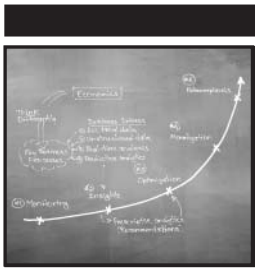
Big Data MBA



Big Data MBA

Driving Business Strategies with Data Science

Bill Schmarzo

WILEY

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Big Data MBA: Driving Business Strategies with Data Science

Published by
John Wiley & Sons, Inc.
10475 Crosspoint Boulevard
Indianapolis, IN 46256
www.wiley.com

Copyright © 2016 by Bill Schmarzo

Published by John Wiley & Sons, Inc., Indianapolis, Indiana

Published simultaneously in Canada

ISBN: 978-1-119-18111-8
ISBN: 978-1-119-23884-3 (ebk)
ISBN: 978-1-119-18138-5 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 646-8600. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permissions.

Limit of Liability/Disclaimer of Warranty: The publisher and the author make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation warranties of fitness for a particular purpose. No warranty may be created or extended by sales or promotional materials. The advice and strategies contained herein may not be suitable for every situation. This work is sold with the understanding that the publisher is not engaged in rendering legal, accounting, or other professional services. If professional assistance is required, the services of a competent professional person should be sought. Neither the publisher nor the author shall be liable for damages arising herefrom. The fact that an organization or website is referred to in this work as a citation and/or a potential source of further information does not mean that the author or the publisher endorses the information the organization or website may provide or recommendations it may make. Further, readers should be aware that Internet websites listed in this work may have changed or disappeared between when this work was written and when it is read.

For general information on our other products and services please contact our Customer Care Department within the United States at (877) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley publishes in a variety of print and electronic formats and by print-on-demand. Some material included with standard print versions of this book may not be included in e-books or in print-on-demand. If this book refers to media such as a CD or DVD that is not included in the version you purchased, you may download this material at http://booksupport.wiley.com. For more information about Wiley products, visit www.wiley.com.

Library of Congress Control Number: 2015955444

Trademarks: Wiley and the Wiley logo are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates, in the United States and other countries, and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc. is not associated with any product or vendor mentioned in this book.

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

About the Author

Bill Schmarzo is the Chief Technology Officer (CTO) of the Big Data Practice of EMC Global Services. As CTO, Bill is responsible for setting the strategy and defining the big data service offerings and capabilities for EMC Global Services. He also works directly with organizations to help them identify where and how to start their big data journeys. Bill is the author of *Big Data: Understanding How Data Powers Big Business*, writes white papers, is an avid blogger, and is a frequent speaker on the use of big data and data science to power an organization's key business initiatives. He is a University of San Francisco School of Management (SOM) Fellow, where he teaches the "Big Data MBA" course.

Bill has over three decades of experience in data warehousing, business intelligence, and analytics. He authored EMC's Vision Workshop methodology and co-authored with Ralph Kimball a series of articles on analytic applications. Bill has served on The Data Warehouse Institute's faculty as the head of the analytic applications curriculum. Previously, he was the Vice President of Analytics at Yahoo! and oversaw the analytic applications business unit at Business Objects, including the development, marketing, and sales of their industry-defining analytic applications.

Bill holds a master's degree in Business Administration from the University of Iowa and a Bachelor of Science degree in Mathematics, Computer Science, and Business Administration from Coe College. Bill's recent blogs can be found at `http://infocus.emc.com/author/william_schmarzo/`. You can follow Bill on Twitter @schmarzo and LinkedIn at `www.linkedin.com/in/schmarzo`.

