Glossary

α-Error occurs when erroneously rejecting a true null hypothesis. Also referred to as *type I error*.

α-Inflation results when multiple tests are conducted simultaneously on the same data. The result is that you are more likely to claim a significant result when this is not so (i.e., an increase or inflation in the type I error).

Acquiescence describes the tendency of respondents from different cultures to agree with statements (e.g., as formulated in a Likert scale item) regardless of their content.

Adjusted coefficient of determination is a modified measure of goodness-of-fit that takes the number of independent variables and the sample size into account. The statistic is useful for comparing regression models with different numbers of independent variables, sample sizes, or both.

Adjusted R^2 See *Adjusted coefficient of determination*.

Agglomerative clustering is a type of hierarchical clustering method in which clusters are consecutively formed from objects. It starts with each object representing an individual cluster. The objects are then sequentially merged to form clusters of multiple objects, starting with the two most similar.

Aggregation is a type of scale transformation in which variables measured at a lower level are taken to a higher level.

Akaike information criterion (AIC) is a relative measure of goodness-of-fit, which can be used to assess the various statistical models such as regression or factor analysis. Compared to an alternative solution with a different number of variables or factors, smaller AIC values indicate a better fit.

American Marketing Association (AMA) is the world's leading association for marketing professionals.

Analysis of variance (ANOVA) is a multivariate data analysis technique that allows testing whether the means of (typically) three or more groups differ significantly on one (one-way ANOVA) or two (two-way ANOVA) metric dependent variable(s). There are numerous extensions to more dependent variables and to differently scaled independent variables.

Anti-image is a measure used in principal component and factor analysis to determine whether the items correlate sufficiently. The anti-image describes

the portion of an item's variance that is independent of another item in the analysis.

Arithmetic mean See *mean*.

Armstrong and Overton procedure is used to assess the degree of non-response bias. This procedure calls for comparing the first 50% respondents with the last 50% with regard to key demographic variables. The concept behind this procedure is that later respondents more closely match the characteristics of non-respondents.

Autocorrelation occurs when the residuals from a regression analysis are correlated.

Average see *mean*.

Average linkage is a linkage algorithm in hierarchical clustering methods in which the distance between two clusters is defined as the average distance between all pairs of objects in the two clusters.

Average marginal effects average change (in percentage) of one variable when another variable increases by one unit.

β-Error occurs when erroneously accepting a false null hypothesis. Also referred to as *type II error*.

Back-translation is a translation method used in survey research in which a survey is being translated and then back-translated into the original language by another person.

Balanced scale describes a scale with an equal number of positive and negative scale categories.

Bar chart is a graphical representation of a single categorical variable indicating each category's frequency of occurrence. Bar charts are primarily useful for describing nominal and ordinal variables.

Bartlett method is a procedure to generate factor scores in principal component analysis. The resulting factor scores have a zero mean and a standard deviation larger than one.

Bayes information criterion (BIC) is a relative measure of goodness-of-fit, which is similar to the AIC. Compared to the AIC, the BIC applies a greater penalty to statistical analysis that have a greater number of variables or factors.

Big data refers to very large datasets, generally a mix of quantitative and qualitative data in very large volumes.

Binary logistic regression is a type of regression method used when the dependent variable is binary and only takes two values.

Bivariate statistics describes statistics that express the empirical relationship between two variables. Covariance and correlation are key measures that indicate (linear) associations between two variables.

Bonferroni correction is a post hoc test typically used in an ANOVA that maintains the familywise error rate by calculating a new pairwise alpha that divides the statistical significance level α by the number of comparisons made (see also *familywise error rate* and *α-Inflation*).

Box-and-whisker plot See *box plot*.

Box plot shows the distribution of a variable. A box plot is a graph representing a variable's distribution and consists of elements expressing the dispersion of the data. Also referred to as *box-and-whisker plot*.

Breusch-Pagan test is used to test for heteroskedasticity in regression analysis.

Canberra distance is a distance measure used in cluster analysis. The Canberra distance is a weighted version of the city-block distance, typically used for clustering data scattered widely around an origin.

Case is an object such as a customer, a company, or a country in statistical analysis. Also referred to as *observation*.

Causal research is used to understand the relationships between two or more variables. Causal research explains how variables relate.

Census is a procedure of systematically acquiring and recording information about all the members of a given population.

Centroid linkage is a linkage algorithm in hierarchical clustering methods in which the distance between two clusters is defined as the distance between their geometric centers (centroids).

Chaining effect is a solution pattern typically observed when using a single linkage algorithm in cluster analysis.

Chebychev distance is a distance measure used in cluster analysis that uses the maximum of the absolute difference in the clustering variables' values.

City-block distance is a distance measure used in cluster analysis that uses the sum of the variables' absolute differences. Also referred to *as Manhattan metric*.

Closed-ended questions is a type of question format in which respondents have a certain number of response categories from which to choose.

Cluster analysis is a class of methods that groups a set of objects with the goal of obtaining high similarity within the formed groups and high dissimilarity between groups.

Clustering variables are variables used in cluster analysis.

