

高等教育规划教材



免费提供
电子课件、课程设计、实验总结

BIG DATA VISUALIZATION

大数据 可视化

王文 周苏 主编



机械工业出版社
CHINA MACHINE PRESS

高等教育规划教材

大数据可视化

王 文 周 苏 主编



机械工业出版社

大数据可视化是一门理论性和实践性都很强的课程。本书针对计算机、信息管理、经济管理和其他相关专业学生的发展需求,系统、全面地介绍大数据可视化的基本知识和技能,详细介绍了数据可视化之美、Excel数据可视化方法与应用、数据引导可视化设计、数据可视化过程、数据可视化组织、Tableau 应用初步、Tableau 数据管理、Tableau 可视化分析、Tableau 仪表盘与故事以及 Tableau 地图分析与发布等内容,共 11 章,各章均配套设计了导读案例、实验与思考等内容,具有较强的系统性、可读性和实用性。

本书为高等院校相关专业“大数据可视化”“数据媒体设计”等课程全新设计编写,是具有丰富实践特色的主教材。还可供有一定实践经验的软件开发人员、管理人员参考和作为继续教育的教材。

本书配套授课电子课件,需要的教师可登录 www.cmpedu.com 免费注册,审核通过后下载,或联系编辑索取(QQ: 2850823885。电话: 010-88379739)。

图书在版编目(CIP)数据

大数据可视化 / 王文, 周苏主编. —北京: 机械工业出版社, 2018.8
高等教育规划教材
ISBN 978-7-111-61514-9

I. ①大… II. ①王… ②周… III. ①数据处理—高等学校—教材
IV. ①TP274

中国版本图书馆 CIP 数据核字(2018)第 267971 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: 郝建伟 责任编辑: 郝建伟

责任校对: 张艳霞 责任印制: 张 博

北京铭成印刷有限公司印刷

2019 年 2 月第 1 版·第 1 次印刷

184mm×260mm·13.75 印张·2 插页·335 千字

0001—3000 册

标准书号: ISBN 978-7-111-61514-9

定价: 45.00 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) 88379833

机工官网: www.cmpbook.com

读者购书热线: (010) 88379649

机工官博: weibo.com/cmp1952

教育服务网: www.cmpedu.com

封面防伪标均为盗版

金书网: www.golden-book.com



图 1-4 萤火虫之路 (<http://quit007.deviantart.com/>)



