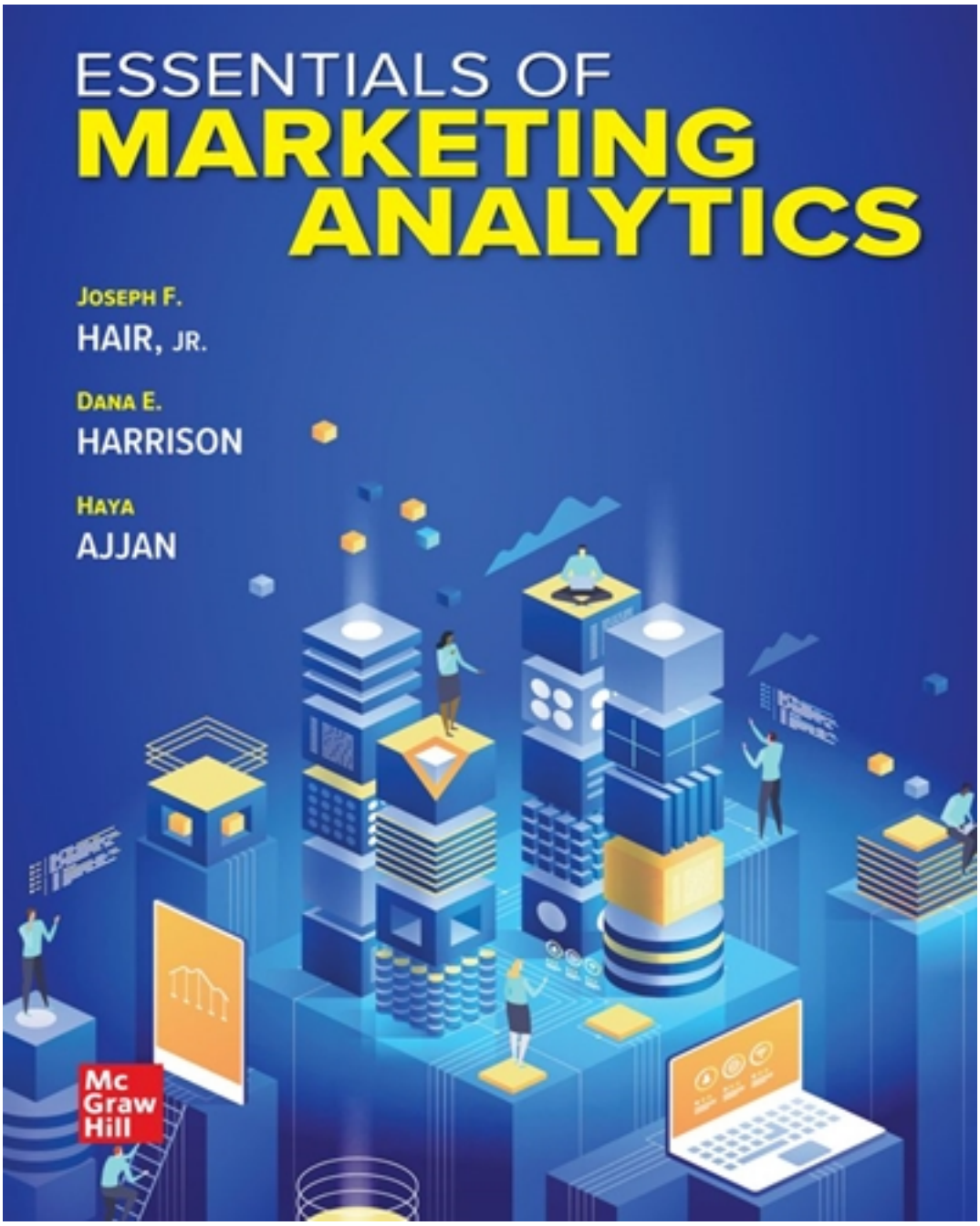
Solutions for Essentials of Marketing Analytics 1st Edition by Hair

CLICK HERE TO ACCESS COMPLETE Solutions



Solutions

Chapter 01
Introduction to Marketing Analytics

Part 01: Overview of Marketing Analytics and Data Management

Chapter 01 Learning Objectives

LO 01-01: Discuss marketing analytics.
LO 01-02: Discuss how to identify the right business problem.
LO 01-03: Identify and compare different data sources.
LO 01-04: Describe different data types.
LO 01-05: Explain the difference between predictors and target variables.
LO 01-06: Differentiate between supervised and unsupervised modeling.
LO 01-07: Investigate the 7-step marketing analytics process.
LO 01-08: Explain the value of learning marketing analytics.

