

商务数据分析

主 编 宋沛军 白美娇 孙国喆
副主编 夏朋举 申 丽 赵露洁
孙晓俊 杨 想 张晓丽
余佩芳



中国传媒大学出版社
Communication University of China Press

· 北 京 ·

图书在版编目（CIP）数据

商务数据分析 / 宋沛军, 白美娇, 孙国喆主编. -- 北京 : 中国传媒大学出版社, 2025. 5.
ISBN 978-7-5657-3952-1

I . F712.3

中国国家版本馆 CIP 数据核字第 2025798BS7 号

商务数据分析

SHANGWU SHUJU FENXI

主 编 宋沛军 白美娇 孙国喆
策划编辑 温晓芳
责任编辑 温晓芳
封面设计 廖 源
责任印制 李志鹏

出版发行 中国传媒大学出版社
社 址 北京市朝阳区定福庄东街 1 号 邮 编 100024
电 话 86-10-65450528 65450532 传 真 65779405
网 址 <http://cucp.cuc.edu.cn>
经 销 全国新华书店

印 刷 三河市海新印务有限公司
开 本 787mm × 1092mm 1/16
印 张 17.5
字 数 404 千字
版 次 2025 年 7 月第 1 版
印 次 2025 年 7 月第 1 次印刷

书 号 ISBN 978-7-5657-3952-1 定 价 49.80 元

本社法律顾问: 北京嘉润律师事务所 郭建平

前言

PREFACE

随着大数据、云计算、人工智能等新一代信息技术的快速发展，商业经济进入数据化运营模式。党的二十大报告指出，要“加快建设制造强国、质量强国、航天强国、交通强国、网络强国、数字中国”。建设数字中国是数字时代推进中国现代化的重要引擎，是构筑国家竞争新优势的有力支撑。在数据化运营过程中会产生大量商务数据，且这些数据往往蕴藏着巨大价值，商家通过科学合理的方式对收集的数据进行分析，可以以这些数据为依据调整运营策略，不断优化市场调研、客户管理、店铺运营和商品销售等业务流程，从而在激烈的市场竞争中赢得一席之地。

商务数据分析的重要作用得到许多商家的认可，尤其是在电子商务应用领域，数据分析甚至已成为店铺运营管理人员的基本工作内容，部分商家还会专门设置数据分析岗位，这使得数据分析人才供不应求。鉴于此，我们立足商务数据分析岗位的实际需求，邀请业界专家编写了本书。

本书推荐学时安排见下表：

项目序号	项目标题	学时
项目一	筹备商务数据分析	8
项目二	使用数据分析平台	6
项目三	采集商务数据	6
项目四	预处理商务数据	4
项目五	可视化商务数据	4
项目六	调研分析市场数据	8
项目七	调研分析竞争对象	4
项目八	调研分析客户数据	8
项目九	统计分析商品数据	8
项目十	分析店铺运行数据	4
项目十一	制作商务数据分析报告	4
总计		64

本书具有以下特点。

1. 素质教育，润物无声

党的二十大报告指出：“育人的根本在于立德。”本书融入党的二十大精神，秉承能力教育与素质教育同向同行的理念，选取既能够对应相关知识点，又能够体现职业素养并与实际应用紧密相关的案例，将能力教育和素质教育完美融合，力求培养高素质、高水准的专业型人才。

2. 全新形态，全新理念

本书秉承“实例教学，讲练结合”的教学理念，采用“项目+任务”的体例形式，对标“1+X”电子商务数据分析职业技能等级标准，将重点知识内容按岗位需求模块化，突出职业能力培养目标，形成了集课、岗、赛、训一体化的新形态教材。

3. 图解详细，实用性强

本书内容全面，既包含商务数据分析的完整知识体系，又包含常用数据分析工具的相关操作，并将实际应用中的操作步骤进行详细图解。同时，本书还对案例实施过程中难以理解的概念和方法加以解释，让读者一目了然，有效提升读者的学习效率。

由于编者水平有限，书中难免存在疏漏和不足之处，敬请广大读者批评指正。

本书配备教学资源，请拨打电话 010-82967726 免费获取。

编者

目录

CONTENTS

项目一

筹备商务数据分析

任务一 认识商务数据	3
任务二 认识商务数据分析	11
任务三 认识分析工作流程	14
任务四 认识数据分析方法	17

项目二

使用数据分析平台

任务一 熟悉线上数据分析平台	29
任务二 熟悉线下数据分析平台	38

项目三

采集商务数据

任务一 认识商务数据采集	48
任务二 选择数据采集工具	55
任务三 实施商务数据采集	60

项目四

预处理商务数据

任务一 清洗采集的数据	79
任务二 整理采集的数据	86

项目五

可视化商务数据

任务一 熟悉商务数据可视化	96
任务二 熟悉可视化图形	103
任务三 分解可视化图表	110
任务四 熟悉识图方法	111

项目六

调研分析市场数据

任务一 采集市场数据	127
任务二 市场数据分析	130
任务三 市场数据分析方式	131
任务四 市场数据分析方法	137
任务五 分析市场数据	140

项目七

调研分析竞争对象

任务一 竞争对象分析的概念	156
任务二 竞争对象分析的步骤	157
任务三 竞争对象分析的方法	159

项目八

调研分析客户数据

任务一 建立客户指标	170
任务二 客户智能画像	175
任务三 客户价值分析模型	190

项目九

统计分析商品数据

任务一 商品定价与成本	199
任务二 商品的生命周期	201
任务三 商品库存分析	212

项目十

分析店铺运行数据

任务一 商品流量分析	235
任务二 商品销售数据分析	243
任务三 网络营销推广	251

项目十一

制作商务数据分析报告

任务一 认识商务数据分析报告	255
任务二 撰写商务数据分析报告	257

参考文献 /271





知识目标

- (1) 熟悉商务数据预处理的常用方法。
- (2) 掌握数据清洗的常用方法。
- (3) 掌握数据整理的常用方法。



技能目标

- (1) 能够根据实际需求判断当前数据存在的问题。
- (2) 能够根据当前数据存在的问题进行针对性的数据清洗。
- (3) 能够根据当前数据的特点, 结合数据分析的需要, 快速进行数据处理。



素养目标

- (1) 培养学生敏锐的问题意识, 能够迅速识别商务数据中存在的缺失、异常、冗余等问题。
- (2) 培养学生严谨认真、一丝不苟的工作态度, 保证数据的准确性和完整性。
- (3) 引导学生保持对新知识、新技术的学习热情, 以适应快速变化的商业环境。



思政园地

对数据进行预处理是一项比较繁杂的工作。在预处理的各个环节中, 既需要知道其原理, 还需要学会动手实操, 才能真正掌握数据清洗的应用。俗话说, “看花容易绣花难”, 任何事情要真正做好, 都要靠实践。实践是检验真理的唯一标准。

爱迪生 12 岁时, 沉迷于科学实验, 孜孜不倦地自学和实验, 16 岁便发明了每小时拍发一个信号的自动电报机。后来, 他又接连发明了自动数票机、第一架实用打字机、二重与四重电报机、自动电话机和留声机等。爱迪生并不满足于此, 1878 年 9 月, 爱迪生决定向电力照明这个堡垒发起进攻。他翻阅了大量有关电力照明的书籍, 决心制造出价钱便宜、经久耐用, 而且安全方便的电灯。

可见, 不管是学习还是研究, 我们一定要笃信躬行。尽信书不如无书, 只有把知识与经验有机结合, 通过实践印证所学, 才能真正学到知识。

任务一

清洗采集的数据

在实际工作中，初次采集到的数据会存在较多的问题，如数据缺失、数据重复等。存在问题的数据无法直接进行分析。因此，就需要进行一系列处理。将数据表中多余、重复的数据筛选出来并删除，将缺失、不完整的数据进行填补，将内容、格式错误的数据纠错或删除等一系列操作过程就叫数据清洗。数据清洗是对数据进行重新审查和校验的过程，目的在于提升数据的质量，确保数据的准确性、完整性和一致性。

一、清洗缺失数据

（一）缺失值清洗的概念

采集到的数据表或数据集中往往会存在某个或某些属性的值不完整或者空缺的情况，如果不进行处理，将会影响到随后的数据分析等使用。因此就需要将这些空缺的值补全，或者将缺失关键数据值的记录清除掉，这一工作就叫清洗缺失数据。

（二）缺失值产生的原因

在实际工作中，造成数据缺失的原因主要有以下 3 种。

（1）信息涉及隐私。有些信息无法获取，如在收集顾客婚姻状况和工作信息时，已婚人士的配偶等是无法获取的信息。

（2）人为原因疏漏。人为原因导致某些信息遗漏或被删除了。

（3）数据存储失败。数据收集或保存失败会造成数据缺失，如数据存储失败、存储器损坏、机械故障等。

（三）缺失值的处理方法

在实际工作过程中，对于缺失数据，通常有以下 2 种处理方法。

（1）取样本均值或众数代替缺失值。

（2）将缺失关键数据的记录删除。

二、清洗重复数据

数据采集来自不同的平台，有的甚至是很多人分工去采集，然后将这些数据汇集到一起的，因此往往会存在大量重复的数据记录。而重复数据会影响数据处理结果的正确性，导致数据分析出现偏差，因此需要将其删除。

在 Excel 等数据处理工具中，经常使用删除重复项、高级筛选、数据透视表、筛选公式等方法，找出重复的数据记录，并将其删除。实际工作中要根据以下具体情况进行处理。

(1)数据值完全相同的多条数据记录。对于数据值完全相同的数据，可以使用去重工具自动识别和删除重复的数据行。

(2)数据主体相同，但一个属性匹配到不同的多个值。对于这种情况，可以选择通过合并多列数据来实现去重。例如，可以将相同的记录合并，或者通过编写脚本来筛选出相同的唯一记录。

三、清洗错误标识符

在采集到的数据中，会存在各种形式的错误标识符，如#DIV/0！、#VALUE！等，这些通常表示单元格中的数据超出了该单元格的宽度，或者日期时间公式产生了一个负值，或者是键入公式的方式错误，或者引用的单元格错误导致的。

数据清洗过程中，清洗错误标识符是一个重要的步骤，它可以提高数据的整体质量和可靠性。常见错误标识符汇总见表4-1。

表 4-1 常见错误标识符汇总

序号	符号	错误原因	解决方案
1	####	单元格宽度不够，或在日期格式列中输入负的日期时间	拉大列宽，修改格式为常规
2	DIV/0！	数据相除运算中除数（分母）为0或者为空值	查看公式是否合理，如果合理，可以使用IFERROR进行规避
3	#VALUE！	数据类型问题导致错误	如果是输入错误就进行纠正；如果是数据类型不一样，就将类型进行统一
4	#N/A	进行函数运算，对应的值不存在	需要验证是否对应的值不存在，如果真的不存在，可以使用IFERROR函数规避
5	#NAME？	函数名输入错误，或单元格引用区域错误，或数值类型错误不符合计算规则，或文本字符引用时未规范使用英文双引号	检查基础数据是否有格式错误；检查函数公式是否有输入错误；检查文本内容是否为正确引用
6	#REF！	无法找到需要引用的单元格或区域	重新梳理函数计算逻辑，更新函数计算单元格或区域
7	#NUM！	超出了限定的数值计算范围	对数据进行拆分或者重新整理计算
8	#NULL！	引用单元格区域出现错误	重新梳理函数公式，检查对应的函数引用区域是否有书写错误

四、清洗格式不一致的数据

由于数据来源复杂，通常采集到的数据会以各种形态存在，如文本、表格、音频、视频等。另外，通过一系列工作将这些数据整合到一起之后，也会在同一列中出现格式不一样的情况。例如，常见的日期格式数据，在同一数据列中会有“2024年7月8日”“20240708”“2024-7-8”等格式。

不同格式的数据放在一起，将会影响数据分析结果的准确性。因此，需要对格式不一致的数据进行清洗。

五、清洗逻辑错误的数据

在采集到的数据中，往往还会存在逻辑错误的值。比如在年龄数据列中会出现大于200的数，在订单表中会出现发货日期早于购买日期的数据等。其中常见的逻辑错误情况有以下2种。

（一）数据中含有不合逻辑的字符

在处理这种数据时，主要是要去除不合逻辑的字符，如头或尾或中间的空格、姓名中的特殊字符和拼写错误等。这种情况需要以半自动校验、半人工方式来找出可能存在的问题，并删除不需要的字符。

（二）数据与字段要求不符

在处理数据中出现的内容和字段不匹配的问题时，需要采用详细识别、分类处理问题的方法。这部分的处理往往需要人工处理，尽量细致地检查，避免遗漏或误删。处理方法包括以下2种。

（1）针对某个字段的特殊值或异常值，可以通过检查业务逻辑或者调查来识别原因，并进行修正。

（2）对于乱码等无法处理的问题，除删除外，也可以先进行分类，再人工处理，提高清洗的效率和准确性。

六、清洗无价值数据

在进行数据预处理的过程中，受主观因素的影响，有时难以正确判断数据的价值，会将有价值的数据清洗掉。因此，在进行数据清洗时必须谨慎处理。清洗无价值数据有以下3点建议。

（一）理解业务需求和数据特点

需要充分理解业务需求和数据特点，明确需要分析的数据内容和目标输出结果，并在此基础上，筛选出与业务需求不相关或无用的数据。

（二）数据分析和统计

通过数据分析和统计，删除无效数据。对于没有明确业务含义、维度稀少或数据量过少的数据，可以考虑将其删除或归类统计，找到有价值的数据。

（三）根据业务需求调整数据量

不同的业务对数据量的需求也不同，如果数据量太大而无法快速有效地进行处理和分

析,可以根据实际需要和数据分析结果调整数据量。例如,可以通过筛选样本数据、区间采样、数据抽样等方法减少数据量,从而提高分析的效率和准确性。

实战演练——清洗



一、清洗缺失数据

(一) 任务描述

为了分析当前上海市的房产市场情况,小王使用八爪鱼采集器从网上抓取了相关数据,其中“开盘日期”和“户型”等字段有部分数据缺失,如图4-1所示,现需要对表中这两个字段的所有缺失数据进行清洗(使用版本为Microsoft Office 2016)。

#	A	B	C	D	E	F	G	H	I	J
	楼盘名	评分	类别	城市	均价	开盘日期	位置	地址	标签	户型
1	招商中旅·揽阅	9.1	普通住宅	上海	待定		闵行	闵行跃进路及永琴路交叉口	待售 普通住宅 精装修	三居(2)四居(1)
2	保利虹桥和颂	8.5	普通住宅	上海	36721	2023年12月27日首开	青浦	青浦毛家角路177弄	在售 普通住宅 精装修	二居(2)三居(2)四居(1)
3	港城悦领	9.2	普通住宅	上海	36025	2023年9月17日加推	浦东	浦东东方竹路55弄	在售 普通住宅 叠拼 精装修	三居(2)四居(3)
4	宝华紫薇花园(公园道)	9.6	普通住宅	上海	待定	预计2024年5月	普陀	普陀康泰路702弄	待售 普通住宅 精装修	二居(3)三居(5)
5	招商中旅·揽阅	9.1	普通住宅	上海	待定		闵行	闵行跃进路及永琴路交叉口	待售 普通住宅 精装修	三居(2)四居(1)
6	中建御璟园	9.1	洋房	上海	58820	二期2024年3月26日即将奉	奉贤	奉贤湖畔路与清朗路交叉口	在售 洋房 普通住宅 联排	三居(2)四居(4)五居(1)
7	招商臻境	9.3	普通住宅	上海	待定		浦东	浦东东至上南路、0901-09-01地	待售 普通住宅 叠拼	三居(3)四居(4)
8	尚湾林语	9.2	普通住宅	上海	60500	2023年12月6日首开	闵行	闵行闵园路1998号	在售 普通住宅 洋房	三居(1)四居(1)
9	中国铁建花语前湾	9.8	普通住宅	上海	68878	三期2024年3月19日加推	闵行	闵行朱建路开市客往东约400米处	在售 普通住宅 精装修	三居(2)四居(2)
10	大华星曜	9.1	普通住宅	上海	40172	2024年3月27日加推	闵行	闵行浦星公路与周航公路交叉口	在售 普通住宅	三居(2)四居(2)
11	保利·光合上城/跃城	9.2	普通住宅	上海	69900	光合上城2023年12月19日	闵行	闵行三莲路168弄	在售 普通住宅 洋房 精装修	三居(9)四居(3)
12	金樾·星海湾	7.7	普通住宅	上海	35500	2023年11月12日加推	浦东	浦东洲桥路及茂海路交汇处	在售 普通住宅 洋房 叠拼	三居(3)四居(1)
13	华发海上郡荟	9.2	普通住宅	上海	49100	首期2024年3月12日开盘	松江	松江长兴西路与城隍庙路交叉口东	在售 普通住宅 叠拼 精装修	三居(2)四居(1)
14	中环金茂府	9.2	普通住宅	上海	637	2024年3月22日首开	宝山	宝山区镇北路南陆路交叉路南100	待售 普通住宅 大平层 精装	二居(1)三居(1)
15	金地嘉峇汇	9.5	普通住宅	上海	46213	三批次2024年3月13日加推	嘉定	嘉定北至树屏路,东至城北路,	在售 普通住宅 叠拼 精装修	三居(2)四居(3)
16	安富中宸院	9.3	普通住宅	上海	69645	三期2024年3月29日加推	浦东	浦东康桥路199弄	待售 普通住宅 精装修	三居(6)
17	绿城·春晓园	9.4	普通住宅	上海	60910	二期2024年3月24日加推	青浦	青浦崧泽大道与金字圩路交汇处	在售 普通住宅 叠拼 精装修	三居(2)四居(2)
18	中環壹号院	9.0	普通住宅	上海	66000	2024年3月27日首开	宝山	宝山东至汇丰路,南至走马塘路,	待售 普通住宅	三居(2)四居(2)
19	绿城·光庭园	9.5	普通住宅	上海	75130	二期2024年3月23日加推	闵行	闵行华展路268弄	待售 普通住宅 精装修,毛坯	四居(1)
20	兴云间	8.8	洋房	上海	50600	2024年1月31日首开	松江	松江蔡家浜东路788弄	在售 洋房 联排 叠拼 毛坯	三居(1)四居(2)五居(1)
21	象屿交控·中环云悦府	9.4	普通住宅	上海	79033	2024年1月23日首开	普陀	普陀莲花路祁连山路交叉口	在售 普通住宅 精装修	三居(2)四居(2)
22	保利·海上瑔悦	9.3	普通住宅	上海	63000	二期2023年11月30日加	宝山	宝山丰翔路与太场路交汇处	在售 普通住宅 精装修	三居(4)四居(1)
23	浦发东望	9.1	普通住宅	上海	63000	2023年11月25日首开	浦东	浦东东至D05C-17地块,西至D05	在售 普通住宅 叠拼	三居(3)四居(1)
24	大华梧桐樾	8.8	普通住宅	上海	41432	2023年12月8日加推	嘉定	嘉定米夏路199弄	在售 普通住宅 叠拼	二居(2)三居(25)四居(8)
25	大华锦绣四季·澜悦	9.1	普通住宅	上海	54198	2023年4月3日加推	宝山	宝山东至南浦路126弄	在售 普通住宅 精装修	三居(2)四居(2)

