

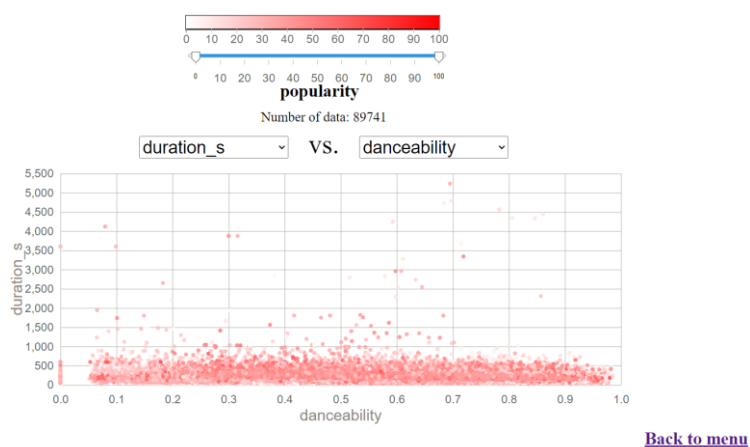
● 目標

觀察不同 popularity 的 track 有哪些不同特性分布，以分析怎樣的歌較受市場歡迎，未來創作新歌曲時，可參考此資料迅速掌握市場口味。我以 Scatter Plot 及 Collapsible Tree 兩張圖呈現資料的視覺化。

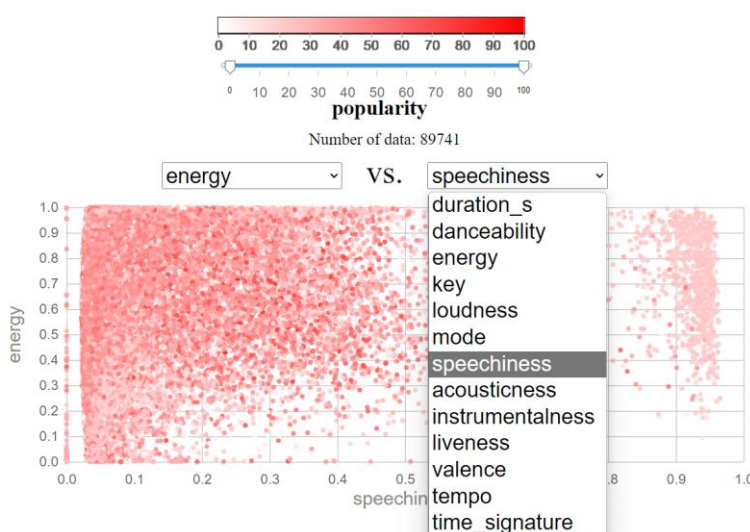
● 資料預處理

Dataset 中共有 114,000 筆 tracks，但其中有很多是重複的，所以先過濾 track_id 相同的資料，取獨有的 89,741 筆資料。

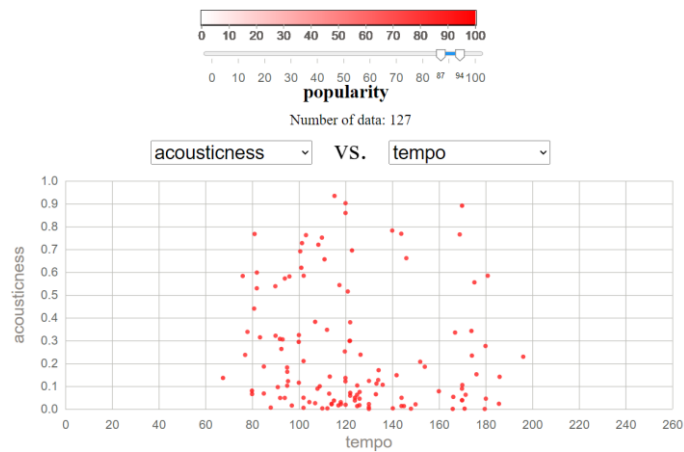
● Scatter Plot



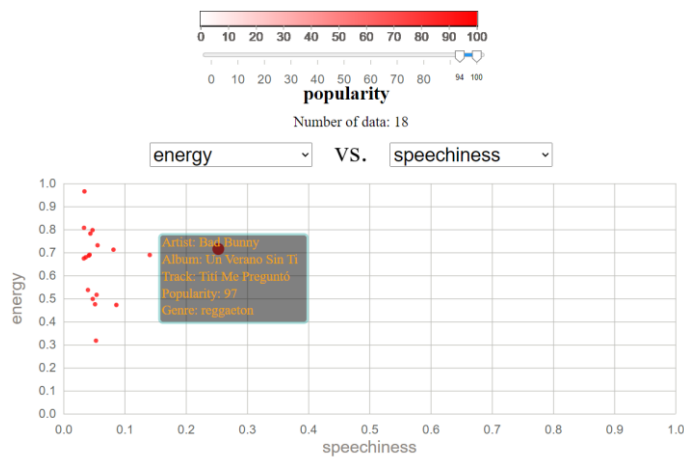
Scatter plot 中每一個點代表一個 track，以不同顏色深度的紅點代表該 track 的 popularity，點越紅該 track 越有名。



x、y 軸有下拉式選單可以分別選擇 track 的不同屬性，以觀察各 popularity 在不同屬性的分布。



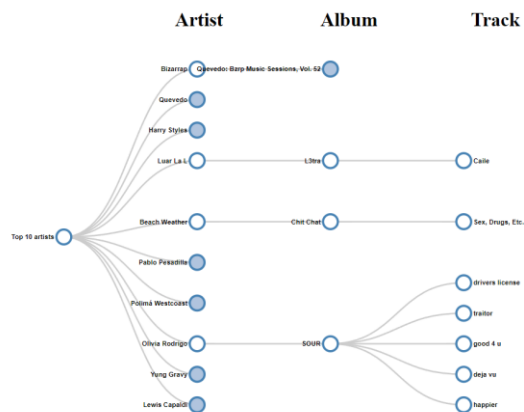
由於資料量大，圖中的點會大量重疊，不利我們觀察，所以上方我製作一個可拖曳的 slider，可讓圖中只顯示特定範圍 popularity 的資料點。



將游標放在點上，會出現該資料點的資訊。由此圖可觀察出，較受歡迎的 track energy 值屬於中高，speechiness 則很低。

● Collapsible Tree

Top 10 artists with at least songs in dataset (Enter a number from 1 to 100)

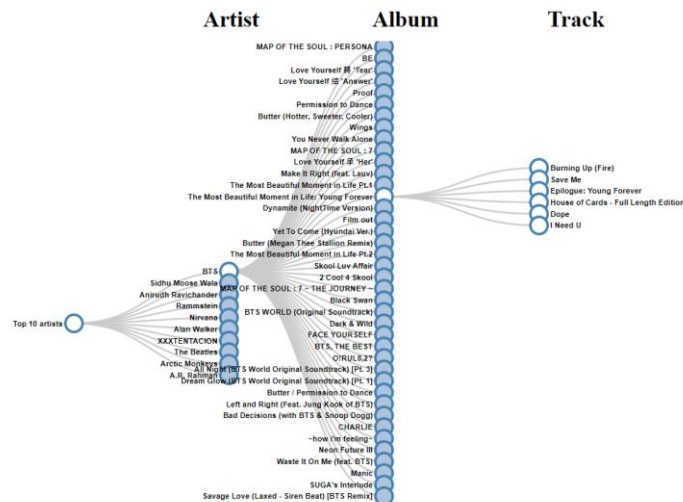


[Back to menu](#)

找出 top 10 artists 並以 collapsible tree 呈現他的所有作品，計算 top 10 artists 時，我先將每個 artist 的所有 tracks 的 popularity 取平均，如果一個 track

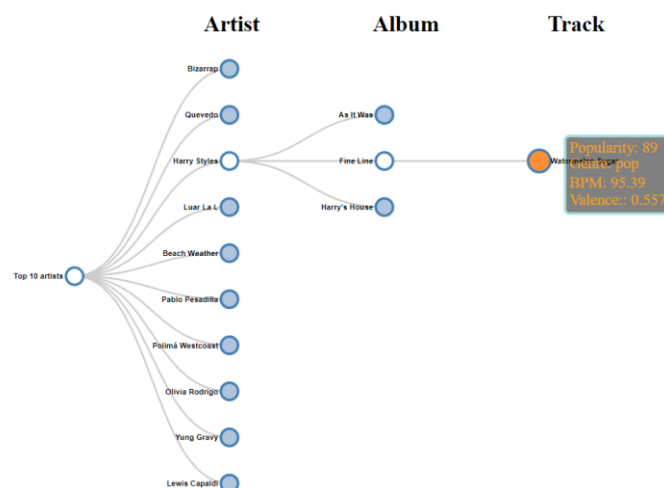
有多個創作者，所有其創作者都會使用到該 track popularity 來他們各自計算平均，再從各 artist 的平均 popularity 中取最高的前 10 名。除了 tree 中最尾端的 node，每個 node 都可以透過點擊展開或收合，藍色實心的 node 為還可以展開的 node，藍框空心的 node 為已經展開的 node。

Top 10 artists with at least songs in dataset (Enter a number from 1 to 100)



從 collapsible tree 的第一張圖中可以觀察出 top 10 artists 的創作量都很少，這並符合直覺，在這種情況計算出的 top 10 artists 可能只是某一首 track 曾經爆紅，而不是該 artist 本人很有名，所以我在圖的上方多加了可讓使用者輸入參數的地方，限制呈現的 top 10 artists 在 dataset 中至少要有幾個 tracks。方塊中允許輸入數字 1~100，並設有防呆裝置，若輸入不在這個範圍會無效。

Top 10 artists with at least songs in dataset (Enter a number from 1 to 100)



將游標放在 tree 中最尾端的 node，圖中會顯示該 track 的資訊。