

Chapter 22

Analyzing by Groups

Chapter Table of Contents

USING GROUP VARIABLES	332
COMPARING GROUPS BY COPYING WINDOWS	334
SETTING DEFAULT GROUP VARIABLES	337
FORMATTING GROUP VARIABLES	339

Chapter 22

Analyzing by Groups

In SAS/INSIGHT software, you can use a *group variable* to process your data separately for each value of the group variable. You can use multiple group variables to process your data separately for each unique combination of grouping values.

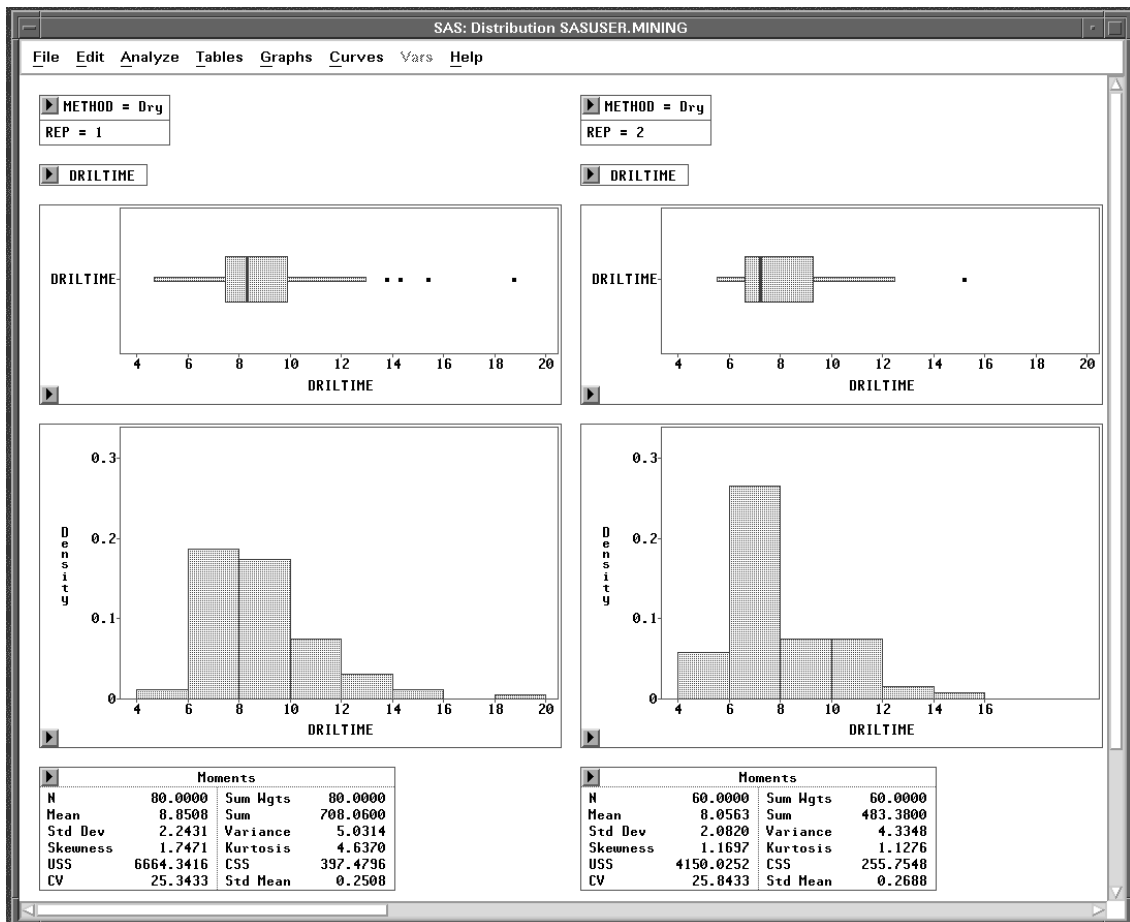


Figure 22.1. Analyzing by Groups

Using Group Variables

You can learn more about the distribution of drilling times by constructing a distribution analysis using group variables.

