

市场调研

应用R软件

周庭锐 著



市场调研

应用R软件

周庭锐 著



中国人民大学出版社

· 北 京 ·

图书在版编目 (CIP) 数据

市场调研: 应用 R 软件/周庭锐著. —北京: 中国人民大学出版社, 2012. 9
(管理者终身学习)

ISBN 978-7-300-15859-4

I. ①市… II. ①周… III. ①市场调研-应用软件 IV. ①F713. 52-39

中国版本图书馆 CIP 数据核字 (2012) 第 213918 号

管理者终身学习

市场调研: 应用 R 软件

周庭锐 著

Shichang Diaoyan: Yingyong R Ruanjian

出版发行	中国人民大学出版社		
社 址	北京中关村大街 31 号	邮政编码	100080
电 话	010-62511242 (总编室)		010-62511398 (质管部)
	010-82501766 (邮购部)		010-62514148 (门市部)
	010-62515195 (发行公司)		010-62515275 (盗版举报)
网 址	http://www.crup.com.cn		
	http://www.ttrnet.com (人大教研网)		
经 销	新华书店		
印 刷	北京东方圣雅印刷有限公司		
规 格	175 mm×250 mm 16 开本	版 次	2012 年 10 月第 1 版
印 张	18.25 插页 2	印 次	2012 年 10 月第 1 次印刷
字 数	321 000	定 价	45.00 元

版权所有 侵权必究

印装差错 负责调换

这本书想带给读者一种完全不同的关于市场调研的学习体验。市场调研是营销学科里一个实践性非常明显的领域，企业通过它来认识自己据以生存的环境，包括政治、经济、社会、行业、竞争对手、顾客以及其他利益相关者的种种意图、行为，甚至隐匿在意图与行为背后的思维逻辑。对于这么“接地气”的一门学问，我们在坊间所见的课本、书籍，却往往令人产生“蜀道难，难于上青天”的无奈感慨，令初窥门径的学子，以及没时间深入钻研的专业经理人望书兴叹，无法具体掌握这门学问的实践方法。

这真是非常可惜的事。一方面，学习的目的就是要学会、学懂，悬在空中的学问再了不起，也与庶民无关；另一方面，学以致用才是实学，中国企业想在竞争激烈的现代市场里胜出，很有必要精通前沿的市场调研方法，来有效掌握市场信息，通过自己的真本事（而不是各种制度规范与行政保护）击败竞争对手，优雅地（而不是无赖地）登上世界级的商业舞台。

市场调研之所以容易陷于窒碍难行的困境，主要是这门学问的学科交叉性使然。首先，市场调研的分析手段高度依赖统计学，而单单统计学这个学科里的数学公式、种种统计分布、复杂的概率计算等，就足以让人望而却步。统计学是市场调研的基础，但是在介绍市场调研的书籍里，不应该到处充斥统计学。这两门学问应该是分工互补而又相互独立的，至少在我个人的认识里，市场调研更趋向于应用，读者只需要知道如何适时适地地运用统计方法就好，而完全不必变成统计专家。

其次，市场调研在方法论上很接近严谨的学术科研，而学术科研从理论构造、模型设计、假说界定，乃至样本选择、抽样方法、变量的信度效度检查、分析手段等，都非常讲究。为了保证研究的严谨性，所设计出来的研究方法可能成本高昂，甚至旷日费时，而企业不可能为了保证学术的严谨性，而增加不可负担的额外成本，或者牺牲稍纵即逝的商机。如何在学术严谨性与实践的可行性之间取得适当平衡，是撰写市场调研书籍的一大挑战。

再次，市场调研是工具性学科，必须直接为行业所用，而在不同行业里我们

所面对的不同企业个体，其数据特征可能是不一样的，所能使用的调研手段与统计方法必须因地制宜，进行必要调整来适应行业的特性。从撰写市场调研书籍的视角看，这也是对作者的一大挑战，因为大多数作者很难同时掌握多种行业知识，更难以同时精通行业里不同企业个体的特殊情况，在撰写市场调研书籍时，往往顾此失彼，或者更严重点，只在云端空谈理论与方法，而不能落下凡尘，与行业实际接轨，这就为读者带来了理解书籍内容的困难。