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

About the Technical Editor

Jeffrey Abbott leads the EMC Global Services marketing practice around big data, helping customers understand how to identify and take advantage of opportunities to leverage data for strategic business initiatives, while driving awareness for a portfolio of services offerings that accelerate customer time-to-value. As a content developer and program lead, Jeff emphasizes clear and concise messaging on persona-based campaigns. Prior to EMC, Jeff helped build and promote a cloud-based ecosystem for CA Technologies that combined an online social community, a cloud development platform, and an e-commerce site for cloud services. Jeff also spent several years within CA's Thought Leadership group, creating and promoting executive-level messaging and social-media programs around major disruptive trends in IT. Jeff has held various other product marketing roles at firms such as EMC, Citrix, and Ardence and spent a decade running client accounts at numerous boutique marketing firms. Jeff studied small business management at the University of Vermont and resides in Sudbury, MA, with his wife, two boys, and dog. Jeff enjoys skiing, backpacking, photography, and classic cars.

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Credits

Project Editor
Adaobi Obi Tulton

Technical Editor
Jeffrey Abbott

Production Editor
Barath Kumar Rajasekaran

Copy Editor
Chris Haviland

Manager of Content Development & Assembly
Mary Beth Wakefield

Production Manager
Kathleen Wisor

Marketing Director
David Mayhew

Marketing Manager
Carrie Sherrill

Professional Technology & Strategy Director
Barry Pruett

Business Manager
Amy Knies

Associate Publisher
Jim Minatel

Project Coordinator, Cover
Brent Savage

Proofreader
Nicole Hirschman

Indexer
Nancy Guenther

Cover Designer
Wiley

Cover Image
©STILLFX/iStockphoto

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Acknowledgments

Acknowledgments are dangerous. Not dangerous like wrestling an alligator or an unhappy Chicago Cubs fan, but dangerous in the sense that there are so many people to thank. How do I prevent the Acknowledgments section from becoming longer than my book? This book represents the sum of many, many discussions, debates, presentations, engagements, and late night beers and pizza that I have had with so many colleagues and customers. Thanks to everyone who has been on this journey with me.

So realizing that I will miss many folks in this acknowledgment, here I go…

I can't say enough about the contributions of Jeff Abbott. Not only was Jeff my EMC technical editor for this book, but he also has the unrewarding task of editing all of my blogs. Jeff has the patience to put up with my writing style and the smarts to know how to spin my material so that it is understandable and readable. I can't thank Jeff enough for his patience, guidance, and friendship.

Jen Sorenson's role in the book was only supposed to be EMC Public Relations editor, but Jen did so much more. There are many chapters in this book where Jen's suggestions (using the Fairy-Tale Theme Parks example in Chapter 6) made the chapters more interesting. In fact, Chapter 6 is probably my favorite chapter because I was so over my skis on the data science algorithms material. But Jen did a marvelous job of taking a difficult topic (data science algorithms) and making it come to life.

Speaking of data science, Pedro DeSouza and Wei Lin are the two best data scientists I have ever met, and I am even more grateful that I get to call them friends. They have been patient in helping me to learn the world of data science over the past several years, which is reflected in many chapters in the book (most notably Chapters 5 and 6). But more than anything else, they taught me a very

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

valuable life lesson: being humble is the best way to learn. I can't even express in words my admiration for them and how they approach their profession.

Joe Dossantos and Josh Siegel may be surprised to find their names in the acknowledgments, but they shouldn't be. Both Joe and Josh have been with me on many steps in this big data journey, and both have contributed tremendously to my understanding of how big data can impact the business world. Their fingerprints are all over this book.

Adaobi Obi Tulton and Chris Haviland are my two Wiley editors, and they are absolutely marvelous! They have gone out of their way to make the editing process as painless as possible, and they understand my voice so well that I accepted over 99 percent of all of their suggestions. Both Adaobi and Chris were my editors on my first book, so I guess they forgot how much of a PITA (pain in the a**) I can be when they agreed to be the editors on my second book. Though I have never met them face-to-face, I feel a strong kinship with both Adaobi and Chris. Thanks for all of your patience and guidance and your wonderful senses of humor!