Clusters are groups of objects with similar characteristics.

Codebook contains essential details of a data file, such as variable names and summary statistics.

Coefficient of determination (R^2) is a measure used in regression analysis to express the dependent variable's amount of variance that the independent variables explain.

Collinearity arises when two variables are highly correlated.

Communality describes the amount of a variable's variance that the extracted factors in a principal component and factor analysis reproduce.

Complete linkage is a linkage algorithm in hierarchical clustering methods in which the distance between two clusters corresponds to the longest distance between any two members in the two clusters.

Components are extracted in the course of a principal component analysis. They are also commonly referred to as factors.

Computer-assisted web interviews (CAWI) See *Web surveys*.

Confidence interval provides the lower and upper limit of values within which a population parameter will fall with a certain probability (e.g., 95%).

Confirmatory factor analysis is a special form of factor analysis used to test whether the measures of a construct are consistent with a researcher's understanding of that construct.

Constant sum scale is a type of scale that requires respondents to allocate a certain total number of points (typically 100) to a number of alternatives.

Constant is a characteristic of an object whose value does not change.

Construct scores are composite scores that calculate a value for each construct of each observation. Construct scores are often computed by taking the mean of all the items associated with the construct.

Construct validity is the degree of correspondence between a measure at the conceptual level and its empirical manifestation. Researchers often use this as an umbrella term for content, criterion, discriminant, face, and nomological validity.

Construct measures a concept that is abstract, complex, and cannot be directly observed by (multiple) items. Also referred to as *latent variable*.

Content validity refers to the extent to which a measure represents all facets of a given construct.

Correlation residuals are the differences between the original item correlations and the reproduced item correlations in a principal component and factor analysis.

Correlation is a measure of how strongly two variables relate to each other. Correlation is a scaled version of the covariance.

Covariance is a measure of how strongly two variables relate to each other.

Covariance-based structural equation modeling (CB-SEM) is an approach to structural equation modeling to test relationships between multiple items and constructs.

Criterion validity measures how well one measure predicts the outcome of another measure when both are measured at the same time.

Cronbach's alpha is a measure of internal consistency reliability. Cronbach's alpha generally varies between 0 and 1 with greater values indicating higher degrees of reliability.

Cross validation entails comparing the results of an analysis with those obtained when using a new dataset.

Crosstabs are tables in a matrix format that show the frequency distribution of nominal or ordinal variables.

Customer relationship management (CRM) refers to a system of databases and software used to track and predict customer behavior.

Data entry errors is a mistake in transcribing data during data entry. Erroneous values that fall outside a variable's standard range can easily be identified by means of descriptive statistics (minimum, maximum, and range).

Degrees of freedom (df) represents the amount of information available to estimate a test statistic. In general terms, an estimate's degrees of freedom are equal to the amount of independent information used (i.e., the number of observations) minus the number of parameters estimated.

Dendrogram visualizes the results of a cluster analysis. Horizontal lines in a dendrogram indicate the distances at which the objects have been merged.

Dependence of observations is the degree to which observations are related.

Dependent variables are the concepts a researcher wants to understand, explain, or predict.

Depth interview is an interview type typically used in exploratory research that allows one-to-one probing to foster interaction between the interviewer and the respondent.

Descriptive research is used to detail certain phenomena, characteristics, or functions. Descriptive research often builds on previous exploratory research.

Discriminant validity ensures that a measure is empirically unique and represents phenomena of interest that other measures in a model do not capture.

Distance matrix expresses the distances between pairs of objects.

Disturbance term see *Residual*.

Divisive clustering is a type of hierarchical clustering method in which all objects are initially merged into a single cluster, which the algorithm then gradually splits up.

Double-barreled questions are survey questions to which respondents can agree with one part but not with the other. Also refers to survey questions that cannot be answered without accepting an assumption.

Duda-Hart index is a statistic used in cluster analysis to determine the number of clusters to extract from the data. The statistic compares the sum of the squares in a pair of clusters to be split both before and after this extraction.

Dummy variables are binary variables that indicate whether a certain trait is present or not.

Durbin-Watson test is a test for autocorrelation used in regression analysis.

Eigenvalue indicates the amount of variance reproduced by a specific component or factor.

Eigenvectors are the results of a principal component analysis and include the factor weights.

Equidistance is indicated when the (psychological) distances between a scale's categories are identical.

Equidistant scale is a scale whose scale categories are equidistant.

Error is the difference between the regression line (which represents the regression prediction) and the actual observation.

Error sum of squares quantifies the difference between the observations and the regression line.

ESOMAR is the world organization for market, consumer, and societal research.

Estimation sample is the sample used to run a statistical analysis.

Eta-squared (η^2) is a statistic used in an ANOVA to describe the ratio of the between-group variation to the total variation, thereby indicating the variance accounted for by the sample data. There are two types of η^2: the model η^2, which is identical to the R^2, and each variable's partial η^2, which describes the percentage of the total variance accounted for by that variable. Eta squared also refers to the percentage of variance explained by a single variable in regression analysis.

