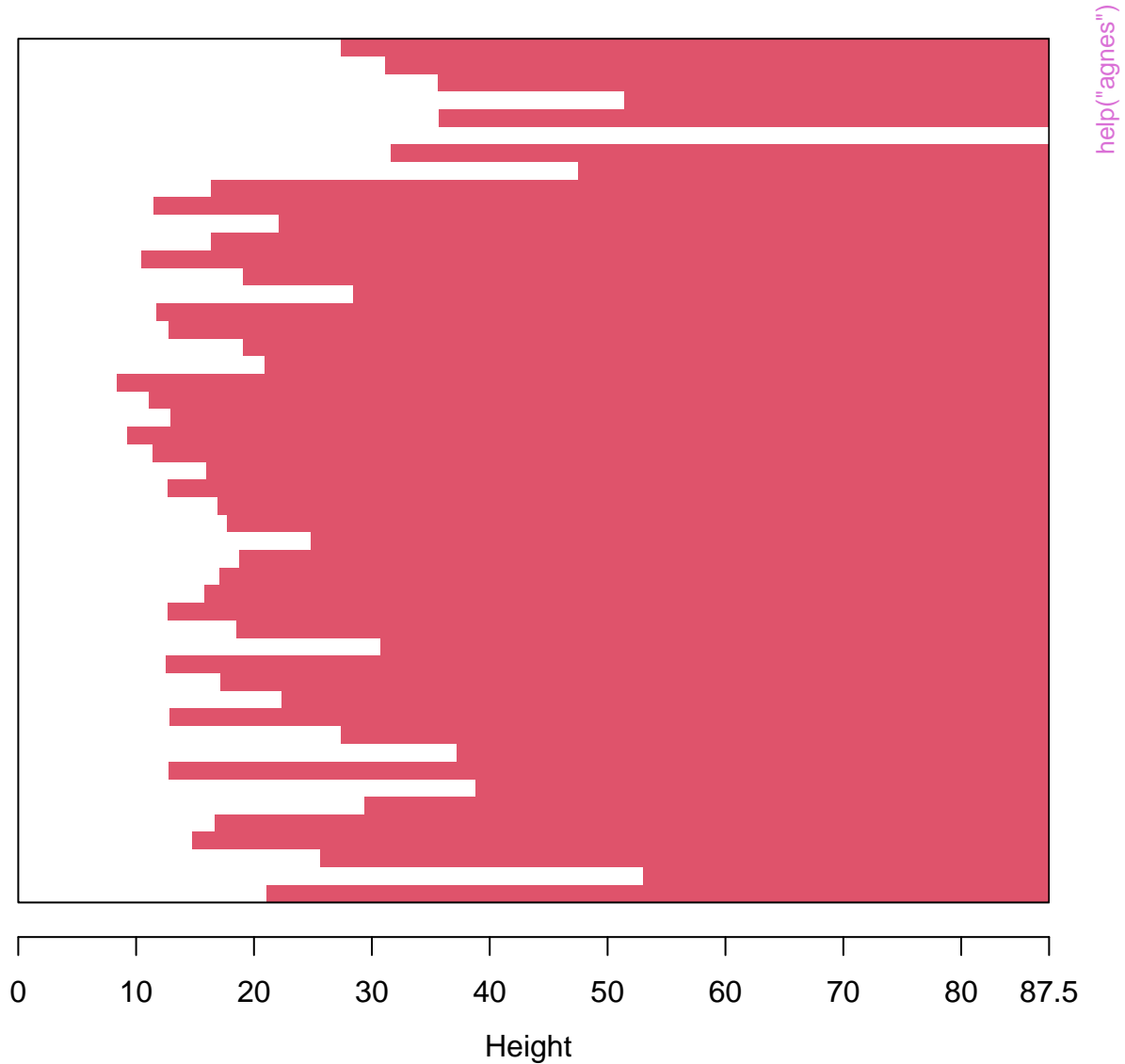
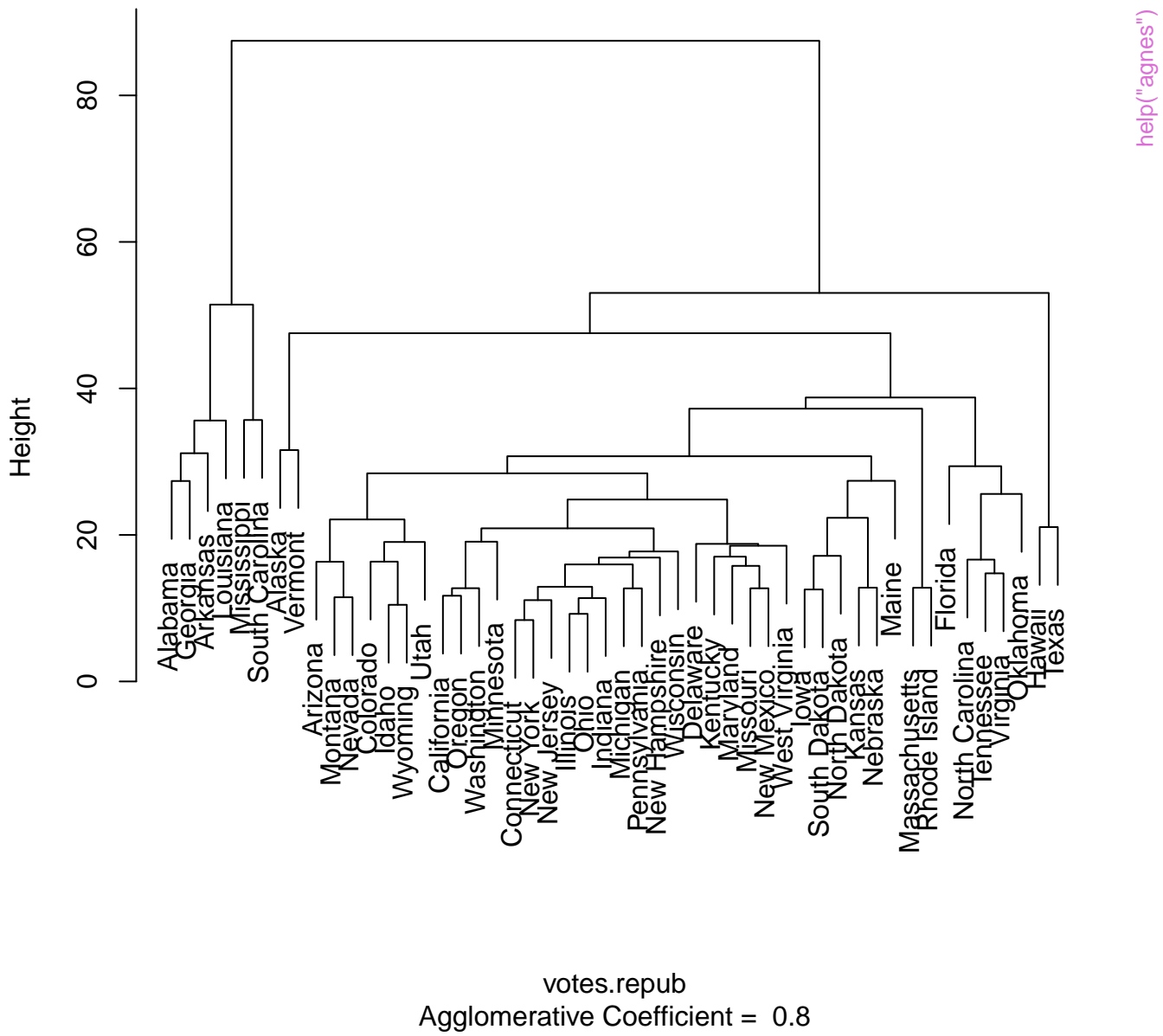


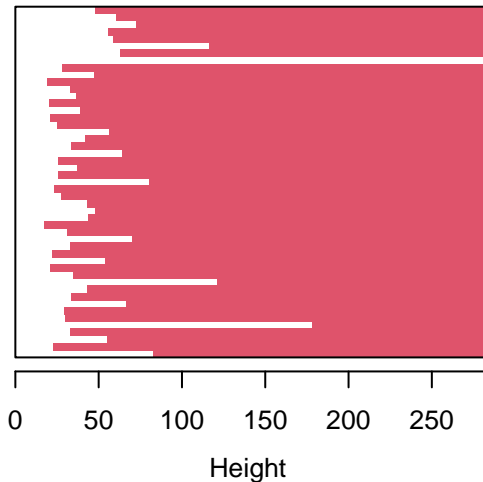
Banner of `agnes(x = votes.repub, metric = "manhattan", stand = TF`



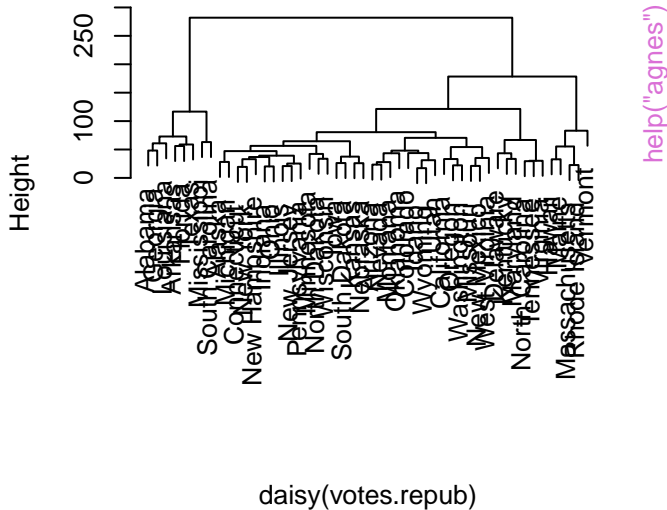
Dendrogram of agnes(x = votes.repub, metric = "manhattan", stand = TRUE)



Banner of `agnes(x = daisy(votes.repub)`, `diss = "complete"`)

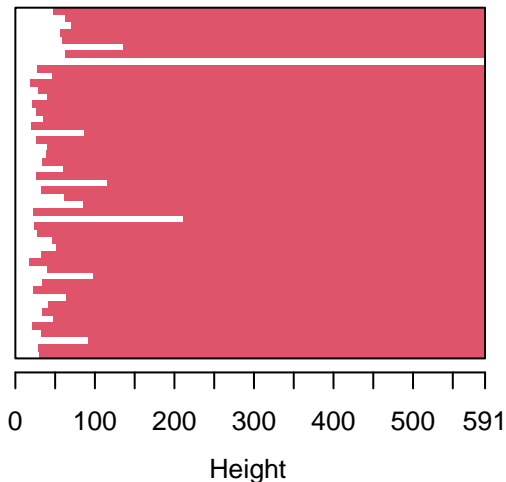


Agglomerative Coefficient = 0.88

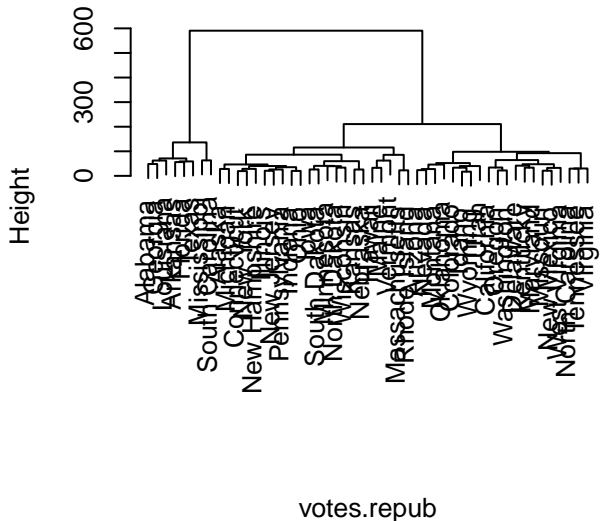


`daisy(votes.repub)`
Agglomerative Coefficient = 0.88

Banner of `agnes(x = votes.repub, n of agnes(x = votes.repub, method = "flexible", n = 1000, seed = 12345))`

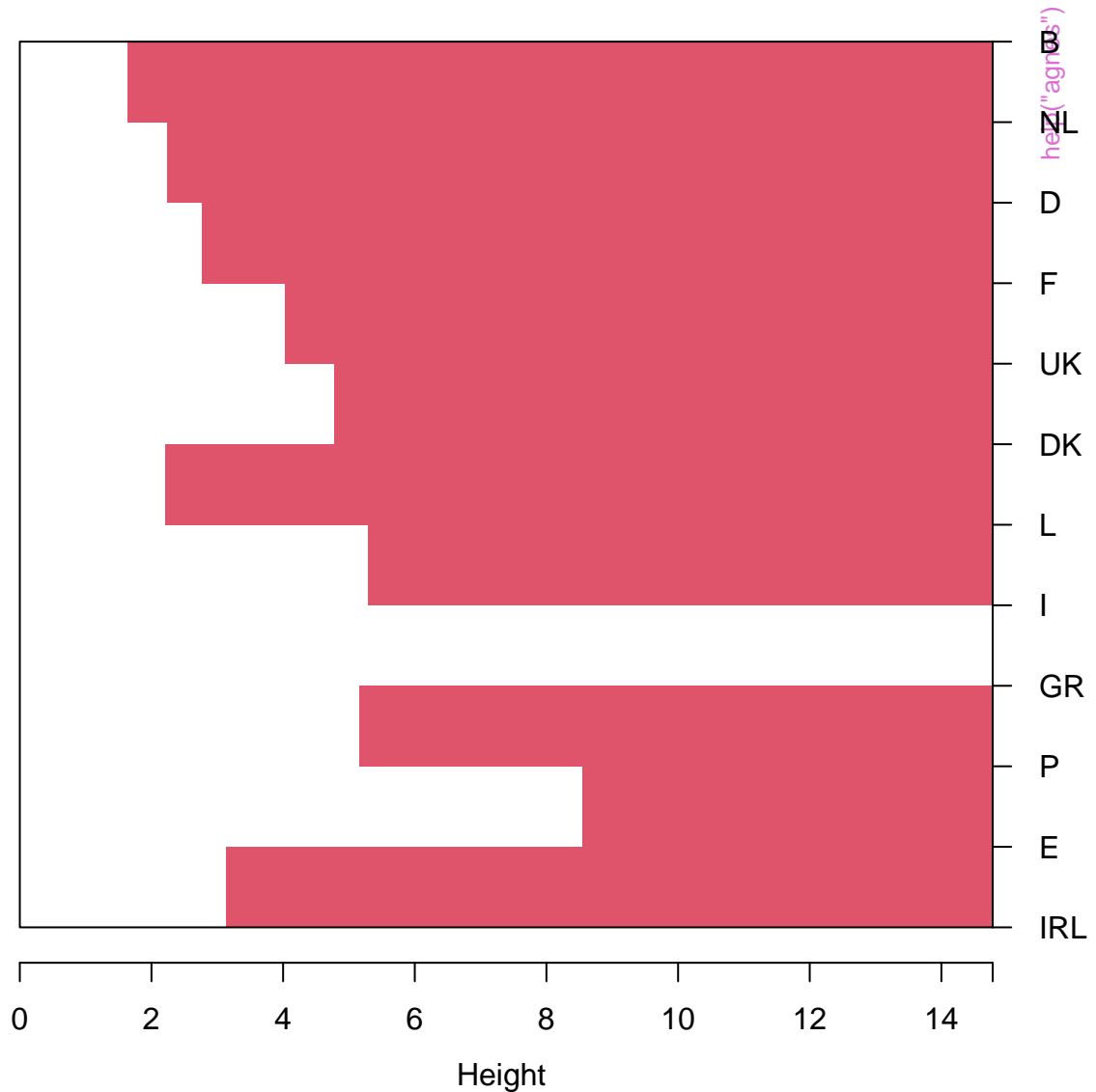


Agglomerative Coefficient = 0.94



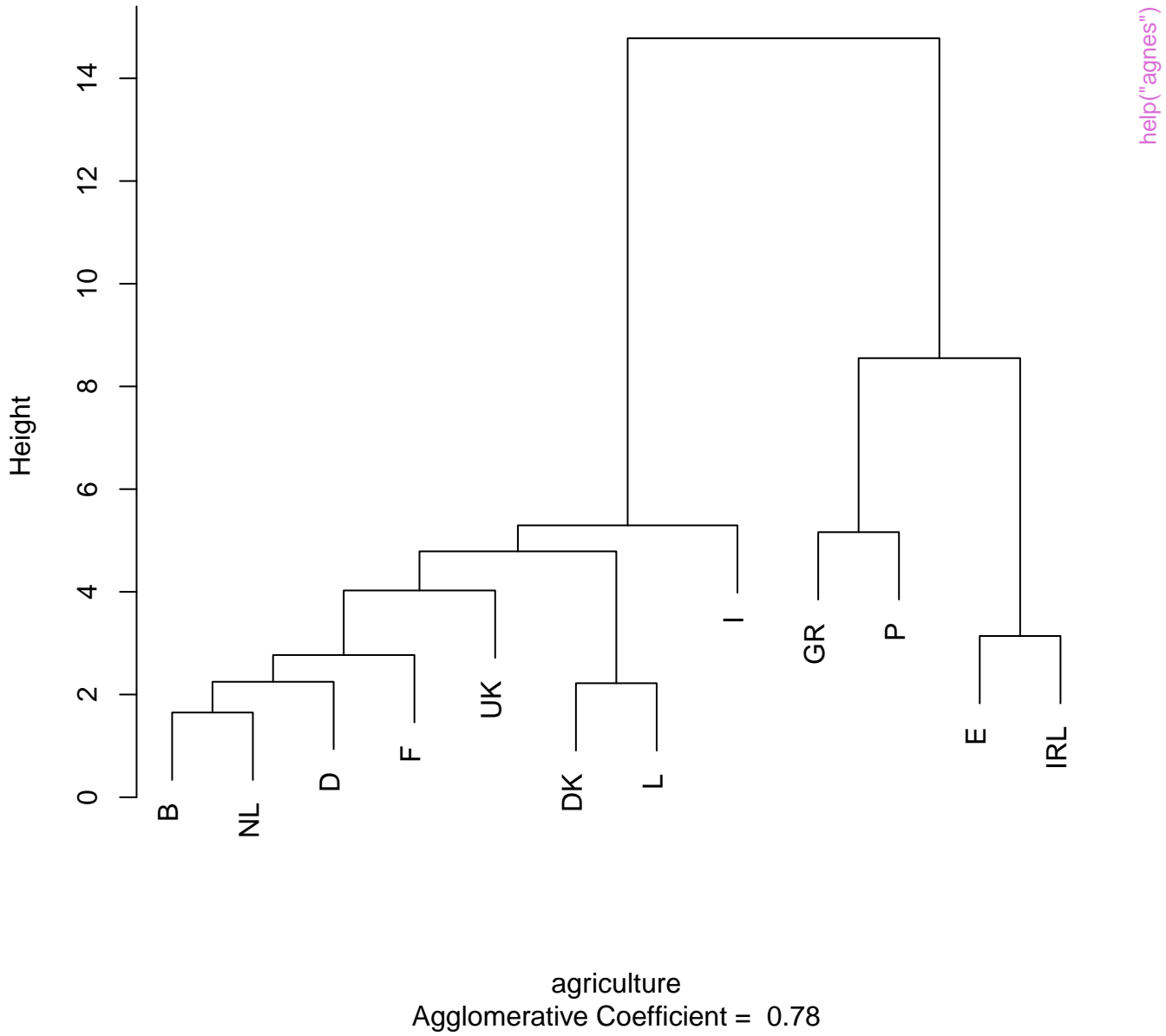
`votes.repub`
Agglomerative Coefficient = 0.94

Banner of agnes(x = agriculture)

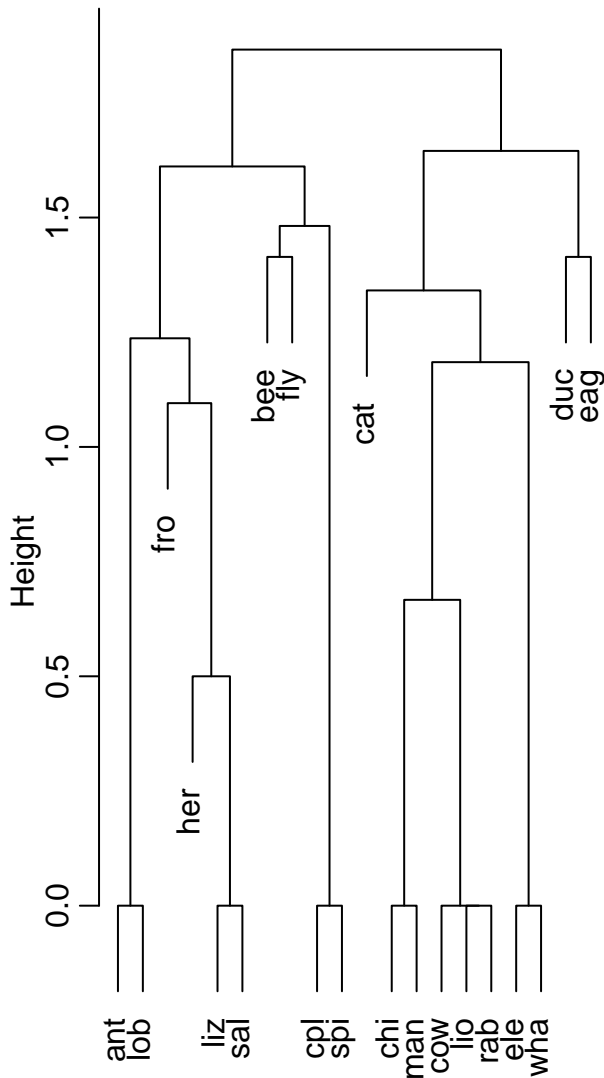


Agglomerative Coefficient = 0.78

Dendrogram of agnes(x = agriculture)



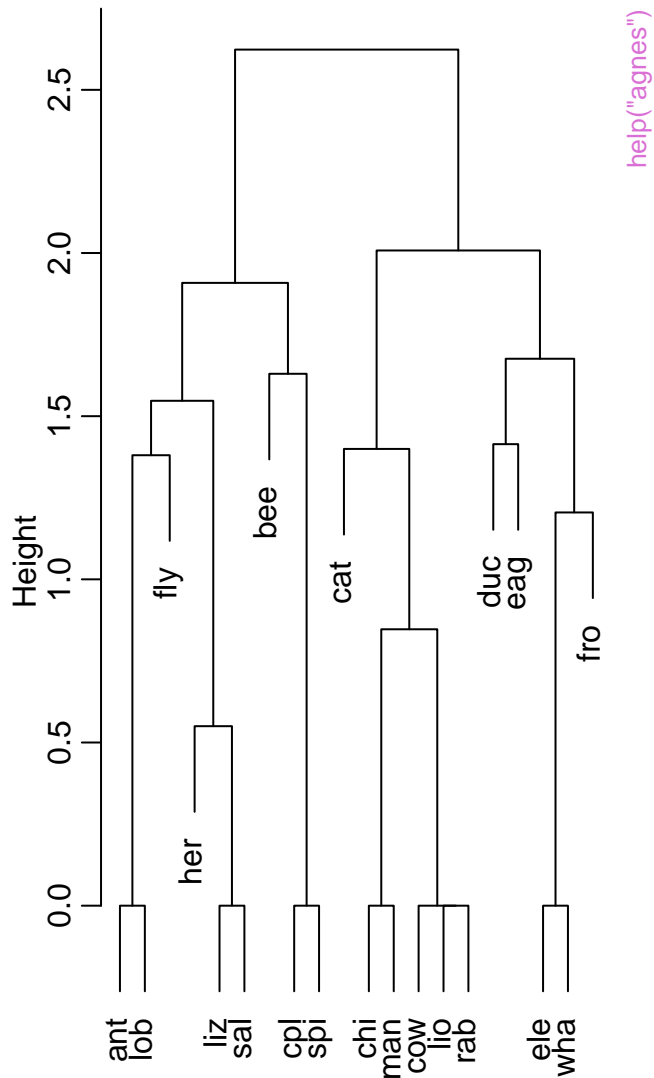
Dendrogram of agnes(x = animals)



animals

Agglomerative Coefficient = 0.77

Dendrogram of agnes(x = animals, method = "gaver

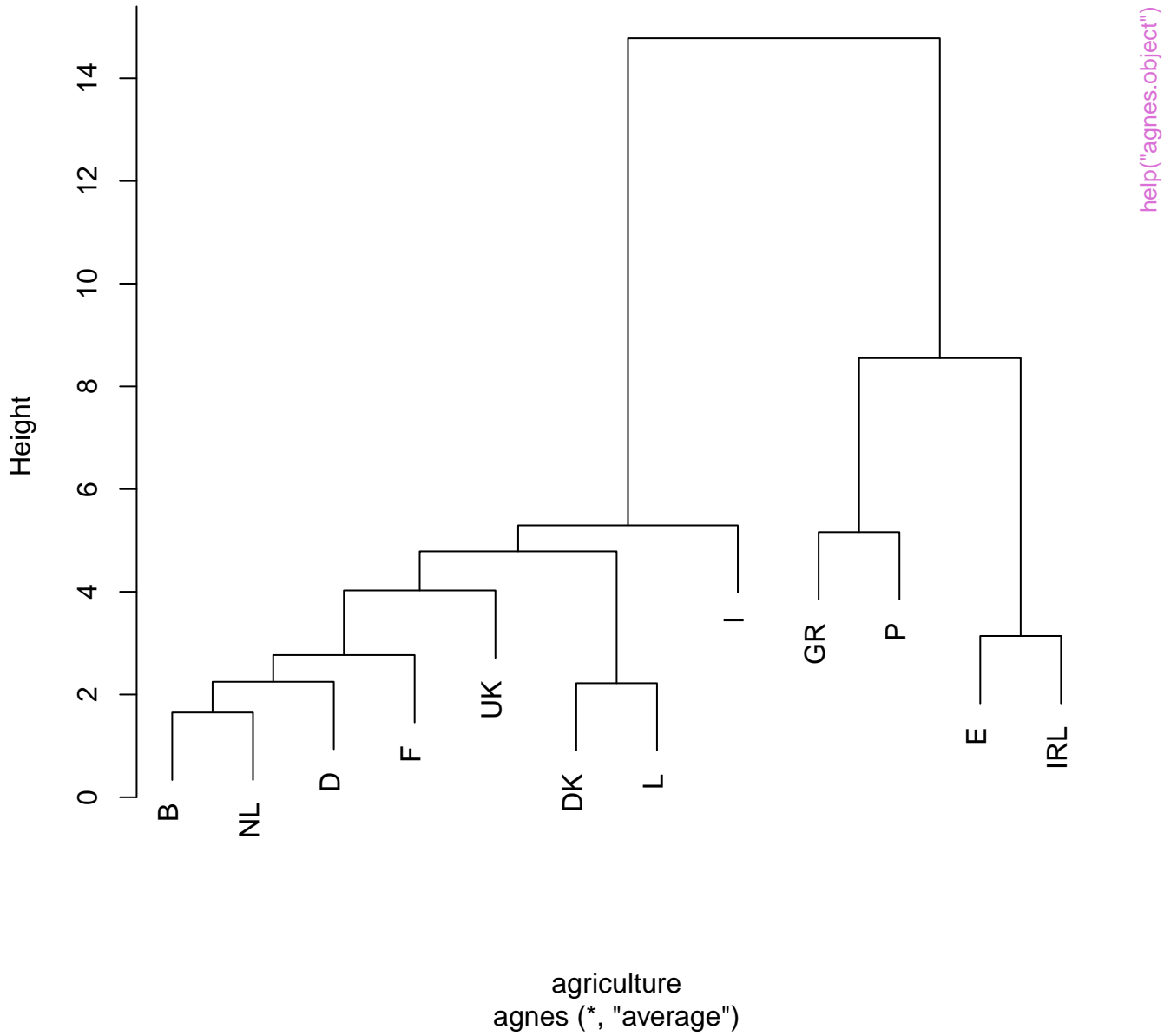


animals

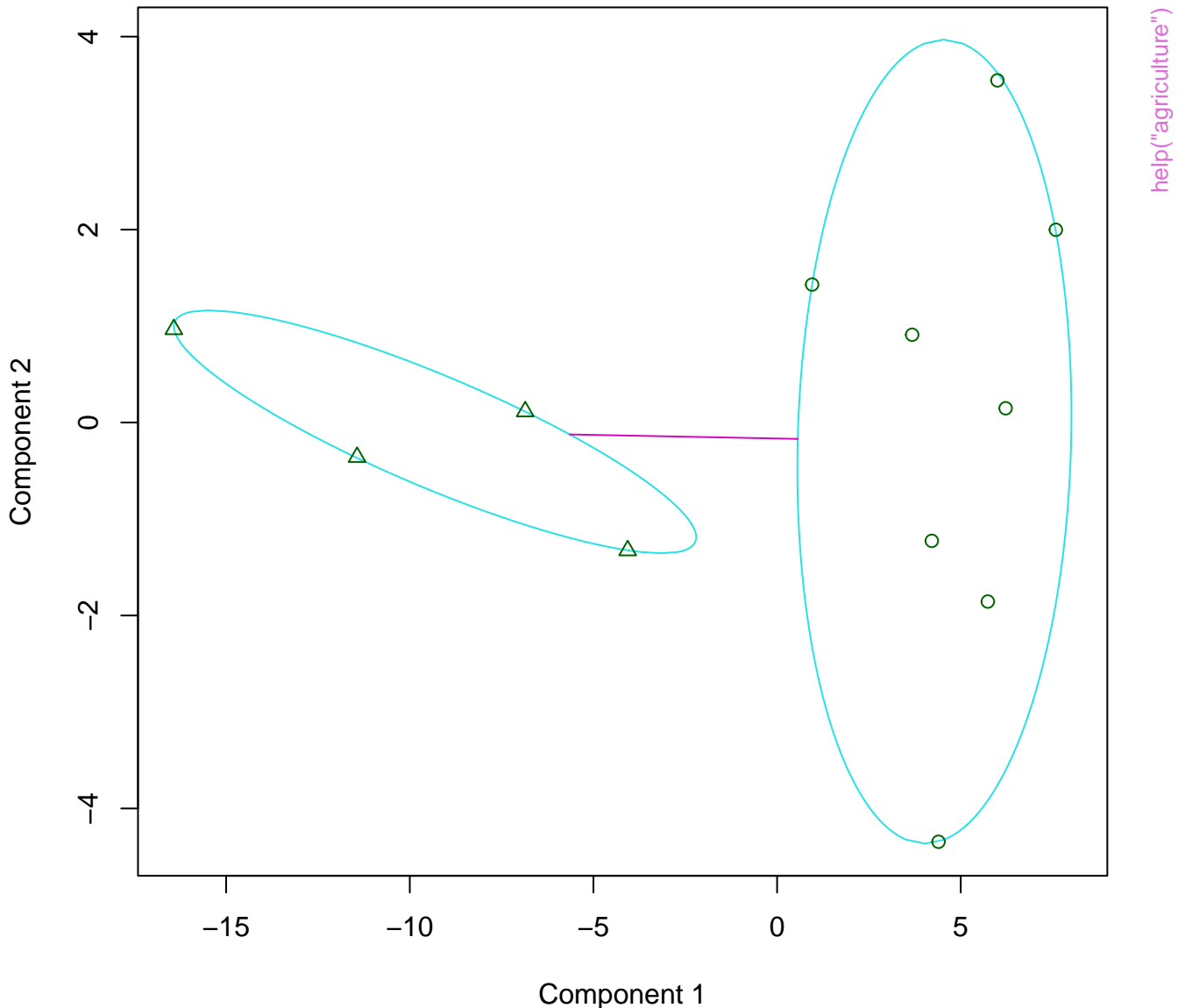
Agglomerative Coefficient = 0.83

help("agnes")

Dendrogram of agnes(x = agriculture)



clusplot(pam(x = agriculture, k = 2))



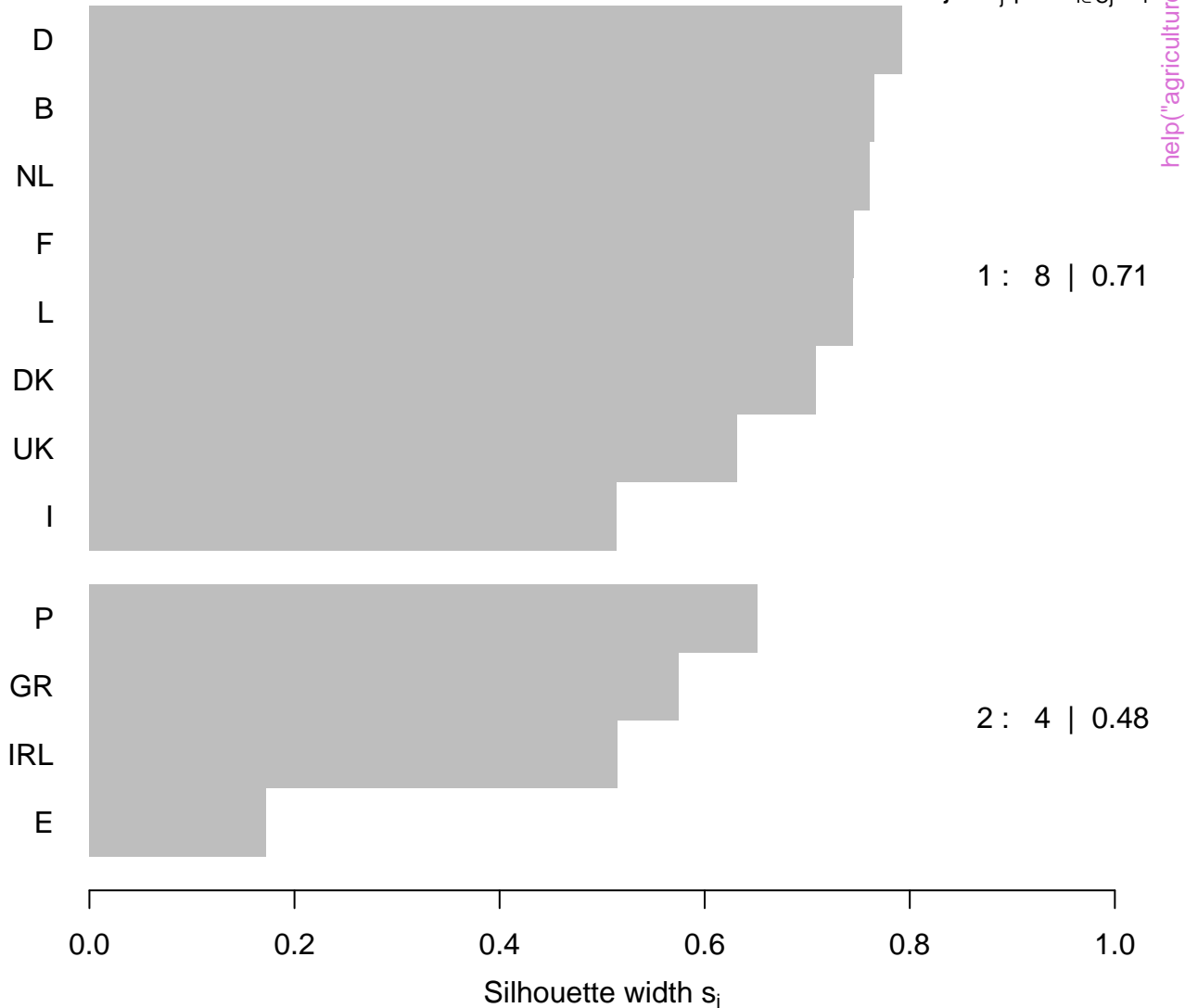
These two components explain 100 % of the point variability.

Silhouette plot of pam(x = agriculture, k = 2)

n = 12

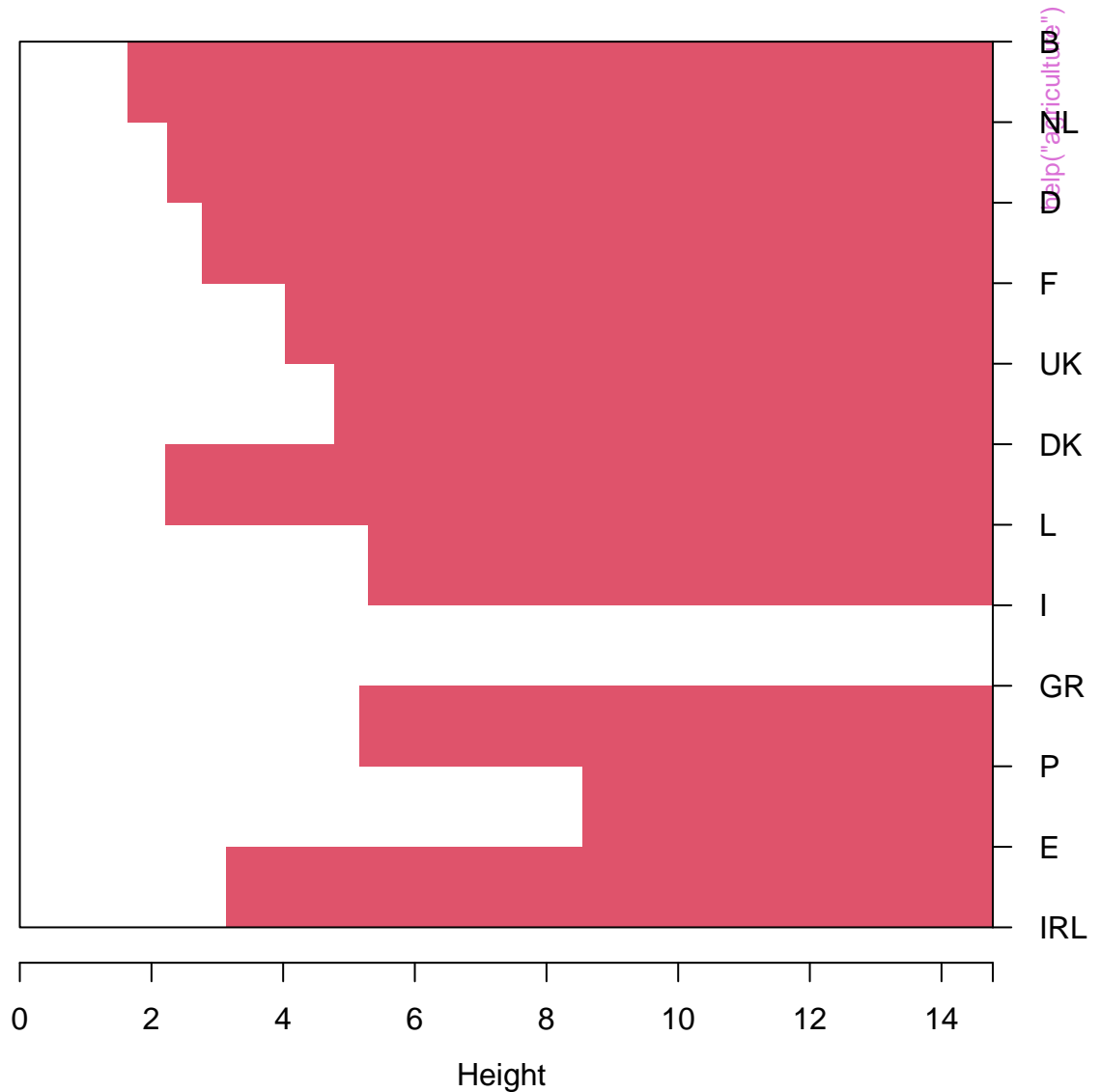
2 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$



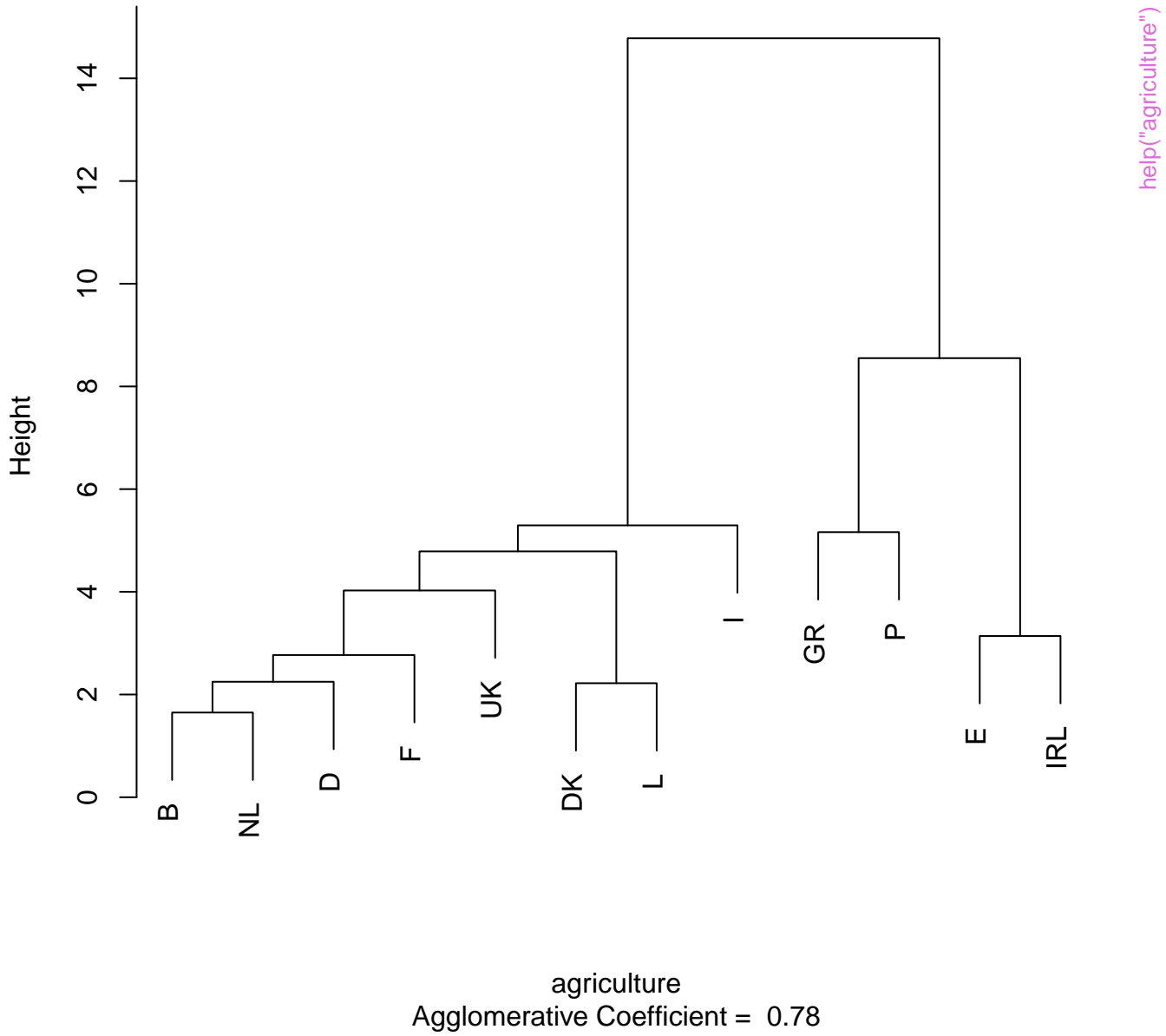
Average silhouette width : 0.63

Banner of agnes(x = agriculture)

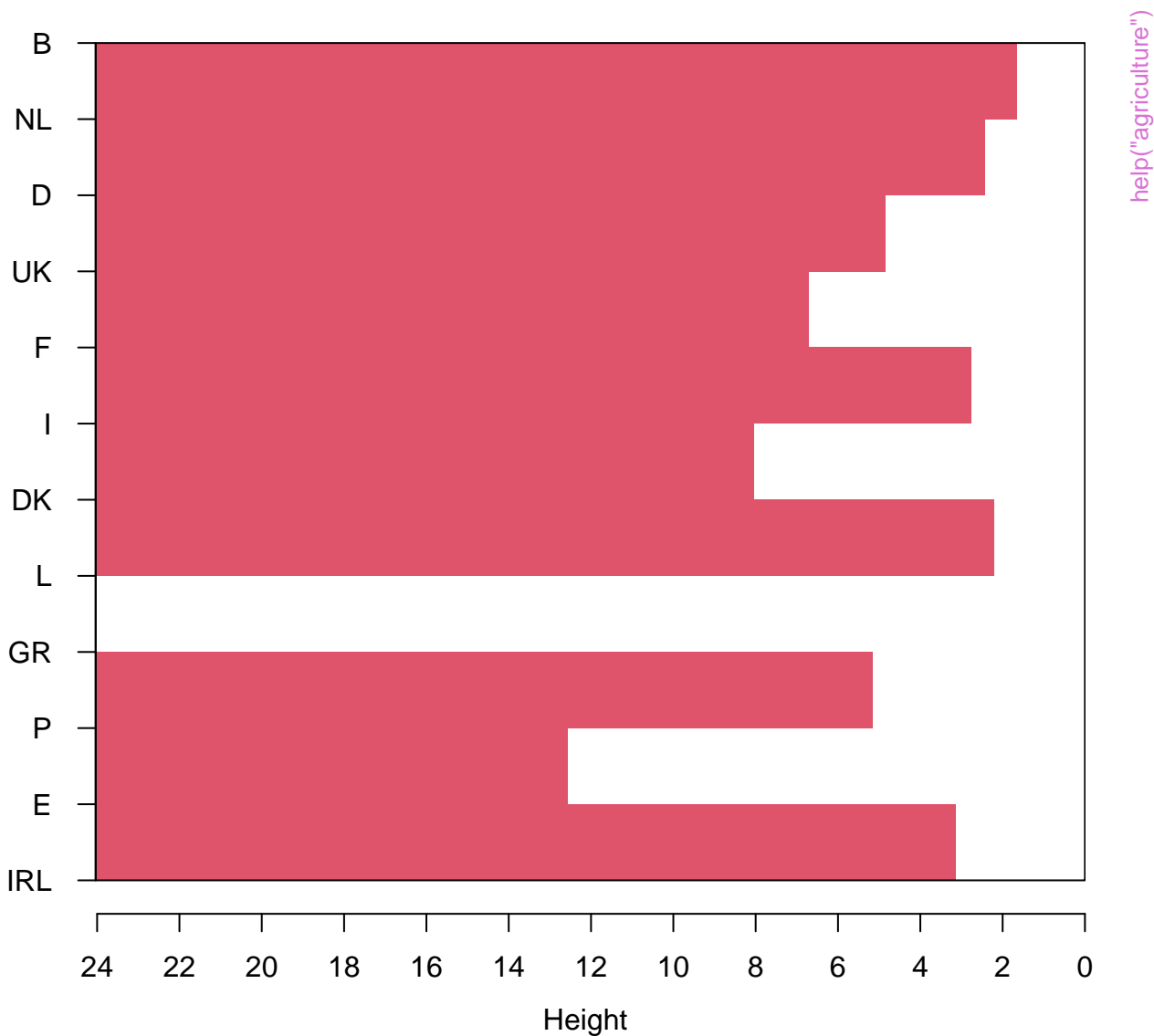


Agglomerative Coefficient = 0.78

Dendrogram of agnes(x = agriculture)

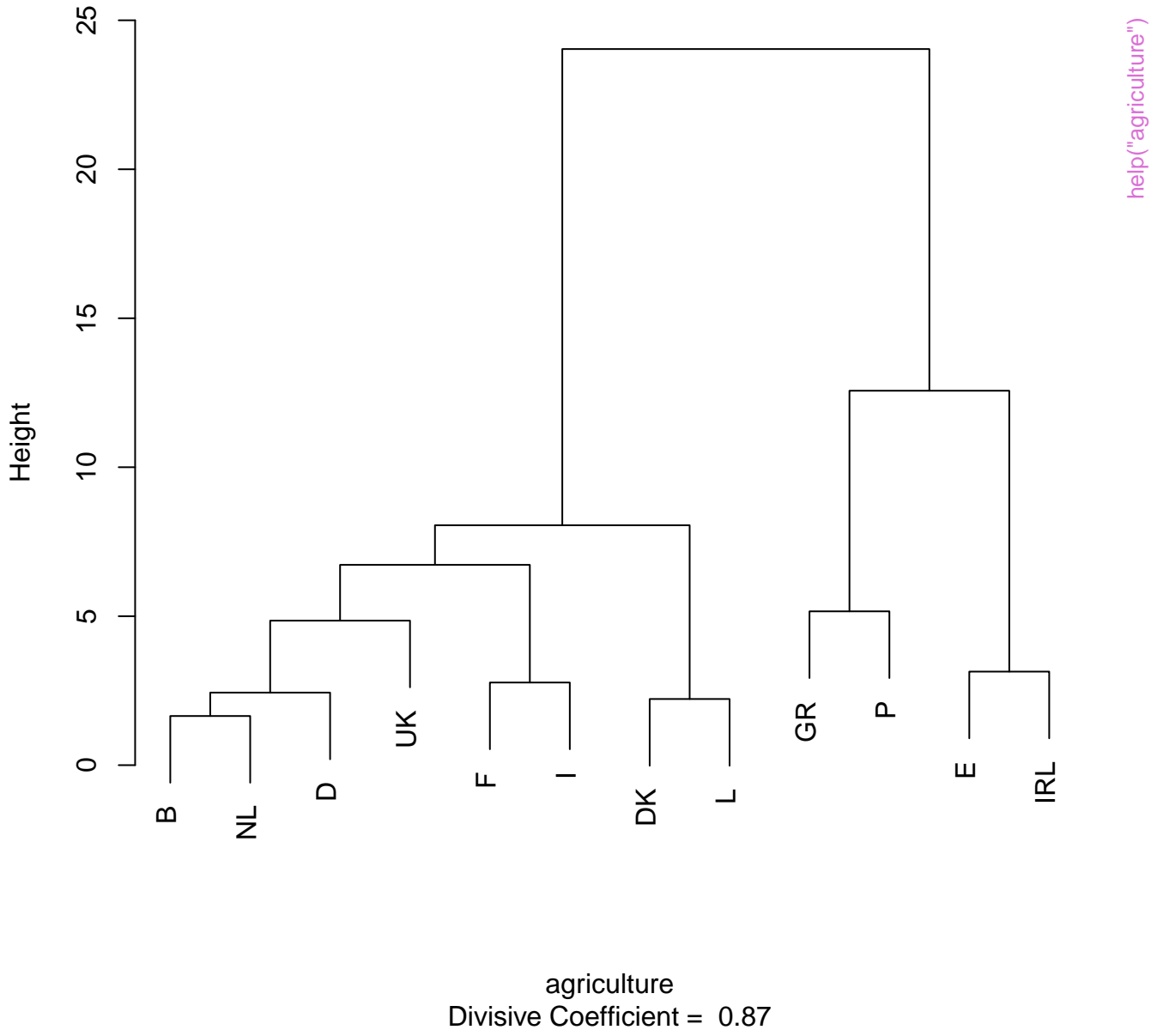


Banner of diana(x = agriculture)

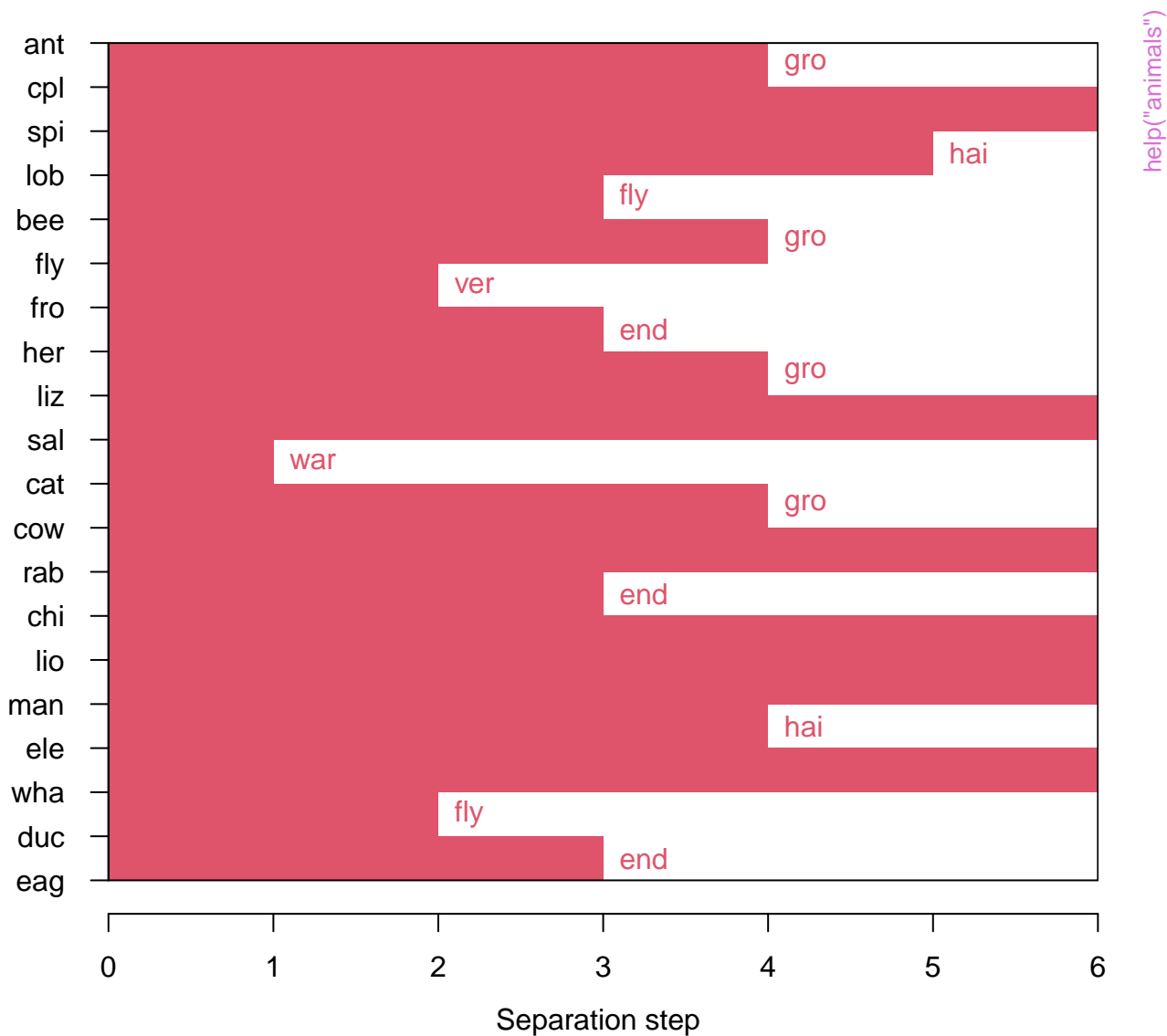


Divisive Coefficient = 0.87

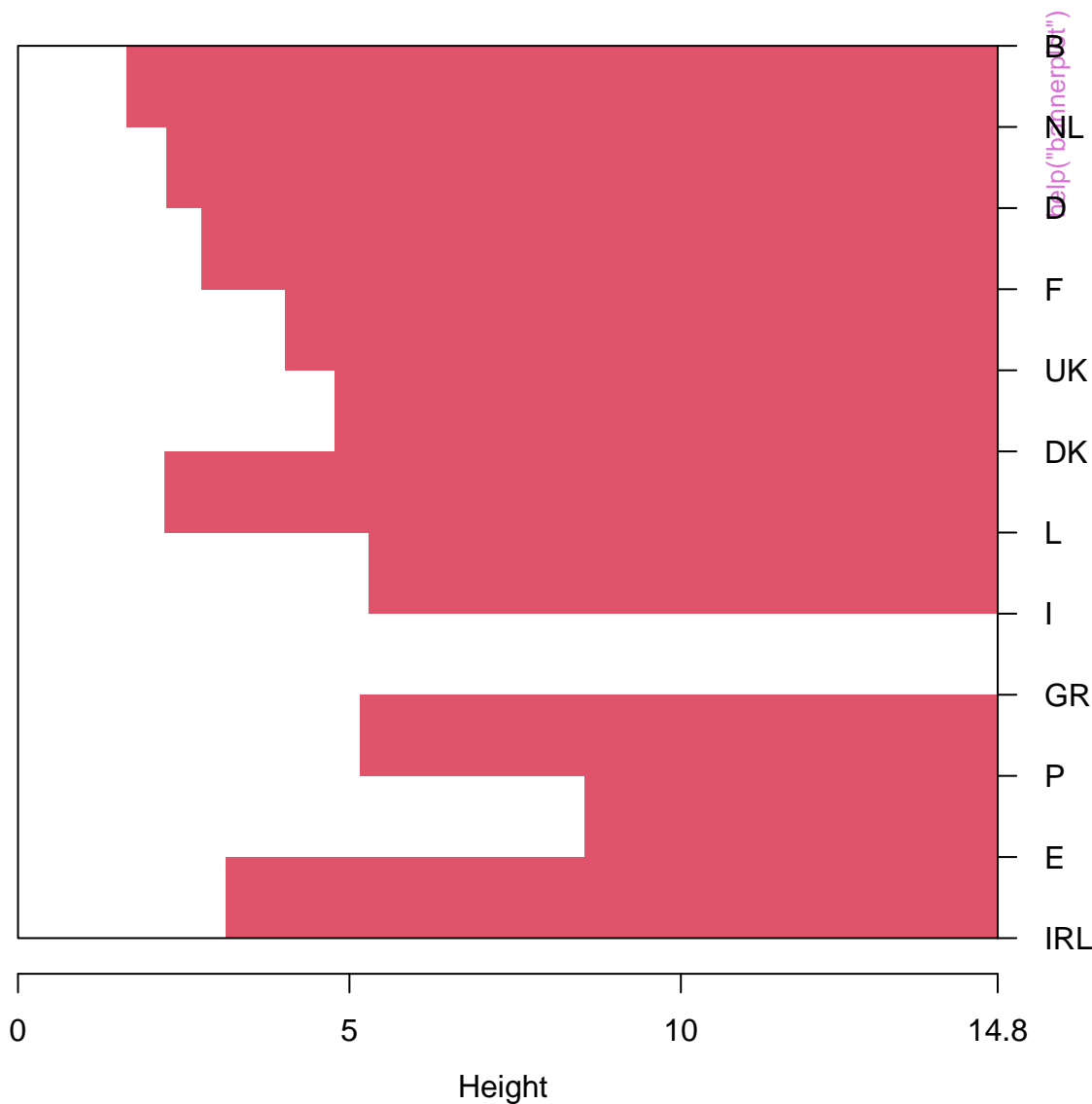
Dendrogram of diana(x = agriculture)

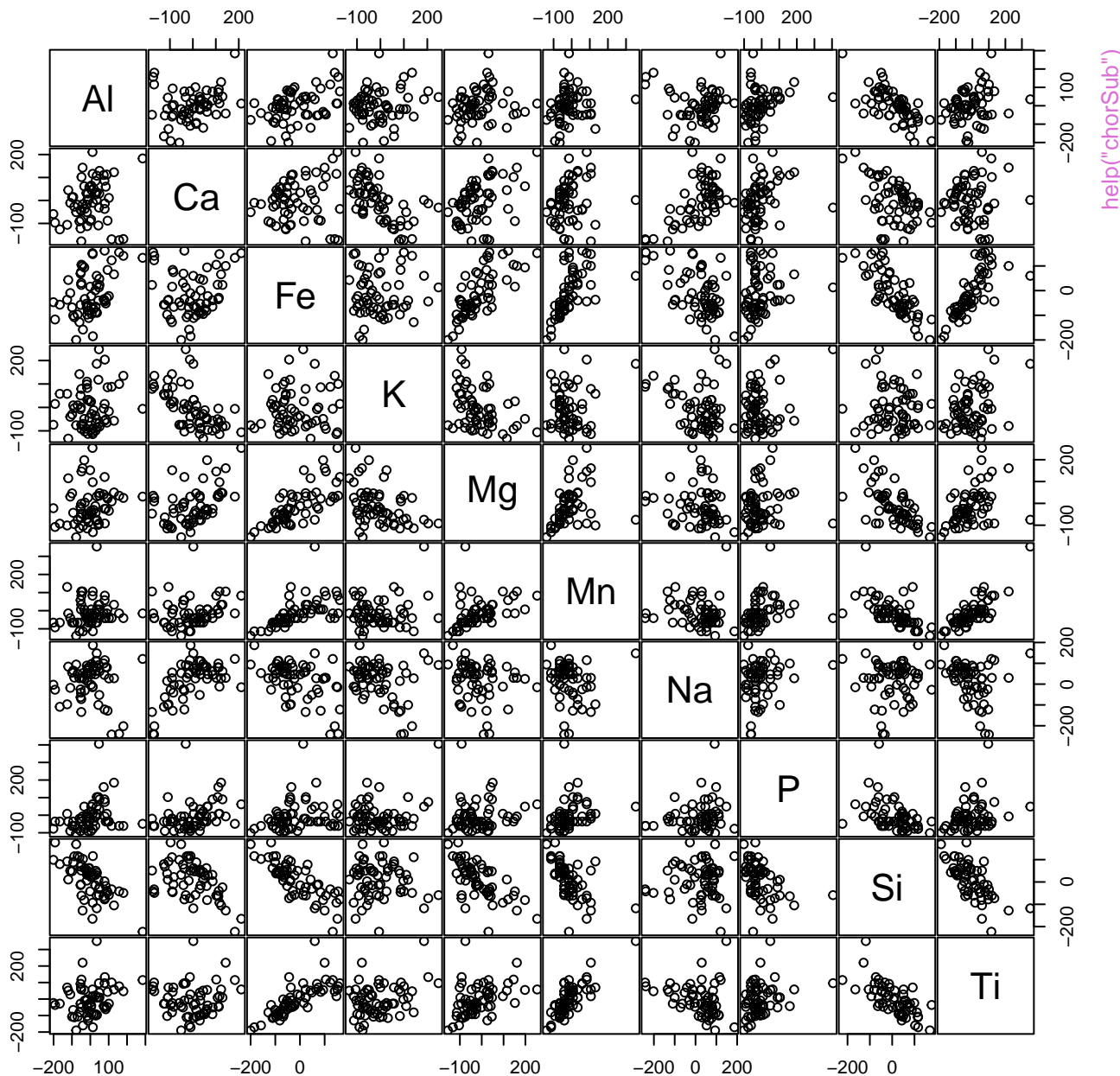


Banner of mona(x = animals)

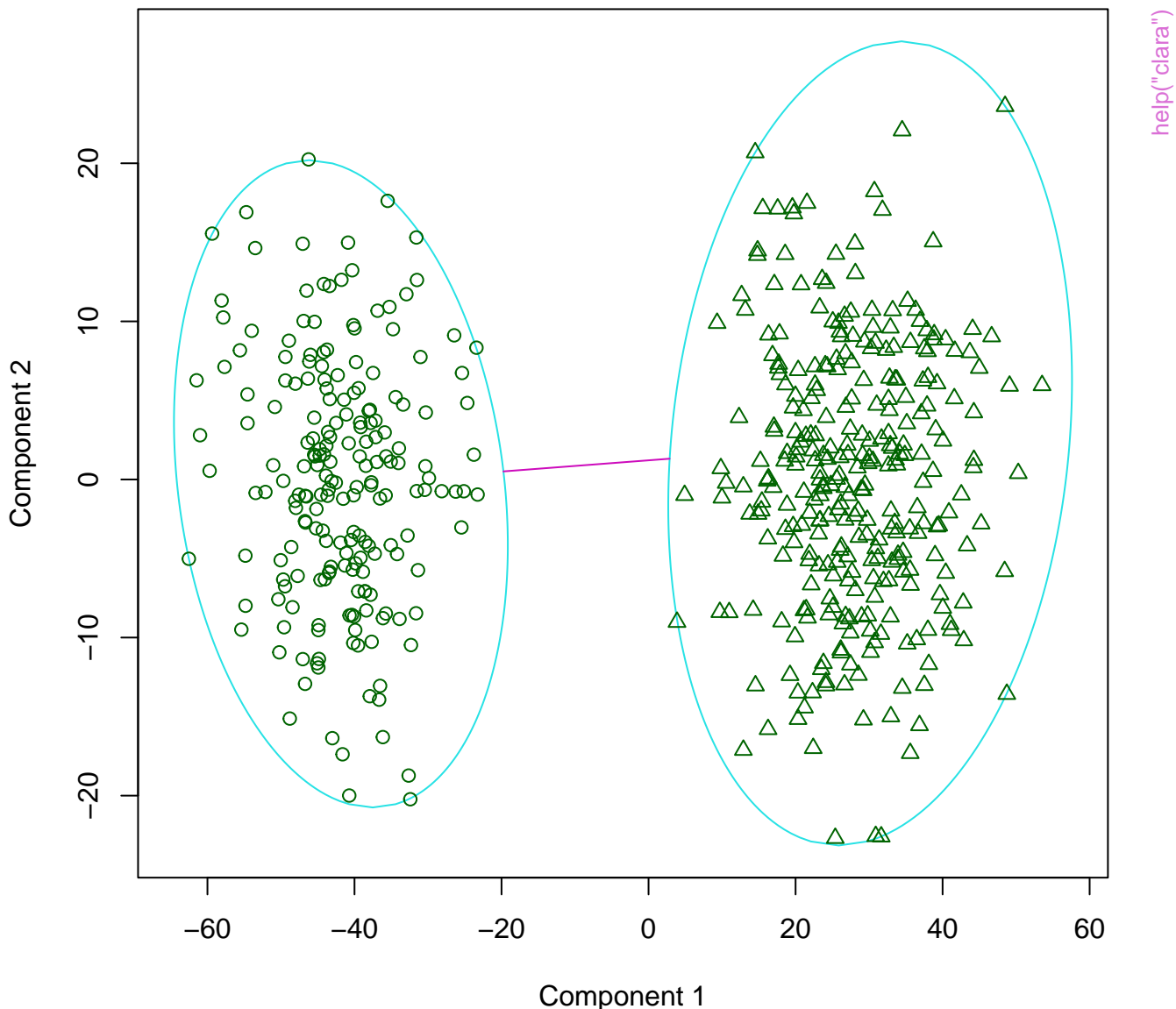


Bannerplot





clusplot(clara(x = x, k = 2, samples = 50))



These two components explain 100 % of the point variability.

help("clara")

Silhouette plot of clara(x = x, k = 2, samples = 50)

n = 44

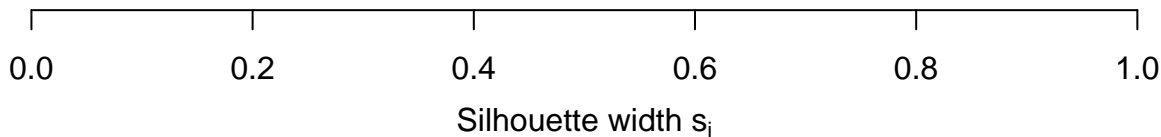
2 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$

help("clara")

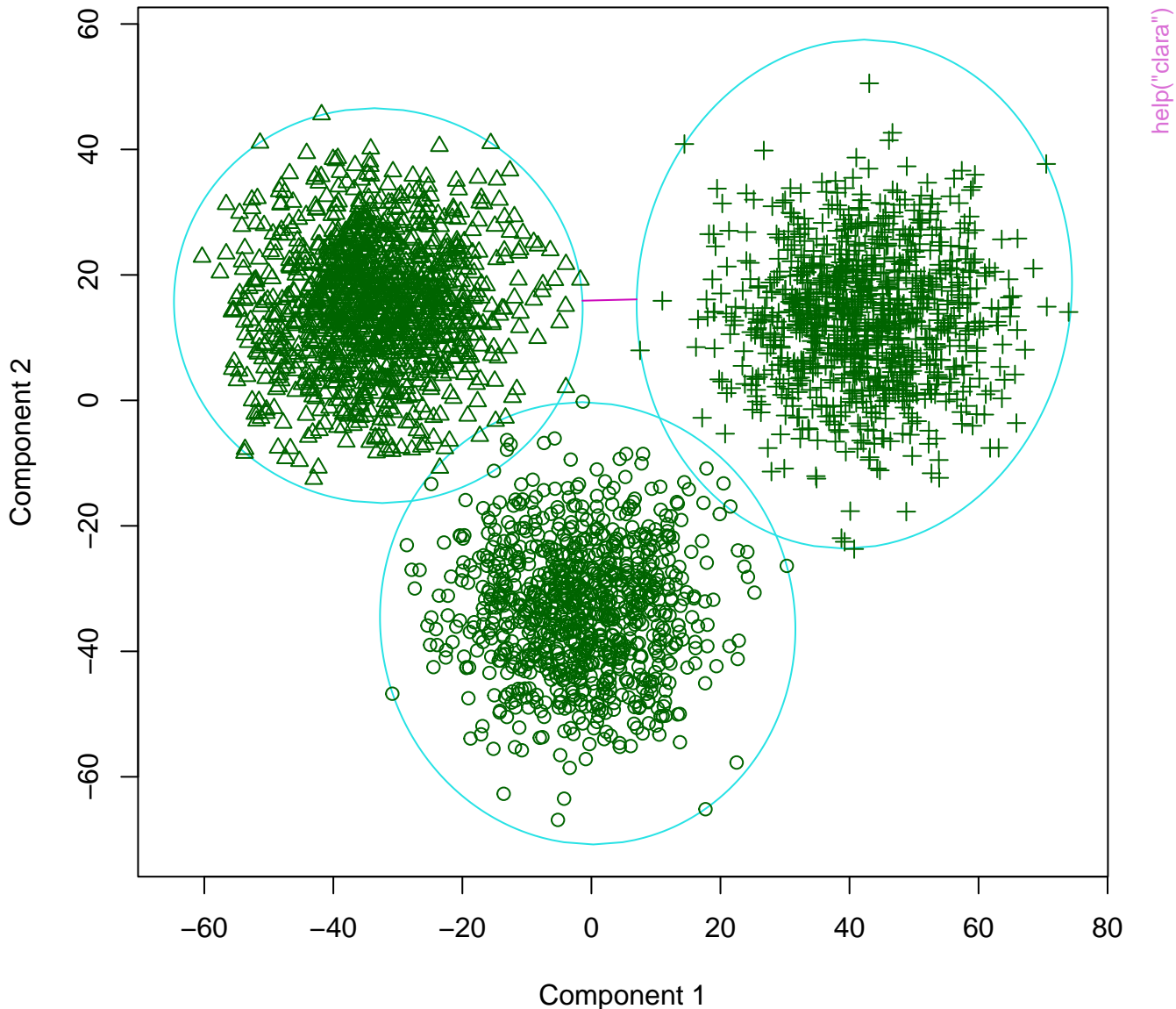
1 : 21 | 0.80

2 : 23 | 0.77



Average silhouette width : 0.78

clusplot(clara(x = xclara, k = 3))



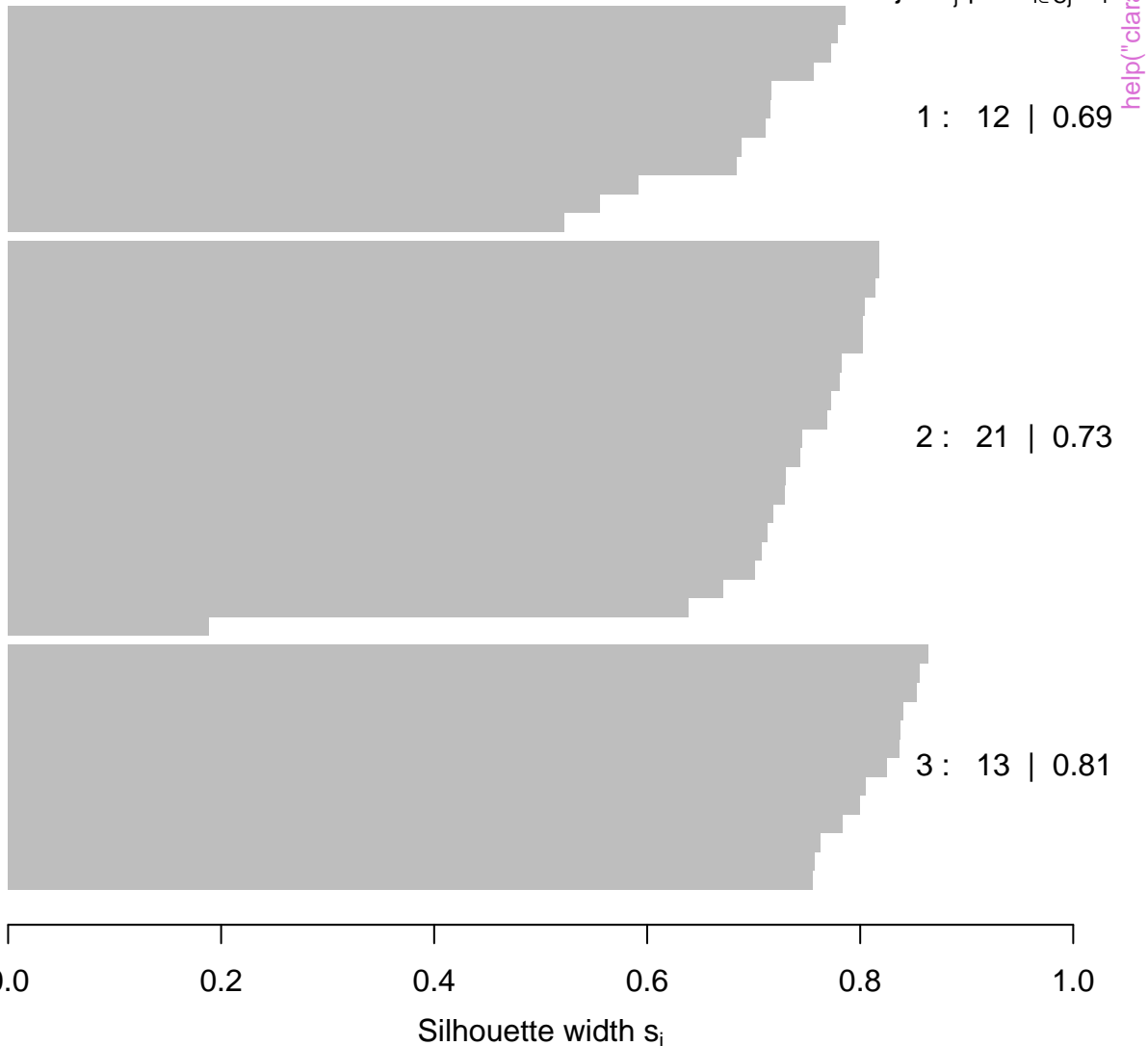
These two components explain 100 % of the point variability.

