DIGITAL ANALYTICS PRIMER

JUDAH PHILLIPS

Author of *Building a Digital Analytics Organization: Create Value by Integrating Analytical Processes, Technology, and People into Business Operations*

Praise for *Digital Analytics Primer*

"The full strengths of an approach that is both practical and intellectual… You'll find plenty of real insight here, not just into how to do digital analytics, but also how to get digital analytics done. That's a distinction that's hard to appreciate until you've walked the walk, not just talked the talk."

—**Gary Angel**, Partner, Ernst & Young, Digital Analytics Center of Excellence

"This is the primer for the modern-day digital marketing practitioner. This book offers practical information on how to develop a process for data collection, reporting, and dashboarding in order to extract actionable insights from large data sets. Judah draws on his experience as practitioner and as an executive to offer real-world advice and guidance on how to win with digital analytics. This is a must-read for anyone working in the analytics field."

—**Jonathan Corbin**, Manager, Finance and Insurance, Adobe

"If you are new to the world of digital analytics, or just need a refresher, I can think of no better place to start than Judah Phillips' *Digital Analytics Primer*. This e-book presents a comprehensive look at the strategies, tactics, and terminology of data-driven businesses and 'big data.' It's definitely recommended reading for today's digital marketers."

—**Kim Ann King**, Chief Marketing Officer, SiteSpect, Inc.

"Whether you're an engineer, entrepreneur, manager, or business analyst, this book brings you up to speed on the latest concepts from data science and statistics, and how each is put into practice in a data-driven business. A must-read for anyone attempting to modernize their business operations and increase value from data and analytics."

—**Kurt Gray**, Director, Analytics Platform, Nokia/Microsoft

"Judah is one of the smartest guys in the digital analytics space. In this book he explains important, fundamental concepts in analytics that will help you win. Learn from his experience and you will be way ahead of your competition."

—**Anil Batra**, Sr. Director, Data and Analytics, Wunderman

Digital Analytics Primer

Judah Phillips

Vice President, Publisher: Tim Moore
Associate Publisher and Director of Marketing: Amy Neidlinger
Executive Editor: Jeanne Glasser Levine
Operations Specialist: Jodi Kemper
Cover Designer: Alan Clements
Managing Editor: Kristy Hart
Senior Project Editor: Lori Lyons
Copy Editor: Apostrophe Editing Services
Proofreader: Sarah Kearns
Indexer: Tim Wright
Compositor: Nonie Ratcliff

© 2014 by Judah Phillips
Publishing as Pearson Education, Inc.
Upper Saddle River, New Jersey 07458

Pearson Education offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales. For more information, please contact U.S. Corporate and Government Sales, 1-800-382-3419, corpsales@pearsontechgroup.com. For sales outside the U.S., please contact International Sales at international@pearsoned.com.

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

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed in the United States of America

First Printing October 2013

ISBN-10: 0-13-355206-3
ISBN-13: 978-0-13-355206-5

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.

For my family, friends, and colleagues who have helped me achieve my goals throughout the years. You know who you are. Thank You—especially Elizabeth, Lilah, Steven, Elyse, Merlin, Siggy, Baron, and Ivy. May they, and you, fine reader, be blessed with more good times, success, and memories past, present, and future.

This page intentionally left blank

Contents

Foreword

I'm an incurable addict when it comes to books. I read constantly. I'm so bad that yes, I usually even read the foreword. Reading a book's foreword is a bit like watching pre-season games in the NFL. You know they're meaningless—it's not the real stuff, but you do it anyway because you just can't help yourself. It's a few more weeks of football.

So when asked to actually write a foreword, it inevitably casts my mind back to all the many different types of forewords that I've read. There are a few basic flavors: the "Here's what you're about to read" variety, the "Anecdotes about the author" style, the "Here's how my own books relate to this theme," and the "Apparently unrelated to anything" musings of someone who doesn't seem to have read the actual book.

By this point, you probably fear that you're deep into that last alternative. Not so. Not only did I read *A Primer for Digital Analytics* with pleasure, it would be utterly foreign to me to pass up an opportunity to talk about analytics and the theory of analytics as embodied in what follows. This is stuff I unrepentantly love.

What Judah's written here is a broad primer for a digital analyst. Like any such text, it contains (implicitly or explicitly) a theory about the topic. Since Judah isn't someone who's reluctant to apply a theory to a problem (as you're going to discover when you read the book), you can bet that he put a lot of thought into the overall structure of the book.

Here's what that structure looks like:

Chapter 1: What Is Digital Analytics?

Chapter 2: The Analytics Value Chain and the P's of Digital Analytics

Chapter 3: Methods and Techniques for Digital Analysis

Chapter 4: Reporting Data and Using Key Performance Indicators

Chapter 5: What Are Analytics Tools?

In Chapter 1, you get the basic vocabulary necessary to talk the talk. Every profession has its own private lingo. This isn't an elitist kind of thing, at least not necessarily. It's just that the details of every profession demand specificity in ways that will never be present in general English. Our profession has this in spades, of course. From cookies to visitors, you'll get clear, careful definitions here. It's just like working in a foreign country—it's hard to succeed in a field until you know its language.

Of course, knowing the language is just the basic price of entry—it doesn't help you to know what to say! That's what the next three sections address. I'm particularly struck by the order and focus of these next three chapters. Putting the analytics value chain as the next step is interesting, and unusual, if not necessarily controversial.

Depending on your natural proclivities (or your level in the enterprise), this step-by-step guide to generating analytics value may appear absolutely fundamental, or it may make you feel like you're back in school. It made me feel a little of both. By nature, I'm a determinedly hands-on analyst, so if you're impatient with this type of step-by-step process flow that exists one level up from the real work, I certainly get that. But after years of doing enterprise measurement and analytics, I have to confess that a truly methodical approach to analysis is both essential and alarmingly uncommon. Far too many folks seem to assume that analysis just happens when you put an analyst in front of a tool and load it with data. Not really. That's like putting teens in a room and assuming that they'll inevitably dance.

But if, like me, you have a tendency to speed-read the abstract stuff, slow down for the next section. If Chapter 2 functions as an introduction to the general flow of analytics projects and might work best for project and analytics managers, Chapter 3 gets to the real heart of the matter: doing analytics. Here you'll find an excellent primer into basic methods of quantitative analysis and types of basic data modeling. I don't want to damn this section with faint praise. A primer on the basic methods of quantitative analysis may sound fairly routine, but in the world of digital analytics, it just isn't. Digital analytics as a profession has been shockingly neglectful of core techniques for using data.

You will search most of the key texts in our field in vain for any hint that statistics and modeling might be part of our profession. Any technique more advanced than cross-tabulation seems to be relegated to the "esoteric" category. That's just wrong. Putting this section smack in the pivotal middle of the book, and delving into a variety of appropriate quantitative methods, just feels right to me. This is the way digital analytics should be approached.

In the next section, we get a nuanced and careful introduction to KPIs and metrics, and the frameworks within which they exist. I like that KPIs are presented in terms of some fundamental frameworks, and that Judah lays out some alternatives here. I think that's more likely to keep people focused on thinking about the problem, not assuming that the author is providing a final solution. That's good because I'm not over-wedded to KPIs, and I don't happen to believe in standard ones, and too often I feel like our profession has treated them as the Holy Grail. "Find the right KPIs and your job is done." That just isn't close to true—and by placing this whole section *after* the two sections on analysis, I think that message comes through loud and clear. No one, I think, will read Chapters 2 and 3 and then come to the section on KPIs and feel that they're the core of digital analytics.

Finally, we come to the tools section. I think most of us would agree that putting tools last is right—partly because you can't think productively about tools until you understand the profession, and also because diving into tools just isn't the path to mastery. You'll get a nice introduction to the types of tools and some key tools in our profession, but this section is clearly written with the knowledge that no subject ages faster than tool talk. Tools change so fast and so frequently that writing specifics is an exercise better left to blogging than books. So instead, in this section, you'll find useful discussions of things like tool evaluation, RFPs, and buy-vs.-build decisions.

It's in areas like this that you see the full strengths of an approach that is both practical and intellectual. Heaven knows Judah isn't afraid of a theory, a method, or a framework—you'll get abundant exposure to many such constructs—but he's not at all removed from the practical exigencies of real-world enterprise measurement. You'll find plenty of real insight here; not just into how to do digital analytics, but also how to get digital analytics done. That's a distinction that's hard to appreciate until you've walked the walk, not just talked the talk.

Enjoy!

—Gary Angel, Principal, Enst & Young LLP

About the Author

Judah Phillips helps people create economic value using data, analytics, and research. He works with leading global companies whose executive and management teams are building, adapting, or reengineering their approach to digital analysis in order to increase profitable revenue, reduce cost, and boost profitability. Judah has worked for some of the world's largest corporations, globally managing business and digital analytics teams, including Sun Microsystems (now Oracle), Reed Elsevier, Monster Worldwide, Nokia, and Karmaloop.

He has lectured at undergraduate and graduate business schools, including Stern School of Business at NYU, Olin School of Business at Babson College, Carrol School of Management at Boston College, D'Amore/McKim School of Business at Northeastern University, and more.

He founded and globalized Digital Analytics Thursdays (DAT) and launched the Analytics Research Organization (ARO). Judah has served on the board of advisors to several Internet and technology companies, and has spoken at more than 40 technology industry events. Judah is the author or editor of several analytics books, including *Building a Digital Analytics Organization*. He lives in Boston and holds MBA and M.S. degrees.

Introduction to This Primer

This primer provides a broad overview of digital analytics in the context of understanding the people, process, and technology necessary to create value from it. It was written to be a stand-alone contribution to the canon of knowledge about business analytics; however, this book provides the basic information necessary to maximize your understanding of my longer book, *Building a Digital Analytics Organization*. This primer contains new content and also some content that also exists in my other book. Overall, this primer can teach you what you need to know when beginning to understand digital analytics or as you continue your immersion. To that end, the following topics are covered:

- Review of the fundamentals of digital analytics and related activities, such as social and mobile analytics, including a set of definitions about key concepts encountered when analyzing digital data across digital experiences.
- A discussion of how social and mobile analytics are different than site analytics, as well as how to start analyzing social data for branding or for direct response.
- Guidelines that form a business foundation for ensuring digital analytics considers Privacy.
- The Analytics Value Chain (AVC) reviews a phased approach for understanding what it means to "do analytics." The AVC explains the process of analytics from defining business requirements to collecting and verifying data to reporting and analysis to analytics communication to prediction and optimization.
- A review of basic statistical concepts in the context of the importance of Tukey's Exploratory Data Analysis (EDA) as well as

an overview of useful analytical techniques and methods, from regression to data visualization.

- An explanation of the importance of Reporting, Dashboarding, and Key Performance Indicators (KPIs), including RASTA Reporting and LIVES Dashboarding, which are acronyms, for the best practices in each. Also covered are descriptions of several KPIs common to digital analytics, such as conversion rate and average order value.
- A practitioner-focused overview of analytics tools—from what they do, to how to select them, to how to justify investment based on the state of your toolset, and common reasons why tools get replaced.

After reading this primer, you can improve your understanding of digital analytics whether you are just learning at a basic level or have experience. This guide can give you the building blocks for forming new ideas for applying digital analysis and digital analytics tools that you can take to the office tomorrow and consider as you approach your digital analytics projects in the future.

Although much of this primer is original material, some of the chapter content is also included in my longer book for analytics professionals titled *Building a Digital Analytics Organization*.

1

What Is Digital Analytics?

