Chapter 24
# Formatting Variables and Values

Chapter Table of Contents

350

Chapter 24
Formatting Variables and Values

Formats determine how variables and values are displayed. In group variables and model effects, formats can also determine how values are used in calculations.

You can use formats to set the width of displayed values, the number of decimal points displayed, the handling of blanks, zeroes, and commas, and other details. The SAS System provides many standard formats for displaying character and numeric values.

In addition, you can use the FORMAT procedure to create your own formats.
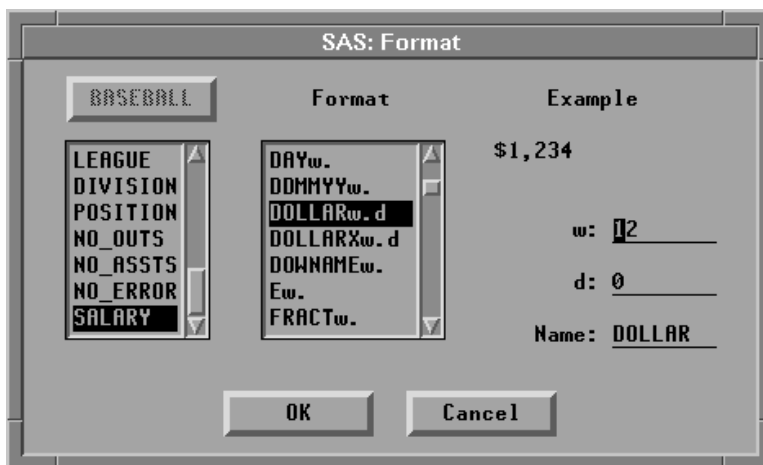
Figure 24.1. Assigning a Format

Assigning Formats

By default, SAS/INSIGHT software displays each variable using the format supplied in your SAS data set. If your data set contains numeric variables with no formats, SAS/INSIGHT software chooses a format based on that variable's values. When you save the data, formats chosen by SAS/INSIGHT software are not automatically saved, but any formats you assign are saved.

You can assign formats by using the **Edit:Formats** menu.

⟹ **Open the BASEBALL data set.**
This data set contains statistics and salaries of major league baseball players.

⟹ **Select the variable SALARY.**

Figure 24.2. **SALARY** Selected
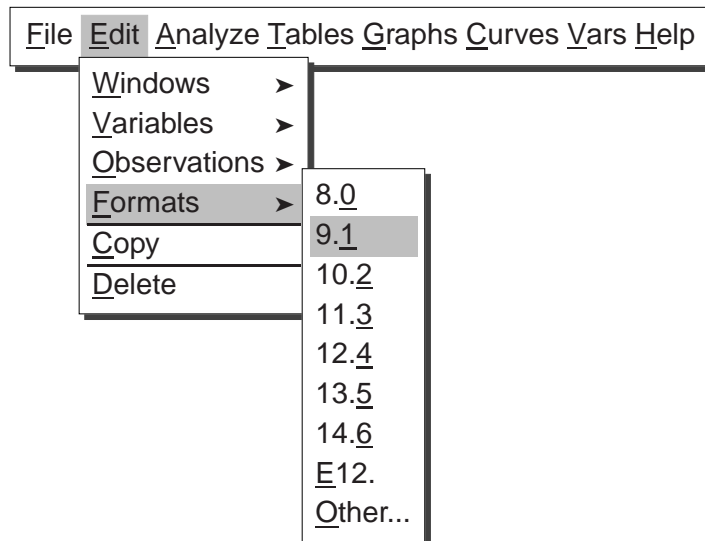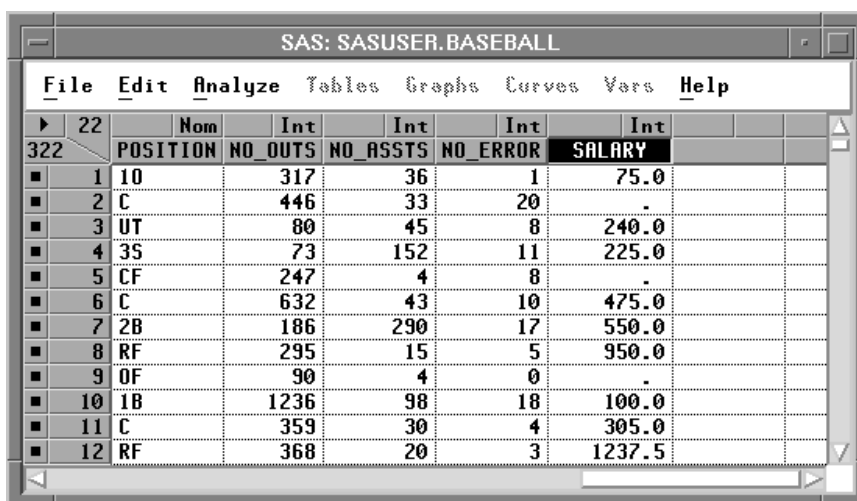
⟹ **Choose Edit:Formats:9.1.**

