

Demand and Market Research for Media and Information Products

9.1 Why Demand Analysis – 234

- 9.1.1 The Importance and Special Problems of Demand Estimation for Media Industries – 234
- 9.1.2 Examples for the Problems in Forecasting Media Demand – 235
- 9.1.3 Limits to Audience and Market Research – 236
- 9.1.4 How Media Companies Organize Their Market Research – 237
- 9.1.5 Case Discussion – 238

9.2 Data Collection – 238

- 9.2.1 The Impact of Collection Methodology – 238
- 9.2.2 Collecting Data from Users – 239
- 9.2.3 Measurement at the Provider (Sell-Side) Level – 243

9.3 Analyzing the Data – 247

- 9.3.1 Transforming Data into Information—Audience Metrics – 247
- 9.3.2 Transforming Information into Knowledge: Qualitative Analysis – 249
- 9.3.3 Quantitative Analysis: “Data Mining” —Overviews of Techniques – 250

9.4 Conclusions and Outlook – 259

- 9.4.1 Case Discussion – 259
- 9.4.2 Challenges in Audience and Market Research – 260
- 9.4.3 Conclusion – 262

9.5 Review Materials – 262

- 9.5.1 Questions for Discussion – 263
- 9.5.2 Quiz – 263

Quiz Answers – 266

9.1 Why Demand Analysis

» “Nobody knows anything”.

William Goldman (Oscar-winning screen-writer for the films *Butch Cassidy and the Sundance Kid*, *All the President’s Men*, *The Stepford Wives*, *A Bridge Too Far*, *The Great Waldo Pepper*, *Marathon Man* and *The Princess Bride*).

Is Goldman, a classic Hollywood insider and guru, correct with his oft-quoted observation that guesswork rules when it comes to understanding and assessing media audiences, their tastes and preferences? In a strict sense, yes. We do not fully know what the users of a new piece of media content (and of media-related technology and platform) want at any given moment, which is why so many media and information projects fail in the marketplace. But perhaps one should define the task more modestly. To be absolutely sure may be impossible, but maybe one can increase the probability a little? To succeed against competitors operating under similar uncertainty, one need not be exactly on target all, or even most of, the time—just a little less wrong than they are. Over time, this leads to a better track record and to success. And this is the subject of this chapter: how media and technology firms can improve the assessment of the demand for their products and services, and thereby become more successful.

9.1.1 The Importance and Special Problems of Demand Estimation for Media Industries

Understanding demand research is always important and always difficult. On the *macro* level of the overall economy, governments and businesses need to know what to expect by way of aggregate national demand, or sectorial demand such as for housing or energy. On the firm-specific *micro* level, every industry and firm wants to know:

- Who are the potential buyers?
- What is the buyer’s willingness to pay?
- What is their price sensitivity?
- What product features are valued?
- What do customers like about competing products?
- How to identify promotional effectiveness.
- How to identify market segments and select target markets.

Market research is particularly important (and difficult) for media and information firms. Recall some of the basic economic characteristics of media, discussed earlier.

High Investment Needs Media content is often expensive to produce and has a short shelf life. A Hollywood film may cost \$100 million in upfront investment in production and marketing cost, yet have an economic life of a few months only. Similarly, infrastructure distribution networks require huge investment far ahead of actual demand.

High Uncertainty In media, an “80–20 rule” often applies, wherein 80% of products do not break even. Of products, 10% account for 90% of the profits, and 2% of products account for 50% of the profits. Such performance does not follow a normal statistical distribution but, rather, an exponential one. The probability of a great demand (a hit) is very low. Conversely, the probability of a low demand is very high. Of course, this will usually be true, but seldom to such an extreme. These are the characteristics of a “winner-takes-all” business.

Preferences Are Unstable Content suppliers must be able to respond rapidly to changing audience tastes. Each discrete media product—such as a film, a book, a song—is unique and hard to evaluate in advance.

Public Good Characteristics Because viewing and sharing cannot be easily accomplished, media products such as broadcast TV or online content are often given away, rather than sold to identifiable users, and the audience’s attention is then sold to advertisers. To monetize such an audience, the media company must be able to identify and quantify it.

Technology Change Digital technology has had a particularly rapid product cycle in recent decades. Consumers often have no experience with new products in advance. As transmission costs drop, media content markets become global and audiences fragment.

Network Effects The product and service preferences of individuals are interdependent on those of others. There is a “network effect.” This leads to extremes of success because users

dynamically influence each other, both by positive encouragement and negative dismissal. When a product or service catches on, it becomes a self-reinforcing process. Conversely, a product that does not generate such positive feedback drops out. Facebook, Twitter, Pinterest and Instagram quickly became popular, but Friendster and MySpace evaporated.

Where the average utility of a product increases as other participants are added, the demand for the product will increase with the number of users. The more people are on the network or share the experience, the more people are willing to pay. This can lead to a highly unusual demand. Whereas, classically, the number of users drops as prices rise, one may now have a situation where, as the number of users rises, people are willing to pay more because the service becomes more valuable to them.

For these and other reasons, demand analysis, while particularly important in the media and information field, is also particularly difficult.

9.1.2 Examples for the Problems in Forecasting Media Demand

The media world is full of false predictions. Statisticians speak of “Type I errors” and “Type II errors.” Type I errors are “false positives”: the wrong decision is taken by a company, which mistakenly accepts a positive but wrong prediction (“hypothesis”)—in this case, that there is sufficient demand for the product, which turns out to be much smaller in actuality. This happens all the time. It has been estimated that startups overvalue the demand for their innovations by a factor of three.¹ Eternal optimism governs media content and technology, and is a force for progress. But it comes with frequent, expensive and, often, fatal flops.

In communications services, when “picture phones” were introduced in 1963 at the New York World’s Fair, the telecom giant AT&T, which had been at the forefront of designing and introducing the technology, estimated that there would be ten million such phones in use in America by 1980. However, the real number in that year was closer to zero. (30 years later, with smartphones, laptops,

tablets, and camera equipped desktop computers, personal video penetration was at the other extreme, at 230 million). Similarly, at one time, mobile phones connected by satellites were widely expected by experts to be the next Big Thing. In 1998, *The Wall Street Journal* gushed that, “the consensus forecast by media analysts is of 30 million satellite phone subscribers by 2006.” The reality, however, was vastly more modest than these experts predicted. Such phones, aside from some subsidized national security applications, are mostly used as rental units on adventure travel, and even that niche has been rapidly contested by alternatives. Other examples are the video devices Tivo and Slingbox. Both were thought of as sure winners, yet struggled to find markets.

Type II errors, in contrast, are “false negatives:” a product should have been picked but was not. The prediction that demand would be low is accepted, but it transpires that this is incorrect and, in actuality, the product is a winner. In 1877, Western Union, the largest telegraph company in the world, believed that there was no market for the new-fangled telephone. It passed on acquiring the patent offered to it by Alexander Graham Bell. Bell had to go it on his own and, within a few years, his company, AT&T, eclipsed Western Union, which then experienced a long and lingering death. A century later, the shoe was on the other foot. AT&T vastly underestimated the prospects of mobile phones, after an expensive McKinsey consulting study commissioned by the company predicted in 1981 that there would be only 900,000 cell phones in use worldwide by the year 2000. AT&T took the advice and left the field to the local phone companies. However, there were more than one billion cell phones by the millennium year. The company had to spend billions of dollars to get back into the business and was too late to be successful. It failed and was acquired for a song by another phone company, SBC, which renamed itself AT&T.

Errors abound by industry insiders who are close to the subject. In 1916, Charlie Chaplin, who went on to an illustrious film career, opined that, “The cinema is little more than a fad. What audiences really want to see is flesh and blood on the stage.”² When TV started to be successful, movie

1 Gourville, John T. “Eager Sellers & Stony Buyers.” *Harvard Business Review* 84, no. 6 (June 2006): 98–106.

2 Tech News and Lifestyle Blog. “Top 50 Failed Technology Predictions of All Time.” Last accessed May 31, 2011. ► <http://data-katalog.com/index.php?newsid=50975>.

mogul Daryl Zanuck, the 20th Century Fox studio chief, was similarly out of touch. “[Television] won’t be able to hold on to any market it captures after the first six months. People will soon get tired of staring at a plywood box every night”. In the computer field, industry experts could be amazingly wrong. “I think there is a world market for maybe five computers,” opined Thomas Watson, Chairman of IBM in 1943. Ten years later, the company was the world’s leading computer manufacturer. A generation later, the president of Digital Equipment Corporation, one of the world’s largest computer manufacturers at the time, predicted that “There is no reason anyone would want a computer in their home.”³ Bill Gates, a founder of Microsoft and the microcomputer industry, may have continued this trend of myopia in observing in 1992 that “640 kilobytes of memory should be enough for anybody.” In 2018, a typical desktop computer has about eight gigabytes, 12,500 times as much.

Consumers regularly opt not to adopt a product even if it is beneficial to them. The approach of behavioral economics has sought explanations. As mentioned in ► Chap. 4 Technology Management in Media and Information Firms, losses have a far greater negative impact on people than similarly sized gains, a phenomenon that behavioral economists have called a “loss aversion.” Consumers overvalue those things they already own over those they do not own by a factor of about three.⁴ This is known as “the Endowment Effect.”

In many situations, both Type I and Type II errors exist, where the producer is over-optimistic and the buyer is over-cautious. In other words, the producer makes the wrong decision to go forward with the product while the buyer makes the wrong decision not to buy it. Thus, when consumers undervalue the existing benefits of an entrenched product by a factor of three while developers overvalue the benefits of their innovation by a factor of three, the result is a mismatch of nine times between what innovators think consumers desire and what consumers actually want—about one order of magnitude.⁵

9.1.3 Limits to Audience and Market Research

In this chapter, we will investigate ways to conduct the estimation of audiences and market demand. However, as we proceed with looking into various techniques, we must also keep asking the questions: when it comes to content rather than products, *should* media companies use demand estimation techniques in the same way as a car manufacturer or an airline? Shouldn’t decisions on media creations by a media company be based on artistic originality, news judgment and public responsibility?

In that vein, a great deal of criticism has been directed toward audience research as a substitute for a creative judgment. Screenwriters and film directors despise audience research that forces them to modify their creation. Editors bristle at the pressure to tailor their front page stories to attract younger demographics. This is not how a business visionary acts, they argue. Steve Jobs did not test market the original Macintosh. Colonel Parker did not Q-test Elvis Presley. Orson Welles did not use a focus group.

Taking one step further, do media owe their audience a special responsibility to go beyond what that audience thinks it wants? Should they have an obligation to cover significant but boring news and difficult but important topics?

More fundamentally, some thinkers believe that the entire exercise of demand estimation by media companies is tautological: media creates its own demand by influencing people and their preferences. They like what they are told to like. There has long been a debate whether peoples’ preferences are shaping the content of media, or, to the contrary, whether media content has been shaping peoples’ preferences. Do “powerful media” or “powerful audiences” determine media content?⁶ Social science and communications research have not resolved this question.

Those who believe that preferences are inherent to the viewer (the “powerful audience”) tend to study it purely empirically, looking at audience behavior, rather than engaging in theories to explain that behavior. George Gallup, the famed

3 Tech News and Lifestyle Blog. “Top 50 Failed Technology Predictions of All Time.” Last accessed May 31, 2011. ► <http://data-katalog.com/index.php?newsid=50975>.

4 Gourville, John T. “Eager Sellers & Stony Buyers.” *Harvard Business Review* 84, no. 6. (June 2006): 98–106.

5 Gourville, John T. “Eager Sellers & Stony Buyers.” *Harvard Business Review* 84, no. 6. (June 2006): 98–106.

6 Livingstone, Sonia M. “The Rise and Fall of Audience Research: An Old Story With a New Ending.” *Journal of Communication* 43, no. 4 (Autumn 1993): 5–12.

pollster, was among the first to research audience preferences in a very practical way. In the academic field, a central figure was Paul Lazarsfeld, who started an institute in audience measurement at Columbia University to study radio listeners.⁷ Lazarsfeld's statistical techniques were adopted by media and by audience measurement companies such as Nielsen. This approach centers on the audience, on audience sub-groups and their behaviors, rather than their motivations.

In such a vein, many, if not most, media firms look at audience preferences and seek to satisfy them in order to be commercially successful. Yet, media firms themselves are self-contradictory. On the one hand, they argue that they only passively serve existing influences, rather than shape them. This could be called the “Nielsen perspective.” Yet, at the same time, they promote themselves to advertisers as being able actively to re-shape consumer preferences, including on their media choices. That could be called the “Madison Avenue” perspective.

Maybe both major perspectives are correct. Media audiences have preferences that can be analyzed as given at a certain moment. This is referred to by the industry as “media research” or “market research.” However, these preferences can also be influenced by means of “media marketing.” This chapter deals with the former; we will deal with the latter in the next chapter.

9.1.4 How Media Companies Organize Their Market Research

To improve on their odds for success, media organizations engage in substantial market research at every step:

- To let the companies know who their audience is, and how it responds to the content provided;
- To let advertisers know who they are reaching;
- To let advertisers know how effective they are.

A great deal of money is at stake. The largest advertisers in the USA in 2016 were Proctor & Gamble with \$4.3 billion, AT&T with \$3.9 billion, General Motors with \$3.5 billion, Comcast with \$3.4 billion

and Verizon with \$2.7 billion.⁸ On the receiving end, 2016 advertising sales by the TV company CBS were about half of its overall revenues, about \$6.1 billion, and a further \$6 billion from licensing fees and affiliate payments that were themselves heavily dependent on other TV companies' advertising sales. For the major newspaper company Tribune Publishing, two-thirds of its \$1.6 billion in revenues were from advertising about \$1 billion. Of Google's (Alphabet) 2018 revenues of \$110 billion, over 90% came from advertising. With such large amounts of money at stake, advertisers, media companies, network platform service providers and technology firms need to understand their markets and audiences.

Over time, large TV firms organized media and audience research in increasingly complex ways, both in-house and with the use of outside measuring services. The major US TV networks have internal research departments of approximately 30 people—about 20 ratings analysts, plus 3–5 people in primary work on strategic and regulatory issues such as violence, children and boycotts, as well as for qualitative studies, including focus groups. These analysts are mostly at company headquarters in New York. The agenda consists of sales research for advertising sales, internal program research, studies for local stations and radio, and work with outside contractors.⁹ About half a dozen analysts are deployed in Los Angeles engaged in program research.