最后，在现代市场发展的进程里，用来记录消费者行为与厂商决策的工具越来越复杂，也越来越高效，传统上使用的深入访谈或问卷调查的市场调研手段，已经不能满足厂商深入理解它所面对的高度复杂的市场的需求。在现代商业环境里，许多商业决策与消费信息，已经被自动截取并存储在各式各样的信息系统里，这些信息大多以数据库数据的形式留存下来，结构严谨，层次分明，数量庞大，而且积累的速度惊人。

例如，零售终端的 POS 收银数据、网页服务器里消费者的网页浏览记录、谷歌地图里的商圈资料、电子商务环境里顾客的论坛讨论与回应，甚至与客户交易往来的电子邮件内容，这些都是厂商认识它所处市场环境的重要信息来源，但是从这些来源所获得的信息，无法通过传统市场调研手段进行处理与解读，面对这些数据，我们需要的是数据挖掘，以及对大数据的统计分析能力，就这个需求来看，目前坊间还没有任何一本市场调研的书籍对此有过任何介绍与讲解。

本书的目的是希望能尝试性地为上述这些困境寻找到一条出路。我的自我期许是，撰写一本“接地气”的市场调研书，让读者阅读之后就能亲自动手做，而且做出来的调研，可以解决中国市场里的许多实际问题。这就是包括明茨伯格在内的许多世界级大师不断提倡的“匠意” (craftsmanship)。“匠意”演变到今天，已经不再是孔子时代的“君子不以小道试之”，而是真正能够富国利民的实学。我希望我的文字能够精确地将实学传达到读者手脑并用的实践里。

之所以能够如此自许，或许是因为我一直是个幸运的人，我既有机会浸淫学术，接受世界一流大学的商学教育，亲近并请益列身学术前沿的良师益友，不同行业里的诸多商界前辈，又愿意赐予我机会，参与他们胼手胝足的商业实践。自我本科时参与人生里第一个企业咨询项目以来，25 年里，除了读研、读博、从英国回台湾创立台湾高校里第一个市场营销系，继之又往来于澳大利亚、新加坡、香港、台湾以及大陆之间担任高校教师，有幸的是，在游走南北半球、体验教学科研的丰富人生之余，还能与企业保持紧密联系，通过对企业提供咨询服务，得以贴近市场，聆听企业的心声。

我真诚相信，只有将理论与实践充分结合，学问才开始发挥作用，才开始为人类作出真实的贡献，也才开始萌生它本来应有的价值。学习市场营销正应如此。市场营销的土壤就是市场，一切脱离真实市场的营销都是空谈。可惜大部分学校只愿把自己当做殿堂，把人锁在高高的阁楼里，除了故纸堆，四壁萧然，久而久之，我们慢慢就遗忘了青草绿地，遗忘了来处，记不起来混着牛羊粪便味儿的花香，究竟该是什么样的气味？

想要与现实接轨，就要把心门打开，让四季的风都吹进来。走到户外，看看我们在阁楼里读到的知识，可以如何增进或者改善我们的生活，看看我们学习到的市场调研方法，究竟在如何解决营销里的真实问题，而这正是我撰写这本书的初衷。在此非常感谢中国人民大学出版社，慷慨地提供这么一个了却初衷的机会。

本书是中国人民大学科学研究基金（中央高校基本科研业务费专项资金资助）项目成果（项目批准号：10XNI012）。

周庭锐

2012 年于北京

第 1 章 市场调研的本质 1

1.1 从一个故事谈起 2

1.2 国内市场调研的现状 3

1.3 市场调研的本质 5

1.4 市场调研的现代意义 6

第 2 章 市场调研必备的基础统计知识 9

2.1 抽样 9

2.2 变量与测量 14

2.3 题项的信度与效度 19

2.4 数据的统计分布 26

2.5 在 R 软件里面的统计分布 32

第 3 章 市场调研在调研些什么 37

3.1 调研策略的选择 37

3.2 对调研目标的初步探索 40

3.3 调研方案的设计 47

第 4 章 观察、探索与证明 51

4.1 观察 51

4.2 探索 68

4.3 证明 76

第 5 章 更犀利的观察 80

5.1 在时空里观察 82

5.2 分类与排序 89

5.3 相关与因果 97

5.4 表现形式 107

第 6 章 更有效的探索 110

6.1 深入访谈 111

6.2 焦点座谈 114

6.3 其他的定性资料收集方法 118

6.4 处理定性资料 118

6.5 定性资料的统计分析 127

第 7 章 证明你的猜想 134

7.1 收集定量数据 137

7.2 定量数据的统计分析 168

第 8 章 真实世界里的市场调研 194

8.1 真实世界里的调研问题 194

8.2 非正态分布的数据形式 197

8.3 广义线性模型 199

8.4 其他的类别数据分析方法 210

8.5 零售环境里的市场调研 218

附录:关于 R 软件 251

后记 283

市场调研的本质

其实我们每天都在进行着市场调研。清晨赶早市的人们嘴里问着菜价，随手“检阅”着青翠欲滴的小黄瓜，这是市场调研；我们想添购一台笔记本电脑，除了问问周围朋友们的意见外，还可能上网搜一下，研究研究最新的机型，比较一下不同品牌之间的性价比，甚至读读论坛上网友们的评价，这是市场调研；当我们旅行来到一个新的城市，才下飞机，便迫不及待地打开手机，马上在诸如微领地、人人网等 Location-Based Service (LBS)^① 上面签到一下，看看所在地附近哪里有好吃好玩的，这也是一种市场调研。

市场调研不是厂商的专利，市场调研是人类获取环境信息、理解生存条件、判断未来趋势，甚至满足好奇心的一种天性。这种天性与生俱来，历史悠远，连《西厢记》里崔莺莺想要自荐枕席之前，都不免要派上小红娘事先打探张生一番。愈到现代，科技的力量愈是无远弗届，人们打探事情的功力一日千里，调研的技术含量与日俱增，通过调研所得来的信息质量也愈来愈精细准确，于是市场调研逐渐变成了一种专业。

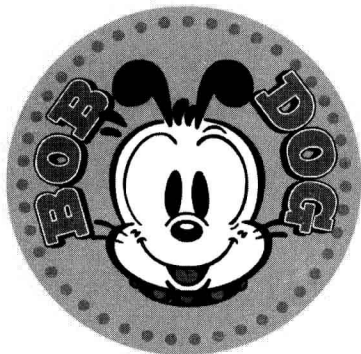
专业的市场调研正在日益明显地影响到我们每一天的生活。例如刚刚说到的 LBS 签到，消费者在搜寻身边的各种消费信息的同时，提供这些服务的厂商，也正虎视眈眈地关注着消费者每分每秒的行为。现代科技提供给厂商更便利的渠道

^① Location-Based Service 是基于地理位置的服务，它是通过移动运营商的无线电通信网络（例如 GSM，CDMA 等）或其他外部定位服务（如全球定位系统 GPS），来获取移动终端用户的位置信息（地理坐标），然后在地理信息系统（GIS）平台的支持下，为用户提供各种相应服务的一种移动增值业务。

来观察消费者行为，同时还自动将这些信息存储于以电脑数据库为主的存储媒介中，以方便后续的分析与统计。许多 LBS 甚至还提供社群网络的功能，鼓励消费者与消费者之间互动，因此厂商所观察到的消费者行为还不仅仅是单一消费者的行为，而是整个社群消费者的互动与行为偏好，我们会在本书的部分章节中探讨这种形式的调研与分析。

1.1 从一个故事谈起

市场调研的专业性并不一定只能依托在最新的现代科技上。上海巴布豆儿童用品有限公司^①的创办人林先生曾经是我的邻居，他告诉过我关于这家公司如何进军上海的故事。巴布豆（LITTLEBOBDOG）其实是日本向阳株式会社在 1987 年注册的卡通造型品牌，林先生大约在 1990 年前后取得了该品牌在大中华区的总代理权，中国台湾地区是他建立巴布豆品牌的第一站，积累了一些经验后，他在 1994 年初来到上海，准备在神州大地大展拳脚。



开业前为了认识这个新市场，他选择住在上海的希尔顿酒店，而且一住就是半年。那个期间他的生活方式单一固定：每天带着地图和计数器逛街、逛商场。他首先将上海分成若干大区，然后标定商圈，选择想要观察的商场，排定每天的逛街路线，定时定点在路口或商场里计算路过客人的类型和数量。据他说，半年时间里，磨穿了 5 双皮鞋。我曾看过当时他那本又厚又脏的笔记本，由日期、时间、天气、地点、商圈类型、客人类型、客人数量等变量组成。

