



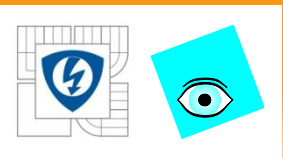
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

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.

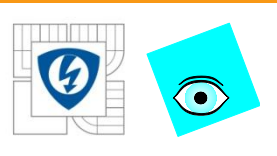


4.3 ODHALENÍ STRUKTURY VE ZNACÍCH A OBJEKTECH

25.2.2010

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



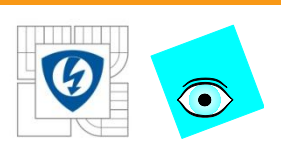


Odhalení struktury ve znacích a objektech

Zdrojová matice dat X má rozměr $n \times m$.

Nejprve je třeba provést ***průzkumovou (exploratorní) analýzu dat:***

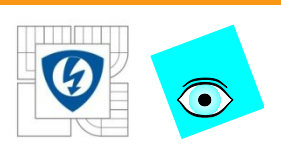
- a) posoudit *podobnost objektů* pomocí rozptylových a symbolových grafů,
- b) nalézt *vybočující objekty*, resp. jejich znaky,
- c) stanovit, zda lze použít předpoklad *lineárních vazeb*,
- d) ověřit *předpoklady o datech* (normalitu, nekorelovanost, homogenitu).



Techniky k určení vzájemných vazeb

se dále dělí zda hledají

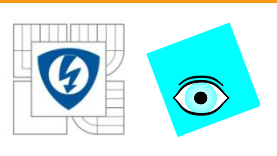
1. *strukturu a vazby ve znacích nebo*
2. *strukturu a vazby v objektech, což je:*
 - a) hledání struktury ve znacích v metrické škále
 - *faktorová analýza FA,*
 - *analýza hlavních komponent PCA a*
 - *shluková analýza.*
 - b) hledání struktury v objektech v metrické škále
 - *shluková analýza.*
 - c) hledání struktury v objektech v metrické i v nemetrické škále: *vícerozměrné školování.*
 - d) hledání struktury v objektech v nemetrické škále:
 - *korespondenční analýza.*



Techniky k určení vzájemných vazeb

Metody vícerozměrné statistické analýzy umožňuje zpracování ***lineárních vícerozměrných modelů***:

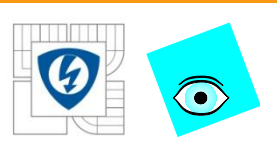
závisle proměnné se uvažují jako lineární kombinace nezávisle proměnných, resp. vazby mezi proměnnými jsou lineární.



Určení struktury a vazeb

Zdrojová matice má rozměr $n \times m$.

1. **Hledání struktury v *proměnných* v metrické škále:**
faktorová analýza a analýza hlavních komponent.
2. **Hledání struktury v *objektech* v metrické škále: *shluková analýza.***
3. **Hledání struktury v *objektech* v metrické i v nemetrické škále: *vícerozměrné škálování.***
4. **Hledání struktury v *objektech* v nemetrické škále: *korespondenční analýza.***



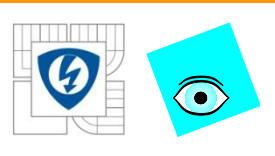
Určení struktury a vazeb

5. *Zpracování lineárních vícerozměrných modelů, kde závisle proměnné se*

1. uvažují jako lineární kombinace nezávisle proměnných resp. vazby mezi proměnnými jsou lineární.

6. *Exploratorní (průzkumová) analýza dat umožňuje*

- posoudit *podobnost objektů* pomocí vizuálních rozptylových diagramů a symbolových grafů,
- nalézt *vybočující objekty*, resp. jejich proměnné,
- stanovit, zda lze použít předpoklad *lineárních vazeb*,
- ověřit *předpoklady o datech*:
- (normalita, nekorelovanost, homogenita).



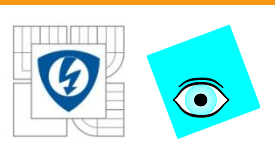
Určení struktury a vazeb

Objekty mohou být

- 1) shodné či různé
- 2) podobné či nepodobné
- 3) úměrné či neúměrné
- 4) směsi

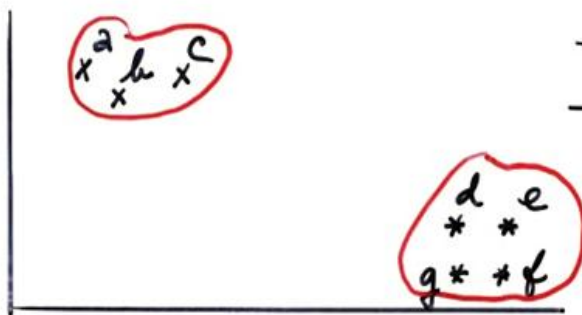
Proměnné mohou být

- 1) shodné či různé
- 2) podobné či nepodobné
- 3) úměrné či neúměrné
- 4) lineární kombinace



Rozličné analýzy vícerozměrných dat

(1)



- objekty a, b, c jsou stejné či podobné
- objekty d, e, f, g jsou stejné či podobné
ale nepodobné s objekty a, b, c

ANALÝZA SKLUPŮ
KLASIFIKAČNÍ ANALÝZA

(2) Objekty a, b, c, ... jsou směsí několika základních objektů i, j, k

$$a = \alpha \cdot i + \beta \cdot j + \gamma \cdot k, \text{ kde } \alpha, \beta, \gamma \text{ jsou koeficienty úměrnosti}$$

jedna proměnná je lineární kombinací ostatních proměnných

FAKTOROVÁ ANALÝZA (FA)
METODA HLAVNÍCH KOMPONENT (PCA)

lineární vztah mezi jednou proměnnou a všemi ostatními proměnnými

KORELAČNÍ ANALÝZA
KOVARIANČNÍ ANALÝZA

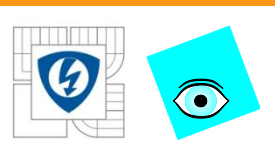
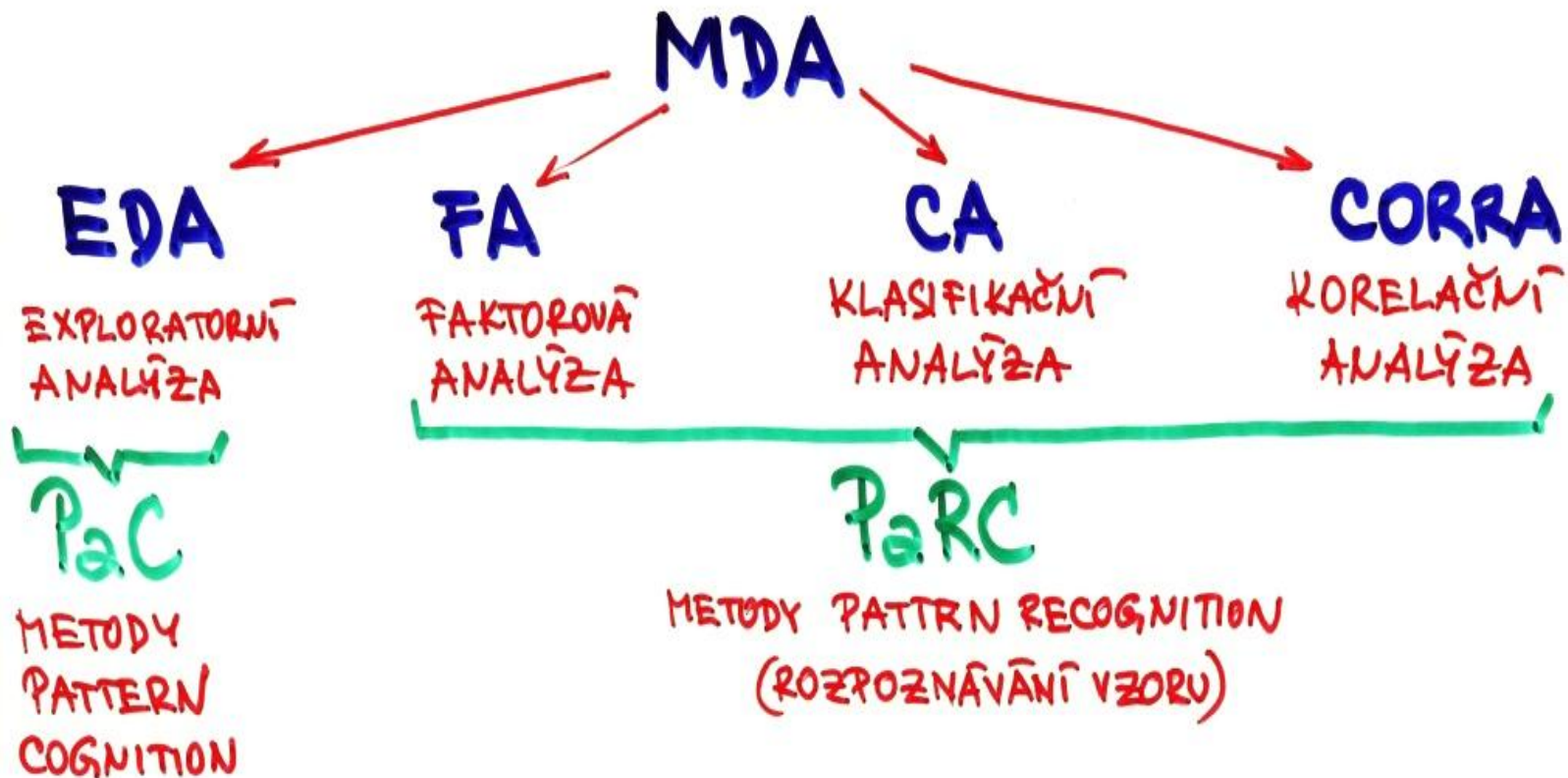
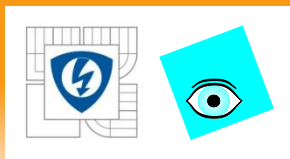


Schéma analýzy vícerozměrných dat



SCHEMA ANALÝZY VÍCEROZMĚRNÝCH DAT

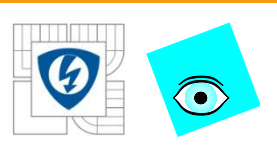


EDA (exploratorní průzkumová analýza dat)

- 1) Znázornění informace v souboru dat
- 2) Detekce anomálií a chyb
- 3) Shlukování objektů a proměnných
- 4) Korelování podobnosti mezi objekty s několika znaky
- 5) Výběr užitečných proměnných

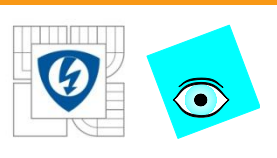
FA (faktorová analýza)

- 1) Redukce šumu naměřených dat
- 2) Detekce základní struktury dat
- 3) Lineární vztah mezi komplexním objektem a a čistými objekty i, j, k , dle vztahu $a = \alpha i + \beta j + \gamma k + \dots$



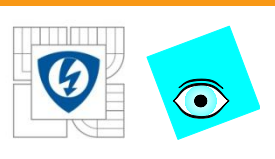
CA (klasifikační analýza)

1. **Klasifikuje** objekty do jedné z několika kategorií určených analýzou shluků či kategoriemi fyzikálního významu. Verifikuje stupeň separace, vzdálenosti mezi objekty.
 2. **Předpovídá** (=predikuje) objekty, jejichž kategorie je neznámá.
 3. **Detekuje** objekty, jež nepatří do žádné z předešlých kategorií (=2. úroveň PaRC)
1. a 2. se nazývá 1. úroveň PaRC: rozeznává, rozpoznává objekty dle jejich původu, povahy, kategorií.
1. **úroveň PaRC**: staví hranice mezi kategoriemi
 2. **úroveň PaRC**: staví hranice mezi jednou kategorií a zbývajícím prostorem.

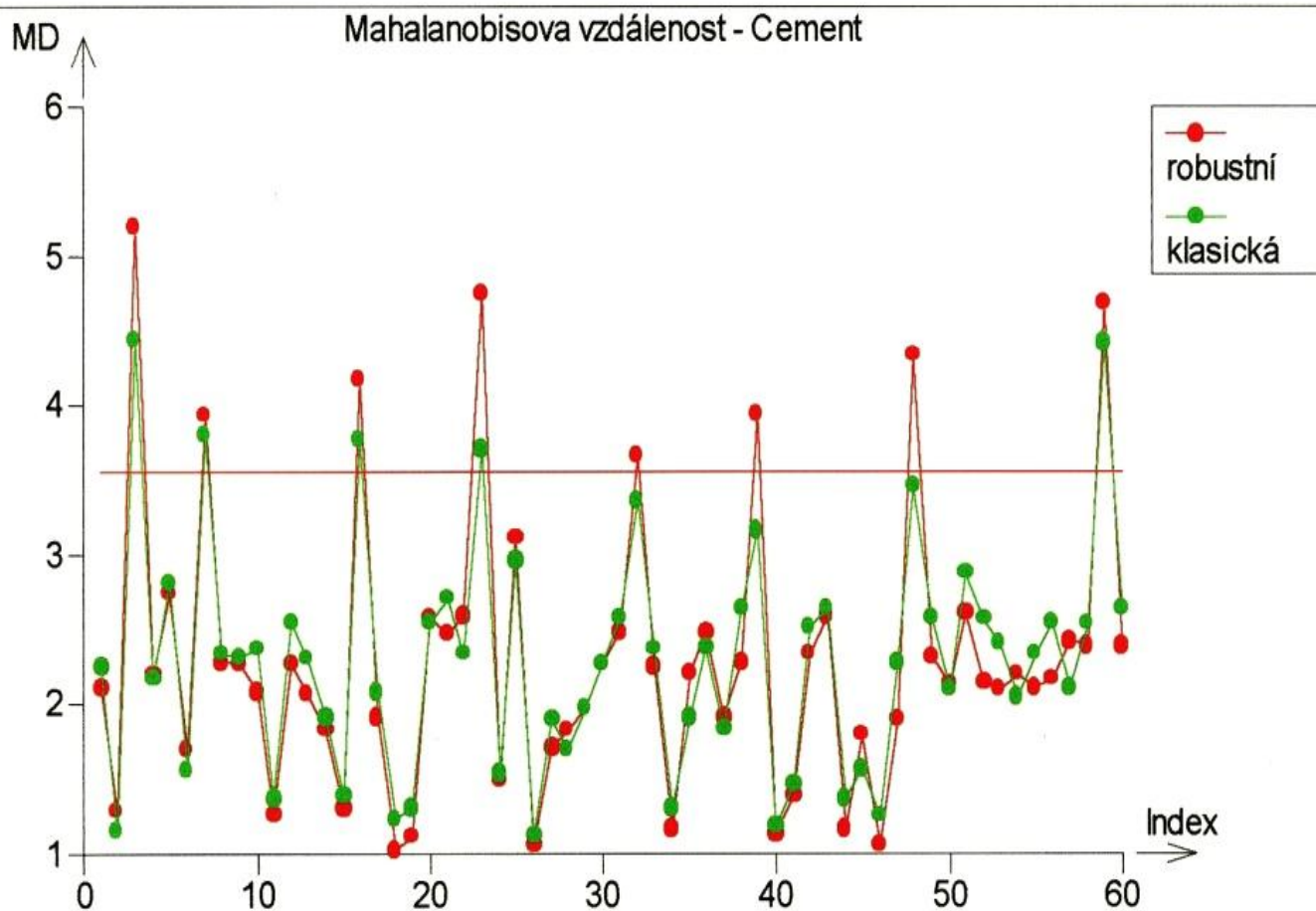


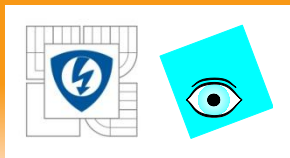
CORRA (korelační analýza)

1. **Objasňuje** hodnoty jedné proměnné pomocí hodnot ostatním proměnných (=3. úroveň PaRC)
2. **Detekuje** vztah mezi jedním blokem proměnných a jiným blokem proměnných (=4. úroveň PaRC)



Mahalanobisova vzdálenost - Cement





PŘÍKLAD 3.1. Zobrazování znaků zdrojové matice dat Lidé

Rozptylovým diagramem je třeba nalézt trendy a vnitřní párové vztahy mezi znaky zdrojové matice *Lidé* a odhalit korelaci dvou znaků. **Který pár vykazuje nejsilnější korelaci tj. nejvyšší hodnotu r .**

Data: použity UNSCRAMBLER, SCAN.

Řešení:

a) 2D- a 3D-rozptylové diagramy: Objekty jsou indikovány pořadovými čísly (obr. 3.1a) a nebo svými jmény (obr. 3.1b):

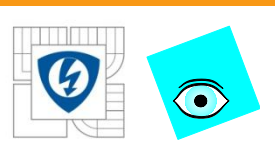
Pár znaku **výška** a **hmotnost** vykazuje $r = 0.96$ a silnou lineární závislost.

Muži (M) jsou vyšší a hmotnější než **ženy (F)**,

Objekt 1 je nejtěžší a nejvyšší osoba,

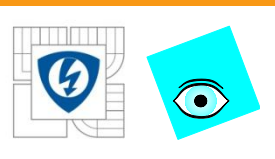
Objekty 12 a 21 jsou umístěny na opačném konci proložené regresní přímky a týkají se nejmenších a nejlehčích osob.

Vedle grafu jsou i **popisné statistiky** v rámečku jako jsou směrnice (**Slope**), úsek (**Offset**), r (**Correlation**), atd.

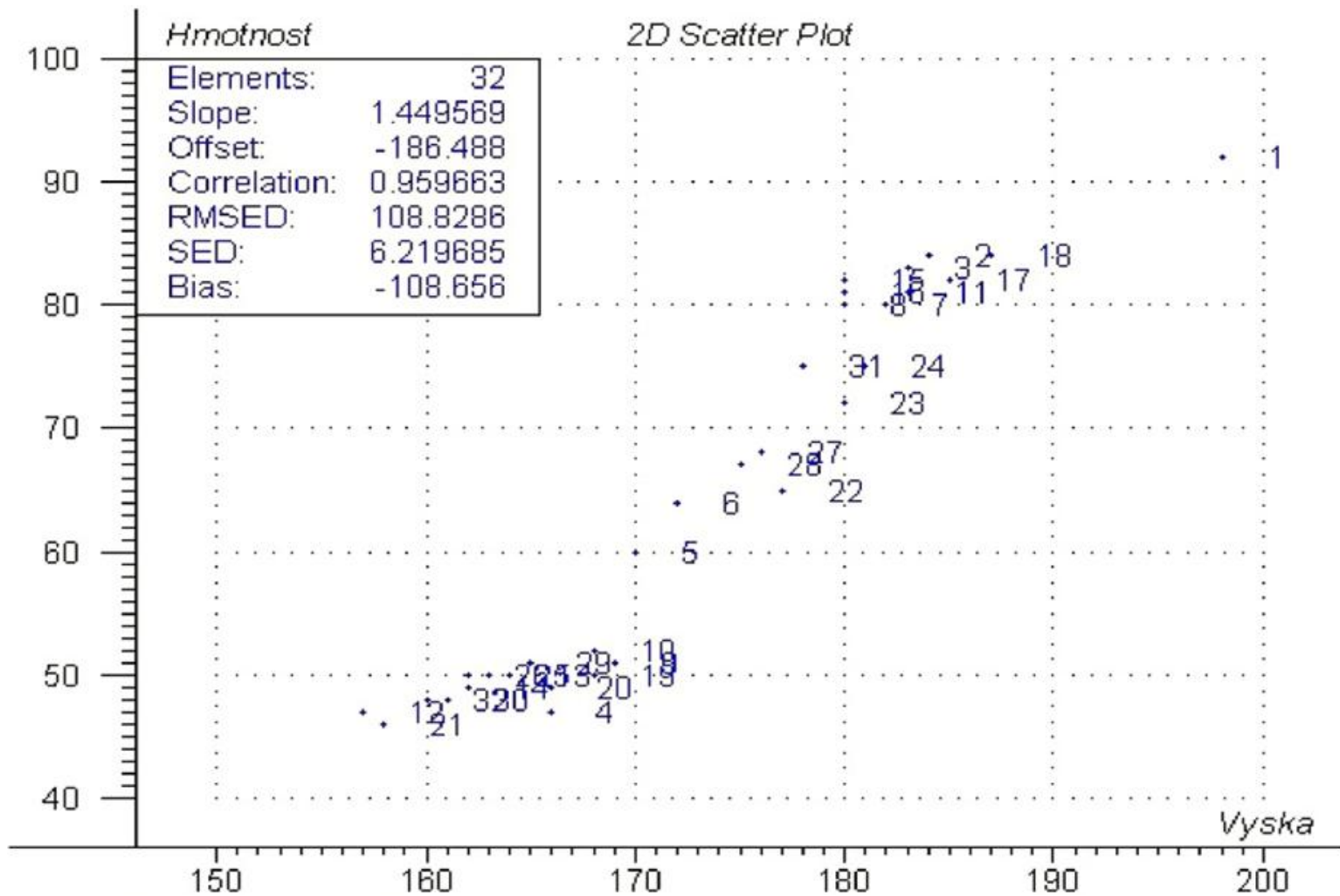


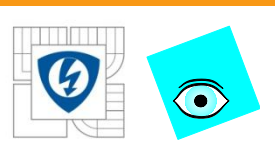
Úprava zdrojové matice dat: Označení osob vložením jejich indexů, např. 1MA, 2MA, atd.

	1	2	3	4	5	6	7	8	9	10	11	12	13
	Osoba	Vyska	Hmotnost	Vlasy	Boty	Vek	Prijem	Pivo	Vino	Sex	Plavani	Puvod	IQ
1MA	1MA	198	92	-1	48	48	45000	420	115	-1	98	-1	100
2MA	2MA	184	84	-1	44	33	33000	350	102	-1	92	-1	130
3MA	3MA	183	83	-1	44	37	34000	320	98	-1	91	-1	127
1FA	1FA	166	47	-1	36	32	28000	270	78	1	75	-1	112
2FA	2FA	170	60	1	38	23	20000	312	99	1	81	-1	110
3FA	3FA	172	64	1	39	24	22000	308	91	1	82	-1	102
4MA	4MA	182	80	-1	42	35	30000	398	65	-1	85	-1	140
5MA	5MA	180	80	-1	43	36	30000	388	63	-1	84	-1	129
4FA	4FA	169	51	1	36	24	23000	250	89	1	78	-1	98
5FA	5FA	168	52	1	37	27	23500	260	86	1	78	-1	100
6MA	6MA	183	81	-1	42	37	35000	345	45	-1	90	-1	105
6FA	6FA	157	47	1	36	32	32000	235	92	1	70	-1	127
7FA	7FA	164	50	1	38	41	34000	255	134	1	76	-1	102
8FA	8FA	162	49	1	37	40	34000	265	124	1	75	-1	108
7MA	7MA	180	82	-1	44	43	37000	355	82	-1	88	-1	109
8MA	8MA	180	81	-1	44	46	42000	362	90	-1	86	-1	113
1MB	1MB	185	82	-1	45	26	16000	295	180	-1	92	1	109
2MB	2MB	187	84	-1	46	27	16500	299	178	-1	95	1	119
1FB	1FB	168	50	1	37	49	34000	170	162	1	76	1	135
2FB	2FB	166	49	1	36	21	14000	150	245	1	75	1	123
3FB	3FB	158	46	1	34	30	18000	120	120	1	70	1	119
3MB	3MB	177	65	-1	41	26	18000	209	160	-1	86	1	120
4MB	4MB	180	72	-1	43	33	19000	236	175	-1	85	1	115
5MB	5MB	181	75	-1	43	42	31000	198	161	-1	83	1	105
4FB	4FB	163	50	1	36	18	11000	143	136	1	75	1	102
5FB	5FB	162	50	1	36	20	11500	133	146	1	74	1	132
6MB	6MB	176	68	-1	42	50	36000	195	177	-1	82	1	96
7MB	7MB	175	67	1	42	55	38000	185	187	-1	80	1	105
6FB	6FB	165	51	1	36	36	26000	121	129	1	76	1	126
7FB	7FB	161	48	1	35	41	31500	116	196	1	75	1	120
8MB	8MB	178	75	-1	42	30	24000	203	208	-1	81	1	118
8FB	8FB	160	48	1	35	40	31000	118	198	1	74	1	129

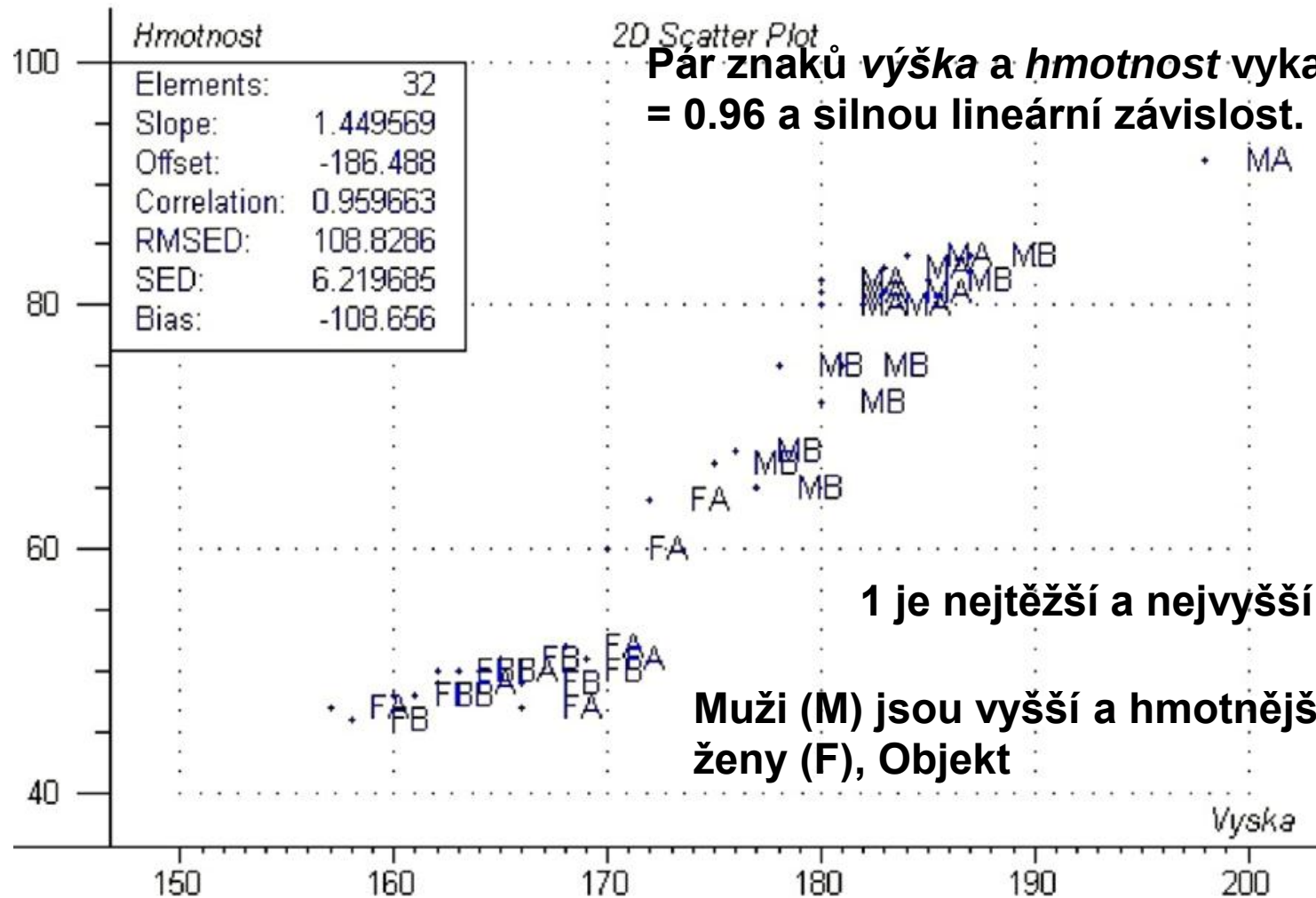


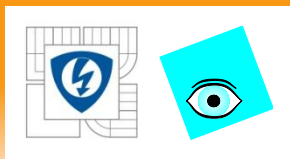
Obr. 3.1a Rozptylový diagram znaků Hmotnost a Výška s objekty označenými svými pořadovými čísly, dat Lidé.





Obr. 3.1b Rozptylový diagram znaků Hmotnost a Výška s objekty označenými svými jmény dat Lidé.



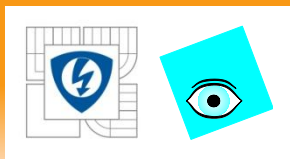


Interpretace a komentáře 2D-diagramů:

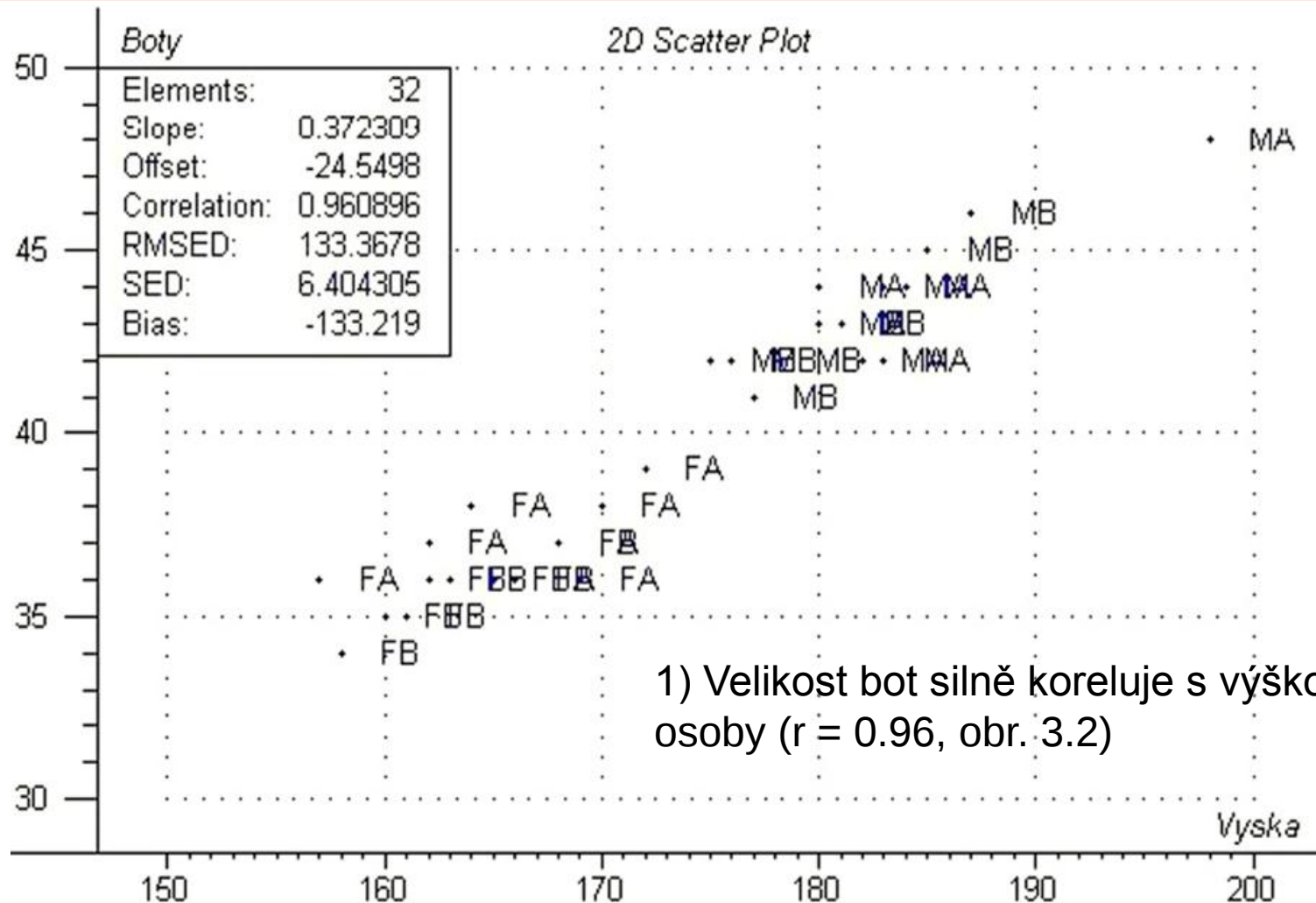
1. Velikost bot silně koreluje s výškou osoby ($r = 0.96$, obr. 3.2),
2. Věk téměř nekoreluje s výškou osoby ($r = 0.24$, obr. 3.3),
3. Příjem rovněž téměř nesouvisí čili nekoreluje s výškou osoby ($r = 0.30$, obr. 3.4),
4. Schopnost plavat souvisí s výškou osoby ($r = 0.96$, obr. 3.5), děti uplavou testovaných 500 m za delší čas než dospělí,
5. Pivo středně silně koreluje s výškou osoby ($r = 0.71$, obr. 3.6),
6. Víno už vůbec nekoreluje s výškou osoby ($r = -0.13$, obr. 3.7),
7. Víno a pivo spolu korelují negativně ($r = -0.65$, obr. 3.8), totiž čím více kdo pije víno tím méně pak pije pivo a naopak.

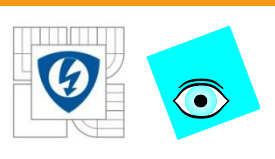
K zobrazení všech kombinací m znaků by se jednalo o $m(m - 1)/2$ kombinací.

To by bylo pro nás značně nepohodlné, a proto budeme hledat nějaký schůdnější způsob.

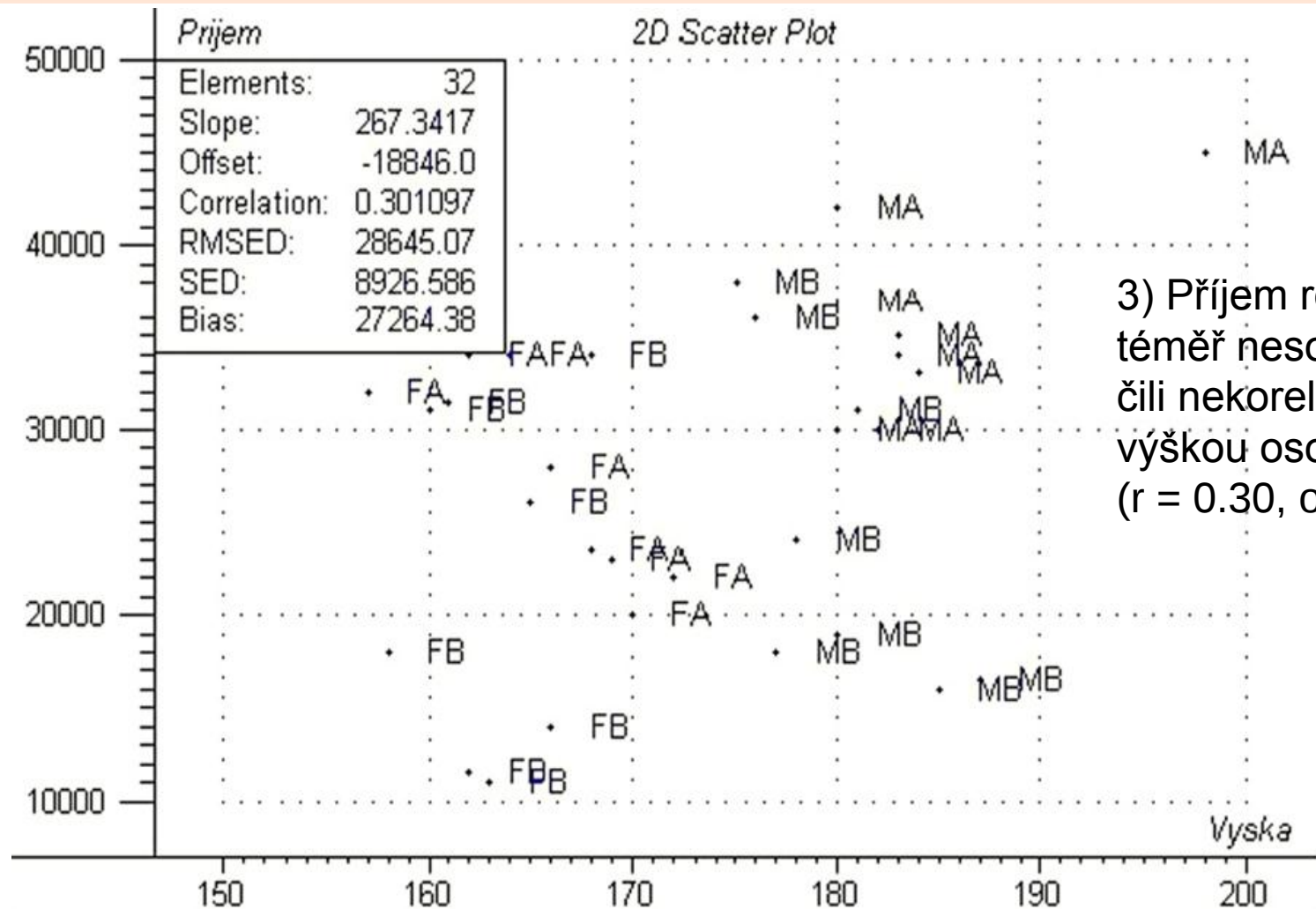


Obr. 3.2 Rozptylový diagram znaku Boty a Výška dat Lide, (UNSCRAMBLER).

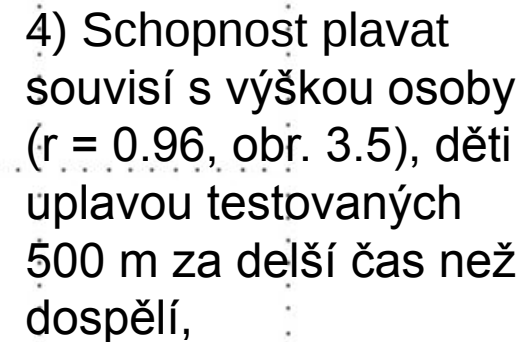


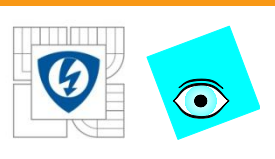


Obr. 3.4 Rozptylový diagram znaků Příjem a Výška dat Lidé, (XJNSCRAMBLER).

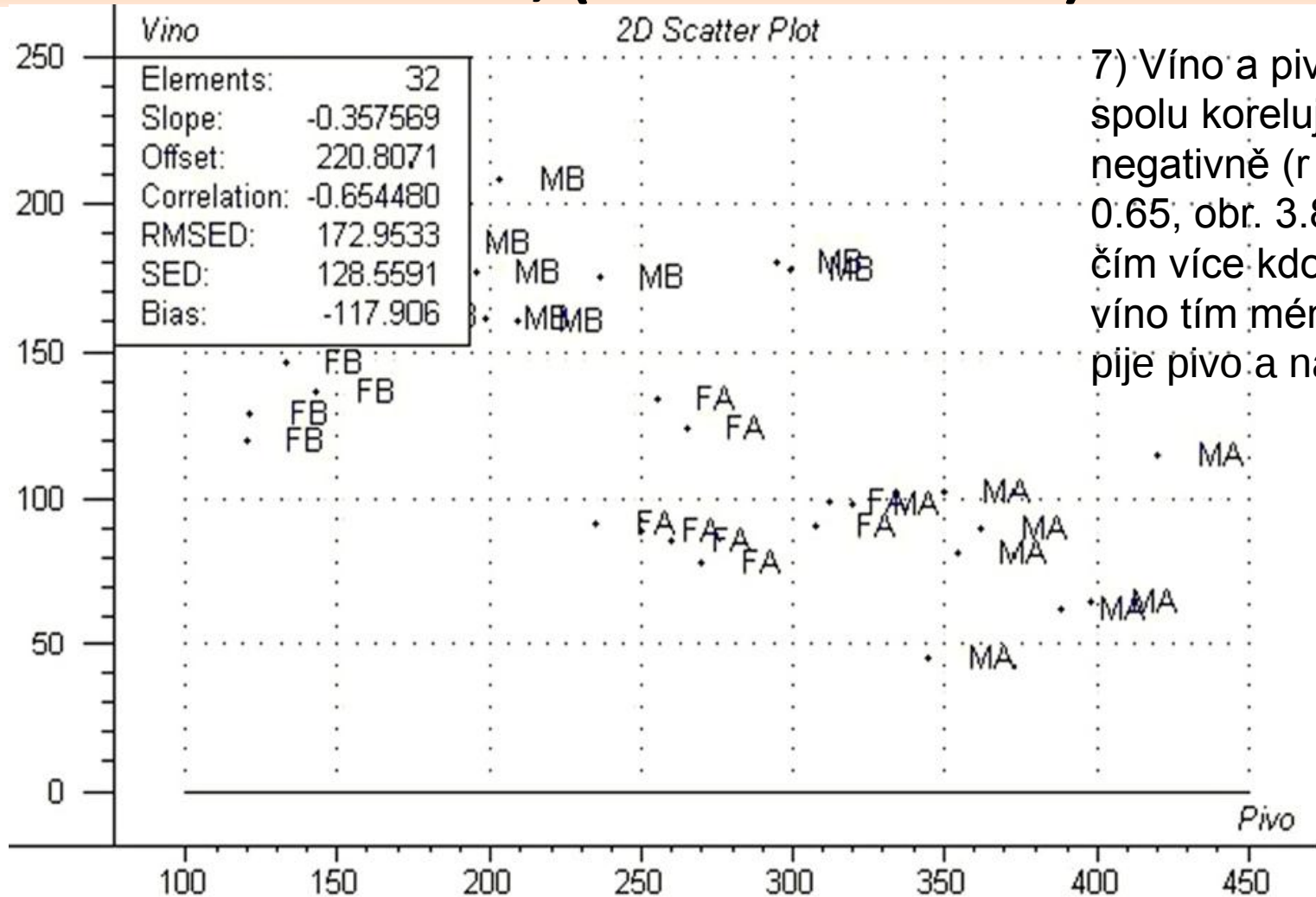


3) Příjem rovněž téměř nesouvisí čili nekoreluje s výškou osoby ($r = 0.30$, obr. 3.4),

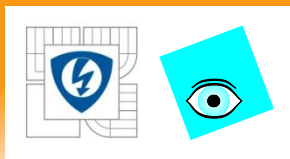




Obr. 3.8 Rozptylový diagram znaků Vín a Pivo dat Lidé, (UNSCRAMBLER).

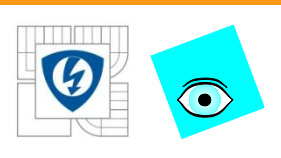


7) Víno a pivo spolu korelují negativně ($r = -0.65$, obr. 3.8), totiž čím více kdo pije víno tím méně pak pije pivo a naopak.



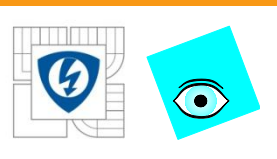
Interpretace a komentáře 3D-diagramů:

- 1) U trojice znaků **výška-hmotnost-velikost boty** (obr. 3.9) je zřejmá *silná pozitivní korelace* všech tří znaků.
- 2) U trojice znaků **výška-hmotnost-schopnost plavat** (obr. 3.10) je zřejmá *silná pozitivní korelace* všech tří znaků.
- 3) Naopak u trojice znaků **hmotnost-pivo-víno** (obr. 3.11) je vidět spíše *slabou negativní korelaci*.
- 4) U trojice **věk-pivo-víno** (obr. 3.12) je zřejmá spíše *slabou negativní korelaci*.
- 5) U trojice znaků **hmotnost-pivo-schopnost plavání** (obr. 3.13) je vidět středně *silná pozitivní korelace*.
- 6) Zatímco trojice **věk-víno-schopnost plavání** ukazuje, že znaky spolu téměř nekorelují.

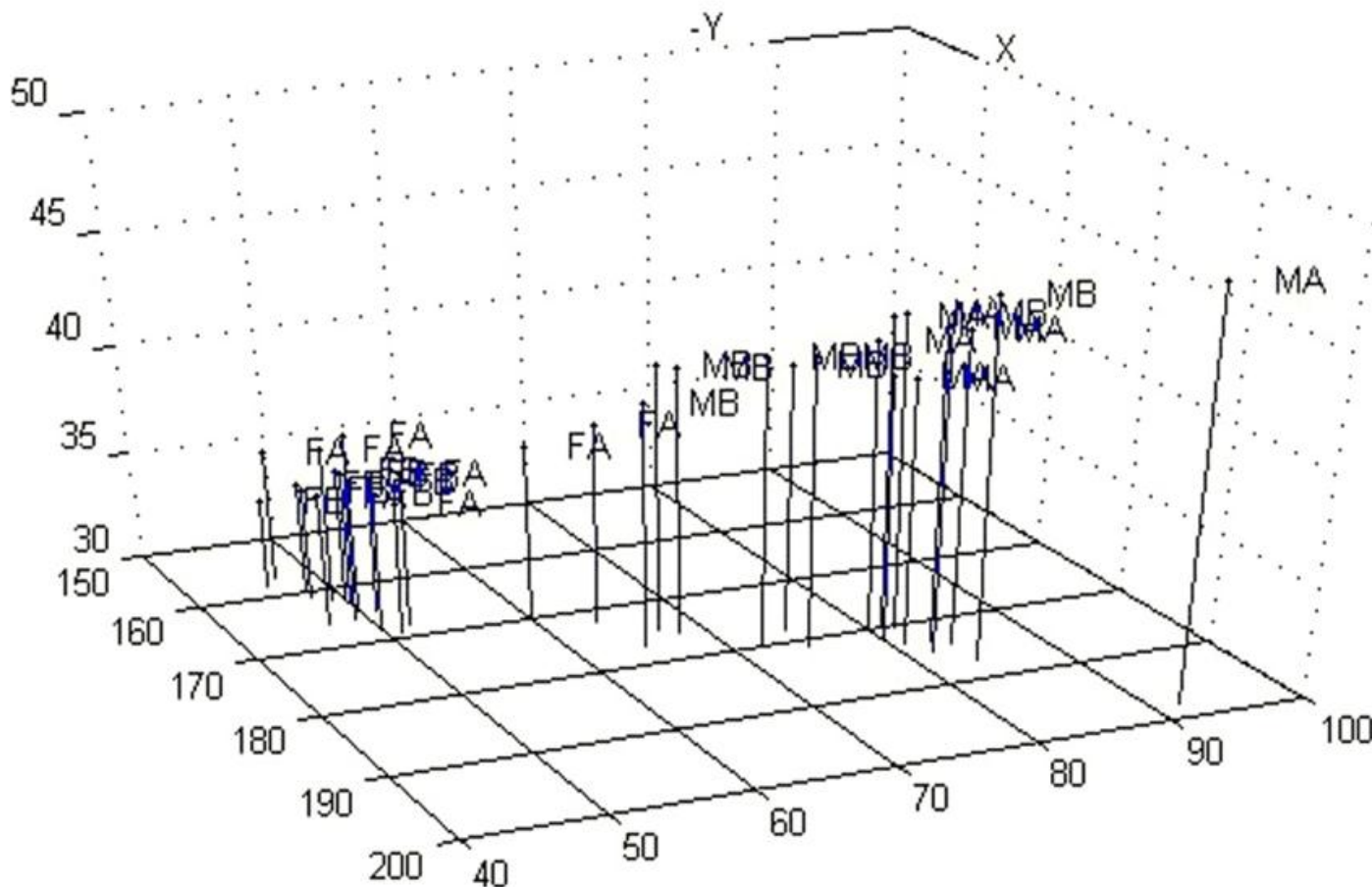


Pak snadno odpovíme na otázky:

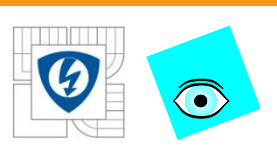
- 1) Je znak schopnost plavat korelovaný se znakem **výška** nebo se znakem **hmotnost**?
- 2) Liší se **muži** od **žen** z hlediska **schopnosti plavání**?
- 3) Existuje v těchto diagramech nějaký *shluk* osob?
- 4) Které osoby se nejvíce *odlišují* od ostatních?
- 5) Které osoby jsou si *podobné* a které jsou si naopak zcela *nepodobné*?
- 6) Dají se najít i jiné kombinace tří znaků, mezi kterými existuje obdobně *silná korelace*?



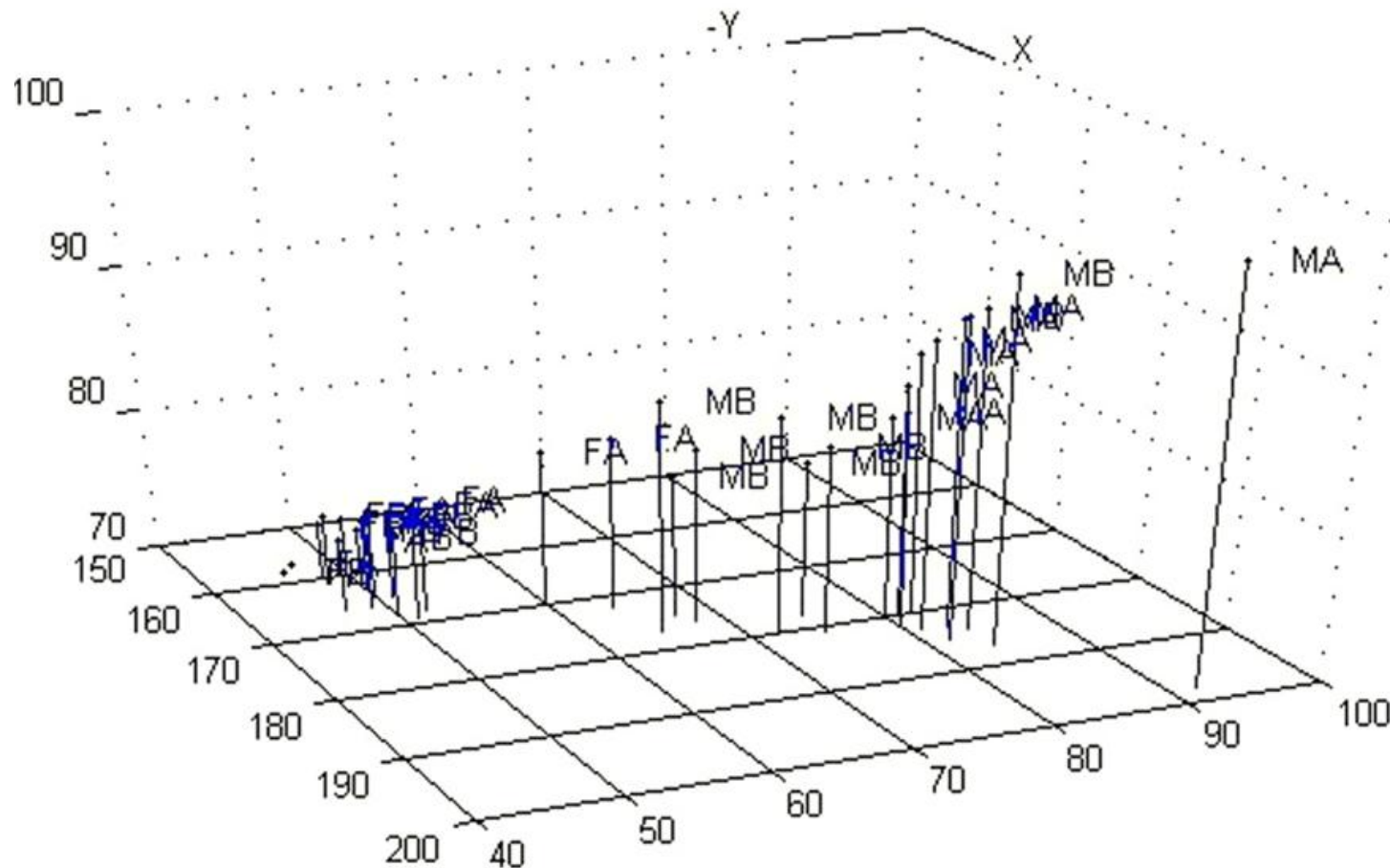
Obr. 3.9 Rozptylový 3D–diagram znaků Výška-Hmotnost-Boty dat Lidé



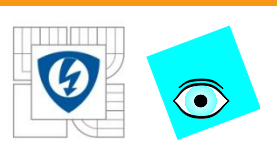
3D Scatter Plot
1) U trojice
znaků **výška-
hmotnost-
velikost boty**
(obr. 3.9)
je zřejmá *silná
pozitivní
korelace* všech
tří znaků.



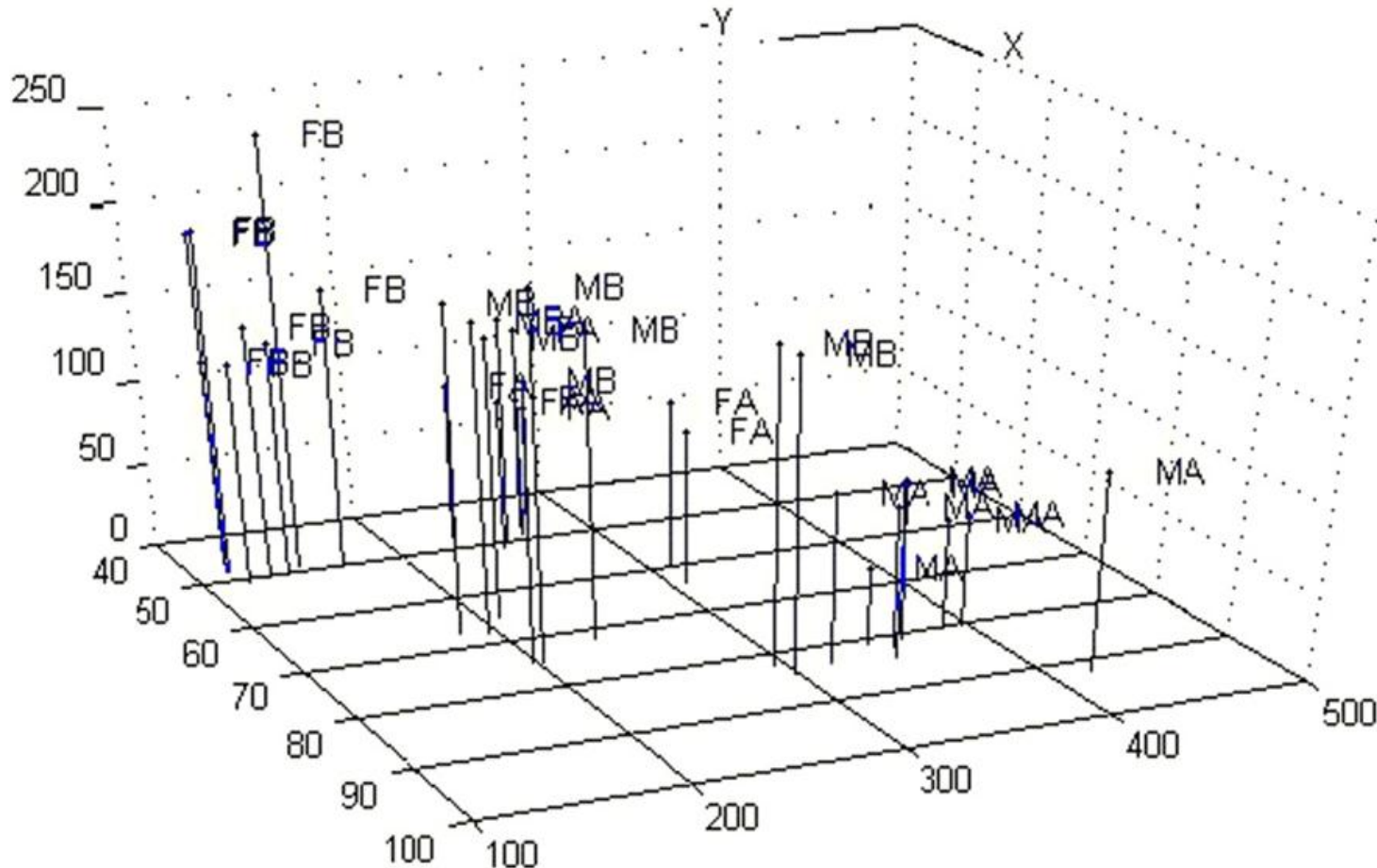
Obr. 3.10 Rozptylový 3D–diagram znaků Výška-Hmotnost-Plavání dat Lidé



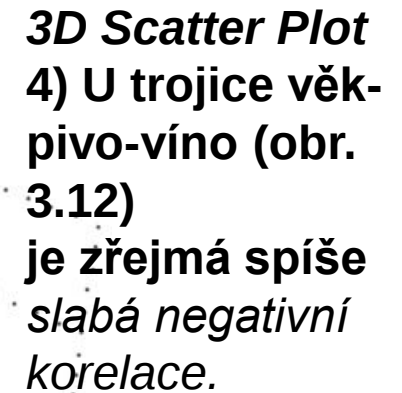
3D Scatter Plot
2) U trojice znaků **výška-hmotnost-schopnost plavat** (obr. 3.10 je zřejmá silná pozitivní korelace všech tří znaků.

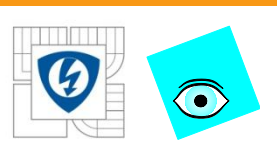


Obr. 3.11 Rozptylový 3D–diagram znaků Hmotnost-Pivo-Víno dat Lidé

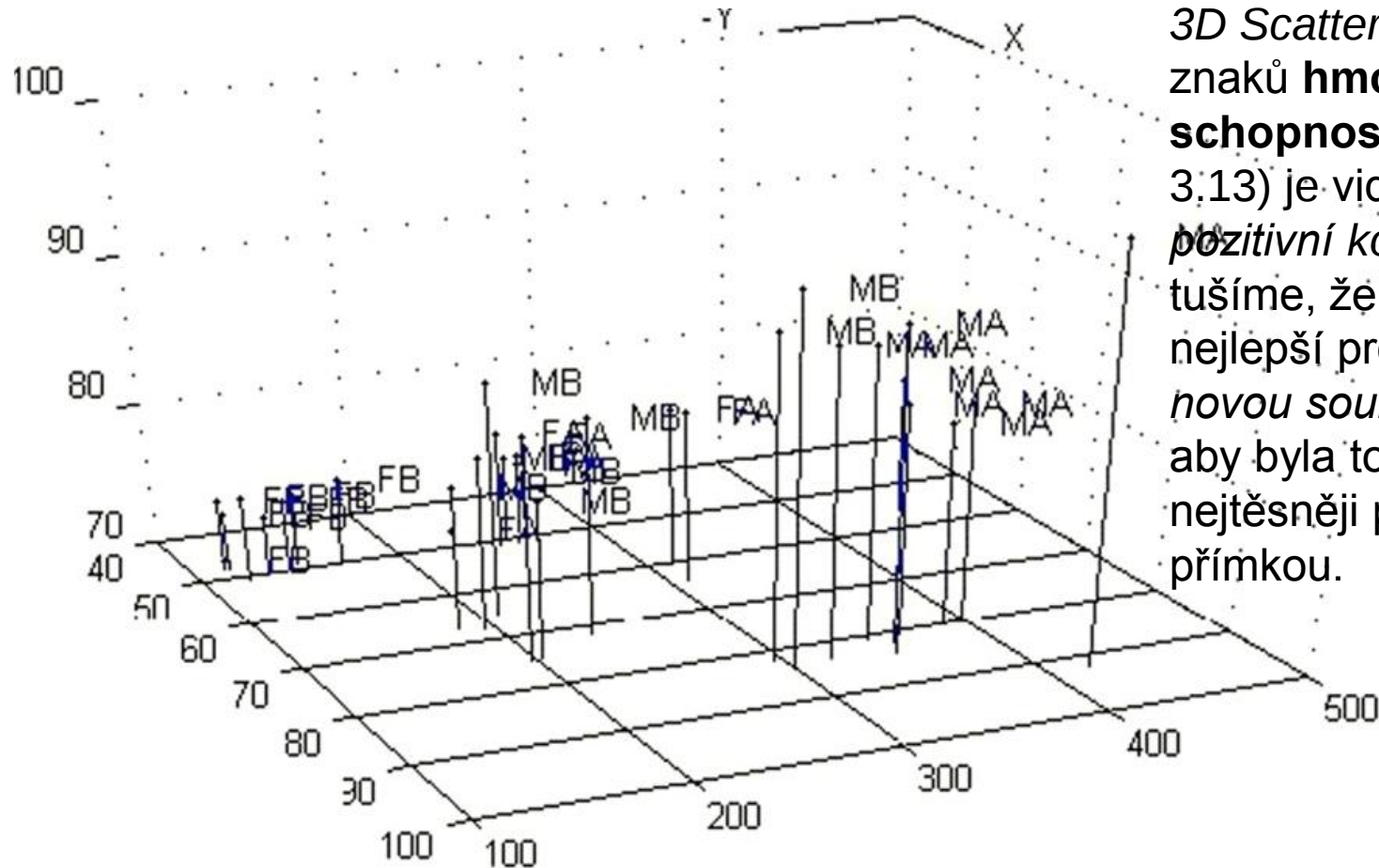


3D Scatter Plot
3) Naopak u
trojice znaků
hmotnost-pivo-
víno (obr 3.11 ,
je vidět spíše
slabou
negativní
korelaci

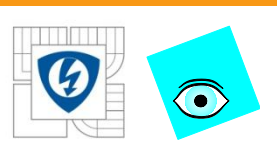




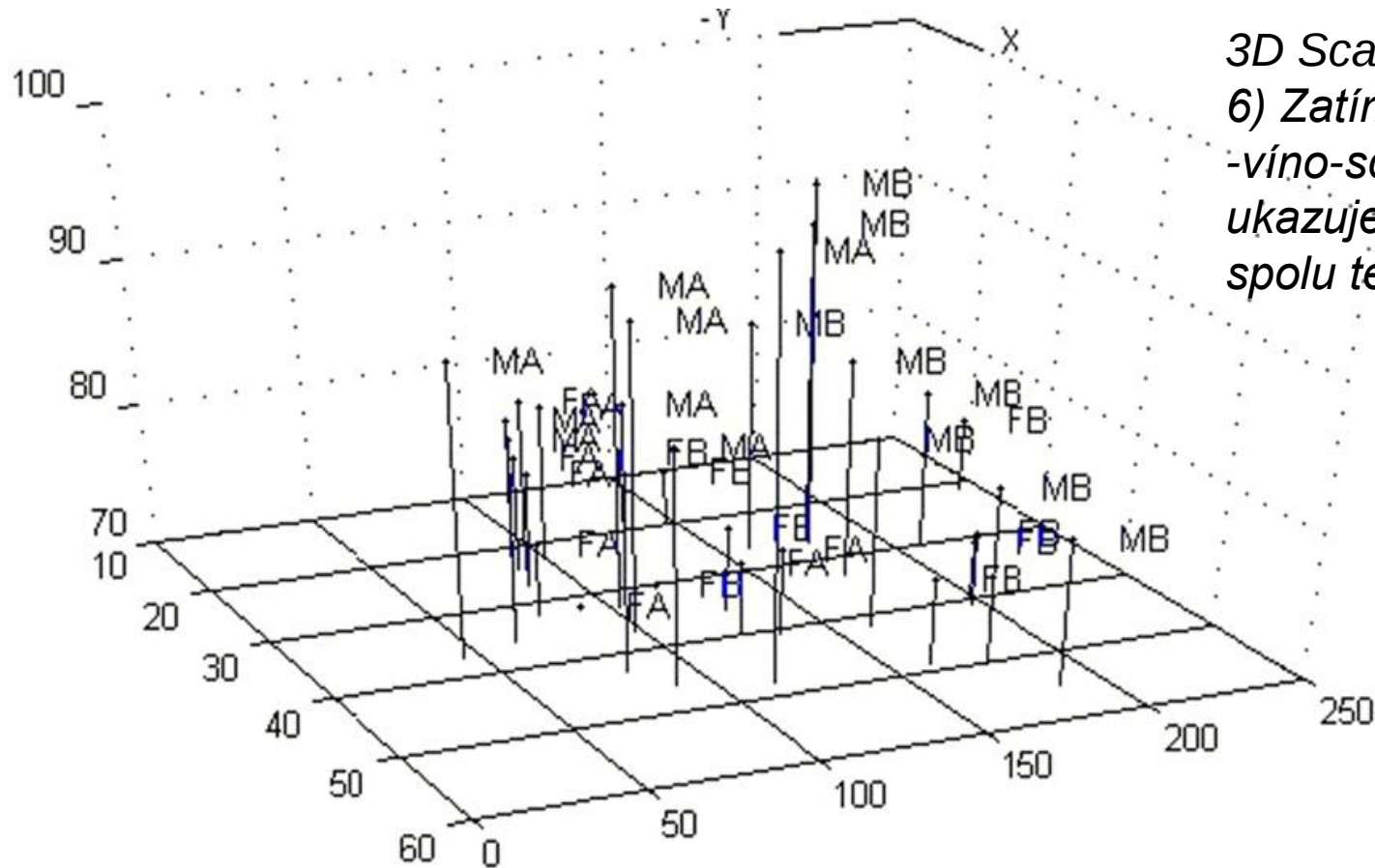
Obr. 3.13 Rozptylový 3D-diagram znaků Hmotnost-Pivo-Plavání dat Lidé



3D Scatter Plot 5) U trojice znaků **hmotnost-pivo-schopnost plavání** (obr. 3.13) je vidět středně *silná pozitivní korelace*. Intuitivně tušíme, že na obr. 3.13 bylo nejlepší proložit objekty *novou souřadnou osu* tak, aby byla totožná s nejtěsněji prokládanou přímkou.

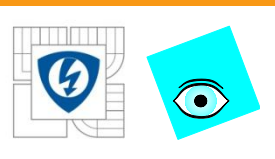


Obr. 3.14 Rozptylový 3D-diagram znaků Věk-Víno-Plavání dat Lidé



3D Scatter Plot

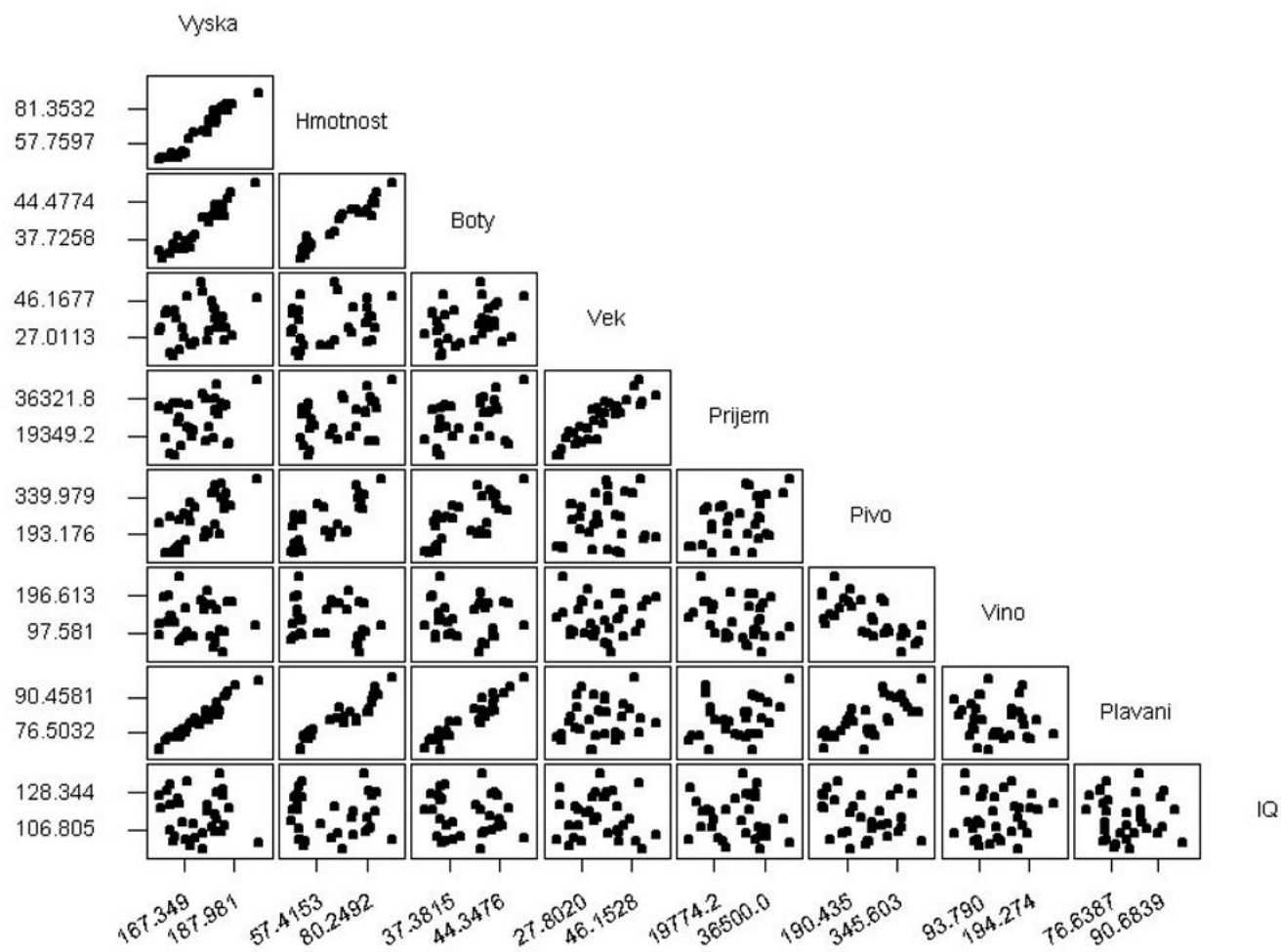
6) Zatímco trojice věk
-víno-schopnost plavání
ukazuje, že znaky
spolu téměř nekorelují



Tabulka 3.1 Matice korelačních koeficientů r znaků zdrojové matice dat Lidé.

Pod korelačním koeficientem je vypočtená hladina významnosti p . Je-li $p < 0.05$, je r statisticky významný a v tabulce vyznačeno **tučně**.

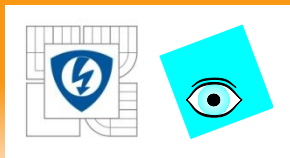
	<i>Vyska</i>	<i>Hmotn.</i>	<i>Boty</i>	<i>Vek</i>	<i>Prijem</i>	<i>Pivo</i>	<i>Vino</i>	<i>Play.</i>
<i>Hmotn.</i>	0.96 0							
<i>Boty</i>	0.961 0	0.969 0						
<i>Vek</i>	0.241 0.184	0.254 0.16	0.323 0.071					
<i>Prijem</i>	0.301 0.094	0.335 0.061	0.354 0.047	0.885 0				
<i>Pivo</i>	0.715 0	0.738 0	0.697 0	0.128 0.484	0.417 0.018			
<i>Vino</i>	-0.138 0.451	-0.197 0.28	-0.09 0.629	0.027 0.884	-0.297 0.099	-0.654 0		
<i>Plavani</i>	0.962 0	0.937 0	0.933 0	0.158 0.387	0.252 0.165	0.725 0	-0.166 0.365	
<i>IQ</i>	-0.122 0.507	-0.03 0.854	-0.12 0.512	-0.08 0.673	-0.107 0.561	-0.107 0.559	0.068 0.713	-0.116 0.526



25.2.2010

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





Směry maximálního rozptylu v rozptylovém diagramu:

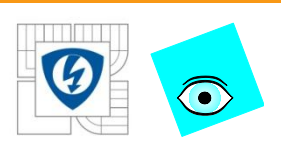
Intuitivně tušíme, že na obr. 3.13 bylo nejlepší proložit objekty *novou souřadnou osu* tak, aby byla totožná s nejtěsněji prokládanou přímkou.

Vhodný počet potřebných os není 3 ale pouze 1.

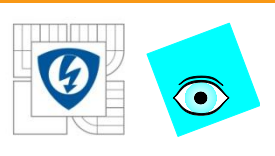
Je to proto, že znaky x_1 , x_2 a x_3 spolu silně korelují a objekty leží na prostorové přímce.

Nejtěsněji objekty prokládaná přímka, a tím pádem i **nová souřadnicová osa** zvaná **PC 1** leží ve směru *největšího rozptylu objektů*.

Souřadnicová osa **PCX** **není** obecně přitom paralelní s žádnou z původních os znaků x_1 , x_2 a x_3 .

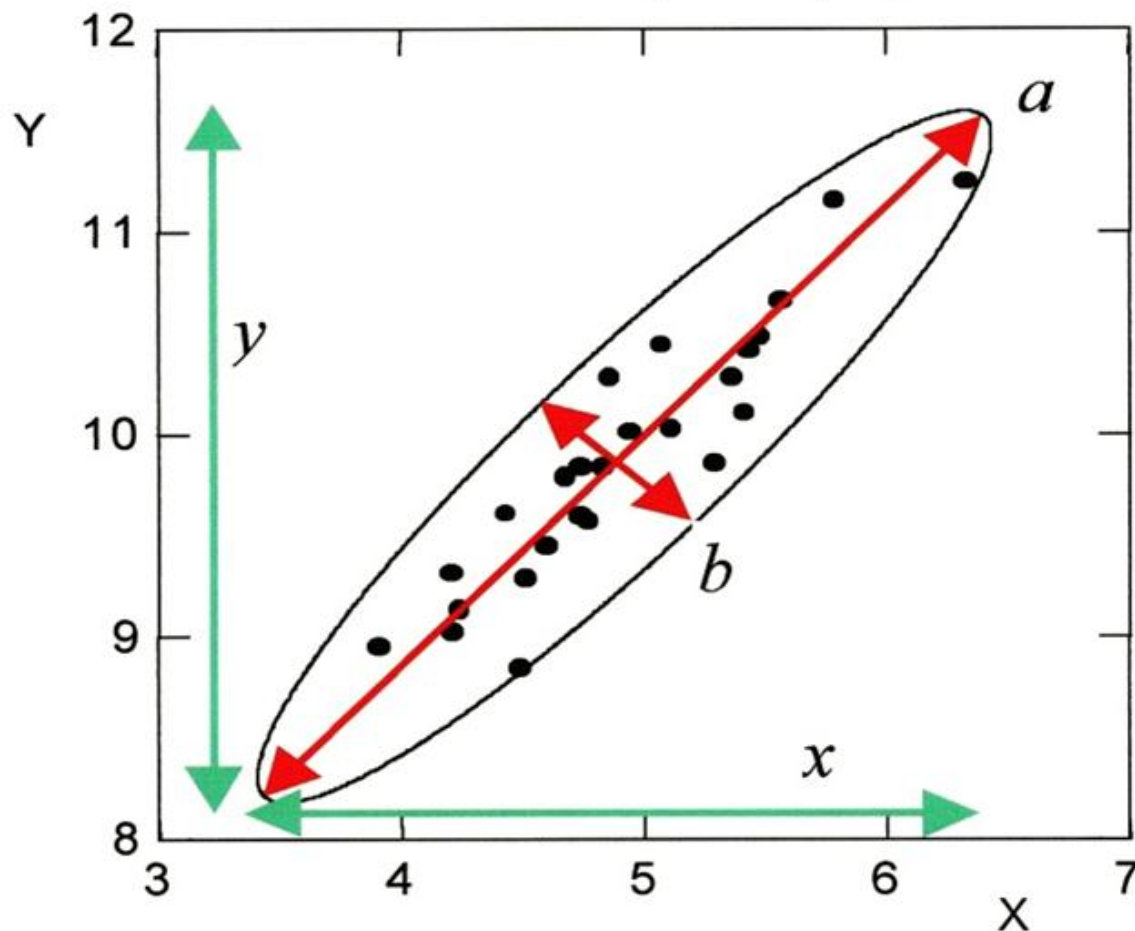


Novou souřadnou osu PCX nazveme ***první hlavní komponentou***.
Bude ležet *ve směru maximálního rozptylu* zdrojové matice dat X .
S touto osou je spjata jakási nová latentní proměnná.
Metoda hlavních komponent PCA poskytne nejenom tuto první komponentu ale také další hlavní komponenty.
Obvykle jedna, dvě, maximálně však tři nové souřadnicové osy $PC1$, $PC2$ a $PC3$ vystihnou dohromady největší podíl proměnlivosti dat.

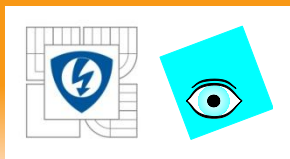


Princip hlavních komponent:

Hlavní komponenty a , b



Pravoúhlým promítnutím dat na osu a získáme nejvíce informace o původních datech na jediné přímce.



První hlavní komponenta proložením nejmenšími čtverci

Je třeba proložit přímkou v prostoru x_1 , x_2 a x_3 umístěnými objekty.

Jak?

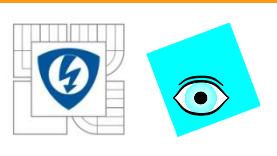
Z každého objektu spustíme k této přímce kolmici
a obdržíme e_i reziduum.

Nejtěsněji proloženou přímkou dostaneme, když suma čtverců reziduí bude dosahovat své minimální velikosti

$$RSS = \sum_{i=1}^n e_i^2 = \text{minimum}$$

Existují dvě kritéria ke konstrukci *první hlavní komponenty*:

1. Směr největšího rozptylu,
2. Metoda nejmenších čtverců MNC.



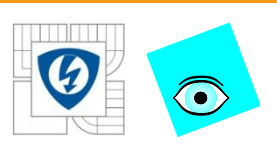
Rozšíření k dalším hlavním komponentám

Může nastat případ, že první hlavní komponenta **PC₁** nestačí dostatečně popsat rozptýlení objektů.

Druhá hlavní komponenta PC₂ je ortogonální vůči první (*je na ni kolmá*) a je umístěna do směru druhého největšího rozptýlení objektů v rovině.

U elipsoidu objektů jsou do jeho prostorových os umístěny tři osy největšího rozptýlení objektů čili osy tří hlavních komponent **PC₁**, **PC₂** a **PC₃**.

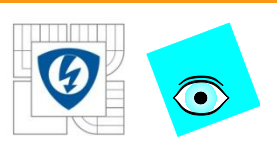
Nové tři osy nekorelují, jsou vzájemně ortogonální tj. **na sebe kolmé**, PCA umožňuje projekci zobecnit na libovolný počet m znaků.



Závěr:

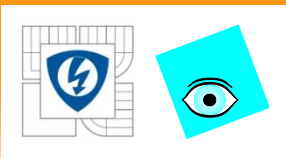
Vyšetřili jsme matici dat v 2D- a 3D-rozptylových diagramech a získali řadu závěrů:

- čím vyšší jsou osoby, tím jsou těžší,
- nebo schopnost plavat koreluje s výškou a hmotností osob.
- existují minimálně dva shluky, jeden pro muže a jeden pro ženy.
- některé znaky ukázaly shluky odpovídající původu osob.



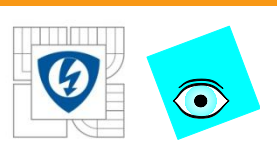
Závěr:

- některé znaky ukázaly shluky odpovídající původu osob.
- některé korelace mohou být však typicky falešné, například plavání proti výšce.
- diskrétní znaky jdou vyšetřit velmi obtížně a zobrazit je nelze.
- i přes veškerou názornost 2D- a 3D-rozptylových diagramů zobrazení všech 9 znaků je časově náročné a pracné.
- budeme proto hledat výhodnější postup především v PCA.



Metoda hlavních komponent – použití modelů s latentními proměnnými

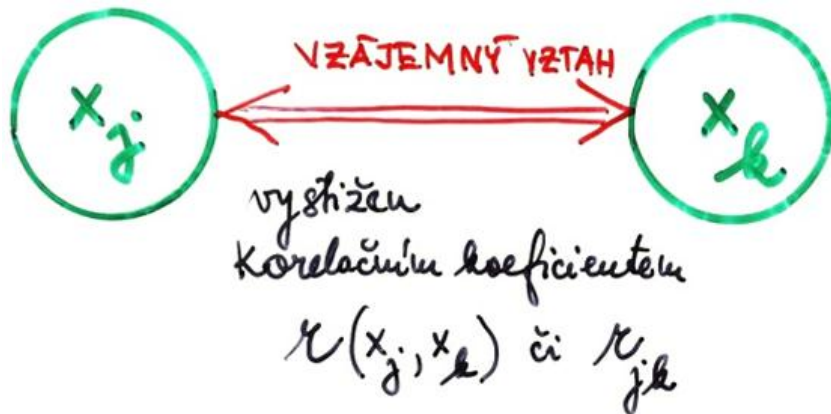
- 1) **Redukce** počtu proměnných, zjednodušení popisů jevů, reakcí, objektů
- 2) **Klasifikace** proměnných cestou shlukování původních proměnných v jiné obecnější
- 3) **Odhad neměřitelných veličin** a zprostředkování veličin jinak experimentálně nedostupných
- 4) **Tvorba, ověření** struktury objektů
- 5) **Transformace** původních proměnných do výhodnějšího tvaru, např. jejich ortogonalizace



Metoda hlavních komponent – použití modelů s latentními proměnnými

Otázka: Proč spolu x_j s x_k souvisí, koreluje, asociuje?

Odpověď (FA, PCA): Existuje „něco“ společného, existuje společná latentní proměnná f (nebo více $f_1, f_2, \dots, f_m$)



Arial 9 **B** *I* U

 Proměnné Případy

% P/L Vypočtená data grafu... Přidat graf...

	1	2	3	4	5	6	7	8	9	10	11	12	13
	Osoba	Vyska	Hmotnost	Vlasy	Boty	Vek	Prijem	Pivo	Vino	Sex	Plavani	Puvod	IQ
MA	MA	198	92	-1	48	48	45000	420	115	-1	98	-1	100
MA	MA	184	84	-1	44	33	33000	350	102	-1	92	-1	130
MA	MA	183	83	-1	44	37	34000	320	98	-1	91	-1	127
MA	FA	166	47	-1	36	32	28000	270	78	1	75	-1	112
MA	FA	170	60	1	38	23	20000	312	99	1	81	-1	110
FA	FA	172	64	1	39	24	22000	308	91	1	82	-1	102
FA	MA	188	88	1	42	35	30000	300	85	1	85	1	140
FA	MA												129
MA	FA												98
MA	FA												100
FA	MA												105
FA	FA	157	47	1	30	32	32000	230	92	1	70	-1	127
MA	FA	164	50	1	38	41	34000	255	134	1	76	-1	102
FA	FA	162	49	1	37	40	34000	265	124	1	75	-1	108
FA	MA	180	82	-1	44	43	37000	355	82	-1	88	-1	109
FA	MA	180	81	-1	44	46	42000	362	90	-1	86	-1	113
MA	MB	185	82	-1	45	26	16000	295	180	-1	92	1	109
MA	MB	187	84	-1	46	27	16500	299	178	-1	95	1	119
MB	FB	168	50	1	37	49	34000	170	162	1	76	1	135
MB	FB	166	49	1	36	21	14000	150	245	1	75	1	123
FB	FB	158	46	1	34	30	18000	120	120	1	70	1	119
FB	MB	177	65	-1	41	26	18000	209	160	-1	86	1	120
FB	MB	180	72	-1	43	33	19000	236	175	-1	85	1	115
MB	MB	181	75	-1	43	42	31000	198	161	-1	83	1	105
MB	FB	163	50	1	36	18	11000	143	136	1	75	1	102
MB	FB	162	50	1	36	20	11500	133	146	1	74	1	132
FB	MB	176	68	-1	42	50	36000	195	177	-1	82	1	96
FB	MB	175	67	1	42	55	38000	185	187	-1	80	1	105
MB	FB	165	51	1	36	36	26000	121	129	1	76	1	126
MB	FB	161	48	1	35	41	31500	116	196	1	75	1	120
FB	MB	178	75	-1	42	30	24000	203	208	-1	81	1	118

Načtení zdrojové matice dat, kde v nultém sloupci musí být označení osoby.

STATISTICA Cz - [Data: 22Lide.sta (13s krát 32F)]

Soubor Upravit Zobrazit Vložit Formát Statistický Data Mining Grafy Nástroje Data Okno Náp

Přidat do sešitu Přidat do protokolu Přidat

Arial 9 B I U

	1	2	3	4	5	6	7	8
	Osoba				Boty	Vek	Prijem	Pivo
1MA	1MA				48	48	45000	420
2MA	2MA				44	33	33000	350
3MA	3MA				44	37	34000	320
1FA	1FA	166	47	1	36	32	28000	270
2FA	2FA	170	60	1	38	23	20000	312
3FA	3FA	172						
4MA	4MA	182						
5MA	5MA	180						
4FA	4FA	169						
5FA	5FA	168						
6MA	6MA	183						
6FA	6FA	157						
7FA	7FA	164						
8FA	8FA	162						
7MA	7MA	180						
8MA	8MA	180						
1MB	1MB	185						
2MB	2MB	187						
1FB	1FB	168						
2FB	2FB	166						
3FB	3FB	158						
3MB	3MB	177						
4MB	4MB	180						
5MB	5MB	181						
4FB	4FB	163						
5FB	5FB	162						
6MB	6MB	176						

1. krok

Základní statistiky a tabulky: 22Lide.sta

Základní výsledky

- Popisné statistiky
- Korelační matice**
- t-test, nezávislé, dle skupin
- t-test, nezávislé, dle proměn.
- t-test, závislé vzorky
- t-test, samost. vzorek
- Rozklad & jednofakt. ANOVA
- Rozklad
- Tabulky četností
- Kontingenční tabulky
- Tabulky vícenásob. odpovědí
- Testy rozdílu: t, %, průměry
- Pravděpodobnostní kalkulátor

OK Sorno Možnosti Otevři data

2. krok

Vyberte proměnné na analýzu

1 - Osoba 13 - IQ

2 - Vyska

3 - Hmotnost

4 - Vlasy

5 - Boty

6 - Vek

7 - Prijem

8 - Pivo

9 - Vino

10 - Sex

11 - Plavani

12 - Puvod

OK Sorno [Svazky]...

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

Vybrat vše Roztáhnout Přiblížit

Vyberte proměnné:

2-13

☐ Pouze odpovídající proměnné

Korelace a parciální korelace: 22Lide.sta

1 seznam proměn. 2 seznamy (obd. matice) Výpočet

První seznam: žádné Sorno

Druhý seznam: žádné

Možnosti Barev. matice

Základní výsledky Detaily

Souhrn: Korelace Grafy

Matice bod. grafů zvolených proměnných

SELECT CASES f y

☐ Vážené momenty

SV =

W1 N1

ChD vynechána

☒ Celé případy

☐ Párově

3. krok

Vyberte 1 nebo 2 seznamy proměnných

1 - Osoba 11 - Plavani

2 - Vyska 12 - Puvod

3 - Hmotnost 13 - IQ

4 - Vlasy

5 - Boty

6 - Vek

7 - Prijem

8 - Pivo

9 - Vino

10 - Sex

OK Sorno

Vybrat vše Podrobn. Přiblížit Vybrat vše Podrobn.

1. seznam proměnných:

2-13

2. seznam proměnných (nepov):

2-13

☐ Pouze odpovídající proměnné

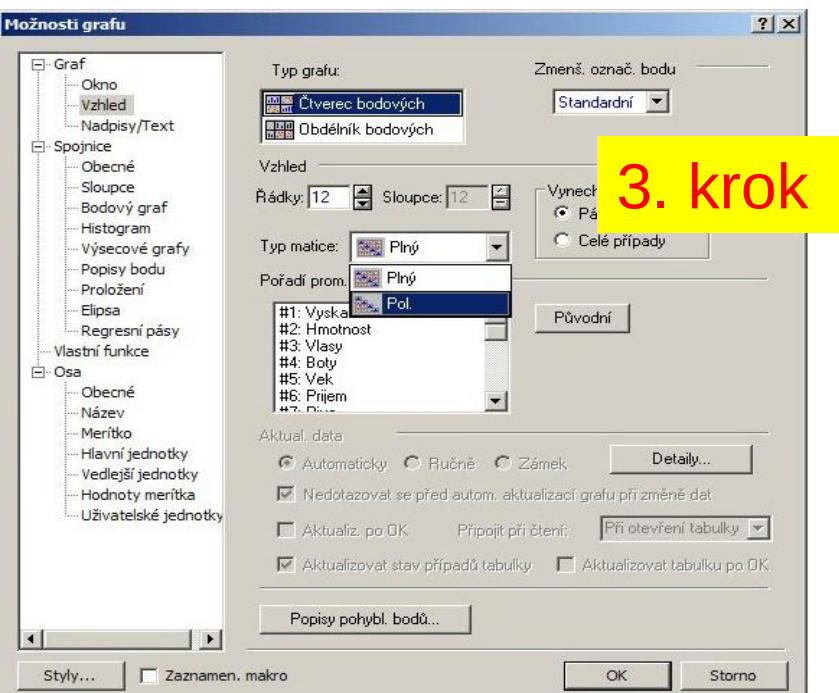
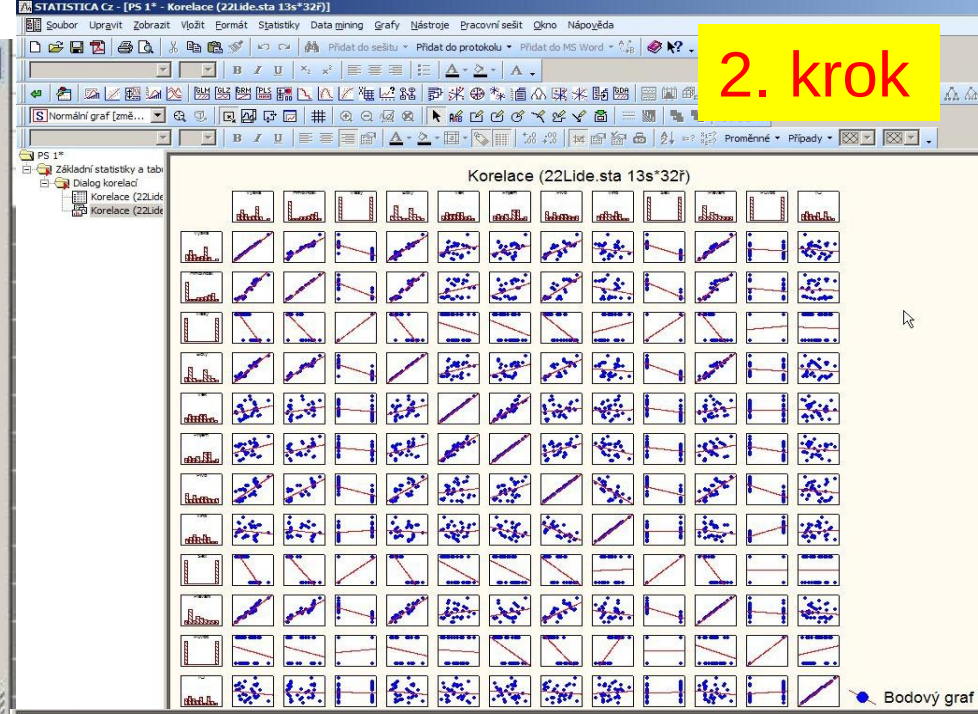
4. krok: výsledek

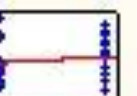
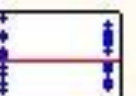
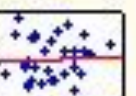
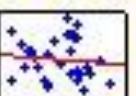
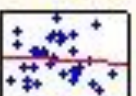
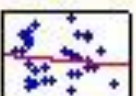
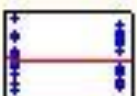
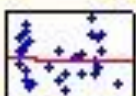
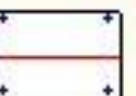
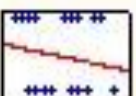
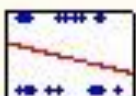
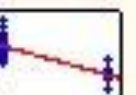
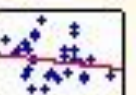
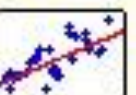
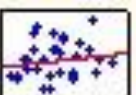
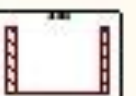
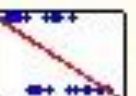
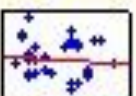
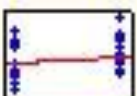
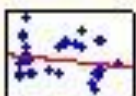
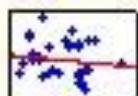
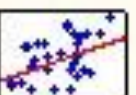
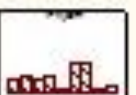
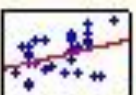
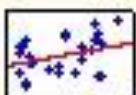
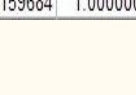
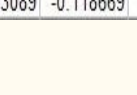
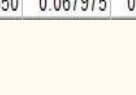
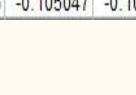
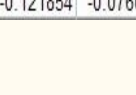
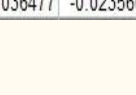
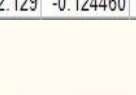
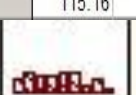
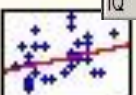
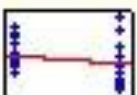
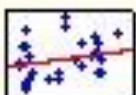
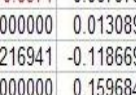
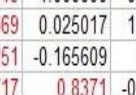
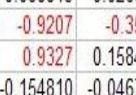
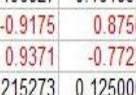
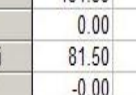
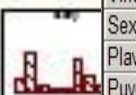
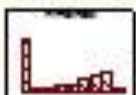
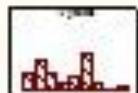
Korelace (22Lide.sta)

Označ. korelace jsou významné na hlad. $p < .05000$

N=32 (Celé případy vynechány u ChD)

Proměnná	Vyska	Hmotnost	Vlasy	Boty	Vek	Prijem	Pivo	Vino	Sex	Plavani	Puvod	IQ
Vyska	1.00	0.9597	-0.8208	0.9609	0.24	0.30	0.7150	-0.14	-0.8776	0.9621	-0.18	-0.12
Hmotnost	0.9597	1.00	-0.8339	0.9691	0.25	0.34	0.7382	-0.20	-0.9175	0.9371	-0.22	-0.04
Vlasy	-0.8208	-0.8339	1.00	-0.8229	-0.20	-0.25	-0.5965	0.16	0.8750	-0.7723	0.12	-0.02
Boty	0.9609	0.9691	-0.8229	1.00	0.32	0.3544	0.6970	-0.09	-0.9207	0.9327	-0.15	-0.12
Vek	0.24	0.25	-0.20	0.32	1.00	0.8849	0.13	0.03	-0.3536	0.16	-0.05	-0.08
Prijem	0.30	0.34	-0.25	0.3544	0.8849	1.00	0.4170	-0.30	-0.32	0.25	-0.4516	-0.11
Pivo	0.7150	0.7382	-0.5965	0.6970	0.13	0.4170	1.00	-0.6545	-0.5369	0.7251	-0.7717	-0.11
Vino	-0.14	-0.20	0.16	-0.09	0.03	-0.30	-0.6545	1.00	0.03	-0.17	0.8371	0.07
Sex	-0.8776	-0.9175	0.8750	-0.9207	-0.3536	-0.32	-0.5369	0.03	1.00	-0.8157	-0.00	0.01
Plavani	0.9621	0.9371	-0.7723	0.9327	0.16	0.25	0.7251	-0.17	-0.8157	1.00	-0.22	-0.12
Puvod	-0.18	-0.22	0.13	-0.15	-0.05	-0.4516	-0.7717	0.8371	0.00	-0.22	1.00	0.16
IQ	-0.12	-0.04	-0.02	-0.12	-0.08	-0.11	-0.11	0.07	0.01	-0.12	0.16	1.00





Korelace (22Lide.sta)
Označ. korelace jsou významné na hlad. p < .05000
N=32 (Celé případy vynechány u ChD)

Proměnná	Průměry	Sm.odch.	Vyska	Hmotnost	Vlasy	Boty	Vek	Prijem	Pivo	Vino	Sex	Plavani	Puvod	IQ
Vyska	173.13	10.057	1.000000	0.9597	-0.8208	0.9609	0.241055	0.301097	0.7150	-0.138053	-0.8776	0.9621	-0.176791	-0.124460
Hmotnost	64.47	15.191	0.9597	1.000000	-0.8339	0.9691	0.254231	0.335403	0.7382	-0.196827	-0.9175	0.9371	-0.215273	-0.036477
Vlasy	0.00	1.016	-0.8208	-0.8339	1.000000	-0.8229	-0.200165	-0.252447	-0.5965	0.164854	0.8750	-0.7723	0.125000	-0.023560
Boty	39.91	3.897	0.9609	0.9691	-0.8229	1.000000	0.322976	0.3544	0.6970	-0.088845	-0.9207	0.9327	-0.154810	-0.121854
Vek	34.44	9.517	0.241055	0.254231	-0.200165	0.322976	1.000000	0.8849	0.128175	0.026822	-0.3536	0.158410	-0.046705	-0.076065
Prijem	27437.50	8929.608	0.301097	0.335403	-0.252447	0.3544	0.8849	1.000000	0.4170	-0.296704	-0.323559	0.251769	-0.4516	-0.105047
Pivo	249.50	90.597	0.7150	0.7382	-0.5965	0.6970	0.128175	0.4170	1.000000	-0.6545	-0.5369	0.7251	-0.7717	-0.107550
Vino	131.59	49.497	-0.138053	-0.196827	0.164854	-0.088845	0.026822	-0.296704	-0.6545	1.000000	0.025017	-0.165609	0.8371	0.067975
Sex	0.00	1.016	-0.8776	-0.9175	0.8750	-0.9207	-0.3536	-0.323559	-0.5369	0.025017	1.000000	-0.8157	-0.000000	0.013089
Plavani	81.50	7.318	0.9621	0.9371	-0.7723	0.9327	0.158410	0.251769	0.7251	-0.165609	-0.8157	1.000000	-0.216941	-0.118669
Puvod	-0.00	1.016	-0.176791	-0.215273	0.125000	-0.154810	-0.046705	-0.4516	-0.7717	0.8371	-0.000000	-0.216941	1.000000	0.159684
IQ	115.16	12.129	-0.124460	-0.036477	-0.023560	-0.121854	-0.076065	-0.105047	-0.107550	0.067975	0.013089	-0.118669	0.159684	1.000000

Úprava zdrojové matice dat:

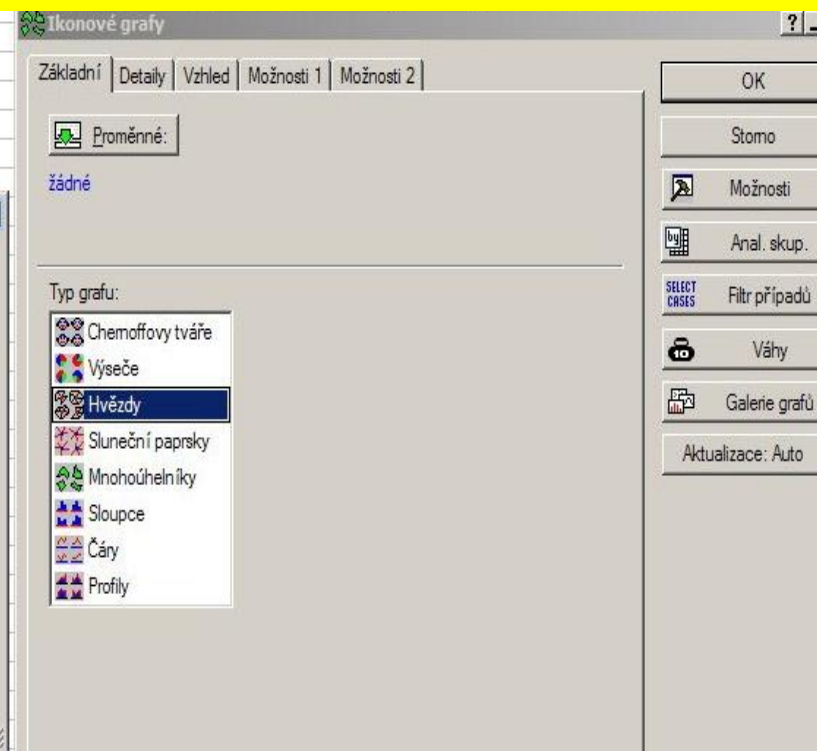
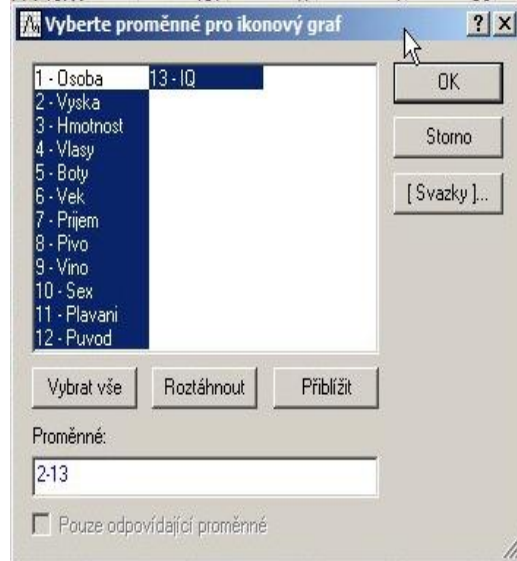
Označení osob vložení jejich indexů, např. 1MA, 2MA, atd.

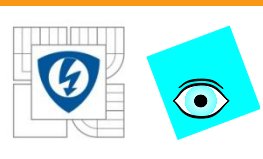
	1 Osoba	2 Vyska	3 Hmotnost	4 Vlasy	5 Boty	6 Vek	7 Prijem	8 Pivo	9 Vino	10 Sex	11 Plavani	12 Puvod	13 IQ
1MA	1MA	198	92	-1	48	48	45000	420	115	-1	98	-1	100
2MA	2MA	184	84	-1	44	33	33000	350	102	-1	92	-1	130
3MA	3MA	183	83	-1	44	37	34000	320	98	-1	91	-1	127
1FA	1FA	166	47	-1	36	32	28000	270	78	1	75	-1	112
2FA	2FA	170	60	1	38	23	20000	312	99	1	81	-1	110
3FA	3FA	172	64	1	39	24	22000	308	91	1	82	-1	102
4MA	4MA	182	80	-1	42	35	30000	398	65	-1	85	-1	140
5MA	5MA	180	80	-1	43	36	30000	388	63	-1	84	-1	129
4FA	4FA	169	51	1	36	24	23000	250	89	1	78	-1	98
5FA	5FA	168	52	1	37	27	23500	260	86	1	78	-1	100
6MA	6MA	183	81	-1	42	37	35000	345	45	-1	90	-1	105
6FA	6FA	157	47	1	36	32	32000	235	92	1	70	-1	127
7FA	7FA	164	50	1	38	41	34000	255	134	1	76	-1	102
8FA	8FA	162	49	1	37	40	34000	265	124	1	75	-1	108
7MA	7MA	180	82	-1	44	43	37000	355	82	-1	88	-1	109
8MA	8MA	180	81	-1	44	46	42000	362	90	-1	86	-1	113
1MB	1MB	185	82	-1	45	26	16000	295	180	-1	92	1	109
2MB	2MB	187	84	-1	46	27	16500	299	178	-1	95	1	119
1FB	1FB	168	50	1	37	49	34000	170	162	1	76	1	135
2FB	2FB	166	49	1	36	21	14000	150	245	1	75	1	123
3FB	3FB	158	46	1	34	30	18000	120	120	1	70	1	119
3MB	3MB	177	65	-1	41	26	18000	209	160	-1	86	1	120
4MB	4MB	180	72	-1	43	33	19000	236	175	-1	85	1	115
5MB	5MB	181	75	-1	43	42	31000	198	161	-1	83	1	105
4FB	4FB	163	50	1	36	18	11000	143	136	1	75	1	102
5FB	5FB	162	50	1	36	20	11500	133	146	1	74	1	132
6MB	6MB	176	68	-1	42	50	36000	195	177	-1	82	1	96
7MB	7MB	175	67	1	42	55	38000	185	187	-1	80	1	105
6FB	6FB	165	51	1	36	36	26000	121	129	1	76	1	126
7FB	7FB	161	48	1	35	41	31500	116	196	1	75	1	120
8MB	8MB	178	75	-1	42	30	24000	203	208	-1	81	1	118
8FB	8FB	160	48	1	35	40	31000	118	198	1	74	1	129



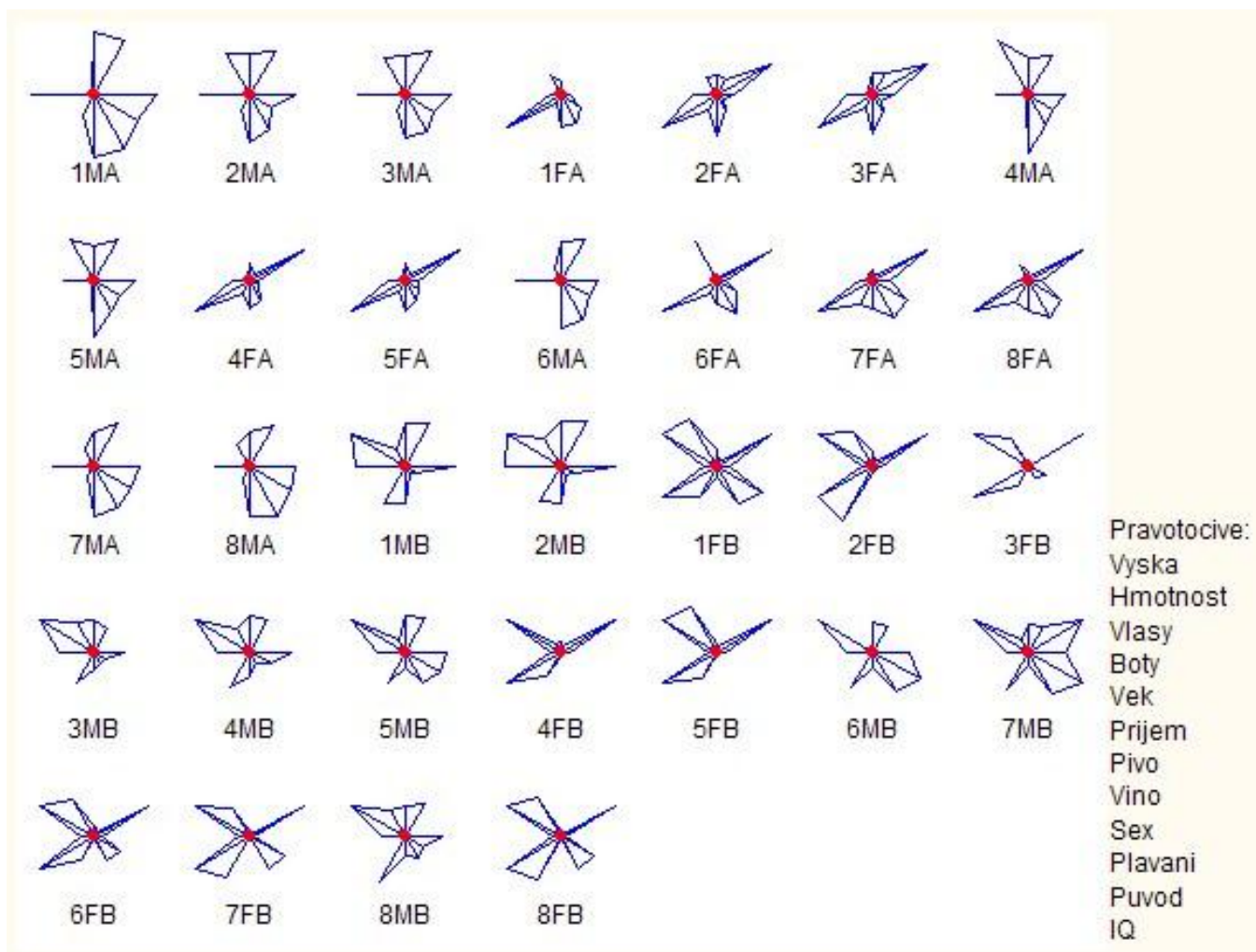
2. krok: postupná volba typu ikonových grafů

FA	2FA	1			
FA	3FA	1			
MA	4MA	182	80	-1	42
MA	5MA	180	80	-1	43
FA	4FA	169	51	1	36
FA	5FA	168	52	1	37
MA	6MA	183	81	-1	42
FA	6FA	157	47	1	36





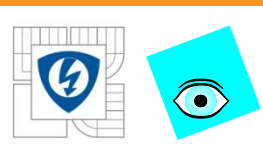
EDA: Symbolové grafy hvězdiček



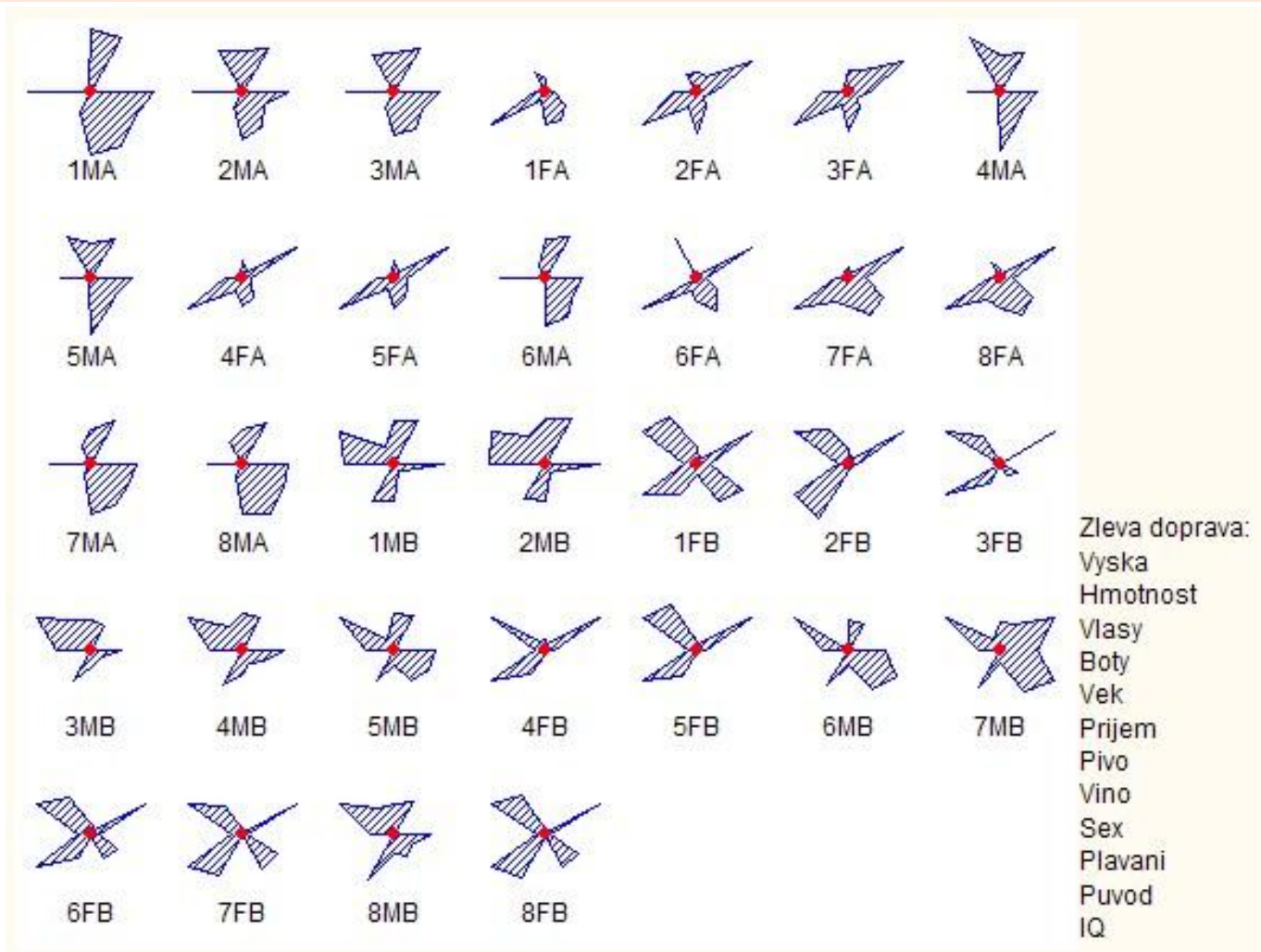
25.2.2010

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ





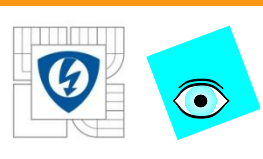
EDA: Symbolové grafy mnohoúhelníků



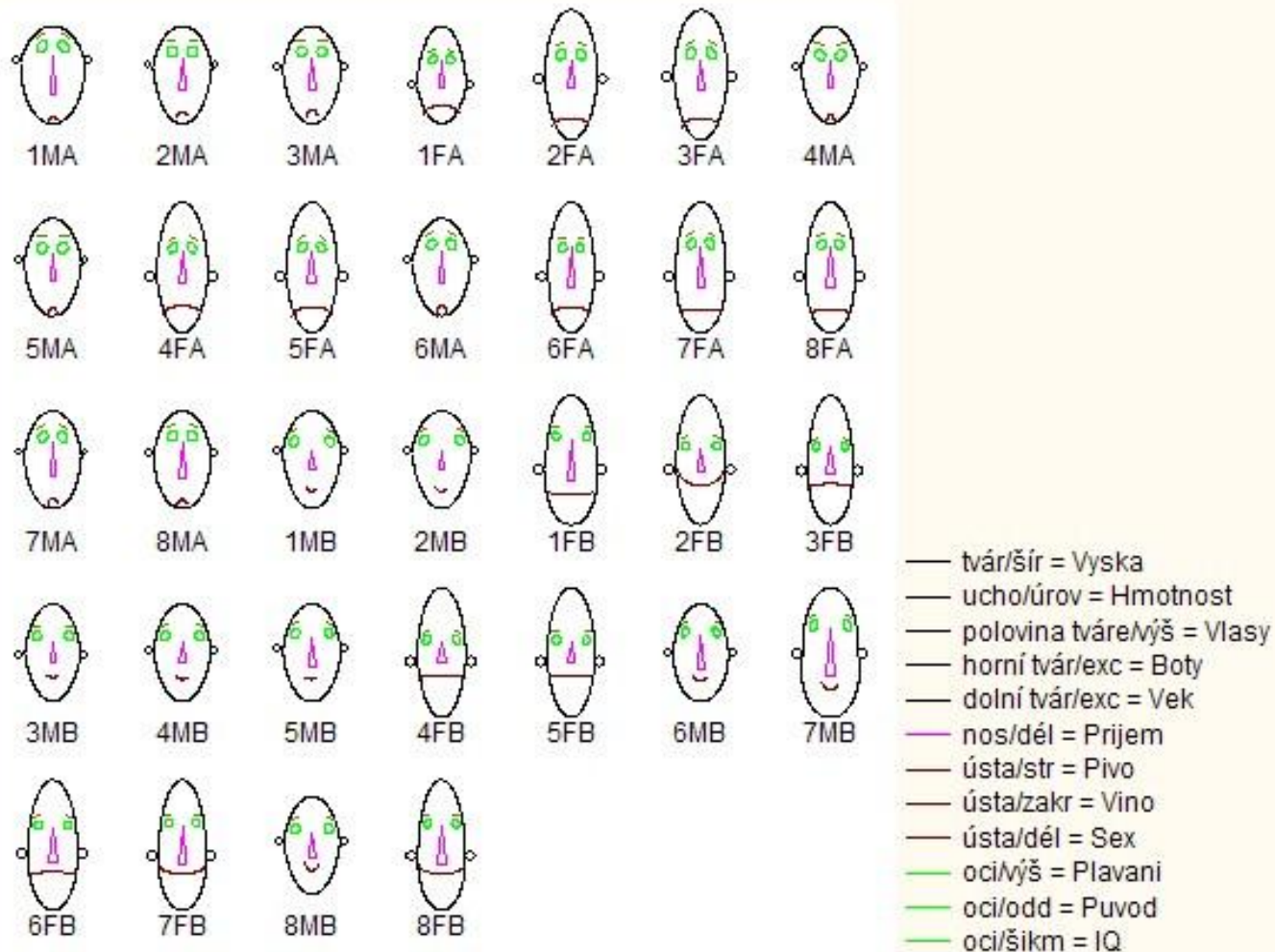
25.2.2010

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



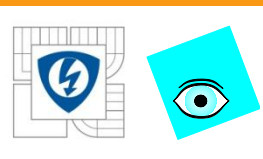


EDA: Symbolové grafy obličejů

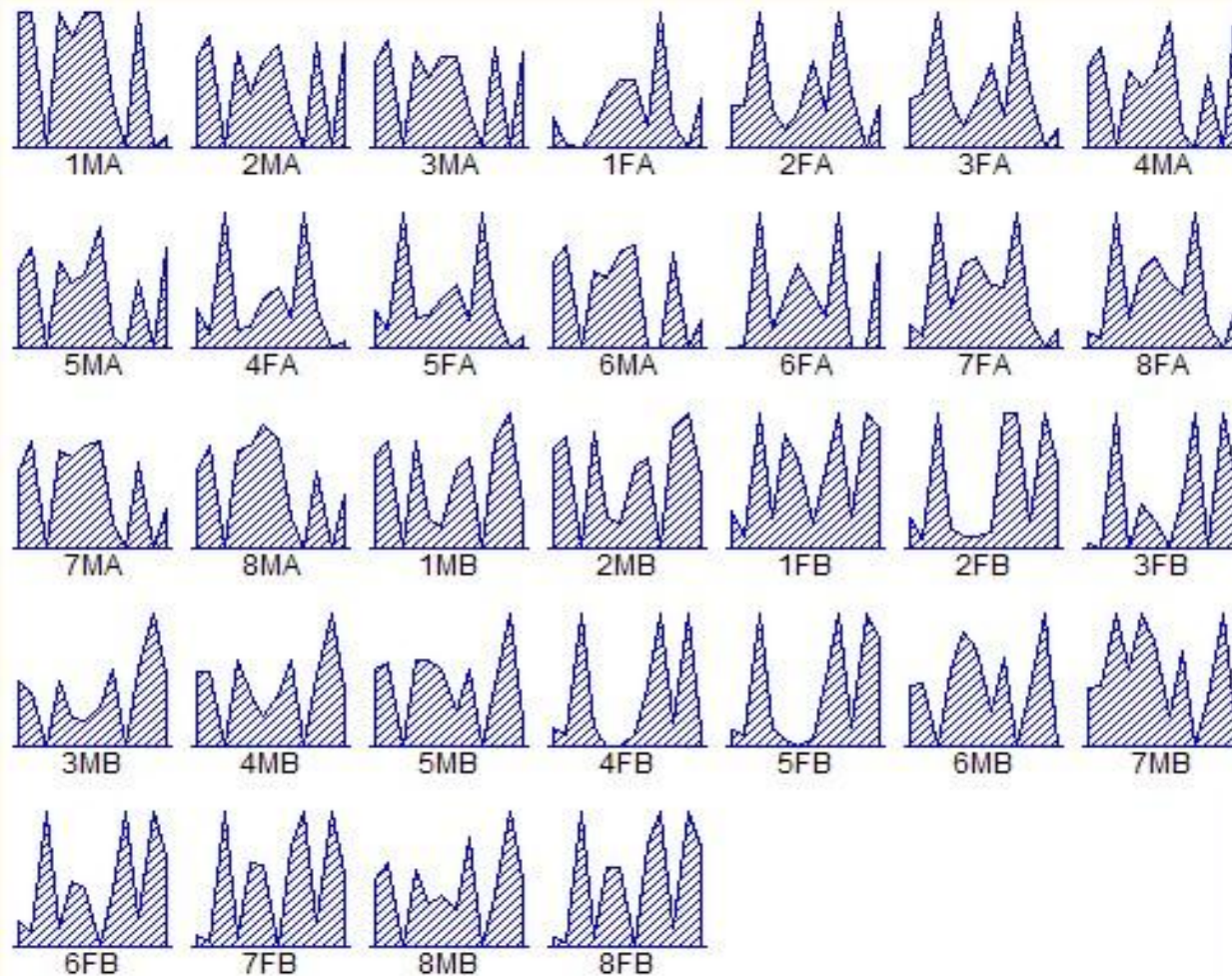


25.2.2010

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



EDA: Symbolové grafy profilů



Zleva doprava:
 Vyska
 Hmotnost
 Vlasy
 Boty
 Vek
 Prijem
 Pivo
 Vino
 Sex
 Plavani
 Puvod
 IQ

25.2.2010

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ



85	-1	140
84	-1	129

2D bodové grafy [?] [-] [X]

Základní | **Detaily** | Vzhled | Kategorizovaný | Možnosti 1 | Možnosti 2

Proměnné: X: **Vyska** Y: **Hmotnost**

Typ grafu:

- Běžný**
- Vícenásobný
- Dvojitě-Y
- Četnost
- Bubliny
- Kvantil
- Voronoi

Proložení

- ☐ Vypnuto
- ☒ **Lineární**
- Polynomiální
- Logaritmické
- Exponenciální
- MNC váž. vzdáleností
- MNC váž. neg. expon.
- Spline
- Lowess

Statistiky

- ☒ **R kvadrát**
- ☒ **Korelace a p (lin. prolož)**
- ☒ **Regresní rovnice**

Elipsa

- ☐ Vyp.
- ☒ **Normální** Koeficient:
- ☐ Rozsah

Regresní pásy

- ☐ Vyp.
- ☒ **Spolehl.** Hladina:
- ☐ Predikce

Dz. zvolené podsk.: **Vyp.**

OK

Storno

Možnosti ▾

Anal. skup.

SELECT CASES Filtr případů

Váhy

Galerie grafů

Aktualizace: Auto ▾

V submenu **Detaily** a pak v bloku **Statistiky** zaškrtnete **R-kvadrát**, **Korelace** a **Regresní rovnice**. V **Elipsa** zvolíme **Normální**. V bloku **Regresní pásy** zvolíme **pás spolehlivosti**.

Základní | Detaily | Vzhled | Kategorizovaný | Možnosti 1 | Možnosti 2

Nadpis

☒ Zobrazit výchozí název

Vl. nadpis:

Zobrazit nahoře ▾

OK

Storno



Možnosti ▾

V bloku Popisy případů zvolíte Jména případů a zaškrtnete Zobrazit popisy případů a Zobrazit textové popisy

☐ Zobrazit poznámku

Možnosti zobrazení

Popisy případů: **Jména přípa** ▾ Proměnná: Osoba

☒ Zobrazit popisy případů ☒ Zobr. textové popisy (data) jako hodnoty osy

Automatická aktualizace

☒ Aktualizovat z tabulky **Auto** ▾

☐ Aktualizovat z dialogu analýzy

☒ Aktualizovat při změně vstup. dat

☐ Neaktualizovat automaticky

Měřítko

Osa: **X** ▾

Typ: **Lineární** ▾

☐ Po vytvoření grafu tento dialog automaticky zavřít



Váhy



Galerie grafů

Aktualizace: Auto ▾

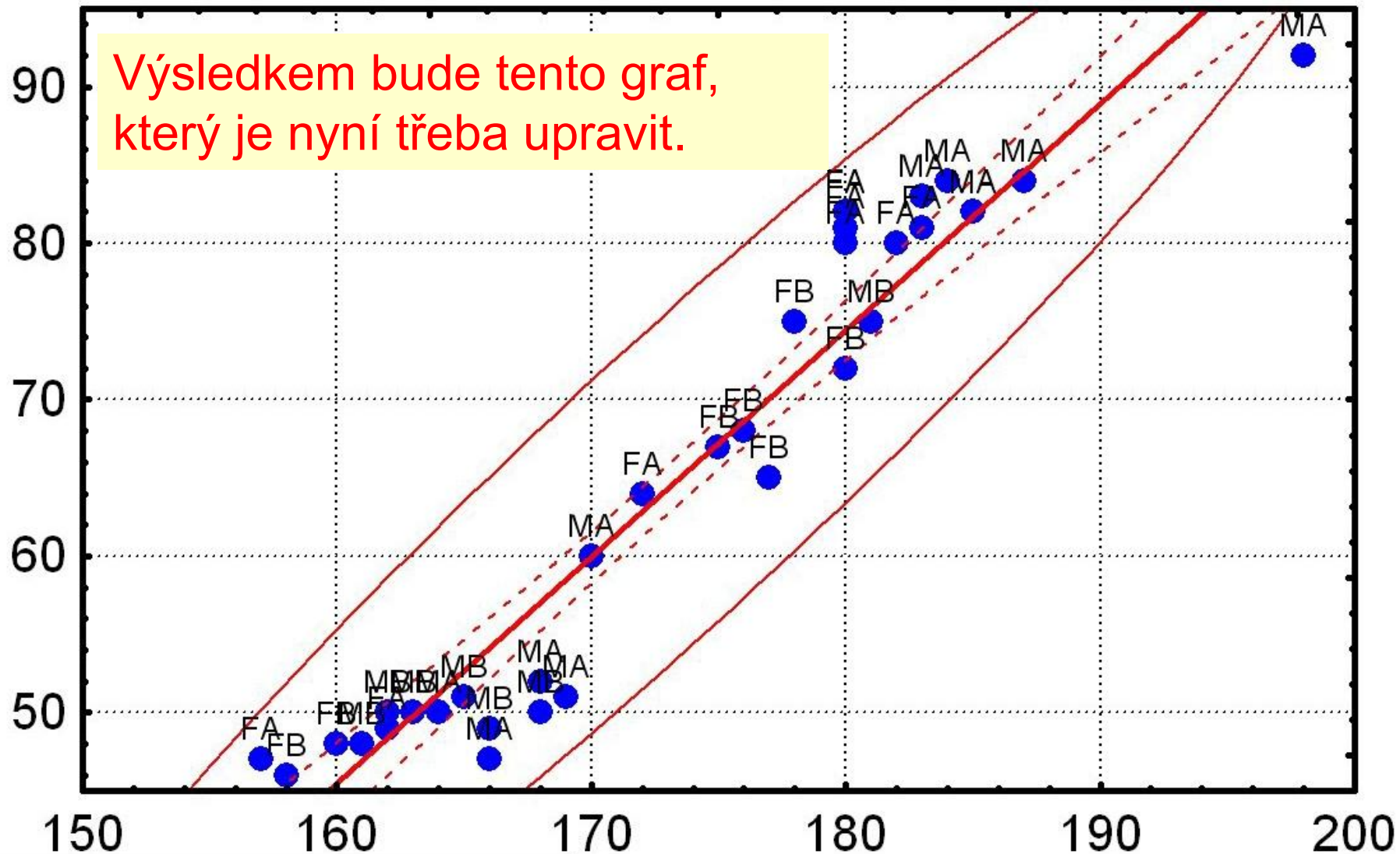
Bodový graf z Hmotnost proti Vyska

22Lide 13v*32c

Hmotnost = -186.488+1.4496*x; 0.95 Int.spol.

Výsledkem bude tento graf,
který je nyní třeba upravit.

Hmotnost



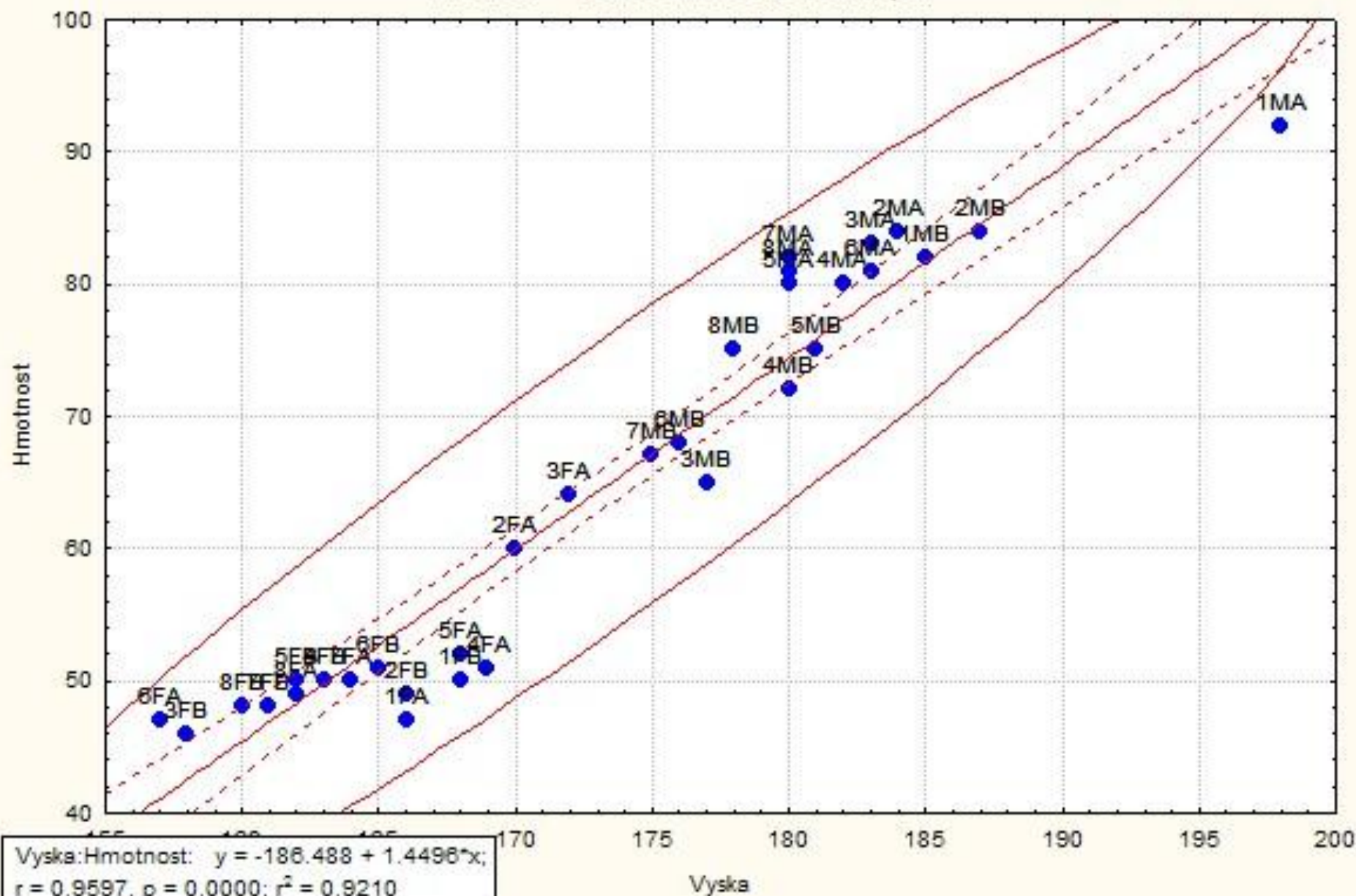
Vyska:Hmotnost: $y = -186.488 + 1.4496x$;
 $r = 0.9597$; $p = 0.0000$; $r^2 = 0.9210$

Vyska

Bodový graf z Hmotnost proti Vyska

22Lide.sta 13v*32c

Hmotnost = -186.488+1.4496*x; 0.95 Int.spol.



$$H_{motnost} = -186.488 + 1.4496 \cdot x; 0.95 \text{ Int.spol.}$$



Vyska

$$\text{Hmotnost} = -186.488 + 1.4496 \cdot x; 0.95 \text{ Int.spol.}$$

Hmotnost

Možnosti grafu

- Graf
- Spojnice
- Vlastní funkce
- Osa
 - Obecné
 - Název
 - Měřitko
 - Hlavní jednotky
 - Vedlejší jednotky
 - Hodnoty měřítka
 - Uživatelské jednotky

Osa: ☒ Kopírovat spec. osy do

Rozsah

Mód:

Minimum: Maximum:

Volby pro automatický rozsah

☐ Začít od nuly

☒ Okraje na hlavní jednotky

Měřítka vícenás. grafu:

Typ měřítka

☐ Reverz. měřítko

Přerušení měřítka

☐ Vše ignorovat

Od: Do: Šířka (všech) přerušení: bodů

☐ Zaznamen. makro

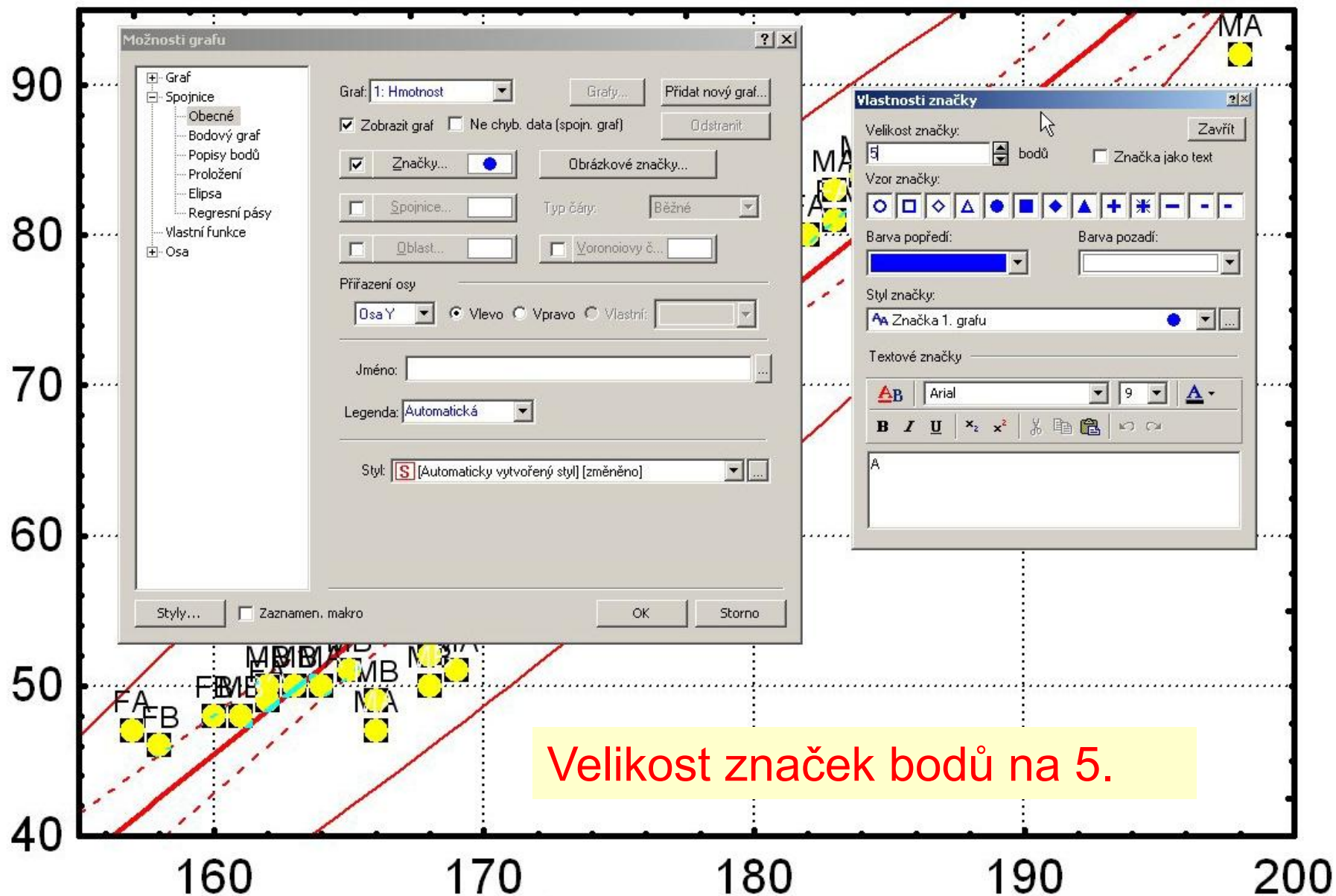
Popisy a mód osy y.

Vyska:Hmotnost: $y = -186.488 + 1.4496 \cdot x$;
 $r = 0.9597$; $p = 0.0000$; $r^2 = 0.9210$

Vyska

$$\text{Hmotnost} = -186.488 + 1.4496 \cdot x; 0.95 \text{ Int.spol.}$$

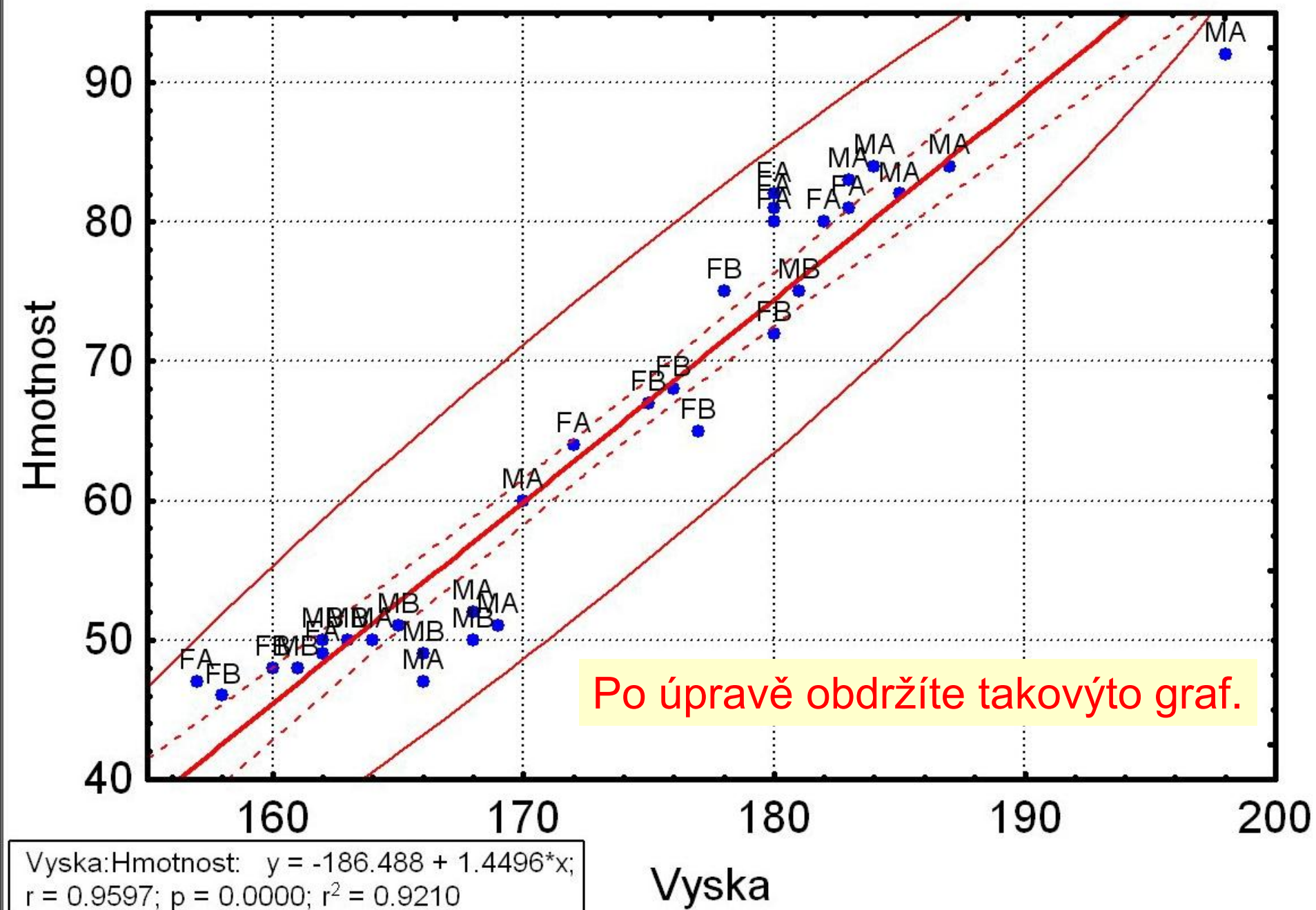
Hmotnost



Vyska:Hmotnost: $y = -186.488 + 1.4496 \cdot x$;
 $r = 0.9597$; $p = 0.0000$; $r^2 = 0.9210$

Vyska

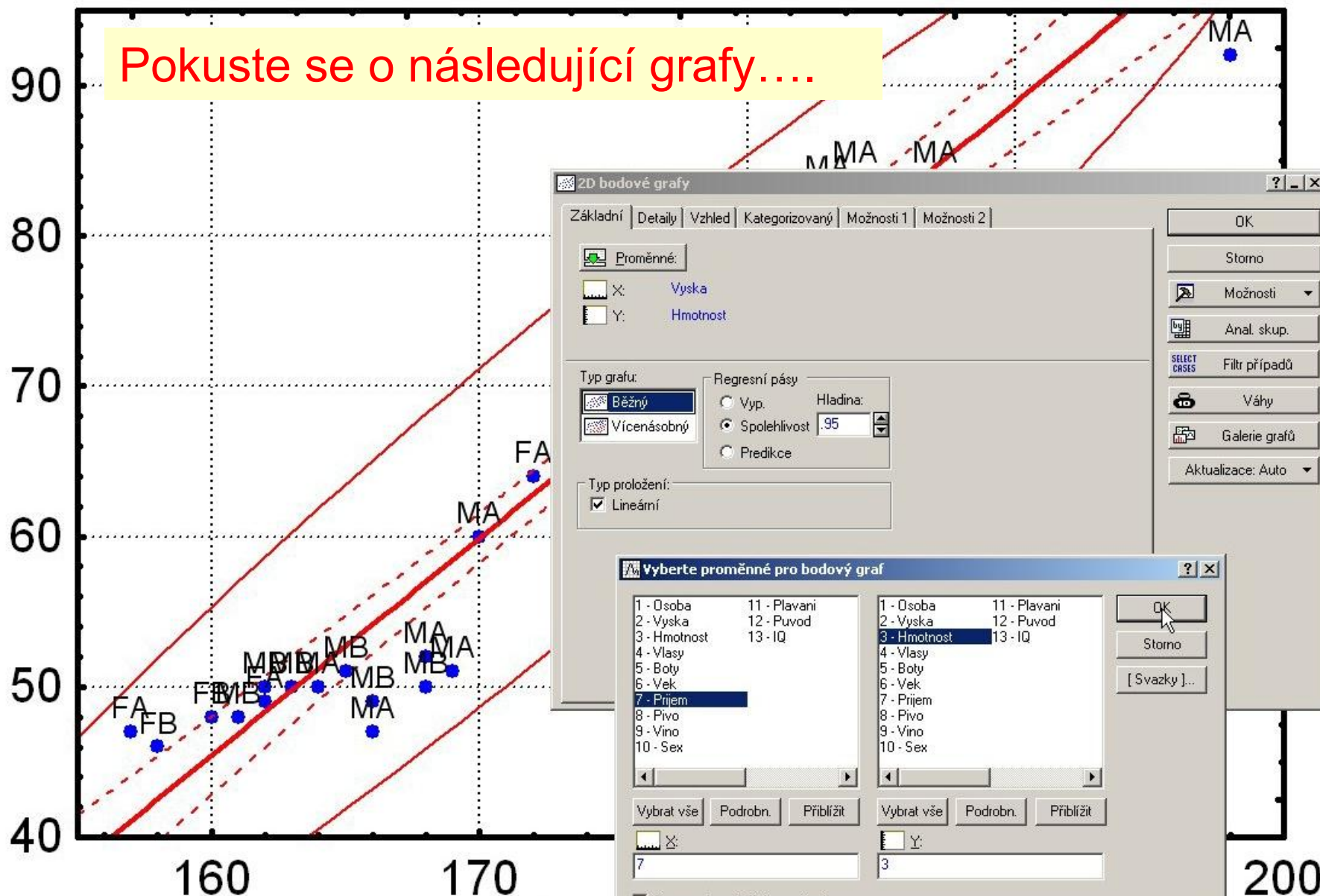
$$Hmotnost = -186.488 + 1.4496 \cdot x; 0.95 \text{ Int.spol.}$$



$$\text{Hmotnost} = -186.488 + 1.4496 * x; 0.95 \text{ Int.spol.}$$

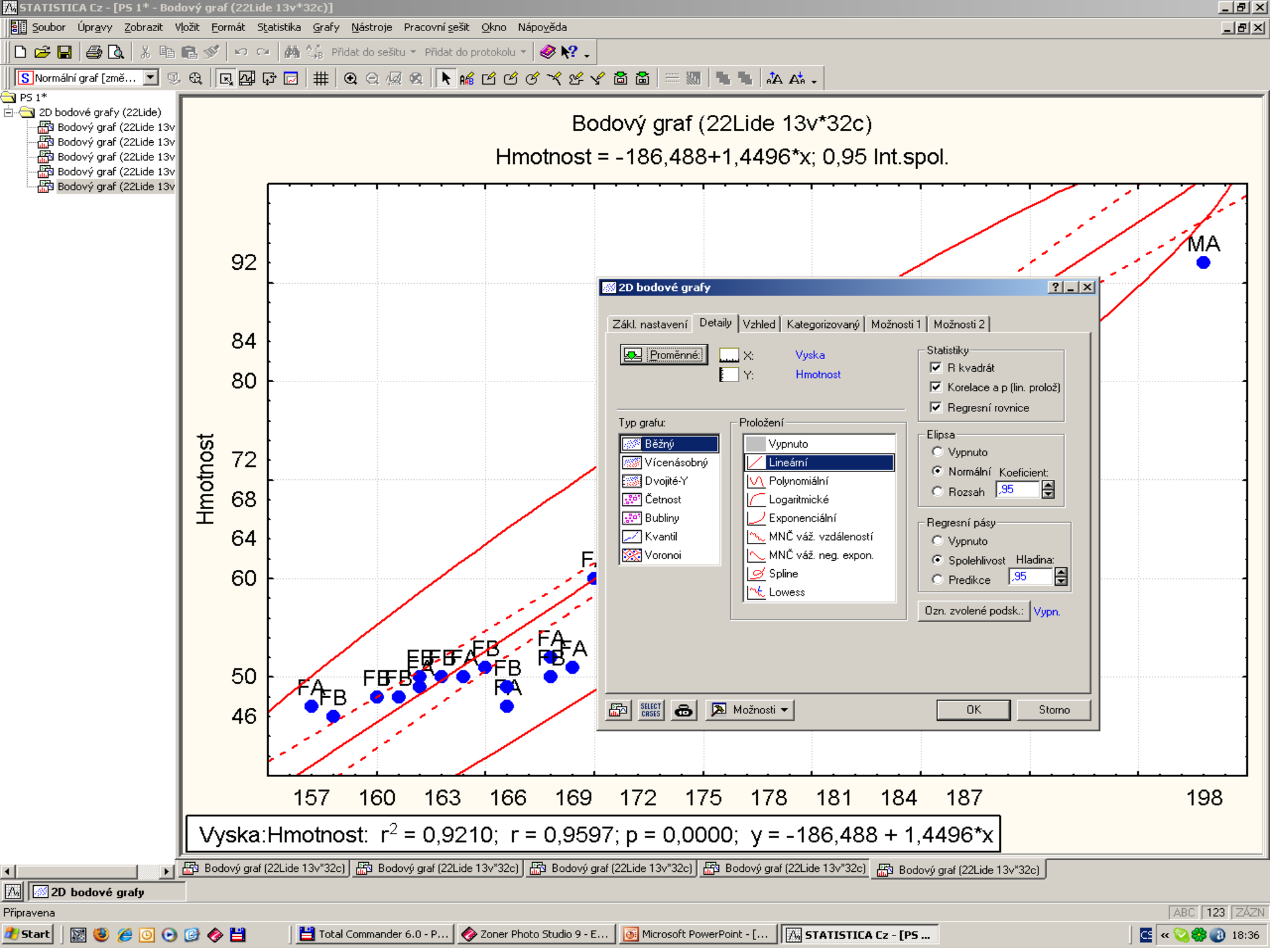
Pokuste se o následující grafy....

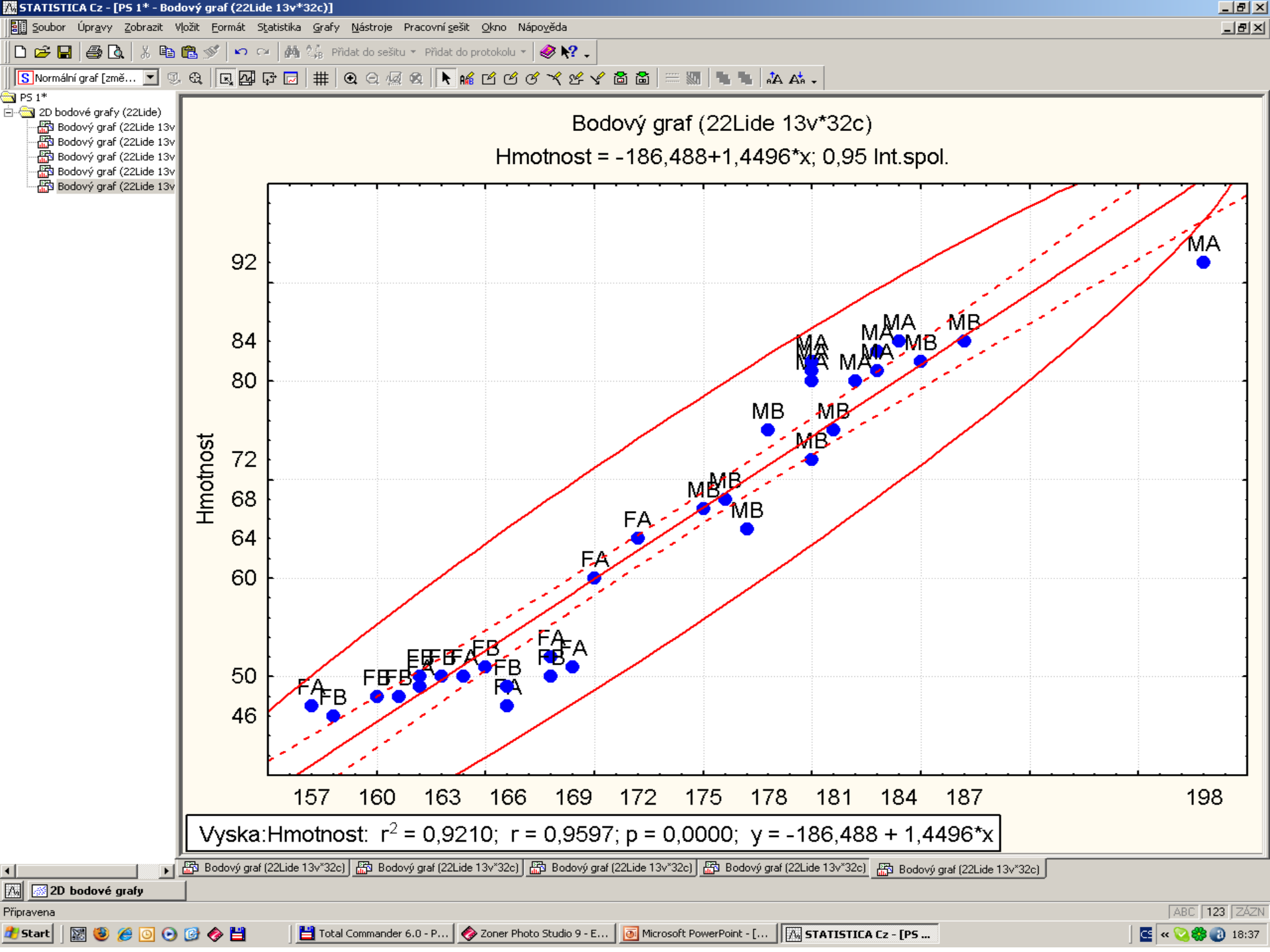
Hmotnost



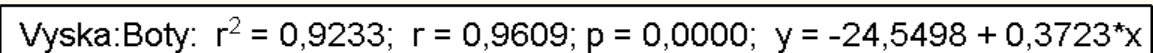
Vyska:Hmotnost: $y = -186.488 + 1.4496 * x;$
 $r = 0.9597; p = 0.0000; r^2 = 0.9210$

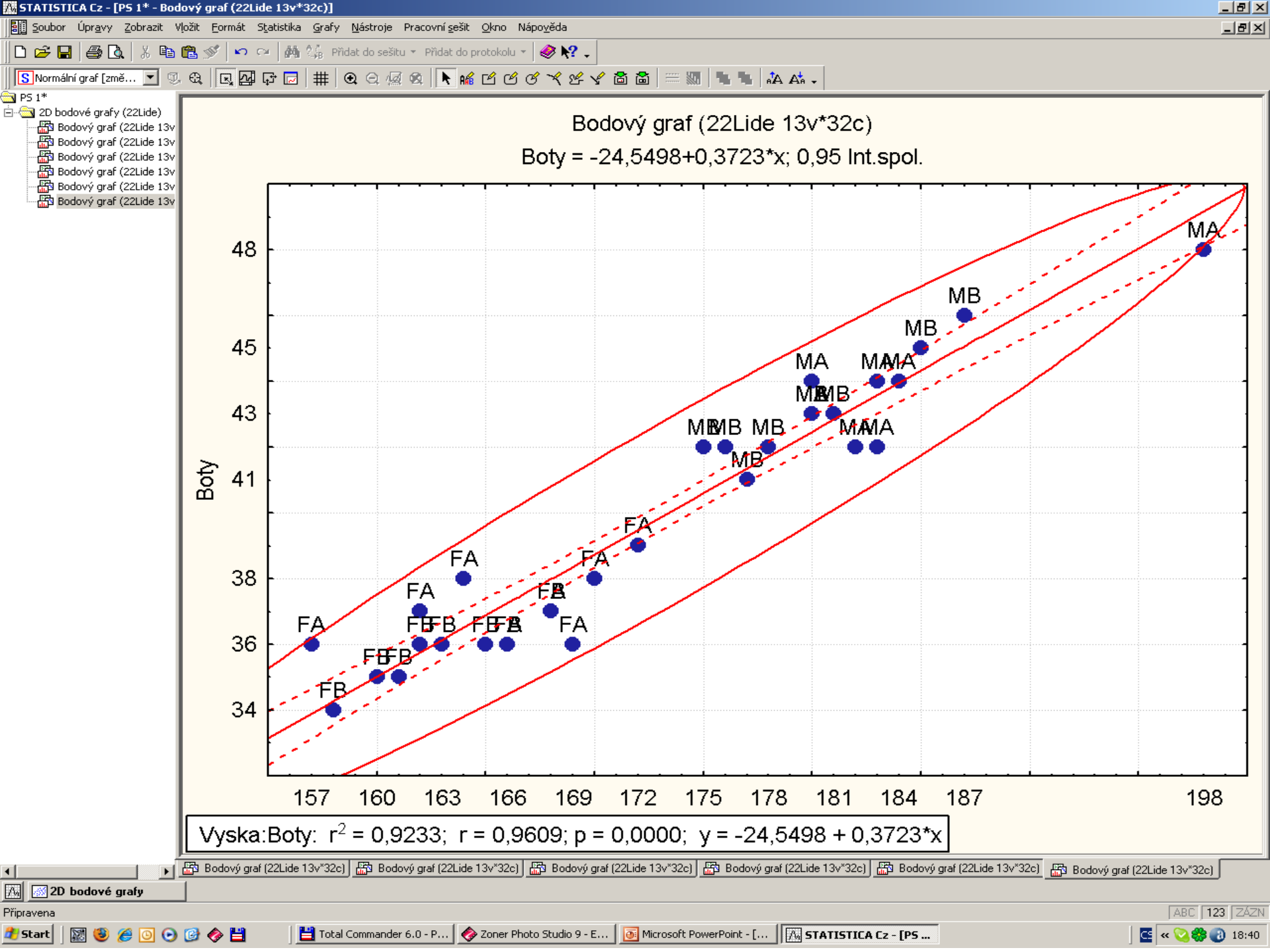
Vyska





- Bodový graf (22Lide 13v)
- Bodový graf (22Lide 13v)
- Bodový graf (22Lide 13v)
- Bodový graf (22Lide 13v)
- Bodový graf (22Lide 13v)
- Bodový graf (22Lide 13v)

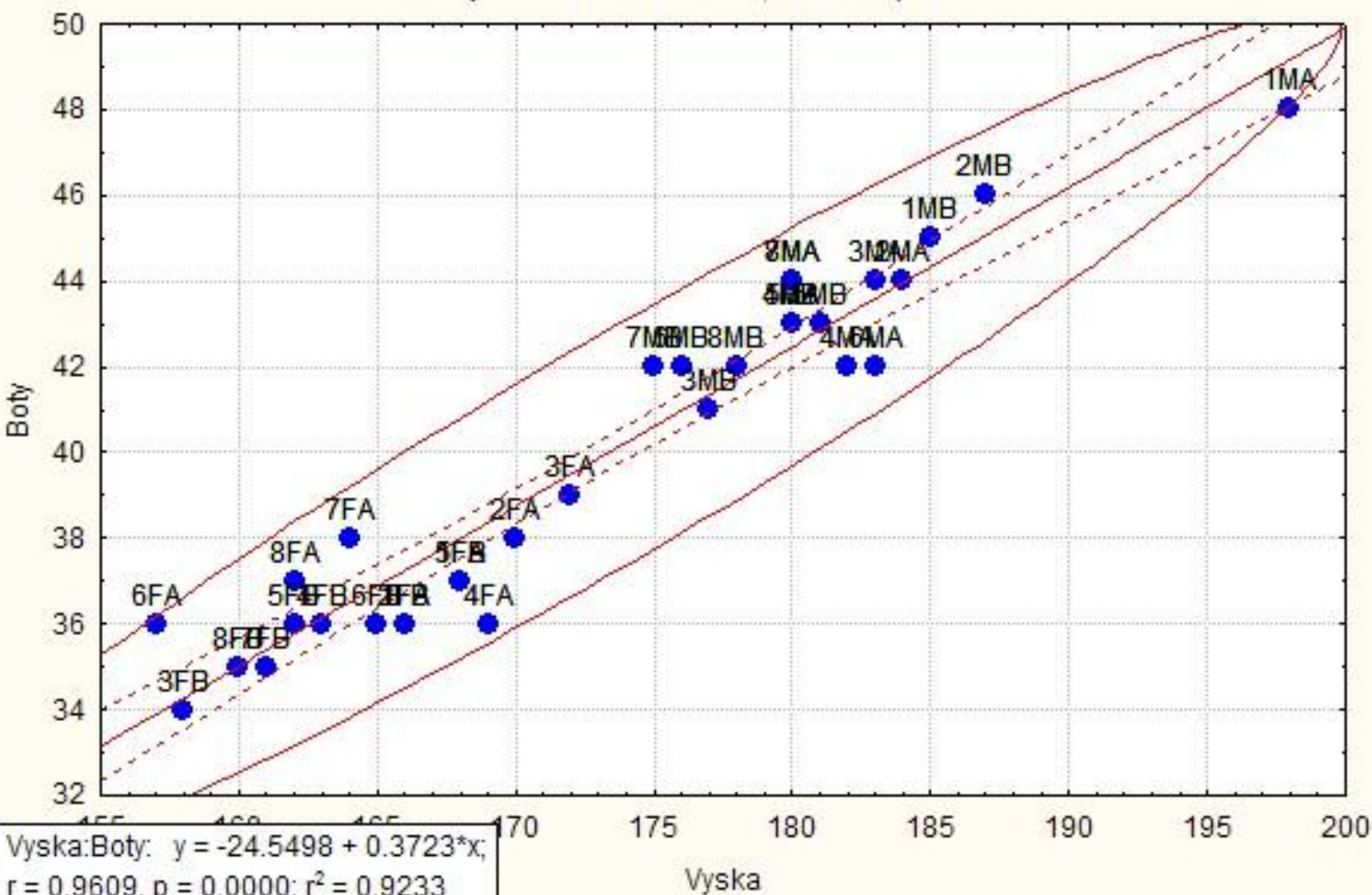


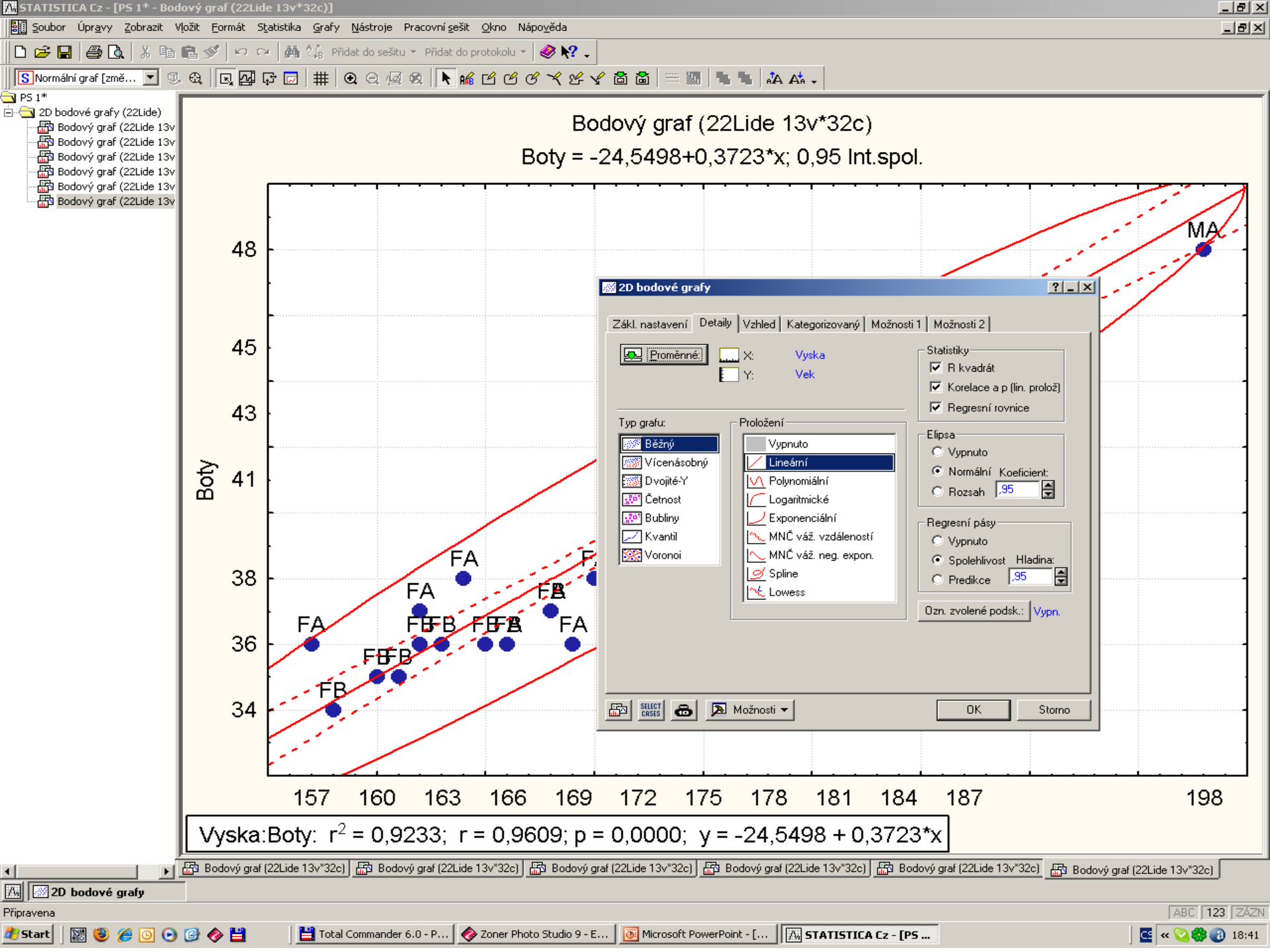


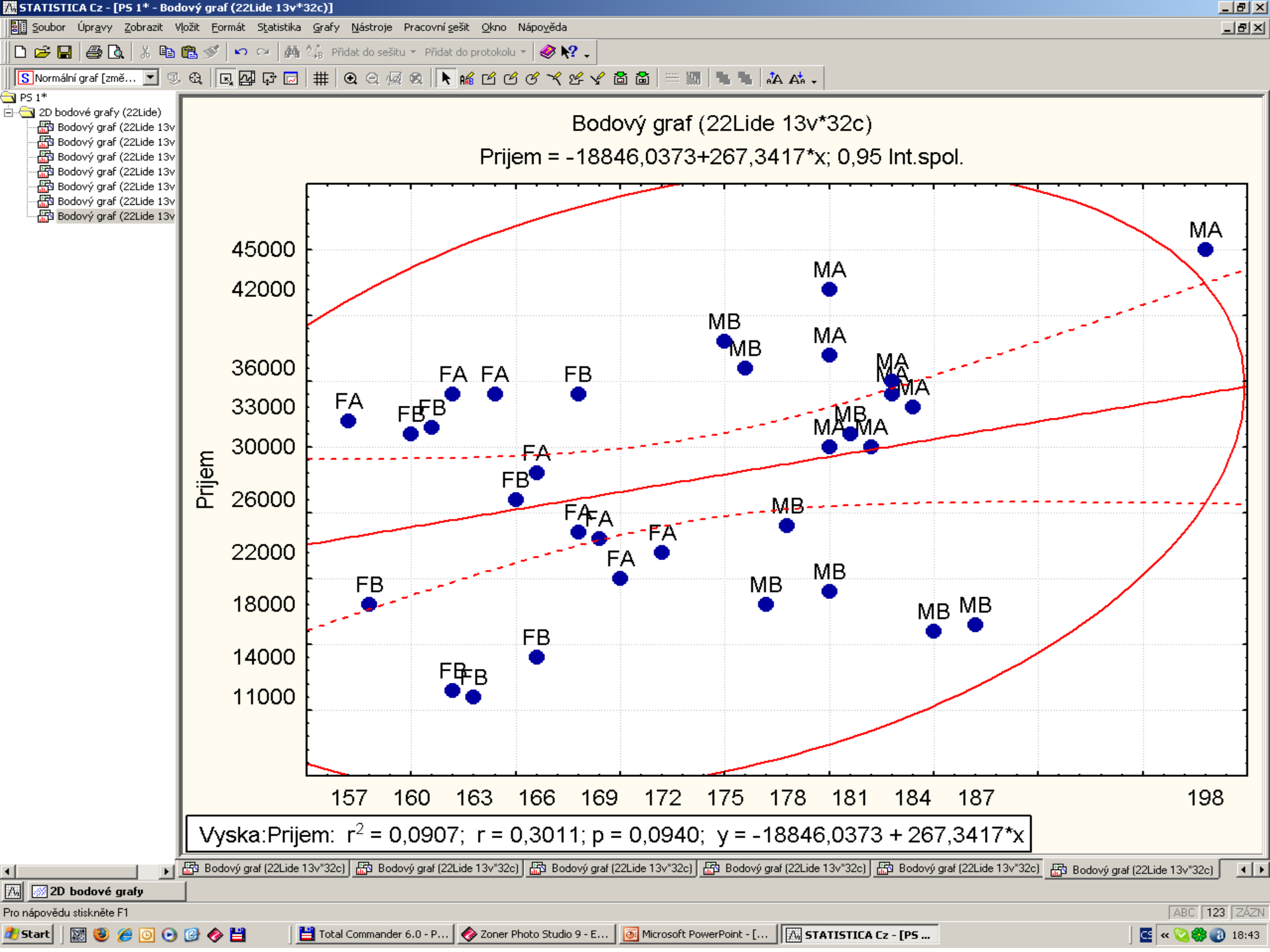
Bodový graf z Boty proti Vyska

22Lide.sta 13v*32c

Boty = -24.5498+0.3723*x; 0.95 Int.spol.



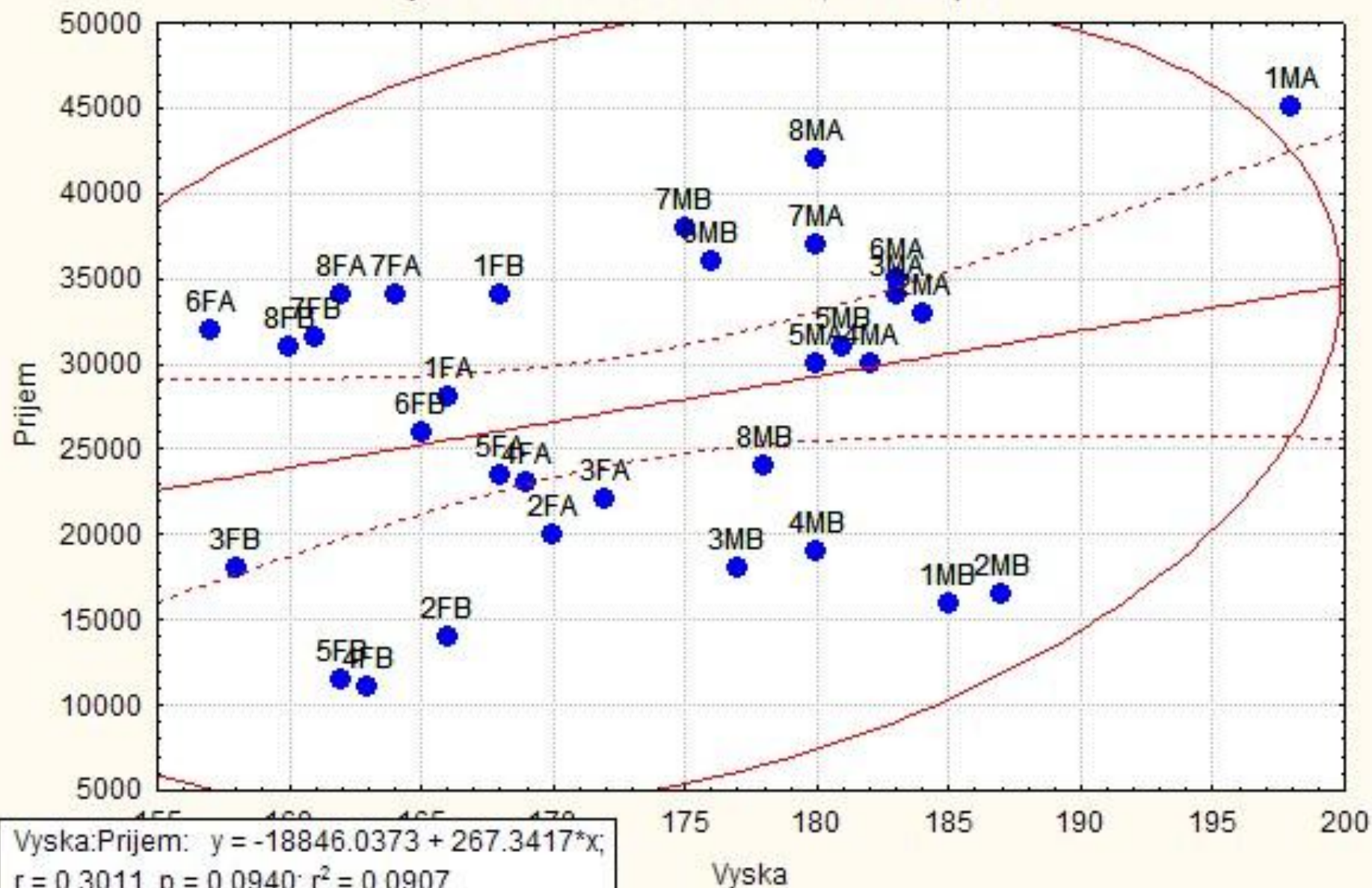




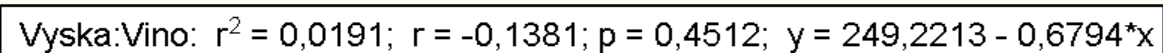
Bodový graf z Prijem proti Vyska

22Lide.sta 13v*32c

Prijem = -18846.0373+267.3417*x; 0.95 Int.spol.



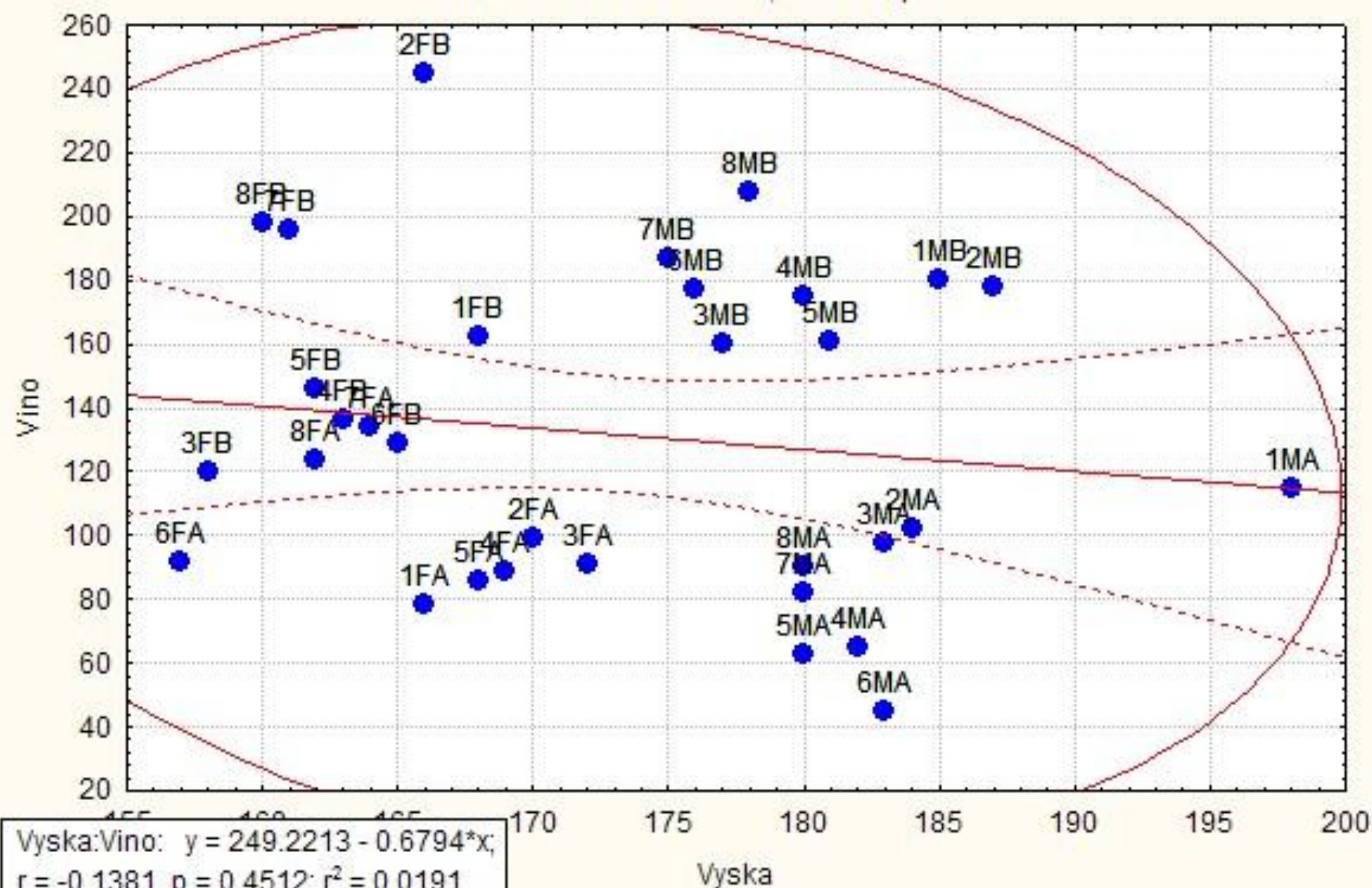
- Bodový graf (22Lide 13v)
- Bodový graf (22Lide 13v)
- Bodový graf (22Lide 13v)
- Bodový graf (22Lide 13v)
- Bodový graf (22Lide 13v)
- Bodový graf (22Lide 13v)

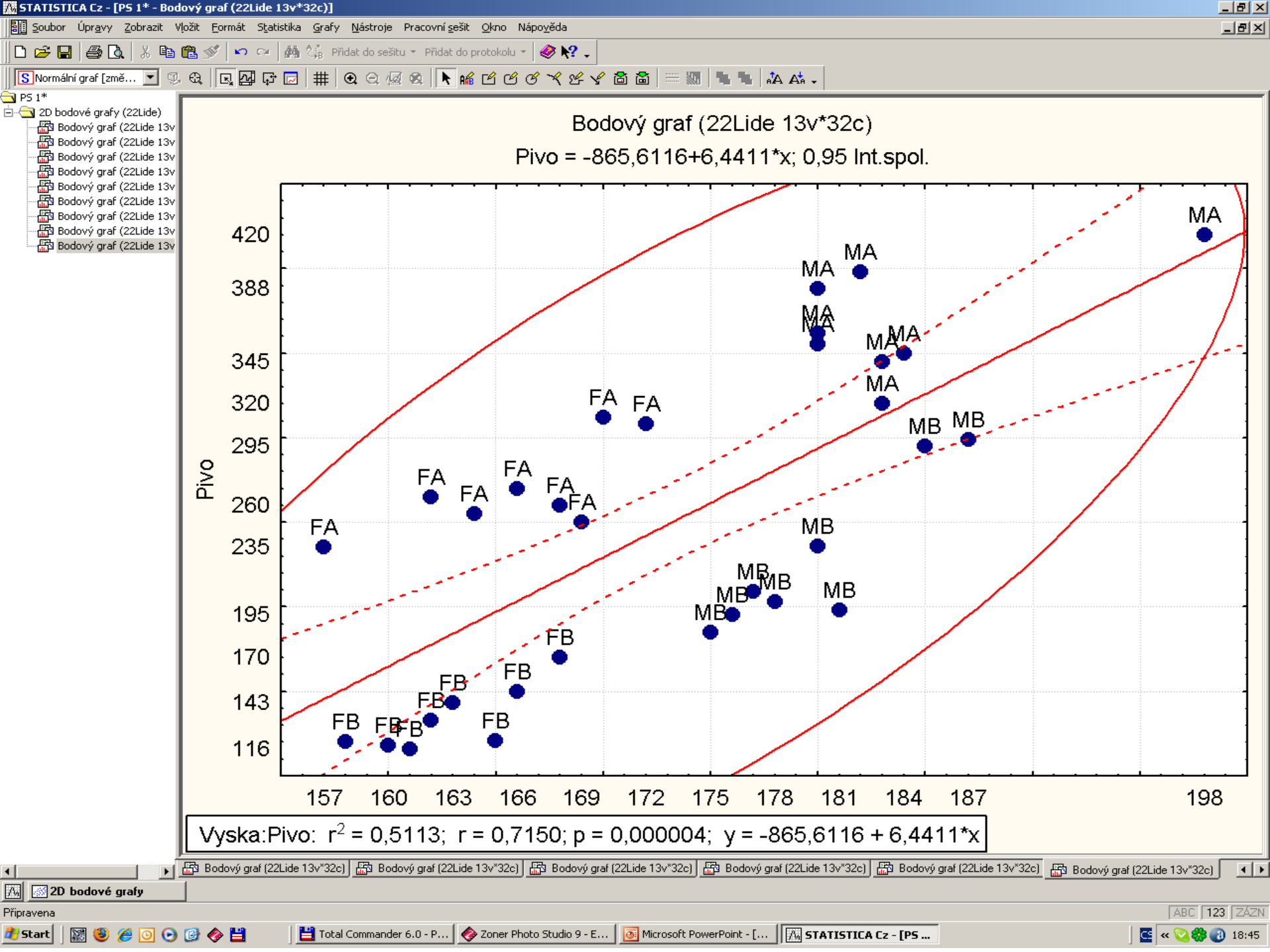


Bodový graf z Vino proti Vyska

22Lide.sta 13v*32c

Vino = 249.2213-0.6794*x; 0.95 Int.spol.

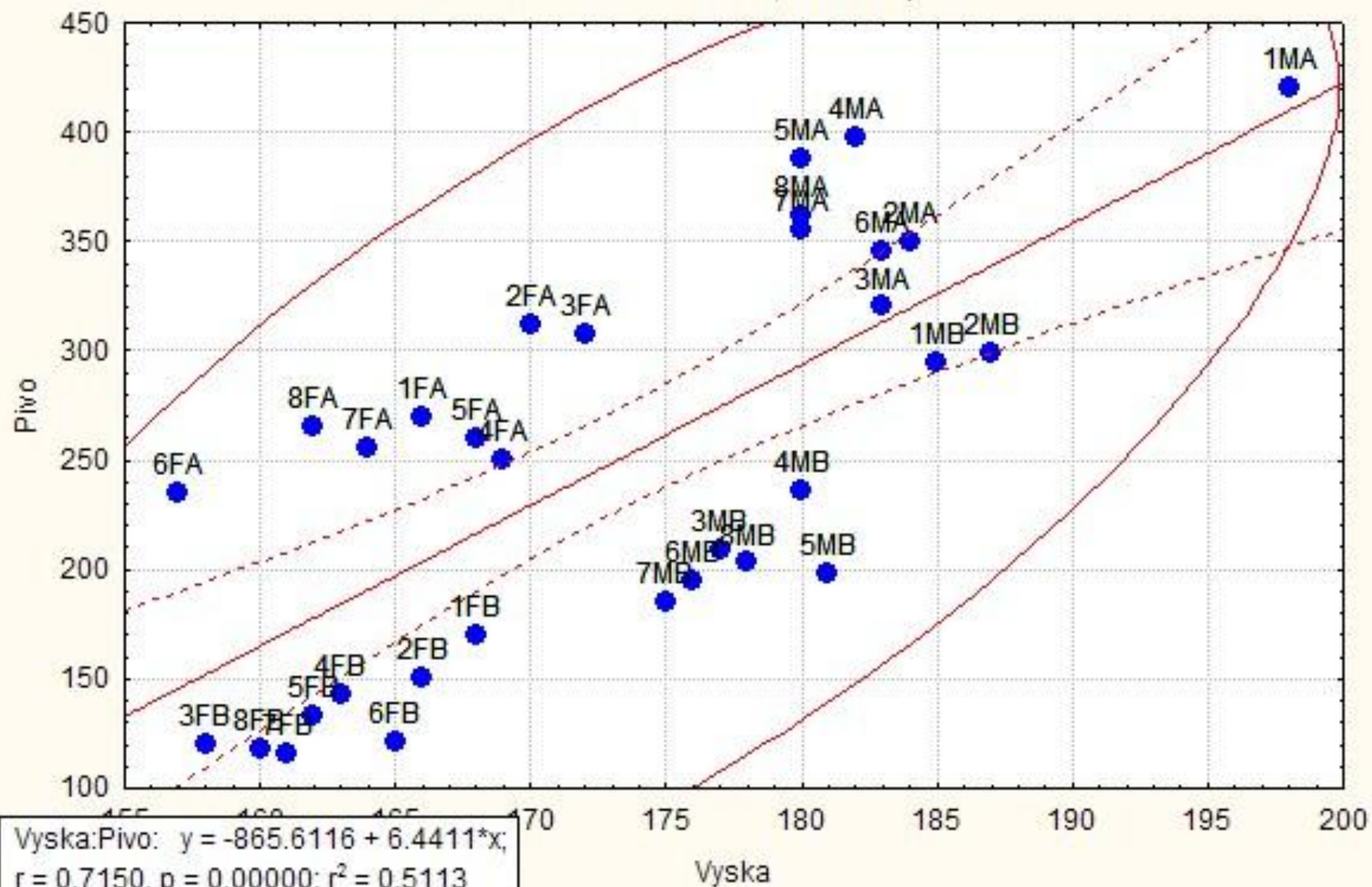


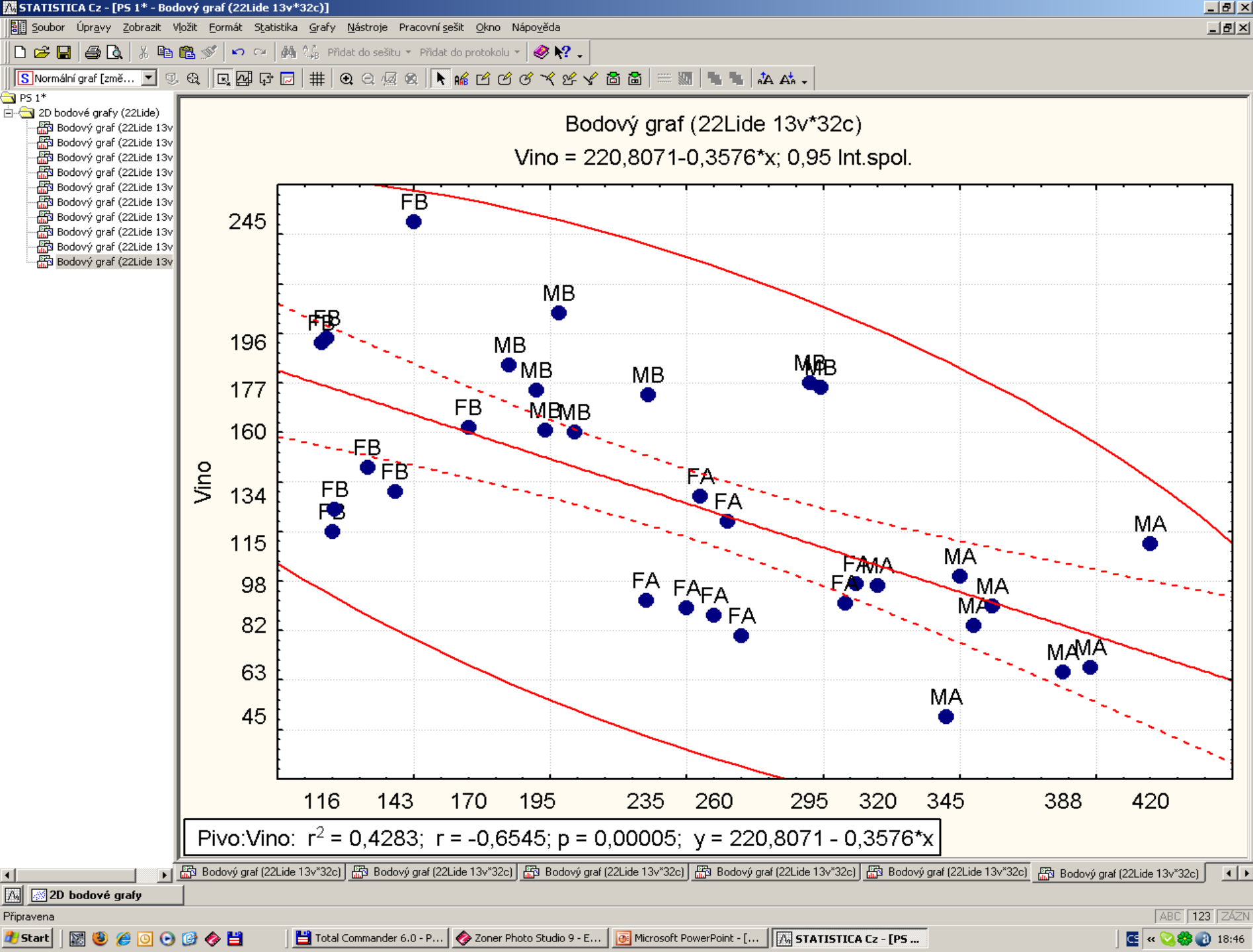


Bodový graf z Pivo proti Vyska

22Lide.sta 13v*32c

Pivo = -865.6116+6.4411*x; 0.95 Int.spol.

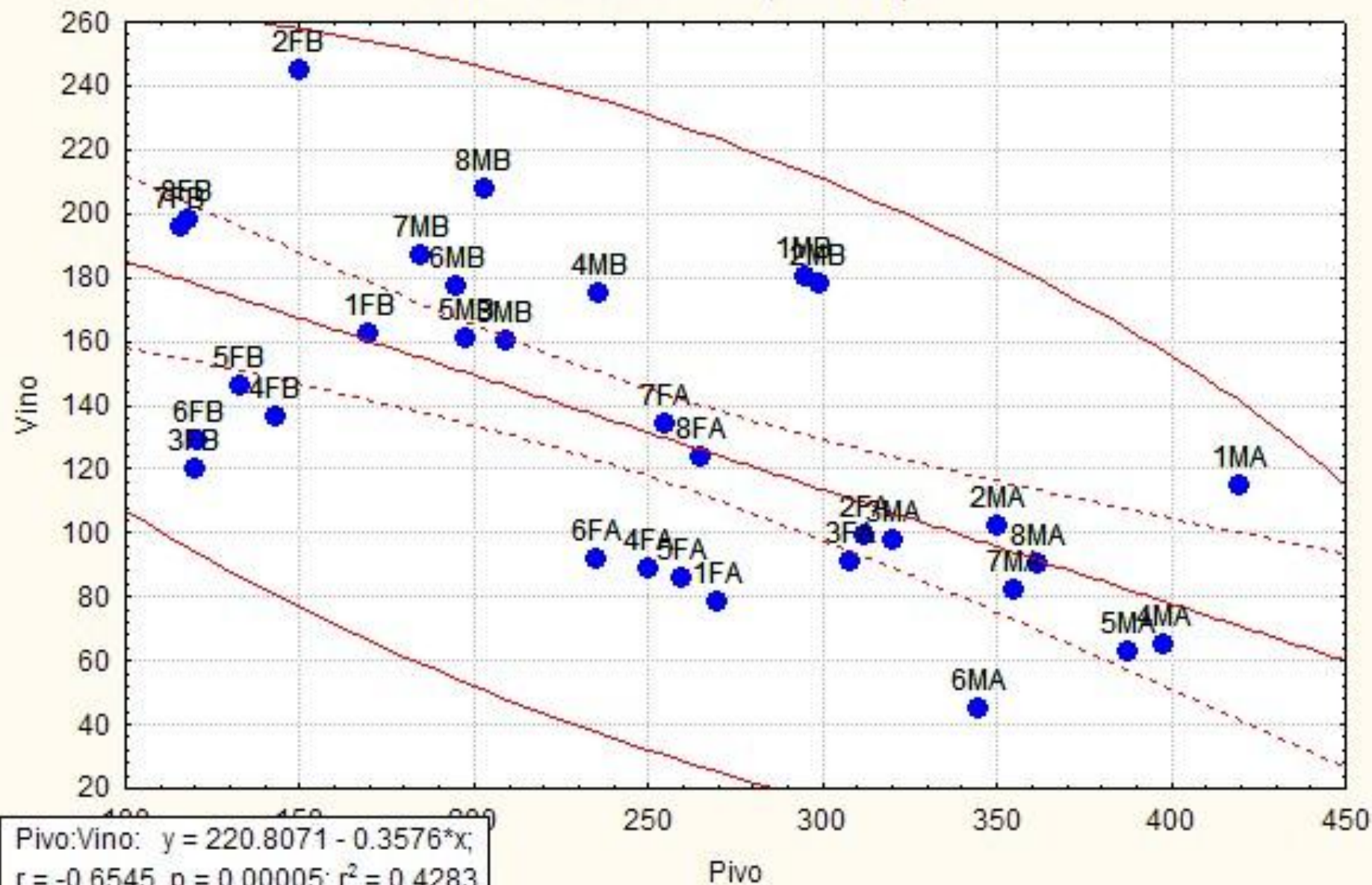




Bodový graf z Vino proti Pivo

22Lide.sta 13v*32c

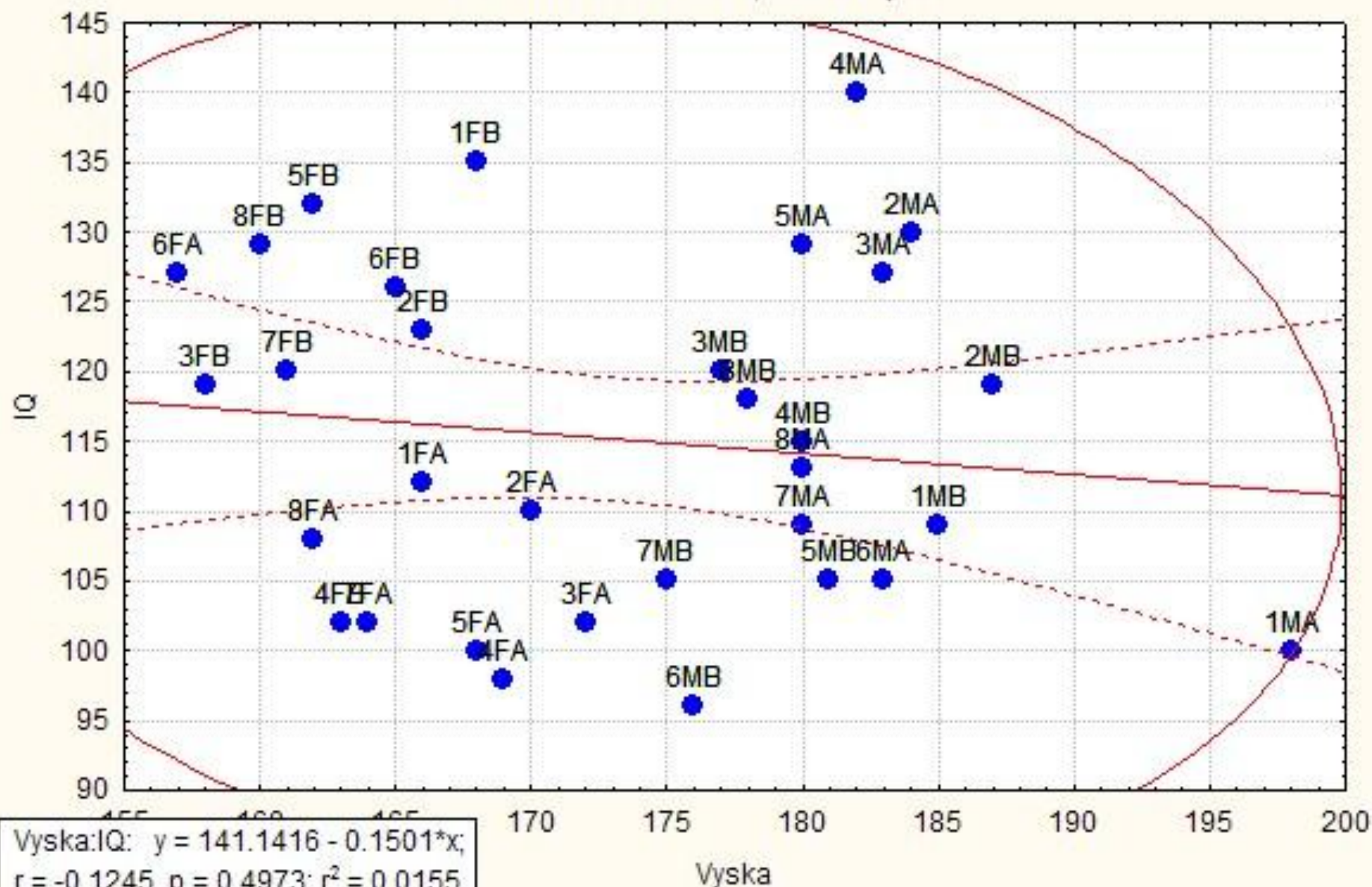
Vino = 220.8071-0.3576*x; 0.95 Int.spol.



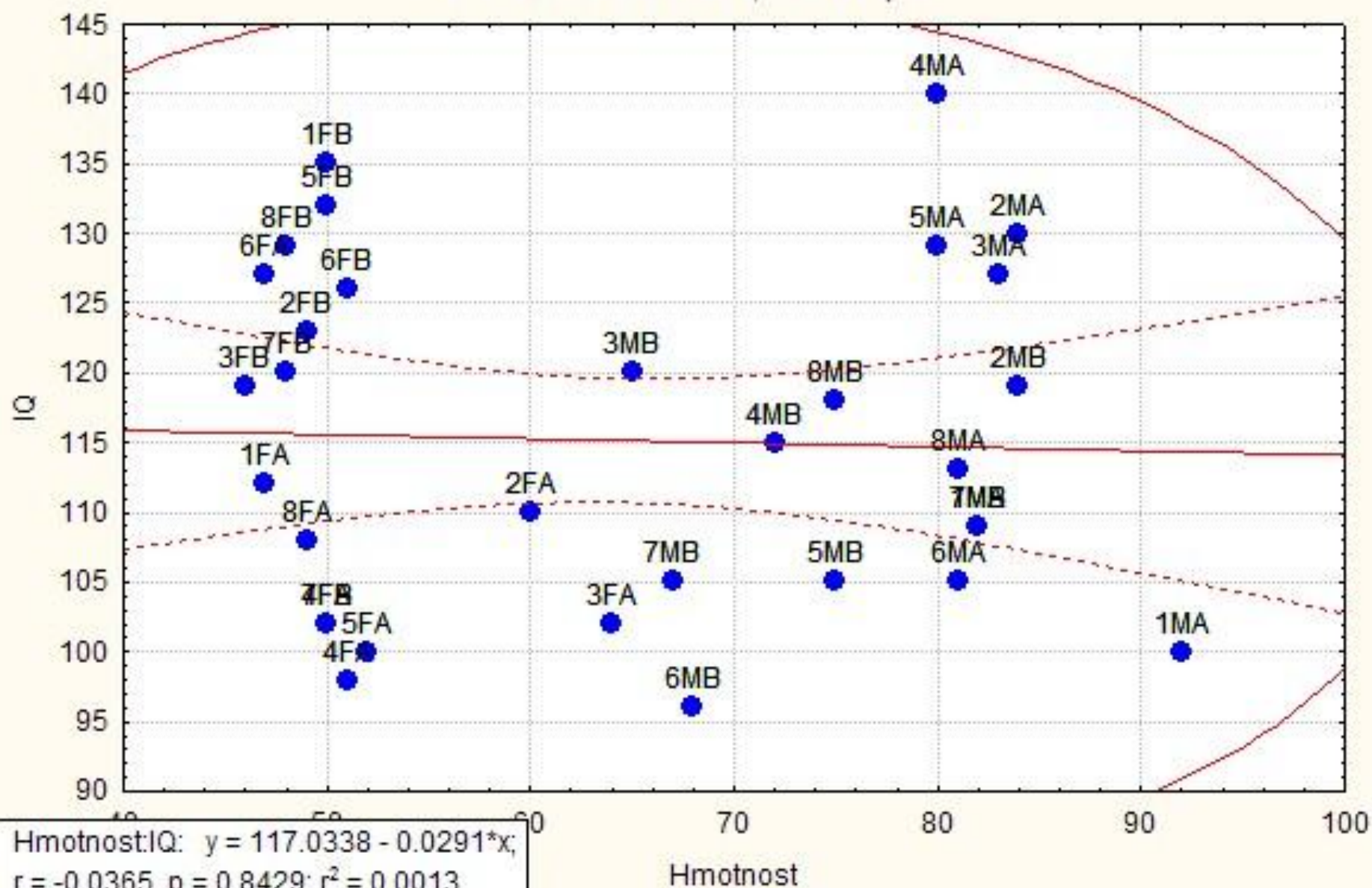
Bodový graf z IQ proti Vyska

22Lide.sta 13v*32c

IQ = 141.1416-0.1501*x; 0.95 Int.spol.



Bodový graf z IQ proti Hmotnost
 22Lide.sta 13v*32c
 $IQ = 117.0338 - 0.0291 \cdot x$; 0.95 Int.spol.



 PS 1*

2D bodové grafy (22Lide)

 Bodový graf (22Lide 13v)

 Bodový graf (22Lide 13v)

 Bodový graf (22Lide 13v)

 Bodový graf (22Lide 13v)

 Bodový graf (22Lide 13v)

 Bodový graf (22Lide 13v)

 Bodový graf (22Lide 13v)

 Bodový graf (22Lide 13v)

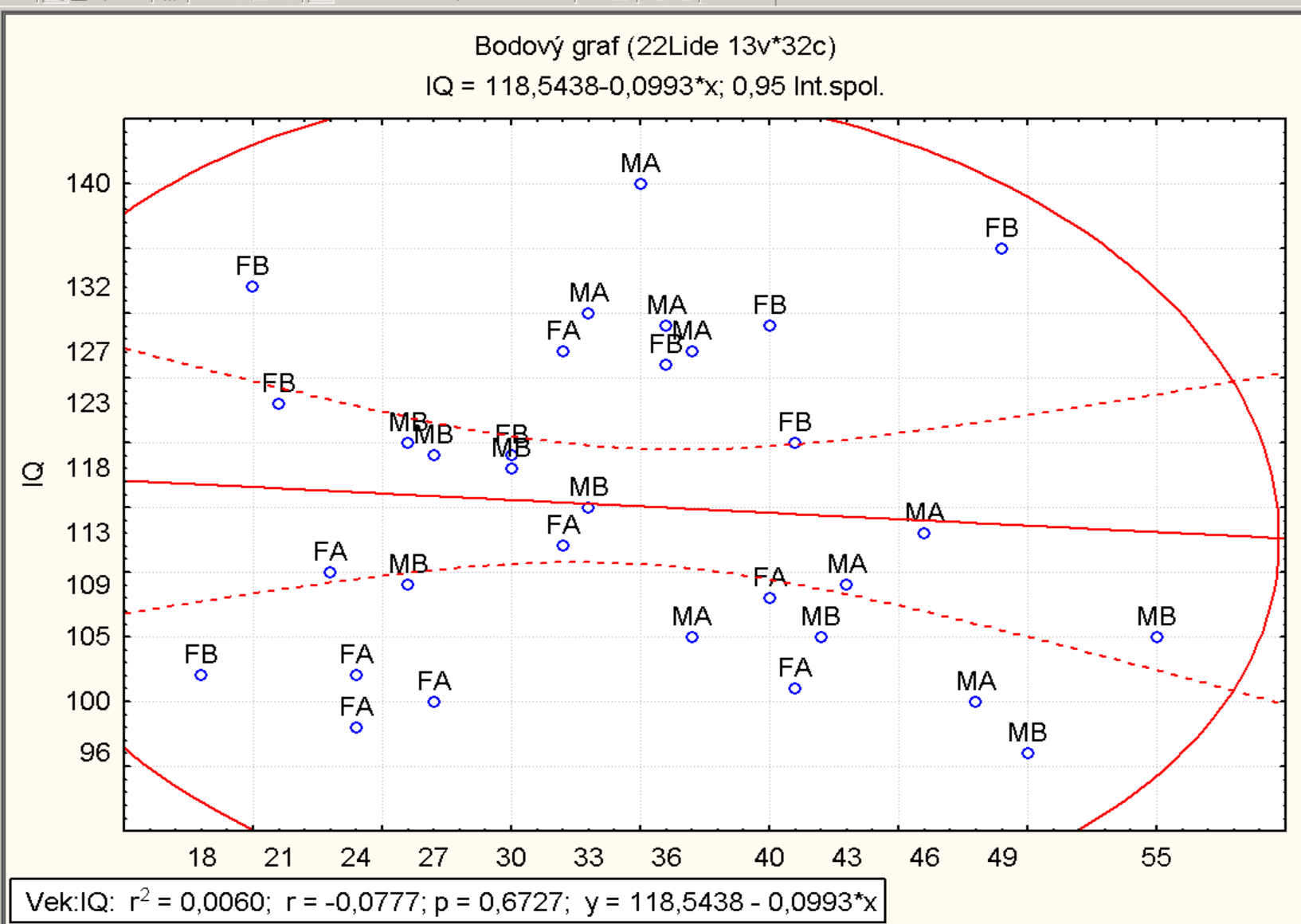
 Bodový graf (22Lide 13v)

- 3D bodové grafy (22Lide)

	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1
	3D bodový graf (22Lide 1

2D bodové grafy (22Lide)

Bodový graf (22Lide 13v
Bodový graf (22Lide 13v
Bodový graf (22Lide 13v
Bodový graf (22Lide 13v

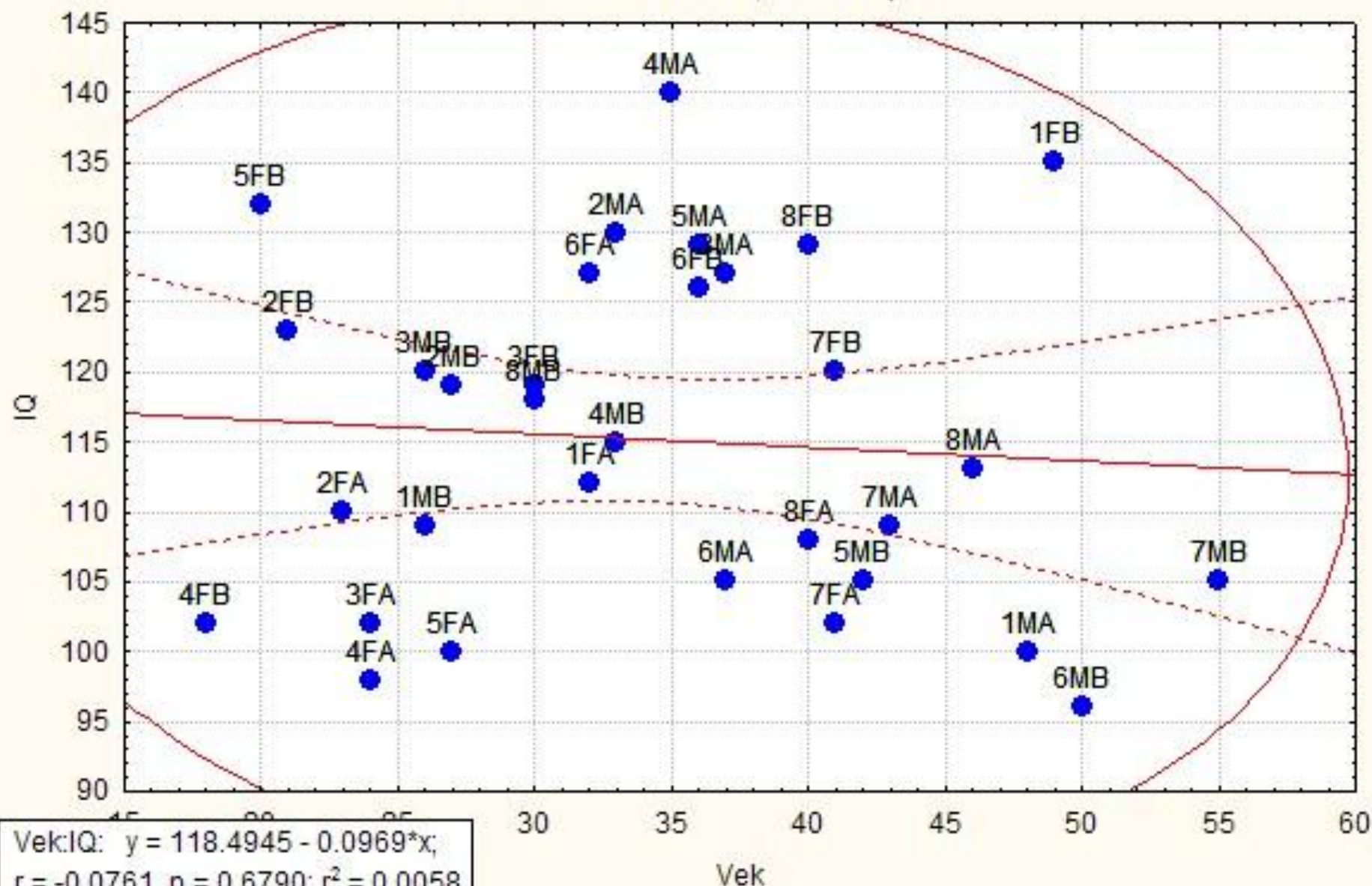


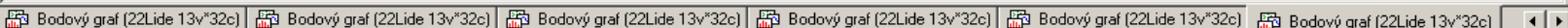
2D bodové grafy

Bodový graf z IQ proti Vek

22Lide.sta 13v*32c

IQ = 118.4945-0.0969*x; 0.95 Int.spol.

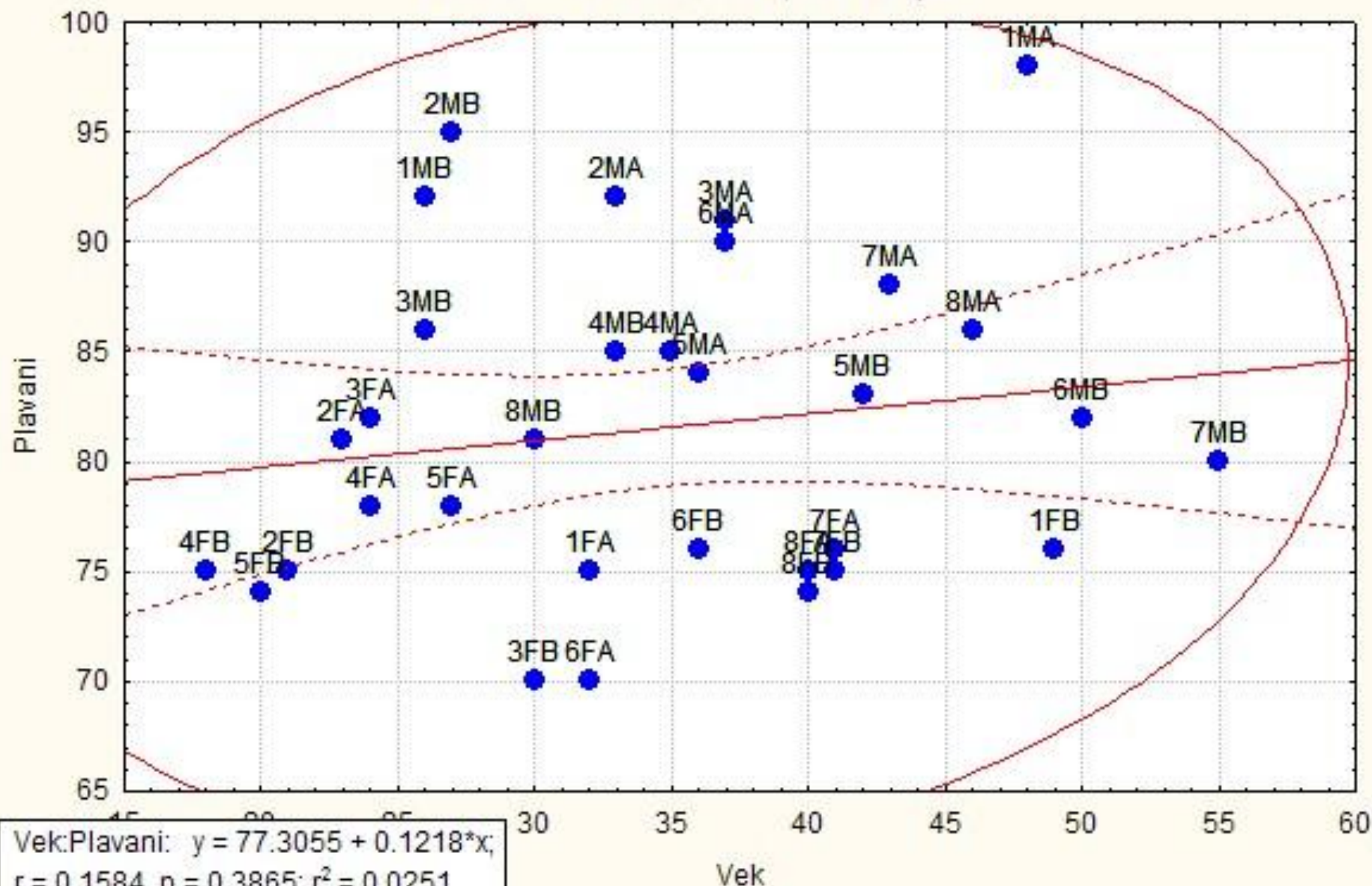




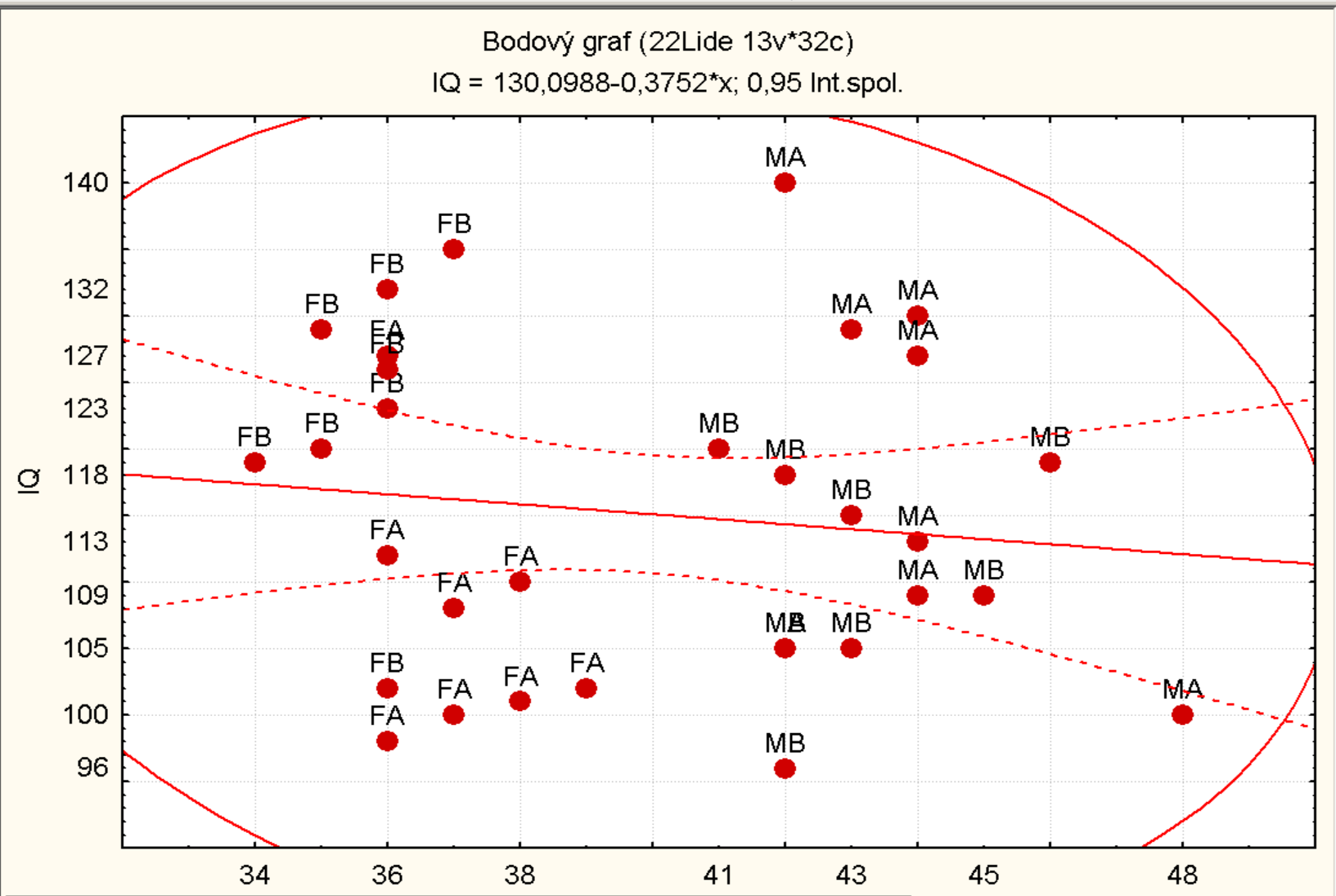
Bodový graf z Plavani proti Vek

22Lide.sta 13v*32c

Plavani = $77.3055 + 0.1218 \cdot x$; 0.95 Int.spol.



Vek:Plavani: $y = 77.3055 + 0.1218 \cdot x$;
 $r = 0.1584$, $p = 0.3865$; $r^2 = 0.0251$

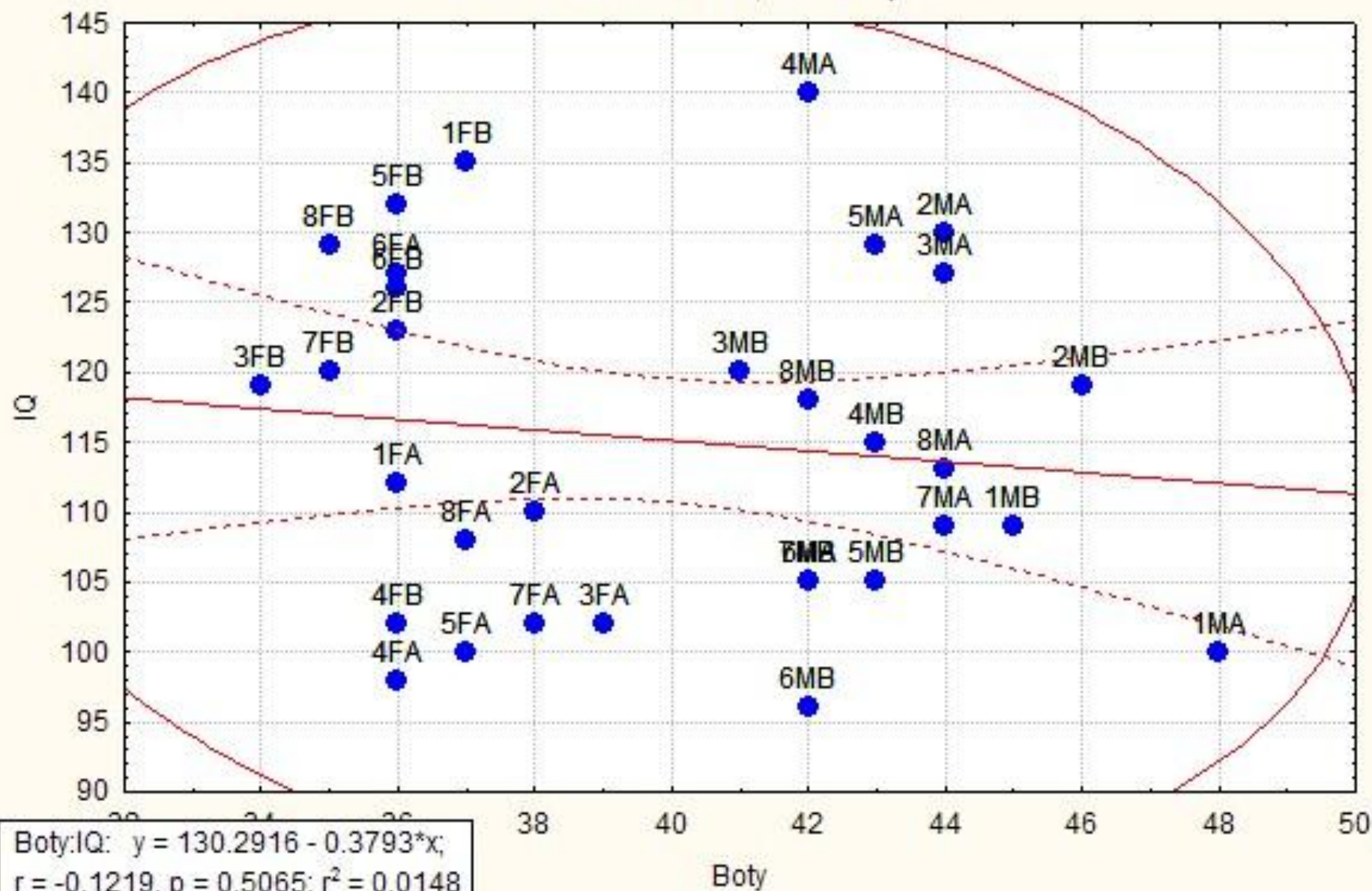


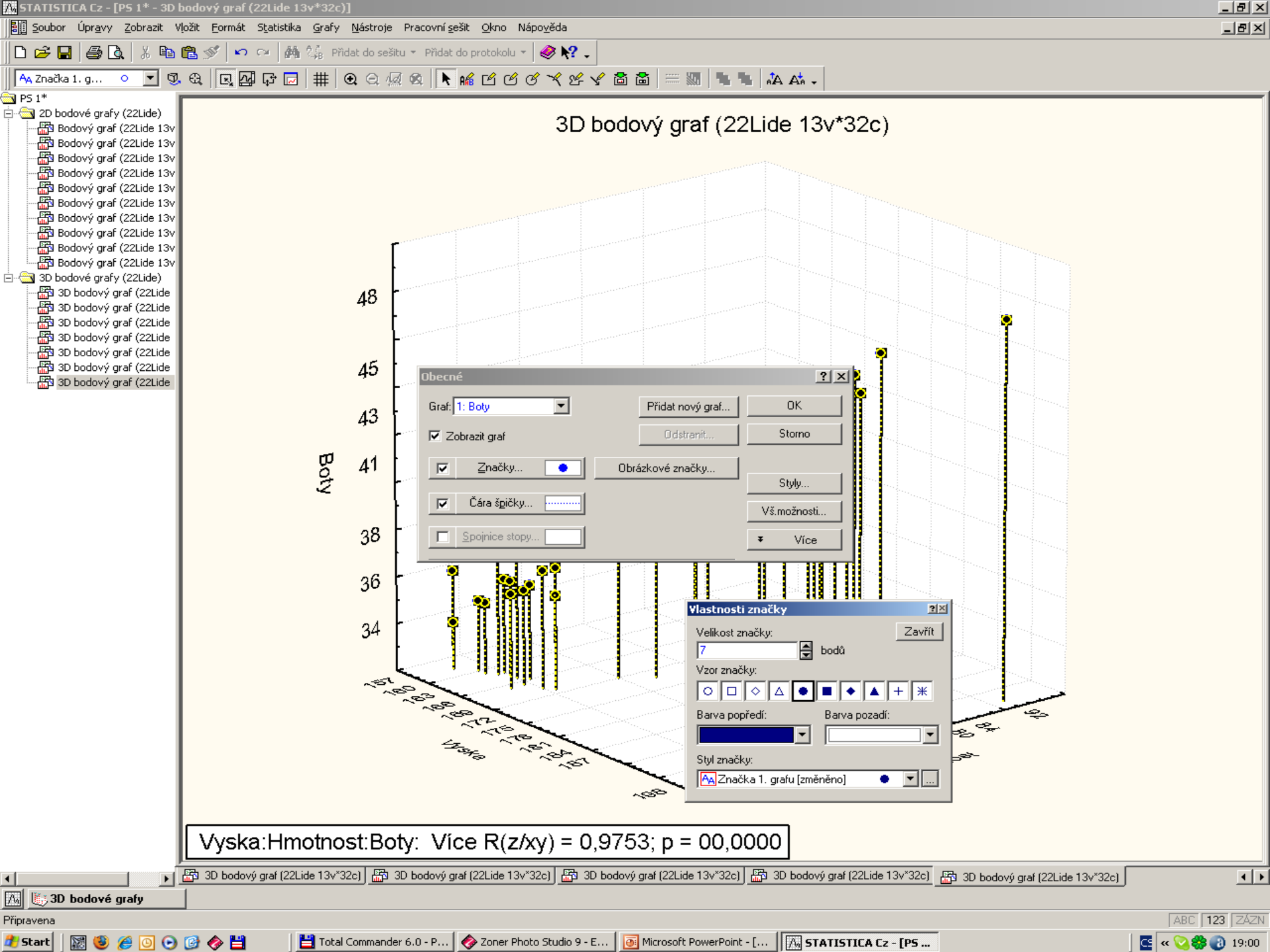
Boty:IQ: $r^2 = 0,0144$; $r = -0,1202$; $p = 0,5123$; $y = 130,0988 - 0,3752 \cdot x$

Bodový graf z IQ proti Boty

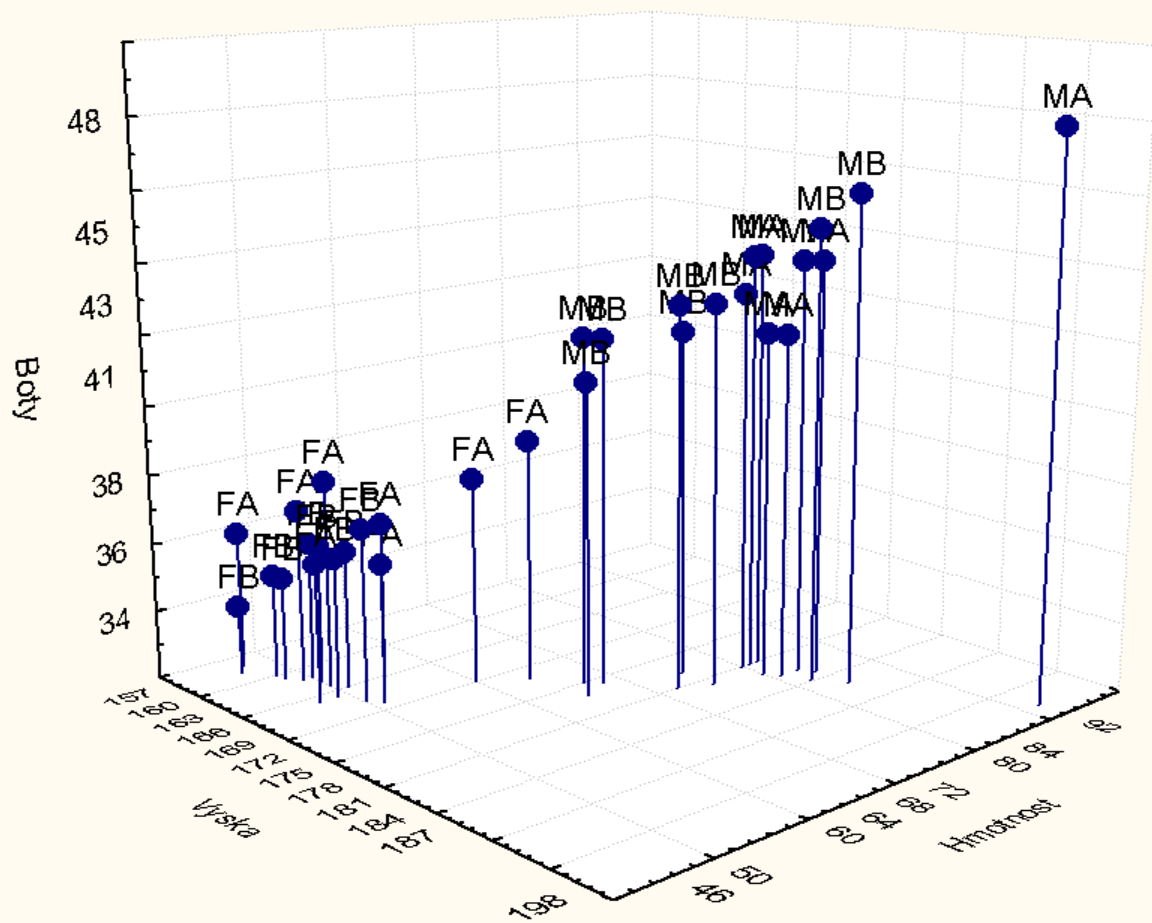
22Lide.sta 13v*32c

IQ = 130.2916-0.3793*x; 0.95 Int.spol.

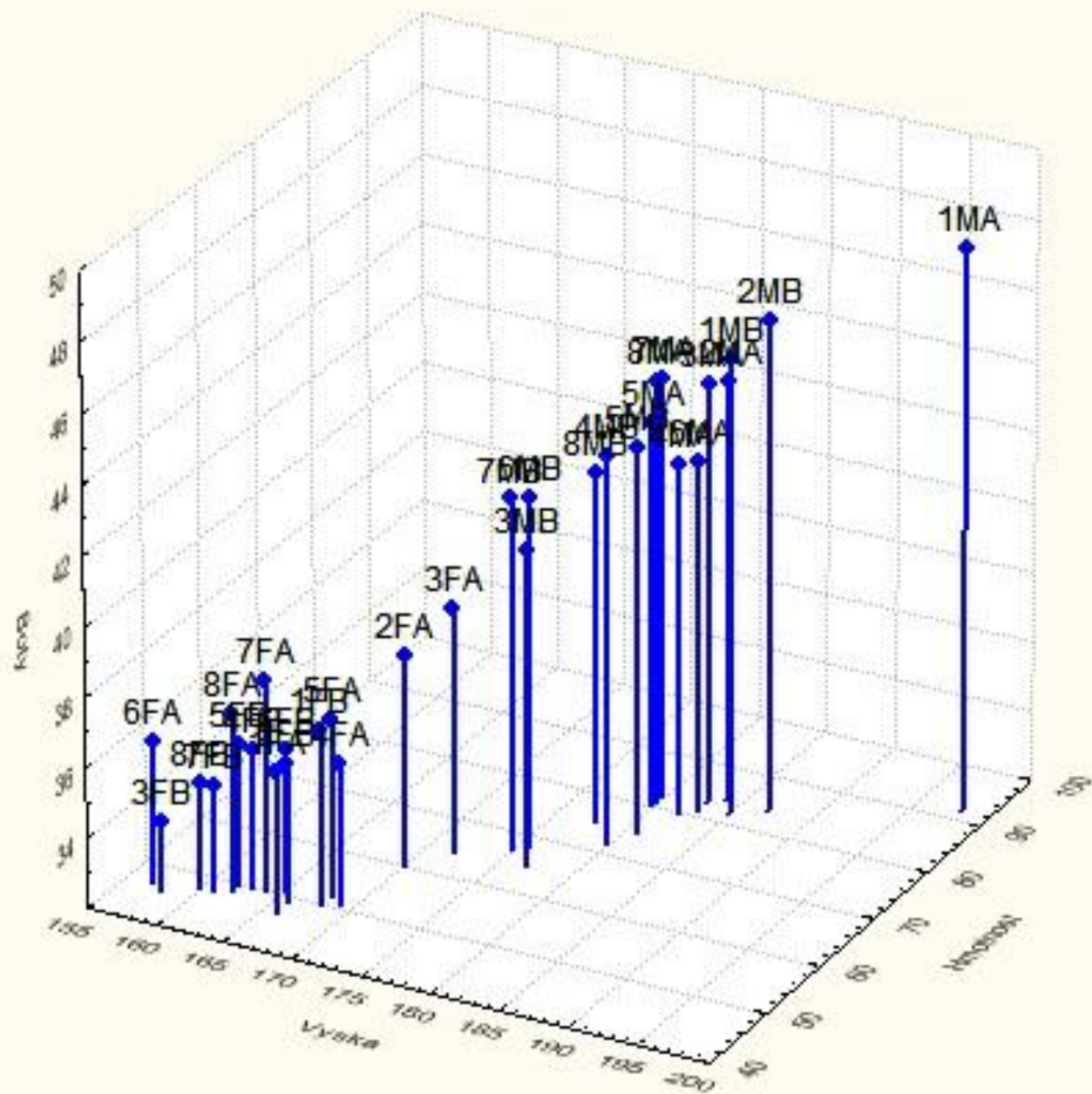


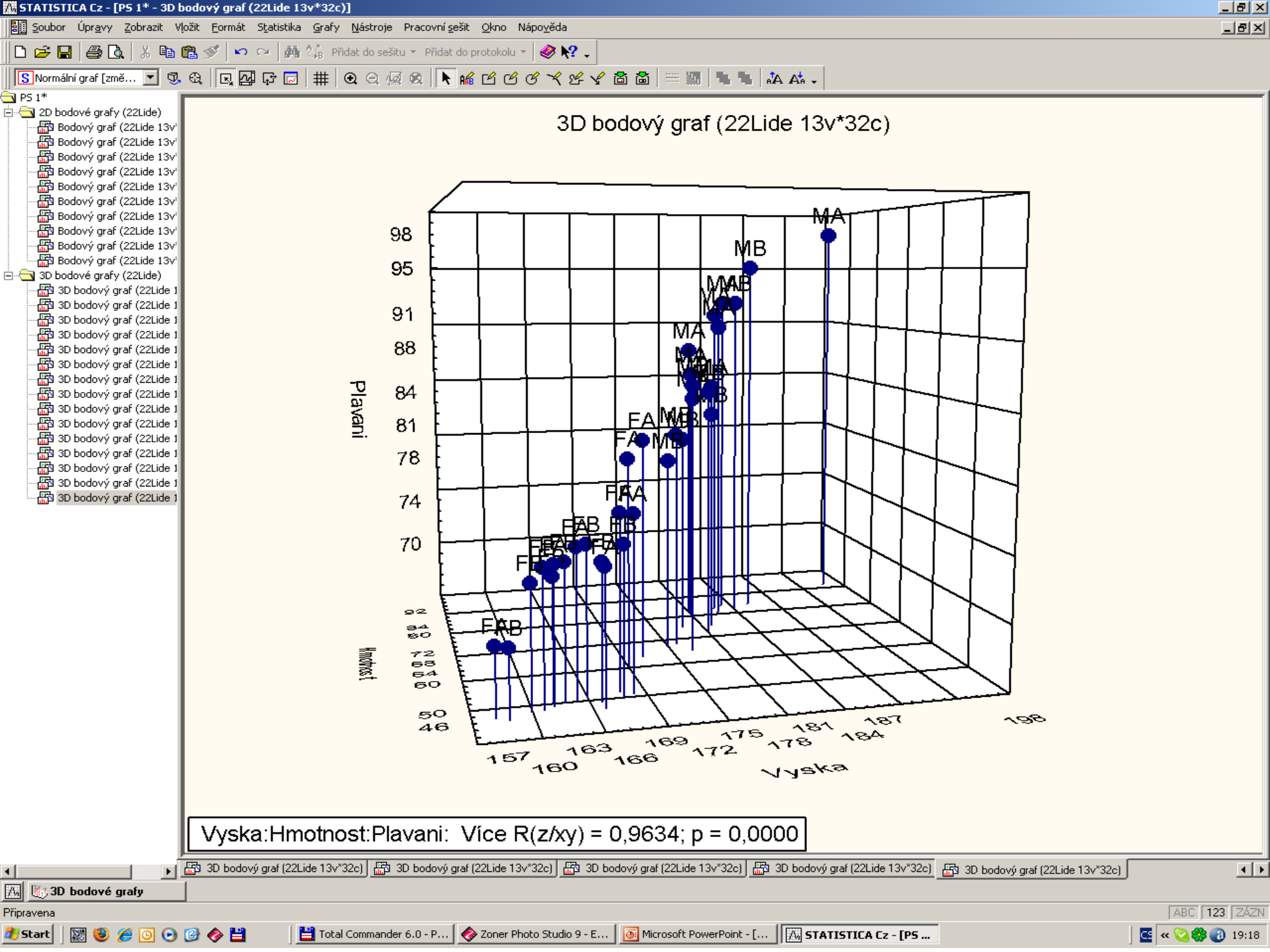


3D bodový graf (22Lide 13v*32c)



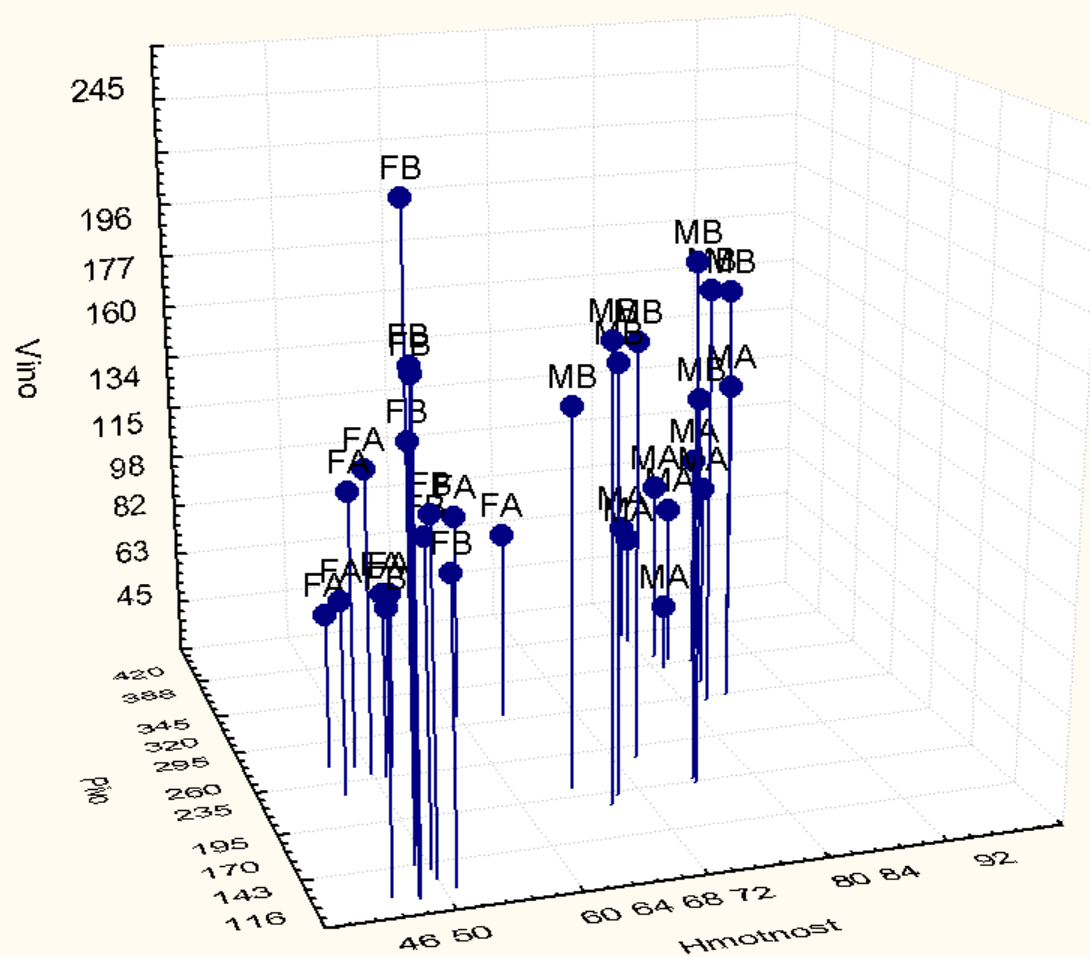
Vyska:Hmotnost:Boty: Více $R(z/xy) = 0,9753$; $p = 00,0000$





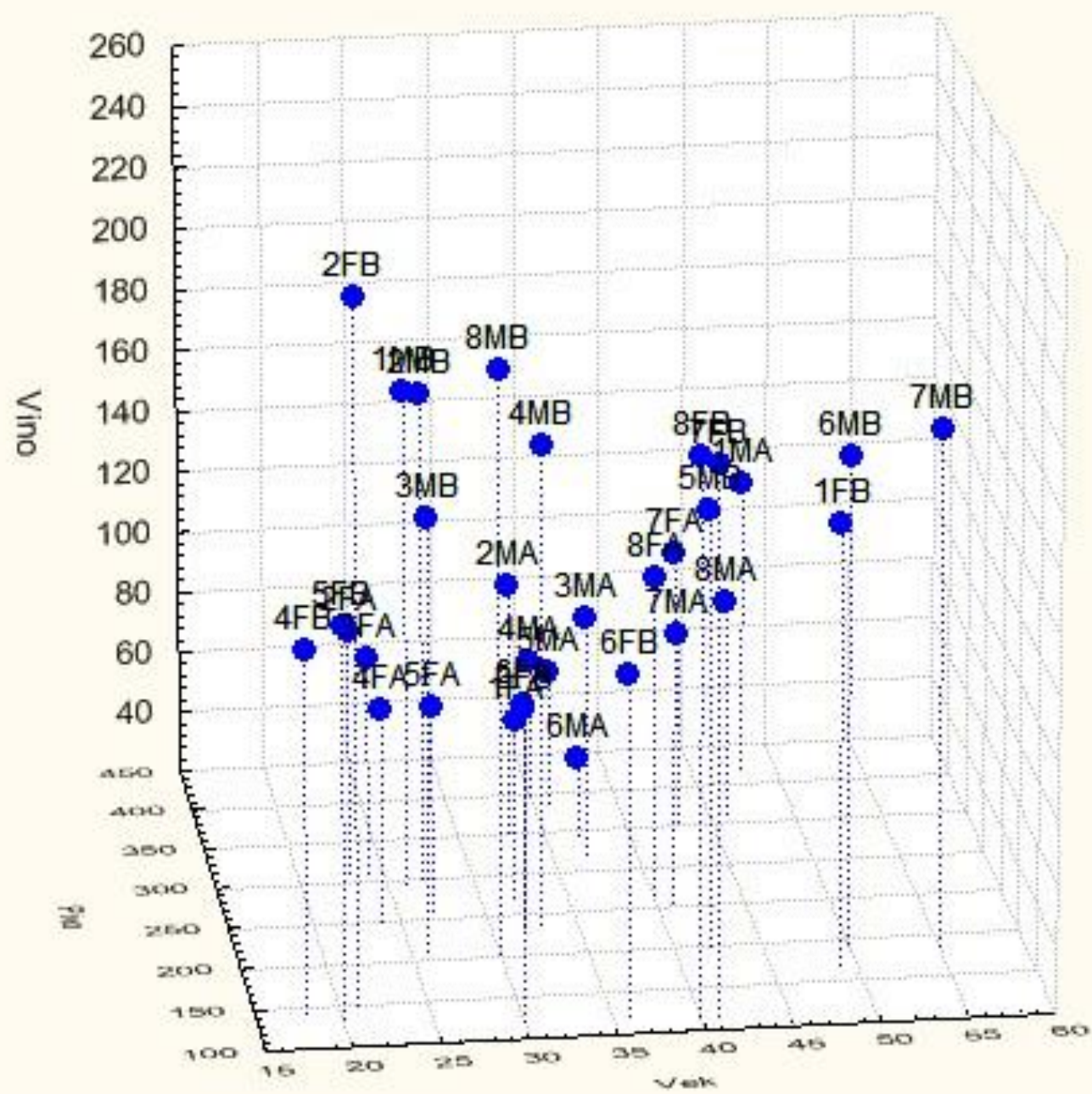
- [illegible]

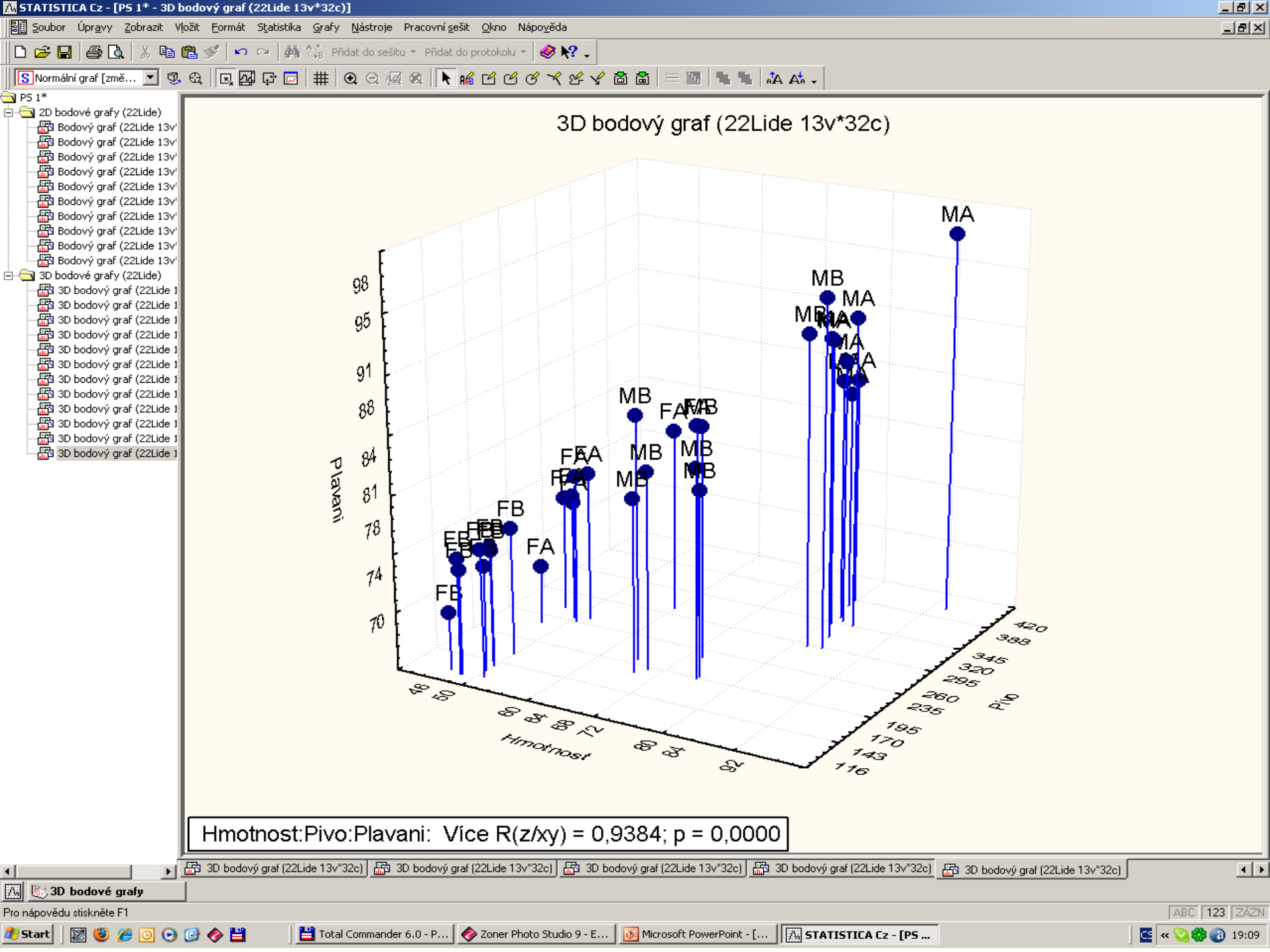
3D bodový graf (22Lide 13v*32c)

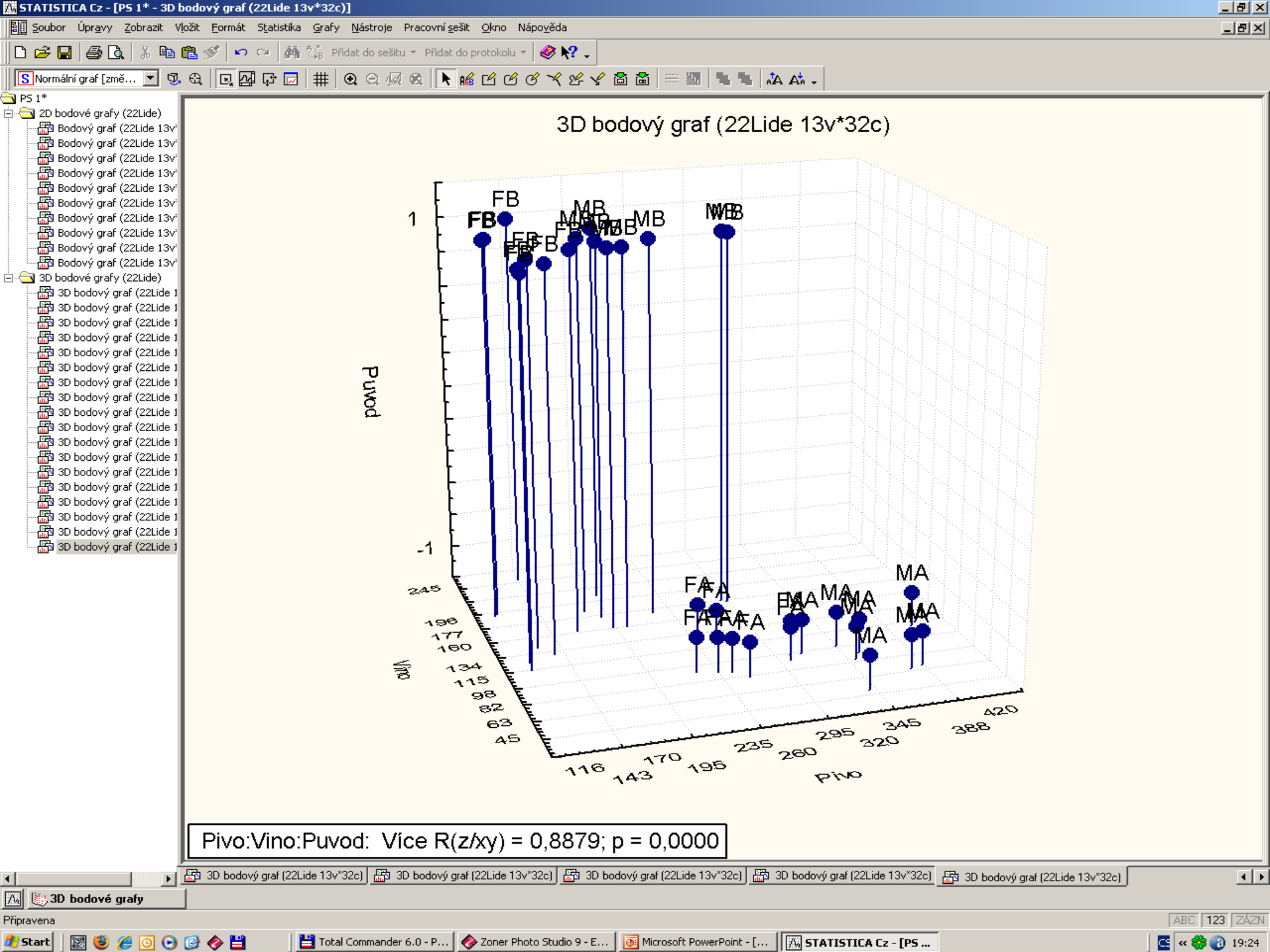


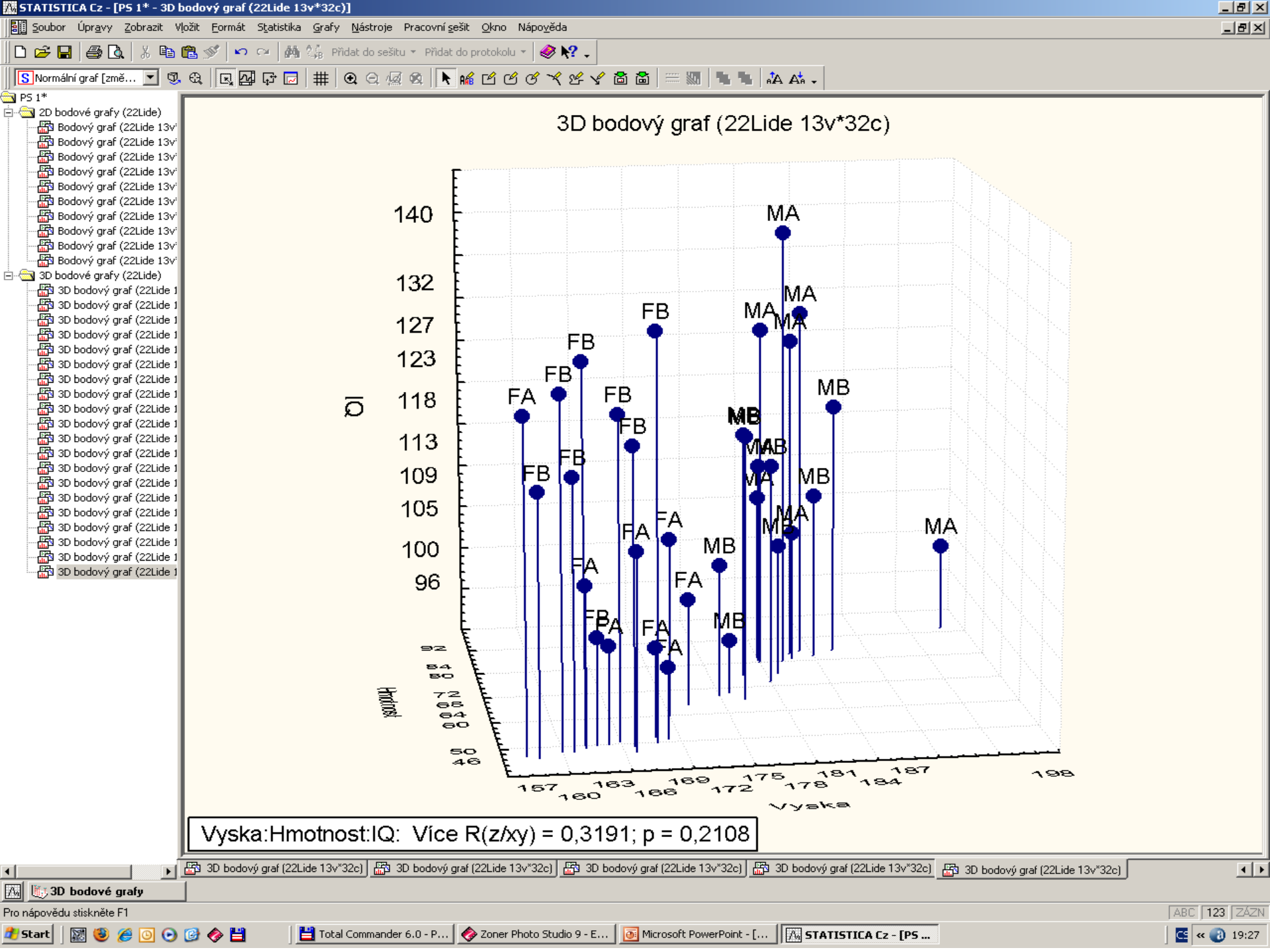
Hmotnost:Pivo:Vino: $Vice\ R(z/xy) = 0,7801; p = 0,000001$

Vek:Pivo:Vino: Více $R(z/xy) = 0,6639$; $p = 0,0002$

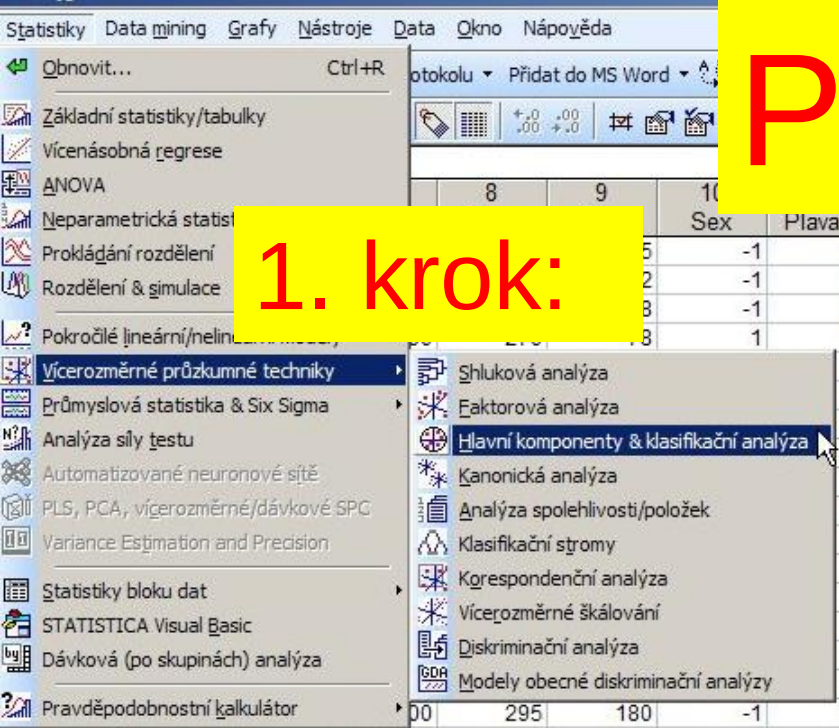




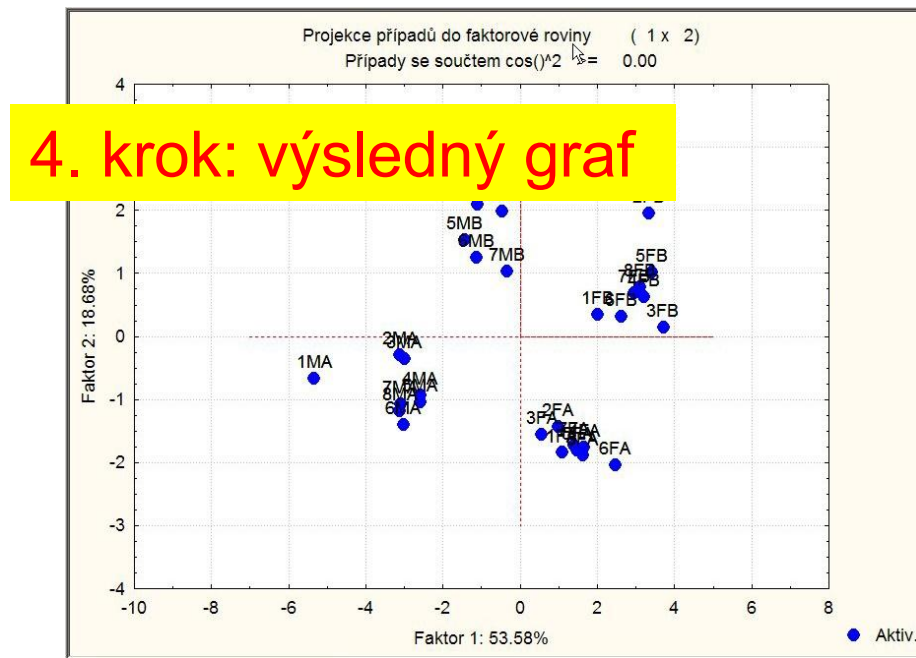
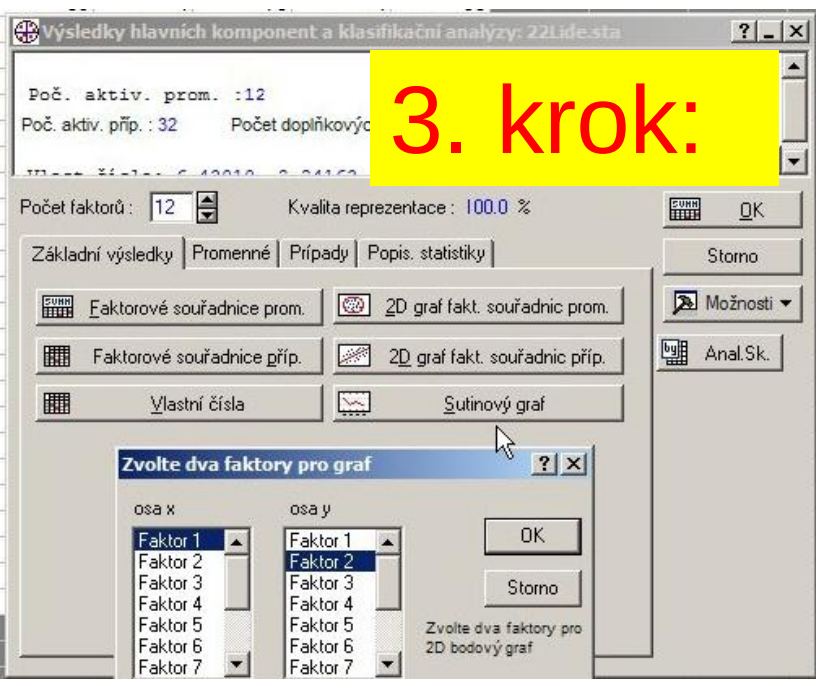
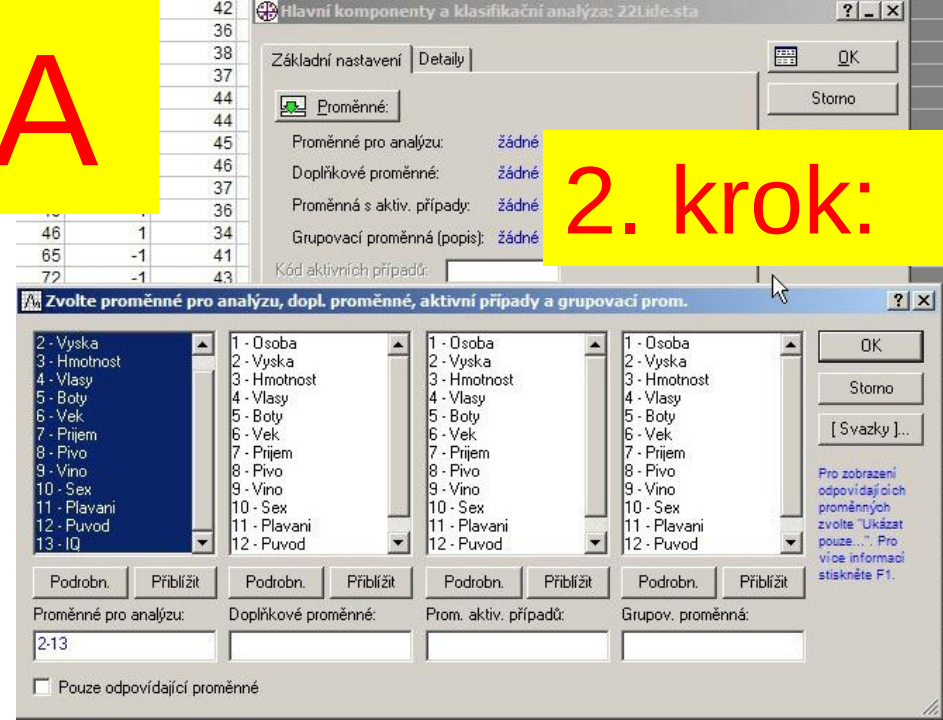


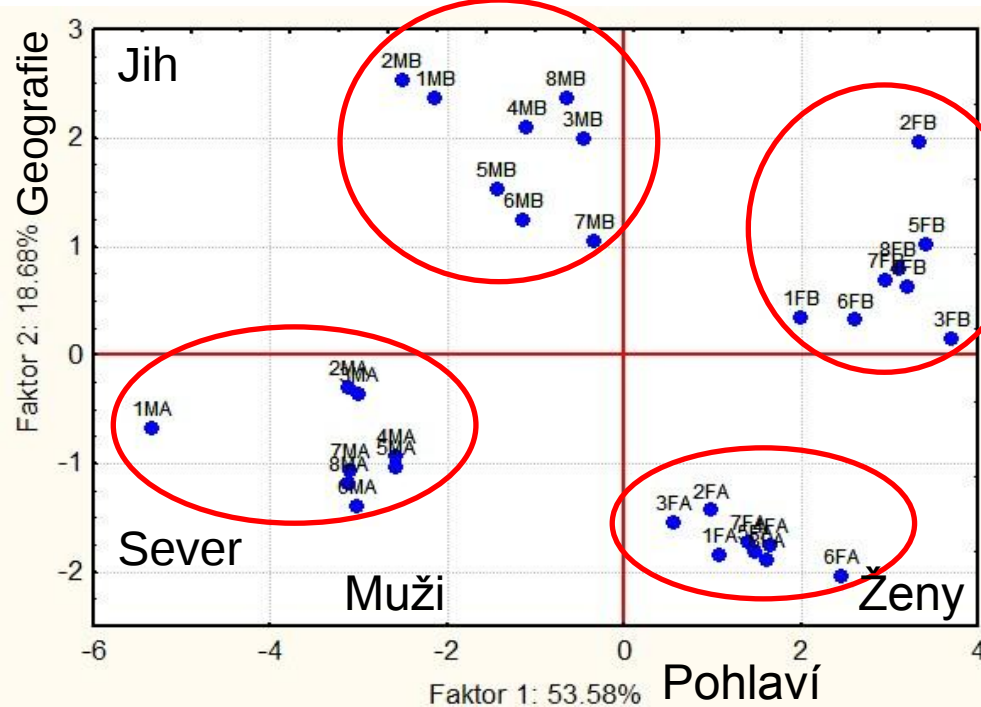


3D bodový graf (22Lide 13v*32c) 3D bodový graf (22Lide 13v*32c) 3D bodový graf (22Lide 13v*32c) 3D bodový graf (22Lide 13v*32c) 3D bodový graf (22Lide 13v*32c)

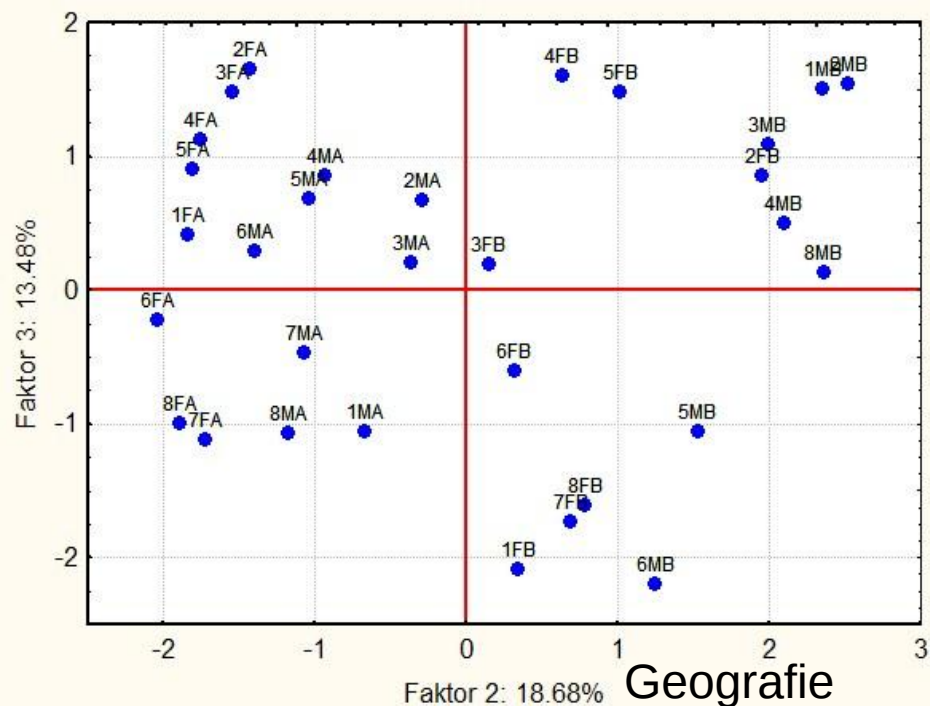
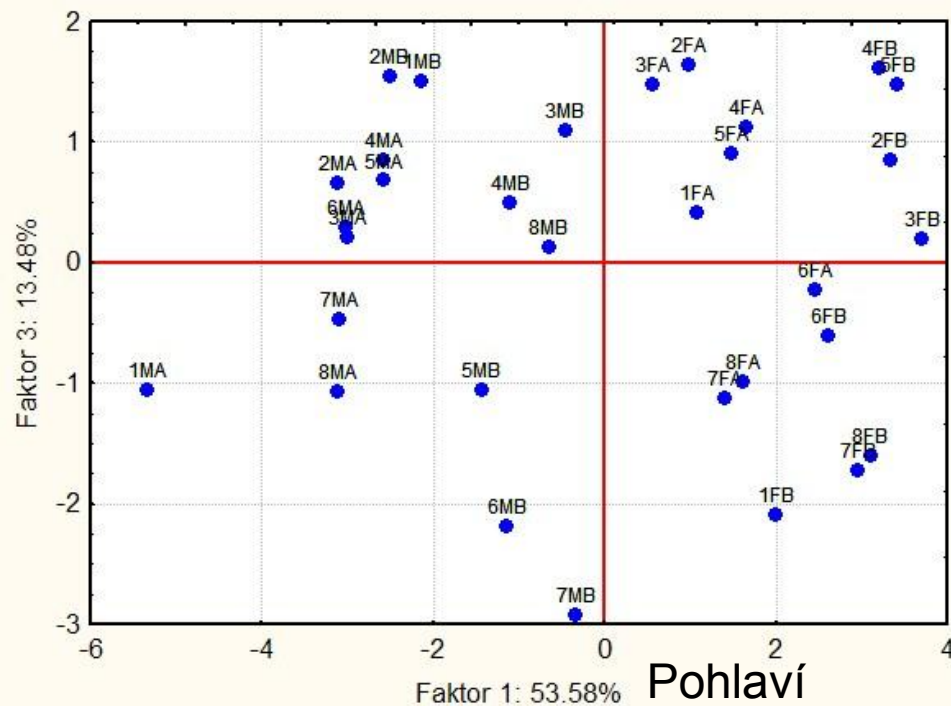


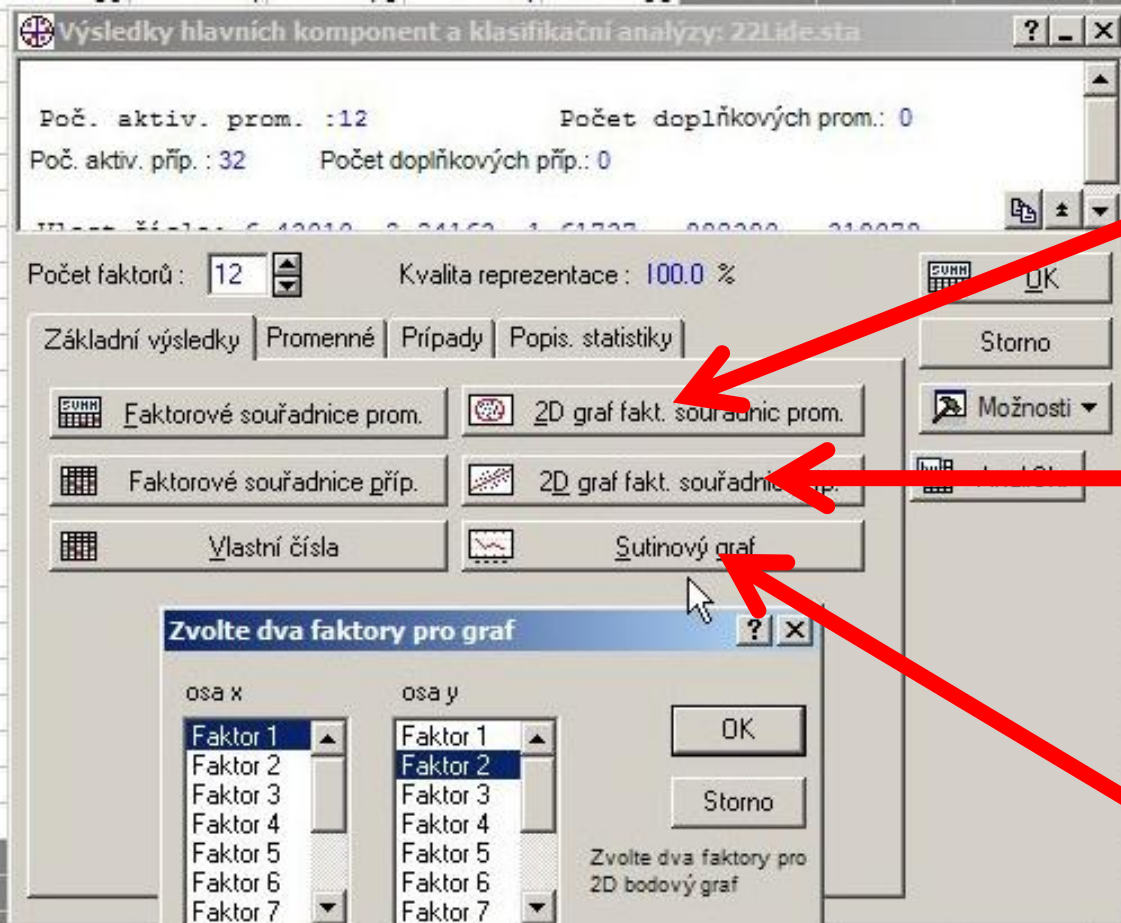
PCA





3 rozptylové
diagramy
komponentního
skóre , tj. grafy
1-2, 1-3, 2-3





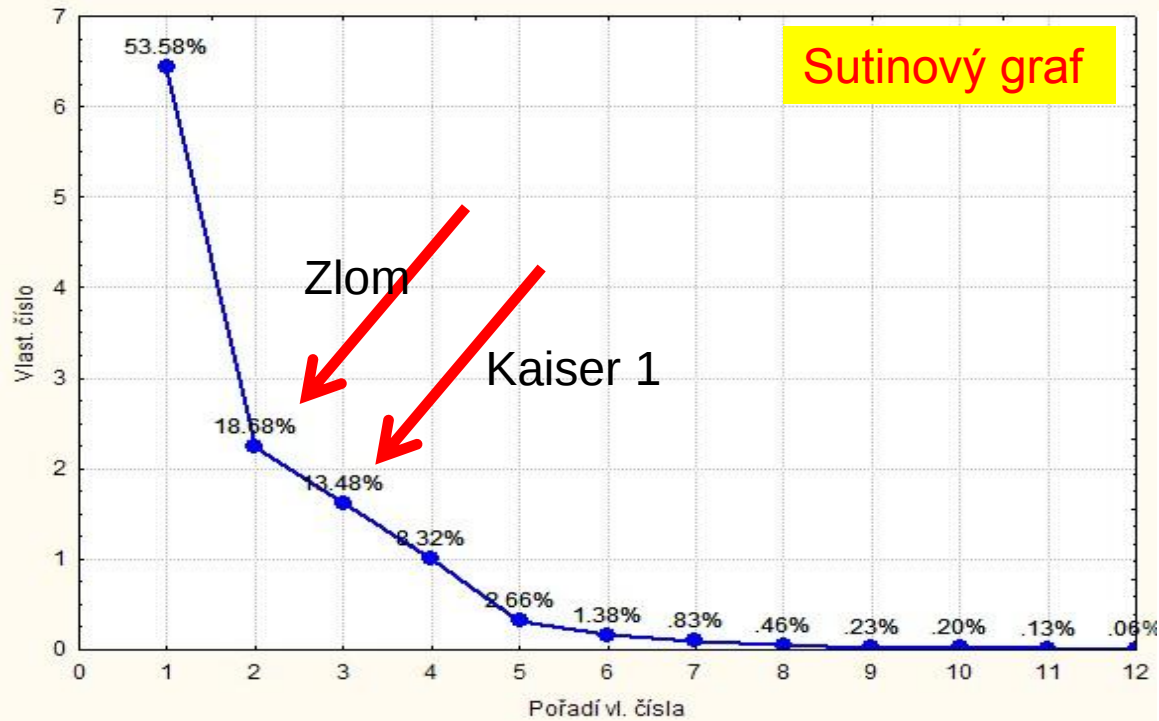
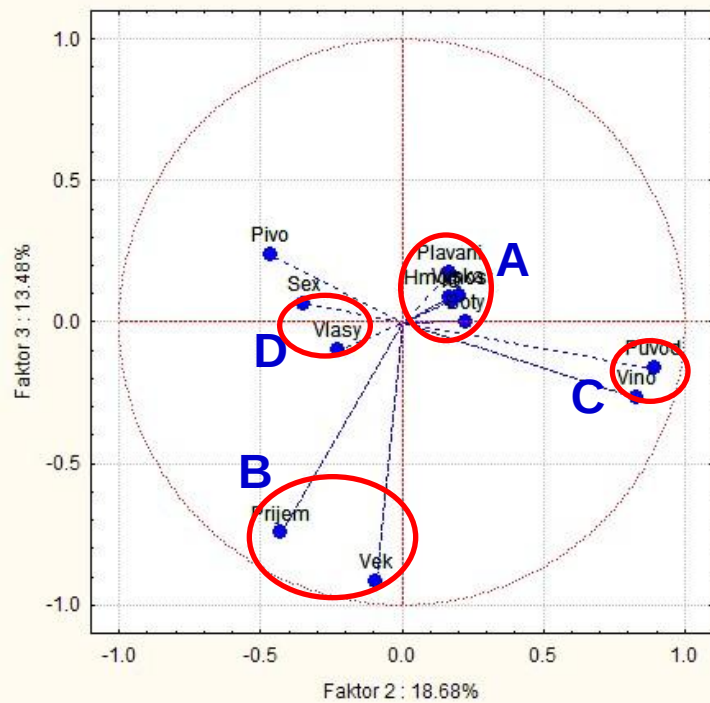
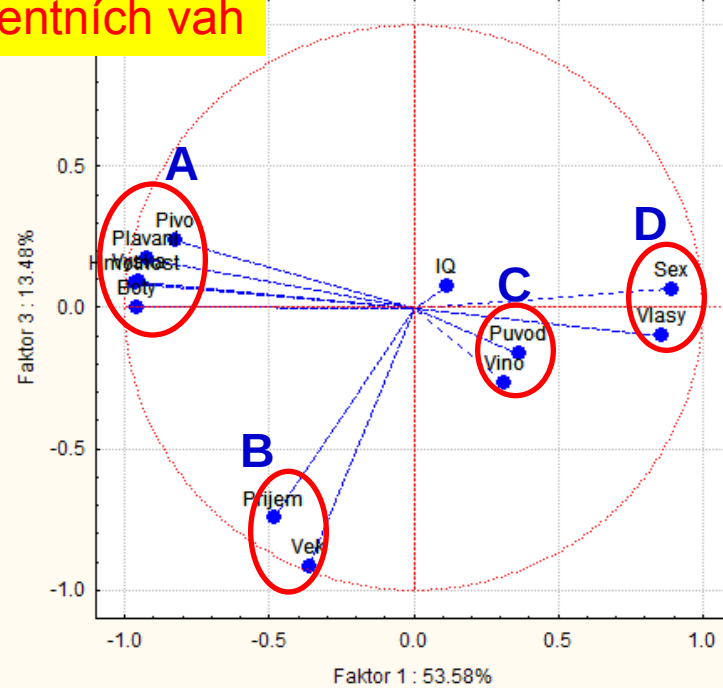
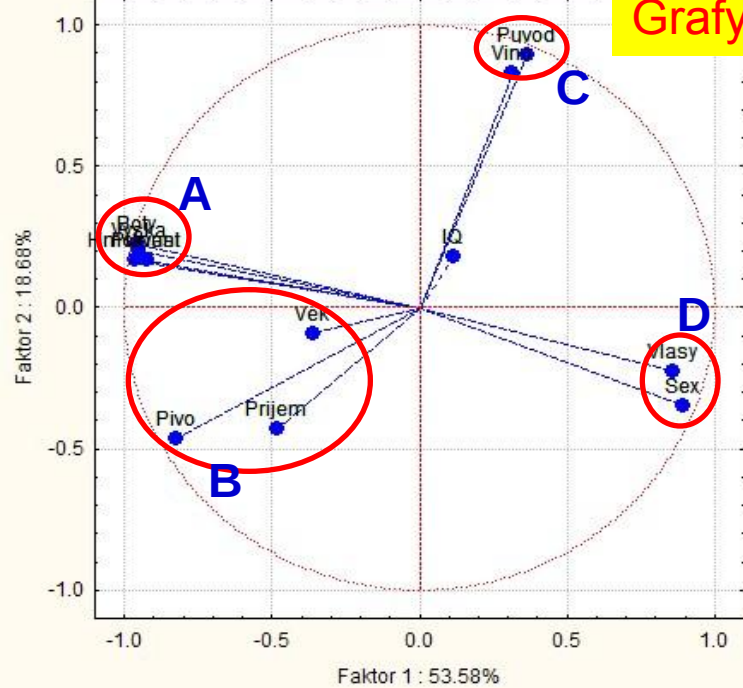
Volba tří grafů
komponentních vah

Volba tří rozptylových
diagramů
komponentního skóre

Volba Cattelova
indexového grafu
vlastních čísel
(=sutinového) grafu

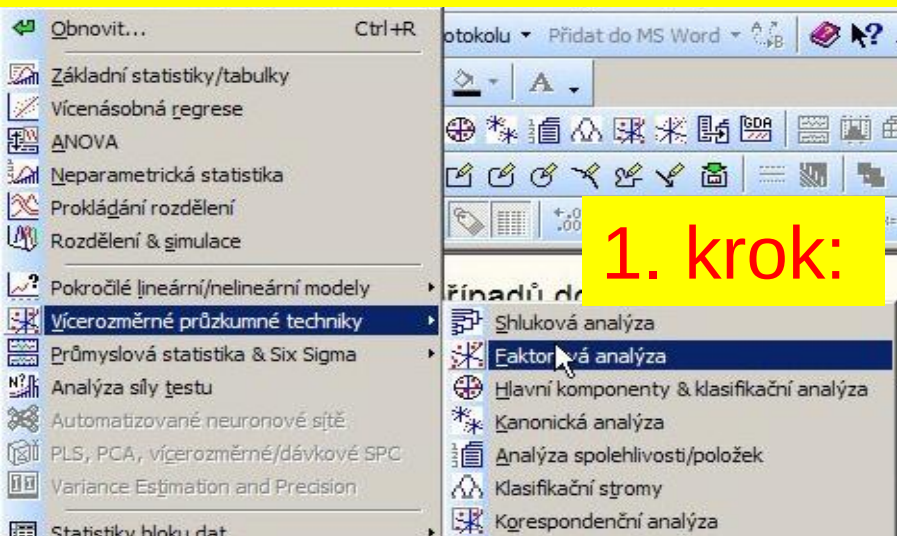
Hledání shluků podobných znaků v
grafech komponentních vah

Grafy komponentních vah



Faktorová analýza (FA):

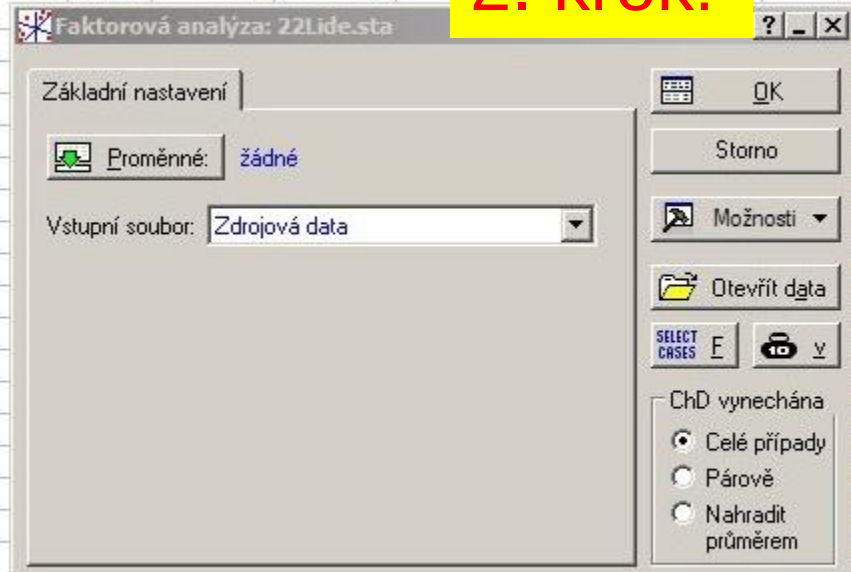
- 1) Graf faktorových vah po rotaci varimax
- 2) Rozptylový diagram faktorového skóre po rotaci varimax



5FA	168	52	1	37
6MA	183	81	-1	42
6FA	157	47	1	36



27	23500	260	86	1
----	-------	-----	----	---



FA

Počet proměnných: 12

Metoda: Hlavní komponenty

$\log(10)$ determinantu korelační matice: -9.3232

Počet extrahovaných faktorů: 3

Vlas. čísla: 6.43010 2.24163 1.61737

Zátěže

Skóre

Popisné statistiky

Shrnutí

Shrnutí

Zákl. výsledky

Výklad rozptylu

Storno



Vlastní čísla

Rotace faktorů: Bez rotace



Shrnutí: Faktorové zátěže



Graf faktorových zátěží, 2D



Možnosti



Anal.Sk.

Zvolte dva faktory pro graf

osa x

osa y

Faktor 1
Faktor 2
Faktor 3

Faktor 1
Faktor 2
Faktor 3

OK

Storno

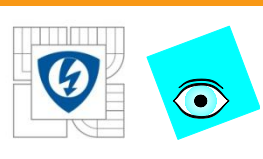
Zvolte dva faktory pro
2D bodový graf

Volba grafu faktorových vah

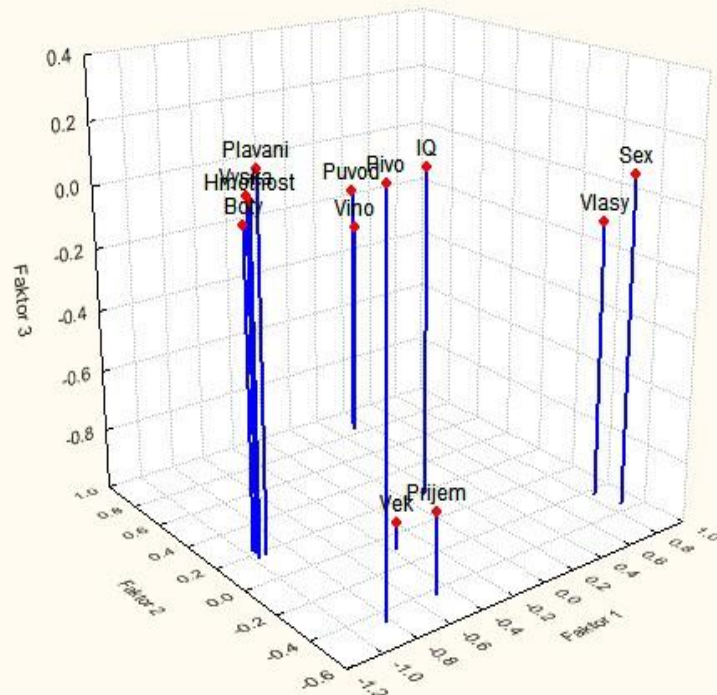
Volba typu rotace

Volba faktorů v grafu

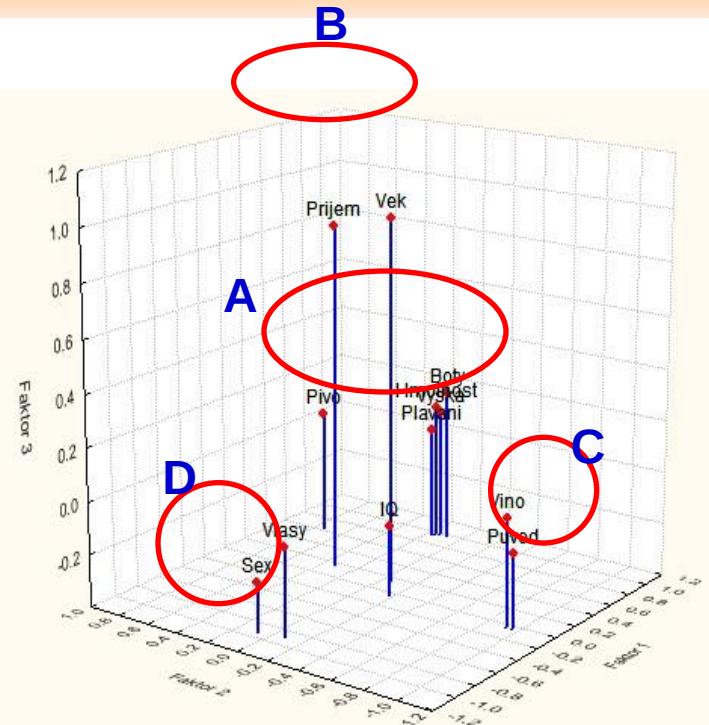
32000	235	92	1	70
34000	255	134	1	76
34000	265	124	1	75
37000	355	82	-1	88
42000	362	90	-1	86
16000	295	180	-1	92
16500	299	178	-1	95
34000	170	162	1	76
14000	150	245	1	75
18000	120	120	1	70
18000	209	160	-1	86
19000	236	175	-1	85



3D graf faktorových vah



Faktorová analýza:
Graf faktorových vah **bez rotace**



Faktorová analýza:
Graf faktorových vah **po rotaci**
Varimaxem normalizovaným

Výsledky faktorové analýzy: 22Lide.sta

Počet proměnných: 12
 Metoda: Hlavní komponenty
 log(10) determinantu korelační matice: -9.3232
 Počet extrahovaných faktorů: 3
 Vlas. čísla: 6.43010 2.24163 1.61737

Zátěže **Skóre** Popisné statistiky Shrnutí
 Zákl. výsledky Výklad rozptylu Sromo

Vlastní čísla

Rotace faktorů: Bortolotti

Shrnutí faktorové zátěže

Graf faktorových zátěží, 2D

32000	235	92	1	70
34000	255	134	1	76
34000	265	124	1	75
37000	355	82	-1	88
42000	362	90	-1	86
16000	295	180	-1	92
16500	299	178	-1	95
34000	170	162	1	76
14000	150	245	1	75
18000	120	120	1	70
18000	209	160	-1	86
19000	236	175	-1	85

Zvolte dva faktory pro graf

osa x osa y

Faktor 1 Faktor 1
 Faktor 2 Faktor 2
 Faktor 3 Faktor 3

OK Sromo

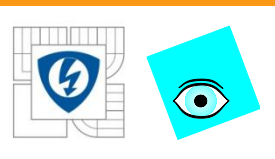
Zvolte dva faktory pro 2D bodový graf

	Faktor. skóre pro 3		
	1 FAKTOR1	2 FAKTOR2	3 FAKTOR3
1MA	-2.10337	-0.44951	-0.83652
2MA	-1.22769	-0.19874	0.52285
3MA	-1.18146	-0.24292	0.16262
1FA	0.42743	-1.23279	0.32023
2FA	0.38680	-0.95855	1.29410

Tabulka faktorového skóre, která se vynese do 2D- nebo 3D-grafu či rozptylového diagramu faktorového skóre

1MB	-0.84052	1.37420	1.18824
2MB	-0.98991	1.68548	1.21436
1FB	0.79134	0.22809	-1.64409
2FB	1.31543	1.30530	0.67030
3FB	1.46358	0.09754	0.15280
3MB	-0.18253	1.32941	0.85759
4MB	-0.43320	1.39869	0.39180
5MB	-0.56724	1.02185	-0.82931
4FB	1.26410	0.42203	1.26431
5FB	1.35025	0.67875	1.16285
6MB	-0.44903	0.83231	-1.72421
7MB	-0.13216	0.69329	-2.30135
6FB	1.03155	0.21238	-0.47405
7FB	1.16669	0.45960	-1.35898
8MB	-0.25720	1.57594	0.10002
8FB	1.22349	0.52505	-1.26331

Volba výstupní tabulky faktorového skóre

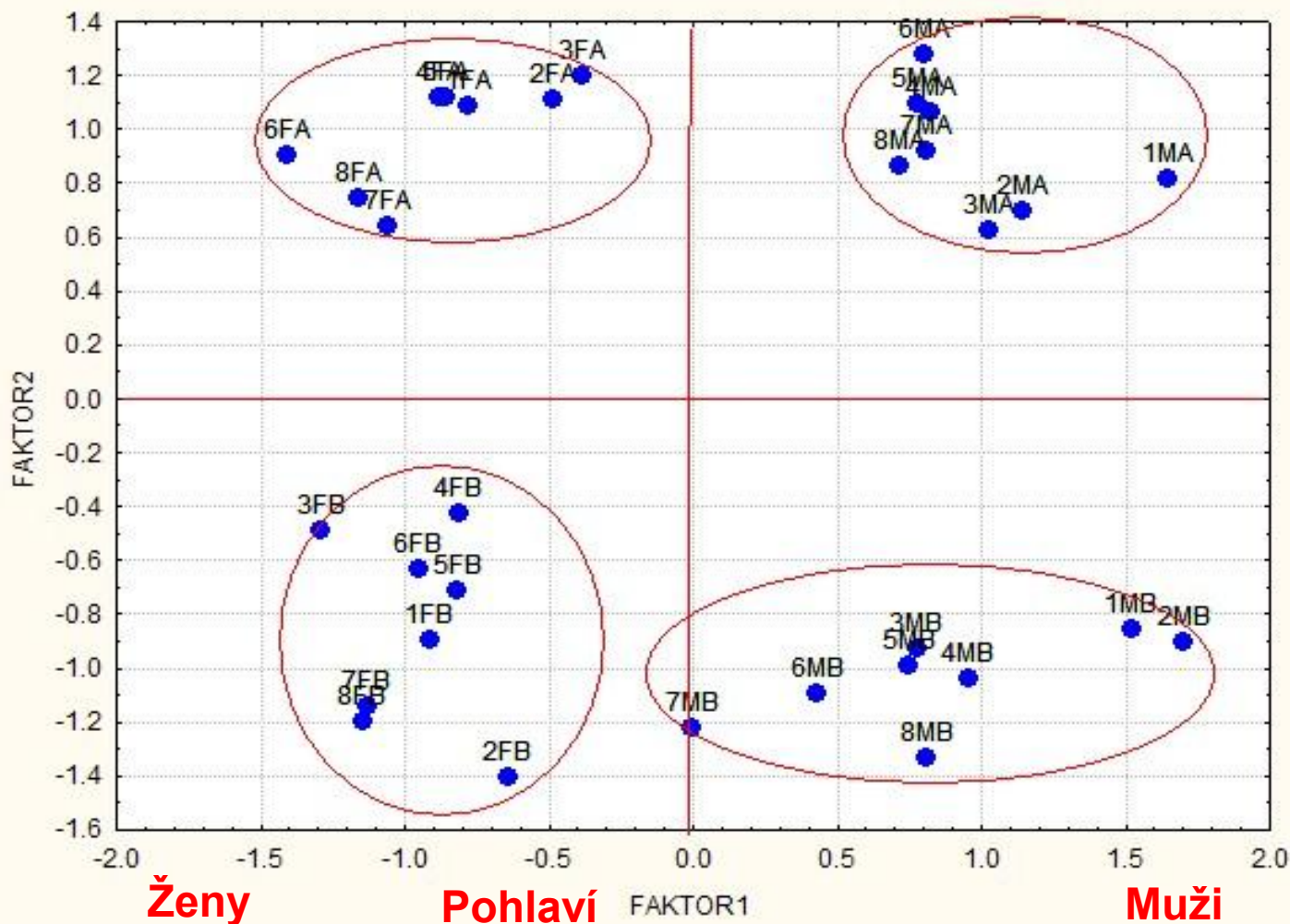


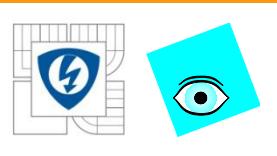
Rozptylový diagram faktorového skóre po rotaci Varimaxem normalizovaným

Sever

Geografie

Jih

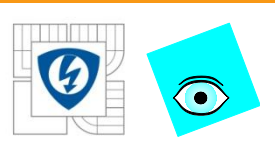




Určení struktury a vazeb

Zdrojová matice má rozměr $n \times m$.

1. **Hledání struktury v *proměnných* v metrické škále:** *faktorová analýza a analýza hlavních komponent.*
2. **Hledání struktury v *objektech* v metrické škále:** *shluková analýza.*
3. **Hledání struktury v *objektech* v metrické i v nemetrické škále:** *vícerozměrné škálování.*
4. **Hledání struktury v *objektech* v nemetrické škále:** *korespondenční analýza.*



5. ***Zpracování lineárních vícerozměrných modelů***, kde závisle proměnné se uvažují jako lineární kombinace nezávisle proměnných resp. vazby mezi proměnnými jsou lineární.
6. ***Exploratorní (průzkumová) analýza dat*** umožňuje
 - a) posoudit *podobnost objektů* pomocí vizuálních rozptylových diagramů a symbolových grafů,
 - b) nalézt *vybočující objekty*, resp. jejich proměnné,
 - c) stanovit, zda lze použít předpoklad *lineárních vazeb*,
 - d) ověřit *předpoklady o datech*: (normalita, nekorelovanost, homogenita).