

5.6 Grafy ve statistické analýze dat

Obsah:

5.6.1 Krabicový graf (Box Plot)

5.6.2 Diagram rozptýlení (Simple Dot Chart)

5.6.3 Vícerozměrný diagram rozptýlení (Multi-Data Dot Chart)

5.6.4 Vícerozměrné proložení histogramu křivkami (Add multiple fitted curves in a Histogram)

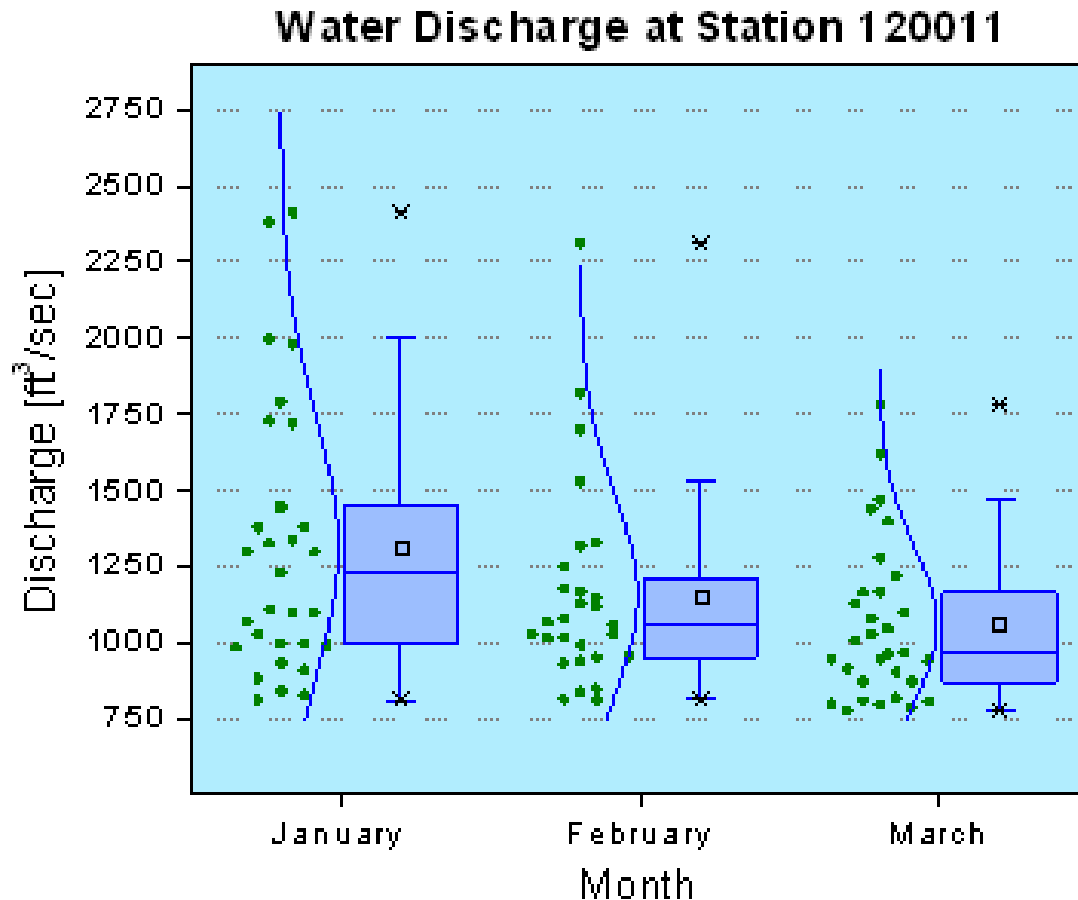
5.6.5 Weibullův pravděpodobnostní graf (Weibull Probability Plot)

5.6.6 Kvantil-quantilový graf (Q-Q Plot)

5.6.7 Matice rozptylového grafu (Scatter Matrix)

5.6.1 Krabicový graf (Box Plot)

Vytvořte diagnostický graf o jednorozměrných datech v projektu **File, Open, Samples, Statistical and Specialized Graphs.opj, Open.**



Kroky:

1. V **Project Explorer** otevřete projekt **Statistical and Specialized Graphs, Statistical, Box Chart**.
2. Aktivujte list **Box Data** a zvolte sloupce **January(Y), February(Y), and March(Y)** a nakreslete je **Plot, Statistics, Box Chart**. Dvoj-klik na jednu krabici a otevře se **Plot Deatils**. V **Border Color** vyberte **None** a klikněte na **Apply**.

3. Poté aktualizujte vlastnosti skupiny dle těchto kroků:

Box záložka:

Type: Box[Right]+Data[Left]

4. **Data záložka:**

Distribution Curve Type: Normal
Automatic Binning: Disabled/Unchecked
Bin Size: 100
Begin: 750
End: 3600
Bin Height: 100

5. **Line záložka:**

Style: Solid
Width: 2
Color: Blue

6. Symbol záložka:

Size:	5
Symbol Color:	Olive
Shape:	Circle
Interior:	Solid

7. Pattern záložka:

Border Color:	Blue
Border Style:	Solid
Border Width:	2
Fill Color:	Custom (R:156, G:190, B:254)
Pattern:	None

8. Percentile záložka:

Size:	5
Edge Color:	Black
Fill Color:	Black

9. V **Plot Details** vyberte vrstvu v levém panelu pro vrstvy **Layer1**. Přejděte na záložku **Background** a nastavte barvu souřadnicemi (**R: 177, G: 237, B: 254**). Kliknutím na **OK** potvrďte nastavení a zavřete dialogové okno.

10. Dvoj-klikem na osu **Y** a vyberte záložku **Scale** a nastavte osy následovně:

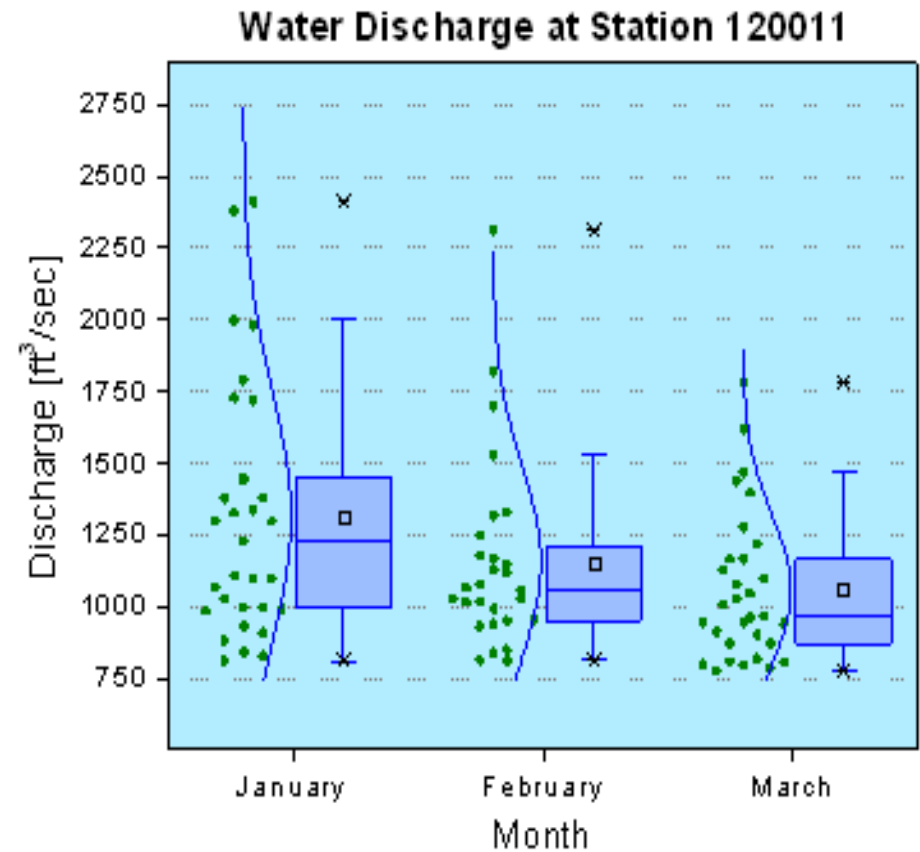
From:	510
To:	2900
Increment:	250
#Minor Ticks:	0

11. Aktivujte **Grid Lines**. Vyberte ikonu **Horizontal** v **Selection** a upravte mřížky pro osu **Y**. Zaškrtněte políčko **Major Grids** a nastavte **Line Color** = **LT Gray** a **Line Type** = **Dot**.

12. Zaškrtněte políčko **Opposite Line** v záložce **Grid Lines** pro obě **Vertical** a **Horizontal** v seznamu **Selection**.

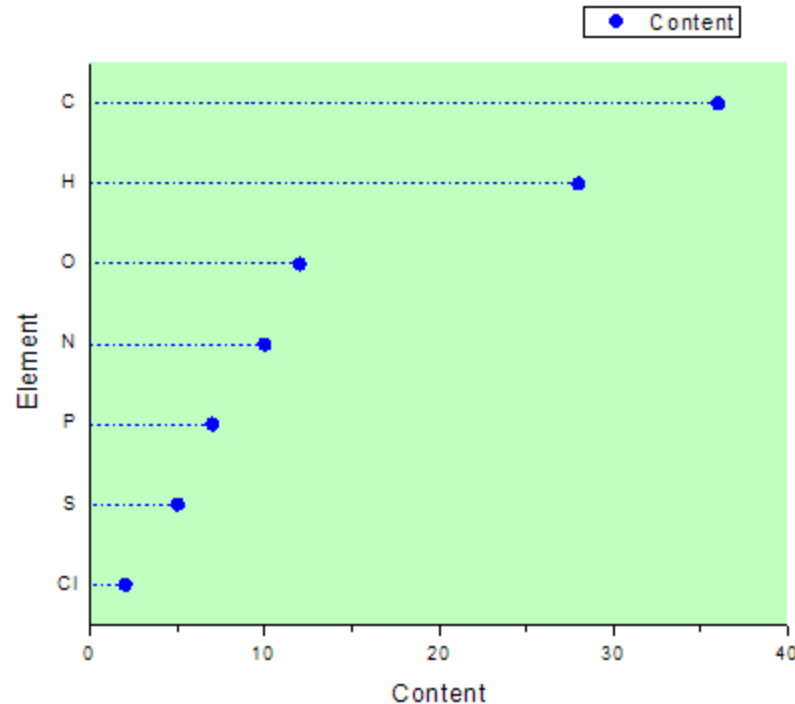
13. Odstraňte legendu a aktualizujte názvy os **X** a **Y** v případě potřeby.

14. Klikněte pravou myší na vrstvu a vyberte **Add/Modify Layer Title** a přidejte název dle vlastní volby.



5.6.2 Diagram rozptýlení (Simple Dot Chart)

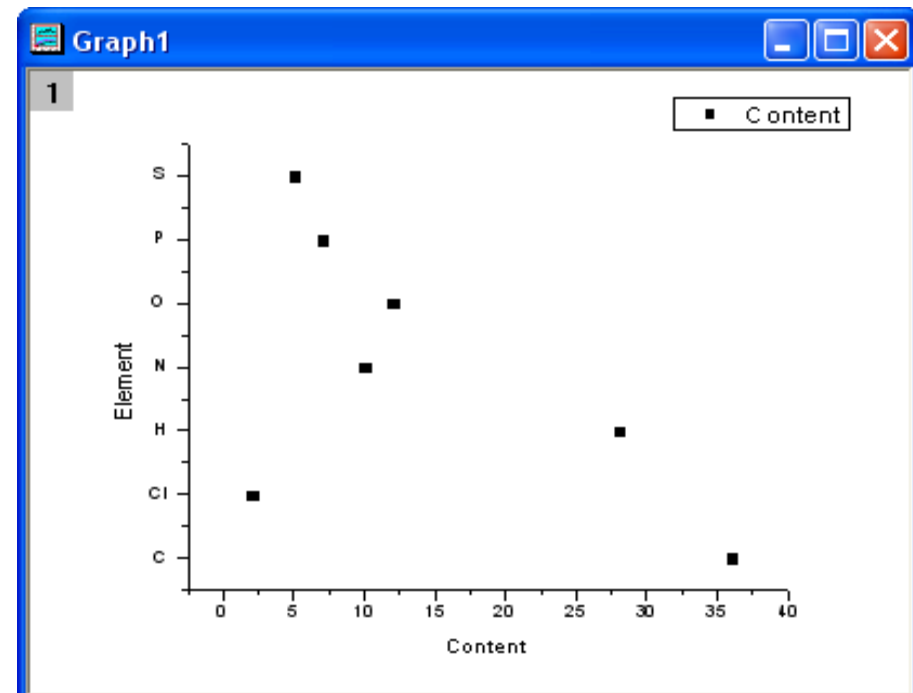
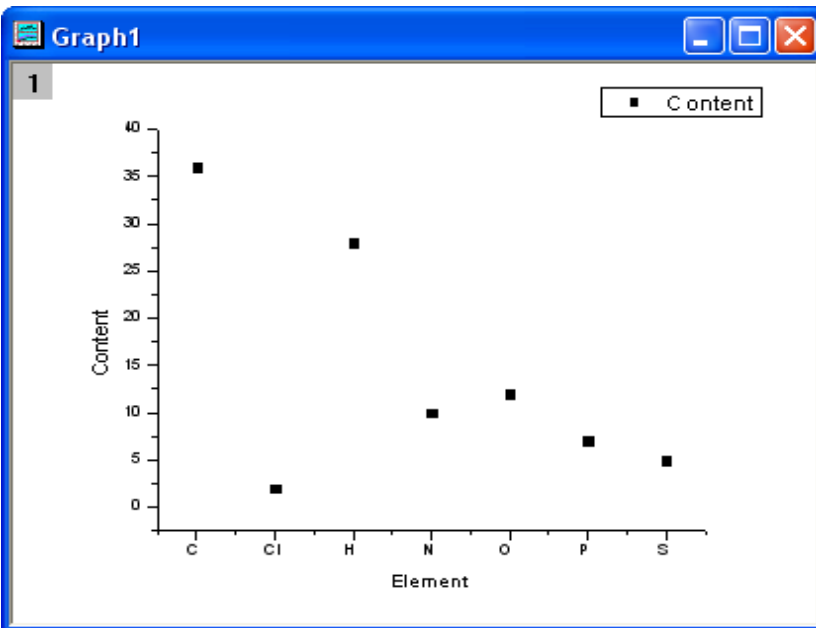
Diagram rozptýlení (**Dot graph**) je statistický diagnostický graf, který se skládá z datových bodů zakreslených v jednoduchém měřítku. Často je využíván jako náhrada za výsečový graf, neboť umožňuje posoudit a porovnat i větší množství bodů.



Kroky:

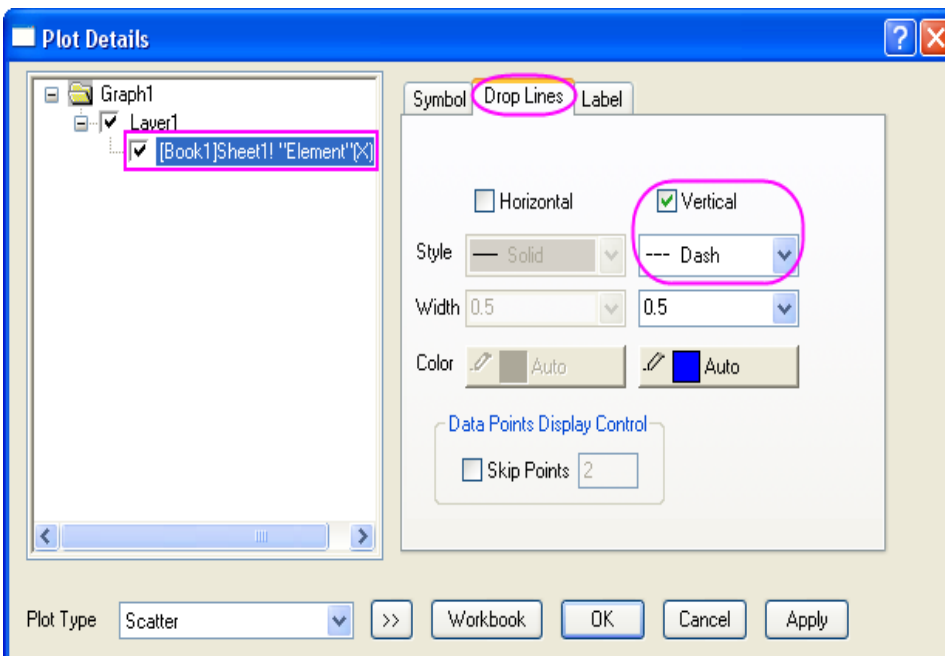
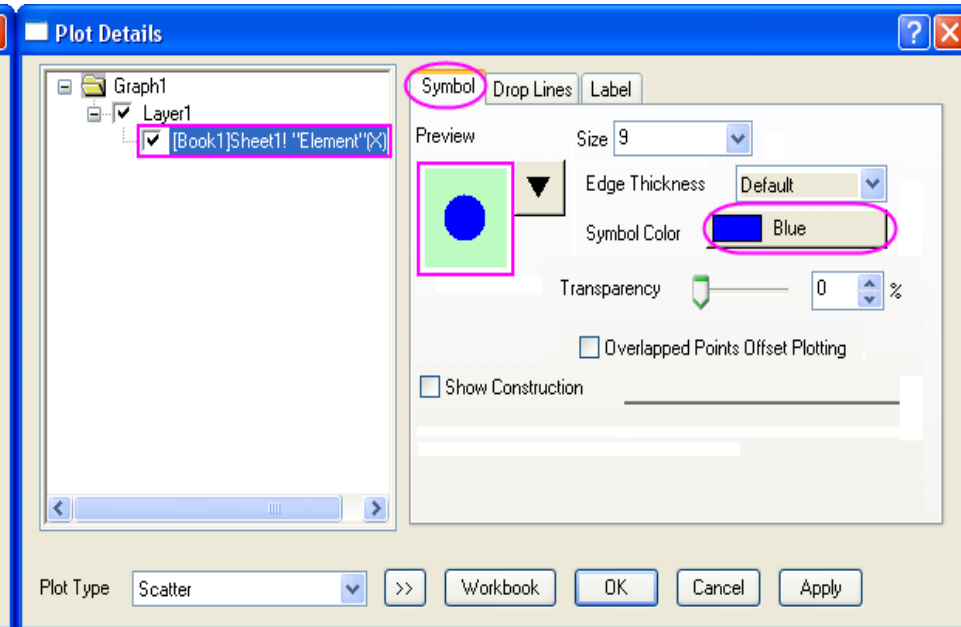
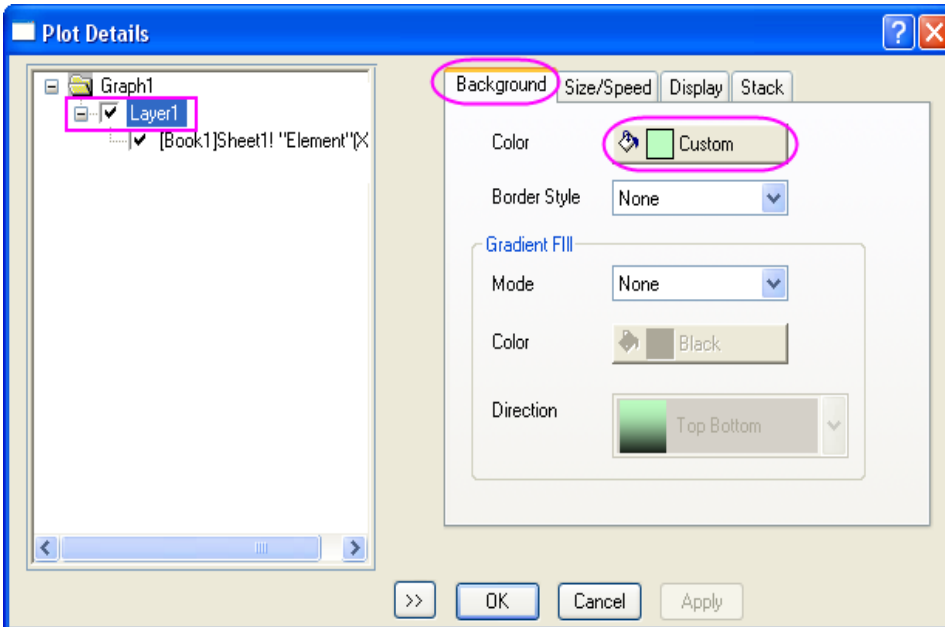
1. Začneme s daty **File, Open, 5.6.2 Simple-Dot-Chart.txt, Open**, která představují obsah různých prvků ve sloučenině:
2. Zvýrazněte **col(A)** a **col(B)**, a vyberte **Plot, Symbol, Scatter** k vytvoření bodového grafu.

	A()	B()
Long Name	Element	Content
1	C	36
2	Cl	2
3	H	28
4	N	10
5	O	12
6	P	7
7	S	5
8		
9		
10		
11		

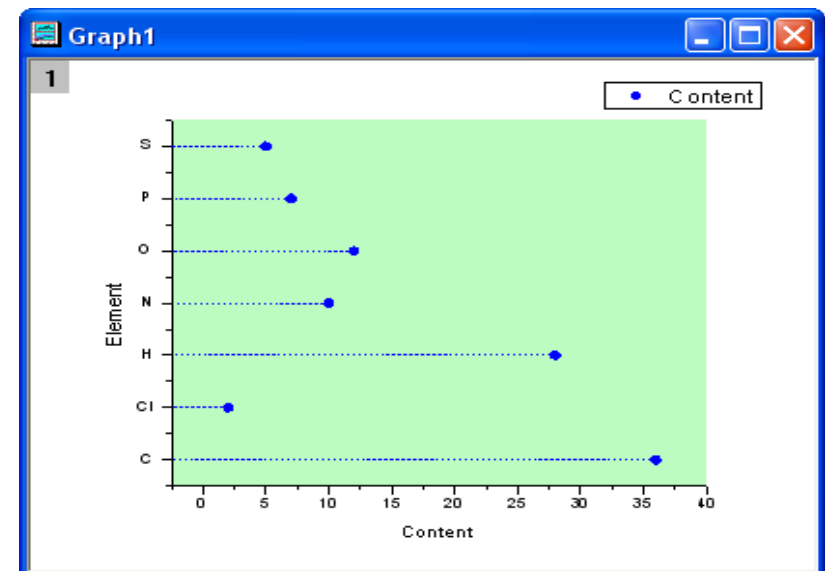


3. Vyberte **Graph, Exchange X-Y Axis** z nabídky.

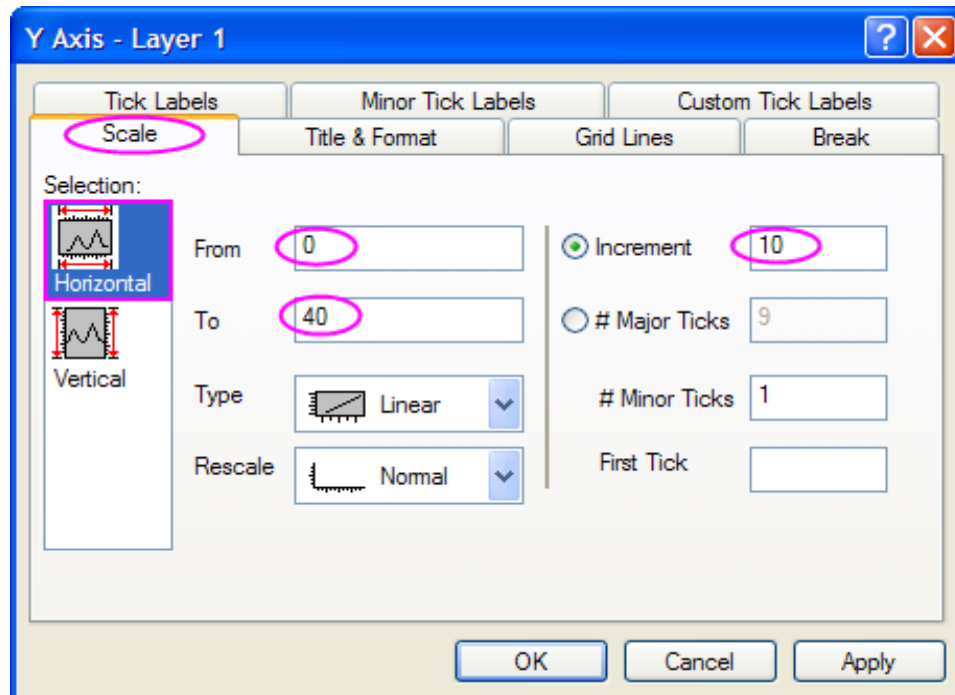
4. Dvoj-klikem levou myší na graf vyvolejte **Plot Details** a změňte symboly a barvu dle obrázku:



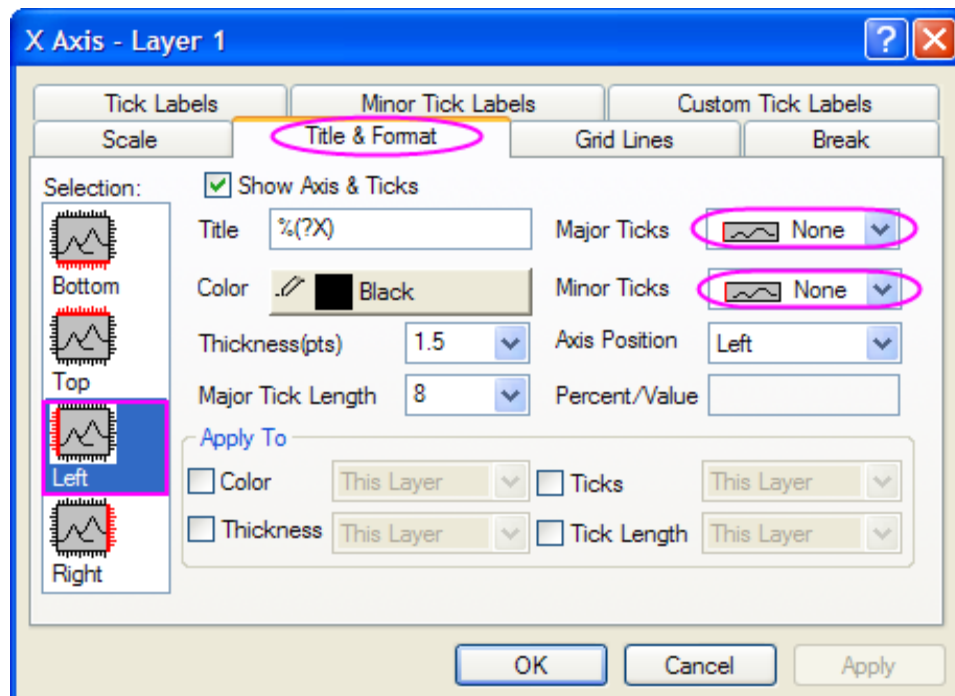
5. Kliknutím na OK zavřete dialog.



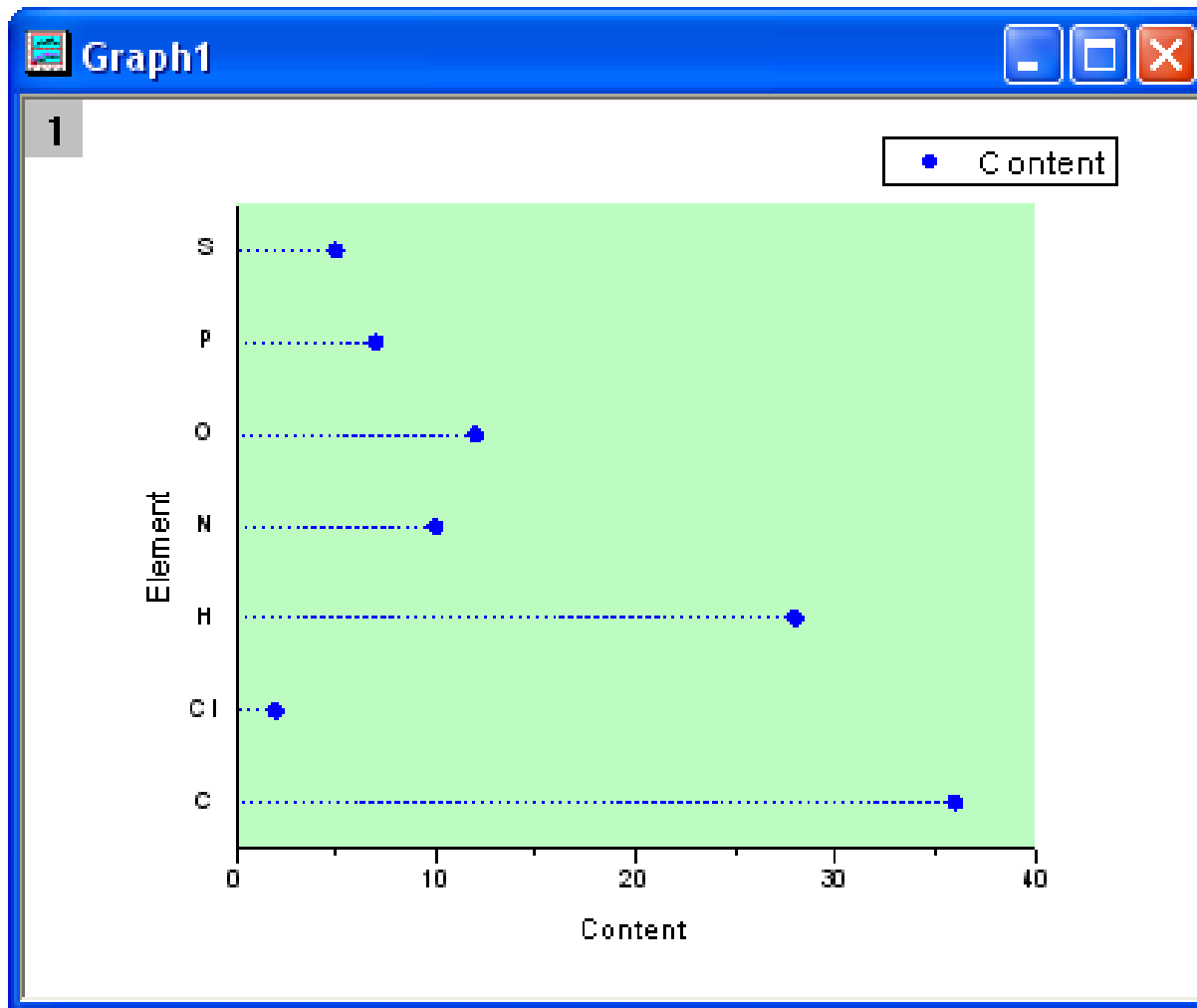
6. Obnovte osy **X** a **Y**. Poklepejte na osu **X** a otevře se dialog **Axis Properties**. Na záložce **Scale** nastavte **From** na **0** a **To** na **40**. Nastavte **Increment** na **10**.



7. V záložce **Title and Format** se ujistěte, že **Left** je vybráno v **Selection** a pak nastavte **Major Ticks** a **Minor Ticks** na **None**. Nakonec klik na **OK**.



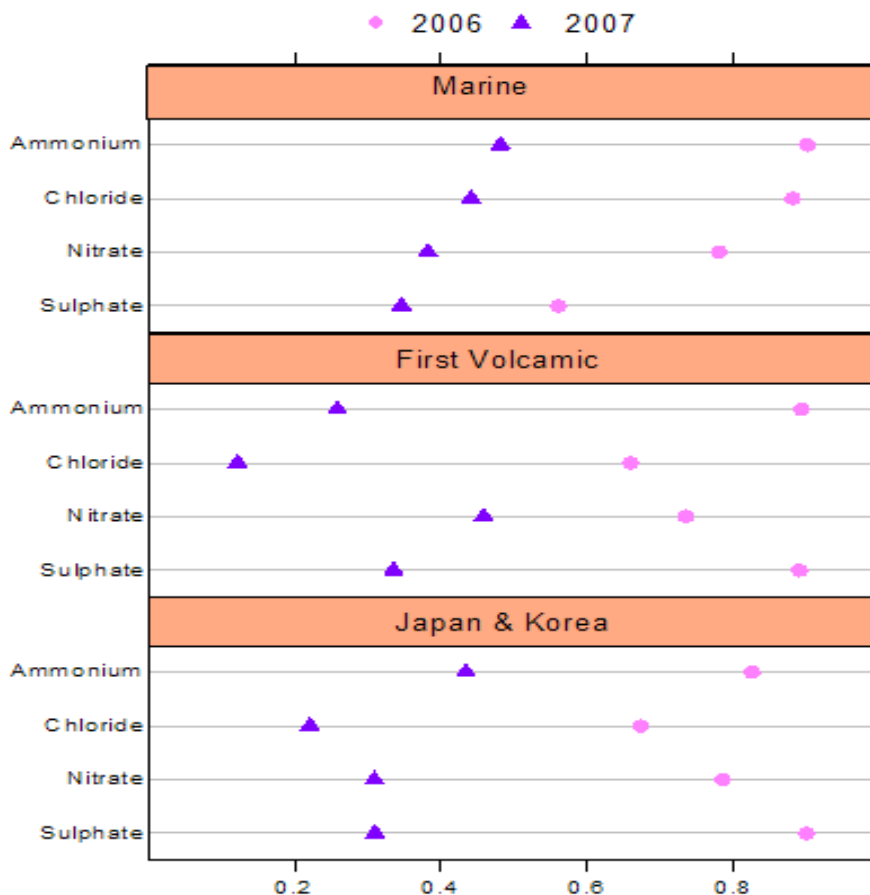
Nyní je rozptylový diagram úspěšně vyneseno a měl by vypadat dle obrázku níže:



5.6.3 Vícerozměrný diagram rozptýlení (Multi-Data Dot Chart)

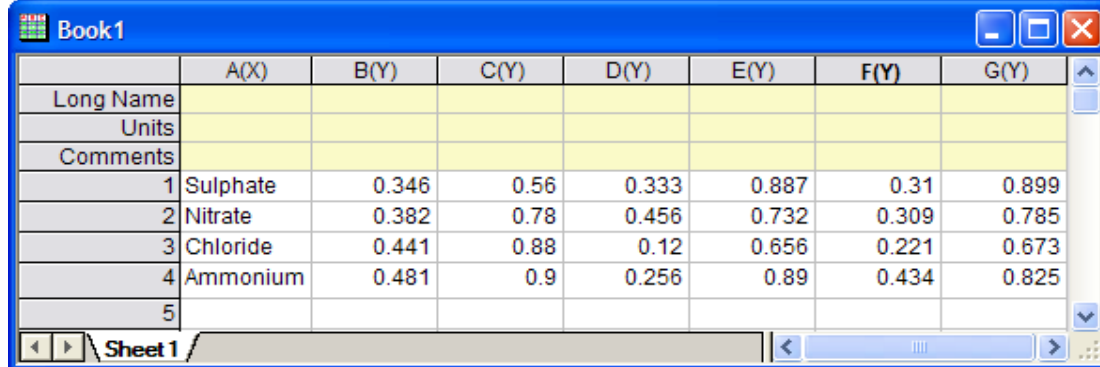
Sulphate	0.346	0.560	0.333	0.887	0.310	0.899
Nitrate	0.382	0.780	0.456	0.732	0.456	0.732
Chloride	0.441	0.880	0.120	0.656	0.221	0.673
Ammonium	0.481	0.900	0.256	0.890	0.434	0.825

