

模块一

电子商务大数据分析基础

早在 20 世纪 80 年代，美国就有人提出了“大数据”的概念，大数据时代的到来也为电子商务带来观念的转变及对数据的新管理模式，使得数据的实际应用能更有效地与企业运营结合。大数据分析可以为企业经营决策提供参考依据，进而为企业创造更多的价值。本模块对电子商务大数据分析的概念、应用、工作流程及职业前景进行大致的介绍，并且讲解获取原始数据的几种重要途径。

第1章

电子商务大数据分析概论

1.1 初识电子商务行业中的大数据

随着网络和信息技术的日益普及，大数据逐步改变甚至颠覆着人类的生活方式，随着数据量呈指数级增长及云计算的诞生，我们步入大数据时代。大数据逐步向各行业渗透、辐射，颠覆着很多特别是传统行业的管理和运营思维。大数据更是触动着电子商务行业管理者的神经，搅动着电子商务行业管理者的思维，大数据在电子商务行业释放出的巨大价值吸引着诸多电子商务行业人士的兴趣和关注。

之所以称之为大数据，是因为电子商务时代的数据管理量很大，以至于无法在可承受的时间范围内用常规软件工具对数据进行捕捉、管理和处理。进行电子商务数据分析的意义在于对数据进行专业化处理，得到一个服务于企业管理决策或提升客户体验的结论。当然，大数据在电子商务行业中的应用有着越来越多的创新性发展。

1.1.1 借助大数据分析优化市场定位

电子商务企业要想在互联网市场站稳脚跟，必须架构大数据战略，对外要拓宽电子商务行业调研数据的广度和深度，从大数据中了解电子商务行业市场的构成、细分市场特征、消费者需求和竞争者状况等众多因素；对内企业想进入或开拓某一区域电子商务行业市场，首先要进行项目评估和可行性分析，决定是否开拓某块市场，最大化规避市场定位不精准给投资商和企业自身带来的毁灭性损失。



迷你案例

Decide.com 是一家预测商品价格并为消费者提供最佳购买时间，哪款产品评价好、值得买建议的创业公司。Decide.com 董事长兼 CEO Mike Fridgen 在一份声明中称：“它很真诚地告诉每一个亲爱的用户，Decide 网站的价格预测不是表面文章，因为所有的价格建议都来自我们专业的数据分析结果。”

不错, Decide.com 通过抓取亚马逊、百思买、新蛋网及全球各大网站上数以十亿计的数据进行分析, 最终将其整合在一个页面中方便用户对比查看, 并且能够预测产品的价格趋势, 帮助用户确定产品的最佳购买时机。

为了证明预测的准确性, Decide.com 每天都会选择 10 个交易, 如果它预测的价格在两个星期内降价, 则 Decide.com 将会自动通知买方, 并支付价格下降的金额 (最多 200 美金)。

注: Decide.com 于 2011 年 6 月成立, 2013 年 9 月被 eBay 收购, 同时对外关闭。



迷你案例

Iron Maiden, 这个颇有争议的英国重金属乐队从 1976 年出道火到现在。其通过对客户上网数据的分析, 定位于歌迷分布密度最大的区域——圣保罗, 并在此举办了一场演唱会, 取得了空前的成功。

市场定位对电子商务行业市场开拓非常重要, 但是, 要想做到这一点, 就必须有足够的信息数据来供电子商务行业研究人员分析和判断, 数据的收集、整理就成了最关键的一步。在传统分析情况下, 分析数据的收集主要来自统计年鉴、行业管理部门数据、相关行业报告、行业专家意见及属地市场调查等, 这些数据多存在样品量不足、时间滞后和准确度低等缺陷, 研究人员能够获得的有效信息量非常有限, 使准确的市场定位存在着数据瓶颈。

互联网时代, 借助数据挖掘和信息采集技术, 不仅能够给研究人员提供足够的样本量和数据信息, 而且能够建立基于大数据的数学模型对企业未来市场进行预测。

1.1.2 借助大数据优化市场营销

今天, 从搜索引擎、社交网络的普及到人手一机的智能移动设备, 互联网上的信息总量正以极快的速度不断暴涨。每天的 Facebook、Twitter、微博、微信、论坛、新闻评论、电子商务平台上分享的各种文本、照片、视频、音频等信息高达几百亿甚至几千亿条, 涵盖商家信息、个人信息、行业资讯、产品使用体验、商品浏览记录、商品成交记录、产品价格动态等海量信息。这些数据通过聚类可以形成电子商务行业大数据, 其背后隐藏的是电子商务行业的市场需求、竞争情报, 闪现着巨大的财富价值。