Ethics are a system of morals and principles which defines a research organization's obligations, for example, with regard to the findings they release being an accurate portrayal of the survey data.

Ethnography is a type of qualitative research in which the researcher interacts with consumers over a period to observe and question them.

Euclidean distance is a distance measure commonly used in cluster analysis. It is the square root of the sum of the squared differences in the variables' values. Also referred to as *straight-line distance*.

Experimental design describes which treatment variables to administer and how these relate to dependent variables. Prominent experimental designs include the one-shot case study, the before-after design, the before-after design with a control group, and the Solomon four-group design.

Experiments are study designs commonly used in causal research in which a researcher controls for a potential cause and observes corresponding changes in hypothesized effects via treatment variables.

Exploratory factor analysis is a type of factor analysis that derives factors from a set of correlated indicator variables without the researcher having to prespecify a factor structure.

Exploratory research is conducted when the researcher has little or no information about a particular problem or opportunity. It is used to refine research questions, discover new relationships, patterns, themes, and ideas or to inform measurement development.

External secondary data are compiled outside a company for a variety of purposes. Sources of secondary data include, for example, governments, trade associations, market research firms, consulting firms, (literature) databases, and social networks.

External validity is the extent to which the study results can be generalized to real-world settings.

Extreme response styles occur when respondents systematically select the endpoints of a response scale.

Face validity is the extent to which a test is subjectively viewed as covering the concept it purports to measure.

Face-to-face interview See *Personal interview*.

Factor analysis is a statistical procedure that uses the correlation patterns among a set of indicator variables to derive factors that represent most of the original variables' variance. Also referred to as *Principal axis factoring*.

Factor loading is the correlation between a (unit-scaled) factor and a variable.

Factor rotation is a technique used to facilitate the interpretation of solutions in principal component and factor analysis.

Factors are (1) independent variables in an ANOVA and (2) the resulting variables of a principal component and factor analysis that summarize the information from a set of indicator variables.

Factor scores are composite scores that calculate a value for each factor of each observation.

Factor variable is a categorical variable used to define the groups (e.g., three types of promotion campaigns) in an ANOVA.

Factor weights express the relationships between variables and factors.

Familywise error rate is the probability of making one or more false discoveries or type I errors when performing multiple hypotheses tests (see also *α-inflation*).

Field experiments are experiments in which the manipulation of a treatment variable occurs in a natural setting, thereby emphasizing the external validity, but potentially compromising internal validity.

Field service firms are companies that focus on conducting surveys, determining samples and sample sizes, and collecting data. Some of these firms also translate surveys or provide addresses and contact details.

Focus groups is a method of data collection in which four to six participants discuss a defined topic under the leadership of a moderator.

Forced-choice scale is an answer scale that omits a neutral category, thereby forcing the respondents to make a positive or negative assessment.

Formative construct is a type of measurement in which the indicators form the construct.

Free-choice scale is an answer scale that includes a neutral choice category. Respondents are therefore not forced to make a positive or negative assessment.

Frequency table is a table that displays the absolute, relative, and cumulative frequencies of one or more variables.

***F*-test of sample variance** see *Levene's test*.

***F*-test** a test statistic used in an ANOVA and regression analysis to test the overall model's significance.

Full service providers are large market research companies, such as The Nielsen Company, Kantar, or GfK, that offer syndicated and customized services.

Gower's dissimilarity coefficient a dissimilarity coefficient used in cluster analysis that works with a mix of binary and continuous variables.

Heteroskedasticity refers to a situation in regression analysis in which the variance of the residuals is not constant.

Heywood cases negative estimates of variances or correlation estimates greater than one in absolute value.

Hierarchical clustering methods develop a treelike structure of objects in the course of the clustering process, which can be top-down (divisive clustering) or bottom-up (agglomerative clustering).

Histogram is a graph that shows how frequently categories derived from a continuous variable occur.

Hypotheses are claims made about effects or relationships in a population.

Inconsistent answers are a respondent's contradictory answer patterns.

Independent samples *t*-test a test using the *t*-statistic that establishes whether two means collected from independent samples differ significantly.

Independent variables are variables that explain or predict a dependent variable.

In-depth interview is a qualitative conversation with participants on a specific topic.

Index consists of a set of variables that defines the meaning of the resulting composite.

Index construction is the procedure of combining several items to form an index.

Indicators See *items*.

Interaction effect refers to how the effect of one variable on another variable is influenced by a third variable.

Intercept is the expected mean value of the dependent variable in a regression analysis, when the independent variables are zero. Also referred to as a constant.

Internal consistency reliability is a form of reliability used to judge the consistency of results across items in the same test. It determines whether the items measuring a construct are highly correlated. The most prominent measure of internal consistency reliability is Cronbach's alpha.

Internal secondary data are data that companies compile for various reporting and analysis purposes.

Internal validity is the extent to which causal claims can be made in respect of the study results.

Interquartile range is the difference between the third and first quartile.

Inter-rater reliability is the degree of agreement between raters expressed by the amount of consensus in their judgment.