Je obdobný jako předešlý diagram rozptýlení.
Několik dat o obsahu prvku v rozličných oblastech v různém čase použijete zde k vytvoření vícerozměrného diagramu rozptýlení



Kroky:

1. Vytvořte nový sešit a zadejte data z obrázku vpravo.

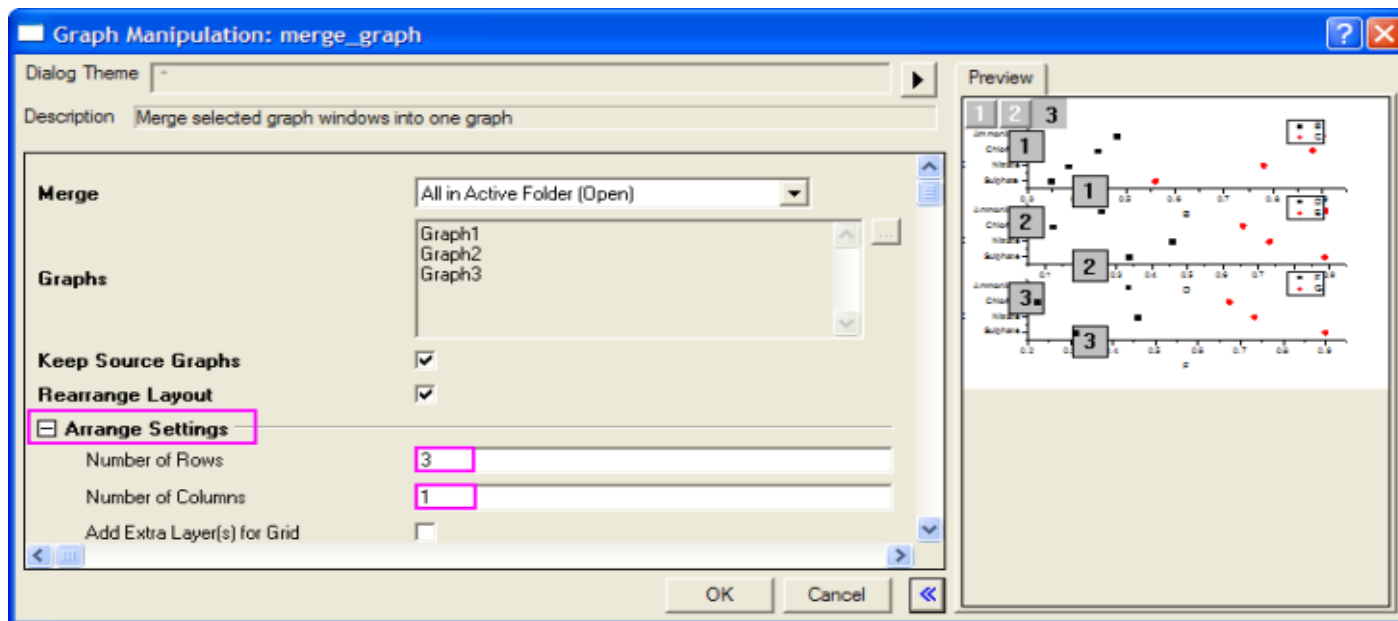


	A(X)	B(Y)	C(Y)	D(Y)	E(Y)	F(Y)	G(Y)
Long Name							
Units							
Comments							
1	Sulphate	0.346	0.56	0.333	0.887	0.31	0.899
2	Nitrate	0.382	0.78	0.456	0.732	0.309	0.785
3	Chloride	0.441	0.88	0.12	0.656	0.221	0.673
4	Ammonium	0.481	0.9	0.256	0.89	0.434	0.825
5							

2. Zvýrazněte **col(B)** a **col(C)** a zvolte **Plot, Symbol, Scatter** a pak **Graph, Exchange XY Axis**.

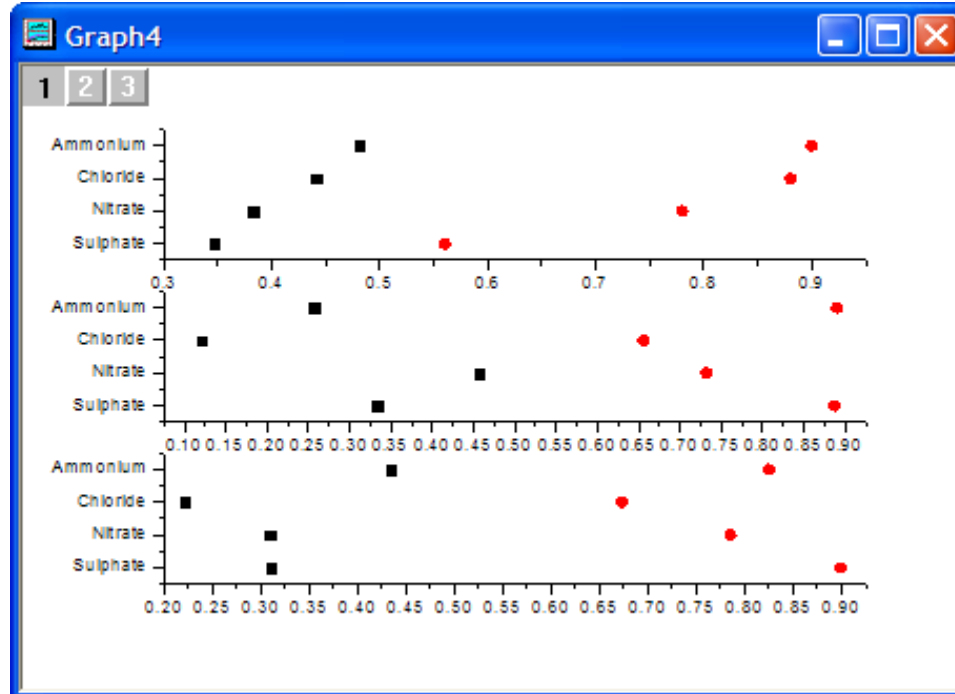
3. Opakujte krok 2 ke tvorbě grafu s **col(D)** a **col(E)** a pak také dalšího grafu s **col(F)** a **col(G)**.

4. Sloučení těchto tří právě vytvořených grafů: **Graph, Merge Graph Windows, Open Dialog**. Rozbalte **Arrange Settings** a položte **Number of Rows** rovné **3** a **Number of Columns** rovné **1**. Klik na **OK**. Nyní máte nový graf, který obsahuje tři vrstvy.



5. Můžete odstranit legendu a popisky os **X, Y** v grafech, a tím vyčistíte graf.

6. Zvolte **Graph, Layer Management**. Vyberte **Layer2** na levé straně. Pak na pravém panelu aktivujte záložku **Link**. V této záložce nastavte **Link** na **1** a osa **Y Axis** na **Straight(1 to 1)**. Klepněte na **Apply**.



Name	Link to
1 Layer1	
2 Layer2	1
3 Layer3	1

Link

Link To: 1

X Axis: None

Y Axis: Straight(1 to 1)

Units: % of Linked Layer

Apply

Use the Apply buttons on each tab to update your preview graph. Use the OK button to apply the changes in preview graph to source graph.

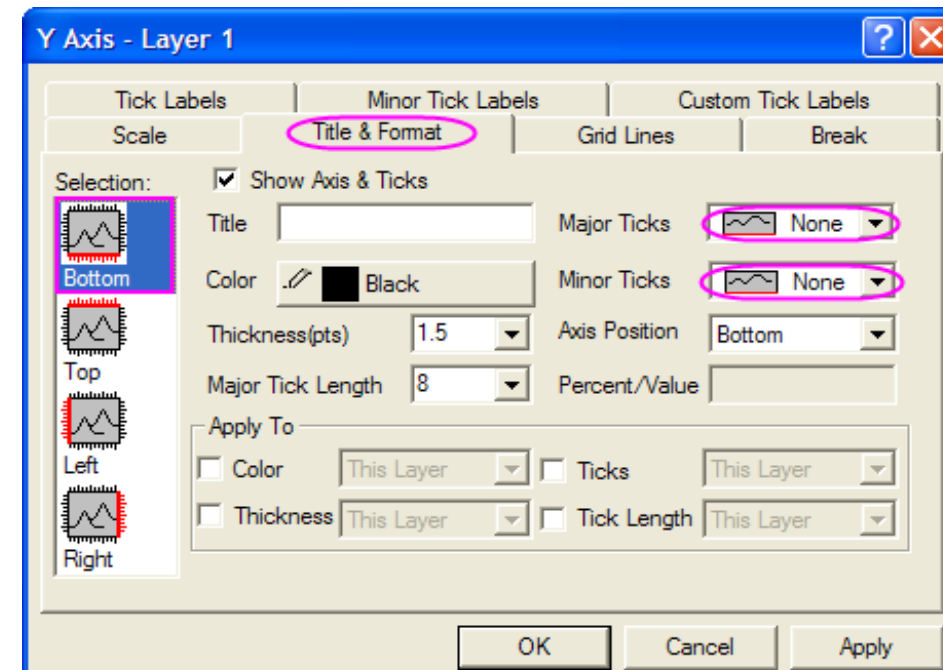
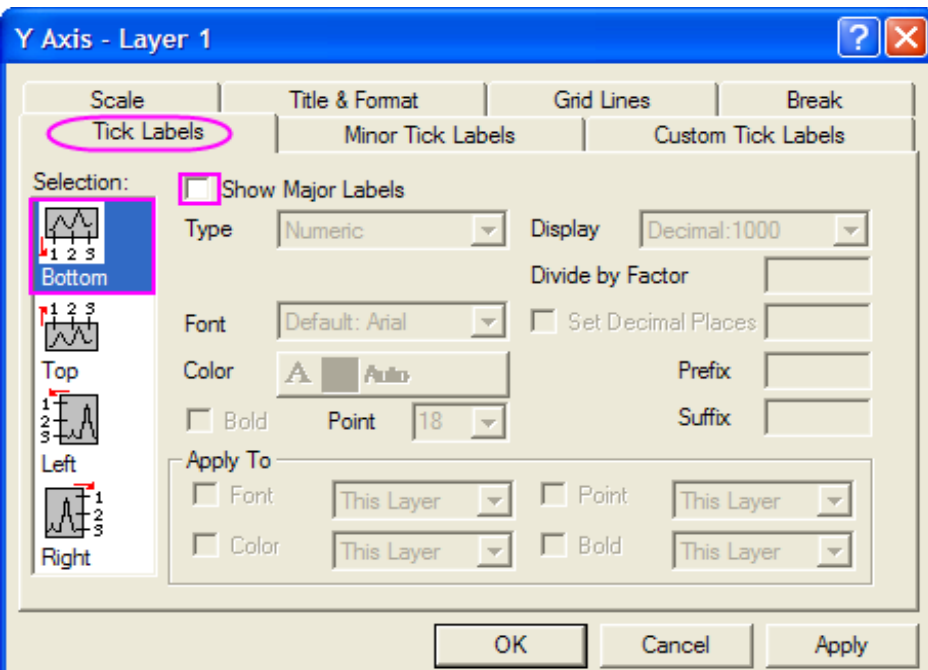
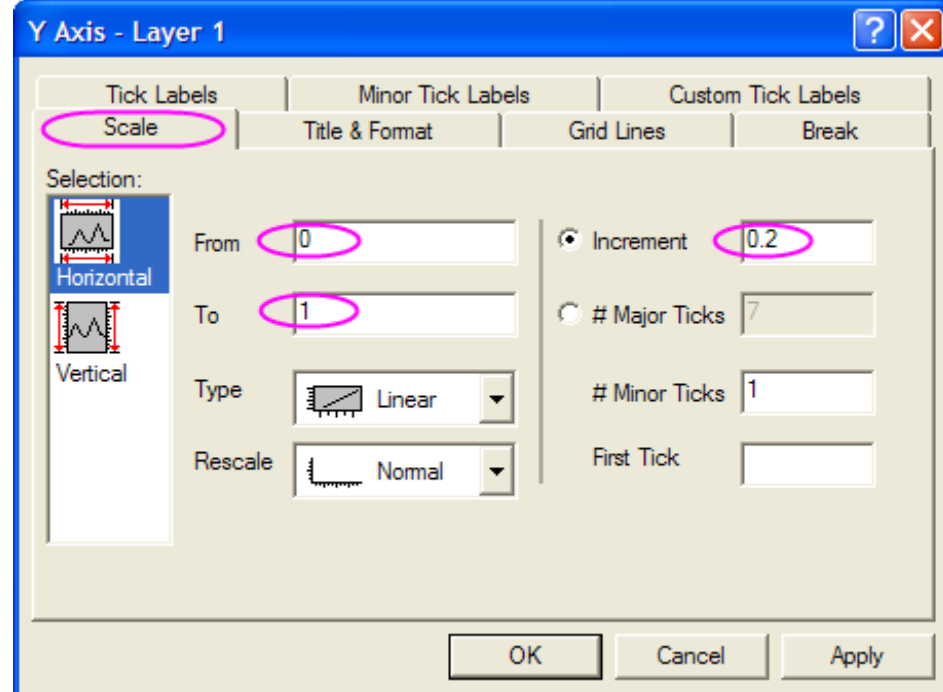
Undo OK Cancel

7. Vyberte **Layer 3** na levé straně. Také v záložce **Link** nastavte na **1** a **Y Axis** na **Straight(1 to 1)**. Poté klik na **OK**.

8. Obnovte osy **X** a **Y**. Aktivujte **Layer 1**, dvojklik na na osu **X** od **Layer 1**. Nastavte **Scale** a **X From** na **0** **To** na **1** a **Increment** na **0,2**.

9. V záložce **Tick Labels** zrušte zaškrtnutí políčka **Show Major Label**

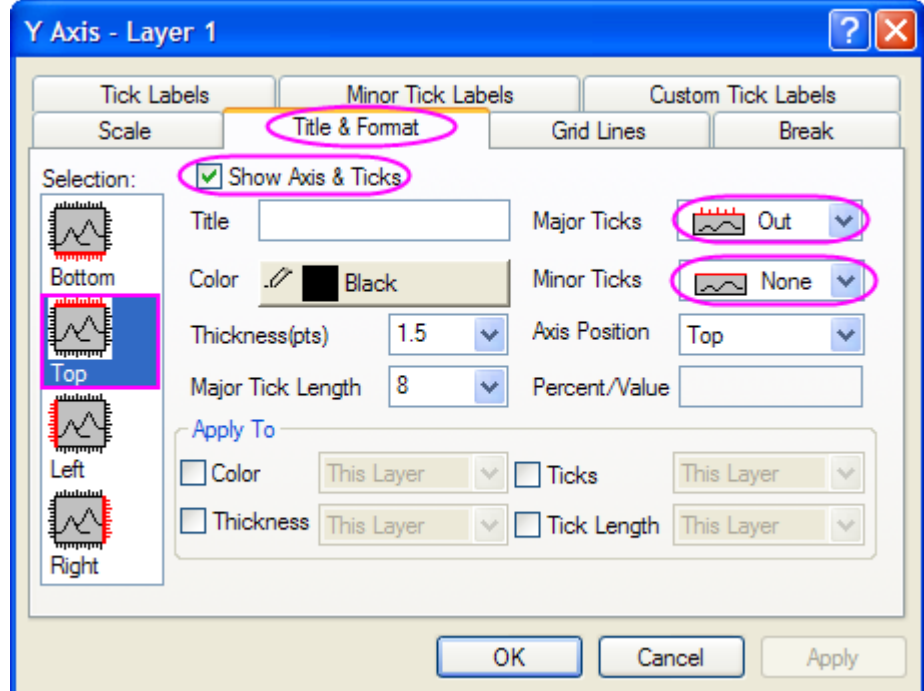
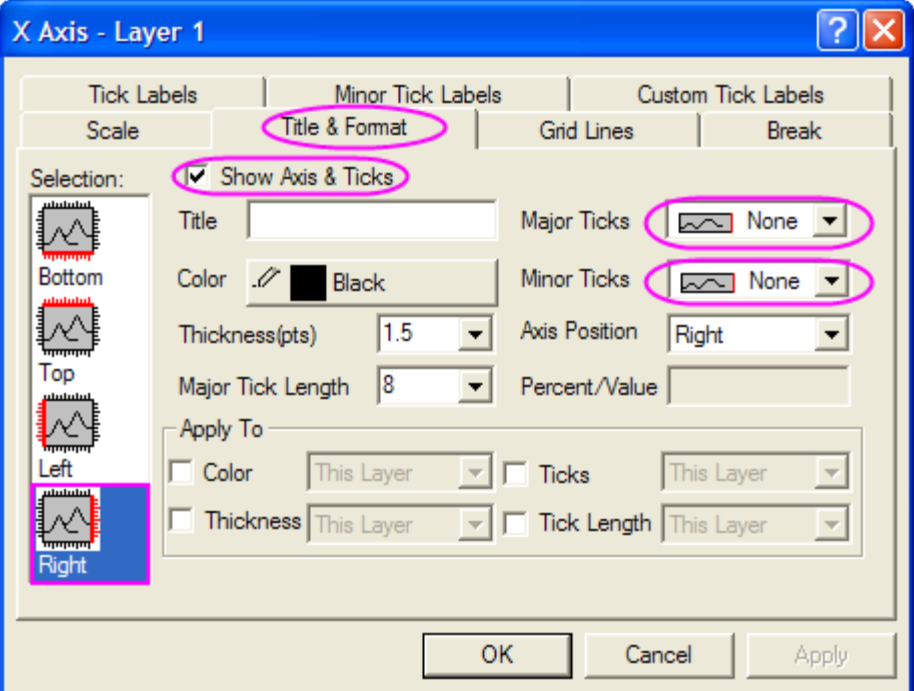
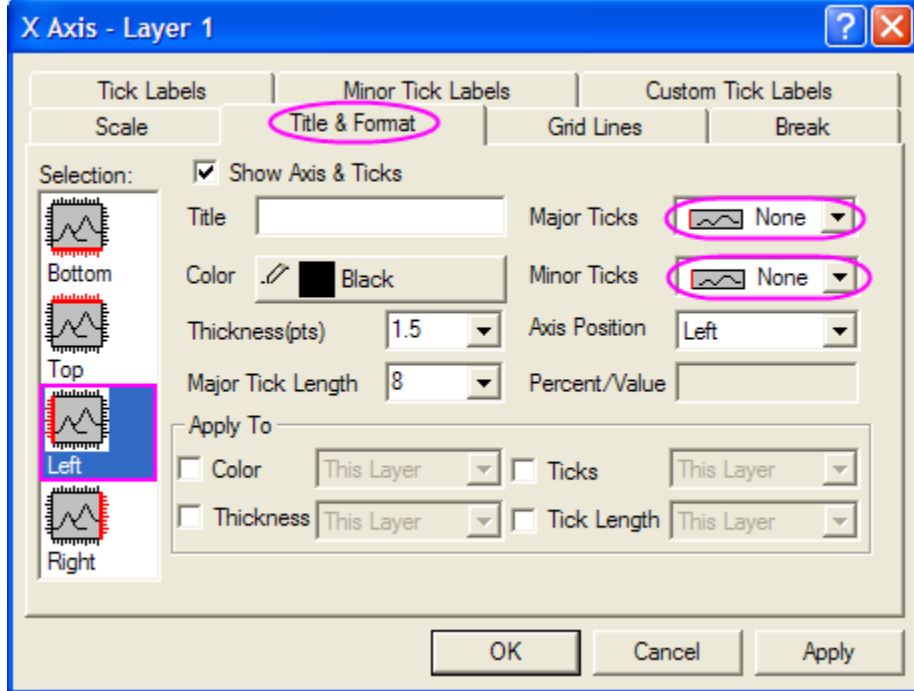
10. V záložce **Title and Format** nastavte **Major Ticks** a **Minor Ticks** na **None**.



