ECOMMERCE ANALYTICS

Analyze and Improve the Impact of Your Digital Strategy

FREE SAMPLE CHAPTER

SHARE WITH OTHERS
Praise for *Ecommerce Analytics*

“Leveraging analytics to improve business results requires first knowing the questions that need to be answered. *Ecommerce Analytics* is the book to read if you are looking to use data to improve your online performance.”

—*Josh James*, Founder and CEO, Domo; Cofounder and Former CEO, Omniture

“With U.S. ecommerce sales set to cross $450 billion in 2017, the opportunity in front of you is immense. Yet, it is likely your conversion rates are stuck at 2%. I’m excited about Judah’s comprehensive tone because it is just what the doc prescribed to help unstick your ecommerce strategies. From cart abandonment to multichannel attribution to lifetime value… You’ll get precise guidance to win big!”

—*Avinash Kaushik*, Digital Marketing Evangelist, Google; author, *Web Analytics 2.0, Web Analytics: An Hour a Day*

“Are you a digital analyst working for an ecommerce company? Or are you an ecommerce marketer and your boss just asked you to provide more data about your digital efforts? If you are, then you should check out *Ecommerce Analytics*!

“Judah does a great job of making digital analytics for ecommerce logical and easy to understand. If you’re a digital analyst familiar with topics like attribution modeling, you’ll love how Judah discusses the specifics for an ecommerce business.

“If you’re new to digital analytics you’ll find the content easy to approach and very actionable. But that doesn’t mean that he cuts corners! Judah is really, really thorough! He takes the time to dive into all the different metrics and analysis techniques that you can perform on your ecommerce business.

“I put *Ecommerce Analytics* on my bookshelf and plan to use it whenever I work with an ecommerce company—you should, too!”

—*Justin Cutroni*, Analytics Evangelist, Google; author of *Google Analytics* and *Performance Marketing with Google Analytics*

“This important book is required reading for anyone who wants to understand how to deliver successful ecommerce analysis and data science. It’s instructive and helpful, unifying the subject matter in a way that is actionable for leadership, managers, technologists, and analysts.”

—*Raj Aggarwal*, Cofounder and CEO, Localytics

“Judah has created a must-read book for all digital analysts. It’s clearly framed and combines a comprehensive understanding of the topic with a practical flavor only the author can bring through decades of experience. This should be a hit in any college analytics class and will be on my graduate analytic course reading list in the future.”

—*Rand Schulman*, Managing Partner, Efectyv Digital; Cofounder, DealSignal; Cofounder, Digital Analytics Association

“In *Ecommerce Analytics*, Judah has delivered a comprehensive survey of the field, covering a broad array of topics important to implementers, analysts, and executives. This book contains a wealth of information that will be valuable in successfully executing an ecommerce strategy.”

—*Bob Page*, Director Emeritus, Digital Analytics Association; Internet entrepreneur
“Judah has done the impossible. In just a few chapters, he helps you understand the opportunity ecommerce retailers have in front of them, the monumental impact ecommerce analysis can have, the challenges you’ll face from a management and socialization perspective, and most importantly, how to navigate all of it and get stuff done.

“This book is a must-read for those who want to elevate their ecommerce stores to the next level, for the data teams who are tired of shouting into the wind, and for the managers who want to make a more meaningful impact on the organization. It is well researched, carefully thought out, and covers all of the bases. Stop reading this and buy the book already; your company will thank you.”

—**Tommy Walker**, Editor-in-Chief, Shopify

“Judah has delivered something critical in this remarkable information age—a broad, comprehensive summary of the essential elements of digital commerce. *Ecommerce Analytics* starts by showing what’s possible, in stark terms: the size and growth of the market, along with metrics of the leaders. From there, it elaborates on the entire taxonomy—defining the end-to-end process, what data to collect—how to analyze it—and then most importantly how to interpret it and optimize from there. He even digs into some of the trickiest bits: attribution, governance, and how to organize the company around digital success. Finally, his consistent emphasis on asking business questions and defining the outcomes and measurements before starting to crunch the numbers makes the subject truly accessible—and actionable.”

—**Sid Probstein**, Founder and CEO, RightWhen; Cofounder and former CTO, Attivio

“Ecommerce Analytics helps you understand each of the critical components needed to collect data, personalize interactions, and move customers through the funnel. It’s the backbone of marketing automation and personalization. This is the best book for anyone who works in ecommerce or who wants a job in that field where they will have to use data to understand and drive the business.”

—**Jonathan Corbin**, Director, Marketo

“Ecommerce: In 2016, everyone’s doing it; few do it well; even fewer share what they’ve learned, and only one makes the top of that list. Judah Phillips is that person, and this is that book. I’ve learned a ton from Judah, and you will, too.”

—**Cesar Brea**, Founder, Force Five Partners; author of *Marketing and Sales Analytics* and *Pragalytics*

“No one understands analytics more than my man Judah. This is a crucial read for anyone trying to step up their analytics game and who wants the tangible results that mastering data brings!”

—**Greg Selkoe**, Founder and CEO, Curateurs.com; Founder and former CEO, Karmaloop

“Ecommerce Analytics is a must-read for anyone involved in ecommerce. Not only will you learn how to do different types of analysis, it also covers the hard topics like data modeling, data integration, attribution, and building data-driven teams. You’ll even get some conversion optimization knowledge, too. You should read it before your competition does.”

—**Bryan Eisenberg**, Cofounder and CMO, IdealSpot; coauthor of the *Wall Street Journal* and *New York Times* bestselling books *Call to Action*, *Waiting for Your Cat to Bark?*, *Always Be Testing*, and *Buyer Legends: The Executive Storytellers Guide*

“This book is a must-read for anyone interested in making the most of their ecommerce-related data. The author does an excellent job making complex concepts accessible as well as teaching the reader how to get to actionable insights.”

—**Jesse Harriott**, CAO, Constant Contact
Ecommerce Analytics

Analyze and Improve the Impact of Your Digital Strategy

Judah Phillips
To Lilah and Elizabeth and my mom and dad, who all shop online and wonder what I do at work.
**Contents at a Glance**

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ecommerce Analytics Creates Business Value and Drives Business Growth</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>The Ecommerce Analytics Value Chain</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>Methods and Techniques for Ecommerce Analysis</td>
<td>27</td>
</tr>
<tr>
<td>4</td>
<td>Visualizing, Dashboarding, and Reporting Ecommerce Data and Analysis</td>
<td>71</td>
</tr>
<tr>
<td>5</td>
<td>Ecommerce Analytics Data Model and Technology</td>
<td>91</td>
</tr>
<tr>
<td>6</td>
<td>Marketing and Advertising Analytics in Ecommerce</td>
<td>103</td>
</tr>
<tr>
<td>7</td>
<td>Analyzing Behavioral Data</td>
<td>119</td>
</tr>
<tr>
<td>8</td>
<td>Optimizing for Ecommerce Conversion and User Experience</td>
<td>133</td>
</tr>
<tr>
<td>9</td>
<td>Analyzing Ecommerce Customers</td>
<td>155</td>
</tr>
<tr>
<td>10</td>
<td>Analyzing Products and Orders in Ecommerce</td>
<td>179</td>
</tr>
<tr>
<td>11</td>
<td>Attribution in Ecommerce Analytics</td>
<td>213</td>
</tr>
<tr>
<td>12</td>
<td>What Is an Ecommerce Platform?</td>
<td>229</td>
</tr>
<tr>
<td>13</td>
<td>Integrating Data and Analysis to Drive Your Ecommerce Strategy</td>
<td>241</td>
</tr>
<tr>
<td>14</td>
<td>Governing Data and Ensuring Privacy and Security</td>
<td>263</td>
</tr>
<tr>
<td>15</td>
<td>Building Analytics Organizations and Socializing Successful Analytics</td>
<td>279</td>
</tr>
<tr>
<td>16</td>
<td>The Future of Ecommerce Analytics</td>
<td>307</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>319</td>
</tr>
<tr>
<td></td>
<td>Index</td>
<td>329</td>
</tr>
</tbody>
</table>
Contents

Chapter 1  Ecommerce Analytics Creates Business Value and Drives Business Growth .......................... 1

Chapter 2  The Ecommerce Analytics Value Chain ................. 9
  Identifying and Prioritizing Demand ............................ 11
  Developing an Analytical Plan .................................. 14
  Activating the Ecommerce Analytics Environment ............. 16
  Preparing and Wrangling Data .................................. 20
  Analyzing, Predicting, Optimizing, and Automating with Data .................................................. 22
  Socializing Analytics ............................................. 23
  Communicating the Economic Impact of Analytics .......... 24

Chapter 3  Methods and Techniques for Ecommerce Analysis ....27
  Understanding the Calendar for Ecommerce Analysis .......... 28
  Storytelling Is Important for Ecommerce Analysis .......... 29
  Tukey’s Exploratory Data Analysis Is an Important Concept in Ecommerce Analytics .................... 31
  Types of Data: Simplified ........................................ 34
  Looking at Data: Shapes of Data ................................. 36
  Analyzing Ecommerce Data Using Statistics and Machine Learning .............................................. 47
  Using Key Performance Indicators for Ecommerce ........... 58

Chapter 4  Visualizing, Dashboarding, and Reporting Ecommerce Data and Analysis ............................... 71
  Understanding Reporting ....................................... 75
  Explaining the RASTA Approach to Reporting ............... 77
  Understanding Dashboarding ................................... 77
  Explaining the LIVEN Approach to Dashboarding ............ 80
  What Data Should I Start With in an Ecommerce Dashboard? ......................................................... 81
  Understanding Data Visualization ............................... 81

Chapter 5  Ecommerce Analytics Data Model and Technology ....91
  Understanding the Ecommerce Analytics Data Model: Facts and Dimensions .................................... 93
  Explaining a Sample Ecommerce Data Model .................. 96
Understanding the Inventory Fact ................................. 97
Understanding the Product Fact ..................................... 98
Understanding the Order Fact ........................................ 98
Understanding the Order Item Fact ................................. 99
Understanding the Customers Fact ................................ 99
Understanding the Customer Order Fact ........................ 100
Reviewing Common Dimensions and Measures in Ecommerce ......................................................... 100

Chapter 6  Marketing and Advertising Analytics in Ecommerce . 103
Understanding the Shared Goals of Marketing and Advertising Analysis .................................................. 105
Reviewing the Marketing Lifecycle ................................. 108
Understanding Types of Ecommerce Marketing ............... 111
Analyzing Marketing and Advertising for Ecommerce .... 112
What Marketing Data Could You Begin to Analyze? ... 116

Chapter 7  Analyzing Behavioral Data ................................. 119
Answering Business Questions with Behavioral Analytics . 123
Understanding Metrics and Key Performance Indicators for Behavioral Analysis ..................................... 124
Reviewing Types of Ecommerce Behavioral Analysis .... 126

Chapter 8  Optimizing for Ecommerce Conversion and User Experience ....................................................... 133
The Importance of the Value Proposition in Conversion Optimization ....................................................... 137
The Basics of Conversion Optimization: Persuasion, Psychology, Information Architecture, and Copywriting . 138
The Conversion Optimization Process: Ideation to Hypothesis to Post-Optimization Analysis .................. 141
The Data for Conversion Optimization: Analytics, Visualization, Research, Usability, Customer, and Technical Data ................................................................. 145
The Science Behind Conversion Optimization ............... 147
Succeeding with Conversion Optimization ..................... 151

Chapter 9  Analyzing Ecommerce Customers ..................... 155
What Does a Customer Record Look Like in Ecommerce? ................................................................. 156
What Customer Data Could I Start to Analyze? .......... 157
## Chapter 12 What Is an Ecommerce Platform?