在电子商务行业市场营销中, 无论是产品、渠道、价格还是客户, 可以说每一项工作都与大数据的采集和分析息息相关, 以下两个方面内容是电子商务行业市场营销工作的重中之重。

- 对外: 通过获取数据并加以统计分析来充分了解市场信息, 掌握竞争者的商情和动态, 知晓产品在竞争群中所处的市场地位, 达到“知己知彼, 百战不殆”的目的。
- 对内: 企业通过积累和挖掘电子商务行业消费者数据, 有助于分析消费者的消费行为和价值趋向, 便于更好地为消费者服务和忠诚客户。

以电子商务行业对消费者的消费行为和趋向分析为例, 企业平时善于积累、收集和整理消费者的消费行为方面的信息数据, 例如:

- 消费者购买产品的花费。
- 选择的产品渠道。
- 偏好产品的类型。
- 产品使用周期。
- 购买产品的目的。

- 消费者的家庭背景。
- 消费者的工作和生活环境。
- 个人消费观和价值观等。

如果企业收集到了这些数据，并建立了消费者大数据库，则可通过统计和分析来掌握消费者的消费行为、兴趣偏好和产品的市场口碑现状；再根据这些总结出来的行为、兴趣爱好和产品口碑现状制定有针对性的营销方案和营销战略，投消费者所好，其带来的营销效应是可想而知的。

1.1.3 大数据助力电子商务企业的收益管理

收益管理是一种起源于 20 世纪 80 年代的谋求收入最大化的新经营管理技术，意在把合适的产品或服务在合适的时间以合适的价格，通过合适的销售渠道出售给合适的顾客，最终实现企业收益最大化目标。要达到收益管理的目标，需求预测、细分市场和敏感度分析是此项工作的 3 个重要环节，而这 3 个环节推进的基础就是大数据。

- 需求预测：通过对建构的大数据进行统计与分析，采取科学的预测方法，通过建立数学模型，使企业管理者掌握和了解电子商务行业的潜在市场需求、未来一段时间每个细分市场的产品销售量和产品价格走势等，从而使企业能够通过价格的杠杆来调节市场的供需平衡，针对不同的细分市场来实行动态的前瞻性措施，并在不同的市场波动周期以合适的产品和价格投放市场，获得潜在的收益。
- 细分市场：为企业预测销售量和实行差别定价提供条件，其科学性体现在通过电子商务行业市场预测来制定和更新价格，使各个细分市场的收益最大化。
- 敏感度分析：通过需求价格弹性分析技术，对不同细分市场的价格进行优化，最大限度地挖掘市场潜在的收入。

大数据时代的来临，为企业收益管理工作的开展提供了更加广阔的空间。需求预测、细分市场和敏感度分析对数据需求量很大，而传统的数据分析大多是采集企业自身的历史数据来进行预测和分析的，容易忽视整个电子商务行业的信息数据，因此预测结果难免存在偏差。企业在实施收益管理的过程中，在自有数据的基础上，依靠一些自动化信息采集软件来收集更多的电子商务行业数据，了解更多的电子商务行业市场信息，将会对制定准确的收益策略、赢得更高的收益起到推进作用。

1.1.4 大数据协助创造客户新的需求

差异化竞争的本质在于不停留在产品原有属性的优化上，而是创造了产品的新属性。满足客户需求是前提，但创造客户新需求才是行业革命的必要条件。

随着网络社交媒体的技术进步，以及论坛、博客、微博、微信、电子商务平台、点评网等媒介在 PC 端和移动端的创新和发展，公众分享信息变得更加便捷自由，而公众分享信息的主动性促进了“网络评论”这一新型舆论形式的发展。微博、微信、点评网、论坛上众多的网络评论形成了交互性的大数据，其中蕴藏了巨大的电子商务行业需求开发价值，这些数据已经受到了电子商务管理者的高度重视。很多企业已把“评论管理”作为核心任务来抓，既可以通过客户评论及时发现负面信息进行危机公关，更核心的是还可以通过这些数据挖掘客户需求，进而改良企业的产品，提升客户体验。

1.2 数据与数据的时代观

数据（data）是事实或观察的结果，是对客观事物的逻辑归纳，是用于表示客观事物的未经加工的原始素材。将为了特定研究而搜集的所有数据称为研究的数据集（data set）。表 1-1 是一个包含某电子商务平台某日 10 笔订单信息的数据集。

表 1-1 某电子商务平台某日 10 笔订单信息的数据集

订单编号	订购日期	客户 ID	产品	单价/元	数量/个	金额/元
119000001	2011/9/1	119970	产品 A	100	6	600
119000002	2011/9/1	128236	产品 D	300	6	1800
119000003	2011/9/1	130887	产品 B	200	10	2000
119000004	2011/9/1	132368	产品 B	200	7	1400
119000005	2011/9/1	133422	产品 A	100	10	1000
119000006	2011/9/1	137401	产品 D	300	9	2700
119000007	2011/9/1	138462	产品 B	200	8	1600
119000008	2011/9/1	146853	产品 A	100	1	100
119000009	2011/9/1	149841	产品 B	200	9	1800
119000010	2011/9/1	151129	产品 A	100	3	300

