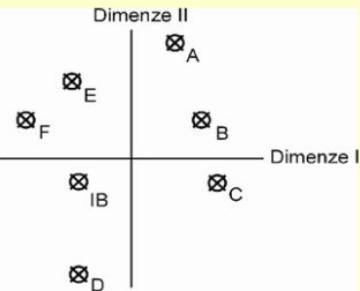


Vícerozměrné škálování objektů (MDS)

Je vytvoření **subjektivní mapy** relativního umístění objektů v rovině grafu, a to na základě vzdáleností či podobnosti mezi objekty, tzv. **matice proximity** (blízkosti).

Cílem: detekovat nejlepší počet souřadnic, které objasní podobnosti či vzdálenosti mezi vyšetřovanými objekty. Mapa objektů v jedné, dvou či tří souřadnicích pak vzájemně porovnává objekty.

Podstata: v FA či PCA jsou podobnosti mezi objekty vyjádřeny v korelační matici, v MDS se analyzuje jakýkoliv druh podobnostní či vzdálenostní matice (proximita).

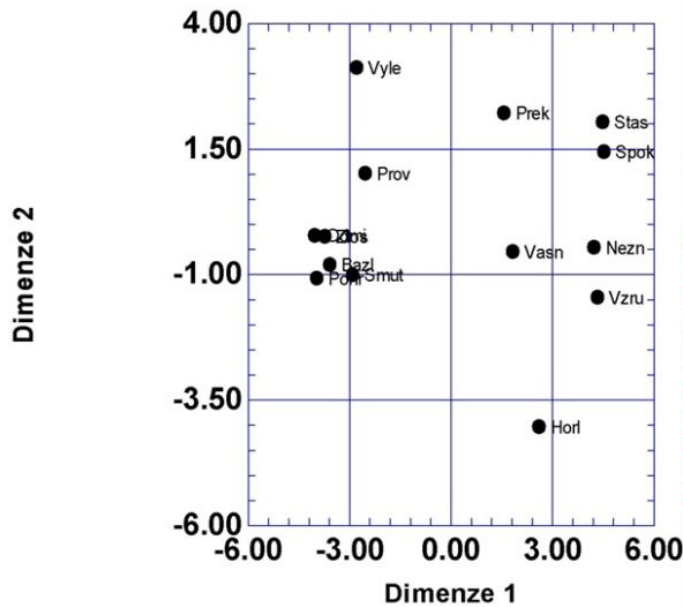


Příklad: matice vzdáleností (proximity) mezi městy.

Řešení: MDS vytvoří mapu, kde poloha se vyjádří dvěma souřadnicemi, sever-jih na y a západ-východ na x.

Postup:

- Počet vhodných souřadnic k vytvoření mapy měst, obvykle 2.
- Vzdálenosti mezi městy na souřadnice polohy města na mapě.
- Otočení souřadnicových os tak, aby se poloha dala co nejlépe vysvětlit.



Dimenze 2

Dimenze 1

Dvojměrný škálovací diagram CMDS podobností u 14 emocí, **stress = 0.060** (NCSS2000).

Posuďte porovnání "každé emoce s každou".

Matice proximity *Emoce*

vyjadřuje vzdálenosti, nepodobnosti dvojic těchto emocí:

- Spok** značí spokojený,
- Vzru** značí vzrušující,
- Prek** překvapený,
- Horl** horlivý,
- Stas** šťastný,
- Vasn** vášnivý,
- Nezn** něžný,
- Pohr** pohrdá,
- Vyle** vylekaný,
- Bazl** bázlivý,
- Prov** provinilý,
- Smut** smutný,
- Zlost** zlostný,
- Odmí** odmítavý.

Intuitivně podobné emoce jako spokojený **Spok** a šťastný **Stas** budou ležet těsně u sebe.

Dimenze 1 bude oddělovat pozitivní emoce od negativních.

Matice proximity *Emoce* obsahuje nepodobnosti dvojice následujících emocí:

Spok značí spokojený, **Vzru** značí vzrušující, **Prek** překvapený, **Horl** horlivý, **Stas** šťastný, **Vasn** vášnivý, **Nezn** něžný, **Pohr** pohrdá, **Vyle** vylekaný, **Bazl** bázlivý, **Prov** provinilý, **Smut** smutný, **Zlost** zlostný, **Odmí** odmítavý.

Kodování: Při dokonalé podobnosti každé emoce s každou je jejich vzdálenost **0**, při jejich nepodobnosti pak vzdálenost **10**.

	Spok	Vzru	Prek	Horl	Stas	Vasn	Nezn	Pohr	Vyle	Bazl	Prov	Smut	Zos	Odmí
Spok	0													
Vzru	7	0												
Prek	7	7	0											
Horl	9	4	9	0										
Stas	1	3	6	9	0									
Vasn	3	3	4	2	1	0								
Nezn	6	2	9	6	7	1	0							
Pohr	9	7	8	9	9	3	9	0						
Vyle	9	6	4	2	9	6	6	9	0					
Bazl	9	6	3	9	9	3	9	4	1	0				
Prov	9	9	9	9	9	3	9	9	2	4	0			
Smut	9	7	3	8	9	3	6	9	2	2	4	0		
Zos	9	9	7	6	9	1	5	2	4	3	9	6	0	
Odmí	9	9	7	9	9	9	9	9	2	1	4	1	3	0

Každý objekt je popsán svými **znaky**:

- znaky subjektivními** (znaky vnímané člověkem, hezký, ošklivý, laciný, drahý, barvy, ...),
- znaky objektivními** (znaky měřitelné fyzikálně).

Matice proximity

Matice vzdáleností objektů *D*: přímo měřená vzdálenost (disimilarita, nepodobnost) mezi objekty d_{ij} .

Matice podobností objektů *S*: podobnost mezi objekty S_{ij} vyjadřují jak blízko se nacházejí dva objekty. Podobnost lze převést do veličiny vzdálenost d_{ij} vztahem

$$d_{ij} = \sqrt{S_{ii} + S_{jj} - 2S_{ij}}$$

Matice znaků objektů *X*: hodnoty znaků (sloupce) pro objekty (řádky) X_{ij} představují proměnné pro jednotlivé objekty a jsou to spíše standardní míry. Z nich se vypočte nejprve korelační matice objektů **R** a nebo matice eukleidovských vzdáleností **D** či matice Mahalanobisových vzdáleností objektů **D**. Z matice znaků pro všechny objekty **X** lze vyčíslit přímo matici vzdáleností **D** vztahem

$$d_{ij} = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2}$$

Vzdálenost d_{ij} mezi dvěma objekty x_i a x_j je eukleidovská (Pythagorova věta).

Maticе proximity

1. Maticе vzdáleností objektů D:

Trojúhelníková **maticе proximity** *Sporty* obsahuje vzdálenosti objektů u dvojice her:

Hokej, Fotbal, Basket, Tenis, Golf, Krokety.

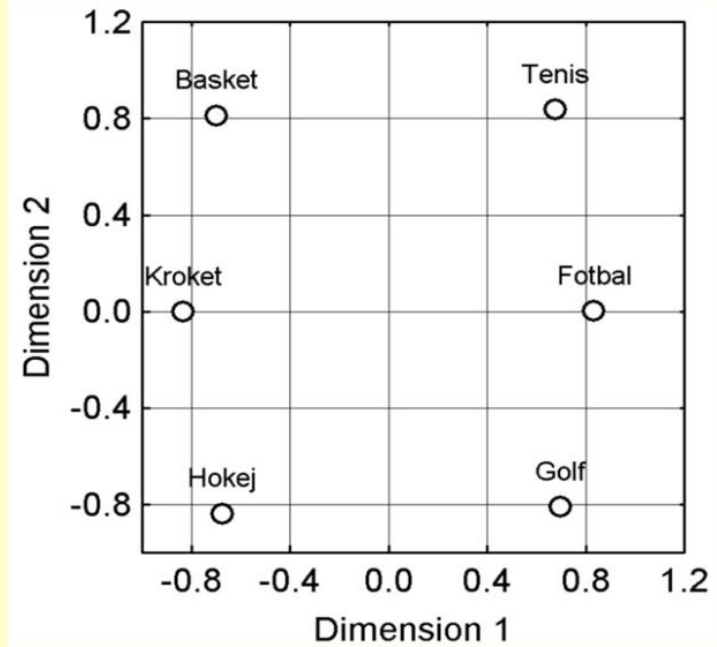
Kodování: Při dokonalé podobnosti je vzdálenost 1, při nepodobnosti 6. Párové vzdálenosti tvoří prvky čtvercové matice, ze které se užije trojúhelníková část nad diagonálou jedniček.

	<i>Hokej</i>	<i>Fotbal</i>	<i>Basket</i>	<i>Tenis</i>	<i>Golf</i>	<i>Krokety</i>
<i>Hokej</i>	1					
<i>Fotbal</i>	2	1				
<i>Basket</i>	3	3	1			
<i>Tenis</i>	4	5	5	1		
<i>Golf</i>	5	6	4	4	1	
<i>Krokety</i>	5	5	6	3	2	1

MDS si vytváří svou „náhodnou proměnnou L “,

tj. *subjektivní souřadnici* L jako řádek pomocné matice L

vhodnou k vzájemnému porovnání objektů na MDS mapě.



Obr. 10.4a Dvozměrný škálovací CMDS diagram vyjadřuje příbuznost či podobnost 6 sportovních her, **stress = 0.100**, (STATISTICA).

2. Maticе korelace objektů R

Pro šest předmětů na vysvědčení u 220 žáků byla vyčíslena korelační matice R .

Prvky **korelační matice** *řvsve* obsahují ve dvojicích tyto předměty:

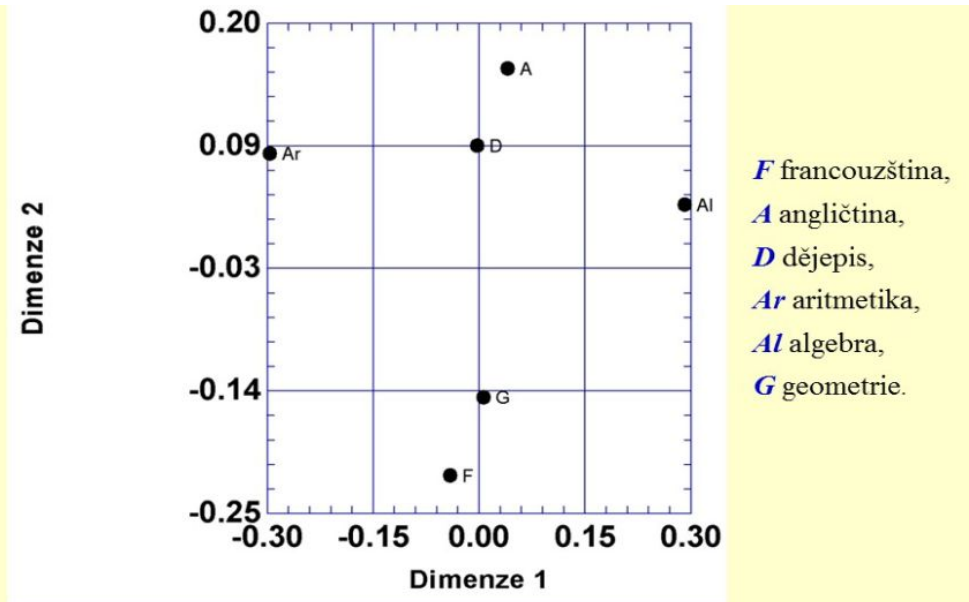
F francouzština, *A* angličtina, *D* dějepis, *Ar* aritmetika, *Al* algebra, *G* geometrie.

	<i>F</i>	<i>A</i>	<i>D</i>	<i>Ar</i>	<i>Al</i>	<i>G</i>
<i>F</i>	0					
<i>A</i>	0.44	0				
<i>D</i>	0.41	0.35	0			
<i>Ar</i>	0.29	0.35	0.16	0		
<i>Al</i>	0.33	0.32	0.19	0.59	0	
<i>G</i>	0.25	0.33	0.18	0.47	0.46	0

MDS si vytváří svou „náhodnou proměnnou L “,

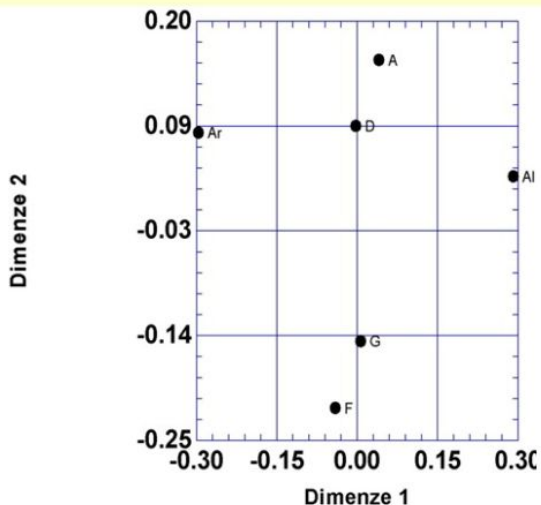
tj. *subjektivní souřadnici* L jako řádek pomocné matice L

vhodnou k vzájemnému porovnání objektů na MDS mapě.

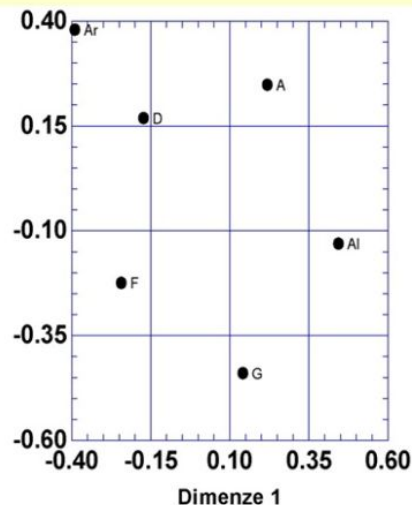


Dvozměrný škálovací diagram CMDS podobnosti 6 předmětů na vysvědčení, **stress = 0.097**, (NCSS2000).

Hledání nejlepší metody škálovací mapy podobností objektů.



Obr. 10.8e Dvojměrný škálovací diagram CMDS podobnosti 6 předmětů na vysvědčení, $stress = 0.097$, (NCSS2000).



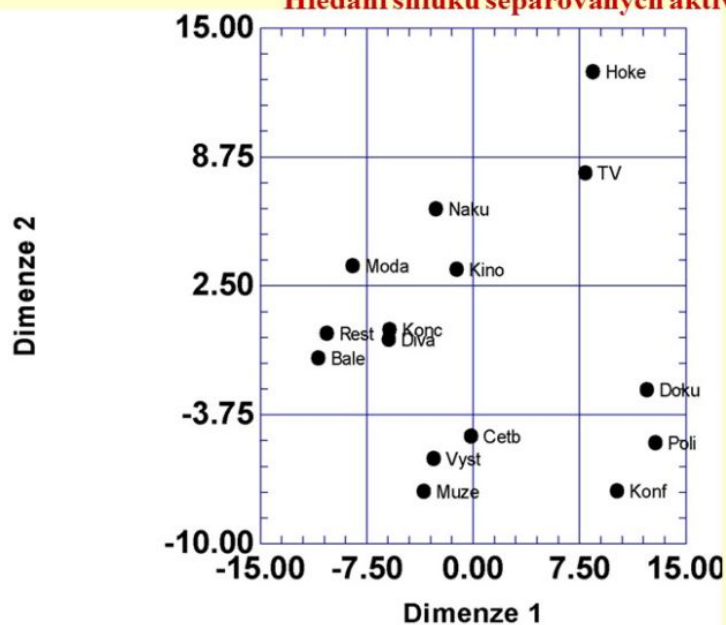
Obr. 10.8f Dvojměrný škálovací diagram NNMDS podobnosti 6 předmětů na vysvědčení, $stress = 0.173$, (NCSS2000).

Matice proximit Relax obsahuje vzdálenosti jednotlivých aktivit relaxace ve dvojicích: *Konc* značí koncert, *Muze* značí muzeum, *Diva* divadlo, *Kino* kino, *TV* televize, *Konf* konference, *Cetb* četba, *Hoke* divák hokeje, *Bale* balet, *Poli* zájem o politiku, *Moda* móda, *Doku* dokumentaristika, *Vyst* zájemce o výstavy, *Naku* zájem o nákupy, *Rest* pobyt v restauraci.

Kodování: Při dokonalé podobnosti je vzdálenost 0, při nepodobnosti pak 25.

	Konc	Muze	Diva	Kino	TV	Konf	Cetb	Hoke	Bale	Poli	Moda	Doku	Vyst	Naku	Rest
Konc	0														
Muze	16	0													
Diva	3	18	0												
Kino	12	12	11	0											
TV	15	21	16	2	0										
Konf	20	10	19	15	12	0									
Cetb	15	12	13	9	19	6	0								
Hoke	21	23	23	19	7	22	20	0							
Bale	7	10	6	18	19	25	15	25	0						
Poli	19	22	25	22	14	8	22	23	25	0					
Moda	9	7	13	15	12	19	20	22	8	25	0				
Doku	22	16	16	19	13	7	13	15	23	13	25	0			
Vyst	7	3	13	12	21	13	10	22	13	12	7	18	0		
Naku	21	22	22	12	23	21	18	18	21	22	9	22	12	0	
Rest	8	8	7	9	21	21	2	22	5	25	9	23	10	8	0

Hledání shluků separovaných aktivit osobní relaxace



Mapa separuje 15 aktivit osobní relaxace a indikuje 4 shluky.

Dimenze 1 se pojmenuje *aktivní relaxace* (jde o intelektuální náročnost aktivit).

Dimenze 2 se pojmenuje *pasivní relaxace* souvisí spíše s mírou vzrušení při zábavě či s akčností zábavy.

Aktivní a pasivní je zde třeba chápat jako protipóly.

CMDS diagram podobnosti 15 aktivit relaxace, $stress = 0.089$, (NCSS2000).

MDS vyčíslí a) **metrické klasické řešení** (CMDS),
b) **nemetrické řešení** (NNMDS)

a vychází buď přímo z experimentálních hodnot X ,
z korelační matice R ,
z matice podobnosti S ,
z matice vzdáleností D .

- Pro n objektů existuje $n(n - 1)/2$ podobností čili vzdáleností mezi páry objektů.

- Když podobnosti objektů nemohou být kvantifikovány, (např. podobnost mezi barvami),
užijeme za vstupní data **pořadová čísla podobností objektů**.

Podobnosti objektů S lze **vzestupně uspořádat do řady**

$$S_{i_1, j_1} \leq S_{i_2, j_2} \leq S_{i_3, j_3} \leq \dots \leq S_{i_m, j_m}$$

S_{i_1, j_1} značí nejmenší podobnost ze všech

Chceme nalézt q -rozměrnou sestavu n objektů tak, že vzdálenosti mezi párem objektů souhlasí s pořadovými čísly uvedených podobností. **Řada pořadových čísel objektů**

$$d_{i_1, j_1}^{(q)} \leq d_{i_2, j_2}^{(q)} \leq d_{i_3, j_3}^{(q)} \leq \dots \leq d_{i_m, j_m}^{(q)}$$

popisuje sestupně řazení vzdáleností párů objektů od největší do nejmenší.

Test těsnosti proložení s využitím statistické míry *stress*: je založen na rozdílu mezi skutečnou vzdáleností dvou objektů d_{ij} a modelem predikovanou vzdáleností dle vzorce

$$stress = \sqrt{\frac{\sum_{k=1}^m (d_{ij} - d_{ij,vyp})^2}{\sum_{k=1}^m d_{ij}^2}}$$

kde $d_{ij,vyp}$ je vypočtená vzdálenost objektů dle modelu MDS.

Je-li hodnota *stress* blízká nule, jeví se proložení jako nejlepší.

Pravidlo: čím menší je hodnota *stress*, tím těsnější proložení mezi vypočtenými a zadanými souřadnicemi objektů bylo dosaženo.

Kruskal navrhl **míru důležitosti *stress(q)***, dle které je souhlas modelu MDS s daty co nejlepší,

$$stress(q) = \sqrt{\frac{\sum_{i<j}^m \sum_j^m (d_{ij}^{(q)} - d_{ij,vyp}^{(q)})^2}{\sum_{i<j}^m \sum_j^m d_{ij}^{(q)2}}}$$

kde $d_{ij,vyp}^{(q)}$ jsou vypočtená vzdálenostní pořadová čísla objektů, která jsou vypočtena dle monotonní funkce podobnosti objektů.

Takano zavedl **přednostní kritérium S_{stress}** pro daný počet souřadnic q

$$S_{stress} = \sqrt{\frac{\sum_{i<j}^m \sum_j^m (d_{ij}^2 - d_{ij,vyp}^2)^2}{\sum_{i<j}^m \sum_j^m d_{ij}^4}}$$

v intervalu 0 až 1, a hodnoty menší než 0.1 se týkají dobré prezentace objektů body nalezeného uspořádání.

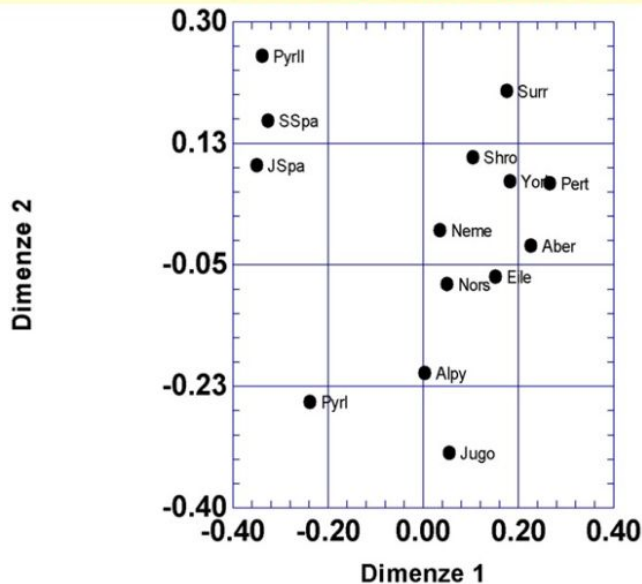
Počet potřebných souřadnic v modelu MDS

Cattellův indexový grafem úpatí relativní velikosti hodnot *stress*, která jsou vyčíslována pro rostoucí počet souřadnic.

Cattellův indexový graf úpatí vlastních čísel formou čarového diagramu

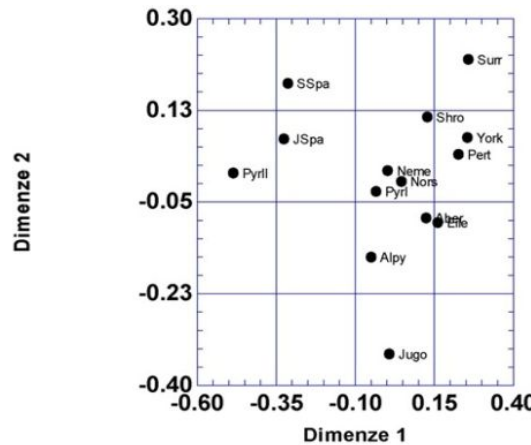
Pořadové číslo souřadnice	Vlastní procento	Jednotlivé procento	Kumulativní procento	Čarový diagram vlastních čísel
1	30.73	54.28	54.28	
2 (Užívané)	12.85	22.69	76.97	
3	6.38	11.27	88.24	
4	1.68	2.97	91.21	
5	0.00	0.00	91.21	
6	-4.98	8.79	100.00	

MDS separuje hryzce do tří shluků:

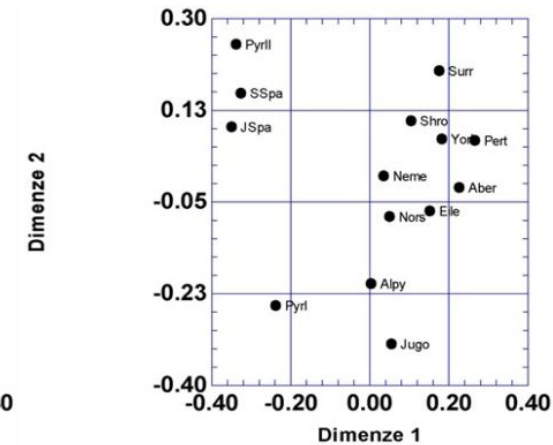


1. **shluk** obsahuje *PyrII*, *SSpa*, *JSpa* a je dobře oddělen od ostatních.
2. **shluk** obsahuje 6 britských populací s německými *Neme* a norskými *Norsk*.
3. **shluk** obsahuje *Alpy*, *Jugo* a *PyrI*.

Kritérium *stress* dosahuje 0.144, a tím ukazuje na těsné proložení dat.



Obr. 10.9e Dvojměrný škálovací **CMDS** diagram podobnosti lebek hryzců 14 míst Evropy, *stress* = 0.093, (NCSS2000).



Obr. 10.9f Dvojměrný škálovací **NNMDS** diagram podobnosti lebek hryzců 14 míst Evropy, *stress* = 0.144, (NCSS2000).

Obr. 10.9e Dvojměrný škálovací **NNMDS** diagram podobnosti lebek hryzců 14 míst Evropy, *stress* = 0.144, (NCSS2000)

Hledání nejlepší metody vícerozměrného škálování při třídění hryzců do shluků, a to CMDS nebo NNMDS.

Porovnání tří metod:

Metoda hlavních komponent PCA, FA třídí znaky, které definují dotyčné rozměry v původním souboru znaků či proměnných. Znaky, které silně korelují, jsou pak zařazeny spolu.

Analýza shluků objektů CLU třídí objekty podle jejich profilu v souboru znaků, ve kterých jsou objekty v těsné blízkosti zařazovány spolu.

Vícerozměrné škálování objektů MDS se liší od analýzy shluků objektů ve dvou klíčových bodech:

- Řešení může být získáno pro každý objekt.
- Neuzivá se znaků ale vypočtených souřadnic v diagramu subjektivní mapy objektů.

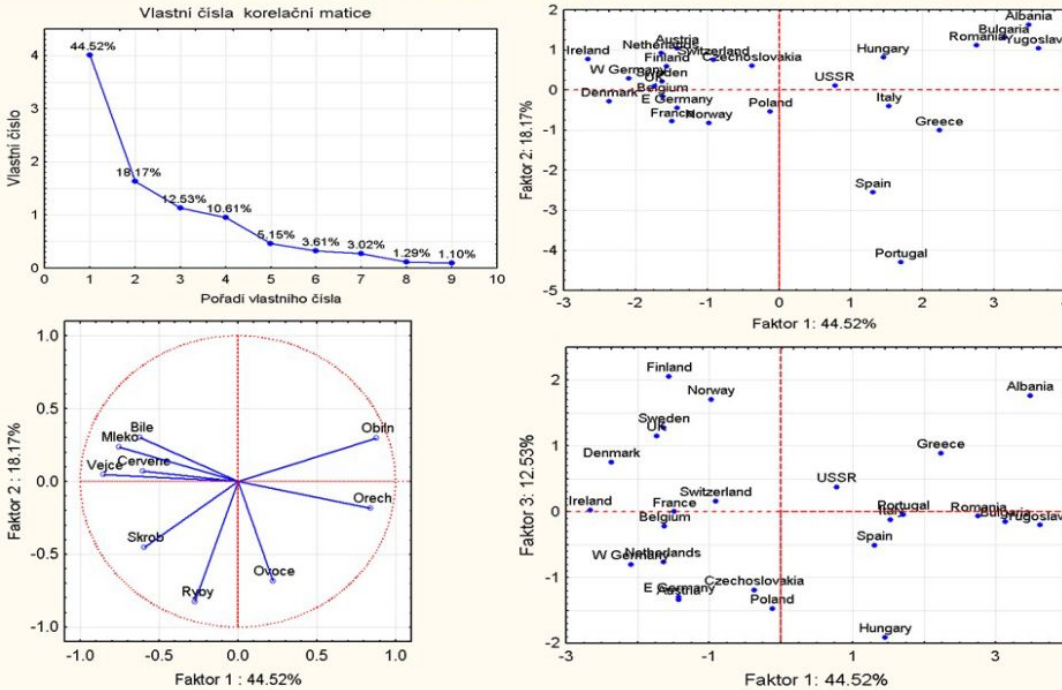
Zdrojová matice dat

Stat	Cervene	Bile	Vejce	Mleko	Ryby	Obiln	Skrob	Orech	Ovoce
Albania	10.1	1.4	0.5	8.9	0.2	42.3	0.6	5.5	1.7
Austria	8.9	14	4.3	19.9	2.1	28	3.6	1.3	4.3
Belgium	13.5	9.3	4.1	17.5	4.5	26.6	5.7	2.1	4
Bulgaria	7.8	6	1.6	8.3	1.2	56.7	1.1	3.7	4.2
Czechoslovakia	9.7	11	2.8	12.5	2	34.3	5	1.1	4
Denmark	10.6	11	3.7	25	9.9	21.9	4.8	0.7	2.4
E Germany	8.4	12	3.7	11.1	5.4	24.6	6.5	0.8	3.6
Finland	9.5	4.9	2.7	33.7	5.8	26.3	5.1	1	1.4
France	18	9.9	3.3	19.5	5.7	28.1	4.8	2.4	6.5
Greece	10.2	3	2.8	17.6	5.9	41.7	2.2	7.8	6.5
Hungary	5.3	12	2.9	9.7	0.3	40.1	4	5.4	4.2
Ireland	13.9	10	4.7	25.8	2.2	24	6.2	1.6	2.9
Italy	9	5.1	2.9	13.7	3.4	36.8	2.1	4.3	6.7
Netherlands	9.5	14	3.6	23.4	2.5	22.4	4.2	1.8	3.7
Norway	9.4	4.7	2.7	23.3	9.7	23	4.6	1.6	2.7
Poland	6.9	10	2.7	19.3	3	36.1	5.9	2	6.6
Portugal	6.2	3.7	1.1	4.9	14	27	5.9	4.7	7.9
Romania	6.2	6.3	1.5	11.1	1	49.6	3.1	5.3	2.8
Spain	7.1	3.4	3.1	8.6	7	29.2	5.7	5.9	7.2
Sweden	9.9	7.8	3.5	24.7	7.5	19.5	3.7	1.4	2
Switzerland	13.1	10	3.1	23.8	2.3	25.6	2.8	2.4	4.9
UK	17.4	5.7	4.7	20.6	4.3	24.3	4.7	3.4	3.3
USSR	9.3	4.6	2.1	16.6	3	43.6	6.4	3.4	2.9
W Germany	11.4	13	4.1	18.8	3.4	18.6	5.2	1.5	3.8
Yugoslavia	4.4	5	1.2	9.5	0.6	55.9	3	5.7	3.2

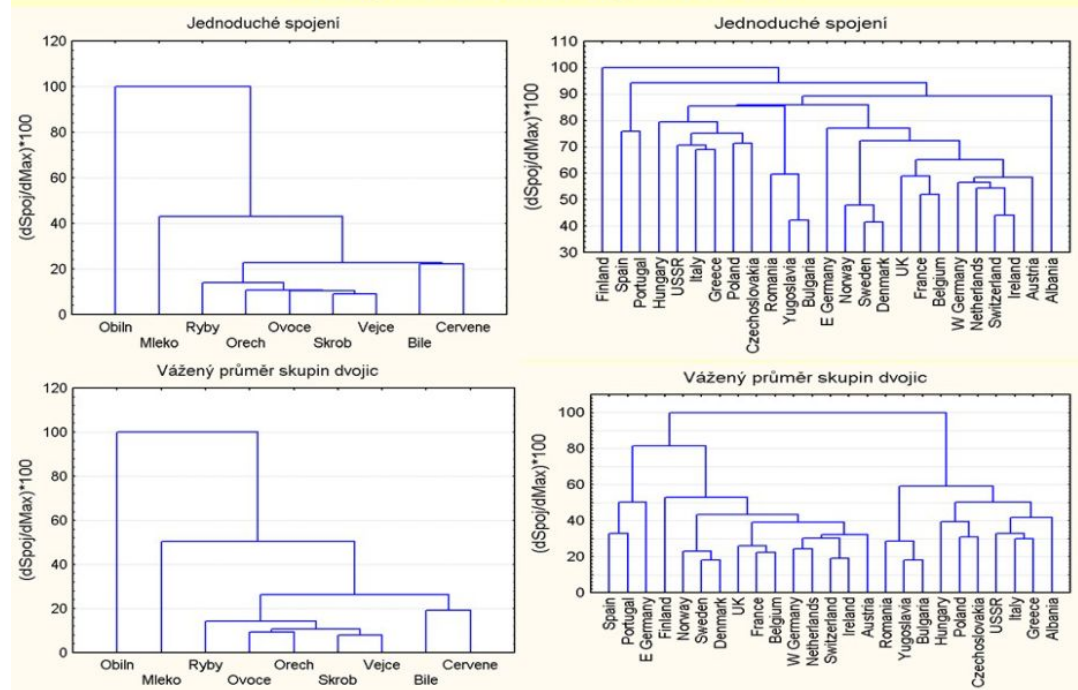
Korelační matice

	Cervene	Bile	Vejce	Mleko	Ryby	Obiln	Skrob
Cervene	1.0000						
Bile	0.1530	1.0000					
Vejce	0.5856	0.6204	1.0000				
Mleko	0.5029	0.2815	0.5755	1.0000			
Ryby	0.0610	-0.2340	0.0656	0.1379	1.0000		
Obiln	-0.4999	-0.4138	-0.7124	-0.5927	-0.5242	1.0000	
Skrob	0.1354	0.3138	0.4522	0.2224	0.4039	-0.5333	1.0000
Orech	-0.3495	-0.6350	-0.5598	-0.6211	-0.1472	0.6510	-0.4743
Ovoce	-0.0742	-0.0613	-0.0455	-0.4084	0.2661	0.0466	0.0844

Metoda hlavních komponent PCA, FA



Analýza shluků objektů CLU

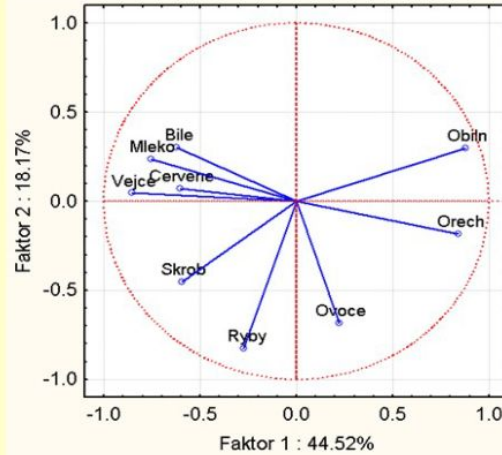
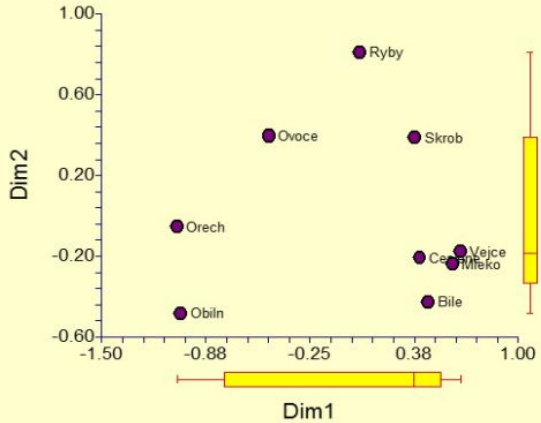


Vícerozměrné škálování objektů MDS

Korelační matice

	Cervene	Bile	Vejce	Mleko	Ryby	Obilin	Skrob	Ovoce
Cervene	1.0000							
Bile	0.1530	1.0000						
Vejce	0.5856	0.6204	1.0000					
Mleko	0.5029	0.2815	0.5755	1.0000				
Ryby	0.0610	-0.2340	0.0656	0.1379	1.0000			
Obilin	-0.4999	-0.4138	-0.7124	-0.5927	-0.5242	1.0000		
Skrob	0.1354	0.3138	0.4522	0.2224	0.4039	-0.5333	1.0000	
Orech	-0.3495	-0.6350	-0.5598	-0.6211	-0.1472	0.6510	-0.4743	1.0000
Ovoce	-0.0742	-0.0613	-0.0455	-0.4084	0.2661	0.0466	0.0844	0.0844

MDS Map



Nalezené řešení a dosažená těsnost proložení

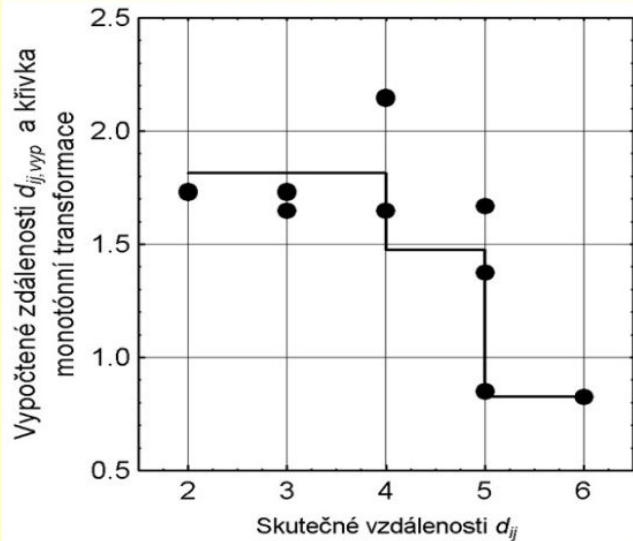
Prvním krokem v určení polohy objektu na mapě je co nejlepší vystižení vyhodnocované podobnosti. Postup určení optimálních poloh objektů v MDS má čtyři fáze:

- Vybere se počáteční sestava podnětů, znaků, vlastností při požadovaném *počátečním počtu souřadnic q*, který je založen na předešlých datech.
 - Vypočtou se vzdálenosti mezi body objektů d_{ij} v počáteční konfiguraci a porovnávají se ve vzdálenostmi $d_{ij,vyp}$ vyčíslenými z jejich odhadů podobnosti S_{ij} . Rozdíly ve vzdálenostech objektů čili v reziduiích vedou k vyčíslení míry těsnosti proložení *stress*.
 - Jestliže *stress* nedosáhne předem zadané hodnoty terminačního kritéria, naleznou se nové konfigurace objektů, pro kterou je *stress* minimalizován. Program určí směry největšího zlepšení *stressu* a přemístí body na mapě v těchto směrech.
 - Po dosažení uspokojivé hodnoty *stress* je počet souřadnic snížen o 1 a proces opakován, dokud není dosaženo nejnižšího počtu souřadnic s přijatelnou těsností proložení *stress*.
- Počet map objektů k interpretaci závisí na počtu zvolených souřadnic, a to pro každou kombinaci souřadnic vždy jedna mapa.

Cílem je co nejlepší těsnost proložení při nejmenším možném počtu souřadnic.

Shepardův diagram je rozptylový diagram vypočtené vzdálenosti pro zvolený počet souřadnic na y -nové ose proti zadaným vzdálenostem na x -ové ose.

Když všechny vypočtené body padnou na schodovitou křivku, je dosaženo těsného proložení. Když ale dojde k odchylkám od křivky, je dosaženo nedostatečného proložení.



Cvičení v programu STATISTICA

	1	2	3	4	5	6	7	8	9	10	11	12
BRAZIL	0,00	4,83	5,28	3,44	4,72	4,50	3,83	3,50	2,39	3,06	5,39	3,17
CONGO	4,83	0,00	4,56	5,00	4,00	4,83	3,33	3,39	4,00	3,39	2,39	3,50
CUBA	5,28	4,56	0,00	5,17	4,11	4,00	3,61	2,94	5,50	5,44	3,17	5,11
EGYPT	3,44	5,00	5,17	0,00	4,78	5,83	4,67	3,83	4,39	4,39	3,33	4,28
FRANCE	4,72	4,00	4,11	4,78	0,00	3,44	4,00	4,22	3,67	5,06	5,94	4,72
INDIA	4,50	4,83	4,00	5,83	3,44	0,00	4,11	4,50	4,11	4,50	4,28	4,00
ISRAEL	3,83	3,33	3,61	4,67	4,00	4,11	0,00	4,83	3,00	4,17	5,94	4,44
JAPAN	3,50	3,39	2,94	3,83	4,22	4,50	4,83	0,00	4,17	4,61	6,06	4,28
CHINA	2,39	4,00	5,50	4,39	3,67	4,11	3,00	4,17	0,00	5,72	2,56	5,06
RUSSIA	3,06	3,39	5,44	4,39	5,06	4,50	4,17	4,61	5,72	0,00	5,00	6,67
USA	5,39	2,39	3,17	3,33	5,94	4,28	5,94	6,06	2,56	5,00	0,00	3,56
YUGOSLAV	3,17	3,50	5,11	4,28	4,72	4,00	4,44	4,28	5,06	6,67	3,56	0,00
Means	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Std.Dev.	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
No.Cases	18,00											
Matice	2,00											

1. Zdrojová matice Nations.sta se načte z databáze EXAMPLES programu STATISTICA. Data jsou tvořena podobnostní maticí (čili maticí proximity), která obsahuje průměry odhadnutých podobností od 18 studentů o vzájemné podobnosti 12 zemí. (1 značí málo podobně, 9 značí silně podobně)

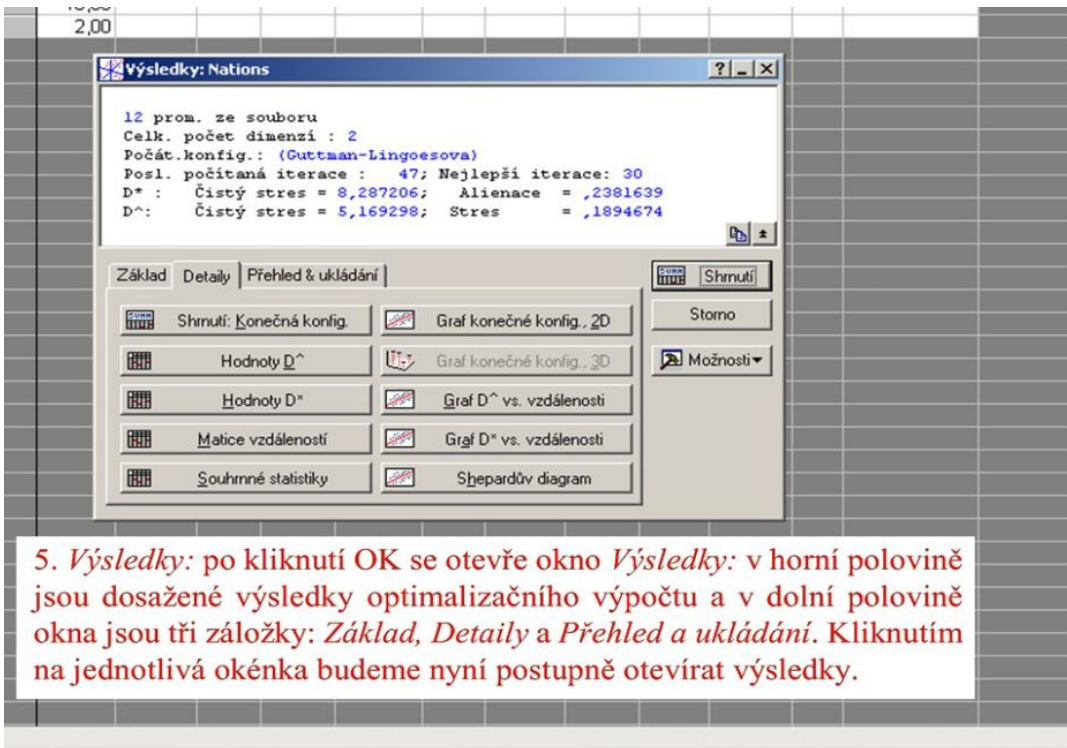
2. Zavolá se Vícerozměrné škálování ze Vícerozměrné průzkumné techniky.

	1	2	3	4	5	6	7	8	9	10	11	12
INDIA	4,50	4,83	4,00	5,83	3,44	0,00	4,11	4,50	4,11	4,50	4,28	4,00
ISRAEL	3,83	3,33	3,61	4,67	4,00	4,11	0,00	4,83	3,00	4,17	5,94	4,44
JAPAN	3,50	3,39	2,94	3,83	4,22	4,50	4,83	0,00	4,17	4,61	6,06	4,28
CHINA	2,39	4,00	5,50	4,39	3,67	4,11	3,00	4,17	0,00	5,72	2,56	5,06
RUSSIA	3,06	3,39	5,44	4,39	5,06	4,50	4,17	4,61	5,72	0,00	5,00	6,67
USA	5,39	2,39	3,17	3,33	5,94	4,28	5,94	6,06	2,56	5,00	0,00	3,56
YUGOSLAV	3,17	3,50	5,11	4,28	4,72	4,00	4,44	4,28	5,06	6,67	3,56	0,00
Means	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Std.Dev.	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
No.Cases	18,00											
Matice	2,00											

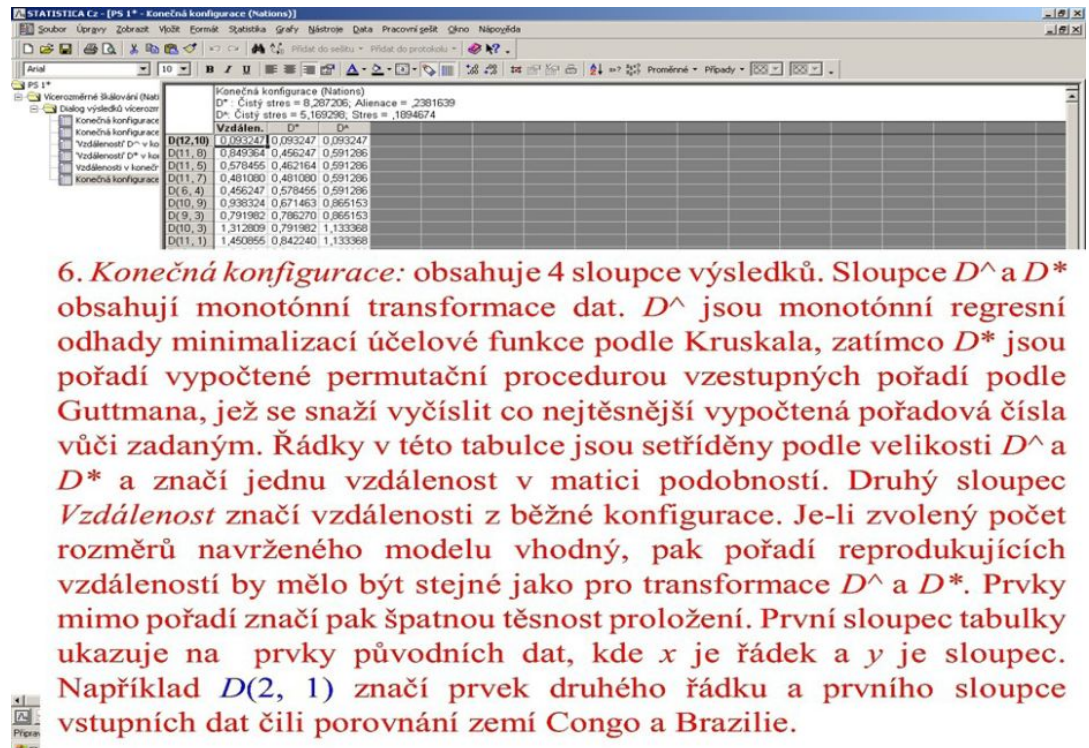
iter.	[dim=2]	D*	D*	D^	d^
s: t:	cosin	krok	čistý stres	alienace	čistý stres
15	1	,968	,485	5,310299	,1920341
16	1	,887	,404	5,304970	,1919377
17	1	,694	,230	5,300321	,1918537
18	1	,711	,197	5,295306	,1917628
19	1	,908	,322	5,284995	,1915760
20	1	,960	,456	5,264813	,1912098
21	1	,898	,411	5,242467	,1908036
22	1	,700	,234	5,228747	,1905538
23	1	,662	,180	5,217968	,1903573
24	1	,877	,281	5,201593	,1900584
25	1	,956	,429	5,182857	,1897157
26	1	,561	,188	5,178150	,1896296
27	1	,665	,168	5,175625	,1895833
28	1	,950	,354	5,170296	,1895187
29	1	,968	,487	5,165822	,1894770
30	1	,698	,247	5,162298	,1894674
30 *			8,287206	,2381639	5,162298
13	1	,697	,184	5,320363	,1922159
14	1	,938	,349	5,315846	,1921343

3. Kliknutím na okénko Proměnné se z matice proximity zadají podobnosti všech 12 zemí (sloupce), a to v okně Vyberte proměnné (objekty) na analýzu a pak OK. Za počet dimenzí se dosadí 2.

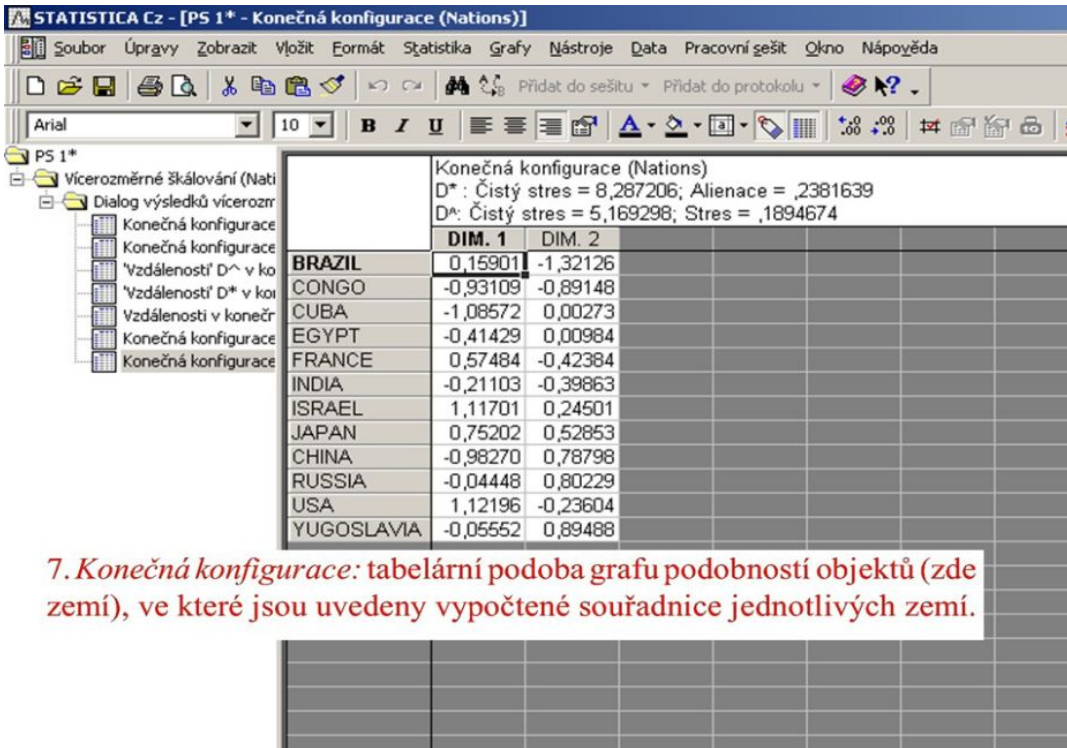
4. Odhadování parametrů: program začne výpočet přes PCA s maticí podobnosti a pak se otevře okno Odhadování parametrů: po počtu iterací metody největšího spádu značeném s následuje ještě několik (obvykle 5) dalších zjemňujících iterací a jejich počet je ve sloupci t. Jsou uvedeny číselné hodnoty veličin těsnosti proložení stress, D^ (metricky) a D* (nemetricky).



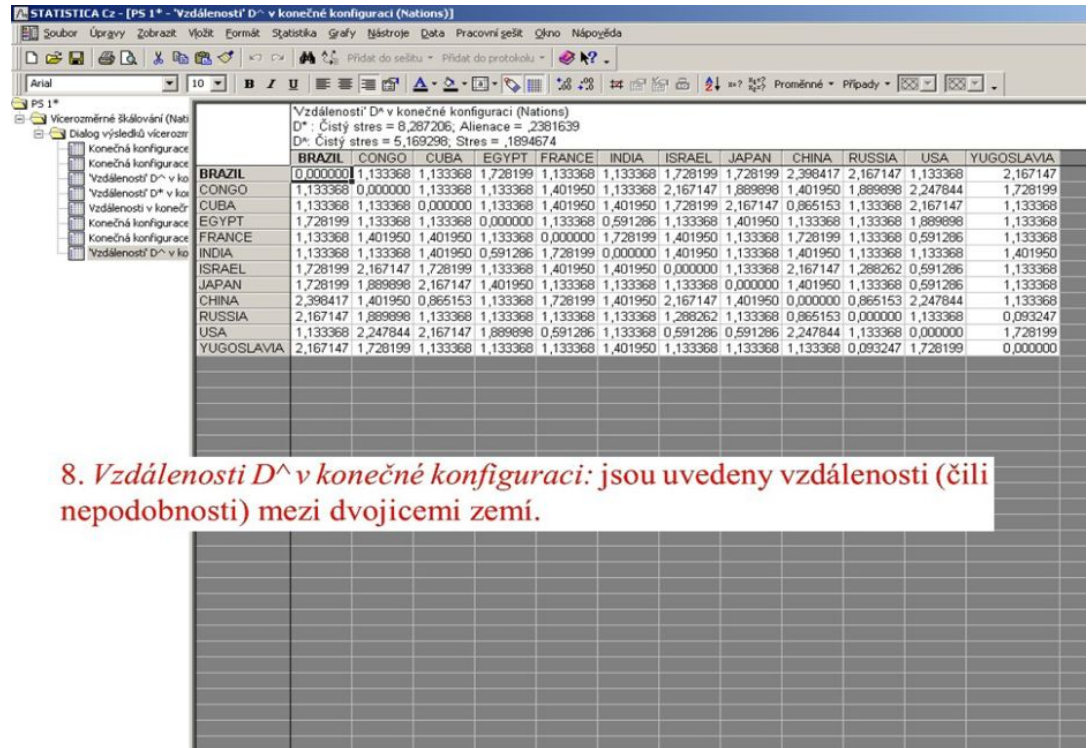
5. *Výsledky*: po kliknutí OK se otevře okno *Výsledky*: v horní polovině jsou dosažené výsledky optimalizačního výpočtu a v dolní polovině okna jsou tři záložky: *Základ*, *Detaily* a *Přehled a ukládání*. Kliknutím na jednotlivá okénka budeme nyní postupně otevírat výsledky.



6. *Konečná konfigurace*: obsahuje 4 sloupce výsledků. Sloupce $D^$ a D^* obsahují monotónní transformace dat. $D^$ jsou monotónní regresní odhady minimalizací účelové funkce podle Kruskala, zatímco D^* jsou pořadí vypočtené permutační procedurou vzestupných pořadí podle Guttmana, jež se snaží vyčíslit co nejtěsnější vypočtená pořadová čísla vůči zadaným. Řádky v této tabulce jsou seříděny podle velikosti $D^$ a D^* a značí jednu vzdálenost v matici podobnosti. Druhý sloupec *Vzdálenost* značí vzdálenosti z běžné konfigurace. Je-li zvolený počet rozměrů navrženého modelu vhodný, pak pořadí reprodukcí vzdáleností by mělo být stejné jako pro transformace $D^$ a D^* . Prvky mimo pořadí značí pak špatnou těsnost proložení. První sloupec tabulky ukazuje na prvky původních dat, kde x je řádek a y je sloupec. Například $D(2, 1)$ značí prvek druhého řádku a prvního sloupce vstupních dat čili porovnáni země Congo a Brazilie.



7. *Konečná konfigurace*: tabelární podoba grafu podobnosti objektů (zde země), ve které jsou uvedeny vypočtené souřadnice jednotlivých zemí.



8. *Vzdálenosti $D^$ v konečné konfiguraci*: jsou uvedeny vzdálenosti (čili nepodobnosti) mezi dvojicemi zemí.

STATISTICA Cz - [PS 1* - Vzdálenosti D* v konečné konfiguraci (Nations)]

D* : Čistý stres = 8,287206; Alienace = ,2381639
D* : Čistý stres = 5,169298; Stres = ,1894674

	BRAZIL	CONGO	CUBA	EGYPT	FRANCE	INDIA	ISRAEL	JAPAN	CHINA	RUSSIA	USA	YUGOSLAVIA
BRAZIL	0,000000	0,936324	0,849364	1,942526	0,969079	1,171765	1,680248	1,836027	2,398417	2,216016	0,842240	2,168776
CONGO	0,936324	0,000000	1,080029	0,907474	1,549258	0,954992	2,155133	1,973432	1,475790	1,989392	2,342291	1,911485
CUBA	0,849364	1,080029	0,000000	0,861000	1,461635	1,561636	1,754018	2,226501	0,786270	0,791982	2,202111	0,872578
EGYPT	1,942526	0,907474	0,861000	0,000000	0,963631	0,578455	0,994077	1,632614	1,312809	1,302817	2,133277	1,340578
FRANCE	0,969079	1,549258	1,461635	0,963631	0,000000	1,911779	1,555804	1,362908	1,714472	0,874488	0,462164	0,968716
INDIA	1,171765	0,954992	1,561636	0,578455	1,911779	0,000000	1,449315	1,276456	1,450855	1,212404	1,336820	1,576863
ISRAEL	1,680248	2,155133	1,754018	0,994077	1,555804	1,449315	0,000000	0,962378	2,220552	1,373663	0,481080	1,288262
JAPAN	1,836027	1,973432	2,226501	1,632614	1,362908	1,276456	0,962378	0,000000	1,415453	1,038965	0,456247	1,342864
CHINA	2,398417	1,475790	0,786270	1,312809	1,714472	1,450855	2,220552	1,415453	0,000000	0,671463	2,340554	0,886758
RUSSIA	2,216016	1,989392	0,791982	1,302817	0,874488	1,212404	1,373663	1,038965	0,671463	0,000000	0,933320	0,093247
USA	0,842240	2,342291	2,202111	2,133277	0,462164	1,336820	0,481080	0,456247	2,340554	0,933320	0,000000	1,817224
YUGOSLAVIA	2,168776	1,911485	0,872578	1,340578	0,968716	1,576863	1,288262	1,342864	0,886758	0,093247	1,817224	0,000000

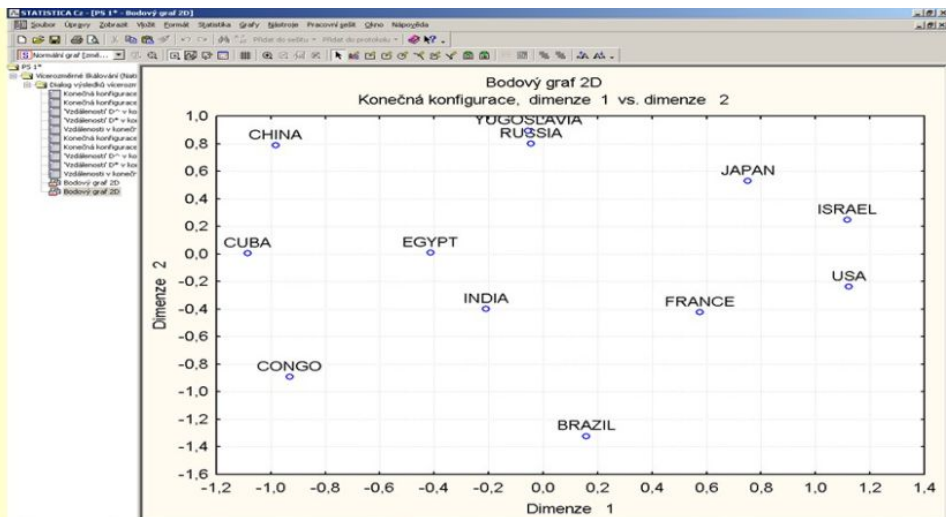
9. Vzdálenosti D* v konečné konfiguraci: jsou uvedena pořadí vzdáleností (čili nepodobností) mezi dvojicemi zemí.

STATISTICA Cz - [PS 1* - Vzdálenosti v konečné konfiguraci (Nations)]

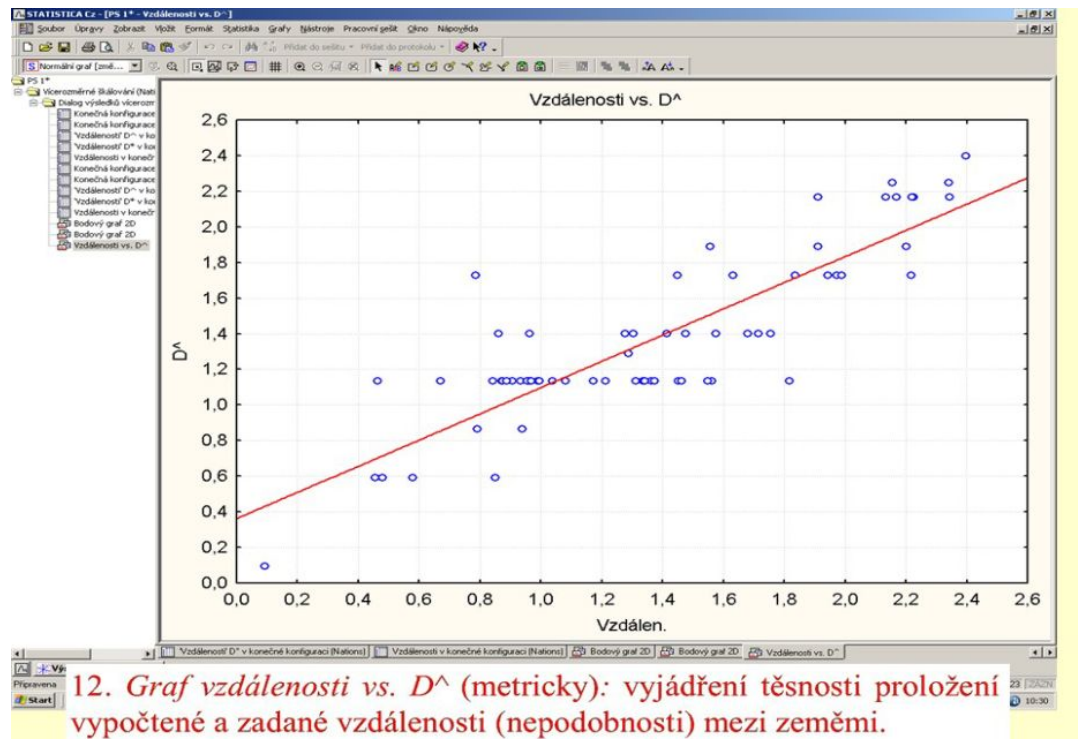
D* : Čistý stres = 8,287206; Alienace = ,2381639
D* : Čistý stres = 5,169298; Stres = ,1894674

	BRAZIL	CONGO	CUBA	EGYPT	FRANCE	INDIA	ISRAEL	JAPAN	CHINA	RUSSIA	USA	YUGOSLAVIA
BRAZIL	0,000000	1,171765	1,817224	1,449315	0,969079	0,994077	1,836027	1,942526	2,398417	2,133277	1,450855	2,226501
CONGO	1,171765	0,000000	0,907474	1,038965	1,576863	0,872578	2,342291	2,202111	1,680248	1,911779	2,155133	1,989392
CUBA	1,817224	0,907474	0,000000	0,671463	1,714472	0,962378	2,216016	1,911485	0,791982	1,312809	2,220552	1,362808
EGYPT	1,449315	1,038965	0,671463	0,000000	1,080029	0,456247	1,549258	1,276456	0,963631	0,874488	1,555804	0,954992
FRANCE	0,969079	1,576863	1,714472	1,080029	0,000000	0,786270	0,861000	0,968716	1,973432	1,373663	0,578455	1,461635
INDIA	0,994077	0,872578	0,962378	0,456247	0,786270	0,000000	1,475790	1,336820	1,475790	1,212404	1,342864	1,302817
ISRAEL	1,836027	2,342291	2,216016	1,549258	0,861000	1,475790	0,000000	0,462164	2,168776	1,288262	0,481080	1,340578
JAPAN	1,942526	2,202111	1,911485	1,276456	0,968716	1,336820	0,462164	0,000000	1,754018	0,842240	0,849364	0,886758
CHINA	2,398417	1,680248	0,791982	0,963631	1,973432	1,415453	2,168776	1,754018	0,000000	0,938324	2,340554	0,933320
RUSSIA	2,133277	1,911779	1,312809	0,874488	1,373663	1,212404	1,288262	0,842240	0,938324	0,000000	1,561636	0,093247
USA	1,450855	2,155133	2,220552	1,555804	0,578455	1,336820	0,481080	0,849364	2,340554	1,561636	0,000000	1,632614
YUGOSLAVIA	2,226501	1,989392	1,362808	0,954992	1,461635	1,302817	1,340578	0,886758	0,933320	0,093247	1,632614	0,000000

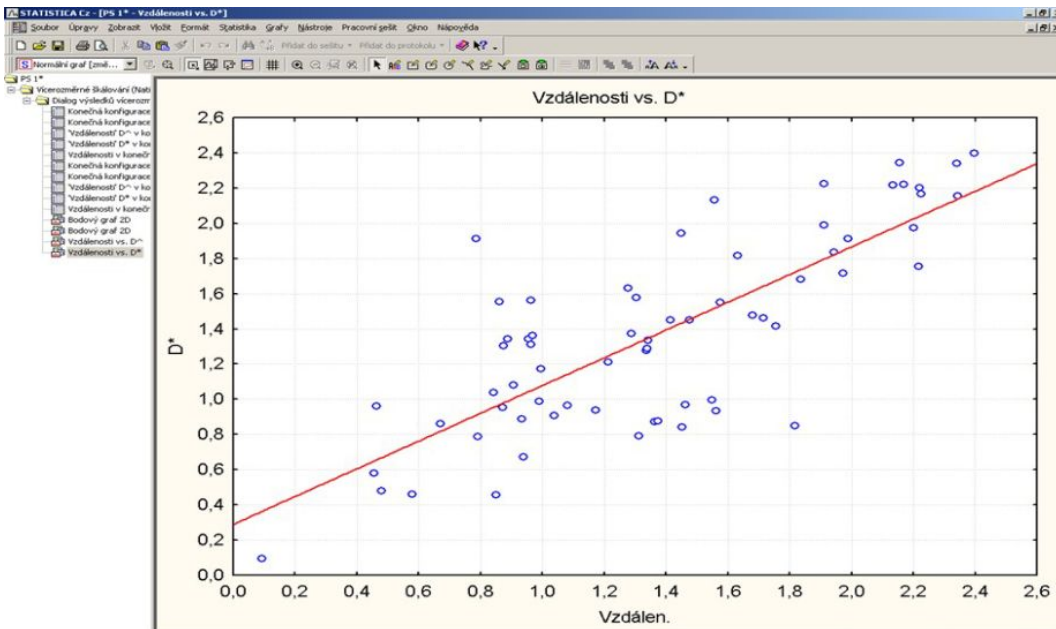
8. Vzdálenosti D v konečné konfiguraci: jsou uvedeny vzdálenosti (čili nepodobnosti) mezi dvojicemi zemí.



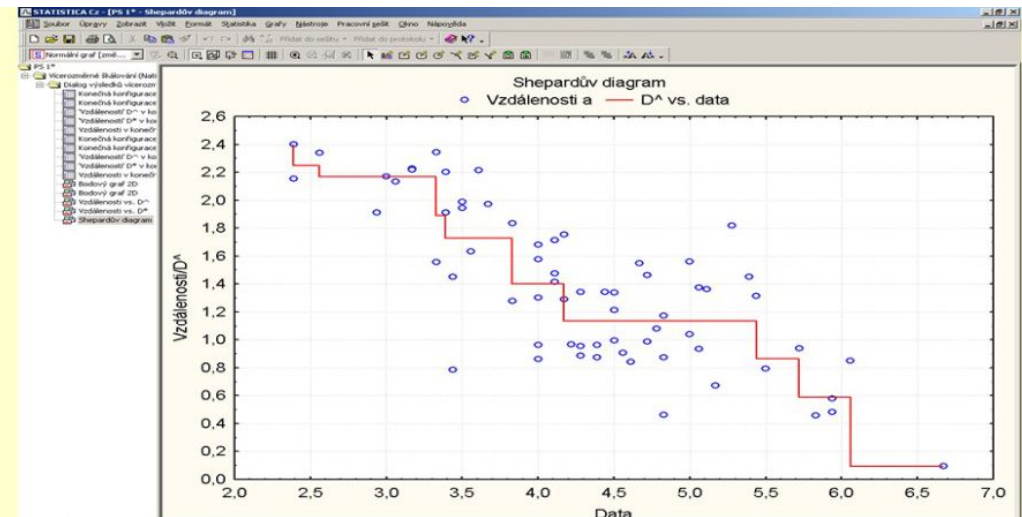
11. Bodový graf podobnosti objektů (zde zemí): souřadné osy grafu lze pojmenovat podobně jako ve faktorové analýze. Osy lze také mírně natočit, abychom dostali "faktorově čisté" objekty v ploše. Když otočíme osy o 45°, budou země lépe vysvětleny jako pro-západní, pro-komunistické, atd. Data byla vytvořena v roce 1970. Lze sledovat také vznik shluků.



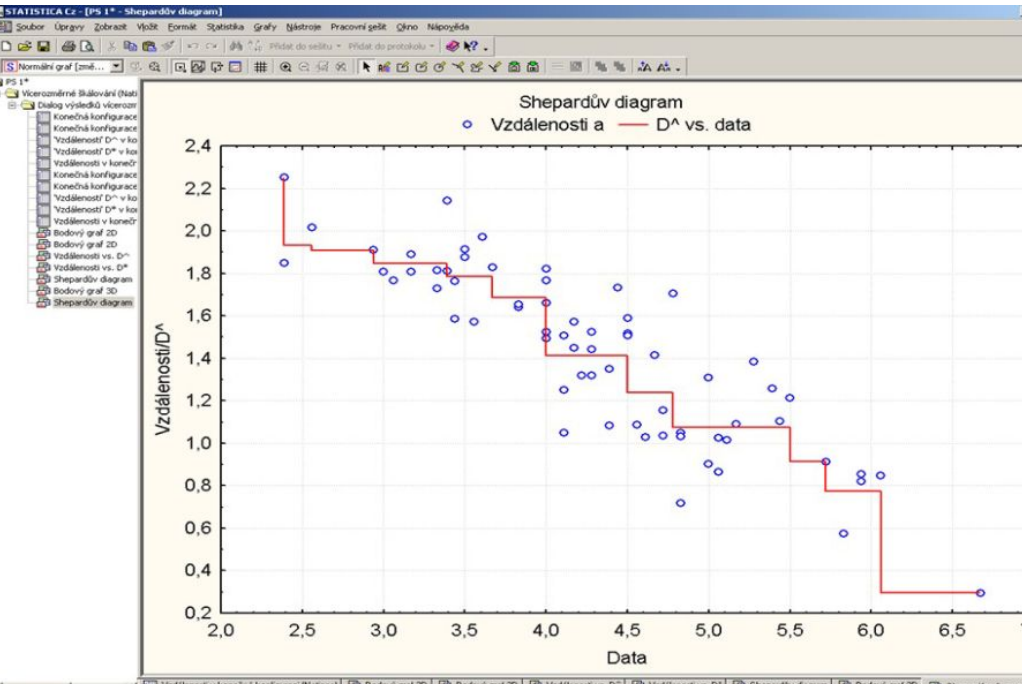
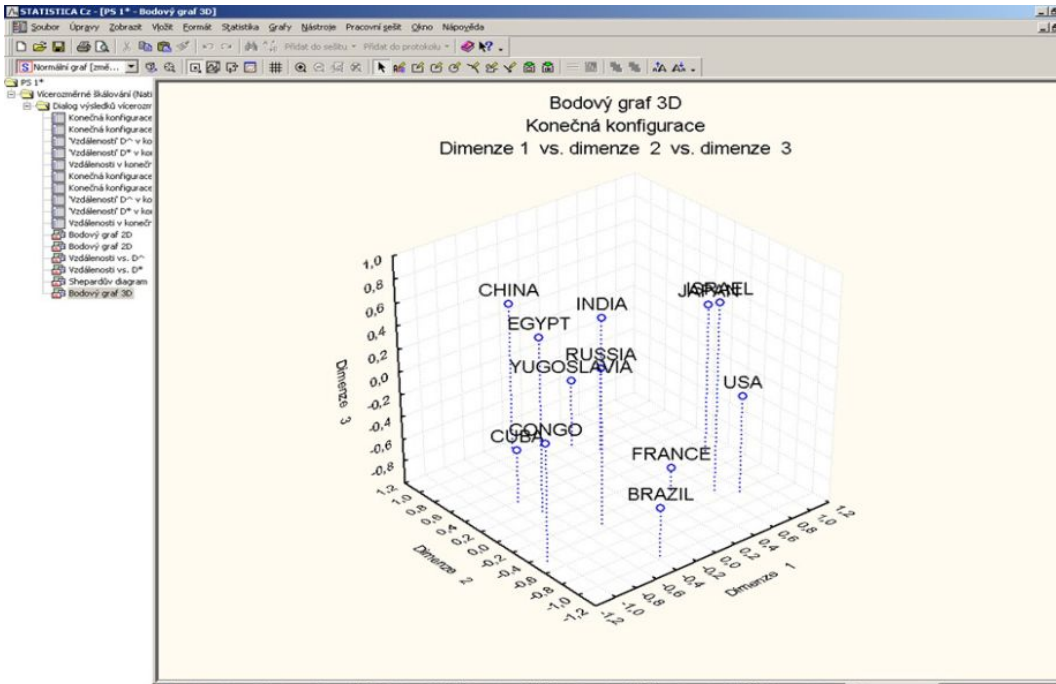
12. Graf vzdálenosti vs. D^ (metricky): vyjádření těsnosti proložení vypočtené a zadané vzdálenosti (nepodobnosti) mezi zeměmi.



13. Graf vzdálenosti vs. D^* (nemetricky): vyjádření těsnosti proložení pořadí vypočteného a zadaného vzdálenosti (nepodobnosti) mezi zeměmi.



14. Shepardův diagram: tento rozptylový graf vstupních dat podobnosti (nebo nepodobnosti) proti zadaným vzdálenostem z dat ukazuje D^* hodnoty čili monotónně transformovaná vstupní data jako schodovitou funkci. Většinu bodů je zde shluknuta kolem schodovité křivky. Proto lze uzavřít, že toto dvoj-rozměrné zobrazení je dostatečné k popisu podobnosti mezi zeměmi.



PŘÍKLAD 10.1 Porovnání podobnosti míčových sportů

Posuďte podobnost a vztah 6 míčových a míčkových sportů, je-li dána tabulka vzájemných podobností (vzdáleností) porovnáním “každého sportu s každým”.

Data: Prvky matice proximit *Sporty* vyjadřují nepodobnosti objektů u dvojic her:

Hokej, Fotbal, Basket, Tenis, Golf, Krokety. **Při podobnosti je 1, při nepodobnosti 6.**

Symetrická čtvercová matice proximit užije pouze horní trojúhelníkovou část.

	Hokej	Fotbal	Basket	Tenis	Golf	Krokety
Hokej	1	2	3	4	5	5
Fotbal	2	1	3	5	6	5
Basket	3	3	1	5	4	6
Tenis	4	5	5	1	4	3
Golf	5	6	4	4	1	2
Krokety	5	5	6	3	2	1

Řešení: NCSS2000, MINITAB a STATISTICA:

MDS vytvoří mapu relativních poloh jednoho sportu vůči druhému.

Řešení je dvojí:

metrické CMDS na základě vzdáleností mezi sporty,

nemetrické NNMDS na základě pořadových čísel seřazených vzdáleností.

2. Optimalizace výpočtu vzdáleností

- Iterační algoritmus minimalizuje účelovou funkci, hledá optimální sestavu vypočtených vzdáleností a vyčíslí **kritérium stress**.
- Při minimalizaci jsou pak vyčíslována rezidua, d_{vyp} (metricky) a d^*_{vyp} (nemetricky).
- Sleduje se jak MDS model prokládá data vzdáleností dané matice proximity.
- Proložení se sleduje kritériem těsnosti proložení **stress**.
- **Stress** je navíc funkcí počtu použitých souřadnic v CMDS modelu.

Vhodný počet použitelných souřadnic

Počet užitých souřadnic	Čtverec rezidui	Stress	Pseudo R2	
1	37.106	0.133	0.000	Počet vzdáleností = 15
2	6.948	0.025	70.730	Průměr vzdáleností = 4.13
3	2.413	0.009	89.830	Suma čtverců vzdáleností = 280.00
4	2.469	0.009	89.600	Suma čtv. vzd. okolo průměru = 23.73

Čtverec rezidui = rezidua jsou mezi vypočtenou a skutečnou vzdáleností sportů.

Stress = odmocnina ze čtverců rezidui dělených odmocninou sumy čtverců vzdáleností.

(Hodnoty pod 0.05 jsou přijatelné a pod 0.01 jsou dobré).

Pseudo R2 ukazuje na procento sumy čtverců vzdáleností pro tento počet souřadnic.

(Nad 80% je velmi nadějně).

Suma čtverců vzdáleností = hodnota užitá ve jmenovateli vzorce pro stress.

Suma čtverců vzdáleností okolo jejich průměru = ve jmenovateli vzorce **Pseudo R2**.

1. Nalezení počtu souřadnic MDS mapy

(určení počtu použitelných souřadnic, ve kterých se zobrazí porovnávané sporty).

- Každá souřadnice představuje rozličný základní faktor.
- Cílem je udržet počet souřadnic na nejmenším počtu (obvykle 2D-graf).
- Kumulativní procento proměnlivosti v datech je kritériem určení počtu souřadnic.
- Kritérium ukazuje, že hrana je u 3 souřadnic:

První dvě souřadnice pokrývají 77% a první tři souřadnice pokrývají 88% proměnlivosti.

(Výhodnější se jeví užít první 2 souřadnice).

Cattelův indexový graf úpatí vlastních čísel formou čarového diagramu

Pořadové číslo souřadnice	Vlastní číslo	Jednotlivé procento	Kumulativní procento	Čarový diagram vlastních čísel
1	30.73	54.28	54.28	
2 (Užívané)	12.85	22.69	76.97	
3	6.38	11.27	88.24	
4	1.68	2.97	91.21	
5	0.00	0.00	91.21	
6	-4.98	8.79	100.00	
Součet	56.62			

3. Těsnost proložení statistickou analýzou rezidui

Indikovány vzdálenosti v těch dvojicích míčových sportů, které nejsou modelem dobře proloženy.

Řádek	Sloupec	Skutečná vzdálenost	Vypočtená vzdálenost	Reziduum	Relativní reziduum, %
1 Hokej	2 Fotbal	2.000	0.823	1.177	58.830
5 Golf	6 Krokety	2.000	2.968	-0.968	-48.390
1 Hokej	3 Basket	3.000	2.773	0.227	7.570
2 Fotbal	3 Basket	3.000	3.259	-0.259	-8.630
4 Tenis	6 Krokety	3.000	1.387	1.613	53.780
1 Hokej	4 Tenis	4.000	3.477	0.523	13.080
4 Tenis	5 Golf	4.000	3.486	0.514	12.840
3 Basket	5 Golf	4.000	4.496	-0.496	-12.410
3 Basket	4 Tenis	5.000	4.981	0.019	0.380
1 Hokej	5 Golf	5.000	5.079	-0.079	-1.580
1 Hokej	6 Krokety	5.000	4.735	0.265	5.310
2 Fotbal	6 Krokety	5.000	5.419	-0.419	-8.380
2 Fotbal	4 Tenis	5.000	4.103	0.897	17.940
3 Basket	6 Krokety	6.000	5.766	0.234	3.900
2 Fotbal	5 Golf	6.000	5.902	0.098	1.630

Počet souřadnic = 2
Suma čtverců vzdáleností = 280.000
Suma čtverců rezidui = 6.948
Stress = 0.025
Pseudo R2 = 70.726

Skutečná vzdálenost = aktuální vzdálenost v datech.

Vypočtená vzdálenost = vypočtenou vzdálenost pro zvolený počet souřadnic.

Reziduum = rozdíl mezi skutečnou vzdáleností a vypočtenou.

Suma čtverců vzdáleností = jmenovatel ve vzorci pro stress.

Suma čtverců rezidui = čísel ve vzorci pro stress.

Stress = míra těsnosti proložení. (Pod 0.05 je přijatelné, pod 0.01 je ještě dobré).

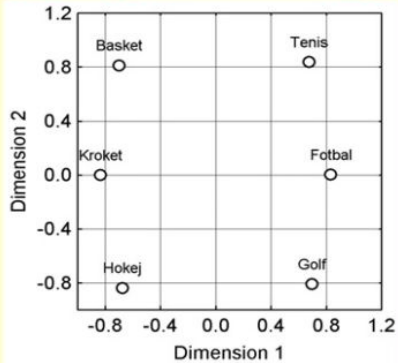
Pseudo R2 = obdoba koeficientu determinace R2 v regresi. Ukazuje na sumu čtverců vzdáleností, korigovanou na průměr a počítanou pro zvolený počet souřadnic. (Nad 80% značí velmi nadějně proložení).

4. Mapa objektů

Představuje: cíl celé MDS analýzy (v tabelární a grafické podobě).

Umožňuje: vysvětlit matici proximit obvykle ve dvojrozměrném rozptylovém diagramu. Protože jsou data škálována, je suma čtverců každého sloupce (každé souřadnice) rovna vlastnímu číslu této souřadnice.

	Souřadnice 1	Souřadnice 2	Souřadnice 3	Souřadnice 4
Hokej	1.9301	-0.6756	0.3818	1.0441
Fotbal	2.6179	-1.1281	-1.1303	-0.4680
Basket	2.1119	2.0914	0.4168	-0.4032
Tenis	-1.4786	-1.3608	1.8070	-0.3940
Golf	-2.3836	2.0059	-0.2743	0.2351
Kroket	-2.7976	-0.9328	-1.2011	-0.014



Obr. 10.4a Dvojměrný škálovací **CMDS diagram** podobnosti 6 sportů, **stress = 0.100**, (STATISTICA).

6. Závěry o metodě MDS:

- Ve škálovacím diagramu CMDS **není pevná orientace** souřadných os.
- Osy lze **otáčet okolo počátku** k dosažení názorné polohy a separované shluky bodů.
- Respondenti v anketě považují **hokej a fotbal** za blízké sporty.
- Respondenti považují **kroket a tenis** za blízké míčkové hry.
- **Fotbal** je však považován za zcela odlišný **od golfu**.
- Z umístění objektů v MDS grafu plyne **podobnost** porovnávaných míčových her.
- Hlavní **rozdíl NNMDS mapy vůči metrické CMDS** mapě spočívá:
 - a) golf a kroket jsou nyní blíže sobě.
 - b) fotbal a basketbal jsou těsněji u sebe, (zde se ale nejeví NNMDS mapa správnější než CMDS mapa).
 - c) I když NNMDS dále zjemňuje CMDS mapu, nebylo dosaženo lepší těsnosti proložení kritéria stress.
- Vedle nalezení podobnosti ve druzích sportu je cílem MDS pojmenovat souřadnice:
 - vertikální y-nová osa rozděluje shora dolů týmové sporty,
 - horizontální x-ová osa spíše individuální sporty od nuly směrem doleva.

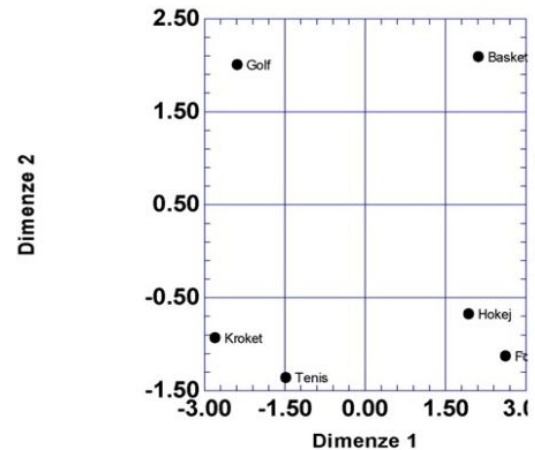
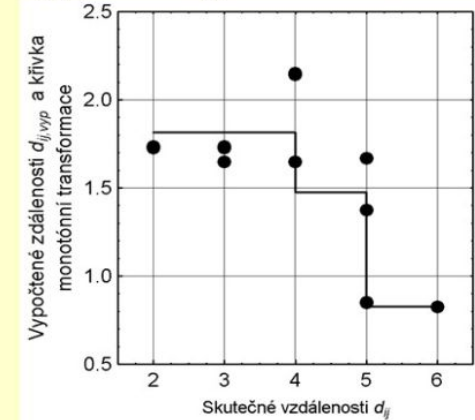
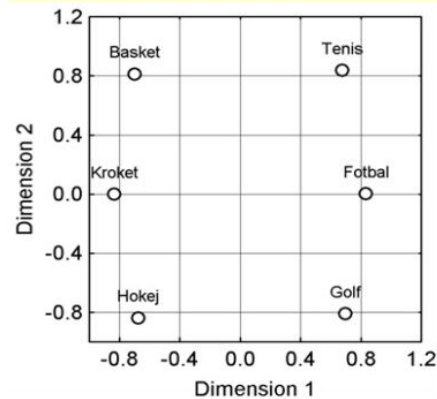
5. Shepardův rozptylový diagram

Zobrazuje: vypočtené vzdálenosti na y-nové ose proti skutečným podobnostem (opak vzdáleností) na x-ové ose a proto je křivka sestupná.

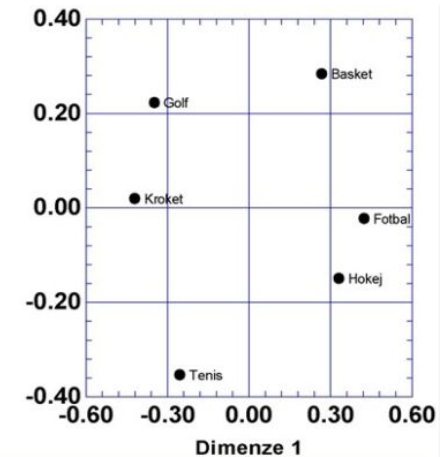
Představuje: křivka představuje spojitou monotónní transformaci $d_{ij,vyp} = f(d_{ij})$ hodnot.

Těsnost proložení: body těsně u křivky značí dobrý MDS model, body od křivky vzdálené nedostatečné proložení.

Pravidlo: je-li těsnost proložení při správně zvoleném počtu souřadnic velmi dobrá, jsou vypočtené vzdálenosti d_{vyp} (metricky) a d^*_{vyp} (nemetricky) v dobré shodě.

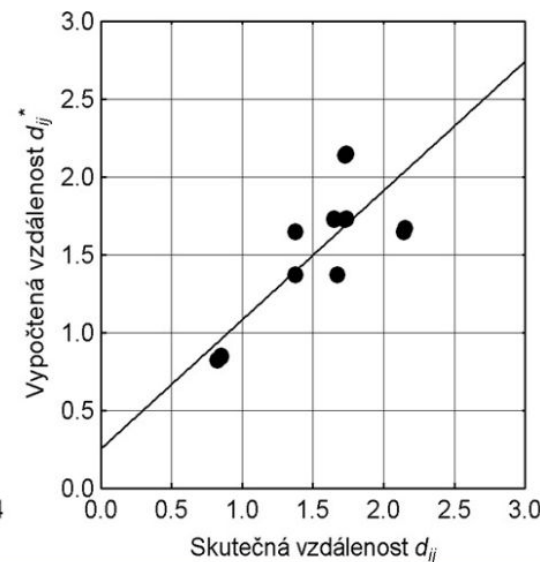
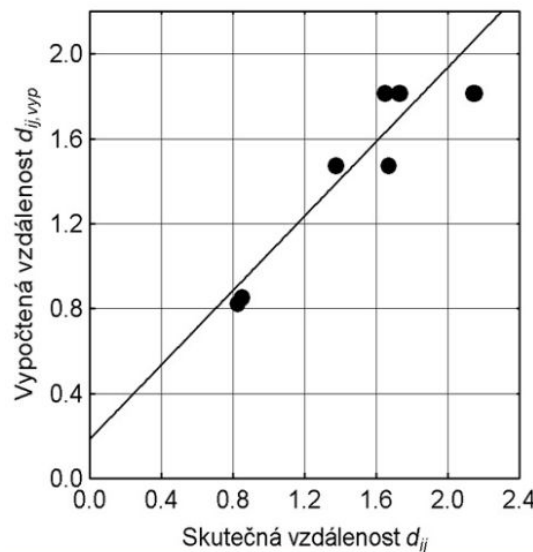
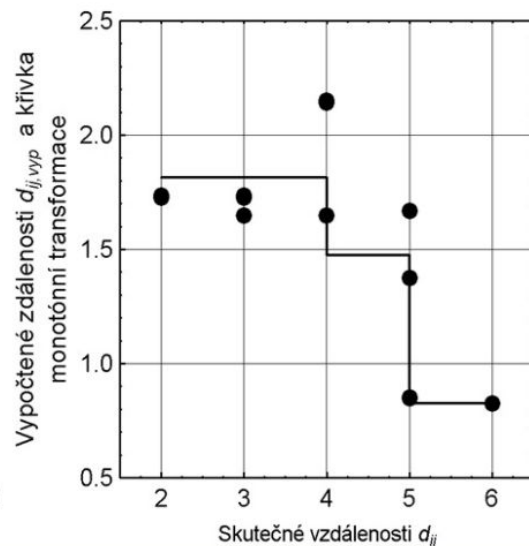
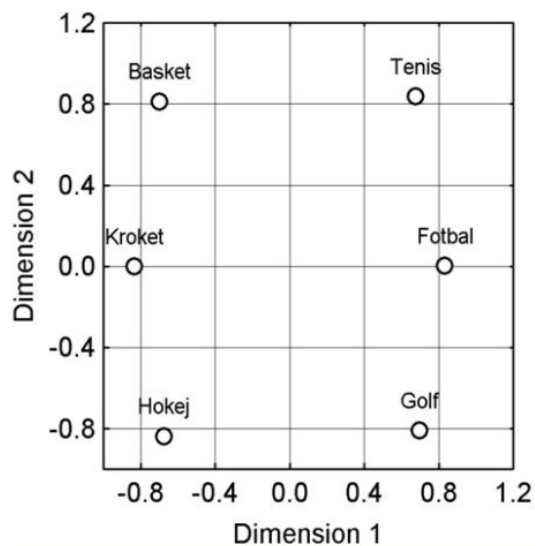


Obr. 10.4e Dvojměrný škálovací **CMDS diagram** podobnosti 6 sportů, **stress = 0.025**, (NCSS2000).



Obr. 10.4f Dvojměrný škálovací **NNMDS diagram** podobnosti 6 sportů, **stress = 0.052**, (NCSS2000).

Závěr: Metrická metoda vícerozměrného **škálování CMDS** dosáhla nejnižší hodnoty kritéria stress, a tím pádem i nejlepšího separování na škálovací mapě podobnosti šesti míčových sportů.



Obr. 10.4a Dvojmrozměrný škálovací CMDS diagram podobnosti 6 sportů, $stress = 0.100$, (STATISTICA).

Obr. 10.4b Shepardův diagram CMDS těsnosti proložení podobnosti 6 sportů, $stress = 0.100$, (STATISTICA).

Obr. 10.4c Těsnost proložení vzdáleností $d_{ij,vyp}$ mezi 6 sporty ve CMDS diagramu, $stress = 0.361$, (STATISTICA).

Obr. 10.4d Těsnost proložení vzdáleností d_{ij}^* mezi 6 sporty v NNMDS diagramu, $stress = 0.993$, (STATISTICA).

STATISTICA Cz - [Data: 101Sporty (6s krát 10ř)]

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Okno Nápoředa

Průběh: Arial 9 B I U

	1	2	3	4	5	6
	okej	otbal	asket	enis	iof	roket
Hokej	0	2	3	4	5	5
Fotbal	2	0	3	5	6	5
Basket	3	3	0	5	4	6
Tenis	4	5	5	0	4	3
Golf	5	6	4	4	0	2
Krokety	5	5	6	3	2	0
Means	0	0	0	0	0	0
StandDeviation	0	0	0	0	0	0
No. Cases	15					
Matic	2					

1. Zdrojová matice 101Sporty se načte z databáze EXAMPLES programu STATISTICA. Data jsou tvořena podobnostní maticí (čili maticí proximity), která obsahuje odhadnuté podobnosti sportů.

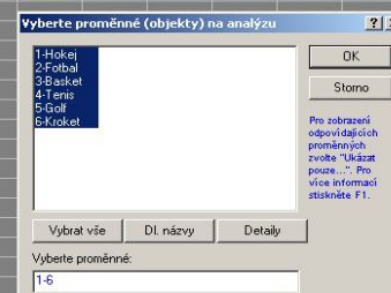
STATISTICA Cz - [Data: 101Sporty (6s krát 10ř)]

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Okno Nápoředa

Průběh: Arial 9 B I U

	1	2	3	4	5	6
	okej	otbal	asket	enis	iof	roket
Hokej	0	2	3	4	5	5
Fotbal	2	0	3	5	6	5
Basket	3	3	0	5	4	6
Tenis	4	5	5	0	4	3
Golf	5	6	4	4	0	2
Krokety	5	5	6	3	2	0
Means	0	0	0	0	0	0
StandDeviation	0	0	0	0	0	0
No. Cases	15					
Matic	2					

2. Zavolá se Vícerozměrné škálování ze Vícerozměrné průzkumné techniky.



STATISTICA Cz - [Data: 101Sporty (6s krát 10P)]

Úpravy Zobrazení Vložit Formát Statistika Grafy Nástroje Data Okno Nápořádání

Přidat do sešitu Přidat do protokolu

1	2	3	4	5	6
okej	otbal	asket	enis	loft	roket
0	2	3	4	5	5
2	0	3	5	6	5
3	3	0	5	4	6
4	5	5	0	4	3
5	6	4	4	0	2
5	5	6	3	2	0
0	0	0	0	0	0
0	0	0	0	0	0
15					
2					

Odhadování parametrů: 101Sporty

iter.	[dim=3]	D*	D*	D*	d*	
s: c:	cosin	krok	čistý stres	alienace	čistý stres	stres
20	1	.167	.0022953	.0079851		
20	1		.0022977	.0079890		
21	1	.167	.0022972	.0079885		
21	1		.0022987	.0079908		
17	0				.0007087	.0044370
18	1	.167	.0006702	.0043145		
19	1	.973	.599	.0006130	.0041265	
20	1	.472	.295	.0006003	.0040836	
21	1	.902	.557	.0005883	.0040424	
22	1	.966	.870	.0005803	.0040148	
23	1	.925	.865	.0005777	.0040061	
24	1	.722	.486	.0005771	.0040037	
25	1	.683	.371	.0005767	.0040025	
26	1	.937	.680	.0005763	.0040009	
27	1	.974	.957	.0005758	.0039994	
28	1	.951	.983	.0005754	.0039980	
29	1	.924	.897	.0005751	.0039969	
30	1	.798	.591	.0005749	.0039962	
30	*		.0021714	.0077664	.0005749	.0039962

Procedura odhadování konvergovala

Storno OK

1	2	3	4	5	6
okej	otbal	asket	enis	loft	roket
0	2	3	4	5	5
2	0	3	5	6	5
3	3	0	5	4	6
4	5	5	0	4	3
5	6	4	4	0	2
5	5	6	3	2	0
0	0	0	0	0	0
0	0	0	0	0	0
15					
2					

Hokej 0 2 3 4 5 5
 Fotbal 2 0 3 5 6 5
 Basket 3 3 0 5 4 6
 Tenis 4 5 5 0 4 3
 Golf 5 6 4 4 0 2
 Krokret 5 5 6 3 2 0
 Means 0 0 0 0 0 0
 StandDeviation 0 0 0 0 0 0
 No.Cases 15
 Matice 2

Výsledky: 101Sporty

6 prom. ze souboru
 Celk. počet dimenzi : 3
 Počet.konfig.: (Guttman-Lingoesova)
 Posl. počítaná iterace : 85; Nejlepší iterace: 30
 D* : Čistý stres = .0021714; Alienace = .0077664
 D^: Čistý stres = .0005749; Stres = .0039962

Základ | Detaily | Přehled & ukládání

Shnutí: Konečná konfig.
 Graf konečné konfig., 2D
 Graf konečné konfig., 3D
 Šperardův diagram

Shnutí Storno Možnosti

5. Výsledky: po kliknutí OK se otevře okno *Výsledky:* v horní polovině jsou dosažené výsledky optimalizačního výpočtu a v dolní polovině okna

1	2	3	4	5	6
okej	otbal	asket	enis	loft	roket
0	2	3	4	5	5
2	0	3	5	6	5
3	3	0	5	4	6
4	5	5	0	4	3
5	6	4	4	0	2
5	5	6	3	2	0
0	0	0	0	0	0
0	0	0	0	0	0
15					
2					

Výsledky: 101Sporty

6 prom. ze souboru
 Celk. počet dimenzi : 3
 Počet.konfig.: (Guttman-Lingoesova)
 Posl. počítaná iterace : 85; Nejlepší iterace: 30
 D* : Čistý stres = .0021714; Alienace = .0077664
 D^: Čistý stres = .0005749; Stres = .0039962

Základ | Detaily | Přehled & ukládání

Shnutí: Konečná konfig. Graf konečné konfig., 2D
 Hodnoty D^ Graf konečné konfig., 3D
 Hodnoty D* Graf D^ vs. vzdálenosti
 Matice vzdáleností Graf D* vs. vzdálenosti
 Souhrnné statistiky Šperardův diagram

Shnutí Storno Možnosti

STATISTICA Cz - [PS 1* - Konečná konfigurace (101Sporty)]

Soubor Úpravy Zobrazení Vložit Formát Statistika Grafy Nástroje Data Pracovní sešit Okno Nápořádání

Přidat do sešitu Přidat do protokolu

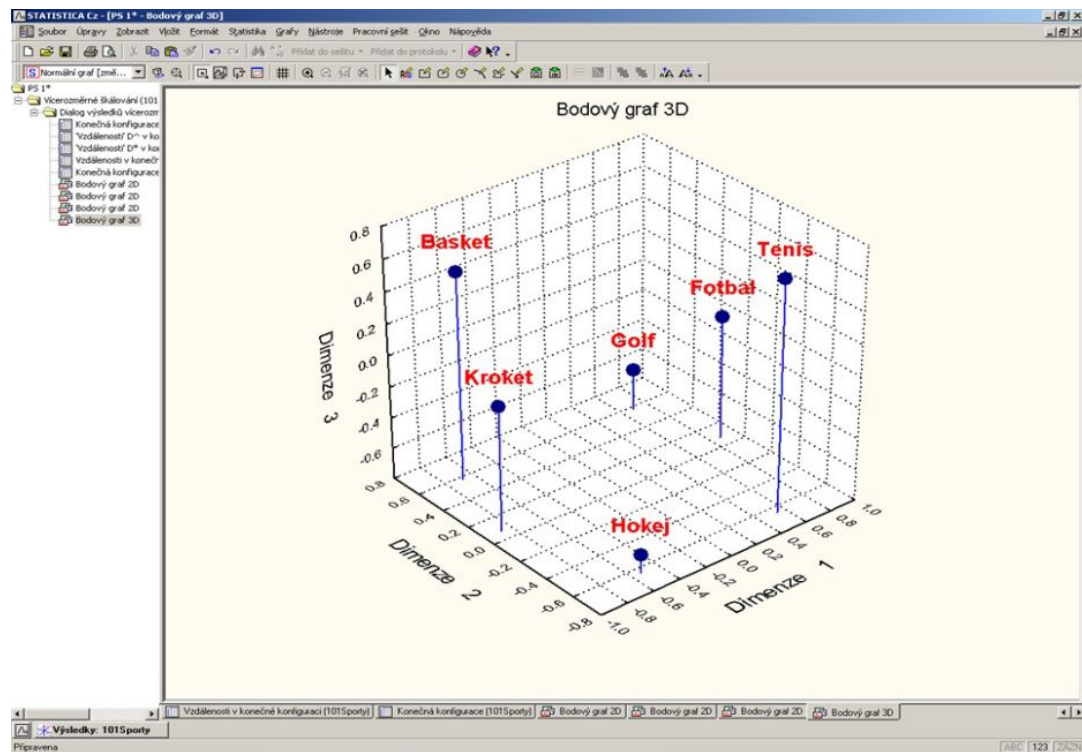
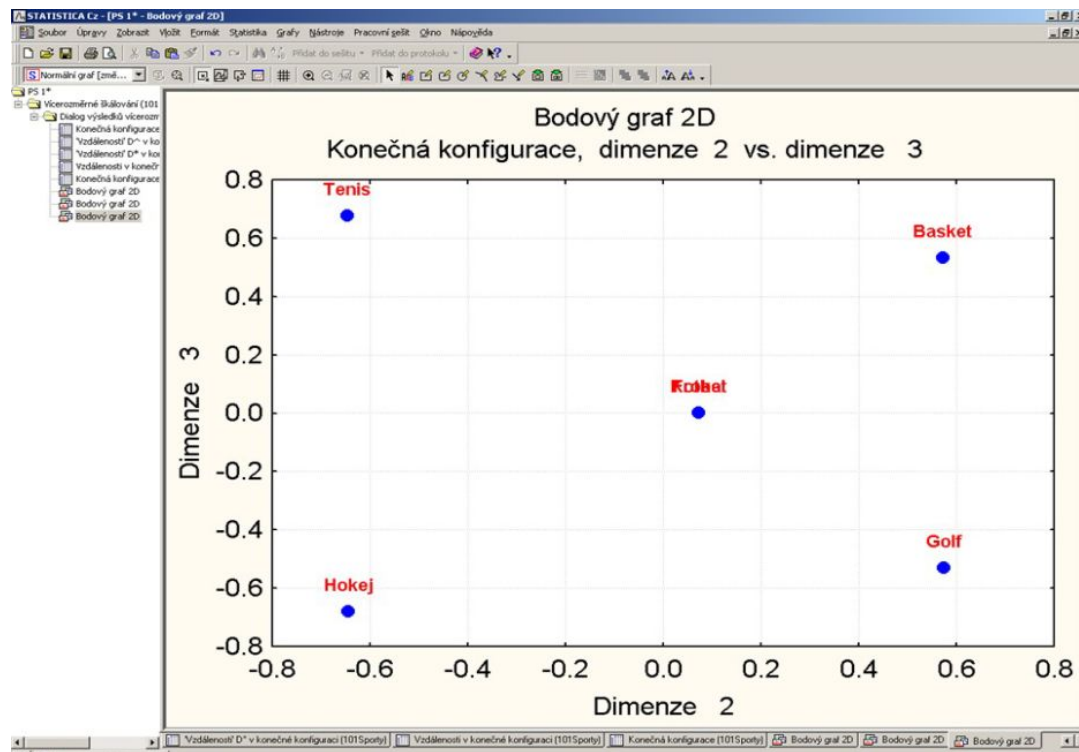
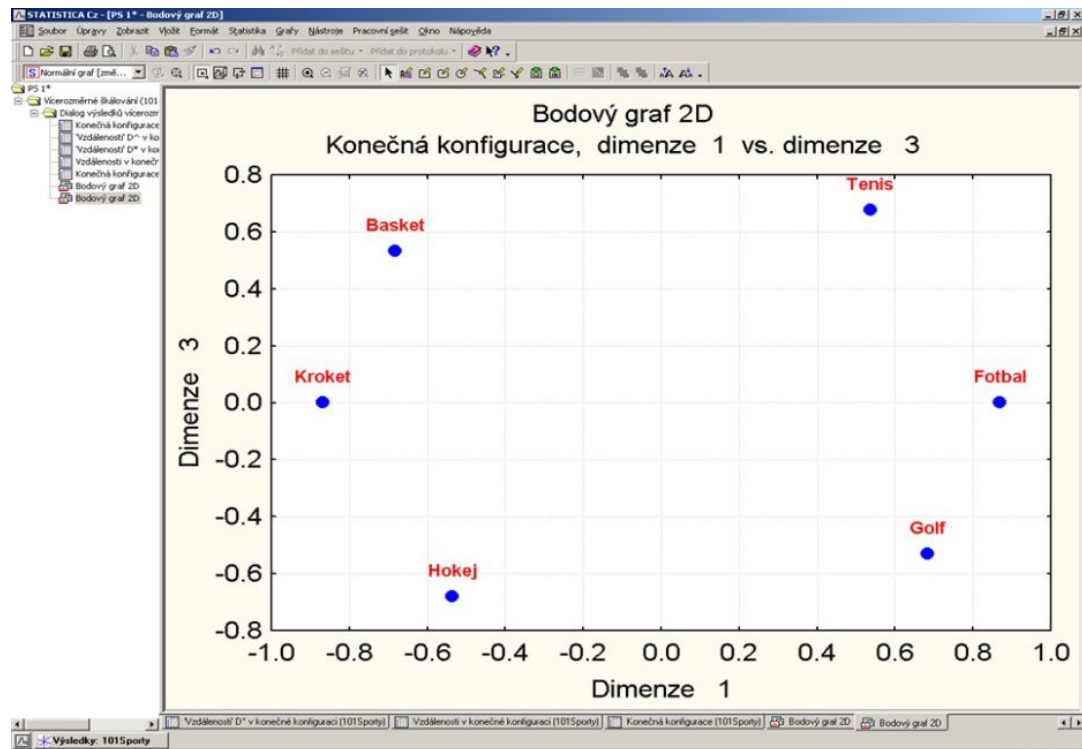
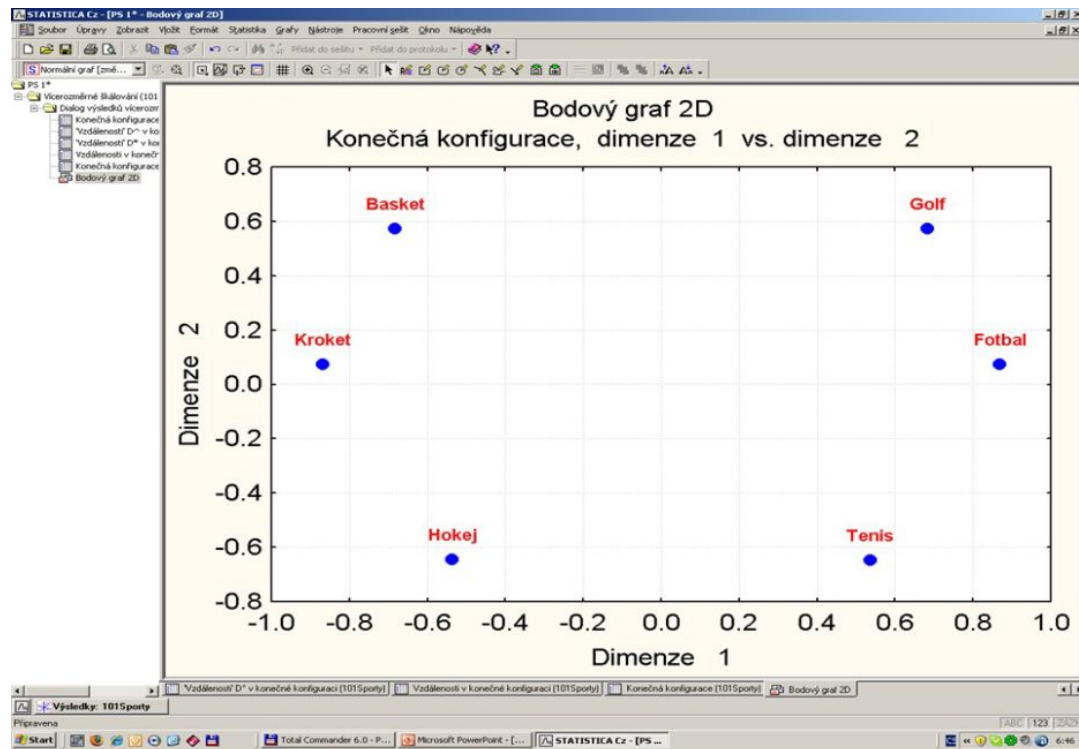
Arial 10 B I U

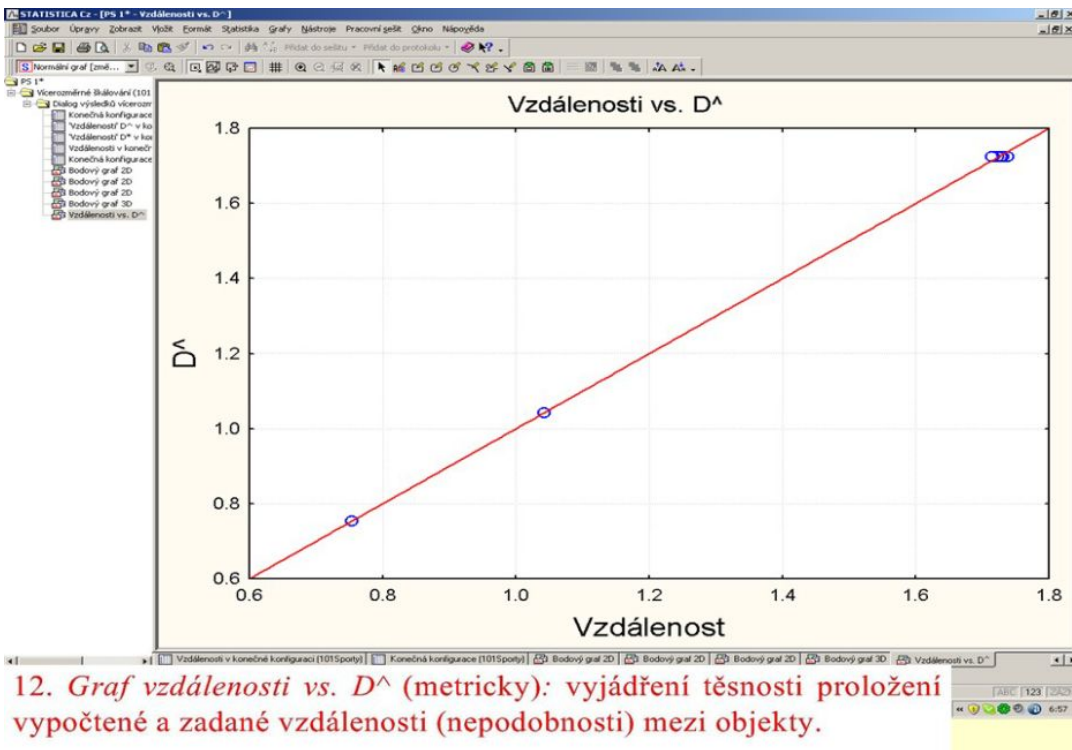
PS 1*
 Vícezměrné škálování (101
 Dialog výsledků vícezmr
 Konečná konfigurace

Konečná konfigurace (101Sporty)
 D* : Čistý stres = .0021714; Alienace = .0077664
 D^: Čistý stres = .0005749; Stres = .0039962

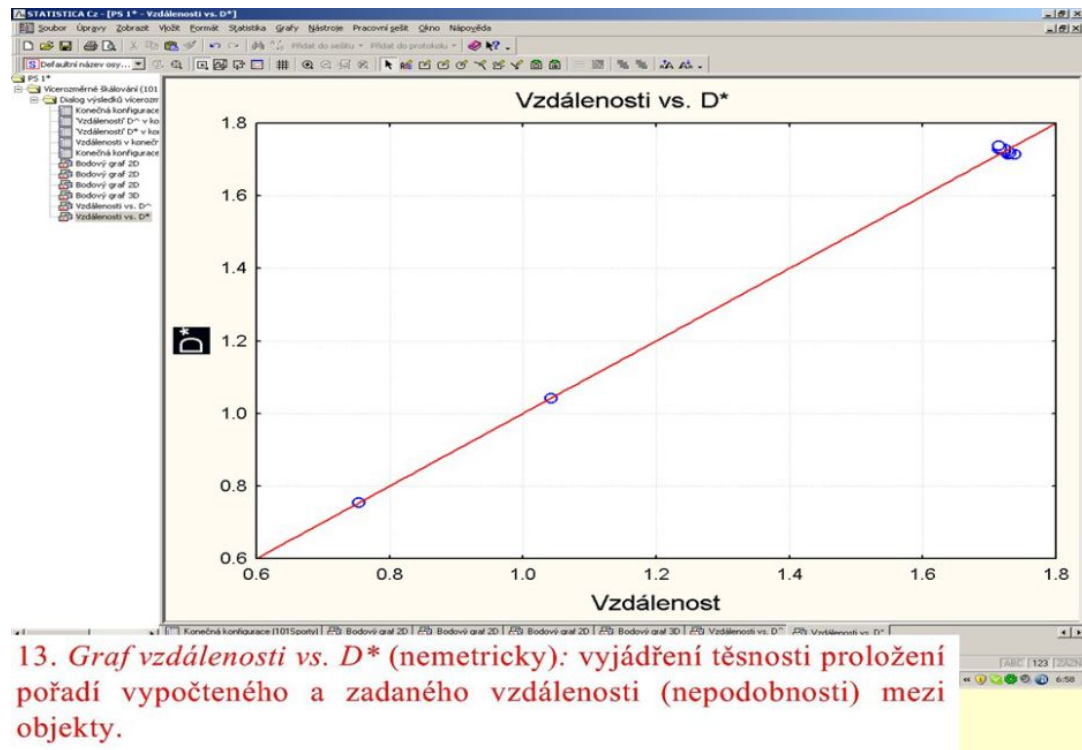
	DIM. 1	DIM. 2	DIM. 3
Hokej	-0.535557	-0.644179	-0.678998
Fotbal	0.869030	0.072771	0.000148
Basket	-0.682309	0.571738	0.533084
Tenis	0.535608	-0.646842	0.676478
Golf	0.682282	0.573812	-0.530889
Krokret	-0.869054	0.072700	0.000177

7. Konečná konfigurace: tabulární podoba grafu podobností objektů, ve které jsou uvedeny vypočtené souřadnice jednotlivých objektů.

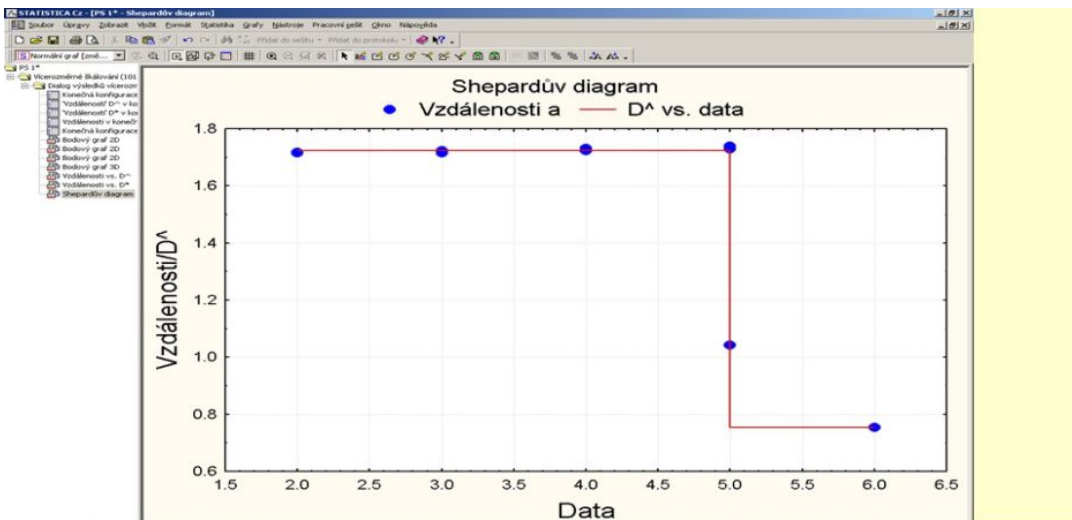




12. Graf vzdálenosti vs. $D^$ (metricky): vyjádření těsnosti proložení vypočtené a zadané vzdálenosti (nepodobnosti) mezi objekty.



13. Graf vzdálenosti vs. D^* (nemetricky): vyjádření těsnosti proložení pořadí vypočteného a zadaného vzdálenosti (nepodobnosti) mezi objekty.



14. Shepardův diagram: tento rozptylový graf vstupních dat podobnosti (nebo nepodobnosti) proti zadaným vzdálenostem z dat ukazuje $D^$ hodnoty čili monotónně transformovaná vstupní data jako schodovitou funkci. Většinu bodů je zde shluknuta kolem schodovité křivky. Proto lze uzavřít, že toto dvoj-rozměrné zobrazení je dostatečné k popisu podobnosti mezi objekty.

STATISTICA Cz - [PS 1* - Konečná konfigurace (101Sporty)]

Konečná konfigurace (101Sporty)
 D^* : Čistý stres = .0021714; Alienace = .0077664
 $D^$: Čistý stres = .0005749; Stres = .0039962

	DIM. 1	DIM. 2	DIM. 3
Hokej	-0.535557	-0.644179	-0.678998
Fotbal	0.869030	0.072771	0.000148
Basket	-0.682309	0.571738	0.533084
Tenis	0.535608	-0.646842	0.676478
Golf	0.682282	0.573812	-0.530889
Krokiet	-0.869054	0.072700	0.000177

PŘÍKLAD 10.2 Podobnost 10 výrobků coca-coly

Je třeba vyhodnotit subjektivní mínění spotřebitelů: anketou posoudit podobnost 10 výrobků coca-coly A až J , a to technikou CMDS a NNMDS.

Data: Data *Cola* obsahují vzájemné porovnání 10 výrobků coca-coly A až J způsobem „každý objekt s každým“ 50 respondenty. Při dokonalé podobnosti byla dvojici přidělena vzdálenost 0, zatímco při naprosté nepodobnosti vzdálenost 100. Z hodnot od 50 respondentů byla vyčíslena střední hodnota a zapsána do čtvercové matice. Užije se pouze trojúhelníková část.

	A	B	C	D	E	F	G	H	I	J
A	0	20	75	60	80	55	80	45	87	12
B	20	0	35	31	70	40	90	80	35	90
C	75	35	0	80	37	20	90	77	50	96
D	60	31	80	0	70	89	55	75	88	89
E	80	70	37	70	0	30	87	25	60	75
F	55	40	20	89	30	0	88	86	10	40
G	80	90	90	55	87	88	0	40	98	27
H	45	80	77	75	25	86	40	0	83	14
I	87	35	50	88	60	10	98	83	0	90
J	12	90	96	89	75	40	27	14	90	0

V MDS mapě objektů neexistuje orientace diagramu, diagramem je totiž možné libovolně otáčet okolo počátku. Důležité jsou relativní polohy objektů vůči sobě a hlavně polohy shluků.

Jednotlivé druhy coca-coly A až J jsou zřetelně rozříděny v rovině. Naprosto stejné nápoje budou mít mezi sebou vzdálenost 0, odlišné 100.

	1 A	2 B	3 C	4 D	5 E	6 F	7 G	8 H	9 I	10 J
	0	20	75	60	80	55	80	45	87	12
	20	0	35	31	70	40	90	80	35	90
	75	35	0	80	37	20	90	77	50	96
	60	31	80	0	70	89	55	75	88	89
	80	70	37	70	0	30	87	25	60	75
	55	40	20	89	30	0	88	86	10	40
	80	90	90	55	87	88	0	40	98	27
	45	80	77	75	25	86	40	0	83	14
	87	35	50	88	60	10	98	83	0	90
	12	90	96	89	75	40	27	14	90	0
Means	0	0	0	0	0	0	0	0	0	0
Std. Dev.	0	0	0	0	0	0	0	0	0	0
No. Cases	45									
Matice	2									

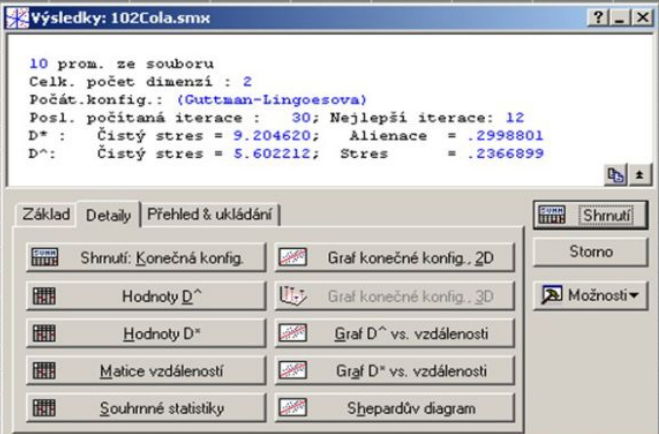
2. Zavolá se Více rozměrné škálování ze Více rozměrné průzkumné techniky.

3. Kliknutím na okénko *Proměnné* se z matice proximity zadají podobnosti všech sportů (sloupce), a to v okně *Vyberte proměnné (objekty) na analýzu* a pak OK. Za počet dimenzí se dosadí 2.

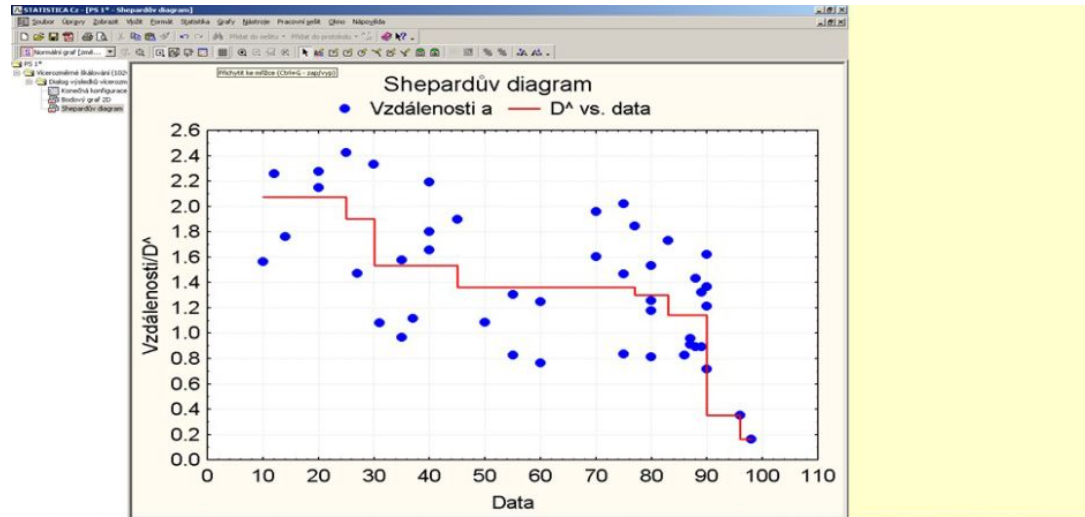
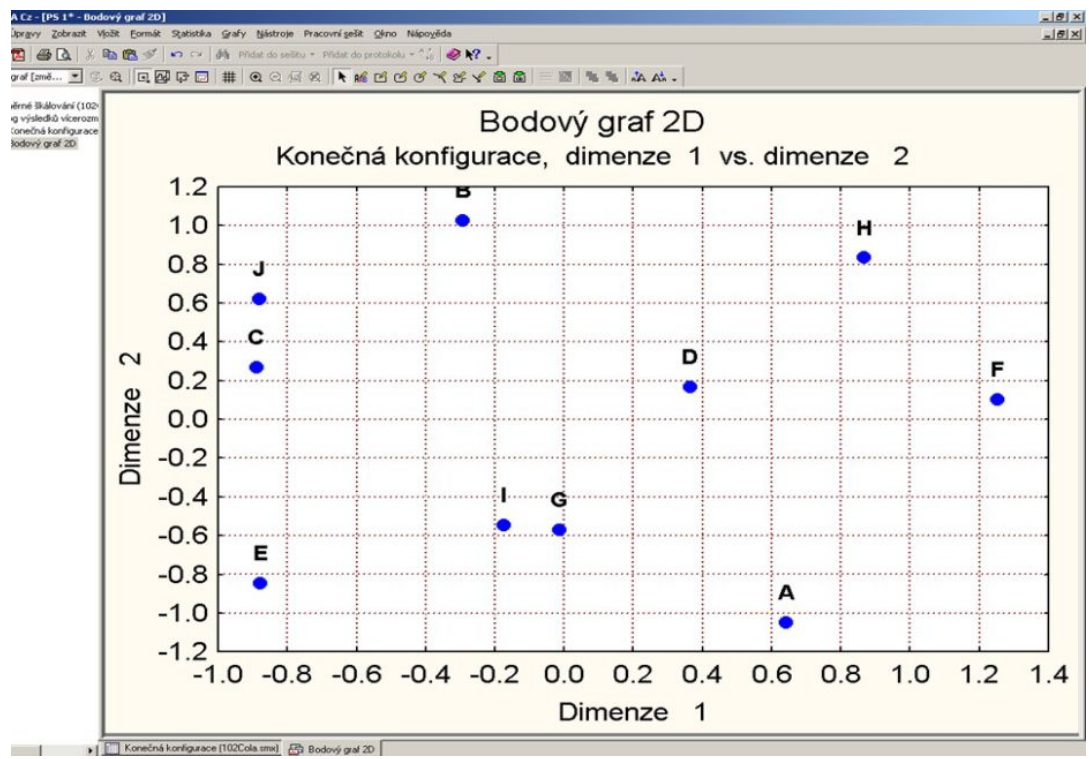
iter.	[dim=2]	D*	D*	D*	d*
s: t:	cosin	krok	čistý stres	alienace	čistý stres
5 2	.100	8.909040	.3127387		
5 2		9.180164	.2994908		
6 1	.100	8.965996	.3138350		
6 2	.100	8.963082	.3137785		
6 2		9.218696	.3001039		
7 1	.100	9.002570	.3145378		
7 2	.100	8.999480	.3144780		
7 2		9.227455	.3002431		
3 0				6.016556	.2452867
4 1	.106			5.851666	.2419022
5 1	.973	.381		5.688605	.2385080
6 1	-.021	.162		5.665922	.2380320
7 1	.783	.238		5.648557	.2376669
8 1	.962	.477		5.626441	.2372012
9 1	.917	.509		5.613093	.2369197
10 1	.844	.411		5.606737	.2367855
11 1	.455	.190		5.603909	.2367258
12 1	.587	.170		5.602212	.2366899
12 *			9.204620	.2998801	5.602212

4. *Odhadování parametrů*: program začne výpočet přes PCA s maticí podobnosti a pak se otevře okno *Odhadování parametrů*: po počtu iterací metody největšího spádu značeném s následuje ještě několik (obvykle 5) dalších zjemňujících iterací a jejich počet je ve sloupci t . Jsou uvedeny číselné hodnoty veličin těsnosti proložení *stress*, D^* (metricky) a D^* (nemetricky).

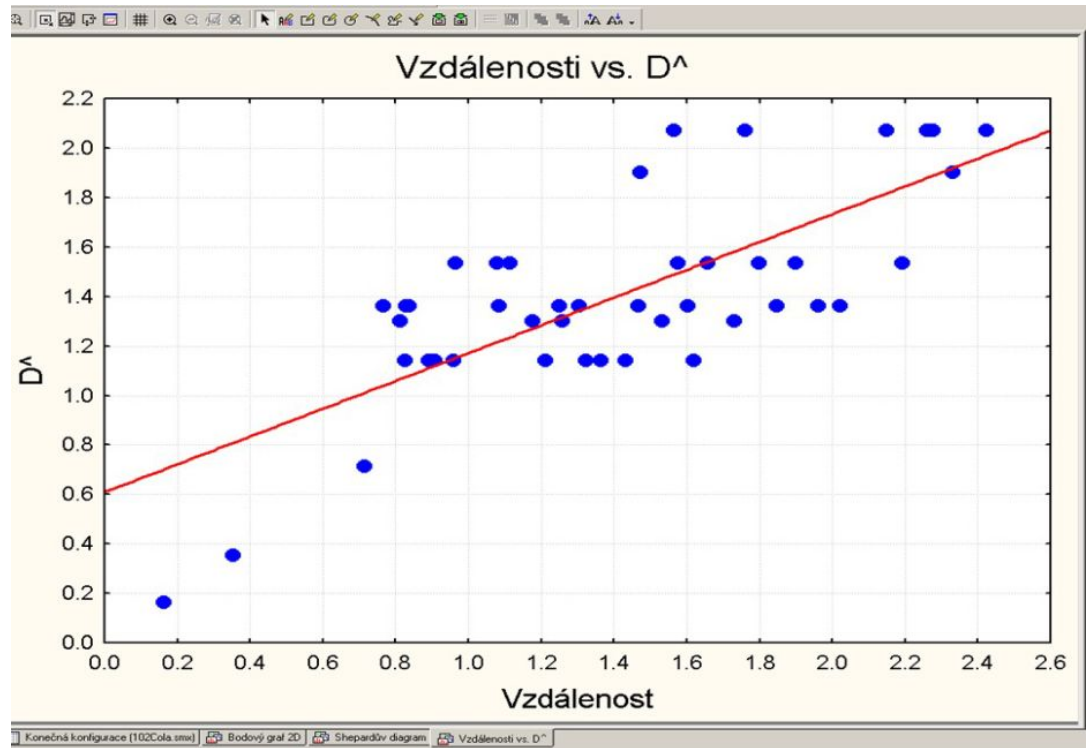
15	10	40
0	98	27
0	83	14
3	0	90
4	90	0
0	0	0
0	0	0

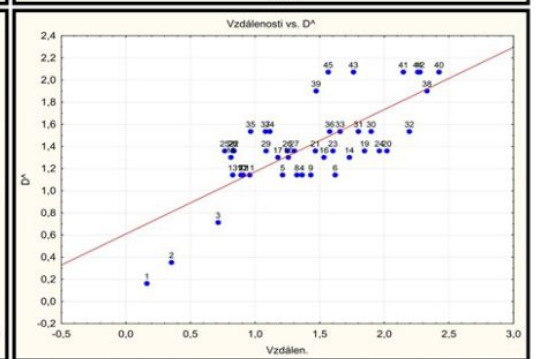
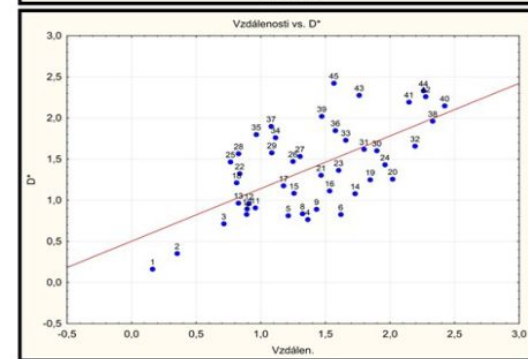
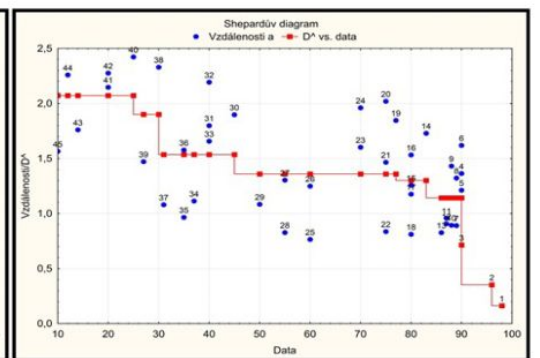
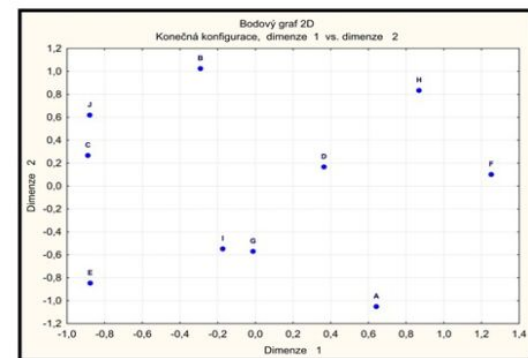
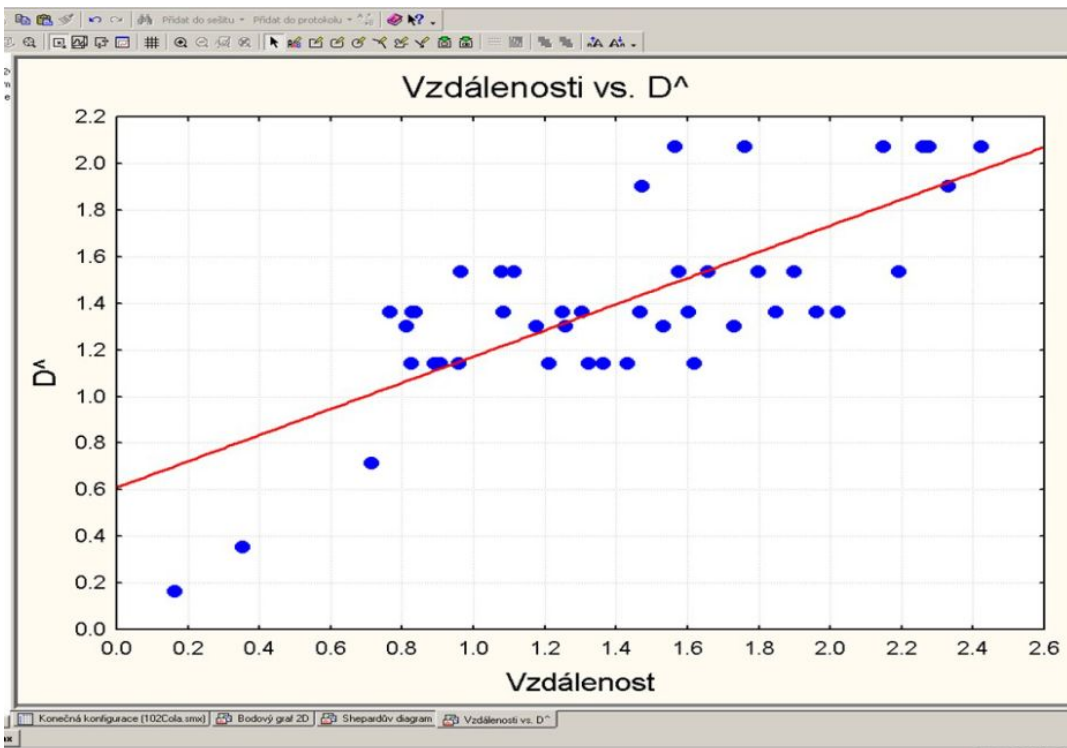


5. Výsledky: po kliknutí OK se otevře okno *Výsledky*: v horní polovině jsou dosažené výsledky optimalizačního výpočtu a v dolní polovině okna jsou tři záložky: *Základ*, *Detaily* a *Přehled a ukládání*. Kliknutím na jednotlivá okénka budeme nyní postupně otevírat výsledky.



14. *Shepardův diagram*: tento rozptylový graf vstupních dat podobnosti (nebo nepodobnosti) proti zadaným vzdálenostem z dat ukazuje $D^$ hodnoty čili monotónně transformovaná vstupní data jako schodovitou funkci. Většinu bodů je zde shluknuta kolem schodovité křivky. Proto lze uzavřít, že toto dvoj-rozměrné zobrazení je dostatečné k popisu podobnosti mezi objekty.





Příklad 10.3 Vícerozměrné škálování příbuznosti 14 lidských emocí

Posuďte podobnost nebo vzájemný vztah 14 lidských emocí, když byla psychology vypracována tabulka vzájemných vzdáleností způsobem porovnání „každé emoce s každou“. Při dokonalé podobnosti emocí je přidělena vzdálenost 0 a při naprosté nepodobnosti vzdálenost 10. Aplikujte CMDS a porovnejte se NNMDS.

Data: Prvky trojúhelníkové matice proximit *Emoce* jsou nepodobnosti dvojice emocí: *Spok* značí spokojený, *Vzru* značí vzrušující, *Přek* překvapený, *Horl* horlivý, *Štas* šťastný, *Vášn* vášnivý, *Něžn* něžný, *Pohr* pohrdá, *Vyle* vylekaný, *Bázl* bázlivý, *Prov* provinilý, *Smut* smutný, *Zlos* zlostný, *Odmí* odmítavý.

	<i>Spok</i>	<i>Vzru</i>	<i>Přek</i>	<i>Horl</i>	<i>Štas</i>	<i>Vášn</i>	<i>Něžn</i>	<i>Pohr</i>	<i>Vyle</i>	<i>Bázl</i>	<i>Prov</i>	<i>Smut</i>	<i>Zlos</i>	<i>Odmí</i>
<i>Spok</i>	0	3	4	6	1	2	3	9	8	8	8	8	8	9
<i>Vzru</i>	7	0	5	2	3	4	2	9	9	8	7	7	8	9
<i>Přek</i>	7	7	0	5	3	5	5	7	3	7	7	5	6	7
<i>Horl</i>	9	4	9	0	6	1	2	7	9	6	9	6	7	7
<i>Štas</i>	1	3	6	9	0	2	4	9	8	9	7	9	9	8
<i>Vášn</i>	3	3	4	2	1	0	2	1	5	8	6	5	2	8
<i>Nezn</i>	6	2	9	6	7	1	0	9	7	8	7	7	9	9
<i>Pohr</i>	9	7	8	9	9	3	9	0	5	4	4	3	2	2
<i>Vyle</i>	9	6	4	2	9	6	6	9	0	3	6	4	3	5
<i>Bázl</i>	9	6	3	9	9	3	9	4	1	0	3	4	3	5
<i>Prov</i>	9	9	9	9	9	3	9	9	2	4	0	4	4	4
<i>Smut</i>	9	7	3	8	9	3	6	9	2	2	4	0	2	2
<i>Zlos</i>	9	9	7	6	9	1	5	2	4	3	9	6	0	2
<i>Odmí</i>	9	9	7	9	9	9	9	9	2	1	4	1	3	0
Means	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Std.Dev.	0	0	0	0	0	0	0	0	0	0	0	0	0	0
No. Cases	91													
Matice	2													

STATISTICA Cz - [Data: 103Emoce (14s krát 18F)]

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Ojno Nápožveda

1 2 3 4 5 6 7 8 9 10 11 12 13 14
pok trzu prek forl stas lasn lezn pohr vyle bazl prov smut zlos odmi

Spok 0 3 4 6 1 2 3 9 8 9 8 8 8 9
Vzru 7 0 5 2 3 4 2 9 9 8 7 7 8 9
Přek 7 7 0 5 3 5 5 7 3 7 7 5 6 7
Horl 9 4 9 0 6 1 2 7 9 6 9 6 7 7
Štas 1 3 6 9 0 2 4 9 8 9 7 9 9 8
Vášn 3 3 4 2 1 0 2 1 5 8 6 5 2 8
Nezn 6 2 9 6 7 1 0 9 7 8 7 7 9 9
Pohr 9 7 8 9 9 3 9 0 5 4 4 3 2 2
Vyle 9 6 4 2 9 6 6 9 0 3 6 4 3 5
Bázl 9 6 3 9 9 3 9 4 1 0 3 4 3 5
Prov 9 9 9 9 9 3 9 9 2 4 0 4 4 4
Smut 9 7 3 8 9 3 6 9 2 2 4 0 2 2
Zlos 9 9 7 6 9 1 5 2 4 3 9 6 0 2
Odmí 9 9 7 9 9 9 9 9 2 1 4 1 3 0
Means 0 0 0 0 0 0 0 0 0 0 0 0 0 0
Std.Dev. 0 0 0 0 0 0 0 0 0 0 0 0 0 0
No. Cases 91
Matice 2

Vícerozměrné škálování: 103Emoce

Základní nastavení Možnosti

Proměnné: žádné

Počet dimenzí: 2

Vyberte proměnné (objekty) na analýzu

1:Spok 13:Zlos
2:Vzru 14:Odmí
3:Přek
4:Horl
5:Štas
6:Vášn
7:Nezn
8:Pohr
9:Vyle
10:Bázl
11:Prov
12:Smut

Vyberte proměnné:
1:14

STATISTICA Cz - [Data: 103Emoce (1-4s krát 18?)]

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Okno Nápožída

Přidat do sešitu Přidat do protokolu

Arial 10 B I U

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	pok	tzru	rek	forl	stas	asn	lezn	ohr	lyle	bazl	prov	mut	tos	ldmi
Spok	0	3	4	6	1	2	3	9	8	9	8	8	8	9
Vzru	7	0	5	2	3	4	2	9	9	8	7	7	8	9
Prek	7	7	0	5	3	5	5	7	3	7	7	5	6	7
Horl	9	4	9	0	6	1	2	7	9	6	9	6	7	7
Stas	1	3	6	9	0	2	4	9	8	9	7	9	9	8
Vasn	3	3	4	2	1	0	2	1	5	8	6	5	2	8
Nezn	6	2	9	6	7	1	0	9	7	8	7	7	9	9
Pohr	9	7	8	9	9	3	9	0	5	4	4	3	2	2
Vyle	9	6	4	2	9	6	6	9	0	3	6	4	3	5
Bazl	9	6	3	9	9	3	9	4	1	0	3	4	3	5
Prov	9	9	9	9	3	9	9	2	4	0	4	4	4	4
Smut	9	7	3	8	9	3	6	9	2	2	4	0	2	2
Zlos	9	9	7	6	9	1	5	2	4	3	9	6	0	2
Odmi	9	9	7	9	9	9	9	2	1	4	1	3	0	0
Means	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Std.Dev.	0	0	0	0	0	0	0	0	0	0	0	0	0	0
No.Cases	91													
Matrice	2													

Odhadování parametrů: 103Emoce

iter.	[dim=4]	D*	D^*	D^*	d^*	
s: t:	cosin	krok	čistý stres	alienace	čistý stres	
15	1	.644	.171		3.257437	-.1289169
16	1	.709	.158		3.254799	-.1288647
17	1	.949	.308		3.250562	-.1287808
18	1	.976	.426		3.246183	-.1286941
19	1	.877	.334		3.243764	-.1286461
20	1	.557	.153		3.242612	-.1286232
20*			6.363432	.1794519	3.242612	-.1286232
3	0				4.015064	-.1431259
4	1	.073			3.846738	-.1400936
5	1	.975	.263		3.562461	-.1348177
6	1	.576	.145		3.489403	-.1334282
7	1	.903	.254		3.407603	-.1318550
8	1	.925	.331		3.350433	-.1307442
9	1	.888	.318		3.317630	-.1301026
10	1	.780	.229		3.300931	-.1297747
11	1	.700	.171		3.290790	-.1295753
12	1	.851	.227		3.280302	-.1293686
13	1	.944	.341		3.268467	-.1291350
14	1	.878	.310		3.261034	-.1289881

Procedura odhadování konvergovala

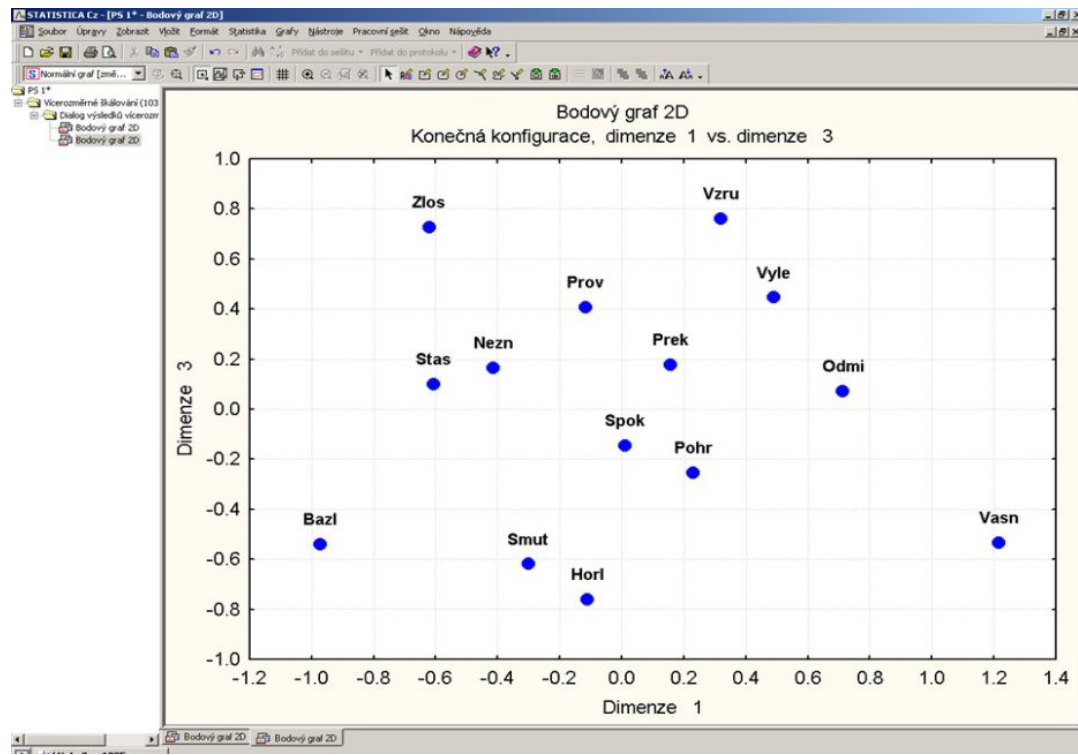
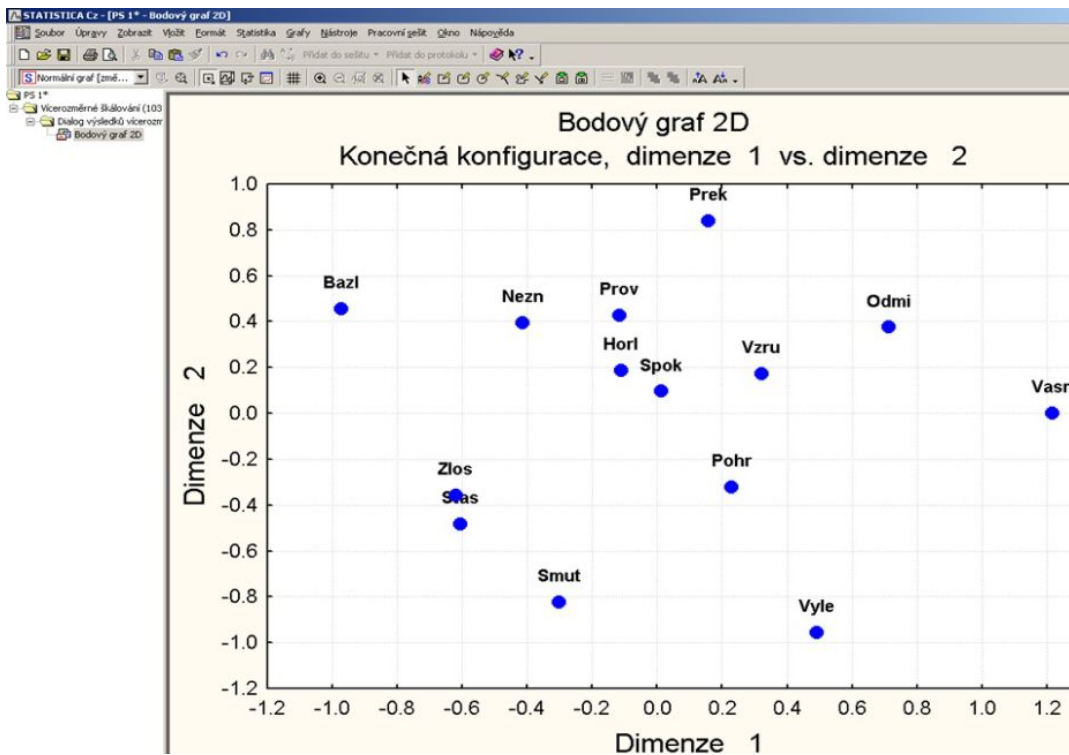
Pohr	9	7	8	9	9	3	9	0	5	4	4	3	2	2
Vyle	9	6	4	2	9	6	6	9	0	3	6	4	3	5
Bazl	9	6	3	9	9	3	9	4	1	0	3	4	3	5
Prov	9	9	9	9	3	9	9	2	4	0	4	4	4	4
Smut	9	7	3	8	9	3	6	9	2	2	4	0	2	2
Zlos	9	9	7	6	9	1	5	2	4	3	9	6	0	2
Odmi	9	9	7	9	9	9	9	2	1	4	1	3	0	0
Means	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Std.Dev.	0	0	0	0	0	0	0	0	0	0	0	0	0	0
No.Cases	91													
Matrice	2													

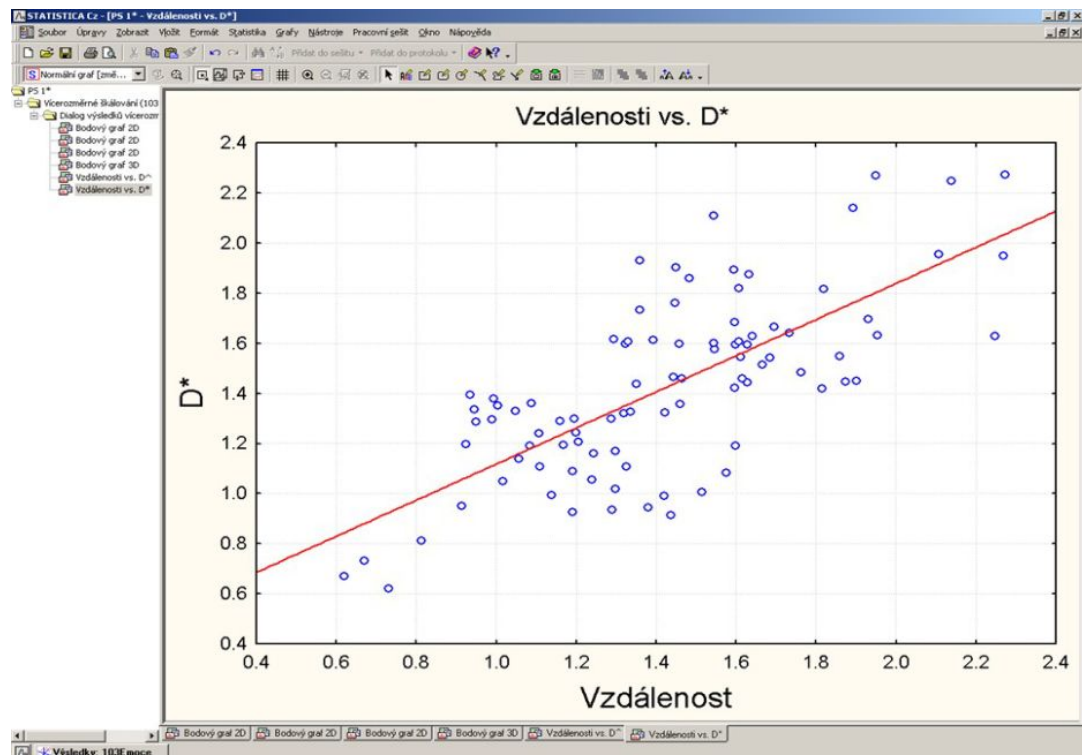
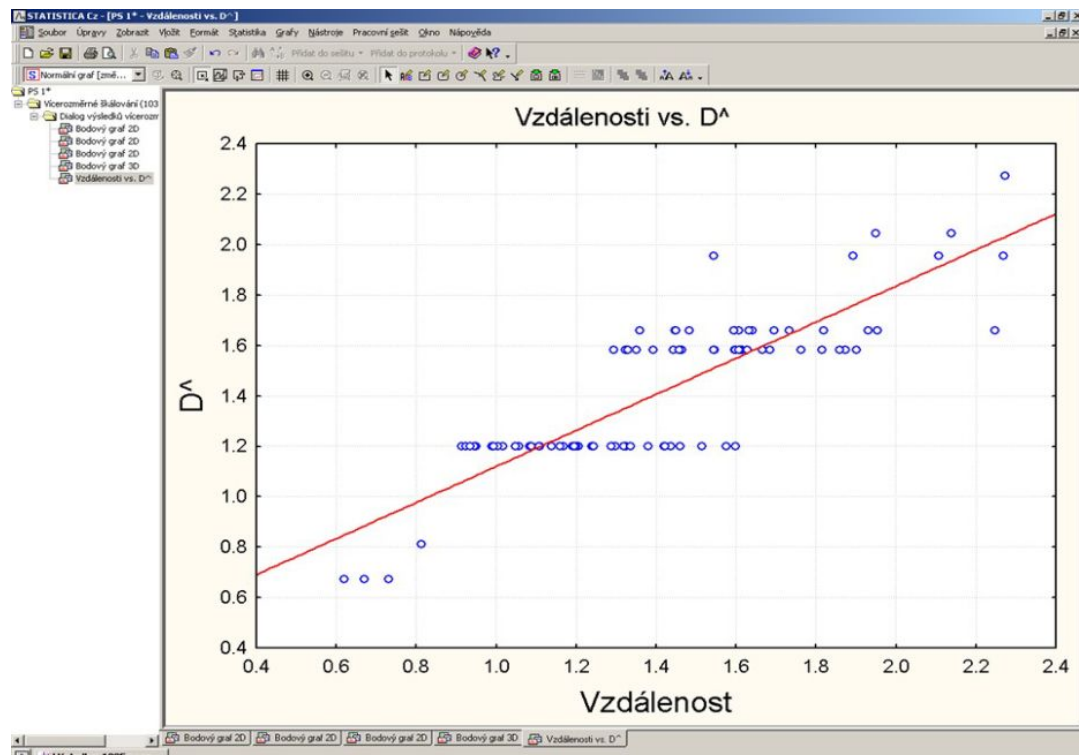
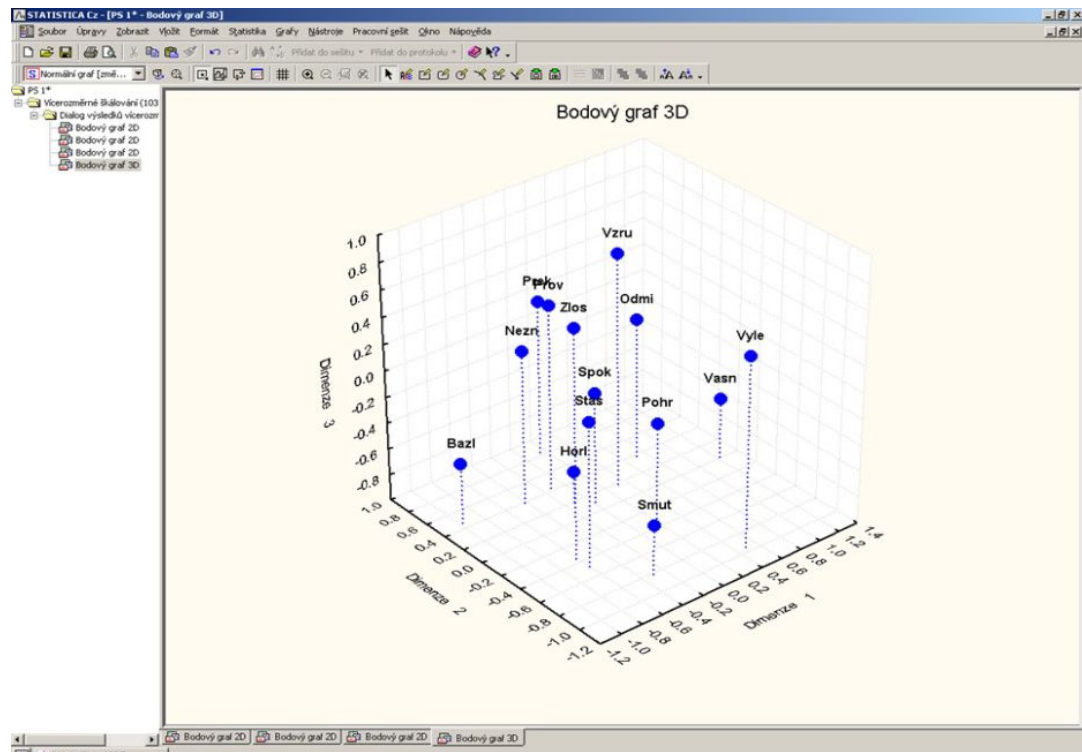
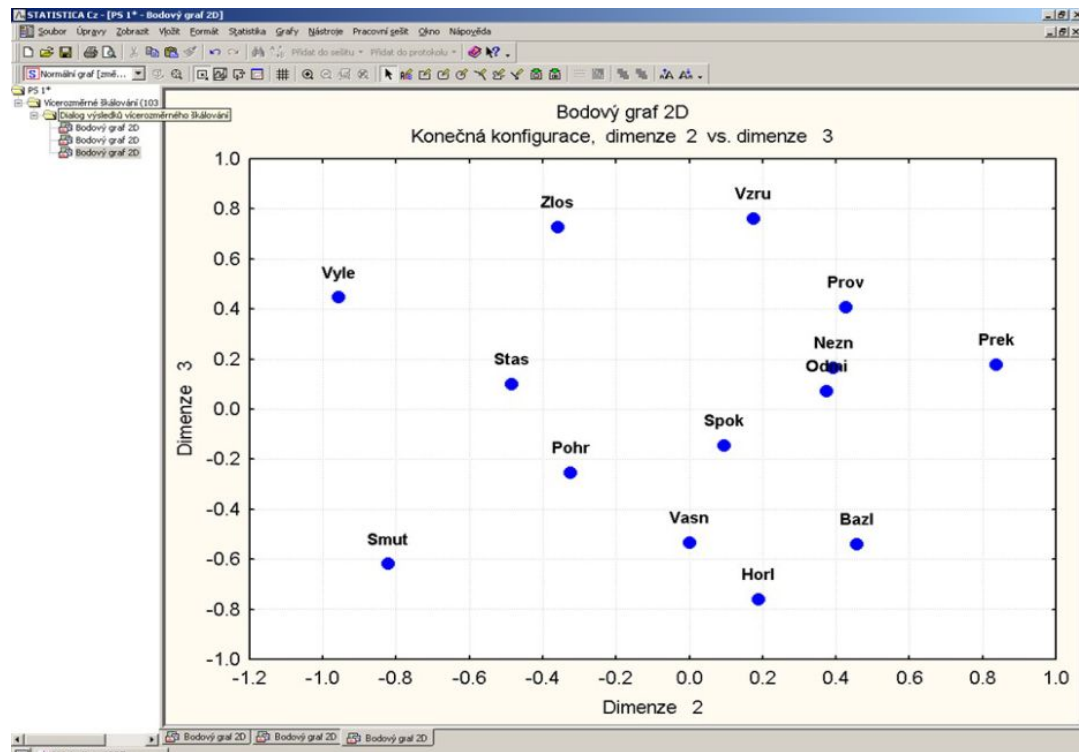
Výsledky: 103Emoce

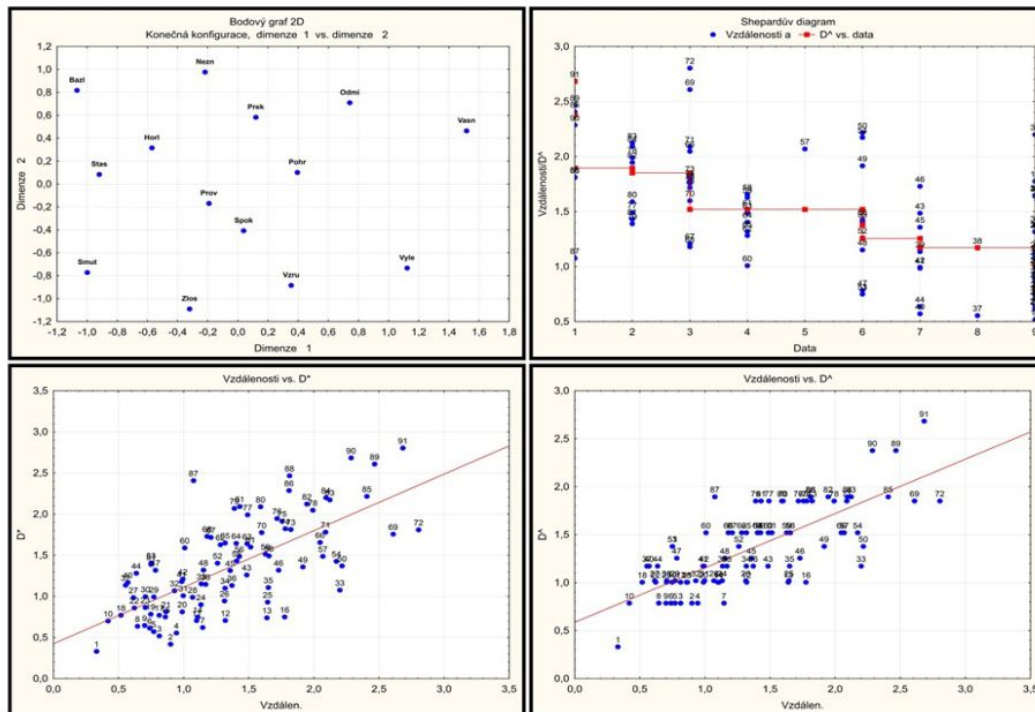
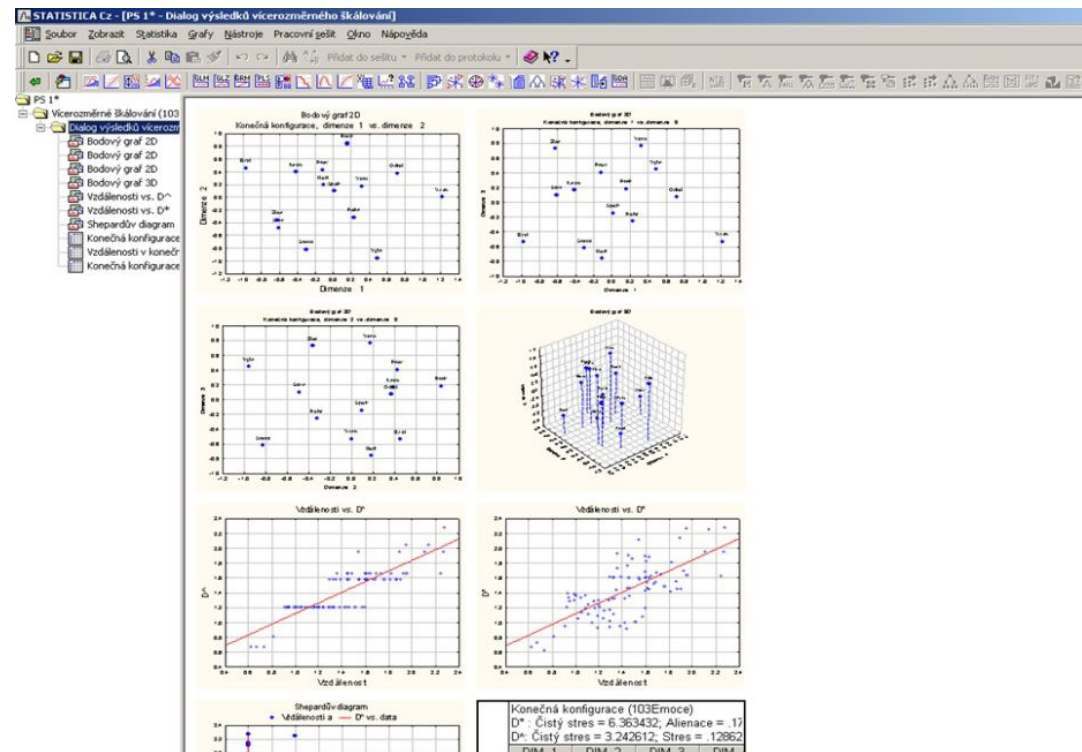
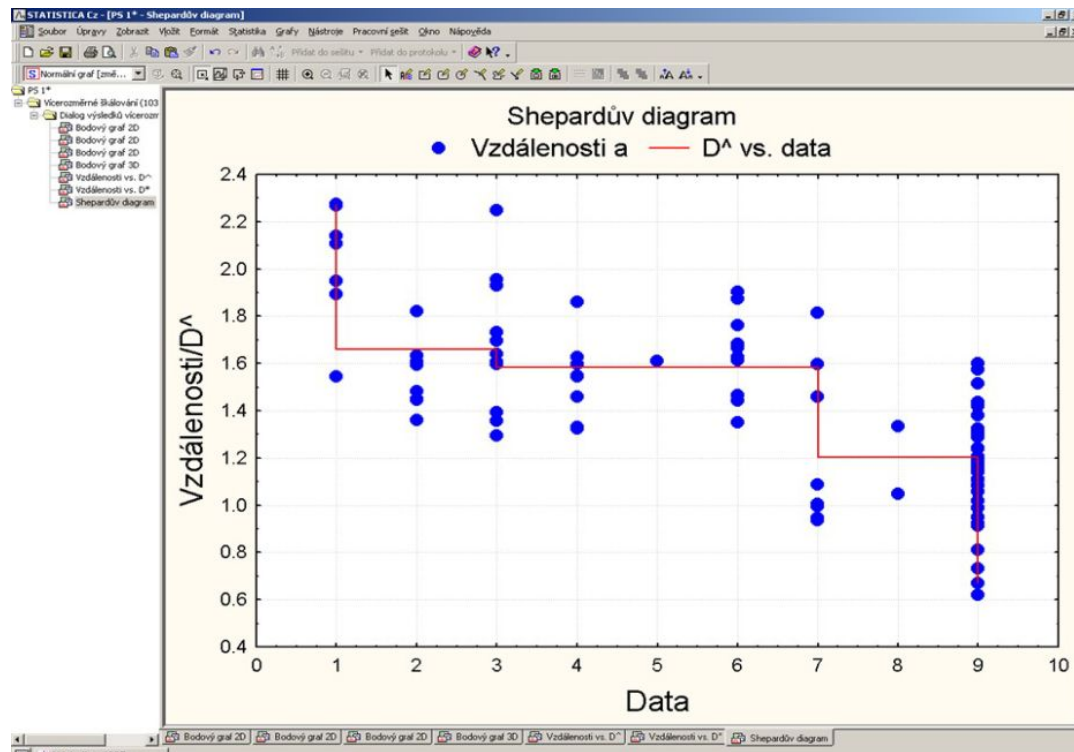
14 prom. ze souboru
 Celk. počet dimenzí : 4
 Počet.konfig. : (minulá konečná)
 Posl. počítaná iterace : 40; Nejlepší iterace: 20
 D* : Čistý stres = 6.363432; Alienace = .1794519
 D^: Čistý stres = 3.242612; Stres = .1286232

Základ Details Přehled & ukládání

Shnutí: Konečná konfig. Graf konečné konfig., 2D
 Hodnoty D^ Graf konečné konfig., 3D
 Hodnoty D* Graf D^ vs. vzdálenosti
 Matice vzdáleností Graf D* vs. vzdálenosti
 Souhrnné statistiky Shepardův diagram





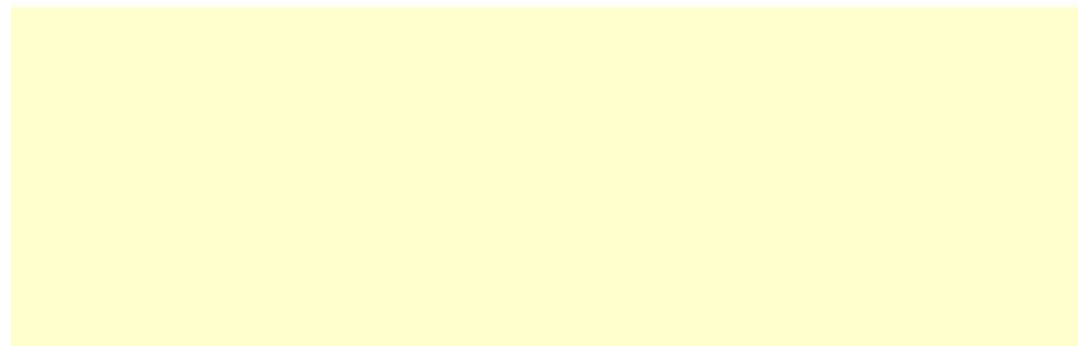


PŘÍKLAD 10.4 Vícezměrné škálování podobnosti aktivit relaxace

Vícezměrným škálováním posuďte podobnost a vzájemné vztahy 15 rozličných aktivit osobní relaxace, které tvoří $15(15 - 1)/2 = 105$ prvků matice nad diagonálou čili párů vzájemného porovnání “každé aktivity s každou”.

Při dokonalé podobnosti dvou aktivit přidělena vzdálenost 0, zatímco při naprosté nepodobnosti vzdálenost 25.

Jako rozhodčí kritérium těsnosti proložení užitje koeficient *stress*.



Data: Prvky trojúhelníkové matice proximit *Relax* vyjadřují vzdálenosti aktivit relaxace ve dvojicích: *Konc* značí koncert, *Muze* značí muzeum, *Diva* divadlo, *Kino* kino, *TV* televize, *Konf* konference, *Cetb* četba, *Hoke* divák hokeje, *Bale* balet, *Polí* zájem o politiku, *Moda* móda, *Doku* dokumentaristika, *Vyst* zájemce o výstavy, *Naku* zájem o nákupy, *Rest* pobyt v restauraci.

	Konc	Muze	Diva	Kino	TV	Konf	Cetb	Hoke	Bale	Polí	Moda	Doku	Vyst	Naku	Rest
Konc	0	16	3	12	15	20	15	21	7	19	9	22	7	21	8
Muze	16	0	18	12	21	10	12	23	10	22	7	16	3	22	8
Diva	3	18	0	11	16	19	13	23	6	25	13	16	13	22	7
Kino	12	12	11	0	2	15	9	19	18	22	15	19	12	12	9
TV	15	21	16	2	0	12	19	7	19	14	12	13	21	23	21
Konf	20	10	19	15	12	0	6	22	25	8	19	7	13	21	21
Cetb	15	12	13	9	19	6	0	20	15	22	20	13	10	18	2
Hoke	21	23	23	19	7	22	20	0	25	23	22	15	22	18	22
Bale	7	10	6	18	19	25	15	25	0	25	8	23	13	21	5
Polí	19	22	25	22	14	8	22	23	25	0	25	13	12	22	25
Moda	9	7	13	15	12	19	20	22	8	25	0	25	7	9	9
Doku	22	16	16	19	13	7	13	15	23	13	25	0	18	22	23
Vyst	7	3	13	12	21	13	10	22	13	12	7	18	0	12	10
Naku	21	22	22	12	23	21	18	18	21	22	9	22	12	0	8
Rest	8	8	7	9	21	21	2	22	5	25	9	23	10	8	0
Means	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Std.Dev.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
No.Cases	105														
Matic	2														

The screenshot shows the STATISTICA interface with a data matrix for '104Relax'. Two dialog boxes are open: 'Vícerozměrné škálování: 104Relax' and 'Vyberte proměnné (objekty) na analýzu'. The first dialog shows 'Počet dimenzí' set to 4. The second dialog shows selected variables: 1-Konc, 2-Muze, 3-Diva, 4-Kino, 5-TV, 6-Konf, 7-Cetb, 8-Hoke, 9-Bale, 10-Polí, 11-Moda, 12-Doku, 13-Vyst, 14-Naku, 15-Rest.

Doku	22	16	16	19	13	7	13	15	23	13	25
Vyst	7	3	13	12	21	13	10	22	13	12	7
Naku	21	22	22	12	23	21	18	18	21	22	9
Rest	8	8	7	9	21	21	2	22	5	25	9
Means	0	0	0	0	0	0	0	0	0	0	0
Std.Dev.	0	0	0	0	0	0	0	0	0	0	0
No.Cases	105										
Matic	2										

The 'Výsledky: 104Relax' dialog box displays the following information:

- 15 prom. ze souboru
- Celk. počet dimenzí : 4
- Počít.konfig.: (Guttman-Lingoesova)
- Posl. počítaná iterace : 73; Nejlepší iterace: 45
- D*: Čistý stres = 6.696388; Alienace = .1718729
- D^: Čistý stres = 4.039824; Stres = .1339954

 The dialog also includes tabs for 'Základ', 'Detaily', and 'Přehled & ukládání', and buttons for 'Shnutí', 'Možnosti', and 'Storno'.

