (i.e., a correlation with an r between -1 and 0) indicates that the variables tend to move in opposite directions; that larger values of one variable are typically associated with smaller values of the other.

The magnitude of r (the closer it is to 1 or to -1) suggests the strength of the relationship.

R^2 measures the proportion of the *Variance* in one variable that is accounted for ("explained") by the other(s) and may take values between 0 and 1
($r^2 = r * r$).

Note: correlations, by themselves, do not demonstrate *Causality*. Without sufficient theoretical justification, the exclusion of potentially intervening variables, etc., the potential for concluding causality exists, when in fact, the relationship may be illogical or *Spurious*.

Correlation Coefficient (r) – The typical product of *Correlation*; a measure of the relationship between two variables. The correlation coefficient r may take on any value between -1 and 1.

Correlation coefficients may be calculated using a number of different formulas; however, the two most common are:

- ✓ the *Pearson Product-Moment* (or just *Pearson*) correlation for correlations of *metric variables*, and
- ✓ the *Spearman* (or *Rank-difference*) *correlation* for correlations involving data measured on an *ordinal scale*.

A *Positive Correlation* (i.e., a correlation with an r between 0 and 1) indicates that the variables tend to move together; that larger values of one variable are typically associated with larger values of the other. A *Negative Correlation* (i.e., a correlation with an r between -1 and 0) indicates that the variables tend to move in opposite directions; that larger values of one variable are typically associated with smaller values of the other.

The magnitude of r (the closer it is to 1 or to -1, and away from 0) suggests the strength of the relationship. (See the *Coefficient of Determination* – r^2 – for another measure of the strength of a correlation.)

Correlation Matrix – a table that contains the *Correlation Coefficients* between all relevant pairs of *Variables* in a *Database*.

Covariate – A third variable (besides the *Independent Variable*) that correlates with the *Dependent Variable*.

Some statistical procedures (e.g., *Analysis of Covariance*) are designed to *Control* for the effects of covariates (so that that only the relationship between the independent and dependent variables are assessed).

Credibility – The trust with which is afforded to sources of knowledge. Some sources have more than others. See, for example, the *Pyramid of Evidence*.

Criterion Variable – Another name for the *Dependent Variable*.

Critical Thinking – Any process that seeks *Objectivity*, typically through the conscientious avoidance of *Bias* (i.e., *Confirmation Bias*). An important practice of good decision makers in that it serves as a guard "against delusion, deception, superstition, and misapprehension of ourselves and our earthly circumstances" (Sumner, 1940).

Cross-tab – The research industry's name for a computer generated *Contingency Table*. For convenience and to save space, most cross-tabs compare a single *Dependent Variable* of interest to a number of *Independent* descriptive variables and/or market *Segments* (*Classification Variables*). The resultant table presents both the absolute frequency and the column percentage for each row-column combination.

Cross-tabs are frequently subjected to a series of statistical analyses of the differences in the proportion in dependent (row) response categories across columns. These tests are typically a series of *Z-Test* of the difference in proportion rather than *Chi-Squared* tests.

Cross-Sectional Study – Research that is conducted all at a single time. This type of study precludes the measurement of change. Also known as a *One-Shot Case Study*. Cross-sectional studies are useful as statistical 'snapshots,' but are largely unhelpful in establishing *Cause-and-Effect* (as it is difficult to measure any change in, say, the *Dependent Variable when it is only measured one time).*

Custom Panel – A *Research Panel* that contains individuals with specific characteristics for study. Custom panels are typically constructed when the desired characteristic is not easy to identify in the general public (or, for example, a broad, *consumer panel*).

Custom panels may be developed as proprietary, for use by a single organization, or for rent to a broad range of companies.

Custom panels are also called *specialty panels*.

Custom Research – Research that is conducted for a single, typically proprietary, purpose. Contrasts with *Syndicated Research*.

Data – A collection of quantified or categorized measurements (*Observations*). (Note: "data" is the plural of "datum.")

Database – The *Data* from a number of cases that are arranged in a systematic manner for easy assessment and analysis.

Convention holds that each row of a database contains the *Observations* for a single *Case*. The columns of a database, consistent across all cases, contain the information collected for a single *Variable*. Thus, the intersection of a row A and a column Z contains the measure of variable Z for case (respondent) A.

Data Dredging – See *Data Mining*, *Fishing Expeditions*, and *P-Hacking*.

Data Mining – Any *A Posteriori* practice used in the attempt to identify either statistical relationships among variables or statistical differences among subgroups (*Segments*) of the *Population*. The practice typically involves the analysis of very large amounts of *Data*. These data reside in *databases* (often *relational databases*) with a large number of both *Variables* and *Cases*.

As with other exploratory data analysis practices, any after-the-fact explanations for statistical relationships uncovered in this manner should be rigorously examined for theoretical and logical consistency in order to ensure that any relationships exposed are not *Spurious*. As a practice for theory generation, any results obtained though data mining must be tested on new data before they could be accepted. See *P-Hacking* and *Fishing Expeditions*.

Data weighting – The practice of adjusting the relative contribution of subsets of the *Sample* to reflect incidence in the *Population*.

For example, if the population being studied contain equal numbers of men and women, yet the sample yield 40% men and 60% women, the data would need to be adjusted by gender to reflect the true 50/50 distribution.

Deductive (Reasoning) – The process of drawing a conclusion about an individual (item, incident, person, etc.) based upon a set of reasons that generically state that, if something is always true for a group, it is true for every individual within the group. Sometimes referred to as "arguing from the general to the specific."

First derived by Aristotle through the development of syllogisms, deductive reasoning evolved into set theory, and has been employed in computer language logic (e.g., 'If/Then' statements). See Parry and Hacker (1991).

Contrasts with *Inductive Reasoning*.

Degrees of freedom (df) – The "degrees of freedom" in statistical calculations refers to the number of items or measures in a group about which knowledge

must be known in order to determine the values of all the measures in the group.

For example, if there are three items measured and their average is known to be five, then we only need to learn the values of two (say 5 and 9) to be able to deduce the value of the third (1); [5 = (5+9+x)/3 and x=1]. This situation would have 2 degrees of freedom.

For calculations of the *t* distribution, the number of degrees of freedom is one less than the number of elements measured; df = n – 1.

For calculations of *chi-square* (X^2), the number of degrees of freedom is one less than the number of categories for the first variable times one less than the number of categories for the second variable; df = (number of rows – 1) * (number of columns – 1).

Knowledge about the number of degrees of freedom is used in other statistical calculations; for example, the *F-ratio* (*ANOVA*).

Demographic Variables – A set of measurable characteristics – *Variables* – of a (human) population that are used for classification and *Segmentation* purposes. They describe population attributes such as gender, age, education, income, number of children, et cetera.

Demographic variables are one type of "*Background Variable*."

Depth Interview – A fairly long, relatively unstructured interview (either by telephone or face-to-face) designed to explore the details of situation, event, or process. Also known as an in-depth interview. A *Qualitative Research* technique.

Descriptive Research – Research that is designed to provide a description of a *Population* or *a segment* of a population through *Descriptive Statistics*. Descriptive research is often undertaken in order to understand various characteristics of the population as it relates to *demographic*, "*technographic*," or "*firmographic*" variables. As descriptive research is typically not *Longitudinal*, it is inappropriate for *Causal Research* or conclusions.

Descriptive Statistics – Statistical measures that summarize or collapse large amounts of data from a number of individual observations (*Cases*) into singular, quantitative characteristics of a group (e.g., a *Mean* or *Proportion*).

The appropriate descriptive statistics for a group or *Variable* depends on the *Level of Measurement* of the observations. When the level of measurement is *categorical* (i.e., *Nominal* or *Ordinal*), the appropriate descriptive statistic is the *proportion*. When the level of measurement is *Metric* (*Interval* or *Ratio*), the appropriate descriptive statistics are measures of *central tendency* (*Mean* and *Median*) and measures of *Dispersion* (*Range*, *Inter-Quartile Range*, *Variance*, *Standard Deviation* and/or *Average Deviation*).

Descriptive statistics also refers to the estimation of *Population Parameters* from *Sample Statistics*; (i.e., the development of *Confidence Intervals* or *Margins of Error* for a proportion).

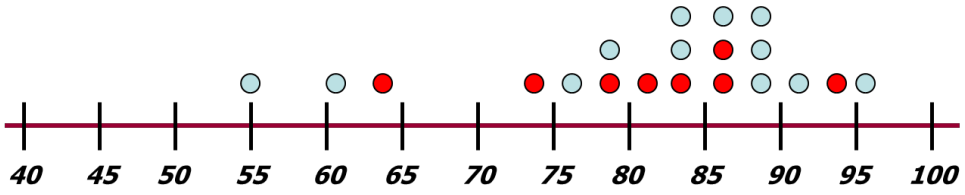
Dichotomous Variable – A *Variable* with two antithetical values (or ranges). So, besides simple yes/no variables, dichotomous variables include scaled questions with ranges of antithetical values. See, as an example, *Likert Scales* (with ranges of 'agreement' and 'disagreement'). *Binary Variables* and *Binomial Variables* are also examples of dichotomous variables.

Discrete Variable – Any variable (either *Metric* or *Categorical*) that may take only a limited number of possible values. For example, the number of bedrooms in a domicile or preferred cola brand. Contrasts with *Continuous Variable*.

Dispersion – Statistical measures that depict how spread out the individual values of a set are. Dispersion may be measured overall (the *range*), about the *median* (the *inter-quartile range*), or about the *mean* (the *variance* and *standard deviation*, and, occasionally, *average deviation*).

Distribution – A collection of *observations* (*data*) into *mutually exclusive* groups based upon the value of a particular *variable*; describing the number or percentage of each. Distributions may be described in either tabular (i.e., data tables), formulaic, or graphical form.

Dot plot – A simple charting technique that presents a single dot or circle for each observation across the scale. Multiple observations with the same value are stacked upon each other. If the number of observations is relatively small, these dots may be colored by subpopulation.



Dot plot: Test Score by Gender

Double-Barreled Question – A type of *Ambiguous Question* where a respondent is asked to answer two questions at the same time. Since the answers to the two questions may be different, the respondent may not know how they should respond.

"How satisfied were you with the wait time and the technical knowledge of the support associate during your recent technical service call?" is a double-barreled question. If the respondent were satisfied with the level of knowledge, but unhappy at having to wait half an hour to receive it, he or she would not know how the question was to be

answered. Consequently, the data collected from double-barreled questions are difficult to analyze effectively.

Double Blind – An *Experiment* in which neither the test subjects nor the principal evaluator knows which subjects are members of the *Control Group* and which are members of the *Experimental Group*. See the *Placebo Effect*.

Dummy Variable – A *Regression* practice in which a *Categorical Variable* is transformed into a number of *Binary Variables* in order to satisfy the *Metric Variable* requirement of *Correlation* or regression (particularly, *Multiple Regression)*.

In practice, a categorical variable with n categories is converted into a series of n-1 "dummy variables," each of which measures the presence (1) or the absence (0) of a particular response category or attribute from the original variable. (Only n-1 categories are necessary because, for an individual observation, if none of the n-1 measured attributes are present – that is, they are all are zeros – then the remaining attribute must be present.)

A multiple regression analysis is then run which includes the series of n-1 dummy variables. The resulting regression formula will contain *Beta Weights* for each of the categorical attributes.

Dyad – A paired unit of measurement. For example, a study of managerial effectiveness that collected data from both managers and their subordinates could combine the data from both into manager-subordinate dyads for analysis.

Effect Size – Within a specific experiment, the magnitude of any change in the dependent variable. Due to the multitude of challenges in translating raw experimental results to the world at large, extreme caution should be used before any claims that such results (in the experiment) suggest that similar results are true for any broader population.

Ethics (General) – Of or relating to the intentional insertion of *Bias* into research, analysis and reporting practices. Ethical concerns include: *Plagiarism*, *Fabrication* (the falsification of data), and the intentional mis-interpretation of results (including such activities as *P-Hacking*).

Ethics (Research) – The primary ethical consideration in research is to "do no harm" to the entities being studied. That is, they should come out of the study no worse than they went into it. (If they are, or could be, exposed to harm in any way, they should be informed of the risk before they participate in the research; see *Informed Consent.*) The types of "harm" that researchers should be concerned with include: physical, financial, social and economic.