1.2.1 个体、变量和观测值

个体（element）是指搜集数据的实体。在表 1-1 所示的数据中，每一笔订单是一个个体，订单编号列为表 1-1 的第 1 列。10 笔订单对应数据集中的 10 个个体。

变量（variable）是一个个体的任意“特征”。表 1-1 所示的数据集中有下列 7 个变量：订单编号、订购日期、客户 ID、产品、单价、数量和金额。通常，我们需要在数据分析报告中对数据变量进行说明。

在一项研究中，对每个个体的每一变量收集测量值，从而得到了数据。对某一特定个体得到的测量值集合称为一个观测值（observation）。如表 1-1 所示，我们看到第一个观测值（119000001）的测量值集合是 2011/9/1、119970、产品 A、100、6 和 600。第二个观测值（119000002）的测量值集合是 2011/9/1、128236、产品 D、300、6 和 1800。10 个个体的数据集有 10 个观测值。

1.2.2 变量的基本分类

类别变量（categorical variable）是取值为事物属性或类别及区间值的变量，也称分类变量（classified variable）或定性变量（qualitative variable）。例如，观测人的性别、公司所属的行业、客户对商品的评价时，得到的结果就不是数字，而是事物的属性，如观测性别的结果是“男”或“女”，公司所属的行业为“制造业”“零售业”“旅游业”等，客户对商品的评价为“很好”“好”“一般”“差”“很差”。人的性别、公司所属的行业、客户对商品的评价等作为变量取的值不是数值，而是事物的属性或事物的类别。此外，学生月生活费支出的档次可能分为 1000

元以下、1000~1500 元、1500~2000 元、2000 元以上 4 档，作为变量的“学生月生活费支出的档次”这 4 档取值也不是普通的数值，而是数值区间，因而类别变量也称为区间值类别变量。人的性别、公司所属的行业、客户对商品的评价、学生月生活费支出的档次等都是类别变量。

类别变量根据取值是否有序通常分为两种：名义（nominal）值类别变量和顺序（ordinal）值类别变量。名义值类别变量也称无序类别变量，其取值是不可以排序的。例如，“公司所属的行业”这一变量取值为“制造业”“零售业”“旅游业”等，这些取值之间不存在顺序关系。再如，“商品的产地”这一变量的取值为甲、乙、丙、丁，这些取值之间也不存在顺序关系。顺序值类别变量也称有序类别变量，其取值可以排序。例如，“客户对商品的评价”这一变量的取值为很好、好、一般、差、很差，这 5 个值之间是有序的。取区间值的变量当然是有序类别变量。只取两个值的类别变量称为二值（binary）类别变量，例如，“人的性别”这一变量取值为男和女。二值类别变量可以看成名义值类别变量，也可以看成顺序值类别变量。

类别变量的观测结果称为类别数据（categorical data）。类别数据也称为分类数据或定性数据。与类别变量相对应，类别数据相应分为名义值类别数据和顺序值类别数据两种。其中，只取两个值的类别数据称为二值类别数据。

数值变量（metric variable）是取值为数字的变量，也称为定量变量（quantitative variable）。例如，“企业的销售额”“上涨股票的数量”“生活费支出”“掷一枚骰子出现的点数”等变量的取值可以用数字来表示，都属于数值变量。数值变量的观测结果称为数值数据（metric data）或定量数据。

数值变量根据其取值的不同，可以分为离散变量（discrete variable）和连续变量（continuous variable）。离散变量是只能取有限个数的变量，而且其取值可以一一列举，如“企业数”“产品数量”等。连续变量是可以在一个或多个区间中取任何值的变量，它的取值是连续的，不能一一列举，如“温度”“零件尺寸的误差”等。当离散变量的取值很多时，也可以将离散变量当作连续变量来处理。

