

Téma č. 10.: Aplikace shlukové analýzy

Článek Ladislava Rabušice Koho Češi nechtějí? (uveřejněn ve Sborníku prací FSS MU Sociální studia 5, 2000) se zabývá touto problematikou:

V roce 1999 proběhlo ve 24 evropských zemích sociologické šetření, v němž měli respondenti za úkol odpovědět na otázku „Můžete prosím z následujícího seznamu vybrat všechny ty, koho byste nechtěl(a) mít za sousedy?“

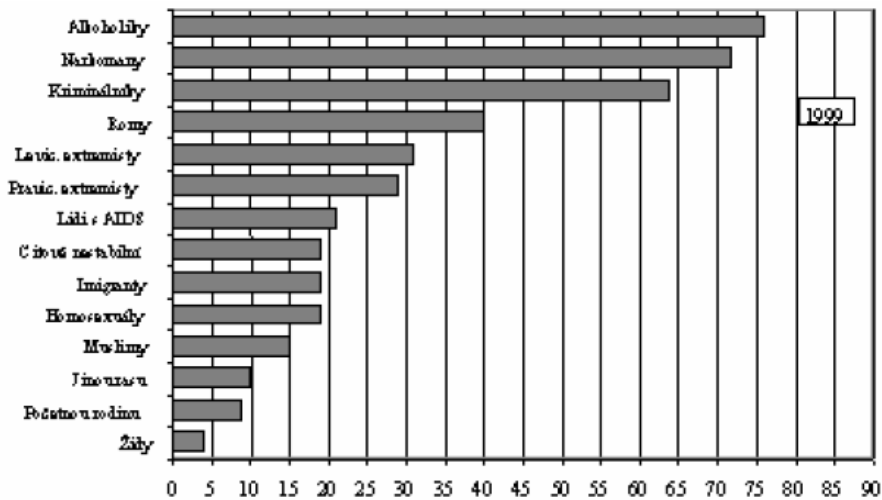
V seznamu byly tyto skupiny osob:

Kriminálníci, osoby jiné rasy, levicoví extrémisté, alkoholici, pravicoví extrémisté, početné rodiny, citově nestabilní lidé, muslimové, imigranti, lidé s AIDS, narkomani, homosexuálové, židé, Romové.

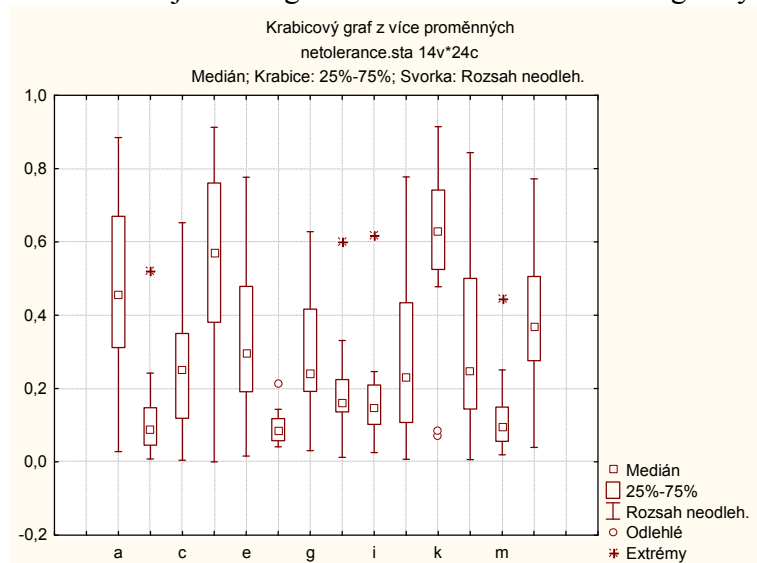
V datovém souboru netolerance.sta jsou zaznamenány relativní četnosti vybraných skupin osob.

V České republice se výzkumu, který proběhl v květnu 1999, zúčastnilo 1908 osob.