- Understanding the Core Components of an Ecommerce Platform ........................................ 232
- Understanding the Business Functions Supported by an Ecommerce Platform .................... 235
- Determining an Analytical Approach to Analyzing the Ecommerce Platform. ....................... 239

## Chapter 13 Integrating Data and Analysis to Drive Your Ecommerce Strategy

- Defining the Types of Data, Single-Channel to Omnichannel ............................................. 243
- Integrating Data from a Technical Perspective ........................................................................ 246
- Integrating Analytics Applications ......................................................................................... 259
- Integrating Data from a Business Perspective ......................................................................... 261

## Chapter 14 Governing Data and Ensuring Privacy and Security

- Applying Data Governance in Ecommerce ............................................................................ 268
- Applying Data Privacy and Security in Ecommerce ................................................................. 272
- Governance, Privacy, and Security Are Part of the Analyst’s Job ......................................... 276

## Chapter 15 Building Analytics Organizations and Socializing Successful Analytics

- Suggesting a Universal Approach for Building Successful Analytics Organizations .............. 280
- Determine and Justify the Need for an Analytics Team ........................................................... 283
- Gain Support for Hiring or Appointing a Leader for Analytics .............................................. 285
- Hire the Analytics Leader ....................................................................................................... 287
- Gather Business Requirements ............................................................................................... 288
- Create the Mission and Vision for the Analytics Team .......................................................... 289
- Create an Organizational Model ............................................................................................ 289
- Hire Staff ............................................................................................................................... 291
- Assess the Current State Capabilities and Determine the Future State Capabilities .............. 291
- Assess the Current State Technology Architecture and Determine the Future State Architecture ......................................................... 292
- Begin Building an Analytics Road Map ................................................................................ 294
- Train Staff ............................................................................................................................. 294
- Map Current Processes, Interactions, and Workflows ............................................................ 295
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Build Templates and Artifacts to Support the Analytics Process</td>
<td>296</td>
</tr>
<tr>
<td>Create a Supply-and-Demand Management Model</td>
<td>296</td>
</tr>
<tr>
<td>Create an Operating Model for Working with Stakeholders</td>
<td>297</td>
</tr>
<tr>
<td>Use, Deploy, or Upgrade Existing or New Technology</td>
<td>298</td>
</tr>
<tr>
<td>Collect or Acquire New Data</td>
<td>298</td>
</tr>
<tr>
<td>Implement a Data Catalog, Master Data Management, and Data Governance</td>
<td>299</td>
</tr>
<tr>
<td>Meet with Stakeholders and Participate in Business Processes, and Then Socialize Analysis on a Regular Cadence and Periodicity</td>
<td>300</td>
</tr>
<tr>
<td>Do Analysis and Data Science and Deliver It</td>
<td>300</td>
</tr>
<tr>
<td>Lead or Assist with New Work Resulting from Analytical Processes</td>
<td>302</td>
</tr>
<tr>
<td>Document and Socialize the Financial Impact and Business Outcomes Resulting from Analysis</td>
<td>303</td>
</tr>
<tr>
<td>Continue to Do Analysis, Socialize It, and Manage Technology While Emphasizing the Business Impact Ad Infinitum</td>
<td>303</td>
</tr>
<tr>
<td>Manage Change and Support Stakeholders</td>
<td>304</td>
</tr>
<tr>
<td>Chapter 16 The Future of Ecommerce Analytics</td>
<td>307</td>
</tr>
<tr>
<td>The Future of Data Collection and Preparation</td>
<td>311</td>
</tr>
<tr>
<td>The Future Is Data Experiences</td>
<td>313</td>
</tr>
<tr>
<td>Future Analytics and Technology Capabilities</td>
<td>314</td>
</tr>
<tr>
<td>Bibliography</td>
<td>319</td>
</tr>
<tr>
<td>Index</td>
<td>329</td>
</tr>
</tbody>
</table>
Foreword

Let’s say that you fit into one (or perhaps more) of the following categories:

- You work at a sophisticated e-commerce company—say you are a new employee at Amazon.com—and you know your employer is very analytical, but you aren’t.
- You are a manager at a company with some e-commerce capabilities, and you have a nagging feeling that your firm could be better at analytics, but you don’t know how to go about it.
- You have an analytics-oriented job, but you need to learn more about e-commerce analytics.
- You work at a brick-and-mortar retailer that has dabbled in e-commerce, but you and your colleagues know you need to get more serious about it.
- You want to start a new e-commerce company.
- You are a student and want a job in e-commerce.

If you resemble any of these people, this book is the answer to your prayers. It is the one-stop shopping destination for everything you need to know about e-commerce analytics. It will provide a short refresher course on statistics. (Perhaps you did not fully engage with your statistics class in college.) It will give you enough content on retail and selling stuff online to get you by in most meetings. It will help you through difficult and esoteric topics like customer behavioral flow analysis, customer lifetime value, abandoned shopping carts, and virtually anything else you will encounter in this field.

As I read this book, I was surprised at how thorough it was and how clearly Phillips presents the key concepts. I was a bit amazed at the breadth of topics represented and that a young guy like Phillips has accumulated the analytical wisdom of the ages in topics like exploratory data analysis and communicating effectively about analytics. But I was truly shocked to read early in the first chapter that this is the only book on e-commerce analytics. Given the importance of e-commerce to the economy and the importance of analytics to e-commerce, I couldn’t believe that there are no other books on the topic. But I skeptically undertook a cursory search on “e-commerce analytics books” and found nothing in the category. There are books
on digital analytics (Phillips himself has written a couple of good ones), specific topics like conversion, and tools that are related to digital and ecommerce analytics (Google Analytics, for example), but literally nothing directly targeted at ecommerce analytics.

So if you have an interest in ecommerce analytics, it’s a no-brainer to buy and read the book. If you are reading this foreword in Amazon’s “Look inside” preview mode, go directly to the shopping cart and buy it. If you have actually found this book in a physical bookstore (though they should probably refuse to carry it since it threatens their business), take it up to the cash register and get out your wallet. I don’t know how much they will sell this thing for, but do you value your career? Do you want a promotion? You should not only buy the book but also send Judah Phillips 10% of your total compensation as a tribute.

You should also stop reading this foreword and get to the real content. But first, one more request. If you’re like many people, your primary interest in this book is in the “hard” advice on ecommerce analytics—statistical methods, techniques for getting data, specific key performance indicators to employ, and so forth. These are certainly important and valuable, and attending to them here makes perfect sense.

However, the book is also replete with “softer” advice that will make you a much better ecommerce analyst. There is, for example, the general philosophy of “getting close to your data” by observing distributions and outliers. There is the focus on telling a good story about your data. Phillips also maintains a strong focus on understanding what your organization wants from analytics and on making analytical results as digestible as possible. These types of wise perspectives and suggestions are just as valuable as any method or tool prescription.

So buy and read this book—every page of it. Take its recommendations to heart. If you want to be successful at ecommerce, you need a heavy dose of analytics. And if you want to be successful at ecommerce analytics, you need a heavy dose of this fine book.

—Thomas H. Davenport, President’s Distinguished Professor of Information Technology at Babson College; Research Fellow, MIT Center for Digital Business; Cofounder, International Institute for Analytics; Senior Advisor to Deloitte Analytics
Acknowledgments

This book is the first text written that is solely dedicated to explaining and exploring the subject of ecommerce analytics. The content is an unprecedented synthesis of experience and ideas, bringing together hard-learned knowledge about the business, technical, organizational, managerial, mathematical, scientific, artistic, and analytical aspects of delivering successful ecommerce analytics (and data science) in companies of all sizes. I wrote this book for analytics practitioners, managers, and leaders, as well as people who work with, who are studying, or who simply want to learn more about ecommerce and how to analyze it. It’s a multidisciplinary text that crosses a range of topics and contextualizes them under the idea of “ecommerce analytics” and thus unifies the content into a cohesive whole. I cover analytical concepts and theory, to applied mathematical and statistical methods, to analytical processes and organizations, to specific types of ecommerce analysis, to data modeling, integration, collection, storage, security, and privacy. I tried to be as ambitious and innovative in writing this text as the ecommerce industry, its leaders, and its workers are daily in their jobs. Thank you for reading my book.

First I’d like to thank Jeanne Glasser-Levine, Betsy Gratner, Cheri Clark, Amy Neidlinger, and the staff at Pearson who helped me produce this book. I’d also like to thank the following people (in random order): Thomas Davenport, Jonathan Corbin, Jeff Bezos, Brian Massey, Sid Probstein, Julio Gomez, Sergio Maldonado, Sean Hickey, Sarabjot Kaur, Peep Laja, Harpreet Singh, Shail Jain, Nick Kramer, Ants Anupold, Dave Weineke, Peter Gibson, Deepak Sahi, Jeff Evernham, Gary Rush, Vishal Kumar, Eric Valyocsik, Mike Taubleb, Dennis Mortensen, Gillian Ahouanwoheke, Kevin Rowe, Patrick McGinley, Ali Benham, Feras Alhou, Eric Fettman, Sergei Dvoynos, Amit Shah, Dave Munn, Jesse Harriott, Bart Schutz, Ton Wesseling, Raj Aggarwal, Jim Levininger, Tommy Walker, Justin Cutroni, Avinash Kaushik, Brian Clifton, Eric Dumain, Cesar Brea, Brett House, Olivier Titeca-Baeport, Bala Iyer, Eric Colsen, Julien Coquet, Bob Mould, Jeff Checko, Bob Page, Jascha-Kaykas Wolff, Gary Angel, Kim Ann King, Eric Hansen, Gary Angel, Rand Schulman, Andrew and Luchy Edwards, Jim Sterne, Mark Zuckerberg, Bryan and Jeffrey Eisenberg, Josh James, and you.
About the Author

Judah Phillips helps companies create value with analytics and data science by improving business performance. Judah has led analytics and data science teams for Fortune 500 companies and has improved their financial performance through the applied analysis of data, the management of analytical and technical resources, and the alignment and optimization of analytics strategy against short-term roadmaps and long-term strategic visions. Judah strongly believes that cutting-edge technology is critical and necessary but often becomes technical overhead unless strategy is aligned with excellence in organizational development, operational management, and delivery execution that is solidly tied to impacting material financial goals. Judah has worked for or been hired as a consultant by Internet companies, media companies, consumer product companies, financial services firms, and various types of agencies.


• He served on various boards of or advised established and start-up technology companies, including global leaders in digital analytics, mobile analytics, ecommerce, mobile apps, and advertising technology.

• He is an Adjunct Professor at Babson College and has guest lectured on analytics and data science at the business schools for New York University, Boston College, Northeastern University, and others.

• He is the former V.I.P. at Harvard Innovation Lab, where he advised Harvard start-ups about analytics and data science.
• He has spoken at more than 70 technology and industry conferences since 2006.

Judah holds a master of science in finance and a master of business administration from Northeastern University and a B.A. from the University of Massachusetts Amherst.
Ecommerce Analytics Creates Business Value and Drives Business Growth

The global ecommerce market is expected to grow at a compounded annual growth rate of 17% from $1.3 trillion in 2014 to $2.5 trillion by the end of 2018. In the United States in the third quarter of 2015, ecommerce generated $87.5 billion and accounted for 7.4% of all retail sales (Rogers 2015). Ecommerce has been growing annually on average between 14% and 15% quarterly since 2014, while retail growth has remained less than 3%. comScore estimated that U.S. consumers spent more than $57 billion online from November 1 through December 31, 2015, up 6% from 2014. On Cyber Monday, U.S. consumers spent more than $2 billion online (comScore 2016). Alibaba, in China, reported more than $14.92 billion in “goods transacted” on one popular shopping day (Denale 2015). Amazon’s Q3 2015 ecommerce revenue grew 23% from 2014 to more than $25 billion (SEC 2015). Frost and Sullivan predict that by 2020, the business to consumer (B2C) ecommerce market will be $3.2 trillion and the business to business (B2B) ecommerce market nearly twice as large at $6.7 trillion. Nearly $10 trillion in ecommerce revenue will occur by 2020. Globally, the United State and China are the largest ecommerce markets, accounting for more than 55% of ecommerce sales in 2015. eMarketer estimates that by 2018, China’s ecommerce marketing annually will be more than $1 trillion, with the U.S. likely reaching $500 billion, followed by the U.K. at $124 billion and Japan at $106 billion (Rogers 2015). Clearly, huge amounts of goods and services are being transacted between businesses to consumers and between businesses and other businesses globally.
Ecommerce analysis will continue to be an important activity for generating such growth and new levels of revenue. Puma generated a 7% conversion lift using analytics. PBS increased conversions and visits by 30% using analytics to track customer events in the funnel. WBC cited a 12% boost in conversion rate through customer segmentation. Watchfinder claimed a 1,300x increase in ROI remarketing based on analytics. Marketo claims a 10x higher conversion rate for personalized campaigns using analytics. BT used conversion testing to increase form completions by more than 60%. Amari Hotels increased online bookings and sales by 44% by using analytics to optimize online advertising (Google 2015). Companies that do ecommerce analysis increase their business performance.

Ecommerce is transacted on pure-play B2C ecommerce sites that have no physical storefront, such as Zulily, eBags, and Wayfair, and by omnichannel B2C retailers, such as Walmart and Staples, that have physical stores. Even pure-play ecommerce sites, like Amazon and Warby Parker, are opening stores. As a result, existing companies that already sell goods and services are now selling online and vice-versa. New companies are almost required by the market to have an online presence. Although some companies that sell physical products use an online presence only for branding to drive offline sales, that’s increasingly rare. Even luxury brands are selling their goods directly to consumers on retail sites. B2B ecommerce is even larger than B2C ecommerce. Major global companies execute ecommerce, such as Ford, GM, Coca-Cola, Chevron, IBM, General Mills, Kraft Heinz, ExxonMobil, General Electric, and Microsoft. The largest 300 B2B ecommerce companies were projected by eMarketer to grow 13.3% this year to $547 billion (from $483 billion in 2014)—figures that easily eclipse the U.S. B2C ecommerce market.

Ecommerce isn’t just about the site anymore. The most popular ecommerce sites have a mobile experience, whether mobile web or mobile app. 30% of U.S. ecommerce sales in 2015 were generated on a mobile device (Brohan 2015). Many ecommerce sites also have physical stores. And in the future, ecommerce will be embedded into “things” and pervasive in Internet-connected devices—with mobile payments just a touch away both online and in-store. Internet Retailer predicts that in 2015 the U.S. mobile commerce sales will total $104.05 billion, which is up 38.7% from $75.03 billion in 2014.
They estimate that mobile commerce in 2015 will grow 2.58 times faster than desktop ecommerce sales, which they predict will grow 15% this year to an estimated $350.64 billion globally. Note that 30% of mobile customers leave an ecommerce site when it is not optimized for mobile (Dorian 2015).

Ecommerce is an extremely competitive space. It takes huge amounts of capital to even try to compete with the major ecommerce players. This competition can create razor-thin margins or revenue that can be driven primarily through discounting. Ecommerce can be considered a zero-sum game. Thus companies are competing by creating digital experiences that enable a person to quickly and easily find and buy. Whether on a desktop, tablet, or mobile device, the companies that are winning in ecommerce make it easy and frictionless to find the product or service desired, understand how it fits the need, and buy it. Then these sites can compel their customers to come back again and again to buy more online and in-store. To do so, ecommerce companies use marketing and advertising that is tightly coupled with a user experience that ladders up to a prospect or customer’s notion of the brand and works to meet their intent. People come to ecommerce experiences with certain goals in mind: to learn more about a product by reading product information and social reviews, to compare prices and promotions, and to purchase products. Ecommerce sites that win at this zero-sum game can match that intent to a product and create commerce.

Leading ecommerce companies use data and analytics to compete—and they use a lot of different data to do so. Data is collected and analyzed about who visits an ecommerce site, when they visit, what pages they view, and what site or source they came from (the referrer or marketing channel). Other information is also collected about user behavior, such as user interactions and events on the site, data related to products viewed, promotions used, pages visited, time spent, the different paths and clickstreams on the site, the queries entered in search, and many other data points, such as the order value, the price of products, the shipping method used, and the payment information. Customer data may be captured or inferred, such as who the customers are or could be, where they live, what they like and their preferences or propensities, what they’ve bought, and other demographic and psychographic information.
The idea of “conversion”—where a prospect transitions to a paying customer—is embedded into the analytical DNA of the world’s leading ecommerce companies. They staff entire teams and run comprehensive programs for conversion testing and optimization. Marketing channels and sources of traffic, such as organic and paid search and various types of online advertising, are measured and tracked. Higher-order consumer research around brand awareness, favorability, and consideration is performed. Customer data is analyzed, segmented, grouped into cohorts, modeled, and understood using financial measures, like the cost of customer acquisition and customer lifetime value. Customer loyalty, retention, satisfaction, and churn are known and optimized. Merchandise, products, orders, and transactions are analyzed from the site to the warehouse through to shipping and fulfillment.

All this different quantitative and qualitative data about the entire ecommerce experience and operations can be captured, measured, and analyzed to improve business performance and make better decisions. Although tracking, measuring, and analyzing all of this data may sound challenging—and it is—it is possible to do. Of course, doing so isn’t easy. It requires investment in people, first and foremost, who understand business, technology, and the process of doing analytics. It also requires investment into different types of analytical tools and technologies, including ecommerce platforms, business intelligence tools, analytical platforms, and data science sandboxes. It might even require the collection of new first-party data, the usage of second-party data, or the purchase of third-party data.

All of this data, the people and teams who work with and analyze it, and the technology supporting it represent powerful assets for ecommerce companies to use to help run their business. But the data must be collected and analyzed effectively and accurately for companies to use it to create better experiences, make better decisions, drive conversion, satisfy and retain customers, and thus increase revenue, growth, profitability, and value. The effective use of ecommerce data and related data requires investing in the analytics value chain—from the technology to the people to the processes, governance, and change management necessary. Doing so can provide a material return on investment from analytics by converting more users
to customers and providing insights that can be used to improve the customer experience. The return from analytical investment can also come from improving marketing operations and tracking the cost and return of marketing and advertising. The impact of merchandising programs can be attributed to sales and other financial metrics. The details of transactions, the metrics around products, and the key performance indicators related to the shopping cart can be understood, benchmarked, and targeted with goals. These methods for competing with ecommerce data are entirely possible if you know how to succeed with ecommerce analytics.

Ecommerce analytics is the phrase used to describe business and technical activities for systematically analyzing data in order to improve business outcomes of companies that sell online. This broad definition incorporates business activities such as the gathering of business requirements, the execution of analytical programs and projects, the delivery and socialization of business analysis, and the ongoing management of the demand and supply of analysis. The range of business stakeholders demanding service in ecommerce companies will run the gamut from the C-suite to the leaders of merchandising, buying, planning, marketing, finance, user experience, design, customer service, inventory, warehousing, fulfillment, and more.

Ecommerce analytics also involves working with IT and engineering teams in the appropriate software and Internet development lifecycle. It requires the analytical team to participate and possibly lead technical activities that are required to deliver or support analysis, such as data collection, extraction, loading, transformation, governance, security, and privacy. Ecommerce analytics can include understanding and doing dimensional data modeling, working with databases, handling data processing, creating and executing querying, determining data lineage, participating in data governance committees, acting as a data steward, working with and defining metadata, and using tools to analyze data, create data visualizations, and do data science and advanced analytics. All of this work occurs within a corporate organization with its own culture and ways of working, into which the analytics team must integrate and learn to support and guide to drive data-informed business outcomes. Successful analytics often requires rethinking and reorganizing the way a company is structured,
including new roles in the C-suite, such as chief analytics officers, chief data officers, and chief data scientists.

Companies that are successful and effective with ecommerce analytics ask business questions that can be answered with data, and then they employ analytical teams that can collect and acquire data, govern and operate analytical systems, manage analytical teams, and generate analysis and data science that inform stakeholders. These companies create value by asking questions, answering them with data, and changing the way they take action as a result. Business questions for ecommerce can include the following:

- **Customers**: What are the characteristics of my most loyal customers? Least loyal?
- **Marketing**: How do customers perceive our company and products?
- **Categories and products**: Which products drive the most sales or highest gross margin? Which products are frequently purchased together?
- **Price and promotions**: What impact do discounts and promotions have on overall sales?
- **Omnichannel sales**: How are my channels performing and how do they complement each other on the path to purchase? How does this differ from attribution?
- **Prospects and customers**: Which prospects should I target to convert into loyal customers? What products or offers would be most effective?
- **Optimization and prediction**: What parts of my site should I test? What products should I order now to match sales forecasts?

Many other questions can be asked to help guide and drive business performance; ecommerce analytics leads to asking a lot of questions. The analysis of ecommerce is complex not only because it crosses both business and technology, but also because it is on the forefront of digital experience and innovation. The site, mobile, and connected ecommerce experiences online in 2016 are innovative, fast, personalized, contextual, and powerful for guiding us to the right
product, at an appealing price, and then leading us through a purchasing process that is easy and frictionless. But in certain cases, the opposite is true. Ecommerce sites and experiences have many opportunities to improve. They may be hard to navigate or may make it difficult to find product information. The trustworthiness of the site may be in question. The experience of selecting products, adding them to the cart, and stepping through the shipping and payment pages may be problematic, confusing, or in the worst case dysfunctional for the device or browser the person is using. In addition, the people working at ecommerce companies may largely be unaware of these problems because they aren’t getting timely, complete, and relevant data and analysis to help improve the experience and increase conversion. Or ecommerce stakeholders might be suspicious of experiential or customer issues but can’t prove them using data. Or there’s the worst case, where no set of unified resources, technologies, or analytics team exists to help stakeholders. What’s missing at these companies that aren’t taking full advantage of the information and insights in their data is solid, focused ecommerce analysis that helps business stakeholders do their jobs better. Whether that job is to merchandise the site, improve the user experience, drive customer acquisition, increase conversion, manage orders and fulfillment, or maximize customer profitability and shareholder value, ecommerce analytics can be a successful competitive advantage.

