以下內容，有時會用簡寫的方式代表一個資料或標籤，例如cam\_id常用來描述CameraID

主要功能

1. 將影像跟影片的exif轉換成各項數值並且儲存於AccessDB
2. 計算有效照片數與相機工作時數，在不同資料組合下算出工作時間與OI，以及其他的搜尋功能

情境

希望的格式

e.g. IMG\_0223.JPG

將圖檔的exif資訊輸出成AccessDB如下

Filename,CreateDate,Subject,HierarchicalSubject

IMG\_0223.JPG,2020/03/15 15:38:10,"JC38, Researcher, 1","1\_Site ID|JC38, 2\_Animal|Human|Researcher, 3\_Number|1",

其中資料會以下列所列的為對應的資料庫資料格式，以及相關的擷取標籤

---Table “file\_record”

SourceFile varchar(50) = File:FileName

DateTimeOriginal datetime = EXIF:CreateDate  
Date datetime = DateTimeOriginal 的日期

Time datetime = DateTimeOriginal 的時間

Site varchar(6) = Camera\_ID 的前半部, 限定英文字母

Plot\_ID varchar(6) = Camera\_ID 的後半部, 字母以後的所有數字(因此也能超過100, 也就是三個數字以上)

Camera\_ID varchar(6) = part from XMP:HierarchicalSubject -> 1\_Site ID

Group varchar(20) = part from XMP:HierarchicalSubject -> 2\_Animal

Species varchar(30) = part from XMP:HierarchicalSubject -> 2\_Animal

Number int = part from XMP:HierarchicalSubject -> 3\_Number

Note varchar(100) = ?

IndependentPhoto int = 參照可選的時間間隔為有效隻次

CreateDate datetime = 產生列的datetime (並非檔案開創時間)

period\_start datetime = 單一資料夾的最早的檔案的開創時間，計算工作時間用

period\_end datetime = 單一資料夾的最晚的檔案的開創時間，計算工作時間用

---

注意! 有時候會有N個以上的動物tag，這種情況要存N筆紀錄

且會提示在視窗中(包含console黑底白字介面 跟圖形化介面)

這些範例數值為(以IMG\_0223.JPG為例)

---

SourceFile = IMG\_0223.JPG

DateTimeOriginal = 2020/3/15 下午 03:38:10 (此為資料庫的DateTime格式)

Date = 2020/03/15

Time =15:38:10 (24小時制)

Site = JC

Plot\_ID = 38

CameraID = JC38 (前面請勿設為數字, 否則Site可能會找不到. 例如12JC12, Site=””, Plot\_ID=”12JC12”)

Group = Human

Species = Researcher

Number = 1 (有時候tag會出現>N, 就取N為最大出沒數. 完全沒有該tag則設為1)

Note =

IndependentPhoto = 0 ~ N (0表示為同一隻; 1以上表示為不同隻, 即有效隻次. 會因時間間隔而有所不同)

CreateDate = 2021/02/06

period\_start = (同上)

period\_end = (同上)

---

註：以下只會拿來參考而不會存到資料表中

--

Subject = 只有勾選的item XMP:Subject

HierarchicalSubject = 連同parent item也會被納入 XMP:HierarchicalSubject

# 範例值如下

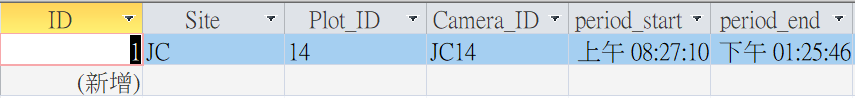
Subject = JC38,Researcher,1

HierarchicalSubject = 1\_Site ID|JC38,2\_Animal|Human|Researcher,3\_Number|1

--

camera\_task\_record Table(相機工作紀錄)

---



---

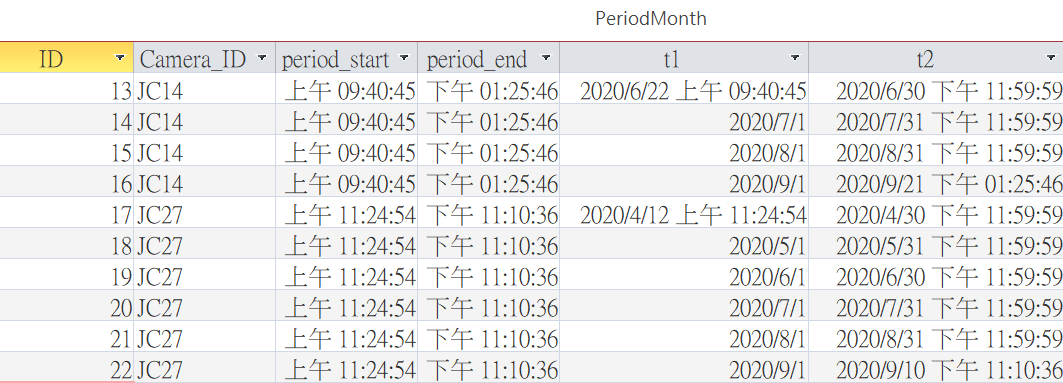
這個資料表會紀錄單一資料夾的最先與最後的檔案開創時間，也就是預先整理出一個單純的時段

主要會給一些查詢使用

PeriodMonth Table由於有的自動相機可能在該月份沒有任何紀錄，但該工作時間還是要算

不幸的是, SQL能夠算一整段時段，卻不能很直接的分成各個月份並計算每個月的時段

因此一定要先由Python來分解成各時段，以利"每月"的Query，如下



之後才由不同的SQL query算出更精確的時段

Query Table = OI值(下SQL Query計算)

---

OI = (主要為一個數值 per 自動相機/母分類/子分類, 看情境, 如後述)

---

程式流程

編輯設定檔來設定抓取條件

A.可選取資料夾, 資料夾下的子資料夾也會被處理連同圖片都會被處理

1. 不同資料夾, 有同檔名的情形，例如**100\_BTCF，101\_BTCF**  
   可用後面的JCXX來辨認為不同自動相機的檔案  
   不同樣區代碼也不同e.g. 惠蓀是C, 陽明山是YC
2. 檔名並非照著時間排序，得從時間標籤來評比
3. 檔案時間(DateTimeOriginal)有時會因為檔案搬移而消失，這時會由多種方式來搜尋其時間戳。可由csv來定義，或從圖檔本身的XMP:CreateDate獲得，其優先順序如下
   1. 讀取csv列表，主要以與資料夾同名的csv檔來當作指定檔案，沒有的話則找第一個csv檔  
      csv欄位應該要有"檔名Filename"以及"CreateDate"，且只能有這兩個欄位  
      才可拿來提供正確的檔案開創時間  
      請勿使用例如JC14\_20200622\_20200921(100RECNX).csv的csv，因為裡面的時間符號很亂  
      因此請盡量將要使用的csv的檔名改成跟資料夾的一樣，以免抓到另一個格式不符的csv  
        
      時間格式以年/月/日 時:分:秒(不一定會有秒)，一定要有分鐘以上的單位，以免影響有效照片數  
      e.g.   
      2020/6/22 09:40:5 -> 2020/06/22 09:40:05  
      2020/06/22 09:40 -> 2020/06/22 09:40:00  
      2020:06:22 09:40:12 -> 2020/06/22 09:40:12 (2021.7.20註:不太正確的特殊格式)  
      只有一個數字會強制當成個位數，沒有秒數則會強制補成0秒  
      但如果出現  
      2020/6/22 09:40:  
      這個是錯誤的時間格式，遇到這種情形，將會再看看有無其他可用的時間資訊  
      (csv檔案可以用記事本開，並將冒號逗點 :, 以逗點 , 取代，或許能夠避免這個情形)
   2. 直接以exiftool所獲得的XMP:CreateDate來當作時間戳  
      如果csv沒有包含該檔案的時間戳，也會在這裡設定時間  
      一般來說exiftool跟時間格式應該不會出錯  
      如果依舊沒有時間資訊的話，則進入下個偵測方式
   3. 如果檔案沒有csv與XMP:CreateDate的時間資訊，則以OCR(Optical Character Recognize)  
      但由於並非100%成功，因此仍舊失敗的話則到下一方式
   4. 使用上一個連續的檔名的有效DateTimeOriginal  
      如果剛好是第一個檔案，則設定為2000/1/1

由於一定要有時間才能排序，因此時間是一定要先找到並定義  
  
如果使用OCR，會出現下列警示資訊  
WARN RCNX0020.AVI has no XMP:CreateDate, use OCR to detect datetime.

WARN result: (2020-06-25 10:29:58)  
如果OCR偵測失敗，則會出現下列資訊  
WARN RCNX0108.AVI has no XMP:CreateDate, use OCR to detect datetime.

WARN 2020-07-09 07217217 not a valid datetime

WARN result: (None)

WARN Still not a valid datetime, use previous datetime for RCNX0108.AVI (2020-06-25 11:18:22)  
主要是提醒使用者，該時間有可能需要修正

B.輸出Access DB檔案(主要為.accdb)

C.Access DB有五種query可供查詢各種情形的OI值，分別為

1. 查詢每月\_自動相機ID
2. 查詢每月\_自動相機ID+母分類
3. 查詢每月\_自動相機ID+母分類+子分類
4. 查詢每月\_母分類+子分類
5. 查詢單一時段\_自動相機ID
6. 查詢單一時段\_自動相機ID+母分類
7. 查詢單一時段\_自動相機ID+母分類+子分類
8. 查詢單一時段\_母分類+子分類
9. 查詢時段\_分小時\_百分比之母分類+子分類  
   e.g.加一個查詢是 輸入時間區間, 譬如2021/1/1-2021/2/28, 然後算每個小時各物種的有效照片/該物種所有的有效照片\*100

譬如水鹿這段時間全部的有效照片是100張, 其中18:00-19:00是10張, 18:00-19:00 這段時間算出來的活動量就是10%

每個query都能輸入時間區間(e.g. 一個季可選則3個月份)，來計算工作時數  
工作時數不會超過時間區間  
注意! ms access db對於datetime的定義

假設輸入2020/1，則會是2020/1/1，且時間會是0時0分0秒

因此如果想要指定2020年1~3月，則需要輸入2020/1/1(同2020/1)跟2020/4/1(同2020/4)來指定

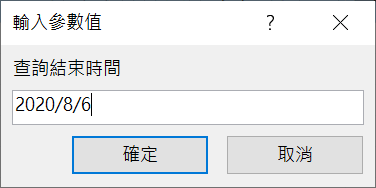
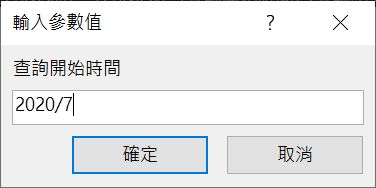
基本上所有的檔案開創時間只要介於2020/1/1到2020/3/31都會被找出來

條件式等同 2020/1/1 00:00:00 <= 開創時間 < 2020/4/1 00:00:00，其中"<="為小於等於之意

假設檔案開創時間超過這個時段，則不會被納入。假設這區間第一筆是2020/1/3，最後一筆是2020/3/3，那就只會計算1/3~3/3的工作時間，但如果搜尋的時間比最後一筆還要來得前面，那麼就會以搜尋時間為準。例如搜尋2020/1~2020/2/28，則工作時間會縮為2020/1/3~2020/2/27(請記得如果只輸入2020/2/28則只會納入該天0時0分0秒的剎那，除非連時間都設定，才有可能涵蓋2/28的大部分)

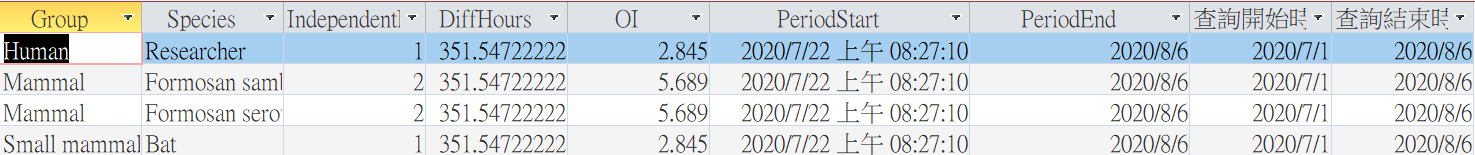
範例 "查詢時段\_母分類+子分類"

