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

Benjamin Chan (chanb@ohsu.edu)

2018-07-17

# 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 |

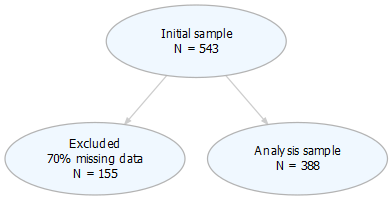
## Warning: package 'bindrcpp' was built under R version 3.4.4

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.

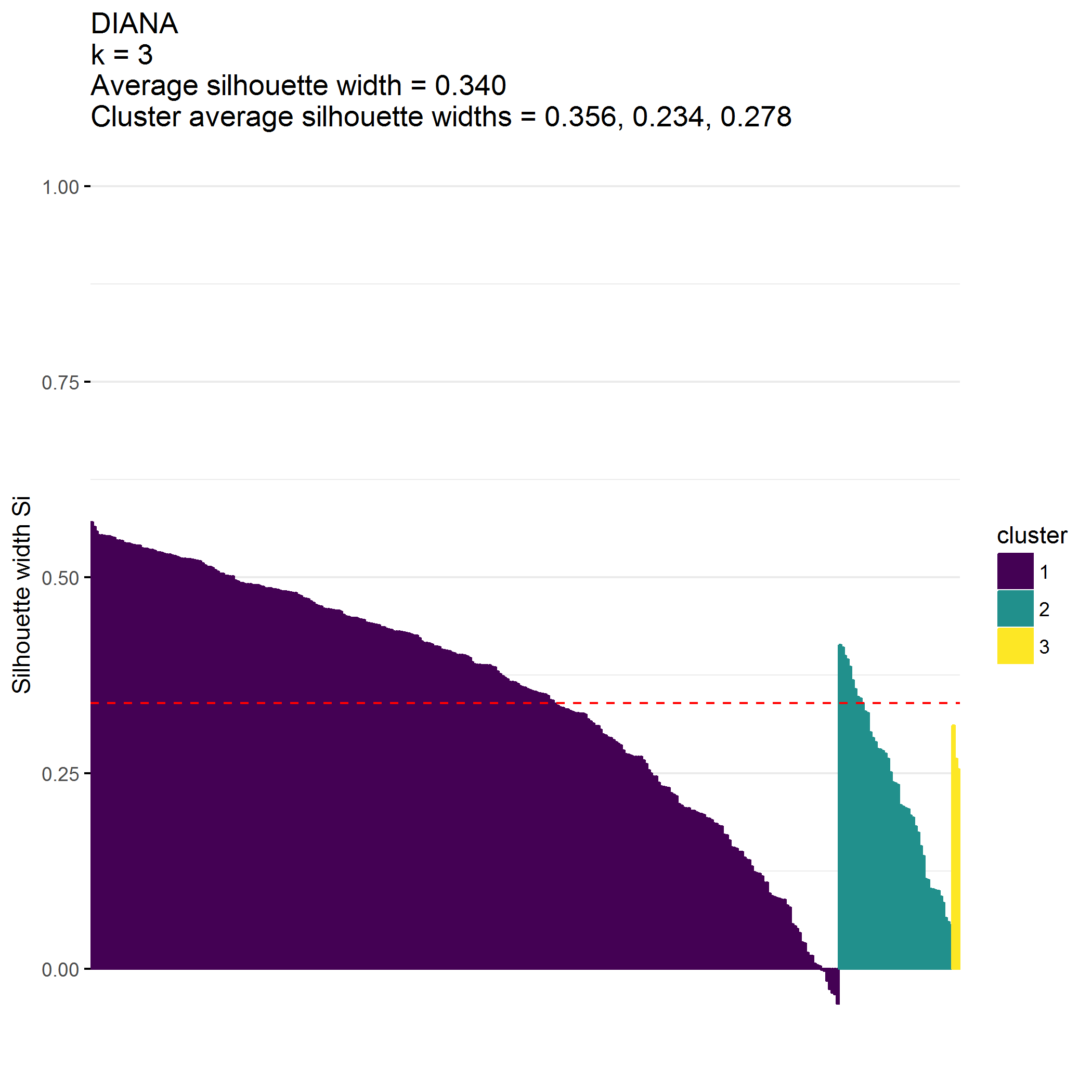
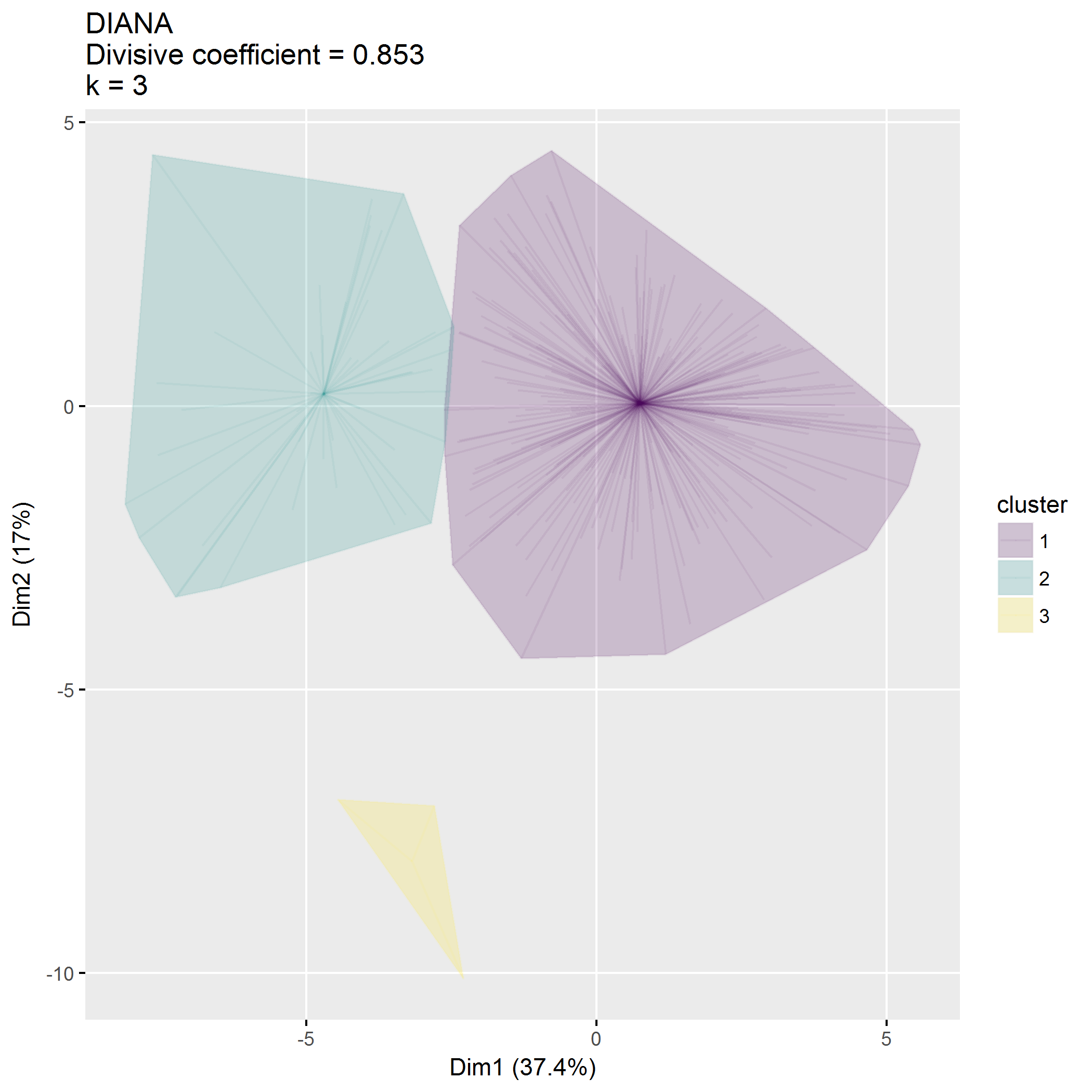
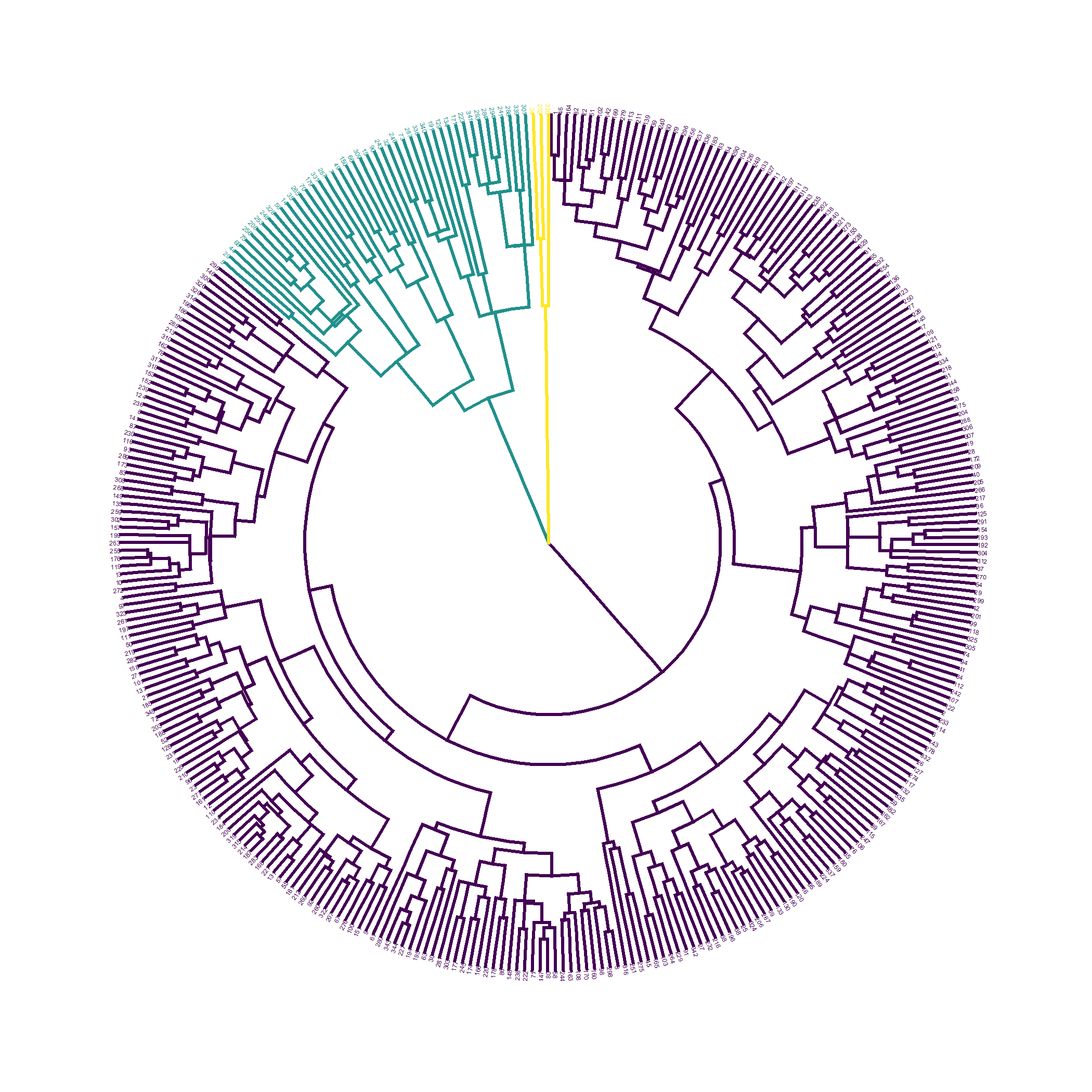
## Cluster on ECBI, MAPS, SEPTI metrics (no demographics)

