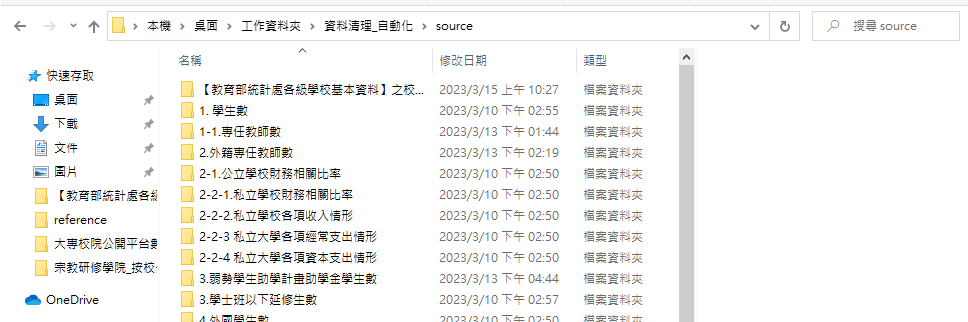
## 資料夾內含物說明 :

1. Output 資料夾 : 存放每個清理完成的檔案。
2. Ptyhon 資料夾 : 存放五個程式碼檔案，這是個分別對應資料清理的五個步驟。
3. Error 資料夾 : 存放過去資料與現在資料不同的檔案。
4. Reference 資料夾 : 存放會與程式碼互動的公式、檔案資料。
5. Source 資料夾 : 存放所有的原始資料來源。根據不同的資料，放入對應的資料夾中。



1. 大專校院公開平台數據計算\_參考版 資料夾 : 存放過去清理好的檔案。

## 程式碼使用說明 :

**步驟一、建構一個用於填寫數值的空白檔案**

以圖 2-1-1師生比，舉例。此步驟會將過去已經清理好的檔案，「大專校院公開平台數據計算\_參考版 資料夾」，其數值全部都改成 0以後輸出成一份EXCEL的其中一個分頁，之後用於與新的資料進行互動，並將新的公開資料之數據，填入這個空白的檔案後輸出，依此類推。

**步驟二、建構一個用於比較當前與過去數據差異的檔案**

程式碼的書寫基本與步驟一相同，但數值會全部保留以後輸出成一份EXCEL的其中一個分頁。

**步驟三、建構由公開資料組成的資料庫**

由於每個需要清理的檔案，都有不同的來源，不同的來源也有不同的清理方式。如果要每個檔案個別處理十分耗時，所以這裡將所有需要使用到的變數，整理出來以後，根據需要的變數，從對應的原始資料中抓取，最後彙整成一個較大的資料集。並以每個學校的學校代碼作為主鑑。

**步驟四、資料填入、檢查**

在建構出資料庫以後，透過調用在 reference 資料夾中，「填值表」的Excel檔案之內容，根據不同的圖型，載入不同的公式，與資料庫中的變數互動。最後可以在，步驟一建構的空白檔案中，如 圖2-1-1師生比，寫入需要計算的最終結果。

接著，將這一份填入好的檔案，與步驟二建構的檔案，進行比較，抓取出數值不同的圖型名稱、欄位名稱、學年度，等資訊，方便後續除錯。並將錯誤資訊放在 error 資料夾。

**步驟五、資料填入、輸出**

資料填入的部分與步驟四類似，差別在於步驟五會輸出為一張一張的圖表，而非以分頁儲存。並且每一個輸出的圖表中，第一個分頁為統計說明、第二頁為計算結果、後續每個分頁為每個學校，每一學年度使用的資料庫資料。

## 填值表使用說明



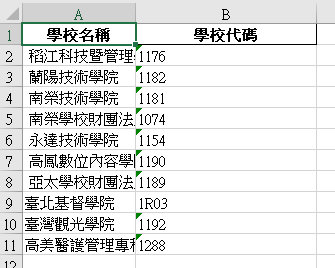
以上為節錄自填值表的部分內容，

1. 變數名稱 : 決定每個圖型最後輸出的以後，檔案內部的欄位名稱。
2. 能否直接計算 : 決定該圖型的每一個學年度，是否計算方式不同。如 速動比率，101-107年與107年以後的計算方法不同，則不能用同一套公式去算每一個學年度的數值。
3. 能否先串接 : 決定該圖型的每一個學校，是否計算方式不同。如 教授平均每週授課時數，具有合併紀錄之學校，與沒有合併紀錄之學校，計算方法不同。不能夠用同一套公式計算。
4. 公式(pandas) : 如果每張圖型的公式有變，請修改這個欄位的內容。

格式 : data[“變數名稱”]，這個變數名稱必須與資料庫的變數名稱相同。

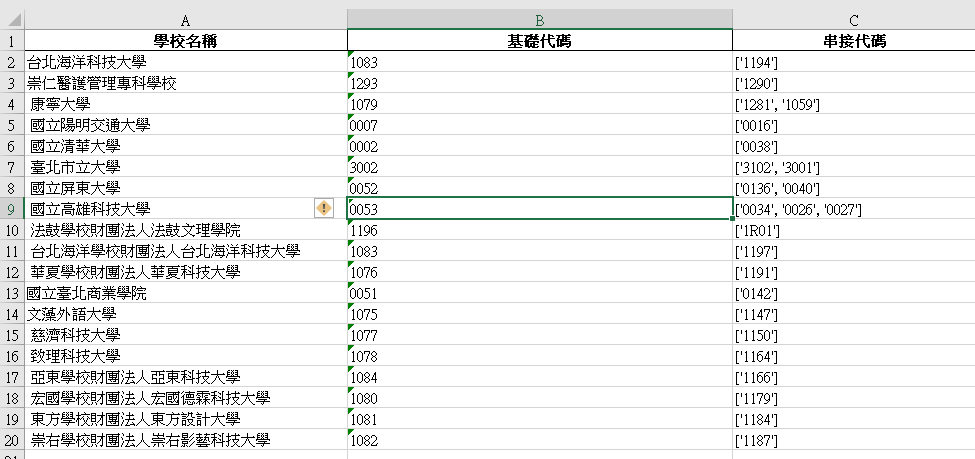
1. 上限為100，如註冊率，上限為100，則該圖型中就算有差過100%的資料，都會以100%輸出。

## 刪除表使用說明



程式會將原始資料中，此處具有學校代碼的學校資料刪除，刪除的學校可能是停辦之類的。

## 串接表使用說明



程式會將，每一筆原始資料中，會將串接代碼中的學校，根據基礎代碼做串接。如 台北海洋大學的學校代碼為，1083。如果，需要與其串接的學校之學校代碼為，1194。而1083為最後輸出的學校代碼，則將1083放在基礎代碼欄位，1194以[“1194”]的方式，放在串接代碼的欄位。

如果有複數學校串接，則以，[“0034”, “0026”, “0027”, …]，書寫。