Úkol 1.: Zaměřte se na ČR. Vytvořte sloupkový diagram tohoto tvaru:

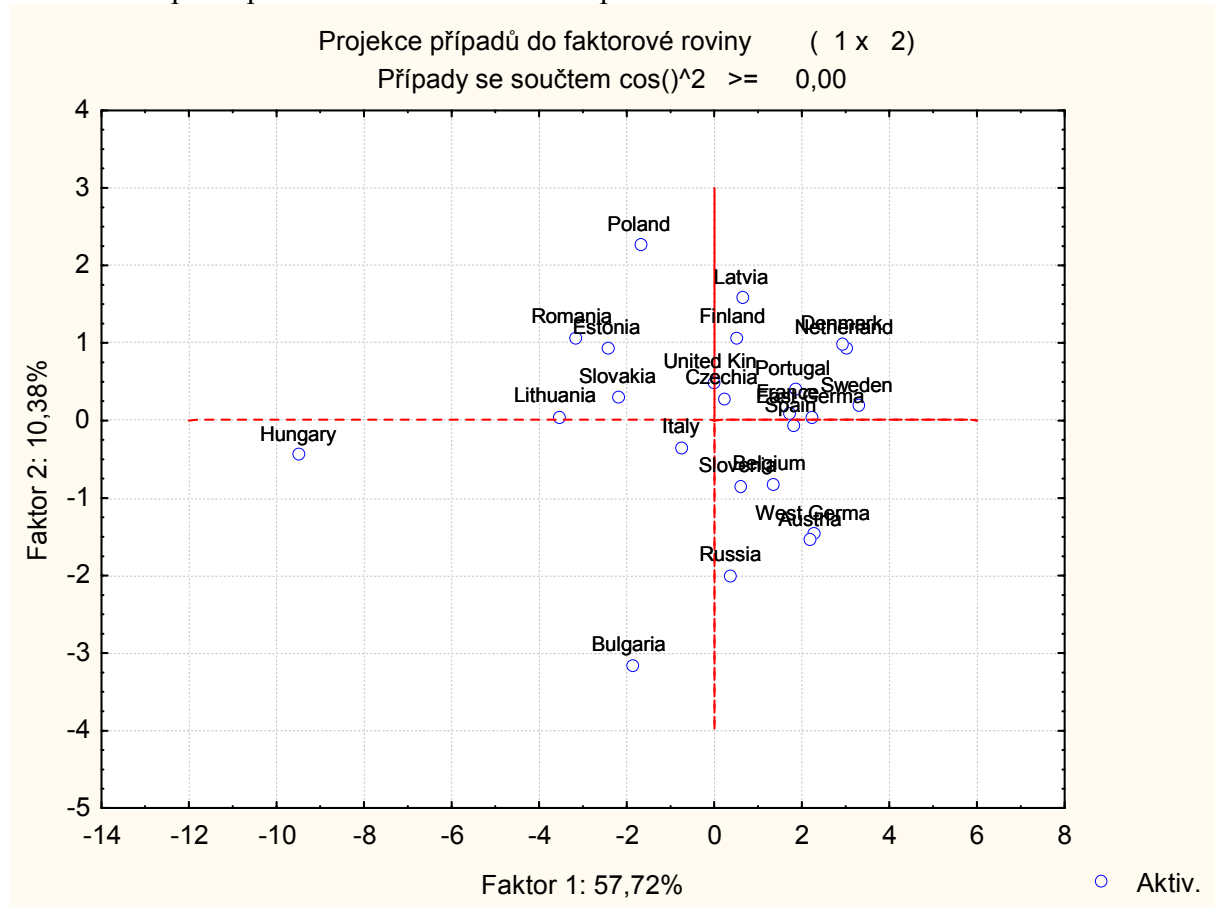


Úkol 2.: Do jednoho grafu nakreslete krabicové diagramy všech 14 proměnných.