11. Klikněte **Left** v **Selection** a zaškrtněte políčko **Show Axis&Tick** a nastavte **Major Ticks** a **Minor Ticks** na **None**.

12. Klikněte **Right** v **Selection** a zaškrtněte políčko **Show Axis&Tick** a nastavte **Major Ticks** a **Minor Ticks** na **None**.

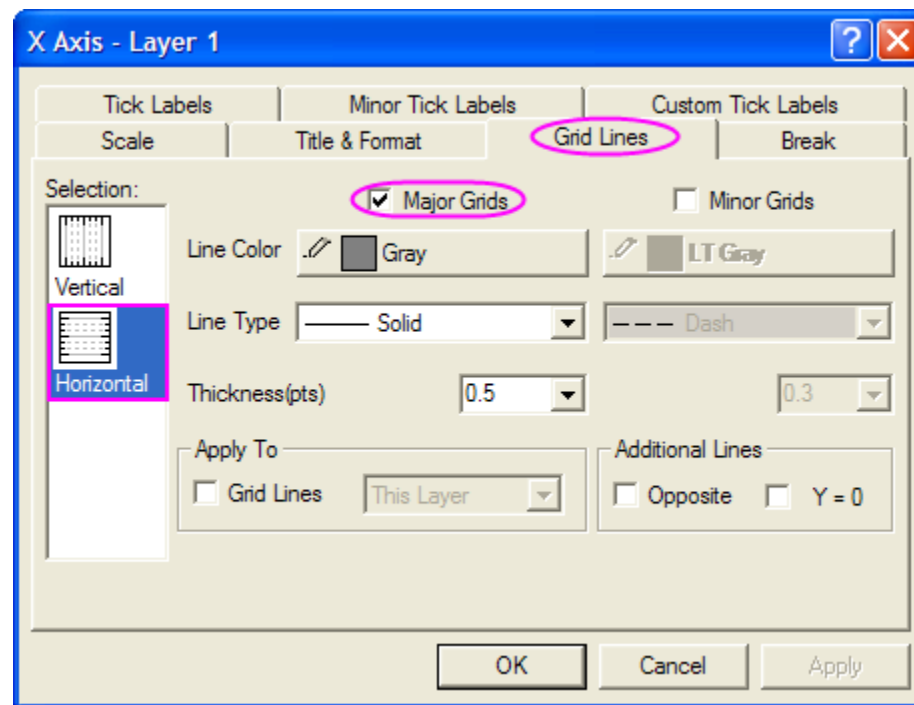
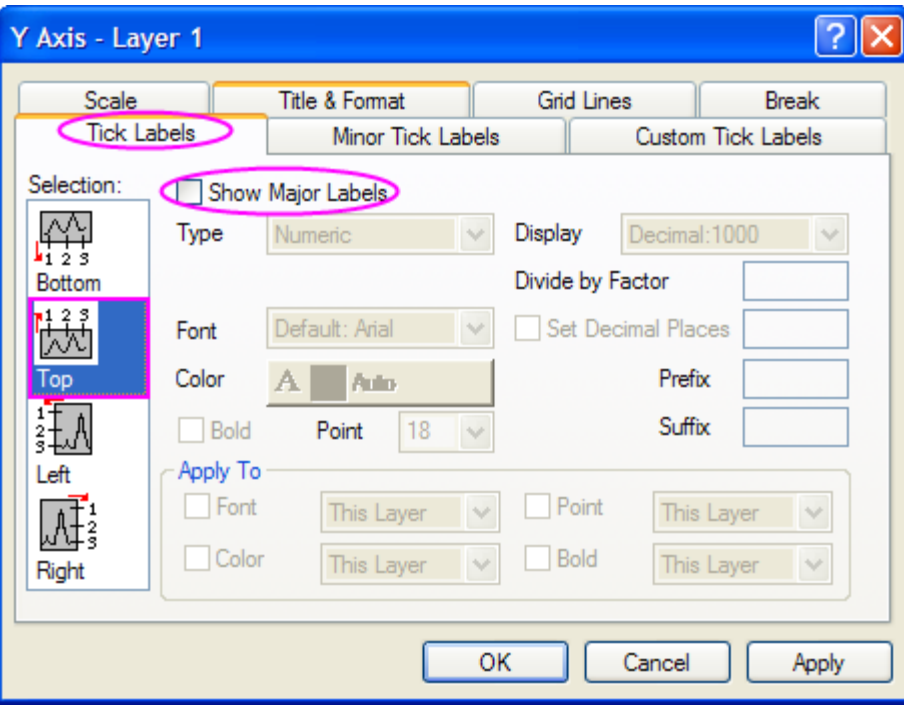
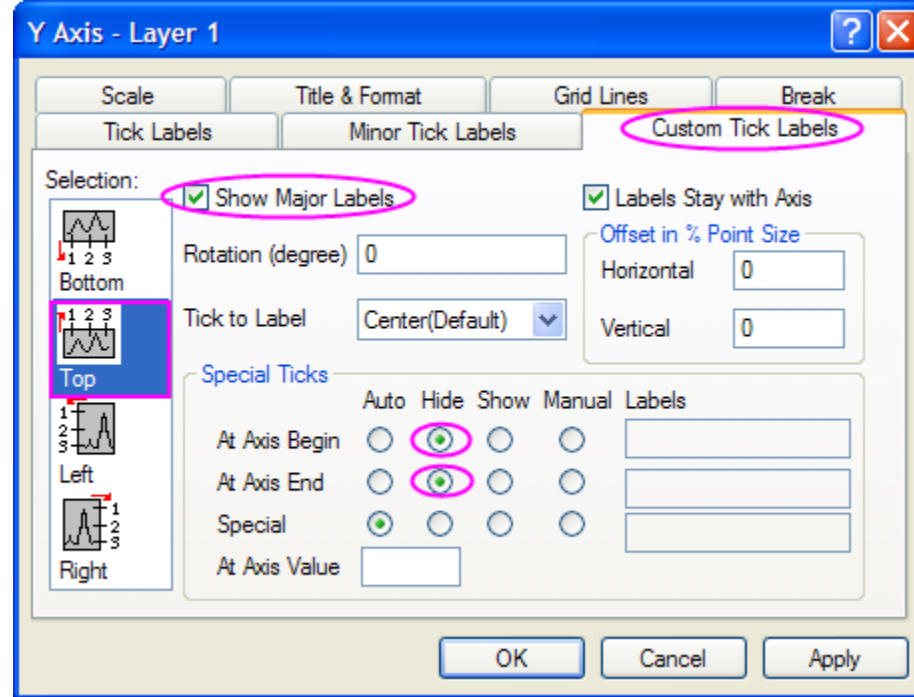
13. Klikněte **Top** v **Selection** a zaškrtněte políčko **Show Axis&Tick** a nastavte **Major Ticks** a **Minor Ticks** na **None**.



14. Vyberte **Custom Tick Labels** a označte **Top** v **Selection**, vyberte přepínač **Hide** s oběma **At Axis Begin** a **At Axis End**.

15. V záložce **Tick Labels** zrušte zaškrtnutí políčka **Show Major Label**.

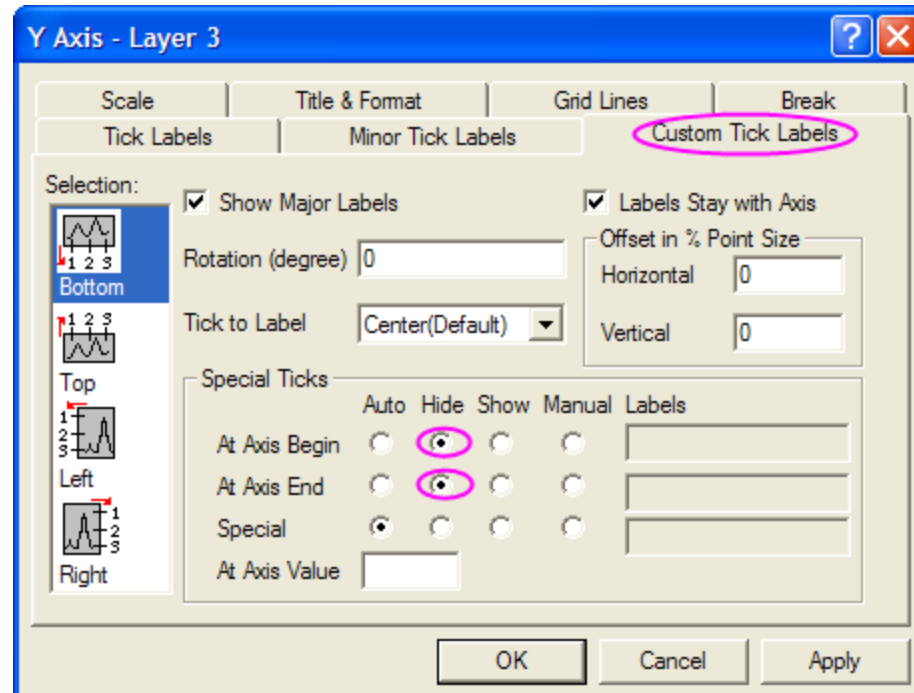
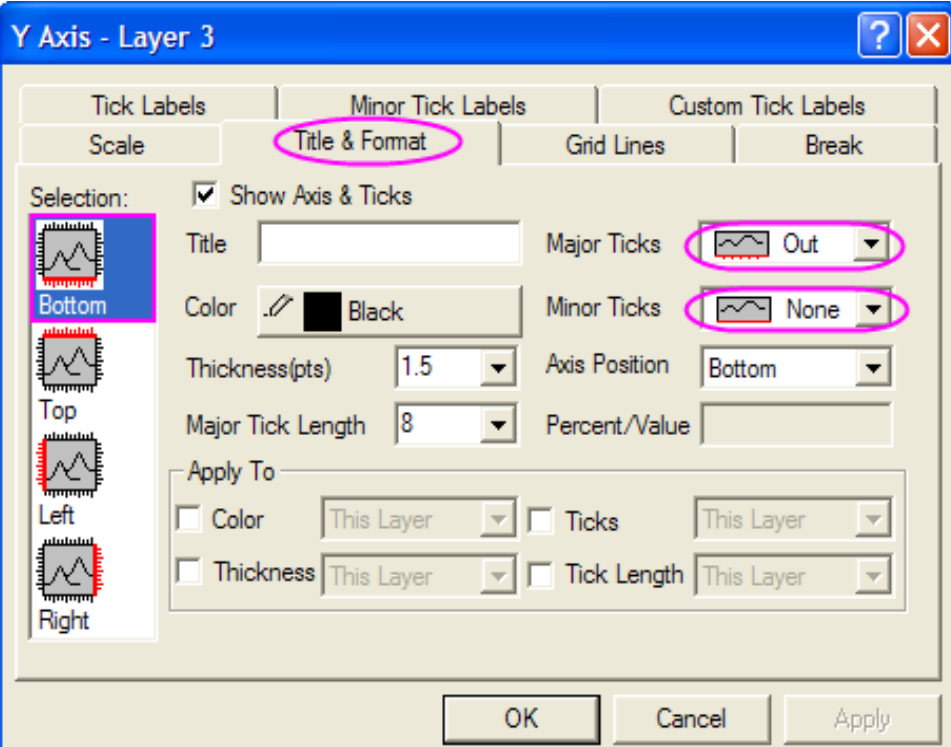
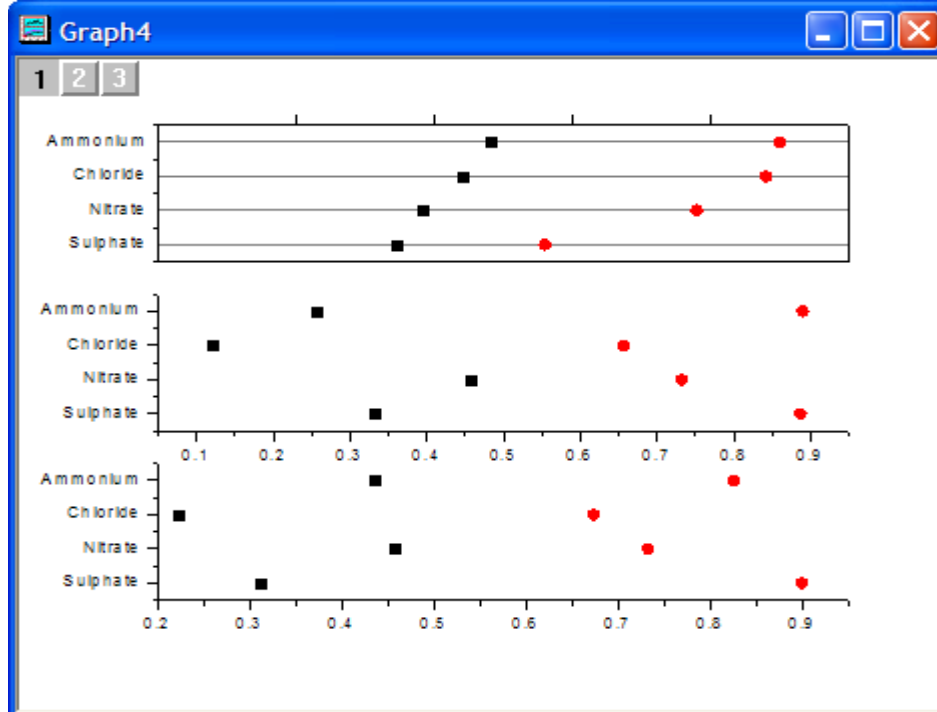
16. V záložce **Grid Lines** vyberte **Horizontal** ve **Selection**, potom vyberte políčko **Major Grid**. Také si vyberte barvu a styl čar mřížky. Klepněte na **OK**.



