ABSTRACT

The CHART procedure produces vertical and horizontal bar charts (also called histograms), block charts, pie charts, and star charts. These charts are useful as a visual representation of the values of a single variable or several variables.

INTRODUCTION

You can control the appearance of charts you produce with the CHART procedure by specifying several factors:

* the type of chart
* the summary measures displayed for the variable whose values are charted
* the way values are grouped
* the line-size, page-size, and form-character options you use.
SPECIFICATIONS

The CHART procedure is controlled by the following statements:

PROC CHART <option-list>;
   BY <variable-list>;
   VBAR variable-list </ <standard-option-list> <VBAR-specific-option-list> >;
   HBAR variable-list </ <standard-option-list> <HBAR-specific-option-list> >;
   BLOCK variable-list </ <standard-option-list> <BLOCK-specific-option-list> >;
   PIE variable-list </ <standard-option-list> <NOHEADER> >;
   STAR variable-list </ <standard-option-list> <NOHEADER> >;

Any number of chart-request statements can follow a PROC CHART statement. The options in the standard-option-list can be used on any CHART procedure statement. These options are discussed in detail in Standard and Statement-specific Options later in this chapter. Table 9.1 shows the options that can be specified in standard-option-list.

Table 9.1 Standard Options for the CHART Procedure

<table>
<thead>
<tr>
<th>Function</th>
<th>Option</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group variables</td>
<td>DISCRETE</td>
</tr>
<tr>
<td>Collect summaries</td>
<td>SUMVAR</td>
</tr>
<tr>
<td>Weight an observation</td>
<td>FREQ</td>
</tr>
<tr>
<td>Locate interval midpoints</td>
<td>MIDPOINTS</td>
</tr>
<tr>
<td>Locate chart midpoints</td>
<td>LEVELS</td>
</tr>
<tr>
<td>Choose a measure to compute and display it</td>
<td>TYPE</td>
</tr>
<tr>
<td>Control output</td>
<td>MISSING</td>
</tr>
<tr>
<td></td>
<td>AXIS</td>
</tr>
</tbody>
</table>
Table 9.2 shows the options that are available for specific types of charts.

**Table 9.2 Summary of Statement-specific Options**

<table>
<thead>
<tr>
<th>Options Grouped by Function</th>
<th>HBAR</th>
<th>VBAR</th>
<th>BLOCK</th>
<th>PIE</th>
<th>STAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Separate into groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GROUP=</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUBGROUP=</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G100</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Request statistical analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FREQ</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFREQ</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERCENT</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPERCENT</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUM</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEAN</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control output</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOLEGEND</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOSYMBOL</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYMBOL=</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASCENDING</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DESCENDING</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOZEROS</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REF=</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOSTATS</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOSPACe</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NOHEADER</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These options are listed again with the appropriate chart-request statement, and then each option is described in *Standard and Statement-specific Options* later in this chapter.

**PROC CHART Statement**

**PROC CHART <option-list>;**

The following options can be used in the PROC CHART statement:

- **DATA=**<SAS-data-set>
  - names the SAS data set to be used by PROC CHART. If the DATA= option is not specified, PROC CHART uses the most recently created SAS data set.

- **FORMCHAR<index-list> =‘formchar-string’**
  - defines the characters used to construct the horizontal and vertical axes and the lower left corner, and defines the symbols used to create the bars, sections, or blocks in the PROC CHART output. The default value is a string six characters long composed of the horizontal and vertical lines, the lower left corner, the cross, the slash, and the asterisk. Any character or hexadecimal string can be used to customize table appearance. If you omit the FORMCHAR= option, the value supplied, if any, with the system option FORMCHAR= is used. You can set the
format characters using the SAS system option FORMCHAR= or the PROC CHART statement option FORMCHAR=. You can set all the format characters or only selected ones. See Formatting Your Output Using the FORMCHAR= Option later in this chapter for more information.

LPI=number
specifies the proportions of PIE and STAR charts. The LPI= value is determined by

(lines per inch / columns per inch) * 10

The default is LPI=6. For example, if you have a printer with 8 lines per inch and 12 columns per inch, specify LPI=6.6667.

**BLOCK Statement**

**BLOCK** variable-list </ <standard-option-list> <BLOCK-specific-option-list>>;

Table 9.3 shows the options that can be used in standard-option-list and BLOCK-specific-option-list.

<table>
<thead>
<tr>
<th>Specify Grouping</th>
<th>Control Printing</th>
<th>Standard Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP=</td>
<td>NOLEGEND</td>
<td>AXIS=</td>
</tr>
<tr>
<td>SUBGROUP=</td>
<td>NOSYMBOL</td>
<td>DISCRETE</td>
</tr>
<tr>
<td>G100</td>
<td>SYMBOL=</td>
<td>FREQ=</td>
</tr>
<tr>
<td></td>
<td>NOHEADER</td>
<td>LEVELS=</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MIDPOINTS=</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MISSING</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SUMVAR=</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TYPE=</td>
</tr>
</tbody>
</table>

In the BLOCK statement, list the variables for which you want block charts. Since each block chart must fit on one output page, there are some restrictions on the number of bars of the BLOCK and GROUP= variables. For example, if you use

```plaintext
proc chart;
   block school / group=class;
run;
```

and there are three different values for SCHOOL and six different values (or groups) for CLASS in each school, you need to specify

```plaintext
option pagesize=66 linesize=76;
```
Table 9.4 shows the maximum number of bars of BLOCK variables for selected LINESIZE= (LS=) specifications that can be represented in a block chart using a 66-line page.

**Table 9.4 Maximum Number of Bars of BLOCK Variables**

<table>
<thead>
<tr>
<th>GROUP= Value</th>
<th>LS=132</th>
<th>LS=120</th>
<th>LS=105</th>
<th>LS=90</th>
<th>LS=76</th>
<th>LS=64</th>
</tr>
</thead>
<tbody>
<tr>
<td>0,1</td>
<td>9</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>5,6</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

If the value of any GROUP= level is longer than three characters, the maximum number of BLOCK levels that can be produced may be reduced by one. BLOCK level values are truncated to 12 characters. If these limits are exceeded, data are represented as a horizontal bar chart. See Standard and Statement-specific Options later in this chapter for details on available options.

**BY Statement**

*BY variable-list;*

A BY statement can be used with PROC CHART to obtain separate analyses on observations in groups defined by the BY variables. When a BY statement appears, PROC CHART expects the input data set to be sorted in order of the BY variables or to have an appropriate index. If your input data set is not sorted in ascending order or is not indexed, you can do one of the following:

- Use the SORT procedure with a similar BY statement to sort the data.
- If appropriate, use the BY statement options NOTSORTED or DESCENDING.
- Create an index on the BY variables you want to use. For more information on creating indexes and using the BY statement with indexed data sets, see Chapter 17, "The DATASETS Procedure."
HBAR Statement

HBAR variable-list / <standard-option-list> <HBAR-specific-option-list>;

Table 9.5 shows the options that can be used in standard-option-list and HBAR-specific-option-list.

<table>
<thead>
<tr>
<th>Specify Grouping</th>
<th>Request Statistics</th>
<th>Control Printing</th>
<th>Standard Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP=</td>
<td>FREQ</td>
<td>NOLEGEND</td>
<td>AXIS=</td>
</tr>
<tr>
<td>SUBGROUP=</td>
<td>CFREQ</td>
<td>NOSYMBOL</td>
<td>DISCRETE</td>
</tr>
<tr>
<td>G100</td>
<td>PERCENT</td>
<td>SYMBOL=</td>
<td>FREQ=</td>
</tr>
<tr>
<td></td>
<td>CPERCENT</td>
<td>ASCENDING</td>
<td>LEVELS=</td>
</tr>
<tr>
<td></td>
<td>SUM</td>
<td>DESCENDING</td>
<td>MIDPOINTS=</td>
</tr>
<tr>
<td></td>
<td>MEAN</td>
<td>NOZEROS</td>
<td>MISSING</td>
</tr>
<tr>
<td></td>
<td>NOSTATS</td>
<td>REF=</td>
<td>SUMVAR=</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TYPE=</td>
</tr>
</tbody>
</table>

The HBAR statement requests a horizontal bar chart for each variable listed. For example, the following statements produce three horizontal bar charts:

    proc chart;
      hbar a x1 x2;
    run;

Each chart occupies one or more output pages, depending on the number of bars; each bar occupies one line.

By default, for horizontal bar charts of TYPE=FREQ, CFREQ, PCT, or CPCT, the CHART procedure prints the following statistics: frequency, cumulative frequency, percentage, and cumulative percentage. If you use one or more of the statistics options, PROC CHART prints only the statistics you request. For example, using TYPE=MEAN, statistics include only frequency and mean. For TYPE=SUM, statistics printed are frequency and sum. See Standard and Statement-specific Options later in this chapter for details.

PIE Statement

PIE variable-list / <standard-option-list> <NOHEADER>;

The following options can be used in standard-option-list:

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AXIS=</td>
<td></td>
</tr>
<tr>
<td>DISCRETE</td>
<td></td>
</tr>
<tr>
<td>FREQ=</td>
<td></td>
</tr>
<tr>
<td>LEVELS=</td>
<td></td>
</tr>
<tr>
<td>MIDPOINTS</td>
<td>=</td>
</tr>
<tr>
<td>MISSING</td>
<td></td>
</tr>
<tr>
<td>SUMVAR=</td>
<td></td>
</tr>
<tr>
<td>TYPE=</td>
<td></td>
</tr>
</tbody>
</table>
The PIE statement requests a pie chart for each variable listed. For example, the following statements produce three one-page pie charts:

```plaintext
proc chart;
   pie a x1 x2;
run;
```

PROC CHART determines the number of slices for the pie in the same way that it determines the number of bars for vertical bar charts. Any slices of the pie accounting for less than three print positions are grouped together into a slice called OTHER.

The pie’s size is determined only by the LINESIZE= and PAGESIZE= system options. By default, the pie looks elliptical if your printer does not print 6 lines per inch and 10 columns per inch. To make a circular pie chart on a printer that does not print 6 lines and 10 columns per inch, use the LPI= option on the PROC CHART statement. For example, if your printer prints 8 lines per inch and 10 columns per inch, specify LPI=8 in the PROC CHART statement.

If a PIE chart is requested for a variable with over 50 levels, a horizontal bar chart is produced instead. See Standard and Statement-specific Options later in this chapter for descriptions of each option.

**STAR Statement**

```plaintext
STAR variable-list <; <standard-option-list> <NOHEADER>>;
```

The following options can be specified in standard-option-list:

- `AXIS=`
- `MIDPOINTS=`
- `DISCRETE` = `MISSING`
- `FREQ=` = `SUMVAR=`
- `LEVELS=` = `TYPE=`

The STAR statement requests a star chart for each variable listed. For example, the following statements produce a one-page star chart for the variable Z:

```plaintext
proc chart;
   star z;
```

The number of points in the star is determined in the same way as the number of bars for vertical bar charts.

If all the data to be charted with a STAR statement are positive, the center of the star represents zero and the outside circle represents the maximum value. If negative values occur in the data, the center represents the minimum. See the AXIS= option for more information about how to specify maximum and minimum values. If a star chart is requested for a variable with over 24 levels, a horizontal bar chart is produced instead. See Standard and Statement-specific Options later in this chapter for descriptions of each option.

Note: If you want different variables to form the rays of the star, use an OUTPUT statement in a DATA step to create new observations having one variable with values equal to the variables you want represented by the rays; in other words, create another variable whose values are the original variable names.
The CHART Procedure

VBAR Statement

VBAR variable-list / <standard-option-list> <VBAR-specific-option-list>;

Table 9.6 shows the options that can be specified in the VBAR statement.

Table 9.6 Options Available in the VBAR Statement

<table>
<thead>
<tr>
<th>Specify Grouping</th>
<th>Control Printing</th>
<th>Standard Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>GROUP=</td>
<td>NOLEGEND</td>
<td>AXIS=</td>
</tr>
<tr>
<td>SUBGROUP=</td>
<td>NOSYMBOL</td>
<td>DISCRETE</td>
</tr>
<tr>
<td>G100</td>
<td>NOSPACE</td>
<td>FREQ=</td>
</tr>
<tr>
<td></td>
<td>SYMBOL=</td>
<td>LEVELS=</td>
</tr>
<tr>
<td></td>
<td>ASCENDING</td>
<td>MIDPOINTS=</td>
</tr>
<tr>
<td></td>
<td>DESCENDING</td>
<td>MISSING</td>
</tr>
<tr>
<td></td>
<td>NOZEROS</td>
<td>SUMVAR=</td>
</tr>
<tr>
<td></td>
<td>REF=</td>
<td>TYPE=</td>
</tr>
</tbody>
</table>

In the VBAR statement, list the variables for which you want vertical bar charts. For example, if you list STUDENTS as a chart variable, then a vertical bar chart is produced with the values of STUDENTS underneath the bars.

The procedure prints one page per chart. Along the vertical axis, PROC CHART describes the chart frequency, the cumulative frequency, the chart percentage, the cumulative percentage, the sum, or the mean. At the bottom of each bar, PROC CHART prints a value according to the value of the TYPE= option, if specified. For character variables or discrete numeric variables, this value is the actual value represented by the bar. For continuous numeric variables, the value gives the midpoint of the interval represented by the bar.

PROC CHART can automatically scale the vertical axis, determine the bar width, and choose spacing between the bars. However, options allow you to choose bar intervals and the number of bars, include missing values in the chart, produce side-by-side charts, and subdivide the bars. If the number of characters per line (LINESIZE=) is not sufficient to display all vertical bars, PROC CHART will produce a horizontal bar chart instead.

Standard and Statement-specific Options

Each of the following options is either a standard option or an option that can be used only in some chart-request statements. The options are presented in alphabetic order.

You can use an option with any of the chart-request statements unless otherwise specified. You need to specify these options following a slash (/), for example,

```plaintext
proc chart;
    hbar year / <standard-option-list> <HBAR-specific-list>;
run;
```

ASCENDING | ASC
prints the bars and any associated statistics in ascending order of size within groups. The ASCENDING option can only be used with the HBAR and VBAR statements.
\textbf{AXIS= \texttt{<min-value > max-value}}

allows you to specify the minimum and maximum values used in constructing the FREQ, PCT, CFREQ, CPCT, SUM, or MEAN axis. If you use the VBAR or HBAR statements and TYPE=SUM or TYPE=MEAN, and if any of the sums or means are less than zero, then a negative minimum value can also be specified in the AXIS= option. Otherwise, the default is a minimum value of zero. Counts or percentages outside the maximum (or minimum) override the AXIS= specification. If the AXIS= option is specified and a BY statement also appears, uniform axes are produced over BY groups. When the AXIS= option appears in a STAR statement, the first value specified is the center (minimum) of the star and the second value is the outside circle (maximum). If only one AXIS= value is specified in a STAR statement, PROC CHART assumes this value is the maximum and zero is the minimum. For example, the following statements produce a star chart for the sums of \textit{X} classified by \textit{A} and scaled from 100 at the center to 200 at the outside circle:

\begin{verbatim}
   proc chart;
   star a / sumvar=x type=sum axis=100 200;
   run;
\end{verbatim}

\textbf{CFREQ}

prints the cumulative frequency. The CFREQ option can only be used with the HBAR statement.

\textbf{CPERCENT}

prints the cumulative percentages. The CPERCENT option can only be used with the HBAR statement.

\textbf{DESCENDING | DESC}

prints the bars and any associated statistics in descending order of size within groups. The DESCENDING option can only be used with the HBAR and VBAR statements.

\textbf{DISCRETE}

is used when the numeric chart variable specified is discrete rather than continuous. If the DISCRETE option is omitted, PROC CHART assumes that all numeric variables are continuous and automatically chooses intervals for them unless the MIDPOINTS= or LEVELS= options are used.

\textbf{FREQ}

prints the frequency of each bar to the side of the chart. The FREQ option can only be used with the HBAR statement.

\textbf{FREQ=variable}

is used when a variable in the data set represents a count (or weight) for each observation. Normally, each observation contributes a value of one to the frequency counts. When the FREQ= option appears, each observation contributes the FREQ= value. If the FREQ= values are not integers, they are truncated to integers. If the values are missing or negative, the contribution is zero. If the SUMVAR= option is specified, the sums are multiplied by the FREQ= value.

\textbf{GROUP=variable}

produces side-by-side charts, with each chart representing the observations having a given value of the GROUP= variable. The GROUP= variable can be character or numeric and is assumed to be discrete. For example, the following statement produces a frequency bar chart for men and women in each department:

\begin{verbatim}
   vbar sex / group=dept;
\end{verbatim}
Missing values for a GROUP= variable are treated as valid levels when a chart is produced. The GROUP= option can only be used with the HBAR, VBAR, and BLOCK statements.

G100

is used in conjunction with the GROUP= option to force the bars and statistics to add to 100% for each group. The G100 option can only be used with the HBAR, VBAR, and BLOCK statements.

LEVELS=number-of-midpoints
specifies the number of bars, blocks, or sections representing each chart variable when the variables given in the VBAR statement are continuous.

MEAN
prints the mean of the observations represented by each bar. The MEAN option can only be used with the HBAR statement.

MIDPOINTS=midpoint-list | OLD
defines the range of values each bar or section represents by specifying the range midpoints. For example, the following statement produces a chart with five bars: the first bar represents the range of data values with a midpoint of 10; the second bar represents the range of data values with a midpoint of 20; and so on.

vbar x / midpoints=10 20 30 40 50;

When the variables given in the VBAR statement are numeric, the midpoints must be given in ascending order, although they need not be uniformly distributed. For example, the following statement produces a chart of X with logarithmic intervals:

vbar x / midpoints=10 100 1000 10000;

A numeric midpoint-list of the form

midpoints=10 to 100 by 5

is also acceptable. For character variables, the MIDPOINTS= option can be specified in any order, which is useful in ordering the bars or in specifying a subset of the possible values. For example, you can give a list of the form

midpoints='JAN' 'FEB' 'MAR'

Without the MIDPOINTS= option, the values are displayed in sorted order.

If you don't use the MIDPOINTS= option to specify midpoints, PROC CHART uses its own internal algorithm to choose midpoints for continuous variables. In previous versions of the SAS System, this algorithm was based on the work of Nelder (1976); now an improved algorithm has been implemented based on the work of Terrell and Scott (1985). If you prefer the old algorithm, you can use it by specifying MIDPOINTS=OLD.

MISSING
specifies that missing values are to be considered as valid levels for the chart variable.

NOHEADER
NOHEADER
suppresses the default header line normally printed at the top of a chart. The NOHEADER option can only be used with the BLOCK, PIE, and STAR statements.
NOSPACE
specifies that if the line size does not allow room for spaces between the bars, PROC CHART can print a vertical bar chart without spaces between bars. If space is still insufficient, a horizontal bar chart is printed instead. This option is used only in the VBAR statement.

NOSTATS
NOSTAT
suppresses printing statistics on a horizontal bar chart. The NOSTATS option can only be used with the HBAR statement.

NOSYMBOL
NOLEgend
is used in conjunction with the SUBGROUP= option to suppress printing of the subgroup legend or symbol table. Both the NOSYMBOL and NOLEgend options can only be used with the HBAR, VBAR, and BLOCK statements.

NOZEROS
suppresses any bar with zero frequency. The NOZEROS option can only be used with the HBAR and VBAR statements.

PERCENT
prints the percentages of observations having a given value for the chart variable. The PERCENT option can only be used with the HBAR statement.

REF=value
draws a single reference line on the response axis. For TYPE=FREQ or TYPE=CFREQ, the REF= option should be a frequency; for TYPE=PCT or TYPE=CPT, the REF= option should be a percent between 1 and 100. For TYPE=SUM or TYPE=MEAN, the REF= option should be a sum or mean. The REF= option can only be used with the HBAR and VBAR statements.

SUBGROUP=variable
subdivides each bar into characters that show the SUBGROUP= variable’s contribution to the bar. For example, the following statement produces a chart with one bar for each department:

vbar dept / subgroup=sex;

The portion of each bar filled in with the character * represents those observations that have a SEX value of *.

The first character of the value is used to fill in the portion of the bar corresponding to the value unless more than one value begins with the same first character. In that case, the letters A, B, C, and so on are used. If the variable is formatted, PROC CHART uses the first character of the formatted value. The subgroup symbols are ordered A through Z and 0 through 9 with the characters in ascending order. You will notice, for example, that Output 9.7 shown earlier in this chapter has the value * at the lower part of the bar chart and * on the upper part. The characters used in the chart and the values they represent are given in a legend at the bottom of the chart.

Missing values for a SUBGROUP= variable are treated as valid levels when a chart is produced.

PROC CHART calculates the height of the bar for each subgroup individually and then rounds each bar’s percentage of the total bar up or
down. Thus, the total height of the bar may be higher or lower than the same bar without the SUBGROUP= option.

If you use both TYPE=MEAN and the SUBGROUP= option, PROC CHART first calculates the mean for each variable listed in the SUMVAR= option, then subdivides the bar into the percentages contributed by each subgroup.

SUM
prints the total number of observations that each bar represents.

SUMVAR=variable
names the variable to collect summaries for means, sums, or frequencies. The SUMVAR= option is useful for producing bar charts showing total expenditures for each department or showing means at each level of an experiment. For example, the following statement produces a chart showing the mean yield for each location:

    vbar location / type=mean sumvar=yield;

The next example charts total expenditures by department:

    vbar dept / sumvar=expend;

If the SUMVAR= option is specified but the TYPE= value is not MEAN or SUM, then TYPE=SUM overrides whatever TYPE= value is specified.

SYMBOL='character-list'
is used, when the SUBGROUP= option is not used, to define the symbol or symbols in the bars or blocks of the chart. The default SYMBOL= value is the asterisk (*). If the SAS system option OVP is set and your printing device supports overprinting, you can specify up to three characters in character-list to produce overprinted charts. For example, the following statements produce a chart of very thick, dark horizontal bars:

    options ovp;
    proc chart;
    hbar dept / symbol='XOA';
    run;

The SYMBOL= option can only be used with the HBAR, VBAR, and BLOCK statements.

TYPE=type
specifies what the bars or sections in the chart represent. If the TYPE= option is omitted, the default TYPE is FREQ. When the SUMVAR= option is specified, the default TYPE is SUM. You can specify one of the following keywords for type:

    CFREQ makes each bar or section represent cumulative frequency, which is the frequency for the group plus all the frequencies that precede it.

    CPERCENT | CPCT makes each bar represent the cumulative percentage of observations of the chart variable, which is the percentage of the group plus the percentages of all the groups that precede it.

    FREQ makes each bar or section represent the frequency with which a value or range occurs for the chart variable in the data.
MEAN makes each bar or section represent the mean of the SUMVAR= variable for observations having the bar’s value. For example, the following statement produces a chart showing the mean sales for each department:

```sas
vbar dept / sumvar=sales type=mean;
```

PERCENT | PCT makes each bar or section represent the percentage of observations of the chart variable having a given value or falling into a given range.

SUM makes each bar or section represent the sum of the SUMVAR= variable for observations having the bar’s value. For example, the following statement produces a chart with one bar or section for each DEPT value:

```sas
vbar dept / sumvar=sales type=sum;
```

The bar height for a given DEPT corresponds to the total of the SALES values for observations having that DEPT value.

**DETAILS**

**Missing Values**

If you use the MISSING option, missing values are not considered as valid levels for the chart variable. The MISSING option is available in all chart-request statements.

Missing values for a GROUP= or SUBGROUP= variable are treated as valid levels when a chart is produced.

**Formatting Your Output Using the FORMCHAR= Option**

The SAS system option FORMCHAR= and the FORMCHAR= option in the PROC CHART statement both define the characters used to print charts. If you set the FORMCHAR= option in the PROC CHART statement, it affects only PROC CHART output; if you set the SAS system option FORMCHAR=, it affects the output of all procedures that use the FORMCHAR= option.

PROC CHART uses 6 of the 20 possible form characters. Table 9.7 shows all format characters available in PROC CHART and gives appropriate definitions.