

MAD - pracovní list 3

Analýza shluků

1. část - Společná práce

Dnes nás čeká jedno z nejnáročnějších témat. Shluková analýza je vlastně celý soubor metod. Narozdíl od většiny ostatních vícerozměrných metod není shluková metoda závislá na předpokladu normality dat či homoskedasticity, na druhou stranu použití vhodné metody (a vůbec opodstatnění jejího použití!) závisí na povaze úlohy a také na zkušenostech a subjektivním uvážení samotného badatele. Navíc je třeba dvou předpokladů: reprezentativnost obdrženého vzorku a vyloučení multikolinearity.

Z webu si stáhněte soubor zdrojový kód **cluster_analysis.R** a otevřete si jej v **R**. Dnes budeme pracovat s daty z databáze softwaru **R**, není tedy třeba stahovat žádná data.

2. část - samostatná práce v R

Vášim dnešním samostatným úkolem je prouzkoušet použití jiných metod a vzdáleností (popřípadě jiný počet předepsaných shluků) ve shlukové analýze daných dat **wine**. K nalezení příslušných příkazů použijte funkci **help**. Grafické výstupy exportujte z **R** a porovnejte.