The top audience research executives in the large media firms hold titles such as “Head of Research”, “President—Research”, “Exec VP—Research”, or “Chief Research Officer”. They report either to the Chief Marketing Officer or to the CEO directly. Audience research departments tend to hire at the BA or Masters level, often with a background in statistics. Some hire MBAs. Experience in digital data analysis and in advertising agencies is helpful.

Starting in the mid- 1980s to mid- 2000s, the major TV companies in America added new cable TV channels. Research departments doubled and tripled, even though the major networks themselves shrank. In time, decentralization took place as the subsidiary cable channels created their own

8 Nanji, Ayaz. “The 10 Biggest Advertisers in the United States.” *Marketing Profs*. July 22, 2016. Last accessed July 11, 2017. ► <http://www.marketingprofs.com/charts/2016/30313/the-10-biggest-advertisers-in-the-united-states>.

9 Scott McDonald, interview with author, July 18, 2012.

7 Zitrom, Daniel. *Media and the American Mind*. Chapel Hill: University of North Carolina Press, 1938, 122–146.

research departments within the same company, e.g. the Disney Channel, ABC and ESPN. There was much expansion, including into digital research and to more program research.¹⁰

Local TV stations, depending on market size, also have market researchers, in particular for ratings analysis. Their work is supported by the TV networks or by outside contractors.

Large magazine group publishers may employ several dozen researchers. Plus the individual magazines may have a researcher or two. At Condé Nast, this function is called the “Research and Insights Group.” Its head reports directly to the CEO and is a member of the Executive Committee.

9.1.5 Case Discussion

Viacom—Golden Age Media (a Hypothetical Case)

The large US media company Viacom, in an effort to target the large and growing retirement-age audience, considers launching a new media endeavor: Golden Years Media (GYM). GYM would offer three products: (a) a cable pay-channel (the Golden Years Channel) that is marketed directly to viewers, (b) a magazine (*Golden Years*), and (c) a website (GY portal). The basic question for the company is whether it should launch Golden Age Media. To answer that question, how could Viacom estimate its audiences, their content preferences, their consumption preferences and their willingness to pay?

A Brief Overview of Viacom

Viacom is an American mass-media company. It owns film and TV production (Paramount), runs more than 160 cable networks

and serves over 700 million viewers worldwide in many dozens of countries and languages. Its channels include MTV, Nickelodeon, BET and Comedy Central. Sumner Redstone, aged 95 in 2018, is the controlling shareholder of Viacom, as well as of the TV network CBS and other media firms, through the family business National Amusements, originally a medium-sized movie theater chain.

Viacom owns a large number of satellite-delivered TV channels that are carried by cable and direct broadcast satellite (DBS) TV operators. Ordered by the target audience ages, they include Noggin (pre-schoolers), Nick Jr. (ages 2–5), NickToons, Nickelodeon (twens), TeenNick (ages 12–24), MTV and MTV2 (12+).¹¹ These channels reach up

to 100 million households. Other Viacom channels target college students and young adults (ages 18–49): mtvU (college), Comedy Central, Spike, VH1 and Logo. These channels reach up to 127 million households. Viacom’s channels targeting mature adults are Nick at Nite (50+), TV Land (18–65) and CMT (Country Music TV) (18–65). These channels reach up to 68 million households. Other Viacom channels that target specific demographics include BET (African American), BET HER (African American women), Logo (LGBT) MTV Tr3s (Latino) and Sundance (film fans).

Viacom now plans to target the 65+ age demographics though three interrelated new media products: the Golden Years channel; the *Golden Years* magazine and the GY portal.

9.2 Data Collection

Good research requires a well-defined question, good data, resourceful interpretation and effective use. The topic of the next section is how to get the data. Later, we will discuss how to analyze it.

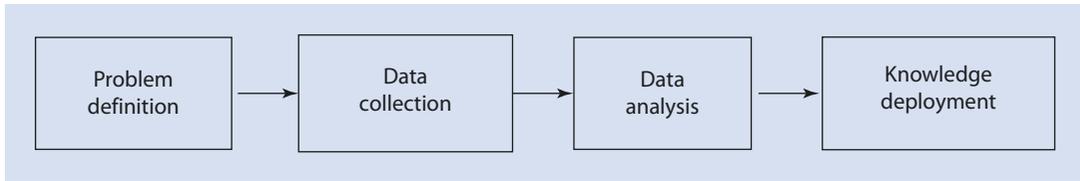
9.2.1 The Impact of Collection Methodology

■ Figure 9.1 shows the stages of data use, from problem definition, to creation and use of knowledge. Data collection is a critical step. The choice

of the methodology and technology of such collection is not an objective science. But such choices affect results and, thus, influence business and creative decisions. Therefore, it is always a battlefield. One main purpose of audience measurement is to be a tool for attracting and charging advertisers and others. Any change in a metering procedure will therefore impact business. In TV, the replacement of a system of paper diaries with an automated system made a significant difference to the results. The People Meter

¹⁰ Scott McDonald, interview with author, July 18, 2012.

¹¹ Viacom. “Viacom Brands.” Last accessed July 13, 2017. ► <http://www.viacom.com/brands/pages/default.aspx>.



■ Fig. 9.1 Collection Methodology

measurement system showed ratings for the major TV networks that were consistently lower by about 4.5 points than the combined TV diary ratings. In contrast, cable TV had a substantial ratings gain. There were also effects on different programming categories. Participation shows were boosted by 5 points in rating, sitcoms by 1.5, and news by 0.2. But medical shows dropped by 4.1 points. All this has revenue impacts. Each ratings point was worth approximately \$140 million per year. The networks' decrease in ratings by 4.5 points could therefore cost them over \$600 million per year. Thus, one can see that ratings technology and ratings methodology affect dollars, euros and yen.

9.2.2 Collecting Data from Users

Broadly speaking, data is primarily collected from one of the two sides of transactions, either from the *buy side* (users, consumers, audiences), or from the *sell side* (providers, producers, distributors). There are several ways to collect data from users.

9.2.2.1 Surveys

Personal surveys are usually conducted by market research firms. Personal interviews can be in-depth, but they are also expensive and need a reliable survey team. The sample is often biased by self-selection of the subjects who agree to participate, and by the accessibility of the subjects. Follow-up research is time-consuming. In some cases, an interview subject is shown aids to memory (such as past copies of a publication, or of an advert), followed by questions. This is known as the “recognition” method. When no such aids are given, it is referred to as “a recall method” interview.¹²

A major problem with personal surveys is the truthfulness of responses. People will often dissemble about their incomes, taste and actual consumption patterns. Or, they can be forgetful. There is also an “interviewer effect” where, for example, the age, gender, attractiveness, or status of the person conducting the survey may have an impact on responses.¹³ In short, personal interviews are time-consuming, expensive and limited in accuracy.

Mail and Phone Surveys

Mail surveys cost less than in-person surveys and their greater anonymity increases candor. However, the low response rates to “junk mail” mean that, on a “per responder” basis, this is not a cheap method after all.

A type of survey, used for decades by magazines, is often deployed by the Internet industry. It is known as “A/B testing.” Example: By comparing the acceptance rates of test marketing that offer two alternative prices, with the other variables held constant, one may find the responsiveness to price. Suppose that, for a listed price of \$17, acceptance was 10.2% lower than when the price was listed at \$15.¹⁴ This is a price reduction of 11.8%. The “price elasticity” (defined as the percentage change in demand divided by the percentage change in price) is, hence, $\eta = -10.2/11.8 = -0.86$. This means that the price sensitivity is slightly inelastic (less than 1).¹⁵ Where demand is inelastic, a higher price point (such as going to \$18 or higher) would raise revenues.

Telephone surveys are cheap, and allow follow-up questions and clarifications. Here, too, there is a self-selection bias. There are legal restrictions on unsolicited phone surveys when they are coupled to a sales pitch.

12 Kim, Hyo Gyoo. “Traditional Media Audience Measurement: Print and Broadcast Media.” 2006. Last accessed July 13, 2017. ► <http://www.columbia.edu/B8210/read24/suggested/Audience3.htm>.

13 Frankel, Martin R., Marc B. Glassman, and Eleanor Singer. “The Effect of Interviewer Characteristics and Expectations on Response.” *Oxford Journal Volume 47*, no.1 (Spring 1983): 68–83.

14 Kobak, James B. *How to Start a Magazine*. New York: M. Evans & Company, 2002.

15 Disregarding the negative sign.

Internet Surveys

The Internet provides convenient and low-cost ways to ascertain consumer views. To recruit respondents, a website may seek volunteers (for a one-time response, or to serve as part of an ongoing panel). It may offer a reward, such as access to otherwise restricted or for-pay content (a “survey-wall”), or a chance to win a reward in a lottery. They may solicit general user feedback such as “likes,” or create polls. Companies seek respondents through advertising or posting on blogs, websites and social media. They can send e-mails with survey instruments to various mailing lists that can be self-created or rented. They can present each online visitor with a survey, or do so selectively by picking a sample, either randomly or according to certain parameters. A number of companies provide relevant software and services, such as SurveyMonkey and Google Surveys. The software permits customizations of various kinds, as well as data analytics of the responses. Disadvantages are that self-selected respondents may be a statistically biased sample in terms of demographics, and that strategic manipulation by multiple responses from a small group are hard to control.

9.2.2.2 Focus Groups and Test Marketing

Market researchers often use group-wide approaches. They may include interactive “focus groups,” or larger-scale product tests.

Focus groups are often used for film in advance of a wide release. There are two types of such group testing: for production, and for marketing. Production previews help filmmakers fine-tune the movie while it is being made, whereas marketing previews study an audience’s reactions to complete films and assess marketing strategy.¹⁶ A typical focus group gathers about a dozen people (more would be unwieldy for interaction) to discuss the product for, maybe, two hours, talking with each other and to a professional moderator. To benefit from the focus group, marketing managers often listen in from the outside, or peek in through one-way mirrors or video screens. Focus groups can be used to help fine-tune a work in progress. Their response provides feedback to producers, creators, marketers and executives, but

is no substitute for artistic judgment. For example, focus groups hated the pilot to the TV series *Seinfeld* which was green-lit anyway and then became one of TV’s most successful shows.

Focus group testing is similarly done by games developers or publishers. A group of the target demographics, typically young men or adolescents, are brought together. They play the game for a while and are then prompted to discuss their experience and observations. This could include whether the game is too easy or too hard, the visuals, the level of excitement, the characters and so on.

Focus groups are highly unpopular with screenwriters and directors. They are held to be responsible (or used as a smoke screen) for media executives watering down artistically exciting approaches. For people with a statistical bent, the focus group approach is laughably imprecise with its sample size, bias and randomness.

A broader outreach uses *test audiences*. Such test screening is done in a theater, typically in New York or Los Angeles. Sometimes, test audiences are asked to fill out a survey questionnaire about what they liked and did not like about the film. This includes elements such as actors, characters, special effect, plot and pacing.¹⁷ Other test screenings use “audience perception analyzers,” which are systems with little hand-held clickers similar to TV remote controls¹⁸ that send out, record and analyze responses, and their intensity, in real time.

Many popular movies have been altered after being shown to test audiences or focus groups. Television producers and directors fear that shows that test poorly will be dropped by networks, and thus forgo a major content development advantage over film. In contrast to film, a series can improve with constant refinements as it goes along. But this potential is lost when cancellations are based on only one or two episodes, and the show has no chance to improve itself and build a following.¹⁹ The opposite criticism is also made, that testing a film serves as a shield to protect managers’ careers if a film opens badly. Then, the marketing executives can point to the earlier poor

16 Friedman, Robert, and Jason Squire. *The Movie Business Book*, 3rd ed. Maidenhead, UK: Open University Press, 2006, 282–298.

17 Marich, Robert. *Marketing to Moviegoers: A Handbook of Strategies Used by Major Studios and Independents*. Burlington, MA: Elsevier Focal, 2005.

18 Conferex. “Excellence in Presentation.” 2008. Last accessed June 2, 2011. ► <http://www.conferex.co.uk/index-3.html>.

19 Albarran, Alan, and Sylvia Chan-Olmsted. *Handbook of Media Management and Economics*. Mahwah, NJ: L. Erlbaum Associates, 2006, 629.

test results and claim that they produced the best release campaign possible for a weak product.²⁰

9.2.2.3 Demand Experiments

One can get data on audiences and buyers by observing their actual behavior or by experiments.

In an *uncontrolled study*, researchers are only observers. Uncontrolled research often uses behavioral data from samples from panels of consumers. Consumers can keep a diary of their activity and purchases, or these are recorded automatically. Customers may get a reward for cooperating.

In contrast, in *controlled studies* researchers can manipulate the important variables to observe their effect. In controlled studies of actual purchases, the researchers may generate price variations while holding other variables constant, such as advertising. This can be useful but takes time and money.²¹

Magazine test marketing serves as a good example of controlled purchase experimentation. Magazine firms may utilize a “dry test,” where product acceptance is tested without it actually yet being published. On the Internet, such experiments have become much easier. If a website wants to find out whether a new design of a webpage increases sales, it can quickly run a controlled experiment. It will show the new page design to, say, every hundredth visitor. Determination of whether the new design increases sales can be made after a few days, which allows for numbers to build and to correct for unusual weather, holidays, etc.²²

Test marketing means launching the media product, e.g. a TV show or a film, with full marketing and advertising efforts in several test cities or regions. The consumer response is then tracked. Such test marketing for films enables decisions about marketing strategies, improvements, fine-tuning or discontinuation. The problem with test marketing is that it is slow. It also leads to a premature exposure of the product to competitors. With the Internet, information spreads rapidly and test audiences are much less isolated than in the past.

In-store purchase experiments can be costly and run into millions of dollars. The cost is high because each additional factor studied requires the use of more stores in order to obtain statistically valid results. When Quaker Oats conducted an in-store experiment that focused on the effect of price alone, the study required the use of 120 stores for three months.²³ For technology products such as TV sets or computers, charging lower prices for experimental purposes can become quite expensive. This, then, leads to the use of laboratory experiments.

A *laboratory experiment* research facility can be set up at a shopping mall or other high-traffic location. It resembles a small store to provide the realism of in-store trials without their high cost and exposure to competitors. Participants and prices are controlled. Consumers may be rewarded by a substantial discount, or other benefits. The overall cost is smaller than for in-store testing and is therefore more popular with electronics products.²⁴ Other laboratory experiments may provide users with play money, observe how they engage with content and with on-shelf retail content such as magazine covers, subject them to advertising messages, followed by product choice decisions, and monitor the impact of the advert. They may also simulate a living room setting to observe user behavior. Such lab in-home simulation aims to provide insight into the typical interactions of consumers with devices and programming in a home-like setting. Research questions may be: How are programming choices made? When and how are available devices used? What else is done while watching? What interactions does a message or content trigger? How can a program stand out from the rest?

