

Q: What is the problem of this result? What can be done to get better result?

Please give an explanation.

A:

1. 這個結果的問題在於，即使第一個群體的死亡比例低於第二個群體，但因為第一個群體的患者數量較多，導致實際死亡人數更高。這可能代表 **k-means** 沒有很好地捕捉到導致死亡的關鍵特徵，或者這些特徵在數據中的變異並不足以對患者進行有效的分類。

2. 為了獲得更好的結果，可以嘗試以兩種方法:

(1) 特徵標準化：確保所有特徵都在相同的尺度上可以增進聚類的性能，因為 **k-means** 算法對特徵的尺度敏感。

(2) 特徵選擇：分析哪些特徵對於預測患者結果最重要，並考慮只使用這些重要特徵進行聚類。

此外，關於 **ST elevation**，他是心電圖中一個重要的特徵，他與死亡與否有密切的關聯，或許提高 **ST elevation** 的權重也是不錯的方法。