他甚至毫不避讳地询问商场专柜的每日营收。当时人心淳朴，“一位商场的售

^① 公司网址：<http://www.bobdog.com.cn/index.html>。

货小姐居然打开收银机，当着我的面，拿起成叠的钞票，直接数清当天实际营收的金额给我看！”林先生说。很显然，林先生是希望能够将商铺位置、商圈状态、消费者特征和商铺的营收结合在一起，作为将来自己选择开店位置的判断依据。换句话说，在他心里是存在一个针对这个调研设计的研究模型的：将商铺营收作为因变量，将商铺位置、商圈状态、消费者特征等作为自变量，然后进行模型估计。

林先生并没有接受过很完整的统计科研训练，因此即使收集了这些宝贵的第一手资料，也没能进行什么深入的统计分析，但是无疑这些资料对于他开展大陆事业帮助很大。他是我所见过的在大陆进行实体投资的台商中，最迅速获得投资回报的商人，同时，他也是我所认识的商人里，愿意拿出半年时间，脚踏实地亲自进行市场调研的唯一一人。

尽管进行了这么深入的市场调研，他在大陆的事业也并非一帆风顺。大陆现阶段的市场经济里存在许多难以想象的“非市场经济”潜规则，这是在台湾长大的林先生所无法想象的。同时巴布豆这个品牌也不断遭遇廉价仿冒品的挑战。更糟糕的是，本地的法律不一定保护正品^①，巴布豆在大陆的真正威胁可能还不是正规的商业竞争，而是大环境里无声无息的某些潜规则。

这个故事告诉我们，市场调研的对象，不应该只是消费者而已。如果林先生能够正确认识市场调研的本质，同时关注市场调研内外几个不同的层面，他在大陆的事业会发展得更好。实际上，外部环境永远是内部循环的最大约束力量。最入门级的战略学里有个分析模型叫做 PEST，主张从政策面（policy）、经济面（economy）、社会面（society）和技术面（technology）分别来解读企业所面对的外部环境，如此一来，市场调研的对象就同时包含了微观的企业内部运营循环和宏观的外部市场环境，这是比较完整的市场调研概念。

1.2 国内市场调研的现状

其实，企业每天都在进行着市场调研。传统上许多企业常会设置一个称为

^① 例如 2009 年福建晋江万泰盛鞋服有限公司使用“巴布狗”（BABUDOG）商标侵犯“巴布豆”（BOBDOG）品牌注册商标，连商标图案都十分近似（同一只卡通狗造型，但是身体动作不一样），这个案例是否明显侵权由读者自行判断，我们国家的工商行政管理总局商标评审委员会判定侵权不成立，人民法院支持商标评审委员会的判定。见北京市第一中级人民法院行政判决书（2010）一中知行初字第 592 号。类似的案例还有 2012 年的深圳唯冠起诉苹果 iPad 的故事，这里就不详谈了。

“政策研究室”的部门，专责解读国家和地方政府出台的政策。由于国内的商业现实更多的是国家宏观政策决定企业的商业机会，所以这样的部门在某种程度上就是一个专责市场调研的单位，为企业制定战略起到非常重要的作用。同时，大部分企业还特别重视对竞争对手的情报收集工作，据我所观察到的案例，有些收集对手情报的手段还相当极端，从派人潜伏渗透到恶性挖角无所不用其极。我想，大部分企业确实明白竞争信息的价值。

除此之外，大部分中国企业在现阶段并不特别费心收集消费面的信息。这是因为这个市场还处于商业勃兴的初级阶段，市场供给还达不到全面饱和，外延式的增长机会远远超过对本业精耕细作所带来的利益。在本业上，基本上只要产品能够以低成本生产出来，加以足够的推广刺激，供给本身就可以创造相当大的市场需求，因此对消费者的深入理解还没能够显现足够突出的商业价值，而对于国家政策的正确解读，对于竞争动向的合理掌握，可以直接生成非常明显而及时的商业利益。

然而，这种情况很快就会改变了，中国已经成为全世界大小厂家的竞技场。此外，市场中的舆论监督力量，也随着各种新旧媒体的兴起与发展逐渐壮大起来，过去不透明的行政垄断力量正在悄悄消退中，自由竞争的市场机制会越来越成熟。因此，无论中国本土厂商高兴不高兴，都必须面对这样一个事实：未来的市场竞争只会越来越激烈、越来越知识密集，及时而有效的市场信息也会越来越重要，最终将成为决定商业胜负的最关键因素。