## [1] 345 18

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

## cluster size ave.sil.width  
## 1 1 297 0.36  
## 2 2 45 0.23  
## 3 3 3 0.28

* Hopkins statistic is 0.288
* Analysis identified clusters
* Divisive coefficient is 0.853
* Average silhouette width is 0.340



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| cluster | n | ECBI\_intensity\_T\_score\_mean | ECBI\_problem\_T\_score\_mean | ECBI\_Opp\_mean | ECBI\_Inatt\_mean | ECBI\_Cond\_mean |
| 1 | 297 | 52.5 | 52.2 | 32.3 | 13.2 | 14.3 |
| 2 | 45 | 62.3 | 64.6 | 43.2 | 15.8 | 24.4 |
| 3 | 3 | 37.7 | 44.3 | 17.7 | 5.3 | 8.0 |

## Error in dots\_values(...): object 'dfXMetricsg' not found

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| cluster | n | SEPTI\_nurturance\_mean | SEPTI\_discipline\_mean | SEPTI\_play\_mean | SEPTI\_routine\_mean |
| 1 | 297 | 38.1 | 24.3 | 32.8 | 29.4 |
| 2 | 45 | 33.4 | 18.8 | 25.0 | 23.5 |
| 3 | 3 | 26.3 | 19.3 | 25.7 | 24.0 |

|  |  |  |  |
| --- | --- | --- | --- |
| cluster | parentRaceWhite | n | pct |
| 1 | 0 | 58 | 0.20 |
| 1 | 1 | 239 | 0.80 |
| 2 | 0 | 20 | 0.44 |
| 2 | 1 | 25 | 0.56 |
| 3 | 1 | 3 | 1.00 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| cluster | n | PCB1\_Total\_mean | PCB1\_CondEmot\_mean | PCB1\_DevHab\_mean |
| 1 | 297 | 65.8 | 47.5 | 18.3 |
| 2 | 45 | 71.9 | 52.6 | 19.3 |
| 3 | 3 | 85.3 | 62.3 | 23.0 |

|  |  |  |
| --- | --- | --- |
| cluster | n | PCB2\_Tot\_mean |
| 1 | 297 | 24.4 |
| 2 | 45 | 24.2 |
| 3 | 3 | 27.3 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| cluster | n | PCB3\_Total\_mean | PCB3\_PCPonly\_mean | PCB3\_Person\_mean | PCB3\_Resource\_mean |
| 1 | 297 | 46.4 | 4.1 | 15.7 | 26.7 |
| 2 | 45 | 50.1 | 4.2 | 17.2 | 28.7 |
| 3 | 3 | 65.7 | 4.7 | 23.3 | 37.7 |

* Cluster 1 () has high positive MAPS scores and high SEPTI scores
* Cluster 2 () has high negative MAPS scores and high ECBI scores
* Cluster 3 () has low ECBI scores

