

新手学

数据分析 (入门篇)

杨群 编著

快速了解数据分析流程，新手学数据分析入门必读。
深度剖析数据分析精华，一本书教你玩转数据分析。

数据
分析
关键点

零基础学数据分析快速上路

59

数据
分析
关键点

深入剖析数据分析方法与技术的精髓

55

典型
案例
解析

助新手快速精通数据分析实战应用之道

108

清华大学出版社



大众创业系列丛书

新手学 数据分析 (入门篇)

杨群 编著

常州大学图书馆
藏书章

清华大学出版社
北京

内 容 简 介

本书共分为10章,内容涉及数据分析工作的各个方面,循序渐进地讲解了数据分析工作的开展流程、技术以及各种注意事项,其主要内容有:认识数据分析、了解数据分析方法论和数据分析方法,数据源的获取、加工、处理、分析、呈现以及最终的报告撰写等。

全书内容详细全面,且注重实践操作,尤其对于想要从事数据分析工作的读者,借助本书可实现快速入门,此外本书也非常适合从事数据分析相关工作不久的新人巩固知识和提升相关技能。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

新手学数据分析:入门篇/杨群编著. —北京:清华大学出版社,2018

(大众创业系列丛书)

ISBN 978-7-302-48658-9

I. ①新… II. ①杨… III. ①数据处理 IV. ①TP274

中国版本图书馆CIP数据核字(2017)第266029号

责任编辑:李玉萍

封面设计:郑国强

责任校对:吴春华

责任印制:沈 露

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦A座 邮 编: 100084

社 总 机: 010-62770175 邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者: 三河市君旺印务有限公司

经 销: 全国新华书店

开 本: 170mm×240mm 印 张: 18.25 字 数: 310千字

版 次: 2018年1月第1版 印 次: 2018年1月第1次印刷

定 价: 45.00 元

关于本书

在这个信息爆炸的时代，我们每天都在与数据打交道，而且这些数据都在直接或间接地影响着我们的生活和工作，特别是在市场现状的调查、行业发展趋势的预测、公司运营决策的制定、生产数据的预测、新产品的研发等领域，更加需要基于科学、严谨的数据分析得到的数据结果，来为决策者制定决策提供可靠的数据来源。

随着大数据时代的到来和发展，数据分析从业人员的需求量也在不断增大，虽然许多人想要从事数据分析工作，但是碍于数据分析技术的繁多与学习难度，最终都止步于这个大门外。有一部分人即使已经涉足或者从事了数据分析工作，但是对于整个数据分析的过程或许还存在迷惑的地方。

其实数据分析并没有那么难，只要依照最初的数据分析目的，运用正确的数据分析方法论和数据分析方法，逐步进行就可以完成数据分析工作。虽然数据分析的方法和技术很多，但是 Excel 作为强大的数据分析工具，几乎可以完成大部分的数据分析任务。对于初、中级的数据分析师而言，已经足够。

让更多的人了解什么是数据分析，并且快速入门，我们依照数据分析的整个过程和思路，精心编写了本书。

特点	说明
典型案例快速精通	本书在创作过程中侧重于实践方面的讲述，摒弃“假、大、空”的套话，并提供了许多典型案例和实操内容，通过案例分析和讲解辅助读者了解数据分析技术的应用，让读者快速精通数据分析工作各个环节的重点和必会知识
全程图解步步详解	本书包含大量的图片、表格和图示，步步详解各种数据分析方法和数据分析技术的实战应用，以帮助读者更快且更熟练地掌握数据分析技术

本书结构

本书作为一本数据分析入门的实用工具书，按照数据获取→数据处理→数据分析→数据呈现→数据报告的线索，为读者详细描绘了数据分析的完整流程以及整个数据分析工作中涉及的各种方法、方法论以及 Excel 技术和注意事项。全书共分为 10 章，主要内容可分为以下 3 个部分。

