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

Benjamin Chan (chanb@ohsu.edu)

2018-07-06

# 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

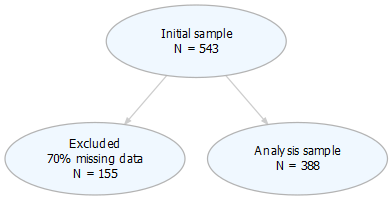
Remove certain predictor variables:

* 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.

## PCB1\_Total PCB2\_Tot PCB3\_Total   
## Min. :18.00 Min. : 6.00 Min. :15.00   
## 1st Qu.:58.00 1st Qu.:22.00 1st Qu.:39.00   
## Median :71.00 Median :25.00 Median :48.00   
## Mean :67.85 Mean :24.53 Mean :47.58   
## 3rd Qu.:81.00 3rd Qu.:28.00 3rd Qu.:57.00   
## Max. :90.00 Max. :30.00 Max. :75.00



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 'ggdendro' 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

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

##   
## ---------------------  
## Welcome to dendextend version 1.8.0  
## Type citation('dendextend') for how to cite the package.  
##   
## Type browseVignettes(package = 'dendextend') for the package vignette.  
## The github page is: https://github.com/talgalili/dendextend/  
##   
## Suggestions and bug-reports can be submitted at: https://github.com/talgalili/dendextend/issues  
## Or contact: <tal.galili@gmail.com>  
##   
## To suppress this message use: suppressPackageStartupMessages(library(dendextend))  
## ---------------------

##   
## Attaching package: 'dendextend'

## The following object is masked from 'package:ggdendro':  
##   
## theme\_dendro

## The following object is masked from 'package:stats':  
##   
## cutree

##   
## 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

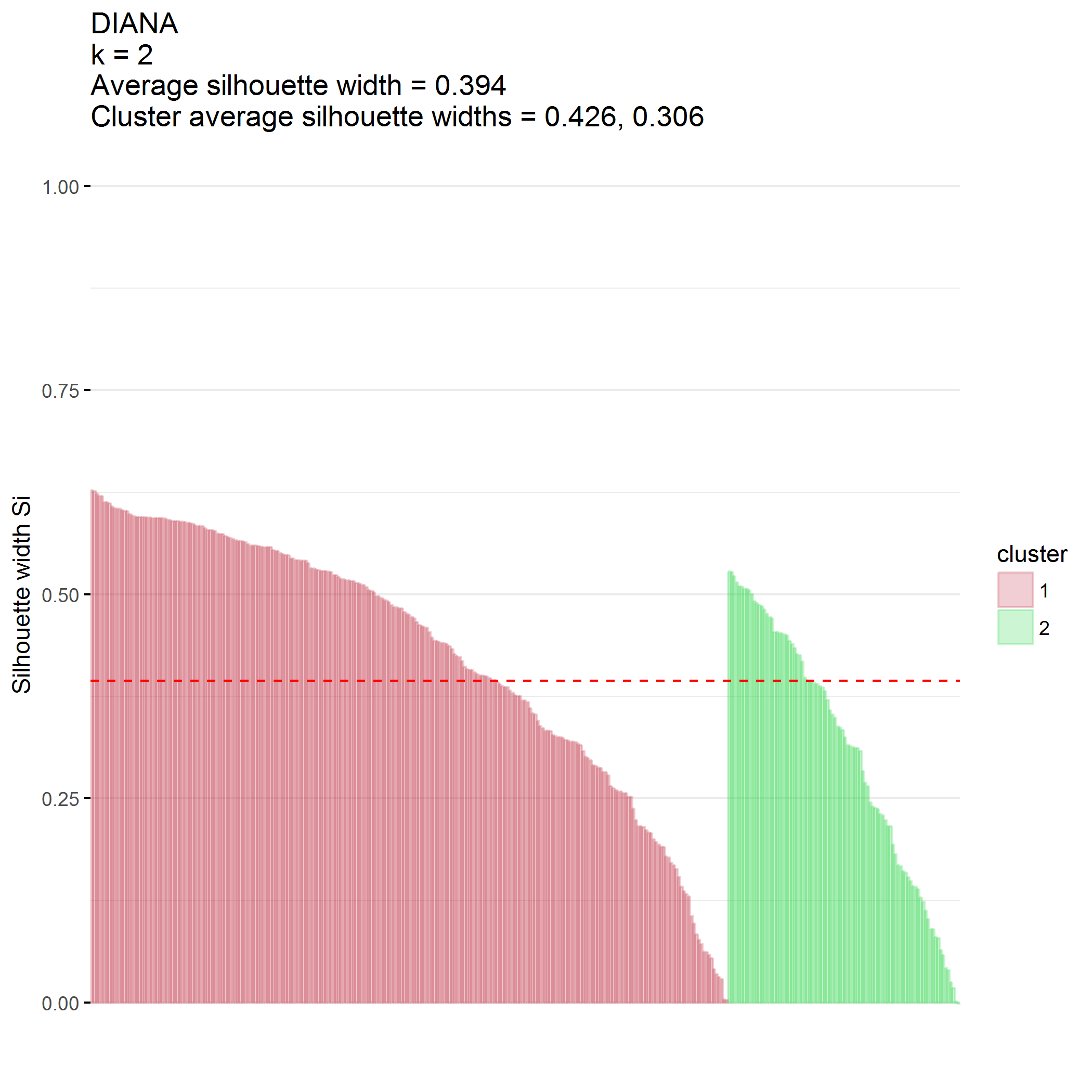
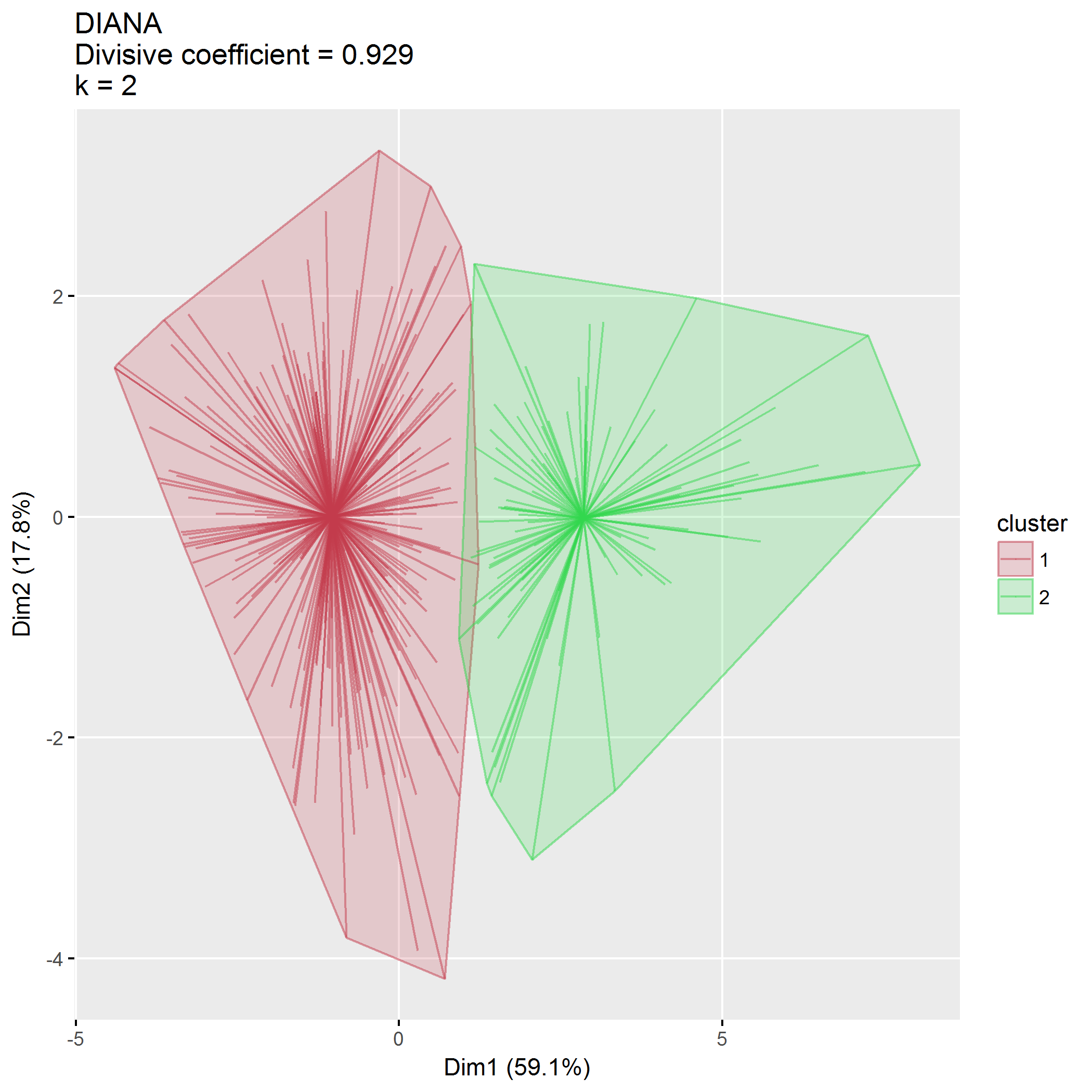
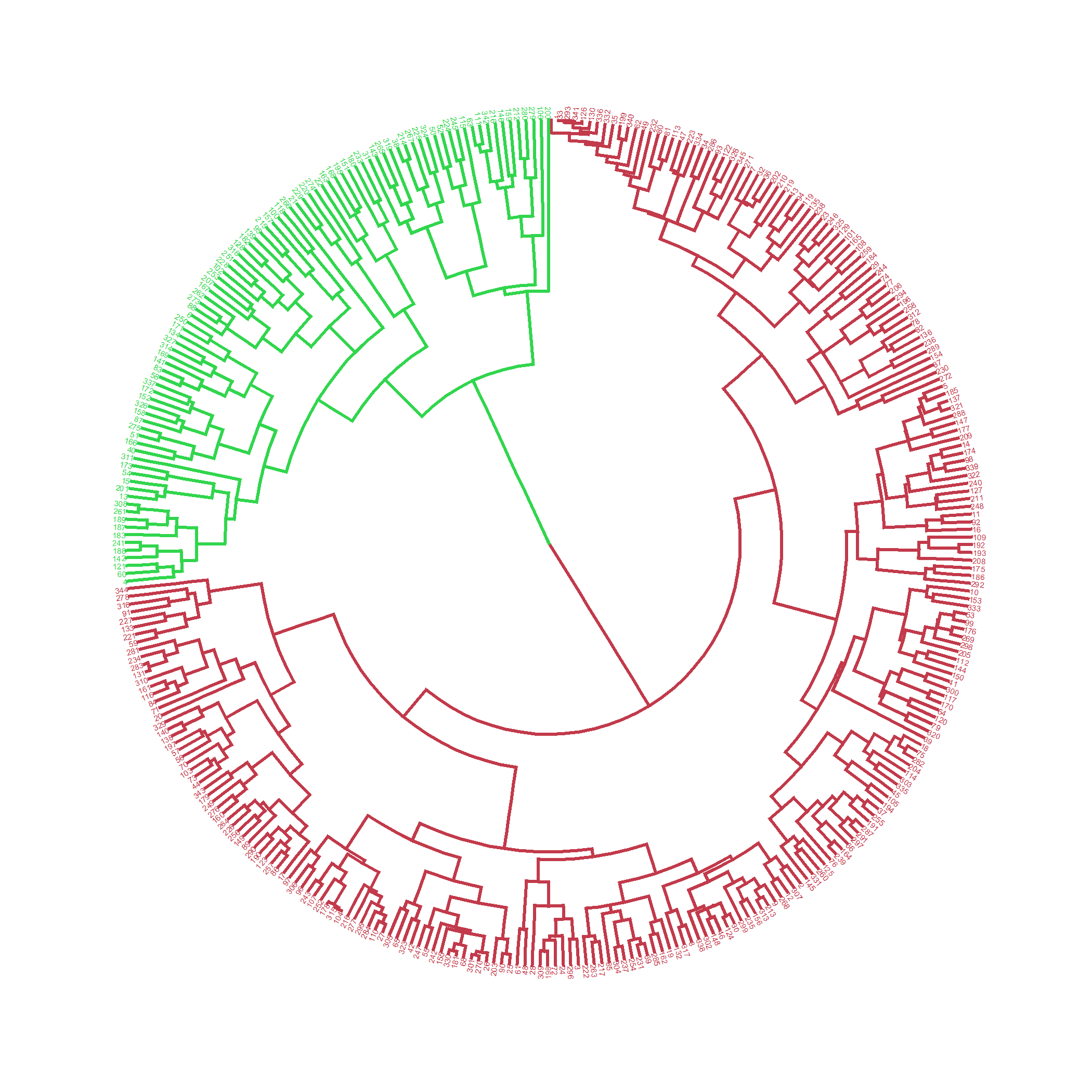
## [1] 345 8

