



管理数据分析

原理、方法、工具及实践

MANAGERIAL ANALYTICS:

AN APPLIED GUIDE TO PRINCIPLES,
METHODS, TOOLS, AND BEST PRACTICES

【美】 麦克·沃森 (Michael Watson) 著
德里克·内尔森 (Derek Nelson) 著
王忠玉 王琼 译

管理数据分析

——原理、方法、工具及实践

[美] 麦克·沃森 (Michael Watson) 著
德里克·内尔森 (Derek Nelson)
王忠玉 王琼 译



机械工业出版社

Authorized translation from the English language edition, entitled *Managerial Analytics: An Applied Guide to Principles, Methods, Tools, and Best Practices*, 9780133407426, by Michael Watson, Derek Nelson, published by Pearson Education, Inc, Copyright © 2014 by Michael Watson and Derek Nelson.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

CHINESE SIMPLIFIED language edition published by PEARSON EDUCATION ASIA LTD. , and CHINA MACHINE PRESS Copyright © 2017.

本书中文简体字版由培生教育出版公司授权机械工业出版社合作出版, 未经出版者书面许可, 不得以任何形式复制或抄袭本书的任何部分。

本书封面贴有 Pearson Education (培生教育出版集团) 激光防伪标签。无标签者不得销售。
北京市版权局著作权合同登记 图字: 01-2015-5831 号。

图书在版编目(CIP)数据

管理数据分析: 原理、方法、工具及实践/(美) 麦克·沃森 (Michael Watson), (美) 德里克·内尔森 (Derek Nelson) 著; 王忠玉, 王琼译. —北京: 机械工业出版社, 2017. 4

书名原文: *Managerial Analytics: An Applied Guide to Principles, Methods, Tools, and Best Practices*

ISBN 978-7-111-56697-7

I. ①管… II. ①麦… ②德… ③王… ④王… III. ①数据处理—应用—管理学—分析方法 IV. ①C931-39

中国版本图书馆 CIP 数据核字(2017)第 089790 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑: 易敏 责任编辑: 易敏 付鑫宇

责任校对: 刘岚 封面设计: 鞠杨

责任印制: 李昂

三河市国英印务有限公司印刷

2017 年 8 月第 1 版第 1 次印刷

169mm × 239mm · 11.5 印张 · 212 千字

标准书号: ISBN 978-7-111-56697-7

定价: 39.80 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

电话服务

网络服务

服务咨询热线: 010-88361066 机工官网: www.cmpbook.com

读者购书热线: 010-68326294 机工官博: weibo.com/cmp1952

010-88379203 金书网: www.golden-book.com

封面无防伪标均为盗版

教育服务网: www.cmpedu.com

想要利用大数据和数据分析获得利润吗？本书是你理想的首选资源。无论你的行业或管理角色怎样，本书都将有助于你以最快的速度开始，并迅速产生价值。

现在，数据分析可以帮助组织做从盈利到拯救生命的任何事情。假如能正确运用它，你的组织将可能凭借数据分析获得巨大的机会和价值。

本书将展示如何做数据分析。本书是专门为每一位想要从数据分析中获得利润，但不知道从哪里开始的经理，每一位想要向客户提供更好的结果的顾问和软件供应商，每一位负责评估分析产品或供应商的专家而设计的。

本书运用现实事例阐明了三种分析形式：描述数据和可视化数据的描述性数据分析，识别趋势和关系的预测性数据分析，利用已知信息和预期目标来优化决策的规范性数据分析。

本书介绍了当今最好的数据分析工具和技术，提供了足够的技术细节来帮助你做出最好的选择。本书适合于学习和研究商业数据分析、大数据在企业中应用的管理人员、高校学生、政府工作者，帮助他们了解如何用数据来解决问题。