開始時需要輸入查詢的時間



注意，由於access db的特性，如果只指定日期，則會以當天的00:00:00作為基準

因此要包含2020/8/5的紀錄，通常都以2020/8/6來指定

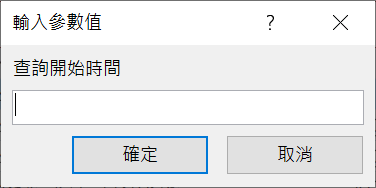


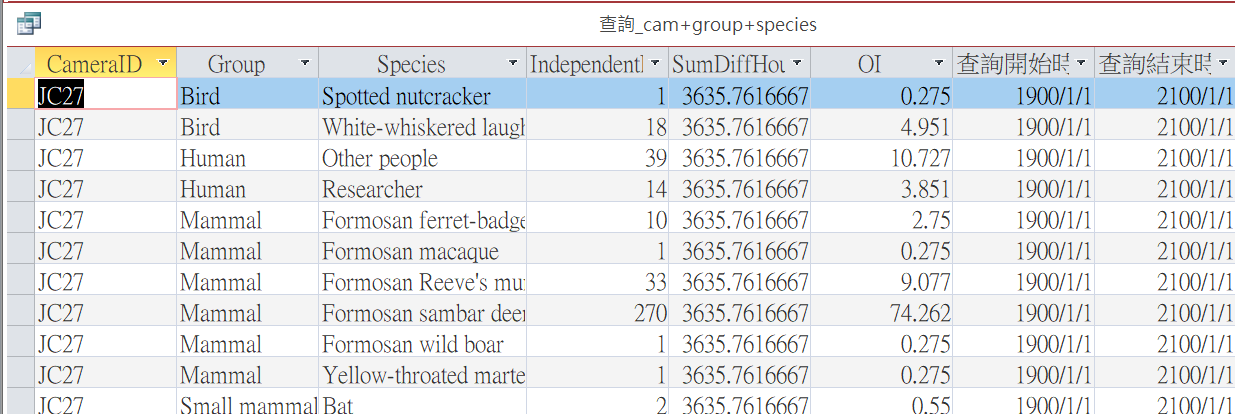
同樣的，PeriodStart跟PeriodEnd也會受限於影像開創時間跟查詢年月

查詢結束時間本身只有指定日期而沒有指定時間，則只有該天的00:00:00而已，要注意

version 2021.3.22 自動設定時段

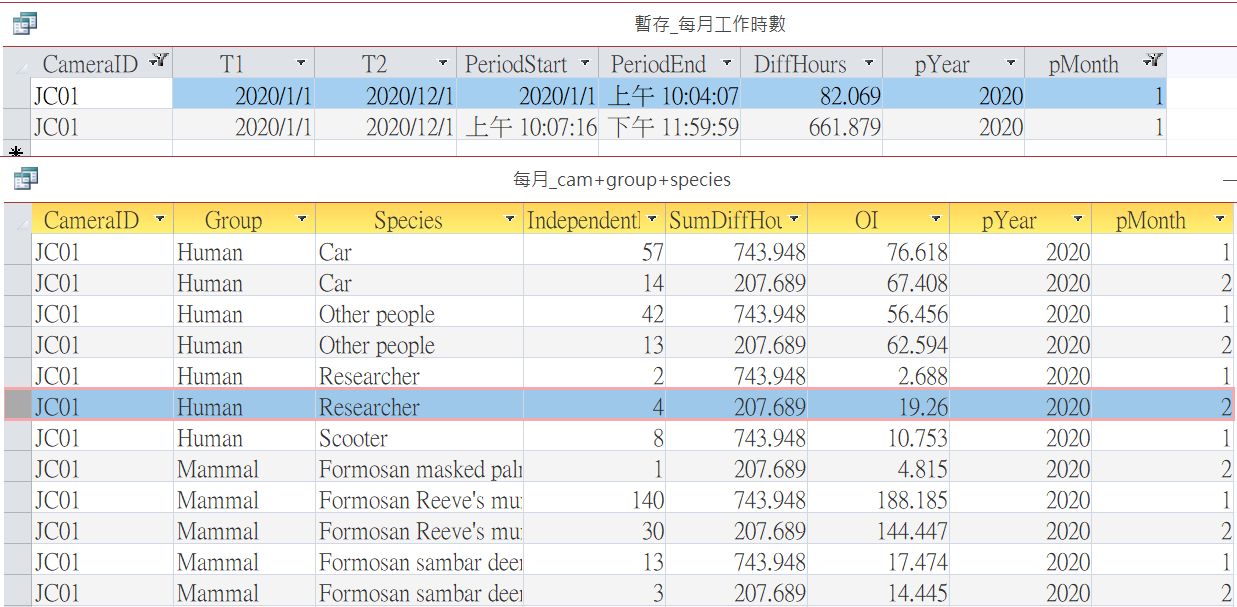
目前可以不輸入任何東西即可搜尋所有時段(預設1900~2100)，方便使用與測試

 不用輸入!



查詢開始時間跟結束時間會自動設為1900~2100年間(所以2100年以後如果有事請不要找我)

暫存table



主要是將長度過長的query分成其他部分，這裡的暫存所整理好的資料表都為其他table所用(本身不會儲存)

有些是為了方便計算，例如T1, T2是輸入的搜尋範圍，可用來比較時間並且直接刪去搜尋範圍以外的部分

PeriodStart跟PeriodEnd可以大致上看出工作時數

這樣應該會比較清楚

已知問題

1. 當adobe bridge以不同順序勾選exif值時，輸出的exif也會照該次順序
   1. 假設點選  
      1\_Site ID|JC38,3\_Numbe, 2\_Animal|Human|Researcher  
      希望能夠排列成  
      1\_Site ID|JC38, 2\_Animal|Human|Researcher, 3\_Number|1  
      而非點選順序
2. 影片的CreateDate可能跟影像的tag不一樣  
   有些影片在複製到電腦時，時間tag值會被更新，而遺失當下錄影的時間
3. 如果沒有Animal tag或Animal tag為unknown，則不應insert到資料表中
4. csv可用來參考，但csv中，CreateDatetime有時會缺少秒數，因此如果沒有秒數會自動補0，e.g.  
   09:7:5 -> 09:07:05  
   09:7 -> 09:07:00  
   09:7: -> 不符合時間格式, 會出錯  
   所以有09:7:這樣的字串，是一定會出錯的，為了預防錯誤的格式所造成的影響

