

# Power Query

具体的可以去关注我的小红书账号：小辛（63885557733），是我从2025年4月开始的学习笔记，虽然断断续续，但是都是干货^^

## 【Power Query 新手入门 | 添加列全攻略】

⭐ 零基础也能快速上手！

- ◆ 一、自定义列 = 灵活公式操作

✓ 路径：添加列 → 自定义列

💡 比如我们有1月、2月、3月三列费用，要算“总费用”，就可以写公式：

[1月]+[2月]+[3月]

生成新列“总费用”，轻松完成。

📌 Tips：用完别忘了点“关闭并上载”，可以直接生成新的工作表，后续原始数据更新后，只要刷新就能同步！

- ◆ 二、条件列 = 类似IF判断

✓ 路径：添加列 → 条件列

💡 比如根据“总费用”判断费用级别，可以设置规则：

如果总费用  $\geq 100$ 万，则为“第一阶”，总费用 $>10$ 万为第二阶，否则为“第三阶”

📌 很像 Excel 的 IF 函数，但操作更清晰直观。

- ◆ 三、索引列 / 复制列 = 管理行号&备份字段

✓ 索引列：添加列 → 索引列，从0或1开始都行，一般从1开始

✓ 重复列：添加列 → 复制列，常用于备份或辅助拆分

📌 常配合后续“合并”、“提取”等操作使用，逻辑更清楚。

- ◆ 四、拆分列 = 文本细拆神器

✓ 路径：转换 → 拆分列

支持：

按分隔符拆分（逗号/空格/下划线都OK）

按字符数拆分（如每2个字切1段）

按自定义位置拆（适用于固定结构文本）

## 案例：

原文： “今天是个好日子”

你想要变成三段 → “今天” 、 "是个"、 “好日子”

就可以按字符索引来拆 (0,2,4)

### ◆ 五、提取列 = 文本处理全能工具

 路径：转换 → 提取

支持提取：

前n个字符

后n个字符

指定范围的字符 (从第几位开始，取几个)

某个分隔符前、后的内容

分隔符之间的文本

## 案例： “生产成本”

想提取 “产成” ， 就可以设置范围：索引号=1，字符=2

## Power Query 数据清洗操作入门②：格式统一

### 为什么学 Power Query？

处理 Excel 数据，格式不统一是不是经常让你头疼？比如：

数值被当成文本，无法计算

一堆没用的列，眼花缭乱

列名太乱，找不到想要的字段

某些值太多种写法，难以统计？

 别担心！ Power Query 这几个功能，帮你一步到位清洗格式：

 常用基础操作合集（建议按这个顺序处理）

### 一、修改列的数据类型

 功能入口：选中列 → 工具栏选择 “数据类型” （文本、十进制、日期等）

 作用：保证后续计算/筛选能正确执行，不会报错！

### 二、删除无用列

 快捷方式：

选中不需要的列 → 右键 【删除列】

或先选中需要的列 → 右键 【删除其他列】

💡 建议：提前规划好分析目标，避免处理冗余数据。

### ✓ 三、重命名列

📌 方法一：双击列名直接修改

📌 方法二：选中列 → 右键 【重命名】

💡 建议：统一列名规范，有助于后续建模或数据透视。

### ✓ 四、替换值

📌 功能入口：选中列 → 【开始】选项卡 → 【替换值】

💡 场景举例：

把“男、男性、M”统一替换为“男”

把空白或“N/A”统一为“缺失”

📌 结语：

Power Query 就像给 Excel 装了个“自动清洗神器”！掌握这些小步骤，数据干净清晰，效率翻倍！

## Power Query 数据清洗操作入门③：拆分列

在清洗数据时，“拆分列”是最常见也最实用的操作之一。比如从“2024/03/31-001”这样的凭证号中拆出日期和编号，或者从“张三-销售部-主管”中提取姓名、部门、职位。

今天这篇，我们就来系统掌握拆分列的三种方式、常见场景、以及“一次拆分 vs 多次拆分”的使用技巧！

### 💡 拆分列的三种方式及适用场景

Power Query 中，点击“开始”→“拆分列”，你会看到多种方式，常用的有：

按分隔符拆分（最常用）

分隔符可以是：逗号、空格、斜杠 (/)、破折号 (-) 等

✓ 适合：从类似“销售部-张三-主管”中拆出各字段

按数字与非数字拆分

✓ 适合：如“产品A123”拆成“产品A”和“123”

