

$$(X, O) = e^{-\frac{x^2}{2a^2}}$$

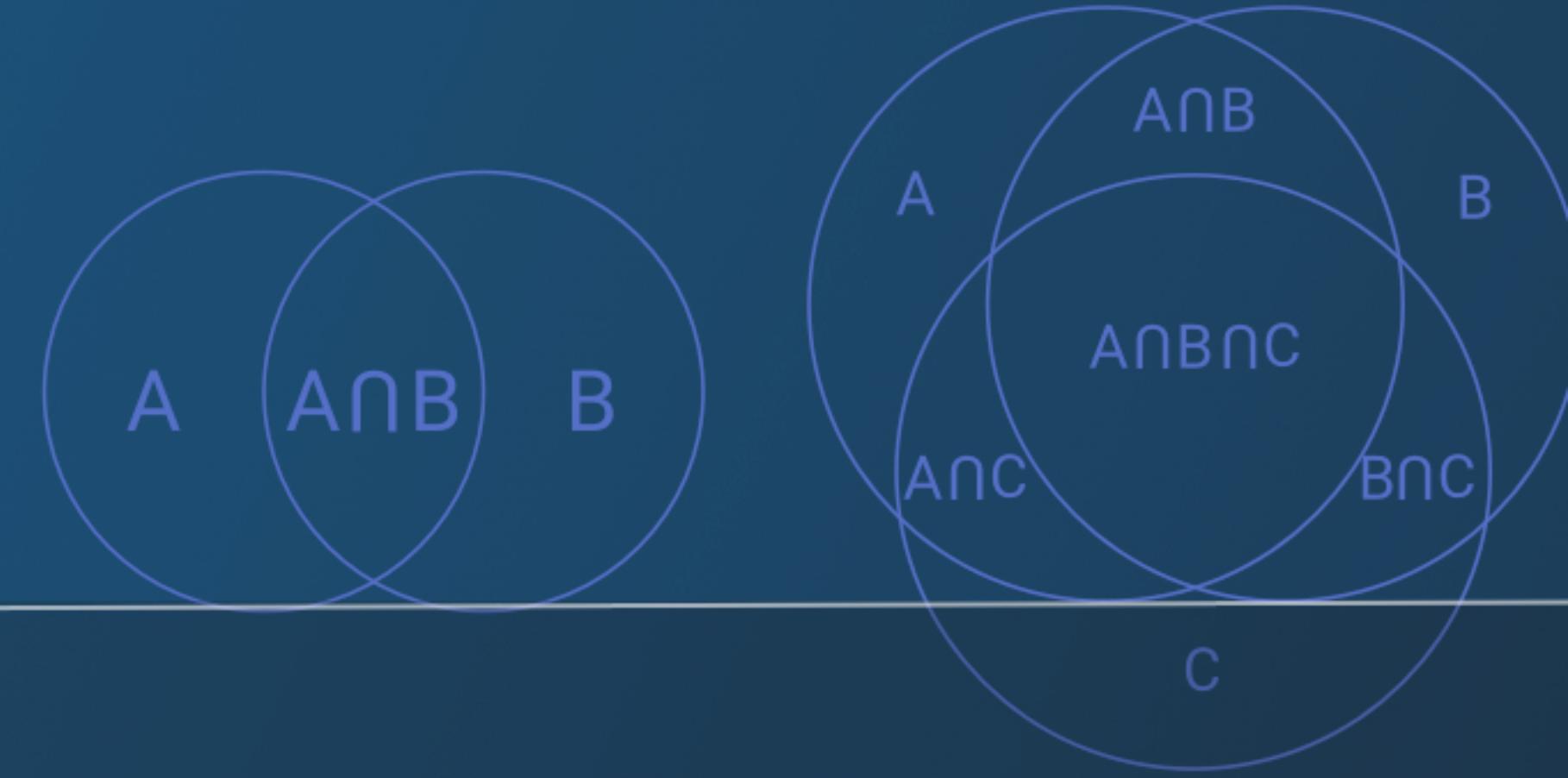
$$x(X, O) = -\frac{x}{\sigma^2} G(X, O) = -\frac{x}{\sigma^2} e^{-\frac{x^2}{2a^2}}$$