因此，实践市场调研的能力会变成商业运作的基础之一。这样的趋势其实已经很明显了，宝洁中国的市场研究部（Consumer and Market Knowledge, CMK）就是其中最知名的一个例子（事实上，宝洁的 CMK 部门被业内人士称为“中国市场调研行业的黄埔军校”！中国许多市场调研公司的创办人、合伙人或者高级调研主管都是从这里历练出来的）。

宝洁 CMK 部门的三个关键词分别是：Hindsight, Insight, Foresight。其中 Hindsight 的含义是通过数据挖掘来整理行业竞争的历史动态，理解消费者在态度与行为上的偏好，分析竞争对手的行为模式，进而掌握整个市场的现况与发展规律。Insight 的含义是通过上述分析，结合公司市场部、财务部、销售部、产品供应部、研发部等各个部门对市场的局部见解，来统合出更全面、更前瞻的市场观点，以便指导公司未来新产品与新事业的发展。Foresight 是指基于前两者的分析与归纳，协助公司更科学、更正确地预测未来，为公司的未来发展提供更可靠的方向性指引。

宝洁的例子很好地总结了企业进行市场调研的基本方法与手段。宝洁打的其实是个正规的、现代化的商业战争，而不是通过某些官本位的垄断性资源来进行一个不公平的商战。当然，如果能够取得某种官本位赏赐的垄断性资源也是种本事，在这个前提下，市场调研自然就不那么重要了，但是官本位赏赐的资源毕竟很有限、不稳定，而且政治风险可能很高。无论如何，本书的立意前提是自由竞争的市场经济，在这个前提下，市场调研对于企业的成长与壮大非常重要。

1.3 市场调研的本质

市场调研的原始产出是资料。通过市场调研手段所取得的资料还需要加工处理，才能够用来指导企业的未来发展。通过市场调研手段取得的市场情报通常是以资料（data）的形式存在的，这些资料的表现形式可能是数字、文字、图像，甚至是影音形式等，其中直接保存为数字形态的为数据记录（record），而能够存储于类似数据库（database）、試算表（spreadsheet）或文本文件（例如 CSV 格式）里的，称为结构化资料（structured data），不能直接以数字形式存储并进行内容检索比对的（例如声音、影像、文稿等），则称为非结构化资料（unstructured data）。

资料经过适当的加工（process）后，被人赋予意义的称为信息（information）。正由于信息是有意义的资料，而赋予资料意义的又是不同的人，所以信息具有专有性（exclusiveness）的特质，即相同的资料，对于不同的人，其意义是不一样的。换句话说，相同的客观资料，基于不同的主观需要，可以获得种种不同的解释，这是市场调研的基本性质之一。市场调研的目的是服务于对它有需要的人群，因此对于资料的解读，自然包含了特定人群的主观需要。

到了信息的层次，市场调研的工作已经基本算是功德圆满了，当然信息还存在一些更高的阶段。一般认为，信息为人们所消化吸收，进而影响人们的行为，则成为知识（knowledge）。知识是信息的实践。站在市场调研的立场，市场调研的分析报告是需要被人“赏识”的，我们需要通过种种沟通与表达的技巧，让市场调研的成果很容易地被传递与理解。这些信息只有被正确接收了，才对人们的决策发挥作用，也才达到了进行市场调研的初始目的。因此，信息的表达方式对于信息能否提升到知识的高度是很重要的，我们需要学习有效整理与表达信息的各种技术。

信息的最高阶段大概就是智慧（wisdom）了。个人是服膺老子的见解：“为

学日益，为道日损。损之又损，以至于无为，无为而无不为。”在知识的层次上，有意义、有结构的信息越多越好，可以增加判断的依据；在智慧的层次上，信息量却是越少越好的。丝竹乱耳，案牍劳形，太多的信息反而增添直觉的负担，压抑了人类的创造力，为清明智慧制造了莫大麻烦。因此，信息需要过滤，需要梳理，需要归纳，需要收敛，才能为有智慧的人类提供最大的帮助。

