

GARY ANGEL

MEASURING THE DIGITAL WORLD

USING DIGITAL ANALYTICS TO DRIVE
BETTER DIGITAL EXPERIENCES

Measuring the Digital World

Using Digital Analytics to Drive Better
Digital Experiences

Gary Angel

Publisher: Paul Boger
Editor-in-Chief: Amy Neidlinger
Executive Editor: Jeanne Glasser
Development Editor: Natasha Wolmers
Cover Designer: Alan Clements
Managing Editor: Kristy Hart
Senior Project Editor: Lori Lyons
Copy Editor: Hansing Editorial
Proofreader: Debbie Williams
Senior Indexer: Cheryl Lenser
Senior Compositor: Gloria Schurick
Manufacturing Buyer: Dan Uhrig

© 2016 by Pearson Education, Inc.
Publishing as FT Press
Upper Saddle River, New Jersey 07458

For information about buying this title in bulk quantities, or for special sales opportunities (which may include electronic versions; custom cover designs; and content particular to your business, training goals, marketing focus, or branding interests), please contact our corporate sales department at corpsales@pearsoned.com or (800) 382-3419.

For government sales inquiries, please contact governmentsales@pearsoned.com.

For questions about sales outside the U.S., please contact international@pearsoned.com.

Company and product names mentioned herein are the trademarks or registered trademarks of their respective owners.

All rights reserved. Printed in the United States of America. This publication is protected by copyright, and permission must be obtained from the publisher prior to any prohibited reproduction, storage in a retrieval system, or transmission in any form or by any means, electronic, mechanical, photocopying, recording, or likewise. For information regarding permissions, request forms, and the appropriate contacts within the Pearson Education Global Rights & Permissions Department, please visit www.pearsoned.com/permissions/.

Printed in the United States of America

First Printing December 2015

ISBN-10: 0-13-419508-6

ISBN-13: 978-0-13-419508-7

Pearson Education LTD.

Pearson Education Australia PTY, Limited.

Pearson Education Singapore, Pte. Ltd.

Pearson Education Asia, Ltd.

Pearson Education Canada, Ltd.

Pearson Educación de Mexico, S.A. de C.V.

Pearson Education—Japan

Pearson Education Malaysia, Pte. Ltd.

Library of Congress Control Number: 2015951825

*We all need a little pushing sometimes.
This book is dedicated to Grace, Isabella, and Ilise
(my wife and daughters) who, knowing I wanted to,
bugged me until I agreed to write a book.
They will probably make fun of my picture on the book jacket,
but given that it is about digital analytics,
none of them will likely ever read it.
Yet without them, it would never have been written.*

Contents

Preface	Measuring the (Digital) World	ix
Chapter 1	Digital Meaning	1
	The Digital Challenge: Our Metrics and Our Measurement Lack Meaning.	1
	The Grocery Store with Invisible Patrons	1
	Getting Digital	4
Chapter 2	Two-Tiered Segmentation	11
	Creating a Foundation for Digital Measurement	11
	Two-Tiered Segmentation	21
	Signal to Noise: Another View on Two-Tiered Segmentation	26
	Understanding Two-Tiered Segmentations.	29
	Sample Two-Tiered Segmentations.	32
Chapter 3	Use Cases and Visit Intent.	39
	Angel's Taco Divina Food Trucks: Designing a Digital Segmentation.	41
	Building Segmentation Rules.	46
	Developing Initial Use Cases.	51
	Creating Behavioral Signatures	56
	Segmentation	63
	Remainder Analysis	71
	Defining Success.	74
Chapter 4	Customer Identity and Taxonomy.	87
	Deepening Our Understanding of Behavior	87
	Functionalism	101
	Building Taxonomies	114
Chapter 5	Website Structure	119
	Postulate 2: Postulate 1 Is Often Wrong	119
	No New Problems—Just Old Mistakes	125
	Selection Limitation	126

	Behavioral Shifting	131
	Controlled Experimentation.	133
	Mixing Your Own Experiments	139
Chapter 6	Attitudes and Behaviors: Mixing a More Powerful Measurement Cocktail.	145
	Understanding Site Behavior	147
	Coloring a Segmentation	149
	Validating Segmentations	151
	Building a Segmentation	154
Chapter 7	Voice of Customer, Digital Marketing, and Success Measurement	165
	Digital Website to Retail Location.	167
	Acquisition Online and via Call Center	167
	Online Broadcast and Show Consumption	168
	Online Lead Generation and Sales Conversion	169
	Product Support with Commerce Elsewhere Online or Offline	169
	Pharma Company Marketing a Drug to Potential Patients.	170
	Sampling and Bias.	171
	Ask Before, Ask During, Ask After	175
	Measuring the Upstream Journey	176
	Targeting Precision.	179
	Offsite Survey	182
	Measuring Onsite (During)	184
	Stability	188
	When to Ask: The Role of Targeted Surveys and Samples	189
	Benchmarking and Cross-Site Comparison.	191
	Measuring Downstream Value (Ask After)	194
	Resurvey Techniques	196
	Sampling Challenges.	200
Chapter 8	Big Data and Measuring the Digital World	203
	What's Really Different about Analyzing Digital Data.	207
	Big Data Use Cases.	216
	Why Everyone Gets It Wrong	218
	Getting by with What You Have	220

- Chapter 9 Omnichannel Analytics 223**
 - The Myth of the Golden Path 225
 - The Implications 228
 - Omnichannel Analytics: Journey Stage 229
 - Building the Measurement Framework for
 - Tracking Journeys 233
 - Recency/Frequency/Monetary (RFM) 237

- Chapter 10 No End in Sight 241**
 - Reviewing the Framework 243
 - Some Humble Time: All Things Not Covered Here 248
 - Looking Forward: Where There’s Still Work to
 - Be Done 251
 - Measurement and All It Means 252

- Index 255**

Acknowledgments

There are countless people who have contributed to and helped shape my understanding of digital measurement over the course of almost two decades. There are too many to count and probably too many to remember. I can't not mention the small group of people who helped me build Semphonic, the digital analytics practice where nearly all of this thinking grew up and matured. Every person on that team is special to me. They helped build a company and practice that proved to be the central endeavor of my adult life, and almost without exception they have been good friends as well.

