



# 前言

## PREFACE

连锁经营被引入中国后,其发展速度可谓惊人,也为传统零售行业的发展提供了新的方向。连锁经营在高速发展的过程中,也引发了一系列问题,最典型的就总部对门店的控制不强,只连不锁。对连锁门店的控制能力缺失进一步说明连锁企业核心能力建设的缺失。从管理的角度来看,能够协调的资源只有人、财、物、信息四个方面。因此,本书围绕连锁管理的核心能力——信息管理能力进行课程设计。

互联网经过不断发展迭代,已经进入人工智能时代。众所周知,互联网对实体经济造成重大的冲击,连锁企业如何融合互联网,成为当代连锁企业必须解决的问题。闭门造车,不顾信息技术发展的企业只有等待破产和倒闭,并且今天的互联网企业已经进行线下资源整合,收购相关零售企业,开启“新零售”发展模式。“新零售”的重点不是线上到线下(Online to Offline, O2O),而是连接消费者购买的持续追踪,并利用消费者的购买数据在线下为消费者搭建购买场景,完成人(消费者)-货(商品)-场(场景)的对接。对消费者购买过程的持续追踪,说到底就是信息分析能力的建设。

因此,本书从连锁零售企业信息分析工作过程中的信息收集、信息处理、数据分析入手,介绍连锁零售业信息分析的具体方法;并从人、货、场三个层面,介绍连锁企业信息分析的内容,具体包括认识连锁企业信息分析,数据的收集、整理与呈现,不确定性分析,推断分析,零售信息分析。本书课时分配建议如下:

内 容	理论板块课时/h	实践板块课时/h
项目一 认识连锁企业信息分析	6	0
项目二 数据的收集、整理与呈现	8	4
项目三 不确定性分析	10	4
项目四 推断分析	10	4
项目五 零售信息分析	10	4
总学时(60 h)	44	16



本书针对高职高专学生的学情特点,重点讨论数据分析模型的应用和实际工作场景的使用,利用生活实例和企业案例形象化地说明数据统计模型在工作场景中的应用。因此,本书具有以下三个特点:

(1)遵循项目导向设计教材内容。本书重新解构信息分析、零售数据分析、统计学、概率论、市场调查学等内容,按照工作过程,将典型工作任务项目化,依据项目成果要求,重新组合以上内容,遵循成果导向,进行教学内容重组,并将我国传统的数字分析技术和数学思维融入课程,强化文化自信。

(2)按照理论和技术两条线组织教学内容。连锁企业信息分析主要分析连锁企业市场、竞争、顾客、商品、价格等经营信息,而这些经营信息由人、货、场、财等信息源组成,通过对信息源产生的数据进行对比、归纳、逻辑推理等分析设计连锁企业经营指标体系,通过指标计算训练学生对数据的敏感性和敏捷性。本书的一个主要任务是训练学生的数据处理能力,数列推导、两位数加减乘除、数独等内容贯穿始终,将数据“三分”(分类、分组、分布)能力作为重点训练能力,并能够在一组数据中找出大数、小数(组距)、众数、中数、均值、方差等差异数据的情况,从而进行分组、分类,通过概率分布,让学生了解数据分布的特点,进而推断未来数据的变化规律(统计分布律),并通过课后习题对知识内容和数据处理能力进行强化,达到学以致用目的。

(3)利用模型帮助学生建设问题处理范式。本书在理论分析中利用问题分析模型(如5W2H分析法、用户行为分析法、逻辑树分析法等)帮助学生建立以问题解决为导向的思维分析模式,从是什么—为什么—怎么样的角度提出问题的解决办法。本书利用概率基本模型帮助学生搭建数据分析模式,通过虚拟数据帮助学生建设人-货-场思维模型,让学生在实际工作中时刻关注门店周围环境的变化,时刻关注市场变化,掌握门店经营数据,做一个有数学思维的高水平管理者。

本书由辽宁现代服务职业技术学院邢学鹏编著。在编写本书过程中,笔者参考了大量国内外的研究成果和观点,在此向相关的作者表示衷心的感谢。

由于编者水平有限,书中难免存在不足之处,敬请广大读者批评指正。

编者

2023年5月



# 目录

## CONTENTS

### 项目一 认识连锁企业信息分析 1

学习目标	1
学习导航	1
项目导入	2
任务一 连锁企业信息分析概述	6
任务二 连锁企业信息分析的基本思路	15
任务三 连锁企业信息分析的方法	24
项目小结	37
过关练习	37

### 项目二 数据的收集、整理与呈现 39

学习目标	39
学习导航	40
项目导入	40
任务一 统计数据的收集	42
任务二 统计数据的整理	51
任务三 频数分布	55
任务四 统计数据的呈现	62
项目小结	71
过关练习	73

### 项目三 不确定性分析 75

学习目标	75
学习导航	76



项目导入	76
任务一 认识概率	80
任务二 概率分布	88
任务三 常用分布数字特征	101
项目小结	110
过关练习	110

## 项目四 推断分析 112

学习目标	112
学习导航	112
项目导入	113
任务一 总体和抽样	113
任务二 参数估计	124
任务三 假设检验	131
项目小结	145
过关练习	145

## 项目五 零售信息分析 147

学习目标	147
学习导航	148
项目导入	148
任务一 连锁企业的行业市场分析	152
任务二 洞察顾客需求	164
任务三 商品分析	175
项目小结	185
过关练习	185

## 参考文献 188

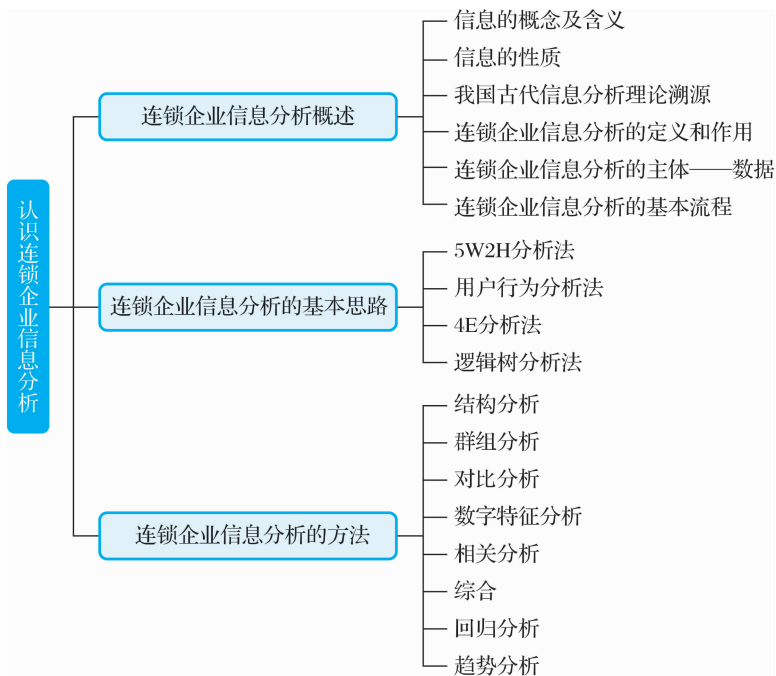
# 项目一

## 认识连锁企业信息分析

### 学习目标

明确信息的概念、连锁企业信息分析的定义,掌握连锁信息分析的基本思路和方法。

### 学习导航





### 数据可以创造竞争优势……也可能做不到<sup>①</sup>

许多高管和投资者理所当然地以为,利用客户数据可以获得无可比拟的强大竞争优势。客户越多,能够收集到的数据就越多。如果公司运用机器学习算法进行数据分析,就可以提供更好的产品,进而吸引更多的客户。客户增多,公司又可以收集更多的数据,最终像拥有强大网络效应的大企业一样,将竞争对手边缘化。虽然想得很美,但事实往往并非如此。大部分情况下,人们过度高估了数据带来的优势。

数据驱动的机器学习创造出的良性循环,乍看之下,与一般的网络效应相似:用户越多,价值越高(如社交媒体平台),最终获得大量的用户,令竞争对手望尘莫及。然而在现实中,一般网络效应持续时间更长、影响更大。要打造强大的竞争力,企业需要同时具备网络效应和数据驱动型学习,但能够同时发展这两项的公司少之又少。不过,即使没有网络效应,只要具备相应的条件,公司也可以利用客户数据驱动型学习建立防御策略以应对竞争。

以数据为基础的公司原本就有很多,并不是近几年才兴起的,如信用机构及信息汇总公司律商联讯(LexisNexis)、汤森路透(Thomson Reuters)和彭博(Bloomberg)等。这些公司受到行业准入门槛保护,因为获取和整理大量数据需要规模经济才能实现,但这些公司的商业模式并不涉及收集、分析并了解服务对象(客户)如何提供更好的产品。

长久以来,企业就开始收集客户信息并加以利用,进而提供更好的产品和服务。在信息技术不发达的情况下,客户信息收集的速度十分缓慢,范围有限,而且很难扩大规模。许多传统制造商若想采用这一策略,就要开展客户调研和焦点小组讨论,研究销售数据和分析策略,因销售数据一般与客户个人无关,调研和焦点小组花费大量时间和金钱只收集到一小部分客户的数据。

云端技术和新科技的发展使公司能够迅速处理和解读大量数据,上述状况也随之发生了巨大的变化。现在的联网产品及服务能够直接收集客户的信息,包括个人的详细资料、搜索行为、内容选择、交流方式、社交动态、行动轨迹以及生活习惯。用机器学习算法分析客户的这种数字化记录之后,公司可以根据用户情况自动调整产品和服务,更加精准地满足客户的需求,乃至实现自动定制化服务。

这样的科技发展使得数据驱动型机器学习比以往公司的客户洞察更加强大,但这同样适用于竞争对手。

若要在客户数据学习方面领先竞争对手,就必须利用客户数据驱动型学习建立防御壁垒,进而形成竞争优势,要确定这种竞争优势是否可持续,公司应当回答以下七个问题。

第一,客户数据如何为产品和服务增值?

客户数据对产品或服务增加的价值越高,建立持久竞争优势的可能性就越高。例如,先进驾驶辅助系统(advanced driver assistance system, ADAS)的一流供应商 Mobileye 为汽车提供防撞辅助系统和车道偏离预警系统。该公司主要向汽车制造商出售其辅助系统,汽车制造商会在将其实际投入应用之前进行大量的测试。辅助系统必须有失效保护,而测试数

<sup>①</sup> 格林芬,忻榕,伊普.中国零售商如何重塑顾客的购物之旅[J],哈佛商业评论,2022(8),有改动。



据对于产品改进必不可少。Mobileye 从数十个服务的汽车制造商处收集相关测试数据,将 ADAS 的准确率提升至 99.99%。

智能电视则恰恰相反,学习客户数据获得的价值较低。现在一些智能电视装有相关软件,可以根据用户的观看习惯和节目的人气提供定制化的电视节目或电影推荐。目前,消费者还不太在意这一功能。消费者购买时主要考虑的因素是电视的屏幕尺寸、画面质量、易用性和耐用性。很显然,这些客户数据无法为智能电视公司增强竞争力。

第二,数据驱动型学习的边际价值下降有多快?

换言之,对于公司而言,客户数据边际价值的临界点(边际价值为零)在哪里?

边际价值下降速度越慢,基于数据的学习带来的竞争力就越强。需要注意的是,回答这一问题时,不要根据其他应用程序的特定衡量标准(如自动应答机器人正确回答问题的比例、电影推荐获得点击的次数等)判断基于数据的学习价值,要重点关注客户的购买意愿。

如果将 Mobileye 的 ADAS 准确度和客户使用率表示为函数关系,会发现部分汽车制造商和中等水平的测试就能够达到 90% 的准确度,但要想达到 99%,就需要汽车制造商增加海量的数据样本并扩大测试范围,而要达到 99.99% 的精确度就变得更加困难了。有时为获得准确的客户信息而损失的成本可能高于测试带来的价值,当然,将行驶安全从 90% 的准确度提高到 99%(乃至 99.99%),价值依然非常高,因为这是生死攸关的事。但实际上,任何一家汽车制造商仅凭一己之力很难获得足够的数据,这也正是 Mobileye 的潜在竞争对手难以超越它的重要原因,由此保障了 Mobileye 在 ADAS 市场的主导地位,并于 2017 年被英特尔以 150 亿美元的高价收购,为其成为汽车智能驾驶领域的领导者奠定了坚实的科技基础。

如果学习客户数据带来的边际价值在已经获得大量客户的情况下依然居高不下,产品及服务就会拥有显著的竞争优势。预测罕见病专用的系统(如 RDMD 公司提供的系统)和百度、谷歌等在线搜索引擎体现出了这一点。微软投入多年时间、数十亿美元开发搜索引擎必应却无法撼动谷歌的地位。搜索引擎和疾病预测系统都需要大量的用户数据,才能持续提供可靠的结果。

用户数据的边际价值会迅速降低的一个反面例子是智能恒温器。智能恒温器只需要几天时间就能掌握用户一天各个时段的温度偏好。在这种情况下,数据驱动型学习无法提供太多的竞争优势。Nest(于 2014 年被谷歌收购)于 2011 年率先推出了第一台可以从用户行为中学习的智能恒温器,现在却面临 Ecobee、霍尼韦尔等公司推出的同类产品的激烈竞争。

第三,用户数据的贬值率如何?

如果数据很快就会过时,那么在其他条件都相同的情况下,竞争对手进入这一市场会更加容易,不必忌惮原来竞争企业积累的数据驱动型学习能力。

多年以来,Mobileye 从各汽车制造商处获得的大量数据,对于产品的当前版本依然有价值。谷歌数十年来积累的搜索引擎用户数据也一样有价值。虽然对某些术语的搜索可能会逐渐减少,用户渐渐开始搜索新事物,但拥有多年历史搜索数据对于服务今天的用户仍然有着重要的借鉴价值。这就意味着在这两个领域的用户数据贬值率较低,这也正是这两家公司基业长青的一个重要原因。





然而在休闲和社交游戏领域,用户数据驱动型学习的价值会很快降低。2009年,星佳(Zynga)推出的种田偷菜游戏《开心农场》(FarmVille)大获成功,休闲社交游戏行业由此开始迅速发展。星佳虽然以高度依赖用户数据分析进行设计而闻名,然而从一个游戏中获得的用户洞察难以用于另一个游戏中:休闲社交游戏只是风靡一时,用户偏好会迅速改变,因此公司很难依靠收集用户数据打造可持续的竞争优势。继《开心农场2》和《城市小镇》(CityVille)等作品之后,星佳无法再推出新的热门社交游戏,在2013年流失了近一半客户。它渐渐地被其他游戏厂商,如Supercell公司[代表作《部落冲突》(Clash of Clans)]、Epic Games公司[代表作《堡垒之夜》(Fortnite)]。星佳的市值在2012年达到高峰(104亿美元),随后就低迷不振,至今6年的大部分时间中都不到40亿美元。

第四,数据是否属于公司独有,无法从其他来源获取,无法轻易复制及破译?