按大写字母变化拆分（高级用法）

✓ 适合：如“FirstNameLastName”变成“First Name Last Name”

⌚ 一次拆分 vs 多次拆分，差别在哪？

以“2024/03/31-001”为例：

## 一次拆分（仅限第一个分隔符）

→ 会拆成：“2024” 和 “03/31-001”

- 适合字段结构简单、只需分成两段

## 重复拆分（每个分隔符都拆）

→ 会拆成：“2024” “03” “31-001”

→ 再拆“-”，得到完整四段

 实战建议：如果数据中包含多个嵌套分隔符，建议多次拆分+逐列检查，而不是一次性全拆，避免字段错位或混乱。

### 示例操作

以“2024/03/31-001”为例：

第一步：按“/”分隔

→ 拆得：2024 03 31-001

第二步：对“31-001”列，再按“-”分隔

→ 最终：2024 03 31 001

 技巧1：每一步拆分后建议重命名字段，保持数据整洁

 技巧2：如有空值或异常行，建议拆分后立刻筛查空值

 拆分后的小细节：

列顺序调整：拖动列头排序，便于后续读取或建模

列名规范：统一命名如“年”“月”“日”“编号”，提升可读性

空值检查：拆分可能导致空列，建议使用“空值筛选”功能处理

### 小结

实战建议：

拆分方式选择：多数用“按分隔符”，但要注意场景

拆分次数：可分步进行，多拆几次更清晰

拆分后处理：重命名字段 + 空值处理 + 顺序调整

## Power Query 数据清洗操作入门4：合并查询

在 Power Query 中，合并查询是连接两个表的强大工具，尤其适合处理类似 VLOOKUP、INDEX+MATCH 的操作逻辑，是数据整合、查找和补充信息的常用方式。

### 一、什么是合并查询？

合并查询是 Power Query 中用于根据一个或多个列，将两个表的数据合并到一个表中的操作，支持多种连接方式，如左外连接、内连接等。

可以类比 Excel 中的查找函数，但更灵活、更稳定。

## 二、操作步骤详解

以“库存表”合并“物料信息表”为例：

1. 选中一个主查询（如库存表），点击「开始」选项卡下的「合并查询」
2. 选择要合并的表（如物料信息表）
3. 指定关键列（如“物料编码”），主表与次表必须都选
4. 选择合并类型
5. 点击「确定」后，会生成一个新的“表格列”
6. 点击该列右上角的「双箭头图标」展开字段，勾选你想要添加的列

### ✓ 注意事项：

关键列的数据类型必须一致！否则匹配不到结果。

可进行多字段匹配（Ctrl 多选列）。

多次合并操作建议清楚命名每次结果，避免混淆。

## 三、合并类型对比

类型类似 Excel 函数说明

1. 左外连接保留左表全部，右表匹配则带入
2. 右外连接 - 保留右表全部，左表匹配则带入
3. 内连接 - 仅保留两表匹配成功的行
4. 完全外连接 - 保留两表全部，匹配不到的显示 null
5. 反连接（仅左或仅右） - 仅保留左（或右）中未匹配的行

建议从“左外连接”入门，因其最贴近日常 VLOOKUP 需求。

## 四、常见使用场景举例

1. 物料表 + 分类表 → 给物料表补充品类信息
2. 销售表 + 客户表 → 根据客户 ID 获取客户地区
3. 订单表 + 汇率表 → 将本币金额转换为统一币种
4. 库存表 + 产品表 + 仓库表 → 连续合并多表，整合出完整库存视图
5. 异常数据匹配 → 用反连接找出未对账的发票或未匹配的记录

⚠ 注意：

与“追加查询”不同，合并强调横向扩展字段，而追加是纵向叠加记录。

## Power Query 数据清洗操作入门⑤：文本提取

在数据清洗中，提取特定格式的文本（如邮箱、数字、日期、编号）是非常常见的需求。Power Query 配合正则表达式，可以实现灵活高效的模式匹配提取，极大提升数据处理效率！

### ●一、适用场景举例

从客户评论中提取邮箱地址

从产品名称中提取产品编号

从混杂文本中提取开头字母/数字/关键字

清洗掉文本中的多余前后缀信息

### ●二、提取文本的常见方法

Power Query 提供了多种方法进行文本提取，这里推荐常用三种：

#### ✓ 1. Text.BeforeDelimiter

场景：提取某个字符前的内容

公式：Text.BeforeDelimiter([列名], "-")

