

电子商务新形态一体化教材
电商实战宝典系列

新型
活页式
教材

商务数据分析与处理

主编 才琳 王玥

SHANGWU SHUJU FENXI YU CHULI



北京邮电大学出版社
www.buptpress.com

服务热线: 400-615-1233

配套精品教学资料包

www.huatengedu.com.cn

商务数据 分析与处理

SHANGWU SHUJU FENXI YU CHULI

策划编辑: 刘建

责任编辑: 许青

封面设计: 黄燕美

ISBN 978-7-5635-6693-8



9 787563 566938 >

定价: 59.80元

电子商务新形态一体化教材
电商实战宝典系列

商务数据 分析与处理

主编 才琳 王玥
副主编 宋艳苹 毕波
李朝红 花景田

SHANGWU SHUJU FENXI YU CHULI



北京邮电大学出版社
www.buptpress.com

内 容 简 介

本书共包括7个项目,分别为商务数据分析概述、数据存储与数据分析工具、制订数据采集与处理方案、市场数据分析与可视化、运营数据分析与可视化、产品数据分析与可视化、数据监控与数据分析报告。

本书既可作为职业院校商务数据分析与应用、电子商务等专业的教材,也可供相关人员参考。

图书在版编目(CIP)数据

商务数据分析与处理 / 才琳, 王玥主编. -- 北京: 北京邮电大学出版社, 2022. 8(2023. 7 重印)

ISBN 978-7-5635-6693-8

I. ①商… II. ①才… ②王… III. ①商务工作—数据处理 IV. ①F715

中国版本图书馆 CIP 数据核字(2022)第 141745 号

策划编辑: 刘 建 责任编辑: 许 青 封面设计: 黄燕美

出版发行: 北京邮电大学出版社

社 址: 北京市海淀区西土城路 10 号

邮政编码: 100876

发 行 部: 电话: 010-62282185 传真: 010-62283578

E-mail: publish@bupt.edu.cn

经 销: 各地新华书店

印 刷: 三河市骏杰印刷有限公司

开 本: 787 mm×1 092 mm 1/16

印 张: 16 插页 1

字 数: 331 千字

版 次: 2022 年 8 月第 1 版

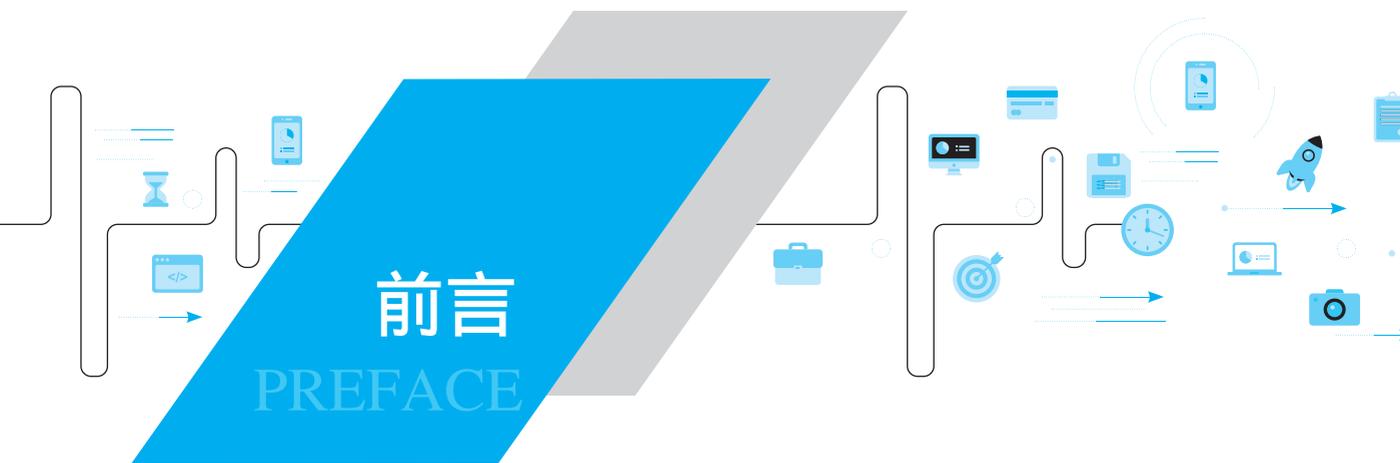
印 次: 2023 年 7 月第 2 次印刷

ISBN 978-7-5635-6693-8

定 价: 59.80 元

• 如有印装质量问题,请与北京邮电大学出版社发行部联系 •

服务电话:400-615-1233



前言

PREFACE

2021年,《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》第五篇“加快数字发展 建设数字中国”指出,要“迎接数字时代,激活数据要素潜能,推进网络强国建设,加快建设数字经济、数字社会、数字政府,以数字化转型整体驱动生产方式、生活方式和治理方式变革”。

中华人民共和国工业和信息化部印发的《“十四五”大数据产业发展规划》指出,大数据产业是以数据生成、采集、存储、加工、分析、服务为主的战略性新兴产业,是激活数据要素潜能的关键支撑,是加快经济社会发展质量变革、效率变革、动力变革的重要引擎。

在数字化时代,数据已经成为企业的生产要素,数据分析的重要性日益凸显。科学有效的数据分析能帮助企业了解自身运营状况、查找问题症结、判断发展趋势、发现商业机会、做出科学决策等。随着科技的发展,企业对数据分析人才的需求量不断增大,数据分析能力成为从业人员的一项重要技能。在此背景下,编者编写了本书。

本书遵循学生的认知规律,围绕数据分析的工作过程,从数据分析的基本认知入手,让学生在学习过程中掌握数据分析的方法与操作技能,养成数据分析的思维习惯,具备用数据分析的思维与方法处理企业运营过程中出现的问题的能力。

本书的学时安排见下表。

项 目	内 容	学 时
1	商务数据分析概述	6
2	数据存储与数据分析工具	10
3	制订数据采集与处理方案	6

续表

项 目	内 容	学 时
4	市场数据分析与可视化	8
5	运营数据分析与可视化	20
6	产品数据分析与可视化	8
7	数据监控与数据分析报告	6
总学时		64

本书的主要特色如下。

1. 项目引领、任务驱动

本书通过项目引领、任务驱动的方式帮助学生掌握理论知识及操作技能。全书本着“必需、够用”的原则讲解理论知识，力求内容全面，图文并茂，通俗易懂。

2. 内容设计与课程思政建设相结合

本书内容设计充分结合课程思政建设要求，编者努力挖掘与内容相适应的思政元素，旨在引导学生树立正确的世界观、人生观和价值观，培养学生的政治素养、爱国情怀，以及遵纪守法和恪守职业道德的基本素养。

3. 活页式装订，使用更方便

本书采用活页式装订，方便学生做注释或拆装使用。

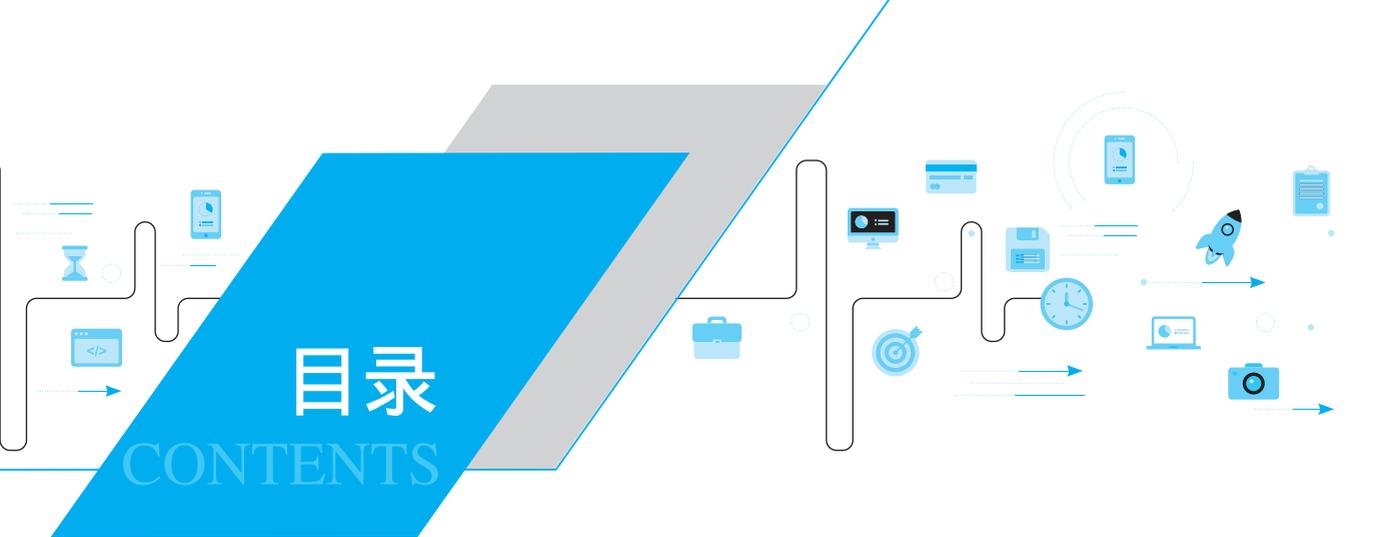
4. 数字资源丰富

本书将微课、例题解答等内容以数字资源的形式体现，学生用手机扫描书中二维码即可实现碎片化学习。

本书由长春职业技术学院才琳、王玥任主编，长春职业技术学院宋艳苹、毕波、李朝红、花景田任副主编。

由于编者水平有限，书中难免存在疏漏之处，恳请广大读者批评指正。

编 者



项目 1	商务数据分析概述	
任务 1.1	认识商务数据分析	4
1.1.1	认识数据	4
1.1.2	数据分析	6
任务 1.2	数据分析指标与流程	10
1.2.1	常用的数据分析指标	10
1.2.2	数据分析流程	14
任务 1.3	数据分析模型	16
1.3.1	5W2H 模型	16
1.3.2	生命周期分析模型	16
1.3.3	4P 策略模型	19
1.3.4	PEST 模型	20
1.3.5	SWOT 分析模型	20
1.3.6	逻辑树分析模型	21
1.3.7	RFM 模型	22
1.3.8	漏斗模型	23
任务 1.4	数据分析方法	24
1.4.1	对比分析法	24
1.4.2	结构分析法	24
1.4.3	同期群分析法	24
1.4.4	聚类分析法	25

项目 2 数据存储与数据分析工具

任务 2.1 数据的存储与查询 31

2.1.1 数据存储的基本概念	31
2.1.2 数据模型与数据库的类型	31
2.1.3 SQL 语言	34
2.1.4 数据存储	35
2.1.5 数据查询	40

任务 2.2 常用数据分析工具 49

2.2.1 Excel 电子表格	49
2.2.2 SPSS	74
2.2.3 Python	85
2.2.4 生意参谋	87
2.2.5 京东商智	93

项目 3 制订数据采集与处理方案

任务 3.1 确定数据分析目标与指标 104

3.1.1 确定数据分析目标	104
3.1.2 确定数据分析指标	104

任务 3.2 选择数据采集渠道及工具 106

3.2.1 数据采集渠道	106
3.2.2 数据采集工具	107

任务 3.3 撰写数据采集与处理方案 111

3.3.1 背景说明	111
3.3.2 明确分析目标	111
3.3.3 确定分析指标	111
3.3.4 明确数据采集渠道及数据采集工具	111

项目 4 市场数据分析与可视化

任务 4.1 行业数据分析 116

- 4.1.1 行业发展分析 116
- 4.1.2 市场需求分析 122
- 4.1.3 目标客户分析 125

任务 4.2 竞争数据分析 129

- 4.2.1 竞店分析 129
- 4.2.2 竞品分析 136

项目 5 运营数据分析与可视化

任务 5.1 客户数据分析 144

- 5.1.1 客户分类 144
- 5.1.2 客户特征分析 145
- 5.1.3 客户忠诚度分析 153
- 5.1.4 新老客户占比分析 155
- 5.1.5 客户行为分析 156

任务 5.2 推广数据分析 162

- 5.2.1 推广渠道分析 162
- 5.2.2 关键词推广分析 170
- 5.2.3 活动推广效果分析 173
- 5.2.4 内容运营分析 176

任务 5.3 销售数据分析 190

- 5.3.1 交易数据分析 190
- 5.3.2 服务数据分析 197

任务 5.4 供应链数据分析	205
5.4.1 采购数据分析	205
5.4.2 物流数据分析	206
5.4.3 库存数据分析	210

项目 6 产品数据分析与可视化

任务 6.1 产品行业数据分析	216
6.1.1 利用百度指数采集	216
6.1.2 利用生意参谋查看淘宝店铺行业数据	217
6.1.3 利用京东商智查看京东店铺行业数据	221
任务 6.2 产品能力数据分析	226
6.2.1 产品获客能力分析	226
6.2.2 产品盈利能力分析	228

项目 7 数据监控与数据分析报告

任务 7.1 数据监控	237
7.1.1 数据监控的含义	237
7.1.2 数据监控的流程	237
7.1.3 数据监控方法	238
7.1.4 制作数据监控报表	239
7.1.5 借助监控工具监控数据	241
任务 7.2 数据分析报告	243
7.2.1 数据分析报告的分类	243
7.2.2 数据分析报告的结构	244
7.2.3 数据分析报告的撰写	246

参考文献	249
-------------	------------



项目

商务数据分析概述

学习目标

- 【知识目标】
- (1) 掌握商务数据分析的基本概念。
 - (2) 熟悉商务数据分析指标。
 - (3) 了解商务数据分析流程。
 - (4) 了解商务数据分析模型。
- 【能力目标】
- (1) 能够实际运用商务数据分析的工作流程。
 - (2) 能够根据数据分析目标选择合适的数据分析指标。
 - (3) 能够根据数据分析目标选择合适的数据分析模型。
- 【素质目标】
- (1) 培养对问题的分析能力和总结归纳能力。
 - (2) 形成认真严谨的工作态度。
 - (3) 树立正确科学的价值观。

●【引导案例】

谷歌的意图

根据搜索研究公司 comScore 的数据，仅一个月的时间，谷歌处理的搜索词条数量就高达 122 亿条。谷歌的体量和规模，使它拥有比其他大多数企业更多的应用大数据的途径。

谷歌搜索引擎本身的设计，旨在让它能够无缝链接成千上万的服务器。如果出现更多的处理或存储需要，抑或某台服务器崩溃，谷歌的工程师们只要再添加更多的服务器就能轻松搞定。将所有数据集合在一起所带来的结果是，企业不仅从最好的技术中获益，还可以从最好的信息中获益。

1. 谷歌存储

谷歌不仅能存储搜索结果中出现的网络链接，还会存储用户搜索关键词的行为。它能够精准地记录下人们进行搜索行为的时间、内容和方式，坐拥人们在谷歌网站进行搜索及经过其网络时所产生的大量机器数据。这些数据能够让谷歌优化广告排序，并将搜索流量转化为盈利模式。谷歌不仅能追踪人们的搜索行为，还能够预测出搜索者下一步将要做什么。用户所输入的每一个搜索请求都会让谷歌知道他在寻找什么。所有人类行为都会在互联网上留下痕迹路径，谷歌占领了一个绝佳的点位来捕捉和分析该路径。换言之，谷歌能在你意识到自己要找什么之前预测出你的意图。这种抓取、存储并对海量人机数据进行分析，然后据此进行预测的能力，就是数据驱动的产品。

2. 谷歌分析

谷歌在搜索之外还有更多获取数据的途径。企业安装“谷歌分析”等产品来追踪访问者在其站点的足迹，而谷歌也可获得这些数据。网站还使用“谷歌广告联盟”将来自谷歌广告客户网的广告展示在其站点。因此，谷歌不仅可以洞察自己网站上广告的展示效果，还可以对其他广告发布站点的展示效果一览无余。

3. 谷歌趋势

既然搜索本身是网民的“意图数据库”，当然也可以根据某一专题搜索量的涨跌预测下一步的走势。谷歌趋势可以预测旅游、地产、汽车的销售走向。此类预测最著名的就是谷歌流感趋势，其跟踪全球范围的流感等病疫传播，依据网民搜索，可分析全球范围内流感等病疫的传播状况。

案例思考：

从以上案例可以发现数据分析在企业中的价值。那么，什么是数据分析，数据分析有哪些应用，怎样进行数据分析呢？

 素质拓展

请阅读下面的信息

2021年11月30日,中华人民共和国工业和信息化部印发《“十四五”大数据产业发展规划》(以下简称《规划》)。

《规划》提出“十四五”时期的总体目标,到2025年我国大数据产业测算规模突破3万亿元,年均复合增长率保持25%左右,创新力强、附加值高、自主可控的现代化大数据产业体系基本形成。

《规划》围绕加快培育数据要素市场、发挥大数据特性优势、夯实产业发展基础、构建稳定高效产业链、打造繁荣有序产业生态、筑牢数据安全保障防线六个方面提出重点任务,设置数据治理能力提升、重点标准研制及应用推广、工业大数据价值提升、行业大数据开发利用、企业主体发展能级跃升、数据安全铸盾六个专项行动。

《规划》提出在原材料、装备制造等4个工业领域率先实施大数据价值提升行动,在通信、金融等12大行业开展大数据开发利用行动。要加强关键核心技术攻关,引导大中小企业融通发展和产业链上下游协同创新,推动大数据领域国家新型工业化产业示范基地高水平建设。加强数据安全治理,提升数据安全保障能力。《规划》还从提升全社会数据思维和加快大数据人才培养两方面进行了重点部署,提出加强大数据知识普及,推广首席数据官制度,运用数据加快组织变革和管理变革。

