

KEEPING UP WITH THE QUANTS

Your Guide to Understanding
and Using Analytics

全球顶尖商业思想家

托马斯·达文波特

智能商业五部曲

成为数据分析师

6步练就数据思维

[美] 托马斯·达文波特 金镇浩◎著

Thomas H. Davenport Jinho Kim

盛杨燕◎译

无须统计背景，也能练就数据力
清楚沟通，有力说服，精准决策

KEEPING UP WITH THE QUANTS

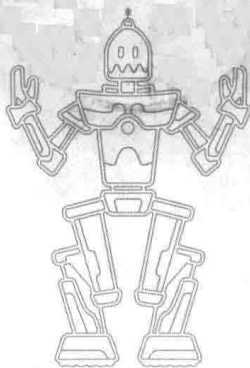
Your Guide to Understanding and Using Analytics

成为数据分析师

6步练就数据思维

[美] 托马斯·达文波特 金镇浩◎著
Thomas H. Davenport Jinho Kim

盛杨燕◎译



浙江人民出版社
ZHEJIANG PEOPLE'S PUBLISHING HOUSE

图书在版编目(CIP)数据

成为数据分析师 / (美) 达文波特, 金镇浩著; 盛杨燕译. —杭州:
浙江人民出版社, 2018.1

ISBN 978-7-213-08622-9

I. ①成… II. ①达… ②金… ③盛… III. ①数据处理—应用—
企业管理 IV. ①F272.7

中国版本图书馆CIP数据核字(2018)第015103号

上架指导: 经济管理 / 智能商业

版权所有, 侵权必究

本书法律顾问 北京市盈科律师事务所 崔爽律师
张雅琴律师

浙江省版权局
著作权合同登记章
图字: 11-2017-308号

成为数据分析师

[美] 托马斯·达文波特 金镇浩 著
盛杨燕 译

出版发行: 浙江人民出版社(杭州体育场路347号 邮编 310006)

