

# 商业数据分析

## 原理、方法与应用

### BUSINESS ANALYTICS

PRINCIPLES, CONCEPTS, AND APPLICATIONS

WHAT, WHY, and HOW



[美] 马克·J.施尼德詹斯 (Marc J. Schniederjans) 著  
达拉·G.施尼德詹斯(Dara G. Schniederjans) 著  
克里斯多夫·M.斯塔基(Christopher M. Starkey) 著  
王忠玉 王天元 王伟 译

 机械工业出版社  
CHINA MACHINE PRESS

# 商业数据分析

——原理、方法与应用

**Business Analytics Principles, Concepts,  
and Applications: What, Why, and How**

马克·J. 施尼德詹斯(Marc J. Schniederjans)

[美] 达拉·G. 施尼德詹斯(Dara G. Schniederjans) 著

克里斯多夫·M. 斯塔基(Christopher M. Starkey)

王忠玉 王天元 王伟 译



机械工业出版社

Authorized translation from the English language edition, entitled Business Analytics Principles, Concepts, and Applications: What, Why, and How, 1<sup>st</sup> Edition, 9780133552188, by Marc J. Schniederjans, Dara G. Schniederjans, Christopher M. Starkey, published by Pearson Education, Inc, Copyright © 2014 by Marc J. Schniederjans, Dara G. Schniederjans, and Christopher M. Starkey.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

CHINESE SIMPLIFIED language edition published by CHINA MACHINE PRESS, Copyright © 2018 by China Machine Press.

本书封面贴有 Pearson Education (培生教育出版集团) 激光防伪标签。  
无标签者不得销售。

北京市版权局著作权合同登记 图字: 01-2015-5832 号。

### 图书在版编目(CIP)数据

商业数据分析: 原理、方法与应用 / (美) 马克·J. 施尼德詹斯等著; 王忠玉, 王天元, 王伟译. —北京: 机械工业出版社, 2018. 6

书名原文: Business Analytics Principles, Concepts, and Applications: What, Why, and How

ISBN 978-7-111-60209-5

I. ①商… II. ①马… ②王… ③王… ④王… III. ①商业信息-数据处理  
IV. ①F713.51

中国版本图书馆 CIP 数据核字(2018)第 128536 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑: 易敏 责任编辑: 易敏 刘静 刘鑫佳

责任校对: 刘秀芝 封面设计: 鞠杨

责任印制: 张博

三河市国英印务有限公司印刷

2018 年 9 月第 1 版第 1 次印刷

169mm × 239mm · 17.25 印张 · 330 千字

标准书号: ISBN 978-7-111-60209-5

定价: 59.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换  
电话服务 网络服务

服务咨询热线: 010-88361066 机工官网: [www.cmpbook.com](http://www.cmpbook.com)

读者购书热线: 010-68326294 机工官博: [weibo.com/cmp1952](http://weibo.com/cmp1952)

010-88379203 金书网: [www.golden-book.com](http://www.golden-book.com)

封面无防伪标均为盗版 教育服务网: [www.cmpedu.com](http://www.cmpedu.com)

针对商业数据分析的教学需求，本书提供了综合而完整的知识体系。本书介绍了最新的商业数据分析理念，阐述了什么是商业数据分析、为什么它如此有价值、如何应用它，并介绍了现在常用的重要统计技术、方法和软件工具。本书阐述了商业分析中的三种数据分析——描述性数据分析、预测性数据分析、规范性数据分析，并指导用户认识理解自己的数据分析项目。本书将基本概念与工具、技术及实际应用结合起来，通过实例，循序渐进地厘清那些企业或其他组织中可以用商业数据分析来解决的共性问题。事实已经证明，商业数据分析的原则、概念和应用提供了行之有效的理论和方法，成功地将它整合到组织当中，能帮助组织有效提升智能决策的竞争优势。

