



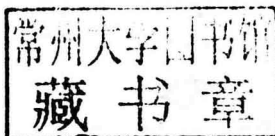
◎ 邵贵平 主编

北京理工大学出版社

BEIJING INSTITUTE OF TECHNOLOGY PRESS

网店数据分析

主 编 邵贵平



 **北京理工大学出版社**
BEIJING INSTITUTE OF TECHNOLOGY PRESS

版权专有 侵权必究

图书在版编目 (CIP) 数据

网店数据分析/邵贵平主编. —北京: 北京理工大学出版社, 2017. 12
ISBN 978-7-5682-4986-7

I. ①网… II. ①邵… III. ①网店—运营管理—数据处理—教材 IV. ①F713.365.2
中国版本图书馆 CIP 数据核字 (2017) 第 282442 号

出版发行 / 北京理工大学出版社有限责任公司

社 址 / 北京市海淀区中关村南大街 5 号

邮 编 / 100081

电 话 / (010) 68914775 (总编室)
(010) 82562903 (教材售后服务热线)
(010) 68948351 (其他图书服务热线)

网 址 / <http://www.bitpress.com.cn>

经 销 / 全国各地新华书店

印 刷 / 三河市天利华印刷装订有限公司

开 本 / 787 毫米 × 1092 毫米 1/16

印 张 / 19

字 数 / 523 千字

版 次 / 2017 年 12 月第 1 版 2017 年 12 月第 1 次印刷

定 价 / 69.80 元

责任编辑 / 王晓莉
文案编辑 / 王晓莉
责任校对 / 周瑞红
责任印制 / 李 洋

图书出现印装质量问题, 请拨打售后服务热线, 本社负责调换

前言

在互联网行业精益化运营的背景下,数据分析已成为网店运营的标配,大家都希望通过精细的数据分析来提高网店运营的效率。可很多网店运营人员对数据分析有一种淡淡疏远的心理,特别是非技术型的网店运营人员更是对数据敬而远之,好比一谈到数据分析,就需要一些高大上的数据分析工具,抽象的建模、函数似的。

事实上,通常意义上的数据分析用不了多少专业的数学知识,大多时候用到的都是非常简单的加、减、乘、除。所以,大学生不要对数据分析有太多的畏难情绪,所谓的数据分析就是指从数据中提取有用的信息,并指导实践。比如说结合数据优化产品的用户体验,通过数据来进行用户画像,通过数据发现产品改进的关键点,以及产品改版、迭代是否在一个正确的方向上。而这些事情,在经过实践之后,都会变得简单且容易上手。

本书作为高等院校电子商务专业学生学习用书,注重理论与实践的结合,强调锻炼学生数据分析的应用能力,着眼于培养具备数据分析技能的网店运营人才。

本书采用面向岗位,项目引导,任务驱动的教学模式设计教学内容,即围绕着高等教育培养生产、管理、服务第一线需要的高等技术应用性专门人才的目标,将网店数据分析岗位的工作内容整合成九个教学项目,针对教学项目设计工作任务,基于工作任务拓展数据分析的应用。书中的内容紧扣岗位操作需要,分析的数据取自真实的网店,发布的工作任务源于网店运营数据分析岗位的实际工作。

对于网店运营人员来说,要想提升销售额,必须做好提高访客数、提高转化率、提高客单价这三项工作。而要提高访客数、转化率和客单价,关键在于对客户、商品、市场行情和竞争对手的分析。