A very special thank you to Professor Mouwafac Sidaoui, with whom I co-teach the Big Data MBA at the University of San Francisco School of Management (USF SOM). I could not pick a better partner in crime—he is smart, humble, demanding, fun, engaging, worldly, and everything that one could want in a friend. I am a Fellow at the USF SOM because of Mouwafac's efforts, and he has set me up for my next career—teaching.

I also what to thank Dean Elizabeth Davis and the USF MBA students who were willing to be guinea pigs for testing many of the concepts and techniques captured in this book. They helped me to determine which ideas worked and how to fix the ones that did not work.

Another special thank you to EMC, who supported me as I worked at the leading edge of the business transformational potential of big data. EMC has afforded me the latitude to pursue new ideas, concepts, and offerings and in many situations has allowed me to be the tip of the big data arrow. I could not ask for a better employer and partner.

The thank you list should include the excellent and creative people at EMC with whom I interact on a regular basis, but since that list is too long, I'll just mention Ed, Jeff, Jason, Paul, Dan, Josh, Matt, Joe, Scott, Brandon, Aidan, Neville, Bart, Billy, Mike, Clark, Jeeva, Sean, Shriya, Srini, Ken, Mitch, Cindy, Charles, Chuck, Peter, Aaron, Bethany, Susan, Barb, Jen, Rick, Steve, David, and many, many more.

I want to thank my family, who has put up with me during the book writing process. My wife Carolyn was great about grabbing Chipotle for me when I had a tough deadline, and my sons Alec and Max and my daughter Amelia were supportive throughout the book writing process. I've been blessed with a marvelous family (just stop stealing my Chipotle in the refrigerator!).

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

My mom and dad both passed away, but I can imagine their look of surprise and pride in the fact that I have written two books and am teaching at the University of San Francisco in my spare time. We will get the chance to talk about that in my next life.

But most important, I want to thank the EMC customers with whom I have had the good fortune to work. Customers are at the frontline of the big data transformation, and where better to be situated to learn about what's working and what's not working then arm-in-arm with EMC's most excellent customers at those frontlines. Truly the best part of my job is the chance to work with our customers. Heck, I'm willing to put up with the airline travel to do that!

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Contents at a Glance

Introduction		xxiii
Part I	Business Potential of Big Data	1
Chapter 1	The Big Data Business Mandate	3
Chapter 2	Big Data Business Model Maturity Index	17
Chapter 3	The Big Data Strategy Document	35
Chapter 4	The Importance of the User Experience	61
Part II	Data Science	83
Chapter 5	Differences Between Business Intelligence and Data Science	85
Chapter 6	Data Science 101	107
Chapter 7	The Data Lake	133
Part III	Data Science for Business Stakeholders	153
Chapter 8	Thinking Like a Data Scientist	155
Chapter 9	"By" Analysis Technique	171
Chapter 10	Score Development Technique	183
Chapter 11	Monetization Exercise	199
Chapter 12	Metamorphosis Exercise	211

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Part IV	**Building Cross-Organizational Support**	**229**
Chapter 13	**Power of Envisioning**	**231**
Chapter 14	**Organizational Ramifications**	**245**
Chapter 15	**Stories**	**255**
Index		**269**

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Contents

Introduction		**xxiii**
Part I	**Business Potential of Big Data**	**1**
Chapter 1	**The Big Data Business Mandate**	**3**
	Big Data MBA Introduction	4
	Focus Big Data on Driving Competitive Differentiation	6
	Leveraging Technology to Power Competitive Differentiation	7
	History Lesson on Economic-Driven Business Transformation	7
	Critical Importance of "Thinking Differently"	10
	Don't Think Big Data Technology, Think Business Transformation	10
	Don't Think Business Intelligence, Think Data Science	11
	Don't Think Data Warehouse, Think Data Lake	11
	Don't Think "What Happened," Think "What Will Happen"	12
	Don't Think HIPPO, Think Collaboration	14
	Summary	14
	Homework Assignment	15
Chapter 2	**Big Data Business Model Maturity Index**	**17**
	Introducing the Big Data Business Model Maturity Index	18
	Phase 1: Business Monitoring	20
	Phase 2: Business Insights	21
	Phase 3: Business Optimization	25
	Phase 4: Data Monetization	27
	Phase 5: Business Metamorphosis	28
	Big Data Business Model Maturity Index Lessons Learned	30
	Lesson 1: Focus Initial Big Data Efforts Internally	30
	Lesson 2: Leverage Insights to Create New Monetization Opportunities	31
	Lesson 3: Preparing for Organizational Transformation	32

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

	Summary	33
	Homework Assignment	34
Chapter 3	**The Big Data Strategy Document**	**35**
	Establishing Common Business Terminology	37
	Introducing the Big Data Strategy Document	37
	Identifying the Organization's Key Business Initiatives	39
	What's Important to Chipotle?	40
	Identify Key Business Entities and Key Decisions	41
	Identify Financial Drivers (Use Cases)	45
	Identify and Prioritize Data Sources	48
	Introducing the Prioritization Matrix	51
	Using the Big Data Strategy Document to Win the World Series	52
	Summary	57
	Homework Assignment	58
Chapter 4	**The Importance of the User Experience**	**61**
	The Unintelligent User Experience	62
	Capture the Key Decisions	63
	Support the User Decisions	63
	Consumer Case Study: Improve Customer Engagement	64
	Business Case Study: Enable Frontline Employees	66
	Store Manager Dashboard	67
	Sample Use Case: Competitive Analysis	69
	Additional Use Cases	70
	B2B Case Study: Make the Channel More Effective	71
	The Advisors Are Your Partners—Make Them Successful	72
	Financial Advisor Case Study	72
	Informational Sections of Financial Advisor Dashboard	74
	Recommendations Section of Financial Advisor Dashboard	77
	Summary	80
	Homework Assignment	81
Part II	**Data Science**	**83**
Chapter 5	**Differences Between Business Intelligence and Data Science**	**85**
	What Is Data Science?	86
	BI Versus Data Science: The Questions Are Different	87
	BI Questions	88
	Data Science Questions	88
	The Analyst Characteristics Are Different	89
	The Analytic Approaches Are Different	91
	Business Intelligence Analyst Engagement Process	91
	The Data Scientist Engagement Process	93
	The Data Models Are Different	96

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

	Data Modeling for BI	96
	Data Modeling for Data Science	98
	The View of the Business Is Different	100
	Summary	104
	Homework Assignment	104
Chapter 6	**Data Science 101**	**107**
	Data Science Case Study Setup	107
	Fundamental Exploratory Analytics	110
	Trend Analysis	110
	Boxplots	112
	Geographical (Spatial) Analysis	113
	Pairs Plot	114
	Time Series Decomposition	115
	Analytic Algorithms and Models	116
	Cluster Analysis	116
	Normal Curve Equivalent (NCE) Analysis	117
	Association Analysis	119
	Graph Analysis	121
	Text Mining	122
	Sentiment Analysis	123
	Traverse Pattern Analysis	124
	Decision Tree Classifier Analysis	125
	Cohorts Analysis	126
	Summary	128
	Homework Assignment	131
Chapter 7	**The Data Lake**	**133**
	Introduction to the Data Lake	134
	Characteristics of a Business-Ready Data Lake	136
	Using the Data Lake to Cross the Analytics Chasm	137
	Modernize Your Data and Analytics Environment	140
	Action #1: Create a Hadoop-Based Data Lake	140
	Action #2: Introduce the Analytics Sandbox	141
	Action #3: Off-Load ETL Processes from Data Warehouses	142
	Analytics Hub and Spoke Analytics Architecture	143
	Early Learnings	145
	Lesson #1: The Name Is Not Important	145
	Lesson #2: It's Data Lake, Not Data Lakes	146
	Lesson #3: Data Governance Is a Life Cycle, Not a Project	147
	Lesson #4: Data Lake Sits Before Your Data Warehouse, Not After It	148
	What Does the Future Hold?	149
	Summary	150
	Homework Assignment	151

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Part III	**Data Science for Business Stakeholders**	**153**
Chapter 8	**Thinking Like a Data Scientist**	**155**
	The Process of Thinking Like a Data Scientist	157
	Step 1: Identify Key Business Initiative	157
	Step 2: Develop Business Stakeholder Personas	158
	Step 3: Identify Strategic Nouns	160
	Step 4: Capture Business Decisions	161
	Step 5: Brainstorm Business Questions	162
	Step 8: Putting Analytics into Action	166
	Summary	168
	Homework Assignment	169
Chapter 9	**"By" Analysis Technique**	**171**
	"By" Analysis Introduction	172
	"By" Analysis Exercise	174
	Foot Locker Use Case "By" Analysis	178
	Summary	181
	Homework Assignment	182
Chapter 10	**Score Development Technique**	**183**
	Definition of a Score	184
	FICO Score Example	185
	Other Industry Score Examples	188
	LeBron James Exercise Continued	189
	Foot Locker Example Continued	193
	Summary	197
	Homework Assignment	197
Chapter 11	**Monetization Exercise**	**199**
	Fitness Tracker Monetization Example	200
	Step 1: Understand Product Usage	200
	Step 2: Develop Stakeholder Personas	201
	Step 3: Brainstorm Potential Recommendations	203
	Step 4: Identify Supporting Data Sources	204
	Step 5: Prioritize Monetization Opportunities	206
	Step 6: Develop Monetization Plan	208
	Summary	209
	Homework Assignment	210
Chapter 12	**Metamorphosis Exercise**	**211**
	Business Metamorphosis Review	212
	Business Metamorphosis Exercise	213
	Articulate the Business Metamorphosis Vision	214
	Understand Your Customers	215
	Articulate Value Propositions	215
	Define Data and Analytic Requirements	216
	Business Metamorphosis in Health Care	223

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

	Summary	226
	Homework Assignment	227
Part IV	**Building Cross-Organizational Support**	**229**
Chapter 13	**Power of Envisioning**	**231**
	Envisioning: Fueling Creative Thinking	232
	Big Data Vision Workshop Process	232
	Pre-engagement Research	233
	Business Stakeholder Interviews	234
	Explore with Data Science	235
	Workshop	236
	Setting Up the Workshop	239
	The Prioritization Matrix	241
	Summary	243
	Homework Assignment	244
Chapter 14	**Organizational Ramifications**	**245**
	Chief Data Monetization Officer	245
	CDMO Responsibilities	246
	CDMO Organization	246
	Analytics Center of Excellence	247
	CDMO Leadership	248
	Privacy, Trust, and Decision Governance	248
	Privacy Issues = Trust Issues	249
	Decision Governance	250
	Unleashing Organizational Creativity	251
	Summary	253
	Homework Assignment	254
Chapter 15	**Stories**	**255**
	Customer and Employee Analytics	257
	Product and Device Analytics	261
	Network and Operational Analytics	263
	Characteristics of a Good Business Story	265
	Summary	266
	Homework Assignment	267
Index		**269**

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Introduction

I never planned on writing a second book. Heck, I thought writing one book was enough to check this item off my bucket list. But so much has changed since I wrote my first book that I felt compelled to continue to explore this once-in-a-lifetime opportunity for organizations to leverage data and analytics to transform their business models. And I'm not just talking the "make me more money" part of businesses. Big data can drive significant "improve the quality of life" value in areas such as education, poverty, parole rehabilitation, health care, safety, and crime reduction.

My first book targeted the Information Technology (IT) audience. However, I soon realized that the biggest winner in this big data land grab was the business. So this book targets the business audience and is based on a few key premises:

- Organizations do not need a big data strategy as much as they need a business strategy that incorporates big data.
- The days when business leaders could turn analytics over to IT are over; tomorrow's business leaders must embrace analytics as a business discipline in the same vein as accounting, finance, management science, and marketing.
- The key to data monetization and business transformation lies in unleashing the organization's creative thinking; we have got to get the business users to "think like a data scientist."
- Finally, the business potential of big data is only limited by the creative thinking of the business users.

I've also had the opportunity to teach "Big Data MBA" at the University of San Francisco (USF) School of Management since I wrote the first book. I did well enough that USF made me its first School of Management Fellow. What I

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

experienced while working with these outstanding and creative students and Professor Mouwafac Sidaoui compelled me to undertake the challenge of writing this second book, targeting those students and tomorrow's business leaders.

One of the topics that I hope jumps out in the book is the power of data science. There have been many books written about data science with the goal of helping people to become data scientists. But I felt that something was missing—that instead of trying to create a world of data scientists, we needed to help tomorrow's business leaders think like data scientists.

So that's the focus of this book—to help tomorrow's business leaders integrate data and analytics into their business models and to lead the cultural transformation by unleashing the organization's creative juices by helping the business to "think like a data scientist."

Overview of the Book and Technology

The days when business stakeholders could relinquish control of data and analytics to IT are over. The business stakeholders must be front and center in championing and monetizing the organization's data collection and analysis efforts. Business leaders need to understand where and how to leverage big data, exploiting the collision of new sources of customer, product, and operational data coupled with data science to optimize key business processes, uncover new monetization opportunities, and create new sources of competitive differentiation. And while it's not realistic to convert your business users into data scientists, it's critical that we teach the business users to *think like data scientists* so they can collaborate with IT and the data scientists on use case identification, requirements definition, business valuation, and ultimately analytics operationalization.

This book provides a business-hardened framework with supporting methodology and hands-on exercises that not only will help business users to identify where and how to leverage big data for business advantage but will also provide guidelines for operationalizing the analytics, setting up the right organizational structure, and driving the analytic insights throughout the organization's user experience to both customers and frontline employees.

How This Book Is Organized

The book is organized into four sections:

- **Part I: Business Potential of Big Data**. Part I includes Chapters 1 through 4 and sets the business-centric foundation for the book. Here is where I introduce the Big Data Business Model Maturity Index and frame the big data discussion around the perspective that "organizations do not

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

need a big data strategy as much as they need a business strategy that incorporates big data."

- **Part II: Data Science**. Part II includes Chapters 5 through 7 and covers the principle behind data science. These chapters introduce some data science basics and explore the complementary nature of Business Intelligence and data science and how these two disciplines are both complementary and different in the problems that they address.
- **Part III: Data Science for Business Stakeholders**. Part III includes Chapters 8 through 12 and seeks to teach the business users and business leaders to "think like a data scientist." This part introduces a methodology and several exercises to reinforce the data science thinking and approach. It has a lot of hands-on work.
- **Part IV: Building Cross-Organizational Support**. Part IV includes Chapters 13 through 15 and discusses organizational challenges. This part covers envisioning, which may very well be the most important topic in the book as the business potential of big data is only limited by the creative thinking of the business users.

Here are some more details on each of the chapters in the book:

- **Chapter 1: The Big Data Business Mandate**. This chapter frames the big data discussion on how big data is more about business transformation and the economics of big data than it is about technology.
- **Chapter 2: Big Data Business Model Maturity Index**. This chapter covers the Big Data Business Model Maturity Index (BDBM), which is the foundation for the entire book. Take the time to understand each of the five stages of the BDBM and how the BDBM provides a road map for measuring how effective your organization is at integrating data and analytics into your business models.
- **Chapter 3: The Big Data Strategy Document**. This chapter introduces a CXO level document and process for helping organizations identify where and how to start their big data journeys from a business perspective.
- **Chapter 4: The Importance of the User Experience**. This is one of my favorite topics. This chapter challenges traditional Business Intelligence reporting and dashboard concepts by introducing a more simple but direct approach for delivering actionable insights to your key business stakeholders—frontline employees, channel partners, and end customers.
- **Chapter 5: Differences Between Business Intelligence and Data Science**. This chapter explores the different worlds of Business Intelligence and data science and highlights both the differences and the complementary nature of each.

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

- **Chapter 6: Data Science 101**. This chapter (my favorite) reviews 14 different analytic techniques that my data science teams commonly use and in what business situations you should contemplate using them. It is accompanied by a marvelous fictitious case study using Fairy-Tale Theme Parks (thanks Jen!).
- **Chapter 7: The Data Lake**. This chapter introduces the concept of a data lake, explaining how the data lake frees up expensive data warehouse resources and unleashes the creative, fail-fast nature of the data science teams.
- **Chapter 8: Thinking Like a Data Scientist**. The heart of this book, this chapter covers the eight-step "thinking like a data scientist" process. This chapter is pretty deep, so plan on having a pen and paper (and probably an eraser as well) with you as you read this chapter.
- **Chapter 9: "By" Analysis Technique.** This chapter does a deep dive into one of the important concepts in "thinking like a data scientist"—the "By" analysis technique.
- **Chapter 10: Score Development Technique.** This chapter introduces how scores can drive collaboration between the business users and data scientist to create actionable scores that guide the organization's key business decisions.
- **Chapter 11: Monetization Exercise**. This chapter provides a technique for organizations that have a substantial amount of customer, product, and operational data but do not know how to monetize that data. This chapter can be very eye-opening!
- **Chapter 12: Metamorphosis Exercise**. This chapter is a fun, out-of-the-box exercise that explores the potential data and analytic impacts for an organization as it contemplates the Business Metamorphosis phase of the Big Data Business Model Maturity Index.
- **Chapter 13: Power of Envisioning**. This chapter starts to address some of the organizational and cultural challenges you may face. In particular, Chapter 13 introduces some envisioning techniques to help unleash your organization's creative thinking.
- **Chapter 14: Organizational Ramifications**. This chapter goes into more detail about the organizational ramifications of big data, especially the role of the Chief Data (Monetization) Officer.
- **Chapter 15: Stories**. The book wraps up with some case studies, but not your traditional case studies. Instead, Chapter 15 presents a technique for creating "stories" that are relevant to your organization. Anyone can find case studies, but not just anyone can create a story.

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Who Should Read This Book

This book is targeted toward business users and business management. I wrote this book so that I could use it in teaching my Big Data MBA class, so included all of the hands-on exercises and templates that my students would need to successfully earn their Big Data MBA graduation certificate.

I think folks would benefit by also reading my first book, *Big Data: Understanding How Data Powers Big Business*, which is targeted toward the IT audience. There is some overlap between the two books (10 to 15 percent), but the first book sets the stage and introduces concepts that are explored in more detail in this book.

Tools You Will Need

No special tools are required other than a pencil, an eraser, several sheets of paper, and your creativity. Grab a chai tea latte, some Chipotle, and enjoy!

What's on the Website

You can download the "Thinking Like a Data Scientist" workbook from the book's website at `www.wiley.com/go/bigdatamba`. And oh, there might be another surprise there as well! Hehehe!

What This Means for You

As students from my class at USF have told me, this material allows them to take a problem or challenge and use a well-thought-out process to drive cross-organizational collaboration to come up with ideas they can turn into actions using data and analytics. What employer wouldn't want a future leader who knows how to do that?

10.1002/9781119238881.fmatter, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/9781119238881.fmatter by Cochrane Canada Provision, Wiley Online Library on [12/01/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License