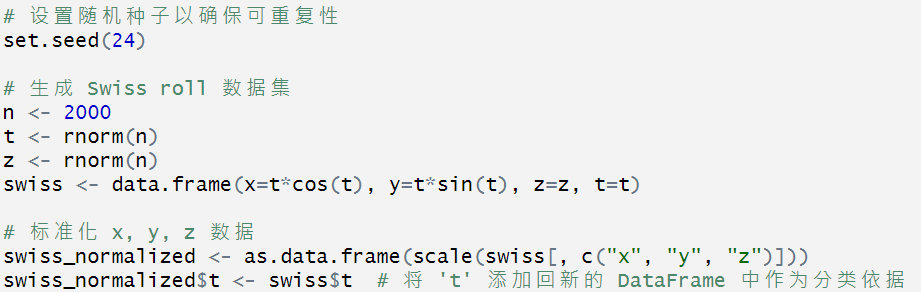
* 資料介紹

1. 生成方式

* t：從標準常態分佈生成。在Swiss roll結構中，𝑡 代表從捲起原點出發沿螺旋結構的距離或者 "長度"。t 的值決定了點在螺旋上的位置，並通過cos(t)和sin(t)影響每個數據點的x和y座標。
* x 和 y 的計算：x=t⋅cos(t), y=t⋅sin(t)

這種計算方式使得數據點按照螺旋的方式排列，隨著 t 增加，每個點圍繞中心向外擴展，形成一個漸進捲起的結構。

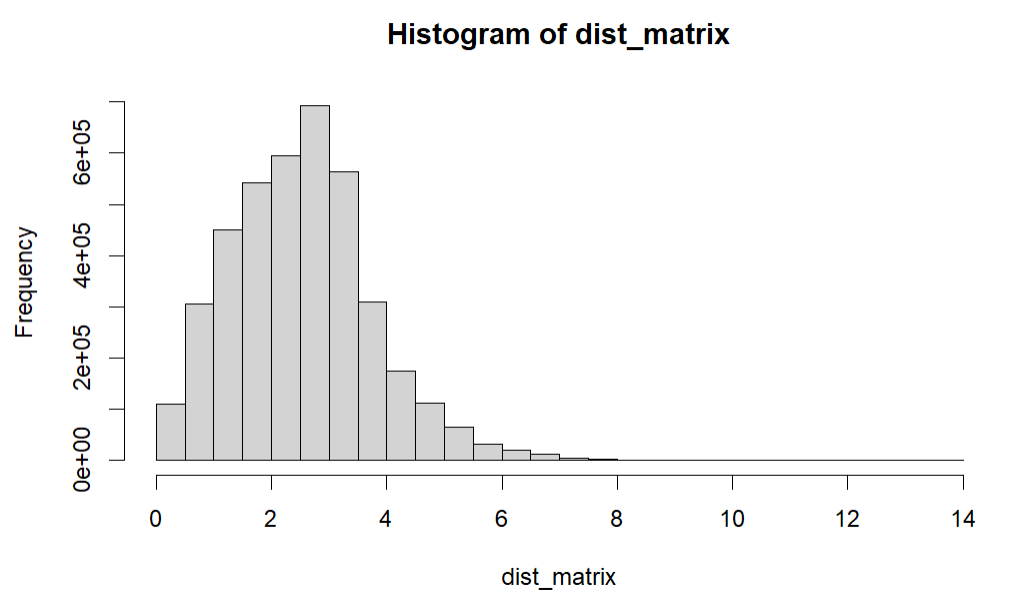
* 參數 z：z也是從0到10的均勻分佈生成的，它代表點在垂直於捲曲平面的方向上的位置。這使得 Swiss roll 有了第三維，增加了數據集的複雜性。
* 將以上基礎建構出的資料集標準化，便是此次分析所使用的資料



