Chapter 32
Histograms and Bar Charts

Chapter Table of Contents

Chapter 32
Histograms and Bar Charts

Bar charts are pictorial representations of the distribution of values of a variable.

You can use bar charts to show distributions of interval or nominal variables. Bar charts of interval variables are also called *histograms*.

You can label the heights of the bars in a bar chart, control the orientation, and control the information shown on the axes. For bar charts of interval variables, you can also control the width and offset of the bars.

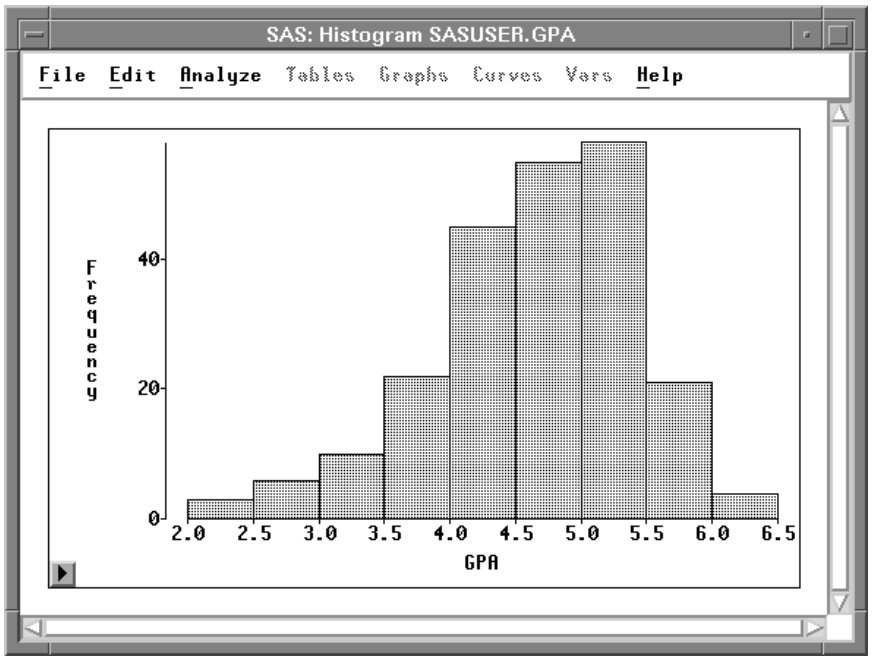

Figure 32.1. Bar Chart

Variables

To create a bar chart, choose **Analyze:Histogram/Bar Chart (Y)**. Bar charts require a **Y** variable. If you have already selected one or more variables, they are assigned the **Y** variable role, and a bar chart is created for each selected variable. If you have not selected any variables, a variables dialog appears.

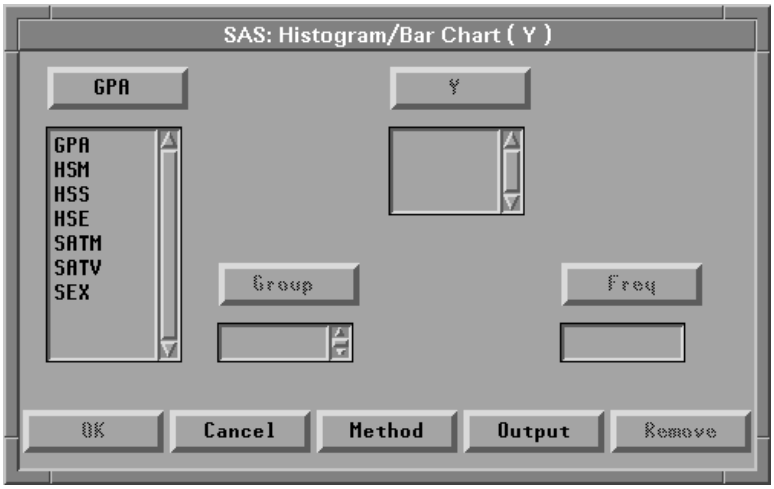

Figure 32.2. Bar Chart Variables Dialog

In the dialog, select at least one **Y** variable. A separate bar chart is created for each **Y** variable you select.

You can select one or more **Group** variables if you have grouped data. This creates one bar chart for each group.

You can select a **Freq** variable. If you select a **Freq** variable, each observation is assumed to represent *n* observations, where *n* is the value of the **Freq** variable.

Method

Observations with missing values for **Y** variables are not used. Observations with **Freq** values that are missing or that are less than or equal to 0 are not used. Only the integer part of **Freq** values is used.

For nominal variables, values that represent less than 4% of the total frequency are grouped together in an "**Other**" category by default. Clicking on the **Method** button in the variables dialog displays the dialog in Figure 32.3. This dialog enables you to change the threshold at which values are grouped into the **Other** category.

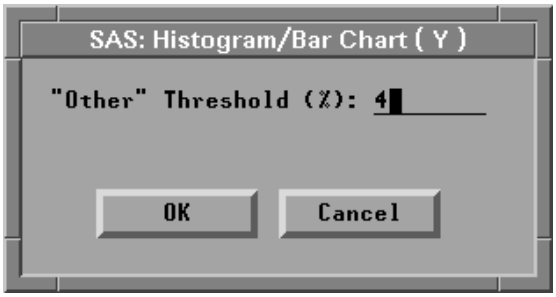

Figure 32.3. Bar Chart Method Options Dialog

For interval variables, values that fall on the boundary between two bars are added to the upper bar. For example, if two bars span ranges (1 to 2) and (2 to 3), the value 2 is considered to fall in the range (2 to 3).

By default, bar width and offset are calculated using an algorithm developed from Terrell and Scott (1985). *Bar width* is the distance along the **Y** axis represented by one bar. *Bar offset* is the distance from the start of the bar to the nearest multiple of the bar width. For example, if a bar starts at 1.2 and has a width of 1, then the offset is 0.2.

Output

For nominal variables, bars are distinguished by different colors. For interval variables, all bars have the same color.

To view or modify output options associated with your bar chart, click on the **Output** button of the variables dialog. This displays the options dialog shown in Figure 32.4.

Figure 32.4. Bar Chart Output Options Dialog

Bar Heights	labels all bars with their heights.
Variable:Names	labels the **Y** axis with variable names.
Variable:Labels	labels the **Y** axis with variable labels.
Variable:Both	labels the **Y** axis with both names and labels.
Orientation: Y Axis Vertical	draws the axis for the Y variable vertically. If this option is turned off, the Y axis is horizontal.
Orientation: Vertical Axis at Left	places the vertical axis at the left side of the chart. If this option is turned off, the vertical axis is at the right side of the chart.
Orientation: Horizontal Axis at Bottom	places the horizontal axis at the bottom of the chart. If this option is turned off, the horizontal axis is at the top of the chart.

You can modify other aspects of the bar chart using the bar chart pop-up menu. Click on the button at the lower left corner of the bar chart to display the pop-up menu.

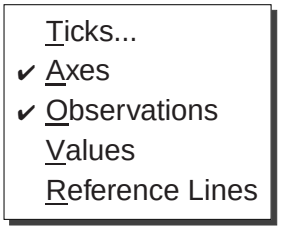

Figure 32.5. Bar Chart Pop-up Menu

Ticks... displays the dialog in Figure 32.6 to set tick values for the variable being charted. In histograms, you can use this menu to set bar width and offset. You can set tick values for the frequency axis by clicking on the **Frequency** label before selecting **Ticks** from the pop-up menu.

Axes toggles the display of axes.

Observations toggles the display of observations (bars). When this menu is toggled off, observations are displayed only if selected.

Values toggles the display of values for bar heights.

Reference Lines toggles the display of lines that indicate the position of major ticks on the frequency axis. This option is not available unless the axes are visible.

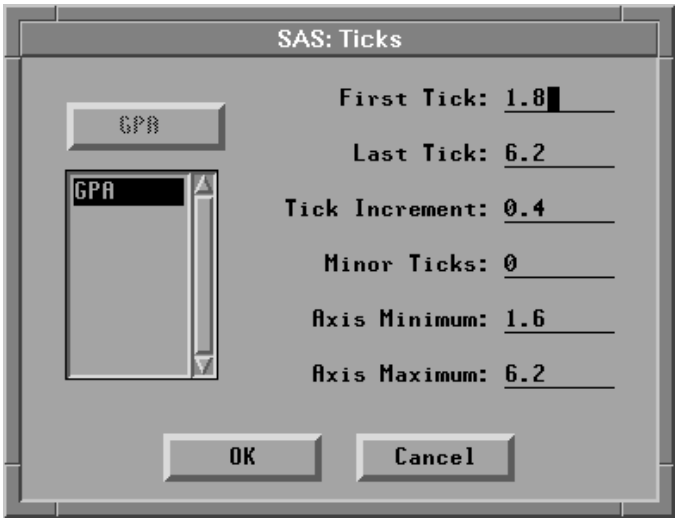

Figure 32.6. Bar Chart Ticks Dialog

You can adjust bar width and offset interactively with the Hand Tool, as described in Chapter 12, “Examining Distributions.” You can also add density curves to bar charts in distribution analyses, as described in Chapter 38, “Distribution Analyses.”

⊕ **Related Reading:** Bar Charts, Chapter 4.

⊕ **Related Reading:** Distributions, Chapter 12, Chapter 38.

References

Terrell, G.R. and Scott, D.W. (1985), "Oversmoothed Nonparametric Density Estimates," *Journal of the American Statistical Association*, 80 (389), 209–214.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc., *SAS/INSIGHT User's Guide, Version 8*, Cary, NC: SAS Institute Inc., 1999. 752 pp.

SAS/INSIGHT User's Guide, Version 8

ISBN 1–58025–490–X

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st printing, October 1999