17. Aktivace **Layer 2**, opakujte kroky 9 až 12 a krok 16.

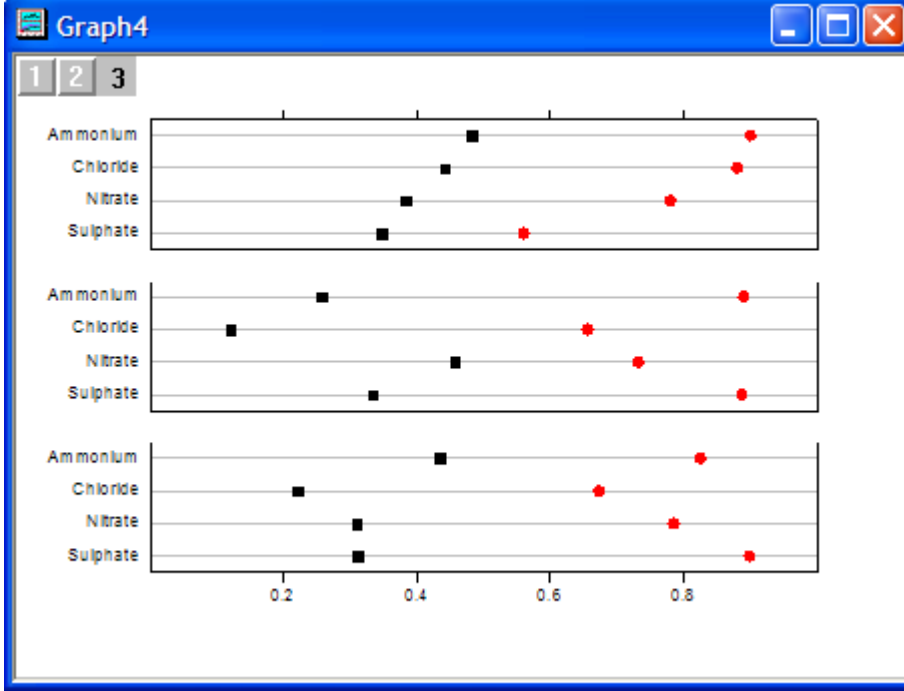
18. Aktivace **Layer 3**, rovněž opakujte kroky 9 až 12 a krok 16 kromě kroku 10. V záložce **Title and Format** nastavte **Major Ticks** na **Out** jako ven a **Minor Ticks** na **None**.

19. Vyberte **Custom Tick Labels**, zvýrazněte **Bottom** v **Selection**, vyberte přepínač **Hide** na **At Axis Begin** a **At Axis End**.



20. Klikněte na **OK**. Graf bude dle obrázku vpravo.

21. Dvoj-klikem na graf se otevře dialog **Plot Details**. Změňte možnosti, jako na následujícím obrázku dole.



Plot Details

Symbol Drop Lines **Group** Offset Label

Group Members

■ Book1_...
● Book1_...

Plot Type: Scatter

>> Workbook OK Cancel Apply

Plot Details

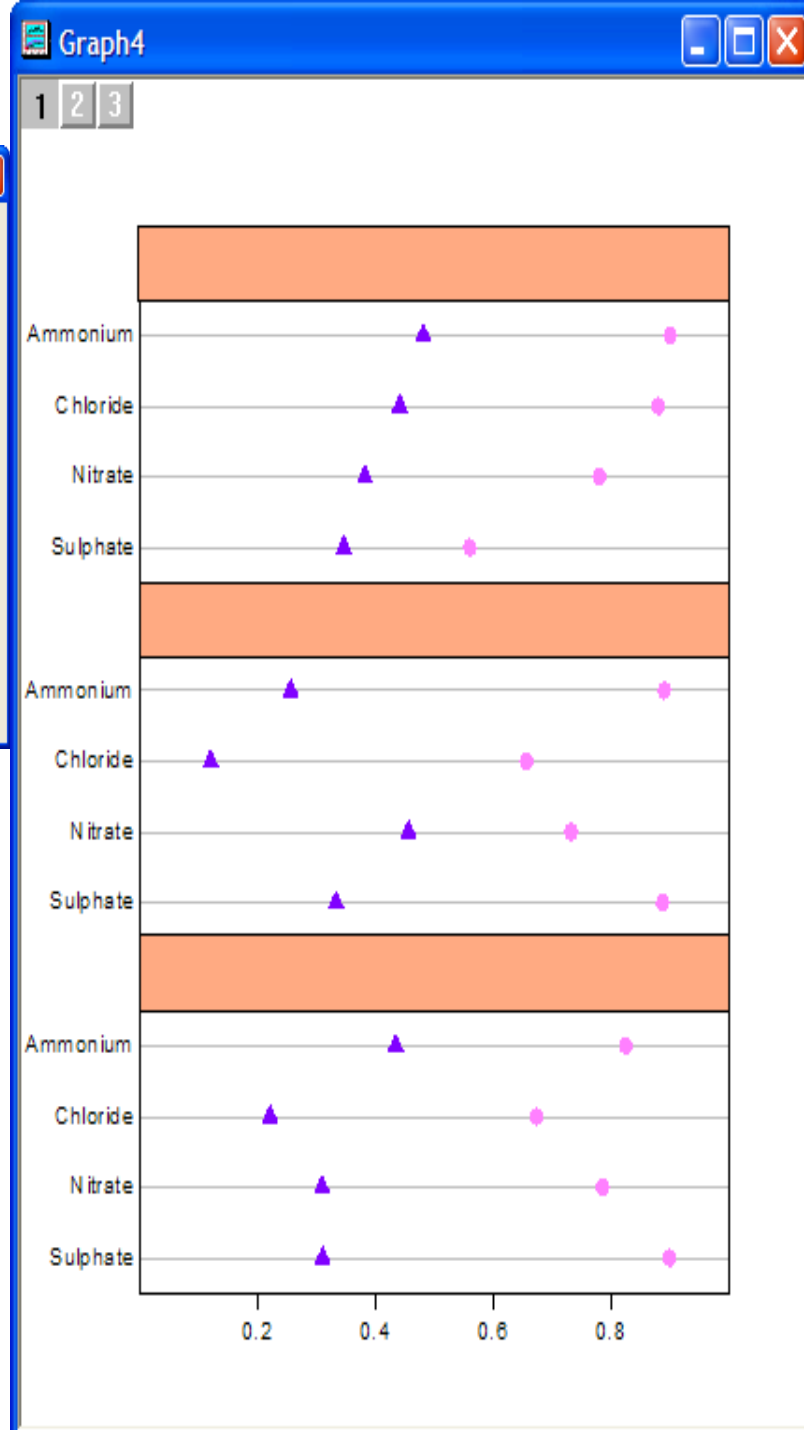
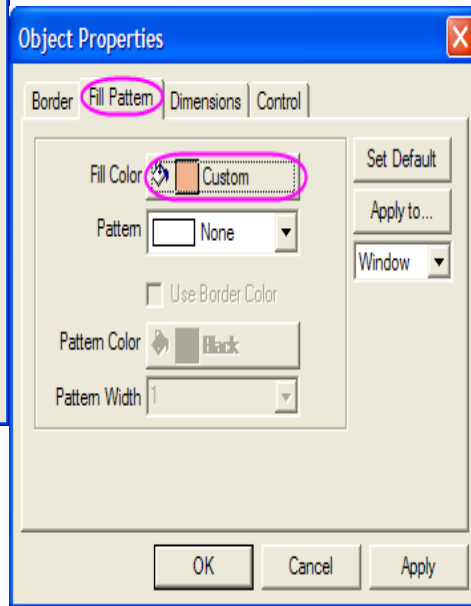
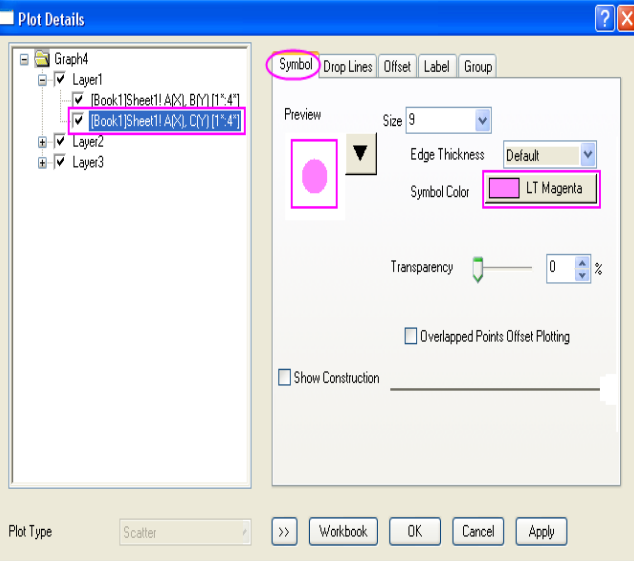
Symbol Drop Lines Offset Label Group

Preview: [Purple Triangle]

Symbol Color: **Violet**

Plot Type: Scatter

>> Workbook OK Cancel Apply



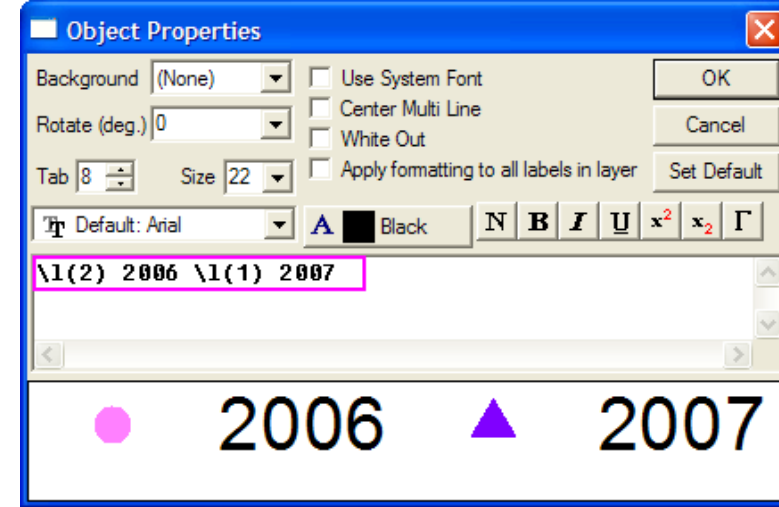
22. Opakujte krok 21 pro **Layer 2** a **Layer 3**. Klepněte na **OK** a uvidíte tento graf vpravo.

23. Vyberte v panelu **Tools** tři obdélníky v grafu. Pro každý obdélník:

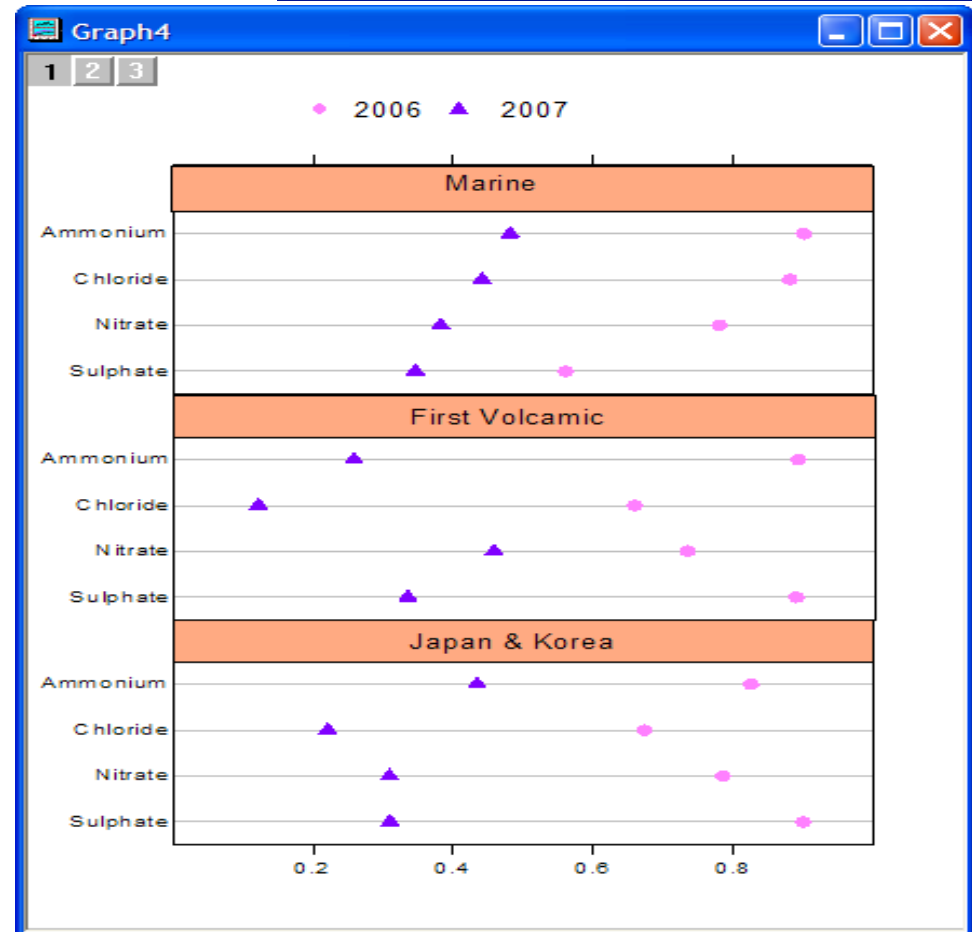
1. Dvoj-klikem na obdélník se otevře dialog **Object properties**.
2. V **Fill Pattern** nastavte požadovanou barvu **Fill Color**.
3. V záložce **Dimensions** upravte velikost a polohu obdélníku, pokud si to přejete.

24. Vyberte nástroj **Text** v Tools-panelu nástrojů a klikněte na příkaz uvnitř obdélníků přidat text, který si přejete.

25. Vyberte **Graph, New Legend**. Přesuňte legendu na vhodné místo, klikněte na něj pravou myší a zvolte **Properties**. Změňte nastavení jak je uvedeno vpravo:



26. Aktivujte horní osy **X** u **Layer 1** a ji přesuňte ji na vhodné místo. Nyní je diagram dokončen dle obrázku vpravo.



5.6.4 Vícerozměrné proložení histogramu křivkami (Add multiple fitted curves in a Histogram)

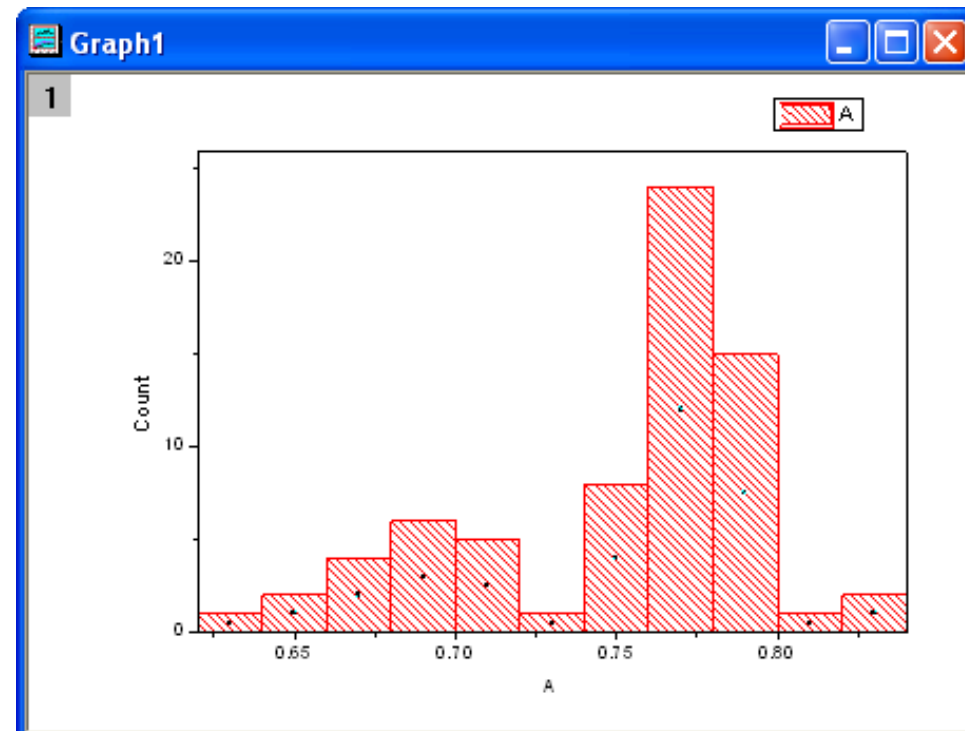
Po nakreslení histogramu umožní Origin překrýt tento distribuční křivkou výběru normálního, logaritmicko-normálního, Poissonova, exponenciální, Laplaceova, nebo Lorentzova. Chcete-li přidat více distribučních křivek do histogramu, postup obsahuje vždy několik kroků.

Kroky:

Nastavte sloupec za osu **Y** (čili označíte sloupec a pak **Column, Set As Y**). Grafem těchto dat bude histogram, a to kliknutím **Plot, Statistics, Histogram** z menu.

A. Frekvence čítač

1. Otevřete dialog **Frequency Count** volbou **Statistics, Descriptive Statistics, Frequency Count**.
2. Kliknutím na **OK**. Bude generován nový list.



B. Proložení píku:

1. Vyberte **col**(počet) a potom otevřete dialog **Peak Analyzer** v **Analyzis, Peak and Baseline, Peak Analyzer**.
2. Na úvodní stránce vyberte **Fit Peaks** za **Goal**, a pak klikněte na **Next**.
3. Na stránce **Baseline Mode** nastavte nulovou linii na **Y = 0**, když je políčko **Custom** zaškrtnuté.
4. Klepněte na **Next** dvakrát a přejděte na stránku **Find Peaks**. Klikněte na **Find** při zaškrtnutí políčka **Enable Auto Find**, aby se našly dva vrcholy.
5. Klikněte na **Next** a znovu otevřete **Fit Peaks**. Klikněte na **Fit Control** ve spodní části této stránky a otevře se dialog **Peak Fit Parameters**.
6. V tomto dialogovém okně je defaultní funkce **Gaussian**, což je ta správná funkce k normalizování dat. Zavřete nyní dialog **Peak Fit Parameters** a vraťte se do dialogu **Peak Analyzer**. Klepnutím na **Finished** dokončete proložení křivkou.

