Brought to you by:

ASTRA

MARKETING (book to read for non attending students)

1° BIG

Written by:
Lia Bassanini

2023-2024 Edition

Find out more at:

astrabocconi.it

Marketing - general non attending

These are the notes of the requested book for non attending students. Our recommendation is to read the book and use this handout for a review.

Chapter 1: Introduction to marketing research

Marketing research: the systematic and objective identification, collection, analysis, dissemination, and use of information for the purpose of assisting management in decision making related to the identification and solution of problems and opportunities in marketing.
In order to identify and solve marketing problems: identification of information needed → collection of data → analysis of data → dissemination of information → use of information

Problem identification research: research undertaken to help identify problems that are not necessarily apparent on the surface and yet exist or are likely to arise in the future.
Problem solving research: research undertaken to help solve specific marketing problems.

Marketing research process: a set of 6 steps that define the tasks to be accomplished in conducting a marketing research study: defining the problem → developing an approach to the problem → formulating a research design → doing field work (collecting data) → preparing and analyzing data → preparing and presenting the report (see figure page 32)

The task of marketing research is to assess the information needs and provide management with relevant, accurate, reliable, valid, current and actionable information. In the absence of sound information, an incorrect management decision can be made (see figure page 34).
Marketing managers make strategic and tactical decisions in the process of identifying and satisfying customers' needs. They make decisions about potential opportunities, target market selections, market segmentation, planning and implementing marketing programs, marketing performance, and control. These decisions are complicated by interactions among the controllable marketing variables of product, pricing, promotion, and distribution (under the control of the marketing manager). Additional complications can come from uncontrollable environmental factors such as general economic condition, technology, public policies and laws, political environment, competition, and social and cultural changes (not under the control of the marketing manager). Another factor is the complexity of various customer groups: consumers, employees, channel members, shareholders, and suppliers.

The decision to conduct marketing research should be guided by a number of considerations, including the costs versus the benefits, the resources available to conduct the research, the resources available to implement the research findings, and management's attitude towards research. Marketing research should be undertaken when the expected value of information it generates exceeds the cost of doing the research and if there are not seriously inhibiting constraints.

Marketing research industry: all internal and external suppliers who provide marketing research services.
Internal suppliers: marketing research departments located within firms.
External suppliers: outside marketing research companies hired to supply marketing research services;
Full-service suppliers: companies that offer the full range of marketing research activities and perform all the six steps of the marketing research process.
Customized services: companies that tailor the research procedures to best meet the needs of each client.

Syndicated services: companies that collect and sell common pools of data designed to serve information needs shared by a number of clients.
Internet services: companies that specialize in conducting marketing research on the Internet, including social media.
Limited-service suppliers: companies that specialize in one or a few steps of the marketing research process.
Field services: companies whose primary service offering is their expertise in collecting the data for research projects.
Qualitative services: services related to facilities, recruitment, and other services for focus groups and other forms of qualitative research such as one-on-one depth interviews.
Technical and analytical services: companies that provide guidance in the development of the research design and computer analysis of quantitative and qualitative data
Data analysis services: firms whose primary service is to conduct statistical analysis of data.
Branded marketing research products: specialized data collection and analysis procedures developed to address specific types of marketing problems.

Marketing information system (MIS): a formalized set of procedures for generating, analyzing, storing, and distributing pertinent information to marketing decision makers on an ongoing basis.
Decision support system (DSS): an information system that enables decision makers to interact directly with both databases and analysis models.

We use the term international marketing research broadly to denote all research conducted in markets other than the domestic market of research commissioning organization. It is more complex to carry out because you have to take into account more environmental factors.

Marketing research has often been described as having 4 shareholders: the marketing researcher, the client, the respondent, and the public. Ethical issues arise when the interests of these shareholders are in conflict and when one or more shareholders is lacking its responsibilities.

Chapter 2: Defining the marketing research problem and developing an approach
Problem definition: a statement of the management decision problem and the marketing research problem. The marketing research problem is broken down into a broad statement and specific components.
Problem-definition process: the process of defining the management decision problem and the marketing research problem (see figure page 56).

The tasks involved in problem definition include: discussion with the decision makers, interviews with industry experts, analysis of secondary data, and qualitative research.
Problem audit: a comprehensive examination of a marketing problem to understand its origin and nature. It involves discussions with the decision maker (DM) on the history of the problem, the alternative course of actions available to the DM, and the information that is needed to answer the DM's questions.
Secondary data: data collected for some purpose other than the problem at hand.
Primary data: data originated by the researcher specifically to address the research problem.
Analyzing available secondary data is an essential step in the problem-definition process and should always precede primary data collection.

Qualitative research: an unstructured, exploratory research methodology based on small samples intended to provide insight and understanding of the problem setting

Environmental context of the problem: consists of the factors that have an impact on the definition of marketing research problem, including past information and forecasts, resources and constraints of the firm, objective of the decision maker, buyer behavior, legal environment, economic environment, and marketing and technological skills of the firm (see figure page 62).
Objectives: the goals of the organization and of the DM that must be considered in order to conduct successful marketing research. The formulation of the management decision problem must be based on a clear understanding of two types of objectives: the organizational objectives and the personal objectives of the DM.
Buyer behavior: a body of knowledge that tries to understand and predict consumers' reactions based on specific characteristics of individuals. The buyer behavior factors that should be considered include: the number and geographical location of the buyers and nonbuyers, demographic and psychographic (lifestyle) characteristics, product consumption habits and the consumption of related product categories, media consumption behavior and response to promotions, price sensitivity, retail outlets patronized, and buyer preferences.
Legal environment: regulatory policies and norms which organizations must operate
Economic environment: composed of purchasing power, income, prices, savings, credits availability, and general economic conditions.

Management decision problem: the problem confronting the DM. It asks what the DM needs to do.
Marketing research problem: problem that entails determining what information is needed and ow it can be obtained in the most feasible way.

The general rule to follow in defining the research problem is that the definition should: allow the researcher to obtain all the information needed to address the management decision problem, and guide the researcher in proceeding the project. Researchers make two common errors in problem definition. The first arises when the research problem is defined too broadly. A broad definition does not provide clear guidelines for the subsequent steps involved in the project. Some examples of overly broad marketing research problem definition are: developing a marketing strategy for the brand, improving the competitive position of the firm, or improving the company's image. The second type of error is to define the marketing research project too narrowly, which can preclude considerations of some courses of action, particularly those that are innovative and not obvious, and it may also prevent the researcher from addressing important components of the management decision problem.
Broad statement of the problem: the initial statement of the marketing research problem that provides an appropriate perspective on the problem.
Specific components of the problem (the second part of the marketing research problem definition): they focus on the key aspects of the problem and provide clear guidelines on how to proceed.

An approach to a marketing research problem should include the following components: analytical framework and models, research questions and hypotheses, and a specification of the information needed.
Objective evidence: unbiased evidence that is supported by empirical findings.
Theory: conceptual scheme based on foundational statements, or axioms, that are assumed to be true.
Analytical framework: statement of the theory as applied to the marketing research problem
Analytical model: explicit specification of a set of variables and their interrelationships designed to represent some real system or process in whole or in part.

Graphical model: an analytical model that provides a visual picture of the relationships between variables.
Research questions (RQs): refined statements of the specific components of the problem.
Hypothesis (H): unproven statement or proposition about a factor or a phenomenon that is of interest to the researcher.
By focusing on each component of the problem, the RQs and H, the research can determine what information should be obtained (see figure page 72).

Chapter 3: Research design, secondary and syndicated data
Research design: framework or blueprint for conducting the marketing research project. It specifies the details of the procedure necessary for obtaining the information needed to structure and/or solve marketing research problems. This process begins by defining the marketing research problem, then developing an approach to the problem, and finally formulating the research design. Typically, a research design involves all or most of the following components or tasks: analysis of secondary and syndicated data, qualitative research, survey and observation research, experimental research, measurement and scaling procedures, design and pretest of a questionnaire or an observation form for data collection, sampling process and sample size determination, and development of a preliminary plan of data analysis.

Research designs are of 2 broad types: exploratory and conclusive.
Exploratory research: has as its primary objective the provision of insights into and comprehension of the problem situation confronting the researcher.
Conclusive research: research designed to assist the decision maker in determining, evaluating and selecting the best course of action to take in a given situation.
Descriptive research: a type of conclusive research that has as its major objective the description of something, usually market characteristics or functions.
Cross-sectional design: type of research design involving the collection of information from any given sample of population elements only once.
Longitudinal design: type of research design involving a fixed sample of population element, that is measured repeatedly. The sample remains the same over time. Thus, it provides a series of pictures that, when viewed together, portray a vivid illustration of the situation and the changes taking place over time.
Causal research: type of conclusive research where the major objective is to obtain evidence regarding cause-and-effect relationships (see figure page 88).

The criteria used for evaluating secondary data consist of specifications (to identify possible sources of bias), error, currency, objective, nature and dependability (by examining the expertise, the credibility, reputation and trustworthiness of the source).

Internal data: data available within the organization for which research is being conducted
External data: data that originate external to the organization for which the research is being conducted.
See figure page 93

Syndicated sources: companies that collect and sell common pools of data designed to serve information needs shared by a number of clients, including competing firms in the same industry (see figure page 97)

Periodic surveys: surveys that collect data on the same set of variables at regular intervals, each time sampling from a new group of respondents (see figure page 98).
Panel surveys: surveys that measure the same group of respondents over time, but not necessarily on the same variables.
A major advantage of surveys is that they are flexible. However, errors might occur because respondents remember incorrectly or feel pressured to give the "right" answer.
Purchase panels: a data-gathering technique in which respondents record their purchases in a diary or on the Internet
Media panels: a data-gathering technique that involves samples of respondents whose television viewing behavior is automatically recorded by electronic devices, supplementing the purchase information recorded in a diary.
Scanner data: data obtained by passing merchandise over a laser scanner that reads the UPC code from the packages.
Volume-tracking data: scanner data that provide information on purchases by brand, size, price, and flavor or formulation
Scanner panels: scanner data collected from panel members who are issued an ID card that enables their purchases to be linked to their identities.

Audit: a data-collection process derived from physical records or inventory analysis. Data are collected personally by the researcher or by representatives of the researcher, and the data are based on counts, usually of physical objects (see figure page 105).
Industry services: secondary data derived from industrial firms and organizational sources and intended for industrial or institutional use.
Single-source data: an effort to combine data from different sources by gathering integrated information on household and marketing variables applicable to the same set of respondents.

Chapter 4: Qualitative research
Quantitative research: a research methodology that seeks to quantify the data and typically applies some form of statistical analysis (see figure page 121).

Direct approach: one type of qualitative research in which the purposes of the project are disclosed to the respondent or are obvious given the nature of the interview.
Indirect approach: a type of qualitative research in which the purposes of the project are disguised from the respondents

Focus group: an interview conducted by a trained moderator among a small group of respondents in an unstructured and natural manner. In order to plan and conduct a focus group, you have to follow 6 steps: specify the objectives of the qualitative research, state the objectives of the focus group, write a screening questionnaire, develop a moderator's outline, conduct the focus group interview, and prepare the focus group report.

Depth interviews: an unstructured, direct, personal interview with a single respondent is questioned by a highly skilled interviewer to uncover underlying motivations, beliefs, attitudes, and feelings on a topic.

Ethnographic research: the study of human behavior in its natural context; it involves observation of behavior and setting along with depth interviews.

Netnography: netnography uses ethnographic techniques but uses data that is naturally found on the Internet and that is generated by online communities.

Chapter 5: Survey and Observation

Survey method: a structured questionnaire given to a population sample and designed to elicit specific information from respondents.
Structured data collection: use a formal questionnaire that presents questions in a prearranged order.
Fixed-alternative questions: questions that require respondents to choose from a set of predetermined answers (see figure page 149).

Survey questionnaires can be administered in four major modes: telephone interviews (traditional or computer assisted), personal interviews (in-home or mall-intercept or computer assisted), mail interviews (mail panels), and electronic (e-mail or internet).

To improve survey response rates: prior notifications, incentives (prepaid or promised), follow-up, and other facilitators of response.

Observation: the recording of people's behavioral patterns and of objects and events in a systematic manner to obtain information about the phenomenon of interest.
Personal observation: an observational research method in which human observers record the phenomenon being observed as it occurs.
Humanistic inquiry: a special form of personal observation in which the researcher is immersed in the system under study.
Mechanical observation: an observational research method in which mechanical devices, rather than human observers, record the phenomenon being observed.

Chapter 6: Experimentation and causal research

Before assuming causality, 3 conditions must be satisfied: concomitant variation, time order of occurrence of variables, and elimination of other possible causal factors. These conditions are necessary but not sufficient to demonstrate causality.
Concomitant variation: a condition for inferring causality which requires that a cause, X, and an effect, Y, occur together or vary together as predicted by the hypothesis under consideration.
The time order occurrence condition states that the causing event must occur either before or simultaneously with the effect.
The absence of other possible causal factors means that the factor or variable being investigated should be the only possible causal explanation.
If all the evidence is strong and consistent, it may be reasonable to conclude that there is a causal relationship.

Field experiment: an experiment conducted in actual market conditions, that is a natural environment.
Laboratory experiment: an experiment conducted in an artificial environment in which the researcher constructs the desired conditions.

Independent variables: variables that are manipulated by the researcher and whose effects are measured and compared.

Dependent variables: variables that measure the effect of the independent variables on the test units.
Random assignment: the process of randomly assigning test units and treatments to the experimental and control groups. It is one of the most common techniques used to control for the effect of extraneous variables on the dependent variable (see page 185).

Internal validity: a measure of the accuracy of an experiment. It measures if the manipulation of the independent variables, or treatments, actually caused an effect on the dependent variable.
External validity: a determination of whether the cause-and-effect relationships found in the experiment can be generalized.

Pre-experimental designs: designs that do not control for extraneous factors by randomization
True experimental designs: experimental designs distinguished by the fact that the researcher can randomly assign test units to experimental groups and also randomly assign treatments to experimental groups.
Statistical design: an experimental design that allows for the statistical control and analysis of extraneous variables.

One-shot case study: a pre-experimental design in which a single group of test units is exposed to a treatment X, and then a single measurement on the dependent variable is taken. Test units are not assigned at random.
One-group pretest-posttest design: a pre-experimental design in which a group of test units is measured twice, before and after exposure to the treatment. Test units are not assigned at random.
Static group: a pre-experimental design in which there are 2 groups: the experimental group (EG), which is exposed to treatment, and the control group (CG). Measurements on both groups are made only after the treatment, and test units are not assigned at random.

Factorial design: a statistical experimental design used to measure the effects of 2 or more independent variables at various levels and to allow interactions between variables.
Interaction: an interaction occurs when the simultaneous effect of 2 or more variables is different from the sum of their separate effects.

Designs that offer the greatest level of internal validity typically are conducted in highly artificial environments that can threaten the generalizability or external validity of the experimental results. Designs that offer tight internal validity (i.e. laboratory experiments) could be used during the early stages of the research effort and, during the later stages of the study, more natural settings (i.e. field experiments) could be used to enable the generalization of results. True experimental and factorial designs are favored in the later stages.

Although experimentation is becoming increasingly important in marketing research, it does have some limitations: time, cost, and administration of an experiment.

Test marketing: an application of a controlled experiment done in limited but carefully selected test markets. It involves a replication of the planned national marketing program for a product in the test markets.

Chapter 7: Measurement and scaling

Measurement: the assignment of numbers or other symbols to characteristics of objects according to a certain set of rules.
Scaling: the generation of a continuum upon which measured objects are located (see figure page 207)

Nominal scale: a scale whose numbers serve only as labels or tags for identifying and classifying objects, with a strict one-to-one correspondence between the numbers and the objects when used for identification.
Ordinal scale: a ranking scale in which numbers are assigned to objects to indicate the relative extent to which some characteristic is possessed. Thus, it is possible to determine whether an object has more or less of a characteristic than another object.
Interval scale: a scale in which the numbers are used to rate objects. The numerical equal distances on the scale represent equal distances in the characteristic being measured.
Ratio scale: the highest measurement scale conveying the most information. It allows the researcher to identify or classify objects, rank-order the objects, and compare intervals or differences. It is also meaningful to compute ratios of scale values.

Comparative scale: one of 2 types of scaling techniques in which there is direct comparison of stimulus objects with one another (see figure page 213).
Noncomparative scales: one of 2 types of scaling techniques in which each object is scaled independently of the other object in the stimulus set (see figure page 217).

Paired-comparison scaling: a comparative scaling technique in which a respondent is presented with 2 objects at a time and asked to select one object in the pair according to some criterion. The data obtained are ordinal in nature.
Rank-order scaling: a comparative scaling technique in which respondents are presented with several objects simultaneously and asked to order or rank them according to some criterion.
Constant sum scaling: a comparative scaling technique in which respondents are required to allocate a constant sum of units such as points, dollars, chits, stickers, or chips among a set of stimulus objects with respect to some criterion.

Continuous rating scale: a non comparative measurement scale in which respondents rate the objects by placing a mark at the appropriate position on a line that runs from one extreme of the criterion variable to the other. The form may vary considerably.
Itemized rating scale: a noncomparative measurement scale having numbers and/or brief descriptions associated with each category. The categories are ordered in terms of scale position.
Likert scale: a noncomparative measurement scale that typically has 5 response categories ranging from "strongly disagree" to "strongly agree", which requires the respondents to indicate a degree of agreement or disagreement with each of a series of statements related to the stimulus objects.
Semantic differential scale: a noncomparative 7-points rating scale with endpoints associated with bipolar labels that have semantic meaning.
Stapel scale: a noncomparative vertical scale for measuring attitudes that consists of a single adjective in the middle of an even-numbered range of values, usually ten.

Measurement error: the variation in the information sought by the researcher and the information generated by the measurement process employed.
Systematic error: error that affects the measurement in a constant way; stable factors that affect the observed score in the same way each time the measurement is made.

Random error: measurement error that arises from random changes or differences in respondents or measurement situations.
Reliability: the extent to which a scale produces consistent results if repeated measurements are made on the characteristics.
Validity: the extent to which differences in the observed scale scores reflect true characteristics being measured rather than systematic or random error.

Chapter 8: Questionnaire and form design

Questionnaire: a structured technique for data collection that consists of a series of questions, written or verbal, that a respondent answers. It has 3 main objectives: translate the information needed into a set of specific questions that the respondents can and will answer (aka that the respondents are both able and willing to answer); uplift, motivate, and encourage the respondent to become involved in the interview, to cooperate, and to complete the interview; minimize response error.
Response error: the type of error that arises when respondents give inaccurate answers or their answers are misrecorded or are analyzed incorrectly.