变量（或数据）的基本分类如图 1-1 所示。

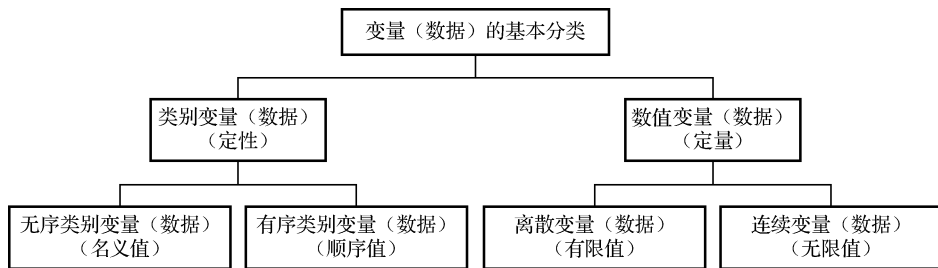


图 1-1 变量（或数据）的基本分类

1.2.3 数据的时代观

人们都说今天是数据的时代，到处都在讨论大数据，每个人都说自己在研究大数据，到处都宣称数据可以产生价值，但是，到底什么是数据？此处提出一个新的概念，即数据的时代观。

什么是数据？这个看似简单的问题却不易回答。如尝试向不同的人请教，常见的答案有两个：一是数据就是信息。这个答案完全正确但是定义太抽象了。数据和信息都是非常抽象的概念。两者相互定义，并不令人满意。二是数据就是数字。这种说法有一定的道理，因为数字是

一种最典型的传统数据。例如，GDP、股市的指数及人的身高、体重、血压等都是数字，也都是数据。因此，我们可以得出数字就是数据的结论。但是反过来，数据就是数字吗？未必。作者认为凡是可以电子化记录的信息其实都是数据。这里的记录不是靠自然人的大脑，而是通过必要的信息化技术和电子化手段。基于此，数据的范畴就大得多了，远不局限于数字。既然涉及电子化记录，下面简要介绍一下记录数据的技术手段。手机、数码照相机、各种工程设备上的探头等都是记录数据的技术手段。这些手段是有时代特征的，在不同时代，记录数据的技术手段是不一样的。这就是所谓的数据时代观。

问：声音是数据吗？

很久以前，声音并不是数据。因为当时没有任何技术手段能够把它记录下来。既然不能记录下来，更谈不上分析，也就不能称之为数据。但是今天，音频设备可以采集声音，然后将其转化为音频数字信号，进而支撑很多有趣的应用，如 iPhone 的 Siri、搜狗的语音输入法、微信的语音翻译，等等。由此可见，在可以记录声音的时代，声音是一种数据，而且是一种具有强烈时代特征的数据。

问：图像是数据吗？

很久以前，图像也不是数据，因为图像不能记录下来。那时图像只能是人们肉眼中看到的这个大千世界，如此美妙！但遗憾的是，图像没法记录。如今，数码成像技术的成熟让所有的图像都能够记录下来，而且分辨率非常高。在此基础上，人们可以做进一步的分析和建模，进而支撑很多有趣的应用。例如，脸部识别、指纹识别、车牌号识别、美图秀秀，以及医学中大量的医学影像分析。由此可见，在可以记录图像的时代，图像是一种数据，而且是一种具有强烈时代特征的数据。

类似的例子还有很多。例如，生物信息技术的进步产生了 Microarray 数据，社交网络的兴起产生了社交链数据，物联网技术的成熟产生了物联网数据。所有这些电子化的记录，都是数据。所有这些数据的产生都依赖于一定的技术手段，都有强烈的时代特征。

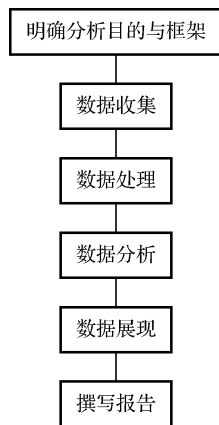


图 1-2 数据分析流程

1.3 数据分析流程

商务数据分析是基于商业目的，有目的地收集、整理、加工和分析数据，提炼有价值信息的过程。图 1-2 展示了数据分析的六大环节，主要包括明确分析目的与框架、数据收集、数据处理、数据分析、数据展现和撰写报告。

1. 明确分析目的与框架

针对数据分析项目，首先要明确数据对象是谁、目的是什么、要解决什么业务问题，然后基于商业的理解，整理分析框架和分析思路。常见的分析目的有减少新客户的流失、优化活动效果、提高客户响应率等。不同项目对数据的要求不同，使用的分析手段也不同。

2. 数据收集

数据收集是按照确定的数据分析和框架内容，有目的地收集、整合相关数据的过程，它是数据分析的基础。数据收集渠道包括内部渠道和外部渠道两类。内部渠道包括企业内部数据库、内部人员、客户调查，以及专家与客户访谈；外部渠道包括网络、书籍、统计部门、行业协会、展会、专

业调研机构等。常见的数据收集方法包括观察和提问、客户访谈、问卷调查、集体讨论、工具软件等。不同类型的数据在收集方式上区别较大，相关内容详见第2章，这里就不展开介绍了。