## NULL

## [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

## .  
## 1 2   
## 253 92



* 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 | PCB2\_Tot\_mean | PCB3\_Total\_mean | PCB3\_PCPonly\_mean | PCB3\_Person\_mean | PCB3\_Resource\_mean |
| 1 | 253 | 74.0 | 53.7 | 20.3 | 25.9 | 51.7 | 4.4 | 17.4 | 30 |
| 2 | 92 | 47.1 | 33.6 | 13.5 | 20.0 | 34.2 | 3.3 | 11.9 | 19 |

## Cluster on ECBI, MAPS, SEPTI metrics

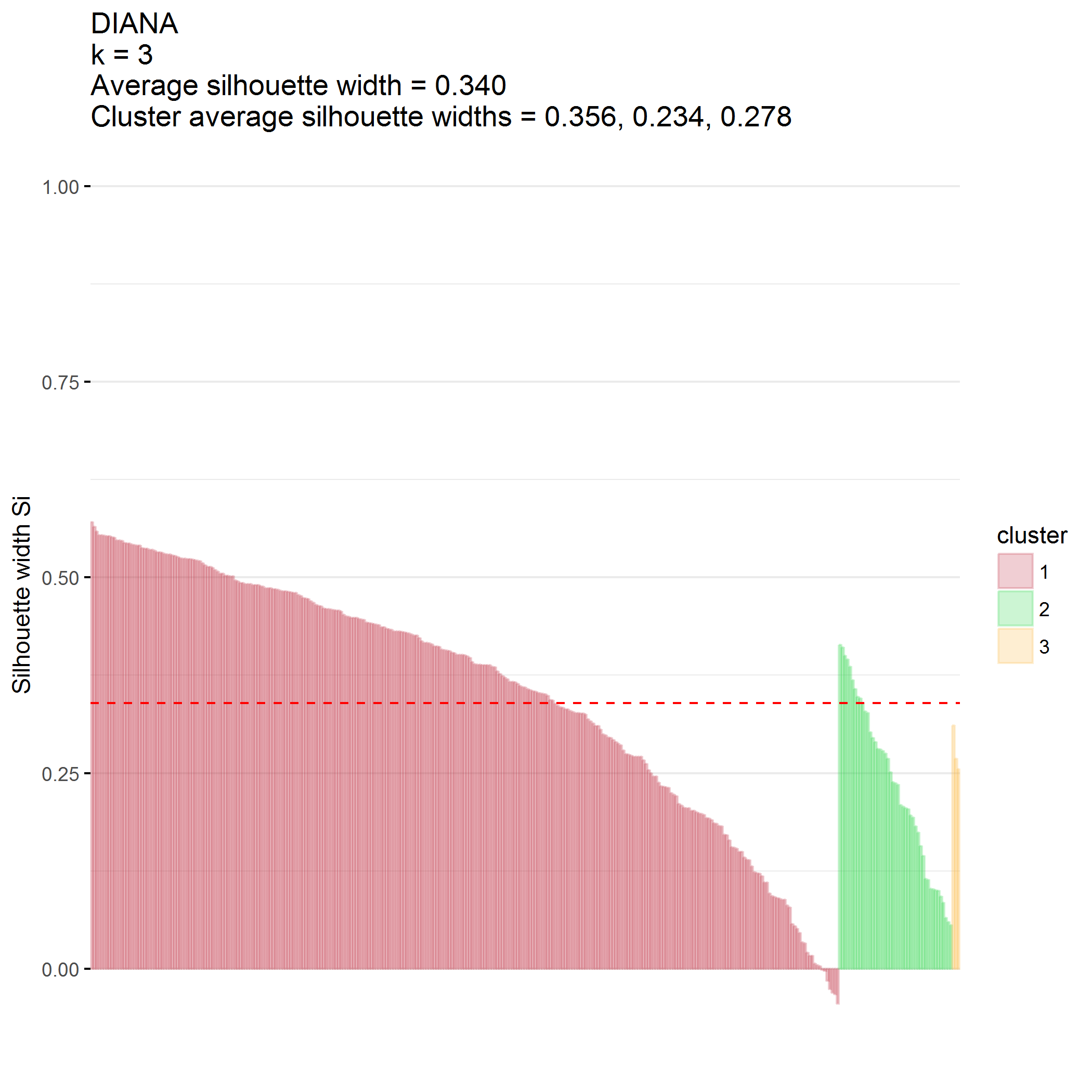
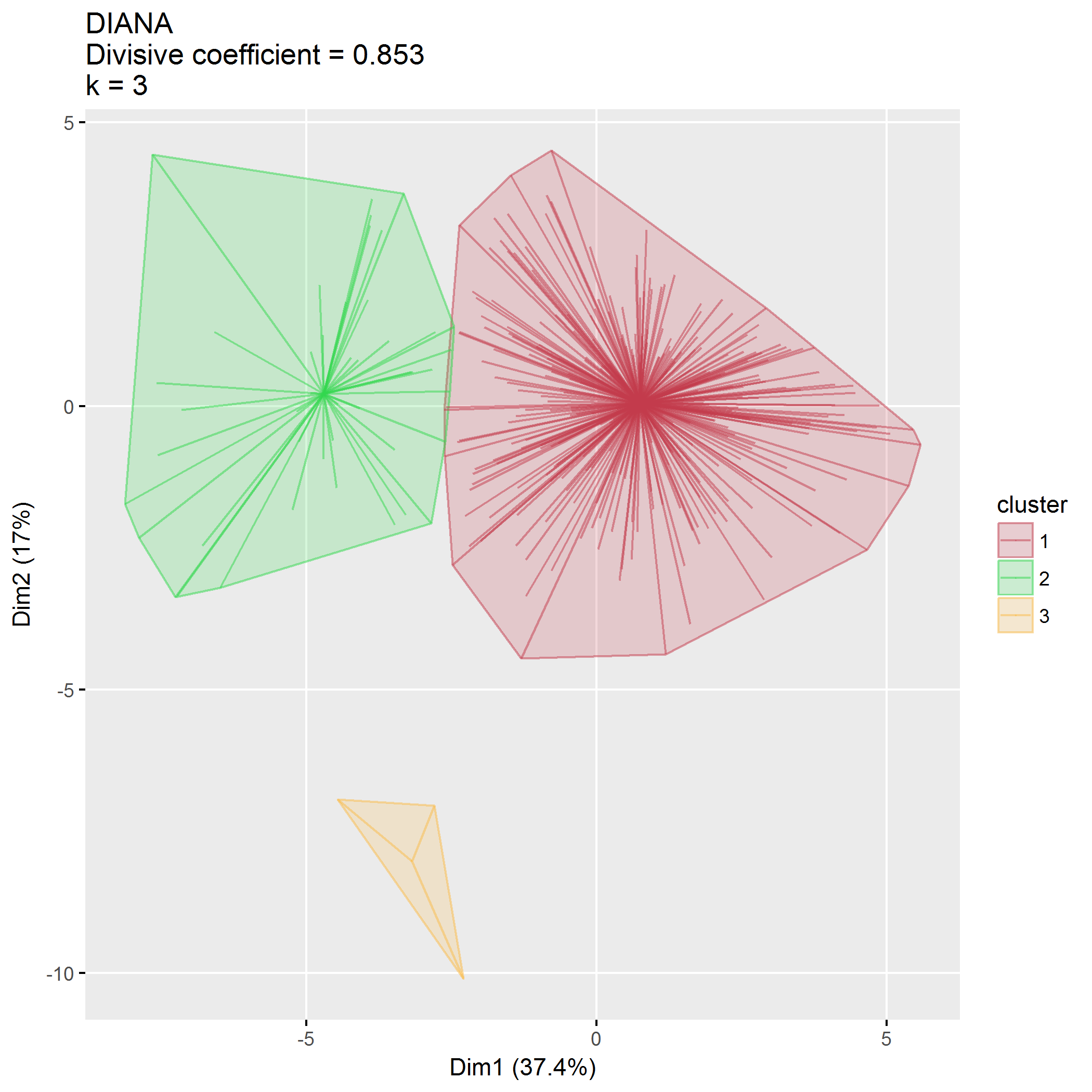
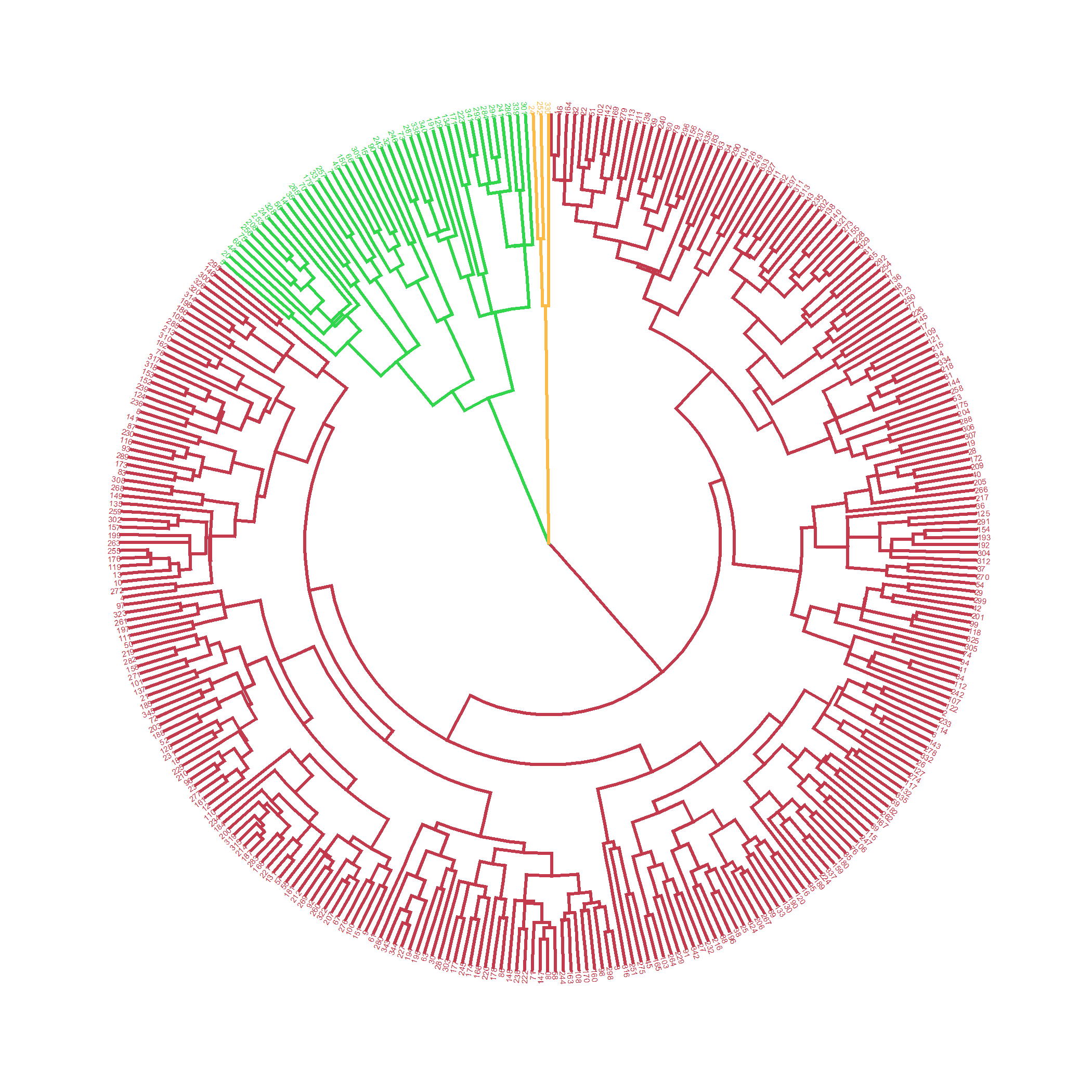
## [1] 345 18