市场部电话:(0571) 85061682 85176516

集团网址: 浙江出版联合集团 <http://www.zjcb.com>

责任编辑: 郦鸣枫

责任校对: 戴文英

印刷: 河北鹏润印刷有限公司

开本: 720mm 965mm 1/16 印张: 14

字数: 169千字 插页: 5

版次: 2018年2月第1版 印次: 2018年2月第1次印刷

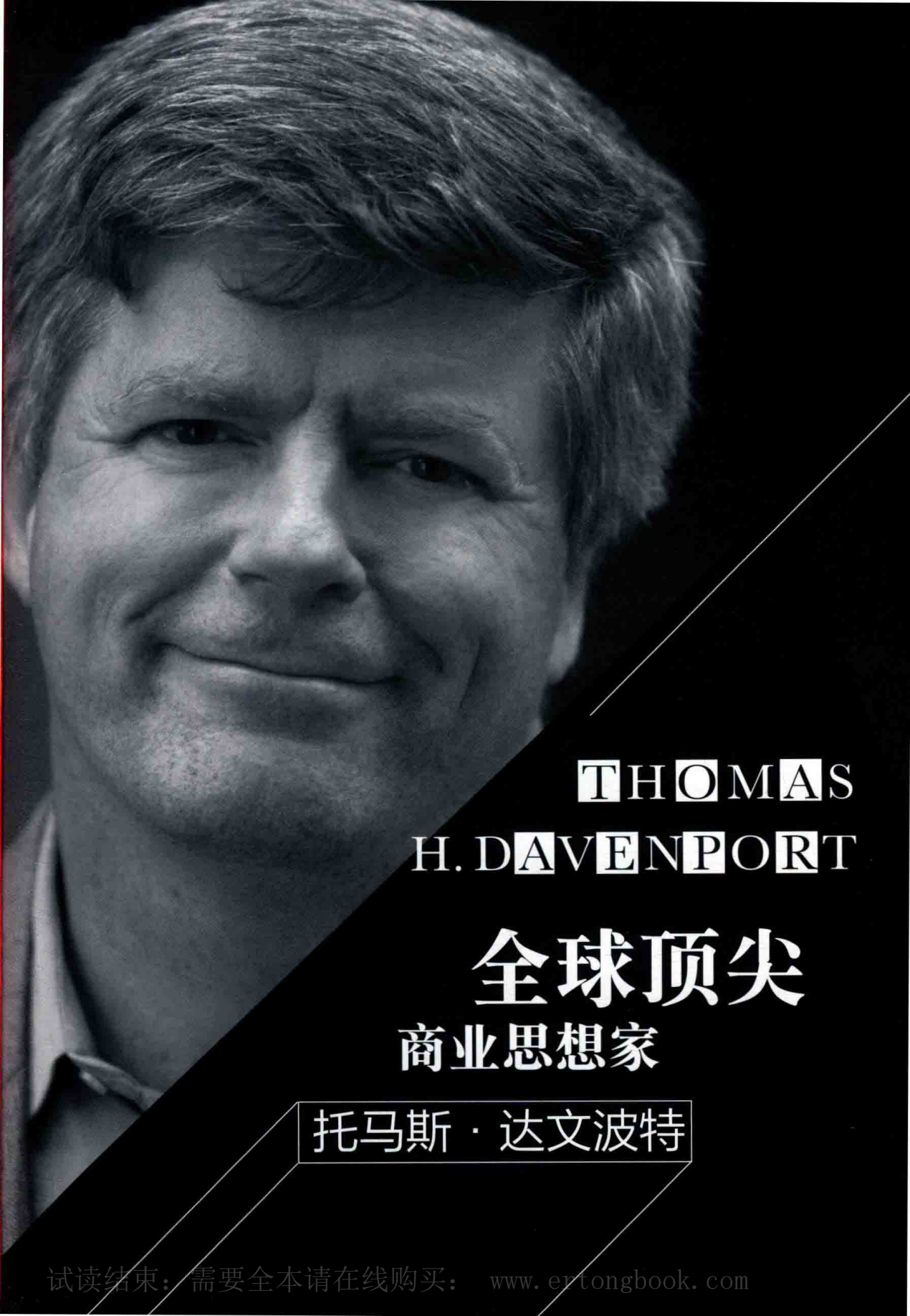
书号: ISBN 978-7-213-08622-9

定价: 62.90元

如发现印装质量问题, 影响阅读, 请与市场部联系调换。

C 滬声文化
heers Publishing

a mindstyle business
与思想有关



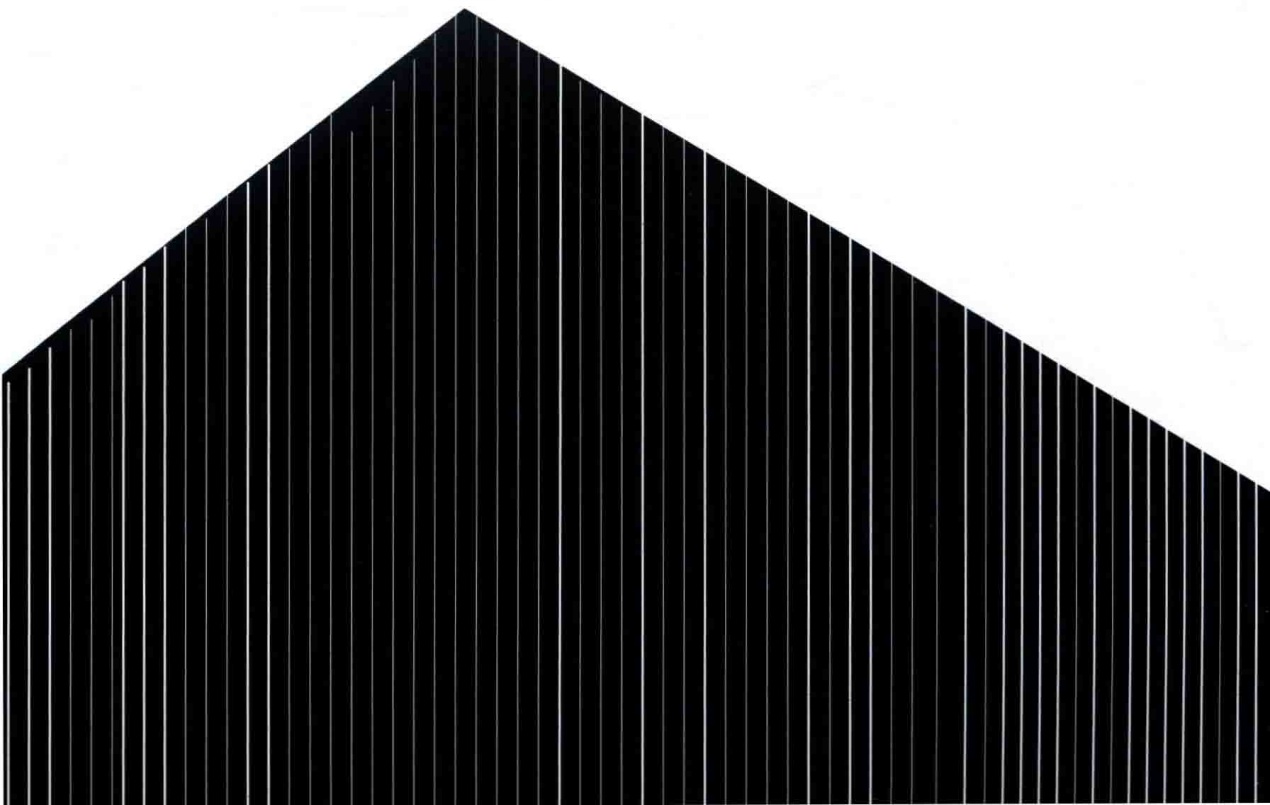
THOMAS
H. DAVENPORT

全球顶尖
商业思想家

托马斯·达文波特

DATA B.

全球商业界炙手可热的
数据分析之父

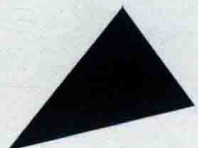


G

在

商业界，曾有这样一句至理名言：如果你问一位CIO，今时今日数据对其企业的意义，那么他会很愿意跟你谈谈“数据分析竞争法”的必要性，以及快速做出正确决策的重要性。这句话源于一本叫作《数据分析竞争法》（*Competing on Analytics*）的著作，随着这本著作诞生的还有两个超级火热的概念：一个是“数据分析竞争法”，一个是“大数据”。

