

The background of the cover features a complex network diagram with numerous nodes and connecting lines, rendered in a light beige color. A large, stylized, light beige graphic element, resembling a wide, shallow trapezoid with dashed lines, frames the main title.

# 大数据时代 预算管理

—— 理论与创新实践研究 ——

DASHUJU SHIDAI YUSUAN GUANLI LILUN YU CHUANGXIN SHIJIAN YANJIU

王晓丽 孟秀蕊 著

吉林人民出版社

## ■ 作者简介

---

**王晓丽**，管理学硕士。现为河北劳动关系职业学院教师，高级会计师。曾拥有多年大型集团企业预算管理实践经验，精通公司预算与业绩评价。现主要从事会计教学与研究，研究方向为财务管理及管理会计。主持多项课题，发表学术论文十余篇。

---

**孟秀蕊**，管理学硕士。现为河北劳动关系职业学院教师，讲师，中级会计师。主要从事会计教学与研究，研究方向为财务管理及管理会计。主持和参与多项课题，发表学术论文十余篇。

# 大数据时代预算管理理论 与创新实践研究

王晓丽 孟秀蕊 著

吉林人民出版社

## 图书在版编目(CIP)数据

大数据时代预算管理理论与创新实践研究 / 王晓丽,  
孟秀蕊著. — 长春: 吉林人民出版社, 2021.5  
ISBN 978-7-206-18129-0

I . ①大… II . ①王… ②孟… III . ①预算管理 - 研  
究 IV . ① F810.3

中国版本图书馆 CIP 数据核字 (2021) 第 101995 号

## 大数据时代预算管理理论与创新实践研究

---

著 者: 王晓丽 孟秀蕊

责任编辑: 王 丹

封面设计: 谢少红

吉林人民出版社出版 发行 (长春市人民大街 7548 号) 邮政编码: 130022

印 刷: 三河市华晨印务有限公司

开 本: 710mm × 1000mm 1/16

印 张: 12.75

字 数: 240 千字

标准书号: ISBN 978-7-206-18129-0

版 次: 2021 年 5 月第 1 版

印 次: 2021 年 5 月第 1 次印刷

定 价: 59.00 元

---

如发现印装质量问题, 影响阅读, 请与印刷厂联系调换。



# PREFACE

## 前 言

预算管理作为管理会计的一种方法，其功能从最初的计划、协调生产发展到现在的兼具控制、激励、评价等功能，成为一种综合贯彻企业战略方针的经营机制，处于企业管理控制系统的核心位置。近年来，我国在会计准则国际趋同、会计信息化建设及企业内部控制等各方面出台了许多新制度与新指引，对于规范我国会计管理及会计国际化起到了非常重要的促进作用，并将管理会计推向了公众视野，使管理会计成为企业与社会关注的热点。全面预算管理作为管理会计的基础，也受到了越来越多的重视，随着大数据时代的到来，全面预算管理迎来了新的发展契机，探究大数据时代企业全面预算管理的发展趋势有着重要的现实意义。

大数据如今已成为移动互联网时代的关键词，越来越多的企业通过各类互联网技术搜集与挖掘高速度、大容量数据进行分析与使用。大数据时代，企业面对杂乱无章的数据海洋，如何高效应用数据，让大数据“活起来”，全面预算管理可谓企业应用大数据最好的方式。例如，针对预计利润表中产品销售收入实施穿透分析，利用全面预算管理的解决方案，能够得到不同产品类别及不同时期等具体数据，有利于企业将预算数据与实际数据展开对比，进而找出解决问题的对策。总之，大数据时代，无论小企业或者大企业，都会在网络上产生信息、关系及行为等方面的数据，大数据也将成为现代企业非常重要的资产；全面预算管理作为现代企业内部管理控制系统核心内容，是企业实现战略、业务、财务、人力协调发展的重要手段。而大数据的深入应用则为企业全面预算管理的实施提供了极为重要的保障与技术支持。