拥有无可替代的独家客户数据,对于打造竞争优势至关重要。例如,投资一家农业合作社,为其提供农作物管理系统,使种植者可以持续监测所有作物的生长状态。这一系统依靠人工智能(Artificial Intelligence, AI)、计算机软件和独家的数据注释技术检测人眼看不到的植物生物学指标,提前判断疾病、营养不良,并分析成因,为种植者提供农作物病虫害防治策略和提高单位产量的建议。随着农作物管理系统被越来越多的种植者使用,其能够学习的性状差异、耕作环境等各种因素范围也越来越广,预测的准确性也就随之提升。相比之下,垃圾邮件过滤器提供商可以更为轻松地获取用户数据,所以有了很多类似的提供商。

科技发展可能会破坏企业借由独家数据获得的地位。其中典型的例子是语音识别软件。多年以前,用户必须自行训练语音识别软件识别自己的嗓音和语言习惯,使用次数越多,识别越准确。Nuance的语音识别解决方案Dragon主导这一市场数年之久。然而近十年来,借助非特定语者语音识别技术,可以使用公开的语音数据进行训练,不用花很多时间就能识别新用户的语音。这种发展使得许多公司都能够提供新的语音识别应用(AI电话客服、自动会议记录、虚拟助理等),对Nuance的核心市场造成的压力日益增加。

第五,基于用户数据的产品改进是否难以被模仿?

即使数据难以被复制,且能提供有价值的用户洞察,但如果借由数据驱动型学习实现的产品改进能被没有相关数据的竞争者轻易照搬,也很难打造持久的竞争优势。

克服这一难题主要受以下两个因素的影响:

其一,相应的改进是否深藏于复杂的生产过程中而难以被复制。流媒体音乐服务商潘多拉(Pandora)公司受益于此。潘多拉公司独家开展“音乐基因组计划(Music Genome Project)”,根据约450个特征为几百万首歌曲分类,为用户提供个性化的音乐电台服务。用户收听潘多拉上的音乐电台越多,为歌曲打分,潘多拉就越能更好地为用户定制曲目。竞争对手无法轻易复制这种个性化服务,因为该服务与音乐基因组计划紧密结合。与之相比,通过学习用户使用多种办公软件(如日程软件Calendly和向众人征求意见确定会议时间的软件Doodle)的情况做出的设计改进则很容易被观察和模仿,因此提供类似软件的公司很多。

其二,通过用户数据获得的洞察是否会迅速变化。变化速度越快,模仿难度就越大。例如,谷歌地图界面的许多设计特点都很容易被模仿(苹果地图和很多其他地图软件都模仿了谷歌地图),但谷歌地图的核心价值是其预测交通状况、推荐最优路线的功能,这一功能利用





了随时变化的实时用户数据,因此很难被复制。只有用户基础能与谷歌比肩的竞争对手(如苹果在美国)才能复制这一功能。苹果地图在美国正在缩小与谷歌地图的差距,但在其他地区的用户基础不足,无法与谷歌匹敌。

第六,来自一位用户的数据是否有助于为同一用户或其他用户改进产品?

在理想的状态下,这两个方面都可以实现,但两者之间的差异很重要。如果来自一位用户的数据能够使产品更加符合这位用户的需求,公司可以提供个人定制服务,提高用户的转移成本,借此留住用户。如果来自一位用户的数据能够更加符合其他用户促进产品改进实现的需求,就可能实现网络效应(也可能实现不了)。这两种改进都有助于提高市场准入门槛,但前者是增强现有用户的黏性,后者则有利于争取新用户。

例如,潘多拉是数字音乐流媒体领域出现的第一个巨头,但随后就被 Spotify 和苹果音乐赶超,而且这两款软件仍在继续成长。如前所述,潘多拉的主要卖点是根据用户喜好定制的个人音乐电台,但这一方向的数据驱动型学习仅限于为一位用户提供更好的服务,在不同用户间并不通用:一位用户的评分使潘多拉了解该用户喜欢的音乐的特征,进而为该用户提供具备这一特征的其他歌曲。与之相比,Spotify 更加重视分享与发现功能,如让用户搜索和收听其他用户的音乐电台,以此实现直接的网络效应,吸引更多的用户。潘多拉的服务至今依然仅限于美国(美国有潘多拉的忠实用户群体),而 Spotify 和苹果音乐已经走向全球。潘多拉于 2019 年 2 月以 35 亿美元的价格被 SiriusXM 收购,而 Spotify 于 2018 年 4 月上市,截至 2019 年 11 月初的市值是 260 亿美元。很明显,通过用户数据驱动型学习实现的定制化可以留住已有的用户,但无法实现网络效应带来的指数级增长。

第七,从用户数据中获得的洞察应用于产品的速度有多快?

学习周期够快,可以将竞争对手甩在身后,平均用户体验中对多项产品同时改进时尤其如此。但如果基于数据的改进需要几年时间或连续几代产品才能实现,那么竞争对手在此期间自行收集用户数据实现创新的可能性就会提升。因此,积极利用从当前用户处获取的数据,加速为当前用户实现产品改进,获得的竞争优势会比关注未来用户更大。前文提到的地图、搜索引擎、农作物管理系统等都可以被迅速升级,将当前的数据驱动型学习成果应用于实践,为当前的用户提供更好的服务。

这方面的一个反例是 LendUp、LendingPoint 等网络贷款提供商。这些公司考查用户的还款记录,以及还款记录与用户资料、行为等多个方面的关联性,以此做出更好的贷款决策。在这里,与当前借款者有关的唯一数据洞察来自以前的借款者,而这部分信息体现在当前借款者的贷款合同和利息上。借款者没有理由关心让提供商未来会从中获益的任何洞察,这是因为目前的贷款合同不会受到任何影响。因此,借款者在考虑寻求贷款时,并不关心还有多少人会向特定的提供商贷款。于是借款者会倾向于一直与原本的提供商合作,因为该提供商比其他提供商更加了解自己的需求,而争夺新客户的竞争依然十分激烈。

以上案例讲述了商业领域一个非常重要的问题,即在数字经济发展的背景下,企业如何有效地利用数据,如何判断用户数据是否能被转化成企业的竞争力,案例中给出了七个判断标准。无论数据如何改变,都是某种商业行为的数据记载,对其背后的规律需要利用一定的方法进行分析、挖掘,这就是信息分析。而连锁企业需要思考的是在数字经济下如何利用好数据。



## 任务一 连锁企业信息分析概述

### 一、信息的概念及含义

信息是常见的一个词,却又是很难定义的一个词,几乎日常生活中的方方面面都会涉及信息。例如,人们买菜时要知道市场信息(行情),孩子上学时要知道学校信息,买股票时要知道股市信息,等等。信息的概念大概历经三个阶段的变迁,即本体论层次、信息论层次和认识论层次。其中,影响较深刻、较广泛的一种说法是本体论、信息论和认识论中关于信息的定义和解释,并在此基础上形成了本体论方法、信息论方法和认识论方法三种科学方法论。

#### (一)本体论中的信息

本体即存在、实在、事实。本体可分为两个部分:一是自然本体,二是社会本体。人们对本体的认识主要是通过本体产生的各种信息,即本体信息。本体信息是对特定领域知识的共同认识,以便实现领域相关应用之间的互相操作和互相转化。人们对本体认识的准确程度取决于本体信息来源的丰富程度及对本体信息的处理能力。

在本体论层次的意义上,即在没有任何约束的条件下,可以将信息定义为事物存在的方式和运动状态的表现形式。这也是最一般意义上所说的信息概念。这里的事物泛指存在于人类社会、思维活动和自然界中一切可能的对象。存在方式是指事物的内部结构和外部联系。运动泛指一切意义上的变化,包括机械的、物理的、化学的、生物的、思维的和社会的运动。运动状态则是指事物在时间和空间上变化所展示的特征、态势和规律。由于宇宙间的一切事物都具有特定的存在方式和运动状态,因而一切事物都在不断地产生信息,而且是不同的信息。在这一层次上定义的信息是最普遍、最广义的信息,可与物质、能量并驾齐驱,称为本体论层次的信息。本体论层次的信息可等同于知识,即确定性信息。其基本特征表现为:本体信息是客观存在的事物、事件或过程;既可以是事物本身,也可以是事物的运动状态及其变化方式;作为客观存在的事物,是一种既定的事实,具有客观性、真实性、确定性和稳定性;可以与人无关而独立存在。

#### (二)信息论中的信息

如今,在通信领域所展开的对信息的研究取得了丰硕的成果。在这里,信息被看作有序程度(或组织程度)的度量和负熵,是用以减少不确定性的事物。这些也是香农、维纳、布里渊等人对信息的共同理解。这些认识比仅仅把信息看作消息或通信的内容要丰富、深刻得多。事实上,这些关于信息的理解也在不断影响着其他领域的研究。例如,有关风险、不确定性与信息之间关系的研究在20世纪40年代就受到许多经济学家的关注,不确定性的减少(信息的获取)与成本、收益被联系起来,取得了突出成果。英国生物学家阿思比把信息定义为“变异度”。他认为,任何一个集合包含的元素的数目以2为底的对数就是该集合的变异度。随着信息论方法的不断成熟和应用,信息开始作为一个通用概念在各学科领域和社会生活的各方面被广泛使用。



信息科学方法论具有相对独立和相对完整的体系结构,并认为信息科学方法论包含三种方法和两个基本法则,即信息系统分析方法(认识复杂事物的方法)、信息系统综合方法(综合复杂系统的方法)、信息系统进化方法(变革复杂系统的方法)、物质-能量-信息三位一体准则、结构功能行为辩证统一准则。

### (三)认识论中的信息

如果考虑到信息的产生、认识、获取和利用离不开人这一主体,并且必须从主体的立场定义信息,本体论层次的信息定义就转化为认识论层次的信息定义。在认识论层次上,可以将信息定义为主体所感知或表述的事物存在的方式和运动状态。在这里,主体所感知的是外部世界向主体输入的信息,主体所表述的则是主体向外部世界输出的信息。显然,在引入了主体这一约束条件后,信息定义的范围变窄了。在认识论层次上,信息主要用于辅助人的行为选择,即决策,涉及人的各种信息活动以及人对信息的理解和把握。而决策恰恰是主体间相互博弈和主体内部博弈的过程。

### (四)信息的含义

认识论信息就是主体所感知或所表述的事物运动的状态与方式。其中,认识所感知的事物运动状态与方式是外部世界向主体输入的信息,可称为感知信息;认识所表述的事物运动状态与方式是主体向外部世界(包括向其他主体)输出的信息,可称为再生信息。认识论信息受主体约束,可以说,没有主体就没有认识论信息。一般而言,在人类所及的有限时空中,本体论信息与认识论信息是可以互相转化的,其转化过程大致与人类认识和改造世界的过程相统一。在认识论层次的信息概念中若再引入认识深度这一约束条件,认识论信息就可以进一步扩展为3个层次:语法信息,即主体所感知或表述的事物运动状态和方式的形式化关系,这是最低层次的认识论信息;语义信息,即主体所感知或表述的事物运动状态和方式的逻辑含义,这是较高层次的认识论信息;语用信息,即主体所感知或表述的事物运动状态与方式相对于某种目的的效用,这是最高层次的认识论信息。一般而言,语法信息、语义信息和语用信息是密不可分的,不可能撇开其中的一个方面而孤立地研究其他方面。当认识和研究一种事物时,人们多是遵循从语法、语义到语用的认识顺序,这可以看作一个认识小循环,是人类认识和改造世界的方法论。

信息是事物的运动状态与方式,这是广义的信息概念,是最普通的层次。由此层层引入约束条件,信息的内涵就变得越来越丰富,适用范围则变得越来越窄。与此同时,信息的概念逐级展开,自然形成了信息的概念体系。

综合以上讨论,对信息的定义应该从认识论的角度出发。其主要意义如下:

(1)信息是为人所用的,至少对有感觉、有记忆能力的生命体信息才有意义。如果没有生命的存在,物体之间存在热辐射、存在吸引力等所谓的信息是没有意义的。

(2)对信息要通过人获取,尽管它是“客观存在的”,或者说信息起码应该是人所看到的那部分事物的状态,而不应该包括未被认识到的那部分事物的状态,尽管人们可以通过已知事物的状态发现未知事物的状态。例如,如果人类是色盲,信息中就不会有颜色的概念,事物的所谓“状态”中也不会有颜色的状态。

(3)信息要以一定的方式被表达出来,其表达方式是按照某种约定。这种约定实际上就是编码,不同的人群有不同的编码,不同的生物种群也有不同的编码。



(4)信息具有交流的功能,包括人与人之间、人与客观世界之间的交流。

(5)信息可以表达事物的运动状态及状态变化的方式,但其表达方式是由人决定的,并且是为了让人理解和认识的,离开了人,信息就没有任何意义。例如,“太阳向我们发信息”,这种说法人们显然不能接受;而“太阳光中含有某种信息”“太阳风暴告诉我们某种信息”等说法是可以接受的。其区别就在于有无人的认识成分。事物的运动状态及状态变化的方式的表达不过是人们对于所认识到的那部分事物的状态和运动规律所做的自以为正确的描述。这是辩证唯物主义认识论的基本原理。

(6)信息具有可挖掘特性,不同的信宿对同一条信息可能有不同的理解。例如,“尼克松要来中国访问”,这从形式上看只是某人访问一件事,但政治人物和商人可能从中发掘出不同的信息。

## 二、信息的性质

虽然从不同的角度出发,信息的含义不同,但是对于信息的基本性质,人们是有共识的。

### (一)普遍性

信息是事物状态和变化的反映。世界是物质的,物质是普遍存在的,物质都处于运动之中,而信息是源于物质及其运动的,并以物质的运动为其存在的条件。因此,物质及其运动的普遍性决定了信息存在的普遍性,信息无处不在、无时不有。

### (二)客观性

信息是对事物的状态、特征及其变化的客观反映。因为事物及其状态是不以人们的意志为转移的客观存在,所以反映这种客观存在的信息同样带有客观性,不能凭借人们的主观臆想来创造,只有客观的信息才具有储存价值。这是信息客观性的一个方面。另一个方面,信息不仅具有客观性,而且信息一旦生成,也就成为一种客观的存在,其时效性会降低,客观性却是不变的。

### (三)动态性

信息反映的总是特定时刻事物的运动状态和方式,当人们将该时刻的信息提取出来之后,事物还在不停地运动,其内容和信息量都会随着时间的变化而不断地被取舍、更新、充实。

### (四)可识别性

信息源于事物及其运动,是事物的一种普遍属性,是事物本质、特征和运动规律的反映。人类可以通过自己的感觉器官或借助各种仪器设备感知、识别事物,事实上就是通过感知、识别事物所发出的信息进而认识事物,可见信息是可以被识别的。

## 三、我国古代信息分析理论溯源

《孙子兵法》曰:“上兵伐谋,其次伐交,其次伐兵,其下攻城。”可见对战事的信息分析尤为重要。《史记·货殖列传》中记载,越王采用计然之策得以复国,而后范蠡用计然之策成为富甲天下的巨商。《计然七策》对环境尤为重视,认为“旱则资舟,水则资车,物之理也”。这也说明经商要时刻关注环境变化,从环境中获取信息,以待发展。

另外,本书为什么被命名为“信息分析”,而没有采用人们耳熟能详的“数据分析”?这是





因为信息分析所涵盖的范围要比数据分析更庞大。首先,进行的预测和决策是未来将要发生的行为或者结果,数据分析不能表达这种行为或者结果。其次,在进行信息分析时,往往直接指向研究的目的性。例如,编制营销计划,需要获取历年销售数据进行分析,而后进行营销计划的编制。信息分析能够完整地概括营销计划编制的全过程,而数据分析仅能说明进行营销计划编制前的数据处理和数据挖掘部分。最后,信息分析强调信息的获得、信息的整理、信息的处理,而数据分析仅是其中的一部分。

## 四、连锁企业信息分析的定义和作用

### (一)连锁企业信息分析的定义

连锁企业信息分析是通过对连锁企业经营数据(本体论)的判断,进行有目的的分析(认识论),从中提取有价值的信息(信息论),形成有目的、有说服力(经营决策)结论的过程。

#### 1. 连锁企业信息定义的分析

对上述定义的理解,需要把握以下三点:

(1)连锁企业经营数据是连锁企业信息分析的主体。首先,信息分析的载体来自连锁经营数据,这些数据既可以是经营过程中生成的数字数据,也可以是其他资源。其次,这些数据主体的应用可以为企业带来巨大的效益。

1998年,7-ELEVEN便利店开始使用一条综合业务数字网(Integrated Service Digital Network, ISDN),将其零售商店与总部联系起来,利用卫星、互联网将商店层面的信息跟踪情况融入供应链层面的信息系统。当顾客提着一篮产品来到收银台前时,店员将顾客的性别和(估计的)年龄输入,然后再扫描他们采购产品的条形码。这些销售数据通过 ISDN 被传递至 SEJ 的总部。同时,数据被商店计算机(Store Computer, SC)系统处理,该系统控制着商店所有的计算机设备和外设。SC 使得商店的经理和 SEJ 的总部可以同时更新和分析 POS 数据。SEJ 的信息反馈循环以两个周期运营——每日和每周。一旦收到 POS 数据,总部按照区域、产品和时间将这些数据集成,并且在次日一早就让所有商店和供应商能够使用它们。商店经理可以根据顾客群体分析所有存货单位(Stock Keeping Unit, SKU)每小时的销售趋势和缺货率。而且 SEJ 的门店系统还会每天固定 5 次收集天气的动态信息。这就是 SEJ 的信息反馈循环的每日周期。在天气方面,SEJ 非常重视天气情况对消费者影响的监测。每个周二的早上,每个现场运营顾问(Operation Field Counselor, OFC)负责 8 家门店,指导门店收集天气、温度、电视广告的有效性、特殊事件(如学校运动会和道路建设)和消费者口味的变化趋势等信息,并为地区经理准备一份每周报告,地区经理在每周二的经理会议上共享该报告,然后制定有针对性的营销措施。

(2)连锁企业信息分析具有目的性。连锁企业信息分析的目的在于了解门店经营环境的变化,为未来的经营决策提供基本依据。利用信息分析,进行选址、优化门店布局、调整品类、制定营销策略、安排销售策略等经营活动。例如,餐饮连锁企业为了优化食材的储量,开始针对日常销售的时间、销售数量进行统计,发现 11:00—14:00 和 16:00—20:00 为销售的高峰期,因此,每天 7:00—9:00 开始进行食材储备,以保证其新鲜度。另外,连锁便利店经营的鲜食产品有同样的特点,连锁便利店合理利用销售规律,既可以提高产品的销售针对性,又可以保障消费者得到最好的产品。因此,进行连锁企业信息分析,必须先明确目的。

(3)信息分析的过程是对数据优化处理的过程。日常经营数据量非常大:一个是统计指



标数量大,另一个是数据观察值大。如果是全国性连锁企业,需要对数据进行优化,目的有两个:一是使数据清晰;二是使数据能够有效地反映变化趋势。例如,某企业近三年的营业额为19 974 562.22元、20 057 869.32元、21 214 547.44元,这三个数据的单位是元,但是可读性差;如果以万元为单位(小数点后一般取两位数),上述数据就变成了1 997.46万元、2 005.79万元、2 121.45万元,营业额的变化趋势也比较明显;如果以百万元为单位,上述数据就变成了19.97百万元、20.06百万元、21.21百万元,营业额的变化趋势依然明显。这就是对数据的优化。根据日常的工作经验,门店人员会有一些的工作预期。例如,销售人员会持续关注当期和前期数据的对比(环比)及当期和去年同期数据的对比(同比),使用这两个对比都是为了说明当期销售数据的变化情况,表现的是相对变化趋势,即当期销售数据是处于上升趋势还是处于下降趋势。这就是对数据进行处理的过程。

## 2. 连锁企业信息分析的相关名词辨析

在实际应用中,为了更好地理解连锁企业信息分析的作用,此处对易混淆的概念进行说明,以便于今后在学习的过程中更深刻地理解连锁企业信息分析的概念。

(1)信息分析与信息技术。信息技术是指以电子计算机和现代通信为主要手段实现信息的获取、加工、传递和利用等功能的技术总和,是获取、存储、传递、处理和分析以及使信息标准化的技术。因此,信息技术是保证信息传递、存储的工具。而信息分析是对数据的分析、挖掘和处理。两者都是方法和应用的学科,但是侧重点不同。

(2)信息分析与信息系统。信息系统是由计算机硬件、网络和通信设备、计算机软件、信息资源、信息用户和规章制度组成的以处理信息流为目的的人机一体化系统。按发展和特点,信息系统可分为数据处理系统、管理信息系统、决策支持系统、专家系统和虚拟办公室五种。在连锁企业,信息系统对商业流通的实际工作和生活场景进行信息化处理,信息分析正是利用信息系统的原生数据进行加工和处理的过程。

## (二)连锁企业信息分析的作用

### 1. 为连锁企业的决策提供支持

随着经济的不断发展,为保证连锁企业在面临互联网的冲击时又好又快地发展,必须重视连锁企业信息化建设,尤其是要做好连锁企业信息分析工作,通过对经营数据的挖掘和分析更好地为消费者服务。另外,应通过信息分析打破连锁企业内部和外部数据衔接的瓶颈,利用信息技术建设工作常模,减少数据重复统计、重复计算的工作量。应采用连锁企业信息分析的方法和结论为企业更好地发展提供决策依据。

### 2. 为连锁企业的研究和开发提供支持

连锁企业信息分析可为产品开发提供思路。连锁企业信息分析为连锁企业提供了历史概况、当前水平和研究动向,提供了产品研发的可能性、存在的问题及解决办法。通过对连锁企业信息分析的流程化和标准化,又为连锁企业的开发工作奠定了基础,保证连锁企业在行业中的竞争力和影响力。

### 3. 为连锁企业开拓市场提供支持

连锁企业门店开拓必须遵循四个“适当”原则,即适当的地点、适当的时间、适当的价格和适当的产品。要做到四个“适当”,必须有充分的市场信息作为保障。市场信息通常包括两个方面:一是连锁企业市场内部产生的与经济活动有关的信息,如市场供需、价格水平、消费者偏好等;二是连锁企业市场外部信息,包括政治、经济、文化等方面。信息分析就是收集





这两个方面的信息,寻找、识别和把握市场机会,选准市场开拓的突破口,规避潜在的市场风险,提升市场开拓的成功率。

## 五、连锁企业信息分析的主体——数据

连锁企业信息分析主要是分析来自门店运营、采购供应、存储配送、销售数据。从产品的运动轨迹看,信息产生于产品原料供应、生产、存储、运输、销售等基本环节,相邻两个环节的数据背后反映着需求和供给之间的关系。连锁企业信息分析就是通过分析各环节之间的数据反映企业经营状况的变化。

### (一)供应链中的数据

供应链是指围绕核心企业,从配套零件开始,制成中间产品以及最终产品,最后由销售网络把产品送到消费者手中的,将供应商、制造商、分销商直到最终用户连成一个整体的功能网链结构。供应链管理的经营理念是从消费者的角度出发,通过企业间的协作,谋求供应链整体的最佳化。成功的供应链管理能够协调并整合供应链中所有的活动,最终使其成为无缝衔接的一体化过程。

### (二)连锁企业门店的运营数据

顾客在与连锁企业门店接触的过程中会通过自己的观察、体验、感受评价门店的服务水平,良好的购买体验成为吸引顾客重要的因素。从顾客走近门店到顾客购买后的体验评价,会产生购买前、购买中、购买后三种类别的数据。

购买前,在顾客走近门店时,最先引起顾客注意的是门店的选址和门店的装修。

购买中,对卖场布局、装潢设计、商品陈列、氛围营造、目标产品检索、结账服务等因素的良好体验会使顾客产生购买的兴趣。

购买后,满意的价格、被满足的购买期望、购买后的良好感受形成顾客的满意和忠诚。

根据上述因素可以构建顾客购房图,如图 1-1 所示。

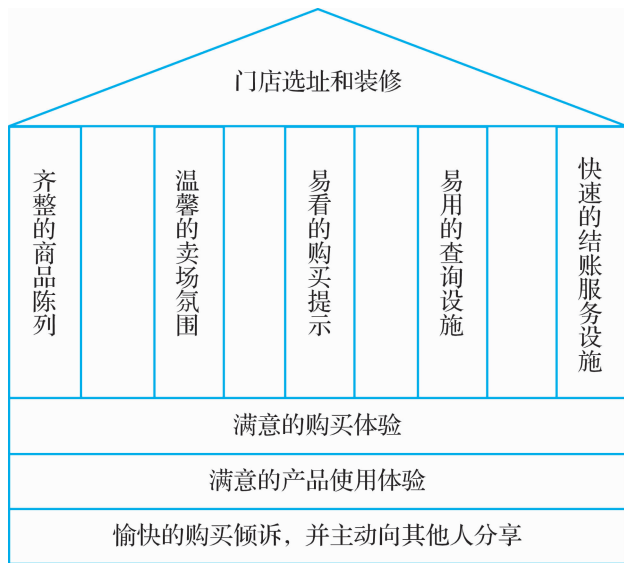


图 1-1 顾客购房图



### 门店是如何吸引顾客的<sup>①</sup>

门店开发顾客,首先要考虑如何才能将顾客吸引到店中。

一位电动车专卖店老板说:“只要把消费者带到店里,如果有 10 个人,我可以卖掉 2 辆车。”销售中有一个平均法则:门店的销售额与门店的顾客数量成正比关系。顾客越多,销售额就越高。如何将消费者吸引到专卖店中呢?这家专卖店的老板总结了引客的 16 个招数,具体包括电动车秀、免费试驾、巡回展销、电动车赛、景点租车、第一报道、赞助新闻、软性文章、横幅大战、团体介绍、网上宣传、安全讲座、以租代销、用户服务、以旧换新、来店就送。

具体来讲,门店要想吸引顾客,要从以下三个方面深入做文章。

#### 一、用闪亮的店面形象吸引顾客

对化妆品营销而言,店面形象就是产品价值。在信息化高速发展的今天,若搞不好店面建设,做不好店内设计,就注定在经营中会失败。

许多门店存在的问题主要表现为以下三个方面。

##### 1. 缺乏醒目的标志

例如,有一个大大的黄色的字母 M,顾客从很远处就可以知道这里有一家麦当劳快餐店;有一个旋转的三色灯箱,顾客知道这里有一家理发店。醒目的招牌是外在吸引顾客的第一要素,随处可见的快餐店、便利店都有明确的店招设计;甚至一些手机配饰经营店都用上了华为、苹果的 logo 说明自己的营业内容。

##### 2. 灯光昏暗

顾客就像飞蛾,具有趋光性,愿意进入明亮干净的地方。例如,化妆品店吸引的是女性顾客,明亮的灯光、清晰的化妆镜、整齐的产品陈列、整洁卫生的店面更容易吸引顾客走进门店;如果灯光不明亮,门店整体的氛围就会阴沉,顾客的购买兴致就不会强烈,营业效果自然就会差。

##### 3. 陈列杂乱

商品陈列时,如果没有进行系统的设计,会导致陈列没有章法;另外,对有些商品无法进行有效的摆放,几乎全部采用堆放的形式。在面积较小的门店内,不要采用堆放陈列,因为这不仅会导致门店内混乱,也会影响顾客动线。

#### 二、用商品吸引顾客

例如,服装店的经营者会把橱窗布置得干净、明亮,把本季度最新款的服装放到橱窗中,告知路过的顾客“我们店有最新款的服装,快进来看看”;餐饮店的经营者会把最早进来就餐的顾客安排在靠近窗户的座位上,暗示路过的顾客“我们店的饭菜好吃,你看有这么多的人在店中就餐,快来尝尝呀”。

门店经营者也可以将新产品、名牌产品、购买频率高的产品陈列在门口的显眼处,并以略低的价格进行销售,以吸引对新产品、名牌产品和购买频率高的产品感兴趣的顾客到店中逛一逛。

<sup>①</sup> 佚名. 如何吸引路过的顾客进店? 有 4 大绝招! [EB/OL]. (2019-07-15)[2022-04-07]. [https://www.sohu.com/a/327040061\\_167454](https://www.sohu.com/a/327040061_167454). 有改动.



### 三、用促销吸引顾客

#### 1. 渲染促销活动信息

可在门店入口处通过卖点广告(POP)、宣传单、条幅等方式介绍店内的促销活动。有调查显示,顾客受到化妆品店面或美容院店面宣传气氛的感染而进入其内的概率为40%。

门店利用海报、POP、旗子、影灯、气球、花束等造势引人注目,效果显著。

#### 2. 开展促销活动吸引顾客

某化妆品专卖店开展“空瓶换购”活动,选取了几款夏季产品(如防晒霜、洗面奶、洗发水、沐浴露等),以震撼价冲破消费者的心理底线。对每款产品象征性地收一至两元钱甚至是一至两分钱,并且承诺使用同类产品的空瓶(洗面奶的空瓶只能用于换洗面奶,以此类推)加对应的钱数就可以买走所需要的产品。例如,对于18元的洗面奶,只要顾客持该洗面奶空瓶,用一分钱就可以将其买走;并且规定对其他品牌的产品可以加几分钱兑换该店的同类产品,目的是把原先消费其他品牌的顾客争取过来,将其变成自己的顾客;注明换购产品的数量,并限定每人换购的数量。该促销活动一经推出,即受到顾客的热烈欢迎,一时顾客盈门。在顾客进店后,可通过后续活动,增加顾客的购买数量,吸引顾客成为会员,最终取得盈利。

### (三) 零售数据

零售数据分析就是从一大堆看似杂乱无章的数据中提取关键经营数据,用来说明目前的经营现状。但是具体而言,从产生的条件和基础来看,零售数据可以总结为三个字,即人、货、场。

#### 1. 人

通常而言,零售数据中的人包括顾客和销售人员两类。对顾客要分析购买原因、购买特征、购后评价等主观购买因素,还要分析购买时间、客单价、购买产品数量等客观行为。对销售人员要了解其劳动效率、劳动贡献、创新能力、努力程度等。

#### 2. 货

货指的是零售中的商品,对商品进行管理是所有零售人员必备的基本技能,商品分析的主要数据来自商品基础数据和销售数据,从而产生以分析结构为主线的分析思路。商品的主要分析数据包括商品结构的广度、宽度、深度,市场覆盖度,类别结构,关联度及库存结构,等等。通过对商品数据的准确把握,可做到动态调整产品结构,增强所经营商品的竞争力并对商品进行合理的配置。

#### 3. 场

这里的场是场所,就是零售的卖场、门店。选址就是要收集商圈促销、商场促销、品牌竞争、顾客在商圈内的行为特征等商业数据,这对企业零售事业发展至关重要。

## 六、连锁企业信息分析的基本流程

连锁企业信息分析遵循一般信息分析工作规律,大体遵循分析目标的明确、信息的收集和整理、信息的分析和提炼、信息产品的形成、信息产品的利用和反馈的思路,如图1-2所示。

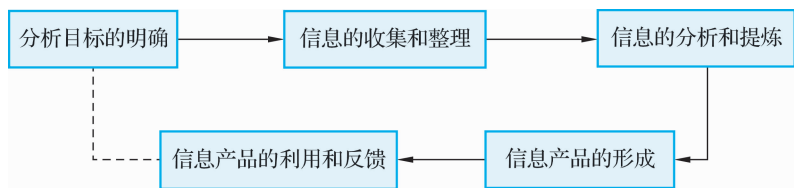


图 1-2 连锁企业信息分析的基本流程

### （一）分析目标的明确

进行连锁企业信息分析之前,要明确了解信息分析的目的是什么,否则将南辕北辙。例如,要制定明年的销售目标,整理当年的销售数据,要做的是把当年的销售额整理出来,然后对比近四年的销售情况,用以说明本年度销售额的变化趋势。另外,通过对销售数据的分月统计分析各月的销售变化情况,说明有无月份的销售额特别不好(或者特别好),并分析原因。

### （二）信息的收集和整理

连锁企业往往通过市场、系统(一般是会员系统)、第三方信息源(报纸、杂志、会议、咨询机构)等渠道收集信息,并对采集的数据进行整理,剔除信息中的噪点(无效信息),如无效的问卷等。一般采用因素分类法、矩阵分析法等将数据按照一定的逻辑关系进行分类,并将整理好的信息文件进行存储,以待进一步分析。

### （三）信息的分析和提炼

再次核对信息,查找信息中的三类无用数据,即残缺数据、错误数据和重复数据。残缺数据是信息中表达不清楚的数据,如缺少日期、姓名、签字等主体关键信息的数据,按照缺少的情况再次寻找补全数据的措施,若依然不能补全,则废弃数据。错误数据往往是表达格式不符合要求或者信息表达有明确错误的信息,可以直接废弃。重复数据是查找到的重复出现的数据,应征询重复的原因,然后进行相关的处理。

应根据分析的目的对关键信息进行提炼。例如,企业分析销售数据,需要将企业的销售额、销售量、利润额等数据进行提取,然后通过搭建分析框架,借助相关理论进行系统分析,得出相关结果。有时结果表现为结论,如近五年销售额增长;有时结果表现为具体数值,如2023年的预测销售额为2 300万元等;有时结果表现为复合结论,即从多个角度得出结论,供企业参考。但是,结果一定是明确表达的内容,不能有推测的内容和模糊的词语出现。

### （四）信息产品的形成

信息产品是信息分析的产物,是对信息分析结果全面的总结。信息产品存在质量的差别。按形式,信息产品可分为消息类产品 and 数据类产品,基本结构都可以分为前言(或者导语)、主体、结尾三个部分。不同的是,对消息类产品要说明背景,对数据类产品要说明数据的出处。无论是哪一类信息产品,描述必须清楚、准确,说明必须清晰、明了,使观者能快速掌握信息产品所表达的主要思想和核心内容。

### （五）信息产品的利用和反馈

连锁企业信息产品形成后,可为连锁企业发展提供指导和支持。因此,在连锁企业信息产品使用的实践中,要把信息分析的结论转化为用户语言,其应符合用户已有的知识结构、



经验、情报意识等习惯,发挥信息产品的最大价值,创造经济效益和社会效益。同时,相关人员在使用信息产品时形成对产品的信息反馈,连锁企业必须重视信息反馈,及时修正信息产品,以保障信息产品的质量。

## 任务二 连锁企业信息分析的基本思路

在工作中,结合管理学、市场营销学、社会学等,连锁企业信息分析形成了一些比较成熟的模式,通常用5W2H分析法、用户行为分析法、4E分析法和逻辑树分析法厘清连锁企业信息分析的基本思路。

### 一、5W2H分析法

#### (一)5W2H分析法的起源

##### 1. 方法的提出

5W2H分析法又称七何分析法,是第二次世界大战中美国陆军兵器修理部首创的。该方法简单、方便,易于理解、使用,富有启发意义,广泛应用于企业管理和技术活动,尤其对于决策和执行性的活动、措施非常有用,同时有助于弥补考虑问题的疏漏。5W2H分析法是用五个以W开头的英语单词和两个以H开头的英语单词对研究的问题进行设问,发现解决问题的线索,寻找解决思路,进行设计构思,从而解决实际工作中的问题,具体如下:

- (1)What—何事? 目的是什么? 做什么工作?
- (2)Why—为何? 为什么要这么做? 理由何在? 造成这样的结果是因为什么?
- (3)Who—何人? 由谁承担? 由谁完成? 由谁负责?
- (4)When—何时? 什么时间完成? 什么时机最适宜?
- (5)Where—何处? 在哪里做? 从哪里入手?
- (6)How—何法? 如何提高效率? 如何实施? 方法怎样?
- (7)How much—何样? 做到什么程度? 数量如何? 质量水平如何? 费用、产出如何?

如果现行的做法或产品经过上述七个问题的审核已无懈可击,便可认为这一做法或产品可取。如果七个问题中有一个答复不能令人满意,则表示在这方面有改进的余地。如果某个方面的答复有独创的优点,则可以扩大这一方面的效用。

##### 2. 方法的具体应用

对于在店铺内长期从事商品销售的人员而言,必须灵活使用上述七个方面分析顾客的购买行为。例如,某顾客到门店购物,第一,要明确顾客买什么(What),可以通过简单的提问明确顾客要购买什么商品、门店能够为顾客提供什么样的商品。第二,要分析顾客购买的原因(Why),了解顾客购买的出发点和目的。第三,要了解顾客,即何人购买(Who),了解顾客是否为商品使用者。第四,要了解消费者为什么在这一时间点(When)来购买,以便于把握推销时机,提高产品的成交率,另外,了解顾客的购买时机,还有利于采用适当的销售方法,以促成交易。第五,要了解消费者为什么到本店(Where)购买、从何处获得本店的信息,提升顾客对门店的文化认识,引导顾客了解产品经营的历史和经营优势。第六,要了解顾客如何(How)购买,即了解顾客的结算方式,从而为顾客提供尽可能多的支付方案,并为顾客





有针对性地制定售后服务方案。第七,要了解顾客的购买价格(How much),掌握顾客的购买承受力,提升商品的成交率。针对顾客购买的 5W2H 分析法具体如图 1-3 所示。

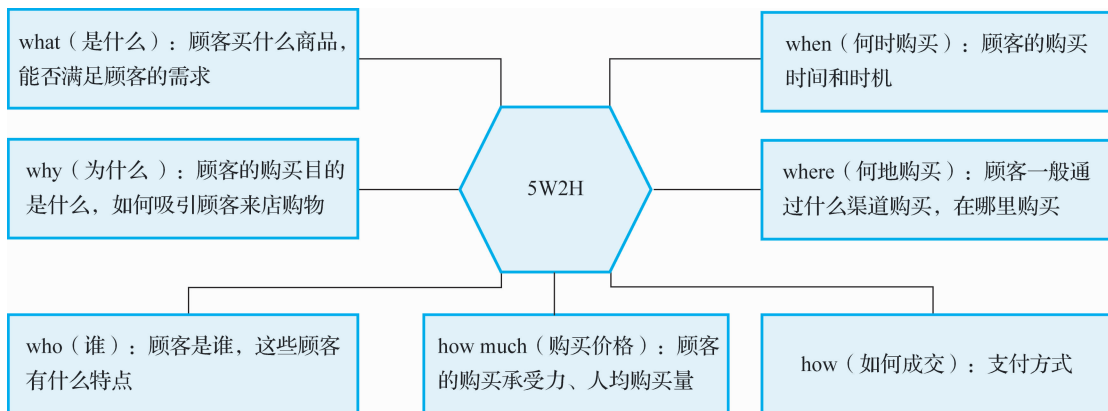


图 1-3 针对顾客购买的 5W2H 分析法

### (二)5W2H 分析法的进阶提问

在实际应用 5W2H 分析法时,针对每个问题并不能一步就能找到准确的原因。例如,顾客为什么要购买产品?往往顾客有着更私密的购买目的,需要调整问题,逐步诱导,追寻真因,这就需要对 5W2H 分析法进行改进。这里介绍一种对 5W2H 分析法的进阶提问方式,以明确 5W2H 分析法的具体应用。

提出疑问对发现问题和解决问题是极其重要的。创造力强的人都具有善于提出问题的能力,因为提出一个好的问题,就意味着问题被解决了一半。提问的技巧好,可以发挥人的想象力;反之,会挫伤想象力。在发明、设计中,对问题不敏感与平时不善于提问有密切的关系。对一个问题刨根问底,有可能发现新的问题。从根本上来看,学会发明、设计首先要学会提问,善于提问。

在实际工作中,要针对某一问题不断地提问,通过提问训练发现问题和解决问题的能力。提问能力进阶训练具体如表 1-1 所示。

表 1-1 提问能力进阶训练

5W2H	一 层 次	二 层 次	三 层 次	四 层 次	结 论
What	何事	为什么做这件事情	有更合适的事情吗	为什么是更合适的事情	定事
Why	何因	为什么是这一理由	有更合适的理由吗	为什么是更合适的理由	定因
Who	何人	为什么是他	有更合适的人吗	为什么是更合适的人	定人
When	何时	为什么在这一时间	有更合适的时间吗	为什么是更合适的时间	定时
Where	何地	为什么在这一地点	有更合适的地点吗	为什么是更合适的地点	定位
How	何法	为什么采用这一方法	有更合适的方法吗	为什么是更合适的方法	定法
How much	何样	为什么需要这些花费	有更合理的花费吗	为什么是更合理的花费	定费





## 知识链接

### 大野耐一的质量提问法

有一次,企业家大野耐一见到生产线上的机器总是停转,虽然修过多次仍不见好转,便上前询问现场的工作人员。

(1-Why)问:“为什么机器停了?”答:“因为超过了负荷,保险丝就断了。”

(2-Why)问:“为什么超负荷?”答:“因为轴承的润滑不够。”

(3-Why)问:“为什么润滑不够?”答:“因为润滑泵吸不上油来。”

(4-Why)问:“为什么吸不上油来?”答:“因为油泵轴磨损、松动了。”

(5-Why)问:“为什么磨损、松动了?”答:“因为没有安装过滤器,混入了铁屑等杂质。”

经过连续五次不停地问“为什么”,找到问题出现的真正原因(润滑油中混入了杂质)和真正的解决方案(在油泵轴上安装过滤器)。由现象推其本质,找到永久性解决问题的方案,即 5-Why。

### (三)5W2H 分析法的具体步骤

5W2H 分析法主要包括问题说明和描述、问题追因、问题分析和执行验证四个方面,具体如图 1-4 所示。

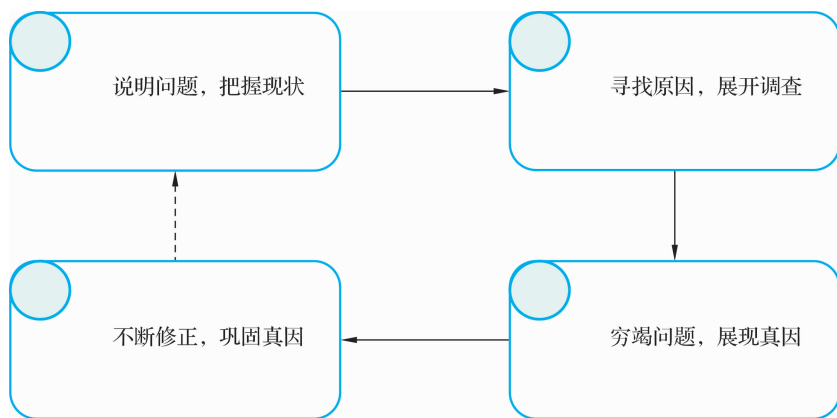


图 1-4 5W2H 法的具体步骤

#### 1. 说明问题, 把握现状

(1) 总体描述问题。应说明目前发生了什么,掌握现场的状况,对现场描述越清晰,进阶提问就越准确,问题解决的方向就越容易被捕捉。问题描述清楚,就能够引导信息使用者走近现场,清楚地知道发生了什么。例如,某个超市出现了床上用品区域的商品陈列不整齐的情况,门店督导向店长反馈巡店意见时说“发现超市内的商品摆放不整齐”。这样的描述无法把门店店长带入具体的工作场景,因为反馈的信息中缺少具体商品的信息、卖场位置的信息,不利于判断哪里出现问题,反而更容易让店长误会是自己的管理出现了问题,门店督导故意刁难店长。所以,应准确地表述问题。例如,“发现超市的第一排和第三排的枕头、被褥



陈列展示不够整齐”。

(2)描述工作标准。在描述问题时,还要考虑是否有工作标准、工作说明等规范性文件的解释说明,用以表示工作质量或工作成果展示的具体要求。这样,通过现场观察对比,才能更清晰地说明事物的现状及应该怎么样。继续分析上例,店长若追问为何商品摆放不整齐,门店督导若没有相关的工作标准作为参考,就无法给予店长有效的回答。

(3)描述工作环境。在对问题进行描述时,要将事发的时间和地点进行准确的说明,说明事发时周围有何人,事情如何发生,周围环境中还有哪些地方发生变化。对问题进行全面的描述,才能让信息使用者更全面地掌握现场的状况。

(4)分析问题,聚焦要点。在听取问题描述后,信息使用者如果能够到达现场,应到现场了解事情的发展和经过,解决问题要秉持“三现”(现场、现物、现实)原则。首先,应把现场看作问题发生的根源、管理水平提升的基石。与其费尽心思追求制度的完美,不如踏踏实实搞好现场,才有“水到渠成”的效果。其次,应对发生问题的对象(现物)进行确认,对哪里出现了问题、造成什么影响都应加以明确。“涓涓细流,汇聚成河”,不注意小事、细节,永远难以找出问题的根源。最后,依据事实和数据(现实)找出问题的根源。

(5)分析频率,说明情况。问题描述还需说明问题出现的频率、频次、范围、影响,用以判断问题的状态,并加以处理。

## 2. 寻找原因,展开调查

首先,识别和确认异常事件发生的直接原因。如果原因明显,验证它;如果原因不明显,考虑潜在的原因和检查类似的事件,以事实为基础确认直接的原因。

其次,建立一个通向根本原因的原因/效果关系链,避免分析问题出现只见树木、不见森林的片面思维方式。

## 3. 穷竭问题,展现真因

应验证采取的措施是否已经阻止问题出现,若是,则找到了真因;若不是,则需要进一步进行处理,可以对目前分析的原因采取针对性的措施,这称为临时性措施。临时性措施实施后,若问题不再出现,则分析到此为止,问题的成因即真因;若问题还是出现,说明问题的真因还没有出现,还需要重复第二步的工作。

## 4. 不断修正,巩固真因

这一步是建立在第三步的基础上的。得出真因后,还需加强对实施措施的巩固和完善,确保有效地解决问题,并反馈问题描述,通过问题描述建立分析判断模型,提升问题解决的有效性。

《麦肯锡传奇》一书中有这么一段描述:“企业倒闭最常见的原因不是因为对正确的问题提出了错误的答案,而是因为对错误的问题提出了正确的答案。我见过太多的企业一次次做出看似最佳却是建立在错误假设之上的决策,结果一点一点地把自己逼进了死路……麦肯锡要帮助客户免遭倒闭的厄运,就必须找准问题……”可见找准问题点是避免决策失误的不二法门。

此外,还需要提问者能准确地把握对方的回答是否为真正的原因,否则很容易被对方的错误回答带入歧途。



## 二、用户行为分析法

### (一)行为事件分析

行为事件分析用来研究某行为事件的发生对企业组织价值的影响以及影响程度。企业借此来追踪或记录的用户行为或业务过程,如用户注册、浏览产品详情页、成功投资、提现等,通过研究与事件发生关联的所有因素挖掘用户行为事件背后的原因、交互影响等。

在日常工作中,运营、市场、产品、数据分析师根据实际工作情况关注不同的事件指标,如最近三个月哪个渠道的用户注册量最高以及变化趋势如何、各时段的人均充值金额、上周产生过购买行为的独立用户数、用户年龄段的分布情况等,行为事件分析起到重要的作用。

行为事件分析具有强大的筛选、分组和聚合能力,逻辑清晰且使用简单,已被广泛应用。简单地说,行为事件分析一般经过事件定义与选择、下钻分析、解释与结论等环节。

### (二)用户留存分析

用户留存分析是一种用来分析用户参与情况/活跃程度的分析模型,考察进行初始行为的用户中有多少人会进行后续的行为。这是用来衡量产品对用户的价值的重要方法。留存分析可有助于回答以下问题:一位新客户在未来的一段时间内是否完成了自己期许完成的行为,如支付订单等;某件社交产品改进了新注册用户的引导流程,期待改善用户注册后的参与程度,如何验证;想判断某件产品的改动是否奏效,如新增了一个邀请好友的功能,观察是否有人因新增功能而多使用该产品几个月。

### (三)用户行为路径分析

用户行为路径是用户在 App 或网站中的访问行为路径。为了衡量网站优化的效果或营销推广的效果,以及了解用户的行为偏好,时常要对访问路径的转换数据进行分析。

以电商为例,买家从登录网站或 App 到支付成功要经过首页浏览、搜索商品、加入购物车、提交订单、支付订单等过程。而用户真实的选购过程是一个交缠反复的过程。例如,提交订单后,用户可能会返回首页继续搜索商品,也可能取消订单,每一个路径背后都有不同的动机。与其他分析模型配合进行深入分析后,能快速地找到用户动机,从而引领用户走向最优路径或者期望中的路径。门店用户行为路径分析已成为设计顾客动线的有效手段。

### (四)用户漏斗分析

用户漏斗分析是流程式数据分析。漏斗分析模型是能够科学地反映用户的行为、状态以及从起点到终点各阶段用户转化率情况的重要分析模型。

用户漏斗分析模型已经广泛应用于流量监控、产品目标转化等日常数据运营与数据分析的工作中。例如,在一个直播服务平台上,用户一般的购物路径为激活 App、注册账号、进入直播间、产生互动行为、购物消费五大阶段。用户漏斗分析模型能够展现出各个阶段的转化率,通过各环节相关数据的比较,能够直观地发现和说明问题所在,从而找到优化的方向。

### (五)用户分群数据分析

用户分群数据分析模型是进行用户画像的关键数据分析模型,是企业进行数据分析、精



细化运营的第一步。用户分群即用户信息标签化,通过用户的历史行为路径、行为特征、偏好等属性将具有相同属性的用户划分为一个群体,并进行后续的分析。可以简单地将用户分为普通用户群和预测用户群。普通分群是依据用户的属性特征和行为特征将用户群体进行分类;预测分群是根据用户以往的行为属性特征,运用机器学习算法预测他们将来会做出某些行为的概率。

### 三、4E 分析法

#### (一)4E 的由来

麦肯锡将市场营销的主要策略分为产品(Product)、价格(Price)、促销(Promotion)、渠道(Place)四个方面,取首字母简称为 4P。连锁零售企业在此基础上根据运营和销售特点总结了零售领域的 4E,即以人为核心,以数字化为基础,实现高度的场景化(Environment)、体验化(Experience),极致的效率化(Efficient),成本效益最优化(Earning),而这一切将依托于运用科技和数据连接消费者、场景、商品实现。

#### (二)4E 的具体分析

##### 1. 场景

连锁门店应该把创造更多元丰富的“场景”作为第一着眼点,让手机成为人类感知的延伸,将用户眼前呈现的场景和内心感受深度匹配、融合,从而实现转化效率的最大化。以往固定零售渠道或电商网络的场景形式单一,不能极大地调动消费者的各种感官,白白错失与用户有效互动和触达的时机。在新零售时代中,利用移动数据所构建的用户画像更加饱满、完整,集合地域标签和时间维度在内的场景数据将渠道和消费者进一步融合,场景成了距离消费者最近的渠道,消费者在场景中最终实现随心随性地购物。

##### 2. 体验

体验将成为最终购买的决定因素。看似差别不大的购物流程,但是体验相差很多。距离远近、颜色明暗、操作难易等都能成为消费者接受或拒绝的理由。新零售时代中,头像、语音识别技术将被广泛采用,利用大数据技术对用户消费数据进行精细化的运营和管理,最终可实现为用户提供极致的智能化体验,从而促进购买、提升转化效果。

##### 3. 效率

在未来,从零售终端到终端中售卖的商品,到购买商品的人群,三者的容量都将发生巨大的增长,营销的边界也将变得更为广阔。围绕人、货、场三者之间运转和匹配效率的最优化将成为新零售营销的另一个重点。

##### 4. 效益

效益就是共赢,不同于以往零售时代追求企业单方面或企业和用户两者的狭义共赢,新零售的最终目的是达成企业、消费者、供应商等商业生态链上所有参与者的利益最大化。应通过 4E 分析,更多地获取顾客购买行为,由此形成以顾客为中心的信息收集路径。

#### (三)4E 提升传统零售效率

从传统零售到新零售,消费者的行为发生巨大的转变,可以用“三达”概括,即到达、送



达、触达。传统零售要比拼产品质量,即所谓“酒香不怕巷子深”,而后竞争加剧,零售更多地强调便利性,于是更多地开店,方便消费者到达,甚至将产品送到消费者的家门口,即实现送达。无论是到达还是送达,仍然没有脱离空间的限制。而数字技术与零售结合后,消费者不再受实体店的限制,只要打开手机,在购物网站上浏览,对所有的商品尽可获取,这就形成了网络和门店之争。如何让顾客进入门店,形成第一时间的触达,是现代连锁零售门店急需破解的难题。因此,在网络上优化商品场景,利用虚拟现实、增强现实等手段将产品的功能属性和链接属性有效地融合在一起,形成新的场景体验,通过更深层次地挖掘产品的卖点,并将其转化为消费者喜闻乐见的表述方式,使消费者身临其境,达到高效的人(消费者)一物(商品)合一,这才是4E分析法的精髓。

#### 四、逻辑树分析法

逻辑树又称为问题树、演绎树或分解树等,是麦肯锡咨询公司提出的分析问题、解决问题的重要方法。首先它的形态像一棵树,把已知的问题比作树干,即考虑哪些问题或任务与已知的问题有关,将这些问题或任务比作逻辑树的树枝,一个大的树枝上还可以继续延伸出更小的树枝,即逐步列出所有与已知问题相关的问题。逻辑树分析法由麦肯锡提出,并且麦肯锡在此基础上不断丰富了分析问题的方法,形成了新麦肯锡五步法,如图1-5所示。

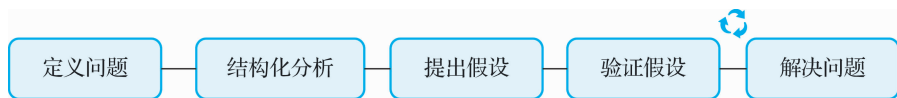


图 1-5 新麦肯锡五步法

##### (一) 定义问题

定义问题是解决问题的根本出发点,遇到困难首先要思辨问题是什么。

##### 【示例:敲钉子的故事】

假设一位客户想外包一个简单的任务——将一根钉子敲入木板中。客户找来了D公司和M公司,想了解它们各自的解决方案。

D公司是行业中有名的专业咨询公司,有很多敲钉子的专家。在仔细观察了钉子和木板之后,D公司的销售代表直奔主题,不无夸张地说:“D公司不光能做这个项目,简直就是为这类项目而生的!”

D公司的销售代表从“为什么是我们”开始并以此为中心展开陈述:“D公司采取世界领先的××技术敲钉子,拥有世界最大且材质最坚硬的锤子。工作人员的手臂最有力,锤子按照符合力学原理的完美运动轨迹落下,保证多快好省地一锤定音(敲钉子)。”

在展示的最后,D公司的销售代表提供了以前100次成功敲钉子的案例,然后低声说:“老客户有折扣。”

客户听得频频点头:D公司的方案有历史数据支持,能够直接解决问题,而且经济实惠。

M公司是一家战略咨询公司,并没有派出销售代表,只派出了一位年轻的合伙人。M公司的合伙人连钉子都没看,开口就问了一个问题:“我们为什么要敲这个钉子?”语惊四座,毕竟在“甲方为王”的时代谁敢对客户的想法产生疑问。

客户回答:“我们要把木板A和木板B连接起来。”





M公司的合伙人追问：“为什么要连接起来？”

客户回答：“我们要做一把椅子。”

合伙人接着发问：“为什么要做一把椅子？”

客户有点不耐烦，回答：“我们要用椅子在新建的会议室里接待客人。”

询问至此，M公司的合伙人停止发问，陷入片刻的思考。经过三个连续“为什么”，他已经更多地了解客户敲钉子的真正原因和诉求。客户敲钉子要解决的问题并不是连接木板，也不只是做椅子，而是要解决“在会议室待客使用什么硬件”的问题。M公司的合伙人与客户接下来的对话都围绕用户群体画像和主要诉求、刚需进行。有哪些因素能使客人在会议室感觉舒适？是文化、舒适度、仪式感或便捷等显性诉求，还是更为隐晦的奢华感等心理满足？客人对会议室硬件的要求是什么？

初步与客户讨论需求之后，M公司的合伙人会站在全球的视野看待问题，与客户分享自己涉猎的具有国际视野的相关方案。会议室的坐具除了椅子之外，还有很多可选方案，如日本的榻榻米、韩国的坐垫、爱斯基摩人铺在地上的兽皮、欧美复古家具或现代简约沙发等，当然，这肯定离不开富含中国元素的庄重典雅的木制座椅，还有带有浓重地方特色的解决方案，如东北火炕等。如果时间允许，M公司的合伙人还会谈谈未来的趋势。

上例中，通过与客户的逐步研讨，最终确定了客户的真正需求，因此有时遇到的问题绝非表面上那么简单，而是要通过深入分析了解表象后面的真正需求。这就是定义问题的真正含义。

日常工作中，界定问题定义是否正确可以采用以下三个标准：

- (1) 正确地进行问题界定后，边界清晰，没有其他问题交叉干扰。
- (2) 正确地解决问题后，相关的所有问题也会得到完美的解决而没有后遗症。定义了正确的问题后，着手解决这一问题时就会发现问题表象发生了根本性的改变。
- (3) 正确地进行问题界定后，验证标准和衡量标准随之被确定。

## (二) 结构化分析

应利用 MECE 原则进行问题结构化分析。MECE 即相互穷尽 (Mutually Exclusive)、完全穷尽 (Collectively Exhaustive)。进行问题分析时，要考虑问题所包含的维度，然后保证各维度之间相互独立、不相交。这类似于概率上所描述的完备事件，所有子事件不相交，所有子事件合并后构成事件总体。因此，利用 MECE 原则，就像对蛋糕进行切割。

### 【示例：波特五力分析模型的 MECE 应用】

假如你打算开设一家大郎烧饼店，可以进行头脑风暴，列举每一个相关的外部实体，形成如下的结果：

- (1) 同品类竞争者：其他卖烧饼的店铺，如太郎烧饼店、老狼烧饼店、直送网店等。
- (2) 上游供应商：卖面粉商、卖炭商等。
- (3) 下游消费者：顾客。
- (4) 替代品：街上的其他饭店、卖面条的店铺、卖西式糕点的店铺、卖包子的店铺等。
- (5) 潜在进入者：正在购买烧饼制作设备的附近的店铺等。

总结出同品类竞争者、上游供应商、下游消费者、替代品和潜在进入者后，可形成波特五力分析模型的经典图谱，如图 1-6 所示。



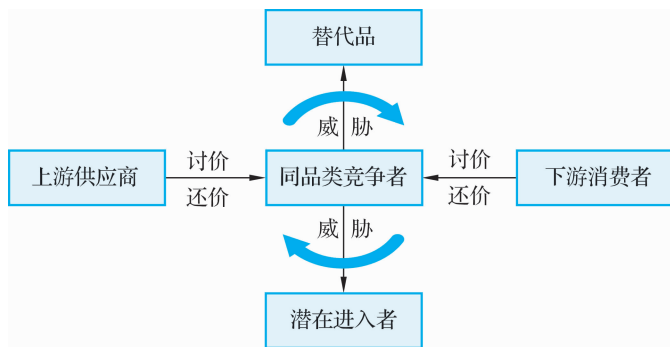


图 1-6 波特五力分析模型

假设得到了真实的数据,可以推导出在中国类似大郎烧饼店的线下早餐餐饮业的综合吸引力较弱,不建议投资。结论是基于议价能力的分析:上游供应商的议价能力很强,原材料都是食品类通货,导致商家的议价空间十分有限;下游消费者由于选择很多,在缺乏品牌力的情况下,商家处于相对弱势,消费者具有较强的议价能力;由于科技含量低,行业的进入门槛低,潜在进入者容易进入;替代品更是琳琅满目,中西糕点和其他早餐品类等都是合格的替代品;本行业的早餐同品类竞争者也相对强势,如果大郎烧饼没有特色,消费者缺乏品牌认知,就很容易被太郎烧饼店或老狼烧饼店蚕食市场份额。波特五力分析模型体系化地指导了企业分析的路径,在短时间内便可得出初步的判断结果。

在使用 MECE 原则时需要注意以下两点:

(1)在确立问题时,通过类似鱼骨图的方法,在确立主要问题的基础上,再逐个往下层层分解,直至所有的问题都被找到。通过问题的层层分解,可以分析出关键问题和初步的解决问题的思路。

(2)结合头脑风暴法找到主要的问题,然后在不考虑现有资源限制的基础上考虑解决问题的所有可能方法。在这一过程中,要特别注意多种方法的结合有可能形成新的问题解决方法,然后再往下分析每种解决问题的方法所需要的各种资源,并通过分析比较找到目前最现实和最令人满意的答案。

### (三)提出假设

当有了清晰且正确的问题定义和全面的结构化分析之后,就到了新麦肯锡五步法的第三步,即提出假设。

提出假设是自上而下方法论的实战落地,要求在项目初期还没有进行实地调研和缺少必要专业知识的情况下,用假设的方式准备可能的方案选项。

头脑风暴是提出假设的关键举措,也是战略咨询公司常用的解题工具之一。它采用一种非正式的讨论方式生成关键的思路或观点。头脑风暴既可以应用于针对问题的整体,也可以应用于聚焦细节层面的主题。头脑风暴应遵循三个原则,即差异、平等和创造性。

#### 1. 差异

差异是对假设的要求,同时是对头脑风暴参与者的要求。头脑风暴鼓励采用不同的视角,因此,迥异的背景直接带来思考角度的多样性。



## 2. 平等

平等能保证头脑风暴参与者破除壁垒,不受限制地阐述自己的看法,在讨论中没有任何“愚蠢”的假设,也没有“更重要”的假设。头脑风暴是一种“对事不对人”的百家争鸣。

## 3. 创造性

头脑风暴要有创造性思维,越是天马行空,越是发散思维,越是跳出框架,就越能挑战已有的模式和原则,带来突破性创新的可能。应鼓励用其他行业的成熟经验跨界解决当前的问题。

提出假设是新麦肯锡五步法中承上启下的关键一环,而且有别于平常的经验导向的思维定式,需要参与者适应和在实践中反复磨炼。这一步完全不同于毫无根据的“拍脑袋”,要有严格的验证假设的过程,及时修正并提出新的假设,形成科学验证的闭环。

### (四) 验证假设

验证假设需要收集的信息量很大。为信息收集而做的验证假设调研工作分为两种,即案头调研和实地调研。

#### 1. 案头调研

案头调研一般关注客户已有的书面资料、互联网公共信息、专业期刊、各种行业报告和内部资料库等,调研者根据这些资料总结出核心观点支持或否定之前提出的假设。

#### 2. 实地调研

实地调研是指在实验室或图书馆之外的地方实地收集一手客观数据,用严谨的逻辑验证之前假设的真伪。

调研收集信息的原则通常是以一手资料为主,应尽可能多地收集一手资料;以二手资料为辅,在使用二手资料时要尽量验证资料来源的真实性和可靠性。如果在实地调研收集数据时有新发现而生成新的假设,要再经过同样严谨的验证,如此循环,直到最终寻找到解决问题的最佳方案。

### (五) 解决问题

经过上述四步后,应仔细验证问题是否得到全面的解决、是否留有后遗症,要对问题提出和解决的过程进行全面的复盘,并保存整个工作过程,保留重要文件以备查,并为下一次的逻辑思考提供参照。

## 任务 三 连锁企业信息分析的方法

连锁企业信息分析的方法有结构分析、群组分析、对比分析、数字特征分析、相关分析、综合、回归分析、趋势分析等。

### 一、结构分析

#### (一) 结构分析的定义

结构分析是在统计分组的基础上计算各组成部分所占的比重,进而分析某一总体现象的内部结构特征、总体的性质、总体内部结构依时间推移而表现出的变化规律性的统计方



法。结构分析的基本表现形式是计算结构指标。结构指标的计算公式为

$$\text{结构指标} = \frac{\text{总体的某一部分}}{\text{总体}} \times 100\%$$

## (二) 结构分析的具体应用

例如,市场占有率是企业某种商品销售量占市场上该商品销售总量的百分比,由此可以看出企业经营的该种商品在市场上的分量。

结构分析是静态分析,表明截止某一时间点部分相对于总体的情况,是对历史的判断。使用结构分析时,需要结合动态数据加以辅助说明。例如,A公司的面包的市场占有率为3%,得不出有用的结论;但是如果知道A公司的面包连续三年销售量增长10%,则可以认为A公司的面包销售情况很好。

## 二、群组分析

### (一) 群组分析的概念

群组分析也称同期群分析,是按某个特征,将数据分为不同的组,然后比较各组的数据。例如,在学校上体育课时,体育老师考虑到男生和女生的运动项目不一样,会将男生分为一组打篮球,将女生分为一组跳绳。这其实是按性别对学生进行了分组。人们使用微信时建立群(工作群、家人群等)都是在进行群组分析。

### (二) 群组分析的应用

随着时间的推移,产品会出现新的版本,对于新版产品的效果如何、产品版本更新后用户是增长了还是流失了等问题,需要将用户按时间分组,然后比较不同组的用户留存率。因此,群组分析常用于分析用户留存率(或流失率)随着时间的推移发生了哪些变化,然后找出用户留下或者离开的原因。

#### 【示例:视频平台用户流失分析】

腾讯、优酷、爱奇艺等视频平台的用户按月付费成为会员才能观看某些视频。用户可以在任意月取消订购,取消订购的用户就是流失的用户。为了分析用户为何流失,可以使用群组分析。

表1-2所示为某视频平台的新增用户数,表格的第1列“分组”是按新用户注册的月份分组,每一行是对应组之后各个月留存下来的用户(数据来源:《精益数据分析》)。

表 1-2 某视频平台的新增用户数

分组	新增用户数										
	当月	1个月后	2个月后	3个月后	4个月后	5个月后	6个月后	7个月后	8个月后	9个月后	10个月后
1月组	150	140	130	125	118	105	102	97	95	95	95
2月组	180	172	160	150	140	130	121	118	118	118	
3月组	200	190	178	169	155	142	135	132	128		
4月组	270	188	175	170	153	144	137	131			



续表

分组	新增用户数										
	当月	1个月后	2个月后	3个月后	4个月后	5个月后	6个月后	7个月后	8个月后	9个月后	10个月后
5月组	350	247	228	216	202	189	178				
6月组	450	307	288	269	258	244					
7月组	225	210	195	180	166						
8月组	235	218	207	197							
9月组	240	224	211								
10月组	250	233									

“1月组”一行中,当月(1月)新增用户数是150人,1个月后有140人留存下来,2个月后有130人留存下来。“2月组”一行中,当月(2月)的新增用户数是180人,1个月后有172人留存下来,2个月后有160人留存下来。

现在计算留存率。对于“1月组”一行,1个月后的留存率=留下的人数(140)/1月组总人数(150)×100%=93.33%,2个月后的留存率=留下的人数(130)/1月组总人数(150)×100%=86.67%。

按照上述方法,可以把每一行的留存率计算出来,就得到了表1-3所示的数据。

表 1-3 某视频平台用户的留存率

分组	新增用户数										
	当月	1个月后	2个月后	3个月后	4个月后	5个月后	6个月后	7个月后	8个月后	9个月后	10个月后
1月组	1	93.33%	86.67%	83.33%	78.67%	70.00%	68.00%	64.67%	63.33%	63.33%	63.33%
2月组	1	95.56%	88.89%	83.33%	77.78%	72.22%	67.22%	65.56%	65.56%	65.56%	
3月组	1	95.00%	89.00%	84.50%	77.50%	71.00%	67.50%	66.00%	64.00%		
4月组	1	69.63%	64.81%	62.96%	56.67%	53.33%	50.74%	48.52%			
5月组	1	70.57%	65.14%	61.71%	57.71%	54.00%	50.86%				
6月组	1	68.22%	64.00%	59.78%	57.33%	54.22%					
7月组	1	93.33%	86.67%	80.00%	73.78%						
8月组	1	92.77%	88.09%	83.83%							
9月组	1	93.33%	87.92%								
10月组	1	93.20%									

当群组分析表格中的数据较多时,直接分析比较困难。这时可以把各组的数据绘制成折线图,这样就可以很容易地发现数据随着时间的推移发生了哪些变化。



把每组的数据绘制成一条折线,横轴是时间,纵轴是留存率,然后比较各组的折线。因为9月组和10月组的数据很少,所以没有将其绘制到图上。从图1-7中可以发现,当1、2、3月组的折线趋于平稳时,4、5、6月组的折线还在继续下行,尤其是4月组的变化明显。

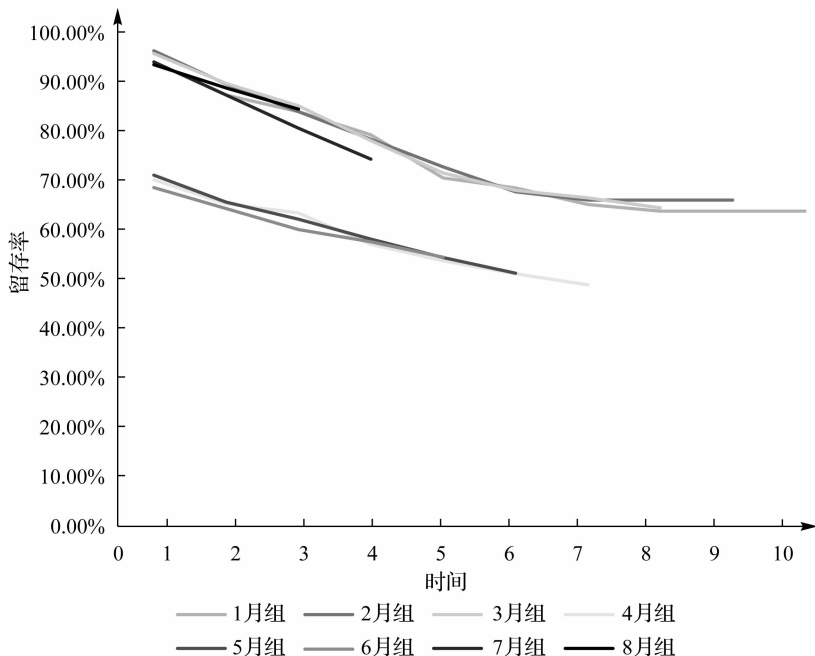


图 1-7 某视频平台用户的留存率

### 三、对比分析

为了了解研究的事物的具体状态或性质,通过对比分析,往往可以得出结论。《道德经》曰:“天下皆知美之为美,斯恶已;皆知善之为善,斯不善已。故有无相生,难易相成,长短相形,高下相倾,音声相和,前后相随。”因此,有对比才有发现。

#### (一)对比的三要素

在使用对比分析时,首先是对两种事物或同一种事物的两种状态可以比较,其次是比较要有一个指标或标准。对比的两种事物,一种是对比的主体,另一种是对比的客体,对比的指标为对比的度量。因此,对比的三要素为主体、客体和度量(指标)。

例如,某连锁店分析销售额,A店的月营业额为100万元,B店的月营业额为120万元,C店的月营业额为130万元;要了解为什么A店的销售额低,A店就是主体,B店和C店就是客体,月营业额就是指标。

#### (二)对比分析的应用

##### 1. 寻找对比量

最常用的对比是大小对比、数量对比,如销售额的对比、人数的对比、时间的对比。使用不同的标准会得出不同的结论。

对比指标的选择称为对比视角,对比视角不同,就会得出不同的结论。例如,小强身高





为 180 cm,小明身高为 185 cm,通过身高对比对两人的高矮可立即判断。当然,针对人的对比还有学习成绩、年龄、生活经历的对比。因此,人力资源专员在进行招聘时会综合考虑各种因素,对比各种指标,用到各种模型。对比随时随地发生,要找到合适的对比视角,针对同样的问题发现不同的结果。

对事物的原始状态描述的变量称为第一变量。例如,直接描述事物的变量有长度、数量、额度、宽度、高度等。汽车在 1 h 内行驶了 60 km,就有了速度,速度是距离和时间的比值,速度称为第二变量。在对比分析中,通过两个变量的比值而得到的变量称为第二变量,常用的有效率、效益等指标。

为了比较复杂的事物,往往一个变量只代表事物的一个侧面。如前所述,小明比小强度高,但是小强比小明长得帅气;有的客户购买力不强,但是影响力大,能够让亲朋好友都来购买。要比较复杂的事物,就要考虑综合指标。

