NCSS2007
ve vícerozměrné statistické analýze

Metoda hlavních komponent

PCA

Analýza hlavních komponent (PCA)

Zaměření metody PCA: PCA jedna z nejstarších a nejvíce používaných metod vícerozměrné analýzy.

Zavedena: Pearsonem již v roce 1901 a nezávisle Hotellingem v roce 1933.

Cílem PCA: zjednodušení popisu lineárně závislých tj. korelovaných znaků, a to rozkladem matice dat do ***matice strukturní*** (= využité hlavní komponenty) a do ***matice šumové*** (= nevyužité hlavní komponenty).

Popis metody PCA:

1) Lineární transformace původních znaků na nové, nekorelované proměnné, zvané *hlavní komponenty*.

2) Základní charakteristikou každé hlavní komponenty je její míra variability čili *rozptyl.*

3) Hlavní komponenty jsou *seřazeny dle důležitosti*, tj. dle klesajícího rozptylu, od největšího k nejmenšímu.

4) *Většina informace* o variabilitě dat je přitom soustředěna do první komponenty a nejméně informace je obsaženo v poslední komponentě.

5) ***Pravidlo:*** má-li nějaký původní znak malý či dokonce žádný rozptyl, není schopen přispívat k rozlišení mezi objekty.

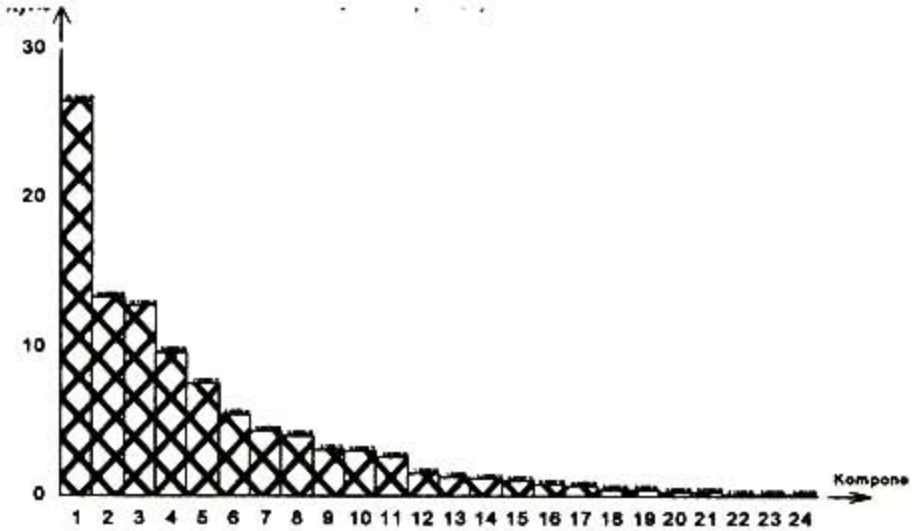
6) Využitím PCA je *snížení dimenze úlohy* čili redukce počtu znaků bez velké ztráty informace, užitím pouze prvních několika hlavních komponent.

7) *Nevyužité hlavní komponenty* obsahují malé množství informace, protože jejich rozptyl je příliš malý.

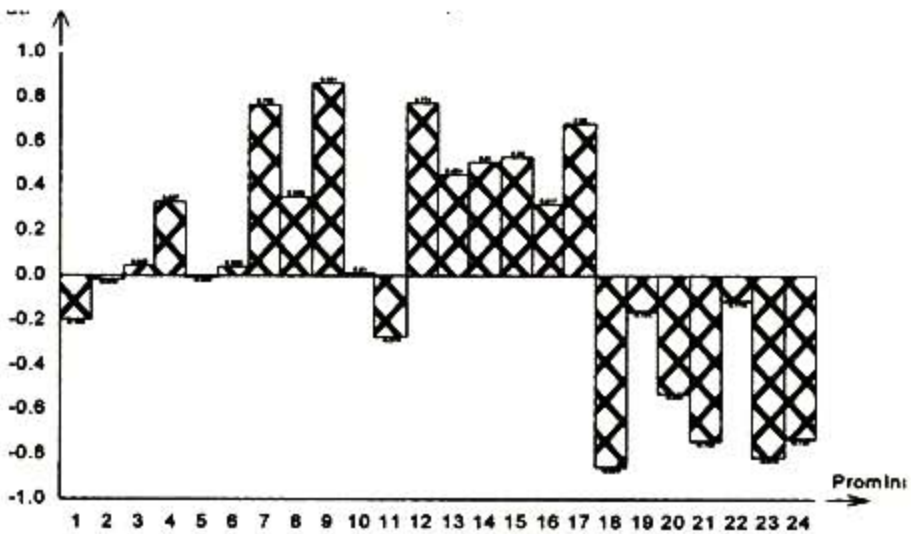
8) Hlavní komponenty jsou nekorelované.

9) První hlavní komponenta je například vhodným ukazatelem jakosti.

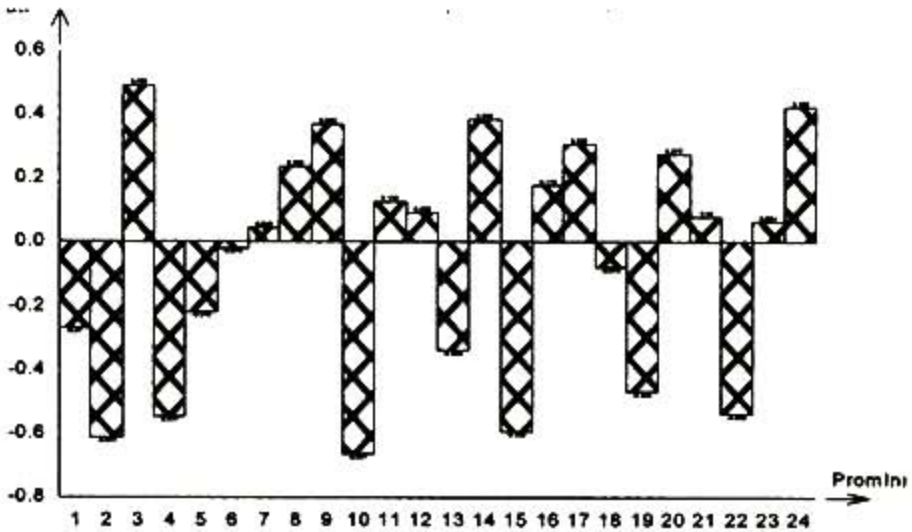
10) První dvě resp. první tři hlavní komponenty se využívají především jako techniky zobrazení vícerozměrných dat v projekci do roviny (nebo do prostoru).



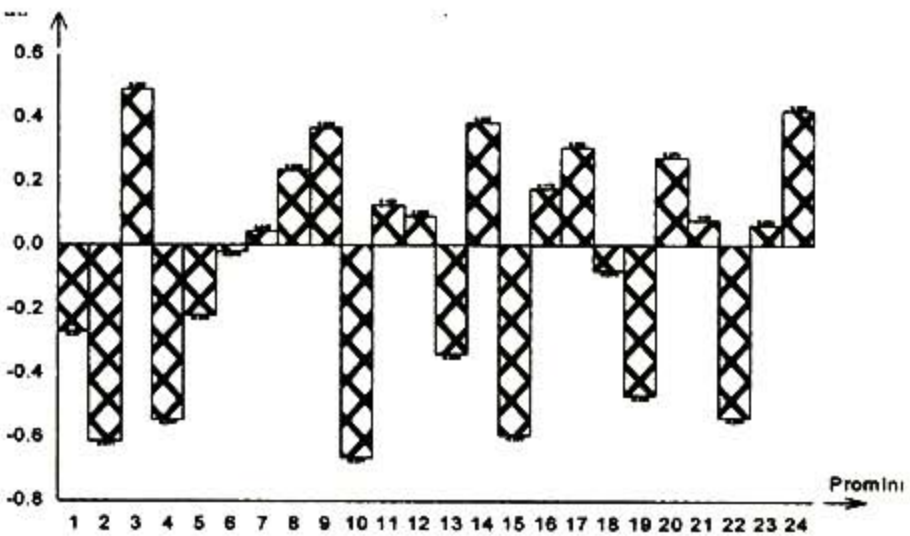
Obr. 1. Sloupcový diagram indexového grafu úpatí pro 38 objektů a 24 původních proměnných zdrojové matice Wine.



Obr. 2. Složení 1. hlavní komponenty z původních proměnných pro 38 objektů a 24 původních proměnných zdrojové matice Wine.



Obr. 3. Složení 2. hlavní komponenty z původních proměnných pro 38 objektů a 24 původních proměnných zdrojové matice Wine.



Obr. 4. Složení 3. hlavní komponenty z původních proměnných pro 38 objektů a 24 původních proměnných zdrojové matice Wine.

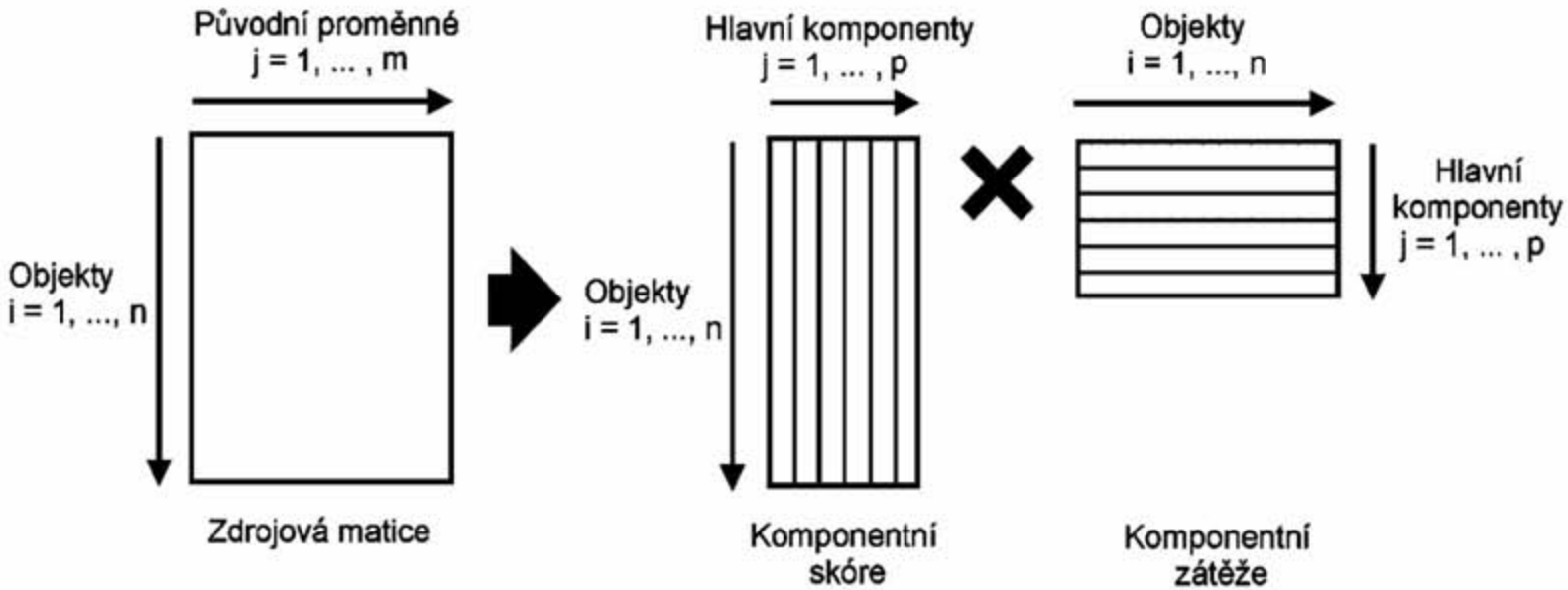
Analýza hlavních komponent (PCA)

Cílem PCA: transformace dat z původních proměnných x_i, j=1, ..., m, do menšího počtu latentních proměnných y_j.

PCA vystihují téměř *proměnlivost* původních proměnných a jsou vzájemně nekorelované.

První hlavní komponenta y_1 popisuje největší část proměnlivosti čili rozptylu původních dat,

Druhá hlavní komponenta y_2 popisuje zase největší část rozptylu neobsaženého v y_1 atd.



Obr. 4.1 Schéma maticových výpočtů v PCA.

- **Maximální počet hlavních komponent:** počet efektivních hlavních komponent se rovná *hodnosti zdrojové matice* $\boldsymbol{X}$.

- $\boldsymbol{X}$ = **Struktura + šum:** všechny hlavní komponenty jsou vzájemně ortogonální a souvisí postupně se snižující hodnotu rozptylu objektů.

Model hlavních komponent má tvar

$$\boldsymbol{X} = \boldsymbol{T}\,\boldsymbol{P}^{\mathrm{T}} + \boldsymbol{E} = \textit{struktura dat} + \textit{šum}.$$

- **Střed modelu:** hlavní komponenty mají společný počátek, který odpovídá *průměrnému objektu* čili těžišti celého shluku objektů (centrování).

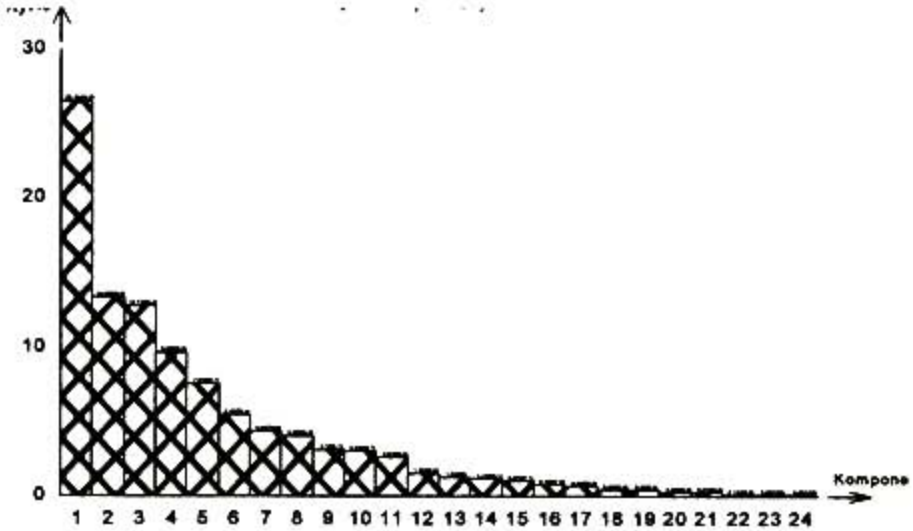
• Komponentní váhy,

zátěže - vztah mezi *X* a *PC*:

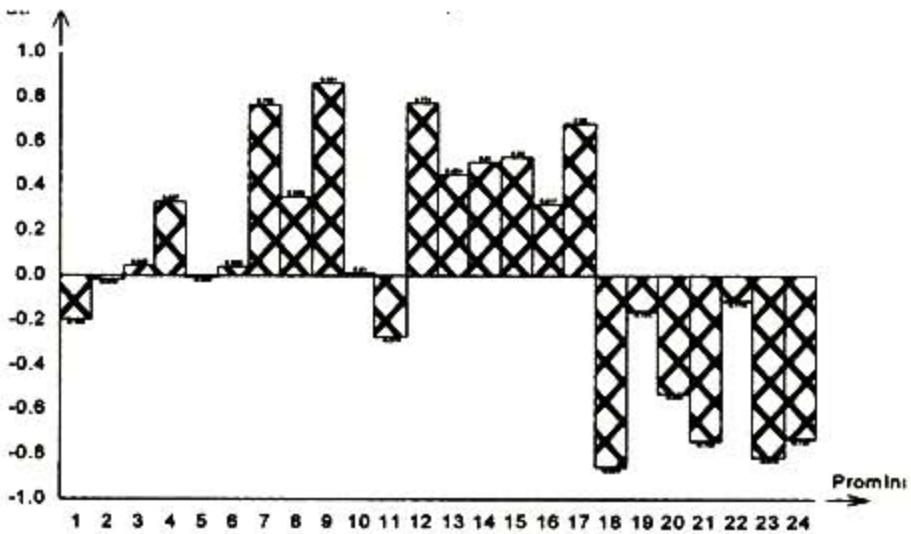
Hlavní komponenta představuje *lineární kombinaci všech m* vektorů v prostoru znaků v *m* rozměrném prostoru a jejich koeficienty se nazývají *komponentní váhy*.