本书属于预算管理方面的著作，由大数据时代概述、预算管理概述、大数

据时代给预算管理带来的机遇、投资决策与资金预测、预算与管控、预算管理在企业管理中的地位与作用、大数据时代预算管理的创新应用案例等组成，全书以预算管理为核心，对大数据背景下的预算管理的具体内容与方法应用等做了详细论述，对预算管理相关方面的研究者与从业人员具有学习和参考价值。

# CONTENTS

## 目 录

### 第一章 大数据时代概述 / 001

第一节 大数据的基本理论 / 001

第二节 大数据时代的特色 / 007

第三节 大数据的应用领域 / 010

### 第二章 预算管理概述 / 015

第一节 预算管理的内涵 / 015

第二节 预算管理的理论基础 / 023

第三节 预算管理的特征 / 027

第四节 预算管理的内容 / 030

第五节 预算管理的运行环境 / 036

### 第三章 大数据时代给预算管理带来的机遇 / 041

第一节 大数据时代预算管理与传统预算管理的联系与区别 / 041

第二节 大数据时代给预算管理带来的机遇和挑战 / 049

第三节 大数据背景下预算管理的发展 / 052

### 第四章 投资决策与资金预测 / 057

第一节 预测、决策与预算的关系 / 057

## 002 \ 大数据时代预算管理理论与创新实践研究

第二节 项目分析与投资决策 / 072

第三节 资金预测与可持续发展 / 080

### 第五章 预算与管控 / 090

第一节 预算编制方法 / 090

第二节 业务预算的编制 / 097

第三节 专门决策预算的编制 / 105

第四节 财务预算的编制 / 106

第五节 财务预算的执行与考核 / 110

第六节 预算管理组织架构与管理流程 / 127

### 第六章 预算管理在企业管理中的地位与作用 / 138

第一节 预算管理控制与战略管理控制 / 138

第二节 预算管理在企业管理控制系统中的地位 / 146

第三节 预算管理在企业管理控制系统中的作用 / 149

第四节 以预算管理为导向的管理控制体系在企业管理中的应用 / 152

### 第七章 大数据时代预算管理的创新应用案例 / 157

第一节 大数据环境下 H 公司全面预算管理系统的研究 / 157

第二节 基于大数据利用的企业财务管理创新研究——以 L 公司为例 / 169

第三节 大数据背景下服饰类电商企业全面预算管理研究——以 H 集团为例 / 184

### 参考文献 / 196

# 第一章 大数据时代概述

## 第一节 大数据的基本理论

### 一、大数据的概念

随着社会化网络的兴起及云计算、移动互联网和物联网等新一代信息技术的广泛应用，全球数据量呈现出前所未有的爆发增长态势。大数据带来的信息风暴正在逐渐改变我们的生活环境、工作习惯和思维方式。我们看到，在商业、经济、医药卫生及其他领域中，决策正日益基于数据和分析而做出，而非仅仅基于经验和直觉。大数据是近年来科学研究的核心所在，其已成为信息时代新阶段的标志，是大型信息系统和互联网的产物，是实现创新驱动发展战略的重要机遇。大数据的发展与应用，将对社会的组织结构、国家治理模式、企业的决策机构、商业的业务策略及个人的生活方式产生深刻的影响。美国政府将“大数据战略”提升为最高国家发展策略，将大数据定义为“新石油”，把对数据的占有与控制作为陆海空权之外的新国家核心能力。

对于“大数据”（Bigdata），研究机构 Gartner 给出了这样的定义：大数据是指那些需要利用新处理方法才能通过数据体现出更强决策力、洞察力和流程优化能力的海量、高增长率和多样化的信息资产。

从认识论的角度来说，科学始于数据。人类历史上的大数据，源于科技领域，确切地说，是源于大科学研究。位于瑞士的欧洲核子研究中心由全球逾 8 000 位物理学家合作兴建的大型强子对撞机，刚试运行数据量即达 25PB/年，可以说他们率先创建了“大数据”的概念。旨在测定人类基因组 30 亿碱基遗传密码的基因组计划，进行个体基因组测定时，数据量即已高达 13PB/年。而此计划后，学界受其鼓舞开展了一系列遗传背景迥异、不同疾病群体及大量其他物种的基因组测序，数据量迅速逼近 ZB 级（是 PB 的百万倍），不约而同地创造了“大数据”概念。今天人们常用的互联网最初就是这些领域的科学家