Outstanding among that truly wonderful group is, of course, my original partner and co-founder in that endeavor, Joel Hadary. No matter the intellectual interest of a journey, it's the people you travel it with that make it enjoyable. It's hard for me to imagine a better partner for building a company and creating a discipline.

Three early employees who have been with me now on this journey for more than a decade must also be mentioned. Paul Legutko, Phil Kemelor, and Jesse Gross have all contributed mightily not just to the practice we built, but also to the thinking and approaches laid out here. Paul has shaped so much of my thinking on analytics—particularly the problems of segmentation, selection, and bias—that's it hard for me to know where my thinking starts and his ends. To Phil I owe much of the broader picture that makes analytics more than number crunching and focuses it on people and problems and how to solve them. Jesse—who began with us right out of college knowing no more of digital analytics than I know of Sanskrit—has been at least a full partner in the exploration of voice of customer and its role in digital measurement.

The voice may be mine, but the ideas are, as often as not, theirs.

When I started in digital analytics, it was a tiny community. Like most tiny communities, its members were oddballs. The best kind of eccentrics. Lacking credentials and degrees, digital analytics was peopled by those willing to explore a discipline that offered neither safe position nor any clear path to success. That's often the best kind of people. It has been both a pleasure and privilege to grow up in that community. To argue, to fight, to learn, and to share with a group of people who actually love what they do and strive to do it better.

Thank you all.

About the Author

Gary Angel currently leads Ernest & Young's (EY) Digital Analytics Practice. EY acquired Gary's previous company, Semphonic, in 2013. As President and co-founder, Gary led Semphonic's growth from a two-person consultancy to one of the leading digital analytics practices in the United States. Voted the most Influential Industry Contributor by the Digital Analytics Association, Gary writes an influential blog (<http://semphonic.blogs.com/semangel>), has published numerous whitepapers on advanced digital analytics practice, and is a frequent speaker at industry events. Over the last two decades, he's helped create and advance the state-of-the-art in digital measurement.

Preface

Measuring the (Digital) World

How do we understand the world around us? It's not so obvious.

Our senses feed a never-ending stream of data into our brains. In its raw form, that data is incredibly complex. Shapes, surfaces, sizes, shades, motions, scents, textures—we absorb them all, seemingly without effort. Even the most powerful supercomputers yet developed cannot begin to compete in real time. But the plastic mind of human babies—with the help, no doubt, of some careful pre-wiring—can learn how to make sense of this data, parse it, and react swiftly and intelligently to this tidal surge of information.

For centuries, philosophers have understood that the very existence of a world outside ourselves is impossible to verify. We know the world only from the inside, from the endless, constant processing we call consciousness and the apparent flow of data that we believe is generated by our eyes, ears, nose, and fingers. That such a thing as the physical world exists is beyond our ability to seriously doubt. But we know it only by interpreting essentially abstract data.

We grasp the world by imposing patterns on it—foundational patterns that have been imprinted and wired deep into our minds in the endless laboratory of nature. Size, shape, color, and motion are just a few of the core building blocks of our human understanding of the physical world. As fast as sense data funnels into the brain, we are contextualizing it, categorizing it along key dimensions, and then measuring and comparing everything.

That's the physical world. Evolution ensures that we know how to understand it, even if we don't know how we know. But the physical world is no longer the place where we spend all our time. We

also live in the digital world, a world with different rules, different types of data, different frames of reference, and different types of measurement.

For the last 17 years, my job has been to measure the digital world. To glean, from the vast streams of data it showers upon us, the fundamental categorizations that matter. To develop the framing devices, dimensions, and measurements that let us understand this digital world with the same ease and power with which we parse the physical world around us. It's important work because we can shape and improve that world only when we understand how it works—not how it works from a programming perspective (although I started out as a programmer), but how it works for people.

These days, we spend immense amounts of time, energy, and money trying to improve the digital world. How well we do that work can determine the success and direction of public policy (healthcare.gov), the health of our love life (eHarmony), and the state of our knowledge about the world (nytimes.com).

This work isn't easy, and it isn't finished. We are still like newborn babes learning to parse the data from our digital sensors.

The digital world is fascinating. And unlike in the physical world, we have no pre-wiring for measuring and understanding the digital realm. Without that huge prebuilt advantage, our interpretations are wrong much more often than they are right. But when we get them right, we at least know what we did.

The pages that follow show you what we've learned so far about how to measure and understand the digital world.

1

Digital Meaning

We have tools dedicated to measuring the digital world. So it's no surprise that we assume the measurements those tools give us are the right ones for the job. They aren't. The standard set of web metrics most digital analytics tools use were developed long before people had even a basic understanding of how to do digital measurement, and mostly before analytics tools that could do much with the data were in widespread use.