新手学数据分析（入门篇）

章节安排	主要内容	作用
第 1 ～ 3 章	本部分讲解数据分析概述、数据分析行业发展、数据分析人才的培养、认识商业数据分析师、了解数据分析的流程以及了解数据分析方法论和数据分析方法	这部分作为本书的开篇内容，为读者详细介绍了有关数据分析的基础知识，让读者对数据分析快速入门
第 4 ～ 7 章	本部分介绍数据分析过程中数据源的准备、加工处理、各种数据分析技术以及数据结果简单的呈现方式	这部分内容全面介绍了整个数据分析工作中的实战操作，掌握这些内容可以完成最基本的数据分析核心过程
第 8 ～ 10 章	本部分介绍如何用透视功能查看数据分析结果、用更专业的图表展示数据以及数据分析报告的撰写技巧和注意事项	这部分内容为本书的提升内容，通过对这部分内容的学习，可以让读者掌握数据分析结果的展示技巧以及数据分析报告的撰写技巧，从而更好、更完整地完成任务

本书读者

本书作为数据分析入门的实用书籍，能帮助想要涉足数据分析行业的读者快速入门，也可以帮助初涉数据分析工作的工作人员更专业地完成工作。此外，对于有过数据分析工作经验的初、中级数据分析师巩固和提升数据分析技术也有一定的指导作用。由于编者经验有限，书中难免会有疏漏和不足之处，恳请专家和读者不吝赐教。

本书作者

本书由杨群编著，参与本书编写的人员有邱超群、罗浩、林菊芳、马英、邱银春、罗丹丹、刘畅、林晓军、周磊、蒋明熙、甘林圣、丁颖、蒋杰、何超等，在此对大家的辛勤工作表示衷心的感谢！由于编者经验有限，书中难免会有疏漏和不足之处，恳请专家和读者不吝赐教。

第 1 章 全面了解数据分析行业 001

1.1 数据分析概述 002

1.1.1 认识数据分析及其分类 002

1.1.2 数据分析的重要性 003

1.2 初步了解数据分析行业 ... 004

1.2.1 数据分析行业的发展历程 004

1.2.2 充分认识大数据时代 006

1.2.3 我国大数据产业存在挑战 011

1.3 数据分析人才的培养 012

1.3.1 大数据时代需要的人才 012

1.3.2 数据分析人才需要具备的能力 013

1.3.3 成为数据分析人才必备的素质 015

1.4 认识数据分析职位 016

1.4.1 数据分析的职位体系 016

1.4.2 数据分析师的工作内容 017

1.4.3 常见数据分析职位的技能要求 018

第 2 章 深入认识数据分析 019

2.1 充分理解数据 020

2.1.1 了解数据形成过程与数据处理 020

2.1.2 理解字段、记录和数据表 021

2.1.3 认识 Excel 处理的数据类型 022

2.2 掌握数据分析的流程 024

2.2.1 第一步：明确数据分析的目的和思路 024

2.2.2 第二步：获取需要分析的数据 025

2.2.3 第三步：对收集的数据进行处理 029

2.2.4 第四步：分析数据以获得有用信息 030

2.2.5 第五步：选择合适的呈现方式 031

2.2.6 第六步：撰写数据分析结果报告 033

2.3 认识数据分析的误区 033

2.4 了解基本的数据分析指标 ... 035

2.4.1 平均数指标 036

2.4.2 频数与频率指标 ... 036

2.4.3 绝对数与相对数指标 037

2.4.4 其他常见数据 分析指标	039
----------------------------	-----

第3章 数据分析方法论和 数据分析方法 ... 041

3.1 数据分析方法论和数据分析 方法概述	042
3.1.1 了解数据分析 方法论	042
3.1.2 了解数据分析方法 ...	043
3.2 经典数据分析方法论详解 ...	044
3.2.1 4P 营销理论：分析公司 整体营运情况	044
3.2.2 用户使用行为理论： 分析用户行为	047
3.2.3 PEST 分析法：分析 宏观环境	048
3.2.4 逻辑树分析法：分析 专项业务问题	050
3.2.5 5W2H 分析法：分析 任何问题	051
3.3 常见的数据分析法模型 ...	054
3.3.1 对比分析法	054
3.3.2 分组分析法	055
3.3.3 交叉分析法	057
3.3.4 综合评价分析法 ...	058

