

100%

内容丰富、权威

使用Access超强的分析能力分析数据

使用自定义的视图更好地展示数据

掌握数据分析的基本技能

使用高级分析技术实现数据分析自动化

宝典丛书

100万

Microsoft Access Data Analysis

[美] Michael Alexander 著

梁普选 刘玉芬 等译

Access 数据分析

宝典

可在网址:

[www.wiley.com/
go/accessdataanalysis](http://www.wiley.com/go/accessdataanalysis)
找到本书所用的示例。



电子工业出版社
Publishing House of Electronics Industry
<http://www.phei.com.cn>

宝典丛书

Access 数据分析宝典

Microsoft Access Data Analysis

[美] Michael Alexander 著

梁普选 刘玉芬 等译

电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING

内 容 简 介

Excel被认为是进行数据分析的首选平台,但是,它缺乏分析过程的透明度,难以分析大规模的数据。而 Access 可避免这些不足,被认为是下一代数据分析工具。

本书从数据分析员的角度介绍了 Access 强大的数据分析能力,涉及的内容包括 Access 的基本知识、Access 的基本和高级分析技术、数据分析的自动化等。本书所介绍的内容始终以数据分析为中心,自始至终贯彻了“数据与展示分离”这一业界的理念,不仅能够使读者养成良好的数据分析习惯,而且能够培养良好的数据分析实战经验。本书用大量的示例一步一步地讲解数据分析方法,这些示例全部来自真实世界,不仅能够巩固所学内容,而且可帮助读者把所学内容快速应用到实际数据分析中。

本书内容全面深入,适合对数据分析感兴趣的读者使用。



Copyright© 2006 by Publishing House of Electronics Industry. Original English language edition copyright© 2006 by John Wiley & Sons, Inc. All rights reserved including the right of reproduction in whole or in part in any form. This translation published by John Wiley & Sons, Inc. 本书中文简体本专有翻译出版版权由美国 John Wiley & Sons, Inc. 授予电子工业出版社及其所属华信卓越公司。未经许可,不得以任何手段和形式复制或抄袭本书内容。该专有出版版权受法律保护,侵权必究。

版权贸易合同登记号 图字: 01-2006-0819

图书在版编目(CIP)数据

Access 数据分析宝典 / (美)亚历山大(Alexander, M.)著;梁普选,刘玉芬等译.
—北京:电子工业出版社,2006.8

(宝典丛书)

书名原文: Microsoft Access Data Analysis

ISBN 7-121-03086-1

I. A... II. ①亚...②梁...③刘... III. 关系数据库-数据库管理系统, Access - 系统分析 IV. TP311.1.138

中国版本图书馆 CIP 数据核字(2006)第 096847 号

责任编辑:梁卫红

印 刷:北京东光印刷厂

装 订:三河鹏成印业有限公司

出版发行:电子工业出版社

北京市海淀区万寿路 173 信箱 邮编:100036

经 销:各地新华书店

开 本:787 × 1092 1 / 16 印张:22.5 字数:576 千字

印 次:2006 年 8 月第 1 次印刷

定 价:39.00 元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系电话:(010)68279077; 邮购电话:(010)88254888。

质量投诉请发邮件至 zlt@phei.com.cn。

服务热线:(010)88258888。

新的一天

今年对于“宝典”丛书来说确实是非同寻常的一年。今年，“宝典”丛书在中国图书市场的发行量累计达到了100万册；今年，“宝典”丛书登陆中国市场已经10年。因而，对于一直关爱这套丛书的读者，对于许许多多为这套丛书曾经付出过辛劳和汗水的翻译人员，对于为这套丛书默默耕耘的工作人员来说，今年都是那么不同凡响的一年，也是我们值得骄傲的一年。

回首这10年，我们有欣慰、有遗憾。俗话说“不经历风雨，怎么见彩虹！”，终于我们等到了这一天。而今我们不仅在销量上取得不俗的成绩，而且成为电子工业出版社的著名品牌之一。

来自美国著名的出版公司Wiley的这套“宝典”丛书，秉承了Wiley图书一贯的水准，内容全面、权威。在世界各地51个国家被译为31种文字，拥有几百万读者。

其作者都是某个计算机专业领域的专家、教授，有些还是某软件的特约测试者。比如Deke McClelland、Alan Simpson和Ellen Finkelstein等知名畅销计算机图书作家，在相关领域都具有很高的声望。中文版“宝典”的译者均经过严格筛选，他们大都是来自于高等院校的教授、学者，计算机领域的高手，不但具有高深的专业知识，同时在英语方面也颇有造诣。我们的编辑队伍，同样是来自于计算机专业的高素质人才。通过这种严格的层层把关，相信最终奉献给读者的将是一部部精品。