大数据产业政策体系日益完善,相关政策内容已经从宏观的总体规划方案逐渐向微观细分领域深入。工业和信息化部、交通运输部、公安部、农业农村部等均推出了关于大数据的发展意见、实施方案、计划等,推动各行业应用大数据。另外,大数据技术攻关政策、安全保障政策、产业关联政策等日益完善,为大数据产业发展提供保障。

请针对上述信息进行思考

大数据为什么能够受到国家的重视?大数据产业的健康有序发展对于促进我国发展有着怎样的作用?搜集相关信息,了解国家和各省(自治区、直辖市)颁布的利于大数据发展的相关政策。

任务

1.1

认识商务数据分析

随着网络技术和信息技术的飞速发展，大数据时代已经来临，数据分析已渗透到人们生活和工作的方方面面，人们可以通过网络查询车次、航班，预订酒店，查看新冠肺炎疫情的出行限制等；数据分析同样深刻影响着不同行业企业的运营与管理，如影视公司可以预测电影票房，经营企业可以预测销售额，收集客户行为，挖掘客户喜好等。就电子商务行业来说，利用数据分析可以发现营销手段的不足、客户体验的不足；利用数据分析可以了解客户的内在需求，为客户进行智能推荐等。只有充分利用海量数据以及数据分析思维和方法，分析和挖掘数据背后的规律，才能使数据分析的价值得到充分发挥，从而有效指导企业的运营管理工作。

1.1.1 认识数据

1. 数据的含义

对于数据，大多数人认为它就是数字。实际上，数据不仅指狭义上的数字，它还是具有一定意义的文字、字母，或者是各种符号的组合，也可以是图形、图像、视频、音频等。数据是可识别的、抽象的符号，是客观事物的属性、数量、位置及其相互关系的抽象表示。

例如，“1 150”只是一个数字，但是这个数字可以表达多个含义。在具体的情境中，“某店铺访客数是 1 150 人”，那么“1 150 人”虽然是数据，但表达不准确，因为店铺的访客数是动态指标，随时间变化而变化。“某店铺 2021 年 12 月 1 日上午 8 时访客数为 1 150 人”，此时的“1 150 人”才是具体的、有准确意义的数据，如图 1-1 所示。



图 1-1 数据所表示的含义

本书中所指的数据，包括数值、文字、声音、图片、视频等多种形式，包含所有用数字化方式记录下来事物发生和发展的状态信息。

2. 大数据的特征

在大数据时代，任何微小的数据都可能产生不可思议的价值。大数据有以下 4 个特点。

(1) 数据量巨大。大数据的首要特征就是数据量巨大，从数据单位 MB、GB、TB 到

PB、EB 的变化，就可以感受到数据的这一特征。随着信息技术的高速发展，数据呈现爆发性的增长。社交网络、移动网络、各种智能工具、服务工具等，都成为数据的来源。中国铁道科学研究院电子所副总工程师兼 12306 技术部主任单某在接受采访时表示：12306 高峰时段 1 秒售出近 700 张火车票。也就是说，在一秒内一辆热门动车组的票全被抢完了。如此高的售票量背后是惊人的访问量。据有关数据显示，12306 高峰日的网络页面浏览量超过 1 500 亿次，相当于全中国人每人每天访问了票务页面 100 多次。在这种情况下，迫切需要智能的算法、强大的数据处理平台和新的数据处理技术来统计、分析、预测和实时处理如此大规模的数据。

对于网店经营者来说，在网店运营过程中，需要明确数据的容量是有很大区别的。同一家店铺，一天、一周、一个月和一个季度的数据量差别很大。因此要根据分析目标，尽量选择更多的数据作为分析基础，数据量越大，分析结果准确性越高，越有参考价值。

例如，某店铺要了解店铺最近的销售情况，如果分析人员只采集一周的销售数据，发现销售额有下降趋势，这显然是片面的，因为数据量太小，分析结果很难有说服力。因此分析人员应该至少采集一个月或者更长时间的数据，然后分析并得出结论。

(2) 数据形式和种类多样。数据来源广泛，数据形式多种多样。人们每天会接收到不同形式的数据，如文字、数字、声音、图像等。其中对数据应用比较广泛的是各平台的推荐系统，如淘宝、京东等各类电商平台，抖音、快手等短视频平台，以及 QQ 音乐、网易云音乐等音乐平台，都会通过对用户的日志数据进行分析，从而推荐用户喜欢的产品或服务。

从形式上看，商务数据更多的是数字形式，然而商务数据类型多样，如流量数据、转化数据、客户数据、销售数据、服务数据等。卖家在进行数据分析时，要根据分析目标，从多个维度选择恰当的数据类型，从多个角度进行分析，才能得出准备结论。

例如，某电商店铺参加了淘宝/天猫平台“双 11”大促活动，对活动效果进行分析时，卖家就要采集流量类、转化类、销售类等数据加以综合分析，才能对活动效果做出准确的总结。

(3) 数据产生高速，具有时效性。互联网无处不在，人们每天的生活、行动轨迹会产生大量数据，同时也在向大数据提供大量的数据资源。数据的产生非常迅速，主要通过互联网传输。对平台来说，时间较远的数据需要清理，只保留一定期限内的数据。很多平台都需要做到实时分析，因此大数据对处理速度要求极高，这样数据的时效性才有保障。数据随时产生，谁的速度更快，谁就有优势。

例如，某电商店铺要了解受客户欢迎的产品价格区间，如果采集店铺近三年的数据进行分析，结果就会产生偏差。因为随着经济的发展、行业的变化、客户消费心理、消费能力等因素的变化，客户能接受的产品价格区间会产生很大不同，因此需要采集最近一段时间的数据，才比较准确。

(4) 数据的价值需要分析与挖掘。这是大数据的核心特征。大量数据中有价值的数据所占比例很小，大数据就是通过从大量不相关的各种类型的数据中，挖掘出对未来趋势与

模式预测分析有价值的数 据，并运用人工智能、数据挖掘等方法深度分析，发现规律，并应用于各个领域，从而提高工作效率、改善社会治理、推进科技发展等。

课堂讨论

观察你所处的环境，有哪些方面是应用了大数据的？这些应用对你的生活产生了怎样的影响？

1.1.2 数据分析

1. 数据分析的含义

数据无处不在，对数据的分析利用已经渗透到人们生活中的方方面面，对企业来说更是如此。对于从事数据分析的人员来说，一定要能够透过数据的表象去发现数据背后的含义。例如，一张精美的广告图片，客户看到的是商品、价格、促销等信息，然而实际上这是相关工作人员根据产品受众、流行趋势等一系列数据分析之后，才精心制作出来，最终呈现在消费者面前的。

数据分析是指用适当的统计、分析方法对收集来的大量数据进行整理、加工，获取有用信息、形成结论的过程。数据分析的目的是把大量数据中的信息提炼出来，从而找出所研究对象的内在规律。数据分析是数学、统计学、计算机科学等相关学科相结合的产物。

2. 数据分析的作用

数据分析的应用越来越广泛，日益受到企业重视。通过对收集到的数据进行有效分析，能够为企业提供决策支持，助力精准营销，从而使企业不断发展壮大。

(1) 了解行业竞争格局。企业进入某个行业之前，先要了解行业的整体情况，通过行业数据分析，能够了解行业规模、市场需求、发展趋势。任何行业领域都存在着市场竞争，通过竞争数据分析，能够明确竞争对手、了解竞争对手的经营管理活动，明确自身在竞争中所处的地位，借鉴竞争对手的经营策略，规划自身的发展方向，使企业保持良好的竞争优势。

(2) 了解企业现状。通过对企业各项业务指标的收集、整理和分析，企业能够明确自身的经营现状，评估企业的运营状态，发现企业在经营管理中存在的问题，查找原因，及

时调整经营策略。

(3) 洞悉用户。通过数据分析可以帮助企业了解用户来源、用户特征、购物偏好、购物行为路径等，进而明确用户分类和用户价值，针对不同用户开展不同的活动，不断吸引新用户，同时提高老用户的忠诚度，建立良好的用户关系。通过用户对产品的需求分析，企业可以改善现在的产品、服务，也可以开发新的产品和服务，以满足用户需求，提升用户体验。

(4) 宏观预测。通过数据分析能够对行业和企业未来的发展趋势做出科学有效的预测，有助于帮助企业制定合理的经营目标，及时调整企业的运营策略。

(5) 通过数据化运营提高工作效率。通过数据分析和综合研判，能将问题数据化、直观化，为运营决策提供更多的数据支撑，避免凭借经验，想当然地进行决策，从而使决策更加科学、更加合理。在现代企业中，数据是科学决策的依据，能提升运营决策的效率；在业务的数据化基础上，能提高运营工作的协作与沟通效率，可以提升运营人员的业务分析效率，提高对业务问题的反应速度和响应效率。

3. 发展前景

随着科学技术的发展，数据分析的应用越来越广泛，数据分析的重要性和价值不断彰显，数据分析行业的发展前景也越来越广阔。

(1) 发展趋势。“大数据”搜索指数趋势如图 1-2 所示。

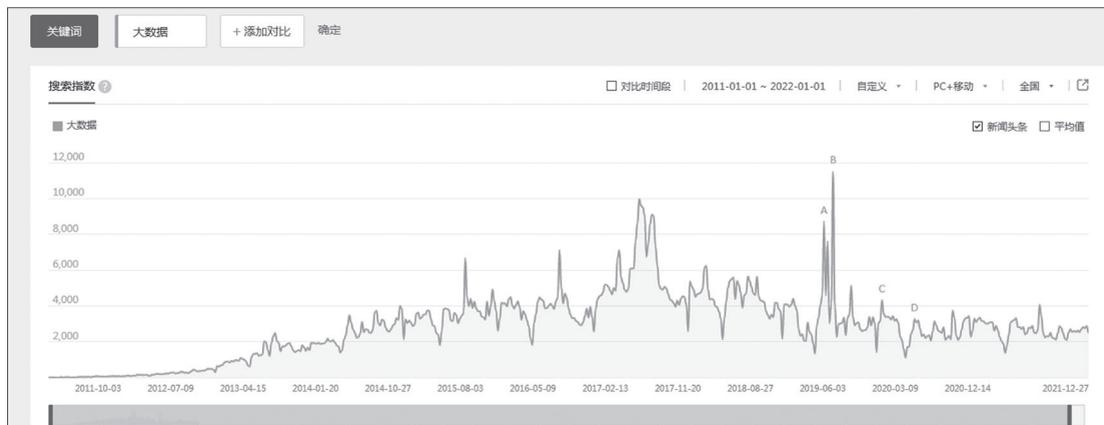


图 1-2 “大数据”搜索指数趋势

图 1-2 展示了 2011 年 1 月 1 日至 2022 年 1 月 1 日的搜索指数变化趋势。

在图 1-2 中，“大数据”的搜索指数从 2011 年开始逐渐提高，增长迅猛，2019 年 4 月搜索指数为 11 474，达到最高值。根据国际权威机构 Statista 的统计和预测，2020 年全球数据产生量达到 47 ZB，而到 2035 年，这一数字将达到 2 142 ZB，全球数据量将迎来更大规模的爆发。

2018 年 1 月，在麦肯锡 (McKinsey) 咨询公司的研究报告 *Analytics Comes of Age* 中，以大数据分析为代表的高级分析技术对于销售与营销、研究与开发 (R&D)、包括分销在内的供应链管理 (SCM)、工作场所管理与运行等领域的营收促进效果最为明显。McKinsey 认

为，大数据和其他分析技术将以 AI 为技术基础，以创建出全新的生态系统。而在具体行业方面，大数据等分析技术在基础材料和高科产业中做出了最有价值的贡献。

波士顿咨询公司（BCG）发现，通过利用大数据，一家分散型的日用百货零售商的销售额将增长 3% ~ 4%。大数据将成为零售业转型的催化剂，并为本地化带来比以往程度更高的精确度。如今，大数据正在被零售商用来增加投资回报率、优化规划路径图、提升销售与交叉销售的业绩，并同时优化商品价格。

数字经济是指通过对大数据的处理与使用，实现经济高质量发展的经济形态。随着我国科技的不断发展，在国家政策的大力支持下，大数据、云计算、物联网、区块链、5G 通信等新兴技术的研发，在我国多个行业的落地使用，中国数字经济规模将进一步发展壮大，其发展前景十分广阔。2017—2022 年我国数字经济总体规模统计预测数据如图 1-3 所示。

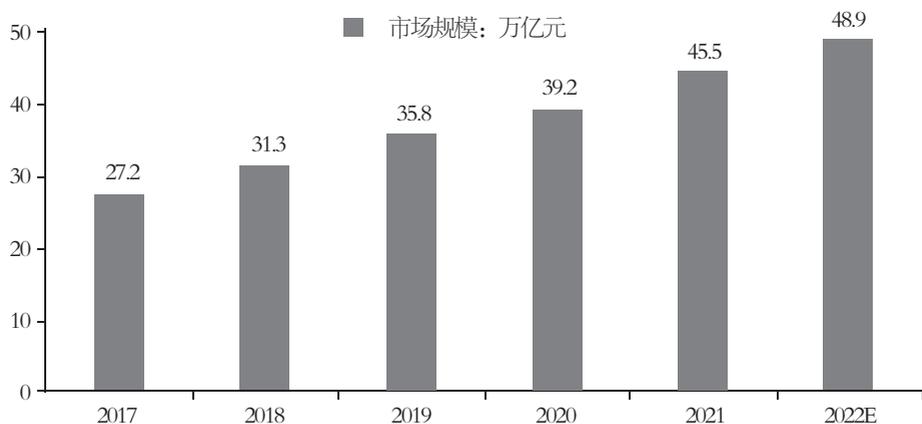


图 1-3 2017—2022 年我国数字经济总体规模统计预测数据

通过图 1-3 中的数据可知，近年来，我国数字经济依然保持蓬勃发展态势，2020 年数字经济规模达到 39.2 万亿元，2021 年达到 45.5 万亿元，中商产业研究院预测，2022 年我国数字经济规模将达 48.9 万亿元。

（2）人才需求。随着企业对数据科学与分析的重视，数据分析已成为企业的常规业务，企业对于数据分析人才的需求持续增长。面对人才需求的不断增长，近年来很多大专院校开设了数据科学、数据分析等相关专业，开发了数据科学与数据分析等课程，以满足市场需求。

某招聘网站发布的《2020 人才资本趋势报告》（以下简称《报告》）指出，随着各种新兴行业崛起，人工智能、大数据岗位、电竞玩家、无人机飞手、带货主播等岗位从无到有，从有到“热”。企业需要数据分析人员具有大数据的采集和存储能力，大数据的分析和转化能力，数据分析师更多地扮演了一个团队大脑或者指挥官的角色，将存储的数据利用起来。企业采集数据、成立数据团队的目的是利用它来进行商业决策，最终变成商业利润，这也是数据分析人员的责任。可以说数据分析师作为时代催生的新职业，蕴藏着极大的发展前景。

《报告》指出，越来越多的职场人开始重视技能为自己职业生涯带来的变化，他们或通过强化已有技能增加本岗位竞争力，或培养新的技能来拓展职业边界。技能的选择和掌握程度已成为职场人能否打破职业天花板，突破能力上限的决定因素之一。

理论上，求职者掌握的技能越多，对于自身能力提高和求职时所能获得的薪资议价空间帮助越大。近年来，超过七成专业技能的竞争力绝对值同比有所上涨。技能在职场中发挥的作用日益明显，更多企业不再单纯依赖求职者的资历，而是以核心技能为导向筛选求职者。部分岗位对数据类技能尤其看重。对数据类技能看重程度提升最快的职位如表 1-1 所示。

表 1-1 对数据类技能看重程度提升最快的职位

排 序	职 位	排 序	职 位
1	营销经理	11	媒介投放
2	项目经理	12	社区运营
3	采购主管	13	BD 经理
4	酒店经理	14	网络推广
5	仓储管理	15	UE 设计师
6	销售顾问	16	游戏运营
7	商家运营	17	海外市场
8	交互设计师	18	商务渠道
9	产品经理	19	活动策划
10	生物制药	20	运营专员

任务

1.2

数据分析指标与流程

在数据分析的过程中，要达到数据分析的最终目标，需要采集、整理与分析目标相适应的指标，并按一定的流程展开分析工作。

1.2.1 常用的数据分析指标

数据分析的核心工作是对各项指标进行分析，然后得出结论。在数据分析的过程中会涉及很多指标，因此要理解关键的常用指标的含义。数据分析指标很多，下面介绍一些常用的数据分析指标。

1. 流量类指标

网站流量能够体现一个网站的用户数量、浏览页面数量的多少。网站流量可以真实反映网站运营状况，能帮助运营人员发现问题，改进策略。

(1) 独立访客。独立访客（unique visitor，UV）是指某站点被多少用户访问过，是衡量网站用户数量的重要指标。可以分时段统计 UV 指标，但一般相同的客户端（同一 IP 地址）在 00:00 ~ 24:00 如果发生多次访问，UV 只被计算 1 次。某网店近两周 UV 数据变化情况如图 1-4 所示。

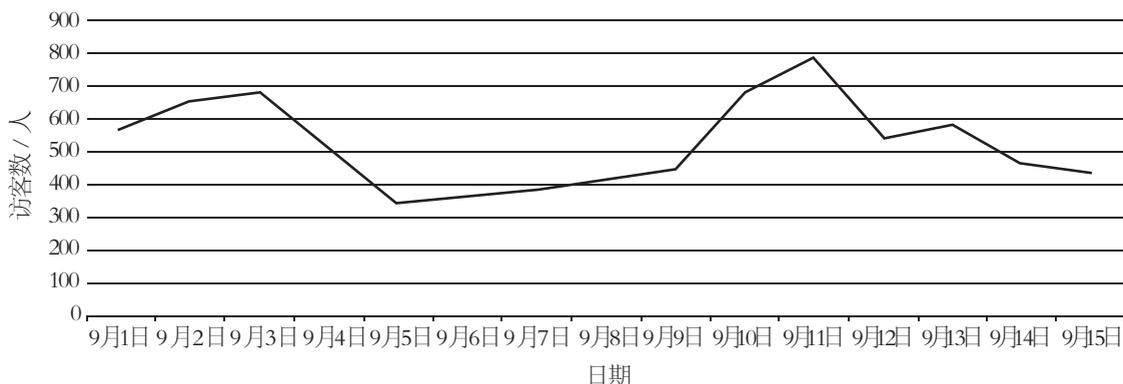


图 1-4 某网店近两周 UV 数据变化情况

(2) 访问量。访问量（page view，PV）即页面浏览量或点击量。用户对每个页面的每次点击都被计算一次，它是重复累计的，同一个页面被同一用户的重复点击浏览也被计入 PV。

(3) 新访客占比。新访客是指首次访问某网站或店铺的客户。新访客占比是新访客数

与总访客数的比值，新访客数占比反映的是网站或店铺吸引新客户的能力。新访客占比高，说明营销推广效果好。其计算公式为

$$\text{新访客占比} = \frac{\text{新访客数}}{\text{总访客数}} \times 100\%$$

(4) 访问深度。访问深度是指用户在浏览网站的过程中每次所浏览的网站页数，也就是用户在每次访问中打开了多少个页面。用户访问深度数值越大，说明用户对网站的内容越满意、用户体验度越好；但是如果用户访问深度过高，可能表明用户在浏览的过程中找不到目标页面，因此访问深度也不是越高越好。用户访问深度可以用来衡量用户访问质量的优劣。其计算公式为

$$\text{访问深度} = \frac{\text{PV}}{\text{访问次数}} \times 100\%$$

【例 1-1】

某服装网店某天共有 3 位访客，小张、小王和小李。小张当天访问网站两次，上午 9 点第一次浏览了网站首页，打开一件连衣裙的详情页，又打开一件女式开衫的详情页，然后离开；下午 2 点小张第二次进入该店铺，浏览首页后离开。小王上午 10 点进入店铺，浏览首页后进入一件女式衬衫详情页，又进入一件半身裙详情页，然后离开。小李下午 1 点进入店铺，浏览首页后离开。



图文

【例 1-1】答案

根据上面的描述统计 UV、PV、访问次数、访问深度。写出答案并扫描二维码核实。

(5) 页面停留时长。页面停留时长是指访客浏览某一页面时所花费的时长。其计算公式为

$$\text{页面停留时长} = \text{进入下一个页面的请求时间戳} - \text{当前页时间戳}$$

页面停留时长越长，说明网站或页面对用户的吸引力越强，越受用户喜爱。反之，说明页面对客户的吸引力小，对网站或店铺来说，需要优化和完善页面信息。需要注意的是，页面停留时长也不是越长越好，用户如果在首页的停留时间过长，说明该页面的引导性可能不够好。

(6) 跳出率。跳出是指仅访问了入口页面就离开网站的情况，跳出率是跳出的访问量与网站总访问量的百分比。跳出率是用来衡量网站质量的指标，反映的是网站的吸引力和用户对网站内容的满意度。当跳出率升高到一定程度时，需要对网站的结构或内容做出优

化和调整。其计算公式为

$$\text{跳出率} = \frac{\text{跳出的访问量}}{\text{网站总访问量}} \times 100\%$$

2. 销售类指标

- (1) 订单量。订单量是指统计时间内用户提交订单的数量。
- (2) 商品销量。商品销量又称销售件数，是指统计时间内销售商品的总数量。

【例 1-2】

某服装网店某天共有 3 位访客，小张、小王和小李。小张这天共提交了 2 个订单，第 1 个订单包含 3 件商品，第 2 个订单包含两件商品。小王提交了 1 个订单，包含 1 件商品。小李提交了 2 个订单，第 1 个订单包含 1 件商品，第 2 个订单包含 3 件商品，那么，这个网店当天共有几个订单？销售了几件商品？

统计某网店的订单量与商品销量。写出答案并扫描二维码核实。



图文
【例 1-2】答案

(3) 成交总额。成交总额 (gross merchandise volume, GMV) 是衡量平台竞争力的核心指标。一般电商平台 GMV 的计算公式为

$$\text{GMV} = \text{销售额} + \text{取消订单金额} + \text{拒收订单金额} + \text{退货订单金额}$$

即 GMV 为已付款订单和未付款订单两者之和。

GMV 虽然不是实际的交易数据，但可以作为参考依据来研究顾客的购买意向、顾客购买之后发生退单的比率、GMV 与实际成交额的比率等。

(4) 销售额。企业或店铺在销售过程中会产生大量的销售数据，销售额是其中非常重要的一个指标。企业要实时关注销售额变化情况，以便及时调整销售策略。对于网店来说，销售额计算公式为

$$\text{销售额} = \text{展现量} \times \text{点击率} \times \text{转化率} \times \text{客单价}$$

或

$$\text{销售额} = \text{访客数} \times \text{转化率} \times \text{客单价}$$

或

$$\text{销售额} = \text{客单价} \times \text{成交客户数}$$

(5) 客单价。客单价是指在一定时间内，每个客户平均购买商品的金额，即平均交易金额。客单价是影响店铺盈利的因素之一，在流量稳定的情况下，客单价越高，销售额就越高。客单价的计算公式为

$$\text{客单价} = \frac{\text{成交金额}}{\text{成交用户数}} \times 100\%$$

(6) 产品页转化率。客户进入网站后，随着购物的进行，可以通过不同的转化率指标来衡量客户在访问页面、加购、结算等方面的实际情况，为企业进一步做好各类转化工作提供数据支持。产品页转化率根据实际情况，有以下两种计算方法。

① 方法一。

$$\text{产品页转化率} = \frac{\text{产品页访问次数}}{\text{总访问次数}} \times 100\%$$

② 方法二。

$$\text{产品页转化率} = \frac{\text{产品页 UV}}{\text{总 UV}} \times 100\%$$

(7) 加购转化率。加购转化率是指把商品加入购物车的访问量占总访问量的比例。根据实际情况，加购转化率有以下两种计算方法。

① 方法一。

$$\text{加购转化率} = \frac{\text{加入购物车的访问量}}{\text{总访问量}} \times 100\%$$

② 方法二。

$$\text{加购转化率} = \frac{\text{加入购物车 UV}}{\text{总 UV}} \times 100\%$$

(8) 结算转化率。结算转化率是指客户进入结算页面的占比情况。其计算公式为

$$\text{结算转化率} = \frac{\text{进入结算页 UV}}{\text{总 UV}} \times 100\%$$

(9) 下单转化率。下单转化率是指产生订单的访问量和总访问量的比值。下单转化率高，说明客户的购买意愿高。其计算公式为

$$\text{下单转化率} = \frac{\text{下单的 UV}}{\text{总 UV}} \times 100\%$$

或

$$\text{下单转化率} = \frac{\text{下单访问量}}{\text{总访问量}} \times 100\%$$

(10) 支付转化率。支付转化率是指已经支付的客户和需要支付的用户数的比值，支付转化率越高说明企业的收益越高。其计算公式为

$$\text{支付转化率} = \frac{\text{已支付的用户数}}{\text{需要支付的用户数}} \times 100\%$$

(11) 利润。利润是收入与成本之差。

(12) 标准产品单位 (standard product unit, SPU)、库存量单位 (stock keeping unit, SKU)。SPU 是商品信息聚合的最小单位，是一组可复用、易检索的标准化信息的集合，该集合描述了某个产品的特性。对电商而言，SKU 是指一款商品，每款都有一个，便于电商品牌识别商品。例如，某款手机有黑、白、灰、红四种颜色，内存分别有 64 GB、128 GB、256 GB 三种大小，那么这个手机的 SKU 数为 12。

分析产品 SKU，可以帮助企业了解客户喜好，为企业制定销售策略、价格策略等方面提供依据，同时根据 SKU 历史数据，可以为进货阶段提供参考。

3. 推广类指标

(1) 展现量。展现量是指在一段时间内广告在搜索结果中出现的次数，也就是广告对用户展示的次数，展现量越大，说明广告被用户点击的概率越大，从而为店铺带来更多流量。

(2) 点击量。点击量指一段时间内某个广告被用户点击的次数，用户每点击 1 次，就记录 1 次；每点击 1 次，企业就要付出相应的费用。点击量是衡量广告投放效果的重要指标。

(3) 点击率。点击率是点击量与展现量之比，点击量越大，点击率越高。点击率越高，说明广告投放效果越好。

(4) 投入产出比。投入产出比是成交总额除以投入成本之值，是衡量投入产出效果的指标。投入产出比越高，说明收益越好；投入产出比低，说明要调整优化推广策略。

1.2.2 数据分析流程

数据分析对企业的决策和发展十分重要，应用越来越广泛。在数据分析过程中，不同部门、不同人员对数据分析的要求不同、目标不同，涉及的数据指标也就有所不同，但数据分析流程是一致的。数据分析流程包括明确分析目标、采集获取数据、数据加工处理、数据分析、数据可视化、撰写数据分析报告等几个方面。

1. 明确分析目标

目标指引方向。只有明确分析目标，才能使数据分析工作更有针对性，不偏离方向，才能为决策者提供科学的、有价值的指导意见。

明确了数据分析的目标之后，要根据目标的大小，决定是否要对总目标进行分解，以厘清思想，逻辑清晰地开展数据分析工作。

2. 采集获取数据

数据分析人员在明确分析目标之后，需要采集、收集、整理数据。数据的获取渠道有直接和间接两种，一般包括但不限于店铺的后台、平台提供的数据分析工具、相关权威网站发布的数据和报告、政府部门、机构协会、媒体发布的数据，统计工具、指数工具，以及市场调查数据等。

3. 数据加工处理

通过数据采集，数据分析人员获得了大量相关数据，这其中并不是所有数据都有意义，需要数据分析人员对无效的、无价值的数据进行删除、清理、转换等加工和处理。

4. 数据分析

数据分析就是对加工处理后的数据，运用恰当的方法、合适的工具进行分析，得出有效结论的过程。

数据分析方法和分析工具在后面的项目会详细介绍，此处不再赘述。

5. 数据可视化

数据可视化是将数据分析的结果用图表的形式直观生动地表现出来，以便管理者、决策者更直观地看到数据分析的结果和重点。数据可视化一般根据数据分析结果所要表达的侧重点而灵活地选择柱形图、饼图、条形图、折线图、漏斗图等图表来展现。

6. 撰写数据分析报告

数据分析报告是对以上几项工作的展现、总结和提升。数据分析报告要求结构严谨、主题明确、图文并茂、有理有据、条理清晰。要形成有价值的结论，并提供支撑结论的数据。决策者和管理者能够通过分析报告找到有价值的信息，以辅助决策。

任务

1.3

数据分析模型

数据分析模型简单地说就是数据分析思路、分析框架。在进行数据分析之前，需要先搭建数据分析模型，根据模型所包含的内容、指标，进行细化分析，并得出结论。常用的数据分析模型有以下几种。

1.3.1 5W2H 模型

5W2H 模型就是由 why、what、who、when、where、how、how much 7 个单词的首字母命名的模型，5W2H 模型主要用于用户行为分析、业务问题专题分析、营销活动分析等。

- (1) why——用户为什么购买产品？
- (2) what——用户购买了什么产品？
- (3) who——购买产品的用户有哪些特点？
- (4) when——用户在何时购买？
- (5) where——用户购买产品的渠道是什么？
- (6) how——用户的购买方式、支付方式是什么？
- (7) how much——用户购买产品的花费是多少？

使用 5W2H 模型可以对用户的购买偏好进行分析。5W2H 分析模型的应用如表 1-2 所示。

表 1-2 5W2H 分析模型的应用

项 目	含 义
why	购买原因
what	购买了哪些产品
who	购买者的年龄、性别、职业等特点
when	购买时间、决策时间
where	获取信息的渠道、购买地域等
how	支付方式
how much	了解程度、花费的时间成本等

1.3.2 生命周期分析模型

生命周期分析模型分为用户生命周期分析模型和产品生命周期分析模型。用户生命周期是指用户从引入到流失的过程，产品生命周期（product life cycle）是指产品从萌芽到退

出市场的过程。

1. 用户生命周期分析模型

用户生命周期是用户从接触产品开始到离开产品的整个过程。用户生命周期包括引入期、成长期、成熟期、休眠期、流失期 5 个阶段。按照用户生命周期来划分用户，能够帮助企业了解不同生命周期用户的需求，从而有针对性地采取运营策略，提高活跃客户数量，减少用户流失。用户生命周期如图 1-5 所示。

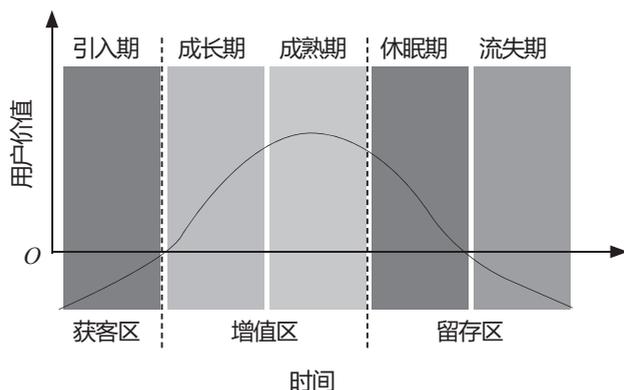


图 1-5 用户生命周期

作为企业的重要资源，客户具有价值和生命周期。客户生命周期模型是指从企业与客户建立业务关系到完全终止关系的全过程，是客户关系水平随时间变化的发展轨迹。对于不同时期的客户，客户为企业带来的利润不同，对应的运营策略也有所不同。

(1) 引入期。企业只能获得基本的利益，客户对企业的贡献不大。

(2) 成长期。客户开始为企业做贡献，企业从客户交易获得的收入大于投入，开始盈利。

(3) 成熟期。客户愿意支付较高的价格，带给企业的利润较大，而且由于客户忠诚度的增加，企业将获得良好的间接收益。

(4) 休眠期和流失期。客户对企业提供的价值不满意，交易量回落，客户利润快速下降。

企业构建用户生命周期模型的目标是增加用户数、提高每个用户月均消费、延迟用户生命周期。企业要判断各个用户目前所处的生命周期，进而进行精细化运营，提高用户价值。

2. 产品生命周期分析模型

产品生命周期是指产品从准备进入市场开始到被淘汰退出市场为止的全部运动过程，即一个产品从诞生到衰退的过程，关注产品的生命周期有利于更好地调整发展策略，促进产品发展。产品生命周期主要由消费者的消费方式、消费水平、消费结构和消费心理的变化所决定，一般分为引入期、成长期、成熟期、衰退期 4 个阶段。产品生命周期如图 1-6 所示。

(1) 引入期。新产品上市，销售缓慢。初期通常利润偏低或为负数，此时，顾客对产品尚不了解，主要的消费者为愿意尝鲜的少数个体，因此销售量很低。为了扩展销路，需要大量的资金用于产品推广、宣传。在这一时期，产品销售额增长缓慢，多数企业可能不仅没有多少盈利，还会有相当的亏损。同时产品刚投入市场，可能存在多种问题，产品质

量和品质也存在进一步提升的空间。

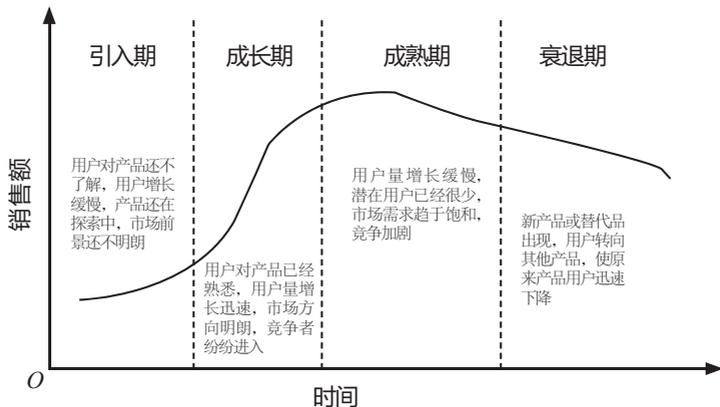


图 1-6 产品生命周期

引入期的产品一般都是利用最小化可行产品（minimum viable product, MVP）来验证猜想，并且根据用户的反馈和数据来对产品进行迭代和调整，在导入期阶段一般只做最基础的核心功能，解决用户最核心的需求，并不很注重交互、视觉和用户体验等，首先保证产品是有用的，其次才是更高层次上的好用和易用。

在引入期阶段一般是不追求用户数量的快速增长的，一方面是因为产品本身的功能和体验并不完善，有许多需要打磨的地方，而早期的服务器也可能承受不了蜂拥而入的用户带来的压力；另一方面由于涌入的用户中很有可能带来大量的非目标用户，这些用户对当前不完善的产品的吐槽和不好的口碑会为后期产品的推广带来很多的隐患。

(2) 成长期。产品经过一段时间已有相当知名度，销售快速增长，利润也显著增加。但由于市场及利润增长较快，容易吸引更多的竞争者。此时顾客对该产品已经有了大致的印象，越来越多的客户愿意试用购买此产品，市场逐步扩大。产品大批量生产，生产成本相对降低，企业的销售额迅速上升，利润也迅速增长。竞争者看到有利可图，将纷纷进入市场参与竞争，使同类产品供给量增加，价格随之下降，企业利润增长速度逐步减慢，最后达到生命周期盈利能力的最高点。

