Select subset of data by XY coordinates

Author: Greta Linse Peterson, Golden Helix, Inc.

Overview

This script takes an upper and lower bound for two numeric columns and creates a subset spreadsheet for the two columns.

Recommended Directory Location

Save the script to the following directory:

*..\Application Data\Golden Helix SVS\UserScripts\Spreadsheet\Plot\

Note: The **Application Data** folder is a hidden folder on Windows operating systems and its location varies between XP and Vista. The easiest way to locate this directory on your computer is to open SVS and go to **Tools >Open Folder > UserScripts Folder**. If saved to the proper folder, this script will be accessible from the spreadsheet **Plot** menu.

Goal

The purpose of this script is to create a subset spreadsheet based on X and Y coordinates either known in advance or determined by looking at an XY Scatter Plot.

In the example below, a subset spreadsheet is needed for the samples with data points in the area outlined in **Figure 1**.



Figure 1: XY Scatter Plot with area of interest highlighted.

Using the Script

Maria - Sheet 1 [200]								
<u>E</u> dit <u>S</u> el	ect Quality Assurance	<u>A</u> nalysis <u>P</u> lot <u>S</u> cripts	<u>H</u> elp					
🔎 🚎 🔢 🖄 🛅 🗠 🗟 🛄 🤎 🔜 💷 ೫ 🗷 🚏 🗡 🖾 🖉								
	R 5	R 6	C 7	R 8	B 9	R 10 🔺		
Label	sbp	dbp	alcohol use	dose	treat	Lab 1		
1	143.0257	95.35056	Low	0	0			
2	142.4812	90.24244	High	0.25	1			
3	145.4304	92.86172	None	0.5	1			
4	144.9249	92.4967	High	0	0			
5	154.1325	103.243	High	0	0			
6	141.7159	94.22128	High	0.25	1			
7	144.5358	92.53854	Med	0.5	1			
8	133.2964	86.04026	None	0	0			
9	145.2573	94.75519	High	0.25	1			
10	147.1132	95.02317	Med	0	0			
11	148.3448	99.46514	Med	0.5	1			
12	143.7099	95.53665	High	0.5	1			
13	142.5077	92.01928	Med	0.25	1	-		
						•		
Maria - Sheet 1								
	a - Sheet idit <u>S</u> el Label 1 2 3 4 5 6 7 8 9 10 11 12 13 ""	a - Sheet 1 [200] dit Select Quality Assurance R 5 Label Sbp 1 443.0257 2 142.4812 3 144.4812 3 144.9249 5 154.1325 6 141.7159 7 144.5388 8 1132.964 9 145.2573 10 147.1132 11 148.3488 12 143.309 13 142.5077 m	a - Sheet 1 [200] dit Select Quality Assurance Analysis Plot Scripts	a - Sheet 1 [200] dit §elect Quality Assurance Analysis Plot §cripts Help R 5 R 6 C 7 label sbp dbp alcohol use 1 143.0257 95.35056 Low 1 2 142.4812 90.2424 High 3 145.4304 92.86172 None 4 144.9249 92.86172 None 4 144.9249 92.4967 High 5 154.1325 103.243 High 6 141.7159 94.22128 High 6 141.7159 94.22128 High 7 144.5358 92.25854 Med 8 133.2964 86.04026 None 9 145.2573 94.75519 High 10 147.1132 95.02317 Med 11 148.3448 99.45514 Med 12 148.37089 95.53655 High 13 142.5077 92.01928 Med	a - Sheet 1 [200] dit §elet Quality Assurance Analysis Plot §cripts Help	a - Sheet 1 [200] dit §elet Quality Assurance Analysis Plot Scripts Help R 5 R 6 C 7 R 8 B 9 Label sbp dbp alcohol use dose treat 1 143.0257 95.35056 Low 00 00 2 142.4812 90.2424 High 0.25 11 3 145.4304 92.86172 None 0.5 11 4 144.9249 92.4967 High 00 5 1154.1325 103.243 High 00 6 141.7159 94.22128 High 0.25 11 8 133.2964 86.04026 None 00 9 145.2573 94.75519 High 0.25 11 8 133.2964 86.04026 None 00 9 145.2573 94.75519 High 0.25 11 10 147.1132 95.02317 Med 00 9 145.2573 94.75519 High 0.25 11 11 148.3448 99.46514 Med 0.5 11 12 143.7099 95.3565 High 0.5 11 13 142.5077 92.01928 Med 0.25 11 14 142.5077 92.01928 Med 0.25 11 14 142.5077 92.01928 Med 0.25 11 15 142.5573 94.5519 High 0.5 11 16 142.5573 94.5519 Migh 0.5 11 17 144.5458 99.45514 Med 0.5 11 18 142.5077 92.01928 Med 0.25 11 19 142.5077 92.01928 Med 0.25 11 19 142.5077 92.01928 Med 0.25 11 10 143.5077 92.01928 Med 0.25 11 10 143.5077 92.01928 Med 0.25 11 10 143.5077 92.01928 Med 0.25 11 10 143.		

1. Open a spreadsheet containing numeric columns, such as in Figure 2.

Figure 2: Spreadsheet containing numeric columns.

2. Select **Plot > Select subset of data by XY coordinates** and enter in the appropriate information to the parameters dialog, see **Figure 3**.

🛃 Select Subset of Data by XY Co 💡 🔀							
R Chi-Squared P	Select Column						
Lower bound for X axis:	.01						
Upper bound for X axis:	.99						
R Chi-Squared expect	Select Column						
Lower bound for Y axis:	.01						
Upper bound for Y axis:	.99						
	ОК	Cancel					

Figure 3: Select subset of data by XY coordinates parameters dialog.

The resulting spreadsheet will be a row subset spreadsheet with only the rows (samples) that met the specified criteria.