

## Téma č. 4.: Provedení shlukové analýzy

**Příklad:** Působení neuroleptik na CNS (Příklad je převzat z knihy Meloun M., Militký J., Hill, M.: Počítačová analýza vícerozměrných dat v příkladech. Academia Praha 2005) Neuroleptika redukují nežádoucí účinky přebytečného dopaminu. Liší se však ve svých účincích. Potlačují např. nervozitu, třes, ospalost apod. Tlumicí účinek neuroleptika se vyjadřuje mediánovou účinnou dávkou ED50 (v mg/kg), která představuje miligramy léčiva na 1 kg hmotnosti organismu utlumené poloviny celkového počtu pokusných zvířat. Nízká hodnota ED50 značí výkonnější léčivo, které působí již v malém množství. Často se však používá převrácená hodnota 1/ED50 (v kg/mg), která vyjadřuje přímou úměru mezi množstvím neuroleptika a a tlumicí aktivitou. Cílem analýzy je provést klasifikaci neuroleptik do shluků, které vykazují podobné účinky.

Datový soubor neuroleptika.sta obsahuje údaje o 20 přípravcích. Byly sledovány tyto proměnné (proměnné  $X_1$  až  $X_4$  jsou převrácenými hodnotami mediánové účinné dávky):

$X_1$  ... název neuroleptika

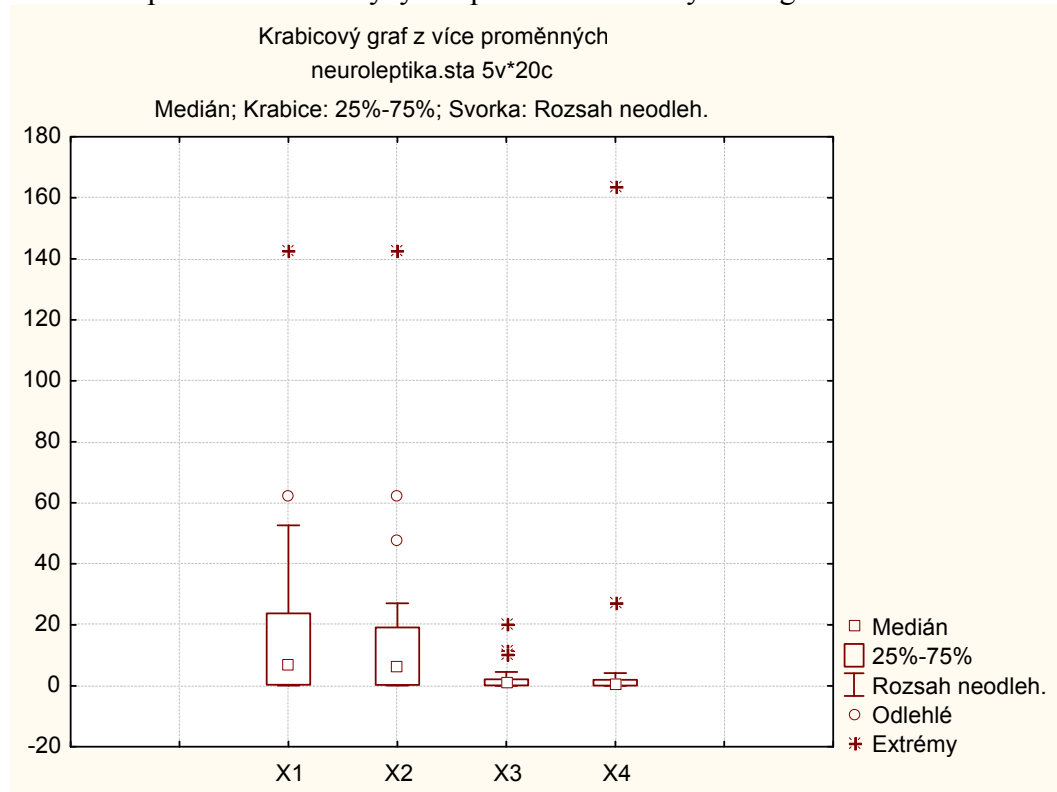
$X_2$  ... potlačení nervozity

$X_3$  ... potlačení stereotypního chování

$X_4$  ... potlačení třesu

$X_5$  ... dávka smrtícího účinku.