This book was written to help both new and experienced analysts succeed with ecommerce analytics. It was also written with the understanding that people who work at ecommerce companies in non-analytical roles, or who are simply interested in the topic, may read this book. Thus it is structured to guide the reader into the topic by first reviewing ways to think about doing ecommerce analytics as part of what I call a “value chain.” Methods and techniques for doing analysis are discussed in detail for both the new reader and experienced analyst. Reporting, dashboarding, and data visualization for ecommerce are explored. Data modeling is reviewed, including a discussion about dimensions, facts, and metrics. Several chapters are dedicated to detailing the what, why, and how of useful types of ecommerce analysis executed for marketing, advertising, behavior, customers, merchandising, orders, and products. The sciences of conversion optimization and attribution are discussed. Guidance on
building effective and high-performing teams is provided. Data governance, security, and privacy of ecommerce data and what the future holds for ecommerce analytics are deliberated. The comprehensive scope of this book offers an experienced practitioner’s perspective and viewpoint into ecommerce analytics across multiple dimensions: business, management, technology, analytics, data science, and the ecommerce domain. Although more content and detail can always be added in future volumes, the broad and ambitious subject matter discussed is unprecedented. This book offers a view into ecommerce analytics that hasn’t before been consolidated nor unified into one source. Whether you read this book as a standalone text or in combination with my other books, *Building a Digital Analytics Organization* and *Digital Analytics Primer*, you will develop and enhance your understanding of ecommerce analytics, the business and competitive opportunities it enables, and how to use analytics to take advantage of them.
Index

Numbers
4-5-4 retail calendar, 28

A
abandon cart rates (KPI), 63-64
abandonment rates (behavioral analytics), 125
ACID (Atomicity, Consistency, Isolation, Durability), 179
acquisition costs (customer analytics), 168-169
acquisition marketing, 112
action logs, 24
activation marketing, 111
ad hoc reporting, 76
administrating data in analytical environment, 19
advertising
  advertising analytics, 103-105, 112-113
    affiliate analysis, 114
    attribution, 115
    audience analysis, 115-116
    campaign analysis, 113
    conversion optimization, 134
    customer journey analysis, 116
Customer Origin reports, 116-117
ecommerce and user experience, 3
email analysis, 114-115
focus groups, 107
goals of, 105-108
interviews, 107-108
marketing goals, 105-108
marketing mix-modeling, 115
necessary data in, 116-117
online display ad unit analysis, 116
panels, 108
projective techniques (stimuli), 108
qualitative analytics, 107-108
SEM, 113
SEO, 113-114
social media analysis, 114
surveys, 108
display advertising and attribution, 218
affiliates
  analysis of, 114
  marketing and attribution, 218
  orders and, 182
Agile environments (software solutions), 248, 250
dashboards, 251
data models, 250
data stewardship, 251
data visualization, 251
reports, 251
scrums, 247-250
AI (artificial intelligence), 317
Aksoy, Lerzan, 172-173
algorithmic (mechanistic) analytics, 302
analytical dashboards, 79
analytics, 4-8
analysis techniques
  action logs, 24
capital budgeting models, 25
confirmatory analysis, 22
descriptive analysis, 22-23
explanatory analysis, 22
exploratory analysis, 22, 31-33
predictive analysis, 22
prescriptive analysis, 22-23
revenue/cost tracking, 24-25
socializing analysis, 23-24
stakeholders and financial impact assessments, 24
analytical environment
activating, 16-20
data administration, 19
data analysis tools, 19-20
data collection, 19-20
data processing, 19
data querying, 19
data sources, 19
data storage, 19
data virtualization, 19
elements of, 18-20
presentation tools, 20
tools versus technology, 16-18
visualization tools, 20
analytical plans, developing, 14-16
asking questions, 6-7
Bayesian statistical analysis and EDA, 33
binomial distributions, 52-53
conversion optimization, 141-143, 145
correlating data, 47-49
data integration, 259-261
data preparation tools, 20-22
defining, 5-6
economic impact of, 24-25
EDA, 31-33
experimenting with data, 53-56
future of, 307-310, 314-315, 318
AI, 317
cataloging data, 311-312
collaboration, 316
conversational speech technology, 316
data collection, 310, 311-313
data experiences, 310, 313-314
data flows, 315
data lakes, 311
data pipelines, 315
data preparation, 312
encryption, 315-316
machine learning, 317
natural language searches, 316
new data sources, 312-313
predictive analysis, 317
profiling data, 311-312
self-service automation, 316
hypothesis testing, 52-53
maps, 294
platforms, 293
road maps, 294
sampling data, 54-56
socializing, 23-24
storytelling and, 29-31
teams, 279-280, 300-301
algorithmic (mechanistic)
analytics, 302
analytics platforms, 293
analytics road maps, 294
architectures, 292
artifacts, 296
build templates, 296
building (universal process), 280-283
business outcomes from, 303
business processes analysis, 300
causal analytics, 301
cataloging data, 299
change management processes, 304-306
checkpoints, 302
communication, 292
current state assessments, 291-293
data governance, 292-293, 299-300
data models, 293
data ranges, 292
data storage, 293
data visualization, 293
delivery systems, 292
descriptive analysis, 301
education, 304
exploratory analysis, 301
financial analysis, 303
funding, 291
future capabilities, 291-292
future state architectures, 292-293
gathering business requirements, 288-289
hiring leadership, 285-288
hiring staff, 291
inferential analytics, 301
interaction maps, 295-296
leadership, 285-288
need for, 283-285
new data collection/acquisition, 298-299
new work, 302-303
newsletters, 304
non-team meetings, 302
office hours, 302-303
organizational models, 289-291
planning committees, 305
postmortems (projects), 305
predictive analysis, 301
prescriptive analysis, 301
process maps, 295
quarterly reviews, 305
results, 292
scope, 291
scrums, 305
self-service automation, 293, 304
SLA, 298
social analysis, 300
sponsorship, 291
staff, 291, 294-295
stakeholder operational models, 297-298
stakeholder support, 304-306
supply/demand management models, 296-297
technology architecture assessments, 292-293
ticketing systems, 302
use/deploy/upgrade strategies, 298
value of, 291-292
vision of, 299
workflow maps, 296
training staff, 294-295
value chain, 9-20, 23
analytics layer (ecommerce platforms), 230
Analytics Tier (ecommerce), 20
Angel, Gary, 305
anonymous customer data and data governance, 266
ANOVA (Analysis of Variance) tests, 39, 49-50
Applied Regression: An Introduction (Quantitative Applications in the Social Sciences), 50
architectures
analytics teams, 292
future state architectures, 292-293
technology architecture assessments, 292-293
area charts, 44
artifacts (analytics teams), 296
Ash, Tim, 149-150
assessments
analytics teams
current state assessments, 291-292
technology architecture assessments, 292-293
financial impact assessments, 24
use/deploy/upgrade strategies, 298
attribution, 213, 217
affiliate marketing, 218
blogs and, 219
bottom-up, 224
budgets and, 214
buyers, determining, 217-220
conversion rates, determining, 217-220
custom-weighted attribution, 222
defining, 213
direct traffic and, 220
display advertising and, 218
earned media and, 219-220
e-mail, 219
engagement mapping, 220-221
first interaction attributes, 221, 223
importance of, 213-214
“last click” modeling, 215-217
last interaction attributes, 221, 223
last nondirect interaction attributes, 221, 223
latency scoring, 222, 224
linear attribution, 221, 223
logistic regression, 222, 224
marketing mix-modeling versus, 214-215
marketing/advertising analytics, 115
media mix-marketing versus, 214-215
owned media and, 219
paid media and, 218-219
paid searches, 218
pathing, 222
position-based attribution, 221, 223-224
profits, determining, 217-220
proportional hazard modeling, 222-223
referral traffic and, 219
retargeting and, 219
INDEX

revenue, determining, 217-220
SEO, 219
social media, 219
software, assessing, 225-227
stacking, 222, 224
starter-player-closer attribution, 221-222, 224
time decay attribution, 221, 223
top-down versus bottom-up approaches, 224-225
types of, 221-224

audiences
  audience analysis, 115-116
  audience data and data integration, 245
  automation (self-service), 293, 304, 316
  averages (KPI)
    Average Revenue Per Order, 184
    examples of, 64
    order value, 64
  awareness/exposure marketing, 111

B

B2B (business to business) data and data integration, 245
B2C ecommerce, 2
Bain and Company, 164
bar charts, 85
batch processing and data integration, 261
Bayesian statistical analysis, 33, 39, 147
behavioral analytics, 119-120
  abandonment rates, 125
  answering questions via, 123-124
  behavioral flow analysis, 127-130
  checkout abandonment rates, 125-126
  content analysis, 131
  conversion rates, 125
  correlating data, 126
  framework of, 120-122
  frequency (time between visits), 124
  KPI, 124-126
  metrics, 124-126
  ordered behavioral analytics, 122
  in-page (on-screen) behavior analysis, 132
  recency (time since last visit/loyalty), 124
  shopping behavior analysis, 130-131
  types of, 126-132
  unordered behavioral analytics, 122
  visits (sessions), 124, 125
behavioral data and data governance, 267
behavioral triggers, predicting by
  (customer analytics), 174-175
  binomial distributions, 52-53
  bivariate data, 34
  block plots, 39
  blogs and attribution, 219
  bounce rate (KPI), 64-65
  brands
    brand analysis, 186, 188-189
    order data models, 183
    product brand analysis, 194-195
  brokers (intermediaries), 74
  budgets and attribution, 214
  build templates, 296
  Building a Digital Analytics Organization, 280
bullet charts, 46
Buoye, Alexander, 172-173
business intelligence data and conversion optimization, 145
business processes analysis, 300
buyers, determining (attribution), 217-220

C

cable/television data and data integration, 245
cadence (social analysis), stakeholders and analytics teams, 300
calendar (retail), 28
campaign analysis, 113
capital budgeting models, 25
catalog (ecommerce platforms), 231
cataloging data, 299, 311-312
categories
  categorical data, 35
  category analysis, 186, 188-189, 195-196
  order data models, 183
causal analytics, 301
center of excellence organizational model (analytics teams), 290
centralized organizational model (analytics teams), 289-290
channels (marketing), 182
charts
  area charts, 44
  bar charts, 88
  bullet charts, 46
  flow charts and conversion optimization, 146
  funnel charts, 46, 146
  Harvey Balls, 43-44
  line charts, 44-45, 88
  network charts, 89
  pie charts, 43-44, 88-89
  radar charts, 88
  sparkline charts, 44-45
  streamgraph charts, 45
  tree charts, 89
  tumbler charts, 46-47
  waterfall charts, 89
checkout
  abandonment rates, 125-126
  checkout path analysis, 186, 189-190
choice modeling and conversion optimization, 148
churn rates (customer analytics), 169-170
classical statistical analysis, 33, 36-38
clouds and data integration, 255-256
CLTV (Customer Lifetime Value), 166-167, 170
clusters
  cluster analysis, 186, 191-192
  clustered histograms, 42-43
  clustering (customer analytics), 174, 175-176
cohort analysis, 165-166
collaboration
  collaborative filtering, predicting by (customer analytics), 175
  future of ecommerce analytics, 316
collecting data
  analytical environment, 19, 20
  analytics teams, 298-299
  data governance, 273
  ecommerce and, 3-5
  future of ecommerce analytics, 310, 311-313
communication and analytics teams, 292
comparisons
  data over time (systemic sampling), 56
  data presentations, 72
competition and ecommerce, 3
competitive intelligence and merchandising analysis, 207
compositions (data) and data presentations, 72
conclusive analysis (marketing analytics), 107
conditional probability, 52
confidence intervals (systemic sampling), 55
confirmatory analysis, 22
consideration marketing, 111-112
consultative organizational model (analytics teams), 290
consumer layer (ecommerce platforms), 230
content
  content analysis, 131, 135
  content management (ecommerce platforms), 231, 238
contextualizing data (data governance), 269
continuous data, 35
controlled data experiments, 53-56
conversational speech technology, 316
conversion
  conversion optimization, 133-134, 138, 150
    analytics, 141-143
    Bayesian statistical analysis, 147
    best practices, 151-153
    business intelligence data, 145
    choice modeling, 148
    content, 135
    copywriting, 140
    customer research, 145-146
    data visualization, 145-146
    digital analytics, 145
    email, 134
    eye tracking, 146
    factorial testing, 149-150
    finger tracking, 146
    flow charts, 146
    flows, 135
INDEX