3. 数据处理

数据处理是指对所收集到的数据进行加工、整理，以便开展数据分析，它是数据分析前必不可少的环节。这个过程在整个数据分析过程中最占时间，在一定程度上取决于数据仓库的搭建和数据质量的保证。数据处理方法主要包括数据清洗和转化等。数据清洗和转化的主要对象包括残缺数据、错误数据和重复数据。

4. 数据分析

数据分析是指通过分析手段、方法和技巧对准备好的数据进行探索、分析，从中发现因果关系、内部联系和业务规律，为企业提供决策参考。到了这个环节，要能驾驭数据、开展数据分析，就要涉及工具和方法的使用。一是要熟悉常规数据分析方法，最基本的要了解如方差、回归、因子、聚类、分类、时间序列等数据分析方法的原理、使用范围、优缺点和结果的解释；二是熟悉数据分析工具，如 Excel、SPSS、Python、R 语言等，以便于进行一些专业的统计分析、数据建模等。

5. 数据展现

俗话说“字不如表，表不如图”。一般情况下，数据分析的结果是通过图、表的方式来呈现的。借助数据展现可视化工具，数据分析师能更直观地表达想呈现的信息、观点和建议。常用的图包括饼图、柱形图、条形图、折线图、散点图、雷达图、金字塔图、矩阵图、漏斗图、帕雷托图等。

6. 撰写报告

最后一个环节是撰写数据分析报告，这是对整个数据分析成果的一个呈现。分析报告可以把数据分析的目的、过程、结果及方案完整地呈现出来，以供企业参考。一份好的数据分析报告，首先需要有一个好的分析框架，并且图文并茂、层次明晰，能够让读者一目了然。结构清晰、主次分明，可以使读者正确理解报告内容；图文并茂，可以令数据更加生动活泼，提高视觉冲击力，有助于读者更形象、更直观地看清楚问题和结论，从而产生思考。另外，数据分析报告需要有明确的结论、建议和解决方案，不仅仅要找出问题，解决问题才是最重要的，否则称不上是好的分析，同时也会失去报告的意义，数据分析的初衷就是为了实现一个商业目的，不能舍本逐末。

1.4 电子商务典型数据分析任务

1.4.1 行业分析

行业分析通常由营销、运营岗位完成，该岗位设置在营销部、运营部，与数据开发部、公司战略管理部等均有配合及合作。行业分析流程包括行业数据采集、市场需求调研、产业链分析、细分市场分析、市场生命周期分析、行业竞争分析，最后撰写市场分析报告。

1. 行业数据采集

根据行业特性确定数据指标筛选范围，做出符合业务要求的数据报表模板；整合该行业数据资源，使用合适的方式采集数据并完成数据报表的制作。

2. 市场需求调研

通过客户行为、行业特性及业务目标要求设计出调研问卷；通过网络调研、深度访谈等方法发放与回收调研问卷；通过 Excel 等数据处理工具对回收问卷进行数据清洗，得到可靠的样本数据。

3. 产业链分析

通过对该行业中供应商、制造商、经销商、客户等环节之间交互关系的分析，绘制交互示意图；通过对前期的市场调研及价值交互关系的分析，制作产业链的合理性评估表。

4. 细分市场分析

根据细分市场历史数据确定相应的优势细分市场，编制优势细分市场列表；根据产品特点和消费者需求关联目标细分市场，编制关联列表；通过定性与定量的分析方法进行匹配度分析，编制匹配度对应列表。

5. 市场生命周期分析

根据市场历史数据判定出该细分市场所处的生命周期；通过行业资讯、领域专家意见，以及历史数据确定该细分市场所处生命周期中的机遇与挑战；根据细分市场所处生命周期给出改善建议。

6. 行业竞争分析

通过网络及纸质等渠道进行同类企业市场信息收集；进行同类企业与企业市场相关性与差异性的分析，编写市场差异性分析报告；通过 SWOT 分析法分析自身企业的机遇与挑战，编制 SWOT 分析图表。

1.4.2 客户分析

客户分析通常由客户运营岗位完成，该岗位设置在运营部，与市场部、品牌部、策划部、客服部、设计部、物流部等均有配合及合作。客户分析流程包括客户数据收集、客户特征分析（客户画像）、客户行为分析、客户价值评估、目标客户精准营销（营销策略制定和资源配置）、销售效果跟踪，最后撰写市场分析报告。

1. 客户数据收集

了解 B 端（企业端）及 C 端（消费者端）的不同客户数据收集；熟悉公司品牌及产品定位、客户定位，熟悉各业务部门的客户数据需求；根据客户的访问、浏览、购买、评价等行为数据对客户数据属性标签进行收集、整理；熟练运用 Excel、CRM、评价分析、舆情监控等客户数据收集分析工具（软件）；利用问卷、调研等数据收集方法收集客户数据，并对数据进行清洗和处理。

2. 客户特征分析（客户画像）

了解 B 端（企业端）及 C 端（消费者端）的客户行为属性区别；根据客户的购买行为、购买地域、购买金额、购买次数等行为对客户进行特征分析；熟悉地域、性别、年龄等客户基础属性，并进行相关归类分析；借助 Excel、CRM 等工具对客户特征进行挖掘分析及梳理。

3. 客户行为分析

对客户的评价行为、购买趋势、购买喜好、营销喜好、产品喜好等行为进行分析；根据客户行为数据分析制定不同渠道的内容模式，挖掘出客户接受度较高的营销方式。

4. 客户价值评估

了解 B 端（企业端）及 C 端（消费者端）的不同，分析客户价值行为；熟悉客户画像、回购率、客单价、地域等客户行为分析的概念和行为价值；了解各业务部门对客户数据的需求，基于需求挖掘客户价值，并进行相关价值评估；有一定的业务分析能力及文案撰写能力，对客户行为特征进行有价值的二次挖掘。

5. 目标客户精准营销（营销策略制定和资源配置）

熟悉 B2B 及 B2C 的平台区别，了解 B 端（企业端）及 C 端（消费者端）不同平台客户精准分析、营销策略及营销规则工具；熟悉 B2C 平台渠道下的（消费者）微博、微信、淘宝、京东等各平台的客户推广营销渠道及推广方法；掌握消费者心理，基于推广渠道，了解短信、电子邮件、自媒体、直播等营销渠道，制订各渠道的精准推广计划；根据制订的计划协调公司相关资源，最终完成营销计划的投放。

6. 销售效果跟踪

熟悉营销回购率、转化率、投资回报率（Return On Investment, ROI）等指标；对各渠道的客户营销数据进行总结、分析、对比，输出各渠道的效果报告，调整不同渠道的客户运营策略；跟踪各渠道的销售效果及 ROI，给各业务部门提出业务建议，并协助各渠道进行客户营销模式的调整。

1.4.3 产品分析

产品分析通常由产品、客服岗位完成，该岗位设置在产品部、运营部、客服部，与设计部、美工部、生产部等均有配合及合作。产品分析流程包括竞争对手分析、客户特征分析、产品需求分析、产品生命周期分析、客户体验分析，最后通过调研报告形成合理化建议，对产品开发及市场走向提出预测。

1. 竞争对手分析

分析目标客户、定价策略、市场占有率等，确定竞争对手；对竞争对手的价格、产品、渠道、促销等方面进行调研，归纳、整理调研数据；通过 SWOT 分析法，得出竞争对手产品及自身产品的优劣势。

2. 客户特征分析

根据研究目的，确定典型客户特征的分析内容；做好客户年龄、地域、消费能力、消费偏好等数据的收集与整理工作；通过 Excel 等工具分析客户数据，赋予不同的人群标签。

3. 产品需求分析

根据典型客户特征分析结果，收集客户对产品需求的偏好；通过整理分析客户需求偏好，提出产品开发的价格区间、功能卖点、产品创新、包装等建议，通过产品的不断升级和迭代，提高客户对产品 & 品牌持久的黏性。

4. 产品生命周期分析

利用 Excel 等工具汇总产品部、运营部、客服部等产品销售数据；密切监控季节、气温、地域等因素对产品销售周期性数据的变化及波动的影响；协助指导采购、生产等部门合理安排采购及生产计划。

5. 客户体验分析

通过客户访谈或工具软件收集，了解客户体验现状；跟踪和分析客户对产品的反馈，监测

产品使用状况并及时提出改进方案；识别客户痛点及产品机会，组织有价值的典型客户参与产品设计，评估产品价值及客户体验。

1.4.4 运营分析

运营分析通常由产品、客服岗位完成，该岗位设置在产品部、运营部、客服部，与设计部、美工部、生产部等均有配合及合作。运营分析流程包括销售数据分析、推广数据分析、客服数据分析。

1. 销售数据分析

通过评估历史销售数据等，进行企业销售目标的定位；通过市场调研，归纳、整理调研数据，设计销售指标；运用 Excel 等数据管理工具或调用平台数据，制定销售业绩、价格体系、区域布局、产品结构、销售业绩异动等指标；通过建立多维报表，明确销售任务；制作销售多维报表，得出整体销售分析指标；通过内部报告系统或数据采集工具，获取销售数据；通过与客服部门沟通，获取销售反馈信息；对数据进行整理和清洗，保证数据的有效性和完整性；对整体销售进行分析，包括销售额分析、销售量分析、季节性分析、产品结构分析、价格体系分析；对销售区域进行分析，包括区域分布分析、重点区域分析、区域销售异动分析；对产品线进行分析，包括产品系列结构分布分析、产品-区域分析；对价格体系进行分析，包括价格体系构成分析、价格-产品分析、价格-区域分析；根据既往数据进行预测，包括总体销售预测、区域销售预测、季节性销售变化预测；对电子商务平台特有的指标，如货品流失率、客单价、存销比等进行分析及预测；对数据可视化方案进行设计，结合业务场景设计出实用的可视化方案；应用可视化方案对已分析出的销售数据结果进行展现。

2. 推广数据分析

通过公司已有的商务推广数据及公司现状、商品维度、外部竞争数据等，确定数据分析的目标；根据数据分析目标及公司现有商务推广数据，制定分析原则和分析策略；根据数据分析目标、分析原则和分析策略，确定详细分析步骤及时间规划；根据整体规划划分阶段目标，通过 Excel 及 PPT 等分析汇报工具，进行分析方案规划；根据具体推广业务和推广方式，对数据进行合并或拆分操作，以便对数据进行分析；根据业务和分析工具，对数据进行标准化、归一化操作或对定性数据进行量化操作；根据现有推广数据，分析各种推广方式、推广渠道对不同人群的推广效果；对不同人群适合的推广方式和推广渠道，提出合理的推广建议；根据现有推广数据，分析各种推广方式、推广渠道的整体效果；对分析出的各种推广渠道整体效果进行图表化展现；对分析出的各种推广方式、推广渠道、面向人群互相结合的效果进行图表化展现。

3. 客服数据分析

根据企业目标、运营过程、历史数据、企业环境，进行分析目标设计；通过调研企业领导及各层次人员，收集历史数据及递增幅度，定义成本、人员留存、销售成功率、人均销售收入、数据可用率、销售成功营销率等指标，以达成提升运营质量、降低成本、开展精准营销等企业目标；将具体的问题抽象成指标，以达成特定目标；通过收集基本数据，计算成本、人员留存、销售成功率、人均销售收入、数据可用率、数据的及时率、销售成功率等指标；通过数据分析工具，分析转化率、响应时间、销售额等指标；分析售前、售中、售后指标；将数据结果以图表的方式展现给客户。

1.5 数据分析岗位的职业规划

1.5.1 职业前景

工资高：数据分析企业的高薪企业主要分布在长江三角洲、珠江三角洲和京津地区。据统计，2018年北京、上海、深圳的就业者月薪在15000元以上，实习期薪资是8000元起步。

晋升快：1年晋升主管，2年晋升经理，3年晋升核心管理团队。

前景好：中国最火的TMT行业，数据将成为企业的核心竞争力，是21世纪最有发展前景的工作。

缺口大：每年新增市场缺口300万人（传统企业“触电”的需要），75%的高校新增数据类专业。

1.5.2 职业素养及成长阶段建议指南

数据分析岗位大体分为3个级别，具体内容如下。

1. 初级分析师

企业需求：初级分析师是数据架构的基础，承担了大多数最基础的工作。通常初级分析师的人员比例不应超过20%。

职能定位：初级分析师的定位是从事数据整理、数据统计和基本数据输出工作，服务的对象包括中、高级分析师和业务方等。

职业要求：

- 数据工具方面，具有基本的Excel操作能力和SQL取数能力；具有基本数据输出能力，包括PowerPoint、电子邮件、Word等软件的使用能力。
- 数据知识方面，理解日常数据体系内涉及的维度、指标、模型。
- 业务知识方面，理解日常业务知识，能把业务场景和业务需求用数据转换和表达出来。

2. 中级分析师

企业需求：中级分析师是数据人员架构中的主干，承担着公司中的专项数据分析工作，通常人员比例为40%~60%。

职能定位：中级分析师的定位是从事数据价值挖掘、提炼和数据沟通落地，服务的对象主要是业务方，除此之外，还可能参与高级分析师的大型项目并独立承担其中的某个环节。

职业要求：

- 数据工具方面，熟练使用数据挖掘工具、网站分析工具。
- 数据知识方面，了解不同算法和模型的差异及最佳实践场景，能根据工作需求应用最佳实践方案。
- 业务知识方面，对业务知识深度理解，有较强的数据解读和应用推动能力。

3. 高级分析师

企业需求：高级分析师及以上职位通常是数据人力架构中的“火车头”，承担着企业数据方向的领导职能，人员比例为20%~40%。

职能定位：高级分析师的定位是进行企业数据工作方向规则体系建设、流程建设、制度建

设等，服务对象通常是企业领导层。

职业要求：

- 能搭建起企业数据体系，并根据企业发展阶段提出适合当前需求的数据职能和技术架构方案。
- 规划出所负责领域内数据工作方向、内容、排期、投入、产出等，并做投入、产出分析和数据风险管理。
- 实时跟进项目进度和落地，并通过会议、汇报、总结、阶段性目标、KPI 等形式做好过程控制和结果控制。

