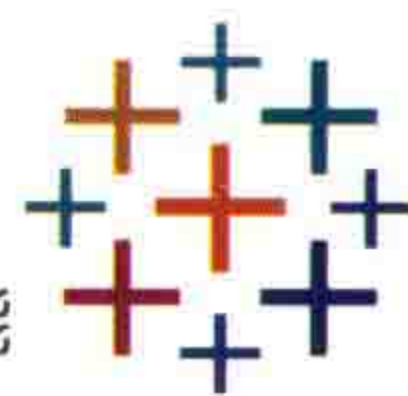


- 本书由专业数据分析师根据新版Tableau 10精心编写
- 详尽介绍Tableau的各项功能及其在数据可视化方面的应用技巧
- 配合丰富案例讲解，特别提供两个热门商业案例分析，快速提高实战技能



—— 邂逅 Tableau 10, 数据依旧, 魅力倍增 ——

Tableau 数据可视化

从入门到精通

王国平 编著

清华大学出版社



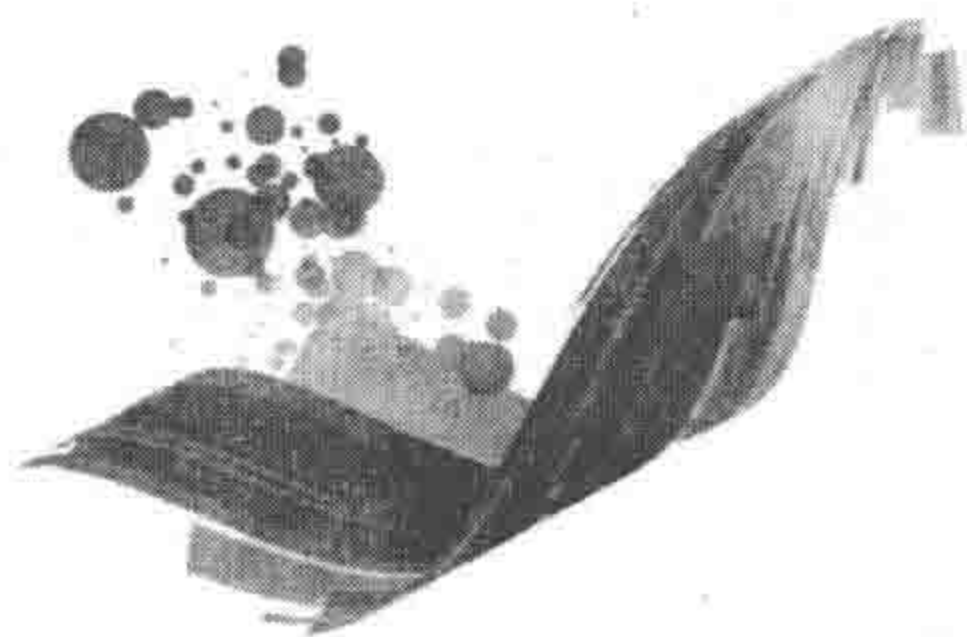


Tableau 数据可视化 从入门到精通

王国平 编著

清华大学出版社

北京

内 容 简 介

本书由专业的数据分析师基于 Tableau 10.1 精心编撰, 详细介绍 Tableau 的数据连接、图形编辑与展示功能, 包括软件的安装与激活、数据类型和运算符、连接数据源、基础操作、数据与图形的导出、连接大数据、基础图表、函数、高级数据操作、地图分析、故事、Tableau Online、Tableau Server 等内容, 并通过网上超市运营、网站流量统计等案例介绍 Tableau 在实际项目中的应用。

本书适合互联网、银行证券、咨询审计、快消品、能源等行业的数据分析师以及媒体、网站等数据可视化从业者使用, 也可供高等院校相关专业的师生及从事数据分析的研究者参考, 还可用作 Tableau 软件的培训教材。

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。侵权举报电话: 010-62782989 13701121933

图书在版编目 (CIP) 数据

Tableau 数据可视化从入门到精通/王国平编著. —北京: 清华大学出版社, 2017
ISBN 978-7-302-47377-0

I. ①T… II. ①王… III. ①可视化软件 IV. ①TP31

中国版本图书馆 CIP 数据核字 (2017) 第 124191 号

责任编辑: 王金柱

封面设计: 王 翔

责任校对: 闫秀华

责任印制: 沈 露

出版发行: 清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦 A 座 邮 编: 100084

社 总 机: 010-62770175 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: 北京国马印刷厂

经 销: 全国新华书店

开 本: 190mm×260mm 印 张: 14.25 字 数: 364 千字

版 次: 2017 年 7 月第 1 版 印 次: 2017 年 7 月第 1 次印刷

印 数: 1~3000

定 价: 49.00 元

前 言

大数据时代正在变革我们的生活、工作和思维。如何让大数据更有意义，使之更贴近大多数人，重要的手段之一就是数据可视化。数据可视化是关于数据视觉表现形式的技术，这种数据的视觉表现形式被定义为一种以某种概要形式抽取出来的信息，包括相应信息单位的各种属性和变量。

Tableau 可以帮助人们将数据转化为可以付诸行动的见解，探索无所不能的可视化分析，只需点击几下即可构建仪表盘，进行即兴分析，与任何人共享自己的工作成果。从全球性企业到早期初创企业和小企业，使用 Tableau 来查看和理解数据的人无处不在，其中既有像联合利华、德勤、UPS、耐克、杜邦、Verizon、T-mobile、BBC、探索频道、美国航空、Zynga、LinkedIn、Facebook、雅虎、苹果、可口可乐等欧美知名企业，也有美国联邦航空管理局、美国陆军等美国政府机构以及康奈尔、杜克、牛津等知名学府，Tableau 在中国市场也有所开拓，中国东方航空是其重要客户。

Tableau 的应用已经遍及商务服务、能源、电信、金融服务、互联网、生命科学、医疗保健、制造业、媒体娱乐、公共部门、教育、零售等各个行业，因此掌握该软件的使用非常必要。

本书的内容

本书基于 Tableau 10.1 编写，详细介绍了 Tableau 的数据连接功能、图形编辑与展示功能。本书共分 16 章，各章内容说明如下：

第 1 章介绍大数据时代的特征和挑战、数据可视化的新特性以及目前主要的可视化软件。此外，还简单介绍了 Tableau 软件，包括 Tableau Desktop、Tableau Online、Tableau Server、Tableau Mobile、Tableau Public 和 Tableau Reader。

第 2 章介绍 Tableau Desktop 的新增功能，以及软件的数据类型、运算符、开始页面、软件的安装与激活。

第 3 章介绍 Tableau Desktop 可以连接的数据源，包括 Excel 文件、文本文件、Access、统计文件等，还将介绍如何连接各类数据库，如 Tableau Server、SQL Server、MySQL、Oracle 等。

第 4 章首先介绍 Tableau Desktop 中维度和度量、连续和离散的概念和操作，然后介绍工作区和工作表，并结合具体的案例进行讲解。

第 5 章介绍 Tableau Desktop 的数据导出，包括数据文件导出、图片文件导出、PDF 文件导出。

第 6 章和第 7 章介绍 Tableau 在大数据方面的应用，主要包括连接基本条件、主要步骤和注意事项等，还将介绍 Tableau 大数据引擎的优化方法，以及如何提升连接性能。

第 8 章介绍如何使用 Tableau 生成一些统计图形，如条形图、饼图、直方图、折线图、散点图、并排图、甘特图等。

第 9 章介绍 Tableau 函数，包括数学函数、字符串函数、日期函数、类型函数、逻辑函数、

聚合函数、直通函数、用户函数、表计算函数等，同时介绍每类函数的用法和例子。

第 10 章介绍 Tableau 的一些高级操作，如表计算、创建字段、创建参数、聚合数据和缺失值处理，并结合具体实例进行讲解。

第 11 章介绍如何使用 Tableau 创建地图，包括设置角色、比较地图、添加字段信息、设置地图选项、创建分布图和自定义地图等。

第 12 章介绍 Tableau 故事的概念，包括如何创建故事、修改故事点、设置格式和演示故事等。

第 13 章介绍 Tableau Online，包括如何导入工作簿、导入数据源、搜索内容、创建用户站点角色、为站点添加用户、导入现有用户、添加和删除项目、为项目添加工作簿等。

第 14 章介绍 Tableau Server，包括安装的系统要求、安装步骤，以及如何配置服务器，包括常规配置、数据连接、通知和订阅、SMAL、OpenID 等。

第 15 章介绍网上超市运营分析案例，主要从客户分析、配送分析、销售分析、利润分析、退货分析和预测分析 6 个方面分析。

第 16 章介绍网站流量分析案例，主要从页面指标分析、访问量分析、浏览量分析、退出量分析和下载量分析 5 个方面分析。

本书的特色

(1) 内容全面，讲解详细

本书详细介绍了 Tableau 大部分常用功能，对于初次学习 Tableau 软件的读者来说帮助较大。书中列出了创建可视化数据分析视图的每一步操作，便于读者的实践练习。

(2) 由浅入深、循序渐进

本书从 Tableau 简介、连接数据源、基础操作、高级操作到 Tableau Online 和 Tableau Server 等，由浅入深，循序渐进，适合各个层次的读者阅读。

(3) 案例丰富，高效学习

本书在介绍功能的同时基本都是结合案例进行操作。同时，为了使读者快速提高数据分析的整体能力，本书的最后两章列举了网上超市运营案例分析和网站流量案例分析两个案例。

本书的读者对象

本书的内容和案例适用于互联网、银行证券、咨询审计、快消品、能源等行业数据分析用户及媒体、网站等数据可视化用户，可供高等院校相关专业学生及从事数据分析的研究者参考使用，也可作为 Tableau 软件培训和自学的教材。

由于编者水平所限，书中难免存在错误和不妥之处，请广大读者批评指正。

编者

2017 年 3 月

目 录

第 1 章	数据可视化及 Tableau 概述	1
1.1	大数据时代的挑战	2
1.2	大数据可视化的难点	2
1.3	可视化技术的新特性	3
1.4	主要的可视化软件	4
1.4.1	Tableau	4
1.4.2	QlikView	4
1.4.3	ECharts	5
1.4.4	HighCharts	5
1.5	Tableau 软件概况	5
1.5.1	Tableau Desktop	6
1.5.2	Tableau Online	7
1.5.3	Tableau Server	8
1.5.4	Tableau Mobile	8
1.5.5	Tableau Public	9
1.5.6	Tableau Reader	9
第 2 章	Tableau Desktop 简介	10
2.1	主要新增功能	11
2.1.1	Tableau 10.1 新增功能	11
2.1.2	Tableau 10.0 新增功能	11
2.1.3	Tableau 9.3 新增功能	12
2.1.4	Tableau 9.2 新增功能	12
2.2	数据类型	13
2.2.1	主要的数据类型	13
2.2.2	更改数据类型	14
2.3	运算符及优先级	15
2.3.1	算术运算符	15
2.3.2	逻辑运算符	15
2.3.3	比较运算符	16
2.3.4	运算符优先级	16
2.4	软件安装	16
2.4.1	软件下载	16
2.4.2	安装步骤	17
2.4.3	软件激活	18
2.5	软件界面简介	19
2.5.1	开始界面	20
2.5.2	数据源界面	21
2.5.3	工作簿界面	22

2.6 文件类型	23
第 3 章 连接数据源	24
3.1 连接到文件	25
3.1.1 Excel 文件	25
3.1.2 文本文件	26
3.1.3 Access	27
3.1.4 JSON 文件	29
3.1.5 统计文件	30
3.1.6 其他文件	31
3.2 连接到数据库	32
3.2.1 Tableau Server	32
3.2.2 SQL Server	32
3.2.3 MySQL	33
3.2.4 Oracle	34
3.2.5 Amazon Redshift	34
3.2.6 更多数据库	35
第 4 章 Tableau 的基础操作	38
4.1 维度和度量	39
4.1.1 维度	39
4.1.2 度量	40
4.2 连续和离散	41
4.2.1 连续字段	41
4.2.2 离散字段	41
4.3 工作区操作	42
4.3.1 “数据”窗格	43
4.3.2 “分析”窗格	44
4.3.3 工具栏	44
4.3.4 状态栏	45
4.3.5 卡和功能区	46
4.3.6 语言和区域设置	46
4.4 工作表操作	47
4.4.1 创建工作表	47
4.4.2 复制工作表	47
4.4.3 导出工作表	48
4.4.4 删除工作表	48
4.5 基础操作案例	49
第 5 章 Tableau 数据导出	52
5.1 数据文件导出	53
5.1.1 导出图形中的数据	53
5.1.2 导出数据源数据	54
5.1.3 导出交叉表数据	55
5.1.4 导出 Access 数据	56

5.2	导出图形文件	57
5.2.1	通过复制导出	57
5.2.2	直接导出图像	59
5.3	导出 PDF 格式文件	60
第 6 章	Tableau 连接到 Hadoop Hive	62
6.1	Hadoop 简介	63
6.1.1	Hadoop 分布式文件系统	63
6.1.2	MapReduce 计算框架	63
6.2	连接基本条件	64
6.2.1	Hive 版本	64
6.2.2	驱动程序	64
6.2.3	启动 Hive 服务	65
6.3	连接主要步骤	65
6.3.1	Cloudera Hadoop	65
6.3.2	Hortonworks Hadoop Hive	66
6.3.3	MapR Hadoop Hive	66
6.4	连接注意事项	67
6.4.1	日期/时间数据	67
6.4.2	已知限制	67
6.5	检验测试连接	68
第 7 章	Tableau 大数据引擎优化	69
7.1	提高连接性能	70
7.2	高级性能技术	72
7.2.1	“筛选器”形式的分区字段	72
7.2.2	分组字段形式的群集字段	72
7.2.3	初始化 SQL	72
7.3	提升数据提取效率	73
7.3.1	存储文件格式	73
7.3.2	数据分区	73
7.3.3	数据分组	74
第 8 章	创建图表	75
8.1	单变量图形	76
8.1.1	条形图	76
8.1.2	饼图	78
8.1.3	直方图	79
8.1.4	折线图	80
8.2	多变量图形	82
8.2.1	散点图	82
8.2.2	甘特图	83
第 9 章	Tableau 函数	85
9.1	数字函数	86

