

## Chapter 24

# Formatting Variables and Values

### Chapter Table of Contents

---

<b>ASSIGNING FORMATS</b> . . . . .	352
<b>CREATING FORMATS</b> . . . . .	359



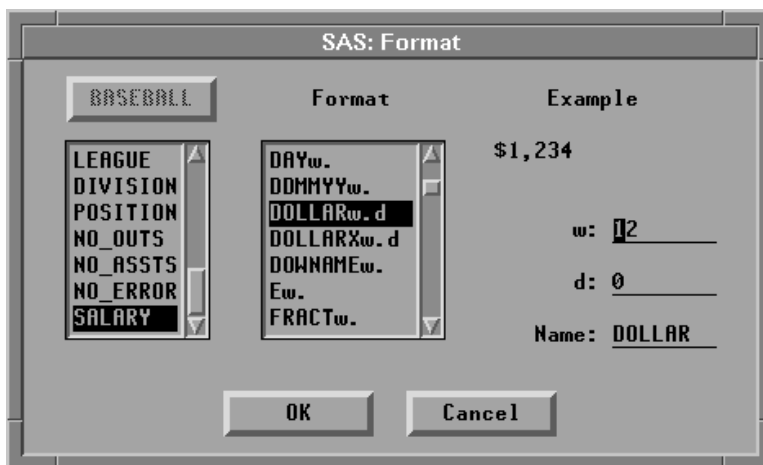
## Chapter 24

# 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.**

	Nom	Int	Int	Int	Int
	POSITION	NO_OUTS	NO_ASSTS	NO_ERROR	SALARY
1	10	317	36	1	75.000
2	C	446	33	20	.
3	UT	80	45	8	240.000
4	3S	73	152	11	225.000
5	CF	247	4	8	.
6	C	632	43	10	475.000
7	2B	186	290	17	550.000
8	RF	295	15	5	950.000
9	OF	90	4	0	.
10	1B	1236	98	18	100.000
11	C	359	30	4	305.000
12	RF	368	20	3	1237.500

Figure 24.2. SALARY Selected

⇒ **Choose Edit:Formats:9.1.**

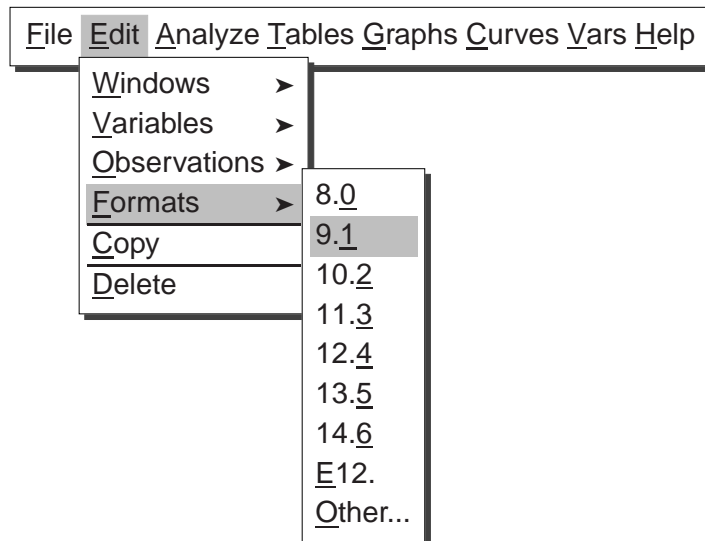
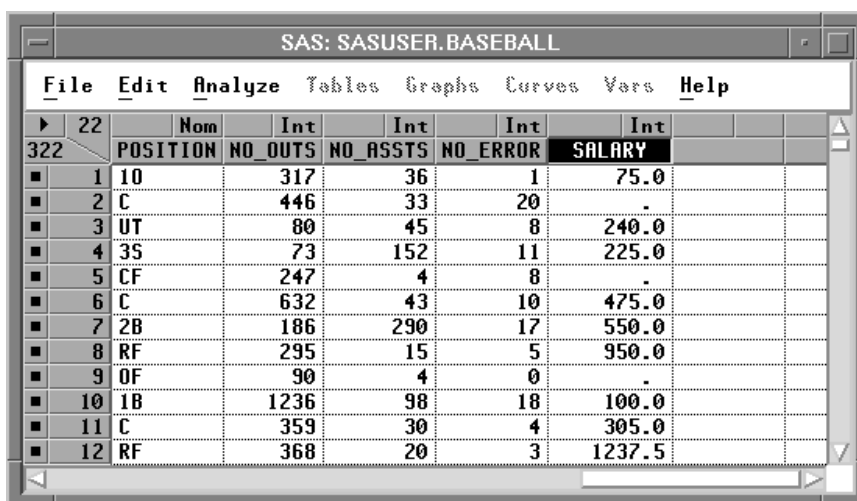