1.5.3 技术类数据管理岗位介绍

技术类数据管理岗位如下。

1. 数据工程师

定位：数据仓库开发、数据库实施，为分析应用及分析系统开发等提供技术支持。

应用技术：

- Oracle、DB2、MySQL、SQL、Linux。
- Java、Python、R、XML、JSON。
- Hadoop、Redis、HBase、ZooKeeper。

岗位要求：

- 掌握数据存储技术的基本概念、原理、方法和技术。
- 了解数据应用系统的生命周期及其设计、开发过程。
- 理解数据分析的基本业务知识、业务场景，并开发、实施具体项目。

2. 算法工程师

定位：针对特定的数据分析业务需求，提供最合适的算法进行数据计算。

应用技术：

- SPSS、Clementine/SAS、EM。
- 回归模型、决策树、协同过滤、聚类算法。
- Mahout、机器学习、自然语言处理。

岗位要求：

- 根据自己对行业及企业业务的了解，独自承担复杂的分析任务，并形成分析报告。
- 相关分析方向包括客户行为分析、广告点击分析、业务逻辑相关分析及竞争环境相关分析。
- 根据业务逻辑变化，设计相应分析模型并支持业务分析工作开展。

3. 架构师

定位：将数据需求转换为规范的开发计划及文本，并制定项目的总体架构，指导整个开发团队实现需求。

应用技术：

- 大型数据库、分布式数据库、内存数据库。
- 各种开发语言（了解其优劣、算法及其用处）。
- EJB 设计模式、J2EE 架构、UDDI、软件设计模式。

- PowerPoint、Power Designer、Mind Manager、Visio。

岗位要求:

- 提供技术支持,在关键时刻攻克最艰巨的技术壁垒。
- 了解业务需求,提出架构方案并搭建技术体系。
- 转化业务需求为技术需求,并确定最终技术架构方案。
- 主导系统全局分析和实施,负责软件架构和关键技术,并协助进行核心技术开发。

思考与实训

亚马逊的数据化经营之路

悉数全球哪家公司从大数据中发掘出了最大价值,一定非亚马逊公司莫属。作为一家“信息公司”,亚马逊公司不仅从每个客户的购买行为中获得信息,而且将每个客户在其网站上的所有行为都记录下来,包括页面停留时间、客户是否查看过评论、每个搜索的关键词、浏览的商品等。这种对数据价值的高度敏感和重视,以及较强的数据分析能力,使得亚马逊公司早已远远超出了同行业采用传统运营方式的公司。

- 亚马逊推荐:亚马逊公司的各个业务环节都离不开“数据驱动”的身影。在亚马逊网站上买过产品的客户可能对它的推荐功能很熟悉,“买过X产品的人,也同时买过Y产品”的推荐功能虽看上去很简单,却非常有效。需要说明的是,这些精准推荐结果的得出过程非常复杂。
- 亚马逊预测:客户需求预测是通过历史数据来预测客户未来的需求。对于书籍、手机、家电这些亚马逊公司内部称为硬需求的产品,预测是比较准的,甚至可以预测出相关产品属性的需求。但是对于服装类等软需求的产品,预测相对复杂和困难。
- 亚马逊测试:亚马逊网站上的某段页面文字只是碰巧出现的吗?其实,亚马逊公司会在网站上持续不断地测试新的设计方案,从而找出转化率最高的那一个。整个网站的布局、字体大小、颜色、按钮,以及其他所有的设计,其实都是在多次审慎测试后得出的最优结果。
- 亚马逊记录:亚马逊公司的移动应用在让客户有一个流畅的、无处不在的体验的同时,也通过收集手机上的数据深入地了解每个客户的喜好信息;更值得一提的是,KindleFire内嵌的Silk浏览器可以将客户的行为数据一一记录下来。

亚马逊公司的数据分析不仅限于以上领域。对亚马逊公司来说,数据是商业决策的指挥棒,数据可显示出什么是有效的、什么是无效的,新的商业投资项目必须要有数据的支撑。

活动组织:

- 1) 教师要求学生认真阅读案例素材。
- 2) 借鉴以上案例,学生形成学习小组,每组确定一家企业,通过搜索工具查找相关资料,结合案例素材,分析讨论此企业在采集数据时所使用的方法及分析数据时所使用的工具;讨论此企业是如何将数据分析结果运用在供应链管理、销售管理等经营活动中的;结合此企业的实例,描述数据分析在传统零售型企业商务活动中的实际意义。
- 3) 邀请一组学生在投影仪上展示自己的分析报告,全体同学互相交流。