## Index

### A

- **action items.** See call to action checklist
- **advanced analytics.** See also data science; predictive analytics
  - customer behavior analytics example, 162
  - database analytics and, 188–189
  - envisioning and, 126, 133
  - insights and, 59, 56
  - intelligent products and, 10
  - network optimization example and, 167
  - new capabilities, 180–181
  - powering key business initiatives and, 156
  - predictive maintenance example and, 163
  - in retail industry, 22
  - solution engineering and, 151
  - transformation of business processes and, 5–6
  - value creation process and, 201
- **advertiser solutions,** 32–33
- **aggregation,** 136–138
- **Amazon**
  - Apache Hive and, 178
  - big data technologies and, 175
  - analysis paralysis, defined, 49
  - analytic algorithms, new, 180–181
  - analytic models, vision workshops and, 129
  - analytic tools, new, 179–180
- **analytics**
  - advanced. See advanced analytics; data science
  - analytic process, evolution of, 38
  - analytics chasm, 54
  - combining big data and. See decision theory
  - customer behavioral analytics example, 162–163
  - embedded, 9
  - formal analytics governance processes, 10
  - how not to leverage, 106–107
  - to improve customer engagement, 108–110
  - in-database analytics, 188–190
  - maturing of business and, 11, 16–17
  - anchoring bias, 59–60
  - Anderson, Chris, 53
  - Anderson, S., 113, 154
- **annual reports**
  - brokerage firm example, 171–172
  - financial services firm example, 168–169
  - importance of, 167–168
  - retail example, 169–171
- **Apache**
  - Apache Software Foundation, 177
  - Hadoop, 176–177
  - Hadoop Distributed File System (HDFS), 176
  - HBase, 178
  - Hive, 178
  - Mahout, 180
  - MapReduce, 176
- **architectural ramifications,** 173–191
  - Apache Hadoop, 176–177
  - Apache HBase, 178
  - Apache Hive, 178
  - cloud computing, 190–191
  - data capabilities of new architectures, 176
  - Hadoop MapReduce, 177
  - in-database analytics, 188–190
  - need for new data architecture, 173–175
  - new analytic algorithms, 180–181
  - new analytic tools, 179–180
  - Pig, 178–179
  - architecture
    - implementing, 199
    - options, 205–206
- **automotive manufacturer example,** 87–89

### B

- **B2B.** See Business-to-business (B2B) companies
- **B2C.** See Business-to-consumer (B2C)
- **Beane, Billy,** 27, 28
- **BI.** See Business Intelligence (BI)
- **BI analysts.** See Business Intelligence (BI) analysts
- **big data**
  - changes made by, 193
  - envisioning process. See vision workshops vs. ERP, 125
impact on targeted business initiatives, understanding, 156–157
summary of observations and strategies. See Big Data Storymap
using effectively, 125
Big Data Business Model Maturity Index, 5–19
business aspirations and, 197–198
Business Insights phase, 7–8, 17
Business Metamorphosis phase, 12–16, 17, 197
Business Monitoring phase, 7, 17, 197
Business Optimization phase, 9–10, 17
Data Monetization phase, 10–11, 17
observations of, 16–17
overview, 5–6
Big Data Envisioning Worksheets
business drivers and, 85
Michael Porter, merchandising examples, 95–97, 101–104
Big Data Storymap, 193–202
Big Data Business Model Maturity Index, 197–198
business opportunities, data growth and, 194–195
goal of, 194
insufficiency of traditional technologies, 195–196
organizational insights, 199–200
value creation process and, 200–201
value journey, 198–199
big data strategy, creating, 65–78
basics of, 65
Big Data Strategy Document, 66–70
customer intimacy example, 67–69
San Francisco Giants example, 73–77
Starbucks example, 70–73
Bonds, Barry, 74
brainstorming. See also ideation workshop
impact of big data and analytics on business initiatives, 156–157
leveraging big data, 134–136
use cases (solution engineering), 157–158
brokerage firm example, 171–172
business impact, 25–35
changes in business, 25–26
customers, products, and operations, understanding of, 26–27
data monetization and, 30–34
management with metrics, 27–30
social media, mobile devices, POS scanners, 1–2
business initiatives
Big Data Strategy Document and, 65, 66, 68, 69
customer satisfaction example, 87–88
drivers and, 81, 85, 86–87
examples of, 127
identifying, 203
identifying for solution engineering, 155
impact of big data and analytics on, 156–157
to improve customer cross-sell effectiveness, 168–169
micro-segmentation example, 89, 90
researching, 127–129
San Francisco Giants and, 73–77
Starbucks and, 71–72
Business Insights phase, 7–8, 9–10, 197
Business Intelligence (BI)
birth of software industry and, 20
vs. data science, 37
organizational impact and, 37–39
solution engineering and, 161
traditional, business growth and, 196
Business Intelligence (BI) analysts
communication of results and, 45
data analytics lifecycle and, 40, 41
data scientists and, 37–39
discovery and, 43
operationalization and, 46
responsibilities, 41
business models, 12–15
Business Model Maturity Index. See Big Data Business Model Maturity Index
changes in, 1, 12–15
Business Monitoring phase, 7
business opportunity
Business Model Maturity Index. See Big Data Business Model Maturity Index
current data-driven revolution, 1–2
data growth and, 194–195
transformation imperative, 2–5
Business Optimization phase, 9–10, 197
business performance management, 7–8
business stakeholders
big data and, 125
collaboration with IT stakeholders, 198–199, 204
business users. See users
business value, Prioritization Matrix and, 141–142, 143, 144, 145
Business-to-business (B2B) companies
customer satisfaction example and, 87–88
merchant effectiveness example, 119–122
new customer experience and, 116
powering small and medium-sized merchant effectiveness, 119–122
predictive maintenance example, 163–164
Business-to-consumer (B2C)
customer lifecycle and, 113
powering retail customer experience example, 116–119
Buyer Power factors, 91–92, 100, 101, 103

