**13. 绘图Ⅰ—通用语法**

图形是展示数据的重要方法，图形的形象直观是数据报表无法替代的。SAS中超过80%的统计分析过程，都可以利用ODS Graphics模块生成图形独立输出。

SAS 9.2版以来，SAS画图模块正变得越来越灵活、精良和易于使用。在一些分析过程步（PROCs） 中，ODS Graphics可以自动的生成一些图形，而不需要额外的代码。这使得用户多了一个选择，即可以使用默认的输出生成图形，也可以自己来创造个性化的图形。

**（一）ODS Graphics概述**

**一、 打开/关闭ODS Graphics**

SAS 9.2开始，将“ODS Graphs”从SAS/GRAPH移到Base SAS 9.3，并增加了新的功能更加强大的绘图过程步：

SGPLOT, SGSCATTER, SGPANEL, SGRENDER.

这些过程步，用很少的代码就能生成高质量的图形。

注：早期SAS是用GPLOT和GCHART过程步绘制图形。

从SAS 9.3开始，ODS Graphics模块默认是关闭的，使用它绘图之前需要先打开，使用如下语句：

ODS GRAPHICS ON;

不用了可以保持其打开状态（影响软件启动速度），若要关闭用：

ODS GRAPHICS OFF;

ODS Graphics相比SAS/GRAPH具有更先进的特性：直接将图形以图片格式输出（默认输出路径是work临时库），在结果查看器中默认以html格式显示。

**（二）PROC SGPLOT过程步**

**一、图形分类**

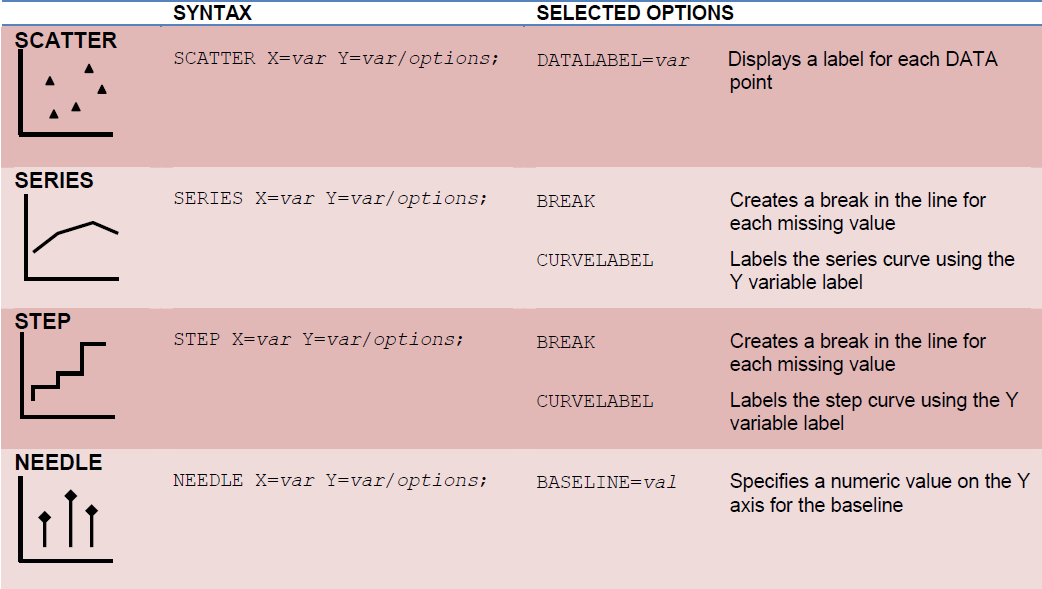