图 1-13 深圳受大面积雷电影响，某日 18 时至次日 0 时共记录到 9119 次闪电

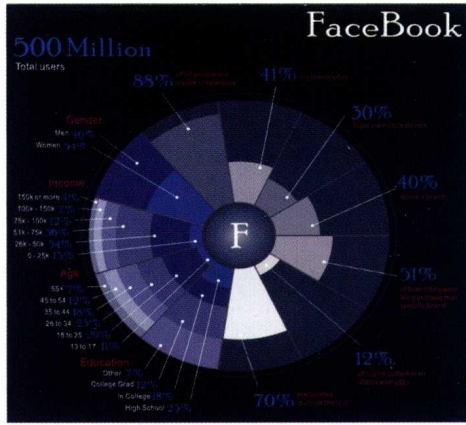


图 1-19 极区图：Facebook vs. 推特

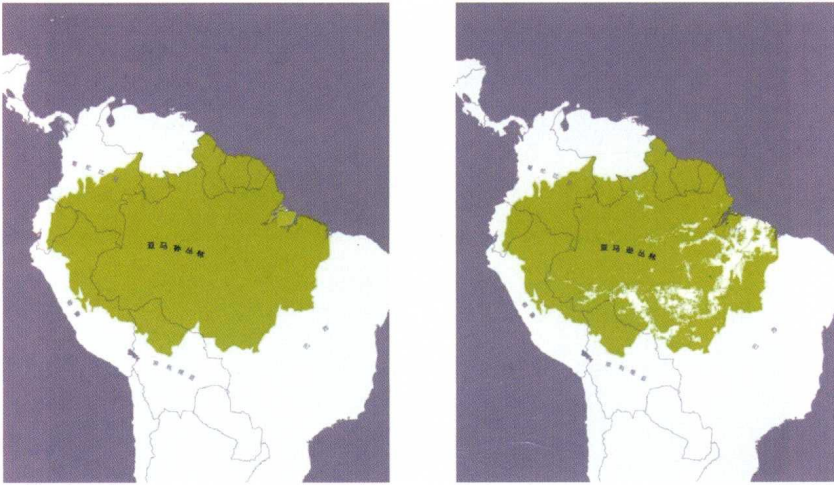


图 2-2 亚马孙丛林 30 年变迁

	A	B	C	D	E
1	系列	系列A	系列B	系列C	系列D
2	店铺A	60%	13%	10%	17%
3	店铺B	49%	24%	16%	11%
4	店铺C	55%	23%	14%	8%

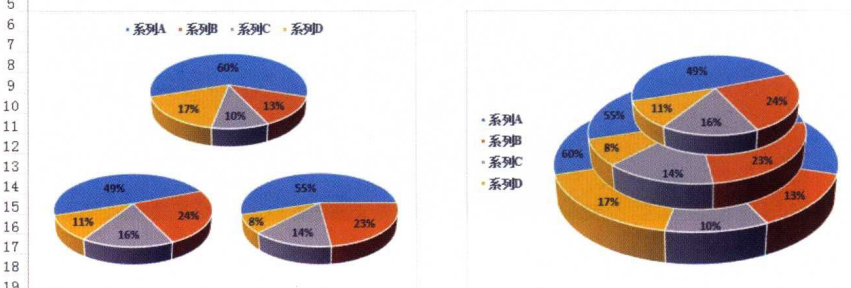


图 3-14 堆叠圆饼图



图 4-14 水循环平面图 (NASA 戈达德航天飞行中心绘制)

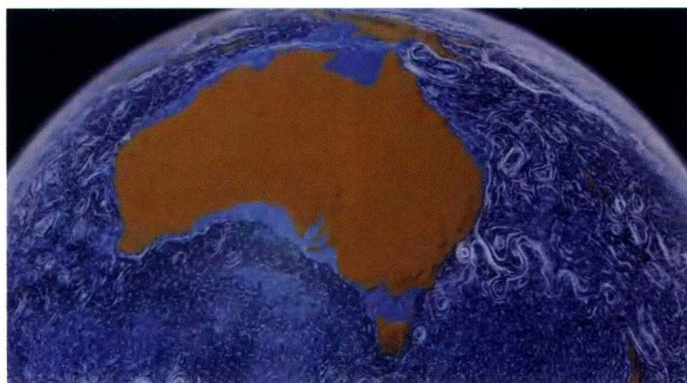


图 4-15 永恒的海洋 (NASA 戈达德航天飞行中心绘制, <http://datafl.ws/2bc>)