Interviewer fraud is an issue in data collection resulting from interviewers making up data or even falsifying entire surveys.

Item non-response occurs when people do not provide answers to certain questions, for example, because they refuse to answer, or forgot to answer.

Items represent measurable characteristics in conceptual models and statistical analysis. Also referred to as *indicators*.

Kaiser criterion is a statistic used in principal component and factor analysis to determine the number of factor to extract from the data. According to this criterion, researchers should extract all factors with an eigenvalue greater than one. Also referred to as *latent root criterion*.

Kaiser-Meyer-Olkin criterion is an index used to assess the adequacy of the data for a principal component and factor analysis. High values indicate that the data are sufficiently correlated. Also referred to as *measure of sampling adequacy* (*MSA*).

KISS principle the abbreviation of "keep it short and simple!" and implies that any research report should be as concise as possible.

***k*-means** is a group of clustering methods that starts with an initial partitioning of all the objects into a prespecified number of clusters and then gradually reallocates objects in order to minimize the overall within-cluster variation.

***k*-means++** is a variant of the *k*-means method that uses an improved initialization process.

***k*-medians** is a popular variant of *k*-means that aims at minimizing the absolute deviations from the cluster medians.

***k*-medoids** is a variant of *k*-means that uses other cluster centers rather than the mean or median.

Lab experiments are performed in controlled environments (usually in a company or academic lab) to isolate the effects of one or more treatment variables on an outcome.

Label switching a situation in which the labels of clusters change from one analysis to the other.

Laddering is an interviewing technique where the interviewer pushes a seemingly simple response to a question in order to find subconscious motives. It is typically used in the means-end approach.

Latent concepts represent broad ideas or thoughts about certain phenomena that researchers have established and want to measure in their research.

Latent root criterion See *Kaiser criterion*.

Latent variable measures a concept that is abstract, complex, and cannot be directly observed by (multiple) items. Also referred to as *construct*.

Levene's test tests the equality of the variances between two or more groups of data. Also referred to as *F-test of sample variance*.

Likert scale is a type of answering scale in which respondents have to indicate their degree of agreement to a statement. The degree of agreement is usually set by the scale endpoints, which range from strongly disagree to strongly agree.

Limited service providers are market research companies that specialize in one or more services.

Line chart is a type of chart in which measurement points are ordered (typically according to their *x*-axis value) and joined with straight-line segments.

Linkage algorithm defines the distance from a newly formed cluster to a certain object or to other clusters in the solution.

Listwise deletion entails deleting cases with one or more missing value(s) in any of the variables used in an analysis.

Little's MCAR test is used to analyze the patterns of missing data by comparing the observed data with the pattern expected if the data were missing completely at random.

Local optimum is an optimal solution when compared with similar solutions, but not a global optimum.

Log transformation is a type of scale transformation commonly used to handle skewed data.

Mahalanobis distance is a distance measure used in cluster analysis that compensates for the collinearity between the clustering variables.

Mail surveys are paper-based surveys sent to respondents via regular mail.

Main effect is the effect of one independent variable (i.e., factor) on the dependent variable, ignoring the effects of all the other independent variables in a two-way ANOVA.

Manhattan metric See *City-block distance*.

Manipulation checks a type of analysis in experiments to check whether the experimental treatment was effective.

Mann-Whitney U test is the nonparametric equivalent of the independent samples *t*-test used to assess whether two sample means are equal or not.

Marginal mean represents the mean value of one category in respect of each of the other types of categories.

Market segmentation is the segmenting of markets into groups (segments) of objects (e.g., consumers) with similar characteristics (e.g., needs and wants).

Market segments are groups of objects with similar characteristics.

Matching coefficients are similarity measures that express the degree to which the clustering variables' values fall into the same category.

Mean is the most common method of defining a typical value of a list of numbers. It is equal to the sum of a variable's values divided by the number of observations. Also referred to as *arithmetic mean* or simply *average*.

Means-end approach a method used to identify the ends consumers aim to satisfy and the means (consumption) they use to do so.

Measure of sampling adequacy (MSA) See *Kaiser-Meyer-Olkin criterion.*

Measurement scaling refers to (1) the level at which a variable is measured (nominal, ordinal, interval, or ratio scale) and (2) the general act of using a set of variables to measure a construct.

Measures of centrality are statistical indices of a typical or average value of a list of numbers. There are two main types of measures of centrality, the *median* and the *mean*.

Measures of dispersion provide researchers with information about the variability of the data (i.e., how far the values are spread out). There are four main types of measures of dispersion: the *range*, *interquartile range*, *variance*, and *standard deviation*.

Median is a value that separates the lowest 50% of values from the highest 50% of values.

Middle response styles a systematic way of responding to survey items describing respondents' tendency to choose the midpoints of a response scale.

Minto principle a guideline for presentations that starts with the conclusion, raising questions in the audience's mind about the way this conclusion was reached. The presenter subsequently explains the steps involved in the analysis.

Missing at random (MAR) is a missing values pattern in which the probability that data points are missing varies from respondent to respondent.

Missing completely at random (MCAR) is a missing values pattern in which the probability that data points are missing is unrelated to any other measured variable and to the variable with the missing values.

Missing data occur when entire observations are missing (survey non-response) or respondents have not answered all the items (item non-response).

Mixed mode is the act of combining different ways of administering surveys.

Moderation analysis involves assessing whether the effect of an independent variable on a dependent variable depends on the values of a third variable, referred to as a moderator variable.

(Multi)collinearity is a data issue that arises in regression analysis when two or more independent variables are highly correlated.

Multi-item construct is a measurement of an abstract concept that uses several items.

Multinomial logistic regression is a type of regression analysis used when the dependent variable is nominal and takes more than two values.

Multiple imputation is a simulation-based statistical technique that replaces missing observations with a set of possible values (as opposed to a single value) representing the uncertainty about the missing data's true value.

Multiple regression is a type of regression analysis that includes multiple independent variables.

Mystery shopping is a type of observational study in which a trained researcher visits a store or restaurant and consumes their products/services.

Nested models are simpler versions of a complex model.

Net Promoter Score (NPS) is a measure of customer loyalty that uses the single question: "How likely are you to recommend our company/product/service to a friend or colleague?"

Nomological validity is the degree to which a construct behaves as it should in a system of related constructs.

Nonhierarchical clustering methods see *Partitioning methods*.

Nonparametric tests are statistical tests for hypothesis testing that do not assume a specific distribution of the data (typically a normal distribution).

Non-probability sampling is a sampling technique that does not give every individual in the population an equal chance of being included in the sample. The resulting sample is not representative of the population.

Nonrandom missing is a missing values pattern in which the probability that data points are missing depends on the variable and on other unobserved factors.

Null and alternative hypothesis the null hypothesis (indicated as H_0) is a statement expecting no difference or no effect. The alternative hypothesis (indicated as H_1) is the hypothesis against which the null hypothesis is tested.

Oblimin rotation is a popular oblique rotation method used in principal component and factor analysis and principal component analysis.

Oblique rotation is a technique used to facilitate the interpretation of the factor solution in which the independence of a factor to all other factors is not maintained.

Observation is an object, such as a customer, a company, or a country, in statistical analysis. Also referred to as *case*.

Observational studies are procedures for gathering data in which the researcher observes people's behavior in a certain context. Observational studies are normally used to understand what people are doing rather than why they are doing it.

Omega-squared (ω^2) is a statistic used in an ANOVA to describe the ratio of the between-group variation to the total variation, thereby indicating the variance accounted for by the data. It is commonly used for sample sizes of 50 or less and corresponds to the adjusted R^2 of regression analysis. Omega squared is also used to indicate effect sizes of individual variables in regression analysis.

One-sample *t*-test is a parametric test used to compare one mean with a given value.

One-tailed tests are a class of statistical tests frequently used when the hypothesis is expressed directionally (i.e., $<$ or $>$). The region of rejection is on one side of the sampling distribution.

One-way ANOVA is a type of ANOVA that involves a single metric dependent variable and one factor variable with three (or more) levels.

Open-ended questions are a type of question format that provides little or no structure for respondents' answers. Generally, the researcher asks a question and the respondent writes down his or her answer in a box. Also referred to as *verbatim items*.

Operationalization is the process of defining a set of variables to measure a construct. The process defines latent concepts and allows them to be measured empirically.

Ordinary least squares (OLS) is the estimation approach commonly used in regression analysis and involves minimizing the squared deviations from the observations to the regression line (i.e., the residuals).

Orthogonal rotation is a technique used to facilitate the interpretation of a factor solution in which a factor's independence is maintained from all other factors. The correlation between the factors is determined as zero.

Outliers are observations that differ substantially from other observations in respect of one or more characteristics.

Paired samples *t*-test is a statistical procedure used to determine whether there is a significant mean difference between observations measured at two points in time.

Parallel analysis is a statistic used in principal component and factor analysis to determine the number of factors to extract from the data. According to this criterion, researchers should extract all factors whose eigenvalues are larger than those derived from randomly generated data with the same sample size and number of variables.

Parametric tests are statistical tests that assume a specific data distribution (typically normal).

Partial least squares structural equation modeling (PLS-SEM) is a variance-based method to estimate structural equation models. The goal is to maximize the explained variance of the dependent latent variables.

Partial sums of squares is a statistic in an ANOVA indicating the additional portion of variance explained when another variable is added to the analysis.

Partitioning method is a group of clustering procedures that does not establish a treelike structure of objects and clusters, but exchanges objects between clusters to optimize a certain goal criterion. The most popular type of partitioning method is *k*-means.

Path diagram is a visual representation of expected relationships tested in a structural equation modeling analysis.

Personal interview is an interview technique that involves face-to-face contact between the interviewer and the respondent. Also referred to as *face-to-face interviews*.

Pie chart displays the relative frequencies of a variable's values.

Population is a group of objects (e.g., consumers, companies, or products) that a researcher wants to assess.