PROC SGPLOT过程步可以绘制16种图形，归为5类：

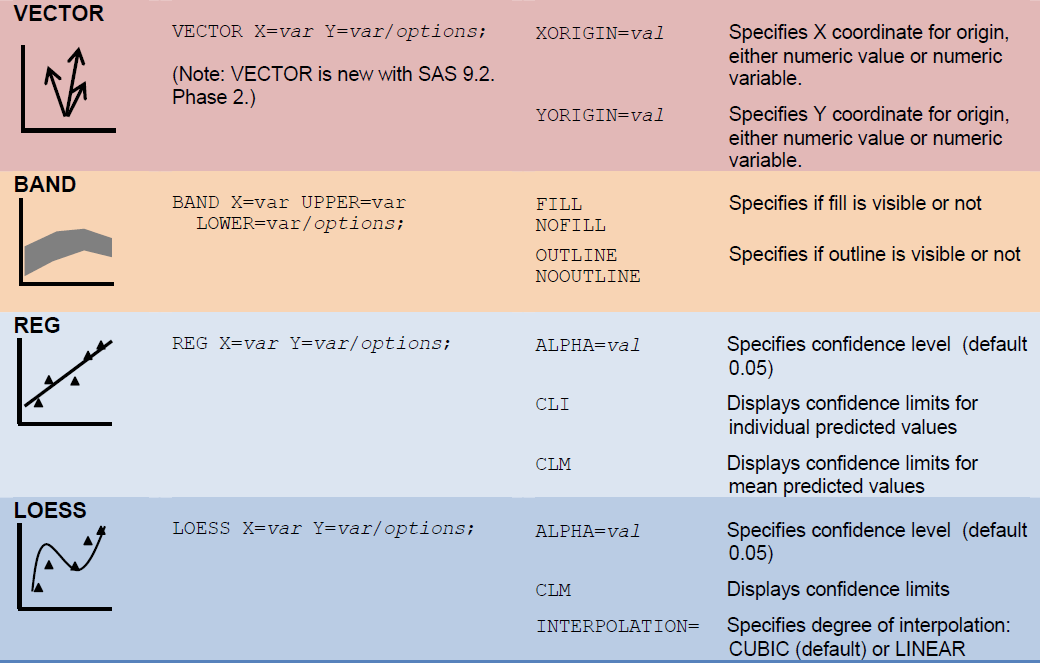
|  |  |  |
| --- | --- | --- |
| **图形类** | **图形名** | **基本语法** |
| X-Y图 | ①散点图（SCATTER） | PLOTNAME X=var Y=var / options;  注：PLOTNAME为SCATTER、SERIES等图形名。 |
| ②时间序列图（SERIES） |
| ③阶梯图（STEP） |
| ④针形图（NEEDLE） |
| ⑤向量图（VECTOR） |
| 带状图 | ⑥带状图（BAND） | BAND X=var UPPER=var  LOWER=var / options; |
| 拟合-  置信图 | ⑦线性回归图（REG） | PLOTNAME X=var Y=var / options; |
| ⑧局部加权回归（LOESS） |
| ⑨B-样条回归（PBSPLINE） |
| ⑩预测椭圆（ELLIPSE） |
| 连续变量分布图 | ⑪盒形图（HBOX or VBOX） | PLOTNAME response-var / options; |
| ⑫直方图（HISTOGRAM） |
| ⑬密度图（DENSITY） |
| 分类变量分布图 | ⑭条形图（HBAR or VBAR） | PLOTNAME category-var / options; |
| ⑮线形图（HLINE or VLINE） |
| ⑯点状图（DOT） |

