Parent and Provider Perceptions of Behavioral Healthcare in Pediatric Primary Care (PI: Andrew Riley; BDP2-262)

Benjamin Chan (chanb@ohsu.edu)

2018-07-12

# Import Andrew’s SPSS data

Map new names to variables.

|  |  |
| --- | --- |
| oldnames | newnames |
| record\_id | id |
| eng\_span | languageSurvey |
| children\_totv\_1 | totalChildren |
| oldest\_middle\_youngest | birthOrder |
| child\_sexv\_1 | childSex |
| child\_age\_years | childAge |
| child\_ethnicity | childEthnicity |
| child\_racev\_1\_\_\_1 | childRaceWhite |
| child\_racev\_1\_\_\_2 | childRaceAsian |
| child\_racev\_1\_\_\_3 | childRaceAfrAm |
| child\_racev\_1\_\_\_4 | childRaceAIAN |
| child\_racev\_1\_\_\_5 | childRaceNHPI |
| child\_racev\_1\_\_\_6 | childRaceOther |
| child\_racev\_1\_\_\_7 | childRaceNoResp |
| related\_child | childRelationship |
| gender | parentGender |
| parent\_sexv\_1 | parentSex |
| parent\_agev\_1 | parentAge |
| parent\_ethnicity | parentEthnicity |
| parent\_race\_\_\_1 | parentRaceWhite |
| parent\_race\_\_\_2 | parentRaceAsian |
| parent\_race\_\_\_3 | parentRaceAfrAm |
| parent\_race\_\_\_4 | parentRaceAIAN |
| parent\_race\_\_\_5 | parentRaceNHPI |
| parent\_race\_\_\_6 | parentRaceOther |
| parent\_race\_\_\_7 | parentRaceNoResp |
| marital\_status | parentMaritalStatus |
| parenting\_situationv\_1 | parentSituation |
| number\_parents | parentsNumber |
| parent\_to\_child\_ratio | parentChildRatio |
| zipcode\_classification\_combined | zipcodeClass |
| zipcode | zipcode |
| community\_type | community |
| distance | distance |
| parent\_educationv\_1 | parentEducation |
| annual\_income | income |
| internet | internet |
| ECBI\_intensity\_raw\_score | ECBI\_intensity\_raw\_score |
| ECBI\_intensity\_T\_score | ECBI\_intensity\_T\_score |
| ECBI\_intensity\_clinical\_cutoff | ECBI\_intensity\_clinical\_cutoff |
| ECBI\_problem\_raw\_score | ECBI\_problem\_raw\_score |
| ECBI\_problem\_T\_score | ECBI\_problem\_T\_score |
| ECBI\_problem\_clinical\_cutoff | ECBI\_problem\_clinical\_cutoff |
| ECBI\_Opp | ECBI\_Opp |
| ECBI\_Inatt | ECBI\_Inatt |
| ECBI\_Cond | ECBI\_Cond |
| MAPS\_PP | MAPS\_PP |
| MAPS\_PR | MAPS\_PR |
| MAPS\_WM | MAPS\_WM |
| MAPS\_SP | MAPS\_SP |
| MAPS\_HS | MAPS\_HS |
| MAPS\_LC | MAPS\_LC |
| MAPS\_PC | MAPS\_PC |
| MAPS\_POS | MAPS\_POS |
| MAPS\_NEG | MAPS\_NEG |
| SEPTI\_nurturance | SEPTI\_nurturance |
| SEPTI\_n\_clinical\_cutoff | SEPTI\_n\_clinical\_cutoff |
| SEPTI\_discipline | SEPTI\_discipline |
| SEPTI\_d\_clinical\_cutoff | SEPTI\_d\_clinical\_cutoff |
| SEPTI\_play | SEPTI\_play |
| SEPTI\_p\_clinical\_cutoff | SEPTI\_p\_clinical\_cutoff |
| SEPTI\_routine | SEPTI\_routine |
| SEPTI\_r\_clinical\_cutoff | SEPTI\_r\_clinical\_cutoff |
| SEPTI\_total | SEPTI\_total |
| SEPTI\_total\_clin\_cutoff | SEPTI\_total\_clin\_cutoff |
| PCB1\_Total | PCB1\_Total |
| PCB1\_CondEmot | PCB1\_CondEmot |
| PCB1\_DevHab | PCB1\_DevHab |
| PCB2\_Tot | PCB2\_Tot |
| PCB3\_Total | PCB3\_Total |
| PBC3\_PCPonly | PCB3\_PCPonly |
| PCB3\_Person | PCB3\_Person |
| PCB3\_Resource | PCB3\_Resource |

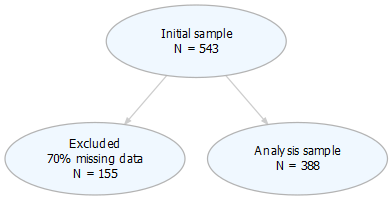
## Warning: package 'bindrcpp' was built under R version 3.4.4

Remove certain predictor variables:

* Clinical cutoffs
* Raw scores
* Total scores

## [1] "ECBI\_intensity\_raw\_score" "ECBI\_intensity\_clinical\_cutoff"  
## [3] "ECBI\_problem\_raw\_score" "ECBI\_problem\_clinical\_cutoff"   
## [5] "SEPTI\_n\_clinical\_cutoff" "SEPTI\_d\_clinical\_cutoff"   
## [7] "SEPTI\_p\_clinical\_cutoff" "SEPTI\_r\_clinical\_cutoff"   
## [9] "SEPTI\_total" "SEPTI\_total\_clin\_cutoff"

Build analysis data set. Exclude if missing any dependent variable, PCB1\_Total, PCB2\_Tot, PCB3\_Total. Exclude rows if there are a high proportion of row-wise NA.



figures/flowChart.png

# Cluster analysis

Use divisive hierarchical clustering (DIANA). See [Divisive Hierarchical Clustering Essentials](http://www.sthda.com/english/articles/28-hierarchical-clustering-essentials/94-divisive-hierarchical-clustering-essentials/).

## Warning: package 'cluster' was built under R version 3.4.4

## Warning: package 'factoextra' was built under R version 3.4.4

## Welcome! Related Books: `Practical Guide To Cluster Analysis in R` at https://goo.gl/13EFCZ

##   
## To cite package 'factoextra' in publications use:  
##   
## Alboukadel Kassambara and Fabian Mundt (2017). factoextra:  
## Extract and Visualize the Results of Multivariate Data Analyses.  
## R package version 1.0.5.  
## https://CRAN.R-project.org/package=factoextra  
##   
## A BibTeX entry for LaTeX users is  
##   
## @Manual{,  
## title = {factoextra: Extract and Visualize the Results of Multivariate Data Analyses},  
## author = {Alboukadel Kassambara and Fabian Mundt},  
## year = {2017},  
## note = {R package version 1.0.5},  
## url = {https://CRAN.R-project.org/package=factoextra},  
## }

Use the **manhattan** metric.