Figure 24.3. Edit:Formats Menu

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.

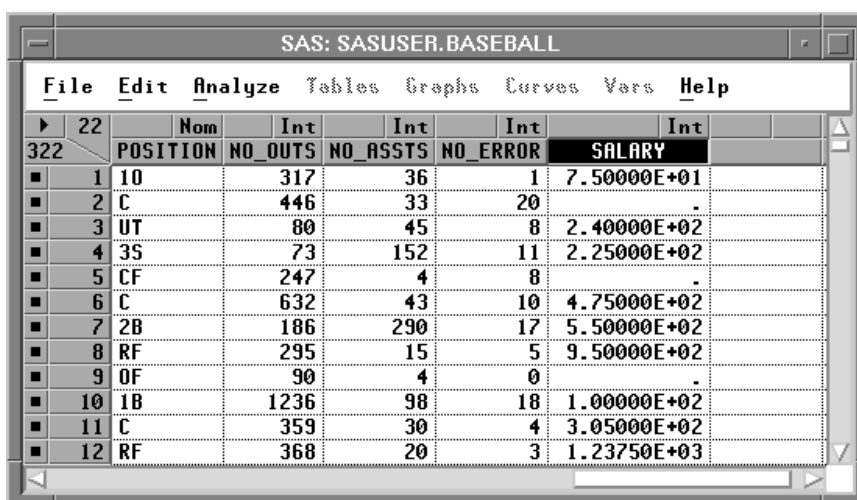


		Nom	Int	Int	Int	Int
	POSITION	NO_OUTS	NO_ASSTS	NO_ERROR	SALARY	
1	10	317	36	1	75.0	
2	C	446	33	20	.	
3	UT	80	45	8	240.0	
4	3S	73	152	11	225.0	
5	CF	247	4	8	.	
6	C	632	43	10	475.0	
7	2B	186	290	17	550.0	
8	RF	295	15	5	950.0	
9	OF	90	4	0	.	
10	1B	1236	98	18	100.0	
11	C	359	30	4	305.0	
12	RF	368	20	3	1237.5	

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.

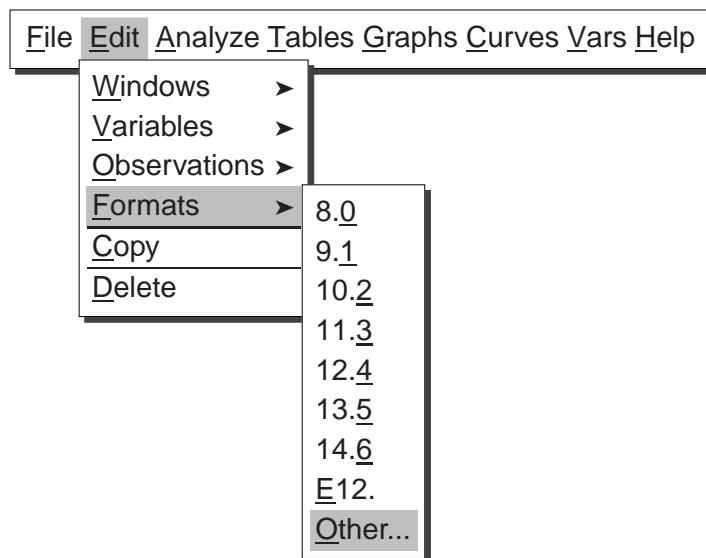


		Nom	Int	Int	Int	Int
	POSITION	NO_OUTS	NO_ASSTS	NO_ERROR	SALARY	
1	10	317	36	1	7.50000E+01	
2	C	446	33	20	.	
3	UT	80	45	8	2.40000E+02	
4	3S	73	152	11	2.25000E+02	
5	CF	247	4	8	.	
6	C	632	43	10	4.75000E+02	
7	2B	186	290	17	5.50000E+02	
8	RF	295	15	5	9.50000E+02	
9	OF	90	4	0	.	
10	1B	1236	98	18	1.00000E+02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

Figure 24.5. Format E12.