Frequentist methodologies, 147
funnel charts, 146
heatmapping, 146
information architectures, 140
measuring, 140-141
mobile experience, 134
mouse tracking, 146
multivariate testing, 148
network data, 146
neuromarketing, 146
online advertising, 134-135
optimal design, 150
persuasion, 138-139
process of, 141-145
promotions, 135
psychology, 139-140, 146
social media, 135
successful optimization, 151-153
surveys, 145-146
Taguchi method, 147-148
technical data, 146
tools, 135-137
types of data used, 145-146
univariate testing, 148
user experience, 134, 139
value propositions, 137-138
conversion rates, 62, 125, 217-220
ecommerce and, 3-4
marketing, 112
testing tools, 24-25
Cool, Bruce, 172-173
copywriting and conversion optimization, 140
correlating data, 47-49, 126
costs
  acquisition costs (customer analytics), 168-169
cost/revenue tracking, 24-25, 68-69
  estimating (pricing analysis), 207
creative testing/optimization
  (merchandising analysis), 200
cross-sell analysis, 186, 192
CSL (Customer Service Level) KPI
  (merchandising analysis), 202-203
current state assessments (analytics teams), 291-293
customers
  anonymous customer data and data governance, 266
  attribution, 217
affiliate marketing, 218
blogs and, 219
budgets and, 214
buyers, 217-220
conversion rates, 217-220
custom-weighted attribution, 222
defining, 213
direct traffic and, 220
display advertising and, 218
earned media and, 219-220
e-mail, 219
engagement mapping, 220-221
first interaction attributes, 221, 223
importance of, 213-214
“last click” modeling, 215-217
last interaction attributes, 221, 223
last nondirect interaction attributes, 221, 223
latency scoring, 222, 224
linear attribution, 221, 223
logistic regression, 222, 224
marketing mix-modeling versus, 214-215
media mix-marketing versus, 214-215
owned media and, 219
paid media and, 218-219
paid searches, 218
pathing, 222
position-based attribution, 221, 223-224
profits, 217-220
proportional hazard modeling, 222-223
referral traffic and, 219
re-targeting and, 219
revenue, 217-220
SEO, 219
social media, 219
software, 225-227
stacking, 222, 224
starter-player-closer attribution, 222-223, 224
time decay attribution, 221, 223
top-down versus bottom-up approaches, 224-225
types of, 221-224
audience data and data integration, 245
cohort analysis, 165-166
conversion, 3-4, 134, 145-146
CSL KPI (merchandising analysis), 202-203
customer analytics, 155-156
acquisition costs, 168-169
calculating CLTV, 166-167
churn rates, 169-170
clustering customers, 174, 175-176
cohort analysis, 165-166
customer records, 156-157
personalizing customer experiences, 178
predicting customer behavior, 174-175, 176-178
questioning customer data, 158-159
RFM analysis, 171-172
scoring customers, 173-174
segmenting customers, 163-164
SOW analysis, 172-173
Tumbler lifecycle, 159-161
types of analytics, 162-163
types of customers, 157-158, 161-162
VoC analytics, 170-171
Customer facts, 96-97, 99-100
customer ID, 182
customer journey analysis, 116
customer management (ecommerce platforms), 231
Customer Order facts, 97, 100
Customer Origin reports, 116-117
customer service, 196-197, 232
ecommerce platforms, 234
geography and orders, 182
Orders Per Customer KPI, 184-185
percentage of new customers KPI, 67-68
personalizing experiences, 178, 313-314
PII and data governance, 266-267
predicting customer behavior, 174-175
repeat customers KPI, 61
VoC analytics, 170-171
customized research data and data integration, 245
custom-weighted attribution, 222

D

dashboards, 77-78
Agile environments (software solutions), 251
analytical dashboards, 79
data governance, 271-272
data visualization and, 78-79
KPI, 78-79
LIVEN method of dashboarding, 80
necessary data in, 81
need for, 71-72
operational dashboards, 79
purpose of, 72-74
strategic dashboards, 78-79
team roles in, 74
data
administration and analytical environments, 19
analysis. See analytics
anonymous customer data and data governance, 266
area charts, 44
audience data, 245
B2B data, 245
bar charts, 88
behavioral data and data governance, 267
binomial distributions, 52-53
bivariate data, 34
block plots, 39
cataloging, 299, 311-312
categorical data, 35
clouds and data integration, 255-256
collection
analytical environment, 19-20
analytics teams, 298-299
data governance, 273
ecommerce and, 3-5
future of ecommerce analytics, 310, 311-313
comparisons, 56, 72
compositions and presentations, 72
contextualizing data (data governance), 269
continuous data, 35
correlation optimization, 145-146

customer analytics, 157-159
dashboards, 77-81
data analysis tools and ecommerce platforms, 232
data integration, 261-262
data layer (ecommerce platforms), 230
Data Tier (ecommerce), 20
digital analytics data, 244
discrete data, 35
distributions and presentations, 72
EDW, 254-257
encryption, 275-276, 315-316
experiences, 310, 313-314
experimenting with data, 53-56
external/internal data components (ecommerce platforms), 234-235
federations and data integration, 258
financial data, 245, 267-268
flow visualizations, 45-47, 315
four plots, 42
governance, 263-264, 268-269
analytics teams, 292, 299-300
anonymous customer data, 266
behavioral data, 267
collecting data, 273
contextualizing data, 269
dashboards, 271-272
data lineage, 274
data privacy, 264-268, 272-278
data security, 265-268, 272-278
discipline for violations, 275
encryption, 275-276
financial data, 267-268
guidelines, 266-268, 276-278
monitoring data, 270-271, 275
ownership of data, 269-270
permissions, 274-275
PII, 266-267, 275-276
quality of data, 270
stewards, 267-268, 270-271
storing data, 273
teams, 293
technology and, 271-272
tools, 271-272
transactional data, 267
UIA data, 277
Harvey Balls, 43-44
heatmapping, 89, 146
histograms, 42-43
hypothesis testing, 52-53
integration
Agile environments (software solutions), 247-251
analytics applications and, 259-261
audience data, 245
B2B data, 245
cable/television data and data integration, 245
clouds, 255-256
customized research data, 245
data virtualization, 239
defining, 241
digital analytics data, 244
EDW, 254-256
federations, 258
financial data, 245
framework of, 246-247
internal data, 244
lakes, 256-258
lead times, 241-242
ODS, 253-254
social data, 244-245
specialized research data, 245
support for, 241-242
syndicated research data, 245
systems integration and, 242-243
types of data, 243-245
Waterfall environments (software solutions), 248-249, 251-253
internal data, 244
internal/external data components (ecommerce platforms), 234-235
interval data, 35
KPI, 58-61, 64-67
abandon cart rates, 63-64
average examples, 64
average order value, 64
bounce rate, 64-65
conversion rates, 62
cost/revenue tracking, 68-69
creating, 58
gross margin, 60-61
inventory turnover, 65
lifetime value, 61
loyalty/time since last visit (recency), 66
page (screen) views, 59
“per” examples, 68-69
percentage examples, 67-68
percentage of new customers, 67-68
percentage of orders with discounts/promotions, 65
rate examples, 62-64
repeat visitors (users/customers), 61
returns, 59-60
revenue, 60
ROI, 65-66
step completion rates, 63
time between visits (frequency), 66-67
visits (sessions), 59
kurtosis, 36
lag plots, 40
lakes, 256-258, 311
line charts, 44-45, 88
lineage, 274
loaders, analytics applications and data integration, 260-261
mean, 36-37, 53
median, 37
mode, 37
modeling, 91-96
Agile environments (software solutions), 250
analytics teams, 293
Customer facts, 96-97, 99-100
Customer Order facts, 97, 100
dimensions, 94-95, 100-102
facts, 94
Inventory facts, 96-97
measures, 94, 100-102
Order facts, 96, 98-99
Order Item facts, 96, 99
primary keys, 95
probability analysis, 52
Product facts, 96, 98
sample data model, 96-97
supply/demand management models, 296-297
techniques for building models, 56-57
monitoring (data governance), 270-271, 275
multivariate data, 34
network charts, 89
neuromarketing data and conversion optimization, 146
nominal data, 35
ODS and data integration, 253-254
ordinal data, 35
outliers, 36, 37-38, 56
ownership of (data governance), 269-270
permissions and data governance, 274-275
PI charts, 43-44, 88-89
PII and data governance, 266-267, 275-276
pipelines, 315
plotting, 38-39
area charts, 44
bar charts, 88
discussions, 39
greavages, 40
heatmap, 59, 146
histograms, 42-43
lag plots, 40
line charts, 44-45, 88
network charts, 89
pie charts, 43-44, 88-89
probability plots, 41
radar charts, 88
radar plots, 40
run sequence plots, 41-42
scatter plots, 41, 89
six plots, 42
sparkline charts, 44-45
spider plots, 40
star plots, 40
streamgraph charts, 45
tree charts, 89
waterfall charts, 89
preparation tools, 20-22
preparing, 20-22, 312
presentations, 71-72, 75-89
privacy, 264-265, 272-273
data collection, 273
data lineage, 274
discipline for violations, 275
encryption, 275-276
guidelines, 266-268, 274, 276-278
monitoring data, 275
permissions, 274-275
policies/procedures, 273
probability analysis, 51-52
probability plots, 41
processing in analytical environments, 19
profiling, 311-312
qualitative data, 34-35
quality of (data governance), 270
quantitative data, 34-35
querying in analytical environments, 19
radar charts, 88
radar plots, 40
random variables, 52
range, 37, 292
regression analysis, 49-51
relationships and presentations, 72
reports, 71-77
run sequence plots, 41-42
sampling, 54
sampling (systemic), 54-56
scatter plots, 41, 89
security, 265-266, 272-273
data collection, 273
data lineage, 274
discipline for violations, 275
encryption, 275-276
guidelines, 266-268, 274, 276-278
monitoring data, 275
permissions, 274-275
policies/procedures, 273
shapes of, 36-47
six plots, 42
social data, 244-245
sources, 19, 312-313
sparkline charts, 44-45
spider plots, 40
standard deviation, 37
star plots, 40
stewardship of, 74, 251, 267-268, 270-271
storage, 19, 273, 293
streamgraph charts, 45
syndicated research data, 245
transactional data and data governance, 267
tree charts, 89
types of, 34-36
UIA data and data governance, 277
univariate data, 34
virtualization, 19, 259
visualization, 81-82

