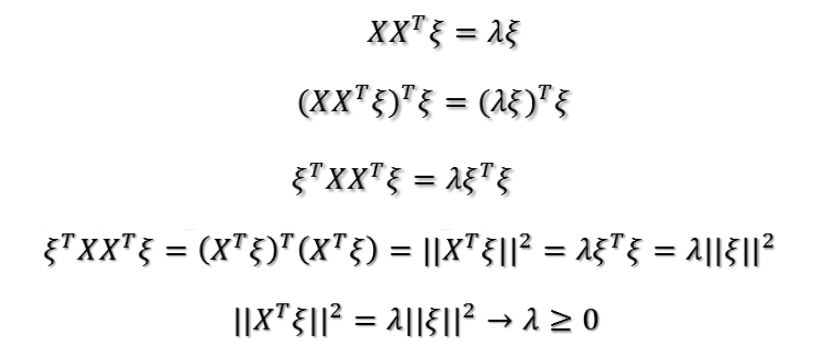
Machine Learning HW #2

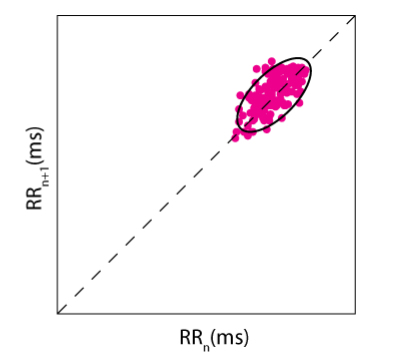
1. 直線 L 的指向向量 \mathbf{w} 就是樣本共變異數矩陣 S 的特徵向量。因為 \mathbf{w}^TS\mathbf{w}=\mu\mathbf{w}^T\mathbf{w}=\mu，欲使 \mathbf{w}^TS\mathbf{w} 有最大值，我們必須選擇對應最大特徵值的特徵向量。實對稱半正定矩陣有非負的特徵值，故可設 S 的特徵值為 \lambda_1\ge\lambda_2\cdots\ge\lambda_p\ge 0。此外，S 的特徵向量構成一個單範正交集 (orthonormal set)，

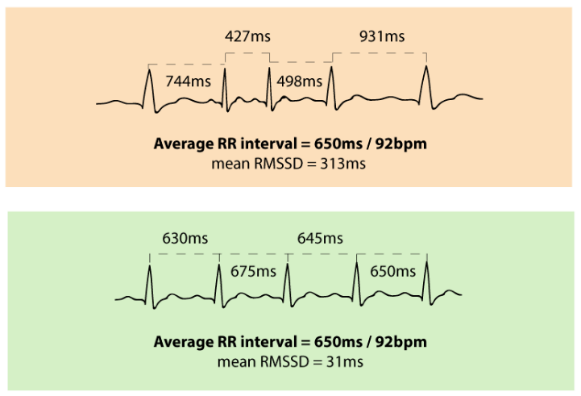
(a) 實對稱矩陣可正交對角化(orthogonally diagonalizable)