Figure 24.3. Edit:Formats Menu

352

This gives the variable **SALARY** a width of nine character positions, including the decimal and one position after the decimal. The actual data values for **SALARY** continue to be stored with double precision.
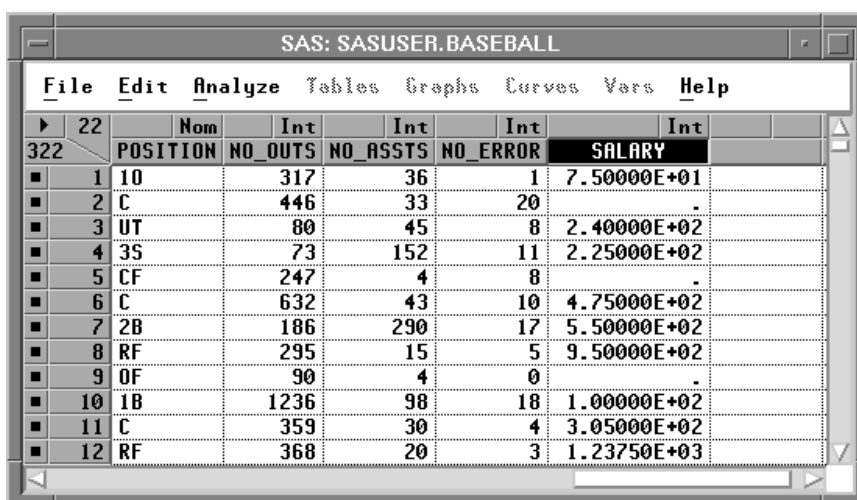
22		Nom	Int	Int	Int	Int		
322		POSITION	NO_OUTS	NO_ASSTS	NO_ERROR	SALARY		
■	1	10	317	36	1	75.0		
■	2	C	446	33	20	.		
■	3	UT	80	45	8	240.0		
■	4	3S	73	152	11	225.0		
■	5	CF	247	4	8	.		
■	6	C	632	43	10	475.0		
■	7	2B	186	290	17	550.0		
■	8	RF	295	15	5	950.0		
■	9	OF	90	4	0	.		
■	10	1B	1236	98	18	100.0		
■	11	C	359	30	4	305.0		
■	12	RF	368	20	3	1237.5		

Figure 24.4. Format 9.1

⟹ **Choose Edit:Formats:E12.**

This gives the variable **SALARY** a width of 12 character positions and expresses each value in exponential notation.

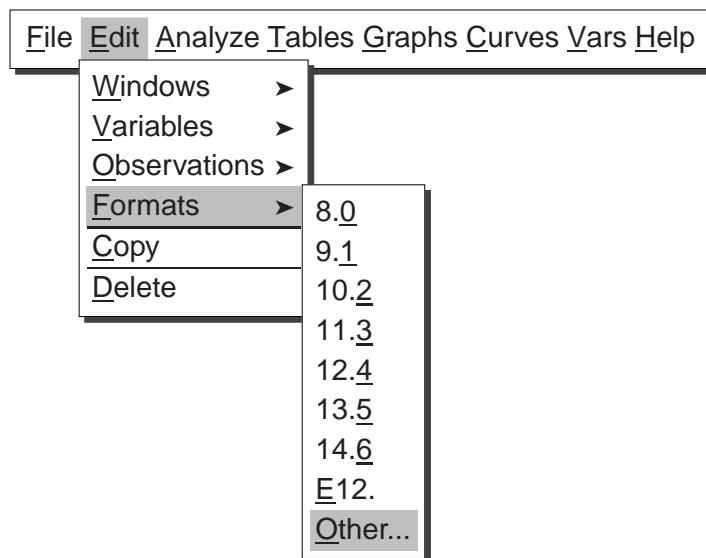| 22		Nom	Int	Int	Int	Int		
322		POSITION	NO_OUTS	NO_ASSTS	NO_ERROR	SALARY		
■	1	10	317	36	1	7.50000E+01		
■	2	C	446	33	20	.		
■	3	UT	80	45	8	2.40000E+02		
■	4	3S	73	152	11	2.25000E+02		
■	5	CF	247	4	8	.		
■	6	C	632	43	10	4.75000E+02		
■	7	2B	186	290	17	5.50000E+02		
■	8	RF	295	15	5	9.50000E+02		
■	9	OF	90	4	0	.		
■	10	1B	1236	98	18	1.00000E+02		
■	11	C	359	30	4	3.05000E+02		
■	12	RF	368	20	3	1.23750E+03		

Figure 24.5. Format E12.

The **Edit:Formats** menu provides quick access to frequently used formats. There are many other standard formats provided by the SAS System.

⟹ **Choose Edit:Formats:Other.**

353

| File | Edit | Analyze | Tables | Graphs | Curves | Vars | Help |

| Windows | ➤ |
Variables	➤
Observations	➤
Formats	➤	8.0
Copy		9.1
Delete		10.2
		11.3
		12.4
		13.5
		14.6
		E12.
		Other...

Figure 24.6.   Edit:Formats Menu

This displays the Format dialog. In the dialog, the fields **w** and **d** specify the width and decimal places to be used by the formats. Note that the **SALARY** variable and the **E12.** format are currently selected.
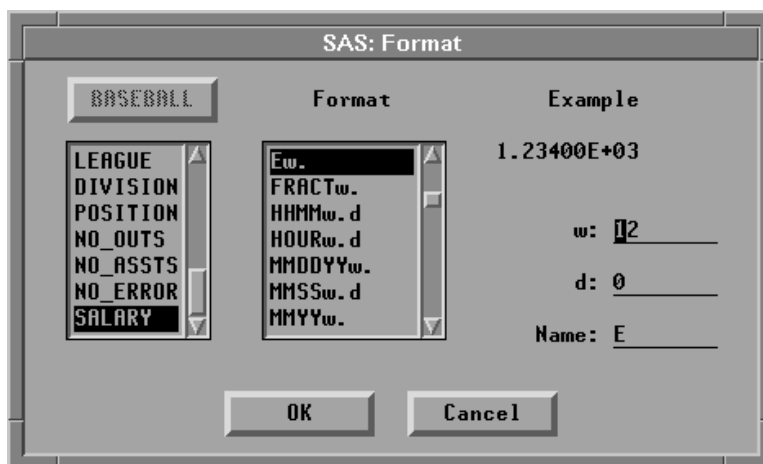
Figure 24.7.   Format Dialog

