



数据可视化分析

Tableau原理与实践

喜乐君 著

作者简介

喜乐君

山东大学政治学本科、教育学硕士，Tableau Desktop QA、Server QA官方技术认证，Tableau官方合作伙伴。多年央企、民企、创业经历，历经人资行政经理、董事长助理、业务经理、项目经理等职，热爱读书，特别是技术、心理学，用心写好一本书，全心全意为客户服务。

“喜乐君”(yupengwu.com)博客作者。

公众号：



[Tableau传道士]

查看本书续篇、最新博客文章，问题交流及企业培训咨询。



“腾讯课堂”视频课

持续更新的Tableau随书视频及高级视频。

版权信息

COPYRIGHT

书名：数据可视化分析：Tableau原理与实践

作者：喜乐君

出版社：电子工业出版社

出版时间：2020年7月

ISBN：9787121391293

字数：690千字

版权方：电子工业出版社有限公司

版权所有·侵权必究

内容简介

本书系统地讲解了Tableau Prep Builder和Tableau Desktop的原理与实践应用。全书以可视化分析、Tableau计算为重点，详细介绍了如何理解数据的层次、如何使用Tableau Prep Builder整理和准备数据、如何使用Tableau Desktop开展敏捷数据分析、Tableau高级互动，特别是深入介绍了Tableau的各种计算，从而以有限的的数据实现无限的业务场景分析。

全书贯穿数据与问题的层次分析方法，并用实例加以说明，不仅适合希望系统学习Tableau的初学者，而且适合Tableau的中高级分析师。

推荐序一

英国数学家和企业家、消费者洞察公司Starcount的首席数据科学家Clive Humby曾经说过：“数据是新的石油！”但是如果数据不能被提炼，就不能被真正利用。就像石油需要被转换成燃料、化工制品等才可以变成有价值的物品驱动盈利的行为，数据必须被整合、分析后才会产生价值。

越来越多的企业领导者认识到数据的价值，并且付出行动，开始向“数据驱动型企业”转型。根据咨询公司麦肯锡的调查，在数据转型的道路上，92%的企业面临挫折，只有8%的企业取得了成功。市场咨询公司IDC和Gartner的研究指出，在推动数据驱动型企业的转型道路上，企业员工的数据素养是企业转型的关键因素之一；所谓的数据素养是指人们读取、处理分析和讨论数据的能力。Gartner预测，数据素养是21世纪人类需要的最重要的技能之一。自2003年诞生以来，Tableau一直致力于“帮助人们看到和理解数据”这一伟大使命，很多数据分析爱好者、Tableau的粉丝也在默默传播数据分析文化。

在国内，有很多学术机构和商业机构编写了不少Tableau主题的书籍，介绍如何使用Tableau进行数据分析。最近，我拜读了“Tableau传道士”喜乐君（吴玉鹏）的《数据可视化分析：Tableau原理与实践》初稿，马上给人眼前一亮的感觉。喜乐君以非技术人员的背景，将数据可视化的原理、过程和方法娓娓道来，以浅显的道理将读者带入了可视化分析的领域，进而用Tableau将这些理论方法变成实践。

在国内众多的图书中，本书创造了众多“第一次”，具有鲜明的特征：

- 国内第一本用DIKW模型解释可视化分析原理的书——在商业用户中，很多人问Tableau和其他厂商有什么区别？当你了解DIKW

模型后，你自然会理解为什么Tableau可视化分析与其他厂商产品有本质区别。

- 国内第一本以可视化心理学研究的成果解释了什么是可视化的前意识属性（Preattentive Attribute）的书，并利用可视化最佳实践将这些元素用于可视化分析。
- 国内第一本详细阐述了Tableau可视化分析全流程的书，并一一讲解可视化分析的平台，真正反映了Tableau可视化数据分析平台的精髓。
- 国内第一本系统介绍Tableau可视化数据准备产品——Tableau Prep Builder的书，方便用户进行数据准备。
- 数据是为了回答业务问题的，本书以业务实例的方式展示了技术人员和业务人员看待数据的不同方式，按照业务人员的分析思维用Tableau将数据分析层层展示，探索数据的价值。
- 本书也展示了大量Tableau使用中的最佳实践和业务模型，方便读者快速应用到实际业务场景中。

感谢喜乐君倾心编写本书，无论你是数据分析的老兵还是数据分析的小白，这本书都会给你提供非常有价值的信息。作为帮助企业推广数据分析文化的同道中人，感谢喜乐君将自己的实践心得整理成书，奉献给读者！感谢喜乐君在践行Tableau的使命——Help People See and Understand Data！祝愿更多的读者通过本书进入数据分析领域，帮助企业成为数据驱动型企业，成为企业数据分析文化的布道者！

Tableau大中华区副总裁 张磊

2020年3月21日于新加坡

推荐序二

欣闻喜乐君要写一本关于Tableau学习方面的书，我很是期待。

我在2007年真正深入接触数据分析，除了Office里面的Excel和Access，公司正在逐步推行BI，在早些年接触过SAP BW、SAP BO等建模和展示工具，直到2014年7月公司升级BI系统，引入Tableau，从此被这款绚丽的可视化工具所吸引。Tableau这些年在国内发展迅猛，版本不断迭代，但我们对它的认识和应用基本还是停留在最初的作图展示阶段。

在Tableau官方的推荐下，2019年夏天，我请喜乐君为我的团队提供了3天的Tableau深层应用培训，这使得我们第一次比较系统地学习了Tableau的理论基础、数据处理以及深层次应用。也正因为这3天的培训，使我对作者对Tableau的掌握和理解程度感到由衷赞叹。因为他并非技术人员出身，而是一个典型的应用人员，通过自己的实践、摸索和自学达到了专家级的水平。

数据分析的难点在哪里？在理解业务的前提下，无非是复杂的数据处理、错综的逻辑表达以及如何直观地展示，而Tableau具备了解决以上问题的所有特点。不过，市面上很多Tableau相关书籍或一些视频课程，都偏于功能介绍，却不能系统地从原理和应用两个方面阐述。

我翻阅了这本书的样稿，除去前面的理论介绍，大致分为3个方面：数据基础处理、数据可视化和高级应用。作者都结合自己的实践经验将其故事化、场景化，通过优雅的文笔、哲学的思维在纸上娓娓道来。通过阅读，你会发现技术并不枯燥，数字亦很有趣。作为由业务人员成长起来的数据分析师，本书作者更加了解业务需求的痛点、使用的难点以及Tableau的核心优点。

我也相信读者通过阅读学习此书，不仅能绘制适合应用场景的图

表，更能学会和理解其原理和方法；还能养成数据思维，进而能在自己的领域达到一定的高度。

作者是一位有深厚修为、深谙哲学且学识渊博的人，本书虽为侧重应用的工具书，但也不乏是一本介绍数据思维培养的难得读本！每一个章节不仅有专业的统计学术语，也有对数据处理原理通俗的解释，这些论述内容都融入了作者哲理性的思维和个人的虔诚。

这是一本可以面向初级、中级和高级各个阶段应用者学习的书籍，我相信每位读者都会有不同的收获！

石家庄以岭药业股份有限公司，营销效率部经理 宋洪涛

2020年4月3日

推荐序三

十年前，作为供应链管理咨询顾问团队成员，我为外企、国企、民企等各类型企业提供数据方案规划时，感受特别深的就是各层级领导和部门关注的KPI差异很大。当我们试图提供一套通用的KPI报告时，公司的各级管理层、经营层和执行层，都会提出层出不穷的新需求。那时我就想，如果能有一个工具，让各层级的人员结合自己的业务经验，确定分析视角，自行完成数据提取、分析展现和预测趋势，那么将会使行业管理变得更轻松。

五年前，我开始深入零售数据分析领域，我和团队列出了25个影响门店销量和利润的关键点，借助数据分析精确指导门店铺货、调货，不仅完全控制了货物损失，更是实实在在地将销售额增长了65%以上，净利润增长自不必说。不过，那时仅靠团队成员从各种报表中手动提取数据，再用Excel分析，不仅工作量巨大，而且需要每天反复地进行数据验证，这对团队中的每个人的能力、体力、精力是考验也是折磨。那时我就想，如果能有一个工具可以即时呈现、灵活验证、固化模型，将会减少非常多的工作量。

很幸运的是，2019年年底，我认识了喜乐君，他向我简单介绍了敏捷BI的代表作Tableau，不足两小时的时间，我就确定它能解决我多年来的数据困扰，能让我的团队发挥更大的价值。

喜乐君并非IT专业，却有深厚的业务背景，正因为如此，他更清楚业务分析的需求点，更明白如何可视化数据，让管理者一目了然，包括各种维度、习惯，所以基于业务目标或者问题的分析，他总是能深入浅出地解析。他那种不骄不躁的沉稳、执着让人佩服，Tableau工具的灵活、便捷也被他演绎得淋漓尽致。当他说想写一本Tableau主题的数据分析的书，帮助所有爱好者、需要者快速掌握并受益时，我就非常期待

这本书的问世。

时隔不足半年，喜乐君便将本书的第一稿寄给了我。他用通俗易懂的方式，特别是大量的图形和示例，向读者展示了他多年来使用Tableau的经验和思考。特别是他总结的报表分析的多个步骤、高级分析的过程，可以帮助公司的数据分析师少走弯路，从而大幅度地提高效率。

衷心地感谢这本书介绍的工具、方法、分析维度给我带来的轻松感，它可以说是让IT部门升级为DT部门并得到企业认可的催化剂。

湛江国联水产开发股份有限公司，信息总监 王强

2020年4月26日

推荐序四

能为我的学生作序，十分开心，因为我看到玉鹏（喜乐君）在继续前行。

我和玉鹏在2007年的山东大学创新教育的课堂上结下师生之缘，至今保持了十多年的师生友谊。玉鹏是法学本科和教育学硕士毕业，对学习创新一直保持着渴望，更难能可贵的是，他能在学习中加入自己的理解、展望。他早早地寄来他的书稿请我过目，这本书有许多与众不同之处。

其一，他在书中展示的不只是操作方法，而更多是自己的理解；其二，理念超前，他从自己的分析经验中努力构建各种分析方法，“营造方法”是发明创造中关键的部分。

玉鹏时常和我沟通他的工作、学习，以及动态性思考问题的方法，呈现在你眼前的这本书，是他最近几年的学习成果。当你读完本书后就会发现，他没有教科书一般地罗列知识点，而是独具匠心地将软件的工具知识和来自经验的所思所想创造性地结合，并用问题分析和层次思维一以贯之。

如今，数据分析师成为21世纪最重要的职业之一，正因如此，越来越多的大学开设了数据相关专业，旨在为企业培养具有丰富的知识基础、创新性的分析能力的从业者。以Tableau为代表的数据可视化分析软件，是这个专业中非常重要的细分领域。不管是大学生，还是企业的从业者，都可以从本书中快速地了解数据可视化分析的思维和原理，以及Tableau的基础知识与应用。

本书中也包含了很多企业真实的分析场景和案例，故有助于读者提前了解企业的真正需求，更有针对性地进行学习，对于尚未进入职场的学生或者初入职场的新人而言非常重要。

期待每位读者都有所收获。

山东大学发明创造研究室主任 王思悦教授

2020年4月

自序 鸟会飞是因为有羽毛吗—— Tableau与笔者的分析之旅

从2017年在婴儿儿偶遇Tableau至今整整三年时间，从昨日的爱好到今朝的工作，仿佛一瞬，又好似半生。如今，笔者完成了之前未曾想象的任务——把笔者的所思所想、所知所悟以出版的方式分享给更多人。

理想主义者总是习惯性地低估困难，写书这件事情尤其如此。累计638张精心制作的插图，有别于博客内容，建立新的体系框架，你我虽隔书相望，但笔者希望每一位读者都能感受到我毫无保留的写作态度与努力。

与此同时，还是想说一下笔者和Tableau的渊源，以此说明笔者如何以文科学历和业务背景从零开始成为今日的“Tableau大使”，这条路每个人都可以走，只需要用心与努力即可抵达。

1.我和Tableau的渊源

毕业后历经国企、创业、私企几番锻炼，2017年回到婴儿儿担任总裁助理，忙里偷闲四处学习，并且获得了“买任何图书均可报销”的公司特权，受领导鼓励，也在公司义务培训Excel、消费心理学等；考虑到公司低效的“PPT数据传统”和自身专业数据分析知识的薄弱，因此私下搜寻各种大数据分析工具，最后被Tableau的灵活、易用和美观所折服。之后陆陆续续为运营、采购、人资等板块做了一些并非成熟的分析。

笔者是典型的“写作型”，因此从学习第一周开始，笔者就陆陆续续记笔记、写博客，纯粹为了帮助自己增强理解，不料三年下来，竟然积累了可观的笔墨。笔者相信“所有的成功都是长期主义的胜利”。数据和

数据分析恰好是不错的风口，而且至少会常年不衰，于是误打误撞进入了这个“陌生但新鲜的行业”。

《经济学人》中曾写道：“21世纪最重要的资源是数据”，但是不经过分析的数据没有价值，如同不经过反省的人生没有意义，而这正是转型期的企业遇到的成长烦恼。笔者决定和Tableau同行，将自己多年的工作经验与笔者对数据的理解融为一体，认真服务每一位客户，同时获得自我的提升。笔者选择了Tableau，之后通过了Tableau Desktop和Server的原厂QA认证，并在参加Tableau峰会时认识了众多的Tableau员工和爱好者，之后开始了开发客户、服务客户的美好旅程。

在服务客户的过程中，笔者不断积累自己的Tableau知识和业务理解，并持续更新博客增强理解并向更多客户传播Tableau文化。笔者从不拒绝客户的任何问题，把它视为最好的收集问题和不断学习的机会——没有什么是学习不能解决的问题，如果有，那就是学艺不精。

2.从所知到所悟

在学习过程中，笔者不断阅读各类数据分析的书籍，并仔细翻阅官方近万页的文档和白皮书。可惜的是，国内外笔者能找到的每一本Tableau主题书籍，只能满足笔者的初学，却不能满足中高级进阶的胃口，总觉得要义未精、框架欠明，如同武林秘籍缺少最后一章，即便各种招式纯熟，却难以在实战面前随心所欲。这种理解上的束缚，阻碍了为客户交付最高品质的培训、实施和咨询。跟随山东大学王思悦老师十年学习，他教给笔者一种处事态度：“和人交往改变自己，和物交道改变对方”，因此，笔者希望重新构建Tableau的知识体系，并希望帮助初学者和高级分析师都能更好地使用Tableau产品。

在克里斯坦森教授《你该如何衡量你的人生》一书的开篇，作者提出了一个让我终生难忘的问题：“鸟会飞是因为有羽毛吗？”笔者曾经以为是，但正如作者所言，人类上千年来一直尝试模仿轻盈的翅膀飞上天，最后，倒是成吨的钢铁飞机实现了。100年前，人类在“流体力

学”和“空气动力学”领域积累了足够的知识，才实现了飞翔的梦想，这就是原理的重要性，重要的不是翅膀，而是如何创造升力。

在瑞·达里奥的《原则》一书中，作者说：“要明白几乎所有‘眼前的情况’都是‘类似情景的再现’，要识别‘类似情景’是什么，然后应用经深思熟虑的原则来应对。”生活如此，工作如此，分析亦是如此。

因此，笔者迫切地希望洞察Tableau“拖拉曳”、可视化，特别是高级计算背后的原理，只有能掌握了原理，笔者才能用最简单的语言，让所有的客户以最低的时间和金钱成本换来最高效的培训和效果。

而通往大彻大悟的道路只有一条，那就是持续的努力和深度的思考相融合的道路。

整个2019年，笔者一方面不断地向Tableau最难的高级计算和高级互动发起总攻，并持续修改博客作为通达明了的明证；另一方面每月组织Tableau公开课程，在分享过程中不断深化自我理解，并在为中原消费金融、以岭药业等客户交付培训的过程中不断总结本书的宏观框架。2019年年底在国联水产的项目中，带着写书的心态为客户额外提供了多天的培训，又获得了本书第5章的关键灵感。如今，笔者的不少Tableau博客文章，特别是关于“LOD详细级别表达式”原理和案例解读系列，几乎可以与官方的介绍文章并驾齐驱。2020年年初，因疫情在家，得以从头重写每一个细节及其思路，并把基础计算和高级计算融为一炉，形成了全新的讲解体系，从而保证初学者也可以快速掌握最高难度的知识环节。

最后，笔者找到了从Excel分析到Tableau数据分析的根本性差异，即层次。客观的数据层次用于描述数据结构和颗粒度，主观的视图层次用于描述业务问题及其相关性，并通过计算的多种分类把二者融为一体。全书都贯穿了“层次分析”的思路，并在高级计算部分得以升华——高级计算的实质就是多层次问题分析。因此，读者在本书中能看到很多全新的内容，比如用DIKW模型理解数据的层次，用层次理解大数据分

析的核心特征，用层次理解数据结构并识别行级别唯一性，以层次理解Tableau的计算并引导如何选择等。

而精心绘制的插图，旨在用可视化的方式增强理解，而非仅仅是文本。并通过二次处理，尽可能提高每一个图片的知识密度。

同时，本书特别推崇集和集动作，大数据分析通常都是某一个样本的分析，集正是保存样本的绝佳功能，随着Tableau 2020.2集控制功能的推动，可以进一步将集作为传递多值变量的媒介，与之相对的是参数作为传递单值变量的媒介。笔者之前把“集、详细级别表达式和表计算”称之为Tableau的“三剑客”，如今有了数据关系，大家不妨以“F4”称之。

3.大数据时代的趋势与业务驱动的数据分析

随着互联网经济的蓬勃发展，大数据时代已经成为不可回避的事实。在经济危机面前，企业更应该追求精益分析驱动的精益成长。

正因为此，敏捷BI已经是大势所趋、不可抵挡。企业成长依赖于在竞争环境中不断做出最优的决策，而决策来自于充分的建立假设并高效地验证，数据分析是连接数据资产与价值决策的纽带，而敏捷BI提高数据的利用效率和企业的决策效率。“分析即选择，决策即择优”，数据分析可以直接创造企业价值，未来已来，所有的企业都将是数据驱动型的组织。

以Tableau为代表的敏捷BI，超越了Excel的局限性，操作灵活，对业务用户足够友好，帮助他们把数据与业务紧密结合，为企业中最庞大的群体打开了一扇进入大数据的窗户。

作为世界首屈一指的敏捷BI和大数据可视化分析平台，Tableau为企业提供了低成本试错、高杠杆收益，且面向业务、模型构建的解决方案。作为文科背景、业务出身、自学成长起来的Tableau分析师，笔者享受了大数据时代的“数据红利”，提前从传统BI切换到了敏捷BI的快车道，如果有朝一日笔者重返业务岗位，就如同手握尚方宝剑必然更加得

心应手。如果企业有更多用户能在Tableau帮助下发挥数据的价值，不仅能在危机面前确保个人的竞争力，而且能为企业创造更多的分析价值。

对于业务分析师而言，Tableau入门容易、使用灵活，因此它适用于企业中的几乎每一位数据用户和业务决策者；同时，Tableau博大精深、足够专业，在可视化样式、互动探索、高级计算等方面有无限空间值得探索，因此不断钻研的Tableau分析师可以为自己构建足够高的技术壁垒，从而捍卫自己的专业领地。这也是笔者的选择和道路，只要努力，人人皆可模仿，没有所谓的“学习力”，需要的只是用心和努力而已。

在这条充满光明的道路上，最大的障碍其实不是工具，而是人和文化。借助于本书，衷心地希望更多的人能熟练Tableau，并建立自己的职业壁垒，节省时间就是拯救个人生命，提高效率就是创造企业利润。

4.致谢

从博客到一本书，这是一年之前笔者还未曾预料的事情；因为疫情在家隔离，一个春天，不料梦想就变成了现实。

特别感谢笔者工作之旅中遇到的每一位同事和领导，以及服务的每一家Tableau客户。特别是山东婴贝儿的领导早年为笔者提供了广泛学习的机缘，中国软件的各位领导对我工作的支持，感谢北投集团、红塔山玉溪卷烟厂、航科院、中国人民银行沈阳分行、河南中原消费金融、以岭药业、石药恩必普、野村综研、大连日信、上海烟草、国联水产股份、青岛啤酒、烟台创迹等众多客户对笔者的信任和支持。

感谢电子工业出版社石编辑为本书付出的努力，她帮笔者实现了而立之年的第一个梦想，也帮助大家目睹了这本书的精彩。

特别感谢2019年陪笔者一起学习的几位朋友：济南公安局于警官、秦皇岛税务局冯伟、沈阳李博、百威啤酒刘洋、婴贝儿史国丽等，他们不远万里来听我不成熟的课程；古人云“教学相长”，在笔者才疏学浅之

时，每一位聆听者都是对笔者的激励。

特别感谢Tableau给我的学习机会，认识了各行各业的企业客户和朋友，结交了众多的Tableau爱好者。

感谢山东大学七年求学历程最重要的导师王思悦教授，追随他学习“发明创造学”前后逾十年，他已年迈，我正年轻，亦师亦友，受益良多。感谢每一位信任与支持笔者的朋友，他们给了笔者诸多勇气。

感谢我的家人，他们给了笔者生活的意义，并陪伴了笔者写书的每一天。

感谢时间，感谢充满坎坷与喜乐的人生。

喜乐君（吴玉鹏）

2020年6月10日

注：

本书主体部分依据Tableau 2020.1版本完成写作和绘图；同时增加了Tableau 2020.2版本的新功能，特别是数据关系、集控制。不同版本之间的界面略有差异，但不影响功能展示和使用，后续重印或修订会逐步更改。

第1篇 从数据到图形：Tableau可视化

第1章 可视化分析：进入大数据时代的理性与直觉之门

关键词：数据金字塔模型、分析与决策、直觉、可视化心理学

数据分析的目的是辅助决策，这就需要数据分析师和业务决策者能快速获取外部数据并高效地做出分析。可视化分析借助位置、颜色、长度、形状、大小等直观可见的方式表达数据，帮助我们更快地识别数据中的关键信息，发现数据背后的逻辑关系，从而做出业务决策。

在本章，笔者将结合实际的业务决策过程，从数据金字塔模型、数据决策过程、可视化心理学等方面，向读者介绍数据分析的关键背景。本书中的“业务分析师”指业务部门的数据分析师，而“IT分析师”指信息部门或者IT部门的数据分析师。

1.1 数据金字塔：从数据到决策有多远？

由于计算机技术的迅速发展，人类从数据稀缺，进入了数据爆炸的时代。伴随着产生的一个问题是：如何将数据、信息转化为知识，从而

有效辅助决策？

在20世纪90年代以前，辅助决策一直未能有效发展起来，直到关系数据库兴起，数据挖掘、可视化分析崭露头角，辅助决策才快速发展起来。1989年，知名咨询机构Gartner的报告中明确提出了BI（Business Intelligence，商业智能）的概念，指“使用基于事实的支持系统支撑商业决策的概念与方法”（Concepts and methods to improve business decision making by using fact-based support systems），自此之后，BI的概念和应用逐步被官方使用。如今，Gartner的“BI和分析平台魔力象限”（Magic Quadrant for Business Intelligence and Analytics Platforms）代表了整个行业的风向标，而我们的主角Tableau则已经连续8年出现在领导者（Leader）象限中，不出意外，Tableau还将持续代表敏捷BI行业的发展方向。

得益于互联网经济的引导和驱使，众多的中国企业开始重视IT软硬件设施、数据收集和存储等方面的投入，企业的数据量也开始呈指数级增长。不过，数据并不意味着价值，分析和决策才能创造价值。笔者最爱的管理大师彼得·德鲁克当年的感慨，特别适合于当下的中国：

“迄今为止，我们的系统产生的还仅仅是数据，
而不是信息，更不是知识。”

那什么是数据、信息和知识呢？涂子沛老师在《大数据》一书中举了一个形象的例子，“185”“奥巴马”等仅仅是孤零零的数据，只有当将这些数据置于特定背景时，比如“奥巴马身高185cm”，相互独立的数据才转化为有效的信息；基于更多的数据就会发现特定的规律，比如“大多数成年美国人的平均身高为185cm”，这样就积累了特定的行业知识。因此，数据仅仅是分析的原材料，知识才是数据分析的最终产品，也是辅助决策的关键依据。

也就是说，数据本身并没有价值，价值来自数据整理、分析和加工的综合过程，而人的智力和经验，是数据分析过程中最重要的“催化剂”。从数据到信息，再从信息到知识，构成了数据金字塔最主要的三个层次。在这三层模型中加入“Wisdom”，笔者称之为“智慧”或者“洞见”，就形成了标准的如同金字塔的“DIKW模型”（见图1-1）。DIKW模型清晰地表述了从数据到信息，从信息到知识的过程，反复积累的知识不断提升了我们的心智和智慧。由于每一次对数据的分析，都是答疑解惑和数据增值的过程，可以理解为“数据密度”在不断增加——一张A4纸放不下上市公司一天的营业数据（Data），却能给投资者展示充满价值的业绩简报（Information）。

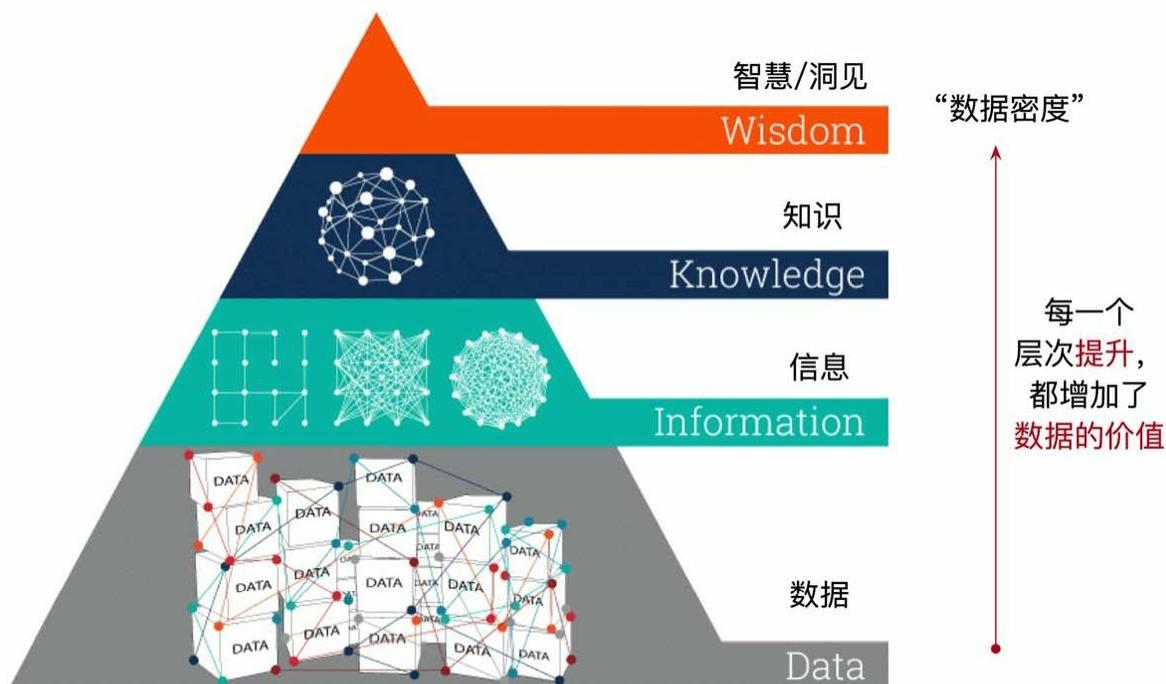


图1-1 DIKW数据金字塔模型

《经济学人》发表的一篇文章中称，“21世纪最有价值的不是石油，而是数据”。在一次给中石油山东公司的员工上大数据分享课时，笔者把数据分析的过程比作石油勘测、挖掘和提炼的过程，以此形象理解数据分析各环节的含义，如图1-2所示。

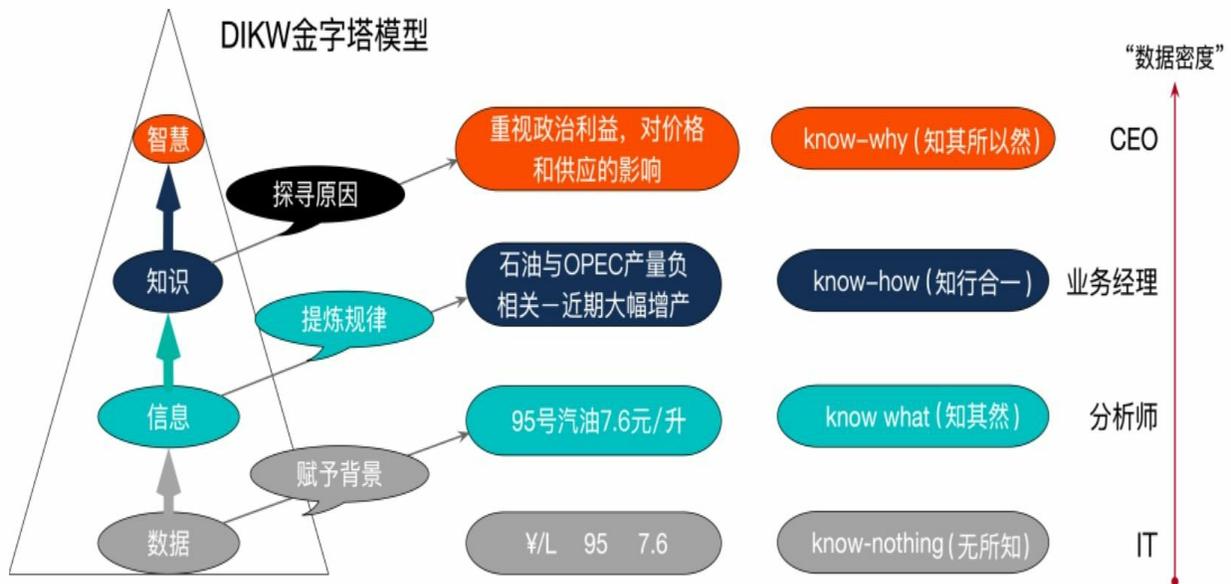


图1-2 DIKW模型各层次关系

- **Data（数据）**：是理解事实的符号，比如数字、单位、程度描述等，在未被整理或者加以理解之前，它是无用的，有人称之为“know-nothing”（无所知）；不经分析的数据如同不经反省的人生，如同地底下未经开采的石油，存在却缺乏意义。计算机用字段描述数据，详见第2章，对应Tableau的基础概念、拖曳逻辑。
- **Information（信息）**：信息是带有逻辑的数据组合，多是结构化表述，比如“95号汽油7.6元/升”；通过信息，我们可以了解数据背后的世界和关系，因此称之为“know-what”（知其然）。从数据到信息的过程如同从地壳中勘测和挖掘石油，这个过程的基础是数据整理和数据准备，详见第3章、第4章，对应Tableau Prep Builder和Tableau Desktop的部分功能。
- **Knowledge（知识）**：知识是在众多数据、信息中增加了主观理解并进一步升华的数据见解，因此因人而异；和信息不同，知识是直接指导业务决策和行动的，因此能直接产生价值，故称之为“know-how”（知行合一）。就像石油经过精炼转化为汽油，汽

油为汽车提供动力，知识是数据分析最关键的产品。

- **Wisdom**（智慧/洞见）：古人说“玄之又玄是为道”，到了这一层，就融合了决策者的深层理解和经验性的洞见，透析数据，不仅知其然，更知其所以然，此乃妙理，即“know-why”（知其所以然）。每个公司总有少数的业务领导和管理者可以通过微小的数据线索判断行业大势，见微知著，预判未来。所有的智慧和洞见背后，是更加抽象和前瞻性的数据逻辑和知识体系。

简而言之，数据分析就是从数据中提取和整理信息，进而总结知识、增进洞见，并指导决策的过程。

在企业中，DIKW模型的每一个层次对应不同的“数据相关者”。如图1-2所示，数据层面对应IT人员（管理和维护数据），信息对应分析师（IT分析师或者业务分析师），知识对应业务经理（基于数据做决策的人），而智慧对应公司高管和CEO（领导业务经理看到数据，也看到未来）。

随着数据爆炸，企业在数据领域的主要焦点从“如何获得更多数据”变成了“如何做出更有助于决策的分析”。而影响决策分析的主要矛盾是“拥有数据的IT分析师不了解业务逻辑与直接做出数据决策的业务经理难以精通数据分析方法”之间的矛盾。正是意识到了这一点，从新兴的互联网公司到传统的医药公司，越来越多的企业正在将数据分析工作从信息部门转向业务部门，甚至在业务部门中成立专门的数据分析团队。

也正因此，Gartner在2019年的BI分析报告中写道：“到2020年，业务部门的数据以及分析专家数量的增长速度将是IT专家增长速度的3倍，这将迫使企业重新考虑其组织模式，以及人力资源管理。”如今，这一预测正在逐步成为现实，经济危机进一步促使企业领导重视数据分析的重要性——借助数据分析，进一步降低决策的试错成本。

对业务分析师和业务经理而言，可视化的数据分析是进入大数据时

代最好的捷径，因为这条道路符合人类直觉决策和理性决策结合的基本逻辑。

1.2 直觉先于理性：可视化的心理学

人类历史从有考古记录开始虽已有三百多万年，但有文字记录却不足一万年。在几百万年的历史中，人类和其他动物一样要随时面临各种突如其来的危险，判断必须迅速而敏捷才能免于死亡。时至今日，每日的穿衣出行、躲避车辆、家庭言语，甚至手机游戏，80%以上的决策依然依赖于眼、耳、鼻、舌、身的感官信号，无须大脑深刻思考，根据经验和习惯做出快速判断，这种方式不仅安全，而且节省能量。著名行为经济学家、诺贝尔奖得主丹尼尔·卡尼曼在《思考：快与慢》一书中深刻揭示人类决策时直觉思考和理性思考二者搭档的过程，并深入分析了这种决策模型在经济运行中的决策行为——经济领域的决策分析，可以视为数据分析的黄金地带，融合了各利益相关方的理性思考与直觉反应。

我们身边有很多关于可视化的经典案例。比如机场、高铁站的公共洗手间普遍都是用“图形”来代表男/女洗手间，用箭头指引路线；在全世界的各个十字路口，用红绿灯指挥交通；世界顶尖的公司或者公益组织，普遍选择用图形作为Logo，比如华为的“花瓣”、苹果公司的“苹果”。这些被广泛使用的标志背后，都是可视化元素的典型应用，是为了更好地顺从消费者的直觉思考过程。

人类进入文明时代后，这样的“直觉决策系统”依然被保留下来，将频繁发生的相似事件前的决策难度降低；而文明社会的标志在于“理性决策”的快速发展，我们不再仅仅依赖眼、耳、鼻、舌、身的直觉，而更依赖逻辑、理性和深刻的常识，超越此前的认知限制，从而可以研发疫苗对抗病毒、发射火箭探索太空。直觉判断和理性思考是人类的决策

体系，也是数据分析的基本过程。

随着大数据时代的数据爆炸，数据噪音越来越多，快速、有效地表达信息，就成了数据分析的关键。交叉表和简单图形的方式逐步失效，分析师必须使用更好的展示方式，而数据可视化是最佳的窗口，它有足够的“知识密度”，且能直观地展示数据重点。直觉判断是理性思考的引路人，可视化是大数据时代的解析语言。

可视化分析的关键是选择最佳的“前注意属性”。现代心理学把颜色、形状等能快速引起心理反应的信号统称为“前注意属性”，它们在我们的潜意识中活动，只需要0.25秒就可以做出识别，因此是可视化分析最佳的引子。这些“前注意属性”主要包括如图1-3所示的几种。

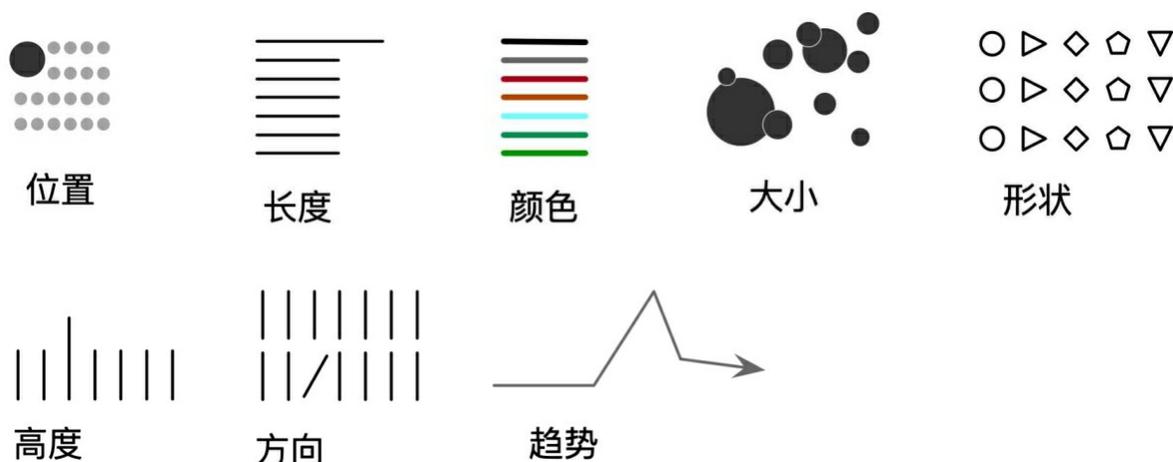


图1-3 常见的“前注意属性”

首选的“前注意属性”是位置，即把最重要的信息放在最关键的地方，比如企业数据简报把营收和增长率放在首页顶部。

“因位置胜利”的经典案例是印度数字（阿拉伯数字）。印度的十进制和0经过阿拉伯人传入欧洲之后，希腊数字就逐渐衰落，关键原因之一，印度数字借助于0和十进制能更简单地表达“大数据”，无须思考即可理解数字大小。举例来说，如果想要表达“1888”，印度数字的逻辑是左边的数字更大，逢10进1。如果用希腊数字表示，则需要 **MDCCCLXXXVIII** 这样的一串字符，希腊数字没有进位制，位置不能

代表大小，只能通过计算来表达（见图1-4）。想象一下，如果用希腊数字来做财务报表.....

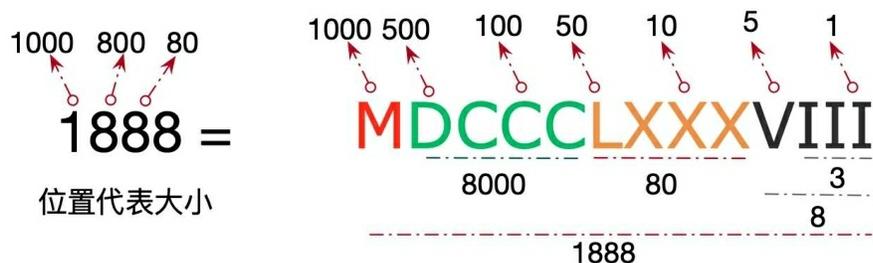


图1-4 印度数字和希腊数字的对比

排在位置之后的常用属性是颜色和大小，在不同场合下大小又可以分为长度、高度，甚至角度、面积等多种形式，分别应用在条形图、柱状图、饼图和树形图中（见第5章）。

在传统的数据分析中，条形图、折线图和饼图是三大基本图表，而大数据分析更强调大数据样本的宏观特征、分布规律和相互关系，所以就有了直方图、盒须图和散点图等高级图形。所有的图形，都是通过长度、颜色、形状等直观方式增加信息展示的深度和层次的，比如图1-5所示的“准时装运趋势”中，使用颜色代表分类，使用面积代表数量，使用条形图的长度代表发货时长，从而清晰地表示数据的相互关系。

准时装运趋势

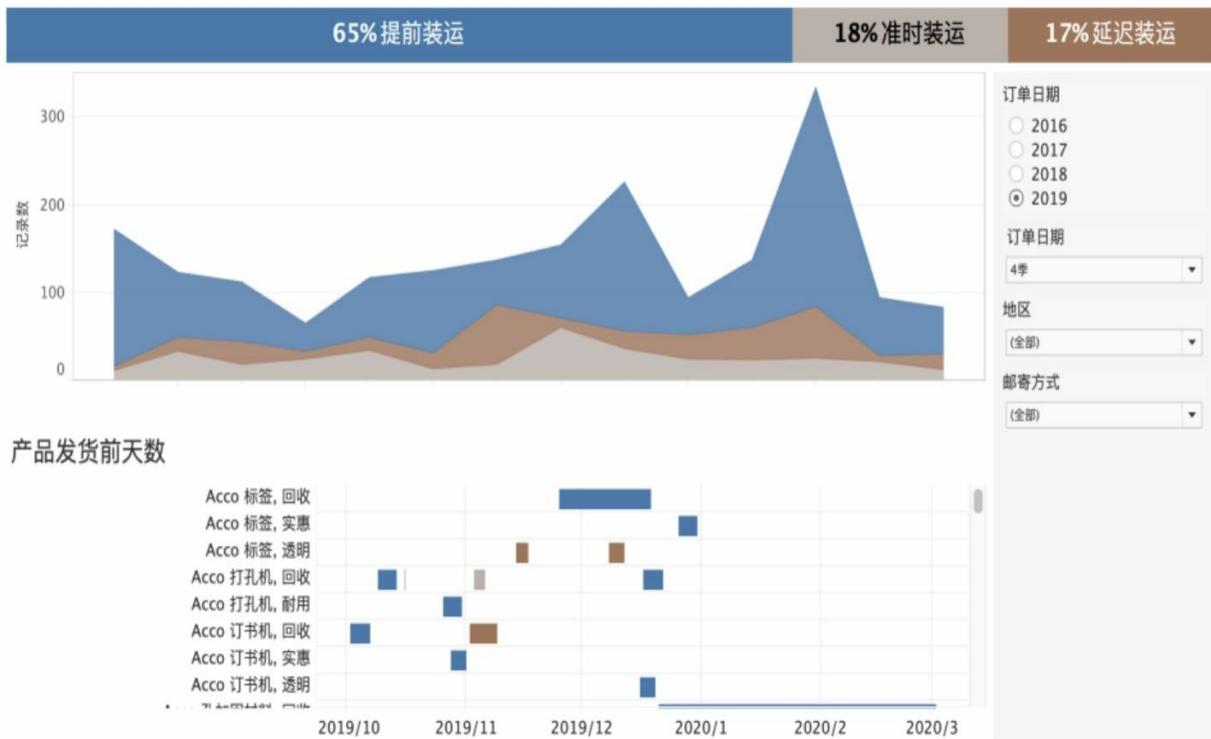


图1-5 基于超市数据的可视化案例

在企业的运营过程中，无时无刻不在做决策，或者为下一个决策做准备。直觉思考虽然快，但是不适合复杂问题；理性思考虽然可靠，但是效率却不够高。可视化分析借助图形要素简化数据表达，从而节省更多的心智资源用于理性思考，实现二者的平衡。

也可以说，可视化的数据分析是艺术与科学的结合，是快思考与慢思考的结合，它努力帮助我们从小多样、快速、庞大的数据中快速识别数据线索，进而借助经验和思考验证假设，从而实现数据辅助决策的目标。

此时，我们就需要一个真正敏捷的数据可视化分析工具，让每一位业务运营者都能随时分析数据，提高决策的效率。可视化的敏捷BI分析，能帮助更多的人，特别是业务分析师和业务经理整理和分析数据，从而提高决策的效率。

Tableau正是顺应了这样的时代大趋势，从单一的可视化分析工

具，快速成长为企业级的数据可视化分析平台。Tableau Prep Builder帮助用户更快、更好地整理数据，Tableau Desktop通过拖曳分析、计算和互动把数据转化为信息和知识，而Tableau Server帮助企业把每个人的知识转化为组织的知识，进一步提高数据分析的价值。持续使用Tableau等BI产品，可以有效帮助我们增进对数据的理解，帮助企业成为“数据驱动组织”（见图1-6）。



图1-6 Tableau主要产品在数据分析过程中的位置

笔者毕业后经历过多个行业和岗位，在国企和零售私企的经验格外记忆犹新。在日益激烈的竞争面前，不管是增加效率还是降低成本，不管是国企还是私企，“拍脑袋”已经是高风险的决策方式，越来越多的决策需要依赖数据分析去证伪或证实。借助商业智能工具，企业就能将数据转化为资产，实现以数据分析来驱动业务进步，在业务增长和市场竞争方面领先于同行。

1.3 Tableau: 大数据时代的“梵高”

我们知道有很多数据存储和分析工具，比如Excel、WPS，还有各种数据库软件，比如SQL Server、Oracle、PostgreSQL等，它们普遍存在于从手机通信录到企业ERP的各种场景中。数据越来越多，分析数据却越来越难，在越来越激烈的市场竞争中，必须由掌握数据逻辑和最终决策的群体接管数据分析过程，技术是数据探索的助力，而经验和知识才是数据分析最好的催化剂。正如Tableau所倡导的一样，“谁提问题谁寻找答案”。

对于拥有数据、拥有经验的业务人员而言，唯一的障碍就是工具了。简单的数据工具无法满足需求，SQL语言过于抽象和艰涩，难以作为基础知识普及。

大势已到，只欠东风，在业务用户逐渐成为分析主力军的时代，我们需要全新的直观、快速、敏捷、易用的数据处理技术。Tableau独创VizQL技术，正是顺应了这样的潮流，快速发展并被行业所普及，它具有SQL查询的综合性功能，又兼顾了业务人员的便捷易用需求。

“Tableau帮助人们查看并理解数据。”

VizQL如同为SQL数据库查询语言封装了一件人人易懂的“新衣”，而把复杂的技术交给了黑箱。如图1-7所示，只需要拖曳动作，就能生成可视化图形、增加分析深度、筛选数据样本、创建计算字段，甚至创建综合仪表板和发布数据故事。本书将在第2章介绍创建简易可视化图表的过程，并在第5章详细介绍创建高级可视化图表的方法。

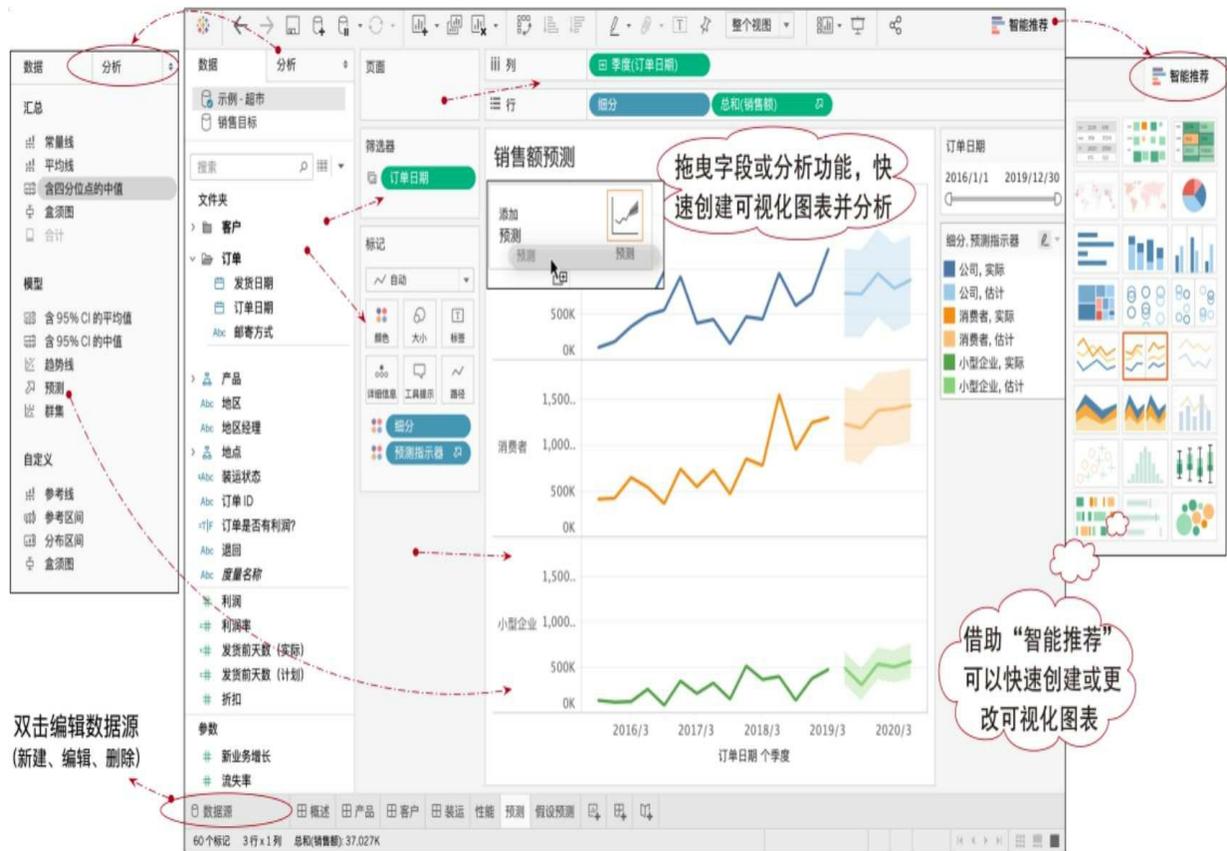


图1-7 通过拖曳字段或者分析功能实现可视化分析

特别是在数据可视化领域，Tableau无愧于“大数据时代的‘梵高’”的美誉。

笔者作为没有技术背景的业务人员，因为工作需要寻遍市面上的各种BI产品，最后选定的Tableau不负期望，帮助笔者极大地提高了零售分析、营销定价等方面的工作效率，也不断积累了本书的素材。如今，Tableau已经从可视化分析工具，逐步发展成为企业级的数据可视化分析平台，并在自然语言、AI分析等方面持续进步。

一方面，Tableau不断拓展和丰富产品线。

如图1-8所示，Tableau在Tableau Desktop可视化工具的基础上新增Tableau Prep Builder，弥补了此前敏捷数据整理的短板（2018年），在Tableau Server基础上扩展了数据治理和大规模管理的平台功能（2019年）。特别是Tableau Prep Builder，笔者对它的热爱甚至要超过早年对