Post hoc tests are a group of tests used for paired comparisons in an ANOVA. Post hoc tests maintain the familywise error rate (i.e., they prevent excessive type I error).

Power of a test represents the probability of rejecting a null hypothesis when it is in fact false. In other words, the power of a statistical test is the probability of rendering an effect significant when it is indeed significant (defined by $1 - \beta$, where β is the probability of a type II error).

Practical significance refers to whether differences or effects are large enough to influence decision-making processes.

Predictive validity measures how well one measure predicts the outcome of another measure when both are measured at a later point in time.

Primary data are data gathered for a specific research project.

Principal axis factoring See *Factor analysis*.

Principal component analysis is a statistical procedure that uses correlation patterns among a set of indicator variables to derive factors that represent most of the original variables' variance. Different from factor analysis, the procedure uses all the variance in the variables.

Principal components are linear composites of original variables that reproduce the original variables' variance as well as possible.

Principal factor analysis See *Factor analysis*.

Probability sampling is a sampling technique that gives every individual in the population an equal chance, different from zero, of being included in the sample.

Profiling is a step in market segmentation that identifies observable variables (e.g., demographics) that characterize the segments.

Projective technique is a special type of testing procedure, usually used as part of in-depth interviews. This technique provides the participants with a stimulus (e.g., pictures, words) and then gauges their responses (e.g., through sentence completion).

Promax rotation is a popular oblique rotation method used in principal component and factor analysis and principal component analysis.

***p*-value** is the probability of erroneously rejecting a true null hypothesis in a given statistical test.

Pyramid structure for presentations See *Minto principle*.

Qualitative data are audio, pictorial, or textual information that researchers use to answer research questions.

Qualitative research is primarily used to gain an understanding of *why* certain things happen. It can be used in an exploratory context by defining problems in more detail or by developing hypotheses to be tested in subsequent research.

Quantitative data are data to which numbers are assigned to represent specific characteristics.

R^2 See Coefficient of determination.

Ramsey's RESET test is a test for linearity used in regression analysis.

Range standardization is a type of scale transformation in which the values of a scale are standardized to a specific range that the researcher has set.

Range is the difference between the highest and the lowest value in a variable measured, at least, on an ordinal scale.

Rank order scale is an ordinal scale that asks respondents to rank a set of objects or characteristics in terms of, for example, importance, preference, or similarity.

Reflective constructs is a type of measurement in which the indicators are considered manifestations of the underlying construct.

Regression method is a procedure to generate factor scores in principal component analysis. The resulting factor scores have a zero mean and unit standard deviation.

Regression sum of squares quantifies the difference between the regression line and the line indicating the average. It represents the variation in the data that the regression analysis explains.

Reliability is the degree to which a measure is free from random error.

Reliability analysis is an important element of a confirmatory factor analysis and essential when working with measurement scales. See *Reliability*.

Research design describes the general approach to answer a research question related to a marketing opportunity or problem. There are three broad types of research design: exploratory research, descriptive research, and causal research.

Residual is the unexplained variance in a regression model. Also referred to as *disturbance term*.

Reverse-scaled items are items whose statement (if a Likert scale is used) or word pair (if a semantic differential scale is used) is reversed when compared to the other items in the set.

Robust regression is a variant of regression analysis used when heteroskedasticity is present.

Russell and Rao coefficient is a similarity coefficient used in cluster analysis.

Sample size is the number of observations drawn from a population.

Sampling error occurs when the sample and population structure differ on relevant characteristics.

Sampling is the process through which objects are selected from a population.

Scale development is the process of defining a set of variables to measure a construct and which follows an iterative process with several steps and feedback loops. Also referred to as *operationalization*, or, in the case of an index, *index construction*.

Scale transformation is the act of changing a variable's values to ensure comparability with other variables or to make the data suitable for analysis.

Scanner data are collected at the checkout of a supermarket where details about each product sold are entered into a database.

Scatter plot is a graph that represents the relationship between two variables, thus portraying the joint values of each observation in a two-dimensional graph.

Scree plot is a graph used in principal component and factor analysis that plots the number of factors against the eigenvalues, resulting in a distinct break (elbow) that indicates the number of factors to extract. Following the same principle, the

scree plot is also used in hierarchical cluster analysis to plot the number of clusters against the distances at which objects were merged.

Secondary data are data that have already been gathered, often for a different research purpose and some time ago. Secondary data comprise internal secondary data, external secondary data, or a mix of both.

Segment specialists are companies that concentrate on specific market segments, such as a particular industry or type of customer.

Self-contained figure is a graph in a market research report that should be numbered sequentially and have a meaningful title so that it can be understood without reading the text.

Self-contained table is a table in a market research report that should be numbered sequentially and have a meaningful title so that it can be understood without reading the text.

Semantic differential scales is a type of answering scale that comprises opposing pairs of words, normally adjectives (e.g., young/old, masculine/feminine) constituting the endpoints of the scale. Respondents then indicate how well one of the word in each pair describes how he or she feels about the object to be rated (e.g., a company or brand).

Sentence completion is a type of projective technique that provides respondents with beginnings of sentences that they have to complete in ways that are meaningful to them.