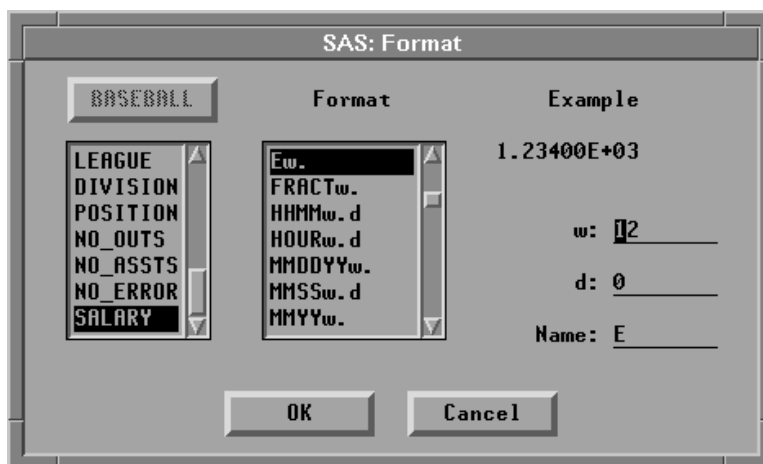
The **Edit:Formats** menu provides quick access to frequently used formats. There are many other standard formats provided by the SAS System.

⇒ Choose **Edit:Formats:Other.**



**Figure 24.6.** Edit:Formats Menu

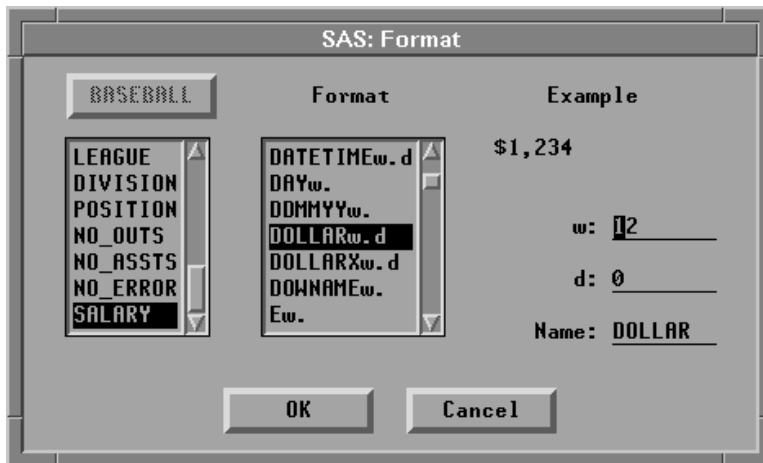
This displays the Format dialog. In the dialog, the fields **w** and **d** specify the width and decimal places to be used by the formats. Note that the **SALARY** variable and the **E12.** format are currently selected.



**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 selected. **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.

The image shows the 'SAS: SASUSER.BASEBALL' data window. The table has columns: POSITION, NO\_OUTS, NO\_ASSTS, NO\_ERROR, and SALARY. The 'SALARY' column is formatted with the DOLLARw.d format, showing values like \$75, \$240, \$225, etc. The table is as follows:

	POSITION	NO_OUTS	NO_ASSTS	NO_ERROR	SALARY
1	10	317	36	1	\$75
2	C	446	33	20	.
3	UT	80	45	8	\$240
4	3S	73	152	11	\$225
5	CF	247	4	8	.
6	C	632	43	10	\$475
7	2B	186	290	17	\$550
8	RF	295	15	5	\$950
9	OF	90	4	0	.
10	1B	1236	98	18	\$100
11	C	359	30	4	\$305
12	RF	368	20	3	\$1,238

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.