为解决海量数据传输而发明的。

传统哲学认识论是以人为主体，而在大数据背景下的认识论主体发生了分化，即认识论主体的意向方和实施方分离，意向方仍然是人类，实施方则由人类变成了机器，意向方和实施方各自承担着自己的需求职责，认识的动机和目的发生了相应的变化，任何人只关注对自己有用的信息，而机器提供可视化分析，形成大数据认知外包的特性。

大数据通过海量数据来发现事物之间的相互关系，通过数据挖掘从海量数据中寻找蕴藏其中的数据规律，并利用数据之间的相互关系来解释过去、预测未来，从而实现新的数据规律对传统因果规律的补充。大数据能预测未来，但作为认识论主体意向方的人类，只关注预测的结果，而忽视预测的解释，这就造成预测能力强、解释能力弱的局面。

大数据模型和统计建模有本质的区别。就科学研究中的地位来说，统计建模经常是经验研究和理论研究的配角、检验者；而在大数据的科学研究中，数据模型就是主角，模型承担了科学理论的角色。就数据类型来说，统计建模的数据通常是精心设计的实验数据，具有较高的质量；大数据中则是海量数据，往往类型繁多，质量参差不齐。就确立模型的过程来说，统计建模的模型是根据研究问题来确定的，目标变量预先已经确定好；大数据中的模型则是通过海量数据确定的，且部分情况下目标变量并不明确。就建模驱动不同来说，统计建模是验证驱动，强调的是先有设计再通过数据验证设计模型的合理性；而大数据模型是数据驱动，强调的是建模过程及模型的可更新性。

大数据思维：是指一种意识，认为公开的数据一旦处理得当就能为千百万人急需解决的问题提供答案。

量化思维：大数据是直觉主义到量化思维的变革，在大数据量化思维中一切皆是可量化的，大数据技术通过智能终端、物联网、云计算等技术手段来“量化世界”，从而将自然、社会、人类的一切状态、行为都记录并存储下来，形成与物理足迹相对应的数据足迹。

全局思维：是指大数据关注全数据样本，大数据研究的对象是所有样本，而非抽样数据，关注样本中的主流，而非个别，这表征大数据的全局和大局思维。开放共享、数据分享、信息公开在分享资源的同时，也在释放善意，取得互信，在数据交换的基础上产生合作，这将打破传统封闭与垄断，形成开放、共享、包容、合作思维。大数据不仅关注数据的因果关系，更关注其相关性，提高了数据采集频度，放宽了数据的精确度，提升了容错率，用概率看待问题，使人们的包容思维得以强化。

关联思维、轨迹思维：每一天，我们的身后都拖着一条由个人信息组成的长长的“尾巴”。我们点击网页、切换电视频道、驾车穿过自动收费站、用信用卡购物、使用手机等行为，这些过去完全被忽略的信息，都通过各种方式被数据化地记录下来，全程实时追踪数据轨迹，管理数据生命周期，保证可靠的数据源头、畅通的数据传递、精准的数据分析、友好可读的数据呈现。

预测思维：预测既是大数据的核心，也是大数据的目标。

从技术上理解，大数据是一次技术革新，对大数据的整合、存储、挖掘、检索、决策生成都是传统的数据处理技术无法顺利完成的，新技术的发展和成熟加速了大数据时代的来临，如果将数据比作肉体，那技术就是灵魂。大数据时代，数据、技术、思维三足鼎立。《大数据时代》的作者维克托认为，大数据使我们真正拥有了决定性的价值资源，它是新的黄金。这里值得注意的是，大数据的意义不在于掌握海量的数据，而是通过数据挖掘等手段对其进行专业的分析来实现数据的“增值”。

大数据可分成大数据技术、大数据工程、大数据科学和大数据应用等领域。目前，人们谈论最多的是大数据技术和大数据应用，工程和科学问题尚未被重视。大数据工程指大数据的规划建设、运营管理的系统工程；大数据科学关注大数据网络发展与运营过程中发现和验证大数据的规律，及其与自然、社会活动之间的关系。