今后我们这些“宝典”人还会将继续秉承“宝典”丛书一贯的“权威、全面、精益求精”的宗旨，力争使每一本书成为您探索计算机领域奥秘的“宝典”。

现在，让我们抛开过去，展望未来，共同期待一个新的开始，新的一天；共同迎接新的100万册，迎接又一个崭新的10年。

电子工业出版社
2004年·仲夏

译者序

在当今商务世界，企业和政府部门每年都要对大量的数据进行分析。传统的数据分析工具Excel因为能够快速完成交互式分析而被许多人认为是数据分析的首选平台。但是，Excel缺乏分析过程的透明度，不能实现数据与展示分离，难以分析大规模的数据集，这些缺陷随着未来数据量的增长而越发明显。Access避免了Excel的这些不足之处。许多工业界的专家认为Access是更好的数据分析工具。面对日益增长的数据量，对数据分析人员来说，Access是优秀的下一代分析工具。

本书从数据分析员的角度介绍了Access强大的数据分析能力，涉及的内容包括Access的基本知识、Access的基本和高级分析技术、数据分析的自动化等。本书所介绍的内容始终以数据分析为中心，自始至终贯彻了“数据与展示分离”这一业界的理念，不仅能够使读者养成良好的数据分析习惯，而且能够培养良好的数据分析实战经验。相信本书能够很快帮助读者掌握Access数据分析技术。

本书用大量的示例一步一步地讲解数据分析方法，这些示例全部来自真实世界，不仅能够巩固所学内容，而且可帮助读者把所学内容快速应用到实际数据分析中。

参加本书翻译的有：杨小蓉、梁普选、刘玉芬、关山、王连、王强、李斌、韩育、付宇光、刘明、李强、丰金兰、刘宝忠、苏飞、唐玲、毛波等。由于时间和水平所限，书中错误在所难免，望读者不吝赐教。

译者

2006年6月

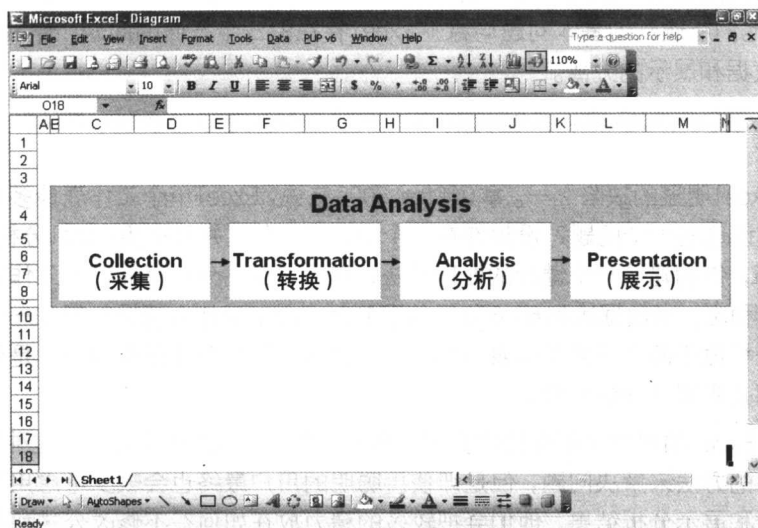
前言

数据分析定义

如果随机问一个人什么是数据分析？多数人会说它是为获得某个问题的答案而对数据进行计算和汇总的过程。从某种意义上说，他们是正确的。但是他们所描述的这些仅仅是数据分析过程中的一小部分。

例如，如果要分析公司上个月的销售收入，那么为完成该分析该做些什么呢？难道仅仅是计算和汇总上个月的销售数据吗？还有，从哪里获取这些销售数据？数据存储在哪里？在获取数据时有必要清理数据吗？如何提供分析结果，按周、日，还是按区域提供结果？问题的关键在于，数据分析过程不仅仅包括数据的计算和汇总。

对数据分析更具有代表性的定义是，为提供有意义的结论而对数据进行系统收集、变换和分析的过程。为更好地理解这个概念，可认为数据分析过程包含如下四个基本动作：采集、转换、分析和展示。



构成数据分析的四个基本动作

- ◆ **采集**：包括数据的收集和存储，也就是说，从何处获得数据，如何检索数据，怎样保存数据，以及在进行某种分析时怎样访问数据。
- ◆ **转换**：保证数据结构统一、无冗余、稳定的过程。这通常需要建立表结构、整理文本、删除空白和使数据字段标准化。
- ◆ **分析**：是指从整体上研究数据源的数据构成以及它们之间的关系。在计算、汇总、分类、比较、对比、检验或测试数据时都需要分析数据。