Ethnographic Research – Typically *Qualitative Research* into the norms, values & social structure (e.g., culture) of a specific group. Ethnographic research

often requires that researchers immerse themselves in the organization in order to eliminate any participant bias in the study (see, for example, *Participant Observation*).

Exhaustive – Accounting for all possible responses. If the range of *Response categories* for a *variable* provides for all possible responses to the question asked, the question is considered to be "exhaustive."

For example, a variable that collects individuals' ages into the following categories is NOT exhaustive.

A. 20 to 39
B. 40 to 59
C. 60 to 79

Note that individuals younger than 20 or 80 or older have no response categories. A more appropriate way to construct the response categories for this particular variable would be:

A. Less than 20
B. 20 to 39
C. 40 to 59
D. 60 to 79
E. 80 or more

Only variables whose response categories are both *mutually exclusive* and exhaustive are appropriate for research and analysis.

Expected Value – A basic calculation of decision science used to assess and compare various options. In its most simple format the following computation is performed for each option:

Expected Value = (*Probability* of success) X (Value of the option)

According to general tenants the option with the highest expected value is the preferred option.

Experiment – In *Causal Research*, the manipulation (or naturally occurring changes) of the value of an *Independent Variable* in order to measure hypothesized changes in the value of a *Dependent Variable*.

There are a variety of ways that experiments are conducted (see *Experimental Design*), however, the more effective experiments are designed in such a way such that only changes in the dependent variable that are attributable to changes in the independent variable are measured or reported. This is typically accomplished through the use of Control Groups (for which the independent variable is not manipulated).

Experimental Design – The methodology used to conduct an *Experiment*. Effective experimental designs are those that best control for threats to *Internal Validity*. They also give attention to threats to *External Validity* (Campbell & Stanley, Cook and Campbell).

Some experimental designs are better than others in ensuring that the conclusions drawn are warranted (i.e., that, barring *Intervening Variables*, the only possible explanation for changes in the *Dependent Variable* are changes in the *Independent Variable*).

In increasing order of the control of threats to internal validity afforded, experimental designs include:

Static Group Comparison:			X	O	Experimental Group
(Posttest only)				O	Control Group
1 Group Pretest/Posttest		O_1	X	O_2	
Pretest/Posttest with Control		O_1	X	O_2	Experimental Group
		O_1		O_2	Control Group
Randomized Pre/Post	(R)	O_1	X	O_2	Experimental Group
(with Control)	(R)	O_1		O_2	Control Group

Other designs, such as the Solomon Four Group design, control threats for some forms of External Validity.

Experimental Group – In *Experimental Designs*, the portion of the *Sample* for whom the *Independent Variable* is changed or manipulated in order to see if there are consequential changes in the *Dependent Variable*.

Compare with *Control Group*.

Exploratory Research – Research studies, predominantly *Qualitative Research*, designed to surface ideas and concepts (ideally, for future quantitative analysis). Designs include *Focus Groups* and *Depth Interviews*.

Typically, Exploratory studies collect information in an unstructured manner from a limited number of respondents, rendering statistical inference impractical, if not dangerous. Exploratory studies should be confirmed using larger samples and quantified data before any decision is made.

Exploratory Data Analysis – Any *a posteriori* statistical procedure that is not a *hypothesis test*; that is, not based on *theory*. A search for relationships among *variables* (or differences between *sub-populations*) that researchers have not previously expected or contemplated.

Exploratory data analysis techniques are more subject to the possibility of *Type I errors* than are techniques driven by theory.

See *fishing expedition* and *data mining*.

External Validity – The extent to which the results of a research study may be effectively translated to other, broader situations (e.g., the world outside the situation in which the experiment or study was conducted).

Researchers are concerned with "threats" to both *internal* and external validity. The most common threats to external validity are:

Reactive effects that occur when the subjects being studied behave (react) differently when they are being measured than they would outside the study environment. For example, when someone knows that they are being observed, they may act differently than they would if they were not observed (e.g., people tend to drive differently – "better" – during driving tests than they do without the observer).

Non-representative samples where the sample studied does not accurately reflect the population about which conclusions are to be drawn.

Pre-test interaction where subsequent measures of subjects are tainted by the subjects' exposure to a *pre-test*.

Extraneous Variables – Variables other than those included in a hypothesis that may be measured (or have some influence on the *Dependent Variable*) during a study. Many of the threats to *Internal Validity* are extraneous variables.

F-ratio – A ratio used in *Analysis of Variance*, *Analysis of Covariance* and certain regression procedures to determine the statistical significance of the test.

Basically, the ratio of the "explained" variance, as in 'attributable to the independent variable' (adjusted for the degrees of freedom in the model), to the unexplained variance (adjusted for the remaining degrees of freedom in the data). Computer program generated F-Ratios have attendant *P-Values* that will assist in determining the extent to which the statistic should be considered *Significant*.

Fabrication – Any situation wherein a researcher alters or manufactures aspects of the research (e.g., data or citations) in order to obtain a desired result or authenticity. Also called 'Data Falsification.' See *Ethics*.

Factor Analysis – A statistical procedure designed to collapse a number of *Variables* into smaller number of 'factors.' If such a practice is effective, then these factors serve better to describe or define differences in subjects.

Factor analysis may be either 'Confirmatory' (designed to test hypothesized relationships, such as the relationships between *Variables* and *Indicators*) or 'Exploratory' (seeking to identify which variables contribute most to the various factors).

Factorial Designs – *Experimental Designs* that measure different levels of two or more *Independent Variables* simultaneously to assess their joint (an/or individual) impact on a *Dependent Variable*.

Faith – A belief in something that cannot be verified through evidence-based approaches. For many scientists, 'faith' is an abomination … a 'four letter word.'

"Firmographic" Variables – The informal term for a series of descriptive measures relating to the classification of businesses (such as industry type, number of employees, technologies employed, ownership structure, number of branches, etc.). Contrasts with *Demographic* variables for people.

Fishing Expedition – The practice of using statistical techniques in the attempt to discover relationships or differences among variables; relationships and differences that have not been previously hypothesized. Typically, researchers will then attempt to construct theory to fit the statistical conclusions (i.e., *a posteriori* reasoning). See, for example, *Data Mining*.

Fishing expeditions are often performed through the visual assessment of an inclusive set of *Cross-Tabs* (in the attempt to identify statistical differences between groups) or of a *Correlation Matrix* (in the search for unusually high *Correlation Coefficients*).

For purists, this practice is generally viewed as methodologically unsound due to the likelihood of uncovering *Spurious Relationships*. If a fishing expedition IS used, follow-up studies with independent (or *Hold-Back*) data must be performed prior to using the results. In any event, extreme caution must be employed to ensure that any theory subsequently postulated also has a logical – and not solely statistical – basis.

Fishing expeditions are also known as *Data Dredging*.

Five Factors Model (FFM) – A model that posits five distinct components (constructs) of personality: Openness (to Experience), Conscientiousness, Extraversion, Agreeableness and Neuroticism (Digman, 1990). Measures of these components are often considered important Psychographic Variables.

Sometimes referred to as 'The Big Five.'

Focus Group – An *Exploratory*, *Qualitative Research* process in which a small group of individuals (10 – 15) are brought together to discuss (focus on) a topic of interest to the researcher.

Focus groups are overseen by a participating moderator who directs the conversation through a 'discussion guide' that covers the topics of interest to the researcher. All participants are expected to engage in the conversation as the interactions among them is considered important. Focus groups sessions are recorded (with participant knowledge) and evaluated through some form of *Content Analysis*.

Due to small sample sizes and the collection of qualitative, impressionistic information, it is unwise to try to draw inferences of the population from a focus group. Rather, as with other exploratory research practices, they are most useful in Hypothesis generation (that could then be tested through more credible processes (e.g., *Controlled Experiments*).

Forced Response – A term that applies to either:

1) a type of question on a survey or questionnaire that requires exactly one response to a *scaled question*. "Radio button" questions (wherein only one response category may be selected) during Internet based research. Or,

2) the requirement that a respondent answer one question on a survey or questionnaire before being allowed to proceed to the next question.

Garbage Can Model (Decision Making) – The idea that situational or temporal aspects of decisions often play a critical role in their success. That is, that every decision is influenced by a variety of unique factors (unique times, players, organizations, wants, needs, etc.) all interacting to result in a particular decision choice. As such, each decision is difficult to assess or understand and harder yet to generalize from.

Geodemographics – Traditionally, Geodemographics has referred to databases that contain both geographic and *Demographic* data. These databases are useful for such business activities as finding the optimal location for a store relative to a target demographic.

Geometric Mean – A measure of *Central Tendency* calculated by multiplying a number of items (n) together and then taking the n^{th} root of the product.

For example, the geometric mean of 4, 8, and 16 is:

$$GM(4, 8, 12) = \sqrt[3]{4 * 8 * 16} = \sqrt[3]{512} = 8$$

The geometric mean is useful normalizing scales of differing magnitudes and for describing proportional growth over a number of observations.

Halo effect – A *Cognitive Bias* in which a positive aspect of an entity (person, corporation, sports team, etc.) influences assessments of other aspects of the entity. A *Heuristic* that should be used only with extreme caution.

Heterogeneous – Generally dissimilar or diverse. When used in reference to a single measure (variable), heterogeneous implies that the distribution of cases or observations covers a rather large *range*. When used in reference to a sample or population, heterogeneous implies that the group varies widely with regard to specific variables of interest (e.g., income, age, education, etc.).

Metric variables with large measures of *dispersion* (i.e., standard deviations) and *categorical variables* with proportionally equal categories would be considered heterogeneous.

In practice, Heterogeneous is a rather relative term that contrasts with *Homogeneous*.

Heteroscedastic – (See *Homoscedastic*)

Heuristic – Sometimes called a "Rule-of-Thumb," any general rule employed during decision making that relies on experience, tradition, or *Cognitive Bias* (rather than *Inductive* or *Deductive* methods) to arrive at a conclusion.

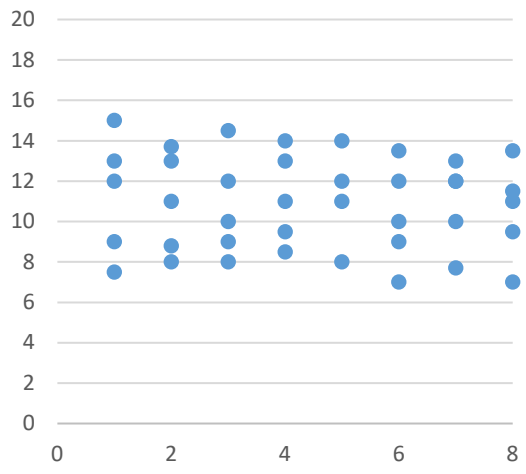
In general, heuristic approaches are more effective in routine and similar situations but become less useful and more error prone as situations become less similar and/or unstructured. A bias in decision making comes when decisions makers fail to understand when the rule simply does not apply.

Hold-Back – A portion of the data not included in the original analysis that may be used to test hypotheses generated *Fishing Expeditions* or any relationship uncovered without being previously hypothesized.

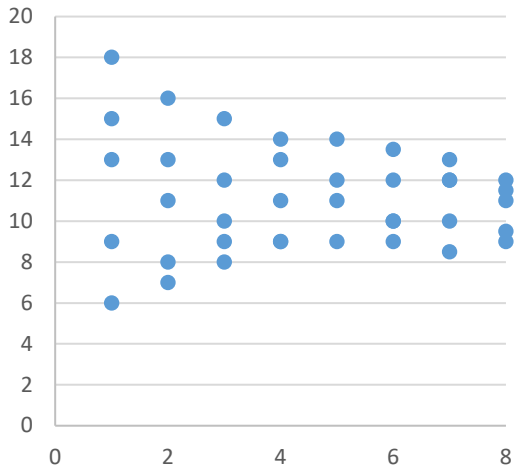
Homogenous – Generally, similar. *Metric variables* with low measures of *dispersion* (i.e., standard deviations) and *categorical variables* with disproportionately high incidence in a single category would be considered homogeneous.

