# **Algorithm Illustration Hierachical Clustering Accuracy**

Step 1: We prepare the data:

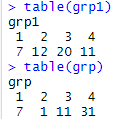
* Cluster data according to different standard
* Divide the processed data into 4 group (set label)

|  |
| --- |
| res.agnes <- agnes(x = USArrests, # data matrix  stand = TRUE, # Standardize the data  metric = "euclidean", # metric for distance matrix  method = "average" # Linkage method  )  # Cut tree into 4 groups  grp <- cutree(res.agnes, k = 4)  head(grp, n = 4) |

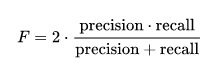
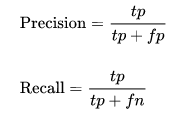
|  |
| --- |
| res.agnes1 <- agnes(x = USArrests, # data matrix  stand = TRUE, # Standardize the data  metric = "manhattan", # metric for distance matrix  method = "complete" # Linkage method  )  # Cut tree into 4 groups  grp1 <- cutree(res.agnes1, k = 4)  head(grp1, n = 4) |

Let's see how the results are different

|  |
| --- |
| table(grp)  table(grp1) |



Step 2: Create a function precision, recall and f-measure. Here are some formulas regarding to calculating:



|  |
| --- |
| measurePrecisionRecall <- function(predict, actual\_labels){  precision <- sum(predict & actual\_labels) / sum(predict)  recall <- sum(predict & actual\_labels) / sum(actual\_labels)  fmeasure <- 2 \* precision \* recall / (precision + recall)    cat('precision: ')  cat(precision \* 100)  cat('%')  cat('\n')    cat('recall: ')  cat(recall \* 100)  cat('%')  cat('\n')    cat('f-measure: ')  cat(fmeasure \* 100)  cat('%')  cat('\n')  } |

Step 3: Use the result in step 1 as predict and real label, then pass two parameters into the function.

|  |
| --- |
| measurePrecisionRecall(grp,grp1) |