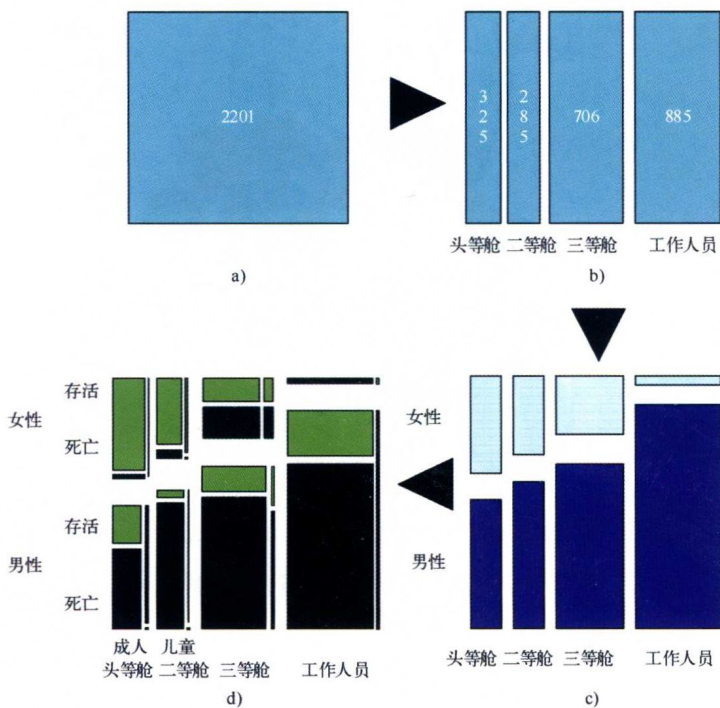


图 5-2 泰坦尼克号事件的镶嵌图生成过程

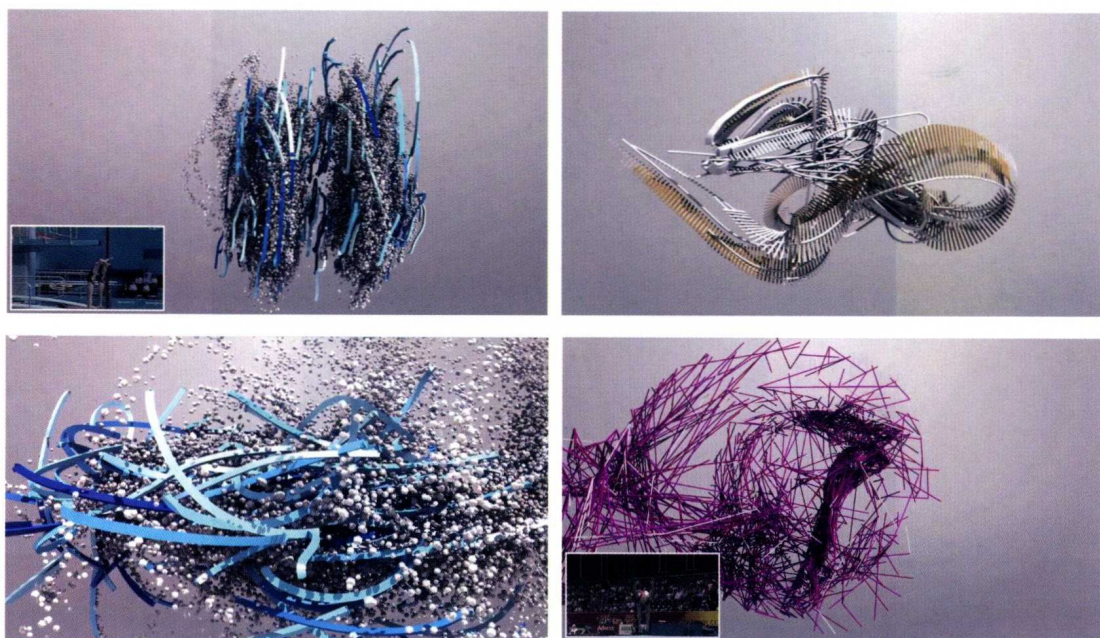


图 4-17 “形态”图（穆罕默德·阿克坦和格约拉）

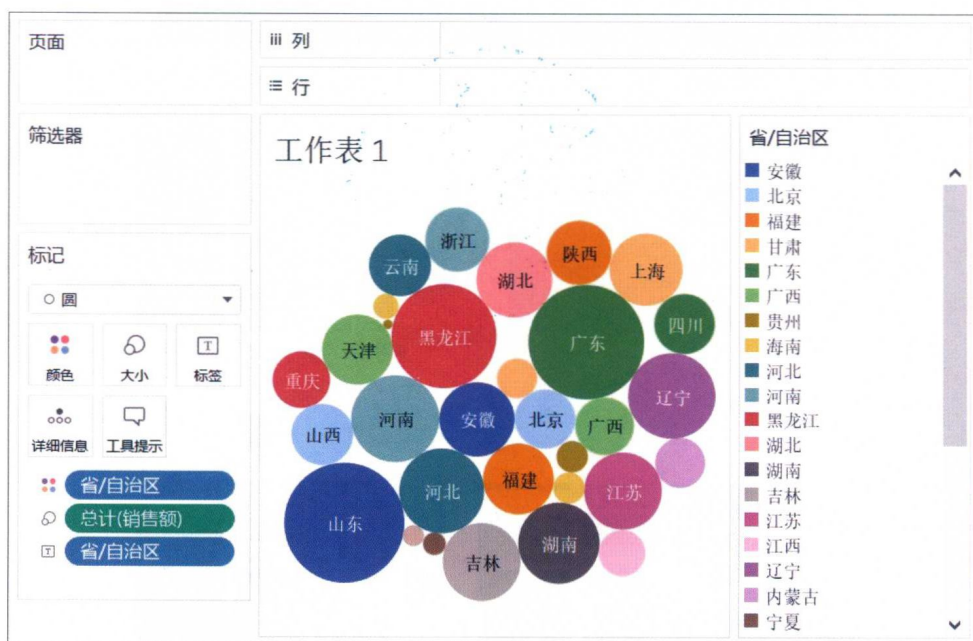


图 9-16 填充气泡图——分省市销售额情况

前 言

大数据 (Big Data) 的力量,正在积极地影响着我们社会生活的方方面面,冲击着各行各业,同时也正在不断地改变我们的学习和日常生活。如今,通过简单、易用的移动应用和基于云端的数据服务,人们能够追踪自己的行为以及饮食习惯,还能提升个人的健康状况。因此,有必要真正理解大数据这个极其重要的议题。

然而,仅有数据是不够的。对于身处大数据时代的企业而言,成功的关键还在于找出大数据所隐含的真知灼见。“以前,人们总说信息就是力量,但如今,对数据进行分析、利用和挖掘才是力量之所在。”

大数据可视化这种新的视觉表达形式是应信息社会蓬勃发展而出现的——因为我们不仅要呈现世界,更重要的是通过呈现来处理更庞大的数据,理解各种各样的数据集合,表现多维数据之间的关联,换句话说,就是归纳数据内在的模式、关联和结构。复杂数据可视化涉及科学、设计与艺术,它的艺术性实际上是使用独特手法展示万千世界的某个局部,从而提出问题。大数据可视化位于科学、设计和艺术等学科的交叉领域,准确地说应该是位于 3 个不同维度的人类活动的交叉领域,蕴藏着无限可能性。

对于在校大学生来说,大数据可视化的理念、技术与应用是一门理论性和实践性都很强的“必修”课程。在长期的教学实践中,我们体会到,坚持“因材施教”的重要原则,把实践环节与理论教学相融合,抓实践教学促进理论知识的学习,是有效地改善教学效果和提高教学水平的重要方法之一。本书的主要特色是:理论联系实际,结合一系列了解和熟悉大数据可视化理念、技术与应用的学习和实践活动,把大数据可视化的相关概念、基础知识和技术技巧融入实践当中,使学生保持浓厚的学习热情,加深对大数据可视化技术的兴趣,认识、理解和掌握大数据可视化技术。