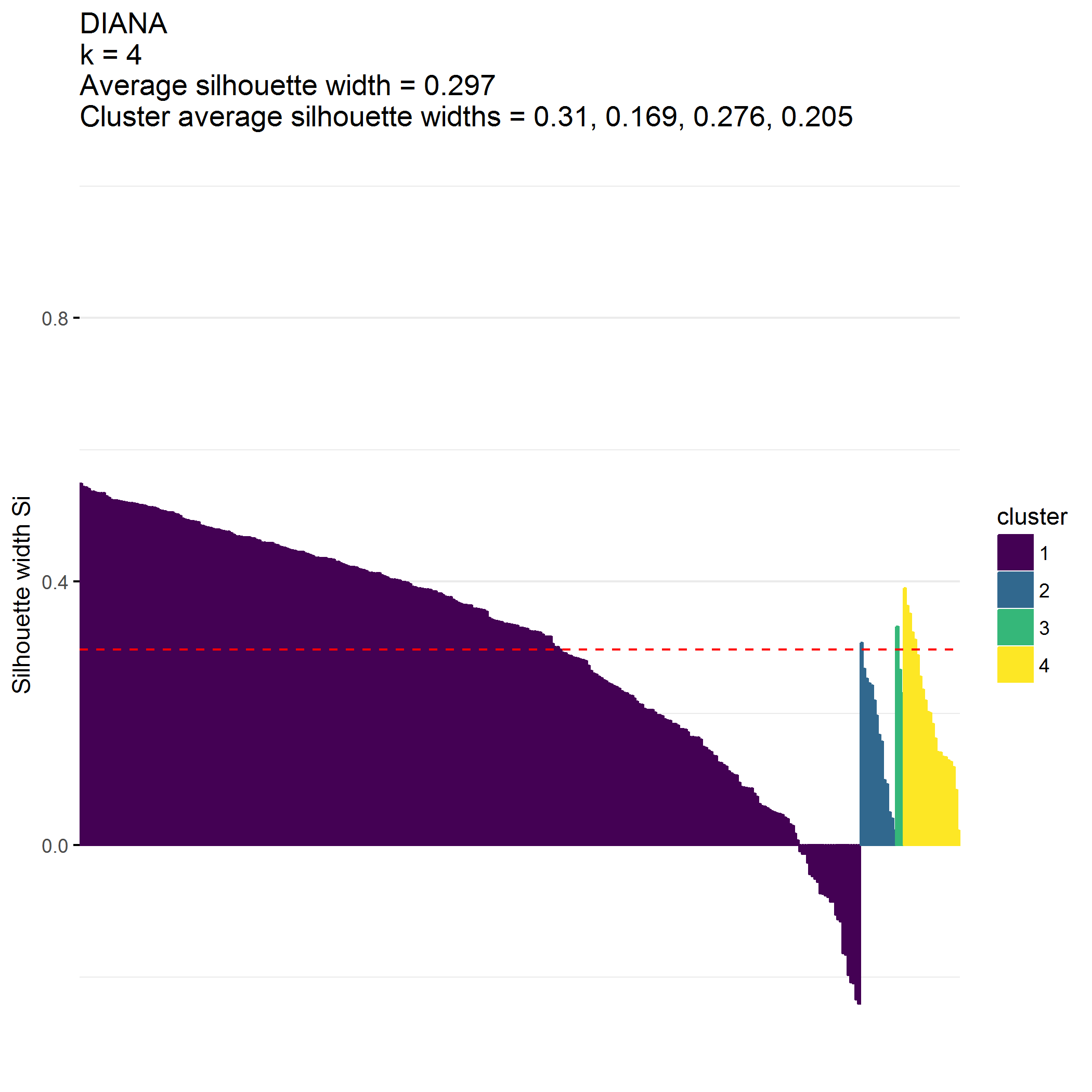
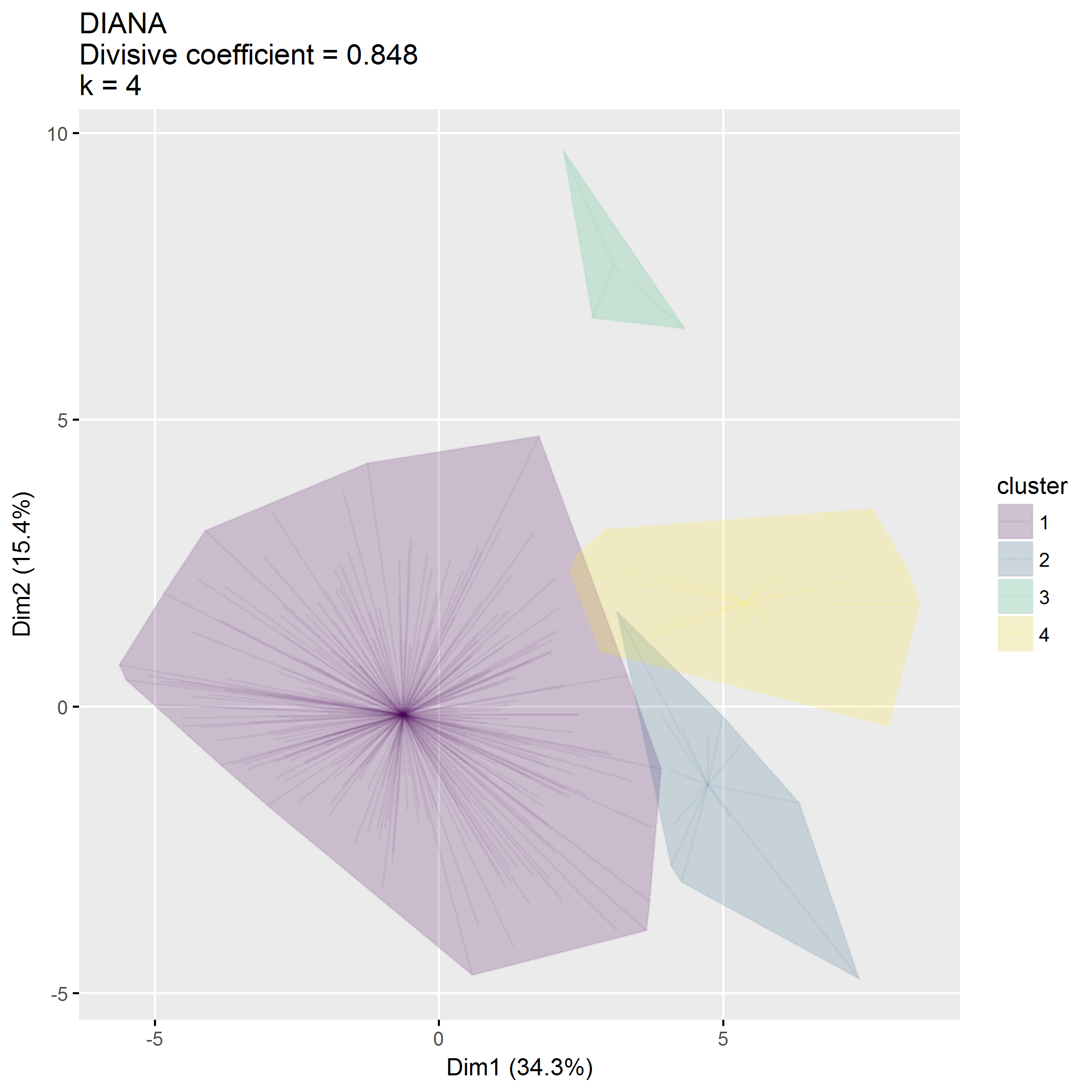