C
call to action checklist, 203–207
business and IT stakeholder collaboration, 204
envisioning process, formalizing, 204–205
existing internal business processes, building from, 206
find new monetization opportunities, 206–207
identify key business initiatives, 203
leverage mockups, 205
understand organizational ramifications, 207
understand technological and architectural options, 205–206
Campaign Managers, as target users, 31
career opportunities, impact of big data and, 37
case studies. See examples
Category Management, 21, 24
churn
aggregated customer usage variables vs., 131
churn rate, defined, 112
customer financial and billing variables vs., 130
Client Promoter Score (CPS) program, solution engineering and, 171–172
cloud computing, 190–191
collaboration between businesses and IT stakeholders, 198–199, 204
communication of results, data scientists and BI analysts and, 45
Competitive Rivalry factors, 91, 100, 101, 103
competitive sentiment analysis, 131
“Competitive Strategy: Techniques for Analyzing Industries and Competitors”, 92, 94
Confusion Matrix, 180
Consumer Package Goods (CPG) POS scanner data and, 19, 20–23
pre-1988, 19–23
CPI and customer sales data, integrating, 132
CPS (Client Promoter Score) program, solution engineering and, 171–172
Crest toothpaste, early marketing data and, 19–20
critical success factors (CSF), 66, 68, 72
CRM (customer relationship management), vs. big data, 125
cross-sell effectiveness, 168–169
customer insights, uncovering and leveraging, 110–115
amount of, 110–112
customer lifecycle process, 112–113
profitability and, 113–115
customers
customer behavioral analytics example, 162–163
customer engagement, improving, 108–110
customer experience, powering (overview), 116
customer intimacy example, 67–69
customer lifecycle management process, 112–113
Customer Loyalty Programs, 22, 24
customer micro-segmentation example, 89–90
Customer Profitability Curve, 114
Customer Profitability Distribution, 154
customer relationship management (CRM), vs. big data, 125
Customer Retention Business Driver, 115
customer satisfaction example, 87–89
dashboards, customer experience and, 11
data monetization and, 30
most important, 80
retail customer experience example, 116–119
Starbucks example, 69–70
support of, Business Insights phase and, 8
understanding of, 25, 26–27
value creation process and, 201

