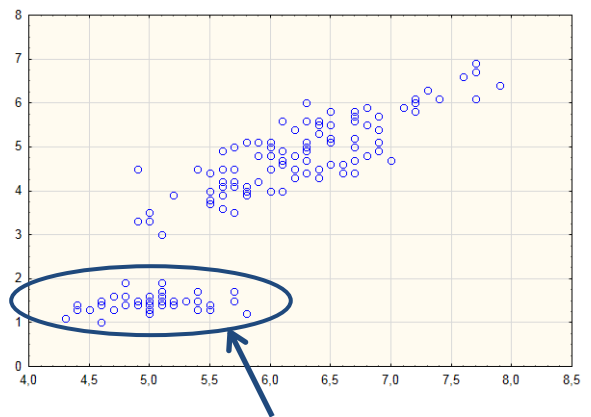


StatSoft Průzkumník grafu

Ukažme si detailně zajímavou a velmi užitečnou funkcionalitu programu *STATISTICA* nazvanou „průzkumník“ grafu.

O co jde

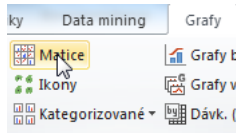
Představte si, že se koukáte na graf (například bodový) a napadne Vás, že by bylo dobré vědět, která pozorování patří do shluku dat, který se Vám v obrázku objevil. Je jisté, že toto je nejpohodlnější zjistit nějak přímo z grafu, který vidíme. Toto byla také základní motivace pro vznik funkcionality průzkumník grafu. Nástroj ale nezůstal pouze u identifikace případů, ale šel mnohem dále, jak se přesvědčíme v následujícím.



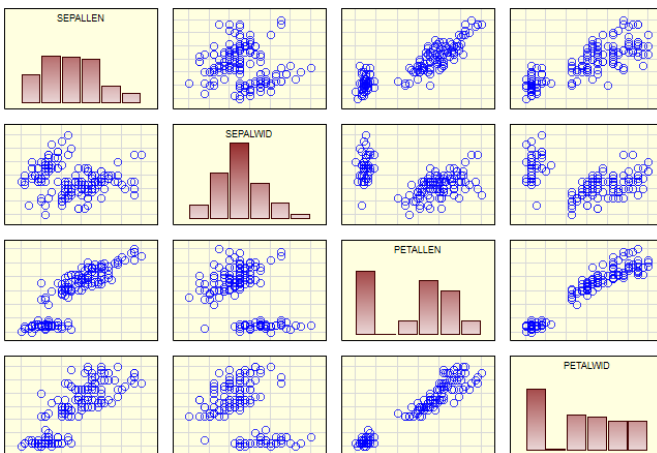
Příklad 1

Pro hraní si s průzkumníkem otevřeme nejdříve soubor [Irisdat.sta](#) (soubor najdete v příkladech souboru *STATISTICA: Soubor-Otevřít příklady-Datasets-Irisdat.sta*). Jedná se o slavný soubor s měřeními okvětních a kališních lístků u tří druhů kosatců.

Vykresleme si nejdříve nějaký graf, kde krásně uvidíme data. Použijeme maticový graf pro všechny naměřené proměnné *Grafy-Matice* (v prostředí nabídek *Pás karet*, jinak se jedná o *Maticové grafy*).



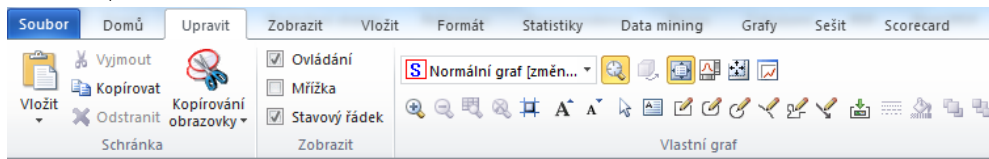
Data: Irisdat* (5s krát 150)					
Fisher (1936) iris data: length & width of sepals and petals.					
	1	2	3	4	5
	SEPALLEN	SEPALWID	PETALLEN	PETALWID	IRISTYPE
1	5,0	3,3	1,4	0,2	SETOSA
2	6,4	2,8	5,6	2,2	VIRGINIC
3	6,5	2,8	4,6	1,5	VERSICO
4	6,7	3,1	5,6	2,4	VIRGINIC
5	6,3	2,8	5,1	1,5	VIRGINIC
6	4,6	3,4	1,4	0,3	SETOSA



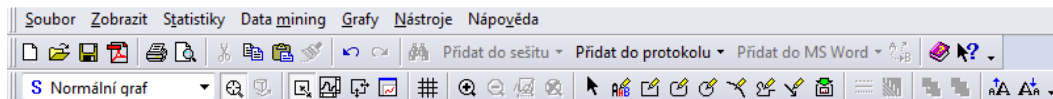
Vidíme přehledně data pro všechny proměnné, jak spolu souvisí, včetně rozdělení jednotlivých proměnných. Můžeme si všimnout, že se v obrázku vyskytují jakési seskupení bodů, které bychom rádi blíže pochopili. Zapněme tedy průzkumníka grafu.

Zapnutí průzkumníka:

1. Při otevřeném grafu v liště **Upravit** (pokud máme aktivní zobrazení nabídek ve formě *Pásu karet*), jedná se o ikonu s lupou se „zaměřovačem“



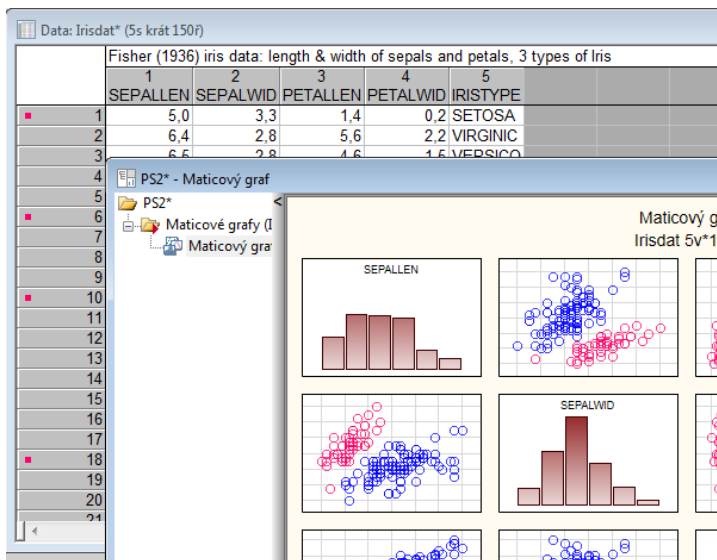
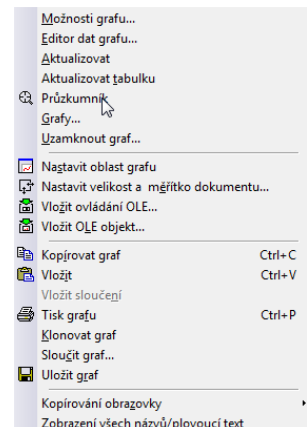
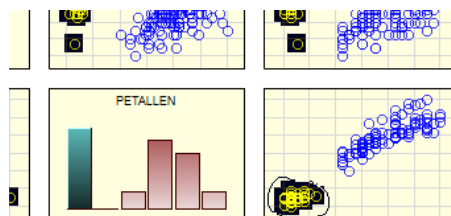
2. Při otevřeném grafu ikona s lupou se „zaměřovačem“ (pokud máme aktivní klasické nabídky)



3. V grafu klikneme pravým tlačítkem myši a vybereme možnost **Průzkumník**

Jakmile je průzkumník zapnut, vyberte možnost *laso* a označte jím v obrázku (držením levého tlačítka myši) body jednoho ze shluků:

Vidíme, že se data označila a že se označila na všech obrázcích zároveň. To ale není vše, pokud se překlikneme na datový soubor, vidíme, že jsou označeny i body v souboru. Už jen toto je velmi lákavá funkcionality, protože ihned vidíme, která data jsme vybrali. Tuto funkcionality jsme použili např. i ve článku [Zpracování chybějících dat a dat mimo rozsah](#). S vybranými body však jde dělat další věci. Vraťme se na graf, zatrhneme *obarvit*, vybereme barvičku a klikneme na *Použít*.

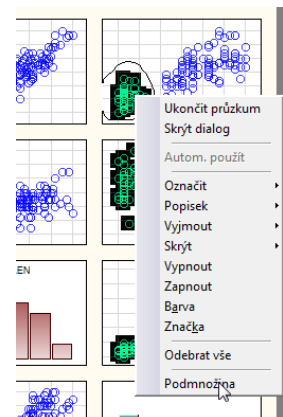


Nyní jsou obarveny body, které jsme si před tím označili. Také se barevná značka „otiskla“ do datového souboru. Pokaždé, kdy budeme vykreslovat body z tohoto datového souboru, budou tyto body v grafu obarveny touto barvou.

Je potřeba poznamenat, že případy datového souboru mohou mít několik vlastností, které jsou právě vidět u čísla či názvu případu, bude o nich řeč později, nyní jsme viděli alespoň značku pro obarvení.

Nyní bychom chtěli nějak shrnout, jaká data jsme vlastně obarvili, vypadá to na jasně oddělenou skupinu. Můžeme se podívat do souboru a zkontrolovat jaký typ kosatce je obarven. Jednoduše však můžeme použít

zajímavou funkcionality průzkumníka a vytvořit přímo podmnožinu označených dat: Pokud máme průzkumníkem označena data, klikneme pravým tlačítkem a vybereme možnost **Podmnožina**, vytvoří se nový datový soubor, ve které jsou pouze vybrané body. V našem případě jsou ve výsledném souboru pouze kosatce typu SETOSA, a to navíc všechny z původního souboru. Dokázali jsme tedy „pouhým okem“ určit skupinu SETOSA, která se tedy od ostatních dvou druhů kosatce liší velikostí okvětních a kališních lístků. Není se tedy



čemu divit, že se tato data používají jako neklasičtější případ pro ukázkou a příklad funkčnosti diskriminačních metod.

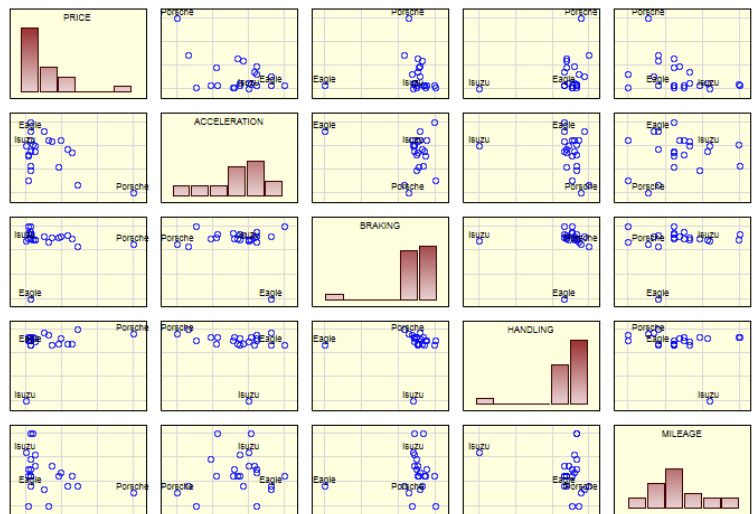
Příklad 2

Další funkcionality vyzkoušíme na dalších datech a to na datech [Cars.sta](#) (soubor najdete opět v příkladech programu *STATISTICA*). Jedná se o data shrnující důležité parametry automobilů. Sloupce jsou standardizovány, takže konkrétní hodnoty nevíme, nicméně rozdíly mezi auty budou samozřejmě stále vidět.

Data: Cars (5s krát 22f)					
	Performance, fuel economy, and approximate price for various				
	1	2	3	4	5
	PRICE	ACCELERATION	BRAKING	HANDLING	MILEAGE
Acura	-0.521	0.477	-0.007	0.382	2.079
Audi	0.866	0.208	0.319	-0.091	-0.677
BMW	0.496	-0.802	0.192	-0.091	-0.154
Buick	-0.614	1.689	0.933	-0.210	-0.154
Corvette	1.235	-1.811	-0.494	0.973	-0.677
Chrysler	-0.614	0.073	0.427	-0.210	-0.154

Opět vykresleme maticový graf pro všechny proměnné.

Vidíme, že některé body jsou výrazné odlehle – jsou jinde než ostatní většina dat. Pokud nás zajímá, která auta představují tyto body, můžeme si navíc data popsat. Otevřme průzkumníka, označme bod mimo ostatní data, zaškrtneme možnost *Popsat* a dáme *Použít*. V obrázku se objevily názvy případů (řádků) – kdybychom neměli u případů názvy, napsalo by se číslo případu. Popsané body vidíte na obrázku vpravo, například je vidět, že Porsche je mimo ostatní data v ceně a zrychlení.



Další zajímavou možností, co s body dělat, je nastavit je na *Vypnuto* (jedna z dalších možností průzkumníka). Vypnutý bod není zobrazen a měřítko os se překreslí a nastaví podle ostatních bodů. Tuto funkcionalitu jistě využijete, pokud odhalíte během vizuální kontroly data, která byla například naměřena chybně nebo neodpovídají účelu, pro který jste graf vykreslovali.

Tuto funkcionalitu jistě využijete, pokud odhalíte během vizuální kontroly data, která byla například naměřena chybně nebo neodpovídají účelu, pro který jste graf vykreslovali.

Funkcionalita průzkumník

Podrobný popis funkcionalit průzkumníka:

Tato oblast určuje, jaký je typ označování

Bod:

Okno:

Laso:

Řez X:

Řez Y:

Hranol (u 3D obrázků):

Řez Z (u 3D):

Průzkumník... (Screenshot of the main dialog box)

Interaktivní | **Rozšířené**

Použít (highlighted)

Skrýt | Konec

Graf 1

Obnovit vše

Automaticky použít

Akce

Normál. | Opak | Přep.

Označit

Popsat

Vypnuto

Obarvit

Značka

Vyber průzkumníka

Bod

Řez X

Okno

Řez Y

Laso

Řez Z

Hranol

Posuvatelny výběr

Uchovat výběr

Automaticky animovat

Animovat

Uložit nastavení

Legenda:

!	Označit	Jsou označeny v grafech jinou značkou, v souboru výkřičníkem
◀	Popsat	Do grafu vypíše jméno případu, pokud není, tak číslo
⊗	Vypnuto	Tento bod nebude vidět v grafu ani nebude použit v analýzách
■	Obarvit	Tento bod bude mít v grafech vždy tuto barvu
▲	Značka	Tento bod bude mít v grafech tuto značku

Annotation 1: Výběr toho, co chceme s označenými body udělat. Poté potřeba aplikovat tlačítkem **Použít**

Annotation 2: Tlačítko aplikuje na označené body to, co je vybráno v oblasti **Akce** a přidá či odebere příznak (označen značkou) před název případu

Annotation 3: Vymaže všechny příznaky u případů, tedy smaže i všechnu práci provedenou průzkumníkem

Annotation 4: Při zaškrtnutí je možné výběr posouvat – aktivní při volbě Okno, Laso a Hranol

Annotation 5: Označené body zůstávají označeny, i když měníme výběr

Annotation 6: Spustí animaci, kdy vybrané okno/laso/řez cestuje v grafu a zvýrazňuje body. Rychlost a krok lze měnit v nově se objevivším okně

Ani... (Screenshot of animation dialog box)

OK

Storno

Přerušit

Původní

X krok:

Y krok:

Čas prodlouhy (s):

0 1 2

Průzkumník... (Screenshot of the extended dialog box)

Interaktivní | **Rozšířené**

Přepínač

Zvýraznit | Odvýraznit

Všechny grafy

Výběr dle skupin:

Označené body

Popsané body

Výjmuté body

Skryté body

Ostatní body

Volba podle rozsahu:

X prom: PRICE

Y prom: ACCELERATION

Min: 0 | Max: 1

X: 0 | 1

Y: 0 | 2

Akce

Normál. | Opak | Přep.

Odznačit

Odstranit

Zapn.

Akce

Normál. | Opak | Přep.

Přepínač - značka

Přepínač

Přepínač

Sekce **Akce** má ještě další 2 varianty, varianta **Opak** značí opačné operace k **Normál.**, tedy postupně značí: odznačit, zrušit popis a zapnout případ (tedy pokud je označen případ s příznakem Vypnuto, pak ho tlačítko **Použít** zruší a případ bude opět v obrázku vidět). Varianta **Přep.** přepíná mezi příznaky, první zatrhávátko po stisknutí **Použít** označí neoznačené a zároveň odznačí označené, druhá volba podobně přepne popis u označených bodů, třetí přepínač vypne či zapne viditelnost bodů.

Průzkumník má ještě rozšířené možnosti (zobrazeny vlevo). Zde můžete zvýraznit, odvýraznit či přepínat skupinu bodů, které splňují podmínky v oblasti **Výběr podle skupin** (výběr se řídí podle příznaků u případů) a **Volba podle rozsahu** (volba podle rozsahu dvou proměnných).

Na obrázku vlevo by tedy byly po stisknutí **Zvýraznit** označeny popsané body, které mají hodnoty v rozmezí 0-1 v proměnné PRICE a 0-2 v proměnné ACCELERATION.