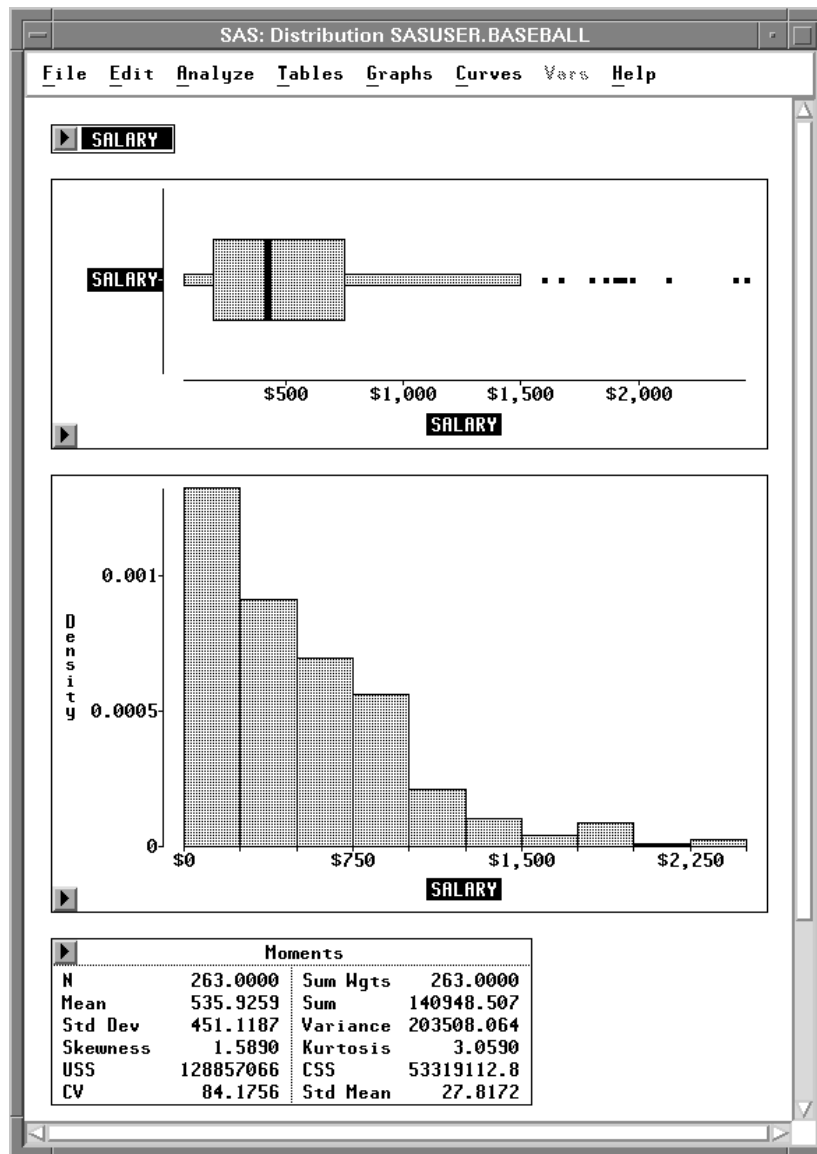
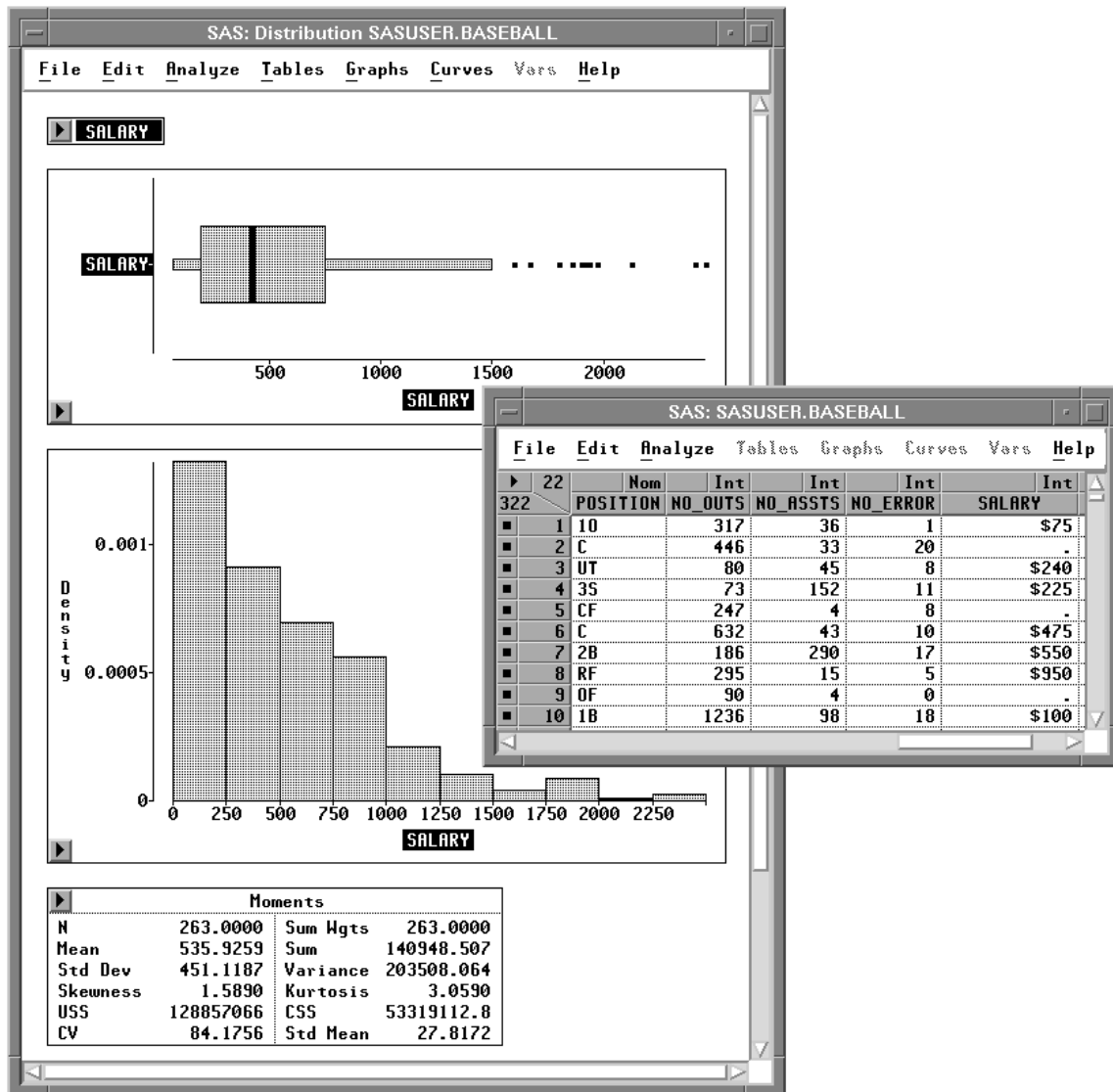