註

1. OI = 有效照片數/相機工作時數\*1000  
   相機工作時數為第一張的時間到最後一張的時間，但如果選定的時間區間沒有覆蓋到，則以時間區間為開頭時間 或/和 結尾時間  
   相機工作時數不足一小時以一小時計?
2. 有效照片定義，一定時間間格內(通常是30分鐘或是1小時)，同一個物種，只算1張有效照片  
   可選的時間間隔為  
   30分鐘 / 1小時 / [其他時段]

建議

1. 如果Microsoft Access沒有開的話，程式必須開啟Access插入一行資料然後關閉，這會導致速度過慢  
   建議在使用程式的時候開啟Access  
   程式跑完了之後，按上面的"全部重新整理"按鈕，即可看到最新的列表
2. 由於AccessDB似乎沒有增加SQL query的API，或許得用手動或用舊有的accdb資料來完成query  
   以下列出可用的query，來直接計算OI值，以便參考

---

查詢\_cam+group+species

PARAMETERS [查詢開始時間] DateTime, [查詢結束時間] DateTime;

SELECT `CameraID`, `Group`, `Species`, `IndependentPhoto`,

`DiffHours`, `OI`, `PeriodStart`, `PeriodEnd`, [查詢開始時間], [查詢結束時間] FROM (

SELECT a.Camera\_ID AS `CameraID`,

a.group AS `Group`,

a.species AS `Species`,

a.period\_start - [查詢開始時間] AS T1,

a.period\_end - [查詢結束時間] AS T2,

IIf([T1] > 0, a.period\_start, [查詢開始時間]) AS `PeriodStart`,

IIf([T2] < 0, a.period\_end, [查詢結束時間]) AS `PeriodEnd`,

SUM(a.IndependentPhoto) AS `IndependentPhoto`,

DateDiff('s', PeriodStart, PeriodEnd)/3600.0 AS `DiffHours`,

Round(IndependentPhoto/DiffHours\*1000, 3) AS `OI`

FROM file\_record AS a

WHERE [查詢開始時間] <= [a].[DateTimeOriginal] AND [a].[DateTimeOriginal] <= [查詢結束時間]

GROUP BY a.Camera\_ID, a.group, a.species, a.period\_start, a.period\_end

) AS C;

---

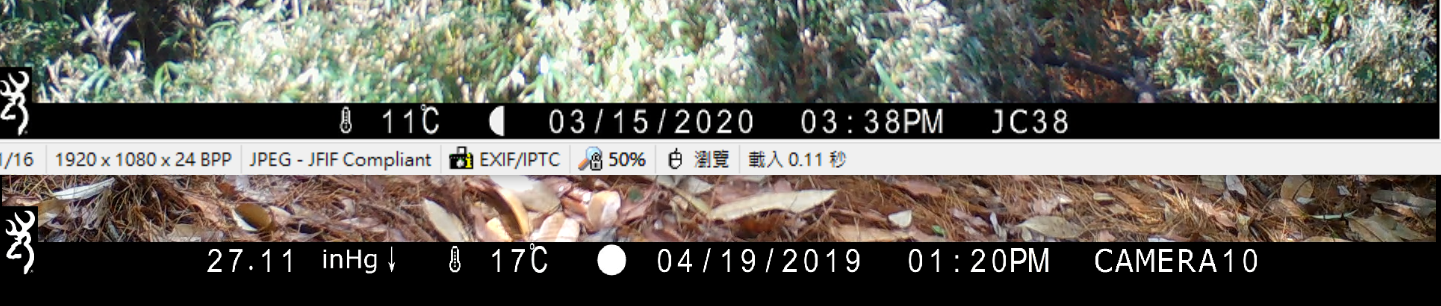
上述依序為  
攝影機ID, 母分類, 子分類, 有效照片總和, 相機工作時數(以小時計), OI值(小數點第3位後4捨5入), 開始工作時間(查詢時間以內), 結束工作時間(查詢時間以內), 輸入的查詢時間