1. 距離矩陣

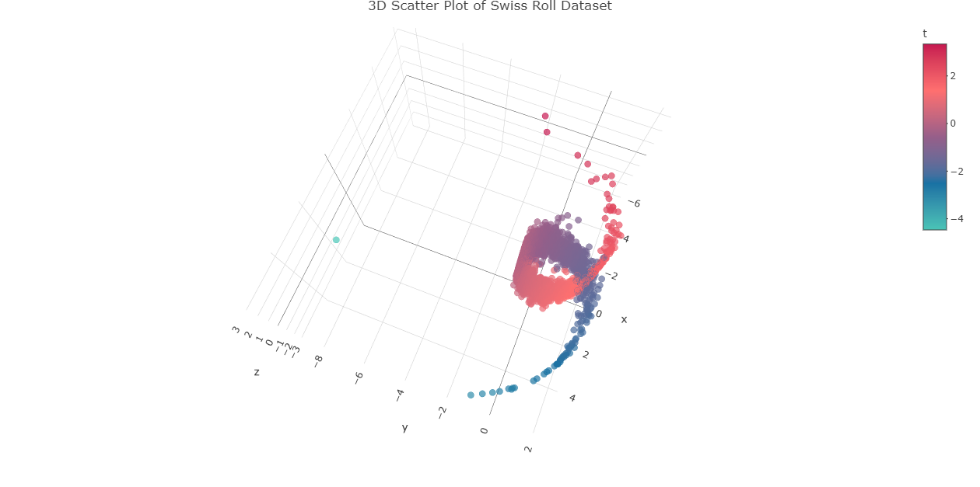
觀察標準化後的資料之距離矩陣，可以看出絕大多數的點距離介於1~5中間，在使用 ”enn” 方法轉換資料時，或許介於此中間的是較為妥當的。

由於kernal-weight方法也有牽涉到距離矩陣，這個直方圖也可以給予其參數的相關資訊。

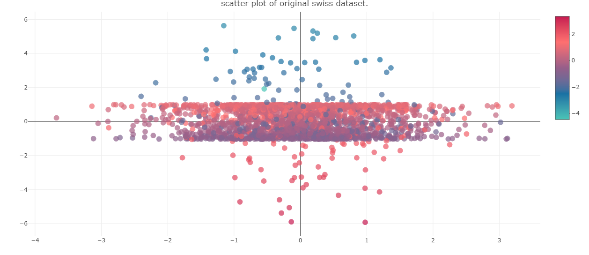
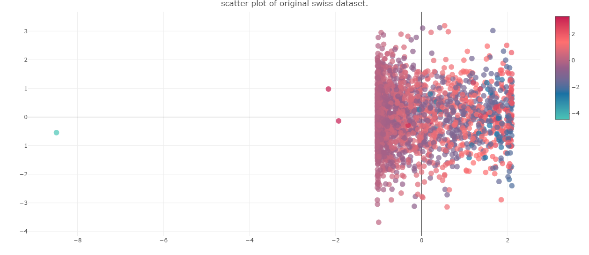
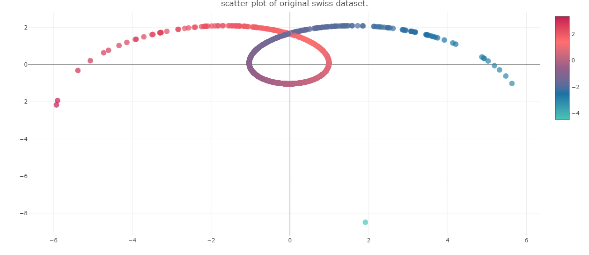


1. 立體圖與二維圖

* 3d：資料呈現捲狀從中心向外旋轉，由於資料建構基礎是以參數t為基準，所以我們將根據參數t的大小作顏色漸變，方便觀察後續的分析結果。

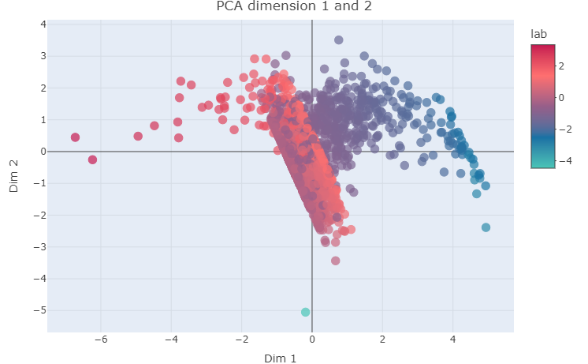


* 2d：左方圖為固定z軸，右方圖為固定y軸，可以看出顏色變化的位置以x,y軸上可觀測到的為主，故降維觀測後的顏色變化應該要與下左圖類似。



* 分析

1. PCA

(mds?)