Digital analytics is a set of business and technical activities that define, create, collect, verify, or transform digital data into reporting, research, analyses, recommendations, optimization, predictions, and optimizations that create business value. The activity of digital analysis, at the highest and best application, helps companies create value by increasing revenue or reducing cost—often by providing a basis of fact for making decisions related to business planning, performance, and strategy. Digital data can be behavioral data about how people use and interact with digital experiences, transactional data, metadata about digital experiences, or data or metadata related to the events, clicks, and interactions within those experiences. Digital analytics can inform predictive modeling, testing and optimization, real-time and "real-timely" automation, controlled experimentation, competitive intelligence, and market research. As such, the analyses performed on digital data is collected and gathered from digital experiences, such as websites and browsers, mobile applications and devices, social networks and socially-enabled devices. Digital data that can be analyzed is also created by digital television and set top boxes, online and digital advertising, interactive billboards, and other fixed and mobile Internet connected displays, like kiosks in retail and public environments. The activities performed in digital analytics require coordinating processes, people, and technology internally within a company and externally from partners and vendors in order to produce analysis that answers business questions, makes recommendations, predictions, and automations based on mathematically and statistically rigorous methods, and enables successful business activities that create value.

Digital analytics also incorporates the upstream activities critical and necessary to analyze data, such as the technical and process work (requirements/questions, data collection, definition, extraction, transformation, verification, reporting, analysis, optimization, prediction, automation, and tool implementation and configuration). By bringing together data from different systems to create cohesive and relevant analysis, digital analysis is used to answer business questions and provide a foundation for fact-based decisions. The best digital analytics teams tell "data stories" to answer the "business questions" asked by stakeholders. The analytical insights in the data-driven answers to business questions provide recommendations, data-oriented guidance, and insightful ideas to management that helps make their company money by reducing costs or increasing profitable revenue (or both). Digital analysts, the people on the digital analytics team, navigate effectively the upstream technical and downstream social and organization processes inherent in executing a data-driven communication function via processes that unify teams across technology and the business. If that sentence is hard to pull apart, read it again when you finish the primer, and your interpretation can be crystal clear.

Digital analytics involve defining, collecting, transforming, processing, governing, analyzing, reporting, predicting, optimizing, communicating, and managing data (and the human and technical resources) related to how people use digital experiences. Digital analysts can look at data from one channel, such as a website, to data across multiple channels individually (multichannel analysis across the mobile and website) and against one channel and others (omnichannel analysis of the site, the mobile site, and the app). IBM estimates that 90 percent of the world's data ever created was created during just the last 2 years. This largely digital data can be thought of as clickstreams—a legacy web analytics term. *Clickstream* is a word that refers to the many deliberate series of actions (often clicks) performed by people, during a visit, when using a digital experience (such as a mobile app). The clickstream includes the screens (pages) viewed on the digital technology (such as a phone or mobile browser) and all the events, interactions, transactions, metadata, and behaviors associated with a visitor visiting and using a digital experience. Although the clickstream contains an abstraction of a person's (visitors) behavior, the Internet's underlying infrastructure makes freely available other information (and more) that can be joined with a clickstream. Such data includes

- The visitor's IP address, mobile phone type, tablet type, browser used, mobile device type and version, and operating system.
- The referrer (such as a search engine) or previous site or screen viewed.
- Keyword information entered into search boxes directly in internal search on your website or in your mobile app. Or, the keyword entered externally in a search engine or clicked on in a paid search campaign to land on your site.
- The duration of time spent per page on a website or per screen when using an application. Time-derivatives include the time spent hovering over an online ad or in the total time a person spent online in a given month.
- Product viewed, left, and abandoned in an online shopping cart, and the products purchased—segmented by marketing campaign.
- Amount of money spent on a transaction during the month in a country, and derivatives like AOV (average order value).
- Number of page views, visits, and unique visitors during a given time period.
- Customer attributes like a credit score, Klout score, LTV (lifetime value), and customer segment.

All these clickstreams and related data are stored in one or more databases and made available at the individual, detailed (raw) or summed-up, aggregate (summary/indexed) levels for querying, reporting, integrating, and using in other systems via a number of methods from flatfiles to web services, to analyzing, mining, predicting, and optimizing. Regardless of the digital data you choose to collect and analyze, it is stored either at your company on servers in a data center or in an infrastructure owned and managed by another company, such as a cloud-based SaaS application or a private cloud-managed service. You may even have one or more analytics appliance or server solely dedicated (single tenant!) for analytics. Companies enhance clickstream data with other related information and meta-data to add value to it, such as demographic information (such as educational level) to understand more about their audience.

Digital analytics is a complex business activity to execute effectively. The definition, collection, reporting, analysis, and communication of digital data is highly cross-functional and requires the support of multiple teams in IT and the business—and such coordination and alignment may need improvement to compete, win, and succeed in the long term with analytics. For example, data collection may be defined by the analytics team, implemented by the development team, tested and verified by the QA and Analytics team, and the reporting and analysis accepted only by business users. Each step in the process of "doing analytics," called the Analytics Value Chain (AVC)—discussed in Chapter 2, "The Analytics Value Chain and the P's of Digital Analytics"—must be aligned with the previous and next step in the process, and the end state must yield accurate data that can be used for analysis. Thus, analysis requires not just technology, but more important, business-oriented analytical people at every step in the analytical process who have solid business and technical skills and have been adequately trained to follow established analytical processes for gathering requirements and defining, collecting, transforming, measuring, reporting, and analyzing data.

Table 1.1 was created by Tom Davenport (the Information and Action row), Harvard and Babson College professor and world-renowned academic and author of *Competing on Analytics* and *Analytics at Work*, and me, Judah Phillips (the Action row and the parenthetical activities). The grid describes how data evolves from being informational in the past to actionable in the future. Each cell identifies questions and potential business activities for digital analytics to understand what happened, what is happening now, and to predict what will happen in the future. In addition, the grid helps you understand "insights" about why things happened, what may happen, and whether what happened is possible with digital data analysis.

Table 1.1 Evolution of Data

	Past	Present	Future
Information	What happened? (Data Mining and Reporting)	What is happening now? (Alerts)	What will happen? (Trending, Extrapolation)
Insight	How and why did it happen? (Data Modeling and Experimental Design)	What's the next best action? (Recommendation)	What are the best and worst things that can happen? (Prediction, Simulation)
Action	How do you leverage what you already know? (Dynamic Interaction/ Profiling)	How do we dynamically modify the site in real-time or in a timely way? (Detection)	How can you apply the data to the future? (Ongoing Optimization)

Digital analytics creates value by informing work activities and ensuring that data is considered across the decisioning process beginning at the beginning: Phase Zero in engineering parlance and "at the beginning of the project" in common diction. Table 1.2 shows in vertical columns ways that analytics can help different functional teams.

Table 1.2 How Analytics Can Work with Different Teams

Marketing	Product	Sales	Executive	IT
Landing Page Optimization	Behavioral Analysis	Customer Value	Dashboarding	Performance Monitoring
Lifetime Value / RFM Models / Customer Segmentation	Search Engine Optimization	Sales Readiness	Scorecarding	Site Usage for Capacity Planning
Search Engine Marketing	Demo/Geo/ Firmagraphic Analysis	Sales Collateral	Custom Research	Disaster Recovery

Marketing	Product	Sales	Executive	IT
Ad and Media Plan Optimization	Funnel and Flow Optimization	RFPs and RFIs	Financial Performance	Infrastructure Enhancements
Social Media Optimization	Application and Product Performance	Customer Usage Information	Competitive Intelligence	Tag Management and QA

Digital analytics includes the work of web analytics or site analytics (and "mobile analytics" and "social analytics") because digital analytics does not regard the experience as a primary focus nor the tool used to collect it as primary concern. Digital Analytics cares about digital data regardless of the source from which it was created, the type of data it is (qualitative or quantitative), and where it exists (external and internal to the company). The term web analytics is an archaic term that refers to site analytics when the Internet was not pervasive and ubiquitous across human life in first-world countries. When web analytics was invented and commercialized in the 1990s, the primary vehicle with which to access the World Wide Web was the browser. A popular browser was Marc Andreessen's Netscape. In this world, which was not highly mobilized or composed of social media, the term web analytics was created. Web analytics is about the site as access, primarily but not exclusively, through the web browser (such as Google Chrome, Mozilla Firefox, Opera, or Internet Explorer). Digital analytics includes but extends far beyond "web analytics."

Today, the Internet is everywhere—in your household appliances, in your pockets on your smart devices, in your automobiles, in stores, and almost everywhere you go—you are hyperconnected and always on. More than 6 billion mobile phones are in use worldwide. The Pew Internet and American Life Project estimates that as of May 2013, 91 percent of Americans have a cell phone, and 56 percent of Americans have a smartphone. Meanwhile, more than 34 percent of Americans have a tablet computer, and 26 percent own an e-reader. All this usage by people creates huge volumes of data—the so-called "big data" that can be collected, mined, and analyzed using analytical methods and data scientists.

Digital analytics is about doing analysis on all this digital data from wherever it is collected on whatever device to create business value from it. Digital data is certainly web data, and it is also other types of digital data: mobile data, social data, customer relationship management (CRM) data, business intelligence (BI) data, and other digital data related to transactions, behaviors, and audiences. Thus, do not think web analytics is digital analytics because it is not. Digital analytics is bigger in size, grander in scope, and more critical to apply within your business than ever before.

How Does Big Data and Data Science Fit into Digital Analytics?

Big data is a phrase that has gained widespread popularity and mainstream acceptance as something new and innovative. McKinsey Global Initiative defines big data as "datasets whose size is beyond the ability of software tools to capture, store, manage, and analyze." Then MGI disclaims its definition as being "intentionally subjective and incorporates a moving definition...." In this book, I emphasize the importance of the criticality of definitions as fundamental to understanding. The fact that a leading think-tank like MGI can't pin down a definition for big data creates both frustration and opportunity for those practicing digital analytics. Other anecdotal definitions for big data overheard by the author include "datasets larger than 5 TB" and "data that can't be handled in Excel" to data requiring more advanced infrastructure and processing platforms, such as Hadoop/Hive and Cassandra.

A few years ago, I worked with big data, which fits MGI's definition because the world's leading database vendor would pause and process and process and process because queries were entered against raw data with billions of rows. IBM now owns the software I was using, but no one called what my team was doing big data; it was just said that "the database can't handle the size of the data" and decided materialized (summary) views of the common queries should be built to help out. Still, we had a lot of slowly changing dimensions and cardinality.

"Big Data" Not New

I postulate that as long as there has been data, we have had big data that has been pushing the limits of software. As a boy, on my Amiga, the 512 k expansion kit was not sufficient to run Space Ace and required 2x the memory (more than 1MB was over $1000). This was big gaming data in the 1980s on the Amiga according to MGI. While I first met the notion of big data with great skepticism, as you may read in my tone, I have come to agree that given the mainstreaming of the Internet and the interweaving and intermeshing of digital into commerce and life across the world, there is more to big data than marketing spin to sell new software products. And without a doubt due to the easy availability of the Internet and the widespread distribution of devices that create digital data, the lower costs of materials, and generally lower barriers of entry to create or work with digital business, the volume and velocity of the available data to analyze is growing more quickly in larger quantity than ever before. Big data is real, and it is here to stay.

The data speaks for itself; MGI claims 15 out of 17 industry sectors have more data than the Library of Congress, which has 235 terabytes of data. Anecdotally, some large companies have claimed to process petabytes per day. The reason for all this big data is the potential for profit. MGI claims a 60 percent increase in retailers' operating margins is possible with big data. Just location-based big data has the potential to create a $600 billion annual consumer surplus. The United Nations claims that more people on Earth have access to cell phones than toilets. Six billion of the world's 7 billion people have access to mobile phones. Only 4.5 billion people have access to working toilets. Meanwhile, 2.5 billion people don't have proper sanitation—big data on mobile devices is more common than the infrastructure used for human sanitation.

As a result of all the big data in the public and private sector, McKinsey estimates that 1,500,000 more data-savvy managers and 140,000–190,000 new roles for analytical talent are needed to support the growth in big data in the future. Yet, the opportunities for analyzing all this data today remain largely untapped. EMC estimates only 3 percent of the data created today is useful for analyses, whereas only .05 percent of the data is actually analyzed. Therefore, 99.05 percent of the data, thus, is not analyzed, and there is opportunity for businesses to create value. Within all this big data, you must remember:

> Big data is little more than a lot of small data.

