

智能会计人才培养新形态系列教材

财务数据可视化

黄璐 赵盈科 主 编

高恺迪 黄镜霖 副主编

清华大学出版社
北 京

内 容 简 介

随着我国信息化进程的不断加速及大数据和人工智能技术的不断发展,财务人员应全面、深入地分析财务数据,帮助企业预测和防范经营过程中可能遇到的风险。本书分两篇,上篇为理论知识篇,以Power BI 为数据分析工具,重点介绍商业智能的基础理论知识和Power BI在数据收集、整理及可视化中的应用;下篇为实战应用篇,通过财务数据案例,详细介绍财务大数据分析和可视化的全过程,并讲解应如何规范地编写企业财务大数据分析报告,以更好地赋能大数据会计人才培养。

本书内容新颖,难易适中,可作为高等院校本科及研究生财会类专业的教材,也可作为企业管理人员的参考书。

本书提供丰富的教学资源,包括但不限于教学课件、操作视频、教学计划、教学大纲。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。举报:010-62782989, beiqinquan@tup.tsinghua.edu.cn。

图书在版编目(CIP)数据

财务数据可视化 / 黄璐, 赵盈科主编. -- 北京:
清华大学出版社, 2024. 8. -- (智能会计人才培养新形态系列教材). -- ISBN 978-7-302-66910-4

I . F231.2-39

中国国家版本馆 CIP 数据核字第 20240R4C92 号

责任编辑:高 岫

封面设计:周晓亮

版式设计:方加青

责任校对:马遥遥

责任印制:宋 林

出版发行:清华大学出版社

网 址: <https://www.tup.com.cn>, <https://www.wqxuetang.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-83470000 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:小森印刷霸州有限公司

经 销:全国新华书店

开 本:185mm×260mm 印 张:11.5 字 数:252 千字

版 次:2024 年 9 月第 1 版 印 次:2024 年 9 月第 1 次印刷

定 价:49.00 元

产品编号:104159-01

前言

在这个信息化飞速发展的时代，互联网、移动互联网、云计算、大数据、人工智能等新技术不断涌现，给企业的经营管理方式带来了革命性的影响。新质生产力，是指以信息技术为核心的现代科技，特别是智能化技术在生产过程中的应用，极大地提高了生产效率和产品质量，改变了生产方式和经济结构。党的二十大报告中指出：“必须坚持科技是第一生产力、人才是第一资源、创新是第一动力，深入实施科教兴国战略、人才强国战略、创新驱动发展战略，开辟发展新领域新赛道，不断塑造发展新动能新优势。”创新驱动发展战略的核心，就是发挥新质生产力的引领作用，加快传统产业的升级改造，促进新兴产业的蓬勃发展。企业管理者们正面临着如何有效利用这些新技术来提升企业竞争力，同时预见并应对经营过程中可能出现的各种风险的重要课题。

在这样的背景下，财务数据可视化成为企业管理中的一个重要趋势。Power BI作为一款强大的数据可视化工具，能够帮助企业将海量的财务数据转化为直观、易懂的图表和报告，从而提高决策效率和准确性。通过Power BI，企业不仅能够实现数据的实时监控和分析，还能够深入挖掘数据背后的业务逻辑，为企业的战略规划和日常管理提供有力支持。

随着人工智能技术与财务会计的深度融合，财务人员的角色也在发生变化。他们不仅是数字的记录者和报告者，还是业务与财务之间的桥梁，为企业的生产经营提供有效的财务支持。Power BI的应用，使得财务人员能够更加专注于数据分析和业务洞察，而不是烦琐的数据整理工作。企业高层对于管理的实时性和精确性要求日益提高，希望能够随时随地地通过移动设备了解企业的财务状况和业务动态，及时地反映企业经营中可能存在的问题，并预见未来的风险。Power BI的移动应用功能正好满足了这一需求，使得企业高层能够在任何时间、任何地点，快速获取企业的关键数据和分析报告。此外，由于企业业务的扩展和市场营销范围的扩大，跨地域、多组织的运作模式成为许多企业的常态。Power BI的多维度分析和报告功能，能够帮助企业及时有效地反映多组织、多业务单元的经营绩效，为企业的战略调整和资源优化提供数据支持。

本书以Power BI为工具，结合企业实际业务流程，详细介绍如何进行财务数据的可视化分析。本书通过设计具体的企业案例，帮助读者深入理解Power BI在业财融合模式下的应用，以及如何将所学知识应用于企业的实际财务业务中。

本书共分为7章，内容涵盖数据导入、数据清洗、数据建模、图表设计、仪表盘创建、报告发布等。同时，为便于教师授课和学生自学，本书提供丰富的教学资源，包括但不限于教学课件、操作视频、教学大纲、教案，读者可通过扫描右侧二维码获取。



教学资源

与市面上现有的同类书籍相比，本书具有以下特点。

(1) 理论够用为度，能力培养为主。本书注重理论与实践的结合，使读者在了解商业智能、数据可视化基础之上，将这些知识应用于实际的数据分析和可视化工作中，从而提高工作效率和决策质量。

(2) 简单易学、易上手。利用Power BI清晰且直观的用户界面，使读者通过拖放等简单操作便可创建数据模型、构建报告和仪表板。教材内容简练不繁杂，易于初学者快速上手。

(3) 结构清晰、内容全面，系统性强。本书按照财务数据可视化的流程设计内容体系，力求全面、系统地介绍Power BI的基本操作、数据分析技术和可视化策略。

(4) 项目驱动学习。通过具体的案例分析，可以使读者借助Power BI软件的各项功能，将所学知识应用于项目中，以实现学习与工作的无缝对接。

(5) 思政引领，紧跟时代发展。本书每一章在案例导入模块配有相关的思政阅读材料，培养学生的社会责任感和家国情怀，同时本书结合财务数据可视化趋势和技术，为读者提供行业信息和发展方向。

(6) 适用于不同层次的读者。无论是Power BI的初学者，还是有一定基础的进阶者，都可以从本书中获得有价值的知识和技能。

本书适合作为高等院校财务管理、会计学、工商管理、信息管理与信息系统等相关专业的本科和研究生教材，也适合作为企业中高层管理人员和信息化主管的参考书使用。我们相信，通过本书的学习，读者将能够更好地利用Power BI这一工具，提升企业的管理水平和竞争力。

本书系2023年广州工商院校级教材建设项目“财务数据可视化(项目编号：2023JC-11)”、广东省本科高校教学质量工程建设项目“数智财会校企联合实验室”、2023年广东省本科高校教学质量与教学改革工程建设项目“多模态视域下财务数据可视化课程改革与实践”、广州工商学院会计学重点学科项目。本书由广州工商学院的黄璐、赵盈科担任主编，由广州工商学院的高恺迪、黄镜霖担任副主编，广州城市理工学院郑磊老师参与了部分章节内容的编写工作。

本书在编写的过程中，得益于多部优秀财务数据可视化教材的启发，也凝聚了编者们的多年的教学经验和心得体会。由于编写水平所限，本书内容难免存在不足之处，恭请广大同仁及读者批评指正。