本书适合高校相关专业学生作为教材使用，也可供想了解大数据、数据分析如何帮助商业运营的管理者参考。

## 推 荐 序

我们正处于一个变革的时代，面对大数据、人工智能（AI）的崛起所带来的巨大机遇与挑战。一方面，我们需要深刻洞察技术变革的趋势，树立大数据时代的商业思维，特别是看清对商业运营管理和商业模式所带来的冲击，并及时做出应对策略，在管理形式和方法上做出响应和调整；另一方面，在商业、经济及其他领域中，管理决策行为将日益基于数据和分析做出，而并非基于经验和直觉，这些正是管理学院或商学院在人才培养中面临的新挑战。

最近，国外一些知名大学的管理学院或商学院纷纷创建了商业数据分析（Business Analytics，又称为商业分析）硕士项目或专业。商业数据分析是融合了商业经济管理、统计学、计算机科学的新兴交叉学科。商业数据分析既注重经济、金融、管理知识的掌握和运用，又关注大数据分析方法及人工智能技术的学习和训练，目的是促使数据分析工作更具有预测潜力，从而提高公司业绩和战略管理能力，最终提高运营效率。简言之，商业数据分析搭建了数据分析与商业管理之间的桥梁，是利用各种分析工具来研究数据（包含大数据），寻找有用信息，进而更好地进行决策的一门新兴交叉学科。

商业数据分析着重培养学生综合运用专业知识解决理论与实际问题的能力。这些能力的获得，通常离不开管理学院或商学院规划的知识体系和实践能力培养，当然也离不开这一领域的优秀教材。

这本书由马克·J. 施尼德詹斯、达拉·G. 施尼德詹斯及克里斯多夫·M. 斯塔基合著，是商业数据分析领域中一本非常优秀的教材。全书以商业数据分析为主线，详细地讲述了商业数据分析三个基本部分——描述性数据分析、预测性数据分析、规范性数据分析的内容。全书结构清晰，观点明确，阐述详尽，语言简练。

正如作者在书中前言所说，“本书为商业数据分析的基础教学而编写，既可服务于大学教育，也可服务于实践者。除了为商业数据分析提供最新的文献和研究外，本书还利用简单的术语及在实践中十分有用的辅助方法，来解释商业数据分析流程。”

这本书具有以下三个特点：①紧跟当代大数据、人工智能变革的发展

前沿，密切追踪业界发展动态；②理论与应用实践更加紧密，在给出有关理论的阐述讲解之后，辅以案例应用，体现学以致用；③知识体系完整，在讲述商业数据分析内容时，突出逻辑主线，深入浅出，图文并茂，并将有关辅助工具的内容放在附录之中。

本书由机械工业出版社正式出版中译本，译者之一王忠玉老师曾翻译过多本经济管理方面的书籍，比如前不久出版的《管理数据分析——原理、方法、工具及实践》中译本（机械工业出版社，2017年出版）也聚焦于商业数据分析领域的，只是侧重点有所不同，将其作为补充阅读更为有益。

李一军

教授、博士生导师

“长江学者”特聘教授

于 哈尔滨工业大学经济与管理学院

## 译者的话

当今，大数据（Big Data）或数据（Data）正以爆炸方式涌现在社会经济生活的方方面面，各行各业都在产生和利用着数据。比如人们在网络上购物、通信、浏览新闻、收听在线音乐、搜索信息，或在网上表达观点，这些行为都会生成原始数据而被逐一记录。当前，人们不仅拥有大量的数据（包括大数据），而且计算机的运算能力也有了快速的提升，这些为利用数据从事商业研究，特别是从数据挖掘和提取有价值的信息，提供了必要的基本条件和良好的环境。

随着大量丰富多彩的数据源不断地涌现的新时代的到来，数据成为一种新的资源。关于如何从数据中挖掘和提取有价值的信息的例子，这里选取2012年沃尔玛依据飓风预报而安排备货的案例来说明。

在飓风 Frances 来临的一周之前，沃尔玛公司的首席信息官（CIO）Linda M. Dillman 督促她的团队，根据几周之前飓风 Charley 来袭期间沃尔玛的销售数据，对新飓风来袭时的销售进行预测。之前的海量销售数据被保存在数据仓库中，公司利用这些数据预测销售，目的是提前备货，进而提高公司的销售额。公司分析人员对那些销售数据进行挖掘分析，发现了一些不同寻常的需求特征。他们发现，人们确实更多地购买了某些特定的产品，但不是主观猜测的手电筒等。比如，他们发现，飓风到来前，草莓馅饼的销售量出现了增长，是平时销量的7倍左右，而最畅销的产品则是啤酒。据此，他们提前配备货物，极大地提高了公司的销售额。

当今，信息技术，尤其是互联网技术、计算机技术的不断进步和发展降低了人们获取数据、存储数据和传输数据的成本，使得越来越多的企业和政府机构等社会组织有能力从自身的业务系统，或者通过互联网等其他途径，获取规模日益庞大的数据。对于公司来说，数据的价值越来越高，人们更加重视对历史数据的积累，进行数据分析，挖掘和提取有价值的信息，以便高效、正确地认识和掌握事物的发展变动趋势，帮助决策者更好地进行决策。

要想认识和理解商业数据分析的理论及方法，一定离不开当前所处的时代背景，更离不开数据科学的基本理论和技术的支持。下面，首先简略介绍有关数据科学的基础知识，然后给出作为数据科学中一个特定的应用

领域——商业数据分析的有关内容。

### 1. 数据科学

什么是数据科学 (Data Science)? 依据维基百科的释义, 数据科学是一门利用数据来学习知识的学科, 其目标是从数据中提取出有价值的信息。它结合了诸多领域中的理论和技术, 包括数学、统计学、计算机科学, 特别是计算机科学领域的数据库、大数据、机器学习、数据挖掘、数据可视化等子领域的相关技术和方法。

从数据科学的内涵来看, 它是一个全新的概念, 试图把统计方法和数据分析方法统一起来, 目的是分析和理解客观现象产生的数据。数据科学的核心任务是从数据中抽取信息、发现知识。其研究对象是各种各样的数据及其特性。

为了认识和理解数据科学的内涵, 首先从数据概念开始介绍。如果从数据科学的研究对象来看, 数据是指利用测量工具所获得的未经处理的事实或数字。这些数据的类型多种多样, 不仅包括传统的结构化数据, 也包括网页、文本、图像、视频、语音等非结构化数据。这样定义的范畴既超出数学关于“数”的定义, 又扩展了统计学关于“数据”的定义。实际上, 比较常见的数据包括以下几种类型:

(1) 表格。这是经典的数据类型。在表格数据中, 通常行表示样本, 列表示特征。

(2) 点集 (Point Set)。许多数据被认为是某空间中的点的集合, 还有点云 (Point Cloud) 数据。

(3) 时间序列。文本、通话和 DNA 序列等都可以看成是时间序列。它们也是某种变量 (通常认为是时间) 的函数。

(4) 图像。可以看成是有两个变量的函数。

(5) 视频。可以看成以时间和空间为坐标的函数。

(6) 网页和报纸。虽然网页或报纸上的每篇文章都可以看成是时间序列, 但整个网页或报纸又具有空间结构。

(7) 网络数据。网络本质上是图, 由节点和联系节点的边构成。

(8) 模糊数据。本质上由各种模糊数构成。

除了前述几种基本数据类型之外, 还存在更高层次的数据, 例如图像集、时间序列集、表格序列等。

数据科学的理论与方法通常包括一组概念、原则、过程、技术/方法及工具, 为其核心任务服务。其中, 概念和基本原则给予人们观察问题、解决问题的一套完整的思想框架, 而大量的数据分析技术/方法和工具则帮助人们切



实实现数据科学的目标。一般来说，研究人员在进行数据分析时，要对数据的产生机制做某些基本假设，例如假设数据都是由某类相应的模型产生的。因此，数据分析的基本问题是找出这个模型，如表1所示。由于数据采集收集过程中不可避免会引入噪声，因此这些模型常常都是随机模型。

表1 常用数据类型及相应的数据模型

数据类型	相应的数据模型
点集	概率分布
时间序列	随机过程
图像	随机场
网络数据	图模型、贝叶斯模型
模糊数据	模糊概率分布

美国雪城大学 (Syracuse University) 信息学院的杰弗里·斯坦顿 (Jeffrey Stanton) 在《数据科学导论》(An Introduction to Data Science) 一书中提出：数据科学是一个新兴领域，它包括对海量数据信息的收集、预处理、分析、可视化、管理及保存等工作。

许多企业和机构之所以收集数据，原因是数据带来的价值或者潜在的价值超出了收集数据和管理数据的成本。通常，特定行业或领域的数据中会蕴含着规律，即数据中包含价值。

## 2. 数据科学的属性和应用过程

从数据科学的交叉特性来看，可将数据科学看成计算机科学、数学和统计学 (含有现代数据分析)、特定领域的实务知识交叉融合而形成的新兴学科。相关的理论和技术源于不同学科的研究方向，相互之间存在比较大的差异，比如研究的基本假设等，而数据科学则试图在此基础上构建和谐自洽的理论体系。

德鲁·康韦 (Drew Conway) 利用韦氏图说明研究数据科学需要的技能，如图1所示，进而揭示出数据科学十分重要的属性——交叉性。可以看出，数据科学的特征之一是多学科的交叉。

从实践应用来看，数据科学研究并不是某一个领域的专项问题，而是多学科综合性研究问题，应用时要具

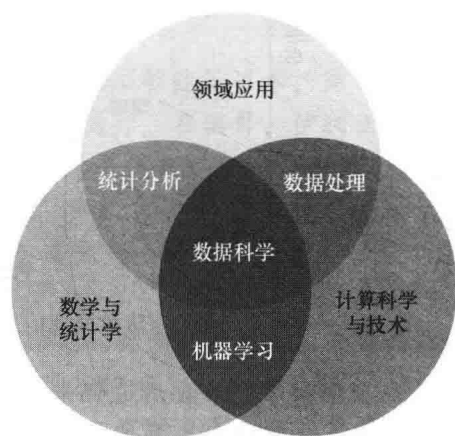


图1 数据科学的多学科交叉属性

备三个条件：第一个条件是底层构架开发或使用能力；第二个条件是程序开发能力；第三个条件是数据建模和解决问题能力。

数据科学家需要什么样的能力呢？美国加州大学伯克利分校统计系的郁彬（Bin Yu）教授提出，一个合格的数据科学家应具备的基本素质和技能可概括为 SDC<sup>3</sup>：

- Statistics (S)：统计学。
- Domain (Science) Knowledge (D)：坚实深厚的科学知识。
- Computing (C)：计算技术。
- Collaboration (Team Work) (C)：团队的合作能力。
- Communication (to Outsiders) (C)：与外界的沟通能力。

从现实来看，数据科学家收集数据、清洗数据、创建数据集、分析数据然后提出新观点，尝试用现有的数据预测未来，帮助提高产品、服务的质量和顾客黏性。更好的质量意味着更能取悦顾客、获得收益。因此，数据科学家最应该具备的三个特质是：①知道如何提出好问题；②理解所要处理数据的结构；③很好地解读这些数据。换句话说，数据科学家必须能够完成从数据到信息、知识、洞察力的过程，如图 2 所示。概括地讲，数据科学家就是提出合适的问题，然后提出有意义的见解来指导决策。

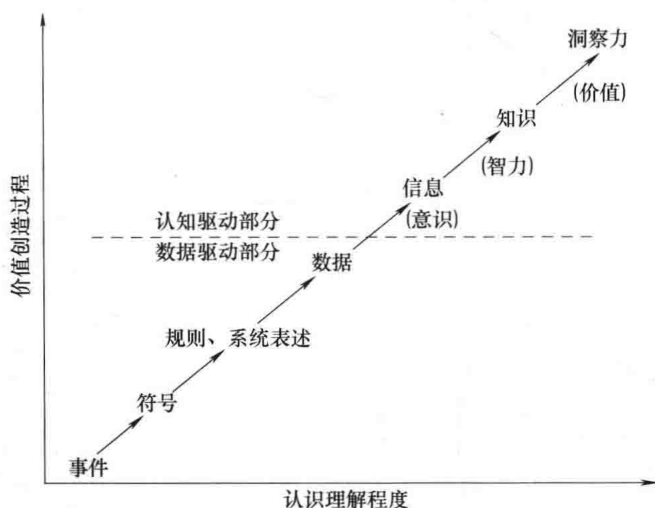


图 2 从数据到信息、知识、洞察力的过程

从数据科学整个应用过程来看，不同阶段工作拥有各自的特定目标和探索工具，由不同的人来完成。例如数据获取与管理由首席数据官负责执行；数据存储与处理由数据构架师负责完成；而数据分析与建模由数据分

析师负责执行；当结合到某个特定领域时，特别是试图发现商业价值和启发并进一步应用时，则由商业数据分析师来执行完成。如图3所示。

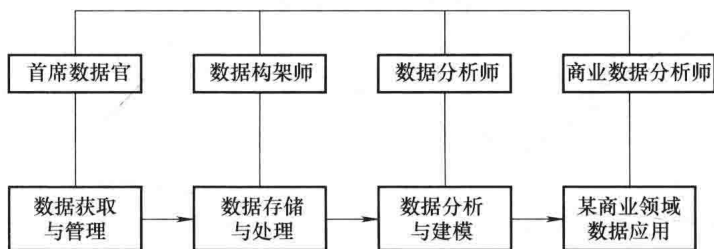


图3 数据应用中各阶段工作的目标任务和相应负责人

对于商业数据分析师和数据分析师来说，他们的工作尽管可能有交叉重叠之处，但是侧重点是不一样的。本质上，商业数据分析师探寻数据中隐藏的商业规律，在商业和IT之间作为中间沟通人，提高管理者的洞察力，进而改进业务流程，提高公司绩效。

### 3. 数据科学专业与商业数据分析专业

大数据时代，数据成为最重要的资源。2017年，中国网民规模达到7.72亿人，互联网普及率为55.8%，手机网民规模达到7.53亿人。2016年12月，工业和信息化部发布了《大数据产业发展规划（2016—2020年）》，这是为贯彻落实《中华人民共和国国民经济和社会发展第十三个五年规划纲要》和《促进大数据发展行动纲要》，加快实施国家大数据战略，推动大数据产业健康快速发展，在行动纲领上提出的指导规划。中国信息通信研究院发布的《中国大数据发展调查报告（2017年）》显示，2016年中国大数据市场规模达168亿元，增速达到45%，预计2017—2020年仍将保持30%以上的增长。

大数据的发展离不开物联网、云计算、人工智能等技术，离不开大数据相关人才。人才是产业发展的重要基石和支撑。在国外，伊利诺伊大学厄巴纳-香槟分校从2011年起举办“数据科学暑期研修班”；哥伦比亚大学从2013年起开设“应用数据科学”课程，并从2013年起开设相关培训项目，从2014年起设立硕士学位，2015年设立博士学位；纽约大学从2013年秋季起设立“数据科学”硕士学位。

加州大学伯克利分校在2016年春季学期的时候，学校首次开设的数据科学基础课程只吸引了300多名学生，而到2017年，已经有1000多名学生注册这门课程。学校一位数据科学教授表示：“在伯克利，从来没有哪门课程发展得如此之快。”

在中国，香港中文大学从2008年起设立“数据科学商业统计”硕士学位；复旦大学从2007年起开设数据科学讨论班，2010年开始招收数据科学博士研究生，并从2013年起开设“数据科学”课程；北京航空航天大学于2012年设立大数据工程硕士学位。

2017年4月，教育部公布高校新增专业——“数据科学与大数据技术”本科专业，有32所高等院校成为第二批成功申请“数据科学与大数据技术”本科新专业的高校。这样，加上此前第一批成功申请该专业的北京大学、对外经济贸易大学和中南大学，已有35所高校成功申请“数据科学与大数据技术”专业。

为了深刻认识大数据、人工智能等新技术对商业运营管理和商业模式所带来的冲击，及时做出应对策略，国内外一些知名大学的管理学院或商学院纷纷创建商业数据分析（Business Analytics，又称为商业分析）硕士项目或专业。比如，美国南加州大学的马歇尔商学院开设了商业数据分析的硕士项目，北京大学光华管理学院设立了商业分析硕士项目等，哈尔滨工业大学经济与管理学院创建了商务数据分析、商务智能辅修专业（学位）。

#### 4. 关于 Business Analytics 的译法

如果针对 Analytics 以一种学术界常见的追根溯源方式来探究，那么通常可将 Analytics 理解为成“n. 分析学，解析学，分析论”。

在大数据时代背景下，Analytics 是否被赋予了新的含义呢？当前，如果从学术角度来考察“Analytics”，一般是指 INFORMS 学术组织（www.informs.org）对其所赋予的新的含义：Analytics is defined as the scientific process of transforming data into insight for making better decisions。中文大意是：分析学被定义成将数据转变成用于更好地决策的洞察力的科学过程。可以发现，这个新定义中，Analytics 的对象是“数据”，对数据的加工过程则是一种科学的过程，最终目标是为人民的决策提供某种参考借鉴的依据，而且突出了“更好地决策”的高级目标。

国外有一本期刊《分析学》（Analytics），将 Analytics 定义为：利用数学、运筹学、统计影响商业决策。

而对于“数据分析”，我们特别欣赏和赞同麦克·沃森（Michael Watson）和德里克·内尔森（Derek Nelson）的观点。他们对“数据分析”所给出的定义为：“数据分析是利用数据获得洞察力，帮助人们更好地做决策的学科集合。它由下面三个分支组成：①描述性数据分析，提供数据描述，报告数据并对数据可视化处理；②预测性数据分析，利用数据预测趋势，识别数据关系；③规范性（指导性）数据分析，根据所拥有的已知数据，

依据未来所希冀的方向，为执行一系列的最佳决策提供指导方针。换句话说，描述性数据分析的目的是对已经发生什么和正在发生什么，提供一种认识理解；预测性数据分析的目的是告诉人们，将要可能发生什么；而规范性数据分析的目的是告诉人们如何应对。”（源自《管理数据分析——原理、方法、工具及实践》，中译本，机械工业出版社2017年出版）

鉴于以上介绍的 Analytics 和“数据分析”的学术含义，将 Analytics 翻译成“数据分析”就非常顺理成章。查询国内最近几年以来已经引进翻译的涉及 Analytics 方面的书籍，比如中译本《数据分析竞争法——企业赢之道》（2009年出版）、《精益数据分析》（2015年出版），以及前不久我们自己的译著《管理数据分析——原理、方法、工具及实践》，也是这样翻译的。本书是一本介绍数据分析方法在管理领域应用的初中级读本，通俗易懂、简明扼要，特别适合管理人员学习阅读，因此，我们将 Business Analytics 译成“商业数据分析”就十分自然了。

## 5. 本书的翻译和出版

本书的翻译和出版，是为了介绍如何将商业数据分析的理论及方法应用于各个行业，是一本学习和研究商业数据分析的基础教材。

本书第一译者曾于2013年10月参加哈尔滨工业大学经济与管理学院的“商务数据分析”新专业的课程建设小组，从那时起就一直关注和积累商业数据分析方面的知识。

本书的翻译和出版要特别感谢哈尔滨工业大学经济与管理学院的李一军教授、胡运权教授、于渤教授，以及学院院长叶强教授多年来给予译者的教诲和关怀。本书翻译工作具体分工如下：哈尔滨工业大学经济与管理学院王忠玉翻译第3、4、6、7、8章和附录B、附录C、附录D；黑龙江大学王天元翻译第5章和附录E、附录F、附录G；绥化学院王伟翻译第1、2章和附录A。最后，由王忠玉统稿。译者还要感谢曾经参与本书翻译和出版工作的其他老师和同学，他们是：哈尔滨工业大学经济与管理学院金融专业硕士研究生温雅欣，上海财经大学硕士研究生黄慧凡，哈尔滨工业大学经济与管理学院学生徐明泽、赵桓、吴春媛、张宇昂、王初旭、邢喆、王欢、王晓楠、高旭彤、张天爱、钟世超、彭娟，以及中央财经大学学生张丽波等。另外，在后期统稿时，哈尔滨工业大学经济与管理学院的陈政同学参与了部分工作。

虽然译者精心钻研和翻译，但仍可能存在纰漏和错误，希望广大读者指正。联系方式：wangzhy@hit.edu.cn。

王忠玉、王天元、王 伟

## 作者介绍

马克·J. 施尼德詹斯 (Marc J. Schniederjans), 内布拉斯加大学林肯分校工商管理学院 C. Wheaton Battey 特聘教授, 在此之前曾经在其他三所大学任教。施尼德詹斯教授是决策科学研究所 (Decision Sciences Institute, DSI) 的研究员, 并在 2014—2015 年担任 DSI 主席。他曾经经营过自己的货车租赁业务。目前, 他是供应管理学会 (ISM)、生产和运营管理学会 (POMS) 和 DSI 的成员。

施尼德詹斯教授主要讲授“运营管理和科学”, 并赢得了许多教学奖, 是金钥匙 (Golden Key) 荣誉会员和 Alpha Kappa Psi 商业兄弟会的荣誉会员。他发表了 100 多篇杂志文章, 写作出版了 20 本书。他最近的一本书名为《重塑供应链生命周期》, 广泛地研究了运营管理和决策科学。他还在学术会议上提交了 100 多篇研究论文。

施尼德詹斯教授不仅在五个期刊编辑委员会任职, 包括《计算机与运筹学》《国际信息与决策科学期刊》《国际服务业信息系统学报》《运营管理》《生产和运营管理学》, 而且担任《运营管理研究》杂志的区域编辑, 以及《国际战略决策科学》杂志和《国际社会系统科学与管理评论杂志: 国际杂志》(韩国) 的副主编。除此之外, 施尼德詹斯教授还担任多个商业和政府机构的顾问和培训师。

达拉·G. 施尼德詹斯 (Dara G. Schniederjans), 罗德岛大学工商管理学院供应链管理助理教授, 讲授“供应商关系管理和运营管理”等课程, 曾在《决策支持系统》《运筹学研究会》《商业过程管理》等期刊上发表过多篇文章。她曾与人合作编写过两本教材和一本大众读物, 也曾参加编写有关定量统计方法的读物。她担任过《国际社会系统科学杂志》社会科学商业伦理专题的客座顾问, 目前还担任 DSI 的网站协调员。

克里斯多夫·M. 斯塔基 (Christopher M. Starkey), 康涅狄格大学的经济学学生, 在管理和生产及运营管理学会会议上提交过论文; 他讲授“微观经济学原理”课程, 也曾经讲授过“宏观经济学原理”课程。他目前的研究兴趣包括宏观经济和货币政策, 以及决策方法。

# 前 言

我们每天都要面对大爆炸式的信息，我们尽力汇总整理信息，并运用信息来帮助决策，但有时候还是会被大量的数据所淹没。这可能导致我们得出错误的结论，同时做出错误的决定。例如一家全球性公司，在从世界各地收集数百万笔交易和客户行为数据时，单独就数据量而言，也会使得查找有用的客户信息这一任务几乎不可能完成。对于这家公司，以及更小的企业来说，解决方案是应用商业数据分析（Business Analytics, BA）。商业数据分析可以帮助人们整理大数据文件（称为“大数据”），在预测未来时，寻找有用的行为模式，并据此配置资源以实现最优化决策。商业数据分析是一个逐步深入的过程，帮助企业收集有用信息，以系统化程序来管理大数据，进而解决问题，找准提高公司绩效的机会。

本书为商业数据分析的基础教学而编写，既可服务于大学教育，也可服务于实践者。除了为商业数据分析提供最新的文献和研究外，本书利用简单的术语及其在实践中十分有用的辅助方法，来解释商业数据分析流程。就整体而言，本书所阐述的统计和定量工具中需要的数学知识不超过高中代数的水平。为了服务于读者需求，我们给出大量实例和案例，以训练读者运用普通的商业数据分析工具和软件。从业者可以发现，对商业数据分析方法进行研究，有益于考察所探讨的专题。大学师生可以发现，每章学习目标和对问题的讨论有助于满足他们的需求。

本书的目标是阐明什么是商业数据分析，为什么商业数据分析十分重要，以及如何应用商业数据分析。为实现这个目标，本书将介绍概念内容、相关软件和一些分析工具。

## 概念内容

本书的前面8章阐述有关的概念内容（参看第1章的1.4节关于本书内容的组织结构）。本书的概念内容远远超出商业数据分析的范围。本书将要解释：对于提供问题的解决方案来说，为什么商业数据分析是十分重要的，如何用它来实现竞争优势，以及如何协调组织更好地利用它。本书解释了在组织中应用商业数据分析时需要用到的管理知识，还有商业数据分析人员应该拥有的技能。本书还阐述了数据管理问题，例如数据收集、业务外包、数据质量及变革管理，因为这些内容和商业数据分析有关。

在建立管理基础来解释商业数据分析是什么和为什么重要之后，剩下各章则关注如何运用商业数据分析。为使这三个过程具体化，我们将商业数据分析解释为具有描述性、预测性以及规范性的分析步骤。对于每一个步骤，本书提供了一系列策略，以及商业数据分析流程的最佳实践指南。

### 相关软件

商业数据分析的绝大部分工作都要使用软件。令人遗憾的是，没有一个软件可以涵盖商业数据分析的所有方面。虽然有很多相关软件，但许多机构各自偏爱其中的一个。为了提供灵活性，本书介绍了多个软件，以供选择。在本书中，我们利用 SPSS®、Excel®和 Lingo®软件进行建模并解决问题。尽管本书提供了使用这些软件系统时的一些输入和指令，但主要内容是其输出。至于对运行软件不感兴趣的人，软件输出的资料也为他们提供了有价值的信息。对于那些在大学其他课程中会教授使用方法的基本软件，本书不再详细培训，建议对商业数据分析感兴趣的人要先了解这些基本软件的使用。

### 分析工具

本书附录对分析工具的内容进行了概述。商业数据分析是综合了统计、管理信息系统（MIS）和定量方法的学科。虽然书中的概念内容概括了如何进行商业数据分析流程，但如何在实际中应用商业数据分析则需要掌握定量化工具。因为一些实践者和大学课程对商业数据分析的技术方面不太感兴趣，所以本书只好在附录中给出这些定量内容。这些附录提供了用于支持各种分析的商业数据分析工具的大量解释。本书所解释和阐明的统计工具包括统计计数（排列、组合、重复）、概率方法（加法准则、乘法准则、贝叶斯定理）、概率分布（二项分布、泊松分布、正态分布、指数分布）、置信区间、抽样方法、一元回归与多元回归、绘图方法与假设检验。尽管管理信息系统已超出本书的范围，但本书仍然利用前面所提到的软件运用来阐明管理信息系统的搜索、聚类 and 典型的数据挖掘方法的应用。此外，本书中所解释和阐明的定量方法工具包括：线性规划、对偶性和灵敏度分析、整数规划、0-1 规划、预测建模、非线性最优化、模拟分析、盈亏平衡分析和决策理论（确定型、风险、不确定型决策、期望机会损失分析，完全信息期望值，不完全信息的价值）。

我们要感谢许多人的帮助，这些人为创作本书提供了所需的支持。首先，我们非常感谢我们的编辑 Jeanne Glasser Levine 和 Pearson 的优秀工作人员的支持。他们使得创作本书变成一种快乐，并与我们一起改进最终产品。和出版商合作编写书籍的几十年使我们认识到，顶级的出版商在图书出版



中发挥了非常大的作用。我们感谢 Alan McHugh，本书的封面（英文版）是他所设计的。他的不断探索精神和创新思想，对我们完成书稿做出了重要贡献。我们还要感谢 Jill Schniederjans 给予的大量编辑帮助。她减少了冗长的话语，增强了内容的趣味性。最后，我们要感谢 Miles Starkey 的帮助，他的存在和魅力提升了我们的精神状态，使我们保持在正确的方向上并按期完成书稿。

尽管许多人为本书写作提供了帮助，但其准确性和完整性由我们负责。对于本书可能存在的所有错误，我们事先表示歉意。

马克·J. 施尼德詹斯 (Marc J. Schniederjans)

达拉·G. 施尼德詹斯 (Dara G. Schniederjans)

克里斯多夫·M. 斯塔基 (Christopher M. Starkey)