9.2.2.4 Automatic Audience Metering

The individualized or group surveys and experiments are usually too slow, costly and unreliable for the mass production of data. In consequence, large-scale data collection was an early goal. TV and radio broadcasting were the first to use automatic metering. In the USA, the rating service company Nielsen became the main authority for TV ratings. Nielsen has over 200 designated TV direct market areas (DMAs), which it “sweeps”

20 Marich, Robert. *Marketing to Moviegoers: A Handbook of Strategies Used by Major Studios and Independents*. Burlington, MA: Elsevier Focal Press, 2005, 27.

21 Holden, Reed, and Thomas Nagle. *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*, 3rd ed. New Jersey: Prentice Hall, 2001.

22 Varian, Hal R. “Kaizen, That Continuous Improvement Strategy, Finds Its Ideal Environment.” *New York Times*. February 8, 2007. ► <http://www.nytimes.com/2007/02/08/business/08scene.html>.

23 Holden, Reed, and Thomas Nagle. *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*, 3rd ed. New Jersey: Prentice Hall, 2001.

24 Holden, Reed, and Thomas Nagle. *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*, 3rd ed. New Jersey: Prentice Hall, 2001.

four times a year. Similar audience measurement companies operate all over the world wherever advertising supported TV or radio is in place.²⁵

Early TV audiences were measured using, primarily, a “diary system.” The viewers recorded their TV viewing in a paper diary—an awkward system for channel-hopping viewers. A second way in which TV audiences were tracked was by telephone surveys. This enabled fast turnaround, but this sample, too, can be biased. These problems encouraged the development of automatic systems for audience monitoring. Nielsen’s People Meter, rolled out in 1987, is an electronic box placed in 5000 randomly selected household and positioned near their TV sets. It measured households viewing instantly and eliminated “lying” responses. Drawbacks are that often there is nobody in the room actually watching. Viewers must identify themselves so as to differentiate between different members of a household, and this requires active cooperation. A solution to this problem is a “passive” meter that records the behavior of the respondents without requiring any effort by the respondent.²⁶ One such method was introduced by Arbitron, (a company subsequently acquired by Nielsen). It created a digital “watermark” (identifier) embedded in recorded music and in the music of advertisements played, and these codes are received and identified by monitors in each market. Analytic software then identifies the time and station, as well as the advertising or music content.²⁷ This system with its content identified through embedded codes, can be used for broadcast TV, satellite and cable, as well as for advertisements, and follows the person wherever they are. Such a “Passive People Meter” (PPM) requires a device worn by the consumer, which then detects and records the content the consumer is exposed to, whatever the program source.²⁸

A problem with the accuracy of TV data collection is that greater audience fragmentation makes results less reliable. The percentage of standard deviation tends to grow as the measured percentages of “ratings” and “shares” become

smaller. For example, a “true” ratings of 6 (6% of TV households) in a sample of 3000 will show as a sample rating between 5.2 and 6.8 (± 0.8) in 95% of the samples taken. The relative error is thus $0.8/6$, or $\pm 14\%$. But the same error at the same 95% confidence level for a “true” rating of only 2 will be ± 0.5 . (This means a relative error of $\pm 25\%$). And for a small cable channel with a “true” rating of $.3(\pm 0.2)$, the relative error is $\pm 65\%$. This means that the samples need to become larger as audiences fragment, and this adds cost and potentially slows down the system.

9.2.2.5 The Measurement of Websites: User-Level Measurements of Internet Activities

The Internet is an excellent survey tool. When it comes to the measurement of Internet usage itself such as of website visitors, here, too, there are two basic ways of how to proceed:

1. To measure user behavior (the “user-centric” approach);
2. To measure websites (the “site-centric” approach);
 - 2.1 A subcategory of website measurement is the “ad-centric” approach, in which clicks on advertisements are counted.

User-level measurement of online activity is an approach based on a sampling of techniques. It uses a large panel of general Internet users, observing their behavior and visits to websites. The users install a software meter on their computers that measures and recalls their behavior. The meter reads the URL in the browser, then counts and forwards data to the web rating company. Data is then matched to websites and reported.²⁹ To automate the process of tracking Internet traffic on the user level, “cookies” or similar tools are used. Cookies are electronic files to tag individual users with unique identifications. It allows websites to recognize individuals.³⁰

comScore (which acquired rival Media Matrix in 2002) claims a panel of two million worldwide Internet users, including one million in the USA who opt in and share their clickstream data on

25 European Audiovisual Observatory. “Sites of the TV Audience Measurement Companies.” August 2001. Last accessed June 1, 2011. ► <http://www.obs.coe.int/db/gavis/00002194.html>.

26 Albarran, Alan, and Sylvia Chan-Olmsted. *Handbook of Media Management and Economics*. Mahwah, NJ: L. Erlbaum Associates, 2006, 632.

27 Baskerville, David. *Music Business Handbook and Career Guide*, 8th ed. Thousand Oaks, CA: Sage Publications, 2005.

28 Arbitron. “The Portable People Meter System.” Last accessed July 11, 2017. ► <http://www.arbitron.com/downloads/esomar2000.pdf>.

29 Coffey, Steve. “Internet Audience Measurement: A Practitioner’s View.” *Journal of Interactive Advertising* 1, no. 2 (Spring 2001): 13.

30 Deck, Cary A. et al. “Tracking Customer Search to Price Discriminate.” *UArk.edu*. June 2003. Last accessed July 11, 2017. ► <http://comp.uark.edu/~cdeck/Tracking%20Customer%20Search%20to%20Price%20Discriminate.pdf>.

webpages visited.³¹ Nielsen's NetRatings uses a group of about 2.5 million in the USA which form part of an opt-in panel of people who give information about themselves such as age, gender and income.³² In time, over 100 web ratings companies emerged; the major firms are Google Doubleclick, Nielsen, Experian, Adobe Analytics, Akamai Real Time Web Monitor, Quantcast, comScore, Hitwise and Amazon's Alexa.

There are several advantages to a user-level approach. It provides demographics and measures actual behavior. There is no conflict of interest in which the website inflates its count in order to raise its advertising revenues. (In some cases, however, the user measurement company also operates large websites, as Google does for example, and provides placement services to advertisers). The user-level approach provides additional tools for analysis. Mouse activity can be measured through a monitoring of the number of clicks, time spent moving the mouse,³³ or total time spent on a webpage. On the other hand, the user-based methodology requires user cooperation by signing up. Hence, incentives must be offered to users.³⁴ The sample thus has some biases.

Another problem of user-centric measurement is its disadvantage to small sites which may get only a few hits and may thus be ignored or undercounted. It also provides poor site diagnostics on which site a user reaches and what the user does there.³⁵ Also, user-centric systems are not welcome at workplace computers. Most large companies have firewalls and security systems, and measurement software cannot be installed. And, as with all panel/sample-based measurements, certain locations and demographics are under-represented. Therefore, the measuring companies often recruit panelists to "balance" panels, and they may also correct for remaining biases through post-sample weighting, i.e. retreating the results. This invites problems.

9.2.3 Measurement at the Provider (Sell-Side) Level

So far, we have discussed the ways in which audiences can be measured on the level of the actual *users*. The other major approach is to focus on the *seller/provider side* and to measure there. This is done in various ways.

9.2.3.1 Surveying Retailers

Bestseller Lists Book bestseller lists are tabulated by newspapers, magazines, or other organizations from actual sales. They are compiled from reports by hundreds of book stores, with the identity and weight given to each store not being disclosed. The system is basically an extensive sampling of retailers. The problem with such lists is that they are self-fulfilling. They determine the location of a book's display inside a bookstore, which has a substantial effect on the book's sales. The list also determines whether or not the book will be discounted in price.

Because it does not take all that many sales to make it to a bestseller list, they have been subject to manipulation. Publishers or authors can "pad" the list by buying their own books in bulk from stores around the United States to get their sales figures up for the lists. The authors of the book *The Discipline of Market Leaders* business consultants Michael Treacy and Fred Wiersema, reportedly spent \$250,000 of their personal money to buy 10,000 copies of their own book. This propelled the book into the bestseller list and to subsequent sales of over 250,000 copies, easily recouping their money. *The New York Times* now places a dagger next to any titles when substantial bulk sales are being reported at individual stores.

9.2.3.2 Self-Reporting

Producer self-reporting is mainly used by newspapers and magazines, and is also part of Internet "site-centric" measurements. This system of central audit bureaus of circulation (ABCs) exists in many countries to audit and verify newspaper and magazine circulation. Before ABCs, publishers exaggerated sales to advertisers. Controversies ensued. Advertisers and advertising agencies then created an institutional method to verify circulation claims. The way the system works is that a publication such as a newspaper or magazine reports its circulation and sales to a central unit, the ABC. Typically, twice a year each magazine and newspaper submits a

31 Delo, Cotton. "Your Guide to Who Measures What in the Online Space." *Advertising Age* 82 (September 18, 2011): 1.

32 Innovateus. "What Is the Nielsen Rating System?" Last accessed July 11, 2017. ► <http://www.innovateus.net/innopedia/what-nielsen-rating-system>; Delo, Cotton. "Your Guide to Who Measures What in the Online Space." *Advertising Age* 82 (September 18, 2011): 1.

33 Brown, David. "Inferring User Interest." *IEEE Internet Computing* 5, no. 6 (November 2001): 35.

34 Cooley, Robert. "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data." *SIGKIDD Explorations* 18, no. 2 (January 2002): 13.

35 McDonald, Scott and James Collins. "Internet Site Measurement Developments and Print." Presented at *Worldwide Readership Research Symposium 2007*, Vienna, October 2007.

statement of its circulation, (a “Publisher’s Statement”). The ABC is responsible for auditing to verify the numbers and keep everyone honest. An ABC board typically consists of advertisers and advertising agency representatives, as well as newspaper and magazine representatives.

Even with audits, there have been regular instances of padding of circulation numbers. In

2004, Belo Corp., owner of the *Dallas Morning News*, as well as of other newspapers and TV stations, was investigated. It was found that the company had falsely reported numbers by counting, for example, unsold papers. This overestimated weekday circulation by 5.1% and Sunday circulation by 11.9%. The company had to refund advertisers \$23 million.

9.2.3.3 Case Discussion

Viacom Golden Years—Self-Reporting

Magazines self-report their circulation, subject to auditing of those numbers. Suppose that the *Golden Years* magazine had a verified test run in the New York State market for a three-month period (■ Table 9.1), accompanied by a regional promotional campaign.

Extrapolating nationally from New York State, the number of

subscribers in the USA would be about 133,000. Such an overall circulation, after strong promotion, is relatively low considering that *AARP the Magazine* has a circulation of 22.5 million.

Another observation is that, in Month 1, the rate of increase of subscribers was rapid, probably because the new product was being

promoted. It grew from 0 to 7000. In Month 2, subscribers increased to 8000, a much slower rate of growth. In Month 3, the increase was only 100. Circulation plateaued, and would likely even decline with the decline of magazines generally, although this decline would be somewhat offset by the rise in the number of retirees.

9.2.3.4 Measuring Film Box Office Sales

The measurement companies in the USA for film audiences are Exhibitor Relations Co. (ERC) and, in particular, Rentrak.³⁶ They record film ticket data by collecting box office attendance figures from theaters (i.e. the retailers) and issuing weekly reports to the news media. To make sure theaters are not misreporting the number of tickets sold, undercover checkers may be used who buy numbered tickets at the first and last shows at randomly selected theaters. The main criticism of this movie audience reporting methodology is that it is inaccurate. Anne Thompson, editor of *Premiere Magazine*, dismissed the numbers as “made up—fabricated—every week.”³⁷ For example, distributors tend to exaggerate the popularity of 3-D films (whose tickets are pricier.) Sunday figures are extrapolated from the Friday–Saturday figures, based on experience.

■ Table 9.1 Golden Years Magazine Circulation Reports (Schematic)

Golden Years magazine Circulation January–June, NEW YORK (3 issues)	Month 1	Month 2	Month 3
<i>Paid</i>	6,000	7,400	7,500
<i>Complimentary</i>	1,000	1,100	1,100
<i>Single copy sales</i>	0	0	0
<i>Total paid and verified circulation</i>	7,000	8,500	8,600

9.2.3.5 Point-of-Sale Measurement

For music, one must distinguish between different approaches to measurement. The first is to measure radio audiences and airplay. This has been discussed in the preceding section on user-level measurement. The second is the

36 Other major Hollywood market research firms are Marketcast (owned by Shamrock Capital). The French owned MPG closed down in 2014.

37 Shaw, David. “Tinseltown Spins Yarns, Media Take Bait.” *Los Angeles Times*. February 12, 2001. Last accessed July 11, 2017. ▶ <http://articles.latimes.com/2001/feb/12/news/mn-24444>.

provider-level measure of sales and of digital uploads.

In the old days of the music sales system, *Billboard Magazine* (or its equivalent in other countries) contacted a sample of selected retailers to create bestseller lists. Reporting was often inaccurate, merely rank-ordered rather than with full numbers, and susceptible to manipulation and favoritism. A vast improvement came about through the point-of-sale (POS) SoundScan System. Developed by Sound Data in 1987 and used by *Billboard* for its charts, it is a computerized data collection system with barcode scanning done by retailers. SoundScan claims to measure 85% of all music sales in the USA. Point-of-sale purchases are tabulated from over 14,000 US retail outlets, as well as mass merchants and non-traditional distributors such as online stores and venues.³⁸ It is also used by performing rights organizations so as to track royalties.

Nielsen acquired SoundScan in 1991. Nielsen also provides the related POS sales services BookScan and VideoScan. Video DVD sales, however, remain a more difficult number to do well. Distributors usually hype a film's initial DVD sales, but do not release periodic sales information thereafter.³⁹ Yet, such DVD sales information is important to actors, directors and writers for royalties and profit information. In consequence, talent agencies and management firms created research teams to check on DVD revenues. There are also specialized companies that work on DVD sales, such as Screen Digest.

9.2.3.6 Tracking Downloads and Uploads

Another way of measuring the sale of music is to measure uploads of music to users. In 2004, *Billboard* magazine introduced the “Billboard Buzz Top 25”, based not on sales or airplay but, rather, on the extent artists were being discussed by a selected sample of users.

Music download purchases from iTunes are known to Apple, of course, but are not reported

by the company outside of a “most popular” iTunes chart. SoundScan tracks online sales as well as online streams of music. Nielsen partnered with major online music sellers and streamers including Apple, Spotify, Beats Music, Google Play and Xbox Music to count songs downloaded and streamed online. In aggregating overall sales, online album sales are counted as a full sale, just as a physical CD or vinyl record. When a single track is purchased, it counts as 1/10 of an album sale. For streaming (i.e. from one of the services above), 1500 tracks streamed from the same album counts as one album sale.⁴⁰ Nielsen tracks video streamed from Netflix and other online video platforms by using the audio watermarks embedded in the videos. These are received by the Nielsen People Meters located in the sample households.

YouTube displays a counter on the page which counts the number of times a video is played. YouTube uses these views to put together “What’s hot” pages showing clips with large interest as well as weekly, monthly, yearly and all-time lists of the most viewed videos. Being highly popular creates prestige and bragging rights. Google also pays highly watched content providers a share of the advertising revenues associated with that content. Thus, for reasons of both ego and money, content providers have an incentive to drive up the visitor count. This can be done by promoting and pushing content, and by presenting the content and its links with sensationalist headlines or pictures to attract click-throughs and to encourage forwarding of the material over social media. This is known as “clickbait.” Websites can also buy views. Content providers on YouTube inflate their views by buying them from service providers such as YView.⁴¹ The individual seeking clicks selects the service requested and its quantity. This can include “likes” or (even more problematically) “dislikes” (\$5.00 for 100 “likes”), comments (\$15.00 for 100 comments), subscribers (\$20.00 for 100 subscribers) and views (\$2.00 for 1000 views). Some sites use bots to browse the selected websites.

38 International Standard Recording Code. “What is SoundScan?” 2010. Last accessed July 11, 2017. ► <https://www.isrc.net/FAQ-Soundscan.php>.

39 Horn, John. “DVD Sales Figures Turn Every Film into a Mystery.” *Los Angeles Times*. April 17, 2005. Last accessed July 11, 2017. ► <http://articles.latimes.com/2005/apr/17/entertainment/et-dvdmoney17>.

40 MacNN. “Billboard, Nielsen tipped to track streaming music, video services.” November 20, 2014. Last accessed July 13, 2017. ► <http://www.electronista.com/articles/14/11/20/billboard.introducing.streaming.plays.digital.sales.into.album.chart/>.

41 Hoffberger, Chase. “I bought myself 60,000 YouTube views for Christmas.” *The Daily Dot*. January 3, 2013. Last accessed July 11, 2017. ► <http://www.dailydot.com/entertainment/how-to-buy-youtube-views/>.

9.2.3.7 Measuring Internet Website Traffic

Site-level measurement is basically a self-reporting system by the website (or by a hired service provider) about its visitor count. This method relies on software that records every time a page is requested.⁴² Site-level measurement has systematic measurement biases. These are the major problems of over-inclusion and under-inclusion. It over-counts because it repeats returning visitors. Users may be viewing from several devices and locations, or clear their browser caches, and would then appear as new users.⁴³ Conversely, the system under-counts cached pages and cannot distinguish multiple users on the same computer.

Websites try to get information about their visitors in several ways:

1. One way is often referred to as “packet sniffing.” It directly extracts usage data from the packets addresses.⁴⁴
2. Another way is setting registration requirements, with users identifying themselves; however, this does not work well because it requires an effort by users.
3. Using “cookies.” As mentioned, these are electronic tags of individual customers with a unique identification. Essentially, it allows a website to recognize an individual.⁴⁵ Many users refuse or delete cookies. Such users get counted as “uniques” by a server-centric system when, in fact, they are repeat visitors.⁴⁶

Site-centric measurement methods are preferred by website owners, in particular because they usually report a higher number of visits to the website.

Also, even small websites with a limited visitor count can show their traffic and its composition to advertisers. But there are other problems

which turn advertisers off. We already described the ways in which “clickbait” is provided and clicks are purchased from service providers who will, for hefty fees, drive up traffic. Other problems are:

- Often, measurements are made by the websites themselves.
- Even if third party companies handle the measurements, they are often biased in favor of their client.
- Demographic information is often poor.
- “Bots” are counted. Although there are filters that cut automated page requests, these filters can be easily defeated
- Websites try to attract traffic through various techniques of “search engine optimization” (SEO), but traffic gained in such a way is often fickle and uninterested.

Major companies in the USA that measure Internet audiences are Nielsen/Net Ratings, Media Metrix, Google DoubleClick, Adobe Omniture and Quantcast.

9.2.3.8 Advertising-Level Measurement

Advert-level measurement is a sub-category of site-level measurement. It counts the number of visitors to an advertisement. Click-through (CTR) software measures whether a user clicked on a particular advertisement and links it to its sponsor. This helps advertisers to measure the actual effect of an advert on many in a way that is unavailable to most TV and print adverts. It creates a payment structure based on success. With a high value placed on a user’s clicking on an advert, abuse was inevitable. Robots were used to create “hits,” and when they were filtered out, the websites that sought to inflate their click rate hired people from low-waged countries to create phony clicks all day.

9.2.3.9 Hybrid Web Measurement

The difference between third-party services (user-centric) and a site’s own server count (server-centric) is not only methodological, but also has concrete economic results. User-centric numbers were smaller and websites were therefore under-compensated by advertisers. This led to hybrid approaches by several companies.

42 McDonald, Scott, and James Collins. “Internet Site Measurement Developments and Print.” Presented at *Worldwide Readership Research Symposium 2007*, Vienna, October 2007.

43 Thielman, Sam. “Watching the TV Watchers: Companies jostle to measure network, online audiences.” *Variety*, January 12, 2001. Last accessed July 11, 2017. ► <http://variety.com/2011/digital/news/watching-the-tv-watchers-1118030108/>.

44 Nirsoft. “Description” 2011. Last accessed July 11, 2017. ► <http://www.nirsoft.net/utils/cports.html>.

45 Deck, Cary A., and Bart Wilson. “Tracking Customer Search to Price Discriminate.” *Electronic Inquiry* 44, no. 2 (April 2006): 280–295.

46 McDonald, Scott, and James Collins. “Internet Site Measurement Developments and Print.” Presented at *Worldwide Readership Research Symposium 2007*, Vienna, October 2007.

9.3 · Analyzing the Data

In 2009, to compromise between the approaches favored by websites and advertisers, comScore introduced a dual system. It added direct (site-centric) measurements which enables websites to count and register traffic directly. Such a combined measurement takes into account actual visits plus extrapolations based on the panelists' behavior. The two data streams are then triangulated using an algorithm to produce an estimate.

9.3 Analyzing the Data

We have looked at how to measure and collect data. The next step, using the flow chart of  Fig. 9.1, is how to organize and interpret it. The first step is to organize the raw data in a manageable way, to transform it into useful information. This is the basic approach of what is generically known as “business intelligence.” It was—and often still is—the major way to deploy data. A second and more ambitious step is to take this organized information and transform it into “knowledge” by subjecting it to analysis. One way to do so is in a *qualitative* fashion, by processing the information as people and organizations typically do, through judgment and experiences. This is “expert-based information processing.” The alternative path is a statistically-based *quantitative* approach of analyzing the data.

We now deal with the first element: the organization of data into information.

9.3.1 Transforming Data into Information—Audience Metrics

Organizing raw data in the media field results in audience metrics. There are several such measures developed, in particular for broadcasting.

9.3.1.1 Broadcast Metrics

Ratings, Shares, and Gross Ratings Points: The rating of a program is its share of the total audience. For example, a TV rating = viewers of a program ($\times 100$) \div TV households. In the United States, there are around 110 million TV households. If 22

million households watch the show *Grey's Anatomy*:

$$\text{Rating} = \frac{22 \times 100}{110} = 20$$

The next metric, “share” is also a market share, but only as applied to the consumers at that given moment. “Share” is the percentage of TV sets in use (or persons viewing) tuned to a program at its time slot, also known as HUT (Households using TV).

$$\text{Share} = \frac{\text{Viewers} \times 100}{\text{HUT}}$$

For example, if 66 million households watch any TV during the *Grey's Anatomy* time slot (= HUT). Then the share = 22 million HH $\times 100/66$ million HH (HUT) = 33.3.

The share is greater than the rating, since it is the percentage of actual watching households rather than of all potential ones. However, that also means that a program aired when hardly anybody is watching TV, such as at 5:00 am on a Sunday morning, may have a tiny audience (and rating) yet a high share.

Reach (or CUME): The CUME or “reach” measures the number of viewers, listeners, readers, or visitors who use a specific media product at least once per week.⁴⁷

Average Quarter Hour Audience (AQH): AQH is the average number of persons who listen to (or watch) a station or site for at least five minutes during a week. Example: Suppose that Radio Station #1 has a CUME of 20,000, which is high for the market, and an AQH of 150, which is low in that market. Interpretation: The station attracts large number of people but does not keep them and therefore has few listeners at any given time. The station seems to promote itself well but does not provide good programming to keep all listeners tuned in.

Average Frequency (AF) of Exposure: The AF is a calculation that uses some of the measure above. It shows how many times an advertisement must be played or shown so that the average user will be

⁴⁷ Search Microservices. “What is reach? - Definition from Whats.com.” Last accessed July 11, 2017. <http://searchsoa.techtarget.com/definition/reach>.

exposed to it a certain number of times, as desired by the advertiser's marketing plan. This frequency is the number of advertising exposures divided by the overall cumulative audience. The number of exposures is the number of advertisements times the average audience:

$$AF = \frac{\text{number of spots per week} \times \text{AQH}}{\text{CUME}}$$

To meet the target AF, one must have the Number of advertisement per week = $(AF \times \text{CUME})/\text{AQH}$.

For example, assume Radio Station #1 has an AQH of 150 and a CUME of 20,000. To obtain a desired AF of 3, the advertiser calculates the number of spots per week using the above equation: $(3 \times 20,000)/150 = (60,000/150) = 400$. The result means that it needs 400 advertising spots per week to reach the average listener three times. On the other hand, Radio Station #2 has an AQH of 2500 and a CUME of 10,000. To obtain an AF of 3, the number of spots per week is: $(3 \times 10,000)/2500 = 30,000/2500 = 12$. This indicates that, on Radio Station #2, one only needs 12 advertising spots per week to reach the station's average listener three times. This will be much cheaper because it is more targeted. However, Radio Station #1 will reach twice as many people overall because it has a higher CUME.

Number of Impressions: This is the number of times an advertisement is viewed, listened to, or displayed.

There are economic measures associated with audience metrics:

Cost per 1000 Impressions (CPM): Cost associated with 1000 views of content, video, audio, and so on.⁴⁸

9.3.1.2 Website Metrics

Websites count "users" or "unique visitors." Unique visitors are based on cookies that identify a user's computer. For online TV, the basic metric is "streams," which means the number of times the video is accessed during a time period (Nielsen uses +3 days from the live initial showing).

For online media, some of the audience metrics are similar to those of TV and radio. Others make more use of the capabilities of interactive media where users can be identified and counted:

- *Clicks.*
- *Click-through Rate (CTR):* Number of clicks divided by number of views of an advert or post.⁴⁹
- *Advert Views:* those may be substantially different from content views;
- *Total Reach:* Includes direct users and those shared by them with friends or as the result of clicking on an advert.
- *Organic Reach:* Total number of people who saw a post in a newsfeed or otherwise, but not as the result of an advert.
- *People Talking About This:* The total number of people generating stories during the past week about the post or content, which includes comments, likes, shares, offer claims, mentions and so on. This kind of metric measures the "buzz"—the number of people talking or writing about something (but not the number of people exposed to these messages).⁵⁰
- *Likes:* Facebook lets its users express their sentiment through "like," "love," "sad," or "angry." However, the "likes" of a company page do not represent how many people actually see the content of the page (number of views).⁵¹
- *# of Retweets:* Number of times an advert/Tweet has been "reshared" by another user.⁵²
- *# of Mentions:* Number of times a second user uses an earlier Twitter handle in one of his posts.
- *Link Shares:* Number of times a URL has been shared.
- *Follower Growth:* Growth of follower base.

49 Google. "Clickthrough rate (CTR): Definition." Last accessed July 13, 2017. ► <https://support.google.com/adwords/answer/2615875?hl=en>

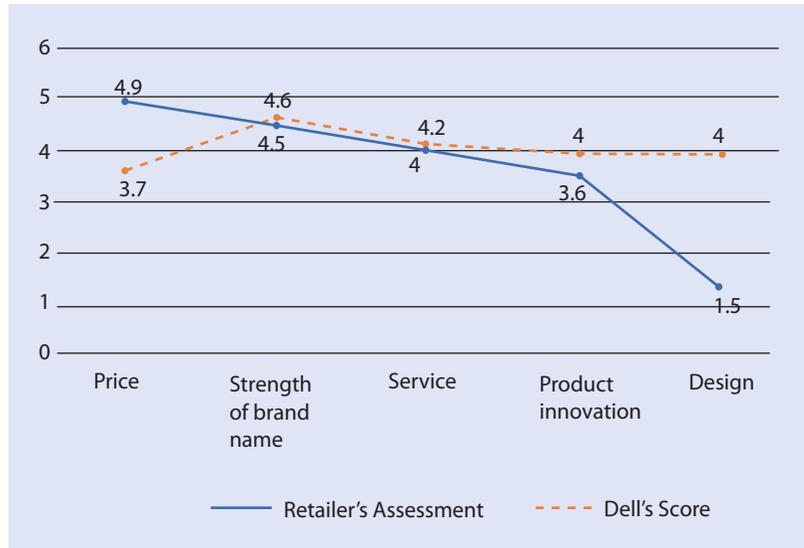
50 Bhandari, Rishi, Jonathan Gordon, and Andris Umblijs. "Getting beyond the buzz: Is your social media working?" *Financial Times*. Last accessed July 11, 2017. ► http://www.ft.com/cms/s/0/811df72c-c477-11e1-a98c-00144feabd00.html?ft_site=falcon&desktop=true#axzz4mYdPU7aR

51 Krug, Sammi. "Reactions Now Available Globally." *Facebook*. February 24, 2016. Last accessed July 11, 2017. ► <https://newsroom.fb.com/news/2016/02/reactions-now-available-globally/>

52 Egeland, Chad. "5 Metrics You Should Be Tracking on Twitter, But Probably Aren't." *Social Media Today*. February 7, 2015. Last accessed July 11, 2017. ► <http://www.socialmediatoday.com/content/5-metrics-you-should-be-tracking-twitter-probably-arent>.

48 AMA. "Dictionary." Last accessed July 11, 2017. ► <https://www.ama.org/resources/Pages/Dictionary.aspx?dLetter=C>.

Fig. 9.2 Comb Analysis: Divergences in Evaluating Product Factors by Dell and its Retail Distributors



There are also economic metrics associated with the audience metrics, such as:

- **Cost per Click (CPC):** Cost associated with each click on an advert.⁵³
- **Cost per Action:** Cost associated with each of the single actions that are monitored.⁵⁴

9.3.2 Transforming Information into Knowledge: Qualitative Analysis

The next step in the use of the data is the transformation of *information* into *knowledge*. One important way to do so is in a non-statistical way—instead, going through the judgment process of experts, both external and internal to the organization. They will make their judgment based on their own internal computation, experience and gut feeling. How can one use these expert judgments in a methodical way? We describe three such approaches: comb analysis, the Delphi methodology and opinion leaders' predictions and evaluations.

9.3.2.1 Expert Surveys: Comb Analysis

Retailers are closer to customers than a manufacturer. They are experts on buyers. “Comb chart analysis” is a technique for seeking the opinions of

the producer for the criteria most important to consumers regarding the selection of a product and comparing it with the opinions of the retailers.⁵⁵ For instance, if the computer manufacturer Dell wants to know why it is selling fewer desktops to the electronics retail chains than HP does, it could use a comb chart analysis. There are three steps. First, researchers ask the retailers to rate (e.g. on a scale of 1–5) the importance to its customers of various purchase criteria such as price, brand, reputation, service, product innovation and so on.

Figure 9.2⁵⁶ shows the retailers' evaluation of importance of Dell's score. The comb analysis indicates that Dell needs to lower its price—the most important purchase criterion. But it can also cut back on design and the cost associated with it.

9.3.2.2 Expert Surveys: The Delphi Methodology

The Delphi methodology was created in the 1950s by the RAND Corporation think tank. The goal is to reach consensus by experts on a certain topic. The Delphi method combines quantitative and qualitative data. It uses a group process consisting of about 10–15 experts. Anonymity is protected. The researchers solicit written responses to questions while preventing direct communication

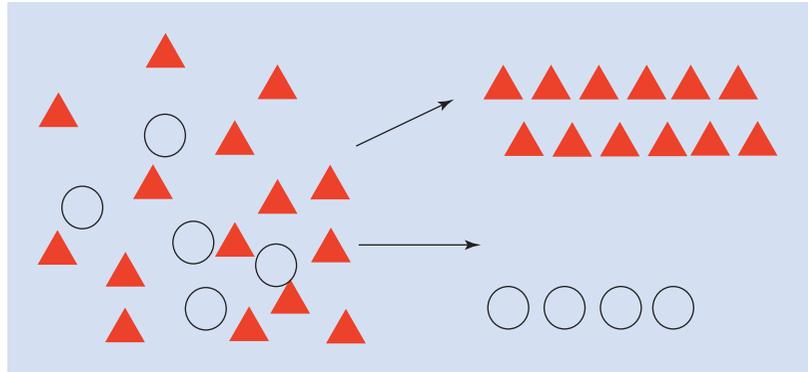
53 Google. “Cost-per-click (CPC): Definition.” Last accessed July 13, 2017. [▶ https://support.google.com/adwords/answer/116495?hl=en](https://support.google.com/adwords/answer/116495?hl=en).

54 Loomer, Jon. “How to Measure Facebook Advertising Success: Monitor These 5 Metrics.” *Jon Loomer Digital*, March 11, 2014. Last accessed July 11, 2017. [▶ http://www.jonloomer.com/2013/09/10/facebook-ads-metrics/](http://www.jonloomer.com/2013/09/10/facebook-ads-metrics/).

55 Koch, Richard. *The Financial Times Guide to Strategy*. London: FT Prentice Hall, 2000, 193.

56 Graph based on Koch, Richard. *The Financial Times Guide to Strategy*. London: FT Prentice Hall, 2000, 193.

■ Fig. 9.3 Organizing and Classifying the Data



among the respondents. In the first round of questions, questions have several answers with scores ranging from one to ten. In the second and subsequent rounds, participants are provided with information on how the group rated the same questions and a summary of comments made by each participant. Then, the participants receive the same questions again. The Delphi rounds continue until a predetermined level of consensus is reached or no new information is gained.⁵⁷

9.3.2.3 Surveying Trendsetters and Opinion Leaders

A third way to access experts is to survey the opinion of critics. Professional critics and reviewers are prime examples for opinion *leaders*, as well as *predictors*, of their respective audiences. Critics often function to represent the tastes of their audience and they are therefore “leading indicators” just as much as they are opinion shapers.⁵⁸

9.3.3 Quantitative Analysis: “Data Mining” —Overviews of Techniques

The next step in the quantitative interpretation of data is the transformation of “information” into “knowledge” through statistical techniques. This is described as “business modeling” or

“business analysis.” More popularly, it is known as “data mining.” This activity is a central part of what has become known (and feared) as “Big Data.”

There are several basic techniques for “data mining” in audience and market research. They are included in the software packages offered by companies such as Oracle, IBM, Microsoft, SAS and Statsoft. The purpose of this discussion is not to make you an expert in the use of these tools, but to give you a broad understanding of what they are and for what purposes they can be used.

9.3.3.1 Organizing and Classifying the Data

In this technique data gets segmented by category; for example, male or female, age, group, watching or not watching and so on, and its behavior is observed (■ Fig. 9.3).⁵⁹ This is similar in concept to the organization of data described above for the creation of basic media metrics.

9.3.3.2 Attribute Importance

This technique is useful for finding the factors that have the most influence on observed behavior (■ Fig. 9.4).⁶⁰ For example, the factors most likely to influence customers to purchase a certain product, or to respond to a marketing campaign. The technique works by considering each attribute and weighting its importance

57 Green, Jeremy et al. “Forecasting market demand for new telecommunications services: An introduction.” *Telematics and Informatics* 19, no. 3 (August 2002). Last accessed July 11, 2017. ► <http://www.sciencedirect.com/science/article/pii/S0736585301000041>.

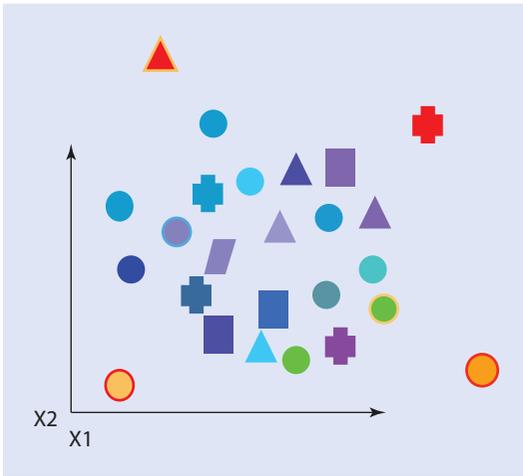
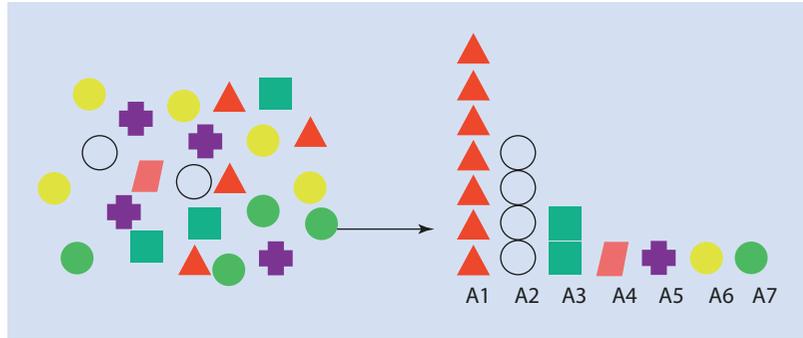
58 Eliasburg, Jehoshua, and Steven M. Shugan. “Film Critics: Influencers or Predictors.” *Journal of Marketing* 61, no. 2 (April 1997): 68–78.

59 Berger, Charlie. “Oracle Data Mining 11g Release 2: Competing on In-Database Analytics.” *Oracle White Paper*. February 2012. Last accessed July 13, 2017. ► <http://www.oracle.com/technetwork/database/options/advanced-analytics/odm/twp-data-mining-11gr2-160025.pdf>.

60 Berger, Charlie. “Oracle Data Mining 11 g Release 2: Competing on In-Database Analytics.” *Oracle White Paper*. February 2012. Last accessed July 13, 2017. ► <http://www.oracle.com/technetwork/database/options/advanced-analytics/odm/twp-data-mining-11gr2-160025.pdf>.

9.3 · Analyzing the Data

■ Fig. 9.4 Attribute Importance



■ Fig. 9.5 Anomaly Detection

against the other attributes. The factors are then ranked from strongest to weakest.

9.3.3.3 Anomaly Detection

This technique identifies unusual cases based on deviation from the norm. For example, an “outlier” point may merit attention (■ Fig. 9.5).⁶¹ It could be an error, flagging a needed change in the data collection. Or, it may help to identify a potential earlier adopter and innovator. Or, it points to someone who abuses the system, such as an identity thief.

9.3.3.4 Clustering

This kind of analysis helps to find groupings within the data which may represent customer segments. What are the customer categories that

are most likely to spend more than \$1000 per year on books? This kind of data analysis may help design different marketing approaches to different user categories.

To identify clusters, algorithms try to find a grouping of data at a point with a small distance among them, which means that they are similar. A “centroid” point represents the most representative point in a cluster (■ Fig. 9.6).⁶² The algorithm searches for “clouds of observations” where observations occur frequently and which are clearly distinct from other clouds by space between them in which no (or very few) observations occur.

9.3.3.5 Association

An association rule could be, “Given past choices of film viewing by a viewer category, a new film x may be expected to be chosen, with a certain statistical confidence.” This kind of analysis can be used to find cross-sell opportunities, with product A often associated with product B and indicating root causes of user behavior (■ Fig. 9.7). The associations that are identified can help design recommendations, promotions and product bundles.

9.3.3.6 Feature Extraction

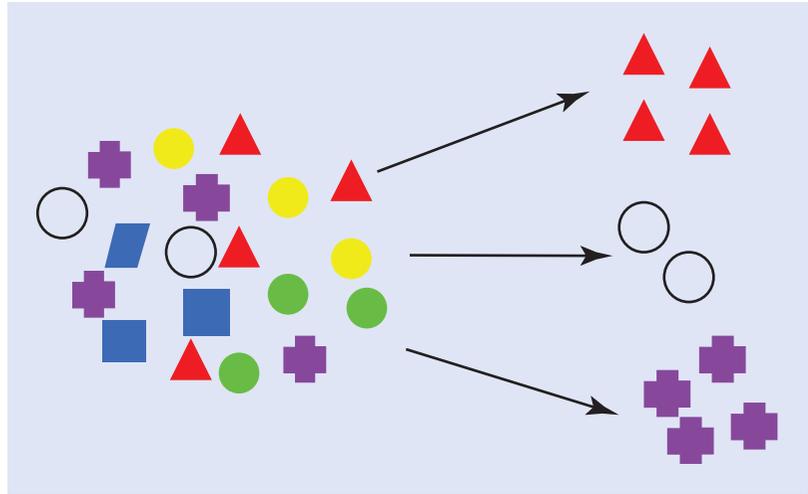
The number of attributes or variables may be large. Statistical technique reduces them to underlying essential attributes. This results in a smaller number of variables (■ Fig. 9.8).⁶³ One such technique for feature extraction is principal component analysis (PCA). PCA aims to identify the main factors, for behavior, and identifies the vari-

61 Berger, Charlie. “Oracle Data Mining 11g Release 2: Competing on In-Database Analytics.” *Oracle White Paper*. February 2012. Last accessed July 13, 2017. ► <http://www.oracle.com/technetwork/database/options/advanced-analytics/odm/twp-data-mining-11gr2-160025.pdf>.

62 Oracle. “Data Mining Concepts.” Last accessed July 13, 2017. ► http://docs.oracle.com/cd/E11882_01/datamine.112/e16808/algo_kmeans.htm#DMCON238.

63 Berger, Charlie. “Oracle Data Mining 11g Release 2: Competing on In-Database Analytics.” *Oracle White Paper*. February 2012. Last accessed July 13, 2017. ► <http://www.oracle.com/technetwork/database/options/advanced-analytics/odm/twp-data-mining-11gr2-160025.pdf>

Fig. 9.6 Clustering



9

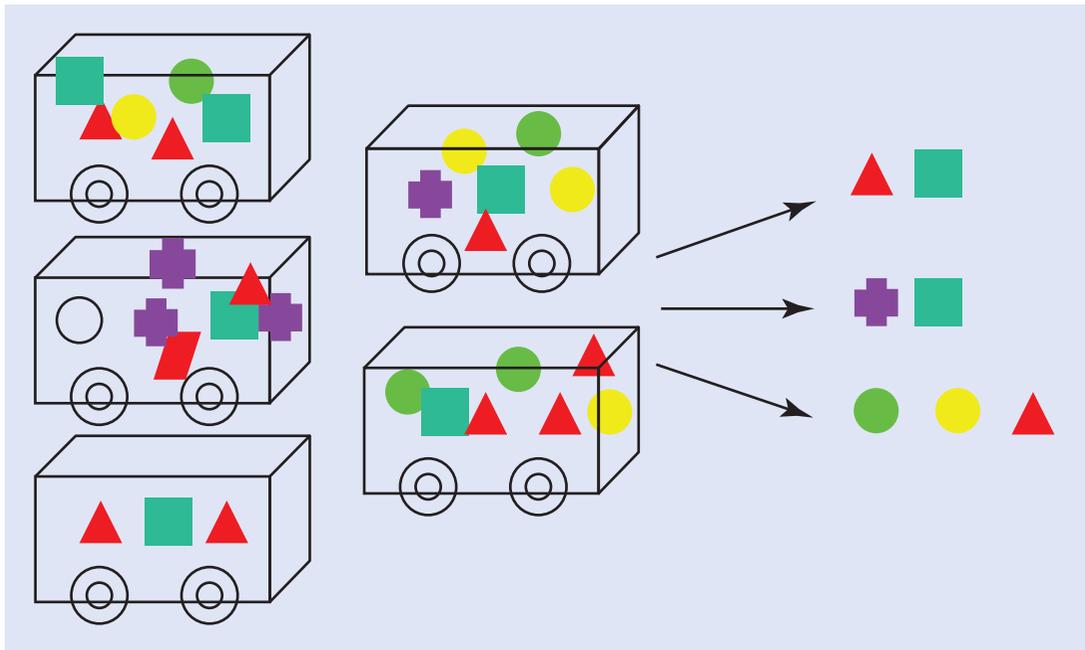


Fig. 9.7 Association

ables of a dataset which explain most of the variance in the data (Fig. 9.8).⁶⁴

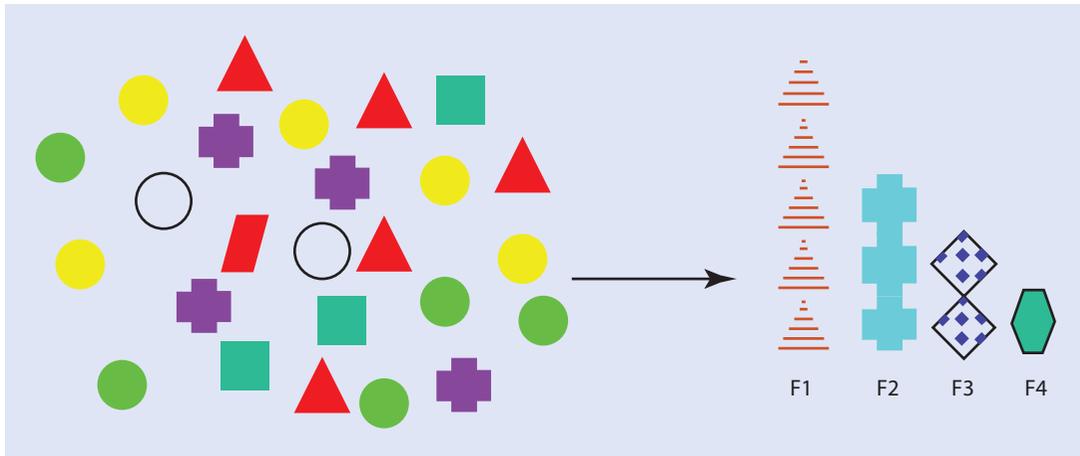
9.3.3.7 Correlation

This approach tries to find the inter-relationship among various factors, such as age and the frequency of watching a particular program. The

strength of the linear association between two variables is quantified by the correlation coefficient.⁶⁵ The correlation coefficient $\rho(x,y)$ takes a value between -1 and 1 : the larger (or smaller) the number, the stronger the correlation of the two variables. A $+1$ or -1 coefficient indicates perfect correlation (all points would lie along a

64 Berger, Charlie. "Oracle Data Mining 11g Release 2: Competing on In-Database Analytics." *Oracle White Paper*. February 2012. Last accessed July 13, 2017. ▶ <http://www.oracle.com/technetwork/database/options/advanced-analytics/odm/twp-data-mining-11gr2-160025.pdf>.

65 Berger, Charlie. "Oracle Data Mining 11 g Release 2: Competing on In-Database Analytics." *Oracle White Paper*. February 2012. Last accessed July 13, 2017. ▶ <http://www.oracle.com/technetwork/database/options/advanced-analytics/odm/twp-data-mining-11gr2-160025.pdf>.



■ Fig. 9.8 Feature Extraction

straight line in this case). A positive correlation shows a positive association, i.e. increasing the value of one variable corresponds to a higher value in the other variable; for example, the older the observed person, the higher the frequency of watching a given program. In contrast, with a negative correlation, the older a person, the lower the frequency of watching the program indicates no association between the variables. A correlation coefficient close to zero.

An example is the question as to whether there is a correlation between movie theater ticket sales and a film's budget. An analysis of almost 12,000 movies looked at each film's budget in relation to how many tickets it sold. While one would assume that a bigger budget—which brings with it better known movie stars, more special effects and so on—would result in high ticket sales, no such correlation was found.⁶⁶

9.3.3.8 Regression and Econometric Estimation

Regressions add more factors than a simple one-on-one correlation. The technique identifies how several factors explain the target variable and also the impact of other exogenous factors (■ Fig. 9.9).⁶⁷

Econometrics is an application of various regression techniques over several variables; typically using either cross-section observations over several people or firms, or a time series analysis over several time periods. It allows the synthesis of large amounts of information. It also provides a framework for a systematic analysis through explicit assumptions, which is known as “modeling.” An example is the explanation of how sales—which is known the “dependent,” or the “explained,” or the “left-hand” variable—are related to several “independent,” or “explanatory,” or “right-hand” variables such as price.

Given adequate data, an econometric technique can identify the key variables that may affect sales—price, advertising effort and so on. It uses “control variables” to adjust for factors that may have affected sales—such as the state of the economy, the growth of population, or the season.⁶⁸ One can also add demographic variables such as age, education and gender, or psychographic variables that reflect a buyer's lifestyle—such as their activities, interests, and opinions (known as AIOs):

$$\text{Unit sales} = a + b_1 \text{ price} + b_2 \text{ advertising} \\ + c_i \text{ control variables} + \mu$$

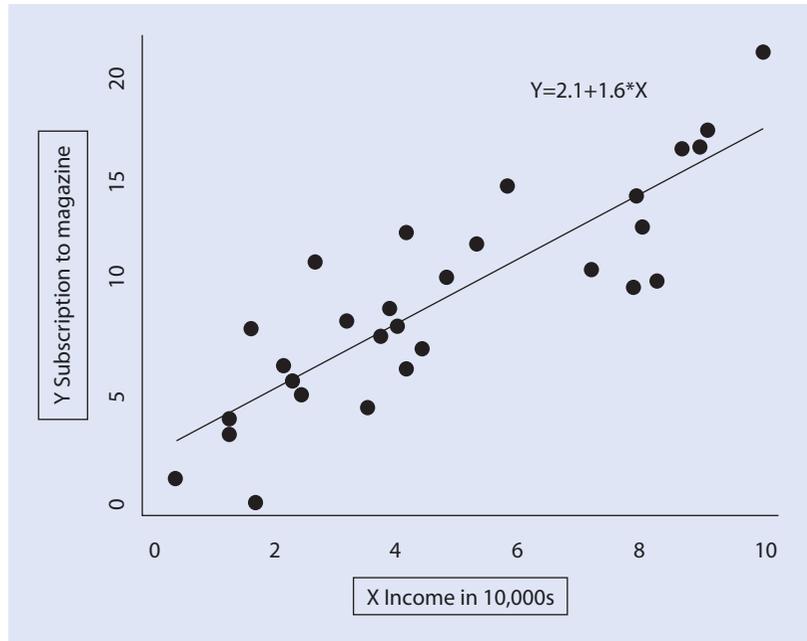
The a and the several b and c coefficients are parameters that are estimated in the regression.

66 Olson, Randy. “Does a bigger film production budget result in more ticket sales?” Randal S. Olson. December 29, 2014. Last accessed July 11, 2017. ▶ <http://www.randalolson.com/2014/12/29/does-a-bigger-film-production-budget-result-in-more-ticket-sales/>.

67 Statsoft. “Elementary Concepts in Statistics.” Last accessed July 11, 2017. ▶ <http://www.statsoft.com/textbook/elementary-concepts-in-statistics/>.

68 Nagle, Thomas T., and Reed K. Holden. *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*, 2nd ed. Hoboken: Prentice Hall, 1995.

Fig. 9.9 Econometric Regression Analysis: Magazine Subscriptions and Income



The μ is the unexplained residual, also known as the “error term.”

The most common form of a regression technique is called ordinary least squares (OLS). One can estimate OLS regression using readily available statistical software packages, e.g. STATA, SAS, EXCEL, or Minitab.⁶⁹ The analyst defines the variables and the model and inserts the data. OLS estimation results in parameter estimates for the coefficients b_1, b_2, \dots, c_i that best fit the data. The “best fit” is defined as the lowest sum of the squares of the difference between the actual value of the data and the value predicted by the equation when it uses the parameter estimates of the parameters.

A statistical measure known as the “R-square” reflects the overall fit of the model: the percentage of the observed variation in the dependent variable that is explained by the independent variables. An R^2 above 0.8 would indicate a fairly good fit of the model. The statistical significance of each individual coefficient is measured by the

“t-statistic.” The rule of thumb when interpreting results is that t-statistics must be greater than 2 (or -2) to be a statistically significant.

Often, relations among variables are not linear but, rather, exponential. For example, higher income may lead to higher consumption, but not in a linear fashion but rather that one consumes more in the same proportions as the increase in income. In such situations, a “logarithmic” model allows the determination of the exponential growth rates and other non-linear relations. For example, one could define a model as:

$$\text{Sales} = a(\text{price})^{b_1} (\text{advertising})^{b_2} (\text{other variables})^{b_i}$$

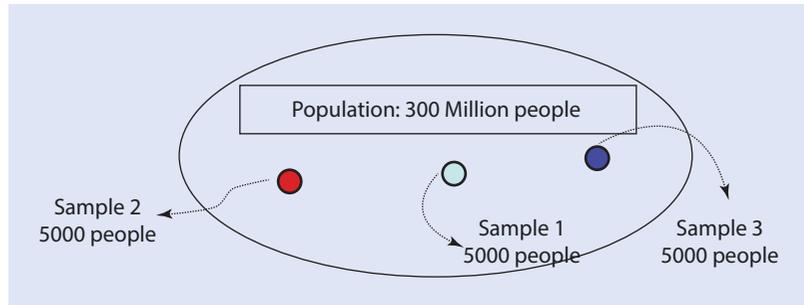
To calculate those parameters, a, b_1, b_2, b_i , one takes the natural logarithms (\ln), which transforms the exponential relation into a linear one, and then find the linear “least square” best fits:

$$\ln \text{sales} = \ln a + b_1 \ln \text{price} + b_2 \text{advertising} + b_i \ln \text{other variables} + u$$

The coefficients of the logarithmic models are mathematically the “elasticities”, or the “sensitivities” of the explained (or left-hand) variable “Sales” with respect to changes in the explanatory

⁶⁹ Princeton University, Data and Statistical Services. “Interpreting Regression Output.” Last accessed July 11, 2017. ► http://ds.princeton.edu/online_help/analysis/interpreting_regression.htm.

■ Fig. 9.10 Sampling



(or right-hand) variables of price, advertising and so on.⁷⁰

As mentioned before, the elasticity of sales with respect to price (price elasticity) is defined as the percentage change in sales as a result of a percentage change in price. This can be expressed as:

$$\text{Price elasticity} = \frac{\% \text{ change in sales}}{\% \text{ change in price}}$$

When demand is said to be elastic, it means that a relatively small difference in price will affect the quantity demanded considerably. In contrast, where elasticity is low, it would take a relatively high change in price to make much difference in the quantity demanded.

Beyond the models described—the linear and the logarithmic models—there are many other specifications for econometric demand estimation that involve increasing complexity.

9.3.3.9 Statistical Inference: Sampling and Statistical Confidence

There are two ways to measure a target market. One can measure the entire group we are interested in. An example of this would an analysis of, say, all women aged 23–30. Alternatively, one could look at a smaller group, a “sample,” as a representative of the larger population.⁷¹ This is

cheaper, faster and more practical. But it may provide incomplete and unrepresentative coverage.

After choosing a sample and measuring its response, the researchers must address the question of how the sample relates to the overall population. Suppose one takes three independent samples of 5000 people from the the overall national population and asks all sample groups the same question: “Did you watch the Golden Age channel last week?” (■ Fig. 9.10). There is an excellent chance that each sample result will differ.

One would expect that the three samples would yield similar but not identical estimates. For example, we could find that for Sample 1: $p = 25\%$, Sample 2: $p = 27\%$ and Sample 3: $p = 24\%$. Which one is right? If we take a large number of samples, we would find that their results are distributed in a normal distribution around a peak, which corresponds to the true number.⁷² But finding that “true peak” will take considerable sampling and thus is not practical. Due to constraints of time and money, researchers typically take one sample and follow up with “inferential statistics” to state how certain they are of the result. 65%, 95% or 99% “confidence levels” are often chosen. One starts with the center of the distribution of the observed sample. One “standard deviation” from the center will include about 65% of all sampling results, two standard deviations will include about 95% and three standard deviations will include 99%. In other words, 95% of samples will lie within about two standard deviations from the observed sample.

70 AmosWeb. “Elasticity and Demand Slope”. 2000–2001. Last accessed July 11, 2017. ► http://www.amosweb.com/cgi-bin/awb_nav.pl?s=wpd&c=dsp&k=elasticity+and+demand+slope.

71 There are several types of samples. In a “random sample” each unit in the population has an equal chance of being selected. A “stratified random sample” is when one defines sub-groups and selects random samples from each sub-population group and pools them. A “convenient sample” uses accessible observations, for example choosing students on-campus for an academic marketing research study by a professor. The “judgment” sample is chosen according to the estimate of someone familiar with the characteristics of the overall population.

72 This follows the “Central Limit Theorem:” the true percentage of people watching a program will be at the peak of the distribution of the multiple samples.

A *z*-score (or “standard score”) is a statistical measure of comparing a single sample result to a data set that is normally distributed. It uses the standard deviation of the sample mean for comparison. A *z*-score can be calculated using the following formula:

$$z = (X - \bar{x}) / SD$$

where *X* is the observed value, \bar{x} is the sample mean and SD is the standard deviation of the sample.

For example, if the sample mean is 0.25, with a standard deviation of 0.02 for 5000 observations, there is a 65% chance that the true population mean is between 0.23 and 0.27 (one standard deviation from 0.25). There is a 95% chance the population mean is between 0.21 and 0.29 (two standard deviations) and a 99% chance it is between 0.19 and 0.31 (three standard deviations). The potential error, due to sample being “off,” is the term *e*:

$$e = z \sqrt{\frac{pq}{n}}$$

where:

- *p* = proportion that answered positively;
- *q* = (1-*p*), those who answered negatively.

Suppose these are the parameters:

- *p* = 0.25 (25% of sample watched the program);
- *q* = 0.75 (75% did not watch);
- *n* = 5000 (sample size);

and we chose a 95% probability, which means, from a table, that *z* = 1.96:

$$e = 1.96 \sqrt{\frac{.25 \times .75}{5000}} = 0.12 \text{ or } 1.2\%$$

If there are 100 million TV households in the USA, then the number of American HHs that watched the Golden Years channel last month, with 95% certainty, lies between 23.8 and 26.2 million (25 million ± 1.2 million).

This method works for questions requiring a yes/no response (e.g. watched/did not watch). However, it does not work well for continuous types of data, such as determining the average amount of time spent watching a particular program. For this type of calculation, one uses a related method.

Similarly to before, one estimates the mean of the population by finding the mean of the sample and adding/subtracting an appropriate margin of error. The confidence interval equation is as follows:

$$u = \bar{X} \pm e$$

where:

- *u*: mean of the population;
- \bar{X} : Mean of the sample;
- *e*: margin of error.

The margin of error (*e*) is determined by the standard deviation (SD) and the sample size *n*. The differences are that, instead of a “*z*-score,” one now looks up a “*t*-score” in a statistical table, and the numerator is the standard deviation:

$$e = t \left(\frac{SD}{\sqrt{n}} \right)$$

$$u = \bar{X} \pm t \left(\frac{SD}{\sqrt{n}} \right)$$

Like the *z*-score, the *t*-score is determined by the level of confidence (e.g. 0.90, 0.95, 0.99) that is desired and the sample size *n*. It can be found in the relevant tables by inputting values for *n* and *r*.

If Nielsen claims its measurement results are within the 95% confidence (*r* = 0.95), what does that mean? It means that if its study was repeated 100 times, then, in 95 of the cases the result would be in the stated interval.

9.3.3.10 Conjoint Analysis

Conjoint analysis—also discussed in ► Chap. 3 Production Management in Media and Information and in ► Chap. 11 Pricing of Media and Information—is another standard market research tool.⁷³ Conjoint analysis is used to measure the tradeoffs people make in choosing among products and services.⁷⁴ The foundation of this technique is the assumption that a product can be disaggregated into individual attributes. For example, a TV set has attributes such as size, price, style, picture sharpness, color and

73 Green, Paul E., and Vithala R. Rao. “Conjoint Measurement for Quantifying Judgmental Data.” *Journal of Marketing Research* 8, no. 3 (August 1971): 355–363.

74 Populus. “Conjoint Analysis.” Last accessed July 11, 2017. ► http://www.populus.com/files/Conjoint%20Introduction_1.pdf.

9.3 · Analyzing the Data

so on. The respondent is asked to choose between different levels of a pair of product attributes.⁷⁵ This enables the researcher to identify the value that a consumer attaches to each product attribute. The data are then analyzed with statistical tools and the utility of a consumer's relative strength of preference for each level of each product attribute is determined. The value of a product, it is assumed in this methodology, is equal to the sum of utilities to the consumer. This enables the researcher to predict the prices which the consumer would pay for various combinations of attributes in a product. It helps to set prices and target buyers, even before the product is developed. The methodology works best for products that are evaluated by consumers based on separable attributes. But this is often not the actual buying experience, because consumers often do not pay much attention to specific product attributes.⁷⁶

This statistical analysis is calculation-intensive, but computer packages are available that generate an optimal set of tradeoffs for each participant. The example concerning an MP3 Player presented in Table 9.2 shows how conjoint analysis attributes importance to products. Acceptance likelihood is calculated by adding up the sums of the

attribute level utilities contained in the product profile.^{77, 78}

9.3.3.11 Diffusion Models

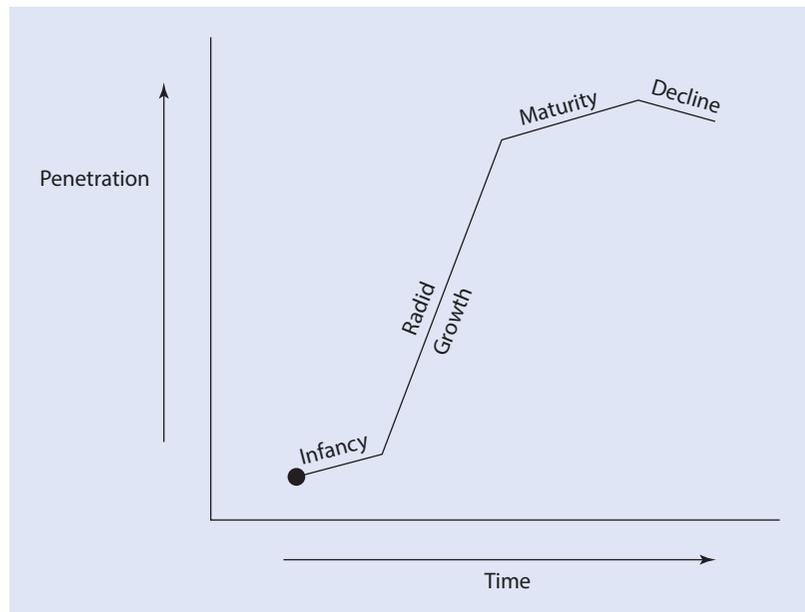
Generally, the adoption of a new product follows an S-curve pattern (Fig. 9.11). Such an S-curve helps to predict demand for a new product. Adoption rises slowly at first. It then accelerates as the pace of adoption picks up. In time, however,

Table 9.2 Conjoint Analysis Example: The Importance of Attributes of MP3 Players (Scale 1–10)

Quality	8.24
Styling	6.11
Price	2.67
User friendliness	7.84
Battery life	4.20
Customer service	5.66

Nagle, Thomas T., and Reed K. Holden. *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*, 2nd ed. Hoboken: Prentice Hall, 1995.

Fig. 9.11 Stages of Product Penetration



75 Nagle, Thomas T., and Reed K. Holden. *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*, 2nd ed. Hoboken: Prentice Hall, 1995.

76 Nagle, Thomas T., and Reed K. Holden. *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*, 2nd ed. Hoboken: Prentice Hall, 1995.

77 Populus. "Conjoint Analysis." Last accessed July 11, 2017. ► http://www.populus.com/files/Conjoint%20Introduction_1.pdf.

78 A conjoint analysis of cellphones is provided in ► Chap. 11, Pricing of Media and Information.

the market becomes saturated and the adoption rate of new users slows down.

$$\text{Cumulative sales} = \frac{a}{1 + be^{-kt}}$$

where t is time, and a , b and k are for saturation level and steepness and e is the base of natural logarithms, about 2.718.⁷⁹ This S-curve of adoption is called a “diffusion,” “logistic” or “epidemic model”. The latter term derives from the notion that a product will spread like a virus as it catches on.⁸⁰ Examples for diffusion are the adoption of smartphones, the audiences of a word-of-mouth hit movie, or the diffusion of Internet Protocol Version 6 (IPv6).

Different S-shapes occur with different parameters a , b , and k . One has to determine, from early data, what the parameters are, for a projection of the rest of the S-curve.

To find the acceleration inflection and the saturation level,⁸¹ and can compare the new product with an earlier and similar product. One can then use a “historical diffusion index” (HDI) to compare a product with a predecessor:

$$\text{HDI}_t = \frac{\text{Penetration } A_t \times 100}{\text{Penetration } B_t}$$

For example, compare the diffusion of DVD players to the diffusion of VHS video cassette recording players 15 years earlier. VHS was in 95% of US households in 2008, which was its maximum market demand. DVD was in 75% of households in 2008. This means that the $\text{HDI} = (75 \times 100) / 95 = 79\%$. Thus, the DVD market in 2008 was still 21% below its potential, measured by VHS. The VCR reached a 75% penetration after 12 years, while DVD took only 6 years. Hence, the DVD penetration rate was two times faster than that of VCR. The VCR took 4 years to rise from 75% to 95% and, assuming that DVDs maintain their penetration pace, it was likely to

take another $4/2 = 2$ years to reach 95%. One can make similar comparisons of DVD with Blu-ray DVD. However, it was quite possible that consumers do not greatly value high-definition over standard-definition quality, or that they shift to on-demand online video, and both diffusion speed and maximum market demand for Blu-Ray may therefore be lower. And that is what, indeed, happened. Thus, the problem with the diffusion approach is that there are many differentiating variables and they make comparisons among products unreliable.

The Bass Diffusion Model is a way to predict the gradual penetration of a product. It is discussed in ► Chap. 10, Marketing of Media and Information. The model has been widely used in marketing forecasts, especially for new products and technology.⁸²

9.3.3.12 Autonomous Data Mining

The various data mining techniques can be applied in two ways: “supervised” and “unsupervised” learning. In supervised learning, the data analyst identifies a target attribute and several variables, sets the questions for the data, and then the program finds the strength of relationship, categories and other features. In contrast, in unsupervised learning the analyst does not specify target attributes; there is no particular association, or even a goal. Instead, the data mining program will find associations and clusters in the data by itself, if such exist.⁸³

9.3.3.13 What to Do with Data Mining?

Beyond methodology, there are two larger questions regarding data mining of audience preferences and behavior:

1. Does it really work?
2. Does it work too well? (The “creepiness” factor).

Examples of the application of data analysis are everywhere. For example, the music industry tracks the performance and potential of music

79 McBurney, Peter, Simon Parsons, and Jeremy Green. “Forecasting market demand for new telecommunications services: An introduction.” *Telematics and Informatics* 19, no. 3 (2002): 225–249.

80 Wilson, Ralph. “The Six Simple Principles of Viral Marketing.” *Web Marketing Today*, February 1, 2005. Last accessed July 11, 2017. ► <http://www.wilsonweb.com/wmt5/viral-principles.htm>.

81 Carey, John, and Martin Elton. “Forecasting demand for new consumer services: Challenges and alternatives.” *New Infotainment Technologies in the Home: Demand-Side Perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates, 35–57.

82 The model makes several simplifying assumptions:

- The market potential remains stable over time.
- The model is binary; that is, it assumes that a customer either adopts or does not adopt an innovation.
- Market and product characteristics have no influence on product diffusion.

83 Berger, Charlie. “Oracle Data Mining 11 g Release 2: Competing on In-Database Analytics.” *Oracle White Paper*, February 2012. Last accessed July 13, 2017. ► <http://www.oracle.com/technetwork/database/options/advanced-analytics/odm/twp-data-mining-11gr2-160025.pdf>.

9.4 · Conclusions and Outlook

instantaneously from data on streaming and downloading. The information generated by digital platforms such as Spotify and Twitter is analyzed for insights. Labels can see which songs perform with various audiences and the behavioral patterns of those audiences.

This being part of marketing operations, stories of accomplishment are well-publicized. Yet, it is important not to be swayed by anecdotal stories of success. In other instances, there have been outright failures.⁸⁴

Google Flu: Google tried to use big data information (specifically, user searches of symptoms) to predict and analyze trends for the flu faster than the Centers for Disease Control (CDC). For a while, the company received highly favorable publicity.⁸⁵ In subsequent years, however, Google's methodology could not identify flu outbreaks reliably.

2016 US Presidential Election: All big data analytics showed that Hillary Clinton was going to win the election, with a high certainty. Yet, this did not happen.⁸⁶

9.4 Conclusions and Outlook

9.4.1 Case Discussion

Viacom Golden Years—Conclusion For the Cable TV Channel Golden Years, What Do We Find?

- A Delphi study confirms the experts' overall view that there is room and demand for such a channel. Delphi experts believed in the potential in a market, niche, and expected about three hours of weekly watching.
- Focus groups show an active support for the concept.
- Nielsen results shows that senior-oriented channels have fairly substantial audiences: Fox News ranked #4 among cable channels, History Channel #5, Nick at Nite #11, AMC #12, TV Land #22.
- A Simmons market survey found in test marketing among viewers of the Golden Years channel a viewing of 120 minutes \pm 3 for a week.
- An econometric analysis shows that the likelihood of subscribing to the Golden Years channel is positive and significant with age and income, with sports and news programming, rural location and a female household. Audience surveys and advertising-based mea-

asures show a strong demand by users of the Golden Years channel (and users of the Golden Years website) for travel, drugs and insurance. These are large markets for advertisers.

- A diffusion analysis shows a likely slow but steady growth in audience.
- Focus groups and surveys (mall, phone and follow-up) confirm demand.
- Conclusion: Viacom should go ahead with the Golden Years channel based on demand analysis, but subject to an analysis of cost.

For the Golden Years Magazine, What is the Evidence?

- Test marketing in New York State shows ABC national audited circulation figures of about 133,000 subscribers, and demand experiments with differentiated subscription offers show an optimal circulation of about 120,000.

Both figures are low, and unlikely to generate a profit.

- Demographic trends (seniors) are favorable but, in general, magazine circulation and advertising are in decline.
- Conclusion: Given the various negatives, and considering that Viacom has no experience as a magazine company, it should not proceed with the magazine. (This assumes that the magazine does not generate significant spillover benefits to the GYC cable channel or to the web portal that would justify it to function in the role of a "loss leader.")

For the Website Golden Years Portal, What is the Evidence?

- From site-based Google Analytics data, Viacom finds that the 65+ years cohort is fairly low in visitor count.
- Conclusion: The trends suggest growth potential. Because the return rates and page views of

84 McGlohon, Mary. "Data Mining Disasters: A report" Last accessed July 11, 2017. ► <http://www.cs.cmu.edu/~mmcgloho/pubs/accidents-sigbovik08.pdf>.

85 Leber, Jessica. "The Failures of Google Flu Trends Show What's Wrong with Big Data." *Fast Company*. March 13, 2014. Last accessed July 11, 2017. ► <https://www.fastcompany.com/3027585/the-failures-of-google-flu-trends-show-whats-wrong-with-big-data>.

86 Lohr, Steve, and Natasha Singer. "How Data Failed Us in Calling an Election." *New York Times*. November 10, 2016. Last accessed July 11, 2017. ► <https://www.nytimes.com/2016/11/10/technology/the-data-said-clinton-would-win-why-you-shouldnt-have-believed-it.html>.

visitors are solid, Viacom should take a chance with growth potential and go ahead with the Golden Years web portal.

Should Viacom Modify the Content of the Golden Years Channel?

- An econometric study shows strong positive correlations of its content with sports and news/documentaries. There were also correlations to romance and a

female audience. These are also a positive effect of late-night programming. Programming decisions should take these preferences into account.

- A conjoint analysis found that consumers base 72.3% of their decision to subscribe on price, 39.6% on the number of films per day, 18.2% on discount coupons, but a negative 3% on easy listening hours. Therefore, Viacom should avoid easy listening hours on the Golden

Years channel and increase the frequency of films per day. Viacom should also lower the cable channel price, since users seem to be very price sensitive.

- For the Golden Years website, the audience research found that there was an unexpectedly large interest in the 55–65 year cohorts. This suggests editorial and pricing policies to attract and keep such viewers who are still outside the core 65+ target group.

9.4.2 Challenges in Audience and Market Research

This should be the golden age of demand research. Many of the data constraints of the past have been lifted, and data processing and distribution have become easy and cheap. But, even with better tools, it is much harder to do demand research today than in the past. It is harder to estimate demand for new products and services in a rapidly changing environment, with fragmented audiences, much greater choice and shorter attention spans.

9.4.2.1 Challenge #1: Coordinating and Integrating the Several Data Flows about Users

Media firms want to know who they reach across the various distribution platforms, and advertisers want to know the total reach of a campaign. Networks providing content need to demonstrate the reach of their program, regardless of when it was watched and on what platform. Industries seek measurements that cross the boundaries of platforms, devices, time, and location of usage. This is technologically complex, expensive to develop and hard to operate reliably in real time.

To be effective in measuring across platforms requires a presence by measuring companies on these platforms, or at least close partnerships. This strengthens measurement companies that operate in several media segments. There are significant economies of scale, economies of scope, and network effects that,

together, generate advantage to market leaders such as Nielsen, ComScore and, potentially, Google.⁸⁷

9.4.2.2 Challenge #2: Individualization

The logic of increasingly targeted audience research is to drill down to the individual. The collection technology is capable of individualization and can track individual media behavior, rather than broad aggregates, and to do so for an individual's overall media consumption. This means a device that people carry with them, probably a smartphone app of some sort, that can identify media watermarks in the digital flow of information, regardless of the platform or location in which they are received.

The trends go in that direction. Cellphones have been used in media measurement beyond their role in measuring their own usage as media devices. Researchers can use specially adapted mobile phones to measure what TV programs consumers watch.

On top of that, smart phones are increasingly used as payment devices. Thus, location, transaction, media consumption and personal information could be correlated. One could then measure, in real-time, the effectiveness of advertising, promotions, and differentiated pricing.

There are significant privacy issues here. There are also important questions on whether individualization means that some consumers will get

⁸⁷ Google. "TV, Digital, & Marketing Mix Modeling." Last accessed July 11, 2017. ► <https://www.google.com/analytics/attribution/features/>.

better deals than others. The airline model of highly differentiated pricing will spread. It also means that some of the content will be individualized, and that the concept of the mass audience served by mass media and measured in highly aggregate ways will become anachronistic.

9.4.2.3 Challenge #3: The Globalization of Media Audience Research

The third challenge is the internationalization of media consumption and supply, and hence of measurements. Nielsen is active in over 100 countries. It also provides a measurement base that is generally accepted by partners to apply to transactions based on those numbers. Similarly, Google's measurements operate for numerous countries' users and websites.

9.4.2.4 Challenge #4: Restrictions on Data Collection and Use

Many countries have enacted privacy laws that restrict the collection and use of individual data. In addition to affecting the practices of audience research, these laws vary greatly. European countries are more concerned with the use of individual data than the USA, or many Asian countries. In the past, the practices of data collection and analysis were shielded by being aggregate and anonymous. With individualization, however, this will cross into personal identification. Thus, as technology is becoming more powerful in collection, interpretation and dissemination, the legal rules will create a significant pushback.

9.4.2.5 Challenge #5: Creation of New Research Methodologies

In the past, data collection methods had been inaccurate and slow. Given the relatively leisurely collection methods, analytical tools were similarly sluggish. On top of that, most of the academic demand models could not even be realistically applied in a business setting. They included variables and information that were not available in a way that could satisfy the rapid operational needs of companies.

Today, we have moved from data shortage to glut, and the main problem is how to analyze the data effectively and rapidly. It is here that social and behavioral scientists need to make progress,

and it is here that media audience practitioners and media academic researchers must collaborate.⁸⁸ In the near future, the tools of online tracking will permit a real-time observation of overall audience, global aggregations, large samples, customer individualization and the tracking of their consumer behavior. This will add power to data collection. But how is that data used? Current methodologies are limited.

Today's audience research instruments, described in this chapter, are probably just the beginning of the development of next generation tools that will utilize, in much more advanced ways:

- Behavioral research;
- Audience instant feedback;
- Trendsetter tracking;
- Cross-cultural sampling;
- Online and mobile tracking;
- Social network analyses;
- Internet-of-things sensor networks.

9.4.2.6 Challenge #6: Semantic Mining and Unstructured Data

The ability to combine structured data with unstructured material—that is, “text”—is an important frontier for media and audience research.⁸⁹ There are no strong methodologies to integrate non-quantitative “fuzzy” information. For example, social media tweets, blogs and so on are analyzed and combined with structured Nielsen data to extract more information about audience views of a program. Companies have been using Twitter and Facebook as means for finding demand for their products as well as the view of their brands.

9.4.2.7 Challenge #7: Integration with Managerial Decision Making

Collecting data and transforming it into information and knowledge is one thing. Making it operationally useful inside media organizations is another. New data products may not be known to or utilized by the media manager and creators, or are resisted since it might affect their content making decision.

88 Noam, Eli. “Research Demands on Demand Research.” In *Telecommunications Demand and Investments: The Road Ahead*. Eds. James Alleman et al. New York: Springer, 2014.

89 Berger, Charlie. “Oracle Data Mining 11 g Release 2: Competing on In-Database Analytics.” *Oracle White Paper*. February 2012. Last accessed July 13, 2017. ► <http://www.oracle.com/technetwork/database/options/advanced-analytics/odm/twp-data-mining-11gr2-160025.pdf>.

9.4.3 Conclusion

Demand analysis becomes more important:

- The greater the uncertainty;
- The greater the upfront investment;
- The greater the economies of scale and network effects;
- The more competitive alternatives there are;
- The shorter the product cycle.

Is the determination of audience preferences—demand analysis—a “bean-counting” by uncreative minds, a tool for pandering to audiences rather than for leading them? A manager should not make the choice between creative “gut” judgment and qualitative analysis. If used effectively, they are complementary.⁹⁰ The avant garde media manager may be three steps ahead of the audience, too far for popular success. The conventional media manager follows the audience by one step in effect, letting audience research make content decisions. The moderately successful media manager is probably one step ahead, using audience research as a starting point. And the successful innovator can be two steps ahead, with a creative understanding of the audience, market and society, plus research to lower the risk.

Reliance on the “gut feeling intuition” of single-minded entrepreneurs and of internal advocates can be the most expensive way to learn. If a film has a cost of \$50 million with a 20% probability grossing \$250 million, then improving the odds to 22% by smarter demand analysis raises expected profits by \$5 million, or 10%.

And, therefore, we should disagree with Goldman’s classic Hollywood adage that, “nobody knows anything.” Of course, nobody knows everything. But one can improve the odds on getting it right, and that is enough for a competitive advantage. Understanding one’s audience is a relatively cheap investment with a high return. Demand analysis is the key to improving the odds.

We are just at the beginning.

9.5 Review Materials

Issues Covered

We have covered the following issues in this chapter:

- Why market and audience research is important, but particularly difficult, for media firms;
- How media companies organize themselves to perform market research;
- How the data collection technology and methodology impacts business decisions;
- The ways to collect data at the user level and from the provider;
- How to apply techniques for data mining analysis;
- The techniques of data mining for demand analysis;
- How audience research is becoming globalized;
- The business impacts of collection methodologies;
- Why new research methodologies will be needed in the future;
- Whether media managers should be driven by demand numbers.

Tools Covered

We used tools for demand and market research:

- Automatic metering;
- Controlled and uncontrolled experiments;
- Personal interviews;
- Panel data use;
- Surveys (mail, phone, Internet);
- Focus groups;
- Mobile research;
- Self-reporting and auditing;
- Click-counting;
- Measuring Internet traffic: site-level measurement, user-level measurement and user-centric measurement;
- Audience metrics (ratings, CUME, AQH, average frequency of exposure);
- Statistical inference, sampling and confidence intervals;

⁹⁰ Holden, Reed, and Thomas Nagle. *The Strategy and Tactics of Pricing: A Guide to Profitable Decision Making*, 3rd ed. Hoboken: Prentice Hall, 2001.

- Correlation coefficient;
- Econometric demand estimation;
- Epidemic models of diffusion;
- Delphi and comb analysis;
- Audience model-building;
- Feature extraction (e.g. principal component analysis);
- Clustering;
- Conjoint analysis;
- Attribute importance;
- Anomaly detection;
- Nielsen and Arbitron methodologies;
- POS measurement into issues.

8. Propose how to conduct demand experiments for a new video game console.
9. If a magazine publisher considered publishing a new business magazine for college students, describe techniques to see if there was a demand for such a magazine. What would be the pros and cons of each technique?
10. How can a website company measure the audience engagement with its product? What are the advantages and disadvantages of various techniques?

9.5.1 Questions for Discussion

1. Since the 1980s, media technology has led to a fragmentation of audiences on many levels. Is there still something that can be called “the mass audience”?
2. *Fashion and Food* is a new magazine on upscale shopping. How can one determine who will be interested in the magazine?
3. How can Internet research enhance a film distributor’s knowledge about the audience and demand for a particular movie?
4. What kind of data is needed to enable a media firm to use regression analysis to predict demand and potential sales for a mobile phone? A book?
5. How can qualitative audience measurements help television broadcasters?
6. Describe how a laptop manufacturer would use the conjoint analysis technique to design and market its product. At which stages should this technique be used?
7. Compare the usefulness of a Delphi survey with a survey of the general public. When would you use each one?

9.5.2 Quiz

1. Which of the following media require the least participant collaboration for audience research?
 - A. Radio.
 - B. Internet.
 - C. Television.
 - D. Books.
2. A banner advertisement is a type of:
 - A. Active advertising exposure.
 - B. Passive advertising exposure.
 - C. Click through.
 - D. Target advertisement.
3. Which of the following is *not* a tool to forecast demand for new products?
 - A. Consumer surveys.
 - B. Producer surveys.
 - C. Comb analysis.
 - D. Z-score.
 - E. Historical analogy test.
4. What is an audience “share”?
 - A. The percentage of time each family member spends watching TV.
 - B. The percentage of TV sets in use (or persons viewing) tuned to a program.
 - C. The number of TVs in a household.
 - D. The amount of time spent watching a TV program.
 - E. The percentage of TV sets per town.

- 9
5. When should management utilize the Delphi methodology?
 - A. To predict the future of a product.
 - B. To compare purchase criteria with the producer's opinions.
 - C. To study "network effects."
 - D. To gain the opinions and suggestions of experienced experts on audience needs, attitudes and priorities.
 6. What method should management use to compare its products with those of a competing firm?
 - A. Metcalfe's law.
 - B. Log-file analysis.
 - C. Comb analysis.
 - D. Audience metrics.
 7. What should the management of a motion pictures firm ask itself in order to estimate demand for a movie?
 - A. Does the genre cater to overlapping audiences?
 - B. What was the demand for movies with the same actors, genre, or director?
 - C. How many times was the movie trailer viewed via the Internet?
 - D. What is the lifecycle for similar movies?
 - E. All of the above.
 8. Which of the following is *not* a problem of user-centric Internet audience measurement?
 - A. Click fraud.
 - B. Meters are "black boxes" and require industry validation.
 - C. Small sites are at a disadvantage.
 - D. Poor site diagnostics.
 9. The management of a new magazine would most likely use which of the following techniques to measure its circulation?
 - A. Folio 400 reports.
 - B. Direct mail test.
 - C. Self-reporting.
 - D. Surveys.
 10. Why is audience research important for media firms?
 - A. To know how many people are utilizing their products and services.
 - B. To know who and how many people are reached by their products and services, and what the audience actually wants.
 - C. To estimate future profits.
 - D. To study audience demographics.
 - E. All of the above.
 11. What are the problems of econometric demand estimation?
 - A. Insufficient and unreliable data.
 - B. Need to assume a specific mathematical model for relationship between price and sale.
 - C. Predicting the future requires assumption that behavior is the same as in the past.
 - D. If specification is incorrect, the results will be incorrect.
 - E. All of the above.
 12. Which statement is right about the problems with econometrics models?
 - A. The models estimate elasticities exactly.
 - B. Econometric models are good at forecasting future performances.
 - C. Results rarely turn out to be statistically significant.
 - D. Econometric models often face methodological problems of how to interpret the results.
 13. Which of the following is not a reason why advertisers distrust site-centric data?
 - A. No demographic profile information is given.
 - B. All page requests, regardless of whether from a work-based or a home-based computer, are included.
 - C. Pages sent from web servers are not necessarily received.
 - D. Returning users are counted.
 - E. A, C and D.

14. Under what circumstances would one apply a conjoint analysis?
- If the company wants to predict the overall brand perception and acceptance based on familiarity and how much the brand is liked.
 - To identify the potential market for a product and contrast the position of competitive products according to their target demographics.
 - To illustrate and predict how a new product will be accepted by the population.
 - If the company wants to vary the product attribute mix according to estimated customer preferences in order to add value to the customer perception of the product.
15. What is the correct order of the data use methodology?
- Data collection -> Problem definition -> Data analysis -> Knowledge deployment;
 - Problem definition -> Data collection -> Data analysis -> Knowledge deployment;
 - Data collection -> Problem definition -> Knowledge deployment -> Data analysis;
 - Knowledge deployment -> Problem definition -> Data collection -> Data analysis.
16. Which of the following is a characteristic of the Conjoint analysis?
- It assumes a rational buying process where a typical buying experience constitutes of aggregating the value of individual attributes to make a purchasing decision.
 - It is calculation-intensive.
 - It is a market research tool.
 - It predicts the prices which the consumer would pay for different attribute combinations.
 - All of the above.
17. What are challenges in audience and market research?
- Different data privacy laws in different countries.
 - Applying and using data resources by media managers.
 - An increasing fragmentation of audiences.
 - All of the above.
18. What data mining technique refers to segmenting data by category?
- Clustering.
 - Organizing and classifying data.
 - Attribute importance.
 - Feature extraction.
19. If two filmmakers would want to fine tune their movie, what data collection would they be most likely to use?
- In-person surveys.
 - Production focus groups.
 - Marketing focus groups.
 - A laboratory experiment.
20. What is Business Intelligence (BI)?
- BI refers to the process of organizing raw data in a manageable way and to transform it into useful information.
 - BI is the analysis of organized information and transforming this information into knowledge.
 - BI describes qualitative data analysis.
 - BI refers to the whole data analysis process that includes raw data transformation, further quantitative analysis tools mixed with qualitative ways of judgment.

Quiz Answers

- ✓ 1. B
- ✓ 2. B
- ✓ 3. C
- ✓ 4. B
- ✓ 5. D
- ✓ 6. C
- ✓ 7. E
- ✓ 8. C
- ✓ 9. B
- ✓ 10. E
- ✓ 11. E
- ✓ 12. D
- ✓ 13. E
- ✓ 14. D.
- ✓ 15. B
- ✓ 16. E
- ✓ 17. D
- ✓ 18. B
- ✓ 19. B
- ✓ 20. A