Odhalení skryté struktury
a vnitřních vazeb dat metodami PCA, FA a CLU
vícerozměrné statistické analýzy

Internal bounds and hidden structure of the metallurgic data with the use of Multivariate Data Analysis MDA

Prof. RNDr. Milan Meloun, DrSc.,

Katedra analytické chemie, Univerzita Pardubice, 532 10 Pardubice,
milan.meloun@upce.cz

Ing. Roman Lisztwan,

Třinecké železárny, 739 70 Třinec, roman.lisztwan@trz.cz

Určení vzájemných vazeb

(a) *struktura a vazby v proměnných*
(b) *struktura a vazby v objektech*

(1) Hledání struktury v proměnných (metrická škála): faktorová analýza FA, analýza hlavních komponent PCA a shluková analýza.

(2) Hledání struktury v objektech (metrická škála): shluková analýza.

(3) Hledání struktury v objektech (metrická i nemetrická škála): vícerozměrné škálování.

(4) Hledání struktury v objektech (nemetrická škála): korespondenční analýza.

(5) Hledání lineárních vícerozměrných modelů (metrická i nemetrická škála): většina metod vícerozměrné statistické analýzy, kde závisle proměnné se uvažují jako lineárn kombinace nezávisle proměnných.

Internal bounds and hidden structure in variables

Internal bounds and hidden structure in objects

1. Search for hidden structure among variables (metric scale): Factor analysis FA, Principal component analysis PCA, Cluster analysis CLU.

2. Search for hidden structure among objects (metric scale): Cluster analysis CLU.

3. Search for hidden structure among objects (metric and nonmetric scale): Correspondence analysis CORRA.

4. Search for hidden structure among variables and objects (metric and nonmetric scale): Multidimensional Data Scaling MDS.

5. Search for linear bounds among variables or objects: regression techniques.

Postup analýzy vícerozměrných dat

1. Standardizace: analýze vždy předchází standardizace čili škálování proměnných.

2. Odhady parametrů polohy, rozptýlení, tvaru a intenzita vztahu mezi proměnnými:
Vyčíslení výběrové střední hodnoty každé proměnné.
Odhad kovarianční matice ***S*** a její normované podoby - korelační matice ***R*.**
Odhadu vícerozměrné šikmosti a vícerozměrné špičatosti.
Matice ***R*** obsahuje Pearsonovy párové korelační koeficienty, které se diskutují.

3. Exploratorní analýza dat EDA:
(a) Podobnost objektů **vizuálními rozptylovými diagramy** (casement plot, draftsman plot, symbolových a profilových grafů (hvězdičky, sluníčka, obličeje, křivky, stromy),

(b) Nalezení **vybočujících objektů** nebo vybočujících proměnných, mnohdy nevhodných k analýze,

(c) Testy **předpokladů lineárních vazeb**,

(d) Testy **předpokladů o datech** (normalitu, nekorelovanost, homogenitu).
Ověřování normality založené na vícerozměrné šikmosti a vícerozměrné špičatosti.

4. Určení vhodného počtu latentních proměnných:

a) **Matice *S*** nebo ***R*** se rozloží na vlastní čísla a vlastní vektory.

b) **Indexový graf úpatí vlastních čísel (Scree plot):** určí vhodný počet latentních proměnných, které ještě dostatečně popisují proměnlivost v datech.

c) Když se latentní proměnné podaří **pojmenovat** a dát jim i fyzikální, biologický či jiný věcný význam, jedná se o faktory. Jinak jde o hlavní komponenty.

5. Určení struktury v proměnných (PCA a FA):

a) Graf komponentních vah (Plot of components weights, loadings): hledání struktury a vzájemných vazeb (korelace) proměnných se provede v grafu

b) Rozptylový diagram komponentního skóre (Scatterplot): hledání struktury v objektech a třídění objektů do shluků.

c) Dvojný graf (Biplot) je přehledným spojením obou předešlých grafů a ukáže interakci objektů a proměnných.

Procedure of Multivariate Data Analysis

1. Standardization of data.

2. Estimates of parameters of location, spread and shape and the intensity of internal bounds among variables:

Estimate of the covariance matrix ***S*** and its normalized form – the correlation matrix ***R*.**
Estimate of multivariate skewness and multivariate kurtosis.
Interpretation of the Pearson correlation coefficients in the correlation matrix ***R***.

3. Exploratory data Analysis EDA:

(a) Similarity with scatter diagrams (casement plot, draftsman plot), the symbol graphs, profile graphs (stars plot, sun-ray plot, Cernoff faces plot, curves plot, trees plot),

(b) Identification of outliers not suitable for MDA analysis.

(c) Tests of assumptions of the linear bounds in data.

(d) Tests of assumptions about the sample data (normality, multicolinearity, homogeneity).

4. Determination of a number of latent variables:

a) **Matrix *S*** or ***R*** is decomposed into eigennumbers and eigenvectors.

b) **Scree plot**: determines a convenient number of latent variables, which still describe a variability in data.

c) Latent variables are interpreted and named according to their biological, chemical or physical meaning.

5. Determination of structure in variables and objects (PCA a FA):

a) **Plot of components weights, loadings**: searches for a hidden structure and internal bounds among variables.

b) **Scatterplot of components scores**: searches for a hidden structures and internal bounds among the objects and classify objects into the clusters.

c) **Biplot:** represents a connection of both previous two plots and shows an interaction of variables and objects.

Multivariate data analysis ***MDA*** deals with ***Objects*** (samples, individuals, molecules, ...) described by ***Variables*** (quantities, parameters, biological activities, etc.).

It searches for relationship among objects, among variables, and among objects and variables.

MDA is based on the latent variables being formed as the linear transformation of the original variables.

The source data matrix contains variables in *m* columns and *n* rows.

Before a data treatment the data are scaled.

The ***principal components analysis*** reduces dimensionality and presents objects in two or three dimensions.

The *plot of components weight* shows hidden structure among variables while the *scatterplot of component score* shows the hidden structure of objects.

Posouzení chemické homogenity v kruhové tyči CrNi oceli

Test of Chemical Homogeneity of Cr-Ni Steel (H404)

Chromniklová ocel 50 × 1000 mm (Poldi Kladno) rozřezána na 32 vzorků 50 × 30 mm. Každý označen číslem při zachování orientace v tyči, k testu homogenity 16 vzorků. Emisní spektroskopií s jiskrovým buzením na každém vzorku analýza na 8 expozicích.

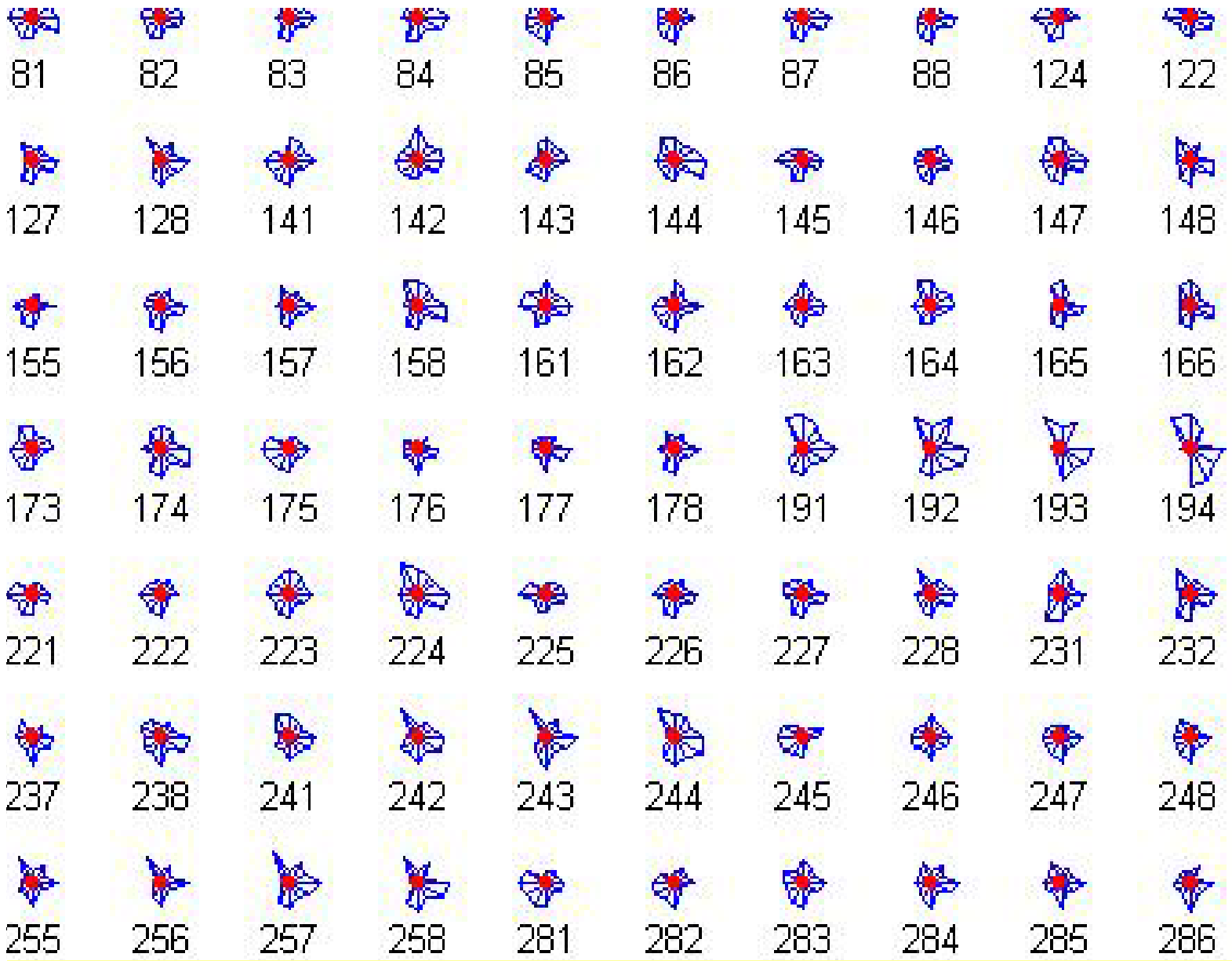
Cílem: k hodnocení homogenity materiálu omezený počet znaků a odhalit trendy v chemickém složení.

Data: Výběr *TYC* má 16 x 8 vzorků v řádcích (content of element): **I** pořadové číslo oceli, **C** obsah C, **MN** obsah Mn, **SI** obsah Si, **P** obsah P, **S** obsah S, **CU** obsah Cu, **CR** obsah Cr, **NI** obsah Ni, **AL** obsah Al, **MO** značí obsah Mo, **TI** obsah Ti, **B** obsah B.

I	**C**	**MN**	**SI**	**P**	**S**	**CU**	**CR**	**NI**	**AL**
	MO	**TI**	**B**						
31	0,0510	0,1531	1,4636	27,1810	28,5060	0,0980	1,7085	0,7280	1,9752
	1,8461	2,5478	0,0620						
32	0,0510	0,1537	1,4523	27,1220	29,2860	0,0990	1,7089	0,7240	1,9524
	1,8466	2,5678	0,0620						
...	...	...	...	...	...	...	...	...	...
	...	...	...						
...	...	...	...	...	...	...	...	...	...
	...	...	...						
...	...	...	...	...	...	...	...	...	...
	...	...	...						
288	0,0500	0,1531	1,4548	27,7460	29,4650	0,0980	1,7127	0,7283	1,9348
	1,8348	2,4939	0						

Exploratory Data Analysis: Stars Plot

To right:
C
MN
SI
P
S
CU
CR
NI
AL
MO
TI
B



To right:
C
MN
SI
P
S
CU
CR
NI
AL
MO
TI
B

Detail of Stars Plot discover similarity of individual steel samples

Korelace (H404)
Označ. korelace jsou významné na hlad. p < .05000
N=128 (Celé případy vynechány u ChD)

Proměnná	C	MN	SI	P	S	CU	CR	NI	AL	MO	TL	B
C	1.00	0.22	0.27	0.31	0.36	0.21	0.03	0.22	-0.14	-0.02	0.05	0.18
MN	0.22	1.00	-0.37	0.85	0.64	0.84	0.78	-0.24	-0.49	-0.71	-0.58	0.64
SI	0.27	-0.37	1.00	-0.23	-0.03	-0.41	-0.71	0.36	0.42	0.34	0.39	0.12
P	0.31	0.85	-0.23	1.00	0.75	0.72	0.73	-0.13	-0.63	-0.68	-0.62	0.67
S	0.36	0.64	-0.03	0.75	1.00	0.55	0.44	-0.02	-0.50	-0.53	-0.35	0.60
CU	0.21	0.84	-0.41	0.72	0.55	1.00	0.73	-0.40	-0.51	-0.62	-0.55	0.54
CR	0.03	0.78	-0.71	0.73	0.44	0.73	1.00	-0.25	-0.56	-0.59	-0.55	0.34
NI	0.22	-0.24	0.36	-0.13	-0.02	-0.40	-0.25	1.00	0.11	0.24	0.22	-0.23
AL	-0.14	-0.49	0.42	-0.63	-0.50	-0.51	-0.56	0.11	1.00	0.57	0.62	-0.37
MO	-0.02	-0.71	0.34	-0.68	-0.53	-0.62	-0.59	0.24	0.57	1.00	0.79	-0.67
TL	0.05	-0.58	0.39	-0.62	-0.35	-0.55	-0.55	0.22	0.62	0.79	1.00	-0.54
B	0.18	0.64	0.12	0.67	0.60	0.54	0.34	-0.23	-0.37	-0.67	-0.54	1.00

Matrix of Pearson´s correlation coefficients expresses internal dependence between two variables.

Red numbers stand for statistically significant correlation.

STATISTICA Cz - [PS 1* - Korelace (H404 13s*128ř)]

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Pracovní sešit Okno Nápověda

PS 1*
- Základní statistiky a tabulky
  - Dialog korelací
    - Korelace (H404)
    - Korelace (H404 13s*

C
MN
SI
P
S
CU
CR
NI
AL
MO
TL
B

Korelace (H404)
Označ. korelace jsou významné na hlad. p < .05000
N=128 (Celé případy vynechány u ChD)

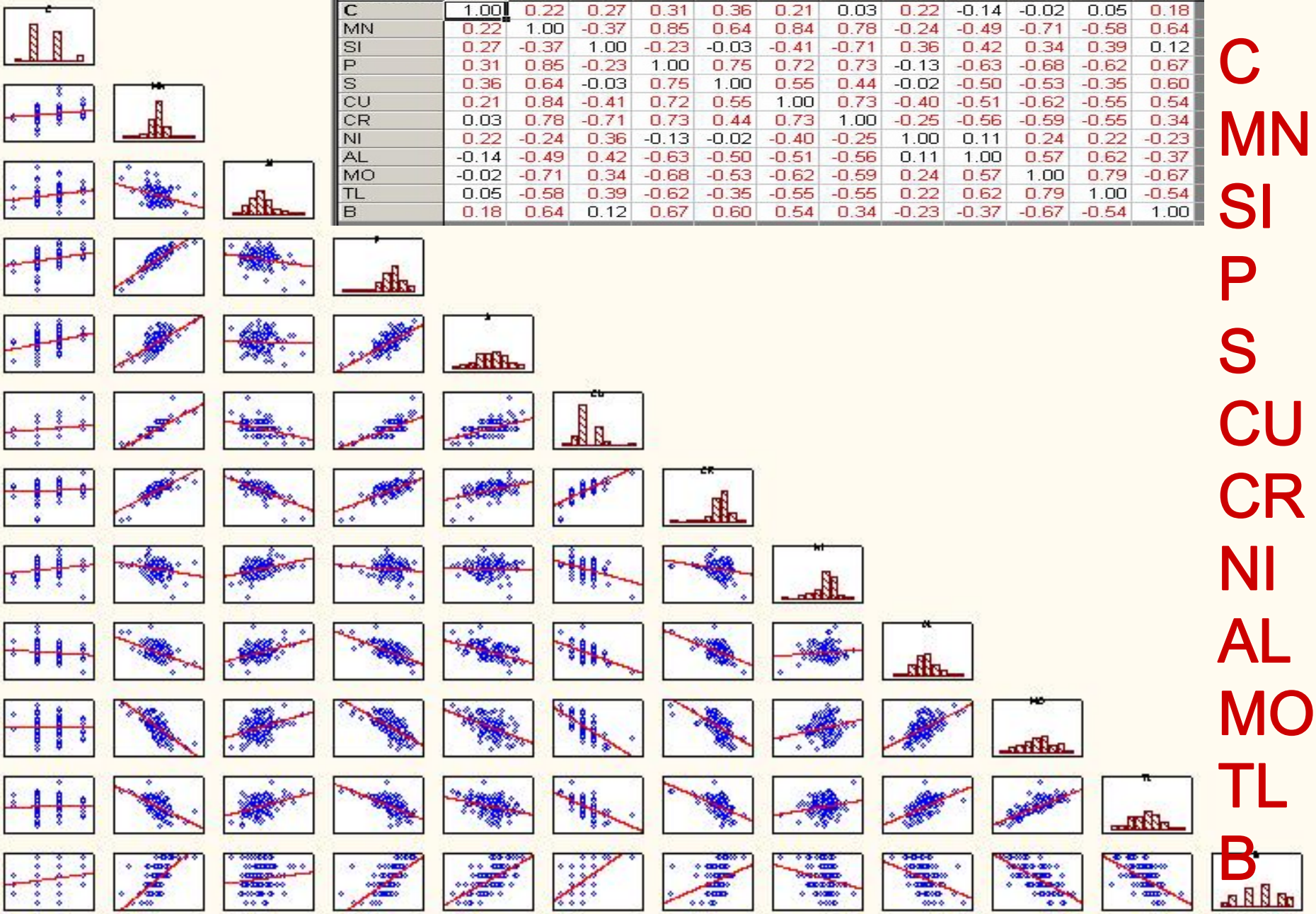
Proměnná	C	MN	SI	P	S	CU	CR	NI	AL	MO	TL	B
C	1.00	0.22	0.27	0.31	0.36	0.21	0.03	0.22	-0.14	-0.02	0.05	0.18
MN	0.22	1.00	-0.37	0.85	0.64	0.84	0.78	-0.24	-0.49	-0.71	-0.58	0.64
SI	0.27	-0.37	1.00	-0.23	-0.03	-0.41	-0.71	0.36	0.42	0.34	0.39	0.12
P	0.31	0.85	-0.23	1.00	0.75	0.72	0.73	-0.13	-0.63	-0.68	-0.62	0.67
S	0.36	0.64	-0.03	0.75	1.00	0.55	0.44	-0.02	-0.50	-0.53	-0.35	0.60
CU	0.21	0.84	-0.41	0.72	0.55	1.00	0.73	-0.40	-0.51	-0.62	-0.55	0.54
CR	0.03	0.78	-0.71	0.73	0.44	0.73	1.00	-0.25	-0.56	-0.59	-0.55	0.34
NI	0.22	-0.24	0.36	-0.13	-0.02	-0.40	-0.25	1.00	0.11	0.24	0.22	-0.23
AL	-0.14	-0.49	0.42	-0.63	-0.50	-0.51	-0.56	0.11	1.00	0.57	0.62	-0.37
MO	-0.02	-0.71	0.34	-0.68	-0.53	-0.62	-0.59	0.24	0.57	1.00	0.79	-0.67
TL	0.05	-0.58	0.39	-0.62	-0.35	-0.55	-0.55	0.22	0.62	0.79	1.00	-0.54
B	0.18	0.64	0.12	0.67	0.60	0.54	0.34	-0.23	-0.37	-0.67	-0.54	1.00

Diagram of correlation matrix shows internal bounds among individual variables.

Start | Microsoft PowerPoint - [... | Total Commander 6.0 - P... | Zoner Photo Studio 9 - E... | STATISTICA Cz - [PS ... | 15:43

Korelace (H404)
Označ. korelace jsou významné na hlad. p < .05000
N=128 (Celé případy vynechány u ChD)

Proměnná	C	MN	SI	P	S	CU	CR	NI	AL	MO	TL	B
C	1.00	0.22	0.27	0.31	0.36	0.21	0.03	0.22	-0.14	-0.02	0.05	0.18
MN	0.22	1.00	-0.37	0.85	0.64	0.84	0.78	-0.24	-0.49	-0.71	-0.58	0.64
SI	0.27	-0.37	1.00	-0.23	-0.03	-0.41	-0.71	0.36	0.42	0.34	0.39	0.12
P	0.31	0.85	-0.23	1.00	0.75	0.72	0.73	-0.13	-0.63	-0.68	-0.62	0.67
S	0.36	0.64	-0.03	0.75	1.00	0.55	0.44	-0.02	-0.50	-0.53	-0.35	0.60
CU	0.21	0.84	-0.41	0.72	0.55	1.00	0.73	-0.40	-0.51	-0.62	-0.55	0.54
CR	0.03	0.78	-0.71	0.73	0.44	0.73	1.00	-0.25	-0.56	-0.59	-0.55	0.34
NI	0.22	-0.24	0.36	-0.13	-0.02	-0.40	-0.25	1.00	0.11	0.24	0.22	-0.23
AL	-0.14	-0.49	0.42	-0.63	-0.50	-0.51	-0.56	0.11	1.00	0.57	0.62	-0.37
MO	-0.02	-0.71	0.34	-0.68	-0.53	-0.62	-0.59	0.24	0.57	1.00	0.79	-0.67
TL	0.05	-0.58	0.39	-0.62	-0.35	-0.55	-0.55	0.22	0.62	0.79	1.00	-0.54
B	0.18	0.64	0.12	0.67	0.60	0.54	0.34	-0.23	-0.37	-0.67	-0.54	1.00



Výklad:

Indexový graf úpatí vlastních čísel

(Scree Plot)

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

Zobrazuje: relativní velikost jednotlivých vlastních čísel.

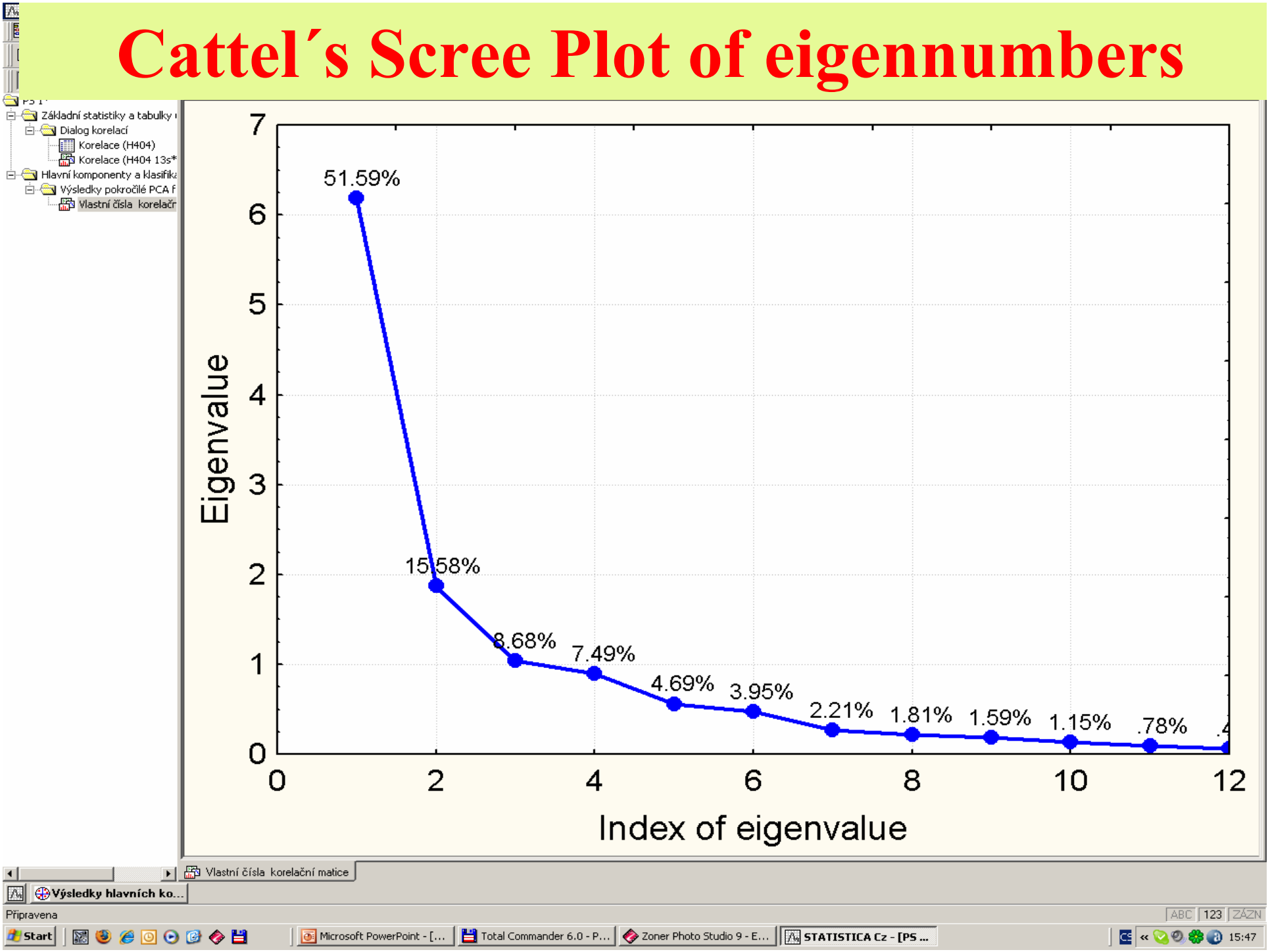
Využití: k určení počtu *A* "užitečných" hlavních komponent. Graf úpatí se jeví nejobjektivnějším kritériem.