本书在编写上的安排顺序是:学习目标、项目导入、知识准备、任务发布、任务拓展、项目结构、同步测试和能力测评。

学习目标包括知识目标、技能目标和基本素养,作为学生学习的指引。

项目导入是用案例引导学生进入本项目学习的情境,让学生在学习理论知识前有一个感性认识。

知识准备是为学生完成工作任务准备的相关理论知识,包括概念、内容、原理、模型、流程、方法、指标、工具等。

任务发布是学生在教师的帮助下完成工作任务,这有助于学生掌握完成工作任务的方法和技巧。

任务拓展要求学生在没有教师的指导下独立完成任务,目的是培养学生独立思考问题、分析问题和解决问题的能力。

项目结构的作用是帮助学生了解整个项目的内容结构。

同步测试用于检验学生对本项目知识点的掌握程度。

能力测评是学生在完成本项目的学习后，对本项目的核心知识点和技能点掌握程度的自评。通过能力测评，教师能够了解学生的学习效果。

本书在编写过程中，借鉴了国内外许多专家学者的学术观点，参阅了大量书籍、期刊和网络资料，在这儿对各位作者表示感谢。本书还得到学校各位同事与北京理工大学出版社姚朝晖和陈圆圆的大力支持，在此一并致以衷心的感谢！

邵贵平
2017 年 12 月

目 录

项目一 网店数据分析认知	(1)
一、项目导入	(1)
二、知识准备	(3)
(一) 网店数据分析概念	(3)
(二) 网店数据分析流程	(6)
(三) 网店数据分析方法	(15)
(四) 网店数据分析工具	(26)
(五) 网店数据分析模型	(32)
三、任务发布	(35)
(一) 运用思维导图绘制淘宝网生意参谋的功能架构	(35)
(二) 基于 AARRR 模型分析云集微店	(38)
四、任务拓展	(40)
(一) 运用思维导图绘制京东智商的功能架构	(40)
(二) 讲解一个大数据分析成功案例	(41)
项目结构	(42)
同步测试	(42)
能力测评	(43)
项目二 网店运营数据分析	(44)
一、项目导入	(44)
二、知识准备	(45)
(一) 网店运营数据分析概念	(45)
(二) 淘宝网运营数据分析工具	(48)
(三) 网店运营数据分析指标	(51)
三、任务发布	(58)
(一) 网店初步诊断	(58)

2 网店数据分析

(二) ROI 计算	(61)
四、任务拓展	(62)
(一) 网店销售额诊断方案设计与实施	(62)
(二) 直通车 ROI 诊断方案设计	(63)
项目结构	(64)
同步测试	(64)
能力测评	(65)
项目三 网店流量数据分析	(66)
一、项目导入	(66)
二、知识准备	(69)
(一) 流量来源分析	(69)
(二) 流量趋势分析	(75)
(三) 流量质量评估	(80)
(四) 流量价值计算	(81)
(五) 爆款引流	(83)
(六) 七天螺旋	(89)
(七) 千人千面	(92)
三、任务发布	(94)
(一) 流量来源对比分析	(94)
(二) SEO 标题优化	(98)
四、任务拓展	(101)
(一) 设置商品上下架时间	(101)
(二) 商品类目优化	(102)
项目结构	(103)
同步测试	(103)
能力测评	(104)
项目四 网店转化率数据分析	(105)
一、项目导入	(105)
二、知识准备	(107)
(一) 电商转化率	(107)
(二) 成交转化漏斗模型	(108)
(三) 转化路径	(111)
(四) 转化率分析指标	(112)
(五) 影响转化率的因素	(116)
(六) 直通车转化分析	(120)
三、任务发布	(124)
(一) 跳失率诊断与优化	(124)
(二) 店铺首页装修因素分析	(131)