$$xx(X, O) = \frac{x^2 - \sigma^2}{\sigma^4} G(X, O) = \frac{x^2 - \sigma^2}{\sigma^4} e^{-\frac{x^2}{2a^2}}$$

$$xxx(X, O) = -\frac{x^3 - x\sigma^2}{\sigma^6} G(X, O) = -\frac{x^3 - x\sigma^2}{\sigma^6} e^{-\frac{x^2}{2a^2}}$$

# Julia 程式語言學習馬拉松

Day 28



$$\ln(x + \sqrt{1+x^2}) + x - \frac{1}{x + \sqrt{1+x^2}} \left( 1 + \frac{x}{\sqrt{1+x^2}} \right)$$

$$\ln(x + \sqrt{1+x^2}) + x - \frac{1}{x + \sqrt{1+x^2}} \left( \frac{\sqrt{1+x^2} + x}{\sqrt{1+x^2}} \right)$$



cupay

陪跑專家 : Andy Tu

# Clustering 分群模型





# 重要知識點



- 認識 Clustering 套件
- 了解 k-means 的用法
- 了解 k-medoids 的用法
- 了解 heirarchical clustering 的用法



# Clustering 套件



- Clustering 套件提供了多種的分群模型：

- K-means
- K-medoids
- Hierarchical Clustering
- Markov Cluster Algorithm
- Affinity Propagation
- DBSCAN
- Fuzzy C-means



# K-means 模型



- K-means 是分群方法中最簡單的，需要事先決定分群的數量 k。
- 這邊以 iris 資料集做示範，需先將資料矩陣轉置，變成特徵-樣本的維度。

```
iris = dataset("datasets", "iris")
X = Matrix(iris[, 1:2])'
k = 3
result = kmeans(X, k)
```



# K-means 模型

- K-means 的分群結果可以由 assignments 函式取得。
- 我們可以試著將分群結果繪製在座標平面上。

```
iris[:cluster] =  
string.assignments(result)  
plot(iris, x="SepalLength",  
y="SepalWidth", color="cluster",  
Geom.point)
```



# K-means 模型



- 如果想取得 K-means 的群中心座標可以由 result.centers 取得。

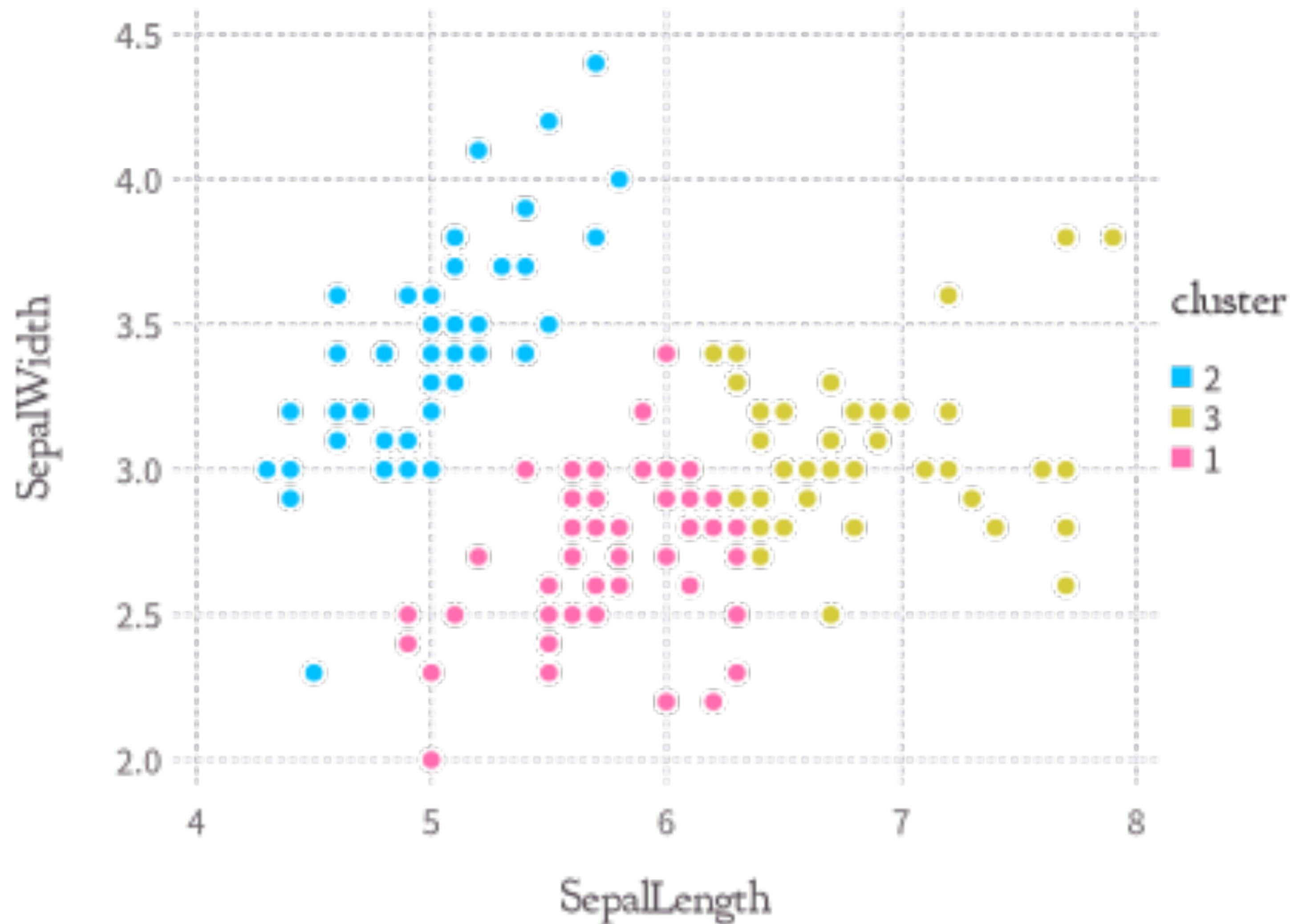
```
result.centers
```

- counts 函式可以取得各群的樣本數量。

```
counts(result)
```



# 對 iris 分群 - k-means





# K-medoids 模型



- K-medoids 同樣需要事先決定分群數目。
- 程式上與 k-means 不同的是，需要事先計算距離矩陣。
- 這邊我們可以先對資料點計算兩兩的距離矩陣 D，再放入 kmmedoids 函式中。

```
k = 3
```

```
result = kmmedoids(D, k)
```



# K-medoids 模型



- 我們同樣可以從 assignments 函式中取得分群結果。
- 並且將分群結果繪製成圖表。

```
iris[:cluster] =  
string.assignments(result)  
plot(iris, x="SepalLength",  
y="SepalWidth", color="cluster",  
Geom.point)
```



# K-medoids 模型

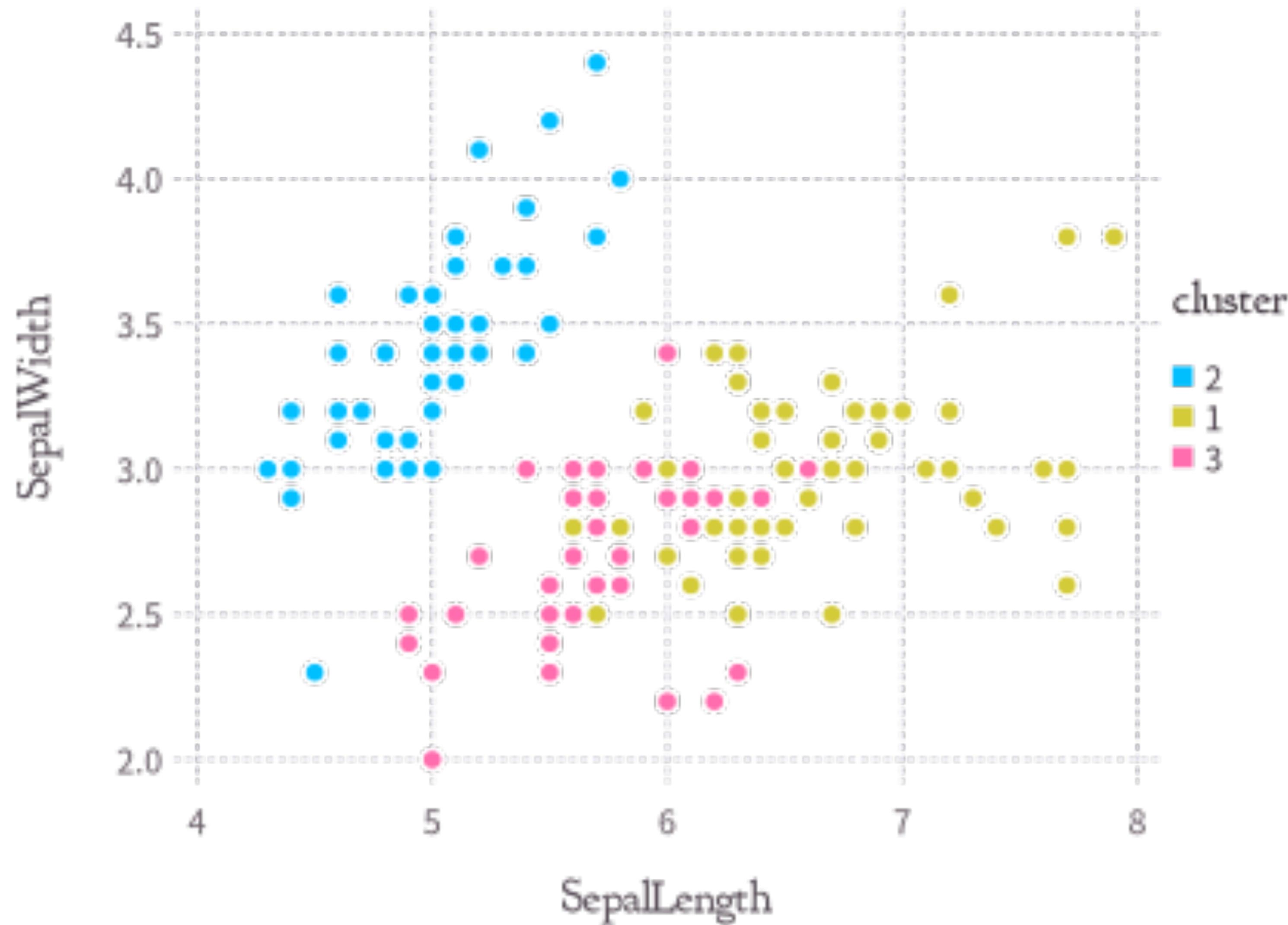


- 要取得群中心可以從 `result.medoids` 取得資料的 index。

```
iris[result.medoids, :]
```



# 對 iris 分群 - k-medoids





# Hierarchical Clustering



- 階層式分群法是不需要事先決定群數的分群方法。
- 在使用之前需先計算資料與資料之間的距離，這邊可以使用 Distance 套件所提供的距離函式。
- 得到距離矩陣  $D$  後，便可以直接放入 `hclust` 函式計算。

```
result = hclust(D, linkage=:single)
```



