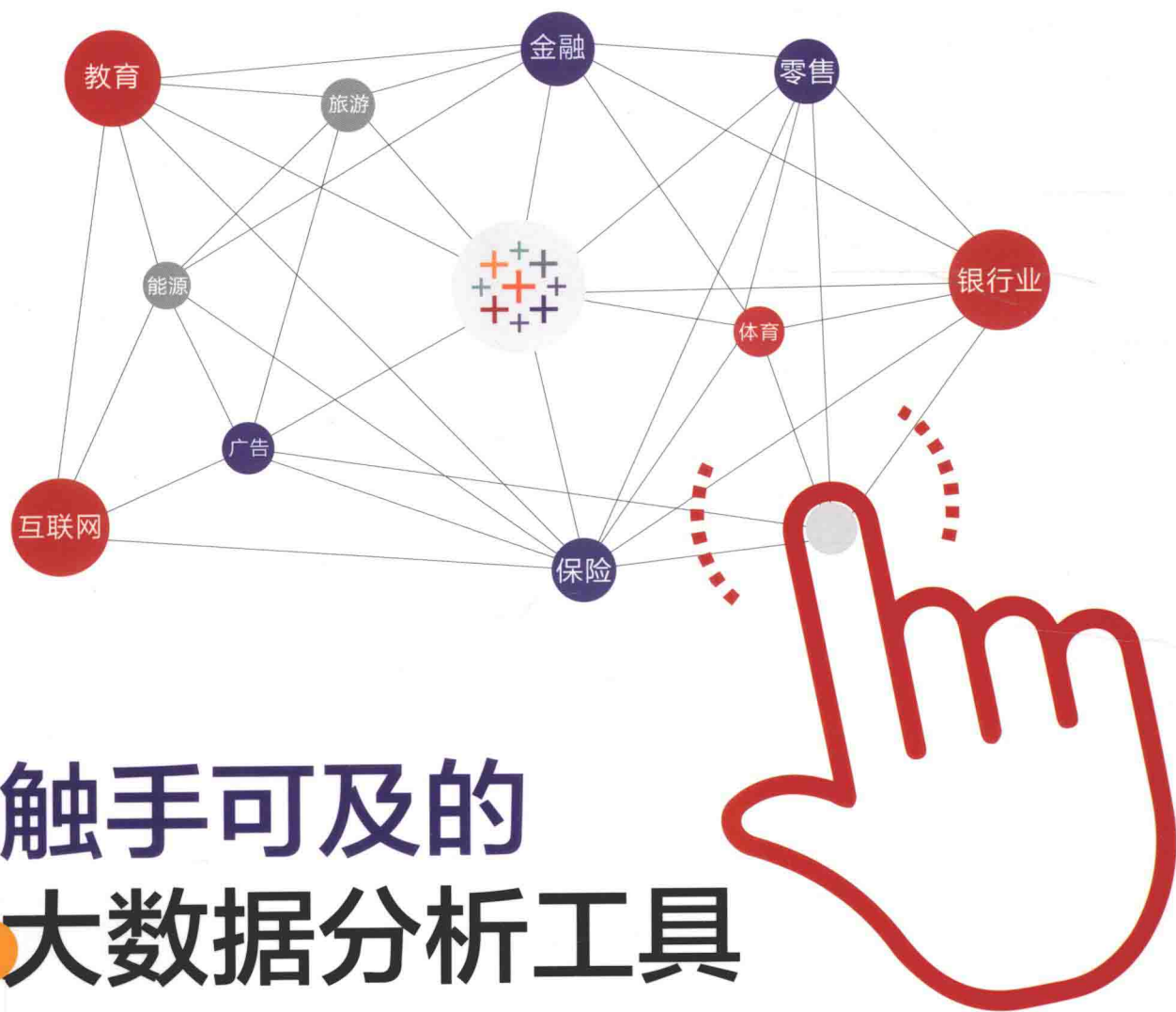


国内首本Tableau著作，最新版本Tableau 9.0
实战经验分享，精选28个案例，涉及15个行业



触手可及的 大数据分析工具 Tableau 案例集

沈浩 王涛 韩朝阳 李健◎著

CDA数据分析师 系列丛书

触手可及的 大数据分析工具



Tableau案例集

沈浩 王涛 韩朝阳 李健◎著



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

本书对 Tableau 的产品、优势、Tableau 9.0 的新特性及其功能做了全面的介绍，并且从 15 个不同行业的案例入手，让你在阅读本书后能够从一个新手成长为能够创建出复杂仪表盘的高手。