In practice, Homogeneous is a rather relative term that contrasts with *Heterogeneous*.

Homoscedastic – (along with *Heteroscedastic*) The extent to which a distribution has consistent variability (in one variable) across the range of observations in another.



A



B

The data in Chart A above would be considered more Homoscedastic than that in Chart B because there the amount of spread is fairly consistent across the observations. Chart B is the opposite, more *Heteroscedastic*, in that the observations for the left side (values 1 and 2) appear to be much more spread-out than the observations at the right (7 and 8).

Homoscedasticity is measured by looking at the *Variance* of the Y (vertical) axis observations for each value on the X axis and then calculating at the ratio of the highest to lowest.

Homoscedasticity is an important assumption for the usefulness of *Linear Regression* interpretations and applications.

Hypothesis – An, as of yet, untested belief about the nature of the world. Business research hypotheses typically fall into two basic categories:

1) statements about a specific characteristic of a population (or the differences between the same characteristic in more than one population) and

2) statements about the relationships between two or more variables within a population.

Only those hypotheses that are capable of being disproved are considered appropriate for scientific testing.

To be useful in research and analysis, there must be good (e.g., logical or theoretical) reason to believe that the hypothesis is true. Otherwise, the research tends to be more prone to errors (typically *Type I errors*). Hypotheses must also be quantifiable and testable if they are to be used for drawing conclusions (i.e., if they are to be tested statistically).

In formal research discussions, hypotheses are called "Alternate" hypotheses (and symbolized H_1, H_2, etc.) in order to distinguish them from *Null hypotheses* (symbolized H_0).

Objective scientific research is the process of testing hypotheses, not "proving" them.

Hypothesis Test – Any of a variety of (typically, statistical) procedures designed to determine if a *Hypothesis* is likely to be true.

More formally, hypothesis tests are designed to determine if the *null hypothesis* (H_0) should be rejected (to effectively manage the probability of making a *Type I error*) in order to determine whether or not there is evidence to support the hypothesis (H_1).

Statistical tests of hypotheses test the null hypothesis to determine whether or not the null hypothesis should be accepted or rejected. The basis of the conclusion is the probability of making a *Type I Error* if the null hypothesis is rejected. If this probability (the *p-value*) is less than

Alpha (α), then the null hypothesis is rejected. If the probability is greater than Alpha, then the null hypothesis is accepted.

If the conclusion is to reject the null hypothesis, then there is considered sufficient statistical evidence to support the (alternate) hypothesis (Reject H_0 = Accept H_1).

If the conclusion is to accept the null hypothesis, then there is insufficient evidence to support the claim of the (alternate) hypothesis (Accept H_0 = Reject H_1).

Since research and analysis are processes of discovery, it is important to note that hypothesis testing is a just that, a test. Research and analysis projects that set out to "prove" a hypothesis are, by their very definition, biased toward that conclusion and are likely to suffer from a lack of *objectivity*.

Impossibility (Statistical) – An oxymoronic statement betraying a lack of understanding of *Probability*. Statistical processes seek not to prove anything, but, rather, to identify the likelihood that something is or is not.

Note: A *P-Value* of .0000 does NOT indicate that it is statistically impossible to obtain the results by chance. The proper interpretation of this P-value is that the likelihood is less than .00005 (and simply rounded to .0000 for ease of reporting).

Indicator – A potentially measurable attribute believed to be associated with a *Concept*. Indicators are utilized to operationalize concepts (they assist in the creation of an *Operational Definition* for the concept). Indicators are converted into *Variables* which can then be used to establish the extent to which a concept is present (at least as operationally defined).

An effective set of indicators will provide *Construct Validity* for a concept.

Inductive (Reasoning) – A branch of logic that holds that if something is true for a representative subset of a *Population*, it is likely true for the entire population. Commonly understood to mean the practice of arguing from the specific to the general.

This branch of reasoning is the foundation of *Inferential Statistics*, such as drawing a conclusion about all of one's customers based upon a market research study of 500 of them. See *Hypothesis Test*.

Inferential Statistics – *Inductive* quantitative processes that are utilized in attempts to determine the extent to which data from a *Sample* may or may not be used to draw conclusions about a *Population*. Inferential Statistics are a critical tool in the *Hypothesis Testing* process.

Informed Consent – The formal acknowledgement, usually in writing, of the potential risks involved as a participant in a study. Ethical standards hold that the participants in a research study should be informed of any possible adverse consequences they might encounter before they are asked to agree to participate in the study.

Instrument – Any research tool used to collect a measurement. For example, survey questionnaires, scales, web traffic counters, thermometers, etc.

Intercept (Regression) – The value of the *Dependent Variable* when the *Independent Variable* is zero, as established through a *Regression* equation.

Regression is short for "Linear Regression", the best estimate of the straight-line relationship between variables. A straight line is defined through the equation y = (ax) + b, and, thus, provides an estimate of the value of the dependent variable (y) given a value of the independent variable (x). The *Intercept* (b), is a constant that represents the value of the dependent variable when the independent variable has a value of zero. Note: The slope of this relationship (another constant, 'a') represents the incremental increase (or decrease) in y for values of x.

For example, a regression equation for estimating income could take the form Income = ($5000 * Number of Years of Education) + $20,000. In this case, $5000 is the slope and $20,000 is the intercept.

Intercept (Sampling) – The act of selecting sample participants (randomly) from a public venue. Traditionally, this referred to as in-person solicitation of participants at such public places as shopping malls or public sidewalks. More recently it can refer to the solicitation of participants at websites (e.g., through "pop-ups").

The inter-quartile range is represented by the "box" in a *Box Plot*.

Internal Validity – The extent to which an *Experimental Design* 'controls' for the measurement of *Extraneous Variables*. (Campbell & Stanley, 1963; Cook & Campbell, 1979). Traditional 'threats to internal validity' include:

History: All variables outside of the experiment itself that may be measured (e.g., the weather or the economy)

Maturation: Changes to test subjects due to the passage of time

Testing: the effects of a *Pretest* (definition 2) on subjects as it influences their responses to subsequent measurements (e.g., the *Post Test*)

Instrumentation: Differences in the *Instruments* used in the Pre Test and the Post Test. (Ideally, there are none.)

Regression: A statistical process by which extreme cases tend to (randomly) migrate to the middle upon subsequent testing.

Selection: Possible differences in results when the *Control Group* and the *Experimental Group* are fundamentally different prior to the Pre Test

Mortality: The inability to measure some subjects during the *Post Test*.

Effective research designs (i.e., controlled, randomized *Experiments*) effectively 'control' for threats to internal validity.

Interquartile Range (IQR) – The *Range* of the middle 50% of the observations (as measured around the *Median*). A simple measure of *Dispersion* wherein smaller values are indicative of *Homogeneity* (observations tend to me more tightly packed around the middle). The interquartile range is visually depicted as the 'box' in a *Box Plot*.

Interval scale – A *Metric* scale, either *Continuous* or *Discrete*, where the difference between consecutive values is constant, but relative only to some arbitrary zero point (e.g., SAT scores or the Celsius scale).

The arbitrary zero point for interval level variables distinguishes them from other *Metric Variables* (i.e., *Ratio* level variables) in that ratio comparisons are not accurate or useful. For example, twenty degrees Celsius is NOT twice as hot as ten degrees (converted to Fahrenheit, these temperatures are, respectively, 50° and 68°).

If a scale is metric and has negative values, it should be considered an Interval scale.

The only *Statistic* that is not available for interval scales that is useful in other metric scales, is the *Geometric Mean*.

Intervening variable – Any unmeasured variable between the hypothesized *Independent* and *Dependen*t variables that may influence the relationship.

Example: While there is a popular notion that 'education *Causes* income,' the intervening variable 'occupation' may be the more precise explanation. That is, education ➔ occupation and then occupation ➔ income (directly from Crossman, 2018).

Inverse relationship – A relationship between two variables, both of which are measured on at least an *ordinal scale*, where higher values of one tend to coincide with lower values of the other. In *correlation* or *regression* procedures inverse relationships are identified by negative *correlation coefficient* (i.e., when $r<0$).

Itemized rating scale – An *Ordinal* Level Scale wherein the values the variable may take are labelled.

Judgment Sampling – The process of building a *Sample* based upon the researcher's belief/perception that those selected are members of the *Population* of interest.

Kurtosis – Generally, a measure of the extent to which a *Distribution* has a sharp or a flat peak. A *Normal Curve* has a kurtosis measure of zero. Distributions with sharper peaks (than a normal distribution) have a positive kurtosis, distributions that are flatter have a negative kurtosis.

LCL (Lower Confidence Limit) – The lower boundary of a *Confidence Interval*.

Laboratory Experiment – The generic name for all *Experiments* NOT conducted in the normal (real world) environment of the *Population*. While such experiments *Control* for *Environmental Variables* (which may be desirable), their absence may result in an incomplete understanding of the phenomena studied.

Latin Squares Design – A type of *Randomized Block Design*.

Leading Question – A question in a questionnaire or survey that intentionally or unintentionally, blatantly or subtly, influences respondents to answer one way or another. Leading questions create *Bias* and are often considered not *Ethical*.

Level of Measurement – The level of measurement refers to the type of information conveyed by a *variable*. It refers to the different values (the "scale") that the variable may take. There are four common levels of measurement:

Nominal: A *Categorical scale* (such as Profession) where there are no quantitative differences between the possible values.

Ordinal: Another *Qualitative* scale that progresses from lowest to highest but does not define a quantitative distance between possible values.

Interval: A *Quantitative* scale wherein the difference between values is constant, but there is an arbitrary zero point (e.g., the Celsius scale for temperature).

Ratio: A quantitative scale where the difference between values is constant and there is an absolute zero point (e.g., the number of employees in a division).

See the individual entries for these levels for further explanation.

Level of Significance – See *Statistical Significance*. The highest probability (as determined by a statistical test) of concluding that the hypothesis is true when it is not (i.e., of making a *Type I error*) due to *Sampling Error*. The level of *alpha (α)* that a statistical test satisfies. Unless otherwise warranted and noted, this level is set at .05 for business and social research.

Likert scale – A basic set of response categories that can be used for any topic simply by asking *Respondents* to state the extent to which they agree with a statement. Likert scales are valuable to researchers because respondents can answer a larger number of questions (in a timeframe) when the response categories stay the same (and they do not have to recalibrate what a '1' means from question to question).

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1	2	3	4	5

While the format above is the one described by its creator, Rensis Likert, 'Likert-like' scales have also been developed with differing numbers of response categories.

Linear Regression – A mathematical/statistical procedure that defines the 'best fit' straight line that describes the relationship between an *Independent Variable* and a *Dependent Variable*. The nature and strength of the relationship is described the *Correlation Coefficient* (R), while the amount of Variance in the dependent variable that is 'explained' by the value of the independent variable is described by the *Coefficient of Determination* (R^2). Values closer to 1.00 suggest that the independent variable can be used to explain (or predict) the value of the dependent variable.

List Broker – An individual or firm that sells or rents lists of individuals typically used for direct marketing campaigns. Given an appropriate research design and target population, these lists may be adopted for research purposes. Such lists typically are collected to identify either individuals or businesses by profession or industry. The information would include some method of contacting the individual; e-mail, telephone, or street address.

Longitudinal Study – Any study where data are collected more than once. Such studies are well suited to *Causal Research* (*Experiments*) in that they permit for the observation of change in the *Dependent Variable* following a change in the *Independent Variable*; and can be useful in attempts to explain cause-and-effect.

Lower Confidence Limit (LCL) – The point that defines the lower limit of a *Confidence Interval*. The LCL combined with the UCL (*Upper Confidence Limit*) define the *Range* of the confidence interval.

Margin of Error – The name for the *Confidence Interval*, for population *Proportions*, such as *categorical* (i.e. *Binomial*) *variables*. The margin of error represents the distance above or below the sample *statistic* that the population *parameter* is expected to fall. (Confidence intervals report this as a range from lowest to highest possible score and make no mention of the sample statistic).

Margins of error are typically reported for more public consumption than are confidence intervals (e.g., for public opinion *polls*). Unless otherwise indicated, margins of error typically identify the 95% confidence interval.

