



万水 Windows 软件精通指南系列

使用 ^{M I C R O S O F T} EXCEL 进行数据分析

Data Analysis Using Microsoft Excel

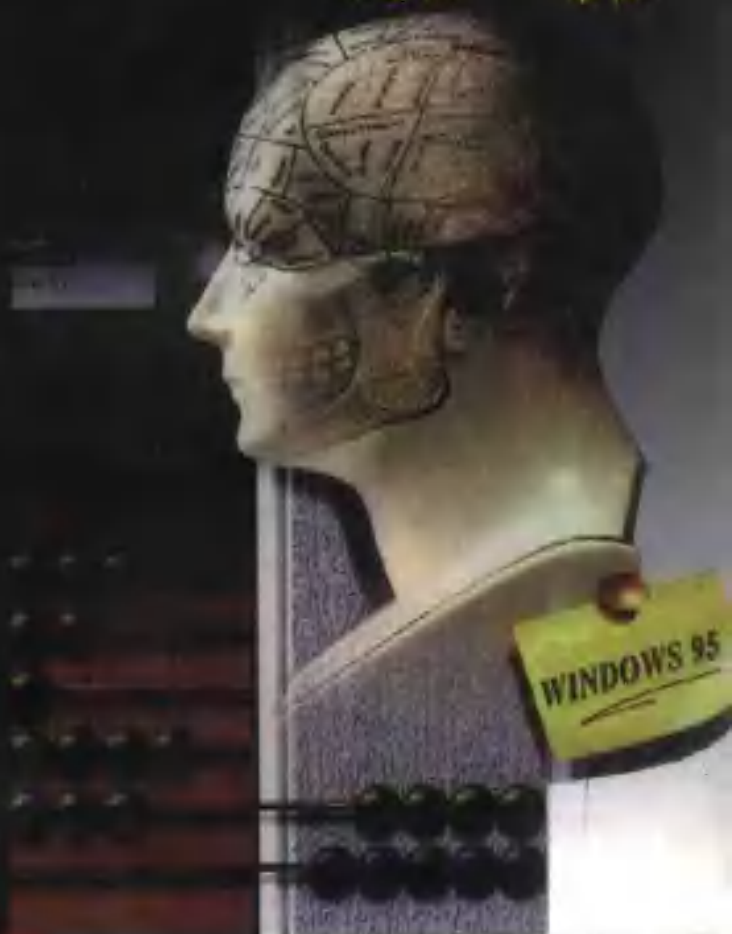
MICHAEL R. MIDDLETON

康博创作室 译
冯志强 审校

本书将教会您如何优化和充分利用当前流行的 Microsoft Excel 7.0 电子表格程序在 Windows 95 版本。

本书通过从如何使用 Excel 7.0 的简洁清晰的操作指令，到示例说明给予读者在数据分析方法方面坚实的实际操作指导。

本书集中讲解统计分析、回归模型和预测方法，以便您在工作中更好地进行商务决策。



中国水利水电出版社

万水 Windows 软件精通指南系列

使用 Microsoft Excel 进行数据分析

〔美〕 Michael R. Middleton 著

康博创作室 译

冯志强 审校

中国水利水电出版社

内 容 简 介

本书讲述了如何运用Excel电子表格软件进行数据分析,全书共分二十章。前三章介绍了Excel的文件、打印和图表的基本功能;第四章到第二十章分别介绍了各种强大的Excel数据分析功能。每章后面均附有适当的练习,以供读者进一步巩固所学内容。

本书适用于广大Excel的用户和自学者,尤其是从事数据分析和商业统计的人员必备的参考资料。

COPYRIGHT © 1997 by Wadsworth, A Division of International Thomson Publishing Inc.

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 any information storage and retrieval system, without permission, in writing, from the Publisher.

本书中文简体字版由中国水利水电出版社出版,未经出版者书面许可,不得以任何方式复制或抄袭本书的任何部分。

图书在版编目(CIP)数据

使用Microsoft Excel进行数据分析 / (美)米德尔顿(Middleton, M.R.)