Figure 24.10. Distribution Analysis, **SALARY** Selected

⇒ 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.

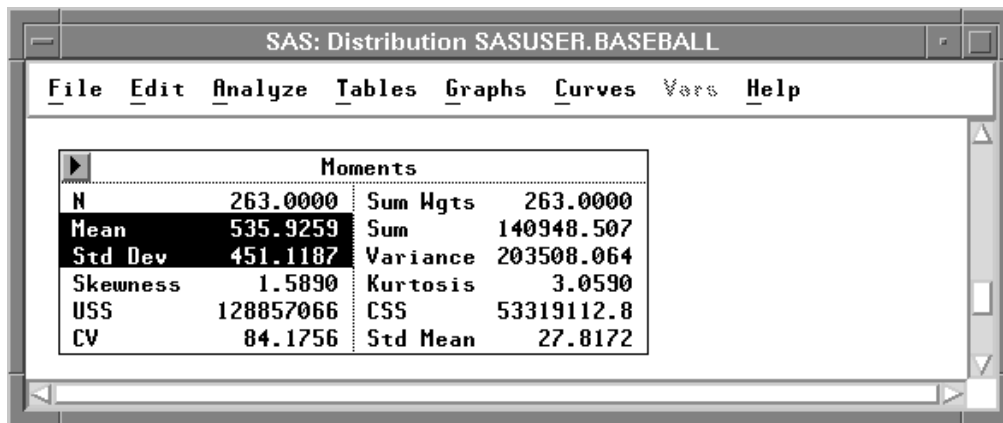




**Figure 24.11. SALARY Axes Formatted**

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.