物联网、云计算、移动互联网、车联网、智能手机、平板电脑、PC 及遍布地球各个角落的各种各样的传感器，无一不是数据来源或者承载的方式。

核心价值在于对于海量数据进行存储和分析。相比现有的其他技术而言，大数据的“廉价、迅速、优化”这三方面的综合成本是最优的。大数据必将是一场新的技术信息革命，我们有理由相信，未来人类的生活、工作也将随大数据革命而产生革命性的变化。

## 二、大数据的产生背景

大数据似乎是在一夜之间悄然而至，并迅速走红。大数据在 2012 年进入主流大众的视野，人们把 2012 年称为“大数据的跨界年度”。经过各方面的分析，大数据之所以进入人们的视野，缘于三种趋势的合力。

第一，随着互联网的发展，许多高端消费公司为了提供更先进、更完美的服务，加大了对大数据的应用。

第二，人们在无形中纷纷为大数据投资。

第三，商业用户和其他以数据为核心的消费产品也开始期待以一种同样便

捷的方式来获得大数据的使用体验。我们在网上看电影、买产品等已经成为现实。既然互联网零售商可以为用户推荐一些阅读书目、电影和产品，为什么这些产品所在的企业却做不到呢？举个例子说，为什么房屋租赁公司不能明智地决定将哪一栋房屋提供给租房人呢？毕竟，该公司拥有客户的租房历史和现有可用租房屋库存记录。随着新技术的出现，公司不仅能够了解到特定市场的公开信息，还能了解到有关会议、重大事项及其他可能会影响市场需求的信息。通过将内部供应链与外部市场数据相结合，公司可以更加精确地预测出可租的房屋类型和可用时间。

类似地，通过将这些内部数据和外部数据相结合，零售商每天都可以利用这种混合式数据确定产品价格和摆放位置。通过考虑从产品供应到消费者的购物习惯这一系列事件的数据（包括哪种产品卖得比较好），零售商就可以提升消费者的平均购买量，从而获得更高的利润。所以，商业用户也成了推动大数据发展的动力之一。

总的来说，大数据的产生既是时代发展的结果，也是利益驱使的结果。当然，那些小公司的发展，乃至个人的服务需求也在为大数据的产生添砖加瓦，只是单个个体的效果不明显，但反映在整个大数据产业中依然是巨大的，其中的道理就不再多说了。

### 三、大数据的特点

数据分析需要从纷繁复杂的数据中发现规律并提取新的知识，是大数据价值挖掘的关键。经过数据的计算和处理，所得的数据便成为数据分析的原始数据，根据所需数据的应用需求对数据进行进一步的处理和分析，最终找到数据内部隐藏的规律或者知识，从而体现数据的真正价值。大数据的分析技术必须紧密围绕大数据的特点开展，只有这样才能确保从海量、冗杂的数据中得到有价值的信息。

维克托·迈尔-舍恩伯格及肯尼斯·库克耶编写的《大数据时代》中，大数据一般具有“4V”特点：Volume（大量）、Velocity（高速）、Variety（多样）、Value（价值）。具体来讲，大数据有如下特点。

#### （一）数据体量巨大

大数据通常指10TB（1TB=1024GB）规模以上的数据量，之所以产生如此巨大的数据量，一是由于各种仪器的使用，使用户能够感知到更多的事物，这样，这些事物的部分甚至全部数据就可以被存储下来；二是由于通信工具的使用，人们能够全时段地相互联系，机器—机器（M2M）方式的出现，使得交

流的数据量成倍增长；三是由于集成电路价格降低，很多电子设备拥有了智能模块，这些智能模块的使用过程中依赖或产生大量的数据存储。

### （二）流动速度快

数据流动速度一般是指数据的获取、存储及挖掘有效信息的速度。计算机的数据处理规模已从 TB 级上升到 PB 级，数据是快速动态变化的，形成流式数据是大数据的重要特征，数据流动的速度快到难以用传统的系统去处理。

### （三）数据种类繁多

随着传感器种类的增多及智能设备、社交网络等的流行，数据类型也变得更加复杂，不仅包括传统的关系数据类型，也包括以网页、视频、音频、E-mail、文档等形式存在的未加工的、半结构化的和非结构化的数据。

### （四）价值密度低

数据量呈指数增长的同时，隐藏在海量数据中的有用信息却没有以相应比例增长，反而使获取有用信息的难度加大。以视频为例，在连续的监控过程中，可能有用的数据仅有一两秒。大数据“4V”特征表明其不仅仅是数据海量，对于大数据的分析将更加复杂，更追求速度，更注重实效。

## 四、大数据的价值

现在的社会是一个高速发展的社会，科技发达，信息流通，人们之间的交流越来越密切，生活也越来越方便，大数据就是这个高科技时代的产物。阿里巴巴创办人马云曾提到，未来的时代将不是 IT 时代，而是 DT 时代，DT 就是 Data Technology，即数据科技。

有人把数据比喻为蕴藏能量的煤矿。煤炭按照性质有焦煤、无烟煤、肥煤、贫煤等，而露天煤矿、深山煤矿的挖掘成本又不一样。与此类似，大数据并不在“大”，而在于“有用”。价值含量、挖掘成本比数量更为重要。对于很多行业而言，如何利用这些大规模数据是赢得竞争的关键。

大数据的价值体现在以下几个方面：①对大量消费者提供产品或服务的企业可以利用大数据进行精准营销；②做小而美模式的中小企业可以利用大数据做服务转型；③面临互联网压力之下必须转型的传统企业需要与时俱进，充分利用大数据的价值。

不过，大数据在经济发展中的巨大意义并不代表其能取代一切对于社会问题的理性思考，科学发展的逻辑不能被湮没在海量数据中。著名经济学家路德维希·冯·米塞斯曾说：“就今日言，有很多人忙碌于资料之无益累积，以致对

问题之说明与解决，丧失了其对特殊的经济意义的了解。”这确实是需要警惕的。

在这个快速发展的智能硬件时代，困扰应用开发者的一个重要问题就是如何在功率、覆盖范围、传输速率和成本之间找到那个微妙的平衡点。企业利用相关数据进行分析，可以降低成本、提高效率、开发新产品、做出更明智的业务决策等。例如，通过结合大数据和高性能的分析，可能发生下列对企业有益的情况：①及时解析故障、问题和缺陷的根源，每年可能为企业节省数十亿美元；②为成千上万的快递车辆规划实时交通路线，躲避拥堵；③分析所有库存，以利润最大化为目标来定价和清理库存；④根据客户的购买习惯，为其推送他可能感兴趣的优惠信息；⑤从大量客户中快速识别出金牌客户；⑥使用点击流分析和数据挖掘来规避欺诈行为。

## 五、大数据的发展前景

如何开发、利用、保护好大数据这一重要的战略资源，是我国当前亟待解决的问题。大数据未来的发展趋势可以从以下几个方面进行。

### （一）开放源代码

大数据获得动力的关键在于开放源代码，帮助分解和分析数据。Hadoop 和 NoSQL 数据库便是其中的赢家，它们让其他技术商望而却步。毕竟，我们需要清楚怎样创建一个平台，既能解开所有的数据，克服数据相互独立的障碍，又能将数据重新上锁。

### （二）市场细分

当今，许多通用的大数据分析平台已投入市场，人们同时期望更多平台的出现，可以运用在特殊领域，如药物创新、客户关系管理、应用性能的监控和使用。若市场逐步成熟，在通用分析平台之上，开发特定的垂直应用将会实现。但现在的技术有限，除非考虑利用潜在的数据库技术作为通用平台（如 Hadoop、NoSQL）。人们期望更多特定的垂直应用出现，把目标定为特定领域的数据分析，这些特定领域包括航运业、销售业、网上购物、社交媒体用户的情绪分析等。同时，其他公司正在研发小规模分析引擎的软件套件，如社交媒体管理工具，这些工具以数据分析作为基础。

### （三）预测分析

建模、机器学习、统计分析和大数据经常被联系起来，用以预测即将发生的事情和行为。有些事情是很容易被预测的，如坏天气可以影响选民的投票率，而有些事很难被准确预测，如中间选民改变投票决定的决定性因素。但是，当数据累加时，我们基本上有能力可以大规模尝试一个连续的基础。网上

零售商重新设计购物车，来探索何种设计方式能使销售利润最大化；根据病人的饮食、家族史和每天的运动量，医生有能力预测未来疾病的风险。当然，在人类历史的开端，我们就已经有各种预测。但是，在过去，许多预测是基于直觉的，没有依靠完整的数据集，或者单单靠的是常识。当然，即便有大量数据支撑你的预测，也不表明那些预测都是准确的。另外，预测分析在许多领域流行起来，如欺诈发现（在外省使用信用卡时会接到的诈骗电话）、保险公司和顾客维系的风险管理。

## 第二节 大数据时代的特色

各种经济时代的区别，不在于生产什么，而在于怎样生产，用什么劳动资料生产。劳动资料不仅是人类劳动力发展的测量器，还是劳动借以进行的社会关系的指示器。马克思按照劳动资料或劳动工具的标准，把人类社会的发展分别称为石器时代、青铜时代、铁器时代、大机器时代。虽然马克思没有看到大数据时代的到来，但大数据作为一种新的劳动资料出现，对生产力的发展有直接的推动作用，这也是为什么如今会被称为大数据时代的原因。大数据时代有着一些小数据时代所没有的基本特征，包括以下三方面。

### 一、数据化：一切皆可“量化”

早在数千年前，人们就已经认识到了记录信息的重要性，于是发挥聪明才智发明了各种办法来记录信息。比如，结绳记事，据说波斯王大流士给他的指挥官们一根打了60个结的绳子，并对他们说：“爱奥尼亚的男子汉们，从你们看见我出征塞西亚人那天起，每天解开绳子上的一个结，到解完最后一个结那天，要是我不回来，就收拾你们的的东西，自己开船回去。”

尼罗河畔的古埃及人开始使用十进制的计数法。在贸易发达的两河流域，苏美尔人开始使用账单、收据、票据等物，可谓是现代会计学的鼻祖。印度人发明了阿拉伯数字，并且创立了“0”的概念。后来阿拉伯人把古希腊的数学融进自己的数学，又把这一简便易写的十进制位值记数法传遍欧洲，逐渐演变成今天的阿拉伯数字。

进入20世纪信息时代，由于计算机技术的发展与进步，数字化的信息储存方式应运而生。数字化指的是把模拟数据转换成用0和1表示的二进制码，这样计算机就可以处理这些数据了。1995年，美国麻省理工学院媒体实验室

的尼古拉斯·尼葛洛庞帝（Nicholas Negroponte）发表了其标志性著作《数字化生存》，主题就是“从原子到比特”，深刻阐释了信息技术、互联网对时代和人们生活的影响和价值。

记录方式的不断革新、信息技术的不断发展都为数据化奠定了良好的基础。过去不可计量、存储、分析和共享的很多东西被数据化了。拥有大量的数据和更多不那么精确的数据为我们理解世界打开了一扇新的大门。

日本先进工业技术研究所的教授越水重臣做了关于人的坐姿的研究，通过在汽车座椅下部安装总共 360 个压力传感器以测量人对椅子施加压力的方式，将人坐着时的身形、姿势和重量分布量化和数据化，产生独属于每个乘坐者的精确数据资料。这个系统能根据人体对座位的压力差异识别出乘坐者的身份，准确率高达 98%。这项技术可以作为汽车防盗系统安装在汽车上，有了这个系统之后，汽车就能识别出驾驶者是不是车主。这个系统不但可以发现车辆被盗，而且可以通过收集到的数据识别出盗贼的身份。通过这些数据，人们还可以利用事故发生之前的姿势变化情况，分析出坐姿和行驶安全之间的关系，在司机疲劳驾驶的时候发出警示或者自动刹车。把一个人的坐姿转化成数据，看似不可能，但这些数据孕育出了前景光明的产业。

通过光学字符识别软件，文字可以被数据化；通过 GPS，方向和位置可以被数据化；通过社交网站，人际交往可以被数据化……当整个世界都被数据化后，我们对世界的认知也在发生变化。数据就是资源，数据就是财富的理念迅速深入人心。一切事物背后都隐藏着未被发掘的数据。数据的真实价值就像漂浮在海洋中的冰山，第一眼只能看到冰山一角，绝大部分则隐藏在表面之下。如果你没有大数据的理念，很可能与很多有价值的数据失之交臂。判断数据的价值，需要考虑到未来它可能被使用的各种方式，而非仅仅考虑其目前的用途，要充分挖掘数据的潜在价值，才能在大数据时代立于不败之地。

## 二、数据可以预测未来

大数据的核心就是预测，通常被视为人工智能的一部分，它是把数学算法运用到海量的数据中来预测事情发生的可能性。比如，半个小时后出发到中关村，请根据当时的路况规划线路。这个问题估计会难倒所有导航软件，难点在于谁也不知道半个小时后中关村大街堵不堵，该不该绕道，但大数据可以。

百度推出大数据预测开放平台，主要是针对有需求的企业提供平台化的预测服务。比如，某个旅游城市的酒店希望预测未来一个月的入住量，以便提前规划房间定价、营销策略和人员分工等工作。酒店只要从其 CRM 系统中导出

过去两年每天的入住量数据，将其上传到百度大数据预测开放平台，并填写行业、地域、关键词信息，提交预测任务，百度大数据预测开放平台就会自动挖掘与酒店入住量相关的因素，如该酒店的百度搜索指数、微博热度、舆情、酒店附近人流量等指标，并结合入住量数据的季节性、周末效应、假日因素、中长期变化趋势，建立大数据预测模型，来准确预测该酒店未来一个月的入住量，从而酒店相关的部门可以对未来一定时间内的客流量做出相对准确的预估，并以此来合理安排相应的资源分配。

这些预测系统之所以能够成功，关键在于它们是建立在海量数据基础之上的。当今社会，随着智能移动设备的不断普及，新的网络基础设施特别是5G网络被引入，大量移动互联应用出现和发展，使得数以亿计的用户每时每刻都在产生数据。比如，使用微博、微信等社交软件，你每天的行为、爱好、交友圈子就被记录下来；登录购物网站购买商品，你的消费习惯，甚至财富能力都被记录下来。当人们的活动、决定、社会关系等都能够被数据记录，再融合物理和虚拟空间数据，个体、城市及社会活动信息就都能够全方位得以呈现了。

### 三、数据之间具有相关性

在小数据时代，先大胆地假设原因，再小心地求证，无疑是人们追寻真理的常用方法。这种由“果”寻“因”的思维方式曾经让人们不断发现和验证一些客观规律的存在，正因如此，才有了科技发达的今天，但是仍有其局限性。一方面，从结果出发去找原因，往往是先提出假设，然而一旦假设出现错误，那么之后所做的大量的求证工作将没有任何意义；另一方面，对于提出的假设，人们也常常选择一些表面上就已经存在或者可以看得出的规律的因素，然后去验证，这样就会漏掉许多表面上看似无关，实则有着深层次关联的、复杂的、让人难以发现的规律。

在大数据时代，我们要更加关注数据之间的相关关系，而不是因果关系。相关关系的核心是量化两个数据值之间的数理关系。相关关系强是指当一个数据值增加时，另一个数据值很有可能也会随之增加。相反，相关关系弱就意味着当一个数据值增加时，另一个数据值几乎不会发生变化。这种新的思维方式有助于我们发现以前看似无关的事件也可能存在一定联系，而这种联系也许蕴含着新的机遇。

例如，中英人寿保险有限公司（Aviva）是一家大型保险公司，它通过德勤咨询公司开发出一种预测模型，通过分析人们经常浏览的网站、经常收看的节目、收入估算等从任何网络上收集的反映生活方式的数据，用来找出更