四、任务拓展	(136)
(一) 点击率诊断与优化	(136)
(二) 商品详情页装修因素分析	(137)
项目结构	(138)
同步测试	(138)
能力测评	(139)
项目五 网店客单价数据分析	(140)
一、项目导入	(140)
二、知识准备	(141)
(一) 认识客单价	(141)
(二) 客单价影响因素	(142)
(三) 客单价分析指标	(149)
三、任务发布	(151)
(一) 用 Apriori 算法做商品关联分析	(151)
(二) 导购路线设计	(155)
四、任务拓展	(157)
(一) 设计与评估提升客单价的促销方案	(157)
(二) 分析小米公司的连带销售策略	(158)
项目结构	(159)
同步测试	(159)
能力测评	(160)
项目六 网店客户数据分析	(161)
一、项目导入	(161)
二、知识准备	(163)
(一) 客户分析概述	(163)
(二) 客户分析指标	(165)
(三) 客户细分	(167)
(四) 客户忠诚度分析	(170)
(五) 客户生命周期	(171)
(六) 访客特征分析	(172)
(七) 访客行为分析	(178)
三、任务发布	(188)
(一) 绘制访客的用户画像	(188)
(二) 基于 RFM 模型细分客户	(191)
四、任务拓展	(198)
(一) SEM 营销中的受众分析	(198)
(二) 退货客户数据分析	(199)
项目结构	(200)
同步测试	(200)
能力测评	(201)

4. 网店数据分析

项目七 网店商品数据分析	(202)
一、项目导入	(202)
二、知识准备	(204)
(一) 商品分析认知	(204)
(二) 商品销售分析	(205)
(三) 商品价格分析	(207)
(四) 商品功能组合分析	(213)
(五) 用户体验分析	(214)
(六) 商品生命周期分析	(216)
(七) 商品毛利分析	(217)
(八) 商品库存分析	(219)
三、任务发布	(220)
(一) 商品价格带分析	(220)
(二) 购物篮分析	(223)
四、任务拓展	(227)
(一) 撰写产品分析报告	(227)
(二) 商品用户需求分析	(228)
项目结构	(229)
同步测试	(229)
能力测评	(230)
项目八 市场行情数据分析	(231)
一、项目导入	(231)
二、知识准备	(233)
(一) 市场行情认知	(233)
(二) 市场供给分析	(233)
(三) 市场需求分析	(235)
(四) 市场价格分析	(240)
(五) 经济周期分析	(240)
(六) 生产要素分析	(242)
三、任务发布	(244)
(一) 电商市场品牌竞争力分析	(244)
(二) 利用百度指数分析市场行情	(249)
四、任务拓展	(254)
(一) 利用阿里指数分析市场行情	(254)
(二) 撰写市场调研报告	(255)
项目结构	(256)
同步测试	(257)
能力测评	(258)

项目九 竞争对手数据分析	(259)
一、项目导入	(259)
二、知识准备	(261)
(一) 竞争对手认知	(262)
(二) 竞争对手分析步骤	(262)
(三) 竞争对手分析的层次和内容	(263)
(四) 竞争对手识别	(265)
(五) 竞争对手分析方法	(267)
(六) 竞争对手数据收集	(272)
(七) 竞争对手跟踪与监测	(275)
三、任务发布	(277)
(一) 竞品分析	(277)
(二) 顾客流失分析	(281)
四、任务拓展	(288)
(一) 竞店分析	(288)
(二) 店铺标杆管理	(289)
项目结构	(290)
同步测试	(290)
能力测评	(291)
参考文献	(292)

网店数据分析认知



学习目标

- 理解网店数据分析的相关概念;
- 熟悉网店数据分析的流程;
- 掌握网店数据分析的方法;
- 熟悉网店数据分析的工具;
- 理解网店数据分析模型。

技能目标

- 具有数据敏感性;
- 善于用数据思考和分析问题;
- 具备收集、整理和清洗数据的能力;
- 具备较好的逻辑分析能力。

知识目标

- 具备运用思维导图绘制数据分析平台功能架构的能力;
- 具备运用 AARRR 模型分析产品的能力;
- 具备讲解数据分析案例的能力。

基本素养

一、项目导入

时趣大数据助力李宁精准跨界明星营销

从李敏镐，到都敏俊，韩潮在中国的热度持续升温，而李宁与韩国少女时代 Jessica（郑秀妍）的跨界合作，让“LI-NING X Jessica”系列一经推出就备受追捧。在李宁首度宣布与 Jessica 合作当日，李宁官方微博创下自身官方微博互动纪录，参与互动的 30% 以上的网友都明确表示具有购买欲望。之后“LI-NING X Jessica”跨界合作产品在李宁官方商城正式开始预售，瞬间就吸引了数以万计的客户。