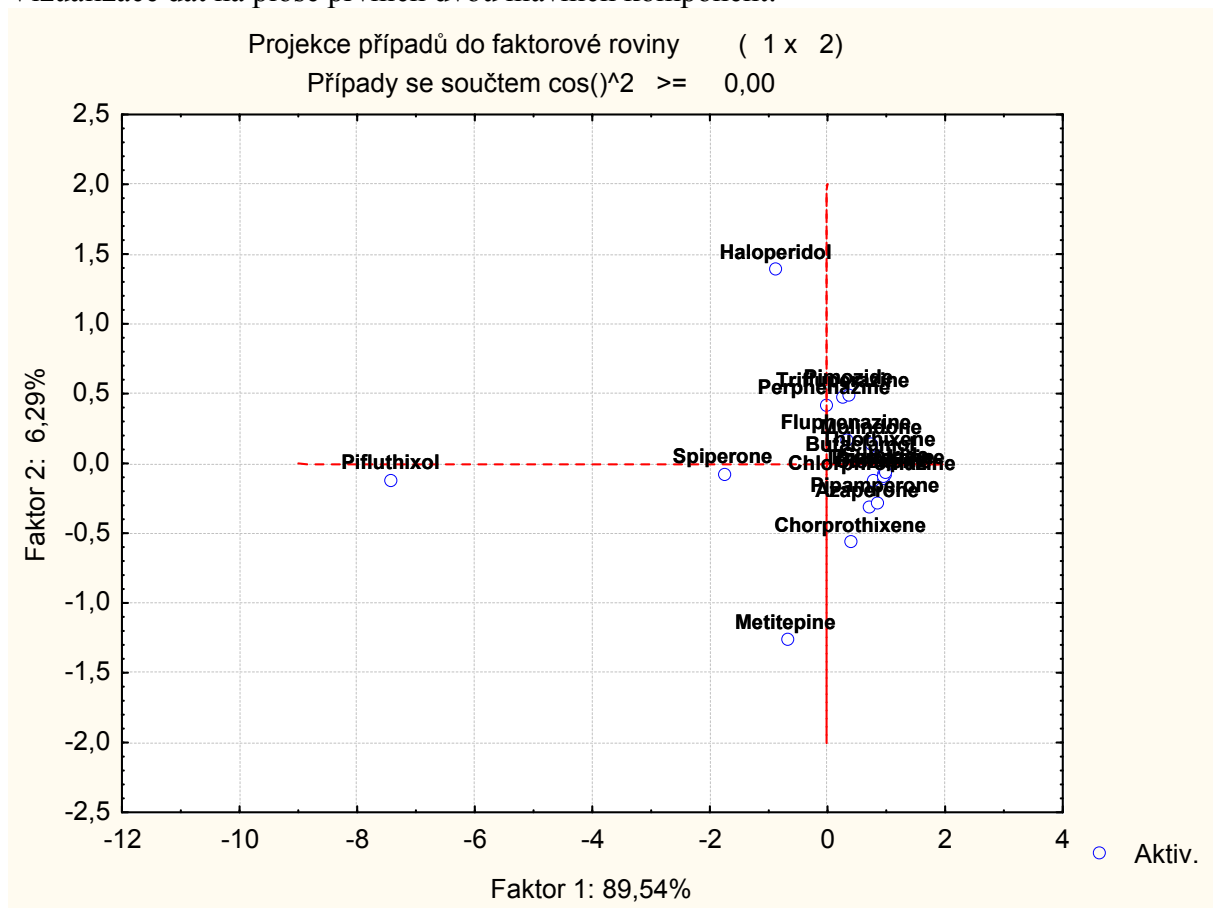
Provedení průzkumové analýzy dat pomocí krabicových diagramů:



Vzhledem k tomu, že proměnné  $X_1$  až  $X_4$  vykazují velmi rozdílnou variabilitu, budeme nadále pracovat se standardizovanými proměnnými.

