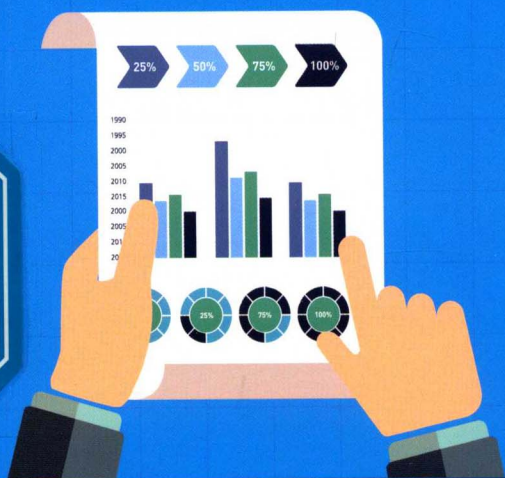


系统 / 专业 / 实战 / 接地气

Broadview®
www.broadview.com.cn

人人 都能学会的 Excel数据分析

罂粟姐姐 著



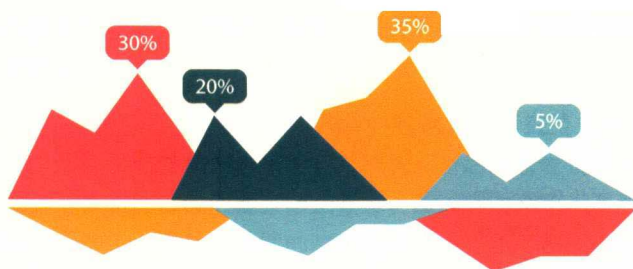
不需要高深理论，不需要复杂编程，
所有案例均从真实项目中提取，
只需使用Excel就能完成数据分析全流程

 中国工信出版集团

 电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

人人 都能学会的 Excel数据分析

罂粟姐姐 著



電子工業出版社

Publishing House of Electronics Industry

北京•BEIJING

内 容 简 介

这是一本有趣、有料又文艺的书。

本书在对百余个市场调研及管理咨询项目进行分析与提炼的基础上，汇集了数据分析师在入门阶段使用 Excel 进行数据处理与分析过程中的常见需求，通过具体项目案例的演示和讲解，将数据分析的核心技能和关键流程手把手地教给读者。

全书共 7 章，内容涉及数据分析的基本概念、数据输入、数据整理、数据处理、数据分析、数据呈现及数据分析报告，结构严谨、条理清晰、图文并茂、旁征博引、通俗易懂，使读者不仅能够读下去，还能读进去。

本书不仅能够帮助在校学生增加就业机会，帮助职场新人提升职场竞争力，帮助市场营销、人力资源、财务管理等人员解决实际问题，也能成为市场调研、管理咨询、数据分析等专业人士的入门指南，还能帮助所有需要使用 Excel 的职场人士提高工作效率。

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

人人都能学会的 Excel 数据分析 / 罍粟姐姐著. —北京：电子工业出版社，2018.1

ISBN 978-7-121-33135-0

I. ①人… II. ①罍… III. ①表处理软件 IV. ① TP391.13

中国版本图书馆 CIP 数据核字 (2017) 第 294847 号

策划编辑：石 倩

责任编辑：徐津平

特约编辑：赵树刚

印 刷：北京虎彩文化传播有限公司

装 订：北京虎彩文化传播有限公司

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：787×980 1/16 印张：15.5 字数：372 千字

版 次：2018 年 1 月第 1 版

印 次：2018 年 6 月第 3 次印刷

定 价：69.00 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) 88254888，88258888。

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

本书咨询联系方式：010-51260888-819，faq@phei.com.cn。

序

作者邀我为本书作序，我倍感压力，一来觉得自己资历尚浅，怕扛不起这面大旗；二来担心在对这本书的理解不够全面和透彻的情况下就妄加凿啄，不仅枉顾了作者的一番盛情，还可能会贻笑大方。然而，当她告诉我这是一本介绍数据分析方法与技巧的书，而且我有幸拜读了本书初稿后，我勇敢地接受了这项任务，以下请容我将其中的缘由详细道来。

我们需要用哲学理解世界，也需要直观地认识世界，数据分析和呈现是我们直观认识世界的有效方式。在 10 余年的专业学习、企业实践和科学研究工作中，我一直试图对比东方管理思想和西方管理理论之间的区别，探索如何将它们巧妙地融合在一起。就我个人理解而言，我们的国学名著中蕴含的丰富的管理思想在理论研究和实践探索中给予我们诸多启发和教育，但这些著作的一个共同而显著的特征是圆融，即每句话都蕴含着丰富的哲学思想，引导人们展开无限的遐想。然而，要在实践中应用这些思想或观点，就需要个人高超的悟性及艺术的取舍。更为尴尬的是，如果你对其中的论述存在疑惑，则只能归咎于自己的理解能力不够；每当你对这些论述提出质疑时，“前辈”们会提出丰富的论据来捍卫原著的权威。所以，我们的国学经典原著生命力顽强，但后人对其的研究只是局限于从不同的角度进行解读，鲜有突破，更谈不上革新。

相比之下，西方管理理论则更为直观和具体，在他们的研究范式中首先限定假设条件，然后提出理论观点，最后通过量化方法来论证其理论。当后来者发现以往理论不足以解释现实或者不能精确地指导实践时，他们就会修正假设，改进现有理论，甚至提出截然不同的观点。西方管理理论就是这样，在后人持续的改进甚至否定过程中不断完善、走向繁荣。在这一过程中，将现实世界量化并使用科学的统计分析方法推理无疑是不可或缺的工具。毫无疑问，我们应该坚守自己的文化自信，大胆地从国学思想宝库中汲取认识和理解世界的哲学思想；同时，我们还应该持有开放的心态和开阔的视野，积极地学习西方“用数据来解读世界”的研究方法，更科学和富有技巧地投入实践、改造世界。

回归日常工作管理及具有自主空间的任何一线工作，核心环节之一就是决策，而决策的本质是选择，选择就需要遵循科学的依据做出判断。比如，西方管理理论中提出了决策树来描述这一过程，他们认为任何一项决策都面临多个选项，在某个决策点上应该选择哪个选项，取决于该选项能够带来的收益及该选项成功的概率，后续决策点上的选择要遵循同样的依据。

也就是说，人们的日常工作乃至生活就是由这样一个又一个“惊险”的选择组成的决策链。数据描述向我们展示了一个更为清晰的世界，对这些数据的统计分析则让我们在纷繁复杂的现象中提纲挈领地掌握事物发展变化的规律。因此，数据统计分析会赋予我们一双“慧眼”，帮助我们在关键决策上“入对行”或者“嫁对郎”，在日常工作中面对“三千弱水”，时常能够明智地选择适合自己的“一瓢”。

我对本书正式出版后的反响充满期待。我认为人们做工作可以分为三种境界：第一，把工作完成，即根据领导或工作职责要求，保质保量地完成任务；第二，将工作做出美感，即工作成果在直观上很协调，给人的感觉很和谐；第三，在工作上追求极致，即将个人价值取向充分地展现在工作中，将工作当作自己亲手打造的艺术品。在与作者的工作协作中，我真切地体会到她是一个追求工作美感和极致的人，她能够将一个凌乱的 PPT 瞬间“妙手回春”，使整体内容更富有逻辑，每一页的空间布局、字体组合、颜色搭配让人赏心悦目，同时在细节上追求完美。当然，她更擅长的还是数据统计分析，多年来市场调研项目经理的工作历练使她在这方面积累了丰富的实践经验；知名高校研究生的学习生涯使她具备了深厚的理论功底和超强的再学习能力；追求美感和极致的态度与习惯使她能够将看似枯燥的数据整理和分析变得花枝招展、妙趣横生，特别是这一点，作者已经在高校课堂和企业培训中做了充分的印证！

愿本书讲授的数据分析方法和技巧为我们打开一扇窗，让我们认识一个更为清晰明亮的世界；愿作者慷慨地和我们分享的这些心得体会赋予我们一双有力的翅膀，助力我们在学习和工作中一路翱翔！

故乐为之序！

上海财经大学 国际工商管理学院 博士

河南财经政法大学 工商管理学院 讲师

河南君友商务咨询有限公司 咨询 & 培训部 总监

田启涛

2017 年 10 月 8 日

前言

职场 6 年，见了很多新人，也合作过很多“老人”，大家都在具体的项目中摸爬滚打，却很少去主动思考数据分析到底该怎么做？为什么要这样做？为什么自己的分析结果得不到客户的认可？为什么自己总在无休止地加班？

我们总是在惯性中工作，在被逼无奈的时候去主动学习，在现在的模式下进行思考，在自己狭隘的格局下做决定，在没有达到预期的结果出现后内疚自责，然后日复一日，却从来没有想过去系统地学习和分析“数据”到底有什么样的力量、如果去挖掘“数据”的力量。

所以就有了这一本书，这一本很特别的书。

它很系统，数据分析从 0 到 1 整个流程，它介绍得清清楚楚，看得懂、学得会。

它很实际，所有案例都是从具体项目中提取的，给你真刀真枪，助你闯荡职场。

它很利索，全部是干货，没有一句废话，偶尔来点儿矫情，还是为了让场景更逼真。

它还很接地气，没有晦涩难懂的统计学理论，也不需要高级的代码编程，一个 Excel 足矣，偶尔来点 SPSS 点缀。

万丈高楼平地起，它就是你需要的“地基”。

所以，它不是一本工具书，也不是一本理论书，更不是一本说教风格的职场鸡汤书。

对于不需要的人，它是一本普通得不能再普通的书；但是对于以下这些人，它就是“黄金屋、千钟粟”。

本书适合下列读者：

- 数据分析爱好者。
- 管理学、经济学等专业的在校大学生。
- 初入职场的菜鸟、小白。
- 从事调研、咨询、研究、分析等工作的专业人士。
- 市场营销、人力资源、财务管理等工作中需要进行数据分析的人士。
- 想提升职场竞争力或者正迷茫的职场“老人”。

.....

本书不适合下列读者：

- 对数据完全不感兴趣的人。
- R、Python 等软件的狂热追求者。
- 统计、数学、IT、金融等专业或行业的精英。
- 数据统计、数据挖掘、计算机科学等方面的专家。

构思一年，撰写一年，整整两年，这本书能够顺利出版，现在你能够拿着它细细品读，它又能够让你有所收获和启迪，我想，这就是缘分。

关于这本书，我要感谢的人很多。

首先，感谢所有的同事和客户，如果没有这么多年的项目经验积累，没有跟你们的交流沟通，没有你们的答疑解惑，就不会有书中那么多的经典案例。

其次，感谢公众号“欢乐 PPT”这个大家庭的所有成员。如果没有这个平台，那么我不可能在 Excel 学习和分享的路上坚持到现在。同时感谢好友海域对第 7 章 PPT 设计的支持。

再次，感谢电子工业出版社的编辑，谢谢你们在万千作者中选择了我，并不厌其烦地指导我；感谢为我写序、推荐语、书评的所有人，谢谢你们对这本书的认可及对我的信任。

最后，感谢我的家人，感谢你们理解并支持我的“任性”。

因知识与能力所限，书中纰漏之处在所难免，欢迎及恳请读者给予批评与指正。

作 者

轻松注册成为博文视点社区用户（www.broadview.com.cn），扫码直达本书页面。

- 下载资源：本书如提供示例代码及资源文件，均可在 [下载资源](#) 处下载。
- 提交勘误：您对书中内容的修改意见可在 [提交勘误](#) 处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- 交流互动：在页面下方 [读者评论](#) 处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/33135>



目 录

第 1 章 知己知彼，百战不殆——什么是数据分析 / 1

- 1.1 数据分析的力量 / 1
- 1.2 初识数据分析 / 3
- 1.3 数据分析的前景 / 12
- 1.4 数据分析的六大误区 / 13

第 2 章 稳扎稳打，步步为营——数据输入 / 18

- 2.1 Excel 基础操作 / 18
- 2.2 Excel 数据类型 / 24
- 2.3 单元格逐个输入 / 27
- 2.4 单元格批量输入 / 34
- 2.5 单元格限制性输入 / 39
- 2.6 调查问卷的录入规则 / 45

第 3 章 去粗取精，去伪存真——数据整理 / 49

- 3.1 数据表格的规范化 / 49
- 3.2 单元格的规范化 / 63
- 3.3 数据分列 / 68
- 3.4 数据提取 / 71
- 3.5 数据合并 / 74
- 3.6 删除重复值 / 80
- 3.7 判断错误值 / 83
- 3.8 相同字段的查找匹配 / 86
- 3.9 保护工作表和工作簿 / 97

第4章 抽丝剥茧，刨根问底——数据处理 / 101

- 4.1 “高级筛选”，数据筛选更智能 / 101
- 4.2 “条件格式”，让数据可视化 / 108
- 4.3 数据分析的利器——数据透视表 / 118
- 4.4 3组9个常用的统计分析函数 / 130
- 4.5 筛选与隐藏状态下的数据统计 / 136
- 4.6 最值计算与排序 / 140
- 4.7 身份证号码中隐藏的秘密 / 143

第5章 透过现象看本质——数据分析 / 149

- 5.1 对比分析法 / 149
- 5.2 交叉分析法 / 153
- 5.3 综合评价分析法 / 154
- 5.4 矩阵分析法 / 157
- 5.5 对应分析 / 161
- 5.6 漏斗分析法 / 169
- 5.7 杜邦分析法 / 173

第6章 用图说话，图个明白——数据呈现 / 176

- 6.1 让图表会说话 / 176
- 6.2 基础图表的使用 / 180
- 6.3 Excel 2016 新增图表的使用 / 189
- 6.4 表格设计与美化 / 195
- 6.5 高级数据分析图表 / 203
- 6.6 数据可视化 / 217
- 6.7 图表验收 / 222

第7章 万事俱备，只欠东风——数据分析报告 / 224

- 7.1 认识数据分析报告 / 224
- 7.2 数据分析报告的结构 / 232
- 7.3 高质量数据分析报告的必备要素 / 238

第1章

知己知彼，百战不殆 ——什么是数据分析

为适应市场发展，公司在市场部新增了2个专职的数据分析岗，一直表现比较突出的老员工安迪任职数据分析主管，刚入职的新员工关关任职数据分析专员，这个80后、90后组合在正式合作的第一天就对数据分析相关问题进行了深入的探讨。

1.1 数据分析的力量

关关：安总，现在人人都在谈论数据分析，数据分析到底有怎样的魔力，让企业和个人都为之疯狂，连公司都开始设数据分析岗？

安迪：关关，其实数据分析并不是一个新鲜事物，它对于赢得战争、科学研究、企业发展、个人发展等都具有非常重要的作用。我给你分享3个数据分析案例，你可以从中看到数据分析是有多么的神奇。

1. “战神”林彪活捉廖耀湘

林彪被毛泽东评价为“无与伦比的常胜元帅”、被蒋介石赞赏为“黄埔最优秀的将军”，可以说是当之无愧的“战神”。江湖上关于林彪的战争传闻很多，最有传奇色彩的可能就是“他利用数据分析活捉廖耀湘”了。

自1948年辽沈战役打响后，每天深夜林彪都在东北野战军前线指挥所里听取军情汇报，由值班参谋读出下属各个纵队、师、团用电台报告的当日战况和缴获情况，而林彪则认真地记录着这些数据。林彪敏锐地从1个数据的微小变化中察觉到异样，面对一脸迷茫的部下提出了3个关键性问题：

“为什么那里缴获的短枪与长枪的比例比其他战斗略高？”

“为什么那里缴获和击毁的小车与大车的比例比其他战斗略高？”

“为什么那里俘虏和击毙的军官与士兵的比例比其他战斗略高？”

没等大家反应过来，林彪大步走向挂满军用地图的墙壁，指着地图上的1个点语气坚定地说：“我断定敌人的指挥所就在这里！”在林彪的指挥下，国民党新六军的指挥所很快就被连锅端了，新六军军长廖耀湘——这位出身黄埔并留学法国著名的圣西尔军校、参加过滇缅战役的名将，根本没有想到自己精心隐蔽的野战司令部这么快就被消灭了，当他得知林彪如何得出判断之后，说：“我服了，败在他手下，我不丢人。”

2. 啤酒与尿布的缘分

这是一个已经被证实为杜撰但仍被传为经典的数据分析案例。

20世纪90年代，美国沃尔玛的超市管理人员分析销售数据时发现了一个令人难以理解的现象：在某些特定的情况下，“啤酒”与“尿布”这两件看上去毫无关系的商品经常会出现同一只购物篮中。

这是因为在美国有婴儿的家庭中，一般是母亲在家中照看婴儿，年轻的父亲前去超市购买尿布，父亲在购买尿布的同时，往往会顺便为自己购买啤酒，这样就会出现啤酒与尿布这两件看上去不相干的商品经常会出现同一只购物篮中的情况。

3. “智商与颜值成正比”的美女领导

2016年有一部很火的电视剧《欢乐颂》，五美之一的Andy被谭宗明聘请到自己的公司出任CFO，员工对这位美女年纪轻轻就出任公司的CFO议论纷纷，甚至猜测Andy与谭宗明有不正当关系。一时间，谣言四起（第2集）。

而Andy则在新公司上班第一天的会议上凭借对数据来源抽丝剥茧般的探究、对国内和国际市场的深入了解，让给自己工作找借口的汇报者虚汗如豆，其镇压效果比拍案斥骂还灵验三分。由此，那些会前议论新领导的员工才算明白：智商和颜值是可以成正比的（第3集）。

安迪：关关，这就是数据分析的力量，它要求我们在工作中拒绝思维懒惰，坚持深度思考，通过长期的数据记录和分析，让这些枯燥的数字形成系统化的数据库，一旦出现规律或者偏差，我们便可以及时提炼出准确信息，找出关键问题所在。

1.2 初识数据分析

关关：安总，既然数据分析如此重要，那到底什么是数据分析、怎么做数据分析呢？

安迪：我们先来看看数据分析的定义、作用和分类，然后再来详细梳理一下数据分析的流程。

数据分析的定义：数据分析是指用适当的统计方法对采集的数据进行分析，通过对数据中所包含信息的集中、整理和提炼，找出数据的内在规律，充分挖掘数据的价值，为生产经营活动提供参考和支撑。

数据分析的作用：一切以数据说话，成为当今和未来经济社会发展的趋势。数据最有说服力，在生活和工作中，我们以详尽的数据来证实己方观点、批驳对方言论，最能起到令人信服的作用，依靠相关的数据来制定市场政策也最为科学。例如，国家层面的“GDP”，行业层面的“百度指数”（见图 1.1）“阿里指数”“微指数”等，企业层面的“市场占有率”“客户满意度”等，个人层面的“信用度”等，都是对数据加以分析、得出结论，然后指导实际行动。总之，现代决策离不开数据。

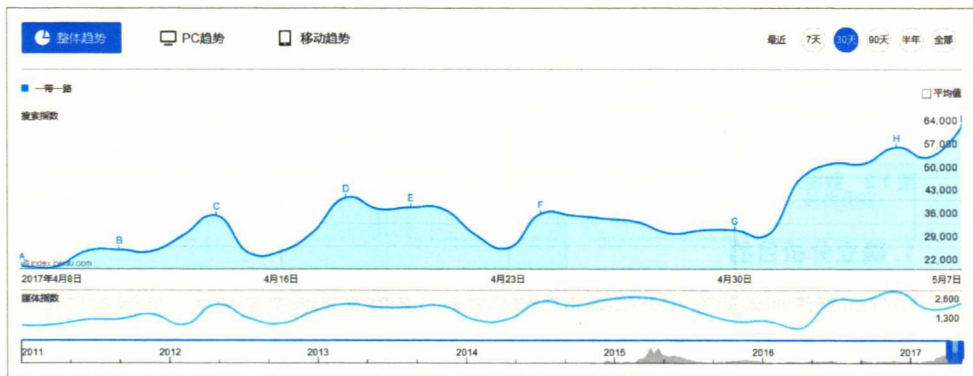


图 1.1 百度指数 - “一带一路”指数

数据分析的分类：在统计学领域，数据分析被分为三类，分别是描述性统计分析、探索性数据分析及验证性数据分析，如图 1.2 所示。

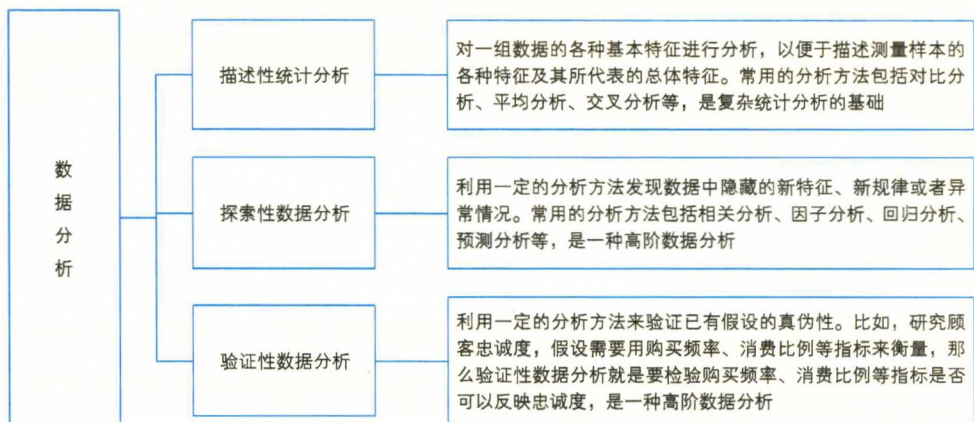


图 1.2 数据分析的分类

数据分析的流程：数据分析有一个标准化流程，从确立分析目的到最终完成分析包括 9 个步骤，每个步骤的质量都对最终数据分析结论正确与否、价值大小起着至关重要的作用，如图 1.3 所示。



图 1.3 数据分析的流程

1. 确立分析目的

做任何事情都需要有一个明确的目的，这个目的既是动力又是方向，数据分析也不例外。在每个数据分析项目开始前，应该明确为什么要开展数据分析？通过数据分析要解决什么问题？确立了分析目的后，每一步都围绕分析目的来进行，既不需要考虑分析模型设计得多么高级，也不需要考虑 Excel 图表或者 PPT 报告做得多么炫酷，只要实现了分析的目的，此次数据分析结果就有指导意义。

2. 明确分析思路

确立了分析目的后，就应该系统梳理并明确分析思路了。目的是终点，要想到达终点就需要一条路径，这条路径就是我们的分析思路，即数据分析的理论和实践支撑、需要从哪几个角度进行分析、采用哪些分析指标等。

数据分析思路需要以管理理论为指导，结合业务实际，搭建分析框架，只有这样才能确保数据分析维度的完整性和分析结果的可靠性。如图 1.4 所示，TCSI 模型是国家工业和信息化部推荐的电信行业用户满意度指数模型，在进行某电信企业用户满意度调研时，就可以以该模型为基础进行优化，既具有普遍的适用性，又具有该企业的针对性。

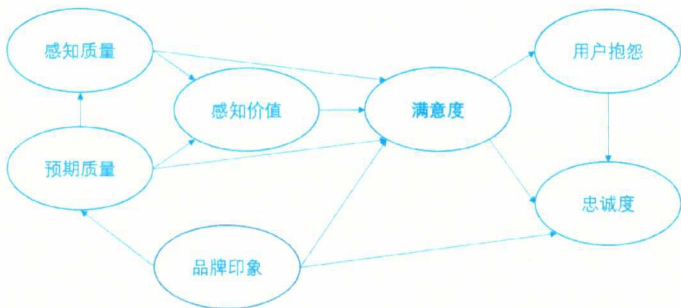


图 1.4 电信行业用户满意度指数 (TCSI) 模型

再比如，进行网站流量分析，必须包括 5 个方面的内容：转化率分析、来源分析、页面数据分析、访客分析和优化分析，如图 1.5 所示。

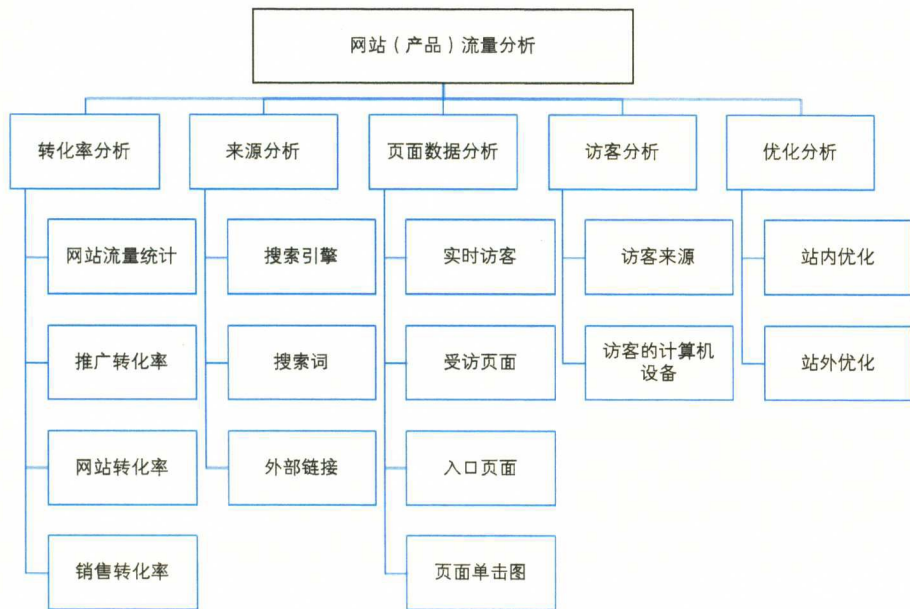


图 1.5 网站（产品）流量分析框架

明确分析目的和分析思路是确保数据分析结果有效的先决条件，它让我们的数据分析不是为了分析而分析，而是为了实现一定的目的并且采用了合理的分析框架而进行的分析。

3. 数据采集

确立了分析框架后就需要采集相关数据了。从数据来源来看，数据主要有一手数据和二手数据两类，一手数据主要是指直接获取的原始数据，二手数据主要是指经过加工整理后的数据。

数据采集方式主要有4种：公开出版物、互联网、数据库系统和市场调查，如图1.6所示。

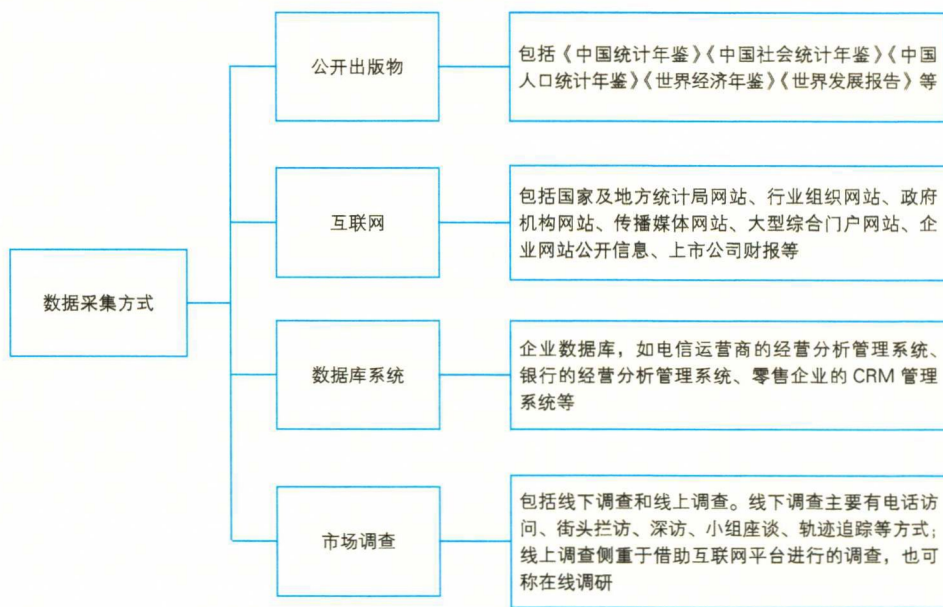


图 1.6 数据采集方式的分类

这里需要引入一个“大数据”的概念。大数据（Big Data）一般是指需要新处理模式才能具有更强的决策力、洞察力和流程优化能力的海量、高增长率和多样化的信息资产。维克托·迈尔-舍恩伯格及肯尼斯·库克耶编写的《大数据时代》一书中指出：大数据是指不用随机分析法（抽样调查）这样的捷径，而采用所有数据进行分析处理。因此，我们可以认为通过“互联网”“数据库系统”等方式采集的数据偏向于“大数据”，而通过“市场调查”等方式采集的数据侧重于“小数据”，即抽样调查数据。

关关：现阶段人人都在谈论“大数据”，就像多年前的“物联网”这个概念一样。我也有疑问：“大数据”与“小数据”最本质的区别是什么呢？到底是炒作性的概念，还是真的有其重要价值呢？

安迪：自“大数据”的概念诞生以来，人们关于“大数据”与“小数据”的探讨不绝于耳。但不可否认的是，“大数据”的价值是不可估量的，我给你讲述一个案例。

2010年10月23日，英国《卫报》发布了一条伊拉克战争日志的数据新闻，利用维基解密提供的数据制作了一幅包含约39.1万条信息的红点数据地图，来描述伊拉克战争中所有人员的伤亡情况，如图1.7所示。图上的每一个红点代表一次伤亡事件，用鼠标单击，则会列出时间、伤亡人数和原因等更详细的信息。没有煽情的语句，没有激烈的言辞，只有沉默的图表，但是战争的残酷却展露无疑。这条新闻一经发布，立即引起朝野震动，推动英国最终做出撤出驻伊拉克军队的决定。

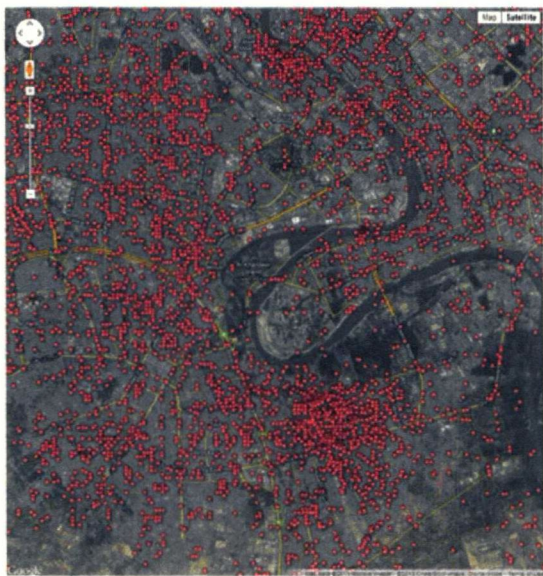


图 1.7 伊拉克战争伤亡情况

安迪：这就是“大数据”带来的震撼，这样的震撼是“小数据”远远达不到的。“大数据”与“小数据”究竟有什么区别呢？我们一起来看一下。

“大数据”与“小数据”的一个明显的区别是，“小数据”侧重于对因果关系的探求，而“大数据”更加关注事物之间的相关关系，也就是从追求“为什么是这样”到追求“世间万物的链接规则”的转变。这种转变颠覆了千百年来人类的思维惯例，对人类的认知和与世界交流的方式提出了全新的挑战。

除了这种观念的转变，“大数据”与“小数据”还有一个重要的区别在用途上，“小数据”在很大程度上停留在过去的状态，用过去的数据证明过去，发现过去存在的问题及为什么会存在这个问题；而“大数据”的核心就是预测，为人类的生活创造前所未有的可量化的维度，使数据从原来停留在说明过去转变为驱动现在甚至未来。

安迪：其实，“大数据”与“小数据”在数据分析工具、数据分析方法、数据解读与呈现等方面也有很多不同，在此就不一一展开了，在以后的工作中我们会慢慢接触到、感受到，今天我们有一个概念性的认识就行了。

关关：那我们市场部的工作是以“大数据”为主还是以“小数据”为主呢？

安迪：这就需要具体问题具体分析了。比如，我们要做市场规模预测，可能需要用到“大数据”；如果需要了解客户需求、客户满意度、客户忠诚度、产品及服务方面存在的问题等，就需要用到“小数据”。

关关：嗯，我明白了，安总。不局限于是“大数据”还是“小数据”，最重要的是能不能实现我们的“分析目的”。

安迪：是的，总结得非常好！我们的工作也会从“小数据”开始，一是因为“大数据”的噪声太多，你暂时还没有消除噪声的能力；二是因为我们身边的“小数据”相对容易获取。相信这7天你就能真正入门咱们的数据分析了。

关关：我一定好好跟安总学习，争取7天后可以独立做数据分析。

安迪：介绍完“数据采集”，我们就进入“数据输入”这个阶段吧。

4. 数据输入

随着企业信息化的发展，在日常工作中需要手动输入的数据越来越少，一般都是从其他软件系统导出的。但是我们会遇到手动输入的情况，如拦截访问的数据，以及临时需要补充的数据。这时候，我们就需要了解分析软件的特点，思考如何将纸质数据科学地录入电子表格。