9.2	字符串函数	88
9.3	日期函数	91
9.4	类型转换函数	93
9.5	逻辑函数	94
9.6	聚合函数	96
9.7	直通函数	97
9.8	用户函数	99
9.9	表计算函数	100
9.10	其他函数	105
9.10.1	模式匹配的特定函数	105
9.10.2	Hadoop Hive 的特定函数	106
9.10.3	GoogleBigQuery 的特定函数	107
第 10 章	Tableau 的高级操作	109
10.1	表计算	110
10.2	创建字段	112
10.3	创建参数	113
10.4	聚合计算	118
10.5	缺失值处理	121
10.6	案例：超市利润额变化	122
第 11 章	创建地图	125
11.1	设置角色	126
11.2	标记地图	126
11.3	添加字段信息	127
11.4	设置地图选项	128
11.5	创建分布图	130
11.6	自定义地图	130
第 12 章	故事	131
12.1	故事简介	132
12.2	创建故事	133
12.3	设置故事格式	135
12.3.1	调整标题大小	135
12.3.2	使仪表板适合故事	136
12.3.3	设置故事格式	136
12.4	演示故事	138
第 13 章	Tableau Online	139
13.1	简介	140
13.2	导航	143
13.3	基础操作	147
13.3.1	导入数据源	147
13.3.2	搜索内容	148

13.3.3	内容操作	149
13.3.4	用户设置	151
13.4	用户和组	151
13.4.1	用户站点角色	151
13.4.2	为站点添加用户	152
13.4.3	导入用户	153
13.4.4	创建组	154
13.4.5	管理用户	156
13.5	项目	156
13.5.1	添加项目	157
13.5.2	移动工作簿	157
13.5.3	删除项目	158
第 14 章	Tableau Server	160
14.1	安装系统要求	161
14.2	软件安装步骤	162
14.2.1	准备安装	162
14.2.2	软件安装	163
14.2.3	在线激活	164
14.3	服务器配置	166
14.3.1	常规配置	166
14.3.2	数据连接	167
14.3.3	服务器	167
14.3.4	SMTP 设置	168
14.3.5	通知和订阅	169
14.3.6	SSL	169
14.3.7	SAML	170
14.3.8	Kerberos	171
14.3.9	SAP HANA	172
14.3.10	OpenID	173
14.4	登录服务器	174
第 15 章	网上超市运营分析	175
15.1	客户分析	176
15.1.1	各省市交易次数	176
15.1.2	各省市利润	177
15.1.3	客户散点图	178
15.1.4	客户交易量排名	178
15.2	配送分析	179
15.2.1	各省市配送情况	179
15.2.2	配送准时性	181
15.2.3	商品发货天数	181
15.2.4	配送延迟商品	182
15.3	销售分析	183
15.3.1	各个省市销售额	183

15.3.2	区域销售额	184
15.3.3	产品细分	184
15.3.4	客户细分	185
15.4	利润分析	186
15.4.1	产品利润分析	186
15.4.2	区域利润分析	187
15.4.3	区域经理业绩	188
15.5	退货分析	188
15.5.1	退货区域分布	189
15.5.2	退货产品数量	190
15.5.3	退货产品类型	190
15.5.4	退货产品名称	191
15.6	预测分析	192
15.6.1	销售额预测	192
15.6.2	区域销售预测	193
15.6.3	区域利润预测	194
第 16 章	网站流量统计分析	196
16.1	页面指标分析	197
16.1.1	页面访问次数	197
16.1.2	平均停留时间	198
16.1.3	区域热门页面	198
16.2	访问量分析	199
16.2.1	访问量地图	199
16.2.2	各省市访问量	201
16.2.3	访问量趋势	201
16.2.4	访问量树地图	202
16.3	浏览量分析	203
16.3.1	浏览量地图	203
16.3.2	各个省市的浏览量	204
16.3.3	区域浏览量	205
16.3.4	浏览量趋势	205
16.4	退出量分析	206
16.4.1	退出量地图	206
16.4.2	各省市退出量	208
16.4.3	区域退出量	208
16.4.4	退出量趋势	209
16.5	下载量分析	210
16.5.1	下载量地图	210
16.5.2	各省市下载量	211
16.5.3	区域下载量	212
16.5.4	下载量趋势	212
附录 A	配置 ODBC 数据源	214
	参考文献	218