- ◆ **展示:** 怎样把分析结果提供给特定的用户使用,即选择什么方式来显示结果。展示分析结果需要考虑下面几个方面:包括所要使用的平台、可见程度以及给予用户修改视图的自由度。

在阅读本书时记住这四个基本要素将有助于读者认识每章所要达到的目标以及怎样用Access来加强数据分析能力。

为什么用 Microsoft Access 进行数据分析

如果问及日常数据分析使用何种软件工具,大多数人会回答Excel。事实上,如果在Amazon.com的搜索功能中输入“data analyze”,则会找到大量有关如何用Excel分析数据的书籍。既然这么多人都认为应该用Excel进行数据分析,那么为什么还要把Access用于数据分析呢?答案是避免Excel的局限性和存在的问题。

用 Excel 分析数据会在何处出错

我必须谦虚地承认,我已经使用Excel多年并且每天还在继续使用它。Excel被认为是进行数据分析的首选平台。在当今的商务世界中,任何不了解Excel的人毫无疑问会掩饰这一事实。Excel能快速完成交互式分析,这使它在工业界独一无二。但是,在给经理、会计师、分析员等提供了多年的咨询经验后,我们不得不接受这样一个简单的事实,他们的分析需求已经超出了Excel的能力。他们所遇到的基本问题均源于Excel存在的三个方面的问题:伸缩性、分析过程的透明性以及数据和展示分离。

伸缩性

伸缩性是Excel明显的缺陷之一。拿Office 2003来说,Excel单个工作表最多能放置65 536行数据。为什么说这是一个问题?设想你在一个小公司工作,并且在用Excel分析每月的交易,随着时间的推移,你构建了一个健壮的处理方法,该方法包括分析存储在该工作表中数据的公式、数据透视表和宏。当数据达到65 536行的限制时,会发生什么情况?新开一个工作表吗?怎样从整体上分析位于两个不同工作表中的两个数据集?原先公式还能用吗?是否得编写一个新的宏?这些都是需要处理的问题。

当然,精通Excel的用户可能会替你找到各种聪明的方法来应付Excel的这种局限性。虽然他们总是用奇异的方法来解决,但是即使再聪明的用户最终也会较少考虑用更有效的方法来完成数据分析和展示分析结果,他们会把较多的精力放在如何在不修改公式和函数的情况下适应Excel。我承认Excel很灵活,精通Excel的用户会把大多数任务放在Excel中完成。但是,当用户只考虑使用Excel时,虽然他们能用难以置信的方法完成任务,但是毫无疑问,他们把自己给束缚住了。

此外,该限制常迫使Excel用户要事先准备好数据。也就是,另外某个人从大型数据库中提取大块数据,然后为在Excel中使用而汇总和改造数据。严肃的分析员难道总是依赖他人提供数据吗?如果给分析员提供访问大量数据的工具,而不是依赖他人提供所需数据,情况会怎样?该分析员对机构不是更有价值吗?无需维护Excel数据,他可以把精力集中于分析精度和展示质

量上。

许多人认为，面对日益增长的数据量，对数据分析人员来说，Access是优秀的下一代分析工具。因为Access表没有预先限制行数，无需为适合Excel而对数据进行汇总和准备，所以分析人员就能够处理更大的数据集。因为许多任务在Excel和Access中是重复的，对两者都精通的分析员能够适应任何分析要求。



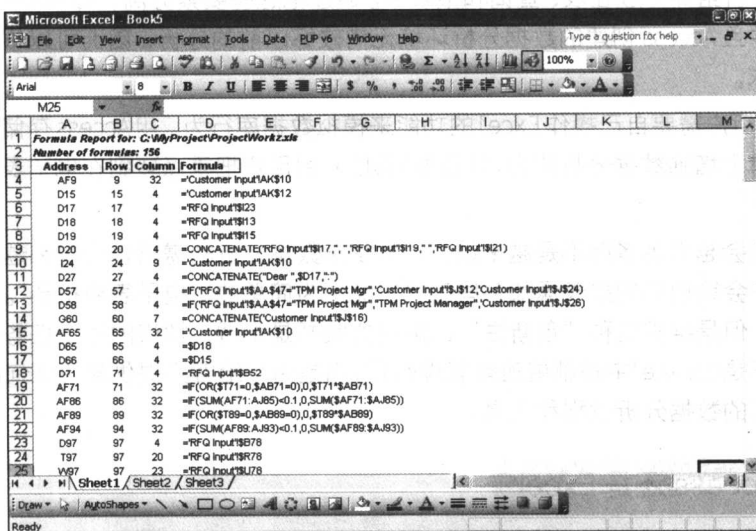
Access表被限制为不超过256列，但是没有行数限制。这并不是说，Access有无限的数据存储能力。每一位(bit)数据都会增长Access数据库文件的大小。Access数据库文件大小不能超过2GB。