编者

2024年6月

目 录

上篇 理论知识

第 1 章 商业智能概论	2
1.1 商业智能基础	3
1.2 商业智能应用	5
1.3 商业智能分析工具	7
复习与思考	9
参考文献	9
第 2 章 数据可视化基础	10
2.1 数据类型与财务大数据	11
2.2 常见的数据分析模型	13
2.3 数据可视化的主要内容	16
2.4 常见的可视化图表	19
复习与思考	20
参考文献	20
第 3 章 Power BI 概论	21
3.1 Power BI 简介	21
3.2 Power BI 安装	25
3.3 Power BI 界面	28
复习与思考	38
参考文献	38
第 4 章 Power BI 应用	39
4.1 数据获取	39
4.2 数据整理	48
4.3 数据建模	73
4.4 常用公式及函数	78

4.5 数据可视化	90
复习与思考	126
参考文献	127

下篇 实战应用

第5章 财务数据分析与可视化	130
5.1 数据获取与整理	130
5.2 资产负债表分析与可视化	135
5.3 利润表分析与可视化	139
5.4 现金流量表分析与可视化	141
5.5 偿债能力分析与可视化	144
5.6 营运能力分析与可视化	145
5.7 盈利能力分析与可视化	146
5.8 杜邦分析与可视化	148
复习与思考	149
参考文献	149
第6章 企业财务与运营数据可视化分析	150
6.1 销售分析与可视化	150
6.2 应收账款分析与可视化	160
6.3 费用分析与可视化	163
6.4 完成度分析与可视化	166
复习与思考	168
参考文献	168
第7章 在线共享发布	169
7.1 分享与协作	170
7.2 移动应用	174
复习与思考	176
参考文献	176

上篇 理论知识

第1章 商业智能概论

【案例导入】

大数据在电商平台的应用

2016年4月，杭州市率先在全国提出建设城市大脑，以城市交通领域为突破口，开启了利用大数据技术破解城市交通拥堵问题的新探索。例如，在2022年高德交通发布的全国50个主要城市拥堵指数排名中，杭州市从2014年的全国第3位下降到2021年的第34位。杭州城市大脑“交通治堵”大数据决策还入选了《2020年联合国电子政务调查报告》中全国三大经典案例之一。在由“摸着石头过河”的局部探索迈向“投着石头过河”的系统实践转变过程中，杭州城市大脑在“交通治堵”领域已经培育出一套集信号灯智能调控、自动监控交通异常事件、特种车辆优先通行等功能为一体的智慧交通大数据决策系统。杭州城市大脑“交通治堵”大数据应用场景是杭州智能交通系统数字化、网联化、协同化发展 to 一定阶段的表现形式，其大数据应用涉及接入交警卡口、视频监控、信号灯、路面车流等多个设备和交通、城管、建设等多个部门及系统的多源异构数据集，彼此具有关联复杂性。在不断优化市内交通出行方案的过程中，其数据空间扩展历经了一个从数据挖掘到数据识别再到数据赋能的过程，在打通跨部门、跨领域、跨层次的数据孤岛、数据壁垒方面积累了丰富的理论和实践经验。近年来，随着杭州城市大脑“交通治堵”经验的不断积累和应用场景的持续创新，潮汐车道时间设置、紧急车辆道路优化等交通细分领域的应用正逐步走向多元化、深层次，越来越多“沉淀”在交通系统中的数据被不断挖掘、识别和连接，用以更加精确地解决城市交通拥堵问题。习近平总书记在浙江考察期间观看了杭州城市大脑的案例展示，对杭州运用城市大脑推动城市治理体系和治理能力现代化方面的创新举措给予充分肯定，并指出“运用大数据、云计算、区块链、人工智能等前沿技术推动城市管理手段、管理模式、管理理念创新，从数字化到智能化再到智慧化，让城市更聪明一些、更智慧一些，是推动城市治理体系和治理能力现代化的必由之路，前景广阔。”



思政拓展

(资料来源：程聪，严璐璐，曹烈冰. 大数据决策中数据结构转变：基于杭州城市大脑“交通治堵”应用场景的案例[J]. 管理世界, 2023, 39(12): 165-185.)

1.1 商业智能基础

根据麦肯锡全球研究院的数据，在进入互联网时代以后，人类几天时间所产生的数据量就相当于此前上百年时间产生的数据量。如此庞大的数据量，使得企业的数据处理及数据分析能力成为了其独一无二的竞争优势资源。通过商业智能，企业能够迅速降低成本，提高净利润；能够实现产品创新，提高竞争力；能够获得用户的行为特征，赚取高附加值，等等。Amazon、Facebook、Google等互联网巨头都在通过大数据及商业智能分析获益。由此可见，大数据时代下商业智能的重要性日益凸显。

1.1.1 大数据时代的来临

我们身处的这个时代，高速发展、科技日新月异、信息瞬息万变。人与人之间的交流日益密切，生活也愈发便捷。大数据，正是这个时代的产物。它不仅正在影响着我们生活和工作的方方面面，更将在未来发挥举足轻重的作用。

麦塔集团(后被高德纳公司收购)分析员道格·莱尼在2001年的一场演讲中指出，数据增长的挑战和机遇有三个方向：大量(volume，数据规模)、高速(velocity，数据输入输出的速度)、多变(variety，数据类型的多样性)，合称“3V”或“3Vs”。2012年，高德纳公司修改了大数据的定义：“大数据是大量、高速及多变的信息资产，需要新型的处理方式去促成更强的决策能力、洞察力与最优化处理。”现今大数据产业中的大部分公司都使用“3V”来描述大数据。

1. 大量

顾名思义，大数据的规模相当庞大。以百度搜索引擎为例，有资料显示，为了在首页向目标用户提供更精准的个性化网站导航和信息推送，百度每天需要挖掘的数据超过1.5PB(1PB=1024TB)，这些数据如果打印出来，将超过5000亿张A4纸；如果把这些纸张连接起来，可绕地球上万圈。到目前为止，人类产生的所有印刷材料的数据仅为200PB，只相当于百度在130多天中挖掘的数据总量。

2. 高速

大数据区别于传统数据的最显著特征是它往往以数据流的形式动态、快速地产生，具有很强的时效性。只有掌控好数据流，才能有效利用这些数据。

3. 多变

传统意义上的数据是结构化的，便于存储和分析，大数据时代需要处理的则是文字、图片、视频、音频、地理位置等种类多样的半结构化和非结构化数据。在现代互联网世界中，半结构化和非结构化数据已经占据了整个数据量的绝大部分，高达95%以上，这无疑对数据的处理能力提出了更高要求。

大数据的难题是在众多数据中有价值的数据并不多，即“数据价值密度低”。所谓的海量数据，其实大多是未经过处理的原始数据，需要经过不断筛选、处理、分析，我们才能获得有效的数据。有价值的数据往往隐藏在看似并不重要的数据背后，

以视频为例，摄像头不间断地监控一小时所获得的视频，其中有用的数据可能仅有一两秒。

总而言之，在大数据时代，人们的一举一动都在产生海量的数据。其中，不仅包含数字的信息，而且隐含着规律和知识。这些规律和知识并不是显而易见的，而是需要通过深度的挖掘与分析才能获得的，商业智能便应运而生。

1.1.2 商业智能技术

商业智能(business intelligence, 简称BI, 又称商业智慧或商务智能)技术, 是现代企业满足大数据管理、分析需求的得力助手。它不仅能高效地展现数据发展变化的趋势, 还能挖掘数据中有价值的信息, 快速完成最新业务数据分析, 推动企业发现更多见解, 为企业管理和长远发展提供决策支持。

1. 商业智能技术发展概况

在日常工作中, 我们每天都会收到同事发来的各种类型的数据, 包括报表、电子表格、含图表的电子邮件。随着企业的发展, 数据越来越多, 在需要时快速找到所需数据变得更加困难。数据太多会增加使用的风险, 具体表现为数据更新不够及时, 大数据维护困难; 业务数据问题不能及时、直观地呈现, 分析维度不能根据业务需求灵活调整; 数据分享安全控制困难。

商业智能领域相关产品技术已经进行了几代更新, 从早期的技术人员利用程序语言定制开发的报表, 到自助式分析工具应用, 再到现在流行的数据分析可视化工具, 总体经历了三代技术更新。

2. 数据可视化解决方案介绍

第三代商业智能技术包括的数据可视化是指将经过采集、清洗、转换、处理过的数据映射为图形、图像、动画等可视化的形式, 并可以进行多维度、多层次的交互分析, 通过直观的可视化方式发现数据蕴含的规律和特征, 从而实现对复杂数据进行深入洞察。

Gartner, 作为全球领先的IT研究与顾问咨询公司, 每年2月都会发布分析和商业智能平台“魔力象限”。目前, 在商业智能可视化分析领域领先的是Microsoft、Tableau和Qlik。

这些商业智能平台共有的特点如下:

- 自助式数据分析能力, 彻底摆脱技术制约;
- 多维度、多层次的深度分析, 支持交叉筛选与交互式分析;
- 炫酷且实用的可视化效果;
- 数据导出功能, 轻松导出可视化图表背后的数据;
- 提供数据分析报告, 用数据讲故事。

商业智能可视化效果如图1-1所示。



图1-1 商业智能可视化效果图

1.2 商业智能应用

1.2.1 商业智能的应用价值

商业智能可以帮助管理者减少收集、处理信息的时间，使管理者能把更多精力用在决策上。商业智能的应用价值主要体现在以下几个方面。

1. 增强业务洞察能力

商业智能可以减少经营者收集数据、获取信息所花费的时间，加速决策过程，使正确的信息于正确的时间在信息系统中流向各类相关人员。决策者通过监控关键绩效指标(key performance indicator, KPI)，掌控业务执行的状况，以便及时调整策略。例如，管理者通过KPI监控销售人员最新的销售信息、任务额信息和任务完成度信息，可随时掌握企业的营收完成情况。

2. 优化企业营销策略

企业通过构建商业智能分析模型，可以深入挖掘消费者行为，从而制定适当的营销策略。“啤酒和尿布”的故事，就是著名的零售企业沃尔玛通过商业智能分析发现了尿布销售额和啤酒销售额具有一定的相关性，于是管理层做出决策，将尿布和啤酒这两种看上去不相关的商品摆放在一起销售，从而提高企业的销售业绩。

3. 提高市场响应能力

企业借助商业智能的大数据整合能力，将行业信息、政策法规等信息融入商业智能系

统，通过适当的模型以预测市场变化，精简流程，确定需要改进的环节，从而适应外部环境的变动。

4. 加强风险管理能力

企业可通过商业智能风险预警模型，发现企业存在的潜在风险，如经营风险、财务风险、纳税风险等。当出现这些风险预警时，企业可随时调整其经营策略来应对、规避、降低各类风险。例如，就贷款业务而言，银行可以应用数据挖掘技术对客户进行信用分析，发现其中的欺诈行为特征，作为有效的预警机制，在源头上为银行减少了潜在的损失。

5. 改善客户关系管理

很多企业正在逐渐由“以产品为中心”转化为“以客户为中心”。企业应用商业智能中的在线分析处理和数据挖掘等技术，对客户的交易记录等相关资料进行处理与挖掘，并对客户行为进行分类，然后针对不同类型的客户制定相应的服务策略。这类应用就叫作“客户智能”。例如，电信企业利用分析模型对客户行为、信用度等进行评估，为不同类型的客户提供有针对性的服务，从而提高客户的满意度和忠诚度。

1.2.2 商业智能的利器——图表

在这个信息激增的时代，人们每天接收到的信息严重过载，在这样的环境中，许多人失去了一字一句阅读的耐心，能被人看完的信息已经不多，能有效传达并被人记住的信息更是少之又少。在美国著名营销专家艾·里斯和杰克·特劳特合著的营销理论经典《定位》一书中提出了提升信息传达效果的解决之道：“在我们这个传播过度的社会里，最好的办法是传送极其简单的信息。你必须把你的信息削减了，好让它钻进人们的头脑，你必须消除歧义、简化信息，如果想延长它给人留下的印象，还得再简化。”

以如今的职场为例，尽管在各种展示、宣讲、汇报中用数据说话已经蔚然成风，但是相信很多职场人士在应用数据时都遇到过以下一些困境。

- 数据过于枯燥，难以引起受众的兴趣。
- 数据过于分散，难以有效利用。
- 数据过于庞大且质量参差不齐，难以挖掘到有价值的信息用于支撑观点或辅助决策。

想要简化数据并快速挖掘有价值的信息，图表是一种非常有效的方式，它能帮助企业更好地分析现状的本质，合理地预测未来，做出科学的经营决策。

图表以视觉化的方式呈现数据，将数据化繁为简，更易于理解和记忆，并使其更具说服力。和数据信息相比，图表具有以下特点。

1. 加速理解

对大脑的相关研究发现，大脑传输的信息中有90%是图像等视觉化的信息，大脑约有一半功能用来处理这类信息。大脑对图像和文本的处理方式也不同，处理图像的速度远远快于处理文本的速度。这就是“一图胜千言”的背后隐藏的科学道理。因此，图表能加快人们对数据信息的理解。

2. 化繁为简

当各类数据间的关系纷繁复杂时，使用图表以简洁的方式呈现数据，能帮助人们发现其中的关联，进而拓展思路，获得更多灵感，赋予数据更多的内容和意义。

3. 强化记忆

有研究发现，彩色图像能够显著提升人们的阅读意愿，由此引发的主动阅读相较于被动阅读，在记忆效果上有着显著的优越性。同时，视觉信号留存的时间也更长久。人们对于亲眼所见和亲身经历的事情能记住90%，通过阅读文字能记住的只有20%，而对于听过的事情则只能记住10%。可见，使用美观的图表能更好地吸引眼球并强化记忆。

4. 说服力强

有研究发现，通过演讲的方式能打动50%的观众，而通过视觉信息的展示能打动观众的比例则提升至67%。因此，合理利用图表能增强报告、演讲等的说服力。

5. 预测未来

预测在选举、体育、娱乐等领域被广泛地应用。在现代社会中，预测是基于过去和当前的历史资料，借助一定的逻辑推理和科学方法，对事物未来可能的发展趋势进行预计和推测。图表是在预测时经常使用的一种工具，对于直观地观察事物发展方向或发展趋势有很大帮助。如图1-2能很直观地反映出2016—2020年我国老龄人口增长趋势。