## Cluster on PCB metrics

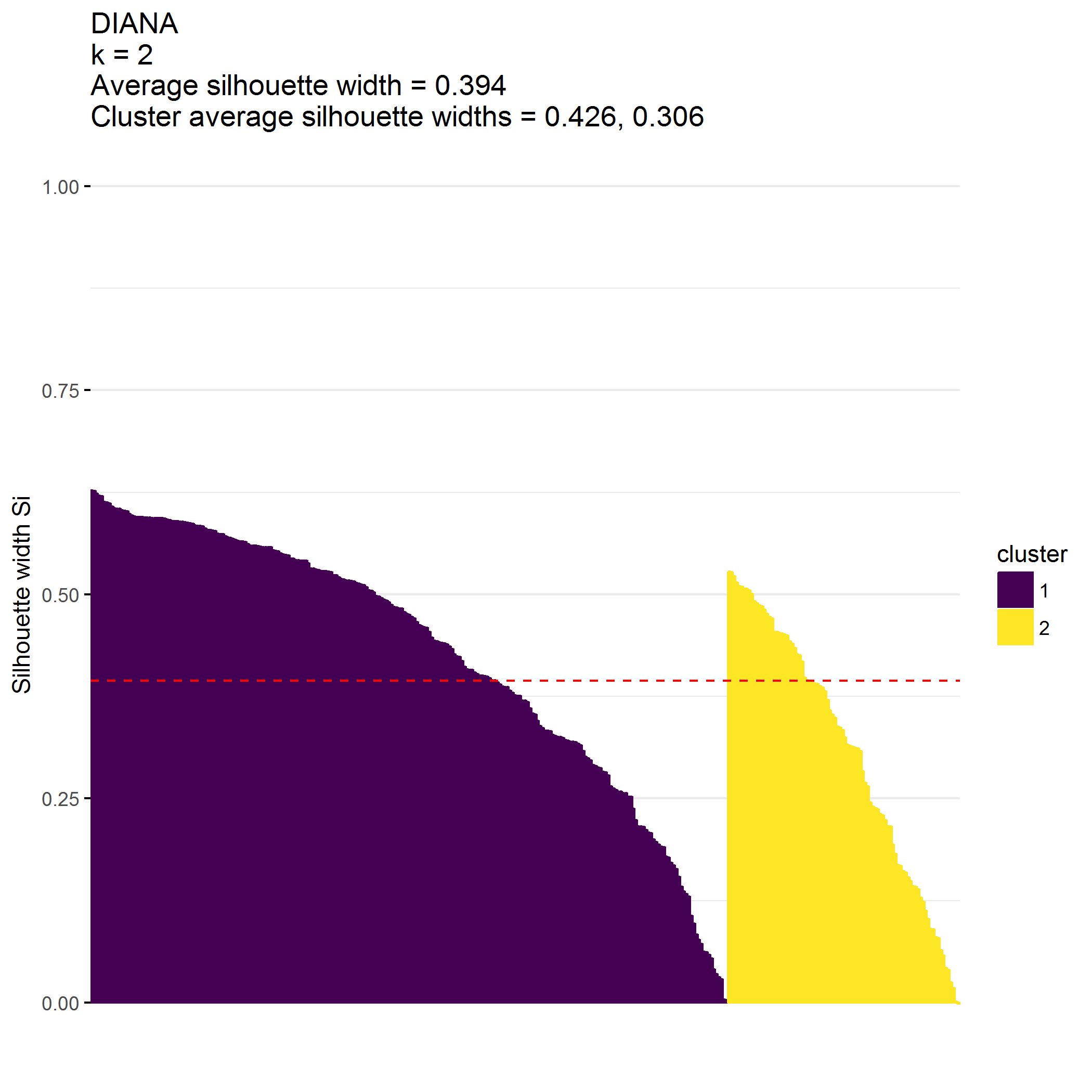
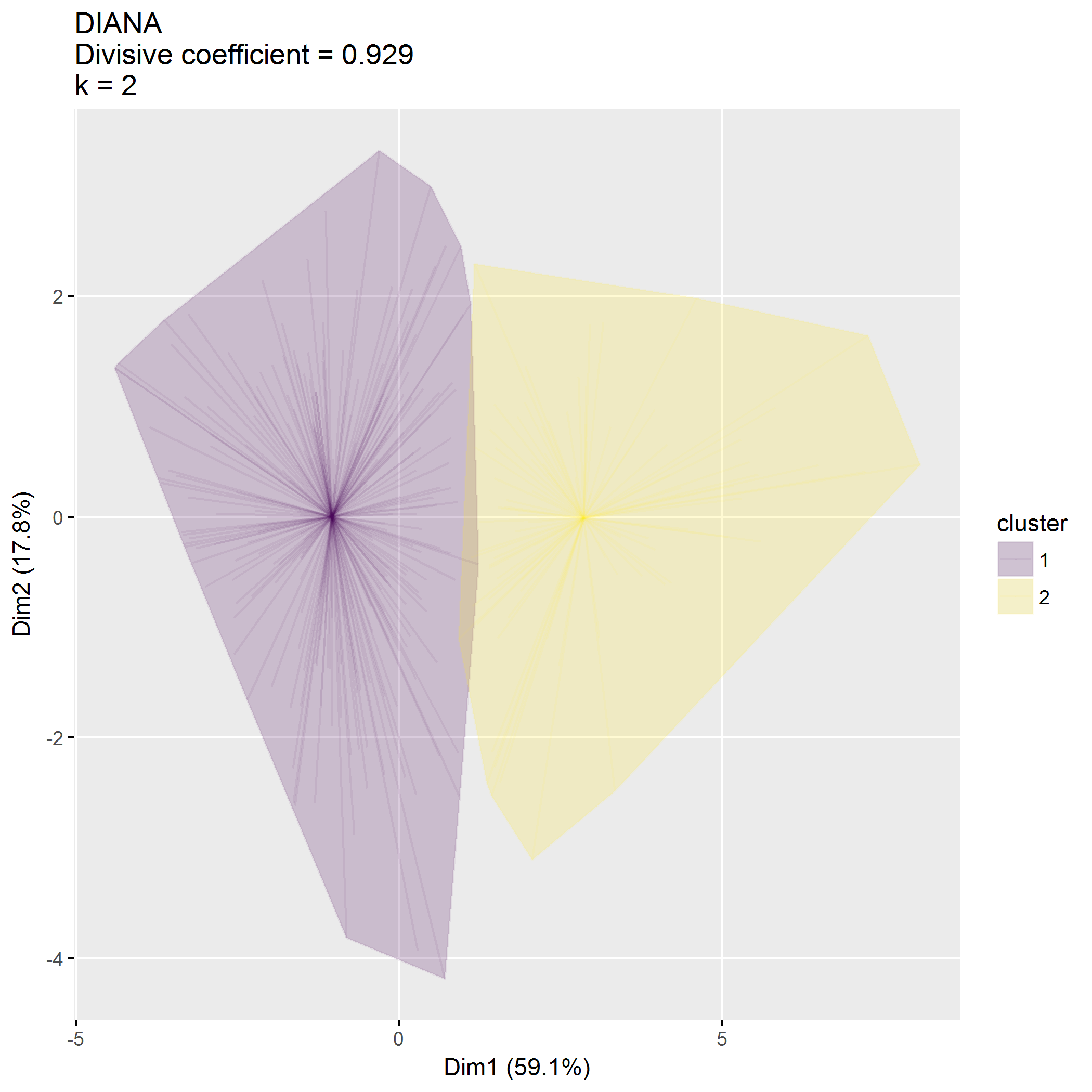
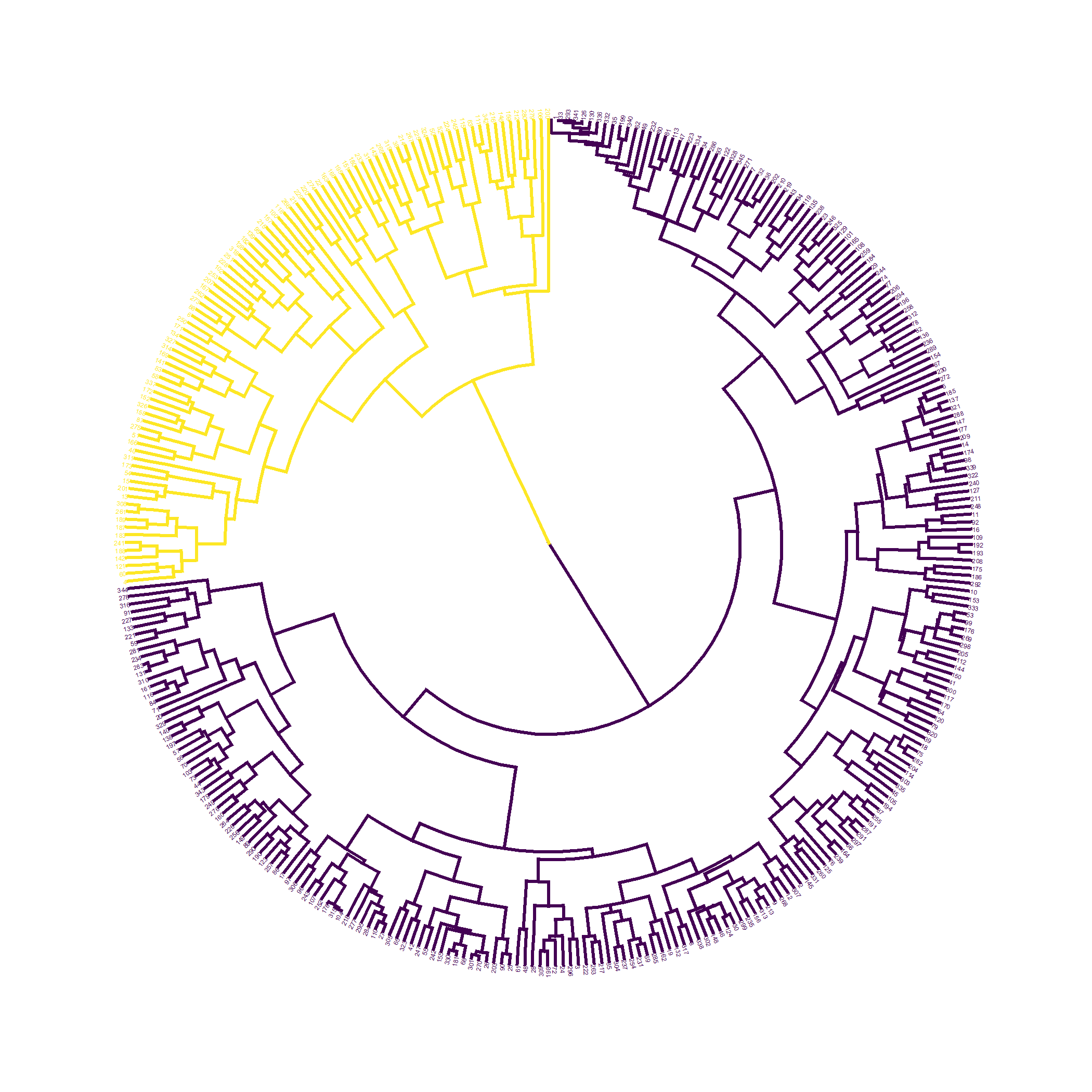
**Clustering on PCB metrics isn’t terrible.**

## [1] 345 8

## [1] "PCB1\_Total" "PCB1\_CondEmot" "PCB1\_DevHab" "PCB2\_Tot"   
## [5] "PCB3\_Total" "PCB3\_PCPonly" "PCB3\_Person" "PCB3\_Resource"

## cluster size ave.sil.width  
## 1 1 253 0.43  
## 2 2 92 0.31

* Hopkins statistic is 0.254
* Analysis identified clusters
* Divisive coefficient is 0.929
* Average silhouette width is 0.394



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| cluster | n | PCB1\_Total\_mean | PCB1\_CondEmot\_mean | PCB1\_DevHab\_mean |
| 1 | 253 | 74.0 | 53.7 | 20.3 |
| 2 | 92 | 47.1 | 33.6 | 13.5 |

|  |  |  |
| --- | --- | --- |
| cluster | n | PCB2\_Tot\_mean |
| 1 | 253 | 25.9 |
| 2 | 92 | 20.0 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| cluster | n | PCB3\_Total\_mean | PCB3\_PCPonly\_mean | PCB3\_Person\_mean | PCB3\_Resource\_mean |
| 1 | 253 | 51.7 | 4.4 | 17.4 | 30 |
| 2 | 92 | 34.2 | 3.3 | 11.9 | 19 |

* Cluster 1 () has high PCB scores on all domains
* Cluster 2 () has low PCB scores on all domains