The Digital Challenge: Our Metrics and Our Measurement Lack Meaning

The most common digital metrics are almost useless. They measure the wrong things in the wrong ways. They fail, at the most basic level, to link what happens in the digital world to our understanding of people's behavior. In this chapter, you identify the basic challenge of digital measurement and analytics, and you see why common metrics and reports can't easily answer the fundamental questions you're likely to ask about the digital world.

The Grocery Store with Invisible Patrons

Imagine a fairly normal grocery store, well stocked with cereal, milk, beer and wine, eggs, ice cream, canned goods, vegetables, fruits,

and, of course, the usual assortment of treats near the register. Now imagine that the patrons and their carts are invisible. You see the door swing open when they arrive. You hear the cash register ring when they depart. You know what they bought. But everything else remains hidden. It would be hard to know how well the store was working and what you could do to make it better. Is it missing items or brands shoppers want? Is the store laid out in a way that makes life easy on customers? Does it maximize their purchasing behavior? Have you allocated the right amount of shelf space to each type of item? What might you do to get an individual customer to spend more or be more loyal?

These are the types of questions that merchandising experts have studied, pondered, and worked on for many years—since well before the digital world ever existed. Interestingly, they found that they could answer some of these questions even when the customers were, for all practical purposes, invisible. Equally interesting, they found that some types of questions are much harder to answer when you don't know who your customers are and that, for many questions, the data might suggest possible answers but rarely provides definitive guidance.

Suppose, for example, that you found that the most purchased items in your store were milk, beer, eggs, and chips. You might be tempted to move all these items together in one place right at the front of the store. That should make it easy for customers to find what they need quickly and efficiently. Is that the way your supermarket is laid out, with the items you buy most right up front?

Chances are, it's almost exactly the opposite. That isn't because you're invisible! Supermarkets work differently for two reasons. We're all deeply cynical consumers, so you probably identified the first reason right away. Groceries aren't set up for your convenience. They often place the things people purchase most at the very back of the store and might even consciously try to locate them far apart. If you've never made an impulse buy at a grocery store, this might seem

odd. But if, like me, you've wandered by the dairy aisle and added some ready-to-bake cookies, or you've thrown a bag of chips next to your beer, it's not too hard to see why this setup works. By trying out different store layouts and measuring how much people buy (their average cart), store designers can maximize total sales. Mind you, most grocery stores count on you to make your decision about where to shop based on other factors than how long it takes you to get your items. They know price, selection, and location are more important than convenience. If a new grocery store opened right next door and had the same selection and same prices, a store might well compete on the convenience of layout. But most stores see their layout as a chance to maximize their profits, not your time.

The second reason grocery stores aren't laid out for your convenience is more interesting and more important than good old profit maximizing. Grocery stores have more than one customer. Guess what? They're all different. When grocery merchandisers began to study what people bought (still without knowing who they were—only what was purchased on the same ticket), they found very distinct patterns. Beer and milk might be two of the most commonly purchased items in a grocery store, but they might not often be purchased together. Chips, on the other hand, go pretty well with that beer. And milk buyers are often looking to add cereal or eggs to their cart. So setting up a grocery with the most purchased items all clustered together might not work particularly well or be particularly convenient for anyone.

What's more, even if a particular setup worked well for you today, it might not tomorrow. When merchandisers could only look at the receipts from each shopper, they had no way to tell how much people's habits and shopping patterns varied. That's a huge hole in their understanding. To get around that, grocery stores created loyalty programs so that, in return for discounts, they could tell what you bought every trip. They found that most people don't shop the same way every time they visit the grocery store. Most of us have regular shopping

expeditions when we buy everything we need and go up and down every aisle. Store layout might not be a big deal when we're traversing every inch of the store (or, in my case, traversing two or three times as I remember things). But we also have visits when we've just run out of beer or milk—or, heaven forbid, both. We might stop to pick up lunch or to shop for specific recipe ingredients (my flour, bag of chocolate chips, vanilla extract, and egg visits). These are very distinct types of visits, and it would be great if the grocery store could make each type of visit perfect (or perfect for the grocer). Stores would love to be able to do that. But it's hard to push those shelves around when you walk in the door.

Let's not forget those chocolate bars and women's magazines perched right at the register. Very few of us go to the grocery store with the express intent of buying a Snickers bar and a *Cosmo*, but many of us are tempted by one or the other. What spot in the grocery store do people *have* to linger at with nothing to do but be tempted? That's where the candy (eye and stomach) goes.

We can learn a lot from those grocery store merchandisers when we start to think about the digital world. The straightest path isn't always the best. The customer's goals and your goals aren't always identical. Not every product is the same, and some products are more position sensitive than others. A store doesn't have one ideal layout because it doesn't have one type of customer, and customers aren't always going to do the same thing anyway. Last, and most important, what people actually do tells us a great deal about who they are and why they are doing those behaviors.

Getting Digital

We are blind to the digital world. Unassisted, we have no way of knowing whether our website is thronged with visitors or as empty as a mall after closing time, whether our cash registers are overflowing

or stubbornly silent, whether our customers are young or old, whether our content is read with rapt attention or is barely and desultorily skimmed. We need eyes and ears to help us see into the digital world. Certain tools have that very function—to track and make visible the otherwise unseen patterns of that world. These digital analytics tools are powerful and rich. They include hundreds or thousands of possible reports and options that seem to expose every aspect of digital behavior. It's all too easy to forget how dependent we are on the exact nature of those tools and to assume that what they show us and the way they show it to us is all there is.

Our natural senses in the physical world have evolved to give us many advantages. We have adopted a deep and abiding faith in what we *see*. Yet even with our physical senses, it's all too easy to forget that the window they provide into the world is a narrow one.

Remember the image of a dress that went viral in early spring 2015? Many people see the dress as black and blue. Others see it as white and gold. If you look long enough or over some period of time, you might see it each way. If you didn't hear about the dress and you can't believe that anyone could see it differently than whichever way happens to strike you, check it out online and show it around. You'll be surprised.

Optical illusions are just one aspect of how our eyes can mislead us. We see color (no matter how much we might disagree about it), but we don't see heat.

Why should we see heat?

Well, why shouldn't we?

Infrared cameras see heat. It's just another wavelength, and for many purposes, seeing heat is far more useful than seeing light (when hunting at night, for example). For that matter, what if we could see radio waves? Hearing and vision seem fundamentally different to us, but each is a set of waves that different tools inside our body use.

What would our radio-wave eyes make of a Madonna song? Probably not much.

The simple fact is this: Our reality is constrained by the tools we experience it with.

What does all this have to do with digital analytics? The digital analytics tools we have are our windows into the digital world. We see only what they can track or think is important. We see what page a user requested from a server, but we don't see how long that page took to load. We see what link a user clicked on, but we usually don't see what part of the page that user scrolled to. We (sometimes) see what website the user came from, but we don't (usually) see what website that user went to. These choices make a profound difference in how we think about the digital world and what we tend to value as important there.

What if our tools aren't very good? What if the events they choose to capture or the ways they choose to show them to us give us only a shadowy impression of the real digital world or what's important inside it?

I've been around since the very beginning of digital analytics. I witnessed firsthand and, in some small ways, even helped shape how those digital tools evolved. Having seen their history, I know that the decisions about what to track in the digital world and how to track it were often ad hoc and shallow.

The first digital analytics tools were built to read weblogs. These logs weren't built to understand and measure the digital world. They were built to create a record of what a web server was doing so that IT professionals might be able to track down operational problems (although they were hardly ever used for that, either). These logs recorded IT-focused information about which content file was requested, when exactly it was processed, what IP address requested it, how much content was sent, and whether the request was successful.

Because those were the fields in the logs, those were the fields we used when we first built digital analytics tools. And being clever folk, we interpolated a whole lot from this bare bones little set of fields. We figured out a way to group the records by the device requesting them (which we promptly anthropomorphized into a human *visitor*). By looking at the time between requests, we could group the requests into batches by creating an arbitrary time limit, and we labeled these batches of requests *visits*. Then we could look at what page a visitor looked at first in that batch and we called that an entry page. We could also look at what page was last in the batch and call that an exit page.

It's important to realize just how arbitrary these decisions were. When a visitor first arrives on a website, that website sometimes records which website the visitor came from—this is called the referring site. By saving the referring site for each batch of records (a visit), you can get a sense of which sites are generating traffic to your pages. But here's a peculiarity: By defining an arbitrary time limit of 30 minutes to group records, we created situations in which a visit sometimes had more than one referring site; in other situations, a visit had a referring site that was the last page the visitor viewed on the same website.

For example, imagine that a visitor searches on Google or Bing, finds your website, and views a page. Then that visitor returns to the search engine, does another search, goes to a different site, and links from there to you within 30 minutes of the first request. You'll have a single visit with two referring sites. This might sound far-fetched, but in certain permutations, it's not uncommon. Many sessions will have multiple visits to Google interspersed with views of your pages.

It's even more likely, especially in our tabbed browser world (which came after all these definitions were created—you remember browsers without tabs, right?), that a visitor will view a page or tab elsewhere, spend some time there, and then return to your website and view another page. Same session? According to our tools, if that happens 25 minutes apart, it is. But if it happens 31 minutes apart, it

isn't. And if it does happen 31 minutes apart, you'll have a whole new visit with a referring site of your own website!

It would have been perfectly plausible (maybe much more plausible) to decide that a batch of records should be separated by a referring site other than your own domain, regardless of time. But that's not the way some early vendors did it, so the definition stuck and became an artifact of truth.

And if a visit is merely a rough-and-ready and poorly defined artifact, then so are the entry page (the first page in a visit), the exit page (the last page in a visit), the visit time (the time between the first and last requests that are part of a visit), and the referring site (the domain recorded in the first record of the visit as the referrer, the site from which the user came)—all based on the way we defined a visit.

As with the words we read on a page, the numbers we see in a tool tend to take on privileged status in our minds. But if the only problem with standard web metrics were a certain sloppiness of definition, our situation wouldn't be all that bad. How much difference can it make whether a visit is defined by 30 minutes of inactivity or a new referring domain? Hard to say.

By showing how arbitrary these standard metrics are in construction, I hope to lessen their privileged status and make it easier to convince you that, not only are they arbitrary, but they are largely misguided.

Web measurement began with weblogs, whose goal was to measure digital assets. This long-ago bias has persisted through every generation of digital analytics tool. The implicit goal of analytics tools is to measure the website or the app. That's missing, if not the whole point, a big part of it. In the digital world, our goal should be to understand our customers, not our digital assets.

Our tools have improved to do that—probably more than our practice has. Digital analytics tools now deliver significant and interesting segmentation capabilities that enable you to define and track

cohorts, segment on fairly complex behaviors, and compare different types of users. They even provide limited capabilities for integrating nondigital data into their reporting.