本书分为五个部分共 14 章：Tableau 使用概述、新手上路、成功晋级、高手秘籍以及实际应用。

Tableau 使用概述包含四章，分别从为何要进行数据可视化、Tableau 的发展历程、Tableau 的产品介绍、Tableau 9.0 的新特性、Tableau 的应用优势以及如何利用 Tableau 进行数据连接和了解工作区几个方面做了详尽的阐述。

新手上路包含两章，该部分以各行业案例为依托，带你从 Tableau 的排序、筛选、分层和分组、参数和函数等方面来全面了解 Tableau 9.0 的新功能。

成功晋级包含四章，分别从不同行业的案例入手介绍 Tableau 的高级功能。

高手秘籍包含三章，该部分介绍如何利用 Tableau 中的颜色和形状进行异常值检测、如何利用背景图像进行货架分析以及如何使用 Tableau 中的超级链接使用 WMS 地图服务和动态调用外部网页。

实际应用自成一章包含八个数据可视化分析实例，全面展示 Tableau 进行数据可视化分析的各种技术细节和实用技巧。

本书对于初次接触 Tableau 软件的读者会有很大帮助，书中对 Tableau 创建可视化视图的每一步操作都有详尽的说明。如果你是 Tableau 软件的老用户，本书中不同行业案例的可视化分析和展现方式也会给你带来全新的思路和视角。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目(CIP)数据

触手可及的大数据分析工具：Tableau 案例集 / 沈浩等著. —北京：电子工业出版社，2015.9
(CDA 数据分析师系列丛书)

ISBN 978-7-121-26938-7

I. ①触… II. ①沈… III. ①可视化软件—教材 IV. ①TP31

中国版本图书馆 CIP 数据核字(2015)第 189077 号

策划编辑：张慧敏

责任编辑：徐津平

特约编辑：顾慧芳

印 刷：北京京科印刷有限公司

装 订：三河市华成印务有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱 邮编 100036

开 本：787×1092 1/16 印张：27.75 字数：575 千字

版 次：2015 年 9 月第 1 版

印 次：2015 年 9 月第 1 次印刷

印 数：4000 册 定价：79.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：(010) 88254888。

质量投诉请发邮件至 zltts@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线：(010) 88258888。

前言

本书写作背景

回顾过去近十年中国商业智能（**Business Intelligence**，以下简称 **BI**）市场的经营与运作，若以一个业内人士的身份来描述这段历程，我想把它归纳为“光环褪去，返璞归真”，**BI** 之于中国，已经渐渐从概念走向了成熟。

这份成熟，从厂商角度来看，我们很容易观察到绝大多数原厂收敛了自己“**BI** 无所不能”的宣传方向，回归“商业智能”的本源：或从技术底层加强自己处理海量数据的能力与速度，或潜心整合收买而来的各套系统，或从数据分析与数据挖掘的模块上融入更多的常用模型，或从操作习惯上更贴近人类天生的惰性……

而从最终用户角度来看，仅凭几套漂亮 **Dashboard** 所描述的美好未来就签合同的场景已经不多见了，越来越多的企业意识到，商业智能只不过是一套提升业务分析、数据展现与信息分享能力的 **IT** 系统，它非常需要规范的数据源，非常需要懂得业务的专业人士和愿意动手获取信息的操作者。

也正是在过去的十年里，我们看到了层出不穷的商业智能新产品。它们不论在开发之初的源代码上有多么不同，在最终使用者的体验上，都有着相似的感觉：轻巧灵便、架构简洁、价格实惠，而 **Tableau** 是其中的佼佼者。不论你曾经买过什么系统：**ERP**、**CRM**、**SCM**、**OA**、**KM**、**her**、…，还是任何品牌：**SAP**、**IBM**、**ORACLE**、**MSTR**、**IBI**、**SAS**、**SPSS**、…，你都可以把 **Tableau** 部署在其中且增色不少，它无意取代，愿作为补充和桥梁。