著:康博创作室译. - 北京:中国水利水电出版社, 1997.11

(万水 Windows 软件精通指南系列)

书名原文: Data Analysis Using Microsoft Excel

ISBN 7-80124-597-0

I. 使… II. ①米…②康… III. 电子表格系统, Excel—应用—数据处理 IV. TP274

中国版本图书馆CIP数据核字(97)第26220号

书 名	使用 Microsoft Excel 进行数据分析
作 者	Michael R. Middleton
译 者	康博创作室
审 校	冯志强
出版、发行	中国水利水电出版社(北京市三里河路6号 100044) 北京万水电子信息有限公司(北京市车公庄西路20号 100044)
排 版	北京万水电子信息有限公司
印 刷	北京牛山世兴印刷厂
规 格	787 × 1092 毫米 16 开本 14.5 印张 316 千字
版 次	1997 年 12 月第 1 版 1997 年 12 月北京第一次印刷
印 数	0001—5000 册
定 价	24.00 元

译者序

Microsoft Excel 是微软公司推出的著名办公室自动化软件,它用于生成和打印各种电子表格,进行一些商务数据的处理。关于如何使用 Excel 的书籍已经出版了很多,但是专门深入讲解 Excel 的数据分析和处理的书籍极为少见,本书正好填补了出版界的这一空白。

本书深入分析和讲解了运用 Excel 进行数据分析和商业统计的过程,结合各种实际的商务工作精选了丰富的典型示例,这些示例的逐步操作说明和屏幕显示均适用于至 Excel 7 for Windows 95 的各种 Windows 版本,同时也适用于 Macintosh 版本。作者深入研究了用 Excel 进行数据处理与统计的特点,摸索了读者学习 Excel 数据处理的过程,根据初学者的特点提供了丰富的处理实例,精心设计了全书的体系结构。在翻译过程中,我们学到了许多有益的知识,并为本书的优秀内容和编排结构所吸引。我们郑重地向读者推荐此书,它定会成为您掌握用 Excel 进行实际工作中的数据处理与统计分析的良师益友。

本书由康博创作室集体翻译而成,参加人员包括:曹康、朱琳、李东升、李文珍、李伟、曹云、崔洪斌、李志诚、李娟、曹锋、李端银、曹永胜、王睿、袁勤勇、邓中亮、蔡汇锦、陈军和赵霞。最后,感谢冯志强先生,他在百忙中抽出宝贵时间认真审校了全书的译稿,并提出了一些很有价值的修改意见。

译者

1997 年 7 月于北京

作者序

本书讲述如何使用 Excel 电子表格软件进行数据分析,适用于想要逐步学习使用 Excel 统计分析工具的学生、管理人员及分析人员。

读者

本书主要是为了那些学习数据分析和商业统计课程的大学生而写的。尽管书中所有示例均是面向商业的,但是逐步的操作说明也适用于其它课程及学术科目中的统计分析。绝大多数学生都认识到了电子表格技能在他们将来开始工作中的重要性,把此书作为教材使用的老师目的在于帮助同学们能熟练使用电子表格。管理人员和技术分析人员也会发现,这些技术对他们进行数据分析以便在公司中作出决策十分有用。

本书试图作为统计教科书的补充,但不能代替教科书。当然,有统计基础知识的管理人员和分析人员可以不用参考统计学教科书而直接使用本书,因为在讲解逐步的操作说明之前,都是先讨论适当的统计理论,然后再对 Excel 的输出进行简要解释。

特点

学生、经理和分析人员都将受益于使用单一的前端界面去访问各种各样的分析工具,这些工具包括图形、数据库、模拟、优化和统计分析工具,这些都在书中给出。

我们主要的目标就是说明用 Excel 分析数据的各种不同方法,更强调用最简单的方法开始。通常,用 Excel 有好几种方法可得到指定的结果,书中首先给出了最简单的方法,通常使用 Analysis Tool 或 Excel Wizard。在许多分析任务中,包括了另一种替补方法的逐步操作说明,这些方法往往是为了提供更高级的结果或者用它来避免在 Tool 或 Wizard 功能中包含不需要的功能。例如,本书说明了三种进行简单回归的方法:(1)在一散布图上的 Insert Trendline 命令;(2)能提供更完善的诊断分析的回归 Analysis Tool;(3)工作表 TREND 函数。

本书还说明了如何利用 Excel 的图形能力来“观看数据”,这可以帮助用户在用数值计算之前选择正确的分析技术。

尽管我们故意使每一例的数据集很小,以便用户可快速将其输入一 Excel 工作表中,但是同样的逐步操作说明均可用于大的或小的数据集。首先输入小数据集,按逐步指令操作,确认了结果正确后,再按相同的步骤来处理用户自己的数据,可能数据集较大。

Excel 的联机帮助中可用的公式没有在本书重复出现。如果 Excel 的文档不完整或者含糊的话,笔者希望能把它们澄清。

本书所用示例中的数据可以电子形式在 Web 上的 www.thomson.com/duxbury.html 中获得。希望得到软盘的教师可通过您的本地 ITP 代理与出版联系。

Excel 的版本

本书中所有的逐步操作说明和屏幕显示均是基于 Windows 95 下的 Excel 7 的。Excel 7 和 Excel 5.0c 的统计分析工具的功能是相同的,在使用 Windows 95 时,Excel 7 和 5.0 的对话框具有相同的外观。Excel 5.0 中的几个错误在 5.0c 版本中已修复,有 5.0 版本注册权的用户都可获得一组完整的 5.0c 的磁盘。

本书所讲的许多技术也适用于 Excel 4,先于版本 4 的 Excel 必须升级,老版本(3,2.1 等)没有完整的统计函数集。

Macintosh 版和 Windows 版的 Excel 5 在文件、对打印对话框和一些命令击键方面有区别。附录中给出了关于使用 Excel 5,Windows 3.1 和 Macintosh 的信息。

Windows 95 版的新特点

由于 Windows 3.1 和 Windows 95 的屏幕显示的主要差别是对话框的外观,所以必须对屏幕显示的逐步操作说明进行修改,以便强调使用鼠标右键和快捷菜单。第十七章包括了一些新内容,即用回归进行二元依赖变量的判别式分析。

课程应用

许多教师继续使用他们惯用的统计教科书中的内容,并安排了本书的相关部分让学生进行计算机分析。其中一些需要学生重复使用逐步指令并递交与书中特定图形相类似的输出结果。另外一部分是在本书或教师的统计教科书每章末尾安排了练习,在使用同样的操作过程的练习前,鼓励或要求学生先重复书中的逐步操作说明。在一些课程中没有安排练习,本书可当作帮助学生完成课程计划的一种资源。

致谢

第一版的成功是一件令人兴奋的事情。在这里首先感谢 Wadsworth Publishing 的 Curt Hinrichs 对本书的鼓励和帮助。感谢在修改本版中提供了有价值的反馈意见的审阅人员,他们包括:Rollins College 的 Harry Kypraios; Fullerton 的 California State University 的 John Lowrence 和 George A. Marcoulides; University of Arizona 的 Lisa Ordonez; Appalachian State University 的 George A. Schieren; North Carolina Agricultural and Technical State University 的 Charles VanDezande; University of San Francisco 的 Steven Widener; 另外,感谢将第一版用于工业课堂教学并提供有益意见的 Bill Anderson 和 Baxter Labs。

同时感谢对第一版的编写做出贡献的审阅者: Indiana University 的 Christian Albright; University of Alabama 的 Michael D. Coneryly; University of Southern California 的 Delores Conway; Los Angeles 的 University of California 的 George Geis; Wilfrid Laurier University 的 Gerald Keller; Boston University 的 Ron Klimberg; Georgia Institute of Technology 的 Tyra Anne Mitchell; University of Tennessee 的 Sharon Neidert; University of Colorado 的 Richard Spinetto; Bucknell University 的 Nancy C. Wedia; Indiana University 的 Wayne L. Winston。

Nancy Padgett 对第一版提供了重要的风格反馈,有几个 University of San Francisco 的 MBA 学生帮助我们确保了各条逐步操作说明的准确性,他们是:Lindsay Sutton, Svarai Luenchavee, Thierry Crifasi 和 Renu Gupta。

最后,欢迎读者多提改进意见。

Michael Middleton
McLaren School of Business
University of San Francisco
2130 Fulton Street
San Francisco, CA 94117
middleton@usfca.edu

目 录

第一章 Excel 介绍	1
1.1 鼠标和菜单的约定	1
1.2 Windows 95 的 Start 按钮	1
1.3 Windows 95 帮助	2
1.4 拷贝文件	4
1.5 启动 Excel	6
1.6 Excel 帮助	7
练习	8
第二章 文件管理和打印	9
2.1 工作簿、工作表和图表	9
2.2 打开文本文件	9
2.3 备份和版本控制	12
2.4 页面设置、预览和打印	14
练习	17
第三章 基本图表	18
3.1 条形图形和柱状图表	18
3.2 饼图	20
3.3 时序数据的折线图表	21
3.4 嵌入式图表和图表工作表	23
3.5 图表放缩 VBA 宏(选读)	24
练习	27
第四章 单变量数值数据	28
4.1 分析工具:DESCRIPTIVE STATISTICS	28
4.2 分析工具之二:直方图	33
4.3 分析工具三:Ranks and Percentile	37
练习	38
第五章 分类数据	39
5.1 单变量表格	39
5.2 双向表	43
练习	47
第六章 双变量的数值数据	50
6.1 XY(散点)图表	51