分析过程的透明度

Excel最吸引人的特点之一是其灵活性。每个单元格可包含文本、数值、公式或用户定义的任何其他数据。事实上，这是Excel成为有效数据分析工具的基本原因之一。用户可以使用命名的范围、公式和宏来创建一个实现一连串计算、单元格链接、格式化汇总等功能的复杂的系统，这些功能的综合使用构成了一个最终的数据分析。

这存在什么问题吗？问题在于这样做缺少分析过程的透明度。也就是说，电子表格中实际发生的情况是很难确定的。使用他人创建的电子表格难以搞清各种计算以及用于执行分析的关系。要搞清执行少量分析的小电子表格已经让人头疼，而对于大型、复杂的包含多个电子表格的工作簿来说，我们实际是不可能弄清它是如何完成数据分析的，我们经常要从头构造数据分析。

甚至大多数Excel插件包的审计工具也没有提供多少帮助。下图显示的是在实际工作簿中使用公式审计工具的结果。它列出了该工作簿中的所有公式。使用该工具的人是想用该工具找到现有的公式并搞清公式的含义。图中第二行显示该工作簿有156个公式。该列表有点帮助，但作用不大。



公式审计工具对破解电子表格没有太多帮助

与 Excel 相比, Access 的规则是非常严格的。我们不能把公式直接放在数据字段中,不能建立数据字段与另一个表的链接。对许多用户来说, Excel 是热情的,它能够让用户做任何事情,而 Access 是难以相处的,除了错误信息,它什么也没有。但是, Access 的严格性也带来了如下的好处。

因为 Access 只允许特定的操作,所以我们能够更容易地理解 Access 中的一组数据正在完成什么任务。如果在分析过程中编辑某个数据集、某个数值或数据集的任何部分发生变化,我们可以很容易地看到查询或事件中调用该操作的结果。这并不是说用户在 Access 中不会犯错误,但是用户肯定不会在分析过程中遇到隐藏的步骤,如隐藏的公式、隐藏的单元格等。

数据和展示的分离

数据应该与展示分离,这是普遍的看法。人们并不希望数据与任何一种数据展示方法紧紧地捆绑在一起。例如,在收到公司的一张发票时,我们并不认为发票上的财务数据是数据的真正来源,它只是数据的展示。这些数据也可以以图表或网页等其他方式和风格展示,但是这些展示方式中的数据不是数据本身的复制。

Excel 对该概念是如何处理的呢?用 Excel 执行数据分析的用户经常把数据、分析和展示混在一起。例如,我们经常看到一个工作簿中有 12 个电子表格(每个电子表格代表一个月)。每个电子表格列出了该月的数据以及公式、数据透视表和汇总。在要求按季度汇总时会发生什么情况?难道要添加更多的公式和电子表格吗?根本问题在于电子表格实际上提供的是与分析展示混合在一起的数据值。数据不应该与特定的展示捆绑,无论这种捆绑多么富有逻辑和多么有用。但是,在 Excel 中,这种情况一直在发生。此外,正如前面所讨论的,因为所有的分析方式和分析阶段是直接电子表格中完成的,所以 Excel 不能有效地提供合适的分析透明度。Excel 单元格可能会存储公式、隐藏单元格或包含与其他单元格的链接。换句话说,数据和展示混在一起使得确定电子表格中真正发生的情况变得困难;而且,要保证编辑以及无法预料的修改不影响以前的分析,还得以手工维护的方式解决,这要花很多的精力。

Access 把分析组件分放在表、查询和报表中。通过把这些要素分离, Access 使数据对数据修改不再敏感,并且 Access 创建的数据分析环境无需破坏以前的分析,因而可以很容易地响应新的数据分析请求。

许多 Excel 用户需要自己操作 Excel 的功能来模拟数据库行为。用 Access 存储数据和处理分析需求将从整体上增强数据分析能力,并且能够让 Excel 用户把重点转到在电子表格中展示分析结果上。