Komponentní váhy informují o vztahu mezi původními *m* znaky a hlavními komponentami.

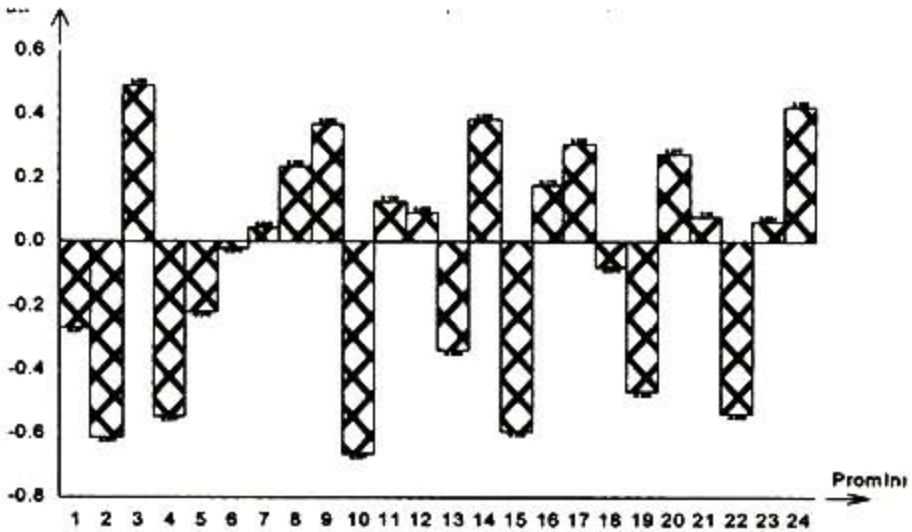
Na *grafu komponentních vah p* pro *PC*1 a *PC*2 jsou místo objektů jejich znaky a lze tak vyšetřovat závislosti a podobnosti mezi znaky.



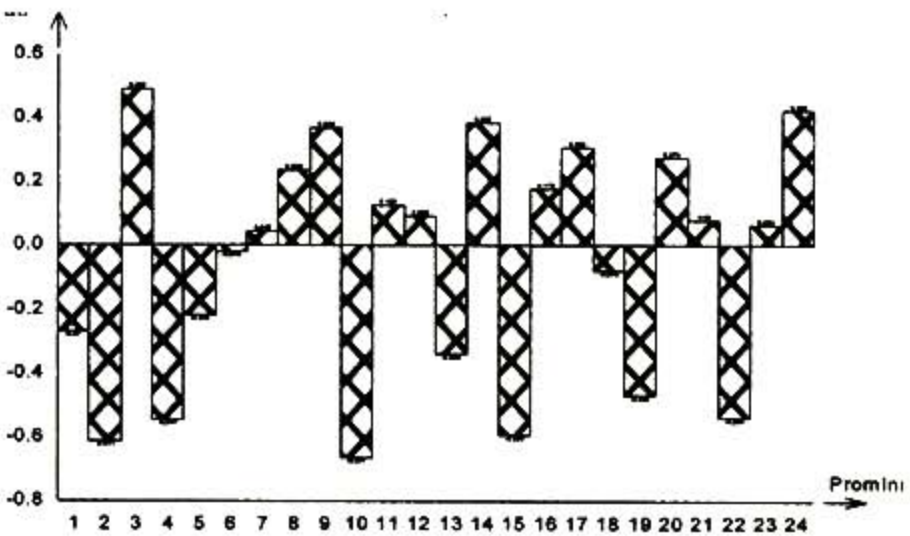
Obr. 1. Sloupcový diagram indexového grafu úpatí pro 38 objektů a 24 původních proměnných zdrojové matice Wine.



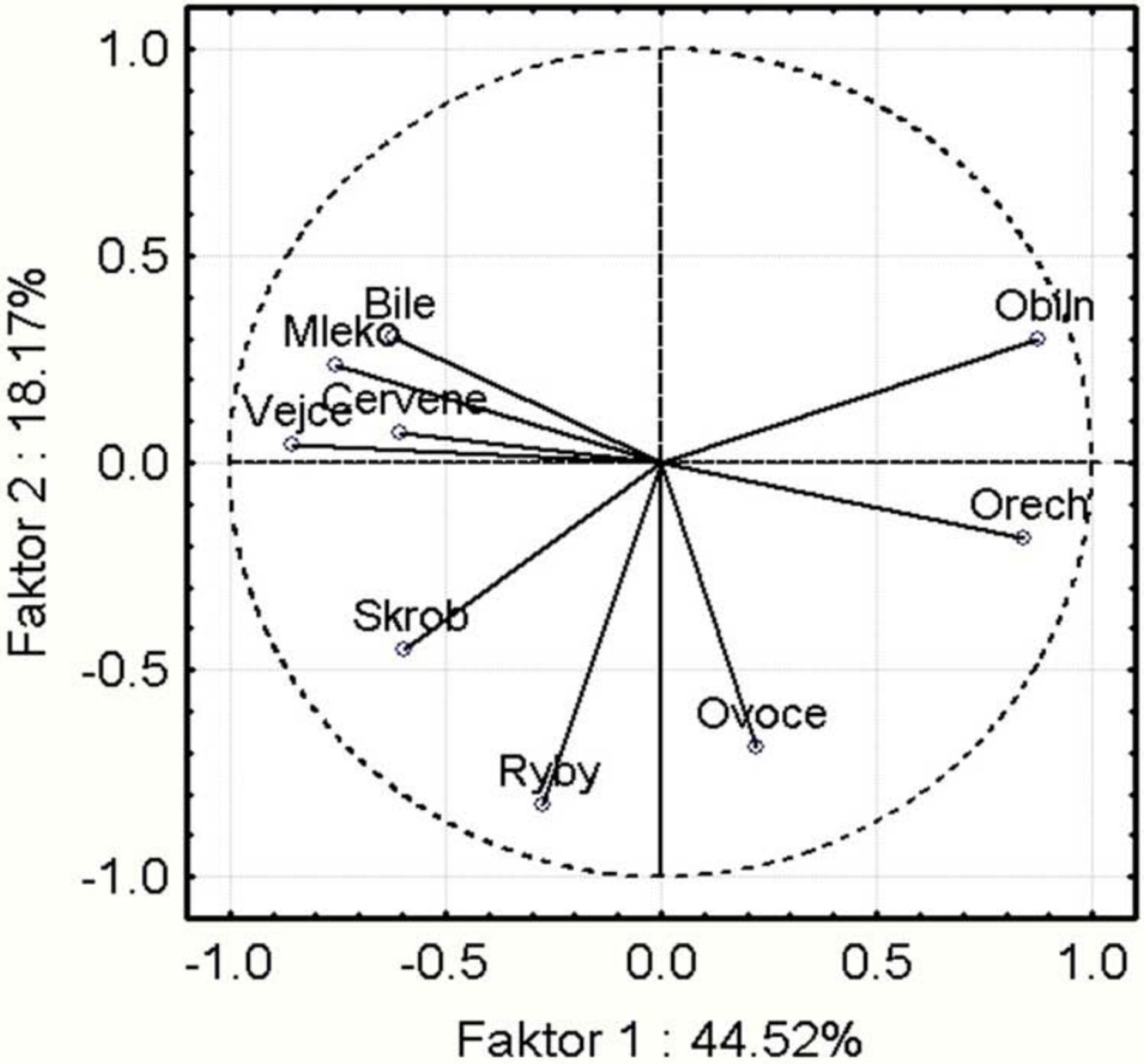
Obr. 2. Složení 1. hlavní komponenty z původních proměnných pro 38 objektů a 24 původních proměnných zdrojové matice Wine.



Obr. 3. Složení 2. hlavní komponenty z původních proměnných pro 38 objektů a 24 původních proměnných zdrojové matice Wine.



Obr. 4. Složení 3. hlavní komponenty z původních proměnných pro 38 objektů a 24 původních proměnných zdrojové matice Wine.


1.0
0.5
0.0
-0.5
-1.0
-1.0
-0.5
0.0
0.5
1.0
Faktor 2 : 18.17%
Faktor 1 : 44.52%
Bile
Mleko
Cervene
Vejce
Obiln
Orech
Skrob
Ovoce
Ryby

Výklad:

Graf komponentních vah, zátěží

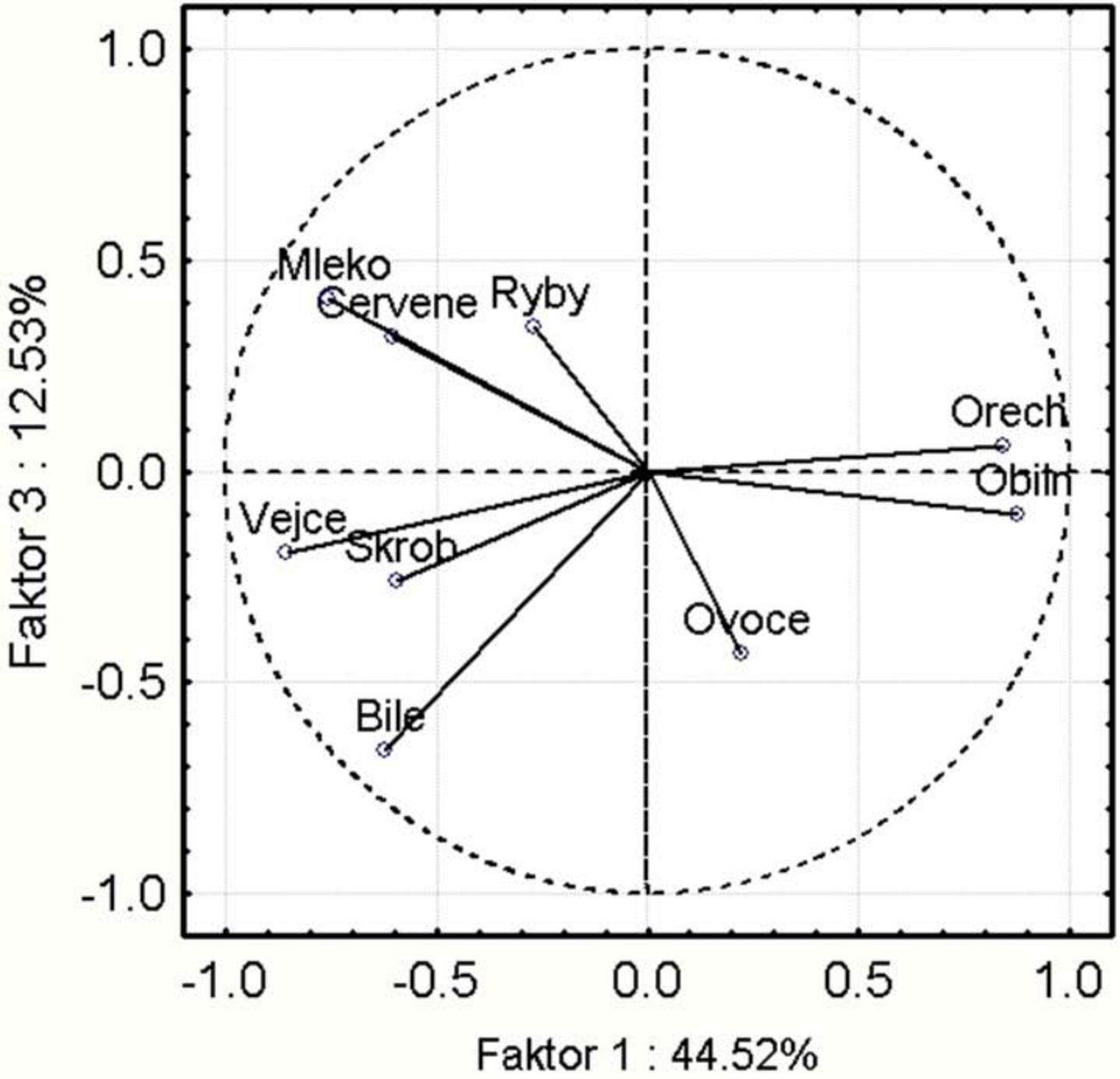
(Plot Components Weights)

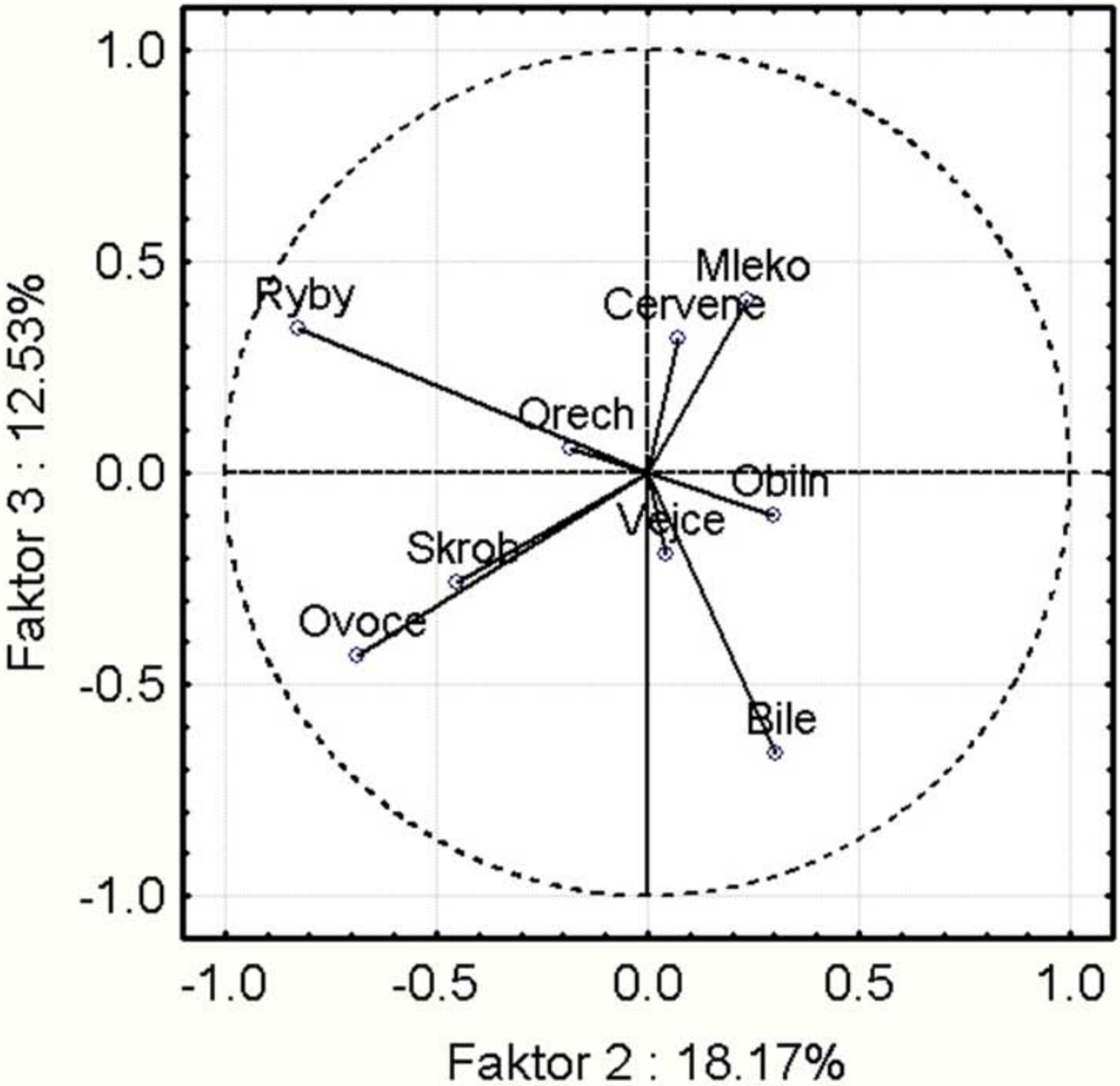
Zobrazí: komponentní váhy

Porovnávají se: vzdálenosti mezi proměnnými.
Krátká znamená silnou korelaci.

Nalezneme: shluk podobných proměnných, jež spolu korelují.

Představuje: most mezi původními proměnnými a hlavními komponentami.


1.0
0.5
0.0
-0.5
-1.0
Faktor 3 : 12.53%
Mleko
Cervene
Ryby
Orech
Obilin
Vejce
Skrob
Ovoce
Bile
-1.0
-0.5
0.0
0.5
1.0
Faktor 1 : 44.52%


1.0
0.5
0.0
-0.5
-1.0
Faktor 3 : 12.53%
Ryby
Mleko
Cervene
Orech
Obiln
Veice
Skrob
Ovoce
Bile
-1.0
-0.5
0.0
0.5
1.0
Faktor 2 : 18.17%

• Komponentní skóre

souřadnice objektů v prostoru hlavních komponent:

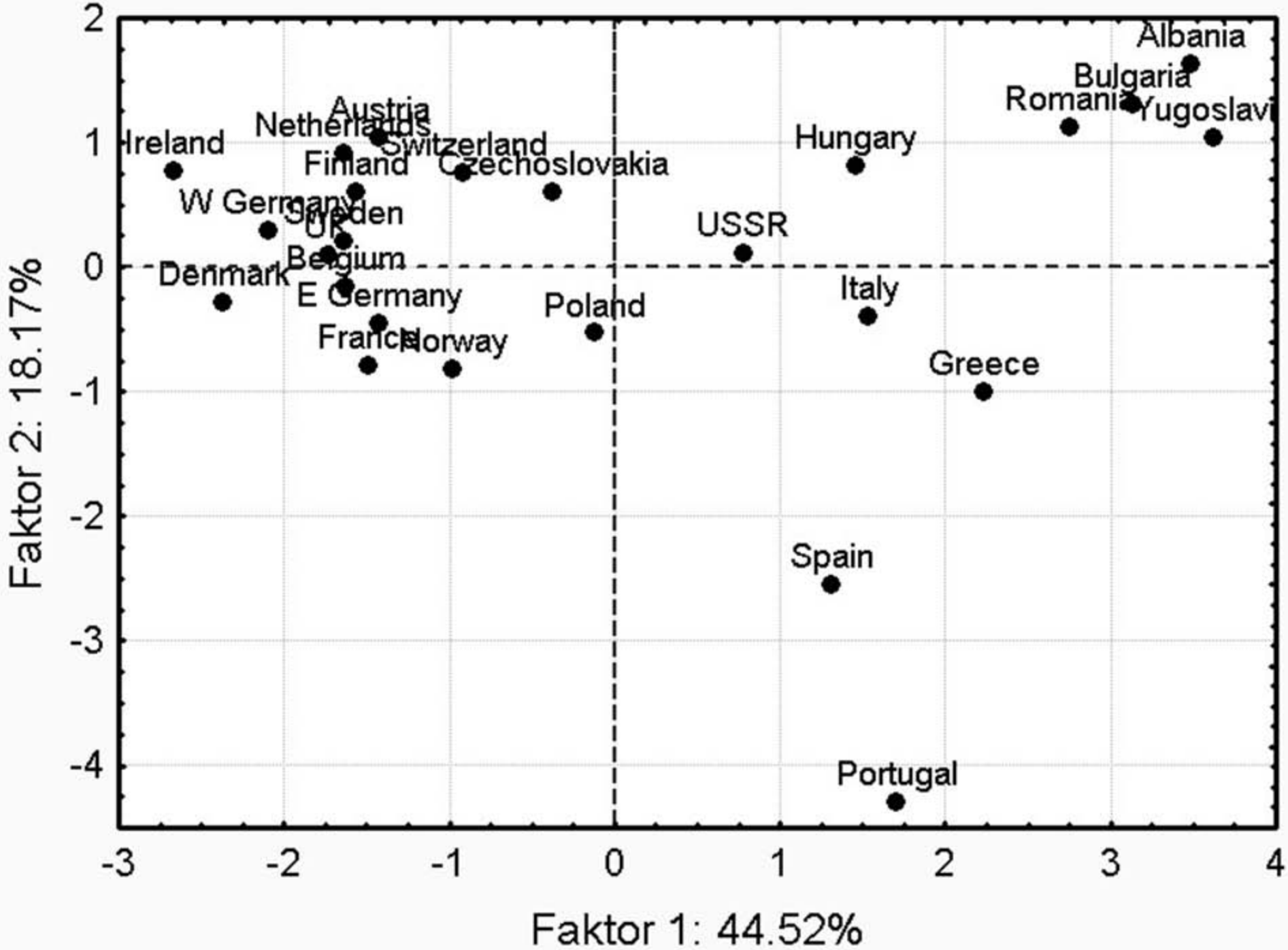
Souřadnice každého objektu na osách hlavních komponent nazýváme *skóre*.

Graf komponentního skóre je zobrazení dvou skórových vektorů vynesených v systému kartézských os jeden proti druhému.

Pravidlo k volbě grafů:

1. Na x-ové ose užijeme vždy stejnou hlavní komponentu (obvykle první) u všech grafů komponentního skóre: t_1 proti t_2, t_1 proti t_3, t_1 proti t_4, t_1 proti t_5, ... atd., takže budeme vyšetřovat ostatní hlavní komponenty proti stále stejné, první.

2. Užijeme první hlavní komponentu, protože vykazuje největší hodnotu rozptýlení pro vyšetřovanou úlohu a vyneseme ji na x-ovou osu.


Albania
Bulgaria
Romania
Yugoslavi
Austria
Netherlands
Switzerland
Ireland
Finland
Czechoslovakia
Hungary
W Germany
Sweden
UK
USSR
Belgium
Denmark
E Germany
Italy
Poland
France
Norway
Greece
Spain
Portugal
2
1
0
-1
-2
-3
-4
-3
-2
-1
0
1
2
3
4
Faktor 2: 18.17%
Faktor 1: 44.52%

Rozptylový diagram komponentního skóre

(Scatterplot)

1. **Umístění objektů:** daleko od počátku jsou extrémy. Objekty nejblíže počátku jsou nejtypičtější.

2. **Podobnost objektů:** objekty blízko sebe si jsou podobné, daleko od sebe jsou si nepodobné.

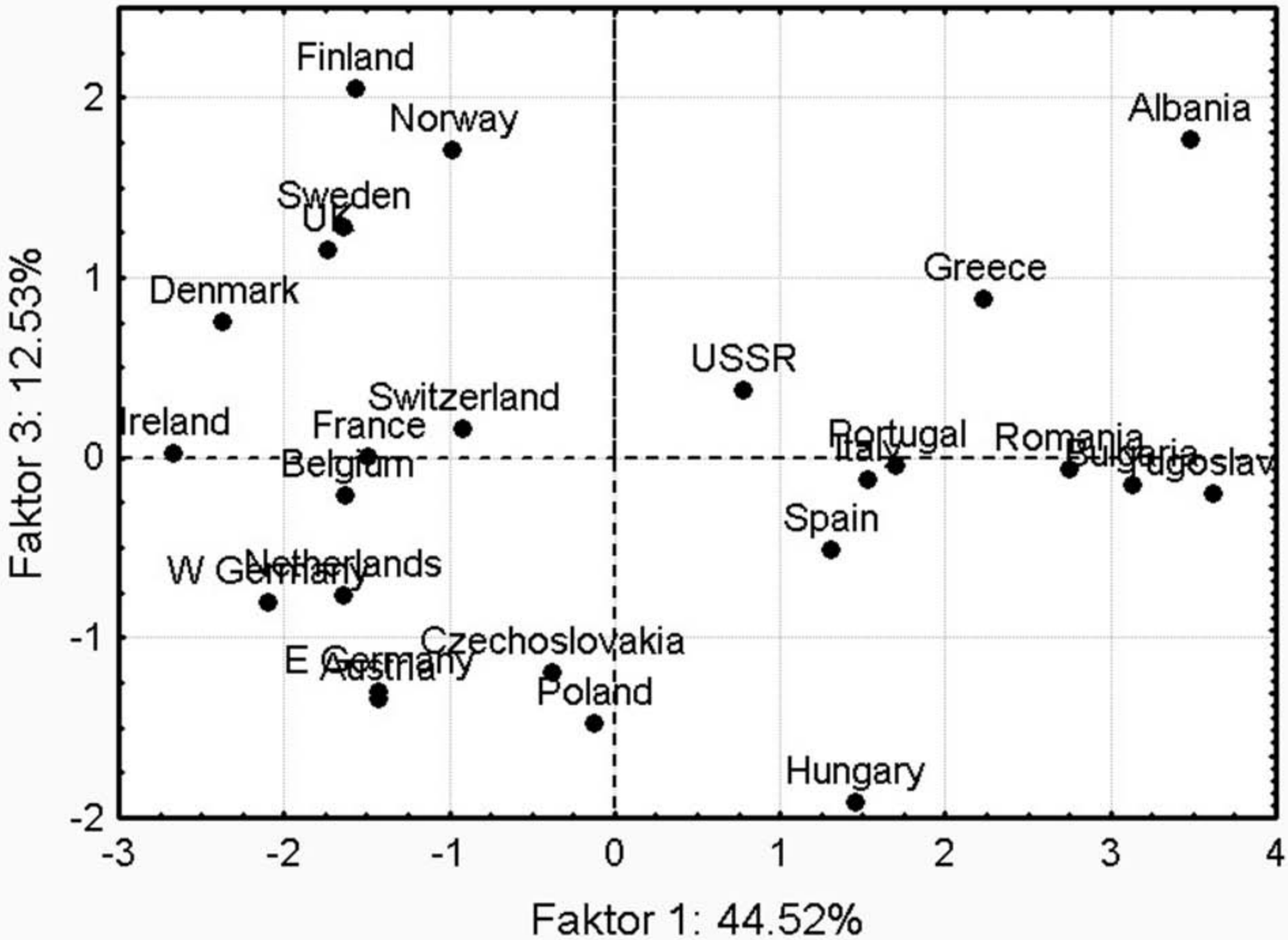
3. **Objekty v shluku:** umístěné zřetelně v jednom shluku jsou si podobné a nepodobné objektům v ostatních shlucích. Jsou-li shluky blízko sebe, znamená to značnou podobnost objektů.

4. **Osamělé objekty:** izolované objekty mohou být odlehlé.

5. **Odlehlé objekty:** ideálně bývají objekty rozptýlené po celé ploše diagramu. V opačném případě je špatný model.

6. **Pojmenování objektů:** výstižná jména objektů slouží k hledání hlubších souvislostí mezi objekty a vystihneme tak jejich fyzikální či biologický vztah.

7. **Vysvětlení místa objektu:** umístění objektu na ploše v diagramu může být porovnáváno s komponentními vahami původních proměnných ve dvojném grafu.


Finland
Norway
Albania
Sweden
UK
Greece
Denmark
USSR
Switzerland
Ireland
France
Portugal
Romania
Italy
Belgium
Spain
W Germany
Netherlands
Czechoslovakia
E Germany
Austria
Poland
Hungary
2
1
0
-1
-2
-3
-2
-1
0
1
2
3
4
Faktor 3: 12.53%
Faktor 1: 44.52%

Rozptylový diagram komponentního skóre
(Scatterplot)

Zobrazuje: *komponentní skóre* čili hodnoty obyčejně prvních dvou hlavních komponent u všech objektů.

Dokonalé rozptýlení objektů v rovině: vede k rozlišení objektů při jejich popisu pomocí y_1 a y_2.

Lze nalézt: shluk vzájemně podobných objektů a dále objekty odlehlé a silně odlišné od ostatních.

Využití: k identifikaci odlehlých objektů, identifikaci trendů, tříd, shluků objektů, k objasnění podobnosti objektů atd.

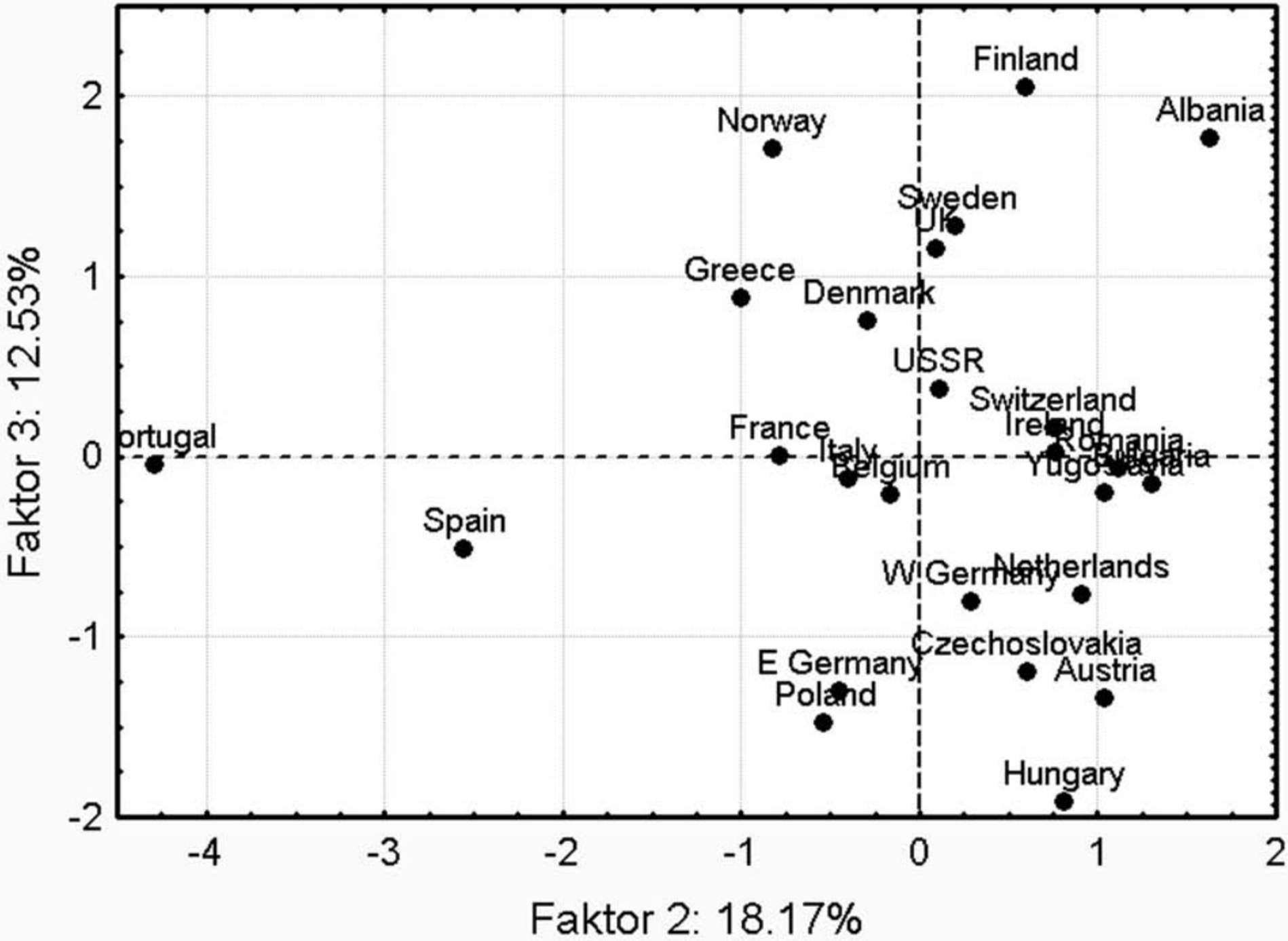
Nemožné analyzovat všechny diagramy: jich je mnoho:
např. $m < n$

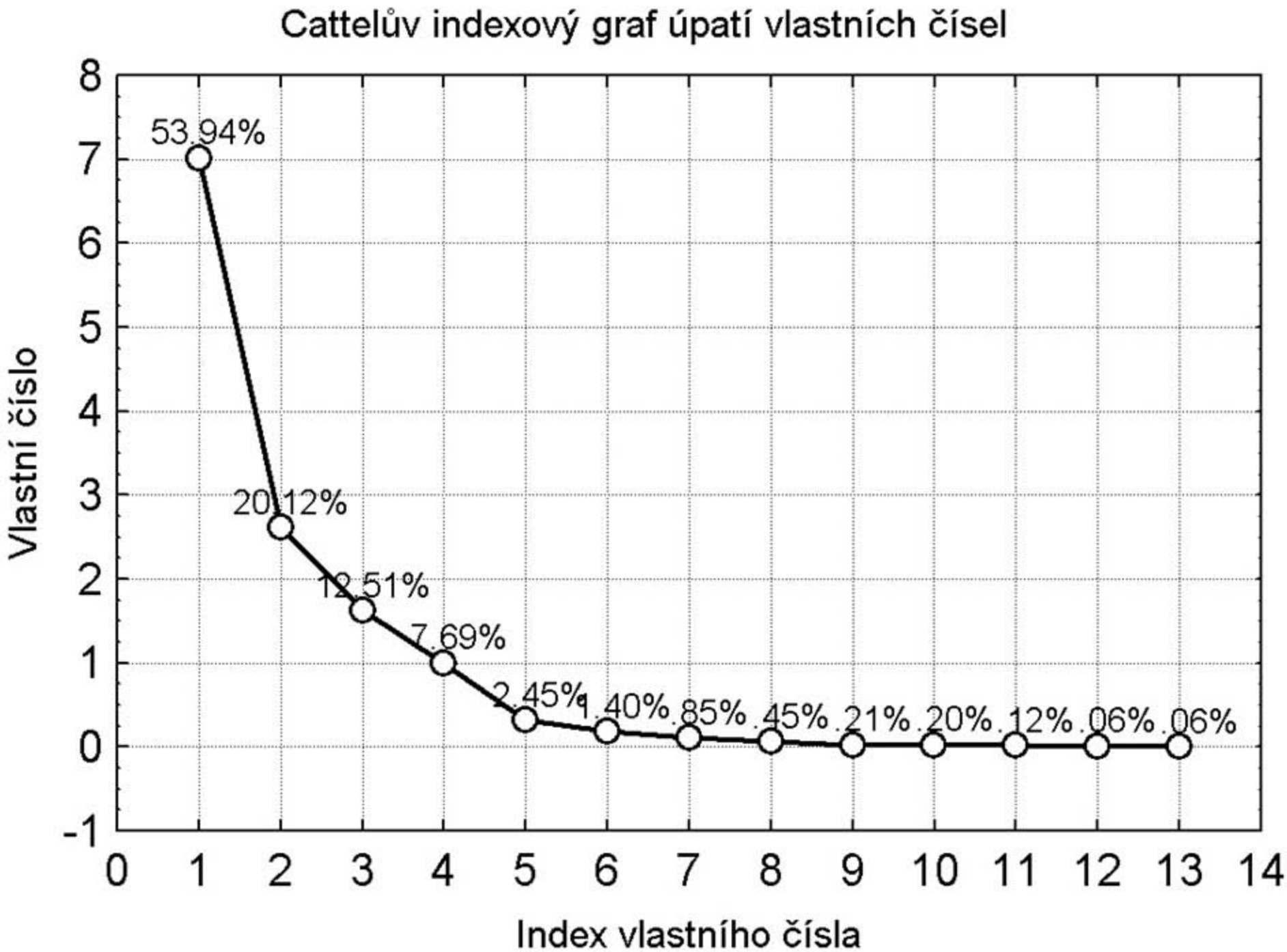
pro $m = 10$ existuje $m(m-1)/2 = 45$ diagramů,

pro $m = 11$ pak 55 diagramů,

pro $m = 12$ pak 66 diagramů, atd.

Vybíráme: diagramy y_1 vs. y_2, y_1 vs. y_3, y_1 vs. y_4 atd. - držíme se první hlavní komponenty y_1, protože v ní bývá nějvětší míra proměnlivosti v datech.


Finland
Albania
Norway
Sweden
UK
Greece
Denmark
USSR
Switzerland
Ireland
Romania
France
Italy
Belgium
Yugoslavia
Bulgaria
ortugal
Spain
W Germany
Netherlands
Czechoslovakia
E Germany
Austria
Poland
Hungary
2
1
0
-1
-2
-4
-3
-2
-1
0
1
2
Faktor 3: 12.53%
Faktor 2: 18.17%


Cattelův indexový graf úpatí vlastních čísel
Vlastní číslo
Index vlastního čísla
53.94%
20.12%
12.51%
7.69%
2.45%
1.40%
.85%
.45%
.21%
.20%
.12%
.06%
.06%
8
7
6
5
4
3
2
1
0
-1
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14

Výklad:

Indexový graf úpatí vlastních čísel

(Scree Plot)

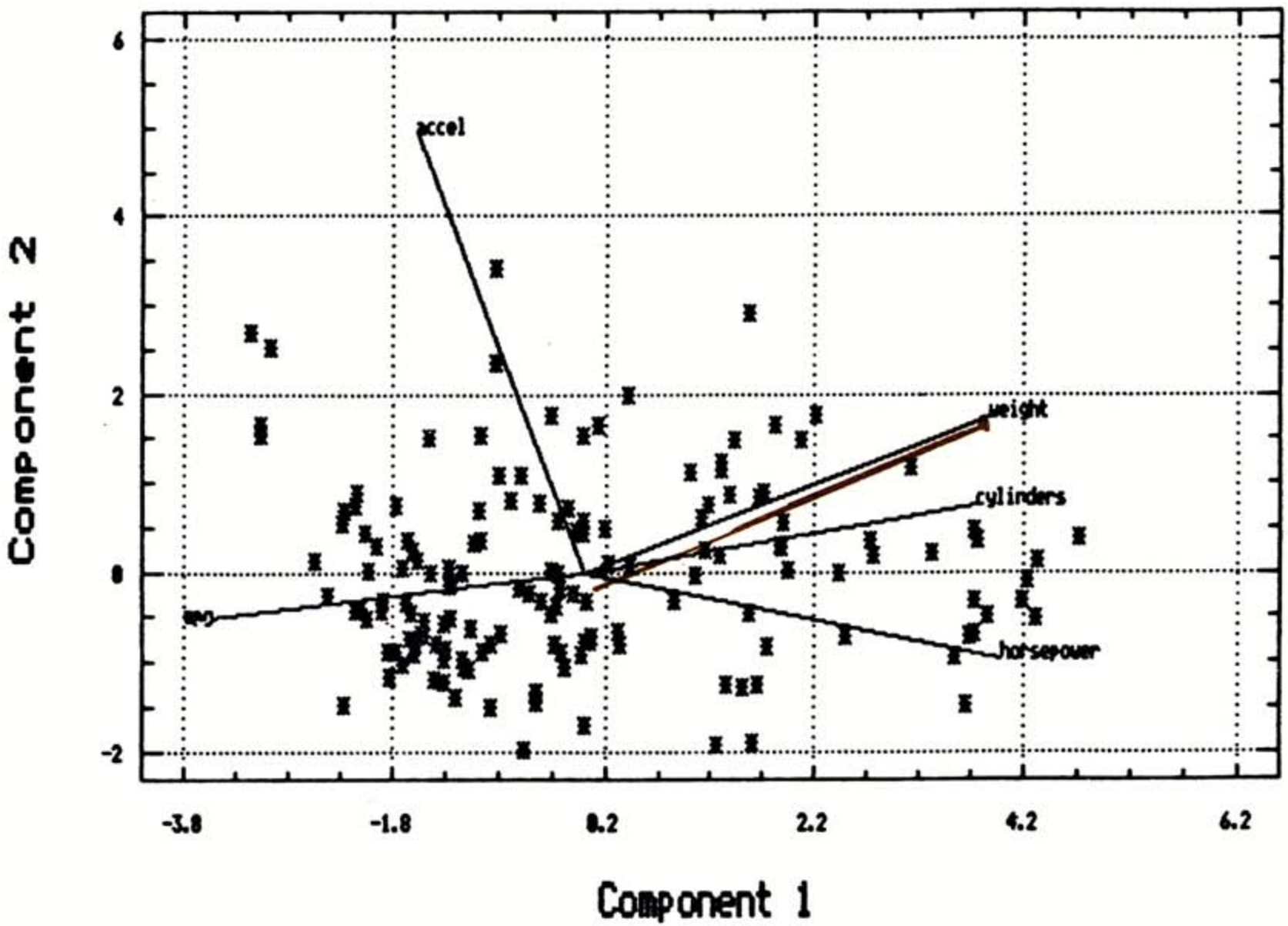
Je to sloupcový diagram vlastních čísel proti indexu *A*.

Zobrazuje: relativní velikost jednotlivých vlastních čísel.

Využití: k určení počtu *A* "užitečných" hlavních komponent. Graf úpatí se jeví nejobjektivnějším kritériem.

Kritérium “1": hrubším kritériem PC, jejichž vlastní číslo je větší než jedna. Graf úpatí se však jeví objektivnějším.

Biplot for First Two Principal Components



Dvojný graf (Biplot) kombinuje oba předchozí grafy.

Úhel mezi průvodiči x_j a x_k: je nepřímo úměrný velikosti korelace mezi proměnnými x_j a x_k: *Čím je menší úhel, tím je větší korelace.*

Délka souřadnice bodu: je úměrná příspěvku x_j do hlavní komponety, čili je úměrná komponentní váze.

Interakce: objekt je v blízkosti proměnné x_j, tzn. objekt "obsahuje" hodně této proměnné.

Umístění vpravo (nahoře) od nuly: Interakce proměnných a objektů vysvětluje umístění objektů vpravo od nuly na ose y_1 (či vlevo od nuly), resp. umístění nahoře od nuly (či dole od nuly) na ose y_2

Diagnostika metody PCA

Maticový graf rozptylových diagramů znaků slouží k získání počáteční informace o datech, zda data potřebují škálování. V PCA postupně provádíme:

1. ***Vyšetření indexového grafu úpatí vlastních čísel*** – z hrany úpatí v tomto diagramu se určí vhodný počet hlavních komponent.

2. ***Výpočet vlastních vektorů*** – vedle číselných hodnot se užívá i názorný čárový diagram hodnot vlastních vektorů, který přehledně informuje o relativním zastoupení původních znaků x_j , $j = 1, ..., m$, v hlavních komponentách.

3. ***Výpočet komponentních vah*** – matice párových korelačních koeficientů obsahující korelace původních znaků s hlavními komponentami. Uživatel nyní vybere pouze prvních k hlavních komponent a vytvoří tak model PCA.

4. ***Vyšetření grafu komponentních vah***.

5. ***Vyšetření rozptylového diagramu komponentního skóre.***

6. ***Vyšetření dvojného grafu***.

7. ***Vyšetření reziduí*** – rezidua objektů a rezidua proměnných by měla prokazovat dostatečnou těsnost proložení.

8. ***Určení významných původních znaků*** – je výhodné vyhledávat významné znaky, protože klasická metoda PCA umožňuje sice redukci počtu hlavních komponent, ale každá komponenta zůstává stále kombinací všech původních znaků.

Příklad 4.1/str. 70

NCSS 2007 Data - [E:\Moje\Učebnice 2005 M+M+H\CD v knížce\DATA\NCSS2002\P202Lide.50]

File Edit Data Analysis Graphics Tools Window Help

MA

	Osoba	Vyska	Hmotnost	Vlasy	Boty	Vek	Prijem	Pivo	Vino	Sex	Plavani	Puvod	IQ	C14	C15	C16	C17	C18	C19
1	MA	198	92	-1	48	48	45000	420	115	-1	98	-1	100						
2	MA	184	84	-1	44	33	33000	350	102	-1	92	-1	130						
3	MA	183	83	-1	44	37	34000	320	98	-1	91	-1	127						
4	FA	166	47	-1	36	32	28000	270	78	1	75	-1	112						
5	FA																		
6	FA																		
7	MA																		
8	MA																		
9	FA																		
10	FA																		
11	MA																		
12	FA																		
13	FA																		
14	FA																		
15	MA																		
16	MA																		
17	MB																		
18	MB																		
19	FB																		
20	FB																		
21	FB																		
22	MB																		
23	MB																		
24	MB																		
25	FB																		
26	FB																		
27	MB																		
28	MB																		
29	FB																		
30	FB																		
31	MB																		
32	FB																		

NCSS: Principal Components Analysis

File Run Analysis Graphics Tools Window Help

Loadings Plot | Storage | Template
Variables | Reports | Scores Plot

Input Variables

Variables: VYSKA-PUVOD

Data Input Format: Regular Data

Covariance Estimation Options

Robust Covariance Matrix Estimation

Robust Weight: 4.0

Missing Value Estimation: None

Maximum Iterations: 1000

Type of Matrix Used in Analysis

Matrix Type: Correlation

Factor (Component) Options

Factor Rotation: None

Factor Selection - Method: Percent of Eigenvalues

Factor Selection - Value: 100

QUICK ACCESS BUTTON:
This user programmable button loads the specified procedure. These buttons can be customized to load any procedure you want.

Click using the left mouse button to load the procedure.

Click using the right mouse button to bring up a window that will let you assign another procedure to this button.

Reset | Guide Me

Opt 3 Template Id:

Variable Info | Sheet1

This is the spreadsheet that lets you enter and edit your data.

NCSS Output - [NCSS: Principal Components Analysis Output]

File Edit View Format Window Help

Eigenvalues

No.	Eigenvalue	Individual Percent	Cumulative Percent	Scree Plot
1	6.419146	58.36	58.36	||||||||||
2	2.223585	20.21	78.57	|||||
3	1.615153	14.68	93.25	|||
4	0.341648	3.11	96.36	|
5	0.165305	1.50	97.86	|
6	0.099908	0.91	98.77	|
7	0.054995	0.50	99.27	|
8	0.027560	0.25	99.52	|
9	0.026044	0.24	99.76	|
10	0.019590	0.18	99.94	|
11	0.007066	0.06	100.00	|

Principal Components Report

Page/Date/Time 4 9/2/2009 8:07:40 AM
Database E:\Moje\Učebnice 2005 M+M+H\... ce\DATA\NCSS2002\P202Lide.S0

Eigenvectors

Factors

Variables	Factor1	Factor2	Factor3	Factor4	Factor5
Vyska	0.375504	0.136612	-0.075982	0.209396	-0.124824
Hmotnost	0.381965	0.107006	-0.067335	0.071354	-0.188040
Vlasy	-0.338957	-0.142959	0.076356	0.671338	-0.482799
Boty	0.377937	0.152420	-0.004811	0.172701	-0.031256
Vek	0.142687	-0.056872	0.722905	-0.028559	-0.165816
Prijem	0.190015	-0.286015	0.592485	0.047354	0.138717
Pivo	0.324611	-0.314807	-0.181770	0.204827	0.243374
Vino	-0.123733	0.563663	0.200652	0.419838	0.641715
Sex	-0.352771	-0.228706	-0.051692	0.320459	0.098946
Plavani	0.365118	0.112928	-0.138806	0.349952	-0.159201
Puvod	-0.143057	0.599416	0.122347	-0.155260	-0.402600

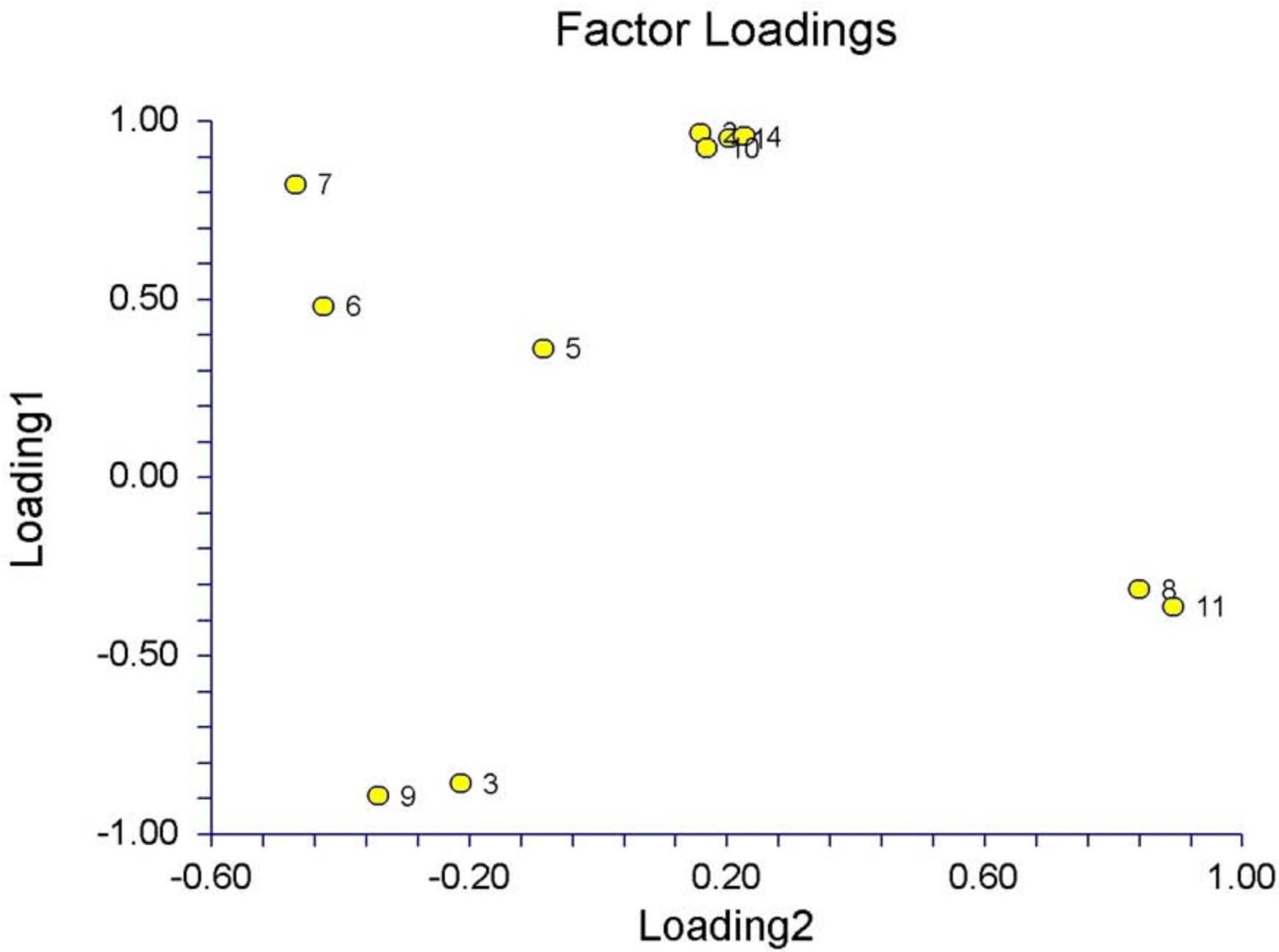
Factors

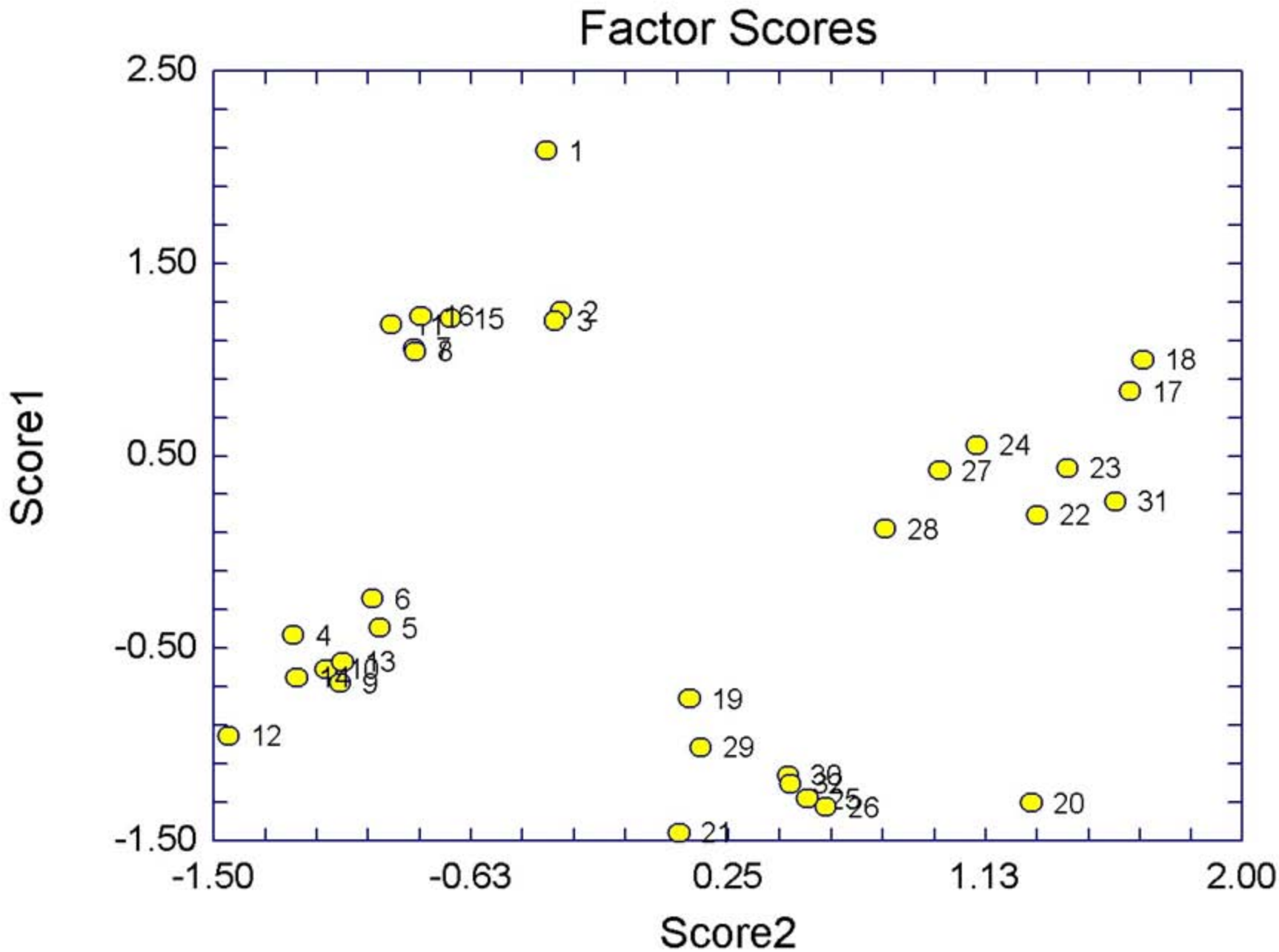
Variables	Factor6	Factor7	Factor8	Factor9	Factor10
Vyska	-0.272277	-0.118754	-0.628249	0.519994	0.139414
Hmotnost	0.240232	0.220100	-0.334306	-0.617979	-0.138333
Vlasy	0.377358	0.081416	-0.017897	0.125235	-0.059182
Boty	0.225158	-0.172031	0.282968	-0.153117	0.757138
Vek	-0.044573	-0.435111	0.193607	0.080468	-0.075481
Prijem	-0.122884	0.434312	-0.216562	-0.145135	0.010058
Pivo	0.182138	-0.566463	-0.030441	-0.127935	-0.426866
Vino	0.123739	0.040392	-0.025290	0.018606	-0.106744
Sex	-0.573098	-0.253379	-0.167812	-0.443739	0.273810
Plavani	-0.507034	0.258605	0.534549	0.020442	-0.276970
Puvod	-0.154793	-0.263994	-0.115270	-0.258851	-0.183238

Factors

Page 4/15 Line 181 Col 1

Start Total ... Doruč... 4 Mi... 4 NC... Zoner... Hledat v počítači 8:08 AM


Factor Loadings
Loading1
Loading2
1.00
0.50
0.00
-0.50
-1.00
-0.60
-0.20
0.20
0.60
1.00
7
6
5
2
4
14
10
8
11
9
3


Factor Scores
Score1
Score2
2.50
1.50
0.50
-0.50
-1.50
-1.50
-0.63
0.25
1.13
2.00
1
2
3
15
16
8
18
17
24
27
23
31
22
28
6
5
4
13
19
29
12
20
21
26

Příklad 4.2/str. 76

NCSS 2007 Data - [E:\Moje\Učebnice 2005 M+M+H\CD v knížce\DATA\NCSS2002\P402Hrach.S0]

Edit Data Analysis Graphics Tools Window Help

	Objekt	Aro	Slad	Med	Bez	Klas	Tvrd	Bel	Bar1	Bar2	Bar3	Slup	Ztr	C14	C15	C16	C17	C18	C19
1	B5	6.48	6.66	4.56	2.2	2.91	3.47	4.72	5.585	5.735	5.985	4.26	3.25						
2	C4	5.75	6.09	3.81	2.32	4.03	3.77	4.17	5.73	5.745	5.325	3.82	3.38						
3	B2	3.94	4.12	2.44	3.63	6.77	5.39	4.77	6.665	5.105	4.595	3.5	3.03						
4	D5	6.6	6.12	4.44	1.93	3.31	4.46	4.86	5.16	5.74	6.665	2.12	3.94						
5	D4																		
6	E2																		
7	B5																		
8	C5																		
9	C2																		
0	A4																		
1	D4																		
2	B1																		
3	D4																		
4	E4																		
5	B1																		
6	B5																		
7	D4																		
8	C2																		
9	A5																		
0	C3																		
1	B2																		
2	B5																		
3	D4																		
4	B1																		
5	A5																		
6	D4																		
7	A1																		
8	B3																		
9	C2																		
0	D3																		
1	B5																		
2	E3																		
3	C3																		
4	A5																		
5	D4																		
6	E4																		
7	B5																		
8	A4																		
9	C3																		
0	C4																		
1	D3																		
2	A1																		
3	A5	6.5	6.68	4.77	2.23	2.09	2.87	5.51	6.375	4.845	4.785	2.71	3.7						
4	D3	5.46	5.41	3.27	2.97	5.15	4.98	3.61	4.305	6.6	5.435	2.37	4.44						
5	A2	3.75	4.3	2.22	4.27	6.1	6.27	4.06	5.14	5.87	4.22	2.23	5.01						
6	C4	5.86	5.27	3.73	2.5	3.86	4.3	4.15	4.495	6.23	6.14	2.11	3.29						
7	A5	6.16	6.97	4.8	2.5	2.87	3.17	4.27	5.315	6.09	5.255	3.35	4.43						
8	B2	3.87	3.88	2.23	4.06	5.99	6.31	4.45	5.53	5.785	4.98	1.93	3.62						
9	B5	6.24	5.8	4.26	2.13	3.24	3.42	5.12	5.94	5.465	4.775	3.11	3.98						

Variable Info | Sheet1

SAVE the memory database to the CURRENTLY specified disk FILE.

NCSS: Principal Components Analysis

File Run Analysis Graphics Tools Window Help

Loadings Plot | Storage | Template
Variables | Reports | Scores Plot

Display the NCSS Navigator Window.

Input Variables

Variables: Aro-Ztr

Data Input Format: Regular Data

Covariance Estimation Options

Robust Covariance Matrix Estimation

Robust Weight: 4.0

Missing Value Estimation: None

Maximum Iterations: 1000

Type of Matrix Used in Analysis

Matrix Type: Correlation

Factor (Component) Options

Factor Rotation: None

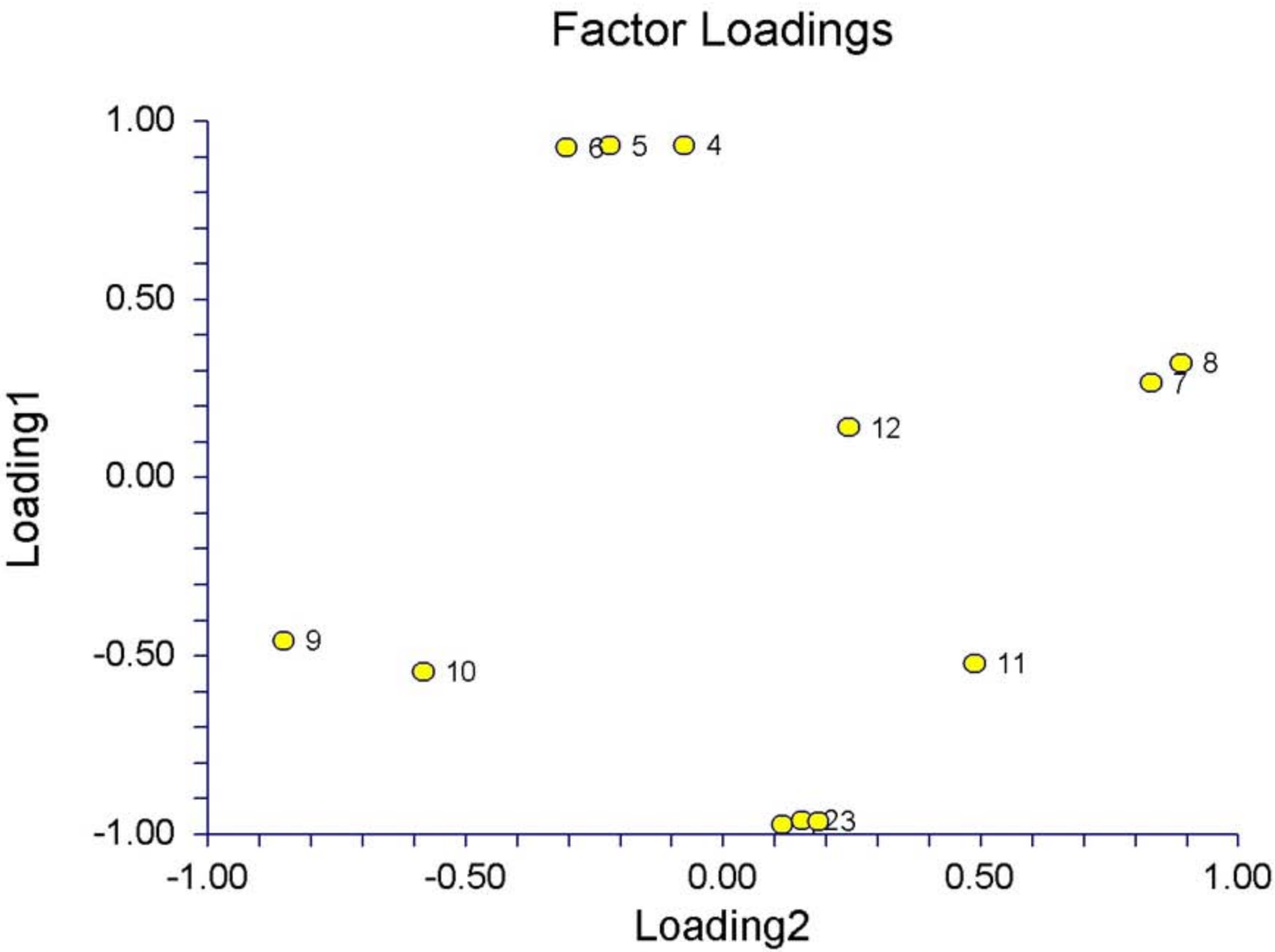
Factor Selection - Method: Percent of Eigenvalues

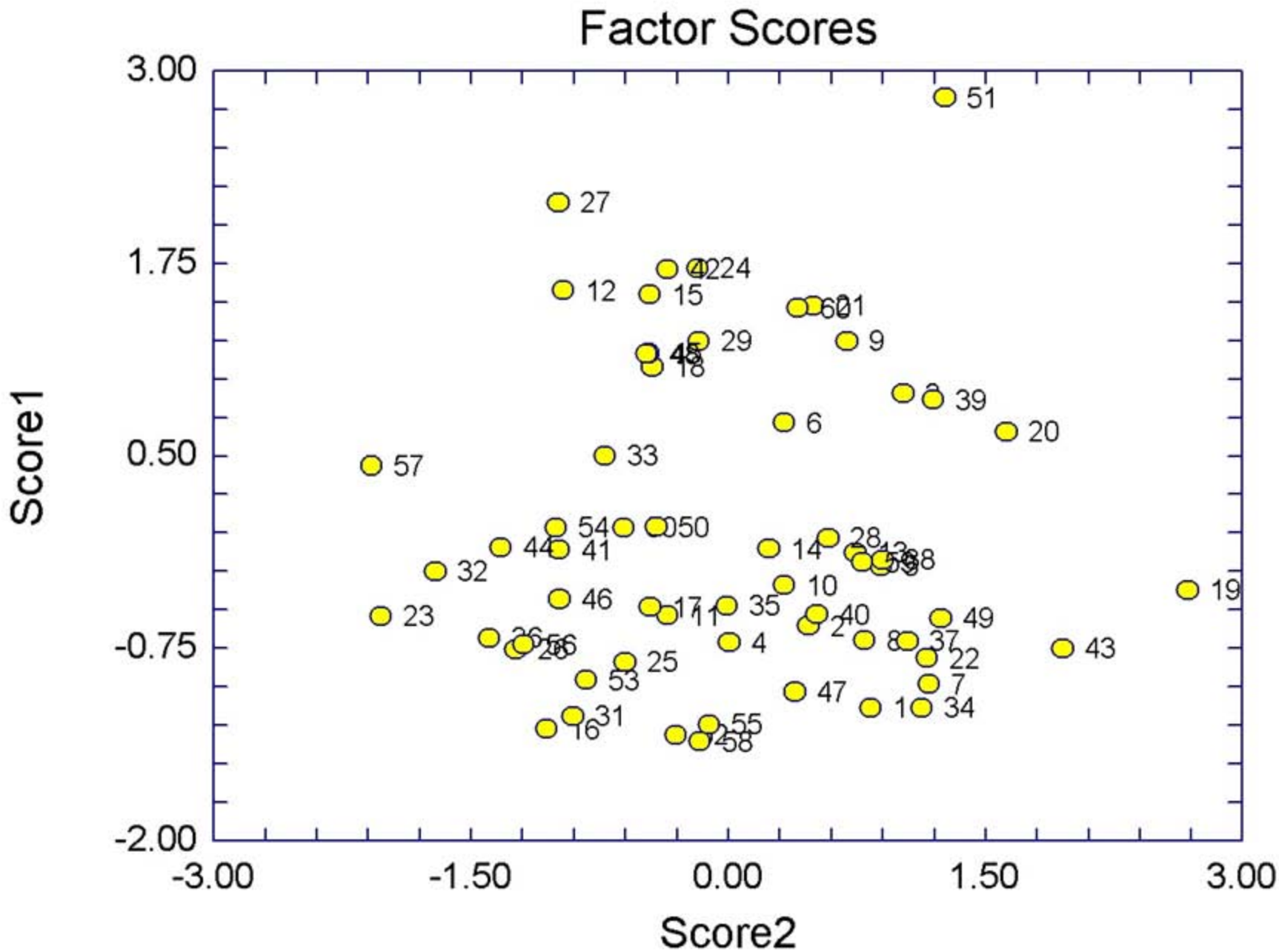
Factor Selection - Value: 100

Opt 2 Template Id:

Reset | Guide Me

Total ... | Doruč... | 4 Mi... | 4 NC... | Zoner... | Hledat v počítači | 8:15 AM


Factor Loadings
Loading1
Loading2
1.00
0.50
0.00
-0.50
-1.00
-1.00
-0.50
0.00
0.50
1.00
6
5
4
8
7
12
9
10
11
23


Factor Scores
Score1
Score2
3.00
1.75
0.50
-0.75
-2.00
-3.00
-1.50
0.00
1.50
3.00

Příklad 4.4/str. 84

NCSS 2007 Data - [E:\Moje\Učebnice 2005 M+M+H\CD v knížce\DATA\NCSS2002\P404Proteiny.S0]

File Edit Data Analysis Graphics Tools Window Help

Albania

	Stat	Cervene	Bile	Vejce	Mleko	Ryby	Obiln	Skrob	Orech	Ovoce	C11	C12	C13	C14	C15	C16	C17	C18	C19
1	Albania	10,1	1,4	0,5	8,9	0,2	42,3	0,6	5,5	1,7									
2	Austria	8,9	14	4,3	19,9	2,1	28	3,6	1,3	4,3									
3	Belgium	13,5	9,3	4,1	17,5	4,5	26,6	5,7	2,1	4									
4	Bulgaria	7,8	6	1,6	8,3	1,2													
5	Czechoslo	9,7	11,4	2,8	12,5	2													
6	Denmark	10,6	10,8	3,7	25	9,9													
7	E German	8,4	11,6	3,7	11,1	5,4													
8	Finland	9,5	4,9	2,7	33,7	5,8													
9	France	18	9,9	3,3	19,5	5,7													
10	Greece	10,2	3	2,8	17,6	5,9													
11	Hungary	5,3	12,4	2,9	9,7	0,3													
12	Ireland	13,9	10	4,7	25,8	2,2													
13	Italy	9	5,1	2,9	13,7	3,4													
14	Netherland	9,5	13,6	3,6	23,4	2,5													
15	Norway	9,4	4,7	2,7	23,3	9,7													
16	Poland	6,9	10,2	2,7	19,3	3													
17	Portugal	6,2	3,7	1,1	4,9	14,2													
18	Romania	6,2	6,3	1,5	11,1	1													
19	Spain	7,1	3,4	3,1	8,6	7													
20	Sweden	9,9	7,8	3,5	24,7	7,5													
21	Switzerlan	13,1	10,1	3,1	23,8	2,3													
22	UK	17,4	5,7	4,7	20,6	4,3													
23	USSR	9,3	4,6	2,1	16,6	3													
24	W Germar	11,4	12,5	4,1	18,8	3,4													
25	Yugoslavia	4,4	5	1,2	9,5	0,6													

NCSS: Principal Components Analysis

File Run Analysis Graphics Tools Window Help

Loadings Plot | Storage | Template
Variables | Reports | Scores Plot

Input Variables

Variables: Cervene-Ovoce

Data Input Format: Regular Data

Covariance Estimation Options

Robust Covariance Matrix Estimation

Robust Weight: 4.0

Missing Value Estimation: None

Maximum Iterations: 1000

Type of Matrix Used in Analysis

Matrix Type:

nent) Options

Factor Selection - Method: Percent of Eigenvalues

Factor Selection - Value: 100

VARIABLE SELECTION BUTTON:
Press this button to display a Variable selection window. You can also double-click directly in the box to the left of this button or simply type the name(s) or column number(s) in that box.

COMMENTS:
There are several ways to specify variables:

1. You can double click in the box to the left or single click this button to bring up a special Variable Selection window.

2. You can enter variable names or column position numbers directly in this box without using the Variable selection window. When entering variables this way, you can enter a range of columns such as 1, 2, 4-7 or X1, X3-X9, X14-X20.

Reset | Guide Me

Variables

Variable Selection List

Stat	C16	C31	C46
Cervene	C17	C32	C47
Bile	C18	C33	C48
Vejce	C19	C34	C49
Mleko	C20	C35	C50
Ryby	C21	C36	C51
Obiln	C22	C37	C52
Skrob	C23	C38	C53
Orech	C24	C39	C54
Ovoce	C25	C40	C55
C11	C26	C41	C56
C12	C27	C42	C57
C13	C28	C43	C58
C14	C29	C44	C59
C15	C30	C45	C60

Ok | Cancel | Help

Select one or more items from this list. Use the <Shift> key to enter contiguous lists. Use the <Ctrl> key to enter disjoint lists.

Variables Selected

Cervene-Ovoce

Select All | Clear | Store as Numbers

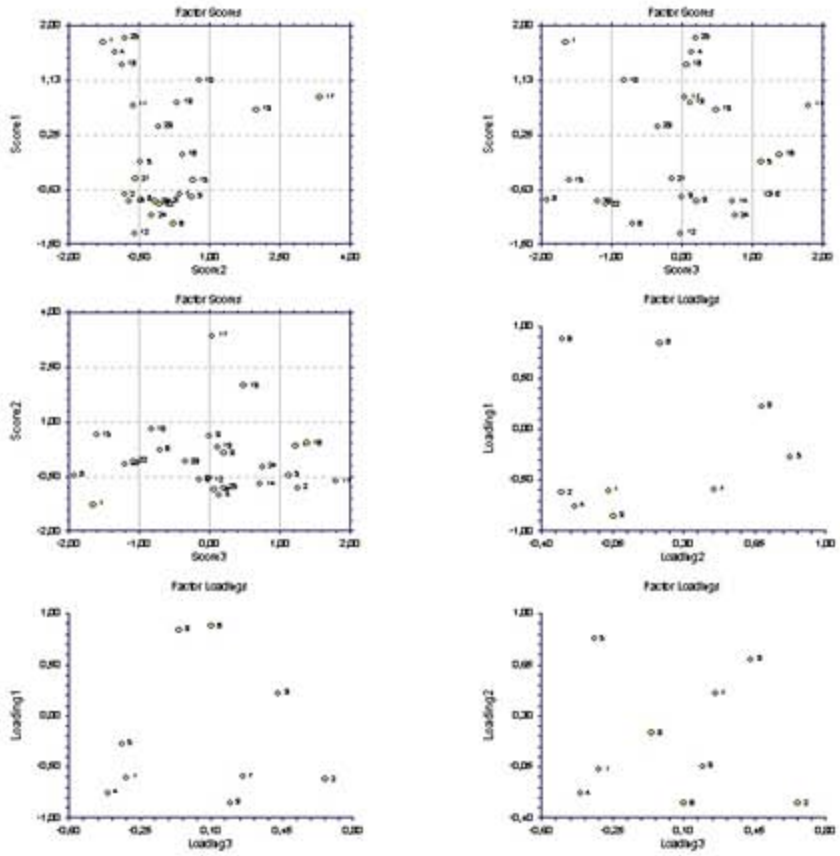
Start | Doručená pošt... | 3 NCSS Execu... | Total Command... | Hledat v počítači | 13:09

NCSS Output - [NCSS: Principal Components Analysis Output]

File Edit View Format Window Help

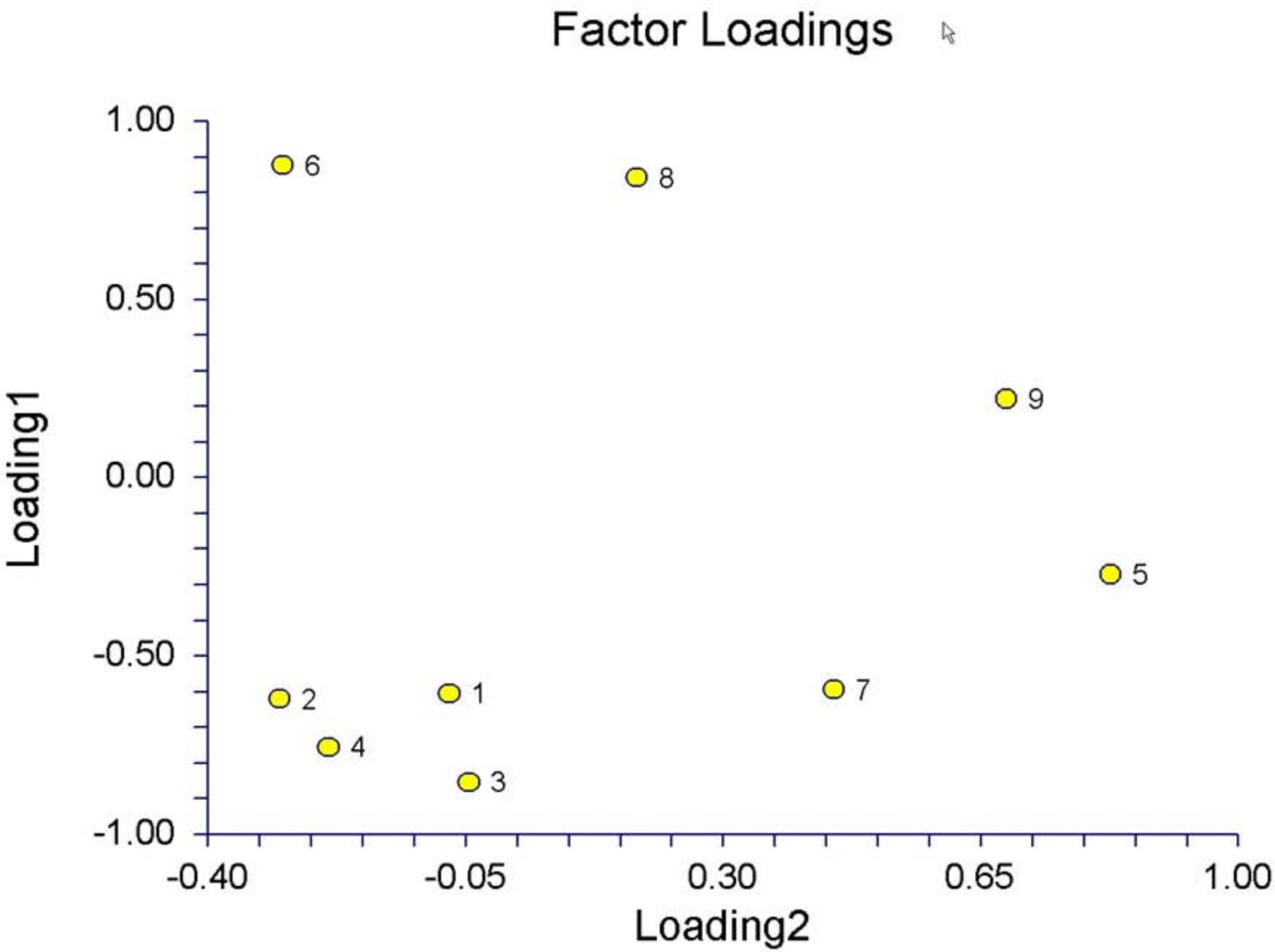
17	0,0892	-0,6012	0,8369
18	-0,2554	0,0789	1,0737
19	-1,1459	-0,6901	-1,5143
20	-1,0384	0,2118	-0,3443
21	0,9830	-1,5521	0,2114
22	-1,2491	0,7015	-0,4191
23	1,1016	0,1526	0,2912
24	-0,8765	-1,0458	-0,0803
25	0,1176	0,5668	0,4740

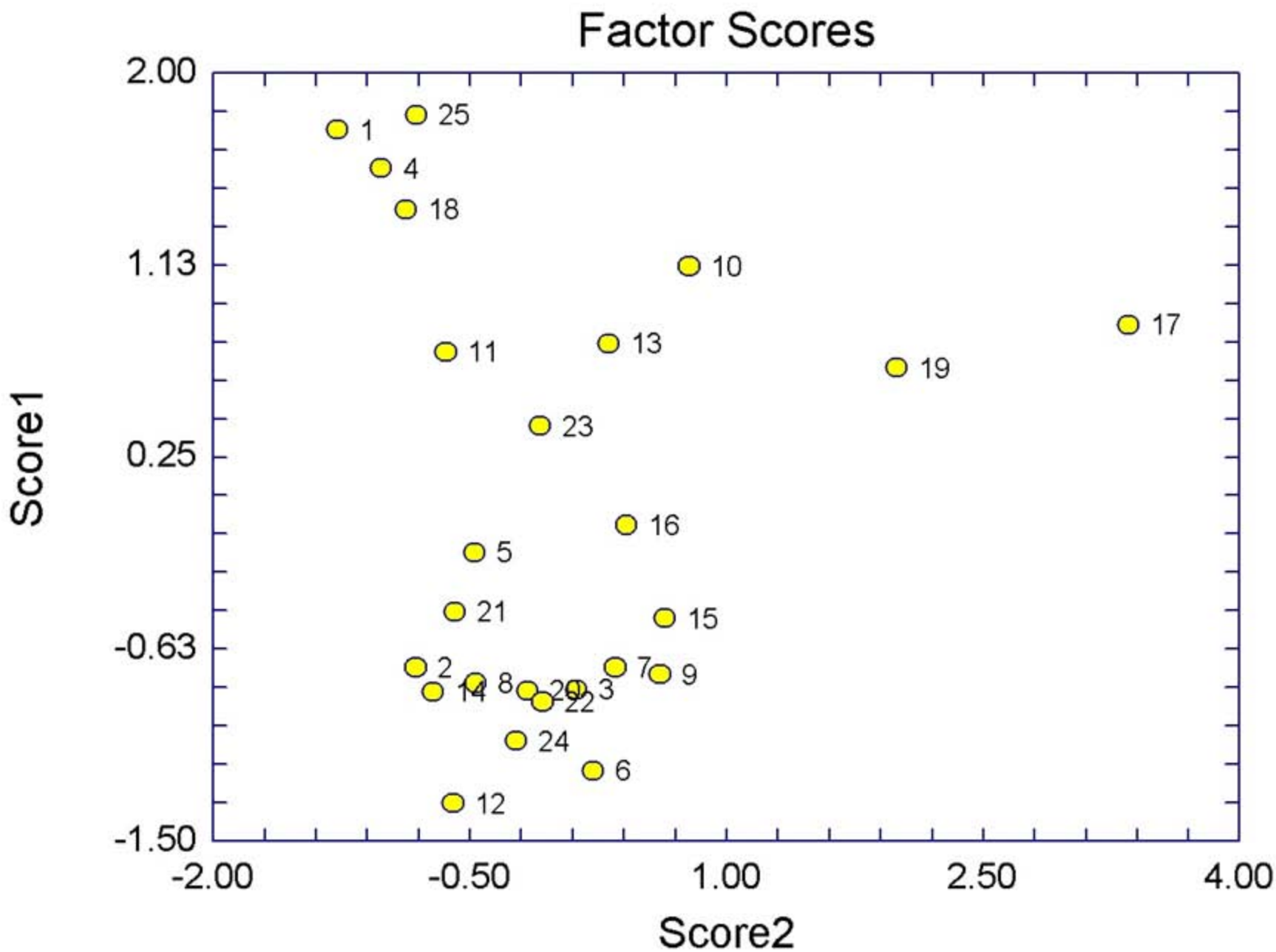
Plots Section



Page 10/10 Line 445 Col 1

Start Doručená pošt... 3 NCSS Exec... Total Command... Hledat v počítači 13:10


Factor Loadings
1.00
0.50
0.00
-0.50
-1.00
Loading1
-0.40
-0.05
0.30
0.65
1.00
Loading2
6
8
9
5
2
1
7
4
3


Factor Scores
Score1
Score2
2.00
1.13
0.25
-0.63
-1.50
-2.00
-0.50
1.00
2.50
4.00
1
25
4
18
10
17
11
13
19
23
16
5
21
15
2
7
9
14
8
20
3
22
24
6
12

Příklad 4.5/str. 87

NCSS 2007 Data - [E:\Moje\Učebnice 2005 M+M+H\CD v knížce\DATA\NCSS2002\P405Guiseppe.S0]

File Edit Data Analysis Graphics Tools Window Help

Atropine

	Objekt	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13	x14	x15	x16	x17	x18
1	Atropine	20	16	29	23	62	33	4	2	13	47	25	42	18	12	0	0	5	8
2	Biperide	91	90	87	68	92	87	73	72	64	85	81	86	68	94	11	40	40	65
3	Caffeine	55	42	52	68	77	54	8	5	13	60	30	51	41	20	13	12	54	57
4	Cocaine	81	82	81	71	87													57
5	Codeine	38	31	44	39	71													21
6	Cyclizin	71	72	80	64	85													54
7	Diazepam	76	79	80	78	85													87
8	Ketamine	77	79	80	76	86													79
9	Lignocaine	77	79	80	73	86													84
10	Lorazepam	47	34	53	77	73													79
11	Mebeveri	85	90	90	65	90													53
12	Methadon	85	84	88	48	89													29
13	Morphine	18	9	15	39	56													18
14	Naloxone	48	40	62	75	79													77
15	Papaverine	68	66	76	79	88													80
16	Pentazoc	72	66	81	65	87													59
17	Phenacet	64	58	62	79	87													84
18	Phenazon	66	53	70	83	86													74
19	Prazepam	81	83	86	83	88													88
20	Procaine	64	60	70	65	82													53

NCSS: Principal Components Analysis

File Run Analysis Graphics Tools Window Help

Loadings Plot | Storage | Template
Variables | Reports | Scores Plot

Run this procedure with the current settings and view the output report and plots.

Input Variables

Variables: x1-x18

Data Input Format: Regular Data

Covariance Estimation Options

Robust Covariance Matrix Estimation

Robust Weight: 4.0

Missing Value Estimation: None

Maximum Iterations: 1000

Type of Matrix Used in Analysis

Matrix Type: Correlation

Factor (Component) Options

Factor Rotation: None

Factor Selection - Method: Percent of Eigenvalues

Factor Selection - Value: 100

Opt 2 Template Id:

Reset | Guide Me

Variable Info | Sheet1

Paste the data contained on the clipboard. This may be data from a cut or copy operation or from another application. It must be tab delimited.

NCSS Output - [NCSS: Principal Components Analysis Output]

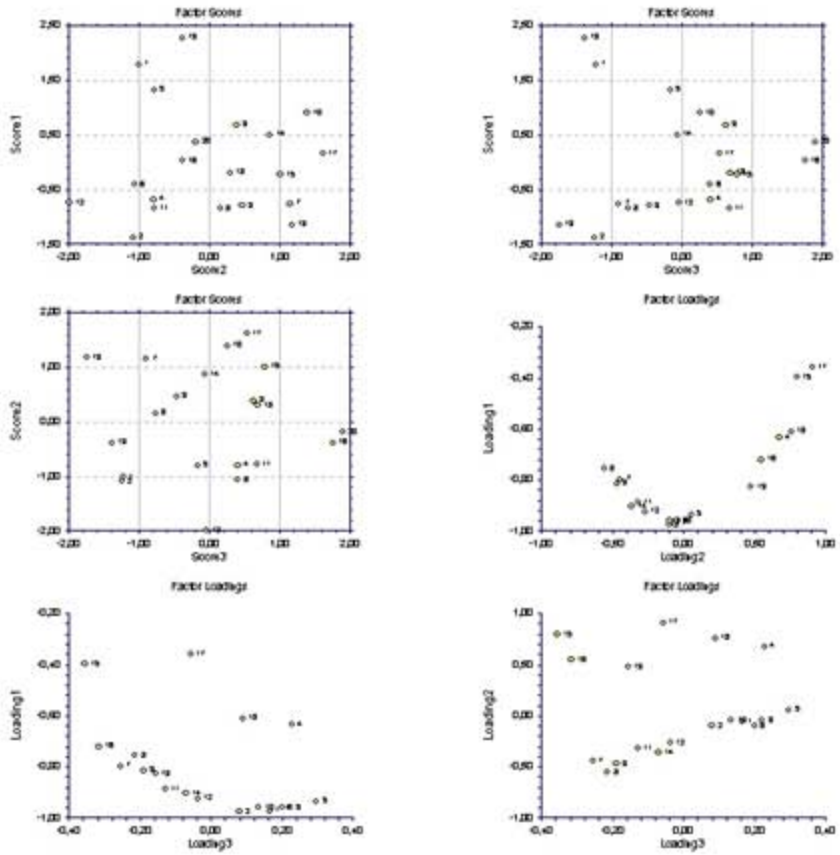
File Edit View Format Window Help

18	0,0562	-0,6232	-0,1885	-1,5294	-0,6220	-0,1620
19	0,0517	-0,1654	0,5658	1,2474	0,8854	0,7163
20	-0,9155	0,3822	-0,2885	0,7986	-0,7211	1,1154

Principal Components Report

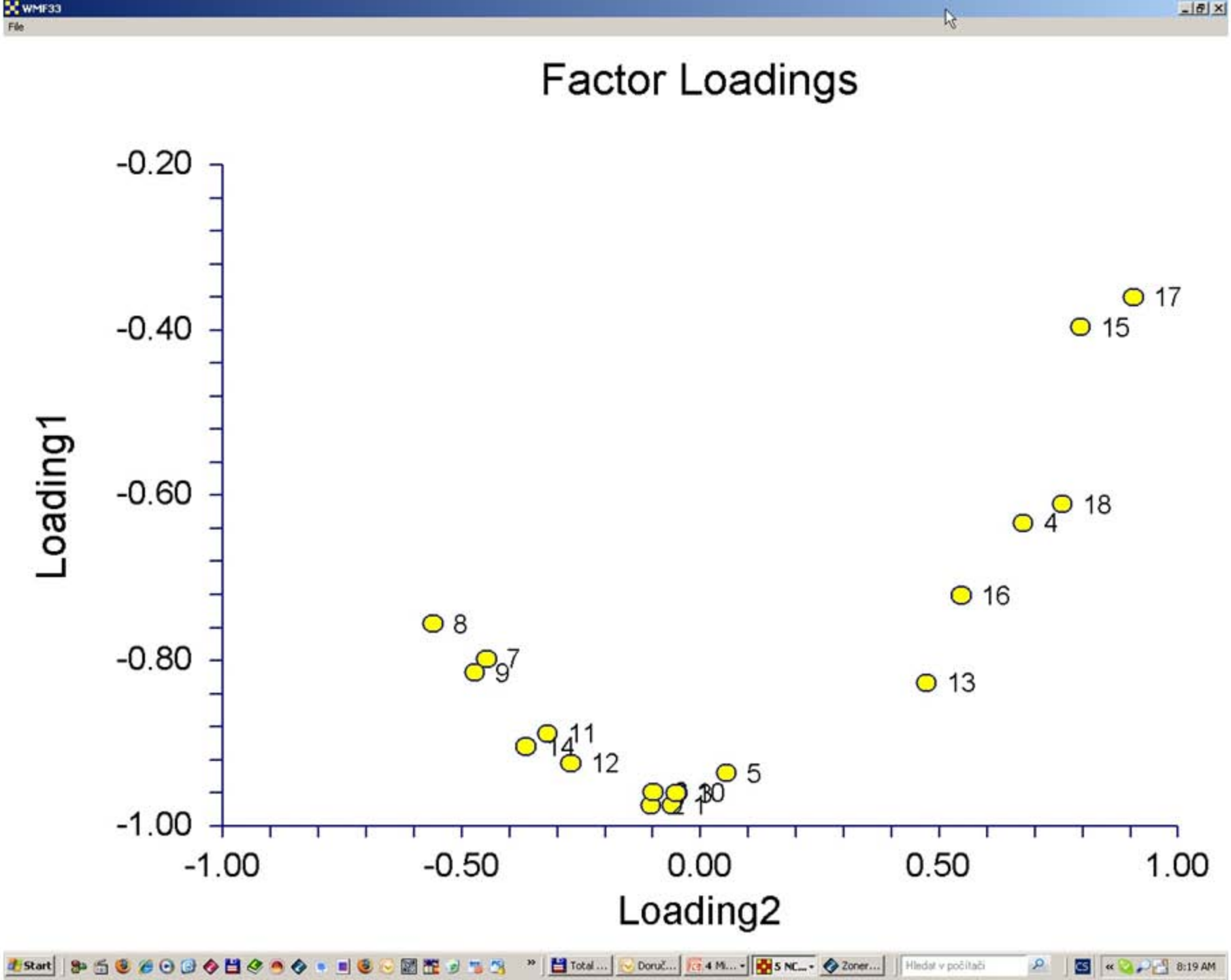
Page/Date/Time 24 31.8.2009 13:12:17
Database E:\Moje\Učebnice 2005 M+M+H\... ATANCSS2002\P405Guiseppe.S0

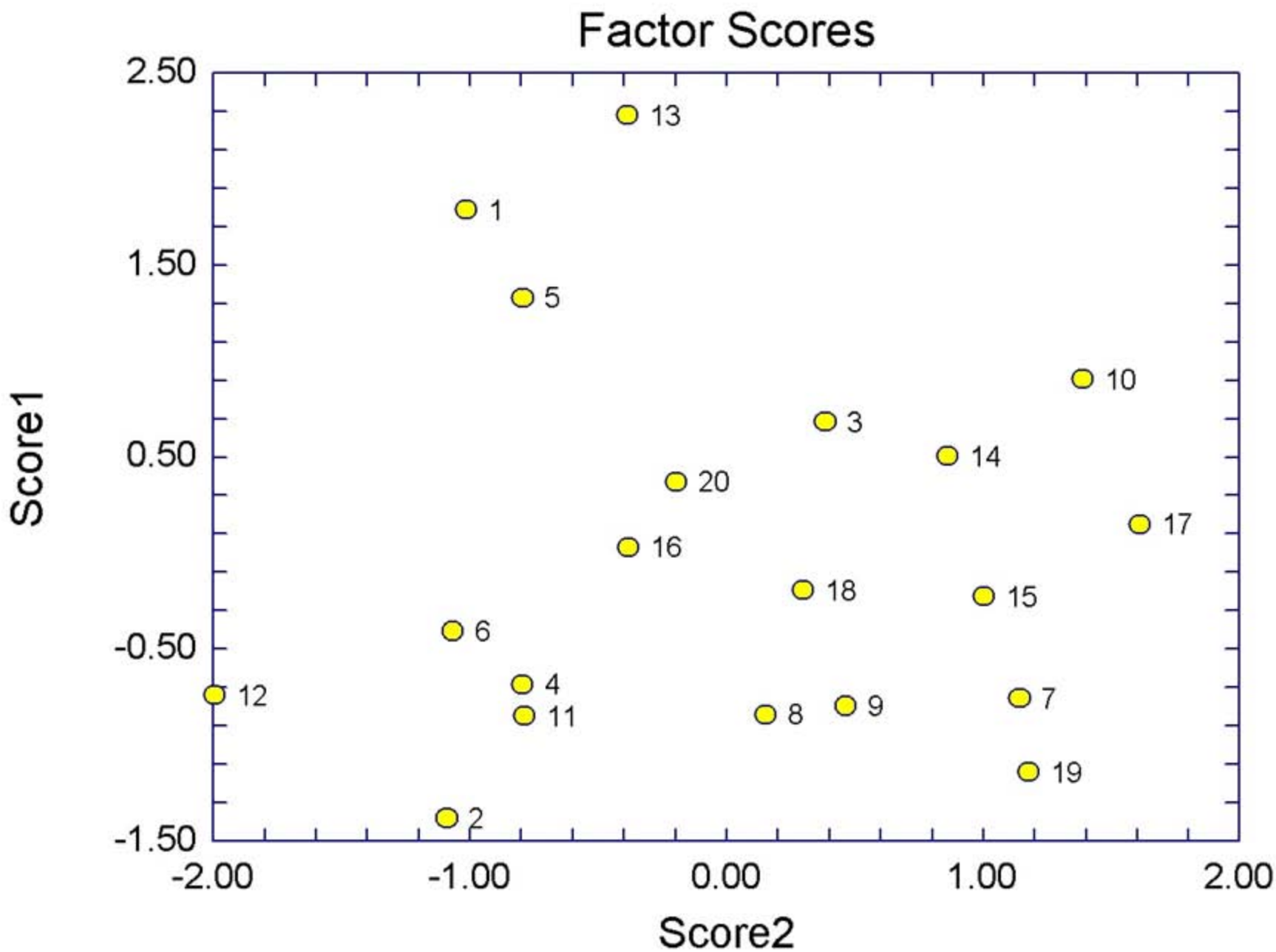
Plots Section



Page 24/24 Line 1159 Col 1

Doručená pošt... 3 NCSS Exec... Total Command... Hledat v počítači 13:12


WMF33
File
Factor Loadings
Loading1
Loading2
-0.20
-0.40
-0.60
-0.80
-1.00
-1.00
-0.50
0.00
0.50
1.00
17
15
18
4
16
8
7
9
13
11
14
12
5
30
Start
Total ...
Doruč...
4 Mi...
5 NC...
Zoner...
Hledat v počítači
8:19 AM


Factor Scores
Score1
Score2
2.50
1.50
0.50
-0.50
-1.50
-2.00
-1.00
0.00
1.00
2.00
13
1
5
10
3
14
20
17
16
18
15
6
4
12
11
8
9
7
19
2

Příklad 4.6/str. 90

NCSS 2007 Data - [E:\Moje\Učebnice 2005 M+M+H\CD v knížce\DATA\NCSS2002\P406Misce.S0]

File Edit Data Analysis Graphics Tools Window Help

	Objekt	Dtela	Stela	Dpkridlo	Dzkridlo	Pruduchy	Dtyk1	Dtyk2	Dtyk3	Dtyk4	Dtyk5	Ptykostn	Dpclnohy	Tibium	Dstehna	Sosak	Kladelko	Ptrnu	Ritotvor
1	1	21,2	11	7,5	4,8	5	2	2	2,8	2,8	3,3	3	4,4	4,5	3,6	7	4	8	0
2	2	20,2	10	7,5	5	5	2,3	2,1	3	3	3,2	5	4,2	4,5	3,5	7,6	4,2	8	0
3	3	20,2	10	7	4,6	5	1,9	2,1	3	2,5	3,3	1	4,2	4,4	3,3	7	4	6	0
4	4	22,5	8,8	7,4	4,7	5													0
5	5	20,6	11	8	4,8	5													0
6	6	19,1	9,2	7	4,5	5													0
7	7	20,8	11,4	7,7	4,9	5													0
8	8	15,5	8,2	6,3	4,9	5													0
9	9	16,7	8,8	6,4	4,5	5													0
10	10	19,7	9,9	8,2	4,7	5													0
11	11	10,6	5,2	3,9	2,3	4													1
12	12	9,2	4,5	3,7	2,2	4													1
13	13	9,6	4,5	3,6	2,3	4													1
14	14	8,5	4	3,8	2,2	4													1
15	15	11	4,7	4,2	2,3	4													1
16	16	18,1	8,2	5,9	3,5	5													1
17	17	17,6	8,3	6	3,8	5													1
18	18	19,2	6,6	6,2	3,4	5													1
19	19	15,4	7,6	7,1	3,4	5													1
20	20	15,1	7,3	6,2	3,8	5													1
21	21	16,1	7,9	5,8	3,7	5													1
22	22	19,1	8,8	6,4	3,9	5													1
23	23	15,3	6,4	5,3	3,3	5													1
24	24	14,8	8,1	6,2	3,7	5													1
25	25	16,2	7,7	6,9	3,7	5													1
26	26	13,4	6,9	5,7	3,4	5													1
27	27	12,9	5,8	4,8	2,6	5													1
28	28	12	6,5	5,3	3,2	5													1
29	29	14,1	7	5,5	3,6	5													1
30	30	16,7	7,2	5,7	3,5	5													1
31	31	14,1	5,4	5	3	5													1
32	32	10	6	4,2	2,5	5													1
33	33	11,4	4,5	4,4	2,7	5													1
34	34	12,5	5,5	4,7	2,3	5													0
35	35	13	5,3	4,7	2,3	5													1
36	36	12,4	5,2	4,4	2,6	5													1
37	37	12	5,4	4,9	3	5													1
38	38	10,7	5,6	4,5	2,8	5													1
39	39	11,1	5,5	4,3	2,6	5													1
40	40	12,8	5,7	4,8	2,8	5													1

Variable Info | Sheet1

This is the spreadsheet that lets you enter and edit your data.

NCSS: Principal Components Analysis

File Run Analysis Graphics Tools Window Help

Loadings Plot | Storage | Template
Variables | Reports | Scores Plot

Input Variables

Variables: Dtela-Phacku

Data Input Format: Regular Data

Covariance Estimation Options

Robust Covariance Matrix Estimation

Robust Weight: 4.0

Missing Value Estimation: None

Maximum Iterations: 1000

Type of Matrix Used in Analysis

Matrix Type: Correlation

Factor (Component) Options

Factor Rotation: None

Factor Selection - Method: Percent of Eigenvalues

Factor Selection - Value: 100

FACTOR SELECTION METHOD:
This option specifies which of the following three methods is used to set the number of factors retained in the analysis. This sets the interpretation of the Factor Selection Value.

--PERCENT OF EIGENVALUES
Specify the total percent of variation that must be accounted for. Enough factors will be included to account for this percentage (or slightly greater) of the variation in the data.

--NUMBER OF FACTORS
Specify the number of factors to keep.

--EIGENVALUE CUTOFF
Specify the minimum eigenvalue amount. All factors whose eigenvalues are greater than or equal to this value will be retained. Older statistical texts suggest that you should only keep factors whose eigenvalues are greater than one.

Reset | Guide Me

Opt 6 Template Id:

Start | Doručená pošt... | 3 NCSS Exec... | Total Command... | Hledat v počítači | 13:13

NCSS Output - [NCSS: Principal Components Analysis Output]

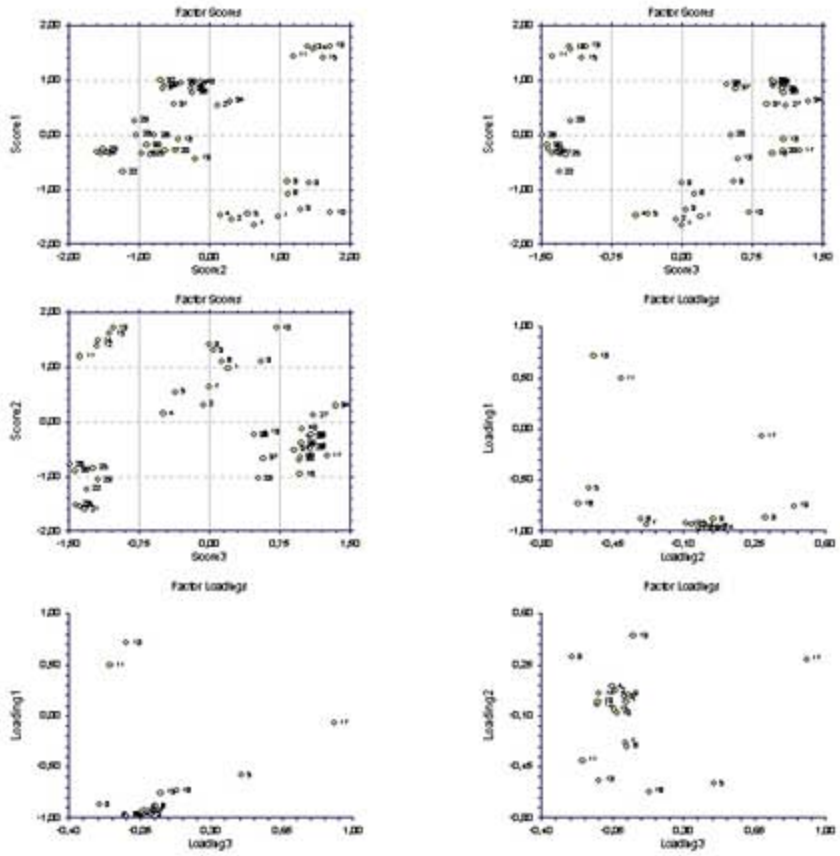
File Edit View Format Window Help

36	-2,1314
37	0,0988
38	-0,5884
39	-1,4358
40	0,6177

Principal Components Report

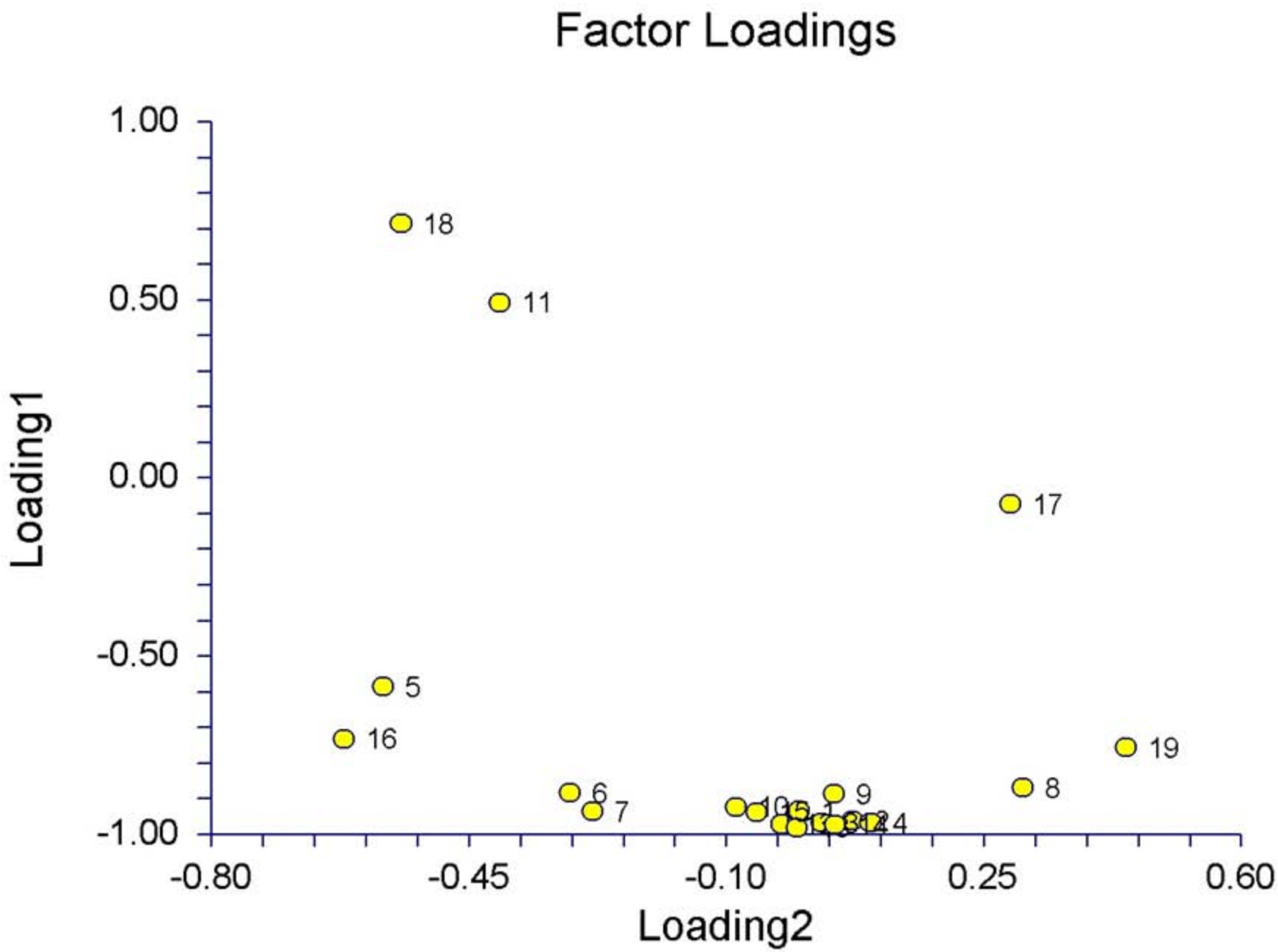
Page/Date/Time 27 31.8.2009 13:13:43
Database E:\Moje\Učebnice 2005 M+M+H\... e\DATA\NCSS2002\P406Msice.S0

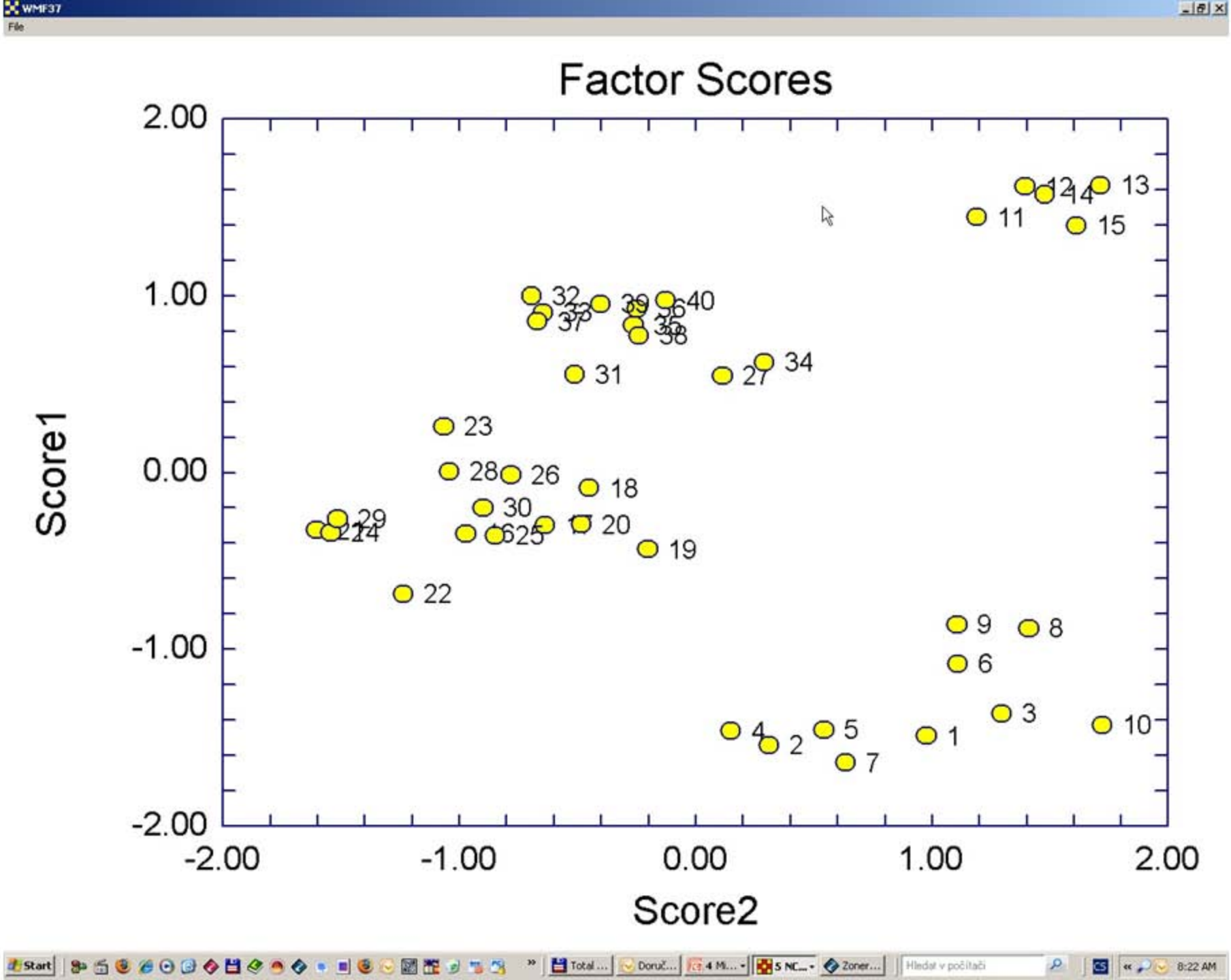
Plots Section



Page 27/27 Line 1341 Col 1

Start Doručená pošt... 3 NCSS Exec... Total Command... Hledat v počítači 13:14


Factor Loadings
Loading1
Loading2
1.00
0.50
0.00
-0.50
-1.00
-0.80
-0.45
-0.10
0.25
0.60
18
11
17
5
16
19
6
7
9
8
4


WMF37
File
Factor Scores
Score1
Score2
2.00
1.00
0.00
-1.00
-2.00
-2.00
-1.00
0.00
1.00
2.00
13
12
14
11
15
32
40
37
38
34
31
27
23
28
26
18
30
29
20
19
22
9
8
6
3
10
4
2
5
1
7
Start
Total ...
Doruč...
4 Mi...
5 NC...
Zoner...
Hledat v počítači
8:22 AM