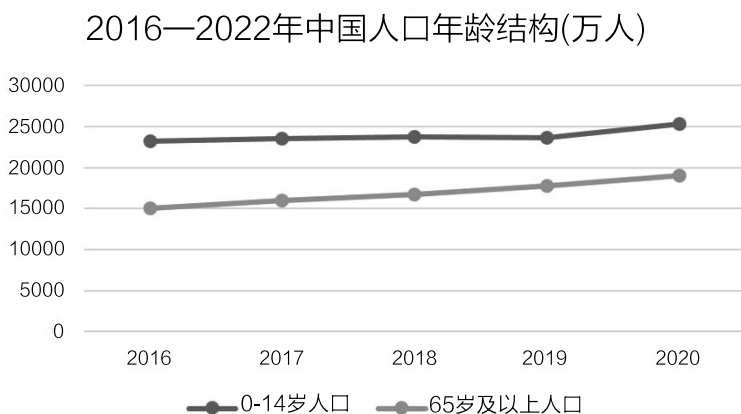


图1-2 2016—2020年中国老龄人口增长趋势

1.3 商业智能分析工具

自助式商业智能分析工具不再只面向IT部门的技术人员，而是面向更多不具备IT背景的业务、财务分析人员。与传统商业智能分析工具相比，自助式商业智能分析工具更灵活，并且更易于使用。下面介绍几种常用的自助式商业智能分析工具。

1. Power BI

Power BI是微软官方推出的可视化数据探索和交互式报告工具。Power BI能让静态数

据转化为动态报表，是让非专业数据分析人员也能有效整合企业数据，并快速准确地进行商业智能分析的数据可视化“神器”。

Power BI应用包括Windows桌面应用程序(Power BI Desktop)、云端在线服务(Power BI Online-Service)和移动端应用(Power BI Mobile)。

2. Tableau

Tableau是一个可视化分析平台。它改变了使用数据解决问题的方式，使个人和组织能够充分利用自己的数据。作为现代商业智能市场的领先产品，Tableau分析平台使人们更加轻松地探索和管理数据，更快地发现和共享各种有价值的见解。

Tableau是斯坦福大学一个计算机科学项目的成果。该项目旨在改善分析流程，并让人们通过可视化技术更轻松地使用数据。自成立以来，Tableau公司一直不断地进行研发投入，开发各种解决方案来帮助需要使用数据的人们更快地找到答案，发现意想不到的见解。Tableau在2019年被Salesforce收购，但其使命不变：帮助人们查看并理解自己的数据。

Tableau以其简单易用、极速高效、视图美观、轻松实现数据融合等优势，帮助人们使用数据推动变革，其家族产品包括Tableau Desktop、Tableau Server、Tableau Online、Tableau Public和Tableau Reader。

3. FineBI

FineBI是帆软软件有限公司推出的一款商业智能产品。业务人员使用该产品可以自主分析企业的信息化数据，帮助企业发现并解决存在的问题，协助企业及时调整策略，做出更好的决策，增强企业的可持续竞争力。FineBI定位于自助式大数据分析工具，能够帮助企业的业务人员和数据分析师开展以问题为导向的探索式分析。

FineBI产品的优势在于：业务人员和数据分析师可以自主制作仪表板，进行探索分析；数据取自业务，并应用于业务，让需要分析数据的人可以自主处理、分析数据。

FineBI的系统构架包括4个部分。

(1) 数据处理。数据处理服务用来对原始数据进行抽取、转换、加载，并为分析服务生成数据仓库FineCube。

(2) 即时分析。业务人员和数据分析师可以快速创建图表使数据可视化，还可以添加过滤条件筛选数据并即时排序，使数据分析更加快捷。

(3) 多维度分析。FineBI可以提供各种分析挖掘功能和预警功能，如任意维度切换、添加，多层钻取、排序、自定义分组、智能关联等。

(4) 仪表盘(dashboard)。仪表盘可以提供各种样式的图表服务，配合各种业务需求展现数据。

4. Smartbi

Smartbi是思迈特软件公司旗下的产品，可以满足客户对企业级报表进行数据可视化分析、自助分析、数据挖掘建模、AI智能分析等方面的需求。Smartbi软件在国内商业智能领域处于领先地位，产品广泛应用于金融、制造、零售、地产等众多行业。

Smartbi产品系列主要包括四大平台。

(1) 大数据分析平台。大数据分析平台可以对接各种业务数据库、数据仓库和其他大数据分析平台，对数据进行加工处理、分析挖掘和可视化展现，满足客户的各种数据分析应用需求，如可视化分析、探索式分析、复杂报表、应用分享等。

(2) 数据化运营平台。数据化运营平台可以为业务人员提供企业级数据分析工具和服务，满足不同类型业务客户的需求，还可以在Excel或者浏览器中实现全自助的数据提取、数据处理、数据分析和数据共享服务，具有很强的实用性。

(3) 大数据挖掘平台。通过深度数据建模，大数据挖掘平台可以为企业提供预测能力支持、文本分析、五大类算法和数据预处理功能，并为客户提供流程式建模、拖曳式操作和可视化配置体验等一站式服务。

(4) SaaS分析云平台。SaaS分析云平台是全新一代云端数据分析平台，可以提供快速搭建数据分析应用的自助式服务，还可以分享深刻见解，提升团队的决策能力。

复习与思考

请从Power BI、Tableau、FineBI、Smartbi官方网站上查找一个商业智能应用的典型案例，仔细研究该案例，并将研究成果制作成PPT与大家分享。

建议：可以从项目背景、项目目标、技术构架、建设方案、项目价值等方面制作PPT，读者也可根据研究案例的特点自行设计。

参考文献

- [1] Jean Piaget. Intellectual evolution from adolescence to adulthood[J]. Human Development, 1972(15): 40-47.
- [2] 任磊, 杜一, 马帅, 等. 大数据可视分析综述[J]. 软件学报, 2014(9): 1909-1936.
- [3] 王志权. 大数据时代与企业财务管理转型[J]. 财务与会计, 2014(6): 74-75.
- [4] 艾达. 数据产品设计[M]. 北京: 电子工业出版社, 2017.
- [5] 朱锴剑. 大数据下企业财务数据可视化的应用现状与趋势探析[J]. 中国总会计师, 2023, (05): 114-116.
- [6] 陈虎, 朱子凝. 数据可视化的财务应用研究[J]. 财会月刊, 2022(16): 120-125.

第2章 数据可视化基础

【案例导入】

从数据可视化看中国的脱贫之路

在改革开放的初期，中国曾是世界上贫困人口最集中的国家。2020年，中国已实现现行标准下农村贫困人口的全面脱贫。世界上没有哪个国家能在如此短的时间内帮助那么庞大的人群摆脱贫困。

1978年，按当时每人每年收入100元的农村贫困标准，中国有2.5亿人口生活在这一水平之下。

到2008年，中国贫困标准提升至每人每年1196元，比1978年翻了近11倍，贫困人口数量减至4007万人。中国贫困标准变化如图2-1所示。

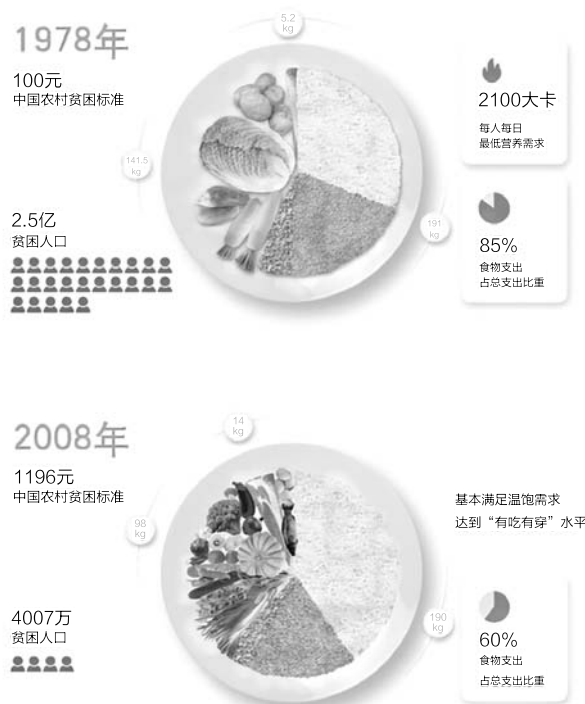


图2-1 中国贫困标准变化图

在我国，对“贫困”的衡量标准是每人每天摄入2100大卡热量。2100大卡换算成一篮子食物，现在大约可以包含1斤粮食、1斤蔬菜、1两肉或1个鸡蛋。1978年，广大农民的餐桌上多是粗粮，鲜有肉蛋，食物支出占总支出比重85%以上。到了2008年，贫困人口饮食结构发生明显改变，可以实现“有吃、有穿”，基本满足温饱，食物支出占比降至60%左右。



思政拓展

(资料来源：国家乡村振兴局. 数据可视化：摆脱贫困，中国这样走过[EB/OL]. https://www.nrra.gov.cn/art/2021/2/24/art_624_187378.html, 2021-02-23.)

2.1 数据类型与财务大数据

大数据本身是一个比较抽象的概念，单从字面来看，它表明数据规模庞大。麦肯锡在其报告《大数据：创新、竞争和生产力的下一个前沿》(*Big data: The next frontier for innovation, competition and productivity*)中给出的大数据定义是：大数据指的是无法通过传统的数据库工具获取、存储、管理和分析的大小超出常规的数据集。它同时强调，并不是说一定要超过特定TB值的数据集才能算是大数据。维基百科对“大数据”的解读是：大数据(big data)，或称巨量数据、海量数据、大资料，指的是所涉及的数据量规模巨大到无法通过人工在合理时间内截取、管理、处理，并整理成为人类所能解读的信息。

大数据具有4V特征：规模性(volume)、多样性(variety)、高速性(velocity)、价值性(value)。随着信息化技术的高速发展，数据开始爆发性增长。大数据的规模性指的是人们对大数据中的数据不再以GB或TB为单位来衡量，而是以PB(1000个T)、EB(100万个T)或ZB(10亿个T)为单位计量。大数据的多样性则体现在数据来源多、数据类型多和数据之间关联性强这三个方面。大数据的高速性是大数据区别于传统数据挖掘最显著的特征，一方面，大数据的数据规模更大；另一方面，大数据对处理数据的响应速度有更严格的要求，包括实时分析而非批量分析，数据输入、处理与丢弃即刻见效，几乎无延迟等。数据的增长速度和处理速度是大数据高速性的重要体现。大数据的价值性体现在其能够从海量、多样且看似互不相关的数据中，挖掘出对未来趋势与模式预测分析有价值的信息。这一过程借助了机器学习方法、人工智能方法或数据挖掘方法进行深度分析。随后，将这些分析结论应用于农业、金融、医疗等多个领域，从而创造出更大的价值。例如，美国社交网站Facebook有10亿用户，网站对这些用户信息进行分析后，广告商可根据分析结果精准投放广告。对广告商而言，10亿用户的数据价值上千亿美元。

2.1.1 数据类型

数据是一种对客观事物的逻辑归纳，是事实或观察的结果。随着科学技术的发展，凡是可以电子化记录的都是数据，如社交网络产生的社交数据、购物网站产生的大量客户及购物数据、物联网技术催生的车联网数据，等等。数据的内涵越来越广泛，不仅包括像GDP(国内生产总值)、股市指数、人口数量等数值型数据，还包括文本、声音、图像、视

频等非数值型数据。数据类型有三种常见的分类方法，分别是按结构属性分类、按连续特征分类与按测量尺度分类。

1. 按结构属性分类

按结构属性分类，数据可分为结构化数据与非结构化数据。它们不仅存储形式不同，在数据处理和数据分析的方法上也大相径庭。

(1) 结构化数据通常是指存储在数据库里，可以用二维表结构来表示的数据。从数据存储角度看，Excel表格数据、SQL Server数据库和Oracle数据库中的数据，都是结构化数据；从应用的角度看，企业ERP系统数据、企业会计信息系统数据、银行交易记录数据等，也是结构化数据，它们大多存储在大型数据库中，用户可以方便地检索、分析和处理。

(2) 非结构化数据通常是指不能用二维表结构来表示和存储的数据。相对于结构化数据而言，非结构化数据没有统一的规则，涉及音(视)频、图片、文本等形式。例如，利用一定手段从网站抓取的新闻数据、某个电影的评价数据等，都需要通过一定的方法，将这些数据量化为结构化数据，才能进行有效的分析。

2. 按连续特征分类

按连续特征分类，数据可分为连续型数据与离散型数据。连续型数据与离散型数据的区别，可以用线、点来区分理解。

(1) 连续型数据是指在一定区间内可以连续取值的数据，如人的身高、体重数值，气温度数，电影票房收益等。

(2) 离散型数据也被称为不连续数据，其取值只能用自然数或整数表达，如硬币的正反面取值、某人的学历取值等。

3. 按测量尺度分类

按测量尺度分类，数据可分为定类数据、定序数据、定距数据和定比数据4类。

(1) 定类数据表现为类别，用于标识数据所描述的主体对象的类别或者属性名称。定类数据只能用来标识事物类别或名称，不区分顺序，无法描述大小、高度、重量等信息，不能进行任何运算，包括比较运算。比如，人的性别分为男性和女性两类，量化后可分别用0和1表示；企业按行业分类，分为旅游业、教育业、制造业、建筑业、金融业等，分别用数字1, 2, 3, 4, 5表示。这些数字只是代号，不能区分大小或进行任何数学运算。

(2) 定序数据表现为类别，但有顺序，也称为序列数据，用于对事物所具有的属性按顺序进行描述。定序数据虽然可以用数字或者序号来排列，但并不代表数据的大小，只代表数据之间的顺序关系。例如，人的受教育程度分为高中毕业、大学本科毕业、硕士研究生毕业、博士研究生毕业，分别用1, 2, 3, 4表示，这些只代表顺序，按照大小正序排列，但不能进行计算。定序数据不仅具有定类数据的特点，可以将所有的数据按照互斥穷尽原则(MECE原则)加以分类，而且各类型之间具有某种意义上的等级差异，从而形成一种确定的排序。

(3) 定距数据是由定距尺度计量形成的, 表现为数值, 可以进行加减运算, 但不能进行乘除运算。定距数据没有绝对零点, 比如温度计的零点是人为指定的, 并不能说 20°C 就是 10°C 的两倍, 但可以说 20°C 比 10°C 高 10°C 。

(4) 定比数据是由定比尺度计量形成的, 表现为数值, 既可以进行加减运算, 也可以进行乘除运算。定比数据代表数据的最高级, 既有测量单位, 也有绝对零点(可以取值为0)。比如, 小明的体重是60千克, 小刚的体重是30千克, 我们可以说小明的体重是小刚体重的2倍。

由此可见, 定类数据和定序数据表现为类别, 属于定性数据; 定距数据和定比数据表现为数值, 属于定量数据。

2.1.2 财务大数据

关于企业财务大数据的界定, 学术界各持不同的观点。本书作者认为, 企业的财务大数据主要包括企业内部数据和企业外部数据两种。

(1) 企业内部数据。企业内部与财务相关的大数据主要来自ERP系统或会计信息系统中的财务、业务数据, 例如用友U8、金蝶K3、SAP等系统中的数据, 以及存储在Access、SQL Server、Oracle等数据库中的数据。在做数据分析时, 我们需要将这些数据从信息系统中导出。

(2) 企业外部数据。企业外部与财务相关的大数据主要包括政策法规文件、行业数据、客户(供应商)数据、国家统计数据等。

2.2 常见的数据分析模型

我们在做数据分析时, 会用到许多分析模型。常见的数据分析模型有以下几种: 用于企业战略分析的SWOT分析模型; 用于外部宏观环境分析的PEST分析模型; 用于外部微观环境分析的波特五力模型; 用于厘清业务问题思路的逻辑树模型; 用于市场营销的4P模型。

2.2.1 SWOT分析模型

SWOT分析法也叫态势分析法, 20世纪80年代初由美国旧金山大学的管理学教授韦里克提出, 经常被用于企业战略制定、竞争对手分析等场合。

SWOT分析模型是产业研究中最常用的分析工具之一, 是对企业内部的优势(strength)与劣势(weakness)、外部环境的机会(opportunity)与威胁(threat)进行综合分析, 并结合企业的经营目标对备选战略方案做出系统评价, 最终制定出一种正确的经营战略, 如图2-2所示。



图2-2 SWOT分析模型

2.2.2 PEST分析模型

PEST 分析模型是战略咨询顾问用来帮助企业分析其外部宏观环境的一种方法。宏观环境又称一般环境，是指影响一切行业和各种企业的各种宏观力量。不同行业和企业会根据自身特点和经营需要对宏观环境因素进行分析，虽然分析的具体内容会有差异，但一般都应对政治(politics)、经济(economic)、社会(society)和技术(technology)这四大类影响企业的主要外部环境因素进行分析。

(1) 政治环境。政治环境主要包括一个国家的社会制度，执政党的性质，政府的方针、政策、法令等。

(2) 经济环境。经济环境主要包括宏观和微观两个方面的内容。宏观经济环境主要指一个国家的人口数量及其增长趋势，国民收入、国内生产总值及其变化情况，以及通过这些指标能够反映的国民经济发展水平和发展速度。微观经济环境主要指企业所在地区或服务地区的消费者的收入水平、消费偏好、储蓄情况、就业程度等因素。这些因素直接决定着企业目前及未来的市场规模。

(3) 社会环境。社会环境包括一个国家或地区的居民受教育程度和文化水平、宗教信仰、风俗习惯、价值观念等。

(4) 技术环境。技术环境除了包括与企业直接相关的技术手段的发展变化外，还包括：①国家对科技开发的投资和支持重点；②该领域技术发展动态和研究开发费用总额；③技术转移和技术商品化速度；④专利及其保护情况等。

2.2.3 波特五力模型

波特五力模型由美国管理学家迈克尔·波特(Michael Porter)于20世纪80年代初提出，是企业制定竞争战略时经常利用的战略分析工具。波特五力分析属于外部环境分析中的微

观环境分析，主要用来分析本行业的企业竞争格局，以及本行业与其他行业之间的关系。

根据波特的观点，一个行业中的竞争，不只是在原有竞争对手之间进行，而是存在5种基本的竞争力量(供应商的议价能力、购买者的议价能力、潜在进入者的威胁、替代品的威胁、行业内现有竞争者的竞争能力)，这5种竞争力量的状况及综合强度决定着行业的竞争激烈程度。5种力量的不同组合变化，最终影响行业利润潜力的变化。波特五力模型如图2-3所示。

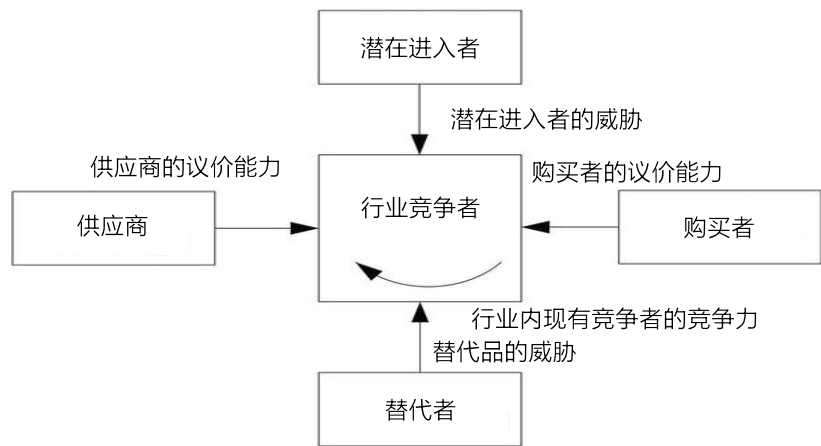


图2-3 波特五力模型

2.2.4 4P模型

4P模型产生于20世纪60年代的美国，是随着营销组合理论的提出而出现的。营销组合实际上有几十个要素，杰罗姆·麦卡锡于1960年在其《基础营销》一书中将这些要素概括为4类：产品(product)、定价(price)、渠道(place)、推广(promotion)，即著名的4P模型(见图2-4)。1967年，菲利普·科特勒在其畅销书《营销管理：分析、规划与控制》中进一步确认了以4P为核心的营销组合方法，具体分析如下。

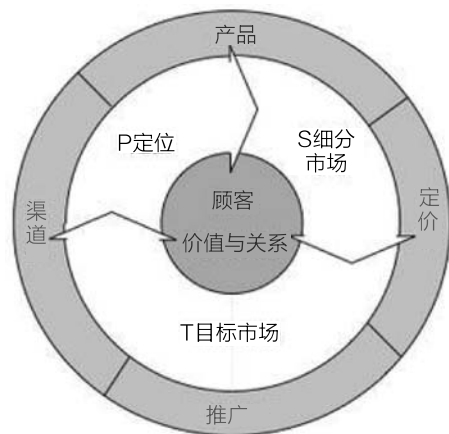


图2-4 4P模型

(1) 产品：注重产品开发，要求产品有独特的卖点，把产品的功能诉求放在第一位。

(2) 定价：根据市场定位，制定价格策略。

(3) 渠道：企业并不直接面对消费者，而是注重经销商的培育和销售网络的建立，企业与消费者的联系是通过分销商来进行的。

(4) 推广：企业注重通过销售行为的改变来刺激消费者，以短期行为(如打折、买一送一等)吸引其他品牌的消费者或促进提前消费，从而获得销售的增长。

2.3 数据可视化的主要内容

可视化主要是借助图形化手段，清晰、有效地传达与沟通信息。数据可视化技术是信息可视化的一个重要分支，它将海量数据转化为图形的形式，让数据更加直观、易懂。数据可视化技术基于大数据、人机交互、计算机图形学等学科，涉及信息可视化、可视分析、虚拟现实等领域。

2.3.1 数据可视化概述

1. 数据可视化的定义

数据可视化是关于数据视觉表现形式的科学技术。

人类从外界获取的信息中，有83%来自视觉，11%来自听觉，6%来自其他。由此可见，视觉是获取信息最重要的通道，超过50%的人脑功能用于视觉的感知，包括解码可视信息、高层次可视信息处理和思考可视符号。可视化是一种映射，可以把客观世界的信息映射为易于被人类感知的视觉模式。这里的视觉模式指的是能够被感知的图形、符号、颜色、纹理等。数据可视化就是将工作中处理的各类数据映射为视觉模式，来探索、解释隐藏在数据背后的信息，在保证信息传递的基础上寻求美感，用数据讲“故事”。因此，数据可视化既是一门科学，又是一门艺术。

2. 数据可视化的过程

(1) 数据获取和格式化：数据可视化技术的第一步是从原始数据源中获取数据，并将其进行格式化和清洗，以便后续的数据处理和分析。

(2) 数据分析和处理：在数据可视化技术中，数据分析和处理是非常重要的一步。通过分析数据可以得出数据的规律、趋势、偏差等，有助于数据决策和预测。

(3) 数据可视化：数据可视化是将数据通过图形形式展示出来的过程。不同的可视化形式可以呈现不同类型的数据信息，如柱状图、饼图、散点图、直方图等。

3. 数据可视化的作用

数据可视化的作用包括数据表达、数据操作和数据分析。

(1) 数据表达是数据可视化最原始的作用。数据表达常见的形式有文本、图表、图像、地图等。有些时候，用可视化形式比文字形式表达更直观，更易于理解。借助有效的图形，可以在较小的空间中呈现大规模的数据。

(2) 数据操作是以计算机提供的界面、接口等为基础来满足人与数据的交互需求的。

当前基于可视化的人机交互技术发展迅速,包括自然交互、可触摸、自适应界面和情景感应等在内的多种新技术极大地丰富了数据操作的方式。

(3) 数据分析的任务通常包括定位、识别、区分、分类、聚类、分布、排列、比较和关联等。将信息以可视化的方式呈献给客户,可以直接提升客户对信息认知的效率,并引导客户从可视化结果中分析、推理出有效信息,帮助人们挖掘数据背后隐藏的信息与客观规律,有助于知识和信息的传播。

2.3.2 数据可视化的基本原理

了解人类视觉的工作原理可以帮助我们更好地理解可视化技术和设计可视化图表。

1. 格式塔原理

格式塔原理产生于1912年,由三位德国心理学家创立,旨在解释人类视觉的工作原理。由他们组成的研究小组观察了许多重要的视觉现象,其中最基础的发现是人类视觉是整体的:我们的视觉系统自动对视觉输入构建结构,并在神经系统层面上感知形状、图形和物体,而非只看到互不相连的边、线和区域。

格式塔原理可以帮助我们理解如何通过视觉认识周围世界的规则,解释在呈现图形元素时人类有组织感知的模式和对象。

格式塔原理由一个基本点、两个假设和5个原则组成。一个基本点是指人类的视觉是整体的。两个假设分别是捆绑假设和关联假设。捆绑假设认为每个复合体都是由基本内容和片段组成的;关联假设则认为,如果任意对象或场景频繁与另一对象或场景一同出现,那么人们通常倾向于在其中一个对象或场景出现时召唤另一个。5个原则分别为相似性、闭合性、接近性、连续性、图形与背景关系原则,具体如下。

- (1) 相似性:如果其他因素相同,那么相似的元素看起来归属同一组。
- (2) 闭合性:如果元素属于封闭图形的一部分,视觉系统通常自动将其感知为一个整体。
- (3) 接近性:距离相近的元素通常被认为属于同一组。
- (4) 连续性:如果定向的单元和组是相互连接在一起的,视觉上通常视为一个整体。
- (5) 图形与背景关系:元素被视为图形(视觉焦点)或背景(图像中的背景)。

2. 视觉编码

视觉编码 (visual coding)是指在个体接收外界信息时,对外界信息的视觉刺激进行编码,如:对颜色、数字、字母、图形等视觉刺激的信息进行编码。

1967年,雅克·贝尔丁(Jacques Bertin)出版的《图形符号学》(*Semiology of Graphics*)一书提出了图形符号与信息的关系,奠定了可视化编码的理论基础。Bertin 视觉编码也叫作 Bertin 视觉变量,包括7个变量,分别是位置、大小、形状、数值、色相、方向和纹理,如图2-5所示。

- (1) 位置(position):根据位置的不同,判断趋势和群组。
- (2) 大小(size):通过图形的大小,反映某一变量取值的大小。
- (3) 形状(shape):形状不同,表示不同的分组。
- (4) 数值(value):指颜色的饱和度数值。同一颜色的饱和度不同,则变量的取值

不同。

- (5) 色相(color): 也指色调, 即颜色。颜色不同, 分组不同。
- (6) 方向(orientation): 表示趋势的不同, 也可用作比较。
- (7) 纹理(texture): 表示分组的不同, 与色相类似。

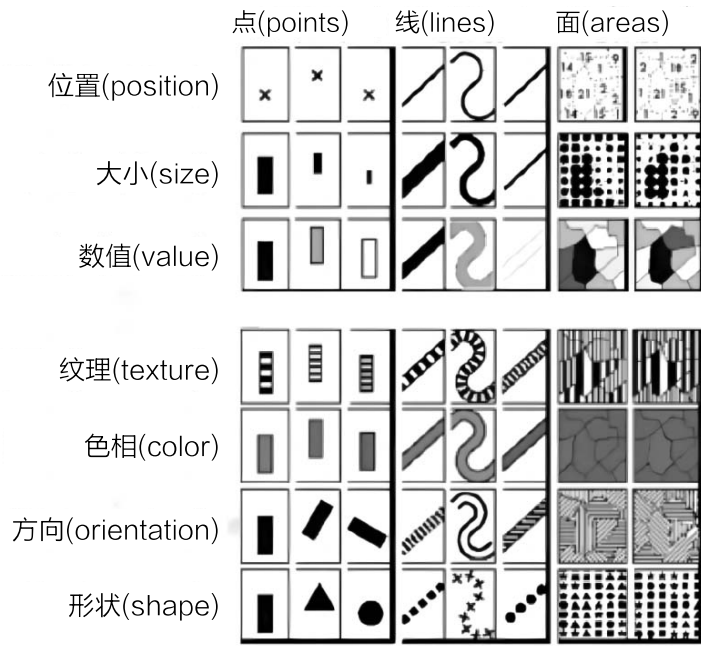


图2-5 Bertin视觉编码

2.3.3 数据可视化的运用场景

数据可视化技术在现代社会中已经被广泛应用在各种领域中。下面列举数据可视化技术的几个应用场景。

- (1) 商业决策: 在商业领域, 数据可视化技术被广泛应用于数据分析和商业决策。通过可视化的方式显示商业数据, 有助于管理者更好地理解业务状况和做出合理的决策。
- (2) 健康医疗: 在健康医疗领域, 数据可视化技术可以被用于监测和分析人体健康状况, 通过可视化的方式可以更好地展示人体的健康信息, 有助于医生更好地理解患者的病情并采取相应的治疗措施。
- (3) 社交网络: 在社交网络领域, 数据可视化技术可以被用于分析用户的行为模式、发现用户的兴趣爱好, 由此进行个性化推荐、朋友圈分析等。
- (4) 城市规划: 在城市规划领域, 数据可视化技术可以被用于分析城市中的流量、污染等数据, 从而指导城市规划和设计, 推动城市的可持续性发展。
- (5) 智能交通: 在智能交通领域, 数据可视化技术可以被用于辅助交通路线规划、拥堵状况监测、交通事故预测等, 提高交通的安全性和便利性。

2.4 常见的可视化图表

在进行数据可视化分析时，我们能选择的图表类型非常多，能否正确地选择图表类型对信息表达有很大的影响。在做图表之前需要了解每个图表的特性，并思考两个问题：

- (1) 数据想要展示什么？
- (2) 想要表达什么？

通常来讲，数据的展示分为比较、序列、描述、构成4种。随着数据可视化的发展及基础图形的扩充，Smartbi专家团整理了一份“图形选择决策树”，可供读者参考，如图2-6所示。

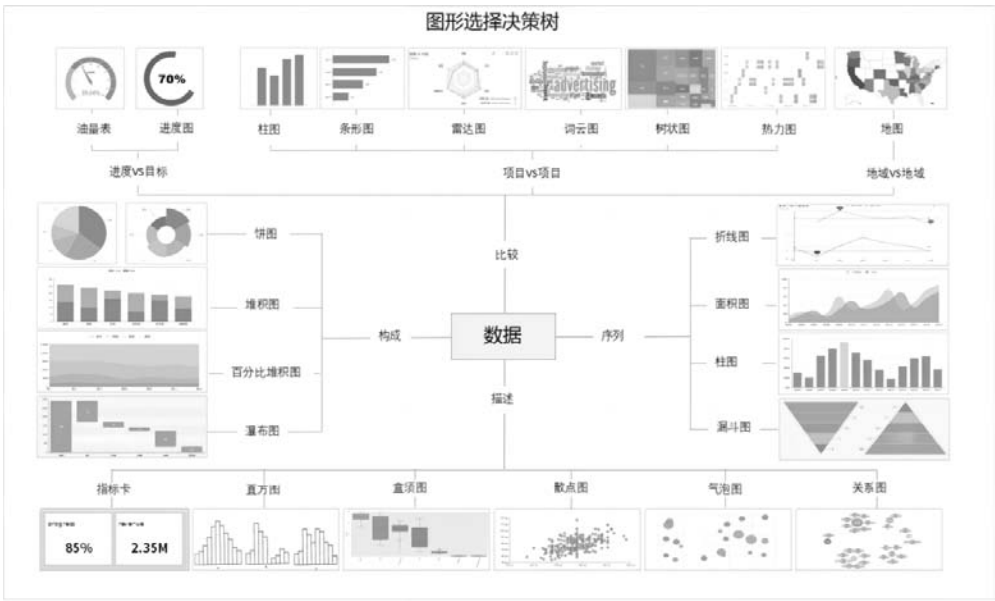


图2-6 图形选择决策树

在进行数据可视化时，用户究竟选择哪种图表更能呈现数据背后的含义，并精准传递数据信息，根据实际运用效果，总结出表2-1，以供选择时参考。

表2-1 图表选择参考表

分类	子分类	图表	解释
比较	实际值与目标值对比	仪表图、码表图	实际值与目标值的比较，关注目标值的完成情况
		进度图	实际值相对于目标值的占比情况
	项目与项目对比	柱形图	适合1~2个维度数据的比较(数据不多时适用)
		条形图	适合1~2个维度数据的比较(数据较多时适用)
		雷达图	适合3个或更多维度数据的比较
		词云图	提取高频文本，过滤低频文本
		树状图	用矩形大小比较同维度下不同的数据
		热力图	通过颜色深浅来表示两个维度数据的大小
	地域与地域对比	地图	不同地域间的数据比较，点越大，数值越大

续表

分类	子分类	图表	解释
序列	连续、有序类别的数据波动(趋势)	折线图 面积图 柱形图	显示随时间变化的数据；折线图和面积图可以用于展示多个维度的变化数据
	各个阶段递减过程	漏斗图	将数据自上而下分为几个阶段，每个阶段的数据都是整体的一部分
描述	关键指标	卡片图	突出显示关键数据
	数据分组差异	直方图	将数据根据差异进行分类
	数据分散	箱线图	展示数据的分散情况(最小值、中位数、最大值等)
	数据相关性	散点图 气泡图	识别变量之间的关系
	人或事物之间的关系	关系图	表示人或事物之间的关系
构成	占比	饼图 环形图 南丁格尔玫瑰图	展示某一维度下不同数据的占比情况
	多类别部分到整体	堆积图 百分比堆积图	展现多个维度下某一维度不同数据的部分和整体情况
	各成分分布情况	瀑布图	表达最后一个数据点的数据演变过程

复习与思考

1. 阐述格式塔原理的一个基本点、两个假设和五个原则各是什么。
2. Bertin 视觉编码或Bertin 视觉变量包括哪些内容？
3. 选择图表进行数据可视化时应考虑哪些因素？
4. 数据类型的分类方法有哪些？如何界定财务大数据？

参考文献

[1] 程佳军, 游宏梁, 汤珊红, 高强, 魏笑, 夏新月. 数据可视化技术在军事数据分析中的应用研究[J]. 情报理论与实践, 2020, 43(09): 171-175.

[2] Lauren Magnuson, Mitchell Data. Visualization: A Guide to Visual Storytelling for Libraries[J]. Technical Services Quarterly, 2017.

[3] 陈红倩. 数据可视化与领域应用案例[M]. 北京: 机械工业出版社, 2019.

[4] 陈虎, 朱子凝. 数据可视化的财务应用研究[J]. 财会月刊, 2022(16): 120-125.

[5] 张超, 肖聪, 朱卫东, 等. 财务智能可视化分析与文献综述[J]. 财会月刊, 2019(03): 24-32.