Agile environments (software solutions), 251
analytics teams, 293
conversion optimization, 145-146
dashboards and, 78-79
data shape and, 36-47
flow visualizations, 45-47
need for, 71-72
plotting data, 38-47
process of, 82-84
purpose of, 73-74, 82
SCREEN method of visualizing data, 84-86
team roles in, 74
tools, 20
types of, 88-89
uses of, 86-87
waterfall charts, 89

Days to Order KPI, 185
decentralized organizational model (analytics teams), 290
delivery systems (analytics teams), 292
demand
estimating (pricing analysis), 205-206
identifying, 11-13
prioritizing, 11-13
dependent variables (systemic sampling), 55
deploy/use/upgrade strategies (analytics teams), 298
derivatives (KPI), 64-67
descriptive analysis, 22-23, 106, 301
development (standard), data shapes, 37
devices (orders), 182
digital analytics, 145, 244
dimensions (data models), 94-95, 100-102
direct traffic and attribution, 220
discipline for data privacy/security violations, 275
discounts
merchandising analysis, 209-210
percentage of orders with discounts/promotions (KPI), 65
discrete data, 35
display advertising and attribution, 218
distributions
binomial distributions, 52-53
data distributions and presentations, 72
Du, R. Y., 172

E
earned media and attribution, 219-220
ecommerce
  Analytics Tier, 20
  B2C ecommerce, 2
  competition, 3
  conversion, 3-4
  data collection, 3-5
  Data Tier, 20
  growth of, 1-2
  improvement, need for, 7
  Integration Tier, 20
  mobile experience, 2-3
  platforms (software), 232, 235-236
    analytics layer, 230
    catalog, 231
    consumer layer, 230
    content management, 231
    content management tools, 238
    customer components, 234
    customer management, 231
    customer service, 232
    data analysis tools, 232
    data components, 234-235
    data layer, 230
    defining, 229
    determining analytical approach, 239-240
    fraud awareness tools, 238-239
    fulfillment, 232
    integration layer, 230
    inventory management, 231-232, 237
    marketing components, 232-233
    marketing tools, 238
    merchandising management, 231-232
    merchandising planning tools, 236-237
    mobile experience, 239
    operations and monitoring layer, 230-231
    ordering functionality, 231
    payment processing, 231
    product management tools, 236
    promotional tools, 237
    search tools, 238
    searches, 231
    security tools, 238-239
    shopping functionality, 231
    sizes of, 229-230
    social media tools, 239
    transactions, 233-234
    user experience, 233
    user interface, 233
  user experience and, 3
  as zero-sum game, 3
economic impact of analytics, 24-25
EDA (Exploratory Data Analysis), 31-32
  Bayesian statistical analysis and, 33
  classical statistical analysis and, 33
  flow visualizations, 45-47
  modeling data, 57
  philosophy of, 32-33
  plotting data, 38-39
    area charts, 44
    block plots, 39
    four plots, 42
    Harvey Balls, 43-44
    histograms, 42-43
    lag plots, 40
    line charts, 44-45
    pie charts, 43-44
    probability plots, 41
    radar plots, 40
    run sequence plots, 41-42
    scatter plots, 41
    six plots, 42
    sparkline charts, 44-45
    spider plots, 40
    star plots, 40
    streamgraph charts, 45
education and analytics teams, 304
EDW (Enterprise Data Warehouses), 254, 256-257
email
  attribution and, 219
  conversion optimization, 134
  email analysis, 114-115
encryption, 275-276, 315-316
engagement mapping and attribution, 220-221
engagement scoring (customer analytics), 173
errors (expected) in systemic sampling, 55

events
  event analysis, 186, 189
  Events/Pages/Steps Per Order KPI, 185

expected errors (systemic sampling), 55

experiences (data), 310, 313-314

explanatory analysis, 22

exploratory analysis, 22, 31-33, 106, 301

Exploratory Data Analysis, 31-33

exposure/awareness marketing, 111

external/internal data components (ecommerce platforms), 234-235

eye tracking and conversion optimization, 146

F

factorial testing and conversion optimization, 149-150

facts (data models), 94
  Customer facts, 96-97, 99-100
  Customer Order facts, 97, 100
  Inventory facts, 96, 97
  Order facts, 96, 98-99
  Order Item facts, 96, 99
  Product facts, 96, 98

federations (data) and data integration, 258

Few, Stephen, 79, 81, 86-87

final price, setting via merchandising analysis, 207-208

finance
  financial analysis, 186-187, 267-268, 303
  financial data and data integration, 245
  financial impact assessments, 24

finger tracking and conversion optimization, 146

first interaction attributes, 221, 223

flows
  conversion optimization, 135
  flow charts and conversion optimization, 146
  future of analytics, 315
  visualizations, 45-47

focus groups and marketing/advertising analytics, 107

four plots, 42

fraud awareness tools (ecommerce platform), 238-239

frequency (time between visits)
  behavioral analytics, 124
  KPI, 66-67
  RFM analysis, 171-172

Frequentist methodologies (conversion optimization), 147

Friedman, Vitaly, 82

fulfillment (ecommerce platforms), 232

functional organizational model (analytics teams), 290

funding analytics teams, 291

funnel analysis, 186, 190-191

funnel charts, 46, 146

future capabilities (analytics teams), 291-292

future of ecommerce analytics, 307-310, 314-315, 318
  AI, 317
  collaboration, 316
  conversational speech technology, 316
  data collection, 310, 311-313
  data experiences, 310, 313-314
  data flows, 315
  data pipelines, 315
  encryption, 315-316
  machine learning, 317
  natural language searches, 316
  predictive analysis, 317
  self-service automation, 316

future state architectures, 292-293

G

game scoring (customer analytics), 174

geography and orders, 182

geovisualizations, 89

goal analysis and orders, 186, 189

governance (data), 263-264, 268-269

analytics teams, 292, 299-300
  anonymous customer data, 266
  behavioral data, 267
  contextualizing data, 269
dashboards, 271-272
data privacy, 264-265, 272-273
data collection, 273
data lineage, 274
discipline for violations, 275
encryption, 275-276
guidelines, 266-268, 274,
276-278
monitoring data, 275
permissions, 274-275
policies/procedures, 273
data security, 265-266, 272-273
data collection, 273
data lineage, 274
discipline for violations, 275
encryption, 275-276
guidelines, 266-268, 274,
276-278
monitoring data, 275
permissions, 274-275
policies/procedures, 273
financial data, 267-268
guidelines, 266-268
monitoring data, 270-271
ownership of data, 269-270
PII, 266-267, 275-276
quality of data, 270
stewards, 267-268, 270-271
storing data, 273
technology and, 271-272, 293
tools, 271-272
transactional data, 267
UIA data, 277
gross margin (KPI metrics), 60-61
growth of ecommerce, 1-2

H
Harvey Balls, 43-44
heatmapping, 89, 146
heuristic analytics and conversion
optimization, 142
hiring (analytics teams)
leadership, 285-288
staff, 291
histograms, 42-43
hub and spoke organizational model
(analytics teams), 290-291
hypothesis testing, 52-53, 143

I
improvement, ecommerce's need
for, 7
inactivity scoring (customer analytics),
174
independent variables (systemic
sampling), 55
inferences (systemic sampling), 56
inferential analytics, 301
information architectures and
conversion optimization, 140
in-page (on-screen) behavior analysis,
132
integration
data integration
Agile environments (software
solutions), 247-251
analytics applications and,
259-261
audience data, 245
B2B data, 245
cable/television data and data
integration, 245
clouds, 255-256
customized research data, 245
data virtualization, 259
defining, 241
digital analytics data, 244
EDW, 254-256
federations, 258
financial data, 245
framework of, 246-247
internal data, 244
lakes, 256-258
lead times, 241-242
ODS, 253-254
process of, 261-262
social data, 244-245
specialized research data, 245
support for, 241-242
syndicated research data, 245
systems integration and,
242-243
types of data, 243-245
Waterfall environments
(software solutions), 248-249,
251-253
integration layer (ecommerce
platforms), 230
Integration Tier (ecommerce), 20
systems integration, 242-243
interaction maps, 295-296
intermediaries (brokers), 74
internal data and data integration, 244
internal/external data components
(ecommerce platforms), 234-235
interval data, 35
interviews and marketing/advertising
analytics, 107-108
inventories
analysis, 200-203
Inventory facts, 96, 97
inventory management (ecommerce
platforms), 231-232
management tools, 237
turnover (KPI), 65
investigators, 74
investment (ROI), 65-66, 210
IRR (Internal Rate of Return), 25
item analysis and orders, 186, 187
orders and, 184-186
page (screen) views, 59
"per" examples, 68-69
percentage examples, 67-68
percentage of orders with discounts/
promotions, 65
rate examples, 62-64
repeat visitors (users/customers), 61
returns, 59-60
revenue, 60, 68-69
ROI, 65-66
shopping behavior analysis, 130-131
step completion rates, 63
visits (sessions), 59
loyalty/time since last visit
(recency), 66
time between visits (frequency),
66-67
kurtosis (data shapes), 36

J-K

Journal of Interactive Marketing, 167
Kamakura, W. A., 172
Keiningham, Timothy L., 172-173
KPI (Key Performance Indicators), 58
abandon cart rates, 63-64
average examples, 64
average order value, 64
behavioral analytics, 124-126
bounce rate, 64-65
conversion rates, 62
cost/revenue tracking, 68-69
creating, 58
CSL (merchandising analysis),
202-203
customers
cost/revenue tracking, 68-69
percentage of new customers,
67-68
repeat customers, 61
dashboards, 78, 79
derivative examples, 64-67
gross margin, 60-61
inventory turnover, 65
lifetime value, 61
metric examples, 59-61
Number of Stock Outs
(merchandising analysis), 202
L
lag plots, 40
Laja, Peep, 142
lakes (data), 256-258, 311
Landing Page Optimization, 149
languages
conversational speech technology,
316
natural language searches, 316
“last click” modeling and attribution,
215-217
last interaction attributes, 221, 223
last nondirect interaction attributes,
221, 223
latency scoring and attribution, 222,
224
Leek, Dr. Jeffrey, 301
leptokurtic kurtosis (data shapes), 36
Lewis-Beck, Michael S., 50-51
lifetime value, 61, 166
line charts, 44-45, 88
lineage of data, 274
linear attribution, 221, 223
LIVEN method of dashboarding, 80
loaders (data) and data integration,
260-261
logistic regression, 51, 222, 224
loyalty
  retention marketing, 112
  scoring (customer analytics), 173
time since last visit (recency) KPI, 66

M
machine learning, 317
MANOVA tests, 49-50
maps
  analytics road maps, 294
  interaction maps, 295-296
  process maps, 295
  workflow maps, 296
margins (KPI metrics), 60-61, 184
markdowns, determining via
merchandising analysis, 209-210
marketing
  ecommerce platforms, 232-233, 238
  marketing analytics, 103-105, 112-113
    acquisition marketing, 112
    activation marketing, 111
    advertising goals and, 105-108
    affiliate analysis, 114
    attribution, 115
    audience analysis, 115-116
    campaign analysis, 113
    channels, 182
    conclusive analysis, 107
    consideration marketing, 111-112
    conversion marketing, 112
    conversion optimization, 143
    customer journey analysis, 116
    Customer Origin reports, 116-117
    descriptive analysis, 106
    ecommerce and user experience, 3
    email analysis, 114-115
    exploratory analysis, 106
    exposure/awareness marketing, 111
    focus groups, 107
    goals of, 105-108
    interviews, 108
    lifecycle of, 108-111
    marketing mix-modeling, 115, 214-215
media mix-marketing, 214-215
  necessary data in, 116-117
  online display ad unit analysis, 116
  panels, 108
  predictive analysis, 106-107
  prescriptive analysis, 107
  projective techniques (stimuli), 108
  qualitative analytics, 107-108
  retention/loyalty marketing, 112
  SEM, 113
  SEO, 113-114
  social media analysis, 114
  surveys, 108
  types of, 111-112
mean
data shapes, 36-37
  sample mean, learning from, 53
  measures (data models), 94, 100-102
Measuring the Digital World, 305
mechanistic (algorithmic) analytics, 302
MECLABS, 155-156
media
  attribution and
    earned media, 219-220
    media mix-marketing, 214-215
    owned media, 219
    paid media, 218-219
    social media, 219, 239
median (data shapes), 37
Median KPI, 184
Mela, C. F., 172
merchandising
  merchandising analysis, 198-200
  creative testing/optimization, 200
  determining types of data for
    analysis, 210-211
  discounts, 209-210
  inventory analysis, 200-203
  markdowns, 209-210
  offer analysis, 203-204
  pricing analysis, 204-208
  promotions, 209-210
  ROI, 210
  sales impact of merchandising, 208-209
merchandising management (ecommerce platforms), 231-232
planning tools, 236-237
mesokurtic kurtosis (data shapes), 36
metadata collection, 20
metrics
behavioral analytics, 124-126
KPI metric examples, 59-61
orders and, 184-186
shopping behavior analysis, 130-131
mobile experience
conversion optimization, 134
ecommerce and, 2-3
ecommerce platforms, 239
mode (data shapes), 37
modeling data, 91-96
Agile environments (software solutions), 250
analytics teams, 293
choice modeling and conversion optimization, 148
dimensions, 94-95, 100-102
facts, 94
Customer facts, 96-97, 99-100
Customer Order facts, 97, 100
Inventory facts, 96, 97
Order facts, 96, 98-99
Order Item facts, 96, 99
Product facts, 96, 98
measures, 94, 100-102
primary keys, 95
probability analysis, 52
sample data model, 96-97
supply/demand management models, 296-297
techniques for building models, 56-57
monetary analysis, 171-172
monitoring data (data governance), 270-271, 275
mouse tracking and conversion optimization, 146
multiple linear regression, 50-51
multivariate data, 34, 148

neuromarketing and conversion optimization, 146
new customers, percentage of (KPI), 66-67
newsletters (analytics teams), 304
next-best product analysis, 186-187, 193
nominal data, 35
nonlinear dependent correlation, 48-49
NPV (Net Present Value), 25
NRF (National Retail Federation) and 4-5-4 calendar, 28
Number of Stock Outs KPI (merchandising analysis), 202

O

ODS (Operational Data Stores) and data integration, 253-254
offer analysis (products), 203-204
office hours (analytics teams), 302-303
online advertising and conversion optimization, 134
online display ad unit analysis, 116
on-screen (in-page) behavior analysis, 132
operational dashboards, 79
operational data store, 253
operations and monitoring layer (ecommerce platforms), 230-231
optimal design and conversion optimization, 150
optimal pricing, determining via pricing analysis, 204-208
optimization/creative testing, 200
Order facts, 96, 98-99
Order Item facts, 96, 99
ordered behavioral analytics, 122
orders
ACID, 179
affiliates, 182
analysis
approaches to, 186-193
brand analysis, 186, 188-189
category analysis, 186, 188-189
checkout path analysis, 186, 189-190
cross analysis, 186, 191-192
cross-sell analysis, 186, 192

natural language searches, 316
network charts, 89
INDEX

determining data for analysis, 183-184
event analysis, 186, 189
financial analysis, 186, 187
funnel analysis, 186, 190-191
goal analysis, 186, 189
importance of, 180-181
item analysis, 186, 187
next-best product analysis, 186-187, 193
promotional analysis, 186, 188
scope of, 180-181
upsell analysis, 186, 192
brands, 183
categories, 183
categories (marketing), 182
customer ID, 182
data model, 180-183
defining, 179, 181-183
devices, 182
geography (customer), 182
KPI, 184

Average Revenue Per Order KPI, 184
Days to Order KPI, 185
Events/Pages/Steps Per Order KPI, 185
Margin KPI, 184
Orders and Products Bought by Marketing Channel KPI, 185
Orders Per Customer KPI, 185
Revenue By Category KPI, 185
Top Products KPI, 186
Top Promotions KPI, 185
Top Shipping Methods KPI, 185
Total Revenue Versus Total Profit KPI, 184
Visits to Order KPI, 185
merchandising, 198-200

creative testing/optimization, 200
determining types of data for analysis, 210-211
discounts, 209-210
inventory analysis, 200-203
markdowns, 209-210
offer analysis, 203-204
pricing analysis, 204-208
promotions, 209-210

ROI, 210
sales impact of merchandising, 208-209
metrics, 184-186
order ID, 181
ordering functionality (ecommerce platforms), 231
orders with discounts/promotions (KPI), 65
price, 183
product ID, 182
product names, 182
promotions, 182
quantity, 183
revenue, 181-182
seasons, 183
shipping, 182
size, 183
SKU, 183
style, 183
taxes, 182
technical information, 182
time, 182
transaction ID, 182

ordinal data, 35
organic searches, 113-114, 219
organizations. See teams (analytics)
outliers
data shapes, 36, 37-38
modeling data, 56
owned media and attribution, 219
ownership of data (data governance), 269-270

P
page (screen) views (KPI metrics), 59
paid media and attribution, 218-219
paid searches and attribution, 218
panels and marketing/advertising analytics, 108
pathing and attribution, 222
path-to-purchase analysis, 186, 189-190
payment processing (ecommerce platforms), 231
Pearson’s correlation, 48
“per” examples (KPI), 68-69
percentages (KPI examples), 66-68
periodicity (social analysis),
stakeholders and analytics teams, 300
permissions and data governance, 274-275
personalizing customer experiences, 178, 313-314
persuasion (conversion optimization), 138-139
pie charts, 43-44, 88-89
PII (Personally Identifiable Information) and data governance, 266-267, 275-276
pipelines (data), 315
planning committees, 305
platforms (ecommerce), 232, 235-236
  analytical approach, determining, 239-240
  analytics layer, 230
  catalog, 231
  consumer layer, 230
  content management, 231
  content management tools, 238
  customer components, 234
  customer management, 231
  customer service, 232
  data analysis tools, 232
  data components, 234-235
  data layer, 230
  defining, 229
  fraud awareness tools, 238-239
  fulfillment, 232
  integration layer, 230
  inventory management, 231-232
  inventory management tools, 237
  marketing components, 232-233
  marketing tools, 238
  merchandising management, 231-232
  merchandising planning tools, 236-237
  mobile experience, 239
  operations and monitoring layer, 231
  ordering functionality, 231
  payment processing, 231
  product management tools, 236
  promotional tools, 237
  search tools, 238
  searches, 231
  security tools, 238-239
shopping functionality, 231
sizes of, 229-230
social media tools, 239
transactions, 233-234
user experience, 233
user interface, 233
platykurtic kurtosis (data shapes), 36
plotting data, 38-39
  area charts, 44
  bar charts, 88
  block plots, 39
  four plots, 42
  geovisualizations, 89
  Harvey Balls, 43-44
  heatmapping, 89, 146
  histograms, 42-43
  lag plots, 40
  line charts, 44-45, 88
  network charts, 89
  pie charts, 43-44, 88-89
  probability plots, 41
  radar charts, 88
  radar plots, 40
  run sequence plots, 41-42
  scatter plots, 41, 89
  six plots, 42
  sparkle charts, 44-45
  spider plots, 40
  star plots, 40
  streamgraph charts, 45
  tree charts, 89
  waterfall charts, 89
policies/procedures and data governance, 273
population (controlled data experiments), 54
position-based attribution, 221, 223-224
postmortems (projects), 305
predicting customer behavior, 174-175
predictive analysis, 22, 106-107, 174-178, 301, 317
preparing data, 20-22, 312
prescriptive analysis, 22-23, 107, 301
presentations
  analytical environment, 20
  dashboards, 77-81
  data visualization, 78-79, 81-89
  need for, 71-72
  purpose of, 72
reports, 75-77
  team roles in, 74
  
prices
  discounts, 209-210
  markdowns, 209-210
  merchandising analysis, 209-210
  order data models, 183
  pricing analysis, 204-208
  ROI, 210

primary keys (data models), 95
privacy (data), 263-265, 272-273
  data collection, 273
  data lineage, 274
  discipline for violations, 275
  encryption, 275-276
  guidelines, 266-268, 274, 276-278
  monitoring data, 275
  permissions, 274-275
  policies/procedures, 273

probability analysis, 51-52
probability plots, 41

procedures/policies and data governance, 273

processes
  business processes analysis, 300
  change management processes, 304-306
  data processes in analytical environments, 19
  maps, 295

products
  analysis, 193
    approaches to analysis, 186-193
    brand analysis, 186, 188-189
    category analysis, 186, 188-189
    checkout path analysis, 186, 189-190
    cluster analysis, 186, 191-192
    cross-sell analysis, 186, 192
    customer service analysis, 196-197
    event analysis, 186, 189
    financial analysis, 186, 187
    funnel analysis, 186, 190-191
    goal analysis, 186, 189
    importance of analysis, 180-181
    item analysis, 186, 187
    next-best product analysis, 186-187, 193
    product brand analysis, 194-195
  product category analysis, 195-196
  product returns analysis, 197
  promotional analysis, 186, 188
  scope of analytics, 180-181
  social media product analysis, 197-198
    types of, 193-194
  up-sell analysis, 186, 192
  defining, 179
  merchandising, 198-200
    creative testing/optimization, 200
    determining types of data for analysis, 210-211
    discounts, 209-210
    inventory analysis, 200-203
    markdowns, 209-210
    offer analysis, 203-204
    pricing analysis, 204-208
    promotions, 209-210
    ROI, 210
    sales impact of merchandising, 208-209
  naming, 182
  Orders and Products Bought by Marketing Channel KPI, 185
  Product facts, 96

Product facts, 98
  product ID, 182
  product management tools, 236
  Top Products KPI, 185-186

profile completeness scoring
  (customer analytics), 173-174

profiling data, 311-312

profits
  attribution and, 217-220
  Total Revenue Versus Total Profit KPI, 184

projective techniques (stimuli) and marketing/advertising analytics, 108

promotions, 182
  conversion optimization, 135
  merchandising analysis, 209-210
  percentage of orders with discounts/promotions (KPI), 65
  promotional analysis, 186, 188
  promotional tools, 237
  Top Promotions KPI, 185
propensities, predicting behavior by (customer analytics), 174-175, 176-178
proportional hazard modeling and attribution, 222-223
psychology (conversion optimization), 139-140, 146