(b) semi-definite半正定的情況下，所有特徵值要大於等於0

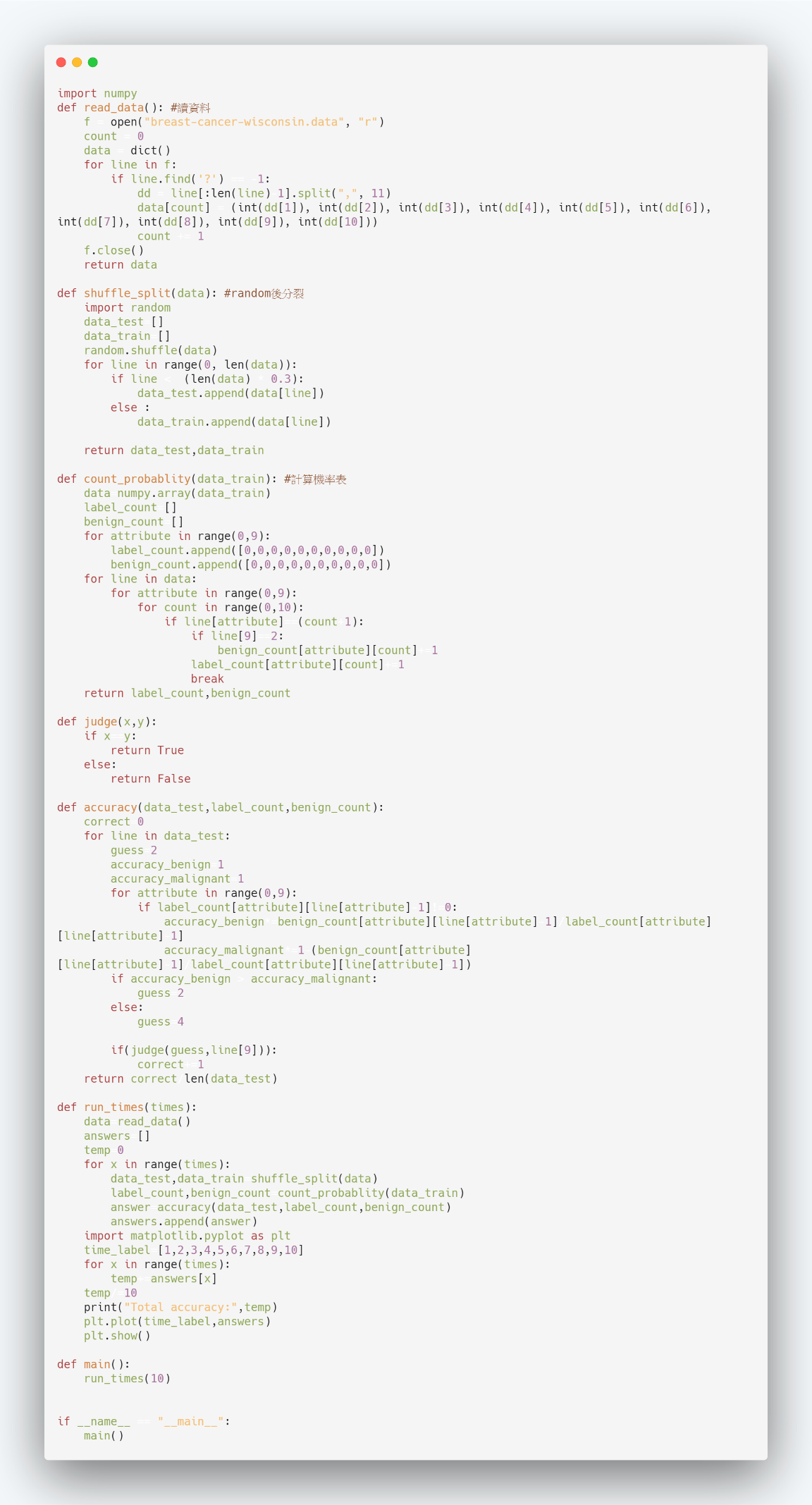
假設的某一個特徵值為，對應的特徵向量為，則

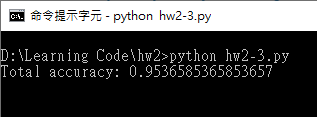
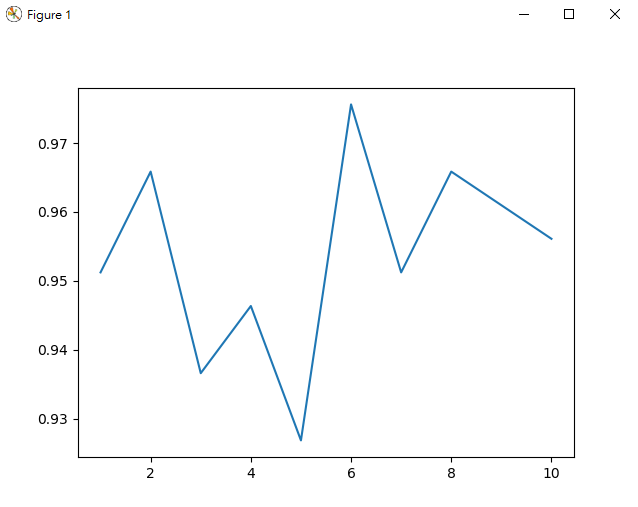


2. 右圖比較像是來自於一個健康的個體，因為他的數據比較集中，比較不會有心律不整的問題。

用平均的RR interval可能看不出來，但是用加權的RMSSD就可以看的出來差別

3. 如下頁

(a)



(b)

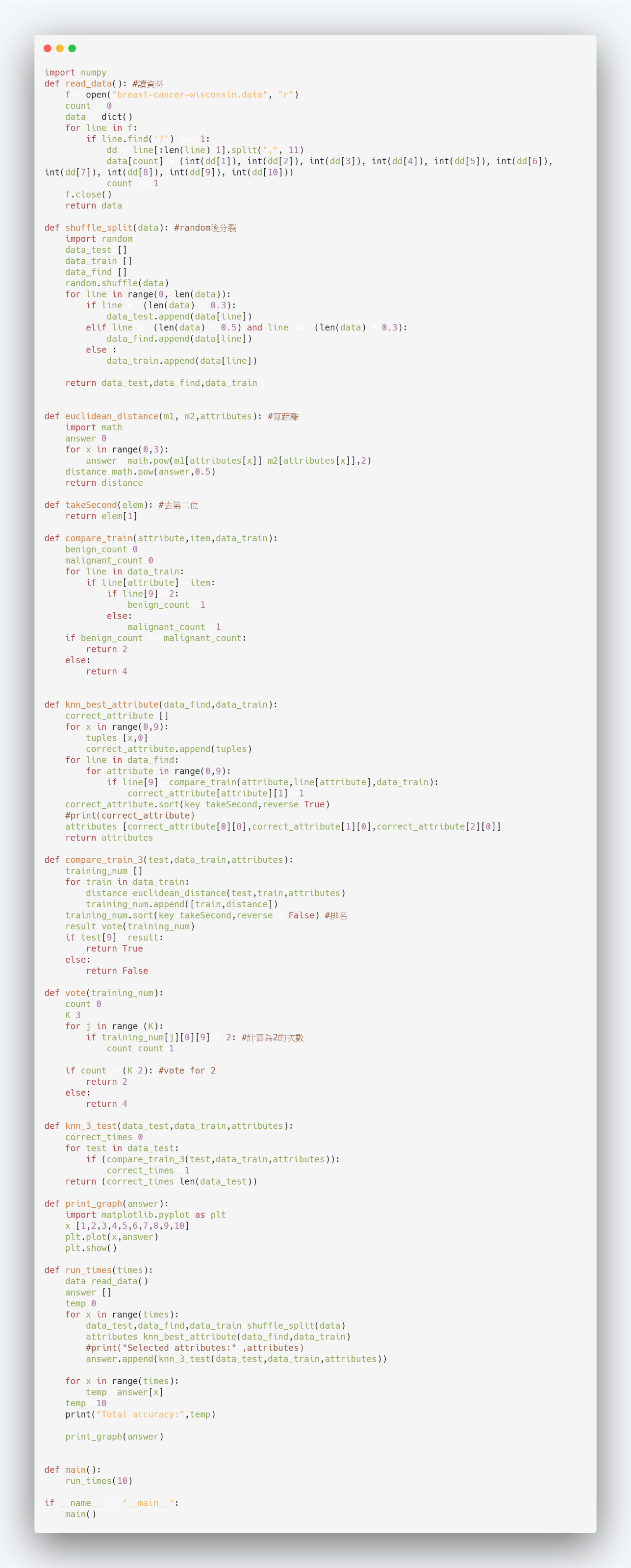
Oversampling: 隨機地複制少數類別的數據。

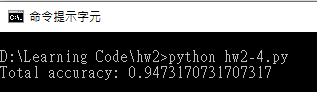
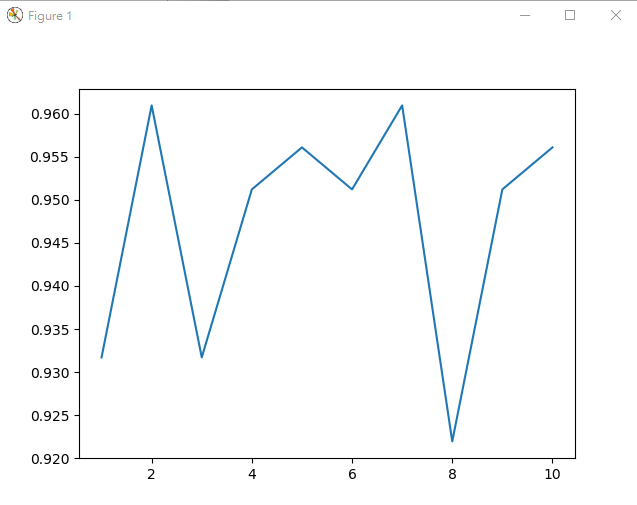
Undersampling: 隨機地刪去多數類別的數據。

Resampling: 不影響本身的數據樣本，把新的數據組合成新的數據。

4. 如下頁



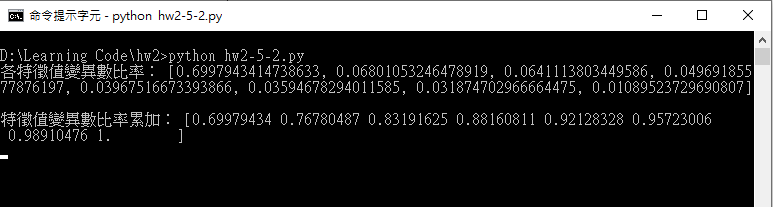
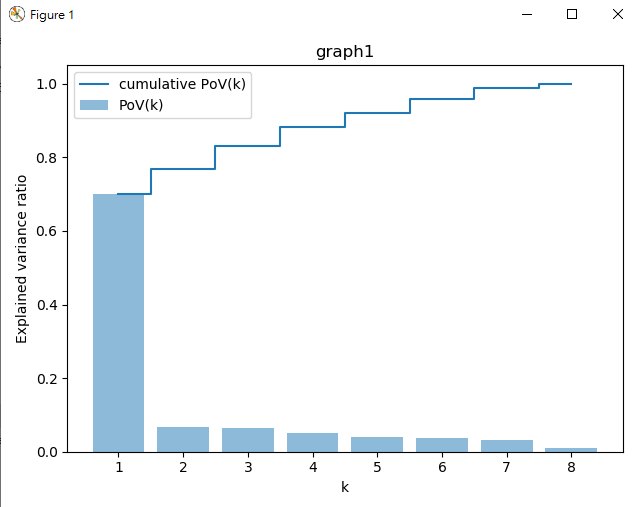


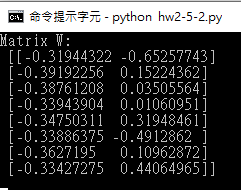
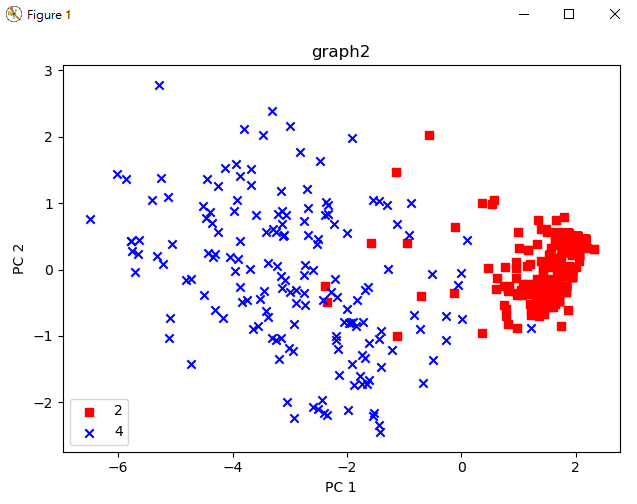


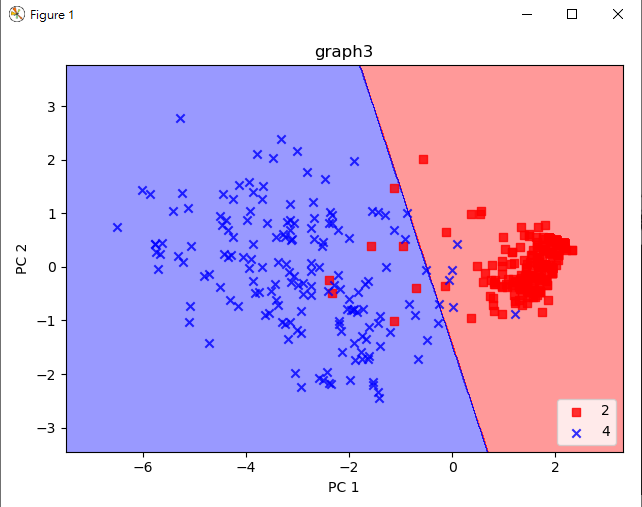
5. (a)如下頁











(b) 因為在維度下降的時候可能會不小心把重要的維度刪除，所以要平均多次一點來避免這個風險。