经过引入期的过渡，产品本身的猜想已经得到了验证，并且也有一定的时间来调整和迭代产品。如果产品本身能够实现自增长，那说明产品是能够满足用户需求的，可以进行运营推广。而如果产品本身就不能实现自增长，说明产品很有可能不被用户所认可，强行利用运营手段来推广产品很有可能会加速产品的死亡过程。

成长期的产品是已经得到了市场的验证，并且已经有一定的用户量和知名度，在这个阶段也会有一些竞争者的加入。在成长期阶段主要的目标就是快速占领市场份额，抢占用户，充分利用先发优势，并且构筑一定的竞争壁垒来阻碍竞争对手的跟进。

(3) 成熟期。此时市场成长趋势减缓或饱和，产品已被大多数潜在购买者所接受，利润在达到顶点后逐渐走下坡路。市场竞争激烈，公司为保持产品地位需投入大量的营销费用。当潜在的客户越来越少，市场需求趋向饱和，销售额增长愈加缓慢，甚至有下降的趋

势，此时被称为成熟期。在这一阶段，竞争逐渐加剧，产品售价降低，促销费用增加，企业利润下降。

(4) 衰退期。在衰退期，产品销售量显著衰退，利润也大幅度滑落。优胜劣汰，市场竞争者也越来越少。更多替代品的出现、受到潮流的影响、抑或是多种多样的原因，顾客的消费习惯发生改变，转向其他产品，从而使原来产品的盈利能力迅速下降。于是，产品逐渐走向衰退，最终退出市场，整个产品生命周期完结。

1.3.3 4P 策略模型

4P 营销理论是产品 (product)、价格 (price)、渠道 (place)、促销 (promotion) 的组合，由于这 4 个英文单词的首字母都是 P，所以简称为 4P。如果再加上策略 (strategy)，就简称为“4P’S”。

1. 产品

产品是根本，产品在具备完备的功能的同时，要有能吸引客户的独特的卖点，要注重产品的外观和颜值。在质量功能相同的情况下，好看的东西更有吸引力。

2. 价格

依据企业定位和品牌战略，根据不同的市场定位，制定不同的价格策略，注重品牌形象和品牌含金量。产品的价格与客户所感知的产品价值相关，产品的定价会影响产品的销售方式。

3. 渠道

传统企业通过经销商来搭建销售网络，企业与消费者的联系是通过分销商来进行的，电商企业采用线上分销。

4. 促销

促销包括品牌宣传 (广告)、促销、打折、公关等一系列的营销行为。无论采取哪种促销方式，一定要确保这种方式是合适的。

4P 营销策略如图 1-7 所示。

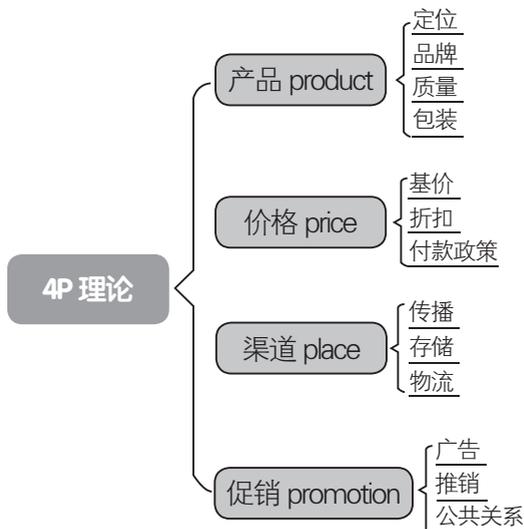


图 1-7 4P 营销策略

1.3.4 PEST 模型

PEST 模型是企业所处宏观环境分析模型。PEST 模型能较好地把握宏观环境的现状及变化趋势，有利于企业对生存发展的机会加以利用，对环境可能带来的威胁及早发现避开。所谓 PEST，是指政治（politics，P），经济（economy，E），社会（society，S），技术（technology，T）。这些是企业的外部环境，一般不受企业掌控，对企业高级管理人员来说，PEST 分析能力是必备的能力之一。

1. 政治环境

政治环境是指一个国家或地区的政治制度、体制、方针政策、法律法规等方面。这些因素影响着企业的经营行为和企业长期的投资行为。企业在制定战略时，要充分了解既有的制度、政策、规定等。

2. 经济环境

经济环境是指国民经济发展的总体情况和所在区域的经济状况，包括宏观和微观两个层面。主要指标包括人口状况、国民生产总值、国民收入、居民消费、就业形势、不同地区的消费者收入水平、消费偏好、储蓄情况等。

3. 社会环境

社会环境是指一定时期整个社会发展的一般状况。主要包括社会道德风尚、文化传统、人口变动趋势、文化教育、价值观念、社会结构等。

4. 技术环境

技术环境是指社会技术总水平及变化趋势，技术发展变迁，技术突破对企业的影响，以及技术对政治、经济社会环境之间的相互作用的表现等。

PEST 模型如图 1-8 所示。



图 1-8 PEST 模型

1.3.5 SWOT 分析模型

SWOT 是指优势（strengths，S），劣势（weaknesses，W），机会（opportunities，O），威

胁 (threats, T)。SWOT 分析模型是对内外部竞争环境和竞争条件的综合分析。SWOT 分析把与研究对象密切相关的内部优势、劣势及外部的机会、威胁等分别列举出来,然后把各种因素相互匹配起来加以分析,从中得出一系列相应的结论。运用这种方法,可以对研究对象所处的情景进行全面、系统、准确的研究,从而根据研究结果制定相应的发展战略、对策等。

SWOT 分析模型如图 1-9 所示。

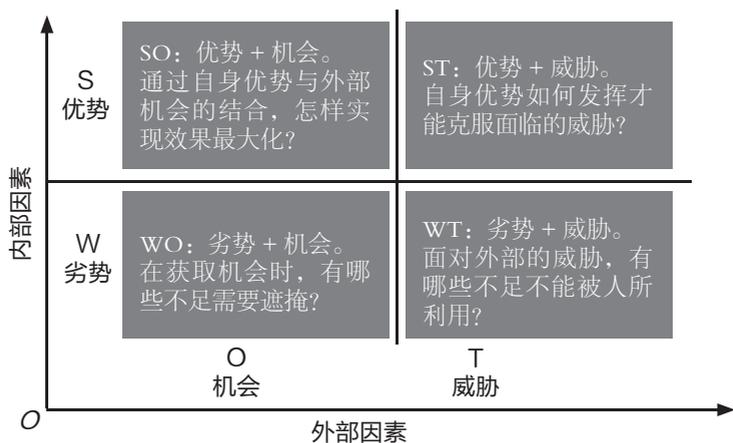


图 1-9 SWOT 分析模型

课堂讨论

SWOT 分析模型从某种意义上来说隶属于企业内部分析,即根据企业自身的条件在既定内进行分析。但是 SWOT 分析模型并不仅仅适用于企业分析,对个人同样适用。例如,我们在求职时,可以利用 SWOT 分析模型分析我们自身的优势、劣势、机会和威胁,从而达到应聘成功的目的。

请运用 SWOT 分析模型对自身进行分析,并把结果写在下面。

1.3.6 逻辑树分析模型

逻辑树又称问题树、演绎树或分解树等。逻辑树是将问题的所有子问题从最高层开始罗列,逐步向下层扩展。逻辑树分析模型是将一个已知问题作为树干,罗列与此相关的项目,每想到一个,就给这个问题加一个“树枝”,并标明这个“树枝”代表是什么项目。一个大的“树枝”上还可以有小的“树枝”,逐步找出与问题相关的所有项目。

逻辑树分析模型条理分明,便于将工作细分,并确定各部分的优先顺序,然后落实责

任。逻辑树分析模型如图 1-10 所示。



图 1-10 逻辑树分析模型

1.3.7 RFM 模型

RFM 模型是衡量客户价值和客户创利能力的重要模型。该模型通过客户的近期购买行为、购买的频率及花费三项指标来分层客户，运用各类运营手段提高不同类型的客户的活跃度、留存率和付费率，提升用户忠诚度和复购率，促进营收增长。RFM 模型常用于交易类企业的用户分群或价值区分。

RFM 模型是基于一个固定的时间点来做分析，因为每个用户在不同时间节点的数据不同，所以不同时间所做的 RFM 分析结果也有所不同。RFM 的具体含义如下。

(1) R (recency)：用户最近一次消费距离现在的时间。最近一次消费与上一次消费时间相隔越短，则 R 值越小，用户价值越高。消费时间越近的客户价值越大，因此企业要通过有效的营销手段来激发用户的购买行为。

(2) F (frequency)：消费频率。消费频率是指用户在统计周期内购买商品的次数，经常购买的用户即老客户，消费频率越高，那么用户活跃度越高，F 值越大，消费频率高的客户价值大于消费频率低的客户。