1.4 市场调研的现代意义

市场调研的主要研究领域是讨论如何通过各种手段，以最高效的方式，收集各种有效资料，整理并分析资料以形成有意义的信息，最终达到创造特定知识的目的。收集资料的手段，除了大家都很熟悉的文献阅读、二手资料收集、深度访谈、焦点座谈、参与式访谈以及问卷调查外，由于现代信息技术的进步，还可以通过各种手段自动收集消费者在消费过程中不知不觉遗留下来的行为轨迹。例如销售时点系统（POS）里的消费记录、信用卡消费记录、各种付款记录、门店里无线射频（RFID）技术的消费行为采集、互联网上的网络浏览记录、论坛发帖、回应与讨论、微博互动、各种文件的上传和下载、网络点击、LBS 签到，甚至网络爬虫^①等。

在现代生活中，那些被自动记录下来的行为轨迹，事实上更能够正确地反映我们的真实生活。可以这么说，我们无所遁逃于这些在我们的生活中自然存在的记录工具或者仪器，而与此同时我们通常却不愿意诚实地回答厂商的问卷调查或者访谈。听过这么一句话：“人们习惯用假名说真话，用真名说假话”，网络上的匿名留言最能真实地反映这一代人的生活。从商业视角来看，厂商对于顾客最感兴趣的有四种信息，但是这四种信息不具备同等的战略价值：（1）顾客心知肚明，并且愿意告诉厂商的信息；（2）顾客心知肚明，但是绝对不愿意与他人分享的信息；（3）顾客心知肚明，却说说明白、讲不清楚的信息；（4）连顾客自己也不清楚的愿望或需要。

通过各种访谈与问卷调查我们或许能够掌握上述四种顾客信息中的第一种，但是拥有这种信息并不能为我们带来任何竞争优势，顾客愿意告诉我们的事，通常不会是秘密，这种信息如果我们可以获得，竞争对手同样可以获得。反倒是后

^① 网络爬虫（又称为网页蜘蛛或网络机器人），是一种按照预先设定的规则（通常是某种正则表达式）所撰写的电脑程序或脚本，能够自动地在互联网里游荡，沿着网页里的链接，一层又一层去穷尽万维网里的网页，并抓取其中的内容。所抓取的内容通常还可以自动进行索引，以方便后续的查询或进一步加工。

三种信息，十分具有战略价值，但都不是那么容易获得的。对后三种信息，厂商通常只能通过实验设计来侦测顾客的真实心意，或者建立消费者行为模型，通过对大量顾客的行为轨迹进行挖掘，从顾客不小心留下的蛛丝马迹中推测出那些顾客不愿说、说不出，或者连自己都不清楚的信息。

从这个角度看，市场调研确实是个技术活儿。就像名侦探柯南^①一样，除了需要大量的事实资料作为分析依据外，还需要无拘无束的洞察力、逻辑清晰的思辨能力，对大量资料进行抽丝剥茧，最后真相大白，水落石出，解开商业世界里的种种谜团。然而那些记录消费者行为轨迹的信息量实在太大了（记录下来的是成千上万人的分分秒秒的思考与选择），或许称之为海量或巨量数据也不为过，柯南式的、依赖人脑进行数据分析已经完全失去了可能，而通过电脑进行数据分析已经变成了一种必然。这些数据的形态大大不同于通过问卷采集来的数据，导致一般的多元统计方法无法适用，这在本书中会专章单独加以讨论。

本书希望能够忠实地展现目前在市场调研方面的世界学术前沿，同时又希望能够避免教材式的枯燥理论宣讲；本书希望能够有机结合中国市场中的现实状况，同时又希望能够涵盖未来中国市场面临的挑战与需要。因此，本书的内容其实超越了目前坊间大部分的市场调研书籍。在本书里，除非必要，我们不具体讨论所涉及方法论里大部分的统计模型，也不进行任何公式推导与理论证明，而是很实际地讨论了真实世界里可以加以应用的统计或数据分析方法，包括一些被认为很前沿，只有学院派的学者才会使用的方法。

我真正想证明的是，这些所谓学院派的方法其实是真正实用的数据分析手段，在实践上不仅具体可行，而且实施起来相当简单。为了协助读者更容易地学会这些工具与技巧，在本章阐述了市场调研的本质之后，在第2章将为读者复习一些在市场调研过程中必备的基础统计知识，包括母体、样本、抽样、变量的测量与信效度检查、描述统计、推论统计以及基本统计分布等概念，这些内容更多的是为了定量调研的统计分析预做准备的，也是后面更高深议题的基础。