C. Přidávání prokládajících křivek:

1. Aktivujte histogram a přidejte vrstvu výběrem vhodného grafu: **Graph, New Layer(Axes), Right-Y.**

2. Klikněte pravou myší na ikonu **Layer 2** a zvolte **Plot Setup** z krátkého menu a otevře se dialog **Plot Setup**.

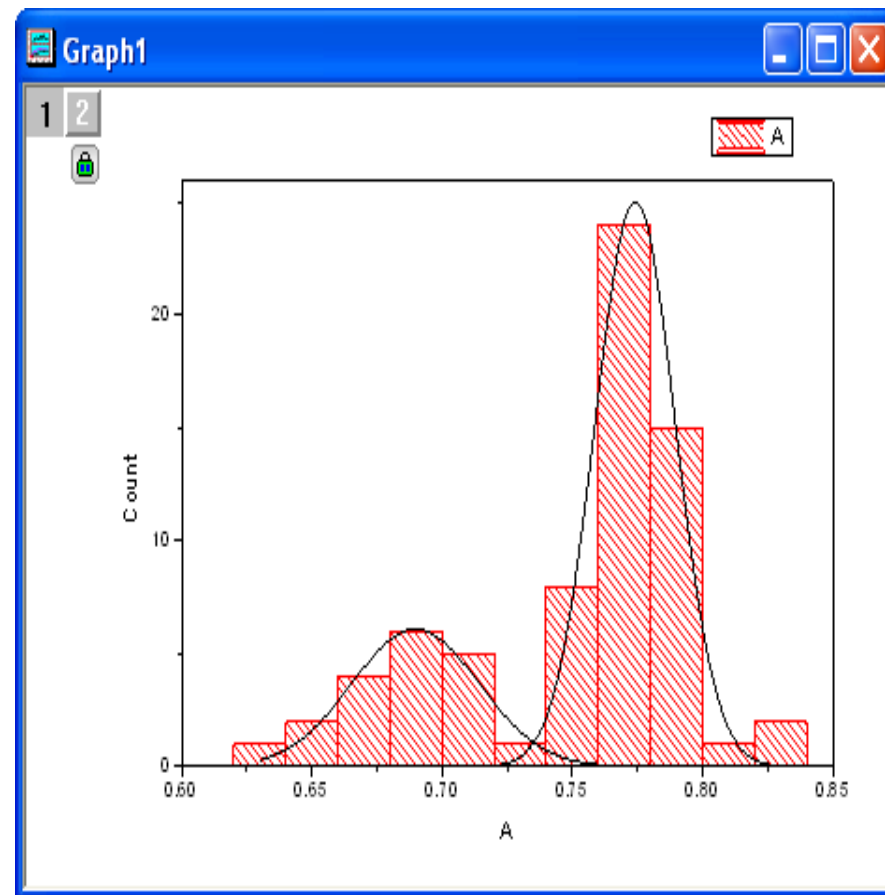
3. Vyberte list **FitPeakCurve1** z horního panelu, pak nastavte **col(A3)** jako **X** a **col(A4)** jako **Y** a přidejte je do vrstvy **RightY** v dolním panelu.

4. Proveďte to stejné pro **col(A5)** a **col(A6)**. Potom oba proložené píky se přidají vrstvy **RightY**.

5. Klikněte na **OK**. Byly přidány dvě proložené křivky do histogramu.

6. Dvoj-klikem na graf se otevře dialog **Plot Details**. Zvolte **RightY** z levého panelu, pak otevřete záložku **Link Axes Scales** v pravém panelu a vyberte položku **Straight(1 to 1)** pro obě osy, jak pro **X Axis Link** tak pro **Y Axis Link**. Kliknutím na **OK** zavřete dialogové okno.

7. Proložené křivky jsou přidány do histogramu ve vhodném měřítku. Výsledkem je graf na obrázku vpravo, když pravá Y-osa byla odstraněna.

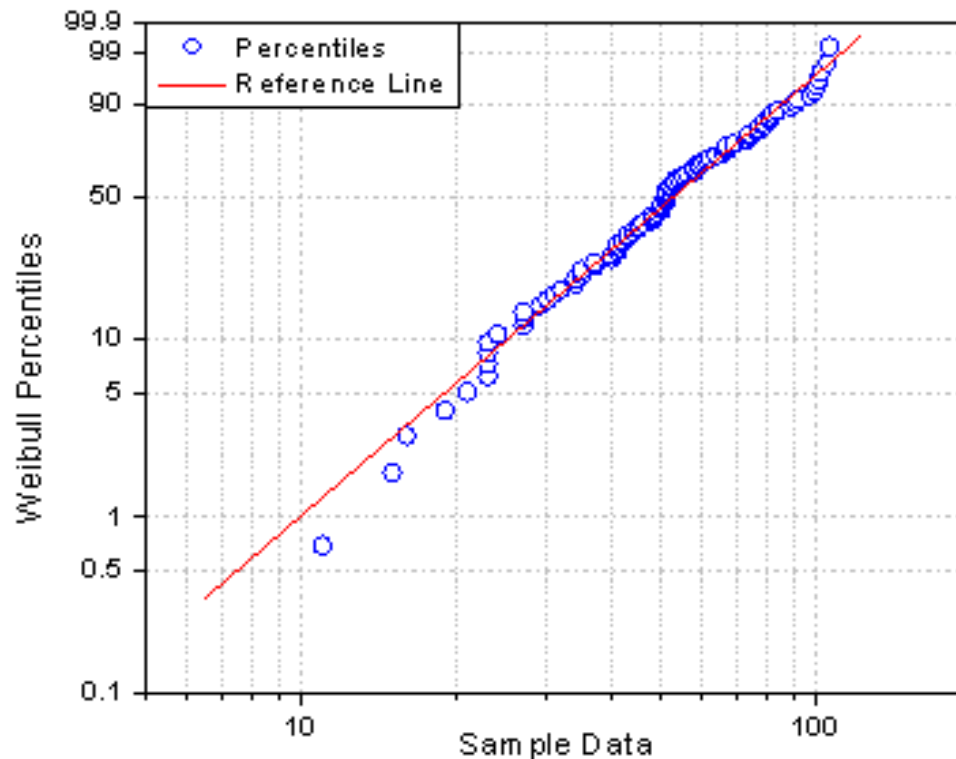


Sample data:

0.631	0.758
0.642	0.758
0.652	0.759
0.662	0.761
0.669	0.761
0.676	0.763
0.677	0.763
0.69	0.763
0.691	0.765
0.696	0.767
0.697	0.768
0.699	0.768
0.699	0.769
0.700	0.769
0.700	0.77
0.708	0.771
0.712	0.771
0.718	0.772
0.731	0.774
0.744	0.775
0.749	0.775
0.751	
0.752	
0.753	

5.6.5 Weibullův pravděpodobnostní graf (Weibull Probability Plot)

Pravděpodobnostní graf vykazuje graf s pozorovaným kumulativním procentem na ose **X** a očekávaným kumulativním procentem na ose **Y**. Weibullův pravděpodobnostní graf se používá k testování, zda soubor dat splňuje Weibullovo rozdělení. Typ X osy a typ Y osy jsou Log10 a Double Log Reciproční. Pokud jsou všechny rozptýlené body blízko či na proložené referenční přímce, lze uzavřít, že datový výběr odpovídá Weibullovu rozdělení.

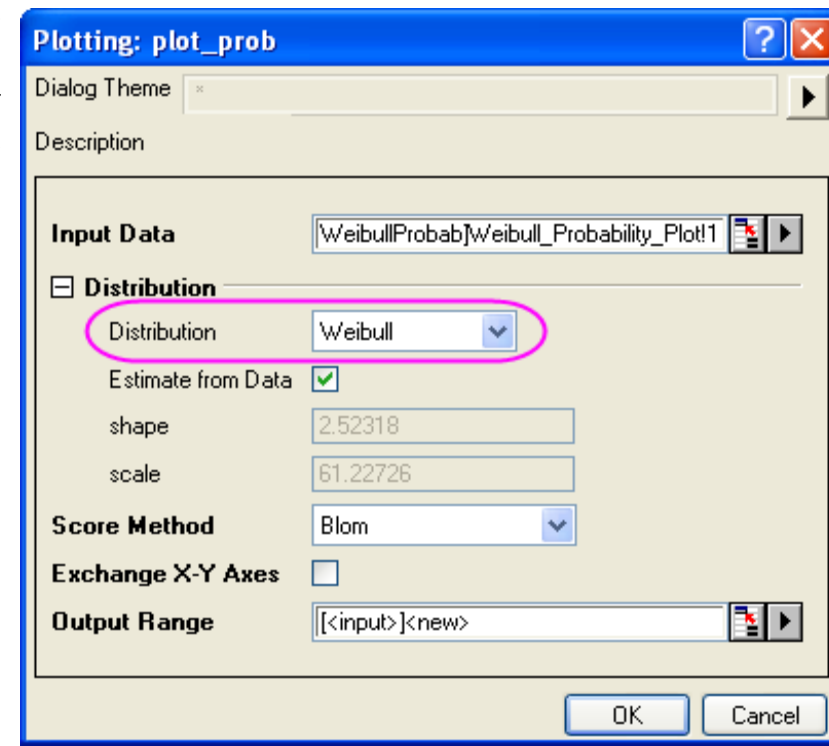
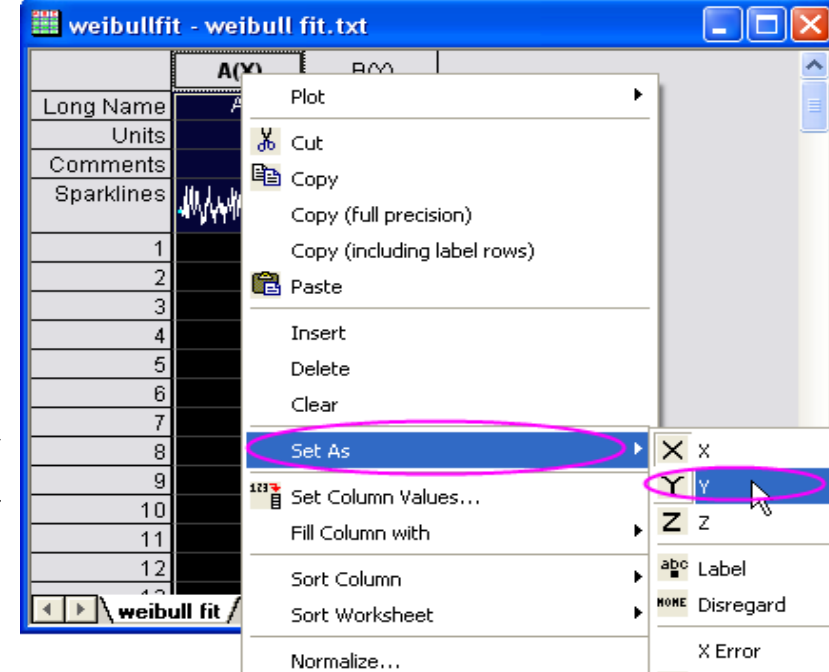


Tutorial je spojen s projektem `\Samples\Statistical and Specialized Graphs.opj`

Kroky:

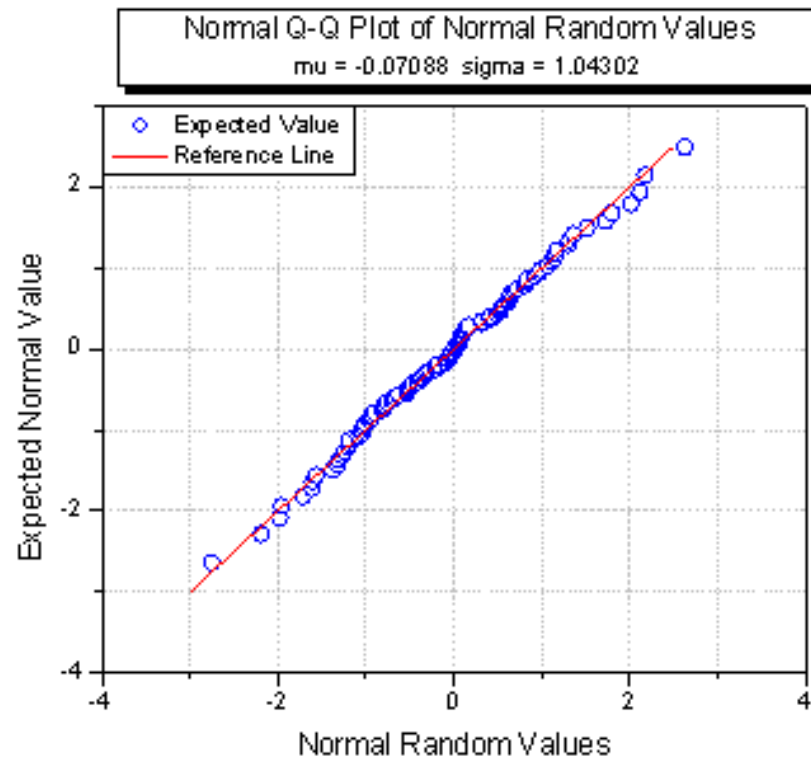
1. Otevřete projekt **Project Explorer** a přejděte do **Statistical and Specialized Graphs, Statistical, Probability, QQ Plot**. Aktivujte list **Probability Plot Data**. Zvýrazněte **col(A)** a klikněte pravou myší a vyberte **Set As: Y**.

2. Udržujte tento sloupec aktivním a zvolte **Plot, Statistics, Probability Plot**. Otevře se dialog **plot_prob**. V tomto dialogu změňte **Distribution** na **Weibull** a potom klikněte na **OK** a vytvoří se Weibullův pravděpodobnostní graf sloupce A.



5.6.6 Kvantil-quantilový graf (Q-Q Plot)

Q-Q plot je grafický diagnostický způsob testování, zda soubor dat vykazuje dané (či právě testované) rozdělení. V Q-Q grafu jsou experimentální hodnoty na ose **X** a očekávané teoretické hodnoty na ose **Y**. Pokud jsou všechny rozptýlené body jsou blízko nebo přímo na prokládané referenční přímce lze uzavřít, že výběr dat sleduje dané a zde testované rozdělení.

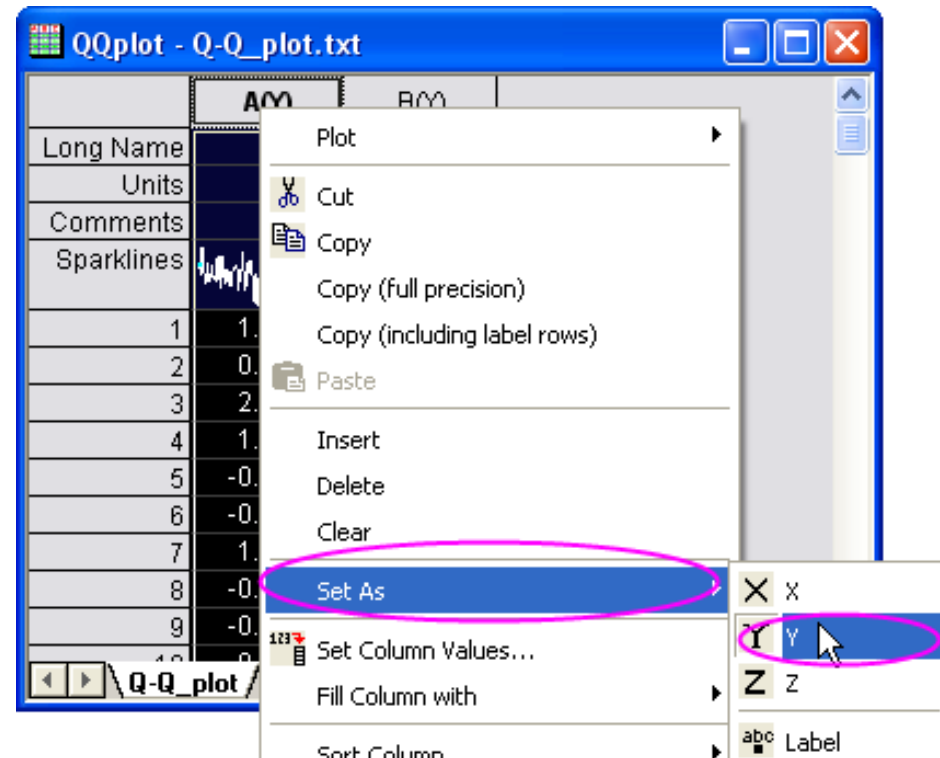
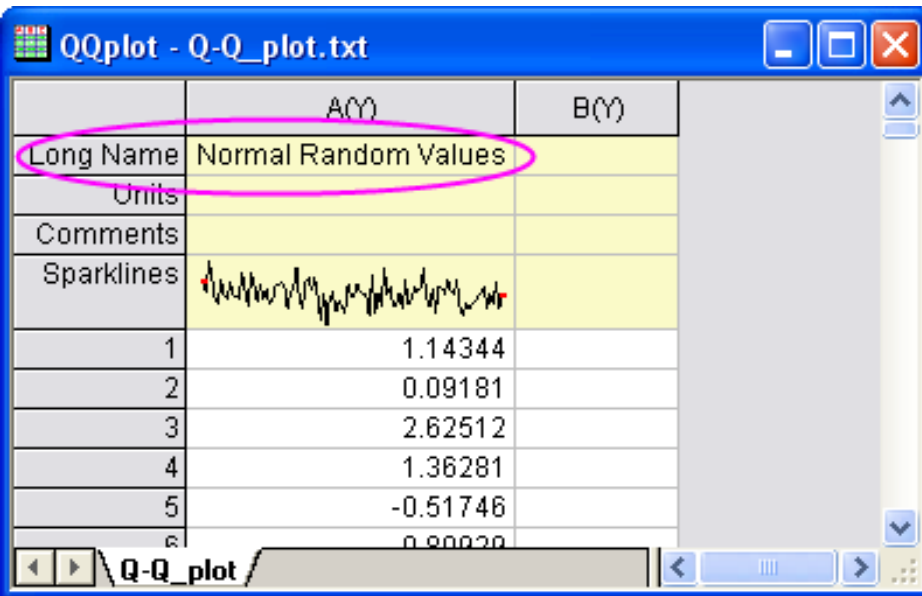