#### ✓ 通过位置索引提取

公式：Text.Middle([列名], 3, 5)，注意不是索引号，是字符位置

#### ✓ 3. 利用正则表达式精准提取（需启用自定义列 + Text.Select, Text.RegexReplace 等函数）

### ●三、案例：提取身份证证前六位

公式：Text.Start([身份证号], 6)

如果身份证证中还有“身份证号：”前缀，可配合 Text.AfterDelimiter 使用：

公式：Text.Start(Text.AfterDelimiter([列名], "身份证号："), 6)

### ●四、小贴士

Power Query 本身并不直接支持 Regex 函数，但在 Power BI 中或搭配自定义函数 可间接使用

多数场景可以通过组合使用内建函数实现复杂提取效果

处理数据时一定要多观察数据模式，设计提取逻辑要根据“是否稳定”决定用哪种方式

### ← 总结

掌握基本的文本提取函数 + 正则表达式，可以覆盖 90% 以上的结构化文本清洗任务。遇到复杂场景时，优先判断数据模式是否稳定，再选择适合的提取方法组合。

## Power Query 数据清洗操作入门⑥：追加查询！

在数据清洗中，经常遇到多个结构相同的表格需要合并成一个，这时候就要用到 Power Query 的「追加查询」。与“合并查询”不同，追加是纵向合并（加行），是表格整合的另一大关键技能

### ✓ 一、什么是追加查询？

追加查询（Append Queries）是指将两个或多个表中的数据纵向整合到一起，前提是各个表的字段结构一致（字段名和顺序最好一致）

类似于 Excel 的「复制粘贴追加在表格底部」，也可以类比数据库中的 UNION 操作

### ✓ 二、追加查询操作步骤

我们以“1月销售数据表” + “2月销售数据表”为例：

在 Power Query 编辑器中点击【开始】→【追加查询】。

有两个选项：

追加为新查询（推荐）→保留原始表不变。

直接追加到现有查询→原查询会被改写。

在弹出的窗口选择要追加的表（两个或多个）。

点击【确定】即可生成一个新查询，表格内容为纵向拼接。

建议统一字段顺序和名称后再追加，避免列错位或数据丢失。

### ✓ 三、字段不一致怎么办？

遇到不同来源的数据格式略有不同时可采取：

字段统一处理：事先用「重命名列」「重新排序列」「添加缺失列」等方式让字段保持一致。

容错式追加：Power Query 默认会识别公共字段，缺失字段以 null 补齐，但这可能引发混乱，建议提前整理。

### ✓ 四、实际应用场景

每月报表合并 → 多个时间段销售/采购数据整合。

多来源数据统一处理 → 不同部门提交的相同模板数据汇总。

财务凭证追加 → 合并不同系统导出的凭证记录。

Power BI 自动刷新 → 结构相同的历史数据定期自动叠加。

拆分表+追加组合 → 先用“按列拆分”处理数据，再“追加”还原为整表。

### ✓ 五、追加 VS 合并 的区别

功能方向目的

追加查询 纵向（加行） 多表拼接为一个大表

合并查询 横向（加列） 关联补充字段、类似查找操作

口诀记忆：合并加列，追加加行

## Power Query 数据清洗操作入门⑦：空值处理

本篇适用于【财务明细】 【库存表】 【客户信息表】等场景

### ✓ 一、什么是空值 (null)

在 Power Query 中，“空值”是以 null 表示的，代表“没有任何值”，而不是“0”或空字符串“”

常见产生空值的原因包括：

- 1、数据源本身缺值
- 2、计算中出现错误
- 3、手动删除了某个单元格数据
- 4、拆分列或合并查询后产生空格

### ✓ 二、如何检查表格中是否有空值

#### ● 方法一：筛选查看

选择某一列 → 点击下拉箭头

如果看到 (null)，表示该列存在空值

#### ● 方法二：条件列判断

可通过添加条件列来标记空值，判断逻辑为：

`if [列名] = null then "空值" else "有值"`

### ✓ 三、常用空值处理方法

#### ● 方法一：替换空值为指定内容

适用场景：财务表中缺失科目名称时统一补“其他”、客户表中缺失手机号时补“未知”等

##### 📍 操作步骤：

选中目标列 → 右键 → 替换值

将“null”替换为“你需要的值”