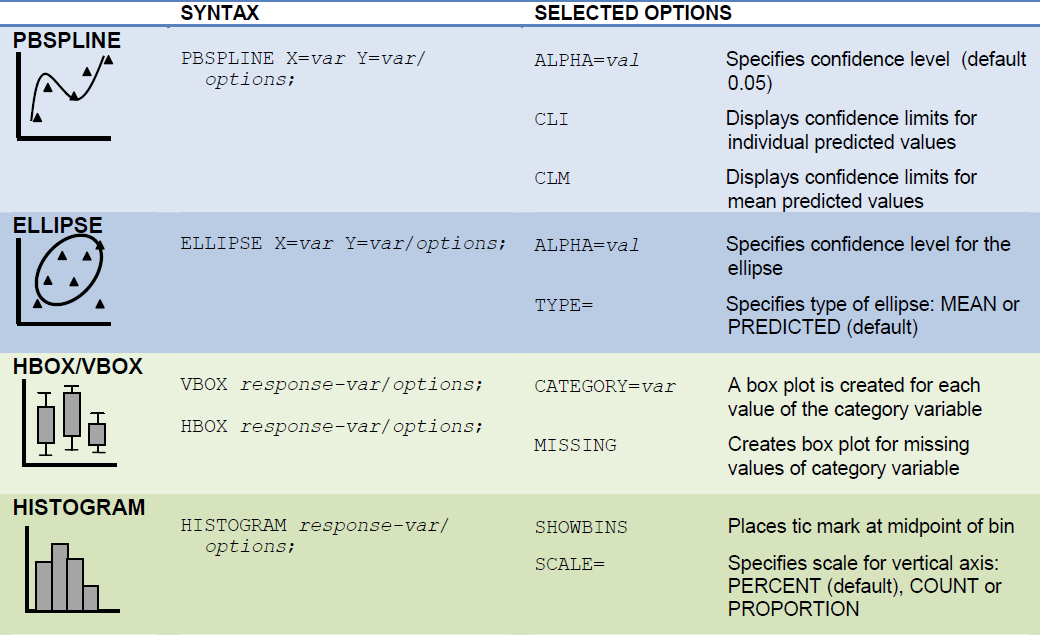
**二、可选参数OPTIONS**

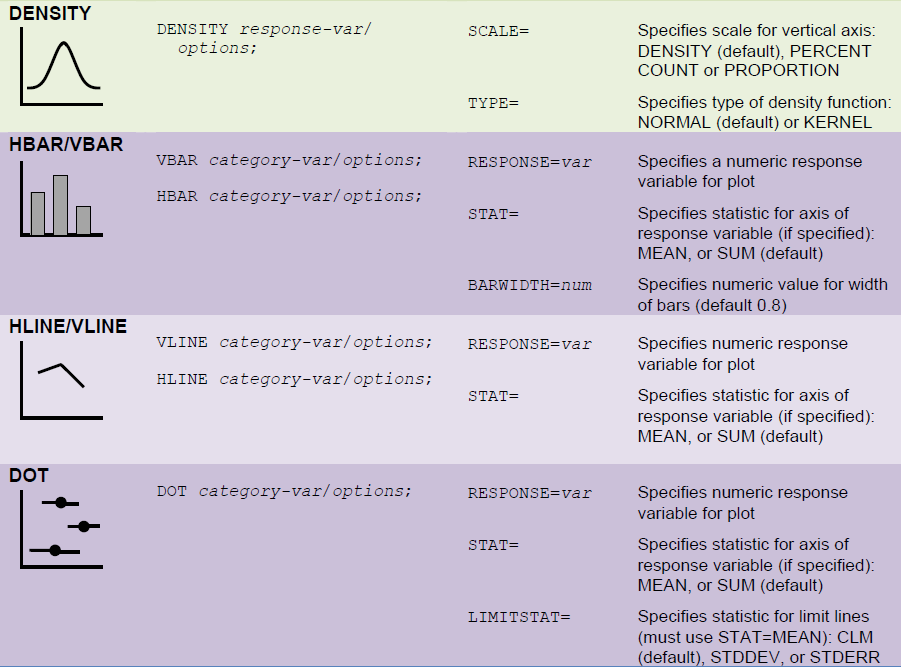
|  |  |  |
| --- | --- | --- |
| **可选参数** | **说明** | **适用图形** |
| RESPONSE=var | 指定数值变量做统计分析 | ⑭-⑯ |
| STAT=统计量 | 对RESPONSE变量做指定的统计分析 | ⑭-⑯ |
| DATALABEL=var | 指定该变量的值作为标签 |  |
| /GROUP=var; | 设置分组变量 | ①-⑨,⑭-⑯ |
| /TRANSPARENCY=value; | 设置透明度：[0,1] | ①-⑥,⑩-⑯ |
| /MARKERS; | 设置标记 | ②-④,⑮ |
| /NOMARKERS; | 设置无标记 | ⑦-⑨ |
| /LEGENDLABEL=’text-str’; | 修改线型说明 | 全部 |
| /FILLATTRS=(attribute=val); | 设置填充属性 | ⑥⑩⑫⑭ |
| /LINEATTRS=(attribute=val); | 设置线型属性 | ②-⑩,⑬,⑮ |
| /MARKERATTRS=(attribute=val); | 设置标记属性 | 1. -④,⑦-⑨,⑮⑯ |

