KPMG

FROM DATA TO INSIGHTS

毕马威亚太区及中国主席**陶匡淳**先生

倾情作序推荐

洞见数据价值

大数据挖掘要案纪实

毕马威中国大数据团队◎著

企业与个人大数据简明读本

清華大學出版社

FROM DATA TO INSIGHTS

洞见数据价值

大数据挖掘思维与实践

FROM DATA
TO INSIGHTS

洞见数据价值

大数据挖掘要案纪实

毕马威中国大数据团队◎著

清華大學出版社
北京

内 容 简 介

本书是国际知名咨询公司毕马威的大数据团队的集大成之作，内容丰富，观点新颖，贴近大众生活、工作与学习实际场景，内容不拘泥于技术阐释，通过寓教于乐的方式，以丰富翔实的案例来解析大数据挖掘，尽量通过常见的场景来阐述数据的价值与意义。

第 1 章介绍大数据在银行业、征信业、审计、传统制造业、互联网行业、舆情监控、影视业、环保产业以及体育产业等多个领域的应用方案和前景。第 2 章重点介绍大数据分析在商业工作和营销推广中的作用。第 3 章介绍大数据挖掘过程中涉及的数据的前期准备工作，重点介绍数据准备工作的要点和诀窍。第 4 章结合业务、生活、娱乐，寓教于乐，介绍大数据的实际应用方式。附录部分介绍一位数据工作者的成长之路，向感兴趣的读者介绍从事数据工作应该具备的素质和掌握的技能。

本书可作为企业管理人员、营销主管、分析人员、IT 人员等理解大数据、应用大数据为企业创造价值的指引，同时，本书也可供统计学、应用数学及计算机专业学者和研究人员参考学习。

图书在版编目（CIP）数据

洞见数据价值：大数据挖掘要案纪实 / 毕马威中国大数据团队著. —北京：清华大学出版社，2018

ISBN 978-7-302-49180-4

Ⅰ. ①洞… Ⅱ. ①毕… Ⅲ. ①数据采集—研究 Ⅳ. ① TP274

中国版本图书馆 CIP 数据核字（2017）第 330716 号

责任编辑：秦 健
封面设计：李召霞
责任校对：胡伟民
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 刷 者：北京亿浓世纪彩色印刷有限公司
装 订 者：三河市金元印装有限公司
经 销：全国新华书店
开 本：148mm × 210mm 印 张：9.75 字 数：245 千字
版 次：2018 年 2 月第 1 版 印 次：2018 年 2 月第 1 次印刷
印 数：1 ～ 4000
定 价：69.00 元

产品编号：073790-01

业界推荐

Praises

谢邦昌教授

台北医学大学管理学院及大数据研究中心院长 / 主任

多年来我一直关注及致力于大数据在各领域的应用及其产生的巨大作用，而毕马威中国大数据团队的著作正是一本绝佳的案例集，它告诉我们数据挖掘如何为行业的最前沿带来价值，面向高维数据，从需求出发，使用丰富的方法，解决多样化的问题，可以看出作者深厚的理论功底和实践经验。一定可以让读者获益匪浅。

车品觉

红杉资本中国专家合伙人

大数据的概念从提出到落地经历了七年时间，作为分析、算法、人工智能的“新石油”，给国家和企业都带来了前所未有的机遇。我一直认为，每一种新技术的实践，最难的环节永远在于固有思想的改变，因此案例变得格外重要，好的案例起着启蒙作用。毕马威中国大数据团队的这本书正好为国内企业填补了包括互联网、金融、媒体、制造业等行业中“存管用”的实例。2018

年作为大数据下半场的开端，我相信大数据将会不断被各种创新终端所重新定义及重构。有趣的是，这种新能源还是在壮大中的婴儿而已。

张波教授

中国人民大学统计学院副院长

在人大统计学院执教多年，我希望高校始终保持着对商界、业界实际应用需求的敏感，因为只有在真实的数据和真实的场景中，理论才能以最快的速度驱动、革新。如今，数据科学的快速发展特别是大数据的热潮正在改变着统计学发展的方向，由我的学生带领的毕马威中国大数据团队的这本《洞见数据价值：大数据挖掘要案纪实》，既有学术的理论与严谨，也有业界的经验与热情。我愿意把这本书推荐给有志于探索数据价值的高校师生和希望用数据解决实际问题的各界精英。

刘政博士

SAS 中国研发中心总经理

这些取自“KPMG 大数据挖掘”微信公众号的文章是由经验丰富的一线咨询团队的成员奉献的、理论联系实际的范文。内容概括面广，针对性强，具有前瞻性，也有很强的实操性。很多文章都是实际经验的积累和创新，对于目前行业中缺少实际经验的数据分析人员来说，具有很好的指导和借鉴作用。

刘贤荣博士

中国建设银行数据管理部副总经理

数据已经成为移动互联时代企业的关键生产要素，无论是传统的制造行业还是新兴的生产性服务业，企业的设计、生产、营销、管理各个方面正与大数据高度融合，尤其是数据密度极高的银行业，大数据能力已经成为竞争的核心能力。古语云，“绝知此事要躬行”，毕马威中国大数据团队依托丰富的项目实战经验，将数据挖掘的专业知识用轻快而不失专业的语言和生动的案例呈现，在故事般愉快的阅读中深刻理解大数据的实际案例操作，读来是一种享受。

施奕明

前海征信副总经理，SAS 中文论坛创始人

第一次看到这个公众号，觉得它也就是繁星中的一颗，人来人往，沧海一粟，并没有太在意。三个月后，我惊讶地发现在并没有推手和炒作的情况下，身边的人都在转发和推荐这个公众号。于是我仔细阅读了每篇文章，确实能够感受到这群小伙伴们的热情、专业和坚持。正是在这种正能量的激励下，我也模仿和学习他们，创建了我们自己类似的公众号，于是在大数据挖掘的前进道路上不再孤单。从某种意义上，这些可爱的小伙伴是我的老师。强烈推荐大家订阅他们的文章，在大数据挖掘内容服务领域他们绝对是一流的先锋，我就是他们的忠实粉丝。衷心希望我们能一

起品读他们的这份不沾染任何商业气息的热情、专业与坚持，在当今这个浮躁的社会中这份珍贵显得如此与众不同。

张磊博士

前 SAS 软件（中国）有限公司首席咨询顾问

从数据到价值，是数据挖掘最令人着迷之处，而其间却要历经种种的曲折反复。毕马威中国大数据团队将理论和实际紧密联系起来，带你领略数据挖掘旅途中的一路风景……

常国珍博士

中银消费金融大数据与 AI 实验室主任

2013 年维克托的《大数据时代》风靡全国，由此带来的大数据喧嚣甚至超过了楼市。曾几何时，没有数据科学家光环的大数据工作者慎言笃行，透过数据洞察商业先机，默默为企业创造价值。本人有幸在大数据元年前与本书部分作者相识、交流，更有幸成为本书的早期读者。本书以“纪实”的形式再现了鲜活的大数据案例，既有知识，又有典故，让作为读者的我收获颇丰。

蒋顺利

北京仁科互动网络技术有限公司市场副总裁

我可能是“KPMG 大数据挖掘”微信公众号最早的粉丝之一。从 2015 年 10 月开始，几乎每篇文章都会仔细阅读。它里

面的文章内容丰富，涵盖的范围广，包括银行业、制造业、互联网、影视等不同的行业，以及营销、风险、反欺诈、分析等不同的应用场景。虽包罗万象，但万变不离其宗，都紧密围绕着“大数据挖掘”这个热门的主题。虽是技术型微信公众号，但不故作高深，文字生动有趣，可读性强。无论是回顾历史还是预测未来，与大数据自然结合，娓娓道来，给人启发。欣闻微信公众号上的文章要结集成书，可喜可贺，也是众多读者的福音。无论是对大数据感兴趣的初学者，还是资深的数据科学家，这都是一本非常难得的大数据读物。

序言 Preface

越是基本的理念，往往越能成为时代的标志、价值的度量，数据正是如此。

2020 年，全球数据总量预计将超过 44 千万亿兆字节，数据之大，正如我们日常呼吸的空气一般不可缺少。这意味着世界上的一切都在产生数据，一切都在使用数据；万事万物都可以数据化，而数据也成为价值的新载体。大数据时代——也就是说，这样一个以数据衡量价值的时代，我们已经置身其中。

今日，“大数据”已不仅仅是一个新兴概念。国家、机构、企业都拥有数据，也急需使用数据，而物联网等理念的兴起，云计算等技术的应用，让我们能够帮助他们探索、挖掘、利用数据中的价值，能够存储、流通、关联、交换、使用大数据，开发每一个环节中的能量。大数据是资源，是一座亟待开掘的金矿。

毫无疑问，大数据拥有巨大的商业潜力和创造力，而这样的潜力和创造力，当然也属于毕马威中国大数据团队。

本书正是毕马威中国大数据团队的心血之作，汇聚了其微信公众号中的精品文章。该大数据团队建立几年以来，我一直非常欣赏团队的业务能力和业绩表现，同样，文如其人，他们在这本书中展现出的实力、动力和潜力也令人赞叹。

学术成果一向是新兴行业的驱动力，而业务能力是学术与实际工作的衔接点：大数据团队是一个集合了这两方面高端人才的

团队，本书中的作品，专注学术、扎根业务，也体现了不俗的行业发展眼光。何况，写这样的文章，一篇或许容易，但能在繁忙的日常工作之余坚持一年之久则殊为不易；如果不是对数据领域有深切的热爱，又怎么会有这样的动力呢？这本著作见证了团队的成长，更是团队潜力的明证。我相信，本书的读者们也一定会有同样的感觉。

大数据时代，谁掌握了数据，谁就掌握了价值，也掌握了世界的节奏。面对数据之潮，毕马威中国一直力求一马当先，而大数据团队正是公司内的行业先驱。“大鹏一日同风起，抟摇直上九万里”，我不由得心生期待，想要看看这样一个出色的团队还能给我和读者们带来怎样的惊喜。

陶匡淳

毕马威亚太区及中国主席

目录 CONTENTS

01

第1章

大数据在各行各业

FROM

DATA

TO

INSIGHTS