

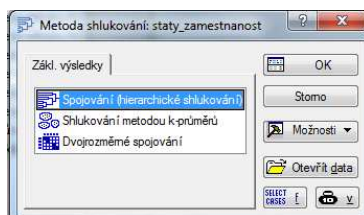
Shluková analýza

Zadání: Metodou shlukování analyzujte strukturu zaměstnanosti ve 26 evropských zemích. Hledejte skupiny států s podobnou strukturou vstupních proměnných. Máte k dispozici tabulku, která pro 26 evropských zemí shrnuje údaje o zaměstnanosti v devíti různých průmyslových odvětvích (příklad je převzat z publikace Hendl, J. (2004) Přehled statistických metod zpracování dat (příklad 13.8, str. 465–467.

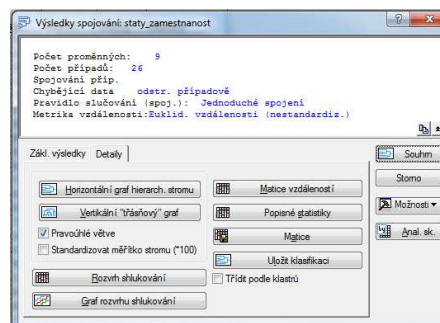
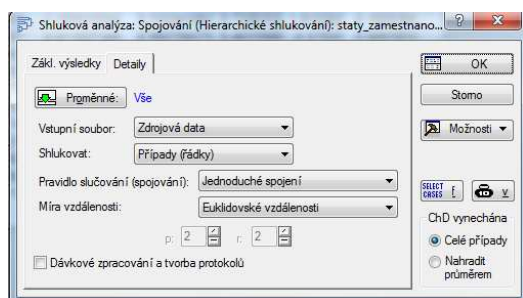
Poznámka: k vypracování použijte soubor **staty_zamestnanost.xls**. Vlastní výpočet proveďte podle dále uvedených instrukcí a interpretaci výsledků podle výše uvedené publikace a poznámek v přednášce

Základní kroky zpracování

- 1) V programu Statistica otevřete soubor **staty_zamestnanost.xls** tak, aby první řádek představoval názvy proměnných a první sloupec názvy případů.
- 2) Tabulku původních dat převedte na tzv. z-skóre (**Data – Standardizace**)
- 3) Spustíte modul pro analýzu hlavních komponent poslopností příkazů: **Statistiky – Vícerozměrné průzkumné techniky – Shluková analýza** a nejprve zvolte metodu hierarchického shlukování



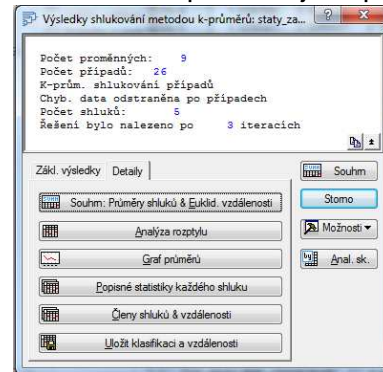
- 4) V zadávacím panelu na záložce **Detaily** zvolte všech devět proměnných jako vstup pro vlastní shlukování a ostatní nastavení upravte podle níže uvedeného obrázku.



- 5) Z panelu, který na záložce **Detaily** nabízí výsledky shlukové analýzy (viz obr. nahoře vpravo) postupně přeneste či zkopírujte do výsledného protokolu a vhodně prezentujte v podobě tabulky či grafu následující dílčí výsledky:
 - Dendrogram v horizontálním tvaru
 - Rozvrh shlukování
 - Graf rozvrhu shlukování
- 6) Na základě informací z přednášky, prezentovaných řešených příkladů a výše uvedených výstupů nalezněte vhodný počet shluků a vytvořené shluky států s obdobnou strukturou zaměstnanosti stručně charakterizujte
- 7) Využijte matici vzdáleností k nalezení státu, který má „typickou strukturu zaměstnanosti“ (je všem ostatním nejbliže) a dále 2–3 států, které jsou všem ostatním „nejvzdálenější“. Pokuste se zdůvodnit

- 8) Stejnou úlohu nyní řešte pomocí metody nehierarchického shlukování (Shlukování pomocí k-průměrů)
- 9) V zadávacím panelu na záložce **Detaily** opět zvolte všechny proměnné. Počet shluků nastavte na základě vaší interpretace výsledků hierarchického shlukování, počet iterací min. 20.
- 10) Z panelu, který na záložce **Detaily** nabízí výsledky shlukové analýzy (viz obr. dole) postupně přeneste či zkopírujte do výsledného protokolu a vhodně prezentujte v podobě tabulky či grafu následující dílčí výsledky:

- Členy shluků & vzdálenosti
- Graf průměrů
- Analýza rozptylu



- 11) Popište, k čemu slouží a co v tomto příkladu ukazuje Graf průměrů a k čemu zde slouží Analýza rozptylu

Závěr: Stručně porovnejte výsledky obou způsobů shlukování