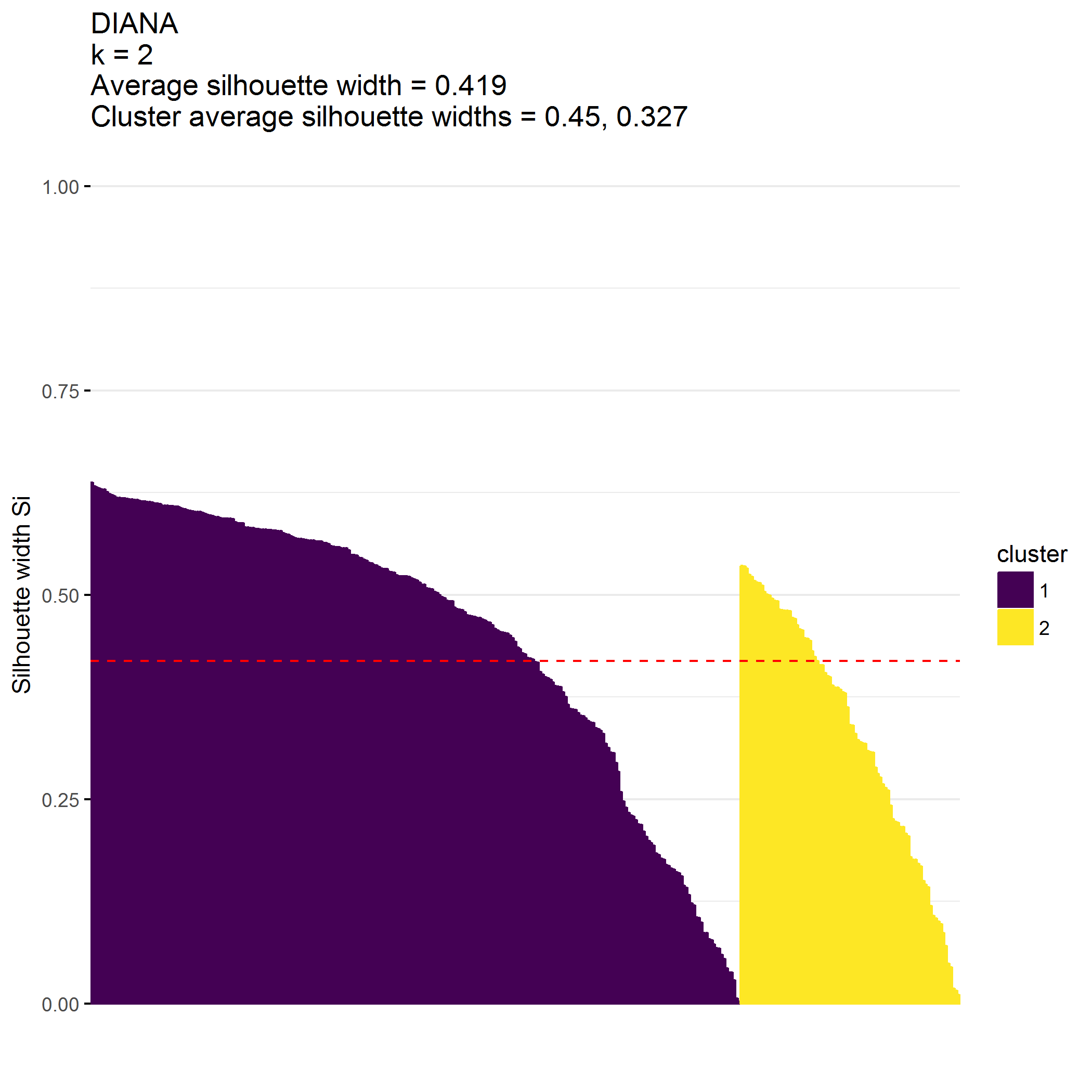
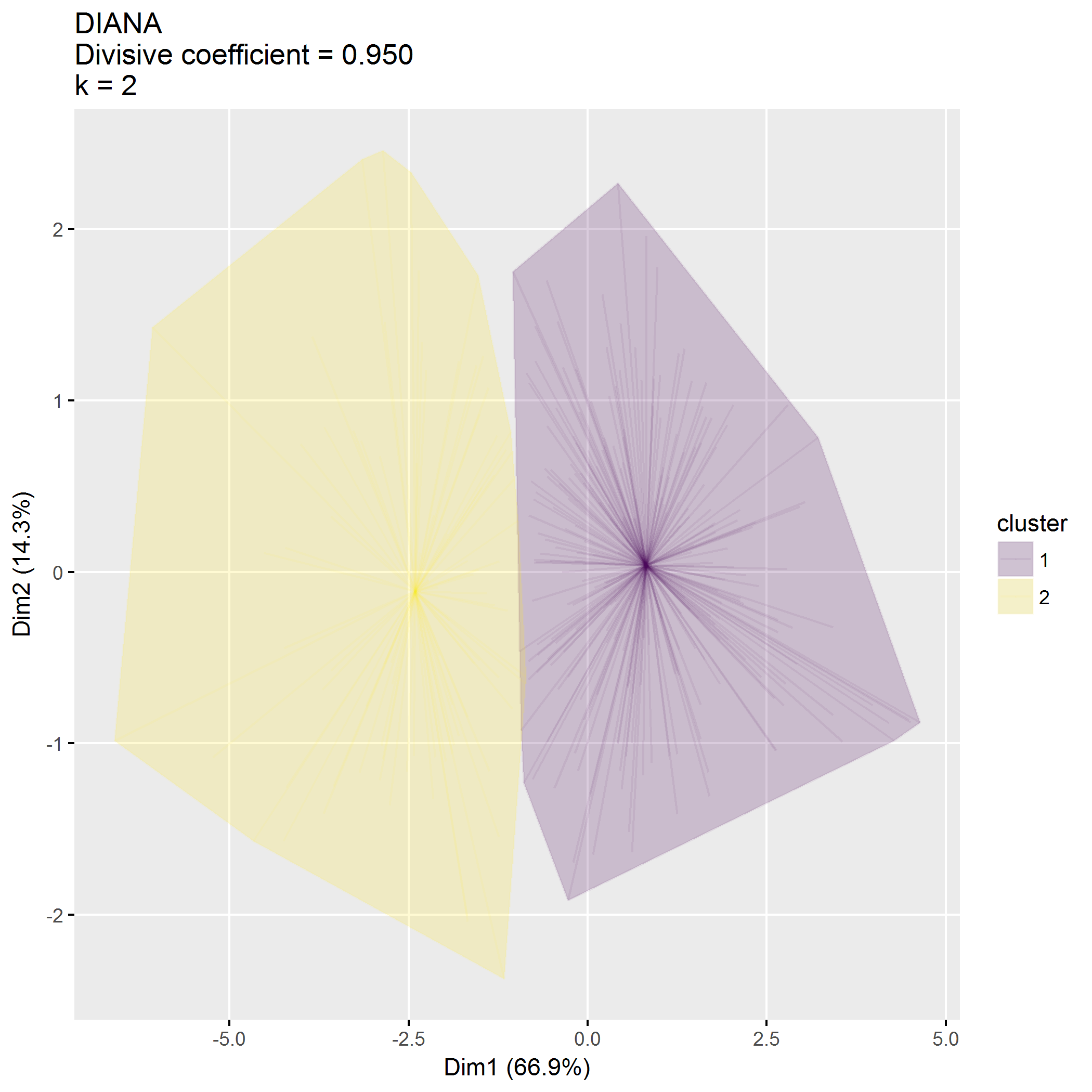
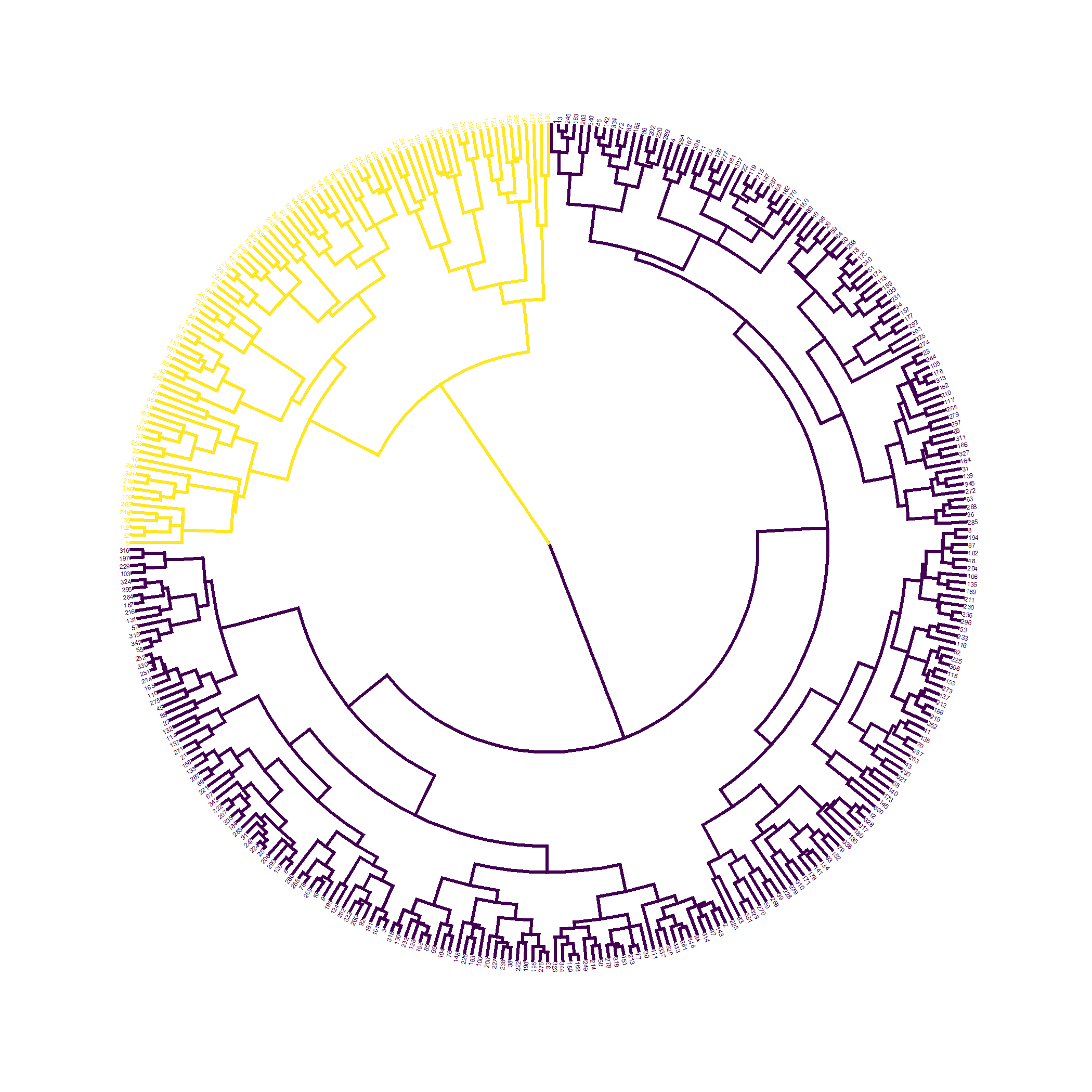
## Cluster on ECBI metrics

## [1] 345 5

## [1] "ECBI\_intensity\_T\_score" "ECBI\_problem\_T\_score"   
## [3] "ECBI\_Opp" "ECBI\_Inatt"   
## [5] "ECBI\_Cond"

## cluster size ave.sil.width  
## 1 1 258 0.45  
## 2 2 87 0.33

* Hopkins statistic is 0.197
* Analysis identified clusters
* Divisive coefficient is 0.950
* Average silhouette width is 0.419



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| cluster | n | ECBI\_intensity\_T\_score\_mean | ECBI\_problem\_T\_score\_mean | ECBI\_Opp\_mean | ECBI\_Inatt\_mean | ECBI\_Cond\_mean |
| 1 | 258 | 50.7 | 49.8 | 30.1 | 12.4 | 13.4 |
| 2 | 87 | 62.4 | 65.4 | 43.7 | 16.9 | 21.9 |

**Clustering on ECBI alone is good**

* Cluster 1 () has low ECBI scores
* Cluster 2 () has high ECBI scores

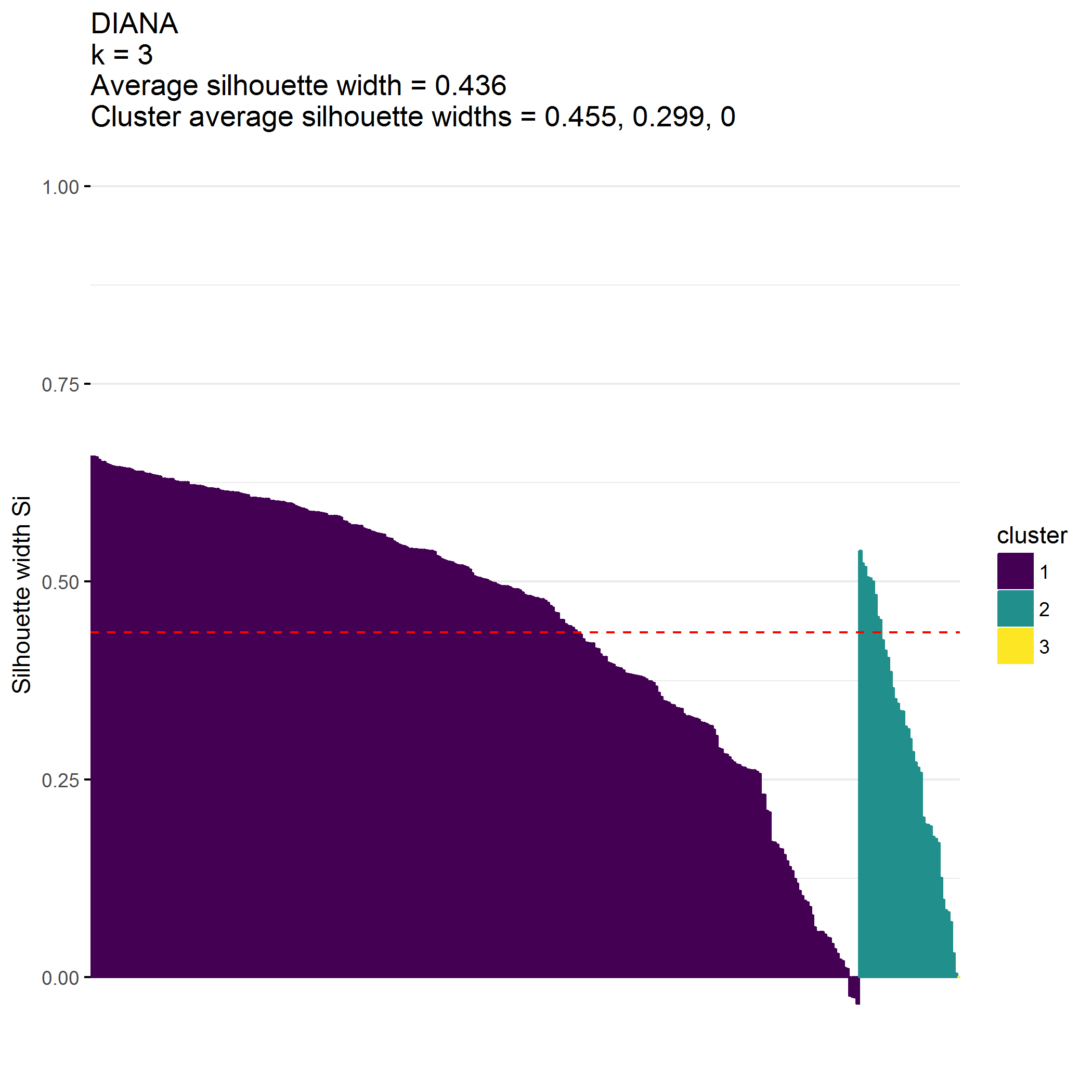
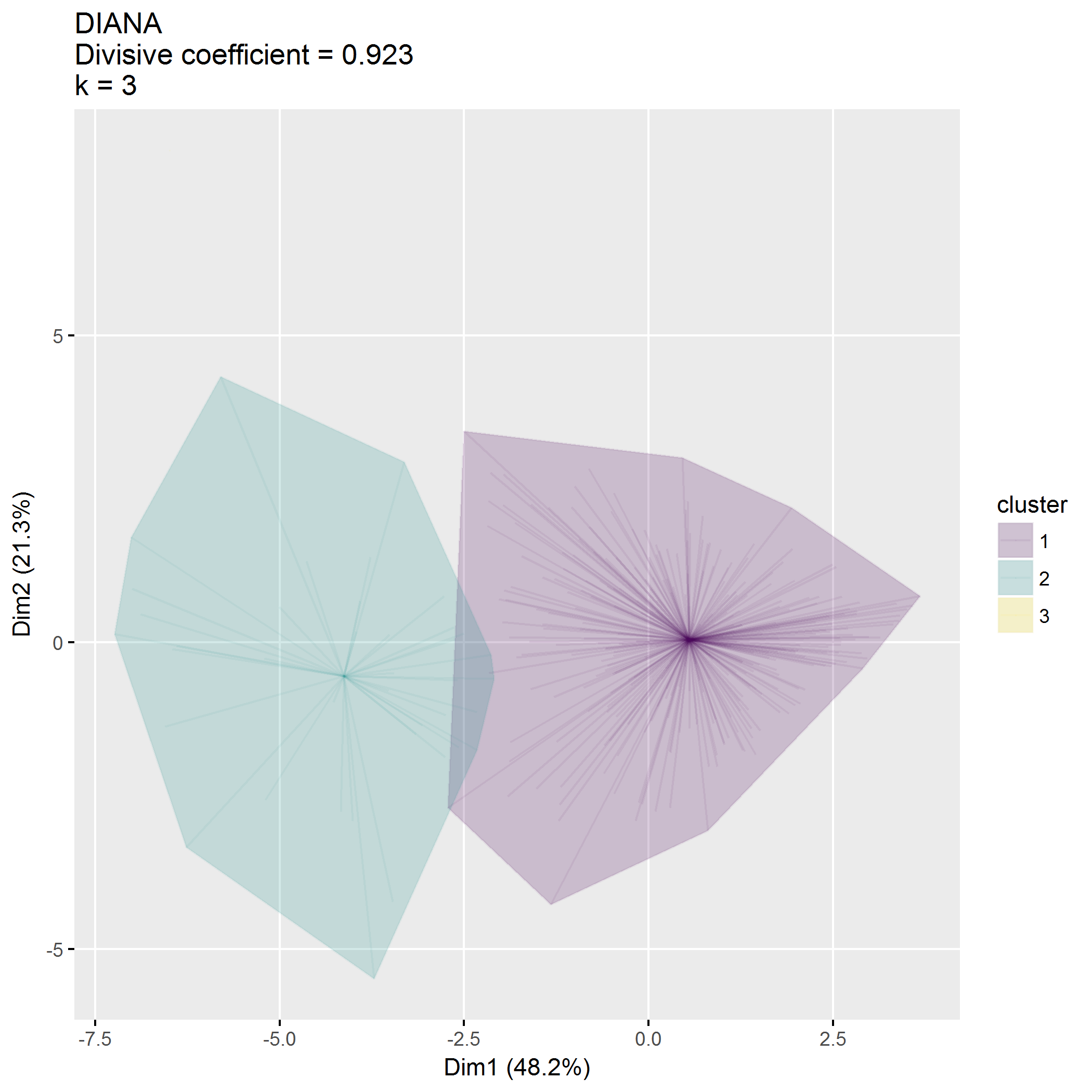
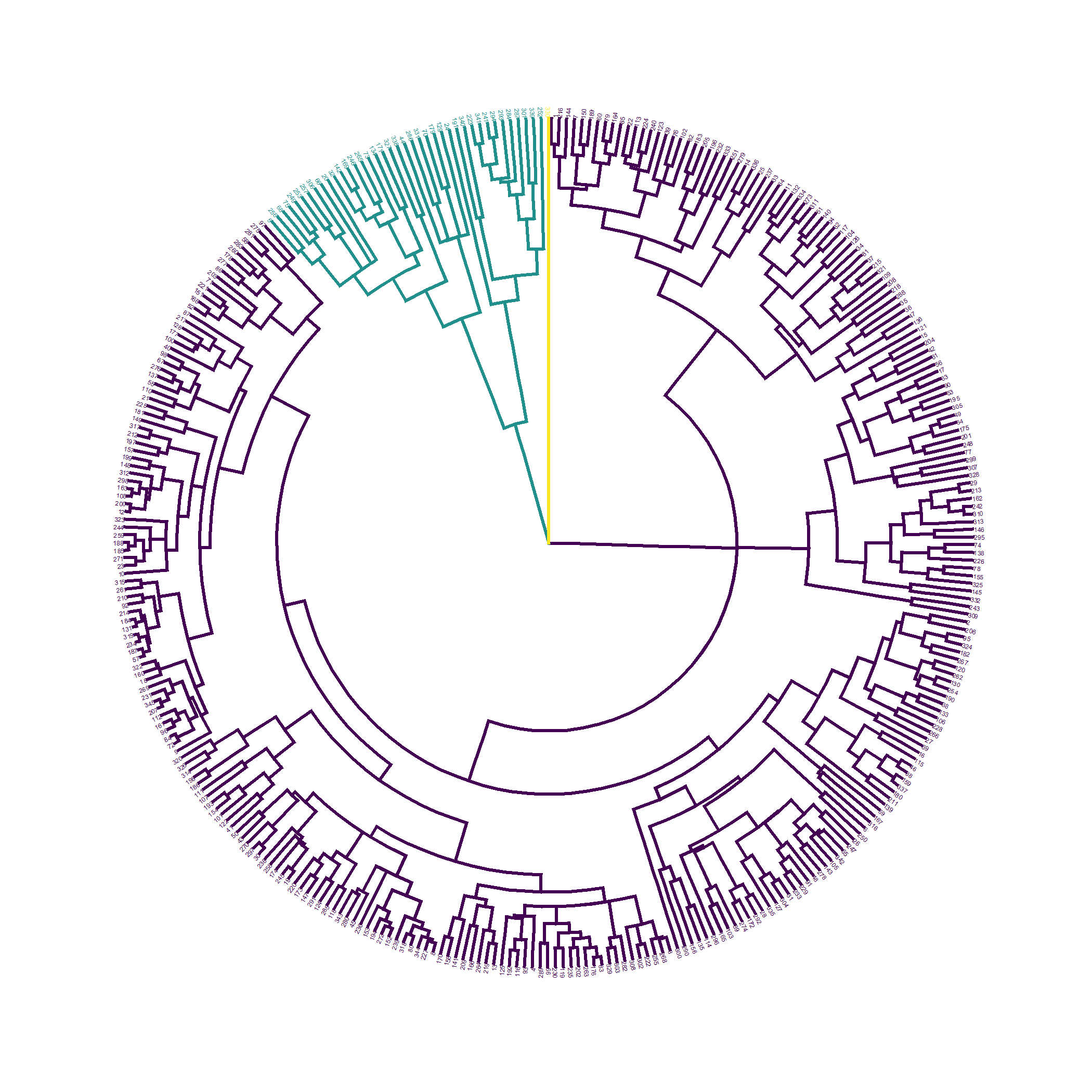
## Cluster on MAPS metrics

## [1] 345 9

## [1] "MAPS\_PP" "MAPS\_PR" "MAPS\_WM" "MAPS\_SP" "MAPS\_HS" "MAPS\_LC"   
## [7] "MAPS\_PC" "MAPS\_POS" "MAPS\_NEG"

## cluster size ave.sil.width  
## 1 1 305 0.45  
## 2 2 39 0.30  
## 3 3 1 0.00

* Hopkins statistic is 0.213
* Analysis identified clusters
* Divisive coefficient is 0.923
* Average silhouette width is 0.436



|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cluster | n | MAPS\_PP\_mean | MAPS\_PR\_mean | MAPS\_WM\_mean | MAPS\_SP\_mean | MAPS\_HS\_mean | MAPS\_LC\_mean | MAPS\_PC\_mean | MAPS\_POS\_mean | MAPS\_NEG\_mean |
| 1 | 305 | 4.1 | 4.6 | 4.7 | 4.5 | 2 | 1.9 | 1.4 | 4.5 | 1.8 |
| 2 | 39 | 3.4 | 3.7 | 4.0 | 3.4 | 3 | 2.6 | 2.6 | 3.7 | 2.7 |
| 3 | 1 | 2.3 | 3.0 | 1.7 | 1.0 | 1 | 1.0 | 1.0 | 2.0 | 1.0 |

**Clustering on MAPS alone is good**

* Cluster 1 () has high *positive* MAPS scores
* Cluster 2 () has high *negative* MAPS scores
* Cluster 3 () is an outlier with low positive and low negative MAPS scores