(3) M (monetary)：消费金额。消费金额是指用户在统计周期内消费的总金额，体现了消费者为店铺创利的多少，消费金额越高，用户价值越大。

把 R、F、M 三个指标按其价值的高低排序，可以将用户分为八类。RFM 模型如图 1-11 所示。

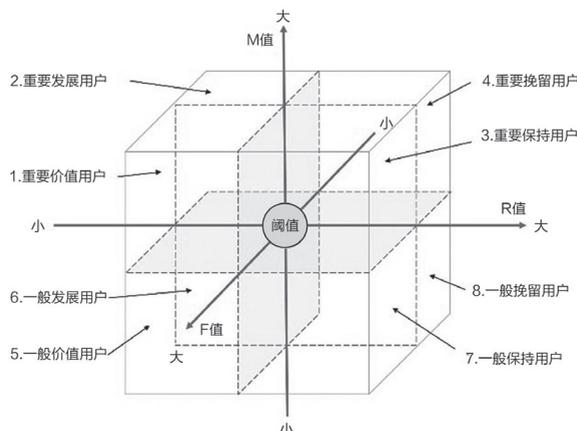


图 1-11 RFM 模型

用户分类的含义如表 1-3 所示。

表 1-3 用户分类的含义

用户分类	R 价值	F 价值	M 价值
重要价值用户	高	高	高
重要发展用户	高	低	高
重要保持用户	低	高	高
重要挽留用户	低	低	高
一般价值用户	高	高	低
一般发展用户	高	低	低
一般保持用户	低	高	低
一般挽留用户	低	低	低

1.3.8 漏斗模型

漏斗模型是做运营数据分析中最基础也是最有效的数据分析模型之一。通过漏斗模型能够明确用户体验过程中每一步的流失和转化情况，从而找到提升转化率的应对策略。

漏斗分析模型是目前关于提升转化率最经典的模型。在营销过程中，展示了将潜在用户逐步变为用户的转化过程。以电商产品为例，用户从搜索商品关键词进来，逐渐经历商品选择、选择商品类型和数量、咨询客服、加入购物车、购买支付等环节。

根据用户不同的搜索记录及以往的购买记录，电商平台会将最适合用户搜索关键词的产品放在第一位，进行千人千面的展示，提升产品转化率，这也是基于漏斗模型的原理。电商产品的漏斗模型如图 1-12 所示。

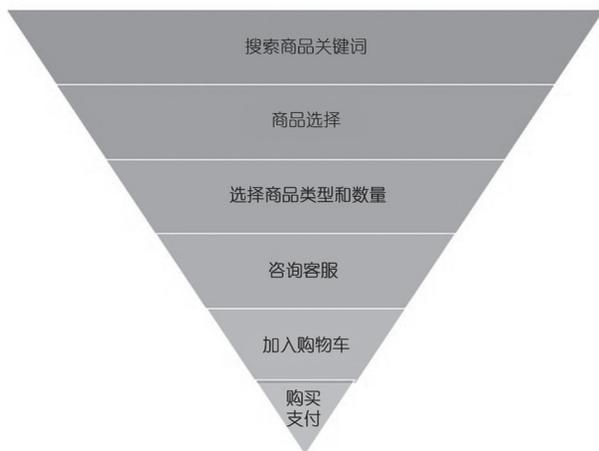


图 1-12 电商产品的漏斗模型

任务

1.4

数据分析方法

数据分析方法是指数据分析工作如何具体的实施，常用的方法有对比分析法、结构分析法、同期群分析法和聚类分析法等。

1.4.1 对比分析法

对比分析法是指将两个或两个以上相关的数据进行对比，来发现、预测研究对象的发展变化规律，找到自身存在的问题，寻求解决方法。对比分析法可分为静态对比和动态对比。静态对比通常指相同时间内，不同部门、不同地域的对比，也叫横向对比；动态对比是指对不同时期数据的对比，也叫纵向对比。

对比分析法可以按时间对比，如本月和上月对比是环比，本年1月和上年1月对比是同比；所有数据同当年的第一个月对比则为定基比。通过这3种方式，可以分析业务增长的速度、能力等信息。

将当前计划完成情况与年度目标进行对比，可以检验是否完成了目标；将自身的指标、业绩与行业平均水平、竞争对手或上一层级的标杆进行对比，可以了解自身在行业内所处的地位和水平，便于更好地制定自己的发展目标。

1.4.2 结构分析法

结构分析法是对各组成部分所占比重进行分析，进而分析某种现象的内部结构特征、总体性质、内部结构随时间变化而表现出的规律性的分析方法。结构分析法的基本表现形式，就是计算结构相对指标。结构相对指标的计算公式为

$$\text{结构相对指标（百分比）} = (\text{总体中某一部分的值} / \text{总体总量}) \times 100\%$$

结构相对指标就是总体中各个部分占总体的比重，因此总体中各个部分的结构相对数之和等于100%。

1.4.3 同期群分析法

同期群分析法是指在规定时间内对具有共同行为特征的用户进行分群。共同行为特征是指在某个时间段内的相似行为，可以按不同时间的新增用户来分类，也可以按不同的行为来分类，如第一次购买时间相同、在同一时间段内使用频率升高等。

同期群分析法侧重于分析在用户生命周期相同阶段的群组之间的差异，通过对性质一致的可对比群体的留存情况的比较，剖析真实的用户行为和用户价值，具体分析哪些因素

影响用户的留存。

同期群分析法十分直观，只用简单的一个图表，就可以说明用户在一段时间内周期的留存或流失变化情况，如表 1-4 所示。

表 1-4 同期群分析

新增	1 个月后	2 个月后	3 个月后	4 个月后	5 个月后
2 000	60%	55%	40%	35%	30%
3 000	75%	55%	35%	30%	
2 000	75%	55%	40%		
3 000	75%	55%			
2 000	75%				

1.4.4 聚类分析法

聚类分析是指将物理或抽象对象的集合分成由类似的对象组成的多个类的分析过程。聚类分析是一种探索性的分析，在分类的过程中，人们不必事先给出一个分类的标准，聚类分析能够从样本数据出发，自动进行分类。聚类分析所使用方法的不同，常常会得到不同的结论。

聚类就是针对大量数据或者样品，根据数据本身的特性研究分类方法，并遵循这个分类方法对数据进行合理的分类，最终将相似数据分为一组，也就是“同类相同、异类相异”。聚类最关键的工作是探索和挖掘数据中的潜在差异和联系。

聚类分析法具有简单、直观的特征，网站分析中的聚类主要分为用户、页面或内容、来源。用户聚类主要体现为用户分群、用户标签法；页面聚类则主要是相似、相关页面组法；来源聚类主要包括渠道、关键词等。

在页面分析中，资讯详情页面、商品页面等都属于同一类页面。简单的分析容易造成跳出率、退出率等指标不准确的问题，通过聚类分析可以获取同类页面的准确数据用于分析场景。

聚类分析法如图 1-14 所示。

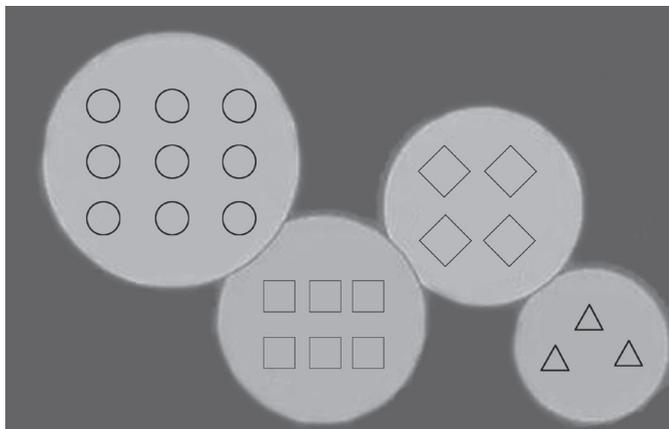


图 1-14 聚类分析法

课堂讨论

仔细分析和对比以上介绍的几种数据分析方法各有什么优点和缺点，请将你的思考写在下面。

课后练习 1

一、单选题

1. 聚类就是针对大量数据或者样品，根据数据本身的特性研究分类方法，并遵循这个分类方法对数据进行合理的分类，最终将相似数据分为（ ），也就是“同类相同、异类相异”。

- A. 一组 B. 两组 C. 三组 D. 四组

2. （ ）是库存量单位，对电商而言，它是指一款商品的编码，便于电商品牌识别商品。

- A. SUV B. SUA C. SPU D. SKU

3. GMV 是（ ）指标的简称。

- A. 成交总额 B. 销量 C. 客单价 D. 客单件

4. 同期群是指在规定时间内对具有共同行为特征的用户进行分群。“共同行为特征”是指在某个时间段内的（ ）。

- A. 同一行为 B. 不同行为 C. 相似行为 D. 以上都不对

5. 对比分析法可以按时间对比，如本月和上月对比是环比，本年1月和上年1月对比是（ ）。

- A. 同比 B. 环比 C. 定基比 D. 以上都不对

6. 漏斗模型是做运营数据分析中最基础也是最有效的数据分析模型之一，通过漏斗模型，能够明确用户体验过程中每一步的流失和（ ）情况，从而找到提升转化率的应对策略。

- A. 引流 B. 提升 C. 丢失 D. 转化