The statement "Our survey results show that 53.5% of the sample will vote for candidate A (with a margin of error of 4%)" communicates exactly the same information as "the 95% confidence interval for the population proportion who will vote for candidate A is from 49.5% to 57.5%." Note that in this example, we cannot conclude that candidate A will win the election because the confidence interval includes the likelihood that the true population proportion who will vote for the candidate is less than 50%.

Maximum – The largest value observed (not the largest value possible). The maximum is a *Descriptive* measure that describes the largest value in a *Range*.

Mean – See *Average*.

Measurement – Any process of either quantifying or categorizing a set of observations according to some consistent rule.

Median – The positional middle of an ordered set of numerical or ordinal values. The median is determined by placing all of the values in ascending (or descending) order and then determining the value at which one half the values are smaller (and half the values larger). If there are an even number of observations, the median is between the middle two observations.

Note: The median is not a quantitative measure, but, rather, identifies the middle position.

Meta-Analysis – The process of combining past studies on the same topic in the attempt to have a stronger evidence to test a *Theory* or proposition or on actual *Effect Size*. Combining studies leads to large sample sizes, and larger sample sizes lead to narrower, more precise *Confidence Intervals* and smaller *Margins of Error*.

Meta-analyses need to *Standardize* the variables combined in order to make results comparable. And some caution need be employed to ensure that the *Operational Definitions* of the *Variables* from the component studies truly isolate the same *Concept*.

Metric Scale or Variable – A variable that is measured at either the *Interval* or *Ratio* level of measurement in which the values of the measure are equidistant apart (such as temperature in degrees or number of houses sold). Arithmetic operations may be performed on metric variables. Consequently, *Statistics* that require such operations (e.g., the *Mean* or the *Standard Deviation*) may be calculated (only) for metric variables.

Metric scales contrast with *Ordinal* and *Nominal* scales where the differences in the values the variable may take are unequal or qualitative. Metric variables may be *Continuous* or *Discrete*.

Minimum – The smallest value observed. The minimum is a *Descriptive* measure that describes the lowest value in a *Range*.

Mode – The category or value from a *Distribution* that has the largest number (or proportion) of observations.

Moving average – A series of *averages* that are calculated of a number of prior periods. As one period gives way to a new one, a new average is calculated, excluding the earliest period and including the newest.

Moving averages are a method of describing general trends while limiting the impact of abrupt or temporary changes. A moving average smoothes out a line in a graph and minimizes the effect of a single observation (i.e., an *Outlier*).

Multicollinearity – A rather pretentious way to say *Collinearity*.

Multiple Regression – A statistical (arithmetic) method used to estimate a *Dependent Variable* (or the formula for calculating such a variable) based upon its linear relationships with a number of *Independent Variables*.

See *Linear Regression*.

Multiple Response Question (variable) – A survey question in which the respondent may select more than 1 response as reflective of their answer. Such questions often carry with them instructions such as "check all that apply." In Internet-based research, multiple response questions are those in which the response categories are selected via "check boxes" (☐) - as opposed to radio buttons (⊙).

For analytical purposes, each response category in a multiple response question is treated as a separate *Binary Variable* and analyzed as to the *Proportion* selecting that particular response.

Multivariate – Referring to more than 2 variables. Multivariate techniques typically use multiple independent variables in attempts to "explain" the value of a dependent variable (e.g., multiple regression).

Mutually Exclusive – The (desirable) aspect of a variable wherein the *Response Categories* are unique (such that no value appears in more than one category).

For example, a variable that collects individuals' ages into the following categories is NOT mutually exclusive.

A. 20 or younger
B. 20 to 40
C. 40 to 60
D. 60 or older

Note that individuals who are 20 (or 40 or 60) years old are represented by two categories; both A. and B.

Only variables whose response categories are both mutually exclusive and *Exhaustive* are appropriate for research and analysis.

Negative Correlation – A (linear) relationship between variables wherein, when one goes down (has a lower value) the other tends to go up (has a higher value). For example, as the temperature goes down the amount of heating oil used goes up.

More formally in statistics, a *Correlation Coefficient* between -1 and 0. Contrasts with *Positive Correlation.*

Nominal Scale – The most basic *Level of Measurement* through which *Variables* record the values of *Observations*. A variable that measures differences based upon a nominal scale will have simple categories to distinguish between different types of observations (e.g., "male" and "female"). By definition, there are no quantifiable differences between the categories themselves.

Non-Experimental Designs – Those ostensibly *Causal Research* designs that do not include a *Control Group* and/or measure different values of the *Independent Variable*. It is difficult to eliminate most threats to *Internal Validity* without a control group. Likewise, it is impossible to make an attribution of *Cause-and-Effect* that does not show different values of the *Dependent Variable* for different values of the independent variable. Hence, true *Experimental Research* designs are preferred over non-experimental designs.

Nonparametric tests (statistics) – Statistical tests wherein assumptions of the *Central Limit Theorem* (i.e., the normal distribution of the sample statistic) are not met (so, basically, tests of nominal and ordinal level variables). *Chi-Square* (for *Categorical* variables) and *Spearman R* (correlation for *Ordinal* variables) are examples of non-parametric tests.

Other non-parametric tests include: Wilcoxon Signed-Rank Test, the Mann-Whitney Test and Kendall's Tau.

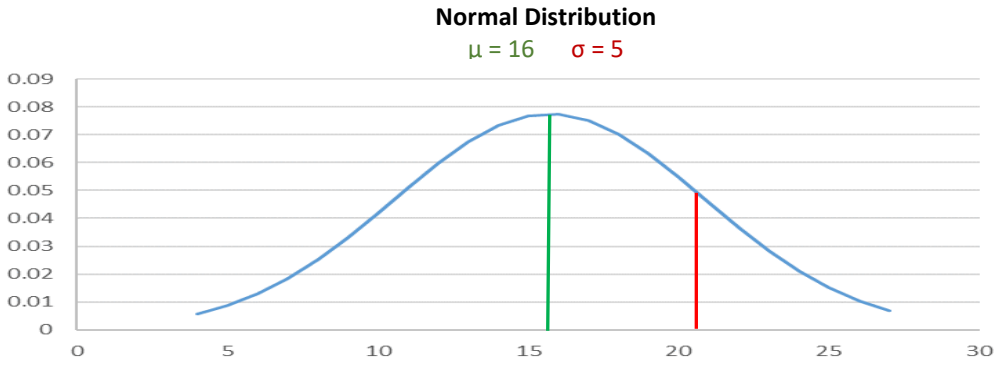
Non-random Sampling – Sampling that does not ensure that every member of the population has an opportunity to be included in a measured sample, and, thus, is subject to *Bias*.

Pure *Random Sampling* can only be possible when every member of the population can be identified (and tagged for inclusion). In the vast majority of research situations this is either impossible or impractical. To minimize bias in these situations it is the responsibility of the researcher ensure to the best degree possible, that the sample fairly represents the population; that all relevant characteristics of the population appear proportionately in the sample. This may be

accomplished through the process of randomly sampling from a large (relevant) *Sampling Frame*.

Non-Response Bias – The bias that results from a sample where the individuals who do not participate are systematically different from those who do participate. This may occur as the result of a potential participant's unwillingness to share their opinions. The flip side of *Self-selection bias*.

Normal Curve (Distribution) – A theoretical probability distribution precisely defined by a complex formula that approximates observed phenomena. Normal curves take on the traditional *Bell-Curve* shape and are both *Symmetric* and Asymptotic (meaning that at the extremes, probabilities approach, but do not achieve, a value of zero).



A normal curve is defined by its mean and standard deviation. Due to its formula (and symmetry), the random probability that an observation falls within a specific *Range* may be calculated. For example:

5% (roughly) of the area under the normal curve is above 1.645 standard deviations above the mean, or

95% of the area under the normal curve occurs between 2 standard deviations below the mean to 2 standard deviations above the mean, which is useful for calculating a 95% *Confidence Interval*. Note that the precise value is 1.96 standard deviations, however, using 2.00 won't get you into trouble.

These areas are easily calculated. See, for example, Lane's (2013a) Z-Table (where 'Z' represents the number of standard deviations away from the mean).

Null Hypothesis (H_0) – The proposition that the *Hypothesis* (H_1; or the alternate hypothesis) is NOT true. Every hypothesis has an attendant null hypothesis and, taken together, they account for all of the possible outcomes (they are *mutually exclusive* and *exhaustive*). See *Hypothesis Test*.

If the hypothesis is that average store sales increased by more than 15%, then the null hypothesis is that average store sales increased by

15% or less. Note that the null hypothesis always contains the equivalence (=, ≤, or ≥).

If the hypothesis is that there is a relationship between advertising and sales, then the null hypothesis is that there is no such relationship.

Statistical tests of hypotheses test the null hypothesis to determine whether or not the null hypothesis is statistically likely or unlikely, and thus, should be accepted or rejected. The basis of the conclusion is the probability of making a *Type I error* if the null hypothesis is rejected. If this probability (the *p-value*) is less than *Alpha* (α), then the null hypothesis is rejected. If the probability is greater than Alpha, then the null hypothesis is accepted.

If the conclusion is to reject the null hypothesis, then there is considered sufficient statistical evidence to support the (alternate) hypothesis (Reject H_0 = Accept H_1).

If the conclusion is to accept the null hypothesis – the technical term is to "fail to reject the null hypothesis" – then there is insufficient evidence to support the claim of the (alternate) hypothesis (Accept H_0 = Reject H_1). Note that this is not the same as saying that there is evidence to conclude that the null hypothesis is true, just that there is insufficient evidence to state with sufficient confidence, that it is not true.

Objectivity – Reality, whether one agrees with it or not. The true nature of things; without *Bias.*

Observation – The process of *measuring* a sample or individual.

Ockham's Razor (Occam's Razor) – Basically, and in general, the proposition that the fewer the assumptions in an argument, the greater the likelihood that it is true. Or, when comparing multiple explanations for a phenomenon, it is probably best to stick with the one that makes the fewest assumptions. See, for example, Occam's Razor (esthermsmth, 2017).

Odds – The *Probability* that something will happen divided by the probability that it will NOT happen.

While the probability that a coin flip will result in a 'heads' is 0.50, the odds that the result is a 'head' is (0.5/0.5=) 1 or 1:1. The odds that a single die roll will result in a 2 is 1:5.

This description defines what is commonly called 'odds for.' If one is interested in determining the odds that something won't happen ('odds against') the calculation is reversed; the probability that is won't happen divided by the probability that it will.

For example, the payout odds in horse racing reflect the amount of money wagered on other horses divided by the amount of money

wagered on the horse of interest (usually rounded). This is an 'odds against' calculation.

Odds Ratio – The ratio of a) the probability of having some trait after some stimuli to b) the probability of having the trait without the stimuli. Odds ratios are used in certain statistical procedures (e.g., logistical regression) as a measure of *Effect Size*.

One-Shot Case Study – A *Cross-sectional Study*. Research that is conducted all at a single time. This type of study precludes the measurement of change.

Open-ended Question – A question designed in such a way as to not limit the responses possible. Open-ended questions are often difficult to quantify and are, thus, more suitable to Qualitative Research than Quantitative.

"What is your favorite television program?" is an open-ended question, as the possible responses are quite large in number and not dictated by the way the question was worded.

Operational Definition – A definition that serves to make a vague *Concept* understandable and consistent for the purpose of a specific conversation or study (e.g. high employee morale or CEO hubris). Such definitions often involve *Indicators* (see *Construct Validity*) and the development of *Variables* so the concept may be used (measured) as a variable during research.

Different studies may use different operational definitions for the same concept, which may lead to conflicting results.

Opt-in – The process by which individuals assent to participate in research, typically *Research Panels*. In order to avoid ethical concerns during the sample selection process (e.g., the impression of email spam), professional research organizations will require prospective research participants to acknowledge and agree to their participation in the research prior to their inclusion in studies.

Order Bias – A phenomena where survey respondents with little or no preference are apt to select options at the beginning of a sequence or list, rather than those at the end of the list. Consequently, those options appear (in analysis) to have greater incidence than are present in the *Population*. Order bias can be controlled for through the use of a *Split Ballot*, where the responses are reordered from one version to another to spread out the impact of such bias.