Kroky

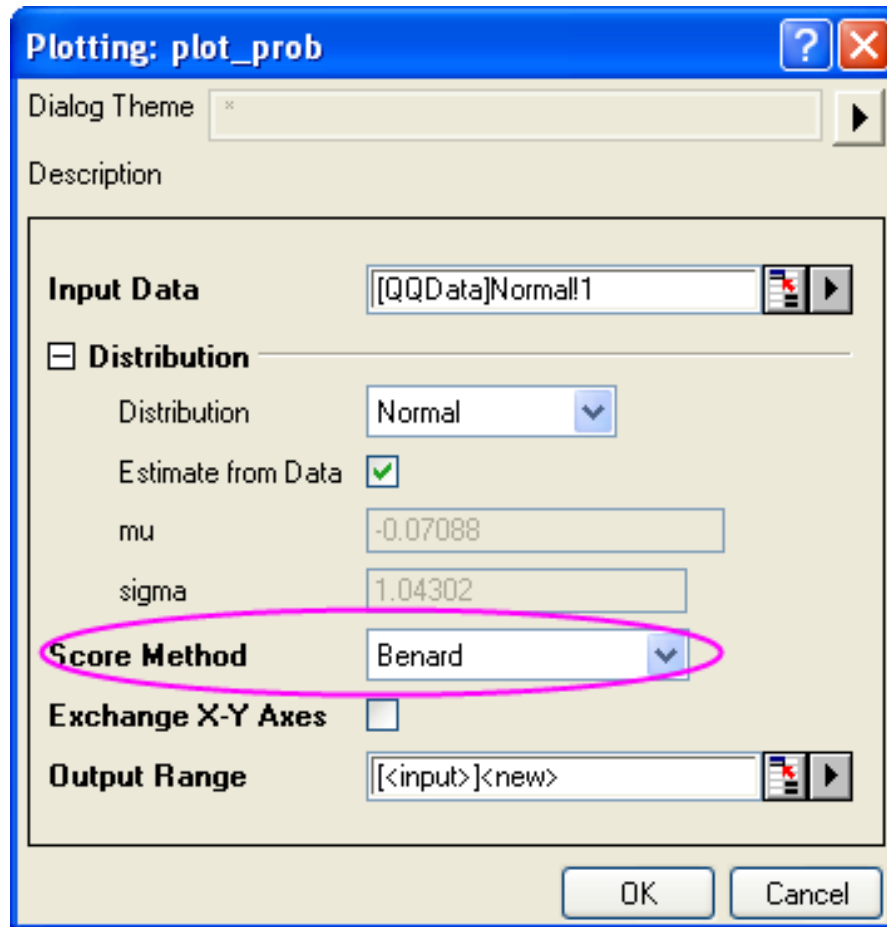
Tutorial je spojen s projektem statistické a specializované grafy: **\Samples\Statistical and Specialized Graphs.opj**.

1. V **Project Explorer** otevřete složku **Statistical and Specialized Graphs, Statistical, Probability, QQ Plot**. Aktivujte list **Q-Q Plot Data**, označte sloupec **A** a klikněte pravou myší a vyberte **Set As: Y**.

2. Přidejte dlouhý název tohoto sloupce, řekněme **Normal Random Values**.



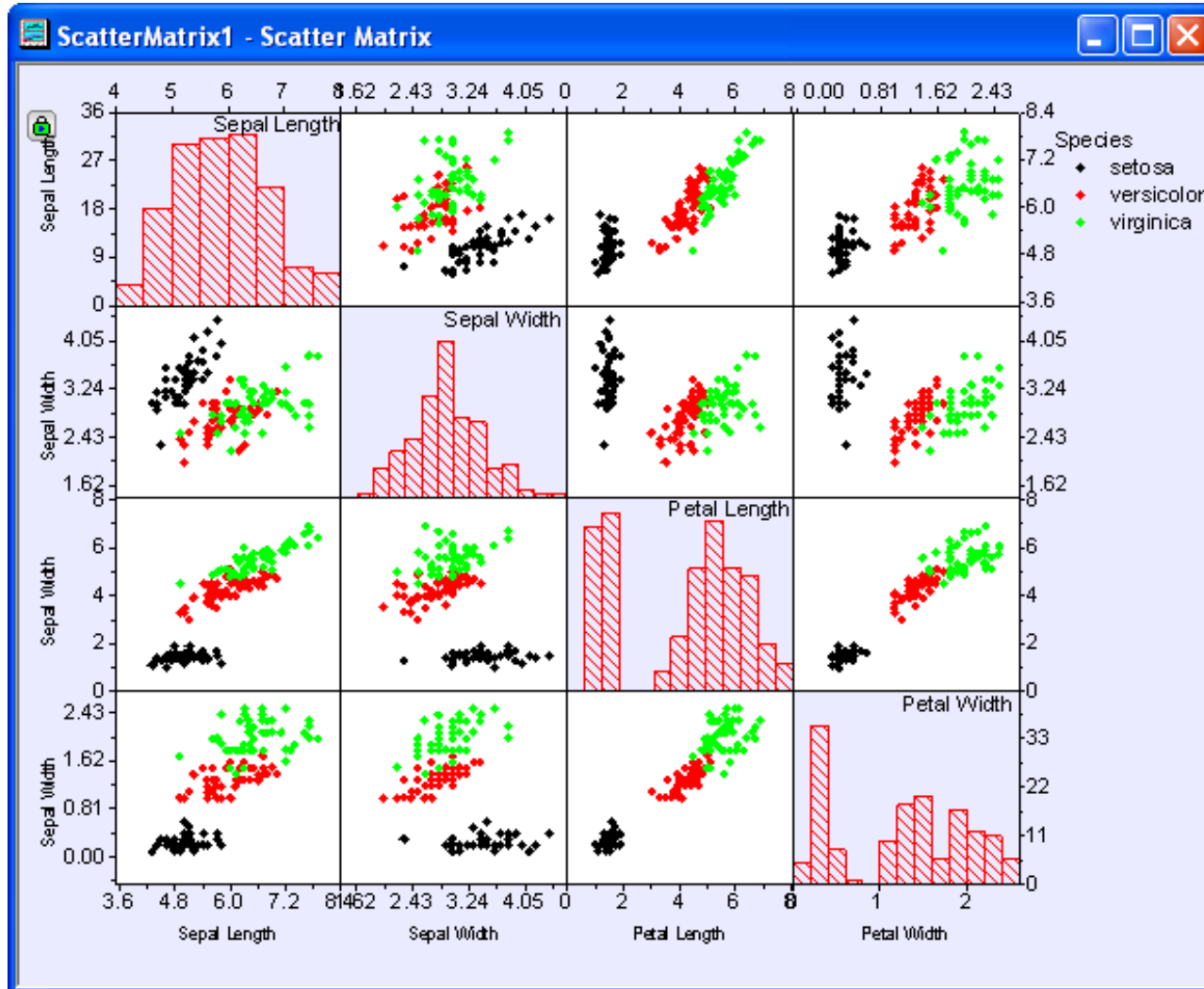
3. Zvýrazněte tento sloupec a vyberte **Plot, Statistics, Q-Q Plot** a otevřete dialog **plot_prob**. V tomto dialogovém okně, změňte **Score Method** na **Benard**.



4. Kliknutím na **OK** vytvořte normální Q-Q graf sloupce A.

5.6.7 Maticový (rozptylový) graf (Scatter Matrix)

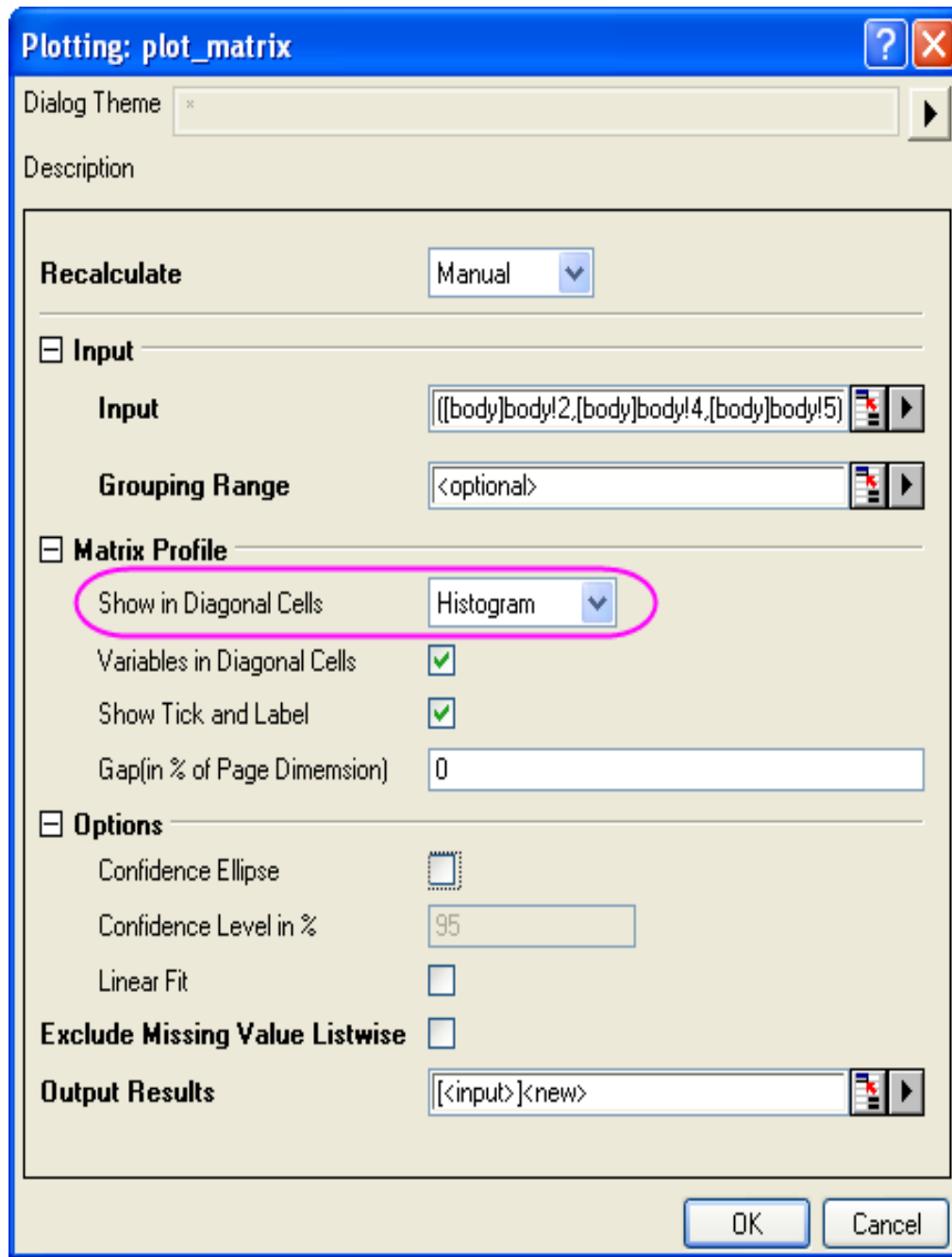
Maticový graf je tvořen z několika dvojic rozptylových ploch proměnných, prezentovaných ve formě matice. Užívá se k analýze, zda jsou proměnné korelovány a zda korelace je pozitivní nebo negativní.



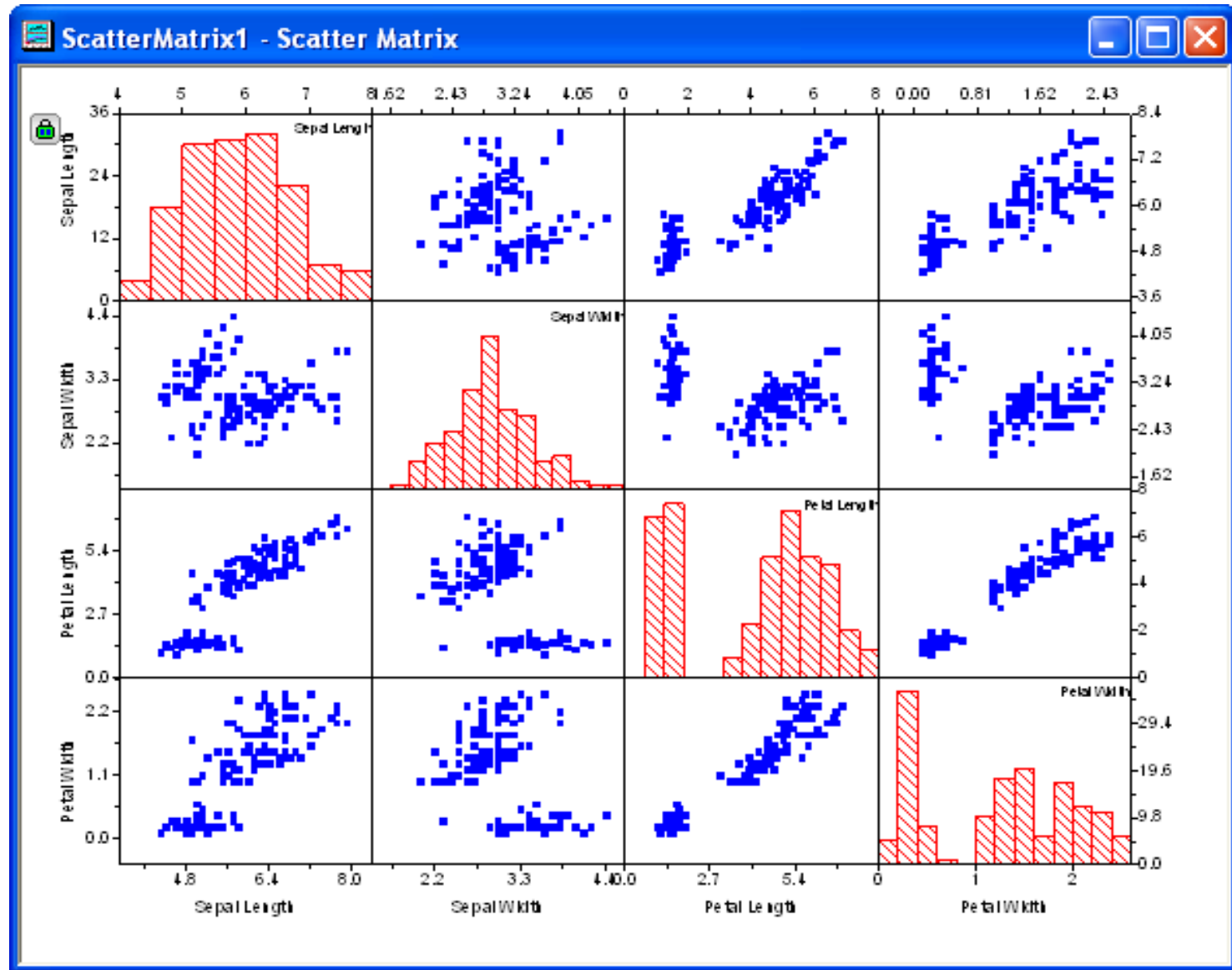
A. Vytvoření maticového grafu rozptýlení

Kroky:

1. Začněte s prázdným listem, vyberte **File, Import, Single ASCII...** a otevřete jednoduchý ASCII dialog ve složce **\Samples\Statistics** a nainportujte soubor **Fisher's Iris Data.dat**.
2. Označte sloupce **(A)** až **(D)** a pak vyberte **Plot, Statistics, Scatter Matrix**.
3. V dialogu vyberte **Histogram** v **Show in Diagonal Cells**.



