

Chapter 28

INSET Statement

Chapter Table of Contents

OVERVIEW	869
GETTING STARTED	870
Displaying Summary Statistics on a Pareto Chart	870
Customizing Labels and Formatting Values	871
Adding a Header and Positioning the Inset	873
SYNTAX	875
Summary of INSET Keywords	876
Summary of Options	877
Dictionary of Options	877
DETAILS	880
Positioning the Inset Using Compass Points	880
Positioning the Inset in the Margins	881
Positioning the Inset Using Coordinates	881

Chapter 28

INSET Statement

Overview

The INSET statement enables you to enhance a Pareto chart by adding a box or table (referred to as an *inset*) of summary statistics directly to the graph. An inset can display statistics calculated by the PARETO procedure or arbitrary values provided in a SAS data set.

Note that an INSET statement by itself does not produce a display but must be used in conjunction with a chart statement. Insets are not available with line printer output, so the INSET statement is not applicable when the LINEPRINTER option is specified in the PROC PARETO statement.

You can use options in the INSET statement to

- specify the position of the inset
- specify a header for the inset table
- specify graphical enhancements, such as background colors, text colors, text height, text font, and drop shadows

When an INSET statement is associated with a chart statement producing a comparative Pareto chart, an inset is produced in each cell of the comparative chart.

Getting Started

This section introduces the INSET statement with examples that illustrate commonly used options. Complete syntax for the INSET statement is presented in the “Syntax” section on page 875.

Displaying Summary Statistics on a Pareto Chart

During the manufacture of a metal-oxide semiconductor (MOS) capacitor, causes of failures were recorded before and after a tube in the diffusion furnace was cleaned. This information was saved in a SAS data set named FAILURE3.

```
data failure3;
  length cause $ 16 stage $ 16 ;
  label cause = 'Cause of Failure' ;
  input stage $ 1-16 cause $ 19-34 counts;
  datalines;
Before Cleaning Contamination 14
Before Cleaning Corrosion 2
Before Cleaning Doping 1
Before Cleaning Metallization 2
Before Cleaning Miscellaneous 3
Before Cleaning Oxide Defect 8
Before Cleaning Silicon Defect 1
After Cleaning Doping 0
After Cleaning Corrosion 2
After Cleaning Metallization 4
After Cleaning Miscellaneous 2
After Cleaning Oxide Defect 1
After Cleaning Contamination 12
After Cleaning Silicon Defect 2
;
```

The following statements generate a Pareto chart from the FAILURE3 data. The MAXNCAT= option is specified to limit the number of Pareto categories to five. An INSET statement is used to display the total count for the categories displayed and the total count for the categories excluded by the MAXNCAT= option.

```
title 'IC Failures';
proc pareto data=failure3;
  vbar cause /
    freq      = counts
    maxncat   = 5
  ;
  inset n nexcl / height = 3;
run;
```

The resulting chart is displayed in Figure 28.1. The INSET statement immediately follows the chart statement that creates the graphical display (in this case, the VBAR statement). Specify the keywords for inset statistics (such as N and NEXCL) immediately after the word INSET. The inset statistics appear in the order in which you specify the keywords. The HEIGHT= option in the INSET statement specifies the text height used to display the statistics in the inset.

A complete list of keywords that you can use with the INSET statement is provided in “Summary of INSET Keywords” on page 876.