And in that simple lesson learned is the key to solving the puzzle of big data. Although the analytical techniques to use on big data are covered later in this primer (and discussed in greater depth in my book *Building a Digital Analytics Organization*), the processing of big data can be technically challenged. Big data (unless aggregated) won't fit in Excel, but big data can be analyzed, insights can be derived, and value can be generated from informed recommendations. Typically, the best way to start creating competency in digital analytics and building a digital analytics organization is allocating at least part-time resources—and often by hiring an experienced practitioner, like myself, to help.

When building the digital analytics organization, you will inevitably encounter a newer label for an analyst: data scientist. Like big data, data science has become another phrase gaining mainstream acceptance to describe a person with a type of analytical talent that merges together data-oriented computer science with statistical skills and business acumen. *Data science* is defined more broadly by the community of Wikipedia users as

> Incorporating varying elements and builds on techniques and theories from many fields, including math, statistics, data engineering, pattern recognition and learning, advanced computing, visualization, uncertainty modeling, data warehousing,

and high performance computing with the goal of extracting meaning from data and creating data products. Data science is a novel term that is often used interchangeably with competitive intelligence or business analytics, although it is becoming more common. Data science seeks to use all available and relevant data to effectively tell a story that can be easily understood by non-practitioners.

Again, like the definition of big data, the definition of data science is evolving, somewhat all-encompassing, and definitely a moving, living definition, which can irk and frustrate someone who has been a data analyst or manager of data analysts. In the rebranding of "analysis" to the conceptually more complex "data science," the industry is calling out the role of big data, applied analytical techniques that are statistically rigorous, an emphasis on technical prowess, and technology competencies that are aligned and focused on solving business issues.

In the other words, the Pollyanna vision of a "data scientist" is a person who can execute analytics from phase 0 to completion, who can manage and execute the Analytics Value Chain (discussed in Chapter 2) up and down all analytical phases. The data scientist can ideally gather and define business requirements, code the data collection, build the database for collecting the data, implement a BI tool, manage data integration, execute applied statistical analysis, and summarize and present results in business language to business stakeholders. These types of cross-functional skills rarely exist in the same person, and even if one person could do all that, success with analytics isn't a result of a cowboy data scientist, but rather success with analytics is the result of cross-functional teamwork in which data science and big data play a role in creating the eventual business value delivered.

The idea of a data scientist who can do it all—as previously described—can also create an unobtainable profile for Human Resources to find. And judging all analytical talent by the set of skills of the "data scientist" would be unfair to apply to current and future practitioners because the inexperienced or new hiring manager or Human Resources professional may think that all analysts can do all data sciences, but that's an unrealistic perspective that will set up the believer for great disappointment. Analytical talent and people vary in their skills, and analytics includes data science and other approaches to analysis. Thus, although data science is real, valuable, and a skill set that can be achieved as part of a large digital analytics team, hiring managers must match the skills of real-world analytics professionals to the requirements of analytics jobs regardless of the notion of data science.

Data scientists have the following characteristics, which are not entirely different than those of digital analysts but can be perceived to be more mathematically rigorous and academically aligned:

- **Academic training.** Most, if not all of the entitled and self-proclaimed data scientists have advanced degrees (usually Master's and not uncommonly PhDs) where they have studied a quantitative discipline that is not necessarily computer science, math, or finance, but could be.
- **Statistical acumen.** Most the data scientists have a solid, and often impressive, knowledge and erudition into statistical data mining and even machine learning. Other data scientists have learned statistics by immersion in research disciplines in which data analyses and the understanding of analytical methods and techniques is necessary, like psychology and social sciences.
- **Experience with database and data extraction, transformation, and loading (ETL).** Whether learned on the job, in self-development, or in a formal educational environment, data scientists know a given level of SQL, understand databases both conceptually and tactically, and know how to work with current database, reporting, and analytics technology. They have some knowledge of ETL. That is not to say that the data scientist does the technical database work, but rather may have strategic oversight and approval over the data model to ensure it meets the needs of the data sciences.

- **Business knowledge.** Considered one of the rarer qualities of a data scientist are the business skills and knowledge and ability to work with business people. The best data scientists understand that professional data science can exist only in data business. They operate with knowledge of business concerns and work toward fulfilling business requirements and goals.

No one person can do it all, and although some experts can come close, the purpose of this primer is to help people start or learn more about doing analytics on data of all sizes using simple conceptual approaches, more complex data science, and other analytical methods and techniques. Even the best data scientist can excel, improve, and create better analysis by working with a team of data scientists, analysts, engineers, business people, and managers within the digital analytics organization.

The Strategic and Tactical Use of Digital Analytics

Digital analytics can help a business in many ways. The two penultimate goals for the highest and best usage of analytics are to create value by generating profitable revenue and reducing cost. Not all companies are ready, however, to make and manage the investment necessary to create a digital analytics organization. After all, there is an opportunity cost in attempting to apply analysis to solve a business challenge, and it is not always obvious when to use analytics.

The right time to use analytics is when, above all else, revenue is at risk and you can improve the situation using data. If revenue is on the line and analytics can make an impact, all analytical resources (within reason) should be immediately applied to the project/issue until that revenue is earned. However, sometimes just running with the ball and providing a quick answer given available data is not possible. Thus, regardless of the criticality and urgency, do not rush the analysis.

Digital analytics, like any fact-based analytical method, requires accurate data. While that sounds obvious, the analytics team needs to ensure the data is accurate, clean, and fit for the analysis. Remember that analytics can be applied only when you have data that contains information that you need. It sounds obvious, right? But many business people bend data to fit a business situation when the data does not fit it. The analytics team must watch out for the misuse of data. Accurate and verifiable data is powerful in environments in which more than one person collaborates or compromises on decisions (that is, cross-functional or highly matrixed companies). The team or person who controls the data has positional power and influence because the performance of others can be judged by the data—and thus, data ownership (as well as analytical interpretation) can be political. The best analytical leaders and talent attempt to reduce politics and concentrate on the making money from data analysis.

The analyst and analytics team should remain outside of political influence. Politics can sometimes lead to the inaccurate presentation of data, or the usage of wrong (and misguided) analytical data to prove a point. First and foremost, analysis, regardless of organizational politics, must be given to someone who can take action or do something useful with it. Data in and of itself is not "actionable," but the people who consider and use the analytical deliverables can take action based on recommendations that result from the data. In that context, do not waste time building reporting or providing analysis to people or teams who can't do anything with it. Sometimes people "just want to look at it." This is called JWaTLIT (juh-what-lit). Avoid JWaTLIT.

Stakeholders ask for a lot of analysis and data requests. As an analytics practitioner, remember there are times in which data can help, and other times, for many reasons, the analytics team is not the right team to help. Sometimes the business has no clear data to help—and cobbling data together without standard definitions from disparate systems sometimes works, but usually fails. Other times, the business questions are not clear or adequately stated, or there can be little to no alignment on business goals or the expected business outcome of the analytics project. In other projects, the digital analytics team may determine that it will take too long to get the data (if it exists). This delay can be due to the data collection of new data and processing of past data (which in big companies could take considerable time) to getting no support at all (or worst case, no response) from the supporting teams on which you are dependent for getting the data. Finally, other reasons exist for remaining cautious about applying data to solve business issues that are related to budgets, other priorities, and available resources.

It can be helpful to ask stakeholders, analysts, managers, and yourself the following questions about any analytical request. The answer to this question can help elucidate the underlying impetus and catalyst for the business question, the resulting action, and even an idea of the type of data and analysis the stakeholder may want to review:

> If you had the data in your hands right now, what business decision would you make?

A Practitioner's Perspective on Answering the What and the Why with Digital Analytics

One commonly expressed idea in web analytics is that quantitative data can tell you only the "what" but not the "why." In other words, you can measure a conversion rate (the what), but digital data can't tell you the underlying cause (the why). The idea being that quantitative data (the what) such as those collected by web analytics tools will never tell you why in the same way the qualitative data from a market research or Voice of Customer (VoC) analytics tool can. A practitioner may hear this what-versus-why notion postulated by experts and salespeople alike.

This idea is only partially true and is a logical difference that no one in the business cares about. As an assertion, business people are likely to be annoyed and think you aren't talented if you say "data only tells you the what, not the why," so avoid saying this (except when complaining to your friends in analytics). Only analytics practitioners and the people vending/selling analytics tools differentiate on this notion of quantitative data as the "what" and qualitative data as the "why." This pedantic point is often not worth making the comparison. Although it might sound smart on the surface, the idea largely doesn't hold up under scrutiny. If you, the analytics practitioner, make the case that analytics data can tell only "what and not "why," you will not only look like you are making an excuse, you will also open yourself up to sharp criticism.

The notion of "why" is largely subjective. Business users might consider the underlying quantitative data that explains the movement in the data to be the "why" and thus answerable by countable data and not text data. Whereas an analyst in pure analytics geek mode may think why can be learned only from a research survey, but that's not entirely true, though the analyst has a good point. For example, if the aggregate sum of revenue dollars from marketing campaigns changes month-over-month, it is likely the campaign performance has changed. Finding out the change in campaign performance month over month is necessary to identify why the data change. With month-over-month change, the simplest reason "why?" is answerable by quantitative data. On the other hand, there are underlying mindset and psychological reasons why people may not engaged with or respond to a marketing campaign. This qualitative data is a deeper reason "why?" and can be only uncovered in market and customer research.

Digital Analytics Concepts

Digital analytics can be understood both logically and conceptually. Little formal education exists for digital analytics; thus, people who have experience in business have learned their skill sets through hard work, tenacity, determination, and effort from hands-on, practical experience. The following section presents a practitioner's viewpoint on the digital analytics concepts. These concepts apply across digital analytics—from site to social to mobile—and represent a set of business-focused definitions that can help analytics practitioners, at all levels, form a foundation for and solidify their existing understanding of digital analytics. No standards apply to digital analytics, like those you would find in other business-focused, quantitative disciplines such as financial accounting. No Generally Accepted Accounting Principles (GAAP) exist for digital analytics data. No Financial Accounting Standards Board (FASB) exists to create standards or definitions for digital data. As a result, there are no standards widespread. Certainly no consensus-based, agreed-upon standards and definitions exist for the concepts covered in this section. Thus, *the definitions and descriptions in this section are not the final say on these concepts nor are they an attempt to be exhaustive or overly formal.* On the other hand, the concepts present ways to understand ideas that are intangible and constantly evolving and yet are based on practical experience. After reviewing these concepts and learning about them, you can apply them in your career to advance your ability to understand, apply, and analyze digital data. As you apply them, you may extend or modify the way you understand these definitions since they are a starting point for forming a viewpoint and understanding of analytics.

Common vocabulary in digital analytics includes the following:

- **Visitor (or unique visitor).** The unique visitor is among the most often cited metrics related to how many "people" visit a site, but it is also a confusing and complicated metric because it can be calculated many different ways by different tools. Because there are no standards, the underlying measurement for a unique visitor is left to the vendor. Thus, different tools measure and count unique visitors differently. As a result, the counts of unique visitors on the same site or app will be counted with different numbers by the same tool.

All tools both on-premise, software-as-a-service (SaaS), and data vendors have similar but different definitions for unique visitors. Companies have attempted to clarify the confusion by naming "unique visitors" with new names such as "unique browsers" or "unique people." At the time of writing this chapter, unique visitors is a mildly differentiated concept understood by digital analysts and those who spend time working with digital analytics (see Table 1.3).