Ordinal Scale – A *Variable* (*Level of Measurement*) scale in which the values that the variable can take (*Response Categories*) vary in an increasing, but unmeasured, magnitude (such that the differences between these values are not defined).

The variable that measured the order of finish for the 1973 Belmont Stakes showed:

1st Secretariat
2nd Twice a Prince
3rd My Gallant
4th Private Smiles
5th Sham

This scale does not indicate (measure) the magnitude of the amount by which Secretariat beat Twice a Prince (31 lengths).

Likewise, an ordinal variable may be used to measure educational attainment on the following scale:

1 Less than High School
2 High School or GED
3 Some College
4 Bachelor Degree
5 Graduate Degree

It would be difficult to quantify the difference in education (say in years) between someone answering "1" on this scale and another answering "2." We can state that all 3s have more education than the 2s, and less than that of the 4s, but we cannot quantify the difference.

The appropriate way to analyze Ordinal Scale variables is to treat each response category as a categorical variable and calculate proportions for each response category, hence averages and standard deviations are less representative of these features. A form of correlational analysis (*Spearmen Correlation*) has been developed to conduct analyses of the relationships between Ordinal variables.

Outlier – Any *Observation* that is a substantial distance (on either an *Interval* or *Ratio level of measurement*) from the bulk of the rest of the observations for a particular variable.

While there is no exact measure/definition for an outlier, some researchers use as a rule-of-thumb three times the distance from the median to the 1st or 3rd quartile (depending on which side of the median the observation falls) or observations more than three S*tandard Deviations* from the *Mean*.

Outliers may be problematic for a number of reasons: 1) they may represent a mismeasurement of the attribute (e.g., a data entry error) and 2) they have a disproportionate impact on the calculation of the *Standard Deviation*. Consequently, they should always be investigated carefully.

P (P-Value) – Technically, the statistical probability that the difference between the test statistic and hypothesized value of the population, as specified by the *Null Hypothesis*, occurred by chance. (The lower the probability, the greater the likelihood that the *Alternative Hypothesis* is an acceptable explanation of the phenomenon tested.) That is, the P-Value is the probability of making a *Type I Error* if the researcher chooses to reject the Null Hypothesis.

The accepted method of Hypothesis testing works from the (preliminary) assumption that the Null Hypothesis is correct. The Null Hypothesis is only rejected if the likelihood (*Probability*) of it being true due to random *Sampling Error* is small. The P-Value tells researchers how low or high that probability is.

In traditional research practices, the P-Value of a study is compared to *Alpha (α)*. If the apparent risk of making a Type I Error (P-Value) is less than the acceptable risk of such an error (Alpha (α)), then the Null Hypothesis is rejected (the Alternative Hypothesis is given credence). Otherwise, the Null Hypothesis should not be rejected.

Practically, the P-Value is used as a guide to establish the extent to which a study gives support to the hypothesis (Alternative Hypothesis).

Typically, in research dealing with people, P-values less than .05 are considered to have *Statistical Significance*. But, out of caution, many researchers prefer a more stringent level i.e., .01 or even .0050. See the discussion of *Alpha (α)*.

P-Hacking – Any of a variety of largely unethical techniques that give the appearance of scientific rigor, while intentionally avoid such rigor.

For one example; knowing that small sample sizes generate larger random variance, a researcher could conduct a large number of small-sample studies and then publicizing only the results from those studies where the 'random' results support the hypothesis, while hiding the results from those studies without such results. (See, for example, Charpentier, 2015)

Pilot Study – Research where an experiment is conducted using a generally small-sample subset of the population for study. Pilot studies are conducted prior to a full-scale experiment or implementation to ensure that the research will conclude with usable data (or actionable results)

Pilot studies serve a multitude of useful purposes in research. Instrument *Pre Tests*, for example, may be used to ensure that surveys capture relevant information. A new product concept may be conditionally tried in only a few *Test Markets*. A proposed organizational change, say in compensation systems, could be tried in only a few, select departments.

The results of the pilot study could then be used to inform a full-blown study or implementation.

Panel – See *Research Panel*.

Parameter – A quantifiable characteristic of a *Population*; for example, the arithmetic *Mean* (μ), the *Proportion* (π), or the *Standard Deviation* (σ) of the population in question.

Only in situations where a *census* has been taken are population parameters known with any precision. Otherwise, parameters must be estimated based upon sample *Statistics*.

Parsimonious – Stingy, frugal. But in research and logic it takes on an aspect of efficiency or not wasteful (see, for example, *Occam's Razor*).

Participant Observation (PO) – A type of *Qualitative Research*, typically used in *Ethnographic Research*, wherein the researcher becomes a part of the group studied to report group incidents, activities and/or behavior.

PO has the benefit of being able to directly witness behavior. Non-participant observation, for example, runs the risk that the subjects of the study would behave differently if they knew they were being observed.

However, PO runs an ethical risk of the researcher driving behavior.

Pearson Correlation (or Pearson Product Moment Correlation) – The standard form for calculating correlations between metric variables.

The standard statistics for the procedure are the correlation coefficient (as measured by "r") and the coefficient of determination (as measured by "r^2").

'People Meter' – Television set-top devices used to record and monitor viewing practices. Used in conjunction with viewer input (e.g., viewer demographics), this information is compiled from a sample of viewers across a media market to estimate viewership size and make-up and then sold to media companies (e.g., networks) and advertisers. (Nielsen Company, 2018)

Some caution is now warranted in using People Meter results as program viewership is migrating away from televisions per se, and on to web-enabled devices.

Phenomenological Research – Any of a variety of research processes that study the effects of a certain phenomenon on subjects (e.g. corporate downsizing survivors). Psychological, and nearly always Qualitative Research, such studies are typically interpretive.

Placebo Effect – The psychological effect of believing that if something is supposed to work, then it actually does. The placebo effect can thwart some types of research (into medicines for example) in that participants will

convince themselves that drug works (e.g., for pain management) regardless of whether or not it actually has any impact.

To combat this effect, a Control Group of participants is administered a drug-free 'placebo' in the same manner (e.g., pill or injection) as the Experimental Group receiving the actual drug. To further combat this problem, respondents should be unaware as to whether they have been administered the actual drug or the placebo (as part of a Blind Experiment.)

Plagiarism – The practice of claiming or inferring that someone else's effort (work, thoughts, writing, etc.) is one's own. Copying without attribution.

Poll – Another name for survey research. Typically, polls are conducted of the general population on issues of social or political interest.

Population – The entire group of interest (for *descriptive* or *inferential* purposes). The population is the group about which conclusions are drawn. Sometimes referred to as the target population.

Positive Correlation – A (typically, linear) relationship between two variables where, when one has higher values, the other tends to as well (and vice versa).

More formally in statistical relationships, a *Correlation Coefficient* between 0 and 1. Contrasts with *Negative Correlation*.

Post Hoc (Theorizing) – The practice of developing theory based upon the analysis of data.

Post hoc theorizing must be tested (on new data) prior to any acceptance. A common problem in an era where confirming studies are rarely conducted, post hoc theories are often untested (and reliance on the original data is, at best, suspect).

Post Test – The *Observation* (*Measurement*) of test subjects after the *Experimental Treatment* has been administered. Compare with *Pre-Test*.

Ideally, but not always, a post test may be used to determine the effect of changes in the *Independent Variable* upon the *Dependent Variable*.

Power – See *Statistical Power*

Pre-test – There are two uses of the term in research and analysis:

1) During the design of a research instrument (such as a questionnaire) a pre-test refers to a trial or test of the instrument on a small number of subjects to ascertain whether or not it measures the concepts desired. That is, whether the respondents understand the questions, concepts and response options.

2) During a research *Experiment*, a pre-test is the observation (measurement) of test subjects prior to the *Experimental Treatment*

(that is, before the independent variable is manipulated). A pre-test provides a baseline from which change may be measured.

For example, an experimental pre-test would be conducted to determine how much customers would pay for a product before a specific feature was added (they would subsequently be measured in a *post test* to determine the difference).

Predictor – Another name for an *Independent Variable* (in a causal relationship).

Premise – In logic and critical thinking, a statement used in the construction of an *Argument*, intended to provide evidence for a *Conclusion*.

Primary data – Original *Data* that is collected by the researchers for the specific purpose of addressing the current research project. Compare with *Secondary Data*.

Probability – The likelihood that an event occurs by random chance, calculated by dividing the number of ways the specific event could occur by the total number of any event that could occur. For example, the probability of drawing an Ace from a (shuffled, randomized) deck of cards is 4 (the number of Aces in the deck) by 52 (the total number of cards in the deck); (4/52 ≃) .077.

Probability Distribution – The distribution of a *Random Variable*. Typically theoretical, probability distributions (such as *Normal Distributions*, *T Distributions* and *Binomial Distributions*) depict the likelihood (probability) of random trial achieving a specific value (or within a specific range of values).

Probability Sampling – Another name for *Random Sampling*.

Proportion – The decimal equivalent of a percentage (e.g., .25 = 25%). Proportions are calculated by dividing the number of items in a particular group by the total number of items.

For example, the proportion of yellow jelly beans in a bag would be calculated by dividing the number of yellow jelly beans in the bag by the total number of jelly beans in the bag.

Proportions always fall within the range between 0 and 1.

Psychographic Variables – Personality related variables such as Agreeableness (and other *Five Factors Model* variables) and/or those related to Interests, Activities and Opinions. These variables are frequently used to *Segment* populations (often for marketing purposes).

Publication Bias – The practice (and ethical concerns thereof) wherein only research that has a positive (and interesting) result is deemed worthy of publication and, therefore, wider dissemination. Thus, research which may be helpful in identifying, say, theoretical dead-ends, may not receive wide dissemination.

Since publication has important consequences for researcher's careers, publication bias can lead to a variety of unethical behavior by authors: swapping *Alternative* and *Null Hypotheses*, adding or removing observations (*Fabrication* or the altering the *Sample Size*), and other, *P-Hacking* type activities. See, for example, O'Boyle, et al (2014).

Pyramid of Evidence – A device used to identify the credibility of information sources. Typically, in increasing levels of credibility, the pyramid has the following levels (from least to most *Credible*):

Expert Opinion
Descriptive Studies (e.g., case studies or descriptive research
Uncontrolled *Longitudinal Studies*
Controlled Longitudinal Studies (e.g., Pre-test Post Test designs)
Randomized, Controlled Longitudinal Studies

Note that many versions of the Pyramid place *Meta-Analyses* at the apex.

Q-sort – A process of assessing the relative values of multiple attributes through participant/respondent sorting different combinations of those attributes into some preferred order.

Qualitative Research – Research that collects information that is difficult to quantify (such as through *Focus Groups* or *Depth Interviews*). Such a process typically requires significant researcher interactions with study subjects and, consequently, the number of subjects actually studied tend to be low (i.e., too small for inferential statistics of any accuracy).

Qualitative research may be helpful in the generation of *Hypotheses*, but should not be used to draw conclusions about a *Population* (as any sample is too small, and likely not representative of the larger population). In such cases, the *Inferential* process is corrupted.

Types of qualitative studies include: *Focus Groups*, *Depth Interviews*, *Phenomenological Studies* and *Ethnographic Research*, among others.

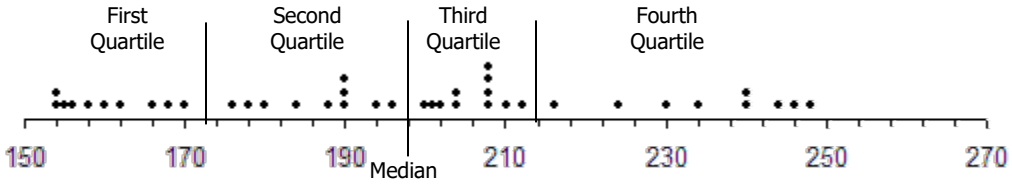
Qualitative Variable –A variable that is measured on either a Nominal or Ordinal scale and, consequently, for which analysis is (largely) restricted to assessments of *Proportions*.

Quantitative Research – Research that collects data from *Qualitative* and/or *Quantitative Variables*.