童书馆

致我的妻子，克里斯汀，
感谢她在我写作此书时的全力支持

麦克·沃森

致我的妻子，布丽奇特，
她像个小女孩一样梦想拥有一本献给她的书 ☺

德里克·内尔森

本书荣获的赞誉

“随着大量的、丰富多彩的数据源不断涌现的新时代来临，关于数据分析的书籍日益增多，其中绝大多数的书籍，要么专注于分析技术层面，要么聚焦于基础商业流程及商业分析层面。当一位管理者想要学习基本技术及其应用时，势必要师从于一位拥有良好沟通能力的数据科学家。本书为管理者提供了一个自己学习相关概念的机会，因此，本书应该摆在那些领导和从事管理数据分析工作的管理人员的书架上。从非技术角度来讲，本书涵盖了数据分析最重要的方法论及概念。每一种方法论都恰当地结合举例说明和实用案例，理解它们不需要任何技术知识。如果我想要在数据分析领域创业，那么这将是我在每天清晨阅读的第一本书。”

——迭戈·克拉比耶 (Diego Klabjan)，
工业工程与管理学教授，
数据分析理学硕士项目主任，
西北大学

“这本《管理数据分析》针对数据分析给出了极好的介绍和概述，对于刚接触这一领域的人来说，非常容易理解，而对于有更多知识背景的读者来说，则提供了一个十分有用的框架。本书是从实践者的角度来写作的，书里面充满了简洁而相关的例子。作者表述了当前数据分析工具的可能边界，并引导读者提出好的问题，避免落入常见陷阱。当读者管理自己的数据分析项目时，知道如何识别隐含的假设。”

——迈克尔·费米尔 (Michael Freimer) 博士，
DemandSignal 首席科学家

“对于不同的读者来说，‘数据分析’术语意味着不同的事物。在连接市场营销的大肆宣传与技术细节之间的缺口方面，沃森和内尔森筑起了桥梁，从而使读者更容易评估以数据分析为基础的解决方案，更好地理解这些方案的潜力。”

——欧文·勒斯蒂格 (Irv Lustig) 博士，
最优化数学软件、商业数据分析和数学科学经理，
IBM 研究中心

关于作者

麦克·沃森现为欧普克斯数据分析机构 (Opex Analytics) 合伙人, 西北大学兼职教授。在欧普克斯数据分析机构中, 他致力于为公司带来崭新的数据分析解决方案。在到欧普克斯数据分析机构之前, 他曾任 IBM 公司中 ILOG 供应链和最优化小组负责人。在西北大学麦考密克学院, 他为工程管理硕士生 (MEM) 讲授运营管理学和管理数据分析, 为西北大学数据分析理学硕士讲授最优化理论。麦克拥有西北大学工业工程与管理科学理学学士和哲学博士学位。

德里克·内尔森现为 OPS Rules 公司高级主任和西北大学兼职教授。在 OPS Rules 公司, 德里克利用数据分析帮助公司提高运营绩效。在此之前, 德里克为 Logic Tools、ILOG 和 IBM 在最优化和供应链软件方面提供咨询服务、产品管理服务和销售服务。在西北大学, 德里克曾为工业工程与管理科学系本科生讲授服务运营管理, 不久^①将任教于工程管理硕士 (MEM) 项目。德里克拥有康奈尔大学运筹学硕士学位。

^① 原书出版时间为 2014 年, 此处说明的是当时的状态。

序 言

第一次听到“数据分析”成为流行语时，我们并没有受到很大的触动。毕竟，我们的职业就是做数据分析研究。我们在本科和研究生阶段，都专注于类似数学最优化、概率论和统计学的专题内容。我们的职业就是整天和数据打交道，分析数据，创建数据库和报表系统，利用各种工具（包括电子表格、数据库、报表工具、统计工具、最优化工具和专业工程工具）来测试、操作数据，同时以有意义的方式理解、认识数据，具体地说，是使我们公司和曾合作过的其他公司据此做出更明智的商业决策。

我们从来没有给自己所做的研究贴上“数据分析”的标签，但是如果有人真的在研究数据分析，那就是我们。

广受欢迎的数据分析术语，人气越来越高涨。这些词似乎很快就火了起来。它被人们频繁地运用，突然出现在各种市场营销活动中，甚至一些公司重新将自己标榜为“数据分析”专家。不可否认，最初我们陷入了常见的陷阱之中，抵制这个词语，认为这不过是“老树发新芽”。我们认为自己不应当被愚弄，数据分析只是我们以前称呼的数学而已。

可是，我们不久后发现，这不仅仅是一个新名词，而且是一个蓬勃发展的趋势。并且，这样的势头能迅速让我们放弃抵抗。一个软件供应商说，如果你要实施公司的解决方案，就必须会运用数据分析；一些硅谷的公司说，每家公司都必须运用数据分析，否则就不会成功。

当我们阅读了更多的文章，采访了更多的人后，我们发现，数据分析这个术语正在被各式各样的人以截然不同的方式运用着。我们也看到，一些软件公司和某些产品看起来绑架了数据分析术语，而且运用它时，仿佛除了他们所做的工作以外任何东西都不是数据分析。从专业角度来看，这和我们的认知不一样。尽管我们知道，我们所做的内容是数据分析，但是我们担心，正是因为有许多人运用这个术语时出现前后不一致的情况，所以人们很难理解它。

渐渐地，我们承认除了自己研究颇深的专题外，或许有更多其他的故事可以研究。关于数据分析，确定曾存在过许多新的领域。不过，在开始探索之前，我们不可以先限定它到底是什么。

换句话说，数据分析领域没有一个得到公认的好的定义和描述。如果

有人想知道，数据分析发展的趋势如何，我们很难给出一个完整的答案。让事情变得更加困惑的是，大数据迅速出现，并成为数据分析的同义词。

我们认为，如果我们处于困惑之中，那么其他人也会如此。当管理者现在被要求运用数据分析和大数据做更多的事情时，他们就需要准确理解那些知识是什么。所以，我们决定写作这本书，目的是帮助管理者和分析师更好地认识数据分析及其发展趋势。

这本书是针对想要知道数据分析及其发展趋势是什么的人而专门设计的。写作本书的目的是帮助需要利用数据分析来提升自己的组织机构，让它们看清大肆宣传的广告，找到清晰的定义，从而更好地实施数据分析的解决方案；本书也面向咨询顾问和软件供应商，以使他们能够帮助其客户取得卓越成果；面向那些需要评估分析解决方案，但可能又不知道从哪里开始研究问题的人。我们希望看到，更多来自各类不同组织机构的人能更好地运用数据分析。我们从管理者需要了解数据分析的视角撰写这本书。由于数据分析是一门技术性很强的学科，所以我们引入了某些技术细节。但是，我们这样做，只是为了帮助管理者获得更好的直觉和洞察力。可是，数据分析的领域非常宽泛，以至于很难在一本书中做到面面俱到，因此我们不得不挑选一些专题并控制深度。当然，不同的人需求可能是不同的，我们尽力提供足够的信息和参考资料，让读者从一个更好的视角来探索更加感兴趣的专题。

我们希望各种各样的人都能从本书中获益：管理者能从更好的视角评估和运营项目；某些领域的专家能知道他们的专业如何更好地应用于其他专业。我们还希望本书能够激发或给予读者一些好的想法，以使所在的组织机构凸显出重要的差异。如果运用得当，数据分析将无所不能，小到帮助企业建立新的经营策略并节省资金，大到帮助医疗机构拯救生命。

目 录

本书荣获的赞誉

关于作者

序 言

第1部分 概 述

第1章 什么是管理数据分析 / 2

- 1.1 关于数据分析含义的疑惑 / 2
- 1.2 什么是数据分析 / 3
 - 1.2.1 运用描述性数据分析的事例 / 5
 - 1.2.2 运用预测性数据分析的事例 / 7
 - 1.2.3 运用规范性数据分析的事例 / 8
- 1.3 什么是新的内容 / 10
- 1.4 哪种数据分析最好 / 12
- 1.5 什么是管理数据分析 / 14

第2章 什么在推动数据分析蓬勃发展 / 19

- 2.1 数据是数据分析的原料 / 19
 - 2.1.1 什么是大数据 / 21
 - 2.1.2 大数据能替代科学吗 / 25
 - 2.1.3 在缺少大数据的条件下可以做数据分析吗 / 26
- 2.2 检验促进数据分析的发展 / 27

第3章 数据分析思维方式 / 30

- 3.1 管理数据盲 / 30
 - 3.1.1 计算能力的错觉 / 31
 - 3.1.2 过滤谬论 / 33
 - 3.1.3 一点改进的益处 / 34
- 3.2 数据分析是一种思维方式 / 35
 - 3.2.1 80/20 法则 / 37
 - 3.2.2 将可变性纳入到数据分析中 / 38
 - 3.2.3 不能仅使用会计数据 / 39
- 3.3 透彻思考数据 / 41
 - 3.3.1 并非所有的数字都是数据 / 41

