

Delta(1.0):	0.160527
--------------------	----------

Průběh spojování do shluků je zobrazen v opačném pořadí. Uvádí se také hladina vzdáleností, na které dochází ke shlukování. To umožňuje vybrat si vhodnou hladinu, na které provedeme výseč shluků. Když vezmeme hladinu vzdáleností rovnou 1, bude výseč u spojení 7 a 8 a povede to na 5 shluků. Podíváme-li se podrobněji, je zřejmé, že jde o 3 shluky a 2 odlehlé objekty. Oba odlehlé objekty jsou přesto nazvány shluky, i když každý shluk obsahuje pouze 1 objekt. **Spoj:** pořadové číslo shluku. Počet shluků však obsahuje také odlehlé hodnoty. **Velikost vzdálenosti:** vzdálenost mezi dvěma spojovanými shluky. Za normálních okolností se tato hodnota monotónně zvyšuje. Slouží k určování vhodného počtu shluků. **Diagram vzdálenosti:** čárový diagram hodnot vzdálenosti. Vybereme takový počet shluků, který vykazuje náhlý skok v diagramu. **Propojené řádky:** jde o čísla objektů čili řádků, propojených v tento aktuální shluk. Všimněte si, že propojení jsou zobrazena v opačném pořadí: v našem příkladě se objekt 7 a 10 spojují jako první, k nim se pak přidává objekt 4 atd. **Kofenetická korelace:** kofenetický korelační koeficient CC mezi skutečnou vzdáleností a dendrogramem predikovanou vzdáleností je založen na dotyčné hierarchické konfiguraci. Hodnoty $CC = 0.75$ a vyšší znamenají, že shlukování je možné považovat za užitečné. **Delta(0.5, 1):** obě kritéria jsou mírami těsnosti proložení. Užijeme je jako rozhodčí kritéria k hledání a rozlišení hledání a mezi technikami shlukování. Čím více je delta blíží nule, tím je shlukovací technika lepší a věrohodnější.

4. Tabulka vzdáleností skutečných a predikovaných v dendrogramu:

První řádek	Druhý řádek	Skutečná vzdálenost	Vzdálenost ve shluku	Rozdíl vzdáleností	Procento rozdílu
1	2	1.202793	1.210352	-0.007559	-0.63
1	3	1.227878	1.210352	0.017526	1.43
1	4	0.910046	1.210352	-0.300306	-33.00
1	5	1.596586	2.054454	-0.457869	-28.68
..
11	12	1.761248	1.609330	0.151919	8.63

Ukazuje skutečné a v dendrogramu predikované vzdálenosti pro každý pár objektů. Obsahuje také jejich rozdíl a rozdíl v procentech. Obvykle se tento oddíl zkracuje, či vůbec vynechává, protože i pro malý počet objektů jeho délka velice narůstá.

5. Hledání nejlepší metody shlukování:

V tabulce je uvedeno postupně 7 metod shlukování a za rozhodčí kritéria jsou použita kofenetický korelační koeficient a obě kritéria delta. Nejlepší technika tvorby dendrogramu vykazuje nejvyšší hodnotu kofenetického korelačního koeficientu a nejnižší hodnotu blízkou nule u kritérií delta. V dané úloze se ukázala jako nejlepší technika Průměrové metody (Group Average, Unweighted Pair-Group).

1. Metoda shlukování: Skupinový průměr, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.891119, Delta(0.5): 0.122037, Delta(1.0): 0.160527,
2. Metoda shlukování: Jednoduchý průměr, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.883564, Delta(0.5): 0.133286, Delta(1.0): 0.173097,
3. Metoda shlukování: Těžiště, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.858393, Delta(0.5): 0.369357, Delta(1.0): 0.454509,
4. Metoda shlukování: Nejbližšího souseda, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.848369, Delta(0.5): 0.294508, Delta(1.0): 0.379311,
5. Metoda shlukování: Nejvzdálenějšího souseda, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.762902, Delta(0.5): 0.283987, Delta(1.0): 0.361304,
6. Metoda shlukování: Median, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.847619, Delta(0.5): 0.297623, Delta(1.0): 0.354229,
7. Metoda shlukování: Wardova metoda, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.687581, Delta(0.5): 0.536833, Delta(1.0): 0.622271,

6. Grafické zobrazení dendrogramu podobnosti objektů:

Dendrogram podobnosti objektů názorně ukazuje tvorbu a rozlišení shluků a představuje vlastně rozhodující výsledek shlukové analýzy vícerozměrných dat.

Player	Player
Chamberlain W	Chamberlain W
Cousy B	Russell B
Russell B	Cousy B
Johnson M	Johnson M
West J	West J
Robertson O	Robertson O
Jordan M	Jordan M
Baylor E	Baylor E
Erving J	Erving J
Bird L	Bird L
Barry R	Barry R
Jabbar K.A.	Jabbar K.A.

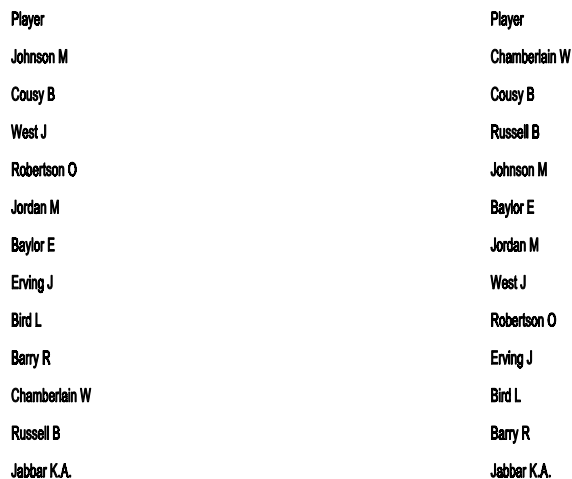
Obr. 4.18a Metoda váženého průměru (Simple Average).

Obr. 4.18b Metoda nevážených průměrů (Group Average).

Player	Player
Chamberlain W	Chamberlain W
Russell B	Russell B
Cousy B	Cousy B
Johnson M	Johnson M
Baylor E	Jordan M
Jordan M	Baylor E
West J	West J
Robertson O	Robertson O
Erving J	Erving J
Bird L	Bird L
Barry R	Barry R
Jabbar K.A.	Jabbar K.A.

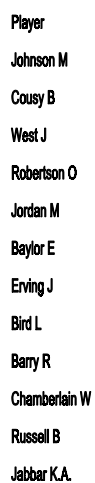
Obr. 4.18c Metoda těžiště (Centroid).

Obr. 4.18d Metoda nejbližšího spojení (Single Linkage).



Obr. 4.18e Metoda nejvdálénéjšiho spojení (Complete Link).

Obr. 4.18f Metoda mediánová (Median).



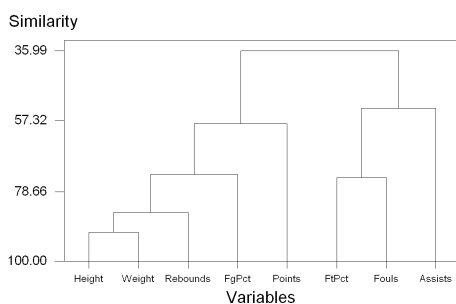
Obr. 4.18g Metoda Wardova (Ward's Minimum Variance).

Interpretace dendrogramu je snadná: objekty blízko sebe jsou propojeny spojovací úsečkou hodně vpravo, mají malou vzdálenost, čili značnou podobnost. Objekty propojené hodně vlevo mají malou podobnost a mezi sebou vykazují velkou vzdálenost, např. hráč Bob Cousy se velice liší od všech ostatních hráčů. Míra podobnosti nebo naopak míra vzdálenosti dvou objektů se může přečíst přímo na ose. Počet vhodných shluků může být snadno určen zakreslením vrislice do diagramu. Vztyčíme-li, například, kolmici v bodě 1.0 na ose vzdáleností x , dostaneme 5 shluků. Jeden shluk obsahuje 2 objekty, jeden shluk obsahuje 7 objektů a tři shluky obsahují každý pouze po 1

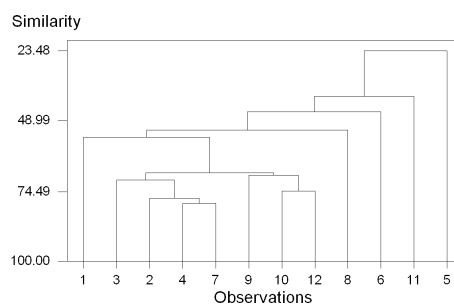
objektu. Nejvhodnější techniku shlukování vybereme na základě dvou rozhodčích kritérií, kofenetické korelace CC a kritéria delta. Na sedmi dendrogramech jsou uvedeny shluky, vytvořené sedmi rozličnými technikami. Protože rozhodčí kritéria našla jako nejlepší techniku průměrovou, budeme tento dendrogram považovat za optimální a v naší analýze za výsledný. Uživatel může porovnat jemně rozdíly mezi dendrogramy a posoudit jejich věrohodnost dle velikosti kofenetické korelace a kritéria delta.

7. Grafické zobrazení dendrogramu podobnosti proměnných:

Dendrogram podobnosti proměnných názorně ukazuje rozlišení proměnných ve shlucích. Jeho interpretace je snadná: proměnné blízko sebe jsou propojeny spojovací úsečkou hodně nízko, mají malou vzdálenost čili značnou vzájemnou podobnost. Proměnné propojené hodně vysoko mají malou podobnost a mezi sebou vykazují velkou vzdálenost. Stejná interpretace podobnosti je i u dendrogramu objektů, kde jsou tentokrát objekty označeny nikoliv svými jmény, ale indexy.



Obr. 4.19a Dendrogram proměnných metodou průměrovou, MINITAB.



Obr. 4.19b Dendrogram objektů (hráčů košíkové) metodou průměrovou, MINITAB.

Vzorová úloha 4.11 Vytvoření dendrogramu objektů neuroleptika

Vytvoření dendrogramu neuroleptik u **Úlohy B4.02 Účinky neuroleptik při tlumení rozličných psychóz**. Neuroleptika redukují nežádoucí účinky přebytečného dopaminu. Liší se však ve svých účincích: potlačují nervozitu, záchvaty, třes, ospalost, parkinsonismus, vynechávání menstruace, vyrážky, zvýšené slinění atd. Účelem je provést klasifikaci neuroleptik do shluků podobných účinků s ohledem na čtyři proměnné.

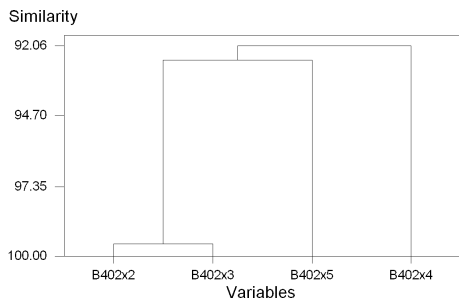
Řešení:

U úlohy uvedeme pro stručnost pouze zkrácený výstup programu NCSS2000. Po výběru optimální metody tvorby dendrogramu uvedeme dendrogram podobnosti proměnných a dendrogram podobnosti objektů. Nejvyšší hodnota kofenetického korelačního koeficientu a nejnižší hodnota kritéria delta svědčí o optimální shlukovací metodě tvorby dendrogramu metodou průměrovou

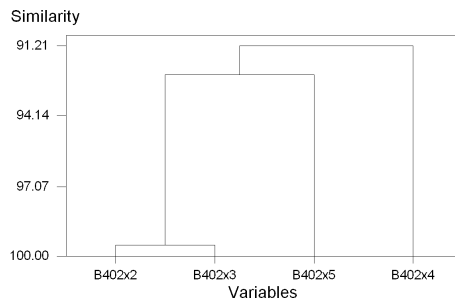
1. Metoda shlukování: Skupinový průměr, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.987356, Delta(0.5): 0.137455, Delta(1.0): 0.125290;
2. Metoda shlukování: Jednoduchý průměr, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.988876, Delta(0.5): 0.177810, Delta(1.0): 0.188781;
3. Metoda shlukování: Těžiště, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.984750, Delta(0.5): 0.175238, Delta(1.0): 0.166599;
4. Metoda shlukování: Nejbližšího souseda, Typ vzdálenosti: Euclid., směrodatná odchylka, Kofenetická korelace: 0.988598, Delta(0.5): 0.474238, Delta(1.0): 0.391993;

5. Metoda shlukování: Median, Typ vzdálenosti: Eucleid., směrodatná odchylka, Kofenetická korelace: 0.984215, Delta(0.5): 0.452308, Delta(1.0): 0.428346;

6. Metoda shlukování: Wardova metoda, Typ vzdálenosti: Eucleid., směrodatná odchylka, Kofenetická korelace: 0.979285, Delta(0.5): 0.549394, Delta(1.0): 0.492716.

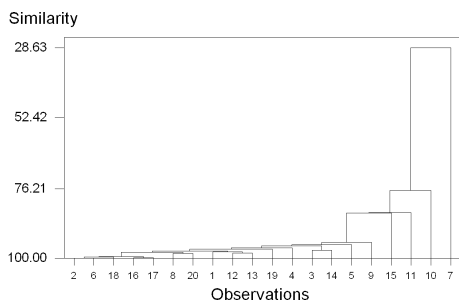


Obr. 4.20a Metoda nejbližšího spojení v dendrogramu proměnných, **MINITAB**.

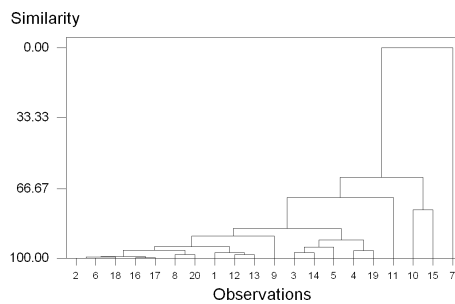


Obr. 4.20b Metoda průměrné v dendrogramu proměnných, **MINITAB**.

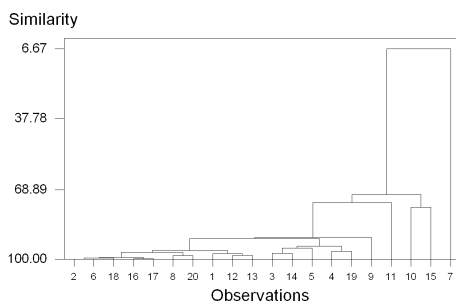
Metoda průměru v dendrogramu podobnosti objektů ukazuje na 6 shluků: první shluk obsahuje 10 objektů 18, 16, 17, 8, 20, 1, 12, 13, druhý shluk 5 objektů 3, 14, 5, 4, 19, třetí shluk 2 objekty 10 a 15 a zbývající tři shluky obsahují vždy po jednom objektu, a to 9, 11 a 7. Jako silně vybočující objekt se jeví objekt 7. Dendrogramy získané ostatními technikami dospěly převážně ke stejným shlukům a podobné struktuře.



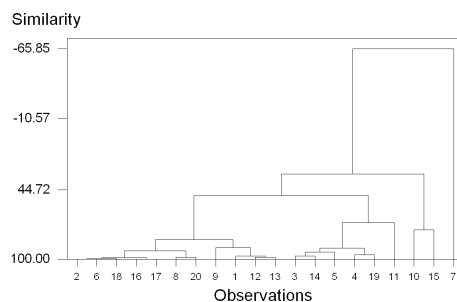
Obr. 4.21a Dendrogram objektů (nejbližšího spojení) metoda nejbližšího spojení, $CC = 0.9886$, $\Delta(0.5) = 0.474$, $\Delta(1.0) = 0.392$.



Obr. 4.21b Dendrogram objektů metoda nejvzdálenějšího spojení, $CC = 0.9828$, $\Delta(0.5) = 0.179$, $\Delta(1.0) = 0.183$.



Obr. 4.21c Dendrogram objektů (neř roleptik) metodou průměrnou, $CC = 0.9889$, $(0.5) = 0.178$, $(1.0) = 0.189$.



Obr. 4.21d Dendrogram objektů (neř roleptik) Wardovou metodou, $CC = 0.979$, $(0.5) = 0.549$, $(1.0) = 0.493$.

Vzorová úloha 4.12 Klasifikace objektů do shluků

Vstupní data jsou z úlohy S4.16 Shluky 12 superhvězd košíkové. Tabulka dat obsahuje informace o osmi hráčských vlastnostech a aktivitách 12 superhvězd košíkové v sezóně 1989. Cílem je najít shluky hráčů podobných vlastností a naopak odhalit jejich aktivity a vlastnost, ve které se hráč neshoduje s ostatními hráči.

Řešení: výklad výstupu programu NCSS2000.

1. Vyčíslení minimálního počtu iterací:

Číslo iterace	Počet shluků	Procento proměnlivosti	Čárový diagram procenta
2	2	65.54	
4	3	46.48	
8	4	29.17	

Oddíl pomáhá určit optimální počet shluků. **Počet vytvořených shluků.** **Procento proměnlivosti:** dává sumu čtverců pro daný počet shluků v tomto řádku ve formě procenta, kdyby vůbec nedošlo ke shlukování. **Čárový diagram procenta proměnlivosti:** grafické zobrazení procenta variace čárovým diagramem.

2. Iterování:

Číslo iterace	Počet shluků	Procento proměnlivosti	Čárový diagram procenta
1	2	72.03	
2	2	65.54	
3	2	66.09	
4	3	46.48	
5	3	49.04	
6	3	46.48	
7	4	31.81	
8	4	29.17	
9	4	29.17	

Oddíl je zvláště užitečný k určení dostatečného počtu náhodných startovacích uspořádání. Když bylo specifikováno dost startovacích uspořádání, obvykle dvě či tři představují optimum pro každý počet shluků. To se pak posuzuje dle dosaženého minima procenta variace. Když se vůbec neobjeví, je třeba zvýšit počet startovacích uspořádání a výpočet opakovat. **Počet shluků:** v tomto uspořádání. **Procento proměnlivosti:** dává sumu čtverců pro počet shluků tohoto řádku vyjádřený jako procento sumy čtverců, když ke shlukování nedojde. Tak, jak se přidává více a více shluků, tato hodnota může poklesnout. Po volbě optimálního počtu shluků, bod selhání tohoto procenta dramaticky poklesne. **Čárový diagram procenta:** dává čárový diagram znázornění procenta variace.

3. Střední hodnoty:

Proměnná	Shluk 1	Shluk 2	Shluk 3
Height	78.25	85.5	77
FgPct	48.6375	54.95	40.75
Points	25.575	27.35	16.75
Rebounds	8.225	17.05	13.9
Počet	8	2	2

Tabulka ukazuje střední hodnoty proměnných pro jednotlivé shluky. Poslední řádek přináší počet objektů v dotyčném shluku.

4. Směrodatné odchylky shluků:

Proměnná	Shluk 1	Shluk 2	Shluk 3
Height	2.171241	0.7071068	6.363961
FgPct	3.357694	1.343503	4.596194
Points	3.770089	3.889087	2.333452
Rebounds	2.544321	8.273149	12.30366
Počet	8	2	2

Tabulka ukazuje směrodatné odchylky proměnných pro jednotlivé shluky. Poslední řádek přináší počet objektů v dotyčném shluku.

5. Testování F -poměru:

Proměnné	SV1	SV2	Mezi průměrnými čtverci	Uvnitř průměrných čtverců	F -test	Spočtená hladina významnosti α
Height	2	9	48.125	8.222222	5.85	0.023532
FgPct	2	9	101.6469	11.31653	8.98	0.007170
Points	2	9	72.7475	13.34056	5.45	0.028096
Rebounds	2	9	75.04459	29.46	2.55	0.132844

Oddíl přináší výsledky jednofaktorové analýzy rozptylu pro každou proměnnou, a to při použití běžně definovaných shluků jako faktoru. Tabulka pomáhá vyšetřit důležitost každé proměnné ve shlukovacím procesu.

6. Vzdálenosti:

Řádek	Shluk	Vzdálenost 1	Vzdálenost 2	Vzdálenost 3
1 Jabbar K.A.	2	2.4609	1.1263	4.0315
2 Barry R.	1	0.9139	3.1499	1.9940
3 Baylor E.	1	1.4427	3.1724	2.2139
4 Bird L.	1	0.8398	1.8867	2.7392
5 Chamberlain W.	2	3.2456	1.1263	4.4712
6 Cousy B.	3	2.9971	5.3790	1.9512
7 Erving J.	1	0.4724	2.4891	2.5912

8 Johnson M.	1	1.6497	2.5426	2.8064
9 Jordan M.	1	1.5532	2.8939	4.0067
10 Robertson O.	1	0.3409	2.9490	2.5629
11 Russell B.	3	3.3878	3.5197	1.9512
12 West J.	1	1.0971	3.6374	2.8439

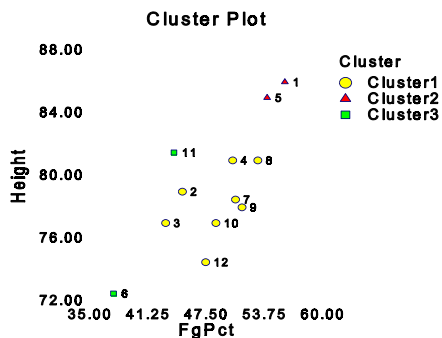
Tabulka přináší relativní vzdálenosti každého objektu ke středu shluku. Tím se prokáže, jak ostře byly shluky vytvořeny: jestliže vzdálenost každého objektu ke středu shluku je mnohem menší než vzdálenost z objektu ke středům ostatních shluků, je uspořádání do shluků výtečné a účinné. Jestliže však nejmenší vzdálenost je téměř stejná jako vzdálenosti ke středům ostatních shluků, je pak otázkou, ke kterému shluku daný objekt vlastně patří. Takové uspořádání pak není příliš žádoucí.

7. Oddíl jednotlivých vzdáleností:

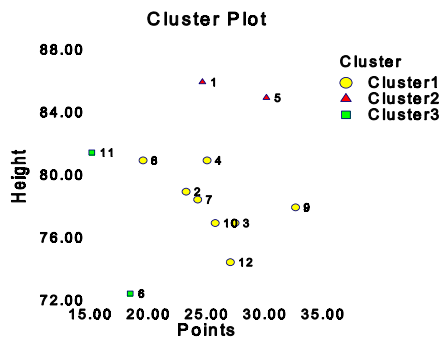
Vzdálenosti pro shluk 1:				
Řádek	Shluk	Vzdálenost 1	Vzdálenost 2	Vzdálenost 3
2 Barry R.	1	0.9139	3.1499	1.9940
3 Baylor E.	1	1.4427	3.1724	2.2139
4 Bird L.	1	0.8398	1.8867	2.7392
7 Erving J.	1	0.4724	2.4891	2.5912
8 Johnson M.	1	1.6497	2.5426	2.8064
9 Jordan M.	1	1.5532	2.8939	4.0067
10 Robertson O.	1	0.3409	2.9490	2.5629
12 West J.	1	1.0971	3.6374	2.8439
Počet = 8				
Vzdálenosti pro shluk 2:				
Řádek	Shluk	Vzdálenost 1	Vzdálenost 2	Vzdálenost 3
1 Jabbar K. A.	2	2.4609	1.1263	4.0315
5 Chamberlain W.	2	3.2456	1.1263	4.4712
Počet = 2				
Vzdálenosti pro shluk 3:				
Řádek	Shluk	Vzdálenost 1	Vzdálenost 2	Vzdálenost 3
6 Cousy B.	3	2.9971	5.3790	1.9512
11 Russell B.	3	3.3878	3.5197	1.9512
Počet = 2				

Vysvětlení tohoto oddílu je stejné jako předešlého, kromě řádku, ve kterém byl zobrazen jediný shluk. Je pak snadnější poznat, který objekt připadá do kterého shluku.

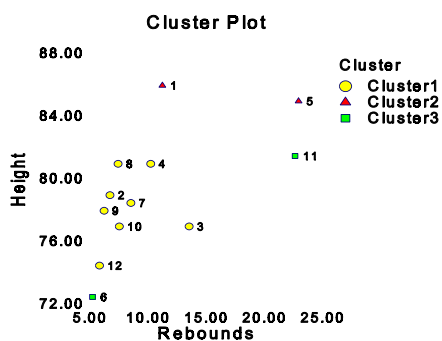
8. Grafy:



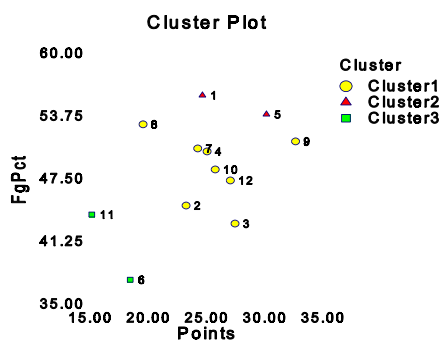
Obr. 4.22a Roštňový diagram proměnných Height-FgPct.



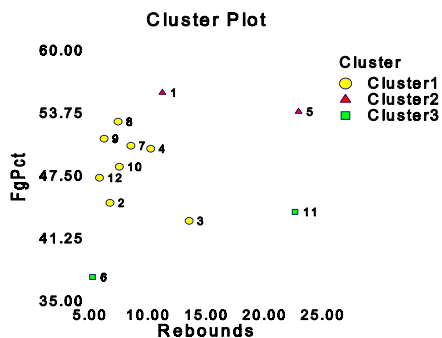
Obr. 4.22b Roštňový diagram proměnných Height-Points.



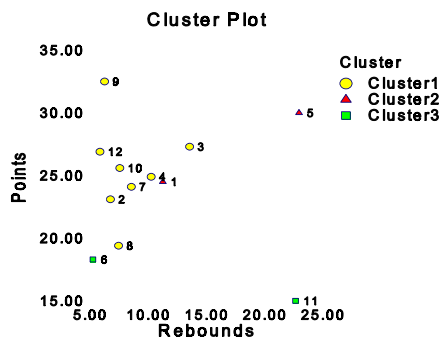
Obr. 4.22c Roštňový diagram proměnných Height-Rebounds.



Obr. 4.22d Roštňový diagram proměnných FgPct-Points.



Obr. 4.22e Roštňový diagram proměnných FgPct-Rebounds.



Obr. 4.22f Roštňový diagram proměnných Points-Rebounds.

Série indexových rozptylových diagramů, vždy pro dvojici proměnných, ukazuje rozdělení objektů do shluků. Diagramy jsou vlastně cílem shlukové analýzy, protože pomohou odhalit odlehle objekty, anomálie a řadu dalších strukturních problémů.

Vzorová úloha 4.13 Odhalení struktury objektů rozličnými metodami shlukování

Vstupní data jsou z úlohy S4.16 *Shluky 12 superhvězd košíkové*. Tabulka dat obsahuje informace o osmi hráčských vlastnostech a aktivitách 12 superhvězd košíkové v sezóně 1989. Cílem je najít shluky hráčů podobných vlastností, a naopak odhalit jejich aktivity a vlastnost, ve které se hráč neshoduje s ostatními hráči.

Řešení: Výklad výstupu programu NCSS2000. Při shlukovací analýze těchto dat bude použito dvou rozličných metod medoidního shlukování programu NCSS2000:

(i) Výstup algoritmu medoidního shlukování za použití Späthovy metody:

1. Titulní stránka: Metoda Späthova

Proměnné:	Height až do Rebounds
Metoda:	Späthova,
Objektivní funkce:	Silueta,
Typ vzdálenosti:	Eukleidovský,
Typ škály:	Směrodatná odchylka,

2. Iterační průběh hledání optimálního počtu shluků:

Počet shluků	Minimalizovat tuto průměrnou vzdálenost	Nastavená průměrná vzdálenost	Maximalizovat tuto průměrnou siluetu
2	35.977405	5.996234	0.135735
2	34.352873	5.725479	0.185579
2	34.862052	5.810342	0.170356
2	36.031237	6.005206	0.101405
3	19.525066	4.881267	0.094407
3	21.106317	5.276579	0.033435
3	19.005957	4.751489	0.045621
3	22.202362	5.550590	-0.026350
4	12.547872	4.182624	-0.013869
4	12.318440	4.106147	0.044989
4	12.210147	4.070049	0.018876
4	14.209356	4.736452	-0.097672
5	9.344940	3.893725	-0.099737
5	10.556815	4.398673	-0.189487
5	8.274123	3.447551	-0.045335
5	8.049819	3.354091	-0.004580

Průběh ukazuje hodnoty účelových funkcí D , D_{adjust} a s pro každou iteraci a proměnný počet shluků metodou Späthovou. Tabulka je zvláště užitečná při zadávání správného počtu opakování: jsou-li totiž dvě či tři shluková uspořádání při stejném požadovaném počtu shluků totožná, potom byl zvolen dostatečně velký zadaný počet opakování. Jinak bylo třeba zvýšit tuto hodnotu a analýzu opakovat. V tomto příkladu dojdeme k závěru, že optimální k je rovno 2. Protože jsme hodnotu siluety $s = 0.185579$ obdrželi pouze jednou, je třeba změnit počet opakování na 10 a výpočet opakovat. **Průměrná vzdálenost:** tato hodnota může být přepočtena tak, že uvede procento vzdálenosti, vztažené vůči maximální vzdálenosti v matici vzdáleností. Tím se zlepší interpretace této veličiny. **Nastavená průměrná vzdálenost:** jde o hodnotu nastavené průměrné vzdálenosti dle vzorce

7 Erving J.	2	1	23.03	40.90	0.4369	
6 Cousy B.	2	1	45.02	68.58	0.3435	
8 Johnson M.	2	1	30.31	45.50	0.3339	
4 Bird L.	2	1	27.19	40.44	0.3275	
Shlukový průměr	2	(7)	28.69	48.18	0.4077	
Celkový průměr		(12)	39.15	48.18	0.1856	
Maximální vzdálenost: 3.012578						

Tabulka přináší cenné informace o každém objektu v řádku, který byl zařazen do shluků. Tabulka je uspořádána podle klesajících hodnot siluety v každém shluku. **Řádek:** číslo nebo jméno objektu v řádku. V tomto výstupu je vysvětlen každý řádek databáze. **Číslo shluku:** číslo, do kterého jsou objekty rozříděny. **Nejbližší souseď:** identifikační číslo nejbližšího shluku vůči tomuto objektu v řádku. Tato informace je užita při výpočtu hodnoty siluety. **Průměrná vzdálenost k němu:** průměrná vzdálenost mezi tímto objektem a ostatními objekty v uvažovaném shluku. Jde o číslo *a* ve výpočtu siluety. **Průměrná vzdálenost souseďa:** průměrná vzdálenost mezi tímto objektem a objekty v nejbližším sousedním shluku. Jde o číslo *b* ve výpočtu siluety. **Silueta:** hodnota *s* musí být kladná a pokud možno vyšší než 0.5, jinak je shluk nevhodný, špatný.

(ii) Medoidního shlukování metodou PAM:

1. Titulní stránka: Metoda PAM (Partition Around Medoids):

Proměnné:	Height až do Rebounds
Metoda:	Kaufmanova-Rousseeuwova,
Objektivní funkce:	Silueta,
Typ vzdálenosti:	Eukleidovská,
Typ škály:	Směrodatná odchylka,

2. Iterační průběh hledání optimálního počtu shluků:

Počet shluků	Minimalizovat tuto průměrnou vzdálenost	Nastavená průměrná vzdálenost	Maximalizovat tuto průměrnou siluetu
2	37.478948	6.246491	0.382164
3	44.870219	11.217555	0.340904
4	31.340599	10.446866	0.270905
5	21.523501	8.968126	0.198372

3. Souřadnice medoidů shluků:

Proměnná	Shluk 1	Shluk 2
Height	86	77
Weight	230	210
FgPet	55.9	48.5
FtPct	72.1	83.8
Points	24.6	25.7
Rebounds	11.2	7.5
Řádek	1 Jabbar K.A	10 Robertson O.

Optimálního počtu shluků je dosaženo pro maximální $F(U)$ a minimální $D(U)$. Po určení optimálního počtu shluků bude řešení studováno detailně. Jelikož zde vyšlo, že optimální jsou 3 shluky, budeme si nadále všimnout jenom výsledků pro 3 shluky a ostatní části výstupu pro počty shluků 2, 4, a 5 vynecháme.

3. Středry (medoidy) shluků:

Proměnná	Shluk 1	Shluk 2	Shluk 3
Červená (Red)	2	14	7
Modrá (Blue)	9	10	2
Řádek	3	10	18

Tato tabulka uvádí středry, čili medoidy, nejbližšího pevného shlukování. Tabulka pomůže vysvětlit podstatu shluků v této úloze. Poslední řádek udává číslo řádku (a někdy také jméno) objektu každého shlukového středru - medoidu.

4. Přehled účasti ve shlucích:

Přehled účasti ve 3 shlucích:						
Row	Shluk	Shluková účast	Suma čtverců účasti	Čárový diagram účasti	Silueta	Diagram siluety
3	1	0.9354	0.8770		0.6907	
5	1	0.8779	0.7782		0.6681	
2	1	0.8728	0.7698		0.6892	
1	1	0.8705	0.7662		0.6509	
4	1	0.8535	0.7395		0.6060	
6	1	0.4227	0.3561		0.1246	
13	1	0.3670	0.3374		-0.0544	
10	2	0.8705	0.7662		0.8282	
8	2	0.8680	0.7622		0.8117	
11	2	0.8556	0.7425		0.8057	
9	2	0.8508	0.7352		0.7905	
12	2	0.8402	0.7188		0.8106	
7	2	0.8231	0.6931		0.7660	
18	3	0.9209	0.8511		0.8627	
21	3	0.8663	0.7595		0.8398	
19	3	0.8652	0.7577		0.8324	
17	3	0.8647	0.7570		0.8210	
15	3	0.8483	0.7312		0.8023	
20	3	0.8258	0.6972		0.8050	
22	3	0.8258	0.6971		0.8145	
16	3	0.8012	0.6617		0.7764	
14	3	0.8000	0.6603		0.7556	

Tabulka přináší informaci o každém objektu (každém řádku). Je rozříděna dle hodnoty siluety v každém shluku. Všimněme si, jak dobře jsou identifikovány dva odlehlé objekty, č. 6 a č. 13. **Řádek:** číslo nebo i jméno objektu, řádku. **Shluk:** číslo shluku, do kterého je objekt klasifikován. **Shluková účast:** jde o maximum shlukové účasti. Je to účast pro shluky, do kterých byl tento řádek (objekt) zařazen při pevném shlukování. **Suma čtverce účasti:** všechny účasti pro daný objekt (řádek) jsou umocněny na druhou a sečteny. Když je objekt úplně přiřazen do jediného shluku, je tato hodnota rovna 1. Když je řádek klasifikován stejnou měrou do každého shluku, bude hodnota $1/K$. Řádky s vysokými hodnotami jsou proto ve shluku blízko středu. Řádky s nízkými hodnotami jsou odlehlé hodnoty. **Čárový diagram čtverce účasti:** čárový diagram sumy čtverců hodnot účasti pomůže odhalit

řádky, které nejsou dobře zařazeny do shluků. **Silueta:** hodnota siluety by měla být kladná a větší než 0.5. **Čárový diagram hodnot siluety:** pomůže rozlišit řádky, objekty, které nejsou dobře zařazeny do shluků.

5. Účast objektů ve shlucích: ukazuje na pravděpodobnost účasti každého řádku v každém shluku.

Řádek	Shluk	Pravděpod. v 1	Pravděpod. v 2	Pravděpod. v 3
1	1	0.8705	0.0572	0.0723
2	1	0.8728	0.0601	0.0671
3	1	0.9354	0.0289	0.0357
4	1	0.8535	0.0613	0.0852
5	1	0.8779	0.0558	0.0663
6	1	0.4227	0.3620	0.2154
7	2	0.0860	0.8231	0.0909
8	2	0.0659	0.8680	0.0661
9	2	0.0673	0.8508	0.0819
10	2	0.0635	0.8705	0.0660
11	2	0.0648	0.8556	0.0797
12	2	0.0745	0.8402	0.0853
13	1	0.3670	0.2821	0.3509
14	3	0.1135	0.0865	0.8000
15	3	0.0803	0.0714	0.8483
16	3	0.0973	0.1015	0.8012
17	3	0.0751	0.0602	0.8647
18	3	0.0415	0.0377	0.9209
19	3	0.0662	0.0686	0.8652
20	3	0.0950	0.0792	0.8258
21	3	0.0694	0.0642	0.8663
22	3	0.0858	0.0884	0.8258