⟹ **Select DOLLARw.d in the Format list.**

Formats are listed alphabetically, so the **DOLLARw.d** format is above the **Ew.** format.
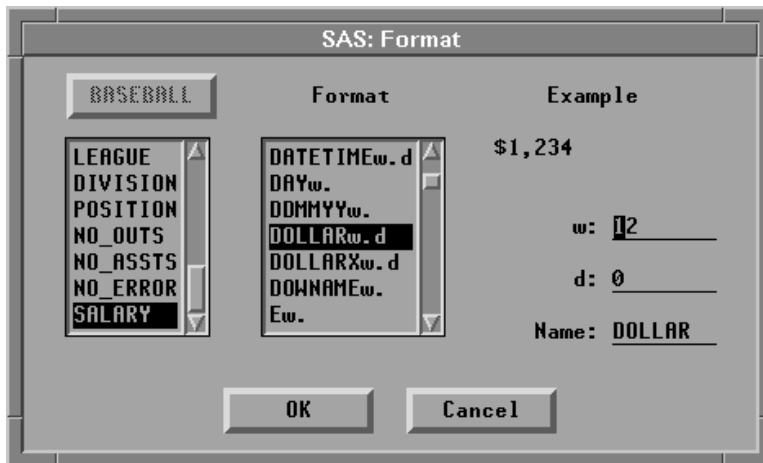
Figure 24.8. Format **DOLLARw.d**

The example in the upper right corner of the dialog illustrates the format you have se-
lected. **DOLLAR** is the standard format for display of currency in the United States.
There is also a **DOLLARX** format sometimes preferred in European countries.

⟹ **Click OK to set the format you prefer.**

Figure 24.9. **SALARY** Formatted

Now the variable **SALARY** uses the format you assigned. By default, this format is
also used for axes in subsequent analyses. You can modify the axes, however, to use
other formats.

⟹ **Choose Analyze:Distribution (Y).**
This creates a distribution analysis of **SALARY**. The box plot and histogram axes
use the format you assigned to the **SALARY** variable in the data window.

⟹ **Select SALARY in the distribution window.**
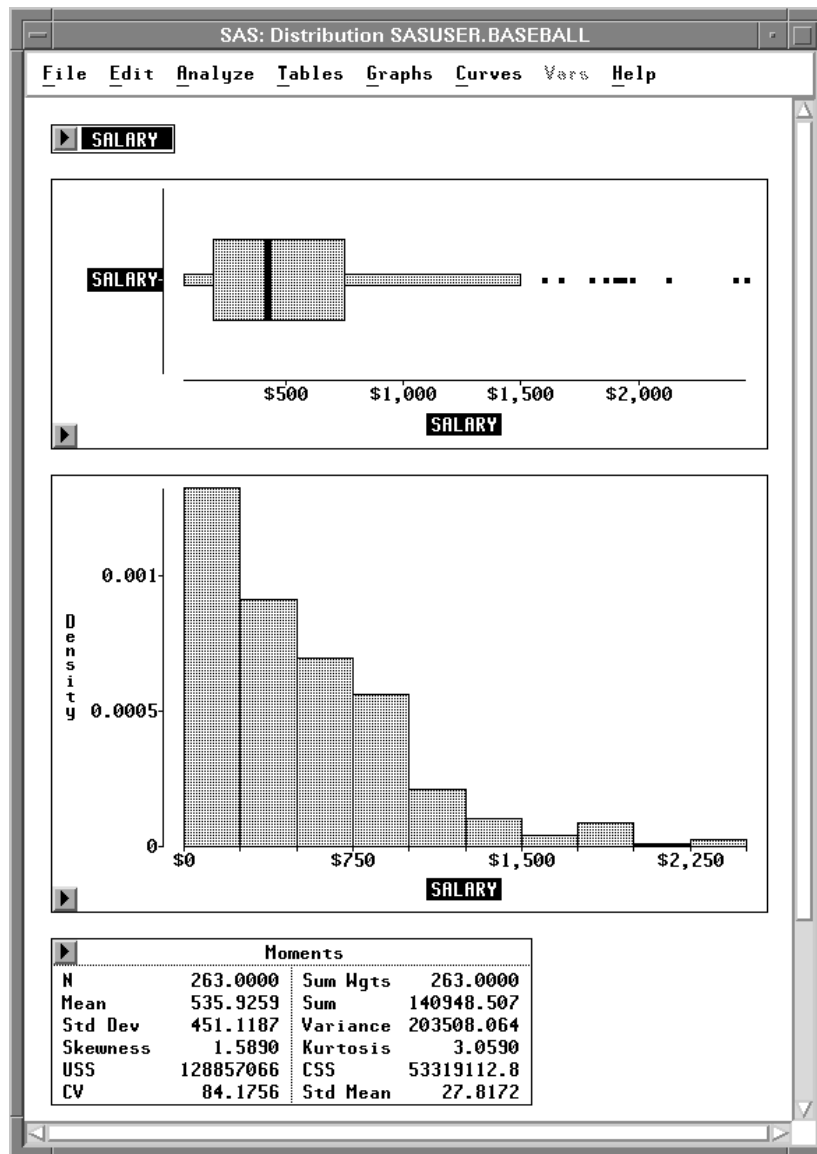
Figure 24.10. Distribution Analysis, **SALARY** Selected

⟹ **Choose Edit:Formats:8.0.**
This assigns the **8.0** format to **SALARY** on axes in the distribution window. In the data window, **SALARY** continues to use the **DOLLAR** format.

356
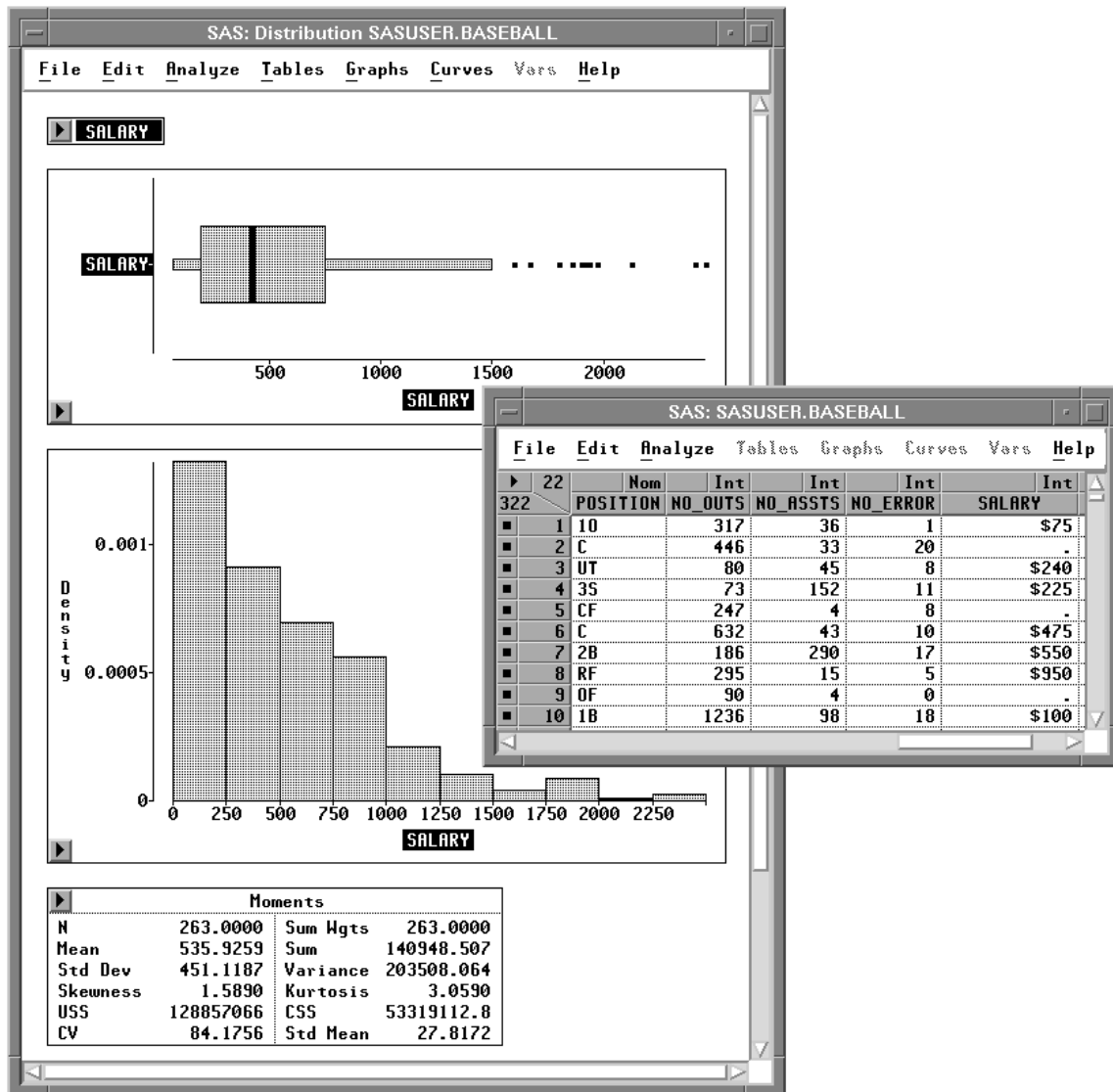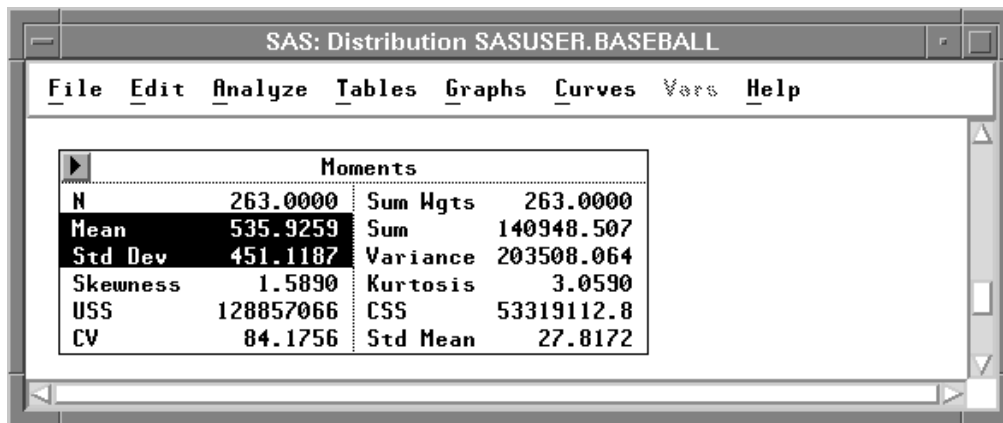
Figure 24.11. **SALARY** Axes Formatted

You can also format individual values in analysis tables. For example, suppose you need to see greater precision for the mean and standard deviation.

⟹ **Select the values for Mean and Std Dev in the Moments table.**

357

Figure 24.12. Moments Table, Values Selected

⟹ **Choose Edit:Formats:14.6.**
Now the mean and standard deviation show six digits after the decimal.

Figure 24.13. Moments Table, Values Formatted

358

Creating Formats

Although there are many formats available in the SAS System, occasionally you will want to create your own. To do this, use the FORMAT procedure.

For example, suppose you want to consider certain groupings of baseball players based on the length of their careers. You can combine the values of **YR⏤MAJOR** into four groups, as follows.

⟹ **Enter PROC FORMAT statements in the Program Editor.**
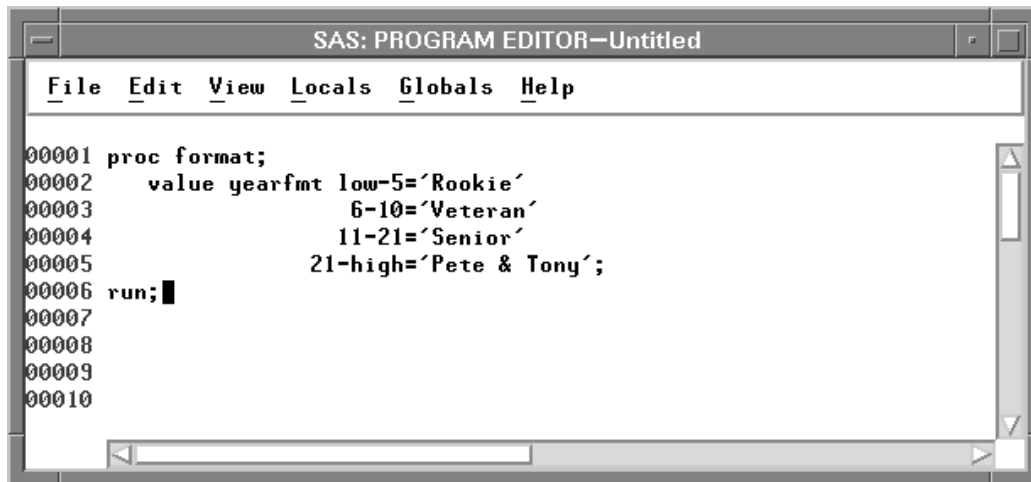
Figure 24.14. Program Editor

⟹ **Choose Run:Submit.**
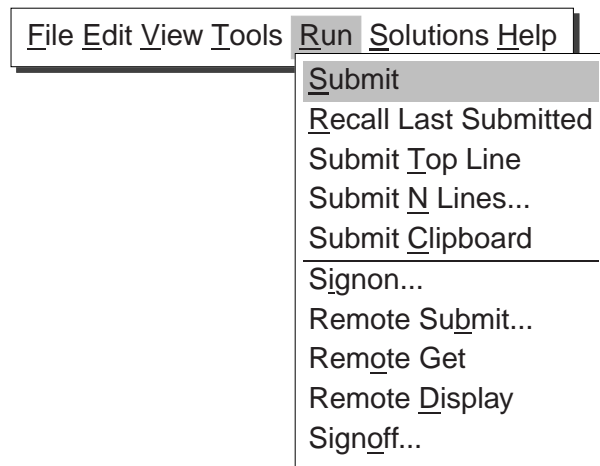
Figure 24.15. Run Menu

⟹ **Select YR⏤MAJOR.**

Figure 24.16. **YR_MAJOR** Selected

$\Longrightarrow$ **Choose Edit:Formats:Other.**
This displays the Format Dialog.

Figure 24.17. Format Dialog

$\Longrightarrow$ **Enter YEARFMT in the Name field.**

$\Longrightarrow$ **Enter 12 in the w field, then press the Return key.**
Now the example in the upper right of the dialog shows a value formatted with
YEARFMT.

Figure 24.18.   YEARFMT Entered

⟹ **Click OK to close the Format dialog.**
Now **YEARFMT** is used to display the values of **YR_MAJOR**.

Figure 24.19.   YEARFMT Assigned

By default, the new format is used to display values wherever you use **YR_MAJOR**.
Formats are not used in calculations except for nominal variables in model effects or
for group variables. In these cases, the format is used to determine the groups. You
can see this use of formats by creating a box plot.

⟹ **Deselect YR_MAJOR in the data window.**

⟹ **Choose Analyze:Box Plot/Mosaic Plot (Y).**
This displays the box plot variables dialog.
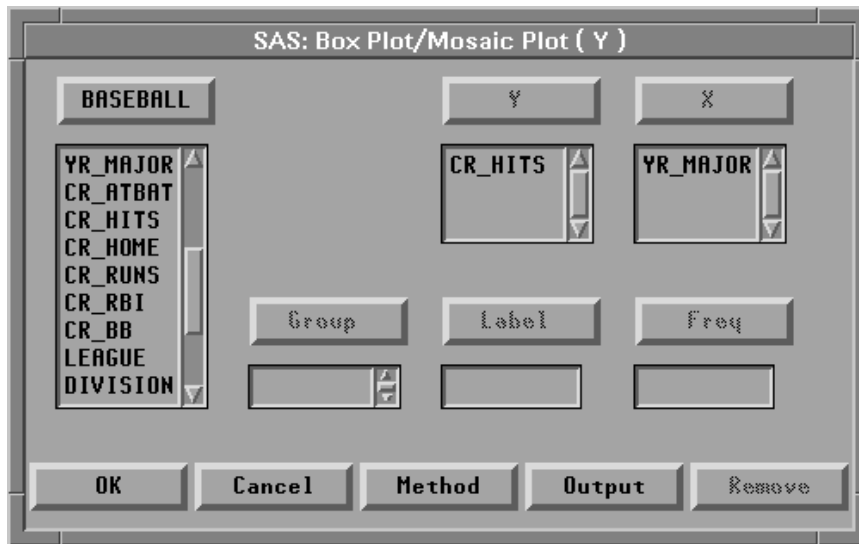
⟹ **Assign YR_MAJOR the X role and CR_HITS the Y role.**

Figure 24.20. Box Plot Variables Dialog

$\Longrightarrow$ **Click the OK button to create the box plot.**
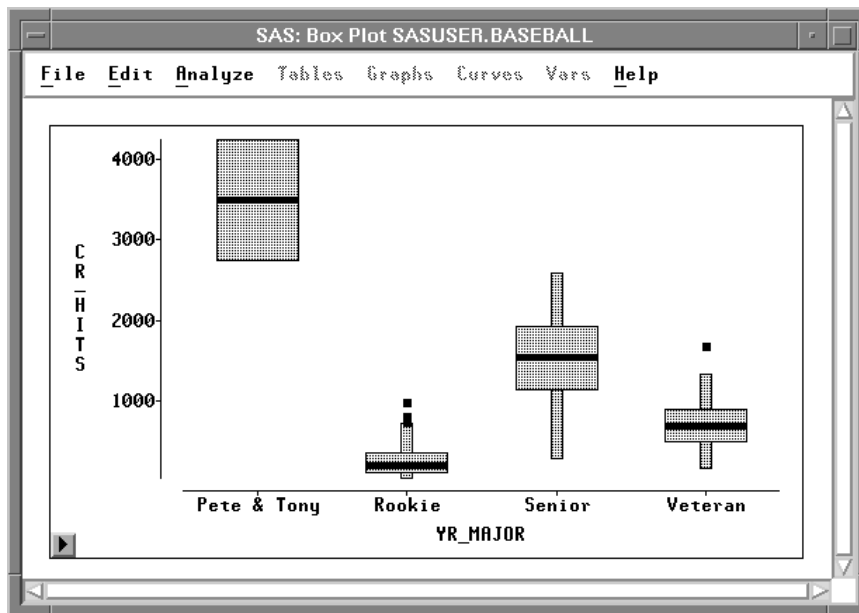

Figure 24.21. Box Plot of **CR_HITS** by **YR_MAJOR**

Since **YEARFMT** defines four formatted values, there are four boxes in the box plot. One of the boxes has no whiskers because it represents only two observations. Pete Rose and Tony Perez, ballplayers of exceptional hitting ability and longevity, are in a class by themselves.

To learn more about SAS formats, refer to *SAS Language Reference: Dictionary*. To learn more about creating your own formats with PROC FORMAT, refer to the *SAS Procedures Guide*.

⊕ **Related Reading:** Box Plots, Chapter 33.

SAS/INSIGHT User's Guide, Version 8