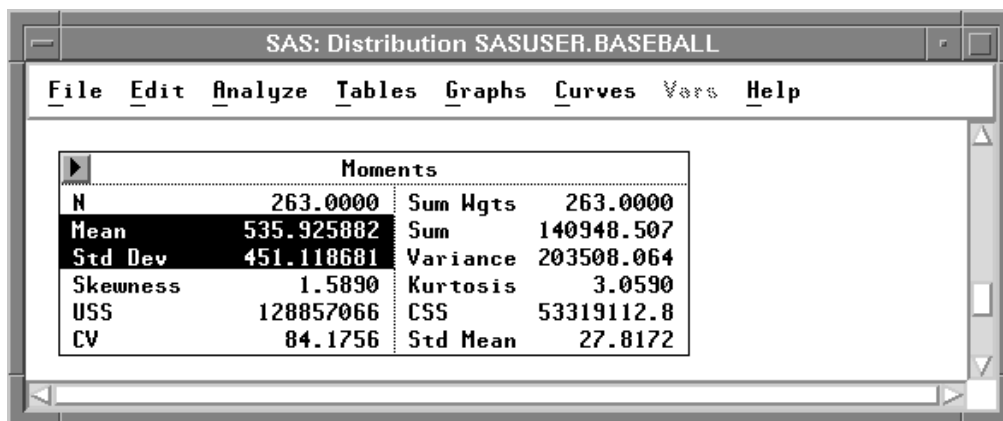
The screenshot shows the SAS: Distribution SASUSER.BASEBALL window. The menu bar includes File, Edit, Analyze, Tables, Graphs, Curves, Vars, and Help. The Moments table is displayed with the following data:

Moments			
N	263.0000	Sum Wgts	263.0000
Mean	535.9259	Sum	140948.507
Std Dev	451.1187	Variance	203508.064
Skewness	1.5890	Kurtosis	3.0590
USS	128857066	CSS	53319112.8
CV	84.1756	Std Mean	27.8172

Figure 24.12. Moments Table, Values Selected

⇒ Choose **Edit:Formats:14.6**.

Now the mean and standard deviation show six digits after the decimal.



The screenshot shows the same SAS: Distribution SASUSER.BASEBALL window, but the Mean and Std Dev values are now formatted to six decimal places:

Moments			
N	263.0000	Sum Wgts	263.0000
Mean	535.925882	Sum	140948.507
Std Dev	451.118681	Variance	203508.064
Skewness	1.5890	Kurtosis	3.0590
USS	128857066	CSS	53319112.8
CV	84.1756	Std Mean	27.8172