站在这两个惊世概念背后的是一位名叫托马斯·达文波特的数据分析师。这位出生于1954年10月17日的美国人，毕业于哈佛大学，曾经先后在哈佛商学院、芝加哥大学和波士顿大学任教，还曾经担任过埃森哲战略变革研究院主任，美国知名商学院巴布森学院著名教授，对知识管理有深入的研究。



如果在全球数据分析师领域进行一次排名，达文波特无疑会成为很多人心中的榜首。在商业分析领域摸爬滚打的35年里，达文波特没有一丝懈怠，多次领先创立顶尖的数据分析法，比如数据分析DELTA模型、成为数据分析师的三原则等。他知道，领先的企业不仅是在收集和存储大量的数据，而且围绕着由数据引发的新观点制定竞争战略，这会使企业获益无穷。

如今，虽然他已经年过花甲，但仍然神采奕奕，一直在从事自己认为最性感的工作——数据分析。他是今时今日大数据时代当之无愧的数据分析之父。

BIG DATA

3次预见商业拐点的
大师级玩家

1983年，达文波特离开教职进入商业研究领域，在短短7年后，便迎来了商业事业的顶峰。他曾3次预见商业范式转型的大拐点，成为当今商业界极富洞见的未来学家。

达文波特第一次预见商业的转型是在1990年，当时他敏锐地注意到，企业要想在市场中取得胜利，就要面向顾客需求，重组业务流程。因此，他开创性地提出了流程再造（reengineering）理念，一时声名鹊起。

达文波特最早发起了知识管理运动，这是他第二次预见商业的未来。俗话说：“拥有100位博士的企业，未必拥有100位博士的知识；拥有一群智商120分以上员工的企业，企业商往往远低于120分。”他认为，没有知识管理的企业，员工进入后只会感受到不断的付出，因此，无论获得多好的待遇，都只能算是“出卖劳动力”。

真正让达文波特在全球商业管理界

成为风云人物的则是他在互联网时代兴起时提出的“注意力经济”概念，这是他第三次预见商业的未来。“注意力经济”研究了CEO们所面临的最大挑战之一：如何赚得和消费新经济时代的企业货币。他成功抓住了企业从“知识时代”到“注意力时代”转型的开创性节点，并出版了同名著作，被IBM前知识管理研究院院长拉里·普鲁萨克（Larry Prusak）重磅推荐。他所在的埃森哲公司，也在当年被《财富》全球500强企业选中作为头号咨询公司。

达文波特一次次对商业时代潮流进行了精准的感知和把握，让无数企业管理者对世界有了理解和思考的框架，而这完全得益于他超强的数据分析能力。



DATA

与克莱顿·克里斯坦森、杰克·韦尔奇 比肩的商业思想家

达文波特凭借睿智的洞见和敏锐的商业思想，为自己赢得了无数荣誉。2000年，他被《CIO》杂志评选为“新经济十大杰出人物”之一。之后，在由他发布的巴布森学院“全球著名商业思想家”排行榜中，他把自己排在了第25位。在这个榜单上，知名战略思想家加里·哈默拔得头筹，畅销书作家托马斯·弗里德曼和微软前董事长比尔·盖茨分别列第二、第三位，达文波特是与克莱顿·克里斯坦森、杰克·韦尔奇比肩的商业思想家。2003年，达文波特被权威的《咨询师》杂志评选为全世界“最顶尖的25名咨询师”之一。

达文波特还是一位知名的商业图书

作家，出版了近20本管理类畅销书，被多个国家引进出版，极负盛名。

作者演讲洽谈，请联系

speech@cheerspublishing.com

更多相关资讯，请关注



湛庐文化微信订阅号

湛庐文化
Cheers Publishing
THE CHINESE WAY 与智者共进

特别
制作

KEEPING
UP
WITH THE
QUANTS

引言

数据决策时代，人人都是分析师

我们生活在一个数据泛滥的时代。数据正以惊人的速度在增长，每个人的下一秒都会被更多的数据包围。我们收集数据的主要目的是完善企业、政府和社会层面的决策制定机制。因此，如果我们无法通过定量分析，利用数据实现更好的决策制定，就是对数据资源的浪费，也有可能造成不良后果。因此，本书致力于为你展示定量分析的运作方式，以及该如何利用定量分析做出更好的决策，即使你没有相关知识背景，也无妨。

数据，创新产品与服务的源泉

数据的力量正在各行各业中崛起。如果你热爱运动，那肯定知道《点球成金》(*Moneyball*) 这部电影，影片主角奥克兰运动家棒球队总经理比利·比

恩 (Billy Beane) 利用球员的表现数据和分析学革新了职业棒球运动。现在, 这种革新已经延伸到了所有的主流体育项目当中。如果你喜欢玩在线游戏, 可能知道星佳 (Zynga) 和美国艺电 (Electronic Arts), 这些社交网络游戏公司正在收集并分析用户所有的游戏行为。那电影呢? 你也许知道奈飞 (Netflix) 可以利用算法预测你可能喜欢的电影, 好莱坞的一些制片人会利用算法来推断什么样的电影投资回报率高, 美国独立电影公司相对论传媒 (Relativity Media) 就是这么做的。

