

k-means and k-prototypes algorithm



Import the dataset *DatasetTelecom*, which contains numerical variables and categorical variables transformed into numerical variables.

To standardize the dataset
`sTelecom=scale(DatasetTelecom)`

To perform a *k-means* analysis with 2 clusters and 50 starting points
`kTelecom2=kmeans(sTelecom,2,nstart=50)`

To visualize the main results
`kTelecom2`

To visualize the clusters
`fviz_cluster(kTelecom2,DatasetTelecom)`

If the number of clusters specified *a priori* ranges between 2 and 4
`AverageSilh=fviz_nbclust(sTelecom, kmeans, method='silhouette', k.max=4)`
`plot(AverageSilh)`

Import the dataset *DatasetStartups3*, which contains numerical variables and categorical variables.

Define the variable *State* as categorical:
`State=DatasetStartups3$State`
`State=as.factor(State)`
`DatasetStartups3$State=State`

After installing and loading the package *clustMixType*, apply the *k-prototypes* algorithm
`kStart2=kproto(DatasetStartups3,2)`

To visualize the main results
`plot(kStart2)`