**三、16种图形样式、绘图语句及可选参数**









**四、设置坐标轴（XAXIS/YAXIS）和基准线（REFLINE）**

1. 设置坐标轴基本语法：

PROC SGPLOT;

XAXIS或YAXIS *options*;

选项包括：

（1）GRID

生成网格线；

（2）LABEL ='text-str'

设置坐标轴标记的文字；

（3）TYPE = axistype

设置坐标刻度的类型（字符变量默认为DISCRETE；数值变量默认为LINEAR；日期时间变量默认为TIME；对数尺度默认为LOG；

（4）VALUES =(values-list)

设置刻度标记，例如

VALUES=(**0** **5** **10** **15** **20**); 等同于 VALUES=(**0** TO **20** BY **5**);

2. 设置基准线（水平或竖直线）

给图形增加基准线，是为了便于观察哪些点在某些重要的取值之上（左）还是之下（右）。

基本语法：

PROC SGPLOT;

REFLINE *values* / *options*;

说明：（1）“values”为一些数值，指定画水平线的位置；

（2）可选参数包括：

AXIS = X或Y——指定水平还是竖直，默认Y（水平）；

LABEL ='text-str'——设置基准线标记的文字；

TRANSPARENCY = n——设置透明度：[0,1]，默认0.8；

示例：

REFLINE **32** **75** / LABEL = ('32 degrees' '75 degrees') TRANSPARENCY = **0.5**;

XAXIS TYPE = DISCRETE;

YAXIS LABEL = 'Average High Temperature (F)';

**五、修改“线型说明”（KEYLEGEND）和添加“文字说明”（INSET）**

SGPLOT和SGPANEL过程步都会自动生成线型说明，但有时候需要修改它或加一些自己的注释。

若想去掉“线型说明”，使用如下语句即可：

**PROC** **SGPLOT** DATA = dataset NOAUTOLEGEND;

1. 修改“线型说明”基本语法：

KEYLEGEND <‘name1’ ‘name2’>/ *options*;

注：若对两条线做说明，需要在前面绘图语句中给图形起名’name1’, ‘name2’，然后在KEYLEGEND语句中使用。

可选项包括：

（1）DOWN = n 和 ACROSS = m

指定“线型说明”占n行m列；

（2）LOCATION =INSID 或 OUTSIDE

指定“线型说明”在坐标轴范围内还是外，默认OUTSIDE；

（3）NOBORDER

“线型说明”是在一个小框中，该参数可以去掉小框；

（4）POSITION =value

指定“线型说明”放置的具体位置：TOP, TOPLEFT,

TOPRIGHT, BOTTOM (默认), BOTTOMLEFT,

BOTTOMRIGHT, LEFT, RIGHT.

2. 添加文字说明

基本语法：

INSET '*text-str1*' ... '*text-strn*' / *options*;

注：多个说明语句，将逐个往下排（第1行，第2行…）；

选项包括：

（1）BORDER

为文字说明加上边框；

（2）POSITION =value

指定“线型说明”放置的具体位置同1.（4）.