第4章 准备数据是数据 分析的第一步 ... 069

4.1 直接获取外部数据源	070
4.1.1 导入文本文件数据 ...	070
4.1.2 导入 Access 数据 ...	072
4.1.3 导入网站数据	073
4.1.4 导入 SQL Server 数据	075
4.1.5 导入 XML 数据 ...	076
4.2 手工录入数据的方法	077
4.2.1 快速录入表格数据 的技巧	077
4.2.2 特殊数据的 输入方法	081
4.3 问卷调查数据的录入要求 ...	085
4.4 手动整理数据要快而准 ...	091
4.4.1 数据来源的 有效性设置	091
4.4.2 数据的编辑与修改 ...	096
4.4.3 数据的批量修改 ...	098
4.5 优化待分析的数据 显示效果	100
4.5.1 利用字体格式 提升专业性	100

- 4.5.2 表格效果的
优化操作 102
- 4.5.3 格式化设置中的
颜色使用原则 107

第 5 章 加工处理数据源 是数据分析的 关键 109

- 5.1 正确理解数据的加工处理 ... 110
 - 5.1.1 数据处理的要求 ... 110
 - 5.1.2 数据处理的步骤 ... 112
- 5.2 数据处理的必备基础知识 ... 114
 - 5.2.1 公式和函数基础 ... 114
 - 5.2.2 使用公式与函数
的方法 116
- 5.3 对数据进行清理与检查 ... 121
 - 5.3.1 处理数据源中的
重复数据 121
 - 5.3.2 检查数据的完整性 ... 125
- 5.4 对数据源进行二次加工 ... 127
 - 5.4.1 在数据源中
抽取数据 127
 - 5.4.2 计算需要的
数据结果 132

第 6 章 利用工具快速 分析数据 135

- 6.1 利用透视功能分析数据 ... 136

- 6.1.1 创建数据
透视表的方法 136
- 6.1.2 合理地设计透视表的
布局和格式 139
- 6.1.3 更改数据透视的
汇总方式 143
- 6.1.4 刷新数据透视表
中的数据 145
- 6.1.5 在数据透视表中
使用计算字段 146
- 6.1.6 使用切片器
分析数据 147

6.2 Excel 数据分析 工具库的应用 150

- 6.2.1 加载 Excel
分析工具库 150
- 6.2.2 数据的描述性
统计分析 152
- 6.2.3 数据的抽样分析 ... 154
- 6.2.4 数据的回归分析 ... 156
- 6.2.5 数据的相关性分析 ... 160
- 6.2.6 数据的假设
检验分析 162
- 6.2.7 数据的预测分析 ... 165

第 7 章 数据结果的简单 呈现方式 169

- 7.1 使用条件格式展示
分析结果 170

- 7.1.1 条件格式在数据分析情况下使用的场合 170
- 7.1.2 用填充色突出显示某个范围的数据 ... 172
- 7.1.3 将前 X% 的数据显示出来 174
- 7.1.4 用图形比较数据大小 175
- 7.1.5 根据关键字将对应的记录突出显示 ... 177
- 7.2 使用迷你图在单元格中分析数据 179
 - 7.2.1 创建迷你图的方法 ... 179
 - 7.2.2 更改迷你图的类型 ... 180
 - 7.2.3 设置迷你图的外观效果 181

第 8 章 透过图表直观查看数据分析结果 ... 183

- 8.1 揭开图表的神秘面纱 184
 - 8.1.1 用图表展示数据的意义 184
 - 8.1.2 掌握图表与数据之间存在的关系 ... 186
 - 8.1.3 数据演变成图表的 5 个阶段 188
 - 8.1.4 了解图表的基本组成部分 189

- 8.2 利用图表展现数据的必会操作 191
 - 8.2.1 创建一个完整图表的步骤 191
 - 8.2.2 图表数据的编辑 ... 197
 - 8.2.3 图表元素的设置 ... 203
- 8.3 优化图表的技巧 209
 - 8.3.1 用图片让数据分析呈现更形象 210
 - 8.3.2 直观区分图表中的正负数 212
 - 8.3.3 断裂折线图的处理方法 215
 - 8.3.4 自动显示图表中的最值数据 217
- 8.4 数据分析中的特殊图表制作 220
 - 8.4.1 制作甘特图 220
 - 8.4.2 制作对称条形图 ... 226

第 9 章 更专业地用图表展示数据 231

- 9.1 根据需要处理细节数据 ... 232
 - 9.1.1 在图表下方添加数据来源 232
 - 9.1.2 处理图表中的冗余数据 235
 - 9.1.3 使用脚注添加说明 ... 237

9.1.4 将数据大的图形 截断展示	238	9.3.5 巧妙处理饼图中的 较小扇区	260
9.1.5 处理数值坐标轴 中的符号	240	第 10 章 最后一步： 撰写数据分析 报告	263
9.2 图表的美化原则	242	10.1 数据分析报告概述	264
9.2.1 图表各组成部分 的文字使用要协调 ...	242	10.1.1 数据分析报告 快速入门	264
9.2.2 不要为了好看而让 图表变得花哨	245	10.1.2 了解数据分析报告的 种类	266
9.2.3 关键数据要突出 显示出来	248	10.1.3 制作数据分析 报告的工具	269
9.2.4 慎用三维立体效果 ...	253	10.1.4 数据分析 报告的生成	271
9.3 常见图表类型的 规范制作要求	255	10.2 数据分析报告的组成 ...	273
9.3.1 柱形图的分类和 数据系列不要太多 ...	255	10.2.1 数据分析 报告的开篇	274
9.3.2 分类标签多而 长首选条形图	257	10.2.2 数据分析 报告的正文	278
9.3.3 排序数据源使 条形图数据展示 更直观	258	10.2.3 数据分析报告 的结尾	281
9.3.4 多折线的情况下 分开做多个图表 ...	259		

第 1 章

全面了解数据分析行业



本章要点

- ◆ 认识数据分析及其分类
- ◆ 数据分析的重要性
- ◆ 数据分析行业的发展历程
- ◆ 充分认识大数据时代
- ◆ 我国大数据产业存在挑战
- ◆ 大数据时代需要的人才
- ◆ 数据分析人才需要具备的能力
- ◆ 成为数据分析人才必备的素质
- ◆ 数据分析的职位体系
- ◆ 数据分析师的工作内容
- ◆ 常见数据分析职位的技能要求



学习目标

在大数据时代，随处都可以听到“数据分析”这个词，什么是数据分析，数据分析的价值是什么，如果要做好数据分析，需要具备什么样的条件……通过阅读本章，你将对数据分析有一个新的认识，对数据分析行业有更全面的了解。

知识要点	学习时间	学习难度
数据分析概述	20 分钟	★★
初步了解数据分析行业	40 分钟	★★★★
数据分析人才的培养	15 分钟	★
认识数据分析职位	25 分钟	★★★

1.1

数据分析概述

数据分析往往会让人联想到 Excel，的确，利用 Excel 可以进行数据分析，但是不能说数据分析就是用 Excel 处理数据。那么什么是数据分析呢？首先来看看数据分析的价值和作用。

1.1.1 认识数据分析及其分类

随着大数据概念的普及，越来越多的人意识到数据分析对经济发展的重要意义。什么是数据分析？通俗地讲，数据分析就是数据加分析，即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。专业地讲，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分析，从中提取有用的信息并形成结论，进而对数据加以详细研究和概括总结。