## NULL

## [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

## .  
## 1 2 3   
## 297 45 3



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

## # A tibble: 3 x 7  
## cluster n ECBI\_intensity\_T\_score~ ECBI\_problem\_T\_scor~ ECBI\_Opp\_mean  
## <fct> <int> <dbl> <dbl> <dbl>  
## 1 1 297 52.5 52.2 32.3  
## 2 2 45 62.3 64.6 43.2  
## 3 3 3 37.7 44.3 17.7  
## # ... with 2 more variables: ECBI\_Inatt\_mean <dbl>, ECBI\_Cond\_mean <dbl>

## Error in inherits(x, "list"): argument "x" is missing, with no default

## # A tibble: 3 x 7  
## cluster n ECBI\_intensity\_T\_score~ ECBI\_problem\_T\_scor~ ECBI\_Opp\_mean  
## <fct> <int> <dbl> <dbl> <dbl>  
## 1 1 297 52.5 52.2 32.3  
## 2 2 45 62.3 64.6 43.2  
## 3 3 3 37.7 44.3 17.7  
## # ... with 2 more variables: ECBI\_Inatt\_mean <dbl>, ECBI\_Cond\_mean <dbl>

## Error in inherits(x, "list"): argument "x" is missing, with no default

## # A tibble: 3 x 11  
## cluster n MAPS\_PP\_mean MAPS\_PR\_mean MAPS\_WM\_mean MAPS\_SP\_mean  
## <fct> <int> <dbl> <dbl> <dbl> <dbl>  
## 1 1 297 4.14 4.59 4.69 4.53  
## 2 2 45 3.65 3.96 4.07 3.64  
## 3 3 3 2.39 2.92 3.22 2.11  
## # ... with 5 more variables: MAPS\_HS\_mean <dbl>, MAPS\_LC\_mean <dbl>,  
## # MAPS\_PC\_mean <dbl>, MAPS\_POS\_mean <dbl>, MAPS\_NEG\_mean <dbl>

## Error in inherits(x, "list"): argument "x" is missing, with no default

## # A tibble: 3 x 6  
## cluster n SEPTI\_nurturance\_mean SEPTI\_discipline\_me~ SEPTI\_play\_mean  
## <fct> <int> <dbl> <dbl> <dbl>  
## 1 1 297 38.1 24.3 32.8  
## 2 2 45 33.4 18.8 25.0  
## 3 3 3 26.3 19.3 25.7  
## # ... with 1 more variable: SEPTI\_routine\_mean <dbl>

## Error in inherits(x, "list"): argument "x" is missing, with no default

## Cluster on parent/child factors

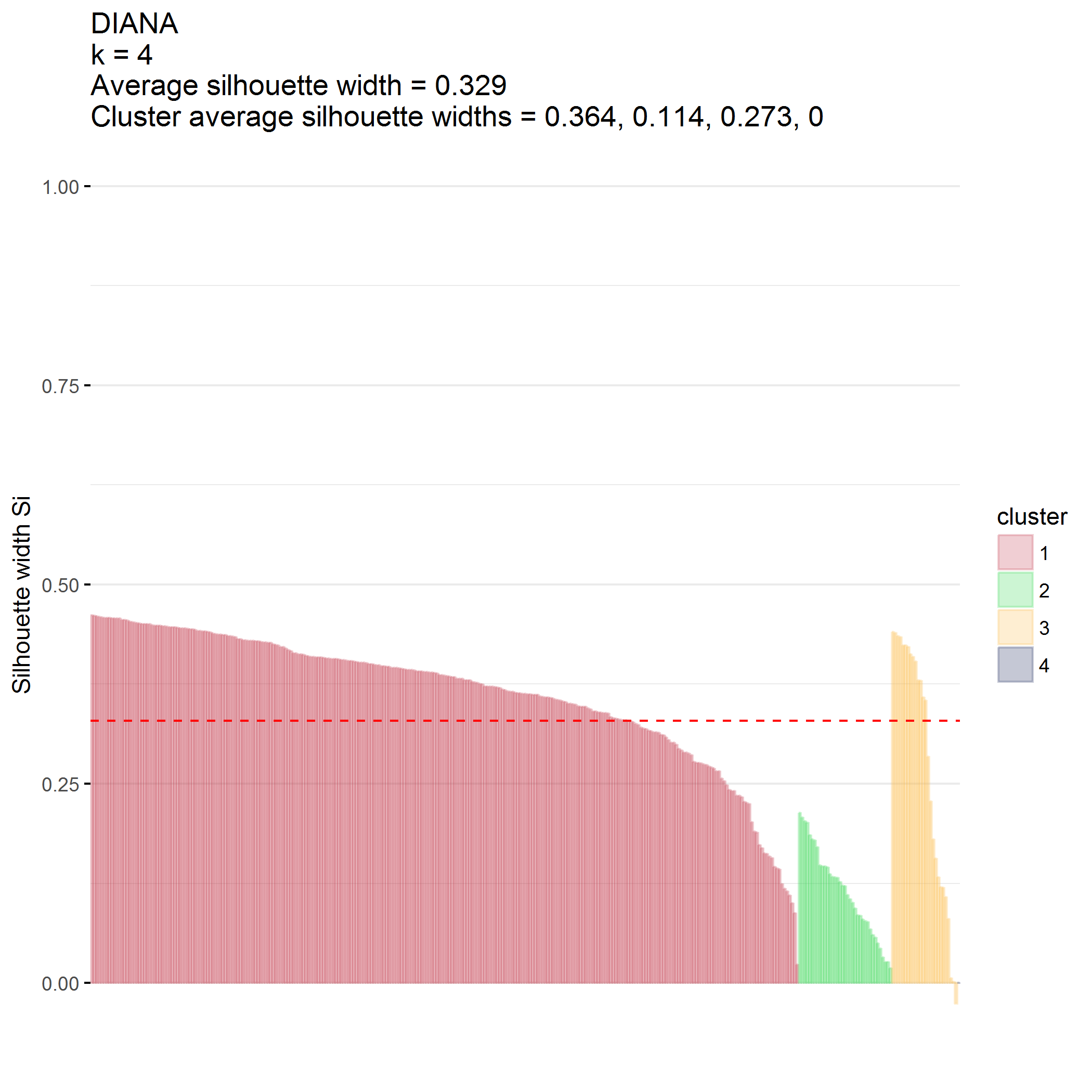
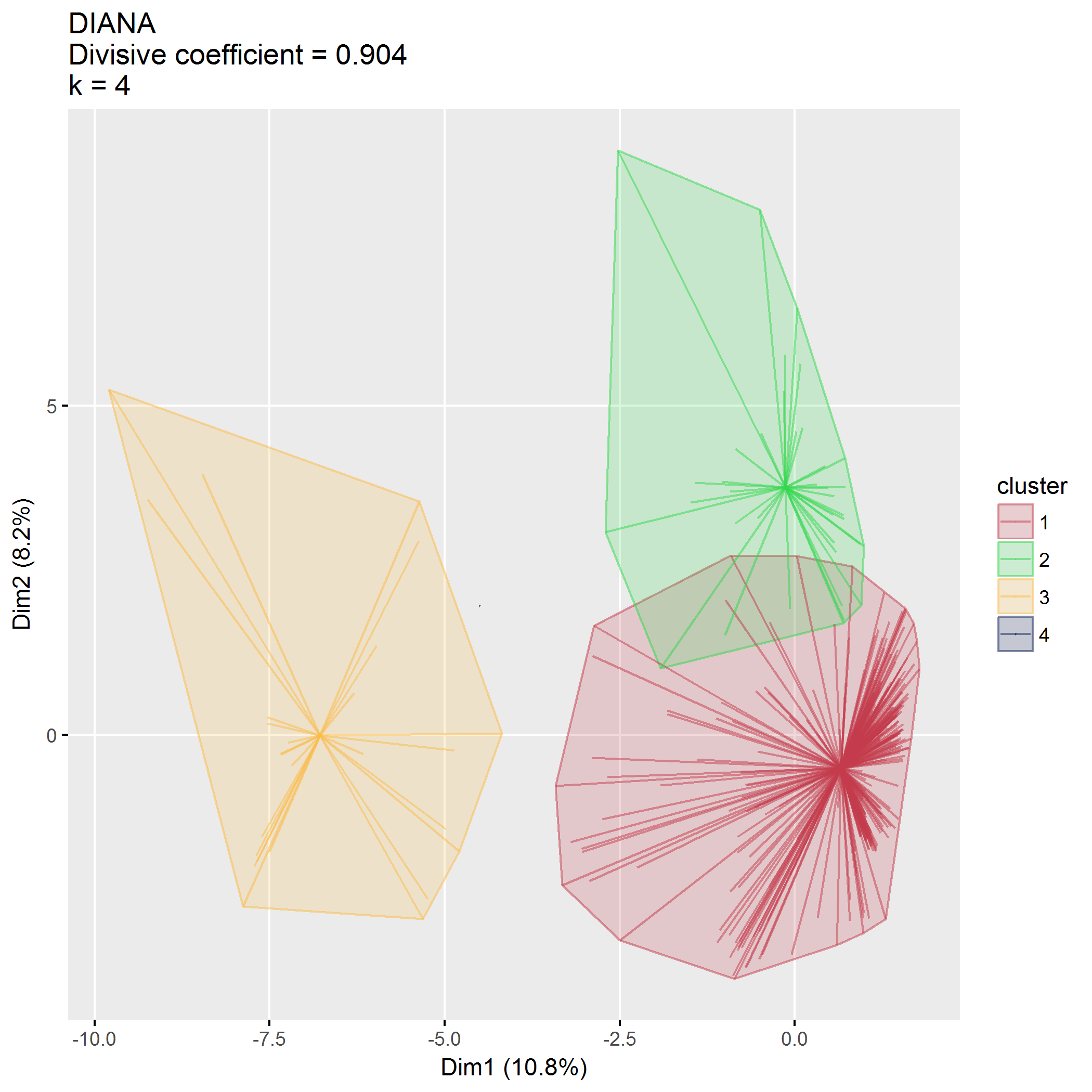
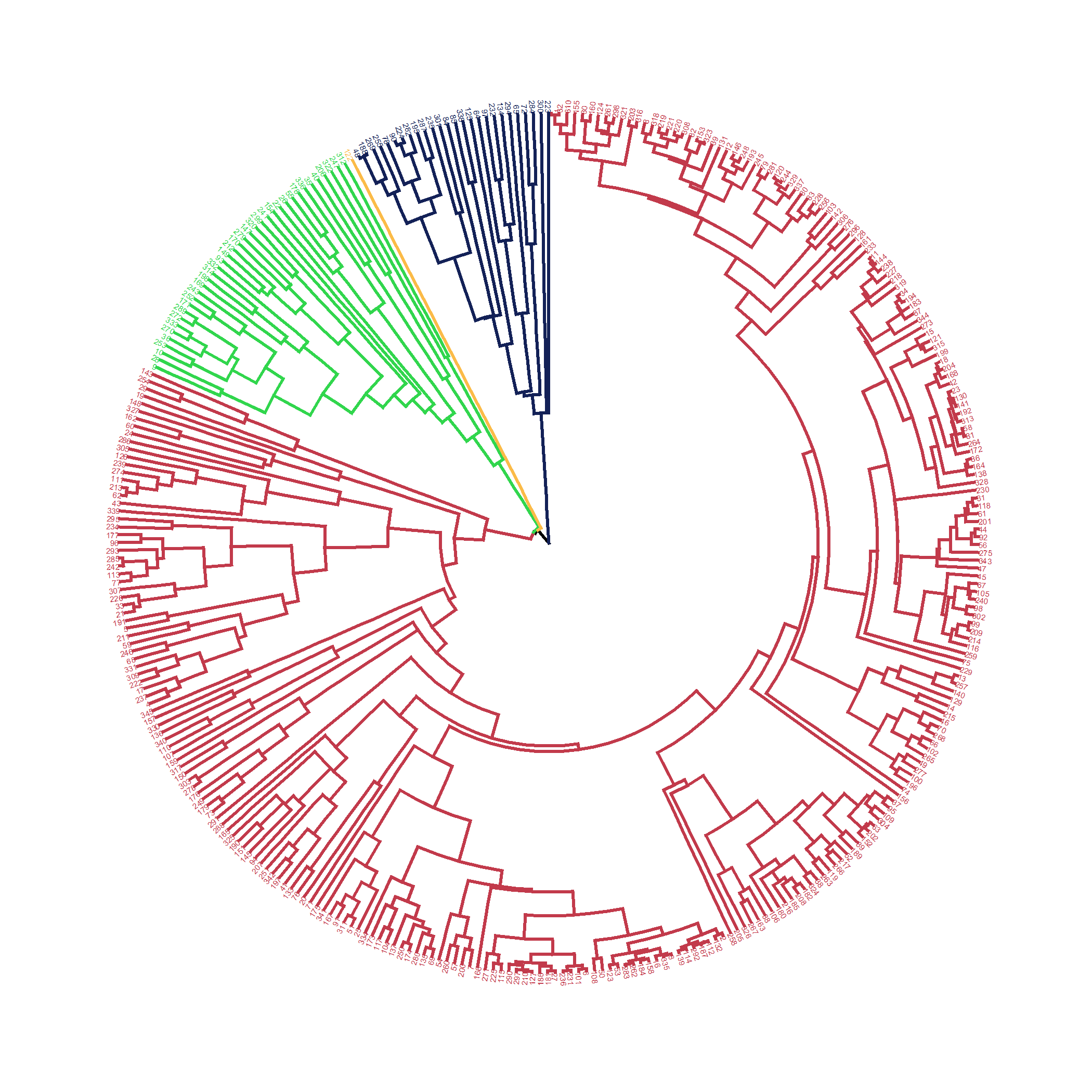
## [1] 345 48

## NULL

## [1] "totalChildren"   
## [2] "birthOrderOnly child"   
## [3] "birthOrderOldest"   
## [4] "birthOrderMiddle"   
## [5] "birthOrderYoungest"   
## [6] "childSexMale"   
## [7] "childAge"   
## [8] "childEthnicityNot Hispanic/Latino"   
## [9] "childEthnicityUnknown"   
## [10] "childEthnicityPrefer not to respond"   
## [11] "childRaceWhite1"   
## [12] "childRaceAsian1"   
## [13] "childRaceAfrAm1"   
## [14] "childRaceAIAN1"   
## [15] "childRaceNHPI1"   
## [16] "childRaceOther1"   
## [17] "childRaceNoResp1"   
## [18] "childRelationshipBiological or adoptive father"   
## [19] "childRelationshipGrandparent"   
## [20] "childRelationshipOther"   
## [21] "parentGenderFemale"   
## [22] "parentGenderTransgender"   
## [23] "parentGenderOther"   
## [24] "parentGenderPrefer not to respond"   
## [25] "parentSexMale"   
## [26] "parentAge"   
## [27] "parentEthnicityNot Hispanic/Latino"   
## [28] "parentEthnicityUnknown"   
## [29] "parentEthnicityPrefer not to respond"   
## [30] "parentRaceWhite1"   
## [31] "parentRaceAsian1"   
## [32] "parentRaceAfrAm1"   
## [33] "parentRaceAIAN1"   
## [34] "parentRaceNHPI1"   
## [35] "parentRaceOther1"   
## [36] "parentRaceNoResp1"   
## [37] "parentMaritalStatusWidowed"   
## [38] "parentMaritalStatusDivorced"   
## [39] "parentMaritalStatusSeparated"   
## [40] "parentMaritalStatusRemarried"   
## [41] "parentMaritalStatusNever married"   
## [42] "parentSituationCouple parenting with spouse or partner in the same household"  
## [43] "parentSituationCo-parenting in separate households"   
## [44] "parentsNumber"   
## [45] "parentChildRatio"   
## [46] "parentEducationVocational school/some college"   
## [47] "parentEducationCollege"   
## [48] "parentEducationGraduate/professional school"

## cluster size ave.sil.width  
## 1 1 281 0.36  
## 2 2 37 0.11  
## 3 3 26 0.27  
## 4 4 1 0.00

## .  
## 1 2 3 4   
## 281 37 26 1



* Hopkins statistic is 0.109
* Analysis identified clusters
* Divisive coefficient is 0.904
* Average silhouette width is 0.329

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| cluster | birthOrder | n | totalChildren\_mean | totalChildren\_median |
| 1 | Only child | 92 | 1.0 | 1.0 |
| 1 | Oldest | 84 | 2.2 | 2.0 |
| 1 | Middle | 38 | 3.6 | 3.0 |
| 1 | Youngest | 67 | 2.4 | 2.0 |
| 2 | Only child | 11 | 1.0 | 1.0 |
| 2 | Oldest | 9 | 2.1 | 2.0 |
| 2 | Middle | 4 | 3.5 | 3.5 |
| 2 | Youngest | 13 | 2.8 | 3.0 |
| 3 | Only child | 9 | 1.0 | 1.0 |
| 3 | Oldest | 3 | 3.0 | 2.0 |
| 3 | Middle | 1 | 2.0 | 2.0 |
| 3 | Youngest | 13 | 2.3 | 2.0 |
| 4 | Youngest | 1 | 5.0 | 5.0 |