📍 提示：文本列常替换为“未知”或“未填”；数字列可以替换为 0 或其他默认值

#### ● 方法二：删除空值行（慎用）

适用场景：清洗掉无效记录（如整行都空，或关键列缺失）

##### 📍 操作步骤：

选择关键列 → 筛选掉 (null) 值即可

多列判断可以组合条件列来设置过滤逻辑

📌 小心误删！特别是“备注列”、“电话列”可为空时，不应直接删除

### ● 方法三：使用填充功能（上填/下填）

适用场景：合并单元格被展开后造成的数据空洞，如“月份”只写第一行、其他行空白

📍 操作步骤：

选择对应列 → 开始选项卡 → 填充 → 向下填充 或 向上填充

📌 示例：

一月 1000

null 800

使用“向上填充”后变为：

一月 1000

一月 800

### ✓ 四、进阶技巧：用条件列处理空值

可以通过“添加列 → 条件列”实现更灵活的替换或标记逻辑

📍 举例：

将空手机号标记为“待补充”，否则保留原手机号

`if [手机号] = null then "待补充" else [手机号]`

### ✓ 五、实战小结与建议

替换为空默认值：文本列：未知 / 数字列：0

删除空值：只在“关键信息缺失”情况下使用

填充空值：合并单元格导致的上下空白

条件列判断空值：提前标记，便于统一后续操作

Power Query 数据清洗操作入门⑧：清洗进阶

### ✓ 一、数据清洗功能

批量清除空格，修改字母大小写，提取数字或文本，增加/去除前后缀等

### ✓ 二、常用的数据类型转换方法

#### ◆ 1. 设置列的数据类型

方法一：在“开始”选项卡 > “数据类型”中选择，如：文本、十进制数、整数、日期等

方法二：点击列名左上角的小图标，快速修改类型

## ◆ 2. 文本转数字 / 日期失败的处理

如果原始数据中包含非数字或非标准日期格式，转换时可能会出错

常见解决方案：

使用“替换值”将非法字符（如“¥”“%”“-”）替换为空

使用“提取数字”功能清洗文本后再转换

对“文本日期”使用`日期.FromText()`（M语言）辅助转换

## ✓ 三、文本清洗进阶技巧

👉 功能入口：转换 > 格式 > 修剪 / 清除/大写/小写/首字母大写/添加前后缀

### ◆ 1. 去除空格与不可见字符

修剪：去除首尾空格

清除：去除所有空格

清除格式字符：清除如`\r\n`等不可见字符（通过“替换值”实现）

### ◆ 2. 改变大小写

全部大写：`Text.Upper([列名])`

全部小写：`Text.Lower([列名])`

首字母大写：`Text.Proper([列名])`

### ◆ 3. 格式化文本内容

添加前缀或后缀

示例：“发票-” & [编号]

拼接字段

示例：[省份] & “-” & [城市]

字符串截取

使用“提取 > 前几个字符 / 后几个字符 / 范围提取”实现灵活获取

## ✓ 四、实战案例分享：整理一列杂乱的价格数据

假设某行为“价格”，内容如下：

¥1,000

1,200 元

-

不详

👉 清洗步骤：

替换“¥”“元”“-”“不详”等为`null`或空字符

使用“提取数字”或Text.Select([列], {"0".."9"})仅保留数字

设置数据类型为“十进制数”

这样我们就把“价格”列变成了干净、可用于计算的数值列！

## ✓ 五、小结

- ◆ 工具功能应用场景
- ◆ 数据类型设置 确保后续分析不会报错
- ◆ 空格清除 杜绝匹配错误与脏数据
- ◆ 大小写转换 格式统一、便于展示
- ◆ 替换字符 清洗非法值、统一数据结构
- ◆ 自定义拼接 生成便于识别的字段

# Power Query 数据清洗操作入门9: 数据去重

 适用场景：清理发票记录、销售明细、报名名单、库存盘点等重复数据问题。

## ✓ 一、基础概念：删除重复≠去除全部重复值

在 Power Query 中，去重操作是“行级别”的处理，你可以选择：

- 1、某一列去重：保留唯一值，常用于名单类数据；
- 2、多列联合去重：确保每个组合唯一，常用于销售数据明细；
- 3、保留重复项：反向操作，提取重复记录用于进一步核查。