本书为高等院校相关专业,尤其是计算机、信息管理、经济管理类专业开设“大数据”相关课程而全新设计编写,是具有丰富实践特色的主教材。还可供有一定实践经验的 IT 应用人员、管理人员参考和作为继续教育的教材。

本书系统、全面地介绍了大数据可视化的基本知识和应用技能,详细介绍了数据可视化之美、Excel 数据可视化方法与应用、数据引导可视化设计、数据可视化过程、数据可视化组织、Tableau 应用初步、Tableau 数据管理、Tableau 可视化分析、Tableau 仪表盘与故事、Tableau 地图分析与发布以及课程设计与实验总结等内容,共 11 章,具有较强的系统性、可读性和实用性。

结合课堂教学方法改革的要求,全书设计了课程教学过程,为每章教学内容都针对性地安排了导读案例和课后实验与思考等环节,要求和指导学生在课前、课后阅读课文及网络搜索浏览的基础上,深入理解课程知识内涵。

本书提供了《教学进度表》,实际执行时,应按照教学大纲编排教学进度,按照校历考

考虑本学期节假日安排，实际确定本课程的教学进度。

本课程的教学评测可以从以下几个方面入手。

- (1) 每章的导读案例（11次）。
- (2) 每章的实验与思考（11次）。
- (3) 课程设计与实验总结（附录）。
- (4) 结合平时考勤。
- (5) 任课老师认为必要的其他考核方法。

与本书配套的教学 PPT 课件等文档可从机械工业出版社教育服务网（www.cmpedu.com）下载，欢迎教师向作者索取为本书教学配套的相关资料并交流：zhousu@qq.com，QQ: 81505050，个人博客：<http://blog.sina.com.cn/zhousu58>。