6.2	分析工具:相关性	52
6.3	分析工具:协方差	53
6.4	多变量的相关性	54
	练习	56
第七章	概率分布	57
7.1	二项分布	57
7.2	其他不连续分布	61
7.3	正态分布	62
7.4	t 分布及其其他连续分布	66
	练习	67
第八章	抽样与模拟	68
8.1	分析工具:抽样	68
8.2	分析工具:随机数产生	70
8.3	用 RAND 函数生成随机数	72
8.4	模拟中心极限作用	72
8.5	模拟随机走动	74
	练习	76
第九章	平均值的单样本推断	77
9.1	正态分布与 t 分布	77
9.2	使用正态分布的置信区间	77
9.3	使用 t 分布的置信区间	79
9.4	假设检验	81
9.5	使用正态分布的假设检验	82
9.6	使用 t 分布的假设检验	83
	练习	87
第十章	质量控制图表	88
10.1	平均值图表	88
10.2	标准偏差图表	93
	练习	95
第十一章	平均值的两样本推断	96
11.1	使用 t 分布的等方差	96
11.2	使用 t 分布的不相等方差检验	101
11.3	使用 t 分布的双样本	103
	练习	106
第十二章	χ^2 检验	108
12.1	正态性的 χ^2 检验	108
12.2	列联表格 P 值	110

12.3	χ^2 统计	113
	练习	114
第十三章	方差分析	116
13.1	单因素方差分析	116
13.2	无重复双因素方差分析	120
13.3	有重复双因素方差分析	122
13.4	方差的二样本 F 检验	126
	练习	128
第十四章	简单线性回归	130
14.1	插入线性趋势线	130
14.2	回归分析工具	133
14.3	回归函数	138
	练习	141
第十五章	简单非线性回归	142
15.1	多项式	143
15.2	对数法	145
15.3	幂函数法	147
15.4	指数法	148
	练习	152
第十六章	多元回归	153
16.1	回归输出的解释	154
16.2	残数的分析	155
16.3	用趋势函数进行预测	156
	练习	158
第十七章	使用分类变量的回归	160
17.1	类别作为解释变量	160
17.2	使用指示器变量回归的解释	162
17.3	多元回归的解释	163
17.4	分类作为依赖变量	164
	练习	166
第十八章	自相关和自回归	169
18.1	线性时序趋势	169
18.2	Durbin-Watson 统计量	171
18.3	自相关性	171
18.4	自回归	173
18.5	自相关性数函数	176
18.6	AR(2)模型	177

练习.....	179
第十九章 时间系列平滑	180
19.1 使用插入趋势线的流动平均值法.....	181
19.2 用数据分析工具求流动平均值.....	183
19.3 指数平滑法工具.....	184
19.4 指数平滑的初始值.....	186
19.5 最佳平滑常数.....	187
练习.....	190
第二十章 时间系列的季节性	192
20.1 使用指示器变量的回归.....	192
20.2 AR(4)模型.....	197
20.3 经典时间系列分解.....	202
练习.....	207
附录 Excel 5, Windows 3.1 和 Macintosh	209
A.1 学习 Windows 3.1	209
A.2 学习 MACINTOSH	210
A.3 学习 Excel 5	210
A.4 使用 Excel 5 的联机帮助	215
练习.....	217

第一章 Excel 介绍

本章内容可使读者熟悉在后序章节中进行数据分析时经常要用到的 Windows 95 和 Excel 的术语和功能。若读者对 Windows 95 和 Excel 不熟悉,应仔细阅读这一章。即使读者已经用过 Excel,也应浏览一下本章,以便复习本章所述的功能。

1.1 鼠标和菜单的约定

下面是用于鼠标操作的约定:

单击:单击鼠标左键并迅速释放。

右击:单击鼠标右键并迅速释放。鼠标右键在 Windows 95 和 Excel 中用于打开快捷菜单。快捷菜单显示当前选定的屏幕对象的选项。本书的指导通常说明为:①单击左键选定目标;②在同一选项上单击右键打开快捷菜单;在许多情况下只需要单击右键即可。

双击:连续快速双击鼠标左键,用 Windows 95 Control Panel 可调整双击速度。

拖动:按住鼠标左键移动鼠标。这一操作可用于选择一个 Excel 工作表中的一组单元或者改变图表或其他图形目标的大小。

Excel 命令包括菜单、子菜单、对话框标签和其他项目中的选项,连续的选项由竖道分开表示。例如,要改变 Excel 的网格线显示,可选择 Tools|Options|View|Gridlines。Tools 在菜单栏上,Options 是 Tools 的菜单选项,View 是 Options 对话框的一个标签,Gridlines 是 View 标签上的一个复选框的选项。

1.2 Windows 95 的 Start 按钮

Windows 95 的初始外观取决于用户对它的设置。通常 Start 按钮将作为任务栏的一部分出现在屏幕左下角,而其他按钮和指示符也可能出现在任务栏上。如果在屏幕上看不到 Start 按钮,那就可能是因为前面的用户选择了任务栏中的“Auto hide”选项,可将鼠标指针慢慢移近显示区的底边,直到 Start 按钮出现。

当单击屏幕底端的 Start 按钮时,Start 菜单就会出现并向上拉出图 1.1 所示内容。这一菜单中可能会包括一些其他选项。

在 Start 菜单中,单击 Programme 可显示已装在用户计算机中的应用程序列表。有时应用程序列表中的项目将会扩展显示出另一个程序列表。可依次打开各层列表直至找到想要运行的程序。

单击 Documents 可显示先前打开的最后 15 个文档,其中包括 Excel 工作簿和 Word 文

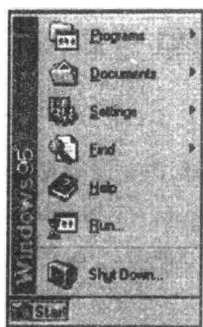


图 1.1 Windows 95 的 Start 菜单

档文件。

单击 Settings 可访问 Control Panel 来改变鼠标选项(例如双击速度)和许多其他 Windows 95 设置。双击位于 Control Panel 中的图标能看到可以改变的设置。Start 菜单中的 Setting 下还有其他选项如 Printer 和 Taskbar 等。

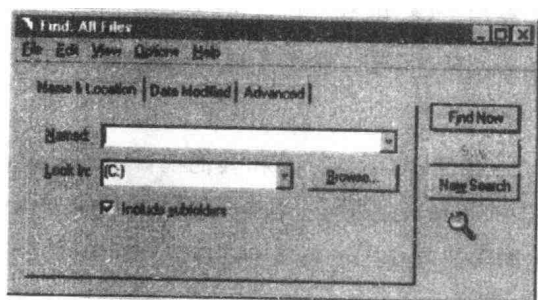


图 1.2 Windows 95 的 Find 对话框

要在计算机中定位一个文件夹、文档或其他文件,可先单击 Start 按钮,再指向 Start 菜单的 Find 项,然后单击 File 或 Folders。接着将出现如图 1.2 所示的 Find 对话框。单击 Named 框,然后键入文件名或文件夹的名字,并单击 Look In 框的箭头或单击 Browse 来确定搜索的位置。如果愿意,可选择 Include Subfolders 框。再单击 Find Now 按钮就可开始寻找。要取消寻找,只须单击位于对话框右上角的小“X”按钮即可。

Start 菜单中另外三个命令是:①Help,用于访问将在下节讲述的联机帮助系统;②Run,通过给出程序路径和程序名来启动程序(一般不如用 Program 命令指向程序方便);③Shut Down,用于完全退出 Windows 95。

1.3 Windows 95 帮助

联机帮助系统是 Windows 95 的一个重要的信息源。要访问帮助,可单击 Start 菜单,再

单击 Start 菜单中的 Help。单击 Contents 标签将出现如图 1.3 所示的主题列表。这一列表可能与用户计算机中的不同,可能包括一个指导性的 Tour:十分钟学会 Windows。

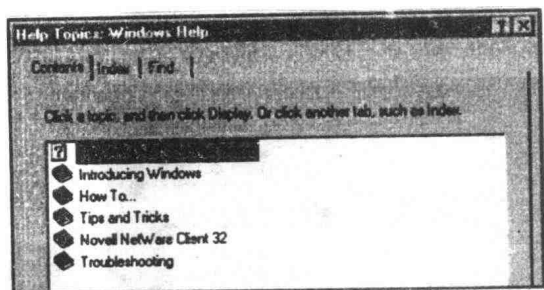


图 1.3 Windows 95 Help 的 Contents 标签

双击一个闭合的书本图标可显示一个主题列表,双击一个打开的书本图标可收起列表。例如,如果依次双击 How to 书本图标,Work with Files and Folders 书本图标和 Copying File or Folder 项,指导将出现在屏幕上的一个独立的窗口中,如图 1.4 所示。

要想在其他窗口中工作时还能看到指导,可单击 Option 按钮并选择 Keep Help on Top。若要退出联机帮助,单击指导的对话框右上角的 X 按钮即可。要返回 Contents 列表,只须单击 Help Topics 按钮。

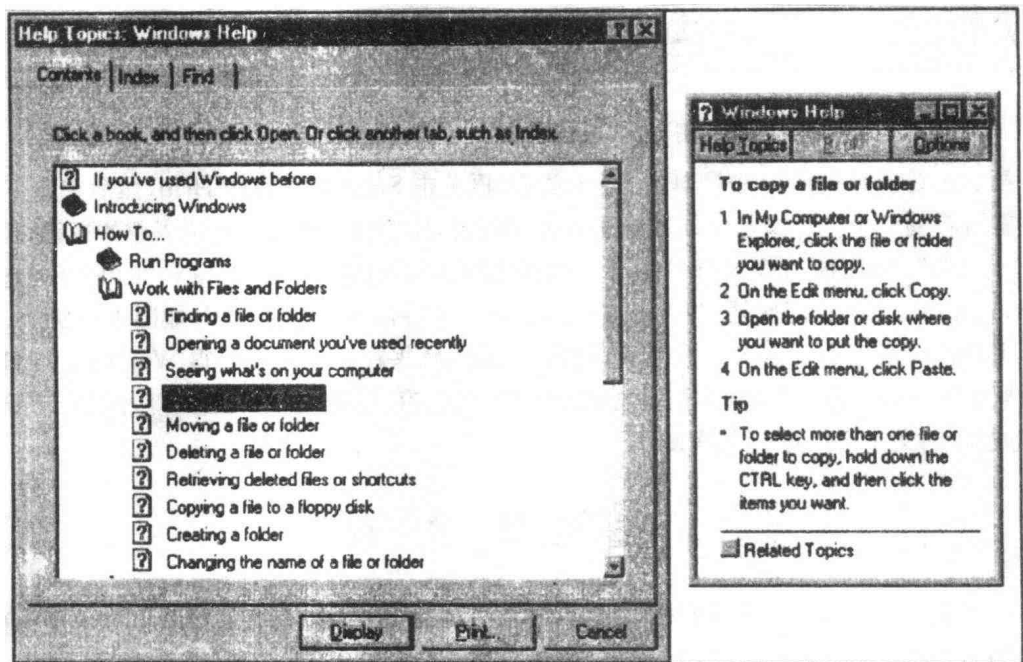


图 1.4 帮助的 Contents 标签和指导

第二种访问帮助的方法是单击 Start 菜单,再单击 Start 菜单中的 Help,然后单击 Index 标签来观看按字母顺序排列的列表的主题。如果用户键入一个感兴趣的单词,如“copying”,索引将显示相关的主题供用户选择。键入“copying”后,选择子标题“files to a floppy disk”,再双击或单击 Display 按钮,指导将显示出来,如图 1.5 所示。

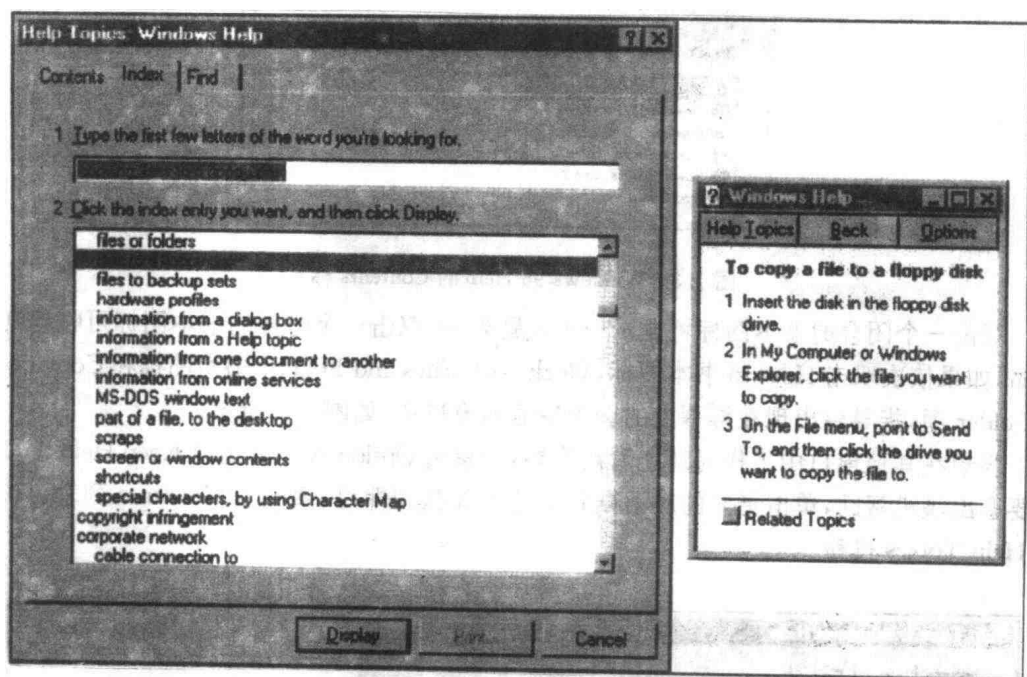


图 1.5 帮助的 Index 标签和指导

第三种访问帮助的方法是单击 Start 菜单,再单击 Start 菜单中的 Help,然后单击 Find 标签就可看到一个包含指定单词或短语的主题列表。如果用户键入一个感兴趣的短语,例如“move a file”,对话框的第 3 部分显示相关主题供用户选择。键入“move a file”后,选择“moving a file or folder”主题,并双击或单击 Display 按钮,指导将会出现,如图 1.6 所示。

用户也可通过单击右上角的问号,然后,再单击感兴趣的选项来获得 Windows 95 Help 对话框中项目的信息。对于大多数的对话框选项,均会有它的解释出现在一个小的文本窗口中,单击小的文本窗口可使之消失。

1.4 拷贝文件

有好几种方法可用来寻找想拷贝的文件,而完成拷贝和粘贴操作也有好几种方法。查找文件的方法之一就是单击 Start 按钮,再指向 Start 菜单的 Find 项,然后单击 Files and Folders,将出现如图 1.2 所示的 Find 对话框。键入文件名,选择适当的选项,然后单击 Find Now,对话框就会扩出显示查找结果。要列出文件的详细信息,可从 View 菜单中选择

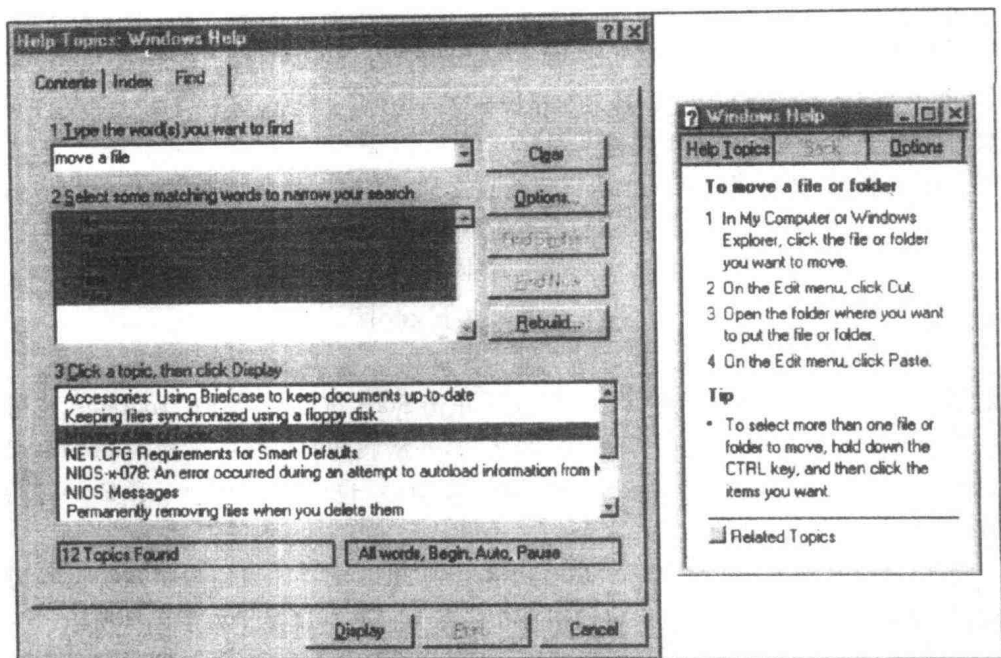


图 1.6 帮助 Find 标签和指导

Details 项,右击文件名,并从快捷菜单中选择 Copy。若不知道要拷贝的文件在何处,这是最好的方法。

第二种查找文件的方法是单击 Start 按钮,再指向 Start 菜单中的 Programs,然后单击 Programs 列表中的 Windows Explorer,用 Explorer 左上角的小下拉式列表框选择相应的驱动器。Explorer 将显示一个大窗口,窗口左边为文件夹的层次结构,右边为选中的文件夹所包含的内容。用加、减按钮扩展或收缩文件夹的层次结构,当发现所要找的文件时,用右键单击文件名,选择快捷菜单中的 Copy 项。这种方法适用于知道所需文件所在的文件夹的层次结构时。

第三种寻找文件的方法是移动桌面上的窗口或者将应用程序最小化直至可看到“My Computer”图标(通常位于屏幕左上角),双击这一图标,再双击相应的驱动器图标,继续双击文件夹的层次结构直至找到文件。这种方法将为每一组文件夹或单个变动的窗口显示一个分离的或重叠的窗口,这取决于 View |Options|Folder 的设置。找到想要的文件后,在文件名上单击右键,从快捷菜单中选择 Copy 就可拷贝文件。这种方法最适用于知道文件所在文件夹的层次结构,并想在桌面上同时显示两个窗口时。

找到并拷贝文件后,还必须将文件粘贴到所需位置中去。可用前边所讲的三种方法之一查找目标文件夹:①单击 Start |Find|Files or Folders;②单击 Start |Programs|Windows Explorer;③双击 My Computer,驱动器,文件夹。找到目标文件夹后,在文件夹图标或文件夹名字上单击右键,然后从快捷菜单中选择 Copy 即可完成拷贝。

如果想从硬盘上拷贝一文件到软盘的缺省目录下,可用前面所述三种方法之一找到文

件,再在文件名上单击右键,然后指向快捷菜单上的 Send To 并单击 Floppy 即可。

拷贝和粘贴文件的命令可用于快捷菜单,也可用于在屏幕顶端普通菜单栏上的 File 选项。第三种拷贝文件的方法是拖放。如果在 Windows Explorer 中排列各级文件夹或者排列 My Computer 的文件夹窗口,以使源文件和目标文件夹均可看见,此时可从一个文件夹中拖动一文件到另一文件夹。如果源文件与目标文件夹在同一磁盘上,文件将会被移动,如果文件夹在另一磁盘上,文件将被拷贝。若只想拷贝文件,而不是移动,在拖放时按住 Control 键即可。

1.5 启动 Excel

启动 Excel 有几种方法。一种常用的方法是单击 Start 按钮,再指向 Programs,也可指向 Office 95 或 Microsoft Office(如果安装的话)之一,并从展开的菜单中选择 Microsoft Excel。另外,如果在计算机桌面上有 Programs 或 Office 工具栏,也可单击 Microsoft Excel 按钮(一个 X 形的图标)。

启动 Excel 后,在 Excel 的菜单栏最右边会看到 Help 项。如果用户将系统设置为启动 Excel 时不带有新工作簿,屏幕显示将如图 1.7 所示,否则,屏幕显示如图 1.8 所示。

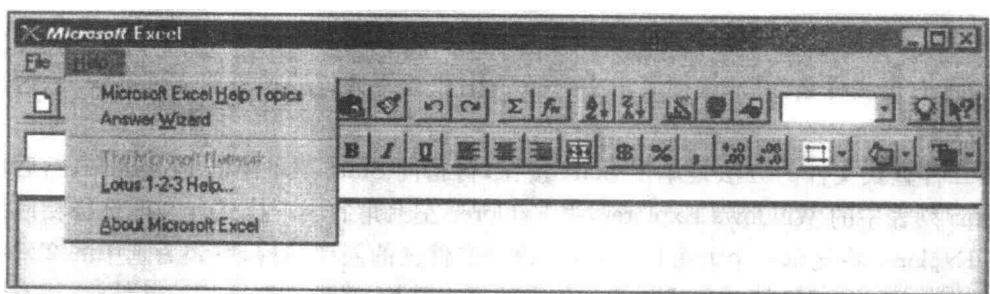


图 1.7 无工作簿的 Excel 打开屏幕

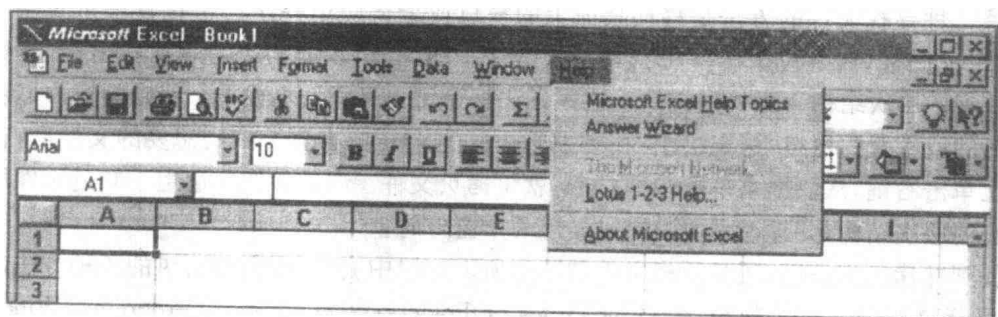


图 1.8 有新工作簿的 Excel 打开屏幕

启动 Excel 后,可用 Excel 的 File 菜单打开一个已存在的工作簿(File | Open)或新建一