Q
QA (quality assurance), Waterfall environments (software solutions), 253
qualitative analytics, 34-35, 107-108, 143
quality of data (data governance), 270
quantitative data, 34-35
quantity (order data models), 183
quarterly reviews (analytics teams), 305
querying data in analytical environments, 19
questions, asking in analytics, 6-7

R
radar charts, 88
radar plots, 40
random variables, 52
range
  analytics teams, 292
data shapes, 37
rank correlation, 48
RASTA method of reporting, 77
rates, KPI rate examples, 62-64
recency (time since last visit/loyalty), 66, 124, 171-172
records (customer) and customer analytics, 156-157
referral traffic and attribution, 219
regression analysis, 49-51
regular histograms, 42
relationships (data) and presentations, 72
repeat visitors (users/customers), KPI metrics, 61
reports, 75-77
  ad hoc reporting, 76
  Agile environments (software solutions), 251
  Customer Origin reports, 116-117
  need for, 71-72
  purpose of, 72-74
  RASTA method of reporting, 77
  sustaining reporting, 76
  team roles in, 74
research data and data integration, 245
results (analytics teams), determining, 292
retail calendar, 28
retargeting and attribution, 219
retention/loyalty marketing, 112
returns
  KPI metrics, 59-60
  product returns analysis, 197
revenue, 181-182
  attribution and, 217-220
  Average Revenue Per Order KPI, 184
cost tracking and, 24-25, 68-69
  Revenue By Category KPI, 185
  revenue by N (KPI metrics), 60
  Total Revenue Versus Total Profit KPI, 184
RFM (Recency, Frequency, and Monetary) analysis, 171-172
road maps (analytics), 294
ROI (return on investment), 65-66, 210
Rosling, Hans, 86
run sequence plots, 41-42

S
sales
  cross-sell analysis, 186, 192
  up-sell analysis, 186, 192
sampling data (systemic), 53-55
  comparisons of data over time, 56
  confidence intervals, 55
  dependent variables, 55
  expected errors, 55
  independent variables, 55
  inferences, 56
  significance testing, 55
scatter plots, 41, 89
scope (analytics teams), 291
scoring customers (customer analytics), 173-174
screen (page) views (KPI metrics), 59
SCREEN method of visualizing data, 84-86
scrums, 249-250, 253, 305
searches
  ecommerce platforms, 231, 238
  paid searches and attribution, 218
  search engines, 113-114, 219
seasons (order data models), 183
security
  data security, 263-266, 272-273
    data collection, 273
    data lineage, 274
data collection, 273
data lineage, 274
discipline for violations, 275
encryption, 275-276
guidelines, 266-268, 274, 276-278
monitoring data, 275
permissions, 274-275
policies/procedures, 273
customer management tools, 236
encryption, 275-276, 315-316
ecommerce platforms, 238-239
encryption, 275-276, 315-316
Seeking phase (Tumbler lifecycle), 159-160
segmentation (customer analytics), 163-164
self-service automation, 293, 304, 316
SEM (search engine marketing), 113
SEO (search engine optimization), 113-114, 219
servers, analytics applications and data integration, 260
sessions. See visits
SFE (secure function evaluation), 315-316
SFTP (SSH file transfer protocol), 261
batch processing and data integration, 261
Sharing phase (Tumbler lifecycle), 160
shipping, 182, 185
shopping
  shopping behavior analysis, 130-131
  shopping functionality (ecommerce platforms), 231
Shopping phase (Tumbler lifecycle), 160
significance testing (systemic sampling), 55
single linear regression, 50-51
six plots, 42
SKU (stock keeping units), 182-183
SLA (service-level agreements), 298
social analysis, 23-24, 300
social data and data integration, 244-245
social media
  attribution and, 219
  ecommerce platforms, 239
  social media analysis, 114, 135, 197-198
software
  Agile environments, 247-251
  attribution software, assessing, 225-227
  ecommerce platforms, 232, 235-236
data analysis tools, 232
data components, 235-236
data layer, 230
defining, 229
determining analytical approach, 239-240
fraud awareness tools, 238-239
fulfillment, 232
integration layer, 230
inventory management, 231-232, 237
marketing components, 232-233
marketing tools, 235
merchandising management, 231-232
merchandising planning tools, 236-237
mobile experience, 239
operations and monitoring layer, 230-231
ordering functionality, 231
payment processing, 231
product management tools, 236
promotional tools, 237
search tools, 238
searches, 231
security tools, 239-239
shopping functionality, 231
sizes of, 229-230
social media tools, 239
transactions, 233-234
user experience, 233
user interface, 233
Waterfall environments, 248-249, 251-253
SOW (share of wallet) analysis, 167, 172-173
sparkline charts, 44-45
specialized research data and data integration, 245
spider plots, 40
sponsorship (analytics teams), 291
SSL (secure sockets layer), 315-316
stacked histograms, 43
stacking and attribution, 222, 224
staff (analytics teams)
hiring, 291
training, 294-295
stakeholders
analytics teams, 297-298, 304-306
financial impact assessments, 24
standard deviation (data shapes), 37
star plots, 40
starter-player-closer attribution, 221-222, 224
step completion rates (KPI), 63
stewards, 74, 251, 267-268, 270-271
stimuli (projective techniques) and marketing/advertising analytics, 108
storing data, 19, 273, 293
storytelling and analytics, 29-31
strategic dashboards, 78-79
streamgraph charts, 45
supply/demand management models, 296-297
surveys, 108, 145-146
sustaining reporting, 76
syndicated research data and data integration, 245
systemic sampling, 54-55
comparisons of data over time, 56
confidence intervals, 55
dependent variables, 55
expected errors, 55
independent variables, 55
inferences, 56
significance testing, 55
systems integration and data integration, 242-243

Taguchi, Genichi, 147-148
taxes, 182
teams (analytics), 279-280, 300-301
algorithmic (mechanistic) analytics, 302
analytics platforms, 293
architectures, 292
artifacts, 296
attending non-team meetings, 302
build templates, 296
building (universal process), 280-283
business outcomes from, 303
business processes analysis, 300
business requirements, gathering, 288-289
cataloging data, 299
causal analytics, 301
change management processes, 304-306
checkpoints, 302
communication, 292
current state assessments, 291-293
data governance, 292-293, 299-300
data models, 293
data ranges, 292
data storage, 293
data visualization, 293
delivery systems, 292
descriptive analysis, 301
education, 304
exploratory analysis, 301
financial analysis, 303
funding, 291
future capabilities, 291-292
future state architectures, 292-293
hiring
leadership, 285-288
staff, 291
inferential analytics, 301
leadership, 285-288
maps, 294, 295-296
need for, 283-285
new data collection/acquisition, 298-299
new work, leading/assisting with, 302-303
newsletters, 304
office hours, 302-303
organizational models, 289-291
planning committees, 305
postmortems (projects), 305
predictive analysis, 301
prescriptive analysis, 301
quarterly reviews, 305
results, determining, 292
scope, 291
scrums, 305
self-service automation, 293, 304
SLA, 298
sponsorship, 291
staff, 291, 294-295
stakeholders, 297-298, 300, 304-306
supply/demand management models, 296-297
team analysis, 303-304
technology architecture assessments, 292-293
ticketing systems, 302
use/deploy/upgrade strategies, 298
value, determining, 291-292
vision of, shaping, 289
technology architecture assessments, 292-293
television/cable data and data integration, 245
templates (build), 296
tendencies, predicting behavior by (customer analytics), 174-175
tests
ANOVA tests, 39, 49-50
conversion optimization, 143-145
creative testing/optimization (merchandising analysis), 200
factorial testing and conversion optimization, 149-150
hypothesis testing, 52-53
MANOVA tests, 49-50
multivariate testing and conversion optimization, 148
significance testing (systemic sampling), 55
UAT, 251, 253
univariate testing and conversion optimization, 148
ticketing systems (analytics teams), 302
time
comparisons of data over time (systemic sampling), 56
orders, 182
time between visits (frequency), KPI, 66-67
time decay attribution, 221, 223
time since last visit/loyalty (recency), KPI, 66
time stamps, 182
Top Products KPI, 185-186
Top Promotions KPI, 185
Top Shipping Methods KPI, 185
total revenue (KPI metrics), 60
Total Revenue Versus Total Profit KPI, 184
tracking revenue/cost, 24-25, 68-69
transactions
data governance, 267
ecommerce platforms, 233-234
transaction ID, 182
tree charts, 89
Tuft, Edmund, 86
Tukey, John, 31-33, 57
tumbler charts, 46-47
Tumbler lifecycle (customer analytics), 159-161
U
UAT (User Acceptance Testing), 251, 253
UIA (Useful, Insightful, Actionable)
data and data governance, 277
univariate data, 34, 148
unordered behavioral analytics, 122
upgrade/use/deploy strategies (analytics teams), 298
up-sell analysis, 186, 192
use/deploy/upgrade strategies (analytics teams), 298
users
repeat users (KPI metrics), 61
user experience
conversion optimization, 134, 139
ecommerce and, 3, 7
ecommerce platforms, 233
user interface (ecommerce platforms), 233
V

validators, 74

value
  - analytics teams, determining, 291-292
  - average order value KPI, 64
  - lifetime value KPI, 61

value chain (analytics), 9
  - analysis techniques, 23-25, 31-33
  - analytical environment, 16-20
  - analytical plans, developing, 14-16
  - data preparation tools, 20-22
  - demand, 11-13
  - economic impact of analytics, 24-25
  - overview of, 9-11
  - phases of, 9

value propositions and conversion optimization, 137-138

variables (random), 52

variance
  - ANOVA tests, 39, 49-50
  - MANOVA tests, 49-50

violations of data privacy/security, 275

virtualizing data, 19, 259

visitors (repeat), KPI metrics, 61

visits (sessions), 59
  - behavioral analytics, 66-67, 124, 125
  - Visits to Order KPI, 185

visualizing data, 81-82
  - Agile environments (software solutions), 251
  - analytics teams, 293
  - conversion optimization, 145-146
  - dashboards and, 78-79
  - data shape, 36-47
  - flow visualizations, 45-47
  - need for, 71-72
  - plotting data, 38-39
    - area charts, 44
    - bar charts, 88
    - block plots, 39
    - four plots, 42
    - geovisualizations, 89
    - Harvey Balls, 43-44
    - heatmapping, 89, 146
    - histograms, 42-43
    - lag plots, 40
    - line charts, 44-45, 88
    - network charts, 89
    - pie charts, 43-44, 88-89
    - probability plots, 41
    - radar charts, 88
    - radar plots, 40
    - run sequence plots, 41-42
    - scatter plots, 41, 89
    - six plots, 42
    - sparkline charts, 44-45
    - spider plots, 40
    - star plots, 40
    - streamgraph charts, 45
    - tree charts, 89
    - waterfall charts, 89

  - process of, 82-84
    - purpose of, 73-74, 82
    - SCREEN method of visualizing data, 84-86
      - team roles in, 74
    - tools, analytical environment, 20
      - types of, 88-89
      - uses of, 86-87

  - VoC (Voice-of-the-Customer) analytics, 170-171

W-X-Y-Z

wallet allocation rule, 172-173

waterfall charts, 89

Waterfall environments (software solutions), 248-249, 251-253

workflow maps, 296

zero-sum game, ecommerce as, 3