Key Terms

Artificial intelligence (AI): Artificial intelligence (AI) is a branch of computer science that is designed to mimic human-like intelligence for certain tasks, such as discovering patterns in data, recognizing objects from an image, understanding the meaning of text, and processing voice commands.

Binary: Binary categorical data can have only two values—for example, yes or no. This can be represented in different ways such as 1 or 0 or "True" and "False." Binary data is commonly used for classification in predictive modeling.

Categorical data: Categorical data exists when values represent a group of categories. Categorical variables can be one of three types: binary, nominal, or ordinal.

Cognitive analytics: Cognitive analytics uses machine learning to understand new data and patterns that have never been identified.

Continuous data: Continuous data includes values with decimals: 1, 1.4, 3.75, . . .

Dependent variable (also **target** or **outcome**)**:** The dependent variable is the target variable (y). Its outcome is impacted by other variables.

Descriptive analytics: Descriptive analytics are a set of techniques used to explain or quantify the past.

Discrete data: Discrete data is measured in whole numbers (integers): 1, 2, 3, . . .

Independent variable: The independent variable is the predictor or feature variable (x). This variable could potentially influence or drive the dependent or outcome variable(s).

Integer: An integer is a whole number.

Interval: Interval data has an equal distance between data points and does not include an absolute zero.

Marketing analytics: Marketing analytics uses data, statistics, mathematics, and technology to solve marketing business problems.

Machine learning: Machine learning is a statistical method of learning that can be trained without human intervention to understand and identify relationships between previously established variables.

Nominal: Nominal categorical data consist of characteristics that have no meaningful order.

Ordinal: Ordinal categorical data represent meaningful values with a natural order but the intervals between scale points may be uneven.

Predictive analytics: Predictive analytics is used to build models based on the past to explain the future.

Prescriptive analytics: Prescriptive analytics identifies the best optimal course of action or decision.

Primary data: Primary data is collected for a specific purpose. For example, companies conduct primary research with surveys, focus groups, interviews, observations, and experiments to address problems or answer distinct questions.

Ratio: Ratio values can have an absolute zero point and can be discussed in terms of multiples when comparing one point to another.

Secondary data: Secondary data relies on existing data that has been collected for another purpose.

SMART principles: SMART principles are used as a goal-setting technique. The acronym stands for specific, measurable, attainable, relevant, and timely.

Structured data: Structured data is made up of records that are organized in rows and columns. This type of data can be stored in a database or spreadsheet format.

Supervised learning: In supervised learning, the target variable of interest is known and is available in a historical dataset.

Testing dataset: A testing dataset is used to evaluate the final selection algorithm on a dataset unique from the training and validation datasets.

Training dataset: The training dataset is the data used to build the algorithm and "learn" the relationship between the predictors and the target variable.

Unstructured data: Unstructured data does not have a predefined structure and does not fit well into a table format (within rows and columns).

Unsupervised learning: Unsupervised learning has no previously defined target variable. The goal of unsupervised learning is to model the underlying structure and distribution in the data to discover and confirm patterns in the data.

Validation dataset: The validation dataset is used to assess how well the algorithm estimates the target variable, and helps select the model that most accurately predicts the target value of interest.

Variables: Variables are characteristics or features that pertain to a person, place, or object. Marketing analysts explore relationships between variables to improve decision making.

Chapter 01 Outline

I. Introduction to Marketing Analytics

A primary responsibility of marketing is to properly manage the wants and needs of customers

- Increasingly, companies are using marketing analytics to satisfy customers and maintain a competitive advantage
  - People-related problems, in turn, frequently stem from the mistaken belief that people are all alike, that they can be treated identically.
  - Like snowflakes, no two people are exactly alike, and everyone differs physically and psychologically from everyone else.
  - The point is that these differences demand attention, so that each person can maximize their potential, organizations can maximize their effectiveness, and society as a whole can make the wisest use of its human resources.
- **How does Expedia, Orbitz, or Hotels.com determine the price to quote when you shop for a hotel room?**
  - Hotel room prices are frequently updated based on demand, seasonality, day of the week, time of the day, and even the type of technology used
    - Orbitz knows Mac users spend as much as 30 percent more a night on hotels
    - So Orbitz shows Mac customers different travel options than Windows users
- **How does Spotify know what songs to suggest to you?**
  - From user-generated playlists, listener preferences, and advanced data analytics
- **How does Stitch Fix achieve the highest-ever rate of purchased items per "Fix" for its female customers?**
  - Their stylists work closely with the analytics algorithm suggestions
  - Then match results with the customer's style
  - Over time, the algorithm learns and becomes more accurate when making clothing suggestions, stocking decision, packing at the warehouse, and shipping

In the rest of this chapter, we describe and explain

- An analytics framework
- The relevant marketing analytics concepts
- Industry best practices

