

## 3.2 Metody s latentními proměnnými a klasifikační metody

### Otázka č. 1

Vypočítejte algoritmem NIPALS 1. latentní proměnnou z matice A[řádek,sloupec]:  
A[1,1]=1, A[2,1]=2, A[3,1]=3, A[1,2]=1, A[2,2]=2, A[3,2]=0, A[1,3]=6, A[2,3]=4,  
A[3,3]=2. Matici před zpracováním standardizujte.

$$\begin{bmatrix} 1 & 1 & 6 \\ 2 & 2 & 4 \\ 3 & 0 & 2 \end{bmatrix}$$

### Standardizace zdrojové matice:

Standardizace matice je postup, při kterém od každého prvku zdrojové matice odečteme aritmetický průměr příslušného sloupce a rozdíl podělíme směrodatnou odchylkou příslušného sloupce.

Standardizace byla provedena pomocí programu OPstat.

Standardizovaná matice:

$$\begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & -1 & -1 \end{bmatrix}$$

První latentní proměnná je vyčíslena na základě vztahu:

$$\mathbf{X} = \mathbf{t}_1 \mathbf{p}_1^T + \mathbf{t}_2 \mathbf{p}_2^T + \dots + \mathbf{t}_A \mathbf{p}_A^T + \mathbf{E}$$

Variabilita všech sloupců je stejná, proto je možné vzít za odhad hlavní komponenty libovolný sloupec, např. první:

$$\mathbf{t}_1^T = [-1 \ 0 \ 1]$$

Po dosazení do vzorce

$$\mathbf{p}_1^T = (\mathbf{t}_1^T \mathbf{t}_1)^{-1} \cdot \mathbf{t}_1^T \mathbf{A}$$

a znormování podle vztahu

$$\mathbf{p}_1^N = (\mathbf{p}_1^T \mathbf{p}_1)^{-\frac{1}{2}} \mathbf{p}_1$$

se získá počáteční odhad vektoru vyjadřujícího míru příspěvků odhadu hlavní komponenty  $t_1$ .

Následujícím dosazením do vztahu:

$$\mathbf{t}_1^T = (\mathbf{p}_1^T \mathbf{p}_p)^{-1} \mathbf{p}_1^T \mathbf{A}^T$$

se získá odhad hlavní komponenty  $t_1$ . Opakováním postupu se získá stabilní rozklad vektorů  $t_1$

a  $p_1$ . Jako konvergenční kritérium se použije vztah:

$$\mathbf{d} = (\mathbf{t}_{\text{nové}} - \mathbf{t}_{\text{staré}})^T (\mathbf{t}_{\text{nové}} - \mathbf{t}_{\text{staré}}) (\mathbf{t}_{\text{nové}}^T \mathbf{t}_{\text{nové}})^{-1}$$

Aby se mohl ukončit výpočet, je potřeba dosáhnou  $\mathbf{d}/N < 10^{-10}$

Stabilní rozklad vektorů:

Vektor první latentní proměnné:

$$\mathbf{p}_1^T = [0,62797 \quad -0,45969 \quad -0,62797]$$

Vektor zátěže:

$$\mathbf{t}_1^T = [0,627963 \quad -0,459701 \quad -0,627963]$$

K výpočtu vektorů byl použit software STATISTICA.

## Otázka č. 2

S použitím vhodných kritérií určete nezbytný počet latentních proměnných, bylo-li z dat určeno: PRESS(0)=S(0)=100, PRESS(1)=20, S(1)=10, PRESS(2)=3.5, S(2)=3.4, PRESS(3)=3.45, S(3)=3.39.

Výpočtem se má stanovit nejmenší signifikantní počet latentních proměnných, které

popisují variabilitu zdrojové matice bez zahrnutí experimentální chyby. Vzhledem k znalosti hodnot PRESS(P) a S(P) můžeme pro výpočet použít test navržený Woldem.

Kritérium: je-li hodnota podílu PRESS(P)/S(P-1) větší jak 0,95, zařazení další latentní proměnné je nevhodné.

$$\text{PRESS}(1)/S(0) = 20/100 = 0,20$$

$$\text{PRESS}(2)/S(1) = 3,5/10 = 0,35$$

$$\text{PRESS}(3)/S(2) = 3,45/3,4 = 1,01$$

Hodnota třetího podílu je větší jak 0,95, proto není čtvrtá latentní proměnná významná .

Nejmenší signifikantní počet latentních proměnných je tedy 3.

### Otázka č. 3

Odhadněte hodnotu chybějícího prvku A[2,2], jestliže výpočtem z nekompletní matice byly určeny vektory p: 0.541 0.423 0.514 0.514 t: -1.340 -0.735 2.076

Na rekonstrukci zdrojové matice X se používají první latentní proměnné popisující podstatnou část variability zdrojové matice bez zahrnutí experimentální chyby. Rekonstrukce predikované zdrojové matice je popsána vztahem:

$$\mathbf{X}^{\text{pred}} = \mathbf{TP}^T$$

Zdrojovou matici lze rovněž rekonstruovat pomocí odpovídající p-latentní proměnné podle vzorce:

$$\mathbf{X}_p^{\text{pred}} = \mathbf{t}_p \mathbf{p}_p^T$$

Protože je zadán vektor latentní proměnné a vektor zátěže, použijeme pro výpočet chybějícího prvku A[2,2] druhý vzorec.

$$\begin{bmatrix} -1.340 \\ -0.735 \\ 2.076 \end{bmatrix} \cdot \begin{bmatrix} 0.541 & 0.423 & 0.514 & 0.514 \end{bmatrix} = \begin{bmatrix} -0,72494 & -0,56682 & -0,68876 & -0,68876 \\ -0,39764 & -0,31091 & -0,37779 & -0,37779 \\ 1,1231 & 0,87815 & 1,0671 & 1,0671 \end{bmatrix}$$

Hodnota chybějícího prvku matice A[2;2] byla vypočtena jako **-0,31091**.

Násobení vektorů bylo provedeno programem OPstat.

### Otázka č. 4

Výpočtem metodou PCA byly určeny vektory p1 : 0.012 0.458 -0.352 0.987 p 2: 0.926 -0.238 0.872 -0.115. Vypočtete komunalitu a vyberte sloupec, který nejlépe charakterizuje celou matici.

Prvek matice zátěží  $\mathbf{p}_{ip}$  příslušejícímu  $i$ -té mu sloupci zdrojové matice je mírou variability tohoto sloupce popsané  $p$ -tou latentní proměnnou. Podíl variability daného sloupce popsané společnými latentními proměnnými lze pak vyjádřit jako součet příspěvků jednotlivých latentních proměnných, tedy

$$h_i^2 = \mathbf{p}^N (\mathbf{p}^N)^T = \sum_{p=1}^P (\mathbf{p}_{ip}^N)^2$$

kde  $h_i^2$  je komunalita pro  $i$ -tý sloupec  
 $P$  ... počet latentních proměnných

$\mathbf{p}_{ip}^N$  jsou zátěže normované podle vztahu: 
$$\mathbf{p}_{ip}^N = \frac{\mathbf{p}_{ip}}{\sum_{i=1}^M \mathbf{p}_{ip}^2}$$

$M$  ... počet sloupců zdrojové matice

#### 4.1 Výpočet normovaných zátěží

Jsou určeny vektory zátěží:  $\mathbf{p1}$ : 0,012 0,458 -0,352 0,987  
 $\mathbf{p2}$ : 0,926 -0,238 0,872 -0,115.

Indexy:  $i$  ... označení sloupce (existují 4 sloupce tj.  $M = 4$ )

$P$  ... počet latentních proměnných (existují 2 latentní proměnné tj.  $P = 2$ )

Podle druhého vzorce vypočítáme normované zátěže pro jednotlivé sloupce při dvou latentních proměnných.

$$\mathbf{p}_{11}^N = 0,0092$$

$$\mathbf{p}_{21}^N = 0,3502$$

$$\mathbf{p}_{31}^N = -0,2691$$

$$\mathbf{p}_{41}^N = 0,7546$$

$$\mathbf{p}_{12}^N = 0,5487$$

$$\mathbf{p}_{22}^N = -0,1410$$

$$\mathbf{p}_{32}^N = 0,5167$$

$$\mathbf{p}_{42}^N = 0,0681$$

## 4.2 Výpočet komunalit

Podle prvního vzorce vypočítáme komunalitu pro všechny 4 sloupce zdrojové matice.

$$h_1^2 = 0,2970$$

$$h_2^2 = 0,1430$$

$$h_4^2 = 0,5741$$

Obecně platí, že čím je komunalita příslušného sloupce větší, tím má sloupec vlastnosti společné s ostatními sloupci zdrojové matice. Z toho vyplývá, vzhledem k hodnotě komunalisty čtvrtého sloupce, že tento sloupec nejlépe charakterizuje zdrojovou matici.

### Otázka č. 5

Vysvětlete, proč vysvětlená variabilita je při výpočtu metodou FA vždy nižší, než při výpočtu metodou PCA.

Vysvětlená variabilita metodou hlavních komponent (PCA) je vždy vyšší než vysvětlená variabilita vypočtená metodou faktorové analýzy (FA). Důvodem této nerovnosti je počet hlavních komponent. Při výpočtu pomocí FA si předem zvolíme počet hlavních komponent. Ty reprodukuje podstatnou, ale nikoliv všechnu variabilitu manifestních proměnných. Jedná se o neúplnou komponentní analýzu. Při výpočtu metodou PCA si počet hlavních komponent sami předem neurčujeme, ale vypočítáme. Vypočítané hlavní komponenty reprodukuje variabilitu přesně. Jedná se o vhodnou ortogonální transformaci, která zachovává beze zbytku původní informaci. V případě PCA hovoříme o úplné komponentní analýze

### Otázka č. 6

Výpočtem metodou kanonických korelací bylo zjištěno:

$$0.297 X_1 + 0.298 X_2 + 0.050 X_3 + 0.256 X_4 = 0.493 Y_1 - 0.213 Y_2$$

$$r_1 = 0.830$$

$$0.006 X_1 - 0.115 X_2 + 0.950 X_3 + 0.056 X_4 = 0.493 Y_1 + 0.213 Y_2$$

$$r_1 = 0.512$$

Vypočítejte skupinový korelační koeficient a interpretujte výsledky.

V metodě kanonické korelace se vypočítá skupinový korelační koeficient podle následujícího vztahu:

$$R_{XY} = \left[ 1 - \frac{|C|}{|C_{XX}| |C_{YY}|} \right]^{\frac{1}{2}}$$

kde: C ... celková kovarianční matice

$C_{XX}$  ... kovarianční matice náhodného vektoru x

$C_{YY}$  ... kovarianční matice náhodného vektoru y

Tento korelační koeficient je složen z kanonických korelačních koeficientů  $r_p$ , pro jejichž vzájemný vztah platí:

$$R_{XY}^2 = 1 - (1 - r_1^2)(1 - r_2^2) \dots (1 - r_p^2)$$

Protože jsou zadány kanonické korelační koeficienty, použijeme pro výpočet druhý vztah.

$$R_{XY}^2 = 1 - (1 - 0,830^2)(1 - 0,512^2) = 0,7704$$

$$R_{XY} = \sqrt{R_{XY}^2} = \sqrt{0,7704} = 0,878$$

Skupinový korelační koeficient má hodnotu 0,7704, což znamená, že 77 % variability dat

jsme vysvětlili kanonickými korelačními koeficienty.

První rovnice vyjadřuje, že růst všech parametrů X má za následek růst obou parametrů Y

(tj. čím budou parametry X vyšší, tím budou vyšší i parametry Y) a obráceně. U parametru

X3 je velmi malá zátěž (0,050). Tento parametr má zanedbatelný vliv na růst či pokles

parametrů Y.

Druhá rovnice udává složitější závislost. Čím budou parametry X1, X3, X4 růst a parametr

X2 klesat, tak bude parametr Y1 růst a parametr Y2 klesat a obráceně, tj. čím budou

parametry X1, X3, X4 klesat a parametr X2 růst, tak bude parametr Y1 klesat a parametr Y2 růst. U parametru X1 je velmi malá zátěž (0,006). Tento parametr má zanedbatelný vliv

na vzrůst a pokles parametrů Y.

### Otázka č. 7

Uveďte nějaký konkrétní příklad vhodný pro zpracování metodou PLS.

Zajištěné vzorky psychotropní látky metamfetaminu (Pervitinu) byly analyzovány metodou infračervené spektrometrie s Fourierovou transformací (FTIR). Naměřená spektra byla vyhodnocena metodou PLS ve stanovených spektrálních oblastech za účelem zjištění podobnosti vzorků.

### Otázka č. 8

Jeden objekt je charakterizován metrickými znaky (2,10), druhý (3,8), třetí (4,9), čtvrtý (10,4) a pátý (11,5). Vypočtěte matici vzdáleností v Euklidově metrice a dokumentujte výpočet shlukování některou z používaných metod. Výsledky interpretujte graficky.

#### Výpočet matice vzdáleností

Blížkost či podobnost objektů se posuzuje na základě míry podobnosti. Jedním z nejjednodušších typů podobností vyjádřené vzdáleností pro metrické proměnné je

Euklidova metrika definovaná vztahem

$$d_E = (X_k, X_l) = \left[ \sum_{p=1}^p (x_{kp} - x_{lp})^2 \right]^{1/2}$$

kde  $d_E$  ... vzdálenost mezi objekty  $X_k$  a  $X_l$

$x_{kp}$  a  $x_{lp}$  ... jsou souřadnice objektů v P-rozměrném prostoru

Dosazením metrických znaků objektů do uvedeného vzorce získáme vzdálenosti mezi jednotlivými objekty:

$$d_E(1,2) = 2,24$$

$$d_E(1,3) = 2,24$$

$$d_E(1,4) = 10,0$$

$$d_E(1,5) = 10,3$$

$$d_E(2,3) = 1,41$$

$$d_E(2,4) = 8,06$$

$$d_E(2,5) = 8,54$$

$$d_E(3,4) = 7,81$$

$$d_E(3,5) = 8,06$$

$$d_E(4,5) = 1,41$$

Získané vzdálenosti uspořádáme do matice:

Matice vzdálenosti znaků má tvar:

|   |       |      |      |      |      |
|---|-------|------|------|------|------|
| 1 | 0,00  |      |      |      |      |
| 2 | 2,24  | 0,00 |      |      |      |
| 3 | 2,24  | 1,41 | 0,00 |      |      |
| 4 | 10,00 | 9,75 | 7,81 | 0,00 |      |
| 5 | 10,30 | 8,54 | 8,06 | 1,41 | 0,00 |

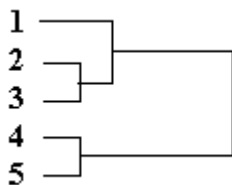
Nejmenší vzdálenost mají prvky 3-2 a 5-4. Z nich je možné vytvořit první shluky, spočítat těžiště nových shluků a opět vypočítat matici vzdáleností:

|     |       |      |      |
|-----|-------|------|------|
| 1   | 0,00  |      |      |
| 2-3 | 2,12  | 0,00 |      |
| 4-5 | 10,12 | 8,06 | 0,00 |

Nejmenší vzdálenost má shluk 2-3 a znak 1. Je tedy možné tyto prvky spojit do shluku a spočítat těžiště nového shluku a opět vypočítat matici vzdáleností:

|       |      |      |
|-------|------|------|
| 1-2-3 | 0,00 |      |
| 4-5   | 8,75 | 0,00 |

Tento proces je možné shrnout do následujícího dendrogramu:





### Otázka č. 9

Popište slovně postup aplikace metod s latentními proměnnými nebo klasifikačních metod na nějakém konkrétním příkladu ze své praxe.

Bylo analyzováno 36 vzorků okvětních částí cannabis a obsahové množství cannabidiolu (CBD), cannabinolu (CBN) a D-9-tetrahydrocannabinolu (THC) bylo vyjádřeno ve formě plochy odezvy detektoru. Pomocí klasifikační metody, shlukové analýzy provedeme klasifikaci vzorků cannabis .

#### Původní naměřená data.

|    | <b>CBD</b> | <b>THC</b> | <b>CBN</b> |
|----|------------|------------|------------|
| 1  | 425        | 6052       | 59         |
| 2  | 129        | 2098       | 14         |
| 3  | 254        | 4860       | 32         |
| 4  | 409        | 1724       | 13         |
| 5  | 547        | 3296       | 15         |
| 6  | 847        | 2760       | 14         |
| 7  | 428        | 3441       | 16         |
| 8  | 1264       | 1439       | 11         |
| 9  | 3376       | 5719       | 552        |
| 10 | 2757       | 4841       | 446        |
| 11 | 2468       | 6094       | 580        |
| 12 | 107        | 3048       | 19         |
| 13 | 28         | 1615       | 11         |
| 14 | 116        | 3047       | 26         |
| 15 | 367        | 889        | 7          |
| 16 | 127        | 1183       | 13         |
| 17 | 105        | 3470       | 42         |
| 18 | 84         | 3125       | 40         |
| 19 | 85         | 3166       | 42         |
| 20 | 108        | 528        | 6          |
| 21 | 67         | 480        | 5          |
| 22 | 129        | 1723       | 25         |
| 23 | 142        | 1947       | 27         |
| 24 | 28         | 306        | 13         |
| 25 | 291        | 954        | 6          |
| 26 | 368        | 906        | 40         |
| 27 | 49         | 1658       | 51         |
| 28 | 593        | 465        | 3          |
| 29 | 107        | 621        | 5          |
| 30 | 145        | 4633       | 24         |
| 31 | 141        | 4409       | 23         |
| 32 | 86         | 1507       | 9          |
| 33 | 82         | 1453       | 9          |
| 34 | 344        | 4803       | 50         |
| 35 | 58         | 1161       | 31         |
| 36 | 147        | 831        | 14         |

### **9.1 Výběr dat.**

Počet objektů převyšuje počet znaků 12x a tudíž lze použít shlukovou analýzu ke zpracování dat.

### **9.2 Průzkumová analýza vícerozměrných dat.**

### **9.3 Úprava dat.**

Data se standardizují, což znamená , že se od základních údajů odečte aritmetický průměr a podělí se směrodatnou odchylkou.

**Tabulka standardizovaných dat**

|    | <b>CBD</b>    | <b>THC</b>   | <b>CBN</b>    |
|----|---------------|--------------|---------------|
| 1  | -0,0534358215 | 2,0460238    | -0,0328006906 |
| 2  | -0,431030061  | -0,236057471 | -0,347221512  |
| 3  | -0,271573034  | 1,3580519    | -0,221453183  |
| 4  | -0,0738463209 | -0,451914424 | -0,354208641  |
| 5  | 0,102194237   | 0,45537737   | -0,340234382  |
| 6  | 0,484891101   | 0,14602088   | -0,347221512  |
| 7  | -0,0496088528 | 0,539065226  | -0,333247253  |
| 8  | 1,01683974    | -0,616404348 | -0,3681829    |
| 9  | 3,71102567    | 1,85383031   | 3,41185408    |
| 10 | 2,92139447    | 1,34708591   | 2,67121837    |
| 11 | 2,55272983    | 2,07026442   | 3,6074937     |
| 12 | -0,459094498  | 0,312242278  | -0,312285865  |
| 13 | -0,559871339  | -0,514824606 | -0,3681829    |
| 14 | -0,447613592  | 0,31166512   | -0,263375959  |
| 15 | -0,127423882  | -0,933841045 | -0,396131417  |
| 16 | -0,433581374  | -0,764156702 | -0,354208641  |
| 17 | -0,46164581   | 0,555802797  | -0,15158189   |
| 18 | -0,488434591  | 0,356683415  | -0,165556148  |
| 19 | -0,487158935  | 0,380346878  | -0,15158189   |
| 20 | -0,457818842  | -1,14219495  | -0,403118546  |
| 21 | -0,510120746  | -1,16989852  | -0,410105676  |
| 22 | -0,431030061  | -0,452491582 | -0,270363089  |
| 23 | -0,41444653   | -0,323208273 | -0,25638883   |
| 24 | -0,559871339  | -1,27032394  | -0,354208641  |
| 25 | -0,224373754  | -0,896325799 | -0,403118546  |
| 26 | -0,126148226  | -0,924029365 | -0,165556148  |
| 27 | -0,533082558  | -0,490006828 | -0,0886977254 |
| 28 | 0,160874423   | -1,17855588  | -0,424079934  |
| 29 | -0,459094498  | -1,08851929  | -0,410105676  |
| 30 | -0,410619562  | 1,22703712   | -0,277350218  |
| 31 | -0,415722187  | 1,09775381   | -0,284337347  |
| 32 | -0,485883278  | -0,57715763  | -0,382157158  |
| 33 | -0,490985903  | -0,608324142 | -0,382157158  |
| 34 | -0,156763975  | 1,32515392   | -0,0956848548 |
| 35 | -0,521601652  | -0,776854169 | -0,228440312  |
| 36 | -0,408068249  | -0,967316187 | -0,347221512  |

#### 9.4 Výběr příslušné míry podobnosti objektů a výpočet matice podobnosti.

Blížkost či podobnost objektů se posuzuje na základě míry podobnosti. Jedním z nejjednodušších typů podobností vyjádřené vzdáleností pro metrické proměnné je

Euklidova metrika definovaná vztahem:

$$d_E(X_k, X_l) = \left[ \sum_{p=1}^p (x_{kp} - x_{lp})^2 \right]^{\frac{1}{2}}$$

kde  $d_E$  je vzdálenost mezi objekty  $X_k$  a  $X_l$

$x_{kp}$  a  $x_{lp}$  jsou souřadnice objektů v  $P$ -rozměrném prostoru

## Výřez tabulky

Euklid. vzdálenosti (THC\_1) – část matice dat

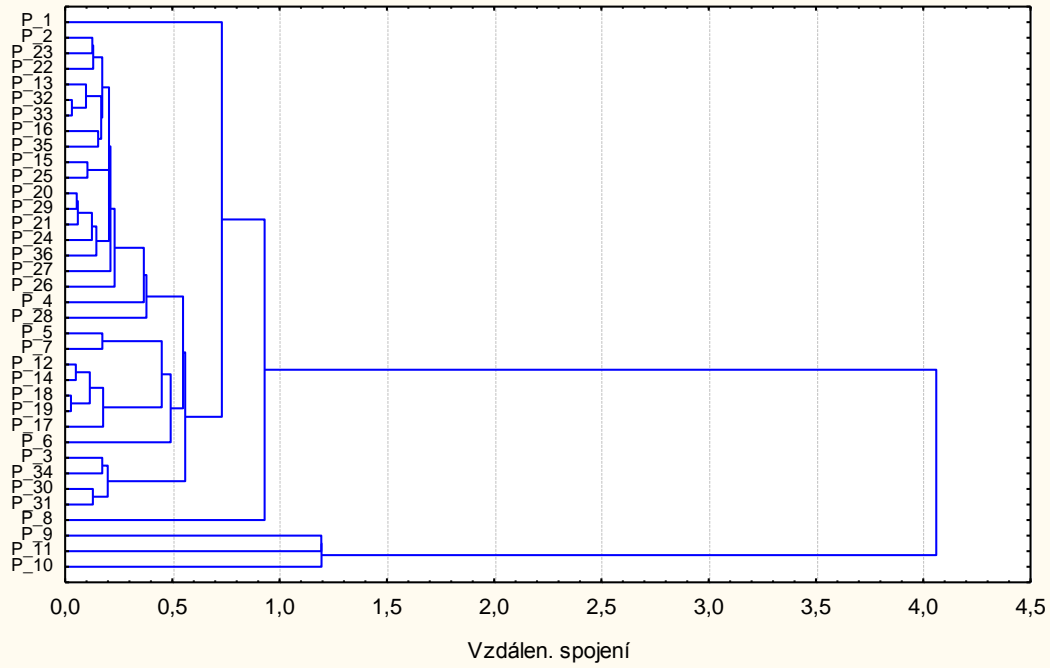
|      | P_2  | P_3  | P_4  | P_5  | P_6  | P_7  | P_8  | P_9  | P_10 | P_11 | P_12 | P_13 | P_14 | P_15 | P_16 | P_17 | P_18 | P_19 |
|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| P_1  | 2,33 | 0,75 | 2,52 | 1,63 | 2,00 | 1,54 | 2,89 | 5,11 | 4,08 | 4,48 | 1,80 | 2,63 | 1,79 | 3,00 | 2,85 | 1,55 | 1,75 | 1,73 |
| P_2  | 0,00 | 1,61 | 0,42 | 0,87 | 0,99 | 0,86 | 1,50 | 5,97 | 4,78 | 5,46 | 0,55 | 0,31 | 0,55 | 0,76 | 0,53 | 0,82 | 0,62 | 0,65 |
| P_3  | 1,61 | 0,00 | 1,83 | 0,98 | 1,43 | 0,86 | 2,36 | 5,41 | 4,31 | 4,81 | 1,07 | 1,90 | 1,06 | 2,30 | 2,13 | 0,83 | 1,03 | 1,00 |
| P_4  | 0,42 | 1,83 | 0,00 | 0,92 | 0,82 | 0,99 | 1,10 | 5,82 | 4,62 | 5,38 | 0,86 | 0,49 | 0,85 | 0,49 | 0,48 | 1,10 | 0,93 | 0,95 |
| P_5  | 0,87 | 0,98 | 0,92 | 0,00 | 0,49 | 0,17 | 1,41 | 5,39 | 4,22 | 4,92 | 0,58 | 1,17 | 0,57 | 1,41 | 1,33 | 0,60 | 0,62 | 0,62 |
| P_6  | 0,99 | 1,43 | 0,82 | 0,49 | 0,00 | 0,66 | 0,93 | 5,24 | 4,06 | 4,86 | 0,96 | 1,24 | 0,95 | 1,24 | 1,29 | 1,05 | 1,01 | 1,02 |
| P_7  | 0,86 | 0,86 | 0,99 | 0,17 | 0,66 | 0,00 | 1,57 | 5,47 | 4,30 | 4,96 | 0,47 | 1,17 | 0,46 | 1,48 | 1,36 | 0,45 | 0,50 | 0,50 |
| P_8  | 1,50 | 2,36 | 1,10 | 1,41 | 0,93 | 1,57 | 0,00 | 5,26 | 4,09 | 5,04 | 1,74 | 1,58 | 1,74 | 1,19 | 1,46 | 1,90 | 1,80 | 1,82 |
| P_9  | 5,97 | 5,41 | 5,82 | 5,39 | 5,24 | 5,47 | 5,26 | 0,00 | 1,20 | 1,19 | 5,80 | 6,18 | 5,76 | 6,08 | 6,18 | 5,64 | 5,72 | 5,70 |
| P_10 | 4,78 | 4,31 | 4,62 | 4,22 | 4,06 | 4,30 | 4,09 | 1,20 | 0,00 | 1,24 | 4,63 | 4,98 | 4,59 | 4,89 | 4,99 | 4,48 | 4,54 | 4,53 |
| P_11 | 5,46 | 4,81 | 5,38 | 4,92 | 4,86 | 4,96 | 5,04 | 1,19 | 1,24 | 0,00 | 5,25 | 5,67 | 5,20 | 5,68 | 5,71 | 5,05 | 5,14 | 5,12 |
| P_12 | 0,55 | 1,07 | 0,86 | 0,58 | 0,96 | 0,47 | 1,74 | 5,80 | 4,63 | 5,25 | 0,00 | 0,84 | 0,05 | 1,29 | 1,08 | 0,29 | 0,16 | 0,18 |
| P_13 | 0,31 | 1,90 | 0,49 | 1,17 | 1,24 | 1,17 | 1,58 | 6,18 | 4,98 | 5,67 | 0,84 | 0,00 | 0,84 | 0,60 | 0,28 | 1,10 | 0,90 | 0,92 |
| P_14 | 0,55 | 1,06 | 0,85 | 0,57 | 0,95 | 0,46 | 1,74 | 5,76 | 4,59 | 5,20 | 0,05 | 0,84 | 0,00 | 1,29 | 1,08 | 0,27 | 0,12 | 0,14 |
| P_15 | 0,76 | 2,30 | 0,49 | 1,41 | 1,24 | 1,48 | 1,19 | 6,08 | 4,89 | 5,68 | 1,29 | 0,60 | 1,29 | 0,00 | 0,35 | 1,55 | 1,36 | 1,38 |
| P_16 | 0,53 | 2,13 | 0,48 | 1,33 | 1,29 | 1,36 | 1,46 | 6,18 | 4,99 | 5,71 | 1,08 | 0,28 | 1,08 | 0,35 | 0,00 | 1,34 | 1,14 | 1,16 |
| P_17 | 0,82 | 0,83 | 1,10 | 0,60 | 1,05 | 0,45 | 1,90 | 5,64 | 4,48 | 5,05 | 0,29 | 1,10 | 0,27 | 1,55 | 1,34 | 0,00 | 0,20 | 0,18 |
| P_18 | 0,62 | 1,03 | 0,93 | 0,62 | 1,01 | 0,50 | 1,80 | 5,72 | 4,54 | 5,14 | 0,16 | 0,90 | 0,12 | 1,36 | 1,14 | 0,20 | 0,00 | 0,03 |
| P_19 | 0,65 | 1,00 | 0,95 | 0,62 | 1,02 | 0,50 | 1,82 | 5,70 | 4,53 | 5,12 | 0,18 | 0,92 | 0,14 | 1,38 | 1,16 | 0,18 | 0,03 | 0,00 |
| P_20 | 0,91 | 2,51 | 0,79 | 1,69 | 1,60 | 1,73 | 1,57 | 6,40 | 5,20 | 5,96 | 1,46 | 0,64 | 1,46 | 0,39 | 0,38 | 1,72 | 1,52 | 1,54 |
| P_21 | 0,94 | 2,55 | 0,84 | 1,74 | 1,65 | 1,77 | 1,62 | 6,45 | 5,25 | 6,00 | 1,49 | 0,66 | 1,49 | 0,45 | 0,42 | 1,75 | 1,55 | 1,57 |
| P_22 | 0,23 | 1,82 | 0,37 | 1,06 | 1,10 | 1,06 | 1,46 | 6,00 | 4,81 | 5,50 | 0,77 | 0,17 | 0,76 | 0,58 | 0,32 | 1,02 | 0,82 | 0,84 |
| P_23 | 0,13 | 1,69 | 0,38 | 0,94 | 1,02 | 0,94 | 1,47 | 5,93 | 4,74 | 5,43 | 0,64 | 0,27 | 0,64 | 0,69 | 0,45 | 0,89 | 0,69 | 0,72 |
| P_24 | 1,04 | 2,65 | 0,95 | 1,85 | 1,76 | 1,88 | 1,71 | 6,49 | 5,30 | 6,05 | 1,59 | 0,76 | 1,59 | 0,55 | 0,52 | 1,84 | 1,64 | 1,66 |
| P_25 | 0,69 | 2,26 | 0,47 | 1,39 | 1,26 | 1,45 | 1,27 | 6,13 | 4,94 | 5,71 | 1,23 | 0,51 | 1,24 | 0,10 | 0,25 | 1,49 | 1,30 | 1,33 |
| P_26 | 0,77 | 2,29 | 0,51 | 1,41 | 1,25 | 1,47 | 1,20 | 5,94 | 4,74 | 5,51 | 1,29 | 0,63 | 1,28 | 0,23 | 0,39 | 1,52 | 1,33 | 1,35 |
| P_27 | 0,38 | 1,87 | 0,53 | 1,17 | 1,23 | 1,16 | 1,58 | 5,98 | 4,79 | 5,45 | 0,84 | 0,28 | 0,82 | 0,68 | 0,39 | 1,05 | 0,85 | 0,87 |
| P_28 | 1,12 | 2,58 | 0,77 | 1,64 | 1,37 | 1,73 | 1,03 | 6,04 | 4,86 | 5,70 | 1,62 | 0,98 | 1,62 | 0,38 | 0,73 | 1,86 | 1,69 | 1,71 |
| P_29 | 0,86 | 2,46 | 0,75 | 1,64 | 1,56 | 1,68 | 1,55 | 6,38 | 5,18 | 5,93 | 1,40 | 0,58 | 1,41 | 0,37 | 0,33 | 1,66 | 1,47 | 1,49 |
| P_30 | 1,46 | 0,20 | 1,71 | 0,93 | 1,41 | 0,78 | 2,33 | 5,57 | 4,45 | 4,96 | 0,92 | 1,75 | 0,92 | 2,18 | 1,99 | 0,68 | 0,88 | 0,86 |

### 9.5 Uvedení výsledků

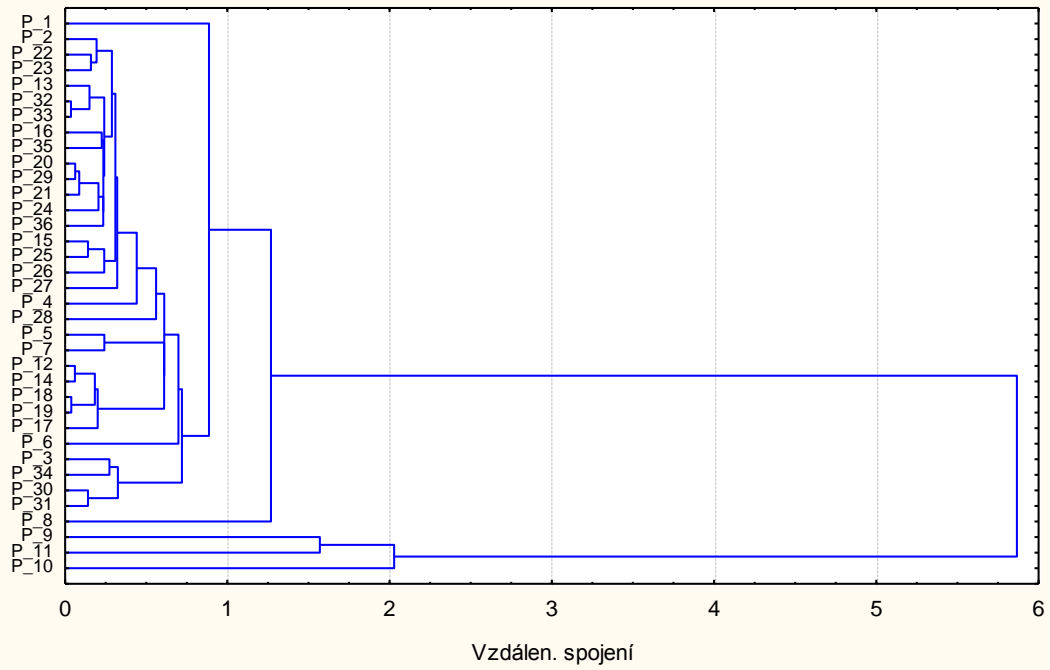
Postup shlukování zaneseme do dendrogramu. Na vodorovné ose je vzdálenost, na které

došlo ke sloučení objektů nebo shluků a na svislé ose jsou uspořádané objekty.