	1 X1	2 X2	3 X3	4 X4
Chlorpromazine	-0,4813	-0,4552	-0,3628	-0,2326
Promazine	-0,5856	-0,5478	-0,5547	-0,2461
Trifluperazine	0,20504	-0,0244	-0,4678	-0,2813
Fluphenazine	-0,0665	-0,0977	-0,2511	-0,2557
Perphenazine	0,20504	0,24762	-0,2002	-0,2282
Thioridazine	-0,5879	-0,5486	-0,5575	-0,2487
Pifluthixol	3,6343	3,68357	3,32907	4,19263
Thiothixene	-0,4664	-0,4251	-0,5663	-0,2757
Chorprothixene	-0,421	-0,4669	0,2942	-0,1713
Spiperone	1,25525	0,85845	1,67551	-0,2619
Haloperidol	0,9631	1,29988	-0,3301	-0,2696
Azaperone	-0,5081	-0,5161	-0,1502	-0,2023
Pipamperone	-0,5854	-0,5486	-0,2455	-0,2742
Pimozide	0,00907	0,05127	-0,5549	-0,2844
Metitepine	-0,1396	-0,2514	1,37686	0,45313
Clozapine	-0,5904	-0,5513	-0,5128	-0,2763
Perlapine	-0,5856	-0,5445	-0,5046	-0,2833
Sulpiride	-0,5937	-0,5527	-0,5748	-0,2851
Butaclamol	-0,293	-0,2844	-0,2939	-0,2844
Molindone	-0,3674	-0,3259	-0,5486	-0,2849

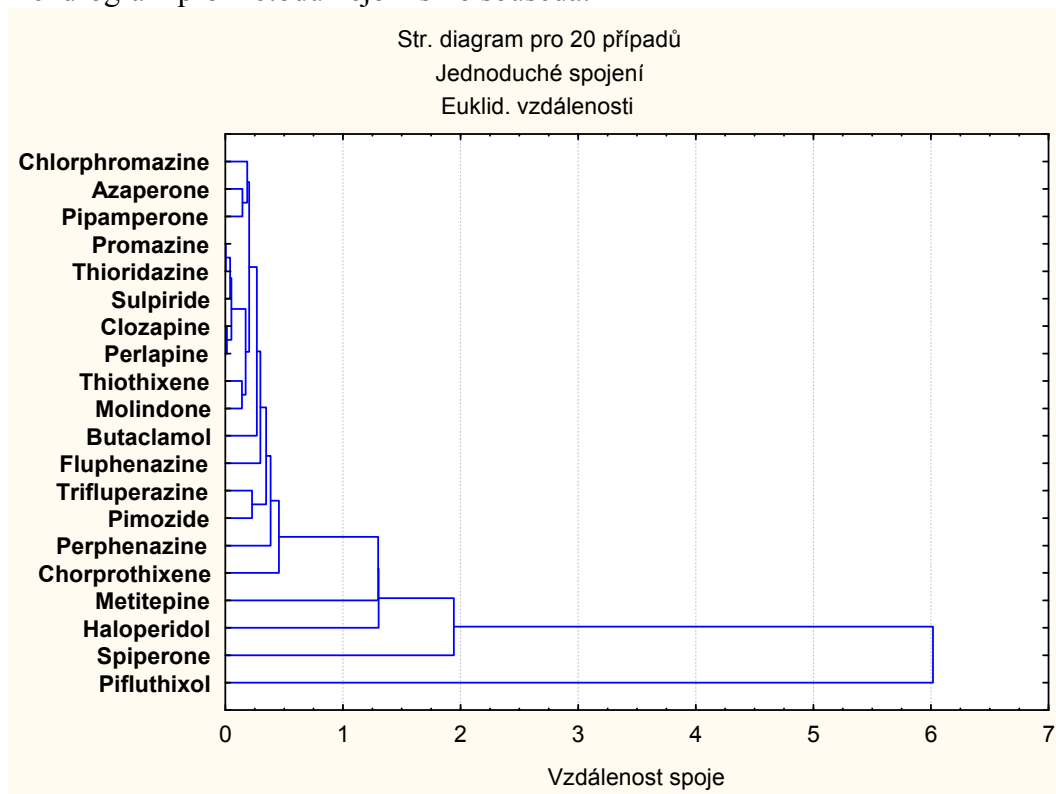
Vizualizace dat na ploše prvních dvou hlavních komponent:



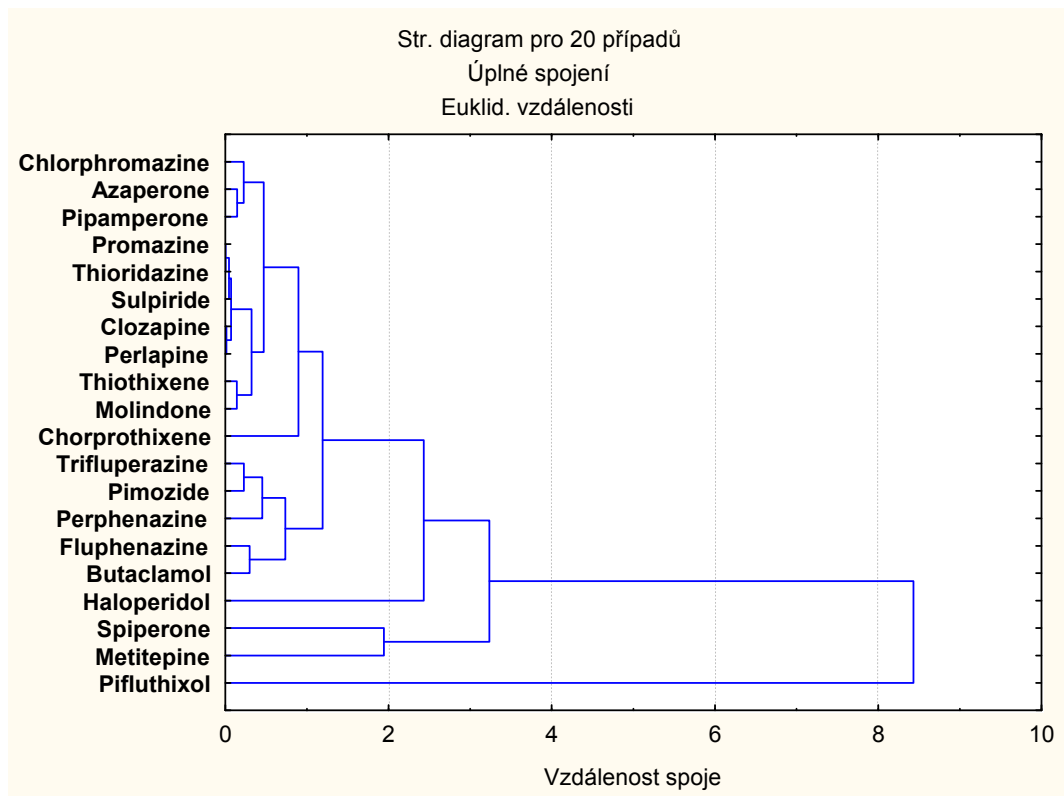
Pro standardizované proměnné  $X_1$  až  $X_4$  provedeme shlukovou analýzu s euklidovskou vzdáleností a čtyřmi metodami: nejbližšího souseda, nejvzdálenějšího souseda, průměrné vazby a Wardovu metodu. Výsledky znázorníme pomocí dendrogramu.