当然，如果读者的统计知识已经齐备，自然可以略过第2章，直奔从第3章开始的市场调研各大主题。从第3章开始，本书首先讨论市场调研究竟在调研些什么，想要达到哪些目的，以及在进行调研之前应该做的一些前期准备工作，例如，调研方案的设计、次级资料与文献的收集。然后我们在第4章初步介绍进行

^① 日本漫画家青山刚昌以侦探推理的情节为主题的漫画作品，剧情中以化名为“江户川柯南”的高中生侦探工藤新一为主角，号称“日本警察的救世主”、“平成年代的福尔摩斯”，很多离奇的案件通过他活跃的逻辑思维得以侦破。

市场调研的几种基本策略，在后续的章节里逐一加以探讨，亦即：观察（第 5 章）、探索（第 6 章）、证明（第 7 章）。

我们会在“观察”的章节里着重介绍图表、数字的表达方式，让观察的视角更加敏锐，让所观察到的事实表述更加清晰明白。接着在“探索”的章节里，集中阐述了定性研究的基本方法，并强调了探索的目标就是发现，而不在于证明。最后在“证明”的章节里主要探讨定量研究方法，从猜想的形成、模型的构造、问卷设计、抽样的实施，到常用的统计分析方法，都进行了比较深入的介绍，并强调证明的目的就是证实猜想的正确性，因此定量研究的前提是必然事先已经存在某些猜想。

为了便于读者边读边学，本书一开始就把数据处理、统计分析技术与市场调研技巧融合在一起。目前国内关于市场调研的书籍与课程，都太关注于对市场调研概念的介绍，而忽略了对动手做的实际指导。本书作者认为，只有亲自动手操作，才是学会市场调研的不二法门（其实也是学习乐趣的真正来源）。可是既然要动手做，那么数据处理与统计分析就是个必经的过程。我们进行统计分析所采用的工具是威力强大而且免费自由软件——R 软件，所有的统计分析脚本都附在章节之中。此外，在本书最后的附录里对 R 软件这个统计编程语言的语法做了一个摘要式的整理。

第 8 章探讨进阶的市场调研议题，尤其是那些违背正态分布假定的数据结构。例如，互联网里面的点击、下载，博客、论坛里的文字或数字资料的抓取，零售门店里的 POS 数据，商场里结合 POS 数据的问卷调查等。这些才是消费者每天的真实生活，但是面向这种真实生活所收集到的数据，多半不服从我们所熟悉的正态分布形式，因此就不能使用常见的多元统计分析方法了。在第 8 章里通过真实生活案例来讨论一些非正态统计分布的数据分析方法，让读者能真正深入实践中去学习市场调研。

市场调研必备的基础统计知识

工欲善其事，必先利其器。我们在这一章里简单复习一下进行市场调研必备的一些基础统计知识。因为是简单复习，所以不会是完整的统计学知识的教授，对统计学特别感兴趣的朋友，或者对统计学的掌握特别没有信心的朋友，建议还是找本优秀的统计学教材，好好地自行研读一番。

在本质上，我们进行市场调研其实完全是可以不懂统计学的。有些市场调研高手独自面对市场，就靠察言观色、望闻问切一番，已经基本把握了市场大概的动向，并不需要精密的数理统计分析，但是这种高手很少，即便有，也只是一人一时独领风骚，既做不成榜样，也不可持续。真正可以长期行之有效的市场调研手段，还是规规矩矩地收集信息，然后进行缜密的数字化分析。

2.1 抽样

谈到收集信息，就必然想到抽样（sampling）。抽样是指通过某些方法，从母体（population）中抽取足够数量的样本（sample），作为我们真实观察的对象。然后针对某些我们特别想要观察的变量，对这些样本进行测量，并以测量的结果来推论母体在这些变量上的情形。具体来说，其实母体才是我们所想要调研的对象，但是由于母体的规模通常过于巨大，我们在经费上、精力上，甚至时间上力有未逮，只好放弃对母体进行普查的想法，改而选取具有代表性的样本，作为真实观察与分析的对象。