Quantitative Variable – A *variable* that is measured on either an *Interval* or *Ratio* level of measurement; a *metric* variable. Arithmetic Descriptive Statistics such as Minimum, Maximum, Range, Average (Mean), Standard Deviation, etc. are typical analyses of quantitative variables.

Quartiles – A division of the respondents/observations into four groups of equal size based upon their ranked values. The first quartile contains the lowest 25% of the population and the second quartile is the next 25%

(from the top of the first quartile to the *median*). The third quartile is the 25% of the observations immediately greater than the median and the fourth quartile is the highest 25% of the observations.



In the *Dot Plot* above, each quartile contains 10 of the 40 observations.

- The first quartile contains the 10 observations from 154 (the *Minimum*) to 173.
- The second quartile contains the 10 observations from 173 to 198 (the median).
- The third quartile contains the 10 observations from 198 (the median) to 211.
- The fourth quartile contains the 10 observations from 211 to 248 (the *Maximum*).

As with the median, there is no arithmetic involved in determining quartiles, they merely represent the location of the various 25% divisions of the observations.

Quintiles – Similar to *Quartiles*, a division of the data into five groups of equal size (20% of the total) based upon their ranked values.

Quota – A predetermined number of respondents to a research study, typically by sub-population. Quotas are often used in research to ensure that sufficient data are collected for useful statistical inference, while limiting data collection costs. When the numerical quotas do not match a population distribution, *Data Weighting* is employed.

Random Digit Dialing (RDD) – A process by which a random 4 digit number (XXXX) is appended to a known Area Code (YYY) and prefix (ZZZ) combination to arrive at random telephone number for survey research:
(YYY) ZZZ-XXXX.

A large number of these random telephone numbers are then called in order to obtain a random sample of households within the geographies covered by the Area Codes (or Area Code-Exchange combination). Typically, these numbers are loaded into a *CATI* system.

Note that there is no guarantee that the RDD telephone number generated will actually be answered by someone researchers would want to interview. There are a large number of other possible outcomes. For example, the telephone number generated could be

"Not in service," a fax or a modem, or be a business line. Combined with the very low response rates for telephone surveys, RDD results in a need to generate (and call) a relatively large number of telephone numbers to obtain sufficient sample size for analysis.

Random Sampling – *Sampling* methods in which every member of the *Population* (or every member of the *Sampling Frame*) has a fair opportunity to be selected for measurement/observation.

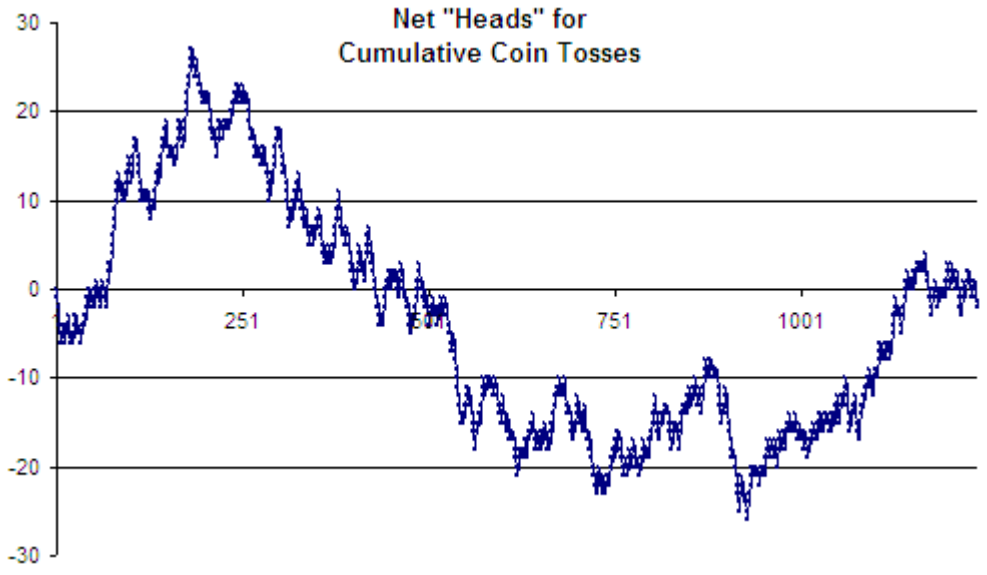
Random sampling is undertaken, where possible, to ensure that the sample is a fair representation of the population.

In situations where it is impossible to identify every member of the population (e.g., Red Bull drinkers), random sampling must be approximated, typically by randomly drawing the sample from as large as possible group of identified members of the population.

Random sampling is typically performed by using a random number generator to select individuals for inclusion in a study. See for example, the '=rand()' and '=randbetween(b,t)' functions of Excel™.

Random Variable – A variable whose value is determined by some random process (for example, the number of people entering a store on any particular day). Random variables are often modelled by *Probability Distributions*.

Random Walk – Any process in which the outcome of one event does not affect the results of the immediate subsequent events. The cumulative effect of these independent events may give the impression of a trend or a pattern, when no such pattern actually exists. It is important to consider the possibility that the appearance of a trend is, in all actuality, a random walk.



The chart above shows a random walk generated by the cumulative difference between heads and tails for 1237 coin tosses. The first toss

was a tail, and the net number of heads was -1. The second toss also resulted in a tail and the net number of heads was now -2. And so on.

At its peak (at toss #185) there were 27 more heads than tails; 106 heads versus 79 tails. At its lowest (toss #928) there were 26 more tails than heads; 477 tails versus 451 heads. However, after the entire 1237 tosses of the coin, there had been only one more tail than head; 619 tails and 618 heads.

It is easy to see that by looking at this chart one might assume that there was a trend for some portion of the tosses, or cyclical behavior for the entire period. Yet, no such phenomenon exists here; this is purely the product of a series of random events.

For a discussion of random walks and stock prices see A Random Walk Down Wall Street (Malkiel, 2015).

Randomized Block Design – An *Experimental Design* for some statistical tests (i.e., *ANOVA*) that allows for the assessment of the impact of possibly confounding *Factors* on the *Dependent Variable*.

Randomized block designs typically group the members of the *Sample* into *Homogeneous* groups (called 'blocks') prior to the random assignment of *Treatments*, in the attempt to better understand the impact, if any, of the Independent Variable.

Range – A measure of *Dispersion* that represents the magnitude of the difference between the *Minimum* and *Maximum*. The range is calculated by simply subtracting the lowest (observed) value from the highest observed value. The range is only appropriately measured for *metric variables*.

Rank-order or Rank-difference correlation – See *Spearman Correlation.*

Ratio Scale – A *Metric Level of Measurement* (for a *Variable*) that has a true zero point and the differences between allowable values is constant throughout.

For example, the exact number of days each employee was absent from work over the past 6 months is a variable that is measured on a ratio scale (0, 1, 2, 3, 4, … 75).

Only variables measured on a ratio scale may be used to draw ratio comparisons between individuals or groups. E.g., an employee who was absent 2 days during the past six months was absent ½ as much as a person who was absent 4 days or the average person in their twenties dines out three times as often as the average person in their forties. This is NOT a feature of the other levels of measurement.

Means (μ, $\bar{X}$) and *Standard Deviations* (σ, s) are the standard *descriptive statistics* for variables that are measured on a ratio scale.

Other statistics used to describe ratio variables include *Skew* and *Kurtosis*.

Rational Model – A model of decision making that holds that individuals make decisions in a rational manner:

1) They clearly understand the problem to be solved (decision made)
2) They identify all possible courses of action to effectively address the issue.
3) They objectively assess these courses of action to determine the magnitude and likelihood of their potential success.
4) They select and implement the course of action with the greatest potential success.

Note: Herbert Simon pointed out the differences between this rational model and the way that people actually behave with his *Bounded Rationality* (Simon, 1978).

Refusal – A potential survey respondent who declines to participate in a study. Refusals are a major source of *non-response bias*. See *Response Rate* for the implications of refusals on the *Sample Size.*

Regression – See *Linear Regression*.

Regression Coefficient – The *Slope* of a *Linear Regression equation*; the incremental change in the *Dependent Variable* per unit of change in the *Independent Variable*, typically, as established through a *Regression* equation.

Regression is short for "Linear Regression", the best estimate of the straight-line relationship between variables. A straight line is defined through the equation y = (ax) + b, and, thus, provides an estimate of the value of the dependent variable (y) given a value of the independent variable (x). The slope of this relationship (the constant 'a') is the *Regression Coefficient* and represents the incremental increase (or decrease) in y for values of x. Note: b, the *Intercept*, is another constant that represents the value of the dependent variable when the independent variable has a value of zero.

For example, a regression equation for estimating income could take the form: Income = ($4000 * Number of Years of Education) + $20,000. In this case, $4000 is the Regression Coefficient.

That is, the magnitude of the change in the *Dependent Variable* for each incremental change in the *Independent Variable*.

Reliability – The extent to which an instrument or practice yields consistent results. Note that consistent ≠ accurate or correct.

There are a variety of imperfect 'measures' of reliability: e.g., inter-observer reliability or test-retest reliability.

Research Panel – a *Sampling Frame* made up of a group of individuals who have consented to participate in a number of research studies. Research panels fall into three broad categories: *consumer panels*, *B2B (Business-to-Business) panels*, and *custom* or *specialty* panels. Panels may be maintained by research firms, for rent to clients, or by individual firms, for their private use. Ethical standards

Response Bias – Any of a number of processes by which a participant's responses fail to accurately represent their true measure of the variable. For example:

- a respondent may not want to hurt the feelings of the questioner, and, hence, answer questions that try to make the questioner feel good about their product or service
- the respondent may fear retaliation for an unfavorable response.
- the respondent may be answering an unrelated question. (A question that asks about the quality of company parking facilities is likely to elicit more positive responses on the day after a substantial pay raise is announced than on the previous day.)
- the respondent does not know (or cannot remember) the accurate response.
- The respondent is indifferent to the accuracy of their response (e.g., out of boredom).

Response Categories – The allowable answer options to a survey item (question). Depending on the *Level of Measurement*, the possible values that the measured variable may take. Note, for Valid research, response categories must be *Mutually Exclusive* and collectively *Exhaustive*.

Response Rate – The proportion of those who are asked to participate in a study to who actually do (i.e., *Refusals*). Since the sample size is determined by the number of actual respondents, and NOT the number invited to participate, response rates are an important consideration when planning studies.

For example, if your desired sample size was 1,000 and the response rate was expected to be 20%, the 5,000 individuals would be needed (1,000/20% = 5,000). On the other hand, if the response rate was expected to be only 5%, then the number invited would need to be much larger: (1,000/5% =) 20,000.

Low response rates are challenging for researchers in that they: 1) raise concerns about the potential for meaningful (research related) differences between those who respond and those who do not, and 2) increase data collection costs, often dramatically.

Response Variable – In many statistical procedures (i.e., *ANOVA*, ANCOVA), the response variable is the *Metric* level *Dependent Variable*. The 'effected' variable in a *Cause and Effect* relationship.

Rule-of-Thumb – See *Heuristic*.

Sample – A group of items (people, days, hotels) that are observed and/or measured during research. Ideally, a sample will closely match the *Population* of interest as to its characteristics; see *Random Sampling*.

Due to the nature of the calculations involved for *Statistical Inferences*, larger samples tend to be more useful (accurate, precise) than are smaller ones. (Boyd, 2006)

Sample Size (n) – The number of entities measured in a study.

The sample size plays a critical role in most inferential statistics. Typically, the larger the sample size, the better the estimation of the population characteristics of interest.

Note: During survey research, the sample size is NOT the number of people sent invitations or contacted by telephone. It is the number of individuals from whom data is collected. (Boyd, 2006)

Sample Bias – Systematic differences between the sample used for evaluation and the population about which conclusions are to be drawn. Sample bias could well result in incorrect conclusions.

For example, an election study that used a land-line based *panel* for a sample of voters would suffer from sample bias. These panels do not have the same make-up as the voting population because young and frugal voters are not well represented (as they rely solely on mobile telephones).

S*elf-selection Bias* (*Non-Response Bias*) is a form of sample bias.

Sampling – The process of selecting members of the *Population* of interest who are to be studied (*observed*). For *Descriptive* and *Causal Research,* it is important – for purposes of accuracy – that the sample be representative of the population; at least as to the characteristics (*variables*) of interest. (Boyd, 2006)

Sampling Distribution – The, typically, hypothetical distribution of a sample statistic (say a sample *Mean* or *Proportion*) containing the statistic for all possible samples (of the same size) drawn from the population.

For example, in a population of three individuals (A, B & C), there are three possible samples of size two (A&B, A&C, B&C). If we took the average age of each of these three pairs, the three resulting values would represent the sampling distribution of the size 2 for this population (and could be plotted).

For large populations, sampling distributions take on shapes best approximated by the *T-Distribution* (for means) or the *Normal Curve*. For proportions, this distribution is approximated by the Normal

Distribution, but is exactly defined by a *Binomial Distribution*. See *Central Limit Theorem* and *Z-Approximation*.

Sampling Error – The (random) difference between a sample *Statistic* and the corresponding population *Parameter*. Generally, the parameter is unknown and, consequently, the amount of sampling error is not known. However, sampling error may be estimated through the *Standard Error*.

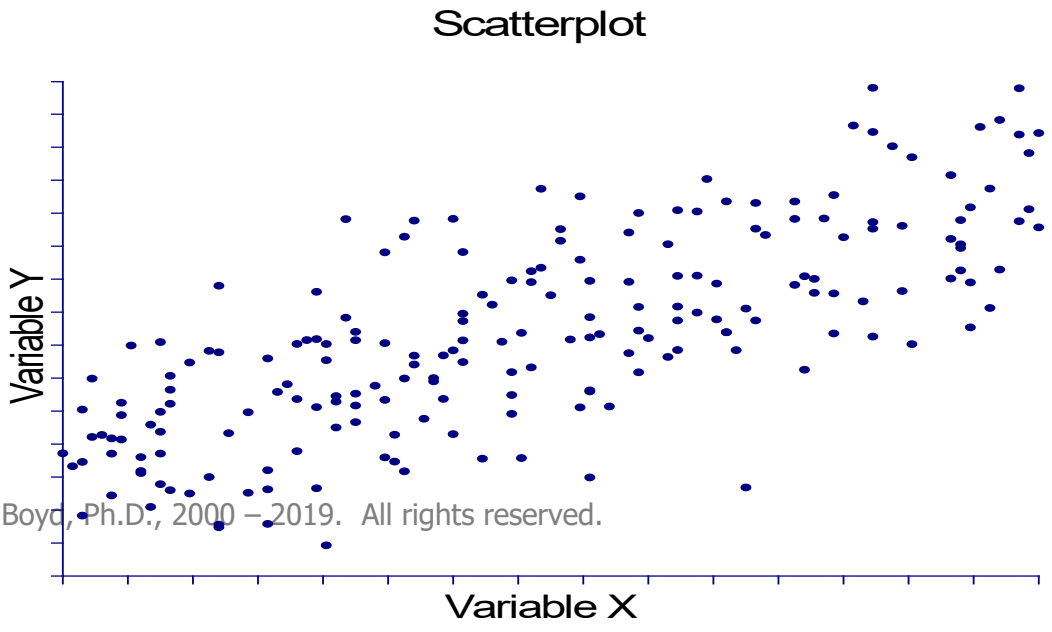
Sampling Frame – An identifiable subset of a *Population* from which a *Sample* is randomly drawn or selected. Sampling frames are typically lists from which a random sample may be drawn.

A "good" sampling frame is one that is representative of the population and not *Biased*. That is, the group from which the sample is selected should not over- or under-represent any segment of the population. Frequently utilized frames (e.g., *Research Panels*) should be sufficiently large as to permit *Random Sampling* (that individual members be not repeatedly selected into the sample). (Boyd, 2006)

Satisficing – The name for the common process described by Simon (1978) wherein the search for solutions (courses-of-action, etc.) stops when one is found that meets a minimal threshold of acceptability. [The 'rational model' of decision making would have this search be exhaustive.] Consequently, decision making tends to be sub-optimal (as better solutions are not uncovered).

Scaled Question – Any question that captures a *Variable*, other than *Nominal*, wherein the response categories take on a pre-established range of values (typically greater than 3). See, for example, *Likert Scale*.

Scatterplot – A *Bivariate* graph or charting technique in which, for each observation, a dot is placed at the intersection of the value for the X variable and the Y variable. The resulting chart may be useful in identifying the nature of the relationship between the variables.



Screening Question – A question employed at the beginning of a survey that is designed to ensure that the *Sample* accurately represents the *Population*. Screening questions are typically employed to screen out individuals who are not members of the population of interest.

Secondary data – Any data that has been repurposed from its original intent or design.

Secondary data is beneficial in that is saves data collections costs. However, caution should be used when using such data to ensure: 1) that it employs substantially similar *Operational Definitions* and 2) that the data are not out-of-date. Contrasts with *Primary Data*.

Segment – Statistically *Homogeneous* subgroups of interest in a population. See *Segmentation*.

Segmentation (segmentation studies) – Research studies designed to identify various subgroups of a population. For example, the group(s) with the highest propensity to purchase a product or service. Segmentation studies typically employ the use of *Psychographic*, *Demographic*, *Technographic*, *Life Style*, and/or *Firmographic* variables to define the relevant segments.

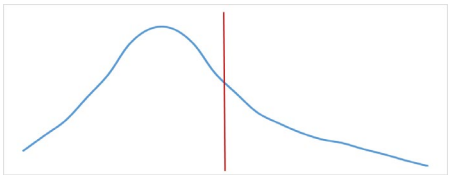
Self-selection Bias – A type of *Sample Bias* in which individuals who have more interest in the research topic or subject are more likely to volunteer to participate in the research than are those with less interest. Consequently, many survey results tend to be biased toward the opinions of those with more extreme views than the general population.

Essentially, this is the same effect as *Non-response bias*.

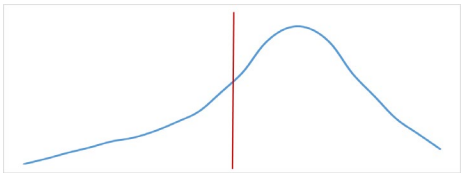
Semantic Difference Scale – A research scale used to capture qualitative/ impressionistic information from subjects wherein the scale is anchored at each extreme by antithetical concepts (e.g., 'modern' and 'old-fashioned') rather than mere negations ('new' and 'old').

Significance – See *Statistical Significance*. For practical purposes, the *Probability* of being incorrect if the *Hypothesis* is deemed to be "true" (on the basis of statistical analysis). Smaller is better (e.g., < 0.05).

Skew – Any of a number of statistical measures intended to show the relative level of *Asymmetry* of a *Metric Variable*. A variable with a positive skew has most of its observations lower than its mean. A skew statistic with a value of zero is *Symmetric*. A variable with a negative skew (value less than zero) would have the bulk of its observations above its *Mean*.



Positive Skew



Negative Skew

Generally, A *Distribution* has a positive skew if its Mean is greater than its *Median* and a negative skew if its mean is less than its median. (If the mean and the median are identical, the distribution is apt to be *Symmetrical* and will have a skew of zero.) As it is a generally relative term, there is no standard definition of what constitutes extreme skew. However, distributions with skew of 0.5 or higher (or -0.5 or lower) are observably skewed.

Skip Pattern – The logical pathways through a questionnaire or survey that account for questions (variables) excluded or included for various respondents.

For example, respondents who have been identified as having a college education through a "level of educational attainment" question may be subsequently asked to identify their specific degree (B.S, B.A, etc.) while those without a college education would not be asked this (they would "skip" this question).

Skip patterns are often initiated through *Branching Questions* (for example, "level of educational attainment" question mentioned above).

Slope (Regression) – The *Regression Coefficient*; the incremental change in the *Dependent Variable* per unit of change in the *Independent Variable*, typically, as established through a *Regression* equation. The 'a' in the formula for a straight line:

$$y = ax + b$$

Spearman Correlation – An attempt to provide a method for relating variables quantitatively when at least one of them is measured on an *Ordinal* (e.g., rank ordered) versus a *Metric* scale. [An example would be attempts to correlate 'satisfaction' and 'intent to purchase' when both are measured on a *Likert Scale*.]

The statistic calculated is ρ (rho), the Spearman Rank Correlation Coefficient. It is analogous to the traditional (Pearson) Correlation Coefficient (r) in that can take values from -1 to +1. Some caution should be used in interpreting any quantitative analysis based upon ordinal scales, as an implied metric is not, in fact, present.

Specialty Panel – See *Custom Panel*.

Split Ballot – Any practice wherein two or more versions of the survey instrument are presented to the sample. Split ballot techniques are typically used to

control for the effects of *Order Bias* by reordering the items for portions of the sample such that its effect is spread amongst all options.

Spurious Relationships – Relationships among variables (e.g., *Correlations*) that appear to exist based upon statistical evidence in a particular data set that in fact do not exist in the general population. The two most typical causes of these false relationships are random attributes of the *Sample* and the presence of *Intervening Variables*.

Standard Deviation (σ or s) – A measure of *Dispersion* of a *Metric* variable. The formula for calculation the standard deviation is slightly different depending upon whether the data are from a *Sample* or a *Population*. The standard deviation is calculated by taking the square root of the *Variance* (see the entry there for the formulas).

Due to the nature of the formulae for calculating the standard deviation that squares the difference between each observation and the mean, *Outliers* (observations farther away from the mean) are given a higher weight. Consequently, the standard deviation is always greater than the other practical measure of dispersion, the *Average Deviation*.

Knowledge of the sample standard deviation (s) is critical for the development of *Confidence Intervals* and statistical *Hypothesis Tests*.

Standard Error (SE) – An estimation of the dispersion of a sample *Statistic* (i.e., the standard deviation of that statistic).

The Standard Error of a sample mean ($SE_{\overline{x}}$) is the *Standard Deviation* of the *Population* divided by the square root of the sample size:

$$SE_{\overline{x}} = \frac{s}{\sqrt{n}}$$

The Standard Error of a sample proportion (SE_p) is the square root of sample proportion times 1 minus the sample proportion divided by the sample size:

$$SE_p = \sqrt{\frac{p * (1-p)}{n}}$$

The standard error is used in the construction of *Confidence Intervals* and for *Tests of Hypotheses*. A 95% confidence interval would extend 1.645 Standard Errors on either side of the sample statistic if a standard normal distribution is used for the estimator. A somewhat larger number would be used if a *t* distribution is assumed.

Note: some caveats are present. For $SE_{\overline{x}}$ the sample size must be greater than 30 (50 is preferable). For SE_p n * p and n * (1 – p) should be greater than 10 (15 is preferable).

Standard Normal Curve – The *Normal Curve* with a *Mean* of zero (0) and a *Standard Deviation* of one (1.00).

Any normal curve may be converted to the standard normal curve by taking each observation and subtracting the *Mean* and then dividing by the *Standard Deviation*. (This process of converting values from a normal curve to those of a standard normal curve is also called a *Z Transformation*.)

The value of using a standard normal curve for estimation and inference is that the proportionate areas under this curve have long been known. In fact, most statistics texts contain a table of the "Area under the normal curve." Such a table describes the proportion of the cases or observations that fall between the mean (0.00) and any number of multiples of the standard deviation away from the mean. For example, 47.5% of all of the observations fall between the mean and 1.96 standard deviations above (or below) the mean of a standard normal curve.

Standardize – Any method for providing a common scale of measurement for variables. Most commonly, a *Z-Transformation (the* process for converting a data set to one with a *Mean* of 0.00 and a *Standard Deviation* of 1.00).

Statistic – A measurable, summary characteristic of a *Sample*; for example, the *Arithmetic Mean* ($\overline{X}$), the *Proportion* ($\hat{p}$), or the *Variance* (s^2).

Also, any value calculated on the basis of such a summary characteristic (e.g., chi-square, Student's t, the F ratio, etc.).

Statistical Power – The ability of a statistical test to accurately identify relationships, etc., when they truly exist.

More formally, statistical power is the ability of a statistical test to correctly reject the *Null Hypothesis* (and thus, the propensity to NOT make a *Type II Error*). It is mathematically calculated (and defined as) 1 – *Beta* (the probability of making a Type II error) and, may take on any value between 0 and 1.