Sequential sums of squares is a statistic in an ANOVA that indicates the additional portion of variance explained when a set of variables is added to the analysis.

Shapiro-Wilk test is a test for normality (i.e., whether the data are normally distributed).

Significance level is the probability that an effect is incorrectly assumed when there is in fact none. The researcher sets the significance level prior to the analysis.

Simple matching coefficient is a similarity coefficient used in cluster analysis.

Simple regression is the simplest type of regression analysis with one dependent and one independent variable.

Single-item constructs is a measurement of a concept that uses only one item.

Single linkage is a linkage algorithm in hierarchical clustering methods in which the distance between two clusters corresponds to the shortest distance between any two members in the two clusters.

Skewed data occur if a variable is asymmetrically distributed. A positive skew (also called right skewed) occurs when many observations are concentrated on the left side of the distribution, producing a long right tail (the opposite is called negative skew or left skewed).

Social desirability bias occurs when respondents provide socially desirable answers (e.g., by reporting higher or lower incomes than are actually true) or

take a position that they believe society favors (e.g., not smoking or drinking).

Social media analytics are methods for analyzing social networking data and comprise text mining, social network analysis, and trend analysis.

Social networking data reflect how people would like others to perceive them and, thus, indicate consumers' intentions. Product or company-related social networking data are of specific interest to market researchers.

Specialized service firms are market research companies that focus on particular products, markets, or market research techniques.

Split-half reliability is a type of reliability assessment in which scale items are divided into halves, and the scores of the halves are correlated.

Split-sample validation involves splitting the dataset into two samples, running the analysis on both samples, and comparing the results.

Stability of the measurement See *Test-retest reliability*.

Standard deviation describes the sample distribution values' variability from the mean. It is the square root of the variance and, therefore, a variant.

Standard error is the sampling distribution of a statistic's standard deviation, mostly from the mean.

Standardized effects express the relative effects of differently measured independent variables in a regression analysis by expressing them in terms of standard deviation changes from the mean.

Standardizing variables have been rescaled (typically to a zero mean and unit standard deviation) to facilitate comparisons between differently scaled variables.

Stata computer package specializing in quantitative data analysis.

Statistical inference is the process of drawing conclusions about populations from data.

Statistical significance occurs when an effect is so large that it is unlikely to have occurred by chance. Statistical significance depends on several factors, including the size of the effect, the variation in the sample data, and the number of observations.

Straight-line distance See *Euclidean distance*.

Straight-lining occurs when a respondent marks the same response in almost all the items.

Structural equation modeling is a multivariate data analysis technique used to measure relationships between constructs, as well as between constructs and their associated indicators.

Survey non-response occurs when entire responses are missing. Survey non-response rates are usually 75%–95%.

Surveys are often used for gathering primary data. Designing surveys involves a six-step process: (1) Determine the survey goal, (2) determine the type of questionnaire required and the administration method, (3) decide on the questions and (4) the scale, (5) design the questionnaire, and (6) pretest and administer the questionnaire.

Suspicious response patterns are issues in response styles in respect of straight-lining and inconsistent answers that a researcher needs to address in the analysis.

Syndicated data are data sold to multiple clients, allowing them to compare key measures with those of the rest of the market.

Telephone interviews allow researchers to collect data quickly and facilitate open-ended responses, although not as well as personal interviews.

Test markets are a type of field experiment that evaluates a new product or promotional campaign under real market conditions.

Test statistic is calculated from the sample data to assess the strength of the evidence in support of the null hypothesis.

Test-retest reliability is a type of reliability assessment in which the researcher obtains repeated measurement of the same respondent or group of respondents, using the same instrument and under similar conditions. Also referred to as *stability of the measurement*.

Ties are identical values in a distance matrix used in cluster analysis.

Total sum of squares quantifies the difference between the observations and the line indicating the average.

Transforming data is an optional step in workflow of data, involving variable respecification and scale transformation.

Treatments are elements in an experiment that are used to manipulate the participants by subjecting them to different situations. A simple form of treatment could be an advertisement with and without humor.

***t*-test** is the most popular type of parametric test for comparing a mean with a given standard and for comparing the means of independent samples (independent samples *t*-test) or the means of paired samples (paired samples *t*-test).

Tukey's honestly significant difference test is a popular post hoc test used in an ANOVA that controls for type I errors, but is limited in terms of statistical power. Often simply referred to as *Tukey's method*.

Tukey's method See *Tukey's honestly significant difference test*.

Two-sample *t*-test is the most popular type of parametric test for comparing the means of independent or paired samples.

Two-tailed tests are a class of statistical tests frequently used when the hypothesis is not expressed directionally (i.e., $\neq$). The region of rejection is on two sides of the sampling distribution.

Two-way ANOVA is a type of ANOVA that involves a single metric dependent variable and two factor variables with three (or more) levels.

Type I error occurs when erroneously rejecting a true null hypothesis. Also referred to as α error.

Type II error occurs when erroneously accepting a false null hypothesis. Also referred to as β error.