d
dark data, 16, 206
dashboards
business insights and, 8
customer experience and, 11
vs. insights-lead analytics user interfaces, 56
intelligent dashboards, business insights and, 7
intelligent merchant dashboards, 119–122
data
access requirements, solution engineering and, 160
acquiring and analyzing, 129–132
amount created, 111
analyzing, 129–132
capabilities of new architectures, 176
dark data, 16, 206
data federation, 183–184
growth of, business opportunities and, 194–195
how not to leverage, 106–107
low-latency data, access to, 83–84
management, solution engineering and, 160
need for new architecture, 173–175
preparation of, data scientists and, 43–44
Proof of Value analytic lab and, 158–159
structured and unstructured, 16, 82–83
data analytics lifecycle, 40–42
data assets
assessing, 32–33
digital media and, 31
inventoring, 11
monetizing, 33–34
transforming, augmenting and repackaging, 32
data monetization
basics of, 33–34
Big Data Strategy Document and, 66, 67, 69, 73
Data Monetization phase, 10–11
monetization impacts of business drivers, 157
opportunities, 30–34
POS scanner data and, 22–23
data science. See also advanced analytics; predictive analytics
BI and, 37–39
communication of results and, 45
methodologies, 50
organizational impact and, 37–39, 40
solution engineering and, 161
team, 48
data scientists
lifecycle and, 42
roles and responsibilities, 41, 42–46, 129, 133
data sources
Big Data Strategy Document and, 66, 67, 69, 73
San Francisco Giants and, 76
solution engineering and, 160
Starbucks and, 73
data warehouse, big data and
ELT, 181–183
Hadoop and, 185–187
MPP architectures, 187–188
schema on read, 184–185
virtualized data sources, 183–184
data warehouse managers, data analytics lifecycle and, 40–41
database administrators, data analytics lifecycle and, 40
Starbucks example of big data strategy, 70–73
transformation by big data (Moneyball), 27–30
video game market, changes in, 12–13
Walmart, transformation based on big data, 3–5
What If analytics cycle, 50–51
experimentation, organizational creativity and,
49–51

F
Facebook
Apache Hive and, 178
big data technologies and, 174–175
data monetization example, 30
facilitation team (vision workshops), 127
Feature Selection, 181
finance
financial services firm example, 168–169
solution engineering and, 153
Five Forces Analysis (Michael Porter)
basics of, 91–93
brainstorming and, 134
Foot Locker example, 101–104
four big data business drivers, 156–157
framing, decision making and, 61–62
fraud reduction example, 166

G
Giants (San Francisco) example, 73–77
Glass, David, 4
Google
advertising and analytics and, 53
big data technologies and, 174–175
data monetization example, 30
government-provided Consumer Price Index (CPI)
data, 132

H
Hadoop
architecture, 178
basics of, 176–177, 179
big data data warehouse and, 185–187
Hadoop Distributed File System (HDFS), 176–
177, 178, 179, 184
Hadoop MapReduce, 177
Hadoop: The Definitive Guide (O’Reilly and Yahoo!
Press), 179
HBase, 178
HDFS (Hadoop Distributed File System), 176–177,
178, 179, 184
high-velocity/low-latency data access, 156, 157
Hive, 178
human decision making, 58–62

I
ideation workshop, 132–139
aggregation, 136–138
brainstorming, 134–136
overview, 132–134
prioritizing big data use cases, 138–139
solution engineering and, 157–158
implementation feasibility, 141–142, 143, 144, 145
Inbound Logistics, 97–98, 99
in-database analytics, 188–190
Information Resources Inc. (IRI), 20
Infoscan, 20
insights
basics of, 56
of customers. See customer insights, uncovering and leveraging
incorporating into apps, 199
operationalizing, 199–200
insights-lead analytics process, 55
instrumentation
organizational creativity and, 49–51
strategy, 160
intelligent dashboards, business insights and, 7
interviews by vision workshop facilitation teams, 128–129

K
Kaplan, R. S., 113, 154
key business initiatives. See business initiatives

L
leveraging
customer insights. See customer insights, uncovering and leveraging
data, how not to leverage, 106–107
leverage mockups, 205
Lewis, Michael, 27
linear scalability, cloud computing and, 190
Long-Term Capital Management (LTCM) example, 58–59
low-latency data access
as business driver, 156, 157
data sources and data access requirements, solution engineering and, 160
low-latency data, access to
customer micro-segmentation example, 90
customer satisfaction example, 88
data monetization and, 81
merchandising examples, 96
predictive maintenance example, 87