Návod: Statistika – Vícerozměrné průzkumné techniky – Shluková analýza - Spojování (hierarchické shlukování) – OK - Proměnné  $X_1$  až  $X_4$ , OK, Detaily - Shlukovat případy (řádky) – Pravidlo slučování: Jednoduché spojení – Míry vzdálenosti: Euklidovské vzdálenosti - OK – Horizontální graf hierarch. stromu. Euklidovská vzdálenost a metoda nejbližšího souseda je nastavena implicitně. Pro další dvě metody změním Pravidlo slučování z Jednoduchého spojení na Úplné spojení resp. Nevážený průměr skupin dvojic resp. Wardova metoda.

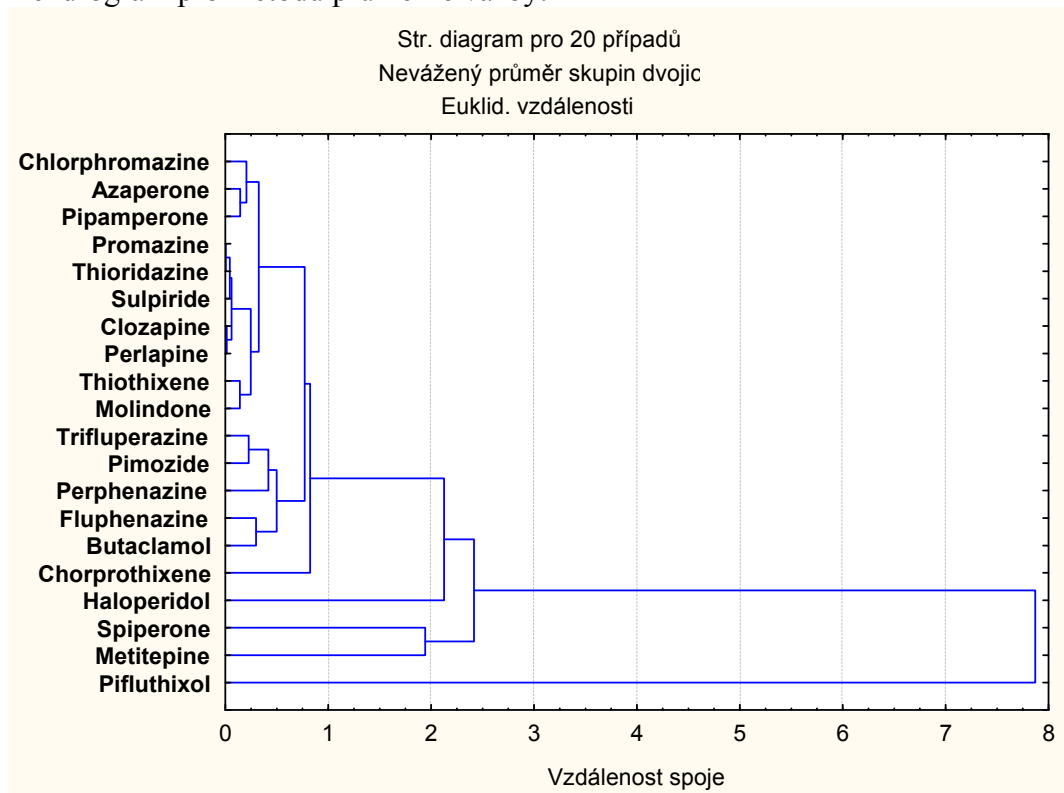
Dendrogram pro metodu nejbližšího souseda:



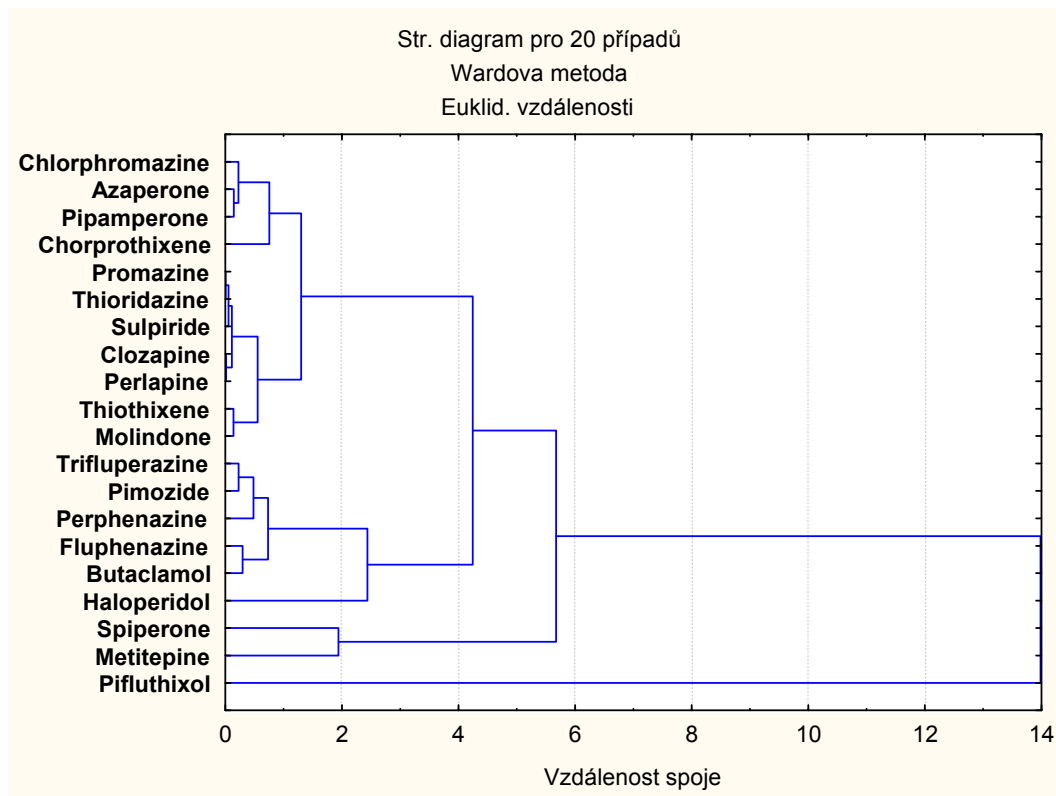
Dendrogram pro metodu nejvzdálenějšího souseda:



Dendrogram pro metodu průměrné vazby:



Dendrogram pro Wardovu metodu:



Ve všech dendrogramech se jako silně vybočující lék jeví Pifluthixol.

Poznámka: Vypočítáme-li pomocí systému MATLAB kofenetické koeficienty korelace, zjistíme, že nejvyšší koeficient poskytne metoda průměrné vazby.

### Provedení metody k-průměrů pro 4 shluky:

Statistiky – Vícerozměrné průzkumné techniky – Shluková analýza – Shlukování metodou k-průměrů – OK – Proměnné  $X_1$  až  $X_4$  – Shlukovat: Případy (řádky), na záložce Details vybereme počet shluků 4 – OK. Na záložce Details vybereme Členy shluků a vzdálenosti. Dostaneme 4 tabulky, které obsahují názvy léků v 1. až 4. shluku a vzdálenosti léků od středu shluku.

1. shluk:

	Členy shluku číslo 1 (neuroleptika.sta) a vzdálenosti od příslušného středu shluku Shluk obsahuje 1 příp.
	Vzdálen.
<b>Pifluthixol</b>	0,00

2. shluk:

	Členy shluku číslo 2 (neuroleptika.sta) a vzdálenosti od příslušného středu shluku Shluk obsahuje 2 příp.
	Vzdálen.
<b>Sziperone</b>	0,485931
<b>Metitepine</b>	0,485931

3. shluk:

	Členy shluku číslo 3 (neuroleptika.sta) a vzdálenosti od příslušného středu shluku Shluk obsahuje 4 příp.
	Vzdálen.
<b>Trifluoperazine</b>	0,224186
<b>Perphenazine</b>	0,139502
<b>Haloperidol</b>	0,549115
<b>Pimozide</b>	0,254224

4. shluk:

	Členy shluku číslo 4 (neuroleptika.sta) a vzdálenosti od příslušného středu shluku Shluk obsahuje 13 příp.
	Vzdálen.
<b>Chlorpromazine</b>	0,013277
<b>Promazine</b>	0,118292
<b>Fluphenazine</b>	0,275511
<b>Thioridazine</b>	0,120093
<b>Thiothixene</b>	0,098881
<b>Chorprothixene</b>	0,336509
<b>Azaperone</b>	0,119668
<b>Pipamperone</b>	0,098279
<b>Clozapine</b>	0,105516
<b>Perlapine</b>	0,100226
<b>Sulpiride</b>	0,129839
<b>Butaclamol</b>	0,129004
<b>Molindone</b>	0,121249

Pro podrobnější průzkum dendrogramu pro metodu průměrné vazby (má nejvyšší kofenetický koeficient korelace) zjistíme, že její výsledky pro 4 shluky jsou poněkud odlišné od výsledků metody k-průměrů.

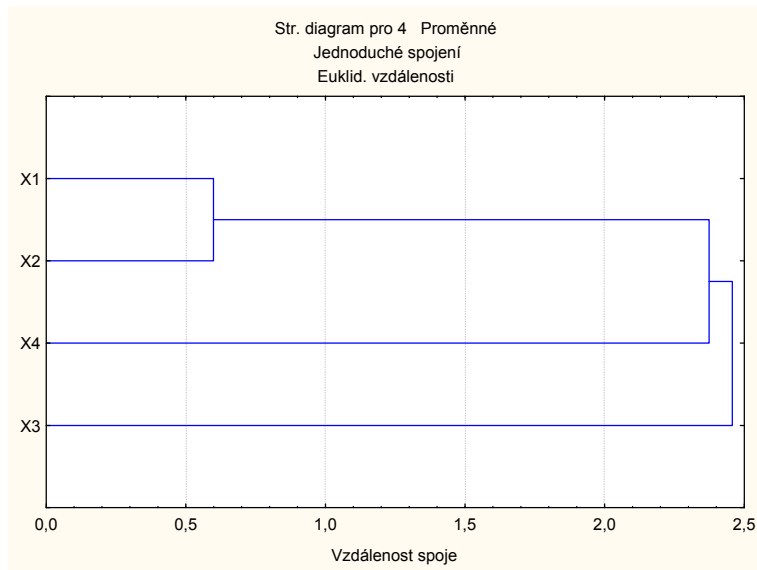
Tabulka ANOVA:

Proměnná	Analýza rozptylu (neuroleptika.sta)					
	Mezisk. SČ	sv	Vnitřní SČ	sv	F	význam. p
X1	17,20068	3	1,799318	16	50,9843	0,000000
X2	17,01859	3	1,981411	16	45,8087	0,000000
X3	18,13750	3	0,862499	16	112,1547	0,000000
X4	18,72720	3	0,272804	16	366,1174	0,000000

Největší vliv na zařazení do shluků má proměnná X<sub>4</sub> vyjadřující dávku smrtícího účinku.

**Pokusíme se ještě o shlukovou analýzu proměnných:**

Dendrogram pro metodu nejbližšího souseda:



Dendrogram pro metodu nejvzdálenějšího souseda, pro metodu průměrné vazby a Wardovu metodu vypadají velmi podobně. Proměnné  $X_1$  (potlačení nervozity) a  $X_2$  (potlačení stereotypního chování) tvoří shluk, k němuž se ve značné vzdálenosti přidává proměnná  $X_4$  (dávka smrtícího účinku). Proměnná  $X_3$  (potlačení třesu) s nimi shluk netvoří.

**Samostatný úkol:** proveďte shlukovou analýzu v systému SPSS.