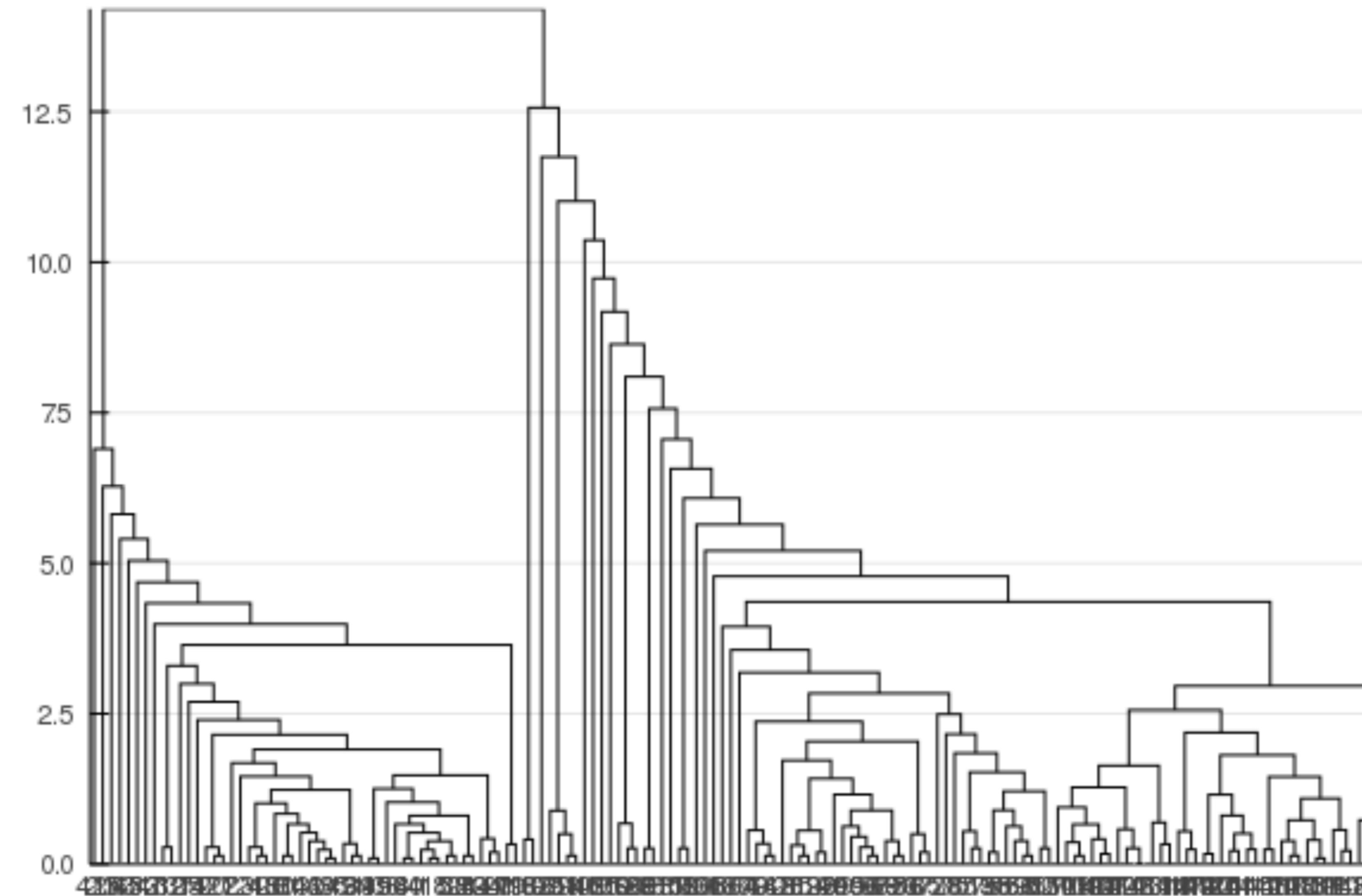
# Hierarchical Clustering



- 目前提供的 linkage 有：
  - :single
  - :average
  - :complete
  - :ward
  - :ward\_presquared



# 對 iris 分群 - hierarchical clustering



# 知識點 回顧

- 認識 Clustering 套件
- 了解 k-means 的用法
- 了解 k-medoids 的用法
- 了解 heirarchical clustering 的用法



## 推薦閱讀

- [Clustering.jl 官方文件](#)





解題時間

請跳出 PDF 至官網 Sample Code  
& 作業開始解題