## Cluster on demographic factors

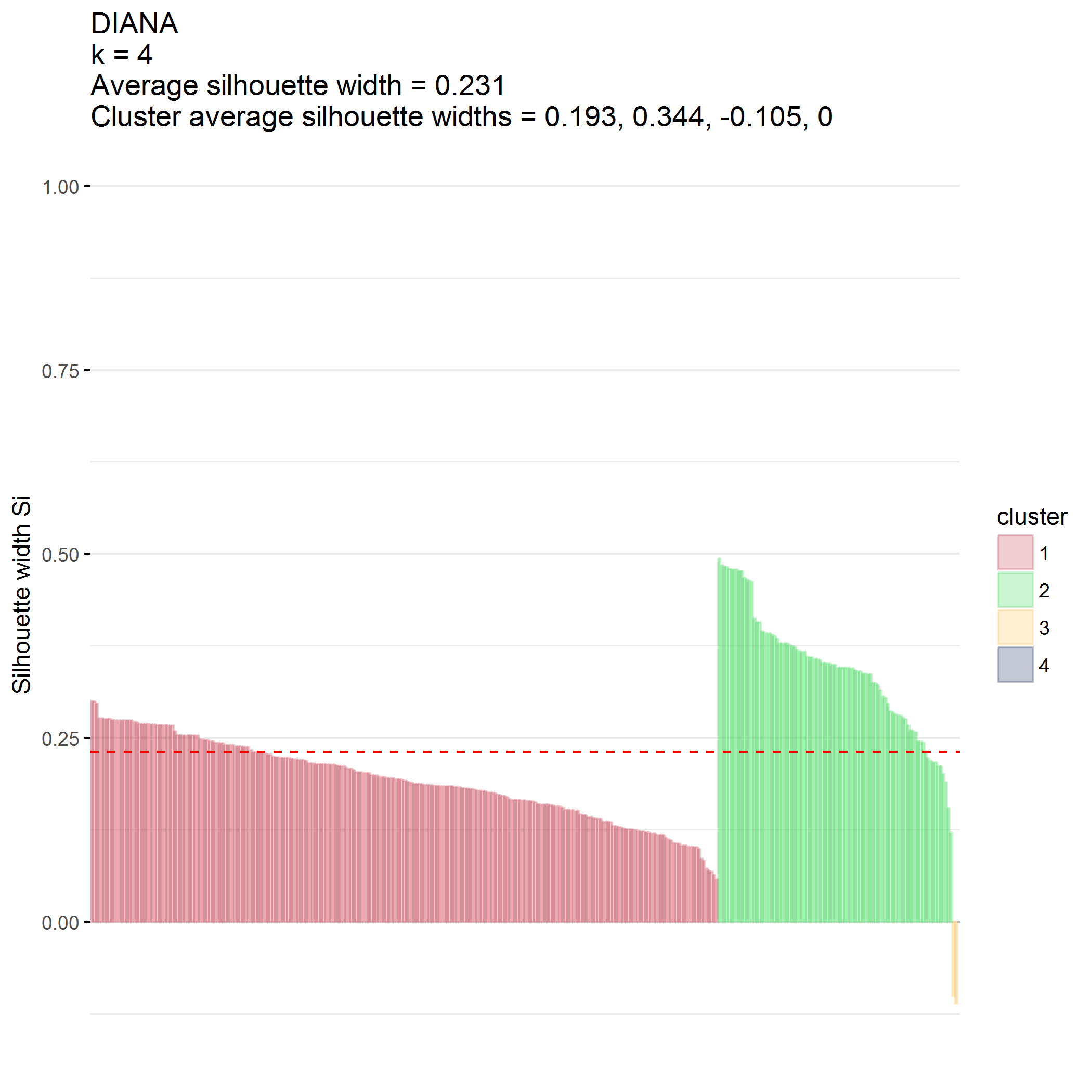
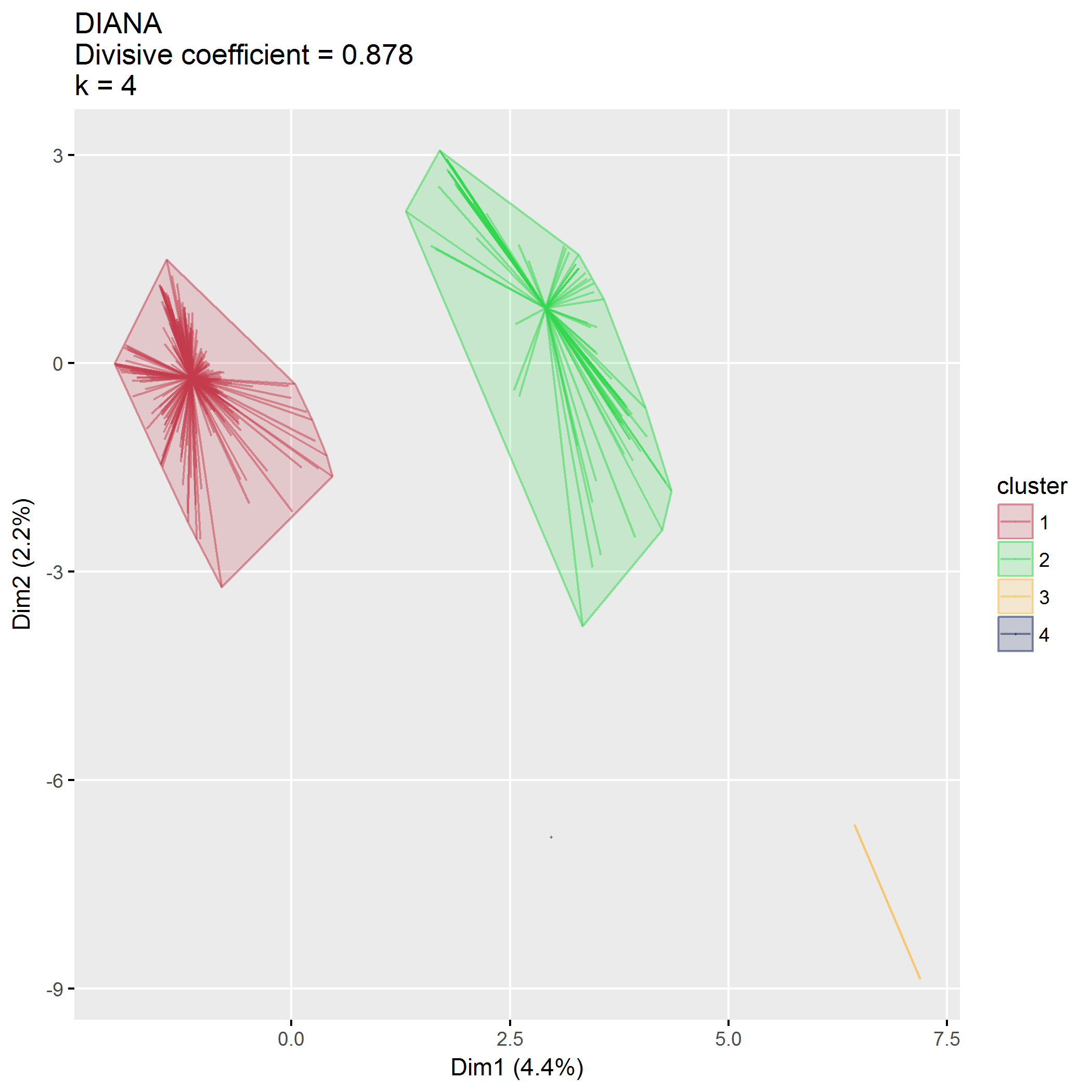
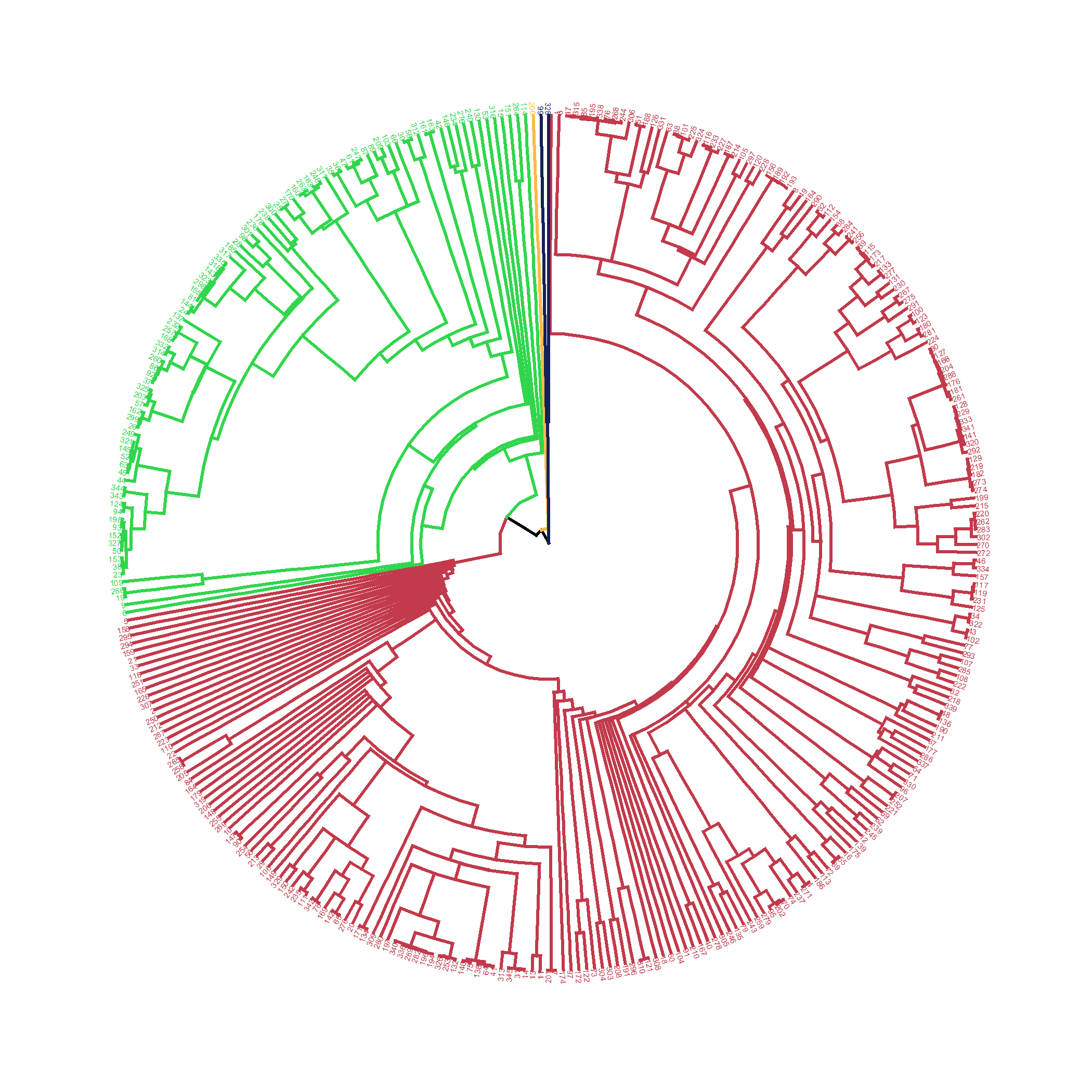
## [1] 345 88

## NULL

## [1] "zipcodeClass1" "zipcodeClass2"   
## [3] "zipcode91020" "zipcode91204"   
## [5] "zipcode91206" "zipcode91210"   
## [7] "zipcode91402" "zipcode97003"   
## [9] "zipcode97006" "zipcode97007"   
## [11] "zipcode97008" "zipcode97023"   
## [13] "zipcode97027" "zipcode97032"   
## [15] "zipcode97034" "zipcode97035"   
## [17] "zipcode97045" "zipcode97056"   
## [19] "zipcode97060" "zipcode97062"   
## [21] "zipcode97068" "zipcode97071"   
## [23] "zipcode97078" "zipcode97086"   
## [25] "zipcode97089" "zipcode97101"   
## [27] "zipcode97116" "zipcode97123"   
## [29] "zipcode97124" "zipcode97140"   
## [31] "zipcode97141" "zipcode97201"   
## [33] "zipcode97202" "zipcode97203"   
## [35] "zipcode97206" "zipcode97209"   
## [37] "zipcode97210" "zipcode97211"   
## [39] "zipcode97212" "zipcode97213"   
## [41] "zipcode97214" "zipcode97215"   
## [43] "zipcode97217" "zipcode97219"   
## [45] "zipcode97220" "zipcode97221"   
## [47] "zipcode97222" "zipcode97223"   
## [49] "zipcode97224" "zipcode97225"   
## [51] "zipcode97227" "zipcode97229"   
## [53] "zipcode97230" "zipcode97232"   
## [55] "zipcode97233" "zipcode97236"   
## [57] "zipcode97239" "zipcode97266"   
## [59] "zipcode97267" "zipcode97321"   
## [61] "zipcode97325" "zipcode97429"   
## [63] "zipcode97527" "zipcode97701"   
## [65] "zipcode97702" "zipcode97703"   
## [67] "zipcode97707" "zipcode97734"   
## [69] "zipcode97738" "zipcode97741"   
## [71] "zipcode97753" "zipcode97754"   
## [73] "zipcode97756" "zipcode97759"   
## [75] "zipcode97760" "zipcode98632"   
## [77] "zipcode98660" "zipcode98683"   
## [79] "zipcode98685" "communitySuburban"   
## [81] "communityRural" "distance"   
## [83] "income$25,001-$49,999" "income$50,000-$79,999"   
## [85] "income$80,000-$119,999" "income$120,000-$149,999"  
## [87] "income$150,000 or more" "internet"

## cluster size ave.sil.width  
## 1 1 249 0.19  
## 2 2 93 0.34  
## 3 3 2 -0.11  
## 4 4 1 0.00

## .  
## 1 2 3 4   
## 249 93 2 1



* Hopkins statistic is 0.041
* Analysis identified clusters
* Divisive coefficient is 0.878
* Average silhouette width is 0.231

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| cluster | zipcodeClass | community | distanceLong | n |
| 1 | 1 | Urban | FALSE | 109 |
| 1 | 1 | Suburban | FALSE | 128 |
| 1 | 1 | Rural | FALSE | 12 |
| 2 | 2 | Urban | FALSE | 13 |
| 2 | 2 | Suburban | FALSE | 29 |
| 2 | 2 | Suburban | TRUE | 1 |
| 2 | 2 | Rural | FALSE | 50 |
| 3 | 2 | Rural | TRUE | 2 |
| 4 | 1 | Rural | FALSE | 1 |

|  |  |  |  |
| --- | --- | --- | --- |
| cluster | zipcodeClass | income | n |
| 1 | 1 | $25,000 or less | 25 |
| 1 | 1 | $25,001-$49,999 | 57 |
| 1 | 1 | $50,000-$79,999 | 58 |
| 1 | 1 | $80,000-$119,999 | 38 |
| 1 | 1 | $120,000-$149,999 | 27 |
| 1 | 1 | $150,000 or more | 44 |
| 2 | 2 | $25,000 or less | 10 |
| 2 | 2 | $25,001-$49,999 | 26 |
| 2 | 2 | $50,000-$79,999 | 31 |
| 2 | 2 | $80,000-$119,999 | 17 |
| 2 | 2 | $120,000-$149,999 | 5 |
| 2 | 2 | $150,000 or more | 4 |
| 3 | 2 | $50,000-$79,999 | 1 |
| 3 | 2 | $80,000-$119,999 | 1 |
| 4 | 1 | $50,000-$79,999 | 1 |

|  |  |  |  |
| --- | --- | --- | --- |
| cluster | zipcodeClass | internet | n |
| 1 | 1 | 0 | 4 |
| 1 | 1 | 1 | 245 |
| 2 | 2 | 0 | 4 |
| 2 | 2 | 1 | 89 |
| 3 | 2 | 1 | 2 |
| 4 | 1 | 0 | 1 |

## Save objects

## size isdir mode  
## data/processed/clusterAnalysis.RData 10004125 FALSE 666  
## mtime  
## data/processed/clusterAnalysis.RData 2018-07-06 16:14:09  
## ctime  
## data/processed/clusterAnalysis.RData 2018-07-06 12:14:44  
## atime exe  
## data/processed/clusterAnalysis.RData 2018-07-06 12:14:44 no