示例：

KEYLEGEND / LOCATION = INSIDE POSITION = TOPRIGHT;

INSET 'Olympics not held in' '1940 and 1944' / POSITION = BOTTOMLEFT;

**六、定制图形属性**

可以设置线型、颜色等图形属性，让图形更具个性化。SGPLOT过程步提供了许多图形属性，用法：跟在SGPLOT的基本绘图语句后面，加上“/ 图形属性设置语句”即可。例如，

**PROC** **SGPLOT**;

SCATTER X = Score Y = HoursOfStudy / MARKERATTRS = (SYMBOL = STAR);

常用的图形属性有：

（1）FILLATTRS = (COLOR =value)

设置填充属性，填充颜色为某值，

COLOR可选项：AQUA, BLACK, BLUE, FUCHSIA,

GREEN, GRAY, LIME, MAROON, NAVY, OLIVE,

PURPLE, RED, SILVER, TEAL, WHITE, YELLOW;

（2）LINEATTRS = (attribute =value)

设置线型属性，attribute包括：颜色COLOR=,

样式PATTERN=, 线宽THICKNESS=val;

PATTERN可选项：SOLID, DASH, SHORTDASH,

LONGDASH, DOT, DASHDASHDOT, or DASHDOTDOT；

（3）MARKERATTRS = (attribute =value)

设置“标记”的属性，attribute包括：COLOR=, SIZE=val,

符号SYMBOL=;

SYMBOL选项包括：CIRCLE, CIRCLEFILLED,

DIAMOND, DIAMONDFILLED, PLUS, SQUARE,

SQUAREFILLED, STAR, STARFILLED,

TRIANGLE, TRIANGLEFILLED

（4）LABELATTRS = (attribute =value)

设置坐标轴的属性，attribute包括：COLOR=, SIZE=,

STYLE=, WEIGHT=;

SYTLE选项：ITALIC（斜体）, NORMAL（默认）

WEIGHT选项：BOLD（加粗）, NORMAL（默认）

（5）VALUEATTRS = (attribute =value)

设置坐标轴刻度标记的属性，同（4）；

注：THICKNESS和SIZE数值的单位默认是PX（相对长度，1像素），其它可选：CM, IN, MM, PCT, PT。

1in = 2.54cm = 25.4 mm = 72pt = 6pc

**七．设置图片格式和输出路径**

作为论文或PPT中使用的图片，经常需要调整图片的大小、类型、分辨率等。SAS图形可以输出为PDF、RTF、LISTING、HTML.

1. 指定图形的输出样式和风格

基本语法：

ODS *输出样式 options*;

说明：

（1）常用的输出样式有：HTML, LISTING, PDF, RTF;

（2）options可选项：

① FILE或GPATH = '输出路径';

“FILE”设置PDF、RTF格式图片的输出路径；

“GPATH”设置LISTING、HTML格式图片的输出路径；

② DPI或IMAGE\_DPI = n

“DPI”设置PDF格式图片的分辨率，默认n=200；

“IMAGE\_DPI”设置HTML, LISTING, RTF格式图片

的分辨率，默认n=100；

③ STYLE = 风格名

指定一种风格模板（见下表）；

|  |  |  |
| --- | --- | --- |
| 色彩 | 风格模板名 | 默认输出样式 |
| Color（彩色） | ANALYSIS |  |
| HTMLBLUE | HTML |
| LISTING | LISTING (graphs only) |
| PRINTER | PRINTER, PDF, PS |
| RTF | RTF |
| STATISTICAL |  |
| 灰色 | JOURNAL |  |
| 黑白 | JOURNAL2 |  |

2. 设置输出图片的格式

基本语法：

ODS GRAPHICS / *options*;

options可选项：

（1）HEIGHT = h和WIDTH = w

设置图片的高和宽，可选单位CM, IN, MM, PT, PX；

默认大小是640×480像素，也可以只设置高或宽，

另一个会自动按比例跟着变化；

（2）IMAGENAME ='filename'