This is all good, and the technology seems to improve continually. But although the capabilities of the tools have improved, the basic views they provide haven't changed much. How many reports in a digital analytics tool tell you anything about customers? For that matter, how many of the digital reports you distribute in your organization have anything to do with customers? And how much do they really help you understand the digital world?

Close your eyes and picture your website. Imagine people of all sorts moving through it. They stop here or there. They go down certain pathways and ignore others. They look at this or that. They make a purchase or head for the exit. Can you see it?

Now open your web analytics tool. Do the reports help you visualize that scene? Do they help you understand who your customers are (beer and chips, or milk and eggs)? Can you find the different types of visits (going down every aisle, or just ran out of something) and see which are most common? Do they help you picture which customers do which visits most often (beer and chips, just ran out, Friday night)? In other words, do they actually help you understand the digital world or do they just confront you with a wall of numbers that elude meaning, even while seeming entirely plausible?

I first started measuring the digital world back when websites were brand new and people still talked about the World Wide Web. I'd spent the previous few years working with a couple large credit card companies, analyzing the way people use their credit cards to create marketing programs (yeah, sorry about all that crappy mail). We used some pretty fancy analysis techniques to group people together, to understand how they used their credit cards and then to classify them. It was pretty easy and powerful. It didn't take analytic genius to know that the cardholder who routinely dropped four figures at

Neiman Marcus was a different beast from that two-digit shopper at the local Walmart.

When I first got my hands on web behavioral data, I ran the same kind of (neural net) analysis and proudly sold the results. But whereas my credit card segmentations had truly been interesting and useful, my digital segmentations looked like some inverted Egyptian monstrosity (see Figure 1.1).

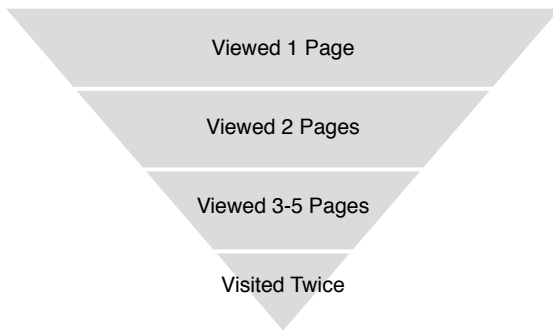


Figure 1.1 Inverted pyramid

Nobody ever got rich targeting “people who viewed 3–5 pages.”

I spent years learning that the digital metrics tools provide aren’t that interesting and that, no matter how powerful a tool I used to study digital behavior, I wouldn’t get interesting results if I picked the wrong type of variables.

So put away your digital analytics tools for a minute. Forget all about page reports, referring sites, average page times, top exit pages, number of visits, average conversion rates, and the whole flavorless cornucopia of web metrics and reports that those tools spit out by default. It’s all garbage in the most literal sense—it takes up mental space and it smells bad.

You’re about to find a better way to understand the digital world.

A

analysis
 experimentation versus, 143
 traditional versus digital data analysis, 207-215

analytics tools. *See* digital metrics

Angel's Taco Divina (ATD) food trucks example, 41-46

arrival articles (page type), 108

articles (page type), 108

ask before, ask during, ask after (VoC strategy), 175-176, 247

 offsite surveys, 182-184

 online surveys
 benchmarking and cross-site comparison, 191-194
 stability of, 188-189
 targeted surveys, 189-191
 types of questions in, 184-188

 prequalification measurement, 176-179

 resurvey techniques, 196-200
 sampling bias in, 200-202

 success analysis, 194-196

 targeting precision, 179-182

ATD (Angel's Taco Divina) food trucks example, 41-46

audience taxonomies, 94-95

B

A/B tests. *See* controlled experimentation

background of digital metrics, 241-243

before-and-after analysis, 137-138

behavior. *See also* visit intent

 mapping, 122-124

 predicting, 90-91

 website structure and, 119-125

behavioral proxies in downstream value measurement, 199-200

behavioral segmentation, 245-246. *See also* voice of customer (VoC)