設定條件為開始日期與結束日期的區間, 其他不會納入  
並以攝影機ID, 母分類, 子分類, 不同資料夾的時間為群組分類

Access DB輸入時間可以只輸入年跟月，例如 2020/6，但這是表示2020/6/1 00:00:00的意思

注意

1. 同品牌的自動相機的日期位置可能不同，需多加注意



位置、編排字樣、時間格式一有不對，我的regular express判斷式就可能會抓錯，因此如有新的格式，請多留意

====

Exif轉換程式使用方式

由於本程式有使用tesseract 來進行數字辨識，因此請先在電腦安裝該程式，詳細可在下列網址中

<https://digi.bib.uni-mannheim.de/tesseract/>

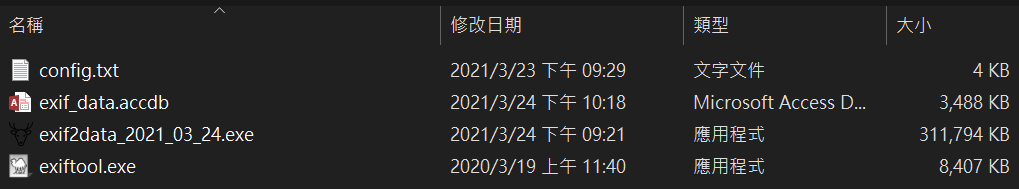
下載64位元的

tesseract-ocr-w64-setup-v5.0.0-alpha.20201127.exe

並安裝

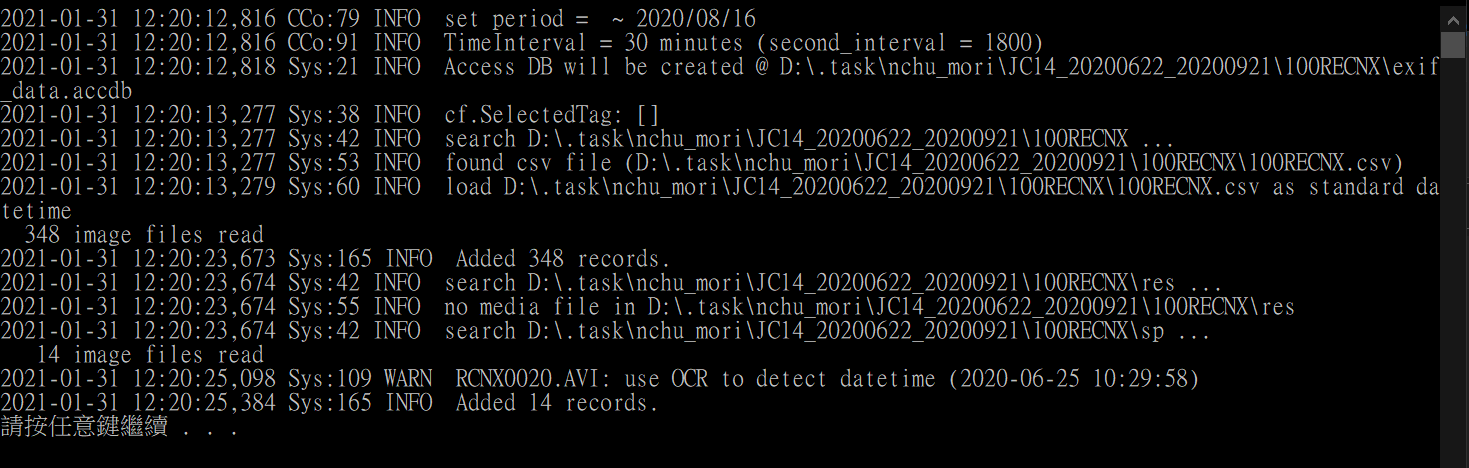
壓縮檔內會有幾個檔案

---



---

其中config.txt是設定檔

exif\_data.accdb為資料表，並含數個query來計算OI值，使用前請確認資料表有無清空(因舊資料會影響OI)

exiftool為抓exif值的外掛程式

exif2data\_年\_月\_日.exe為主程式，後面為release版本

一開始可用記事本讀取config.txt並且編輯 一些預設值

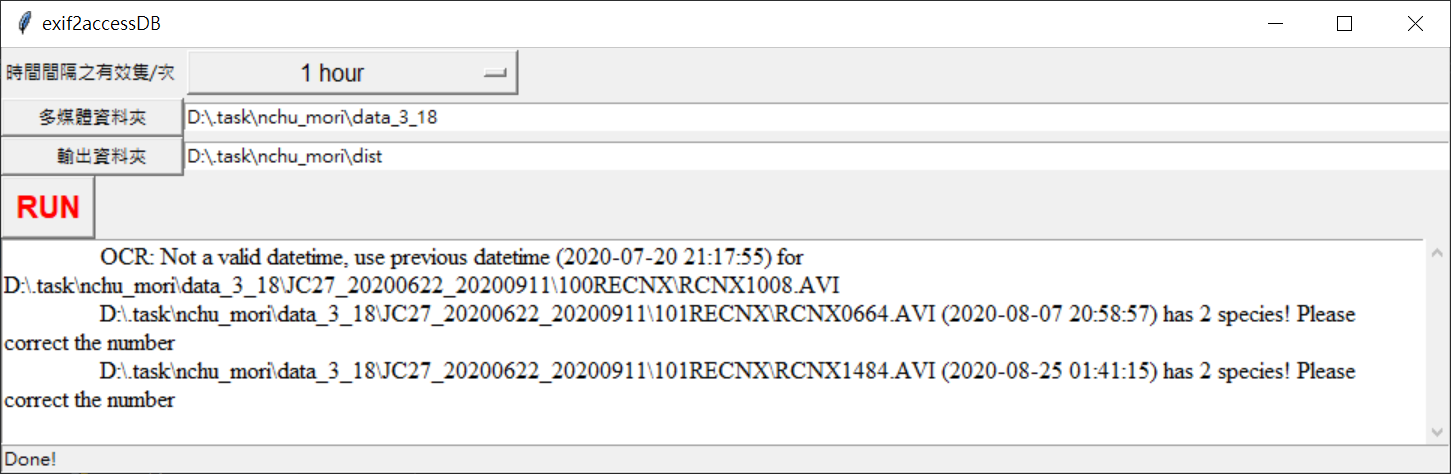
時間間隔的DefaultTimeInterval

多媒體資料夾路徑PathInput

以及輸出路徑PathOutput (建議先將上面的exif\_data.accdb放在此路徑，因為該檔案已經有OI的query)

這樣以後開主程式可以方便一點

設定完config後，即可直接執行主程式GUI(Graphical User Interface)



可以直接用選取的方式來選擇時間間隔、輸出入資料夾

一開始程式會讀取config.txt並輸入預設值

選好就能RUN了

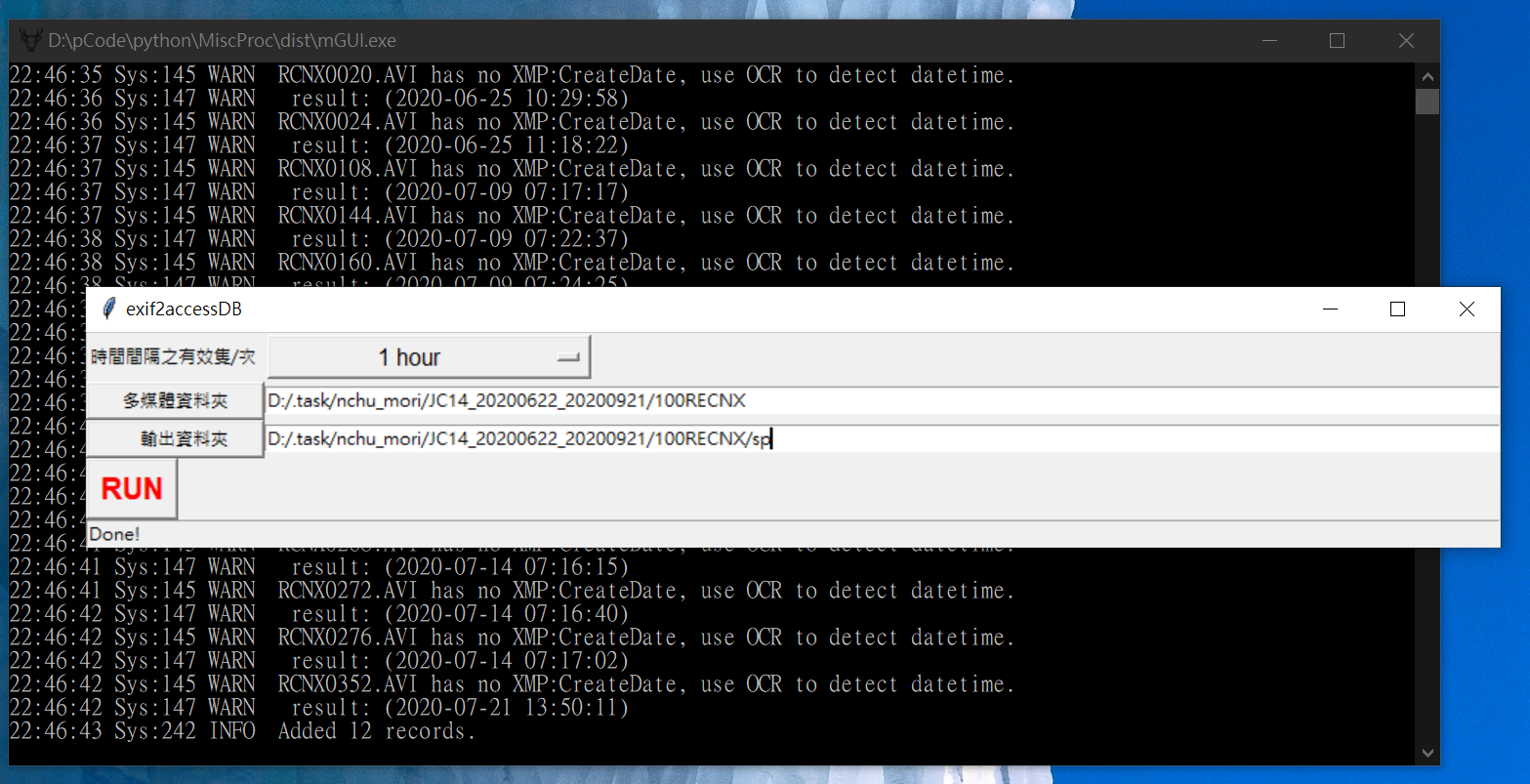
在RUN之後，重要訊息會顯示在GUI的下方，用來警示使用者，部分的紀錄可能需要對照與修改

例如OCR依舊沒法找到時間，一個檔案發現兩個動物tag等等

這個GUI比之前的單一console還要直覺，改變一些設定時會比較簡單，而不用一直開config.txt來改

不過GUI還是會隨著一個console，用以觀看訊息

如下圖的黑色畫面



具體的程式主要流程

當選定多媒體資料夾100RECNX，程式就會搜尋該資料夾的所有多媒體檔案，包含所有的子資料夾，並計算當次所有檔案的最先與最後的開創時間，用以排序以及計算有效隻次

如果個別的資料夾有100RECNX.csv或其他.csv檔，就會以該列表的Datetime當作重要參考，如果沒有該檔案，才會使用OCR等影像辨識的方式來抓取Datetime

注意csv檔名盡量跟資料夾檔名一樣，因為會先以該檔名來當作參考對象

執行程式時，可看看console有什麼異常訊息  
例如  
2021-01-31 12:04:32,871 CCo:79 INFO set period = ~ 2020/08/16

