

刘红阁 王淑娟 温融冰◎著

人人都是数据分析师

Tableau应用实战

第2版



拖放成图



探寻统计规律



便捷数据管理



内置智能推荐



海量数据读取



交互式地图



中国工信出版集团



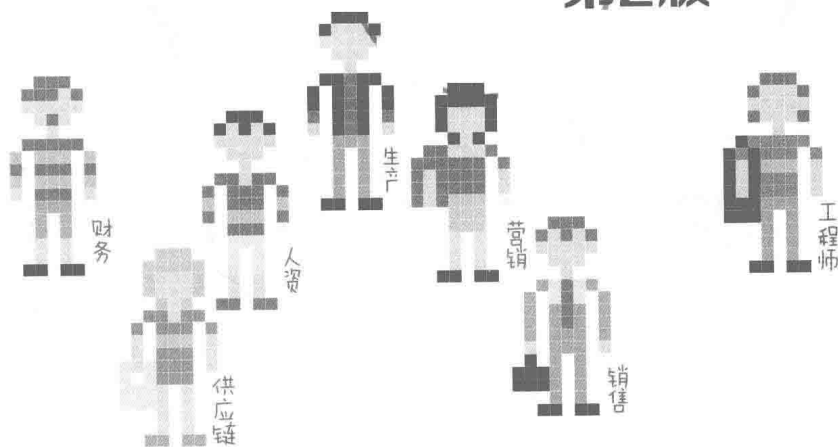
人民邮电出版社
POSTS & TELECOM PRESS

刘红阁 王淑娟 温融冰◎著

人人都是数据分析师

Tableau应用实战

第2版



人民邮电出版社

北京

图书在版编目 (C I P) 数据

人人都是数据分析师 : Tableau应用实战 / 刘红阁, 王淑娟, 温融冰著. — 2版. — 北京 : 人民邮电出版社, 2019. 1

(图灵原创)

ISBN 978-7-115-50119-6

I. ①人… II. ①刘… ②王… ③温… III. ①可视化软件—数据分析 IV. ①TP317.3

中国版本图书馆CIP数据核字(2018)第265520号

内 容 提 要

本书基于 Tableau 10.5 最新版本编写, 详细介绍了 Tableau 的数据连接与编辑、图形编辑与展示功能, 包括数据连接与管理、基础与高级图形分析、高级数据操作、基础统计分析、如何与 R 集成进行高级分析、分析图表整合以及分析成果共享等主要内容。同时, 书中以丰富的实际案例贯穿始终, 对各类方法、技术进行了详细说明, 方便读者快速掌握数据分析方法。

本书适用于互联网、银行证券、咨询审计、快消品、能源等行业数据分析用户以及媒体、网站等数据可视化用户。

-
- ◆ 著 刘红阁 王淑娟 温融冰
责任编辑 张 霞
责任印制 周昇亮
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
涿州市京南印刷厂印刷
 - ◆ 开本: 800×1000 1/16
印张: 19 彩插: 4
字数: 502千字 2019年1月第2版
印数: 16 801—20 800册 2019年1月河北第1次印刷
-

定价: 79.00元

读者服务热线: (010)51095186转600 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147 号