数据种类各不相同。有的数据因为事务性用途被收集和管理, 比如, 企业和机构通过跟踪员工上班时间和剩余假期所得的数据。当企业收集到大量数据之后, 就希望能读懂这些数据, 并在此基础上做出决策。同样地, 也可以利用分析学来研究与人力资源相关的事务性数据, 企业可能会提出“下一年度有多少员工可能会退休”或者“员工休完所有的假期和其年度绩效考核之间是否存在关联”之类的问题。

不过, 数据和分析学的作用并不局限于完善内部决策的制定。像谷歌、Facebook、亚马逊和 eBay 等许多立足于互联网的企业, 都在利用收集网上交易数据所形成的大数据来支撑决策制定以及为客户提供新产品和服务。无论你是想实现更富成效的内部决策, 还是想为客户提供更具价值的产品, 分析学必不可少, 它会对数据进行汇总、分析, 并找出其中的含义和内在关联。要读懂并挖掘出数据的价值, 必须借助数学或统计分析, 简单地说, 就是分析学。

KEEPING UP
WITH THE
QUANTS人人都是
分析师

数据分析的本质

一般情况下，我们所说的分析是指，使用大量数据、统计和定量分析、解释和预测模型以及基于事实的管理来推动决策过程与实现价值增生。

根据分析的方法和目的，分析可以被划分为描述性分析（descriptive analytics）、预测性分析（predictive analytics）和规范性分析（prescriptive analytics）。描述性分析包括数据收集、整理、制表、制图以及描述正要研究的事物的特征，这类分析以往被称为“报告”。描述性分析可能非常有用，但它不能解释某种结果出现的原因或者未来可能会发生的事情。

预测性分析不仅可以对数据特征和变量（可以假定取值范围的因素）之间的关系进行描述，还可以基于过去的的数据预测未来。预测性分析首先会确定变量之间的关联，然后基于这种已知关联预测另一种现象出现的可能性，比如在看到某个广告后，一位消费者可能会去买产品的可能性。虽然预测性分析中的预测是基于变量之间的关联做出的，但这并不代表预测性分析都需要明确因果关系。事实上，准确的预测并不一定需要基于因果关系。

规范性分析是更高层次的分析，如实验设计和优化等。就像医生会在处方上建议患者采取什么行动一样，实验设计试图通过做实验给出某些事情发生的原因。为了能够在因果关系研究中信心满满地做出推断，研究人员必须妥善处理一个或多个独立的变量，并有

效控制其他无关的变量。如果处于实验环境下的测试组的表现大大优于对照组，决策制定者就应该立即推广这种实验环境。

优化是规范性分析采用的一种方法，指试图识别出一个特定变量与另一个变量之间理想的关系水平。例如，我们可能会对识别最有可能让产品实现高收益的价格感兴趣。同样地，优化这种方法能够识别出使零售企业最大限度避免缺货情况的库存水平。

根据分析采用的方法以及收集和分析的数据类型，我们可以将分析分为定性分析（qualitative analysis）和定量分析（quantitative analysis）。定性分析的目的是深入了解某种现象出现的根本原因和诱因。非结构化数据通常是从少数非代表性案例中收集而来，并进行了非统计性的分析。定性分析是分析的最初阶段，它通常是探索性研究的有效工具。定量分析是指通过统计、数学或计算的方法对现象进行系统的实证研究。通常情况下，结构化数据是从大量典型案例中收集而来，并进行统计分析。

为了服务于研究者的不同研究目的，存在以下几种类型的分析：

- 统计学：收集、整理、分析、说明和呈现数据的学科。
- 预测：根据已有数据，预测一些感兴趣的变量在未来某个特定时间点的情况。
- 数据挖掘：通过使用算法和统计技术，自动或半自动地提取大量数据中未知的有趣模式。
- 文本挖掘：用类似于数据挖掘的方式从文本中得出模式和趋势的过程。

- 优化：在同时满足约束条件的情况下，按照某些标准，利用数学方法来寻找最优的解决方案。
- 实验设计：给各组随机分配被试，然后使用测试组和对照组来推导出特定结果中存在的因果关系。