我要感谢中国传媒大学的沈浩教授在 2007 年向博易智讯推荐了 **Tableau**，让我们一直关注和学习这个产品，并且最终成为 **Tableau** 中文版推出时的第一批“代言人”。作为国内最早专注于玩数据的资深教师，沈老师有一双特殊的慧眼，总能在第一时间发现任何实现“从数据到信息”的新技术。

Tableau 在中国的第一批客户，不仅包括眼明手快的年轻精英们，而且包括了许多近花甲之年的老教授、研究员和管理者。它作为一款商业智能与数据分析软件，能够满足企业与科研绝大多数的分析需求。相比于专业的 **SPSS Statistics** 与 **SAS** 等，它入

门简单且功能强大；相比于各种品牌的大型 IT 平台，它易于实施与部署，通过拖、拉、点击几步操作即可完成，无需太多的专业背景，每个人都可以轻松掌握，真正实现了“亲自上阵”，“所见即所得”。为了帮助更多的人了解 Tableau 能够实现哪些图表的制作与分析，我们与沈浩教授共同编写了本书。

本书是国内实际意义上的第一本 Tableau 中文教材，文中的案例部分选自 Tableau 原创的演示数据，绝大多数案例由国内的真实案例修改而成。为了方便广大的 Tableau 采购单位或个人爱好者学习与掌握，你可以加入 QQ 群 202542742 和众多学友一起交流；同时欢迎大家访问我们创建的 Tableau 爱好者交流地：www.tableauhome.com.cn，我们将有更多的精彩内容与你分享。

书中案例均采用 9.0 版本完成，Tableau 9.0 的发布进一步印证了未来五年的数据应用趋势：快、自助、移动端、云模式共享，人人都能玩数据……

这是我们最喜闻乐见的。我们的名字“Bizinsight”是“Business + Insight”的合集，我们为“商业洞察”而来。这意味着，没有 IT 背景的商业用户借助 Tableau 也能够第一时间查看数据、分析数据和分享数据。这是我们撰写本书的初衷，也是我们在过去十年与未来十年的使命。

如何阅读本书

本书以实例讲解为主，目的是让大家熟悉制图所需的每一个步骤，掌握每一项技能。你可以从头开始完整地读一遍，不过如果你已经有想法在酝酿了，也可以只挑选最感兴趣的几章来读。所有的章节都经过了精心的组织，案例是相互独立的。如果读者对数据领域还比较陌生，那么阅读最开始的几章应该会有帮助；这几章介绍了处理数据的方法、需要关注的重点以及各种可用的作图方法，便于读者了解如何获取数据，如何规范格式并为可视化做准备。之后的几章会根据不同的数据类型和侧面分别介绍各种可视化技巧。请记住，永远都要让数据说话。

不管你选择何种阅读方式，我都强烈建议你在阅读时打开电脑，和我们一起逐步完成每一个范例。本书案例丰富，几乎涵盖了所有的行业，为了便于读者找到需要的案例，我们已将案例及其相关信息列在了“本书案例汇总表”中。祝大家阅读开心！