站在巨人的肩上
Standing on Shoulders of Giants



iTuring.cn

站在巨人的肩上
Standing on Shoulders of Giants



iTuring.cn



图 2-5 维度与度量转换

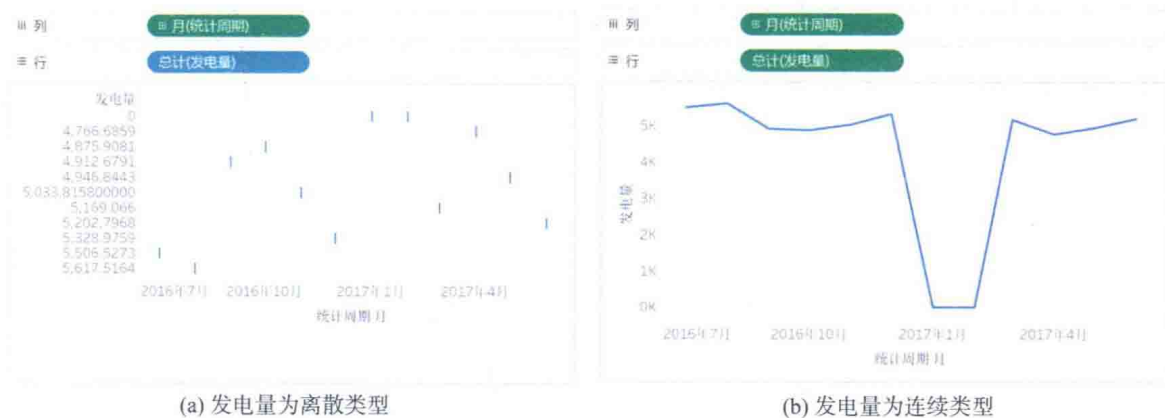


图 2-6 离散和连续类型

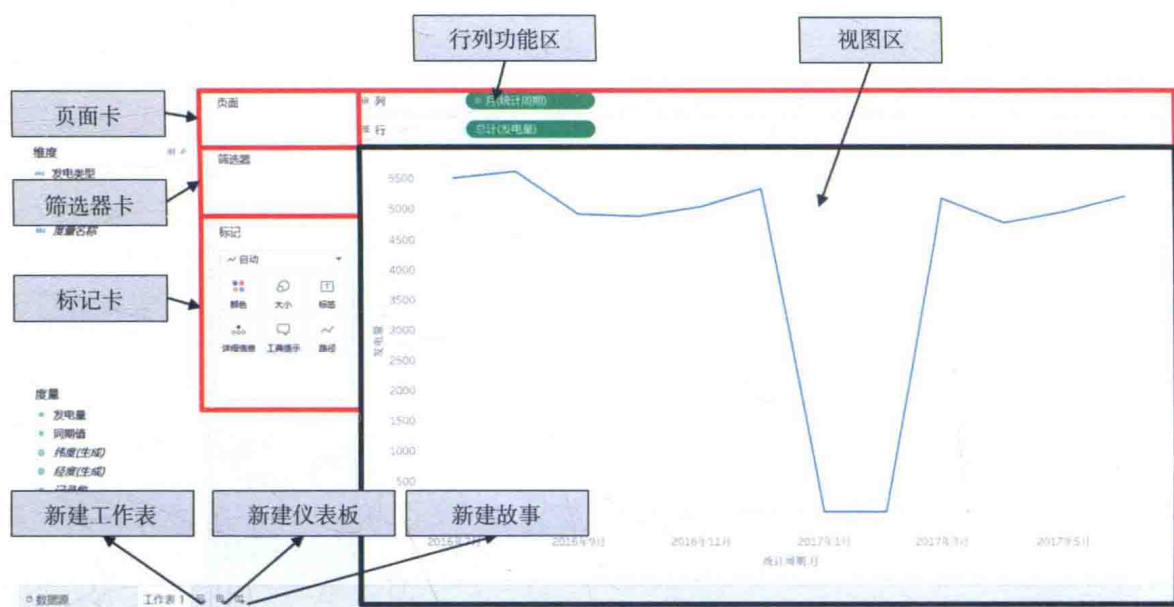


图 2-8 认识视图工作区



图 2-16 颜色图例

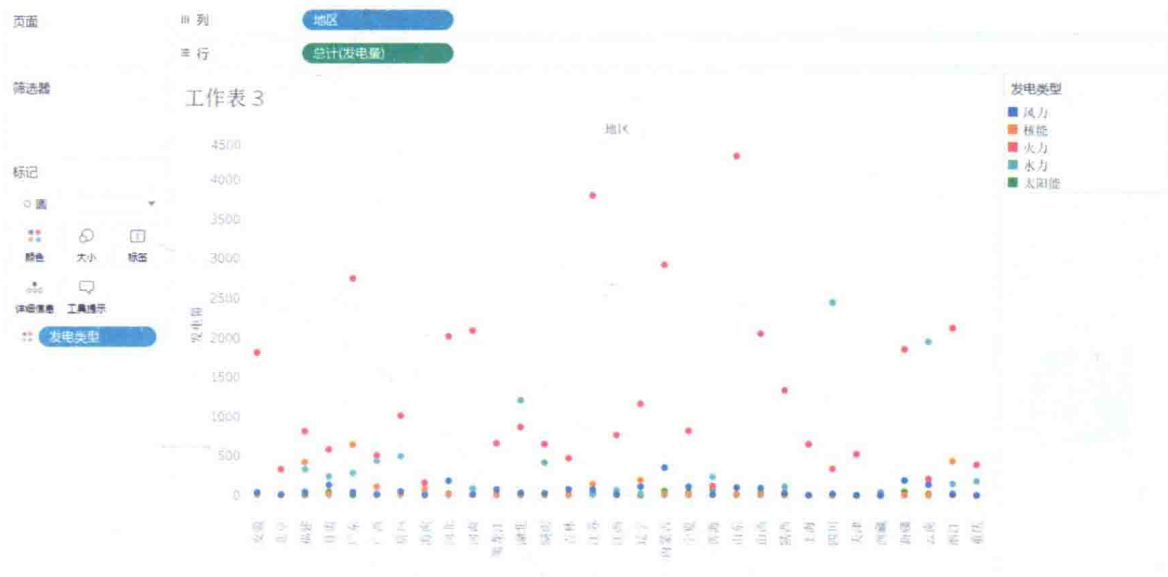


图 2-23 依据颜色分解细化的视图



图 3-46 以 C 指标为主数据源构建视图



图 3-47 构建数据融合

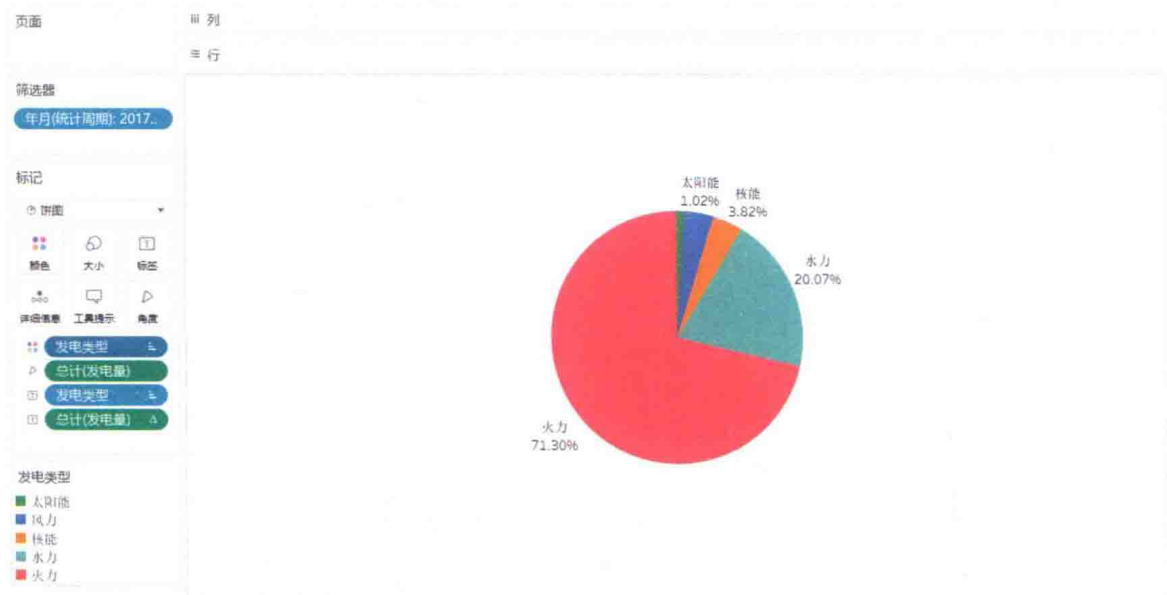


图 4-18 添加标签及排序后的饼图



图 4-33 压力图——2017 上半年部分地区累计发电量与同比变化情况

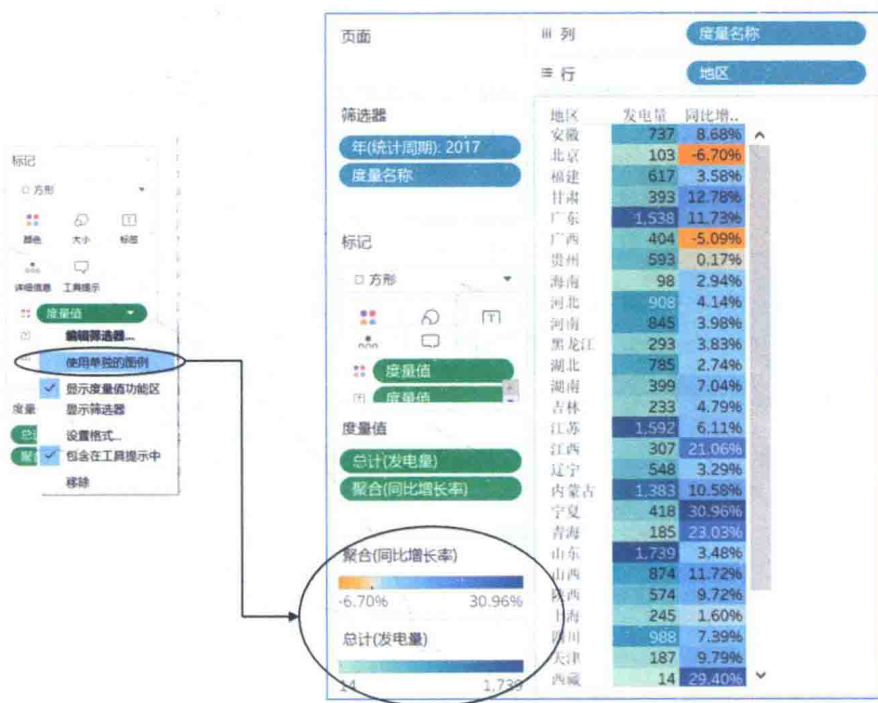


图 4-38 突显表——2017 年上半年不同地区累计发电量及同比增长情况

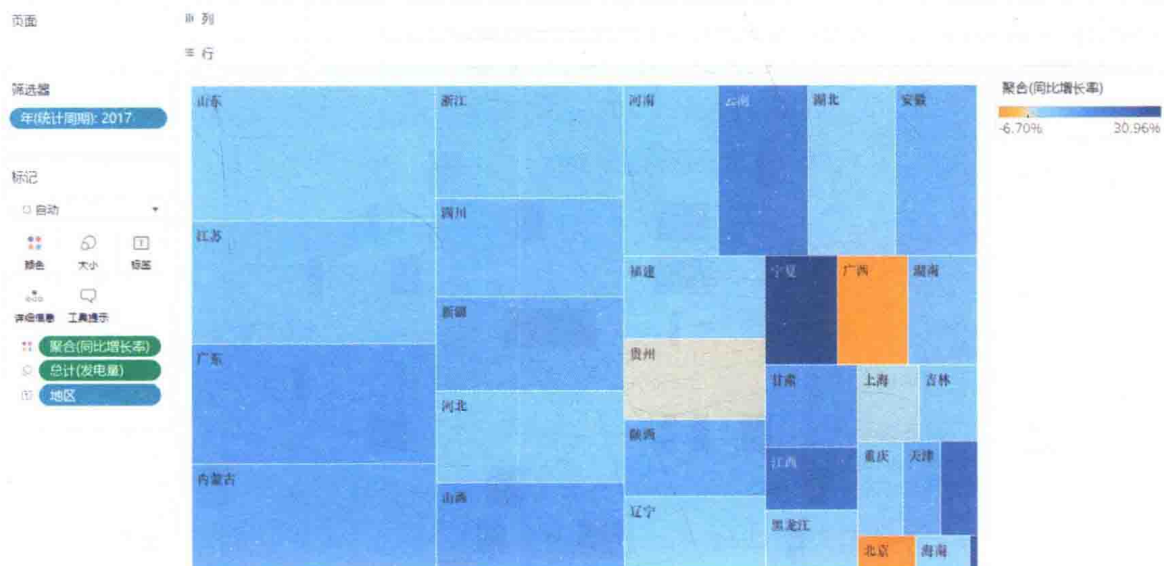


图 4-39 树地图——2017 年上半年不同地区累计发电量及同比增长情况