图 4-1 上海市某时段的房产市场情况

(二) 主要步骤



步骤 1: 打开原始数据表格,选中数据区域。



步骤 2: 选中所有空值单元格。在“开始”选项卡下的“编辑”功能组中单击“查找和选择”按钮,点击“定位条件”命令,在弹出的“定位”对话框中,选中“空值”,如图4-2所示,点击“定位”后,所有的空值即可被一次性选中。

#	A	B	C	D	E	F	G	H	I	J
	楼盘名	评分	类别	城市	均价	开盘日期	位置	地址	标签	户型
1	招商中旅·揽阅	9.1	普通住宅	上海	待定		闵行	闵行跃进路及永琴路交叉口	待售 普通住宅 精装修	三居(2)四居(1)
2	保利虹桥和颂	8.5	普通住宅	上海	36721	2023年12月27日首开	青浦	青浦毛家角路177弄	在售 普通住宅 精装修	二居(2)三居(2)四居(1)
3	港城悦领	9.2	普通住宅	上海	36025	2023年9月17日加推	浦东	浦东东方竹路55弄	在售 普通住宅 叠拼 精装修	三居(2)四居(3)
4	宝华紫薇花园(公园道)	9.6	普通住宅	上海	待定	预计2024年5月	普陀	普陀康泰路702弄	待售 普通住宅 精装修	二居(3)三居(5)
5	招商中旅·揽阅	9.1	普通住宅	上海	待定		闵行	闵行跃进路及永琴路交叉口	待售 普通住宅 精装修	三居(2)四居(1)
6	中建御璟园	9.1	洋房	上海	58820	二期2024年3月26日即将奉	奉贤	奉贤湖畔路与清朗路交叉口	在售 洋房 普通住宅 联排	三居(2)四居(4)五居(1)
7	招商臻境	9.3	普通住宅	上海	待定		浦东	浦东东至上南路、0901-09-01地	待售 普通住宅 叠拼	三居(3)四居(4)
8	尚湾林语	9.2	普通住宅	上海	60500	2023年12月6日首开	闵行	闵行闵园路1998号	在售 普通住宅 洋房	三居(1)四居(1)
9	中国铁建花语前湾	9.8	普通住宅	上海	68878	三期2024年3月19日加推	闵行	闵行朱建路开市客往东约400米处	在售 普通住宅 精装修	三居(2)四居(2)
10	大华星曜	9.1	普通住宅	上海	40172	2024年3月27日加推	闵行	闵行浦星公路与周航公路交叉口	在售 普通住宅	三居(2)四居(2)
11	保利·光合上城/跃城	9.2	普通住宅	上海	69900	光合上城2023年12月19日	闵行	闵行三莲路168弄	在售 普通住宅 洋房 精装修	三居(9)四居(3)
12	金樾·星海湾	7.7	普通住宅	上海	35500	2023年11月12日加推	浦东	浦东洲桥路及茂海路交汇处	在售 普通住宅 洋房 叠拼	三居(3)四居(1)
13	华发海上郡荟	9.2	普通住宅	上海	49100	首期2024年3月12日开盘	松江	松江长兴西路与城隍庙路交叉口东	在售 普通住宅 叠拼 精装修	三居(2)四居(1)
14	中环金茂府	9.2	普通住宅	上海	637	2024年3月22日首开	宝山	宝山区镇北路南陆路交叉路南100	待售 普通住宅 大平层 精装	二居(1)三居(1)
15	金地嘉峇汇	9.5	普通住宅	上海	46213	三批次2024年3月13日加推	嘉定	嘉定北至树屏路,东至城北路,	在售 普通住宅 叠拼 精装修	三居(2)四居(3)
16	安富中宸院	9.3	普通住宅	上海	69645	三期2024年3月29日加推	浦东	浦东康桥路199弄	待售 普通住宅 精装修	三居(6)
17	绿城·春晓园	9.4	普通住宅	上海	60910	二期2024年3月24日加推	青浦	青浦崧泽大道与金字圩路交汇处	在售 普通住宅 叠拼 精装修	三居(2)四居(2)
18	中環壹号院	9.0	普通住宅	上海	66000	2024年3月27日首开	宝山	宝山东至汇丰路,南至走马塘路,	待售 普通住宅	三居(2)四居(2)
19	绿城·光庭园	9.5	普通住宅	上海	75130	二期2024年3月23日加推	闵行	闵行华展路268弄	待售 普通住宅 精装修,毛坯	四居(1)
20	兴云间	8.8	洋房	上海	50600	2024年1月31日首开	松江	松江蔡家浜东路788弄	在售 洋房 联排 叠拼 毛坯	三居(1)四居(2)五居(1)
21	象屿交控·中环云悦府	9.4	普通住宅	上海	79033	2024年1月23日首开	普陀	普陀莲花路祁连山路交叉口	在售 普通住宅 精装修	三居(2)四居(2)
22	保利·海上瑔悦	9.3	普通住宅	上海	63000	二期2023年11月30日加	宝山	宝山丰翔路与太场路交汇处	在售 普通住宅 精装修	三居(4)四居(1)
23	浦发东望	9.1	普通住宅	上海	63000	2023年11月25日首开	浦东	浦东东至D05C-17地块,西至D05	在售 普通住宅 叠拼	三居(3)四居(1)
24	大华梧桐樾	8.8	普通住宅	上海	41432	2023年12月8日加推	嘉定	嘉定米夏路199弄	在售 普通住宅 叠拼	二居(2)三居(25)四居(8)
25	大华锦绣四季·澜悦	9.1	普通住宅	上海	54198	2023年4月3日加推	宝山	宝山东至南浦路126弄	在售 普通住宅 精装修	三居(2)四居(2)

图 4-2 定位空值单元格



步骤 3: 选择处理方式。定位到空值后,可以选择“补齐数据”“删除记录”或者“不

处理”。例如，需要数据补齐，直接输入需要补充的内容，按“Ctrl+Enter”快捷键，进行批量填充（如用“默认值”将数据补齐），如图4-3所示。

#	A	B	C	D	E	F	G	H	I	J
1	楼盘名	评分	类别	城市	均价	开盘日期	位置	地址	标签	户型
2	招商中庭·悦园	9.1	普通住宅	上海	待定	默认值	闵行	闵行跃进路及永翠路交叉口	待售 普通住宅 带装修	三层(2) 四层(1)
3	保利虹桥御邸	8.5	普通住宅	上海	36721	2023年12月27日首开	青浦	青浦毛家角路177弄	在售 普通住宅 带装修	二层(2) 三层(2) 四层(1)
4	融城悦庭	9.2	普通住宅	上海	36025	2023年9月17日加推	浦东	浦东东方路55弄	在售 普通住宅 叠拼 带装修	二层(2) 四层(3)
5	宝华紫宸花园(公园道)	9.6	普通住宅	上海	待定	预计2024年5月	普陀	普陀景泰路702弄	待售 普通住宅 带装修	一层(3)
6	招商中庭·悦园	9.1	普通住宅	上海	待定	默认值	闵行	闵行跃进路及永翠路交叉口	待售 普通住宅 带装修	三层(2) 四层(1)
7	中建御璟园	9.1	洋房	上海	58820	二期2024年3月26日即将奉	奉贤	奉贤湖湾路与清韵路交叉口	在售 洋房 普通住宅 联排	三层(2) 四层(4) 五层(1)
8	招商瑞境	9.3	普通住宅	上海	待定	默认值	浦东	浦东东至上南路、0901-09-01地	待售 普通住宅 叠拼	三层(3) 四层(4)
9	尚湾林语	9.2	普通住宅	上海	60500	2023年12月6日首开	闵行	闵行沪闵路1998号	在售 普通住宅 洋房	一层(1) 四层(1)
10	中国铁建花语前湾	8.8	普通住宅	上海	68878	二期2024年3月19日加推	闵行	闵行朱建路开市客往东约400米处	在售 普通住宅 带装修	默认值
11	大华星曜	9.1	普通住宅	上海	40172	2024年3月27日加推	闵行	闵行浦星公路与同航公路交叉口	在售 普通住宅	三层(2) 四层(2)
12	保利·光合上城/辰城	9.2	普通住宅	上海	69900	光合上城2023年12月19日	闵行	闵行三连路168弄	在售 普通住宅 洋房 带装修	二层(9) 四层(3)
13	金港·星海湾	7.7	普通住宅	上海	35500	2023年11月12日加推	浦东	浦东洲陆路及茂海路交汇处	在售 普通住宅 洋房 叠拼	三层(3) 四层(1)
14	华茂海上郡荟	9.2	普通住宅	上海	49100	首期2024年3月12日开盘	松江	松江长兴西路与城隍路交叉口东	在售 普通住宅 叠拼 带装修	二层(7) 四层(3)
15	中环金茂府	9.2	普通住宅	上海	637	2024年3月22日首开	宝山	宝山环镇北路南陆路交叉口南104	待售 普通住宅 大平层 带装修	一层(1) 三层(1)
16	金地壹峯汇	9.3	普通住宅	上海	46213	三批次2024年3月13日加推	嘉定	嘉定北至城隍庙、东至城北路、	在售 普通住宅 叠拼 带装修	三层(4) 四层(3)
17	安高申宸院	9.3	普通住宅	上海	69645	二期2024年3月29日加推	浦东	浦东银泰路199弄	待售 普通住宅 带装修	三层(6)
18	绿城·春晓园	9.4	普通住宅	上海	60910	二期2024年3月24日加推	青浦	青浦崧泽大道与金字圩路交汇处	在售 普通住宅	三层(2) 四层(2)
19	中環置地中心	9.0	普通住宅	上海	66000	2024年3月27日首开	宝山	宝山东至汇丰路、南至走马塘路、	待售 普通住宅	
20	绿城·沁園	9.5	普通住宅	上海	75130	二期2024年3月23日加推	闵行	闵行华展路268弄	待售 普通住宅 带装修 毛胚	四层(1)
21	兴云间	8.8	洋房	上海	50600	2024年1月31日首开	松江	松江蔡家浜东路788弄	在售 洋房 联排 叠拼 毛胚	三层(1) 四层(2) 五层(1)
22	象屿文控·中环云悦府	9.4	普通住宅	上海	79033	2024年1月23日首开	普陀	普陀连克路祁连路交汇处	在售 普通住宅 带装修	三层(2) 四层(2)
23	保利·海上观悦	9.3	普通住宅	上海	63000	二期2023年11月30日加	宝山	宝山丰翔路与大场路交汇处	在售 普通住宅 带装修	三层(4) 四层(1)
24	浦发东望	9.1	普通住宅	上海	62000	2023年11月25日首开	浦东	浦东东郊D05C-17地块、西至D05	在售 普通住宅	三层(3) 四层(1)
25	大华枫庭	8.8	普通住宅	上海	41432	2023年12月18日加推	嘉定	嘉定米夏路199弄	在售 普通住宅 叠拼	一层(2) 三层(25) 四层(8)
26	大华锦绣四季·澜悦	9.1	普通住宅	上海	54198	2023年4月3日加推	宝山	宝山电台南路126弄	在售 普通住宅 带装修	三层(2) 四层(2)
27	十化·安峯	6.6	普通住宅	上海	164000	招商2020年12月6日开盘	普陀	普陀东安石盘支路 西至东安石盘	在售 普通住宅 联排	三层(14)

图 4-3 填充空值单元格

二、清洗重复数据

(一) 任务描述

某店铺类目结构月表数据中存在重复数据，如图4-4所示，现需要对表中重复数据进行清洗，以便后续进行数据分析。

工作簿2 - Excel										
文件 开始 插入 页面布局 公式 数据 审阅 视图 告诉我您想要做什么...										
G6 24,407.98										
	A	B	C	D	E	F	G	H	I	J
1	统计日期	一级类目	二级类目	叶子类目	访客数	支付商品数	支付金额	支付买家数	支付转化率	
2	2019-06-01 ~ 2019-06-30	彩妆/香水/美妆工具	美甲工具	美甲工具	17,070	119	17,453.78	613	3.59%	
3	2019-07-01 ~ 2019-07-31	其他	邮费	邮费	6	1	23.00	4	66.67%	
4	2019-03-01 ~ 2019-03-31	彩妆/香水/美妆工具	指甲油/美甲产品(新)	指甲彩妆	22,780	71	42,469.75	1,054	4.63%	
5	2019-04-01 ~ 2019-04-30	彩妆/香水/美妆工具	指甲油/美甲产品(新)	指甲彩妆	31,052	78	59,541.03	960	3.09%	
6	2019-05-01 ~ 2019-05-31	彩妆/香水/美妆工具	美甲工具	美甲工具	9,724	117	24,407.98	685	7.04%	
7	2019-04-01 ~ 2019-04-30	其他	邮费	邮费	11	1	41.00	3	27.27%	
8	2019-03-01 ~ 2019-03-31	彩妆/香水/美妆工具	美甲工具	美甲工具	46,043	103	46,332.33	1,567	3.40%	
9	2019-03-01 ~ 2019-03-31	其他	邮费	邮费	9	1	10.00	2	22.22%	
10	2019-07-01 ~ 2019-07-31	彩妆/香水/美妆工具	指甲油/美甲产品(新)	指甲彩妆	13,908	64	35,198.00	370	2.70%	
11	2019-06-01 ~ 2019-06-30	彩妆/香水/美妆工具	指甲油/美甲产品(新)	指甲彩妆	19,284	72	56,310.06	584	3.03%	
12	2019-06-01 ~ 2019-06-30	其他	邮费	邮费	3	0	0.00	0	0.00%	
13	2019-07-01 ~ 2019-07-31	其他	邮费	邮费	6	1	23.00	4	66.67%	
14	2019-02-01 ~ 2019-02-28	其他	邮费	邮费	26	1	728.80	6	23.08%	
15	2019-05-01 ~ 2019-05-31	彩妆/香水/美妆工具	美甲工具	美甲工具	9,724	117	24,407.98	685	7.04%	
16	2019-02-01 ~ 2019-02-28	彩妆/香水/美妆工具	美甲工具	美甲工具	68,930	95	40,806.06	1,508	2.19%	
17	2019-07-01 ~ 2019-07-31	彩妆/香水/美妆工具	美甲工具	美甲工具	9,861	119	15,459.15	467	4.74%	
18	2019-05-01 ~ 2019-05-31	其他	邮费	邮费	15	1	141.00	3	20.00%	

图 4-4 某店铺类目结构月表

(二) 主要步骤

步骤 1: 设置重复项筛选列。选中工作表中的数据区域，在“数据”选项卡下的“数据工具”功能组中，单击“删除重复项”，在弹出的“删除重复项”对话框中，选择要删除的列，注意“统计日期”“一级类目”“二级类目”“叶子类目”应同时选中，否则会造成误删，如图4-5所示。

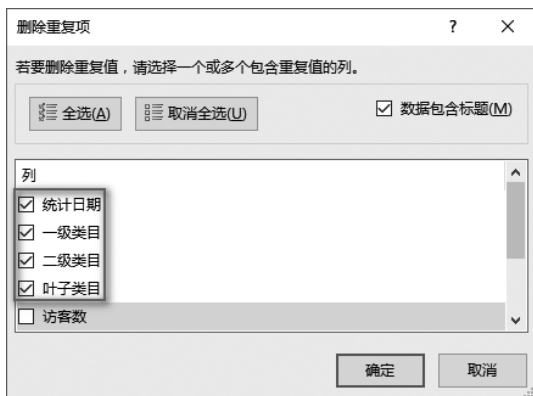


图 4-5 设置重复项筛选列

步骤 2: 删除重复项。单击“确定”按钮，完成重复项删除，将显示一条消息，指出有多少重复值被删除，多少唯一值被保留，如图 4-6 所示。

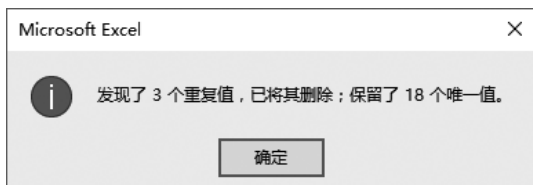


图 4-6 删除重复项结果提示

三、清洗错误标识符

(一) 任务描述

某公司商品采购表中，由于部分商品退货数量不清楚，退货率一栏显示错误标识符“#VALUE!”，如图 4-7 所示，现需要将退货率一栏中的错误标识符清洗掉。

供应链数据统计 - Excel									
文件 开始 插入 页面布局 公式 数据 审阅 视图 告诉我想做什么...									
G25									
A	B	C	D	E	F	G	H	I	J
1	采购日期	商品类别	商品名称	颜色	采购数量	采购单价	采购金额	不合格商品退货数量	退货率
2	2019/7/1	单肩包	真皮复古水桶包	棕色	20	208.0	4160.0	0	0.00%
3	2019/7/1	钱包/卡包/手零钱包	纯色定制长款女士钱包	浅黄色	20	69.0	1380.0	2	10.00%
4	2019/7/1	钱包/卡包/手零钱包	大容量多卡位名片卡包	香芋紫	30	9.9	297.0	4	13.33%
5	2019/7/1	钱包/卡包/手零钱包	大容量多卡位名片卡包	苹果绿	30	9.9	297.0	0	0.00%
6	2019/7/1	钱包/卡包/手零钱包	迷你简约零钱包	橄榄绿	20	79.0	1580.0	3	15.00%
7	2019/7/1	单肩包	真皮复古水桶包	黑色	20	208.0	4160.0	1	5.00%
8	2019/7/1	钱包/卡包/手零钱包	迷你简约零钱包	橡皮粉	20	79.0	1580.0	0	0.00%
9	2019/7/1	双肩包	港风软皮旅行包	黑色	50	128.0	6400.0	5	10.00%
10	2019/7/1	钱包/卡包/手零钱包	大容量多卡位名片卡包	雅致灰	30	9.9	297.0	2	6.67%
11	2019/7/1	单肩包	时尚简约链条小包	黑色	20	119.0	2380.0	5	25.00%
12	2019/7/5	钱包/卡包/手零钱包	纯色定制长款女士钱包	浅绿色	50	69.0	3450.0	未采集	#VALUE!
13	2019/7/5	钱包/卡包/手零钱包	纯色定制长款女士钱包	浅黄色	80	69.0	5520.0	未采集	#VALUE!
14	2019/7/5	钱包/卡包/手零钱包	纯色定制长款女士钱包	浅粉色	80	69.0	5520.0	未采集	#VALUE!
15	2019/7/5	钱包/卡包/手零钱包	迷你简约零钱包	橡皮粉	50	79.0	3950.0	未采集	#VALUE!
16	2019/7/5	单肩包	时尚简约链条小包	红色	10	119.0	1190.0	未采集	#VALUE!
17	2019/7/5	钱包/卡包/手零钱包	迷你简约零钱包	橄榄绿	50	79.0	3950.0	未采集	#VALUE!
18	2019/7/5	斜挎包	欧美印花女士包	印花	30	138.0	4140.0	未采集	#VALUE!

图 4-7 某公司商品采购表

(二) 主要步骤

步骤 1: 修改退货率计算公式。将鼠标定位在出现错误标识符的单元格 I12 中，在公式编辑栏中，修改原公式“=H12/E12”为“=IFERROR(H12/E12, '/')”，如图 4-8 所示。

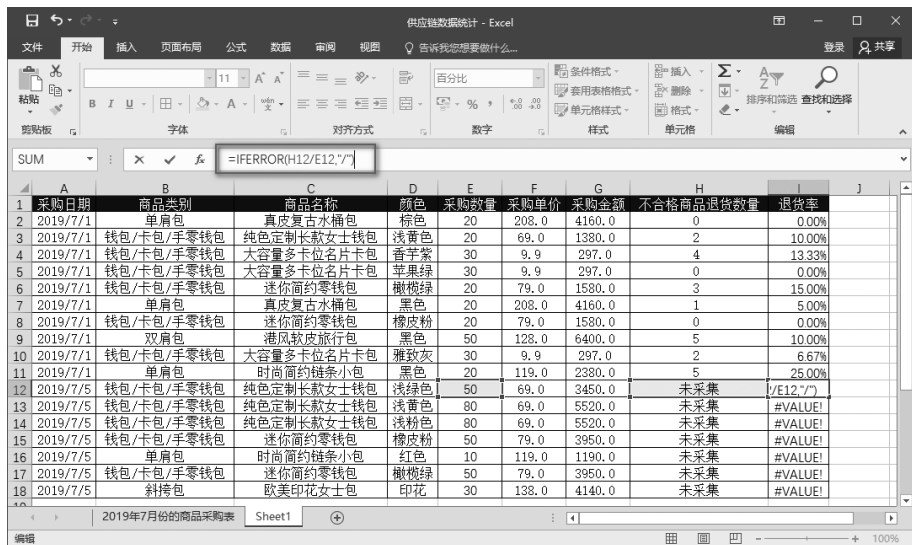


图 4-8 修改退货率计算公式

步骤 2: 填充退货率各单元格。确认输入后，I12 单元格中的“#VALUE!”被修改为“/”，然后拖动 I12 单元格的填充柄，将其他发生错误的单元格进行填充修改，如图 4-9 所示。

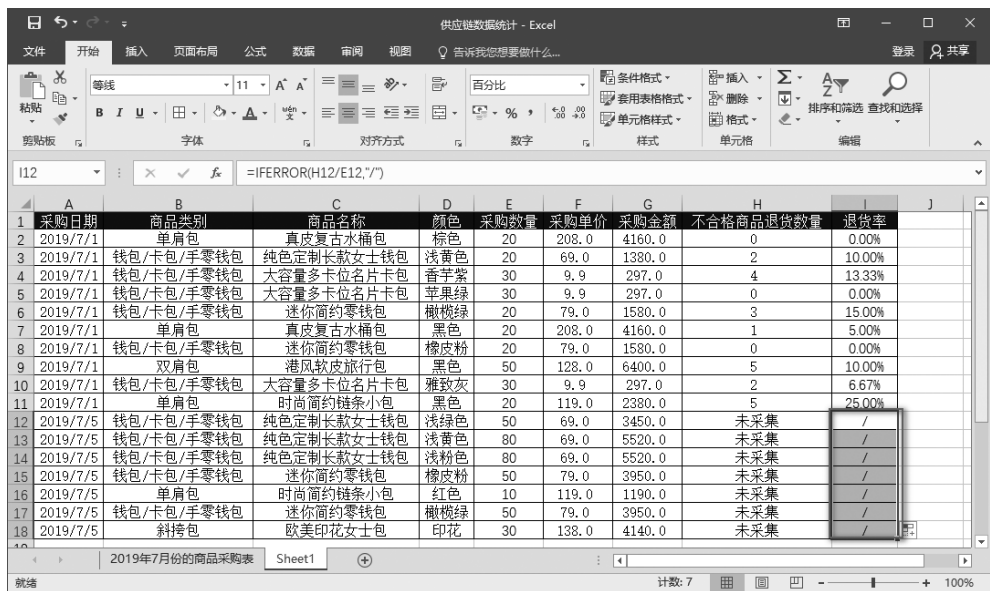


图 4-9 填充退货率各单元格

任务二

整理采集的数据

采集到的数据往往会纠缠在一起，不便于直接进行下一步数据分析，需要使用字段拆分、数据转换、字段匹配、数据抽取等方法进行数据整理。

一、字段拆分

字段拆分是指将一个包含多个数据项的字段拆分成多个字段，以便后续的数据处理和分析。

字段拆分在数据清洗和数据整理过程中很常见，特别是面对刚采集到的原始数据时，字段可能存在合并、重复或格式错误的情况，需要将其拆分为单独的字段进行分析。例如，从网络上抓取到运营相关岗位的需求数据，图中“名称”列的数据，每单元格中有“工作地点”“岗位描述”等内容独立的数据，在具体开展数据分析工作之前，需要将这样的数据进行拆分，如图 4-10 所示。

	A	B	C	D	E	F
1	名称	薪水	关键词	fl	comp_icons	cate
2	魏都城区 运营 速卖通假发POP店铺	3500~7000元/月	加班补助	安徽贝达荣发制品有限公司	1年	速卖通运营
3	魏都城区 速卖通假发运营	3500~7000元/月	五险一金	安徽贝达荣发制品有限公司	1年	速卖通运营
4	魏都城区 运营助理	3000~4000元/月		许昌壮志文化传媒有限公司	1年	运营专员/助理
5	魏都城区 本地生活服务运营	5000~8000元/月		河南庚泰科技有限公司	1年	运营专员/助理
6	魏都城区 无须经验无晚班运营	面议	加班补助	山东伊艺文化传媒有限公司	1年	运营专员/助理
7	万达广场 运营助理	3000~6000元/月		许昌丝弦发制品有限公司	2年	运营专员/助理

图 4-10 某地区运营相关岗位需求数据

二、数据转换

数据转换是指将需要进行计算的字段通过特殊处理，对数据的格式或结构进行转换，做好数据处理的前期准备。

数据转换包括数据表的行列互换、字符与数值转换等。例如，某公司的商品采购数据表，根据不同的数据分析需求，需要将数据进行行列转换。原本作为字段名称的“产品 1”“产品 2”……经过转换，就变成了字段值，如图 4-11 所示。

	A	B	C	D	E	F	G	H
1	产品名称	产品1	产品2	产品3	产品4	产品5		
2	单件采购成本（元）	9.9	7.9	6.9	5.9	8.9		
3	预估年推广成本（元）	35000	37000	46000	25000	50000		
4	直通车推广成本（元）	10000	10000	15000	5000	20000		
5	预估年物流成本（元）	58000	55000	60000	48000	62000		
6	快递费	5000	6000	7000	7000	6000		
7	预估年销量（件）	30000	28000	20000	15000	24000		
8	销售价格（元/件）	15.9	13.9	15.9	13.9	17.9		
9								
10								
11								
12								
13	产品名称	单件采购成本（元）	预估年推广成本（元）	直通车推广成本（元）	预估年物流成本（元）	快递费	预估年销量（件）	销售价格（元/件）
14	产品1	9.9	35000	10000	58000	5000	30000	15.9
15	产品2	7.9	37000	10000	55000	6000	28000	13.9
16	产品3	6.9	46000	15000	60000	7000	20000	15.9
17	产品4	5.9	25000	5000	48000	7000	15000	13.9
18	产品5	8.9	50000	20000	62000	6000	24000	17.9

图 4-11 数据表行列转换示意图

1. 数据行列互换

在 Excel 工具中，可以使用粘贴命令中的转置进行数据行列互换。

2. 数值转字符

在 Excel 表格中输入数据时，会默认使用数值型数据。如果数字太长，会变成科学记数法，不利于查看数据，甚至导致数据出错。对于这种情况，可以使用数据分列等方法进行处理。

3. 字符转数值

字符转数值可以采用数据分列的方法处理，也可以直接选中要转换的数据列，点击数据列前出现的提醒符号，在选项中点击“转换为数据”，即可将文本型字符转换为数值。

三、字段匹配

字段匹配是指根据某个关键字段或条件，在不同的数据集中查找并获取相应的值或信息。字段匹配可以在同一数据表中根据对应的值进行条件筛选，也可以在一个数据表中查找另一个数据表中对应的数据。