⇒ **Choose Analyze:Distribution (Y).**

This displays the distribution variables dialog.

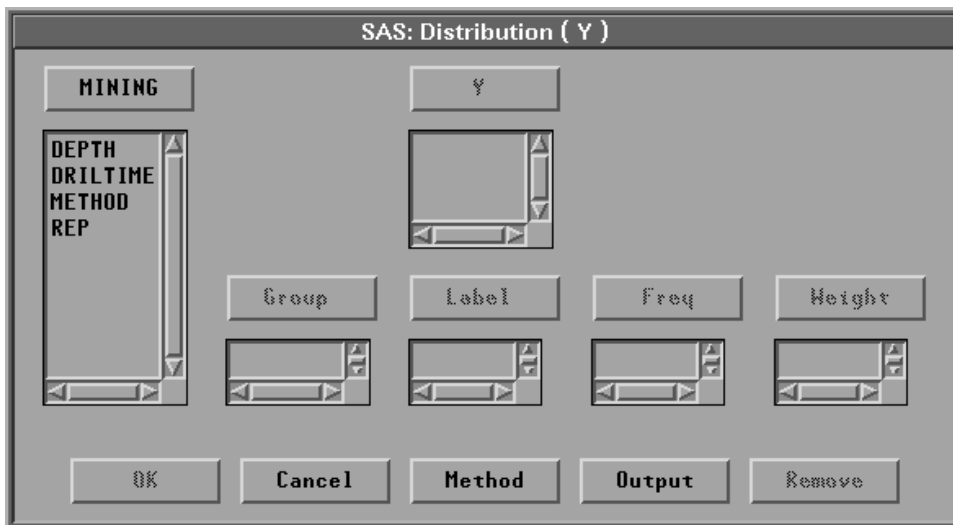


Figure 22.2. Distribution Variables Dialog

⇒ **In the dialog, select DRILTIME, then click the Y button.**

This assigns **DRILTIME** the required **Y** role.

⇒ **Select METHOD and REP, and click the Group button.**

This assigns **METHOD** and **REP** the **Group** role. You can scroll the **Group** list to see both variables. Because there are two values for **Method** and three values for **Rep**, this produces six groups.

⇒ **Click OK to create the distribution window, as shown in Figure 22.3.**

The distribution window shows detailed information on the distributions, including box plots, histograms, moments, and quantiles. At the top of the distribution window is a table indicating the unique combination of values of the two group variables. You can scroll the distribution window to the right to see other levels.

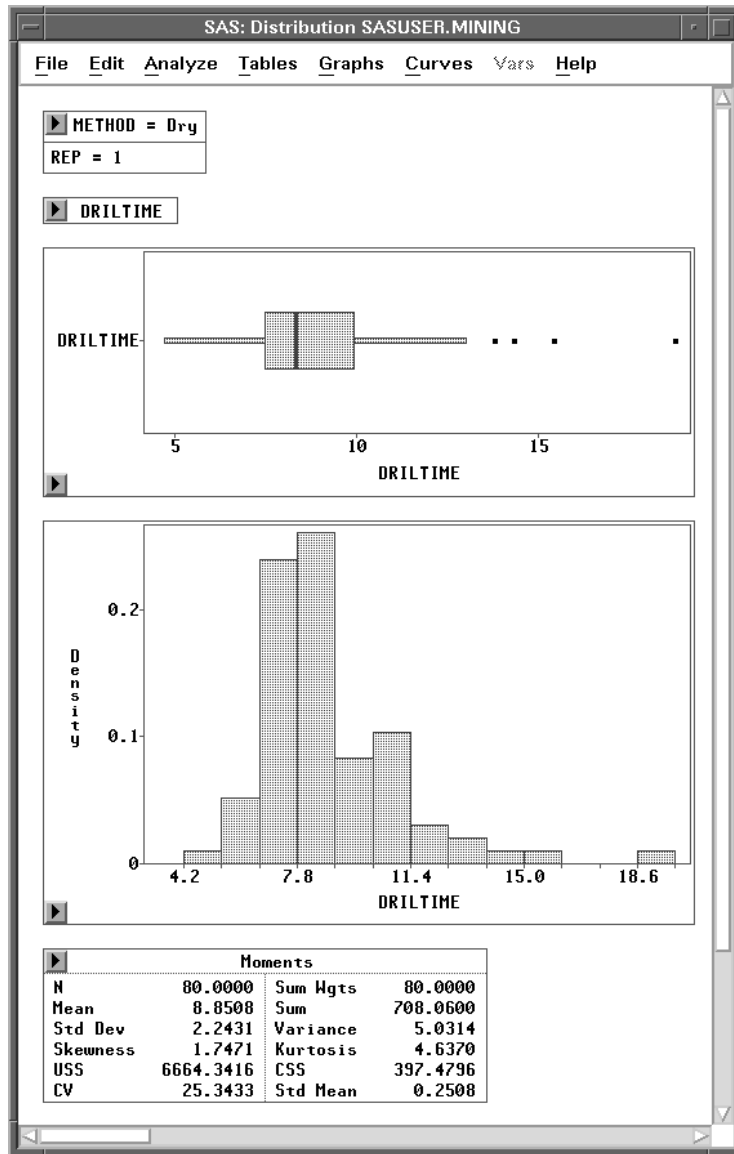


Figure 22.3. Distribution Window with Group Variables

Comparing Groups by Copying Windows

Because there are six groups, it is difficult to compare two groups side by side. Also, the axes are scaled to fit the data, so by default graphs use different axes.

To compare two groups side by side using the same axes, you can create a copy of the distribution analysis, set tick marks, and align the axes.

⇒ Choose **Edit:Windows:Copy Window in the Distribution analysis**.

This creates a copy of the distribution analysis.

⇒ Move the two analyses side by side.

Now you can scroll the windows horizontally to compare any two groups. Figure 22.4 shows the first and last groups side by side.

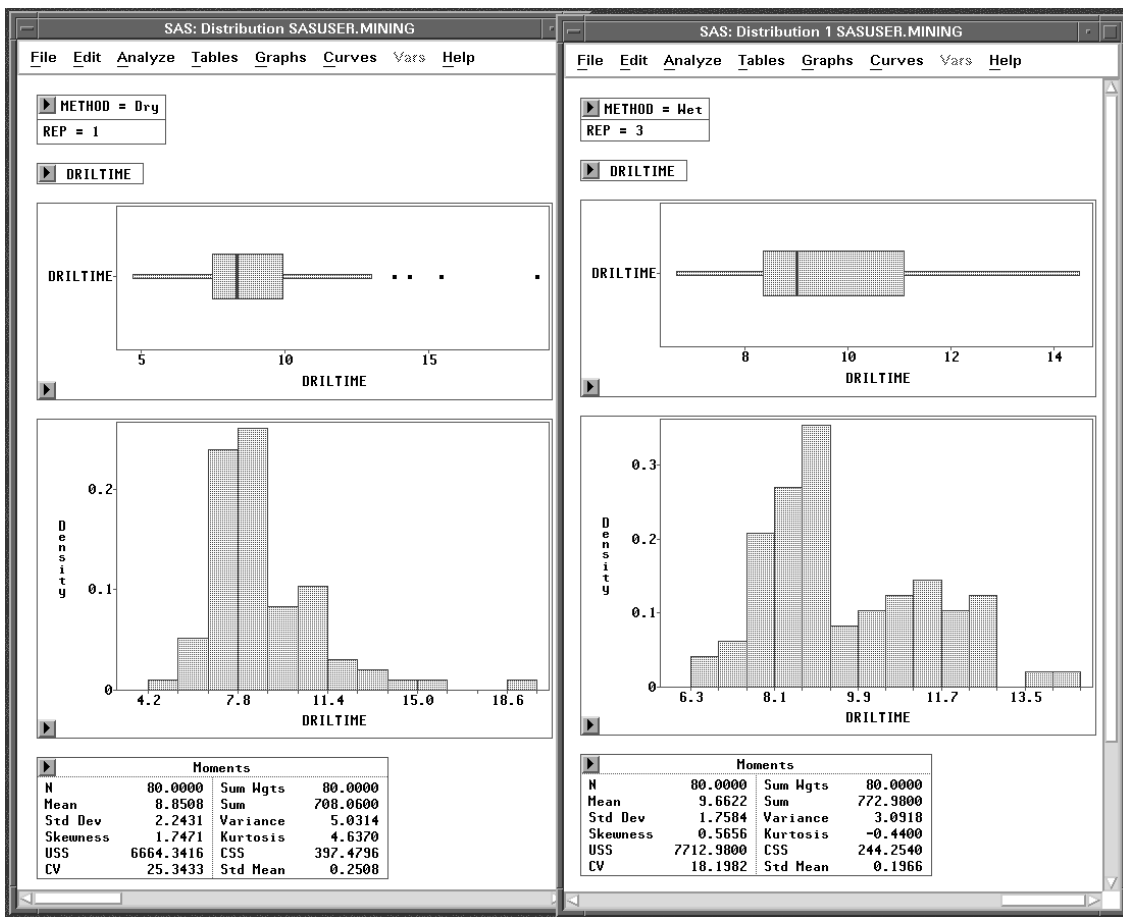


Figure 22.4. Comparing Distribution Analyses

The **Moments** and **Quantiles** tables below the histograms present statistics you can compare. The box plots and histograms, however, are difficult to compare because they use different axes. You can customize the axes with the following steps.

- ⇒ **Select DRILTIME in the first distribution window.**
- ⇒ **Choose Ticks from the histogram pop-up menu in the first window.**
This displays the **Ticks** dialog.
- ⇒ **Make the adjustments shown in the following figure, and click the OK button.**
This scales the **DRILTIME** axis for all histograms.

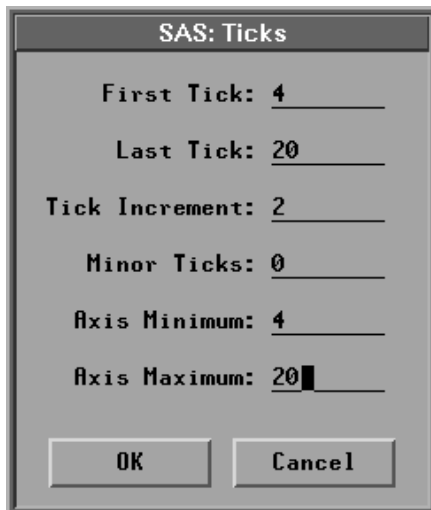


Figure 22.5. Ticks Dialog

- ⇒ **Repeat these steps for the box plots in the first window.**
This scales the **DRILTIME** axis for all box plots.
- ⇒ **Repeat these steps for the second window.**
Now you can compare box plots and histograms in both windows.

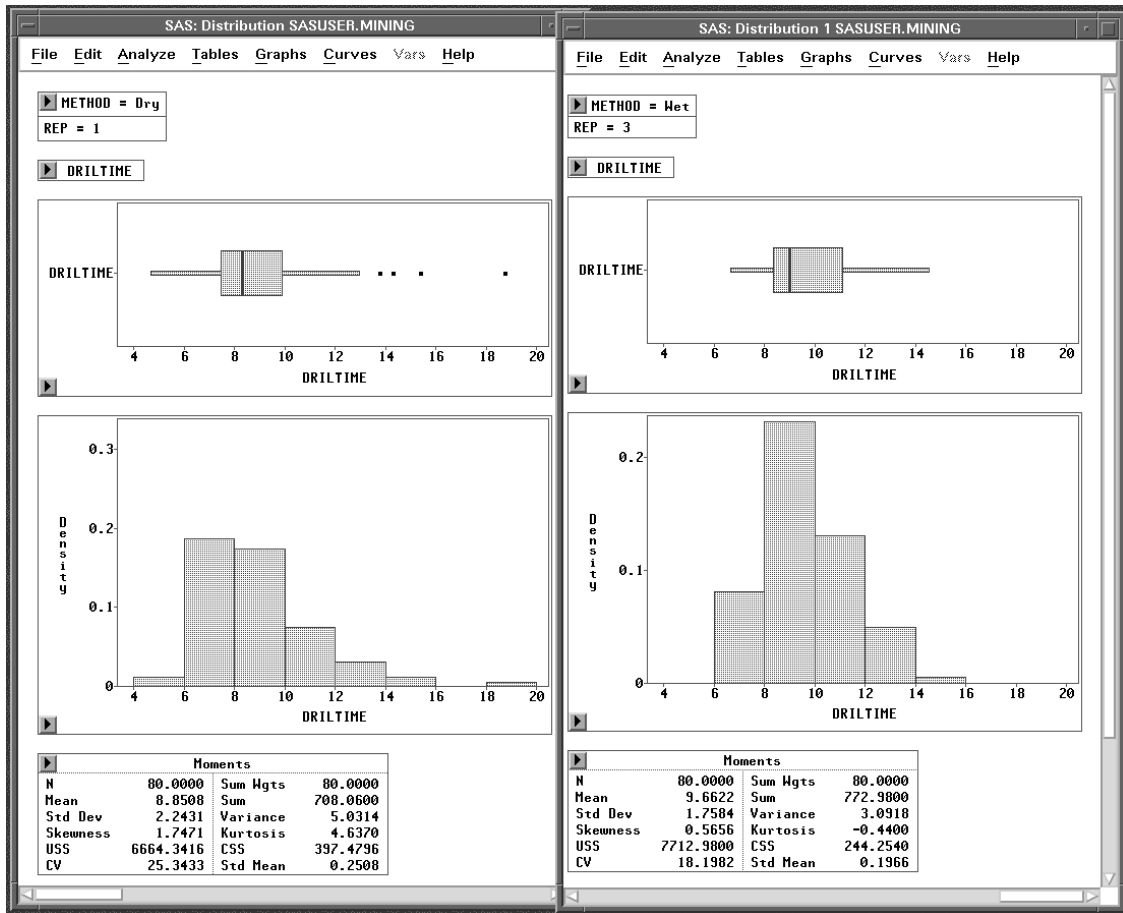


Figure 22.6. Comparing Distribution Analyses

† **Note:** By default, axes in SAS/INSIGHT software are scaled to fit the data. You can choose **Edit:Windows:Align** in any analysis window to align axes that use the same variable. Aligning affects only the axis scale, not the tick marks. When aligning histogram axes as in the preceding example, you should use the Ticks dialog to give histogram bars the same width and position.

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

Setting Default Group Variables

Often you will want to assign **Group** roles to the same group variables throughout a SAS/INSIGHT session. You can save time by setting default **Group** roles in the data window so that you do not have to set them in every variables dialog.

To set default **Group** roles for **SASUSER.MINING**, follow these steps.

⇒ **Choose Define Variables from the data pop-up menu.**

This displays the Define Variables dialog.

⇒ **In the dialog, click on METHOD, then click on Group under Default Role.**

This assigns the **Group** role to the **METHOD** variable.

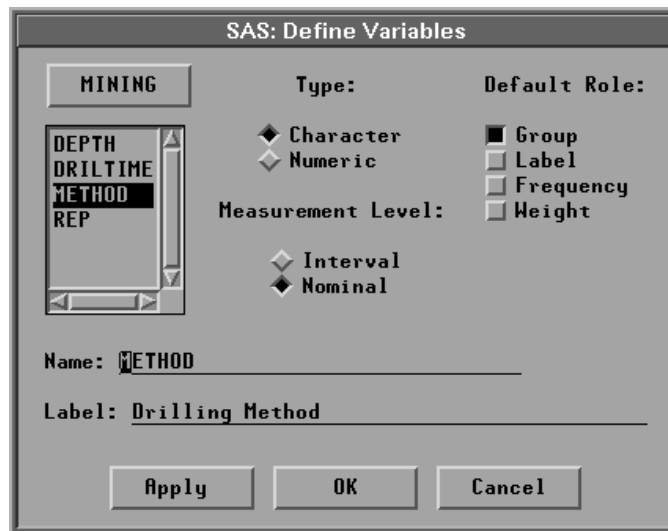


Figure 22.7. Assigning Group Role

⇒ **Click the Apply button.**

This assigns the **Group** role to **METHOD** but leaves the Define Variables window open so that you can assign roles to other variables as well.

⇒ **Click on REP, then click on Group under Default Role.**

This assigns the **Group** role to the **REP** variable as well.

⇒ **Click the OK button to close the dialog.**

The **Group** role now appears above both **METHOD** and **REP** in the data window.

	4	Int	Int	Group	Nom	Group	Int
480		DEPTH	DRILTIME	METHOD		REP	
■	1	5	7.61	Wet		1	
■	2	5	8.68	Wet		2	
■	3	5	8.61	Wet		3	
■	4	5	7.25	Dry		1	
■	5	5	7.07	Dry		2	
■	6	5	4.90	Dry		3	
■	7	10	8.16	Wet		1	
■	8	10	8.13	Wet		2	
■	9	10	7.71	Wet		3	
■	10	10	8.55	Dry		1	
■	11	10	6.62	Dry		2	
■	12	10	5.07	Dry		3	

Figure 22.8. Two Group Roles Assigned

† **Note:** Order is significant. The order in which you assign roles is the order in which your group variables are used in analyses.

⇒ Choose **Analyze:Box Plot/Mosaic Plot (Y)**.

Notice that the **Group** roles are already assigned. Only **METHOD** is visible, but you can scroll the **Group** list to see **REP**.

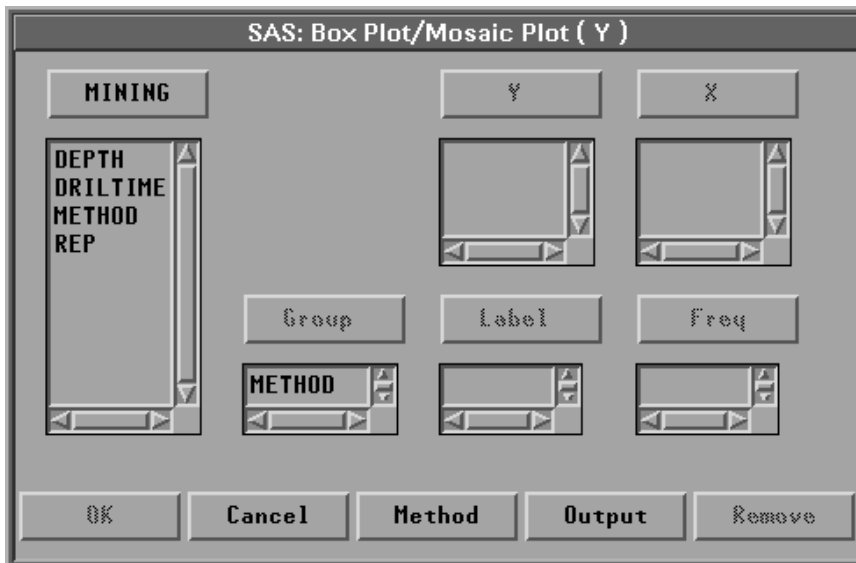


Figure 22.9. Box Plot Variables Dialog

Now every analysis you create will use the default **Group** roles you assigned in the data window. If you want to create an analysis without these variables, you can select them in the variables dialog and click the **Remove** button.

Formatting Group Variables

Usually, SAS formats in SAS/INSIGHT software determine only how data are visually displayed. Group variables, however, can use SAS formats to combine different values into a larger group. For example, suppose you are interested only in approximate depths, not in the exact values of **DEPTH**. You can use a format to combine the values of **DEPTH** into three groups:

- $\text{DEPTH} \leq 100$
- $100 < \text{DEPTH} \leq 300$
- $300 < \text{DEPTH}$

Once you have assigned this format to **DEPTH**, you can assign **DEPTH** a **Group** role and use it as described earlier in this section. Each use of **DEPTH** creates three groups containing values in the three ranges you specified.

⊕ **Related Reading:** Formats, Chapter 24.

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.