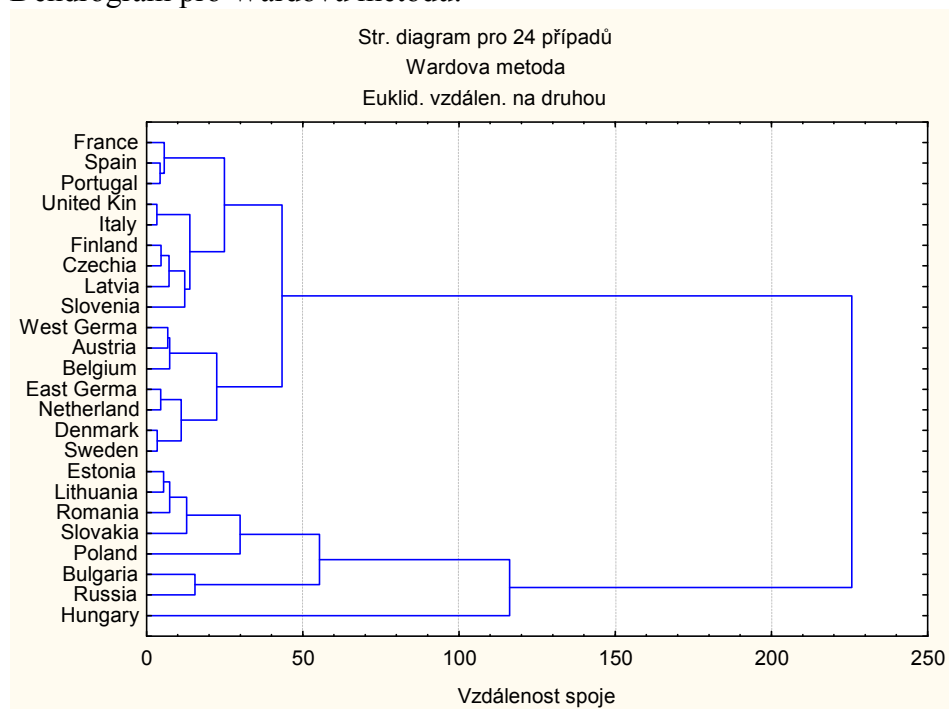
Vzhledem k velmi rozdílné variabilitě proměnných se jeví vhodnější pracovat se standardizovanými daty.

Úkol 3.: Na ploše prvních dvou hlavních komponent znázorněte rozmístění zemí.



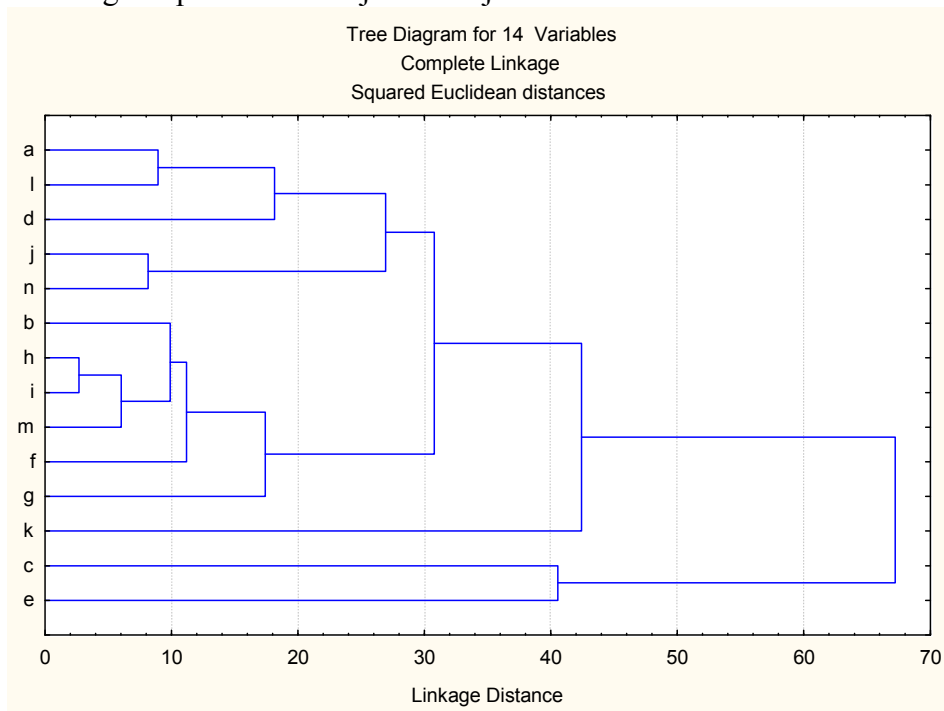
Úkol 4.: Použijte metodu nejbližšího souseda, nejbzdálenějšího souseda, metodu průměrné vazby a Wardovu metodu pro nalezení shluků zemí podobných z hlediska tolerance. Výsledky znázorněte pomocí dendrogramů.

Dendrogram pro Wardovu metodu:



Úkol 5.: Určete 5 shluků navzájem si podobných zemí.

Úkol 6.: Proveďte shlukovou analýzu pro proměnné.
Dendrogram pro metodu nejvzdálenějšího souseda:



Úkol 7.: Použijte metodu k-průměrů k nalezení 5 shluků navzájem si podobných zemí a uložte skupinovou příslušnost do datového souboru. K určení významnosti jednotlivých proměnných proveďte analýzu rozptylu. Nakreslete graf průměrů všech 5 shluků a pokuste se o interpretaci.

Graf průměrů:

