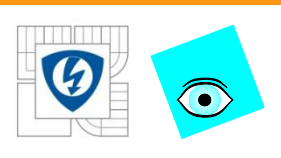


Počítačová analýza vícerozměrných dat v oborech přírodních, technických a společenských věd

Prof. RNDr. Milan Meloun, DrSc. (Univerzita Pardubice, Pardubice)

20.-24. června 2011

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky.

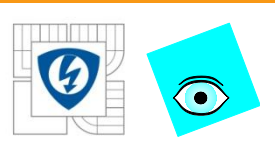


4.6 KANONICKÁ KORELAČNÍ ANALÝZA CCA

24.2.2010

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



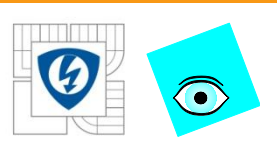


Zaměření metody CCA

Navržena v 1935 Hotellingem při hledání lineární kombinace jedné skupiny znaků $\mathbf{x} = (x_1, \dots, x_q)$, korelující s lineární kombinací druhé skupiny $\mathbf{y} = (y_1, \dots, y_p)$.

Cíl: Hledají se *lineární kombinace* znaků obou skupin, tj. hypotetických kanonických proměnných, které vedou k maximálním vzájemným korelacím.

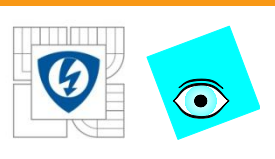
Vznikne *nový souřadnicový systém* vzájemně ortogonálních složek.



Zaměření metody CCA

V 1. kroku se hledá lineární kombinace $\mathbf{x}$ a lineární kombinace $\mathbf{y}$, jejichž korelace je maximální. Tyto lineární kombinace tvoří první složky souřadnicových systémů *kanonických proměnných* pro $\mathbf{x}$ a $\mathbf{y}$.

V dalších krocích se hledají *další lineární kombinace* $\mathbf{x}$ a $\mathbf{y}$ čili kanonické proměnné, které mají maximální vzájemnou korelaci a přitom jsou nekorelované s prvními kanonickými proměnnými.



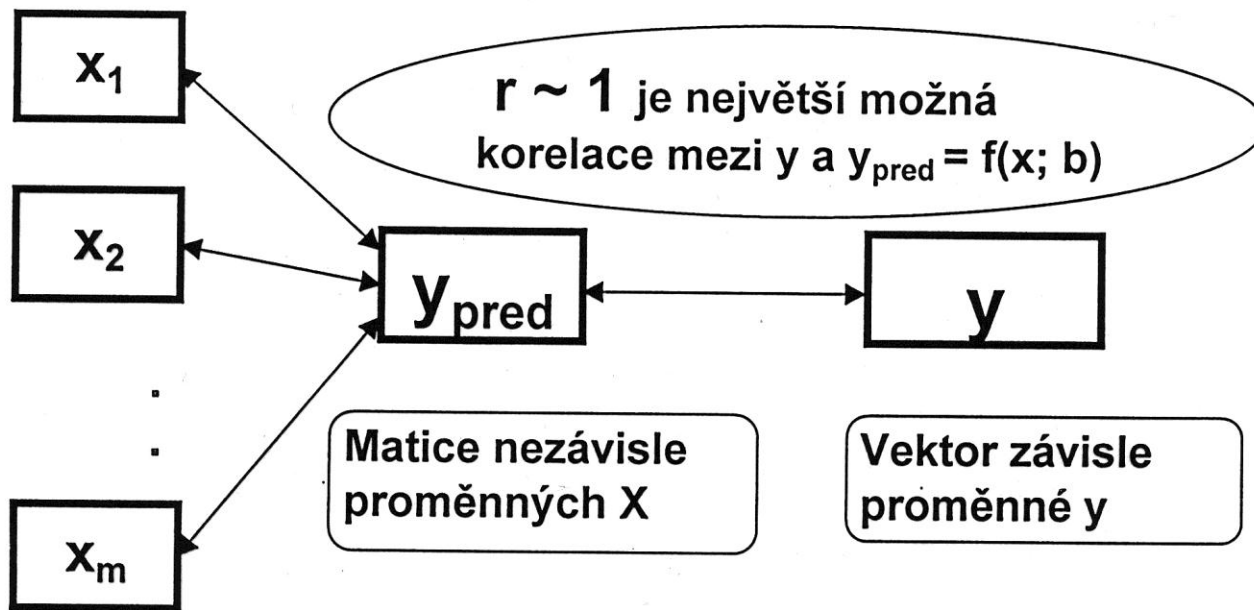
Zaměření metody CCA

Soubor nezávisle proměnných x

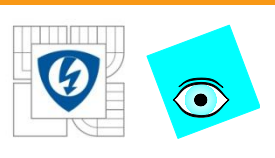
Závisle proměnná y

$$y_{\text{pred}} = b_1 x_1 + b_2 x_2 + \dots + b_m x_m$$

y



nejlepší odhad
parametrů $b_1, b_2, \dots, b_m$



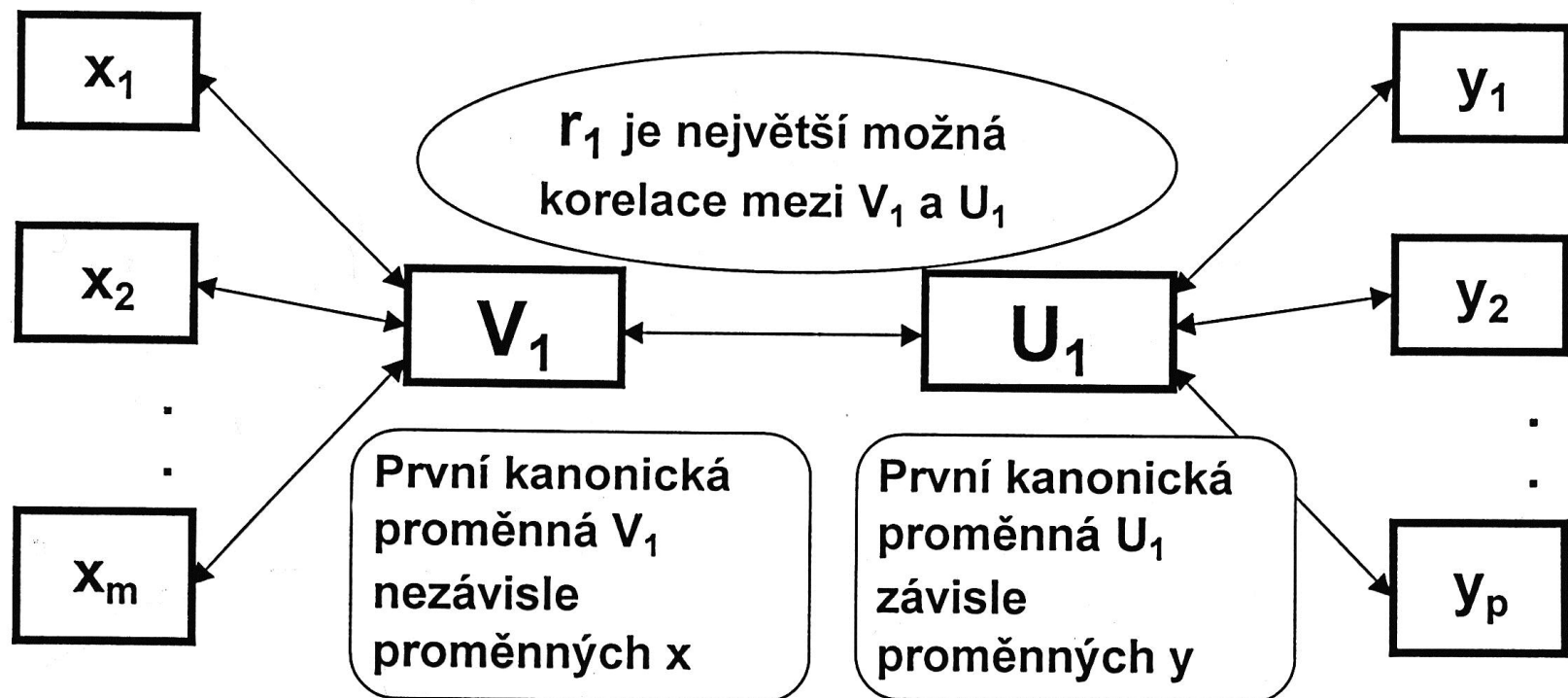
Zaměření metody CCA

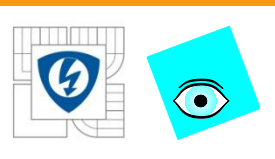
Soubor nezávisle proměnných x

$$V_1 = b_1 x_1 + b_2 x_2 + \dots + b_m x_m$$

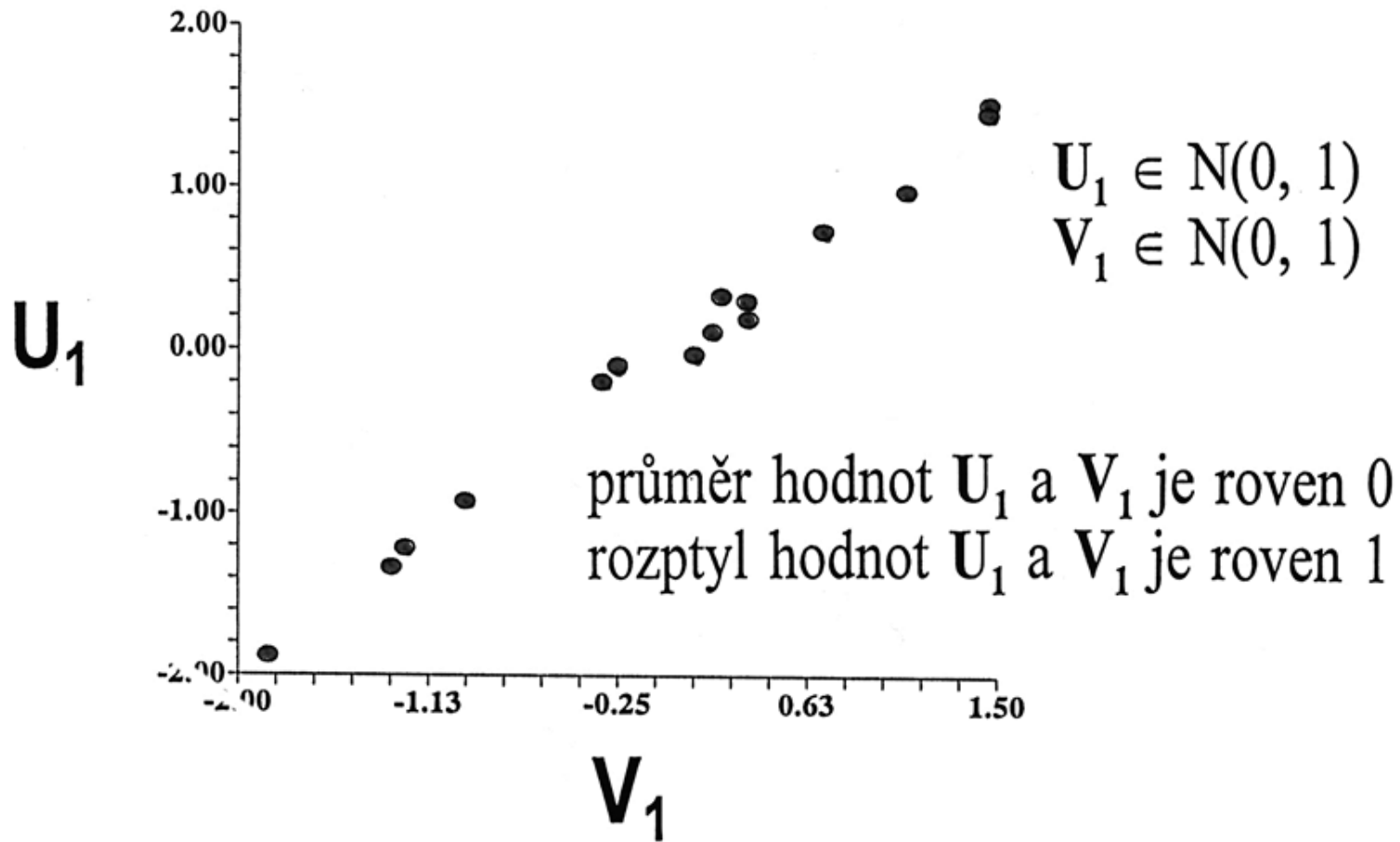
Soubor závisle proměnných y

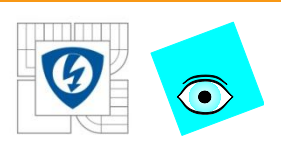
$$U_1 = a_1 y_1 + a_2 y_2 + \dots + a_p y_p$$





Nalézt nejlepší odhady koeficientů a a b , aby
pro všech n objektů vykazovaly U_i a V_i



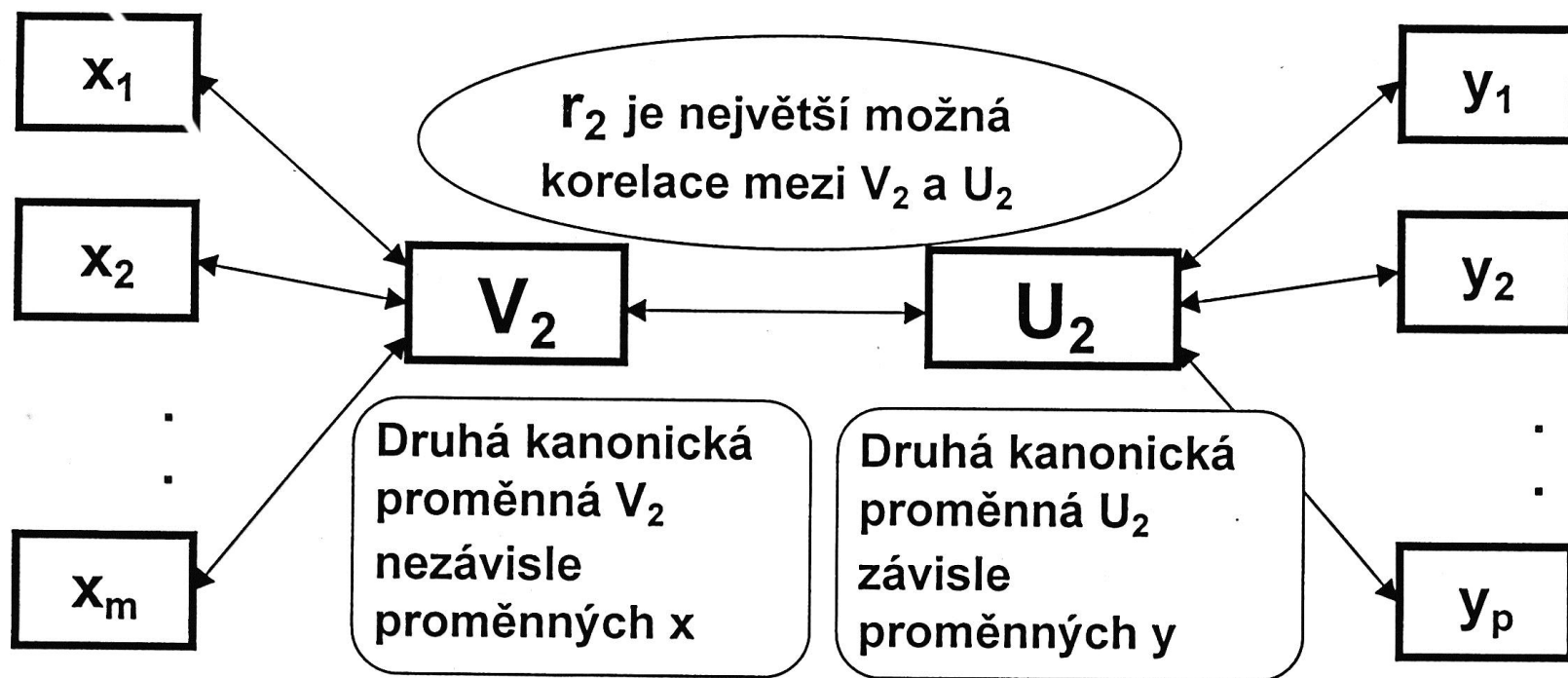


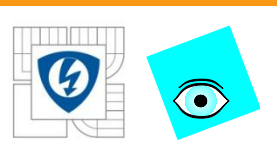
Soubor nezávisle proměnných x

$$V_2 = b_1 x_1 + b_2 x_2 + \dots + b_m x_m$$

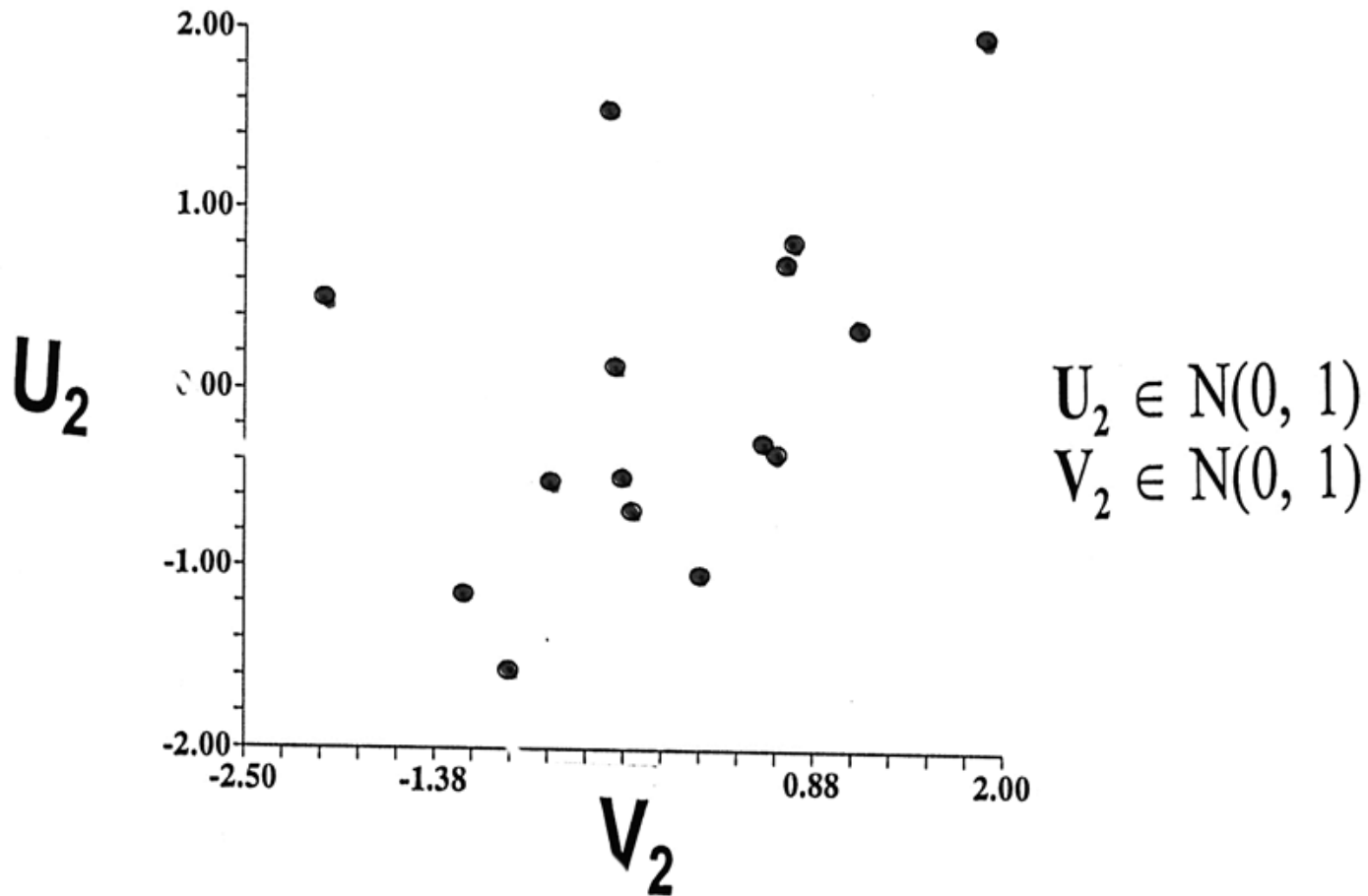
Soubor závisle proměnných y

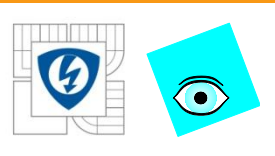
$$U_2 = a_1 y_1 + a_2 y_2 + \dots + a_p y_p$$





Nalézt nejlepší odhady koeficientů a a b , aby pro všech n objektů vykazovaly U_2 a V_2 druhý



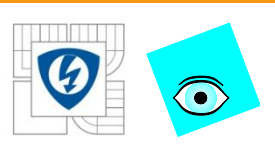


Podstata metody CCA

Vytváříme **vztah** mezi skupinami znaků $\mathbf{x} = (x_1, x_2, \dots, x_p)^T$ a $\mathbf{y} = (y_1, y_2, \dots, y_q)^T$. Lze předpokládat, že všechny znaky jsou **centrované**.

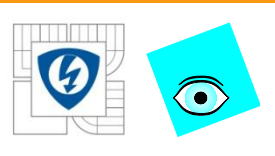
V kanonické korelační analýze se tvoří *(levá) kanonická proměnná* $U_1 = a_1 y_1 + a_2 y_2 + \dots + a_q y_q$ ze znaků $\mathbf{y}$ a *(pravá) kanonická proměnná* $V_1 = b_1 x_1 + b_2 x_2 + \dots + b_p x_p$ ze znaků $\mathbf{x}$.

Hledají se koeficienty a_i a b_i tak, aby kanonické proměnné U_1 a V_1 vykazovaly maximální **párový korelační koeficient** R .



Podstata metody CCA

V druhém kroku se pak hledají další kanonické proměnné U_2 a V_2 které mají druhý největší korelační koeficient ρ_2 a za podmínky, že U_2 a V_2 jsou nekorelované s prvními kanonickými proměnnými U_1 a V_1 .



Korelace R mezi U_1 a V_1 se nazývá první kanonická korelace.

Levá skupina vstupních proměnných,
znaků objektů:

$y_1, y_2, \dots, y_g$

Pravá skupina vstupních proměnných,
znaků objektů:

$x_1, x_2, \dots, x_p$

Kanonická analýza

nové kanonické proměnné

$R(x_j, y_i)$
kanonické zátěže

$$a_{10} + a_{11} \cdot y_1 + \dots + a_{1g} \cdot y_g =$$

kanonické váhy

$\dots =$

U_1

U_2

$\vdots$

U_n

levá
skupina

$$R(U_1, V_1)$$

$$R(U_2, V_2)$$

kanonická
korelace

$$R(U_n, V_n)$$

V_1

V_2

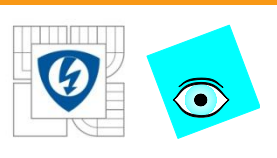
$\vdots$

V_n

pravá
skupina

$$= b_{10} + b_{11} \cdot x_1 + \dots + b_{1p} \cdot x_p$$

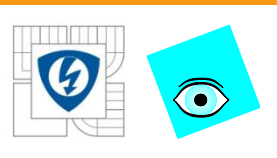
kanonické váhy



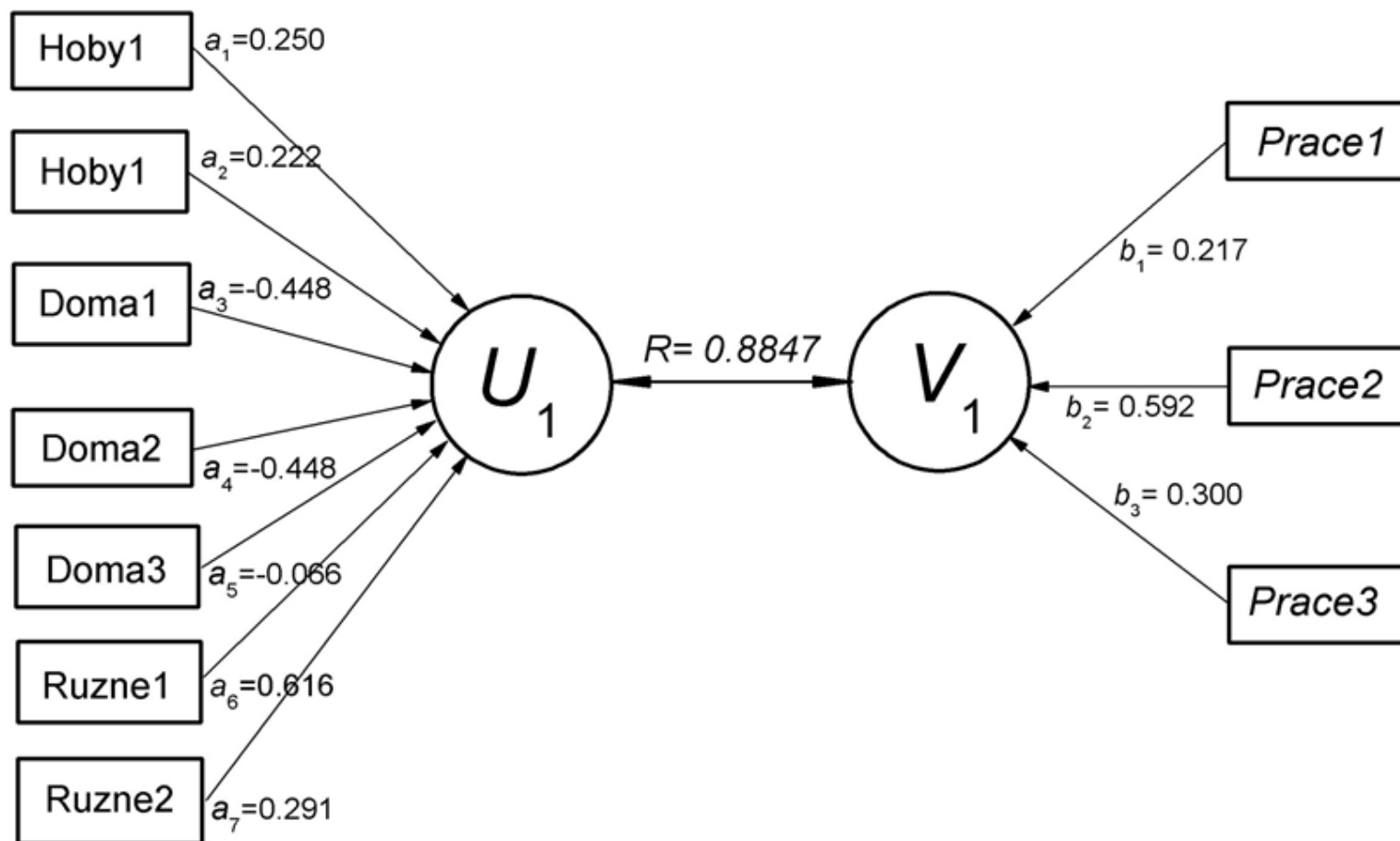
První kanonická korelace.

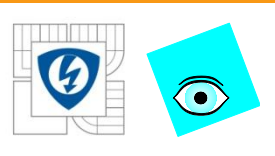
Je to největší možná korelace mezi lineárními kombinacemi znaků y a lineárními kombinacemi znaků x .

Představuje analogii k vícenásobnému korelačnímu koeficientu ve vícenásobné lineární regresi. Rozdíl je v tom, že zde existuje několik znaků y a je nutné navíc hledat optimální lineární kombinaci mezi nimi.



První pár kanonických korelačních proměnných U_1 a V_1 dostatečně popisuje závislost 7 znaků Hoby1, Hoby2, Doma1, Doma2, Ruzne1, Ruzne2 na 3 znacích Prace1, Prace2, Prace3 zdrojové matice dat Spokojenost (STATISTICA, ORIGIN)





První kanonická korelace.

Druhá kanonická proměnná V_2 je lineární kombinací $\mathbf{x}$ a jí odpovídající druhá kanonická proměnná U_2 je lineární kombinací $\mathbf{y}$.

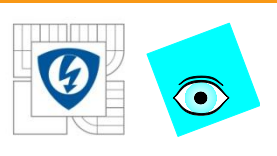
Koeficienty **a** a **b** jsou nyní vybrány tak, že současně platí:

V_2 je nekorelované s V_1 a U_1

U_2 je nekorelované s V_1 a U_1 .

V_2 a U_2 mají maximální možnou korelaci **R** , která je ale **nutně menší** než u V_1 a U_1 .

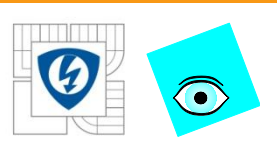
V dalších krocích se určují **další** kanonické proměnné U_3 a V_3 , U_4 a V_4 atd.



První kanonická korelace.

Maximální počet kanonických korelací je roven menšímu ze dvou čísel p a q , když p je počet znaků $\mathbf{x}$ a q je počet znaků $\mathbf{y}$.

Interpretace: Kanonické proměnné je nutno interpretovat jako hypotetické proměnné, které je často obtížné fyzikálně vysvětlit.

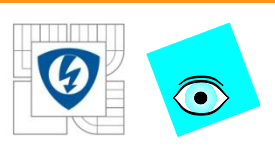


Důležitost každé kanonické proměnné

Důležitost každé kanonické proměnné se hodnotí ze dvou hledisek:

- a) Určuje se **intenzita vztahu** mezi kanonickou proměnnou U a původním znakem y a dále kanonickou proměnnou V a znaky x .
- b) Vyjadřuje se **intenzita vztahu** mezi oběma kanonickými proměnnými V a U .

CCA předpokládá pouze **lineární závislost** mezi proměnnými: je třeba vyšetřit grafy každého páru proměnných a prověřit linearitu a odlehlé body.

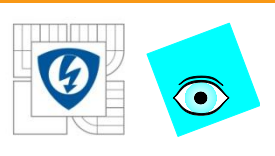


Výpočet: výběrové kanonické proměnné získáme ze dvou datových matic

– X rozměru $(n \times p)$ a Y rozměru $(n \times q)$, když $p \leq q$.

Společná výběrová kovarianční matice je pak pro případ, že jsou všechny znaky centrované (v odchylkách od průměrů jednotlivých znaků), rovna

$$S = \begin{pmatrix} S_{xx} & S_{xy} \\ S_{yx} & S_{yy} \end{pmatrix}$$



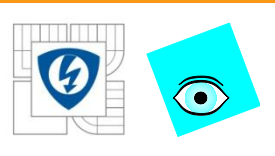
Kovarianční matice

Kovarianční matice znaků $\mathbf{x}$ je rovna $\mathbf{S}_{xx} = \frac{1}{n-1} \mathbf{X}^T \mathbf{X}$ a znaků $\mathbf{y}$ je rovna $\mathbf{S}_{yy} = \frac{1}{n-1} \mathbf{Y}^T \mathbf{Y}$.

Matice kovariancí znaků $\mathbf{x}$ a $\mathbf{y}$ je rovna $\mathbf{S}_{xy} = \frac{1}{n-1} \mathbf{X}^T \mathbf{Y}$.

Platí, že $\mathbf{S}_{yx} = \mathbf{S}_{xy}^T$.

Rozkladem matice $\mathbf{S}_{xx}^{-1} \mathbf{S}_{xy} \mathbf{S}_{yy}^{-1} \mathbf{S}_{yx} = \mathbf{V} \boldsymbol{\lambda} \mathbf{V}^T$ na vlastní čísla $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ a vlastní vektory $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_p$ jsou vyčísleny potřebné koeficienty $\mathbf{a}$ koeficienty $\mathbf{b}_i = \mathbf{V}_i$.



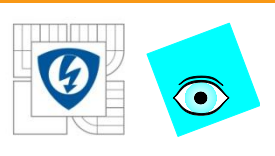
Vysvětlení pojmů

Koeficienty $\mathbf{a}_i$ se vyčíslí vztahem $\mathbf{a}_i = \frac{1}{\sqrt{\lambda_i}} \mathbf{S}_{yy}^{-1} \mathbf{S}_{xy} \mathbf{b}_i$.

Kanonické proměnné jsou reprezentovány **kanonickými skóre** $\mathbf{V} = \mathbf{XB}$ a $\mathbf{U} = \mathbf{YA}$, kde $\mathbf{B}$ obsahuje jako sloupce vlastní vektory $\mathbf{b}_i = \mathbf{V}_i$ a $\mathbf{A}$ obsahuje jako sloupce koeficienty $\mathbf{a}_i$.

K interpretaci kanonických proměnných se počítají také **matice zátěží** podle $\mathbf{L}_x = \mathbf{S}_{xx} \mathbf{B}$ a $\mathbf{L}_y = \mathbf{S}_{yy} \mathbf{A}$.

Pracuje se obvykle s **normovanými znaky**, pak se místo $\mathbf{S}$ užijí **matice korelační** $\mathbf{R}$.

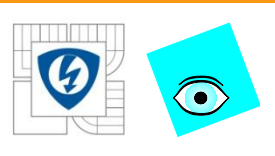


Vysvětlení pojmů

Používá se pouze **malý počet** kanonických proměnných, a to menší než $\min(p, q)$.

Pro normované znaky jsou L_x a L_y přímo korelačními koeficienty mezi znaky x a kanonickými proměnnými V , respektive mezi y a U .

Pro centrované znaky se získají korelační matice vynásobením $D_x^{-1} L_x$, respektive $D_y^{-1} L_y$, kde D matice obsahující na diagonále směrodatné odchylky znaků.



Test významnosti kanonických korelací

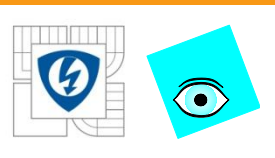
V CCA se vyčísľují:

1. Koeficienty pro všechny kanonické proměnné,
2. Kanonické korelace,
3. Kanonické proměnné pro všechny prvky výběru, zvané *kanonická skóre*.

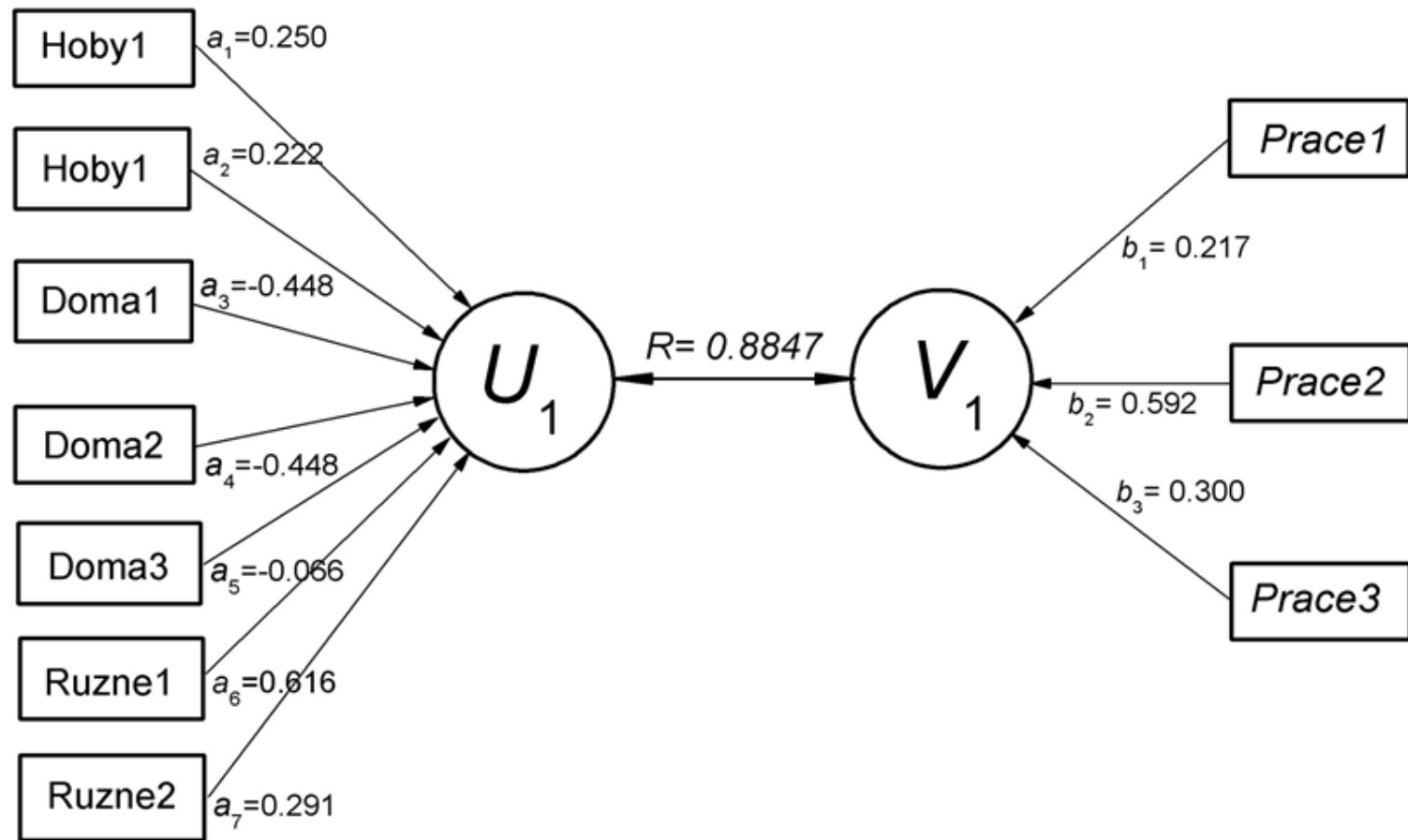
Test nulové hypotézy H_0 : „*k* nejmenších kanonických korelací souboru je rovno nule“.

Dva testy jsou vhodné: Bartlettův χ^2 -test a aproximativní F-test.

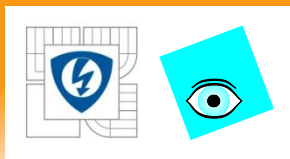
Velká hodnota χ^2 a velká hodnota F jsou indikátorem, že ne všechny korelace mezi proměnnými v souboru jsou nulové.



Test významnosti kanonických korelací



První pár kanonických korelačních proměnných U_1 a V_1 dostatečně popisuje závislost 7 znaků *Hoby1*, *Hoby2*, *Doma1*, *Doma2*, *Ruzne1*, *Ruzne2* na 3 znacích *Prace1*, *Prace2*, *Prace3* zdrojové matice dat *Spokojenost* (STATISTICA, ORIGIN).



Vysvětlení kanonických proměnných

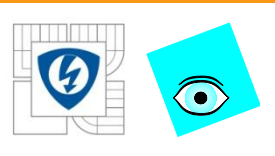
Korelace mezi kanonickými proměnnými a původními znaky:

vysvětlují kanonické proměnné, když některé ze znaků ať už ze skupiny x , nebo ze skupiny y jsou navzájem silně vnitřně korelované.

Nazývají se **zátěže kanonických proměnných** nebo **koeficienty kanonické struktury**:

- jsou to **korelace mezi** každým původním znakem a kanonickou proměnnou, - jsou **užitečné v pochopení** vztahu mezi znaky a kanonickými proměnnými.
- je-li skupina znaků v jedné kanonické proměnné **nekorelovaná**, kanonické zátěže jsou rovny standardizovaným koeficientům kanonické proměnné.
 - -jsou-li některé ze znaků **silně vnitřně korelovány**, pak zátěže a koeficienty jsou zcela rozličné.

Například: dva znaky x jsou vysoce kladně korelovány a každý je pozitivně korelován s kanonickou proměnnou. Pak je možné, že jeden koeficient kanonické proměnné bude kladný a jeden záporný, zatímco kanonické zátěže jsou obě kladné, jak by se dalo očekávat.



Analýza redundance

Průměrná hodnota čtverců kanonických zátěží pro první kanonickou proměnnou V_1 poskytuje podíl rozptylu znaků, který je vysvětlen první kanonickou proměnnou.

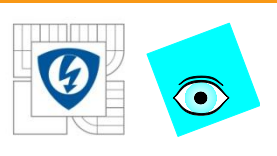
Obdobně pro U_1 a znaky y .

Analýza redundance (nadbytečnosti): podíl rozptylu takto objasněného může být malý, i když jde o vysokou kanonickou korelaci, a to v důsledku jednoho či dvou znaků, které mají hlavní vliv na kanonickou proměnnou.

Redundanční index je mírou průměrného podílu rozptylu ve skupině znaků y , který nebyl zahrnut do souboru V .

Je analogický R^2 vícenásobné lineární regresi.

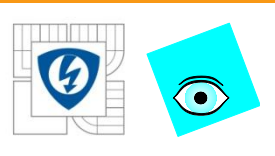
Podobně je možné získat i podíl rozptylu ve skupině znaků x , jež je objasněn všemi proměnnými U .



Graf závislosti skóre kanonických proměnných U_i na V_i :

Stupeň rozptýlení bodů po ploše je důležitým ukazatelem. Graf je užitečný v odkrývání neobvyklých případů ve výběrech jako jsou odlehlé body.

Graf U_1 na V_1 může vést ve zdánlivě nelineární rozptylový diagram nebo eliptický tvar typický pro dvojrozměrné normální rozdělení.

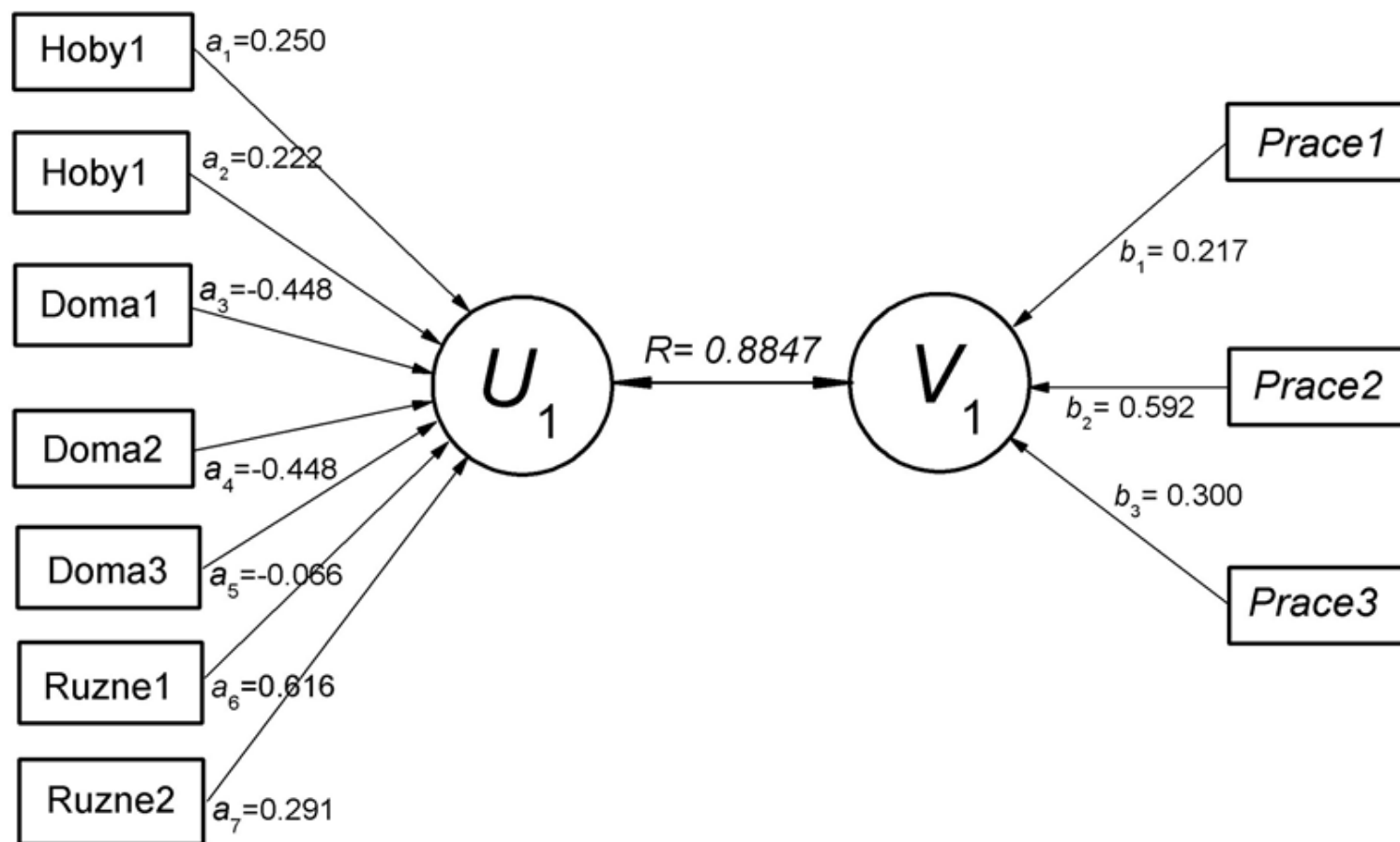
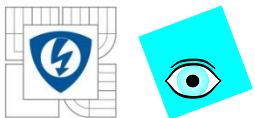


Postup diagnostikování CCA

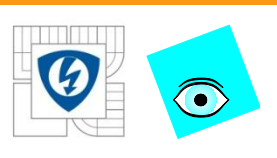
Je-li počet znaků q ve skupině y roven 1, pak se CCA redukuje na vícenásobnou lineární regresi.

CCA kvantifikuje obecně sílu vztahu mezi dvěma skupinami znaků.

Kanonické proměnné mohou být vysvětlovány podobným způsobem, jako byly vysvětlovány hlavní komponenty nebo faktory.

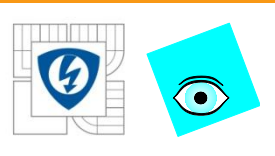


První pár kanonických korelačních proměnných U_1 a V_1 dostatečně popisuje závislost 7 znaků *Hoby1*, *Hoby2*, *Doma1*, *Doma2*, *Ruzne1*, *Ruzne2* na 3 znacích *Prace1*, *Prace2*, *Prace3* zdrojové matice dat *Spokojenost* (STATISTICA, ORIGIN).



PŘÍKLAD 6.1 Spokojenost v práci ovlivňuje některé domény spokojenosti v životě

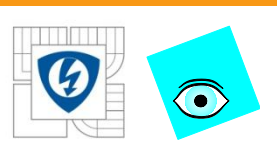
Byl analyzován dotazník o **10** fiktivních doménách spokojenosti v práci, doma v rodině a v zálibách a koníčcích v odpovědích vyplněný **100** respondenty. Cílem je vyšetřit, jak ovlivňují **3 domény** spokojenosti v práci na jedné straně zbývajících **7 domén** spokojenosti na druhé straně. Pokusíme se spokojenost v práci chápat jako nezávisle proměnnou V a spokojenost v ostatních jako závisle proměnnou U .



Příklad pokračování

- **Data:** Zdrojová matice dat *Spokojenost* je analyzována programy STATISTICA a NCSS2002.
- **Řešení:** Kanonická korelační analýza je založena na korelační matici znaků.

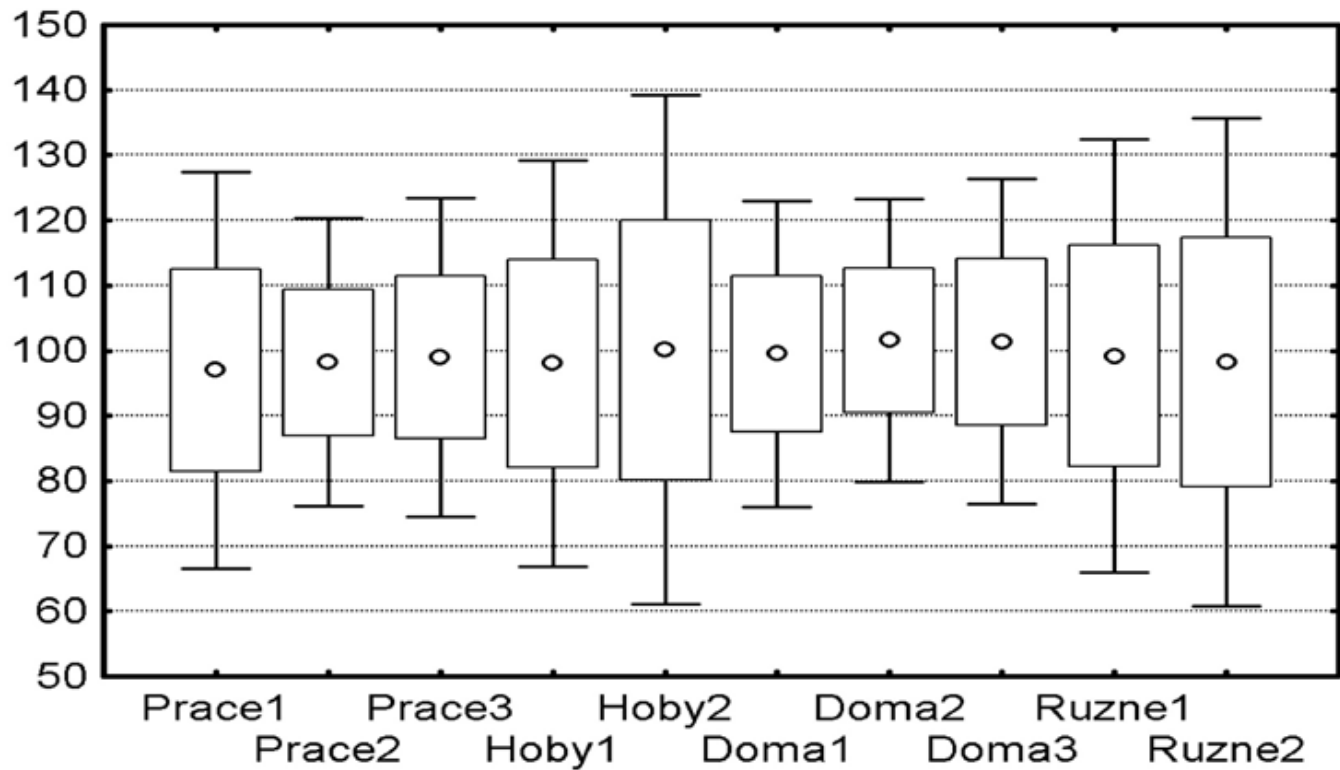
1. Exploratorní analýza dat EDA: Abychom vyšetřili rozdělení znaků a našli vybočující objekty, je vhodné sestavit dva EDA-grafy pro všechny znaky, a to krabicový graf a maticový diagram.

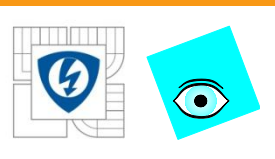


Krabicové grafy 10 znaků pro 100 respondentů dat Spokojenost (STATISTICA).

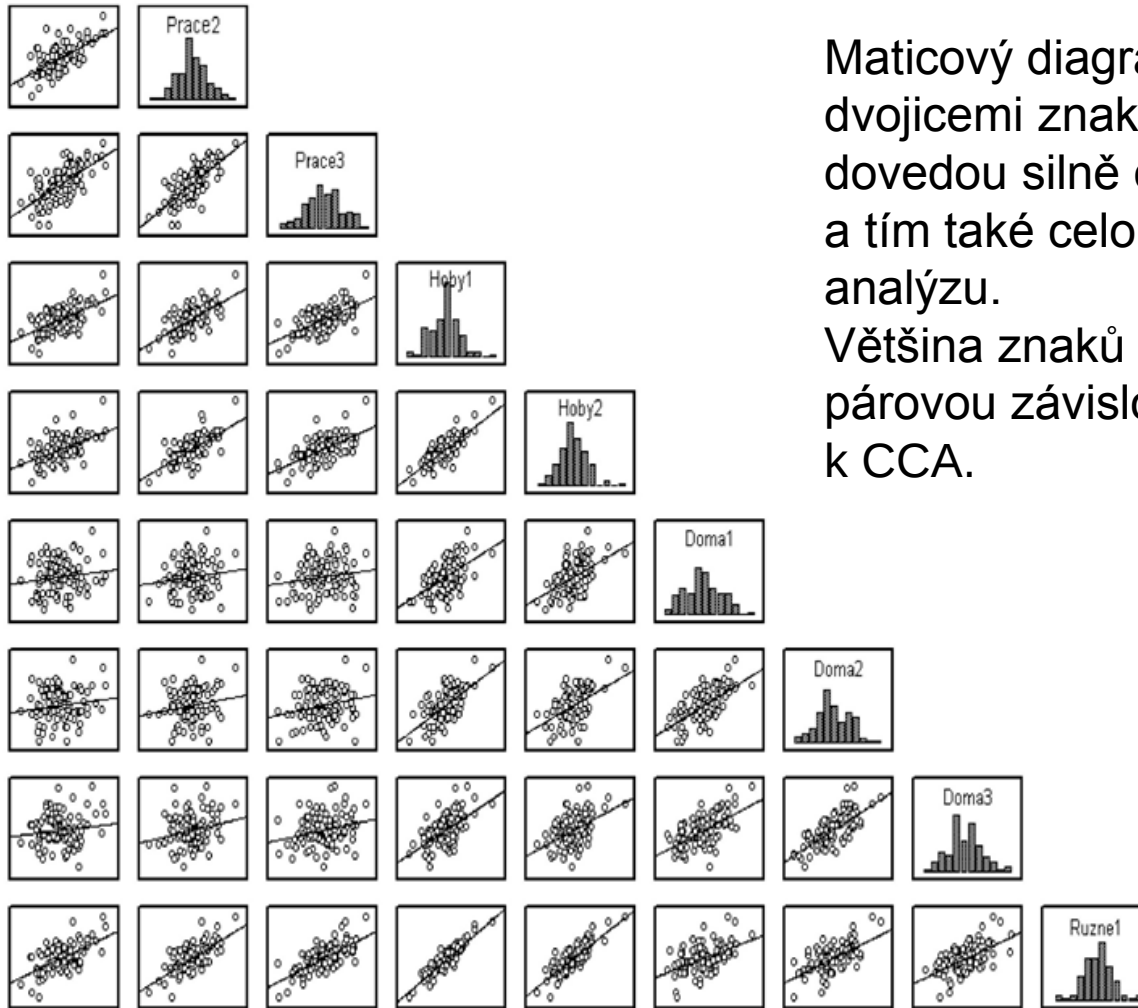
Největší proměnlivosti v datech je dosaženo u znaků *Hoby2*, *Prace1*, *Ruzne1*, *Ruzne2*.

I ostatní znaky vykazují dostatečnou proměnlivost.





Maticový diagram 10 znaků pro 100 respondentů dat Spokojenost (STATISTICA).



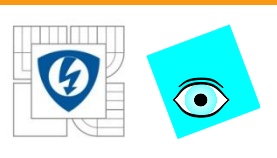
Maticový diagram dat indikuje vazby mezi dvojicemi znaků a odlehlé objekty, které dovedou silně ovlivnit korelační koeficient, a tím také celou kanonickou korelační analýzu. Většina znaků vykazuje dostatečnou párovou závislost a proto data jsou vhodná k CCA.



Korelační matice

	<i>Prace1</i>	<i>Prace2</i>	<i>Prace3</i>	<i>Hoby1</i>	<i>Hoby2</i>	<i>Doma1</i>	<i>Doma2</i>	<i>Doma3</i>	<i>Ruzne1</i>	<i>Ruzne2</i>
<i>Prace1</i>	1									
	<i>P= ---</i>									
<i>Prace2</i>	0.6474	1.0000								
	<i>P=0.0</i>	<i>P= ---</i>								
<i>Prace3</i>	.6526	.7319	1.0000							
	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>							
<i>Hoby1</i>	.5981	.6885	.6369	1.0000						
	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>						
<i>Hoby2</i>	.5211	0.6978	0.63	.8047	1.0000					
	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>					
<i>Doma1</i>	.1428	.1434	.1636	0.5364	.5059	1.0000				
	<i>P=.156</i>	<i>P=.155</i>	<i>P=.104</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>				
<i>Doma2</i>	.1451	.1819	.2383	.6343	.4959	.6577	1.0000			
	<i>P=.150</i>	<i>P=.070</i>	<i>P=.017</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>			
<i>Doma3</i>	.1378	0.236	.2546	.5828	.4824	.5900	.7306	1.0000		
	<i>P=.172</i>	<i>P=.018</i>	<i>P=.011</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>		
<i>Ruzne1</i>	.6113	.7086	.6979	.9045	.8110	.4984	0.6436	.5859	1.0000	
	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>	
<i>Ruzne2</i>	.5489	.6848	.6706	.8432	0.7558	.4247	0.5934	0.5177	0.8414	1
	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>

Ve druhém řádku jsou vypočtené hladiny významnosti *P*. **Platí pravidlo:** je-li $P < \alpha = 0.05$, je hodnota korelačního koeficientu statisticky významná (vyznačena **tučně**).



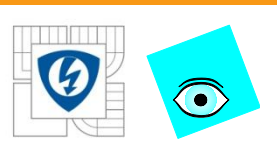
3. Zadání kanonických proměnných

Volba znaků a zařazení do obou obou kanonických proměnných, U a V , je v CCA čistě formální.

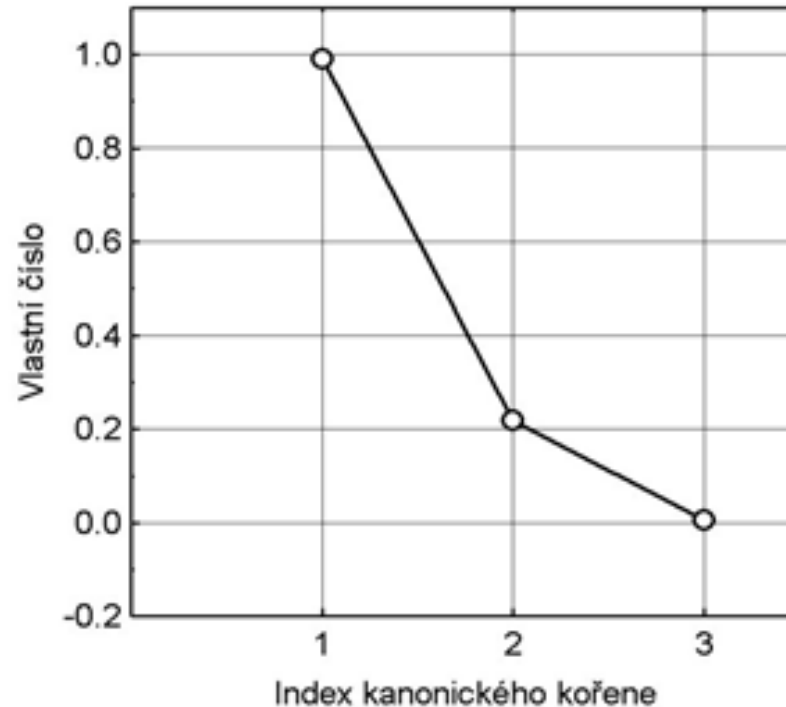
Obě U a V jsou totiž symetrické a stejné, z jejich povahy a obsahu nelze rozlišit.

Znaky do obou proměnných U a V lze libovolně zaměnit a obě proměnné jsou zcela rovnocenné.

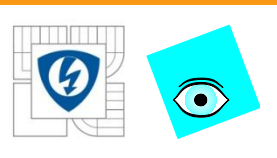
Kritériem počtu vybraných dvojic U a V je Cattelův indexový diagram úpatí vlastních čísel, který zde potvrzuje, že jeden kanonický kořen bude k popisu dat postačovat.



Cattelův indexový graf úpatí vlastních čísel



Cattelův indexový diagram úpatí vlastních čísel při kanonické korelaci dat 10 znaků pro 100 respondentů dat *Spokojenost* (STATISTICA).

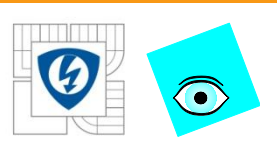


Rozptyl vyextrahovaný

	Pravý soubor nezávislých kanonických proměnných, V	Levý soubor závislých kanonických proměnných, U
Počet znaků	3	7
Rozptyl extrahovaný	100.000%	54.1271%
Celková redundance	61.5661%	33.2973%
Znaky: 1	<i>Prace1</i>	<i>Hoby1</i>
2	<i>Prace2</i>	<i>Hoby2</i>
3	<i>Prace3</i>	<i>Doma1</i>
4		<i>Doma2</i>
5		<i>Doma3</i>
6		<i>Ruzne1</i>
7		<i>Ruzne2</i>

Rozptyl vyextrahovaný značí průměrné množství proměnlivosti, vyextrahovaného ze znaků v obou souborech kanonickými proměnnými.

Kanonické proměnné **extrahují** 100% rozptylu na pravé straně rovnice ze tří znaků *Prace 1*, *Prace2*, *Prace3* a 54.13% rozptylu na levé straně rovnice.

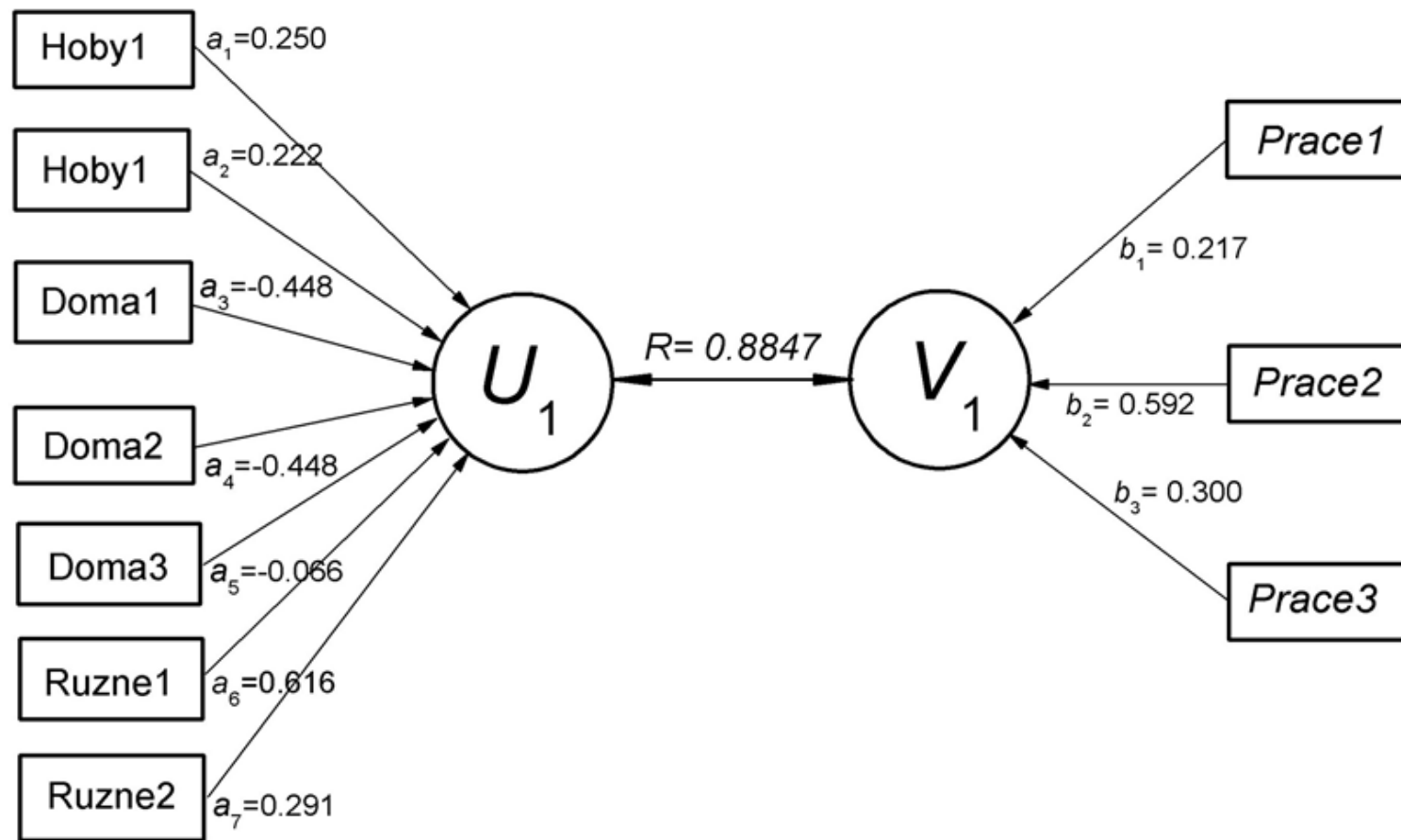
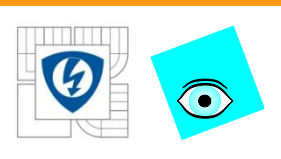


Celková redundance

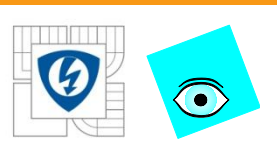
Celková redundance ukazuje na velikost celkové korelace mezi znaky na pravé straně rovnice 61.6% a na levé straně 33.3%.

Hodnoty **jsou založeny** na všech kanonických kořenech, totiž znacích na levé straně, jež se týkají domén spokojenosti v práci.

Podobně je 33.3% rozptylu nepracovních domén spokojenosti v životě na levé straně rovnice.



První pár kanonických korelačních proměnných U_1 a V_1 dostatečně popisuje závislost 7 znaků *Hoby1*, *Hoby2*, *Doma1*, *Doma2*, *Ruzne1*, *Ruzne2* na 3 znacích *Prace1*, *Prace2*, *Prace3* zdrojové matice dat *Spokojenost* (STATISTICA, ORIGIN).

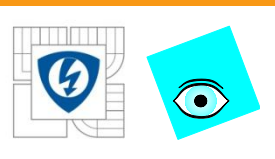


4. Test významnosti kanonických kořenů:

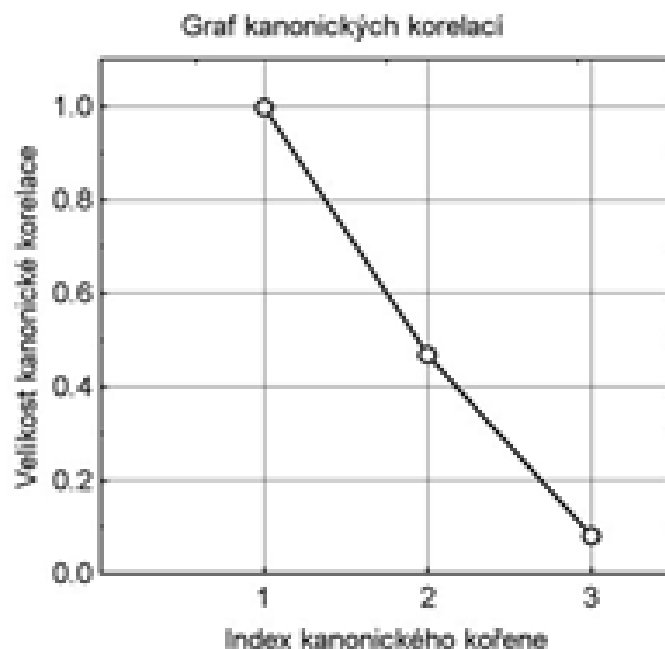
vyšetřuje, zda všechny tři kanonické kořeny U_1 a V_1 , U_2 a V_2 , U_3 a V_3 jsou statisticky významné.

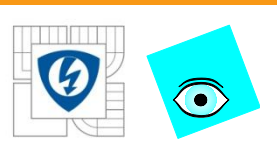
Kanonický korelační koeficient R představuje v datech pouze první kořen U_1 a V_1 čili nejsilnější a nejvýznamnější kanonickou korelaci.

Maximální počet kanonických kořenů, které mohou být z dat extrahovány je roven 3.



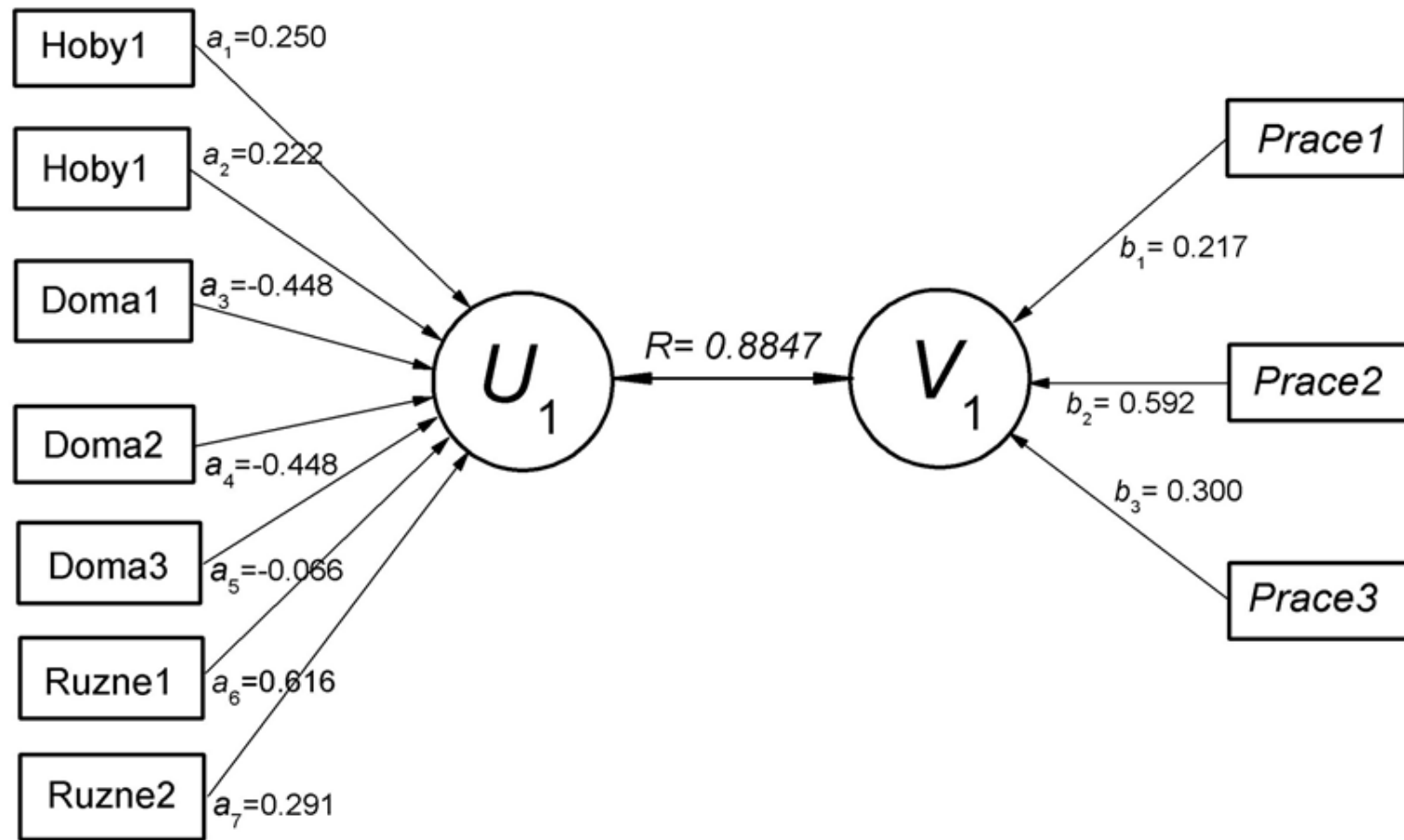
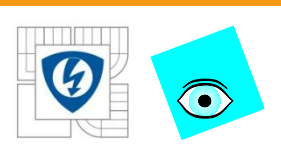
Hledání potřebných kanonických kořenů z grafu kanonických korelací pro 10 znaků a 100 respondentů dat Spokojenost (STATISTICA).



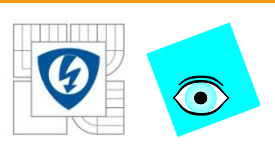


Pokračování

$R = 0.8847$ je statisticky vysoce významný, protože má vypočtenou hladinu významnosti **P** menší než $\alpha = 0.05$ a může být vysvětlen jako jednoduchá korelace mezi váženou sumou znaků, odpovídajícími první a nejvýznamnější kanonické proměnné **U_1** a **V_1**



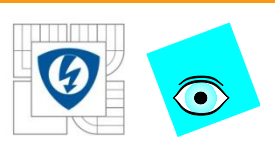
První pár kanonických korelačních proměnných U_1 a V_1 dostatečně popisuje závislost 7 znaků *Hoby1*, *Hoby2*, *Doma1*, *Doma2*, *Ruzne1*, *Ruzne2* na 3 znacích *Prace1*, *Prace2*, *Prace3* zdrojové matice dat *Spokojenost* (STATISTICA, ORIGIN).



Testování kanonických kořenů

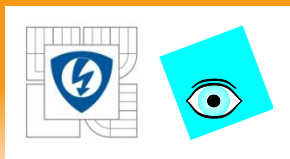
Diskutujme tři řádky *testování kanonických kořenů*:

1. V **prvním řádku** nebyl odstraněn žádný kanonický kořen a testy jsou vysoce statisticky významné.
2. V **druhém řádku** byl první a nejvýznamnější kořen U_1 a V_1 odstraněn a testy už nejsou statisticky významné. Můžeme proto testování ukončit a uzavřít, že statisticky významný je pouze první kanonický kořen U_1 a V_1 .
3. Když by byl i druhý test statisticky významný, postup by pokračoval dál až do takového řádku, kdy by už statistické testy vycházely statisticky nevýznamné.



Data

Vypuštěný kořen	Kanonické R	Kanonické D	χ^2	df	P	$Lambda$
Žádný	0.885	0.783	153.58	21	0.000	0.193
První U_1 a V_1	0.271	0.073	10.85	12	0.542	0.890
První U_1 a V_1 a druhý U_2 a V_2	0.197	0.039	3.72	5	0.591	0.961



5. Struktura kanonických faktorů a redundance:

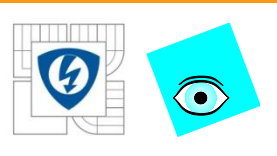
Nadále budeme uvažovat pouze první kanonický kořen U_1 a V_1

Jak vysvětlíme tento kanonický kořen U_1 a V_1 ?

Jak koreluje tento kanonický kořen U_1 a V_1 se znaky v obou souborech?

Vyčíslíme proto korelace mezi znaky v každém souboru s kanonickým kořenem. Tyto korelace se také nazývají *strukturní koeficienty*.

Nesmíme přitom zapomenout, že kanonická proměnná je v každém souboru tvořena jako vážená suma znaků.



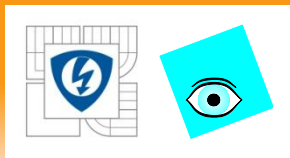
Faktorová struktura v pravém souboru:

pouze první kanonický kořen je statisticky významný a pouze ten budeme vysvětlovat.

Tři znaky spokojenosti *Prace1*, *Prace2* a *Prace3* vykazují podstatně vyšší zátěže prvního kanonického faktoru čili vysoké korelace právě s tímto faktorem.

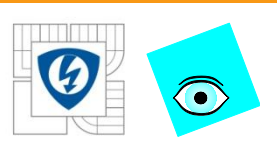
Mírou redundance je průměrné množství rozptylu vyčísleného pro každý znak prvního kanonického kořene.

Sečtou se čtverce vah kanonických faktorů a vydělí se počtem znaků ve výběru, zde to je 3.



Zátěže čili korelace s kanonickou proměnnou V

	<i>Kořen 1, V_1</i>	<i>Kořen 2, V_2</i>	<i>Kořen 3, V_3</i>
<i>Prace1</i>	0.796	-0.575	0.186
<i>Prace2</i>	0.953	0.101	-0.287
<i>Prace3</i>	0.875	0.217	0.432

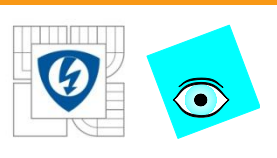


Tabulka extrahovaného rozptylu na pravé straně:

první kanonický kořen U_1 a V_1 extrahuje průměr okolo 77% rozptylu ze tří položek spokojenosti v práci.

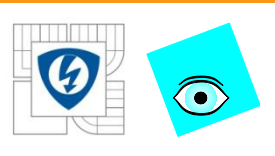
Když vynásobíme hodnotu sdíleného rozptylu mezi kanonickými proměnnými v obou souborech tj. R^2 , obdržíme redundanci.

U znaků na levé straně znaků spokojenosti s nepracovními doménami se vyčíslí okolo 60% rozptylu položek spokojenosti v práci, založených na prvním kanonickém kořenu.



Extrahovaný rozptyl a redundance u kanonické proměnné V

	Rozptyl	Redundance
<i>Kořen 1, V_1</i>	0.769	0.602
<i>Kořen 2, V_2</i>	0.129	0.010
<i>Kořen 3, V_3</i>	0.101	0.004

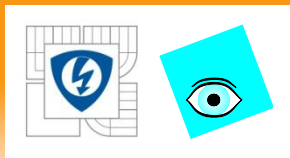


Faktorová struktura v levém souboru.

První kanonický kořen *Kořen 1*, U_1 se vyznačuje vysokými vahami u znaků spokojenosti *Hoby1*, *Hoby2*.

U spokojenosti doma v rodině *Doma1*, *Doma2*, *Doma3* jsou váhy mnohem nižší.

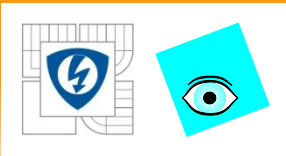
Významné kanonické korelace mezi znaky v obou souborech založené na prvním kanonickém kořenu jsou asi výsledkem vztahu mezi znaky spokojenost v práci, spokojenost se zálibami a koníčky a ostatní spokojeností v životě.



Zátěže čili korelace s kanonickou proměnnou U

Když budeme uvažovat spokojenost v práci jako nezávisle proměnnou, můžeme říci, že spokojenost v práci ovlivňuje spokojenost v zálibách a koníčcích ale mnohem méně spokojenost doma v rodině.

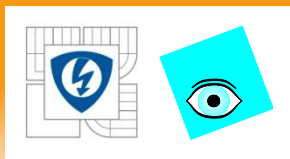
	Kořen 1, U_1	Kořen 2, U_2	Kořen 3, U_3
<i>Hoby1</i>	0.824	-0.100	0.055
<i>Hoby2</i>	0.809	0.278	-0.150
<i>Doma1</i>	0.187	-0.017	0.237
<i>Doma2</i>	0.238	0.253	0.456
<i>Doma3</i>	0.278	0.414	0.173
<i>Ruzne1</i>	0.861	0.049	0.328
<i>Ruzne2</i>	0.821	0.245	0.241



Tabulka extrahovaného rozptylu na levé straně.

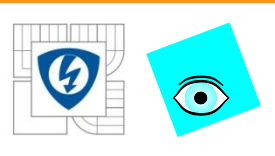
První kanonický kořen vy čí sluje 42% rozptylu znaků na levé straně a okolo 33% rozptylu ostatních spokojeností, založených na prvním kanonickém kořenu.

Je třeba si všimnout, že tato čísla jsou "zatlačena dolů" relativním nedostatkem korelace mezi kanonickou proměnnou a znaky spokojenosti doma v rodině.



Extrahovaný rozptyl a redundance u kanonické proměnné U

	Rozptyl	Redundance
<i>Kořen 1, U_1</i>	0.417	0.326
<i>Kořen 2, U_2</i>	0.055	0.004
<i>Kořen 3, U_3</i>	0.069	0.003

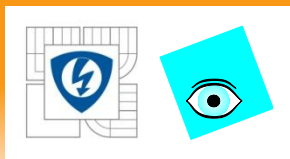


6. Kanonické skóre:

Kanonické proměnné $\mathbf{U}$ a $\mathbf{V}$ představují vážené součty znaků v každém souboru, $\mathbf{U}_1 = a_1 \mathbf{y}_1 + a_2 \mathbf{y}_2 + \dots + a_q \mathbf{y}_q$ a $\mathbf{V}_1 = b_1 \mathbf{x}_1 + b_2 \mathbf{x}_2 + \dots + b_p \mathbf{x}_p$

Váhy a_1, a_2, a_3 a b_1, b_2, b_3 , zobrazené v tabulce kanonických vah odpovídají standardizovaným (Z-transformace) znakům v obou souborech.

Užijeme tyto váhy k výpočtu skóre pro kanonické proměnné.

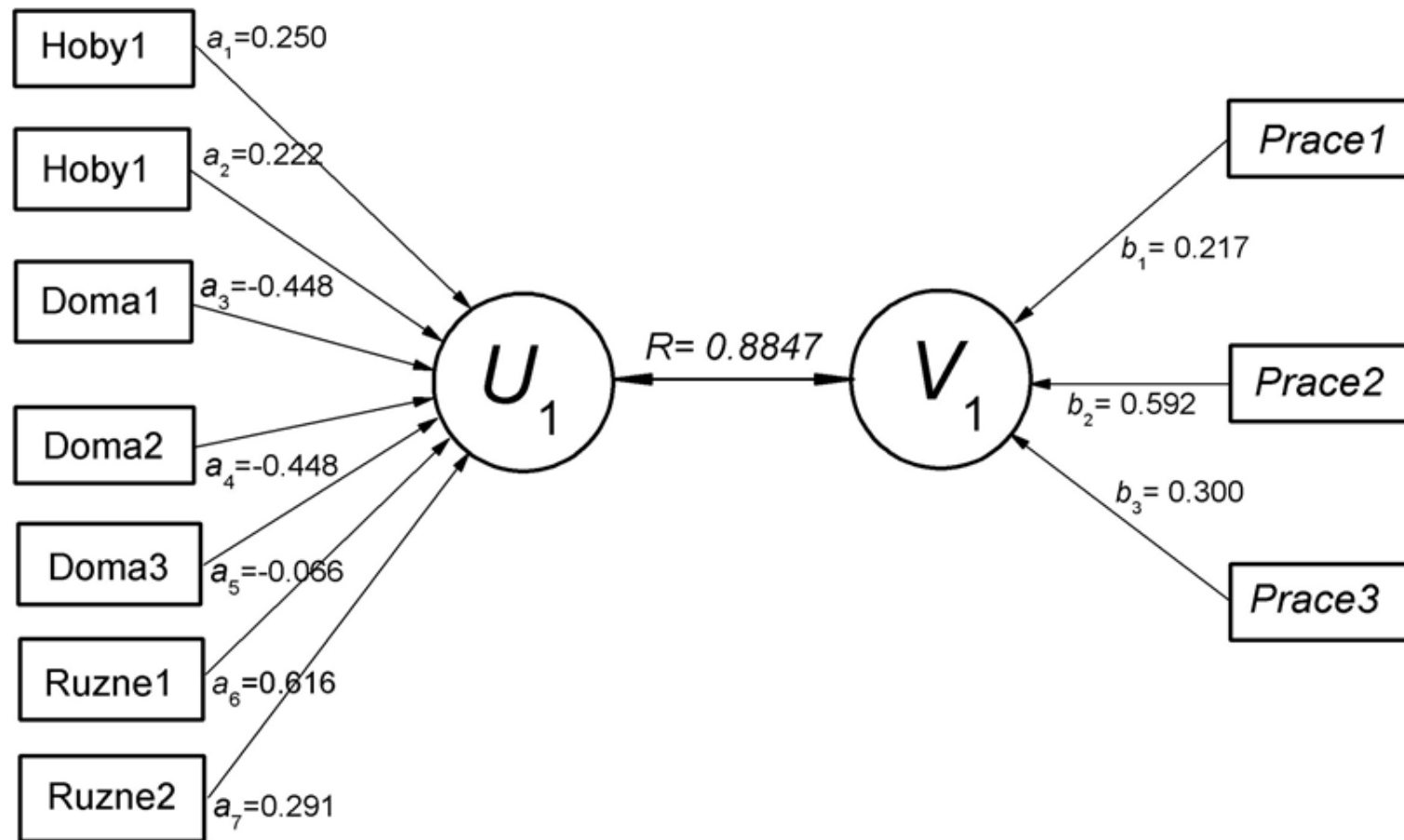
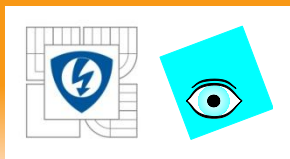


Váhy znaků a_1, a_2, a_3 k výpočtu skóre u kanonické proměnné U

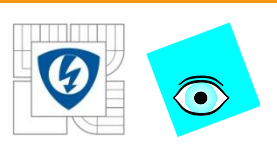
	Kořen 1, U_1	Kořen 2, U_2	Kořen 3, U_3
<i>Hoby1</i>	0.250	-1.636	-1.291
<i>Hoby2</i>	0.222	1.005	-1.096
<i>Doma1</i>	-0.156	-0.434	0.310
<i>Doma2</i>	-0.448	0.326	0.457
<i>Doma3</i>	-0.066	0.700	-0.317
<i>Ruzne1</i>	0.616	-0.361	1.737
<i>Ruzne2</i>	0.291	0.798	0.458

Váhy znaků b_1, b_2, b_3 k výpočtu skóre u kanonické proměnné V

	Kořen 1, V_1	Kořen 2, V_2	Kořen 3, V_3
<i>Prace1</i>	0.217	-1.359	0.244
<i>Prace2</i>	0.592	0.371	-1.388
<i>Prace3</i>	0.300	0.832	1.289



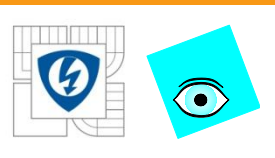
První pár kanonických korelačních proměnných U_1 a V_1 dostatečně popisuje závislost 7 znaků *Hoby1*, *Hoby2*, *Doma1*, *Doma2*, *Ruzne1*, *Ruzne2* na 3 znacích *Prace1*, *Prace2*, *Prace3* zdrojové matice dat *Spokojenost* (STATISTICA, ORIGIN).



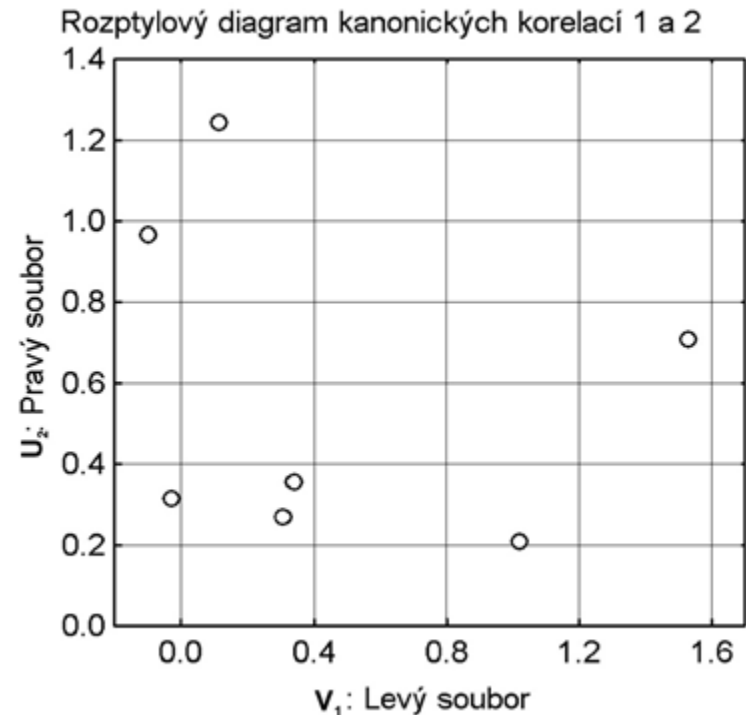
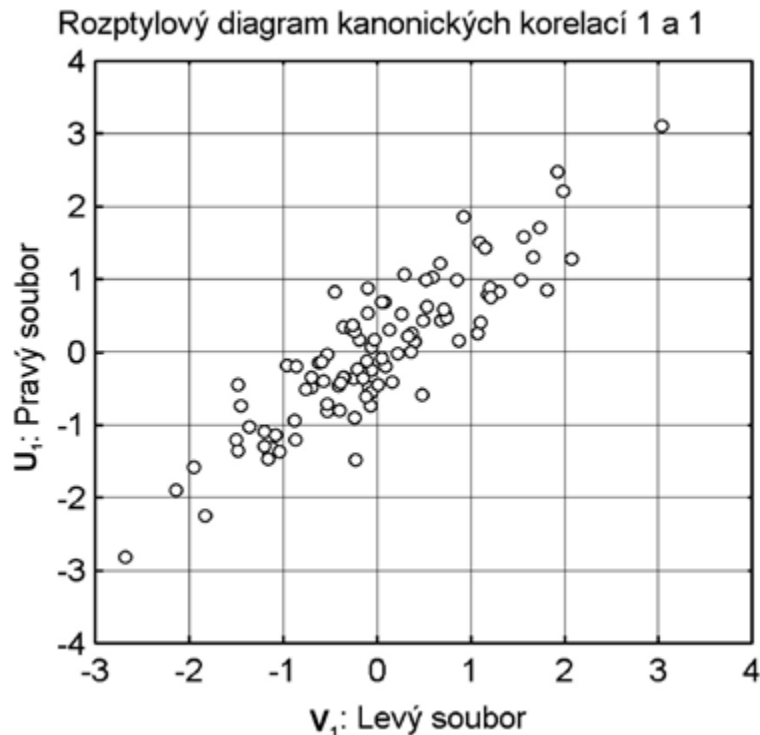
7. Grafy kanonických skóre:

Zobrazíme kanonické skóre znaků na pravé straně čili **V** proti skórum znaků na levé straně čili **U** a nenacházíme žádné vybočující objekty a ani rezidua po regresi neindikují nelineární trendy.

Objekty leží na přímce a nikoliv na křivce tvaru písmen S nebo U.

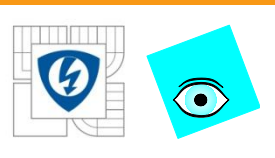


Rozptylové diagramy

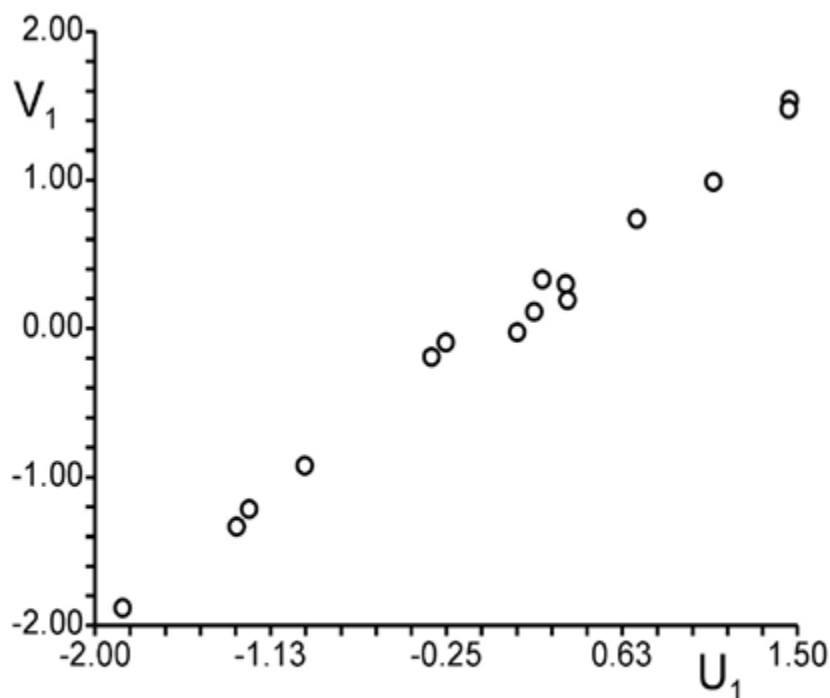


Rozptylový diagram kanonických korelací $V_1 = f(U_1)$ pro 10 znaků a 100 respondentů dat *Spokojenost* (STATISTICA).

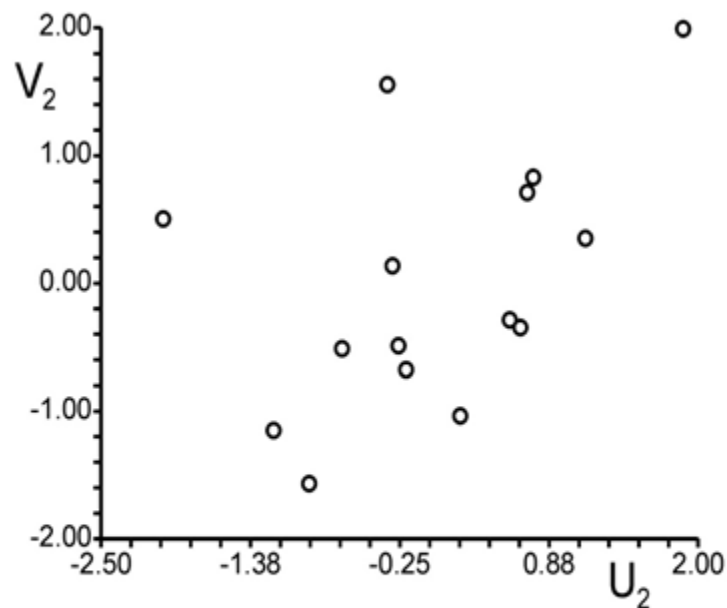
Rozptylový diagram kanonických korelací $U_1 = f(V_1)$ pro 10 znaků a 100 respondentů dat *Spokojenost* (STATISTICA).



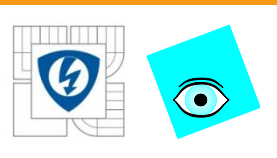
8. Vyšetření shluků objektů: je zde možnost indikovat shluky objektů na obrázku z programu NCSS2002.



Rozptylový diagram kanonických proměnných U_1 a V_1 pro 10 znaků a 100 respondentů dat *Spokojenost* (NCSS2002).



Rozptylový diagram kanonických proměnných U_2 a V_2 pro 10 znaků a 100 respondentů dat *Spokojenost* (NCSS2002).



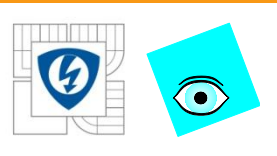
Závěr:

spokojenost v práci významně ovlivňuje spokojenost při zábavě a i celkově.

Spokojenost doma v rodině se nejeví příliš ovlivněna spokojeností v práci.

V praxi by měla být tato dotazníková úloha s různými skupinami respondentů opětovně vyhodnocena touto metodou.

Tím by se potvrdilo, zda kanonická faktorová struktura vede k věrohodné a spolehlivé interpretaci prvního kanonického kořene, tzv. *replikovatelnosti*.



1. Cíle kanonické korelační analýzy

Data tvoří dvě skupiny znaků.

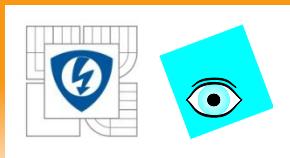
Každá skupina má jiný teoretický význam, jedna může být definovaná jako znaky nezávisle proměnné a druhá jako závisle proměnné.

CCA umožní řešení těchto úloh:

a) Rozhodnout, zda skupiny znaků, představujících obvykle rozličná měření na témže předmětu, **jsou nezávislé jedna na druhé**, nebo naopak určit regresní vztah, který existuje mezi oběma skupinami.

b) Nalezení skupiny znaků, aby lineární kombinace každé skupiny byly **maximálně korelovány**.

c) **Objasnění povahy** kteréhokoliv ze vztahů, jež existují mezi skupinami znaků **x a y** , stanovením relativního příspěvku každého znaku do kanonické funkce.



2. Formulace úlohy kanonické korelační analýzy

Kritická je **velikost výběru** a dostatečný počet pozorování na jeden znak, a to alespoň 10 pozorování na 1 znak.

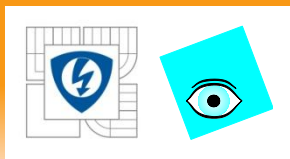
Láká zahrnovat **hodně znaků**, neuvažujícíe důsledky malého výběru zakrytí smysluplného vztahu.

Malé výběry budou indikovat vždy statistickou významnost, protože bude nalezena dokonalá lineární kombinace díky malému počtu stupňů volnosti.

Klasifikace znaků na závislé a nezávislé je **čistě formální** a není vůbec důležitá.

Kanonická korelační analýza váží obě kanonické proměnné tak, aby **maximalizovala korelaci** a nedává žádný zvláštní důraz na některou z kanonických proměnných.

Složení kanonické proměnné je důležité: musí existovat **smysluplný základ** původních znaků, ze kterých budou vytvářeny kanonické proměnné.



2. Předpoklady kanonické korelační analýzy

Korelační koeficient má smysl, pokud mezi dvěma znaky existuje **lineární vztah**.

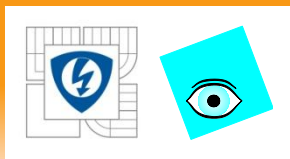
Při nelineárním je třeba vhodně transformovat.

CCA může být použita **pro metrické znaky** bez splnění předpokladu normality.

Normalita je požadována k provedení testů.

CCA může být použita **i pro nenormálně** rozdělené znaky.

Heteroskedasticita měla by být potlačena.



4. Nalezené řešení a dosažená těsnost proložení

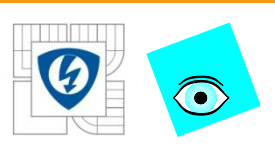
Síla vztahu mezi páry kanonických proměnných je vyjádřena v kanonickém korelačním koeficientu.

Čtverec kanonických korelací představuje velikost sdíleného rozptylu mezi dvěma kanonickými proměnnými. Tyto čtverce nazývají **kanonické kořeny**.

Rutinou je analyzovat kanonické proměnné o statisticky významných korelačních koeficientech.

Nevýznamné kanonické proměnné nemají vztahy mezi znaky vysvětleny.

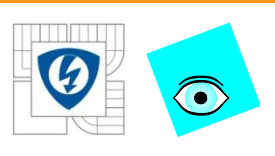
Nestačí užití jednoduchého kritéria $H_0: \rho_i = 0$.



Pokračování

Jsou doporučována **tři kritéria rozhodování** vysvětlení kanonické proměnné:

- a) hladina významnosti,
- b) velikost kanonické korelace a
- c) míra redundance.

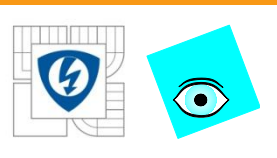


Vysvětlení kritérií rozhodování

Hladina významnosti: vypočtená hladina významnosti P je uvažována za přijatelnou, pokud je rovna nebo menší než zvolená hodnota $\alpha = 0.05$, $P < \alpha$ pak je testovaný **korelační koeficient statisticky významný**.

Velikost kanonických korelací: by měla být vzata v úvahu, i když neexistuje obecný návod o vhodné velikosti kanonických korelací.

Míra redundance sdíleného rozptylu: i když mírou redundance je sdílený rozptyl, může vést tento ukazatel k chybným vysvětlením, protože čtverec kanonických korelací představuje rozptyl sdílený lineárními kombinacemi skupin znaků a ne rozptyl určený ze skupin znaků.

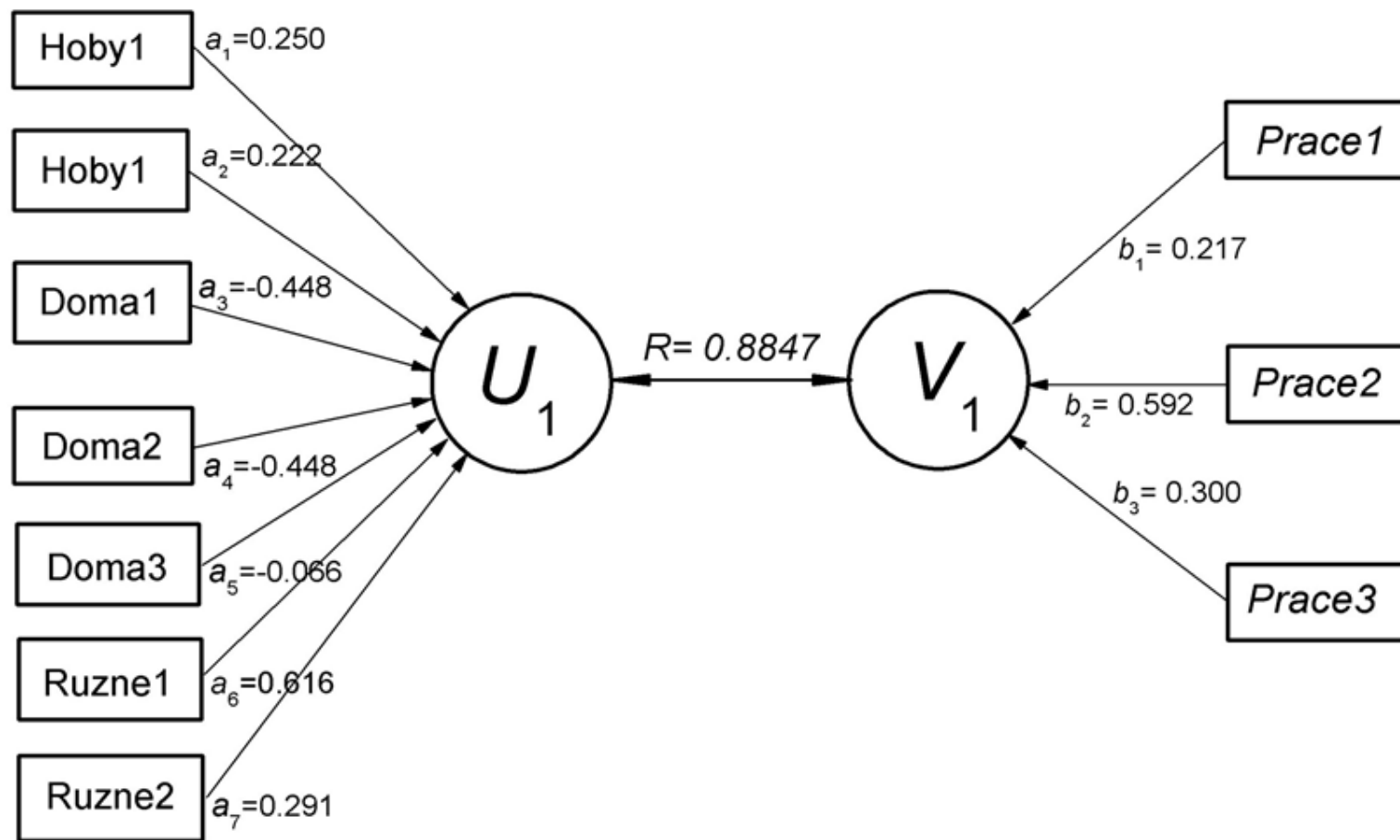
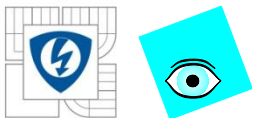


Stewartův-Loveův index redundancy

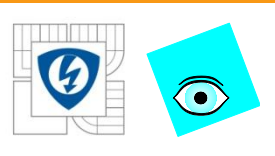
vyčísluje velikost rozptylu v jedné skupině znaků, který může být objasněn rozptylem ve druhé skupině znaků.

Je součinem dvou složek: **První** je mírou sdíleného rozptylu mezi původním znakem a kanonickou proměnnou a vypočte se jako aritmetický průměr čtverců zátěží L_i^2

Druhou složkou je podíl rozptylu v kanonické závislé proměnné, jenž může být objasněn kanonickou nezávislou proměnnou. Její čtverec korelace mezi kanonickou nezávislou proměnnou a kanonickou závislou proměnnou, který je znám jako **čtverec kanonického korelačního koeficientu R^2** .



První pár kanonických korelačních proměnných U_1 a V_1 dostatečně popisuje závislost 7 znaků *Hoby1*, *Hoby2*, *Doma1*, *Doma2*, *Ruzne1*, *Ruzne2* na 3 znacích *Prace1*, *Prace2*, *Prace3* zdrojové matice dat *Spokojenost* (STATISTICA, ORIGIN).



5. Interpretace výsledků

Vysvětlení kanonických proměnných a určení relativní důležitosti každého z původních znaků v kanonických proměnných třemi metodami:

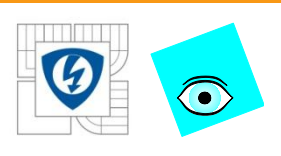
a) **Kanonické váhy**: vysvětlení kanonických proměnných se týká vyšetření **znaménka a velikosti kanonické váhy** charakterizované pro každý znak v parametrech **a , b** .

Původní znaky s **velikými váhami** přispívají více do kanonických proměnných.

Podobně znaky s **vahami opačného znaménka** vykazují inverzní vztah s každým znakem.

Znaky s váhami **stejného znaménka** vykazují přímý vztah.

Malá váha může znamenat, nevýznamný znak v kanonické proměnné.

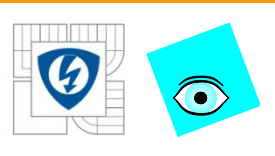


b) **Kanonické zátěže** také zvané **korelace kanonické struktury**, měří lineární korelaci mezi původním znakem a kanonickými proměnnými.

Vyjadřují rozptyl, který sdílejí původní znaky s kanonickými proměnnými.

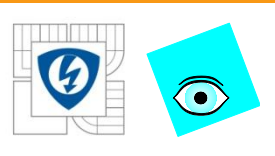
Může být vysvětlen jako **faktorová zátěž** vyčíslená jako relativní příspěvek každé původní proměnné do každé kanonické funkce. Každá kanonická proměnná se uvažuje **odděleně** a vyčísluje se korelace skupin znaků vůči kanonické proměnné.

Také **kritéria k určování významnosti korelací** kanonické struktury jsou stejná jako u faktorových zátěží.



c) ***Kanonické křížové zátěže***: týká se korelací každého znaku y se znaky x a opačně.

Konvenční zátěže **korelují** s původními znaky. Kanonické křížové zátěže vyjadřují lépe vztah mezi znaky x a y .



6. Ověření výsledků

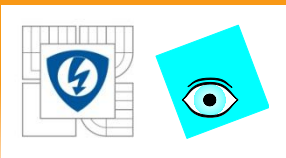
CCA by měla ověřit, že výsledky nejsou pouze specifické pro daný výběr dat, ale že se dají zobecnit.

Postup: vytvoří se dva dílčí výběry dat a provede se analýza s každým výběrem odděleně.

Pak jsou porovnány kanonické funkce, zátěže kanonických proměnných atd.

Když je nalezen vysoký rozdíl, je třeba uvažovat o jiném způsobu vyšetřování dat.

Kanonická korelace pouze **maximalizuje korelace**, takže se kanonické váhy a zátěže podstatně změní, je-li jeden znak odebrán z kterékoliv kanonické proměnné.



7. Diagnostikování problémů kanonické korelační analýzy

CCA má tato **kritická úskalí**:

a) Výběr dat by měl být **reprezentativní a náhodný**.

b) **Špatná spolehlivost** měření může vést k nižším odhadům korelací mezi znaky.

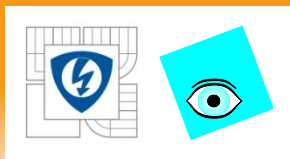
c) Mají se vyhledat **odlehle hodnoty** pomocí rozptylových diagramů znaků.

d) Když znaky **přispívají pouze málo**, jsou obvykle kandidáty na odstranění.

e) Je třeba ověřit kanonickou korelaci **vyšetřením koeficientů**, zda kanonická korelace je dost vysoká.

f) **Liší-li se** podstatně kanonické koeficienty od kanonických zátěží, tj. mají-li třeba rozdílné znaménko, pak je nutné je objasnit.

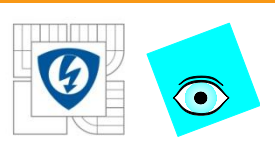
g) Jelikož kanonická korelace užívá obou skupin znaků **y** a souboru **x** , je celkový počet znaků použitých velký, což může způsobit problém u **úloh s dírami v datech**.



PŘÍKLAD 6.1 Spokojenost v práci ovlivňuje některé domény spokojenosti v životě

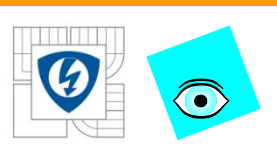
Byl analyzován dotazník o **10** fiktivních doménách spokojenosti v práci, doma v rodině a v zálibách a koníčcích v odpovědích vyplněný **100** respondenty. Cílem je vyšetřit, jak ovlivňují **3 domény** spokojenosti v práci na jedné straně zbývajících **7 domén** spokojenosti na druhé straně. Pokusíme se spokojenost v práci chápat jako nezávisle proměnnou V a spokojenost v ostatních jako závisle proměnnou U .

- *Data*: Zdrojová matice dat *Spokojenost* je analyzována programy STATISTIC A a NCSS2002.
- *Řešení*: Kanonická korelační analýza je založena na korelační matici znaků.

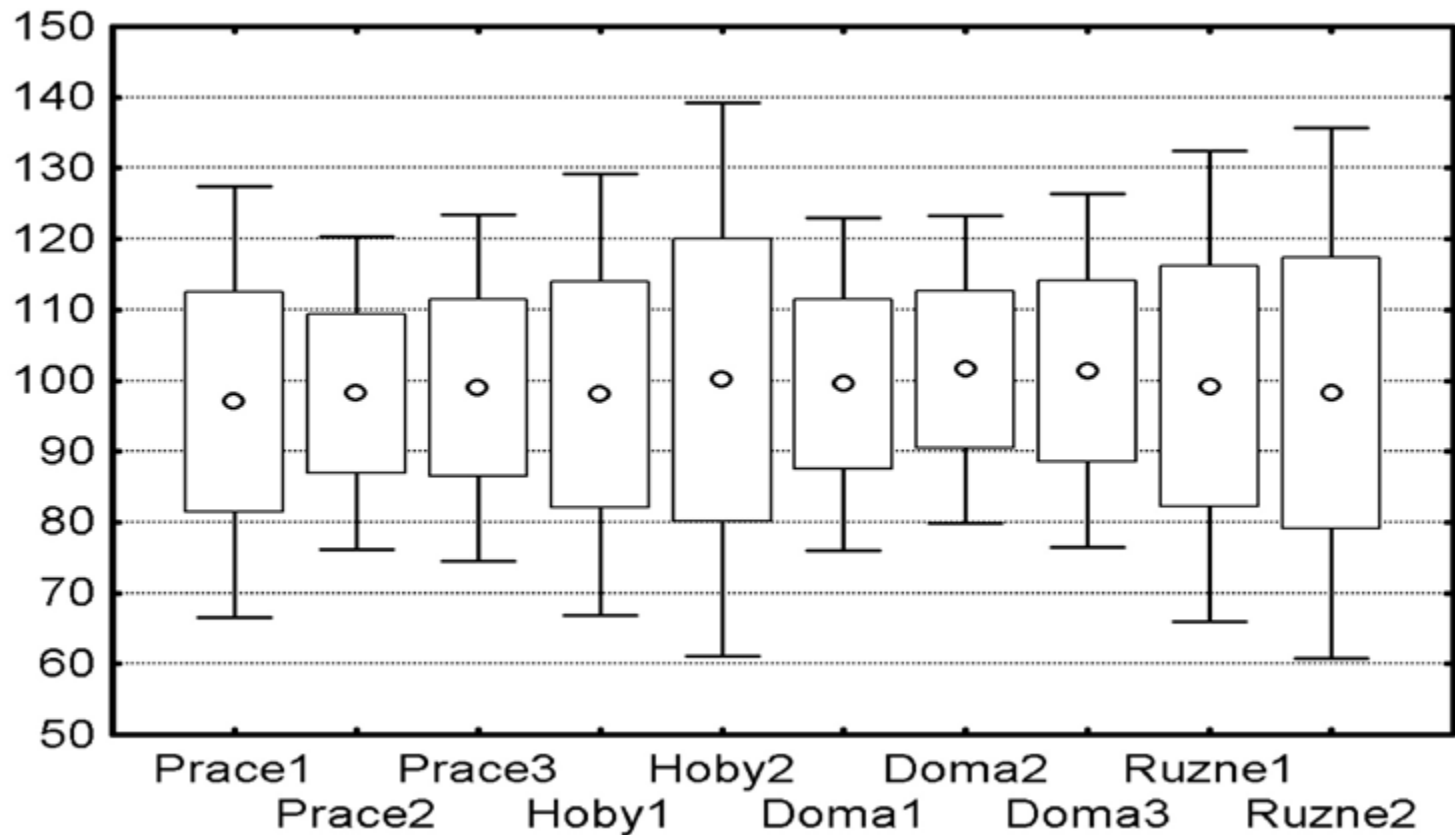


1. Exploratorní analýza dat EDA:

Abychom vyšetřili rozdělení znaků a našli vybočující objekty, je vhodné sestavit dva EDA-grafy pro všechny znaky, a to krabicový graf a maticový diagram.

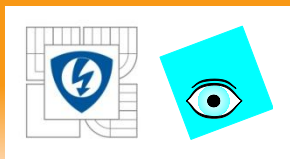


Krabicové grafy 10 znaků pro 100 respondentů dat Spokojenost (STATISTICA).



Největší proměnlivosti v datech je dosaženo u znaků *Hoby2*, *Prace1*, *Ruzne1*, *Ruzne2*.

I ostatní znaky vykazují dostatečnou proměnlivost.

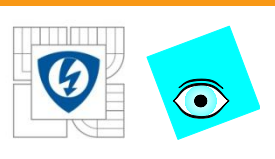


2. Korelační koeficienty všech párů znaků zdrojové matice dat:

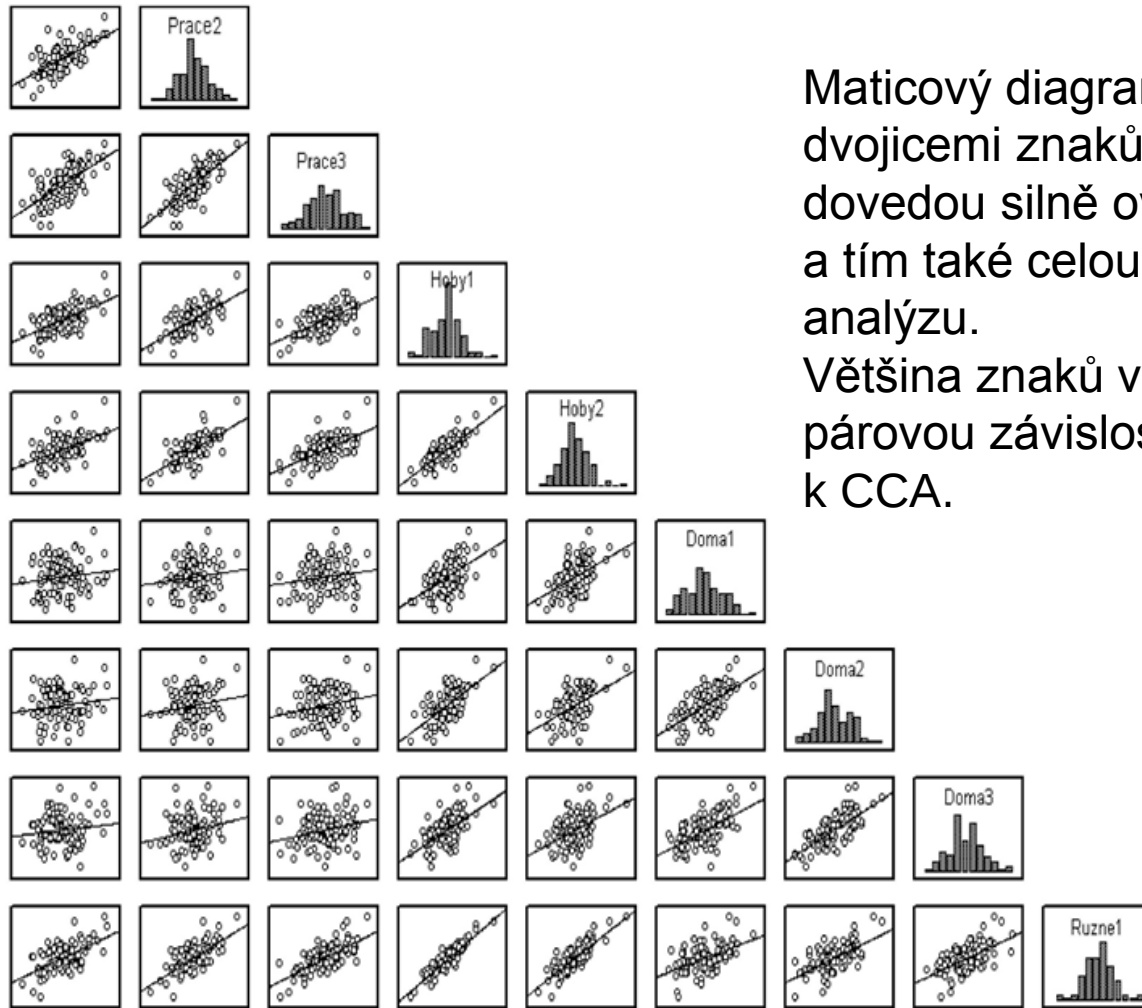
Vyčíslíme parametry polohy a rozptýlení všech 10 znaků.

Největší proměnlivost v datech vykazuje *Hoby2*, *Prace1*, *Ruzne1*, *Ruzne2*.

	Průměr	Směrodatná odchylka
<i>Prace1</i>	97.03	15.52
<i>Prace2</i>	98.17	11.27
<i>Prace3</i>	98.94	12.49
<i>Hoby1</i>	98.03	15.94
<i>Hoby2</i>	100.11	19.94
<i>Doma1</i>	99.51	11.98
<i>Doma2</i>	101.60	11.06
<i>Doma3</i>	101.37	12.73
<i>Ruzne1</i>	99.16	16.97
<i>Ruzne2</i>	98.22	19.11

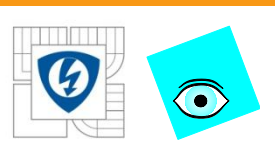


Maticový diagram 10 znaků pro 100 respondentů dat Spokojenost (STATISTICA).



Maticový diagram dat indikuje vazby mezi dvojicemi znaků a odlehlé objekty, které dovedou silně ovlivnit korelační koeficient, a tím také celou kanonickou korelační analýzu.

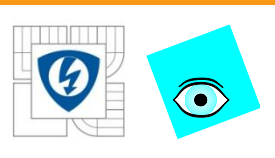
Většina znaků vykazuje dostatečnou párovou závislost a proto data jsou vhodná k CCA.



Korelační matice

	<i>Prace1</i>	<i>Prace2</i>	<i>Prace3</i>	<i>Hoby1</i>	<i>Hoby2</i>	<i>Doma1</i>	<i>Doma2</i>	<i>Doma3</i>	<i>Ruzne1</i>	<i>Ruzne2</i>
<i>Prace1</i>	1									
	<i>P= ---</i>									
<i>Prace2</i>	0.6474	1.0000								
	<i>P=0.0</i>	<i>P= ---</i>								
<i>Prace3</i>	.6526	.7319	1.0000							
	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>							
<i>Hoby1</i>	.5981	.6885	.6369	1.0000						
	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>						
<i>Hoby2</i>	.5211	0.6978	0.63	.8047	1.0000					
	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>					
<i>Doma1</i>	.1428	.1434	.1636	0.5364	.5059	1.0000				
	<i>P=.156</i>	<i>P=.155</i>	<i>P=.104</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>				
<i>Doma2</i>	.1451	.1819	.2383	.6343	.4959	.6577	1.0000			
	<i>P=.150</i>	<i>P=.070</i>	<i>P=.017</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>			
<i>Doma3</i>	.1378	0.236	.2546	.5828	.4824	.5900	.7306	1.0000		
	<i>P=.172</i>	<i>P=.018</i>	<i>P=.011</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>		
<i>Ruzne1</i>	.6113	.7086	.6979	.9045	.8110	.4984	0.6436	.5859	1.0000	
	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>	
<i>Ruzne2</i>	.5489	.6848	.6706	.8432	0.7558	.4247	0.5934	0.5177	0.8414	1
	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P=0.0</i>	<i>P= ---</i>

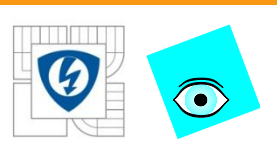
Ve druhém řádku jsou vypočtené hladiny významnosti P . **Platí pravidlo:** je-li $P < \alpha = 0.05$, je hodnota korelačního koeficientu statisticky významná (vyznačena **tučně**).



Kovarianční matice:

10 znaků zdrojové matice dat *Spokojenost*, týkající se 100 respondentů. Na diagonály jsou uvedeny rozptyly znaků.

	<i>Prace1</i>	<i>Prace2</i>	<i>Prace3</i>	<i>Hoby1</i>	<i>Hoby2</i>	<i>Doma1</i>	<i>Doma2</i>	<i>Doma3</i>	<i>Ruzne1</i>	<i>Ruzne2</i>
<i>Prace1</i>	240.9									
<i>Prace2</i>	113.2	127								
<i>Prace3</i>	126.5	103	155.9							
<i>Hoby1</i>	147.9	123.6	126.7	254.0						
<i>Hoby2</i>	161.3	156.8	156.9	255.7	397.6					
<i>Doma1</i>	26.6	19.3	24.5	102.4	120.8	143.5				
<i>Doma2</i>	24.9	22.7	32.9	111.8	109.4	87.1	122.4			
<i>Doma3</i>	27.2	33.9	40.5	118.3	122.5	90.0	102.9	162.2		
<i>Ruzne1</i>	161.0	135.5	147.8	244.6	274.4	101.3	120.8	126.6	287.9	
<i>Ruzne2</i>	162.8	147.5	160	256.8	288	97.2	125.4	126	272.9	365.2



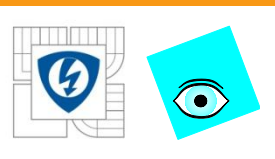
3. Zadání kanonických proměnných:

Volba znaků a zařazení do obou obou kanonických proměnných, U a V , je v CCA čistě formální.

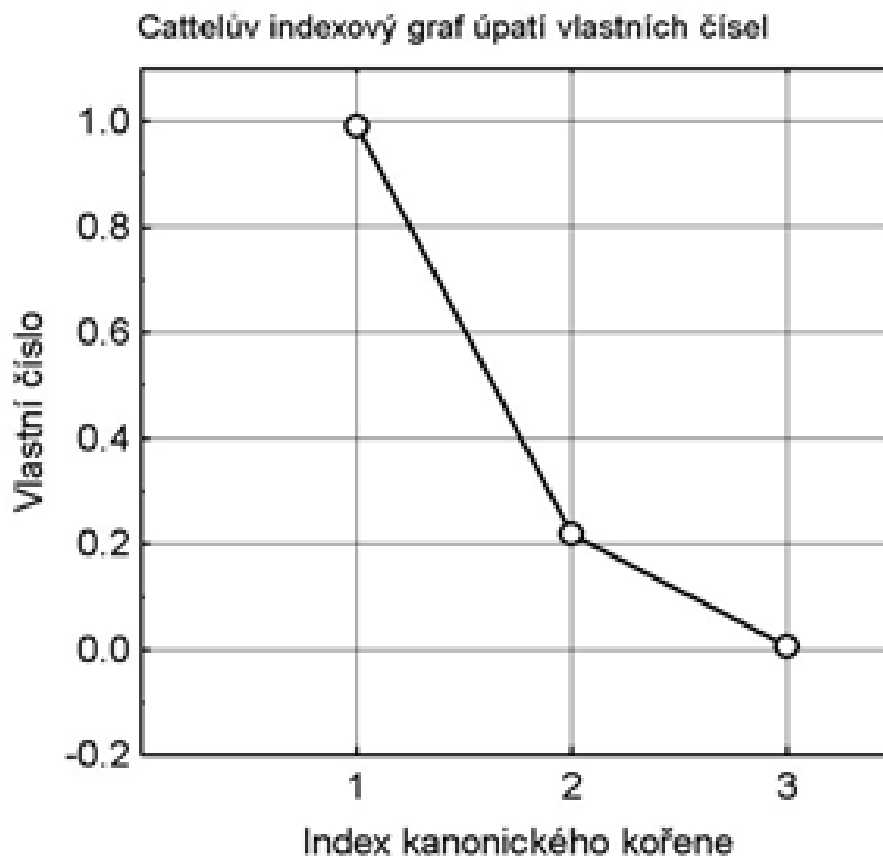
Obě U a V jsou totiž symetrické a stejné, z jejich povahy a obsahu nelze rozlišit.

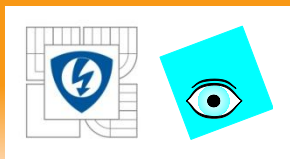
Znaky do obou proměnných U a V lze libovolně zaměnit a obě proměnné jsou zcela rovnocenné.

Kritériem počtu vybraných dvojic U a V je Cattelův indexový diagram úpatí vlastních čísel, který zde potvrzuje, že jeden kanonický kořen bude k popisu dat postačovat.



Cattelův indexový diagram úpatí vlastních čísel při kanonické korelaci dat 10 znaků pro 100 respondentů dat Spokojenost (STATISTICA).



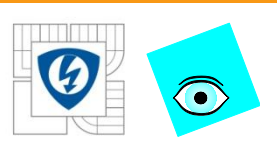


Rozptyl vyextrahovaný

Rozptyl vyextrahovaný značí průměrné množství proměnlivosti, vyextrahovaného ze znaků v obou souborech kanonickými proměnnými.

Kanonické proměnné **extrahují** 100% rozptylu na pravé straně rovnice ze tří znaků *Prace1*, *Prace2*, *Prace3* a 54.13% rozptylu na levé straně rovnice.

	Pravý soubor nezávislých kanonických proměnných, <i>V</i>		Levý soubor závislých kanonických proměnných, <i>U</i>	
Počet znaků	3		7	
Rozptyl extrahovaný	100.000%		54.1271%	
Celková redundance	61.5661%		33.2973%	
Znaky:	1	<i>Prace1</i>		<i>Hoby1</i>
	2	<i>Prace2</i>		<i>Hoby2</i>
	3	<i>Prace3</i>		<i>Doma1</i>
	4			<i>Doma2</i>
	5			<i>Doma3</i>
	6			<i>Ruzne1</i>
	7			<i>Ruzne2</i>

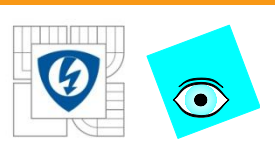


Celková redundance

Celková redundance ukazuje na velikost celkové korelace mezi znaky na pravé straně rovnice 61.6% a na levé straně 33.3%.

Hodnoty **jsou založeny** na všech kanonických kořenech, totiž znacích na levé straně, jež se týkají domén spokojenosti v práci.

Podobně je 33.3% rozptylu nepracovních domén spokojenosti v životě na levé straně rovnice.

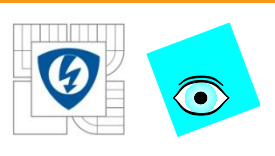


4. Test významnosti kanonických kořenů:

vyšetřuje, zda všechny tři kanonické kořeny U_1 a V_1 , U_2 a V_2 , U_3 a V_3 jsou statisticky významné.

Kanonický korelační koeficient R představuje v datech pouze první kořen U_1 a V_1 čili nejsilnější a nejvýznamnější kanonickou korelaci.

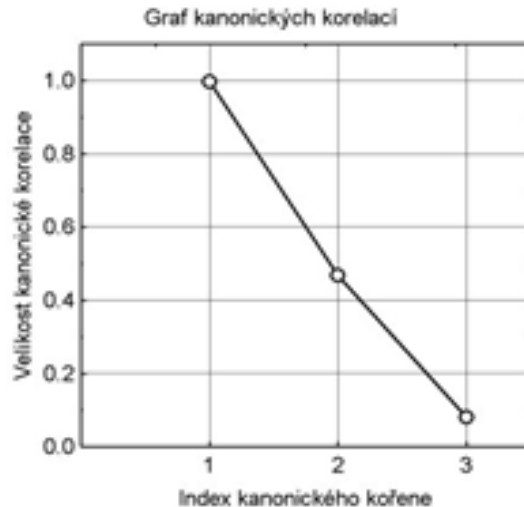
Maximální počet kanonických kořenů, které mohou být z dat extrahovány je roven 3.

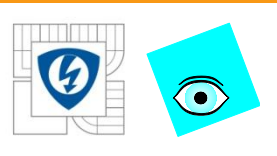


Graf kanonických korelací

Hledání potřebných kanonických kořenů z grafu kanonických korelací pro 10 znaků a 100 respondentů dat *Spokojenost* (STATISTICA).

$R = 0.8847$ je statisticky vysoce významný, protože má vypočtenou hladinu významnosti **P** menší než $\alpha = 0.05$ a může být vysvětlen jako jednoduchá korelace mezi váženou sumou znaků, odpovídajícími první a nejvýznamnější kanonické proměnné **U_1** a **V_1**

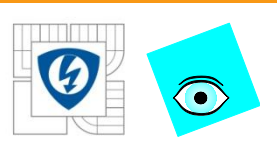




Diskutujme tři řádky testování kanonických kořenů:

1. V prvním řádku nebyl odstraněn žádný kanonický kořen a testy jsou vysoce statisticky významné.
2. V druhém řádku byl první a nejvýznamnější kořen U_1 a V_1 odstraněn a testy už nejsou statisticky významné. Můžeme proto testování ukončit a uzavřít, že statisticky významný je pouze první kanonický kořen U_1 a V_1
3. Když by byl i druhý test statisticky významný, postup by pokračoval dál až do takového řádku, kdy by už statistické testy vycházely statisticky nevýznamné.

Vypuštěný kořen	Kanonické R	Kanonické D	χ^2	df	P	$Lambda$
Žádný	0.885	0.783	153.58	21	0.000	0.193
První U_1 a V_1	0.271	0.073	10.85	12	0.542	0.890
První U_1 a V_1 a druhý U_2 a V_2	0.197	0.039	3.72	5	0.591	0.961



5. Struktura kanonických faktorů a redundance:

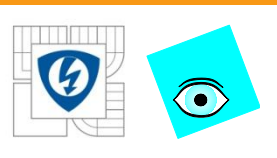
Nadále budeme uvažovat **pouze první kanonický kořen**
 U_1 a V_1

Jak vysvětlíme tento kanonický kořen U_1 a V_1
Jak koreluje tento kanonický kořen U_1 a V_1 se znaky v
obou souborech?

Vyčíslíme proto korelace mezi znaky v každém souboru
s kanonickým kořenem.

Tyto korelace se také nazývají **strukturní koeficienty**.

Nesmíme přitom zapomenout, že kanonická proměnná
je v každém souboru tvořena jako vážená suma znaků.



Faktorová struktura v pravém souboru:

pouze první kanonický kořen je statisticky významný a pouze ten budeme vysvětlovat.

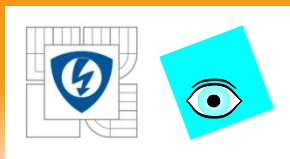
Tři znaky spokojenosti *Prace1*, *Prace2* a *Prace3* vykazují podstatně vyšší zátěže prvního kanonického faktoru čili vysoké korelace právě s tímto faktorem.

Mírou redundancy je průměrné množství rozptylu vyčísleného pro každý znak prvního kanonického kořene.

Sečtou se čtverce vah kanonických faktorů a vydělí se počtem znaků ve výběru, zde to je 3.

Zátěže čili korelace s kanonickou proměnnou V

	Kořen 1, V_1	Kořen 2, V_2	Kořen 3, V_3
<i>Prace1</i>	0.796	-0.575	0.186
<i>Prace2</i>	0.953	0.101	-0.287
<i>Prace3</i>	0.875	0.217	0.432



Tabulka extrahovaného rozptylu na pravé straně:

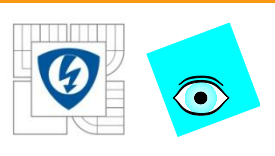
první kanonický kořen U_1 a V_1 extrahuje průměr okolo 77% rozptylu ze tří položek spokojenosti v práci.

Když vynásobíme hodnotu sdíleného rozptylu mezi kanonickými proměnnými v obou souborech tj. R^2 , obdržíme redundanci.

U znaků na levé straně znaků spokojenosti s nepracovními doménami se vyčíslí okolo 60% rozptylu položek spokojenosti v práci, založených na prvním kanonickém kořenu.

Extrahovaný rozptyl a redundance u kanonické proměnné V

	Rozptyl	Redundance
<i>Kořen 1, V_1</i>	0.769	0.602
<i>Kořen 2, V_2</i>	0.129	0.010
<i>Kořen 3, V_3</i>	0.101	0.004



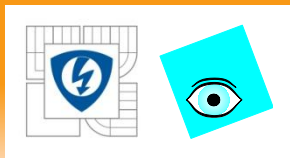
Faktorová struktura v levém souboru.

První kanonický kořen *Kořen 1*, U_1 se vyznačuje vysokými vahami u znaků spokojenosti *Hoby1*, *Hoby2*.

U spokojenosti doma v rodině *Doma1*, *Doma2*, *Doma3* jsou váhy mnohem nižší.

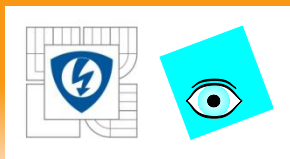
Významné kanonické korelace mezi znaky v obou souborech založené na prvním kanonickém kořenu jsou asi výsledkem vztahu mezi znaky spokojenost v práci, spokojenost se zálibami a koníčky a ostatní spokojeností v životě.

Když budeme uvažovat spokojenost v práci jako nezávisle proměnnou, můžeme říci, že spokojenost v práci ovlivňuje spokojenost v zálibách a koníčcích ale mnohem méně spokojenost doma v rodině.



Zátěže čili korelace s kanonickou proměnnou U

	<i>Kořen 1, U_1</i>	<i>Kořen 2, U_2</i>	<i>Kořen 3, U_3</i>
<i>Hoby1</i>	0.824	-0.100	0.055
<i>Hoby2</i>	0.809	0.278	-0.150
<i>Doma1</i>	0.187	-0.017	0.237
<i>Doma2</i>	0.238	0.253	0.456
<i>Doma3</i>	0.278	0.414	0.173
<i>Ruzne1</i>	0.861	0.049	0.328
<i>Ruzne2</i>	0.821	0.245	0.241



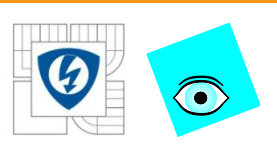
Tabulka extrahovaného rozptylu na levé straně.

První kanonický kořen vyčísluje 42% rozptylu znaků na levé straně a okolo 33% rozptylu ostatních spokojeností, založených na prvním kanonickém kořenu.

Je třeba si všimnout, že tato čísla jsou "zatlačena dolů" relativním nedostatkem korelace mezi kanonickou proměnnou a znaky spokojenosti doma v rodině.

Extrahovaný rozptyl a redundance u kanonické proměnné U

	Rozptyl	Redundance
<i>Kořen 1, U_1</i>	0.417	0.326
<i>Kořen 2, U_2</i>	0.055	0.004
<i>Kořen 3, U_3</i>	0.069	0.003

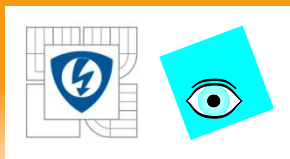


6. Kanonické skóre:

Kanonické proměnné $\mathbf{U}$ a $\mathbf{V}$ představují vážené součty znaků v každém souboru, $\mathbf{U}_1 = a_1 \mathbf{y}_1 + a_2 \mathbf{y}_2 + \dots + a_q \mathbf{y}_q$ a $\mathbf{V}_1 = b_1 \mathbf{x}_1 + b_2 \mathbf{x}_2 + \dots + b_p \mathbf{x}_p$

Váhy a_1, a_2, a_3 a b_1, b_2, b_3 , zobrazené v tabulce kanonických vah odpovídají standardizovaným (Z-transformace) znakům v obou souborech.

Užijeme tyto váhy k výpočtu skóre pro kanonické proměnné.

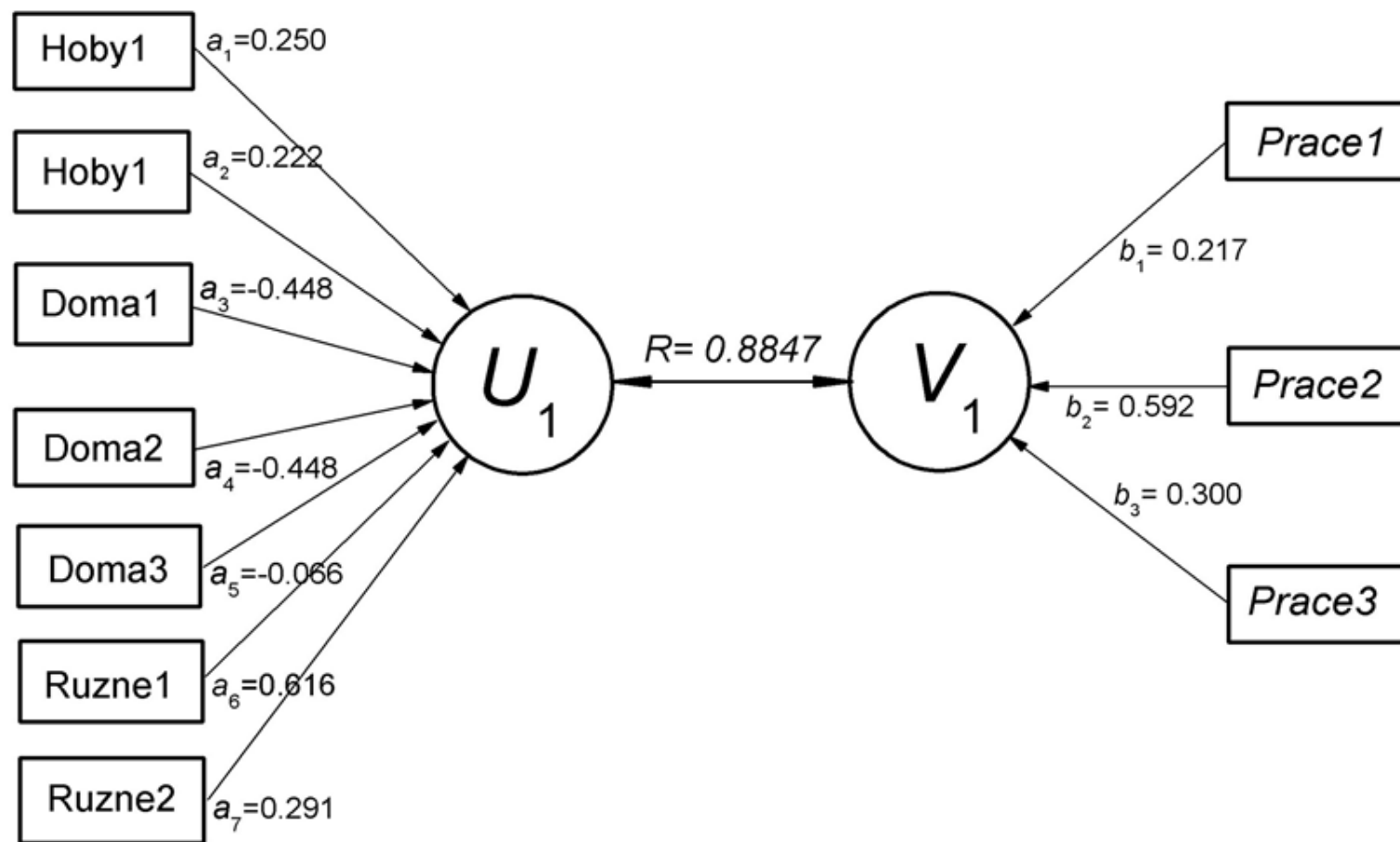
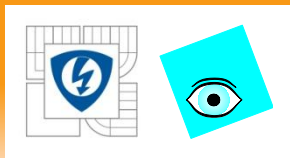


Váhy znaků a_1, a_2, a_3 k výpočtu skóre u kanonické proměnné U

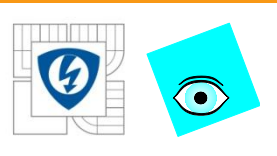
	Kořen 1, U_1	Kořen 2, U_2	Kořen 3, U_3
<i>Hoby1</i>	0.250	-1.636	-1.291
<i>Hoby2</i>	0.222	1.005	-1.096
<i>Doma1</i>	-0.156	-0.434	0.310
<i>Doma2</i>	-0.448	0.326	0.457
<i>Doma3</i>	-0.066	0.700	-0.317
<i>Ruzne1</i>	0.616	-0.361	1.737
<i>Ruzne2</i>	0.291	0.798	0.458

Váhy znaků b_1, b_2, b_3 k výpočtu skóre u kanonické proměnné V

	Kořen 1, V_1	Kořen 2, V_2	Kořen 3, V_3
<i>Prace1</i>	0.217	-1.359	0.244
<i>Prace2</i>	0.592	0.371	-1.388
<i>Prace3</i>	0.300	0.832	1.289



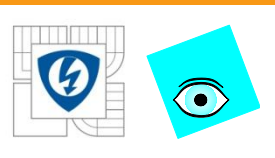
První pár kanonických korelačních proměnných U_1 a V_1 dostatečně popisuje závislost 7 znaků *Hoby1*, *Hoby2*, *Doma1*, *Doma2*, *Ruzne1*, *Ruzne2* na 3 znacích *Prace1*, *Prace2*, *Prace3* zdrojové matice dat *Spokojenost* (STATISTICA, ORIGIN).



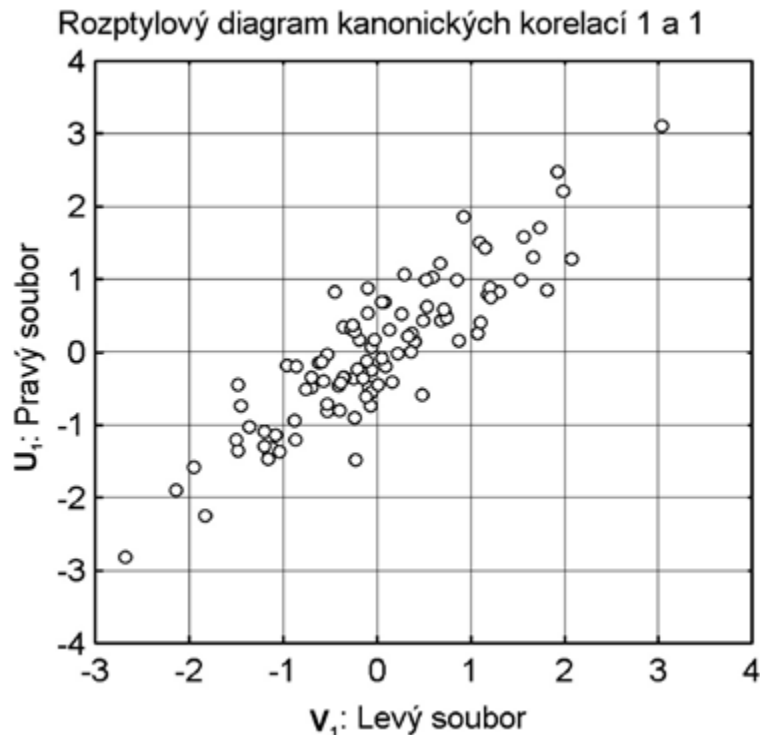
7. Grafy kanonických skóre:

Zobrazíme kanonické skóre znaků na pravé straně čili **V** proti skóřům znaků na levé straně čili **U** a nenacházíme žádné vybočující objekty a ani rezidua po regresi neindikují nelineární trendy.

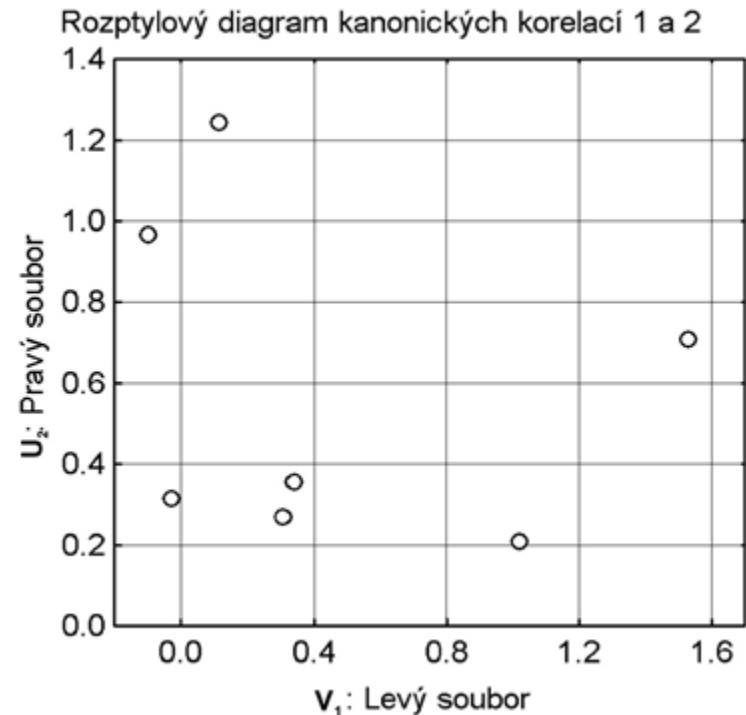
Objekty leží na přímce a nikoliv na křivce tvaru písmen S nebo U.



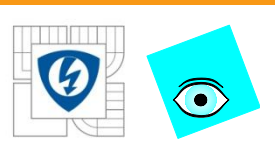
Rozptylové diagramy



Rozptylový diagram kanonických korelací $V_1 = f(U_1)$ pro 10 znaků a 100 respondentů dat *Spokojenost* (STATISTICA).

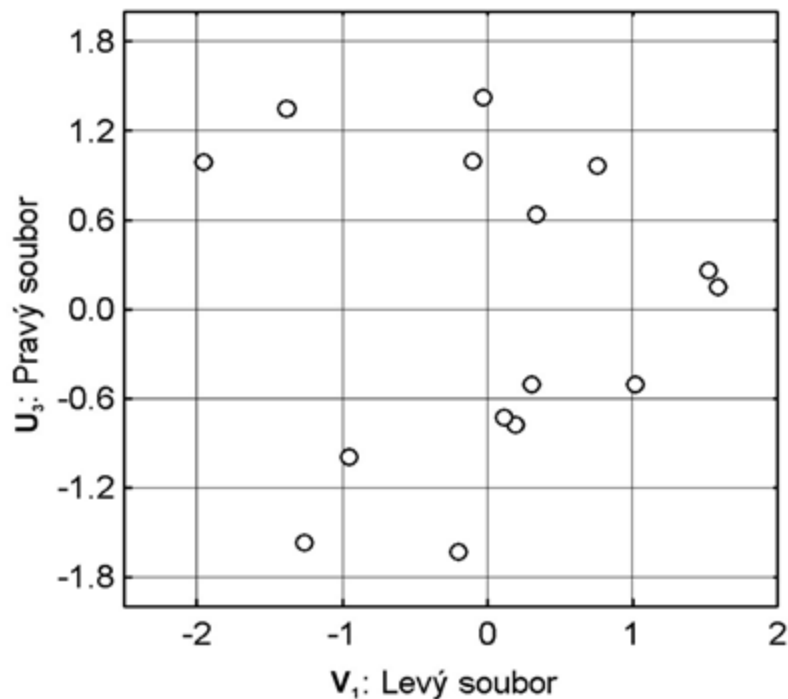


Rozptylový diagram kanonických korelací $U_1 = f(V_1)$ pro 10 znaků a 100 respondentů dat *Spokojenost* (STATISTICA).



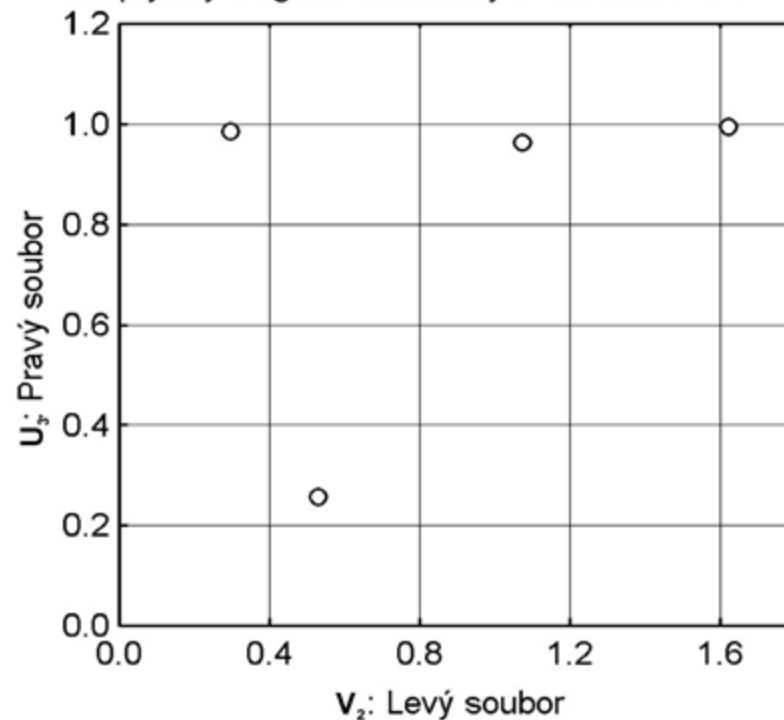
Rozptylové diagramy

Rozptylový diagram kanonických korelací 1 a 3

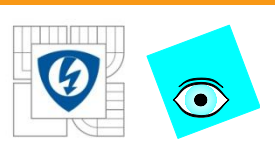


Rozptylový diagram kanonických korelací $V_1 = f(U_3)$ pro 10 znaků a 100 respondentů dat *Spokojenost* (STATISTICA).

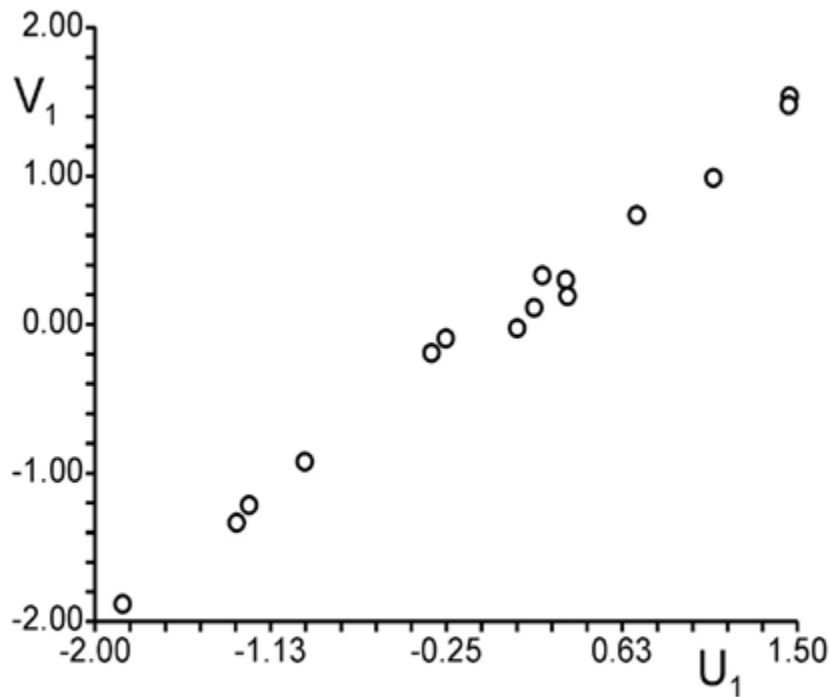
Rozptylový diagram kanonických korelací 2 a 3



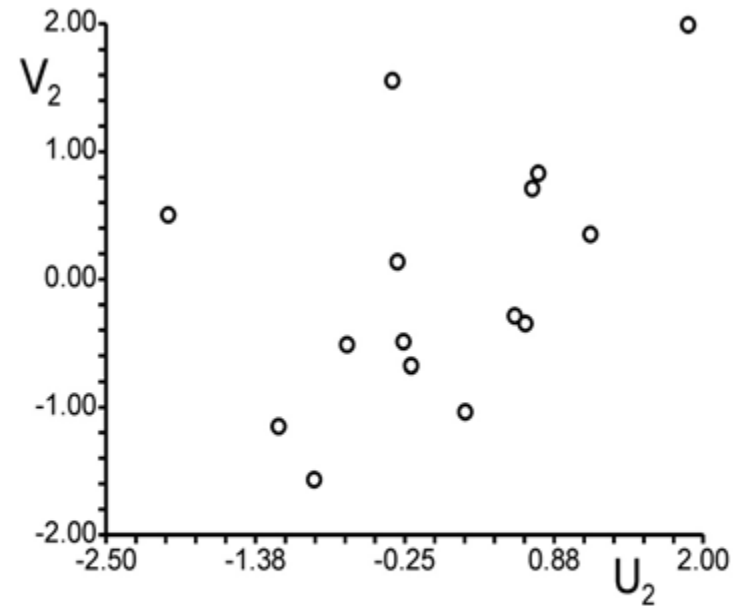
Rozptylový diagram kanonických korelací $U_2 = f(V_3)$ pro 10 znaků a 100 respondentů dat *Spokojenost* (STATISTICA).



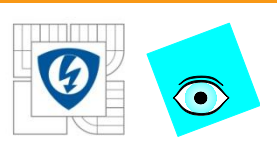
8. Vyšetření shluků objektů: je zde možnost indikovat shluky objektů na obrázku z programu NCSS2002.



Rozptylový diagram kanonických proměnných U_1 a V_1 pro 10 znaků a 100 respondentů dat *Spokojenost* (NCSS2002).



Rozptylový diagram kanonických proměnných U_2 a V_2 pro 10 znaků a 100 respondentů dat *Spokojenost* (NCSS2002).



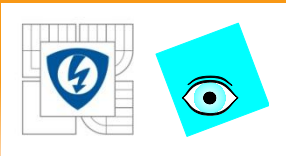
Závěr:

spokojenost v práci významně ovlivňuje spokojenost při zábavě a i celkově.

Spokojenost doma v rodině se nejeví příliš ovlivněna spokojeností v práci.

V praxi by měla být tato dotazníková úloha s různými skupinami respondentů opětovně vyhodnocena touto metodou.

Tím by se potvrdilo, zda kanonická faktorová struktura vede k věrohodné a spolehlivé interpretaci prvního kanonického kořene, tzv. *replikovatelnosti*.



CVIČENÍ V PROGRAMU STATISTICA

24.2.2010

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



Soubor obsahuje proměnné založené na 2 faktorech										
	Zaměstnání	Zaměstnání	Zaměstnání	4 Hobby 1	5 Hobby 2	6 Bydlení 1	7 Bydlení 2	8 Bydlení 3	9 Různé 1	10 Různé 2
1	105,126	101,659	115,060	100,998	95,184	100,281	101,667	85,553	104,035	110,278
2	77,049	72,933	77,485	72,744	61,563	93,854	95,392	88,609	70,115	72,000
3	86,017	82,206	78,889	77,951	91,705	86,773	108,070	93,348	86,021	70,688
4	91,425	106,107	95,640	90,901	111,466	100,248	86,080	93,822	101,224	82,665
5	113,714	92,029	99,079	79,277	98,416	104,013	83,271	69,621	82,820	70,022
6	86,806	87,817	67,663	93,662	77,997	99,822	97,275	108,622	91,400	79,776
7	95,067	94,505	98,081	94,513	97,422	93,694	99,181	96,398	90,732	86,707
8	113,500	104,607	105,572	101,008	102,275	87,427	96,664	86,577	93,057	112,702
9	104,549	97,299	94,074	88,538	98,112	97,785	99,585	99,761	99,399	105,908
10	104,635	97,908	85,823	82,486	90,447	104,688	95,076	99,695	77,630	62,031
11	102,064	87,010	94,687	79,203	68,482	78,995	86,430	78,822	89,729	87,553
12	109,428	94,937	104,396	119,293	112,988	122,931	114,816	102,109	116,339	101,363
13	89,994	77,392	100,771	97,026	111,107	107,660	116,858	114,730	88,715	76,495
14	89,983	91,259	89,216	81,250	100,519	94,902	96,723	94,411	88,419	89,670
15	96,946	103,043	97,493	98,515	96,833	92,778	110,072	95,162	102,368	105,517
16	93,210	93,843	87,345	104,256	115,323	114,286	111,617	111,455	95,701	105,724
17	70,905	65,903	74,231	60,108	47,088	79,806	94,799	88,352	54,658	58,235
18	114,854	98,083	105,634	90,237	101,011	98,951	87,921	88,913	91,297	92,224
19	100,315	92,518	87,609	103,273	88,647	108,401	103,565	82,486	102,326	97,573
20	111,355	104,373	101,852	108,083	121,884	99,954	96,688	89,369	102,148	114,451
21	108,798	97,876	91,457	99,552	110,486	104,375	95,028	85,508	94,534	91,802
22	74,723	105,615	102,246	90,169	93,522	86,003	94,308	106,233	85,310	87,673
23	93,878	90,580	95,486	86,063	92,444	107,834	97,738	105,225	87,353	89,168
24	77,936	94,987	98,586	100,325	100,906	96,458	95,406	95,111	104,175	115,847
25	77,515	100,672	102,301	99,828	106,570	109,101	116,680	105,655	106,498	107,616
26	99,123	110,273	120,531	115,784	118,914	109,747	108,690	119,019	112,253	120,002
27	92,469	94,971	94,383	101,190	95,078	105,823	102,772	116,442	98,448	91,185
28	92,767	107,075	94,019	111,294	146,450	111,862	112,545	119,948	121,325	113,347
29	76,859	83,319	67,787	54,604	66,491	80,830	76,364	81,024	48,978	50,056
30	87,305	85,485	82,823	90,371	89,675	98,095	108,588	123,270	88,390	92,863
31	90,115	98,804	93,923	102,658	117,803	113,116	100,333	108,156	100,138	99,204
32	107,975	105,773	114,251	132,672	152,624	121,607	132,884	128,478	145,244	147,192
33	118,087	96,854	93,445	96,845	74,378	95,036	104,447	106,562	105,696	94,970
34	79,567	91,437	92,243	100,236	91,127	117,613	112,114	108,547	102,015	107,751
35	94,406	97,257	95,566	91,766	98,602	85,128	101,343	93,761	93,623	85,325
36	121,235	114,295	113,353	118,700	122,608	112,353	110,939	123,047	117,641	123,765
37	102,719	111,096	121,558	83,284	107,952	84,165	76,385	80,054	93,209	87,839
38	91,074	105,772	101,445	103,061	123,072	116,930	112,745	106,480	113,189	105,960
39	94,635	81,728	96,747	102,308	95,434	99,814	111,435	110,644	108,013	93,038
40	114,468	115,632	115,656	121,383	127,679	106,550	113,191	111,581	127,931	128,899
41	76,794	83,167	95,768	69,323	76,995	85,357	101,615	100,099	80,356	80,744
42	69,571	99,519	107,377	100,471	109,374	109,876	105,846	105,623	100,425	99,726

1. Zdrojová matice *Factor.sta* se načte z databáze *EXAMPLES* programu STATISTICA a obsahuje 100 řádků sledovaných případů respondentů a 10 sloupců faktorů a prediktorů.

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Okno Nápořěda

Obnovit... Ctrl+R

Analýza skupin

Základní statistiky/tabulky

Vícerozměrná regrese

ANOVA

Neparametrická statistika

Prokládání rozdělení

Pokročilé lineární/nelineární modely

Vícerozměrné průzkumné techniky

Průmyslová statistika & Six Sigma

Analýza síly testu

Neuronové sítě

Vytěžování dat

QC Data mining & Analýza hlavních příčin

Text & Document Mining, Web Crawling

Statistiky bloku dat

STATISTICA Visual Basic

Pravděpodobnostní kalkulátor

Shluková analýza

Faktorová analýza

Hlavní komponenty & klasifikační analýza

Kanonická analýza

Analýza spolehlivosti/prvků

Klasifikační stromy

Korespondenční analýza

Vícerozměrné škálování

Diskriminační analýza

Modely obecné diskriminační analýzy

Soubor obsahuje proměnné založené

	Zaměstnaní	Zaměstnaní	Zaměstnaní						
	1	2	3						
1	105,126	101,659	115,06						
2	77,049	72,933	77,48						
3	86,017	82,206	78,86						
4	91,425	106,107	95,64						
5	113,714	92,029	99,07						
6	86,606	87,817	67,66						
7	95,067	94,505	98,06						
8	113,500	104,607	105,57						
9	104,549	97,299	94,07						
10	104,635	97,908	85,82						
11	102,064	87,010	94,66						
12	109,428	94,937	104,39						
13	89,994	77,392	100,77						
14	89,983	91,259	89,21						
15	96,946	103,043	97,49						
16	93,210	93,843	87,34						
17	70,905	65,903	74,23						
18	114,854	98,083	105,634	90,237	101,011	98,951			
19	100,315	92,518	87,609	103,273	88,647	108,401			
20	111,355	104,373	101,852	108,083	121,884	99,954			
21	108,798	97,876	91,457	99,552	110,486	104,375			
22	74,723	105,615	102,246	90,169	93,522	86,003			
23	93,878	90,580	95,486	86,063	92,444	107,834			
24	77,936	94,987	98,586	100,325	100,906	96,458			
25	77,515	100,672	102,301	99,828	106,570	109,101			
26	99,123	110,273	120,531	115,784	118,914	109,747			
27	88,488	84,874	84,888	104,488	85,878	105,888			
28									
29									
30									
31									
32									
33									
34									
35									
36									
37									
38									
39									
40									
41									
42									
43									
44									
45									
46									
47									
48									
49									
50									
51									
52									
53									
54									
55									
56									
57									
58									
59									
60									
61									
62									
63									
64									
65									
66									
67									
68									
69									
70									
71									
72									
73									
74									
75									
76									
77									
78									
79									
80									
81									
82									
83									
84									
85									
86									
87									
88									
89									
90									
91									
92									
93									
94									
95									
96									
97									
98									
99									
100									

7 Bydlení 2 8 Bydlení 3 9 Různé 1 10 Různé 2

101,667 85,553 104,035 110,278

95,392 88,609 70,115 72,000

108,070 93,348 86,021 70,688

86,080 93,822 101,224 82,665

70,022

79,776

86,707

112,702

05,908

62,031

87,553

01,363

76,495

89,670

05,517

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,701 105,724

94,799 88,352 54,658 58,235

87,921 88,913 91,297 92,224

103,565 82,486 102,326 97,573

96,688 89,369 102,148 114,451

95,028 85,508 94,534 91,802

94,308 106,233 85,310 87,673

97,738 105,225 87,353 89,168

95,406 95,111 104,175 115,847

116,680 105,655 106,498 107,616

108,690 119,019 112,253 120,002

111,817 111,455 95,70

99,822	97,275	108,622	91,400	79,776
93,694	99,181	96,398	90,732	86,707
87,427	96,664	86,577	93,057	112,702
97,785	99,585	99,761	99,399	105,908
104,688	95,076	99,695	77,630	62,031
78,995	86,430	78,822	89,729	87,553
122,931	114,816	102,109	116,339	101,363
107,600	116,850	114,720	99,715	76,405

Kanonická analýza: Factor.sta [?] [-] [X]

Základní nastavení

 **Proměnné:** žádné

Vstupní soubor: Zdrojová data

☐ Přehled popisných statistik a korelační matice

 **OK**

 **Storno**

 **Možnosti** ▾

 **Otevři Data**

SELECT CASES  

ChD vynechána

☒ Celé případy

☐ Nahradit průměrem

Zvolte proměnné pro kanonickou analýzu [?] [X]

1-Zaměstnání 1
2-Zaměstnání 2
3-Zaměstnání 3
4-Hobby 1
5-Hobby 2
6-Bydlení 1
7-Bydlení 2
8-Bydlení 3
9-Různé 1
10-Různé 2

OK

Storno

Pro zobrazení odpovídajících proměnných zvolte "Ukázat pouze...". Pro více informací stiskněte F1.

Vybrat vše **DI. názvy** **Detaily**

Vyberte proměnné:

1-10

☐ Ukázat pouze odpovídající proměnné

3. Kliknutím na *Proměnné* se otevře okno *Zvolte proměnné pro kanonickou analýzu*, které naplníme 10 znaky a klikneme OK.

84,165	76,385	80,054	93,209	87,839
116,930	112,745	106,480	113,189	105,960
99,814	111,435	110,644	108,013	93,038

82,486	90,447	104,688	95,076	99,695	77,630	62,031
79,203	68,482	78,995	86,430	78,822	89,729	87,553
119,293	112,988	122,931	114,816	102,109	116,339	101,363
97,026	111,107	107,660	116,850	114,720	88,715	76,405
81,250	100,519					
98,515	96,833					
104,256	115,323					
60,108	47,088					
90,237	101,011					
103,273	88,647					
108,083	121,884					
99,552	110,486					
90,169	93,522					
86,063	92,444					
100,325	100,906					
99,828	106,570					
115,784	118,914					
101,190	95,078					
111,294	146,450					
54,604	66,491	66,636	76,364	61,624	46,376	36,636
90,274	88,675	88,685	100,500	122,370	88,200	82,660

Kanonická analýza: Factor.sta [?] [-] [X]

Základní nastavení

 Proměnné: **Vše**

Vstupní soubor: **Zdrojová data** ▼

☒ Přehled popisných statistik a korelační matice

OK

Storno

 Možnosti ▼

 Otevři Data

SELECT CASES  

ChD vynechána

☒ Celé případy

☐ Nahradit průměrem

10. 4. Záložku *Přehled popisných statistik a korelační matice* zaškrtneme
13. a OK.
91

100,236	91,127	117,613	112,114	106,547	102,015	107,751
91,766	98,602	85,128	101,343	93,761	93,623	85,325
118,700	122,608	112,353	110,939	123,047	117,641	123,765
83,284	107,952	84,165	76,385	80,054	93,209	87,839
103,061	123,072	116,930	112,745	106,480	113,189	105,960
102,308	95,434	99,814	111,435	110,644	108,013	93,038
121,383	127,679	106,550	113,191	111,581	127,931	128,899
69,323	76,995	85,357	101,615	100,099	80,356	80,744
100,471	109,374	109,876	105,846	105,623	100,425	99,726

2	101,008	102,275	87,427	96,664	86,577	93,057	112,702
4	88,538	98,112	97,785	99,585	99,761	99,399	105,908
3	82,486	90,447	104,688	95,076	99,695	77,630	62,031
7	79,203	68,482	78,995	86,430	78,822	89,729	87,553
6	119,293	112,988	122,931	114,816	102,109	116,339	101,363
1	97,026	111,107	107,660	116,858	114,730	88,715	76,495
6	81,250	100,519					
3	98,515	96,830					
5	104,256	115,320					
1	60,108	47,088					
4	90,237	101,011					
9	103,273	88,640					
2	108,083	121,880					
7	99,552	110,480					
6	90,169	93,520					
6	86,063	92,440					
6	100,325	100,900					
1	99,828	106,570					
1	115,784	118,910					
3	101,190	95,070					
9	111,294	146,450					
7	54,604	66,491	80,830	76,364	81,024	48,978	50,056

Přehled popisných statistik: Factor.sta ? _ X

ChD vynechána párově

100 příp. zvoleny, 0 s chybějíc. daty
100 platné příp. akcept.

Zákl. výsledky | Detaily | Matice

Průměry & sm. odchylky

Korelace

OK

Storno

Možnosti ▼

5. Otevře se okno *Přehled popisných statistik*, ve kterém je uvedeno, že řádky chybějících dat v dírách byly vynechány. Okno má tři záložky: *Zákl. výsledky*, *Detaily* a *Matice*.

6	91,766	98,602	85,128	101,343	93,761	93,623	85,325
3	118,700	122,608	112,353	110,939	123,047	117,641	123,765
3	83,284	107,952	84,165	76,385	80,054	93,209	87,839
5	103,061	123,072	116,930	112,745	106,480	113,189	105,960
7	102,308	95,434	99,814	111,435	110,644	108,013	93,038
3	124,232	127,272	122,552	112,124	114,524	127,224	122,222

	Průměry a směrodatné odchylky (Factor.sta)						
proměnná	Průměry	Sm.odch.					
Zaměstnání 1	97,0296	15,52089					
Zaměstnání 2	98,1689	11,26746					
Zaměstnání 3	98,9430	12,48538					
Hobby 1	98,0252	15,93673					
Hobby 2	100,1081	19,94109					
Bydlení 1	99,5083	11,97813					
Bydlení 2	101,6029	11,06155					
Bydlení 3	101,3717	12,73475					
Různé 1	99,1629	16,96749					
Různé 2	98,2238	19,11151					

6. Okno dialogu *Popisných statistik a korelace* začíná tabulkou *Průměry a směrodatné odchylky* všech 10 znaků.

[illegible]

The screenshot shows the STATISTICA software interface. The title bar reads 'STATISTICA Cz - [PS 1* - Korelace (Factor.sta)]'. The menu bar includes 'Soubor', 'Úpravy', 'Zobrazit', 'Vložit', 'Formát', 'Statistika', 'Grafy', 'Nástroje', 'Data', 'Pracovní sešit', 'Okno', and 'Nápověda'. The toolbar contains various icons for file operations, editing, and statistical analysis. The left sidebar shows a project tree with 'PS 1*' expanded, containing 'Kanonická analýza (Factor.sta)' and 'Dialog popisných statistik', which in turn contains 'Průměry a směrodatné odchylky' and 'Korelace (Factor.sta)'. The main window displays a correlation matrix titled 'Korelace (Factor.sta)'. The matrix has 11 rows and 11 columns, with the first column labeled 'proměnná' and the first row labeled 'Zaměstnaní 1'. The data is as follows:

proměnná	Zaměstnaní 1	Zaměstnaní 2	Zaměstnaní 3	Hobby 1	Hobby 2	Bydlení 1	Bydlení 2	Bydlení 3	Různé 1	Různé 2
Zaměstnaní 1	1,000	0,647	0,653	0,598	0,521	0,143	0,145	0,138	0,611	0,549
Zaměstnaní 2	0,647	1,000	0,732	0,689	0,698	0,143	0,182	0,236	0,709	0,685
Zaměstnaní 3	0,653	0,732	1,000	0,637	0,630	0,164	0,238	0,255	0,698	0,671
Hobby 1	0,598	0,689	0,637	1,000	0,805	0,536	0,634	0,583	0,904	0,843
Hobby 2	0,521	0,698	0,630	0,805	1,000	0,506	0,496	0,482	0,811	0,756
Bydlení 1	0,143	0,143	0,164	0,536	0,506	1,000	0,658	0,590	0,498	0,425
Bydlení 2	0,145	0,182	0,238	0,634	0,496	0,658	1,000	0,731	0,644	0,593
Bydlení 3	0,138	0,236	0,255	0,583	0,482	0,590	0,731	1,000	0,586	0,518
Různé 1	0,611	0,709	0,698	0,904	0,811	0,498	0,644	0,586	1,000	0,841
Různé 2	0,549	0,685	0,671	0,843	0,756	0,425	0,593	0,518	0,841	1,000

7. Kanonická korelace je založena na korelační matici znaků. Prvním krokem proto bude vyčíslení *Korelační matice*, pokud ovšem korelační matice nebyla načtena ve vstupních datech jako jedna možnost inputu. Všimněte si, že později budete zadávat proměnné pro dva výběry z těchto znaků, které jsme zde specifikovali.

8. Kliknutím na okénko *Kovariance* se otevře kovarianční matice.

nce (Factor.s	Zaměstnání 1	240,898	113,218	126,467	102,961	123,640	156,781	19,349	22,665	33,860	135,476	147,462
vý graf	Zaměstnání 2	113,218	126,956	102,961	123,640	156,781	19,349	22,665	33,860	135,476	147,462	162,832
	Zaměstnání 3	126,467	102,961	155,885	126,737	156,860	24,471	32,911	40,488	147,846	160,023	160,023
	Hobby 1	147,947	123,640	126,737	253,979	255,728	102,391	111,819	118,288	244,574	256,818	256,818
	Hobby 2	161,282	156,781	156,860	255,728	397,647	120,840	109,388	122,491	274,404	288,026	288,026
	Bydlení 1	26,551	19,349	24,471	102,391	120,840	143,476	87,149	90,002	101,298	97,226	97,226
	Bydlení 2	24,918	22,665	32,911	111,819	109,388	87,149	122,358	102,923	120,802	125,446	125,446
	Bydlení 3	27,235	33,860	40,488	118,288	122,491	90,002	102,923	162,174	126,592	126,010	126,010
	Různé 1	160,998	135,476	147,846	244,574	274,404	101,298	120,802	126,592	287,896	272,852	272,852
	Různé 2	162,832	147,462	160,023	256,818	288,026	97,226	125,446	126,010	272,852	365,250	365,250

Přehled popisných statistik: Factor.sta

ChD vynechána párově

100 příp. zvoleny, 0 s chybějíc. daty
100 platné příp. akcept.

Zákl. výsledky Details Matice

Průměry & sm. odchylky Krabicový graf proměnných
Korelace Maticový graf korelací
Kovariance ☐ SD=Součet čtverců/N

OK Storno Možnosti

Vyberte proměnné na krabicový graf

1-Zaměstnání 1
2-Zaměstnání 2
3-Zaměstnání 3
4-Hobby 1
5-Hobby 2
6-Bydlení 1
7-Bydlení 2
8-Bydlení 3
9-Různé 1
10-Různé 2

OK Storno

Vybrat vše Dl. názvy Details

Vyberte proměnné:
1-10

☐ Ukázat pouze odpovídající proměnné

9. Kliknutím na *Krabicový graf proměnných* se otevře další okno, nadepsané *Vyberte proměnné na krabicový graf*, ve kterém zadáme znaky k této EDA analýze a OK.

Hobby 1	120,840	120,840	120,840	120,840	120,840	120,840	120,840	120,840	120,840	120,840	120,840
Hobby 2	161,282	156,781	156,860	255,728	397,647	120,840	109,388	122,491	274,404	288,0	288,0
Bydlení 1	26,551	19,349	24,471	102,391	120,840	143,476	87,149	90,002	101,298	97,2	97,2
Bydlení 2	24,918	22,665	32,911	111,819	109,388	87,149	122,358	102,923	120,802	125,4	125,4
Bydlení 3	27,235	33,860	40,488	118,288	122,491	90,002	102,923	162,174	126,592	126,0	126,0
Různé 1	160,998	135,476	147,846	244,574	274,404	101,298	120,802	126,592	287,896	272,8	272,8
Různé 2	162,832	147,462	160,023	256,818	288,026	97,226	125,446	126,010	272,852	365,2	365,2

Přehled popisných statistik: Factor.sta [?] _ X

ChD vynechána párově

100 příp. zvoleny, 0 s chybějícími daty
100 platné příp. akcept.

Zákl. výsledky Details Matice

Průměry & sm. odchylky Krabicový graf proměnných
Korelace Maticový graf korelací
Kovariance ☐ SD=Součet čtverců/N

OK Storno Možnosti ▾

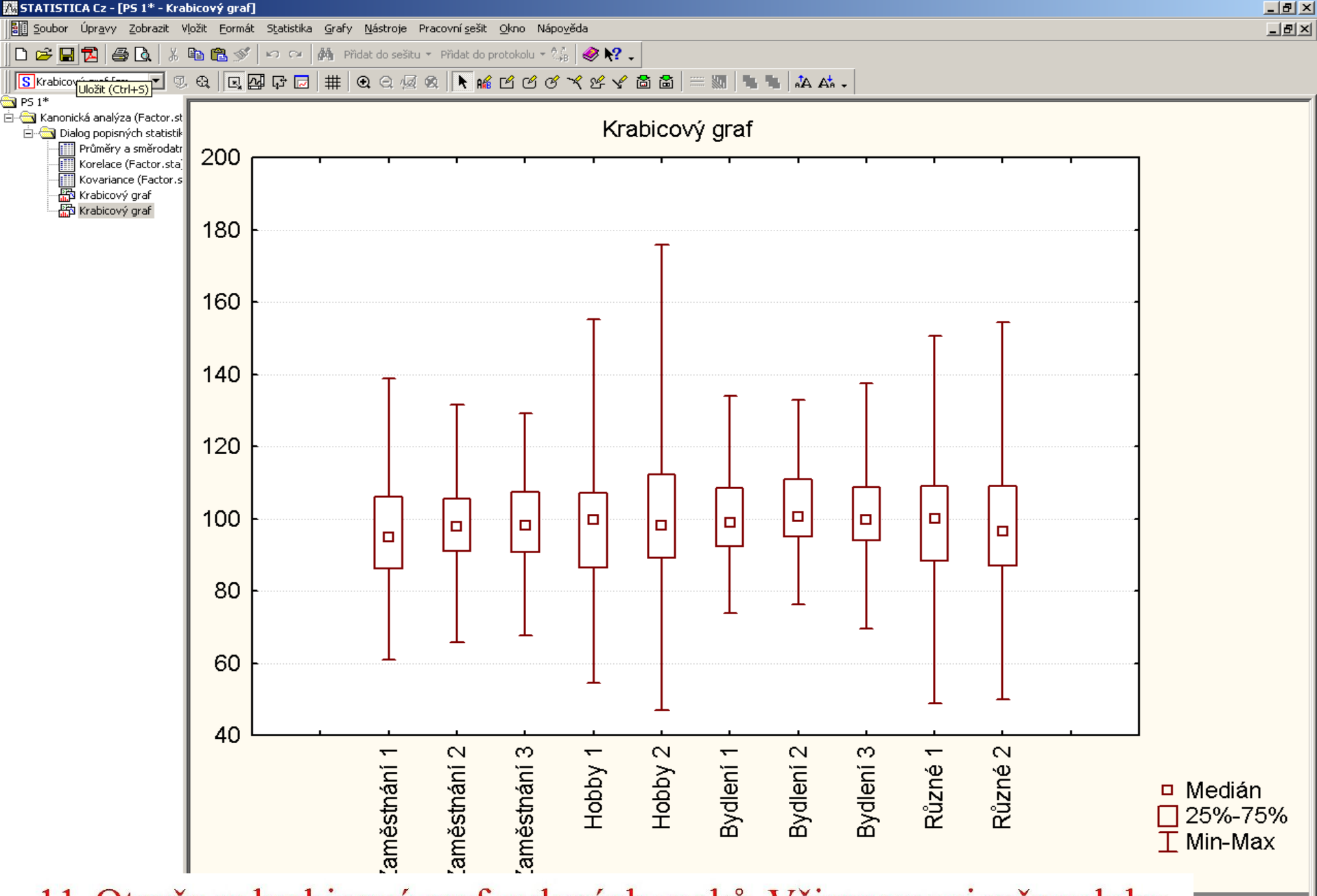
Typ krabicového grafu: [?] X

☒ Medián/Kvartily/Rozpětí
☐ Průměr/SmCh/SmOdch
☐ Průměr/SmOdch/1.96*SmOdch
☐ Průměr/SmCh/1.96*SmCh

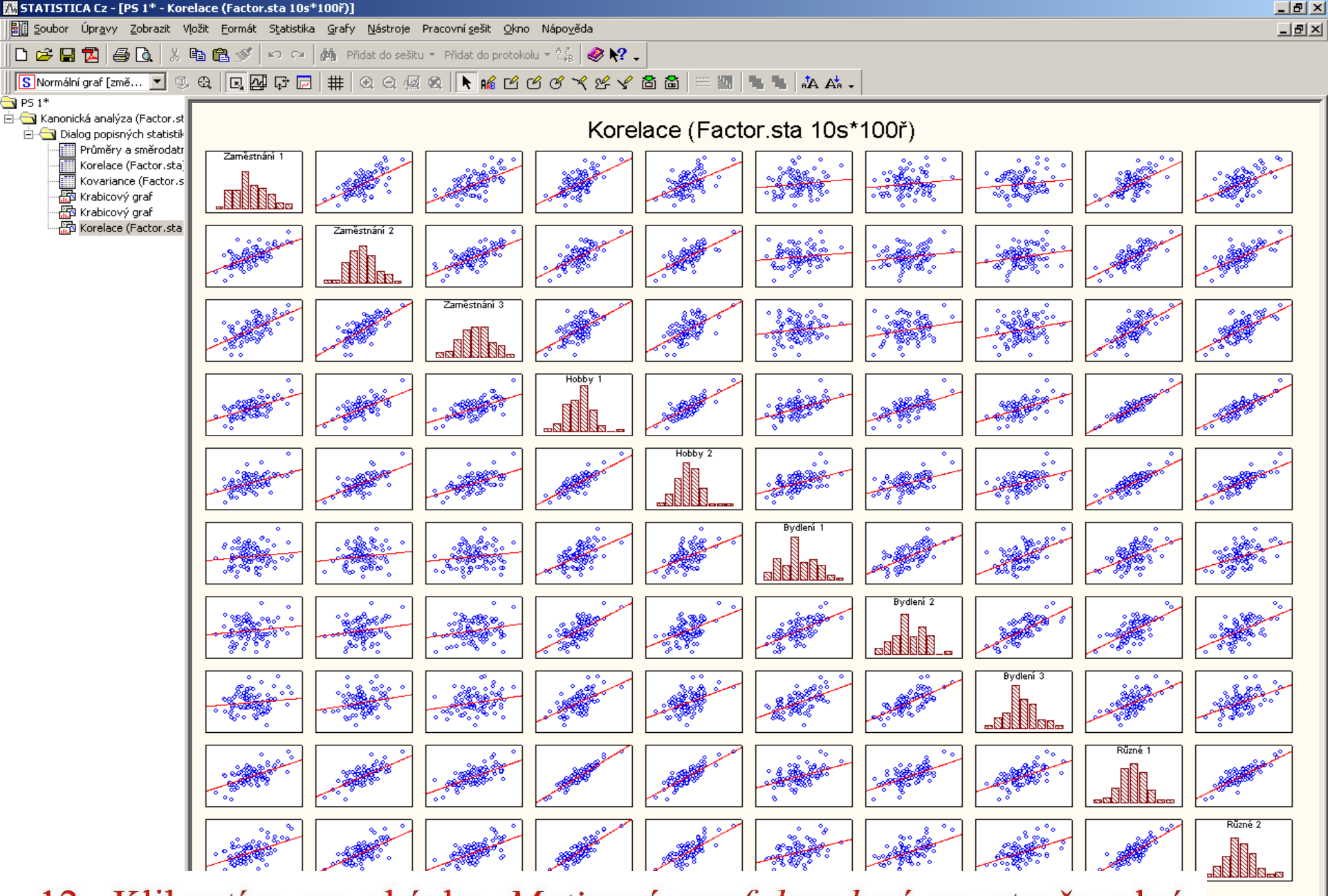
OK Storno

Ostatní typy krabicových grafů a grafů odlehklých a extrémních hodnot najdete v nabídce Grafy.

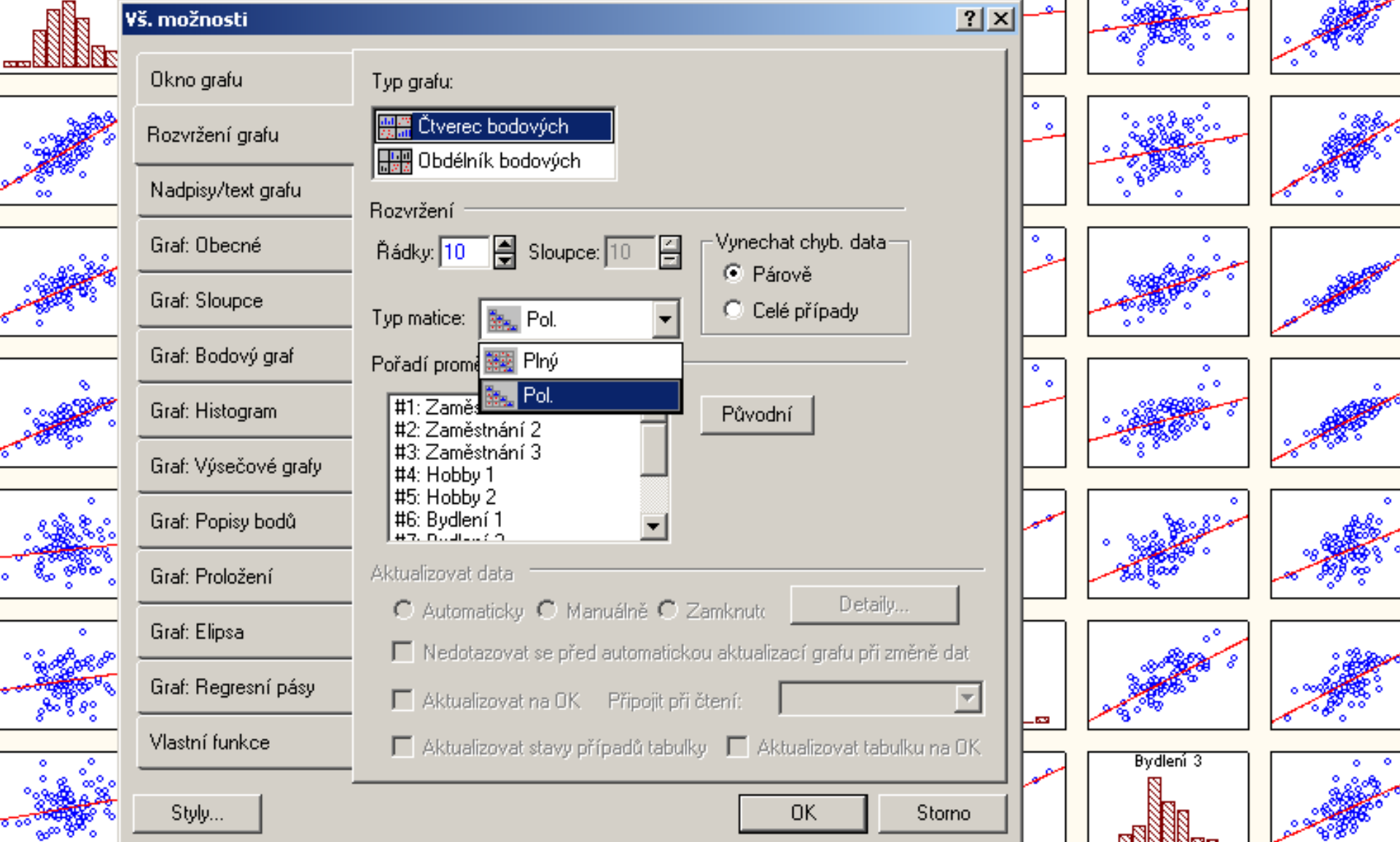
10. Otevře se okno *Typ krabicového grafu*, ve kterém zaškrtneme jeden ze čtyř nabídnutých typů krabice.



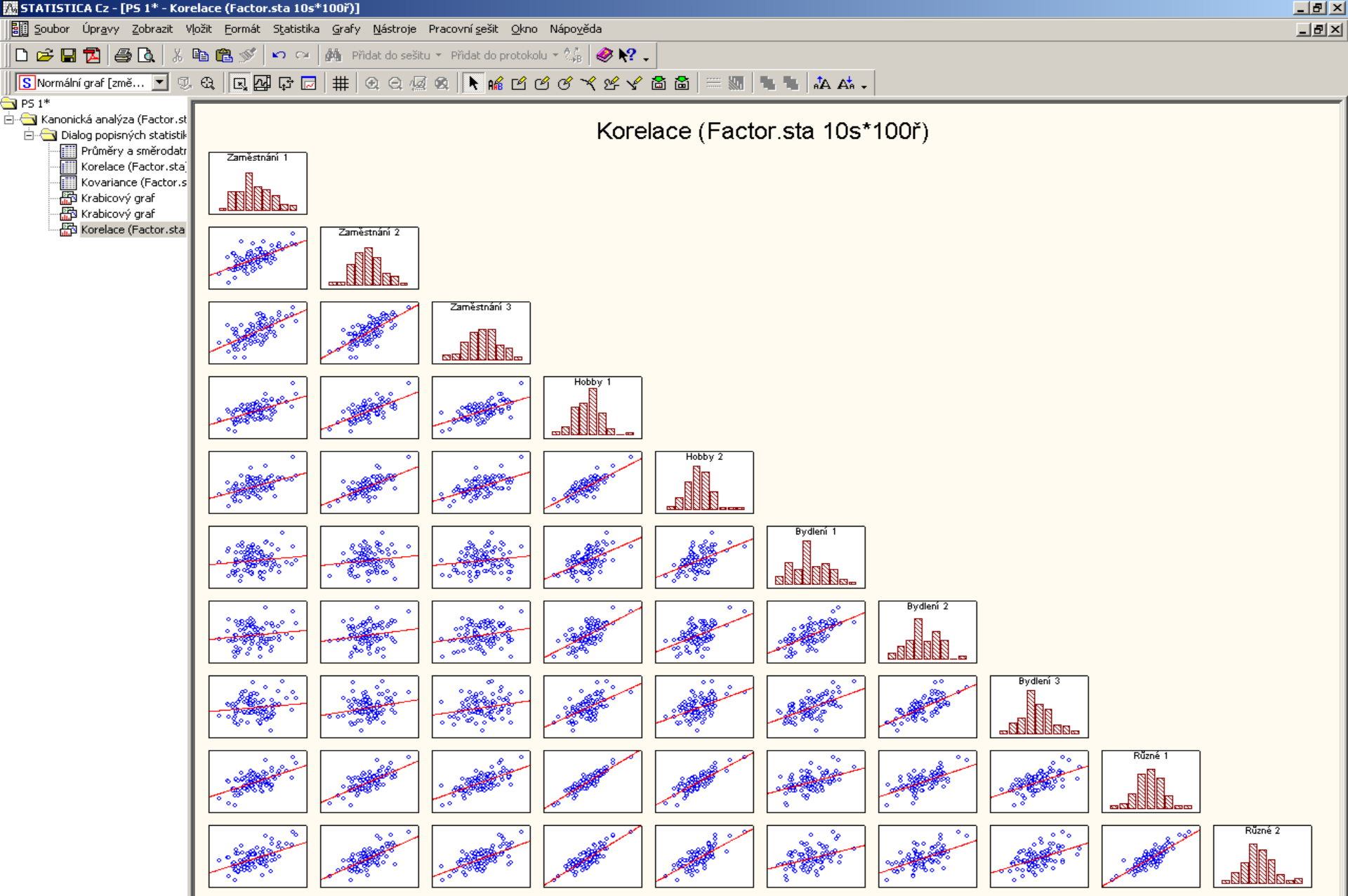
11. Otevře se krabicový graf zadaných znaků. Všimneme si měř polohy, rozptýlení a tvaru. Vně vousů jsou indikovány vybočující body.



12. Kliknutím na okénko *Maticový graf korelací* se otevře plný korelační diagram.



13. Kliknutím kamkoli na ploše se otevře okno *Všeobecné možnosti* a v záložce *Rozvržení grafu* klikneme v *Typu matice* na volbu *Pol.*



14. Zobrazí se dolní polovina maticového diagramu korelace znaků a na diagonále jsou histogramy proměnlivosti znaků.

Hobby 2	0,52110	0,69778	0,63003	0,80469	1,00000	0,50591	0,49591	0,48235	0,81101
Bydlení 1	0,14282	0,14337	0,16363	0,53638	0,50591	1,00000	0,65774	0,59003	0,49842
Bydlení 2	0,14514	0,18185	0,23830	0,63431	0,49591	0,65774	1,00000	0,73065	0,64364
Bydlení 3	0,13779	0,23598	0,25465	0,58284	0,48235	0,59003	0,73065	1,00000	0,58587
Různé 1	0,61135	0,70863	0,69789	0,90447	0,81101	0,49842	0,64364	0,58587	1,00000
Různé 2	0,54894	0,68479	0,67063	0,84320	0,75577	0,42471	0,59340	0,51775	0,84101
Průměry	97,02958	98,16886	98,94298	98,02522	100,10813	99,50832	101,60288	101,37174	99,16212
Sm.odch	15,52089	11,26746	12,48538	15,93673	19,94109	11,97813	11,06155	12,73475	16,96712

Přehled popisných statistik: Factor.sta

ChD vynechána párově

100 příp. zvoleny, 0 s chybějíc. daty

100 platné příp. akcept.

Zákl. výsledky Details Matice

Matice

OK

Storno

Možnosti

15. Okno *Přehled popisných statistik* přináší v záložce *Matice* v číselné podobě korelační matici.

STATISTICA Cz - [Data: Tabulka4 (10s krát 14ř)]

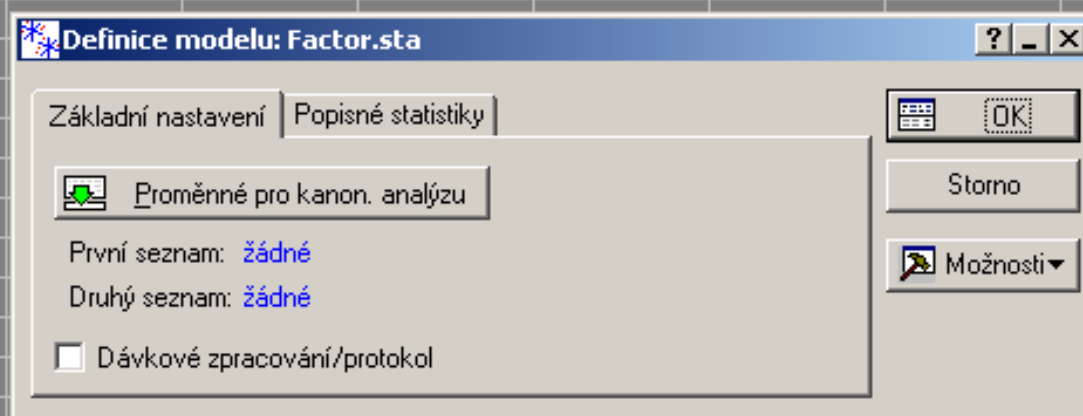
Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Okno Nápořěda

Arial 10 **B** *I* U [Text alignment icons] [Color fill icon] [Color border icon] [Color text icon] [Number format icons] [Language icons] [Proměnné] [Případy]

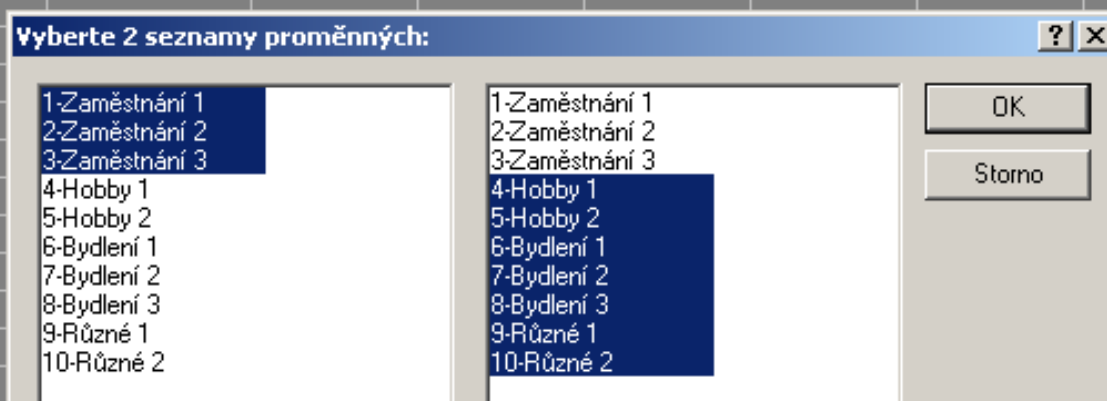
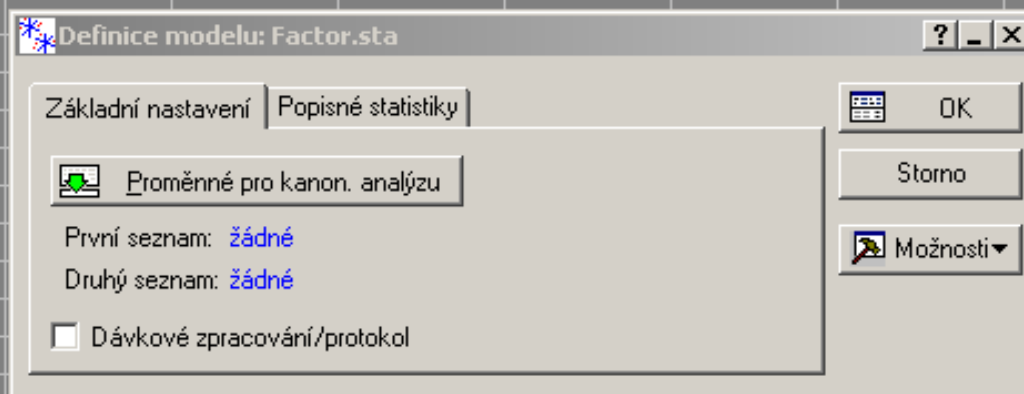
	Factor.sta									
	1	2	3	4	5	6	7	8	9	10
	Zaměstnání 1	Zaměstnání 2	Zaměstnání 3	Hobby 1	Hobby 2	Bydlení 1	Bydlení 2	Bydlení 3	Různé 1	Různé 2
Zaměstnání 1	1,00000	0,64740	0,65262	0,59812	0,52110	0,14282	0,14514	0,13779	0,61135	0,54894
Zaměstnání 2	0,64740	1,00000	0,73189	0,68855	0,69778	0,14337	0,18185	0,23598	0,70863	0,68479
Zaměstnání 3	0,65262	0,73189	1,00000	0,63695	0,63003	0,16363	0,23830	0,25465	0,69789	0,67063
Hobby 1	0,59812	0,68855	0,63695	1,00000	0,80469	0,53638	0,63431	0,58284	0,90447	0,84320
Hobby 2	0,52110	0,69778	0,63003	0,80469	1,00000	0,50591	0,49591	0,48235	0,81101	0,75577
Bydlení 1	0,14282	0,14337	0,16363	0,53638	0,50591	1,00000	0,65774	0,59003	0,49842	0,42471
Bydlení 2	0,14514	0,18185	0,23830	0,63431	0,49591	0,65774	1,00000	0,73065	0,64364	0,59340
Bydlení 3	0,13779	0,23598	0,25465	0,58284	0,48235	0,59003	0,73065	1,00000	0,58587	0,51775
Různé 1	0,61135	0,70863	0,69789	0,90447	0,81101	0,49842	0,64364	0,58587	1,00000	0,84142
Různé 2	0,54894	0,68479	0,67063	0,84320	0,75577	0,42471	0,59340	0,51775	0,84142	1,00000
Průměry	97,02958	98,16886	98,94298	98,02522	100,10813	99,50832	101,60288	101,37174	99,16294	98,22383
Sm.odch.	15,52089	11,26746	12,48538	15,93673	19,94109	11,97813	11,06155	12,73475	16,96749	19,11151
Poč.příp	100,00000									
Matice	1,00000									

16. Korelační matice znaků obsahuje Pearsonovy korelační koeficienty dvojic znaků a na diagonále jedničky. Matice je čtvercová a symetrická podle diagonály. Lze ji v této formě i uložit a použít jinde jako vstupní data k další statistické analýze některých vícerozměrných metod (např. MDS).

,23598	0,25465	0,58284	0,48235	0,59003	0,73065	1,00000	0,58587	0,51775
,70863	0,69789	0,90447	0,81101	0,49842	0,64364	0,58587	1,00000	0,84142
,68479	0,67063	0,84320	0,75577	0,42471	0,59340	0,51775	0,84142	1,00000
,16886	98,94298	98,02522	100,10813	99,50832	101,60288	101,37174	99,16294	98,22383
,26746	12,48538	15,93673	19,94109	11,97813	11,06155	12,73475	16,96749	19,11151



17. Kliknutím na OK *Přehledu popisných statistik* se otevře okno *Definice modelu* se dvěma záložkami: *Základní nastavení* a *Popisné statistiky*. Otevřeme okénko *Proměnné pro kanonickou analýzu* a zadáme znaky do prvního a druhého seznamu proměnných.



18. V otevřeném okně *Vyberte 2 seznamy proměnných* jde totiž o levý výběr proměnných (*První seznam proměnných*) a pak o pravý výběr proměnných (*Druhý seznam proměnných*). Zdá se výhodnější rozlišovat slova *levý* a *pravý* než *první* a *druhý*. Do levého seznamu zadáme první tři znaky a do pravého seznamu pak zbývajících sedm znaků. Samozřejmě zadání můžeme i obrátit záměnou levého seznamu za pravý.

STATISTICA C2 - [PS 1* - Dialog výsledků kanonick...

Soubor Zobrazit Statistika Grafy Nástroje Pracovní sešit Okno Nápořád

PS 1*

Kanonická analýza (Factor.st

Dialog popisných statistik

Průměry a směrodat

Korelace (Factor.sta

Kovariance (Factor.s

Krabicový graf

Krabicový graf

Korelace (Factor.sta

Dialog výsledků kanonick

Test chí-kvadrát po c

Kanonické váhy, levé

Kanonické váhy, prav

Struktura faktorů, L :

Získaný rozptyl (Prop

Struktura faktorů, P

Získaný rozptyl (Prop

VI. čísla (Factor.sta)

Test chí-kvadrát po odstranění post.

Kořen odstraněný	Kanonic. R	Kanonic. R-kvad.	Chí-kv.	sv
0	0,884705	0,782703	153,5785	21
1	0,271080	0,073484	10,8516	12
2	0,197374	0,038957	3,7153	5

Kanonické váhy, levá sada (Fact

Proměnná	Kořen 1	Kořen 2	Kořen 3
Zaměstnání 1	0,217021	-1,35890	0,24350
Zaměstnání 2	0,592485	0,37139	-1,38769
Zaměstnání 3	0,300124	0,83222	1,28861

Kanonické váhy, pravá sada (Factor.:

Proměnná	Kořen 1	Kořen 2	Kořen 3
Hobby 1	0,249966	-1,63580	-1,29069
Hobby 2	0,221850	1,00467	-1,09591
Bydlení 1	-0,156475	-0,43401	0,30966
Bydlení 2	-0,447900	0,32580	0,45678
Bydlení 3	-0,066126	0,69965	-0,31653
Různé 1	0,615556	-0,36114	1,73651
Různé 2	0,290859	0,79787	0,45774

Struktura faktorů, L sada (Factor

Proměnná	Kořen 1	Kořen 2	Kořen 3
Zaměstnání 1	0,796461	-0,575348	0,186078
Zaměstnání 2	0,952643	0,100725	-0,286928
Zaměstnání 3	0,875390	0,217192	0,431880

Získaný rozptyl (Proporce), levá sada

Faktor	Získaný rozptyl	Reddnce.
Kořen 1	0,769395	0,602208
Kořen 2	0,129448	0,009512
Kořen 3	0,101157	0,003941

Struktura faktorů, P sada (Factor.sta

Proměnná	Kořen 1	Kořen 2	Kořen 3
Hobby 1	0,823914	-0,099577	0,055376
Hobby 2	0,808859	0,277944	-0,149747
Bydlení 1	0,186554	-0,017177	0,236525
Bydlení 2	0,238228	0,253169	0,456321
Bydlení 3	0,278220	0,414321	0,173396
Různé 1	0,861283	0,048749	0,328389
Různé 2	0,820765	0,245223	0,241010

Získaný rozptyl (Proporce), pravá sada

Proměnná	Získaný rozptyl	Reddnce.
Kořen 1	0,416787	0,326221
Kořen 2	0,055104	0,004049
Kořen 3	0,069379	0,002703

VI. čísla (Factor.sta)

Kořen	Kořen 1	Kořen 2	Kořen 3
Hodnota	0,782703	0,073484	0,038957

Dialog popisných statistik kanonických

Definice modelu: Fact...

Pro nápovědu stiskněte F1

19. Protože jsme v definici modelu zaškrtnuli návětší Dávkové zpracování/protokol, udělal se nyní bleskově protokol 8 tabulek výsledků analýz CCA, které si nyní podrobně prohlédneme a interpretujeme.

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

10
 B
I
U

10
 B
I
U

cká analýza (Factor.st
log popisných statistik
Průměry a směrodatr
Korelace (Factor.sta
Kovariance (Factor.s
Křabicový graf
Křabicový graf
Korelace (Factor.sta
log výsledků kanonick
Test chí-kvadrát po c
Kanonické váhy, levě
Kanonické váhy, prav
Struktura faktorů, L :
Získaný rozptyl (Prop
Struktura faktorů, P
Získaný rozptyl (Prop

Souhrn kanonické analýzy (Factor.sta)			
Kanonické R: ,88471			
Chi2(21)=153,58 p=0,0000			
N=100		L	P
		sada	sada
Počet proměnných		3	7
Získaný rozptyl		100,000%	54,1271%
Celková redundance		61,5661%	33,2973%
Proměnné:	1	Zaměstnání 1	Hobby 1
	2	Zaměstnání 2	Hobby 2
	3	Zaměstnání 3	Bydlení 1
	4		Bydlení 2
	5		Bydlení 3
	6		Různé 1
	7		Různé 2

20. *Souhrn kanonické analýzy*: obsahuje celkové kanonické $R = 0.88471$, a tím vysokou statistickou významnost této korelace, protože $p = 0.0000$ je menší než zadaná hladina významnosti 0.05. Toto R lze vysvětlit jako jednoduchou korelaci mezi váženou sumou skóre v každém souboru s vahami odpovídajícími prvnímu a nejdůležitějšímu kanonickému kořeni.

STATISTICA Cz - [PS 1* - Test chí-kvadrát po odstranění post. kořenů (Factor.sta)]

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

Arial 10 B I U

PS 1*

- Kanonická analýza (Factor.sta)
 - Dialog popisných statistik
 - Průměry a směrodat.
 - Korelace (Factor.sta)
 - Kovariance (Factor.sta)
 - Krabicový graf
 - Krabicový graf
 - Korelace (Factor.sta)
 - Dialog výsledků kanonické analýzy
 - Test chí-kvadrát po odstranění post. kořenů (Factor.sta)
 - Kanonické váhy, levé
 - Kanonické váhy, pravé
 - Struktura faktorů, L
 - Získaný rozptyl (Procenta)
 - Struktura faktorů, P
 - Získaný rozptyl (Procenta)
 - Vl. čísla (Factor.sta)
 - Souhrn kanonické analýzy
 - Vl. čísla (Factor.sta)

Kořen odstraněný	Kanonic. R	Kanonic. R-kvad.	Chí-kv.	sv	p	První lambda
0	0,884705	0,782703	153,5785	21	0,000000	0,193486
1	0,271080	0,073484	10,8516	12	0,541689	0,890422
2	0,197374	0,038957	3,7153	5	0,591097	0,961043

21. *Test Chi2 po odstranění postupných kořenů:* testuje se statistická významnost všech tří kanonických kořenů.

STATISTICA Cz - [PS 1* - Struktura faktorů, L sada (Factor.sta)]

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

Arial 10 B I U

PS 1*

- Kanonická analýza (Factor.sta)
 - Dialog popisných statistik kanonických korelací
 - Průměry a směrodatné odchylky (Factor.sta)
 - Korelace (Factor.sta)
 - Kovariance (Factor.sta)
 - Krabicový graf
 - Krabicový graf
 - Korelace (Factor.sta 10s*100ř)
 - Dialog výsledků kanonických korelací
 - Test chí-kvadrát po odstranění post. kořenů (F)
 - Kanonické váhy, levá sada (Factor.sta)
 - Kanonické váhy, pravá sada (Factor.sta)
 - Struktura faktorů, L sada (Factor.sta)**
 - Získaný rozptyl (Proporce), levá sada (Factor.sta)
 - Struktura faktorů, P sada (Factor.sta)
 - Získaný rozptyl (Proporce), pravá sada (Factor.sta)
 - VI. čísla (Factor.sta)
 - Souhrn kanonické analýzy (Factor.sta)
 - VI. čísla (Factor.sta)

Proměnná	Kořen 1	Kořen 2	Kořen 3
Zaměstnání 1	0,796461	-0,575348	0,186075
Zaměstnání 2	0,952643	0,100725	-0,286926
Zaměstnání 3	0,875390	0,217192	0,431880

22. *Struktura faktorů, Levá sada*: ukazuje jak korelují kanonické kořeny se znaky levé sady. Tyto korelace se nazývají **strukturní koeficienty**. Pouze první kanonický kořen vykazuje významnou korelaci se znaky čili vysoké strukturní koeficienty.

STATISTICA Cz - [PS 1* - Získaný rozptyl (Proporce), levá sada (Factor.sta)]

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

Arial 10 B I U

PS 1*

- Kanonická analýza (Factor.sta)
 - Dialog popisných statistik kanonických korelací
 - Průměry a směrodatné odchylky (Factor.sta)
 - Korelace (Factor.sta)
 - Kovariance (Factor.sta)
 - Krabicový graf
 - Krabicový graf
 - Korelace (Factor.sta 10s*100ř)

Faktor	Získaný rozptyl	Reddnce.
Kořen 1	0,769395	0,602208
Kořen 2	0,129448	0,009512
Kořen 3	0,101157	0,003941

23. *Získaný rozptyl (proporce), Levá sada:* první kanonický kořen V_1 a U_1 vyčíslí 77% rozptylu tří znaků spokojenosti v zaměstnání. Znaký na levé straně vyčíslí okolo 60% rozptylu znaků spokojenosti v zaměstnání pomocí prvního kanonického kořene V_1 a U_1 čili **redundance**, což je tedy průměrné množství rozptylu vyčísleného pro každý znak prvním kanonickým kořenem.

PS 1*

 Kanonická analýza (Factor.sta)

Dialog popisných statistik kanonických korelací

Průměry a směrodatné odchylky (Factor.sta)

Korelace (Factor.sta)

Kovariance (Factor.sta)

 Krabicový graf

 Krabicový graf

Korelace (Factor.sta 10s*100ř)

 Dialog výsledků kanonických korelací

Test chí-kvadrát po odstranění post. kořenů (F)

Kanonické váhy, levá sada (Factor.sta)

	Struktura faktorů, P sada (Factor.sta)					
Proměnná	Kořen 1	Kořen 2	Kořen 3			
Hobby 1	0,823914	-0,099577	0,055376			
Hobby 2	0,808859	0,277944	-0,149747			
Bydlení 1	0,186554	-0,017177	0,236525			
Bydlení 2	0,238228	0,253169	0,456321			
Bydlení 3	0,278220	0,414321	0,173396			
Různé 1	0,861283	0,048749	0,328389			
Různé 2	0,820765	0,245223	0,241010			

24. *Struktura faktorů, Pravá sada*: ukazuje jak korelují kanonické kořeny se znaky *pravé sady* formou **strukturních koeficientů**. Nejvíce koreluje první kanonický kořen V_1 a U_1 se znaky Hobby1, Hobby2, Různé1 a Různé2. Zbylé dva kanonické kořeny korelují málo, statisticky nevýznamně.

STATISTICA Cz - [PS 1* - Získaný rozptyl (Proporce), pravá sada (Factor.sta)]

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

Arial 10 B I U

PS 1*

- Kanonická analýza (Factor.sta)
 - Dialog popisných statistik kanonických korelací
 - Průměry a směrodatné odchylky (Factor.sta)
 - Korelace (Factor.sta)
 - Kovariance (Factor.sta)
 - Krabicový graf
 - Krabicový graf
 - Korelace (Factor.sta 10s*100ř)
 - Dialog výsledků kanonických korelací

Získaný rozptyl (Proporce), pravá sada (Factor.sta)	
Proměnná	Získaný rozptyl
Kořen 1	0,416787
Kořen 2	0,055104
Kořen 3	0,069379

25. Získaný rozptyl (proporce), Pravá sada: první kanonický kořen V_1 a U_1 vyčíslí 42% rozptylu znaků v pravé sadě a okolo 33% rozptylu znaků založených na V_1 a U_1 .

- Získaný rozptyl (Proporce), pravá sada (Factor.sta)
- Vl. čísla (Factor.sta)
- Souhrn kanonické analýzy (Factor.sta)
- Vl. čísla (Factor.sta)

	Kanonické váhy, levá sada (Factor.sta)
--	--

Proměnná	Kořen 1	Kořen 2	Kořen 3
Zaměstnání 1	0,217021	-1,35890	0,243500
Zaměstnání 2	0,592485	0,37139	-1,387699
Zaměstnání 3	0,300124	0,83222	1,28861

26. *Kanonické váhy u kanonického skóre, Levá sada:* Jde o váhy a_1, a_2, a_3 při výpočtu jednotlivých kanonických kořenů. Váhy odpovídají Z-standardizovaným hodnotám znaků v levé sadě. Tyto váhy se užijí k výpočtu kanonického skóre.

STATISTICA Cz - [PS 1* - Kanonické váhy, pravá sada (Factor.sta)]

Soubor Úpravy Zobrazit Vložit Formát Statistika Grafy Nástroje Data Pracovní sešit Okno nápověda

Přidat do sešitu Přidat do protokolu

Arial 10 B I U

PS 1*

- Kanonická analýza (Factor.sta)
 - Dialog popisných statistik kanonických korelací
 - Průměry a směrodatné odchylky (Factor.sta)
 - Korelace (Factor.sta)
 - Kovariance (Factor.sta)
 - Krabicový graf
 - Krabicový graf
 - Korelace (Factor.sta 10s*100ř)
 - Dialog výsledků kanonických korelací
 - Test chí-kvadrát po odstranění post. kořenů (F

Kanonické váhy, pravá sada (Factor.sta)			
Proměnná	Kořen 1	Kořen 2	Kořen 3
Hobby 1	0,249966	-1,63580	-1,29069
Hobby 2	0,221850	1,00467	-1,09591
Bydlení 1	-0,156475	-0,43401	0,30966
Bydlení 2	-0,447900	0,32580	0,45678
Bydlení 3	-0,066126	0,69965	-0,31653
Různé 1	0,615556	-0,36114	1,73651
Různé 2	0,290859	0,79787	0,45774

27. Kanonické váhy u kanonického skóre, Pravá sada: Jde o váhy b_1, b_2, b_3 při výpočtu jednotlivých kanonických kořenů. Váhy odpovídají Z-standardizovaným hodnotám znaků v pravé sadě. Tyto váhy se užijí k výpočtu kanonického skóre.

PS 1*

Kanonická analýza (Factor.sta)

Dialog popisných statistik kanonických korelací

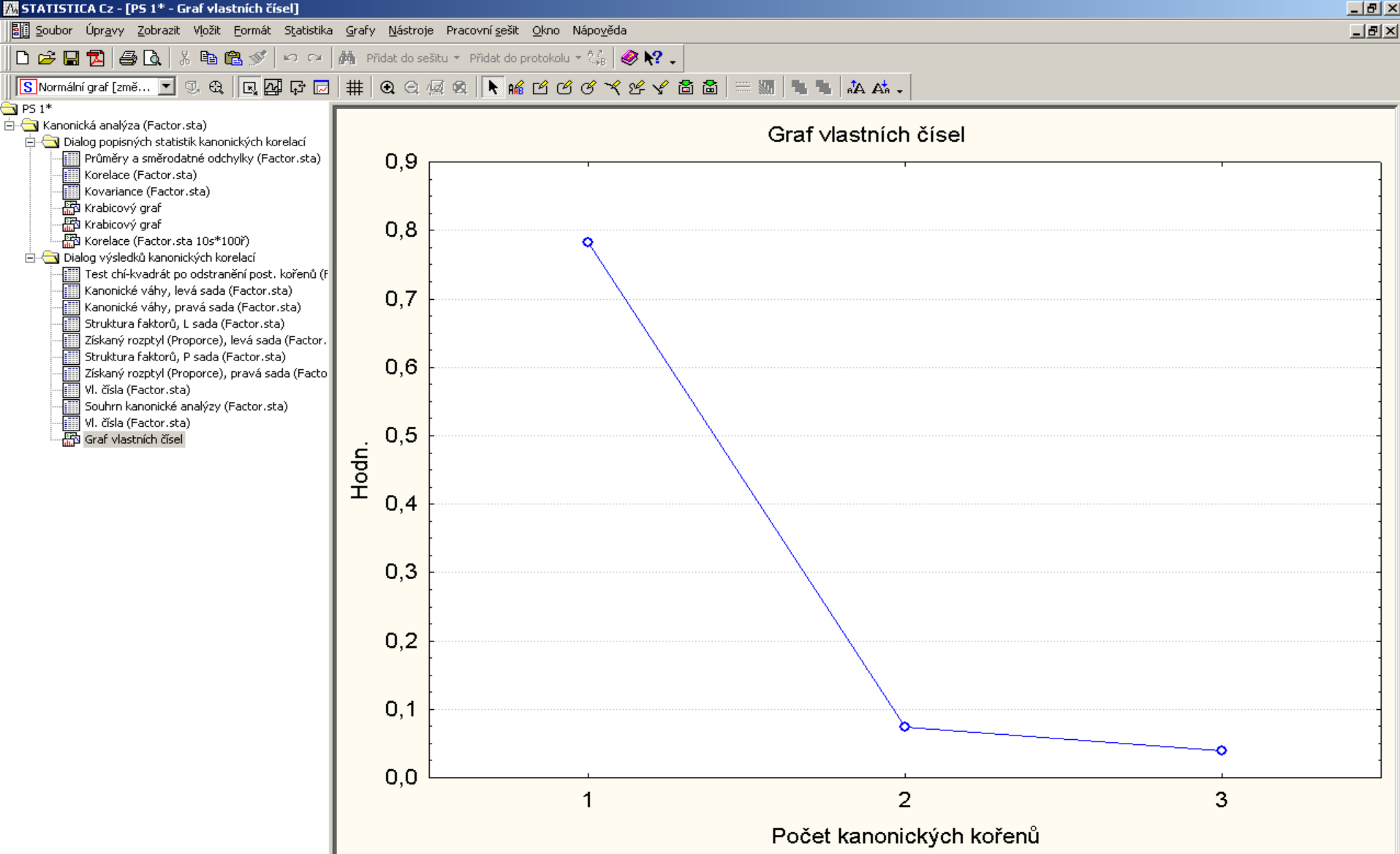
Průměry a směrodatné odchylky (Factor.sta)

Korelace (Factor.sta)

Vl. čísla (Factor.sta)			
Kořen	Kořen 1	Kořen 2	Kořen 3
Hodnota	0,782703	0,073484	0,038957

28. Vlastní čísla u kanonických kořenů: jsou vyčíslena vlastní čísla pro jednotlivé kanonické kořeny.

- Kanonické váhy, levá sada (Factor.sta)
- Kanonické váhy, pravá sada (Factor.sta)
- Struktura faktorů, L sada (Factor.sta)
- Získaný rozptyl (Proporce), levá sada (Factor.sta)
- Struktura faktorů, P sada (Factor.sta)
- Získaný rozptyl (Proporce), pravá sada (Factor.sta)
- Vl. čísla (Factor.sta)
- Souhrn kanonické analýzy (Factor.sta)
- Vl. čísla (Factor.sta)

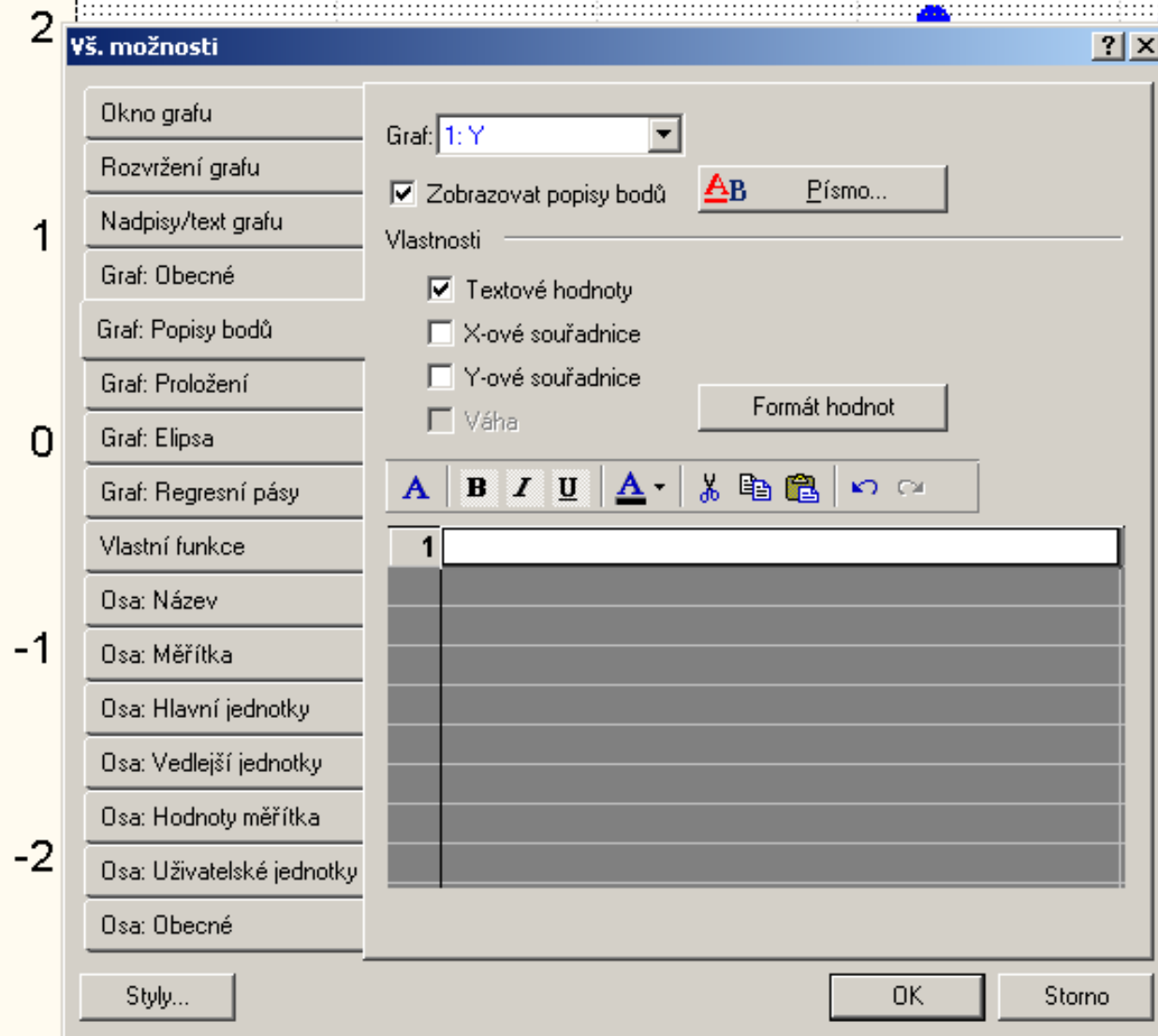


29. *Graf vlastních čísel v závislosti na počtu kanonických kořenů: je to Cattelův indexový graf vlastních čísel a ukazuje, že první kanonický kořen dostatečně popisuje korelaci znaků obou sad.*

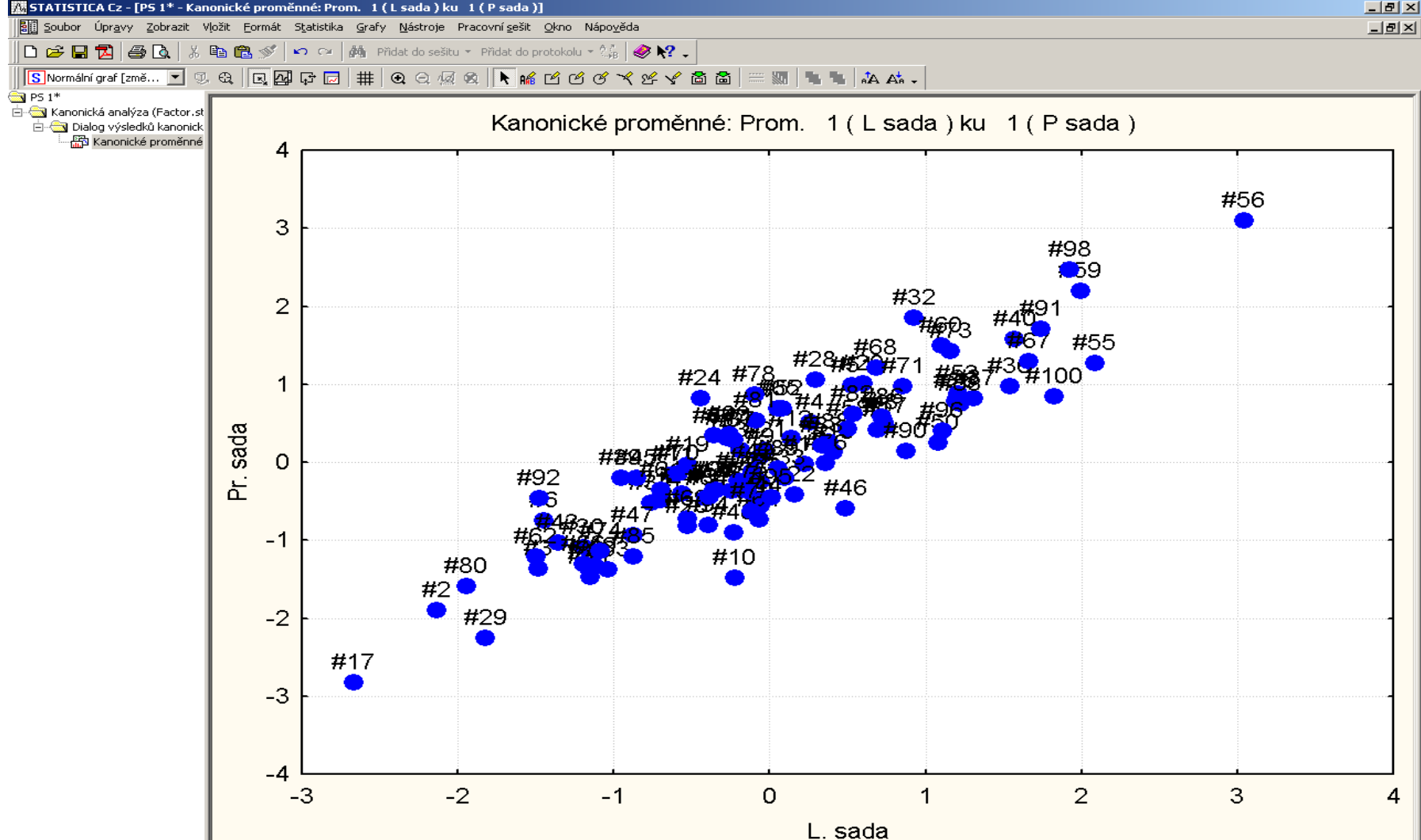
Kořen odstraněný	Korelace, levá i pravá sada (Factor.sta)						
	Hobby 1	Hobby 2	Bydlení 1	Bydlení 2	Bydlení 3	Různé 1	Různé 2
Zaměstnání 1	0,598122	0,521100	0,142818	0,145137	0,137792	0,611346	0,548943
Zaměstnání 2	0,688546	0,697782	0,143365	0,181851	0,235979	0,708630	0,684794
Zaměstnání 3	0,636948	0,630030	0,163629	0,238302	0,254645	0,697894	0,670633

30. *Korelace levé a pravé sady znaků*: odkrývá korelační koeficienty mezi znaky levé a pravé sady.

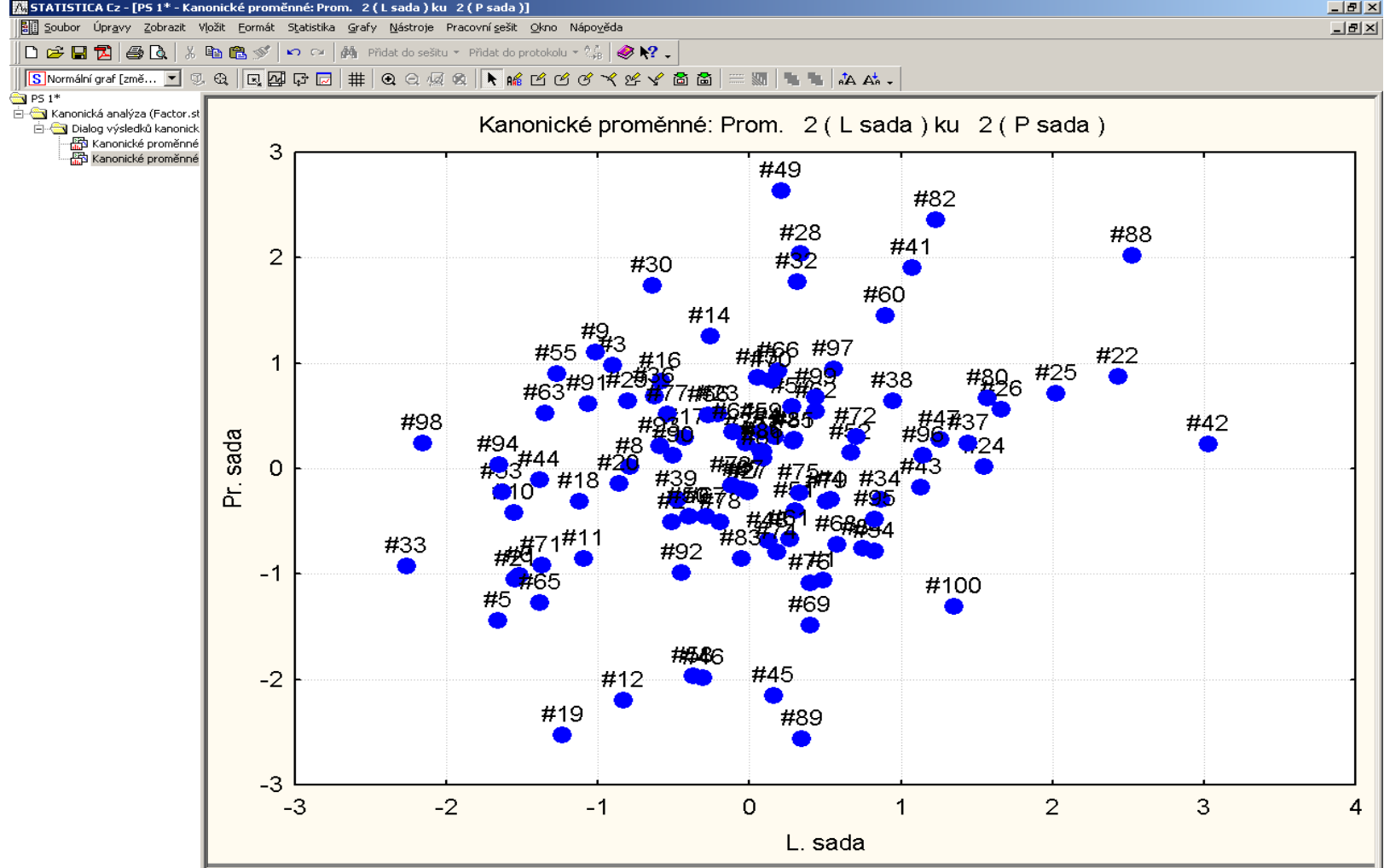
-  Kanonické váhy, levá sada (Factor.sta)
-  Kanonické váhy, pravá sada (Factor.sta)
-  Struktura faktorů, L sada (Factor.sta)
-  Získaný rozptyl (Proporce), levá sada (Factor.sta)
-  Struktura faktorů, P sada (Factor.sta)
-  Získaný rozptyl (Proporce), pravá sada (Factor.sta)
-  Vl. čísla (Factor.sta)
-  Souhrn kanonické analýzy (Factor.sta)
-  Vl. čísla (Factor.sta)
-  Graf vlastních čísel
-  Test chí-kvadrát po odstranění post. kořenů (Factor.sta)
-  Korelace, levá sada (Factor.sta)
-  Korelace, pravá sada (Factor.sta)
-  Korelace, levá i pravá sada (Factor.sta)



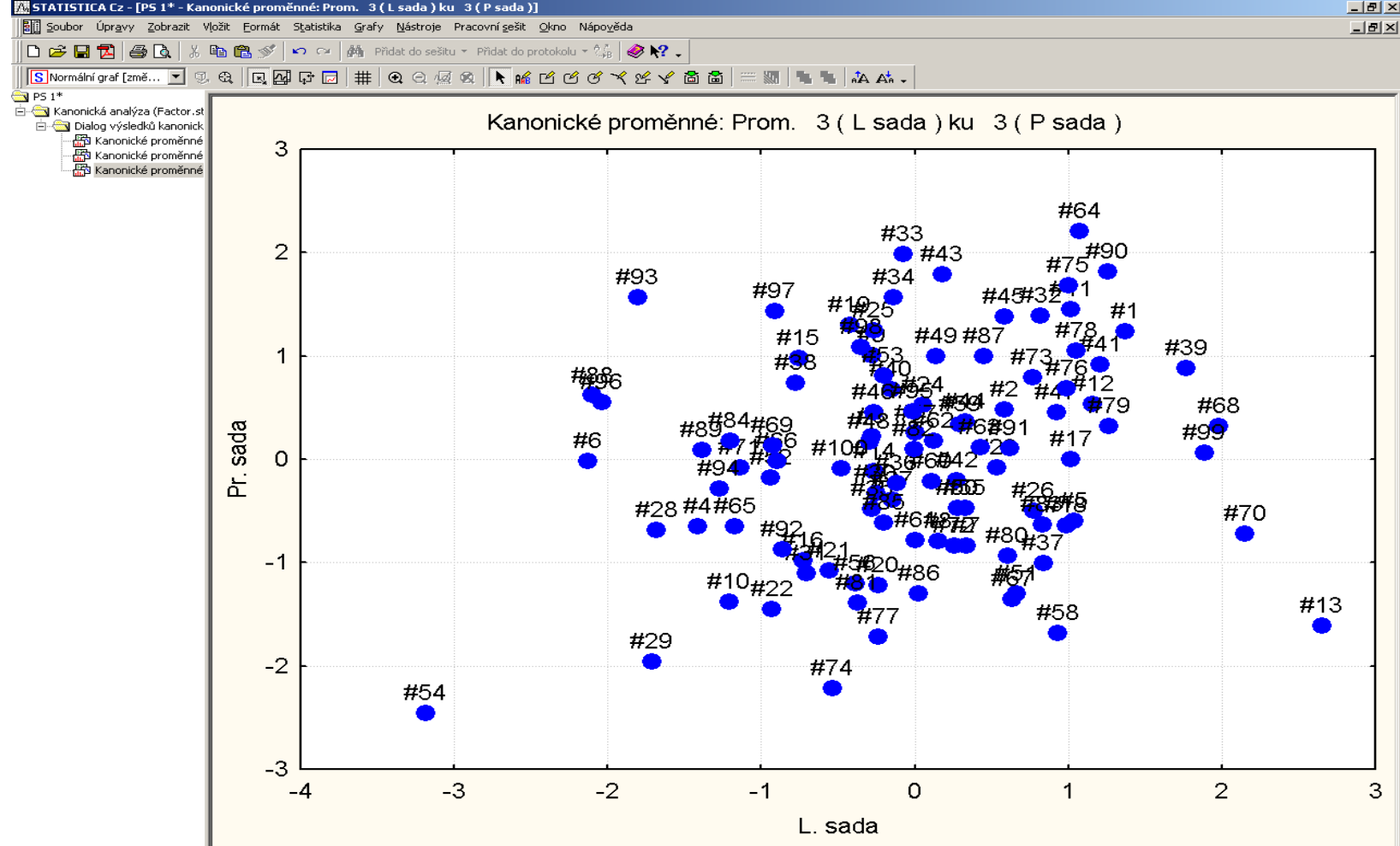
31. Nastavení popisu bodů v grafu kanonického skóre: nastavíme *Zobrazovat popisy bodů* v připravovaném grafu kanonického skóre.



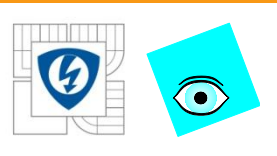
32. *Graf kanonického skóre respondentů pro Levá sada (1) proti Pravá sada(1):* ukazuje na lineární závislost (přímkovou) kanonického skóre pravé sady V_1 na kanonickém kořeni levé sady U_1 . V grafu lze indikovat vybočující body (respondenty), shluky respondentů a také nelinearitu grafu.



33. *Graf kanonického skóre respondentů pro Levá sada (2) proti Pravá sada (2): dosahuje stejných závěrů jako u předešlého grafu. Ukazuje na ne tolik zřetelnou lineární závislost (přímkovou) kanonického skóre pravé sady V_2 na kanonickém kořeni levé sady U_2 . V grafu lze indikovat vybočující body (respondenty), shluky respondentů a také nelinearitu grafu.*



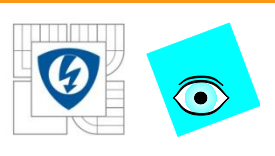
33. *Graf kanonického skóre respondentů pro Levá sada (3) proti Pravá sada (3): dosahuje stejných závěrů jako u předešlého grafu. Ukazuje na ne tolik zřetelnou lineární závislost (přímkovou) kanonického skóre pravé sady V_3 na kanonickém kořeni levé sady U_3 . V grafu lze indikovat vybočující body (respondenty), shluky respondentů a také nelinearitu grafu.*



PŘÍKLAD 6.2 Testy IQ kanonickou korelační analýzou

Bylo vyšetřeno 15 respondentů pěti rozličnými testy a vyčíslena hodnota inteligenčního kvocientu IQ (dohromady šest původních znaků) za účelem zjištění hodnoty výsledného inteligenčního kvocientu. Jde o korelaci vždy mezi dvojicí testů, když je k dispozici 15 dvojic vážených průměrů. Pokuste se vyšetřit tři vybrané testy v závislosti na prvních třech testech, čili pokuste se je popsat závislostí $(\text{TEST4}, \text{TEST5}, \text{IQ}) = f(\text{TEST1}, \text{TEST2}, \text{TEST3})$.

- **Data:** Každý z testů obsahoval 10 bodovaných otázek (0 až 100 bodů). Matice TEST1 až TEST5 a IQ byly velikosti 15 x 10. Kanonická korelační analýza je použita pro korelaci 15 hodnot váženého průměru z 10 bodovaných odpovědí každého testu s 15 hodnotami váženého průměru 10 bodovaných odpovědí jiného testu.



Data

<i>Objekt</i>	<i>Test1</i>	<i>Test2</i>	<i>Test3</i>	<i>Test4</i>	<i>Test5</i>	<i>IQ</i>
1	83	34	65	63	64	106
2	73	19	73	48	82	92
3	54	81	82	65	73	102
4	96	72	91	88	94	121
5	84	53	72	68	82	102
6	86	72	63	79	57	105
7	76	62	64	69	64	97
8	54	49	43	52	84	92
9	37	43	92	39	72	94
10	42	54	96	48	83	112
11	71	63	52	69	42	130
12	63	74	74	71	91	115
13	69	81	82	75	54	98
14	81	89	64	85	62	96
15	50	75	72	64	45	103

STATISTICA Cz - [Data: 61IQ.sta (10s krát 15ř)]

Soubor Upravit Zobrazit Vložit Formát Statistky Data mining Grafy Nástroje Data Okno Nápověda

Přidat do sešitu Přidat do protokolu Přidat do MS Word

Arial 10 B I U

	1	2	3	4	5	6	7	8	9	10
	Objekt	Test1	Test2	Test3	Test4	Test5	IQ	Prom8	Prom9	Prom10
1	1	83	34	65	63	64	106			
2	2	73	19	73	48	82	92			
3	3	54	81	82	65	73	102			
4	4	96	72	91	88	94	121			
5	5	84	53	72	68	82	102			
6	6	86	72	63	79	57	105			
7	7	76	62	64	69	64	97			
8	8	54	49	43	52	84	92			
9	9	37	43	92	39	72	94			
10	10	42	54	96	48	83	112			
11	11	71	63	52	69	42	130			
12	12	63	74	74	71	91	115			
13	13	69	81	82	75	54	98			
14	14	81	89	64	85	62	96			
15	15	50	75	72	64	45				

STATISTICA Cz - [Data: 61IQ.sta (10s krát 15ř)]

Soubor Upravit Zobrazit Vložit Formát Statistky Data mining Grafy Nástroje Data Okno Nápověda

Přidat do sešitu Přidat do protokolu Přidat do MS Word

Arial 10 B I U

	1	2	3	4	5	6	7	8	9	10
	Objekt	Test1	Test2	Test3	Test4	Test5	IQ	Prom8	Prom9	Prom10
1	1	83	34	65	63	64	106			
2	2	73	19	73	48	82	92			
3	3	54	81	82	65	73	102			
4	4	96	72	91	88	94	121			
5	5	84	53	72	68	82	102			
6	6	86	72	63	79	57	105			
7	7	76	62	64	69	64	97			
8	8	54	49	43	52	84	92			
9	9	37	43	92	39	72	94			
10	10	42	54	96	48	83	112			
11	11	71	63	52	69	42	130			

Vyberte proměnné pro kanonickou analýzu

1 - Objekt
 2 - Test1
 3 - Test2
 4 - Test3
 5 - Test4
 6 - Test5
 7 - IQ
 8 - Prom8
 9 - Prom9
 10 - Prom10

[Vše] [Svazky...]

Vybrat vše Roztáhnout Přiblížit

Vyberte proměnné:

2-7

Kanonická analýza: 61IQ.sta

Základní nastavení

☒ Proměnné: žádné

Vstupní soubor: Zdrojová data

☐ Přehled popisných statistik a korelační matice

STATISTICA Cz - [Data: 61IQ.sta (10s krát 15ř)]

Soubor Upravit Zobrazit Vložit Formát Statistky Data mining Grafy Nástroje Data Okno Nápověda

Přidat do sešitu Přidat do protokolu Přidat do MS Word

Arial 10 B I U

	1	2	3	4	5	6	7	8	9	10
	Objekt	Test1	Test2	Test3	Test4	Test5	IQ	Prom8	Prom9	Prom10
1	1	83	34	65	63	64	106			
2	2	73	19	73	48	82	92			
3	3	54	81	82	65	73	102			
4	4	96	72	91	88	94	121			
5	5	84	53	72	68	82	102			
6	6	86	72	63	79	57	105			
7	7	76	62	64	69	64	97			
8	8	54	49	43	52	84	92			
9	9	37	43	92	39	72	94			
10	10	42	54	96	48	83	112			
11	11	71	63	52	69	42	130			
12	12	63	74	74	71	91	115			
13	13	69	81	82	75	54	98			
14	14	81	89	64	85	62	96			
15	15	50	75	72	64	45				

Vyberte 2 seznamy proměnných:

2 - Test1
 3 - Test2
 4 - Test3
 5 - Test4
 6 - Test5
 7 - IQ

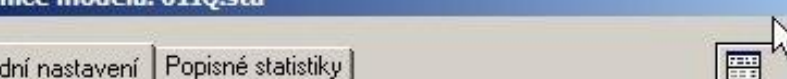
2 - Test1
 3 - Test2
 4 - Test3
 5 - Test4
 6 - Test5
 7 - IQ

Vybrat vše Podrobn. Přiblížit

První seznam proměnných: 2-4

Druhý seznam proměnných: 5-7

☐ Pouze odpovídající proměnné



Definice modelu: 61IQ.sta

Základní nastavení | Popisné statistiky

 Proměnné pro kanon. analýzu

První seznam: Test1-Test3

Druhý seznam: Test4-IQ

☐ Dávkové zpracování/protokol

OK

Storno

 Možnosti ▼

	1	2	3	4	5	6	7	8	9	10			
	Objekt	Test1	Test2	Test3	Test4	Test5	IQ	Prom8	Prom9	Prom10			
1	1	83	34	65	63	64	106						
2	2	73	19	73	48	82	92						
3	3	54	81	82	65	73	102						
4	4	96	72	91	88	94	121						
5	5	84	53	72	68	82	102						
6	6	86	72	63	79	57	105						
7	7	76	62	64	69	64	97						
8	8	54	49	43	52								
9	9	37	43	92	39								
10	10	42	54	96	48								
11	11	71	63	52	69								
12	12	63	74	74	71								
13	13	69	81	82	75								
14	14	81	89	64	85								
15	15	50	75	72	64								

Výsledky kanonické analýzy: 61IQ.sta

Kanonické R: .9955997
 Chi-kvadr.: 52.37490 sv = (9) p = .0000000
 Počet platných případů: 15

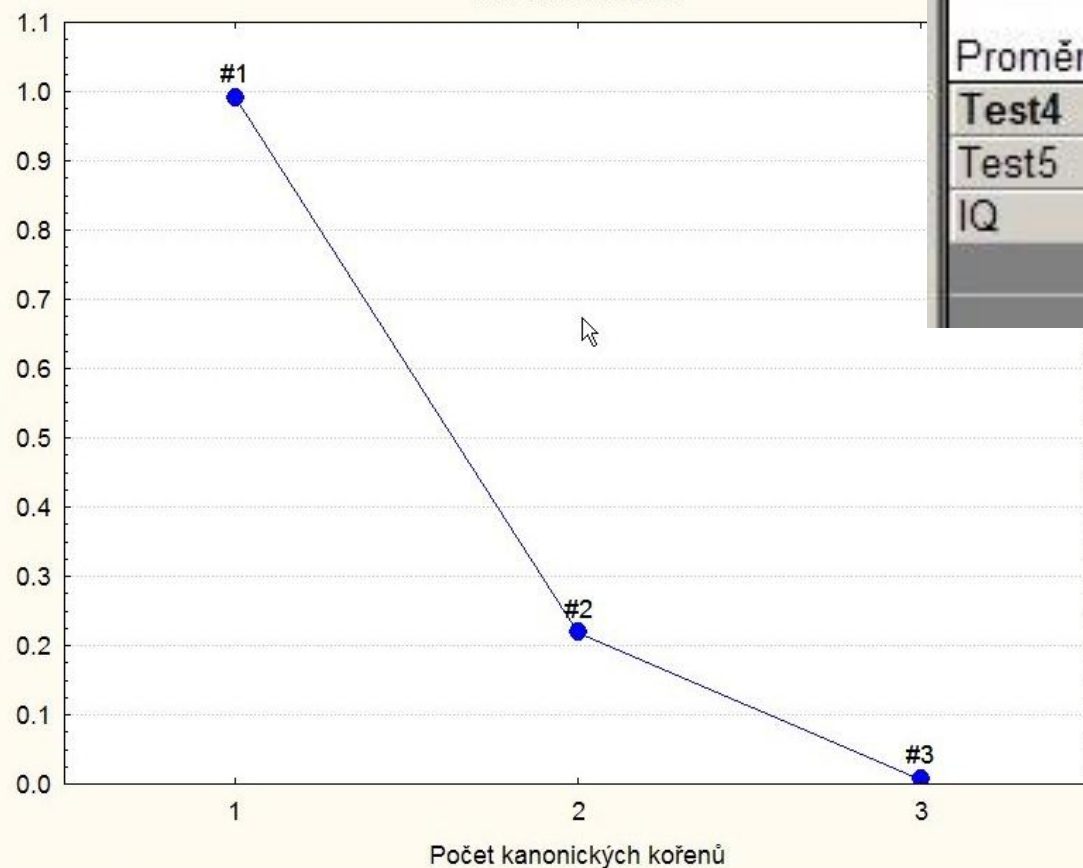
	Počet	Získaný	Celk. redundance
	prom.	rozptyl	při dané druhé sadě
Levá sada:	3	100.00000000%	42.760176532%
Pravá sada:	3	100.00000000%	44.454583851%

Struktury faktorů Kanon. skóre
 Základní nastavení Kanon. faktory

Souhrn: Kanonické výsledky

Souhrn
 Storno
 Možnosti ▾
 Anal. sk.

Graf vlastních čísel



Proměnná	Kanonické váhy, pravá sada (61IQ.sta)		
	Kořen 1	Kořen 2	Kořen 3
Test4	-1.02137	-0.104989	0.37086
Test5	0.00599	-0.990267	0.22402
IQ	0.06536	-0.229775	-1.05024

Kanonické analýzy (61IQ.sta)]

Formát Statisticky Data mining Grafy Nástroje Data Pracovní

Přidat do sešitu Přidat do protokolu

Arial 10

B I U

PS 1*

Kanonická analýza (61IQ.sta)

Dialog výsledků kanonické analýzy

Graf vlastních čísel

Souhrn kanonické analýzy

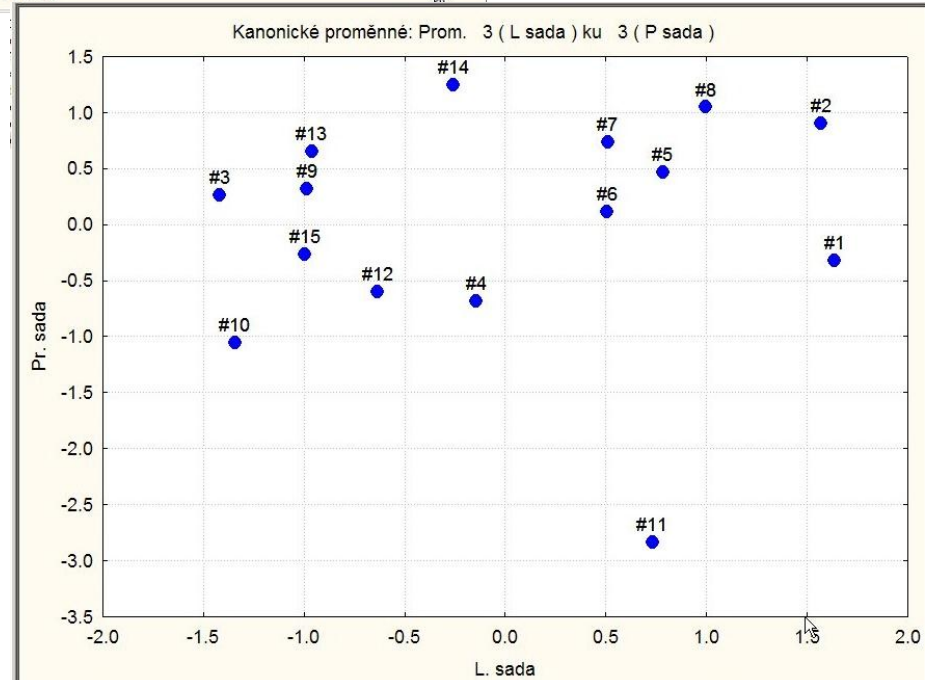
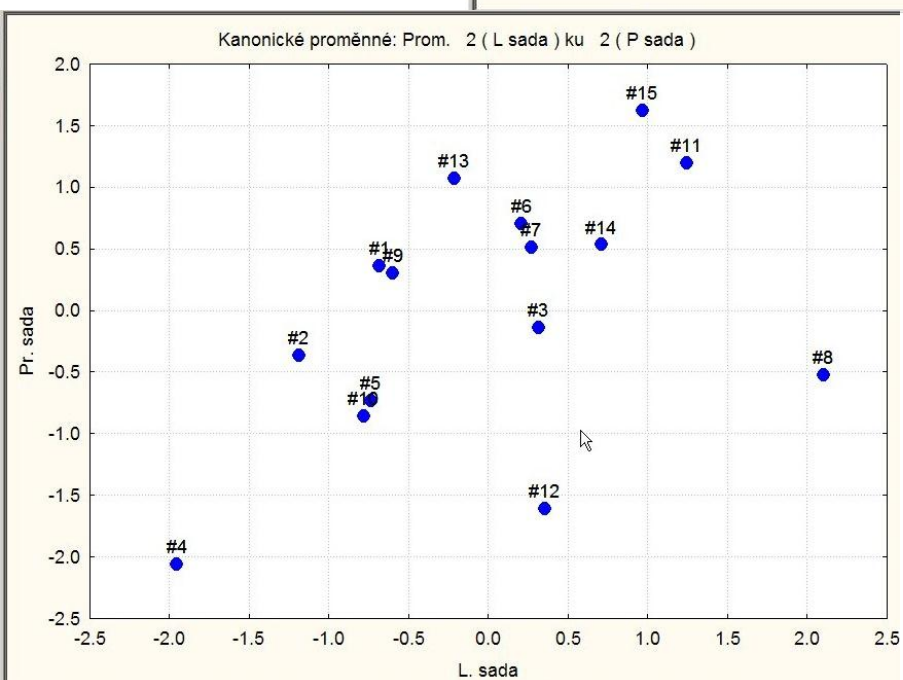
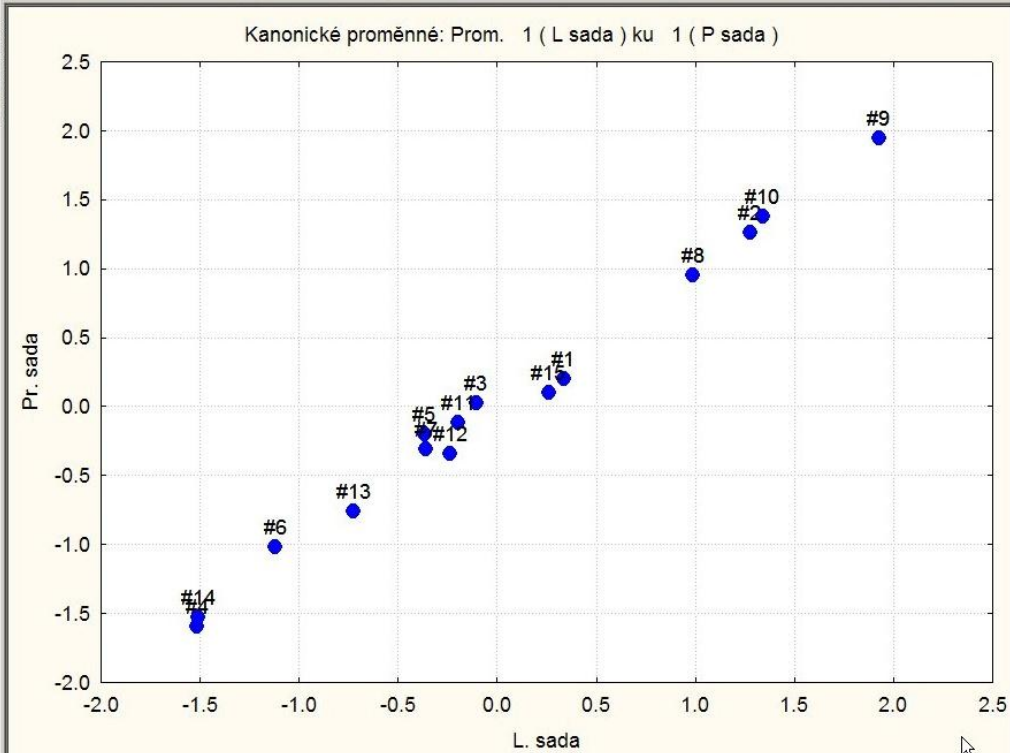
Souhrn kanonické analýzy (61IQ.sta)

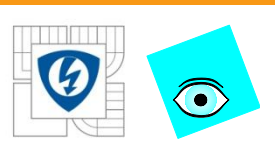
Kanonické R: .99560

Chi2(9)=52.375 p=0.0000

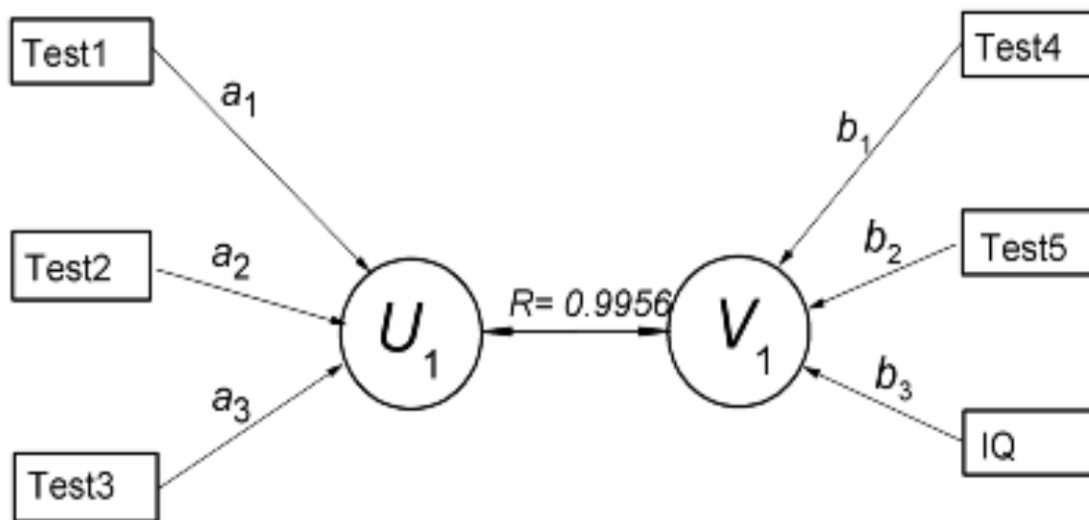
N=15

	L sada	P sada
Počet proměnných	3	3
Získaný rozptyl	100.000%	100.000%
Celková redundance	42.7602%	44.4546%
Proměnné:	Test1	Test4
	Test2	Test5
	Test3	IQ

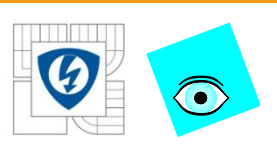




Obr. 6.8b

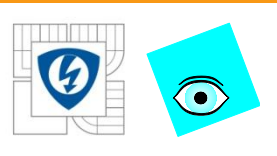


Obr. 6.8b První pár kanonických korelačních proměnných U_1 a V_1 dostatečně popisuje závislost 3 znaků $Test1$, $Test2$, $Test3$ na 3 znacích $Test4$, $Test5$, IQ zdrojové matice dat $IQ(STATISTICA, ORIGIN)$.

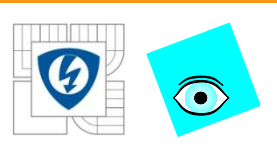


PŘÍKLAD 6.3 Ovlivnění subjektivních postojů a názorů žen svým zdravotním stavem

V roce 1975 bylo dotazníkem vyšetřeno v anglicky hovořící oblasti San Fernando Valley na předměstí Los Angeles 465 žen věku od 20 do 59 let. Psychologický průzkum byl zaměřen na zdravotní, demografické a subjektivně zbarvené postoje a stavy. Cílem byl bylo odhalit míru znaků osobního postoje vztažené ke svým zdravotním charakteristikám. Je třeba analyzovat tuto zdrojovou matici dat a odhalit existují-li zde nějaké skryté vztahy.

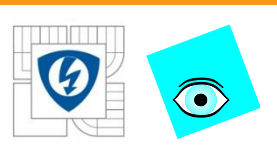


- *Ve zdrojové matici dat Canon každá z respondentek bodově odpověděla na 9 otázek.*
- ***V 1. souboru proměnné U jsou 4 znaky:*** *Postženy značí míru důležitosti postavení ženy (Attrole), Domin značí míru dominantního postavení v rodině (Control), Manžel je hodnotící postoj ke svému manželství (Attmar), Sebehod značí míru sebehodnocení (Esteem). Větší čísla těchto položek ukazují negativnější, nepříznivější postoj, například zvyšující se konzervativní postoj ke správné roli ženy, dále třeba zvýšené pocity bezmocnosti ovlivnit něčí osud (vnější jsou opakem vnitřního ovlivnění) nebo narůstající nespokojenost se svým manželstvím a stále větší sebepodceňování.*



Data pokračování

- ***V 2. souboru proměnné V je 5 znaků:*** zdravotního stavu respondentky: *Mental se týká mentální zdraví (Menheal), Fyzické se týká fyzického zdraví (Physheal), Lékař vyjadřuje časté návštěvy u lékaře (Timedrs), Léky vyjadřuje oblibu v užívání léků (Attdrug), Drogy značí frekvence užívání psychotropních léků (Druguse).*
- *Bodově vyšší hodnoty vyjadřují horší mentální a fyzické zdraví, horší zdravotní stav přináší častější návštěvy lékaře a větší snahu užívat léky a také jejich větší spotřebu. Jsou-li hodnoty transformovány, je znak označen koncovkou L.*



Data pokračování

<i>Objekt</i>	<i>Lékař</i>	<i>Léky</i>	<i>Fyzické</i>	<i>Mental</i>	<i>Sebehod</i>	<i>Domin</i>	<i>Manžel</i>	<i>Drogy</i>	<i>Postženy</i>
1	1	8	5	8	16	5	36	3	42
2	3	7	4	6	17	5	21	0	38

	1 Objekt	2 Lekar	3 LekarL	4 Leky	5 Fyzicke	6 FyzickéL	7 Mental	8 Sebehod	9 Domin	10 Manzel	11 ManzelL	12 Drogy	13 DrogyL	14 Postzeny
1	1	1	0	8	5	1.60943791	8	16	5	36	3.58351894	3	1.09861229	42
2	2	3	0.47712125	7	4	1.38629436	6	17	5	21	3.04452244	0		38
3	3	0		8	3	1.09861229	4	21	6	20	2.99573227	3	1.09861229	44
4	4	13	1.11394335	9	2	0.69314718	2	17	6	24	3.17805383	5	1.60943791	31
5	5	15	1.17609126	7	3	1.09861229	6	8	6	15	2.7080502	24	3.17805383	29
6	6	3	0.47712125	8	5	1.60943791	5	18	7	28	3.33220451	3	1.09861229	26
7	7	2	0.30103	7	5	1.60943791	6	14	6	27	3.29583687	0		44
8	8	0		7	4	1.38629436	5	18	8	18	2.89037176	0		48
9	9	7	0.84509804	7	5	1.60943791	4	12	5	12	2.48490665	2	0.69314718	32
10	10	4	0.60205999	8	3	1.09861229	9	18	7	53	3.97029191	10	2.30258509	24
11	11	15	1.17609126	9	6	1.79175947	3	15	5	11	2.39789527	18	2.89037176	43
12	12	0		6	3	1.09861229	5	20	6	16	2.77258872	1	0	45
13	13	2	0.30103	6	3	1.09861229	10	14	7	17	2.83321334	3	1.09861229	52
14	14	13	1.11394335	8	6	1.79175947	9	16	6	20	2.99573227	29	3.36729583	45
15	15	2	0.30103	5	3	1.09861229	2	10	6	15	2.7080502	0		41
16	16	2	0.30103	8	3	1.09861229	4	13	7	12	2.48490665	1	0	35
17	21	1	0	8	3	1.09861229	1	10	7	18	2.89037176	0		37
18	22	2	0.30103	6	7	1.94591015	8	13	6	11	2.39789527	14	2.63905733	46
19	23	5	0.69897	8	4	1.38629436	6	18	6	21	3.04452244	19	2.94443898	30
20	24	5	0.69897	10	7	1.94591015	13	17	9	14	2.63905733	7	1.94591015	37
21	25	3	0.47712125	6	4	1.38629436	12	18	9	26	3.25809654	1	0	38
22	26	4	0.60205999	5	2	0.69314718	3	11	8	25	3.21887582	3	1.09861229	26
23	27	2	0.30103	8	3	1.09861229	4	13	6	21	3.04452244	3	1.09861229	34
24	28	0		8	5	1.60943791	14	26	8	38	3.63758616	7	1.94591015	46
25	29	13	1.11394335	9	7	1.94591015	10	18	5	32	3.4657359	26	3.25809654	33
26	30	7	0.84509804	9	8	2.07944154	11	11	7	23	3.13549422	12	2.48490665	24
27	31	2	0.30103	8	6	1.79175947	9	18	9	21	3.04452244	2	0.69314718	38
28	32	12	1.07918125	9	9	2.19722458	12	16	6	25	3.21887582	35	3.55534806	30
29	33	2	0.30103	6	3	1.09861229	8	16	7	35	3.55534806	2	0.69314718	41
30	34	5	0.69897	9	7	1.94591015	9	9	9	19	2.94443898	21	3.04452244	34
31	35	4	0.60205999	7	8	2.07944154	4	13	6	19	2.94443898	7	1.94591015	37
32	36	6	0.77815125	6	8	2.07944154	14	16	6	16	2.77258872	2	0.69314718	24

Kanonická analýza: 63CanonLsta

Základní nastavení

Proměnné: žádné

Vstupní soubor: Zdrojová data

☐ Přehled popisných statistik a korelační matice

OK

Storno

Možnosti

Otevři Data

SELECT CASES

ChD vynechána

☒ Celé případy

☐ Nahradit průměrem

Zvolte proměnné pro kanonickou analýzu

1 - Objekt 13 - DrogyL

2 - Lekar 14 - Postzeny

3 - LekarL

4 - Leky

5 - Fyzicke

6 - FyzickéL

7 - Mental

8 - Sebehod

9 - Domin

10 - Manzel

11 - Manzell

12 - Drogy

OK

Storno

[Svazky]...

Pro zobrazení odpovídajících proměnných zvolte "Ukázat pouze...". Pro více informací stiskněte F1.

Vybrat vše Roztáhnout Přiblížit

Vyberte proměnné:

2-14

☐ Pouze odpovídající proměnné

Kanonická analýza: 63CanonLsta

Základní nastavení

Proměnné: Lekar-Postzeny

Vstupní soubor: Zdrojová data

☐ Přehled popisných statistik a korelační matice

OK

Storno

Možnosti

Otevři Data

SELECT CASES

ChD vynechána

☒ Celé případy

☐ Nahradit průměrem

Vyberte 2 seznamy proměnných:

2 - Lekar	12 - Drogy	2 - Lekar	12 - Drogy
3 - LekarL	13 - DrogyL	3 - LekarL	13 - DrogyL
4 - Leky	14 - Postzeny	4 - Leky	14 - Postzeny
5 - Fyzicke		5 - Fyzicke	
6 - FyzickéL		6 - FyzickéL	
7 - Mental		7 - Mental	
8 - Sebehod		8 - Sebehod	
9 - Domin		9 - Domin	
10 - Manzel		10 - Manzel	
11 - Manzell		11 - Manzell	

První seznam proměnných: 8-9 11 14
 Druhý seznam proměnných: 3-4 6-7 13

☐ Pouze odpovídající proměnné

Definice modelu: 63CanonL.sta

Základní nastavení | Popisné statistiky

 Proměnné pro kanon. analýzu

První seznam: Sebehod-Domin Manzell Postzeny

Druhý seznam: LekarL-Leky FyzickéL-Mental DrogyL

☐ Dávkové zpracování/protokol

Výsledky kanonické analýzy: 63CanonL.sta

Kanonické R: .3853026
 Chi-kvadr.: 82.75612 sv = (20) p = 0.000000
 Počet platných případů: 355

	Počet	Získaný prom.	Získaný rozptyl	Celk. redundance při dané druhé sadě
Levá sada:	4	100.00000000%	7.372501282%	
Pravá sada:	5	61.521542430%	4.246736707%	

Struktury faktorů Kanon. skóre
 Základní nastavení Kanon. faktory

Souhrn: Kanonické výsledky

Storno

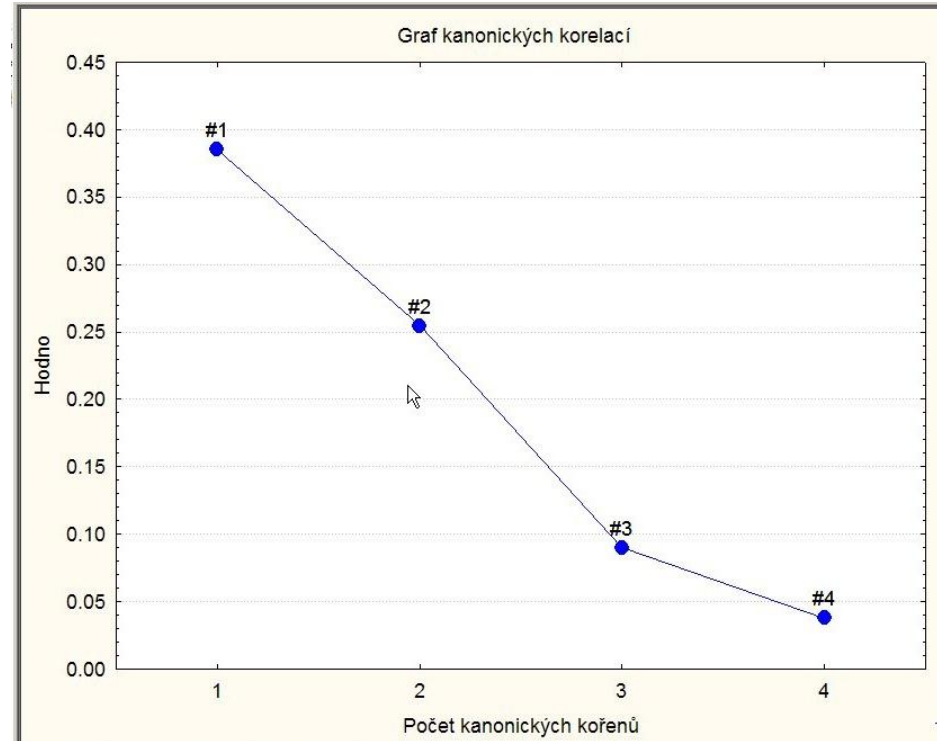
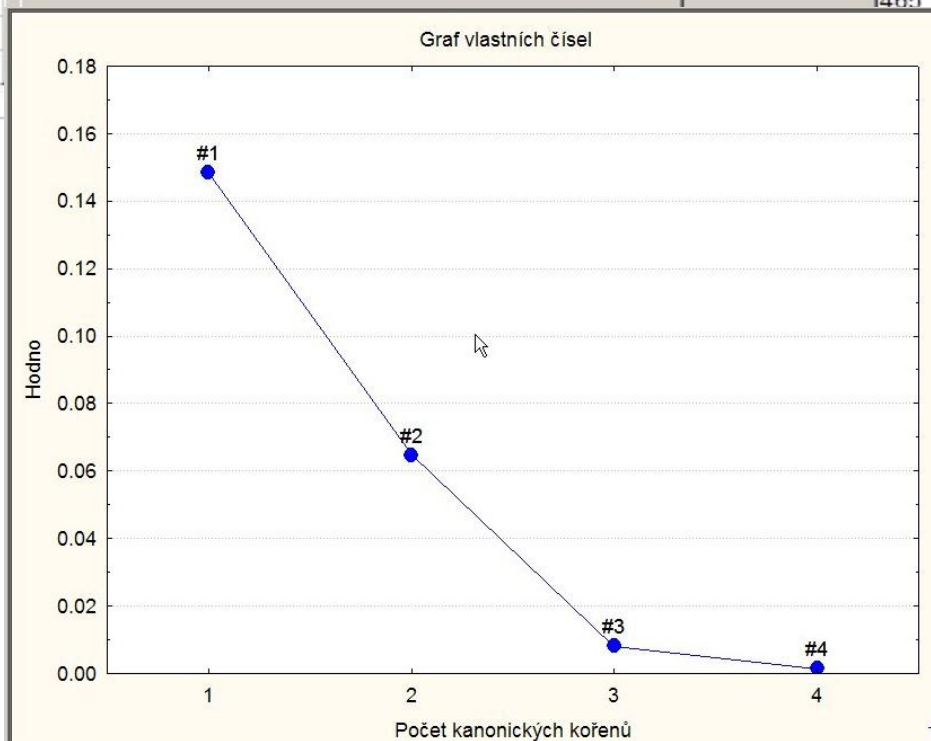
Možnosti

Anal. sk.

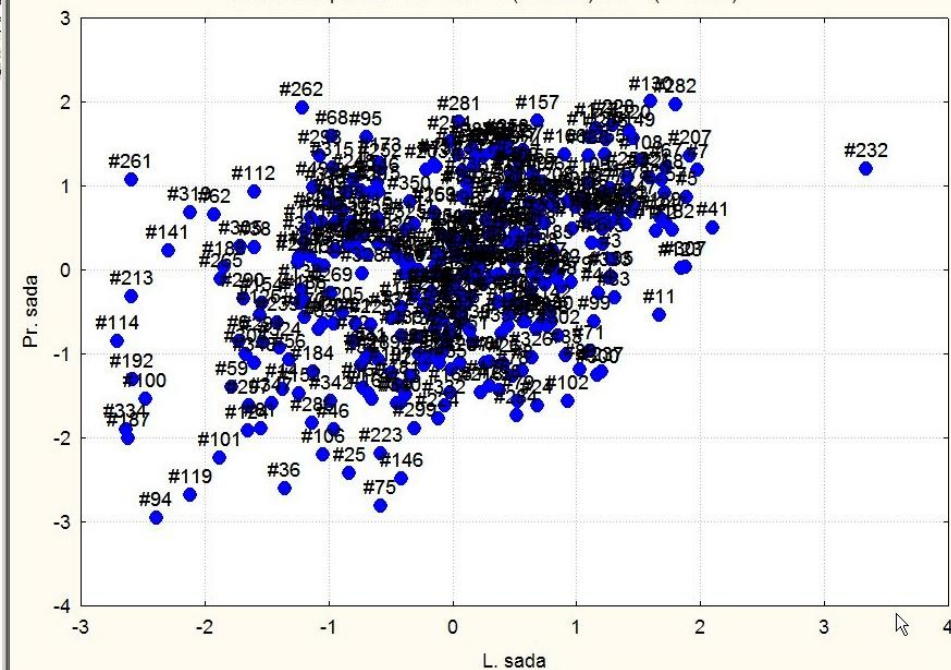
Souhrn kanonické analýzy (63CanonL.sta)
 Kanonické R: .38530
 Chi2(20)=82.756 p=0.0000

	L sada	P sada
N=355		
Počet proměnných	4	5
Získaný rozptyl	100.000%	61.5215%
Celková redundance	7.37250%	4.24674%
Proměnné:	1	Sebehod
	2	Domin
	3	Manzell
	4	Postzeny
	5	

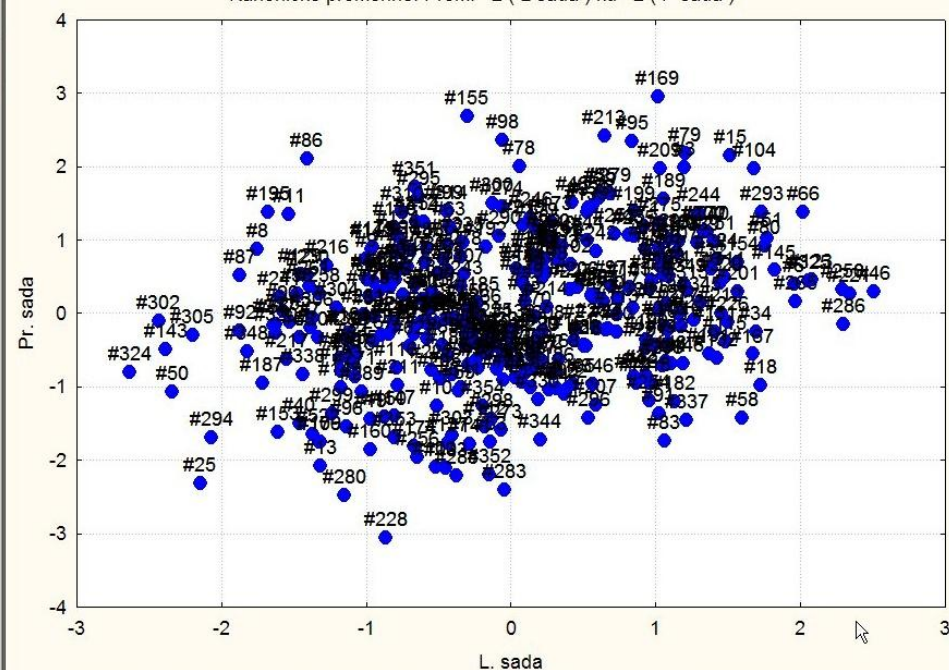
	LekarL	Leky	FyzickéL	Mental	DrogyL



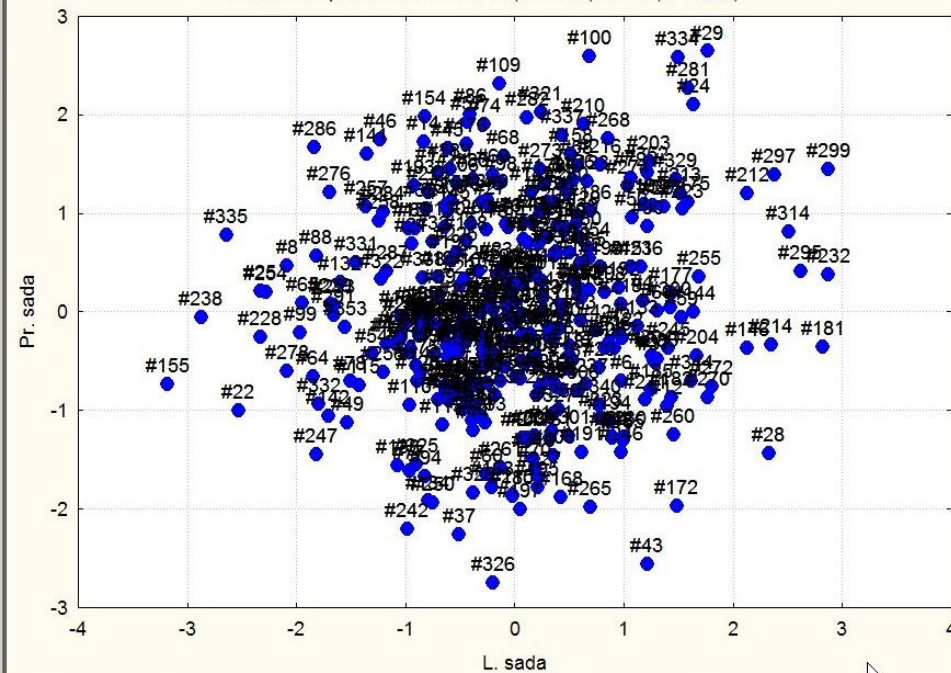
Kanonické proměnné: Prom. 1 (L sada) ku 1 (P sada)



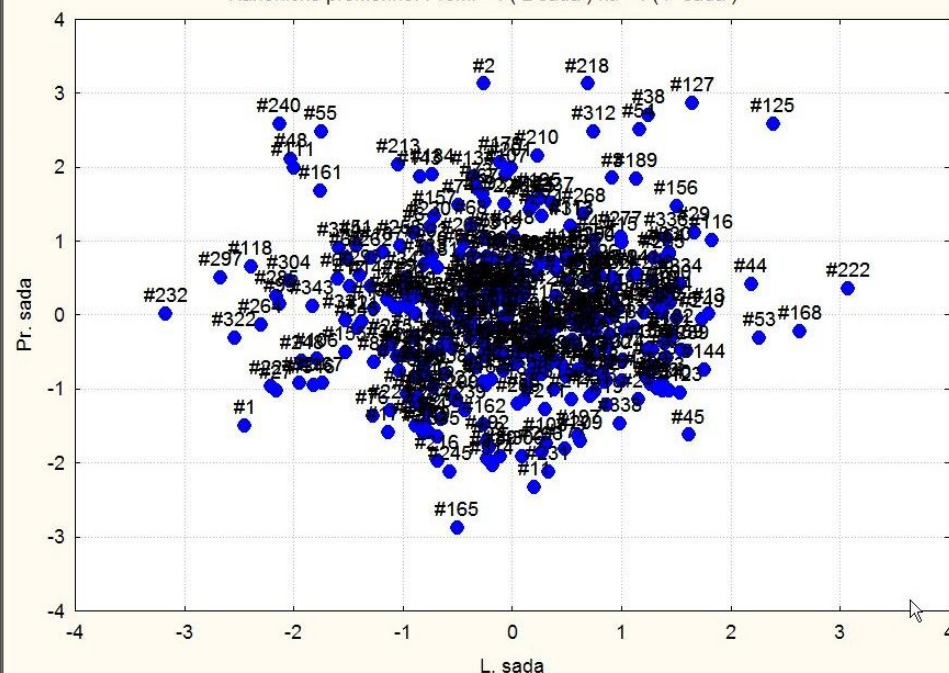
Kanonické proměnné: Prom. 2 (L sada) ku 2 (P sada)

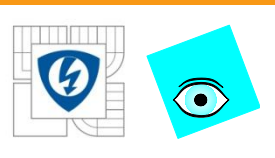


Kanonické proměnné: Prom. 3 (L sada) ku 3 (P sada)

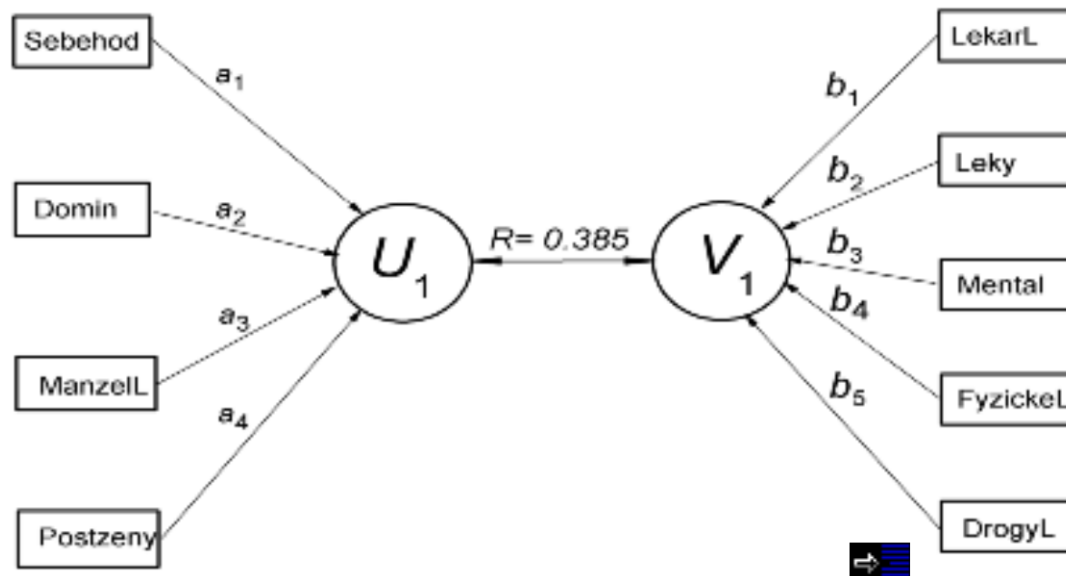


Kanonické proměnné: Prom. 4 (L sada) ku 4 (P sada)

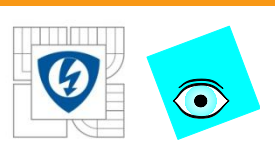




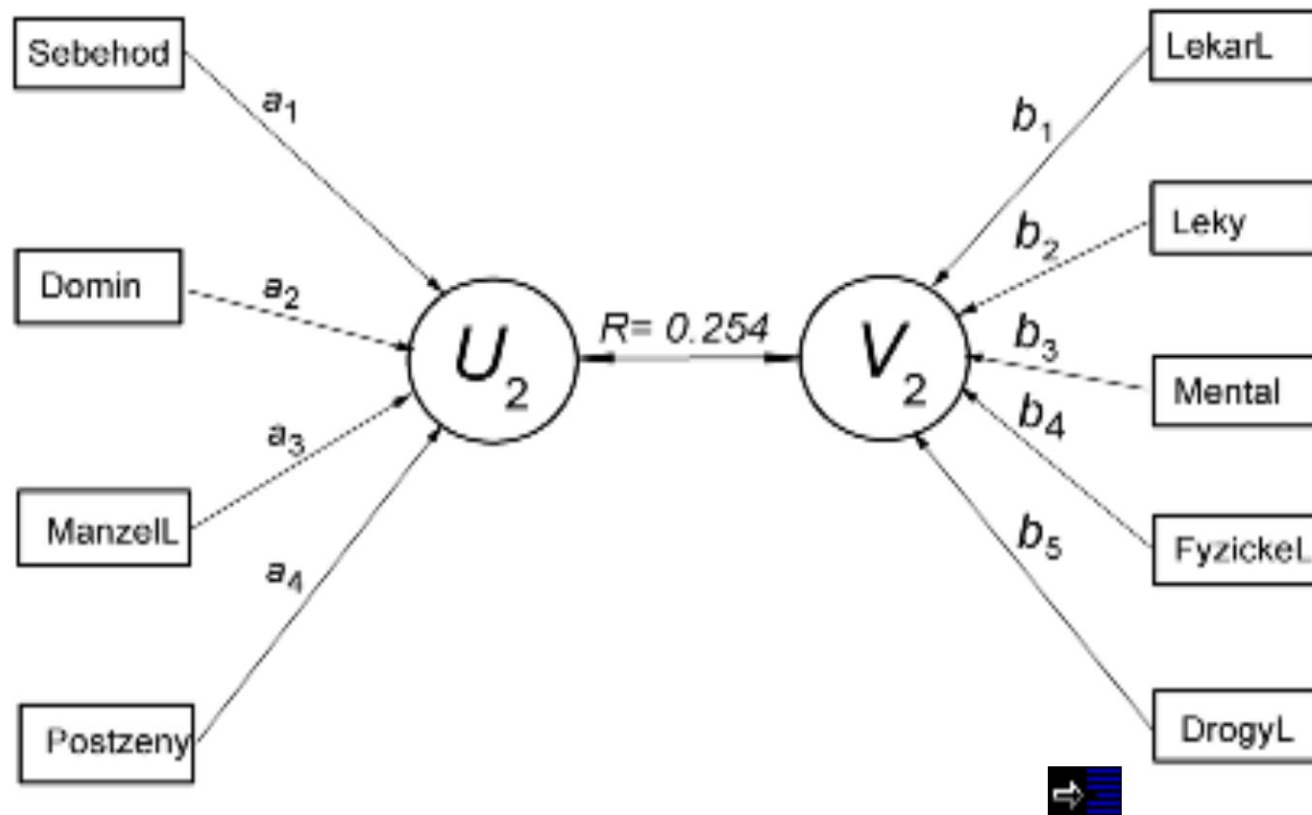
Obr. 6.13c



Obr. 6.13c První pár kanonických korelačních proměnných U_1 a V_1 dostatečně popisuje závislost 4 znaků: Sebehod, Domin, Manžell, Postženy na 5 znacích: LékařL, Léky, FyzickéL, Mental, DrogyL: zdrojové matice dat: CanonL (STATISTICA, ORIGIN).



Obr. 6.13d



Obr. 6.13d Druhý pár kanonických korelačních proměnných U_2 a V_2 dostatečně popisuje závislost 4 znaků Sebehod, Domin, Manžell, Postzeny na 5 znacích LékařL, Léky, FyzickéL, Mental, DrogyL zdrojové matice dat CanonL (STATISTICA, ORIGIN).