用明星效应来扩大销售额，这点我们一点也不陌生，但所用明星对产品目标消费者影响力的大小，常常难以把控。影视歌星、体育达人、网络红人……各类明星不计其数，要选哪一个似乎是企业主面临的一个比世界杯赔率更大的博彩。所有的成功都不是一次单纯的偶

2. 网店数据分析

然，李宁与 Jessica 的跨界合作是大数据下的商业产物，是基于社交媒体的数据收集、处理的一次精准营销。

时趣是怎样通过大数据分析帮助李宁选定合作明星 Jessica 的，又是如何利用数据进行后续的精准营销的呢？

首先，基于大数据的人群洞察——找到品牌与消费者的最优连接者。

作为中国最大的运动品牌之一，李宁需要清晰地把控自己新品系列的调性和受众。李宁把“90后”、女性这些未来消费主力军，作为它们主要的产品战略对象。基于这个战略，李宁联手合作伙伴大数据专家——时趣，在社交媒体上进行了全方位的人群洞察，使其在年轻化、女性群体身上找到了自己新品系列的基础调性：小清新、运动、时尚、阳光。如何让新品信息精准抵达目标消费者，在大范围造势，这需要一个连接，而李宁利用时趣基于社交媒体的卓越的大数据能力，找到了这个具有与其新产品系列属相相似的“连接”——韩国顶级女团少女时代成员 Jessica。

其次，匹配明星信息——确保信息精准抵达。

时趣数据中心通过采集、清洗、存储、计算并整合新浪微博海量微博内容数据及相关客户数据、关系数据等，目前已积累了自己庞大数量的活跃客户数据、关系数据、微博数据、标签数据，并保持每日新增客户数据及微博数据。显然，只有积累足够的客户数据，才能分析出客户的喜好与购买习惯，及时并且全面地了解客户的需求与想法，做到“比客户更了解客户自己”。因此，在纵向上，时趣基于数据对当红艺人社交网络影响力进行了详细的分析与总体评价；在横向上，时趣从性别、年龄、地域、兴趣标签、语义情感等几个维度上把 Jessica 粉丝的集中倾向属性和李宁新产品系列调性进行了综合匹配，确保了其推送信息可以精准抵达目标消费者，并且这个抵达过程速度极快，范围很广，单位抵达成本低得惊人。

最后，建立预测响应机制——优化后续营销活动设计。

要做到“精准营销”就必须准确地预计客户需求，因此，时趣不仅对大量历史数据进行了挖掘与分析，还建立了相应的预测响应机制，根据客户在社交媒体上的活动建立数据收集模型，通过模型完成数据的加工和分析，为品牌下一步的产品策划与营销提供更加有意义的数据参考。“LI-NING X Jessica”相关话题中，“Jessica 参与设计”“Jessica 行程”等相关内容被多次提及。根据这一数据信息，李宁建立了李宁首尔工作室，随继推出了“型自首尔”系列，邀请 Jessica 亲自参与设计。同时，李宁在其官方商城为“Jessica & Krystal 近距离接触”的活动造势，使新产品售卖热度继续上升。LI-NING X 与 Jessica 的合作效果如图 1-1 所示。