Building on this foundation, you will continue to work through practical exercises and develop the mindset of a marketing analyst

A. Marketing Analytics Defined

Marketing analytics uses data, statistics, mathematics, and technology to solve marketing business problems

- It involves modeling and software to drive marketing decisions
- Today, marketing analytics is practical and available to a large audience due to
  - The availability of large amounts of data
  - Improvements in analytics techniques
  - Substantial increases in computer processing power
  - Affordability

Marketing analytics is one of the fastest growing fields of analytics applications, due to

- The increase in user-generated data from social media and mobile apps
- And multiple search and shopping channels accessible by customers
  - Phone, in-store, online
- Marketers can use insights from analytics to increase company performance through various marketing capabilities
  - Pricing
  - Product development
  - Channel management
  - Marketing communications
  - Selling
    - For example, a restaurant may apply marketing analytics to optimize the selection of new locations

Marketing analytics is increasingly being applied in numerous industries and functional departments, and the impact and benefits are evident

- Exhibit 1-1 compares the interest in marketing analytics to analytics use in other business functions – based on Google Trends
- A large amount of marketing data exists, which explains the interest
  - Yet, many companies are "data rich but information poor" due to lack of marketing analytics skills
- Many organizations historically made decisions based on intuition or opinion
  - Data analytics techniques provide an opportunity to bridge the gap between information and insights
- As technology continues to improve and dominate innovate processes, analytics will become a ubiquitous part of everything we do
  - To prepare for this, we explain how to
    - Creatively approach a problem
    - Comprehend the essence of communication and collaboration
    - Understand key elements of project management
    - Complete a successful project

B. Analytics Levels and Their Impact on Competitive Advantage

Analytics involves techniques as simples as descriptive statistics and visualization

- As organizations adopt more advanced techniques, such as predictive, prescriptive, and AI methods
    - Higher data management and analysis maturity are required to achieve a competitive advantage, as depicted in Exhibit 1-2

Descriptive analytics are a set of techniques used to explain or quantify the past

- Examples: data queries, visual reports, descriptive statistics, such as mean, median, variance, standard deviation
- This type of information is essential to summarize questions related to how many and how often situations occur
    - How many customers use a mobile app each day?
    - How often do they visit a website within the same month?
- Can be especially helpful when marketers collect data from a survey as averages and trends can be beneficial in
    - Reinforcing existing practices
    - Determining how the company might improve the customer's experience

Predictive analytics is used to build models based on the past to explain the future

- Mathematical models examine historical data to predict new values, needs, and opportunities
    - For example, historical customer sales data can predict future sales
    - Target Corporation collects data from customer purchases to classify customers as "pregnant" or "not pregnant" and pregnant customers are sent special offers
    - Zillow develops predictive models from publicly available city housing data to provide a "Zestimate"

Prescriptive analytics identifies the best optimal course of action or decision

- For example
    - UPS maps drivers through a city using optimized routes that reduce left turns
    - Airlines optimize flight and crew scheduling to maintain productivity, reduce costs, and increase customer satisfaction
    - Price optimization is used by Amazon to develop pricing strategies
        - The company has reported changing prices more than 2.5 million times a day
        - Done to influence customer behavior and maximize revenue
    - Kellogg Company uses optimization models to forecast sales and determine what should be produced and shipped on a daily basis
        - They also use optimization modeling to improve supply chain infrastructure
        - Kellogg estimates it saves over $475 million a year using the modeling

Artificial intelligence (AI) and **cognitive analytics** are designed to mimic human-like

intelligence for certain tasks, such as

- Discovering patterns in data
- Recognizing objects from an image
- Understanding the meaning of text
- Processing voice commands
    - This type of analytics uses **machine learning** to understand new data and patterns
        - This method produces tasks that are often beyond the reach of a human
        - The techniques "learn" over time by updating the algorithm as new information becomes available

Using technology powered by AI

- Olay skincare doubled its sales conversion rate
    - Customers upload personal photos and Olay recommends specific products
    - Done in the privacy of the customer's home
- Mtailor uses customer images to create a customized clothing fit
    - The app determines sizes and recommends sizes
    - They claim their method is more accurate than a professional tailor
- In both cases, AI engages the customer to obtain data and then produces personalized product recommendations

Today's AI technology enables almost any company to augment human capabilities in developing customer solutions

- Hitachi uses AI named "H" to discover patterns typically undetected by humans
    - The H process generates customer solutions
    - And selects the best options to improve
        - Operations at call centers
        - Retail sales
        - Financing alternatives
        - Warehouse management
        - And more
- Applications like "H" can easily automate and improve customer interactions to increase sales and customer loyalty

II. Defining the Right Business Problems

One of the most important initial steps of marketing analytics is defining the right business problem to pursue

- Understanding business problems requires deep knowledge of the customer's journey
    - From how they search
    - To where they purchase
    - And how satisfied they are with products and services

- Problem identification uncovers strategic business opportunities
  - To improve market share
  - To establish a better customer relationship
  - Or position the enterprise to take advantage of innovation

How do you arrive at the right business problem?