撰写人：何业文（本书组稿和审稿人）

博易智讯（北京）信息技术有限公司 创办人

十年专注于数据挖掘与数据可视化，经济学、统计学、传播学背景，一个创数据的人……

本书案例汇总表

序号	15 个不同的行业	行业	案例名称	章节	页码	所讲解的关键功能
1	①	零售行业	销售数据分析	第 5 章	42	主要功能
2	②	教育行业	教育水平评估	第 6.1 节	132	折线图，地图
3	③	互联网	网站内容评估	第 6.2 节	146	单点图，条形图
4	④	银行业	投资分析	第 6.3 节	151	图形合并
5	⑤	金融保险	索偿分析	第 7.1 节	166	地图
6	⑥	房地产业	识别与预测	第 7.2 节	171	地图
7	⑤	金融保险	欺诈检测	第 8.1 节	176	趋势图
8	⑦	制造行业	生产分析	第 8.2 节	179	子弹图，甘特图，趋势图
9	⑧	能源行业	资源组合分析	第 8.3 节	189	趋势图，辅助曲线
10	⑨	人力资源	继任规划	第 9.1 节	196	仪表板
11	⑧	能源行业	资产监控	第 9.2 节	201	仪表板
12	③	互联网	网站客户细分	第 10.1 节	207	气泡图，面积图，地图
13	①	零售行业	零售业客户细分	第 10.2 节	212	气泡图，地图，交叉表
14	⑩	网络游戏业	游戏客户洞察	第 10.3 节	217	图形合并
15	⑤	金融保险	识别与预测	第 11.1 节	228	双叶图
16	⑧	能源行业	门户创建	第 11.2 节	243	钻取突显功能
17	⑪	广告行业	网络广告投放分析	第 11.3 节	250	曲线拟合
18	⑫	体育行业	NBA 赛事分析	第 12.1 节	261	个性化背景设计
19	①	零售行业	货架图分析	第 12.2 节	265	个性化背景设计
20	①	零售行业	超级链接	第 13 章	276	超级链接
21	⑥	房地产业	楼市分析	第 14.1 节	292	综合
22	⑬	旅游行业	骑行路线分析	第 14.2 节	311	综合
23	②	教育行业	教育行业发展指标分析	第 14.3 节	321	综合
24	⑭	交通行业	空难统计	第 14.4 节	337	综合
25	③	互联网	网站访问	第 14.5 节	365	综合
26	⑫	体育行业	世界杯赛	第 14.6 节	373	综合
27	⑮	环境工程	太平洋风季	第 14.7 节	391	综合
28	⑭	交通行业	伦敦巴士路线	第 14.8 节	406	综合

致谢

首先，感谢广大 Tableau 爱好者对此书的关注，因为你们的殷殷期望，使得本书有了最新的 Tableau 9.0 版本。

其次要感谢我的学生们。自从 2006 年我发现了 Tableau 这款数据可视化工具，就立刻将这款软件应用于电视新闻学院的教学当中，我们将它作为描述性统计分析和数据可视化的重要工具。近十年来，一部分毕业生成长为国内最早从事数据新闻写作的专业人士，另一部分毕业生成为了企业内部应用 Tableau 实现商业分析的统计工程师。可以说，与你们的教学互动是我不断前进的动力，让我在追求技术时髦的道路上不知疲倦。

本书的出版要特别感谢 2003 级的毕业生何业文，她在追求商业洞察力的道路上是最具行动力的实践者。她组建了一家公司 Bizinsight 来实现自己对数据挖掘与数据可视化的追求。本书的完成是我和 Bizinsight 的几位技术工程师王涛、韩朝阳、李健集体劳动的成果。

还要感谢两位伙伴，他们都曾属于 Bizinsight，并为此书的编写做出了贡献，分别是：最富有学习精神的包明明，以及富有闯劲的汪尚，他们正在追求数据分析的道路上探索自己的方向。

此外，本书第 5 部分的编写得益于 2014 年中国传媒大学数据可视化小组的同学们，他们是富有创新精神和学习能力的一群年轻人，我也期待着未来十年，从他们当中出现下一位沈老师。

我还要感谢此书的技术编辑，分别是罗玮、王磊和闵亮，他们纠正了书中的一些技术性错误，并使语言描述的逻辑更加严谨，且通俗易懂。

最后，感谢电子工业出版社，感谢出版社的编辑校对人员，使此书成功出版。

再次感谢大家！

沈浩博士

中国传媒大学电视新闻学院教授

中国传媒大学调查统计研究所所长

第三届全国社会媒体处理大会“社会媒体处理与社会科学”领域主席

本书的答疑服务

为了更好地服务于 Tableau 用户和爱好者，加强作者和广大读者之间的交流，我们提供网上答疑服务：

网站地址：www.tableauhome.com.cn

QQ 群：202542742

资源索取：

书中全部实战案例所需的数据源文件，请通过电子工业出版社的指定链接下载。

目录

第 1 部分 Tableau 使用概述

第 1 章 数据可视化.....	2
1.1 用数据讲故事.....	2
1.2 数据不只是数字.....	3
1.3 在数据中寻找什么.....	3
1.4 本章小结.....	5
第 2 章 Tableau 概述.....	6
2.1 Tableau 的发展历程.....	6
2.2 Tableau 产品简介.....	7
2.2.1 Tableau Desktop.....	7
2.2.2 Tableau Server.....	9
2.2.3 Tableau Online.....	9
2.2.4 Tableau Reader.....	9
2.2.5 Tableau Public.....	9
2.3 Tableau 9.0 版本新特性.....	10
2.4 本章小结.....	16
第 3 章 Tableau 应用优势.....	17
3.1 简单易用.....	17
3.2 极速高效.....	18
3.3 美观交互的视图与界面.....	20
3.4 轻松实现数据融合.....	22
3.5 简便的管理.....	23
3.6 灵活的配置.....	23