 building segmentation rules, 154-160

 closed list versus open-ended questions, 147-149

 coloring with VoC data, 149-151

 future of, 252

 in omnichannel analytics, 235

 validating with VoC data, 151-153

behavioral shifting, 120-122, 131-133

 controlled experimentation and, 133-139

behavioral signatures, creating, 56-63

benchmarking, 191-194

bias

 online surveys and, 171-175

 in resurveys, 200-202

big data

 digital data as, 203-207

 Four V's, 205-207, 218-220

 lack of measurement tools, 220-221

 in omnichannel analytics, 235-236, 248

 use cases, 216-218

brand engagement in success criteria, 76-79

brand knowledge of users, 54

brand value questions (online surveys), 186

broadcast shows, streaming shows versus, 168-169

brokerage sites (two-tiered segmentation), 32

C

call center conversions, website visitors and, 167-168

campaigns, optimizing, 179-182, 249-250

car analogy (functionalism), 101-103

card files (page type), 108
 CEM (customer experience management) tools, 128-129
 channel sales, product support sites and, 169-170
 classifications, 91-93, 244-245. *See also* segmentation
 collecting metadata, 114-117
 strategies for, 93-96
 closed list questions, open-ended questions versus, 147-149
 closed-loop problems in success criteria, 81-85
 closers (page type), 105
 clustering in behavioral segmentation, 155-157
 CMSs (content management systems), creating taxonomies, 114
 coloring behavioral segmentation with VoC data, 149-151
 comment cards, 129-130
 competitive positioning questions (online surveys), 186
 completers (page type), 106-107
 Conan the Librarian analogy, 87-90
 conflating use cases, 49-50
 consumption patterns, 53
 content descriptions
 for page components as metadata, 97-98
 taxonomies, 91-93
 classification strategies, 93-96
 content management systems (CMSs), creating taxonomies, 114
 content sites (two-tiered segmentation), 33
 controlled experimentation, 133-139
 variables to hold constant, 139-143
 conversion rate
 failures in digital metrics, 14-15
 signal to noise, 28
 converters (page type), 106
 convincers (page type), 105
 core demographic and life stage questions (online surveys), 187
 correlations in mapping behavior, 122-125
 cost-per-click, 195
 cross-site comparison, 191-194
 customer experience management (CEM) tools, 128-129

customer service sites (two-tiered segmentation), 37
 customers. *See also* voice of customer (VoC)
 engagement life-cycle, 29-31
 lifetime value (LTV), 84-85
 RFM (recency/frequency/monetary), 237-240
 signal to noise, 26-29
 two-tiered segmentation
 elements of, 21-26
 samples of, 32-37
 visit intent, 31, 244
 building segmentation rules, 46-51, 157-160
 closed list versus open-ended survey questions, 147-149
 Conan the Librarian analogy, 87-90
 creating behavioral signatures, 56-63
 designing use cases, 41-46
 grocery store analogy, 1-4
 initial use cases, 51-54
 limitations of segmentation tools, 63-65
 mutual exclusivity of use case, 65-69
 natural world versus digital world, 39-40
 in omnichannel analytics, 229-232
 remainder analysis, 71-74
 success criteria for segments, 74-85
 validation of segmentation rules, 70-71
 website structure and, 119-125

D

dashboards (page type), 108
 data of week in use cases, 54
 decision-making style questions (online surveys), 188
 deep navigation, selection limitation and, 128-129
 demand elasticity of use cases, 134-136
 designing use cases, 41-46

digital data
 as big data, 203-207
 Four V's of big data and, 218-220
 lack of measurement tools, 220-221
 traditional analysis versus, 207-215

digital metrics
 combining with non-digital
 channels, 166, 170-171, 246-247.
See also omnichannel analytics
lead generation and sales
conversion, 169
online and call center
conversions, 167-168
product support and channel
sales, 169-170
streaming versus broadcast
shows, 168-169
success via third-party, 170
website and retail location
sales, 167

difficulty of, 243-244
 failures in, 1, 11-21
 future of, 251-252
 grocery store analogy, 1-4
 history of, 241-243
 importance of, 252-254
 limitations of, 4-10, 248-250
 Document Object Model (DOM),
 creating taxonomies, 114
 dominant behavior, 43-44
 downstream value measurement,
 194-196
 resurvey techniques, 196-200
sampling bias in, 200-202

E

ecommerce, retail sales and, 167
 elasticity of use cases, 134-136
 emotional self-categorization
 questions (online surveys), 187
 engagement
 as behavioral proxy, 199-200
 in success criteria, 76-79
 engagement life-cycle, two-tiered
 segmentation versus, 29-31
 engagers (page type), 104
 entry pages, 7-8
 error theory, Four V's of big data and,
 218-220
 exclusivity of use cases, 65-69
 exit pages, 7-8

experience level of users, 54
 experimentation
 analysis versus, 143
 controlled experimentation, 133-139
variables to hold constant,
139-143

F

failures in digital metrics, 1, 11-21
 feature value questions (online
 surveys), 187
 first click, 43-44
 focus of users in use cases, 54
 Four V's in big data, 205-207, 218-220
 functional taxonomies, 94
 functionalism
 car analogy, 101-103
 importance of, 108-110
 page types, 104-108
 as segmentation cue, 113
 within use cases, 110-112
 website components, 103-104
 future of digital metrics, 251-252

G-H-I

gaming success metrics, 80-81
 goals in success criteria, 74-76
 GPS navigation analogy, 122-124
 grocery store analogy, 1-4

history of digital metrics, 241-243

informers (page type), 104-105
 initial use cases, 51-54
 in-store sales, website visitors and,
 167
 intent. *See* visit intent
 intercept surveys. *See* online surveys
 internal search analysis, selection
 limitation and, 126-128
 Internet of Things (IoT), 217

journey mapping, 223-225
 ask before, ask during, ask after
 (VoC strategy), 175-176, 247
benchmarking and cross-site
comparison, 191-194
offsite surveys, 182-184

prequalification measurement, 176-179
resurvey techniques, 196-200
sampling bias in resurveys, 200-202
stability of online surveys, 188-189
success analysis, 194-196
targeted surveys, 189-191
targeting precision, 179-182
types of questions in online surveys, 184-188

building use cases, 233-236
 limitations of, 166
 path analysis, 225-228
 touchpoint optimization, 228-229
 with VoC data, 171

journey stage of users, 54
 qualification levels in, 166

K-L

knowledge level of users, 54
 KPIs (key performance indicators)
 failures in digital metrics, 13-21
 signal to noise, 26-29

large segments, splitting, 47-49
 lead generation, sales conversion and, 169
 library analogy
 selection limitation, 120-121
 visit intent, 87-90
 lifetime value (LTV), 84-85
 list size, as metadata, 98-99
 look-to-book metrics, signal to noise, 27
 LTV (lifetime value), 84-85

M-N

manufacturing sites (two-tiered segmentation), 35
 mapping behavior, 122-124
 metadata, 91-93, 244-245
 classification strategies, 93-96
 collecting, 114-117
 elements of, 97-101
 metrics. *See* digital metrics
 multiple-page consumption, 78
 mutual exclusivity of use case, 65-69

navigational structure. *See* website structure
 noise. *See* signal to noise
 non-digital channels, combining with digital metrics, 166, 170-171, 246-247. *See also* omnichannel analytics
 lead generation and sales conversion, 169
 online and call center conversions, 167-168
 product support and channel sales, 169-170
 streaming versus broadcast shows, 168-169
 success via third-party, 170
 website and retail location sales, 167

NPS (net promoter score), failures in digital metrics, 15-18

O

offsite surveys, 182-184
 omnichannel analytics, 247-248
 building use cases, 233-236
 journey mapping, 223-225
 path analysis, 225-228
 touchpoint optimization, 228-229
 RFM (recency/frequency/monetary), 237-240
 two-tiered segmentation in, 229-232

one-and-dones group, 71-72
 online banking sites (two-tiered segmentation), 36
 online surveys, 130, 146-147
 benchmarking and cross-site comparison, 191-194
 building, 160-163
 closed list versus open-ended questions, 147-149
 offsite surveys, 182-184
 resurvey techniques, 196-200
 sampling bias in, 200-202
 sampling and bias, 171-175
 stability of, 188-189
 targeted surveys, 189-191
 types of questions in, 184-188

online visitors. *See* website visitors
 open-ended questions, closed list questions versus, 147-149

optimizing
 campaigns, 179-182, 249-250
 touchpoints, 228-229

P

page components
 content descriptions for, 97-98
 as metadata, 97
 page length, as metadata, 99-100
 page metadata. *See* metadata
 page taxonomies, 91-92
 page types in functionalism, 104-108
 path analysis, 225-228
 paths of navigation. *See* website structure
 pattern of events
 in big data use cases, 216-218
 in digital data, 212-215
 lack of measurement tools, 220-221
 in omnichannel analytics, 235-236
 pop-up surveys. *See* online surveys
 predicting behavior, 90-91
 prequalification measurement,
 176-179
 in downstream value measurement,
 200
 product knowledge questions (online surveys), 186
 product status, as metadata, 98
 product support sites, channel sales and, 169-170
 product taxonomies, 93
 proxies in success criteria, 81-85
 purchase funnel, two-tiered segmentation versus, 29-31
 push date, as metadata, 100

Q-R

qualification levels in journey stage of users, 166
 questions in online surveys
 stability of, 188-189
 types of, 184-188
 race car analogy (functionalism), 101-103
 reassurers (page type), 105
 recency/frequency/monetary (RFM), 237-240

referring site, 7-8, 45
 remainder analysis, 71-74, 236
 repeat consumption, 77
 response rates for online surveys, 162
 resurvey techniques, 196-200
 sampling bias in, 200-202
 retail sales, website visitors and, 167
 revenue, failures in digital metrics, 18-20
 RFM (recency/frequency/monetary), 237-240
 routers (page type), 104

S

sales conversion, lead generation and, 169
 sales stage taxonomies, 95
 sample size for online surveys, 160-162
 sampling
 bias in resurveys, 200-202
 changing nature of, 191-194
 online surveys and, 171-175
 scrolling, 78-79
 searches, internal search analysis, 126-128
 season in use cases, 54
 segment definitions in success criteria, 79
 segmentation. *See also* classifications; two-tiered segmentation
 behavioral segmentation, 245-246
building segmentation rules, 46, 154-160
closed list versus open-ended questions, 147-149
coloring with VoC data, 149-151
future of, 252
validating with VoC data, 151-153
 classifications and, 96
 hierarchy, 67
 in functionalism, 113
 in omnichannel analytics, 233-236
 selection limitation, 120-122, 126-131
 self-selection, 125-126
 controlled experimentation and, 133-139
 sensor data (big data use case), 216

sequence. *See also* journey mapping
 in digital data, 208-211
 lack of measurement tools,
 220-221
 in omnichannel analytics, 235-236
 session replay tools, 128-129
 signal to noise, 26-29
 site taxonomies, 92-93
 site topology. *See* website structure
 site walkthroughs, 52-53
 small segments, discarding, 47
 social media, selection limitation
 research, 130-131
 social sharing, 77-78
 social sites (two-tiered segmentation),
 34
 splitting large segments, 47-49
 streaming shows, broadcast shows
 versus, 168-169
 structure. *See* website structure
 subsetters (page type), 105
 success analysis
 combining digital and non-digital
 data, 170-171
 lead generation and sales
 conversion, 169
 in omnichannel analytics, 236
 online and call center
 conversions, 167-168
 product support and channel
 sales, 169-170
 streaming versus broadcast
 shows, 168-169
 via third-party actions, 170
 website and retail location
 sales, 167
 downstream value measurement,
 194-196
 success criteria for segments, 74-85
 success rates, 53-54
 surveys. *See* online surveys

T

targeted surveys, 189-191
 targeting precision, 179-182
 taxonomies, 91-93, 244-245
 classification strategies, 93-96
 creating, 114-117
 functionalism
 car analogy, 101-103
 importance of, 108-110

page types, 104-108
 as segmentation cue, 113
 within use cases, 110-112
 website components, 103-104
 third-party actions, success analysis
 via, 170
 thrashing sessions, 128-129
 threads (page type), 108
 thresholds in behavioral signatures,
 60-63
 time between events
 in digital data, 211-212
 lack of measurement tools,
 220-221
 in omnichannel analytics, 235-236
 time of day, in use cases, 54
 time on page, 45, 78-79
 time on streaming content, 78
 time-series analysis, 210
 tools, limitations of, 63-65. *See also*
 digital metrics
 topic taxonomies, 94
 topology. *See* website structure
 touchpoints
 in omnichannel analytics, 229-232
 optimizing, 228-229
 traditional analysis, digital data
 analysis versus, 207-215
 two-tiered segmentation, 244
 elements of, 21-26
 engagement life-cycle versus, 29-31
 journey mapping, 223-225
 limitations of, 166
 path analysis, 225-228
 touchpoint optimization,
 228-229
 limitations of, 145-146
 in omnichannel analytics, 229-232
 RFM (recency/frequency/
 monetary), 237-240
 samples of, 32-37
 brokerage sites, 32
 content sites, 33
 customer service sites, 37
 manufacturing sites, 35
 online banking sites, 36
 social sites, 34
 self-selection and, 125-126
 signal to noise, 26-29
 visit intent, 31, 244
 building segmentation rules,
 46-51

creating behavioral signatures, 56-63
designing use cases, 41-46
initial use cases, 51-54
limitations of segmentation tools, 63-65
mutual exclusivity of use case, 65-69
natural world versus digital world, 39-40
remainder analysis, 71-74
success criteria for segments, 74-85
validation of segmentation rules, 70-71

U

unit-of-work, 69
update date, as metadata, 100
upstream journey
 offsite surveys, 182-184
 prequalification measurement, 176-179
 targeting precision, 179-182
use cases
 for big data, 216-218
 building segmentation rules, 46-51
 conflating, 49-50
 creating behavioral signatures, 56-63
 designing, 41-46
 downstream value measurement and, 198-199
 elasticity, 134-136
 functionalism within, 110-112
 identifiers, 54
 initial use cases, 51-54
 limitations of segmentation tools, 63-65
 mutual exclusivity, 65-69
 for omnichannel analytics, 233-236
 remainder analysis, 71-74
 self-selection in, 125-126
 success criteria for segments, 74-85
 validation of segmentation rules, 70-71
user experience in online surveys, 162-163

V-W

validation
 of behavioral segmentation with VoC data, 151-153
 building segmentation rules, 46, 70-71
variables, holding constant, 139-143
variety (big data), 205-207, 219
velocity (big data), 205-207, 218-219
veracity (big data), 205-207, 219-220
visit intent, 31, 244
 building segmentation rules, 46-51, 157-160
 closed list versus open-ended survey questions, 147-149
 Conan the Librarian analogy, 87-90
 creating behavioral signatures, 56-63
 designing use cases, 41-46
 grocery store analogy, 1-4
 initial use cases, 51-54
 limitations of segmentation tools, 63-65
 mutual exclusivity of use case, 65-69
 natural world versus digital world, 39-40
 in omnichannel analytics, 229-232
 remainder analysis, 71-74
 success criteria for segments, 74-85
 validation of segmentation rules, 70-71
 website structure and, 119-125
visit time, 8
visitors. *See website visitors*
voice of customer (VoC)
 ask before, ask during, ask after strategy, 175-176, 247
benchmarking and cross-site comparison, 191-194
offsite surveys, 182-184
prequalification measurement, 176-179
resurvey techniques, 196-200
sampling bias in resurveys, 200-202
stability of online surveys, 188-189
targeted surveys, 189-191
targeting precision, 179-182
types of questions in online surveys, 184-188

- building segmentation rules, 154-160
- coloring segmentation, 149-151
- future of, 252
- journey mapping, 171
- omnichannel use cases, 234-235
- online surveys, 146-147
 - benchmarking and cross-site comparison, 191-194*
 - building, 160-163*
 - closed list versus open-ended questions, 147-149*
 - sampling and bias, 171-175*
 - stability of, 188-189*
 - targeted surveys, 189-191*
 - types of questions in, 184-188*
- selection limitation and, 129-131
- validating segmentation, 151-153
- when to use, 165
- volume (big data), 205-207, 218**

- website components, 103-104**
- website structure, 119, 246**
 - behavior and, 119-125
 - behavioral shifting, 131-133
 - controlled experimentation, 133-139
 - selection limitation, 126-131
 - self-selection, 125-126
- website visitors. *See also* voice of customer (VoC)**
 - call center conversions and, 167-168
 - channel sales and, 169-170
 - engagement life-cycle, 29-31
 - lead generation and sales
 - conversion, 169
 - retail sales and, 167
 - signal to noise, 26-29
 - streaming versus broadcast shows, 168-169
 - two-tiered segmentation
 - elements of, 21-26*
 - samples of, 32-37*
 - visit intent, 31, 244
 - building segmentation rules, 46-51, 157-160*
 - closed list versus open-ended survey questions, 147-149*
 - Conan the Librarian analogy, 87-90*
 - creating behavioral signatures, 56-63*
 - designing use cases, 41-46*
 - initial use cases, 51-54*
 - limitations of segmentation tools, 63-65*
 - mutual exclusivity of use case, 65-69*
 - natural world versus digital world, 39-40*
 - in omnichannel analytics, 229-232*
 - remainder analysis, 71-74*
 - success criteria for segments, 74-85*
 - validation of segmentation rules, 70-71*
 - website structure and, 119-125*