1. Laplacian Eigenmap

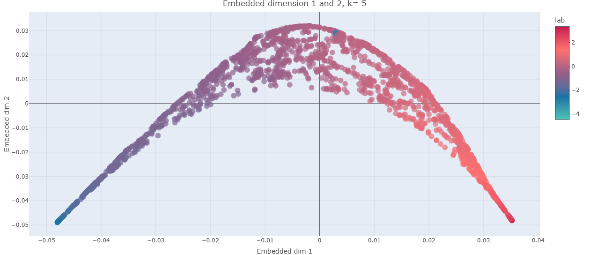
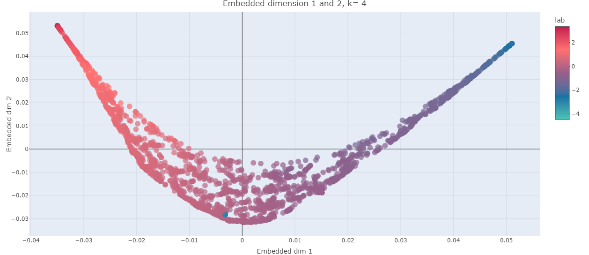
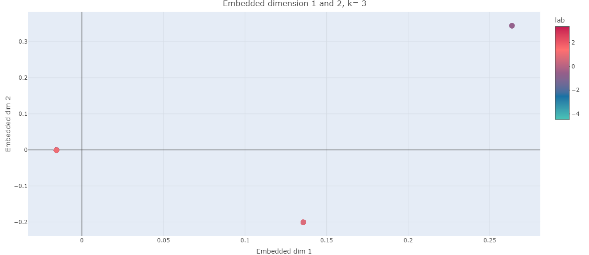
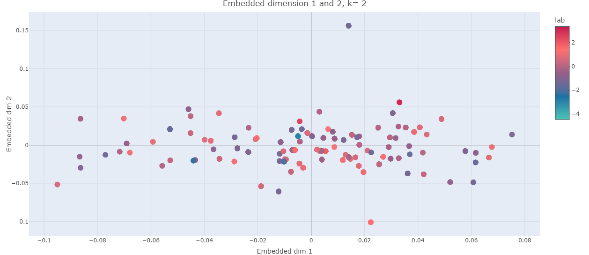
制定實驗計劃如下：

|  |  |  |  |
| --- | --- | --- | --- |
| 實驗次數 | 是否採用熱核 | 建圖方法 | 觀察的維度 |
| 1 | 否 | knn | Dim1, Dim2 |
| 2 | 否 | knn | 比較維度 |
| 3 | 是 | knn | Dim1, Dim2 |
| 4 | 是 | knn | 比較維度 |
| 5 | 否 | enn | Dim1, Dim2 |
| 6 | 否 | enn | 比較維度 |
| 7 | 是 | enn | Dim1, Dim2 |
| 8 | 是 | enn | 比較維度 |

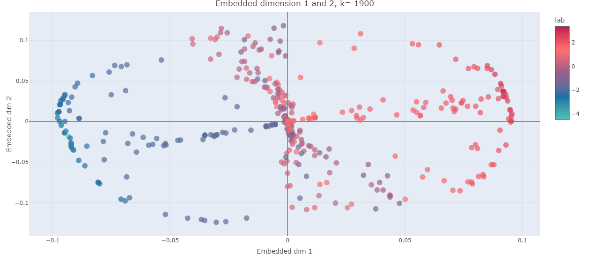
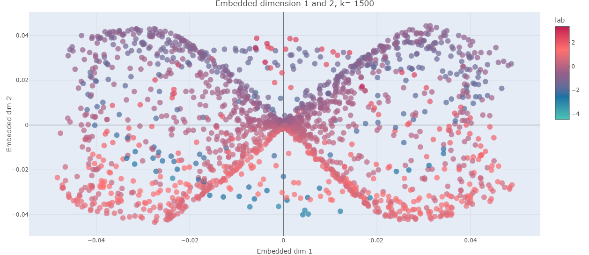
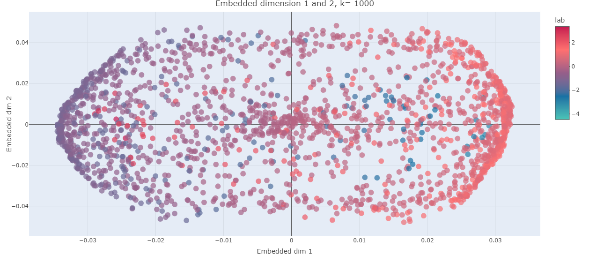
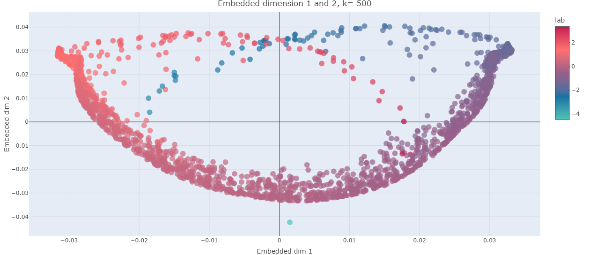
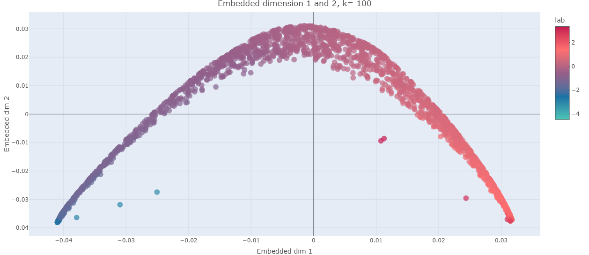
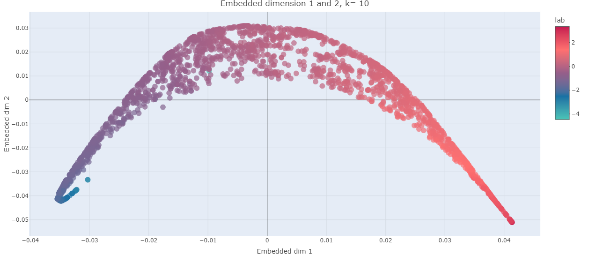
1. Knn without weighted (dim1&2)

嘗試 k = 1, 2, 3, 5, 10, 100, 1000, 1500, 2000：

* k=1 ~3時，很容易出現點重合在一起的情況。因為當2個或3個點是彼此的k近鄰，把它們映射到同一個點就可以完美地保持局部平滑，但當k大於等於4後此情況消失，代表當具有四個比較點時，轉換更能夠保留點與點的相對距離關係且同時保留局部平滑性質，但仍有些許不足。



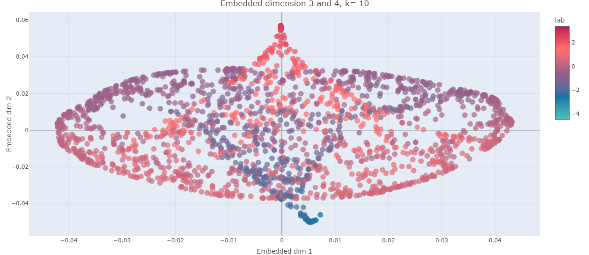
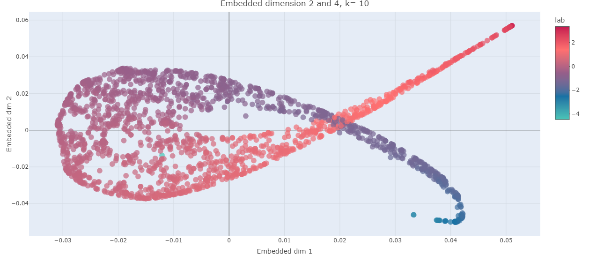
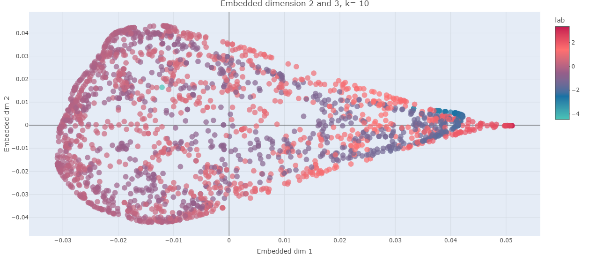
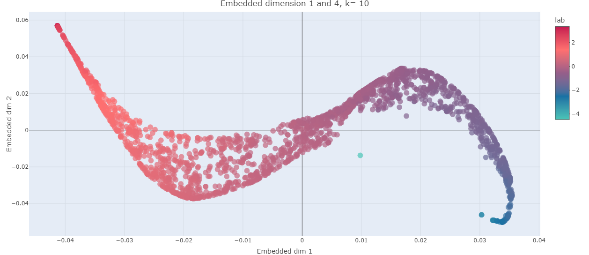
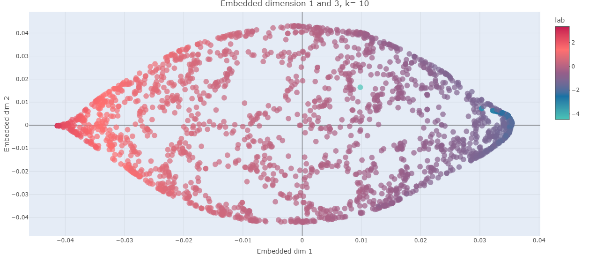
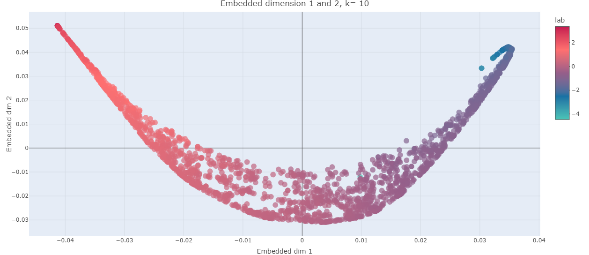
* 由於 Swiss data為全連通圖型，且資料點共2000個，當 k 具有一定大小內，都還是仍能夠在轉換過後維持原先的顏色變化 (即相近點保持原有的相近性質)
* 但當k值接近四分之一時，轉換保留的性質從顏色變化的方向轉變成整體圖型的大方向旋轉
* 再往上的k就無法保留具體的相對位置



1. Knn without weighted (dim)

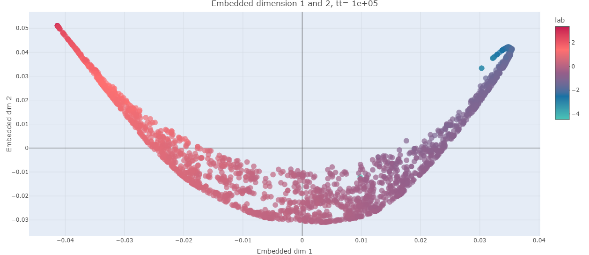
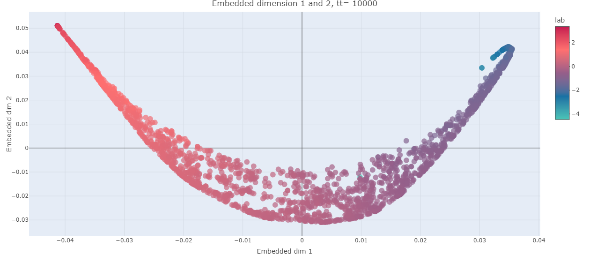
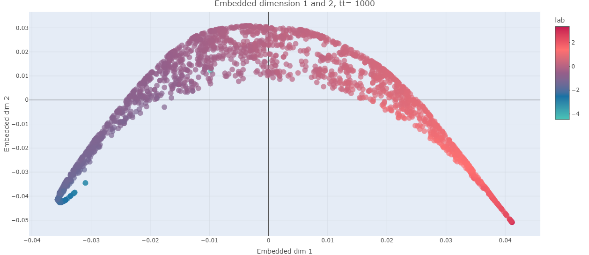
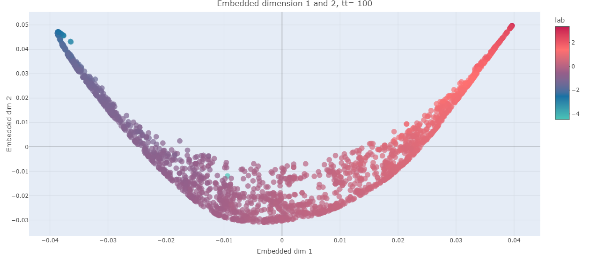
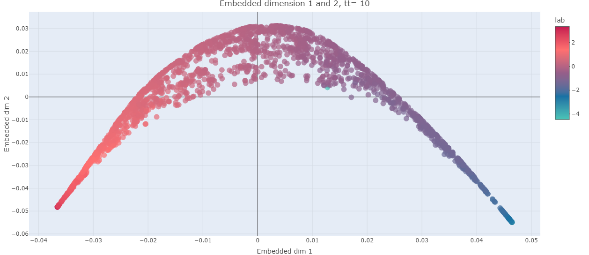
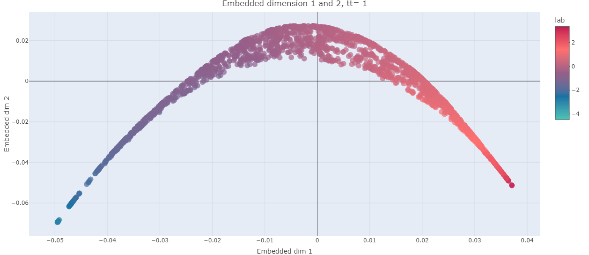
根據上點，我們選擇k=100作為合適的參數大小，並比較不同

可以看出第一維度(顏色變化???)、第二維度(



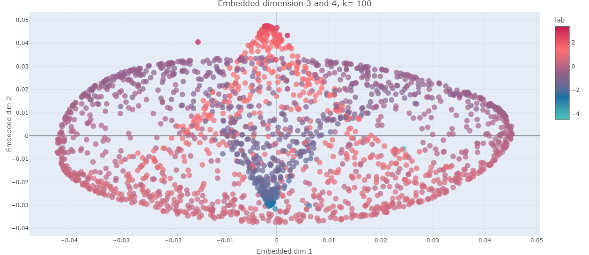
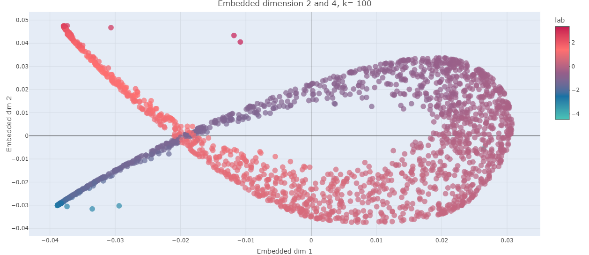
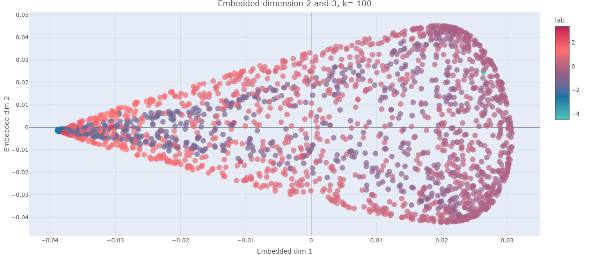
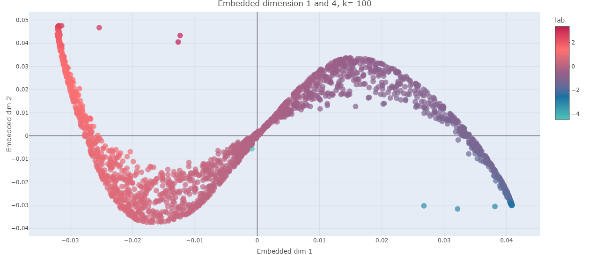
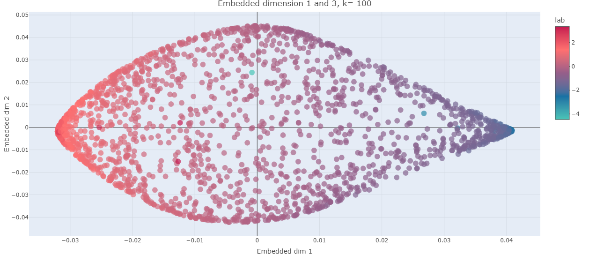
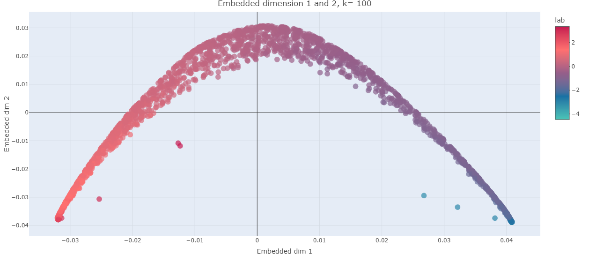
1. Knn with weighted(dim1&2)

固定 k = 100，嘗試tt= 1, 10, 100, 1000, 1e5, 1e7，無明顯差異，可知全連通圖型若是已選擇有效的k值，weighted本身影響不大，僅會影響一些未連通的邊陲點。



1. Knn with weighted(dim)

固定 k = 100，tt=100，嘗試觀察不同維度之結果(與無weighted版本相同)



1. enn without weighted(dim12)