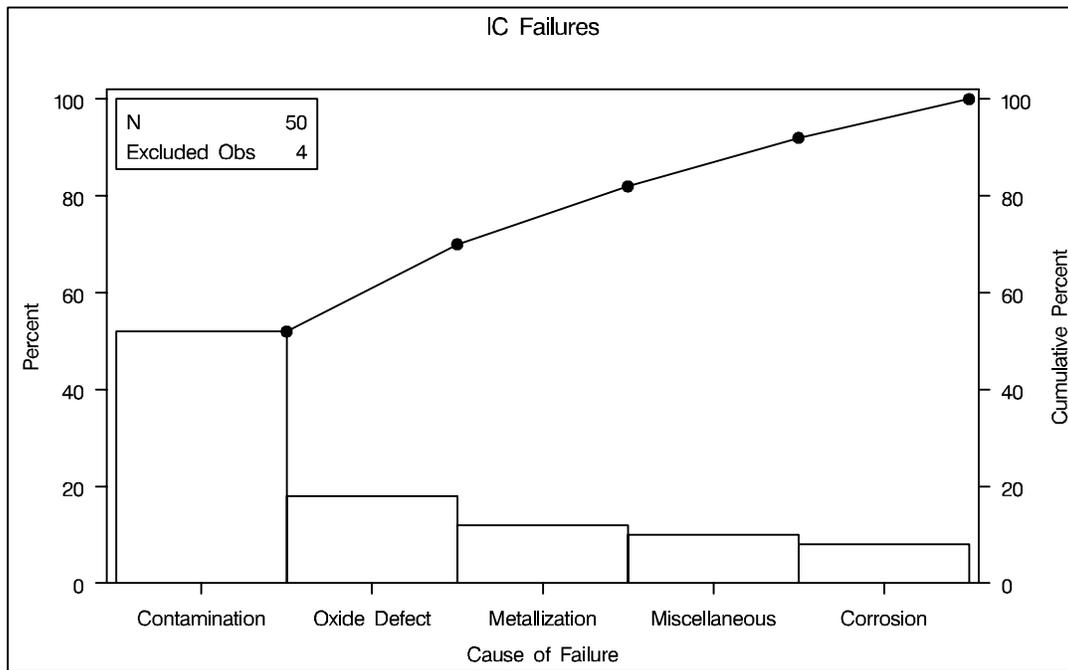


Figure 28.1. A Pareto Chart with an Inset

The following examples illustrate options commonly used for enhancing the appearance of an inset.

Customizing Labels and Formatting Values

By default, each inset statistic is identified with an appropriate label, and each numeric value is printed using an appropriate format. However, you may want to provide your own labels and formats. For example, in Figure 28.1 the default label used for the NEXCL statistic is rather long. The following statements produce a comparative Pareto chart with insets using a shorter label for the number of excluded observations. A format with one decimal place is also specified for each statistic. Note that a single INSET statement produces an inset in each cell of the comparative Pareto chart.

Part 7. The CAPABILITY Procedure

```

title 'Comparison of IC Failures';
proc pareto data=failure3;
  vbar cause /
    class    = stage
    freq     = counts
    maxncat  = 5
    classkey = 'Before Cleaning'
  ;
  inset n (3.1) nexcl='N Excl' (3.1) / height = 3;
run;

```

The resulting chart is displayed in Figure 28.2. You can provide your own label by specifying the keyword for that statistic followed by an equal sign (=) and the label in quotes. The label can have up to 24 characters.

The format 3.1 specified in parentheses after the N and NEXCL keywords displays those statistics with a field width of three and one decimal place. In general, you can specify any numeric SAS format in parentheses after an inset keyword. You can also specify a format to be used for all the statistics in the INSET statement with the FORMAT= option. For more information about SAS formats, refer to Chapter 14 of *SAS Language Reference: Dictionary*.

Note that if you specify both a label and a format for a statistic, the label must appear before the format.