M
MADlib, 180
Mahout, 180
management
Category Management, 21, 24
data warehouse managers, 40–41
Markdown Management, 21–22
with metrics, 27–30
senior management roles, 47–49
Map() procedure, 177
MapMyRun, 10
MapReduce
basics of, 176–177
Pig and, 178
Markdown Management, 21–22
Market Basket Analysis, 21
marketing
Business Insights phase and, 8
marketing effectiveness example, 164–165
solution engineering and, 153
Massively Parallel Processing (MPP) databases, 187
matrix
Prioritization Matrix, 140–141
prioritization process and, 138–139
prioritizing, 142–143
Media Planners and Buyers, as target users, 31
merchandising example, 94–103
merchandising example and. See merchandising example
Value Chain Analysis and, 97–101
merchant effectiveness example, 119–122
metrics, management with, 27–30
Michael Porter’s value creation models, 91–103
brainstorming and, 134
Five Forces Analysis, 91–93
merchandising example and. See merchandising example
Value Chain Analysis, 93–94
mobile devices. See also Smartphone apps
impact of, 1–2
mobile apps drive consumer experience, 117–118, 120
mockups, envisioning process and, 145–149, 205
model planning and building, 44–45
models. See also Michael Porter’s value creation models
analytic models (vision workshops), 129
data modeling, solution engineering and, 160–161
predictive solution engineering and, 163, 164
monetization. See also data monetization of data assets, 33–34
monetization opportunities, new, 206–207
money, how it is made, 153–155
Moneyball example, 27–30
monitoring, Business Metamorphosis phase, 7
Moore, Gordon, 188
Moore’s Law, 188
MPP architectures, 187–188

N
network optimization example, 166–167
New Market Entrants factors, 92, 100, 102, 103
Nielsen store audits in 70’s and 80’s, 19
Nintendo, 13

O
Oakland A’s example, 27–30
Online Transaction Process (OLTP)-centric
relational data base technologies, 173–174
operationalizing, data scientists and BI analysts and, 46
operations
new understanding of, 25, 26–27
operational process outlined in book, 200
solution engineering and, 153
organizational impact
BI and, 37–39, 40
data analytics lifecycle, 40–42
data science and, 37–39, 40
data scientist’s roles and responsibilities, 42–46
introduction to, 37–39
new organizational roles, 46–49
organizational creativity, liberating, 49–51
understanding, 207
Outbound Logistics, 98, 99
overconfidence, decision making and, 58–59

P
Pig, 178–179
platform- or ecosystem-centric business models, 12–15
point-of-sale (POS) scanner data
CPG-Retail industry transformation and, 22
data monetization and, 22–23
effects of, 1, 19, 23
Infoscan and, 20
lessons learned from, 23–24
Walmart and, 4, 20, 22, 24
Porter, Michael E... See Michael Porter’s value creation models
predictive analytics. See also advanced analytics;
data science
basics of, 176
as business driver, 81, 156, 157
customer micro-segmentation example and, 90
customer satisfaction example, drivers, 88–89
data monetization and, 81
integrating with old processes, 206
integration or adoption of as business driver #4, 84–85
merchandising examples, 97
predictive maintenance example, drivers and, 87
predictive maintenance example drivers, 86–87
solution engineering, 163–164
President’s Letter to the Shareholders, importance of, 168
Price and Yield Optimization, 21
primary activities (Value Chain Analysis), 93, 94
Prioritization Matrix
basics of, 140–142
end result, 145
process of, 142–143
taps, 143–145
use cases and, 158
prioritizing
matrix traps, 143–145
process of, 140–142
use cases, 138–139
Procter & Gamble
Category Management and, 24
eyear marketing data and, 19–20
Product and Technology Developments factors, 92, 100, 101, 103
product to platform or ecosystem companies, 13–16
product-centric business models, changes in, 12–15
products
intelligent, 10
new understanding of, 25, 26–27
profit
customer insights, using to drive profitability, 113–115
from customers over time, 115
most important customers and, 80–81
Proof of Value
in Business Optimization phase, 9
process, 158–159

Q
quarterly 10-Q financial filings, importance of, 167

R
R, 179
railroad company example, 86–87
Random Forest, 180
range of influence profitability, 80–81
real-time data access
solution engineering and, 160
as value creation driver, 83–84
recommendations
basics of, 56–57
as final stage of vision workshops, 140
to merchants, 120–121
personalized shopping recommendations, 118
Reduce() procedure, 177
retail example, annual reports and, 169–171
risk aversion, decision making and, 60
roadmaps, in solution engineering, 159–160
RStudio, 179