在 Excel 工具中，常用的数据匹配方法有以下 8 种。

1.VLOOKUP 函数

VLOOKUP 函数可用于在垂直方向上查找匹配的值。它可以在一个范围内查找某个关键字段的值，并返回该字段对应的值或信息。

VLOOKUP 函数的语法为“VLOOKUP (lookup_value,table_array,col_index_num,[range_lookup])”。其中，“lookup_value”表示要查找的关键字段的值；“table_array”表示要查找的数据表；“col_index_num”表示返回值所在的列号；“range_lookup”为可选参数，设置为

FALSE时表示精确匹配，设置为TRUE或省略时则表示模糊匹配。

2. INDEX MATCH 函数

INDEX MATCH 函数是一种更灵活、强大的数据匹配方法，它结合了INDEX和MATCH两个函数，可以按行或列查找并返回匹配值。INDEX函数用于返回数据表中指定位置的值，而MATCH函数用于在一个范围内查找匹配的值得位置。通过在垂直和水平方向上组合使用这两个函数，可以实现更复杂的数据匹配需求。

3. SUMIF 函数

SUMIF函数用于根据某个条件筛选数据并进行求和。它可以根据指定的条件，在一个范围内筛选出满足条件的数据，并对其对应的数值进行求和。SUMIF函数的语法为“SUMIF(range,criteria,[sum_range])”。其中，“range”是要筛选的范围；“criteria”是筛选的条件；“sum_range”是可选参数，表示要求和的范围。例如，可以使用SUMIF函数来统计某个字段满足条件的数据的总和。

4. 标准筛选和高级筛选

Excel还提供了标准筛选和高级筛选功能，用于根据条件筛选数据。标准筛选可以通过选择某列中的值来快速筛选数据，而高级筛选则可以根据自定义的条件进行更灵活的筛选，还可以将结果复制到其他位置。这些筛选功能可以帮助用户快速找到匹配的数据。

5. 数据透视表

数据透视表是Excel中一种强大的数据分析工具，可以用于对大量数据进行汇总、分组和分析。通过创建数据透视表，可以轻松实现数据的匹配和汇总。可以根据需要选择字段作为行、列和值，然后根据指定的条件和计算方式对数据进行匹配和汇总。

6. 自动筛选功能

Excel的自动筛选功能可以帮助用户快速筛选出符合特定条件的数据。可以通过在表头上点击筛选按钮，根据某个字段的值进行筛选。在弹出的筛选框中，可以选择指定字段的匹配条件，并根据需要进行多个条件的组合筛选。

7. 条件格式化

Excel的条件格式化功能可以根据一定的条件对数据进行样式的设置。通过设置条件，可以将匹配的数据标记出来，使它们在整个数据表中更加突出。可以根据具体需求设置颜色、字体、边框等样式，以便更好地显示匹配的数据。

8. 数据验证

数据验证是一种对数据进行限制和检查的功能，可以在选定的单元格范围内设置验证条件，如只允许输入特定的数值范围、从预设的列表中选择等。通过设置数据验证规则，可以确保输入的数据满足一定的条件，避免错误或无效数据的出现。

四、数据抽取

数据抽取是指从采集到的原始数据中提取所需的数据子集或特定数据项的过程。数据抽取是数据预处理的重要步骤，它为后续的数据分析和建模提供了基础。

在Excel工具中，常用的数据抽取函数有LEFT、RIGHT、WEEKDAY等。

1.LEFT函数

LEFT函数主要用于获取字符串左边指定个数的字符。LEFT函数的语法为“LEFT(text,[num_chars])”。其中,“text”为要取得给定值的文本数据源;“num_chars”表示需要从左开始算提取几个字符数,其中每个字符按1计数。

2.RIGHT 函数

RIGHT 函数主要用于获取字符串右边指定个数的字符。RIGHT 函数的语法为“RIGHT (text,[num_chars])”。其中,“text”为要取得给定值的文本数据源;“num_chars”表示需要从右开始算提取几个字符数,其中每个字符按 1 计数。

3.WEEKDAY函数

WEEKDAY 函数主要用于返回代表一周中第几天的数值，是一个 1 到 7（或 0 到 6）之间的整数。WEEKDAY 函数的语法格式为“WEEKDAY(date,type)”。其中，“date”为日期；“type”表示返回值是从 1 到 7 还是从 0 到 6，以及从星期几开始计数，如省略则返回值为 1 到 7，且从星期日起计。

实战演练——字段拆分和匹配



一、字段拆分

(一)任务描述

数据表“针织裙销量TOP220”中，“地址”列包含了省份和城市两项内容，如图4-12所示，现在需要将省份和城市拆分开。

A	B	C	D	E	F	G	H	I
商家标题	价格	收货人数	昵称名	平台	店铺类型	地址	评论数	收藏人数
夏季短款针织小坎裙女修身显瘦外搭裙开衫开薄款披肩白色小外套	24.99	2092	xiaolx24	淘宝	非金店铺	广东 广州	2,815	60,282
法式外穿针织小坎裙白色带肩带春秋款高级感半截小开衫小外套	49	2089	南苑久南	淘宝	非金店铺	广东 东莞	6,304	23,326
春秋外搭裙针织防晒衫女夏季宽松薄款防晒衣长袖针织罩衫小衣	39	2066	喜蝶诗旗舰店	天猫	旗舰店	广东 深圳	3,910	37,402
小披肩夏季搭配针织的上衣冰丝针织开衫女薄款冰丝防晒衫外搭针织衫	29.97	2033	xiaolx24	淘宝	非金店铺	广东 广州	1,669	8,978
春秋薄款针织领开衫女短款空调衫外搭防晒衣针织冰丝防晒衣外套	19.98	1026	xiaolx24	淘宝	非金店铺	广东 广州	2,272	50,697
披肩夏季薄款搭配针织的上衣冰丝针织开衫女短款防晒衫外搭针织衫	39.8	1011	冰衣女装吧	淘宝	非金店铺	广东 东莞	350	112,004
夏季防晒开衫外搭针织衫女夏季短款防晒衫小披肩空调衫薄款外套	39.8	1011	猪猪齐美齐	淘宝	非金店铺	广东 东莞	3,100	32,374
你家halo lala!邻家富家女/季可拆卸蝴蝶结连两件针织连衣防晒外套	49.8	1095	韩韩韩美宝	淘宝	非金店铺	广东 广州	658	49,639
防晒开衫女薄款披肩夏季搭配吊带裙外搭吊带针织衫针织上衣小外套	39	1067	尚蒂伊旗舰店	天猫	旗舰店	广东 广州	1,018	11,525
法式针织开衫女春季显瘦小坎裙长袖带袖开衫搭配吊带裙披肩外搭开衫	49	1061	曼妮奴旗舰店	淘宝	非金店铺	广东 东莞	651	6,173
冰丝防晒开衫防晒衫女外穿夏季空调衫吊带裙外搭针织衫小外套	39.9	1001	所珍服饰店	天猫	旗舰店	广东 广州	177	6,601
碎花连衣裙女夏季2022新款超仙收腰吊带裙针织开衫针织上衣两件外套	39.8	1096	小宅女购物	淘宝	非金店铺	广东 广州	1,298	35,490
早秋短款冰丝修身显瘦白色连衣长裙小个子蓬蓬裙夏季针织开衫外套季	48.98	1035	小魔童定制女装	淘宝	金店铺	广东 广州	3,464	25,143
Dogsgin 02in 2022年春夏减龄超短款上衣蓬蓬显瘦针织连衣外套	41	1092	czh最佳女友	淘宝	金店铺	广东 广州	228	19,399
夏季薄款冰丝短袖开衫女外穿两面穿空调衫搭配针织上衣短款宽松外套	38.8	1084	笑笑88887	淘宝	金店铺	河北 邢台	825	37,317
复古纯欲温柔风穿搭显瘦薄碎花气质吊带长裙长裙长款防晒开衫	88	1074	赵越先生	淘宝	金店铺	河北 邢台	298	35,579
POLO领短袖开衫女春秋2022新款辣妹黑灰色内搭针织包臀短款	39.8	1056	空白格_依	淘宝	非金店铺	广东 广州	6,706	29,040
夏季薄款防晒开衫女短款长袖防晒衫披肩上衣空调衫防晒开衫外套	39	1050	vicky1990923	淘宝	金店铺	广东 东莞	832	37,100
夏季薄款针织衫短款宽松轻薄防晒女秋外搭针织开衫防晒空调衫外套	39	1067	旗旗旗舰店	天猫	旗舰店	广东 广州	181	4,287
法式小开衫针织薄款女外穿夏季薄款空调衫吊带裙长款防晒开衫	39.9	1001	芯莎旗舰店	天猫	旗舰店	广东 广州	330	5,434
法式小坎裙防晒开衫防晒开衫夏季薄款防晒带肩带修身显瘦高腰小坎裙	48	1043	丹豆旗舰店	天猫	旗舰店	河北 邢台	562	7,000
23年新款 不规则针织开衫吊带裙两件外套夏季2022新款气质温柔上衣	138	996	沐沐大码	淘宝	金店铺	广东 广州	2,153	9,844
法式薄款法式系带针织开衫里裙修身显瘦包腰小个短款针织连衣裙	45	984	荃荃mccan	淘宝	金店铺	广东 广州	2,927	27,387

图 4-12 针织裙销量数据表

(二)主要步骤

步骤 1: 为拆分内容准备空间。观察将要拆分的内容, 看拆分后将多出来几列, 在将要拆分的列后边插入相应的空白列, 如图 4-13 所示。

A	B	C	D	E	F	G	H	J
	商品标题	价格	收货人数	掌柜名	平台	店铺类型	地址	评论数 收藏人数
1	夏季短款针织防晒衫女修身显瘦半衫外搭配裤子春秋薄款披肩白色小外套	24.98	2002	xiaol24	淘宝	金牌店铺	广东 广州	2,818 60,282
2	法式外穿针织衫女短款吊带裙外搭上衣春秋高领显瘦小开衫小外套	49	2089	南境久萌	淘宝	金牌店铺	广东 东莞	6,304 23,226
3	春秋外搭针织衫防晒开衫女夏季薄款披肩开衫袖长针织衫上衣	39	2006	喜薇诗服饰店	天猫	旗舰店	广东 深圳	3,910 37,402
4	小披肩外搭配裤子上的衣春秋外穿针织衫女款短款防晒防雨防晒上衣	29.97	2033	xiaol24	淘宝	非金牌店铺	广东 广州	1,669 8,978
5	春秋短款针织衫女短款防晒开衫外搭配裤子春秋防晒外套	19.98	1026	xiaol24	淘宝	金牌店铺	广东 广州	2,272 50,697
6	春秋外穿针织衫女春秋外穿针织衫女短款防晒开衫小开衫春秋上衣	39.8	1040	南境久萌	淘宝	金牌店铺	广东 东莞	3,325 23,226
7	春秋外穿针织衫女春秋外穿针织衫女短款防晒开衫小开衫春秋上衣	39.8	1011	南境久萌	淘宝	金牌店铺	广东 东莞	3,100 32,374
8	春秋外穿针织衫女春秋外穿针织衫女短款防晒开衫小开衫春秋上衣	49.8	1095	韩韩美衣	淘宝	非金牌店铺	广东 广州	658 49,629
9	防晒开衫女薄款披肩夏季短袖开衫外搭外搭半衫春秋针织衫上衣小外套	38	1067	南寄伊服旗舰店	天猫	旗舰店	广东 广州	1,018 11,525
10	法式针织开衫女春秋夏季小开衫显瘦短款两件针织连衣春秋外套	49	1061	曼奴奴服饰有限公司	淘宝	非金牌店铺	广东 东莞	651 6,175
11	冰丝薄款针织开衫防晒外套女夏季薄款配裤子带帽外搭春秋上衣	39.9	1001	所珍旗舰店	天猫	旗舰店	广东 广州	177 6,601
12	碎花女短款外穿春秋2022新款短款上衣带裤子针织开衫外搭外套	39.98	1096	衣衣衣衣衣衣	淘宝	金牌店铺	广东 广州	1,398 35,490
13	春秋外穿针织衫女春秋外穿针织衫女短款防晒开衫小开衫春秋上衣	39.8	1033	南境久萌	淘宝	金牌店铺	广东 东莞	3,404 23,226
14	DoggyRun 8mm mm 2022年春夏显瘦短款上衣修身显瘦针织连衣	41	1092	cah旗舰店	天猫	金牌店铺	广东 广州	228 19,499
15	新款春秋针织开衫防晒衫女两面穿春秋短款上衣带裤子春秋外套	86	1084	笑哭88887	淘宝	金牌店铺	河北 邢台	845 31,377
16	古纯晚温丝来衣套装女显瘦早晚薄花气质吊带连衣春秋针织服饰	38.88	1074	赵越先生	淘宝	金牌店铺	广东 广州	269 35,579
17	POLO领个性感打枣开衫女春季2022新款显瘦黑色上衣针织开衫配裤子	38.8	1056	空白格_依衣	淘宝	金牌店铺	广东 广州	6,766 29,040
18	夏季冰丝防晒开衫女短款防晒开衫配裤子上衣春秋针织衫外搭	59	1050	vicky1990923	淘宝	金牌店铺	广东 东莞	832 37,100
19	夏季针织开衫短款显瘦+轻薄小披肩女针织开衫防晒开衫空调外套	39	1067	南境旗舰店	天猫	旗舰店	广东 广州	181 2,287
20	春秋小开衫开衫女春秋外搭外搭外搭外搭外搭外搭外搭外搭外搭	39.8	1067	南境旗舰店	天猫	旗舰店	广东 广州	330 2,287
21	春秋小开衫开衫女春秋外搭外搭外搭外搭外搭外搭外搭外搭外搭	48	1043	南境旗舰店	天猫	旗舰店	河北 邢台	569 7,000

图 4-13 为拆分内容准备空间

步骤2: 拆分数据列。选中要拆分的数据列，在数据选项卡中找到并点击分列按钮，选择“分隔符号”，如图 4-14 所示，选择“空格”，如图 4-15 所示，点击“下一步”，直到完成。

图 4-14 选择用分隔符号分隔每个字段

[illegible]

图 4-15 选择用空格作为拆分标志

步骤3: 给拆分后各数据列添加字段名。地址列拆分后，分为省份和城市两列，分别给这两列命名，如图 4-16 所示。

商品标题	价格	收货人数	掌柜名	平台	店铺类型	省份	城市	评论数	收藏人数
夏季短袖针织防晒衫女修身薄款短袖开衫薄款披肩白色小外套	24.98	2002	xiaols24	淘宝	非金牌店铺	广东	广州	2,818	60,382
法式外穿针织小坎肩纯色吊带裙外搭罩衣春秋高领感半截小开衫女潮	49	2089	南城久尚	淘宝	非金牌店铺	广东	东莞	6,304	23,226
春秋外搭配裙子防晒开衫女夏季宽松薄款披肩袖冰丝针织罩衫上衣	39	2006	喜蝶诗旗舰店	天猫	旗舰店	广东	深圳	3,910	37,402
小披肩夏季配裙子的上衣冰丝针织开衫女薄款短款防晒吊带外搭罩衫	29.97	2033	xiaols24	淘宝	非金牌店铺	广东	广州	1,669	8,978
春秋薄款V领针织开衫女短款空调衫披肩外搭配裙子冰丝防晒罩衣外套	19.98	1026	xiaols24	淘宝	非金牌店铺	广东	广州	2,272	50,697
披肩夏季薄款配裙子的上衣冰丝针织开衫女短款防晒吊带外搭空调衫	38	1040	衣淘女装吧	淘宝	非金牌店铺	广东	东莞	350	112,904
夏季防晒开衫外搭冰丝针织衫女短款配吊带裙小披肩空调衫薄款外套	39.8	1011	猪宝贝齐	淘宝	非金牌店铺	广东	东莞	3,100	32,374
你shalo lala甜酷富家女/夏季可拆卸蝴蝶结假两件针织连衣裙薄款	49.8	1095	韩韩美衣	淘宝	非金牌店铺	广东	广州	658	49,629
防晒开衫女薄款披肩夏季搭配吊带裙外搭罩衫冰丝针织上衣小外套	38	1067	尚帝伊旗舰店	天猫	旗舰店	广东	广州	1,018	11,525
法式针织开衫女春季夏天小坎肩长袖配吊带裙外搭罩衫短款披肩外套	49	1061	曼织奴服饰有限公司	淘宝	非金牌店铺	广东	东莞	651	6,175
冰丝薄款针织防晒开衫女夏季薄款配吊带裙外搭罩衫空调上衣	39.9	1001	所珍旗舰店	天猫	旗舰店	广东	广州	177	6,601
碎花连衣裙女春夏2022年新款超仙收腰吊带裙子针织开衫外套两件套	39.98	1096	小七女大购物	淘宝	非金牌店铺	广东	广州	1,295	35,490
早秋纯棉风高腰修身吊带约会连衣裙小个子蓬蓬裙针织开衫外套春季	48.98	1035	小魔鹿定制女装	淘宝	金牌店铺	广东	广州	3,464	25,413
DoggyQin 碧nun nun/2022年春夏减龄短款上衣修身显瘦针织连衣裙	41	1092	czh最佳摄影	淘宝	金牌店铺	广东	广州	228	14,399
夏季新款冰丝短袖针织开衫女两面穿空调衫配裙子上衣短款宽松外套	86	1084	笑笑88887	淘宝	金牌店铺	河北	邢台	825	31,377
复古纯棉温柔风套装女收腰显瘦碎花气质吊带连衣裙长袖针织开衫薄	38.88	1074	赵露先生	淘宝	金牌店铺	广东	广州	269	35,579
POLO领性感打底连衣裙女春季2022新款辣妹黑色内搭针织包臀短裙子	39.8	1056	空白格_依	淘宝	非金牌店铺	广东	广州	6,766	29,040
夏季冰丝防晒开衫女短款长袖薄外搭配裙子披肩上衣空调衫针织外套	59	1050	vicky1990923	淘宝	金牌店铺	广东	东莞	832	37,100
夏季配裙子罩衫短款透气轻薄小披肩女冰丝针织开衫防晒罩衫外套	39	1067	淘派旗舰店	天猫	旗舰店	广东	广州	181	4,287
冰丝小针织开衫薄款披肩夏季吊带裙外搭空调衫短款长袖防晒外套	39	1001	芯妍晴旗舰店	天猫	旗舰店	广东	广州	330	5,439
法式小坎肩针织防晒开衫披肩夏季薄款配吊带裙修身外搭高腰长袖	48	1043	丹豆旗舰店	天猫	旗舰店	湖北	鄂州	562	7,004

图 4-16 拆分后的效果

二、字段匹配

(一)任务描述

现有一张商品信息表，如图 4-17 所示，其“供货商”一列的内容需要从供货商信息表，如图 4-18 所示中进行查找，从而匹配到正确数据。商品信息表中的数据量很大，如果一个个地查找，然后进行数据填充，工作量太大且消耗时间长，会影响工作进度，请采用快捷有效的方法完成此任务。

A	B	C	D	E	F	G	H	I
1	店铺商品信息							
2	商品代码	商品名称	供货商	时间	单价	数量	合计金额	
3	BIE45-93	休闲鞋		2024.3	¥ 199.00	3		
4	BIE45-93	休闲鞋		2024.3	¥ 199.00	3		
5	BIE45-93	休闲鞋		2024.3	¥ 199.00	3		
6	BIE45-93	休闲鞋		2024.3	¥ 199.00	3		
7	BSW45-23	板鞋		2024.3	¥ 129.00	4		
8	BSW45-23	板鞋		2024.3	¥ 129.00	4		
9	BSW45-23	板鞋		2024.3	¥ 129.00	4		
10	BSW45-23	板鞋		2024.3	¥ 129.00	4		
11	GJD23-56	板鞋		2024.1	¥ 99.00	8		
12	IE56-13	休闲鞋		2024.2	¥ 299.00	3		
13	IE56-13	休闲鞋		2024.2	¥ 299.00	3		
14	IE56-13	休闲鞋		2024.2	¥ 299.00	3		
15	IE56-13	休闲鞋		2024.2	¥ 299.00	3		
16	IOE12-45	板鞋		2024.2	¥ 168.00	5		
17	IOE12-45	板鞋		2024.2	¥ 168.00	5		

图 4-17 商品信息表

A	B	C	D	E	F	G	H	I
1	供货商信息表							
3	供货商代码	供货商名称	产品代码	产品名称	联系人	联系方式	城市	银行卡号
4	KDN-51	达××服装厂	GJD23-56	板鞋	张成	152****4565	上海	4581231265484580
5	FD-26	恩××鞋业厂	JSE89-02	运动鞋	林小然	155****5689	广州	4367125000125495
6	IEJ-32	索××鞋业	JFA58-03	板鞋	范冰	183****4542	深圳	6222801254953245
7	ODK-16	科××鞋业有限公司	RDS12-12	板鞋	夏雨雨	189****7851	珠海	4270202156354521
8	EDS-25	开××鞋业有限公司	IE56-13	休闲鞋	李心洁	178****4544	深圳	4270305612548626
9	GPQ-12	金××鞋服有限公司	IOE12-45	板鞋	赵大鹏	187****4578	厦门	6222801452354521
10	EIJ-10	拓××时尚鞋业有限公司	BIE45-93	休闲鞋	何浩新	153****4581	上海	404009123220212122
11	FDD-26	佳××鞋业有限公司	BSW45-23	板鞋	李小红	183****4575	杭州	552559451236548921
12	IEJ-12	乐××服装有限公司	XEW16-45	板鞋	吴军	181****7454	广州	427020152345678521
13	ICE-55	美××鞋服有限公司	SIE20-45	运动鞋	孙红	131****4854	广州	458123568512354854

图 4-18 供货商信息表

(二) 主要步骤

步骤 1: 选准匹配字段。观察两个工作表，找到可匹配的字段。通过观察发现，可以通过产品代码匹配到准确的供货商名称。

步骤 2: 调整字段顺序。在使用 VLOOKUP 函数的时候，源数据中的匹配字段必须放在结果字段之前。所以，需要将产品代码字段调整到供货商名称字段之前，如图 4-19 所示。

	A	B	C	D	E	F	G	H	I
1	供货商信息表								
3	供货商代码	产品代码	供货商名称	产品名称	联系人	联系方式	城市	银行卡号	银行
4	KDN-56	GJD23-56	达××服装厂	板鞋	张成	152****4565	上海	4581231265484580	交行
5	FD-21	JSE89-02	恩××鞋业厂	运动鞋	林小然	155****5689	广州	4367125000125495	建行
6	IEJ-32	JFA58-03	索××鞋业	板鞋	范冰	183****4542	深圳	6222801254953245	建行
7	ODK-16	RDS12-12	科××鞋业有限公司	板鞋	夏雨雨	189****7851	珠海	4270202156354521	农行
8	EDS-25	IEN56-13	开××鞋业有限公司	休闲鞋	李心洁	178****4544	深圳	4270305612548626	工行
9	GPQ-12	IOE12-45	金××鞋服有限公司	板鞋	赵大鹏	187****4578	厦门	6222801452354521	建行
10	EIJ-10	BIE45-93	拓××时尚鞋业有限公司	休闲鞋	何洁新	153****4581	上海	404009123220212122	工行
11	FDD-26	BSW45-23	佳××鞋业有限公司	板鞋	李小红	183****4575	杭州	552559451236548921	农行
12	IEJ-12	XEW16-45	乐××服装有限公司	板鞋	吴军	181****7454	广州	427020152345678521	工行
13	ICE-55	SIE20-45	美××鞋服有限公司	运动鞋	孙红	131****4854	广州	458123568512354854	交行

图 4-19 调整字段顺序

步骤 3: 设置 VLOOKUP 函数。在店铺商品信息表 D3 单元格中输入“=VLOOKUP(B3, 供货商!\$B\$3:\$C\$13, 2, FALSE)”，如图 4-20 所示。

注意：B3 代表我们要查找的值。因为本次数据匹配是从一个数据表中查找另一个数据表的数据，所以在数据区域前要加上数据表的名称；为了保证数据区域在后续操作中不变，所以要加上绝对引用符号“\$”。

	A	B	C	D	E	F	G	H	I
1	店铺商品信息								
2	商品代码	商品名称	供货商	时间	单价	数量	合计金额		
3	BIE45-93	=VLOOKUP(B3, 供货商!\$B\$3:\$C\$13, 2, FALSE)			¥ 199.00	3			
4	BIE45-93	休闲鞋		2024.3	¥ 199.00	3			
5	BIE45-93	休闲鞋		2024.3	¥ 199.00	3			
6	BIE45-93	休闲鞋		2024.3	¥ 199.00	3			
7	BSW45-23	板鞋		2024.3	¥ 129.00	4			
8	BSW45-23	板鞋		2024.3	¥ 129.00	4			
9	BSW45-23	板鞋		2024.3	¥ 129.00	4			
10	BSW45-23	板鞋		2024.3	¥ 129.00	4			
11	GJD23-56	板鞋		2024.1	¥ 99.00	8			
12	IFN56-13	休闲鞋		2024.2	¥ 299.00	3			

图 4-20 编辑 VLOOKUP 函数

步骤 4: 复制字段匹配。在店铺商品信息表 D3 单元格匹配到正确的供货商信息之后，需要将这一字段匹配应用到所有供货商单元格。选中店铺商品信息表 D3 单元格，用鼠标双击 D3 单元格右下角的小点，即完成任务，如图 4-21 所示。

	A	B	C	D	E	F	G	H
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

图 4-21 复制匹配好的字段



课后练习

从素材空间获取小周所在企业 2019 年中大促的销售数据表(某企业销售数据表), 打开后的效果如图 4-22 所示。

	A	B	C	D	E	F	G	H	I	J	K
1	商品ID	商品分类	商品名称	浏览量	访客数	成交客户数	支付件数	商品单价	优惠金额	销售额	销售成本
2	20190030	其他	偏远地区邮费补拍	5571	4058	3663	20778	1.00	0.00	20,778.00	20,778.00
3	20190031	四轮推车	高景观婴儿推车可坐可躺四轮宝宝伞车	17812	14950	4252	4933	538.00	240,650.00	2,407,304.00	947,802.00
4	20190032	四轮推车	轻便折叠橡胶大轮婴儿手推车	15575	13529	5030	5664	208.00	56,640.00	1,121,472.00	431,130.00
5	20190033	四轮推车	超轻便单手折叠避震儿童手推车	9017	6877	1808	2020	336.80	40,400.00	639,936.00	329,926.00
6	20190034	四轮推车	简易小巧单手收实儿童婴儿手推车	10334	8259	2173	2381	296.00	23,810.00	680,966.00	295,910.00
7	20190035	四轮推车	便携式简易折叠大口袋双向儿童推车	15920	12960	5511	6312	299.00	63,120.00	1,824,168.00	644,490.00
8	20190036	三轮推车	儿童三轮脚踏车1-2-3-4岁轻便溜娃神器	14309	12877	5243	5284	258.00	52,840.00	1,310,432.00	507,930.00
9	20190037	四轮推车	豪华避震新生儿童手推车	15043	13379	3322	3974	468.00	79,480.00	1,780,352.00	922,634.00
10	20190038	三轮推车	溜娃神器手推车轻便折叠简易	18794	15979	7481	8102	136.90	40,510.00	1,068,653.80	519,194.00
11	20190039	三轮推车	轻便折叠可上飞机儿童婴儿手推车	17373	14459	4468	4713	148.00	23,565.00	673,959.00	335,289.00

图 4-22 某企业销售数据表

现需要对本次活动中各个商品的实际销售情况和盈利情况进行数据整理, 因为用于偏远地区邮费补拍的商品不计入, 请将其作为无价值数据清洗掉。



扫码做题