Unbalanced scale describes a scale with an unequal number of positive and negative scale categories.

Uniqueness is a statistic used in principal component analysis and factor analysis that indicates the proportion of a variable's variance that the factors do not capture. The uniqueness equals 1 – communality.

Unit of analysis is the level at which a variable is measured. Typical measurement levels include that of the respondents, customers, stores, companies, or countries.

Univariate statistics are statistics that describe the centrality and dispersion of a single variable.

Unstandardized effects express the absolute effects that one-unit increases in the independent variables have on the dependent variable in a regression analysis.

Validation sample is a random subsample of the original dataset used for validation testing.

Validity is the degree to which a researcher measures what (s)he wants to measure. It is the degree to which a measure is free from systematic error.

Variable represents a measurable characteristic whose value can change.

Variable names should be clear and short so that they can be read in the dialog boxes (e.g., *loyalty1*, *loyalty2*, etc.).

Variable respecification involves transforming data to create new variables or to modify existing ones.

Variance a measure of dispersion computed by the sum of the squared differences of each value and a variable's mean, divided by the sample size minus 1.

Variance inflation factor (VIF) quantifies the degree of collinearity between the independent variables in a regression analysis.

Variance ratio criterion is a statistic used in cluster analysis to determine the number of clusters. The criterion compares the within- and between-cluster variation of different numbers of clusters.

Varimax rotation is the most popular orthogonal rotation method used in principal component analysis and factor analysis.

Verbatim items See *Open-ended questions*.

Visual aids include overhead transparencies, flip charts, or slides (e.g., PowerPoint or Prezi) that help emphasize important points and facilitate the communication of difficult ideas in a presentation of market research results.

Visual analogue scale is a type of answering scale in which respondents use levers that allow scaling on a continuum. This scale does not provide response categories.

Ward's linkage is a linkage algorithm in hierarchical clustering methods that combines those objects whose merger increases the overall within-cluster variance by the smallest possible degree.

Web surveys are less expensive to administer and can be fast in terms of data collection, because they can be set up very quickly. Also referred to as *computer-assisted web interviews* (*CAWI*).

Weighted average linkage is a variant of the average linkage algorithm used in cluster analysis that weights the distances according to the number of objects in the cluster.

Welch correction is a statistical test used in an ANOVA to assess the significance of the overall model when the group variances differ significantly and the groups differ in size.

White's test is a statistical test that detects the presence of heteroskedasticity in regression analysis.

Wilcoxon matched-pairs signed-rank test is the nonparametric equivalent of the paired samples t-test.

Wilcoxon signed-rank test is the nonparametric equivalent of the independent samples t-test.

Wilcoxon test is the nonparametric equivalent of the one sample t-test.

Workflow is a strategy to keep track of the entering, cleaning, describing, and transforming of data.

z-standardization is a type of scale transformation in which the values of a scale are standardized to a zero mean and unit standard deviation.

z-test is any statistical test for which the distribution of the test statistic under the null hypothesis can be approximated by a normal distribution.

Index

A
Acquiescence, 72, 101
Adaptive questioning, 68
Adjusted R^2, 189, 234
Agglomerative clustering, 322, 327
Aggregating data, 33
Aggregation, 122
Agresti, A., 118
Akaike information criterion (AIC), 235, 280
α error, 158
α-inflation, 180
Alternative hypothesis, 156
American Customer Satisfaction Index, 60
Analysis of variance (ANOVA), 166
Anti-image, 272
Arithmetic mean, 113
Armstrong, J.S., 385
Autocorrelation, 230
Average, 113
Average linkage, 323, 324

B
Back-translation, 72
Balanced scale, 77
Bar chart, 111
Bartlett method, 284
Bayes information criterion (BIC), 235, 280
Before-after design, 88
Before-after experiment with a control
      group, 89
Before measurement effect, 89
β error, 158
Between-group mean squares, 186
Between-group variation, 182–183
Big data, 57
Binary logistic regression, 220
Bivariate graphs, 115
Bivariate statistics, 117

Bonferroni correction, 187
Box-and-whisker plot, 112
Box plot, 112
Breusch-Pagan, 228
Bubble plot, 116

C
Calinski/Harabasz pseudo F, 341, 356
Canberra distance, 335
Case, 28
Causality, 18
Causal research, 18, 86
Census, 43
Centroid linkage, 323, 324
Chaining effect, 324
Chebychev distance, 334
χ^2-test, 108, 116
CIA World Fact Book, 54
City-block distance, 334
Closed-ended questions, 71
Cluster analysis, 314
Clustering variables, 314
Clusters, 314
Cluster sampling, 45
Codebook, 123
Coefficient of determination, 232
Collinearity, 221, 276
Communality, 276, 277
Company records, 53
Complete linkage, 323, 324
Composite measure, 29, 121
Computer-assisted personal
      interviews (CAPI), 66
Computer-assisted self-interviews (CASI), 66
Computer-assisted telephone interviews
      (CATI), 67
Computer-assisted web interviews (CAWI), 67
Confidence interval, 189