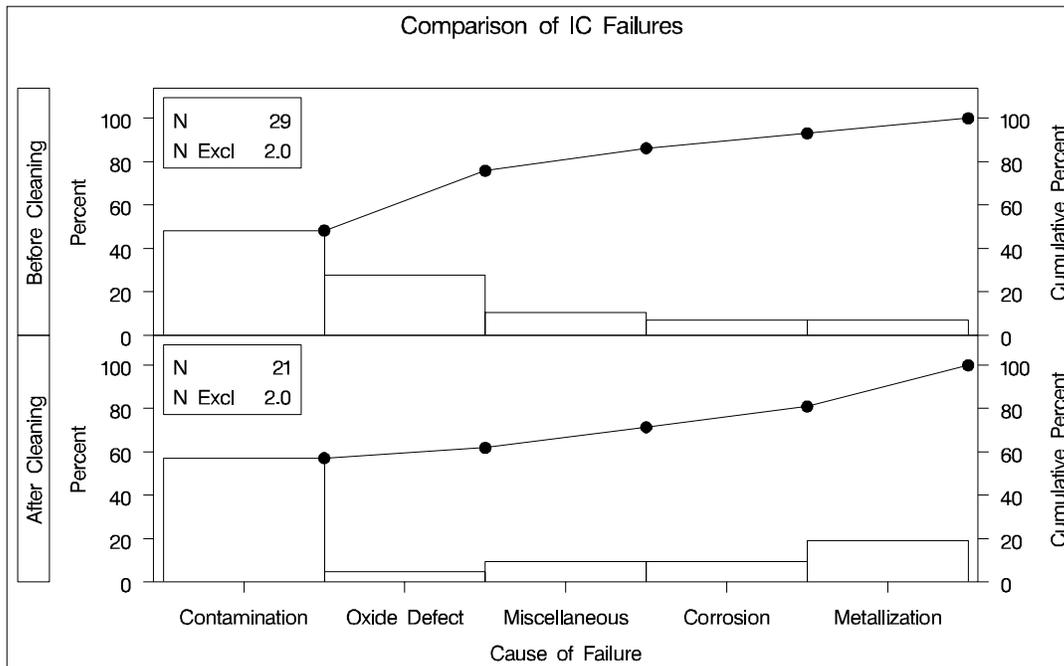


Figure 28.2. Customizing Labels and Formatting Values in an Inset

Adding a Header and Positioning the Inset

In the previous examples, the insets are displayed in the upper left corners of the plots, which is the default position for insets added to Pareto charts. You can control the inset position with the POSITION= option. In addition, you can display a header at the top of the inset with the HEADER= option. The following statements create a data set to be used with the INSET DATA= keyword and produce the horizontal Pareto chart shown in Figure 28.3:

```

data location;
  length _LABEL_ $ 10 _VALUE_ $ 12;
  input _LABEL_ _VALUE_ &;
datalines;
Plant      Santa Clara
Line      1
;

title 'IC Failures';
proc pareto data=failure3;
  hbar cause /
    freq      = counts
    maxncat   = 5;
  inset data = location n nexcl /
    height    = 3
    position  = rm
    cshadow   = black
    header    = 'Count Summary';
run;

```

The header (in this case, *Count Summary*) can be up to 40 characters. Note that a longer list of inset statistics is requested. Consequently, POSITION=RM is specified to position the inset in the right margin so that it does not interfere with features of the chart. For more information about positioning, see “Details” on page 880. The CSHADOW= option is used to display a drop shadow on this inset. The *options*, such as HEADER=, POSITION= and CSHADOW=, are specified after the slash (/) in the INSET statement. For more details on INSET statement options, see “Dictionary of Options” on page 877.

Note that the contents of the data set LOCATION appear before other statistics in the inset. The position of the DATA= keyword in the keyword list determines the position of the data set’s contents in the inset.

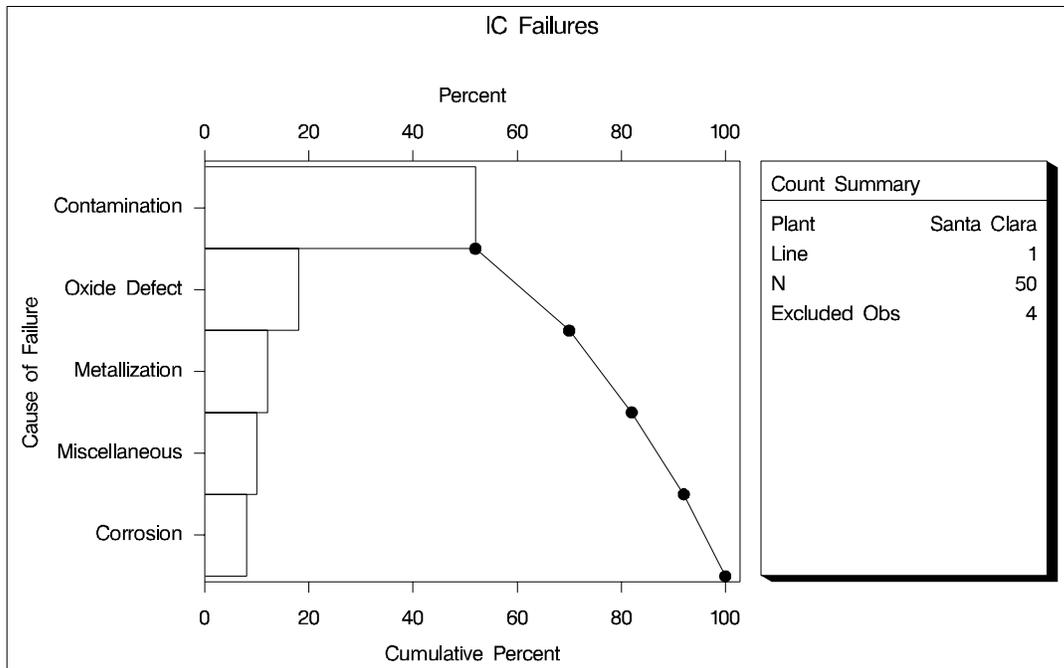


Figure 28.3. Adding a Header and Repositioning the Inset

Syntax

The syntax for the INSET statement is as follows:

```
INSET keyword-list < / options >;
```

You can use any number of INSET statements in the PARETO procedure. Each INSET statement produces a separate inset and must follow one of the chart statements. When the chart statement produces a comparative Pareto chart, an inset appears in every cell produced by the chart statement. The statistics are displayed in the order in which they are specified. The following statements produce a vertical Pareto chart with insets in the upper left and upper right corners, and a horizontal comparative Pareto chart with insets in each cell.

```
proc pareto data=failure3;
  vbar cause / maxncat = 5 other = 'Others';
    inset nothercat / pos = nw;
    inset nother / pos = ne;
  hbar cause / class = stage classkey = 'Before Cleaning';
    inset n / pos = ne;
run;
```

The statistics displayed in an inset are computed for a specific process variable using observations for the current BY group and CLASS= variable level, if applicable. For example, in the following statements there are two process variables (TOMATO and SQUASH), a BY variable (YEAR), and two CLASS= variables (FERT and PEST). If there are three different years (levels of YEAR), then a total of six comparative Pareto charts are produced: three for each process variable. In addition, if there are two different levels of FERT and three of PEST, each comparative Pareto chart contains six cells. Each cell contains an inset with statistics computed for a particular process variable, year, and combination of FERT and PEST values.

```
proc pareto data=plants;
  by year;
  vbar (tomato squash) / class = (fert pest);
  inset n;
run;
```

The components of the INSET statement are described as follows.

keyword-list

can include any of the *keywords* listed in “Summary of INSET Keywords” on page 876. The DATA= *keyword* requires an operand specified immediately after it, naming the data set containing data to be displayed.

The NOTHERCAT and NOTHER statistics are zero if the OTHER= option is not specified. The NEXCL statistic is zero if the OTHER= option *is* specified.

By default, inset statistics are identified with appropriate labels, and numeric values are printed using appropriate formats. However, you can provide customized labels and formats. You provide the customized label by specifying the *keyword* for that statistic followed by an equal sign (=) and the label in quotes. Labels can have up

to 24 characters. You provide the numeric format in parentheses after the *keyword*. Note that if you specify both a label and a format for a statistic, the label must appear before the format. For an example, see “Customizing Labels and Formatting Values” on page 871.

options

appear after the slash (/) and control the appearance of the inset. For example, the following INSET statement uses two appearance *options* (POSITION= and CTEXT=):

```
inset n nothercat nother / position=ne ctext=yellow;
```

The POSITION= option determines the location of the inset, and the CTEXT= option specifies the color of the text of the inset.

See “Summary of Options” on page 877 for a list of all available *options* and “Dictionary of Options” on page 877 for detailed descriptions. Note the difference between *keywords* and *options*; *keywords* specify the information to be displayed in an inset, whereas *options* control the appearance of the inset.

Summary of INSET Keywords

All keywords available with the PARETO procedure’s INSET statement request a single statistic in an inset, except for the DATA= keyword.

The DATA= keyword specifies a SAS data set containing (label, value) pairs to be displayed in an inset. The data set must contain the variables _LABEL_ and _VALUE_. _LABEL_ is a character variable whose values provide labels for inset entries. _VALUE_ can be character or numeric, and it provides values displayed in the inset. The label and value from each observation in the DATA= data set occupy one line in the inset. Figure 28.3 shows an inset containing entries from a DATA= data set.

Table 28.1. Summary Statistics

N	sample size
NOTHERCAT	number of categories merged to form OTHER= category in restricted Pareto chart
NOTHER	number of observations in OTHER= category
NEXCL	observations excluded in restricted Pareto chart
SUMWGTS	sum across all categories of weighted frequencies
DATA=	(label, value) pairs from <i>SAS-data-set</i>

Summary of Options

The following table lists the INSET statement options. For complete descriptions, see “Dictionary of Options,” which follows this section.

Table 28.2. INSET Options

CFILL= <i>color</i> BLANK	specifies color of inset background
CFILLH= <i>color</i>	specifies color of header background
CFRAME= <i>color</i>	specifies color of frame
CHEADER= <i>color</i>	specifies color of header text
CSHADOW= <i>color</i>	specifies color of drop shadow
CTEXT= <i>color</i>	specifies color of inset text
DATA	specifies data units for POSITION=(<i>x</i> , <i>y</i>) coordinates
FONT= <i>font</i>	specifies font of text
FORMAT= <i>format</i>	specifies format of values in inset
HEADER= <i>'quoted string'</i>	specifies header text
HEIGHT= <i>value</i>	specifies height of inset text
NOFRAME	suppresses frame around inset
POSITION= <i>position</i>	specifies position of inset
REFPOINT=BR BL TR TL	specifies reference point of inset positioned with POSITION=(<i>x</i> , <i>y</i>) coordinates

Dictionary of Options

The following entries provide detailed descriptions of options for the INSET statement. Terms used in this section are illustrated in Figure 28.4.

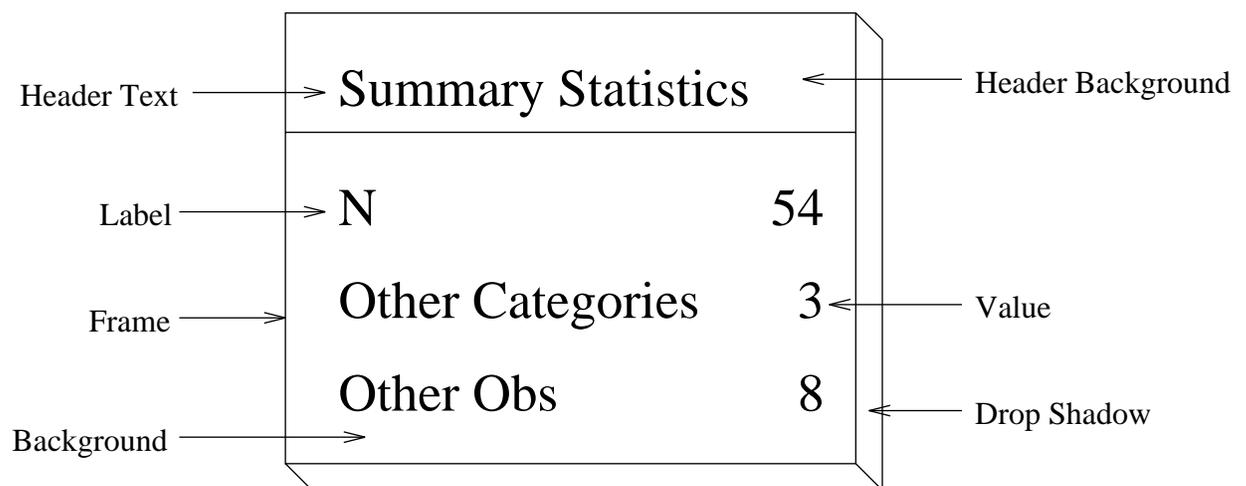


Figure 28.4. The Inset

CFILL=*color* | BLANK

specifies the color of the background (including the header background if you do not specify the CFILLH= option).

If you do not specify the CFILL= option, then, by default, the background is empty. This means that items that overlap the inset (such as subgroup data points or control limits) show through the inset. If you specify any value for the CFILL= option, then overlapping items no longer show through the inset. Specify CFILL=BLANK to leave the background uncolored and also to prevent items from showing through the inset.

CFILLH=*color*

specifies the color of the header background. By default, if you do not specify a CFILLH= color, the CFILL= color is used.

CFRAME=*color*

specifies the color of the frame. By default, the frame is the same color as the axis of the plot.

CHEADER=*color*

specifies the color of the header text. By default, if you do not specify a CHEADER= color, the CTEXT= color is used.

CSHADOW=*color*

CS=*color*

specifies the color of the drop shadow. See Figure 28.3 on page 874 for an example. By default, if you do not specify the CSHADOW= option, a drop shadow is not displayed.

CTEXT=*color*

CT=*color*

specifies the color of the text. By default, the inset text color is the same as the other text on the plot.

DATA

specifies that data coordinates are to be used in positioning the inset with the POSITION= option. The DATA option is available only when you specify POSITION= (*x*, *y*), and it must be placed immediately after the coordinates (*x*, *y*). For details, see the entry for the POSITION= option or “Positioning the Inset Using Coordinates” on page 881. See Figure 28.7 on page 882 for an example.

FONT=*font*

specifies the font of the text. By default, the font is SIMPLEX if the inset is located in the interior of the plot, and the font is the same as the other text displayed on the plot if the inset is located in the exterior of the plot.

FORMAT=*format*

specifies a format for all the values displayed in an inset. If you specify a format for a particular statistic, then this format overrides the format you specified with the FORMAT= option.

HEADER= *'string'*

specifies the header text. The *string* cannot exceed 40 characters. If you do not specify the HEADER= option, no header line appears in the inset.

HEIGHT=*value*

specifies the height of the text.

NOFRAME

suppresses the frame drawn around the text.

POSITION=*position***POS=***position*

determines the position of the inset. The *position* can be a compass point keyword, a margin keyword, or a pair of coordinates (x, y). You can specify coordinates in axis percent units or axis data units. For more information, see “Details” on page 880. By default, POSITION=NW, which positions the inset in the upper left (northwest) corner of the display.

REFPOINT=BR | BL | TR | TL**RP=BR | BL | TR | TL**

specifies the reference point for an inset that is positioned by a pair of coordinates with the POSITION= option. Use the REFPOINT= option with POSITION= coordinates. The REFPOINT= option specifies which corner of the inset frame you want positioned at coordinates (x, y). The keywords BL, BR, TL, and TR represent bottom left, bottom right, top left, and top right, respectively. See Figure 28.8 on page 883 for an example. The default is REFPOINT=BL.

If you specify the position of the inset as a compass point or margin keyword, the REFPOINT= option is ignored. For more information, see “Positioning the Inset Using Coordinates” on page 881.

Details

This section provides details on three different methods of positioning the inset using the POSITION= option. With the POSITION= option, you can specify

- compass points
- keywords for margin positions
- coordinates in data units or percent axis units

Positioning the Inset Using Compass Points

You can specify the eight compass points N, NE, E, SE, S, SW, W, and NW as keywords for the POSITION= option. The following statements create the display in Figure 28.5, which demonstrates all eight compass positions. The default is NW.

```
proc pareto data=failure3;
  vbar cause / freq = counts;
  inset n / height=3 cfill=blank header='NW' pos=nw;
  inset n / height=3 cfill=blank header='N ' pos=n ;
  inset n / height=3 cfill=blank header='NE' pos=ne;
  inset n / height=3 cfill=blank header='E ' pos=e ;
  inset n / height=3 cfill=blank header='SE' pos=se;
  inset n / height=3 cfill=blank header='S ' pos=s ;
  inset n / height=3 cfill=blank header='SW' pos=sw;
  inset n / height=3 cfill=blank header='W ' pos=w ;
run;
```

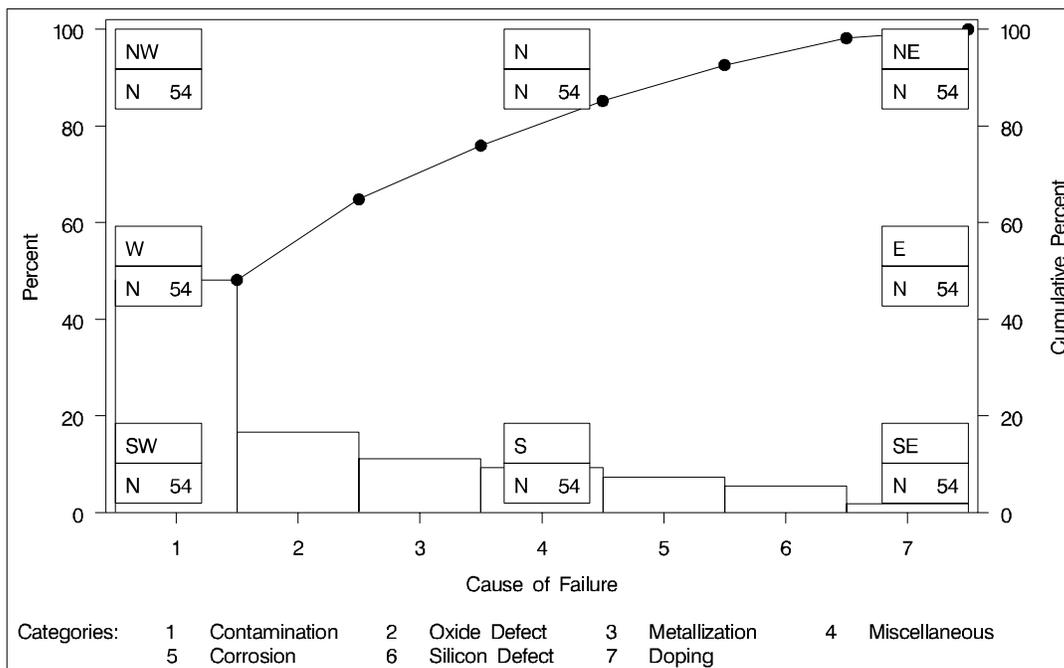


Figure 28.5. Insets Positioned Using Compass Points

Positioning the Inset in the Margins

Using the INSET statement you can also position an inset in one of the four margins surrounding the plot area using the margin keyword LM, RM, TM, or BM, as illustrated in Figure 28.6.

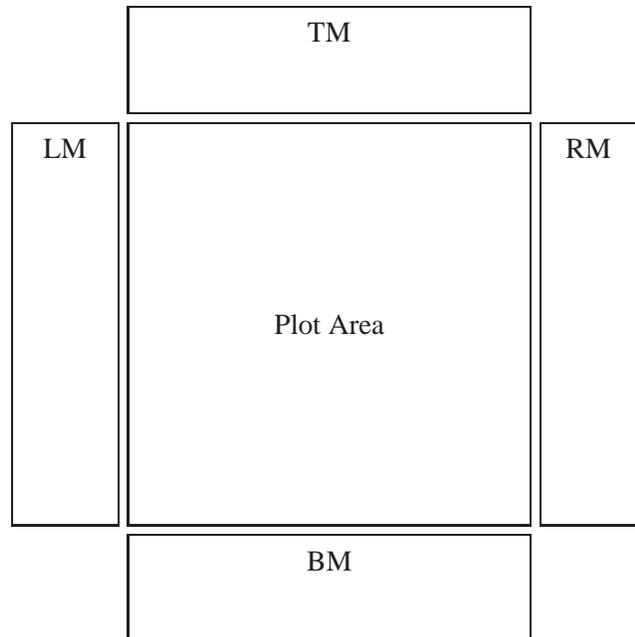


Figure 28.6. Positioning Insets in the Margins

For an example of an inset placed in the right margin, see Figure 28.3 on page 874. Margin positions are recommended if a large number of statistics are listed in the INSET statement. If you attempt to display a lengthy inset in the interior of the plot, it is likely that the inset will collide with the data display.

Insets associated with a comparative Pareto chart cannot be positioned in the margins.

Positioning the Inset Using Coordinates

You can also specify the position of the inset with coordinates: `POSITION= (x, y)`. The coordinates can be given in axis percent units (the default) or in axis data units.

Data Unit Coordinates

If you specify the DATA option immediately following the coordinates, the inset is positioned using axis data units. Data units along the category axis (the horizontal axis on a vertical bar chart or the vertical axis on a horizontal bar chart) are based on category numbers. Categories are numbered from left to right (vertical bar chart) or top to bottom (horizontal bar chart), starting with 1.

Part 7. The CAPABILITY Procedure

For example, the following statements produce the Pareto chart displayed in Figure 28.7:

```
proc pareto data=failure3;
  vbar cause / freq = counts;
  inset n / header = 'Position=(3,60)';
  position = (3,60) data;
  cfill = blank;
  height = 3;
run;
```

The bottom left corner of the inset is lined up with the tick mark for the third category on the horizontal axis and at 60 on the vertical axis. By default, the specified coordinates determine the position of the bottom left corner of the inset. You can change this reference point with the REFPOINT= option, as in the next example.

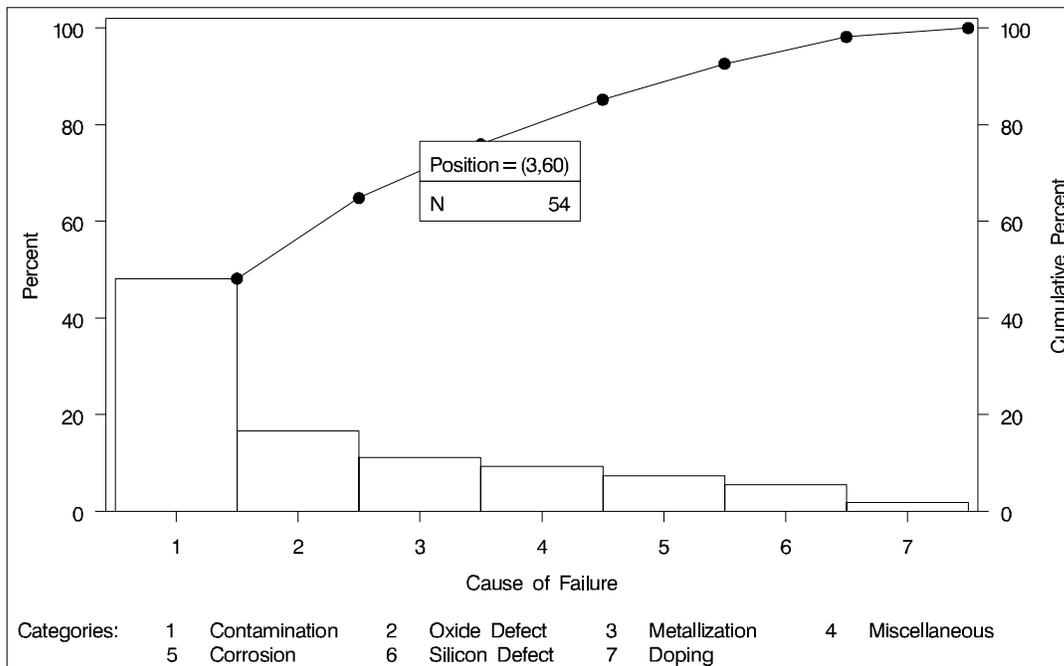


Figure 28.7. Inset Positioned Using Data Unit Coordinates

Axis Percent Unit Coordinates

If you do not use the DATA option, the inset is positioned using axis percent units. The coordinates of the bottom left corner of the display are (0, 0), while the upper right corner is (100, 100). For example, the following statements create a horizontal Pareto chart with two insets, both positioned using coordinates in axis percent units.

```

proc pareto data=failure3;
  vbar cause / freq    = counts
              maxncat = 5;
  inset n / position = (5,25)
        header   = 'Position=(5,25)'
        height   = 3
        cfill    = blank
        refpoint = tl;
  inset nexcl / position = (95,95)
             header   = 'Position=(95,95)'
             height   = 3
             cfill    = blank
             refpoint = tr;
run;

```

The display is shown in Figure 28.8. Notice that the REFPOINT= option is used to determine which corner of the inset is to be placed at the coordinates specified with the POSITION= option. The first inset has REFPOINT=TL, so the top left corner of the inset is positioned 5% of the way across the horizontal axis and 25% of the way up the vertical axis. The second inset has REFPOINT=TR, so the top right corner of the inset is positioned 95% of the way across the horizontal axis and 95% of the way up the vertical axis. Note also that coordinates in axis percent units must be *between* 0 and 100.

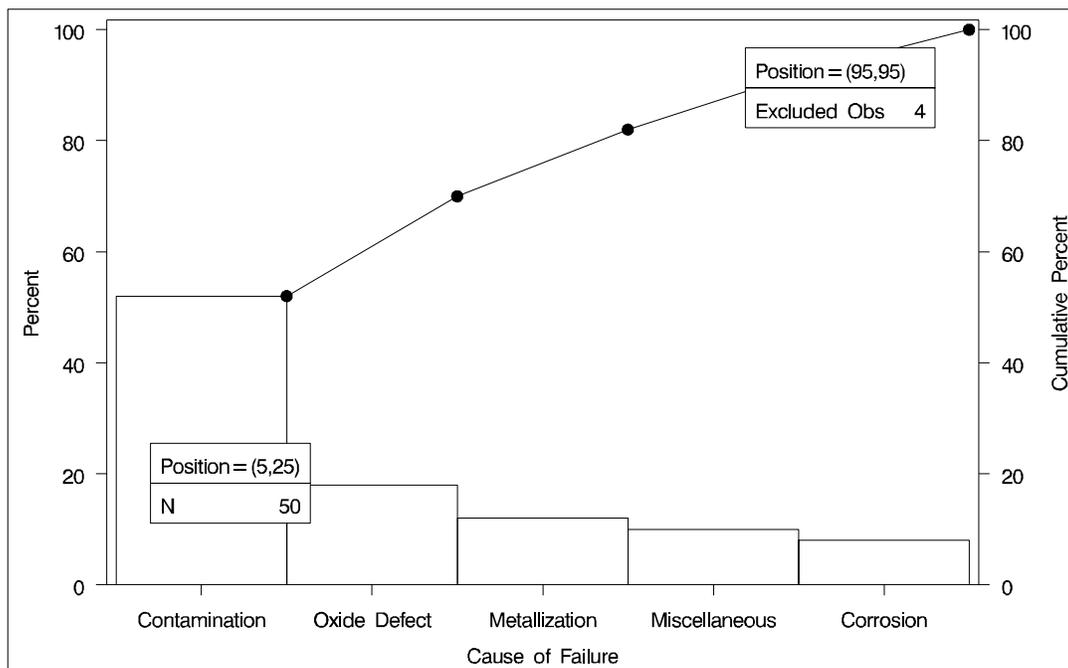


Figure 28.8. Inset Positioned Using Axis Percent Unit Coordinates

The correct bibliographic citation for this manual is as follows: SAS Institute Inc., *SAS/QC[®] User's Guide, Version 8*, Cary, NC: SAS Institute Inc., 1999. 1994 pp.

SAS/QC[®] User's Guide, Version 8

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

ISBN 1-58025-493-4

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, by 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 in the USA and other countries. [®] indicates USA registration.

IBM[®], ACF/VTAM[®], AIX[®], APPN[®], MVS/ESA[®], OS/2[®], OS/390[®], VM/ESA[®], and VTAM[®] are registered trademarks or trademarks of International Business Machines Corporation. [®] 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.