设置图片的文件名；

（3）OUTPUTFMT = filetype

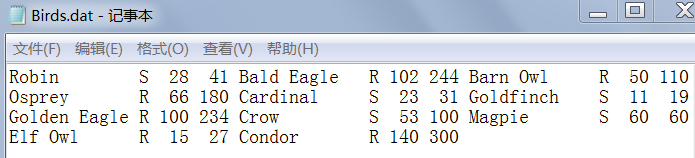
指定图片的输出格式（后缀名）：BMP, GIF, JPEG,

PDF, PNG, PS, SVG, TIFF等；

（4）RESET

恢复默认设置，若没有RESET，将自动给图片依次命名；

**例1** 关于鸟类的数据（C:\MyRawData\Birds.dat），变量包括名称、种类（S-鸣禽，R-猛禽）、身长（cm）、翼展（cm），注意每行有多组数据：



读入数据，绘制散点图，输出样式是LISTING样式，风格是JOURNAL，并设置图片格式类型等。

**代码**：

**data** wings;

infile 'c:\MyRawData\Birds.dat';

input Name $12. Type $ Length Wingspan @@;

**run**;

\* Plot Wingspan by Length;

ODS LISTING GPATH ='c:\MyGraphs' STYLE = JOURNAL;

ODS GRAPHICS / RESET

imagename = 'BirdGraph'

outputfmt = BMP

HEIGHT = **2**IN WIDTH = **3**IN;

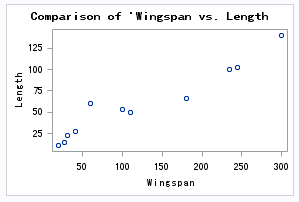
**proc** **sgplot** data = wings;

scatter X =Wingspan Y = Length;

title 'Comparison of ''Wingspan vs. Length';

**run**;

运行结果：



注：图片保存为C:\MyGraphs\BirdGraph1.bmp，事先需要路径C:\MyGraphs存在才行。

**（三）PROC SGPANEL过程步**

SGPLOT与SGPANEL过程步用法基本相同，区别在于：SGPLOT是绘制“单图”；SGPANEL过程步绘制“多子图”。

基本语法：

PROC SGPANEL data = 数据集;

PANELBY *分组变量* / *options*;

*绘图语句*;

说明：

（1）SGPANEL过程步必须与“PANELBY语句”一起使用；

（2）设置坐标轴时用“COLAXIS、ROWAXIS”，不同于SGPLOT用“XAXIS、YAXIS”；

（3）可选项：

① ROWS = n 和 COLUMNS = m

设为放置n行m列（n×m个）子图；

② MISSING

设为包含缺失值；

③ NOVARNAME

删除子图标题的变量名；

④ SPACING = n

设置子图行列间的距离=n个像素，默认n=0；

⑤ UNISCALE =value

设置坐标轴共同的刻度标记；

**例2** 还是 **例1** 关于鸟类的数据用 GSPANEL 过程步绘制翼展（X轴）和身长（Y轴）的散点图，要求按照种类分组。

代码：

**data** wings;

infile 'c:\MyRawData\Birds.dat';

input Name $12. Type $ Length Wingspan @@;

**run**;

\* Plot Wingspan by Length;

**proc** **format**;

value $birdtype

'S' = 'Songbirds'

'R' = 'Raptors';

**run**;

**proc** **sgpanel** data = wings;

panelby Type / NOVARNAME SPACING = **5**;

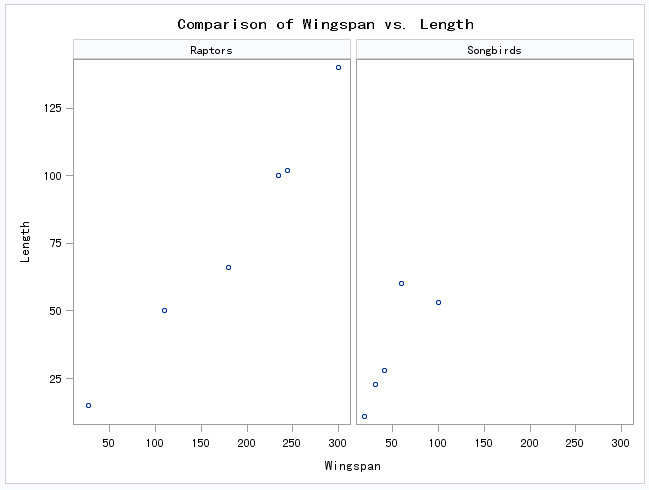
scatter X = Wingspan Y = Length;

format Type $birdtype.;

title 'Comparison of Wingspan vs. Length';

