

轉 Excel 表頭合併儲存格

PRNASK prn2xlsx:表頭(標題/日期/頁數)被資料欄硬切散落 → 合併為單一儲存格

日期:2026-07-15 模組:PRNASK (prn2xlsx.py)

一、訴求與現象

PRNASK 把 PRN 報表轉成 .xlsx 時,是用資料區的 ===== 分隔線推出欄位邊界,再依邊界切欄。但表頭列(標題、日期、頁數)的文字位置不對齊資料欄,於是被硬切散落到多個儲存格。以「庫位別盤點表」為例:標題「§庫位別盤點表§」被切到 B1、C1;「115年07月12日」被切到 A3、B3;「頁數: 001」被切到 I3、J3。客戶希望這些表頭欄位合併成單一儲存格。

表頭欄位	修正前	修正後
§庫位別盤點表§	拆 B1、C1	合併 B1:C1
115年07月12日	拆 A3、B3	合併 A3:B3
頁數: 001	拆 I3、J3	合併 I3:J3

二、踩坑與陷阱

1. 表頭不能沿用資料欄邊界:資料列靠 ===== 邊界切欄是對的,但表頭文字位置不同,沿用同一組邊界一定切散。必須對表頭改用「文字本身的分段」來判斷。
2. 單一空白不可拆:「頁數: 001」中間只有一個空白,若用一般空白切會被拆成「頁數:」與「001」。故分段規則用「2格以上空白」才切,單一空白視為欄位內含。
3. 顯示寬度 vs 位元組:中文佔 2 個顯示格,換算欄位起訖與 col_break 區間要用 display width,不能用字元數或位元組,否則跨欄判斷會偏。
4. exe 優先、須重建:fw_prnout 有 exe 就用 prn2xlsx.exe(沒有才跑 python .py),所以改 .py 後必須用 32-bit Python 依 prn2xlsx.spec 重建 exe,再部署到 HBSTK/build、PRNASK/build、update_pack/EXE 三處。

三、解決方法

只改「表頭列(第一條 ===== 之前的 is_title 列)」的處理,資料列與欄頭列完全不動:

步驟	做法
分段	<code>_title_segments()</code> :用正則 <code>\S+(?: \S+)*</code> 以「2 格以上空白」切出邏輯欄位, 記錄每段的起訖 <code>display</code> 位置
對應欄號	<code>_dpos_to_col()</code> :把 <code>display</code> 位置對回 <code>col_break</code> 區間 → 1-based Excel 欄號
寫入 + 合併	每段寫進其起始欄, 若跨多欄則 <code>ws.merge_cells</code> 合併該段涵蓋的欄
樣式	主標題(第一列)放大置中, 其餘表頭列(日期/頁數)靠左

如此每個表頭邏輯欄位都成為單一(或合併)儲存格, 不再散落。實測「庫位別盤點表」得到 B1:C1、A3:B3、I3:J3 三處合併, 欄頭(倉庫/零件號碼/…/良品金額)與資料列完好。

四、使用技術與用法

技術 / 工具	用途
<code>openpyxl</code> <code>merge_cells</code>	把表頭欄位涵蓋的多欄合併成單一儲存格
正則 2 格以上空白分段	<code>\S+(?: \S+)*</code> :單一空白保留在欄位內(頁數: 001 不拆), 2 格以上才切欄
<code>display width</code> 換算	中文=2 半形=1;段落起訖與 <code>col_break</code> 區間比對都用顯示寬度
<code>pymupdf(fitz)</code> / <code>openpyxl</code> 驗證	程式化讀 <code>xlsx</code> 的 <code>merged_cells</code> 與各格值, 核對合併範圍與內容正確
32-bit PyInstaller 重建	E:\Python312-32 依 <code>prn2xlsx.spec</code> 重建 <code>prn2xlsx.exe</code> , 部署 HBSTK/PRNASK build + <code>update_pack/EXE</code>

變更檔案:HBSTK/build/prn2xlsx.py(+同步 PRNASK/build/prn2xlsx.py):新增 2 個 helper + `is_title` 分支;重建 32-bit `prn2xlsx.exe` 部署三處。