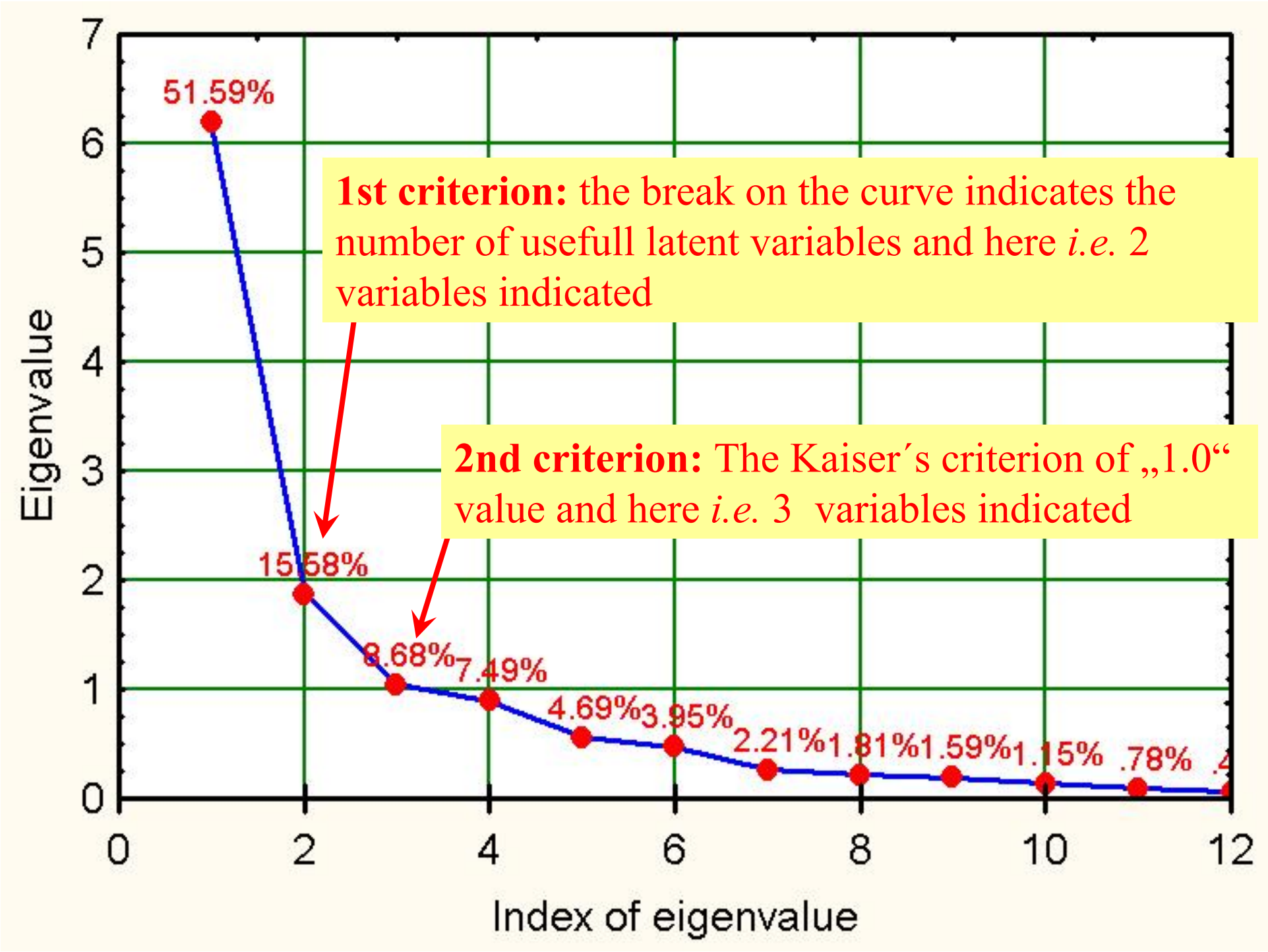
Kritérium "1": hrubším kritériem PC, jejichž vlastní číslo je větší než jedna. Graf úpatí se však jeví objektivnějším.

Cattel´s Scree Plot of eigennumbers

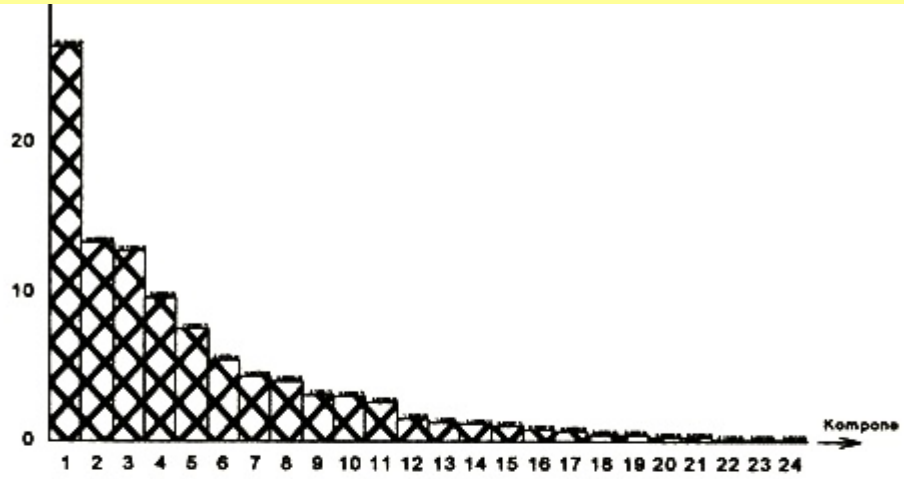
is the columns diagram of eigennumbers plotted against the index and determines a number of usefull latent variables.

Cattel´s Scree Plot of eigennumbers

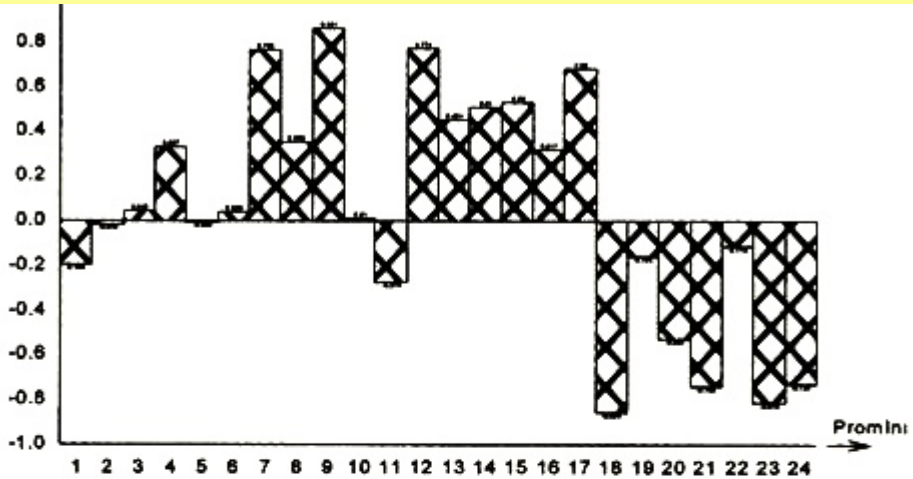



1st criterion: the break on the curve indicates the number of usefull latent variables and here i.e. 2 variables indicated
2nd criterion: The Kaiser´s criterion of „1.0“ value and here i.e. 3 variables indicated
51.59%
15.58%
8.68%
7.49%
4.69%
3.95%
2.21%
1.81%
1.59%
1.15%
.78%
Eigenvalue
Index of eigenvalue
0
1
2
3
4
5
6
7
0
2
4
6
8
10
12

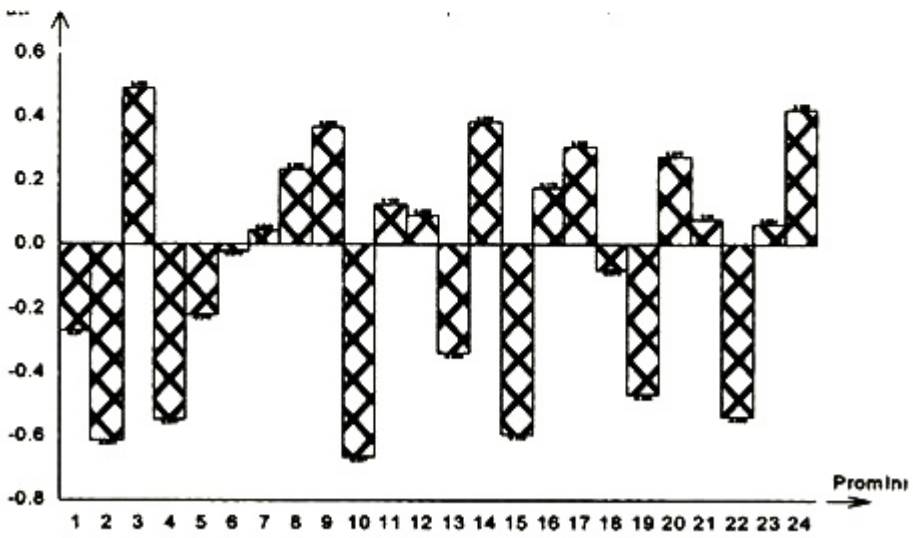
Latent variables composed from the original variables



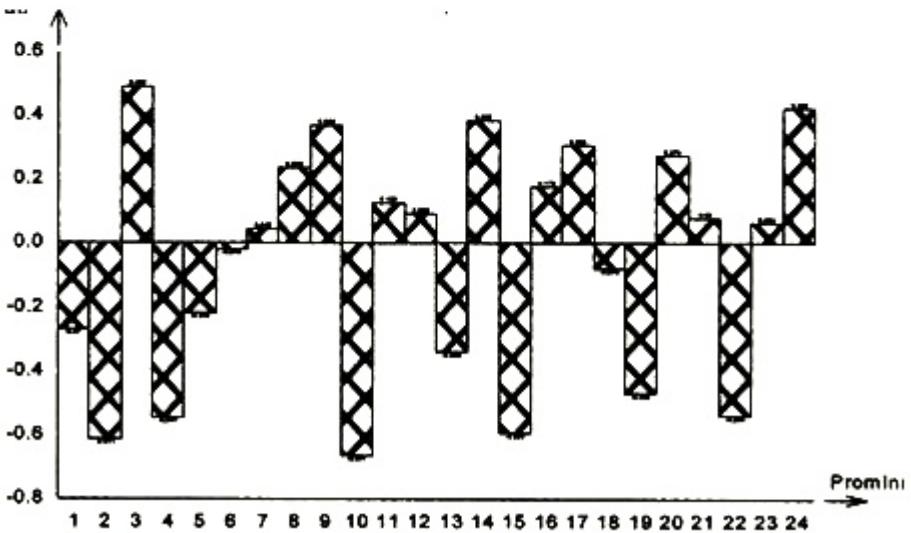
The Cattel´s Scree plot



The first principal component



The second principal component



The third principal component

Výklad:

Graf komponentních vah, zátěží

(Plot Components Weights)

