

Cvičení 8

1. Načtěte data popisující morfometrii bruslařek. Vyberte z nich data popisující samečky 6. instaru. Na tato data aplikujte k-means partitioning. Určete optimální počet clusterů. Proveďte k-means s tímto počtem a následně vykreslete do PCA diagramu, tak, že k-means klasifikaci vyjádříte barvou a skutečnou druhovou příslušnost zakleslete obálkou. Zkuste to ještě jednou pro skutečný počet druhů.
2. Načtěte data o pakomárech ze Svatky. Na základě logaritmovaných ($\ln(x+1)$) druhových dat vypočítejte Bray-Curtis matici nepodobností. Na získané distanční matici vypočítejte 3 následující shlukové analýzy a porovnejte jejich dendrogramy:
 1. single linkage (Nearest Neighbour)
 2. complete linkage (Farthest Neighbour)
 3. average linkage (Unweighted Pair Group Method with Arithmetic mean)
3. Nechejte si zobrazit kofenetické vzdálenosti proti původním (Bray-Curtis) vzdálenostem a vykreslete je ve scatter-plotu proti původním vzdálenostem. Přidejte do grafů přímkou odpovídající $y = x$. Porovnejte grafy a zamyslete se nad tím, co říkají

R funkce

kmeans k-means partitioning

hclust klastrová analýza

cophenetic matice kofenetických vzdáleností mezi vzorky.vegan:
vegan:

cascadeKM k-means partitioning provedené přes sérii k, diagnostika pomocí Calinski - Harabasz kritéria pomocí `plot(cascadeKM.objekt)`