本书是浙江安防职业技术学院 2018 年度教材建设项目“数据技术与应用专业系列教材”的建设成果之一。本书的编写得到了浙江安防职业技术学院、浙江商业职业技术学院、新疆大学科学技术学院、浙江大学城市学院等多所院校的支持，张泳、张丽娜、张健、吴林华参与了本书的部分编写工作，在此一并表示感谢！

周 苏

课程教学进度表

(20 —20 学年第 学期)

课程号：_____ 课程名称：_____ 大数据可视化 _____ 学分： 2 周学时： 2
 总学时： 32 其中理论学时（课内）： 32 （课外）实践学时： (24)
 主讲教师： _____

序号	校历周次	章节（或实验、习题课等） 名称与内容	学时	教学方法	课后作业布置
1	1	前言与第1章 数据可视化之美	2	导读案例 课堂教学	
2	2	第1章 数据可视化之美	2		实验与思考1
3	3	第2章 Excel 数据可视化方法	2		实验与思考2
4	4	第3章 Excel 数据可视化应用	2		实验与思考3
5	5	第4章 数据引导可视化设计	2		实验与思考4
6	6	第5章 数据可视化过程	2		实验与思考5
7	7	第6章 数据可视化组织	2		实验与思考6
8	8	第7章 Tableau 应用初步	2		实验与思考7
9	9	第8章 Tableau 数据管理	2		实验与思考8
10	10	第9章 Tableau 可视化分析	2		
11	11	第9章 Tableau 可视化分析	2		实验与思考9
12	12	第10章 Tableau 仪表盘与故事	2		
13	13	第10章 Tableau 仪表盘与故事	2		实验与思考10
14	14	第11章 Tableau 地图分析与发布	2		
15	15	第11章 Tableau 地图分析与发布	2		实验与思考11
16	16	课程设计、学习与实验总结	2		课程设计、学习与实验总结

填表人（签字）：_____

日期：_____

系（教研室）主任（签字）：_____

日期：_____

目 录

前言

课程教学进度表

第 1 章 数据可视化之美	1
【导读案例】南丁格尔“极区图”	1
1.1 数据与可视化	3
1.1.1 数据是什么	3
1.1.2 数据的可变性	3
1.1.3 数据的不确定性	5
1.1.4 数据的背景信息	6
1.1.5 打造最好的可视化效果	7
1.2 数据与图形	7
1.2.1 地图传递信息	7
1.2.2 数据与走势	8
1.2.3 视觉信息的科学解释	9
1.2.4 图片和分享的力量	10
1.2.5 实时可视化	10
1.3 数据可视化的运用	11
1.3.1 数据可视化的方法	12
1.3.2 数据可视化的挑战	12
1.3.3 传统的数据分析图表	13
1.4 可视化分析与编程工具	15
1.4.1 Microsoft Excel	15
1.4.2 Google Spreadsheets	16
1.4.3 Tableau	16
1.4.4 针对特定数据的工具	16
1.4.5 可视化编程工具	17
【实验与思考】熟悉大数据可视化	19
第 2 章 Excel 数据可视化方法	22
【导读案例】亚马孙丛林的变迁	22
2.1 Excel 的函数与图表	24
2.1.1 Excel 函数	25
2.1.2 Excel 图表	26
2.1.3 选择图表类型	27