**run**;

运行结果：



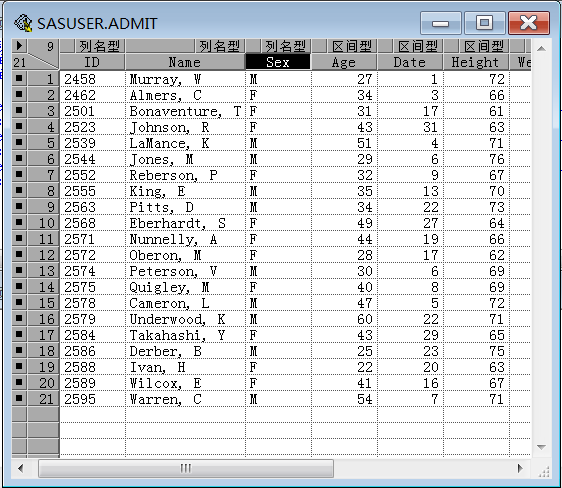
注：不需要事先按“panelby”变量排好序。

**（四）用 INSIGHT 绘制统计图形（窗口操作）**

【解决方案】——【分析】——【交互式数据分析】，

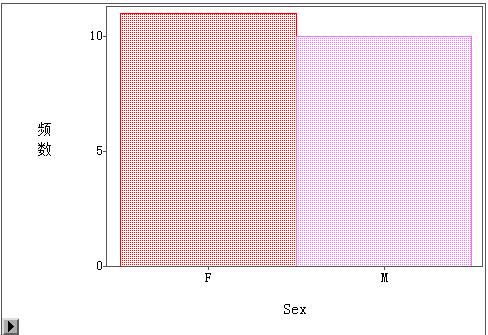


选择要分析的数据集，然后【打开】，



**例3** 绘制关于性别的条形图：

选定要分析的变量，比如Sex，点【分析】——【直方图/条形图】，得到男女频数的条形图：

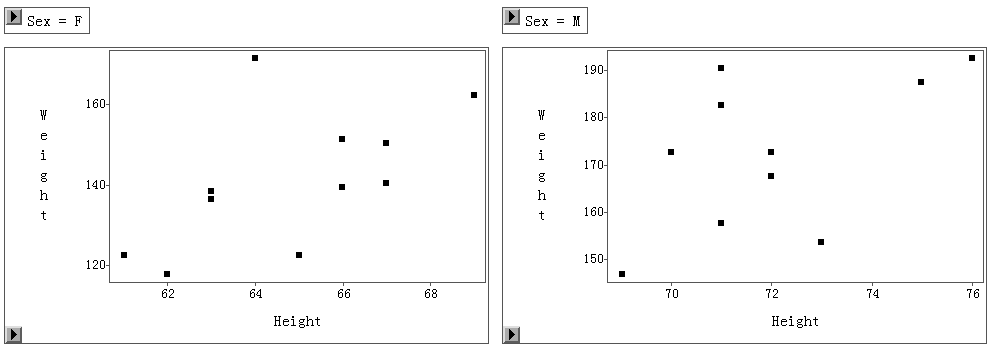


**例4** 绘制体重和身高的散点图，并按男女性别分组考虑。

先不要选定变量，点【分析】——【散点图】，



设定Y变量为 weight，X变量为 Height，分组变量为 Sex，标签变量随意。点【确定】得到



绘制其它图形也是类似的。