日常生活中,在衡量物价变化时,大米价格上涨,玉米、小米、糯米、面粉的价格下降,就不能认为物价上涨,评价物价波动需要用衡量物价的综合指标——消费者物价指数(Consumer Price Index,CPI)或工业价格指数——生产者物价指数(Producer Price Index,PPI)等指标衡量物价的波动。在企业管理中,也会采用一些综合的评价指标进行对比。例如,可使用平衡计分卡(Balanced Score Card,BSC)根据财务指标、客户指标、成长指标和流程指标分析公司的业绩,进而制定战略。

在对数据进行综合对比分析时,除了简单地直接对比数据外,还要构建一些可以重复使用或者在某个部门、某个业务领域、某个情境下进行评测的指标。这些指标背后可以是多个数据综合分析的结果,也可以是某个业务指标的合集。在实践中,经过不断检验的指标才能成为公司持续使用的指标。一个综合指标用得越久,就越完善,并且可以体现出公司管理的特色。

## 2. 对比的层次和维度

从过程看,对比可分为与自己比和与别人比。与自己比包括与自己的过去比和与自己设定的目标比。与别人比则复杂得多。例如,对两个同行业的企业可以从规模指标、速度指标、效率指标、效益指标和综合指标五个维度进行比较。其中,规模指标包括市场份额、销售额、资源规模等;速度指标包括市场增长率、开店率等;效率指标包括人均产值、人均销售额、人均利润等;效益指标包括顾客满意度、利润率、员工满意度等;综合指标就是将以上指标进行综合加工,形成特定的算法。

## 四、数字特征分析

分析单一特征(或属性)的一组数据时,常用的数字特征有一次数值代表值(最大值、最小值)和二次数值代表值(频数、平均数、期望、方差)。

### (一)一次数值代表值

在进行数据处理的过程中,首先要做的就是对数据进行初次分析,就是通过找最大值、最小值、众数观察、研究事物的特点。例如,要研究某课程的考试成绩,首先要看最高分数、最低分数和出现频率最多的分数。

### (二)二次数值代表值

经过对数据的初步观察后,可利用数据统计分析工具进行分析,主要是对数据进行频数



统计、观察数据的集中度和离散度。

### 1. 频数统计

对于一组数据,考察不同数值出现的频数或者数据落入指定区域内的频数,可以了解数据的分布状况。数据频数分析主要通过频数分布函数和直方图等工具进行。

例如,某超市统计蔬菜日销售额,调查了 50 天的销售额,具体如表 1-4 所示。

表 1-4 某超市 50 天蔬菜销售额统计

单位:元

某超市 50 天蔬菜销售额									
880	1 180	1 100	860	1 420	1 210	1 260	1 180	1 510	1 050
1 230	1 420	1 510	1 180	1 170	1 100	1 510	1 160	1 070	1 010
1 210	1 140	1 030	1 270	1 190	1 370	1 420	1 230	1 370	1 080
1 070	1 160	830	960	1 320	1 460	930	1 190	810	1 050
1 230	860	1 340	870	940	970	980	1 130	1 120	1 013

基本分析方法为:进行初次分析,可知最大值为 1 510、最小值为 810;以 100 为间距,可将数据分为 810~910、910~1 010、1 010~1 110、1 110~1 210、1 210~1 310、1 310~1 410、1 410~1 510、1 510 等八个组(左闭右开),各组数据如表 1-5 所示。

表 1-5 某超市 50 天蔬菜销售额频数统计

单位:元

分组	810~910	910~1 010	1 010~1 110	1 110~1 210	1 210~1 310	1 310~1 410	1 410~1 510	1 510
频数	6	6	8	12	4	4	7	3

通过频数分析,可以看出 50 天内 1 110~1 210 这组销售额出现的频数最多。一般情况下,统计频数采用直方图,如图 1-8 所示。

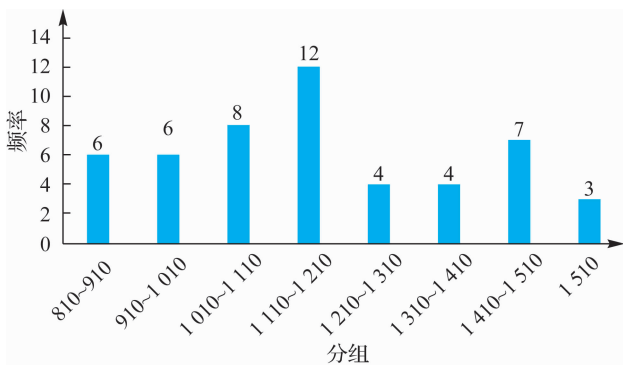


图 1-8 某超市 50 天蔬菜销售额频数统计

### 2. 平均分析法

平均分析法就是用平均数衡量总体在一定时间和地点某一数据的一般水平。平均数据比总量指标更具有说服力,更有助于运营者预测发展趋势和规律。平均数据有算术平均值、几何平均值、对数平均值等,其中最常用的是算术平均值,算术平均值=总体各数据的总和/数据个数。

平均指标是同质总体中各个单位某一指标的平均数字,是一个非常重要的基础性指标,



可以广泛应用于不同的行业和不同的事物。可以综合使用平均分析法与前面所列的方法。

### 3. 方差和标准差

方差和标准差是测度数据变异度最重要、最常用的指标。方差是各个变量值与其算术平均值的离差平方的平均数。对方差的计量单位和量纲不便于从经济意义上进行解释,所以实际统计工作中多用方差的算术平方根——标准差测度总体的离散程度。标准差又称均方差,具有量纲,与变量值的计量单位一致。方差和标准差的计算方法可分为简单平均法和加权平均法。另外,对于总体和样本,方差和标准差的公式略有不同。

总体方差的计算公式为

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

总体标准差的计算公式为

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

总体方差和总体标准差在对各个离差平方平均时是除以数据的个数或总频数,而样本的方差在对各个离差平方平均时是用总离差平方和除以样本数据的个数或总频数减 1。

样本方差的计算公式为

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

样本标准差的计算公式为

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

### 4. 离散系数

离散系数中的全距、平均差、方差和标准差都是反映变量值离散程度的绝对指标。其数值不仅取决于变量值的差异程度,还受变量值平均水平的影响,而且与原变量值的计量单位有关。对原变量值采用不同的计量单位,其离散程度指标的数值也不同。因此,对于平均水平不同或计量单位不同的不同组别的变量值,是不能直接用上述变异指标进行比较的。为了消除变量值水平不同和计量单位不同对离散程度的影响,需要计算离散系数。对离散系数通常是用标准差计算的。因此,离散系数也称为标准差系数,是标准差与其相应的算术平均值的比值,是反映变量值离散程度的相对指标,用  $V_\sigma$  表示,计算公式为

$$V_\sigma = \frac{\sigma}{\bar{x}}$$

## 五、相关分析

### (一) 认识相关分析

有时研究的问题只有一种数据,如人的身高;但是,还有另外一些问题的研究需要多种



数据,如身高和体重之间的关系。当研究两种或者两种以上的数据之间的关系时,就要用到相关分析。如果两种数据之间有关系,称为有相关关系;如果两种数据之间没有关系,称为没有相关关系。

相关关系的计算公式为

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

其中, $\bar{x}$ 与 $\bar{y}$ 为对应两个变量数列的均值。

出现以下三种情况时可以进行相关分析:

(1)在研究两种或者两种以上数据之间的关系,或者某个事件受到其他因素影响的问题时,可以使用相关分析。在职场中经常会遇到的一个问题是子产品对整体产品所具有的贡献度或者影响度。例如,对于“微信读书”这款产品,评估“想法”子模块的用户留存对整体产品留存的影响度,就可以使用相关分析,研究子产品和整体产品的关系。

(2)在解决问题的过程中,相关分析有助于扩大思路,将视野从一种数据扩大到多种数据。例如,在分析“为什么销量下降”的过程中,可以研究哪些因素(如产品价格、售后服务等)和销量有关系。使用相关分析,可以知道哪些因素影响销量、哪些因素对销量没有影响,从而快速锁定问题出现的原因。

(3)相关分析有助于研究者或者相关人员加深对问题的理解。在实际工作中,因为数据分析的结果需要得到其他人的理解和认可,需要参与研究的人员相互沟通。很多分析方法使用起来很复杂,需要相关知识的辅助,才能有助于人们正确地认识和理解问题。

## (二)相关分析的应用

表 1-6 记录了 20 名学生为考试花费的学习时间和取得的成绩,现在想知道学习时间和取得的成绩这两种数据之间有什么关系。

表 1-6 学习时间和取得的成绩

学习时间/h	0.5	0.75	1	1.25	1.5	1.75	1.75	2	2.25	2.5
取得的成绩/分	10	22	13	43	20	22	33	50	62	48
学习时间/h	2.75	3	3.25	3.5	4	4.25	4.5	4.75	5	5.5
取得的成绩/分	55	75	62	73	81	76	64	82	90	93

用散点图可以更直观地看出学习时间和取得的成绩之间的关系,如图 1-9 所示。

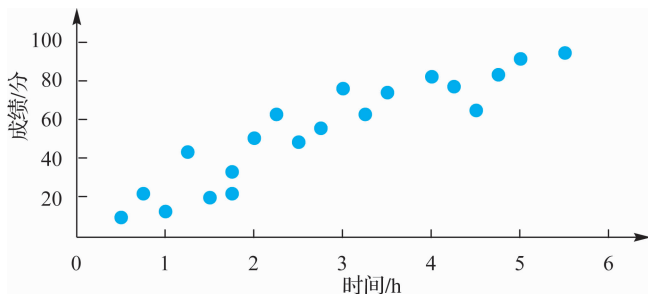


图 1-9 学习时间和取得的成绩之间的关系



每位学生的数据在图中是一个点。例如,图中的最高点表示某位学生的学习时间是 5.5 h,成绩是 93 分。相关系数是用来衡量两种数据之间的相关程度的指标,通常用字母  $r$  表示。相关系数数值的正负可以反映两种数据之间的相关方向,也就是两种数据在变化过程中是同方向变化还是反方向变化。相关系数的取值范围是  $-1\sim 1$ ,  $-1$ 、 $0$  和  $1$  三个值是相关系数的极值。

(1) 如果相关系数  $r=1$ ,数据点都在一条直线上,表示两种数据之间完全正相关,两种数据是同方向变化,即数据  $x$  的值越大,数据  $y$  的值也会越大。

(2) 如果相关系数  $r=-1$ ,数据点都在一条直线上,表示两种数据之间完全负相关,两种数据是反方向变化,即数据  $x$  的值越大,数据  $y$  的值反而会越小。

(3) 如果相关系数  $r=0$ ,表明两种数据之间不是线性相关,但有可能是其他形式的相关。

(4) 如果相关系数  $r>0$ ,说明两种数据是正相关,是同方向变化,即一种数据的值越大,另一种数据的值也会越大。如果相关系数  $r<0$ ,说明两种数据是负相关,是反方向变化,也就是一种数据的值越大,另一种数据的值反而会越小。

相关系数  $r$  的绝对值越大,说明两种数据的相关程度越高,如图 1-10 所示。

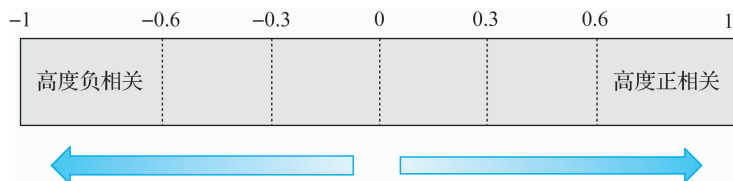


图 1-10 相关关系判断

判断相关系数的数值并没有统一的规定,一般将相关系数分为三部分:如果相关系数的绝对值为  $0\sim 0.3$ ,就认为是低度相关;如果相关系数的绝对值为  $0.3\sim 0.6$ ,就认为是中度相关;如果相关系数的绝对值为  $0.6\sim 1$ ,就认为是高度相关。根据这一大致的相关系数分类,就可以知道两种数据的相关程度。

## 六、综合

综合是与分析法相对立的一种方法。综合是在分析的基础上进行的,是指人们在思维过程中将与事物有关的片面、分散、众多的各个要素(情况、数据、素材等)结合起来考虑,从错综复杂的现象中探索它们之间的相互关系,从整体的角度把握事物的本质和规律,通观事物发展的全貌和全过程,获得新的知识、新的结论的一种逻辑方法。综合把对事物的各个要素之间的认识统一为整体的认识,从整体上把握事物的本质和规律。它不是主观地、任意地把事物的各个要素简单地捏合在一起,而是按照各个要素在事物内部的有机联系从总体上把握事物。它不是抽象地从外部现象的联系上理解事物,而是抓住事物的本质,即抓住事物在总体上相互联系而又矛盾的特殊性,研究这一矛盾如何制约事物丰富多彩的属性、如何在事物的运动中展现出整体的特征。在信息分析中,综合的结果往往是新的发现。

### (一) 综合的基本步骤

(1) 明确综合的目标。

(2) 把握被分析出来的事物的各个组成要素。





(3)确定各个要素的结构形式和有机联系。

(4)将各部分重新组合成一个整体,从整体的角度把握事物的本质和方向,从而获得新的知识和结论。

在信息分析中,综合是一种行之有效的方法。综合可以将分散、片面、内容各异的有关信息(情况、数据、素材等)按特定的目的汇集、整理、归纳和提炼,从而形成系统、全面、新颖的知识和结论。例如,可持续发展是当今世界各国经济和社会发展过程中普遍关心的问题,可持续发展通常涉及资源、环境、人口、资本和技术等方面。

## (二)综合的分类

在信息分析中,人们常用的综合主要有简单综合和系统综合。

### 1. 简单综合

简单综合是对与研究课题有关的信息(情况、数据、素材等)进行汇集、整理和归纳,可细分为以下两种:

(1)从部分到整体的综合。首先把分散在各种信息资源中不同行业的同类信息加以综合处理;然后把来自不同层次、不同地区的信息素材集中在一起,进行整体性研究,概括出比较完整、准确、有层次的信息成果。

(2)从整体到部分再到整体的综合。首先明确研究的整体,然后把分散在各种信息资源中的有关同一行业的不同信息加以归纳、综合,进行整体性分析和研究,形成信息成果。

### 2. 系统综合

系统综合是从系统论的观点出发,对与研究对象有关的信息进行时间与空间、纵向与横向等方面的综合研究。系统综合强调思维要从“以实物为中心”逐渐过渡到“以系统为中心”。其理论出发点是任何系统内部各要素间都是相互联系、相互作用的,任何系统都有一定的结构层次,并处于一定的环境中,受到环境的制约,系统的目标决定着系统的功能。系统综合不是简单的信息汇集、整理和归纳,而是一个创造性的、深入认识研究课题的过程。

按研究对象,系统综合法可以进一步细分为学科信息系统综合法、企业信息系统综合法、政策信息系统综合法等。

## (三)分析与综合的关系

分析与综合是辩证统一的关系,具体如下:

(1)分析与综合是相互依存的关系。综合离不开分析。正确地综合必须了解在现实统一体中存在哪些特殊的规定及最根本的性质,而要做到这些,就必须依赖于对客观事物的周密分析。同理,分析离不开综合,任何统一体中矛盾的各方面都是相互联系、相互依赖的,要对客观事物有一个全面正确的认识,就必须懂得统一体内部各个方面是如何统一的、如何能够统一的,为此必须加以综合。

(2)分析与综合是相互渗透的关系,分析之中有综合。事实上,任何分析都要从某种整体出发,总是以对客观的整体性认识为指导,是综合指导下的分析。

(3)分析与综合在一定条件下可以相互转化。当分析进行到充分成熟的阶段时,分析的方法就转化为综合的方法。随着认识的深化,综合又被分解,新的分析开始,实现着综合向分析的转化。在信息分析中,分析与综合总是结合在一起被使用的。没有分析的综合或者没有综合的分析,很难保证信息分析成果的质量。



## 七、回归分析

### (一) 回归分析的基本概念

在连锁企业经营实践中,各种经济变量相互联系、相互制约。通过相关分析,可以分析现象之间相关关系的方向和相关关系的密切程度。但相关分析不能判断现象之间具体的数量变动的依存关系,也不能根据相关系数估计或预测因变量可能发生的数值。因此,为了探求企业经济变量之间的具体数量变动关系,一般在相关分析的基础上进行回归分析。回归分析是对具有相关关系的两个或两个以上变量之间数量变化的一般关系进行测定,确定因变量和自变量之间数量变动关系的数学表达式,以便对因变量进行估计或预测的统计分析方法。显然,相关分析的主要任务是研究变量间相关关系的表现形式和密切程度,而回归分析是在相关分析的基础上进一步研究现象之间的数量变化规律。

### (二) 回归分析的主要内容

#### 1. 确定自变量和因变量

应根据研究目的和现象之间的内在联系确定自变量和因变量。现象之间除了有相关关系,还存在着因果关系。作为原因的变量为自变量,作为结果的变量为因变量;或者说影响因素为自变量,被影响因素为因变量。进行回归分析时,首先应该从理论出发进行定性分析,然后根据现象的内在联系确定变量之间的因果关系,从而确定自变量和因变量。

#### 2. 确定数学表达式

可根据现象之间的内在影响机制或通过对具体变量数据描点进行分析,找出合适的回归分析模型,再通过计算求出模型的待估参数得到回归方程。确定回归方程最常用的方法是最小二乘法。

#### 3. 对回归分析模型进行评价与诊断

得到具体的回归方程以后,要对其进行统计检验。例如,对回归方程计算一些检验统计量,以对回归方程的代表性及拟合程度进行评价,检验、判断回归分析模型的基本假设是否合理,并做相应的改进。

#### 4. 确定因变量的估计值

应根据给定的自变量数值确定因变量的数值。回归方程可以用于统计估计或预测,即根据给定的自变量数值估计因变量的数值或置信区间,以及利用回归模型进行回归控制。

### (三) 简单回归分析模型——简单线性回归方程

在回归分析中,许多相互关联的两个变量之间存在着线性关系。例如,消费者的家庭消费支出与家庭收入之间基本上是一种线性相关关系。虽然在很多情况下影响因变量的因素不止一个,但在实际工作中,往往因客观条件的限制或者出于研究的目的,需要突出其中一个最重要的因素,即只研究某一个自变量对因变量的影响。这是对经济过程的一种抽象,抓住主要矛盾,得出最有意义的结论。简单线性回归分析是所有回归分析的基础,多元回归分析和非线性回归分析都是从简单回归分析的基础理论上延伸发展出来的。

在相关分析中,通过计算相关系数,可以判断两个变量之间直线相关的紧密程度,但不能说明它们之间因果的数量关系。简单线性回归就是对具有显著线性相关的两个变量间数量变化的一般关系进行测定,拟合一个直线回归方程,以便进行估计或预测的统计方法。



简单线性回归分析中只有一个因变量和一个自变量,这是变量最少、最简单的回归方程。它在平面坐标图上表现为一条直线,故也称为简单直线回归方程。简单线性回归方程的理论模型与估计模型可分别被描述成理论模型和估计模型。

理论模型的计算公式为

$$y = \alpha x + \beta$$

估计模型的计算公式为

$$y = ax + b$$

式中: $\alpha$ 、 $\beta$ 为回归参数或待定系数; $a$ 、 $b$ 为相应的估计值。

$a$ 、 $b$ 的值确定后,估计的直线方程就确定了。具体的参数估计应用最小二乘法,得到如下的参数估计结果:

$$\begin{cases} a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \\ b = \bar{y} - a\bar{x} \end{cases}$$

由此可以确定两个参数。

## 八、趋势分析

趋势分析是根据社会经济发展过程和规律进行引申和外扩,预测其发展趋势的方法。研究信息分析的终极目标是对未来发展趋势进行预测。对未来趋势发展预测和分析的方法有经验预测法、类比预测法、惯性预测法、逻辑关系预测法和时间序列预测法。根据本书的研究内容,主要针对两种时间序列预测法(移动平均法、指数平滑法)进行简单的介绍。

### (一)移动平均法

以门店销售额为例。从门店的历史经营数据中选取近  $n$  期数据,再从这些数据中选取  $m$  ( $m < n/2$ ,  $m$  一旦确定,就不能再改变)期数值。移动平均法是以  $m$  期数值为基础求平均值,并不断向后移动,连续计算观测值平均数,以最后一组平均数作为未来销售预测值的一种方法,具体的计算公式为

$$F_t = (A_{t-1} + A_{t-2} + A_{t-3} + \cdots + A_{t-m})/m$$

式中: $F_t$ 为对下一期的预测值; $m$ 为移动平滑期,表示选取的求平均值的数值的个数; $A_{t-1}$ 为前期实际值,以此类推, $A_{t-m}$ 为前第  $m$  期实际值。

用更加直接的方式表述以上公式,即“预测值=最后一个  $m$  期平均值”。为了使预测值更能反映销售量的变化趋势,可以对上述结果进行修正。修正求出的预测值为

$$\text{预测值} = \text{最后一个 } m \text{ 期平均值} + \text{趋势值}$$

$$\text{趋势值} = \text{最后一个移动平均值} - \text{上一个移动平均值}$$

例如,某连锁门店 2021 年 1—9 月的销售额如表 1-7 所示,试预测 10 月的销售额。

表 1-7 某连锁门店 1—9 月的销售额

单位:万元

月份	1	2	3	4	5	6	7	8	9
销售额	340	270	300	303	295	296	266	301	277

首先用移动平均法预测 10 月的销售额。确定  $m$  值, $m$  要小于  $n/2$ 。因此,可以确定为  $m=4$ ,就是从 9 月开始(包括 9 月在内)向前找四个销售额,计算平均值。设 10 月的销售额



为  $Q$ , 则  $Q = (277 + 301 + 266 + 296) \div 4 = 285$  万元, 即根据移动平均法预测 10 月的销售额为 285 万元。

然后用修正移动平均法预测 10 月的销售额。根据上述公式, 需要对上一期移动平均值进行计算。设  $Q_1$  为 6—9 月的移动平均值,  $Q_2$  为 5—8 月的移动平均值, 趋势值为  $k$ , 计算得  $k = Q_1 - Q_2 = -4.5$  万元。设修正预测值为  $Q_0$ , 则  $Q_0 = Q_1 + k = 285 + (-4.5) = 280.5$  万元。

## (二) 指数平滑法

指数平滑是移动平均法的一种, 其特点是对实际发生的数值给予不同的权重, 即给予近期的数据更大的权重, 给予远期的数据更小的权重, 基本计算公式为

$$S_t = \alpha y_{t-1} + (1 - \alpha) S_{t-1}$$

式中:  $S_t$  为  $t$  期的预测值或者  $t$  期的平滑预测值;  $y_{t-1}$  为  $t-1$  期的实际发生值;  $S_{t-1}$  为  $t-1$  期的平滑预测值;  $\alpha$  为平滑常数。

若求  $t$  期预测值, 需要将  $t-1$  期的观测值全部平滑处理一遍, 平滑后的预测值与实际值之间就有差异, 主要原因是平滑常数  $\alpha$ 。  $\alpha$  的取值符合如下规律:

- (1) 当时间序列呈现出稳定的水平时,  $\alpha$  的取值为 0.1~0.3。
- (2) 当时间序列呈现出波动且波动较大时,  $\alpha$  的取值为 0.3~0.5。
- (3) 当时间序列有上升或下降的趋势时,  $\alpha$  的取值为 0.6~0.9。

实际应用中, 应对  $\alpha$  取若干值进行测试, 选择误差最小的。

例如, 某企业 2013—2021 年的销售额如表 1-8 所示, 试利用指数平滑法预测 2022 年的销售额,  $\alpha$  取值为 0.1、0.6、0.9。

表 1-8 某企业 2013—2021 年的销售额

单位: 万元

年份	2013	2014	2015	2016	2017	2018	2019	2020	2021
销售额	4 000	4 700	5 000	4 900	5 200	6 600	6 200	5 800	6 000

计算过程如表 1-9 所示。

表 1-9 计算过程

单位: 万元

年份	销售额	$\alpha=0$	$\alpha=0.1$		$\alpha=0.6$		$\alpha=0.9$	
	$X$	$S$	$S$	$X-S$ 的绝对值	$S$	$X-S$ 的绝对值	$S$	$X-S$ 的绝对值
2013	4 000	4 566.67 [一般方法: 当观测期小于 15 时, 取前三个数的算数平均值, 即 (4 000+4 700+5 000)÷3=4 566.67]	4 510.00	510.00	4 227.67	226.67	4 056.67	56.67
2014	4 700		4 529.00	171.00	4 510.67	189.33	4 635.67	64.33
2015	5 000		4 576.10	423.90	4 804.27	195.73	4 963.57	36.43
2016	4 900		4 608.49	291.51	4 861.71	38.29	4 906.36	6.36
2017	5 200		4 667.64	532.36	5 064.68	135.32	5 170.64	29.36
2018	6 600		4 860.88	1 739.12	5 985.87	614.13	6 457.06	142.94
2019	6 200		4 994.79	1 205.21	6 114.35	85.65	6 225.71	25.71
2020	5 800		5 075.31	724.69	5 925.74	125.74	5 842.57	42.57
2021	6 000		5 167.78	832.22	5 970.30	29.70	5 984.26	15.74
			合计	6 430.01	合计	1 640.56	合计	420.11



因此,选取误差最小的平滑常数 $\alpha=0.9$ ,由此预测 2022 年的预测值为

$$S_{2022}=0.9 \times 6\,000 + 0.1 \times 5\,984.26 = 5\,998.43 \text{ 万元}$$

## 项目小结

本项目主要介绍了连锁企业信息分析的主要内容、分析思路和分析方法。在介绍连锁企业信息分析的内容时引入了信息的概念,介绍了信息的综合特性,并以此为依据提出连锁企业信息分析的概念。根据对连锁企业信息分析的认识,介绍了连锁企业信息分析的基本思路。信息分析首先要找到问题,找到问题后又必须具备分析问题的逻辑思维能力。因此,本项目从 5W2H 分析法入手,培养学生分析问题的能力,然后根据连锁企业门店面对的运营和管理的客体,从顾客方面、运营方面、管理方面提出用户行为分析法、4E 分析法和逻辑树分析法,着重通过以上 4 种方法阐明连锁企业信息分析必须从问题切入,发现问题是一切分析的开始。而后对连锁企业经营的具体业务提出具体的八种信息分析的方法。

## 过关练习

### 一、数字思维培养

1. 14, 16, 18 的算术平均值是 16。 ( )
2. 某公司以往 3 年分析销售额依次为 450 万元、570 万元、680 万元,依据此可以推断当年的销售额应高于 680 万元。 ( )
3. 某小学 60 名 11 岁的学生身高(单位:cm)如下:  
126 149 143 141 127 123 137 132 135 134 146 142 135 141 150  
137 144 137 134 139 148 144 142 137 147 138 140 132 149 131  
139 142 138 145 147 137 135 142 151 146 129 120 143 145 142  
136 147 128 142 132 138 139 147 128 139 146 139 131 138 149  
其中,身高最高的是 151 cm。 ( )
4. 若有个自然数能够同时被 5 和 2 整除,那么这个数的个位一定是 0。 ( )
5. 如果某产品的市场规模达到 30 亿元,A 企业占据 20% 的市场份额,则 A 企业的销售额为 6 亿元。 ( )
6. 某成衣厂对 9 名缝纫工进行技术评比。9 名工人的得分恰好是等差数列,9 人的平均得分是 86 分,前 5 名工人的得分和是 360 分。前 7 名工人的得分和是多少?
7. 已知数列 5, 11, 19, 29, 请推断下一个数为多少。
8. 数独如图 1-11 所示。

	2	3	4
3	4	1	
2	1	4	
	3	2	1

图 1-11 数独





在空格处填上数字 1~4,使得色彩相同的矩形内 1~4 不重复,并且行列内 1~4 不重复。

## 二、逻辑思维培养

将下列短语重新排序,用以表达清楚的句意。

1 有效的,2 我们,3 信息,4 处理,5 进行,6 将,7 面目,8 高效地,9 真正的,10 事物,11 才能,12 看到。

## 三、基础知识过关

1. 在没有任何约束的条件下,可以将信息定义为事物的存在方式和运动状态的表现形式,这是信息的( )概念。

- A. 本体论                      B. 信息论                      C. 认识论                      D. 方法论

2. 连锁企业信息分析是对连锁企业经营管理数据中有价值的结论提取的过程,这是对信息概念的( )认识。

- A. 本体论                      B. 信息论                      C. 认识论                      D. 方法论

3. 顾客购买房图中,( )是让顾客形成正反馈的基础。

- A. 顾客满意                      B. 购买后愉快的体验  
C. 门店位置                      D. 门店的装潢和布局

4. ( )是解决问题的根本出发点。

- A. 定义问题                      B. 分析问题                      C. 提出假设                      D. 检验假设

5. 日常生活中问题界定清晰有三个标准,下面不为该标准的是( )。

- A. 边界清晰,无交叉干扰                      B. 问题解决后没有后遗症  
C. 问题确定,标准随之确定                      D. 问题确定,解决方案随之确定

6. 对比有三个要素,分别是主体、客体和( )。

- A. 结果                      B. 指标                      C. 对象                      D. 形态

7. 相关系数为零,说明研究的两个变量之间( )。

- A. 线性正相关                      B. 线性负相关                      C. 不相关                      D. 非线性相关