3.7 本章小结..... 24

第 4 章 Tableau 功能介绍..... 25

4.1 数据连接..... 25

4.1.1 数据文件连接 25

4.1.2 数据库连接 27

4.2 了解 Tableau 工作区 29

4.3 本章小结..... 39

第 2 部分 新手上路

第 5 章 创作第一个仪表板..... 42

5.1 排序..... 42

5.2 分层与分组..... 46

5.2.1 分层 46

5.2.2 分组 50

5.3 参数设置..... 52

5.4 语法操作..... 55

5.4.1 主要功能函数简介 55

5.4.2 快速表计算简介 62

5.5 可视化..... 67

5.5.1 地图 67

5.5.2 条形图 73

5.5.3 线形图 77

5.5.4 饼图 79

5.5.5 复合图 80

5.5.6 嵌套条形图 81

5.5.7 动态图 84

5.5.8 热图 86

5.5.9 突显表 89

5.5.10 散点图 90

5.5.11 气泡图 93

5.5.12 甘特图 93

5.5.13 标靶图 95

5.5.14 盒须图 98

5.5.15 瀑布图	100
5.5.16 直方图	101
5.5.17 帕累托图	103
5.5.18 填充气泡图	107
5.5.19 文字云	110
5.5.20 树地图	111
5.6 设计动态仪表板	113
5.6.1 新建一个仪表板	113
5.6.2 创建动作	118
5.6.3 使用仪表板的注意事项	123
5.7 作品分享方案	123
5.8 本章小结	131
第 6 章 实战演练	132
6.1 教育水平评估	132
6.1.1 学校教育水平评估	132
6.1.2 城市教育水平评估	139
6.2 网站内容评估	146
6.3 投资分析	151
6.4 本章小结	164
第 3 部分 成功晋级	
第 7 章 巧用地图	166
7.1 索偿分析	166
7.2 估值分析	171
7.3 本章小结	175
第 8 章 美化趋势图	176
8.1 欺诈检测	176
8.2 生产分析	179
8.3 资源组合分析	189
8.4 本章小结	195

第 9 章 设计动态仪表板 196

9.1 继任规划 196

9.2 资产监控 201

9.3 本章小结 206

第 10 章 客户细分 207

10.1 网站客户细分 207

10.2 零售业客户细分 212

10.3 游戏客户洞察 217

10.4 本章小结 226

第 4 部分 高手秘籍

第 11 章 生“动”形“象” 228

11.1 识别与预测 228

11.2 门户创建 243

11.3 网络广告投放分析 250

11.4 本章小结 260

第 12 章 设计个性化背景 261

12.1 NBA 赛事分析 261

12.2 货架图分析 265

12.3 本章小结 275

第 13 章 超级链接 276

13.1 使用 Web 地图服务 277

13.2 淘宝网数据数据分析 281

13.3 本章小结 290

第 5 部分 实际应用

第 14 章 实际应用案例分析 292

14.1 中国楼市降温的分析 292

14.2 中国最美八条骑行路线的展示 311

14.3 中国教育水平发展指标历史数据统计报告 321

14.4 空难信息统计 337

14.5	Tableau 官网各版块访问情况	365
14.6	2014 年巴西世界杯小组赛各球队球员数据统计	373
14.7	2014 年太平洋台风季分析	391
14.8	制作“伦敦巴士线路数据”视图	406
附 A 录	Tableau 安装	415
附 B 录	Tableau 函数汇总	419

第 1 部分

Tableau 使用概述

数据可视化已经不是什么新鲜的话题，但是它的价值却与日俱增，不仅是从杂乱的数字到美观的视图的飞跃，更是从杂乱的、难以“看透”的数据信息到直观易懂的企业决策信息的蜕变。数据可视化在提升企业形象的同时，提高了企业的收益，被称为企业难题的“美丽杀手”。