2.2	整理数据源	29
2.2.1	数据提炼	29
2.2.2	抽样产生随机数据	31
2.3	数理统计中的常见统计量	33
2.3.1	比平均数更稳定的中位数和众数	33
2.3.2	概率统计中的正态分布和偏态分布	34
2.3.3	应用在财务预算中的分析工具	36
2.4	改变数据形式引起的图表变化	37
2.4.1	用负数突出数据的增长情况	37
2.4.2	重排关键字顺序使图表更合适	38
	【实验与思考】体验 Excel 数据可视化方法	39
第 3 章	Excel 数据可视化应用	40
	【导读案例】包罗一切的数字图书馆	40
3.1	直方图：对比关系	44
3.1.1	以零基线为起点	44
3.1.2	垂直直条的宽度要大于条间距	45
3.1.3	慎用三维效果的柱形图	46
3.1.4	用堆积图表示百分数	47
3.2	折线图：按时间或类别显示趋势	48
3.2.1	减小 Y 轴刻度单位增强数据波动情况	48
3.2.2	突出显示折线图中的数据点	49
3.2.3	通过面积图显示数据总额	50
3.3	圆饼图：部分占总体的比例	51
3.3.1	重视圆饼图扇区的位置排序	51
3.3.2	分离圆饼图扇区强调特殊数据	52
3.3.3	用半个圆饼图刻画半期内的数据	53
3.3.4	让多个圆饼图对象重叠展示对比关系	54
3.4	散点图：表示分布状态	54
3.4.1	用平滑线联系散点图增强图形效果	55
3.4.2	将直角坐标改为象限坐标凸显分布效果	55
3.5	侧重点不同的特殊图表	56
3.5.1	用子弹图显示数据的优劣	56
3.5.2	用温度计展示工作进度	58
3.5.3	用漏斗图进行业务流程的差异分析	59
	【实验与思考】熟悉 Excel 数据图表的分析作用	60
第 4 章	数据引导可视化设计	62
	【导读案例】拿破仑东征莫斯科及撤退	62
4.1	可视化对认知的帮助	65
4.1.1	科学可视化	65

4.1.2	七个数据类型	65
4.1.3	七个基本任务	67
4.2	新的数据研究方法	67
4.3	信息图形和展示	69
4.4	走进数据艺术的世界	70
4.5	掌握可视化设计组件	71
4.5.1	视觉隐喻	72
4.5.2	坐标系	75
4.5.3	标尺	76
4.5.4	背景信息	77
4.5.5	整合可视化组件	78
	【实验与思考】大数据可视化的领军企业 Tableau	78
第5章	数据可视化过程	82
	【导读案例】关于泰坦尼克号的“镶嵌图”	82
5.1	分析数据，指导视觉探索	84
5.1.1	你拥有什么数据	85
5.1.2	关于数据，你了解什么	85
5.1.3	应该使用哪种可视化方式	86
5.1.4	你看到了什么，有意义吗	86
5.2	分类数据的可视化	86
5.2.1	整体中的部分	87
5.2.2	子分类	87
5.2.3	看清数据的结构和模式	88
5.3	时序数据的可视化	88
5.3.1	周期	89
5.3.2	循环	90
5.4	空间数据的可视化	91
5.5	让可视化设计更清晰	92
5.5.1	建立视觉层次	92
5.5.2	增强图表的可读性	93
5.5.3	允许数据点之间进行比较	94
5.5.4	描述背景信息	95
	【实验与思考】绘制新的泰坦尼克事件镶嵌图	96
第6章	数据可视化组织	98
	【导读案例】德克萨斯大学体系的透明化	98
6.1	可视化组织的快速发展	100
6.1.1	什么是数据驱动	100
6.1.2	新的互联网环境	101
6.1.3	更好的数据工具	102

6.1.4	更透明的组织	103
6.1.5	竞争新态势：有样学样	103
6.1.6	元数据和源数据	104
6.2	典型的可视化组织——Netflix	104
6.2.1	创办 Netflix	104
6.2.2	Netflix 自我颠覆	105
6.2.3	大数据整合战略的构成	106
6.2.4	Netflix 文化灌输	106
6.3	创业公司的数据可视化	108
6.3.1	Wedgies 的创业	108
6.3.2	用户体验至高无上	109
6.3.3	应用开源工具	109
6.4	可视化组织的四层架构	110
6.5	建立可视化组织	112
6.5.1	数据提示	113
6.5.2	设计提示	114
6.5.3	技术提示	116
6.5.4	管理提示	116
	【实验与思考】建立数据可视化组织	117
第 7 章	Tableau 应用初步	119
	【导读案例】数据分析的五大思维方式	119
7.1	Tableau 概述	122
7.1.1	Tableau 可视化技术	122
7.1.2	Tableau 主要特性	123
7.2	Tableau 产品线	124
7.3	下载、安装与注册	125
7.4	Tableau 工作区	127
7.4.1	工作表工作区	128
7.4.2	仪表盘工作区	129
7.4.3	故事工作区	130
7.4.4	菜单栏和工具栏	130
7.5	Tableau 数据	132
7.5.1	数据角色	133
7.5.2	字段类型	134
7.5.3	文件类型	135
7.6	创建视图	135
7.6.1	行列功能区	136
7.6.2	标记卡	137
7.6.3	筛选器	139

