

```

nclust.R <- function(){
# Determines "optimal" number of clusters via
# average silhouette measures (see Julian Izenman (2008)
# "Modern Multivariate Statistical Techniques, Regression, Classification
# and Manifold Learning", Springer Verlag, pp 424-428).
# Requires libraries "factoextra" and "cluster".

# read data en load libraries
library("cluster")
library("factoextra")
data("USArrests")

# standardize data
df <- scale(USArrests)

# compute value of average silhouette measure (width) for different values of k (= number of clusters) .
# Then plot k (= number of clusters) vs. value of value of average silhouette measure
fviz_nbclust(df, pam, method = "silhouette")
ggsave("k= number of clusters vs average-silhouette width.pdf",device = "pdf")
}

```

Optimal number of clusters