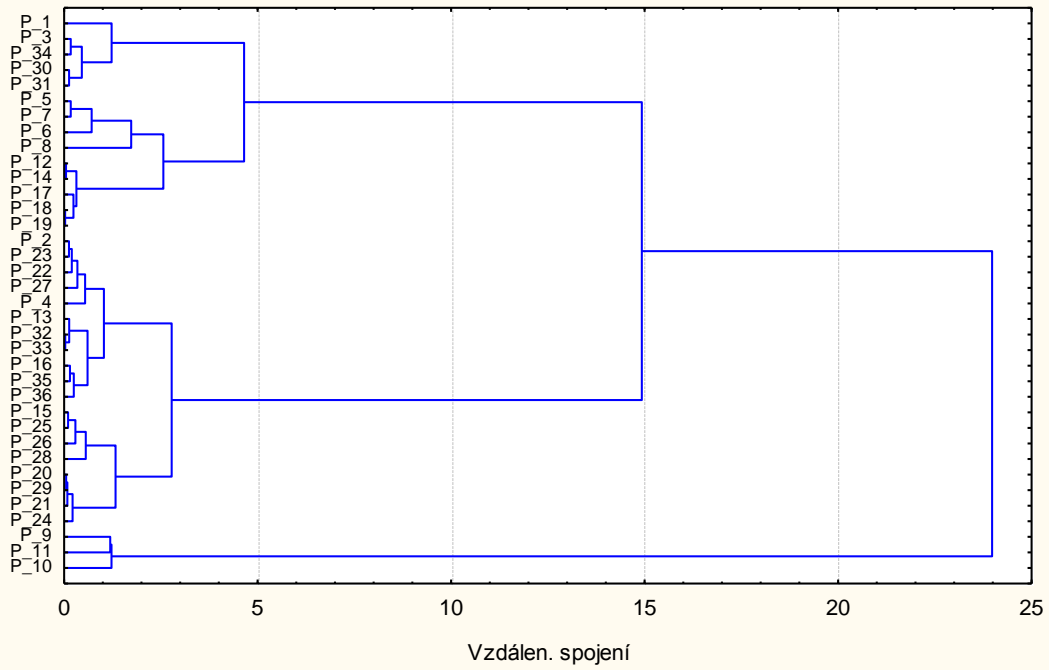
Str. diagram pro 36 případů  
Jednoduché spojení  
Euklid. vzdálenosti



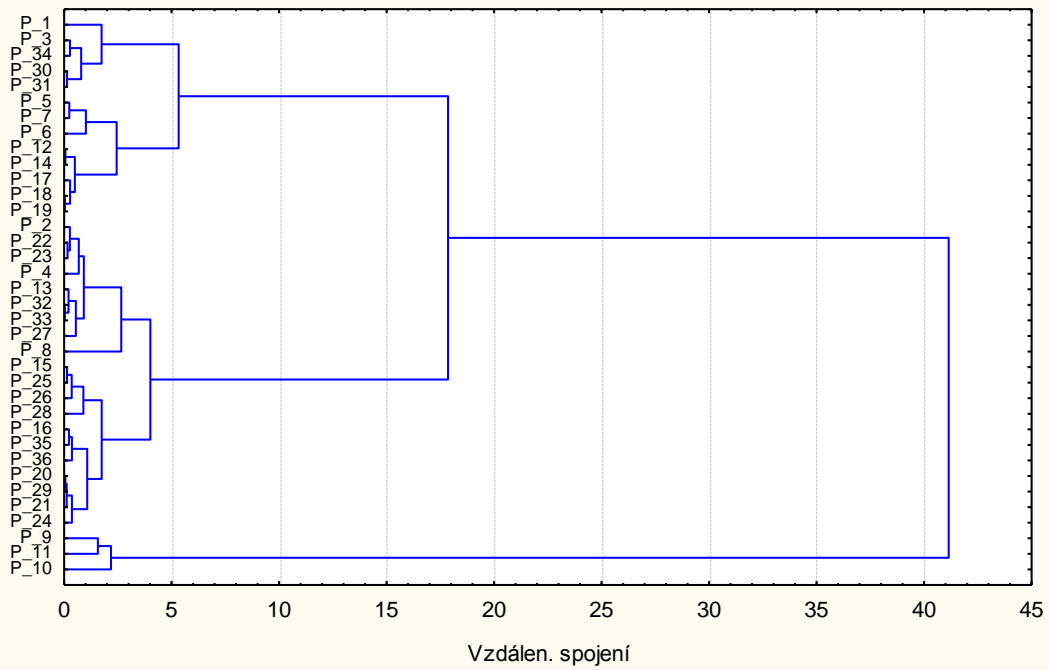
Str. diagram pro 36 případů  
Jednoduché spojení  
Blokové vzdálenosti (Manhattan)



Str. diagram pro 36 případů  
Wardova metoda  
Euklid. vzdálenosti



Str. diagram pro 36 případů  
Wardova metoda  
Blokové vzdálenosti (Manhattan)



Za využití programu STATISTICA 8 bylo celkem 36 objektů seskupeno do jednoho shluku.

Za účelem shlukování bylo použito několik postupů a lze konstatovat, že byly získány obdobné výsledky. Vyhodnocením výsledků se jeví optimální čtyři shluky:

1 shluk: 2-22-23-13-32-33-16-35

2 shluk: 20-29-21-24-36-22-26

3 shluk: 12-14-18-19-17

4 shluk: 3-34-30-31