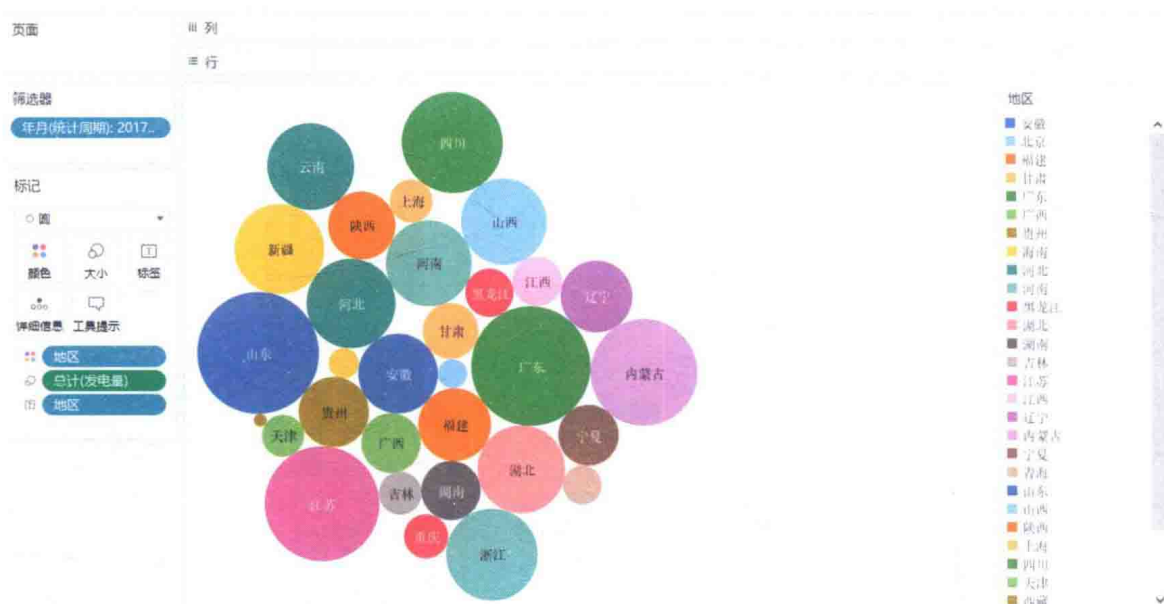


图 4-40 填充气泡图——2017 年 6 月不同地区发电量情况



图 4-49 编辑“延期天数”的颜色

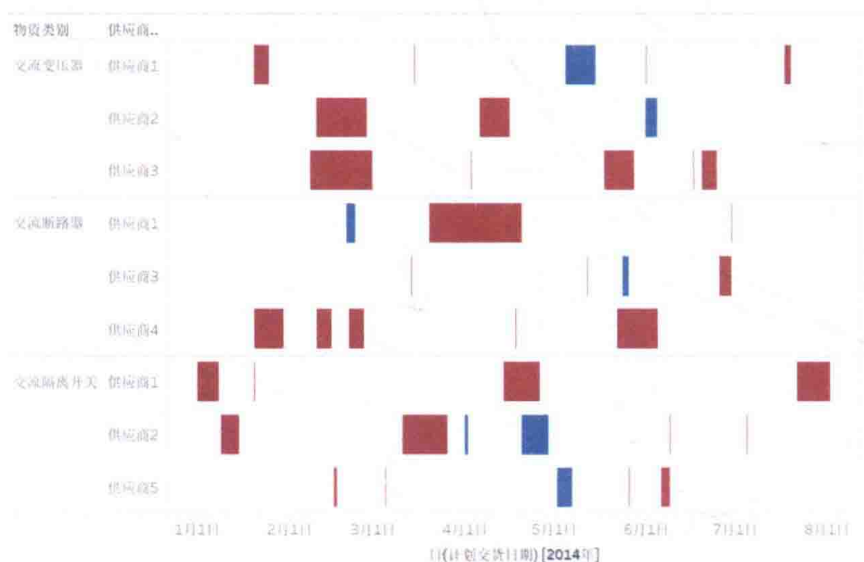


图 4-50 供应商及时供货情况分析

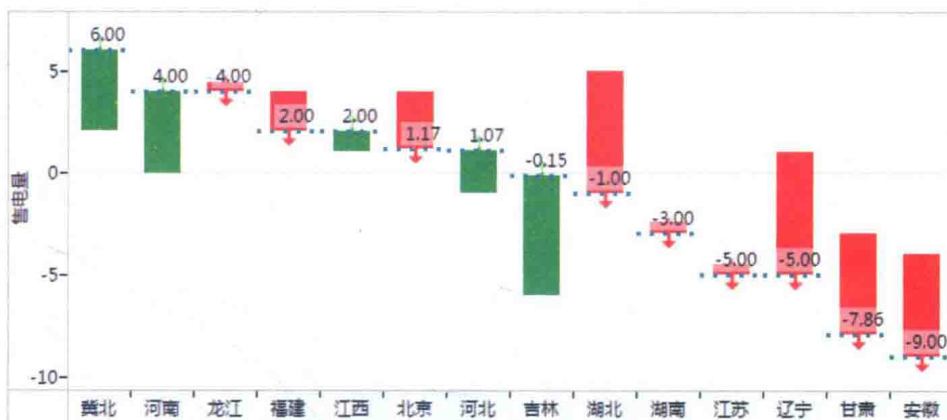


图 6-24 “变化排序”图

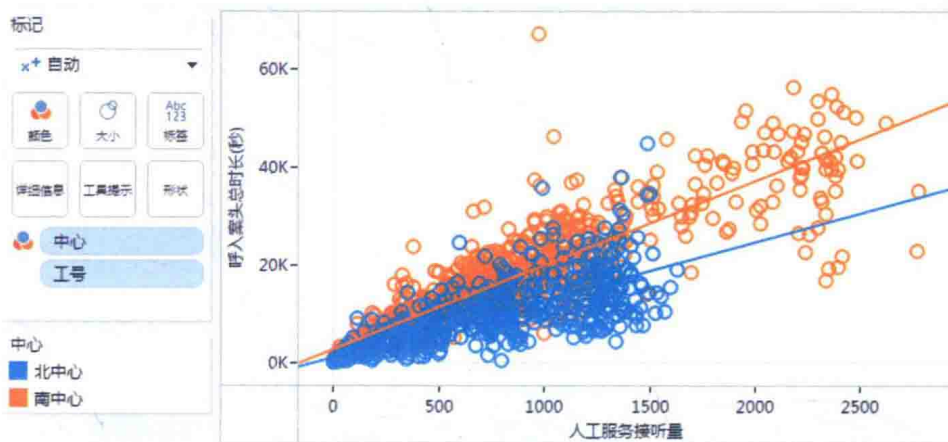


图 7-11 按颜色绘制趋势线

Tableau 用户评价

“Tableau 能够真正地帮助我们快速发展和创造价值。这种帮助无时无刻不体现在生产中。”

——Will Bishop, 特斯拉汽车公司高级测试工程师

“Tableau 帮助我们以更少的资源清理和整塑用于报告的数据。因为实现了报告自动化，所以我们这个小小的团队可以将大量资源集中在将数据分析转变成见解上。”

——Secil Watson, 富国银行集团执行副总裁

“我们完全钟情于 Tableau，钟情于它的承诺：让组织中的每个人都能使用现成、可操作的数据，而不仅限于分析师和高级用户。”

——David Baudrez, 思科公司欧洲、中东和非洲地区商业洞察主管

“如果只是一次又一次地重复同一件事情，那么我永远无法跳出固有的思维模式。Tableau 让我们以一种完全不同的方式来查看数据。”

——Kevin King, 可口可乐装瓶公司报告与分析主管

“使用 Tableau 之后，去年我们的收入增加了 2 亿美元，增幅达到了 2%。Tableau 很实用，使用也很简单，而且比其他商业智能工具反应快。要想在航空业或者运输业赚到钱，就得用 Tableau。”

——James Pu, 中国东方航空公司网络与营收部高级主管

“如果没有 Tableau 系统，我们现在可能还在使用电子表格软件，那样就需要委派专人进行数据分析，不仅偶尔的人为错误难以避免，而且当分析员没有上班时，我们可能就无法获得当天的报告。”

——Leon Zhang, 汉堡王（中国）信息技术部副总监

序 一

AI（Artificial Intelligence，人工智能）是第四次工业革命的核心驱动力，促进着各行各业的变革。多年来，百度在 AI 领域持续投入，成效显著，目前已构建起对内赋能重要业务的完整 AI 技术布局，并通过 AI 开放平台对外开放包含无人驾驶、语音、图像、视频、增强现实、自然语言处理等在内的多项 AI 核心能力，为汽车、零售、教育、金融和政务等行业提供解决方案和产品。

在坚定不移地推进“夯实移动基础，决胜 AI 时代”的百度企业战略落地进程中，我们借助 AI 技术，在“搜索+信息流”双引擎上，大商业体系不断创造智能投放、智能创意、智能客服和跨屏全场景的新营销方式，构建线上线下打通的 AI 新营销生态闭环，为客户赋能，建设商业内容生态，实现对用户需求的精准满足。

在人工智能时代，数据是核心生产力，所有行业都会从数据的采集、分析、推理、预测及控制中受益，数据产生的价值将会给每个人带来更智慧的生活。而通往智慧未来的另一扇门，就是直接将数据和简单易用的分析工具提供给千千万万的普通人。

机器智能擅长处理的是过去曾发生、重复性、具有模式（pattern）的场景，而在 VUCA 时代，我们遇到的更多是不确定、易变、复杂、模糊的问题，这些新问题需要每个个体整合手边数据，运用人类智慧，分析、判断、决策以及跟进反馈。而且这种问题往往不能一劳永逸，而是呈现出序列式、演进式的势态，需要人们循序渐进地寻找解决方案。以 Tableau 为代表的敏捷性分析工具的价值正在于此。

自助式分析，交互式操作，数据在用户的简单拖放中转化为各个领域的决策智慧。麦克卢汉曾说：媒介即信息。新颖的媒介本身就蕴藏着更多的信息。从这种角度看，Tableau 是典型代表，它将数据分析技术民主化，传播、引入到你我之中，形成群体智慧，与作为客体的机器智能一起，在未来和谐共处。

刘红阁博士是大商业体系商业分析团队的负责人，具有多年丰富的数据分析领域实践经验，带领团队同学不断从点滴入手，持续提升百度大商业体系的数据化运营和科学决策水平，同时也一直关注和研究数据分析工具与技术领域的前沿与创新。唯是这般努力，才有《人人都是数据分析师：Tableau 应用实战》新版问世。我为她和其他两位作者的付出感到欣慰，也希望这本书能够切实地帮助到广大非技术出身的商业分析师。人人勤动手，皆可做分析。

郑子斌

百度副总裁、百度搜索公司 CTO

序 二

古谚有云“一图胜万言”，这在不同文化中均广为人知。有人经考证认为，其最早来自记者弗莱德·巴纳德（Fred R. Barnard）1927年发表在一本贸易杂志上的文章；然而另一些人认为它源于中国的孔子，这位2000多年前的智者曾讲过“书意能达万言”。无论出处如何，专家们已经一次又一次地发布研究成果来证实这个结论：人类主要是视觉动物。当我写这篇序言时，我仍惊叹于这种认知在中国是多么地深入人心，从远古时代雄伟的绘画到象形文字、形声字，再到今天的简体字，都有着广泛的体现。

而 Tableau 有幸基于这样的智慧创造应用产品，来帮助更多的人读懂他们的数据。我们十多年前起源于斯坦福大学的这个愿望，已经惠及150多个国家3万多家机构及成千上万的个人。当今世界，人类前所未有地连接在一起，而未来5年内，将会产生多如繁星的数据，那么能够理解、利用和掌握自身的数据将变成一项最基本的生存技能。

许多机构已经开始面对这个人人自助分析、数据驱动决策的现实。少数比较成功，但更多疲于追赶。Tableau 亚太区的一项市场调研发现，那些能够满足业务人员自助开展数据读取和分析需求的企业，其业绩比采取传统做法的企业好两倍。但是尽管知晓此好处，75%的业务人员表明他们自助开展数据分析的需求仍未充分满足或完全未满足。

中国经济正在突飞猛进地发展，在知识与数字化时代，越来越多的国有企业、私营企业甚至公共服务机构都面临着同样的挑战。培训员工应用自助式的分析工具发现海量数据中的价值已变得势在必行。

而这正是本书的关键所在！尽管市面上已经有了传统的商业智能解决方案，但是面对海量的数据和整合这些解决方案的复杂性，很多中国企业仍疲于应对。本书通过真实的案例，阐述了一个完全不同于以往的数据分析方法论。它展示了领先企业如何让商业智能不再局限于少数技术人员，让多数人都掌握自助分析，读懂数据，创造更大的价值。

我非常感谢本书的3位作者，他们持续地为客户带来行业洞见，并先见地倡导客户应用 Tableau 来满足数据分析的需求。

JY Pook
Tableau 亚太区副总裁

序 三

随着数字化技术的广泛应用和信息化水平的持续提升，我们正在逐步迈入大数据时代。全球众多国家已经充分认识到这一发展趋势，纷纷将大数据上升为国家战略。在企业层面，数据已成为一种新型的重要战略资产，越来越多的企业开始强化数据资产管理，建立专门的组织或明确相关的责任部门，大力拓展数据分析与应用。数据分析已经成为一门显学，也正在发展成为一项独立的业务，发挥着越来越重要的作用。

在埃森哲为客户提供的多种咨询服务中，数据分析咨询数量自 10 年前开始逐渐变多。在国内，随着诸多企业信息系统建设的不断发展，近几年利用数据、信息进行业务监测管控与管理优化的需求日益显现。然而由于对数据分析这一新鲜事物认识不足，经常出现这样的情况：尽管在服务器、数据库、分析软件与技术工具的投资颇多，但却难以取得实用的效果。最为明显的认识偏差是没有将数据分析视为一项业务能力，有的人把数据分析等同于报表、报告等结果性的东西，另外有些人又把它视为一些分析技术、工具与 BI 系统。实际上，数据分析有自己独特的策略、组织、流程、技术组成与要求，并且各要素之间需要相互支撑与配合，才能持续产生满足需要的信息与见解。不止如此，数据分析作为一项业务，具有明显不同于一般流程性业务的显著特征：频繁变化。随着企业战略发展、管理焦点与数据支撑度的变化调整，数据分析与使用需求也要不断变化才能满足需要，传统流程性业务的很多做法在数据分析业务上就不能一成不变。一言以蔽之，数据分析是一种随需而变的新型业务。

数据分析业务变化多，频度高，处理数据量日益庞大，这给企业提出了严峻的挑战，尤其是在数据处理与分析支撑技术层面。我们很多客户在这方面决策有误，采用了不那么有效的技术路线，引入或建设了对数据分析工作支撑度有限的技术与系统，最终限制了数据分析能力与作用结果。在给客户提供数据分析咨询的过程中，埃森哲先后使用过 SAS、SPSS、BI、BO、Excel 等多种复杂程度各异的工具与系统。基于我们的使用体会与客户反馈，从以最小化的投入来最大化满足当前阶段企业数据分析需求的角度看，Tableau 不失为一个值得推荐的选择。

Tableau 既可以作为一个加强版的类 Excel 工具，供数据分析人员对数据进行交互式、可视化和挖掘，提高分析工作效率，给出有意义的分析结论；又可以将有意义的分析路径、成果进行发布，共享给其他关联或管理人员进行浏览、查询与使用，实现分析人员与业务人员的协同。在数据分析领域，Tableau 在国外领先公司中得到了广泛应用，并且在国内的市场规模也日益壮大。当然，数据分析涉及的环节与任务众多，包含数据获取、数据存储、数据传输、数据操作、数据分析、数据展示等，要求也各不相同。Tableau 并不是万能药，在有些数据处理工作上，还

需要与其他工具和技术结合，才能够实现最大的成效。

本书 3 位作者基于自身的项目经验与使用体会，结合实际的应用案例，总结了 Tableau 的功能特点与使用技巧，写作了本书，并且后续还会继续完善。对于希望或正在建设数据分析能力、开展数据分析业务的企业、组织或个人，本书可以提供不一样的选择与体验，拓宽对数据分析业务和工具的认识，帮助感兴趣的使用者更好地应用 Tableau 工具开展工作。

王靖

埃森哲大中华区董事总经理