* 在對**”路線11站點與使用次數”**之矩陣做PCA降維時發現需要使用到第11 PC才能夠有超過60%的解釋力，由於維度較高，不便畫在圖形上做觀察，因此相較於需要事先判斷分成 k 群的k-means分群方法，我們更傾向於使用Hierarchical Clustering 的分群方式，進行下一步的分析。
* 在觀察**”路線11站點與使用次數”**之矩陣時發現，有部分乘客所使用的站點是相同的，因此我們將其視為一種**”搭車模式”**並將具有相同搭車模式之乘客累加起來，此舉所附帶的好處便是能夠進一步降低資料的維度。
* 在對整理過後的**”路線11站點與搭車模式”**之矩陣做PCA時發現，此時只需要使用到第3 PC就能夠有超過60%的解釋力，因此我們先將DATA投影在由PC1-PC3所建構出的3D圖上，再從不同角度去觀察，判斷分成幾群，並使用k-means分群方法，去觀察會將哪先車站分到不同的群裡面，同時我們也有使用Hierarchical Clustering 的分群方式兩相對照，觀察分群的結果有何不同。
* 對**” 路線11站點與搭車模式”**之矩陣做完分群之後，我們將分群結果對照**”各站點使用量”**之長條圖，發現有部分使用量非常多的站點並沒有被分出來，因此我們猜測 :

1. 可能是因為對**”搭車模式”**做分群，而該站點之使用量是由多種不同搭車模式提供，造成此站點沒有被區分出來。
2. 是否因為不同的族群(老人、學生)搭車習慣不一樣導致此結果。
   * 依照乘車票種進一步將**”路線11站點與搭車模式”**分為

**”路線11(敬老票)站點與搭車模式”**與

**”路線11(學生票)站點與搭車模式”**進行分析。

* 乘客刷卡記錄只占總搭乘率之20%
* 每天公車發車時間相同，但透過畫出一整年非假日乘客搭乘公車刷卡記錄的點圖，並按照公車行經站點排序，發現乘客搭車時間誤差明顯偏高，初步猜測是因為每日的路況都不盡相同，而造成搭乘時間有較大的變動。