- 3.3.2 能否信任数据 / 42
- 3.3.3 应该对数据做简单的测试 / 45
- 3.3.4 就地取材 / 46
- 3.4 数据科学家的兴起 / 48

第2部分 数据分析工具

第4章 机器学习 / 52

- 4.1 什么是机器学习 / 52
- 4.2 监督机器学习算法 / 55
 - 4.2.1 分类与k最近邻算法 / 56
 - 4.2.2 分类与决策树 / 58
 - 4.2.3 推荐系统 / 61
 - 4.2.4 回归分析 / 62
- 4.3 无监督机器学习算法 / 66
 - 4.3.1 聚类分析与k均值算法 / 66
 - 4.3.2 关联规则与购物篮算法 / 69
- 4.4 对模型过度拟合与欠拟合的评注 / 70
- 4.5 其他机器学习算法与总结 / 71

第5章 描述性数据分析 / 74

- 5.1 基于数据库的描述性数据分析 / 76
 - 5.1.1 数据库基础 / 77
 - 5.1.2 数据建模 / 82
 - 5.1.3 学习SQL语言, 更好地认识数据 / 84
 - 5.1.4 什么是NO-SQL / 86
 - 5.1.5 结构化数据与非结构化数据 / 88
 - 5.1.6 数据仓库与数据集市 / 90
 - 5.1.7 仪表盘和平衡计分卡: 实用的及时报告 / 93
 - 5.1.8 运用OLAP和数据立方体拓展描述性数据分析 / 95
 - 5.1.9 何时打破关系数据库的规则 / 97
 - 5.1.10 实时数据与自动预警 / 98
- 5.2 运用数据可视化做描述性数据分析 / 100
- 5.3 运用描述性统计做描述性数据分析 / 107
- 5.4 运用机器学习做描述性数据分析 / 109

第6章 预测性数据分析 / 111

- 6.1 回归预测 / 111
- 6.2 机器学习与集成模型 / 112

6.3 A/B 检测 / 114

6.4 模拟 / 118

第7章 案例分析：《点球成金》与最优化 / 123

第8章 规范性数据分析（又称最优化） / 128

8.1 什么是最优化 / 131

8.2 最优化 = 目标、约束条件、选择 + 数据 / 131

8.2.1 目标 / 132

8.2.2 约束条件 / 136

8.2.3 选择 / 140

8.2.4 数据 / 141

8.3 TLC + D 应用：每个人都爱吃比萨 / 143

8.3.1 目标 / 144

8.3.2 约束条件 / 144

8.3.3 选择 / 144

8.3.4 数据 / 144

8.3.5 数学公式 / 144

8.4 最优化算法的类型 / 145

8.4.1 精确算法 / 146

8.4.2 近似算法 / 147

8.4.3 启发式算法 / 148

8.4.4 假设分析 / 151

第3部分 结 论

第9章 收益管理 / 154

第10章 实施数据分析的终极技巧 / 161

10.1 它只是弓箭手，而不是弓箭本身 / 161

10.2 总结 / 162

非传统参考书目和扩展阅读 / 163

译后记 / 167

第 1 部分 概 述

第 1 章 什么是管理数据分析

1.1 关于数据分析含义的疑惑

数据分析领域（及其常用同义词大数据）已经引发了许多企业管理者的无限遐想。

有些公司经常公开地声称：他们热衷于数据分析。通用电气（General Electric）最近宣布，为了充分利用自己的工业机器所产生的数据而大举进军“数据分析”。¹IBM 和俄亥俄州立大学宣布成立数据分析研发中心，预计雇用多达 500 人来做数据分析研究工作。而谷歌则大力宣传自己内部的高级搜索功能的数据分析，并为其他公司提供跟踪自己网站的服务。这份所罗列的公司一览表及他们所新公布的数据分析能力，将会继续更新发展。

对咨询公司和软件公司数据分析服务和产品所进行的推销越激动，越是会心存疑惑地认为，这些公司除了他们的营销信息里面包括“数据分析”这个词之外，并没有改变任何东西。但是，这样的怀疑观点并不能反驳如下的事实：很明显，对数据分析的需求是相当多，否则这些公司不会尝试在第一时间跟进。