## ✓ 二、实际操作讲解

### 1. 删除某列重复值（保留唯一）

示例：只保留每位客户唯一的一条记录

操作步骤：

选中目标列（如“客户名称”）；

菜单栏选择【主页】>【删除行】>【删除重复项】；

此时，系统将保留每个客户名称的第一条记录。

### 2. 多列联合去重

示例：客户 + 产品组合的唯一订单

操作步骤：

同时选中“客户名称”和“产品名称”两列；

执行【删除重复项】，仅保留每组客户+产品组合的首条记录。

## 补充说明：

多列去重是“行的去重”，行中所有选定列内容均相同的才会被认定为重复行。

### 3. 保留重复项（反向操作）

菜单栏选择【主页】>【保留行】>【保留重复项】；

示例：查出哪些员工多次打卡或发票重复

操作步骤：

选中关键列（如“员工编号”）；

菜单栏点击【保持重复项】；

Power Query 会保留所有重复的行（至少出现两次），唯一行会被排除。

### 三、注意事项

Power Query 的删除重复项默认保留第一条记录；

如果你希望保留最后一条记录，需要使用索引列辅助排序；

去重操作不会修改原始数据结构，仅处理符合条件的行。

### 四、实战建议与技巧

业务场景建议用法补充技巧：

1、客户名单去重 对“客户名称”列去重 可添加索引列辅助判断首/末次出现

2、销售明细数据去重 对“客户+产品”列组合去重 常结合“销售日期”排序后再操作

3、异常重复检查 保留重复项 可视为“异常记录提取”步骤

### 小结

删除重复值是数据清洗中最常见的需求之一，但去重粒度、顺序、是否保留重复项都可能影响分析结果。在 Power Query 中合理设置去重策略，能有效提高数据的准确性和可靠性。

## Power Query 数据清洗操作入门⑩：分组与聚合

在日常数据处理中，我们经常需要按某列分类统计，比如各地区销售额、每月订单数等。这些都可以通过 Power Query 的“分组与聚合”功能轻松实现。

### 一、什么是“分组”与“聚合”？

分组（Group By）：将数据按一列或多列分类归组

聚合（Aggregate）：对每组数据执行统计计算，如计数、求和、平均值等

举个例子：对“客户地区”进行分组，并统计每个地区的订单总数，就是典型的分组+聚合。

### 二、基础分组操作：单列分组统计

## 操作步骤：

选择目标字段（如“地区”）

在功能区选择「主页」→「分组依据」

设置新列名称（如“订单数”）、操作类型（如“计数行”）

## 适用场景：

各门店订单数统计

每个班级学生人数

按部门统计员工数量

## 三、自定义聚合字段：不只“计数”这么简单！

Power Query 支持多种聚合方式，常用包括：

聚合操作说明

总和（Sum） 统计每组数值总和

平均值（Average） 求每组的平均值

最小值（Min） 取每组的最小数值

最大值（Max） 取每组的最大数值

所有行（All Rows） 保留每组原始数据，可进行后续操作

## 举例应用：

每月销售额总和（按月份分组 + 聚合“销售额”）

每客户平均订单金额（按客户分组 + 平均“订单金额”）

## 四、多列分组：更精细的分组逻辑

## 操作步骤：

按 Ctrl 选中多个列

执行「分组依据」

选择相应聚合方式

## 应用场景：

每年每省份的销售总额

每部门每职位的平均薪资

同一商品在不同门店的库存总量

## 五、保留所有行：后续还能“继续清洗”

选择聚合方式为「所有行」后，会生成一个“表格列”，其中保留了每组的原始数据。

## 实用技巧：

可用于组内排序、进一步筛选

支持展开字段，继续进行列级计算

🧠 举个例子：

按“部门”保留所有员工信息 → 在每组内再筛选工资最高的员工 → 找出“各部门之星”

## ✓ 六、实操演练建议

建议你自己动手试试这些：

建一个订单表，字段含：客户名、地区、订单金额、订单日期

尝试分组统计：

每地区订单总额

每客户平均订单金额

每月订单笔数

用“所有行”功能提取组内最大值

## Power Query 数据清洗操作入门⑪：数据建模

Power Query 不只是处理单一表格数据的工具，它在多表对接与数据建模方面同样强大。尤其是当你需要将来自多个来源（如销售表、客户表、产品表）整合时，规范化字段并建立正确的“关系”就显得非常关键

