

# 呂一昕



## 聯絡資訊

✉ yhsamlu@gmail.com

☎ 0972-733-147

@YiHsinLu (Sam Lu)

@yi\_hsin.lu

Yi-Hsin Lu

## 技術能力

### 程式語言

R  
PostgreSQL  
Python  
LaTeX

### 軟體 & 工具

視覺化  
(e.g. plotly, t-SNE, ...)  
資料分析  
(e.g. PCA, tidyverse, ...)  
機器學習  
(e.g. Random Forest, ...)  
網頁設計  
(e.g. HTML, CSS, ...)

### 語言能力

中文  
英文

## 學歷

### 碩士

應用數學系

📅 民國 110 年 2 月至民國 112 年 6 月  
📍 國立東華大學  
👤 指導教授：曹振海 教授  
📄 論文題目：爵士音樂家資料視覺化

### 大學

應用數學系 統計科學組

📅 民國 103 年 9 月至民國 108 年 9 月  
📍 國立東華大學

## 關於我

- 對於資料科學有著熱情和憧憬，在中研院擔任研究助理和研究所約三年，培養出豐富的資料分析、處理經驗。
- 我能運用統計、數學的背景和專業知識在工作和專案上，為此提供不同角度的靈感並提供其理論背景。
- 我不害怕挑戰新事物，擅長團隊合作、與人和睦、溝通良好，學生時代，無論是課堂或專案的團隊，還是課外的社團和球隊，都是值得信賴的夥伴。

## 專案經驗

### 爵士音樂家資料視覺化

#### 碩士論文

📅 民國 110 年 6 月至民國 112 年 6 月  
📍 國立東華大學

👤 指導教授：曹振海 教授

📄 論文全文(英文撰寫：Visualization for Jazz Musicians)

#### 📊 視覺化圖形成果

- 我們所建構的視覺化圖形用點的顏色和距離在平面中呈現音樂家之間在樂器、演奏風格和活躍年代的相似程度。
- 類別變數處理：  
資料的轉換結合 One-hot encoding 和 Jaccard，並將不同資訊的變數分開處理，可讓圖形中音樂家之間的關係更明顯，並顯示出更多資訊。
- 多值變數 (Multi-value Variable) 處理：  
我們運用三原色的調色處理多值變數上色的困難，舉例而言，大多數爵士音樂家都會使用多種樂器，導致顏色過於複雜而失去上色意義，但三原色的調色可以展現音樂家所使用樂器在不同種類之間的比例。
- 與過去的差別：  
我們呈現出全體音樂家之間的關係，有別於以往兩兩之間的呈現，另外距離和顏色也賦予相似程度的意義。

### 排球賽事與球員的資料視覺化

#### 個人興趣專案

📅 民國 112 年 10 月至今  
📍 家中進行

#### 📊 視覺化圖形：賽事結果。(包含球員數據表現的地圖)

- 我運用 HTML、CSS 和 JavaScript 建構賽事呈現的網頁，每場賽事都利用長條圖呈現兩支隊伍的數據差異。
- 利用球員在本場的比賽數據，經過降維，將球員的各項數據壓縮至平面，更直觀的呈現球員在不同面向的表現。
- 與過去的差別：  
我利用視覺化呈現比賽結果，有別於以往只有數據或資料矩陣，甚至在最頂級的世界聯賽依舊如此，其中也不容易看出球員和隊伍雙方的表現。

## 專案經驗

### 音樂風格 (統計機器學習)

#### 課程專案

📅 民國 111 年 9 月至民國 112 年 1 月  
📍 國立東華大學

👤 授課教授：曹振海 教授

- 我們運用各種統計方法和機器學習的模型針對音源檔中的各項數值進行曲風的分類。

- 建立模型：**

資料的變數眾多，運用並融合課程方法，經過挑選變數、擴增變數、標準化、變數交互作用處理，再經過不同模型的測試和融合，最後以 Random Forest 為主的模型顯示出最高的準確性。

### S&P Global Rating Xpress 資料庫

#### 資料庫管理

📅 民國 111 年 8 月至民國 112 年 1 月  
📍 國立東華大學

👤 專案主持人：朱至剛 教授

- 我們結合 R 和 PostgreSQL 建立出容易管理以及更新資料的程式。
- 資料庫使用說明**：從線上到資料文字檔 (包含 R 和 SQL 的原始碼)
- 資料更新說明書**：如何更新資料表。

## 研討會

### 第 32 屆南區統計研討會

暨 2023 年中華機率統計學會年會及學術研討會

📅 民國 112 年 6 月 26 日至民國 112 年 6 月 30 日  
📍 國立東華大學

- 發表人 (題目：爵士音樂家資料視覺化)
- 頒獎手
- 學習：區間變數的處理和上色的方法

### 2022 AI & Data Science Workshop

📅 民國 111 年 1 月 24 日至民國 111 年 1 月 26 日  
📍 國立東華大學

- 學習：文字探勘和機器學習的結合應用

### 第 30 屆南區統計研討會

暨 2021 年中華機率統計學會年會及學術研討會

📅 民國 110 年 10 月 30 日至民國 110 年 10 月 31 日  
📍 國立高雄大學

- 學習：以視覺化方法 SNE (Stochastic Neighbor Embedding) 為基礎的變形方法，針對不同的資料型態而改變，例如：t-SNE (t-distributed Stochastic Neighbor Embedding)。

## 工作經驗

### 研究助理 (服役前)

#### 資訊科學研究所

📅 民國 108 年 9 月至民國 109 年 1 月  
📍 中央研究院  
👤 指導教授：何建明 教授

總體經濟變數的資料分析與資料庫管理

## 社團經驗

### 隊長

#### 應用數學系男子排球隊

📅 民國 111 年 8 月至民國 112 年 6 月  
📍 國立東華大學

- 位置：舉球員 (球隊的大腦、司令官)

### 教練

#### 應用數學系女子排球隊

📅 民國 112 年 1 月至民國 112 年 6 月  
📍 國立東華大學

### 團長

#### 國樂社

📅 民國 105 年 9 月至民國 106 年 7 月  
📍 國立東華大學

- 管理樂團、學習指揮技巧
- 樂器：高音阮、大阮 (彈撥樂器)

## 特殊經歷

### 從右撇子到左撇子

#### 右手手指斷裂

📅 民國 107 年 9 月至民國 108 年 6 月  
📍 國立東華大學

即將畢業的那年，發生在一次的排球比賽中，我全力的撲球不慎導致右手指骨斷裂，嚴重的影響慣用右手的我，我不禁懷疑是否能完成最後一年的學業，甚至一度萌生休學的念頭。

我並不想輕易放棄，幸運的是小時候曾經因為好玩，練過左手，對於非慣用手的寫字、拿筷子、刷牙等動作並不陌生。那時，我不但必須用左手跟上課堂的速度、完成的作業和考試，還要額外付出更多時間練習把字寫得清楚。最後，我透過持續的努力，順利完成學業。