- Begin by understanding the intent and business considerations behind the question
  - Suppose your boss tasks you with determining how to entice first-time users back who sign up for the new mobile app
    - After listening to the relevant issues and interviewing relevant stakeholders
    - You determine the company should not limit the investigation to attracting new customers
    - But should expand it to encompass how to retain loyal customers
  - If you had proceeded with the original project focus of returning visitors to the mobile app, it would have overlooked important customer relationships
- Marketers must incorporate relevant stakeholder inputs through discovery methods to collectively understand the business problem and align projects to achieve objectives
  - Discover begins with the traditions six questions: what, who, where, when, why, and how
  - Exhibit 1-3 provides samples of discovery questions and how they might be useful

In an effort to define the right business problem, it can be useful to follow the SMART analytics principles

- The **SMART principles** can be used as a goal-setting technique
  - The acronym stands for Specific, Measurable, Attainable, Relevant, and Timely
    - See Exhibit 1-4
  - The project's goals should be specific and clearly defined
  - Second, the project should be trackable and the outcomes measurable
  - Third, project goals should be reasonable to achieve
  - Fourth, the project should solve the analytics problem and align with the business objectives
  - Fifth, the project should be completed in a timely manner
- Following the SMART analytics goal-setting technique is important, but equally important is
  - Examining the potential success of the analytics project and whether it makes a valuable impact
    - To do so, the opinions of the most powerful stakeholders should be included when developing project goals and success measures
      - As well as in evaluating the results
- When the SMART analytics goals are identified, it is time to focus on understanding

the data requirements – we begin by looking at data sources

III. Data Sources

Data sources consist of both primary and secondary data

- **Primary data** is collected for a specific purpose
  - Surveys, focus groups, interviews, observations, and experiments address problems or answer distinct questions and achieve specific objectives
- **Secondary data** relies on existing data collected for another purpose
  - It can be useful in formulating ideas about how to ask the right questions
    - Or to design future data collection initiatives
  - Internal and external secondary data sources can be useful in exploring current business questions
  - Sources of secondary data include
    - Public datasets
      - Google launched Google Dataset Search in 2018
      - Includes data from NASA, NOAA, Harvard's Dataverse, GitHub, Kaggle, and other sources
    - Online sites
      - Browsing behavior, purchase history, and social media posts have become increasingly popular sources of data
    - Mobile data
      - Most mobile apps track data
      - Restaurants and clothing retailers developed apps that record customer purchase behaviors and geographic locations
    - Channel partners
      - Multiple companies often operate within a distribution channel
      - These channel partners include suppliers, wholesalers, distributors, or retailers – each collecting data unique to their business
      - But the data frequently provides value to other partners in the channel
    - Commercial brokers
      - Companies collecting and selling both public and private data to a wide range of customers emerged in recent years
      - Other companies purchase this data to create customer profiles used to target a wide variety of target segments
    - Corporate information
      - Many companies constantly collect and store data
      - Integrating data across functional areas enables companies to better understand customer interactions based on a holistic view of transactions
    - Government sources
      - This is an important source of secondary data collected by local, state, and

federal government agencies
- More than 200,000 datasets are searchable by topic – see Exhibit 1-5
- The following are directly applicable to marketing analysts
  - The U.S. Census Bureau includes data on population, economy, housing, and geography
  - Consumer complaint data covers financial products and services
  - Demographic statistics by ZIP code, gender, ethnicity, and citizenship
  - Fruit and vegetable prices for many commonly consumed products are available at the Department of Agriculture
  - ZIP code data showing tax return data by state and ZIP code level

IV. Data Types

A. Types of Data

- Data is facts and figures collected, organized, and presented for analysis and interpretation
- Data is available in two main forms: structured and unstructured

Structured Data

- **Structured data** is made up of records organized in rows and columns
- This type of data can be stored in a database or spreadsheet format
- Includes numbers, dates, and text strings stored in a defined structure
- Easy to access and analyze using descriptive, predictive, prescriptive, and AI data analytics techniques

Unstructured Data

- **Unstructured data** includes text, images, videos, and sensor data
- No defined structure and content does not fit into a table format
- Examples of this type of data include
  - Voice recording from customer service calls
  - Text, images, and videos
  - Social media conversations
  - The Internet of Things sensor data
- Unstructured data requires advanced analytics techniques such as AI to prepare and analyze – may be converted to structured data before analysis
- More companies are collecting unstructured data
  - Technology has advanced to efficiently support manipulation and exploration of this type of data

- Both structured and unstructured data are important in marketing analytics

B. Data Measurement

Data measurement can be categorized based on the type and means for collection
  - See Exhibit 1-6
- The two main types of data measurement we explore in the remaining chapters are numerical and categorical – see Exhibit1-7

Numerical
  - Data are considered quantitative if numeric and arithmetic operations can be applied
    - For instance, sales data or visits to websites are numerical as they can be summed or averaged
  - Numerical data can be either *discrete* (**integer**) or *continuous* in nature
    - **Discrete data** is measured as whole numbers: 1, 2, 3…
      - Number of items purchased on a retailer's website is discrete
    - **Continuous data** can include values with decimals: 1, 1.4, 2, 2.5, 3.75…
      - Time spent on a retailer's website would be continuous

Categorical
  - **Categorical data** exist when values are selected from a group of categories
    - A common example might be marital status
  - This data can only be summarized by calculating the proportion and count of occurrences across and within categories
  - Categorical variables can be one of three types
    - **Binary** categorical data can have two values – for example, yes or no
      - Commonly used for classification in predictive modeling
      - Examples include
        - Whether a person has purchased a product or not
        - Whether a person uses Twitter or not
    - **Nominal** categorical data consist of characteristics that have no meaningful order – each category's characteristics are equally meaningful
      - The characteristics reflect the state of being
      - If the customer's country is asked for, it could be any number of responses
      - If marital status is inquired about it could be several options
    - **Ordinal** categorical data represent meaningful values
      - They have a natural order, but the intervals between scale points may be uneven
      - Customers may be asked if they prefer a product more or less than another product
      - Or customers may be asked to rank products in order of preference
  - Categorical variables require special consideration in preparation for modeling
    - Their preparation will be discussed in a later chapter

C. Metric Measurement Scales

Scales can also be metric

- Metric scales can be measured as **intervals** or **ratios**
  - Both of these scales possess meaningful, constant units of measure
  - And the distance between each point of the scale are equal
- However, there is a difference between these scales
  - Interval variables do not include an absolute zero
  - Ratio scales have an absolute zero point and can be discussed in terms of multiples
    - Sales of $0 is nothing and sales of $100 is twice as much as $50
    - Zero indicates a lack of weight, and 50 pounds is half of 100 pounds

V. Predictors versus Target Variable

A. Types of Variables

- **Variables** are characteristics or features that pertain to a person, place, or object
- Marketing analysts explore relationships between variables to improve decision making – consider a simple example
  - Does the weather impact customer ice cream purchases?
    - Weather conditions would be the **independent variable**
      - Or what influences or drives the **dependent**, **target**, or **outcome variable** which is ice cream sales
    - Warmer weather increases ice cream sales
    - What other variables might impact ice cream sales?
      - Though Exhibit 1-8 only uses two variables
      - Companies may use multiple variables

VI. Modeling Types: Supervised Learning versus Unsupervised Learning

Depending on the business problem, different types of algorithms can be used –this book, we focus on two types: *supervised learning* and *unsupervised learning*

- **Supervised learning** suggests the target variable of interest is known (ice cream sales; click or no click) and available in a historical dataset divided into three datasets
  - **Training dataset**
    - This is the data used to build the algorithm and "learn" the relationship between the predictors and the target variable
  - **Validation dataset**
    - The resulting algorithm is then applied to this dataset to assess how well it estimates the target variable
      - And to select the model that most accurately predicts the target value

    - **Testing dataset**
        - If many different algorithms are being compared, a testing dataset should be used to evaluate the final selected algorithm
- The final selected algorithm is then applied to predict the target variable using new unlabeled data where the outcomes are not known, as shown in Exhibit 1-9
    - When the target variable is continuous
        - Supervised learning is referred to as *prediction*
            - A retailer wants to understand buying behavior, specifically the purchase amount to create personalized offers
            - Using labeled data, a model is created to predict the purchase amount of customers against various products
                - Historical data showing how much the customer spent on each product
            - The model is then used to predict the purchase amount of new customers with similar characteristics
    - When the target variable is categorical (typically binary-buy/no buy)
        - Supervised learning is called *classification*
            - Consider a large U.S. bank whose objective is to acquire new credit cardholders based on a special promotion
                - Historical data includes records of customers who qualified (or not) for past credit card offers after receiving special promotions
            - An algorithm would be trained and validated on labeled data
            - Then used to predict who should be targeted in the new promotion
    - Exhibit 1-10 shows other examples of supervised learning applications using a variety of predictors
- **Unsupervised learning** has no previously defined target variable
    - The goal of unsupervised learning is to model the underlying structure and distribution in the data to discover and confirm patterns
        - Examples of these techniques include
            - Using association analysis, Sephora's website offers "you may also like product X" based on past purchases
            - Amazon indicates "others who bought this item also bought products X, Y, and Z
        - This may also involve cluster analysis to group customers into sets based on loyalty using purchase history, annual amount spent, and other key variables

Supervised and unsupervised learning can be used together to gain more insights

- For example
    - After conducting the unsupervised learning to determine customer loyalty sets
    - Supervised learning can be used to predict purchase amounts for each segment
- Whether the algorithm is supervised or unsupervised, the modeling must represent how real-world problems begin

  - Starting with a business problem
  - And working toward a solution that makes a business impact
- Modeling steps are discussed in detail in the following section

VII. The 7-Step Marketing Analytics Process

Modeling is the beginning of the journey, not the end – see Exhibit 1-11

- The 7-step marketing analytics process is iterative and continuously evolves to develop and manage improvements in the modeling cycle
- Each step plays an important role in achieving a successful outcome

A. Step 1: Business Problem Understanding

Most marketing analytics models are developed when a business identifies a problem

- The idea is to develop a model using analytics to understand the problem better and design a solution
- A key element is to question whether the problem is the correct problem
  - Exactly what are you trying to understand and solve?
  - How will the stakeholder(s) use the results?
  - Who will be affected by the results?
  - Is this a single, short-term problem or an ongoing situation?

B. Step 2: Data Understanding and Collection

The marketing analyst's first job is to identify where the data is stored, its format, and how it can be combined to understand the question at hand, typically includes

- Examining internal and external databases
- Then talking with key data owners and stakeholders
- Observing and understanding organizational processes to determine if the identified problem is the actual problem or a symptom of another problem

Once a better understanding of the problem is established, the analyst typically samples data from the selected database to obtain records for the analysis

- The marketing analyst may use SQL code (Chapter 2) to examine past purchases and returns of customers
- The Practitioner Corners sections of the book illustrate how marketing analysts collect, clean, and prepare data for analysis
  - Basic tasks in the overall process of identifying and solving business problems

Marketing analysts must have a good understanding of the types and sources of data

- The Statistical Research Group (SRG) engaged in solving problems during WWII
  - They tried to determine the placement of additional armor on American bombers

- They hired a mathematics and statistics expert, Abraham Wald
- Prior to hiring Wald, the SRG examined data on damage suffered by bombers – see Exhibit 1-12
  - They noticed the damage was mostly in the fuselage and decided to include more protection in this area
  - Wald noticed the data was from a single source – airplanes who returned
    - No data was available from destroyed planes
    - He obtained information from pilots who parachuted out of planes that failed to return
- If Wald had not questioned or been aware of the data's origin, the decision on armor placement could have been catastrophic or possibly ineffective

Clearly, this is an important step today as well

- What if Apple failed to test the strength of iPhone screens when dropped?
- What if Tyson did not track its chicken products so salmonella outbreaks can be tracked and resolved?

C. Step 3: Data Preparation and Feature Selection

Data in different formats is combined in this step

- To do this
  - The unit of analysis – customer, transaction, subscription plan
  - And the target and predictor variables are identified
  - The data columns of the target and predictor variables is then visually and statistically examined
    - For example, scatterplots can reveal patterns
  - Data is cleaned by identifying and determining how to deal with missing values, data errors, or outliers
  - Data from different data sources is merged so data is measured consistently and can be used to develop the models
- Other features are further refined in this step, such as adjusting date formats
  - Predictors with a strong relationship with the target variable yet unique and not highly related are included to improve the reliability and accuracy of the model
  - Predictors might be eliminated, but they could also be transformed to improve the measurement – for example
    - If the focus of the problem is mobile phones priced less than or equal to $200 and those greater than $200
    - Rather than examine continuous monetary values, the feature can be changed to a binary variable
      - =<$200 or >$200
- Understanding the meaning of each variable and its unit of analysis is essential here

D. Step 4: Modeling Development

Steps 1 through 3 represent about 80 percent of the analyst's time, but serve as an important foundation for the rest of the steps in the process

- In step 4, the analyst uses analytical skills
- In this step, the analyst selects the method to use
  - Choice depends on the target variable and problem
  - Possible options are classification, prediction, clustering, or association
- If the problem is supervised, the analyst needs to partition the data into datasets
  - Training
  - Validation
  - Testing
- The analyst will decide on appropriate modeling techniques
  - Such as regression and neural network – explained in later chapters
  - More than one modeling technique is usually used, each with a variety of features
- Different models should be tried to find the one giving accuracy, speed, and quality
  - The model should be simple, practical, and useful
  - Some commonly encountered analytics problems can be solved using decision trees and regression analysis
    - Not as accurate, but simple to understand and apply

E. Step 5: Model Evaluation and Interpretation

This step ensures the modeling is accurately performed and provides the best predictions for future applications

- The model is evaluated to identify the algorithm providing the best solution
  - The algorithm is initially run on the validation dataset
  - If the validation shows high accuracy
    - The model can be recommended to predict new cases and address the problem

F. Step 6: Model and Results Communication

This modeling step provides a set of recommendations

- It is key for the analyst to present the model in a way other people can understand, particularly management
- A good approach is to collaborate with key stakeholders early in the process
  - If executives and managers have been providing feedback from the beginning, they are more likely to understand and support the recommended model
- A full understanding of the model is another important consideration
  - Whether simple or complex, the model should be explainable in straightforward terms with the appropriate visualization of results

G. Step 7: Model Deployment

The model is not finished until it has been implemented and running on real-time records to offer decisions or actions

- Only after full buy-in will the model add real value in making decisions
- This step involves other key stakeholders – IT specialists or the sales team
  - These stakeholders should be trained to implement the system

A key consideration throughout the 7-step marketing analytics modeling process is to evaluate the ethical dimensions of the analysis

- Are the privacy and anonymity of the subjects being protected?
- Does a bias exist in the data that could impact the analytics results?
- Are the model results accurate?
- At times, the model may be correct, but the objective is unfair to some subjects or unrealistic in its predictions
- Another issue is that the data, features, data cleaning, and the model are determined by analysts
  - Ethical training, ethical codes, and clear guidelines should be established and communicated to everyone working on developing the analytical model

VIII. Setting Yourself Apart

Marketing analytics is essential for students interested in distinguishing themselves in the job market

- The number of postings for analytics-enabled jobs will continue to rise
- The average salary for analytics positions is $99,000
- Demand and salary will continue to grow worldwide

The remaining chapters discuss marketing problems using a variety of powerful analytics software

- The software shown in Exhibit 1-13 is useful in many careers, particularly marketing, and are currently used by many organizations globally to solve complex problems
- Exhibit 1-14 displays the top software and the relative usage

This book provides the foundation and skills for successfully using analytical methods in marketing

- Chapter cases cover a variety of tools and techniques to provide hands-on experiences
- Application of the analytics methods serve as a toolbox for approaching different problems
- Exhibit 1-15 provides an overview of the chapters by topic, modeling type, software, and coverage of the four major areas of analytics
  - Descriptive, predictive, prescriptive, and AI/cognitive

Sample Answers to Discussion and Review Questions

1. **What is marketing analytics?**

   Marketing analytics uses data, statistics, mathematics, and technology to solve marketing business problems. It uses modeling and software to drive marketing decision making. Today's availability of large amounts of data, improvements in analytics techniques, substantial increases in computer processing power, and affordability have made marketing analytics more practical and available to a much larger audience. Until recently, many organizations made decisions based on intuition or opinion versus data-driven knowledge. Data analytics techniques provide an excellent opportunity to bridge the gap between information and insights.

2. **How are companies using marketing analytics to make strategic marketing decisions?**

   Companies increasingly need to differentiate products and services, optimize processes, and understand the drivers for business performance, and marketing analytics can help them do so. Marketers can use insights from analytics to increase company performance through pricing, product development, channel management, marketing communications, and selling. Restaurants have begun applying marketing analytics to optimize location selection. The restaurant chain Roy Rogers Franchise Co. uses advanced analytics to expand into new markets, determine their next site locations, and forecast sales. Their machine learning platform integrates internal and external data to ensure restaurant locations match the needs and wants of the geographical area.

3. **Name several external data sources that might be helpful to marketers.**

   Google Dataset Search includes data from NASA, NOAA, Harvard's Dataverse, GitHub, Kaggle, and other sources. Social media content can be mined for words or phrases associated with a product or company. Companies can gather data collected by distribution channel partners. Commercial brokers collect and sell both public and private data including socioeconomic status, health interests, and political views. The government provides more than 200,000 datasets searchable by topic. The U.S. Census Bureau provides data on population, economy, housing, and geography. Another dataset compiles consumer complaints about financial products and services. There are demographic statistics by ZIP code, gender, ethnicity, and citizenship. A dataset collects fruit and vegetable prices for many commonly consumed items through the Department of Agriculture. There is also a dataset for tax return data by state and ZIP code level.

4. **How might a company use structured and unstructured data to better understand customers?**

Structured data includes numbers, dates, and text strings stored in a clearly defined structure of rows and columns. Unstructured data includes text, images, videos, and sensor data and does not fit well into a table format. Unstructured data required advanced analytics techniques such as AI to prepare and analyze. Companies can combine unstructured and structured data to better understand their customers. Combining information from conversations on support sites with information from individual customer accounts deepens customer understanding. Technology has improved the combination of these kinds of data structures but user knowledge are lagging. AI and machine learning will be increasingly used to help combine and analyze this data.

5. Define a target variable.

Variables are characteristics or features that pertain to a person, place, or object. Marketing analysts explore relationships between variables to improve decision making. Say an analyst is investigating the relationship between two variables: store lighting and total sales per customer. The store lighting is considered the independent variable or what influences or drives the dependent, target, or outcome variable – sales per customer. The target variable is sales per customer and variables other than lighting may be affecting sales per customer. Companies can use multiple variables at the same time as inputs to systems that process data and use it to predict target variables.

6. Discuss the difference between supervised and unsupervised learning.

Both are types of algorithms used to address business problems. Supervised learning suggests the target variable of interest is known and available in a historical dataset. The historical dataset, or labeled data, is divided into a training dataset, a validation dataset, and an optional testing dataset. Unsupervised learning has no previously defined target variable. The goal of unsupervised learning is to model the underlying structure and distribution in the data to discover and confirm patterns.

7. What are the steps of the marketing analytics process?

The 7-step marketing analytics process is iterative and continuously evolves to develop and manage improvements in the modeling cycle. Each step plays an important role in achieving a successful outcome. The steps in the process are:

- Step 1: business problem understanding
    - A key element here is to question whether the problem the business is presenting is, in fact, the correct problem
- Step 2: data understanding and collection
    - This step includes examining internal and external databases and talking with key data owners and stakeholders
    - It is also determined if the identified problem is the actual problem or a

symptom of an underlying problem.

- Step 3: data preparation and feature selection
  - Data in different formats is cleaned and combined in this step
  - Understanding the meaning of each variable and its unit of analysis is an essential task in this step
- Step 4: modeling development
  - In this step, the analyst selects the method to use
  - Different models should be tried to identify the one providing the best accuracy, speed, and quality
- Step 5: model evaluation and interpretation
  - This step evaluates the model to identify the algorithm providing the best solution
- Step 6: model and results communication
  - The analyst needs to present the model in a way that other people can understand, particularly management
- Step 7: model deployment
  - The model must now be implemented and run on real-time records to offer decisions or actions

Critical Thinking and Marketing Applications

1. **Visit www.data.gov. Click on Consumer, then click on Data. How many datasets are currently located on this website for free? Select one dataset and develop a scenario where the data might be helpful for a marketing manager. Discuss how exploring the data could guide the marketing manager in making more informed decisions.**

   There are 218,078 datasets on the site and 113 with the topic of "Consumer," as of November 2020. Suppose you work for a company that supplies tents for weddings and other gatherings. Top management is thinking about branching out into making awnings for patios and decks. You locate the Characteristics of New Housing dataset which supplies you with the information that a third of newly constructed single-family homes included a patio and a porch while only eight percent had no outdoor features. You could explore the data further looking for income levels in your area or home improvements. You could also look at weather patterns. Sometimes exploring the data may spark thoughts on other products you would not have thought of otherwise. For example, in looking at the new construction for single-family homes, you notice nearly 90 percent have a central air conditioning unit. Maybe adapt the tents you now produce to fit various air conditioning units for winter storage. Or pool covers. Looking at the data can confirm theories or bring to light customer aspects you had not considered.

2. Develop two questions that an airline company might be interested in answering. Describe types of unstructured and structured data that might be important to answering the questions. What data sources might be helpful?

How many times a month does a customer report lost baggage and does this affect the customer from flying with our airline in the future?

The airline can use structured data to find how often in a month a customer reports lost luggage, this should be an actual number. Repeat customers are tracked using frequent flyer programs or other customer loyalty programs, which would be structured, accessible data. As for unstructured data, the airline could use video of the customer interaction when reporting the lost luggage to gauge the customer's emotional reaction, listening for threats of "I'll never fly this airline again." The company could also mine social media sites for search terms including their name and lost luggage for any rants (or compliments) about their airline. The airline could also collect primary data by surveying customers in the frequent flyer program to determine how many have reported lost luggage with the airline. Meshing the internal data and external data may require advanced analytics techniques using AI to prepare and analyze.

How many personal injuries occurred during the boarding and departing processes each month and how can we increase safety during these times?

Structured data would include all recorded injuries during boarding and departing. Unstructured data could be interviews conducted with the injured customer, video footage of the accident, and written injury reports. Analysts charged with increasing safety during these times can use prior incidents to predict, and prevent future injuries. An external source that may prove useful here is the NTSB Aviation Safety News, available as an RSS feed.