- 第 1 章 数据可视化
- 第 2 章 Tableau 概述
- 第 3 章 Tableau 应用优势
- 第 4 章 Tableau 功能介绍

第 1 章

数据可视化

数据可视化越来越成为企业核心竞争力的一个重要组成部分。从数字可视化到文本可视化，从条形图、饼图到文字云，从数据的可视化分析到企业的可视化平台建设。然而，这一块国内人才紧缺，只有为数不多的高校刚开始培养这方面的人才。因此，就需要企业自己培养可视化人才。为了快速掌握数据可视化技术，企业选择一款易于学习的可视化工具是非常必要的，而 Tableau 这款软件操作简单易学，而且功能强大、计算迅速，还具有文字云、回归分析、时间序列预测等功能。因此，本书主要基于 Tableau 介绍数据可视化技术。在这一章，我们主要对数据可视化的概念进行简要说明。

1.1 用数据讲故事

如果一开始我们不知道自己想了解什么，或者不知道有什么可以去了解，那么数据就是枯燥的。它不过是数字和文字的堆砌，除了冰冷的数值之外没有任何意义。而统计和可视化的好处就在于能帮助我们观测到更深层次的东西。事实上，数据是现实生活的一种映射，其中隐藏着许多故事，在那一堆堆的数字之间存在着实际的意义、真相和美学。而且和现实生活一样，有些故事非常简单直接，有些则颇为迂回费解。有些故事只会出现在教科书里，其他一些则题材新奇。讲故事的方式完全取决于你自己，不论你的身份是统计学家、程序员、设计师还是数据研究者。

数据的故事无处不在，企业运营、新闻报道、艺术等，都在用数据讲故事。你准备好怎么讲了吗？

1.2 数据不只是数字

数据不只是数字，还可以是文字、图片、视频、声音等。我们不仅可以用可视化来展现数字特征，还可以用来表达人的情感。就像我们可以从一个人的文章或评论中来分析他所用的词汇是消极的还是积极的，并用可视化的形式来展现这个人的情感；也可以通过产品类别的名称的大小来表示这些产品的销售额或利润大小等，如图 1-2-1 所示。

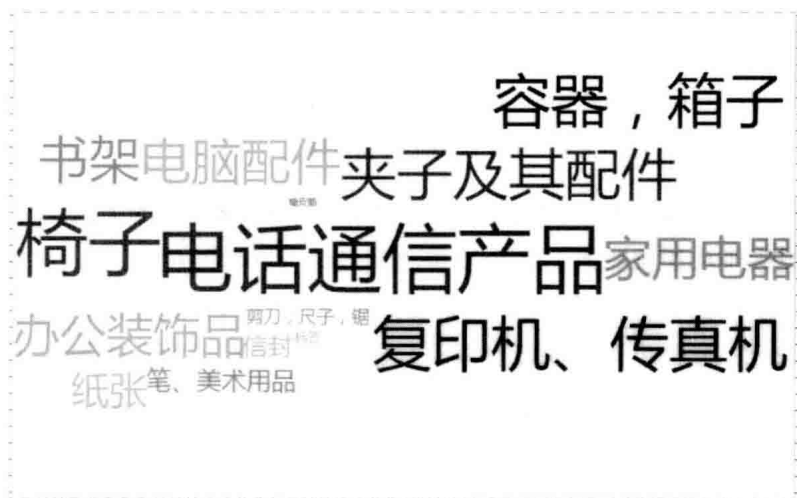


图 1-2-1 可视化展示

1.3 在数据中寻找什么

我们通过数据可视化，是为了从中寻找什么呢？三个方面：模式、关系和异常。不管图形表现的是什么，我们都要留心观察这三者。

模式，即数据中的某种规律。比如机场每月的旅客人数随着时间推移变化不定，通过几年的数据的对比，我们可能发现旅客人数存在着季节性或周期性的变化，某些月份的旅客数量一直偏低，某些月份的旅客数量则一直偏高。又比如，我们可以分析某家网站不同时间各个板块的访问量，等等。如图 1-3-1 所示。