在统计学领域中，数据分析被划分为 3 种类型，分别是描述性数据分析、探索性数据分析以及验证性数据分析，各种分析类型的介绍如图 1-1 所示。

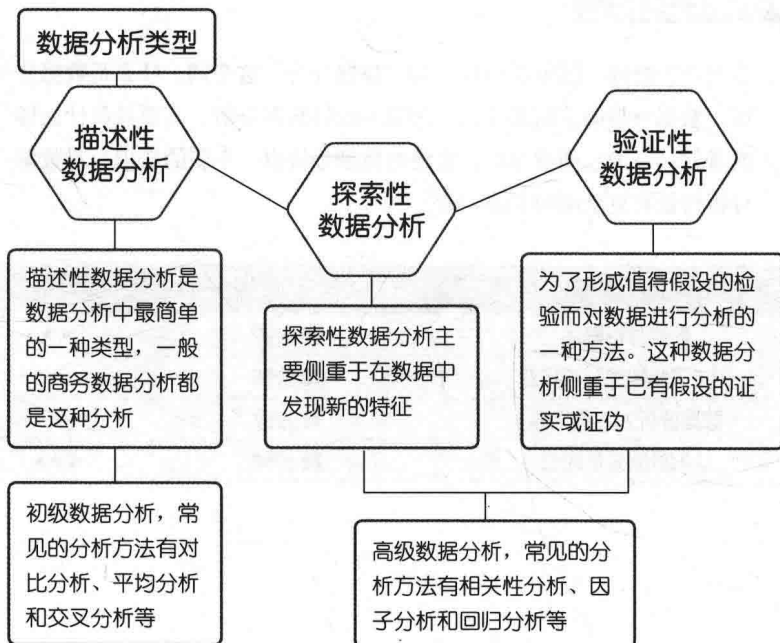


图 1-1

1.1.2 数据分析的重要性

常言道：数据是最有说服力的。也有人说：“不以数据分析为基础的执行都是耍流氓”，虽然这句话听起来不雅，但是道理是很明确的。下面来看两个小案例。

案例陈述

【例 1】

在 20 世纪 90 年代，美国沃尔玛超市的管理人员在分析销售数据时发现了一个令人难以理解的现象：在某些特定的情况下，“啤酒”与“尿布”这两件看上去毫无联系的商品却经常出现在同一个购物单中，后来经过调查发现，这种现象主要发生在年轻的爸爸身上。

因为在美国有婴儿的家庭中，妈妈通常是在家中照顾婴儿，而买尿布的事情则落在了年轻的爸爸身上，所以当爸爸去超市购买尿布时都会顺便为自己购买啤酒，所以看上去毫无关系的两件商品，在这种年轻爸爸这个群体的存在下发生了必然的联系。针对这种现象试想，如果爸爸们在超市中只能买到这两件商品中的一件，则很有可能他会放弃在此处购物而另寻他处。

于是，沃尔玛超市的管理人员尝试在卖场将啤酒与尿布摆放在相同的区域中，让年轻的爸爸可以很方便地同时找到这两件商品，并完成购物，超市对这种调整获得的最直接的益处就是提升了商品的销售收入。

【例 2】

一个人围绕刷牙发生的数据有以下一些：平均一年买几次牙膏，喜欢什么口味，什么价位，什么品牌，什么年龄段……这些数据看似平凡简单，但是这些数据对销售牙膏的人来说有很大的帮助。例如，“平均一年买几次牙膏”这项数据，通过对这项数据进行分析，可以了解到人们使用牙膏的周期，从而可以指导商场或者超市的采购人员配合使用周期来确保牙膏的有效供应。又如，对“口味”“价位”和“品牌”这些数据进行分析，可以帮助销售者判断出客户的需求、消费水平以及对品牌的信任度等。这样的分析结果对采购者采购商品时有很大帮助。

虽然这两个例子很小，但是透过这两个案例得到的结论是一样的，无论是企业的发展还是整个行业的发展，数据分析都是运营的核心。在现代这个网络发达的信息社会，各行各业对信息资料的需求已经达到了垂涎三尺的地步，因此，很多的媒体机构和行业机构每年都会对数以千万的数据进行分析和报道。因为大家都知道，数据分析的终极核

心价值在于资源优化配置。

下面具体来看看数据分析的重要性具体表现在哪些方面。

(1) **能客观地对实际情况进行真实反映。**数据分析是在实事求是原则的指导下，对大量的数据进行加工处理和研究，最终做出科学的判断并编写成数据分析报告，因此该报告比一般的数据报表更能集中、全面和正确地反映客观情况。

(2) **企业运营重要的监督手段。**因为数据分析部门掌握了公司大量的统计数据及信息资料，因此能更全面和准确地对公司当前的运营情况和未来的发展进行把握，从而能较好地监督检查企业相关部门方针政策的执行以及各项任务和指标的完成情况。

(3) **实现管理科学化的有效手段。**可以更好地从量化的角度开展分析研究，让领导和有关部门客观、全面地认识该公司经济活动的历史、现状及其发展趋势，并以数据为基础来进行科学化的管理，并参与到重要决策的制定过程中。

(4) **有利于数据资料的深度开发。**数据分析部门通过普查、抽查和调查等多种形式来收集资料，并对数据进行加工整理和深入分析，而且对分析的结果还可以进行多层次的开发和利用，最终得到更多有用的信息，让公司的决策人更充分地透过数据看到本质，从而可以更好地做出科学、正确的决策。

(5) **有利于各部门更高效地完成工作。**庞大的数据库一般是杂乱无章的，尤其从市场中获得的调查数据，信息更是凌乱分散。仅仅从表面很难看出数据之间内在的联系，而利用数据分析可以让数据变得可视化和直接，不仅利于工作人员理解和管理，而且可以让各项工作进行得更加清晰和有条理。

初步了解数据分析行业

1.2

在国外，数据分析很早就被应用到各个领域，而且还成立了相应的行业组织，培养了专业的数据分析人员。在我国，数据分析行业起步相对较晚，下面就来了解一下国内数据分析行业的发展情况。

1.2.1 数据分析行业的发展历程

我国的数据分析行业从无到有，直至今天的不断发展壮大，主要经历了兴起期、

发展期、成型期和迅速发展壮大期这4个阶段。

1. 第一阶段：兴起期

2003年年底，根据国家财政部和国家发改委《关于规范长期投资项目数据分析方法及与国际接轨的总体精神》，信息产业部电子行业职业技能鉴定指导中心正式设立“项目数据分析师”考试培训认证项目，并制定我国项目数据分析师人才培养管理规则以及考核管理办法。

次年的1月1日，首批项目数据分析师在深圳诞生，这标志着项目数据分析师全国考试培训试点工作正式拉开序幕。随后，这项培训工作在深圳、北京、成都和沈阳试点后，又迅速推广到上海、广州、西安、天津、南京、南宁和厦门等城市开展。这就意味着“项目数据分析师”的人才培养战略在全国全面启动。

2. 第二阶段：发展期

2005年，全国首家项目数据分析师事务所通过国家工商局的正式批准设立，于2005年4月取得营业执照。此后，在西安、深圳、成都和北京等地的项目数据分析师事务所相继诞生。项目数据分析师事务所的出现，不仅是我国数据分析行业的一个里程碑，更是我国数据分析行业开始进入不断发展新时期的标志。

3. 第三阶段：成型期

2006年，全国各地政府和行业协会纷纷发文表示支持数据分析行业的发展，为项目数据分析人才的培养提供了大力帮助，且所有开考地区的省级报纸、杂志、电视台和网络等各种媒体对项目数据分析师进行了全面的报道，从此，“数据分析师”和“项目分析师”等岗位在各种人才招聘会上出现在大家的视野中。

2007年，数据分析师的课程体系得到进一步完善，全国的项目数据分析师学员人数已达到千人，而全国近10个省市组建了近40家专业的项目数据分析师事务所。同年7月，为了项目数据分析全国行业协会的筹备需要，电子行业职业技能职业鉴定指导中心正式下文批准成立全国“项目数据分析师专家委员会”。

在这两年的时间中，项目数据分析师职业和专业数据分析师事务所的出现，标志着数据分析行业已经全面成型，项目数据分析师和专业事务所开始在数据分析所涉及的各个领域发挥巨大的作用。

4. 第四阶段：迅速发展壮大期

随着企业的不断成长和行业规模的日渐扩大，在 2008 年，数据分析行业发展迈进了新的时期，当年 1 月，国家发展和改革委员会培训中心与项目数据分析师考试培训认证中心达成战略合作意向，共同推广项目数据分析师培训项目。

同年 6 月，经国资委审核同意，国家民政部正式批准中国商业联合会数据分析专业委员会（以下简称数据分析委员会）成立，随后在 12 月 21 日，数据分析委员会成立大会暨数据分析行业研讨会在北京隆重召开，标志着中国数据分析行业步入了一个迅速发展壮大的阶段。

次年，数据分析行业的培训全面开展，在 4—12 月，数据分析委员会在北京举办了首次面向全国会员的继续教育面授课程，来自全国几十家事务所及各地 140 名会员和学员参加了此次继续教育培训，实现了会员与专家面对面的交流学习，也为会员的从业和创业提供了最为直接的引导和帮助。

自数据分析委员会成立后，全国各地纷纷成立新的事务所，并向委员会备案，而且委员会颁发的资质已得到社会各界的认可。

在 2010 年，我国项目数据分析师事务所迅速增加，在 4 月，首届中国数据分析业峰会在京举办，在国家发展和改革委员会及相关领导的支持和监督下，各事务所代表共同签署了行业自律宣言，并启动了行业首个社会公益服务平台——项目数据分析服务平台，开始面向社会开放公益性服务职能。

1.2.2 充分认识大数据时代

早在 2008 年，大数据 (Big Data) 这个概念就已被提出，随着“云计算”“互联网”和“物联网”的快速发展，大数据也吸引了越来越多人的关注，成为社会热点之一。数据时代的到来，标志着人类迈入商务智能化时代。

1. 大数据概述

现如今，无论是技术人员、咨询人士还是各行各业的精英人士都在探讨“大数据”，那么，多大的数据才叫大数据呢？其实在业内，大数据至今尚无确切和统一的定义，维基百科给出了一个定性的描述：大数据是指无法使用传统和常用的软件技术和工具在一定时间内完成获取、管理和处理的数据集。

随着大数据的不断火热，对大数据特点描述的说法也很多，但是业内最普遍描述

的大数据必须具备容量 (Volume) 巨大、处理速度 (Velocity) 快、数据类型多样 (Variety) 以及商业价值 (Value) 高这 4 个特点, 具体的特点描述如图 1-2 所示。

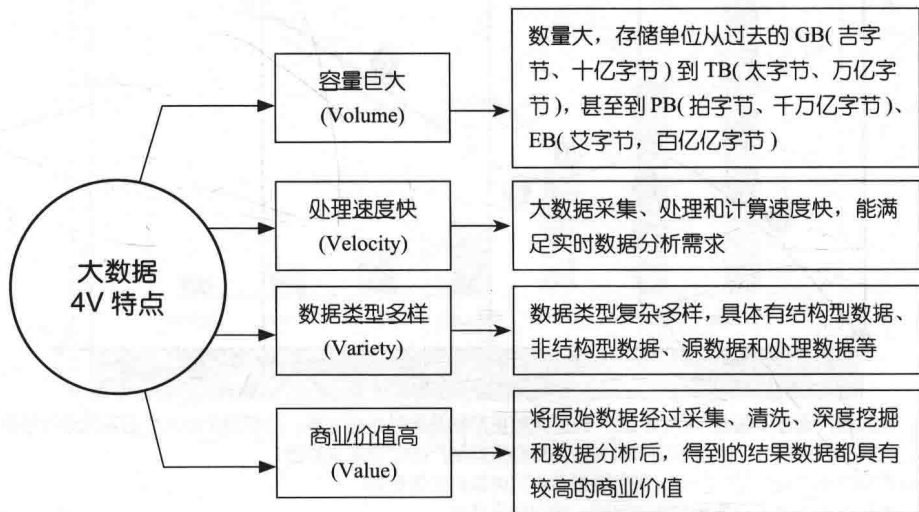


图 1-2

除了以上的 4V 特点以外，大数据还有一个显著的特点，即数据是在线的 (Online)，这是大数据区别于传统数据的最大特征，因为只有在线的数据，才能随时调用和计算，这是互联网高速发展背景下的必然特点。

离线的磁盘数据，其商业价值远远不及在线数据的商业价值大。比如，对于现在比较流行的各种打车工具，客户的数据和司机的数据都是实时在线的，这样的数据才有意义。

知识补充 | 大数据新解

随着社会的发展，“大数据”一词的重点已经不再仅仅局限于对数据规模的定义，如今提到“大数据”，人们联想到的是：①新时代的标志；②给传统计算技术和信息技术带来的技术挑战和困难；③处理大数据需要的新技术和新方法；④大数据分析和应用带来的新发明、新服务和新的发展机遇。

2. 大数据的行业现状

大数据的发展阶段大致可以划分为 4 个阶段，分别是探索期、市场启动期、高速发展期和应用成熟期。而在我国，40% 的企业没有大数据平台部署和大数据应用；24% 的企业已经开始部署大数据平台但还未实现应用；36% 的企业已经实现大数据应用，因此，我国的大数据产业正处于高速发展期，如图 1-3 所示。