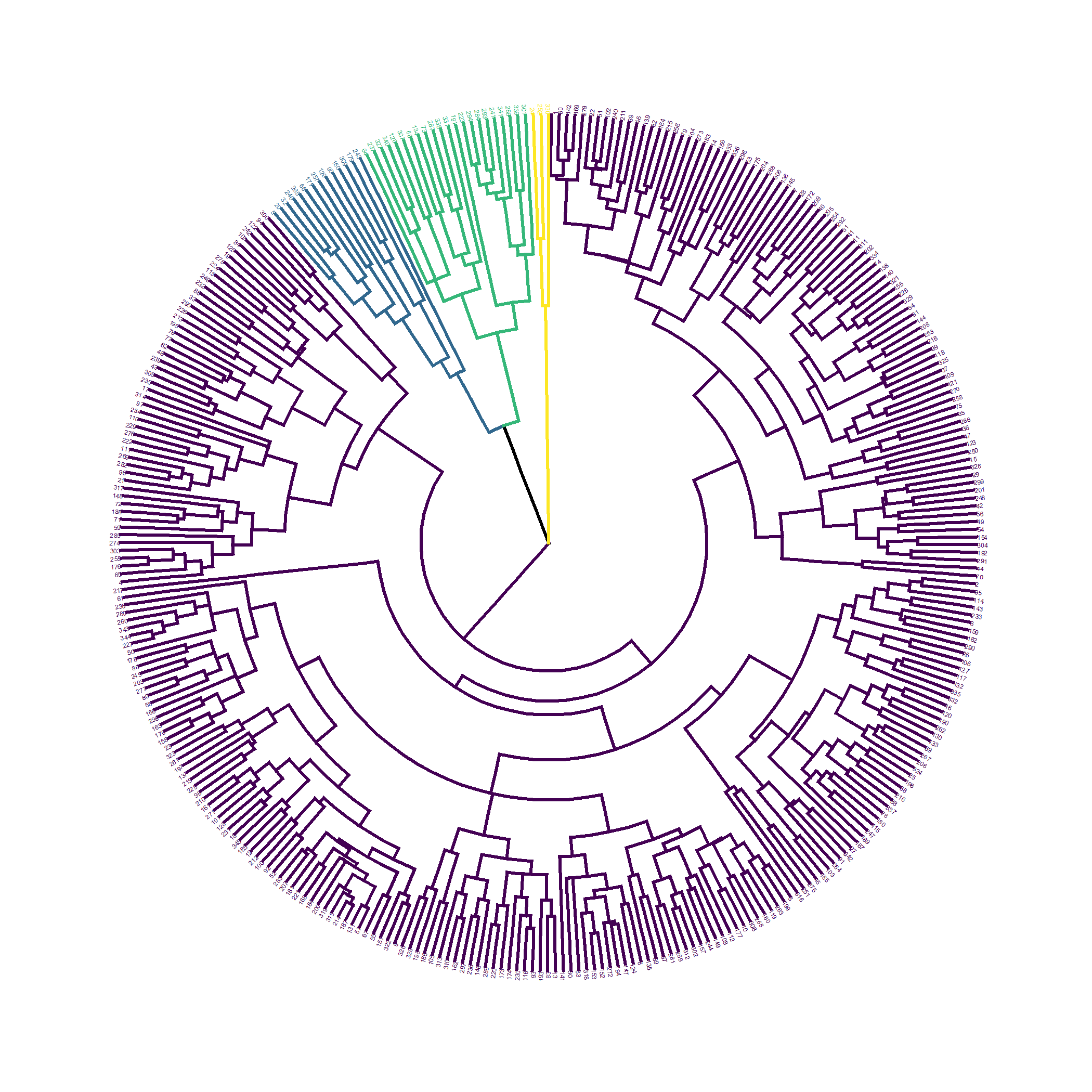
## Cluster on ECBI, MAPS, SEPTI metrics and parent’s race