### ✓ 一、多表建模的常见场景

在实际工作中，我们经常会遇到以下多表结构：

销售明细表：订单号、客户ID、产品ID、销售金额，作为主表记录销售行为

客户信息表：客户ID、客户名称、区域，作为辅助表提供客户维度信息

产品信息表：产品ID、产品名称、品类，作为辅助表提供产品维度信息

👉 目标：把这些表格按照共同字段（如客户ID、产品ID）合并，从而构建“宽表”或用于建立 Power BI 的星型模型。

### ✓ 二、合并查询建立维度关联

在 Power Query 中，合并查询是实现多表对接的关键步骤

#### ✓ 操作步骤：

依次导入多个表格

选择“合并查询”

选择主表与辅助表的关联字段（如客户ID）

选择匹配方式

展开需要添加的列

举个例子：

销售表：包含 “客户ID”

客户表：包含 “客户ID” 和 “客户名称”

合并后，销售表就能“带上”客户名称字段

### 三、字段命名与字段规范

建议在合并前进行以下字段整理：

- 1、重命名字段，避免重复列名
- 2、统一数据类型 比如 “客户ID” 在两个表中都应是文本或数值，避免类型不一致导致匹配失败
- 3、保持字段命名一致性 例如所有主键字段都使用 “ID” 结尾，便于识别与后续模型构建

### 四、什么时候使用合并 vs 追加？

功能适用场景举例：

合并查询：多表横向关联，字段增加，例如销售表中添加客户信息

追加查询：多表纵向拼接，行数增加，例如月度销售数据追加成全年销售数据

### 五、实用技巧与注意事项

使用 “新建查询” 保留原始数据，避免合并操作影响原表

在展开列时，尽量只勾选所需字段，减少冗余字段

合并后检查是否有 null 值，排查未成功匹配的记录

### 总结

Power Query 的合并查询功能是连接多表、构建宽表模型的关键工具。通过规范字段、正确选择连接方式，你可以轻松完成数据建模，为后续可视化与分析打好基础

Power Query 数据清洗操作入门 ⑫：典型业务

本篇复盘式练习——用 Power Query 清洗三大典型业务数据

### 一、发票数据清洗

原始问题：

来自多个系统的发票数据字段命名不一，含有缺失值、无效字符、日期格式混乱

典型操作流程：

### 1、合并查询

将多个来源的发票数据合并为统一结构，使用合并查询和追加查询

## 2、标准化字段名与格式

重命名为统一格式（如“发票号码”“开票日期”）

## 3、修改数据类型

将日期文本转为真实日期，去除异常字符

使用 Text.Select 去除特殊字符或提取数字等

## 4、处理缺失值

替换 null 为默认值或“缺失”

删除发票号码为空的记录

👉 效果：整洁有序的发票清单、无格式错误、可用于比对/统计

## ✓ 二、销售明细清洗

原始问题：

销售数据来自手工输入，商品名称中夹带品牌/型号，字段混乱，销售员拼写不规范

🛠 典型操作流程：

### 1、拆分商品名称字段

用“空格”或“特殊符号”拆分列，提取品牌/型号

用 Text.BetweenDelimiters 提取中间信息

### 2、标准化销售人员名称

替换不规范拼写，使用替换值/自定义函数批量统一名称

### 3、新增分类字段

添加条件列：根据产品关键词打标签（如“数码类”“办公类”）

### 4、添加索引列

为后续分组准备，便于统计客户购买顺序等

👉 效果：字段结构清晰，能直接用于分类分析与销售跟踪

## ✓ 三、库存日志清洗

原始问题：

仓库记录每日入库/出库操作，需计算净变化、识别异常出入库（如出库>当前库存）

🛠 典型操作流程：

### 1、分组

按产品、日期分组，计算总入库/总出库量

### 2、新增字段计算净变动量

用添加自定义列计算：净变化 = 入库量 - 出库量

### 3、识别异常记录

添加条件列：若出库量 > 入库累计，标记为异常

### 4、去重与排序

删除重复操作记录

按产品+日期排序，确保时间顺序处理正确

👉 效果：得到可追踪的库存波动表，异常操作一目了然

## ✓ 四、清洗流程模板

以下是一个适用于多数业务表格的通用清洗步骤模板

- 1、数据源整合：合并、追加、连接
- 2、格式统一：重命名、改数据类型、删除无用列
- 3、字段拆分/合并/提取
- 4、缺失值处理与去重
- 5、逻辑字段构造：添加列、条件列、自定义列
- 6、排序、过滤、分组聚合
- 7、结果输出：加载至 Excel 或 Power BI