

# Overview Tables

## Clustering methods

Method	Data	Cluster concept	clustering type	number of clusters?
k-means	Euclidean	homogeneous, spherical	partition	gap, ASW
Single Linkage	dissimilarities	separated	hierarchy	ASW
Average Linkage	dissimilarities	homogeneous, separated	hierarchy	ASW
Complete Linkage	dissimilarities	homogeneous	hierarchy	ASW
Ward	Euclidean	homogeneous, spherical	hierarchy	gap, ASW
pam	dissimilarities	homogeneous	partition	ASW
Gaussian mixture	Euclidean	elliptical	probabilistic	BIC
t mixture	Euclidean	elliptical, heavy tails	probabilistic	BIC
skew Gaussian mixture	Euclidean	skew	probabilistic	BIC
skew t mixture	Euclidean	skew, heavy tails	probabilistic	BIC
Latent class mixture	categorical	local independence	probabilistic	BIC
funFEM	functional	Gaussian basis coefs (subspace)	probabilistic	BIC

## Dissimilarities

Dissimilarity	Data	Distance?	Characteristic
Euclidean $L_2$	Euclidean	yes	Euclidean geometry, rotation invariant
Manhattan $L_1$	Euclidean	yes	all variables same weight
Maximum $L_\infty$	Euclidean	yes	maximum distance variable
Mahalanobis	Euclidean	yes	affine invariant (incl. scaling)
Simple Matching	categorical	yes	all categories treated equally
Jaccard	binary	yes	dominated by presences not absences
Correlation (1)	variables	no	max. dissimilarity $\rho = -1$
Correlation (2)	variables	no	max. dissimilarity $\rho = 0$
Gower	mixed type	no (with missings)	Manhattan aggregation, missing values

## Other methods

Method	Data	Task
Principal Components	Euclidean	dimension reduction, visualisation
Data scaling	Euclidean	making variables comparable
Multidimensional Scaling	dissimilarities	maps to Euclidean, visualisation
Adjusted Rand Index	clusterings	similarity between clusterings
Pairs plot	Euclidean	visualisation
Heatplot	general (shown for binary)	visualisation
Parallel coordinates plot	mostly Euclidean, time series	visualisation
gap, ASW, BIC results plots	clustering indexes	visualisation
Dendrogram	clustering hierarchy	visualisation
Silhouette plot	pointwise silhouette widths	visualisation
B-spline basis	functional	smooth data representation
functional PCs	functional	dimension reduction, visualisation