中間有個INFO就表示程式目前的執行狀態log紀錄

因此如果出現WARN(警告)或ERRO(即error錯誤)就要注意了

如果出現ERRO請複製整段ERRO的訊息並告知我

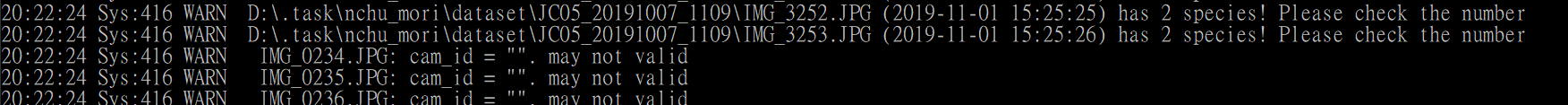
正常狀況應該會如下

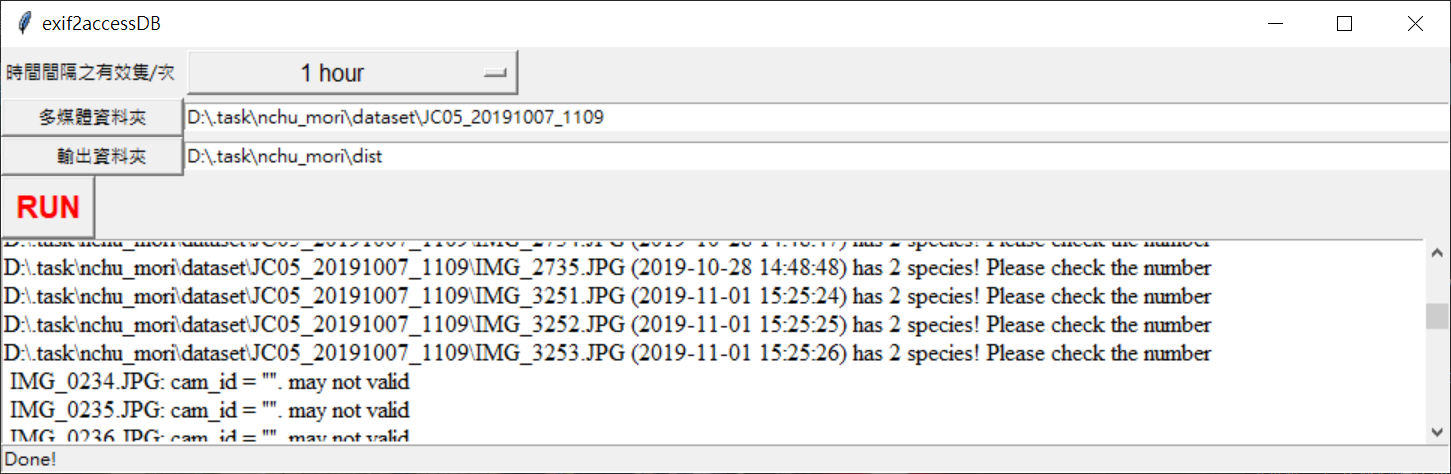
這個基本上只有INFO跟WARN(因為用了OCR)，沒有錯誤發生

這時候就能開啟PathOutput路徑中的exif\_data.accdb來看結果

如果有單一檔案標註兩個以上的動物種類的話，或是有未標上cam\_id的檔案

則會出現警示訊息(GUI也會出現)，包含檔名以及紀錄的時間，以供對照



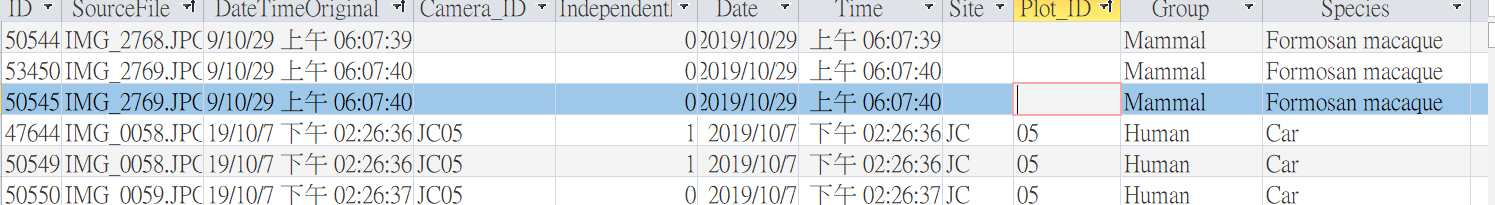


這時可到accdb查看該有效隻數的情形，並適度作修改

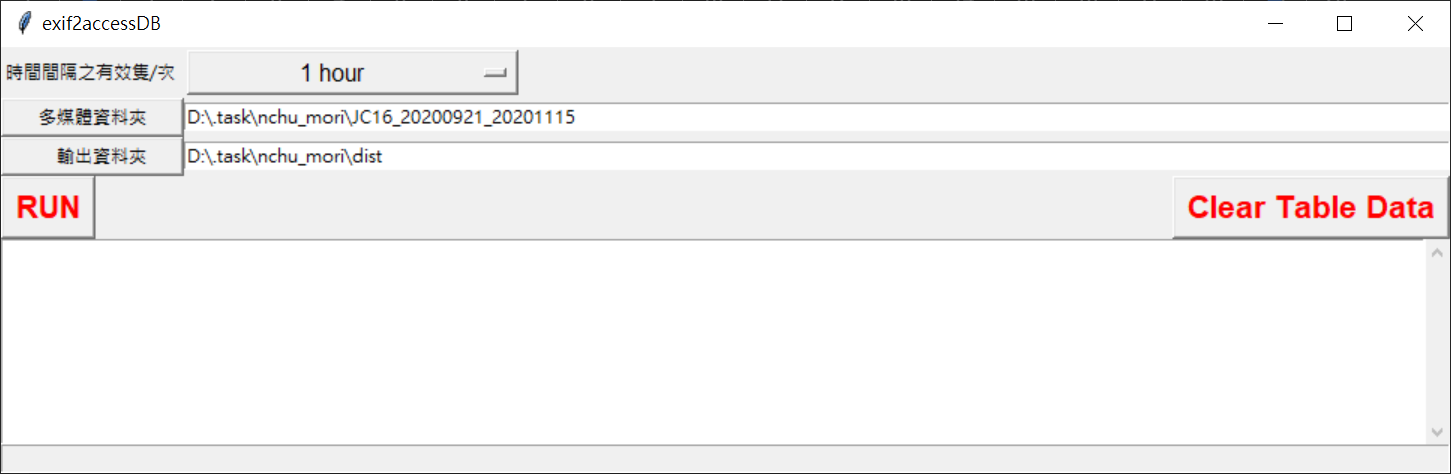




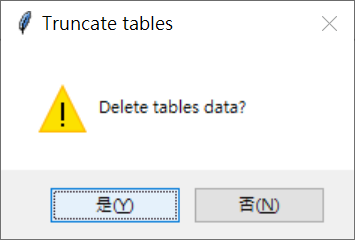
對於無cam\_id標籤的檔案，也能讓使用者了解需處理的範圍，或直接對accdb修改