## [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 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| cluster | n | PCB1\_Total\_mean | PCB1\_CondEmot\_mean | PCB1\_DevHab\_mean |
| 1 | 306 | 66.0 | 47.7 | 18.2 |
| 2 | 14 | 73.1 | 53.3 | 19.8 |
| 3 | 3 | 85.3 | 62.3 | 23.0 |
| 4 | 22 | 71.8 | 51.8 | 20.0 |

|  |  |  |
| --- | --- | --- |
| cluster | n | PCB2\_Tot\_mean |
| 1 | 306 | 24.4 |
| 2 | 14 | 24.4 |
| 3 | 3 | 27.3 |
| 4 | 22 | 23.9 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| cluster | n | PCB3\_Total\_mean | PCB3\_PCPonly\_mean | PCB3\_Person\_mean | PCB3\_Resource\_mean |
| 1 | 306 | 46.3 | 4.1 | 15.7 | 26.6 |
| 2 | 14 | 50.9 | 4.1 | 17.1 | 29.6 |
| 3 | 3 | 65.7 | 4.7 | 23.3 | 37.7 |
| 4 | 22 | 52.5 | 4.1 | 18.1 | 30.2 |

* 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

## Compare clusterings

|  |  |  |
| --- | --- | --- |
| clusterMetrics | clusterMetricsDemog | n |
| 1 | 1 | 292 |
| 1 | 4 | 5 |
| 2 | 1 | 14 |
| 2 | 2 | 14 |
| 2 | 4 | 17 |
| 3 | 3 | 3 |

## Save objects

Bind study ID, id, to cluster ID, cluster.

## mtime size  
## data/processed/clusterAnalysis.RData 2018-07-17 08:29:54 5239484

## mtime size  
## data/processed/clusterCrosswalk.csv 2018-07-17 08:29:54 3091

## mtime size  
## data/processed/clusterCrosswalk.sav 2018-07-17 08:29:54 8768

Test SPSS data file.

all.equal(dfCrosswalk, read\_sav(f))

## Warning: Column `id` has different attributes on LHS and RHS of join

## Warning: Column `clusterMetrics` has different attributes on LHS and RHS of  
## join

## Warning: Column `clusterMetricsDemog` has different attributes on LHS and  
## RHS of join

## [1] TRUE

str(dfCrosswalk)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 345 obs. of 3 variables:  
## $ id : num 1 3 4 5 6 7 8 9 10 11 ...  
## $ clusterMetrics : num 1 1 1 1 2 1 2 1 1 1 ...  
## $ clusterMetricsDemog: num 1 1 1 1 2 1 1 1 1 1 ...

str(read\_sav(f))

## Classes 'tbl\_df', 'tbl' and 'data.frame': 345 obs. of 3 variables:  
## $ id : atomic 1 3 4 5 6 7 8 9 10 11 ...  
## ..- attr(\*, "format.spss")= chr "F8.2"  
## $ clusterMetrics : atomic 1 1 1 1 2 1 2 1 1 1 ...  
## ..- attr(\*, "format.spss")= chr "F8.2"  
## $ clusterMetricsDemog: atomic 1 1 1 1 2 1 1 1 1 1 ...  
## ..- attr(\*, "format.spss")= chr "F8.2"