To design a questionnaire you have to: specify the information needed, specify the type of interviewing method, determine the content of individual questions, overcome the respondent's inability and/or unwillingness to answer, decide on the question structure, determine the question wording, arrange the questions in proper order, choose the form and layout, reproduce the questionnaire, and pretest the questionnaire.
Double barreled question: a single question that attempts to cover 2 issues. Such questions can be confusing to respondents and can result in ambiguous responses.
Unstructured questions: open-ended questions that respondents answer in their own words.
Structured questions: questions that prespecify the set of responses alternatives and the response format. A structured question could be multiple-choice, dichotomous, or scale.
Dichotomous question: a structured question with only 2 response alternatives, such as yes or no.
Leading question: a question that gives the respondent a clue about what answer is desired or leads the respondent to answer in a certain way.
Acquiescence bias: result of some respondents' tendency to agree with the direction of a leading question (yea-saying).
Basic information: information that relates directly to the marketing research problem.
Classification information: socioeconomic and demographic characteristics used to classify respondents.
Identification information: information obtained in a questionnaire that includes the respondent's name, postal address, e-mail address, and phone number.
Funnel approach: a strategy for ordering questions in a questionnaire in which the sequence starts with general questions that are followed by progressively specific questions in order to prevent specific questions from biasing general questions.
Branching questions: questions used to guide an interviewer (or respondent) through a survey by directing the interviewer (or respondent) to different spots on the questionnaire depending on the respondent's answers.
Pretesting: the testing of the questionnaire on a small sample of respondents for the purpose of improving the questionnaire by identifying and eliminating potential problems.

Chapter 9: Sampling design and procedures

Population: the aggregate of all elements that share some common set of characteristics and that compromise the universe for the purpose of the marketing research problem.

Census: a complete enumerations of the elements of a population or study objects

Sample: a subgroup of elements of the population selected for participation in the study.

Sampling errors: errors resulting from the particular sample selected being an imperfect representation of the population of interest.

Nonsampling errors: errors that can be attributed to sources other than sampling, such as errors in problem definition, approach, scaling, questionnaire design, survey methods, field work, and data preparation and analysis.

The sampling design process includes 5 steps: define the target population, determine the sampling frame, select a sampling technique, determine the sample size, and execute the sampling process.

Target population: the collection of elements.

Element: the object that possesses the information sought by the researcher and about which inferences are to be made.

Sampling unit: the basic unit containing the elements of the population to be sampled.

Sampling frame: a representation of the elements of the target population. It consists of a list or set of directions for identifying the target population.

Sample size: the number of elements to be included in a study.

Nonprobability sampling: sampling techniques that do not use chance selection procedures. Rather, they rely on the personal judgment of the researcher (see figure page 276).

Probability sampling: a sampling procedure in which each element of the population has a fixed probabilistic chance of being selected for the sample (see figure page 276).

Convenience sampling: a nonprobability sampling technique that attempts to obtain a sample of convenient elements. The selection of sampling units is left primarily to the interviewer.

Judgmental sampling: a form of convenience sampling in which the population elements are selected based on the judgment of the researcher.

Quota sampling: a nonprobability sampling technique that is a 2-stage judgmental sampling. The first stage consists of developing control categories or quotas of population elements. In the second stage, sample elements are based on convenience or judgment.

Snowball sampling: a nonprobability sampling technique in which an initial group of respondents is selected randomly. Subsequent respondents are selected based on the referrals or information provided by the initial respondents. This process may be carried out in waves by obtaining the referrals from referrals.

Simple random sampling: a probability sampling technique in which each element in the population has a known and equal probability of selection. Every element is selected independently of every other element, and the sample is drawn by a random procedure from a sampling frame.

Systematic sampling: a probability sampling technique in which the sample is chosen by selecting a random starting point and then picking every ith element in succession from the sampling frame.

Stratified sampling: a probability sampling technique that uses a 2-step process to first partition the population into subpopulations, or strata. Then elements are selected from each stratum by a random procedure.
Cluster sampling: first, the target population is divided into mutually exclusive and collectively exhaustive subpopulations called clusters. Then a random sample of clusters is selected based on a probability sampling technique such as simple random sampling. For each selected cluster, either all the elements are included in the sample or a sample of elements is drawn probabilistically.
Area sampling: a common form of cluster sampling in which the clusters consist of geographic areas such as counties, housing tracts, blocks, or other area descriptions.

The choice between nonprobability and probability sampling should be based on considerations such as the nature of the research, relative magnitude of non sampling versus sampling errors, and statistical and operational considerations. In exploratory research, the findings are treated as preliminary, and the use of probability sampling may not be warranted. On the other hand, in conclusive research, where the researcher wishes to use the results to estimate overall market shares or the size of the total market, probability sampling is favored.

Chapter 10: Data collection and preparation

All field work involves the selection, training, and supervision of people who collect data. Their validation and evaluation is also part of the process.
The researcher should: develop job specifications for the project, taking into account the mode of data collection; decide what characteristics the field workers should have; and recruit appropriate individuals.
Training ensures that all interviewers administer the questionnaire in the same manner so that the data can be collected uniformly.
Probing: a motivational technique used when asking survey questions to induce the respondents to enlarge on, clarify, or explain their answers and to help the respondents focus on the specific content of the interview.
Sampling control: an aspect of supervision that ensures that the interviewers follow the sampling plan strictly rather than select sampling units based on convenience or accessibility (see figure page 299).

Once data has been collected one needs to check for acceptable questionnaires, edit, code, and transcribe the data.
Editing: a review of the questionnaire within the objective of increasing accuracy and precision.
Coding: assigning a code to represent a specific response to a specific question along with the data record and column position that the code will occupy.
Consistency checks: the part of the data-cleaning process that identifies data that are out of range, are logically inconsistent, or have extreme values. Data with values not defined by the coding scheme are inadmissible.
Missing responses: values of a variable that are unknown because these respondents did not provide unambiguous answers to the question or their answers were not properly recorded. Also called missing values.
Casewise deletion: a method for handling missing responses or values in which cases or respondents with any missing values are discarded from the analysis.
Pairwise deletion: a method for handling missing responses or values in which all cases, or respondents, with any missing values are not automatically discarded; rather, for each calculation, only the cases or respondents with complete responses are considered.

Variable respecification: the transformation of data to create new variables or the modification of existing variables so that they are more consistent with the objective of the study.
Standardization: the process of correcting data to reduce them to the same scale by subtracting the sample mean and dividing by the standard deviation. A standardized variable will have a mean of 0 and a standard deviation of 1.
Recoding: recoding involves redefining the values of the variables and includes forming categories or redefining the categories of a categorical variable.

Chapter 11: Data analysis: frequency, distribution, hypothesis testing, and cross-tabulation
Frequency distribution: a mathematical distribution with the objective of obtaining a count of the number of responses associated with different values of one variable and to express these counts in percentage terms.
Measures of location: a statistic that describes a location within a data set. Measures of central tendency describe the center of the distribution.
Mean: a measure of the central tendency given as the average.
Mode: a measure of central tendency given as the value that occurs the most in a sample distribution.
Measures of variability: statistics that indicate the distribution's dispersion.
Range: the difference between the largest and smallest values of a distribution.
Variance: the mean squared deviation of all the values from the mean.
Standard deviation: square root of the variance.

The following steps are involved in hypothesis testing: Formulate H_0 and H_1, select appropriate test, choose level of significance α, collect data and calculate the test statistic, determine probability associated with the test statistic, compare with the level of significance $\alpha(\alpha/2)$ reject or not H_0 draw marketing research conclusions.
Null hypothesis: a statement suggesting no expected difference or effect. If the null hypothesis is not rejected no changes will be made.
Alternative hypothesis: a statement suggesting some difference or effect is expected. Accepting the alternative hypothesis leads to changes in opinions or actions.
One-tailed test: a test of the null hypothesis where the alternative hypothesis is expressed directionally.
Two-tailed test: a test of the null hypothesis where the alternative hypothesis is not expressed directionally but is bidirectional.
Test statistic: a measure of how close the sample has come to the null hypothesis.
Normal distribution: bell-shaped and symmetric. Its measures of central tendency are identical.
Z-test: a hypothesis test using the standard normal distribution.
Type I error: an error that occurs when the sample results lead to the rejection of a null hypothesis that is, in fact, true. Also known as alpha (α) error.
Level of significance: the probability of making a type one error, denoted by α.
Type II error: an error that occurs when the sample results lead to nonrejection of a null hypothesis that is, in fact, false. Also known as beta (β) error.
Power of a test: the probability of rejecting the null hypothesis when it is in fact false is of 1- β
P-value: the probability of observing a value of the test statistic as extreme as, or more extreme than, the value actually observed, assuming that the null hypothesis is true.
Cross-tabulation: a statistical technique that describes two or more variables simultaneously and results in tables that reflect the joint distribution of 2 or more variables that have a limited number of categories or distinct values.

Chi-square statistic: the statistic used to test the statistical significance of the observed association is a cross-tabulation. It assists in determining whether a systematic association exists between the 2 variables.
Chi-square distribution: a skewed distribution whose shape depends solely on the number of degrees of freedom. As the number of degrees of freedom increases, the chi-square distribution becomes more symmetrical.
Phi coefficient: a measure of the strength of association in the special case of a table with 2 rows and 2 columns.
Contingency coefficient: a measure of the strength of association in a table of any size, whose maximum value depends on the size of the table.
Cramer's V: a measure of the strength of association used in tables larger than 2x2, whose value ranges from 0 to 1.
see figure page 346

Chapter 12: Data analysis: hypothesis testing related to differences, correlation, and regression
Parametric tests: hypothesis-testing procedures that assume the variable of interest are measured on at least an interval scale (see figure page 359).
T-test: a hypothesis test using the t distribution, which is used when the mean is known, the standard deviation is unknown and is estimated for the sample (see figure page 360).
T statistic: a statistic that assumes the variable has a symmetric, bell-shaped distribution; the mean is known (or assumed to be known); and the population variance is estimated from the sample.
Standard error: the standard deviation of the mean or proportion
T distribution: a symmetric bell-shaped distribution that is defined by n-1 degrees of freedom.
Independent samples: two samples that are not experimentally related. The measurement of one sample has no effect on the values of the other sample (see figure page 364).
F test: A statistical test of equality of the variances of 2 populations.
F statistic: a statistic that is calculated as the ratio of two sample variances by dividing the larger sample variance by the smaller sample variance.
F distribution: a frequency distribution that depends on 2 sets of degrees of freedom: the degrees of freedom in the numerator and the degrees of freedom in the denominator.
Paired samples: in hypothesis testing, the observations are paired so that the 2 sets of observations relate to the same respondents.
Paired-sample t test: a test for differences in the means of paired samples.
Product moment correlation (r): a statistic summarizing the strength of linear association between 2 metric variables.
Regression analysis: a statistical procedure for analyzing associative relationships between a metric dependent variable and one or more metric independent variables.
Bivariate regression: a procedure for deriving a mathematical relationship, in the form of an equation, between a single metric dependent variable and a single metric independent variable.
Multiple regression: a statistical technique that simultaneously develops a mathematical relationship between a single metric dependent variable and 2 or more metric independent variables.
Beta coefficient: used to denote the standardized regression coefficient. Also known as beta weight.
Coefficient of multiple determination: in multiple regression, the strength of association is measured by the square of the multiple correlation coefficient, R^2, which is called the coefficient of multiple determination.
Adjusted R^2: the value of R^2 adjusted for the number of independent variables and the sample size. The value of adjusted R^2 cannot be higher than R^2