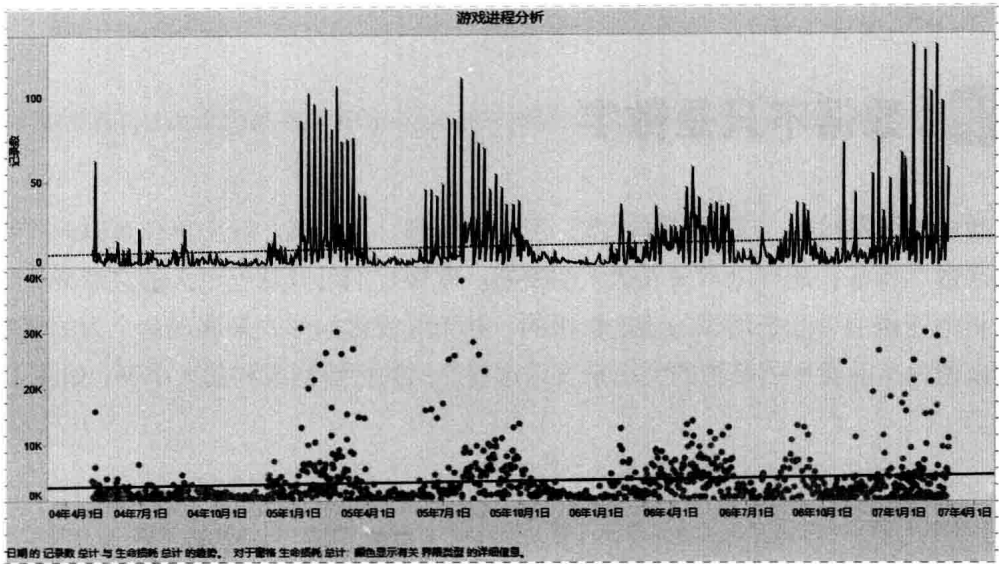


图 1-3-1 数据规律

关系，即指各影响因素之间的相关性，也指各个图形之间的关联。在统计学中，它通常代表关联性和因果关系。多个变量之间应该存在着某种联系。比如，在散点图中，我们可以观察两个坐标轴的两个字段之间的相关关系，是正相关还是负相关，或是不相关。如此，我们可以依次找到与因变量具有较强相关的自变量，从而确定主要的影响因素。比如我们研究网站访问的目标完成情况与访问量、转化率等的关系，如图 1-3-2 所示。

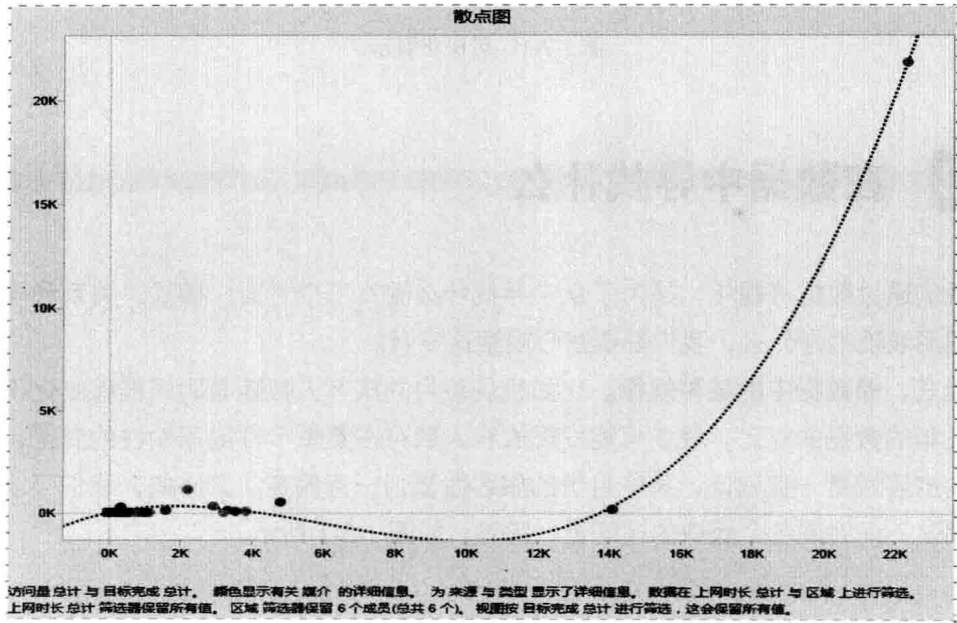


图 1-3-2 关联性分析