Generally, statistical power measures above .5 are considered noteworthy and those above .8 are considered sufficient for drawing conclusions.

Statistical Significance – The likely probability that any effect/result noted is the result of *Sampling Error* (and, thus, not a confirmation of any *Hypothesis*). Basically, the *Probability* of making a *Type I Error* if the *Null Hypothesis* is rejected.

Lower levels of significance lend support to rejecting the null hypothesis; but 'How low' is low enough is a decision for the researcher. See the discussion at *P-Value*.

Stratified Sampling – The process of separating the *Sampling Frame* into *Homogeneous* groups (based upon some *Factor*) and then *Random Sampling* within (or the random assignment of *Treatments* to) each group. Such a process allows for the determination as to whether the factor is a source of *Variance* in the *Dependent Variable*. Stratified sampling is analogous to the *Randomized Block Design* in some Analysis of Variance (*ANOVA*) procedures.

Student's *t* – See *t Test*.

Symmetric – A *Distribution* where one side is the mirror image of the other. For example, a *Normal Distribution* is symmetric, while a *Skew*ed distribution is not (it is *Asymmetric*). In Symmetric distributions, the *Mean* and the *Median* share the same value.

Syndicated Research – Research conducted by a research firm and sold to a variety of customers. Syndicated research comes in a number of forms (e.g., *Tracking Studies*, or '*People Meters*').

Syndicated research typically costs less than *Custom Research*, in that the costs may be amortized over the range of customers. Given that others see the data, these studies tend to be rather generic in nature.

Tabs – See *Cross-tabs*.

Target Population – The *Population* of interest for the current research. The group about which conclusions are to be drawn or decisions made.

Technographic Variables – An informal term for a series of descriptive measures that identify individuals' or corporate use of and comfort with various types of technological products (typically "high technology").

Termination – The process of ending a survey prior to its completion.

Respondent termination is problematic in that often insufficient information is collected from the respondent (resulting in "missing data" which makes analysis difficult). It also incurs data collection costs.

Researcher termination is the process of ending the data collection process from an individual where the data is irrelevant to the study; such as Screening out respondents who are not part of the target *Population*.

Test Market – A subset of the broader market used to *Beta Test* a product or service prior to a full roll-out. Test markets serve as real life (true population) samples that are better able to assess the product acceptance than in-house or laboratory studies. The results test market studies may be used to inform on the likelihood of success of a nationwide product introduction (with its associate manufacturing, marketing and distribution expenses).

Test markets are typically defined as geographies defined by media markets (e.g., the Portland television market) or by other geographic delineators (e.g., franchisees served by the Rhode Island distribution center).

Theory – A logical explanation of a phenomenon or concept which has not been disproved. A plausible explanation of cause and effect. If true, the theory would clarify and help to explain the relationships between variables.

Theories should be trusted to be true only to the extent to which they have been tested, preferably through *Experiments.* That is, some theories have been repeatedly tested and found credible (e.g., speciation through evolution), others, not so much (e.g., 'Intelligent design,' which, by its very nature, cannot be tested. See, for example, Rovelli (2014, page 211) and Tyson (2012).

Threats to Validity – Aspects of the research process that make it difficult to draw conclusions regarding the question at hand. For example, using a *Biased* instrument, or a non-representative *Sample*, or the accidental measurement of some extraneous factor. See *Internal* and *External Validity*.

Tracking Study – A series of studies that measure the same variable(s) on a regular basis (such a Presidential favorability or consumer confidence).

Treatment – Different levels or values of the *Independent Variable* in an *Experiment*. Differing treatments are tested in order to establish which, if any, is more effective.

Trial – An *Experiment*.

t Test – (More formally, **Student's t**) Any test of *Significance* that uses one of a series of distributions that theoretically describe the *Sampling Distribution* of statistics such as the sample mean. T-distributions are similar in shape to a Normal Curve, but are flatter (e.g., more spread out). As the number of objects considered (e.g., sample size) increases, the distribution approximates more closely a normal curve. Traditionally, when the number of objects considered surpassed 30, researchers simply used the normal distribution as a surrogate for t-distribution. Today, however, computing power permits the more accurate use of t in virtually all situations.

Type I Error – Sometimes called a False Positive error, an incorrect research conclusion that supports the *hypothesis* when that hypothesis is not true. More formally, a Type I error is the "rejection of the *Null Hypothesis*" when, in fact, the null hypothesis is true.

Since the truth or falsehood of a hypothesis is unknown and our tests of hypotheses are imperfect (due to *Sampling Error*, etc.), there is

always the chance that statistical conclusions are incorrect. If they are incorrect in such a way as to conclude that the hypothesis is true when it is not, the name of that error is a Type I error.

The acceptable probability of making a Type I error during research and analysis is called *Alpha (α)*. The probability of making a Type I Error, as calculated from a sample, is the *P-Value*.

Type II Error – Sometimes called a False Negative error, the incorrect research conclusion that the *Hypothesis* (i.e., the alternate hypothesis) is not true when, in fact, that hypothesis is true. More formally, a Type II error is the acceptance of the *Null Hypothesis* when, in fact, the null hypothesis is not true.

Since the truth or falsehood of a hypothesis is unknown and our tests of hypotheses are imperfect (due to *Sampling Error*, etc.), there is always the chance that statistical conclusions are incorrect. If they are incorrect in such a way as to conclude that the hypothesis is false when it is not, the name of that error is a Type II error.

The probability of making a Type II error during research and analysis is called *Beta (β)*.

Note: *Statistical Power* is calculated as 1 - β.

Type III Error – Solving the wrong problem (Mitroff and Featheringham, 1974). Caution should be taken during the formulation stage of decision making or problem solving to ensure that the appropriate concepts and relationships are addressed.

UCL (Upper Confidence Limit) – The upper boundary of a *Confidence Interval*.

Unbiased – Basically, *Objective*, not subject to *Bias*. That is, the data collected accurately represent the attitudes/perspectives, etc., of the population about which conclusions are to be drawn.

Univariate – Referring to a single, isolated variable. A univariate analysis would only provide descriptive statistics of the variable in question. Compare with *Bivariate* and *Multivariate*.

Unobtrusive – Research conducted in such a manner as to not interfere or influence the individuals being studied. People behave differently when they know they are being observed, so any information obtained from someone who is aware of it is likely *Biased*.

Unobtrusive measures include: *Participant Observation*, hidden cameras, and the like. Unobtrusive measures may raise serious ethical concerns, as those being observed may engage in behavior that they prefer not be made public. See *Ethics.*

Upper Confidence Limit (UCL) – The point that defines the upper limit of a *Confidence Interval*. The UCL combined with the LCL (*Lower Confidence Limit*) define the *Range* of the confidence interval.

Validity (Instrument) – The extent to which an *Instrument* (e.g., a questionnaire) accurately measures a *Concept*. See *Content Validity*.

Validity (Research Design) – General, *Internal Validity* (the extent to which the research is designed to isolate and measure any changes in the independent variable) and *External Validity* (the usefulness of relating the research results to the problem/decision in question). See Campbell & Stanley (1963) or Cook & Campbell (1979).

Variable – A value that changes from one observation to the next.

Variability – See *Dispersion*.

Variance (σ^2 or s^2) – A measure of *Dispersion* of a *Population* (σ^2) that is calculated by summing the square of the difference between each observation and the *Mean* and then dividing the sum by the number of observations.

$$\sigma^2 = \frac{\sum(X-\mu)^2}{N}$$

When the variance of a *Sample* (s^2) is being calculated, the sum is divided by the number of observations minus 1 (to account for the uncertainty inherent in a sample).

$$s^2 = \frac{\sum(X-\bar{x})^2}{n-1}$$

In both cases, the square root of the variance is calculated to derive the *Standard Deviation* (which is the square root of the variance).

Variance Report – Any report intended to show the difference between the expected (projected) and the observed values of some measure (such as units sold). Should not be confused with calculations of statistical *Variance*.

Weighting – See *Data Weighting*

Z Approximation – With (very) large sample sizes, the *Binomial Distribution* and the *Normal Distribution* converge; begin to take on very similar shapes with similar probabilities. In times past, when binomial calculations were cumbersome, it was more convenient to use the easier calculated Z (normal) statistic as an approximation of this binomial distribution. As computers today can easily calculate exact binomial probabilities, that method is preferred. For example, Excel™ has the function Binom.dist.range to calculate such probabilities.

Z Distribution – The distribution of the random variable that describes the values of the *Standard Normal Curve*.

Z-Score (or just Z) – The result of *Z Transformation*.

Z-Table – A table of the probabilities of random events occurring in a range (typically) either 1) between the *Mean* and a specified number of *Standard Deviations* (as represented by the *Z-Score*) above or below the mean or 2) above a specific Z-score. Today, these probabilities are more easily determined through quick, on-line applications. See, for example, Z-Table Calculations (Lane, 2013a).

Z Transformation – A simple calculation that converts a value from any *Normal Curve* into a value on the *Standard Normal Curve*.

$$Z_i = \frac{X_i - \mu}{\sigma}$$

The resulting measure (Z-Score) is the number of *Standard Deviations* the particular observation or value is away from the *Mean*. This value may be assessed using a *Z-Table*.

Source Material

Boyd, Paul C. (2002), "Sampling Concepts". MBA Faculty Conference Papers & Journal Articles. Paper 2. http://scholarsarchive.jwu.edu/mba_fac/2

Campbell, D. T. and Stanley, J. C. (1963). Experimental and Quasi-Experimental Designs for Research. Houghton Mifflin.

Charpentier, Arthur (2015). "P-HACKING, OR CHEATING ON A P-VALUE," Freakonometrics.com, https://freakonometrics.hypotheses.org/19817

Cook, T.D. and Campbell, D. T. (1979). Quasi-experimentation: Design and Analysis Issues for Field Settings. Rand-McNally Education.

Crossman, A (2108), How Intervening Variables Work in Sociology, ThoughtCo, Feb 1.

Digman, J.M. (1990), Personality Structure: The Emergence of the Five-Factor Model, Annual Review of Psychology, 41, 1, pp. 417- 440.

Esthermsmth (2017), "Occam's Razor," in Learning Theories, September 9, 2017, https://www.learning-theories.com/occams-razor.html. [Requested citation format]

Hayward, M.L.A., and Hambrick, D.C. (1997). Explaining the Premiums Paid for Large Acquisitions: Evidence of CEO Hubris, ASQ (Administrative Sciences Quarterly), 42 March, pp. 103 – 127.

Lane, David M. (2013), Hyperstat, Z-Table Calculations, http://davidmlane.com/hyperstat/z_table.html

Malkiel, Burton H. (2015, 1984). A Random Walk Down Wall Street (11th edition), W.W. Norton.

McCleary, R. and Hay, R.A., Jr. (1980), Applied Time Series Analysis for the Social Sciences, Sage Publications.

Mitroff, I.I. and Featheringham, T.R. (1974), On Systemic Problem Solving and the Error of the Third Kind, *Behavioral Science*, Vol.19, No.6.

Nielsen Company (2018), http://www.nielsen.com/us/en.html

O'Boyle E.H. Jr., Banks, G.C., and Gonzalez-Mule, E. (2014, 2017), The Chrysalis Effect: How Ugly Initial Results Metamorphize Into Beautiful Articles, Journal of Management, Vol 43, 2 (pp. 376-399).

Parry, W.T. and Hacker, E.A. (1991), Aristotelian Logic, State University of New York Press.

Rovelli, Carlo (2014, 2016 translation). Reality Is Not What It Seems: The Journey to Quantum Gravity (pp. 210-211). Penguin Publishing Group.

Simon, Herbert A. (1978) Rational Decision-Making in Business Organizations, Nobel Memorial Lecture, The Nobel Foundation. http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1978/simon-lecture.pdf

Sumner, William Graham (1940). Folkways: A Study of the Sociological Importance of Usages, Manners, Customs, Mores, and Morals. Ginn and Co., pp. 632-633.

Tyson, Neil deGrasse (2015, revised 2016), What Science Is – and How and Why it Works, The Huffington Post, Impact; original published/posted 11/8/2015.

Wikipedia (accessed Feb., 2019), List of cognitive biases (https://en.wikipedia.org/wiki/List_of_cognitive_biases).