



前言

大家好，我是一位社會人士，有幸能參加第一屆TBIA生物開放資料數據松的競賽，最終得到人氣獎的肯定。由於我非常喜愛鳥類，心裡始終有個疑問，我們在網路上看到的鳥類新聞，媒體對於鳥類物種的報導是否有其偏好？或是物種被報導的頻率，其實與整體物種數量的多寡，是否有相關性呢？抱持著這樣的問題意識，我透過python爬蟲爬取三年內的聯合新聞網鳥類新聞，並比對分析生物多樣性資料庫裡的鳥類物種消長，以下內容即是針對這個問題得到的成果。

資料選擇與清理方式

準備資料表單如下：

- +備資料表+如+.

 - 1.一份「鳥類學名」csv檔案:透過爬蟲維基頁面後，整理成表單，供後續新聞資料分析時，判斷新聞是否為鳥類新聞，與抓取鳥種關鍵字。
 - 2.一份「台灣地名」csv檔案:透過爬蟲維基頁面後，整理成表單，供後續新聞資料分析時，判斷鳥種出現之區域，抓取區域關鍵字。
 - 3.一份「聯合新聞網(含鳥類新聞連結)」csv檔案:透過爬蟲udn網站後，整理成一份包含鳥類新聞的連結表單。
 - 4.一份「聯合新聞網(資料清洗完成)」csv檔案:透過爬取如上的連結，整理成一份包含「標題、報導媒體、內文、報導時間、鳥種關鍵字」的表單。

2021-2023年 聯合新聞網 & 生物多樣性資料庫

我們可以對照下方聯合新聞網與生物多樣性資料庫的圖表，兩者皆列出了2021到2023年、排名前15名左右的鳥種關鍵字。在2021年中，前15名、有5名是重疊的，分別為麻雀、赤腹鷹、灰面鶲鷺、鷗鷺、黃頭鷺。之後2022年兩者共同的關鍵字減少到3名，分別為麻雀、赤腹鷹、鷗鷺，到了2023年兩者重疊的只剩一個鳥種「麻雀」。從2021年到2023年的三年間，我們可以看到聯合新聞網在2021年報導的鳥種頻率，原先與此鳥種的數量多寡是有正相關的，可是這個趨勢逐年遞減，到了2023年，已看不出來相關性的存在。



2021-2023年 聯合新聞網 鳥種、地域文字雲

從台灣鳥種文字雲可以看到，最大版面由「黑面琵鷺」獨領風騷，由於是季節性的候鳥與「瀕危」物種，台灣在台南七股溼地有黑面琵鷺保護區，每到入冬此嬌客便是鳥類新聞偏愛的題材。另外，「小燕鷗」在桃園、宜蘭、花蓮皆不時有保育淨灘活動，也是屬於媒體喜愛報導的鳥種。接下來可以看到「水雉」、「領角鴞」、「鷗鷺」也是媒體常見報導的鳥類。

從新聞地域文字雲可以看到一個有趣的現象，原來台灣的媒體報導了不少的「中國」鳥類新聞，接著是日本和美國。而台灣本地的鳥類新聞，偏好的報導區域是嘉義、台南、台北。



2021-2023年 聯合新聞網 國家偏好鳥種