未来的数据会越来越多而不是越来越少,对复杂数据分析的需求也会越来越多而不是越来越少。高级用户会给他们的报表添加某些工具以免成为简单的“电子表格机器”。Excel 几乎可以做任何事情,但是维护这种“创新性”的解决方案却是单调、烦琐的手工任务。数据分析中的迷人的部分不是在 Excel 中按部就班地管理数据,而是为了给客户id提供解决任何问题的最好方案,去创建完美的数据分析过程和工具。

确定使用 Access 还是 Excel

数据分析成功的关键并不是放弃 Excel 而仅使用 Access,而在于精通这两种应用程序、对项目的评估能力以及确定满足分析需求的最佳平台。确定使用 Excel 还是 Access 有什么简便的规

则吗？答案是否定的，但是每个项目中的某些关键特征可指导使用何种应用程序。这些特征包括数据大小、数据结构、数据演变、分析功能的复杂性以及共享处理的可能性。

数据大小

数据集的大小是要考虑的最明显的特征。尽管 Excel 能够处理 65 000 多行数据，但是如果数据集达到 50 000 行，考虑使用 Access 通常是不错的。这是因为 Access 和 Excel 处理数据的基本方法不同。在打开某个 Excel 文件时，为保证快速处理和访问数据，整个文件被载入到内存中。缺点是：即使在电子表格中做最小的改动，Excel 处理数据也需要大量的内存。可以注意到，在试图对有許多公式的大数据集应用自动筛选时，Excel 反应很慢，只是在状态栏指示“正在计算”。数据集越大，Excel 处理数据的效率越低。而 Access 的行为不同于 Excel。在打开某个 Access 表时，Access 似乎是打开了整个表，但是实际上 Access 只是一次打开了一小部分数据。这使打开数据集所需内存最少，从而保证最有效地使用内存和处理大数据集。此外，Access 能够用索引对相当大的数据集进行快速查找、排序、筛选和查询。

数据结构

如果所分析的数据仅驻留在一个与其他表没有关系的表中，Excel 是满足分析需求的较好的选择。但是，如果数据涉及一系列相关的表（如客户、订单、发票），那么可考虑使用 Access。Access 是关系数据库，它是为处理相互交互的数据集而设计的，能够保证数据完整性、防止数据冗余、在数据集之间进行有效的比较和查询。第 1 章要会详细地学习有关表关系的概念。

数据演变

Excel 是快速分析始终恒定数据的理想选择，即为获得更有价值的数据子集而对临时数据集进行整理。数据透视表的结果是这种一次性数据整理的完美示例。但是，如果用可能会演变和增长的数据构造长期的分析过程，Access 是更好的选择。在 Excel 中，许多分析过程起初规模较小、运行得很好，但是随着时间的推移，这些过程的规模和复杂性都在上升，直至达到 Excel 的极限。因此，在确定哪一种平台是更好的选择时，要有远见和考虑未来的需求。

功能复杂性

在实际生活中，有太多的分析项目，即使在达到 Excel 功能的极限时，仍然强行把这些项目放在 Excel 中处理。一个工作簿所包含的分析过程封装了多个电子表格、多个数据透视表以及进行求和、求均值、计数、查找以及与其他工作簿链接的许多公式，这种情况我们看到了多少次？事实上，当基于 Excel 的分析过程过度复杂时就会难以管理、维护和转换。对拥有复杂的、多步骤分析过程的项目考虑使用 Access。

共享处理

尽管让多个用户工作于网络中某个中央电子表格是可能的，但是试图让他们协调工作和管理中央电子表格是很困难和有局限性的。数据冲突、数据丢失、由于锁定而排除用户以及脆弱的数据完整性只是这些问题中的一些少量的示例。可以考虑对共享处理使用 Access。由于诸多的原因，Access 更适合共享环境。其中一些原因是，Access 能够让多用户并行输入和更新数据、内置对数据冲突的保护、防止数据冗余和防止数据输入错误。

Access 简介

许多有经验的管理员、会计师和分析员开始认识到，仅仅因为 Excel 能完成某些任务并不意味着 Excel 是完成这些任务的最好的方法。他们在第一次打开 Access 时，所熟悉的第一个对象是 Access 表。事实上，Access 表与 Excel 电子表格如此相似，大多数 Excel 用户试图像使用电子表格那样使用它。但是，当他们认识到不能在表中直接输入公式和重复 Excel 的大多数行为和功能时，他们想知道从哪里着手学习使用 Access。

很多老练的 Excel 用户在发现 Access 的行为和外观不同于 Excel 时，他们中的许多人认为 Access 学起来太难或学习 Access 要花很长时间。但是，Access 中的数据存储和管理的许多概念都是用户已经熟悉了。所有 Excel 用户都要学习这些概念，以便执行分析和展示分析结果。为在以后的自动化数据处理方面节省大量的时间，让我们先花很少一点时间了解一下 Access 的构成。

本书将要学习用 Access 完成数据分析的各种技术，这些技术不同于在 Excel 中所用的技术。这里从 Excel 专家的角度简要介绍了 Access，重点放在 Access 的大方面。如果提到的某些 Access 术语是新的或不是很熟悉，读者要有耐心，这些术语还会在本书相关章节中作深入的介绍。

表

读者最熟悉的无疑是 Access 表。表的外观几乎与电子表格相同，它们有相似的单元格、行和列。但是，第一次试图在表的某个“单元格”中输入公式时，则会发现 Access 表并不拥有 Excel 的灵活性、多用途特点（Excel 单元格几乎可以完成任何功能），而 Access 表只是存储数据的地方。

数据分析和数据整理都在其他地方完成，而表仅仅存储数值或文本。因此，数据永远不会与特定的分析或展示绑定在一起。表中的数据保持其原始格式，而数据的分析和展示则留待用户确定。如果 Excel 用户仅仅使用 Access 的表，他仍然能够极大地提高效率和生产率。

查询

你可能听说过但从未使用过 Access 查询。考虑如下情况：在 Excel 中，当使用自动筛选或 Vlookup 公式或子汇总时，实际上是为得到答案而基于一组数据运行查询。查询是为得到答案而提出的一个问题。查询答案可以是一个数据项、是/否逻辑值或多行数据。在 Excel 中，查询数据的概念是模糊的，因为 Excel 中的查询可以是各种不同的功能形式，如公式、自动筛选和数据透视表。

在 Access 中，查询是一个有自己功能的实际对象。为保证数据不与特定的分析捆绑在一起，查询与表是分离的。用 Microsoft Access 增强数据分析成功的关键在很大程度上取决于创建各种简单和复杂查询的能力。

报表

报表是 Microsoft Access 中功能非常强大的组件，它能够以各种风格展示数据。Access 报表很好地印证了本书的一个重要的观点：数据应该与分析 and 展示相分离。报表是分析和展示一组特定数据的外壳，而不是数据本身。报表不关心数据信息来自何处。只要报表接收到所需的数据就可以准确和清晰地展示这些信息。

Access报表优缺点并存。一方面，它可以提供清晰的PDF样式的报表，对于展示发票和套信用信函非常理想。另一方面，Access报表对一次性的数据显示不如Excel。不过，Access可很容易地针对各种报表样式进行配置。

宏和VBA

就像Excel有宏和VBA功能一样，Microsoft Access也有相同的功能。这是Microsoft Access数据分析的强大功能和灵活性之所在。无论是在自定义函数、批量分析，还是在自动化方面，宏和VBA都可以增强用其他方法难以实现的灵活性。

本书学习内容

选择本书的读者可能工作于数据密集型的环境中。数据分析是工作任务中的一个完整的部分。读者可能会觉得始终没有掌握Access中的某些强大的功能。如果是这样的话，那么要始终相信自己的能力。事实上，用Access完成分析需求有助于简化分析过程、提高生产率和分析超越Excel限制的更大的数据集。

通过本书的学习，读者会认识到，Access不仅仅是用于存储数据和构造部门级应用的数据库程序，它还拥有强大的数据分析功能；这些功能易于学习并且可以容易地应用于多种机构和数据系统。

前3章学习Access的基本知识，对超出Excel限制的大数据集执行功能强大的数据分析。前9章用Access的高级特性给数据分析增加深度和广度。至本书结束时，读者将学会创建自己的定制函数、执行批量分析和开发自动化过程。

本书涉及和未涉及的内容

- ◆ **本书介绍Access的全部内容吗？** 市场上的许多书详细地介绍了Access的方方面面，但是本书不属于这类书。尽管本书确实介绍了Access的基础知识，但是它始终以数据分析为中心，并且本书是从数据分析员的角度编写的。如果读者是第一次使用Access，本书会为读者掌握好Access基本知识打下坚实基础，为日常使用Access留下有价值的技巧。
- ◆ **本书介绍数据管理理论和最佳实践吗？** 本书既不是介绍数据管理理论和最佳实践的书，也不介绍高级商业智能概念。本书更侧重技术方面；Access可以为常见的数据分析问题提供功能强大的解决方案，本书介绍把Access用做分析工具的具体使用步骤。读完本书，读者将能够用合适的方法分析大量的数据，快速把数据分放在各种视图中，实现冗余分析自动化，节省时间和提高生产率。
- ◆ **本书学习统计分析吗？** 尽管本书没有用专门的章节完整介绍统计分析的各种技术，但是本书确实涉及统计理论、方法和最佳实践。