7.6.4	页面	139
7.6.5	智能显示	140
7.6.6	度量名称和度量值	141
7.7	创建仪表板	142
	【实验与思考】熟悉 Tableau 数据可视化设计	143
第 8 章	Tableau 数据管理	145
	【导读案例】Tableau 案例分析：世界指标——人口	145
8.1	Tableau 数据架构	146
8.1.1	数据连接层	146
8.1.2	数据模型层	147
8.2	数据连接	147
8.2.1	连接文件数据源	148
8.2.2	连接服务器数据源	149
8.2.3	组织数据	149
8.2.4	实现多表联接	150
8.3	数据维护	151
8.3.1	刷新数据源	151
8.3.2	关闭数据源	151
	【实验与思考】熟悉 Tableau 数据管理操作	151
第 9 章	Tableau 可视化分析	154
	【导读案例】Tableau 案例分析：世界指标——医疗支出	154
9.1	条形图与直方图	155
9.1.1	条形图与直方图的区别	155
9.1.2	条形图	156
9.1.3	直方图	157
9.2	饼图	159
9.3	折线图	160
9.4	压力图与突显表	161
9.4.1	压力图	161
9.4.2	突显表	162
9.5	树地图	164
9.6	气泡图与圆视图	165
9.6.1	气泡图	165
9.6.2	圆视图	166
9.7	标靶图	167
9.8	甘特图	168
9.9	盒须（箱线）图	169
9.9.1	创建盒须图	170
9.9.2	图形延伸	171

【实验与思考】熟悉 Tableau 数据可视化分析	172
第 10 章 Tableau 仪表板与故事	173
【导读案例】Tableau 案例分析：世界指标——旅游业	173
10.1 创建仪表板	174
10.1.1 创建仪表板	174
10.1.2 向仪表板中添加视图	175
10.1.3 添加仪表板对象	178
10.1.4 从仪表板中移除视图和对象	179
10.1.5 仪表板 Web 视图安全选项	179
10.2 布局容器	180
10.3 组织仪表板	181
10.3.1 平铺和浮动布局	181
10.3.2 显示和隐藏工作表的组成部分	182
10.3.3 重新排列仪表板视图和对象	182
10.3.4 设置仪表板大小	182
10.3.5 了解仪表板和工作表	182
10.4 故事工作区	183
10.4.1 故事工作表	183
10.4.2 创建故事	185
10.4.3 调整标题大小	186
10.4.4 “设置故事格式”窗格	186
10.4.5 更新与演示故事	187
【实验与思考】熟悉 Tableau 仪表板与故事操作	187
第 11 章 Tableau 地图分析与发布	193
【导读案例】Tableau 案例分析：世界指标——商业	193
11.1 分配地理角色	194
11.2 设置地理信息	195
11.3 导出和发布数据（源）	195
11.3.1 通过将数据复制到剪贴板导出数据	196
11.3.2 导出数据源	197
11.3.3 发布数据源	198
11.4 导出图像和 PDF 文件	198
11.4.1 复制图像	199
11.4.2 导出图像	199
11.4.3 打印为 PDF	199
11.5 保存和发布工作簿	200
11.5.1 保存工作簿	200
11.5.2 保存打包工作簿	200
11.5.3 将工作簿发布到服务器	200

11.5.4 将工作簿保存到 Tableau Public 上	201
【实验与思考】熟悉 Tableau 地图分析与发布	201
附录 课程设计与实验总结	203
附录 A 课程设计	203
附录 B 课程学习与实验总结	204
B.1 课程的基本内容	204
B.2 实验的基本评价	206
B.3 课程学习能力测评	206
B.4 大数据可视化学习与实验总结	207
B.5 学习与实验总结评价（教师）	207
参考文献	208