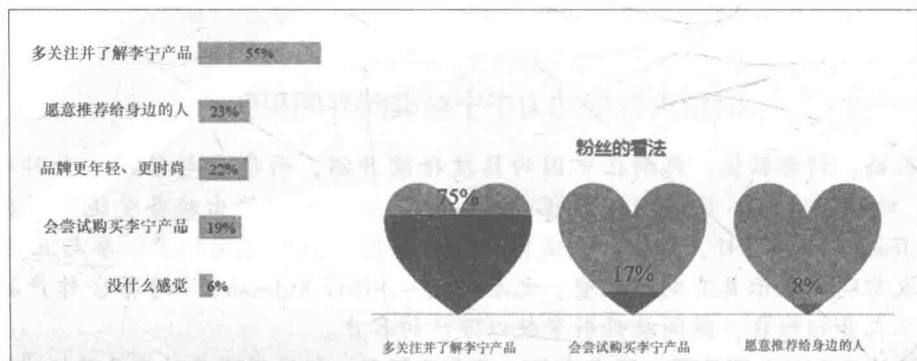


图 1-1 LI-NING X 与 Jessica 的合作效果

总的来说,时趣已经具备通过大数据战略打破行业边界的能力,对庞大、复杂的客户数据进行挖掘、追踪、分析,对不同客户群体进行聚合,获得更为完整的客户或客户群体的模型,打造个性化、精准化、智能化的产品营销解决方案,以个性化营销和主动营销打破传统无差异的、被动的产品服务营销方式。

思考:

1. 大数据分析在李宁的“LI-NING X Jessica”系列产品推广过程中起到哪些作用?
2. 如何评价李宁与明星这次合作的营销效果?请列出具体指标。

二、知识准备

网店数据分析是指对网店经营过程产生的数据进行分析,在研究大量的数据的过程中寻找模式、相关性和其他有用的信息,可以帮助商家更好地适应变化,做出更明智的决策。

(一) 网店数据分析概念

1. 数据分析

数据分析是指收集、处理数据并获取信息的过程。具体地说,数据分析是指在业务逻辑的基础上,运用简单有效的分析方法和合理的分析工具对获取的数据进行处理的一个过程。

数据分析是指用适当的统计方法对收集来的大量第一手资料和第二手资料进行分析,以求最大化地开发数据资料的功能,发挥数据的作用,是为了提取有用信息和形成结论而对数据加以详细研究和概括总结的过程。

(1) 数据分析的目的。数据分析的目的是把隐藏在一大批看起来杂乱无章的数据中的信息集中、萃取和提炼出来,以找出所研究对象的内在规律。在实际生活中,数据分析可帮助网店经营者进行判断和决策,以便他们采取适当的策略与行动。例如网店经营者准备开拓一个新的市场,则需要充分了解竞争对手的市场状况、市场潜力及销售预测,从而发现市场机会,找到突破口,这些都有赖于数据分析。

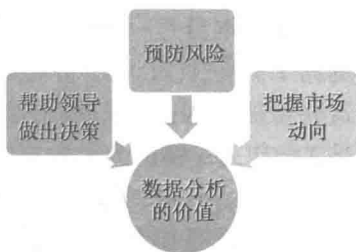


图 1-2 数据分析的价值

(2) 数据分析的价值。数据分析的价值包含三个方面(图 1-2),一是帮助领导做出决策;二是预防风险;三是把握市场动向。数据分析可以反馈网店需要改进的地方、出现的问题以及做得好的地方。

(3) 数据分析的作用。数据分析在网店日常经营分析中,具有以下三个作用:

① 现状分析,提供网店现阶段整体运营情况,其中包括各项经营指标的完成情况,以及网店各项业务的构成,其中包括各项业务的发展以及变动情况,用于衡量网店经营现状。现状分析的结果表现为各种形式的日常通报,如日报、周报、月报、季报、年报等。

② 原因分析,确定网店所存在问题的原因,针对原因制定出相应的解决方案。现状分析可以帮助网店经营者了解店铺的整体运营情况,同时发现运营中存在的问题,而找出问题就要进行分析。例如本月店铺销售额环比下降了 10%,是什么原因导致的?是店铺流量减少了,还是转化率出现了问题,抑或是客单价降低了?通过原因分析找到根源所在,这样才能真正解决问题。

③ 预测分析,对网店未来的发展趋势做预测,便于网店制订运营计划。例如网店经营

者一般都会根据近几个月销售额的变动趋势预测下个月的销售额，作为店铺的运营目标以及对员工考核的依据。

(4) 数据分析的应用。数据分析有极广泛的应用范围。在产品的整个生命周期内，数据分析过程是质量管理体系的支持过程，包括从产品的市场调研到售后服务以及最终处置都需要适当进行数据分析，以提升有效性。例如一家企业通过市场调查，分析所得的数据，以此来判定市场动向，从而制订出合适的生产及销售计划。在网店经营过程中，可从获取的用户需求数据中推导出用户偏好，用于开发新的产品或提供新的服务；可从售后客户反馈数据中分析出用户的需求，用于开展更高效的市场营销活动。因此，数据分析有着广泛的应用范围，在淘宝店铺运营过程中起着积极的意义。

(5) 数据分析的分类。在统计学领域，一般把数据分析分为三类：EDA (Exploratory Data Analysis, 探索性数据分析)、CDA (Confirmatory Data Analysis, 验证性数据分析) 和描述性数据分析。

探索性数据分析是指对已有的数据在尽量少的先验假定下进行探索，侧重于在数据之中发现新的特征。EDA 讲究的是从客观数据出发，探索其内在的数据规律性，让数据自己说话。

验证性数据分析是指进行分析之前一般都有预先设定的模型，它侧重于已有假设的证实或证伪。

描述性数据分析是对一组数据的各种特征进行分析，以便于描述测量样本的各种特征及其所代表的总体的特征。描述性统计分析的项目很多，常用的如平均数、标准差、中位数、频数分布、正态或偏态程度等。这些分析是复杂统计数据分析的基础。

2. 大数据

大数据本身是一个比较抽象的概念，单从字面来看，它表示数据规模的庞大。目前大数据尚未有一个公认的定义，专家通常从大数据的特征出发，试图通过这些特征的阐述和归纳给出其定义。在这些定义中，比较有代表性的是 3V 定义，即认为大数据需满足三个特点：规模性 (Volume)、多样性 (Variety) 和高速性 (Velocity)。维基百科对大数据的定义简单明了：大数据是指利用常用软件工具捕获、管理和处理数据所耗时间超过可容忍时间的数据集。



知识链接

数据分析从小数据开始

大数据最引人注目的是“大”，但是有时从小一些的数据集开始分析也是可以的，直到你能够分辨出你所寻找的是什么。如果你的研究对象为开业小于一年的零售店，那么你不需要一个大的 IT 系统来获取、贮存以及分析所有信息。如果你想要获得更愿意购买你的商品的顾客统计资料，那么很多免费的资源（例如 Google Adwords 以及其相关的分析工具）可以帮助你达到目的。一旦你明白数据分析如何使你受益以及使你明白去获得什么，你将会非常确定地想要更大规模以及更可信的方法。

3. 云计算

Google 作为大数据应用最为广泛的互联网公司之一, 2006 年率先提出“云计算”的概念。云计算是一种大规模的分布式模型, 通过网络将抽象的、可伸缩的、便于管理的数据能源、服务、存储方式等传递给终端客户。根据维基百科的说法, 狭义云计算是指 IT 基础设施的交付和使用模式, 指通过网络以按照需求量和易扩展的方式获得所需资源; 广义云计算是指服务的交付和使用模式, 指通过网络以按照需求量和易扩展的方式获得所需服务。目前云计算可以认为包含三个层次的内容: 服务 (IAAS)、平台即服务 (PAAS) 和软件即服务 (SAAS)。国内的阿里云与云谷公司的 XENSYSTEM, 以及在国外已经非常成熟的 INTEL 和 IBM 都是云计算的忠实开发者和使用者。

云计算是大数据的基础平台与支撑技术。如果将各种大数据的应用比作一辆辆“汽车”, 那么支撑起这些“汽车”运行的“高速公路”就是云计算。正是云计算技术在数据存储、管理与分析等方面的支撑, 才使大数据有用武之地。

4. 数据可视化

数据可视化是指将数据分析结果用简单且视觉效果好的方式展示出来, 一般运用文字、表格、图标和信息图等方式展示。Word、Excel、PPT、水晶易表等都可以作为数据可视化的展示工具。现代社会已经进入一个速读时代, 好的可视化图表可以清楚地表达数据分析结果, 大大节约了思考的时间。

数据分析的使用者有大数据分析专家和普通客户, 他们对大数据分析最基本的要求就是数据可视化, 因为数据可视化能够直观地呈现大数据的特点, 让数据自己说话, 让观者直接看到结果, 这就是数据可视化的作用。

5. 数据挖掘

大数据分析的理论核心就是数据挖掘。各种数据挖掘的算法基于不同的数据类型和格式能更加科学地呈现出数据本身的特点, 能更快速地处理大数据。如果采用一种算法需要花好几年才能得出结论, 那么大数据的价值也就无从说起了。可视化是给人看的, 数据挖掘是给机器看的。集群、分割、孤立点分析还有其他的算法可以使我们深入数据内部去挖掘价值。这些算法不仅能够处理大数据的数据量, 而且能够在一定程度上满足处理大数据的速度要求。

数据挖掘的重点不在数据本身, 而在于如何能够真正地解决数据运营中的实际商业问题。但是, 要解决商业问题, 就得让数据产生价值, 就得做数据挖掘。

6. 语义引擎

非结构化数据与异构数据等的多样性带来了数据分析新的挑战与困难, 需要一系列的工具去解析、提取、分析数据。语义引擎需要被设计成能够从文档中智能提取信息, 使之能从大数据中挖掘出特点, 通过科学建模和输入新的数据, 预测未来的数据。

7. 数据质量

在数据分析和数据挖掘之前, 首先必须保证高质量的数据, 完成数据质量的处理工作, 即对数据的集成和处理。因此, 更好的数据意味着更好的决策, 否则就是 GIGO (Garbage in, Garbage out)。所以, 数据分析的前提就是保证数据质量。

数据质量的处理工作主要包括两方面: 数据集成和数据清洗, 关注的对象主要有原始数据和元数据两方面。

知识链接

元数据

元数据是指信息的信息，描述信息的属性信息。一个信息的元数据可以分为三类：固有性元数据，是指事物固有的与事物构成有关的元数据；管理性元数据，是指与事物处理方式有关的元数据；描述性元数据，是指与事物本质有关的元数据。

以摄像镜头为例，镜头的固有性元数据包括品牌、参数、类型、重量、光圈、焦距等信息；镜头的管理性元数据包括商品类型、上架时间及库存情况；镜头的描述性元数据包括用途、特色，如人文纪实和人像摄影。

8. 数据预测分析

预测分析是大数据技术的核心应用，如电子商务网站利用数据预测顾客是否会购买推荐的产品，信贷公司利用数据预测借款人是否会违约，执法部门用大数据预测特定地点发生犯罪的可能性，交通部门利用数据预测交通流量等。预测是人类本能的一部分，只有通过大数据分析才能获取智能的、有价值的信息。越来越多的应用涉及大数据，大数据的属性决定了其存储的复杂性。大数据预测分析突破了预测分析一直是象牙塔里统计学家和数据分析师的工作的性质。随着大数据的出现，并整合到现有的 BI、CRM、ERP 和其他关键业务系统，大数据预测分析将起到越来越重要的作用。

数据预测分析可帮助网店做出正确而果断的业务决策，让客户更开心，同时避免灾难的发生，这是众多数据分析者的终极梦想，但是预测分析也是一个困难的任务。实施成功的预测分析有赖于数据质量、数据分析师、数据分析工具等。

知识链接

数据分析师应具备的能力

对于数据分析师来说，熟悉业务逻辑和掌握数据分析的方法与工具同样重要。数据分析师不可能在没有理解业务逻辑的基础上，提出一个基于数据的解决方案，因为每个行业都有不同的趋势、行为和驱动因素。事实上，一个不懂业务逻辑的数据分析师，其分析结果不会产生任何使用价值；而一个只懂业务逻辑而不精通数据分析方法与工具的数据分析人员，充其量只能算数据分析爱好者。现实情况是，懂数据分析的人很多，懂业务逻辑的人更多，但既懂业务又懂数据分析称得上“数据分析师”的人却非常少。

（二）网店数据分析流程

网店数据分析流程一般分六步走，具体为：明确分析目的和思路、收集数据、数据处

理、数据分析、数据展现与报告撰写，如图1-3所示，具体步骤如下：



图1-3 数据分析流程

1. 明确分析目的和思路

识别数据分析需求，明确数据分析目标，是确保数据分析过程有效性的首要条件。在进行数据分析之前，就应该冷静思考在数据分析过程中想要获得什么。例如，是想要更精确地确定网店的客户群吗？是想要扩大网店的客户群吗？或者是为了评估产品改版后的效果是否比之前有所提升吗？是为了找到产品迭代的方向吗？是想要进行科学的排班以至于不必在闲时浪费人力、在忙时缺少人手吗？明确通过数据分析要获得什么是至关重要的，就如同要明确这个过程挑战是什么以及如何确定是否能够达到目的。

明确了数据分析的目的，接下来需要确定数据分析思路。数据分析思路是指运用营销和管理的相关技术及方法，结合实际业务将数据分析目的层层分解，形成一个结构化的数据分析框架。数据分析框架的构建有助于提高数据分析维度的完整性和分析结果的有效性。

2. 收集数据

当通过数据分析来揭示变化趋势时，数据量越大越好。对于任何类型的统计分析，样本量越大，所得到的结果越精确。仅仅追踪公司一周的销售数据的价值是很难看出未来发展趋势的，3个月的会好一些，6个月的更佳。即使无法确定所寻找的是什么，也要确保所收集的数据包含的信息尽可能详尽和精确。试着弄清楚获得所需最优数据的途径，然后开始收集。如果没有数据，就不能够进行分析。

收集数据是如何将数据记录下来的环节。在这个环节中需要着重说明的是两个原则，即全量而非抽样，以及多维而非单维。今天的技术革命和数据分析2.0主要就体现在这两个层面之上。

(1) 全量而非抽样。

由于系统分析速度以及数据导出速度的制约，在非大数据系统支撑的公司中，做数据分析的人员也很少能够做到对数据的全量收集和分析。但在未来这将成为一个问题。

(2) 多维而非单维。

在数据维度上，要进行全面细化，将交互过程的时间、地点、人物、原因、事件全面记录下来，并将每一个板块进行细化，时间可以从起始时间、结束时间、中断时间、周期间隔时间等细分；地点可以从地市、小区、气候等地理特征、渠道等细分；人物可以从多渠道注册账号、家庭成员、薪资、个人成长阶段等细分；原因可以从爱好、人生大事、需求层级等细分；事件可以从主题、步骤、质量、效率等细分。通过这些细分维度，增加分析的多样性，从而挖掘规律。

有目的地收集数据是确保数据分析过程有效的基础，需要对收集数据的内容、渠道、方法进行策划，主要考虑：

- ① 将识别的数据分析需求转化为更具体的要求，如评价供方时，需要收集的数据可能包括其过程能力、测量系统不确定性等相关数据；
- ② 明确由谁在何时何处，通过何种渠道和方法收集数据；
- ③ 记录表应便于使用；
- ④ 采取有效措施，防止数据丢失和虚假数据对系统的干扰。