数据分析这个术语甚至进入了大众媒体。最常看到的情况是，它出现在大众媒体的商业部分甚至是报纸的首页。电影《点球成金》（基于同名书改编的电影）向我们展示了，聪明人如何使用统计数据来帮助奥克兰运动家职业棒球队以一种低成本方式来创建团队并赢得胜利，这可能是关于数据分析最能吸引公众眼球的精彩力作。那些密切关注政治的人士，也可能读到过美国前总统奥巴马竞选连任的原因就是大量运用数据分析来精准定位最有可能投票给他的选民。这些进军主流文化（体育和政治）的尝试促进了公众对数据分析的兴趣。

为了解决雇主的招聘需求和满足学生的就业意愿，一些大学开始提供数据分析相关的专业学位。你很可能已经在主流出版物中看到了关于对数据分析技能市场需求的探讨。²

可是，数据分析到底是什么呢？公司和记者更愿意使用这个词，而不是告诉你它是什么。倘若没有定义，你怎样知道你所做的就是数据分析呢？

你如何将数据分析应用于你的具体情况呢？你怎样知道，你的供应商确实卖给你一个数据分析解决方案呢？

然而，每当一个词具有如此好的内涵却难以给出定义时，就会出现那样的风险：当下一个流行词出现时，前一个流行词将被人们所遗忘。当首席执行官想要更多的数据分析项目时，经理就将这个数据分析词语放入到他们所有项目的标题上，此时本质上并没有改变他们所做的任何工作。供货商意识到，他们可以用数据分析来重新命名他们一直在做的事物，进而会卖得更好。如果某个批次的产品与数据分析这个词语有联系，则出售那些产品的公司就没有动力来澄清，如果它对你有利，为什么还要澄清数据分析这个术语呢？吸引人们这样做的部分原因在于，对很多人来讲，数据分析这个词语具有神秘色彩，而这使其听起来很复杂，进而带来了巨大回报，所以，我们最好这样做！

如果你想要正确地运用数据分析，那么如何知道从哪里开始？如何知道在何时、何地利用数据分析可以获得价值？实施某个数据分析项目、或者某个方案是否会比其他的项目或方案更好？

与数据分析同时出现使用的大数据这个术语，无疑增加了理解上的混乱。大数据术语也是一个没有很明确定义的术语。例如，大数据和大数据集有怎样的差异呢？还有，数据分析是否只适用于大数据？数据分析和大数据的意思是一样的吗？

这本书将为你排除困惑，并清晰地给出数据分析和大数据的定义。你将会看到，由这个新定义所产生的领域，因此你将对这样的术语有一个深刻的理解。

我们认为数据分析绝不仅仅是一个流行词汇或一种时尚潮流。数据分析经由有能力的人出色地实施，将会为你的公司或组织机构带来巨大的价值。数据分析适用于不同形态、不同规模的公司和组织机构。数据分析几乎适用于各种大公司、小公司，非营利组织机构和教育研究所，医疗保健和医药机构、政府机关、科学研究所、执法部门和军队。数据分析项目可以由组织机构的负责人、部门经理，甚至单独个体来牵头执行。但是，数据分析只有在你和你的组织的人知道它是什么，对它的内容可以清晰地沟通，并能正确运用它时，才能带来价值。

1.2 什么是数据分析

数据分析这个词由来已久，可是，现在是什么又让它流行起来呢？

硅谷的互联网公司起到了推波助澜的作用。这些公司运用“数据分析”这一术语意指跟踪谁在点击你的网站、他们访问了哪个页面、他们在买什么，等等。但是，很难想象，如果它的性能只是跟踪网站信息，那么它是如何迅速扩散到如此地步的。

销售报告系统（用行业术语来讲就是商业智能软件）的公司对数据分析这个术语的传播所起到的作用也不可小觑。这些公司声称，数据分析能以更简单和有效的方法来报告公司的数据。然而，如果它只是提供报告而已，那么数据分析也不会如此受欢迎。

我们认为，当哈佛商业评论发表了一篇名为“数据分析竞争法”的文章时，数据分析就迈入了主流的商业术语的范畴。此文是由托马斯·达文波特（Thomas Davenport）于2006年1月发表（同时出版了同名书籍）³的。在这篇文章中，达文波特强调像亚马逊、万豪、哈拉斯娱乐公司以及第一资本金融公司是通过在一系列活动中开展各种行业强度分析，取得了所在领域的主导地位。

对于数据分析术语来说，这是一个引爆点。实际上，这篇文章表明，人们应用数据分析可以解决更广泛的问题。它不只是一个细分领域（例如跟踪网站访问者，或创建更好的报告）。它的益处远不止于此。

但是，我们仍然没有真正弄明白，数据分析的定义究竟是什么。达文波特的文章给出了企业解决具体问题的例子。例如，万豪酒店利用数据分析来设定房间的最优价格，而第一资本使用数据分析来测验不同的价格、促销方式和捆绑式服务吸引什么样的客户群体。但文章中只是将数据分析定义为“收集、分析和使用数据。”换句话说，依据这篇文章，数据分析就是通过分析数据做出更好的决策。这个定义，尽管是正确的和严谨的，但是没有给予更多的指导。管理者不总是谈论如何用数据来做决策吗？是的，可是如今可以利用的数据量更为庞大。但是，仅仅是利用大量数据来做出更多的决策，这件事并不能帮助我们定义数据分析。

达文波特文章的贡献并不在于它对分析的定义。相反，这篇文章开启了数据分析的蓬勃发展运动。也就是说，它提出了如下的想法：许多人利用各种不同的工具来解决众多问题，而所有这些工具都被称为“数据分析”。

这篇文章将数据分析的思想看成是一个探索的领域（如同计算机科学或化学领域）。作为佐证的是，许多大学已经开始提供数据分析专业学位。致力于数据分析的专业机构，诸如 INFORMS，也开始促进和不断完善这一领域。⁴

来自于学术界的学者和认真思考的商界思想家给出了数据分析的定义。当人们利用数据更好地做决策时，这个定义会将各种不同的目标进行细化分类。⁵我们将沿用数据分析的这一新兴定义，而且本书自始至终这样使用。综上所述，以下是数据分析的定义：

数据分析是利用数据获得洞察力，帮助人们更好地做决策的学科集合。它由下面三个分支组成：描述性数据分析，提供数据描述，报告数据并对数据进行可视化处理；预测性数据分析，利用数据预测趋势，识别数据关系；规范性（或称指导性）数据分析，根据所拥有的已知数据，依据未来所希冀的方向，为执行一系列的最佳决策提供指导方针。

换句话说，描述性数据分析的目的是对“已经发生什么”和“正在发生什么”提供一个认识理解；预测性数据分析的目的是告诉人们，将有可能发生什么；而规范性数据分析的目的是告诉人们如何应对。数据分析的这三个分支每一个都可以进一步细分，在本书后面将阐述这一点，同时各种不同的工具和技术都可应用于每一个分支。

数据分析的这个定义将经受住时间的检验。这个定义不仅充分定义了这个术语的意义，同时也允许其未来进一步发展。新的细分分支很可能会出现，新的工具也将得到拓展，而且人们将会有新的见解。这个定义为人们提供了理解和安排那些新变化的方法。

本书将对数据分析的这些领域中的每一个分支进行探究，提升读者的洞察力。请记住，虽然短短的几段内容可能涵盖数据分析的某个分支或某个分析工具，但这却有可能是一个大学的某些系、一些专业机构或一些公司致力于研究的分支领域或工具。我们绝不可以低估这些议题的重要性。

为了进一步帮助理解数据分析的分支内容，让我们来看看几个例子。

1.2.1 运用描述性数据分析的事例

最早的名副其实的描述性数据分析实例就是，在1854年伦敦致命性疾病霍乱暴发时期，约翰·斯诺所做的研究工作。当时虽然有许多现成的数据可以利用，但是这些数据却无助于人们控制疫情。因此，他决定将死亡的情况在地图上标出，以便查看是否能获得新的启发（参看图1.1）。在这张地图上，小黑点代表死于霍乱的人所在的住所。当斯诺和居住于伦敦的其他人看到地图时，他们清楚地发现，这些代表死亡的小黑点都聚集在某个水泵附近⁶。这种处理数据的方法，帮助人们将霍乱暴发的源头锁定了在特定的抽水泵上。