虽然此处给出了一系列常用的分析方法，但在使用的过程中，会不可避免地出现相当大的重叠。例如，回归分析（regression analysis）是预测性分析中最常用的方法，与此同时，它也是统计学、预测和数据挖掘中常用的方法。此外，时间序列分析（time series analysis）是用于分析数据随时间变化的一种具体统计方法，在统计学和预测中也经常被用到。

上文提到的支撑人力资源决策的事务性数据是定量分析的结构化数据，其数据量相对较小，即使是大型企业，可能也就只有 1TB 或 2TB。过去，分析就是发生在这样的环境中，因此，就让我们将其称为“小数据”吧。过去，分析学别无选择，必须针对小数据。

现如今，令大企业、非营利性组织和小型初创公司激动不已的是大数据。大数据是大量的非结构化数据，其来源极为广泛，可能来自互联网上的在线交谈，或者摄像机的连续镜头，或者医院患者的 DNA 分析。这类数据量非常庞大，甚至达到拍字节（PB）的程度。比如，谷歌每天要处理 24 PB 的互联网数据；美国电话电报公司（AT&T）每天要转换 30 PB 的语音和数据通信数据。如今，通过新型软件和硬件技术，我们可以分析大数据并挖掘出其蕴涵的意义。

数据即未来

大数据一词指的是数据量异常庞大且多种多样的非结构化数据。

以下是一些例子：

- 社交网络游戏公司星佳每天要处理超过 1 PB 的游戏数据。
- YouTube 用户每天观看超过 10 亿小时的视频。
- Twitter 用户每月进行 320 亿次搜索。
- 2011 年，谷歌用户每天进行近 50 亿次搜索。
- 2009 年，全球每天发送的短信超过 25 亿条。
- 2010 年，世界各地的手机使用量达到 50 亿台。
- 经过全面分析，一个人类基因组包含大约 1TB 的数据。
- 评估一头奶牛健康状况的无线传感器每年会传输 200 MB 的数据。
- 2008 年，连接到互联网的设备的数量超过了全球人口的总量。
- 据思科系统公司（Cisco Systems）估计，20 种常见的家用电器产生的上网流量超过了所有互联网用户在 2008 年间产生的上网流量之和。
- 麦肯锡咨询公司预测，在美国的各行各业，职工超过

千人的企业存储的平均数据比美国国会图书馆（Library of Congress）的数据还要多。

大数据和分析学会改变各行各业的商业职能。任何组织或个人只要抢先一步掌握大数据，就会奠定至关重要的竞争优势，就像在小数据时代占据先机进行分析的人能在竞争者中脱颖而出一样。因此，企业和组织机构必须抓住大数据的历史机遇。

大数据的潜能之所以能得到激发，是因为计算和数据收集设备的普及，而传感器和微型处理器即将无处不在。基本上，任何机器或电子设备都会遗留下有关性能、位置和状态的数据，而这些设备及其用户通过在互联网上的沟通又催生了一个海量数据源，当所有这些数据与由其他媒介，如无线和有线电话、电报、卫星等媒介生成的数据结合起来时，数据的未来岂止于大。

如果所有这些数据都可用，就意味着基本上所有的企业和组织行为都可以被视为大数据问题或者大数据行为。在制造业内，大部分机器设备已经配置了一个或多个微处理器，大数据的氛围越来越浓烈。随着消费者的选择日趋多样化，市场推广早就成为一种大数据行为。谷歌也将其无人驾驶汽车项目定性为大数据项目。

许多企业的首席执行官曾公开表示，分析性思维和决策制定是企业成就和个人财富的来源，其中凯撒娱乐集团（Caesars Entertainment）的首席执行官加里·拉夫曼（Gary Loveman）就曾说过：“我们思考了吗，或者说我们知道结果了吗？”亚马逊首席执行官杰夫·贝佐斯（Jeff Bezos）曾说过：“我们对数据来者不拒！”领英联合创始人兼董事长里德·霍夫曼（Reid