Silhouette plot of clara(x = xclara, k = 3)

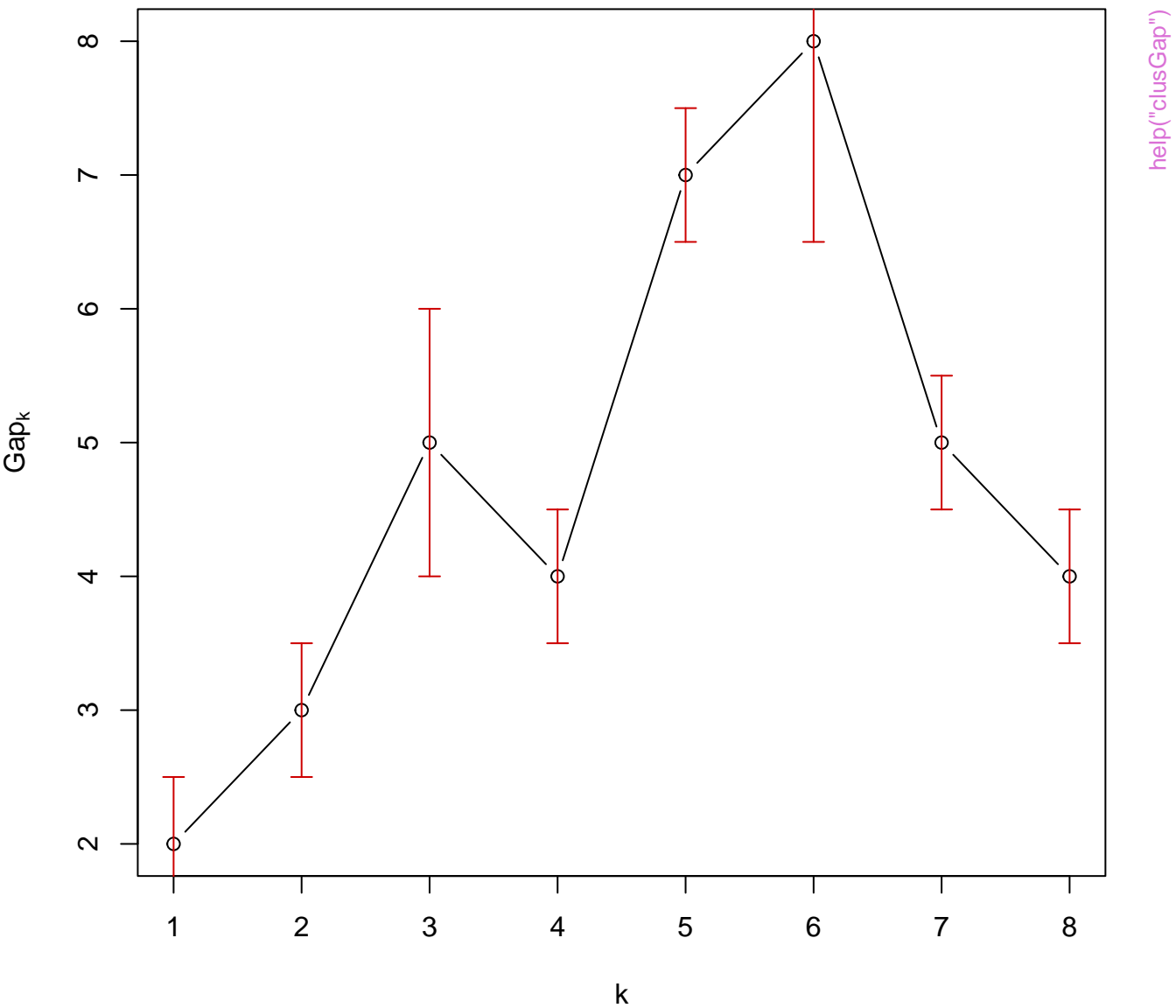
n = 46

3 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$

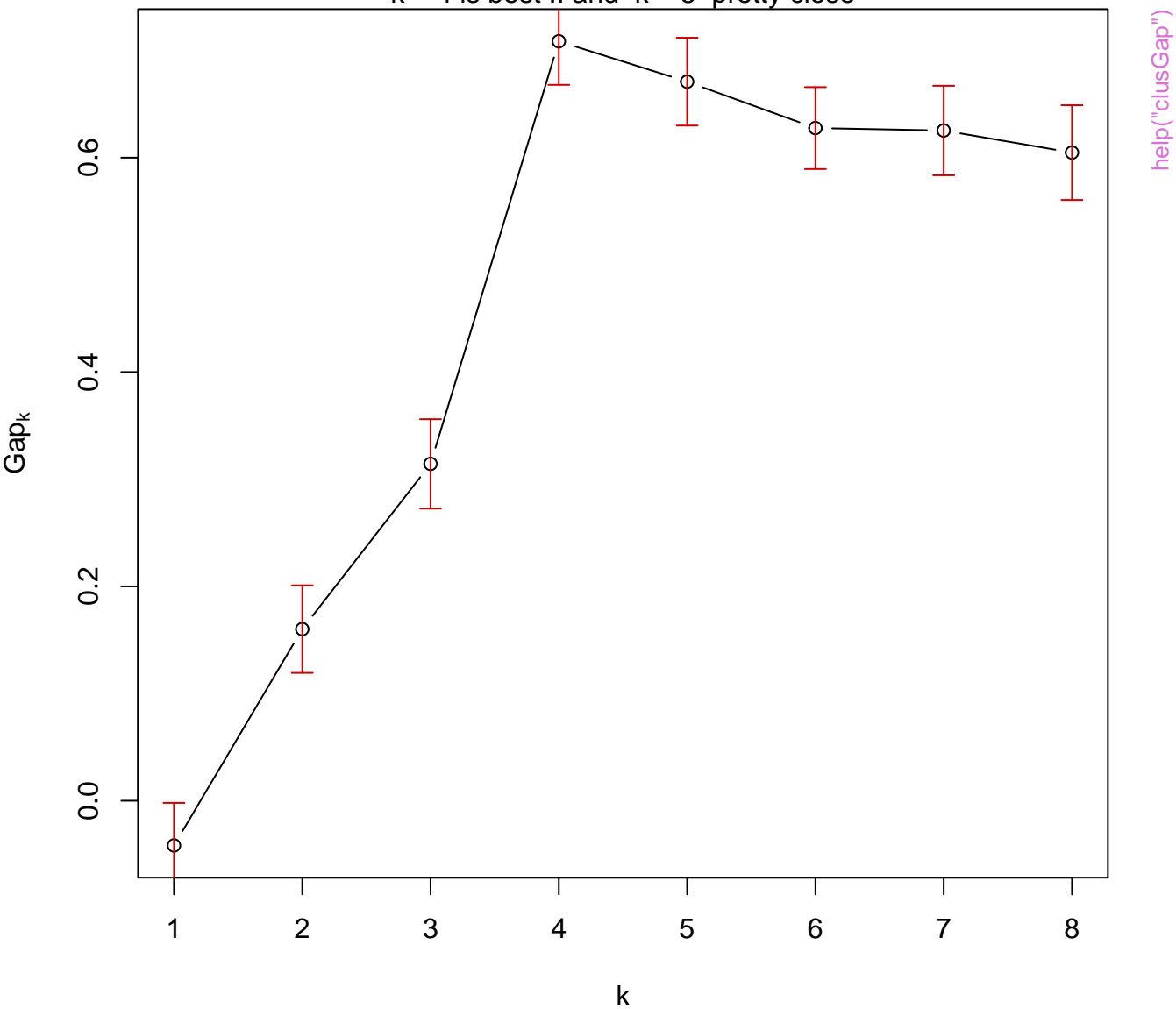


NULL

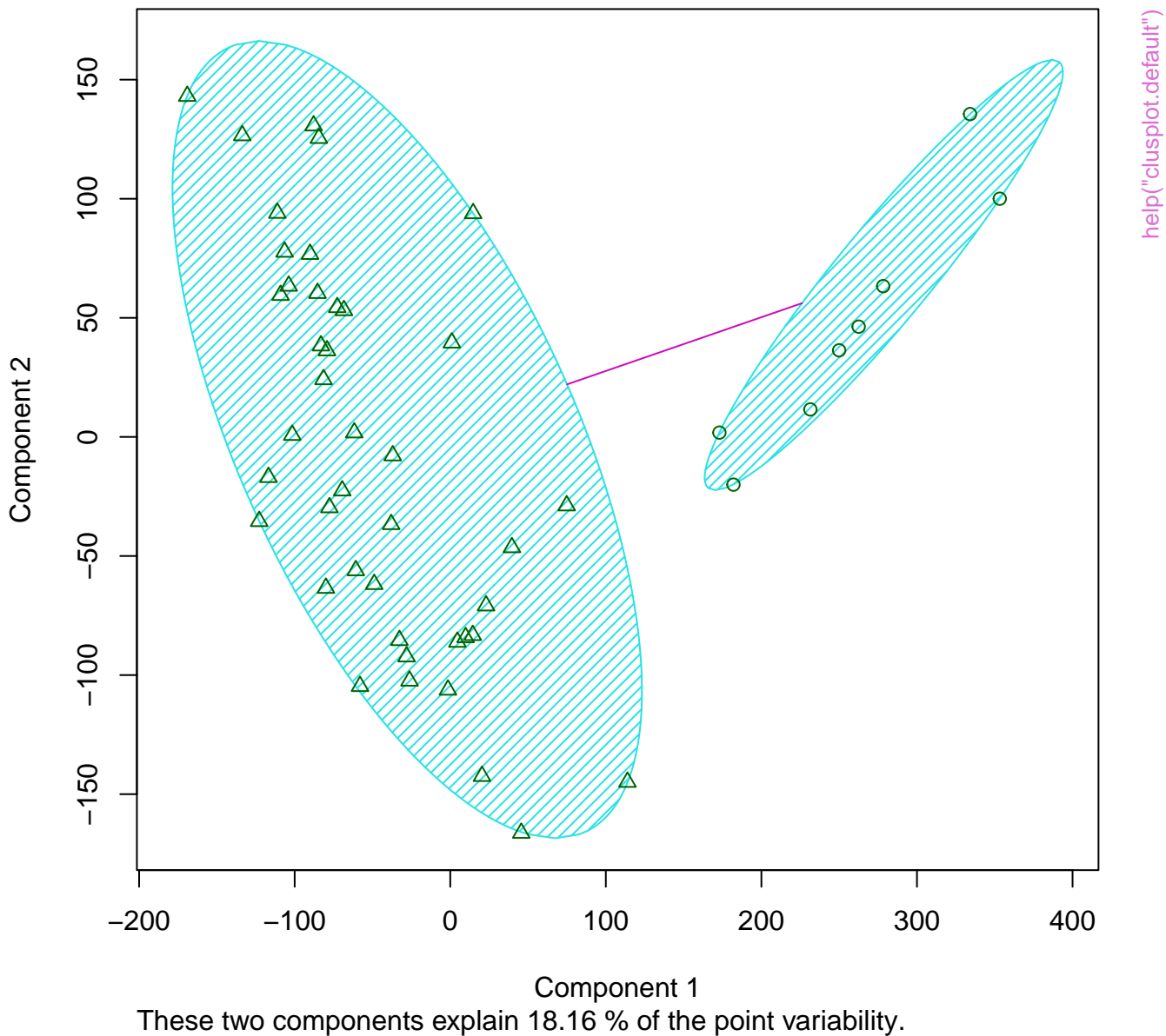


Gap statistic for the 'ruspini' data

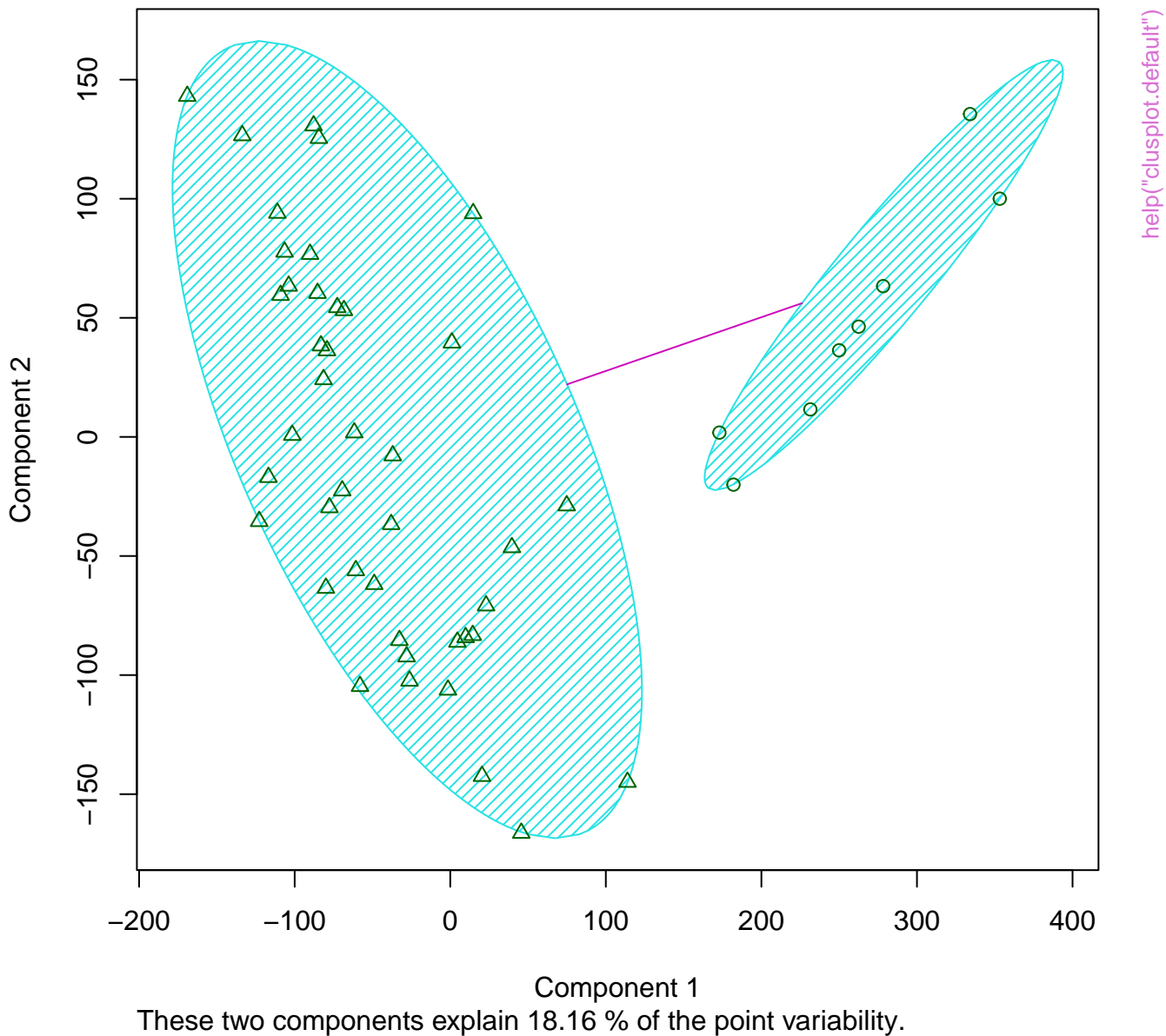
k = 4 is best .. and k = 5 pretty close



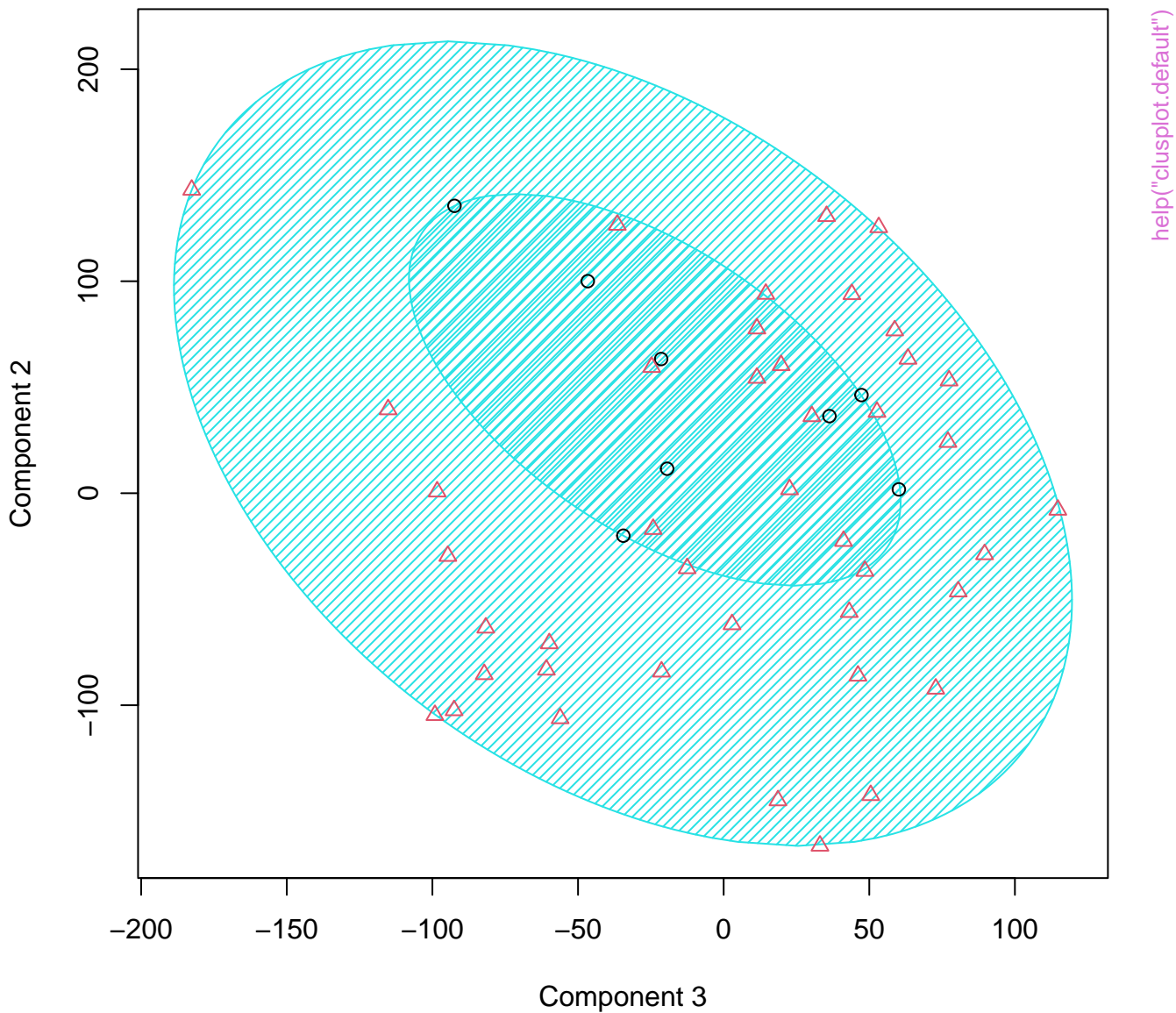
clusplot(pam(x = votes.diss, k = 2, diss = TRUE))



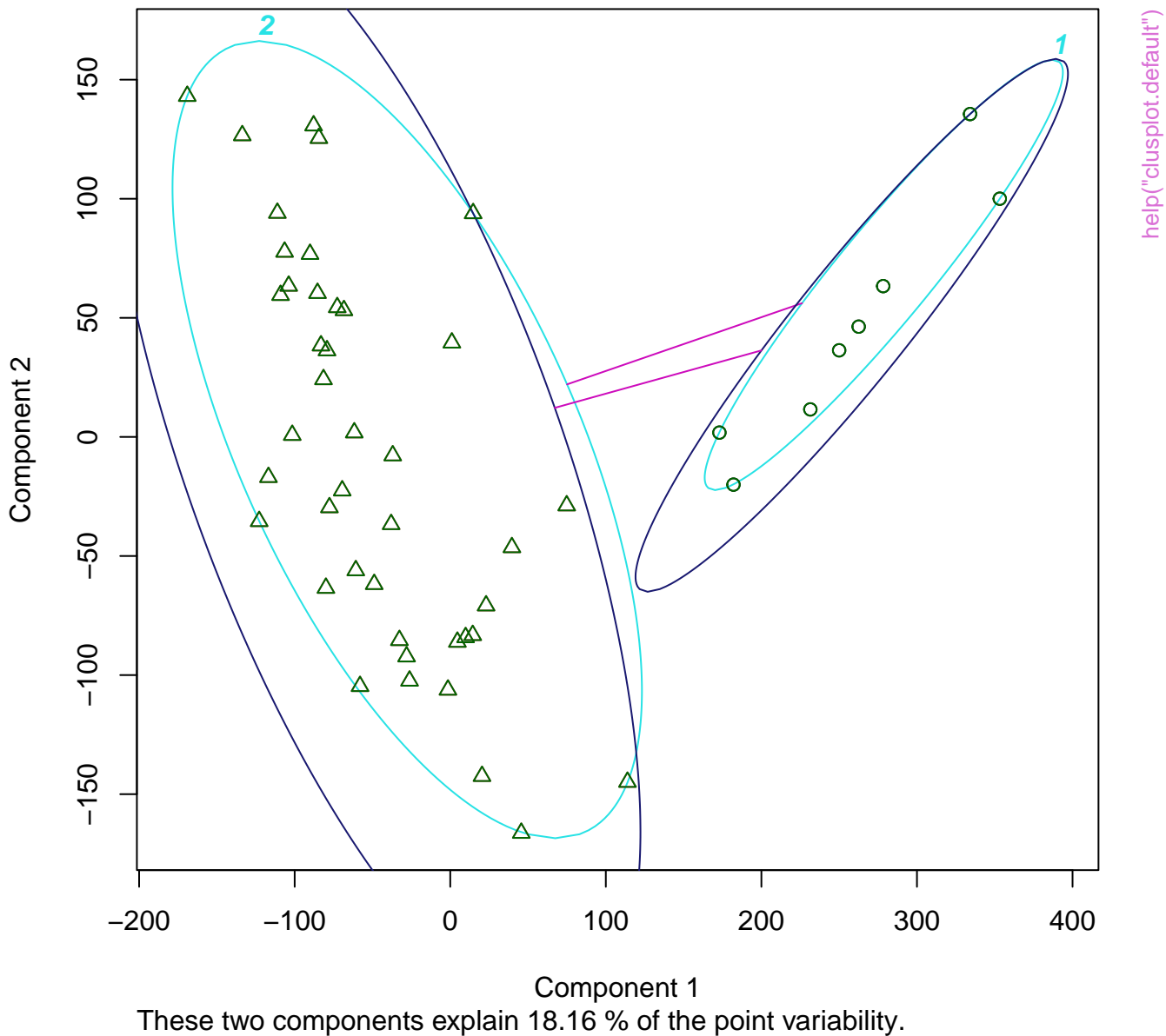
CLUSPLOT(votes.diss)



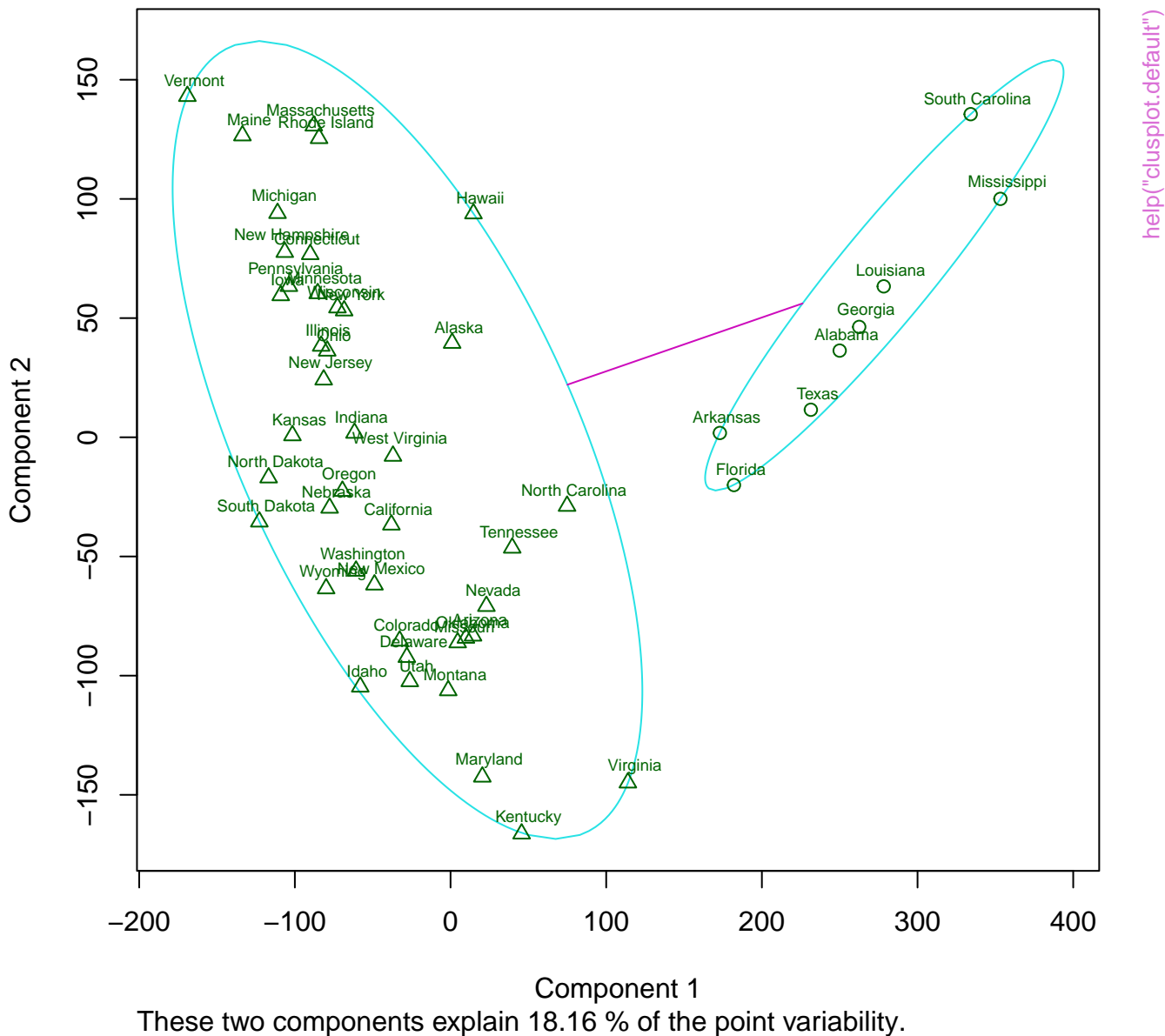
clusplot(pam(x = votes.diss, k = 2, diss = TRUE))



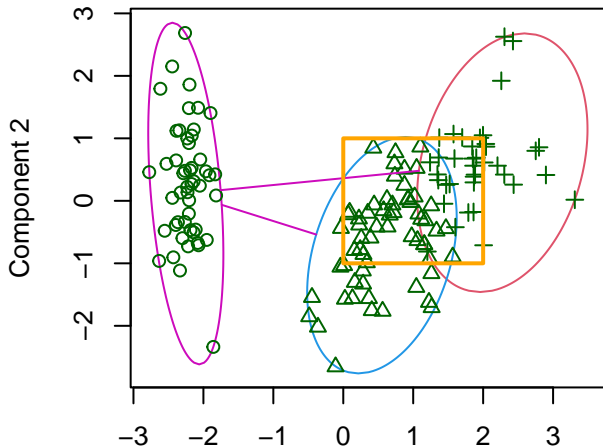
clusplot(pam(x = votes.diss, k = 2, diss = TRUE))



CLUSPLOT(votes.diss)



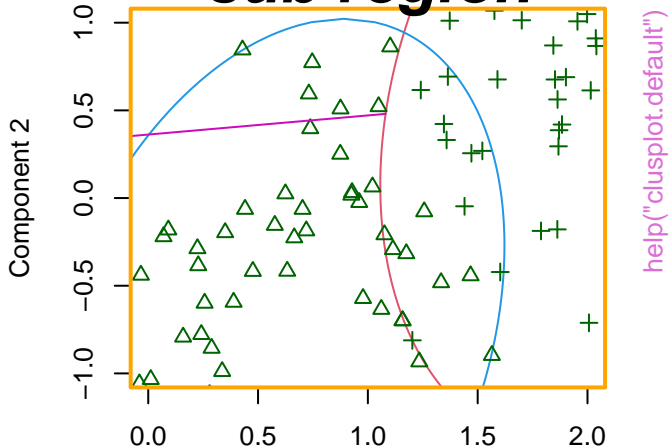
CLUSPLOT(iris.x)



Component 1

These two components explain 95.81 % of the

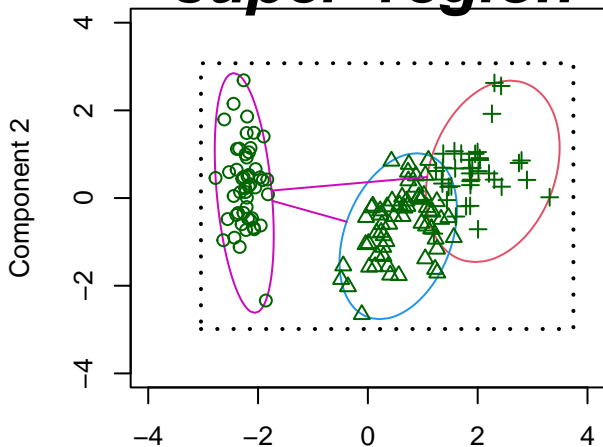
CLUSPLOT(iris.x)
sub region



Component 1

These two components explain 95.81 % of the

CLUSPLOT(iris.x)
'super' region

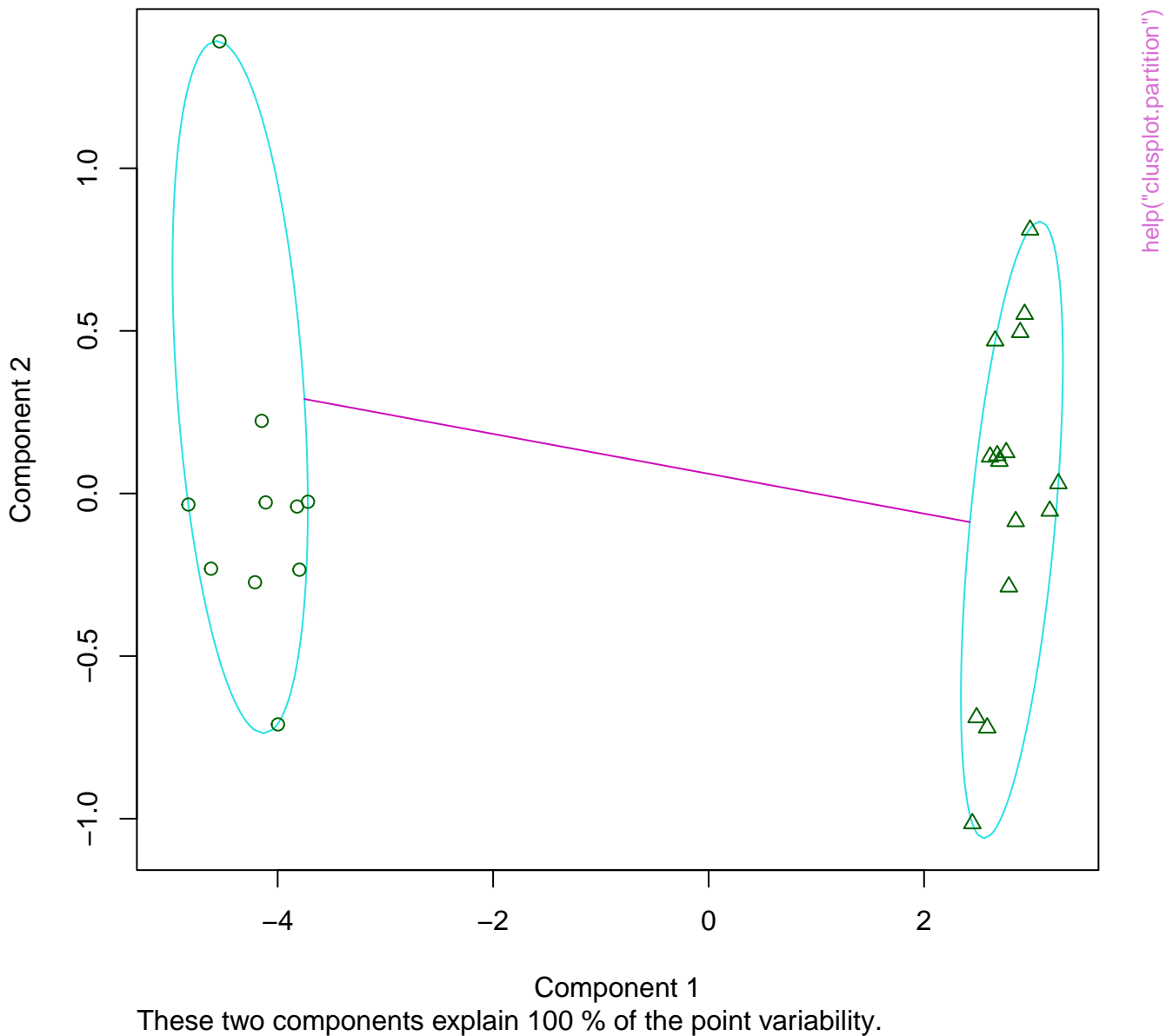


Component 1

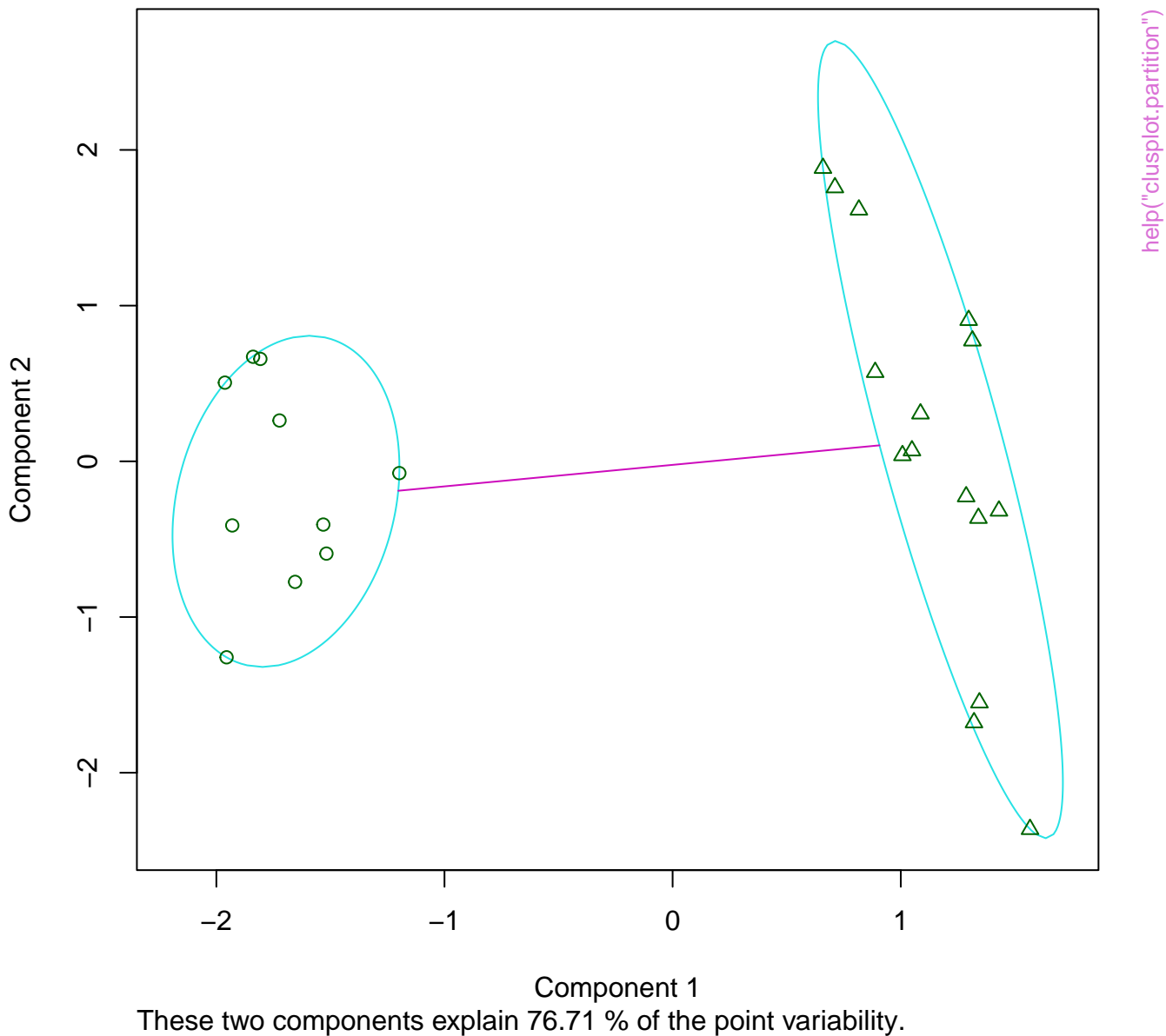
These two components explain 95.81 % of the

help("clusplot default")

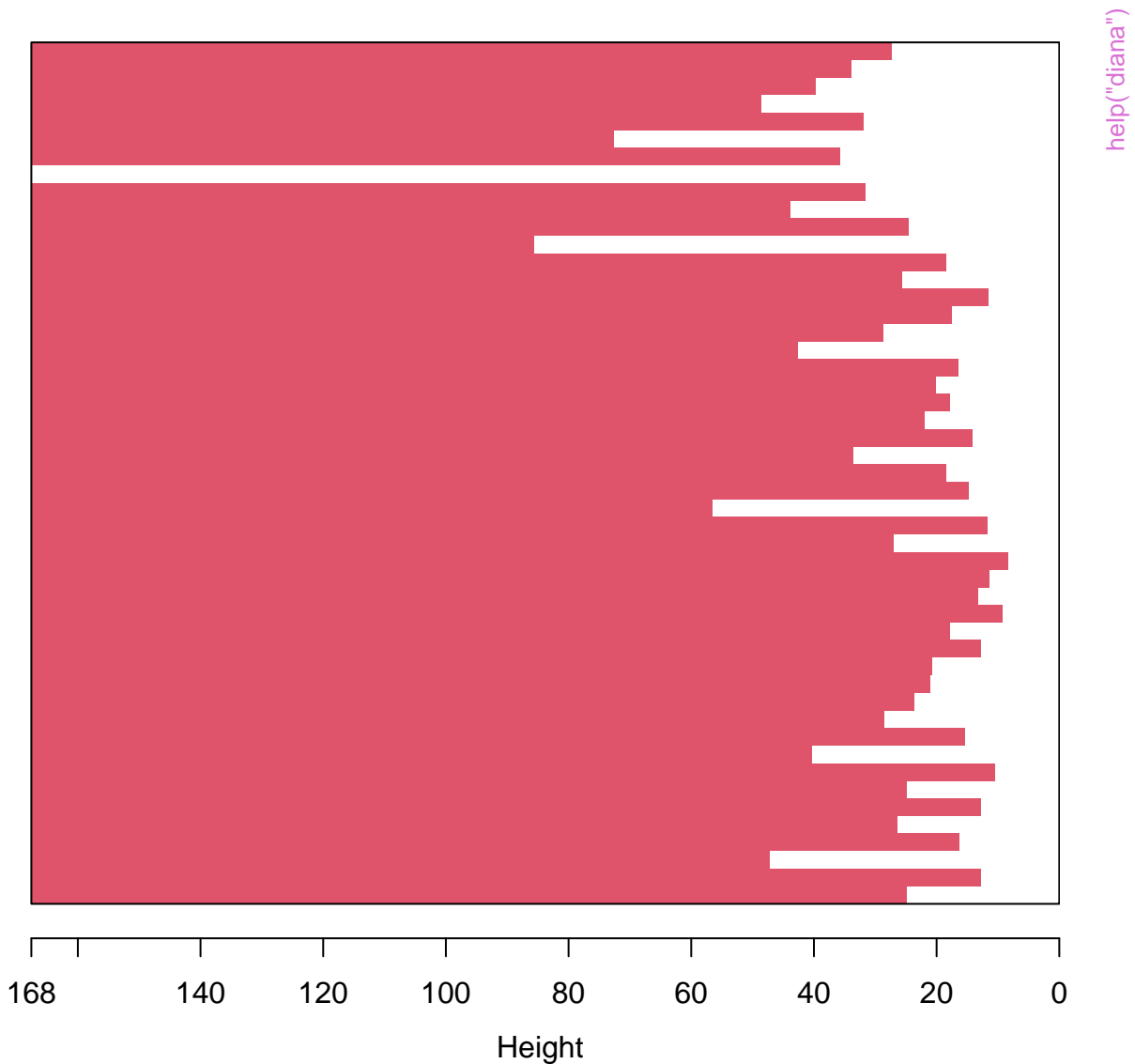
```
clusplot(pam(x = x, k = 2))
```



clusplot(pam(x = x4, k = 2))

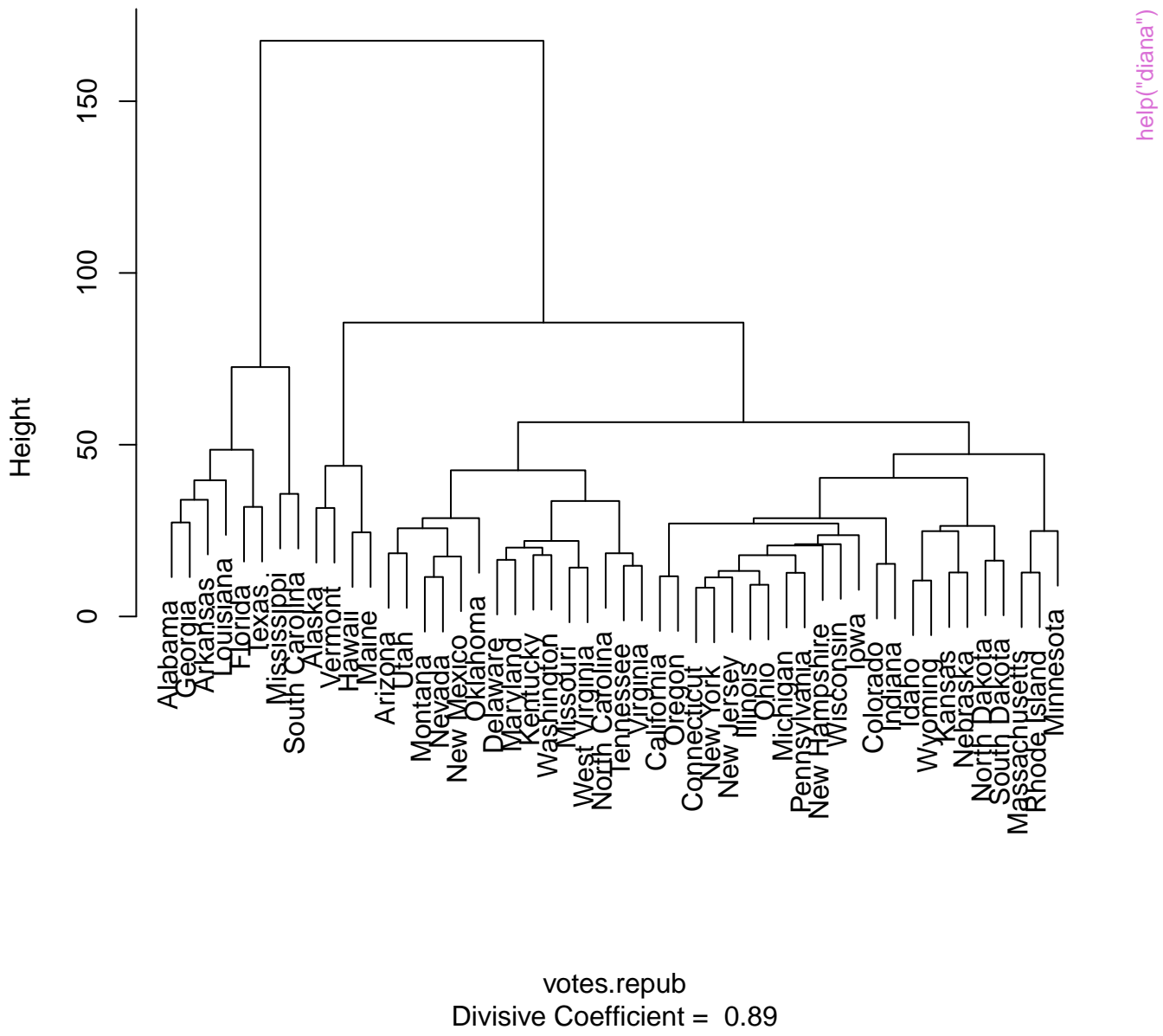


Banner of `diana(x = votes.repub, metric = "manhattan", stand = TR`

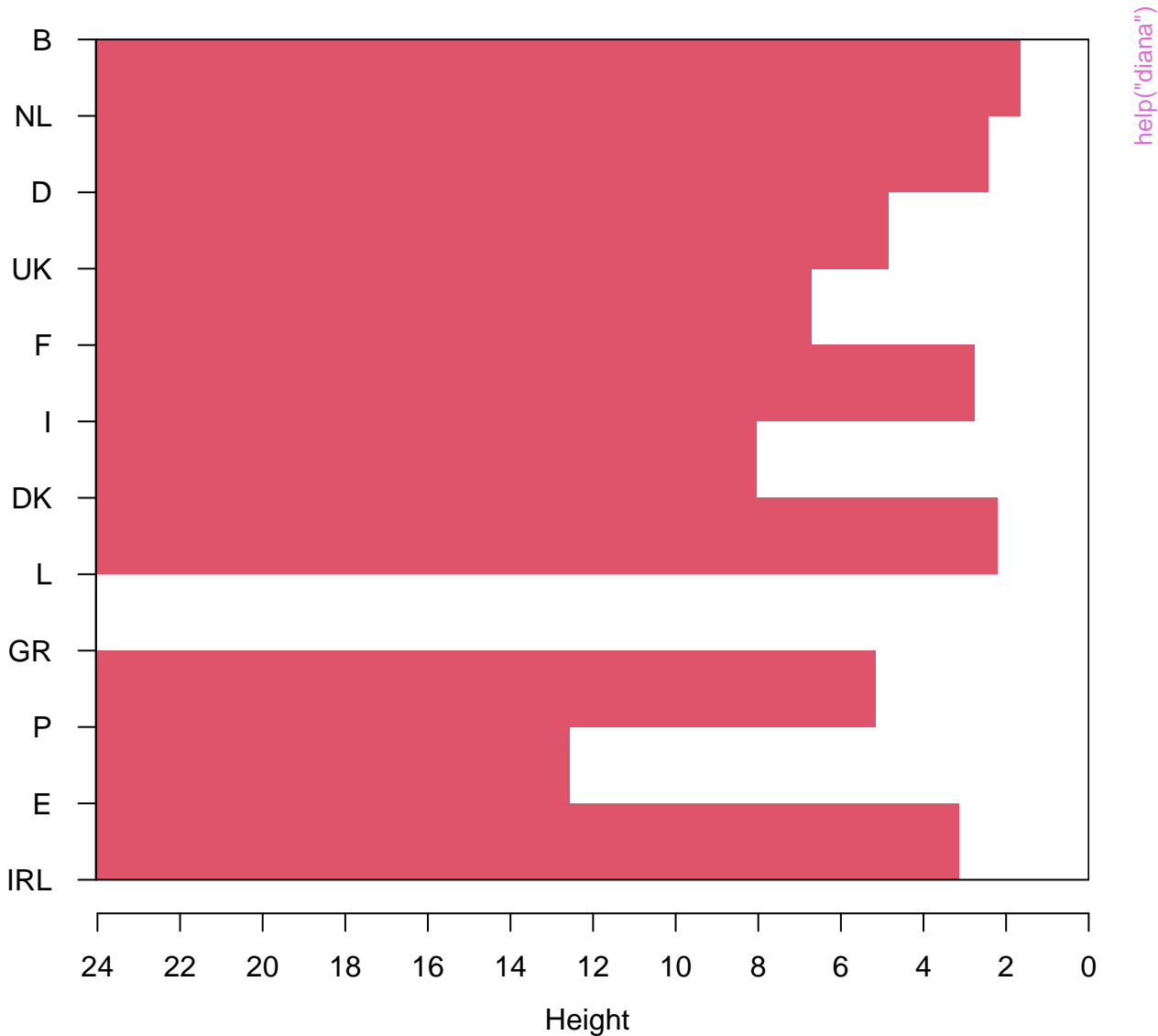


Divisive Coefficient = 0.89

Dendrogram of `diana(x = votes.repub, metric = "manhattan", stand = TRUE)`

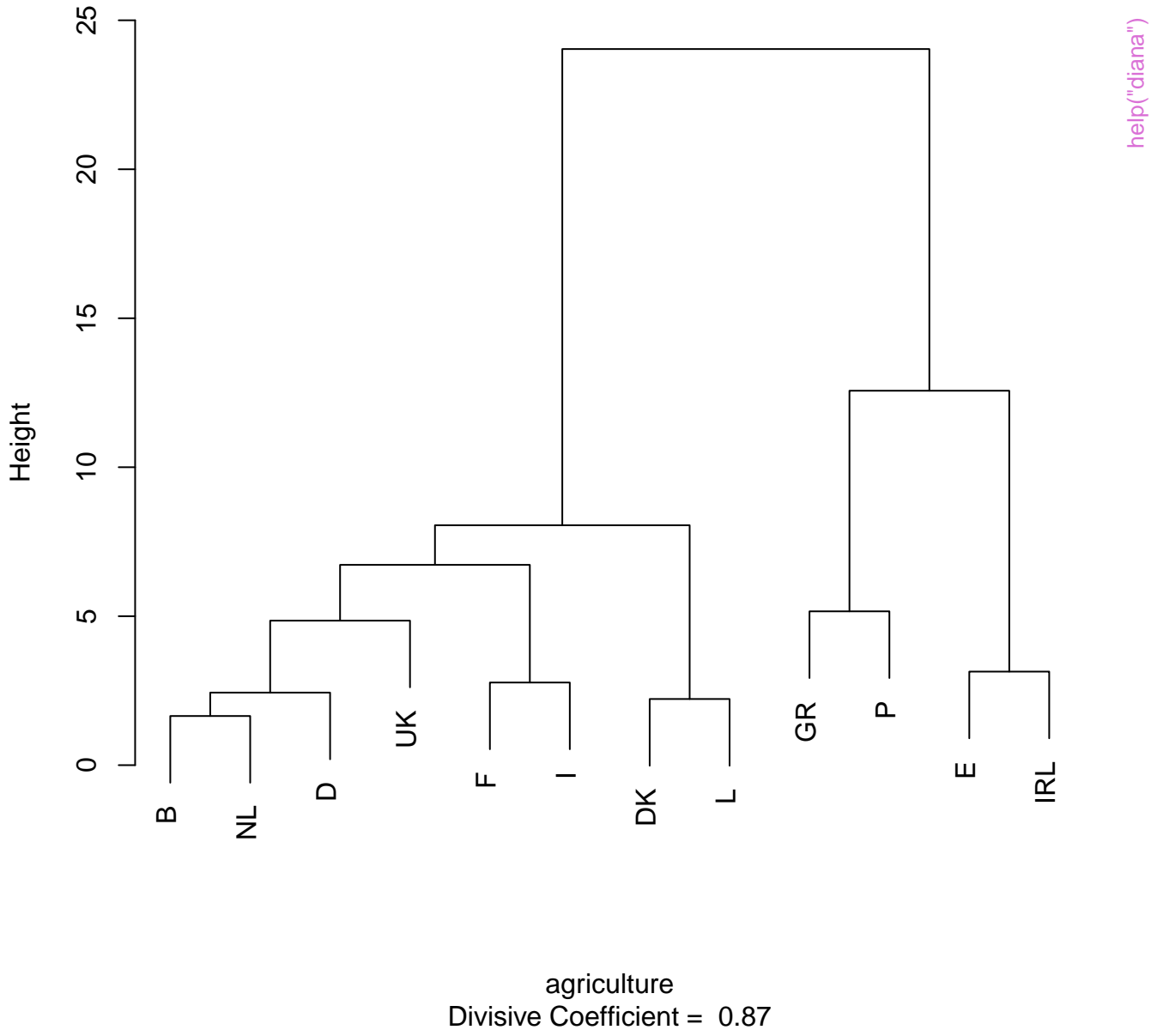


Banner of diana(x = agriculture)

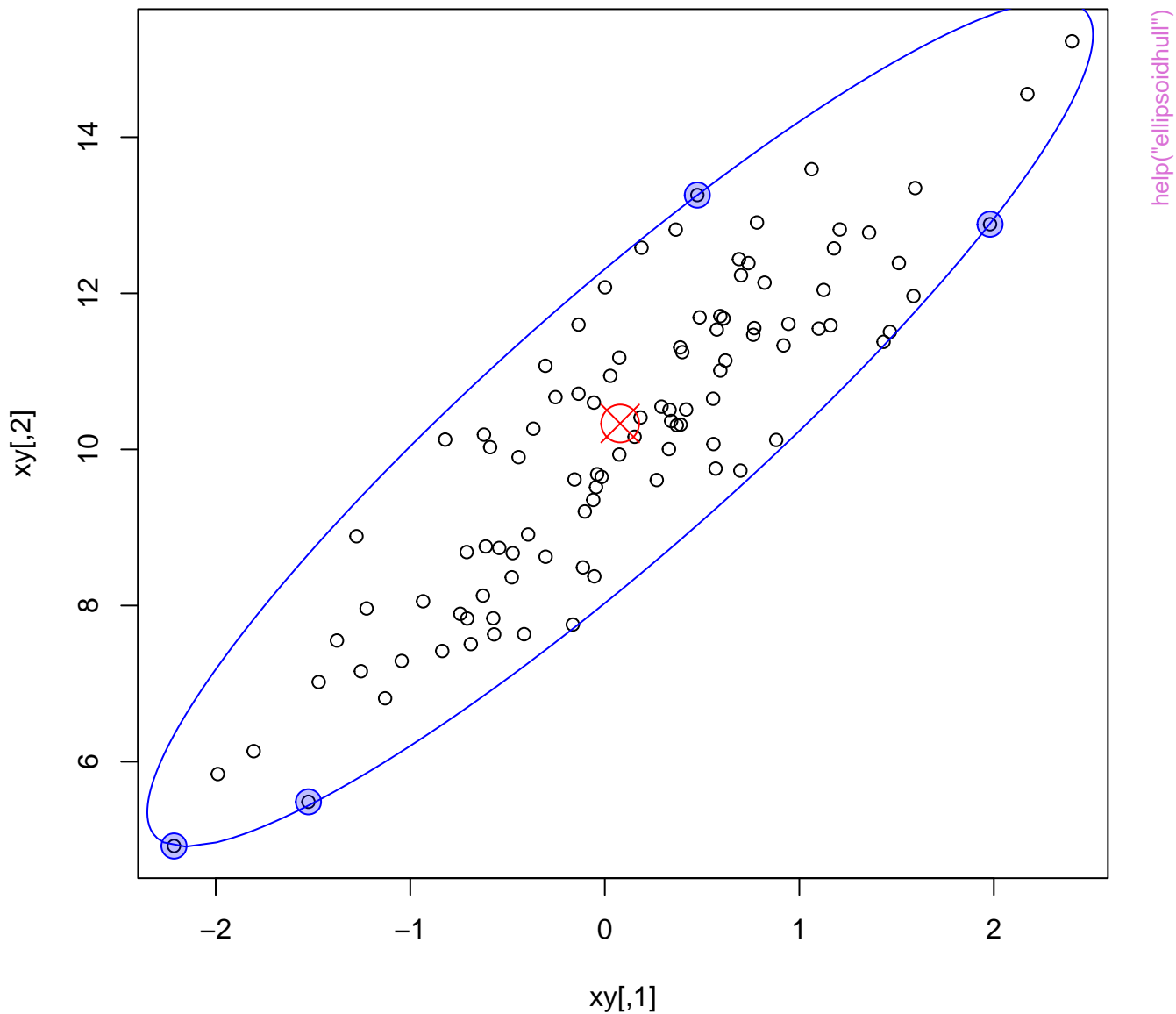


Divisive Coefficient = 0.87

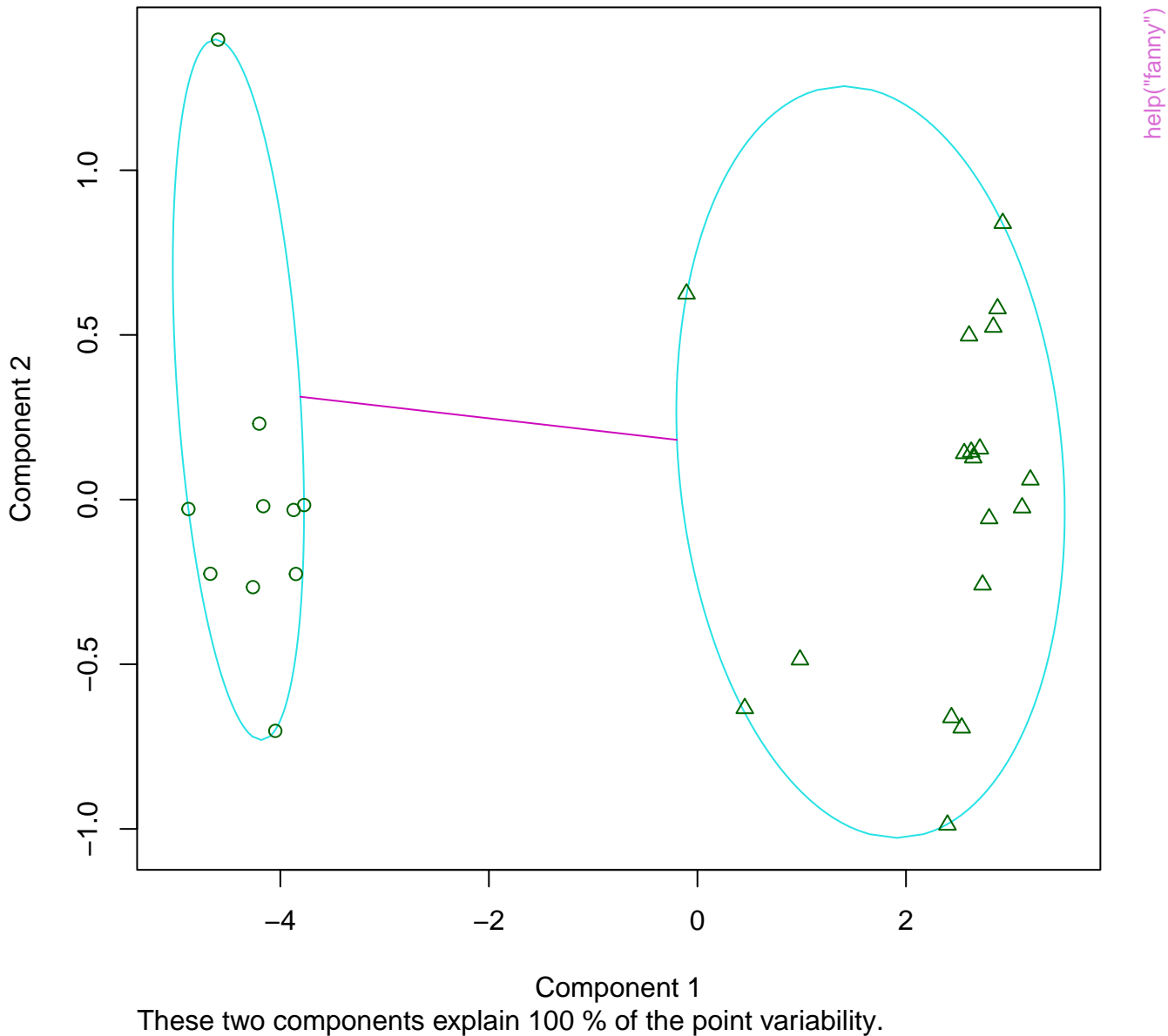
Dendrogram of diana(x = agriculture)



ellipsoidhull(<Gauss data>) -- 'spanning points'



clusplot(fanny(x = x, k = 2))

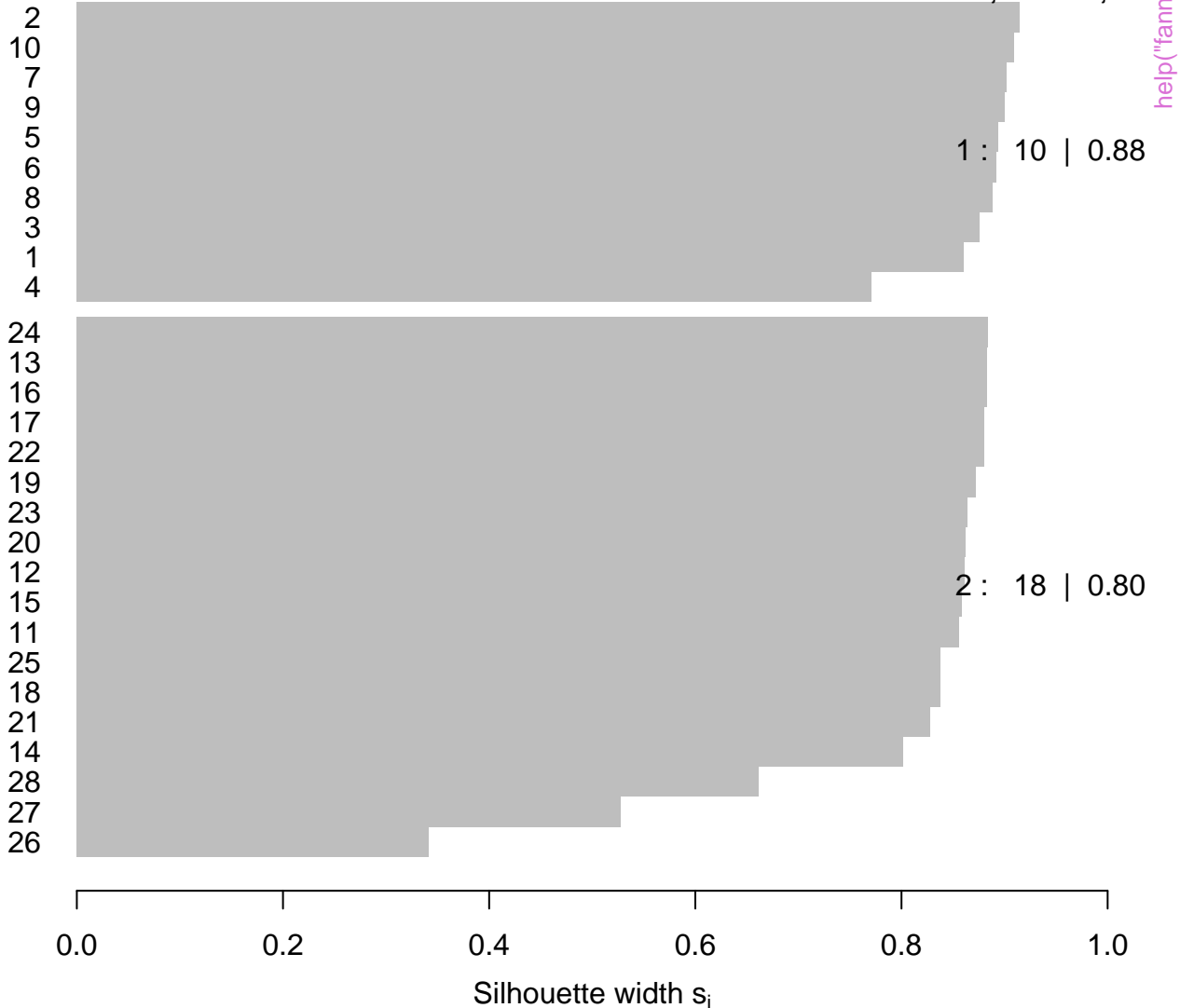


Silhouette plot of fanny(x = x, k = 2)

n = 28

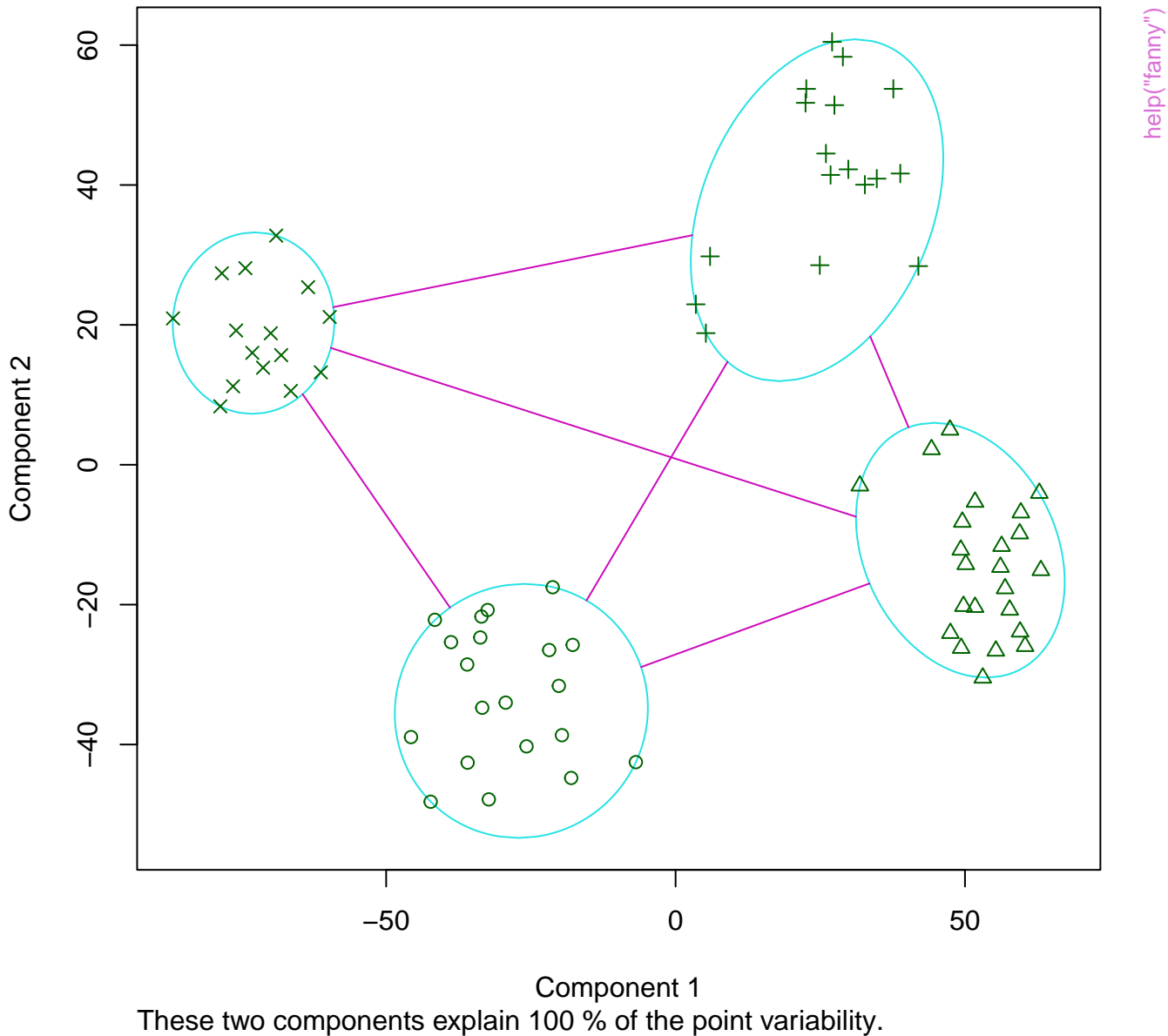
2 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$

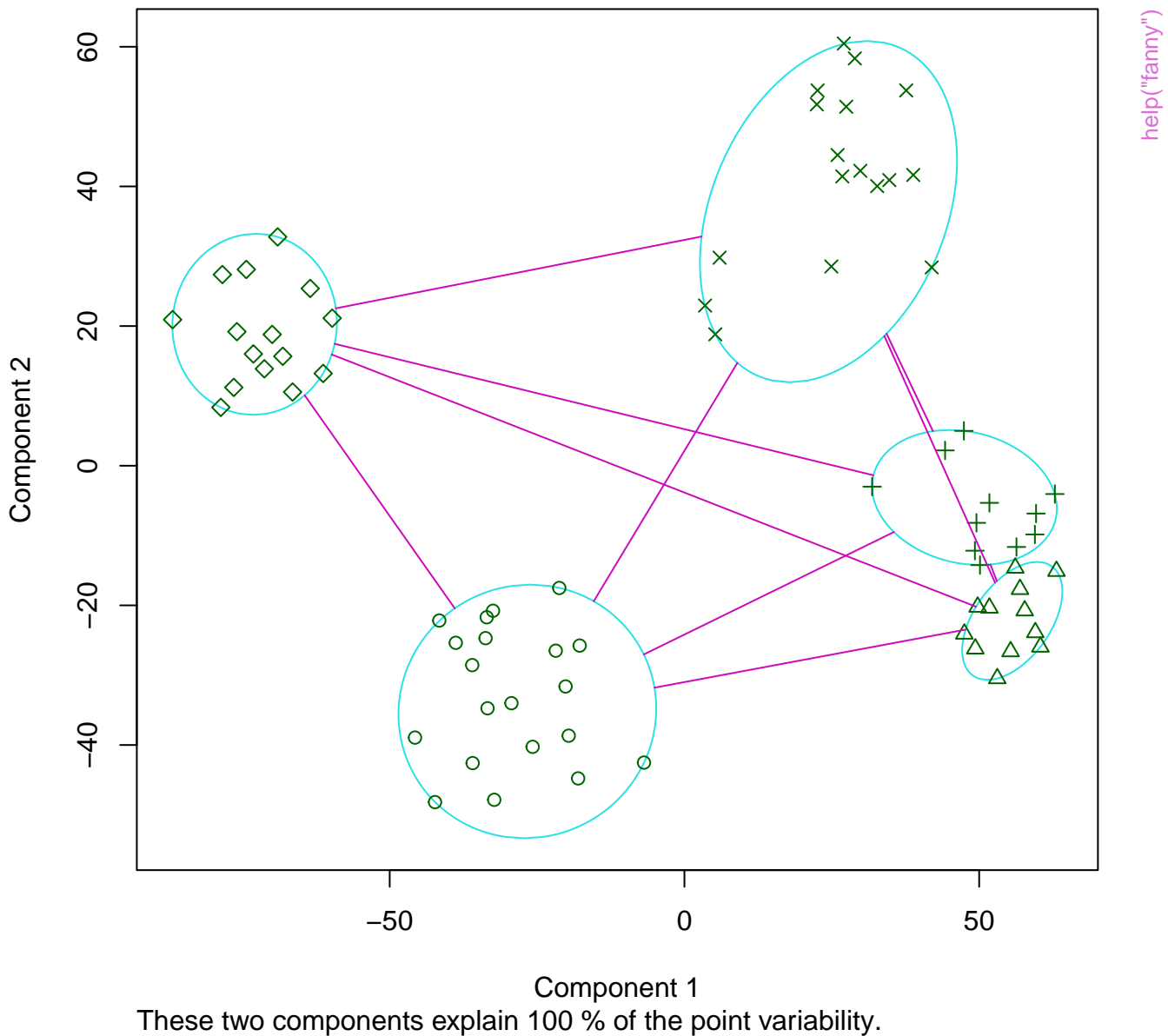


Average silhouette width : 0.83

clusplot(fanny(x = ruspini, k = 4))



clusplot(fanny(x = ruspini, k = 5))

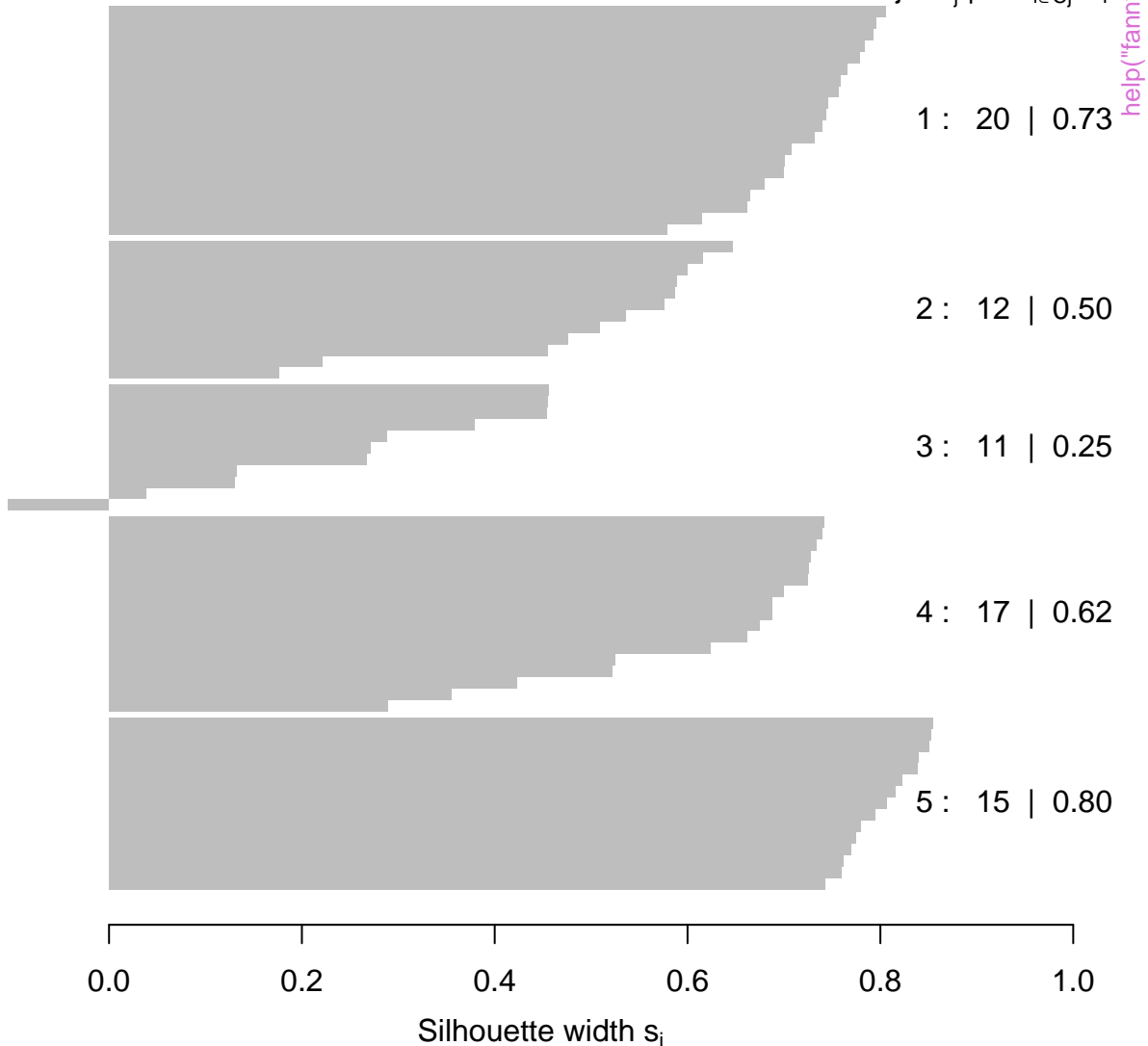


Silhouette plot of fanny(x = ruspini, k = 5)

n = 75

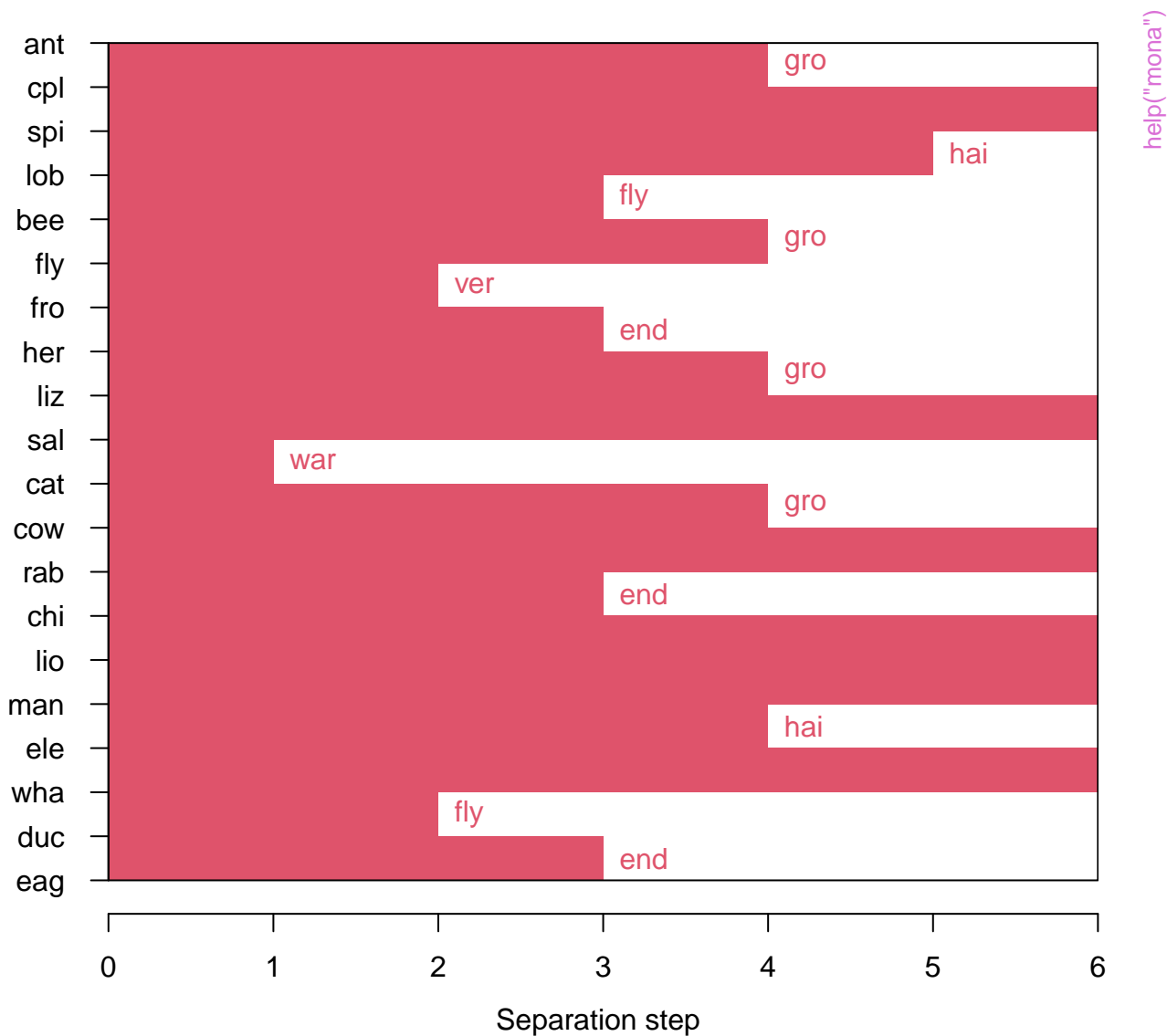
5 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$

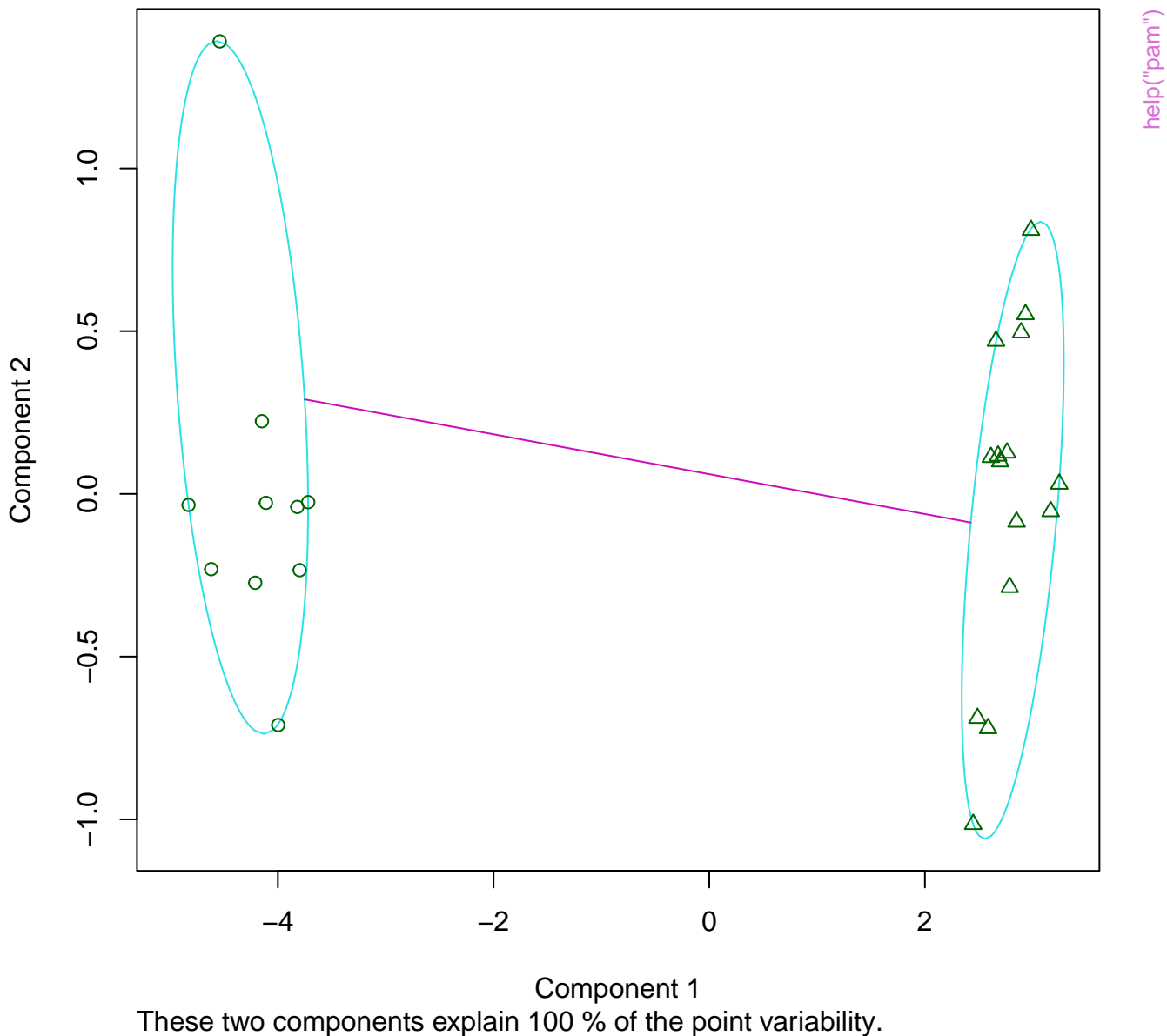


Average silhouette width : 0.61

Banner of mona(x = animals)



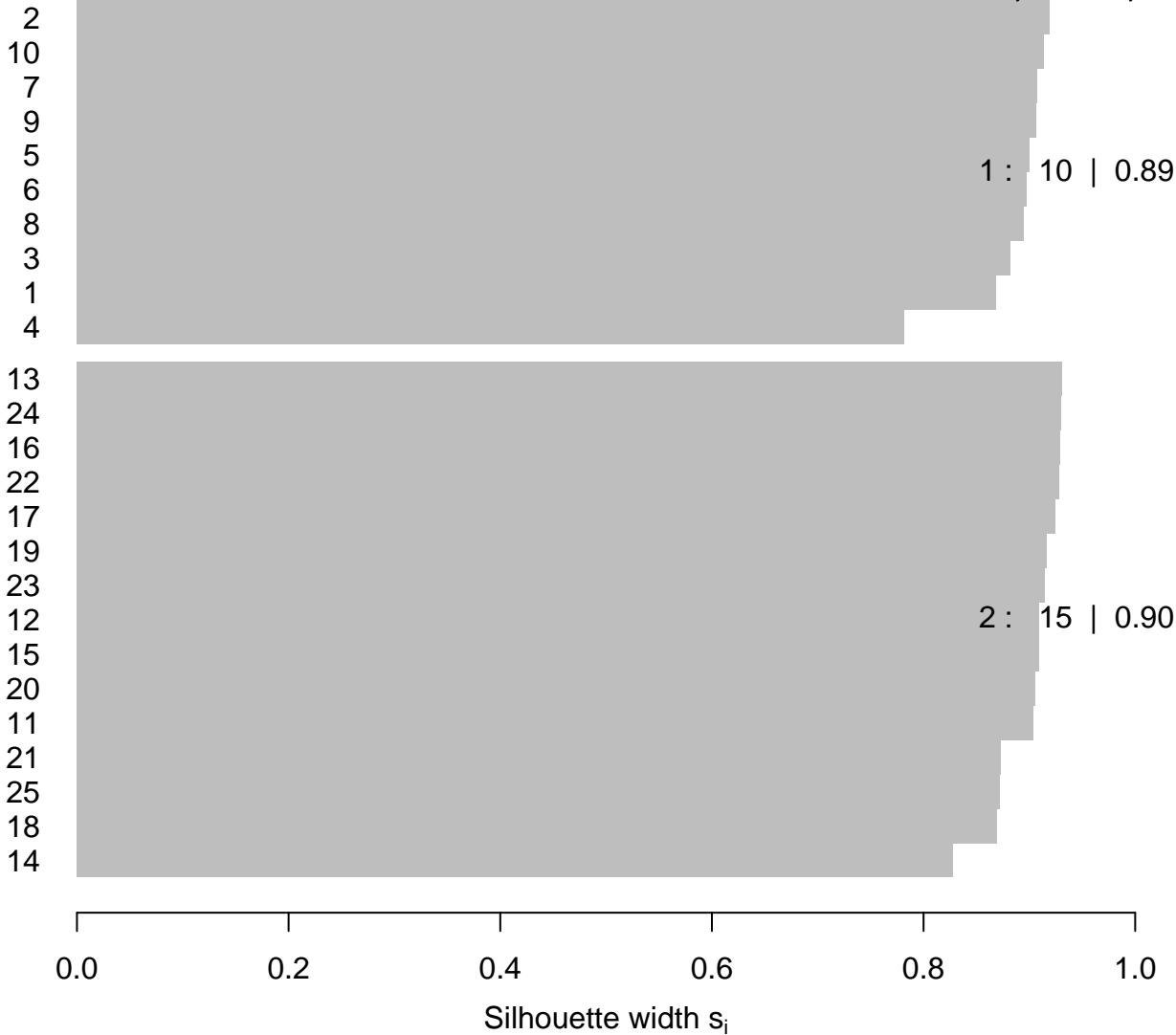
```
clusplot(pam(x = x, k = 2))
```



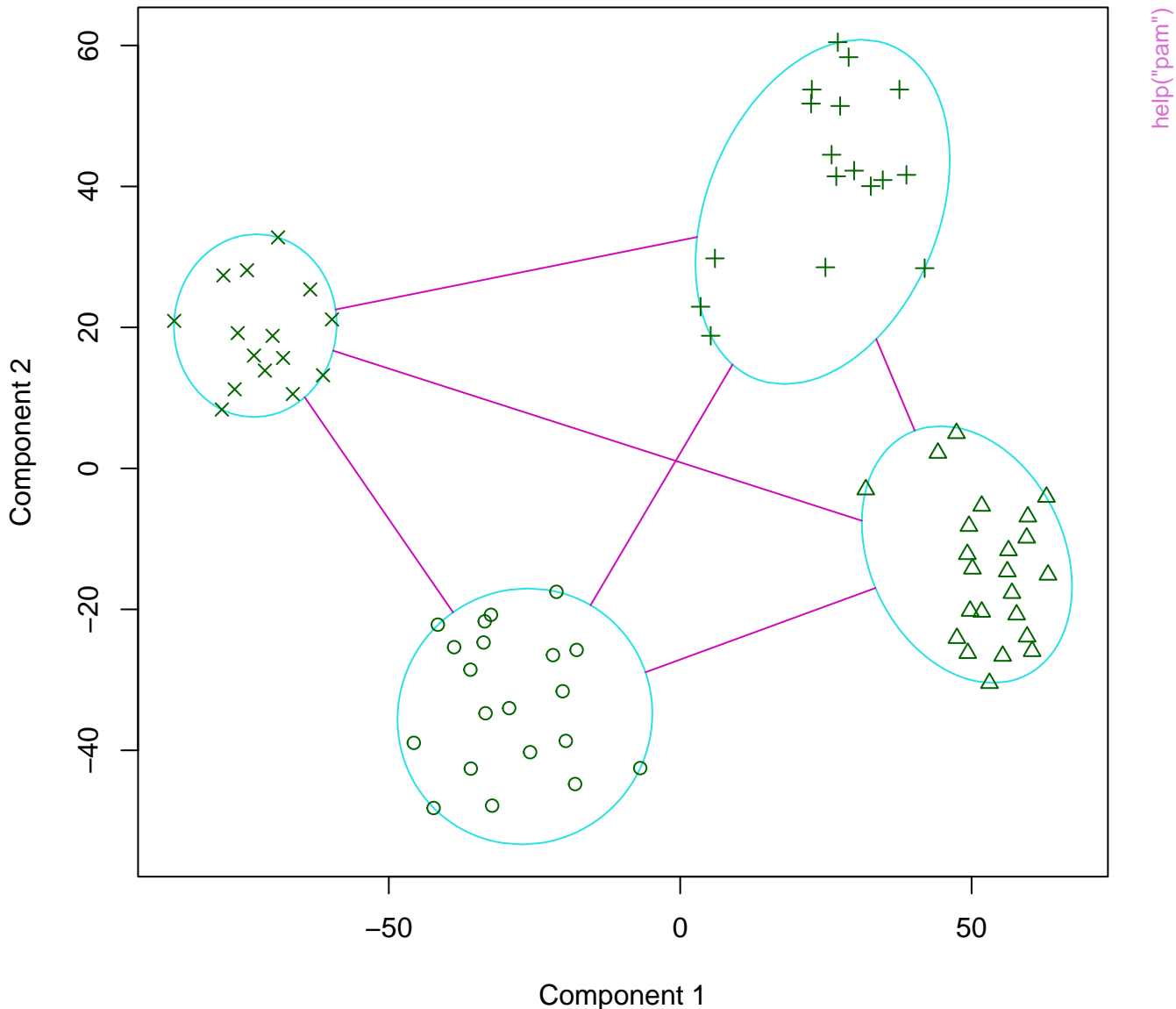
Silhouette plot of pam(x = x, k = 2)

n = 25

2 clusters C_j
 $j : n_j \mid \text{ave}_{i \in C_j} s_i$
help("pam")



clusplot(pam(x = ruspini, k = 4))



These two components explain 100 % of the point variability.

Silhouette plot of pam(x = ruspini, k = 4)

n = 75

4 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$

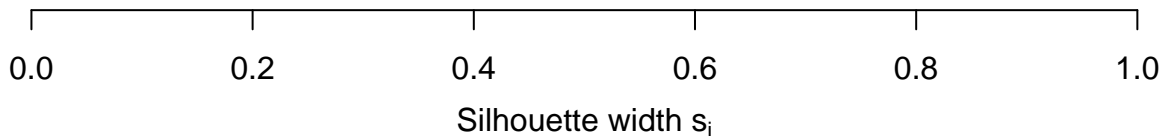
help("pam")

1 : 20 | 0.73

2 : 23 | 0.75

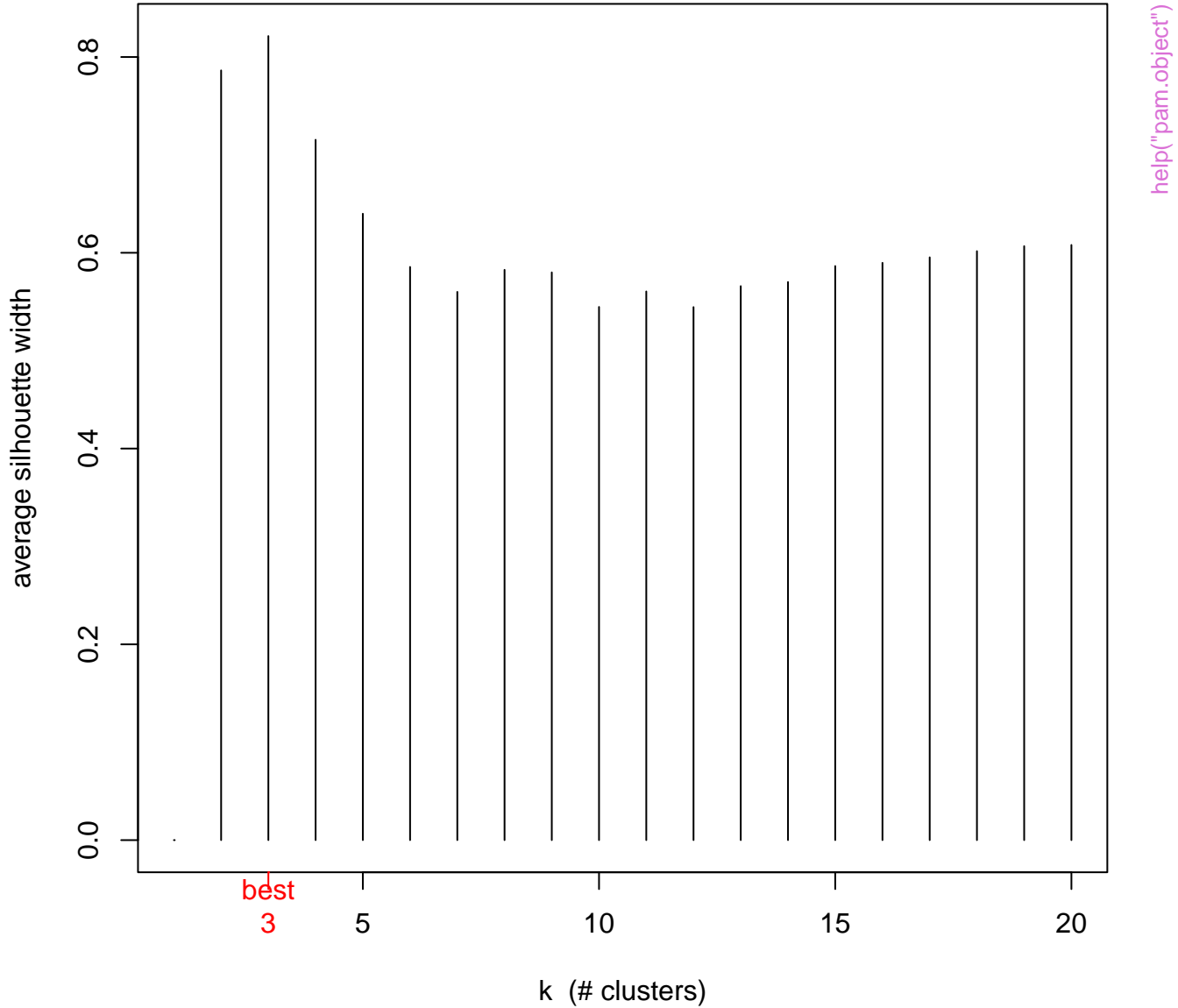
3 : 17 | 0.67

4 : 15 | 0.80

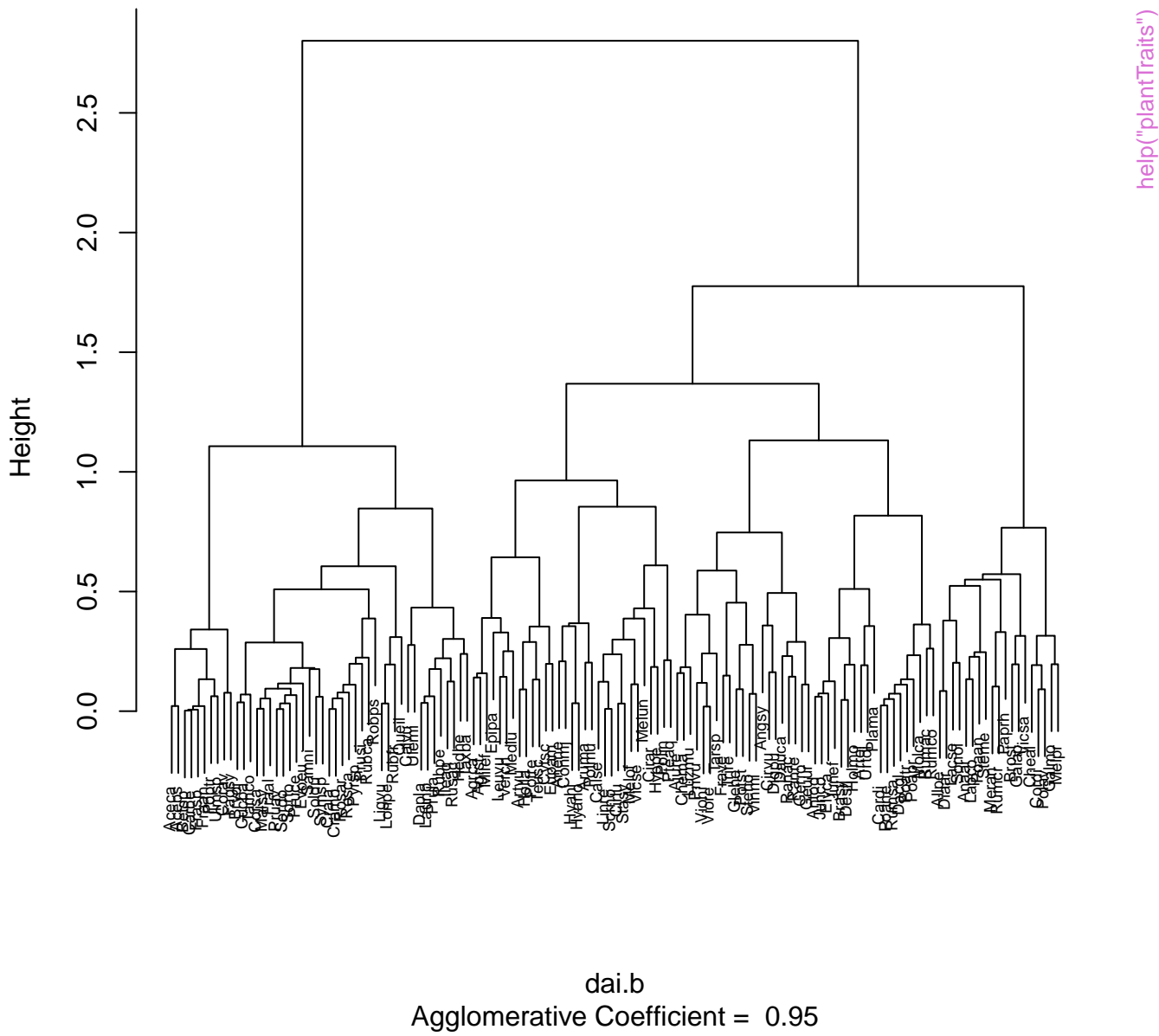


Average silhouette width : 0.74

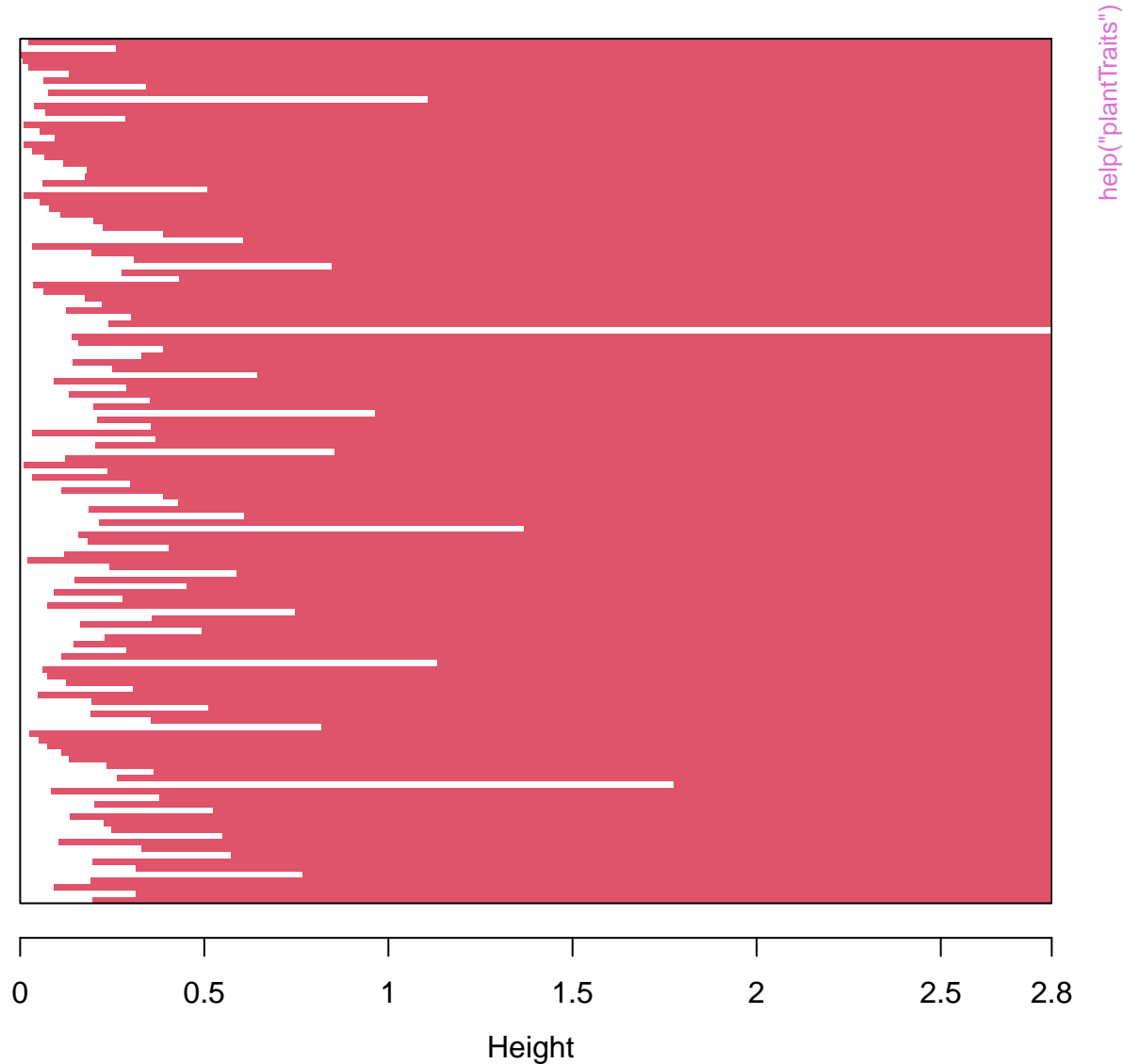
pam() clustering assessment



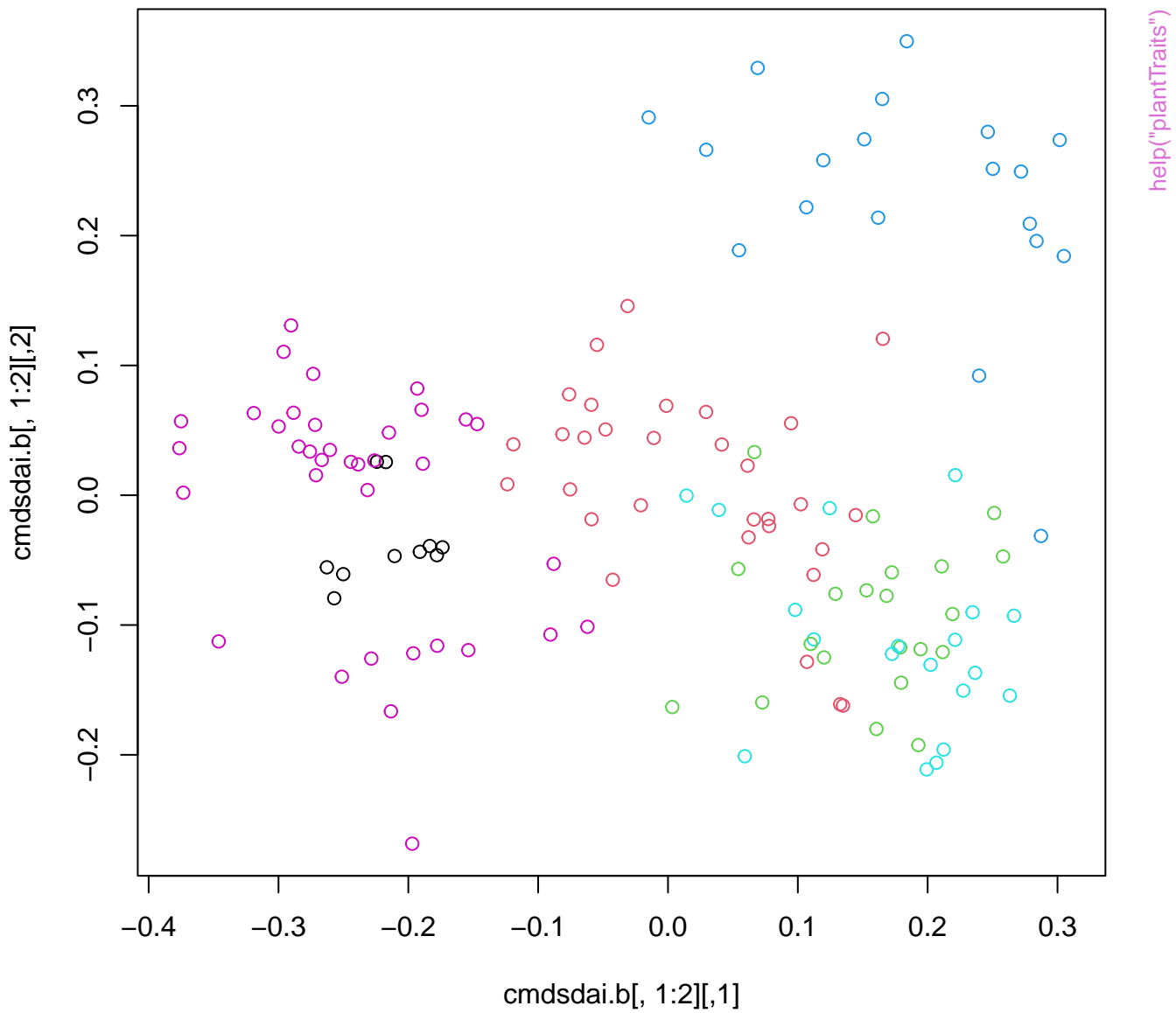
Dendrogram of `agnes(x = dai.b, method = "ward")`



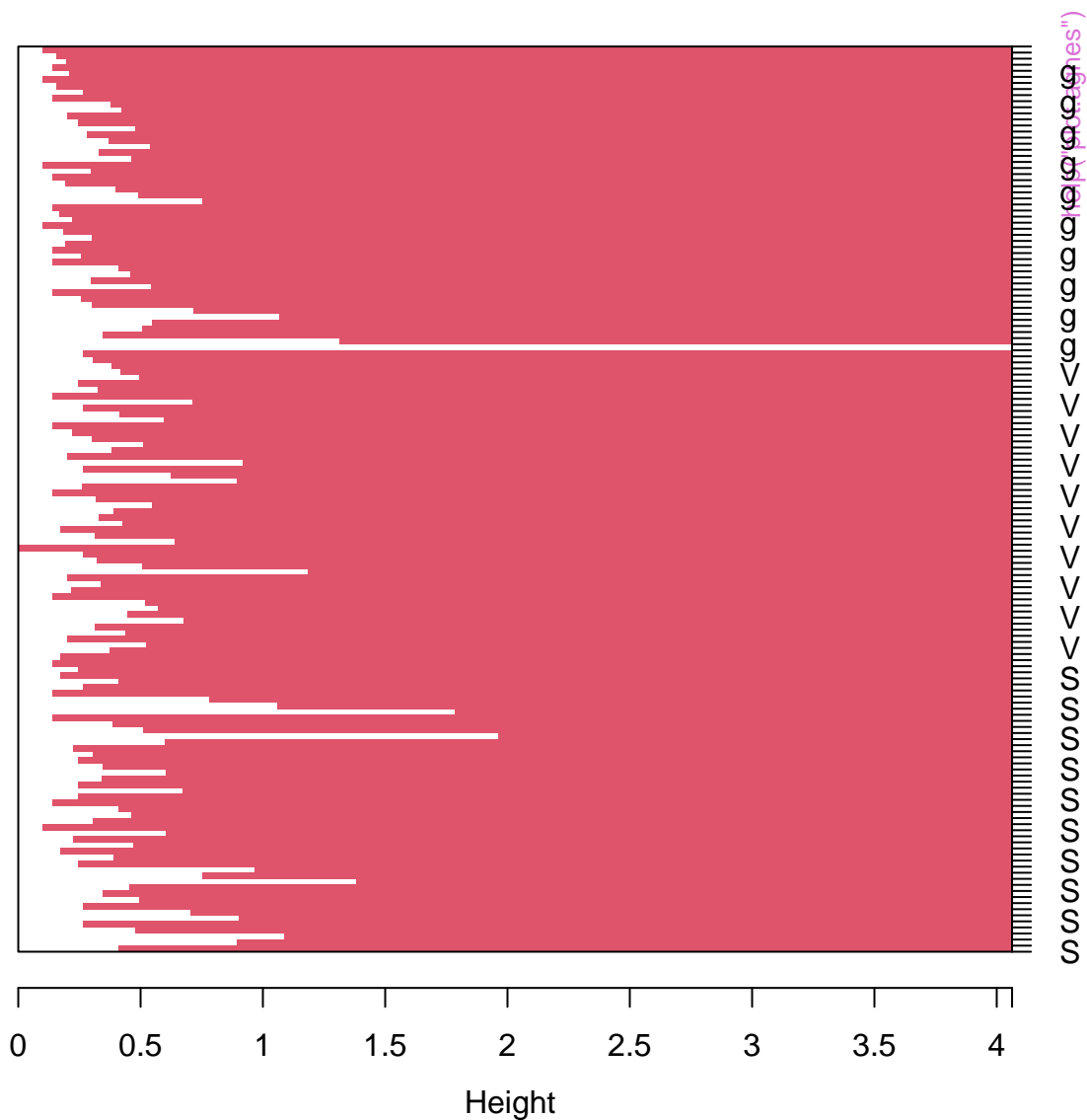
Banner of `agnes(x = dai.b, method = "ward")`



Agglomerative Coefficient = 0.95

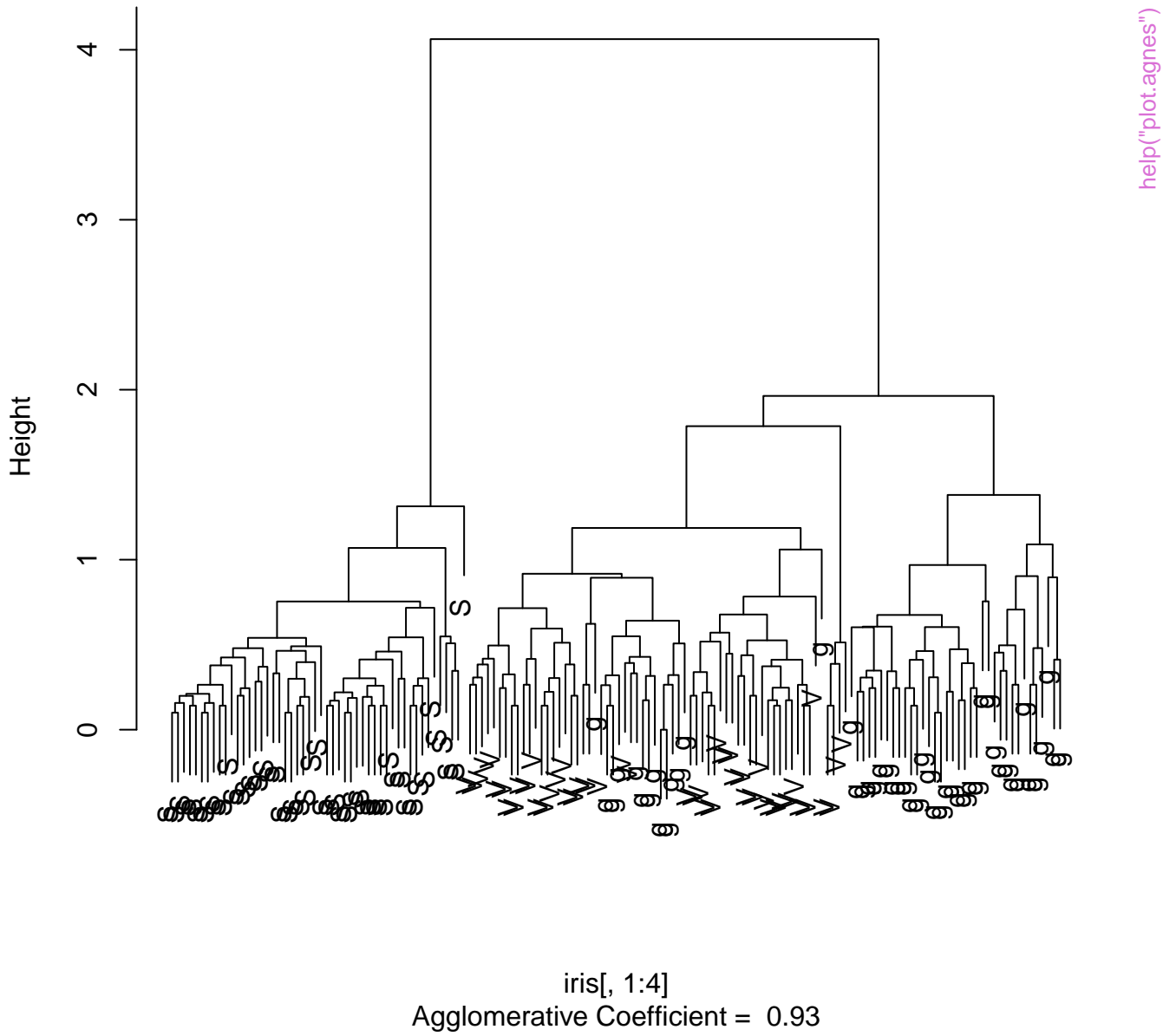


Banner of `agnes(x = iris[, 1:4])`

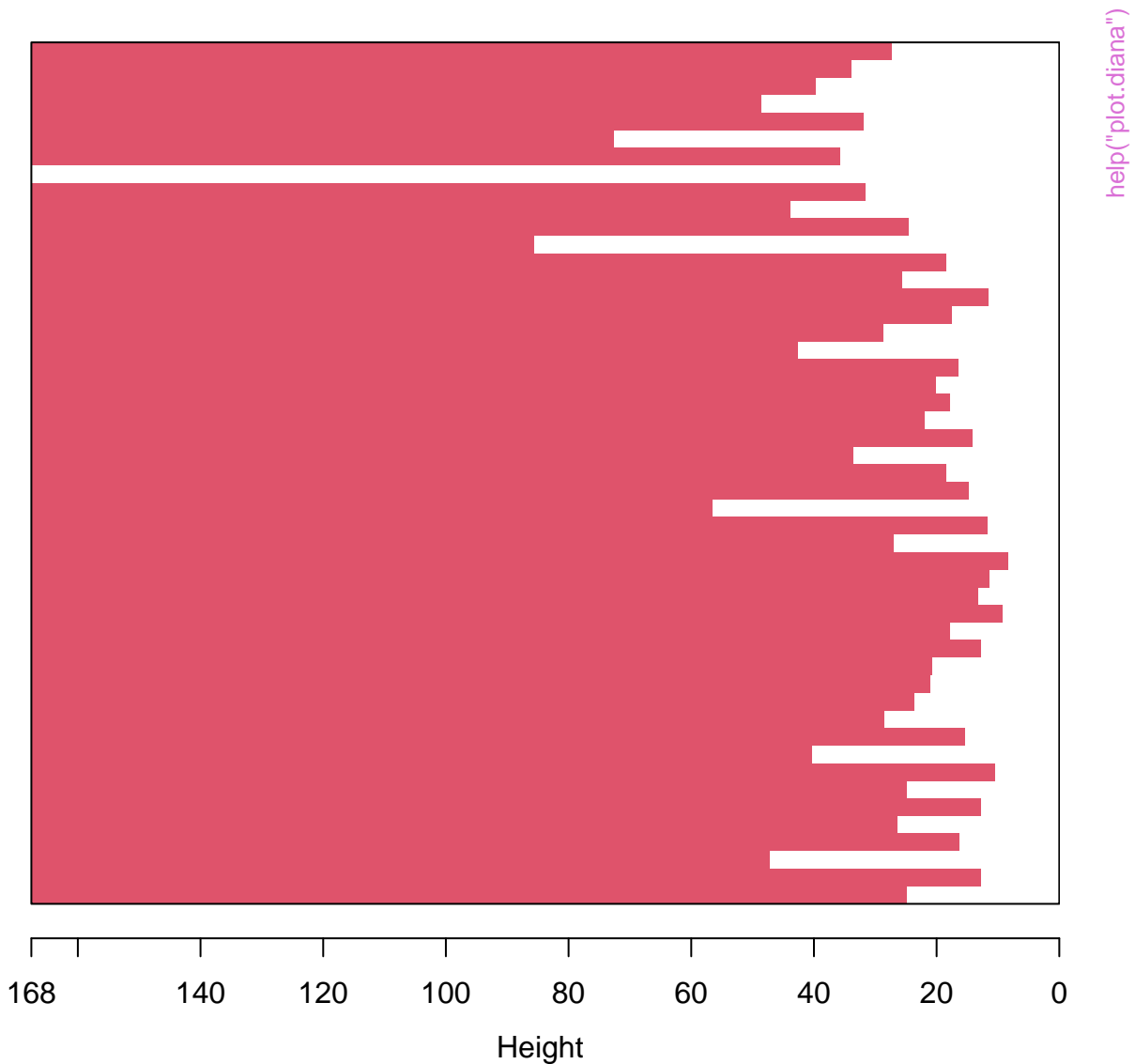


Agglomerative Coefficient = 0.93

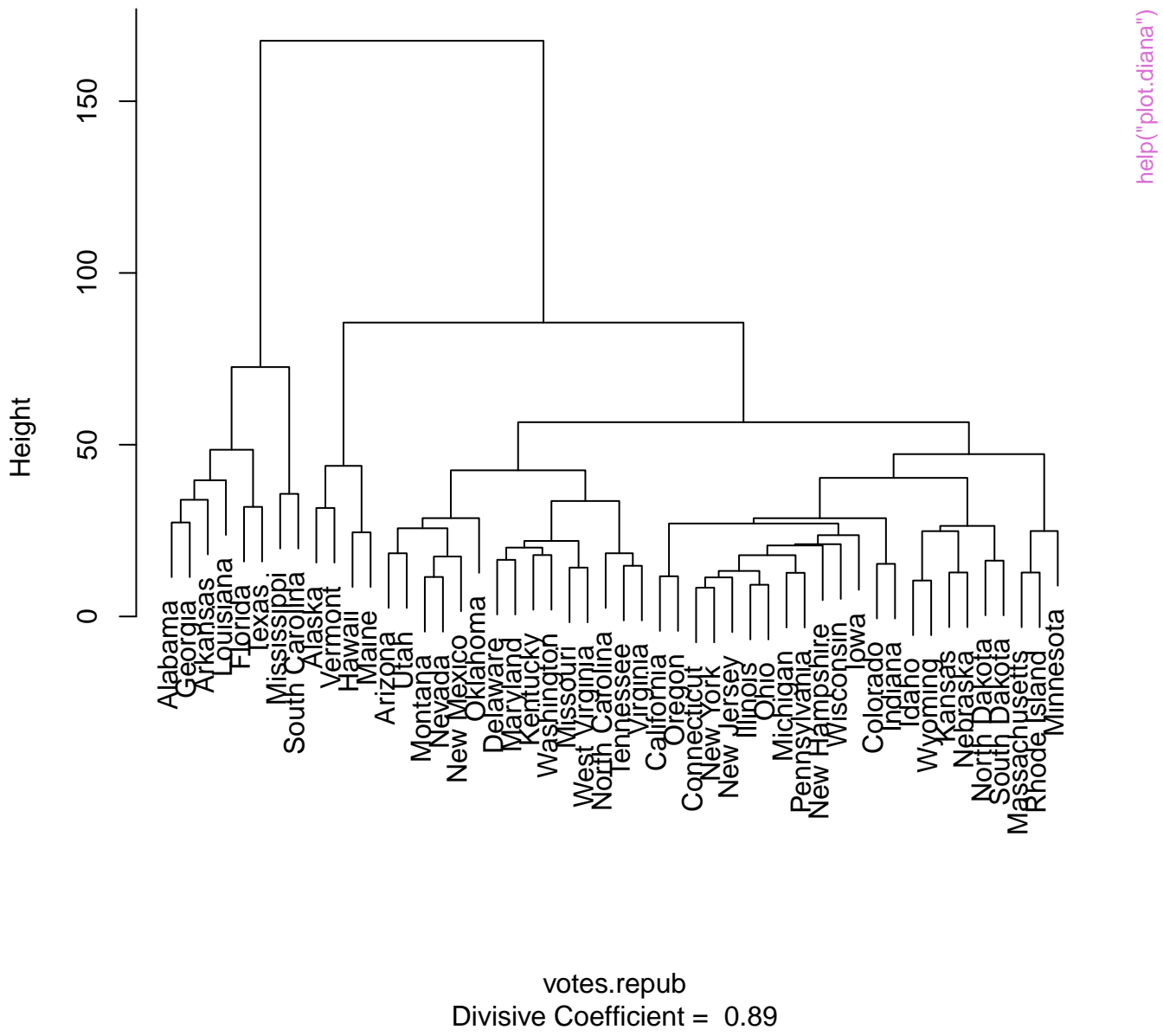
Dendrogram of `agnes(x = iris[, 1:4])`



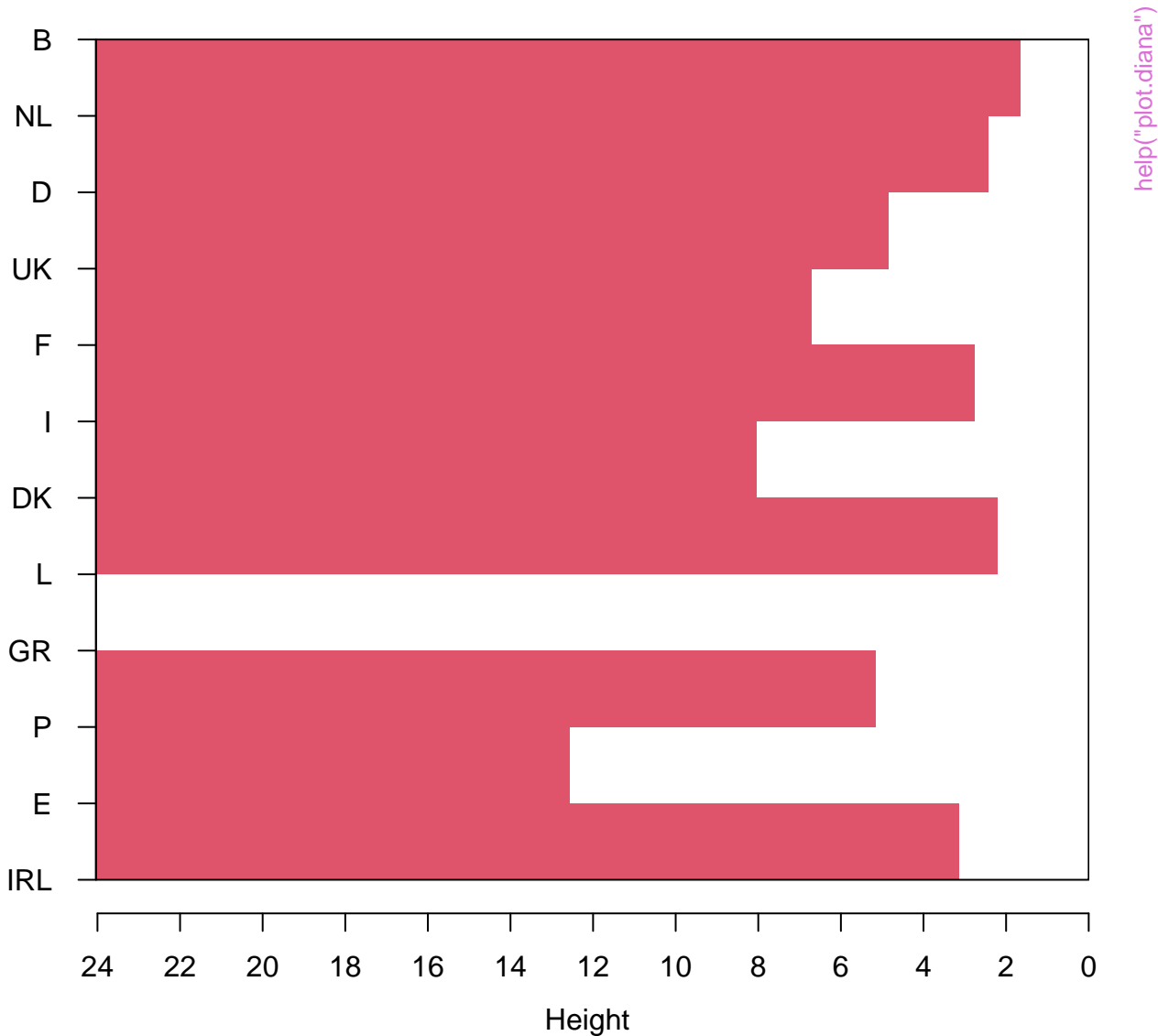
Banner of `diana(x = votes.repub, metric = "manhattan", stand = TR`



Dendrogram of `diana(x = votes.repub, metric = "manhattan", stand = TRUE)`

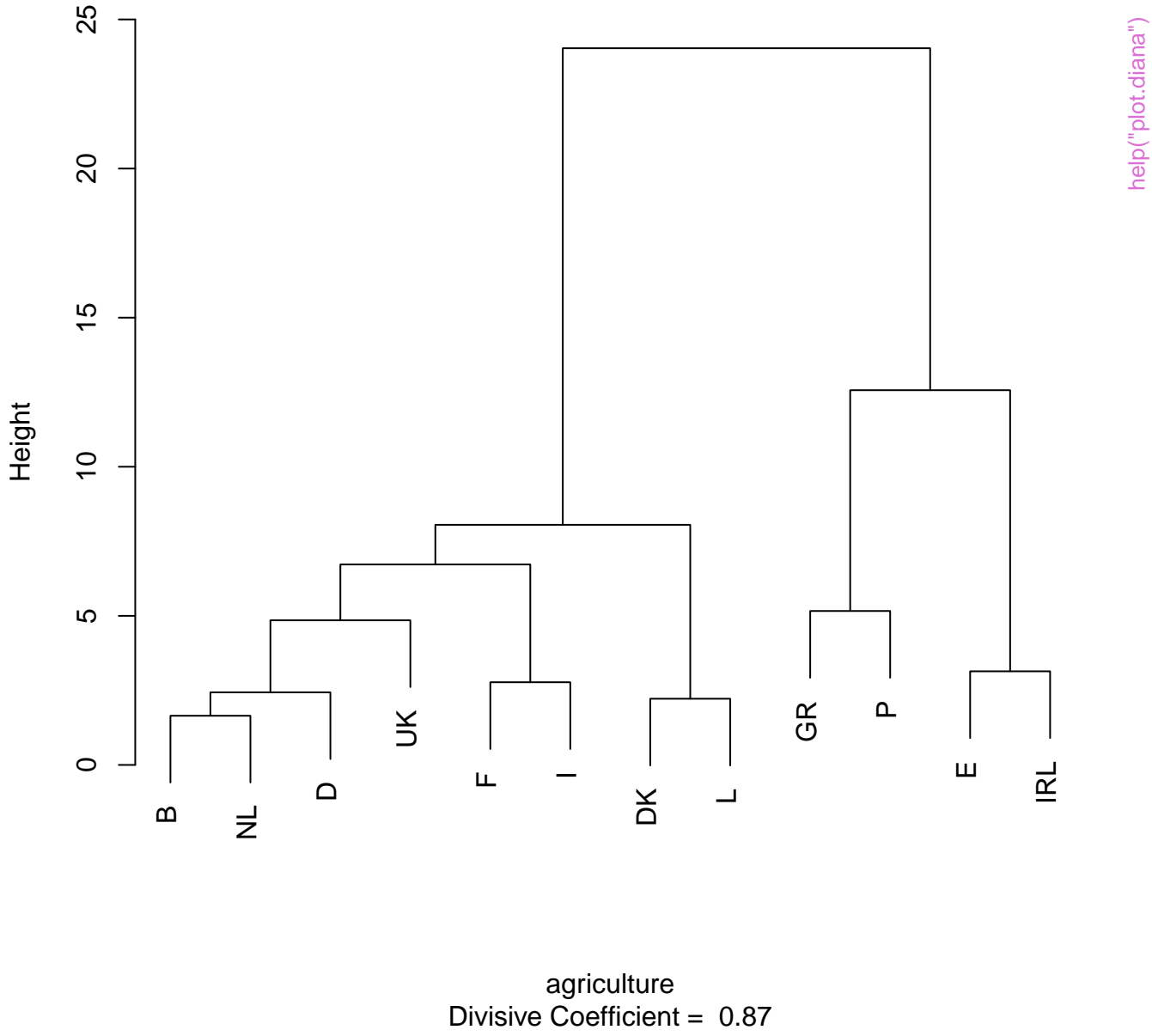


Banner of diana(x = agriculture)

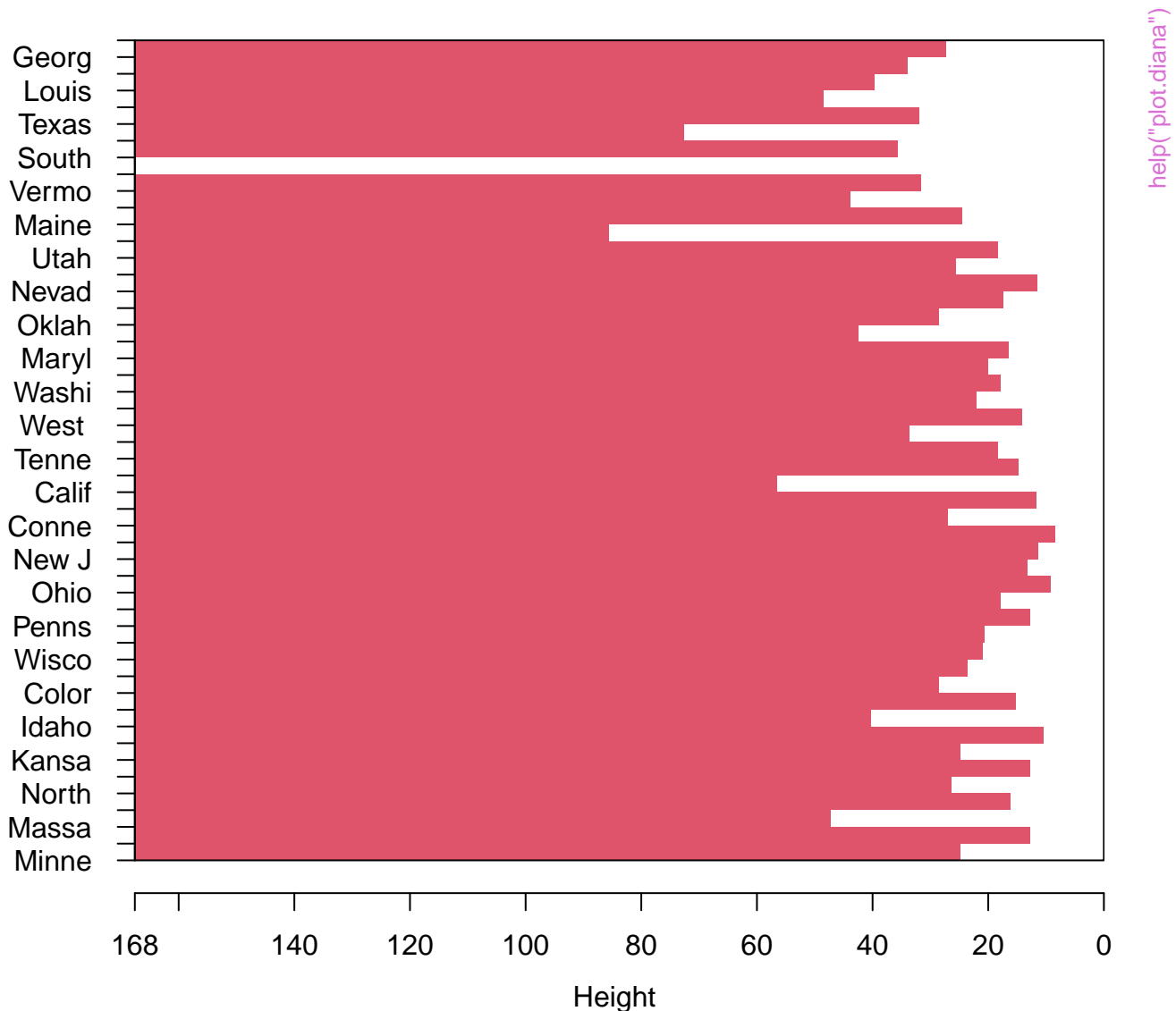


Divisive Coefficient = 0.87

Dendrogram of diana(x = agriculture)

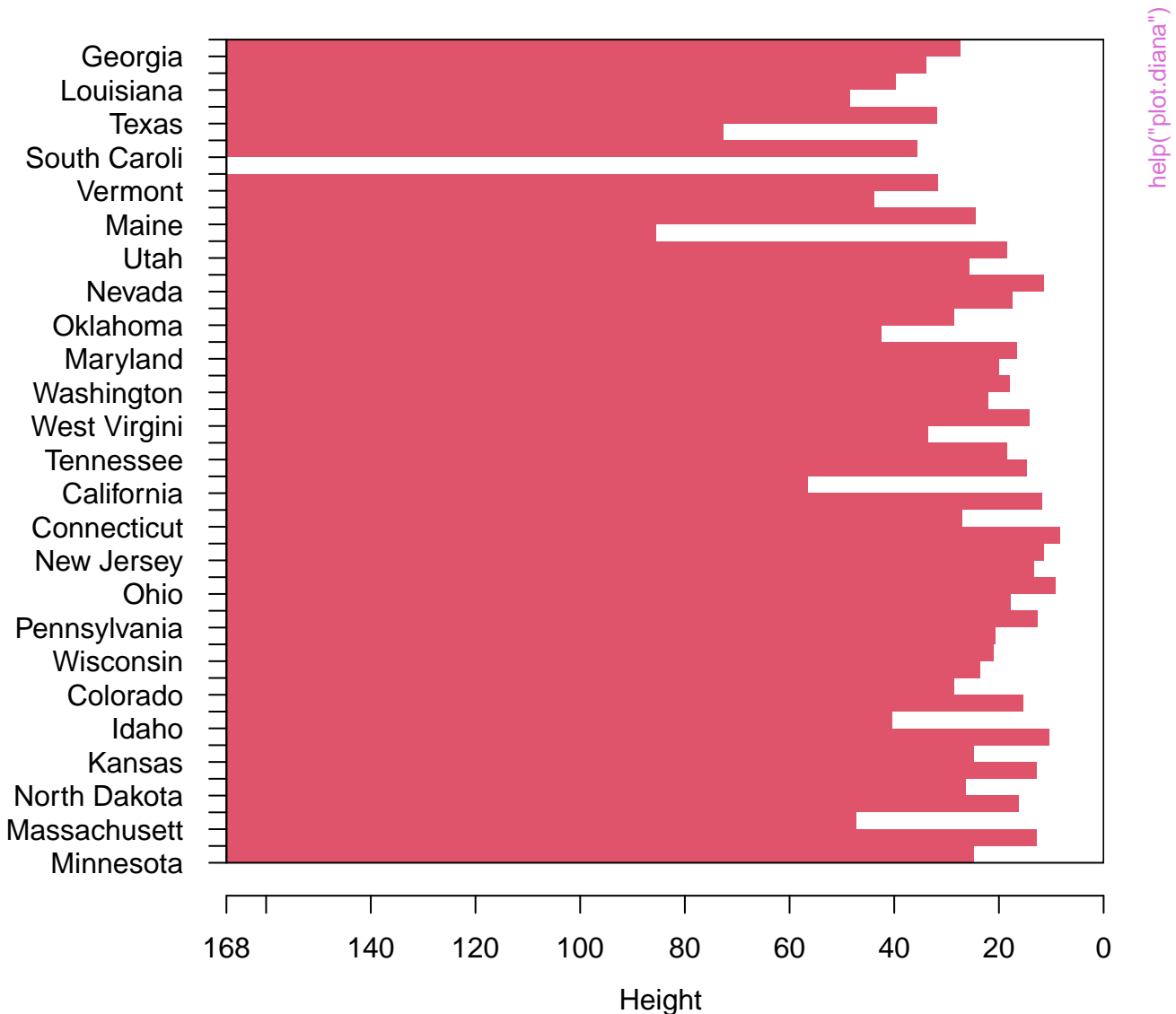


Banner of `diana(x = votes.repub, metric = "manhattan", stand = T`



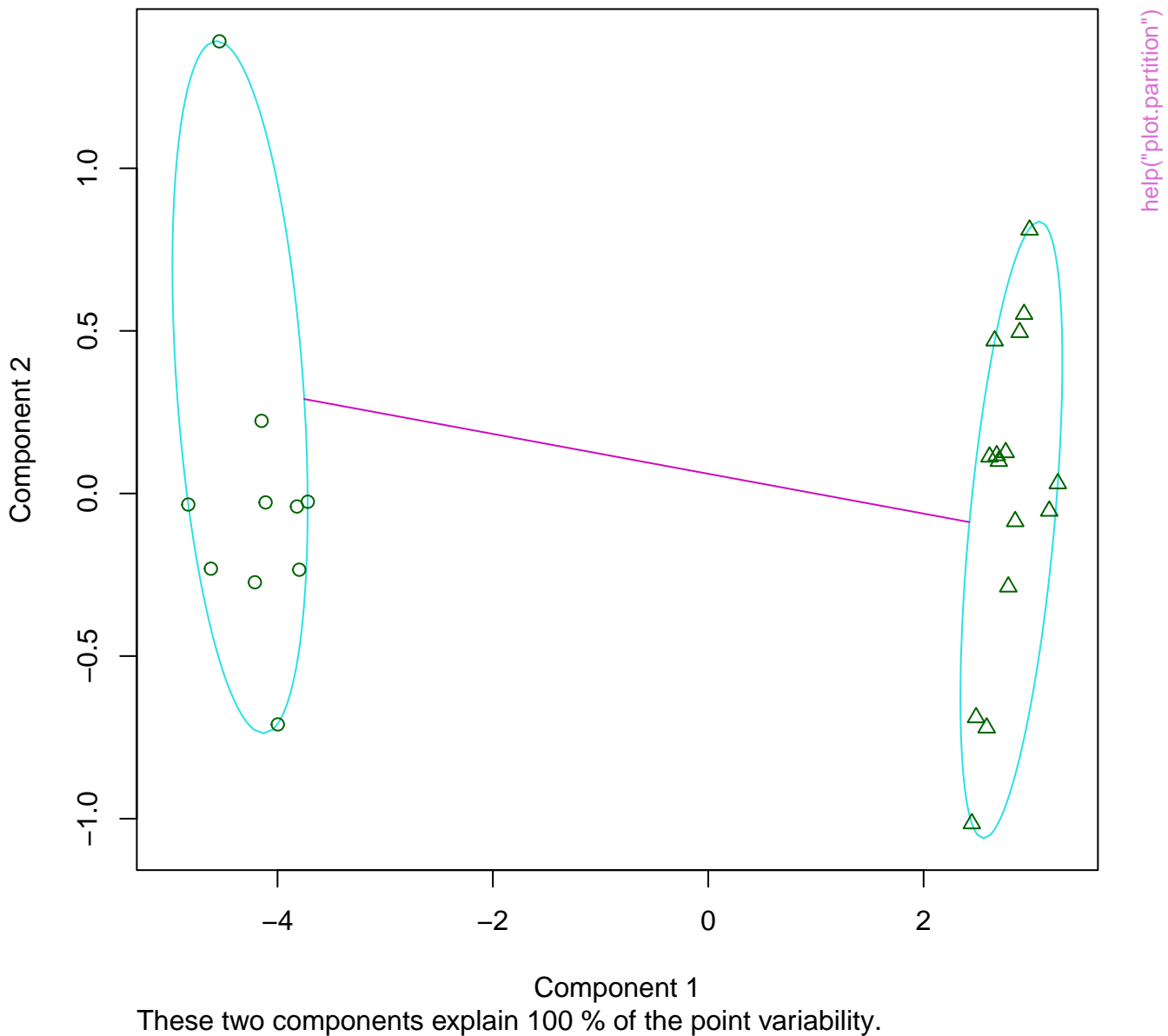
Divisive Coefficient = 0.89

Banner of `diana(x = votes.repub, metric = "manhattan", star`



Divisive Coefficient = 0.89

clusplot(pam(x = x, k = 2))

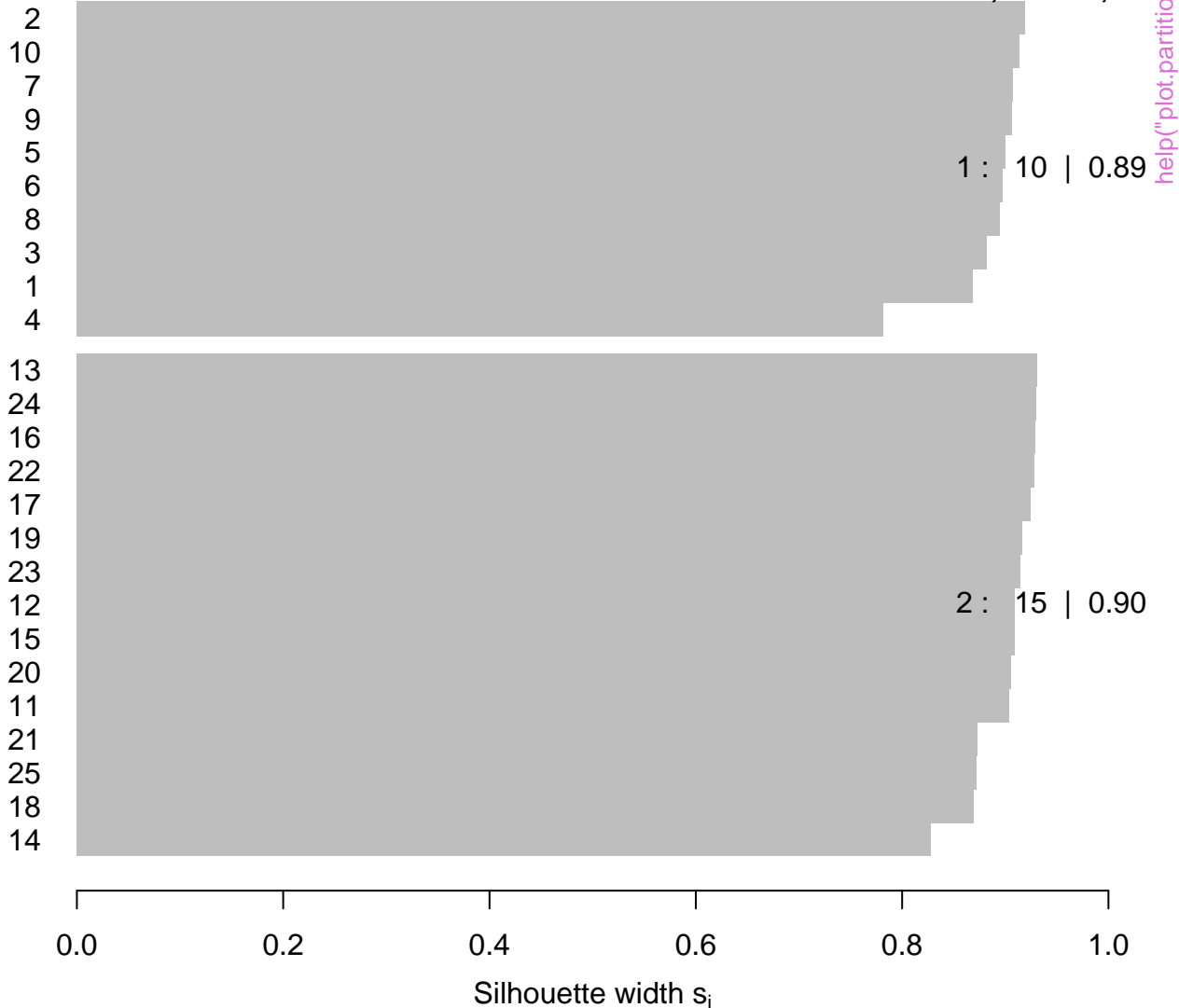


Silhouette plot of pam(x = x, k = 2)

n = 25

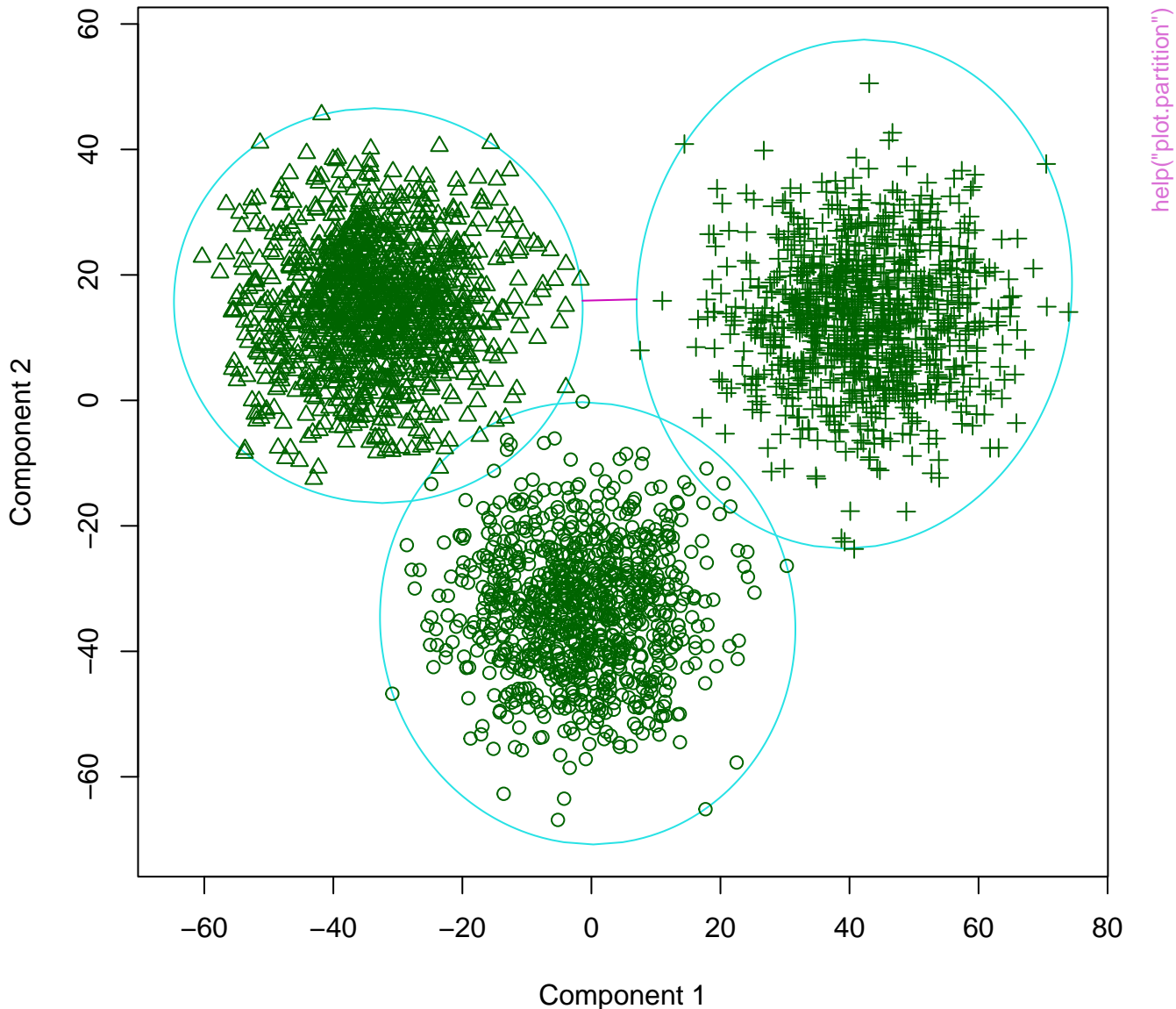
2 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$



Average silhouette width : 0.9

clusplot(clara(x = xclara, k = 3, keep.data = FALSE))



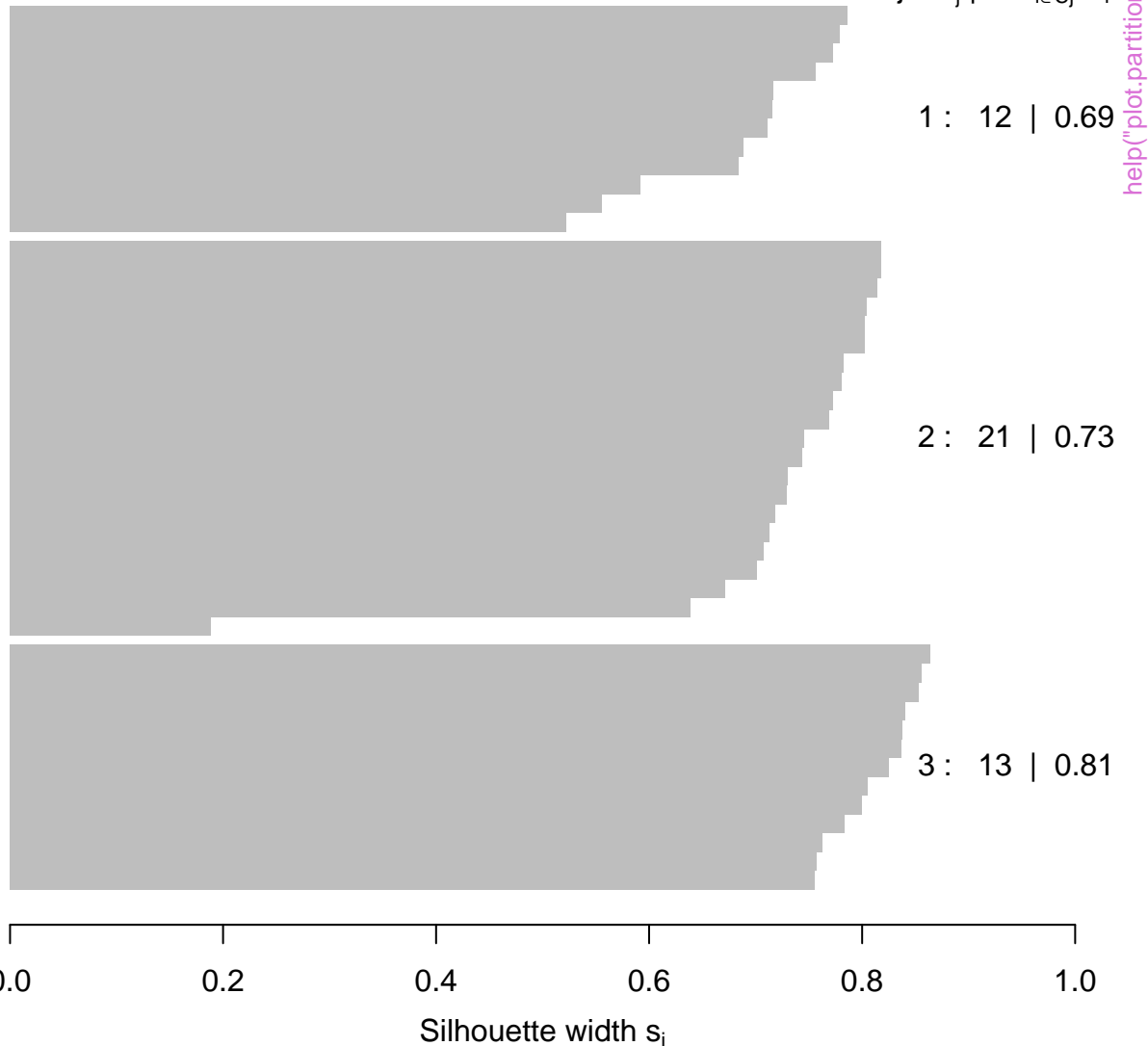
These two components explain 100 % of the point variability.

Silhouette plot of clara(x = xclara, k = 3, keep.data = FALSE)

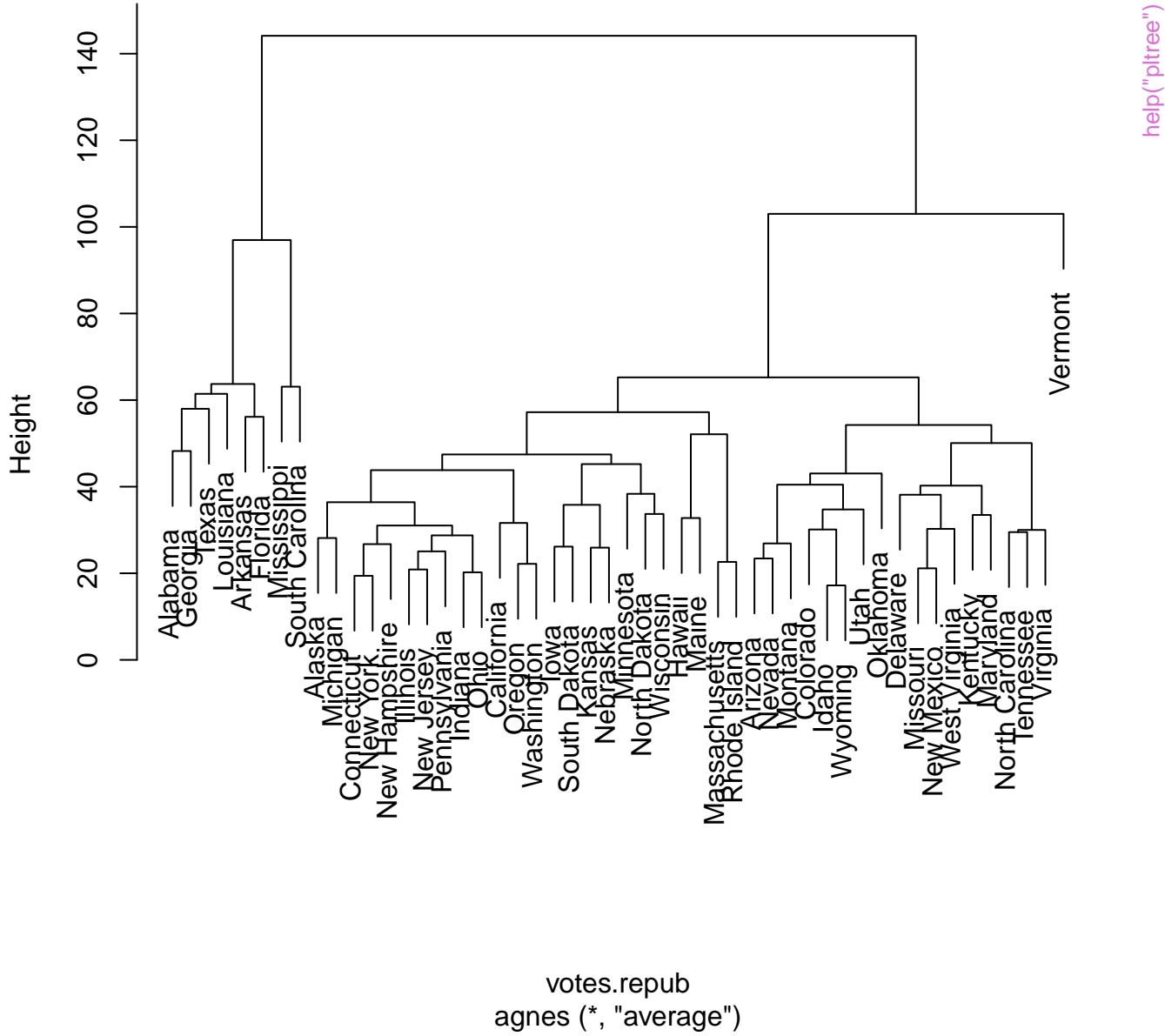
n = 46

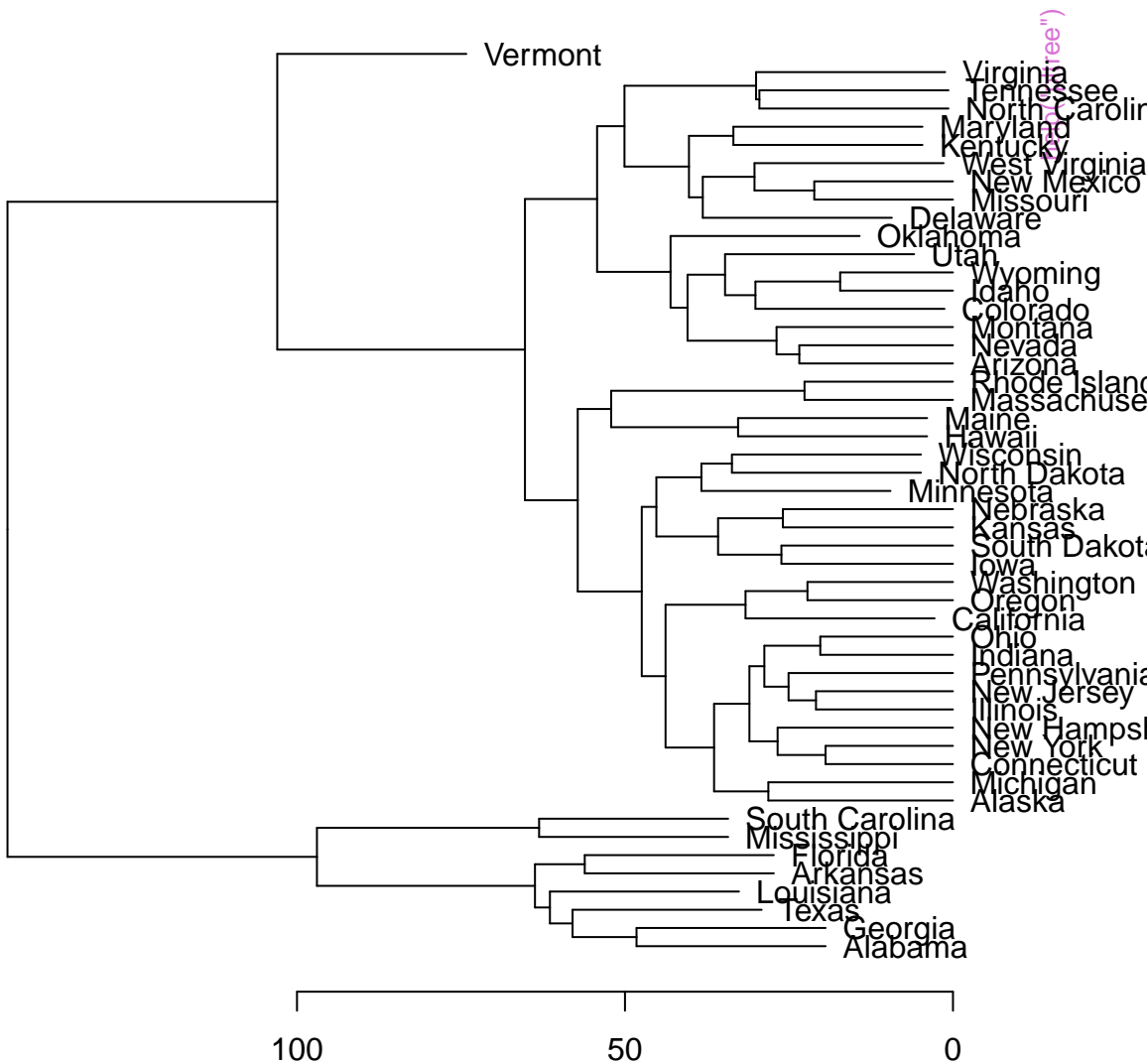
3 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$

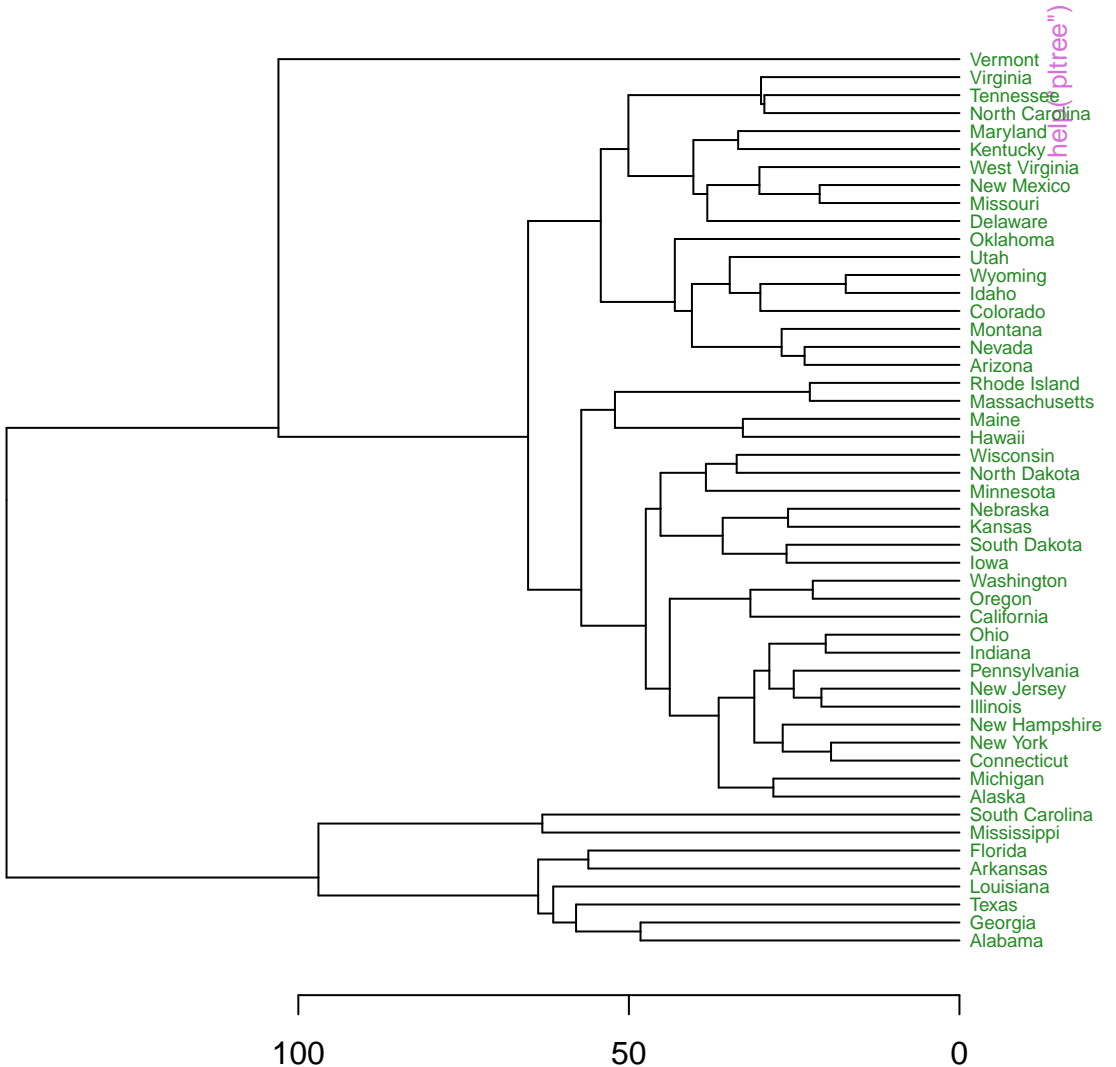


Dendrogram of agnes(x = votes.repub)

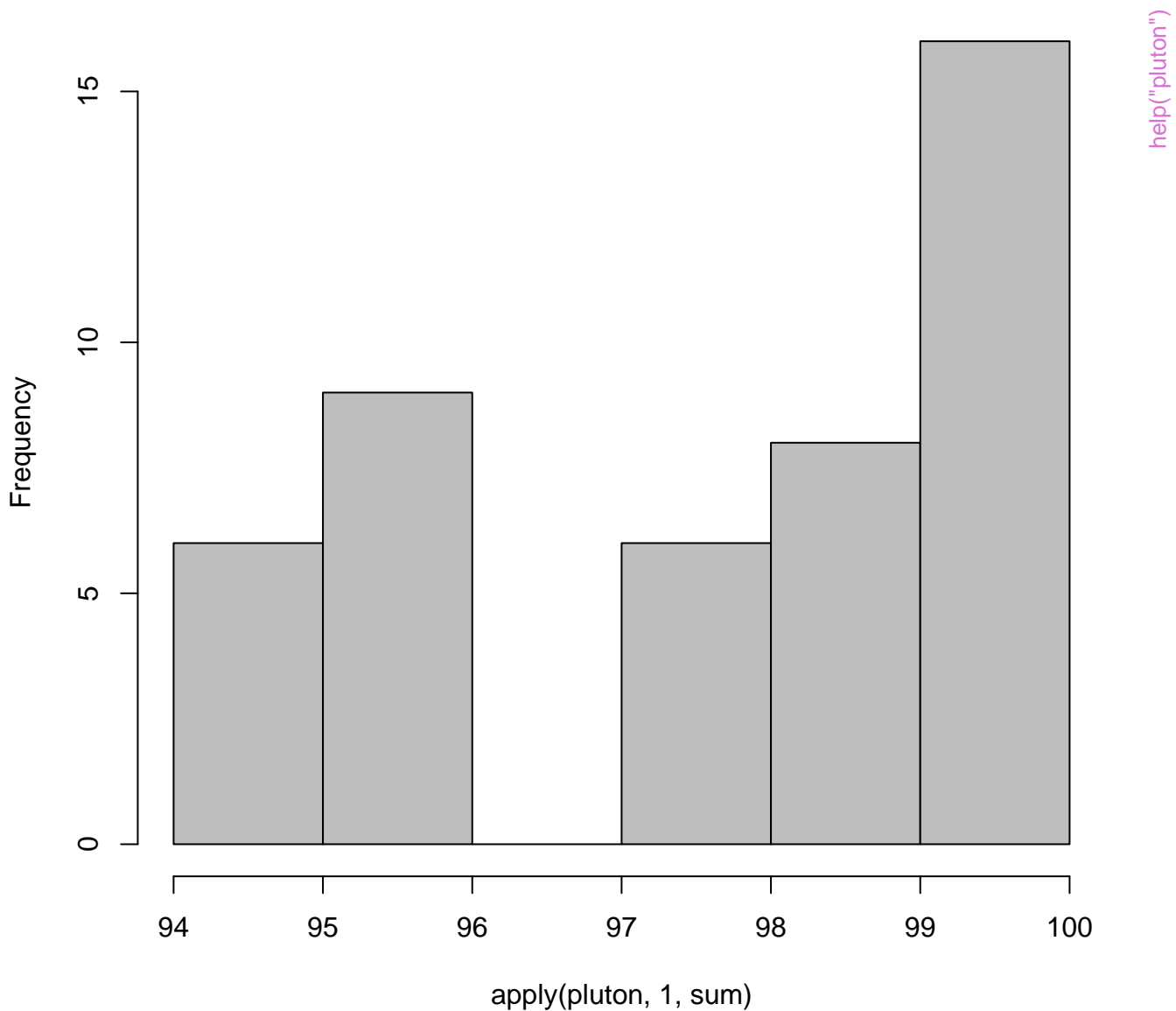


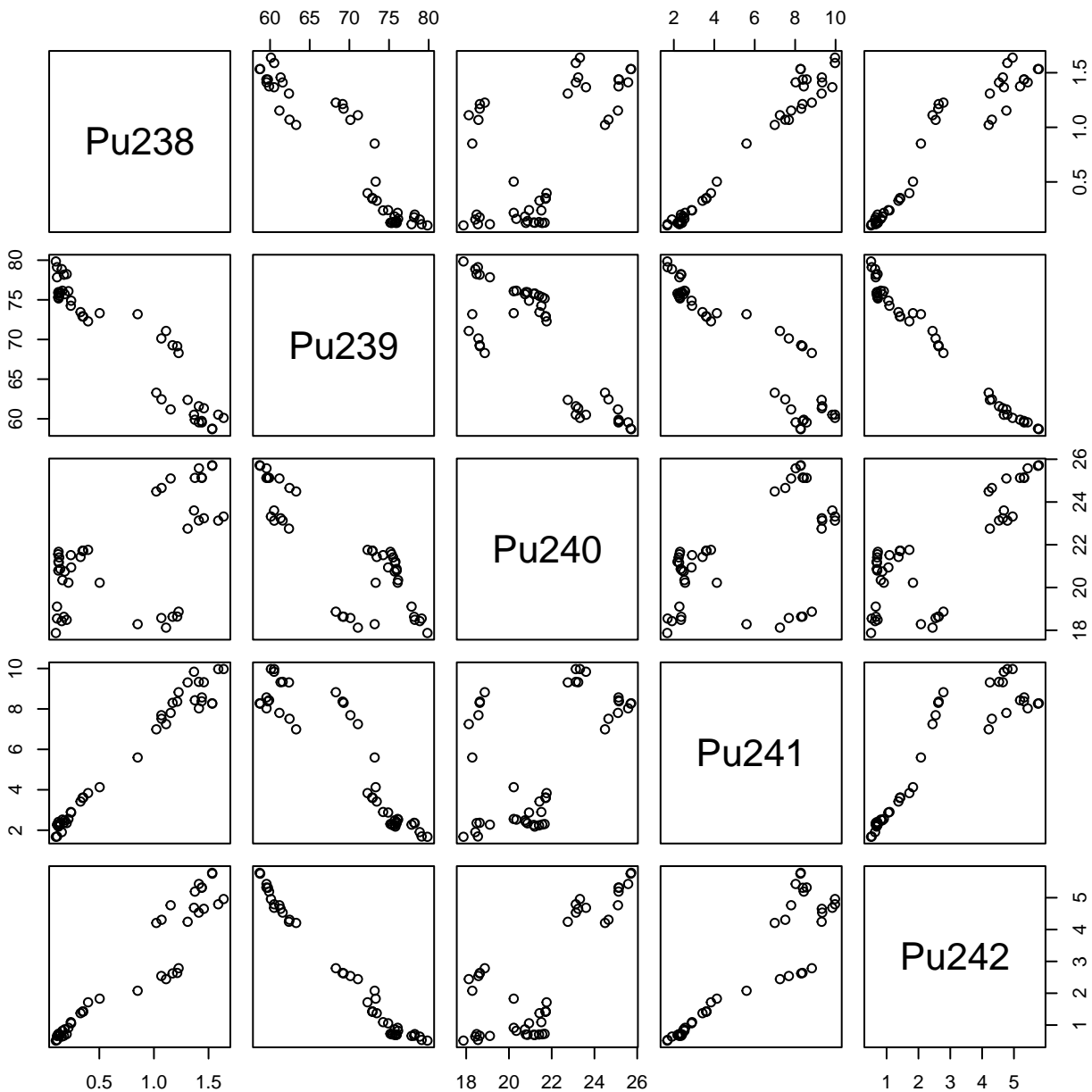


agnes(x = votes.repub)



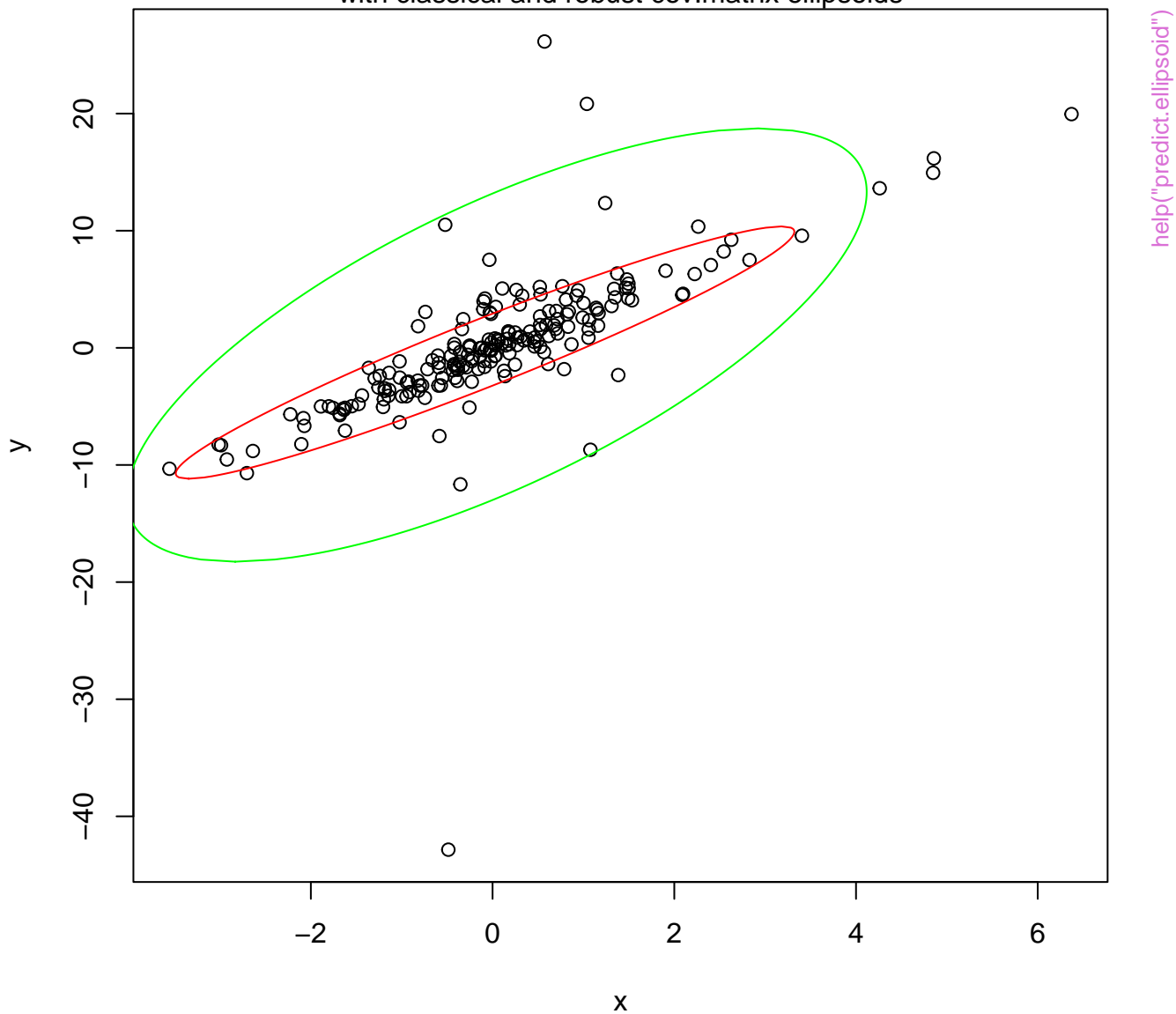
Histogram of `apply(pluton, 1, sum)`



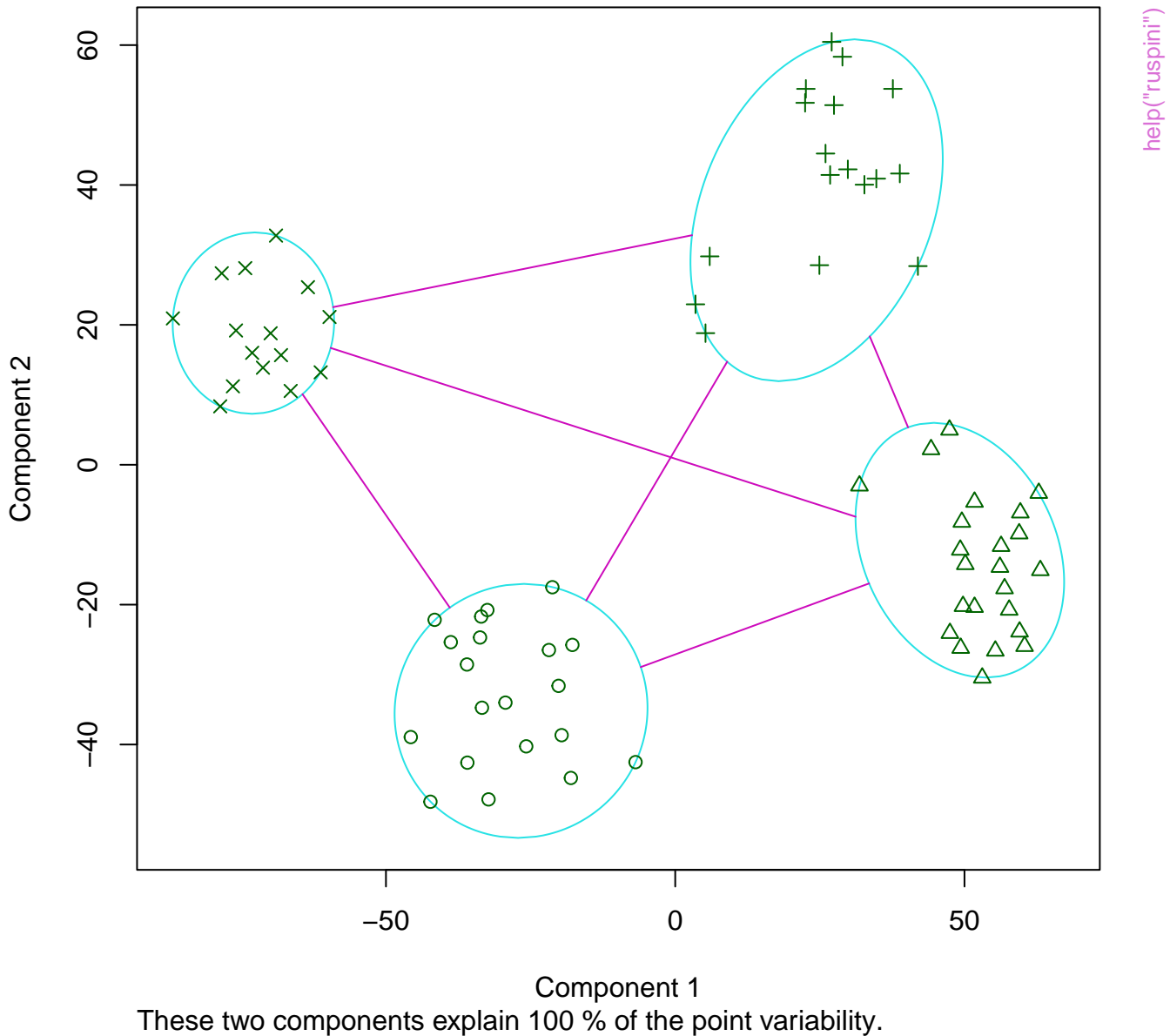


non-normal data (N=200)

with classical and robust cov.matrix ellipsoids



clusplot(pam(x = ruspini, k = 4))

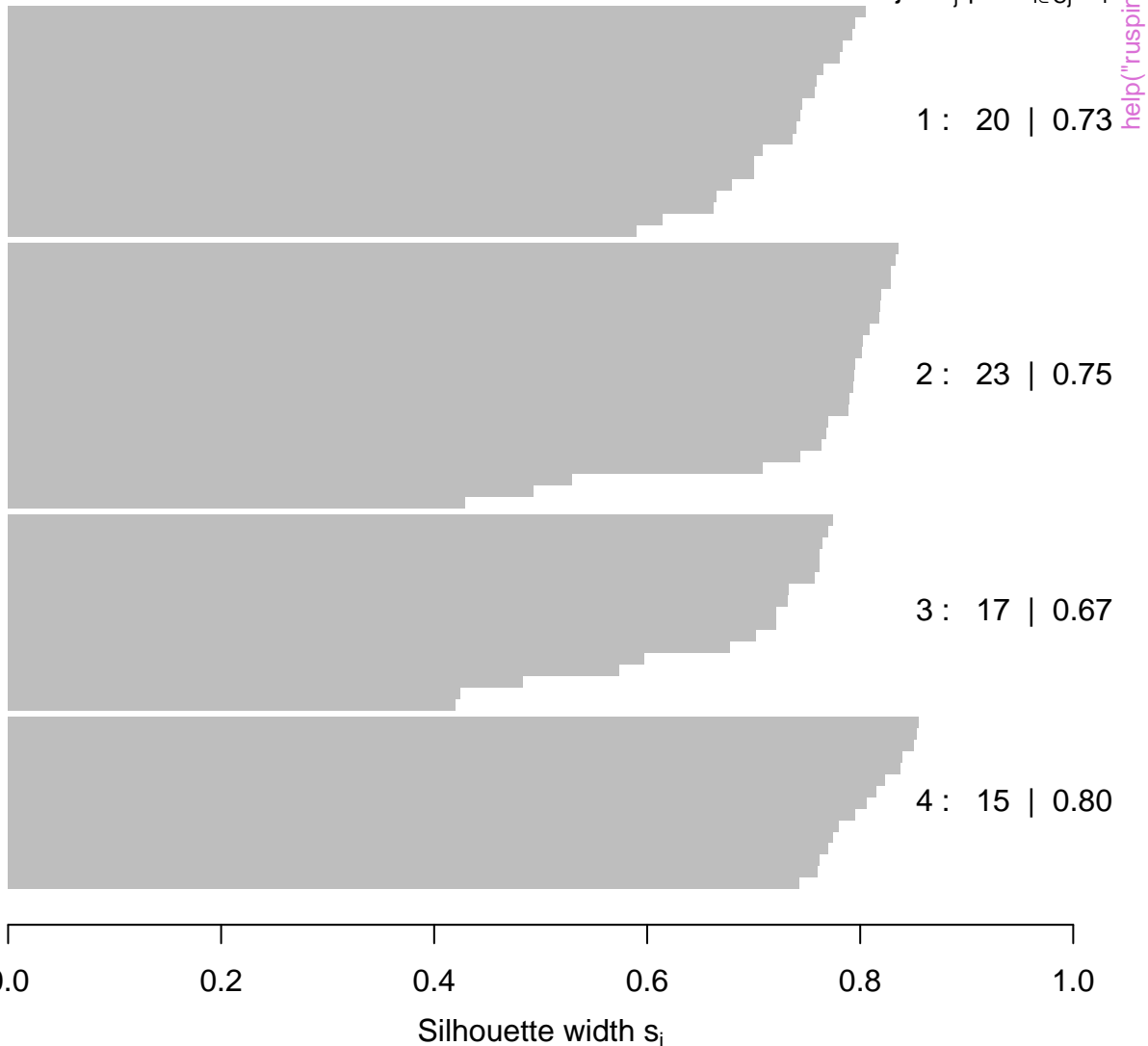


Silhouette plot of pam(x = ruspini, k = 4)

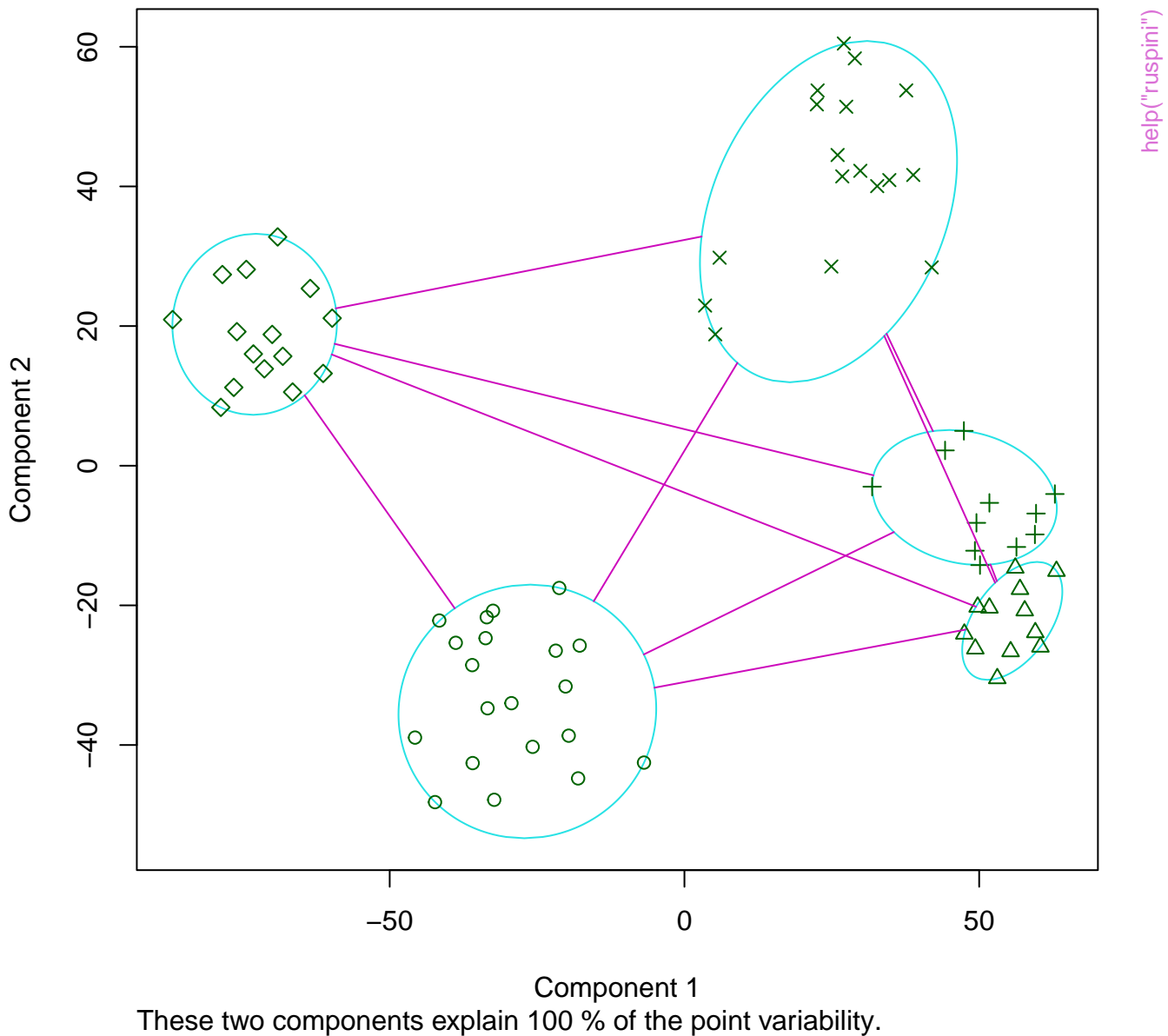
n = 75

4 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$



clusplot(fanny(x = ruspini, k = 5))

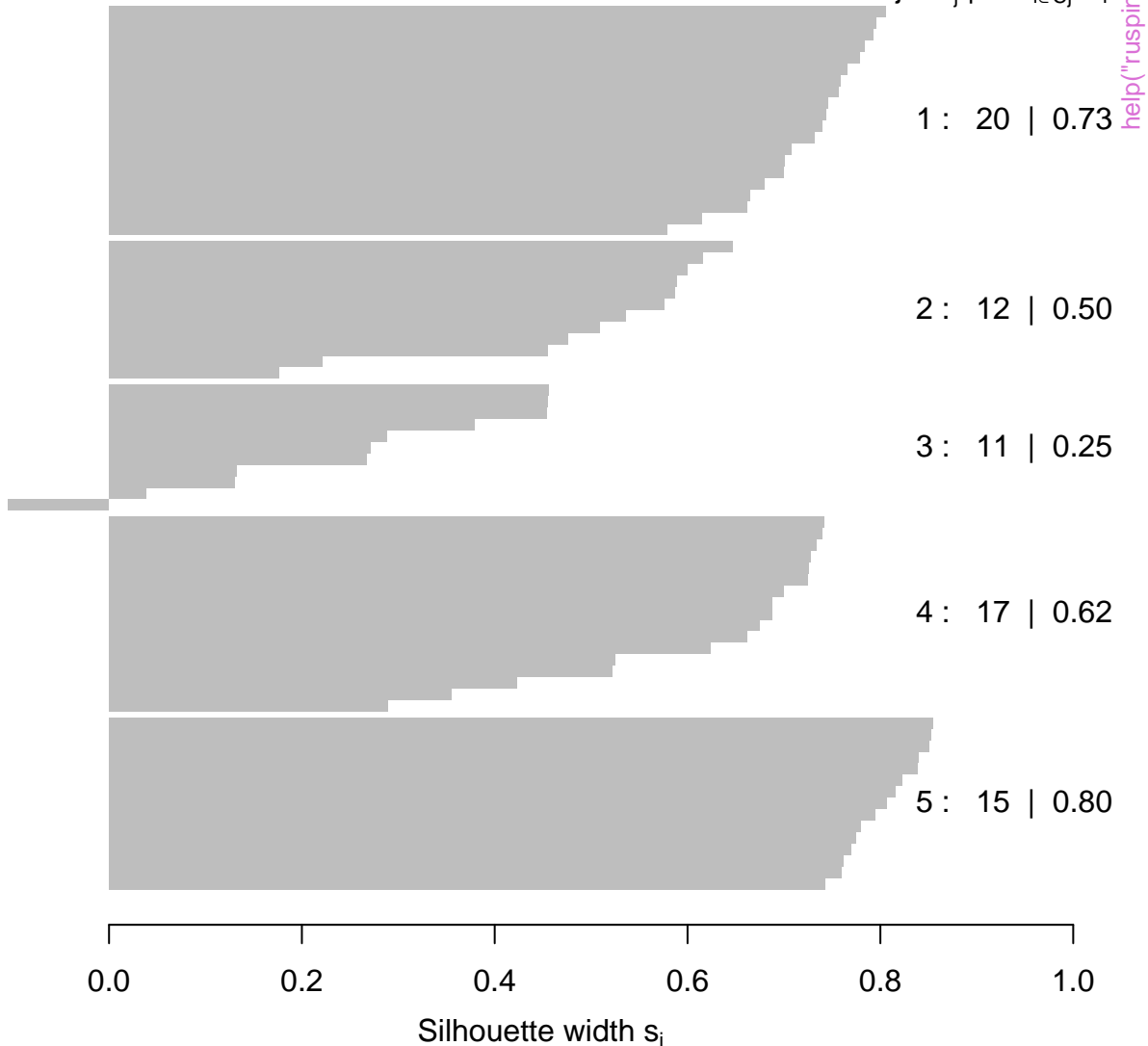


Silhouette plot of fanny(x = ruspini, k = 5)

n = 75

5 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$



Silhouette plot of pam(x = ruspini, k = 4)

n = 75

4 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$

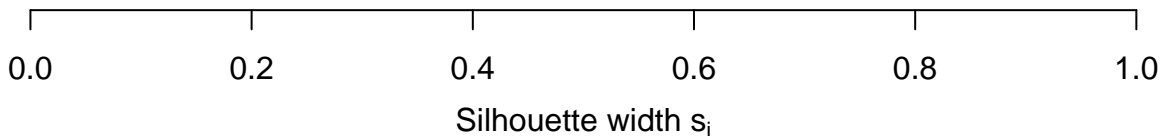
[help\("silhouette"\)](#)

1 : 20 | 0.73

2 : 23 | 0.75

3 : 17 | 0.67

4 : 15 | 0.80



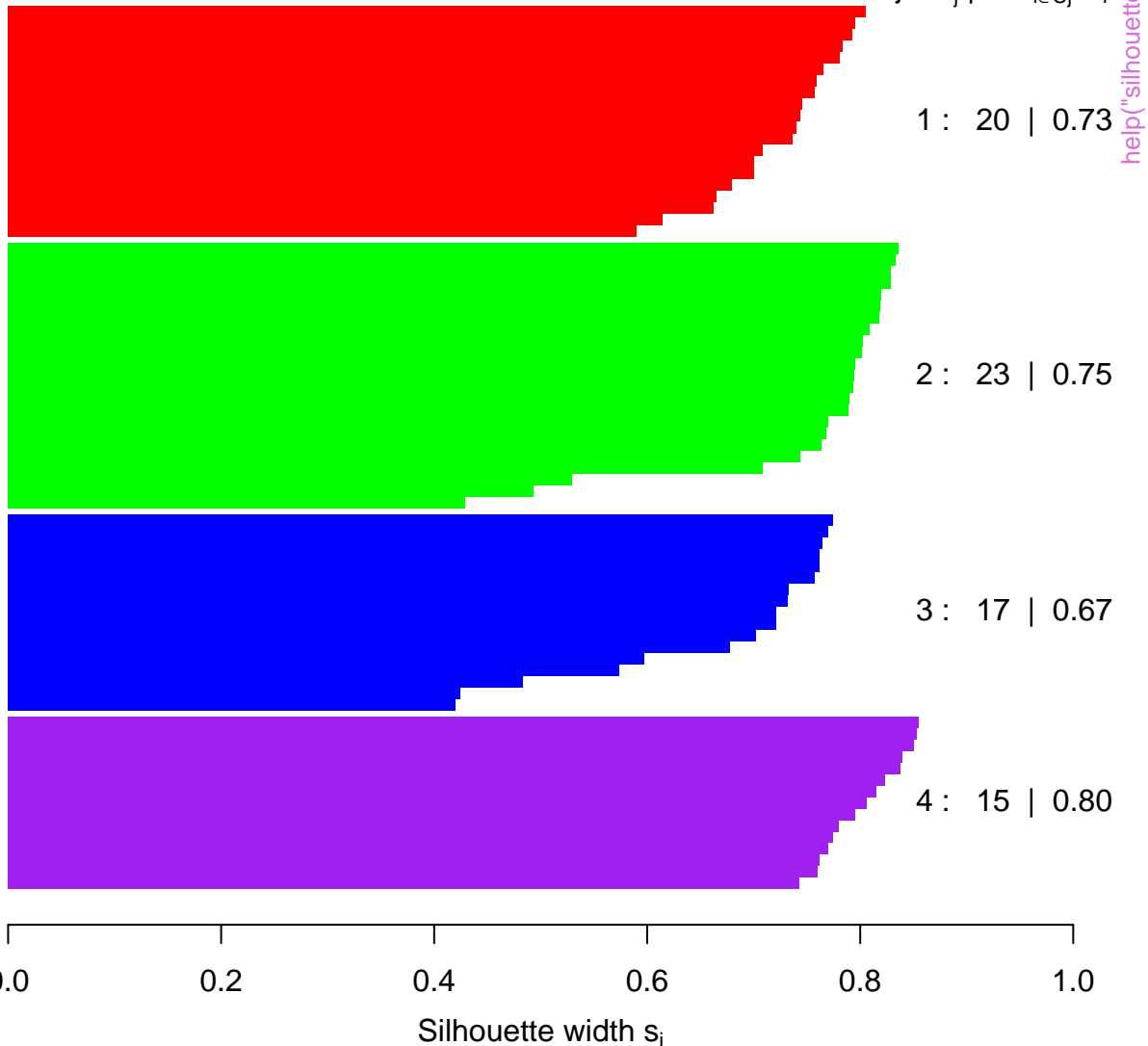
Average silhouette width : 0.74

Silhouette plot of pam(x = ruspini, k = 4)

n = 75

4 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$



Silhouette plot of (x = pr4\$clustering, dist = dist(ruspini, "canberra"

n = 75

4 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$

10
3
14
15
4
2
12
17
19
20
26
25
28
31
23
24
36
29
39
21
41
43
54
57
55
49
59
53
45
44
48
67
70
71
66
68
62
75
61

1 : 20 | 0.47

2 : 23 | 0.67

3 : 17 | 0.73

4 : 15 | 0.66

[help\("silhouette"\)](#)

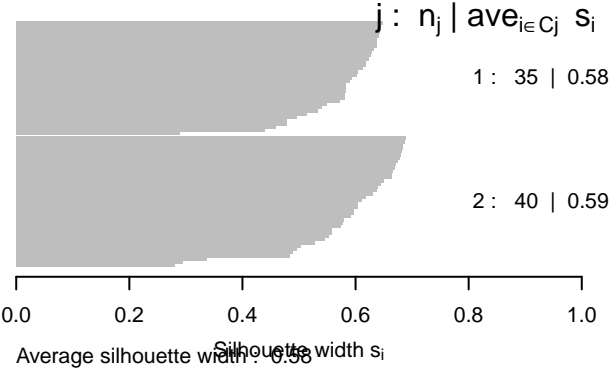


Silhouette width s_i

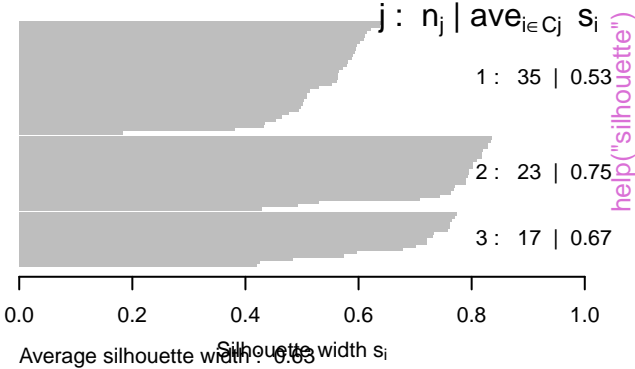
Average silhouette width : 0.63

PAM(Ruspini) as in Kaufman & Rousseeuw, p.101

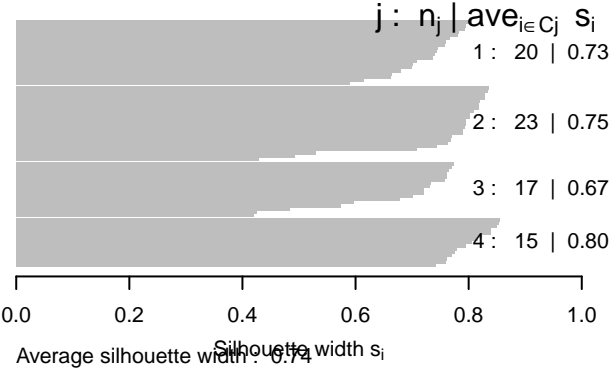
k = 2



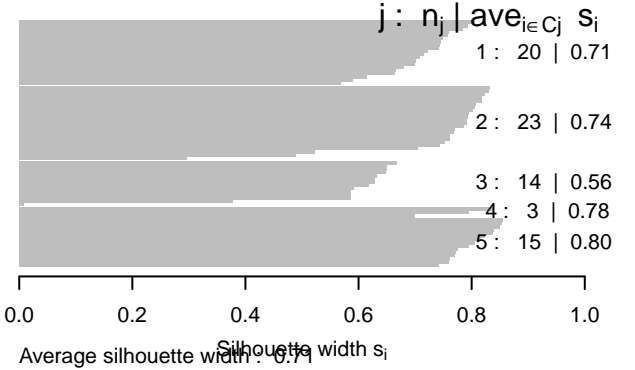
k = 3



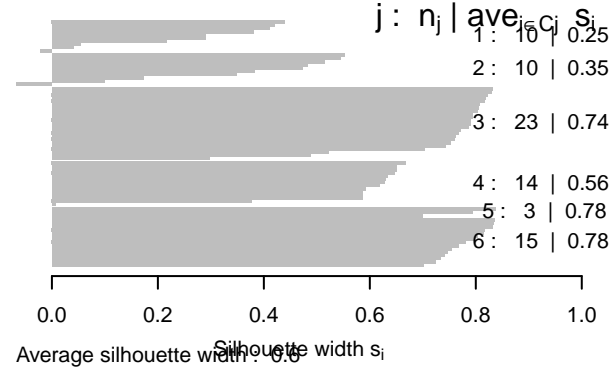
k = 4



k = 5

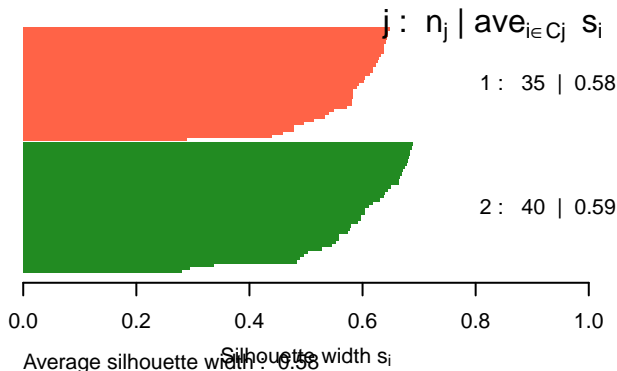


k = 6

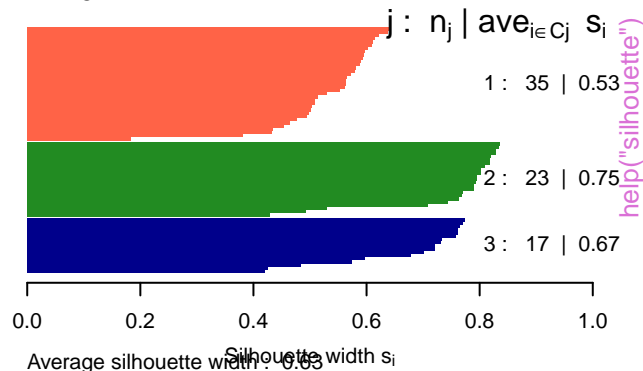


help("silhouette")

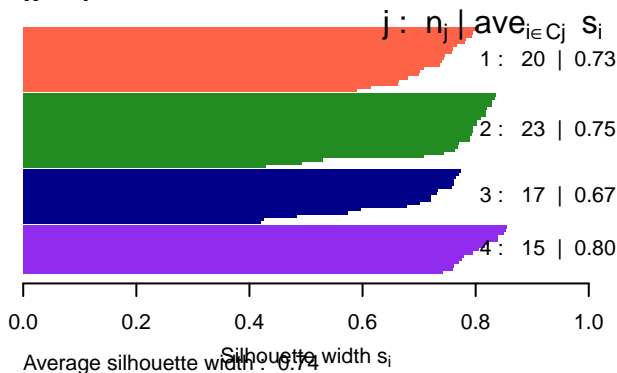
k = 2



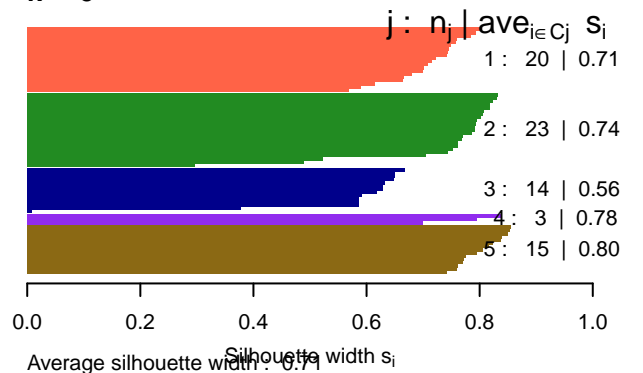
k = 3



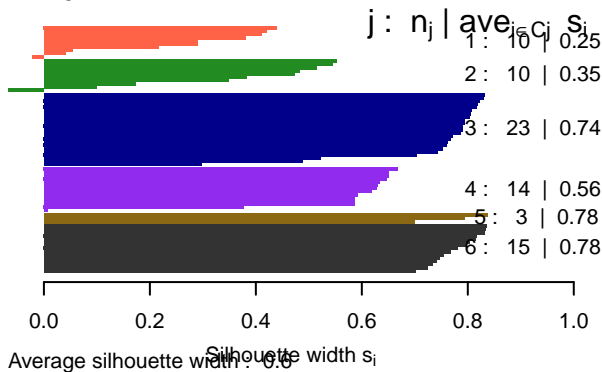
k = 4



k = 5



k = 6



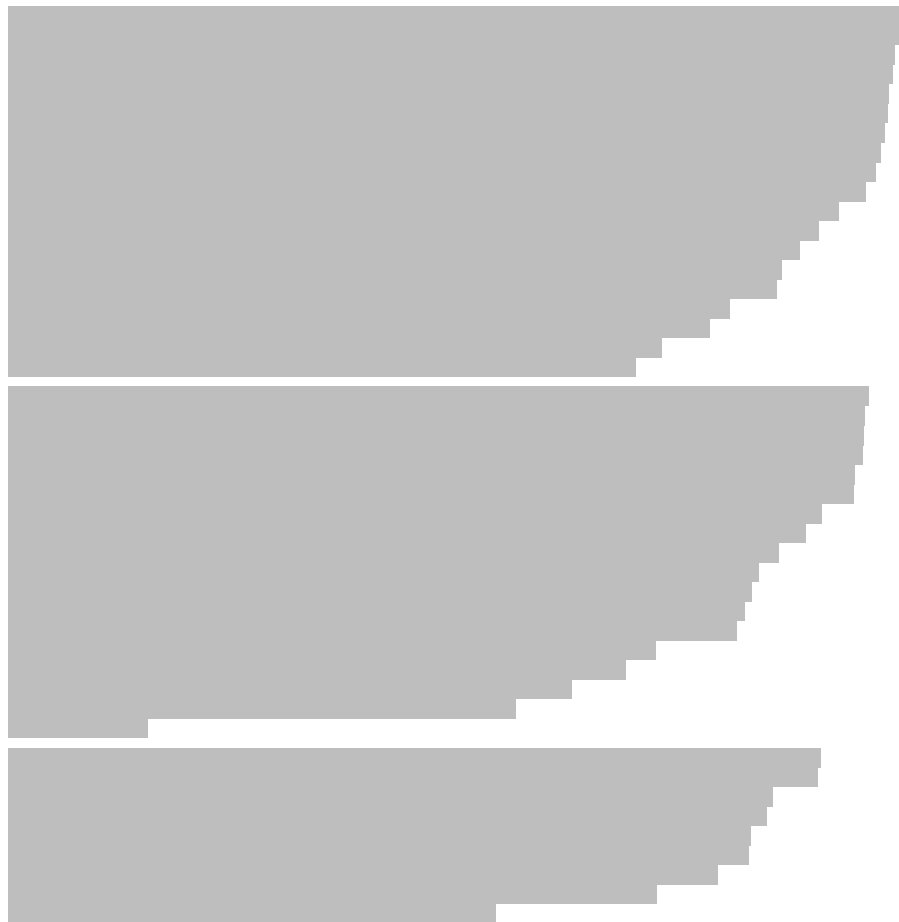
help("silhouette")

Silhouette plot of clara(x = xc1k, k = 3)

n = 46

3 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$



1 : 19 | 0.74

2 : 18 | 0.65

3 : 9 | 0.65

0.0

0.2

0.4

0.6

0.8

1.0

Silhouette width s_i

Average silhouette width : 0.69

help("silhouette")

plot(silhouette(clara(.), full = TRUE))

n = 1000

3 clusters C_j

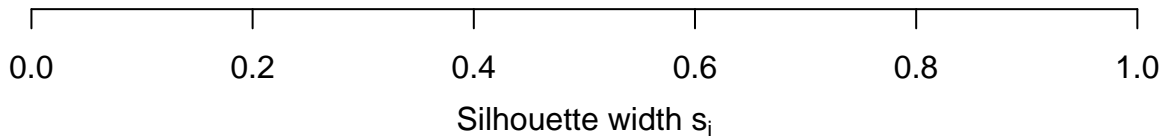
$j : n_j \mid \text{ave}_{i \in C_j} s_i$

help("silhouette")

1 : 367 | 0.68

2 : 306 | 0.67

3 : 327 | 0.71



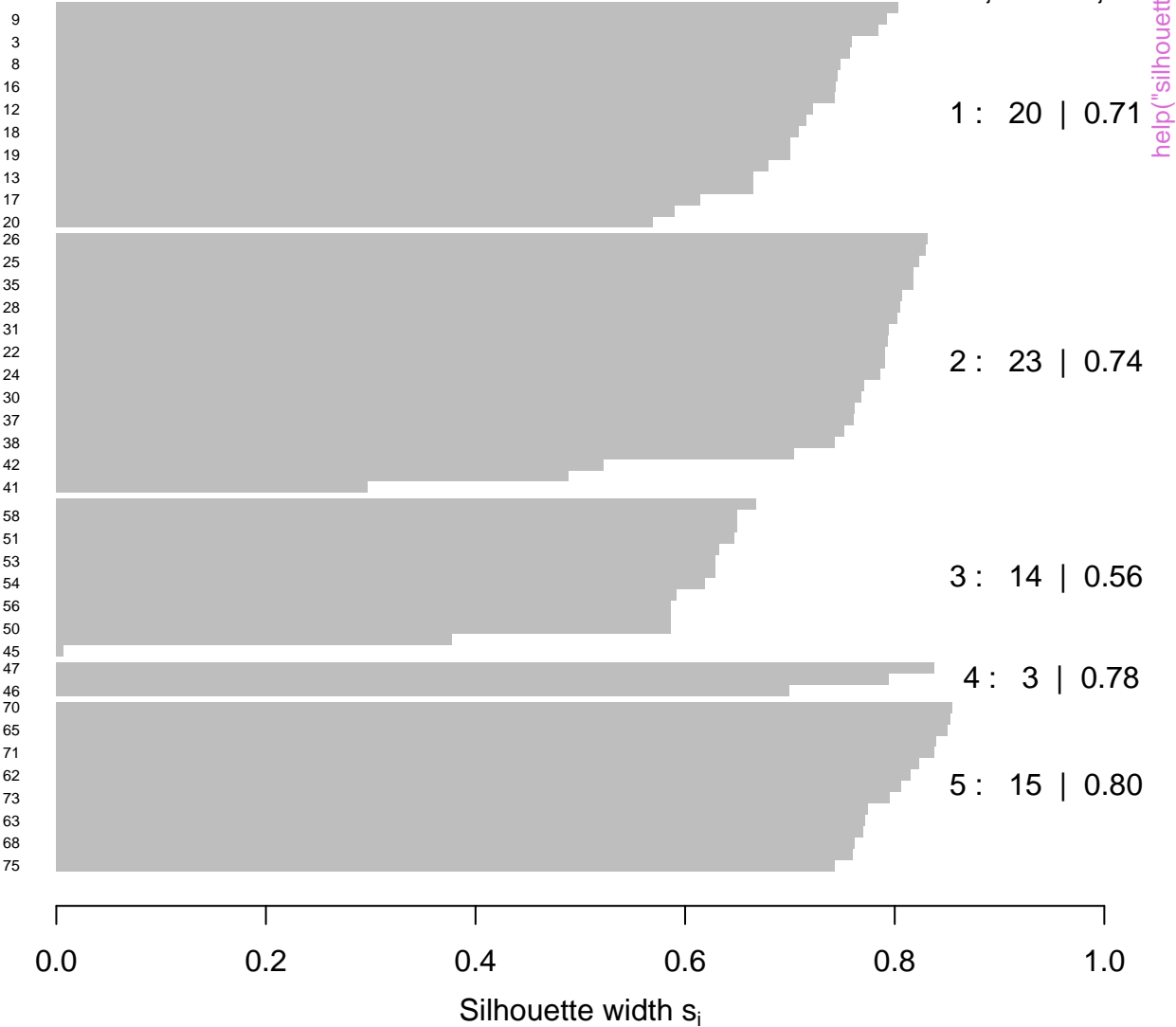
Average silhouette width : 0.69

Silhouette plot of (x = cutree(ar, k = 5), dist = daisy(ruspini))

n = 75

5 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$



Silhouette plot of (x = cutree(ar, k = 2), dist = daisy(ruspini))

n = 75

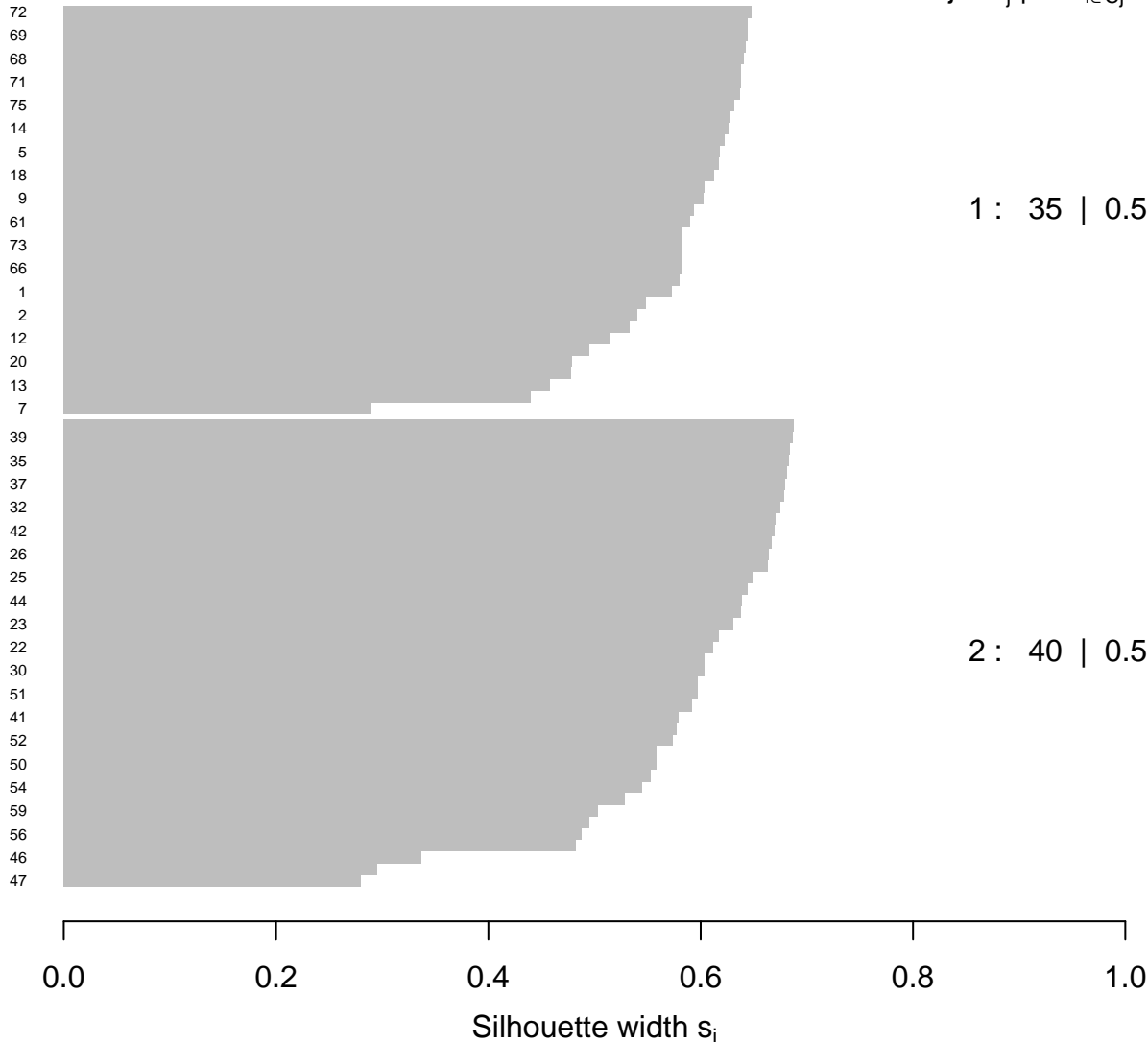
2 clusters C_j

$j : n_j \mid \text{ave}_{i \in C_j} s_i$

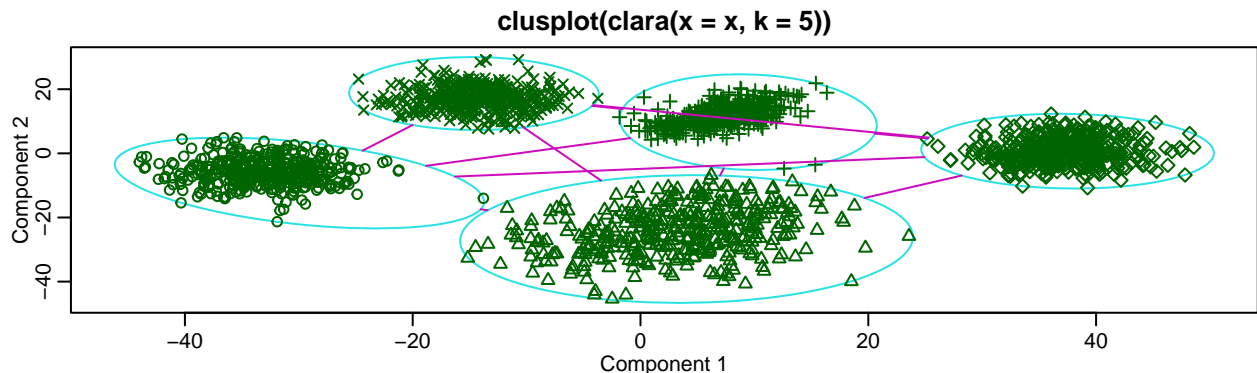
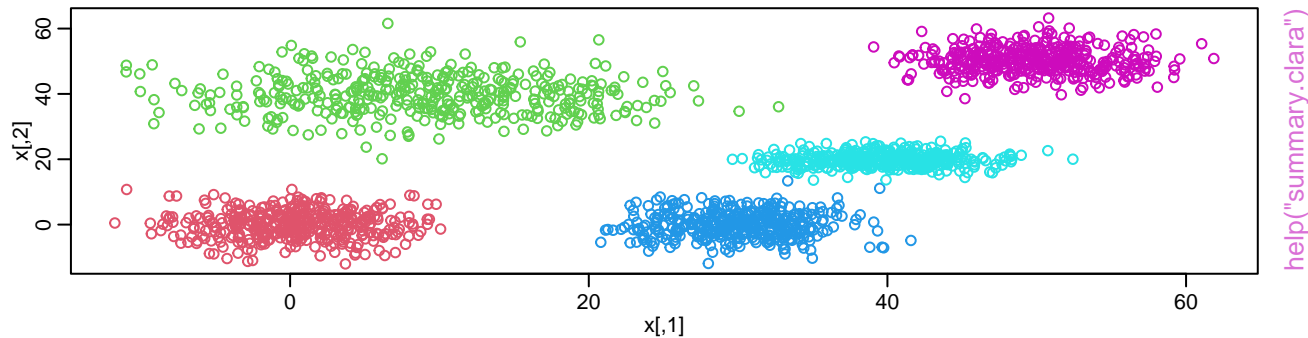
help("silhouette")

1 : 35 | 0.58

2 : 40 | 0.59



Average silhouette width : 0.58



These two components explain 100 % of the point variability.

Silhouette plot of clara(x = x, k = 5)