S
sales and marketing
identifying effective, 201
private label sales, increasing, 153, 170–171
solution engineering and, 153
San Francisco Giants example, 73–77
schema on read, 184–185
schema on write, 184
scope of vision workshops, 127
secondary activities (Value Chain Analysis), 93–94
senior management roles, 47–49
sentiment analysis, 181
“The Service Profit Chain”, 115
Smartphone apps
driving consumer experience and, 117–118, 120
for marketers and manufacturers, 10
mockup using, 146–148
social media
impact of, 1
impact of business initiative and, 88–89, 90
merchandising example and, 96, 99
mockup (PowerPoint), 205
traditional data management tools, 23, 40, 174
Trade Promotion Effectiveness, 21
traditional analytic process, 55
traditional data management tools, 23, 40, 174
traditional data warehouse. See also data warehouse, big data and
business growth and, 196
business monitoring and, 7
challenges created by, 2
transformation fueled by big data, 2–5
traps in decision making, 58–62
Twitter, data monetization and, 30

U
unstructured data
basics of access to, 82–83
as business driver, 156, 157
customer micro-segmentation example and, 90
customer satisfaction example, drivers, 88
data monetization and, 81
integrating with structured data, 169
merchandising examples, 96
predictive maintenance example, drivers and, 86
use cases
breaking down business initiatives into, 157–158
Proof of Value analytic lab and, 158–159
use cases, identifying, 125–150
documenting new steps, 139–140
prioritization matrix process, 142–143
prioritization matrix traps, 143–145
prioritization process, 140–142
prioritizing big data use cases, 138–139
user experience mockups, 145–149
using big data effectively, 125
vision workshops. See vision workshops
user experience (UX), 105–123
analytics to improve, 108–110
customer insights. See customer insights,
uncovering and leveraging
key decisions for, 107–108
mockups, envisioning process and, 145–149
new customer experience overview, 116
retail customer experience example, 116–119
small and medium-sized merchant effectiveness example, 119–122
solution engineering and, 154, 161
teams, 46–47
unintelligent user experience, 106–107
users
data analytics lifecycle and, 40, 41
new user interface, 56–58
user-centric design defined, 47

V
Value Chain Analysis (Michael Porter)
basics of, 93–94
brainstorming and, 134
value creation drivers
basics of, 81
customer micro-segmentation example and, 89–90
customer satisfaction example and, 87–89
driver #1 basics, 82
driver #2 basics, 82
driver #3 basics, 83–84
driver #4 basics, 84–85
envisioning worksheet: merchandising example and, 95–97
predictive maintenance example and, 86–87
value creation models. See Michael Porter’s value creation models
value creation process, 79–104
basics of, 79–81
big data and, 200–201
customer micro-segmentation example, 88–90
customer satisfaction example, 87–89
envisioning basics, 79
low-latency data, access to, 83–84
merchandising example. See merchandising example
predictive analytics, 84–85
predictive maintenance example, 86–87
structured data, access to, 82
unstructured data, access to, 82–83
value creation drivers. See value creation models
value creation models. See Michael Porter’s value creation models
value journey, 198–199
video game market, changes in, 12–13
virtual data warehouse, 183–184
virtualized data sources, 183–184
vision workshops
basics of, 125, 126
benefits of, 126
brainstorming. See ideation workshop
data, acquiring and analyzing, 129–132
documenting results of, 139–140
ideation workshop. See ideation workshop
mockups, envisioning process and, 145–149
researching business initiatives, 127–129
solution engineering and, 157–158

W
Walmart
big data impact on, 3–5
POS scanner data and, 4, 20, 22, 24
Walton, Sam, 4
Wavelet Transformation, 181
web clicks, 1
websites for downloading
Big Data Storymap, 202
sample worksheet, 85
websites for further information
“The 12 greatest entrepreneurs of our time”, 4
amount of data created every minute, 111
analytic algorithms, 180–181
Apache Hadoop technologies, 179
Apache Mahout, 180
Apache Software Foundation, 177
MADlib, 180
MapMyRun, 10
R and RStudio, 179
training and certification, 199
White, Tom, 179
worksheets (Big Data Envisioning Worksheets)
business drivers and, 85
Michael Porter, merchandising examples, 95–97,
101–104

Y
Yahoo!
big data technologies and, 174–175
customer usage variables vs. churn and, 131
data monetization example, 30
user experience and, 46

Z
zones (matrix), use cases and, 143–144