Zobrazí: komponentní váhy

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

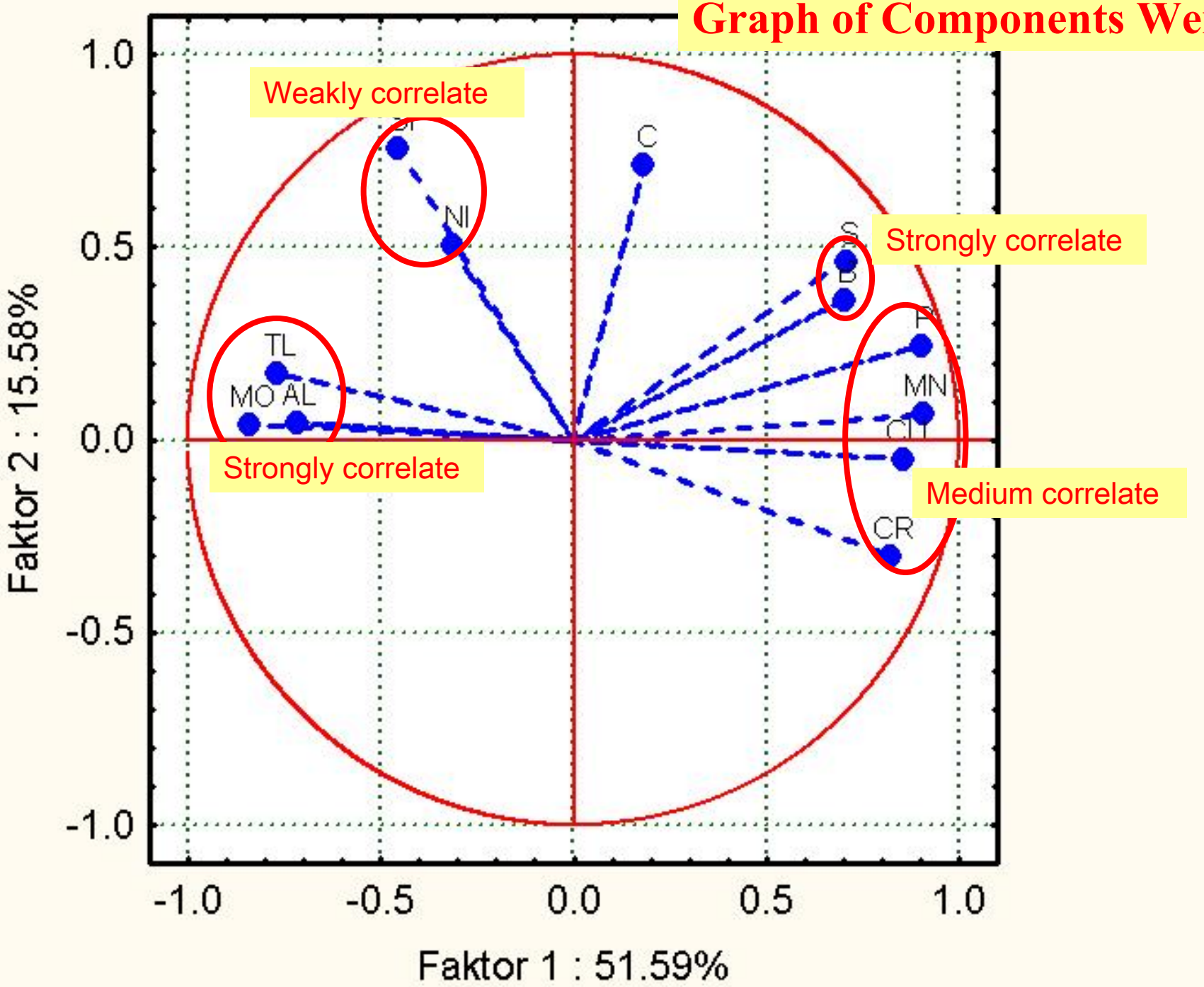
Nalezneme: shluk podobných proměnných, jež spolu korelují.

Představuje: most mezi původními proměnnými a hlavními komponentami.

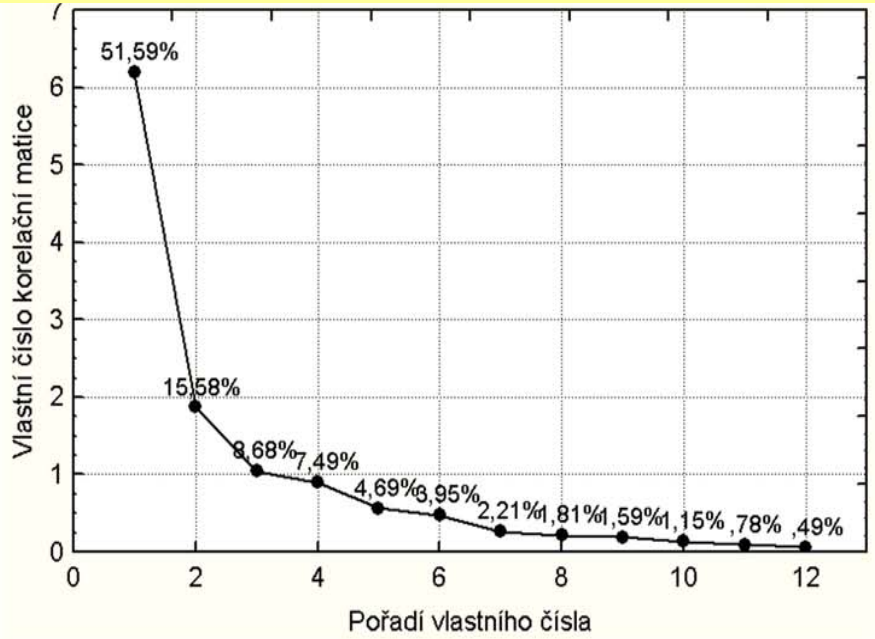
Graph of Components Weights

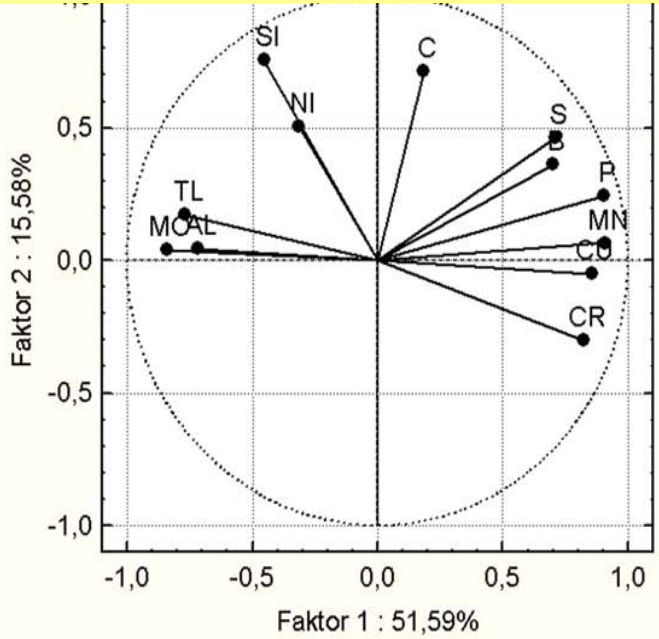
shows clusters of original variables which correlate

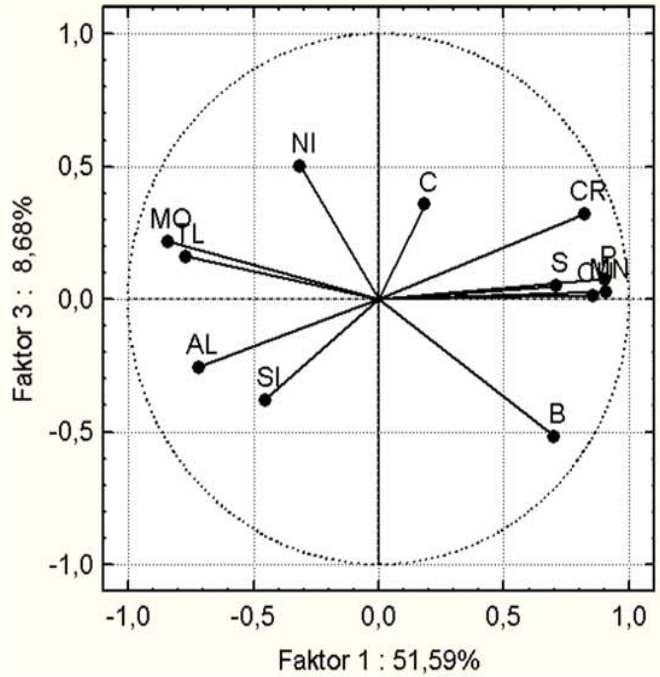
Graph of Components Weights

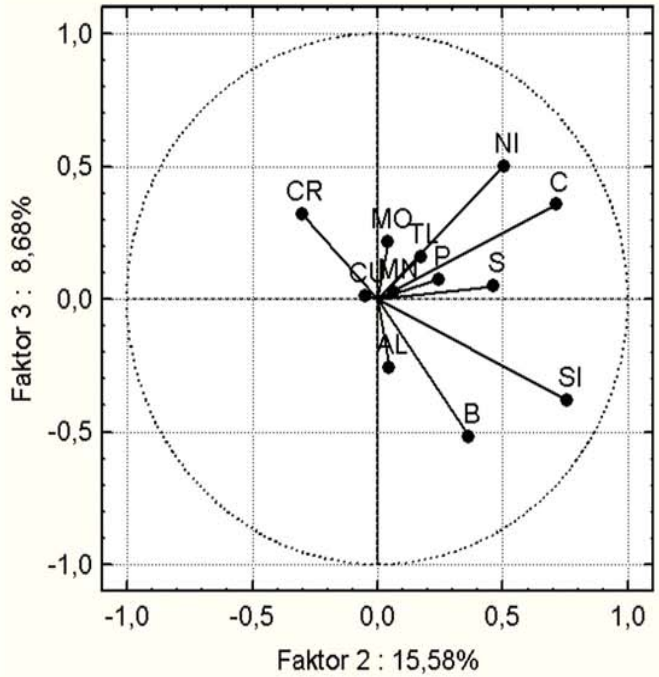


Correlation and internal bounds of elements in Cr-Ni steel and tested homogeneity in PCA graphs









Proměnné ve skupinách (**Al, Mo, Ti**) a (**Mn, P, S, Cu, B**) je možno nahradit jedinou proměnnou ze středu skupin, která bude reprezentovat chování celé skupiny.

Největší přínos pro danou komponentu mají proměnné, které se v grafu nacházejí u souřadnice dané komponenty a co nejdále od nuly.

Největší přínos k 1. hlavní komponentě mají **Mn, Cu, P** a **Cr**. Ale nejmenší přínos mají **Mo, Ti** a **Al**.

Největší přínos k 2. hlavní komponentě mají **Si, Ni, C.** Ale nejmenší přínos mají **Cu, Cr** a **Al**.

První dvě hlavní komponenty popisují 68,5 % variability dat, první tři 77,5 % a čtyři 84,9 %.

Po zredukování počtu proměnných vyjde procento vysvětlené variability pro první dvě hlavní komponenty 80,5 % a první tři 93 %.

Je možné odstranit Mn, protože vykazuje podobné vlastnosti jako Cr.

K **popisu variability** zdrojové matice postačují tři hlavní komponenty.

Grafy naznačují nekorelovanost zbývajících proměnných.
Příspěvky jednotlivých proměnných do hlavních komponent jsou:

$$y_1 = -0{,}510*Mo - 0{,}578*Si - 0{,}960*C + 0{,}630*Cr$$

$$y_2 = 0{,}242*Mo - 0{,}272*Si - 0{,}911*C - 0{,}192*Cr$$

PCA zredukovala počet proměnných na 4 základní, postačující k testování homogenity materiálu:

1) **C - nebezpečí tvorby karbidů s Cr, Mn, Si a Mo**, což vede k místnímu poklesu obsahu uvedených prvků, zvláště chromu a vzniku mezikrystalové koroze. Platí pravidlo: **Čím nižší obsah C, tím vyšší kvalita materiálu zvláště pro antikorozní účely**.

2) **Si - feritotvorný prvek**, zvyšuje stabilitu feritu, zvyšuje odolnost proti oxidaci ve vysoce oxidačních prostředích vznikem vrstvy SiO2 na povrchu materiálu.

3) **Cr - feritotvorný prvek**, hlavní účel-odolnost proti korozi, oxidaci a opotřebení.

4) **Mo - feritotvorný prvek**, odolnost proti korozi, oxidaci a zvýšení žáropevnosti.

Komunalita představuje váhu s jakou jednotlivé proměnné přispívají ke konstrukci latentních proměnných (nejvíc přispívají **Cr** a **Si** a nejméně **C** a **Mo**).

K vysvětlení 85 % variability dat je zapotřebí minimálně 3 faktorů.
Ty společně vysvětlují 93 % variability dat.

1. faktor je silně korelován především s **Cr, Si** a **Mo**,
2. faktor proměnnou **C**,
3. faktor hlavně proměnnou **Mo**.

Na celkové společné vysvětlené variabilitě se nejvíce podílí **C** dále **Mo**.

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.

Scatterplot of principal components score

1. **Location of objects:** leverages (extremes) are far from the origin
2. **Similarity of objects:** close objects are similar, far objects are unsimilar
3. **Objects located in clusters:** objects located in cluster are very similar

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.

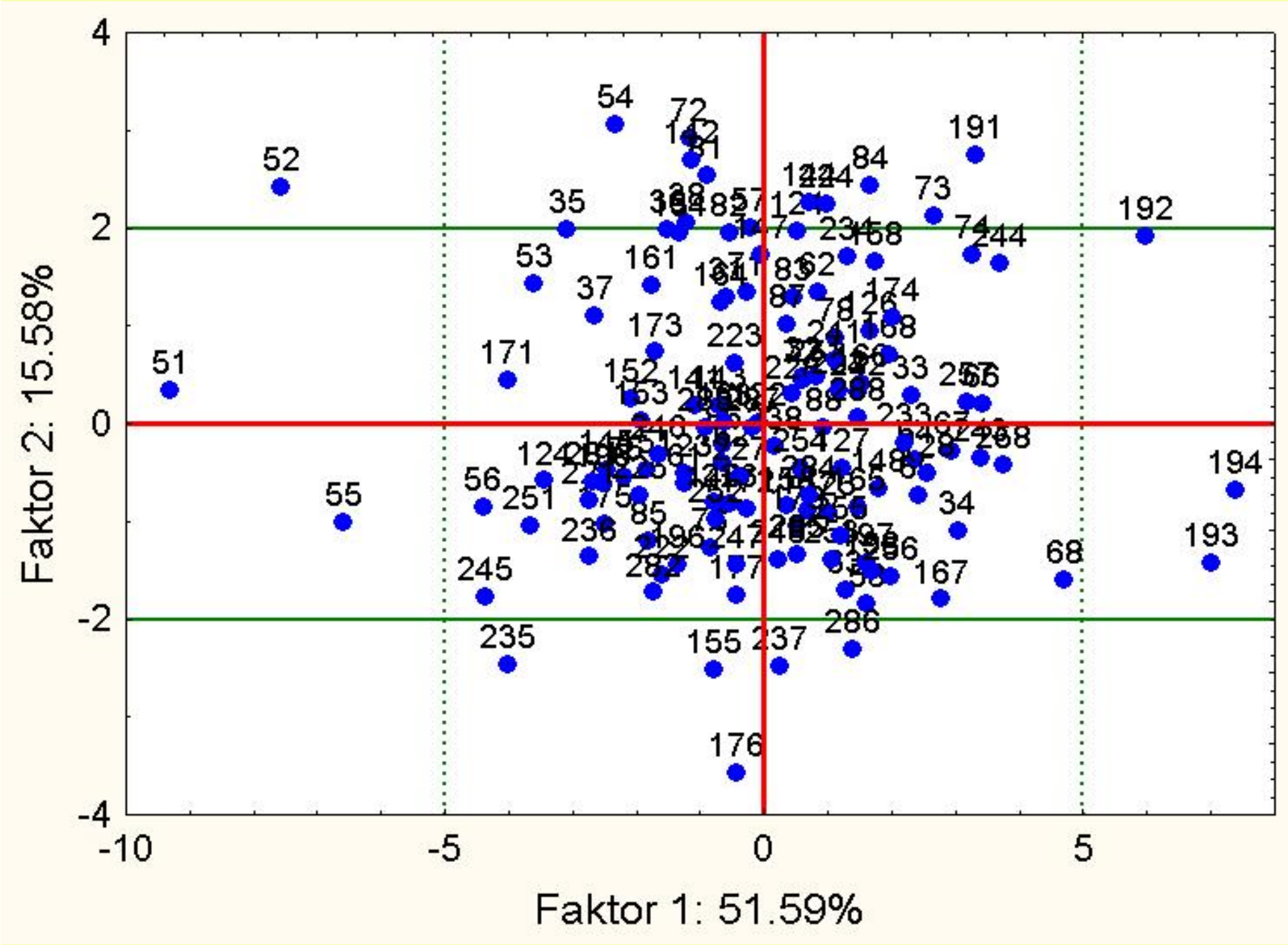
7. **Vysvětlení místa objektu:** umístění objektu na ploše v diagramu může být porovnáváno s komponentními vahami původních proměnných ve dvojném grafu.

4. Secluded objects: isolated or secluded objects can be outlying.

5. Faraway and distant objects: can be outlying.

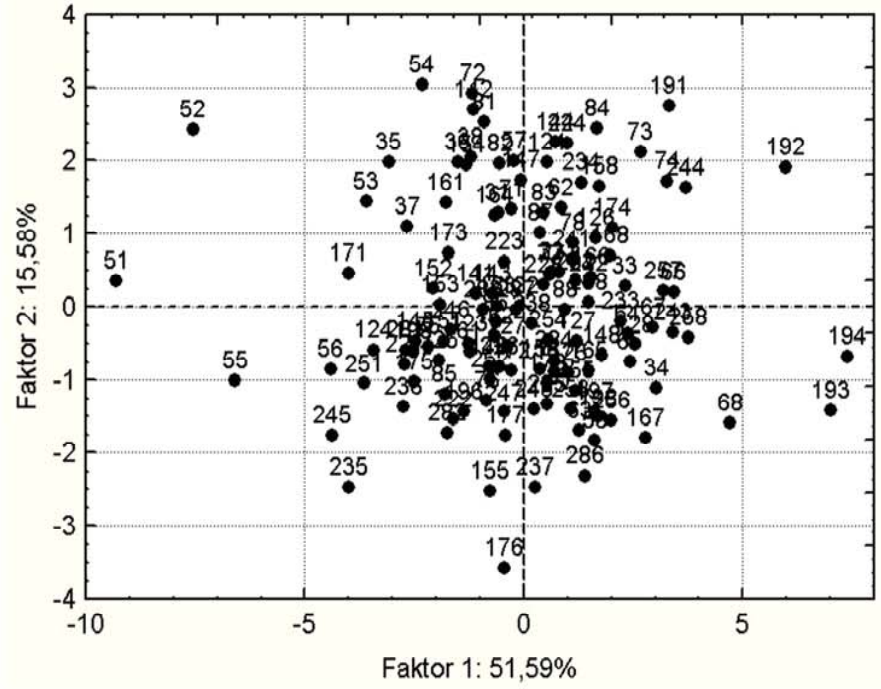
6. Name of objects: can help to give names to principal components.

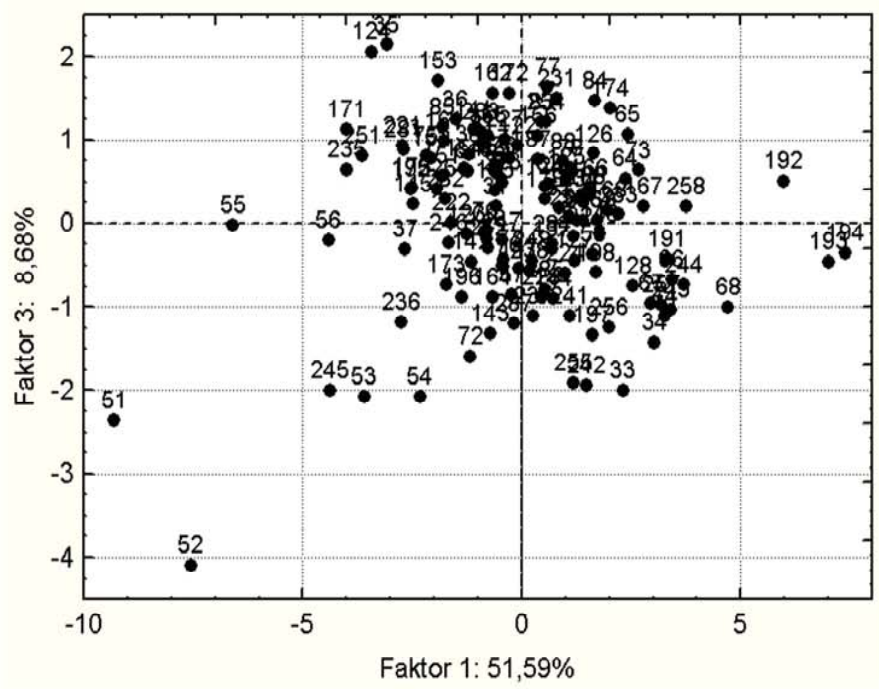
7. Interpretation of object location: location of objects in space can be compared in biplot with the location of variables and interpreted together.

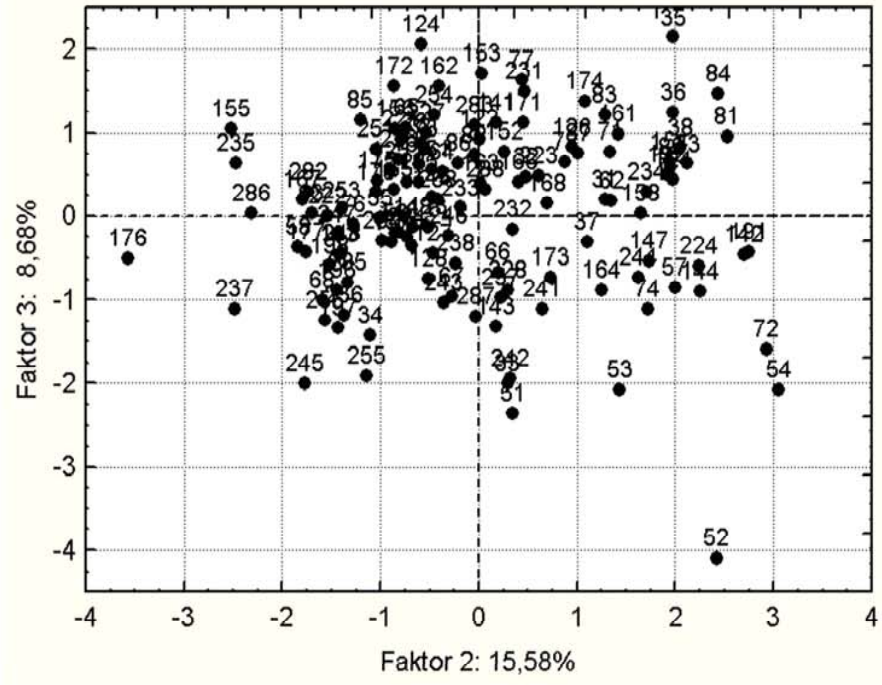

Faktor 2: 15.58%
Faktor 1: 51.59%
4
2
0
-2
-4
-10
-5
0
5
52
51
55
54
72
191
192
194
193
68
167
73
84
35
53
37
161
173
171
223
174
244
245
235
155
237
286
176
56
251
236
34
33

V bodovém diagramu komponentního skóre jsou proti sobě vyneseny první dvě hlavní komponenty, což díky jejich ortogonalitě usnadňuje geometrickou interpretaci jednotlivých objektů po stránce jejich strukturních vazeb.

Objekty 51, 52, 55, 191, 192, 193 a 194 jsou strukturně vzdáleny od ostatních objektů i od sebe samých.


Faktor 2: 15,58%
Faktor 1: 51,59%


Faktor 3: 8,68%
Faktor 1: 51,59%


Faktor 3: 8,68%
Faktor 2: 15,58%

Závěr: Faktorově nejčistší vychází proměnná **C** (druhý faktor, po rotaci třetí), což znamená, že je nejméně závislá a podobná ostatním proměnným.

Stejně vychází **Cr** (v 1. faktoru). **Si** a **Mo** vycházejí po rotaci faktorově čisté (v 1. faktoru a ve 2. faktoru).

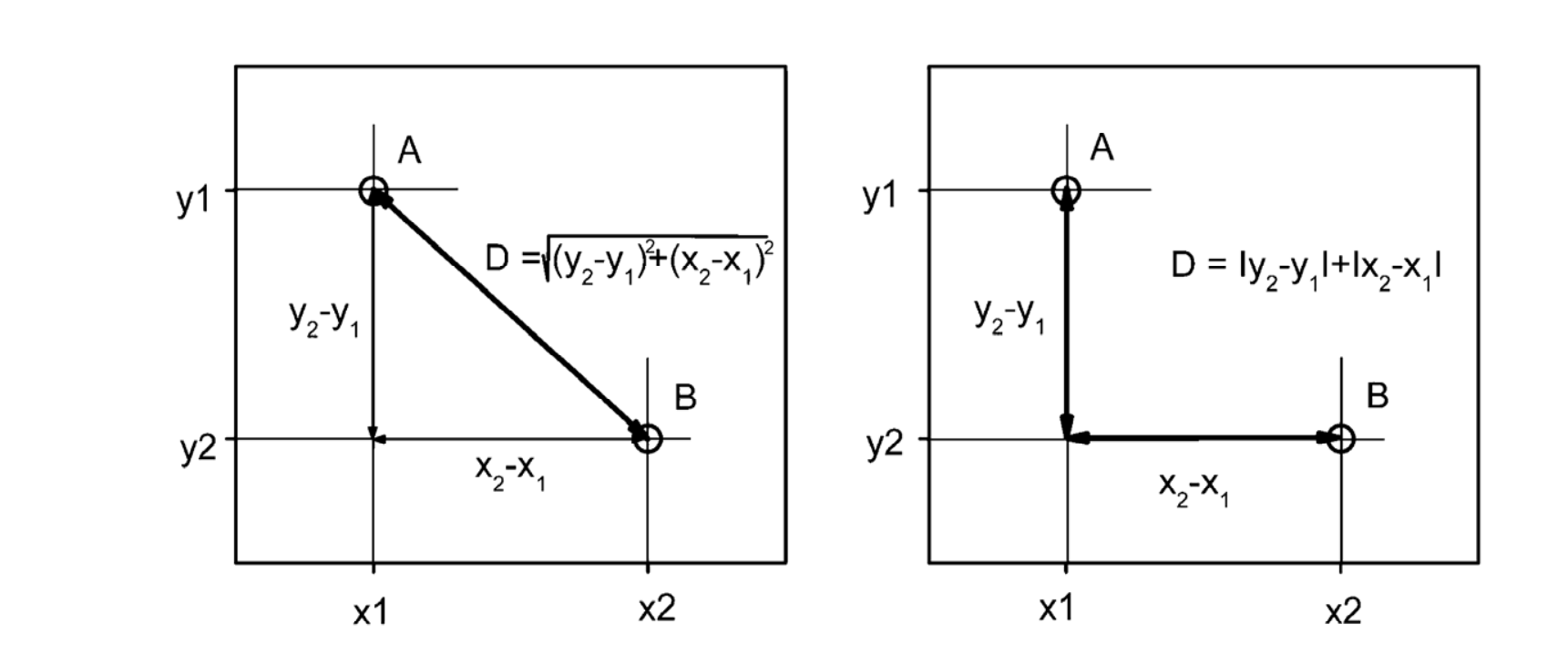
Tím, že proměnná **C** jeví zcela odlišné vlastnosti od ostatních proměnných je zapotřebí dalšího samostatného faktoru k vysvětlení jejího chování.

K vysvětlení proměnných bez **C** by stačily pouze dva faktory.

Čím odlišnější vlastnosti mezi jednotlivými proměnnými, tím více latentních proměnných je zapotřebí k popisu chování proměnných.

Clustering means searching for similar objects or similar variables.

When objects are known to exist in several categories, clustering techniques can be used as classification techniques.



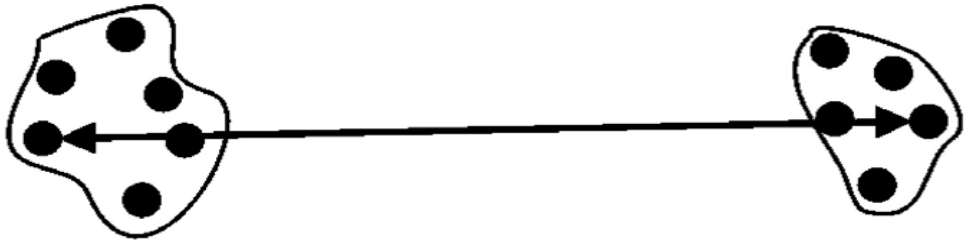
A fundamental concept of clustering techniques is ***similarity*.**

Similarity of objects and variables is considered on base on *Mahalonobis distance* in the *m*-dimensional space or *Euclidean distance* in the *m*-dimensional space.

Metoda nejblizšího souseda

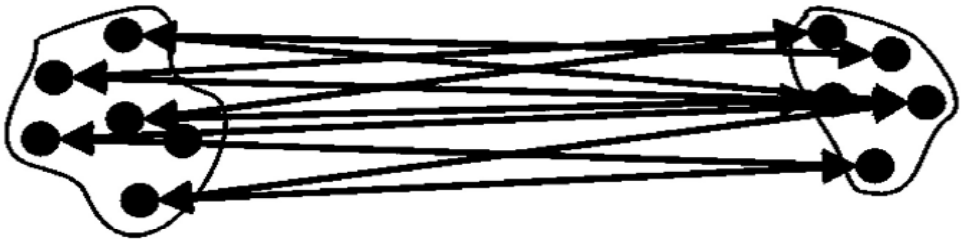
Method of the nearest neighbour

Metoda nevzdálenějšího souseda



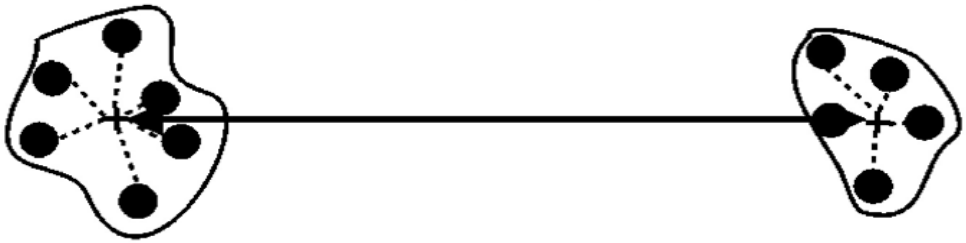
Method of the farthest neighbour

Metoda průměrné vzdálenosti

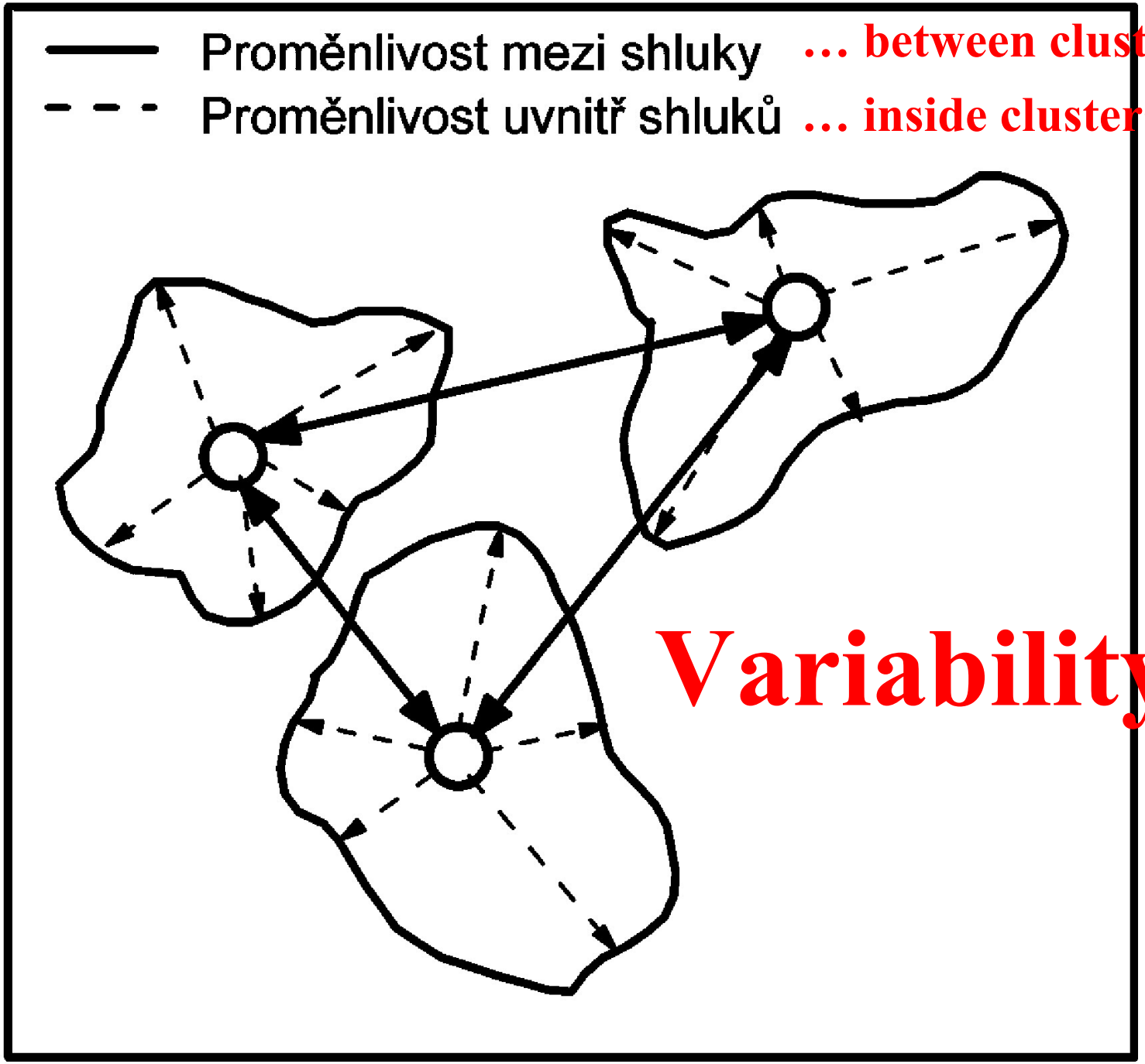


Method of the average distance

Metoda Wardova



Ward method



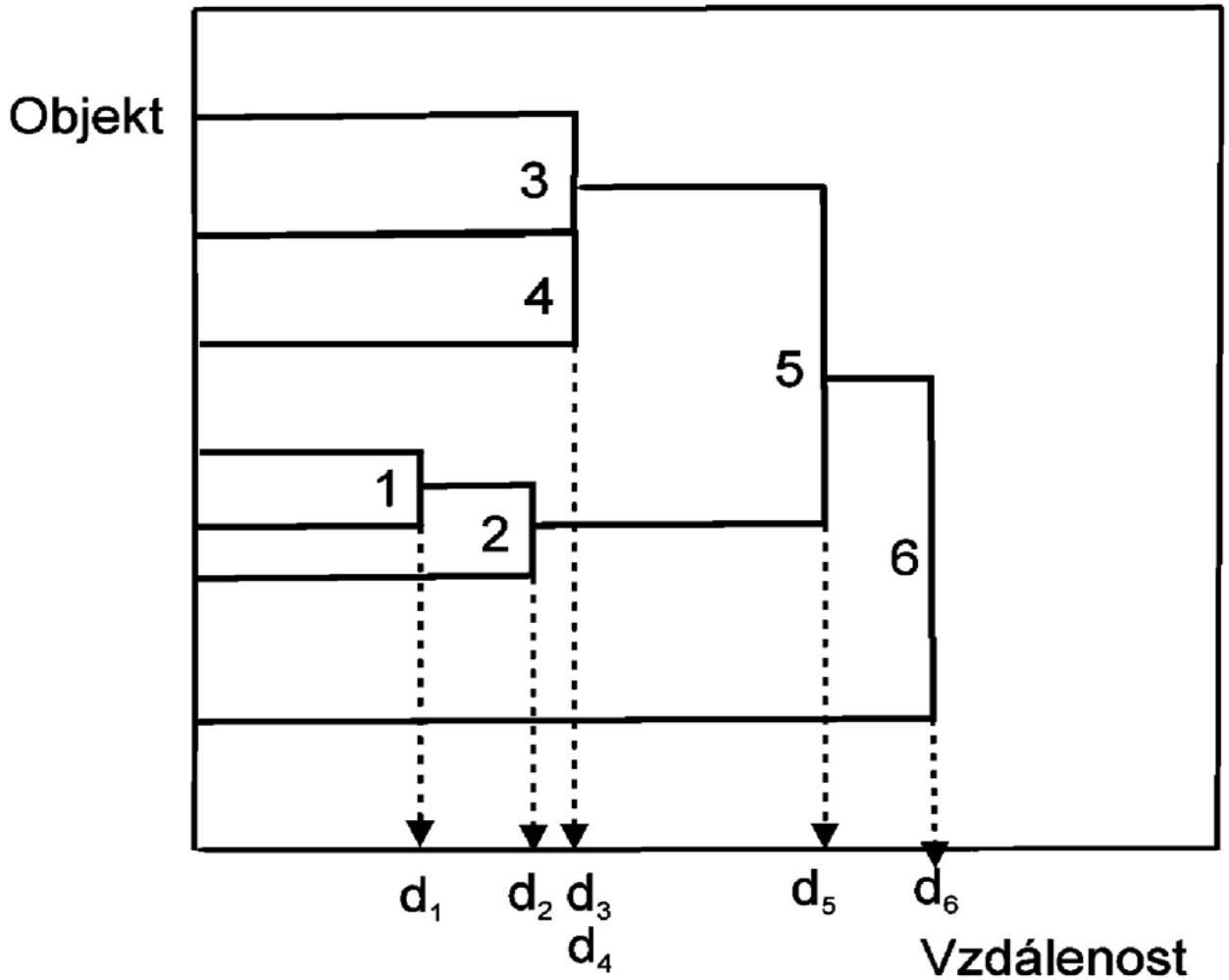
… between clusters

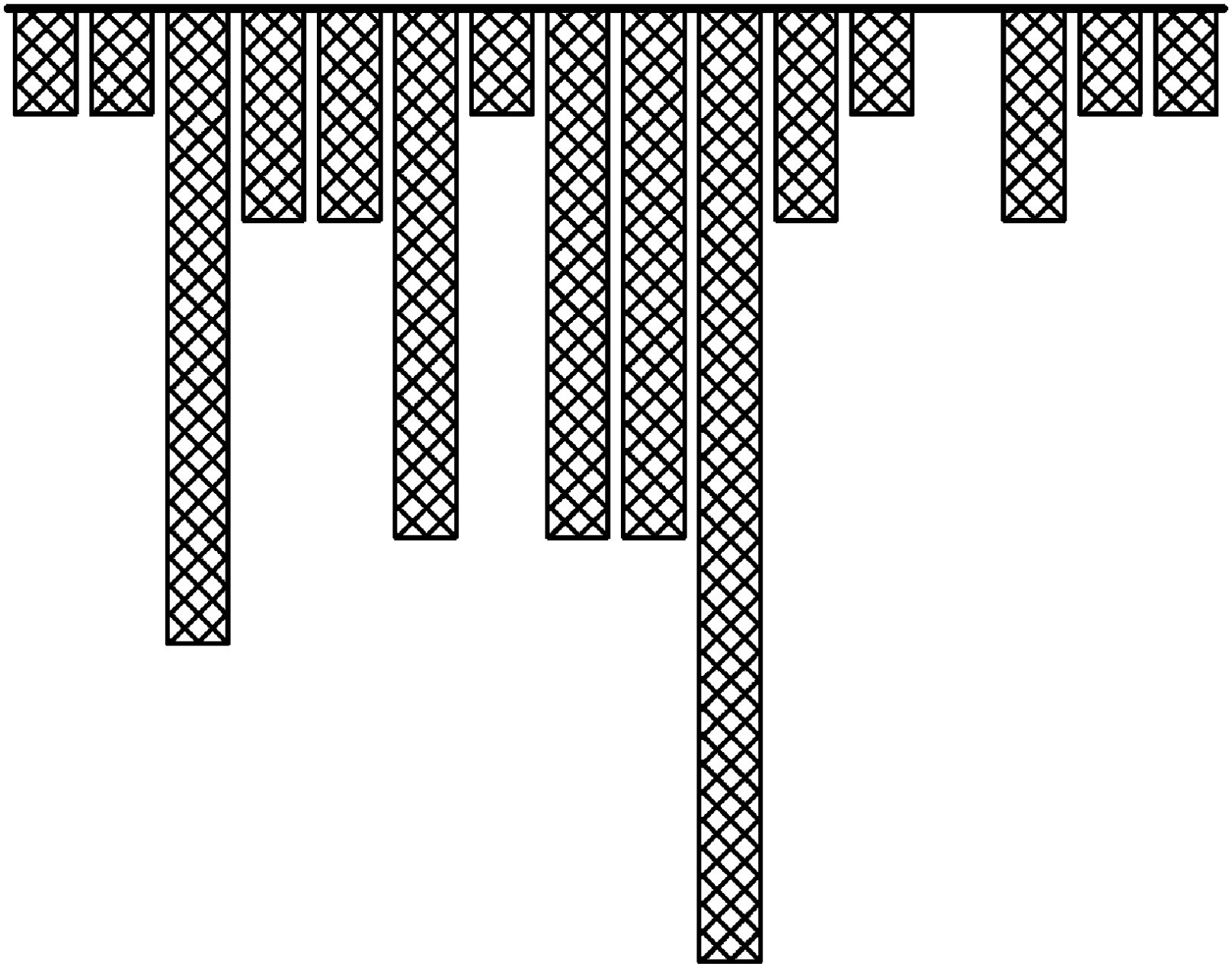
… inside clusters

Variability …

The **cluster analysis** leads to clusters which are plotted in **dendrogram**.

Dendrogram of variables, dendrogram of objects.





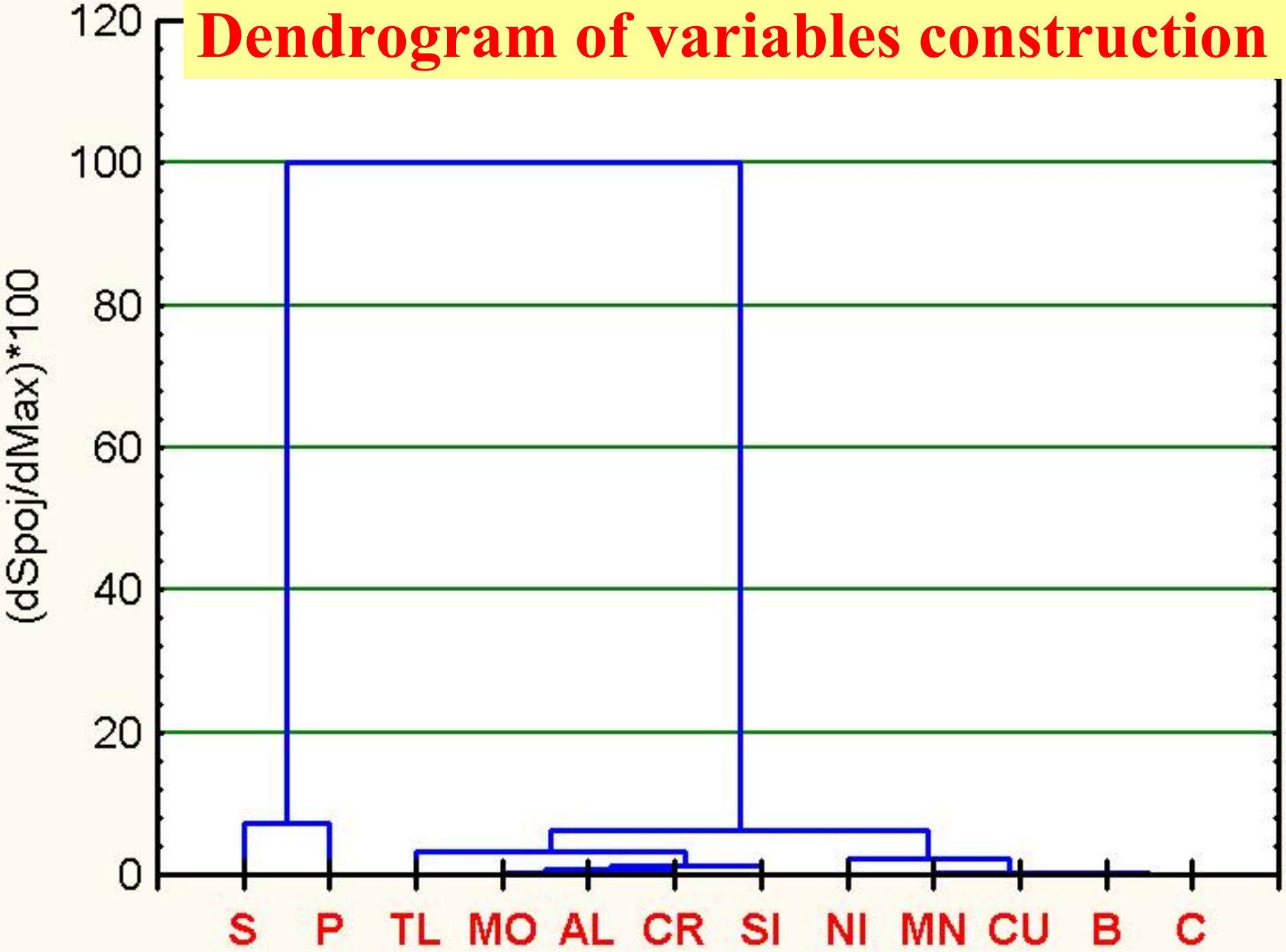
Rampouch type dendrogram

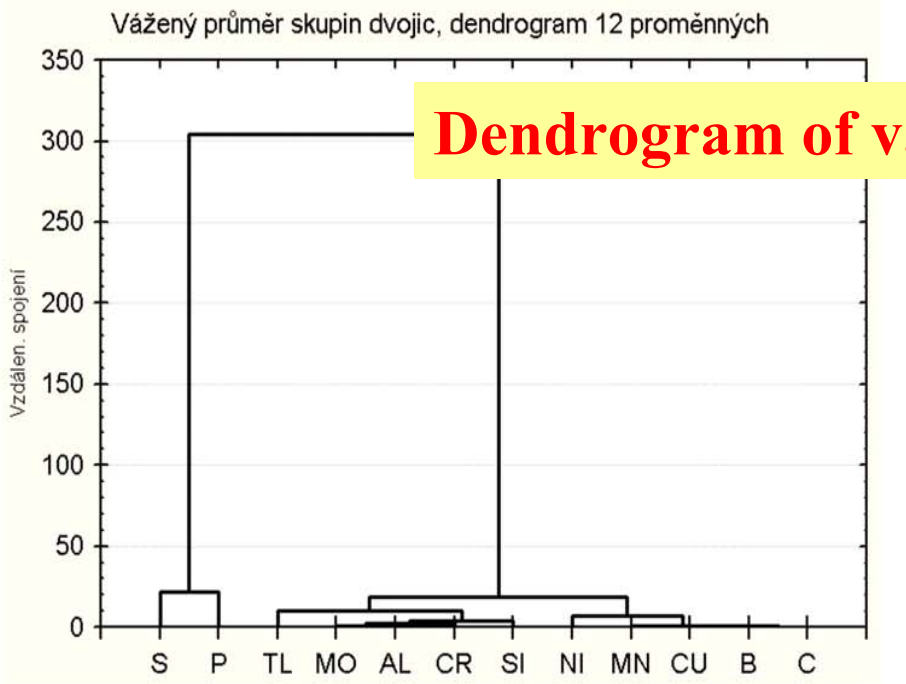
The ***agglomerative hierarchical methods*** start with many clusters as objects. Clusters are progressively linked to form bigger clusters, until a single big cluster, of all the objects, is obtained.

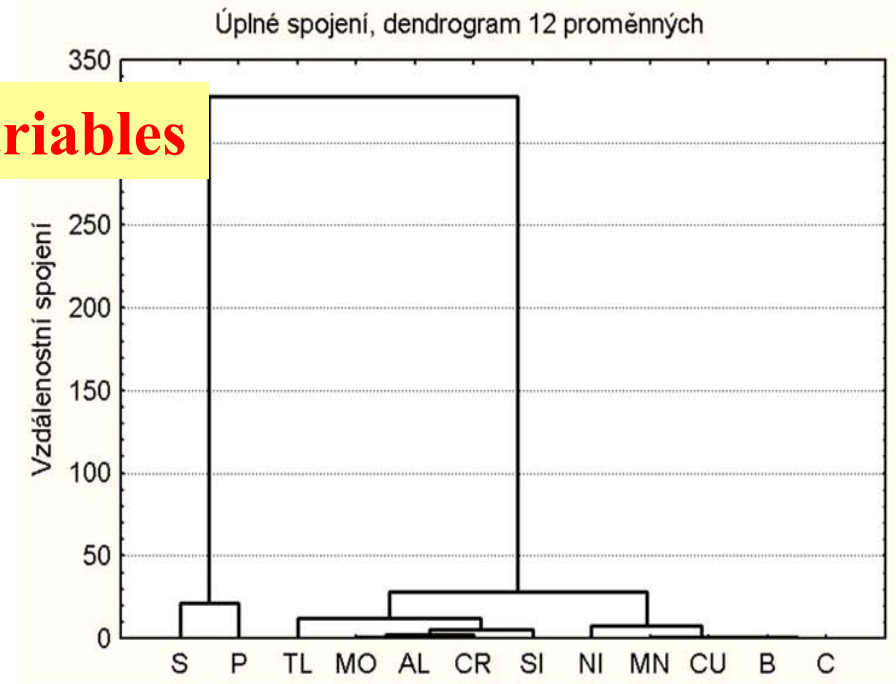
Divisive methods start instead with a single big cluster, and this is progressively subdivided into smaller cluster, until as many clusters as objects have been obtained.

Both statistical techniques are demonstrated on the analysis and classification of the homogeneity of steel material.

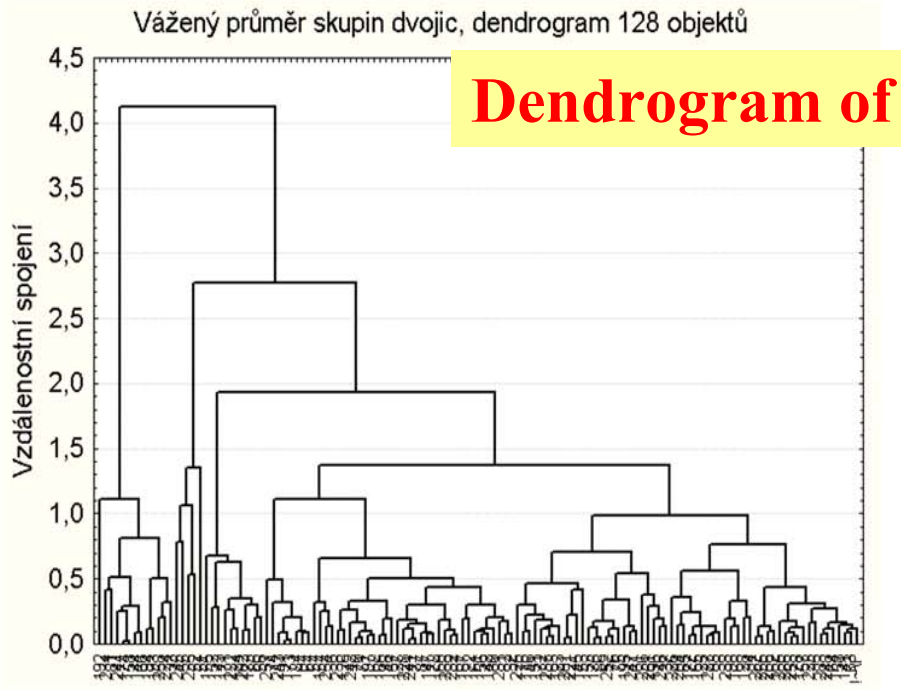
Dendrogram of variables construction

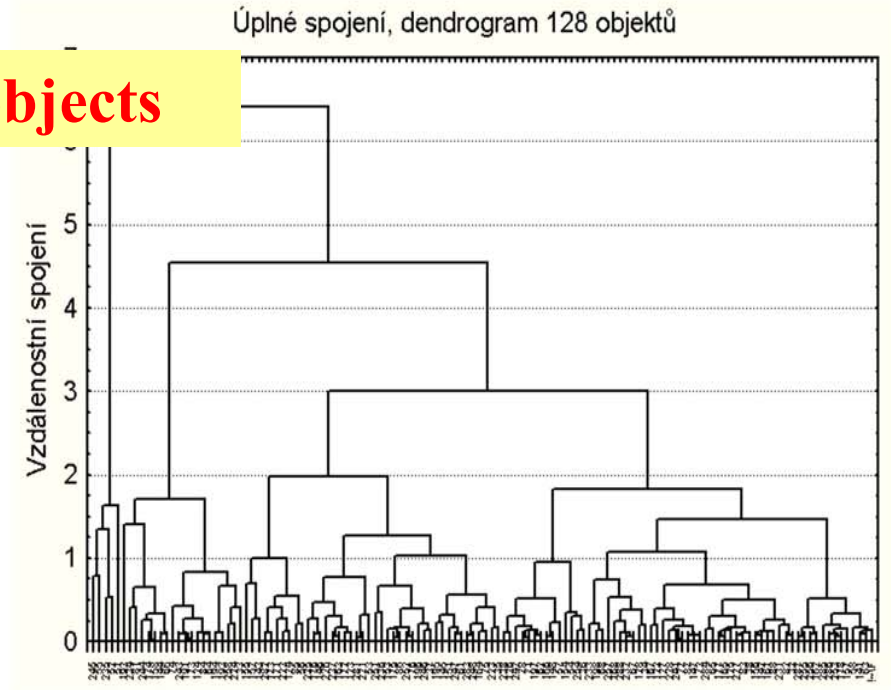






Dendrogram of variables





Dendrogram of objects

Existuje určitá rozdílnost v chemickém složení mezi jednotlivými místy:

Strukturně odlehlé body se jeví 51, 52, 55 a 192, což ukázaly i symbolové a PCA grafy.

Do výpočtu byly zahrnuty všechny proměnné. Determinant korelační matice je velice blízký nule (-0,0046), proto je možno předpokládat multikolinearitu (závislost mezi proměnnými).

Párové korelační koeficienty a příslušné kovariance naznačují významnou míru vztahu mezi některými proměnnými.

Korelační analýza potvrdila významné korelace mezi proměnnými z rozptylových grafů.

Prvky ve skupinách (**Al, Mo** a **Ti**) a (**Mn, P, S, Cu, Cr** a **B** vykazují v rámci skupiny pozitivní korelace a naopak meziskupinově negativní korelace. **C, Ni** a **Si** (pouze **Cr**) se jeví jako nezávislé.

PCA vede ke snížení počtu proměnných na 4 prvky **C, Si, Cr** a **Mo**, které dostatečně reprezentují chování všech proměnných.

C se jeví jako faktorově nejčistší tzn., že je nejméně závislý na ostatních proměnných, i když s určitou vazbou na **Si**, který tvoří jistý předěl mezi **C** a **Mo**.

Cr a **Mo** vykazují navzájem opačné chování.

Shlukovou analýzou byly identifikovány odlehlé body 51, 52, 55 a 192.

Byla indikována **nehomogenita** mezi analyzovanými místy ve vzorcích.

Mezi vzorky samotnými nebyl nalezen žádný koncentrační trend.

Doporučená literatura:

[1] Siotani M., Hayakawa T., Fujikoshi Y.: **Modern Multivariate Statistical Analysis**, A Graduate Course and Handbook. American Science Press, Columbia 1985.

[2] Johnson R.A., Wichern D.W.: **Applied Multivariate Statistical Analysis**, Prentice Hall, 1998.

[3] Meloun M., Militký J. , Forina M.: **Chemometrics for Analytical Chemistry,** Volume 1. **PC-Aided Statistical Data Analysis**, Ellis Horwood, Chichester 1992.

[4] Brereton R. G. **Multivariate Pattern Recognition in Chemometrics**, Illustrated by Case Studies, Elsevier 1992.

[5] Krzanowski W. J.: **Principles of Multivariate Analysis**, A User's Perspective, Oxford Science Publications, 1988.

[6] Meloun M. , Militký J.: **Statistické zpracování experimentálních dat**, Plus Praha 1994, Academia Praha 2004 (v tisku).

[7] Everitt B. S., Dunn G.: **Applied Multivariate Data Analysis**, Arnold, London 2001.

[8] Meloun M. , Militký J.: **Kompendium statistického zpracování dat**, Academia Praha 2002.

Je třeba studovat a na sobě pracovat!

Intenzivní týdenní kurzy (35 hodin) a

2-leté Licenční studium (280 hodin)

na Univerzitě v Pardubicích

Sledujte **http://meloun.upce.cz**

Milan Meloun - Mozilla Firefox

Soubor Úpravy Zobrazit Historie Záložky Nástroje Nápověda

Zpět Vpřed Obnovit Zastavit Domů http://meloun.upce.cz/ Google

Centrála Firefoxu Přehled zpráv

Home Personal Photo Gallery Research Lectures Papers Algorithms Data Sets WWW Visits

MILAN MELOUN

Prof. RNDr. Milan Meloun, DrSc.

professor of analytical chemistry and chemometrics

Address: Department of Analytical Chemistry
University of Pardubice
Čs. legií 565
532 10 Pardubice
Czech republic

Telephone: + 420-46 603 7026
Fax: + 420-46 603 7068

E-mail: milan.meloun@upce.cz

Naše doporučené knihy:

Základní učebnice, 2004

Nová učebnice vícerozměrných statistik, 2005

Kompendium úloh 2002 a **Nové vydání 2006**

Formy dalšího vzdělávání:

Intenzivní kurz v červnu 2007

Nové 12. licenční studium od září 2007

11. licenční studium od ledna 2006

Sledujte http://meloun.upce.cz

Hotovo

Start Doručená pošta... Total Command... meloun@melou... Milan Meloun ... LSWWW99 - Wi... Microsoft Power... 9:11

Milan Meloun - Mozilla Firefox

Soubor Úpravy Zobrazit Historie Záložky Nástroje Nápověda

Zpět Vpřed Obnovit Zastavit Domů http://meloun.upce.cz/ Google

Centrála Firefoxu Přehled zpráv

Home Personal Photo Gallery Research Lectures Papers Algorithms Data Sets WWW Visits

Univerzita Pardubice, Fakulta chemicko-technologická

nabízí týdenní intenzivní kurz na téma

Statistické zpracování dat

1. Týdenní kurz počítačové analýzy dat

Po hodině teorie následuje vždy hodina praktických úloh na počítači.

Kurz je s novou učebnicí Kompendium .....+ CD.!

Termín nejbližšího kurzu 3330/PK370002/81: **4. - 8. června 2007**

Zaplatit na číslo účtu: 37030561/0100, variabilní symbol: 3700023330

.

Termíny kurzů
Místo kurzu
Cena kurzu
Organizátor kurzu
Popis kurzu
Organizační informace
Sylabus kurzu
Obsah úloh
Objednávka kurzu

Termín: Každoročně vždy v měsících únoru, červnu a září (dotaz se na termín nejbližšího kurzu

Sledujte http://meloun.upce.cz

Milan Meloun - Mozilla Firefox

Soubor Úpravy Zobrazit Historie Záložky Nástroje Nápověda

Zpět Uložit (Ctrl+S) Obnovit Zastavit Domů http://meloun.upce.cz/ Google

Centrála Firefoxu Přehled zpráv

Home Personal Photo Gallery Research Lectures Papers Algorithms Data Sets WWW Visits

UNIVERZITA PARDUBICE,

Fakulta chemicko-technologická, Katedra analytické chemie

Projekt dvouletého licenčního studia chemometrie na téma

Statistické zpracování dat

Vedoucí licenčního studia a odborný garant: **Prof. RNDr. Milan Meloun, DrSc.**
Katedra analytické chemie, Fakulta chemicko-technologická,
Univerzita Pardubice, nám. Čs. Legií 565, 53210 Pardubice,
IČO: 216 275, **DIČ:** 248-00216275
Telefon: 466037026, **Fax:** 466037068, **E-mail:** milan.meloun@upce.cz, **http:**//meloun.upce.cz

1. soustředění nového 12. licenčního studia č. 3330/LS340005/45	**koncem měsíce září 2007**

6. soustředění nového 11. licenčního studia č. 3330/LS360001/45:	**28. května- 1. června 2007**

Přehled základních informací k licenčnímu studiu:

Charakter studia	Organizace studia	Vyučující (foto)	Plán studia	Sylaby předmětů	Přihláška do adresáře	Absolventi, Foto1, 2, Diplomky
Návody úloh	Otázky ke zkoušce	Vzory semestr. prací	Potřebná literatura	Potřebný software	Zákony Murphyho	Promoce licen. studia

Charakter studia

Přihlášky a smlouvu do léta 2007,

Zaměření licenčního studium **Statistické zpracování dat při managementu jakosti** je zaměřeno na zvyšování odborné úrovně pracovníků kontrolních laboratoří tak, aby znalosti, technika práce a především způsob zpracování výsledků chemických zkoušek, experimentálních dat byly srovnatelné s laboratořemi zemí Evropské unie.

Určení: licenční studium je určeno pro stávající i budoucí pracovníky kontrolních laboratoří OTK, OKŘJ, dále pracovníky zdravotnických, veterinárních, vodohospodářských laboratoří, potravinářské a zemědělské inspekce, chemických, potravinářských, farmaceutických a zemědělských výrob. Dále pro pracovníky laboratoří kontroly životního prostředí, všech odvětví průmyslu, energetiky a zemědělství s důrazem na využití moderní instrumentální techniky a především zpracování výsledků pomocí matematicko-statistických metod na osobním počítači a s využitím nejmodernějšího programového vybavení.

cena 37.500,- Kč.

Přihlášky do licenčního studia: mohou se přihlásit pracovníci, kteří buď již pracují v oboru anebo se v rámci rekvalifikace potřebují seznámit s

Děkuji za pozornost!

Thank you for your attention!

http://meloun.upce.cz