本书的组织方式

第1部分 Access 数据分析基本知识

第1部分简要介绍Access。这部分首先学习Access的基本知识，然后学习本书后续章节所

需的重要的查询技巧。这部分涉及的话题包括关系数据库、查询基础知识以及汇总查询、操作查询和交叉表查询。

第2部分 基本分析技术

第2部分介绍 Access 的基本分析工具和技术。第3章介绍数据转换，用示例说明怎样为满足需求而对原始数据进行清理和改造。第4章深入地介绍如何在分析中创建和利用自定义的计算；还介绍如何使用日期，包括把日期用于简单的日期计算和执行高级时间分析。第5章讨论一些条件分析技术，这种技术能够给分析过程添加逻辑。

第3部分 高级分析技术

第3部分介绍许多高级技术，这些技术能够把分析技术提高一个层次。第6章介绍 SQL 语句的基本知识，还介绍子查询和域聚合函数。第7章介绍能够用子查询和域聚合函数完成的许多高级统计分析。第8章深入地介绍 Access 数据透视表和数据透视图的用法。

第4部分 数据分析自动化

第4部分介绍的内容超越了手工分析的范畴，把读者带进自动化世界。第9章深入地介绍宏是如何通过运行批量分析帮助我们提高生产率的。第10章介绍应用程序所使用的 Visual Basic (Visual Basic for Application, VBA)，并介绍 VBA 是如何帮助我们提高数据分析能力的。第11章介绍 Excel 自动化幕后的一些技术。第12章介绍构造和维护数据库的最佳实践，以及获取 Access 帮助的方法，该方法能够让读者沿正确方向深入了解某方面的内容。

关于本书示例

可在网址：www.wiley.com/go/accessdataanalysis 找到本书所用的示例。

目 录

第 1 部分 Access 数据分析基本知识	1
第 1 章 Access 基本知识	2
1.1 Access 表的基本知识	2
1.1.1 什么是表	2
1.1.2 在数据表试图中打开表	3
1.1.3 识别表的重要要素	3
1.1.4 在设计视图中打开表	3
1.1.5 数据类型	4
1.1.6 准备建表	5
1.1.7 创建表的不同方法	6
1.1.8 用设计视图创建表	6
1.2 表的高级概念	8
1.2.1 字段属性	8
1.2.2 主键	10
1.3 获取外部数据	12
1.3.1 导入	12
1.3.2 链接	12
1.3.3 有关导入数据的注意事项	12
1.4 关系数据库的概念	13
1.4.1 Excel 和普通文件格式	14
1.4.2 把数据分在不同的表中	15
1.4.3 外键	15
1.4.4 关系类型	16
1.5 查询的基本知识	17
1.5.1 什么是查询	18
1.5.2 创建选择查询	18
第 2 章 Access 中的查询	29
2.1 汇总查询	29
2.1.1 创建汇总查询	29
2.1.2 探讨合计函数	32
2.2 操作查询	36
2.2.1 为什么使用操作查询	37
2.2.2 生成表查询	37
2.2.3 删除查询	39
2.2.4 追加查询	43
2.2.5 更新查询	47
2.3 交叉表查询	50
2.3.1 使用交叉表查询向导	51

2.3.2	手工创建交叉表查询	55
2.3.3	自定义交叉表查询	57
第 2 部分	基本分析技术	61
第 3 章	Access 数据转换	62
3.1	查找和删除重复的记录	62
3.1.1	定义重复记录	62
3.1.2	查找重复记录	63
3.1.3	删除重复记录	65
3.2	常见的数据转换任务	67
3.2.1	填充空字段	67
3.2.2	连接	68
3.2.3	改变大小写	70
3.2.4	删除字符串中前导和尾部空格	72
3.2.5	查找和替换特定文本	73
3.2.6	在字符串中的关键位置添加指定的文本	74
3.2.7	用字符标记解析字符串	77
第 4 章	计算和日期	81
4.1	在分析中使用计算	81
4.1.1	常见的计算场合	81
4.1.2	用表达式生成器构造计算	84
4.1.3	常见的计算错误	87
4.2	在分析中使用日期	90
4.2.1	简单的日期计算	90
4.2.2	使用函数的高级分析	91
第 5 章	执行条件分析	99
5.1	使用参数查询	99
5.1.1	参数查询的工作方法	100
5.1.2	参数查询的规则	100
5.1.3	使用参数查询	101
5.2	使用条件函数	105
5.2.1	Iif 函数	105
5.2.2	Switch 函数	110
第 3 部分	高级分析技术	113
第 6 章	子查询和域聚合函数	114
6.1	理解 SQL	114
6.1.1	SELECT 语句	115
6.1.2	ORDER BY 子句	116
6.1.3	WHERE 子句	116
6.1.4	GROUP BY 子句	117
6.1.5	HAVING 子句	117
6.1.6	AS 子句	117