Figure 24.13. Moments Table, Values Formatted

## 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.**

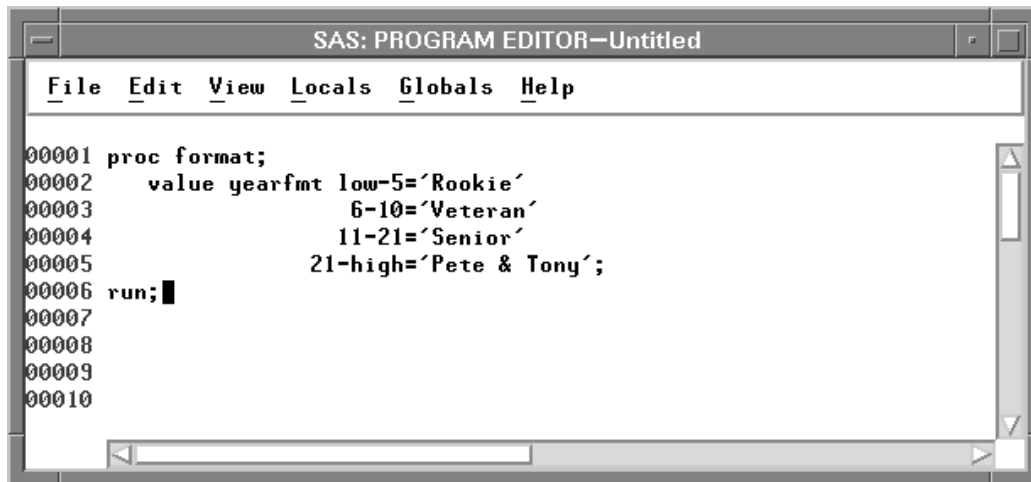


Figure 24.14. Program Editor

⇒ **Choose Run:Submit.**

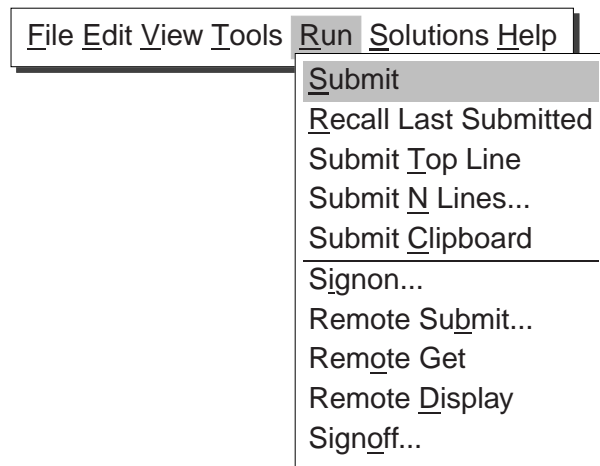


Figure 24.15. Run Menu

⇒ **Select YR\_MAJOR.**

	Int	Int	Int	Int	Int	Int	Int
	NO_HOME	NO_RUNS	NO_RBI	NO_BB	YR_MAJOR	CR_ATBAT	CR_HITS
1	2	27	25	33	1	216	54
2	1	30	29	14	1	293	66
3	7	29	27	30	13	3231	825
4	1	31	15	22	4	926	210
5	11	40	58	24	11	4513	1134
6	7	24	38	39	14	3449	835
7	1	67	27	36	7	1775	506
8	21	72	88	38	7	3754	1077
9	4	25	19	27	19	7117	1981
10	29	54	88	43	6	1750	412
11	2	28	26	22	6	999	236
12	40	107	108	69	6	2325	634

Figure 24.16. YR\_MAJOR Selected

⇒ Choose **Edit:Formats:Other**.

This displays the Format Dialog.

**SAS: Format**

Dataset: **BASEBALL**

Format List: **YR\_MAJOR**, CR\_ATBAT, CR\_HITS, CR\_HOME, CR\_RUNS, CR\_RBI, CR\_BB

Format List: **w.d**, BESTw., COMMAw.d, COMMAXw.d, DATEw., DATETIMEw.d, DAYw.

Example: **12**

Fields: **w:** 2, **d:** 0, **Name:** w.d

Buttons: **OK**, **Cancel**

Figure 24.17. Format Dialog

⇒ Enter **YEARFMT** in the **Name** field.

⇒ 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**.

The image shows a SAS data window titled 'SAS: SASUSER.BASEBALL'. It displays a table with 12 rows of data. The columns are: NO\_HOME, NO\_RUNS, NO\_RBI, NO\_BB, YR\_MAJOR, CR\_ATBAT, and CR\_H. The YR\_MAJOR column uses the YEARFMT format, displaying values like 'Rookie', 'Senior', and 'Veteran'.

	NO_HOME	NO_RUNS	NO_RBI	NO_BB	YR_MAJOR	CR_ATBAT	CR_H
1	2	27	25	33	Rookie	216	
2	1	30	29	14	Rookie	293	
3	7	29	27	30	Senior	3231	
4	1	31	15	22	Rookie	926	
5	11	40	58	24	Senior	4513	1
6	7	24	38	39	Senior	3449	
7	1	67	27	36	Veteran	1775	
8	21	72	88	38	Veteran	3754	1
9	4	25	19	27	Senior	7117	1
10	29	54	88	43	Veteran	1750	
11	2	28	26	22	Veteran	999	
12	40	107	108	69	Veteran	2325	

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.

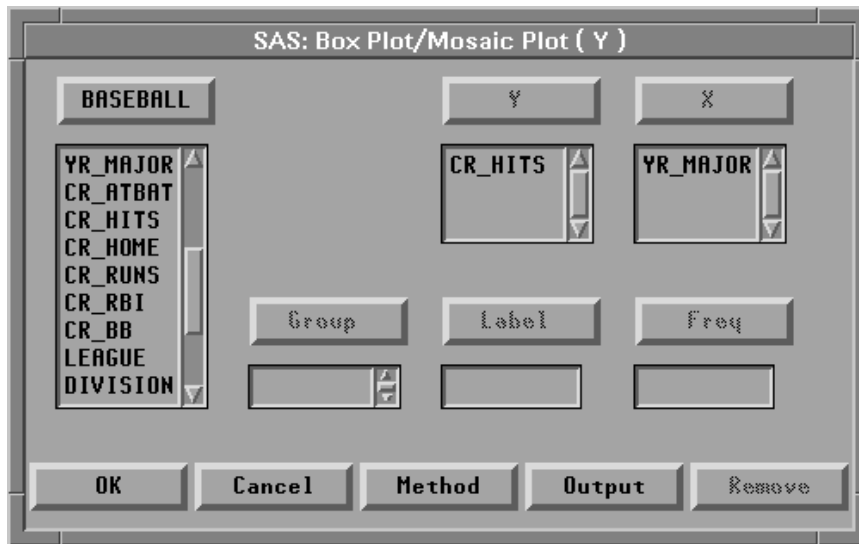


Figure 24.20. Box Plot Variables Dialog

⇒ Click the **OK** button to create the box plot.

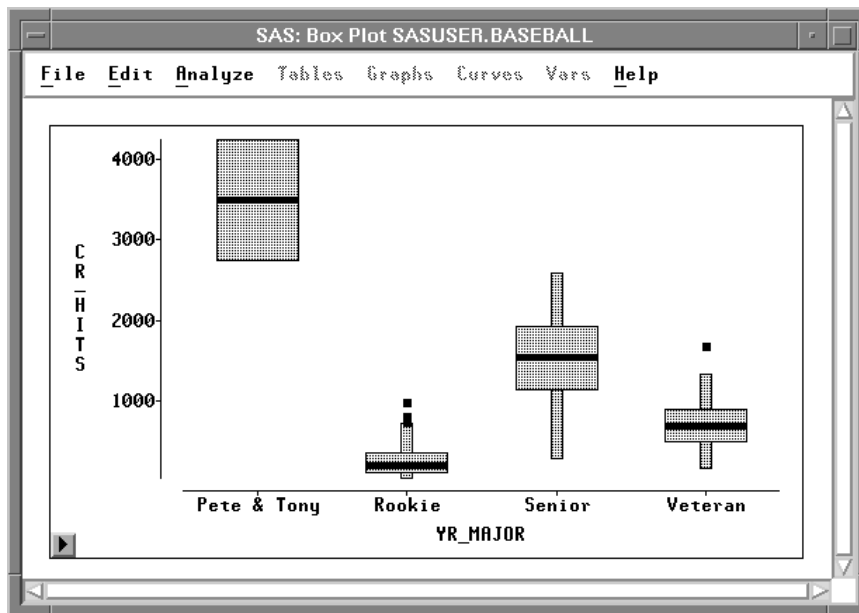


Figure 24.21. Box Plot of **CR\_HITS** by **YR\_MAJOR**

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.

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**

Copyright © 1999 by SAS Institute Inc., Cary, NC, USA.

ISBN 1-58025-490-X

All rights reserved. Printed in the United States of America. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

**U.S. Government Restricted Rights Notice.** Use, duplication, or disclosure of the software by the government is subject to restrictions as set forth in FAR 52.227-19 Commercial Computer Software-Restricted Rights (June 1987).

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

1st printing, October 1999

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries.® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

The Institute is a private company devoted to the support and further development of its software and related services.