Table 1.3 The Unique Visitor Defined

Tool	How Unique Visitor Is Defined	Example of Tool
Digital Analytics	A count of deduplicated cookies during the time period. In other words, if you stay at the beach house for all 30 days during an entire month, you are 1 "monthly unique visitor." If the owner of the rental property was asked by his account "how many people stayed at the house?," the correct answer is 1. In that same scenario, you stayed every day, correct? If the manager asks "how many people stayed at the house each day," you could say, "30 daily visitors" stayed during the month—because you did stay 30 times—once per day. You were "30 daily unique visitors" but only "1 monthly unique visitor." In digital analytics, the time period of the uniqueness is important—and as such, you would never add up a daily unique visitor count to monthly count. Although cookie-based unique visitors are the most common measurement, it is possible to count uniqueness in other ways, using login, query string parameters, lookups, and other derivatives.	Google Analytics, Omniture/Adobe, and Webtrends

Tool	How Unique Visitor Is Defined	Example of Tool
Audience Measurement	Audience measurement uses mostly black box data collections methods that are not transparent, and yet are similar to the algorithms within analytical software. Audience measurement tools combine data collected directly from sites with data collected from a panel of people who choose to have their digital behavior (such as web surfing) monitored in exchange for some incentive, such as free software. Audience measurement companies take the data collected from their sample panel about digital behavior, such as website visitation (and much more), and use proprietary statistical methods to estimate the size of an audience to a website or other digital asset. The estimate is based on the behaviors noticed in the panel (and perhaps on the site) such that a cookie deletion factor can be identified and used to reduce the count of unique visitors to people. These estimates may or may not be combined with cookie-level visitor estimates derived from on-site measurement. This approach to audience measurement, sometimes called hybrid, estimates unique visitors based on unifying together both panel-based data and census/site/cookie-based data.	comScore, Nielsen, Compete, Quantcast, and Google

In the table, unique visitors are a different concept based on the tool because multiple vendors use nonstandard data collection and unaligned measurement methods—few of which are transparent and standardized. The net result is that different measurements are named the same name, which in the case of unique visitors creates and maintains confusion across the industry—from executive to analyst.

Remember that the goal of the unique visitor metrics is *to accurately count people who are exposed to and engage with a* *digital experience to drive profitable global commerce.* In that context, you might hear people, unique cookies, unique browsers, unique people, absolute unique visitors, absolute visitors, or just plain ole visitors. All these are an attempt to quantify mostly the same concept. What can an executive do to cut through the noise?

- **Ask for the definition.** Critical as can be are data definitions for digital analytics concepts. Make no mistake that definitions are not standards. There are few standards. Determine how your vendor identifies unique visitors and compare that to your needs and the offerings of competitors.
- **Ask for time-series trends of your site.** View a time-series trend of the last two years of your site. Then compare all the data sources together. If your vendor can't give that to you, look at the longest time period possible. Notice and ask any questions about strange data movements in unique visitors and why they happened.
- **Ask questions until you understand.** All vendors who are worth your budget should easily and accurately express their data definition for unique visitors (or unique whatever) with no frills or gunk. And they should qualify why they measure the way they do. If vendors are not transparent, are unclear, or can't explain why they measure how they measure, then move on. And repeat the first three tasks until you find the right data vendor for you.

- **Visit.** If visitors visit a site, a visit is the number of times a visitor goes to a site. The same concept applies to any digital experience. When you open and close a mobile app, you just visited it, and the analytics tool counts one visit. When you log in to Facebook, you have just had one visit to Facebook. If you go to your house five times a day in between shopping, taking the kids to sports, going out to get take-out, and other errands, you had five visits to your house but were just one visitor to your house. Get the nuance between visits and visitors? Visitors have one or more visits to a digital experience. A visit is something that a visitor does, and a visit is recorded when the visitors starts to use a digital experience. All clicks, interactions, events, page views, and conversion happen in the context of a visit. You will have the same or more numbers of visits than visitors. If you have one monthly unique visitor to your website, and that unique visitor visits the site five times during the month, your analytics tool would show one visitor and five visits—assuming the definition for the data supported the assertions in this use case. (Remember, there are no standards for any of this stuff.)

- **Page views (web) and screen views (mobile).** A visitor views pages on a screen during a visit. Page views are simply the number of pages that were viewed during a visit by a visitor (or summed cumulatively over some time period). The emerging vernacular on the mobile device is screen views (since pages do not exist). When you visit Facebook and click on five pages, you are one visitor who has had one visit during which there were five page views. If you did the same activity on your mobile phone, you would still have visitors with 1 visit, but you would have had five screen views (not page views). If you go to a coffee shop once per day every day during a 30-day month and read a 10-page newspaper during every visit, you did the following in digital analytics logic: 1 monthly unique visitor, 30 visits, and 300 page views. In the same sense, when you go to a website (or mobile site or other web-enabled experience, such as an in-store kiosk) and you view content, it is generally said that you "viewed a page," hence the metric "page views."

  Dynamic content, such as those created by Rich Internet Applications (RIAs), does not create traditional page views. This type of digital experience has been increasing since 2008. For the more advanced reader, RIAs and Asynchronous JavaScript over eXtensible Markup Language (AJAX) have resulted in a decoupling of the client and server that traditionally held together in HTML 1.0 and the evolution. With HTML5, FLEX, AIR, FLASH, AJAX, RUBY, and other emerging technologies, the traditional web analytics concept of the page view does not always exist. As a result, vendors have created data collection methods that allow for the creation of virtual page views (often called Events) in these dynamic environments.

In mobile experiences or interactive experiences that are web-enabled (such as set-top boxes and interactive billboards), the construct of a page view is insufficient to express the object around which the behavior is centered. For certain digital experiences, in which viewing of content is technology-agnostic and independent of device or browser, the term screen view is used—or as described next, the event.

- **Events (or interactions or goals).** Events are abstractions of behavior that happen on a page and are collected as behavioral data in analytics tools. An event could be a newsletter sign-up or an interaction with faceted search and so on. The event model describes conceptually that, quite basically, behaviors happen within the page that do not necessarily create a communication back to the web server. In other words, certain activities, actions, clicks, and interactions have meaning and need to be identified as "events" on the page; hence, digital analytics tools enable the collection of custom-defined events (often collected via JavaScript or a custom API call). Because of RIA, AJAX, and HTML5, a digital event model is necessary for digital measurement. Without events, it would be impossible to collect what happens on the page at the level below (or in place of) the page view. For example, on the page, you can fill out a form, click a button, start and stop a video, select from a list, or click on this or that. All these things that occur on the page or the screen, within the frame you see, are events. Events can be a click, toggle, search, form field, or any analyst-definable expression of action that occurs on a page.

  Events are subordinate activities to the page view (or screen view). Anything you interact with—click, swipe, pick, select, rub, push, point, or touch with your finger, pad, mouse, or any other input device—in any web-enabled or programmed interactive digital environment can be an event.

- **Cookies.** Information stored by technology on your machine, such as a smart device, laptop, server, Internet-enabled television, and so on. Not all machines accept cookies, but they are common. Cookies should be searched for using your favorite search engine and read about on Wikipedia. The usage of cookies is far outside of the scope of this primer but should be understood deeply by analytics practitioners. Understanding cookies is useful for comprehending the mechanics and technicalities of how visitors and visits are counted by both analytics tools and audience measurement tools. That said, you can track behavior in aggregate and at the visitor level without using cookies (for example, IP address or login name or other parameter persistent or not). Here are the important facts for a manager:
  - **First and third-party cookies.** Can be set either by a site or another company. If the site you are on sets the cookie, it is considered a first-party cookie. If you are on a site and another company, such as an advertiser whose advertising banner you see, sets a cookie from a different site (domain), the cookie is a third-party cookie. In other words, if you go to a site and the site sets a cookie, the cookie is a first-party cookie. If some company is setting cookies on a site you are visiting (often without you having any idea), the cookie is a third-party cookie. Third-party cookies are often considered suspicious, so many companies attempt to set first-party cookies—or use tricks to make it look like a third-party cookie is a first-party cookie.
  - **Cookies are set to expire by the creators.** Cookies can last just during the visit (called *session cookies*) or for many years (called *persistent cookies*).
  - **People delete cookies—all the time.** And not just because they visit adult-themed sites, but also because people and countries have privacy concerns about any tracking linked to an object that can remain on a machine. Security, disk cleanup, system utilities, and antivirus software may all delete cookies. That said, people could delete and even block

cookies by choice. When persistence of cookies is necessary to count uniqueness, the result of one person (or device) having many cookies is that the person looks like more than one unique visitor. Thus, cookie deletion impacts your ability to count unique visitors.

- **Because people delete cookies frequently, the count of unique visitors in site-side, SaaS, and other tools that use cookies (primarily) as a means to count unique visitors are overstated.** This overstatement due to cookie deletion is thought to be a weakness to web analytics tools—and is used to compete against web analytics tool vendors by audience measurement vendors. In fairness to web analytics vendors, the estimate of unique visitors and other similarly named concepts in audience measurement tools is also likely inaccurate due to various types of measurement error.
- **A manager should worry less about cookie deletion—and leave that to the technical analysts on the team.** It is easy for a smart and curious manager in analytics or the boardroom to want to express their technical acumen by criticizing cookie-based measurement and commenting on their understanding cookies and cookie deletion on digital data. Although it's a good idea to become educated, do it as a lunch session with the team, not when it's critical to get the job done. The bottom line with cookies is like visitors: Question what impact they have on the data, and make a judicious decision on whether the digital data has integrity due to the impact of cookies on measurement.

- **Sessionization.** For a manager, this is a fancy word used by analysts to refer to the underlying process of how visits are recognized, processed, and counted in the technical context of first- and third-party persistent and session cookies and other technical nuances. A session is typically another word for a visit; thus, the process of sessionization is a complicated way to describe counting visits. Note that years ago, a session was a set of different visits across more than one web site, but I don't hear that definition too frequently anymore.

- **Visitorization.** For a manager, this is another fancy word, like sessionization, used by analysts to refer to the underlying process of how visitors are recognized, processed, and counted in the technical context of first- and third-party persistent and session cookies and other technical nuances. When an analytics tool runs its underlying algorithm to deduplicate cookies and determine the count of unique visitors, it is said that visitorization occurs.
- **Conversion (and conversion rate).** A conversion is a value-generating transition in a digital experience. For example, a conversion might be an order for a product on an e-commerce site. The conversion rate is the percentage expression of how many people ordered. For example, if 10,000 people came to a site and 200 bought a product (converted), the conversion rate is 2 percent. Conversions are often broken down into different types: site conversion, landing page conversion, and even concepts like macro (site), micro flow conversion across pages or events), or scenario conversion (shopping cart conversion).
- **Deduplication.** Specific to cookies and the measurement of visitors and visits, deduplication expresses how multiple cookies (or other objects used to sessionize/visitorize) are reduced to a single count to not overstate the count of visits or unique visitors. For example, if a cookie is set every day for a logged-in user on a website, it is possible to know the 30 cookies tied to the logged-in visitor. And instead of counting those 30 cookies as 30 unique visitors, deduplication occurs, saying "These 30 cookies are duplicated for just 1 visitor; therefore, don't count 30 visitors, count only 1." Deduplication reduces multiple cookie values to one value and enables more accurate aggregate counts of visitors.
- **Uniqueness.** Similar to deduplication, the idea of reducing to one visitor across multiple sites/domains or across multiple devices (such as exposure to mobile, TV, radio, site, and QR code) exposed to one or more messages (generally advertising-based) is called uniqueness. Omnichannel analytics attempts to

understand uniqueness across more than one channel. If you have one visitor to Facebook on mobile and the same visitor using the site, to express uniqueness you would deduplicate the visitor on Facebook across both mobile and site usage to accurately count (or estimate) deduplicated visitors and their visits. Given the potential need for off-site and cross-domain tracking, site-side analytics tools can estimate uniqueness across sites on which they are installed, but not beyond those sites. Thus, uniqueness across multiple sites (not owned by one company) and digital experiences are often provided by audience measurement tools.

- **Attribution.** Entire books have been written about attribution. Attribution for the manager refers to understanding the source of visitors and visits to a digital experience. Attribution is most often cited in the context of marketing budgets (and justifying them). Traditional analytics does attribution modeling and calls it many different names.

  Attribution might refer to understanding the impact of paid, owned, and earned media and its impact on the site. For example, how many sales did you have from Google? Or did email or social media cost you more for customer acquisition? What was the return on media spent for every dollar you spent? Attribution can help answer these business questions.

  Common attribution models include last-click (where attribution is given to the last click before the conversion); first-click (where attribution is given to the first click that even sent the person to the site before the conversion); weighted (where first and last click is weighted, such as 70 percent first and 30 percent last); event-driven (where certain events across or within channels are proxied and enumerated to identify influence on conversion); algorithmic (where machine learning and statistical algorithms are applied, such as Shapley Values); to other proprietary types of digital attribution. See the discussion in Chapter 3.

Paid, Earned, and Owned Media in Digital Analytics

Digital analytics helps you understand the direct response from these channels (and in some cases, even the impact on brand) across **Paid**, **Earned**, and **Owned** digital experiences. In digital analytics, you can segment paid, earned, and owned media into the follow categories (entirely my framework). Note the following framework is not exhaustive. This is a common framework that can serve as a foundation that can be built upon as you advance your understanding of paid, earned, and owned. The labels and definitions used to define the segments of paid, earned, and owned media will likely shift over time and represent a best effort to define them in 2014:

- **Paid media.** Paid media is just what it sounds like. You or your company pays to advertise for various consumer and shopping goals in the mind of the current or prospective customer.
  - **Online marketing.** The set of channels that compose an online marketing function, including creating, defining, managing, strategizing, and delegating the work required to advertise to current and potential customers, for example, all digital traffic from promotional or coupon codes entered online.
  - **Display and (re) targeted advertising.** Banners ads and other ad units of all sizes served on Internet-enabled devices, primarily the Internet. This segment could be further segmented into underlying campaign types, such as ad network, DSP, ad exchange, retargeters, and so on.
  - **Affiliate marketing.** Marketing resulting from relationships in which some form of capital is exchanged. For example, affiliate marketing generally occurs after you purchase or even abandon a purchase on the Internet.
  - **Paid search.** The purchases of keyword(s) on search engines to display a relevant link/offer to a current or new customer.

Paid search are the links that display for keywords because they have been bought.

- **Owned.** If you have ever "built it because they will come," you know about owned media. It's what happens when you spend your capital and invest to create something fully or partially owned by your company in which you, the analyst, have some sort of stake—whether shares or salary.
  - **Sites.** Traffic from websites whether domains or subdomains that your company has created.
  - **Ad networks.** Loosely identifies every technology that in some way whether automated or manually buy, sell, match, target, or automate Internet-enabled advertising, including exchanges and DSPs and all types of targeting. Some companies own their own ad networks, whereas others pay for ad networks (and in this case, ad networks should go under paid or owned media as appropriate).
  - **CRM emails.** Customer Relationship Management activities and the digital outbound and inbound campaigns and other "customer touches" for establishing, maintaining, or enhancing customer relationships. CRM includes outbound marketing to named customers.
  - **Automated agents.** In 2013, it is entirely possible and not uncommon to receive automated offers for products or content that result from simple customer activities, such as opt-ins or complex machine-learned, algorithm processes—that customers may even consider spooky. These types of automated inbound marketing vehicles—from apps to emails—can be segmented here.
- **Earned.** Earned media is the good or bad content, branding, perception, mindset, awareness, social commentary and posting, and (dis)favorability and (dis)satisfaction that results from your business activities, including your products, your PR, and even your marketing both online and offline. Earned media is a

major contributor to the consumer perception of brand and is related to brand equity.

- **Organic search (SEO).** Although some would argue SEO is a paid activity, the ranking of your site and pages will happen whether you attempt to "game" the system or hire an SEO firm. The best SEO is always awesome, relevant content—not simply complex magic and monitoring from an agency or SEO firm. That said, SEO is altogether something that can be gamed, using SEO vocabulary in a black hat (if you do it, you can get penalized), gray hat (if Google finds out, they may not like it), and white hat (the SEO is acceptable) methods. Analytics enables you to know how you are doing—and even can tell you if you are using a black hat method for SEO when you should be using a white hat method. If you do SEO, it's earned media. But if you pay for someone else to do SEO, you may want to categorize it as paid media.
- **Social media.** Facebook, Twitter, Tumblr, YouTube, Slideshare, Pinterest, Foursquare, Groupon, and LinkedIn—much of the online world lurks or participates at levels in social media from highly engaged to barely listening. This segment captures all the traffic to your site that you earned from social sites and media.
- **Blogs.** Blogs are social media but are mentioned because blogs often exist outside of corporate products that provide an online user experience. It is also valid, of course, to call out other types of digital social media as necessary for your business goals.

- **Other sites.** Any site that links to your site can be measured by digital analytics. In some cases, technical externalities that are entirely outside of your or IT's control can prevent measurement of traffic from other sites that link into you sites—even the sites on which you spend money to advertise.
- **Direct/typed/bookmarked/unknown.** When traffic is from an "unknown source," it is frequently mislabeled at typed or

direct traffic. Although the traffic may be typed, direct, or bookmarked, analytics tools cannot measure this data. Traffic labeled direct, none, typed, bookmarked, and so on is more accurately described as "traffic from an unknown source." Use caution when applying this data to things such as brand and so on because this traffic isn't necessarily just from people who type in or bookmark your brand as a result of brand equity and top-of-mind awareness.

By understanding the common digital analytics concepts presented and their related and derivative definitions, you attain a vocabulary that enables you to understand, work with, analyze digital data, and communicate digital analysis that can help create business value.

Digital Analytics for Social Media: Social Analytics

Social media analytics (SMA) is a type of digital analytics. SMA represents a shift in the way analysts, brands, agencies, and vendors think about data, analytics, and research. Social media is a shift in the way people use the Internet and how digital data is measured. SMA helps you make sense of the shift in social media to focus on what is most important: generating economic value through analysis of the meaningful and relevant signals in social media data.

Social media, in the context of this chapter, refers to the massive directional change across the globe, in how people are creating, producing, editing, sharing, exchanging, conversing, collaborating, befriending, and consuming information. Social media, such as blogging, video sharing on sites like YouTube, Twitter, Facebook, Pinterest, Tumblr, and so on have become pervasive and persuasive in media and within social constructs (such as the family) during the last several years in ways only hypothesized by Futurists decades ago.

Early ideas on social media date back well before the dot.com implosion or early ideation—to radical thinkers like Ted Nelson who created the term *hypertext* (as part of his Xanadu project) and coined the term *intertwingularity* in the 1970s:

> *Intertwingularity* is a term coined to express the complexity of interrelations in human knowledge.
>
> Nelson wrote in *Computer Lib/Dream Machines* (Nelson 1974, p. DM45):
>
> EVERYTHING IS DEEPLY INTERTWINGLED. In an important sense, there are no "subjects" at all; there is only all knowledge, since the cross-connections among the myriad topics of this world simply cannot be divided up neatly.

I bring up Ted Nelson (and encourage you to review his work) because it is relevant to social media, which can expose you to seemingly wild concepts and technologies that can cause you to question their business validity. You can read other early thoughts about social media such as *The Cluetrain Manifesto* by Doc Searls or *The Tipping Point* by Malcolm Gladwell. These books, concepts, and social constructs help to form a basis for understanding social media, its underpinnings, and how to create economic value from it.

If Nelson's concept of intertwingularity appears irregular, odd, or simply dumb to you, then gather up your courage. SMA is a niche field of digital analytics with many oddly named concepts (for example, velocity, centeredness, and betweenness) that can (or cannot) be applicable to your business. I congratulate you on questioning the business applicability and validity of social media, but there is business value in social media. However, it requires rethinking and learning new ideas.

Social media analytics requires learning new vocabulary. Just like Facebook redefined existing vocabulary, such as "friends" and "like," and Twitter redefined the words "tweet" and "twitter," social media vocabulary runs the gamut from academic (betweenness—the amount of vector distance between nodes on a social graph) to slightly absurd (*com*versions—the number of conversions attributed to social media communication); thus, it is helpful to identify some, but not all, of the vocabulary helpful for understanding social media and social media analytics. It is likely you may have heard, used, defined, or criticized the usage of some of the (buzz) words discussed next. That said, social media is like the Internet in the mid-1990s; people are making up new words (neologisms and portmanteaus) quite regularly to explain new concepts. Even the names of social media sites are often neologistic: SnapChat, Tumblr, Liveleak, Facebook, Socialcast, and Pinterest.

Although new ideation in innovative digital spaces can be met with curmudgeon criticism from those entrenched in traditional media (or with vested interest in maintaining the digital status quo), it is important to keep an open (but critical) mind about social media and toward its vocabulary, concepts, and constructs. Social media has certainly mainstreamed—and the jargon will filter out (brick and click, anyone)—but for now, it is helpful to speak the same language to established shared economies and social markets that can help to generate digital commerce.

The following social media vocabulary is helpful for social media analytics:

- **Listening.** Refers to the process by which consumers, brands, or business actively monitors social media data using a social media tool or technology (free or paid) to learn and find out what's said in general or in specific, by whom, where, and why. Listening may be as simple as reviewing a Twitter feed or more complex to automatically categorizing massive volumes of incoming, text-based, social media data from multiple social media sites into logical categories using text and sentiment analysis.
- **Engagement.** Refers to the ongoing and continuous process of actively participating and communicating across one or more social media channels either in private (Facebook and Google Plus) or public (Twitter, outside of private accounts, and YouTube). Engagement for businesses could take the form of offering merchandise or promotions only on social media, responding to criticisms, actively messaging about new

products, or casually participating in the dialog speaking to customers on sites popular for customers of the brand. Engagement, at its most optimal, is strategic and planned.

- **Participating.** When your business starts and stops intermittently or keeps shifting social media strategies, channel switches, and generally is not consistent across social media, you and your business are not *engaging but are simply participating.* Participation can be planned, but if haphazardly periodic, it's not engagement.
- **Lurking.** Brands (and consumers) that spend resources (and thus time) passively observing what's going on in relation to their brand, products, and PR across one of more social media channels, without ever participating or engaging, are said to be lurking. Lurking means viewing social media content and knowing what's being said but not saying anything or participating. Lurkers lurk.
- **Social media platform.** The digital experiences that enable socialization on the Internet. The primary example of the social media platform is Facebook or Twitter. A platform is a strategic concept that is fundamental to social media business models. Although many social sites aspire to be platforms, only a few actually are platforms where other companies build on top of those platforms to create consumer ecosystems. Take Facebook, for example, where companies such as Zynga create billion-dollar businesses on top of the Facebook platform.
- **Social media tool.** Any tool, free or paid, that claims to perform a social media analytics function is a social media tool. There are thousands of social media tools—from simple online applications to complex software.
- **Word of Mouth (WOM).** WOM marketing is a concept known to global marketers. WOM is applicable to social media; thus, WOM measurement models have usefulness in social media analysis. After all, social media is inherently WOM communication but done online virtually and not in offline physical settings.

- **Virality.** The idea that concepts can be rapidly communicated quickly across great distances (and cultures) helps you understand what is meant by *virility* and *viral marketing*. When digital content "goes viral," it means that there has been a sudden and significant increase in viewing of that content in a relatively short period of time. For example, a video gets posted on Reddit and then gets 5 million downloads in 2 days. That video, thanks to Reddit, went viral.
- **Social bookmarking and linking.** Sites such as bit.ly and tinyurl.com to emailing, to SMS, to applications and widgets that send content (via links or text) to people you know (your "friends!") is what is commonly known as social bookmarking and linking. These sites and tools help social media users record and locate social experiences.
- **Social sharing.** When a digital experience enables content to be shared within or across a social experience, social sharing is supported. The set of icons and bookmarks—sometimes known as *story tools*—on sites that enable you to share on Twitter, Like on Facebook, or Post on Reddit are examples of icon-based sharing tools. People around the world are sharing whatever they can think of—from pictures of their vacation on Photobucket, to pictures of their children on Facebook, to pictures of a conference on Twitter, to random images found on the web on Pinterest.
- **Privacy.** The concept of privacy and how it is conceived, regulated, and legislated within your country is germane to social media. Scott McNealy, founder of Sun Microsystems (acquired by Oracle), said, "You have no privacy, get over it?" Was he right, wrong, or both right and wrong? Does it matter and so what? Social media is redefining the notion of privacy. How global governments, businesses, and consumers consider and respond to questions about the impact of social media on privacy are important for social media analytics.

- **Copyright laws.** Social sharing enables the sending of content of all types (from video to audio to text) in near real-time across tens of thousands of miles. In mere seconds, copyrighted data that cost millions to create can be copied, pirated, and sold. Consumers may or may not be aware of copyrights when using and sharing social media. Thus, the region, country, local, and jurisdictional laws, precedent, and rules that guide copyrights and frame the sale and sharing of digital content are important to consider for social media analytics. Legislation in the United States, like the failed SOPA, are attempts to address copyright concerns. Creative Commons is another way.
- **Influencers.** People who have a strong impact on the perception of your company, brand, products, or services care are *influencers*. Influencers have a given level of popularity and trust given to them by others. In social media, the opinion and sentiment of influencers expressed on social media has value. Influencers may have millions of Twitter followers or widely read blogs, but regardless, influencers are people who have the power to impact your business (both positively and negatively) due to their ability to impact considerably large or other influential audience with their social media handles/accounts, blogs, and other digital experiences.
- **Social networking.** A macro term from the sum collection of activities across social media, including social sharing, engaging, transacting, bookmarking, and linking done across social media platforms and measured across social media tools. Social network sites include the Big 3 (Twitter, Facebook, and Google Plus) and include hundreds of other sites from Weibo.com to Tumblr or Vkontake, Polyvore, and Nacer.
- **Social applications.** A new term, as of 2013, that refers to mobile applications that are primarily based on social and online collaboration over mobile devices regardless of physical location, which are also GPS location-aware. Social applications can also take some behavioral or other data input to create a social experience—for example, FourSquare.

Is Social Media Analytics About Brand or Direct Response?

Social media are business activities, and like all business activities, the work must produce profit and payback within a given period of time to continue to exist. Thus, like all other business activities, the digital analytics team is often called upon to validate the need for the overhead of social media and provide a fact-based business justification tied to analytical performance and profitability. Social media, which is often tied to Marketing or PR, must answer this question to qualify the type of analytics necessary to measure the social media campaign/product/experience and measure the business outcome.

It pains me greatly when I hear professionals—whether experienced, globally, or at Internet companies—reference either lightly or heavily that social media is an activity that needs no business return. And that social media is just a cost of doing business in today's world. That's not only naive, it is also wrong and dismisses the considerable ability of today's digital analytics teams and the tools they use to tie social media investment and work directly or indirectly back to financial measures. Social media, when deployed with measurement implemented to business goals, can provide the inputs for financial modeling, such as return on advertising spend (ROAS), return on marketing investment (ROMI), and other more general forms of financial impact analysis, such as return on investment (ROI), net present value (NPV), and internal rate of return (IRR) analysis. These analyses can help to identify to shareholders and management whether social media investment is paying back, and if so, when, and if not, if it will and an estimate of when.

Whether a business should be listening and engaging in social media is an "it depends" answer. Although the author of this chapter strongly believes that all businesses should judiciously use social media, it is obvious that every business has a level of social media listening and engagement that is right for them. Such decisioning and the impetus behind them are beyond the scope of this book, but to frame social media analytics, a business must define whether social media is to be used for one or both of the following:

- **Brand.** Social media channels (and their data) can be used to attract, message, encourage favorability, and promote satisfaction, act as a variable for media mix and other marketing models that help to identify brand equity, impact, and other higher-order consumer and shopping concepts. Social media analytics can provide input for qualitative research and VoC analysis.
- **Direct response.** Social media channels (and their data) can be used to engage directly and convert new and existing customers via promotions, offers, games, advertising, and other ways to directly interact and engage over the social channel. It is even possible to target types of customers from which you want a response and track them from social media to your website or other digital experience (such as a mobile application). Social media data can provide input for quantitative analysis and conversion measurement from direct response.

When determining a measurement strategy for social media, the strategy must be tightly aligned with the overall business goal of the social media campaign. Tying direct response to profitability is less challenging than it is tying the impact of digital advertising campaign on Facebook to a change in favorability or satisfaction based on indirect measurement. It is not uncommon for a digital analytics team to create frameworks, reporting, and analysis to support both social media branding and direct response activities and campaigns.

Social Media Brand Analytics

Social media is certainly influential to the establishment, creation, enhancement, and sustainability of brand and brand equity in the 21st century. Social media is essential to brand identification, awareness, and the resulting equity. If you or your brands aren't on social media, do you actually exist? Social media data would say you don't exist because you and your business have no traces.

Whether you are already a measurer and analyst of social media—or if you aspire to manage and understand social media analytics—it is an important activity to agree on definitions. A definition, and certainly not the only and an arguable definition, of brand follows:

> **Brand**—the sum culmination of all the inputs, thoughts, feelings, and experiences whether real, virtual, or imagined as perceived by an entity about another entity.

It is important to emphasize the words *entity* and *sum culmination* in the preceding definition because when measuring the notion of brand in social media, you must consider that brands are not just corporate, but personal and associated with people beyond celebrities. *Sum culmination* is simply the aggregate, the totality, and, is the purest measurement of brand, which is so ephemeral, yet it is what measurers and analysts both traditional and digital try to manage every day.

Consider the impact of social media reviews for hotels on Tripadvisor or Expedia. How do you read them? Do you believe them? How have such reviews impacted your decisions about commerce or influenced your conclusions about the brand?

The *Digital Analytics Framework for Social Media Brand Analytics* begins by understanding brand measurement attempts to quantify behavior in numbers that are qualitative in much of its nature which is often, at best, an educated estimate and, at worst, value destroying. Consider the following inputs when creating a social media analysis strategy:

1. Answer the following questions:
   - Do I want to impact the brand awareness, favorability, satisfaction, and equity?
   - Do I want to measure the behavior and outcomes of behavior across time?
   - Do I want to measure new or existing customers?
   - What financial metric (such as profit) do I hope to move through my brand analysis?

2. **Identify and document a full set of social media activities for the brand.** This task can be complex if social media activities are not centralized. Consult with your marketing department, PR, and finance team. This activity is not always straightforward or simple—and time should be budgeted for aggregating, synthesizing, and preparing this data for analysis.
3. **Identify and document the social media events and actions within the digital experience that you want to measure.** These events could be sharing content, engaging in commerce, or having additional conversations to direct responses from traffic referred from the social media site.
4. **Identify the financial model and relevant business justification for the social media campaigns.** Digital and traditional analysts must act socially when executing a social media analytics plan. The analytics team must put the data and its analysis in appropriate and relevant business context and do its best to tie it to financial measures.
5. **Assign values by tying the social media campaigns and each action to an estimated revenue value in the context of your brand.** Several methods exist for associating a monetary value with an event—from using spreadsheet and manual data entry to more elegant approaches in which behavioral data and events, transactional events, and financial value are integrated and reportable in a unified way directly into unified databases and reporting.
6. **Analyze the aggregated impact of total social media brand activities and segment them by customer segments, by each campaign, by each action, and so on.** Start with large, simple indicators of macro-level performance—the "KEY" Key Performance Indicators (KPIs). Then segment the totals into breakdowns and detailed segments that make sense both to your business and within the context of your industry.
7. **Compare the aggregate and segmented revenue from your social media brand activity to the estimated cost of that activity to calculate profit.** By associating both revenue

and cost data, however proxied, the analytics team can begin to create a model for estimating financial impact, in terms of profit, from social media.

8. **Communicate the results to stakeholders in the context of the initial financial model and business justification that underpinned social media strategy.** Again, the importance of having a socially-intelligent analytics team as well as a social analytics team that is intelligent about socializing with people is paramount. You and your team will have to leave their offices, cubes, and homes and communicate analytical results, outcomes, and recommendations to stakeholders and clients—or they will do it virtually. Regardless, the ability for an analyst to tell data stories in a way that makes sense (politically, emotionally, and organizationally) for communicating the analysis is a critical skill.
9. **Use the data collected for social media to impact other marketing models.** For example, brand trackers, media mix, attribution, and propensity models using statistically valid and rigorous methods. Social data can be used to inform brand trackers, brand equity studies, and predictive analysis.

Successful execution of the framework I've presented here is entirely possible to implement in a highly customized, value-generating way for your global, national, and local business—by yourself, your team, or with external help. Your mileage, however, will vary—and I presented this model to give you a sense of what is possible, not to tell you what to do. The science and art in the preceding framework is entirely in the execution and experience of the analytics leadership and team in the context of business questions and goals.

Social Media Direct Response Analytics

Many advocates of social media believe social media is responsible for a significant part of direct and indirect revenue for new and existing brands. Although this may be true for the cornucopia of new and interesting social media companies that hope to attract members, qualify their audience, sell it to advertisers, and thus create cash flow and value from a social audience, the impact of social media from direct response for large global companies (such as those traded and scrutinized by the capital markets) is not as large and can even be marginal or immaterial. In that marginality lies a huge opportunity to analyze the existing data from new opportunities for incremental revenue.

Although undoubtedly some legacy companies have had success driving new and incremental revenue from social media advertising and other activities in the social channel, only approximately $10 million of all the dollars generated on Cyber Monday in 2012 (the Monday after Thanksgiving in the United States) came from directly attributable social media. That said, revenue is being and can be generated on social media—especially from targeted and retargeted advertising on social networks—and that amount of revenue will only continue to increase

Direct response is any social media activity, beyond brand campaigning, which involves targeting (or retargeting) an entity and compelling that entity to perform a wanted action that creates value, such as conversion.

The vocabulary in the preceding definition is precise with several critical words: target entity, compelling, wanted action, and create value. It is critical to understand that social media direct response analytics involve:

1. **Targeting or retargeting entities.** Such as customers, segments, cohorts, and individuals—and even slicing and dicing these entities into their constituent attributes, such as demographics, psychographics, firmagraphics, or financial graphics (for example, Facebook's FBX).
2. **Compelling entities.** Via social media listening, participating, and engaging. The social media team must aspire to influence positively the business by creating compelling conversations,

content, and social experiences (both online and offline) that help to generate economic value by increasing brand equity or revenue (or reducing cost).

3. **Wanted actions.** The goals and steps in the conversion funnel, and events and actions in the social media experience that create the economic value. Wanted actions for direct response must tie to financial metrics, such as revenue and profit.
4. **Create value.** Either reduce cost or increase revenue such that it materially impacts the business in ways that can be measured and validated.

The *Digital Analytics Framework for Social Media Direct Response Analytics* articulates a series of steps that are necessary for measuring the digital direct response from social media within digital experiences. If you are not already collecting and analyzing social media data (and even if you are), any social media analytics effort for direct response must be coordinated with IT.

This framework involves executing through the following series of steps:

1. **Answer the following questions:**
   - **a.** Why does the social media exist for the direct response?
   - **b.** What is the goal the social media hopes to persuade the person exposed to it? (That is, what do you want them to do after exposure?)
   - **c.** What are the goals, funnels, and conversions within the social media channel itself (that is, in the advertising or on Facebook directly) and within the digital experience that is controlled by your business (such as your mobile application or website)?
2. **Identify and document a full set of social media activities that are used for direct response.** Consult with your marketing department, PR, and finance team. This activity is not always straightforward or simple—and time should be

budgeted for aggregating, synthesizing, and preparing this data for analysis.

3. **Identify and document the social media events and actions within the digital experiences that you want to measure as conversion, funnels, events, interactions, and goals.** For example, the events you might choose to track could include behaviors such as sharing content or initiating other conversations (new or related), the rate of sharing across and on other social media channels, buying a product, or requesting more information about a service.
4. **Identify the financial model and relevant business justification for the social media direct response campaigns.** This step includes gathering and making available for analysis all relevant cost data to understand the total cost of the social media, including headcount, overhead, fixed, and variable. Make sure to socialize the full allocated or apportioned cost and get the finance team's buy-in.
5. **Assign revenue values to the social media campaigns and each action and event within it.** Work with finance to identify a value for each activity in the social media direct response. The response is often a purchase, click, or other similarly measurable event. The way you do this work can be simple; use spreadsheets to do it manually or to automate the work into more complex integrated environments that use data warehousing to bring together financial data (with other data, such as digital behavioral data for analytics).
6. **Analyze the direct revenue of the total social media direct response activity for each campaign and each action.** By combining the financial event and response data with the behavioral data, you can estimate the actual business value—in actual dollars—of the direct response and upstream behaviors in social media.
7. **Compare the revenue of the total social media direct response activity for each campaign and each action to the cost to generate the profit.** This activity appears easy on

the surface—and can be, but be careful to validate the results and socialize them to key (but not all) stakeholders for their feedback after multiple cycles of data validation and analytical verification as you build the model and during user acceptance testing (UAT) before the production release.

8. **Communicate the results to stakeholders in the context of the initial financial model and business justification that underpinned the social media strategy.** The social in social media analytics is not just in the data, tools, and platforms. The idea of "being social" is necessary for analytics teams that must work in a consultative and collaborative fashion with their stakeholders. In that sense, the best analytics teams socialize analysis. See my book, *Building a Digital Analytics Organization*, for more information.
9. **Use the data collected for social media to impact other marketing models.** This data and analysis could be helpful for attribution, optimization, and predictive modeling.

Like the framework for Brand Social Analytics, the framework for Direct Response is complicated but can be easily adapted for small business as well as global companies traded and scrutinized by the capital market. Both frameworks are abstract enough so that they can be pulled apart or expanded to become part of the analytics process you create and customize specifically for your business.

The Incrementality of Social Media

The easiest way to illustrate the idea of incrementality in social media is with an anecdote. Pepsi has millions of fans on Facebook. How many of those Fans first were exposed to Pepsi's brand on Facebook or, as a result of seeing the Pepsi brand or direct response advertising on Facebook, went out and ordered Pepsi online or immediately at their local market? I would hypothesize not many. Data from audience measurement firms supports my view that new customers for global beverages are not a significant portion of the customers who like Pepsi on Facebook.

Pepsi's Facebook page does not receive as many monthly page views as fans who have liked it over time. There are millions of Likes, but only thousands of page views. But there is brand equity from the existence of this page—and maybe even a level of direct response to promotions placed on it. Also, because of the Like, whenever Pepsi wants, the company can message to Fans directly within its News Feed (as long as the person allows it via Facebook privacy). What that indicates at a high-level is that Pepsi has an opportunity for direct response but an even larger existing opportunity for brand. Thus, Pepsi, which is already one of the top-of-mind beverage brands for a social media exposed person on Facebook (because they already favor and have been satisfied with the beverage enough to Like it), has as excellent opportunity for going beyond consumer influence and use social media for shopper influence offline in-store. In other words, while Pepsi can build brand equity, they can also encourage people with coupons and discounts to buy Pepsi in the physical world. Thus, social media can help to inform insights into both brands and shoppers.

Thus, when tracking and measuring the impact and outcomes of social media using analytics, it is absolutely essential to consider not only Brand and Direct Response, but also the impact social media has on existing new purchases from brand advocates. This type of incrementality of both brand equity and revenue impact on existing sales is something that consumer brands must measure of the social channel on the business. This type of analysis is suitable for the digital analytics team.

Pepsi isn't selling a material amount of Pepsi directly on or even indirectly from Facebook (but has likely had some success with social media campaigns and promotional offers executed online and offline), but Pepsi is thriving on social media because it understands that social media is critical for established, mature global brands that sell in multiple countries, cultures, generations.

Thus, when measuring and tracking social media, measure the incrementality of social media—of incremental signups, registrations, downloads, white papers, members, and orders and attempt to tie the impact directly back to the bottom line using one of frameworks presented in this primer.

Digital Analytics for Mobile: Mobile Analytics

Mobile analytics, like social analytics and web analytics, is a type of digital analytics. Mobile analytics has much in common with site analytics—and it has certain uniqueness, peculiarity, nuances, subtly, and technologies that are specific to "mobile." Mobile analytics may include social analytics that occur on a mobile device. Any portion of the marketing funnel or tumbler can be touched by a mobile device, and any mobile application can touch other digitally enabled applications that include social features. See my book, *Building a Digital Analytics Organization*, for a description of the tumbler. Thus, although mobile analytics is similar to site (web) analytics in terms of data collection, basic concepts, measurement, reporting, and analytical approach(es), it has complexity in that mobile analytics must take into account two different constructs for mobile behavior:

- **Mobile sites.** Slimmed HTML (5), WAP, WML, or other Internet-enabled user experiences that can be rendered on any mobile device, which does not require a special installation of a mobile application, for engaging a visitor. These sites are most often indicated with an "m" in front of the top-level domain such as "m.somesite.com."
- **Mobile applications.** A stand-alone user experience delivered across one or more closed or open, proprietary systems, such as iOS, Android, and Windows, on which a digital experience is presented for engagement by a visitor. *Angry Birds* is an application. iTunes (the application, not the site) is also an application.

Measuring mobile applications requires following a similar decisioning methodology and process discussed in the Analytics Value Chain (reviewed in the next chapter). Analysis of mobile begins by asking the following questions:

- **Are you measuring a mobile site and/or mobile application?** Mobile sites tend to be slimmed-down websites with a user experience that fits the mobile screen. Functionality on mobile sites may be reduced when compared to the website. As a result, applied analytics on mobile sites can be similar to websites—and should take metrics, reports, and analytical approached from site analysis. Mobile applications can mimic and replicate site functionality, but often stand alone in terms of features, flows, and content from counterpart web or mobile sites. Mobile applications may also stand alone with no site counterpart. Some applications can spawn a browser when being used, which represents another type of mobile analysis—one where the crossover from application to site must be measured.
- **Why does the mobile site and application exist?** Critical and crucial to any measurement effort, and thus analysis, is defining the goal of the analysis. Tie the goals to finance and value creation.
- **Should you measure the same metrics across mobile experiences as you do for other experiences?** Mobile analytics does contain some different vocabulary, but not many beyond those custom metrics developed by your team for your business needs. Reconsider the metrics you measure online when determining an approach to mobile analytics, and reuse them as necessary.
- **Do existing data definitions, data, reports, and analytical approaches already exist in your company (or elsewhere) that can be reused?** Relevant tacit and explicit knowledge from experience in other projects should be applied to any mobile project. The analytics team might leverage existing

reporting, dashboarding, and formatting for analysis and customize to mobile business requirements.

- **Does the team have sufficient resources to engage in a mobile data collection, reporting, and analysis effort?** Mobile analytics requires new and different processes. For example, the data collection required may require specialized implementation using custom code and referencing embedded libraries.
- **Has the analytics team set expectations and gained approval from the other teams necessary for analytical success when executing a mobile analytics project (such as the business, IT, engineering, and QA)?** Mobile analysis, like any digital analysis, requires business requirements, data definitions, data collection, which needs alignment and coordination from the analytics team and sufficient resources. Senior management buy-in and support for mobile analytics will likely be necessary given the resourcing required. Due to the increasing importance of mobile to many sites, it is not uncommon for senior leadership to be involved at some level in a mobile analytics project. You also should consider the impact of measurement and how you may need to change measurement in the future—given submission and approval necessary when deploying mobile apps to application stores.
- **What are the key business questions that stakeholders who can actually make or influence changes to the mobile site and/or application need answered by the analytics team?** Like other digital analytics efforts, the capability of the analytics team to get the data, reporting, and analysis in the hands of the people who can action on it in a timely manner is important; thus, you must consult those stakeholders and determine their business questions and goals for mobile analysis. The resulting analysis then can be created to answer these questions and help the stakeholders.

- **What are the best possible ways for communicating answers to the business questions about mobile?** Communication strategies for mobile analysis should use existing processes—or if the mobile project is stealth or more specific to a functional area, like product or marketing, analysis may need communication in specialized ways, such as custom mobile-only reports, self-service environments for querying data, and custom analyses for "what if" and "what else" business questions.

A larger business question for the analytics team to help answer is, "Should my company be building a mobile application, a mobile site, or both?" For more elaboration and information about working with other teams and communicating analytics, please see my book, *Building a Digital Analytics Organization*.

Understanding Mobile Analytics Concepts

Mobile analytics has concepts that, like other digital analytics concepts, have no consensus-based standards and no widely agreed-upon definitions. Although mobile measurement definitions and standards can exist in the future based on industry agreement across vendors, consultants, practitioners, and global industry associations, they don't now; however, the Mobile Marketing Association (MMA) is attempting to do so. Here is a list of help mobile concepts and their definitions from real-world practical experience doing mobile analytics:

- **Screen view (or page view).** A more contemporary and modern term for viewing of content on the screen of a mobile device that may or may not be connected to the Internet at the time of view. While many businesses are still reconciling and struggling with web analytics concepts and are muddled in the technical quagmire of tagging—and many others are still learning how to understand the difference between visitors and visits and what page views are (were) in the world or RIA, AJAX, HTML5, and other rich, dynamic client-side experiences. Thus, while the analytics industry can be ready to adopt this term, the general stakeholder who consumes analysis may not be ready to use a new word to describe a page view on the mobile device screen.

- **Scan-through.** The transition in the user experience when scanning from a Quick Response (QR) code embedded either online or offline, digital, or traditional media to land on the resulting digital experience. A scan-through from a QR code to mobile app or site is homologous to the click-through from a marketing campaign to a landing page on a website.
- **API.** Application programming interface. Mobile data collection often requires the use of APIs specific to the vendor technology implemented directly into site code. APIs are used to collect data in mobile applications and RIA experiences when JavaScript is not possible for data collection. Mobile APIs are restricted only to a technology choice by the vendor. Software Development Kits (SDKs) are provided by mobile analytics vendors for data collection in languages such as Objective C, JSON, Web Services, REST, and so on. APIs can be public or private. Public APIs can be used by customers or partners, whereas private APIs are used by internal resources.
- **Crossover.** When a mobile site or application links back to a nonmobile site such that the mobile experience crosses over to the website and renders a similar, but different experience in a browser. For example, some mobile applications may require registration on the website to use the mobile application. In this case, user registration and sign-up may occur after clicking a link in the mobile application and open Safari or Chrome. This crossover from mobile to browser is necessary to measure whether this user experience is preferred or disdained.
- **Stitching.** As mobile devices move across geographies, the signal links to different towers and antenna on the same network. This movement can present challenges to tying the activity on the device to the mobile phone as it hops across these multiple towers. Given the issues with linking together what people do on mobile devices as those devices travel across distances, a

new vocabulary word was created for when the mobile analytics tool maintains tracking across hops. This word is stitching.

Other digital analytics and social analytics concepts defined in this chapter are relevant to mobile analytics and should be reviewed. Think how to apply the concepts outside of this chapter to mobile analytics. For example, mobile sites and applications have visits, visitors, events, interactions (like swipes), and conversions.

Why Mobile Analytics and How Is It Different from Site Analytics?

If mobile analytics shares concepts both technically and analytically with digital analytics, can similar analytical approaches and frameworks apply?

If the answer to that question is "yes," then why is it important to differentiate mobile analytics as a separate channel in digital analytics? First, the same way as social analytics, mobile analytics requires specialized knowledge and skills. In addition, mobile analytics is different from other channels for digital analytics, such as advertising, search, social, and sites for the following reasons:

- **Mobile analysis often involves location-based segmentation.** The ability to segment your mobile analytics data by country, region/province, city, state, ZIP code, DMA (Direct Marketing Area), MSA (Metropolitan Statistical Area), and other geographic constructs is necessary—and will likely be asked.
- **The data collection on mobile platforms and in mobile experiences tends to use more complex programming languages and methods for collecting data.** In most cases, mobile data collection involves using a Software Development Toolkit (SDK) with an API. It may even be necessary to add libraries or other code to your mobile application to cache data when the application is offline.

- **Mobile data collection and the integrity and accuracy of data (and thus the veracity and relevancy of any analysis).** Can be easily impacted by externalities related to carriers and wireless providers, such as signal strength or available bandwidth.
- **Mobile applications and other rich experiences may have features and capabilities that can be access offline (for example, Instapaper).** Applications enable richer and different functionality that is available on mobile websites.
- **Mobile analytics may require analysis that is almost if not totally as comprehensive as reporting and analysis for site reporting.** Be careful setting expectations on mobile analytics because stakeholders may want identical rigor and depth of analysis on the mobile experience as on the main site experience. In these cases, allocate resources where revenue is at risk and to the digital channel with the most profitability.
- **Mobile analytics are measured by internal tools deployed and used by analytics or technical resources in brands, and are also measured by external tools, such as audience measurement firms and third-party data vendors with a full complement of professional services.** Although vendors of digital analytics platforms, such as IBM, Webtrends, and Omniture, have features and capabilities for mobile analytics, a class of software exists named mobile analytics software developed specifically for the mobile channel. Explore standalone software or SaaS mobile analytics vendors, such as Localytics and Flurry, as well as the larger analytics vendors who offer mobile analytics as a complement to their large digital analytics offerings.
- **Mobile applications are not searchable—thus sentiment analysis and text mining may not be possible.** Android, iOS, and RIM all enable different levels of openness on their platform. For example, you can't search Google.com for text and content in iPhone mobile applications, can you? Android and

iOS have different levels of openness on their platforms, which impacts the way mobile analytical data has to be collected.

- **Mobile applications may have restrictions on the type of data collection. iOS has more stringent terms for mobile data collection than other platforms.** Collecting mobile data for aggregating the data and reselling it (audience measurement) can have obstacles on certain mobile platforms. Instrumenting mobile applications to collect data that helps the company improve a product or sell advertising is generally always allowed.
- **Mobile experiences can incorporate other digital analytics channels.** It is not uncommon for a mobile application to spawn a browser for enabling some functionality, which, for one reason or another, requires a browser and website. This use-case is called "crossover" and refers to the transition between the mobile application and site.
- **Mobile experiences contain advertising, which requires specialized measurement.** Just like mobile analytics is a subset of digital analytics, mobile advertising has its own theory and nuance when compared to online display advertising. It is important to incorporate data about mobile advertising into media mix models.

The Importance of Measuring Mobile

Mobile is part of digital analytics. It is simply another channel to consider in the media mix, just like television or radio. The importance of mobile is completely understated and not macromyopic in certain areas. One of those areas is advertising. Figure 1.1 shows the time spent on mobile compared to the time spent on other channels in the context of advertising spend in those channels.

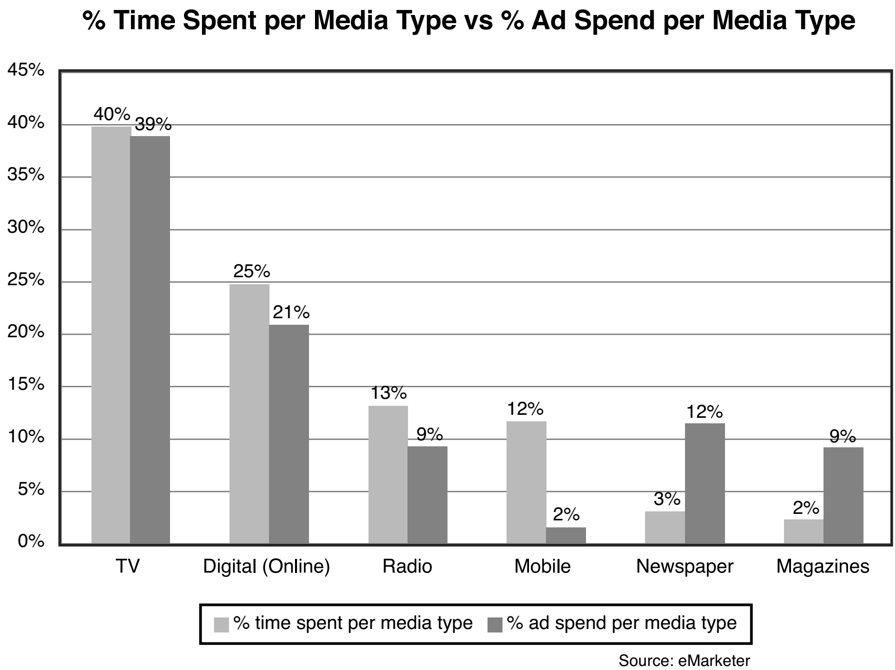

Figure 1.1 Ad Spend versus Time Spent per Media Type in 2012.

Notice the large gaps for the bars in the histogram for print (newspapers and magazines) and mobile. You could hypothesize from this data that advertising spent on the mobile channel will increase to close the gap between time spent on the mobile channel. The ubiquity of mobile devices, the affordability of wireless technology, the importance of mobile technology for global commerce, and the widespread availability of mobile networks and infrastructure across the Earth indicate that more capital will flow into the mobile channel—and especially into advertising. The gap between ad money spent on mobile and time spent on mobile will narrow to look like the relationship between money and time on television or radio.

In that gap between the small monetization of mobile and the large amount of time people spend using mobile devices exists a huge opportunity. In the closing of the gap lay potential huge revenue opportunities for companies innovating with big data science, statistical and machine learning, and data mining the mobile channel. This huge gap and revenue opportunity, and new market that simply can't be ignored. The mobile market worldwide is growing—and the analysis of the mobile channel when tied back to revenue, profit, and reduced expense is absolutely required to compete in the mobile space. Those companies who understand "why mobile analytics is important" will have a competitive advantage against competitors as they innovate the future mobile business, advertising, and digital experiences that will create economic value.

Digital Analytics Team: People Important for Analytical Success

The right people and leadership are absolutely critical to the success of any analytics team. Without people who not only understand the technical and the mathematical, but also the organizational and social components of telling stories using data, an analytics program can never run at full power, if it runs at all.

In the many years running analytics programs in large, globally distributed multinational companies or consulting with them, I found the wrong mix of people to be a recipe for disaster and stress for all involved. Analytics managers taking a role in which they must manage people to carefully consider the business requirements for the job and whether the current team (if you inherit a team) or the team you build must be based on skill sets and personalities that fit business goals and into the intrapersonal dynamics of the people on the team. In rare cases, you must fire people, so just do it.

Digital analytics professionals with experience managing and executing across the analytical value chain are uncommon (even rare) to find. When you have them, do your best to retain their talents. The key attributes of successful digital analysts include multidimensionality, numeracy, intellectual curiosity, and many more discussed in my book *Building a Digital Analytics Organization*. Although people's skills and expertise run the gamut from linear to quadratic, it is possible, even with an acute insufficiency of qualified digital analysts, to find the right mix of skills to build a digital analytics team and execute successfully—it just takes an investment of time and money to find these people.

Digital Analytics and Privacy

As business leaders and data analysts and scientists, it is not only your goal, but it is also your responsibility to protect every person's analytical data from misuse and abuse. Data analysts are largely left to self-regulate their analytical activities in much the same way the analytics tool vendors are left to self-regulate their innovations. That is why you must apply ethical principles to analytics, to ensure privacy and the application of analytics in a way that positively promotes, protects, and safeguards both global commerce and human society. In light of the 2013 acknowledgement that the U.S. NSA's PRISM application exists, it seems necessary for a global, societal conversation and consensus-driven alignment to occur for setting the boundaries of digital data collection, storage, and analysis. Digital analysts can be at the forefront of safeguarding online society in a manner that respects human and civil rights, self-reliance, and individualism while sustaining liberty, freedom, privacy, autonomy, and advancing the destinies of people.

Consider several rules when dealing with digital data, whether behavioral, transactional, qualitative, quantitative, mobile, social, video, first or third party, and private or anonymous digital data and metadata:

- **Be absolutely transparent about what data you collect and how you collect it by creating and frequently updating a Privacy and Data Usage policy and prominently displaying it on your site.** Write it in English, not legalese, and keep it simple, comprehendible, and summarized. If needed, link it to a more formal legal document.

- **Understand and provide, on request, a list of the tracking and measurement technologies currently deployed on your site.** Such a simple idea is hard to execute and deliver—especially at globally distributed enterprises—but smart companies should create and maintain a list of all social media (and digital) tracking and measurement technologies deployed on the site and have that list ready for review when requested.
- **Publish a simple metadata document that people both externally and internally can review that describes the digital data being collected and how it will be used.** For every technology deployed, the vendor should be providing a document answering the following questions: 1) What is this technology? 2) What data is being collected? 3) How is the data being used? and 4) How do I view, modify, and prevent my data from being collected? These answers can be used to craft your policy and privacy statements relevant to analytics that support it.
- **Create formalized governance around measurement, tracking, and advertising technologies and involve cross-functional representatives from teams across your company.** Companies that use analytics should have a governance council driven by their business, not the technology side. Teams from research, analytics, legal, marketing, sales, and technology should participate to ensure that best practices for protecting consumer privacy are practiced.
- **Enable easy and logical "opt-out" and, in the best case, only allow tracking and targeting to be "opt-in" for social media—and all digital tracking in general.** Don't create digital experiences that automatically apply new features and changes to everyone without gaining or at least attempting to gain their consent.
- **Eliminate all unnecessary data collection while regularly reviewing the data you have collected and delete unneeded data.** So much data can be collected, but little is actually useful, insightful, and actionable (UIA). Figure out

what UIA is generating profitable revenue; then determine what to do with the remaining data (from archiving it to deleting it).

- **Don't exploit new technologies in tricky ways that attempt to circumvent a user's choice or perception of privacy.** In other words, do not use technologies to reset or force the persistence of cookies after the user deletes them. Do not use hacks to store cookies forever when doing analytics and any other digital activity. Don't try to get around what the customer wants by technical wizardry. Don't violate the perception of privacy or civil rights or create situations in which your work could be considered unlawful, unethical, immoral surveillance.
- **Make your voice heard by writing your government, such as a senator or congressperson.** The heart of a democracy is the citizen's voice—whether you live in a democratic, free society or are reading this text in more creative ways in a country where freedom is limited. Imagine the potential for understanding and alignment that could be achieved if the thousands of readers of this book wrote an email, made a phone call, or advocated in the public domain positively for our industry and what we consider safeguards on and appropriate usage of digital data and analytics. If you do speak up and voice your opinion about the progression and evolution of your livelihood and business, no one is going to do it for you—but many will try to regulate your work as you choose to remain silent.

Digital analytics requires professionals who advocate for the people being tracked worldwide. Many of these people do not live in countries, or like I do, in the United States where we (or at least I) have a reasonable expectation of privacy. Although some may consider it beyond the scope of an analyst's role to focus on the ethical, proper, and private use of data, the preservation of privacy and anonymity—especially as we progress into the brave new world of the always connected, pervasive Internet we have today—is not only necessary but also critical for the preservation of freedom, liberty, autonomy, and greater personal power, education, and intelligence.

This page intentionally left blank

This page intentionally left blank

Index

B

E

S

T

U

V

W

X-Y-Z