6.1.7	理解连接	118
6.1.8	SQL 中的联合查询	120
6.1.9	SELECT TOP 和 SELECT TOP PERCENT 语句	120
6.1.10	SQL 中的操作查询	122
6.2	用子查询增强分析能力	123
6.2.1	为什么要使用子查询	124
6.2.2	子查询的基本规则	124
6.2.3	在子查询中使用 IN 和 NOT	126
6.2.4	子查询与比较运算符一起使用	127
6.2.5	把子查询用做表达式	127
6.2.6	使用相互关联的子查询	128
6.2.7	在操作查询中使用子查询	129
6.3	域聚合函数	132
6.3.1	域聚合函数剖析	133
6.3.2	理解不同的域聚合函数	133
6.3.3	检查域聚合函数的语法	134
6.3.4	使用域聚合函数	136
第 7 章	Access 中的描述统计	141
7.1	描述统计的基本知识	141
7.1.1	描述统计和汇总查询一起使用	141
7.1.2	确定排位、众数和中值	142
7.1.3	从数据集中提取随机样本	146
7.2	高级描述统计	148
7.2.1	计算百分等级	148
7.2.2	确定记录的四分位数	149
7.2.3	创建频数分布	151
第 8 章	用数据透视表和数据透视图分析数据	156
8.1	在 Access 中使用数据透视表	156
8.1.1	数据透视表剖析	156
8.1.2	创建基本的数据透视表	159
8.1.3	创建带有明细数据的高级数据透视表	162
8.1.4	把数据透视表视图设置为默认视图	164
8.1.5	数据透视表选项	165
8.2	在 Access 中使用数据透视图	173
8.2.1	数据透视图的结构	173
8.2.2	创建基本的数据透视图	175
8.2.3	设置数据透视图的格式	175
第 4 部分	数据分析自动化	179
第 9 章	调度和运行批量分析	180
9.1	Access 宏简介	180
9.1.1	创建第一个宏	180
9.1.2	重要的宏操作	183

9.2	建立和管理批量分析	185
9.2.1	数据库的组织方式	185
9.2.2	建立基本的批量分析	189
9.2.3	构造条件宏	191
9.2.4	宏循环	195
9.3	调度宏使其每夜运行	199
9.3.1	使用 AutoExec 宏调度任务	199
9.3.2	用命令行调度任务	204
第 10 章	用 VBA 增强数据分析能力	206
10.1	创建和使用自定义函数	206
10.1.1	创建第一个自定义函数	207
10.1.2	创建接受参数的自定义函数	209
10.2	用窗体控制分析过程	212
10.2.1	数据从窗体传递给查询的基本知识	213
10.2.2	用窗体增强自动化功能	216
10.2.3	枚举组合框列表项	218
10.3	在幕后处理数据	221
10.3.1	用 RunSQL 语句处理数据	221
10.3.2	使用 RunSQL 语句的高级技术	223
第 11 章	用 Excel 自动化报告结果	227
11.1	Excel 自动化的基本知识	227
11.1.1	关于 Access 与 Excel 的绑定	227
11.1.2	创建第一个 Excel 自动化过程	228
11.1.3	把 Access 数据发送给 Excel	230
11.2	高级自动化技术	233
11.2.1	使用 Excel 代码	233
11.2.2	优化宏产生的代码	238
第 12 章	查询性能、数据库损坏及其他	241
12.1	优化查询性能	241
12.1.1	理解 Microsoft Jet 的查询优化程序	241
12.1.2	优化查询性能的步骤	241
12.2	处理数据库的损坏	243
12.2.1	数据库损坏的特征和征兆	244
12.3	获取 Access 帮助	247
12.3.1	请求帮助时的位置问题	247
12.3.2	使用目录而不是关键词搜索	248
12.3.3	联机帮助好于脱机帮助	248
12.3.4	联机资源的各种知识库	249
附录 A	数据分析员函数参考	250
附录 B	Access VBA 基础	282
附录 C	数据分析员错误参考	291