第 1 章

数据可视化及 Tableau 概述

“让每个人都成为数据分析师”是大数据时代的要求，数据可视化的出现恰恰从侧面缓解了专业数据分析人才的缺乏。Tableau、Qlik、Microsoft、SAS、IBM 等 IT 厂商纷纷加入数据可视化的阵营，在降低数据分析门槛的同时，为分析结果提供更炫的展现效果。为了进一步让大家了解如何选择适合的数据可视化产品，本书将围绕这一话题展开，希望能对正在选型中的个人和企业有所帮助。

数据可视化是技术与艺术的完美结合，它借助图形化的手段，清晰有效地传达与沟通信息。一方面，数据赋予可视化意义；另一方面，可视化增加数据的灵性，两者相辅相成，帮助企业从信息中提取知识、从知识中收获价值。

数据可视化技术允许利用图形、图像处理、计算机视觉以及用户界面，通过表达、建模以及对立体、表面、属性、动画的显示，对数据加以可视化解释。Tableau 数据可视化软件为用户在数据可视化方面提供了行之有效的方法，重视的人越来越多。

1.1 大数据时代的挑战

大数据的出现正在引发全球范围内技术与商业变革的深刻变化。在技术领域，以往更多依靠模型的方法，现在可以借用规模庞大的数据，用基于统计的方法，使语音识别、机器翻译这些技术领域在大数据时代取得新进展。

既有技术架构和路线已经无法高效处理如此海量的数据。对于相关组织来说，如果投入巨大而采集的信息无法通过及时处理与反馈，就会得不偿失。可以说，大数据时代对人类的数据驾驭能力提出了新挑战，也为人们获得更为深刻、全面的洞察能力提供了前所未有的空间与潜力。

大数据时代主要有 4 个挑战：

第一个挑战是数据量大。

大数据的起始计量单位是 PB（1000TB）、EB（100 万 TB）或 ZB（10 亿 TB）。目前，企业面临数据量的大规模增长，预测到 2020 年，全球数据量将扩大 50 倍。如今，大数据的规模尚在不断变化，单一数据集的规模范围从几十 TB 到数 PB 不等。

第二个挑战是数据类型繁多。

包括网络日志、音频、视频、图片、地理位置信息等，多种类型的数据对数据处理能力提出了更高要求。数据多样性的增加主要由新型多结构数据和多种数据类型（包括网络日志、社交媒体、互联网搜索、手机通话记录及传感器网络等）造成。其中，部分传感器安装在火车、汽车和飞机上，每个传感器都增加了数据的多样性。

第三个挑战是数据价值密度低。

大数据非常复杂，有结构化的，也有非结构化的，增长速度飞快，单条数据的价值密度极低。此外，随着物联网的广泛应用，信息感知无处不在。信息海量，但价值密度较低，如何通过强大的机器算法更迅速地完成数据的价值“提纯”，是大数据时代亟待解决的难题。

第四个挑战是高速性。

描述的是数据被创建和移动的速度。在高速网络时代，通过实现软件性能优化的高速电脑处理器和服务器的，创建实时数据流已成为流行趋势。企业不仅需要了解如何快速创建数据，还必须知道如何将数据快速处理、分析并返回给用户，以满足用户的实时需求。

1.2 大数据可视化的难点

大数据具有多层结构，意味着会呈现多变的形式和类型。相较于传统的业务数据，大数据存在不规则和模糊不清的特性，造成很难甚至无法使用传统应用软件进行分析。传统业务数

据随时间演变已拥有标准的格式，能够被标准商务智能软件识别。目前，企业面临的挑战是处理并从各种形式呈现的复杂数据中挖掘价值。

传统数据可视化工具仅将数据加以组合，通过不同展现方式提供给用户，用于发现数据之间的关联信息。近年来，随着云和大数据时代的来临，数据可视化产品已经不再满足于使用传统数据可视化工具对数据仓库中的数据抽取、归纳并简单的展现。新型数据可视化产品必须满足互联网爆发的大数据需求，必须快速收集、筛选、分析、归纳、展现决策者所需要的信息，并根据新增数据进行实时更新。

中国传媒大学新闻学院沈浩教授说过，“随着非结构和半结构化数据的增长，数据可视化的发展需要迎合多类型数据，词云、泡泡图、热图等形式的出现更加贴合新数据类型。”另外，在展现形式上，数据可视化工具还应该满足直接发布到云端、移动端的需求。

阿里巴巴数据平台事业部资深开发工程师宁朗说过，“数据可视化是大数据和大智慧之间的桥梁，大数据将数据变为设计师，每个人都可以利用。”

Splunk 中国区高级售前工程师崔玥说过，“如同 Windows 重新定义了操作系统，数据可视化重新定义了数据分析，将数据从晦涩的代码中脱离出来，通过简单的图形界面和大众更易接受的方式提供一个展现、监控数据的平台，让数据分析工作更简单。”

QlikView 南北亚区售前经理张子斌说过，“数据可视化利用人类发现复杂数据中的异常、模式、趋势甚至相关性的天然能力，这是我们无法用数据的行和列做到的。”好的数据可视化伴随内存关联技术、移动和社交商业探索能力，能让使用者自由、高效地挖掘数据以找出重要规律并做出决策。

1.3 可视化技术的新特性

数据可视化的历史可以追溯到 20 世纪 50 年代计算机图形学的早期，人们利用计算机创建了首批图形图表。到了 1987 年，一篇题目为《Visualization in Scientific Computing》（科学计算中的可视化，即科学可视化）的报告成为数据可视化领域发展的里程碑，它强调了基于计算机可视化技术新方法的必要性。

随着人类采集数据种类和数量的增长、计算机运算能力的提升，越来越多高级计算机图形学技术与方法应用于处理和可视化这些规模庞大的数据集。20 世纪 90 年代初期，“信息可视化”成为新的研究领域，旨在为许多应用领域对于抽象异质性数据集的分析工作提供支持。

当前，数据可视化是一个既包含科学可视化，又包含信息可视化的新概念。数据可视化是可视化技术在非空间数据上的新应用，使人们不再局限于通过关系数据表观察和分析数据信息，还能以更直观的方式看到数据与数据之间的结构关系。

数据可视化是关于数据视觉表现形式的研究。这种数据视觉表现形式被定义为一种以某种概要形式抽取出来的信息，包括相应信息单位的各种属性和变量。

数据可视化技术的基本思想是将数据库中每一个数据项作为单个图元元素表示，是由大量数据构成的数据图像，同时将数据的各个属性值以多维数据的形式表示，可以从不同维度观察数据，从而对数据进行更深入的观察和分析。

在大数据时代，数据可视化工具必须具有以下 4 种新特性。

- **实时性：**数据可视化工具必须适应大数据时代数据量的爆炸式增长需求，必须快速收集、分析数据，并对数据信息进行实时更新。
- **简单操作：**数据可视化工具满足快速开发、易于操作的特性，能满足互联网时代信息多变的特点。
- **更丰富的展现：**数据可视化工具需具有更丰富的展现方式，能充分满足数据展现的多维度要求。
- **多种数据集成支持方式：**数据的来源不局限于数据库，数据可视化工具将支持团队协作数据、数据仓库、文本等多种方式，并能够通过互联网进行展现。

数据可视化的思想是将数据库中每一个数据项作为单个图元元素，通过抽取的数据构成数据图像，同时将数据的各个属性值加以组合，并以多维数据的形式通过图表、三维等方式展现数据之间的关联信息，使用户能从不同维度和不同组合对数据库中的数据进行观察，从而对数据进行更深入的分析 and 挖掘。

1.4 主要的数据可视化软件

1.4.1 Tableau

Tableau 是桌面系统中最简单的商业智能工具软件。Tableau 没有强迫用户编写自定义代码，新控制台也可以完全自定义配置。在控制台上，不仅能够监测信息，还提供了完整的分析能力。Tableau 控制台灵活，具有高度动态性。

Tableau 简单、易用、快速，一方面归功于产生自斯坦福大学的突破性技术。Tableau 是集复杂的计算机图形学、人机交互和高性能的数据库系统于一身的跨领域技术，其中最耀眼的莫过于 VizQL 可视化查询语言和混合数据架构。另一方面在于 Tableau 专注于处理最简单的结构化数据，即已整理好的数据——Excel、数据库等，结构化数据处理在技术上难度较低，这就使得 Tableau 有精力在快速、简单和可视上做出更多改进。

2014 年 3 月，IT168 网站进行了一项有关数据可视化的调查，已经部署数据可视化的企业仅为 15%，有 56% 的企业计划 1~2 年内部署相关应用。从企业部署可视化的目的来看，排在前三位的分别为：通过可视化发现数据的内在价值（36%）、满足高层领导的决策需要（30%）和满足业务人员的分析需要（25%），仅有 9% 的企业选择更美观的展现效果。

针对 Tableau、Qlik、TIBCO Software、SAS、Microsoft、SAP、IBM 和 Oracle 八家数据可视化产品和服务提供商的调查，分别从知名度、流行度和领导者三个角度进行分析。从知名度来看，8 家厂商几乎不分先后，只有微小的差距；从流行度来看，SAP、IBM 和 SAS 占据前三位，所在比例分别为 19%、18% 和 17%；从领导者来看，Tableau 以 40% 的优势遥遥领先。

1.4.2 QlikView

QlikView 是一个完整的商业分析软件，开发者和分析者可以使用 QlikView 构建和部署强

大的分析应用，各种终端用户可以高度可视化、功能强大和创造性的方式互动分析重要业务信息。

QlikView 是一个具有完全集成 ETL 工具向导驱动的应用开发环境，拥有强大的 AQL 分析引擎和一个高度直觉化、使用简单的用户界面。QlikView 让开发者能从多种数据库里提取和清洗数据，建立强大、高效的应用，并能被移动用户和每天的终端用户修改后使用。

QlikView 是一个可升级的解决方案，完全利用了基础硬件平台，用上亿数据记录进行业务分析。QlikView 由开发工具（QlikView Local Client）、服务器组件（QlikView Server）、发布组件（QlikView Publisher）以及其他应用接口（SAP\Salesforce\Informatica）组成，服务器支持多种发布方式，如 AJAX 客户端、ActiveX 客户端，还可以与其他 CS/BS 系统进行集成。

1.4.3 ECharts

ECharts（Enterprise Charts）是商业级数据图表，一个纯 JavaScript 的图表库，可以在 PC 和移动设备上流畅运行，兼容当前绝大部分浏览器，底层依赖轻量级的 Canvas 类库 ZRender，提供直观、生动、可交互、可高度个性化定制的数据可视化图表。创新的拖曳重计算、数据视图、值域漫游等特性大大增强了用户体验，赋予用户对数据进行挖掘、整合的能力。

ECharts 支持折线图（区域图）、柱状图、散点图（气泡图）、K 线图、饼图（环形图）、雷达图、和弦图、力导向布局图、地图、仪表盘、漏斗图、事件河流图等 12 类图表，同时提供标题、详情气泡、图例、值域、数据区域、时间轴、工具箱等 7 个可交互组件，支持多图表、组件的联动和混搭。

1.4.4 HighCharts

HighCharts 界面美观，由于使用 JavaScript 编写，因此不需要像 Flash 和 Java 一样需要插件才可以运行，而且运行速度快。另外，HighCharts 有很好的兼容性，能够完美支持当前大多数浏览器。

HighCharts 是纯 JavaScript 编写的图表库，能够很简单、便捷地为 Web 网站或 Web 应用程序添加交互性图表，并且免费供个人学习、个人网站和非商业用途使用。HighCharts 支持的图表类型主要有曲线图、区域图、柱状图、饼状图、散状点图和综合图表等。

1.5 Tableau 软件概况

Tableau 公司成立于 2003 年，是由斯坦福大学的三位校友 Christian Chabot（首席执行官）、Chris Stole（开发总监）以及 Pat Hanrahan（首席科学家）在远离硅谷的西雅图注册成立的。其中，Chris Stole 是计算机博士，Pat Hanrahan 是皮克斯动画工作室的创始成员之一，曾负责视觉特效渲染软件的开发，两度获得奥斯卡最佳科学技术奖，至今仍在斯坦福担任教授职位，教授计算机图形课程。

Tableau 公司主要面向企业数据提供可视化服务，是一家商业智能软件提供商。企业运用 Tableau 授权的数据可视化软件对数据进行处理和展示，不过 Tableau 的产品并不局限于企业，

其他机构甚至个人都能很好地运用 Tableau 软件进行数据分析工作。数据可视化是数据分析的完美结果，能够让枯燥的数据以简单友好的图表形式展现出来。可以说，Tableau 在抢占细分市场，也就是大数据处理末端的可视化市场，目前市场上并没有太多这样的产品。同时，Tableau 还为客户提供解决方案服务。

1.5.1 Tableau Desktop

“所有人都能学会的业务分析工具”，这是 Tableau 官方网站上对 Tableau Desktop 的描述。确实，Tableau Desktop 的简单、易用程度令人发指，这也是软件的最大特点。使用者不需要精通复杂的编程和统计原理，只需要把数据直接拖放到工具簿中，通过一些简单的设置就可以得到想要的可视化图形。

Tableau Desktop 的学习成本很低，使用者可以快速上手，这无疑对日渐追求高效率 and 成本控制的企业来说具有巨大吸引力，特别适合日常工作中需要绘制大量报表、经常进行数据分析或需要制作图表的人使用。简单、易用并没有妨碍 Tableau Desktop 拥有强大的性能，它不仅能完成基本的统计预测和趋势预测，还能实现数据源的动态更新。Tableau Desktop 的开始页面如图 1-1 所示。

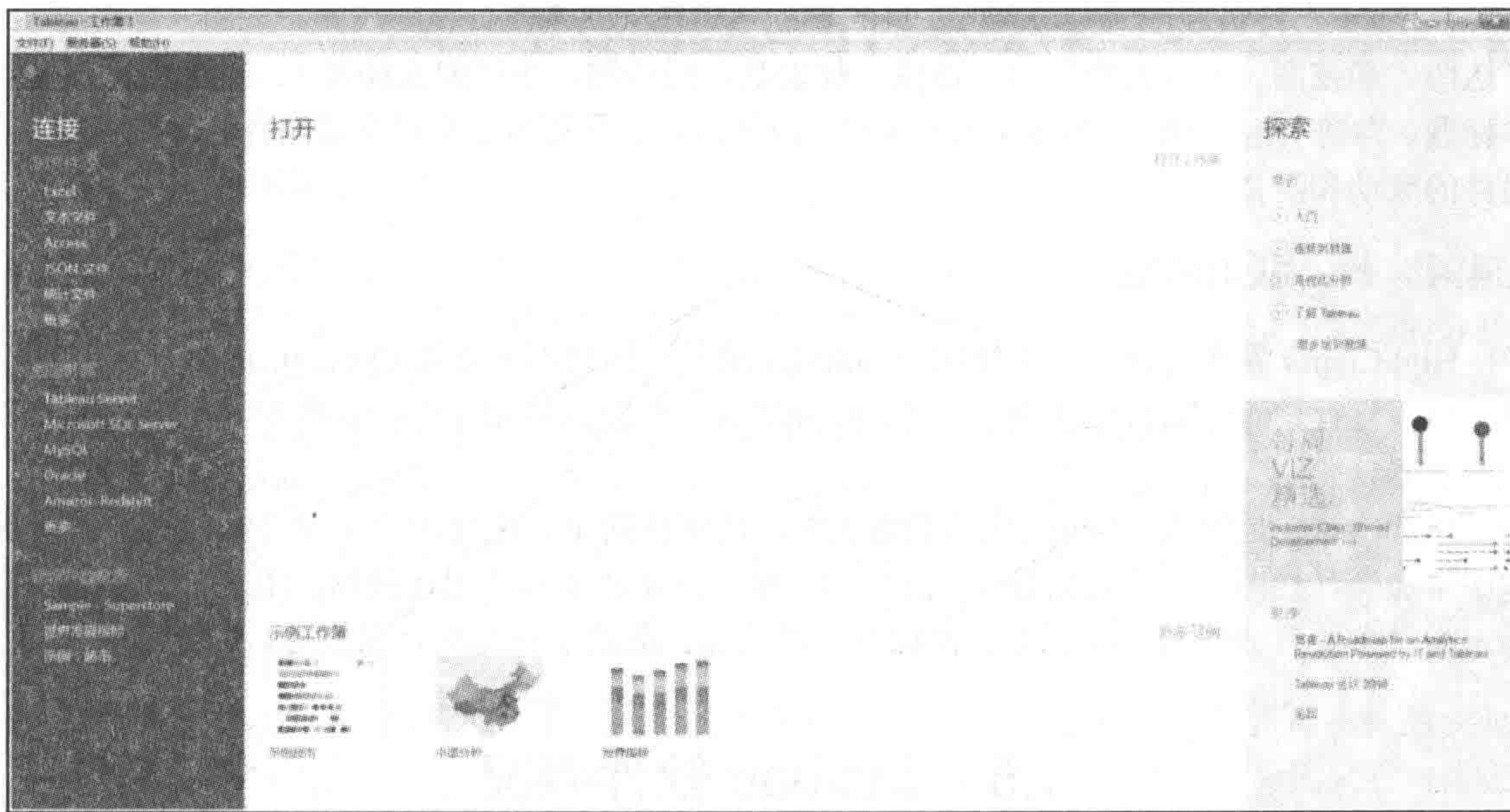


图 1-1 Tableau Desktop 的开始页面

Tableau Desktop 不同于 SPSS，SPSS 作为统计分析软件，比较偏重于统计分析，使用者需要有一定数理统计基础，虽然功能强大且操作简单、友好，但输出的图表与办公软件的兼容性 & 交互方面有所欠缺。Tableau Desktop 是一款完全的数据可视化软件，专注于结构化数据的快速可视化，使用者可以快速进行数据可视化并构建交互界面，用来辅助人们进行视觉化思考，并没有 SPSS 强大的统计分析功能。

总之，快速、易用、可视化是 Tableau Desktop 最大的特点，能够满足大多数企业、政府机构数据分析和展示的需要，以及部分大学、研究机构可视化项目的要求，而且特别适合企业