



csv文件导入前数据标准化流程

从 PDM 系统中导出一系列原始数据文件，例如 `x2c-v3-c.csv`，将 `process_csv.py` 放进这些文件所在的文件夹，双击运行即可。以下是脚本的功能说明。



概述

本脚本用于对指定的多个 `.csv` 文件进行统一格式化处理。主要功能包括：

- 自动识别文件编码（支持多种常见编码）
- 提取指定字段
- 按优先级重排字段顺序
- 删除“图号1(o)”为空的数据行
- 以 UTF-8 编码保存回原文件



支持处理的文件列表及字段要求

文件名	需保留字段
<code>x2c-c.csv</code>	查看字段
<code>x2c-v2-c.csv</code>	查看字段
<code>x2c-v3-c.csv</code>	查看字段
<code>x2c-x.csv</code>	查看字段
<code>x2c-v2-x.csv</code>	查看字段
<code>x2c-v3-x.csv</code>	查看字段



字段重排优先级

以下字段将被优先排列在前：

["描述", "图号1(o)", "喉深", "喉宽", "变压器"]

其余字段按原始顺序排在其后。



文件处理流程详解

1. 自动识别文件编码

尝试以下编码方式读取文件，直到成功为止：

```
ENCODINGS = ['utf-8-sig', 'utf-8', 'gbk', 'gb2312', 'latin1', 'cp1252']
```

若所有编码均失败，则抛出异常提示用户检查文件编码。

2. 读取并校验数据结构

- 若文件不存在：跳过处理。
- 若文件内容为空或表头缺失：提示警告。
- 若缺少所需字段：列出缺失字段并终止当前文件处理。

3. 数据筛选与清洗

- **删除空值行**：剔除“图号1(o)”为空的数据行。
- **字段重排序**：根据预设优先级重新排列字段顺序。

4. 写入更新文件

- 将处理后的数据写回**原文件路径**。
- 统一使用 `UTF-8` 编码保存。
- 原文件内容将被覆盖，请确保已备份重要数据。

处理目录

脚本默认处理**当前工作目录**下的所有目标 **CSV 文件**。如未找到任何匹配文件，程序会提示错误。

示例字段清单详情

X2C-C.csv

```
[  
    "描述", "图号1(o)", "X2C-C型分类", "行程", "电极臂端部", "加压力", "电极臂",  
    "变压器", "重量", "变压器放置方向", "法兰P.C.D", "托架方向", "托架角度",  
    "MOTOR厂家", "托架个数", "齿轮箱型式", "托架材料", "齿轮箱行程", "换枪装置",  
    "喉深", "有无平衡", "喉宽", "水路", "握杆伸出长度", "偏心", "偏心方向",  
    "偏心是否回到中心面", "导向方式"  
]
```

X2C-V2-C.csv

```
[
  "描述", "图号1(o)", "X2C-V2-C型分类", "行程", "电极臂端部", "加压力", "电极臂",
  "变压器", "重量", "变压器放置方向", "法兰P.C.D", "托架方向", "托架角度",
  "MOTOR厂家", "托架个数", "齿轮箱型式", "托架材料", "齿轮箱行程", "换枪装置",
  "喉深", "有无平衡", "喉宽", "水路", "握杆伸出长度", "偏心", "偏心方向",
  "偏心是否回到中心面", "导向方式"
]
```

X2C-V3-C.csv

```
[
  "描述", "图号1(o)", "分类", "行程", "电极臂端部", "加压力", "变压器",
  "重量", "变压器放置方向", "法兰盘P.C.D", "托架方向", "托架角度",
  "MOTOR厂家", "齿轮箱型式", "托架材料", "齿轮箱行程", "换枪装置",
  "喉深", "有无平衡", "喉宽", "水路", "偏心距离", "偏心方向"
]
```

X2C-X.CSV

```
[
  "描述", "图号1(o)", "X2C-X型分类", "静臂偏心", "静电极臂端部", "动臂偏心",
  "动电极臂端部", "行程", "加压力", "电极臂", "变压器", "重量", "变压器放置方向",
  "法兰P.C.D", "托架方向", "托架角度", "MOTOR厂家", "托架个数", "齿轮箱型式",
  "托架材料", "齿轮箱行程", "换枪装置", "喉深", "有无平衡", "喉宽",
  "支轴到驱动中心距离", "水路", "静电极臂前部长", "静电极臂前部高",
  "动电极臂前部长", "动电极臂前部高", "偏心是否回到中心面"
]
```

X2C-V2-X.CSV

```
[
  "描述", "图号1(o)", "X2C-V2-X型分类", "静臂偏心", "静电极臂端部", "动臂偏心",
  "动电极臂端部", "行程", "加压力", "电极臂", "变压器", "重量", "变压器放置方向",
  "法兰P.C.D", "托架方向", "托架角度", "MOTOR厂家", "托架个数", "齿轮箱型式",
  "托架材料", "齿轮箱行程", "换枪装置", "喉深", "有无平衡", "喉宽",
  "支轴到驱动中心距离", "水路", "静电极臂前部长", "静电极臂前部高",
  "动电极臂前部长", "动电极臂前部高", "偏心是否回到中心面"
]
```

X2C-V3-X.CSV

```
[  
    "描述", "图号1(o)", "分类", "静电极臂端部", "动电极臂端部", "行程", "加压力",  
    "变压器", "重量", "变压器放置方向", "法兰盘P.C.D", "托架方向", "托架角度",  
    "MOTOR厂家", "齿轮箱型式", "托架材料", "齿轮箱行程", "换枪装置", "喉深",  
    "有无平衡", "喉宽", "水路"  
]
```



运行环境依赖

- Python >= 3.7
- 标准库模块：
 - os
 - csv
 - sys
 - time

无需安装第三方库，开箱即用。




使用建议

- 推荐在测试目录中先行验证脚本行为。
- 若需保留原始数据，请手动备份文件后再运行脚本。
- 可配合批处理脚本实现多目录自动化处理。



输出日志格式

每一步操作都会输出带时间戳的日志信息，便于追踪进度和排查问题：

```
[17:30:00] 正在处理文件：X2C-C.csv  
[17:30:01]  文件 X2C-C.csv 已成功更新...
```



启动方式

直接运行脚本即可：

```
python process_csv.py
```

执行完成后会提示：

```
按回车键退出...
```


防止窗口自动关闭。



总结

该脚本是一个高度定制化的 CSV 数据清理工具，专为焊接设备图纸数据整理设计。其优势在于：

- 智能编码识别
 - 强大的字段控制能力
 - 安全的数据过滤机制
 - 低依赖、易部署
-

 如需进一步了解该脚本执行前后需遵循的步骤，请联系：

 gaoj@obara.com.cn