2021/5/12 新增clear tables data，直接清空table資料

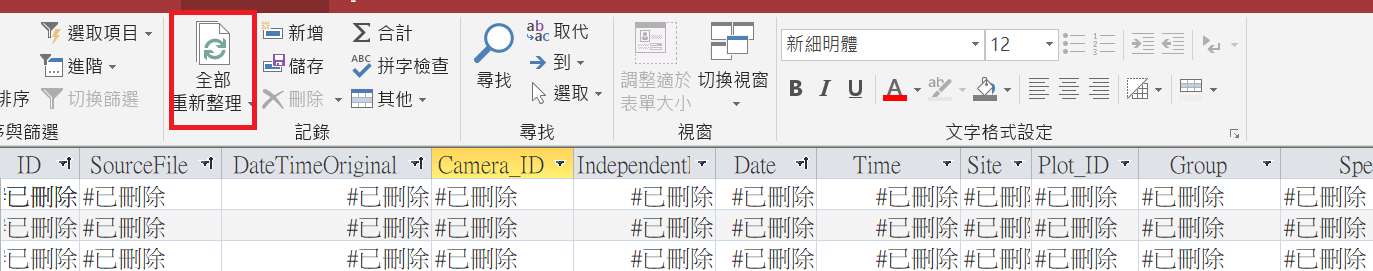


按下之後會跳出警示視窗

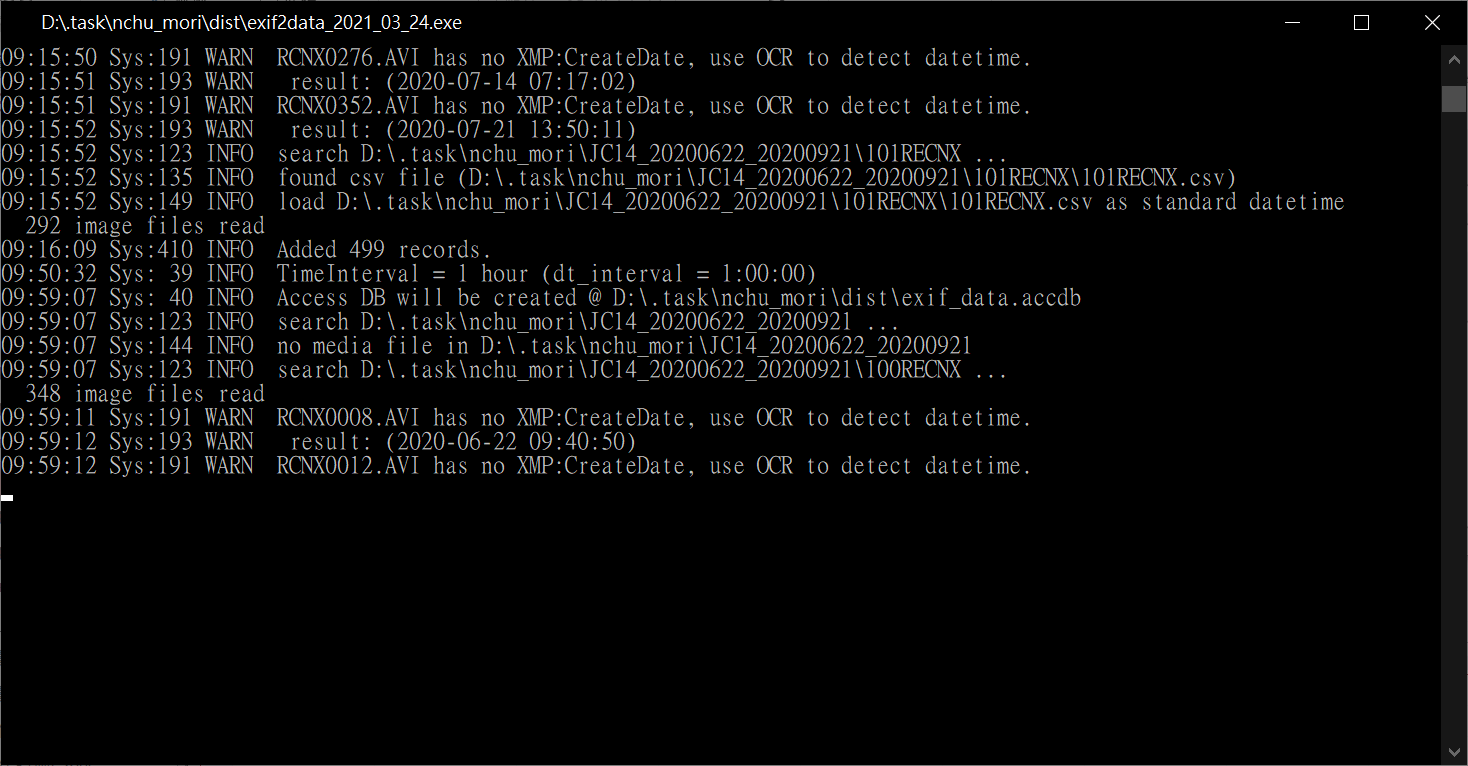


如果選yes的話，就會將table的資料都刪掉，省去手動刪除的麻煩

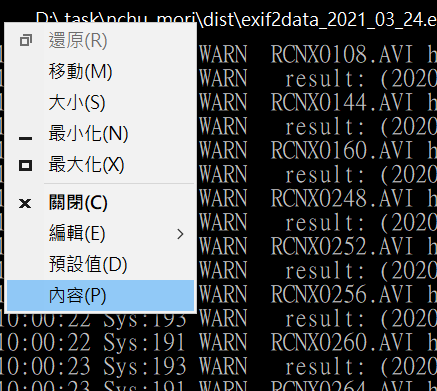
不過已顯示的table的data會變成已刪除，這時按上面的重新整理即可更新畫面



有時候console會因不明原因freeze(不會動)，這時請在console按一下any key(e.g.方向鍵)，就會再繼續運算



這個有方法可以避免，就是在左上角標題欄按下右鍵->內容



然後不要打勾"快速編輯模式"，確定之後應該就不會再freeze，大概或許可能應該...



如果要使用新的一輪，一定要先把"所有的"資料表的紀錄都刪除，以免舊紀錄影響工作時數、有效隻次、OI值

因為query也會參考PeriodMonth跟camera\_task\_record這些table，所以也必須清空

That is it!

========

參考資料:

Exiftool: <https://exiftool.org/>

CamtrapR:

https://cran.r-project.org/web/packages/camtrapR/index.html

<https://cran.r-project.org/web/packages/camtrapR/vignettes/camtrapr3.html>

==========

更新

---

忽略沒有species的tag，如果影像中本身就有勾選species，則還是會存成一個row在試算表中√

針對所有的子資料夾，如果有兩個以上的Camera ID (1\_Site ID)，為自動分別新增不同的camera\_task\_record跟PeriodMonth的時段資料√

個別的整體工作時間也會包含沒有動物的多媒體檔案√

將csv設為第一優先的時間設定√

解決exif標籤太長導致4096 block無法囊括所有資料，在解碼時會出現的錯誤√

N個以上的動物tag，這種情況要存N筆紀錄，且會提示在視窗中√

修正百分比的query√

增加每月的query(工作時數為限制內的時段, 同物種的總和)√

不輸入搜尋時間，時間自動設定成最前 或 最後√

所有小數點最多顯示到第三位√

有CAM的query，工作時間需要個別計算√

---