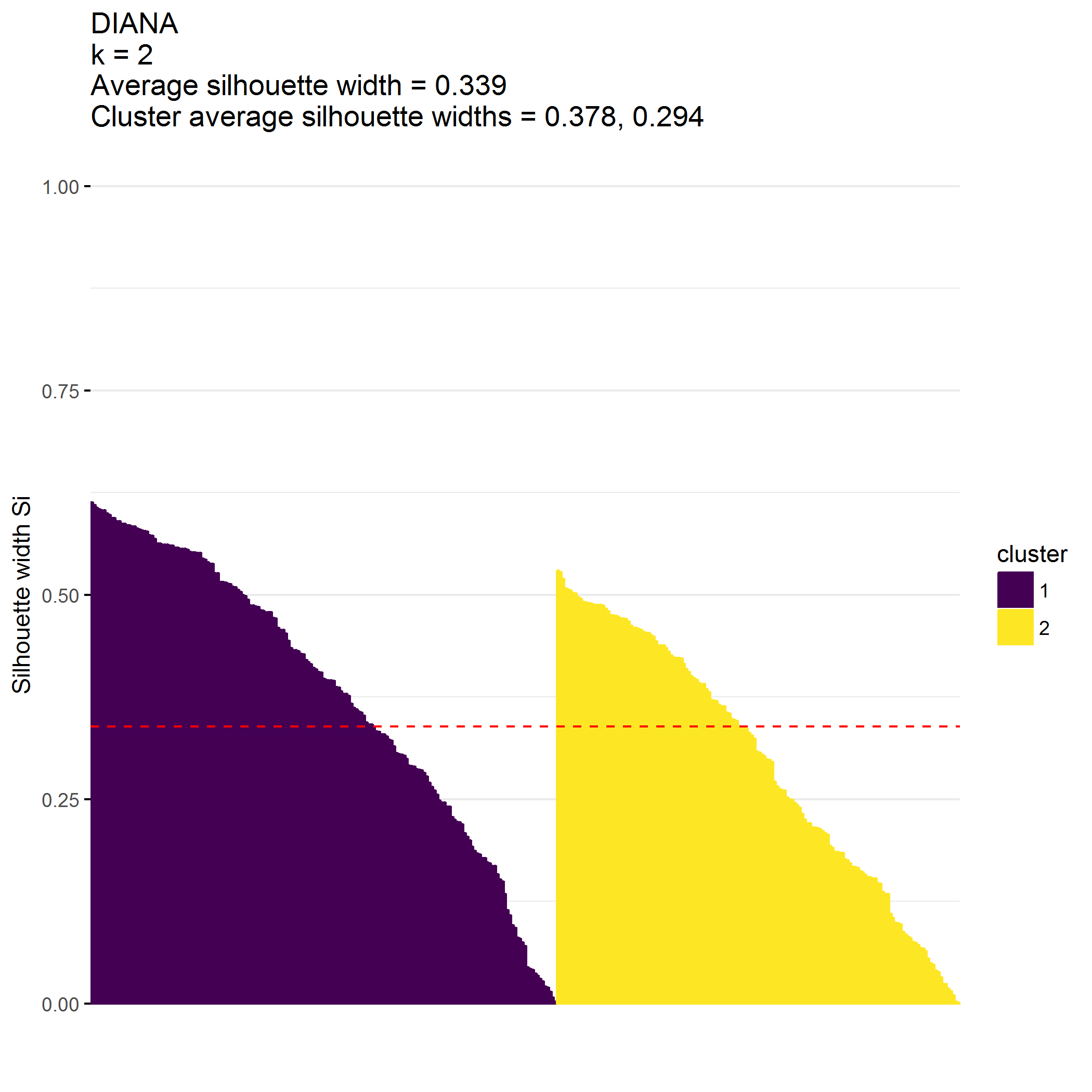
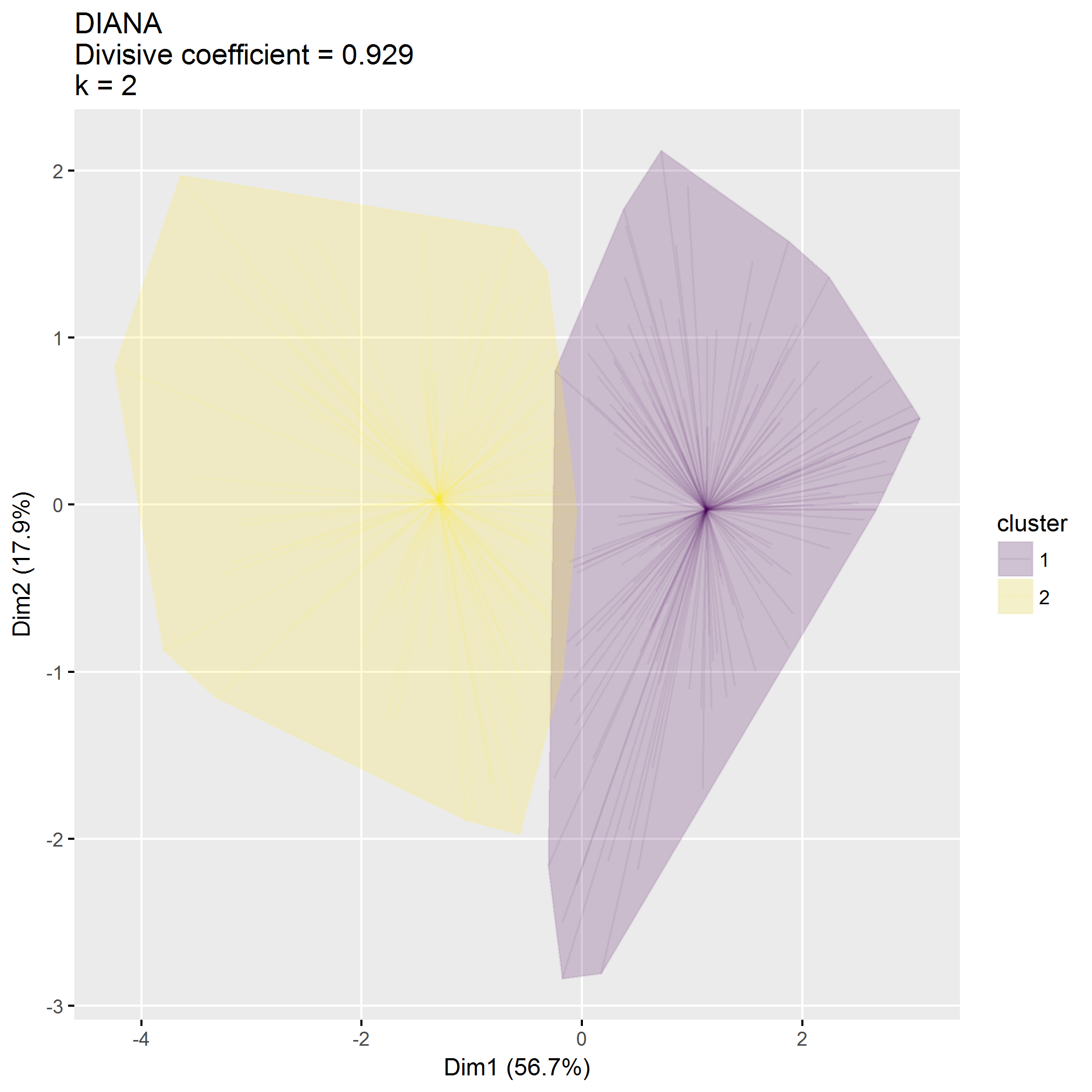
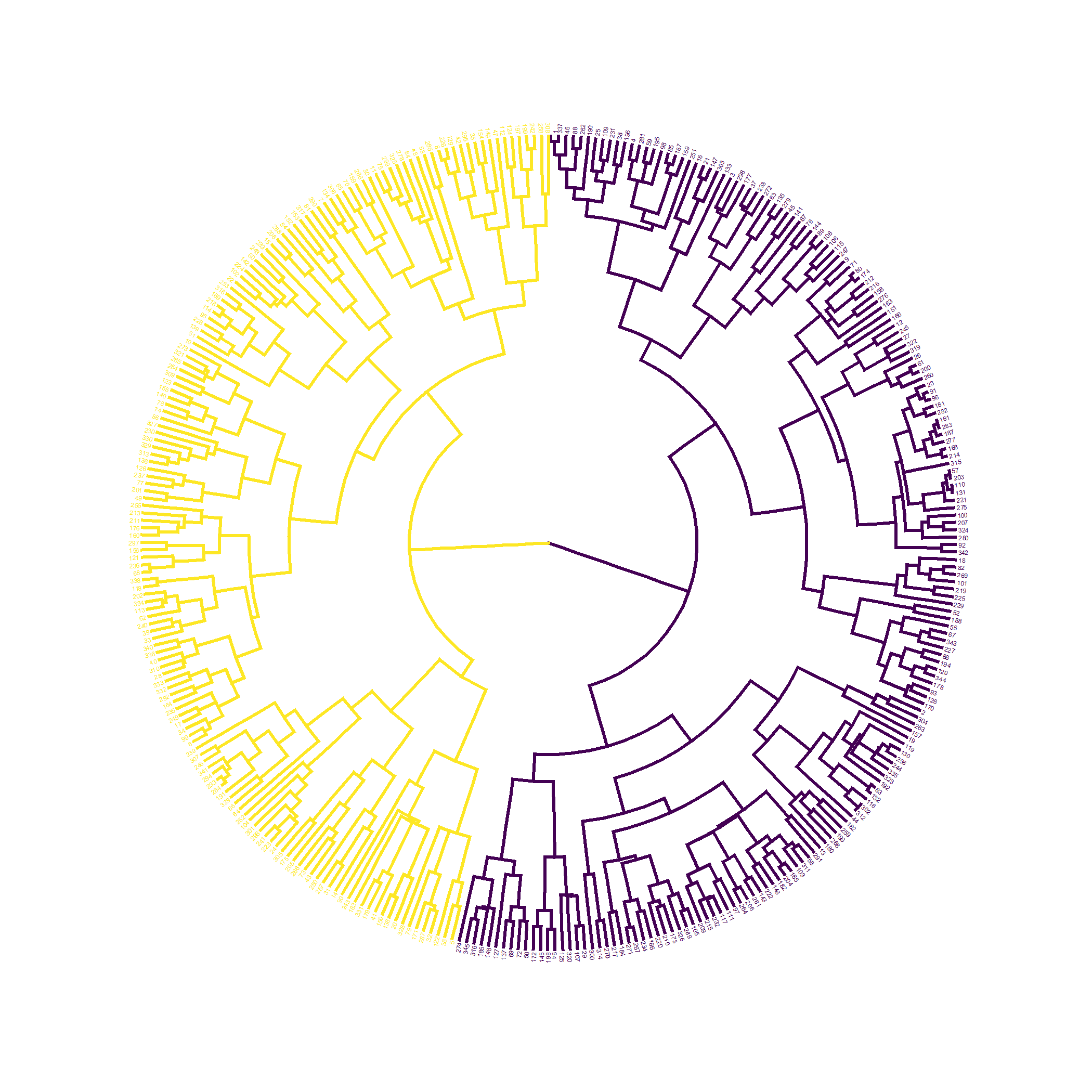
## Cluster on SEPTI metrics

## [1] 345 4

## [1] "SEPTI\_nurturance" "SEPTI\_discipline" "SEPTI\_play"   
## [4] "SEPTI\_routine"

## cluster size ave.sil.width  
## 1 1 185 0.38  
## 2 2 160 0.29

* Hopkins statistic is 0.285
* Analysis identified clusters
* Divisive coefficient is 0.929
* Average silhouette width is 0.339



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| cluster | n | SEPTI\_nurturance\_mean | SEPTI\_discipline\_mean | SEPTI\_play\_mean | SEPTI\_routine\_mean |
| 1 | 185 | 39.6 | 26.8 | 35.8 | 31.5 |
| 2 | 160 | 34.8 | 19.8 | 27.0 | 25.3 |

**Clustering on ECBI, MAPS, and SEPTI metrics isn’t terrible.**

* Cluster 1 () has high SEPTI scores
* Cluster 2 () has low SEPTI scores

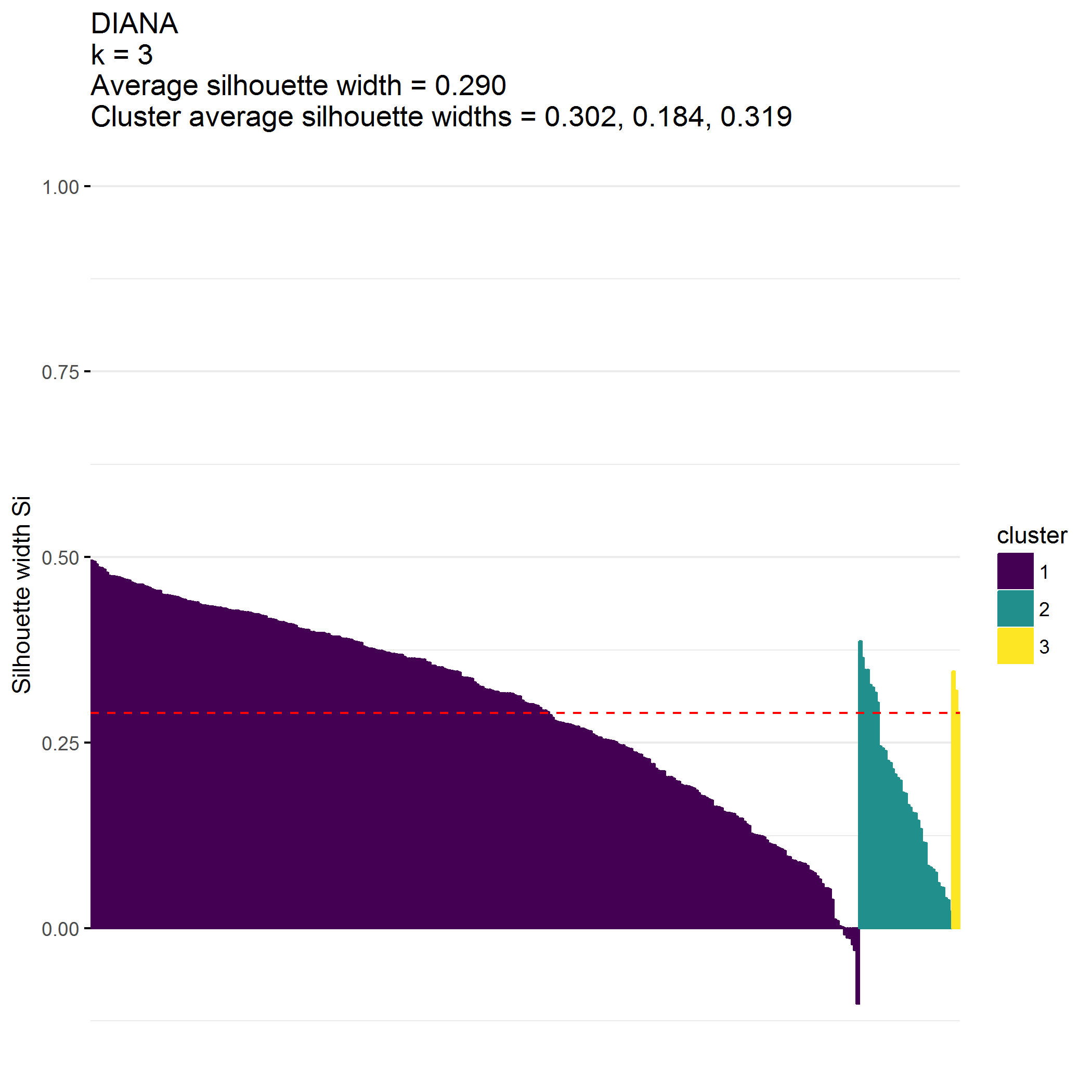
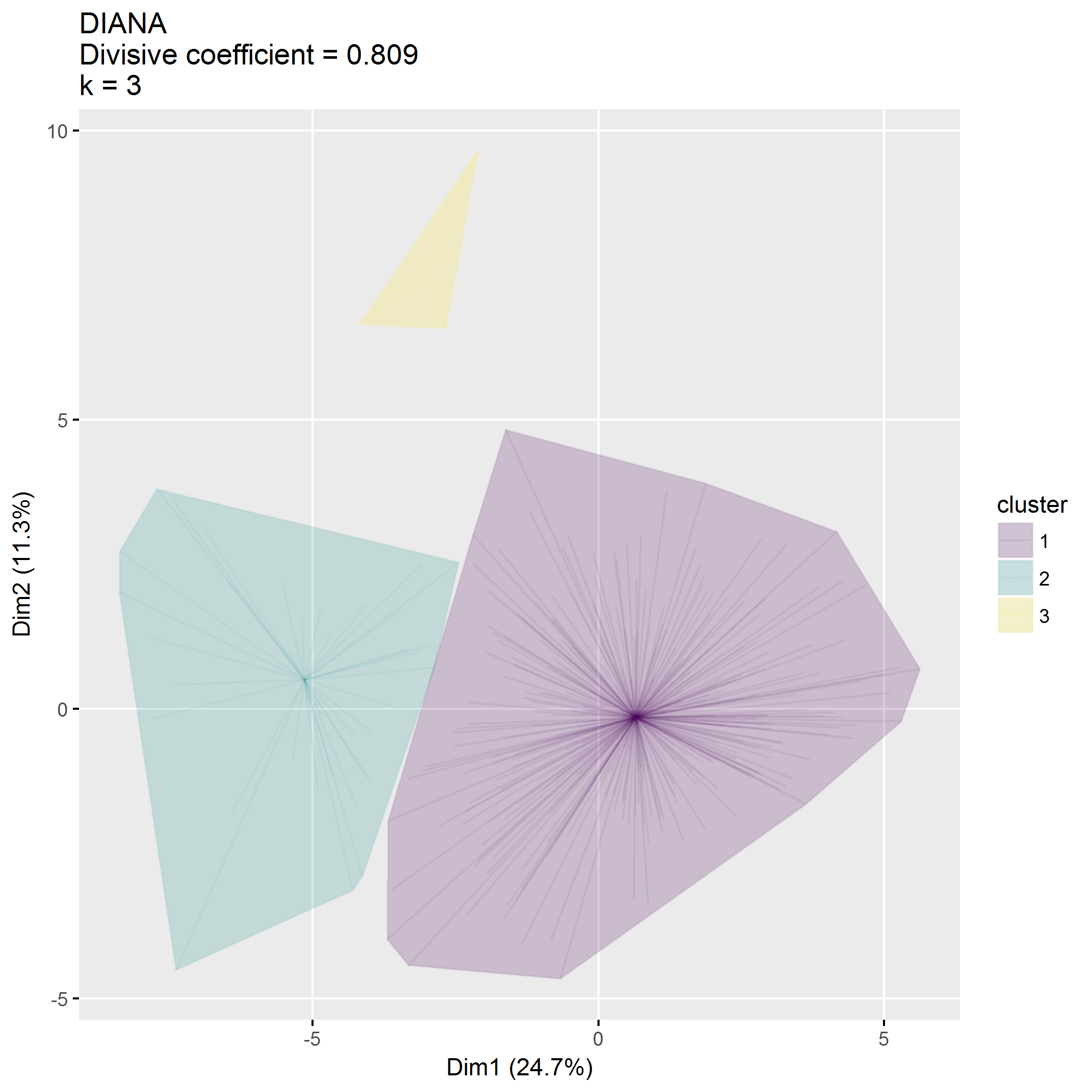
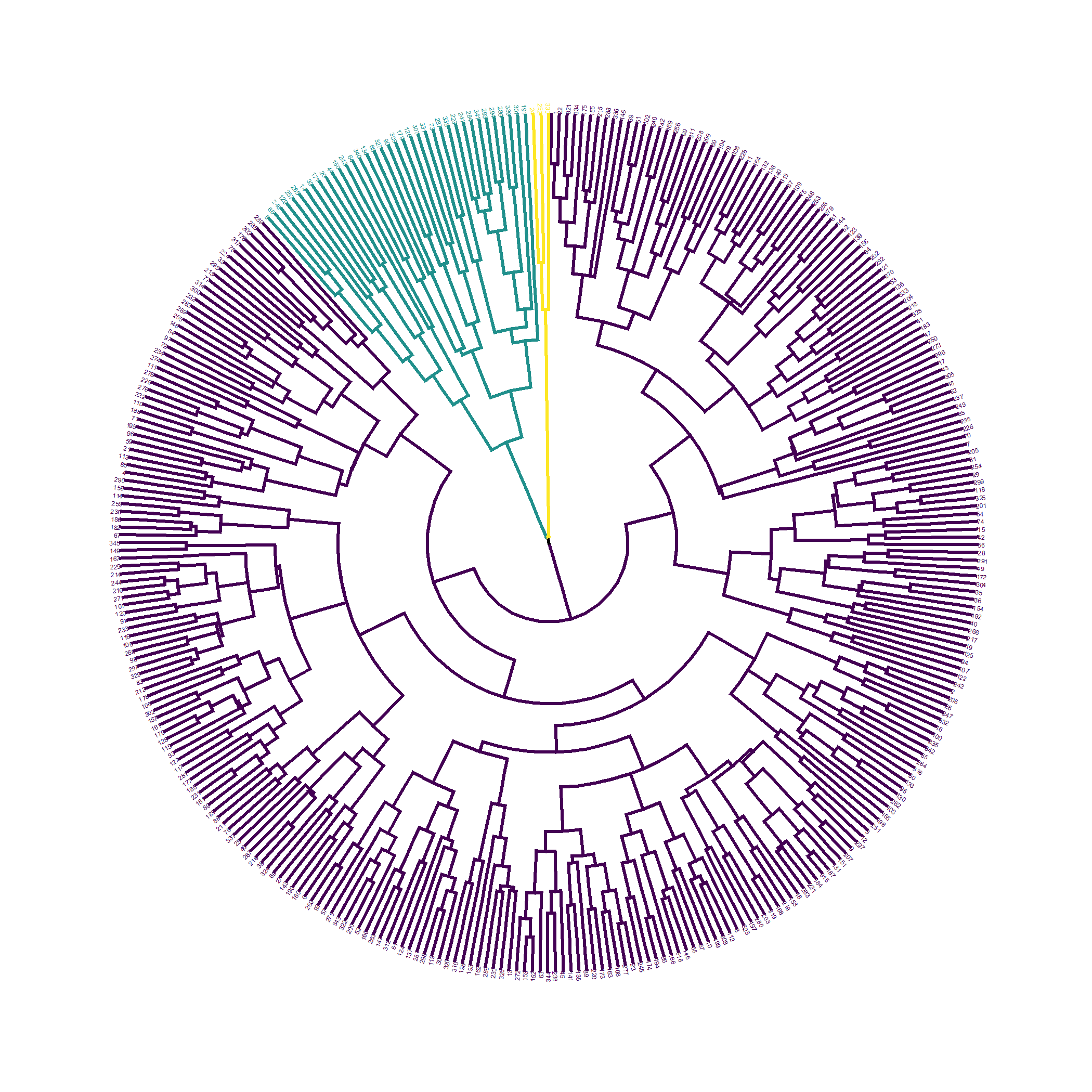
## Cluster on ECBI, MAPS, SEPTI metrics and demographics

## [1] 345 28

## [1] "parentRaceWhite0"   
## [2] "parentRaceWhite1"   
## [3] "parentEducationVocational school/some college"  
## [4] "parentEducationCollege"   
## [5] "parentEducationGraduate/professional school"   
## [6] "income$25,001-$49,999"   
## [7] "income$50,000-$79,999"   
## [8] "income$80,000-$119,999"   
## [9] "income$120,000-$149,999"   
## [10] "income$150,000 or more"   
## [11] "ECBI\_intensity\_T\_score"   
## [12] "ECBI\_problem\_T\_score"   
## [13] "ECBI\_Opp"   
## [14] "ECBI\_Inatt"   
## [15] "ECBI\_Cond"   
## [16] "MAPS\_PP"   
## [17] "MAPS\_PR"   
## [18] "MAPS\_WM"   
## [19] "MAPS\_SP"   
## [20] "MAPS\_HS"   
## [21] "MAPS\_LC"   
## [22] "MAPS\_PC"   
## [23] "MAPS\_POS"   
## [24] "MAPS\_NEG"   
## [25] "SEPTI\_nurturance"   
## [26] "SEPTI\_discipline"   
## [27] "SEPTI\_play"   
## [28] "SEPTI\_routine"

## cluster size ave.sil.width  
## 1 1 305 0.30  
## 2 2 37 0.18  
## 3 3 3 0.32

* Hopkins statistic is 0.311
* Analysis identified clusters
* Divisive coefficient is 0.809
* Average silhouette width is 0.290



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| cluster | n | ECBI\_intensity\_T\_score\_mean | ECBI\_problem\_T\_score\_mean | ECBI\_Opp\_mean | ECBI\_Inatt\_mean | ECBI\_Cond\_mean |
| 1 | 305 | 52.9 | 52.9 | 32.7 | 13.4 | 14.7 |
| 2 | 37 | 61.3 | 62.0 | 42.2 | 15.1 | 23.6 |
| 3 | 3 | 37.7 | 44.3 | 17.7 | 5.3 | 8.0 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cluster | n | MAPS\_PP\_mean | MAPS\_PR\_mean | MAPS\_WM\_mean | MAPS\_SP\_mean | MAPS\_HS\_mean | MAPS\_LC\_mean | MAPS\_PC\_mean | MAPS\_POS\_mean | MAPS\_NEG\_mean |
| 1 | 305 | 4.1 | 4.6 | 4.7 | 4.5 | 2.0 | 1.9 | 1.4 | 4.5 | 1.8 |
| 2 | 37 | 3.5 | 3.8 | 3.9 | 3.6 | 3.0 | 2.8 | 2.4 | 3.7 | 2.7 |
| 3 | 3 | 2.4 | 2.9 | 3.2 | 2.1 | 1.5 | 1.9 | 1.9 | 2.7 | 1.8 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| cluster | n | SEPTI\_nurturance\_mean | SEPTI\_discipline\_mean | SEPTI\_play\_mean | SEPTI\_routine\_mean |
| 1 | 305 | 38.1 | 24.2 | 32.6 | 29.4 |
| 2 | 37 | 32.2 | 17.9 | 25.1 | 22.7 |
| 3 | 3 | 26.3 | 19.3 | 25.7 | 24.0 |

|  |  |  |  |
| --- | --- | --- | --- |
| cluster | income | n | pct |
| 1 | $25,000 or less | 29 | 0.10 |
| 1 | $25,001-$49,999 | 72 | 0.24 |
| 1 | $50,000-$79,999 | 82 | 0.27 |
| 1 | $80,000-$119,999 | 53 | 0.17 |
| 1 | $120,000-$149,999 | 24 | 0.08 |
| 1 | $150,000 or more | 45 | 0.15 |
| 2 | $25,000 or less | 4 | 0.11 |
| 2 | $25,001-$49,999 | 11 | 0.30 |
| 2 | $50,000-$79,999 | 9 | 0.24 |
| 2 | $80,000-$119,999 | 2 | 0.05 |
| 2 | $120,000-$149,999 | 8 | 0.22 |
| 2 | $150,000 or more | 3 | 0.08 |
| 3 | $25,000 or less | 2 | 0.67 |
| 3 | $80,000-$119,999 | 1 | 0.33 |

|  |  |  |  |
| --- | --- | --- | --- |
| cluster | parentEducation | n | pct |
| 1 | High school or less | 37 | 0.12 |
| 1 | Vocational school/some college | 57 | 0.19 |
| 1 | College | 124 | 0.41 |
| 1 | Graduate/professional school | 87 | 0.29 |
| 2 | High school or less | 4 | 0.11 |
| 2 | Vocational school/some college | 11 | 0.30 |
| 2 | College | 12 | 0.32 |
| 2 | Graduate/professional school | 10 | 0.27 |
| 3 | High school or less | 1 | 0.33 |
| 3 | College | 2 | 0.67 |

|  |  |  |  |
| --- | --- | --- | --- |
| cluster | parentRaceWhite | n | pct |
| 1 | 0 | 54 | 0.18 |
| 1 | 1 | 251 | 0.82 |
| 2 | 0 | 24 | 0.65 |
| 2 | 1 | 13 | 0.35 |
| 3 | 1 | 3 | 1.00 |

**Clustering on ECBI, MAPS, and SEPTI metrics and a limited set of demographic factors is good.**

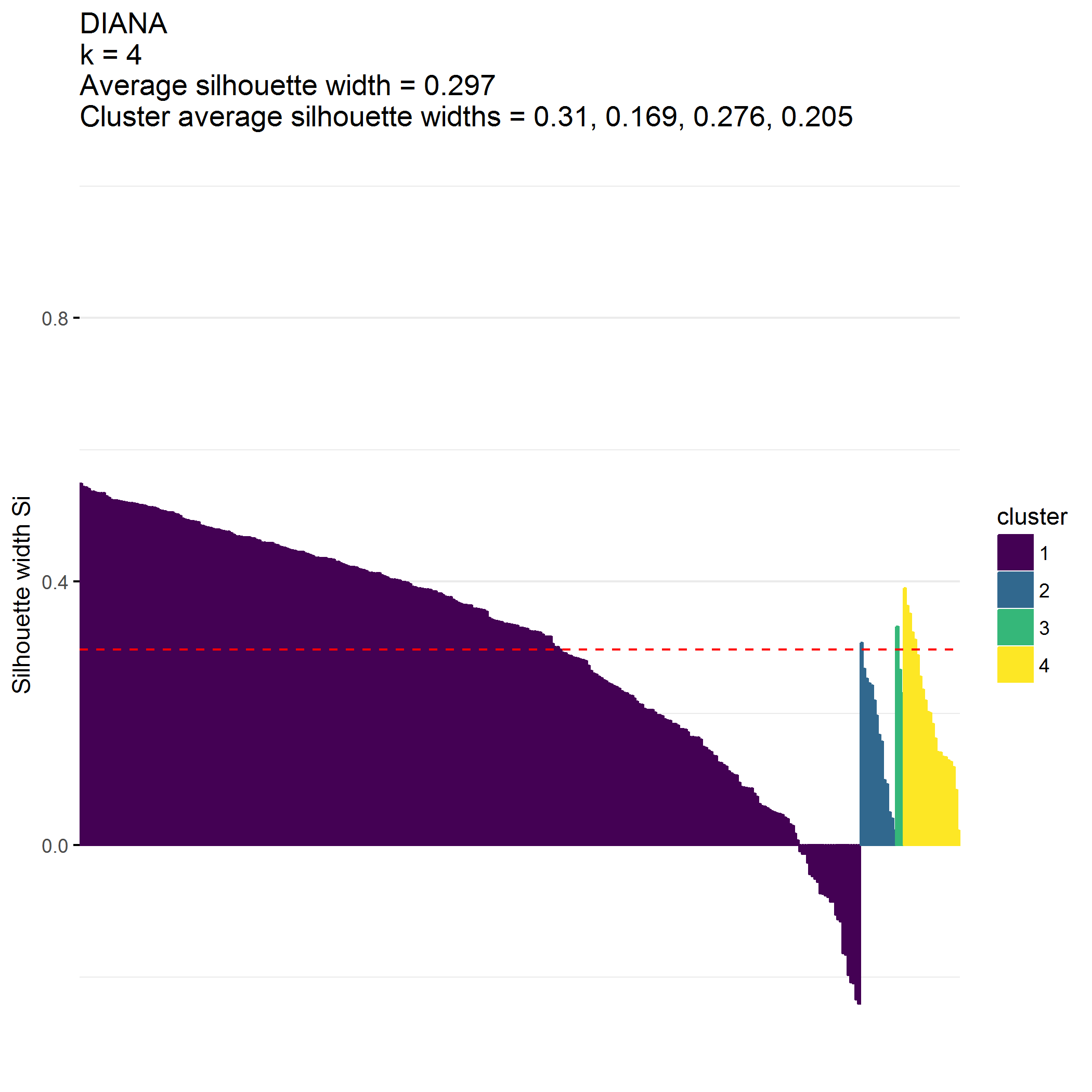
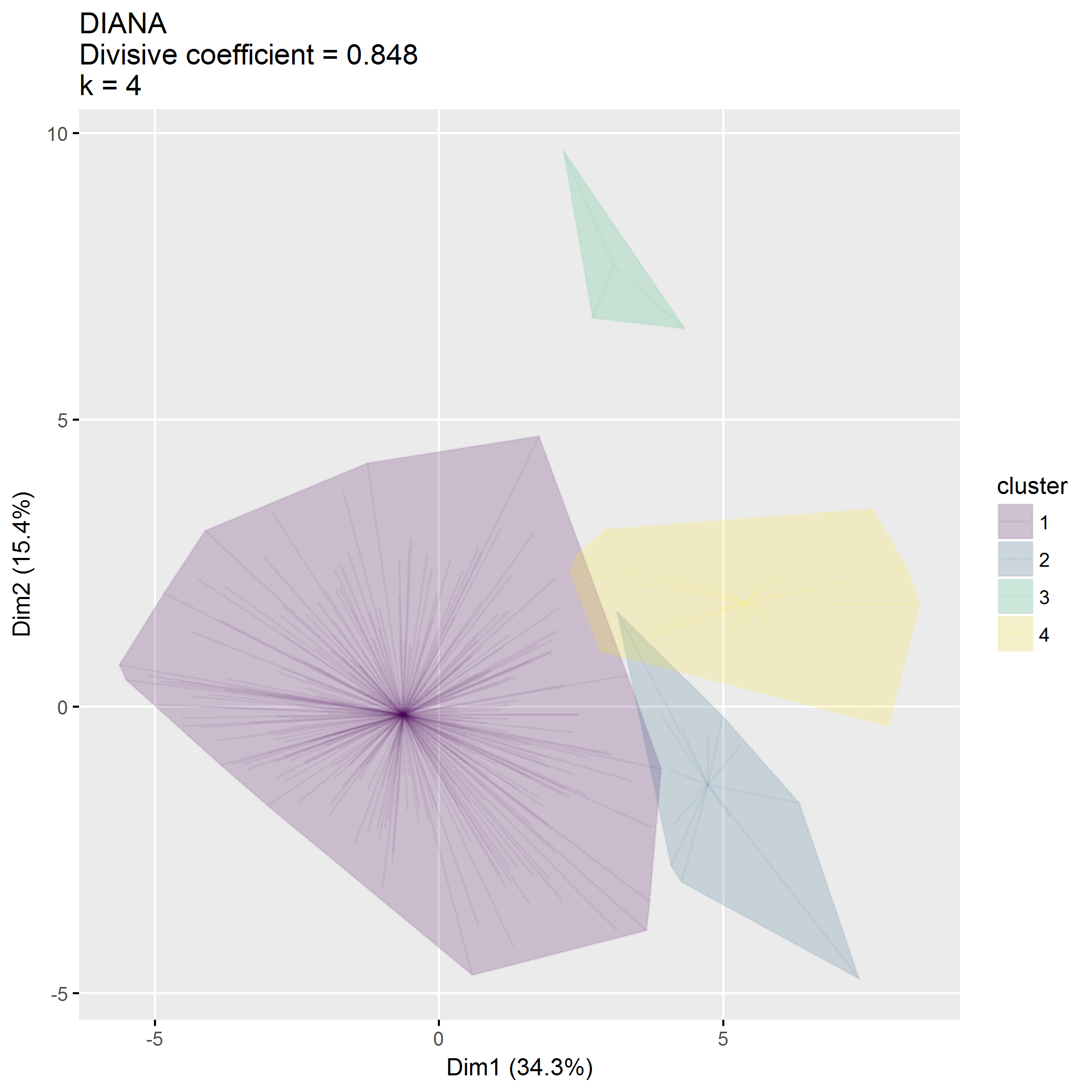
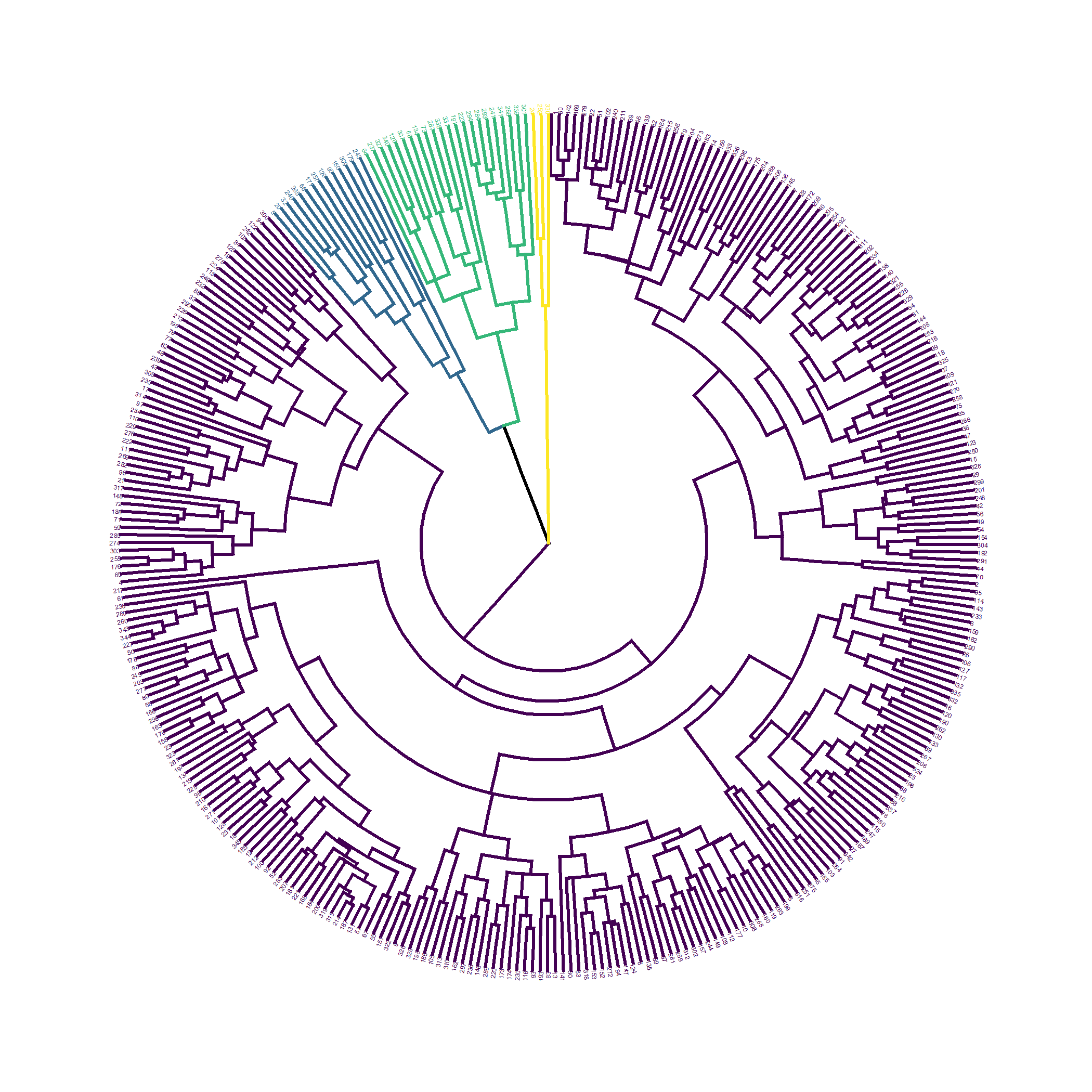
* Demographic factors considered
  + Income
  + Parent education level
  + Parent race, White/Non-White
* Cluster 1 ()
  + Majority White
  + Lower ECBI scores than Cluster 2
  + Higher positive MAPS scores than Cluster 2, lower negative MAPS scores than Cluster 2
  + Higher SEPTI scores than Cluster 2
* Cluster 2 ()
  + Majority non-White
  + High ECBI scores
  + Low positive MAPS scores, high negative MAPS scores
  + Low SEPTI scores
* Cluster 3 ()
  + Is a small outlier cluster
  + Middle income
  + Low ECBI scores
  + Low positive MAPS scores, low negative MAPS scores
* Removing income and parent education level produces better defined clusters
  + **See further analysis below**

## [1] 345 20

## [1] "parentRaceWhite0" "parentRaceWhite1"   
## [3] "ECBI\_intensity\_T\_score" "ECBI\_problem\_T\_score"   
## [5] "ECBI\_Opp" "ECBI\_Inatt"   
## [7] "ECBI\_Cond" "MAPS\_PP"   
## [9] "MAPS\_PR" "MAPS\_WM"   
## [11] "MAPS\_SP" "MAPS\_HS"   
## [13] "MAPS\_LC" "MAPS\_PC"   
## [15] "MAPS\_POS" "MAPS\_NEG"   
## [17] "SEPTI\_nurturance" "SEPTI\_discipline"   
## [19] "SEPTI\_play" "SEPTI\_routine"

## cluster size ave.sil.width  
## 1 1 306 0.31  
## 2 2 14 0.17  
## 3 3 3 0.28  
## 4 4 22 0.21

* Hopkins statistic is 0.289
* Analysis identified clusters
* Divisive coefficient is 0.848
* Average silhouette width is 0.297



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| cluster | n | ECBI\_intensity\_T\_score\_mean | ECBI\_problem\_T\_score\_mean | ECBI\_Opp\_mean | ECBI\_Inatt\_mean | ECBI\_Cond\_mean |
| 1 | 306 | 52.9 | 53.0 | 32.7 | 13.4 | 14.7 |
| 2 | 14 | 63.9 | 68.8 | 46.3 | 16.6 | 22.4 |
| 3 | 3 | 37.7 | 44.3 | 17.7 | 5.3 | 8.0 |
| 4 | 22 | 59.5 | 56.0 | 39.0 | 14.4 | 24.4 |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cluster | n | MAPS\_PP\_mean | MAPS\_PR\_mean | MAPS\_WM\_mean | MAPS\_SP\_mean | MAPS\_HS\_mean | MAPS\_LC\_mean | MAPS\_PC\_mean | MAPS\_POS\_mean | MAPS\_NEG\_mean |
| 1 | 306 | 4.1 | 4.6 | 4.7 | 4.5 | 2.0 | 1.9 | 1.4 | 4.5 | 1.8 |
| 2 | 14 | 4.0 | 4.1 | 4.1 | 3.9 | 3.2 | 2.6 | 2.4 | 4.0 | 2.7 |
| 3 | 3 | 2.4 | 2.9 | 3.2 | 2.1 | 1.5 | 1.9 | 1.9 | 2.7 | 1.8 |
| 4 | 22 | 3.2 | 3.6 | 3.8 | 3.4 | 2.8 | 2.9 | 2.3 | 3.5 | 2.7 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| cluster | n | SEPTI\_nurturance\_mean | SEPTI\_discipline\_mean | SEPTI\_play\_mean | SEPTI\_routine\_mean |
| 1 | 306 | 38.1 | 24.2 | 32.6 | 29.3 |
| 2 | 14 | 34.1 | 15.9 | 21.8 | 21.6 |
| 3 | 3 | 26.3 | 19.3 | 25.7 | 24.0 |
| 4 | 22 | 30.8 | 18.9 | 27.0 | 23.5 |

|  |  |  |  |
| --- | --- | --- | --- |
| cluster | parentRaceWhite | n | pct |
| 1 | 0 | 53 | 0.17 |
| 1 | 1 | 253 | 0.83 |
| 2 | 0 | 4 | 0.29 |
| 2 | 1 | 10 | 0.71 |
| 3 | 1 | 3 | 1.00 |
| 4 | 0 | 21 | 0.95 |
| 4 | 1 | 1 | 0.05 |

* Demographic factors considered
  + Parent race, White/Non-White
* Cluster 1 ()
  + Majority White
  + Lower ECBI scores than Clusters 2/4
  + Higher positive MAPS scores than Clusters 2/4, lower negative MAPS scores than Clusters 2/4
  + Higher SEPTI scores than Clusters 2/4
* Cluster 2 ()
  + Majority White, *more similar to Cluster 1*
  + High ECBI scores, *more similar to Cluster 4*
  + Low positive MAPS scores, high negative MAPS scores, *more similar to Cluster 4*
  + Low SEPTI scores, *more similar to Cluster 4*
* Cluster 3 ()
  + Is a small outlier cluster
  + Middle income
  + Low ECBI scores
  + Low positive MAPS scores, low negative MAPS scores
* Cluster 4 ()
  + Majority non-White
  + High ECBI scores
  + Low positive MAPS scores, high negative MAPS scores
  + Low SEPTI scores

## Save objects

## size isdir mode  
## data/processed/clusterAnalysis.RData 5216263 FALSE 666  
## mtime  
## data/processed/clusterAnalysis.RData 2018-07-12 10:20:42  
## ctime  
## data/processed/clusterAnalysis.RData 2018-07-06 12:14:44  
## atime exe  
## data/processed/clusterAnalysis.RData 2018-07-06 12:14:44 no