4. Klepnutím na **OK** zavřete dialogové okno. Graf by měl vypadat takto:



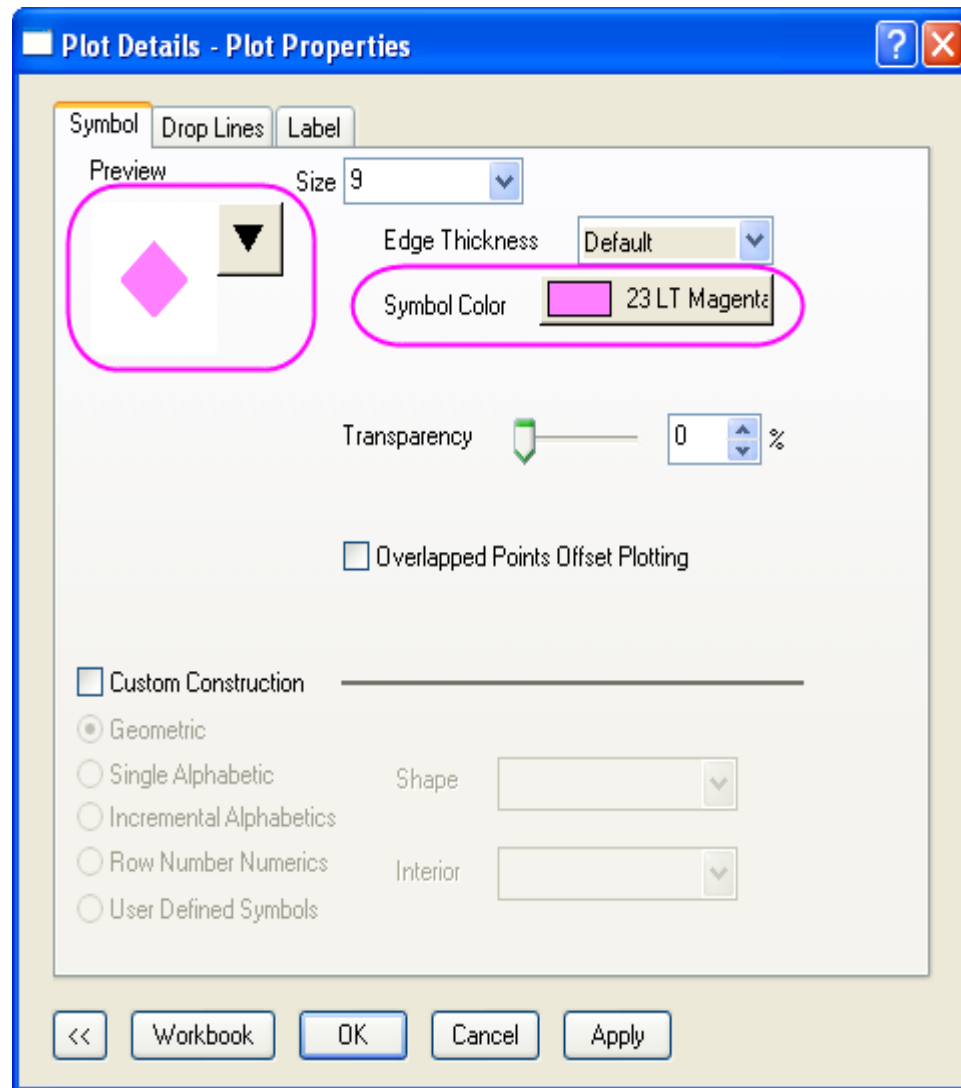
List **PlotData1** pro maticový graf je vytvořen ve stejném sešitu.

B. Přizpůsobení maticového grafu rozptýlení

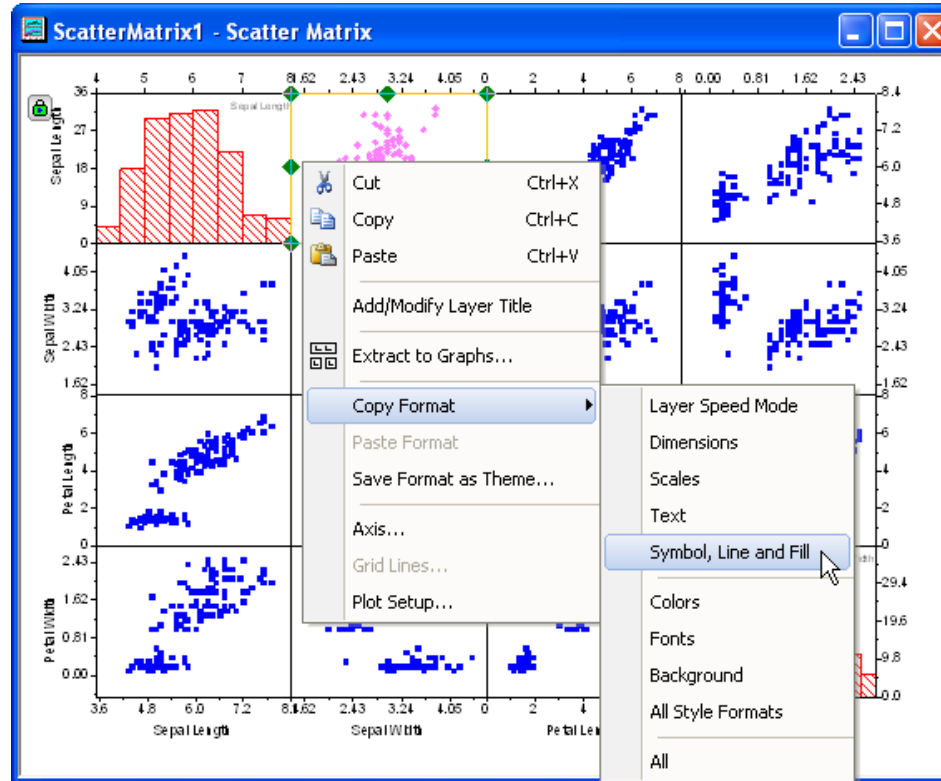
Existuje více vrstev v bodovém maticovém grafu rozptýlení. Tato část ukáže, jak lze přizpůsobit barvu pozadí, typ a barvu grafu a popisky os maticového grafu

Kroky:

1. Dvojklikem na vrstvu s výjimkou diagonálních buněk otevřete dialogové okno **Plot Details**. Určete typ a barvu symbolu, jak je znázorněno na následujícím obrázku, a pak klikněte na **OK**.



2. Kliknutím vyberte vrstvu, kterou je třeba aktualizovat, pak klikněte pravou myší na graf a zvolte **Copy Format, Symbol, Line and Fill**.

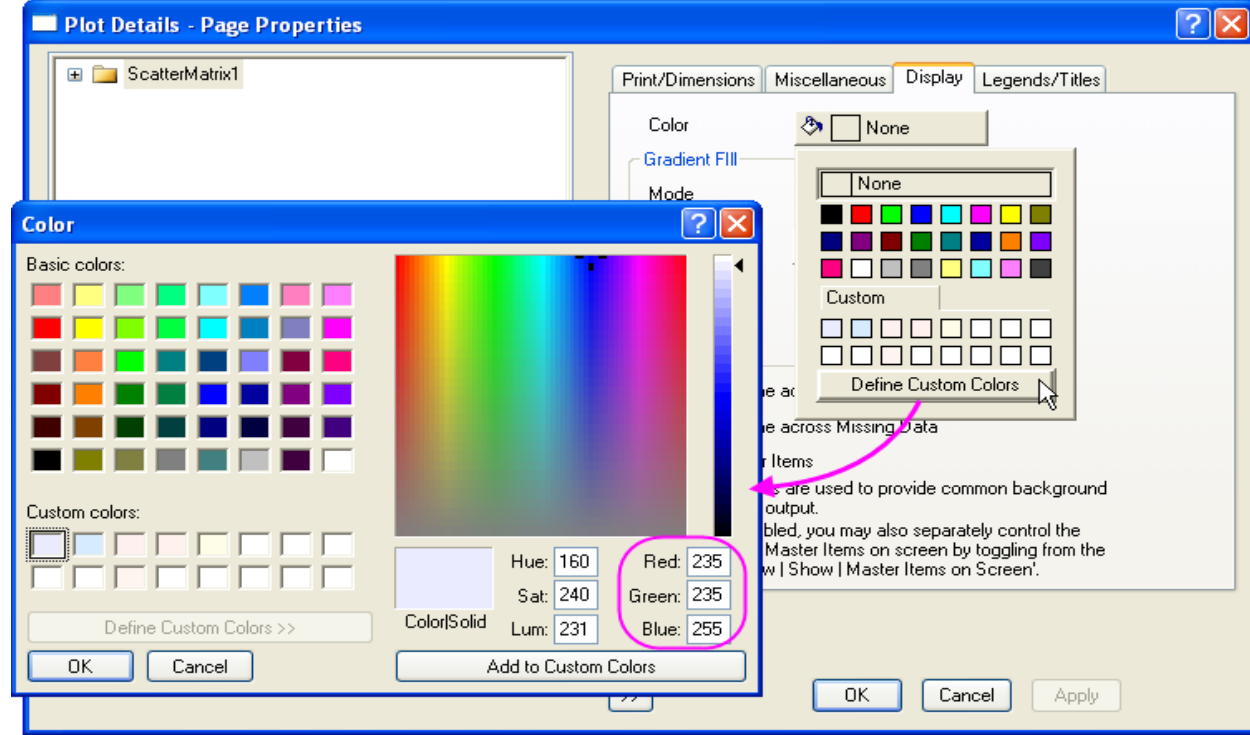


Poté klikněte na bílou plochu grafu, pak klikněte pravou myší a vyberte příkaz **Paste Format**. Můžete použít stejnou metodu pro kopírování barvy symbolu do dalších vrstev.

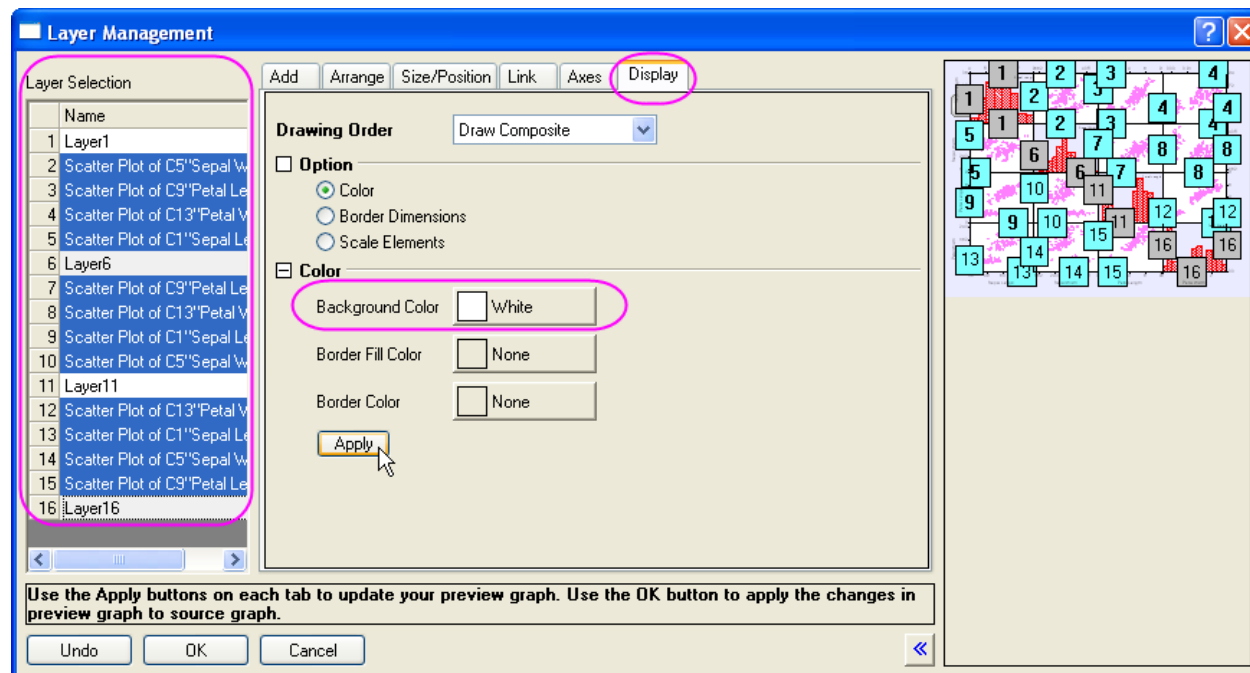
3. Klikněte pravým tlačítkem myši na tuto vrstvu znovu a vyberte položku **Copy Format, Colors**. Poté klikněte na bílou plochu grafu, klikněte pravou myší a vyberte **Paste Format**.

4. Zvolte **Format, Page Properties** k otevření dialogu **Plot Details**. Přejděte na kartu **Display**, klikněte na **Define Custom Colors** pro volbu barev.

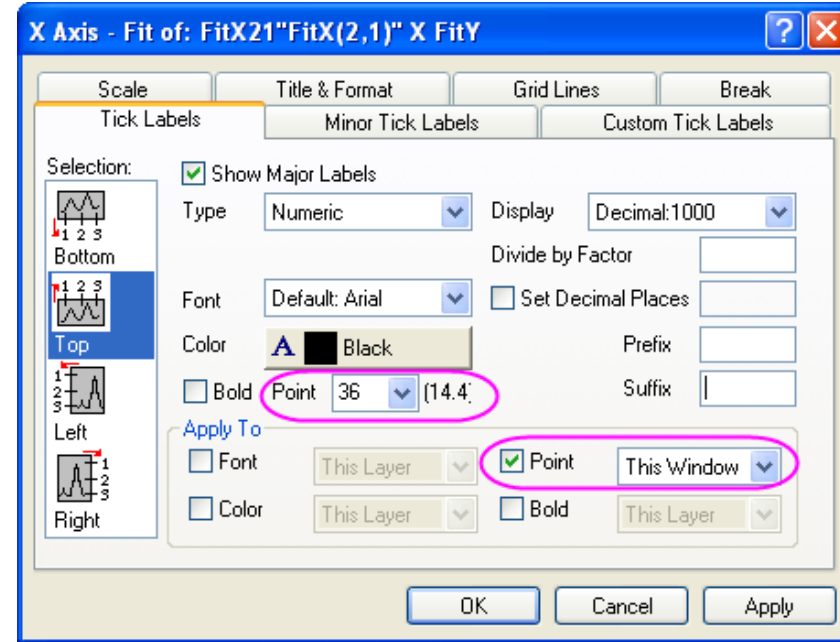
5. V dialogovém okně **Color** určete barvu jako **Red = 235, Green = 235, Blue = 255**. Pak přidejte vlastní barvu a klikněte na **OK**.



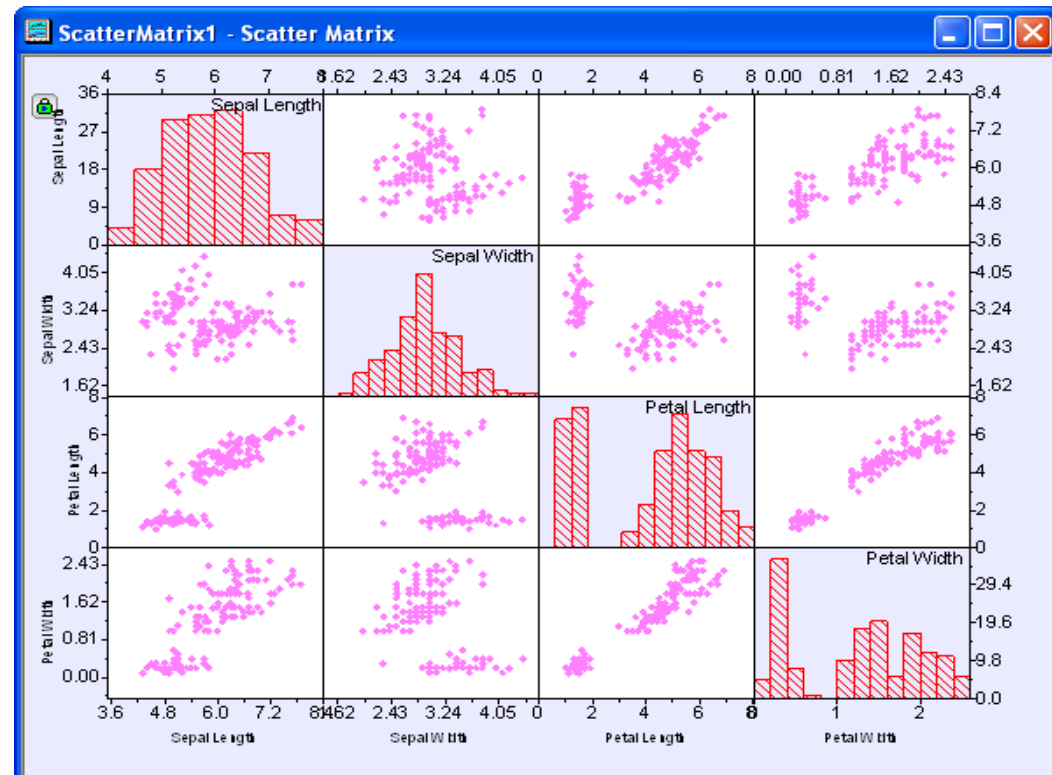
6. Chcete-li určit barvu pozadí vrstev rozptylových grafů, vyberte **Graph, Layer Management**, který otevře dialog a přejděte na záložku **Display**. Označte všechny vrstvy uvedené ve **Layer Selection** kromě diagonálních buněk, a pak určete **White** jako **Background Color**. Klikněte na **Apply**, aby se náhled změnil a objevil v pravém panelu dialogu. Poté klepněte na **OK** v dialogovém okně.



7. Dvoj-klikem na popisek ticku v grafu se otevře dialogové okno **Axis**. V záložce **Tick Labels** přiřadíte **Point** hodnotu **36**. Poté vyberte **Point** v části **Apply** a vyberte **This Window** seznamu drop-down.



8. Zvýrazněte proměnné v diagonálních buňkách a použijte **Size** na **22** ve **Format** k nastavení velikosti na **36**. Graf bude vypadat takto:



C. Přidat rozsah seskupení

1. Chcete-li přidat rozsah seskupení

Grouping Range, klikněte na zelenou ikonu zámku v levém horním rohu. Pak zvolte **Change Parameters**, aby se přivedly zpět k **Plotting, plot_matrix**.

2. Klikněte na trojúhelníkové tlačítko vedle volby **Grouping Range**. Klikněte na **Select Columns** a otevřete dialog **Column Browser** a potom zvolte sloupce **E** (částice) jako rozsah skupiny. Klepněte na **OK**.

3. Klepněte na **OK** v dialogu **Plotting, plot_matrix**. Váš konečný graf by měl vypadat takto:

